Universidade Estadual de Campinas

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA

Departamento de Matemática

Dissertação de Mestrado

PI-Álgebras

por

Alcindo Teles Galvão [†]

Mestrado em Matemática - Campinas - SP

Orientador: **Plamen Emilov Kochloukov**

[†]Este trabalho contou com o apoio financeiro do CNPq.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Galvão, Alcindo Teles

G139p PI-Álgebra / Alcindo Teles Galvão -- Campinas, [S.P. :s.n.], 2003.

Orientador : Plamen Emilov Kochloukov

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Álgebra não-comutativa. 2. Anéis (Álgebra). 3. Ideais (Álgebra). I. Kochloukov, Plamen Emilov. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

PI-Álgebras

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por **Alcindo Teles Galvão** e aprovada pela comissão julgadora.

Campinas, 12 de maio 2003.

Prof. Dr. Plamen Emilov Kochloukov

Banca examinadora:

Prof. Dr. Plamen Emilov Kochloukov

Prof. Dr. Antonio Giambruno.

Prof. Dr. Guilherme A. De La Rocque Leal.

Prof. Dr. Antonio José Engler (Suplente).

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP como requisito parcial para obtenção do título de **Mestre em Matemática**.

Dissertação de Mestrado defendida em 12 de maio de 2003 e aprovada pela Banca Examinadora composta pelos Profs. Drs.

Prof (a). Dr (a). PLAMEN EMILOV KOCHLOUKOV

Prof (a). Dr (a). ANTONIO GIAMBRUNO

Prof (a). Dr (a). GUILHERME AUGUSTO DE LA ROCQUE LEAL

Aos meus pais
José Ramos Galvão e
Maria José Teles dos Santos,
e a minha tia Nancy Galvão.

Agradecimentos

- Agradeço a Jeová Deus.
- Aos professores e Amigos, José Adonai Pereira Seixas e Amauri da Silva Barros.
- Aos professores da minha graduação, dos quais destaco Adroaldo Vasconcelos Dorvillé.
- À minha família. Pais, filhos e minha esposa Luciana Cardoso Barros, pelo apoio constante e por suportarem minha ausência.
- Aos colegas do IMECC, entre os quais destaco, Marcos Vergès, Rogério Casagrande, Elder, Erhan, Humberto, Maurício, Lidermir, Bianca, Karine, Kuo Poling, Mércio, Fernando de Maringá, Gilberlândio, Roger, Lucas, Edson Arrazola, Edson Licurgo, Evandro, Lucélia, Everaldo, Paulo Cesar, Benaia, Ercílio, Dirceu, Jones, Paula, Marcelo, Sérgio, Marcos, Luciano, Gilmar de Goiás e Gilmar do Paraná. Aos alagoanos Juliana, Rinaldo, Wagner, Vânio, Givaldo, José Barros, Selma e José Carlos.
- Além dos colegas do IMECC já citados, devo agradecimentos especiais aos alagoanos Gastão e Fernando. À mineirinha Vanessa, ao goiano Rosemberg e ao paraense Augusto.
- Ao pessoal da secretaria de pós-graduação, Cidinha, Tânia e Ednaldo.
- A todos os meus professores no mestrado e especialmente ao professor Marco Antônio Teixeira.
- Por fim, ao professor Plamen Emilov Kochloukov, pela amizade, pela extrema competência e pela paciência.

Resumo

Esta dissertação introduz as primeiras noções para o estudo combinatório da teoria de álgebras que satisfazem identidades polinomiais (resumidamente PI-álgebras), bem como alguns dos seus resultados mais importantes.

Apresentamos o teorema de Kaplansky e o teorema de Regev sobre produto tensorial de PI-álgebras. Além disso, descrevemos alguns resultados devidos a Amitsur e o teorema sobre identidades mínimas em álgebras matriciais conhecido como teorema de Amitsur e Levitzki. Consideramos também polinômios centrais e o teorema de Posner, o teorema sobre a altura, de Shirshov, incluindo o problema de Kurosh.

No final da dissertação desenvolvemos os métodos descobertos por Razmyslov, que o levaram a descrever uma base para as identidades polinomiais satisfeitas pela álgebra de Lie das matrizes de ordem dois com traço zero, e em seguida, para a álgebra (associativa) das matrizes de ordem dois.

Abstract

This dissertation introduces the first notions of the combinatorial study of the theory of algebras that satisfy polynomial identities (the so-called PI-algebras), as well as some of their most important results.

We present the theorems due to Kaplansky and Regev, about the tensor product of PI-algebras. Besides, we describe some results due to Amitsur and the theorem about minimum identities in matricial algebras known as Amitsur and Levitzki's theorem. We also consider central polynomials and Posner's theorem, and Shirshov's height theorem, including Kurosh's problem.

At the end of the dissertation we develop the methods discovered by Razmyslov which led him to the description of a basis for the polynomial identities satisfied by the Lie algebra of the traceless matrices of order two and, afterwards, for the (associative) algebra of all second order matrices.

Conteúdo

Introdução

A teoria das álgebras com identidades polinomiais é uma teoria relativamente recente, tendo seu maior desenvolvimento ocorrido nos últimos cinqüenta anos. Antes desse período podemos destacar trabalhos como os de M. Dehn (1922), W. Wagner (1936) e M. Hall (1943), em sua maioria motivados pela geometria.

Após 1945 intensificaram-se os estudos sobre as identidades polinomiais em álgebras. Naquela época surgiram resultados devidos a I. Kaplansky e N. Jacobson. O teorema de Amitsur e Levitzki, é um dos resultados mais importantes para essa teoria e foi demonstrado em 1950.

São vários os matemáticos que contribuíram para o estudo das identidades polinomiais em álgebras. Podemos citar ainda alguns como Higman, Nagata, Shirshov, Regev, Razmyslov, Procesi, Rowen, Cohn, Posner, Vaughan-Lee, Herstein, Formanek, Kostrikin, Specht, Zelmanov, Bahturin, entre outros.

Quanto ao texto, tentamos torná-lo o mais simples possível. Os dois primeiros capítulos fornecem as definições básicas, que estão destacadas para que possam ser consultadas rapidamente. No inicio de cada capítulo tecemos comentários sobre os aspectos históricos relacionados aos resultados a serem apresentados.

No capítulo 3, introduzimos polinômios multilineares, mostramos o teorema sobre a densidade, e em seguida o teorema de Kaplansky. O capítulo 4 considera produtos tensoriais e o teorema de Regev. Em seguida expomos alguns dos resultados clássicos de Amitsur.

O capítulo 6 está dedicado ao teorema de Amitsur e Levitzki, onde nós seguimos a demonstração dada por Rosset. No Capítulo 7, definimos polinômios centrais e discutimos o teorema de Posner. O Capítulo 8 considera uma das aplicações mais importantes da combinatória na PI teoria, o teorema de Shirshov sobre a altura. No último capítulo estudamos as identidades satisfeitas pela álgebra matricial de ordem 2, $M_2(K)$, sobre um corpo K de característica 0. Seguindo a demonstração de Razmyslov, exibimos uma base finita para as identidades satisfeitas pela álgebra de Lie sl_2, para as identidades fracas em $(M_2(K), sl_2(K))$, e em seguida mostramos que as identidades de $M_2(K)$ também admitem uma base finita.

Campinas, maio de 2003.

Capítulo 1

Preliminares

Nas primeiras seções colocaremos para consulta rápida as definições elementares que, porém, não nos sentimos a vontade de suprimí-las. O texto é simples e tem por objetivo deixar claro os conceitos que usaremos nos capítulos posteriores.

1.1 Grupos

Definição 1.1.1 *Um conjunto M, não vazio, será chamado* ***monóide*** *se tivermos uma operação de $M \times M$ em M que associa dois elementos a e b em M a outro elemento ab em M, tal que*

- *vale $a(bc) = (ab)c$ para todo a, b e c em M,*
- *existe um elemento $e \in M$ tal que $ea = ae = a$ para todo $a \in M$.*

Definição 1.1.2 *Um subconjunto de um monóide M que é um monóide com a operação induzida de M é chamado* ***sub-monóide*** *de M.*

Definição 1.1.3 *Um conjunto G, não vazio, será chamado* ***grupo*** *se tivermos uma operação de $G \times G$ em G que associa dois elementos a e b em G a outro elemento ab em G, tal que*

- *vale $a(bc) = (ab)c$ para todo a, b e c em G,*
- *existe um elemento $e \in G$ tal que $ea = ae = a$ para todo $a \in G$,*
- *para todo $a \in G$ existe um elemento $a^{-1} \in G$, tal que $aa^{-1} = a^{-1}a = e$.*

Definição 1.1.4 *Se G é um grupo e*

- *vale $ab = ba$ para todo a e b em G,*

então G é chamado ***grupo abeliano*** *ou* ***grupo comutativo.***

Definição 1.1.5 *Um subconjunto de um grupo G que também é um grupo com a operação induzida de G, é chamado* ***subgrupo de*** *G.*

Definição 1.1.6 *Sejam G e H, grupos. Uma aplicação $\varphi : G \to H$ é um* ***homomorfismo de grupos*** *se para quaisquer a e b em G a seguinte condição é satisfeita:*

$$\varphi(ab) = \varphi(a)\varphi(b).$$

Se φ é bijetora dizemos que φ é um ***isomorfismo de grupos****, e desta forma diremos que G e H são isomorfos. Quando temos $\varphi : G \to G$, dizemos que φ é um endomorfismo do grupo G e, neste caso, se φ é bijetora então ela é um automorfismo de G.*

O grupo que mais usaremos será o grupo simétrico formado pelas permutações dos elementos do conjunto $N = \{1, 2, \ldots, n\}$ que denotaremos por S_n.

1.2 Anéis

Definição 1.2.1 *Um* ***anel*** *é um conjunto R, não vazio, dotado de duas operações binárias (isto é, ambas as operações são de $R \times R$ em R); adição, que associa a e b a $a + b$, e multiplicação, que associa a e b a ab tais que:*

- *R é um grupo abeliano em relação a adição,*
- *vale $a(b + c) = ab + ac$ e $(b + c)a = ba + ca$ para todo a, b e c em R.*

Costuma-se geralmente denominar os anéis dependendo dos axiomas que a multiplicação satisfaz.

Definição 1.2.2 *Se R é um anel e*

- *existe um elemento $1 \in R$, $1 \neq 0$, tal que $a1 = 1a = a$ para todo a em R,*

então R é chamado ***anel com unidade*** *e 1 é chamado elemento neutro da multiplicação ou simplesmente unidade de R.*

Definição 1.2.3 *Se R é um anel tal que*

- *vale $a(bc) = (ab)c$ para todo a, b e c em R,*

então R é dito um ***anel associativo****.*

Sempre que nos referirmos simplesmente a anéis, estaremos pensando em anéis associativos com unidade. Caso haja necessidade de diferenciá-los, faremos isso em cada caso.

Definição 1.2.4 *Se R é um anel e*

- *vale $ab = ba$ para todo a, b em R,*

então R é dito um ***anel comutativo****.*

Definição 1.2.5 *Se R é um anel e*

- *se $ab = 0$ implicar que $a = 0$ ou $b = 0$ para todo a, b em R,*

então R é dito um ***domínio***.

Definição 1.2.6 *Se R é um anel com unidade tal que*

- *para todo elemento $a \in R$, $a \neq 0$, existe um elemento $a^{-1} \in R$ tal que $aa^{-1} = a^{-1}a = 1$,*

então R é chamado ***anel com divisão*** *e a^{-1} é chamado inverso multiplicativo de a.*

Definição 1.2.7 *Um anel com divisão comutativo K é chamado* ***corpo***.

Definição 1.2.8 *Um* ***subanel*** *é um subconjunto S de um anel R tal que ele próprio é um anel com as operações induzidas por R.*

Definição 1.2.9 *Sejam R um anel e I um subgrupo aditivo de R. Se $ax \in I$ (respectivamente $xa \in I$) para todo $a \in R$ e para todo $x \in I$, dizemos que I é um* ***ideal*** *à esquerda (respectivamente, à direita) de R. Se I for ideal à esquerda e à direita de R, diremos que I é um ideal bilateral de R.*

Definição 1.2.10 *Um ideal I de um anel R é dito* ***primo*** *se, para quaisquer dois ideais I_1 e I_2 de R, a inclusão $I_1 I_2 \subseteq I$ implicar que $I_1 \subseteq I$ ou $I_2 \subseteq I$.*

Definição 1.2.11 *Um anel R é dito* ***semi-primo*** *se para todo ideal $I \neq 0$ de R temos que $I^2 \neq 0$.*

Definição 1.2.12 *Um anel R é dito* ***primo*** *se $IJ \neq 0$ para todos os ideais não nulos I, J de R.*

Definição 1.2.13 *Um anel R é dito* ***noetheriano à esquerda*** *se ele satisfaz uma das condições equivalentes abaixo:*

1. *todo ideal I à esquerda de R é finitamente gerado, isto é, $I = Ry_1 + \cdots + Ry_n = (y_1, \ldots, y_n)$, $y_i \in R$;*
2. *toda cadeia ascendente, $I_1 \subseteq I_2 \subseteq \ldots \subseteq I_n \subseteq \ldots$ de ideais à esquerda de R é estacionária, isto é, existe um natural n tal que $I_n = I_{n+1} = \cdots$;*
3. *toda família, não vazia, $\mathcal{F}$ de ideais à esquerda de R tem um elemento maximal, isto é, existe $J \in \mathcal{F}$ tal que se $J \subseteq I$ e $I \in \mathcal{F}$ então $I = J$.*

Definição 1.2.14 *Um anel R é dito* ***artiniano*** *se ele satisfaz uma das condições equivalentes abaixo:*

1. *todo ideal I à direita de R é finitamente gerado, isto é, $I = y_1 R + \cdots + y_n R = (y_1, \ldots, y_n)$, $y_i \in R$;*
2. *toda cadeia descendente, $I_1 \supseteq I_2 \supseteq \ldots \supseteq I_n \supseteq \ldots \supseteq$ de ideais à direita de R é estacionária, isto é, existe um natural n tal que $I_n = I_{n+1} = \cdots$;*

3. *toda família, não vazia, $\mathcal{F}$ de ideais à direita de R tem um elemento minimal, isto é, existe $J \in \mathcal{F}$ tal que se $J \supseteq I$ e $I \in \mathcal{F}$ então $I = J$.*

Uma demonstração das equivalências citadas nas definições anteriores pode ser encontrada em [2].

Definição 1.2.15 *Sejam R_1 e R_2, anéis. Uma aplicação $\varphi : R_1 \to R_2$ é um* ***homomorfismo de anéis*** *se para quaisquer a e b em R_1, as seguintes condições são satisfeitas:*

$$\varphi(a+b) = \varphi(a) + \varphi(b) \text{ e } \varphi(ab) = \varphi(a)\varphi(b).$$

Se φ é bijetora dizemos que φ é um ***isomorfismo de anéis****, e desta forma diremos que R_1 e R_2 são isomorfos. Quando temos $\varphi : R \to R$, dizemos que φ é um endomorfismo do anel R e, neste caso, se φ é bijetora então ela é um automorfismo de R.*

1.3 Álgebras

Definição 1.3.1 *Um espaço vetorial A sobre um corpo K é chamado uma* ***álgebra*** *sobre K (ou uma K-álgebra) se A é dotado de uma operação, chamada multiplicação, que a dois elementos a e b em A associa um elemento ab em A, tal que para todo a, b em A e para todo α em K*

- $(a+b)c = ac + bc$,
- $a(b+c) = ab + ac$,
- $\alpha(ab) = (\alpha a)b = a(\alpha b)$.

Claramente, a noção de álgebra generaliza grosseiramente a noção de espaço vetorial e de anel.

Definição 1.3.2 *Se A é uma álgebra e*

- *existe um elemento $1 \in A$, tal que $a1 = 1a = a$ para todo a em A,*

então A é chamada ***álgebra com unidade*** *e 1 é chamado elemento neutro da multiplicação ou simplesmente unidade de A.*

Definição 1.3.3 *Se A é uma álgebra tal que*

- *vale $a(bc) = (ab)c$ para todo a, b e c em A,*

então A é chamada uma ***álgebra associativa****.*

Sempre que nos referirmos simplesmente a álgebra, estaremos pensando em álgebras associativas com unidade. Se houver necessidade de diferenciá-las, faremos isso em cada caso.

Definição 1.3.4 *Se A é uma álgebra e*

- *vale* $ab = ba$ *para todo* a, b *em* A,

então A *é chamada uma* ***álgebra comutativa.***

Definição 1.3.5 *Se* A *é uma álgebra com unidade tal que*

- *para todo elemento* $a \in A$, $a \neq 0$, *existe um elemento* $a^{-1} \in A$ *tal que* $aa^{-1} = a^{-1}a = 1$ *onde* 1 *é a unidade de* A,

então A *é chamada* ***álgebra com divisão*** *e* a^{-1} *é chamado inverso multiplicativo de* a.

Definição 1.3.6 *Um subespaço* S *de* A *é uma* ***subálgebra*** *se é fechado com respeito a multiplicação, i.é.,*

$$s_1 \, e \, s_2 \in S \implies s_1 s_2 \in S.$$

Definição 1.3.7 *Seja* A *uma álgebra (ou um anel). O conjunto*

$$Z = \{z \in A \mid za = az \ \ para \ todo \ a \in A\}$$

é chamado ***centro de A****. Os elementos de* Z *são ditos centrais. Usaremos convenientemente as notações* Z, $Z(A)$, *ou* $Z(R)$ *conforme precisarmos do centro de uma álgebra ou de um anel.*

Definição 1.3.8 *Diremos que* A *é uma* ***álgebra de Lie*** *se para todo* a, b *e* c *em* A *temos* $aa = 0$ *que é a anti-comutatividade e* $a(bc) + b(ca) + c(ab) = 0$ *que é a identidade de Jacobi.*

Definição 1.3.9 *Uma álgebra associativa* A *é dita uma* ***nil álgebra*** *se para todo* $a \in A$ *existe um número* n *natural tal que* $a^n = 0$. *O menor natural com essa propriedade é chamado índice de nilpotência de* a. *Se existir um número fixo* n *tal que* $a^n = 0$ *para todo* $a \in A$ *então* A *é chamada uma nil álgebra de índice limitado* n.

