Universidade Federal de São Paulo

Instituto de Ciência e Tecnologia

Mestrado em Matemática Pura e Aplicada

COMUTADORES E IMAGENS DE POLINÔMIOS NÃO COMUTATIVOS

Pedro Henrique Da Silva Dos Santos

São José dos Campos

22 de março de 2022

Pedro Henrique Da Silva Dos Santos

COMUTADORES E IMAGENS DE POLINÔMIOS NÃO COMUTATIVOS

Dissertação apresentada à Universidade Federal de São Paulo – Instituto de Ciência e Tecnologia como parte dos requisitos para a obtenção do título de Mestre em Matemática Pura e Aplicada.

Orientador:
Prof. Dr. Thiago Castilho de Mello

São José dos Campos

22 de março de 2022

Santos, Pedro

Comutadores e imagens de polinômios não comutativos / Pedro Henrique Da Silva Dos Santos. – São José dos Campos, 2022.

xii, 104.

Dissertação (Mestre) – Universidade Federal de São Paulo, Instituto de Ciência e Tecnologia. Programa de Pós-Graduação em Matemática Pura e Aplicada.

Título em inglês: Commutators and images of noncommutative polynomials.

1. Imagens de polinômios. 2. Conjectura de Lvov-Kaplansky. 3. Conjectura de Mesyan.

Universidade Federal de São Paulo

Instituto de Ciência e Tecnologia

Programa de Pós-Graduação em Matemática Pura e Aplicada

Chefe do Departamento: Prof. Dr. Marcelo Cristino Gama

Coordenador do Programa: Prof. Dr. Pedro Levit Kaufmann

Apoio Financeiro: Capes

Pedro Henrique Da Silva Dos Santos

COMUTADORES E IMAGENS DE POLINÔMIOS NÃO COMUTATIVOS

Presidente da banca:

Prof. Dr. Thiago Castilho de Mello

Banca examinadora:

Profª. Drª. Manuela da Silva Souza

Prof. Dr. Felipe Yukihide Yasumura

Prof. Dr. Willian Versolati França

A meus pais.

AGRADECIMENTOS

Agradeço primeiramente a Deus que concedeu inúmeras bênçãos, conhecimentos e oportunidades, para que eu pudesse estar aqui.

A minha família. Especialmente, minha mãe Maria Do Carmo e meu pai Claudinei Freire, que juntos enfrentaram tantas dificuldades para que eu pudesse estar aqui hoje.

A todos os professores do ICT-UNIFESP, pelo apoio e incentivo dentro e fora da sala de aula, que certamente colaboraram para a minha formação.

Agradeço ao meu orientador, professor Thiago Castilho de Mello, pelas palavras de apoio, correções e paciência ao longo desses anos. E cuja experiência foi inestimável na formulação das questões de pesquisa e na metodologia.

Agradeço a professora Manuela Souza e os professores Felipe Yasumura e Willian França membros da Banca Examinadora por terem aceitos ao convite e sugerido valiosas correções ao texto.

E a CAPES pelo apoio financeiro.

Não ame pela beleza, pois um dia ela acaba. Não ame por admiração, pois um dia você se decepciona. Ame apenas, pois o tempo nunca pode acabar com um amor sem explicação

Madre Teresa de Calcutá

RESUMO

Neste trabalho de dissertação iremos abordar definições e resultados básicos sobre PI-álgebras e estudar as imagens de polinômios multilineares sobre a álgebra das matrizes. Apresentaremos resultados de Herstein que concentra seus estudos nas estruturas dos anéis de Jordan e Lie de anéis associativos. Também estudaremos alguns resultados de Brešar e Vitas sobre a relação entre o espaço linear gerado pelos comutadores em A, e espaço linear gerado pela imagem de f em A. E estabelecemos alguns resultados do tipo Waring para imagens de polinômios.

Palavras-chave: Imagens de polinômios, Conjectura de Lvov-Kaplansky, conjectura de Mesyan.

Abstract

In this dissertation work we will approach the basic results on PI-algebra and we will study images of multilinear polynomials on matrix algebra. We present the results of Herstein concerning the structures of the Jordan rings and Lie rings. We also study some results from Brešar and Vitas to study the relationship between the linear span of commutators in A, and the linear span of the image of f in A. We establish some Waring type results for images of polynomials.

Keywords: Images of polynomials, Lvov-Kaplansky conjecture, Mesyan conjecture.

SUMÁRIO

Capítulo 1

Introdução

Seja $\mathbb{K}$ um corpo e $\mathbb{K}\langle X_1, \ldots, X_m\rangle$ a álgebra livre em $\{X_1, \ldots, X_m\}$ sobre $\mathbb{K}$. A imagem de um polinômio não comutativo $f \in \mathbb{K}\langle X_1, \ldots, X_m\rangle$ na $\mathbb{K}$-álgebra A é o conjunto

$$f(A) = \{f(a_1, \ldots, a_m) | a_1, \ldots, a_m \in A\}.$$

Dizemos que f é um polinômio multilinear se for da forma

$$f(x_1, \ldots, x_m) = \sum_{\sigma \in S_m} \lambda_\sigma x_{\sigma(1)} \ldots x_{\sigma(m)} \text{ com } \lambda_\sigma \in \mathbb{K}.$$

A conjectura de L'vov-Kaplansky[1] afirma que a imagem de um polinômio multilinear na álgebra das matrizes $M_n(\mathbb{K})$ é um espaço vetorial. É sabido que o espaço vetorial gerado pela imagem de um polinômio em uma álgebra A é um ideal de Lie de A. Em [13] Herstein provou que os ideais de Lie de $M_n(\mathbb{K})$ só podem ser um dos quatro espaços vetoriais: $\{0\}$, o espaço das matrizes escalares $\mathbb{K}$, o espaço das matrizes de traço zero $sl_n(\mathbb{K})$ ou a álgebra inteira $M_n(\mathbb{K})$. Portanto a conjectura se resume aos quatro subespaços citados.

Em [17], os matemáticos Alexei Kanel-Belov; Sergey Malev e Louis Rowen provaram a conjectura acima para matrizes 2×2 (sobre um corpo $\mathbb{K}$ quadraticamente fechado). Em 2016, os mesmos três autores estudaram o caso para matrizes 3×3 (novamente sobre um corpo $\mathbb{K}$ quadraticamente fechado), e obtiveram, além de $\{0\}$, $\mathbb{K}$, $sl_n(\mathbb{K})$ e $M_n(\mathbb{K})$, outros conjuntos como candidatos, como, por exemplo, conjuntos densos em $M_3(\mathbb{K})$ (ver [18]). Contudo, não conseguiram apresentar algum polinômio cuja imagem fosse algum destes candidatos.

Mesmo sendo extensivamente estudada por vários matemáticos, a conjectura está atualmente sem solução até mesmo para $n = 3$.

Em 1936 Shoda demonstrou, que para corpos de característica 0, toda matriz com traço zero pode ser expressa como um comutador, ou, equivalentemente, que a imagem do polinômio $f(x,y) = xy - yx$ em $M_n(\mathbb{K})$ é igual a $sl_n(\mathbb{K})$ [26]. Posteriormente, no ano de 1957, os matemáticos Albert e Muckenhoupt, generalizaram este teorema para corpos quaisquer [3], o que hoje conhecemos como Teorema de Shoda, Albert e Muckenhoupt.

Por outro lado, Mesyan generalizou, de certa forma, o teorema de Shoda, Albert and Muckenhoupt [25], mostrando que todo polinômio multilinear diferente de zero de grau no máximo 3, com coeficientes em $\mathbb{K}$, contém as matrizes de traço zero.

Ele conjecturou ainda, o que pode ser visto como um enfraquecimento da conjectura de Lvov-Kaplansky, a chamada conjectura de Mesyan. Ela afirma que sobre um corpo $\mathbb{K}$, dados $m \geqslant 1$ e $f(x_1, \cdots, x_m) \in K\langle x_1, \ldots, x_m\rangle$ um polinômio multilinear não nulo, temos que se $n \geqslant m-1$, então $(f(M_n(\mathbb{K}))$ contém $sl_n(\mathbb{K})$. Essa conjectura foi provada para $m = 3$ pelo próprio Mesyan em [25]. Para $m = 4$ uma solução foi dada em [10], entretanto, a demonstração deste resultado continha um erro, que foi corrigido em [12].

Um análogo da conjectura de Mesyan no caso de dimensão infinita, foi abordado em [28], onde o autor demonstra que tal conjectura é verdadeira para álgebra de matrizes finitárias (ou seja, matrizes infinitas com um número finito de entradas não nulas).

Em [29], o autor provou que se A é uma álgebra tendo uma derivação interna sobrejetora, então $f(A) = A$ para todo polinômio multilinear diferente de zero. Como um exemplo de tal álgebra A é $End_{\mathbb{K}}(V)$, a álgebra de endomorfismos de um espaço vetorial V de dimensão infinita. Esse resultado pode ser visto como a solução de uma versão de dimensão infinita da conjectura de L'vov-Kaplansky.

Além do apresentado acima, diversas variações da conjectura de Lvov-Kaplansky foram abordadas, considerando subálgebras de $M_n(K)$, ou mesmo álgebras não-associativas. Para um levantamento mais detalhado dos resultados a respeito da conjectura de L'vov-Kaplansky e suas variações recomendamos a leitura de [19].

O problema clássico de Waring, proposto por Edward Waring em 1770 e resolvido por David Hilbert em 1909, pergunta se para todo inteiro positivo k existe um inteiro positivo $g(k)$ tal que todo inteiro positivo pode ser expresso como uma soma de $g(k)$ $k-$ésimas potências de k. Por exemplo, se $k = 3$, estamos tratando de cubos, e o problema de Waring seria encontrar um inteiro positivo $g(k)$, tal que todo número inteiro n pudesse ser representado pela soma de no máximo $g(k)$ cubos.

Várias extensões e variações deste problema foram estudadas por diferentes grupos de

matemáticos. Um deles é o problema de Waring para grupos simples finitos, resolvido em 2011, por Larsen, Shaleve e Tiep [22]. Eles provaram que dado um monômio $w = w(x_1, \ldots, x_m) \neq 1$ e denotando por $w(G)$, onde G é um grupo, a imagem de w, pela aplicação que mapeia de G^m para G induzida por w, que $w(G)^2 = G$ para todo grupo simples finito não abeliano G de ordem suficientemente alta.

Portanto é natural buscar análogos desses resultados para álgebras simples de dimensão finita, com polinômios não comutativos, ou seja, elementos de álgebras livres desempenhando o papel de palavras, que são os elementos de grupos livres.

Uma área de pesquisa ainda ativa é o problema de matrizes de Waring, que pergunta sobre o número de somatórios necessários para expressar uma dada matriz em $M_n(C)$, onde C é um anel comutativo, como uma soma de $k-$ésimas potências de algumas matrizes em $M_n(C)$(ver [21], [23]). Isto é, obviamente, uma extensão do problema de Waring. Pode-se observar que problemas do tipo Waring em potências k-ésimas foram estudados em vários anéis e álgebras, não apenas em $M_n(C)$ (como por exemplo em [24, 27]).

Nesta dissertação faremos um apanhado de diversos resultados sobre imagens de polinômios em álgebra. Em especial, o maior interesse será estudar a relação entre o subespaço gerado pelo polinômio comutador e por um polinômio não comutativo f arbitrário (assumiremos que se a álgebra não tem unidade, o polinômio tem termo constante nulo).

No Capítulo 2, apresentaremos as definições e resultados básicos para estabelecer os principais resultados nesta dissertação. Os resultados destas dissertação foram retiradas das referências [9] e [11] .

O Capítulo 3, o principal da dissertação, é dividido em 5 seções. Inicialmente apresentamos os resultados de Mesyan que motivaram sua conjectura (Ver [13]). Após esta primeira etapa, apresentaremos um resultado inédito que sugere que a conjectura pode ser enunciada para um limitante melhor (Teorema 3.3) Observamos que a demonstração deste resultado foi obtido pelo autor da dissertação, apesar de poder ser provada com outros métodos apresentados na dissertação.

Em um segundo momento, as duas seções seguinte serão apresentadas os resultados de Vitas [28], que provam que a conjectura de Mesyan é verdadeira para álgebras de matrizes finitárias.

Dando continuidade, em [9] vemos que quando $f \in K\langle X\rangle$ é um polinômio não constante e não comutativo consideramos a questão de quando $span f(A)$, subespaço vetorial gerado pela imagem de f em A, é igual a A. E posteriormente analisemos a relação entre $[A, A]$, o espaço vetorial gerado pelos comutadores em A, e $span f(A)$.

Já o capítulo 4 apresentamos resultados do tipo Waring. Mais especificamente, mostramos que, sob suposições adequadas, todo elemento (ou pelo menos todo comutador) de uma álgebra A é uma soma/diferença ou uma combinação linear de um número fixo de elementos de $f(A)$. Dentre os resultados apresentados, destacamos, por exemplo, o Corolário 4.44, que afirma que se f não é identidade polinomial nem polinômio central para $M_n(\mathbb{C})$, então toda matriz de traço zero é a soma de 4 matrizes de $f(A) - f(A)$ (ver [9].

Capítulo 2

Conceitos Preliminares

O objetivo principal desde capítulo é estabelecer algumas notações, definições e resultados básicos que serão utilizadas ao longo de toda a dissertação. No texto, $\mathbb{K}$ denotará um corpo e quando não mencionado, consideraremos todas os espaços vetoriais e álgebras definidos sobre $\mathbb{K}$.

2.1 Anéis

Nesta secção apresentaremos os primeiros passos necessários para construir uma teoria geral da estrutura de anéis associativos. As principais referências são os livros [8, 11]

Definição 2.1. Um conjunto não vazio A equipado com duas operações binárias $+ : A \times A \to A$ (adição) e $\cdot : A \times A \to A$ (multiplicação) é chamado de anel, se $(A,+)$ é um grupo abeliano e valem as seguintes propriedades para todo $a,b,c \in A$:

1. $(a+b)\cdot c = a\cdot c + b\cdot c$
2. $a\cdot(b+c) = a\cdot b + a\cdot c$.

Em geral, se não houver ambiguidade, denotaremos $a\cdot b$ simplesmente por ab, e quando as operações estiverem claras pelo contexto, denotaremos $(A,+,\cdot)$ simplesmente por A.

Definição 2.2. Um anel A será dito:

1. **Associativo:** se $(ab)c = a(bc)$, para quaisquer $a,b,c \in A$;
2. **Comutativo:** se $ab = ba$, para quaisquer $a,b \in A$;

3. **Unitário**: (ou com unidade), se existir $1_A \in A$ tal que $1_A a = a 1_A = a$, para todo $a \in A$.

Definição 2.3. Um anel $(D, +, \cdot)$ é chamado de domínio de integridade ou simplesmente por domínio se satisfaz:

$$\forall\, x, y \in D - \{0\},\ \text{tem-se que o produto}\ x \cdot y \neq 0$$

Além disso, um anel $(D, +, \cdot)$ é de divisão se é unitário e satisfaz:

$$\forall\, x \in D - \{0\},\ \exists\ y \in D\ \text{ tal que } x \cdot y = y \cdot x = 1$$

Observação 2.4. A menos que seja dito o contrário, estaremos supondo durante o texto que os anéis são associativos e com unidade.

Definição 2.5. Um anel de divisão comutativo e unitário é chamado de corpo.

Exemplo 2.6. Um conjunto que será um dos objetos fundamentais de nosso estudo e vai ser abordado outras vezes, é o conjunto das matrizes $n \times n$ denotada por $M_n(\mathbb{K})$ em que $\mathbb{K}$ é um corpo. Esse conjunto é um anel com as operações usuais de adição e multiplicação de matrizes. Observamos que tal anel é associativo com unidade, porém não é comutativo para $n > 1$.

Definição 2.7. Seja $(A, +, \cdot)$ um anel. Um subconjunto não-vazio $S \subseteq A$, é chamado de subanel de A, se S é um anel com as operações induzidas pelas operações de A. O Centro $Z(A)$ de um anel A é o subanel definido por $Z(A) = \{a \in A;\ \text{tal que } ab = ba\ \forall b \in A\}$.

Observação 2.8. Se A é um anel comutativo, então $Z(A) = A$.

Observação 2.9. Uma matriz $n \times n$ é dita unitária denotado por e_{ij}, se é uma matrizes com entrada (i, j) igual à 1 e zero nas demais entradas. O produto de tais matrizes é dado por: $e_{ij} \cdot e_{kl} = \delta_{jk} \cdot e_{il}$. Onde

$$\delta_{jk} = \begin{cases} 0, & \text{se } j \neq k \\ 1, & \text{se } j = k \end{cases}$$

Logo para o produto $e_{ij} \cdot e_{kl}$ é zero se $j \neq k$, ou é e_{il} se $j = k$.

Exemplo 2.10. Um fato conhecido é que dado $n \in \mathbb{N}$ tem-se $Z(M_n(K)) = \{\alpha I_n;\ \text{tal que } \alpha \in K\}$. Onde I_n é a matriz identidade $n \times n$.

De fato, tome $X \in \{\alpha I_n | \alpha \in K\}$, então $X = \alpha I_n$ para algum $\alpha \in K$. Note que para qualquer $A \in M_n(K)$ temos que $X \cdot A = \alpha A = A \cdot X$ logo $x \in Z(M_n(K))$ e portanto $\{\alpha I_n;\ \text{tal que } \alpha \in K\} \subseteq Z(M_n(K))$.

Por outro lado, seja $Y = (a_{ij}) \in Z(M_n(K))$, suponha primeiro que existe uma entrada não diagonal diferente de zero. Sem perda de generalidade seja $a_{ij} \neq 0$ com $i < j$, seja e_{jj} a matriz unitária, é fácil verificar que $Ye_{jj} \neq e_{jj}Y$, isso porque na entrada (i, j) da matriz Ye_{jj} temos o elemento $a_{ij} \neq 0$, mas na matriz $e_{jj}Y$ a entrada (i, j) é zero.

Logo deve ocorrer de $a_{ij} = 0$ para todo $i \neq j$. Portando Y é uma matriz diagonal.

Resta mostrar que é múltiplo da identidade. Suponhamos por absurdo que isso não seja verdade, então existem em Y, y_{ii} e y_{jj} tais que $y_{ii} \neq y_{jj}$ logo

$$Y \cdot e_{ij} = \sum_{k=1}^{n} y_{kk}e_{kk}e_{ij} = y_{ii}e_{ij} \neq y_{jj}e_{ij} \sum_{k=1}^{n} y_{kk}e_{kk}e_{ij} = Y \cdot e_{ij}$$

Absurdo, já que Y está no centro de $M_n(K)$, logo Y é múltiplo da identidade.

Observação 2.11. Algumas álgebras possuem o centro trivial, isto é, centro cujo único elemento é 0, como é o caso da álgebra das matrizes com apenas a primeira linha não nulo.

Definição 2.12. Seja I um subconjunto não vazio de A, tal que dados $x, y \in I$ tem-se que $x - y \in I$. O subconjunto I é dito ser um ideal (bilateral) do anel A se ocorrer de $AI \subseteq I$ e $IA \subseteq I$, em outras palavras, se ai e $ia \in I$ para quaisquer $a \in A$ e $i \in I$. Podemos definir ainda ideias unilaterais. No caso de valer apenas $IA \subseteq I$, então I é dito ser um ideal à direita e respectivamente à esquerda se ocorrer apenas de $AI \subseteq I$.

Observação 2.13. Quando o anel A é comutativo, as definições de ideal bilateral e unilaterias coincidem.

Exemplo 2.14.

1. $\{0\}$ e A são subanéis e ideais triviais do anel A. Os ideais não triviais de A são chamados de ideais próprios de A.

2. Seja $n \geqslant 0$ um número inteiro. Temos que o subconjunto $n\mathbb{Z} = \{zn | z \in \mathbb{Z}\ \}$ é um ideal bilateral do anel dos inteiros.

Definição 2.15. Um anel A que só possui ideais triviais é chamado de anel simples.

Observação 2.16. Seja I um ideal à esquerda (respectivamente à direita) do anel A. Se I contém um elemento inversível, então $I = A$.

Exemplo 2.17.

1. Se K é um corpo então K é um anel simples.

   De fato, seja I um ideal não nulo de K, tome $0 \neq k \in I$. Como K é um corpo tem-se que $k^{-1} \in K$. Mas I é um ideal, pela observação anterior temos que $I = K$.

2. Seja A um anel, e $x \in A$. O conjunto $xA = \{xa; a \in A\}$ é um ideal à direita de A, chamado ideal gerado por a. Note que xA subanel de A.

   De forma análoga definimos $Ax = \{ax; a \in A\}$, que neste caso é um ideal à esquerda de A.

3. $AaA = \{$somas finitas de elementos da forma $\alpha_1 \cdot a \cdot \alpha_2$ onde $\alpha_1, \alpha_2 \in A\}$ é um ideal de A.

Definição 2.18. Para qualquer subconjunto S de um anel A, dizemos que o ideal de A gerado pelo conjunto S é o menor ideal que contém S. Em outras palavras, o ideal gerado por S é o conjunto formado por todas as somas finitas de elementos da forma s, as, sa, asb com $s \in S$ e $a, b \in A$.

Definição 2.19. Um ideal I do anel A é dito ser semiprimo se $I \neq A$ e tivermos

$$aAa \subseteq I \text{ com } a \in A \Longrightarrow a \in I.$$

Dizemos que o anel A é semiprimo se o ideal(0) é semiprimo.

Definição 2.20. Um anel A chama-se anel primo se, para quaisquer $a, b \in A$, tais que $aAb = 0$, temos $a = 0$ ou $b = 0$.

É fácil ver que todo anel primo é um anel semiprimo, basta tomar o caso particular $b = a$. Por outro lado a recíproca é falsa. Considere o anel $M = \mathbb{Z} \times \mathbb{Z} = \{(a, b) |\ a, b \in \mathbb{Z}\}$, note que este anel é semiprimo, de fato, tome (a_1, a_2) um elemento qualquer de M, para qualquer $(m_1, m_2) \in M$ tem-se que

$$(a_1, a_2)(m_1, m_2)(a_1, a_2) = (0, 0) \Longrightarrow a_1 m_1 a_1 = 0, a_2 m_2 a_2 = 0 \Longrightarrow a_1 = 0, a_2 = 0$$

Por outro lado, este anel não é primo. De fato, note que $(1, 0)M(0, 1) = (0, 0)$, mas claramente $(1, 0)$ e $(0, 1)$ são não nulos.

Proposição 2.21. *Um anel A é um anel primo, se e somente se, para todos ideais I e J de A tais que $IJ = 0$, implica em $I = 0$ ou $J = 0$.*

Demonstração.

($\Rightarrow$) Sejam I e J ideais de A tais que $IJ = 0$. Como $AJ \subseteq J$, então $IAJ \subseteq IJ = 0$. Se $I \neq 0$, chame de x um elemento não nulo de I e $b \in J$. Como $xAb = 0$, por definição temos que $b = 0$, ou seja, $J = 0$.

($\Leftarrow$) Sejam $a, b \in A$ tais que $aAb = 0$. Logo tem-se que, o produto dos ideais $AaA \cdot AbA = 0$, e por hipótese obtemos que algum deles é nulo. Suponha, sem perda de generalidade, que $AaA = 0$, isso implica que os ideiais Aa e aA satisfazem $Aa \cdot A = A \cdot aA = 0$, novamente por hipótese, tem-se que $Aa = aA = 0$. Logo, $\mathbb{Z}a$ é um ideal e $(\mathbb{Z}a)A = 0$ e consequentemente tem-se que $\mathbb{Z}a = 0$ o que implica que, $a = 0$. □

Observação 2.22. Note que a Proposição 2.21 também é válido para os casos dos conjuntos I, J serem apenas ideias unilaterais. Além disso, através dessa proposição é fácil ver que todo anel simples é um anel primo.

Proposição 2.23. *Se I é um ideal de um anel primo A, então I também é um anel primo*

Demonstração. Seja $a, b \in A$, tal que $aIb = 0$. Observe que para qualquer $\alpha_1, \alpha_2 \in A$, tem-se que $\alpha_1 i \alpha_2 \in I$, para todo $i \in I$. Logo

$$aAIAb = 0 \Longrightarrow aAuAb = 0 \ \forall u \in I.$$

Como A é um anel primo então $a = 0$ ou $uAb = 0$. Se ocorrer de $a = 0$ então não há mais nada a mostrar. Caso ocorra de $uAb = 0$, então por definição de anéis primos tem-se que $u = 0$ ou $b = 0$, tomando $u \neq 0$, tem-se que $b = 0$. Portanto I é um anel primo. □

Definição 2.24. Dada uma matriz $A \in M_n(\mathbb{K})$, dizemos que uma S é uma submatriz de A, se S é uma matriz obtida de A eliminando alguma(s) das suas linhas e/ou colunas. Qualquer submatriz quadrada cuja diagonal faça parte da diagonal de A é chamada submatriz principal.

Observação 2.25. Se toda submatriz principal de $A \in M_n(\mathbb{K})$ é uma matriz múltipla da identidade, então A também é uma matriz múltipla da identidade.

De fato, seja $A = (a_{ij}) \in M_n(\mathbb{K})$, e considere a a_{n-1} a submatriz de A definida por

$$a_{n-1} = \begin{bmatrix} a_{n-1,n-1} & a_{n-1,n} \\ a_{n,n-1} & a_{n,n} \end{bmatrix}$$

Como toda submatriz principal de $A \in M_n(\mathbb{K})$ é uma matriz múltipla da identidade, então ocorre

de $a_{n-1,n-1} = a_{n,n}$. Por outro lado definindo

$$a_{n-2} = \begin{bmatrix} a_{n-2,n-2} & a_{n-2,n-1} & a_{n-2,n} \\ a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,n-2} & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

Donde segue que $a_{n-2,n-2} = a_{n-1,n-1} = a_{n,n}$. Indutivamente conclui-se que $a_{1,1} = \ldots = a_{n-1,n-1} = a_{n,n}$, portanto A é uma matriz múltipla da identidade.

Definição 2.26. Dada uma matriz $A \in M_n(\mathbb{K})$, definimos seu rank (ou posto), como sendo o número de linhas não nulas em sua forma escalonada.

Observação 2.27.

1. Qualquer matriz de quadrado igual a zero em $M_n(\mathbb{K})$, tem rank no máximo $\lfloor n/2 \rfloor = \max\{m \in \mathbb{Z}; m \leqslant n/2\}$.

   De fato, seja $A \in M_n(\mathbb{K})$ uma matriz de quadrado igual a zero, e considere T a transformação linear representada pela matriz A. Se $y \in Im(T)$ então $y = Ax$ para algum $x \in K^n$, e logo $Ay = A(Ax) = A^2x = 0$ e portanto $Im(T) \subseteq ker(T)$. Consequentemente pelo teorema do núcleo e imagem tem-se que

   $$\begin{aligned} dim(K^n) &= ker(A) + rank(A) \\ n &\geqslant rank(A) + rank(A) \\ \lfloor n/2 \rfloor &\geqslant rank(A) \end{aligned}$$

2. Duas matrizes de quadrado igual a zero têm a mesma rank, se e somente, se forem semelhantes.

   De fato, suponha primeiro que duas matrizes A, B são semelhantes, isto é, existe matriz invertível M tal que $A = MBM^{-1}$. Temos então que

   $$rank(B) \geqslant rank(MBM^{-1}) = rank(A).$$

   De forma análoga, obtemos

   $$rank(A) \geqslant rank(M^{-1}AM) = rank(B).$$

   Portanto $rank(A) = rank(B)$.

   Por outro lado, suponhamos agora que duas matrizes A e B de quadrado zero têm o mesmo rank. Devemos mostrar que A e B são semelhantes.

De fato, note que o polinômio minimal de A e B deve ser $p_A(x) = x^2$ ou $p_A(x) = x$ e $p_B(x) = x^2$ ou $p_B(x) = x$ respectivamente. Observe que o caso polinômio minimal for $p_A(x) = x^2$ e $p_B(x) = x$, teríamos que $rank(A) \neq rank(B)$, já que $B = 0$ e $A \neq 0$, o caso em que $p_B(x) = x^2$ e $p_A(x) = x$ é análogo, logo os polinômios minimais devem ser os mesmo.

O caso em que $p_A(x) = x$ e $p_B(x) = x$ é trivialmente verdadeiro. Logo o caso interessante é quando $p_A(x) = x^2$ e $p_B(x) = x^2$. Representando A em sua forma de Jordan temos que ela possui k blocos de tamanho 2 e r de tamanho 1. Já a matriz B apresenta l matrizes de tamanho 2 e s matrizes de tamanho 1.

Note que $n = 2k + r = 2l + s$, como A e B tem o mesmo rank e matrizes nilpotentes tem autovalores nulo, suas formas de Jordan tem apenas zero na diagonal principal. Segue que $k = l$ o que implica que $r = s$.

Portanto A e B são semelhantes.

O conceito de módulo, pode ser entendido como uma generalização de espaço vetorial, uma vez que nos espaços vetorias os escalares são elementos de um corpo, já nos módulos os escalares pertencem há algum anel associativo. Mais precisamente:

Definição 2.28. Um módulo M sobre um anel unitário A (abreviadamente, um A-módulo) é um grupo comutativo com operação de adição, em conjunto com uma lei de composição externa

$$\begin{aligned} A \times M &\to M \\ (a, m) &\mapsto am, \end{aligned}$$

Que satisfaz

1. $1m = m, \ \ m \in M$;
2. $a(m+n) = am + an, \ \ a \in A, \ m, n \in M$;
3. $(a+b)m = am + bm, \ a, b \in A, m \in M$;
4. $a(bm) = (ab)m, \ a, b \in A, \ m \in M$.

Dizemos que M é um A-módulo livre se possui uma base.

Observação 2.29. De maneira semelhante a subespaço vetorial, dizemos que um subgrupo N do grupo abeliano $(M, +)$ é um submódulo do A-módulo M, se para todo $a \in A$ e $n \in N$ temos

$an \in N$. Com isso um módulo M sobre um anel A é chamado simples ou irredutível se não é o módulo nulo e se seus únicos submódulos sobre A são 0 e M.

Exemplo 2.30.

1. Seja $\mathbb{K}$ um corpo. De acordo, com as observações anteriores, um $\mathbb{K}$-módulo coincide com um $\mathbb{K}-$espaço vetorial.

2. Seja D um domínio de integridade, e M um D-módulo. Tem-se que

$$T(M) = \{m \in M; \text{ tal que } \exists\ d \neq 0 \in D \text{ com } dm = 0\}$$

é um submódulo chamado de Torção de M.

Seja N um subconjunto de A, o anulador de N é definido como

$$Ann(N) = \{a \in A; \text{ tal que } an = 0,\ \forall\ n \in N\}.$$

Observação 2.31.

1. Se N é um A-módulo, então $Ann(N)$ é um ideal à esquerda de A.

2. Se N_1 e N_2 são dois módulos isomorfos, então vale a igualdade $Ann(N_1) = Ann(N_2)$.

Iremos agora introduzir conceito de radical de um anel.

Definição 2.32. O radical de Jacobson de um anel A, denotado por $J(A)$, é o ideal que consiste naqueles elementos em A que anulam todos os A-módulos simples. Isto é

$$J(A) = \{a \in A;\ \ aM = 0; \text{onde } \ M \text{ é um módulo simples}\}$$

Lema 2.33. *Seja R um anel unitário, e $y \in R$. São equivalentes:*

1. *$y \in J(R)$;*

2. *y pertence a intersecção dos ideais maximais de R;*

3. *$1 - xy$ tem um inverso à esquerda, para todo $x \in R$.*

Demonstração.

(1) $\Rightarrow$ (2) Se $m \subseteq R$ é um ideal à esquerda maximal, então $M = R/m$ é um R-módulo à esquerda simples. Daí, $yM = 0$, isto é, $yR \subseteq m$. Em particular, $y \in m$.

(2) $\Rightarrow$ (3) Seja y um elemento que está na intersecção de todos os ideais maximais de R, e suponha por absurdo que existe $x \in R$ tal que $1-xy$ não possui um inverso à esquerda. Logo, o ideal à esquerda $R(1-xy)$ é diferente de R, portanto, está contido em algum ideal maximal m. Como $y \in m$, segue que $1 = xy + (1-xy) \in m$, uma contradição.

(3) $\Rightarrow$ (1) Seja M um R-módulo simples tal que, $ym \neq 0$, para algum $m \in M$, isso implica que $R(ym) = M$ e portanto existe $x \in R$ tal que $x(ym) = m$. Consequentemente, tem-se que $(1-xy)m = 0$, além disso, como $(1-xy)$ é inversível à esquerda, segue que $m = 0$ (absurdo).

□

Observe que $J(R) = \cap Ann(M)$, onde esta interseção passa por todos os R-módulos irredutíveis M. Uma vez que $Ann(M)$ são ideais bilaterais de R, vemos que $J(R)$ é um ideal bilateral de R.

Observação 2.34. Seja R um anel, o radical de Jacobson $J(R)$ é a intersecção de seus ideais à esquerda maximais.

Exemplo 2.35. Se $\mathbb{K}$ é um corpo e $A = UT_n(\mathbb{K})$ é o anel de todas as matrizes triangulares superiores $n \times n$ com entradas em $\mathbb{K}$, então $J(A)$ consiste em todas as matrizes triangulares superiores com zeros na diagonal principal.

Definição 2.36. Um anel A é dito ser semisimples se $J(R) = 0$.

Exemplo 2.37. O radical de Jacobson de qualquer corpo, e do anel $\mathbb{Z}$, é $\{0\}$.

2.2 Produto Tensorial

Iniciamos esta secção com a definição de produtos tensoriais bem como alguns resultados elementares.

Definição 2.38. Sejam V e W dois espaços vetoriais com bases $\{v_i, \text{ com } i \in I\}$ e $\{w_, \text{ com } j \in J\}$, respectivamente ($I$ e J são conjuntos de índices). O produto tensorial $V \otimes_k W$ de V por W é um espaço vetorial com

com base $\{v_i \otimes w_j | i \in I, j \in J\}$. Onde vale

$$\left(\sum_{i\in I} \alpha_i v_i\right) \otimes \left(\sum_{j\in J} \beta_j w_j\right) = \sum_{i\in I}\sum_{j\in J} \alpha_i\beta_j (v_i \otimes w_j),$$

na qual α_i e β_j são escalares e as somas são finitas.

Observação 2.39.

1. Denotamos $V \otimes_k W$ quando queremos enfatizar sobre qual corpo estamos efetuando o produto tensorial. Quando não existir dúvidas quanto isso, o produto tensorial de V por W sobre $\mathbb{K}$ será denotado simplesmente por $V \otimes W$.