Definição 1.3.10 *Uma álgebra associativa* A *é dita uma* ***álgebra nilpotente*** *se existe um número natural fixo* n *tal que o produto de quaisquer* n *elementos de* A *é zero. O menor natural com essa propriedade é chamado índice de nilpotência de* A.

Definição 1.3.11 *Sejam* A *uma álgebra e* I *um subespaço de* A. *Se* $ax \in I$ *(respectivamente* $xa \in I$*) para todo* $a \in A$ *e para todo* $x \in I$, *dizemos que* I *é um* ***ideal*** *à esquerda (respectivamente, à direita) de* A. *Se* I *for ideal à esquerda e à direita de* A, *diremos que* I *é um ideal bilateral de* A.

Definição 1.3.12 *Sejam* A *uma álgebra e* I *um ideal à esquerda (à direita ou bilateral) de* A. *Se não existe nenhum ideal à esquerda (à direita ou bilateral) de* A *contido em* I *e que seja distinto de* (0) *e do próprio* I *então* I *é chamado um* ***ideal simples (ou minimal)*** *à esquerda (à direita ou bilateral) de* A.

Definição 1.3.13 *Seja* X *um conjunto não vazio, cujos elementos são chamados letras ou símbolos.*

1. *X será chamado* ***alfabeto.***

2. *Uma seqüência finita de letras de X, será chamada* ***palavra*** *e o conjunto dessas palavras será denotado por X^*.*

3. *O número de letras de cada palavra $w \in X^*$ será chamado,* ***comprimento da palavra*** *w. A palavra de comprimento zero será chamada palavra vazia.*

4. *Uma palavra $v \in X^*$ será dita uma* ***sub-palavra*** *de uma palavra $w \in X^*$ se existirem as palavras w_1 e w_2 em X^* tais que $w = w_1 v w_2$.*

5. *Dadas $v = a_1 \dots a_m$ e $w = b_1 \dots b_n$, palavras em X^*, a operação de $X^* \times X^*$ em X^* definida por $vw = a_1 \cdots a_m b_1 \cdots b_n$ é chamada* ***concatenação*** *ou* ***justaposição.***

6. *Podemos definir uma relação de ordem em X^* da seguinte forma:*
   sejam $v = x_{i_1} \cdots x_{i_l}$ e $w = x_{j_1} \cdots x_{j_p}$ palavras em X^*, então diremos que $v \leq w$ se um dos três seguintes casos ocorre:

   (i) $v = w$, ou seja, $l = p$ e $i_t = j_t$ para $t = 1, 2, \dots, l$,

   (ii) existe $r \in \{1, 2, \dots, l\}$ tal que $i_t = j_t$ para $1 \leq t \leq r-1$ e $i_r < j_r$,

   (iii) $l < p$ e $i_t = j_t$ para $1 \leq t \leq l$.

   Diremos também que $v < w$ se $v \leq w$ e $v \neq w$. Esta relação de ordem denomina-se ***ordem lexicográfica.***

Neste contexto surgem as seguintes definições:

Definição 1.3.14 *Sejam $X = \{x_1, x_2, \dots\}$ um alfabeto enumerável e K um corpo. Consideraremos para essas definições, um espaço vetorial sobre K com base $\{w, 1\}$ composta pelos monômios w (definidos à frente) sobre X e a unidade em K, onde a multiplicação é a concatenação.*

1. *A soma $p(X) = p(x_1, \dots, x_n) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_m w_m$, onde $\alpha_i \in K$ e $w_i = x_1^{k_1} \cdots x_n^{k_n} \in X^*$ com cada $k_j \geq 0$, é chamada* ***polinômio.*** *No caso em que $m = 1$, $p(X)$ é um* ***monômio.*** *Cada x_j é dita uma indeterminada, incógnita ou variável de $p(X)$.*

2. *O* ***grau*** *de um monômio αw com $\alpha \in K$, $\alpha \neq 0$, e $w \in X^*$, que denotaremos por $gr(\alpha w)$, é o comprimento da palavra w. O grau de um polinômio $p(X)$ será o grau máximo de seus monômios.*

3. *Diremos que um monômio αw com $\alpha \in K$ e $w \in X^*$ tem* ***tipo*** *$[n_1, \dots, n_k]$ se a palavra w contém x_i exatamente n_i vezes, $n_k \neq 0$ e $n_i = 0$ para todo $i > k$. Diremos que o número n_i é o grau do monômio αw em x_i.*

4. *Diremos que um* ***polinômio é homogêneo*** *em x_i com grau n_i se todos os seus monômios têm grau n_i em x_i. Diremos também que um polinômio é homogêneo do tipo $[n_1, \dots, n_k]$ se todos os seus monômios são do mesmo tipo, $[n_1, \dots, n_k]$.*

5. *Se agruparmos os monômios do mesmo tipo em qualquer polinômio p, ele se escreve como uma soma de polinômios homogêneos. Estes polinômios serão chamados as* ***componentes homogêneas*** *de p.*

6. *Um polinômio homogêneo em X_i com grau 1 em X_i é chamado* ***polinômio linear*** *em X_i. Um polinômio homogêneo do tipo $[n_1, \ldots, n_k]$, onde $n_i = 0$ ou $n_i = 1$ para cada i, é chamado* ***polinômio multi-linear.***

Observação 1.3.15 1. *O conjunto M, dos monômios sobre um alfabeto X, dotado da concatenação e que possua um elemento unidade 1, tal que $1w = w1 = w$ para cada monômio $w \in M$, é um monóide. Escrevemos $M\{X\}$ e o chamamos monóide livre.*

2. *Podemos determinar uma ordem em $M\{X\}$. Dados dois monômios $w_1 = x_{i_1} \ldots x_{i_k}$ e $w_2 = x_{j_1} \ldots x_{j_u}$, diferentes do monômio vazio, em $M\{X\}$, dizemos que $w_1 < w_2$ se $x_{i_1} < x_{j_1}$ ou, indutivamente, se $x_{i_1} = x_{j_1}$ e temos $x_{i_2} \ldots x_{i_k} < x_{j_2} \ldots x_{j_u}$. Diferente da ordem lexicográfica, essa ordem não é total. Por exemplo, $w_1 = x_1x_2$ e $w_2 = x_1x_2x_3$ não são comparáveis. É conveniente considerar 1 como correspondendo ao monômio vazio e assim 1 terá grau zero. Também é conveniente considerar $1 < w$ para qualquer monômio $w \in M\{X\}$.*

Vamos agora citar alguns exemplos de álgebras e subálgebras.

Exemplo 1.3.16 1. *Uma extensão de corpos $L \mid K$ é uma álgebra sobre K, associativa, comutativa, com unidade $1 \in L$.*

2. *O conjunto $K[x]$ dos polinômios com uma indeterminada e com coeficientes em K, é uma álgebra sobre K, associativa, comutativa, com unidade $1 \in K[x]$.*

3. *Os polinômios $K[x_1, \ldots, x_n]$ com n indeterminadas e com coeficientes em K, também é uma álgebra sobre K, associativa, comutativa, com unidade $1 \in K[x_1, \ldots, x_n]$.*

4. *O conjunto, $K\{x_1, x_2, \ldots\} = K\{X\}$, dos polinômios com indeterminadas em um alfabeto enumerável X, associativas, não comutativas, e com coeficientes em K, é uma álgebra sobre K, associativa, não comutativa, com unidade $1 \in K\{X\}$.*

5. *O conjunto $M_n(K)$ das matrizes $n \times n$ com entradas em K, é uma álgebra sobre K, associativa, não comutativa para $n > 1$, cuja unidade é a matriz identidade I de ordem n.*

6. *O conjunto $End_K(V)$ dos operadores lineares de um espaço vetorial V, com $\dim V > 1$, é uma álgebra sobre K, associativa, não comutativa, cuja unidade é o operador identidade de V.*

7. *O subconjunto $T_n(K)$ de $M_n(K)$, formado por todas as matrizes triangulares superiores é uma álgebra sobre K, associativa, não comutativa, com a mesma unidade $I \in M_n(K)$.*

8. *O subconjunto $sl_n(K)$ de $M_n(K)$, formado pelas matrizes com traço zero e com a multiplicação $[r_1, r_2] = r_1r_2 - r_2r_1$, $r_1, r_2 \in sl_n(K)$, é uma álgebra de Lie, sobre K.*

9. *O subconjunto* $S_n(K)$ *de* $M_n(K)$, *formado pelas matrizes simétricas com a multiplicação* $s_1 \circ s_2 = \frac{1}{2}(s_1 s_2 + s_2 s_1)$, *é uma álgebra sobre* K, *não associativa, comutativa, com unidade* $I \in M_n(K)$.

10. *O subconjunto* $O_n(K)$ *de* $M_n(K)$, *formado pelas matrizes anti-simétricas com a multiplicação* $[r_1, r_2]$ *é uma álgebra de Lie, sobre* K.

11. *Seja* $KG = \{\sum_{g \in G} \alpha_g g \mid \alpha_g \in K\}$ *o espaço vetorial com base* $\{g \mid g \in G\}$ *onde* G *é um grupo finito. A multiplicação em* KG *será dada por* $(\sum_{g \in G} \alpha_g g)(\sum_{h \in G} \beta_h h) = \sum_{g,h \in G} \alpha_g \beta_h gh$; $\alpha_g, \beta_h \in K$. *Aqui* gh *é o produto de* g *e* h *em* G. KG *é uma álgebra associativa com unidade sobre* K *que é chamada, álgebra de grupo. A unidade da álgebra de grupo é o elemento* $1e$, *onde* e *é a unidade do grupo* G.

Naturalmente, as operações não definidas nos exemplos anteriores são as usuais. Vale observar que usaremos o mesmo símbolo para unidade da álgebra e do corpo a menos que seja necessário diferenciá-las.

Definição 1.3.17 *Sejam* A *e* B *duas álgebras sobre um corpo* K. *Uma aplicação* $\varphi : A \to B$ *é um* ***homomorfismo de álgebras*** *se para quaisquer* a *e* b *em* A *e* α *em* K *as seguintes condições são satisfeitas:*

$$\varphi(a+b) = \varphi(a) + \varphi(b), \quad \varphi(a.b) = \varphi(a).\varphi(b) \text{ e } \varphi(\alpha a) = \alpha\varphi(a).$$

Se φ *é bijetora dizemos que* φ *é um* ***isomorfismo de álgebras****, e desta forma diremos que* A *e* B *são isomorfas. Quando temos* $\varphi : A \to A$, *dizemos que* φ *é um endomorfismo da álgebra* A *e, neste caso, se* φ *é bijetora, então ela é um automorfismo de* A.

Capítulo 2

PI-Álgebras

Aqui introduziremos os conceitos de PI-álgebras, módulos, álgebras livres e T-ideais, além de outros relacionados a esses.

2.1 PI-Álgebras

Daqui em diante a notação $K\{X\}$ representará a álgebra associativa com unidade (não comutativa) dos polinômios a várias variáveis em um alfabeto enumerável X com coeficientes em um corpo K (como no exemplo (1.3.16)).

Definição 2.1.1 *Seja A uma álgebra associativa sobre um corpo K e $p(x_1, \ldots, x_n)$ um polinômio, não identicamente nulo, em $K\{X\}$. Se $p(a_1, \ldots, a_n) = 0$ para todo $a_1, \ldots, a_n$ em A, então $p(x_1, \ldots, x_n) = 0$ é uma* ***identidade polinomial*** *de A e A é dita uma* ***álgebra com identidade polinomial****, resumidamente,* ***PI-álgebra****.*
Futuramente, por abuso de linguagem, poderemos dizer simplesmente que p é uma identidade polinomial de A.

Antes dos exemplos, de PI-álgebras, vamos definir alguns polinômios que aparecerão com freqüência nas identidades polinomias.

Definição 2.1.2 *O polinômio*

$$s_n(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} sign(\sigma) x_{\sigma(1)} \ldots x_{\sigma(n)}$$

é chamado ***polinômio standard*** *de grau n. Aqui $sign(\sigma)$ é o sinal da permutação σ do grupo simétrico S_n.*

Exemplo 2.1.3 *O polinômio*

$$s_2(x_1, x_2) = [x_1, x_2] = x_1 x_2 - x_2 x_1$$

é o polinômio standard para $n = 2$ e é chamado polinômio ***comutador****.*

Definição 2.1.4 *O polinômio*

$$e_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \ldots x_{i_k}$$

onde $k \leq n$ e as variáveis são não comutativas, é chamado ***polinômio simétrico*** *de grau k em n variáveis.*

Definição 2.1.5 *O polinômio*

$$p_k(x_1, \ldots, x_n) = x_1^k + \cdots + x_n^k$$

é chamado ***soma de potências*** *de grau k em n variáveis.*

Vamos agora a alguns exemplos de PI-álgebras.

Exemplo 2.1.6 1. *Naturalmente, se A é uma álgebra associativa e comutativa então*

$$s_2 = [x_1, x_2] = x_1x_2 - x_2x_1 = 0$$

é uma identidade polinomial de A. Assim, toda álgebra associativa e comutativa é uma PI-álgebra. Observamos que toda álgebra comutativa (não necessariamente associativa) satisfaz a identidade s_2.

2. *A álgebra das matrizes 2×2 sobre um corpo K, denotada por $M_2(K)$, satisfaz a identidade de Hall*

$$p(x_1, x_2, x_3) = [x_1, [x_2, x_3]^2] = x_1(x_2x_3 - x_3x_2)^2 - (x_2x_3 - x_3x_2)^2x_1 = 0.$$

Isto verifica-se pois se a e b pertencem a $M_n(K)$, $n \in \mathbb{N}$, então o traço de $[a, b]$ é zero e se a pertence a $M_2(K)$ e o traço de a é zero então $a^2 = \lambda I$ onde $\lambda \in K$ e I é a matriz identidade 2×2. Esta última matriz comuta com todas as matrizes.

3. *A álgebra das matrizes, $T_n(K)$, triangulares superiores $n \times n$ sobre um corpo K, é uma PI-álgebra pois satisfaz a seguinte identidade polinomial:*

$$p(x_1, \ldots, x_{2n}) = [x_1, x_2][x_3, x_4] \cdots [x_{2n-1}, x_{2n}] = 0$$

pois se a e b pertencem a $T_n(K)$ então $[a, b]$ ainda pertence a $T_n(K)$ e tem diagonal principal nula. Além disso, o produto de n matrizes de $T_n(K)$ com diagonais principais nulas é a matriz nula.

4. *Toda nil álgebra associativa com índice limitado n é uma PI-álgebra com identidade polinomial*

$$p(x) = x^n = 0.$$

5. *Toda álgebra nilpotente, com índice de nilpotência n, é uma PI-álgebra com identidade polinomial*

$$p(x_1, \ldots, x_n) = x_1 \ldots x_n = 0.$$

6. *Toda álgebra associativa A, sobre um corpo K, de dimensão finita n, é uma PI-álgebra pois satisfaz a identidade polinomial*

$$s_{n+1}(x_1, \ldots, x_{n+1}) = \sum_{\sigma \in S_{n+1}} sign(\sigma) x_{\sigma(1)} \ldots x_{\sigma(n+1)} = 0.$$

É fácil ver, da definição do polinômio standard que, como ele é multi-linear e anti-simétrico ele se anula quando calculado sobre elementos linearmente dependentes. Particularmente, isso ocorre quando dois de seus argumentos são iguais. Assim, sejam $\mathcal{B} = \{e_1, \ldots, e_n\}$ uma base de A e $a_1, \ldots, a_{n+1}$ elementos de A. Podemos escrever cada a_i como combinação linear de elementos de $\mathcal{B}$. Temos então que $s_{n+1}(a_1, \ldots, a_{n+1})$ é uma combinação linear com coeficientes em K de elementos da forma $s_{n+1}(e_{i_1}, \ldots, e_{i_{n+1}})$ onde cada $e_{i_k} \in \mathcal{B}$. Sendo assim, algum e_{i_k} se repetirá, e portanto $s_{n+1}(e_{i_1}, \ldots, e_{i_{n+1}}) = 0$ implicando que $s_{n+1}(a_1, \ldots, a_{n+1}) = 0$.

2.2 Módulos

Definição 2.2.1 *Seja R um anel. Um grupo aditivo abeliano M é dito um R-**módulo** (à esquerda) se existir uma operação de $R \times M$ em M associando (a, m) a am tal que*

1. $(ab)m = a(bm)$,
2. $(a + b)m = am + bm$,
3. $a(m + n) = am + an$,

para todo $m, n \in M$ e $a, b \in R$.
Além disso, se R tem unidade 1 e $1m = m$ para todo $m \in M$ então M é dito um R-módulo unitário.

Definição 2.2.2 *Seja R um anel. Um subconjunto N de um R-módulo M, é chamado um **submódulo** de M se N é um R-módulo com as operações induzidas por M.*

Definição 2.2.3 *Seja R um anel e sejam M e N R-módulos. Uma aplicação $\varphi : M \to N$ é um **homomorfismo de R-módulos** se para quaisquer r em R e m em M a seguinte condição é satisfeita:*

$$\varphi(r.m) = r\varphi(m).$$

*Se φ é bijetora dizemos que φ é um **isomorfismo de R-módulos**, e desta forma diremos que M e N são isomorfos. Quando temos $\varphi : M \to M$, dizemos que φ é um endomorfismo do R-módulo M e, neste caso, se φ é bijetora, então ela é um automorfismo de M.*

Definição 2.2.4 *Um R-**módulo** à esquerda (direita) é dito **fiel** se $rM = (0)$ ($Mr = (0)$), com $r \in R$, implicar que $r = 0$.*

Definição 2.2.5 *Um R-**módulo** M é dito **irredutível** se $MR \neq 0$ e os únicos submódulos de M são (0) e o próprio M. Caso contrário, M é dito redutível.*

Definição 2.2.6 *Um* ***anel*** R *é dito* ***primitivo*** *se ele tem um módulo fiel irredutível.*

Definição 2.2.7 *Diremos que um* ***ideal*** I *de um anel* R *é* ***primitivo*** *se* R/I *é primitivo.*

Definição 2.2.8 *A interseção* $J(R)$ *de todos os ideais primitivos de* R *é chamada* ***radical de Jacobson*** *do anel* R. *Se* R *não tem ideais primitivos, definimos* $J(R) = R$.