2. O produto tensorial é único a menos de isomorfismo.

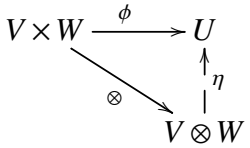
3. É fácil ver que se V e W tem dimensão finita n e m respectivamente, a dimensão do produto $V \otimes W$ é o produto $n \cdot m$.

Exemplo 2.40. Seja $\mathbb{K}[x], \mathbb{K}[y]$ respectivamente os espaços vetoriais dos polinômios nas variáveis x e y sobre $\mathbb{K}$, temos que $\mathbb{K}[x] \otimes \mathbb{K}[y] \simeq \mathbb{K}[x, y]$. De fato, considere a aplicação linear

$$\begin{aligned} \phi : \mathbb{K}[x] \otimes \mathbb{K}[Y] &\longrightarrow \mathbb{K}[x, y] \\ (x^i \otimes y^i) &\mapsto x^i y^i \end{aligned}$$

é um isomorfismo.

Proposição 2.41. *Seja $(V \times W; \otimes)$ o produto tensorial de V e W, U um espaço vetorial e $\phi : V \times W \to U$ um aplicação bilinear. Então existe uma única aplicação linear $\eta : V \otimes W \to U$ tal que o diagrama a seguir é comutativo.*



Demonstração.

Seja $\{v_i \text{ com } i \in I\}$ e $\{w_j \text{ com } j \in J\}$, as bases respectivamente de V e W, onde I e J são conjuntos de índices. Sabemos que $\{(v_i \otimes w_j); i \in I, j \in J\}$ forma uma base para $V \otimes W$. A função linear η será, portanto, o único mapa linear de $V \otimes W$ em U tal que

$$\text{Para todo } i \in I, j \in J \text{ tem-se que } \eta((v_i \otimes w_j)) = \phi(v_i, w_j)$$

E estende-se linearmente definindo $\eta((a \otimes b) + (c \otimes d)) = \eta(a \otimes b) + \eta((c \otimes d))$. Portanto, o diagrama acima comuta. □

Observação 2.42. Essa resultado é conhecido como Propriedade Universal do produto tensorial.

Teorema 2.43. *Sejam V, W, V_1, W_1 quatro espaços vetoriais sobre $\mathbb{K}$. Seja $f : V \to V_1$ e $g : W \to W_1$ aplicações lineares. Existe uma única aplicação linear $V \otimes W$ em $V_1 \otimes W_1$, chamado de produto tensorial de f e g e denotado por $f \otimes g$, de tal forma que*

$$(f \otimes g)(x \otimes y) = f(x) \otimes g(y), \text{ para qualquer } x \in V, y \in W.$$

Demonstração. Chame

$$\begin{aligned} \phi : V \times W &\longmapsto V_1 \otimes W_1 \\ (x, y) &\mapsto f(x) \otimes g(y) \end{aligned}$$

Note que ϕ é bilinear ja que f e g são lineares por hipótese e a aplicação do produto tensorial $V_1 \otimes W_1$ é bilinear. Pela propriedade universal existe uma aplicação linear única $\eta : V \otimes W \to V_1 \times W_1$ tal que o diagrama a seguir comuta

$$\begin{array}{ccc} V \times W & \xrightarrow{b} & V_1 \otimes W_1 \\ & \searrow^{\otimes} & \uparrow \eta \\ & & V \otimes W \end{array}$$

Logo $\eta(x \otimes y) = \phi(x, y)$, para qualquer $(x, y) \in V \times W$. Portanto, usando a definição de ϕ, obtemos que $\eta = f \otimes g$. □

Exemplo 2.44. Seja $n, m \in \mathbb{N}$, $V = \mathbb{K}^n$ e $W = \mathbb{K}^m$. Então $V \otimes W \cong K^{nm}$ é um produto tensorial de V e W cujo aplicação bilinear ϕ é dado por

$$\begin{aligned} \phi : V \times W &\longmapsto \mathbb{K}^{nm} \\ \left((x_i)_{i=1}^n, (y_j)_{j=1}^m\right) &\mapsto (x_i y_j)_{1 \leqslant i \leqslant n; 1 \leqslant j \leqslant m}. \end{aligned}$$

2.3 Álgebras

Nesta secção introduzimos os conceitos de álgebras, subálgebra, homomorfismos de álgebras e álgebras livres, além de algumas propriedades e exemplos relevantes, que terão grandes utilidades ao longo do texto.

Definição 2.45. Uma álgebra A sobre um corpo $\mathbb{K}$ é um espaço vetorial sobre $\mathbb{K}$, munido de uma operação extra sobre os seus vetores $\star$, chamada de multiplicação, tal que para qualquer escalar $\alpha \in \mathbb{K}$ e quaisquer $a; b; c \in A$, tem-se

1. $(a+b)\star c = a\star c + b\star c$;

2. $a\star(b+c) = a\star b + a\star c$;

3. $(\alpha a)\star b = a\star(\alpha b) = \alpha(a\star b)$

Observação 2.46. Para simplificar a notação, vamos escrever ab, em vez de $a\star b$, para todo $a; b \in A$. Além disso, dizemos que um subconjunto é uma base da álgebra A se for uma base de A como espaço vetorial. Analogamente a dimensão da álgebra A será a mesma dimensão do espaço vetorial A.

Note que uma álgebra além de ser um espaço vetorial, apresenta a estrutura de um anel. Portanto alguns conceitos definidos no contexto de anéis tem análogos no contexto de álgebras.

Definição 2.47. Dizemos que uma álgebra A é:

1. associativa: se $(ab)c = a(bc)$, para quaisquer $a; b; c \in A$;

2. comutativa: se $ab = ba$, para quaisquer $a; b \in A$;

3. unitária: (ou com unidade) se o produto possui elemento neutro;

4. de Lie: se para quaisquer $a, b, c \in A$ valem

$$a^2 = aa = 0 \text{ (anticomutatividade)},$$
$$(ab)c + (bc)a + (ca)b = 0 \text{ (identidade de Jacobi.)}$$

Em todo o texto trabalharemos com álgebras associativas. Consequentemente, o termo álgebra deverá ser entendido como álgebra associativa.

Exemplo 2.48. (**Álgebra das matrizes**) Seja $n \in \mathbb{N}$, o espaço vetorial $M_n(\mathbb{K})$, munido da multiplicação usual de matrizes, é uma álgebra unitária, cuja a unidade é a matriz identidade.

Definição 2.49. Seja A uma álgebra.

1. Um subconjunto B de A é dito ser uma subálgebra, se com relação às operações de A, também é uma álgebra. Em outras palavras, B deve ser fechado com respeito à adição, multiplicação e multiplicação por escalar de A.

2. Um subconjunto $I \subseteq A$ é um ideal (bilateral) de A, se I é um subespaço vetorial de A, e além disso, $AI \subseteq I$ e $IA \subseteq I$, ou seja, se $ai \in I$ e $ia \in I$ para quaisquer $a \in A$ e $i \in I$.

Definição 2.50. Dizemos que um elemento a de um álgebra A, é nilpotente se $a^m = 0$ para algum inteiro m. Um ideal é dito nil se cada elemento é nilpotente. Por último, dizemos que um ideal I é nilpotente se existe um número inteiro n, tal que $a_1 \cdot a_2 \cdots a_n = 0$, para todo $a_1, \cdots, a_n \in I$.

Exemplo 2.51. Uma matriz A estritamente triangular superior é nilpotente. Basta ver que o polinômio característico de A é da forma $p(\lambda) = \lambda^n$, e pelo teorema de Cayley-Hamilton, temos que $p(A) = A^n = 0$.

Exemplo 2.52. Sejam A uma álgebra e X um subconjunto não nulo de A. Chame de B_X o subespaço vetorial de A gerado por $\{x_1 x_2 \cdots x_l \text{ tal que } l \in \mathbb{N}, x_i \in X\}$. Temos que B_X é multiplicativamente fechado. Portanto, B_X é uma subálgebra de A.

Definição 2.53. Uma subálgebra de A gerada por um subconjunto X, denotada por S_X, é a menor subálgebra que contém X, isto é, S_X é a intersecção de todas as subálgebras de A que contém X.

Observação 2.54. Das propriedades de espaço vetorial sabemos que um subespaço vetorial gerado por conjunto de vetores é o menor subespaço com tal propriedade. Logo pelo exemplo 2.52 é fácil ver que B_X é uma subálgebra de A gerada por um subconjunto X.

O próximo resultado nos permite obter uma estrutura de álgebra a partir de um espaço vetorial.

Teorema 2.55. *Sejam V um espaço vetorial e B uma base de V. Dada $f : B \times B \to V$ uma função qualquer, existe uma única aplicação bilinear $F : V \times V \to A$ que estende f*

Se V e W são $\mathbb{K}$-álgebras, então $V \otimes_{\mathbb{K}} W$ também é uma álgebra com multiplicação $(v' \otimes w')(v'' \otimes w'') = (v'v'') \otimes (w'w'')$, onde $v', v'' \in V$ e $w', w'' \in W$. Segue do teorema acima que para definir uma estrutura de álgebra $V \otimes W$, onde V e W são espaços vetoriais, basta definir a operação multiplicação nos elementos $v_i \otimes w_j$, onde v_i estão na base de V e w_j pertencem a base de W.

Se V e W são álgebras associativas então $V \otimes W$, é uma álgebra associativa, pela proposição anterior. Além disso, se A e B são álgebras com unidade, a unidade da álgebra $A \otimes B$ será $1_A \otimes 1_B$.

Definição 2.56. Sejam A e B duas álgebras. Dizemos que uma transformação linear $T : A \to B$ é um homomorfismo de álgebras, se $T(xy) = T(x)T(y)$, para quaisquer $x, y \in A$. Se A e B possuem unidade, exigimos que $T(1) = 1$.

Observação 2.57. No caso de um homomorfismo $f : A \to B$ ser bijetivo, diremos que f é um isomorfismo de álgebras, neste caso A e B são álgebras isomorfas e denotaremos $A \simeq B$.

Seja $\varphi : A \to B$ um homomorfismo. Denotamos o núcleo de φ por $N(\varphi) = \{a \in A; \text{tal que } \varphi(a) = 0\}$ e a imagem de φ por $Im(\varphi) = \{\varphi(a); \text{ onde } a \in A\}$. É de fácil verificação que $N(\varphi)$ é um ideal de A e a $Im(\varphi)$ é uma subálgebra de B.

O núcleo do homomorfismo $\varphi : A \to B$, também pode ser denotado por $ker(\varphi)$.

Agora podemos enunciar o Teorema dos Isomorfismos.

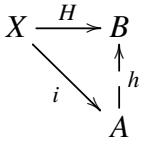
Teorema 2.58. ***(Teorema dos Isomorfismos)*** *Seja $f : A \to B$ um homomorfismo de álgebras. Então a álgebra quociente $A/N(f)$ é isomorfa a $Im(f)$.*

Exemplo 2.59. Sejam A uma álgebra e I um ideal de A. A aplicação $\pi : A \to A/I$, definida por $\pi(a) = a + I$, é um homomorfismo de álgebras, chamado de projeção canônica.

Definiremos agora álgebras livres em uma classe de álgebras e construiremos a álgebra livre na classe das álgebras associativas e com unidade.

Definição 2.60. Seja $\mathcal{V}$ uma classe de álgebras. Dizemos que uma álgebra $A \in \mathcal{V}$ é livre na classe $\mathcal{V}$ livremente gerada por X, se existe conjunto X gerador de A e para cada álgebra $B \in \mathcal{V}$ e cada função $h : X \to B$ existe um único homomorfismo $H : A \to B$ estendendo h.

Observação 2.61. Quando uma álgebra $A \in \mathcal{V}$ é livre na classe $\mathcal{V}$ livremente gerada por X, o diagrama abaixo comuta, isto é



onde $i : X \to A$ é a função inclusão.

Exemplo 2.62. Seja $\mathbb{K}[x_1, x_2, \cdots, x_n]$ a álgebra dos polinômios associativos e comutativos nas variáveis $x_1, x_2, \ldots, x_n$. Temos que esta álgebra é uma álgebra livre na classe de todas as álgebras associativas e comutativas, livremente gerada por $\{x_1, \ldots, x_n\}$.

De fato, é fácil ver que está álgebra é gerada pelo conjunto $X = \{x_1, \cdots, x_n\}$. Seja B uma álgebras associativa, comutativa e unitária e $b_i \in B$.

Para cada função

$$h : X \longrightarrow B$$
$$x_i \mapsto b_i$$

Existe um único homomorfismo

$$H : \mathbb{K}[x_1, \ldots, x_n] \longmapsto B$$
$$f(x_1, \ldots, x_n) \mapsto f(b_1, \ldots, b_n)$$

que satisfaz $H(x_i) = b_i$.

Tentaremos construir um exemplo um pouco mais geral que se refere a uma álgebra livre na classe de todas as álgebras associativas com unidade.

Seja $X = \{x_1, x_2, \cdots\}$ um conjunto não vazio e enumerável, cujos elementos iremos chamar de variáveis não comutativas. Definimos uma palavra em X de comprimento n como sendo $x_{i1} \cdot x_{i2} \cdots x_{ij}$. Quando um palavra tiver tamanho 0 chamamos simplesmente de 1. Alem disso, duas palavras $x = x_{i1} \cdot x_{i2} \cdots x_{in}$ e $y = x_{j1} \cdots x_{jk}$ são iguais se, $k = n$ e se todas as variáveis correspondentes são iguais.

Seja $K\langle X\rangle$ o espaço vetorial que tem como base o conjunto de todas as palavras em X. Os elementos de $K\langle X\rangle$ são chamados de polinômios não comutativos, ou simplesmente polinômios.

Munindo este espaço com a multiplicação

$$(x_{i1}x_{i2} \cdots x_{in}) \cdot (x_{j1}x_{j2} \cdots x_{jm}) = x_{i1}x_{i2} \cdots x_{in}x_{j1}x_{j2} \cdots x_{jm}$$

chamada de concatenação, este se torna torna-se uma álgebra associativa unitária.

Proposição 2.63. *A álgebra $K\langle X\rangle$ é livre, livremente gerada por X na classe das álgebras associativas unitárias.*

Demonstração. Seja B uma álgebra associativa unitária e seja

$$h : X \longmapsto B$$
$$x_i \to b_i.$$

Entao definimos,

$$H : K\langle X\rangle \to B$$
$$p(x_1, \cdots, x_k) \to p(b_1, \cdots, b_n)$$

que é um homomorfismo e satisfaz $H(x_i) = b_i$.

□

Existem vários tipos de polinômios que têm grande importância para o estudos de identidades de uma álgebra, dentre elas destacamos os polinômios multihomogêneos e polinômios multilineares.

Definição 2.64.

1. Um monômio $m(x_1,\ldots,x_n)$ tem grau k em x_i, se a variável x_i aparece exatamente k vezes em $m(x_1,\ldots,x_n)$.

2. Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é dito homogêneo em x_i, se todos os seus monômios têm o mesmo grau em x_i.

3. Seja $m(x_1,\ldots,x_n)$ um monômio de $\mathbb{K}\langle X\rangle$, definimos o multigrau de m como sendo a n-upla $(a_1,\ldots,a_n)$, onde a_i é o grau na variável x_i.

4. Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é dito multihomogêneo de multigrau $(a_1,\ldots,a_n)$, se f é homogêneo de grau a_i em x_i, para todo i.

5. Um polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é dito ser multilinear, se é multihomogêneo de multigrau $(1,1,\ldots,1)$.

Observação 2.65. Quando o polinômio $f(x_1,\ldots,x_n) \in \mathbb{K}\langle X\rangle$ é multilinear, f pode ser escrito como:

$$\sum_{\sigma\in S_n} \alpha_\sigma x_{\sigma(1)}x_{\sigma(2)}\cdots x_{\sigma(n)} \quad \text{onde } \alpha_\sigma \in \mathbb{K} \text{ e } S_n \text{ denota o grupo simétrico.}$$

Exemplo 2.66. O polinômio $f(x_1,x_2,x_3) = x_1^2x_2^2 + x_2x_1x_3x_1$ é homogêneo em x_1, já que os seus dois monômios apresentam grau 2 em x_1, mas não é multihomogêneo.

Definição 2.67. Sejam A uma álgebra, I um ideal (bilateral) de A. Sobre A definimos a relação de congruência $(mod\ I)$, para qualquer $a,b \in A$

$$a \equiv b(mod\ I) \Longleftrightarrow a-b \in I$$

Observação 2.68. É fácil ver que esta relação é de equivalência. Logo denotaremos por $a+I$ a classe de equivalência de $a \in A$ e A/I o conjunto de todas as classes de equivalência.

Definição 2.69. Consideremos no espaço vetorial quociente A/I com as operações de soma e multiplicação por escalar usuais para todo $a,b \in A$ e $\alpha \in \mathbb{K}$ e $\alpha(a+I) = \alpha a + I$, para todo $a,b \in A$ e $\alpha \in \mathbb{K}$. Podemos considerar a operação produto $(a+I)(b+I) = ab+I$ para $a,b \in A$. Este produto está bem definido, pois não depende da escolha dos representantes das classes laterais, e torna A/I uma álgebra, conhecida por álgebra quociente de A por I.

Lema 2.70. *Sejam V um espaço vetorial e U um subespaço de V . Sejam $c_0, c_1, \ldots, c_n \in V$ tais que*

$$\sum_{i=0}^{n} \lambda^i c_i \in U$$

para pelo menos $n+1$ escalares distintos $\lambda \in \mathbb{K}$. Então, $c_i \in U$ para todo i.

Demonstração. Sejam $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ distintos entre si tais que $\sum_{i=0}^{n} \lambda_l^i c_i \in U$ para $l \in \{0, \ldots, n\}$. Logo para cada l, tem-se que

$$\sum_{i=0}^{n} \lambda_l^i \overline{c}_i = 0 \text{ em } V/U.$$

Note que as equações acima podem ser representadas matricialmente como

$$\begin{pmatrix} 1 & \lambda_0 & \lambda_0^2 & \cdots & \lambda_0^n \\ 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^n \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^n \end{pmatrix} \cdot \begin{pmatrix} \overline{c}_0 \\ \overline{c}_1 \\ \overline{c}_2 \\ \vdots \\ \overline{c}_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Desde que os escalares λ_l são distintos entre si, então a matriz de Vandermonde acima possui determinante não nulo. Portanto, $\overline{c}_i = 0$ em V/U e consequentemente prova o lema. □

As definições de álgebras simples e de álgebras primas são totalmente análogas as de anéis simples e anéis primos respectivamente.

Definição 2.71. Dizemos que A é um álgebra central se o centro de A for $\mathbb{K}$. Dizemos que A é central simples se for central e simples.

Uma álgebra simples central, (ou simplesmente (ASC)) sobre um corpo $\mathbb{K}$, é uma álgebra associativa de dimensão finita, que é simples, e para a qual o centro é exatamente $\mathbb{K}$.

Proposição 2.72. *Seja $\mathbb{K}$ um corpo.*

1. *Se A e B são $\mathbb{K}$-álgebras centrais, então $A \otimes B$ é uma álgebra central.*
2. *Se A é simples central e B é uma álgebra que é simples, então $A \otimes B$ é simples.*
3. *Se A e B são álgebras simples centrais, então $A \otimes B$ é simples central.*

Demonstração.

1. Suponha que $z = \sum a_i \otimes b_i \in A \otimes B$ está no centro de $A \otimes B$, com $a_i's$ e $b_i's$ elementos da bases de A e B respectivamente. Como $z \in C(A \otimes B)$ temos que

$$\sum aa_i \otimes b_i = (a \otimes 1)\left(\sum a_i \otimes b_i\right) = \left(\sum a_i \otimes b_i\right)(a \otimes 1) = \sum a_i a \otimes b_i$$

então $aa_i = a_i a$ para cada i e cada $a \in A$ isso mostra que $a_i \in C(A) = \mathbb{K}$, consequentemente

$$z = \sum 1 \otimes a_i b_i = 1 \otimes \sum a_i b_i = 1 \otimes b; \text{ com } b \in B.$$

Como $1 \otimes b$ comuta com cada $1 \otimes b'$ e como $B \to 1 \otimes B$ é um isomorfismo, isso mostra que $b \in Z(B) = \mathbb{K}$. Consequentemente $z \in \mathbb{K}$.

2. Seja I um ideal bilateral não trivial em $A \otimes B$. Seja $x = a_1 \otimes b_1 + \ldots + a_n \otimes b_n$ um elemento não trivial de I onde, os b_i são linearmente independentes sobre $\mathbb{K}$. Suponha, além disso, que n é mínimo para qualquer elemento não trivial de I. Claramente $a_1 \neq 0$ então Aa_1A é um ideal não trivial em A, logo $Aa_1A = A$. Segue que se multiplicarmos u à esquerda e à direita por elementos adequados de $A \otimes 1$ e adicionar elementos de I, podemos obter um elemento de I da forma

$$u = 1 \otimes b_1 + \ldots + a_n \otimes b_n.$$

De fato, digamos que $\alpha_1 a_1 \alpha_2 = 1$, com $\alpha_1, \alpha_2 \in A$ temos que

$$u = \alpha_1 x \alpha_2 - (\alpha_1 a_2 \alpha_2 \otimes b_2 + \ldots + \alpha_1 a_n \alpha_2 \otimes b_n) + (a_2 \otimes b_2 + \ldots a_n \otimes b_n).$$

Considere o elemento

$$(a \otimes 1)u - u(a \otimes 1) = (aa_2 - a_2 a) \otimes b_2 + \ldots + (aa_n - a_n a) \otimes b_n.$$

Como este também é um elemento de I, a minimalidade de n nos diz que ele é zero, então $aa_i = a_i a$ para cada i. Portanto, cada $a_i \in Z(A) = \mathbb{K}$ e como acima $u = 1 \otimes b$ para um algum $0 \neq b \in B$. Por outro lado, temos

$$I \supseteq (1 \otimes B)(1 \otimes b)(1 \otimes B) = 1 \otimes BbB = 1 \otimes B$$

já que B é simples. Então $I \supseteq (A \otimes 1)(1 \otimes B) = A \otimes B$. Consequentemente $I = A \otimes B$.

3. Segue do item 1 e 2.

□

Proposição 2.73. *Se A é uma álgebra simples, então $Z(A)$ é um corpo e consequentemente A é uma álgebra central e simples sobre $Z(A)$.*

Demonstração. Basta mostrar que todo elemento não nulo de $Z(A)$ possui um inverso. De fato, como A tem unidade, então existe, $x \in Z(A)$ um elemento não nulo. Chame de I o ideal gerado por x, como A é simples e $I \neq 0$, segue que $I = A$.

Portanto, existe $y \in A$, tal que $xy = 1$. Como x é elemento do centro, então $xy = yx = 1$ e logo $y = x^{-1}$. Resta mostrar que $y \in Z(A)$. Para isso, note que para todo $a \in A$, tem-se que $ya = ya(xy) = (yx)ay = ay$. □

Note que qualquer álgebra simples é uma álgebra simples central sobre seu centro. O próximo teorema extraído de [8, p. 184], relaciona uma álgebra unitária simples com a sua respectiva dimensão sobre o centro.

Lema 2.74. *Uma PI-álgebra A unitária e simples é uma álgebra de dimensão finita sobre seu centro Z.*

A suposição de que A é unitário é redundante. Isso porque uma PI-álgebra simples tem um centro diferente de zero (Ver [8, p. 186]), do qual se deduz facilmente que é unitária.

Definição 2.75. Seja A uma álgebra. Para $a, b \in A$ definimos as aplicações E_a, $D_b : A \to A$, chamado mapa de multiplicação à esquerda e mapa de multiplicação á direita respectivamente, definido por

$$E_a(x) = ax \quad \text{e} \quad D_b(x) = xb.$$

Além disso, é interessante observar que E_a e D_b comutam entre si.

Observação 2.76. Note que o conjunto

$$M(A) = \{E_{a1}D_{b1} + \cdots + E_{an}D_{bn} | \text{ onde } a_1, b_i \in A; n \in \mathbb{N}\}$$

é subálgebra de $End(A)$. Além disso, a álgebra $M(A)$ é chamada de álgebra multiplicativa de A.

Um fato curioso é que dado $f \in M(A)$, existem elementos $a_i, b_i \in A$ tais que

$$f(x) = \sum_{i=1}^{n} a_i x b_i \;\; x \in A.$$

Contudo esses elementos não são únicos, f pode ser escrito como uma soma de outros elementos da forma $E_a D_b$. O fato de que esses elementos podem ser expressos de maneiras diferentes por meio de mapas de multiplicação à esquerda e à direita pode criar algum incômodo. O seguinte lema mostra que em álgebras centrais simples esse problema é controlável.

Lema 2.77. *Seja A uma álgebra simples central, e sejam $a_i, b_i \in A$ tais que $\sum_{i=1}^{n} E_{a_i} D_{b_i} = 0$. Se $a_i's$ são linearmente independentes, então cada $b_i's = 0$. Da mesma forma, se os b_i são linearmente independentes, então cada $a_i = 0$.*

Demonstração. As duas afirmações do lema são análogas, então consideramos apenas o caso em que os $a_i's$ são linearmente independentes. Suponha por absurdo que $b_n \neq 0$. Como A é simples, o ideal gerado por b_n é igual a A. Ou seja, $\sum_{j=1}^{m} w_j b_n z_j = 1$ para algum $w_j, z_j \in A$. Portanto,

$$0 = \sum_{j=1}^{m} D_{z_j} \left(\sum_{i=1}^{n} (E_{a_i} D_{b_i}) \right) D_{w_j} = \sum_{i=1}^{n} E_{a_i} \left(\sum_{j=1}^{m} (D_{w_j b_j z_j}) \right) = \sum_{i=1}^{n} E_{a_i} D_{c_i},$$

onde $c_i = \sum_{j=1}^{m} w_j b_i z_j$, então $c_n = 1$. Isso implica claramente que $n > 1$. Podemos assumir que n é o menor número natural para o qual o lema não é válido. Como

$$0 = \left(\sum_{i=1}^{n} (E_{a_i} D_{c_i}) \right) D_x - D_x \left(\sum_{i=1}^{n} (E_{a_i} D_{c_i}) \right) = \sum_{i=1}^{n-1} E_{a_i} D_{x c_i - c_i x}.$$

Para cada $x \in A$, segue-se que $x c_i - c_i x = 0$. Consequentemente, c_i pertence ao centro de A. Como A é central, segue que $c_i \in \mathbb{K}$. Mas então

$$0 = \sum_{i=1}^{n} E_{a_i} D_{c_i} = E_{c_1 a_1 + \cdots + c_n a_n},$$

com $c_n = 1$. O que contradiz a independência linear dos $a_i's$. Portanto $b_n = 0$, como queríamos demonstrar.

□

No próximo lema, suponha que álgebra A tem dimensão finita n, logo a dimensão de $End(A)$ é n^2.

Lema 2.78. *Se A é uma álgebra simples central de dimensão finita, então $M(A) = End(A)$.*

Demonstração. Seja $\{u_1, \cdots, u_n\}$ uma base de A. Pelo lema anterior implica que $E_{u_i} D_{u_j}$ são

linearmente independente, com $i, j \in \{1, \ldots, n\}$. De fato, podemos escrever

$$\sum_{i,j=1}^{d} \lambda_{ij} E_{u_i} D_{uj} \quad \text{como} \quad \sum_{i=1}^{d} E_{u_i} D_{b_i} \text{ onde } b_i = \sum_{j=1}^{d} \lambda_{ij} u_j$$

Portanto dimensão $M(A) \geqslant n^2$ que é igual a dimensão $End(A)$. Obrigatoriamente temos que $M(A) = n^2$ e logo $M(A) = End(A)$. □

Veremos agora um corolário que reafirma o teorema de Wedderburn para álgebras sobre corpos algebricamente fechados, para os quais esses teoremas assumem formas extremamente simples. O seguinte resultado foi provado em 1892 por T. Molien e sua demonstração pode ser encontrada, em [8, p. 45].

Teorema 2.79. *Seja A uma álgebra semiprima de dimensão finita diferente de zero sobre um corpo algebricamente fechado $\mathbb{K}$. Então*

$$A \simeq M_{n_1}(\mathbb{K}) \times \cdots \times M_{n_r}(\mathbb{K})$$

para algum $n_1, \cdots, n_r \in \mathbb{N}$. Mais que isso, se A é prima então $r = 1$

Vejamos agora, um teorema sobre álgebras simples centrais de dimensão finita resultante dos lemas anteriores.

Teorema 2.80. *Seja A uma álgebra central simples de dimensão finita. Se $\overline{\mathbb{K}}$ é o fecho algébrico de $\mathbb{K}$, então $A \otimes \overline{\mathbb{K}} \simeq M_n(\overline{\mathbb{K}})$, em que $n = \sqrt{dim(A)}$.*

Demonstração. Seja I um ideal de $A \otimes \overline{\mathbb{K}}$ e um elemento não nulo $u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_d e_d \in I$, com $\alpha_i \in \mathbb{K}$ e $u_i \in A$ elementos da base de A. Sem perda de generalidade suponha que $\alpha_1 \neq 0$. Tome $\phi \in End(A)$ tal que $\phi(e_1) = 1$ e $\phi(e_i) = 0$ $i \geqslant 2$. Pelo Lema 2.78 existe $a_i, b_i \in A$ tal que

$$\phi = \sum_{i=1}^{r} E_{a_i} D_{b_i}.$$

Portanto

$$\sum_{i=1}^{r} a_i u b_i = \sum_{j=1}^{d} \alpha_j \left(\sum_{i=1}^{r} a_i e_j b_i \right) = \sum_{j=1}^{d} \phi(e_j) = \alpha_1 \Longrightarrow \alpha_1 \in I$$

Mas pela Observação 2.16, temos que $I = A$, logo A é uma álgebra simples e portanto é prima. Pelo Teorema 2.79 tem-se que $A \simeq M_n(\overline{\mathbb{K}})$. Onde $dim_{\mathbb{K}}(A) = dim_{\overline{\mathbb{K}}} A \otimes \overline{\mathbb{K}} = n^2$, e assim segue a igualdade $n = \sqrt{dim(A)}$. □

Seja A uma $\mathbb{K}$-álgebra. Então, o conjunto $\mathbb{K}\times A$ torna-se uma $\mathbb{K}$ álgebra, que denotamos por A^*, se definirmos adição, multiplicação escalar e produto da seguinte forma:

$$(\lambda,x)+(\eta,y)=(\lambda+\eta,x+y)$$
$$\eta(\lambda,x)=(\eta\lambda,\eta x)$$
$$(\lambda,x)(\eta,y)=(\lambda\eta,\eta x+\lambda y+xy)$$

Consideramos A como uma subálgebra de A^* por meio da incorporação $x\to(0,x)$. Observe que A é na verdade um ideal de A^*. Uma observação crucial para nós é que A^* é uma álgebra unitária. Na verdade, $(1,0)$ é a sua unidade.

Definição 2.81. A álgebra A^* é chamada de unitização de A.

Esta construção destina-se principalmente a álgebras (e anéis) sem unidade. A ideia por trás disso é poder reduzir alguns problemas de álgebras gerais para álgebras unitárias; isso nem sempre funciona. Em princípio, pode-se construir A^* mesmo quando A é unitário. Isso pode parecer um pouco artificial à primeira vista, especialmente porque a unidade de A é diferente da unidade de A^* . No entanto, as construções de novos anéis e álgebras dos antigos às vezes se revelam inesperadamente úteis, digamos, na busca de contraexemplos.

A unitização de A não preserva todas as propriedades de A. Por exemplo, a simplicidade definitivamente não é preservada, pois A é um ideal de A^* . Se A é um álgebra unitária prima não nula, então A^* não é prima pois $I=\{(\lambda,-\lambda)|\lambda\in\mathbb{K}\}$ é um ideal de A^* tal que $IA=AI=0$. O caso não unitário é diferente.

Lema 2.82. *Se A é uma álgebra prima sem unidade, então A^* também é prima.*

Demonstração. Seja $(\lambda,a),(\mu,b)\in A^*$ que satisfaz $(\lambda,a)A^*(\mu,b)=0$. Então $(\lambda,a)(1,0)(\mu,b)=0$, e consequentemente tem-se que, $\lambda=0$ ou $\mu=0$.

Consideremos o caso em que $\lambda=0$, pois o caso em que $\mu=0$ pode ser tratado de forma semelhante. Podemos supor que $a\neq 0$, de $(0,a)(0,x)(\mu,b)=0$ inferimos que $\mu ax+axb=0$ para cada $x\in A$.

Podemos agora também assumir que $\mu\neq 0$, pois caso contrário $aAb=0$ e portanto $b=0$, como desejado. Definindo $e=-\mu^{-1}b$, temos que $ax=ax$ para todo $x\in A$. Assim $a(xy)e=axy=(axe)y$ para todo $x,y\in A$. Podemos reescrever isso como

$$ax(y-ye)=0=ax(y-ey).$$

Como A é primo, segue que $y = ye$ e $y = ey$ para todo $y \in A$. Isso contradiz a suposição de que A não é unitária. □

2.4 Identidades Polinomiais

O principal objetivo desta secção será definir e estudar algumas propriedades das identidades polinomiais sobre uma álgebra A. As principais referência deste capítulo são os livros [11] e [8].

Definição 2.83. Sejam A uma álgebra e $f = f(x_1, x_2, \ldots, x_n) \in \mathbb{K}\langle X\rangle$. Dizemos que f é uma identidade polinomial de A (ou que A satisfaz f), se $f(a_1, \ldots, a_n) = 0$, para quaisquer $a_1, \ldots, a_n \in A$. A álgebra A é dita ser uma PI-álgebra se satisfaz alguma identidade polinomial não nula.

Observação 2.84. Como $\mathbb{K}\langle X\rangle$ é uma álgebra livre, é fácil verificar que, $f \in \mathbb{K}\langle X\rangle$ é uma identidade polinomial para A, se e somente se, f estiver no núcleo de todos os homomorfismos $\mathbb{K}\langle X\rangle \to A$.

Exemplo 2.85. Toda álgebra comutativa é uma PI-álgebra, já que satisfaz o polinômio $p(x_1, x_2) = x_1x_2 - x_2x_1$.

Exemplo 2.86. Toda álgebra A de dimensão finita n satisfaz o polinômio standard

$$s_{n+1} = \sum_{\sigma \in S_{n+1}} (-1)^{\sigma} x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(n+1)}$$

onde S_n é o grupo das permutações dos números naturais $\{1, \cdots, n\}$ e $(-1)^{\sigma}$ é o sinal da permutação σ.

De fato, note que como o polinômio em questão é multilinear, basta verificar que o mesmo se anula para elementos da base de A. Chame de $B = \{u_1, u_2, \cdots, u_n\}$ uma base de A, ao substituirmos cada variável por um elemento da base, algum u_k aparecerá pelo menos duas vezes em cada monômio. Sem perda de generalidade digamos que u_1 aparece duas vezes na variável x_l e x_m, onde $1 \leqslant l, m \leqslant n+1$. Então, para cada $\sigma \in S_{n+1}$ os monômios associados às permutações σ e $(lk)\sigma$ fornecem o mesmo resultado mas com o sinal trocado, e portanto a soma de tais parcelas se anulam.

Como número de permutações é sempre par ($(n+1)!$ permutações) todos os monômios estão sendo contados. Temos então que $s_{n+1}(u_{x_1}, \ldots, v_{x_{n+1}}) = 0$ com $u_i's \in B$ e portanto A satisfaz $s_{n+1}(x_1, \ldots, x_{n+1})$. Em particular, $M_n(\mathbb{K})$ satisfaz s_{n^2+1}. Entretanto, esta também satisfaz uma identidade standard de grau menor, como vemos no próximo exemplo.

Exemplo 2.87. A álgebra $M_n(\mathbb{K})$ das matrizes de ordem n satisfaz a identidade standard de grau $2n$. Este resultado é conhecido como teorema de Amitsur-Levitzki. Além disso, $M_n(\mathbb{K})$ não satisfaz qualquer identidade polinomial de grau inferior a $2n$ e qualquer identidade polinomial de grau $2n$ de $M_n(\mathbb{K})$ é um múltiplo escalar de S_{2n}.

Se A é uma álgebra associativa e a, b são elementos desta álgebra, definimos o comutador de a, b por

$$[a, b] = ab - ba.$$

E o produto de Jordan de a e b como

$$a \circ b = ab + ba.$$

Já o comutador de comprimento $n \geq 3$ como sendo

$$[a_1, \ldots, a_{n-1}, a_n] = [[a_1, \ldots, a_{n-1}], a_n]$$

Se X e Y são subespaços de uma álgebra A, escrevemos $[X, Y]$ para a espaço vetorial gerado por todos os comutadores da forma $[x, y]$ onde $x \in X$ e $y \in Y$. UM subespaço vetorial L de A é chamado de ideal de Lie de A se $[L, A] \subseteq L$.

Exemplo 2.88. O espaço das matrizes sem traço, denotado por $sl_n(\mathbb{K})$ é igual a $[M_n(\mathbb{K}), M_n(\mathbb{K})]$.

De fato, observe que o conjunto das matrizes unitárias

$$\{e_{i,j} \text{ tal que } i, j = 1, \ldots, n, i \neq j\} \bigcup \{e_{i,i} - e_{i+1,i+1} \text{ tal que } i = 1, \ldots, n-1\}$$

é uma base de $sl_n(\mathbb{K})$. Já que $e_{i,j} = [e_{i,j}, e_{j,j}]$ para $i \neq j$, e por outro lado, $e_{i,i} - e_{i+1,i+1} = [e_{i,i+1}, e_{i+1,i}]$, segue que $sl_n(\mathbb{K}) \subseteq [M_n(\mathbb{K}), M_n(\mathbb{K})]$. A inclusão contrária segue do fato que o traço do comutador de duas matrizes é igual a zero. Portanto, a igualdade nos garante que toda matriz de traço zero pode ser escrita como combinação linear de comutadores.

Se V é um espaço vetorial e W é um subconjunto de V , então $span(W)$ denotará o subespaço de V gerado por W.

No próximo teorema provaremos que o subespaço gerado pela imagem de um polinômio é um ideal de Lie.

Teorema 2.89. *Sejam $\mathbb{K}$ um corpo infinito, A uma álgebra sobre $\mathbb{K}$, $L_1, \ldots, L_m$ ideais de Lie de A e $f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$. Então, $span(f(L_1, \ldots, L_m))$ é um ideal de Lie de A.*

Demonstração. Escreva $f = f_0 + f_1 + \ldots + f_k$, em que f_i é a componente homogênea de grau i de f na variável x_1. Logo,

$$f(\lambda a_1, a_2, \ldots, a_m) = \sum_{i=0}^{k} \lambda^i f_i(a_1, \ldots, a_m)$$

quaisquer que sejam $\lambda \in \mathbb{K}$ e $a_j \in L_j, j \in \{1, \ldots, m\}$.

Pelo Lema 2.70 , cada $f_i(a_1, \ldots, a_m) \in span(f(L_1, \ldots, L_m))$. Portanto, podemos supor que f é multihomogêneo de multigrau $(k_1, \ldots, k_m)$. Segue que existe $h \in K\langle X\rangle$ multilinear tal que

$$f = h(x_1, \ldots, x_1, \ldots, x_m, \ldots, x_m)$$

em que cada x_i ocorre k_i vezes.

Observamos também que $f(a_1 + \lambda a_1', a_2, \ldots, a_m)$ será uma soma da forma $\sum_{i=0}^{k} \lambda^i c_i$ em que $a_1, a_1' \in L_1$, $a_2 \in L_2, \ldots, a_m \in L_m$. Em particular,

$$\begin{aligned} c_1 =& h(a_1', a_1, \ldots, a_1, \ldots, a_m, \ldots, a_m) \\ &\ldots + h(a_1, \ldots, a_1, a_1', \ldots, a_m, \ldots, a_m) \end{aligned}$$

Novamente pelo Lema 2.70 temos que $c_1 \in span(f(L_1, \ldots, L_m))$.

Para qualquer elemento $b \in A$, temos

$$\begin{aligned} [f(a_1, \ldots, a_m), b)] =& h([a_1, b], a_1, \ldots, a_1, \ldots, a_m, \ldots, a_m) \\ &+ \cdots + h(a_1, \ldots, a_1, [a_1, b], \ldots, a_m, \ldots, a_m) \\ &+ \cdots + h(a_1, \ldots, a_1, \ldots, [a_m, b], a_m \ldots, a_m) \\ &+ \cdots + h(a_1, \ldots, a_1, a_m, \ldots, a_m, [a_m, b]). \end{aligned}$$

Portanto, a soma das k_1 primeiras parcelas da equação anterior pertence a $span(f(L_1, \ldots, L_m))$. O c_1 produzido por $f(a_1, a_2 + \lambda a_2', a_3, \ldots, a_m)$ nos garante que a soma das próximas k_2 parcelas da equação anterior também pertence à $span(f(L_1, \ldots, L_m))$. Continuando com esse processo, obtemos $[f(a_1, \ldots, a_m), b)] \in span(f(L_1, \ldots, L_m))$.

□

Definição 2.90. Sejam A uma álgebra e $f(x_1, \ldots, x_n) \in \mathbb{K}\langle X\rangle$. Dizemos que f é um polinômio central para A, se f tem termo constante nulo e $f(z_1, \ldots, z_n) \in Z(A)$ para quaisquer $z_1, \cdots, z_n \in A$.

Observação 2.91. Dizer que f é um polinômio central para A significa dizer que $[f, g]$ é uma

identidade para A para todo $g \in \mathbb{K}\langle X\rangle$

Exemplo 2.92. O polinômio $f(x_1,x_2) = [x_1,x_2]^2$ é central para $M_2(\mathbb{K})$.

De fato, tem-se que o polinômio característico de uma matriz M é dado por $p(\lambda) = \lambda^2 - \mathrm{tr}(M)\lambda + det(M)$ e então pelo teorema de Cayley-Hamilton temos que

$$p([x_1,x_2]) = [x_1,x_2]^2 + det([x_1,x_2])Id = 0$$

já que $\mathrm{tr}([x_1,x_2]) = 0$, logo

$$[x_1,x_2]^2 = -det([x_1,x_2])Id,$$

que é um múltiplo da matriz identidade e, portanto, um elemento do centro de $M_n(\mathbb{K})$.

Corolário 2.93. *A álgebra $M_n(\mathbb{K})$ não possui polinômios centrais de grau inferior a $2n$.*

Demonstração. Suponha que $f(x_1,\ldots,x_m)$ seja um polinômio central multilinear para $M_n(\mathbb{K})$. Então pela observação anterior, o comutador $g = [f(x_1,\ldots,x_m),x_{m+1}]$ é uma identidade polinomial para $M_n(\mathbb{K})$. Como consequência do teorema de Amitsur-Levitszky, $m+1 \geqslant 2n$, o que significa que $m \geqslant 2n-1$. Se $m = 2n-1$, então g é uma identidade polinomial de grau $2n$, e deve ser um múltiplo escalar de S_{2n}, mas escrevendo g como uma soma de monômios diferentes de zero nos dão no máximo $2(2n-1)!$ somas, enquanto em S_{2n} temos $(2n)!$ somas. Uma contradição. Logo, $m \geqslant 2n$. □

A seguir abordaremos algumas propriedades sobre imagens de polinômios multilineares sobre as matrizes, que serão úteis quando tratarmos sobre as conjecturas de Lvov-Kaplansky e a conjectura de Mesyan.

Proposição 2.94. *Seja A uma álgebra e*

$$f(x_1,\ldots,x_m) = \sum_{\sigma\in S_n} \alpha_\sigma x_{\sigma(1)}\cdots x_{\sigma(m)}$$

um polinômio multilinear. Sejam também $a_1,\ldots,a_m \in A$. Então

1. *$f(A)$ é fechado por multiplicação por escalar A, isto é, para qualquer $\alpha \in \mathbb{K}$ temos $\alpha f(a_1,\ldots,a_m) = f(\alpha a_1,\ldots,a_m)$;*
2. *$f(A)$ é fechado para conjugações por elementos invertíveis, ou seja, para qualquer $b \in A$ invertível, tem-se que $bf(a_1,\ldots,a_m)b^{-1} = f(ba_1b^{-1},\ldots,ba_mb^{-1})$*
3. *Se $\sum_{\sigma\in S_m}\alpha_\sigma \neq 0$ então $f(A) = A$.*

Demonstração. A demonstração dos itens 1 e 2 são triviais. Para provar item 3, basta observar que dado $X \in A$ temos que

$$X = \left(\sum_{\sigma \in S_n} \alpha_\sigma \right)^{-1} f(X, 1, 1, \ldots, 1) \in f(A)$$

□

Se $f(x_1, \ldots, x_m) \in \mathbb{K}\langle x_1, \ldots, x_n \rangle$ é multilinear, então, para cada $l \in \{1, \ldots, m\}$, a derivada parcial formal de f em relação à variável x_l (que é um polinômio multilinear) será $f(x_1, \ldots, x_{l-1}, 1, x_{l+1}, \ldots, x_m)$. Observamos que se A é uma $\mathbb{K}$-álgebra com unidade, então a imagem de cada derivada parcial de f sobre A está contida em $f(A)$.

O seguinte corolário descreve as imagens de polinômios multilineares de grau 2 sobre a álgebra das matrizes.

Corolário 2.95. *Seja $\mathbb{K}$ um corpo qualquer e $f(x_1, x_2) \in \mathbb{K}\langle X \rangle$ um polinômio multilinear. Então, $f(M_n(\mathbb{K}))$ é 0, $sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$.*

Demonstração. Seja $f(x_1, x_2) = \alpha x_1 x_2 + \beta x_2 x_1 \in \mathbb{K}\langle X \rangle$. Se $\alpha + \beta \neq 0$, então pela Proposição 2.94 parte 3, temos que $f(M_n(\mathbb{K})) = M_n(\mathbb{K})$.

Caso contrário, se $\alpha + \beta = 0$, então podemos ter $\alpha = \beta = 0$ e, trivialmente tem-se que $f(M_n(\mathbb{K})) = 0$, ou ainda que $\alpha = -\beta \neq 0$, o que nos fornece $f(x_1, x_2) = \alpha[x_1, x_2]$, pelo Teorema de Shoda, Albert and Muckenhoupt, $f(M_n(\mathbb{K})) = sl_n(\mathbb{K})$. □

Definição 2.96. Um ideal I de $\mathbb{K}\langle X \rangle$ é um T-ideal se é invariante por todos os endomorfismos de $\mathbb{K}\langle X \rangle$, ou seja, $f(g_1, \ldots, g_n) \in I$ para quaisquer $f(x_1, \ldots, x_n) \in I$ e $g_1, \ldots, g_n \in \mathbb{K}\langle X \rangle$.

Se S é um PI-álgebra sobre $\mathbb{K}$, as identidades de S constituem um T-ideal diferente de zero em $\mathbb{K}\langle X \rangle$, chamamos este ideal de ideal de identidades de S, e denotamos por $T(S)$. Será dito ainda que as álgebras A e B são PI-equivalentes se $T(A) = T(B)$.

A seguir, considere Q um T-ideal de $K\langle X \rangle$.

Proposição 2.97. *Seja $f(x_1, \ldots, x_n) \in Q$, tal que*

$$f(x_1, \ldots, x_n) = \sum_{i=0}^{n} f_i(x_1, \ldots, x_n)$$

onde cada f_i é um componente homogêneo de x_1 e i é o seu grau. Temos então que $f_i \in Q$ para todo i.

Demonstração. Seja $V = \langle f \rangle^T$ o T-ideal de $K\langle X\rangle$ gerado por f. Nós escolhemos $n+1$ elementos diferentes $\alpha_0, \ldots, \alpha_n$ de K. Como V é um T-ideal,

$$f(\alpha_j x_1, x_2, \ldots, x_m) = \sum_{i=0}^{n} \alpha_j^i f_i(x_1, x_2, \ldots, x_m) \in V,\ j = 0, 1 \ldots n.$$

Pelo Lema 2.70, obtemos que cada $f_i(x_1, \ldots, x_m)$ também pertence a V

□

A seguir demonstraremos um resultado que associa cada conjunto de identidades a um T-ideal.

Proposição 2.98. *O conjunto $T(A)$ das identidades de uma álgebra A é um T-ideal de $\mathbb{K}\langle X\rangle$. Reciprocamente, se I é um T-ideal de $\mathbb{K}\langle X\rangle$, então existe alguma álgebra A tal que $T(B) = I$.*

Demonstração. Mostraremos primeiro que $T(A)$ é um T-ideal. Seja ϕ um endomorfismo de $\mathbb{K}\langle X\rangle$ e f um elemento de $T(A)$, devemos mostrar que f é invariante por esse endomorfismo. De fato, tome $\Omega : \mathbb{K}\langle X\rangle \to A$ um homomorfismo qualquer. Note que $\Omega(\phi(f)) = (\Omega \circ \phi)(f)$ mas $(\Omega \circ \phi)$ é um homomorfismo de álgebras e $f \in T(A)$ logo $\phi(f) \in N(\Omega)$ e portanto $\phi(f) \in T(A)$.

Por outro lado, seja I um T-ideal de $\mathbb{K}\langle X\rangle$, consideremos a álgebra quociente $B = \mathbb{K}\langle X\rangle / I$ e a projeção canônica $\pi : \mathbb{K}\langle X\rangle \to B$. Se $f \in T(B)$, então $f \in ker(\pi)$. Como $ker(\pi) = I$, temos que $f \in I$ e consequentemente $T(B) \subseteq I$.

Para mostrar a inclusão contrária, tome $f(x_1, \cdots, x_n) \in I$ e $g_1, \cdots, g_n \in \mathbb{K}\langle X\rangle$, segue que $f(g_1, \cdots, g_n) \in I$ e daí $f(g_1 + I, \cdots, g_n + I) = f(g_1, \cdots, g_n) + I = 0 + I$. Assim, $I \subseteq T(B)$ e disso segue o resultado. □

É fácil ver que a intersecção de uma família qualquer de T-ideais é um T-ideal. Logo a definição abaixo faz sentido.

Definição 2.99. Seja S um subconjunto de $\mathbb{K}\langle X\rangle$. O T-ideal gerado por S, denotado por $\langle S\rangle^T$, é a interseção de todos os T-ideais de $\mathbb{K}\langle X\rangle$ que contém S. Dessa forma, $\langle S\rangle^T$ é o menor T-ideal contendo S.

Proposição 2.100. *Se $S \subseteq \mathbb{K}\langle X\rangle$ e $\langle S\rangle^T$ é o T-ideal gerado por S, então $\langle S\rangle^T$ é exatamente o subespaço de $\mathbb{K}\langle X\rangle$ gerado por*

$$M = \{g_0 f(g_1, \cdots, g_n) g_{n+1} \text{ tal que } f \in S;\ g_0, \cdots, g_{n+1} \in \mathbb{K}\langle X\rangle\}$$

Demonstração. Chame de V_1 o subespaço de $\mathbb{K}\langle X\rangle$ gerado por M. Para mostrar que $\langle S\rangle^T \subseteq V_1$, mostraremos que V_1 é um T-ideal. De fato, tome $g \in M$ e $\varphi \in End(\mathbb{K}\langle X\rangle)$, temos que $\varphi(g) \in M$, logo V_1 é um T-ideal que contém S, mas por definição $\langle S\rangle^T$ é o menor T-ideal com tal propriedade. □

É fácil ver que o produto de identidades polinomiais continua sendo uma identidade polinomial para uma álgebra A. O objetivo agora é mostrar que sobre a álgebra das matrizes, o inverso é verdade pelo menos em uma das parcelas.

Definição 2.101. Seja f uma identidade polinomial de uma álgebra A, dizemos que f é estável (para A) se f é uma identidade de $C \otimes A$ para cada álgebra comutativa C.

Lema 2.102. *Seja $f = f(x_1, \cdots, x_n)$ uma identidade de uma álgebra A. Se f é uma identidade de $\mathbb{K}[\omega_1, \cdots, \omega_n] \otimes A$ para cada $n \in \mathbb{N}$, então f é estável para A.*

Demonstração. Seja C uma álgebra unitária comutativa. Como cada álgebra comutativa é uma subálgebra de uma álgebra unitária comutativa (pela unitização), é suficiente mostrar que f é uma identidade de $C \otimes A$. Chame

$$\overline{x_i} = \sum_{j=1}^{m_i} c_{ij} \otimes x_{ij} \ \in C \otimes A \ \text{ onde } \ i = 1, \cdots, n$$

E seja

$$\tilde{x_i} = \sum_{j=1}^{m_i} \omega_{ij} \otimes x_{ij} \in F[\Omega] \otimes A, \ \text{ onde } \ \Omega = \{\omega_{ij} \mid i = 1, \ldots, n; j = 1, \cdots, m_i\}.$$

Por hipótese temos que $f(\tilde{x_1}, \cdots, \tilde{x_n}) = 0$. Portanto, se φ é o homomorfismo de $F[\Omega]$ para C, que leva ω_{ij} para c_{ij}, então $f(\overline{x_1}, \cdots, \overline{x_n}) = (\varphi \otimes id)(f(\tilde{x_1}, \cdots, \tilde{x_n})) = 0$ □

O próximoz teorema assim como o lema anterior foi extraído de [8, p. 150].

Teorema 2.103. *Seja $\mathbb{K}$ um corpo infinito. Então, cada identidade polinomial de uma $\mathbb{K}$-álgebra A é estável.*

Demonstração. Seja $f = f(x_1, \ldots, x_n)$ uma identidade de A. Escolha $s \in \mathbb{N}$ e tome elementos arbitrários

$$\tilde{x_i} = \sum_j p_{ij} \otimes x_{ij} \in \mathbb{K}[\omega_1, \ldots, \omega_s] \otimes A, \ \ i = 1, \ldots, n. \tag{2.1}$$

Pelo lema anterior é suficiente mostrar que $f(\tilde{x_1},\ldots,\tilde{x_n}) = 0$. Suponha que isso não seja verdade. Note primeiro que todo elemento $0 \neq x \in U \otimes V$ pode ser escrito como $x = \sum_{i=1}^{d} u_i \otimes v_i$ onde $u_1,\ldots,u_d$ são vetores linearmente independentes em U e $v_1,\ldots,v_d$ são vetores linearmente independentes em V. Sabendo disso, podemos escrever

$$f(\tilde{x_1},\ldots,\tilde{x_n}) = \sum_k q_k \otimes a_k, \text{ onde } q_1 \neq 0 \text{ e } a_k's \text{ são l.i.}$$

Já que $\mathbb{K}$ é infinito, pelo Teorema 2.111, existem $\lambda_1,\ldots,\lambda_s \in \mathbb{K}$ tal que $q_1(\lambda_1,\ldots,\lambda_s) \neq 0$. Vamos definir $\alpha_{ij} := p_{ij}(\lambda_1,\ldots,\lambda_s)$ e $\beta_k := q_k(\lambda_1,\ldots,\lambda_s)$; assim, $\beta_1 \neq 0$. Substituindo ω_i por λ_i na equação 2.1 obtemos

$$f\left(\sum_j \alpha_{1j} \otimes x_{1j}, \cdots, \sum_j \alpha_{nj} \otimes x_{nj}\right) = \sum_k \beta_k \otimes a_k.$$

Como $\alpha_{ij}, \beta_k \in F$ temos que para todo i

$$\sum_j \alpha_{ij} \otimes x_{ij} = 1 \otimes \left(\sum_j \alpha_{ij} x_{ij}\right) \quad \text{e} \quad \sum_k \beta_k \otimes a_k = 1 \otimes \left(\sum_k \beta_k \alpha_k\right).$$

A identidade acima pode ser reescrita como

$$1 \otimes f\left(\sum_j \alpha_{1j} x_{1j}, \cdots, \sum_j \alpha_{nj} x_{nj}\right) = 1 \otimes \left(\sum_j \beta_k \alpha_k\right).$$

O que é uma contradição. A saber, o lado esquerdo é 0, uma vez que f é uma identidade de A, enquanto o lado direito não é 0, uma vez que $\beta_1 \neq 0$ e os $a_k's$ são linearmente independentes. □

Observação 2.104. O ideal $T(M_n(\mathbb{K}))$ é evidentemente também o ideal de identidades do anel das matrizes $M_n(R)$ de ordem n sobre qualquer anel comutativo R que possui unidade. Agora, se A é qualquer álgebra simples e central de ordem n^2 sobre seu centro F, pode-se encontrar uma extensão $\overline{F}$ de F tal que $A \otimes \overline{F} \cong M_n(\overline{F})$. Em vista desses fatos, o lema anterior implica que o ideal de identidades de qualquer álgebra simples central A de ordem n^2 sobre seu centro também está em $T(M_n(\mathbb{K}))$.

Proposição 2.105. *A única álgebra de divisão e com dimensão finita sobre um corpo algebricamente fechado $\mathbb{K}$ é o próprio $\mathbb{K}$.*

Demonstração. Seja D uma álgebra de dimensão finita sobre $\mathbb{K}$ e de dimensão finita. Dado

$a \in D$, temos que as potências de a são linearmente dependentes sobre $\mathbb{K}$, seja

$$f(a) = a^n + c_1 a^{n-1} + \cdots + c_n = 0 \ \ (c_i \in \mathbb{K})$$

o polinômio de menor grau que se anula em a. Como $\mathbb{K}$ é algebricamente fechado, $f(x)$ tem uma raiz $\lambda \in \mathbb{K}$, e consequentemente $f(x)$ escreve-se como $f(x) = (x-\lambda)g(x)$, onde g é um polinômio de grau $n-1$.

Temos que $f(a) = (a-\lambda)g(a)$, e pela minimalidade de n, tem-se que $g(a) \neq 0$, logo $a - \lambda = 0$, o que implica que $a \in \mathbb{K}$ e portanto $D = \mathbb{K}$. □

Definição 2.106. Chamamos um corpo $\mathbb{H}$ de corpo de decomposição de uma álgebra simples central A, se $A \otimes \mathbb{H}$ for isomorfo a um anel de matriz sobre o corpo $\mathbb{H}$.

Exemplo 2.107. Cada álgebra simples central A de dimensão finita tem um corpo de decomposição.

De fato, chame $\overline{\mathbb{K}}$ o fecho algébrico de $\mathbb{K}$ e A uma álgebra simples central sobre o corpo $\mathbb{K}$. Temos que $A \otimes \overline{\mathbb{K}}$ é uma álgebra simples e central sobre $\overline{\mathbb{K}}$. O teorema de Wedderburn afirma que existe uma $\overline{\mathbb{K}}$-álgebra de divisão D de dimensão finita tal que $A \otimes \overline{\mathbb{K}} \cong M_n(D)$.

Uma vez que $\overline{\mathbb{K}}$ é algebricamente fechado, não há alguma álgebra de divisão de dimensão finita (Proposição 2.105), logo deve ocorrer de $D = \overline{\mathbb{K}}$. Portanto o fecho algébrico de $\mathbb{K}$ é um corpo de decomposição de A.

O próximo teorema já mencionado anteriormente, é um clássico, conhecido como teorema de Amitsur, ele afirma que o produto de dois polinômios é uma identidade da álgebras das matrizes apenas quando um deles é uma identidade. Sua demonstração pode ser encontrada em [4].

Teorema 2.108. *Seja $M_n(\mathbb{K})$ o anel de todas as matrizes quadradas de ordem n sobre $\mathbb{K}$. Se $M_n(\mathbb{K})$ satisfaz a identidade $f(x)g(x) = 0$ com $x = (x_1, \ldots, x_n)$, então ou $f(x) = 0$ ou $g(x) = 0$ é uma identidade polinomial.*

Neste momento faremos um breve estudo das matrizes genéricas, e enunciaremos alguns resultados importantes sobre elas.

Definição 2.109. Sejam $n \geqslant 2$ um inteiro, $\mathbb{K}$ um corpo infinito e $\mathbb{K}[y_{pq}^l \mid p, q \in \{1, 2, \cdots, n\}, l \in \mathbb{N}]$ uma álgebra de polinômios com variáveis comutativos. Chamamos de matrizes genéricas

elementares as matrizes

$$y_i = \begin{bmatrix} y_{11}^i & \cdots & y_{1,n}^i \\ y_{21}^i & \cdots & y_{2,n}^i \\ \vdots & \ddots & \vdots \\ y_{n1}^i & \cdots & y_{n,n}^i \end{bmatrix}$$

A subálgebra $GM_n(\mathbb{K})$ de $M_n(\mathbb{K}[y_{pq}^l])$ gerada por tais matrizes é chamada de álgebra das matrizes genéricas.

Exemplo 2.110. Denotamos A_m a subálgebra de $GM_n(\mathbb{K})$ gerada pelas primeiras m matrizes genéricas $y_1, \ldots, y_m$. Tomando $n = m = 2$ e chamando $y_{pq}^1 = a_{pq}$ assim como $y_{ij}^2 = b_{ij}$ temos que A_2 é gerada pelas matrizes

$$y_1 = \begin{bmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{bmatrix} \text{ e } y_2 = \begin{bmatrix} b_{11} & b_{21} \\ b_{21} & b_{22} \end{bmatrix}$$

Teorema 2.111. *Seja $\mathbb{K}$ um corpo infinito e k um subconjunto infinito de $\mathbb{K}$, seja $n \in \mathbb{N}$ e $f \in \mathbb{K}[x_1, x_2, \cdots, x_n]$, o anel de polinômios em n indeterminados sobre k. Então $f = 0$, se e somente se, $f(c_1, c_2, \cdots, c_n) = 0$ para todos os $c = (c_1, c_2, \cdots, c_n) \in \mathbb{K}^n$.*

Demonstração. Faremos a prova por indução. Para $n = 1$, sabemos que qualquer polinômio em uma variável tem no máximo m raízes, onde m denota o grau do polinômio diferente de zero.

Então suponha que o resultado é válido para polinômios $f \in \mathbb{K}[x_1, x_2, \cdots, x_{n-1}]$, objetivo é mostrar que é válido também até n.

Seja

$$f = \sum_{i=0}^{n} f_i(x_1, \cdots, x_{n-1}) X_n^i \text{ onde } f_i \in \mathbb{K}[x_1, x_2, \cdots, x_{n-1}]$$

Agora escolha $c_1, c_2, \cdots, c_{n-1} \in k$. Então $f(c_1, \cdots, c_{n-1}, X_n)$ é um polinômio em uma variável que se anula para todos os $c_n \in k$, implicando que $f_i(c_1, \cdots, c_{n_1}) = 0$ para todos os i. Isso é verdadeiro para todos $(c_1, c_2, \cdots, c_{n-1}) \in k^{n-1}$. Assim, por hipótese, todos $f_i = 0$, ou seja, $f = 0$. A afirmação inversa é verdadeira trivialmente. □

No próximo teorema iremos provar que a álgebra das matrizes genéricas é isomorfa a $\mathbb{K}\langle X\rangle$ quociente $T(M_n(\mathbb{K}))$

Teorema 2.112. *Se $\mathbb{K}$ é um corpo infinito, então*

$$GM_n(\mathbb{K}) \simeq \mathbb{K}\langle X\rangle / T(M_n(\mathbb{K}))$$

Demonstração. Considere o homomorfismo sobrejetor

$$\varphi : \mathbb{K}\langle X\rangle \to GM_n(\mathbb{K})$$
$$x_i \mapsto y_i$$

Pelo teorema do isomorfismo basta mostrar que $ker(\varphi) = T(M_n(\mathbb{K}))$. De fato, tome $f \in ker(\varphi)$, então $f(y_1,\ldots,y_n) = \varphi(f) = 0$. Chame de $A_i = (a^l_{pq}) \in M_n(\mathbb{K})$ com $i \in 1,\ldots,m$ matrizes quaisquer.

Note que $\mathbb{K}[y^l_{pq}]$ é livre na classe das álgebras associativas comutativas (Exemplo 2.62), logo existe um homomorfismo

$$H : \mathbb{K}[y^{(l)}_{pq}] \to \mathbb{K}$$
$$y^l_{pq} \to a^l_{pq}$$

E podemos induzir, através deste, outro homomorfismo

$$\overline{\psi} : GM_n(\mathbb{K}) \to M_n(\mathbb{K})$$
$$y_i \to A_i$$

Com $i \in \{1,\ldots,m\}$ e $\overline{\psi}(y_l) = 0$ para $l > m$. Segue que

$$f(A_1,\ldots,A_m) = \overline{\psi}(f(y_1,\ldots,y_m)) = 0$$

Logo $ker(\varphi) \subseteq T(M_n(\mathbb{K})$;

Por outro lado, seja $g(x_1,\ldots,x_n) \in T(M_n(\mathbb{K}))$. Então para quaisquer matrizes A_i, temos que $0 = g(A_1,\ldots,A_m) = \overline{\psi}(g(y_1,\ldots,y_m))$. Observe que $\overline{\psi}(g(y_1,\ldots,y_m))$ é uma matriz, onde cada entrada da matriz é um polinômio comutativo com variáveis y^i_{pq}, que se anulam em $\mathbb{K}$, mas pelo Teorema 2.111,isso implica que $g(x_1,\ldots,x_n) = 0$. Consequentemente tem-se que $g \in ker(\varphi)$ como queríamos demonstrar. □

2.5 Estruturas de Lie e Jordan em álgebras simples

Nesta secção trataremos das estruturas de um anel associativo simples e alguns de seus subconjuntos como anéis de Jordan e de Lie. Os resultados desta secção foram extraídos do livro [16] e serão úteis para os próximos resultados. Começamos a estudar certos aspectos da teoria dos anéis associativos, iniciando com o seguinte lema.

Lema 2.113. *Se L é um ideal de Lie de A, então $N(L) = \{x \in A; [x, A] \subseteq L\}$ é uma subálgebra e um ideal de Lie de A e*

$$Id([L, L]) \subseteq Id([N(L), N(L)]) \subseteq N(L),$$

onde $Id(\star)$ é o ideal gerado por $\star$.

Demonstração. Seja $x, y \in N(L)$ e $a \in A$, temos que

$$[xy, a] = [x, ya] + [y, ax] \in L. \tag{2.2}$$

Portanto, $xy \in N(L)$, então $N(L)$ é uma subálgebra de A, uma vez que $N(L)$ é fechado em relação a soma. Segue-se da definição de $N(L)$ que $[N(L), A] \subseteq L$. Como $L \subseteq N(L)$, então em particular temos que $N(L)$ é um ideal de Lie de A.

Como $L \subseteq N(L)$, basta provarmos que $Id([N(L), N(L)]) \subseteq N(L)$. Além disso, uma vez que $N(L)$ é um ideal de Lie de A, então $[N(L), N(L)]$ também é. Como

$$Id([N(L), N(L)]) = [N(L), N(L)] + A[N(L), N(L)] = [N(L), N(L)] + [N(L), N(L)]A$$

é suficiente provar que $A[N(L), N(L)] \subseteq N(L)$. De fato, sejam $a \in A$ e seja $x, y \in N(L)$, note que $a[x, y] = [ax, y] - [a, y]x$. Como $N(L)$ é um ideal de Lie e uma subálgebra de A, ambos os termos do lado direito estão em $N(L)$. Donde segue que $a[x, y] \in N(L)$ e portanto $Id([N(L), N(L)]) \subseteq N(L)$. □

Definição 2.114. Um subconjunto não vazio J da álgebra A é considerado um ideal de Jordan de A se J for um subespaço de A com a propriedade que $ja + aj \in J$, para todo $j \in J$ e $a \in A$.

O objetivo será determinar os ideais de Jordan e de Lie da álgebra A no caso em que A é uma álgebra simples. Começaremos com os ideais da Jordan, já que são mais fáceis de caracterizar.

Lema 2.115. *Se J é ideal de Jordan de A então para todo $j_1, j_2 \in J$ e $a \in A$ tem-se que $(j_1 j_2 + j_2 j_1)a - a(j_1 j_2 + j_2 j_1) \in J$.*

Demonstração. Como $j_1 \in J$, para qualquer $a \in A$ tem-se que $j_1(aj_2 - j_2 a) + (aj_2 - j_2 a)j_1 \in J$, com $j_2 \in J$. Contudo

$$\begin{aligned} j_1(aj_2 - j_2 a) + (aj_2 - j_2 a)j_1 = ((j_1 a - aj_1)j_2 + j_2(j_1 a - aj_1) \\ + (a(j_1 j_2 + j_2 j_1) - (j_1 j_2 + j_2 j_1)a)). \end{aligned}$$

Mas note que o lado esquerdo e a primeira parcela do lado direito da igualdade estão em J, donde segue que o lema de fato é verdadeiro.

□

Teorema 2.116. *Seja A uma álgebra tal que $2x = 0$ implica em $x = 0$, e suponha que, além disso, A não tem ideais nilpotentes diferentes de zero. Então, qualquer ideal de Jordan diferente de zero contém um ideal associativo diferente de zero de A*

Demonstração. Seja $J \neq 0$, um ideal de Jordan de A e suponha que $a, b \in J$. Pelo lema anterior, para qualquer $x \in A$, tem-se que $xe - ex \in J$, onde $e = ab + ba$. Contudo, como $e \in J$ então $xe + ex \in J$, consequentemente $(xe - ex) + (xe + ex) = 2xe \in J$ para todo $x \in A$.