Definição 2.2.9 *Um anel* R *é dito* ***semi-primitivo*** *se* $J(R) = 0$.

Podemos então escrever definições análogas quando tivermos uma K-álgebra A no lugar do anel R.

Definição 2.2.10 *Seja* A *uma* K*-álgebra. Então, um* K*-espaço vetorial* M *é dito um* A***-módulo*** *(à esquerda) se existir uma operação de* $A \times M$ *em* M *associando* (a, m) *a* am *tal que, para todo* a *e* b *em* A, m *e* n *em* M *e* α *em* K *as seguintes condições são satisfeitas:*

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

Além disso, se A *tem unidade* 1 *e* $1m = m$ *para todo* $m \in M$ *então* M *é dito um* A*-módulo unitário.*

Claramente poderíamos ter dado definições à direita mas, o nosso uso mais comum será nessa forma. A menos que mencionemos o contrário, A-módulos e R-módulos, para nós, serão unitários e à esquerda.

Definição 2.2.11 *Seja* A *uma álgebra com unidade. Um subconjunto* N *de um* A*-módulo* M *é um* ***submódulo*** *de* M *se* N *é um* A*-módulo com as operações induzidas por* M.

Definição 2.2.12 *Seja* A *uma álgebra com unidade sobre um corpo* K *e sejam* M *e* N *dois* A*-módulos sobre* K. *Uma aplicação* $\varphi : M \to N$ *é um* ***homomorfismo de módulos*** *se para quaisquer* m *e* n *em* M, α *em* K *e* a *em* A *as seguintes condições são satisfeitas:*

$$\varphi(m + n) = \varphi(m) + \varphi(n), \quad \varphi(\alpha m) = \alpha\varphi(m) \text{ e } \varphi(a.m) = a\varphi(m).$$

Se φ *é bijetora dizemos que* φ *é um* ***isomorfismo de módulos****, e desta forma diremos que* M *e* N *são isomorfos. Quando temos* $\varphi : M \to M$, *dizemos que* φ *é um endomorfismo do módulo* M *e, neste caso, se* φ *é bijetora, então ela é um automorfismo de* M.

Proposição 2.2.13 *Se* A *é uma álgebra com unidade e* $\varphi : M \to N$ *é um homomorfismo de módulos , então* $\ker\varphi$ *(o núcleo de* φ*) é um submódulo de* M *e* $\varphi(M)$ *é um submódulo de* N. *Além disso, se* $\ker\varphi = \{0\}$ *então* φ *é injetora.*

Poderíamos ter dado uma proposição análoga à proposição anterior, para cada uma das estruturas algébricas citadas até aqui. O mesmo acontece em relação a algumas definições, como a seguinte, que podem ser facilmente adaptadas a outras estruturas.

Definição 2.2.14 *Seja N um sub-módulo do A-módulo M. Então $M/N = \{x+N \mid x \in M\}$ herdará uma estrutura de A-módulo definida por $a(x+N) = ax + aN$. M/N é chamado o* ***módulo quociente*** *de M por N.*

Definição 2.2.15 *Seja M um A-**módulo** à esquerda(direita). M será dito **fiel** se $aM = (0)$ $(Ma = (0))$, com $a \in A$, implicar que $a = 0$.*

Definição 2.2.16 *Seja A uma álgebra com unidade. Um A-**módulo** M é dito **irredutível** se os únicos submódulos de M são (0) e o próprio M. Caso contrário, M é dito redutível.*

Definição 2.2.17 *Uma álgebra A é chamada **primitiva** se ela tem um módulo fiel irredutível.*

Definição 2.2.18 *Uma álgebra A, associativa com unidade que é uma soma direta de seus ideais simples à esquerda é chamada **álgebra semi-simples**.*

Definição 2.2.19 *Uma álgebra A tal que $A^2 \neq 0$ e que seus únicos ideais são (0) e a própria A, é chamada **álgebra simples**.*

2.3 Álgebras Livres em uma Variedade

Definição 2.3.1 *Seja R um anel associativo e comutativo com elemento unidade 1. Um R-módulo A é chamado uma álgebra sobre o anel R se nele está definida uma operação multiplicação conectada com a operação de módulo pelas relações*

$$(a+b)c = ac + bc, \quad a(b+c) = ab + ac \quad e \quad \alpha(ab) = (\alpha a)b = a(\alpha b)$$

para a, b, $c \in A$ e $\alpha \in R$.

Naturalmente será comum escrevermos R-álgebras quando quisermos dizer uma álgebra sobre o anel R. Nestas condições, o anel R poderá ser chamado anel de escalares ou anel de operadores. Pode também ocorrer que, no lugar do anel R, tenhamos um corpo K. Então fala-se de uma K-álgebra, referindo-se a uma álgebra sobre K e chama-se K o corpo de escalares. Todo anel é uma álgebra sobre $\mathbb{Z}$. Assim, o conceito de uma álgebra sobre um anel combina o conceito de um anel e uma álgebra sobre um corpo.

Dada uma classe de K-álgebras que possuem uma propriedade P, tais como associatividade, comutatividade etc.; nós podemos, muitas vezes, estar interessados na álgebra "livre" (definida à frente) em relação a propriedade P. Tal álegbra livre é o objeto universal com a propriedade P. No nosso caso, interessam as álgebras associativas.

No próximo teorema e na definição de álgebra livre vamos considerar um alfabeto X.

Teorema 2.1 *Sejam A uma K-álgebra e ϕ alguma função de X em A. Então ϕ pode ser estendida de maneira única a um homomorfismo da álgebra $K\{X\}$ na álgebra A.*

Demonstração. Digamos que $\phi(x_i) = a_i$ e consideremos $f(x_1, \ldots, x_n) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n \in K\{X\}$ onde os $w_i = x_{i_1}^{k_{i_1}} x_{i_2}^{k_{i_2}} \ldots x_{i_m}^{k_{i_m}}$ são palavras sobre X. Vamos inicialmente estender ϕ ao conjunto, X^*, dessas palavras. Definiremos $\varphi : X^* \to A$ por $\varphi(w_i) = \phi(x_{i_1}^{k_{i_1}})\phi(x_{i_2}^{k_{i_2}}) \cdots \phi(x_{i_m}^{k_{i_m}})$. Agora obteremos a extensão final definindo $\overline{\phi} : K\{X\} \to A$ por $\overline{\phi}(f) = \alpha_1\varphi(w_1) + \alpha_2\varphi(w_2) + \cdots + \alpha_n\varphi(w_n)$. Resta-nos apenas verificar que $\overline{\phi}$ é um homomorfismo. ■

Tendo em vista o teorema anterior obtemos a definição de uma álgebra livre numa dada classe $\mathcal{C}$.

Definição 2.3.2 *Seja $\mathcal{C}$ uma classe de álgebras e seja $A \in \mathcal{C}$ uma álgebra gerada por um conjunto X. A álgebra A será chamada uma álgebra livre na classe $\mathcal{C}$, livremente gerada por X, se para qualquer álgebra $R \in \mathcal{C}$, toda função de X em R pode ser estendida a um único homomorfismo de A em R. A cardinalidade $\mid X \mid$ do conjunto X é chamada o posto de A.*

Exemplo 2.3.3 *Para qualquer alfabeto X a álgebra polinomial $K[X]$ é livre na classe das álgebras associativas comutativas unitárias.*

O próximo exemplo foi demonstrado no teorema (2.1).

Exemplo 2.3.4 *Para todo alfabeto X, a álgebra $K\{X\}$ que tem como base o conjunto de todas as palavras*

$$x_{i_1} \cdots x_{i_n} \in X^*, n = 1, 2 \ldots .$$

e com multiplicação definida por

$$(x_{i_1} \cdots x_{i_m})(x_{j_1} \cdots x_{j_n}) = x_{i_1} \cdots x_{i_m} x_{j_1} \cdots x_{j_n};$$

com x_{i_k}, x_{j_l} em X, é livre na classe das álgebras unitárias associativas. Se considerarmos o subespaço de $K\{X\}$ gerado pelas palavras de comprimento maior ou igual a 1, nós obteremos uma álgebra livre não unitária associativa que é livre na classe das álgebras associativas.

Será útil agora darmos o conceito de variedades de álgebras.

Definição 2.3.5 *Uma classe de álgebras $\mathcal{C}$, não vazia, é chamada uma variedade se ela satisfaz os seguintes axiomas:*

1. *se $C \in \mathcal{C}$ e existe um homomorfismo de B em C injetor, então $B \in \mathcal{C}$;*
2. *se $C \in \mathcal{C}$ e existe um homomorfismo de C em B sobrejetor, então $B \in \mathcal{C}$ e;*
3. *se $C_\alpha \in \mathcal{C}$, com α em um conjunto de índices I, é uma família de álgebras, então $\prod_{\alpha \in I} C_\alpha \in \mathcal{C}$. Aqui $\prod$ é o produto direto (possivelmente infinito).*

Dada uma classe de álgebras $\mathcal{C}$, ela gera uma variedade $\overline{\mathcal{C}}$ construída da seguinte maneira: $B \in \overline{\mathcal{C}}$ se e somente se existem uma família de álgebras $C_\alpha \in \mathcal{C}$, um homomorfismo injetor de $C \in \mathcal{C}$ em $\prod C_\alpha$ e um homomorfismo sobrejetor de C em B. Vamos então a seguinte proposição:

Proposição 2.3.6 *A classe $\overline{\mathcal{P}}$ é uma variedade contendo a classe $\mathcal{P}$. Assim, se $\mathcal{C}$ é uma variedade contendo $\mathcal{P}$ então $\mathcal{C}$ contém $\overline{\mathcal{P}}$.*

Demonstração. É claro que $\overline{\mathcal{P}} \supseteq \mathcal{P}$ e se a variedade $\mathcal{C} \supseteq \mathcal{P}$ então $\mathcal{C} \supseteq \overline{\mathcal{P}}$. Resta apenas a verificação direta de que $\overline{\mathcal{P}}$ é uma variedade. ■

Na próxima definição, dada uma álgebra C, nós indicaremos por $\overline{C}$ a variedade que ela gera.

Definição 2.3.7 1. *Se $\overline{C} \supseteq \overline{B}$ nós diremos que B é uma especialização de C.*

2. *Se $\mathcal{V}$ é uma variedade e $\mathcal{V} = \overline{C}$ nós diremos que C é uma álgebra geradora em $\mathcal{V}$.*

3. *Se $\overline{C} = \overline{B}$ nós diremos que C é equivalente a B.*

Proposição 2.3.8 *Seja $U \subseteq K\{X\}$ um conjunto de polinômios. A classe $\mathcal{C}$ das álgebras A, tais que para todo homomorfismo $\varphi : K\{X\} \to A$ temos $\varphi(U) = 0 \in A$, é uma variedade.*

Demonstração. Devemos verificar os axiomas da definição de variedades de álgebras (2.3.5).

1. Seja $A \in \mathcal{C}$ e considere um homomorfismo $B \to A$ injetor. Se $\varphi : K\{X\} \to B$ é um homomorfismo, então a composição $\psi : K\{X\} \to B \to A$ é tal que $\psi(U) = 0$. Como $B \to A$ é injetor, teremos $\varphi(U) = 0$.

2. Seja $A \in \mathcal{C}$ e considere um homomorfismo $\phi : A \to B$ sobrejetor. Se $\widetilde{\varphi} : K\{X\} \to B$ é um homomorfismo, então nós podemos determinar um homomorfismo $\varphi : K\{X\} \to A$ pela composição $\varphi = \phi^{-1} \circ \widetilde{\varphi}$. Assim $\varphi(U) = 0$ implicará que $\widetilde{\varphi}(U) = 0$ pois, caso contrário, teríamos $p \in U$ tal que $\widetilde{\varphi}(p) = b \in B$ com $b \neq 0$. Vamos verificar que não existe tal p. Como ϕ é sobrejetor, existiria $c \in A$ tal que $\phi(c) = b$ e além disso, $c \neq 0$. Mas, desta forma $\varphi(p) = \phi^{-1} \circ \widetilde{\varphi}(p) = \phi^{-1}(\widetilde{\varphi}(p)) = \phi^{-1}(b) = c \neq 0$. Absurdo pois $\varphi(U) = 0$.

3. Se $C_\alpha \in \mathcal{C}$, $\alpha \in I$, é uma família de álgebras em $\mathcal{C}$, então considere o produto $\prod C_\alpha$ e um homomorfismo $\varphi : K\{X\} \to \prod C_\alpha$. Para cada projeção $\pi_\alpha : \prod C_\alpha \to C_\alpha$ temos que $\pi_\alpha(\varphi(U)) = 0$. Segue-se daí que $\varphi(U) = 0$ pois a interseção de $\pi_\alpha^{-1}(C_\alpha)$ é $0 \in \prod C_\alpha$.

Desta forma concluímos a demonstração. ■

Até agora, colocamos a definição de variedade de álgebras a partir de uma classe qualquer de álgebras. Mas, nosso interesse é mais específico. Pretendemos tratar da variedade das álgebras associativas satisfazendo identidades polinomiais. Sendo assim, segue-se a definição mais apropriada aos nossos interesses. Ela será dada tendo em vista a proposição (2.3.8).

Definição 2.3.9 1. *Seja $\{f_i(x_1, \ldots, x_{n_i}) \in K\{X\} \mid i \in I\}$ um conjunto de polinômios na álgebra associativa livre $K\{X\}$. A classe $\mathcal{C}$ de todas as álgebras satisfazendo as identidades polinomiais $f_i = 0, i \in I$, é chamada a variedade de álgebras associativas determinadas (ou definidas) pelo sistema de identidades polinomiais $\{f_i = 0 \mid i \in I\}$.*

2. *A variedade $\mathcal{V}_1$ é chamada uma sub-variedade de $\mathcal{V}$ se $\mathcal{V}_1 \subset \mathcal{V}$.*

3. *O conjunto $T(\mathcal{V})$ de todas as identidades polinomiais satisfeitas pela variedade $\mathcal{V}$, é chamado o T-ideal (ou ideal verbal) da variedade $\mathcal{V}$.*

4. *Diremos que o T-ideal $T(\mathcal{V})$ é gerado como um T-ideal pelo conjunto de identidades $\{f_i = 0 \mid i \in I\}$ que determinam a variedade $\mathcal{V}$.*

5. *Usaremos a notação $T(\mathcal{V}) = \langle f_i \mid i \in I \rangle^T$ e diremos que o conjunto $\{f_i = 0 \mid i \in I\}$ é uma base de identidades polinomiais para $\mathcal{V}$.*

6. *Os elementos de $T(\mathcal{V})$ são chamados conseqüências das identidades polinomiais que estão na base de $\mathcal{V}$.*

7. *Dada uma álgebra A qualquer, denotaremos por $T(A)$ o T-ideal das identidades polinomiais de A.*

Exemplo 2.3.10 1. *A classe de todas as álgebras comutativas é uma variedade definida pela identidade $[x_1, x_2] = 0$.*

2. *A classe de todas as álgebras associativas é uma variedade definida pelo conjunto vazio de identidades polinomiais.*

3. *A classe das álgebras A com o produto trivial $xy = 0$ para $x, y \in A$, é uma variedade.*

Definição 2.3.11 *Para um conjunto fixado Y, a álgebra $F_Y(\mathcal{V})$ na variedade $\mathcal{V}$, é chamada uma álgebra relativamente livre de $\mathcal{V}$ (ou uma álgebra $\mathcal{V}$-livre), se $F_Y(\mathcal{V})$ é livre na classe $\mathcal{V}$ e é livremente gerada por Y.*

Veremos agora a existência de álgebras relativamente livres e que duas dessas álgebras, com mesmo posto, são isomorfas. No que segue, pretendemos denotar a álgebra relativamente livre de posto $m = \mid Y \mid$ por $F_m(\mathcal{V})$. Se Y é um conjunto infinito enumerável usaremos a notação $F(\mathcal{V})$ no lugar de $F_\infty(\mathcal{V})$.

Proposição 2.3.12 *Sejam $\mathcal{V}$ uma variedade definida por $\{f_i = 0 \mid i \in I\}$, Y um conjunto qualquer e J o ideal de $K\{Y\}$ gerado $\{f_i(g_1, \ldots, g_{n_i}) \mid g_i \in K\{Y\}; i \in I\}$. Então:*

1. *A álgebra $F = K\{Y\}/J$ é uma álgebra relativamente livre na variedade $\mathcal{V}$ com conjunto de geradores livres $\overline{Y} = \{y + J \mid y \in Y\}$.*

2. *Duas álgebras relativamente livres de mesmo posto são isomorfas.*

Demonstração. 1. Vamos primeiro mostrar que $F \in \mathcal{V}$. Seja $f_i(x_1, \ldots, x_n)$ uma das identidades que determinam $\mathcal{V}$ e sejam $\overline{g_1}, \ldots, \overline{g_n}$ elementos arbitrários de F. Logo $\overline{g_j} = g_j + J$ com $g_j \in K\{X\}$. Então $f_i(g_1, \ldots, g_n) \in J$ e assim $f_i(\overline{g}_1, \ldots, \overline{g}_n) = 0$. Isto mostra que $f_i(x_1, \ldots, x_n) = 0$ é uma identidade polinomial de F. Portanto $F \in \mathcal{V}$.

Agora vamos mostrar que F é uma álgebra relativamente livre em $\mathcal{V}$, livremente gerada por $\overline{Y}$. Seja A uma álgebra qualquer em $\mathcal{V}$ e seja $\phi : \overline{Y} \to A$ uma função arbitrária. Definamos uma função $\theta : Y \to A$ por $\widetilde{\theta}(y) = \phi(\overline{y})$ e estendamos $\widetilde{\theta}$ a um homomorfismo $\theta : K\{Y\} \to A$. Isto é sempre possível pois $K\{X\}$ é uma álgebra associativa livre. Para

provarmos que ϕ pode ser estendida a um homomorfismo de F em A, basta mostrar que $J \subset \ker(\theta)$. Seja então $f \in J$, isto é,

$$f = \sum_{i \in I} u_i f_i(g_{i_1}, \ldots, g_{i_{n_i}}) v_i, \text{ onde } g_{i_j}, u_i, v_i \in K\{Y\}.$$

Para quaisquer $a_1, \ldots, a_{n_i} \in A$ o elemento $f_i(a_1, \ldots, a_{n_i})$ é igual a zero em A e isto implica que $\theta(f) = 0$, ou seja, $J \subset \ker(\theta)$. Concluímos que F é isomorfo a $F_{\overline{Y}}(\mathcal{V})$ que é a álgebra relativamente livre em $\mathcal{V}$, livremente gerada por $\overline{Y}$.

2. Sejam $Y = \{y_i \mid i \in I\}$ e $Z = \{z_i \mid i \in I\}$ tais que $\mid Y \mid = \mid Z \mid$ e sejam $F_Y(\mathcal{V})$ e $F_Z(\mathcal{V})$ as álgebras relativamente livres correspondentes. Como ambas são relativamente livres, podemos definir homomorfismos

$$\phi : F_Y(\mathcal{V}) \to F_Z(\mathcal{V}) \quad e \quad \psi : F_Z(\mathcal{V}) \to F_Y(\mathcal{V})$$

por $\phi(y_i) = z_i$ e $\psi(z_i) = y_i$. Visto que as composições $\psi \circ \phi$ e $\phi \circ \psi$ são as identidades em Y e Z respectivamente, obtemos que ϕ e ψ são isomorfismos. ■

Observação 2.3.13 *Segue-se da prova da proposição anterior que o T-ideal de $K\{X\}$ gerado por $\{f_i \mid i \in I\}$ consiste de todas as combinações lineares de $u_i f_i(g_{i_1}, \ldots, g_{i_{n_i}}) v_i$, onde $g_{i_j}, u_i, v_i \in K\{Y\}$.*

Teorema 2.2 *Existe uma correspondência biunívoca π entre os T-ideais de $K\{X\}$ e as variedades de álgebras associativas. π é uma correspondência de Galois, ou seja, para quaisquer dois T-ideais, T_1 e T_2, a inclusão $T_1 \subset T_2$ é equivalente a inclusão $\pi(T_1) \supset \pi(T_2)$.*

Demonstração. Para todo T-ideal T definiremos $\mathcal{V} = \pi(T)$ a variedade determinada pelas identidades polinomiais pertencentes a T. Esta correspondência é sobrejetiva pois para toda variedade temos um T-ideal. Sejam $T_1 \neq T_2$ dois T-ideais e $\pi(T_i) = \mathcal{V}_i$ para $i = 1, 2$. Então, existe um polinômio $f(x_1, \ldots, x_n)$ que está em $T_1 \setminus T_2$ (ou em $T_2 \setminus T_1$). Note que, $f(x_1, \ldots, x_n) = 0$ é uma identidade polinomial para $\mathcal{V}_1$ e não é uma identidade polinomial para a álgebra relativamente livre $F(\mathcal{V}_2) = K\{X\}/T_2 \in \mathcal{V}_2$. Logo $\mathcal{V}_1 \neq \mathcal{V}_2$ e π é injetiva.

Para vermos que π é uma correspondência de Galois, basta notar que $\mathcal{V}_1 \supset \mathcal{V}_2$ se, e somente se todas as identidades polinomiais de $\mathcal{V}_1$ são satisfeitas também por $\mathcal{V}_2$, ou seja, $T(\mathcal{V}_1) \subset T(\mathcal{V}_2)$ ■

Observação 2.3.14 *Se $\mathcal{V}_1 \subset \mathcal{V}_2$, então $T(\mathcal{V}_1) \supset T(\mathcal{V}_2)$ e nós podemos considerar as identidades polinomiais de $\mathcal{V}_1$ módulo $T(\mathcal{V}_2)$. Então se conhecemos as identidades polinomiais de $\mathcal{V}_2$ e queremos estudar as identidades polinomiais de $\mathcal{V}_1$, podemos trabalhar na álgebra relativamente livre $F(\mathcal{V}_2)$.*

Corolário 2.3.15 *Se $\mathcal{V}$ é uma variedade de álgebras, $K\{X\}$ é uma álgebra livre em um conjunto infinito de variáveis e $T(\mathcal{V})$ é o T-ideal associado a $\mathcal{V}$, então $K\{X\}/T(\mathcal{V})$ é geradora em $\mathcal{V}$, ou seja, $\overline{K\{X\}/T(\mathcal{V})} = \mathcal{V}$.*

Demonstração. Claramente $\overline{K\{X\}/T(\mathcal{V})} \subseteq \mathcal{V}$; por outro lado, estas duas variedades determinariam o mesmo T-ideal $T(\mathcal{V})$ em V se $\overline{K\{X\}/T(\mathcal{V})} = \mathcal{V}$. ■

Observação 2.3.16 *Sejam $\mathcal{V}$ uma variedade de álgebras, $V = K\{X\}$ uma álgebra livre e $T(\mathcal{V})$ o T-ideal associado a $\mathcal{V}$. Claro que $\overline{V/T(\mathcal{V})} \subseteq \mathcal{V}$; mas não necessariamente teremos $\overline{V/T(\mathcal{V})} = \mathcal{V}$ se o conjunto X de variáveis for finito. Se $|X| = 1$ temos que todas as álgebras na variedade são comutativas.*

Proposição 2.3.17 *Seja $\varphi : R \to R'$ um homomorfismo de anéis comutativos, então qualquer R'-álgebra pode ser considerada como uma R-álgebra.*

1. *Se A e B são R'-álgebras e se A é uma especialização de B na classe das R'-álgebras associativas, então A é também uma especialização de B na classe das R-álgebras associativas.*

2. *Se $U \subseteq R'\{X\}$ é um T-ideal, a pré-imagem de U em $R\{X\}$ também é um T-ideal.*

3. *Se A é uma R'-álgebra, o T-ideal de identidades polinomiais de A com coeficientes em R é a pré-imagem do T-ideal das identidades polinomiais de A com coeficientes em R'.*

Demonstração.

1. Se C e D são R'-álgebras e $\phi : C \to D$ é um homomorfismo, então C e D são R-álgebras e ϕ é um homomorfismo de R-álgebras pois a variedade das R-álgebras gerada por B contém a variedade das R'-álgebras gerada por B.

2. Considere o homomorfismo de álgebras $\widetilde{\varphi} : R\{X\} \to R'\{X\}$ induzido por φ. Seja $U \subseteq R'\{X\}$ um T-ideal e considere $\widetilde{\varphi}^{-1}(U)$. Se $\psi : R\{X\} \to R\{X\}$ é um endomorfismo, temos $a_i = \widetilde{\varphi} \circ \psi(x_i)$. Existe um único endomorfismo de R'-álgebras $\overline{\psi} : R'\{X\} \to R'\{X\}$ tal que $\overline{\psi}(x_i) = a_i$. Veja que o diagrama

$$\begin{array}{ccc} R\{X\} & \xrightarrow{\psi} & R\{X\} \\ \downarrow{\scriptstyle \widetilde{\varphi}} & & \downarrow{\scriptstyle \widetilde{\varphi}} \\ R'\{X\} & \xrightarrow{\overline{\psi}} & R'\{X\} \end{array}$$

   é comutativo. Como U é um T-ideal, $\overline{\psi}(U) \subseteq U$ pois

$$\widetilde{\varphi}(\psi(\widetilde{\varphi}^{-1}(U))) = \overline{\psi}\widetilde{\varphi}(\widetilde{\varphi}^{-1}(U)) = \overline{\psi}(U) \subseteq U,$$

   onde $\psi(\widetilde{\varphi}^{-1}(U)) \subseteq \widetilde{\varphi}^{-1}(U)$.

3. Se A é uma R'-álgebra e ela for considerada como uma R-álgebra via um homomorfismo $\varphi : R \to R'$, então qualquer homomorfismo $R'\{X\} \to A$ se decompõe unicamente como $R\{X\} \to R'\{X\} \to A$ pois $f \in R\{X\}$ anula-se em A se e somente se $\widetilde{\varphi}(f)$ é uma identidade polinomial de A.

Terminamos assim a demonstração. ■

Definição 2.3.18 1. *Uma variedade $\mathcal{V}$ de R-álgebras é uma variedade própria se ela não contém a classe de todas as R/I-álgebras para qualquer ideal $I \neq R$.*

2. *Um T-ideal U de $R\{X\}$ é um T-ideal próprio se ele não está contido em $I\{X\}$ para qualquer $I \neq R$.*

Proposição 2.3.19 *Se o conjunto das variáveis é infinito, então uma variedade $\mathcal{V}$ é própria se, e somente se seu correspondente T-ideal é próprio.*

Existe uma outra maneira de ver a condição 2 da definição (2.3.18). Se $f(x) \in R\{X\}$, podemos considerar o ideal gerado pelos coeficientes deste polinômio. Indicaremos este ideal por $c(f)$, o conteúdo de f.

Proposição 2.3.20 *Um T-ideal U é um ideal próprio se, e somente se existe um $f \in U$ com $c(f) = R$.*

Demonstração. Se existe $f \in U$ com $c(f) = R$, então U é próprio. Para a recíproca, veja que se U é próprio então os coeficientes dos elementos de U geram R. Sejam $f_1, \ldots, f_s \in U$ tais que seus coeficientes geram R. Escolhemos potências suficientemente grandes de uma variável, digamos x_1 e teremos $f = f_1 x_1^{h_1} + f_2 x_1^{h_2} + \cdots + f_s x_1^{h_s} \in U$ com os monômios $f_i x_1^{h_i}$ e $f_j x_1^{h_j}$ não similares para $i \neq j$. Segue-se que $c(f) = \sum_i c(f_i) = R$. ∎

Definição 2.3.21 *1. Uma álgebra A é chamada uma álgebra com identidade polinomial, abreviadamente PI-álgebra, se o T-ideal das identidades polinomiais de A é próprio.*

2. Uma identidade polinomial $f(x_1, \ldots, x_s) = 0$ de A é não trivial se $c(f)A \neq \{0\}$.

3. Uma identidade polinomial $f(x_1, \ldots, x_s) = 0$ de A é própria se $c(f)a = 0$, $a \in A$, implica $a = 0$.

Um resultado que nós podemos provar é que se uma álgebra A satisfaz uma identidade polinomial própria, então ela é uma PI-álgebra. Para o momento, vamos provar apenas um resultado auxiliar.

Proposição 2.3.22 *Se $f(x_1, \ldots, x_s) = 0$ é uma identidade própria para uma álgebra A, então ela também é uma identidade própria para a álgebra $R\{X\}/U$; onde U é o T-ideal das identidades de A.*

Demonstração. Como U é o T-ideal das identidades de A, e $f(x_1, \ldots, x_s) = 0$ é uma identidade polinomial de A então $f(x_1, \ldots, x_s) = 0$ é uma identidade polinomial de $R\{X\}/U$. Seja $g \in R\{X\}$ tal que $c(f)g \in U$; nós queremos mostrar que $g \in U$, ou seja, que g é uma identidade polinomial de A. Para isso vejamos que, se $a \in A$ então $c(f)g(a) = 0$ pois $c(f)g \in U$. Assim as hipóteses dadas sobre f implicam que $g(a) = 0$; segue-se que $g \in U$. ∎

Definição 2.3.23 *Dois monômios são ditos equivalentes se eles são compostos pelas mesmas variáveis. Essa característica determina uma relação de equivalência. Um polinômio é dito uniforme se todos os seus monômios são equivalentes.*

Dado qualquer polinômio $f(x_1, \ldots, x_s)$ nós sempre podemos agrupar monômios equivalentes e escrever, de forma única, $f(x_1, \ldots, x_s) = \sum f_i(x_1, \ldots, x_s)$ onde os $f_i(x_1, \ldots, x_s)$ são uniformes.

Proposição 2.3.24 *Sejam U um T-ideal, $f(x_1, \ldots, x_s) \in U$ e a decomposição canônica de $f(x_1, \ldots, x_s)$ em polinômios uniformes dada por $f(x_1, \ldots, x_s) = \sum_{i=1}^{k} f_i(x_1, \ldots, x_s)$. Então os $f_i(x_1, \ldots, x_s) \in U$ para todo i.*

Demonstração. Vamos denotar as variáveis que realmente aparecem em $f(x_1, \ldots, x_s)$ como $x_1, \ldots, x_h$. Faremos uma indução sobre k. Se $k = 1$ o polinômio é uniforme. Em outros casos, existe ao menos uma variável, digamos x_1, que aparece em algum dos $f_i(x_1, \ldots, x_s)$, digamos para $i = 1, 2, \ldots, r$, e não aparece nos outros; $i = r+1, r+2, \ldots, k$. Façamos $x_1 = 0$. Visto que U é um T-ideal, $f(0, x_2, \ldots, x_h) \in U$. Agora temos $f(0, x_2, \ldots, x_h) = \sum_{i=r+1}^{k} f_i(x_1, \ldots, x_r)$. Então

$$\sum_{i=1}^{r} f_i(x_1, \ldots, x_s) = f(x_1, \ldots, x_h) - f(0, x_2, \ldots, x_h)$$

e $\sum_{i=r+1}^{k} f_i(x_1, \ldots, x_s)$ pertence a U. Pela indução, suas componentes uniformes estão em U e a proposição está demonstrada. ■

Corolário 2.3.25 *Se uma álgebra A satisfaz uma identidade não trivial, então ela satisfaz uma identidade não trivial uniforme.*

De forma natural mostraríamos agora que se uma álgebra associativa A satisfaz uma identidade polinomial de grau n, então A também satisfaz uma identidade polinomial multilinear de grau n. Mas, como esse resultado será usado na demonstração do teorema de Kaplansky, ele será o objetivo da próxima seção.

Capítulo 3

Teorema de Kaplansky

3.1 Multi-linearização de Polinômios

Definição 3.1.1 *Sejam $p = p(x_1, \ldots, x_n)$ um polinômio em $K\{X\}$ e $y_1, \ldots, y_k$ em $X - \{x_1, \ldots, x_n\}$. Para cada $i = 1, \ldots, n$ definiremos o polinômio pL_i^k por*

$$pL_i^k(x_1, \ldots, x_{i-1}, y_1, \ldots, y_k, x_{i+1}, \ldots, x_n) = p(x_1, \ldots, x_{i-1}, y_1 + \cdots + y_k, x_{i+1}, \ldots, x_n) -$$

$$- \sum_{q=1}^{k} p(x_1, \ldots, x_{i-1}, y_1 + \cdots + \widehat{y_q} + \cdots + y_k, x_{i+1}, \ldots, x_n) +$$

$$+ \sum_{1 \leq q_1 < q_2 \leq k} p(x_1, \ldots, x_{i-1}, y_1 + \cdots + \widehat{y_{q_1}} + \cdots + \widehat{y_{q_2}} + \cdots + y_k, x_{i+1}, \ldots, x_n) + \cdots$$

$$\cdots + (-1)^{k-1} \sum_{1 \leq q_1 < q_2 < \ldots < q_{k-1} \leq k} p(x_1, \ldots, x_{i-1}, \widehat{y_{q_1}} + \widehat{y_{q_2}} + \cdots + \widehat{y_{q_{k-1}}} + y_{q_j}, x_{i+1}, \ldots, x_n)$$

onde $\widehat{y_q}$ indica que retiraremos essa parcela da soma e assim, na última somatória, um y_{q_j} será a parcela que vai restar após cada escolha dos $y_{q_1}, \ldots, y_{q_{k-1}}$. Poderíamos também ter escrito a última somatória como $\sum_{q=1}^{k} p(x_1, \ldots, x_{i-1}, y_q, x_{i+1}, \ldots, x_n)$. Chamaremos L_i^k de operador de linearização.