Com isso, temos que $(2xe)y + y(2xe) \in J$, com $y \in A$, como $2yxe \in J$, segue que $2xey \in J$, ou seja, $2AeA \subseteq J$. Agora $2AeA$ é um ideal de A, restando mostrar apenas que esse ideal é diferente de zero.

De fato, suponha por absurdo que esse ideal é nulo, então por hipótese temos que $AeA = 0$. Em particular, Ae é um ideal bilateral e $(Ae)^2 = 0$. Como A não tem ideais nilpotentes, então $e = 0$ e consequentemente temos para todo $a, b \in J$ que $ab + ba = 0$.

Seja $j \in J$ não nulo, para todo $x \in A$, tomando $b = jx + xj \in J$, temos que $j(jx + xj) + (jx + xj)j = 0$, ou seja, $j^2x + xj^2 + 2jxj = 0$. Por outro lado, seja $c \in J$, então $0 = cc + cc = 2c^2$, donde segue que $c^2 = 0$.

Das duas equações anteriores segue que $2jxj = 0$, para todo $x \in A$, o que implica em $jAj = 0$. Então $jA \neq 0$ é um ideal nilpotente à direita de A, contrariando a hipótese. Logo todo ideal de Jordan contém um ideal não nulo de A.

□

O próximo teorema é um clássico para a nilpotência de álgebras. Foi provado pela primeira vez em 1953 pelo matemático japonês Masayoshi Nagata sobre um corpo de característica 0 e então em 1956 o britânico Graham Higman o provou para o caso geral. Muito mais tarde, descobriu-se que este teorema fora estabelecido pela primeira vez em 1943 pelos russos Dubnov e Ivanov, mas seu artigo foi ignorado pela comunidade matemática ocidental.

Este teorema tem muitas aplicações não apenas na teoria das PI-álgebras, mas também na teoria de invariantes (comutativa e não comutativa) e na teoria da estrutura dos anéis.

Teorema 2.117. *Seja R uma álgebra associativa sobre um corpo F e suponha que R satisfaz*

a identidade polinomial $x^k = 0$. Então R é nilpotente, ou seja, existe um inteiro $d = d(k)$ dependendo de k tal que $x_1 \cdots x_d = 0$, para todo $x_1, \cdots, x_d \in R$.

Sua demonstração será omitida mas pode ser encontrada em [11].

Definição 2.118.

1. Uma álgebra associativa A sobre um anel comutativo R é definida como uma álgebra nilpotente, se e somente, se existir algum inteiro positivo n tal que $y_1 y_2 \cdots y_n = 0$ para todo $y_1, y_2, \ldots, y_n \in A$. O menor desses n é chamado de índice de nilpotência da álgebra A.

2. Uma álgebra é dita ser nil se cada elemento é nilpotente.

As álgebras nilpotentes são trivialmente álgebras nil, enquanto álgebras nil podem não ser nilpotentes.

Por outro lado, o próximo lema, extraído de [16] afirma que em certas condições uma álgebra A possui um ideal nilpotente não nulo.

Lema 2.119. *Seja A uma álgebra e $0 \neq I$ um ideal à direita de A. Suponha que existe um inteiro $n > 1$, tal que $a \in I$ implica que $a^n = 0$. Então A tem um ideal nilpotente diferente de zero.*

Demonstração. Seja I ideal à direita não nulo, e suponha que existe $n > 1$ tal que $a^n = 0$ para todo $a \in I$. Como I é também uma álgebra associativa, pelo teorema de Nagata Higman existe um N tal que $I^N = 0$, isto é, $x_1 \cdots x_n = 0$, para todo $x_1, \ldots, x_d \in I$.

Defina $J := I + AI$, note que J é um ideal bilateral de A que contém I (e portanto é não nulo). Note que, J^N está contido em $A(I^N) + I^N$ que é zero. De fato, faremos uma prova por indução sobre n. Para $n = 2$, temos que

$$\begin{aligned} J^2 &= I^2 + IAI + AI^2 + AIAI \\ &\subseteq I^2 + I^2 + AI^2 + I^2 = I^2 + AI^2 \end{aligned}$$

Supunha que o resultado é válido para todo k até $n-1$, ou seja, $J^{n-1} \subseteq A(I^{n-1}) + I^{n-1}$. Devemos

mostrar que o mesmo vale para n. De fato,

$$\begin{aligned} J^n &= J^{n-1}J \\ &\subseteq (I^{n-1} + AI^{n-1})(I + AI) \\ &= I^n + I^{n-1}AI + AI^n + AI^{n-1}AI \\ &\subseteq I^n + I^n + AI^n + AI^n \\ &= I^n + AI^n \end{aligned}$$

Logo $J^N = 0$ e J é o ideal nilpotente não nulo pedido. □

Note que a demonstração apresentada é mais simplificada do que a original. Isso decorre do fato que estamos alterando a hipótese de R ser uma álgebra associativa, ao invés de um anel, como é tratado em [16]. Com isso, podemos aplicar o teorema de Nagata-Higman, que simplifica substancialmente a demonstração.

Definição 2.120. A característica de um corpo $\mathbb{K}$, denotada por $char(\mathbb{K})$, é definida como o menor número de vezes que se deve usar a identidade multiplicativa 1, em uma soma para obter a identidade aditiva 0. Se esta soma nunca atingir a identidade aditiva, diz-se que o corpo tem característica zero.

Observação 2.121. Uma álgebra A simples de centro nulo e $char(A) = 2$, não possui ideais de Lie comutativos diferentes de zero. De fato, chame de U um ideal de Lie comutativo e considere $u \in U$ e $x \in A$, como U é um ideal de Lie comutativo temos,

$$ux + xu \in U \Longrightarrow u(xu + ux) = (xu + ux)u \Longrightarrow u^2x = xu^2 \Longrightarrow u^2 \in Z(A).$$

Como $Z = 0$, implica que $u^2 = 0$. Com isso temos que

$$0 = (ux - xu)^2 = uxux - ux^2u + xuxu.$$

Multiplicando à direita por ux obtemos que $(ux)^3 = 0$. Portanto, uA é um ideal nilpotente de índice 3. Então pelo Lema 2.119, A tem um ideal nilpotente diferente de zero, absurdo já que A é simples. Portanto a álgebra A não possui ideais de Lie comutativos diferentes de zero.

Lema 2.122. *Seja A uma álgebra sem ideais nilpotentes não nulo em que $2x = 0$ implica em $x = 0$, para $x \in A$. Suponha que $U \neq 0$ é ideal de Lie e subálgebra de A. Então ou $U \subseteq Z(A)$ ou U contém um ideal não nulo de A.*

Demonstração. Suponhamos primeiro que U é não comutativo, então existem $x, y \in U$ tal que

$xy - yx \neq 0$. Além disso, como U é um ideal de Lie, então para $r \in A$ tem-se que $x(yr) - (yr)x \in U$.

Por outro lado $y(xr - rx) \in U$, já que tanto y como $(xr - rx)$ pertencem a U e U é subalgebra de A. Como

$$x(yr) - (yr)x = (xy - yx)r + y(xr - rx)$$

temos que $(xy - yx)A \subseteq U$. Com isso, tem-se que para qualquer $r, s \in A$

$$((xy - yx)r)s - s((xy - yx)r) \in U,$$

de onde obtemos que

$$A(xy - yx)A \subseteq U.$$

É fácil ver que $A(xy - yx)A$ é um ideal de U. Resta mostrar que de fato esse ideal é diferente de zero. Suponha por absurdo que $A(xy - yx)A = 0$, então $(A(xy - yx))^2 = 0$. Assim, $A(xy - yx)$ é ideal nilpotente, contrariando a hipótese. Logo U contém ideal diferente de zero.

Suponhamos agora que U é comutativo, queremos mostrar que $U \subseteq Z(A)$. Dado $a \in U$ e $x \in A$, temos que $ax - xa \in U$ que comuta com a. Sejam $x, y \in A$

$$a(a(xy) - (xy)a) = (a(xy) - (xy)a)a$$

Expandindo $a(xy) - (xy)a$ como $(ax - xa)y + x(ay - ya)$ pela equação anterior, e lembrando que a comuta com $(ax - xa)$ temos

$$\begin{aligned} a(axy - xya) - (axy - xya)a &= a((ax - xa)y + x(ay - ya)) - ((ax - xa)y + x(ay - ya))a \\ &= a(ax - xa)y + ax(ay - ya) - (ax - xa)ya - x(ay - ya)a \\ &= (ax - xa)ay + ax(ay - ya) - (ax - xa)ya - xa(ay - ya) \\ &= (ax - xa)(ay - ya) + (ax - xa)(ay - ya) \\ &= 2(ax - xa)(ay - ya) = 0 \end{aligned}$$

Por hipótese, temos que $2x = 0$ implica em $x = 0$, para todo $x \in A$, então $(ax - xa)(ay - ya) = 0$. Tomando $y = x$ trivialmente temos $(ax - xa)(ax - xa) = 0$. Vamos mostrar que $(ax - xa)A(ax - xa) = 0$. De fato tome $c \in (ax - xa)A(ax - xa)$, então $c = (ax - xa)\alpha(ax - xa)$,

com $\alpha \in A$. Logo

$$\begin{aligned}(ax-xa)\alpha(ax-xa) &= (ax-xa)\alpha(ax-xa) - (ax-xa)(ax-xa)\alpha \\ &= (ax-xa)\cdot[\alpha,(ax-xa)] = [\alpha,(ax-xa)](ax-xa) \\ &= \alpha(ax-xa)^2 - (ax-xa)\alpha(ax-xa) \\ &= -(ax-xa)\alpha(ax-xa)\end{aligned}$$

Como $(ax-xa)\alpha(ax-xa) = -(ax-xa)\alpha(ax-xa)$ implica que $2(ax-xa)\alpha(ax-xa) = 0$ e logo $(ax-xa)A(ax-xa) = 0$. Mas observe que tomando I como ideal gerado por $(ax-xa)\alpha$, é fácil ver que $I^2 = 0$, mas por hipótese A não tem ideais nilpotentes não nulos. Logo $(ax-xa) = 0$ implicando que $a \in Z(A)$. □

Observe que na última parte da prova do Lema 2.122 também provamos o seguinte resultado.

Corolário 2.123. *Seja A é uma álgebra sem ideal nilpotente diferente de zero em que $2x = 0$ implica $x = 0$. Se $a \in A$ comuta com todos os $ax - xa$ com $x \in A$ então a pertence ao centro de A.*

O lema anterior também implica imediatamente no seguinte corolário

Corolário 2.124. *Seja A é uma álgebra simples de características diferente de 2, então qualquer ideal de Lie de A que também é um subálgebra de A deve ser o próprio A ou está contido no centro de A.*

Teorema 2.125. *Seja A uma álgebra simples de característica diferente de 2, e seja U um ideal de Lie de A. Então $U \subseteq Z(A)$ ou $[A,A] \subseteq U$.*

Demonstração. Pelo Lema 2.113, sabemos que $N(U)$ é uma subálgebra além de ser um ideal de Lie de A, então pelo corolário anterior temos que $N(U) \subseteq Z(A)$, ou $N(U) = A$. Se $N(U) = A$ então por definição temos que $[A,A] \subseteq U$, se $N(U) \subseteq Z(A)$, então como $U \subseteq N(U)$, obtemos $U \subseteq Z(A)$.

□

Corolário 2.126. *Se A é uma álgebra simples não comutativa de características diferente de 2, então a subálgebra gerada por $[A,A]$ é A*

Demonstração. Qualquer subgrupo aditivo de A que contém $[A,A]$ é um ideal de Lie de A, trivialmente por definição. Consequentemente qualquer subálgebra gerada por $[A,A]$, também

é um ideal de lie. Pelo Corolário 2.124, essa subálgebra é igual a própria álgebra A ou está contido em $Z(A)$.

Suponha por absurdo que esta subálgebra está contida no centro $Z(A)$, logo para $a \in A$ temos que a comuta com todos $ax - xa$ com $x \in A$. Pelo Corolário 2.123, tem-se que $a \in Z(A)$, isto é, $A \subseteq Z(A)$. Como assumimos que A não é comutativo, isso é descartado, portanto a subálgebra gerado por $[A, A]$ é A.

□

Gostaríamos agora de resolver o problema do Teorema 2.125 e do Corolário 2.126, mesmo quando A tem a característica igual a 2. Observe que a característica de A não entrou na discussão na passagem do teorema em diante. Logo a primeira questão que certamente vem a cabeça é se o corpo tiver características 2, o Corolário 2.124 falha? Certamente falha em $M_2(\mathbb{K})$, onde, $\mathbb{K}$ é um corpo de característica 2. Tome

$$U = \left\{ \begin{bmatrix} a & b \\ b & a \end{bmatrix}; \text{ tal que } a, b \in \mathbb{K} \right\}$$

É fácil ver que U embora seja um ideal de Lie e subálgebra de A, não está no centro $Z = Z(A)$ da álgebra, muito menos é igual a própria álgebra.

Mesmo assim, podemos melhorar as hipóteses do Teorema 2.125. Suponha que A é uma álgebra simples de característica 2 e que U é um ideal de Lie e subálgebra de A, com $A \neq U$ e $U \nsubseteq Z$. De fato, pela demonstração do Lema 2.122, e sendo a álgebra simples, U deve ser comutativo. Isto é, dado u e $v \in U$, então $uv + vu = 0$.

Seja $a \in U$, então $as + sa \in U$ para todo $s \in A$, consequentemente $a(as + sa) = (as + sa)a$, isso implica que $a^2 \in Z$. Como para qualquer $r \in A$, $ar + ra \in U$ temos analogamente que para todo $r \in A$ que $(ar + ra)^2 \in Z$. De fato:

$$\begin{aligned} (ar+ra)((ar+ra)s + s(ar+ra)) &= ((ar+ra)s + s(ar+ra))(ar+ra) \Longrightarrow \\ (ar+ra)^2 s + (ar+ra)s(ar+ra) &= (ar+ra)s(ar+ra) + s(ar+ra)^2. \end{aligned}$$

Suponha primeiro que $Z = (0)$ então $a^2 = 0$, e $(ar + ra)^2 = 0$, o que implica que $0 = (ar + ra)^2 ar = (arar + arra + rara)ar = (ar)^3$ de onde nós obtemos que $(ar)^3 = 0$. Porém, temos então que aA é um ideal a direita de A na qual o cubo de todo elemento é zero, e pelo Lema 2.119, a álgebra A teria um ideal nilpotente diferente de zero, o que é impossível já que álgebra é simples.

Portanto podemos assumir $Z \neq 0$ e que existe um elemento $a \in U$, com $a \notin Z$, tal que $a^2 \neq 0 \in Z$ e $(ar + ra)^2 \in Z$ para todo $r \in A$.

Para responder completamente qual deve ser a estrutura de A, apresentaremos mais um teorema, contudo para isso, precisamos de mais algumas definições e lemas, que serão abordados logo abaixo.

Definição 2.127. Seja A um anel e M um $R-$módulo. Se $Ann(M) = 0$, então M é chamado de módulo fiel.

Definição 2.128. Um anel R é dito ser primitivo se tem um módulo fiel e irredutível. Diremos que uma álgebra A é primitiva, se a estrutura de anel subjacente a A é um anel primitivo.

Seja A uma álgebra primitiva e M um módulo simples e fiel de A. Considere o conjunto $C(M)$ como conjunto de todos os endomorfismos de M. Note que para todo $\alpha \in A$, tem-se que E_α definido por $E_\alpha(m) = \alpha m$, chamada de multiplicação à esquerda, é um endomorfismo do grupo aditivo de M. Além disso, chamamos de $E(M) = End_R(M)$ o conjunto de todos os endomorfismos do R-módulo à esquerda M.

Um lema que cumpre um papel importante na teoria das representações de grupo é o conhecido Lema de Schur, em homenagem ao matemático russo-alemão Issai Schur. Mostraremos agora que módulos simples dão origem a anéis de divisão. O Lema de Schur pode ser apenas uma observação direta, mas extremamente útil.

Lema 2.129. *(Lema de Schur) Se M é um módulo simples, então o anel de endomorfismo anel $E(M)$ é um anel de divisão.*

Demonstração. Seja $f \in E(M)$, $f \neq 0$. Então $N(f)$ é um submódulo de M diferente de M, e $im(f)$ é um submódulo de M diferente de 0. Como M é simples, a única possibilidade é que $N(f) = 0$ e $im(f) = M$. Assim, f é um automorfismo e, portanto, um elemento invertível em $E(M)$. □

Chamamos de $C(M) = \{\psi \in E(M)$ tal que $E_\alpha \circ \psi = \psi \circ E_\alpha$, $\forall \alpha \in R\}$. Pelo Lema de Schur, $C(M)$ é um anel de divisão e consequentemente um corpo. Podemos assim, considerar M como um espaço vetorial sobre $C(M)$ onde $\alpha(m)$ para $m \in M$ e $\alpha \in C(M)$ é interpretado como a ação de α como um elemento de $E(M)$ aplicado em m.

Definição 2.130. A álgebra A é dita ser densa em M, se para cada inteiro positivo n, todo subconjunto linearmente independente $\{v_1, \cdots, v_n\}$ de M sobre a álgebra de divisão $D = C(M)$ e todo subconjunto arbitrário $\{w_1, \cdots, w_n\}$ de M, há um elemento $r \in A$ tal que $w_i = v_i r$ para todos $i = \{1, \cdots, n\}$.

O resultado básico que surge a partir de toda essa teoria é um conhecido teorema, chamado de Teorema da Densidade de Jacobson e Chevalley. Sua demonstração será omitida, mas é encontrada em vários livros de álgebra como [15, p. 41].

Teorema 2.131 (Teorema da Densidade)**.** *Seja A uma álgebra primitiva e M um módulo fiel e simples. Se $D = C(M)$ então A é denso em M sobre D.*

Agora que temos o que é necessário, podemos voltar a questão anterior, onde um corpo tem características 2, além de U ser ideal de Lie e subálgebra de A, com $A \neq U$ e $U \nsubseteq Z$. O próximo teorema será fundamental para entendermos quais devem ser as hipóteses mínimas para que a álgebra A satisfaça os Teorema 2.125 e o Corolário 2.126. Antes dele, um pequeno lema que será útil para o teorema em questão.

Lema 2.132. *Seja A uma álgebra simples com $Z(A) \neq (0)$, então A é uma álgebra com unidade.*

Demonstração. Suponha que o centro Z de A seja não nulo, então existe $x \in Z$ tal que $I_d(x)$ (ideal gerado por x) é o próprio A, já que A é simples. Como $x \in I_d(x)$, existe $y \in A$, tal que $xy = x$.

Vamos mostrar que esse y é unidade de A. De fato, tome $z \in A$, existe $a \in A$ tal que $z = ax$. Logo temos que

$$yz = y(ax) = y(xa) = (xy)a = xa = z$$
$$zy = (ax)y = a(xy) = a(x) = ax = z$$

Logo y é a unidade de A. □

O teorema abaixo apresenta um resultado bem conhecido, faz parte do resultado central do artigo [20] para aqueles que apresentarem interesse.

Teorema 2.133. *Seja A um anel de divisão com centro Z, e suponha que para cada $x \in A$, existe alguma potência (dependendo de x) que está em Z, então A é um anel comutativo; em outras palavras, para todo $x \in A$, se existir $n = n(x)$ tal que $x^n \in Z$, então A é comutativo.*

O teorema abaixo será fundamental para que Teorema 2.125 possa valer para um caso mais geral.

Teorema 2.134. *Seja A uma álgebra simples de característica 2 e suponha que existe $a \in A$ com $a \notin Z$ tal que $a^2 \in Z$ e $(ax + xa)^4 \in Z$ para todo $x \in A$. Então* $\dim_Z A = 4$.

Demonstração. Se $Z = 0$, então $a^2 = 0$ e $(ax + xa)^4 = 0$, consequentemente $(ax)^5 = a(ax + xa)^4 x = 0$, para todo $x \in A$. Então o ideal à direita aA satisfaz $u^5 = 0$ para todo elemento $u \in aA$, pelo Lema 2.119 isso não é possível em uma álgebra simples.

Por outro lado, se $Z \neq 0$ pelo lema anterior, tem-se que $1 \in A$. Se $a^2 = 0$ tomando $b = a + 1$ temos que $b^2 = a^2 + 2a + 1 = 1$, além disso $(bx + xb) = (ax + xa + 2x) = (ax + xa)$ logo $(bx + xb)^4 \in Z$ para todo $x \in A$. Portanto podemos assumir que $a^2 = \alpha \neq 0$ em Z. Seja $Z' = Z(\sqrt{\alpha})$ então $A' = A \otimes Z'$ é simples já que A é simples. Além disso, em A' temos que $(ax' + x'a)^4 \in Z'$ para todo $x' \in A'$.

Note que a dimensão de A' sobre Z' é igual a dimensão de A sobre Z. Basta ver que, se $B = \{v_1, \cdots, v_n\}$ é uma base de A, então $B' = \{v_1 \otimes 1, \cdots, v_n \otimes 1\}$ é uma base de A'. Consequentemente para provar o teorema é suficiente provar em A'. Temos também que $b = a/\beta$, onde $\beta \in Z'$, $\beta^2 = \alpha$ satisfaz $b^2 = 1$ e $(bx' + x'b)^4 \in Z$. Consequentemente sem perda de generalidade, podemos assumir que $a \in A$ tal que $a \notin Z$, $a^2 = 1$ e $(ax + xa)^4 \in Z$ para todo $x \in A$.

Observe que como A é simples, então tomando $V = M = A$ temos as hipóteses mínimas para aplicar o teorema da densidade. Mas primeiro note que $(a + 1)^2 = 0$ e $a + 1 \neq 0$, V tem dimensão maior que 1 sobre D, já que $a \neq 1$. Pelo teorema da densidade, existe um $v \in V$ tal que v, va são linearmente independentes sobre D.

Se para algum $w \in V$, v, va e $w(1 + a)$ são linearmente independentes sobre D, interpretando a a transformação que age pela direita temos que

$$\begin{aligned} a(v) &= va \\ a(va) &= va^2 = v \\ a(w(1 + a)) &= w(a + 1). \end{aligned}$$

Logo a matriz dessa transformação linear é

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

pelo teorema da densidade, como $v, va, w(a + 1)$ são linearmente independentes existe uma

transformação linear x, tal que

$$x(v) = va$$
$$x(va) = 0$$
$$x(w(1+a)) = 0.$$

Logo a matriz dessa transformação linear é

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

e consequentemente a transformação $(ax+xa)$ induz a matriz

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Mas note que $(ax+xa)^4 \notin Z$, já que não é uma matriz escalar. De fato, $(ax+xa)^4$ induz a matriz

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

o que é um absurdo, logo para qualquer $w \in V$, tem-se que $v, va, w(a+1)$ são linearmente dependentes sobre D.

Suponhamos então que V tem dimensão maior que 2 sobre D, então existe $w \in V$ tal que v, va, w são linearmente independentes sobre D. Observe que a matriz da transformação linear neste caso, não é maior que 3×3, pois como v, va e $w(a+1)$ são linearmente dependentes, isso significa que wa pertence ao subespaço gerado por v, va, e w.

Desta forma, a matriz dessa transformação caso tivesse tamanho maior que 3, seria nula abaixo das 3 primeiras linhas, e seria uma matriz triangular em blocos, ou seja, o subespaço gerado por v, va e w seria invariante pela ação de a. Portanto, basta considerar a matrizes de tamanho 3. Neste caso representado pela matriz

$$\begin{bmatrix} 0 & 1 & \alpha \\ 1 & 0 & \beta \\ 0 & 0 & \varepsilon \end{bmatrix}$$

novamente aplicado o teorema da densidade, existe um $x \in A$ tal que induz a matriz

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Por outro lado, a a transformação $(ax+xa)$ induz a matriz

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 0 \end{bmatrix}$$

além disso, $(ax+xa)^4$ é dado pela mesma matriz acima, que claramente não é uma matriz escalar. Logo V tem dimensão 2 sobre D, restando mostrar que D é comutativo.

Seja

$$a = \begin{bmatrix} \alpha & \beta \\ \varepsilon & \delta \end{bmatrix}$$

como $a^2 = I_d$, temos que

$$\alpha^2 + \beta\varepsilon = 1$$
$$\alpha\beta + \beta\delta = 0$$
$$\varepsilon\beta + \delta^2 = 1.$$

Dessas equações sabemos que não pode ocorrer de $\alpha = 0$ e $\varepsilon = 0$ ao mesmo tempo. Por outro lado seja $\eta \in D$, usando que $(ax+xa)^4 \in Z$, onde

$$x = \begin{bmatrix} 0 & \eta \\ 0 & 0 \end{bmatrix}.$$

Note que embora Z é um subanel de D, também está identificando as matrizes com entradas em Z. Além disso, note que

$$(ax+xa)^4 = \begin{bmatrix} \eta\alpha & \alpha\eta+\eta\delta \\ 0 & \varepsilon\eta \end{bmatrix}^4 \in Z.$$

Isso mostra que para todo $\eta \in D$, tem-se que $(\eta\alpha)^4 \in Z$ e $(\delta\eta)^4 \in Z$. Se $\alpha \neq 0$, então $\eta\alpha \in D$ pode ser interpretado como η, logo para todo $x \in D$ tem-se que $x^4 \in Z$.

Então pelo Teorema 2.133, logo $\alpha \neq 0$ implica que D seja comutativo. O caso $\delta \neq 0$ é

análogo. Portanto D é um corpo e consequentemente a dimensão de A sobre D é $2^2 = 4$. □

Gostaríamos de apontar que um teorema mais geral realmente é válido, a saber: se A é uma álgebra simples com um elemento $a \notin Z$ tal que $(ax - xa)^n \in Z$ para todo $x \in A$ e $n \in \mathbb{N}$ então A é 4-dimensional sobre Z.

Uma vez que a hipótese do Teorema 2.134 é precisamente a que leva à suposição de que o Teorema 2.125 e o Corolário 2.124 era falso, obtemos, uma versão mais geral do Teorema 2.125 como o esperado.

Teorema 2.135. *Se A é uma álgebra simples e U é ideal de Lie de A, então $U \subseteq Z$ ou $[A, A] \subseteq U$, exceto se A é de características 2 e tem dimensão 4 sobre o seu centro Z.*

Assim como foi no caso do Teorema 2.125, temos uma consequência direta do Teorema 2.135 é o seguinte corolário.

Corolário 2.136. *Se A é álgebra simples não comutativa então a subálgebra gerada por $[A, A]$ é A.*

Capítulo 3

Relações entre comutadores e imagens de polinômios.

Neste capítulo estudaremos a conexão entre as matrizes de traço zero $sl_n(\mathbb{K})$, com as imagens de polinômios multilineares sobre $M_n(\mathbb{K})$. Em particular, apresentamos algumas evidências que nos permitem enunciar a conjectura de Mesyan para um novo limitante, que por sua vez, não pode ser melhorada.

Em seguida provaremos que para cada inteiro positivo d existe um inteiro positivo s, tal que $f(M_s(\mathbb{K}))$, a imagem de f em $M_s(\mathbb{K})$, contém todos as matrizes $d \times d$ de traço zero. E por último, sendo A uma álgebra, consideramos a questão de quando $spanf(A)$, subespaço vetorial gerado pela imagem de f em A, é igual a A. E ainda, a relação entre $[A, A]$, o espaço vetorial gerado pelos comutadores em A, e $spanf(A)$

3.1 Um novo limite para a conjectura de Mesyan

O objetivo principal desta secção é apresentar evidências que nos permitam enunciar a conjectura de Mesyan em um cenário mais geral. Os resultados deste seção podem ser encontrados em [12].

Seja $\mathbb{K}$ um corpo e seja $M_n(\mathbb{K})$ a álgebra de matrizes $n \times n$ sobre $\mathbb{K}$. A conjectura de Lvov-Kaplansky afirma que a imagem de um polinômio multilinear em $M_n(\mathbb{K})$ é um espaço vetorial. Tal conjectura é equivalente a imagem de um polinômio multilinear em $M_n(\mathbb{K})$ ser $\{0\}$, $\mathbb{K}$, $sl_n(\mathbb{K})$ ou toda a álgebra $M_n(\mathbb{K})$.

Um enfraquecimento da conjectura de Lvov-Kaplansky é a chamada conjectura de Mesyan.

Conjectura 3.1. *Seja $\mathbb{K}$ um corpo, $n \geqslant 2$ e $m \geqslant 1$ inteiros e $f(x_1, \ldots, x_m)$ um polinômio*

multilinear diferente de zero em $K\langle x_1,\ldots,x_m\rangle$. *Se* $n \geqslant m-1$, *então a imagem de* f *contém todas as matrizes de traço zero em* $M_n(\mathbb{K})$.

A conjectura citada acima é baseada no seguinte resultado:

Proposição 3.2. *Sejam* $\mathbb{K}$ *um corpo,* $n \geqslant 2$ *e* $m \geqslant 2$, *inteiros em que* $char(K) \nmid n$ *e* $f(x_1,\cdots,x_m) \in K\langle X\rangle$ *um polinômio multilinear não nulo. Se* $n \geqslant m-1$, *então o* $\mathbb{K}$*-subespaço, denotado por* $span(f(M_n(K))$ *contém* $sl_n(K)$.

Na verdade, uma vez que assumimos que a conjectura de Lvov-Kaplansky é verdadeira, a frase "a imagem de f em $M_n(\mathbb{K})$ contém $sl_n(\mathbb{K})$"é equivalente a "f não é uma identidade polinomial nem um polinômio central de $M_n(\mathbb{K})$ ". Por outro lado, é sabido que $M_n(\mathbb{K})$ não tem identidades polinomiais ou polinômios centrais de grau $m \leqslant n+1$, e isso faz com que conjectura de Mesyan seja um caso particular da conjectura de Lvov-Kaplansky. Em particular, um contra-exemplo à conjectura Mesyan é um contra-exemplo para a conjetura de Lvov-Kaplanksy.

Seja f é um polinômio de grau $m \leqslant 2n-1$ então pelo Corolário 2.93, f não é um polinômio central nem uma identidade polinomial para $M_n(\mathbb{K})$ e por nossa suposição $f(M_n(\mathbb{K}))$ é $sl_n(\mathbb{K})$ ou $M_n(\mathbb{K})$ que em ambos os casos, contém $sl_n(\mathbb{K})$.

O fato acima sugere que a conjectura de Mesyan pode ser enunciada de uma forma mais geral, isto é, para $n \geqslant (m+1)/2$. Além disso, este limite não pode ser melhorado uma vez que S_{2n}, o polinômio standard de grau $2n$, é uma identidade polinomial de grau $2n$ para $M_n(\mathbb{K})$. Apresentamos agora mais uma evidência de que a conjectura deve ser afirmada neste contexto. Vamos provar uma versão mais geral da Proposição 3.2.

Teorema 3.3. *Sejam* $\mathbb{K}$ *um corpo,* $n \geqslant 2$ *e* $m \geqslant 2$, *inteiros em que* $char(K) \nmid n$ *e* $f(x_1,\cdots,x_m) \in K\langle X\rangle$ *um polinômio multilinear não nulo. Se* $m \leqslant 2n-1$, *então o* $\mathbb{K}$*-subespaço* $span(f(M_n(\mathbb{K}))$ *contém* $sl_n(\mathbb{K})$.

Lembre-se de que, para polinômios de grau $m = 2$, a conjectura de Lvov-Kaplanksy é uma consequência da Proposição 2.94 e dos resultados de Shoda e Albert e Muckenhoupt. Além disso, para $m \geqslant 3$, temos que $\frac{m+1}{2} \leqslant m-1$, o que mostra que o teorema acima é uma generalização da Proposição 3.2. Antes de provar o teorema acima, devemos primeiro lembrar o seguinte lema técnico, extraído de [5].

Proposição 3.4. *Sejam* D *um anel de divisão,* $n \geqslant 2$ *um inteiro e* $A \in M_n(D)$ *uma matriz não central. Então,* A *é semelhante a uma matriz em* $M_n(D)$ *com no máximo uma entrada não nula sobre a diagonal principal. Em particular, se* A *possui traço zero e não é central, então* A *é semelhante a uma matriz em* $M_n(D)$ *com apenas zeros na diagonal principal.*

Demonstração. A demonstração será realizada em três casos: primeiro caso para $n = 2$, o segundo caso para $n = 3$ e o terceiro para $n \geqslant 3$

1º caso: $n = 2$.

Seja

$$A = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in M_2(D)$$

temos então que

$$\begin{bmatrix} 1 & u \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} 1 & -u \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x+uz & -(x+uz)u+y+uw \\ z & -zu+w \end{bmatrix}. \tag{3.1}$$

Se $z \neq 0$, basta tomar $u = -xz^{-1}$. Por outro lado, se y é não nulo, então tomando $u = y^{-1}x$, temos que a matriz

$$P = \begin{bmatrix} 1 & 0 \\ u & 1 \end{bmatrix}$$

é tal que PAP^{-1} é matriz pedida.

Por último, se $z = y = 0$, sabendo que A não é múltiplo da matriz identidade, deve ocorrer de $(x - w) \neq 0$. Logo existe $v = x(x - w)^{-1}$ tal que a matriz

$$P = \begin{bmatrix} 1 & u \\ v & uv+1 \end{bmatrix}$$

é tal que

$$PAP^{-1} = \begin{bmatrix} -vw+vx+x & w-x \\ -v^2w-vw+v^2x+vx & vw+w-vx \end{bmatrix}.$$

Portanto A é semelhante a uma matriz com no máximo uma entrada não nula sobre a diagonal principal.

2º caso: $n = 3$.

Seja

$$A = \begin{bmatrix} x & y & z \\ q & w & t \\ u & v & s \end{bmatrix} \in M_3(D)$$

sendo A não central, então A não é um múltiplo da matriz identidade I_d, logo pela Observação 2.25, pelo menos uma de suas submatrizes principais não é central. Suponha primeiro que a matriz abaixo é não central

$$E = \begin{bmatrix} w & t \\ v & s \end{bmatrix}.$$

Seja I_d a matriz identidade de $M_n(D)$

$$I_d = \begin{pmatrix} & & & coluna(i) & & coluna(j) & & \\ 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & 0 & \dots & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \ddots & 1 \end{pmatrix} \tag{3.2}$$

chamamos de I_{ij} a matriz identidade com as colunas i e j trocadas entre si de lugar. Isto é

$$I_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \dots & 1 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \tag{3.3}$$

logo

$$I_{23}I_{12}AI_{12}I_{23} = \begin{bmatrix} w & t & z \\ v & s & u \\ y & z & x \end{bmatrix}.$$

Caso a outra submatriz seja não central, isto é

$$F = \begin{bmatrix} x & z \\ u & s \end{bmatrix} \text{ é não central, então}$$

$$I_{23}AI_{23} = \begin{bmatrix} x & z & y \\ u & s & v \\ q & t & w \end{bmatrix}.$$

Portanto, a matriz F é a primeira submatriz principal de A, que é não central, pelo primeiro caso existe uma matriz $P \in M_2(D)$ invertível tal que a entrada $(1,1)$ de PFP^{-1} é nula. Desde que a matriz

$$\begin{bmatrix} P & 0 \\ 0^t & 1 \end{bmatrix} A \begin{bmatrix} P^{-1} & 0 \\ 0^t & 1 \end{bmatrix}$$

possui a entrada $(1,1)$ igual a 0, em que 0 é o vetor coluna e 0^t é o vetor linha, segue que podemos supor

$$A = \begin{bmatrix} 0 & x & y \\ z & a & b \\ q & c & d \end{bmatrix}.$$

Se tivermos

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

não central, então novamente pelo caso anterior, conseguimos A conjugada à uma matriz com as entradas $(1,1)$ e $(2,2)$ nulas. Suponha agora

$$A = \begin{bmatrix} 0 & x & y \\ z & \alpha & 0 \\ q & 0 & \alpha \end{bmatrix}$$

com $\alpha \in Z(D)$.

Se $z \neq 0$, tem-se que

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & x & y \\ z & \alpha & 0 \\ q & 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & -a & -b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} az+bq & \star & \star \\ z & -za+\alpha & -zb \\ q & -qa & -qb+\alpha \end{bmatrix}$$

tomando $a = -qz^{-1}$ e $b = 1$, então basta aplicarmos o primeiro caso à submatriz principal no canto direito inferior. Caso $q \neq 0$, então devemos tomar $a = 1$ e $b = -zq^{-1}$.

De modo análogo, também podemos reduzir a primeira entrada a zero, quando $x \neq 0$ ou $y \neq 0$. E quando todos são nulos, isto é, $x = y = z = q = 0$, tem-se que

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & \star & \star \\ \star & 0 & \star \\ \star & \star & 2\alpha \end{bmatrix}$$

como queríamos.

3º caso: $n \geqslant 3$.

Como por hipótese a matriz A é não central, podemos estender de forma análoga os argumentos utilizados no caso anterior afim de garantir que exista um maior número natural não nulo m tal que A seja conjugada à

$$\begin{bmatrix} B & \star \\ \star & C \end{bmatrix}$$

em que é B uma matriz com apenas zero na diagonal principal de ordem $m \times m$ e C é uma matriz de ordem $k = n - m$. Afirmamos que deve ocorrer de $k = 1$, pois caso contrário poderíamos escolher a submatriz principal 3×3 formada pela última linha e coluna de B e pelas duas primeiras linhas e colunas de C. Pelo segundo caso, conseguiríamos produzir uma matriz conjugada à A começando com uma matriz com apenas zero na diagonal de ordem $m+1$, contradizendo a maximalidade de m.

□

Observação 3.5. Observe que a hipótese de que A seja uma matriz não central é importante para a validade do teorema, já que caso a matriz A seja múltiplo da identidade e de traço zero, a conclusão não é verdadeira.

De fato, considere a matriz identidade I_d de tamanho 2×2, sobre um corpo $\mathbb{K}$ de característica 2. Então é claro que I_d possui traço zero, porém a única matriz semelhante à I_d é ela própria, ou seja, não existem matrizes com apenas zeros na diagonal principal que sejam semelhantes matriz identidade em $M_2(\mathbb{K})$.

Agora estamos prontos para apresentar uma prova do Teorema 3.3.

Demonstração. Seja

$$f(x_1, \cdots, x_m) = \sum_{\sigma \in S_m} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(m)},$$

onde $\alpha_\sigma \in \mathbb{K}$. Renomeando as variáveis se necessário podemos supor sem perda de generalidade que $\alpha_{(1)} \neq 0$. Note que pela Proposição 2.95 temos que $sl_n(\mathbb{K}) \subseteq f(M_n(\mathbb{K}))$ sempre que $m \leqslant 2$, portanto podemos assumir que $m \geqslant 3$.

Sejam $i, j \in \{1, \ldots, n\}$ tal que $i \neq j$. A demonstração será dividida em dois casos:

1º caso: m é par.

Neste caso, temos que $m = 2k$, com k um inteiro maior que 1, então temos que $n \geqslant k + \frac{1}{2}$. Como k e n são inteiros tem-se que $n \geqslant k+1$, portanto $n-2 \geqslant k-1$. Com isso, podemos

encontrar $k-1$ elementos distintos em $\{1,\cdots,n\}-\{i,j\}$ que denotaremos por $l_1,\cdots,l_{k-1}$. Portanto,

$$f(e_{ii},e_{ij},e_{jl_1},e_{l_1l_1},e_{l_1l_2},e_{l_2l_2}\cdots,e_{l_{k-2}l_{k-1}},e_{l_{k-1}j})=\alpha_1 e_{ij}.$$

Como $\alpha_1\neq 0$, e e_{ij} com i,j distintos é uma matriz cuja diagonal contém apenas zeros, temos que , $span(f(M_n(\mathbb{K}))$ contém todas as matrizes cuja diagonal principal é composta por apenas zeros.

Seja A uma matriz com traço zero, devemos mostrar que $A\in span(f(M_n(\mathbb{K}))$. De fato, como $car(K)\nmid n$ então A é uma matriz não central, e consequentemente pela Proposição 3.4 tem-se que existe $B\in span(f(M_n(\mathbb{K}))$ tal que $A=PBP^{-1}$ e usando a Proposição 2.94, parte 2, concluimos que $A\in span(f(M_n(\mathbb{K}))$.

2º caso: m é ímpar.

Então m é da forma $m=2k-1$, onde k é um número inteiro e $k\geqslant 2$. Como $2n\geqslant m+1$, temos $n-2\geqslant k-2$. Logo, podemos encontrar $k-2$ elementos distintos em $\{1,\cdots,n\}-\{i,j\}$, que denotaremos por $l_1,\cdots,l_{k-2}$. Portanto,

$$f(e_{ii},e_{ij},e_{jl_1},e_{l_1l_1},e_{l_1l_2},e_{l_2l_2}\cdots,e_{l_{k-3}l_{k-2}},e_{l_{k-2}l_{k-2}},e_{l_{k-2}j})=\alpha_1 e_{ij}.$$

Novamente pelo mesmo motivo do caso anterior, temos que Se $A\in sl_n(\mathbb{K})$ então pelo Proposição 3.4 e Proposição 2.94 temos que $A\in span(f(M_n(\mathbb{K}))$.

Com isso em qualquer um dos casos provamos que $span(f(M_n(\mathbb{K}))$ contém todas as matrizes de traço zero. □

Agora, reenunciamos a conjectura de Mesyan à luz do Teorema 3.3 e da discussão do início desta seção.

Conjectura 3.6. *(Conjectura de Mesyan reformulada). Seja $\mathbb{K}$ um corpo, $n\geqslant 2$ e $m\geqslant 1$ inteiros, e $f(x_1,\ldots,x_m)$ um polinômio multilinear diferente de zero em $K\langle x_1,\cdots,x_m\rangle$. Se $m\leqslant 2n-1$, então $sl_n(\mathbb{K})\subseteq f(M_n(\mathbb{K}))$.*

3.2 Polinômios parcialmente comutativos admissíveis

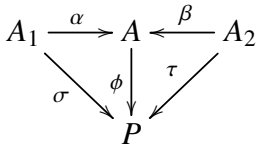
Apresentamos um conceito básico porém fundamental para o estudo de imagens de polinômios em álgebras de matrizes finitárias, que é o de coproduto. Esta seção é baseada no artigo

[28].

Definição 3.7. Sejam A_1 e A_2 álgebras com unidade sobre um anel comutativo K. Então, uma K-álgebra A com unidade é chamado de coproduto de A_1 e A_2 sobre K se:

1. Existem homomorfismos da álgebra $\alpha : A_1 \to A$ e $\beta : A_2 \to A$ tal que $A_1^\alpha \cup A_2^\beta$ gera A como uma K-álgebra. Onde A_1^α é a imagem de A_1 por α e A_2^β é a imagem de A_2 por β.

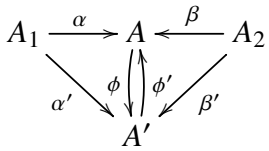
2. Para qualquer K-álgebra P com unidade e homomorfismos $\sigma : A_1 \to P$ e $\tau : A_2 \to P$ existe um homomorfismo $\phi : A \to P$ tal que $\phi \circ \alpha = \sigma$ e $\phi \circ \beta = \tau$. Ou seja, o diagrama



   sempre pode ser completado.

Denotamos coproduto de A_1 e A_2 por $A_1 \coprod A_2$.

Observação 3.8. O coproduto de A_1 e A_2 é único a menos de isomorfismo. De fato, seja A e A' dois coprodutos de A_1 e A_2 sobre $\mathbb{K}$, então A e A' são isomorfos pelo isomorfismos ϕ e ϕ' conforme indicado a seguir pelo diagrama:



De fato, o homomorfismos $\phi' : A' \to A$, é tal que $\alpha = \phi' \circ \alpha'$, assim como $\beta = \phi' \circ \beta'$. Por outro lado o homomorfismos $\phi : A \to A'$ satisfaz $\alpha' = \phi \circ \alpha$, assim como $\beta' = \phi \circ \beta$. Note que $\phi' \circ \phi \circ \alpha = \phi' \circ \alpha' = \alpha$ e $\phi' \circ \phi \circ \beta = \phi' \circ \beta' = \beta$, a mesma que a identidade. Como o mapa da propriedade universal é único, temos que $\phi' \circ \phi = I_d$. O caso $\phi \circ \phi' = I_d$ é análogo.

É imediato de (1) e (2) que ϕ é determinado exclusivamente por σ e τ.

Seja $\mathbb{K}$ um corpo, seja $n \in \mathbb{N}$, e seja $\Omega = \{w_1, \ldots, w_m\} \subseteq \mathbb{N} - \{1, \ldots, n\}$ um conjunto finito (a razão para tal notação é que iremos alterar Ω no decorrer da prova). Por

$$\mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle \coprod \mathbb{K}[U_w; w \in \Omega]$$

denotamos o coproduto da álgebra livre $\mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle$ e a álgebra de polinômios comutativos $\mathbb{K}[U_w; w \in \Omega]$.

A seguir mostramos a existência do coproduto de A_1 e A_2 sobre $\mathbb{K}$ pela seguinte construção natural. A ideia é encontrar uma álgebra gerada por A_1 e A_2 (garantindo assim (i)) e então fatorar um ideal apropriado desta álgebra (assim garantindo (ii). Seja

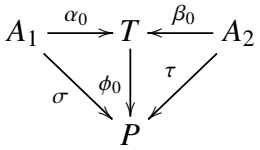
$$T = A_1 \oplus A_2 \oplus (A_1 \otimes A_1) \oplus (A_1 \otimes A_2) \oplus (A_2 \otimes A_1) \oplus (A_2 \otimes A_2) \oplus (A_1 \otimes A_1 \otimes A_1) \oplus \ldots$$

seja a álgebra tensorial de A_1 e A_2, e seja I o ideal de T gerado por todos os elementos da forma

$$a_1 \otimes b_1 - a_1 b_1; \quad a_2 \otimes b_2 - a_2 b_2; \quad 1_{A_1} - 1_{A_2} \tag{3.4}$$

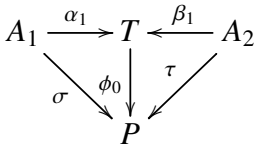
onde $a_1, b_1 \in A_1, a_2, b_2 \in A_2$, afirmamos que $A = T/I$ é um coproduto de A_1 e A_2 sobre $\mathbb{K}$. Sejam α_0 e β_0 os respectivos mapas de inclusão $\alpha_0 : A_1 \to T$ e $\beta_0 : A_2 \to T$ e seja v o homomorfismo natural de T em T/I. Definimos $\alpha_1 : A_1 \to T$ por $\alpha_1 = a_0 v$ e $\beta_1 : A_2 \to T$ por $\beta_1 = \beta_0 v$.

Propriedade (i) é então claro. Agora considere os homomorfismos $\sigma : A_1 \to P$, $\tau : A_2 \to P$, onde P é uma $\mathbb{K}$-álgebra. Primeiro completamos o diagrama



De fato, basta definir ϕ_0 em cada soma direta $T_1 \otimes T_2 \otimes \ldots \otimes T_m$ de T , onde cada T_i é A_1 ou A_2, e então estende por linearidade. O mapa $\chi : T_1 \otimes T_2 \otimes \ldots \otimes T_m \to P$ dado por $\chi(t_1 \otimes t_2 \otimes \ldots \otimes t_m) = t_1^p \otimes t_2^p \otimes \ldots \otimes t_m^p$ onde $p = \sigma$ se $t_i \in A_1$ ou $p = \tau$ se $t_i \in A_2$. Note que está função é linear, e pela natureza da multiplicação em T é fácil ver que ϕ_0 é um homomorfismo de álgebra.

Além disso, aplicando ϕ_0 aos geradores da Equação 3.4, fica claro que ϕ_0 mapeia I para 0. Como resultado, ϕ_0 pode ser elevado a um homomorfismo de álgebra $\phi : T/I \to P$ definindo $(\bar{t})^\phi = t^{\phi_0}$ onde $\bar{t} = t + I$ com $t \in T$. A comutatividade do diagrama acima produz a comutatividade de



o que mostra que a propriedade (ii) é válida. Com isso mostramos a existência e unicidade de um coproduto de A_1 e A_2.

Definição 3.9. Chamamos os elementos do coproduto acima de polinômios parcialmente comutativos. Podemos pensa-los como polinômios nas variáveis X_i, U_w, onde U_w comutam entre si, mas não comutam com X_i.

Seja A uma álgebra com unidade sobre $\mathbb{K}$, seja $x_1, \ldots, x_n \in A$, e seja $u_{w_1}, \ldots, u_{w_m} \in A$ elementos que comutam entre si. Por

$$E_{V_{x_1,\ldots,x_n;u_{w_1},\ldots,u_{w_m}}} : \mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle \coprod \mathbb{K}[U_w; w \in \Omega] \longrightarrow A$$

denotam o homomorfismo da álgebra enviando X_i para x_i e U_w para u_w.

Para um polinômio parcialmente comutativo $f \in \mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle \coprod \mathbb{K}[U_w; w \in \Omega]$, defina sua imagem na álgebra A como

$$f(A) = \{E_{V_{x_1,\ldots,x_n;u_{w_1},\ldots,u_{w_m}}}(f) \mid x_i, u_{w_j} \in A, u_{w_j}u_{w_k} = u_{w_k}u_{w_j}, \forall\, j,k\}$$

Observe que no caso em que f é um polinômio não comutativo, ou seja, um elemento da subálgebra $\mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle$ de $\mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle \coprod \mathbb{K}[U_w; w \in \Omega]$, esta noção da imagem de f coincide com a definição usual.

Diremos que as sequências $a^i = (a^i_1, \ldots, a^i_{k_i}), i = 1, \ldots, n$, contendo os elementos de Ω formam uma partição de Ω, se forem estritamente crescentes (ou seja, $a^i_1 < \ldots < a^i_{k_i}$) e, para cada $w \in \Omega$, existem i e j tal que $a^i_j = w$.

Por $|a^i|$ denotamos o comprimento da sequência, ou seja, k_i. Definimos $\mathcal{A}_\Omega$ como o conjunto de todas as n-tuplas de sequências $a = (a^1, \ldots, a^n)$ tais que $a^1, \ldots, a^n$ formam uma partição de Ω.

Para qualquer $a \in \mathcal{A}_\Omega$, se $k_i > 0$, definimos

$$X^a_i = [U_{a^i_1}, \ldots, U_{a^i_{ki}}, X_i].$$

Por outro lado, se $k_i = 0$ então definimos $X^a_i = X_i$. Estendemos essa definição definindo

$$(X_{i_1}X_{i_2}\ldots X_{i_k})^a = X^a_{i_1}X^a_{i_2}\ldots X^a_{i_k}$$

para todos os $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Generalizamos a noção de um polinômio multilinear como segue.

Definição 3.10. Um polinômio parcialmente comutativo

$$f \in \mathbb{K}\langle X_1, X_2, \ldots, X_n\rangle \coprod \mathbb{K}[U_w; w \in \Omega]$$

é admissível se for da forma

$$f = \sum_{\sigma \in S_n} \sum_{a \in \mathcal{A}} \lambda_\sigma^a \left(X_{\sigma(1)} X_{\sigma(2)} \dots X_{\sigma(n)}\right)^a$$

onde $\lambda_\sigma^a \in \mathbb{K}$.

Exemplo 3.11. Polinômios multilineares não comutativos são exatamente os polinômios parcialmente comutativos admissíveis para $\Omega = \emptyset$. Para dar um exemplo diferente, seja $n = 2$ e $\Omega = \{3,4\}$. Então

$$\mathcal{A}_\Omega = \{((3,4),\emptyset),\ ((3),(4)),\ ((4),(3)),\ (\emptyset,(3,4))\}$$

e os polinômios admissíveis em $\mathbb{K}\langle X_1, X_2\rangle \coprod \mathbb{K}[U_3; U_4]$ são da forma

$$\begin{aligned} f = {} & \lambda_{id}^{((3,4),\emptyset)}[U_3,U_4,X_1]X_2 + \lambda_{12}^{((3,4),\emptyset)}X_2[U_3,U_4,X_1] \\ & + \lambda_{id}^{((3),(4))}[U_3,X_1][U_4,X_2] + \lambda_{12}^{((3),(4))}[U_4,X_2][U_3,X_1] \\ & + \lambda_{id}^{(4),(3))}[U_4,X_1][U_3,X_2] + \lambda_{id}^{((4),(3))}[U_3,X_2][U_4,X_1] \\ & + \lambda_{id}^{(\emptyset,(3,4)}X_1[U_3,U_4,X_2] + \lambda_{12}^{(\emptyset,(3,4)}[U_3,U_4,X_2]X_1. \end{aligned}$$

Por definição, o espaço vetorial de polinômios parcialmente comutativos admissíveis são gerador por $(X_{\sigma(1)} \dots X_{\sigma(n)})^a$. O seguinte lema afirma que esses elementos realmente formam sua base. Sua prova é muito semelhante àquela de [29, Proposition 2.4], então nós a omitimos.

Lema 3.12. *Para qualquer $n \in \mathbb{N}$ e um conjunto finito $\Omega \subseteq \mathbb{N} - \{1,\dots,n\}$*

$$\{(X_{\sigma(1)},\dots,X_{\sigma(n)})^a \mid \sigma \in S_n\ a \in \mathcal{A}_\Omega\}$$

é um conjunto linearmente independente.

3.3 Imagens de polinômios multilineares em matrizes finitárias

Seja $\mathbb{K}$ um corpo infinito e seja $d \in \mathbb{N}$. Nesta seção, o principal objetivo é provar o Teorema 3.17, que afirma que para qualquer polinômio parcialmente comutativo admissível f diferente de zero existe um $s \in \mathbb{N}$ tal que $sl_d(\mathbb{K}) \subseteq f(M_s(\mathbb{K}))$. A principal referência desta seção é o artigo [28].

Aqui, a inclusão de uma álgebra de matriz menor em uma maior deve ser entendida como

$$M_d(\mathbb{K}) = \begin{bmatrix} M_d(\mathbb{K}) & 0 \\ 0 & 0 \end{bmatrix} \subseteq M_s(\mathbb{K})$$

A prova do teorema é por indução no número de variáveis não comutativas $x_1, \ldots, x_n$. Antes de considerar o caso base, provamos o seguinte lema.

Lema 3.13. *Para qualquer $d \in \mathbb{N}$ cada matriz $a \in sl_d(\mathbb{K})$ existe uma matriz invertível $p \in M_{d+1}(\mathbb{K})$ tal que $pap^{-1} \in M_{d+1}(\mathbb{K})$ é uma matriz oca, isto é, uma matriz tendo apenas zeros na diagonal principal.*

Demonstração. Prosseguimos por indução em d. O lema é obviamente verdadeiro para $d = 1$, então assuma que $d > 1$ e que o lema é verdadeiro para $d - 1$. Seja $a \in sl_d(\mathbb{K})$ uma matriz $d \times d$ com traço 0.

Primeiro, considere o caso em que existe um vetor diferente de zero $v \in \mathbb{K}^d$ que não é um autovetor de a. Então podemos estender o conjunto linearmente independente $\{v, av\}$ para uma nova base B do espaço $\mathbb{K}^d$; seja $q \in M_d(\mathbb{K})$ a matriz de mudança de base (da base canônica para a nova). Então

$$[a]_B = qaq^{-1} = \begin{bmatrix} 0 & x^t \\ y & b \end{bmatrix}$$

para algum $x, y \in \mathbb{K}^{d-1}$ e $b \in M_{d-1}(\mathbb{K})$. Como $tr(b) = tr(a) = 0$, pela hipótese de indução, existe uma matriz $r \in M_d(\mathbb{K})$ tal que $rbr^{-1} \in M_d(\mathbb{K})$ é uma matriz oca. Seja O matriz nula de tamanho conveniente, assim para

$$p = \begin{bmatrix} 1 & O \\ O & r \end{bmatrix} \begin{bmatrix} q & O \\ O & 1 \end{bmatrix} \in M_{d+1}(\mathbb{K})$$

temos que

$$
\begin{aligned}
pap^{-1} &= \begin{bmatrix} 1 & O \\ O & r \end{bmatrix}\begin{bmatrix} q & O \\ O & 1 \end{bmatrix}\cdot\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}\cdot\begin{bmatrix} q^{-1} & O \\ O & 1 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & r^{-1} \end{bmatrix} \\
&= \begin{bmatrix} 1 & O \\ O & r \end{bmatrix}\begin{bmatrix} qaq^{-1} & O \\ O & 0 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & r^{-1} \end{bmatrix} \\
&= \begin{bmatrix} 1 & O \\ O & r \end{bmatrix}\begin{bmatrix} 0 & x^t & \\ y & b & O \\ & O & 0 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & r^{-1} \end{bmatrix} \\
&= \begin{bmatrix} O & x^t \\ O & rb \end{bmatrix}\begin{bmatrix} 1 & O \\ O & r^{-1} \end{bmatrix} \\
&= \begin{bmatrix} O & x^t r^{-1} \\ O & rbr^{-1} \end{bmatrix}
\end{aligned}
$$

onde O é matriz nula. Portanto a matriz $pap^{-1} \in M_{d+1}(\mathbb{K})$ é oca.

Agora, suponha que todos os vetores diferentes de zero sejam autovetores de a. Isso implica que a é um múltiplo escalar da matriz identidade. Sem perda de generalidade, podemos assumir $a = I_d$. Como $0 = tr(a) = d$, a característica de $\mathbb{K}$ divide d. Seja $e_1, \ldots, e_{d+1} \in \mathbb{K}^{d+1}$ a base canônica de $\mathbb{K}^{d+1}$. Os vetores $f_i = e_i + e_{d+1}, i = \{1, \ldots, d\}$ e $f_{d+1} = e_{d+1} - \sum_{j=1}^{d} e_j$ formam uma base, uma vez que

$$e_i = -\sum_{\substack{j=1 \\ j\neq i}}^{d+1} f_j$$

para $i = \{1, \ldots, d\}$ e $e_{d+1} = \sum_{j=1}^{d+1} = f_j$. Seja $p \in M_{d+1}(\mathbb{K})$ a matriz de mudança de base (da base canônica para a nova). Como $ae_i = e_i$ para $i = 1, \ldots, d$ e $ae_{d+1} = 0$, temos

$$af_i = e_i = -\sum_{\substack{j=1 \\ j\neq i}}^{d+1} f_j$$

para $i = 1, \ldots, d$ e

$$af_{d+1} = -\sum_{i=1}^{d} e_i = \sum_{i=1}^{d}\sum_{\substack{j=1 \\ j\neq i}}^{d+1} f_j = (d-1)\sum_{i=1}^{d} f_i + df_{d+1}$$

Como a $Char(\mathbb{K})$ divide d, segue que $d = 0$ em $\mathbb{K}$, logo

$$af_{d+1} = \sum_{i=1}^{d} f_j$$

Portanto, a matriz $pap^{-1} \in M_{d+1}(\mathbb{K})$ é oca. □

O próximo resultado é de grande importância para estabelecer a base de nossa indução, consideramos os polinômios comutativos admissível sobre o corpo infinito $\mathbb{K}$.

Lema 3.14. *Para um polinômio parcialmente comutativo admissível diferente de zero*

$$f \in \mathbb{K}\langle X_1\rangle \coprod [\mathbb{K}|w \in \Omega]$$

tem-se que $sl_d(\mathbb{K}) \subseteq f(M_{d+1}(\mathbb{K}))$.

Demonstração. Seja $\Omega = \{w_1, \ldots, w_m\}$ com $w_1 < \ldots < w_m$. Como $\mathcal{A}_\Omega = \{(\underline{w})\}$ com $\underline{w} = (w_1, \ldots, w_m)$. Então

$$f = \lambda X_1^{(\underline{w})} = \lambda[U_{w1}, \ldots, U_{wm}, X_1]$$

para algum $\lambda \in \mathbb{K}$ diferente de zero. Seja $a \in sl_d(F)$ uma matriz de traço zero. Pelo Lema 3.13, existe uma matriz invertível $p \in M_{d+1}(\mathbb{K})$ tal que $pap^{-1} \in M_{d+1}(\mathbb{K})$ é oca.

Como $\mathbb{K}$ é infinito, podemos tomar $u = diag(\alpha_1, \ldots, \alpha_{d+1}) \in M_{d+1}(\mathbb{K})$ com $a_i \in \mathbb{K}$ dois a dois distintos. Assim, para qualquer $y = (y_{ij}) \in M_{d+1}(\mathbb{K})$, temos que $[u, y]_{ij} = (\alpha_i - \alpha_j)y_{ij}$. Em particular tem-se que

$$[\underbrace{u, \ldots, u}_{m}, x] = (\alpha_i - \alpha_j)^m y_{ij}.$$

Tomando $x_{ij} = \frac{a_{ij}}{(\alpha_i - \alpha_j)^m}$ existe uma matriz $x = (x_{ij}) \in M_{d+1}(\mathbb{K})$ tal que

$$pap^{-1} = [\underbrace{u, \ldots, u}_{m}, x].$$

Portanto, a imagem de f contém

$$E_{V_{\lambda^{-1}p^{-1}xp;p^{-1}up,\ldots,p^{-1}up}}(f) = \lambda p^{-1}[\underbrace{u, \ldots, u}_{m}, \lambda^{-1}x]p = a$$

isso mostra que $sl_d(f) \subseteq f(M_{d+1}(F))$ □

Antes de fazer o passo de indução, vamos provar dois lemas. A primeira é apenas uma observação elementar envolvendo as matriz unitárias e_{ij}.

Lema 3.15. *Seja A uma álgebra com unidade, seja $k \in \mathbb{K}$, e seja*

$$v = \sum_{i=1}^{k} e_{i,i+1} + e_{k+1,1} \in M_{k+1}(A)$$

então

$$[\underbrace{v, \ldots, v}_{j}, e_{k+1,1}] = \sum_{s=0}^{j} (-1)^s \binom{j}{s} e_{k+1-j+s,1+s}$$

para cada $j = 0, 1, \ldots, k$. Em particular, para $a \in A$, temos

$$[\underbrace{v, \ldots, v}_{k}, ae_{k+1,1}] = diag(a, \star, \ldots, \star)$$

Demonstração. Prosseguimos por indução em j. O lema é obviamente verdadeiro para $j = 0$, então assuma que $0 < j \leqslant k$ e que o lema é verdadeiro para $j - 1$. Usando a hipótese de indução, temos

$$\begin{aligned}
[\underbrace{v, \ldots, v}_{j}, e_{k+1,1}] &= [v, [\underbrace{v, \ldots, v}_{j-1}, e_{k+1,1}]] \\
&= \left[v, \sum_{s=0}^{j-1} (-1)^s \binom{j-1}{s} e_{k+1-(j-1)+s,1+s} \right] \\
&= \sum_{s=0}^{j-1} (-1)^s \binom{j-1}{s} [v, e_{k+2-j+s,1+s}]
\end{aligned}$$

Note que pela Observação 2.9, tem-se que

$$\begin{aligned}
v e_{k+2-j+s,1+s} &= e_{k+1-j+s,1+s} \\
e_{k+2-j+s,1+s} v &= e_{k+2-j+s,2+s}
\end{aligned}$$

Consequentemente temos que

$$\begin{aligned}
[\underbrace{v, \ldots, v}_{j}, e_{k+1,1}] &= \sum_{s=0}^{j-1} (-1)^s \binom{j-1}{s} (e_{k+1-j+s,1+s} - e_{k+2-j+s,2+s}) \\
&= \sum_{s=0}^{j-1} (-1)^s \binom{j-1}{s} e_{k+1-j+s,1+s} + \sum_{s=0}^{j-1} (-1)^{s+1} \binom{j-1}{s} e_{k+2-j+s,2+s} \\
&= \sum_{s=0}^{j-1} (-1)^s \binom{j-1}{s} e_{k+1-j+s,1+s} + \sum_{s=1}^{j} (-1)^s \binom{j-1}{s-1} e_{k+1-j+s,1+s}
\end{aligned}$$

Note que para o caso $s = 0$, temos que o somatório acima de escreve como $e_{k+1-j,1}$. Logo

$$[\underbrace{v,\ldots,v}_{j},e_{k+1,1}] = e_{k+1-j,1} + \sum_{s=1}^{j}(-1)^s\left(\binom{j-1}{s}+\binom{j-1}{s-1}\right)e_{k+1-j+s,1+s} + (-1)^j e_{k+1,1+j}$$

Usando o fato que

$$\binom{j-1}{s}+\binom{j-1}{s-1} = \binom{j}{s}$$

obtemos a conclusão do lema.

□

O próximo lema nos permitirá reduzir o número de variáveis não comutativas de maneira adequada.

Lema 3.16. *Seja $f \in \mathbb{K}\langle X_1,\ldots,X_n\rangle \coprod [\mathbb{K}|w \in \Omega]$ (com $n > 1$) um polinômio admissível da forma*

$$f = \sum_{\sigma\in S_{n-1}}\sum_{j=1}^{n}\sum_{a\in\mathcal{A}_\Omega}\lambda^a_{\sigma,j}(X_{\sigma(1)}\ldots X_{\sigma(j-1)}X_nX_{\sigma(j)}\ldots X_{\sigma(n-1)})^a$$

Seja $k \in \mathbb{N}$ tal que para cada $a \in \mathcal{A}_\Omega$, $|a^n| < k$ implica $\lambda^a_{\sigma,j} = 0$ para todo $\sigma \in S_{n-1}$ e todo $j \in \{1,\ldots,n\}$. Seja $\overline{w} = (w_1,\ldots,w_k)$ parte de alguma partição de Ω. Defina $\widetilde{\Omega} = \Omega - \{w_1,\ldots,w_k\}$ e seja A uma álgebra com unidade arbitrária. Então, o polinômio parcialmente comutativo

$$g \in \mathbb{K}\langle X_1,\ldots,X_{n-1}\rangle \coprod\coprod [\mathbb{K}|w \in \tilde{\Omega}]$$

definido por

$$g = \sum_{\sigma\in S_{n-1}}\sum_{j=1}^{n}\sum_{\widetilde{a}\in\widetilde{\mathcal{A}_\Omega}}\lambda^{\widetilde{a}}_{\sigma,j}(X_{\sigma(1)},\ldots X_{\sigma(j-1)})^{\widetilde{a}}U_n(X_{\sigma(1)}\ldots X_{\sigma(j-1)})^{\widetilde{a}}$$

onde $\widetilde{\mathcal{A}_\Omega} = \{\widetilde{a} \in \mathcal{A}_\Omega$ tal que $\widetilde{a}^n = \overline{w}\}$, satisfaz $g(A)e_{11} \subseteq f(M_{k+1}(A))$.