Proposição 3.1.2 *Seja A uma álgebra associativa sobre um corpo K. Sejam G e H subgrupos do grupo aditivo de A. Se para todo polinômio $p = p(x_1, \ldots, x_n) \in K\{X\}$ e quaisquer elementos $a_1, \ldots, a_n$ em G temos que $p(a_1, \ldots, a_n) \in H$, então para quaisquer $a_1, \ldots, a_{i-1}, b_1, \ldots, b_k, a_{i+1}, \ldots, a_n$ em G temos que $pL_i^k(a_1, \ldots, a_{i-1}, b_1, \ldots, b_k, a_{i+1}, \ldots, a_n)$ pertence a H. Em particular, se p é uma identidade polinomial de A, então pL_i^k também é.*

Demonstração. Sejam $a_1, \ldots, a_{i-1}, b_1, \ldots, b_k, a_{i+1}, \ldots, a_n \in G$, então temos que qualquer somatório $\sum$ dos b_j pertence a G. Logo, por hipótese,

$$p(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n) \in H$$

para qualquer somatório $\sum$ dos b_j. Como, pela definição (3.1.1),

$$pL_i^k(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n)$$

é uma soma com parcelas da forma

$$p(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n) \in H$$

temos que

$$pL_i^k(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n) \in H.$$

Particularmente, se p é uma identidade polinomial de A, então

$$p(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n) \in H = \{0\}$$

para todo $\sum$ dos b_j. Logo

$$pL_i^k(a_1, \ldots, a_{i-1}, \sum, a_{i+1}, \ldots, a_n) \in H = \{0\}$$

para todo $\sum$ dos b_j, ou seja, pL_i^k é uma identidade polinomial. ■

Lema 3.1.3 *Seja $g : A^n \to A$ uma função em n variáveis que está definida sobre uma K-álgebra A e é linear em cada argumento. Então para quaisquer $a_1, \ldots, a_k \in A$, onde $k \geq n$,*

$$g(a_1 + \cdots + a_k, \ldots, a_1 + \cdots + a_k) -$$

$$- \sum_{q=1}^{k} g(a_1 + \cdots + \widehat{a_q} + \cdots + a_k, \ldots, a_1 + \cdots + \widehat{a_q} + \cdots + a_k) +$$

$$+ \sum_{1 \leq q_1 < q_2 \leq k} g(a_1 + \cdots + \widehat{a_{q_1}} + \cdots + \widehat{a_{q_2}} \cdots + a_k, \ldots, a_1 + \cdots + \widehat{a_{q_1}} + \cdots + \widehat{a_{q_2}} \cdots + a_k) + \cdots$$

$$\cdots + (-1)^{k-1} \sum_{q=1}^{k} g(a_q, \ldots, a_q) =$$

$$= \begin{cases} \sum_{\sigma \in S_n} g(a_{\sigma(1)}, \ldots, a_{\sigma(n)}), & se \quad k = n \\ 0, & se \quad k > n. \end{cases}$$

Demonstração. Seja $k \geq n$. Usando a linearidade da função g, nós podemos remover todas as somas do argumento desta função. Então o lado esquerdo da igualdade a ser demonstrada fica representado como uma combinação linear de elementos da forma $g(a_{j_1}, \ldots, a_{j_n})$ com coeficientes inteiros. Agora, se existirem s diferentes índices entre os j_i com $s < k$ então a soma dos coeficientes para cada $g(a_{j_1}, \ldots, a_{j_n})$ onde isso ocorre é igual a soma alternada

$$1 - \binom{k-s}{1} + \binom{k-s}{2} - \cdots + (-1)^{k-s} \binom{k-s}{k-s} = (1-1)^{k-s} = 0$$

Devemos notar também que s não excede n. Além disso k, por hipótese, não é menor do que n, ou seja, $s \leq n \leq k$. Assim, $s \geq k$ só ocorrerá quando $s = n = k$. Neste caso, os coeficientes dos $g(a_{j_1}, \ldots, a_{j_n})$ são iguais a 1 e a soma destes é $\sum_{\sigma \in S_n} g(a_{\sigma(1)}, \ldots, a_{\sigma(n)})$. ■

Proposição 3.1.4 *Para quaisquer p e p' em que o operador L_i^k está definido, $(p+p')L_i^k = pL_i^k + p'L_i^k$. Se $p = p(x_1, \ldots, x_m)$ é um monômio de grau n em x_i e*

$$g = g(x_1, \ldots, x_{i-1}, y_1, \ldots, y_n, x_{i+1}, \ldots, x_m)$$

é um monômio linear em $y_1, \ldots, y_n \in X - \{x_1, \ldots, x_m\}$ tal que

$$p(x_1, \ldots, x_m) = g(x_1, \ldots, x_{i-1}, x_i, \ldots, x_i, x_{i+1}, \ldots, x_m),$$

então

$$pL_i^k(x_1, \ldots, x_{i-1}, z_1, \ldots, z_k, x_{i+1}, \ldots, x_m) =$$

$$= \begin{cases} \sum_{\sigma \in S_n} g(x_1, \ldots, x_{i-1}, z_{\sigma(1)}, \ldots, z_{\sigma(n)}, x_{i+1}, \ldots, x_m), & se \quad k = n \\ 0, & se \quad k > n. \end{cases}$$

Demonstração. A linearidade do operador L_i^k segue diretamente da definição e demonstra o primeiro resultado. Fixando então $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ nós podemos considerar o monômio g como uma função das n variáveis $y_1, \ldots, y_n$. Aplicando o lema (3.1.3) obteremos o segundo resultado desejado. ■

Veja que usando a proposição (3.1.4) nós podemos mostrar que se o grau de x_i em p for k_i, $i = 1, 2, \ldots, n$, então $pL_1^{k_1}L_2^{k_2}\cdots L_n^{k_n}$ é multi-linear.

Definição 3.1.5 *Se $p = p(x_1, \ldots, x_n)$ é um polinômio homogêneo do tipo $[k_1, \ldots, k_n]$, então o polinômio multi-linear $pL_1^{k_1}L_2^{k_2}\cdots L_n^{k_n}$ é chamado linearização completa de p.*

Teorema 3.1 *Sejam A uma álgebra (não necessariamente associativa) sobre um corpo K e $p = p(x_1, \ldots, x_n) \in K\{X\}$, $p \neq 0$, que se anula pela substituição de quaisquer elementos de um subgrupo H do grupo aditivo da álgebra A. Então a linearização completa de quaisquer de suas componentes homogêneas com grau maximal também se anula em H.*

Demonstração. Seja $p = p_1 + p_2 + \cdots + p_s$ a decomposição de p em uma soma de componentes homogêneas e também seja, sem perda de generalidade, o grau de p_1 o maximal. Veja que as componentes homogêneas são polinômios uniformes. Pela proposição (2.3.24), elas se anulam em H. Particularmente, a de grau maximal também se anula em H. Reenumerando as variáveis, se necessário, assumimos que a componente homogênea p_1 do polinômio p tem o tipo $[k_1, \ldots, k_n]$ onde $k_i \neq 0$ para $i = 1, 2, \ldots, n$. As demais componentes homogêneas, p_2 por exemplo, têm tipo $[m_1, \ldots, m_n]$, onde a desigualdade estrita $m_j < k_j$ vale para algum j. Desta forma, $m_j < k_j$, e pela proposição (3.1.4) temos $p_2L_1^{k_1}L_2^{k_2}\cdots L_n^{k_n} = 0$. Conseqüentemente,

$$p_1L_1^{k_1}L_2^{k_2}\cdots L_n^{k_n} = pL_1^{k_1}L_2^{k_2}\cdots L_n^{k_n}.$$

Pela proposição (3.1.2) nós concluímos que a linearização completa

$$p_1L_1^{k_1}L_2^{k_2}\cdots L_n^{k_n}$$

da componente p_1 anula-se em H. Isto prova o teorema. ■

Proposição 3.1.6 *Se uma álgebra associativa A satisfaz uma identidade polinomial de grau n, então A também satisfaz uma identidade polinomial multi-linear de grau n.*

Demonstração. Esta proposição é conseqüência direta do teorema (3.1). ■

3.2 Teorema Sobre a Densidade de Um Anel – Teorema de Jacobson

Definição 3.2.1 *Se R é um anel e M é um R-módulo então* $A(M) = \{x \in R \mid Mx = (0)\}$.

Lema 3.2.2 *$A(M)$ é um ideal bilateral de R. Além disso, M é um $R/A(M)$-módulo fiel.*

Demonstração. Que $A(M)$ é um ideal à direita de R decorre imediatamente da definição de R-módulo. Para vermos que também é um ideal à esquerda tomemos $r \in R$ e $a \in A(M)$, então $M(ra) = (Mr)a \subset Ma = (0)$. Logo $ra \in A(M)$. Isto prova que $A(M)$ é um ideal bilateral de R.

Vamos mostrar agora que $A(M)$ é um $R/A(M)$-módulo fiel. Para $m \in M$ e $r + A(M) \in R/A(M)$ temos $m(r + A(M)) = mr$ e se $r + A(M) = r' + A(M)$ então $r - r' \in A(M)$. Assim, $m(r - r') = 0$ para todo $m \in M$. Logo $mr = mr'$, ou seja $m(r + A(M)) = mr = mr' = m(r' + A(M))$, nos mostra que a ação de $R/A(M)$ em M está bem definida. A verificação de que isto define uma estrutura de um $R/A(M)$-módulo em M é simples verificação dos axiomas. Finalmente, para vermos que M é um $R/A(M)$-módulo fiel, observemos que se $m(r + A(M)) = 0$ para todo $m \in M$ então, $mr = 0$. Daí, $r \in A(M)$, ou seja, somente o zero de $R/A(M)$ anula todo M. ∎

Definição 3.2.3 *Se M é um R-módulo e $a \in R$, definiremos a aplicação T_a de M em M por $T_a(m) = am$ para todo $m \in M$. Como M é um R-módulo, T_a é um endomorfismo do grupo aditivo de M.*

Definição 3.2.4 *Diremos que $E(M)$ é o conjunto de todos os endomorfismos do grupo aditivo de M.*

Não é difícil ver que $E(M)$ é um anel.

Definição 3.2.5 *Chamaremos* $C(M) = \{\psi \in E(M) \mid T_a\psi = \psi T_a \forall a \in R\}$ *de anel centralizador de R em M.*

Teorema 3.2 (Lema de Schur) *Se M é um R-módulo irredutível então $C(M)$ é um anel com divisão.*

Demonstração. Precisamos apenas mostrar que todo elemento não nulo em $C(M)$ tem um inverso em $C(M)$. Veja que se $\theta \in C(M)$ e existe θ^{-1} então, como $\theta T_a = T_a\theta$ nós temos imediatamente que $\theta^{-1}T_a = T_a\theta^{-1}$, ou seja, $\theta^{-1} \in C(M)$. Vamos então provar a existência do θ^{-1} em $C(M)$. Suponha que $\theta \in C(M)$, $\theta \neq 0$, se $W = M\theta$ então para todo $r \in R$, $Wr = WT_r = (M\theta)T_r = (MT_r)\theta \subset M\theta = W$. Conseqüentemente W é um submódulo de M. Visto que $\theta \neq 0$, pela irredutibilidade de M, nós deduzimos que $W\theta = M$. Além disso $\ker\theta$ é um submódulo de M mas não é todo M, pois $\theta \neq 0$, daí temos que $\ker\theta = (0)$ e assim θ é um monomorfismo. Conclusão, θ é sobrejetivo e monomorfismo. Logo θ^{-1} existe em $E(M)$. ∎

Relembremos a definição:

Definição 3.2.6 *Um anel R é dito anel primitivo se ele tem um módulo fiel irredutível.*

Sejam R um anel primitivo e M um R-módulo fiel irredutível. Pelo lema de Schur (3.2), $C(M)$ é um anel com divisão. Podemos considerar M como um espaço vetorial à direita sobre $C(M)$, onde $m\alpha$, para $m \in M$ e $\alpha \in C(M)$, é interpretado como a ação de α como um elemento de $E(M)$ em m.

Definição 3.2.7 *Diremos que um anel R é um anel denso em M (ou R age densamente em M) se para todo n natural e para $v_1, \ldots, v_n$ em M que são linearmente independentes sobre $C(M)$ e para quaisquer n elementos $w_1, \ldots, w_n$ em M, existir um elemento $r \in R$ tal que $w_i = v_i r$ para $i = 1, 2, \ldots, n$.*

O próximo teorema foi demonstrado por Jacobson e Chevalley. A nossa demonstração segue o texto no livro de I. N. Herstein [13].

Teorema 3.3 (Teorema Sobre a Densidade) *Sejam R um anel primitivo e M um R-módulo fiel irredutível. Então R é um anel denso de transformações lineares em M sobre $C(M)$.*

Demonstração. Primeiro mostraremos que para provar o teorema é suficiente mostrar que dado um subespaço V de M sobre $C(M)$ com dimensão finita e um $m \in M$, $m \notin V$, então nós podemos determinar $r \in R$ com $Vr = (0)$ mas $mr \neq 0$. Veja que se podermos determinar tal r então $mrR \neq (0)$. Assim, pela irredutibilidade de M, $mrR = M$. Poderemos então determinar um $s \in R$ com mrs arbitrário e $Vrs = (0)$. Dados $v_1, \ldots, v_n \in M$ linearmente independentes sobre $C(M)$ e $w_1, \ldots, w_n \in M$, seja V_i o espaço gerado sobre $C(M)$ pelos v_j com $j \neq i$. Visto que $v_i \notin V_i$, nós poderemos determinar um $t_i \in R$ tal que $v_i t_i = w_i$ e ainda $V_i t_i = (0)$. Logo, se $t = t_1 + \cdots + t_n$ nós vemos que $v_i t = w_i$ para todo $i = 1, 2, \ldots, n$. Isto é exatamente a densidade de R em M.

Agora vamos provar o que nos propomos, que dado $V \subset M$, subespaço de dimensão finita sobre $C(M)$, e $m \in M$, $m \notin V$, então existe $r \in R$ tal que $Vr = (0)$ mas $mr \neq 0$. Faremos isso por indução sobre a dimensão de V sobre $C(M)$, ou seja, suporemos que existe tal r sempre que se considera um subespaço de M com dimensão sobre $C(M)$ menor do que a dimensão de V. Vamos então escrever $V = V_0 + wC(M)$ com $w \notin V_0$. Daí, $\dim V_0 = \dim V - 1$. Pela nossa hipótese de indução, definindo $A(V_0) = \{x \in R \mid V_0 x = (0)\}$, dado $y \in M$, $y \notin V_0$, existe $r \in A(V_0)$ tal que $yr \neq 0$. Em outras palavras, se $mA(V_0) = (0)$ então $m \in V_0$. Pelo lema (3.2.2), $A(V_0)$ é um ideal à direita de R e visto que $w \notin V_0$; $wA(V_0) \neq (0)$, então, como $wA(V_0)$ é um submódulo de M, $wA(V_0) = M$. Vamos supor por absurdo que sempre que for dado $m \in M$, $m \notin V$, e $Vr = (0)$ então $mr = 0$. Para mostrarmos que isso não é possível definiremos $\tau : M \to M$ por $\tau(x) = ma$, que faz sentido pois se $x \in M$ então $x = wa$ para algum $a \in A(V_0)$. Verificaremos ainda que τ está bem definida. Tomemos $x = 0$, então $0 = x = wa$, e assim a anula V_0, pois $a \in V_0$; e a anula w. Logo a anula V. Mas, estamos tomando por hipótese $ma = 0$. Isto mostra que $x = \tau(x) = ma = 0$, conseqüentemente τ está bem definida. Claramente $\tau \in E(M)$, mais ainda, se $x = wa$ com $a \in A(V_0)$ então para qualquer $r \in R$, visto que $ar \in A(V_0)$, $xr = (wa)r = w(ar)$ e então $\tau(xr) = m(ar) = (ma)r = \tau(x)r$. Isto nos mostra que τ está em $C(M)$. Assim, para $a \in A(V_0)$, $ma = \tau(wa) = \tau(w)a$ implica que $(m - \tau(w))a = 0$ para todo $a \in A(V_0)$. Pela hipótese de indução $(m - \tau(w)) \in V_0$, logo $m \in V_0 + \tau(w) \subset V_0 + wC(M) = V$. Com esta contradição a prova está terminada. ■

3.3 Teorema de Kaplansky

Teorema 3.4 *Se C é um anel comutativo com unidade então a álgebra $M_n(C)$ não satisfaz nenhuma identidade polinomial de grau menor do que $2n$.*

Demonstração. Já vimos, pela proposição (3.1.6), que poderemos considerar apenas identidades multi-lineares. Sendo assim, denote por $e_{i,j} \in M_n(C)$ a matriz elementar, com 1 na entrada i, j e zero nas outras. Então $e_{i,j}e_{k,l} = \delta_{j,k}e_{i,l}$ onde $\delta_{j,k}$ é o delta de Kronecker para a posição j, k. Considere as primeiras k matrizes entre as $2n-1$ matrizes elementares $e_{1,1}, e_{1,2}, e_{2,2}, e_{2,3}, e_{3,3}, \ldots, e_{n-1,n-1}, e_{n-1,n}, e_{n,n}$. Multiplicando-as consecutivamente na ordem em que elas aparecem nesta lista, obteremos uma matriz elementar $e_{i,j}$ não nula mas, em qualquer outra ordem seu produto será nulo. Assim os valores de um polinômio com grau menor do que $2n$ nestas k matrizes elementares são não nulos. Conseqüentemente, $M_n(C)$ não satisfaz nenhuma identidade polinomial de grau menor que $2n$. ■

Definição 3.3.1 *Usaremos $Hom_N(M, M)$ para indicar o conjunto dos endomorfismos do módulo M que deixam o sub-módulo N de M, fixo. Ou seja,*

$$Hom_N(M, M) = \{\sigma : M \to M \mid \sigma \text{ é um endomorfismo e } \sigma(a) = a \text{ para todo } a \in N\}.$$

Definição 3.3.2 *Uma álgebra A é dita central simples sobre um corpo K, se A é uma álgebra simples tendo K como seu centro.*

Teorema 3.5 (Kaplansky) *Seja A uma álgebra primitiva sobre um corpo L satisfazendo uma identidade polinomial f de grau d. Então A é uma álgebra central simples de dimensão n^2 sobre seu centro Z, com $2n \leqslant d$.*

Demonstração. Nós podemos assumir que f é multi-linear pela proposição (3.1.6). Seja M um módulo fiel irredutível, $C(M)$ o centralizador deste módulo e K o subcorpo maximal de $C(M)$. Nós identificaremos A com o anel dos endomorfismos de M que A induz e formaremos o anel

$$AK = \{\sum a_i k_i \mid a_i \in A \text{ e } k_i \in K\} \subseteq Hom_K(M, M).$$

O módulo M é um AK-módulo irredutível, pelo lema de Schur (3.2), seu centralizador é um anel com divisão. Ainda mais, seu centralizador é K. Uma demonstração disso encontra-se, por exemplo, em ([13] pp. 94). O anel AK satisfaz f visto que f é multi-linear. Também notamos que L é naturalmente imerso em K porque nós podemos pensar nos coeficientes de f em K. Nós afirmamos que $2\dim_K M \leq d$ e queremos provar que $AK \simeq Hom_K(M, M)$, i.é., AK é isomorfo a $Hom_K(M, M)$ e $\dim_Z A = \dim_K Hom_K(M, M) = n^2$ com $n = \dim_K M$ e assim $2n \leq d$. Assumamos então que W é subespaço de M com dimensão finita tal que $2\dim_K M > d$. Logo, pelo teorema de Jacobson, teorema (3.3), existe uma subálgebra T de AK que induz em W o anel completo dos endomorfismos $Hom_K(W, W)$ (ver a definição de anel denso). Como $Hom_K(W, W)$ satisfaz f, a álgebra T também satisfará f. Isto contradiz o teorema (3.4). ■

Capítulo 4

Produto Tensorial de PI-Álgebras

Antes de tratarmos dos teoremas de Regev e Amitsur—Levitzki, vamos relembrar alguns conceitos relativos ao produto tensorial e dar uma tratamento ao produto tensorial de PI-Álgebras. A prova do teorema de Regev [25], ou teorema do produto tensorial, foi dada por ele em 1972. A nossa apresentação seguiu o texto de Rowen [27].