Demonstração. Seja $\Omega = \{w_1,\ldots,w_k,w_{k+1},\ldots,w_m\}$. Considere

$$x_1,\ldots,x_{n-1},u_n,u_{w_{k+1}},\ldots,u_{w_m} \in A$$

de modo que $u_n,u_{w_{k+1}},\ldots,u_{w_m}$ comutam entre si. Para provar o lema, temos que encontrar $\overline{x}_1,\ldots,\overline{x}_n,\overline{u}_{w_1},\ldots,\overline{u}_{w_m} \in M_{k+1}(A)$ de modo que

$$E_{V_{\overline{x}_1\ldots\overline{x}_n;\overline{u}_{w_1}\ldots\overline{u}_{w_m}}}(f) = E_{V_{x_1\ldots x_{n-1};u_n,u_{k+1}\ldots u_{w_m}}}(g)e_{11} \tag{3.5}$$

e $\overline{u}_{w_1}, \ldots, \overline{u}_{w_m}$ comutam. Defina

$$\begin{aligned}
\overline{x}_i &= x_i e_{11};\ \text{para } i = 1, \ldots, n-1 \\
\overline{x}_n &= u_n e_{k+1,1} \\
\overline{u}_{w_j} &= v,\ \text{para } j = 1, \ldots, k \\
\overline{u}_{w_l} &= u_{w_l} I,\ \text{para } l = k+1, \ldots, m
\end{aligned}$$

onde v é a matriz do lema anterior. As matrizes $\overline{u}_{w_1}, \ldots, \overline{u}_{w_m}$ comutam entre si, uma vez que o mesmo ocorre com os elementos $u_{w_{k+1}}, \ldots, u_{w_m}$. Afirmamos que a equação 3.5 é válida. Fixe $\sigma \in S_{n-1}$ e $j \in \{1, \ldots, n\}$ e tome um $a \in \mathcal{A}_\Omega$. Considere a expressão

$$E_{V_{\overline{x}_1 \ldots \overline{x}_n; \overline{u}_{w_1} \ldots \overline{u}_{w_m}}} \left(\lambda^a_{\sigma,j} X^a_n \right)$$

Se $\eta = |a^n| < k$, então, por hipótese, $\lambda^a_{\sigma,j} = 0$ e esta expressão é zero. Se a sequência a^n contém w_j para $j > k$, então

$$E_{V_{\overline{x}_1 \ldots \overline{x}_n; \overline{u}_{w_1} \ldots \overline{u}_{w_m}}} \left(X^a_n \right) = [\overline{u}_{a^n_1}, \ldots, \overline{u}_{a^n_\eta}, \overline{x}_n]$$

uma vez que $\overline{u}_{w_j} = u_{w_j} I$ comuta com todos os $\overline{u}_{w_s}$ e $\overline{x}_n$, segue que

$$E_{V_{\overline{x}_1 \ldots \overline{x}_n; \overline{u}_{w_1} \ldots \overline{u}_{w_m}}} \left(X^a_n \right) = 0.$$

Portanto, a expressão acima pode ser diferente de zero apenas se a^n contiver pelo menos k elementos de $\{w_1, \ldots, w_k\}$, ou seja, $a^n = \overline{w}$. Neste caso, pelo lema 3.15 tem-se que

$$E_{V_{\overline{x}_1 \ldots \overline{x}_n; \overline{u}_{w_1} \ldots \overline{u}_{w_m}}} \left(X^a_n \right) = [\underbrace{v, \ldots, v}_{k}, u_n e_{k+1,1}] = diag(u_n, \star, \ldots, \star).$$

Para tal a e $i = 1, \ldots, n-1$, temos

$$\begin{aligned}
E_{V_{\overline{x}_1 \ldots \overline{x}_n; \overline{u}_{w_1} \ldots \overline{u}_{w_m}}} \left(X^a_i \right) &= [\overline{u}_{a^i_1}, \ldots, \overline{u}_{a^i_\eta}, \overline{x}_i] \\
&= [u_{a^i_1} I, \ldots, u_{a^i_\eta} I, x_i e_{11}] \\
&= [u_{a^i_1}, \ldots, u_{a^i_\eta}, x_i] e_{11}.
\end{aligned}$$

Consequentemente

$$
\begin{aligned}
&E_{V_{\bar{x}_1\ldots\bar{x}_n;\bar{u}_{w_1}\ldots\bar{u}_{w_m}}}(f) = \\
&= \sum_{\sigma\in S_{n-1}}\sum_{j=1}^{n}\sum_{a\in\mathcal{A}_\Omega} E_{V_{\bar{x}_1\ldots\bar{x}_n;\bar{u}_{w_1}\ldots\bar{u}_{w_m}}}\left((X_{\sigma(1)},\ldots,X_{\sigma(j-1)})^a\right) \\
&\cdot E_{V_{\bar{x}_1\ldots\bar{x}_n;\bar{u}_{w_1}\ldots\bar{u}_{w_m}}}\left(\lambda^a_{\sigma,j}X^a_n\right) E_{V_{\bar{x}_1\ldots\bar{x}_n;\bar{u}_{w_1}\ldots\bar{u}_{w_m}}}\left((X_{\sigma(1)},\ldots,X_{\sigma(j-1)})^a\right) \\
&= \sum_{\sigma\in S_{n-1}}\sum_{j=1}^{n}\sum_{\widetilde{a}\in\widetilde{\mathcal{A}_\Omega}} E_{V_{x_1\ldots x_{n-1};u_{w_{k+1}}\ldots u_{w_m}}}\left((X_{\sigma(1)},\ldots,X_{\sigma(j-1)})^{\widetilde{a}}\right)e_{11} \\
&\cdot \lambda^{\widetilde{a}}_{\sigma,j}diag(u_n,\star,\ldots,\star)E_{V_{x_1\ldots x_{n-1};u_{w_{k+1}}\ldots u_{w_m}}}\left((X_{\sigma(1)},\ldots,X_{\sigma(j-1)})^{\widetilde{a}}\right)e_{11} \\
&= E_{V_{x_1\ldots x_{n-1};u_{w_{k+1}}\ldots u_{w_m}}}(g)e_{11}
\end{aligned}
$$

□

Agora estamos em posição de provar nosso teorema principal da seção.

Teorema 3.17. *Seja $\mathbb{K}$ um corpo infinito e seja $d\in\mathbb{N}$. Para cada polinômio parcialmente comutativo admissível f diferente de zero, existe um $s\in\mathbb{N}$ tal que $sl_d(F)\subseteq f(M_s(F))$.*

Demonstração. Procedemos por indução em n, ou seja, o número de variáveis não comutativas $X_1,\ldots,X_n$. O caso em que $n=1$ foi considerado no Lema 3.14.

Seja $n>1$ e assuma que o teorema é verdadeiro para todos os polinômios parcialmente comutativos admissíveis não nulos em $n-1$ variáveis não comutativas. Podemos escrever $f\in F\langle X_1,\ldots,X_n\rangle\coprod[mathbbK|w\in\Omega]$ como

$$f=\sum_{\sigma\in S_{n-1}}\sum_{j=1}^{n}\sum_{a\in\mathcal{A}_\Omega}\lambda^a_{\sigma,j}(X_{\sigma(1)}\ldots X_{\sigma(j-1)}X_nX_{\sigma(j)}\ldots X_{\sigma(n-1)})^a$$

para alguns $\lambda^a_{\sigma,j}\in\mathbb{K}$, nem todos zero. Suponha que o teorema não seja verdadeiro, ou seja,

$$sl_d(F)\nsubseteq f(M_s(F));\ \text{para todo } s\in\mathbb{N}$$

Seja k o menor inteiro não negativo tal que $\lambda^a_{\sigma,j}\neq 0$ para algum $\sigma\in S_{n-1}, j=1,\ldots,n$ e $a\in\mathcal{A}_\Omega$ com $|a^n|=k$. Observe que k satisfaz a suposição do Lema 3.16.

Seja $a^n=\underline{w}=(w_1,\ldots,w_k)$ o n-ésimo componente de uma partição, visto que $\lambda^a_{\sigma,j}\neq 0$ para algum $\sigma\in S_{n-1}$ e $j=1,\ldots,n$. Nosso objetivo é provar que para cada $i=1,\ldots,n$, tem-se que

$$\sum_{j=1}^{i}\lambda^{\widetilde{a}}_{\sigma,j}=0 \tag{3.6}$$

para algum $\sigma \in S_{n-1}$ e cada $\widetilde{a} \in \widetilde{\mathcal{A}_\Omega} = \{\widetilde{a} \in \mathcal{A}\Omega \mid \widetilde{a}^n = \underline{w}\}$. Isso implica $\lambda^{\widetilde{a}}_{\sigma,j} = 0$ para todo $\sigma \in S_{n-1}$, todo $\widetilde{a} \in \widetilde{\mathcal{A}_\Omega}$ e todo $j = 1, \ldots, n$ o que contradiz nossa escolha de w.

Seja $g \in \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle \coprod [\mathbb{K}|w \in \widetilde{\Omega} \cup \{n\}]$, onde $\widetilde{\Omega} = \Omega - \{w_1, \ldots, w_k\}$, definido por

$$g = \sum_{\sigma \in S_{n-1}} \sum_{j=1}^{n} \sum_{\widetilde{a} \in \widetilde{\mathcal{A}_\Omega}} \lambda^{\widetilde{a}}_{\sigma,j} (X_{\sigma(1)}, \ldots, X_{\sigma(j-1)})^{\widetilde{a}} U_n (X_{\sigma(1)}, \ldots, X_{\sigma(n-1)})^{\widetilde{a}}$$

Pelo Lema 3.16, temos $g(A)e_{11} \subseteq f(M_{k+1}(A))$ para uma álgebra com unidade arbitrária A. Uma vez que g não tem a variável X_n, podemos substituir $\widetilde{a}$ pela $(n-1)-upla$ obtida tomando os primeiros $n-1$ componentes de $\widetilde{a}$. Essas uplas são exatamente os elementos de $\mathcal{A}_{\widetilde{\Omega}}$. Por isso,

$$g = \sum_{\sigma \in S_{n-1}} \sum_{j=1}^{n} \sum_{a \in \mathcal{A}_\Omega} \lambda^{a}_{\sigma,j} (X_{\sigma(1)}, \ldots, X_{\sigma(j-1)})^{a} U_n (X_{\sigma(j)}, \ldots, X_{\sigma(n-1)})^{a}$$

onde $\mathcal{A}_{\widetilde{\Omega}}$ contém $(n-1)uplas$ e $\widetilde{a}$ é a $upla$ obtida adicionando a sequência w ao final de a. Seja

$$\pi : \mathbb{K}\langle X_1, \ldots, X_{n-1}\rangle \coprod \mathbb{K}[U_w|w \in \widetilde{\Omega} \cup \{n\}] \longrightarrow F\langle X_1, \ldots, X_{n-1}\rangle \coprod \mathbb{K}[U_w|w \in \widetilde{\Omega}]$$

o homomorfismo que envia U_n para 1 e fixa cada $X_1, \ldots, X_{n-1}$ e U_w remanescente. Temos que

$$\pi(g) = \sum_{\sigma \in S_{n-1}} \sum_{a \in \mathcal{A}_\Omega} \left(\sum_{j=1}^{n} \lambda^{\widetilde{a}}_{\sigma,j} \right) (X_{\sigma(1)} \ldots X_{\sigma(j-1)} X_{\sigma(j)} \ldots X_{\sigma(n-1)})^{a}$$

Obviamente, $\pi(g)(A) \subseteq g(A)$ vale para toda álgebra unitária A. Afirmamos que $\pi(g) = 0$. Na verdade, se isso não fosse verdade, então uma vez que $\pi(g)$ é um polinômio parcialmente comutativo admissível em $n-1$ variáveis não comutativas, segue da hipótese de indução que existe um $s \in \mathbb{N}$ tal que

$$\begin{aligned} sl_d(\mathbb{K}) \subseteq \pi(g)(M_s(\mathbb{K})) \subseteq g(M_s(\mathbb{K})) = g(M_s(\mathbb{K}))e_{11} \\ \subseteq f(M_{k+1}(M_s(\mathbb{K}))) = f(M_{(k+1)s}(\mathbb{K})) \end{aligned}$$

o que contradiz nossa suposição inicial. Agora, o Lema 3.12 implica na Equação 3.6 para $i = n$. Usando a igualdade

$$\lambda^{\widetilde{a}}_{\sigma,j} = -\sum_{j=1}^{n-1} \lambda^{\widetilde{a}}_{\sigma,j},$$

temos que

$$g = \sum_{\sigma\in S_{n-1}} \sum_{a\in\mathcal{A}_\Omega} \sum_{j=1}^{n} \lambda^{\widetilde{a}}_{\sigma,j} (X_{\sigma(1)} \dots X_{\sigma(j-1)})^a [U_n, (X_{\sigma(j)} \dots X_{\sigma(n-1)})^a]$$

Para $a \in \mathcal{A}_\Omega$, denote por $a \triangleleft_i n$ a partição de $\Omega' = \widetilde{\Omega} \cup \{n\}$ obtida adicionando n ao início da sequência a^i.

O conjunto $\mathcal{A}'_\Omega$ está em correspondência bijetiva com a união disjunta $\cup_{i=1}^{n-1} \mathcal{A}'_\Omega$ via $a' = a \triangleleft_i n$ (para alguma permutação fixa σ). Por definição, temos

$$\left[U_n, X^a_{\sigma(i)}\right] = X^{a\triangleleft_i n}_{\sigma(i)}$$

e, assim, usando a fórmula $[X, YZ] = [X,Y]Z + Y[X,Z]$ várias vezes,

$$\begin{aligned}
&[U_n, (X_{\sigma(j)} \dots X_{\sigma(n-1)})^a] = \\
&= \sum_{i=j}^{n-1} (X_{\sigma(j)} \dots X_{\sigma(i-1)})^a [U_n, (X^a_{\sigma(i)}] (X_{\sigma(i+1)} \dots X_{\sigma(n-1)})^a \\
&= \sum_{i=j}^{n-1} (X_{\sigma(j)} \dots X_{\sigma(n-1)})^{a\triangleleft_i n}
\end{aligned}$$

Portanto

$$\begin{aligned}
g &= \sum_{\sigma\in S_{n-1}} \sum_{a\in\mathcal{A}_\Omega} \sum_{j=1}^{n-1} \lambda^{\widetilde{a}}_{\sigma,j} (X_{\sigma(1)} \dots X_{\sigma(j-1)})^a \sum_{i=j}^{n-1} (X_{\sigma(j)} \dots X_{\sigma(n-1)})^{a\triangleleft_i n} \\
&= \sum_{\sigma\in S_{n-1}} \sum_{a\in\mathcal{A}_\Omega} \sum_{j=1}^{n-1} \sum_{i=j}^{n-1} \lambda^{\widetilde{a}}_{\sigma,j} (X_{\sigma(1)} \dots X_{\sigma(n-1)})^{a\triangleleft_i n}
\end{aligned}$$

Alterando a ordem de somatório e usando a correspondência bijetiva mencionada, obtemos

$$\begin{aligned}
g &= \sum_{\sigma\in S_{n-1}} \sum_{i=1}^{n-1} \sum_{a\in\mathcal{A}_\Omega} \left(\sum_{j=1}^{i} \lambda^{\widetilde{a}}_{\sigma,j} \right) (X_{\sigma(1)} \dots X_{\sigma(n-1)})^{a\triangleleft_i n} \\
&= \sum_{\sigma\in S_{n-1}} \sum_{a'\in\mathcal{A}'_\Omega} (X_{\sigma(1)} \dots X_{\sigma(n-1)})^{a'}
\end{aligned}$$

Aqui, $\sigma(i)$ é o componente unicamente determinado por a' que contém n, a é o $(n-1)$ t-upla obtida pela omissão de n em $\sigma(i)$-ésimo componente da sequência a', e $\widetilde{a}$ é a n t-upla obtida de a como antes - i e $\hat{a}$ portanto depende apenas de σ e a'. Temos $g = 0$, caso contrário, pela hipótese de indução (g é um polinômio parcialmente comutativo admissível em $n-1$ variáveis

não comutáveis), existiria um $s \in \mathbb{N}$ tal que

$$sl_d(F) \subseteq g(M_s(F)) \subseteq f(M_{(k+1)s}(F))$$

Agora, o Lema 3.12 implica na equação 3.6 para todo $i = 1, \ldots, n-1$. Como o caso $i = n$ foi estabelecido anteriormente, isso conclui a prova. □

Uma vez que polinômios multilineares não comutativos são exemplos especiais de polinômios parcialmente comutativos admissíveis, o teorema 3.17 também é válido quando $f \neq 0$ é um polinômio não-comutativo qualquer.

Seja $M_\infty(\mathbb{K})$ a álgebra de todas as matrizes finitárias, denotando por $sl_\infty(\mathbb{K})$ o espaço de todas as matrizes finitárias de traço zero, temos então o seguinte corolário para o teorema.

Corolário 3.18. *Seja $\mathbb{K}$ um corpo infinito e seja f um polinômio multilinear diferente de zero. Então $sl_\infty(\mathbb{K}) \subseteq f(M_\infty(\mathbb{K}))$.*

3.4 Quando span $f(A)$ é igual a A?

Um exemplo importante de uma álgebra associativa sobre um corpo de característica zero que coincide com seu comutador é a álgebra de operadores diferenciais com coeficientes polinomiais [6], no entanto, esta álgebra não tem identidades polinomiais. Por outro lado, se o corpo apresenta característica positiva, então surgem algumas identidades não triviais, embora a coincidência com o comutador é perdida.

Entretanto, no caso de característica positiva $p > 0$, é possível construir uma álgebra associativa que coincida com seu comutante. Tal álgebra é fácil de obter indutivamente como segue: O comutador de uma álgebra matricial é formado pelas matrizes com traço zero, e o traço de uma matriz consistindo de p blocos idênticos na diagonal principal é igual a zero. Portanto, toda álgebra A de dimensão finita admite, para todo $x \in A$, uma extensão de dimensão finita por elementos z, t tal que $[z, t] = x$. No entanto, a álgebra obtida dessa maneira não possui identidades não triviais.

Então uma questão natural se manifesta, existe uma PI-álgebra A coincidente com seu comutador $[A, A]$? O principal teorema do artigo [6] responde a esta pergunta.

Lema 3.19. *Se A é uma PI-álgebra associativa então $A \neq [A, A]$.*

Nosso primeiro teorema mostra que a condição $span f(A) = A$ valer para qualquer polinômio

é equivalente à condição que isso vale para o polinômio comutador $[X_1, X_2]$. Antes de declara-lo, registramos mais algumas definições e observações.

Definição 3.20. Seja $f = f(x_1, ..., x_m) \in F\langle X\rangle$ qualquer polinômio. Definimos $\hat{f}$ como

$$\hat{f}(x_1, ..., x_{2m}) = [f(x_1, ..., x_m), f(x_{m+1}, ..., x_{2m})].$$

Observe que $\hat{f}$ uma identidade de A, se e somente se, $f(A)$ é um conjunto comutativo. Finalmente, ressaltamos que, se A é unitário, $f(A)$ é invariante sob conjugação, isto é, $af(A)a^{-1} = f(A)$, para cada elemento invertível $a \in A$. Na verdade, isso decorre de

$$af(a_1, a_2, \cdots, a_m)a^{-1} = f(aa_1a^{-1}, aa_2a^{-1}, \cdots, aa_ma^{-1}),$$

onde $a_1, \cdots, a_m$ são elementos arbitrários em A.

O seguinte lema, é um resultado da teoria de Lie dos anéis associativos de Herstein e será usado em várias demonstrações.

Lema 3.21. *Seja A uma álgebra sobre um corpo infinito F e seja $f \in F\langle X\rangle$.*

1. *Se $[A, A] \nsubseteq span f(A)$, então $\hat{f}$ é uma identidade de uma imagem homomórfica não nula de A.*

2. *Se A é álgebra simples e $\hat{f}$ é uma identidade de A, então f é uma identidade polinomial ou é um polinômio central de A.*

Demonstração. Seja $L = span f(A)$ um ideal de Lie da álgebra A.

1. Seja $I = id([L, L])$ o ideal de A gerado por $\hat{f}(A)$ (ou seja, o ideal gerado por $[L, L]$). Note que $[A, I] \subseteq L$. De fato, pelo lema 2.113 temos que $I \subseteq N(L)$, e consequentemente temos que

   $$[A, I] \subseteq [A, N(L)].$$

   Por outro lado, pela definição de $N(L)$ temos $[N(L), A] \subseteq L$ e portanto

   $$[A, I] \subseteq L.$$

   Como por hipótese $[A, A] \nsubseteq L$, tem-se que I é um ideial próprio de A, em outras palavras $I \neq A$. Claro que $\hat{f}(A) \subseteq I$, e tem-se que $\hat{f}$ é uma identidade polinomial de A/I. E sabemos que todo quociente de A é imagem homomórfica de A.

2. É fácil ver que $\hat{f}$ é uma identidade de A, se e somente se, $f(A)$ é um conjunto comutativo. Assim, a condição de que $\hat{f}$ é uma identidade de A significa que L é um ideal de Lie comutativo de A. Nosso objetivo é provar que L está contido no centro Z de A.

   Dividiremos essa demonstração em dois casos. Suponhamos no primeiro caso que a característica do corpo deve ser diferente de 2 ou que A não tenha dimensão 4 sobre seu centro Z.

   De acordo com o Teorema 2.135, temos que $L \subseteq Z$ ou $[A, A] \subseteq L$. Caso interessante é quando $[A, A] \subseteq L$. Chame S a subálgebra gerada por L, pelo corolário 2.136, temos que S é igual a A. Note que mostrar $L \subseteq Z$ é mostrar que $[L, A] = 0$.

   De fato, seja $x \in [L, A]$, então x é combinação linear de elementos da forma $[l, a]$, com $l \in L$ e $a \in A = S$, como $a \in S$ então $a = g(\alpha_1, \alpha_2, \cdots, \alpha_n)$ onde $g \in K\langle X \rangle$ e $\alpha_1, \cdots, \alpha_n \in L$, sem perda de generalidade tome a como sendo o monômio $a = \alpha_1 \cdots \alpha_m$. Note que $\hat{f}$ ser identidade para A implica em $[L, L] = 0$, e consequentemente $[L, \alpha_i] = 0$, para todo $i \in \{1, \cdots, m\}$.

   Pela hipótese de indução temos que $[l, \alpha_1 \cdots \alpha_{m-1}] = 0$ e consequentemente

$$\begin{aligned}
[l, a] &= [l, \alpha_1 \cdots \alpha_m] \\
&= \alpha_1 \cdots \alpha_{m-1}\cancel{[l, \alpha_m]} + \cancel{[l, \alpha_1 \cdots \alpha_{m-1}]}\alpha_m \\
&= 0
\end{aligned}$$

   No segundo caso, suponha que F tenha a característica 2 e que A tenha dimensão 4 sobre Z. Seja K o fecho algébrico de Z e seja $S = K \otimes_Z A$ a extensão escalar de A a K.

   Em 1905 Joseph Wedderburn provou que todo anel simples com dimensão finita sobre um anel de divisão é isomorfo a um anel de matriz $n \times n$ sobre um anel de divisão D, onde ambos n e D são unicamente determinados, um fato que hoje é conhecido como Teorema de Wedderburn, sabendo disso temos que $S \simeq M_2(K)$ (Ver Teorema 2.80) . Além disso pelo Teorema 2.103 temos que $\hat{f}$ é uma identidade de S.

   Assim, $V = span f(M_2(K))$ é um ideal de Lie comutativo de $M_2(K)$, além disso, como f é invariante sob conjugação temos adicionalmente que $aVa^{-1} \subseteq V$ para toda matriz invertível $a \in M_2(K)$.

   Suponha por absurdo que $V \nsubseteq Z(M_2(K))$, então V contém uma matriz não-escalar v. Se v é uma matriz diagonal, então, ao comuta-la com a matriz unitária e_{12}, vemos que V também contém matrizes não-diagonais. De fato, sem perda de generalidade, podemos

supor

$$v = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \text{ tal que } \beta \neq 0$$

Comutando v com e_{11} duas vezes, vemos que

$$w = \begin{bmatrix} 0 & \beta \\ \gamma & 0 \end{bmatrix} \in V$$

Por outro lado, $[w, e_{21}] = \beta * I_2$, portanto V contém a matriz identidade. Assim,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & \beta \\ \gamma & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \gamma & \beta+\gamma \\ \gamma & \gamma \end{bmatrix} \in V$$

a partir do qual se infere facilmente que V contém a matriz unitária e_{12}. Consequentemente,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot e_{12} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = e_{21} \in V$$

O que é uma contradição, já que $[e_{12}, e_{21}] = id_2 \neq 0$, em outras palavras esses elementos não comutam. Portanto, V está contido no centro de $M_2(K)$. Isso implica que $f(A) \subseteq Z$.

□

Teorema 3.22. *Seja A uma álgebra sobre um corpo infinito F. Se $[A,A] = A$, então $span f(A) = A$ para todo polinômio não constante $f \in F\langle X\rangle$.*

Demonstração. Pelo Lema 3.19, não existe uma PI-álgebra A diferente de zero que coincide $[A,A]$. Como por hipótese $[A,A] = A$ significa que nenhuma imagem homomórfica diferente de zero de A é uma PI-álgebra. Pela contrapositiva do Lema 3.21, temos que $[A,A] = A \subseteq span f(A)$. Como $span(f(A)) \subseteq A$ é trivial, segue que $span f(A) = A$, como desejado.

□

3.5 Quando span $f(A)$ contém $[A,A]$?

Segue agora uma variação do teorema anterior.

Teorema 3.23. *Seja A uma álgebra unitária sobre um corpo F com $char(F) = 0$. Se $1 \in [A,A]$, então $[A,A] \subseteq span f(A)$ para cada polinômio não constante $f \in F\langle X\rangle$.*

Demonstração. Suponha que o teorema não seja válido. Então, pelo Lema 3.21 parte (a), existe um ideal próprio de A tal que $S = A/I$ é uma PI-álgebra. Seja M um ideal maximal de A contendo I. Então $R = A/M$ é uma PI-álgebra simples.

Para mostrar que R é uma PI-álgebra, observe existe um o homomorfismo

$$\begin{aligned} \varphi : A/I &\longrightarrow R \\ a+I &\mapsto a+M \end{aligned}$$

que é sobrejetor, logo R é imagem homomórfica de A/I que é uma PI-álgebra. Por outro lado R é simples. De fato, chame $\pi : A \to R$ a projeção canônica e seja $J \neq 0$ um ideal de R. Então $\pi^{-1}(J) \subseteq A$ é um ideal de A (estritamente) contendo M, mas pela maximalidade de M temos obrigatoriamente que $\pi^{-1}(J) = A$. No entanto, isso implica que $J = R$. Portanto R é simples.

Consequentemente, pelo Lema 2.74, tem-se que R é de dimensão finita sobre seu centro Z. Obviamente, $1 \in [A,A]$ implica $1 \in [R,R]$. Seja K o fecho algébrico de Z e $S = K \otimes_Z R$ a extensão escalar de R. Observe que $1 \in [S,S]$.

No entanto, S é isomorfo a $M_n(K)$ para algum $n \geqslant 1$, então para cada elemento de S podemos considerar o seu traço. O traço de 1 é $n \neq 0$ (uma vez que $char(F) = 0$), enquanto o traço de qualquer elemento em $[S,S]$ é igual a 0. Esta contradição mostra que o $[A,A] \subseteq spanf(A)$.

□

O próximo exemplo mostra que a suposição de que $char(F) = 0$ é realmente necessária.

Exemplo 3.24. Suponha que $char(F) = p$ onde p é um número primo. Seja $A = M_p(F)$. Então a matriz identidade $I_d \in M_p(F)$ tem o traço 0 e pelo teorema de Shoda Albert e Muckenhoupt, tem-se que, é um comutador. No entanto, para todo f sendo uma identidade polinomial ou um polinômio central de A tem-se que $[A,A] \nsubseteq spanf(A)$.

Concluímos esta secção com dois resultados que fornecem uma imagem mais completa de quando $spanf(A)$ contém $[A,A]$. Primeiro, registramos um corolário do Lema 3.21 que decorre da contrapositiva dos dois itens do lema.

Corolário 3.25. *Seja A uma álgebra simples sobre um corpo infinito F. Se $f \in F\langle X\rangle$ não é uma identidade polinomial nem é um polinômio central de A, então $[A,A] \subseteq spanf(A)$.*

Observamos que uma consequência deste corolário é que faz sentido enunciar a conjectura de Mesyan para um limitante mais justo.

Proposição 3.26. *Sejam $\mathbb{K}$ um corpo, $n \geqslant 2$ e $m \geqslant 2$, inteiros em que $car(K) \nmid n$ e $f(x_1, \cdots, x_m) \in K\langle X\rangle$ um polinômio multilinear não nulo. Se $m \leqslant 2n-1$, então o $\mathbb{K}$-subespaço $span(f(M_n(\mathbb{K}))$ contém $sl_n(\mathbb{K})$.*

Nosso último teorema desta secção é um aprimoramento desse corolário. Chamamos de ideal do comutador de A, denotado por $C_{[A,A]}$ como ideal gerado por todos os comutadores em A.

Teorema 3.27. *Seja A uma álgebra unitária sobre um corpo infinito F, e seja $f \in F\langle X\rangle$. As duas declarações a seguir são equivalentes:*

1. *f não é uma identidade polinomial nem é um polinômio central de qualquer imagem homomórfica diferente de zero de A.*

2. *$[A,A] \subseteq spanf(A)$ e A é igual ao seu ideal de comutador $C_{[A,A]}$.*

Demonstração.

Chame de $X = A/C_{[A,A]}$

(1) $\Longrightarrow$ (2)

Observe que X é uma álgebra comutativa, logo todo polinômio é uma identidade ou um polinômio central de X. Assim, basta mostrar que $[A,A] \subseteq spanf(A)$. Suponha que isso não seja verdade. Então, pelo Lema 3.21 parte (a), existe um ideal próprio I de A tal que $\hat{f}$ é uma identidade de A/I. Pegue um ideal maximal M que contém I. Então $\widetilde{f}$ é uma identidade da álgebra simples A/M, e então o Lema 3.21 parte (b) nos diz que é $\hat{f}$ uma identidade ou um polinômio central de A/M, o que contradiz (1).

(2) $\Longrightarrow$ (1)

Suponha por absurdo que exista um ideal próprio I de A tal que f seja uma identidade polinomial ou um polinômio central de A/I. Isso significa que $[f(A), A] \subseteq I$, que junto com $[A,A] \subseteq spanf(A)$ resulta em $[[A,A],A] \subseteq I$. Escolha um ideal maximal M que contenha I. Então $R = A/M$ é uma álgebra simples que satisfaz $[[R,R],R] = \{0\}$. Ou seja, cada comutador em R está no centro Z de R.

Em particular, $[x,y]x = [x,yx] \in Z$, e assim $[x,y]^2 = [[x,y]x,y] = [[x,yx],y] = 0$ para todo os $x, y \in R$. Mas pelo Proposição 2.73 o centro de uma álgebra simples é um corpo.

Tome $x, y \in R$ tais que $[x,y] \in Z$. Como Z é um corpo, existe $[x,y]^{-1} \in Z$ tal que $[x,y][x,y]^{-1} = 1 \Longrightarrow [x,y]^2[x,y]^{-1} = [x,y] \Longrightarrow [x,y] = 0$. Portanto R é comutativo.

Tome $a+M$ e $b+M \in R$, como R é comutativo então $ab+M = ba+M \Leftrightarrow ab-ba \in M$, Ou seja, $[A,A] \subseteq M$, contradizendo o fato que $C_{[A,A]} = A$. □

Fazemos dois comentários finais sobre o Teorema 3.27 . Primeiro, a necessidade da hipótese em (2) que $A = C_{[A,A]}$ é evidente a partir do caso em que A é comutativo. Em segundo lugar, a afirmação (1) pode ser formulada de forma equivalente, pois o ideal gerado por $[f(A), A]$ é igual a A; por outro lado, a afirmação (2) diz respeito apenas ao espaço gerado de $f(A)$.

Capítulo 4

Resultados do tipo Waring para imagens de polinômios

Seja A uma álgebra e seja f um polinômio não constante e não comutativo, neste capítulo estabelecemos alguns resultados do tipo Waring para imagens de polinômios. Mostramos que se C é uma álgebra unitária e comutativa sobre um corpo $\mathbb{K}$ de característica 0, A é a álgebra da matriz $M_n(C)$ e o polinômio f não é uma identidade polinomial nem um polinômio central de $M_n(\mathbb{K})$, então cada comutador em A pode ser escrito como uma diferença de dois elementos, cada um dos quais é uma soma de 7788 elementos de $f(A)$ (se $C = \mathbb{K}$ é um corpo algebricamente fechado, então 4 elementos bastam).

4.1 Polinômios localmente linearmente dependentes

O objetivo da primeira secção é fornecer algumas ferramentas necessárias nas demonstrações dos resultados principais. No entanto, como essas ferramentas são interessantes por si mesmas, iremos trabalhar com resultados mais gerais do que seria necessário para os resultados principais da secção.

Os resultados desta seção foram extraídos do artigo [9], além de algumas definições vistas em [8].

Seja A uma álgebra sobre um corpo $\mathbb{K}$. Dizemos que os polinômios não comutativos $f_1, \ldots, f_s \in \mathbb{K}\langle X_1, \ldots, X_m\rangle$ são localmente linearmente dependentes, se para qualquer $a = (a_1, \ldots, a_m) \in A^m$, os elementos $f_1(a), \ldots, f_s(a)$ são linearmente dependentes sobre $\mathbb{K}$. Caso contrário, dizemos que eles são localmente linearmente independentes.

É fácil ver que se $\{f_1 \ldots, f_n\}$ é linearmente dependente de forma usual então também é

localmente linearmente dependente. O inverso não é verdadeiro. Por exemplo, um único polinômio f é localmente linearmente dependente ,se e somente se, f é uma identidade polinomial de A.

Da mesma forma, se o centro de A consiste em múltiplos escalares da unidade, então $\{1, f\}$ são localmente linearmente dependente, se e somente se, f é uma identidade polinomial ou um polinômio central de A.

O próximo exemplo mostra que mesmo em um conjunto de polinômios linearmente independente na definição usual, pode-se obter um conjunto localmente linearmente independente.

Exemplo 4.1. Seja $M_n(\mathbb{K})$ a álgebra das matrizes $n \times n$, os polinômios linearmente independentes $\{1, X, \ldots, X^n\}$ são localmente linearmente dependentes em $M_n(\mathbb{K})$.

De fato, seja $p(\lambda) = x^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_0$ o polinômio característico de $A \in M_n(\mathbb{K})$. Pelo Teorema de Cayley-Hamilton, tem se que $p(A) = A^n + \alpha^{n-1}A^{n-1} + \ldots + \alpha_0 I_d = 0$, logo é conjunto localmente LD .

Chame de c_s o $s-$ésimo polinômio de Capelli, ou seja,

$$c_s(x_1, \cdots, x_s, y_1, \cdots, y_{s-1}) = \sum_{\sigma \in S_s} (-1)^{\sigma} x_{\sigma(1)} y_1 x_{\sigma(2)} y_2 \cdots y_{s-1} x_{\sigma(s)}$$

Consideramos um polinômio especial que desempenha um papel importante na teoria das identidades polinomiais. Vamos começar com um exemplo. Considere o polinômio

$$h(x_1, x_2, x_3, y) = x_1 y x_2 x_3 - x_1 y x_3 x_2 + x_2 y x_3 x_1 - x_2 y x_1 x_3 + x_3 y x_1 x_2 - x_3 y x_2 x_1$$

É fácil ver que ao substituir a variável x_1 pela váriavel x_2, obtemos $h(x_2, x_2, x_3, y) = 0$, assim como $h(x_1, x_2, x_1, y) = 0$ e $h(x_1, x_2, x_2, y) = 0$

Definição 4.2. Um polinômio multilinear $f = f(x_1, \ldots, x_n, y_1, \ldots, y_r) \in \mathbb{K}\langle X \rangle$ é dito ser alternado em $x_1, \ldots, x_n$, se f torna-se zero sempre que se substitui x_j por x_i, com $1 \leqslant i < j \leqslant n$

Substituindo x_1 por $x_1 + x_2$, no polinômio alternado $f(x_1, x_2, \ldots, x_n, y_1, \ldots, y_r) = 0$ tem-se que

$$\begin{aligned} f(x_1 + x_2, x_1 + x_2, \ldots, x_n, y_1, \ldots, y_r) = f(x_1, x_1, \ldots, x_n, y_1, \ldots, y_r) + f(x_1, x_2, \ldots, x_n, y_1, \ldots, y_r) \\ + f(x_2, x_1, \ldots, x_n, y_1, \ldots, y_r) + f(x_1, x_1, \ldots, x_n, y_1, \ldots, y_r)_0 \end{aligned}$$

O que implica que

$$f(x_1, x_2, \ldots, x_n, y_1, \ldots, y_r) = -f(x_2, x_1, \ldots, x_n, y_1, \ldots, y_r).$$

Observação 4.3. Note que f muda de sinal se trocarmos qualquer par de variáveis x_i e x_j. O nome alternado vem desta condição.

A principal razão para considerar polinômios alternados é sua conexão com a noção de dependência linear.