4.1 Produto Tensorial de Módulos

Definição 4.1.1 *Seja C um anel comutativo. Sejam A e B dois C-módulos e P um grupo abeliano. Então $\psi : A \times B \to P$ é chamada bilinear (sobre C, ou C-linear) se para todo $a_i \in A$, $b_i \in B$ e $c \in C$ tivermos $\psi(a_1 + a_2, b_1) = \psi(a_1, b_1) + \psi(a_2, b_1)$, $\psi(a_1, b_1 + b_2) = \psi(a_1, b_1) + \psi(a_1, b_2)$ e $\psi(ca_1, b_1) = \psi(a_1, cb_1)$.*

Definição 4.1.2 *Seja*

$$\mathbb{Z}(A \times B) = \{ \sum_{(a,b) \in A \times B} \alpha_{(a,b)}(a, b) \mid \alpha_{(a,b)} \in \mathbb{Z} \text{ é quase sempre zero} \},$$

onde, quase sempre zero, quer dizer que $\alpha_{(a,b)} \neq 0$ para um conjunto finito de pares ordenados (a, b). Veja que $\mathbb{Z}(A \times B)$ é o $\mathbb{Z}$-módulo livremente gerado por $A \times B$ que denotaremos por $(AB)^+$. Definiremos então $A \otimes_C B = (AB)^+ / I$ onde I é o subgrupo de $\mathbb{Z}(A \times B)$ gerado por todos os elementos nas formas $(a_1 + a_2, b_1) - (a_1, b_1) - (a_2, b_1)$, $(a_1, b_1 + b_2,) - (a_1, b_1) - (a_1, b_2)$ e $(ca_1, b_1) - (a_1, cb_1)$ para todo $a_i \in A$, $b_i \in B$ e $c \in C$. Vamos escrever $a \otimes b$ para a imagem canônica de (a, b) em $A \otimes_C B$.

Poderemos escrever $A \otimes B$ em vez de $A \otimes_C B$ nos lugares onde não existir ambiguidade sobre C. Note que, assim construído $A \otimes B$ é único como grupo abeliano.

Proposição 4.1.3 *A aplicação canônica $A \times B \to A \otimes B$, dada por $(a, b) \to a \otimes b$, é bilinear. Além disso, para toda aplicação bilinear $\psi : A \times B \to P$ onde P é um grupo abeliano, existe um homomorfismo de grupos induzido, $\overline{\psi} : A \otimes B \to P$, tal que $\overline{\psi}(a \otimes b) = (a, b)$ para todo $a \in A$ e $b \in B$.*

Demonstração. Estendendo ψ a um homomorfismo de grupos $\psi : (AB)^+ \to P$ por $\psi(\sum(a_i, b_i)) = \sum \psi(a_i, b_i)$ teremos $\psi(I) = 0$. Isto prova o que queríamos. ■

Podemos caracterizar $A \otimes B$ pela propriedade dada na proposição anterior. Ela será usada, de forma implícita, na demonstração de outros resultados sobre produto tensorial.

Proposição 4.1.4 *Sejam A' e B' C-módulos e $\psi_1 : A \to A'$ e $\psi_2 : B \to B'$ homomorfismos de C-módulos. Então existe um homomorfismo de grupos, denotado por $\psi_1 \otimes \psi_2 : A \otimes B \to A' \otimes B'$ tal que $(\psi_1 \otimes \psi_2)(a \otimes b) = \psi_1(a) \otimes \psi_2(b)$ para todo $a \in A$ e $b \in B$.*

Demonstração. Defina $\psi : A \times B \to A' \otimes B'$ por $\psi(a, b) = \psi_1(a) \otimes \psi_2(b)$. Veja que ψ é bilinear, assim, pela proposição (4.1.3), ψ induz um homomorfismo de grupos $A \otimes B \to A' \otimes B'$ com as propriedades naturais. ■

Corolário 4.1.5 *Temos que $A \otimes B$ é um C-módulo com relação às operações dadas por $c\sum(a_i \otimes b_i) = \sum(ca_i \otimes b_i)$. Então, com as notações e afirmações como na proposição (4.1.4), $\psi_1 \otimes \psi_2$ é um homomorfismo de módulos.*

Demonstração. Dado $c \in C$, defina $\psi_c : A \to A$ por $\psi_c(a) = ca$. Agora defina $c\,(\sum(a_i \otimes b_i))$ por $(\psi_c \otimes 1)\,(\sum(a_i \otimes b_i)) = \sum_i(ca_i \otimes b_i)$. Esta operação é distributiva porque $\psi_c \otimes 1$ é um homomorfismo de grupos, assim $A \otimes B$ é um C-módulo. Então

$$(\psi_1 \otimes \psi_2)(c(a \otimes b)) = \psi_1(ca) \otimes \psi_2(b) = c\psi_1(a) \otimes \psi_2(b) = c(\psi_1 \otimes \psi_2)(a \otimes b),$$

mostrando que $\psi_1 \otimes \psi_2$ é um homomorfismo de módulos. ■

4.2 Produto Tensorial de Álgebras

Teorema 4.1 *Se A e B são C-álgebras, então $A \otimes B$ é uma C-álgebra com a multiplicação induzida por $(a_1 \otimes b_1)(a_2 \otimes b_2) = (a_1a_2 \otimes b_1b_2)$.*

Demonstração. Fixando $a_2 \in A$ e $b_2 \in B$, defina ψ_1 e ψ_2 por $\psi_1(a_1) = a_1a_2$ e $\psi_2(b_1) = b_1b_2$ para todo $a_1 \in A$ e $b_1 \in B$. Claramente ψ_1 e ψ_2 são homomorfismos de módulos, assim temos $\psi_1 \otimes \psi_2 : A \otimes B \to A \otimes B$, definido pela multiplicação à direita por $a_2 \otimes b_2$. Fazendo isso para todo $a_2 \in A$ e $b_2 \in B$, nós agora inverteremos o procedimento, fixando $x \in A \otimes B$ e definindo $\psi_x : A \times B \to A \otimes B$ por $\psi_x(a, b) = x(a \otimes b)$ para todo $a \in A$ e $b \in B$. Então ψ_x induz $\overline{\psi_x} : A \otimes B \to A \otimes B$ correspondendo a multiplicação à esquerda por x. Agora é fácil ver que $A \otimes B$ é uma C-álgebra. ■

Nós agora podemos retomar a proposição (4.1.4) num contexto mais apropriado.

Proposição 4.2.1 *Se A, A', B e B' são C-álgebras e $\psi_1 : A \to A'$ e $\psi_2 : B \to B'$ são homomorfismos de C-álgebras, então $\psi_1 \otimes \psi_2$ também é um homomorfismo de C-álgberas.*

Demonstração. Segue-se do corolário (4.1.5). ■

Proposição 4.2.2 *Existe um isomorfismo $A \otimes B \to B \otimes A$ dado por $\sum a_i \otimes b_i \mapsto \sum b_i \otimes a_i$.*

Demonstração. Defina a aplicação $\psi : A \times B \to B \otimes A$ por $\psi(a, b) = b \otimes a$; $\overline{\psi}$ é o isomorfismo desejado. ■

Proposição 4.2.3 *Se A_1 e A_2 são C_1-álgebras e A_2 e A_3 são C_2-álgebras, então existe um isomorfismo $(A_1 \otimes_{c_1} A_2) \otimes_{c_2} A_3 \to A_1 \otimes_{c_1} (A_2 \otimes_{c_2} A_3)$ tal que $(a_1 \otimes a_2) \otimes a_3 \to a_1 \otimes (a_2 \otimes a_3)$.*

Demonstração. Fixemos $a \in A_3$ e definamos a C_1-bilinear aplicação $\psi_a : A_1 \times A_2 \to A_1 \times A_2$ por $\psi_a(a_1, a_2) = a_1 \otimes (a_2 \otimes a)$. Teremos $\overline{\psi_a} : A_1 \otimes A_2 \to A_1 \otimes (A_2 \otimes A_3)$. Agora definamos a C_2-bilinear aplicação $\psi : (A_1 \otimes A_2) \times A_3 \to A_1 \otimes (A_2 \otimes A_3)$ por $\psi(\sum_i a_{1_i} \otimes a_{2_i}, a) = \overline{\psi_a}(\sum_i a_{1_i} \otimes a_{2_i})$; $\overline{\psi}$ é o isomorfismo desejado. ■

Para quaisquer $I_1 \subseteq A$ e $I_2 \subseteq B$, vamos usar a notação $I_1 \otimes I_2$ para o conjunto $\left\{\sum_j a_{1j} \otimes a_{2j} \mid a_{ij} \in I_i\right\}$. É bom que sejamos cuidadosos para não vermos isso como um produto tensorial de álgebras sem unidade, porque as duas noções não são as mesmas. Por exemplo: para $A = \mathbb{Z}$, $B = \mathbb{Z}/2\mathbb{Z}$ e $I = 2\mathbb{Z}$, teremos que $A \otimes_{\mathbb{Z}} B$ é isomorfo a B e $I \otimes B = 0$, mas, como álgebras sem 1, $I \otimes B \neq 0$.

Proposição 4.2.4 *Seja $\psi : A \to \overline{A}$ um homomorfismo. Então, sendo $\psi' = \psi \otimes 1 : A \otimes B \to \overline{A} \otimes B$, teremos* $\ker \psi' = \ker \psi \otimes B$.

Demonstração. Seja $I = \ker \psi \otimes B$. Claramente $I \subseteq \ker \psi'$, assim ψ' induz uma aplicação $\overline{\psi'} : (A \otimes B)/I \to \overline{A} \otimes B$ cuja inversa nós podemos construir agora. Existe uma aplicação bilinear $\overline{A} \otimes B \to (A \otimes B)/I$ dada por $(\overline{a}, b) \to (a \otimes b) + I$. Esta é o homomorfismo inverso ψ^{-1}. ■

Definição 4.2.5 *Uma álgebra A é dita uma extensão central de uma subálgebra A' de A se $A = Z(A)A'$.*

Teorema 4.2 *Se B é uma álgebra comutativa, então $A \otimes B$ é uma extensão central de $A \otimes 1$, implicando que $A \otimes B$ satisfaz as mesmas identidades polinômias que A.*

Demonstração. Vamos escrever $A \otimes 1$ para $\{a \otimes 1 \mid a \in A\} \subseteq A \otimes B$. Faremos do mesmo modo para $1 \otimes B$. Observe que $A \otimes 1$ e $1 \otimes B$ são C-subálgebras de $A \otimes B$. Além disso $[A \otimes 1, 1 \otimes B] = 0$, assim $A \otimes B$ é uma extensão de $A \otimes 1$ pois $(a \otimes 1)(1 \otimes b) - (1 \otimes b)(a \otimes 1) = 0$ implica que $(a \otimes 1)(1 \otimes b) = (1 \otimes b)(a \otimes 1) = (a \otimes b)$ para todo $a \in A$ e $b \in B$ e daí $A \otimes B = (1 \otimes B)(A \otimes 1)$ e $Z(A \otimes B) = (1 \otimes B)$. Teremos então homomorfismos canônicos de C-álgebras $A \to A \otimes 1$ e $B \to 1 \otimes B$ dados por $a \mapsto a \otimes 1$ e $b \mapsto 1 \otimes b$. Disto segue-se o teorema. ■

O próximo lema é na verdade um corolário da proposição (3.1.6)

Lema 4.2.6 *Seja A uma PI-álgebra. Seja C uma subálgebra central de A e seja L um anel comutativo contendo C. Então $A \otimes_C L$ é uma PI-álgebra.*

Demonstração. Pela proposição (3.1.6), A satisfaz alguma identidade polinomial multilinear, digamos $f(x_1, \ldots, x_n)$. Sejam $x_1, \ldots, x_n \in A \otimes_C L$. Escreva $x_i = \sum_j r_{i,j} \otimes l_{i,j}$ com $r_{i,j} \in A$ e $l_{i,j} \in L$. Então

$$f(x_1, \ldots, x_n) = \sum_{j_1, \ldots, j_n} f(r_{1,j_1}, \ldots, r_{n,j_n}) \otimes l_{1,j_1} \cdots l_{n,j_n} = 0$$

Isto conclui a demonstração do lema. ■

4.3 Exemplos de Produtos Tensoriais

Nesta seção, vamos buscar um pouco mais de familiaridade com a construção de produtos tensoriais. No que se segue consideraremos o anel R infinito.

Proposição 4.3.1 *Os anéis de polinômios $R[x]$ e $R \otimes_{\mathbb{Z}} \mathbb{Z}[x]$ são isomorfos.*

Demonstração. Defina a aplicação bilinear $\psi : R \times \mathbb{Z}[x] \to R[x]$ por $\psi\left(r \sum_{i=1}^{d} n_i x^i\right) = \sum_{i=1}^{d} n_i r x^i$ para $r \in R$ e $n_i \in \mathbb{Z}$. Então ψ induz um homomorfismo $\overline{\psi} : R \otimes_{\mathbb{Z}} \mathbb{Z}[x] \to R[x]$. Mas, claramente todo elemento de $R \otimes \mathbb{Z}[x]$ tem a forma $\sum_i r_i \otimes x^i$. Se $\overline{\psi}(\sum_i r_i \otimes x^i) = 0$ então $\sum_i r_i x^i = 0$, implicando que cada $r_i = 0$; disto segue-se que $\ker \overline{\psi} = 0$, assim $\overline{\psi}$ é um isomorfismo. ■

Recordemos o conceito de anel de frações; seja S um subconjunto de um anel R fechado para a multiplicação e $1 \in S$. Consideremos a relação de equivalência em $R \times S$ dada por: $(r_1, s_1) \sim (r_2, s_2)$ se, e somente se existir $s \in S$ tal que $(r_1 s_2 - r_2 s_1)s = 0$. Indicaremos por rs^{-1} a classe de equivalência de (r, s). O conjunto $R_S = \{rs^{-1} \mid r \in R, s \in S\}$ é um anel chamado anel de frações de R em relação a S.

Proposição 4.3.2 *Se S é um sub-monóide de $Z(R) - 0$, então R_S e $R \otimes_{Z(R)} Z(R)_S$ são isomorfos como anéis.*

Demonstração. Defina $\psi : R \times Z(R)_S \to R_S$ por $\psi(r, zs^{-1}) = rzs^{-1}$ para $r \in R$, $z \in Z(R)$ e $s \in S$. Então ψ induz um homomorfismo de anéis $\overline{\psi} : R \otimes Z(R)_S \to R_S$. Mas, qualquer elemento de $R \otimes Z(R)_S$ tem a forma $r \otimes 1s^{-1}$, $r \in R$ e $s \in S$. Se $r \otimes 1s^{-1} \in \ker \psi$ então $rs^{-1} = 0$, implicando que $s_1 r = 0$ para algum $s_1 \in S$; assim $0 = rs_1 \otimes (s_1 s)^{-1} = r \otimes s^{-1}$, provando que ψ é um isomorfismo. ■

Proposição 4.3.3 *Se R e R' são anéis comutativos infinitos e R é um R'-módulo, então $R \otimes_{R'} M_n(R')$ e $M_n(R)$ são isomorfos.*

Demonstração. Defina a R'-bilinear aplicação $\psi : R \times M_n(R') \to M_n(R)$ por

$$\psi\left(r, \sum_{i,j=1}^{n} \alpha_{ij} e_{ij}\right) = \sum_{i,j=1}^{n} \alpha_{ij} r e_{ij}, \text{ para } \alpha_{ij} \in R' \text{ e } r \in R.$$

Isto induz um homomorfismo $\overline{\psi}_n : R \otimes_{R'} M_n(R') \to M_n(R)$, cuja inversa é dada por $\sum_{i,j=1}^{n} r_{ij} e_{ij} \mapsto \sum_{i,j=1}^{n} r_{ij} \otimes e_{ij}$. Logo, $\overline{\psi}_n$ é o isomorfismo desejado. ■

Na próxima proposição vamos escrever Z no lugar de $Z(R)$, $End_{Z_S}(R_S)$ é o conjunto dos endomorfismos de R_S como Z_S-módulo, $End_Z(R)$ é o conjunto dos endomorfismos de R como Z-módulos e $\dim_Z R < \infty$ quer dizer que R é gerado sobre Z por um conjunto finito de elementos.

Proposição 4.3.4 *Se $\dim_Z R < \infty$ e S é um sub-monóide de Z, então $End_{Z_S}(R_S)$ é isomorfo a $(End_Z R) \otimes_Z Z_S$.*

Demonstração. Definiremos uma aplicação bilinear $(End_Z(R)) \times Z_S \to (End_{Z_S}(R_S))$ por $(\beta, z) \mapsto z\beta$ onde $z\beta$ é o homomorfismo que envia rs^{-1} a $\beta(r)zs^{-1}$. Desta forma, nós temos um homomorfismo $\psi : (End_Z(R)) \otimes Z_S \to End_Z(R_S)$. Por outro lado, dado β em $End_Z(R_S)$, suponha $R = \sum_{i=1}^{m} r_i Z$ para apropriados $r_i \in R$. Então $\beta(r_i) = x_i s^{-1}$ para apropriados $x_i \in R$, $1 \leq i \leq m$ e $s \in S$, assim $s\beta \in End_Z(R)$. Claramente a aplicação $\beta \mapsto s\beta \otimes s^{-1}$ é ψ^{-1}. Logo ψ é um isomorfismo. ■

O produto tensorial tem a seguinte conexão com extensões: se R é um anel qualquer tal que $R = AB$, onde A e B são anéis com $[A, B] = \{ab - ba \mid a \in A, b \in B\} = \{0\}$, então existe um homomorfismo $\psi : A \otimes B \to R$ dado por $\psi(\sum_i a_i \otimes b_i) = \sum_i a_i b_i$. Por essa razão, R tem as mesmas identidades polinomiais que $A \otimes B$.

4.4 O Teorema de Regev

Teorema 4.3 (Regev) *O produto tensorial de duas PI-álgebras é uma PI-álgebra.*

Este teorema, na realidade, está sendo tratado em todo o capítulo. Mas, aqui nós nos aproximaremos ainda mais deste resultado de Regev que abriu caminho para muitas investigações em PI-álgebras. Uma contribuição importante na demonstração desse teorema foi dada por Latyshev que melhorou a demonstração original, com base nas ideias de Regev.

Nós pretendemos dar a prova do teorema de Regev, reduzindo-o a uma afirmação combinatorial. Claramente é suficiente mostrar que o produto tensorial de quaisquer duas PI-álgebras relativamente livres é uma PI-álgebra ou equivalentemente, se I_1 e I_2 são T-ideais de $K\{X\}$, então temos que $(K\{X\}/I_1) \otimes_K (K\{X\}/I_2)$ é necessariamente uma PI-álgebra. Isto nos conduz a estudar T-ideais ou, mais precisamente, os polinômios multi-lineares nestes T-ideais.

Denotaremos por P_k o conjunto de todos os polinômios multi-lineares nas variáveis $x_1, \ldots, x_k$ em $K\{X\}$, isto é, $P_k = \{f(x_1, \ldots, x_k) \in K\{X\} \mid f \text{ é } multi-linear\}$. Claramente P_k é um espaço vetorial gerado por $\{x_{\pi_1} x_{\pi_2} \cdots x_{\pi_k}\}$ onde $\pi \in S_k$ e denotaremos $P_0 = K$. Observe que $\dim_K P_k = k!$, ou seja, a dimensão de P_k sobre K é $k!$.

Definição 4.4.1 *Se I é um T-ideal de $K\{X\}$, defina $I_k = I \cap P_k$ e $C_k(I) = \dim_K P_k/I_k$. Assim, $C_k(I)$ é chamado a k-ésima codimensão de I.*

Claro que $C_k(I) \leq k!$. Veja então, o quanto é importante quando temos $C_k(I)$ muito menor que $k!$.

Lema 4.4.2 (Regev) *Sejam K um corpo e $I^{(1)}$ e $I^{(2)}$ dois T-ideais de $K\{X\}$ tais que para algum n, $C_n(I^{(1)})C_n(I^{(2)}) < n!$. Então $R = (K\{X\}/I^{(1)}) \otimes_K (K\{X\}/I^{(2)})$ satisfaz uma identidade polinomial multi-linear de grau n com coeficientes em K.*

Demonstração. Para $j = 1, 2$ escreva $d_j = C_n(I^{(j)})$. Então existe um conjunto $\{h_i^j(x_1, \ldots, x_n) \mid 1 \leq i \leq d_j\}$ que gera $P_n/I^{(j)}$. Para cada $\pi \in S_n$, existe $m_{i,\pi}^{(j)}$ em K tal que $x_{\pi_1} \cdots x_{\pi_n} - \sum_{i=1}^{d_j} m_{i,\pi}^{(j)} h_i^{(j)} \in I_n^{(j)}$. Escreveremos formalmente $f(x_1, \ldots, x_n) = \sum_{\pi \in S_n} t_\pi x_{\pi_1} \cdots x_{\pi_n}$ e tentaremos encontrar elementos $t_\pi \in K$, não nulos, tais que f seja uma

identidade de R. Tomemos arbitrariamente $r_{1j}, \ldots, r_{nj}$ em $K\{X\}/I^{(j)}$. Então, para toda $\pi \in S_n$ temos

$$r_{\pi_1,j} \cdots r_{\pi_n,j} = \sum_{i=1}^{d_j} m_{i,\pi}^{(j)} h_i^{(j)}(r_{1j}, \ldots, r_{nj}),$$

assim

$$f(r_{11} \otimes r_{12}, \ldots, r_{n1} \otimes r_{n2}) = \sum_{\pi \in S_n} t_\pi (r_{\pi_1 1} \cdots r_{\pi_n 1}) \otimes (r_{\pi_1 2} \cdots r_{\pi_n 2})$$

$$= \sum_{\pi \in S_n} \sum_{i=1}^{d_1} \sum_{u=1}^{d_2} t_\pi m_{i,\pi}^{(1)} h_i^{(1)}(r_{11}, \ldots, r_{n1}) \otimes m_{u,\pi}^{(2)} h_u^{(2)}(r_{12}, \ldots, r_{n2})$$

$$= \sum_{i=1}^{d_1} \sum_{u=1}^{d_2} (\sum_{\pi \in S_n} t_\pi m_{i,\pi}^{(1)} m_{u,\pi}^{(2)}) h_i^{(1)}(r_{11}, \ldots, r_{n1}) \otimes h_u^{(2)}(r_{12}, \ldots, r_{n2}).$$

Logo, nós necessitamos somente determinar inteiros t_π relativamente primos tais que

$$\sum_{\pi \in S_n} t_\pi m_{i,\pi}^{(1)} m_{u,\pi}^{(2)} = 0$$

para $1 \leq i \leq d_1$ e $1 \leq u \leq d_2$. Mas, nós temos $d_1 d_2$ equações lineares e homogêneas formando um sistema em $n!$ indeterminadas, que obrigatoriamente tem uma solução não trivial porque, por hipótese, $d_1 d_2 < n!$. Tal solução pode ser escolhida de números racionais, e em seguida eliminamos o denominador comum. ■

Observação 4.4.3 *Se k é inteiro, existe $n \in \mathbb{N}$ com $n! > k^n$. Então nós podemos reduzir nossa demonstração a mostrar que, para um T-ideal I de $K\{X\}$, se $K\{X\}/I$ satisfaz uma identidade polinomial de grau d então, para algum número $m(d)$ independente de n, $C_n(I) \leq m(d)^n$ para todo n.*

Definição 4.4.4 *Dado $\pi \in S_n$, definiremos $\rho(\pi)$ como sendo o maior inteiro k, tal que existam $i_1 < i_2 < \ldots < i_k \in \{1, \ldots, n\}$ com $\pi(i_1) > \pi(i_2) > \ldots > \pi(i_k)$.*

Exemplo 4.4.5 *Se* $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 2 & 4 & 1 & 6 \end{pmatrix} \in S_6$ *então $\rho(\pi) = 4$ porque $1 < 2 < 3 < 5$ e $\pi(1) > \pi(2) > \pi(3) > \pi(5)$.*

Definição 4.4.6 *Para uma permutação $\pi \in S_n$ nós construiremos um par de tabelas $T_1(\pi)$ e $T_2(\pi)$ da seguinte maneira. Tomamos o primeiro elemento de $T_1(\pi)$, $t_{11} = 1$, e o primeiro elemento de $T_2(\pi)$ $u_{11} = \pi(1)$. Por indução, se t_{1j} existe, é o menor k tal que, $t_{1,j-1} < k \leq n$ e $\pi(k) > u_{i,j-1}$. Então, tomamos $u_{ij} = \pi(k)$. Onde não podermos determinar o próximo t_{ij}, por terem acabado os possíveis valores de k, partiremos para a segunda linha de $T_1(\pi)$ e $T_2(\pi)$ tomando t_{21} o menor $k \in \{1, 2, \ldots, n\}$ que não aparece na primeira linha de $T_1(\pi)$, $u_{21} = \pi(t_{21})$ e, se $j > 1$, então t_{2j} é o menor k que não está na primeira linha de $T_1(\pi)$ tal que $t_{2,j-1} < k \leq n$ e $\pi(k) > u_{2,j-1}$; tomamos $u_{2j} = \pi(k)$. Desta forma terminamos a segunda linha, continuamos com a terceira, e assim por diante. Estas tabelas são chamadas tabelas de Amtisur.*

Exemplo 4.4.7 *Considere a permutação* $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 5 & 1 & 2 & 6 \end{pmatrix} \in S_6$. *Vejamos a construção de* $T_1(\pi)$ *e* $T_2(\pi)$ *passo a passo. Tomamos* $t_{11} = 1$ *e assim* $u_{11} = \pi(t_{11}) = 4$. *Veja que* $\pi(2) = 3 < 4 = u_{11}$, $\pi(3) = 5 > 4 = u_{11}$. *Logo temos* $t_{12} = 3$ *e* $u_{12} = 5$. *Repetindo a idéia anterior,* $\pi(4) = 1 < 5 = u_{12}$, $\pi(5) = 2 < 5 = u_{12}$, $\pi(6) = 6 > 5 = u_{12}$. *Logo temos* $t_{13} = 6$ *e* $u_{13} = 6$. *Acabaram as opções, não podemos continuar o processo. Partiremos para a segunda linha tomando* $t_{21} = 2$ *que é o menor inteiro pertencente a* $\{1, 2, 3, 4, 5, 6\}$ *que não aparece na primeira linha. Assim,* $u_{21} = \pi(t_{21}) = \pi(2) = 3$. *Os outros inteiros que não aparecem na primeira linha são* 4 *e* 5. *Veja então que* $\pi(4) = 1 < 3 = u_{21}$ *e* $\pi(5) = 2 < 3 = u_{21}$. *Desta forma, essa linha para aqui. Continuamos na terceira linha,* $t_{31} = 4$ *e* $u_{31} = \pi(t_{31}) = 1$. *Visto que* $\pi(5) = 2 > 1 = u_{31}$ *completamos as terceiras linhas das tabelas* $T_1(\pi)$ *e* $T_2(\pi)$ *por* $t_{32} = 5$ *e* $u_{32} = \pi(t_{32}) = 2$. *Obtemos*

$$T_1(\pi) = \begin{pmatrix} 1 & 3 & 6 \\ 2 & & \\ 4 & 5 & \end{pmatrix} \text{ e } T_2(\pi) = \begin{pmatrix} 4 & 5 & 6 \\ 3 & & \\ 1 & 2 & \end{pmatrix}.$$

Para estudar $\rho(\pi)$, podemos usar a tabela de Amitsur para π. Observe que escrevendo duas tabelas da forma descrita na definição (4.4.6), quando terminadas, $T_1(\pi)$ e $T_2(\pi)$ têm n entradas cada.

Exemplo 4.4.8 *Seja* $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 5 & 3 & 2 & 6 \end{pmatrix} \in S_6$, *então* $\rho(\pi) = 3$ *e*

$$T_1(\pi) = \begin{pmatrix} 1 & 3 & 6 \\ 2 & 4 & \\ 5 & & \end{pmatrix} \text{ e } T_2(\pi) = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 3 & \\ 2 & & \end{pmatrix}.$$

Teorema 4.4 (Dilworth) (veja [6], [27]) $\rho(\pi)$ *é o número de linhas em* $T_1(\pi)$.

Demonstração. Seja d o número de linhas em $T_1(\pi)$. Se $i_1 < \ldots < i_k$ e $\pi(i_1) > \ldots > \pi(i_k)$ então $i_1, \ldots, i_k$ estão em linhas distintas, implicando que $\rho(\pi) \leq d$. Analogamente, construiremos uma seqüência $i_d, \ldots, i_1$, da seguinte forma: $i_d = t_{d_1}$ e indutivamente, dado i_{m+1} da linha $m+1$, tomemos $i_m = t_{mj}$ com j maximal, tal que $t_{mj} < i_{m+1}$. Note que $u_{mj} > \pi(i_{m+1})$ pois, se isto não ocorre, durante a construção das tabelas, nós deveríamos ter colocado i_{m+1} na linha m de $T_1(\pi)$. Assim $i_1 < \ldots < i_d$ e $\pi(i_1) > \ldots > \pi(i_d)$ implicando que $\rho(\pi) \geq d$. Então $\rho(\pi) = d$. ■

Antes de aplicarmos o teorema de Dilworth para concluirmos a demonstração do teorema do produto tensorial, observaremos a conexão existente com codimensões (veja 4.4.1).

Observação 4.4.9 *Observe que a restrição a* P_n *da ordem parcial do monóide* $M\{X\}$ *dada em (1.3.15) coincide com a ordem lexicográfica.*

Teorema 4.5 (Latyshev) *Se* I *é um* T*-ideal de* $C\{X\}$ *tal que* $C\{X\}/I$ *satisfaz uma identidade polinomial* g *de grau* d, *então* $C_n(I) \leq$ *a cardinalidade de* $\{\pi \in S_n \mid \rho(\pi) < d\}$.

Demonstração. Seja $S = \{X_{\tau(1)} \dots X_{\tau(n)} \mid \tau \in S_n \text{ e } \rho(\tau) < d\}$ e seja A o C-subespaço de $C\{X\}$ gerado por S. Nós teremos terminado se $A + I_n = V_n$, assim assuma que existe algum $\pi \in S_n$ tal que $h = X_{\pi(1)} \dots X_{\pi(n)} \notin A + I_n$. Tome π tal que h é o menor possível (sob a ordem dada na observação (1.3.15)). Visto que $\rho(\pi) \geq d$, existem $i_1 < \dots < i_d$ com $\pi(i_1) > \dots > \pi(i_d)$. Escrevamos $h = h_1 X_{\pi(i_1)} h_2 X_{\pi(i_2)} \dots h_2 X_{\pi(i_d)} h_{d+1}$ para alguns $h_i \in M\{X\}$. Então $h - h_1 g(X_{\pi(i_1)} h_2, X_{\pi(i_2)} h_3, \dots, X_{\pi(i_d)} h_{d+1})$ é uma soma de monômios de menor ordem e assim, por indução, está em $A + I_n$. Mas $h_1 g(X_{\pi(i_1)} h_2, \dots) \in I_n$, implicando que $h \in A + I_n$, contrariando à hipótese inicial. ■

Para concluirmos a prova do teorema do produto tensorial, nós necessitamos apenas avaliar a cardinalidade de $\{\pi \in S_n \mid \rho(\pi) < d\}$.

Observe que, para todo $\pi \in S_n$ tal que $\rho(\pi) < d$ podemos ver $T_1(\pi)$ (e $T_2(\pi)$) como funções de $\{1, \dots, n\}$ em $\{1, \dots, (d-1)\}$ (visto que os elementos de cada linha crescem). Como $T_1(\pi)$ e $T_2(\pi)$ determinam π, a cardinalidade de $\{\pi \in S_n \mid \rho(\pi) < d\}$ é menor do que $((d-1)^n)^2 = (d-1)^{2n}$. Em particular, no teorema (4.5), $C_n(I) < (d-1)^{2n}$.

Em vista da observação (4.4.3), nós concluímos a prova do teorema de Regev para o produto tensorial de PI-álgebras.

Capítulo 5

Teoremas de Amitsur

Lema 5.0.10 *Se R é um anel simples e seu centro $Z \neq \{0\}$, então Z é um corpo.*

Demonstração. Veja [13], pp. 46–47. ■

Quando quisermos nos referir a um anel com identidade polinomial diremos, resumidamente, PI-anel.

Teorema 5.1 (Kaplansky) *Se R é um PI-anel simples então R é isomorfo a $M_t(D)$ para algum anel com divisão D que tem dimisão finita sobre o centro de R.*

Demonstração. Como R é simples, o teorema sobre a densidade implica que existe um anel com divisão D tal que R é isomorfo a $M_t(D)$ para algum número natural t, ou para cada t existe um subanel de R que é uma cópia isomorfa de $M_t(D)$. Vamos considerar a última afirmação verdadeira e mostrar que isso leva a um absurdo.

Seja K o centro de D. Teremos que R tem um subanel isomorfo a $M_t(K)$. Pela proposição (3.1.6), R satisfaz uma identidade polinomial multi-linear f. Assim, $M_t(K)$ também satisfará f. Segue-se do teorema (3.4) que o grau de f é maior ou igual a $2t$. Visto que isso ocorrerá para todo t, temos uma contradição. Então R é isomorfo a $M_t(D)$. ■

Teorema 5.2 *Seja R um anel simples de dimensão finita sobre seu centro $Z \neq \{0\}$. Então a dimensão de R sobre Z é um quadrado perfeito,* $\dim_Z R = n^2$.