Lema 4.4. *Seja A uma álgebra e $a_1,\dots,a_n \in A$ elementos linearmente dependentes. Se um polinômio multilinear $f = f(x_1,\dots,x_n,y_1,\dots,y_r) \in K\langle X\rangle$ é alternado em $x_1,\dots,x_n$, então $f(a_1,\dots,a_n,b_1,\cdots,b_r)$ =0 para todo $b_1,\dots,b_r \in A$.*

Demonstração. Como $a_1,\dots,a_n \in A$ são elementos linearmente dependente, sem perda de generalidade chame $a_n = \sum_{i=1}^{n-1} \lambda_i a_i$ para $\lambda_i \in \mathbb{K}$. Consequentemente tem-se que

$$f(a_1,\dots,a_n,b_1,\cdots,b_r) = \sum_{i=1}^{n-1} \lambda_i f(a_1,\dots,a_{n-1},a_i,x_1,\dots,x_r) = 0$$

Já que f é alternado. □

Duas famílias de polinômios alternados são de especial importância.

Exemplo 4.5.

1. O primeiro já visto nos teorema anteriores é o polinômio de Capelli. De fato, basta observar que para $n = 2$ tem-se que

   $$c_2(x_1,x_2,a) = x_1 a x_2 - x_2 a x_1$$

   que é claramente um polinômio alternado em x_1 e x_2. E para $n > 2$ tem-se que

   $$c_n(x_1,\dots,x_n,y_1,\dots,y_{n-1}) = \sum_{i=1}^{n} (-1)^{i-1} x_i y_1 c_{n-1}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n,y_2,\dots,y_{n-1})$$

   é alternado em x_1,...,x_n.

2. Já o segundo exemplo, consiste em polinômios que se alternam em todos as variáveis. Seja $n \geqslant 2$. O polinômio

   $$s_n(x_1,\dots,x_n) = \sum_{\sigma\in S_n} (-1)^{\sigma} x_{\sigma(1)} x_{\sigma(2)} \dots x_{\sigma(n)}$$

   é chamado de polinômio standard de grau n. Observe que

   $$s_n(x_1,\dots,x_n) = c_n(x_1,\dots,x_n,1,\dots,1).$$

Seja $R \neq 0$ uma álgebra prima. Denote por τ o conjunto de todos os ideais diferentes de zero de R.Veja que τ é fechado em produtos e, portanto, também em interseções finitas.

Dotamos o conjunto de todos os pares (f, I), onde $I \in \tau$ e $f : I \longrightarrow R$ é um homomorfismo de módulos, pela seguinte relação, que é prontamente vista como equivalência: $(f, I) \sim (g, J)$, se f e g coincidem em algum $K \in \tau$ tal que $K \subseteq I \cap J$. Escreva $[f, I]$ para a classe de equivalência determinada por (f, I), e denote por $Q(R)$ o conjunto de todas as classes de equivalência, equipadas com adição e multiplicação

$$[f_1, I_1] + [f_2, I_2] = [f_1 + f_2, I_1 \cap I_2]$$
$$[f_1, I_1] \cdot [f_2, I_2] = [f_1 \circ f_2, I_2 I_1]$$

Denotaremos $f_1 \circ f_2$ simplesmente por $f_1 f_2$

Observação 4.6. Note que $f_1 f_2$ de fato está definido em $I_2 I_1$, já que $f_2(I_2 I_1) = f_2(I_2) I_1 \subseteq I_1$. É fácil ver que essas operações estão bem definidas, de fato, assumimos que $(f_1, I_1) \sim (g_1, J_1)$ e $(f_2, I_2) \sim (g_2, J_2)$, ou seja, exite $K_i \in \tau$ tal que $K_i \subseteq I_i \cap J_i$ onde $f_i(x) = g_i(x)$ com $x \in K_i$ e $i = \{1, 2\}$. Então $(f_1 + f_2)(x) = (g_1 + g_2)(x)$ com $x \in K_1 \cap K_2 \in \tau$ e por outro lado $(f_1 f_2)(x) = (g_1 g_2)(x)$ com $x \in K_2 K_1 \in \tau$.

Então $(f_1 + f_2, I_1 \cap I_2) \sim (g_1 + g_2, J_1 \cap J_2)$ e $(f_1 f_2, I_2 I_1) \sim (g_1 g_2, J_2 J_1)$. Logo verifica imediatamente que $Q(R)$ é um anel com elemento zero $[0, R]$ e unidade $[id_R, R]$.

Definição 4.7. Quando os τ for conjunto de ideias a direita, então denotaremos por $Q_d(R)$ o anel $Q(R)$. Neste caso $Q_d(R)$ será chamado de anel quociente à direita de Martindale.

Observação 4.8. O anel quociente à esquerda de Martindale $Q_e(R)$ é construído analogamente através dos homomorfismos do R-módulo à esquerda. Em geral $Q_d(R) \neq Q_e(R)$

Teorema 4.9. *Seja $R \neq 0$ um anel primo, e seja τ o conjunto de todos os ideais diferentes de zero de R. O anel $Q_d(R)$ tem as seguintes propriedades:*

1. *$Q_d(R)$ é um anel unitário contendo R como um subanel.*
2. *Para cada $q \in Q_d(R)$ existe $I \in \tau$ tal que $qI \subseteq R$.*
3. *Para cada $q \in Q_d(R)$ e $I \in \tau$, $qI = 0$ implica em $q = 0$.*
4. *Se $I \in \tau$ e $f : I \longrightarrow R$ é um homomorfismo de R-módulo a direita, então existe elemento $q \in Q_d(R)$ tal que $f(x) = qx$ para todo $x \in I$.*

Demonstração.

1. Já sabemos que $Q_d(R)$ é um anel unitário. Observe que

$$\begin{aligned} i : R &\hookrightarrow Q_d(R) \\ r &\mapsto [E_r, R] \end{aligned}$$

onde E_r é a função de multiplicação à esquerda, isto é, $E_r(x) = rx$ para todo $x \in R$.

Note primeiro que i é uma função injetora. De fato, tome $a \neq b$, dois elementos em R, e suponha por absurdo que suas imagens são iguais, logo

$$[E_a, R] = [E_b, R] \Longrightarrow E_a(x) = E_b(x) \text{ com } x \in K \subseteq R$$

com K um ideal não nulo de R. Mas então por definição da função multiplicação tomando $0 \neq x \in K$ tem-se

$$(a-b)x = 0 \text{ como } R \text{ é primo, segue que } (a-b) = 0 \Longrightarrow a = b$$

Logo i é uma inclusão. É fácil ver também que i é um homomorfismo. De fato, seja $a, b \in R$ tem-se que

$$\begin{aligned} i(a+b) = [E_{a+b}, R] = [E_a, R] + [E_b, R] = i(a) + i(b) \\ i(ab) = [E_{ab}, R] = [E_a, R][E_b, R] = i(a)i(b) \end{aligned}$$

Portanto identificando R com sua cópia isomórfica $i(R)$, podemos considerar R como um subanel de Q_d.

2. Seja $q = [f, I] \in Q_d(R)$. Para cada $x \in I$, já vimos que podemos representar cada elemento de R como imagem $i(R)$. Portanto

$$qx = [f, I][E_x, R] = [fE_x, RI] = [E_{f(x)}, R] = f(x) \in R$$

3. Seja $q = [f, I] \in Q_d(R)$, com $I \in \tau$, tal que $qI = 0$, chamando i um elemento do ideal I, pelo item anterior tem-se que $qi = f(i) = 0$ logo $f(I) = 0$, consequentemente $q = [f, I] = 0$.

Seja $J \in \tau$, um ideal qualquer, tal que $qJ = 0$, note que $(qI)J \subseteq qJ$. De fato, seja $y \in (qI)J$, temos que $y = qij$, com $i \in I$ e $j \in J$. Seja $[E_i, R]$ e $[E_j, R]$ as classes de equivalência de

i e j respectivamente. Temos que

$$\begin{aligned}([f,I][E_i,R])[E_j,R] &= ([fE_i,IR])[E_j,R] \\ &= [fE_iE_j,IR] \\ &= [f,I][E_{ij},R]\end{aligned}$$

Logo $y \in qJ$ e portanto $(qI)J \subseteq qJ$, então $(qI)J = 0$. Como R é primo e $qI \subseteq R$, implica que $qI = 0$ e portanto $q = 0$ como queríamos demonstrar.

4. Se $I \in \tau$ e $f : I \to R$ é um homomorfismo do R-módulo à direita, então pela demonstração do item (2), tomando $q \in Q_d(R)$ tem-se que $f(x) = qx$ para todo $x \in I$.

□

Outras propriedades de $Q_d(R)$ podem ser extraídas a partir Teorema 4.9. Por exemplo, o item (3) pode ser estendido como segue abaixo

Teorema 4.10. *Para cada $q_1, q_2 \in Q_d(R)$ e $I \in \tau$, $q_1 I q_2 = 0$ implica em $q_1 = 0$ ou $q_2 = 0$. Consequentemente, $Q_d(R)$ é um anel primo.*

Demonstração. Pelo item (2) do Teorema 4.9, podemos escolher $I_i \in \tau$ tal que $q_i I_i \subseteq R$, $i = \{1,2\}$. Portanto, podemos concluir que $(q_1 I_1)I(q_2 I_2) \subseteq (q_1 I q_2)I_2 = 0$, o que implica em $(q_1 I_1)I(q_2 I_2) = 0$, mas observe que pela Proposição 2.23, temos que $q_1 I_1 = 0$ ou $q_2 I_2 = 0$. Pelo item (3), obtemos $q_1 = 0$ ou $q_2 = 0$, como queríamos. □

Seja R é um anel primo diferente de zero. O centro de $Q_r(R)$ é chamado de centróide estendido de R. O centróide estendido de R será denotado por C. Como de costume, $Z(R)$ representará o centro de R e, τ representará o conjunto de todos os ideais diferentes de zero de R.

Lema 4.11. *Se $q \in Q_d(R)$ é tal que $qr = rq$ para todo $r \in R$, então $q \in C$. Portanto $Z(R)$ é um subanel de C.*

Demonstração. Seja $w \in Q_d(R)$. Devemos mostrar que o comutador $[q,w] = 0$. Pelo item (2), do Teorema 4.9, existe $I \in \tau$ tal que $wI \subseteq R$. Tome $x \in I$, por hipótese, temos que $qx = xq$, assim como $q(wx) = (wx)$q. Logo

$$qwx = (wx)q = w(xq) = w(qx) = wqx \Longrightarrow [q,w]x = 0$$

Assim, $[q,w]I = 0$ e consequentemente $[q,w] = 0$ pelo item (3) do Teorema 4.9. □

Ja vimos que os elementos em $Q_d(R)$ correspondem aos homomorfismos do R-módulo à direita, então os elementos em C correspondem aos homomorfismos do R-módulo à esquerda e à direita.

Lema 4.12. *Se $f : I \to R$, onde $I \in \tau$, é um homomorfismo de R-módulo, então existe $\lambda \in C$ tal que $f(x) = \lambda x$ para todo $x \in I$.*

Demonstração. Uma vez que f é homomorfismo de R-módulos à direita, pelo item (4) do Teorema 4.9, existe $q \in Q_d(R)$ tal que $f(x) = qx, x \in I$. Por outro lado, f também é homomorfismo de R-módulo à esquerda, em particular para todo $r \in R$ e $x \in I$ tem-se que $rx \in I$ e logo

$$q(rx) = f(rx) = rf(x) = r(qx) \Longrightarrow [q,r]x = 0$$

Portanto $[q,r]I = 0$, e pelo item (3) do teorema 4.9 tem-se que $[q,r] = 0$, aplicando o lema anterior com $\lambda = q$ chega-se no resultado pedido. □

Observação 4.13. Seja R uma álgebra sobre um corpo $\mathbb{K}$. Cada $\alpha \in \mathbb{K}$ dá origem ao homomorfismo do R-módulo à direita e à esquerda definindo

$$H : R \to R$$
$$x \mapsto \alpha x$$

Pelo o lema 4.12, pode se considerar $\mathbb{K}$ como um subcorpo de C.

Teorema 4.14. *O centróide estendido C de um anel primo diferente de zero R é um corpo.*

Demonstração. Seja λ um elemento não nulo de C. Devemos mostrar que λ é invertível. De fato, seja $I \in \tau$, tal que $\lambda I \subseteq R$, pelo teorema 4.9, parte (3) tem-se que $\lambda I \in \tau$. Note que

$$f : \lambda I \to R$$
$$\lambda x \mapsto x$$

está bem definido. De fato, $\lambda x = 0$ implica em $\lambda I_x = 0$, onde I_x é o ideal gerado por $\{x\}$ e consequentemente $x = 0$ (item (3) do teorema). Como f é homomorfismo do R-módulo á direita e esquerda, o lema 4.12 afirma que existe um $\lambda' \in C$ tal que $f(y) = \lambda' y$ para todo $y \in \lambda I$.

Chame $y = \lambda x_1$ com $x_1 \in I$. Por um lado, $f(\lambda x_1) = x_1$ por definição, e por outro lado, $f(\lambda x) = \lambda' \lambda x_1$ o que implica que

$$\lambda' \lambda x_1 = x_1 \Longrightarrow (\lambda' \lambda - 1)x_1 = 0$$

Ou seja, $(\lambda'\lambda - 1)I = 0$ o que implica pelo Teorema 4.9 que $\lambda'\lambda = 1$. E portanto λ é invertível. □

Observação 4.15. Podemos, portanto a partir do teorema anterior, considerar $Q_d(R)$ como uma álgebra sobre o corpo C.

Lema 4.16. *Dados $q_1, \ldots, q_n \in Q_r(R)$ então existe $I \in \tau$ tal que $q_i I \subseteq R$, para todo $i = 1, \ldots, n$.*

Demonstração. Pelo teorema 4.9 parte (2) tem-se que para todo $i \in \{1, 2, \ldots, n\}$ e para cada $q_i \in Q_d(R)$ existe um ideal $I_i \in \tau$ tal que $q_i I_i \subseteq R$. Tomando $I = I_1 \cap \ldots \cap I_n$ tem-se o resultado desejado. □

Iremos agora analisar a seguinte condição, seja D um anel de divisão, e $a, b \in D$ tais que $axb = bxa$ para todo $x \in D$. Uma possibilidade natural onde isso ocorre é quando a e b são linearmente dependentes sobre o centro de D. De fato, se a, b são linearmente dependentes sobre o centro $Z(D)$, então existe $\beta \in Z(D)$ tal que $a = \beta b$, então

$$axb = (\beta b)xb = bx(\beta b) = bxa.$$

Vejamos que esta é a única possibilidade.

Lema 4.17. *Sejam $a, b \in Q_r(R)$ e seja $I \in \tau$. Se $axb = bxa$ para todo $x \in I$, então a e b são linearmente dependentes sobre C.*

Demonstração. Pelo Lema 4.16 existe $J \in \tau$ tal que $aJ \subseteq R$ e $bJ \subseteq R$, note que por definição das classes de equivalências podemos substituir o papel de I por $I \cap J$, vemos que não há perda de generalidade ao assumir que $aI \subseteq R$ e $bI \subseteq R$.

Suponha que $a \neq 0$. Então $aI \neq 0$ pelo item 3 do Teorema 4.9 e, portanto, tomando $K = IaI \in \tau$, podemos afirmar que a aplicação,

$$\begin{aligned} f : K &\longrightarrow R \\ \sum_i (x_i a y_i) &\longmapsto \sum_i (x_i b y_i) \end{aligned}$$

está bem definida. Suponha que $\sum x_i a y_i = 0$. Multiplicando à direita por zb, $z \in R$, usando $a(y_i z)b = b(y_i z)a$, e que R é primo, segue que

$$\left(\sum_i x_i b y_i\right) za = 0 \Longrightarrow \sum_i x_i b y_i = 0$$

Por outro lado, f é um aplicação de K para $bI \subseteq R$, além de ser um homomorfismo de R-modulo, e pelo Lema 4.12, existe $\lambda \in C$ tal que $f(y) = \lambda y$ para todo $y \in K$. Logo pela construção de f

tem-se que $f(xay) = xby$, por outro lado, $f(xay) = \lambda xay$, em particular tem-se que

$$xby = \lambda xay \ \ \forall x, y \in I \implies I(b - \lambda a)I = 0$$

Pelo Teorema 4.10 tem-se que $(b - \lambda a) = 0$ e consequentemente $b = \lambda a$, logo a, b são linearmente dependente sobre C como queríamos.

□

O próximo lema trata de uma condição um pouco mais geral em comparação ao lema anterior. Além disso, sua demonstração pode ser reduzida ao caso do Lema 4.17.

Lema 4.18. *Sejam $a_i, b_i \in Q_d(R)$ e $I \in \tau$ tais que*

$$\sum_{i=1}^{n} a_i x b_i = 0 \ \ \textit{para todo } x \in I$$

Se $a_1, \ldots, a_n$ são linearmente independentes sobre C, então cada $b_i = 0$. Da mesma forma, se $b_1, \ldots, b_n$ são linearmente independentes sobre C, então cada $a_i = 0$.

Demonstração. Suponhamos que $a_1, \ldots, a_n$ são linearmente independentes sobre C, devemos mostrar que $b_i = 0$ para todo $i \in \{1, \ldots, n\}$. Por indução, o caso $n = 1$ é exatamente o caso tratado no Teorema 4.10.

Suponha por absurdo que $b_n \neq 0$, pelo Teorema 4.9 parte (2), existe $J \in \tau$ tal que $b_n J \subseteq R$, então $x(b_n y) \in I$ para todo $x \in I$ e $y \in J$, consequentemente tem-se que

$$\sum_{i=1}^{n} a_i (x b_n y) b_i = 0.$$

Como o último termo, $a_n(xb_ny)b_n$, é igual a $-\sum_{i=1}^{n-1}(a_i x b_i) y b_n$, podemos reescrever essa identidade como

$$\sum_{i=1}^{n-1} a_i x (b_n y b_i - b_i y b_n) = 0 \ \text{ para todo } x \in I, y \in J$$

Pela de hipótese de indução, temos que $b_n y b_i - b_i y b_n = 0$ para todo $y \in J$ e $i = 1, \ldots, n-1$. Pelo lema 4.17, b_i e b_n são linearmente dependentes sobre C, logo existem $\lambda_i \in C$ tal que $b_i = \lambda_i b_n$, com $\lambda_n = 1$. Logo

$$\sum_{i=1}^{n} a_i x b_i = \sum_{i=1}^{n} a_i x (\lambda_i b_n) = \sum_{i=1}^{n} (\lambda_i a_i) x b_n$$

Como $b_n \neq 0$ e pelo Teorema 4.10, tem-se que $\sum_{i=1}^{n} \lambda_i a_i = 0 \implies \lambda_i' s = 0$, absurdo já que $\lambda_n \neq 0$. □

O próximo teorema mostra que a dependência linear dos elementos em $Q_r(R)$ pode ser caracterizada por meio de uma identidade envolvendo os polinômios de Capelli.

Teorema 4.19. *Os elementos $a_1, a_2, \ldots, a_m \in Q_r(R)$, $m \geqslant 2$, são linearmente dependentes sobre C, se e somente se, $c_m(a_1, \ldots, a_m, x_1, \ldots, x_{m-1}) = 0$, para todo $x_i \in R$.*

Demonstração.

($\Rightarrow$) Uma vez que c_m é um polinômio alternado multilinear e os elementos $a_1, a_2, \ldots, a_m \in Q_r(R)$, $m \geqslant 2$, são linearmente dependentes, segue diretamente do lema 4.4 que $c_m = 0$.

($\Leftarrow$) A parte "se" será provada por indução em m. O caso $m = 2$ é exatamente o conteúdo do Lema 4.17 (com $I = R$), logo a hipótese de indução nos diz que podemos assumir que o resultado é válido para todo k tal que $2 \leqslant k \leqslant m-1$. Devemos mostrar que o mesmo é valido para m.

Suponha que existem $x_2, \ldots, x_{m-1} \in R$ tais que $c_{m-1}(a_1, \ldots, a_{m-1}, x_2, \ldots, x_{m-1}) \neq 0$. Pelo exemplo 4.5, tem-se que $c_m(a_1, \ldots, a_m, x_1, \ldots, x_{m-1}) = 0$ pode ser escrito como

$$\sum_{i=1}^{m} (-1)^{i-1} a_i x_1 c_{m-1}(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m, x_2, \ldots, x_{m-1}) = 0$$

e, portanto, a dependência linear de $a_1, \ldots, a_m$ segue do Lema 4.18. □

Observação 4.20. Pelo teorema anterior as matrizes $b_1, \cdots, b_s \in Q_r(M_n(\mathbb{K})) = M_n(\mathbb{K})$ são linearmente dependentes, se e somente se, para todo $y_1, \ldots, y_{s-1} \in A$ tem -se

$$c_s(b_1, \cdots, b_s, y_1, \cdots, y_{s-1}) = 0$$

Observe que isso implica que os polinômios $f_1, \cdots, f_s$ são localmente linearmente dependentes, se e somente se

$$c_s(f_1, \ldots, f_s, y_1, \ldots, y_{s-1}) = 0$$

Observação 4.21. Seja $\mathbb{K}$ um corpo infinito e seja $A = M_n(\mathbb{K})$. Suponha que $f_1, \ldots, f_s, g_1, \ldots, g_t \in \mathbb{K}\langle x_1, \ldots, x_m\rangle$ são tais que para cada $a \in A^m$, tem-se que $f_1(a), \ldots, f_s(a)$ ou $g_1(a), \ldots, g_t(a)$ são linearmente dependentes. Afirmamos então que $f_1, \cdots, f_s$ ou $g_1, \ldots, g_t$ são localmente linearmente dependente. Na verdade, nossa suposição pode ser afirmada como

$$c_s(f_1(a), \ldots, f_s(a), y_1, \ldots, y_{s-1}) = 0$$

para todo $y_1, \ldots, y_{s-1} \in A$ ou

$$c_t(g_1(a), \ldots, g_t(a), y_1, \ldots, y_{t-1} = 0)$$

para todo $y_1, \ldots, y_{t-1} \in A$. Isso implica que o produto dos polinômios

$$c_s(f_1, \ldots, f_s, Y_1, \ldots, Y_{s-1}) = 0 \quad \text{e} \quad c_t(g_1, \ldots, g_t, Y_1, \ldots, Y_{t-1}) = 0$$

É uma identidade de A. Estamos, portanto, em posição de aplicar o teorema clássico de Amitsur (Teorema 2.108) . Logo pela observação anterior tem-se que $f_1(a), \ldots, f_s(a)$ ou $g_1(a), \ldots, g_t(a)$ são linearmente dependentes.

Como a álgebra livre $\mathbb{K}\langle X\rangle$ é um domínio, a independência linear de $f_1, \ldots, f_s \in \mathbb{K}\langle X\rangle$ implica na independência linear de $hf_1, \ldots, hf_s$ para todo $h \in \mathbb{K}\langle X\rangle$ diferente de zero. O próximo teorema apresenta resultado semelhante, só que para as álgebras das matrizes $n \times n$.

Teorema 4.22. *Seja $\mathbb{K}$ um corpo infinito, seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$, e seja $h, f_1, \ldots, f_s \in \mathbb{K}\langle x_1, \ldots, x_m\rangle$. Suponha que h não seja uma identidade polinomial de A. Se $f_1, \ldots, f_s$ são localmente linearmente independente, então o mesmo ocorre com $hf_1, \ldots, hf_s$.*

Demonstração. O caso $s = 1$ segue do teorema de Amitsur acima mencionado. Portanto, podemos assumir que $s > 1$ e $hf_1, \ldots, hf_{s-1}$ são localmente linearmente independentes. Suponha por absurdo que $hf_1, \ldots, hf_s$ são localmente linearmente dependentes.

De acordo com a observação acima, tem-se que

$$c_s(hf_1, \ldots, hf_s, Y_1, \ldots, Y_{s-1})$$

é uma identidade polinomial de A, isto é

$$\sum_{i=1}^{s} (-1)^{i-1}\, hf_i Y_1 c_{s-1}(hf_1, \ldots, hf_{i-1}, hf_{i+1}, \ldots, hf_s, Y_2, \ldots, Y_{s-1}) =$$
$$= h \sum_{i=1}^{s} (-1)^{i-1}\, f_i Y_1 c_{s-1}(hf_1, \ldots, hf_{i-1}, hf_{i+1}, \ldots, hf_s, Y_2, \ldots, Y_{s-1})$$

é uma identidade (se s = 2, deve-se entender que $c_1(x) = x$). Como h não é uma identidade, o teorema de Amitsur implica que o último fator,

$$g = \sum_{i=1}^{s} f_i Y_1 \left((-1)^{i-1} c_{s-1}(hf_1, \ldots, hf_{i-1}, hf_{i+1}, \ldots, hf_s, Y_2, \ldots, Y_{s-1})\right)$$

é uma identidade. Seja $\overline{a} \in A^m$ e suponha que

$$h(\overline{a}) f_1(\overline{a}), \ldots, h(\overline{a}) f_{s-1}(\overline{a})$$

são linearmente independentes. Pela Observação 4.20, existem $y_2, \ldots, y_{s-1} \in A$, tais que

$$b_s = (-1)^{s-1} c_{s-1}(h(\overline{a})f_1(\overline{a}), \ldots, h(\overline{a})f_{s-1}(\overline{a}), y_2, \ldots, y_{s-1}) \neq 0$$

Uma vez que o polinômio g é uma identidade polinomial, temos que

$$f_1(\overline{a})y_1b_1 + f_2(\overline{a})y_1b_2 + \ldots + f_{s-1}(\overline{a})y_1b_{s-1} + f_s(\overline{a})y_1b_s = 0$$

para todo $y_1 \in A$ e algum $b_1, \ldots b_{s-1} \in A$. Como $b_s \neq 0$ isso implica que $f_1(\overline{a}), \ldots, f_s(\overline{a})$ são linearmente dependentes (Lema 4.18). Assim, mostramos que para cada $\overline{a} \in A^m$, tem-se que

$$h(\overline{a})f_1(\overline{a}), \ldots h(\overline{a})f_{s-1}(\overline{a}) \quad \text{ou} \quad f_1(\overline{a}), \ldots, f_s(\overline{a})$$

são linearmente dependentes. Consequentemente pela Observação 4.21

$$hf_1, \ldots, hf_{s-1} \quad \text{ou} \quad f_1, \ldots, f_s$$

são localmente linearmente dependentes. No entanto, isso contradiz nossas suposições iniciais.

□

Corolário 4.23. *Seja $\mathbb{K}$ um corpo infinito e seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$. Se $f \in \mathbb{K}\langle X\rangle$ não é uma identidade polinomial de A, então existe um $k \in \mathbb{N}$, com $k \leqslant n$ tal que $1, f, \cdots, f^k$ são localmente linearmente dependente, mas $f, \ldots, f^k$ são localmente linearmente independente. Em particular, $f(A)$ contém uma matriz invertível.*

Demonstração. O Teorema de Cayley-Hamilton nos diz que $1, f, \ldots, f^n$ são localmente linearmente dependentes . Seja $k \leqslant n$ o menor inteiro positivo tal que $1, f, \ldots, f^k$ são localmente linearmente dependente.

Então $1, f, \ldots f^{k-1}$ são localmente linearmente independentes e, pelo teorema anterior, tomando $h = f$ tem-se que $f, \ldots, f_k$ também são. Portanto, existe um $a \in f(A)$ tal que $\lambda_0 1 + \lambda_1 a + \ldots + \lambda_k a^k = 0$ para algum $\lambda_i \in \mathbb{K}$ com $\lambda_0 \neq 0$. Temos que

$$a^{-1} = -\frac{1}{\lambda_0}(\lambda_1 I_d + \lambda_2 a + \ldots \lambda_n a^{n-1})$$

Portanto, a matriz a é invertível. □

Se f é um polinômio central para $A = M_n(\mathbb{K})$, então $f(A)$ consiste apenas em matrizes invertíveis e (possivelmente) 0. O exemplo a seguir (veja [30]) mostra que isso também pode ser válido para polinômios não centrais.

Exemplo 4.24. Seja $f = [X_1, X_2]^3$ e $A = M_2(\mathbb{K})$. Como é bem conhecido pelo Exemplo 2.92, $[X_1, X_2]^2$ é um polinômio central para $M_2(\mathbb{K})$, na verdade, $[a,b]^2 = -det([a,b])I_d$ para todo $a, b \in M_2(\mathbb{K})$, (onde I_d é matriz identidade 2×2). Chame $\lambda = det([a,b])$

Portanto, $f(a,b) = -\lambda[a,b]$, e consequentemente, ou $f(a,b) = 0$ ou $f(a,b)$ é invertível.

Nosso próximo objetivo é provar um teorema relativo à multiplicidade algébrica de autovalores de matrizes em $f(A)$. Mas primeiro, algumas preliminares. Chame ad_α a aplicação linear definida por $ad_\alpha(x) = [\alpha, x]$ (também conhecida como função adjunta). Note que

$$ad_\alpha^k(x) = \underbrace{(ad_\alpha \circ ad_\alpha \cdots ad_\alpha)}_{\text{k vezes}}(x) = \sum_{i=0}^{k}(-1)^i\binom{k}{i}\alpha^{k-i}x\alpha^i$$

Fazemos a prova por indução. É fácil ver que para $k = 1$ o resultado é válido. Suponha que o resultado é verdadeiro para todo $n \leqslant k$, devemos mostrar que vale para $k+1$.

$$\begin{aligned} ad_\alpha^{k+1}(x) &= ad_\alpha^k(\alpha x - x\alpha) \\ &= \sum_{i=0}^{k}(-1)^i\binom{k}{i}\alpha^{k-i}(\alpha x - x\alpha)\alpha^i \\ &= \sum_{i=0}^{k}(-1)^i\left(\binom{k}{i-1}\alpha^{k+1-i}x\alpha^i + \binom{k}{i}\alpha^{k+1-i}x\alpha^i\alpha^{i+1}\right) \\ &= \sum_{i=0}^{k}(-1)^i\binom{k+1}{i}\alpha^{k+1-i}x\alpha^i \end{aligned}$$

Como queríamos demonstrar. Isso mostra que $u^s = 0$ implica $ad_u^{2s-1} = 0$.

Lema 4.25. *Sejam $\mathbb{K}$ um corpo qualquer, $A = M_n(\mathbb{K})$ com $n \geqslant 2$, $\frac{n}{2} < s \leqslant n$, e $u \in M_s(\mathbb{K})$ uma matriz nilpotente, e seja $a \in A$ uma matriz da forma*

$$\begin{bmatrix} u & 0 \\ 0 & \star \end{bmatrix}.$$

onde, 0 e $\star \in M_{n-s}(\mathbb{K})$. Então, para cada $k \geqslant 2s-1$, $ad_a^k(A)$ não contém matrizes invertíveis.

Demonstração. Seja

$$x = \begin{bmatrix} x' & \star \\ \star & \star \end{bmatrix} \in A.$$

onde $x' \in M_s(\mathbb{K})$. Um cálculo fácil mostra que, para qualquer $r \geqslant 1$

$$ad_a^r(x) = \begin{bmatrix} ad_u^r(x') & \star \\ \star & \star \end{bmatrix}$$

Conforme apontado acima, $u^s = 0$ implica $ad_u^{2s-1} = 0$. Portanto,

$$ad_a^{2s-1}(x) = \begin{bmatrix} 0 & \star \\ \star & \star \end{bmatrix}$$

Isso implica que

$$ad_a^k(x) = \begin{bmatrix} 0 & \star \\ \star & \star \end{bmatrix}$$

para cada $k \geqslant 2s-1$. A matriz do canto superior direito tem tamanho $s \times (n-s)$. Como $s > n-s$, suas linhas são linearmente dependentes. Portanto, $ad_a^k(x)$ não é invertível em A. □

Observação 4.26. Em [14], Herstein afirma que se $char(\mathbb{K}) = 0$, então a observação de que $u^s = 0$ implica em $ad_u^{2s-1} = 0$, tem um inverso. Ou seja, se $b \in A = M_n(\mathbb{K})$ e $d \geqslant 1$ são tais que $ad_b^d = 0$ em A, então existe uma matriz escalar α tal que $u = b - \alpha$ satisfaz $u^{[(d+1)/2]} = 0$. Observe que isso em particular mostra que $ad_b^{2s} = 0$ implica $ad_b^{2s-1} = 0$.

Teorema 4.27. *Seja $\mathbb{K}$ um corpo com $char(\mathbb{K}) = 0$ e seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$. Se $f = f(x_1, \ldots, x_m) \in \mathbb{K}\langle X \rangle$ não é uma identidade polinomial nem um polinômio central de A, então $f(A)$ contém uma matriz tal que a multiplicidade algébrica de qualquer um de seus autovalores não exceda $n/2$.*

Demonstração. Suponha que o teorema seja falso. Então, para qualquer $b \in f(A)$ existe um autovalor $\lambda \in \overline{\mathbb{K}}$, o fecho algébrico de $\mathbb{K}$, cuja multiplicidade algébrica é $s > n/2$. Seja $p \in M_n(\mathbb{K})$ uma matriz invertível tal que pbp^{-1} é da forma

$$\begin{bmatrix} \lambda + u & 0 \\ 0 & \star \end{bmatrix}$$

onde $u \in M_s(\mathbb{K})$ é uma matriz nilpotente. Pelo lema anterior tem-se que

$$ad_{pbp^{-1}}^{2n-1}\left(M_n(\overline{\mathbb{K}})\right) = ad_{pbp^{-1}-\lambda}^{2n-1}\left(M_n(\overline{\mathbb{K}})\right)$$

não contém matrizes invertíveis. Portanto $ad_b^{2n-1}(A)$ também não contém matrizes invertíveis. Ou seja, nenhuma das matrizes na imagem do polinômio

$$ad_f^{2n-1}(X_{m+1}) = [f, [\ldots [f, [f, X_{m+1}]] \ldots]$$

é invertível em A. Portanto, $ad_f^{2n-1}(X_{m+1})$ é uma identidade de A pelo Corolário 4.23.

Seja d o menor inteiro positivo tal que $ad_f^d(X_{m+1})$ é uma identidade polinomial de A. Ou seja, $ad_b^d = 0$ em A para cada $b \in f(A)$. Estamos agora em posição de usar o resultado da Observação 4.26 acima mencionado. Para qualquer $b \in f(A)$ existe uma matriz escalar α tal que $v = b - \alpha$ satisfaz $v^{(d+1)/2} = 0$. Isso implica que

$$ad_b^{d-1}(x) = ad_v^{d-1}(x) = (-1)^{(d-1)/2}\binom{d-1}{\frac{d-1}{2}} v^{(d-1)/2} x v^{(d-1)/2}$$

para cada $x \in A$. Portanto $ad_b^{d-1}(x)$ não é invertível. Em outras palavras, nenhuma das matrizes na imagem do polinômio $ad_f^{d-1}(X_{m+1})$ é invertível. Consequentemente, o corolário 4.23 nos diz que $ad_f^{d-1}(X_{m+1})$ é uma identidade, o que contradiz a escolha de d. □

Observação 4.28. Um polinômio é dito 2-central para $M_n(F)$ se f^2 for central, mas f não é. O exemplo mais simples é $[X_1, X_2]$ que é 2-central para $M_2(F)$. Acontece que polinômios 2-centrais para $M_n(F)$ existem para vários n pares. Obviamente, qualquer matriz na imagem de tal polinômio tem no máximo dois autovalores. Por outro lado, a partir do teorema 4.27, inferimos que a imagem sempre contém uma matriz com dois autovalores de multiplicidade algébrica exatamente $\frac{n}{2}$.

4.2 Um lema sobre conjuntos invariantes sob conjugação

O primeiro objetivo desta secção é provar que para qualquer comutador nas álgebras das matrizes pode ser escrito como a soma de 22 elementos de quadrado zero. Começamos, no entanto, com uma consideração puramente algébrica de álgebras de matriz $M_n(B)$.

Os três próximos resultados, foram extraídos de [2].

Lema 4.29. *Seja B uma álgebra unitária. Então, cada elemento em $M_2(B)$ pode ser escrito como $e_1e_2 + e_3e_4 - e_5 - e_6$ para alguns idempotentes $e_i \in M_2(B) (i = 1, \ldots, 6)$.*

Demonstração. Temos que

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & a_{11} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ a_{22} & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -a_{12} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ -a_{21} & 1 \end{bmatrix}$$

e todas as matrizes que aparecem no lado direito são idempotentes. □

Proposição 4.30. *Seja B uma álgebra unitária e $n \geqslant 2$. Então, cada elemento em $M_n(B)$ pode ser escrito como $e_1e_2+e_3e_4+e_5e_6+e_7-e_8-e_9-e_{10}-e_{11}$ para alguns idempotentes $e_i \in M_n(B)$ e $i = 1,\dots 11$.*

Demonstração. Se n é par, então $n = 2k$ para algum k inteiro, e como $M_n(B) \cong M_2(M_k(B))$ basta aplicar o lema anterior, para chegar no resultado desejado. Suponhamos agora que n é ímpar, então é da forma $n = 2k+1$ para algum $k \geqslant 1$. Seja $a = (a_{ij}) \in M_n(B)$ e defina

$$a' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}; \quad a'' = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{bmatrix}$$

Note que

$$a' = \begin{bmatrix} 1 & a_{11} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Além disso, temos ainda que

$$a'' = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

e todas as matrizes do lado direito dessas duas identidades são idempotentes. Uma vez que $a - a' - a''$ está em uma subálgebra de $M_n(B)$ isomorfa a $M_2(M_k(B))$, satisfaz a conclusão do lema anterior. Assim, $a = (a - a' - a'') + a' + a''$ pode ser escrito na forma desejada. □

Lema 4.31. *Seja B uma álgebra unitária e $n \geqslant 2$. Então, cada comutador em $M_n(B)$ pode ser escrito como a soma de 22 elementos de quadrado zero.*

Demonstração. Se e é um elemento idempotente e x é um elemento arbitrário, então o comutador $[e,x]$ é a soma de dois elementos de quadrados zero, a saber

$$[e,x] = ex(1-e) + (e-1)xe$$

Se e' for outro elemento idempotente, então $[ee,x]$ pode ser escrito como a soma de quatro elementos quadrados-zero:

$$[ee',x] = ee'x(1-e) + (e-1)e'xe + e'xe(1-e') + (e'-1)xee'.$$

Tome um comutador qualquer $[y,x] \in M_n(B)$, pela proposição anterior, $y \in M_n(B)$ pode ser escrito como $y = e_1e_2 + e_3e_4 + e_5e_6 + e_7 - e_8 - e_9 - e_{10} - e_{11}$ onde $e_i's$ são matrizes idempotentes. Logo

$$[y,x] = [e_1e_2,x] + [e_3e_4,x] + [e_5e_6,x] + [e_7,x] - [e_8,x] - [e_9,x] - [e_{10},x] - [e_{11},x]$$

Pela duas equações anteriores tem-se que $[y,x]$ é escrito como soma de 22 elementos de quadrado zero. □

O próximo lema revela a ideia principal na qual esta secção se baseia. O resultado será utilizado para $T = f(A)$, mas enunciamos para um subconjunto qualquer T invariante por conjugação, pois o resultado pode ser de interesse independente

Lema 4.32. *Seja $\mathbb{K}$ um corpo com $char(\mathbb{K}) \neq 2$, seja B uma álgebra unitária e seja $A = M_n(B)$ com $n \geqslant 2$. Se um subconjunto T de A for invariante sob conjugação por elementos invertíveis em A, então:*

1. *Qualquer elemento da forma $[t,u]$, onde $t \in T$ e u é um elemento tal que $u^2 = 0$ em A, encontra-se em $T - T = \{t - t'$, tal que $t,t' \in T\}$.*
2. *Qualquer elemento da forma $[t,[x,y]]$, com $t \in T$ e $x,y \in A$, é uma soma de 22 elementos de $T - T$.*
3. *Se T contiver elementos t_1,t_2, de modo que $[t_1,t_2]$ é invertível em A e, para algum $k \geqslant 1$, cada elemento em A é uma soma de k comutadores e um elemento central, então cada comutador em A é uma soma de $1936k^2 + 22k$ elementos de $T - T$.*
4. *Se $B = K$ é um corpo algebricamente fechado e T contém uma matriz t tal que a multiplicidade algébrica de qualquer um de seus autovalores não exceda $n/2$, então todos os elementos de quadrados igual zero em A pertence em $T - T$.*

Demonstração.

1. Como $u^2 = 0$, temos que $(1 - \frac{u}{2})^{-1} = 1 + \frac{u}{2}$, portanto

$$[t,u] = \left(1 - \frac{u}{2}\right) t \left(1 - \frac{u}{2}\right)^{-1} - \left(1 - \frac{u}{2}\right)^{-1} t \left(1 - \frac{u}{2}\right)$$

   O que mostra que $[t,u] \in T - T$.

2. Isso segue de (a) e do lema 4.31, que afirma que todo comutador em A é uma soma de 22 elementos de quadrados zero.

3. A prova é baseada na identidade

$$[w,z] = [[t_2, w[t_1,t_2]^{-1}], t_1 z] - [[t_2, w[t_1,t_2]^{-1} t_1], z] + [t_1, z[t_2, w[t_1,t_2]^{-1}]]$$

   a qual pode ser verificada por um cálculo direto. Esta identidade mostra que tomando $x_1 = w[t_1,t_2]^{-1}$, $x_2 = t_1 z$, $x_3 = w[t_1,t_2]^{-1} t_1$, $x_4 = z$ e $x_5 = z[t_2, w[t_1,t_2]^{-1}]]$ o comutador $[w,z]$ de quaisquer dois elementos w e z de A pode ser escrito como

$$[[t_2,x_1],x_2] + [[t_2,x_3],x_4] + [t_1,x_5]$$

   para alguns $t_i \in T$ e $x_i \in A$. Usando o item anterior (parte 2), junto com nossa suposição de que cada elemento em A é uma soma de k comutadores e um elemento central, vemos que cada comutador $[t,x]$, com $t \in T$ e $x \in A$, é uma soma de 22k elementos de $T - T$. Consequentemente, cada comutador $[w,z]$ é uma soma de

$$2 \cdot (22k)^2 + 2(22k)^2 + 22k = 1936k^2 + 22k$$

   elementos de $T - T$.

4. Sejam $\lambda_1, \ldots, \lambda_r$, $r \geqslant 2$, autovalores distintos de t. Como T é invariante sobre conjugação, assumimos que

$$t = \begin{bmatrix} t_{\lambda_1} & 0 & \cdots & 0 \\ 0 & t_{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{\lambda_r} \end{bmatrix}$$

   onde t_{λ_i} é uma matriz triangular superior tendo λ_i como diagonal. Por hipótese, o tamanho de qualquer t_{λ_i} não excede $n/2$. Sem perda de generalidade, podemos supor que t_{λ_1} e t_{λ_r} têm tamanho maior ou igual a $t_{\lambda_2}, \ldots, t_{\lambda_{r-1}}$.

Seja

$$u = \begin{bmatrix} 0 & d \\ 0 & 0 \end{bmatrix}$$

onde a matriz do canto superior esquerdo (resp. inferior direito) tem tamanho $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$ (resp. $\lfloor (n+1)/2 \rfloor \times \lfloor (n+1)/2] \rfloor$) e d é $\lfloor n/2 \rfloor \times \lfloor (n+1)/2 \rfloor$ matriz com entradas arbitrárias d_i na diagonal principal e zeros em outro lugar (a última coluna de d é, portanto, zero se n for ímpar). Escreva

$$t = \begin{bmatrix} t_1 & * \\ 0 & t_2 \end{bmatrix}$$

onde t_1 é de tamanho $[n/2] \times [n/2]$ e t_2 é de tamanho $[(n+1)/2] \times [(n+1)/2]$. Especificamente,

$$t_1 = \begin{bmatrix} t_{\lambda_1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & t_{\lambda_{j-1}} & 0 \\ 0 & \cdots & 0 & t'_{\lambda_j} \end{bmatrix}; \; t_2 = \begin{bmatrix} t''_{\lambda_j} & \cdots & 0 & 0 \\ 0 & t_{\lambda_{j+1}} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & t_{\lambda_r} \end{bmatrix}$$

onde

$$t_{\lambda_j} = \begin{bmatrix} t'_{\lambda_j} & * \\ 0 & t''_{\lambda_j} \end{bmatrix}$$

(Note que t''_{λ_j} pode ser zero; de fato, se $r = 2$, então t''_{λ_j} é zero, já que t_{λ_r} tem tamanho máximo de $[n/2]$). Note ainda que

$$[t,u] = \begin{bmatrix} 0 & t_1 d - d t_2 \\ 0 & 0 \end{bmatrix}$$

e $t_1 d - d t_2$ é uma matriz de tamanho $[n/2] \times [(n+1)/2]$ com zeros abaixo da diagonal principal e entradas da diagonal principal iguais ao produto de d_i e a diferença de dois autovalores distintos de t. Outra propriedade de que precisamos é que $d_i = 0, i = k, \ldots, [n/2]$, implica que a i-ésima linha de $t_1 d - d t_2, i = k, \cdots [n/2]$ é zero.

Tudo isso mostra que podemos escolher d de forma que $t_1 d - d t_2$ tenha qualquer rank entre 0 e $[n/2]$. Obviamente, $[t,u]$ é uma matriz de quadrado zero de mesmo rank que

a matriz $t_1 d - d t_2$. Observe que $u^2 = 0$, portanto $[t,u]$ reside em $T - T$ pelo item (a) deste lema. Uma vez que qualquer matriz quadrada zero em A tem rank no máximo $[n/2]$ e duas matrizes de quadrado zero têm o mesma rank se e somente se forem semelhantes (observação 2.27), a conclusão desejada de que $T - T$ contém todas as matrizes de quadrado zero, segue do fato de que esta conjunto é invariante sob conjugação por matrizes invertíveis.

□

Observação 4.33. O número 22 no item (2) do lema anterior pode não ser o menor possível. Em particular, se $B = \mathbb{K}$, então ele pode ser substituído por 4, como veremos na próxima secção. Consequentemente, o número $1936k^2 + 22k$ no item (3) pode ser, neste caso, substituído por $2(4)^2 + 2(4)^2 + 4 = 68$, desde que $char(K) = 0$ ou não divida n. Isso ocorre porque toda matriz com traço zero pode ser escrita como um comutador [26, 3], portanto, podemos tomar $k = 1$ sob essas suposições.

4.3 Teoremas principais e seus corolários

Vamos agora considerar a situação em que $T = f(A)$. Como acima, nós escrevemos

$$f(A) - f(A) = \{t - t' \text{ tal que } t, t' \in f(A)\}.$$

Combinando alguns resultados das secções anteriores, podemos agora provar nosso primeiro resultado principal.

Teorema 4.34. *Seja $\mathbb{K}$ um corpo infinito com $char(\mathbb{K}) \neq 2$, seja $n \geqslant 2$, seja $f = f(X_1, \ldots, X_m) \in \mathbb{K}\langle X\rangle$ um polinômio que não é uma identidade polinomial nem é um polinomial central de $M_n(\mathbb{K})$, seja $k \geqslant 1$, e B uma $\mathbb{K}$-álgebra com unidade, tal que cada elemento em $A = M_n(B)$ é uma soma de k comutadores e um elemento central. Então, todo comutador em A é uma soma de $1936k^2 + 22k$ elementos de $f(A) - f(A)$.*

Demonstração. Pelo Lema 3.21 parte (2), o polinômio $\hat{f}$ não é uma identidade de $M_n(\mathbb{K})$. Portanto, o Corolário 4.23 nos diz que $\hat{f}(M_n(\mathbb{K}))$ contém uma matriz invertível. Ou seja, existem $t_1, t_2 \in f(M_n(\mathbb{K}))$ de forma que $[t_1, t_2]$ é invertível em $M_n(\mathbb{K})$.

Como $M_n(\mathbb{K})$ é uma subálgebra (unitária) de A, podemos considerar t_1 e t_2 como elementos de $f(A)$ cujo comutador $[t_1, t_2]$ é invertível em A. Finalmente, como $f(A)$ é invariante sob conjugação, o teorema segue do Lema 4.32 item (3) aplicando a $T = f(A)$. □

Iremos abordar agora um pequeno resultado extraído de [8] sobre anéis (ou álgebras) de endomorfismo de infinitas somas diretas e produtos de cópias de um módulo fixo .

Proposição 4.35. *Seja R um anel, N um R-módulo à direita, Ω um conjunto infinito e $M = \oplus_\Omega N$ ou $M = \prod_\Omega N$. Então $End_R(M) = [x, End_R(M)]$ para algum $x \in End_R(M)$. Se $\Omega = \mathbb{N}$, então x pode ser considerado o operador de deslocamento.*

O próximo teorema, apresenta um resultado sobre comutadores em anéis de matriz. E será útil para os próximos corolários.

Teorema 4.36. *Seja R um anel com unidade e n um inteiro positivo. Então existem matrizes $X, Y \in M_n(R)$ de modo que para toda matriz $A \in M_n(R)$ de traço 0, tem-se que $A \in [X, M_n(R)] + [Y, M_n(R)]$. Especificando, escrevendo e_{ij} para as matrizes unitárias, pode-se tomar*

$$X = \sum_{i=1}^{n-1} e_{i+1,i} \quad e \quad Y = e_{nn}$$

Demonstração. Seja $A = (a_{ij})$ e $X = \sum_{i=1}^{n-1} e_{i+1,i}, Z = \sum_{i=1}^{n-1} e_{i,i+1}$, temos que

$$ZX = e_{11} + e_{22} + \ldots + e_{n-1n-1}$$

Consequentemente

$$ZX = I - e_{nn}. \tag{4.1}$$

Também para $l \in \{0, 1, \ldots, n-1\}$ tem-se que

$$e_{nn}X^l AZ^l e_{nn} = e_{nn}\left(\sum_{i=1}^{n-1} E_{i,i+l}\right) e_{nn}e_{nn-1}Ae_{n-ln} = a_{n-ln-l}e_{nn}$$

Portanto

$$e_{nn}X^l AZ^l e_{nn} = a_{n-ln-l}e_{nn} \tag{4.2}$$

Agora tome

$$C = A + XAZ + \ldots + X^{n-1}AZ^{n-1}$$

Então pela equação (4.1), tem-se que

$$[CZ, X] = CZX - XCZ = C(I - e_{nn}) - XCZ$$

Além disso, como $X^n = 0$ tem-se que $C - XCZ = A$, consequentemente $[CZ, X] = A - Ce_{nn}$. Notamos que Ce_{nn} é uma combinação linear de $e_{1n}, e_{2n}, \ldots, e_{n-1,n}$. De fato, pela equação 4.2 e

uma vez que o traço de A é igual a zero, tem-se que

$$e_{nn}Ce_{nn} = (a_{nn} + a_{n-1,n-1} + \ldots + a_{11})e_{nn} = 0$$

Definindo $Y = e_{nn}$, temos para cada $i \in \{1, 2, \ldots, n-1\}$, $e_{in} = e_{in}Y - Ye_{in}$. Portanto, $Ce_{nn} = [Ce_{nn}, Y]$ e, portanto, $A = [CZ, X] + [Ce_{nn}, Y]$. □

Corolário 4.37. *Seja $\mathbb{K}$ o corpo de característica 0, seja $f \in \mathbb{K}\langle X\rangle$ um polinômio que não é uma identidade nem um polinômio central de $M_n(\mathbb{K})$, e seja C uma álgebra comutativa unitária. Então, todo comutador em $A = M_n(C)$ é uma soma de 7788 elementos de $f(A) - f(A)$.*

Demonstração. Como $\mathbb{K}$ tem a característica 0, podemos escrever todo $x \in A$ como a soma da matriz de traço zero $x - \frac{tr(x)}{n}Id$ e a matriz $\frac{tr(x)}{n}Id$ que está no centro de A. Pelo Teorema 4.36, todo matriz de traço zero é a soma de dois comutadores. Portanto, as condições do Teorema 4.34 são atendidas para $k = 2$. □

Observação 4.38. Para muitas álgebras comutativas, toda matriz sem traço em $M_n(C)$ é na verdade um comutador. Então tomando $k = 1$ o caso do Teorema 4.34 se aplica, de modo que o número 7788 pode ser substituído por 1958. Se C = F, então vemos na observação 4.33 que esse número pode ser reduzido ainda mais para 68.

Corolário 4.39. *Seja $\mathbb{K}$ um corpo infinito com $char(\mathbb{K}) \neq 2$, seja V um espaço vetorial de dimensão infinita sobre $\mathbb{K}$, e seja $f \in \mathbb{K}\langle X\rangle$ um polinômio não constante. Então, cada elemento em $A = End_F(V)$ é uma soma de 1958 elementos de $f(A) - f(A)$.*

Demonstração. Escolha um $n \geqslant 2$ tal que f não é uma identidade polinomial nem um polinômio central de $M_n(\mathbb{K})$ (ver Exemplo 2.87). Como V é isomorfo a V^n, a soma direta de n copias de V, $End_F(V)$ é isomorfo a $M_n(End_F(V))$. Pela Proposição 4.35, cada elemento em $End(_F(V))$ é um comutador , logo podemos aplicar novamente o Teorema 4.34 para $k = 1$. □

Para o tipo de problema tratado aqui, a álgebra $End_F(V)$ com V de dimensão infinita parece ser mais fácil de lidar do que a álgebra matricial $M_n(\mathbb{K})$. Na verdade, todos os seus elementos são comutadores e não tem identidades polinomiais, portanto, não há necessidade de distinguir entre diferentes polinômios. Pode-se então perguntar se $f(End_F(V))$ é realmente igual a $End_F(V)$ para qualquer polinômio não constante f. O próximo exemplo mostra que isso não é verdade.

Exemplo 4.40. Seja V um espaço vetorial de dimensão infinita enumerável sobre um corpo $\mathbb{K}$, e seja $A = End_F(V)$. Tome qualquer base $\{e_1, e_2, \ldots\}$ de V. Seja $l \in A$ um operador de deslocamento à esquerda, isto é, definido por $l(e_1) = 0$ e $l(e_n) = e_{n-1}$ para cada $n > 1$. Afirmamos

que l não é igual ao quadrado de um elemento em A. Suponha que isso não seja verdade. Seja $h \in A$ tal que $l = h^2$. Escreva

$$h(e_1) = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_s e_s$$

onde $\lambda_i \in \mathbb{K}$. Como $l = h^2$ comuta com h, então tem-se que $(l \circ h)(e_1) = h(l(e_1)) = 0$, consequentemente

$$(l \circ h)(e_1) = \lambda_2 e_1 + \ldots + \lambda_s e_{s_1} = 0$$

Como temos uma combinação linear dos elementos da base, segue que $\lambda_2 = \cdots = \lambda_s = 0$ e segue que $h(e_1) = \lambda_1 e_1$. No entanto, como $h^2(e_1) = l(e_1) = 0$, isso só possível quando $h(e_1) = 0$. Continuamos examinando $h(e_2)$.Sejam $\mu_1, \ldots, \mu_n \in \mathbb{K}$ tais que

$$h(e_2) = \mu_1 e_1 + \mu_2 e_2 + \ldots + \mu_d e_d$$

Como

$$l(h(e_2)) = h(l(e_2)) = h(e_1) = 0$$

Temos que

$$\mu_2 e_1 + \ldots + \mu_d e_{d-1} = 0$$

Portando $\mu_2, \ldots, \mu_d = 0$ e por consequência $h(e_2) = \mu_1 e_1$. Mas então

$$e_1 = l(e_2) = h(h(e_2)) = \mu_1 h(e_1) = 0$$

Uma contradição.

Mostramos assim que l não está na imagem do polinômio $f(x) = x^2$. Em particular, $f(End_{\mathbb{K}}(V)) \neq End_{\mathbb{K}}(V)$. Observamos que isso ainda é verdade se V é de dimensão finito. De fato, uma matriz nilpotente $a \in M_n(\mathbb{K})$ de nil índice máximo não pode ser escrita como b^2 para algum $b \in M_n(F)$. Para provar isso, observe que $a = b^2$ implica $b^{2n} = a^n = 0$ e, portanto, $b^n = 0$, o que leva à contradição de que $a^{n-1} = b^{2n-2} = 0$.

Teorema 4.41. *Seja $\mathbb{K}$ um corpo algebricamente fechado com $char(\mathbb{K}) = 0$, seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$, e seja $f \in \mathbb{K}\langle X \rangle$ um polinômio que não é uma identidade nem um polinômio central de A. Então $f(A) - f(A)$ contém todas as matrizes de quadrado zero em A.*

Demonstração. Observe que pelo Teorema 4.27, $f(A)$ contém uma matriz tal que a multiplicidade algébrica de qualquer um de seus autovalores não exceda $n/2$. Por outro lado, pelo Lema 4.32, parte 4 segue que todos os elementos de quadrados igual zero estão $f(A) - f(A)$. □

Proposição 4.42. *Toda matriz nilpotente pode ser escrito como uma soma de duas matrizes de quadrado zero.*

Demonstração. De fato, seja Y uma matriz nilpotente, note que Y tem autovalor nulo e seu polinômio característico é da forma $p_T(x) = x^n$. Seja J a matriz de Jordan de Y, tem-se que

$$\begin{aligned} J &= \sum_{i=1}^{n-1} \alpha_i e_{i,i+1}, \ \ \alpha \in \{0,1\} \\ &= \sum_{i=1}^{n-1} \alpha_{2i} e_{2i,2i+1} + \sum_{i=1}^{n-1} \alpha_{2i-1} e_{2i-1,2i} \end{aligned}$$

Logo J é escrito como a soma de duas matrizes de quadrado zero, e consequentemente Y é escrito como uma soma de duas matrizes de quadrado zero.

□

Teorema 4.43. *Seja F um corpo com $char(F) = 0$, e $A \in M_n(F)$ uma matriz quadrada com traço zero. Então, A é a soma de quatro matrizes de quadrado zero.*

Demonstração. Seja A uma matriz de traço zero de $M_n(F)$. O caso quando $A = 0$ é óbvio e nós o descartamos de agora em diante.

Suponha primeiro que A seja uma matriz escalar. Sem perda de generalidade, podemos assumir que $A = I_n$. Como $tr(A) = 0$, obtemos que F tem característica positiva, o que é absurdo.

Logo o único caso a se considerar é quando A não é uma matriz escalar. Note que pela Proposição 3.4, A é semelhante a uma matriz com apenas zeros na diagonal principal. Seja $A = PBP^{-1}$, tal que B é a matriz com apenas zeros na diagonal principal.

Dividindo B na soma de sua parte triangular estritamente superior U e sua parte triangular estritamente inferior L, temos que $A = PUP^{-1} + PLP^{-1}$. Portanto se decompõe na soma de duas matrizes nilpotentes.

Pela proposição anterior, cada matriz nilpotente se escreve como uma soma de duas matrizes de quadrado zero. Logo A é escrito como soma de quatro matrizes de quadrado zero.

□

Corolário 4.44. *Seja $\mathbb{K}$ um corpo algebricamente fechado com $char(\mathbb{K}) = 0$, seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$, e seja $f \in \mathbb{K}\langle X \rangle$ um polinômio que não é uma identidade nem um polinômio central de A. Então toda matriz de traço zero em A é uma soma de quatro matrizes de $f(A) - f(A)$*

Demonstração. Basta aplicar o Teorema 4.43 que afirma que toda matriz de traço zero é uma soma de quatro matrizes quadradas zero, e em seguida o Teorema 4.41 que afirma que todas as matrizes de quadrado zero estão contidas em $f(A) - f(A)$. □

Dizemos que os polinômios $f, g \in \mathbb{K}\langle X\rangle$ são ciclicamente equivalentes se $f - g$ é uma soma de comutadores em $\mathbb{K}\langle X\rangle$

Corolário 4.45. *Seja $\mathbb{K}$ um corpo algebricamente fechado com $char(\mathbb{K}) = 0$, seja $A = M_n(\mathbb{K})$ com $n \geqslant 2$, e seja $f \in \mathbb{K}\langle X\rangle$ um polinômio que não é ciclicamente equivalente a uma identidade de A e não é um polinômio central de A. Então, cada matriz em A é uma combinação linear de nove matrizes de $f(A)$.*

Demonstração. Por [7, Corolário 4.7], existe um $a \in f(A)$ cujo traço não é 0. Escreva $x \in A$ como

$$x = \frac{tr(x)}{tr(a)}a + \left(x - \frac{tr(x)}{tr(a)}a\right)$$

e aplique o Corolário 4.44 à matriz sem traço $x - \frac{tr(x)}{tr(a)}a$.

□

Este corolário também é válido para corpos que não são fechados algebricamente, mas temos que substituir o número 9 por $2 \cdot 68 + 1 = 137$ (ver Observação 4.38).

Exemplo 4.46. Seja $char(\mathbb{K}) = 0$. Se $f = [X_1, X_2] + \frac{1}{n}$, então toda matriz em $f(M_n(\mathbb{K}))$ tem traço igual a 1. Consequentemente, apenas matrizes cujo traço é um múltiplo inteiro de 1 pertencem ao subgrupo aditivo gerado por $f(M_n(\mathbb{K}))$. Isso mostra que, ao contrário de outros resultados desta seção, o envolvimento de combinações lineares com coeficientes no corpo $\mathbb{K}$ é necessário no corolário 4.45. Além disso, também mostra que os principais teoremas e seus corolários devem envolver diferenças (ao invés de apenas somas) de elementos da imagem de um polinômio.

Na verdade, uma soma de matrizes de $f(M_n(\mathbb{K}))$ nunca é uma matriz sem traço. Pode ser mais esclarecedor dar um exemplo de um polinômio com termo constante zero: se $g = [X_1, X_2]^4$, então uma soma de matrizes de $g(M_2(\mathbb{R}))$ nunca é uma matriz com traço negativo. De fato, como $[x_1, x_2]^2$ é central (Exemplo 2.10), tem-se que $[x_1, x_2]^2 = \lambda I_d$, logo $[x_1, x_2]^4 = \lambda^2 I_d$.

É conhecido que um polinômio central de $M_n(\mathbb{K})$ é uma identidade de $M_k(\mathbb{K})$ para todo $k < n$. Portanto, se um polinômio não é uma identidade nem um polinômio central de $M_2(\mathbb{K})$,

então o mesmo vale para $M_n(\mathbb{K})$ para todo $n \geqslant 2$. O próximo resultado, portanto, segue facilmente do corolário 4.44, então o enunciamos sem prova.

Corolário 4.47. *Seja $\mathbb{K}$ um corpo algebricamente fechado de característica 0 e seja $f \in \mathbb{K}\langle X\rangle$ um polinômio que não é uma identidade nem um polinômio central de $M_2(\mathbb{K})$. Se*

$$A = \prod_{n=2}^{\infty} M_n(\mathbb{K})$$

(ou seja, A é o produto direto de todas as álgebras de matriz $M_n(\mathbb{K})$ com $n \geqslant 2$), então cada comutador em A é uma soma de quatro elementos de $f(A) - f(A)$. Em particular, $[A, A] \subseteq span f(A)$.

Os resultados acima mostram basicamente que para a álgebra $A = M_m(\mathbb{K})$, existe um inteiro positivo n tal que todo comutador (ou mesmo todo elemento) em A pode ser expresso por n elementos da imagem de qualquer polinômio f que satisfaça as restrições necessárias. O ponto principal é que n é independente de f bem como do tamanho das matrizes. No entanto, qual é o n mínimo? Mais precisamente, pode-se fazer a seguinte pergunta.

Questão 4.48. *Qual é o menor número que pode substituir o número que aparece nos Corolários 4.37, 4.39, 4.44 e 4.45 respectivamente (ou seja, 7788, 1958, 4 e 9, respectivamente)?*

Nesta secção dedicamos essencialmente à prova da existência deste número, mas a sua determinação fica como um problema em aberto. Os números de quatro dígitos acima foram obtidos por um método um tanto grosseiro, segundo Bresar em [9], esse número pode ser substancialmente reduzidos. O método que deu origem aos números 4 e 9 era mais sofisticado, mas mesmo assim é difícil acreditar que sejam os menores possíveis.

REFERÊNCIAS

[1] Dniester notebook: unsolved problems in the theory of rings and modules. In *Non-associative algebra and its applications*, M. V. Kochetov, V. T. Filippov, V. K. Kharchenko, and I. P. Shestakov, Eds., vol. 246 of *Lect. Notes Pure Appl. Math.* Chapman & Hall/CRC, Boca Raton, FL, 2006, pp. 461–516. Translated from the 1993 Russian edition [MR1310114] by Murray R. Bremner and Mikhail V. Kochetov and edited by V. T. Filippov, V. K. Kharchenko and I. P. Shestakov.

[2] ALAMINOS, J., EXTREMERA, J., VILLENA, A., BREŠAR, M., AND ŠPENKO, Š. Commutators and square-zero elements in banach algebras. *The Quarterly Journal of Mathematics 67*, 1 (2016), 1–13.

[3] ALBERT, A. A., AND MUCKENHOUPT, B. On matrices of trace zeros. *Michigan Math. J. 4* (1957), 1–3.

[4] AMITSUR, S. A. The T-ideals of the free ring. *J. London Math. Soc. 30* (1955), 470–475.

[5] AMITSUR, S. A., AND ROWEN, L. H. Elements of reduced trace 0. *Israel J. Math. 87*, 1-3 (1994), 161–179.

[6] BELOV, A. Y. No associative PI-algebra coincides with its commutant. *Sibirsk. Mat. Zh. 44*, 6 (2003), 1239–1254.

[7] BREŠAR, M., AND KLEP, I. Values of noncommutative polynomials, lie skew-ideals and tracial nullstellensätze. *Mathematical Research Letters 16*, 4 (2009), 605–626.

[8] BREŠAR, M. *Introduction to noncommutative algebra*. Universitext. Springer, Cham, 2014.

[9] BREŠAR, M. Commutators and images of noncommutative polynomials. *Adv. Math. 374* (2020), 107346, 21.

[10] BUZINSKI, D., AND WINSTANLEY, R. On multilinear polynomials in four variables evaluated on matrices. *Linear Algebra Appl. 439*, 9 (2013), 2712–2719.

[11] DRENSKY, V. *Free algebras and PI-algebras*. Springer-Verlag Singapore, Singapore, 2000. Graduate course in algebra.

[12] FAGUNDES, P. S., DE MELLO, T. C., AND DOS SANTOS, P. H. S. The mesyan conjecture: a restatement and a correction. *arXiv preprint 2111.13698* (2021).

[13] HERSTEIN, I. On the lie and jordan rings of a simple associative ring. *American Journal of Mathematics 77*, 2 (1955), 279–285.

[14] HERSTEIN, I. N. Sui commutatori degli anelli semplici. *Rend. Sem. Mat. Fis. Milano 33* (1963), 80–86.

[15] HERSTEIN, I. N. *Noncommutative rings*. The Carus Mathematical Monographs, No. 15. Published by The Mathematical Association of America; distributed by John Wiley & Sons, Inc., New York, 1968.

[16] HERSTEIN, I. N. *Topics in ring theory*. The University of Chicago Press, Chicago, Ill.-London, 1969.

[17] KANEL-BELOV, A., MALEV, S., AND ROWEN, L. The images of non-commutative polynomials evaluated on $2\times$ 2 matrices. *Proceedings of the American Mathematical Society 140*, 2 (2012), 465–478.

[18] KANEL-BELOV, A., MALEV, S., AND ROWEN, L. The images of multilinear polynomials evaluated on 3×3 matrices. *Proc. Amer. Math. Soc. 144*, 1 (2016), 7–19.

[19] KANEL-BELOV, A., MALEV, S., ROWEN, L., AND YAVICH, R. Evaluations of noncommutative polynomials on algebras: methods and problems, and the L'vov-Kaplansky conjecture. *SIGMA Symmetry Integrability Geom. Methods Appl. 16* (2020), Paper No. 071, 61.

[20] KAPLANSKY, I. A theorem on division rings. *Canad. J. Math. 3* (1951), 290–292.

[21] KATRE, S., AND GARGE, A. Matrices over commutative rings as sums of k-th powers. *Proceedings of the American Mathematical Society 141*, 1 (2013), 103–113.

[22] LARSEN, M., SHALEV, A., AND TIEP, P. H. The waring problem for finite simple groups. *Annals of mathematics* (2011), 1885–1950.

[23] LEE, J. Integral matrices as diagonal quadratic forms. *Linear and Multilinear Algebra 66*, 4 (2018), 742–747.

[24] LIU, Y.-R., AND WOOLEY, T. D. Waring's problem in function fields.

[25] MESYAN, Z. Polynomials of small degree evaluated on matrices. *Linear Multilinear Algebra 61*, 11 (2013), 1487–1495.

[26] SHODA, K. Einige Sätze über Matrizen. *Jpn. J. Math. 13*, 3 (1937), 361–365.

[27] VASERSTEIN, L. Waring's problem for algebras over fields. *Journal of Number Theory 26*, 3 (1987), 286–298.

[28] VITAS, D. Images of multilinear polynomials in the algebra of finitary matrices contain trace zero matrices. *Linear Algebra Appl. 626* (2021), 221–233.

[29] VITAS, D. Multilinear polynomials are surjective on algebras with surjective inner derivations. *J. Algebra 565* (2021), 255–281.

[30] ŠPENKO, V. On the image of a noncommutative polynomial. *J. Algebra 377* (2013), 298–311.