Demonstração. Pelo teorema anterior, R é isomorfo a $M_t(D)$ para algum anel com divisão D que tem dimensão finita sobre o centro de R. Seja K o subcorpo maximal de D. Então K contém Z e $R \otimes_Z K$ é isomorfo a $M_t(D) \otimes_Z K$. Denotaremos $\dim_K D = m$. Então $D \otimes_Z K$ é isomorfo a $M_m(K)$ (veja [13] pp. 94–96). Agora $\dim_K(D \otimes_Z K) = \dim_Z D$; mas, visto que $D \otimes_Z K$ é isomorfo a $M_m(K)$, temos que $\dim_K(D \otimes_Z K) = m^2 = \dim_Z D$. Conseqüentemente, a dimensão de $R \otimes_Z K$ é um quadrado perfeito. ■

Definição 5.0.11 *O inteiro n do teorema (5.2) é chamado grau de R.*

Definição 5.0.12 *Se D é um anel com divisão com centro Z, então um corpo $K \supset Z$ é dito um corpo de decomposição de D se $D \otimes_Z K$ é isomorfo a um anel de transformações lineares em um espaço vetorial sobre K.*

Lema 5.0.13 *Seja R um PI-anel simples satisfazendo uma identidade polinomial multi-linear de grau d. Então $d \geq 2n$, onde n é o grau de R.*

Demonstração. Note que pelo teorema de Kaplansky (3.5), R é de dimensão finita sobre seu centro e, desta forma, o teorema (5.2) pode ser aplicado. Seja K o corpo de decomposição de R (Definição (5.0.12)). Então $R \otimes_Z K$ e $M_n(K)$ são isomorfos como anéis. Pelo lema (4.2.6), $M_n(K)$ também satisfaz a identidade polinomial multi-linear de R. Mas, pelo teorema (3.4), $M_n(K)$ não satisfaz nenhuma identidade multi-linear de grau menor do que $2n$. Disto segue-se que $d \geq 2n$. ■

Teorema 5.3 (Amitsur) *Seja R um PI-anel semiprimo. Seja t uma indeterminada central sobre R, denotamos por $R[t]$ o anel dos polinômios com coeficientes em R na variável t. Então, a interseção de todos os ideais maximais do anel $R[t]$ é zero.*

Demonstração. Veja (em [13] pp. 150) que $R[t]$ é semi-primitivo, isto é, a interseção de todos os ideais primitivos é zero. Segue-se que a interseção de todos os ideais maximais de $R[t]$ é zero, pois todo ideal primitivo está contido em algum ideal maximal. ■

Teorema 5.4 (Amitsur) *Seja R um PI-anel semi-primo. Então existem um inteiro n e um anel comutativo C, tais que R é isomorfo a um subanel de $M_n(C)$. Todo PI-anel semi-primo é obtido como subanel do anel das matrizes sobre um anel comutativo.*

Demonstração. O anel R tem uma cópia isomorfa no anel dos polinômios $R[t]$. Pelo teorema (5.3), a interseção de todos os ideais maximais de $R[t]$ é zero. Trocando R pelo anel polinomial $R[t]$, nós podemos assumir que a interseção de todos os ideais maximais é zero.

Digamos que R satisfaz uma identidade polinomial multi-linear de grau d. Denote por $\{M_\alpha\}$ a família de todos os ideais maximais de R. Fixe um ideal maximal $M_\alpha \in \{M_\alpha\}$. Como vimos no teorema (5.2) e no lema (5.0.13), os PI-anéis R/M_α têm uma cópia isomorfa em alguma álgebra de matrizes $M_{n_\alpha}(C_\alpha)$, onde C_α é algum corpo e $d \geq 2n_\alpha$. Note que o núcleo do homomorfismo $R \to R/M_\alpha \subseteq M_{n_\alpha}(C_\alpha)$ é o ideal M_α. Os homomorfismos $R \to M_{n_\alpha}(C_\alpha)$ induzem o homomorfismo $\Phi : R \to \prod_\alpha M_{n_\alpha}(C_\alpha)$. Visto que o núcleo de $R \to M_{n_\alpha}(C_\alpha)$ é justamente M_α e a interseção de todos os ideais maximais é zero, segue-se que Φ é injetiva. Escolha um número maximal n_0 entre os n_α, seja $n = n_0!$. Note que as matrizes $m \times m$ têm uma cópia isomorfa nas matrizes $n \times n$, basta considerarmos as matrizes $n \times n$ cujas linhas e colunas, após a m-ésima, são zeros. Por exemplo, matrizes 2×2 podem ser imersas em matrizes 4×4 via o homomorfismo

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Isto mostra que $M_{n_\alpha}(C_\alpha)$ tem uma cópia isomorfa em $M_n(C_\alpha)$. Conclusão, R tem uma cópia isomorfa em $\prod_\alpha M_n(C_\alpha)$ que é isomorfa a $M_n(\prod_\alpha C_\alpha) = M_n(C)$, onde $C = \prod_\alpha C_\alpha$ é um anel comutativo. Ou, ainda mais precisamente, um produto de corpos. ■

Capítulo 6

Teorema de Amitsur e Levitzki

Temos seis provas para o teorema de Amitsur e Levitzki, a original dada em 1950 [1], a de Kostant em 1958 [18], Swan em 1963 [32] e [33], Razmyslov em 1974 [23] ou [8], Rosset em 1976 [26] e, Szigeti, Tuza e Révész em 1993 [34].

Nós escolhemos para apresentar por sua simplicidade e utilidade devido ao uso de conceitos, técnicas e resultados interessantes para a teoria das PI-álgebras, a prova de Rosset.

Lema 6.0.14 *Se s_{2k} é uma identidade polinomial de $M_k(\mathbb{Q})$ então também é de $M_k(K)$ onde K é um corpo qualquer.*

Demonstração. Como s_{2k} é multi-linear, basta mostrar que s_{2k} se anula quando calculado sobre matrizes pertencentes a uma base de $M_k(K)$. Escolhemos a base usual de $M_k(K)$ formada pelas matrizes e_{ij}. Então podemos considerar somente as matrizes de $M_k(K_0)$ onde K_0 é o subcorpo primo de K. Se a característica de K for zero o corpo primo de K é isomorfo a $\mathbb{Q}$ e se a característica de K for um número primo p, tal corpo primo será isomorfo a $\mathbb{Z}_p$. Como $\mathbb{Z}_p$ é um quociente de $\mathbb{Z}$ e $\mathbb{Z} \subset \mathbb{Q}$, basta considerar K de característica zero, e $K = \mathbb{Q}$. Sejam $r_p = \sum_{i,j=1}^{k} \alpha_{ij}^{(p)}) e_{ij}$ onde $\alpha_{ij}^{(p)} \in K$ e $p = 1, \ldots, 2k$, matrizes em $M_k(K)$. Temos

$$s_{2k}(r_1, \ldots, r_{2k}) = \sum_{\sigma \in S_{2k}} (-1)^{\sigma} r_{\sigma(1)} \cdots r_{\sigma(2k)} = \sum_{\sigma \in S_{2k}} (-1)^{\sigma} (\sum_{i,j}^{k} \alpha_{ij}^{(\sigma(1))} e_{ij} \cdots \sum_{i,j}^{k} \alpha_{ij}^{(\sigma(2k))} e_{ij}).$$

Mas, veja que

$$e_{ij} e_{kl} = \begin{cases} e_{il}, & \text{se} \quad j = k \\ 0, & \text{se} \quad j \neq k. \end{cases}$$

Logo $s_{2k}(r_1, \ldots, r_{2k})$ é uma combinação linear de $s_{2k}(e_{i_1 j_1}, \ldots, e_{i_{2k} j_{2k}})$ que é zero pois, por hipótese, s_{2k} é uma identidade de $M_k(\mathbb{Z}) \subset M_k(\mathbb{Q})$. ■

Para a prova de Rosset é necessário conhecermos um pouco sobre a álgebra de Grassmann.

Definição 6.0.15 *Sejam K um corpo e n um número natural. A álgebra de Grassmann (ou álgebra exterior) E de um espaço vetorial com dimensão n sobre K é definida como o quociente $E = K\{X_k\}/I$ onde I é o ideal de $K\{X_k\}$ gerado por todos os elementos da forma $(\alpha_1 x_1 + \cdots + \alpha_n x_n)^2$ com $\alpha_i \in K$ e $X_k = \{x_1, \ldots, x_k\}$ é um conjunto ordenado finito. Denotamos $1 = 1 + I$, $e_i = x_i + I$ para $1 \leq i \leq n$ e a multiplicação de dois elementos em E por $a \cdot b$.*

Proposição 6.0.16 *Na álgebra de Grassmann de um espaço vetorial V de dimensão n sobre um corpo K, valem as seguintes propriedades:*

(i) $e_i \cdot e_i = 0$ *para qualquer* i *em* $\{1, \ldots, n\}$;

(ii) $e_i \cdot e_j = -e_j \cdot e_i$ *para quaisquer* i *e* j *em* $\{1, \ldots, n\}$;

(iii) $B = \{e_{i_1} \ldots e_{i_k} : 1 \leq k \leq n \text{ e } 1 \leq i_1 < \ldots < i_k \leq n\} \cup \{1\}$ *é uma base de* E. *Nos próximos ítens, por* $k = 0$, *entenderemos* $e_{i_1} \ldots e_{i_k} = 1$;

(iv) para cada k *em* $\{0, 1, \ldots, n\}$, *definiremos* E_k *como o subespaço de* E *gerado por* $\{e_{i_1} \ldots e_{i_k} \mid 1 \leq i_1 < \ldots < i_k \leq n\}$. *Então* $E = E_0 \oplus E_1 \oplus \cdots \oplus E_n$;

(v) para cada $k \in \{0, 1, \ldots, n\}$, $\dim E_k = \binom{n}{k}$ *e* $\dim E = 2^n$;

(vi) E_0 *é isomorfo a* K *como álgebras e* E_1 *é isomorfo a* V *como espaços vetoriais;*

(vii) o centro $Z(E) = \{a \in E \mid a.b = b.a \text{ para todo } b \in E\}$, *pode ser escrito como* $Z(E) = E_0 \oplus E_2 \oplus E_4 \oplus \cdots \oplus E_m$ *onde* $m = n$ *se* n *é par ou* $m = n - 1$ *se* n *é ímpar.*

Lema 6.0.17 *A álgebra de Grassmann* E *satisfaz a identidade polinomial*

$$[x_1, x_2, x_3] = [x_1, [x_2, x_3]] = 0.$$

Demonstração. Seja E a álgebra de Grassmann. Como $[x_1, [x_2, x_3]]$ é multi-linear, basta verificar que a identidade é satisfeita para uma base qualquer de E. Sejam $r_1 = e_{i_1} \ldots e_{i_l}$, $r_2 = e_{j_1} \ldots e_{j_m}$ e $r_3 = e_{k_1} \ldots e_{k_n}$ elementos quaisquer da base de E descrita acima. Assim, temos

$$[r_2, r_3] = r_2 r_3 - r_3 r_2 = r_2 r_3 - (-1)^{lm} r_2 r_3.$$

Logo, se lm é par, então $[r_2, r_3] = 0$, conseqüentemente $[r_1, r_2, r_3] = [r_1, [r_2, r_3]] = 0$ e não há mais nada a provar. Sendo assim, vamos considerar l e m ímpares e entao $[r_2, r_3]$ será não nulo, mas sempre será uma combinação linear de elementos da base de E com comprimentos pares. Logo, pertencerá ao centro e portanto comuta com r_1, ou seja, $[r_1, r_2, r_3] = [r_1, [r_2, r_3]] = 0$. ■

Lema 6.0.18 *Sejam* K *um corpo,* k *e* l *dois números naturais e* $r_1, \ldots, r_l$ *matrizes de* $M_k(K)$. *Se* l *é par então* $tr(s_l(r_1, \ldots, r_l)) = 0$.

Demonstração. Se r e s são matrizes de $M_k(K)$, é fácil ver que $tr(rs - sr) = 0$. Seja $\sigma \in S_l$, onde S_l é o grupo das permutações de l elementos, considere $\tau = \sigma \circ (123, \ldots, l) \in S_l$. As paridades de σ e τ são distintas pois como l é par, $(123, \ldots, l)$ é ímpar. Façamos então $r = r_{\sigma(1)}$ e $s = r_{\sigma(2)} \ldots r_{\sigma(l)}$. Logo, $sr = r_{\sigma(2)} \ldots r_{\sigma(l)} r_{\sigma(1)}$ e como $\tau(1) = [\sigma \circ (123 \ldots l)](1) = \sigma(2)$, $\tau(2) = [\sigma \circ (123 \ldots l)](2) = \sigma(3)$,..., $\tau(l-1) = \sigma(l)$, temos $sr = r_{\sigma(2)} \ldots r_{\sigma(l)} r_{\sigma(1)} = r_{\tau(1)} \ldots r_{\tau(l-1)} r_{\tau(l)}$. Assim $tr(sign(\sigma) r_{\sigma(1)} \ldots r_{\sigma(l)} + sign(\tau) r_{\tau(1)} \ldots r_{\tau(l)}) = tr(sign(\tau) rs + sign(\sigma) sr) = 0$ pois $sign(\sigma) = -sign(\tau)$ (Lembre-se que $sign(\sigma)$ é o sinal da permutação $\sigma \in S_l$). Desta forma $tr(s_l(r_1, \ldots, r_l)) = 0$. ■

Antes do próximo resultado, lembre-se das definições e notações de polinômio simétrico e soma de potências de grau k a n variáveis. Veja as definições (2.1.4) e (2.1.5).

Para simplificar usaremos apenas (e_k) representando $e_k(x_1, \ldots, x_n)$ e (p_k) para representar $p_k(x_1, \ldots, x_n)$.

Teorema 6.1 (Fórmulas de Newton) *Sejam K um corpo e, n e k números naturais. Então se $k \leq n$*

$$(p_k) - (p_{k-1})(e_1) + (p_{k-2})(e_2) + \cdots + (-1)^{k-1}(p_1)(e_{k-1}) + (-1)^k k(e_k) = 0$$

e se $k > n$ então

$$(p_k) - (p_{k-1})(e_1) + (p_{k-2})(e_2) + \cdots + (-1)^{n-1}(p_{k-n+1})(e_{n-1}) + (-1)^n(p_{k-n})(e_n) = 0.$$

Demonstração. Considere o polinômio

$$f(y) = (y-x_1)(y-x_2)\ldots(y-x_n) = y^n-(e_1)y^{n-1}+(e_2)y^{n-2}-\cdots+(-1)^{n-1}(e_{n-1})y+(-1)^n(e_n),$$

façamos $x_n = 0$ em ambos os membros desta equação e teremos

$$(y - x_1)(y - x_2)\ldots(y - x_{n-1})y = y^n - [e_1]_0 y^{n-1} + [e_2]_0 y^{n-2} - \cdots + (-1)^{n-1}[e_{n-1}]_0 y$$

onde cada $[e_i]_0(x_1, \ldots, x_n) = (e_i)(x_1, \ldots, x_{n-1}, 0)$. Simplificando por y temos

$$(y - x_1)(y - x_2)\ldots(y - x_{n-1}) = y^{n-1} - [e_1]_0 y^{n-2} + \cdots + (-1)^{n-2}[e_{n-2}]_0 y + (-1)^{n-1}[e_{n-1}]_0$$

e assim

$$(y - x_1)(y - x_2)\ldots(y - x_{n-1}) = y^{n-1} - (e_1)y^{n-2} + \cdots + (-1)^{n-2}(e_{n-2})y + (-1)^{n-1}(e_{n-1}).$$

Fazendo então $y = X_i$ temos,

$$X_i^{n-1} - (e_1)X_i^{n-2} + \cdots + (-1)^{n-2}(e_{n-2})X_i + (-1)^{n-1}(e_{n-1}) = 0,$$

ou ainda

$$X_i^n - (e_1)X_i^{n-1} + \cdots + (-1)^{n-1}(e_{n-1})X_i + (-1)^n(e_n) = 0.$$

Multiplicando por X_i^{k-n} para $k \geq n$, obtemos

$$X_i^k - (e_1)X_i^{k-1} + \cdots + (-1)^{n-1}(e_{n-1})X_i^{k-n+1} + (-1)^n(e_n)X_i^{k-n} = 0.$$

Somando as equações enquanto i percorre {1,2,...,n} obtemos a segunda fórmula do teorema,

$$(p_k) - (p_{k-1})(e_1) + (p_{k-2})(e_2) + \cdots + (-1)^{n-1}(p_{k-n+1})(e_{n-1}) + (-1)^n(p_{k-n})(e_n) = 0. \quad (6.1)$$

Para demonstrarmos a primeira fórmula, consideremos $r = n - k$ e examinemos o polinômio homogêneo simétrico

$$f_{k,n}(x_1, \ldots, x_n) = (p_k) - (p_{k-1})(e_1) + \cdots + (-1)^{k-1}(p_1)(e_{k-1}) + (-1)^k k(e_k)$$

de grau $k \leq n$. Veja que se $r = 0$ então $n = k$, assim, substituindo n por k em (6.1) e considerando que $p_0(x_1, \ldots, x_k) = k$ teremos que $f_{k,n} = 0$. Desta forma a primeira fórmula é verdadeira para $r = 0$. Tomemos $r > 0$ e consideremos, por hipótese de indução, que $f_{k,n} = 0$ para todo número natural menor que r. Novamente faremos $x_n = 0$ e observaremos que $[e_k]_0 = (e_k)$ e $[p_k]_0 = (p_k)$ onde $[p_k]_0 = (p_k)(x_1, \ldots, x_{n-1}, 0)$. Concluiremos assim que

$$f_{k,n}(x_1, \ldots, x_n, 0) = [p_k]_0 - [p_{k-1}]_0[e_1]_0 + \cdots + (-1)^{k-1}[p_1]_0[e_{k-1}]_0 + (-1)^k k[e_k]_0 =$$

$$= f_{k,n-1}(x_1, \ldots, x_{n-1}) = 0$$

pois $n - 1 - k = r - k < r$ incorre na hipótese de indução. A relação $f_{k,n}(x_1, \ldots, x_n, 0) = 0$ implica que $f_{k,n}$ é divisível por x_n, ou seja, $f_{k,n} = x_n f_1$. Como $f_{k,n}$ é simétrico obtemos

$$f_{k,n}(x_1, \ldots, x_n) = (e_n)(x_1, \ldots, x_n)g(x_1, \ldots, x_n),$$

que só é possível se $g = 0$ pois o grau de (e_n) é n e o grau de $f_{k,n}$ é k, logo é menor do que n. Portanto $f_{k,n} = 0$ e está provada a fórmula

$$(p_k) - (p_{k-1})(e_1) + (p_{k-2})(e_2) + \cdots + (-1)^{k-1}(p_1)(e_{k-1}) + (-1)^k k(e_k) = 0.$$

Isto demonstra o teorema. ■

Lema 6.0.19 *Sejam K um corpo de característica zero e B uma álgebra associativa e comutativa sobre K. Se $a \in M_n(B)$ e $tr(a^k) = 0$ para todo $k = 1, 2, \ldots, n$, então $a^n = 0$.*

Demonstração. Seja $a \in M_n(B)$ uma matriz com entradas $a_{i,j} \in B$. Considere a álgebra comutativa $K[a_{i,j} \mid i, j = 1, \ldots, n]$, ela é imagem homomorfa da álgebra dos polinômios comutativos $K[x_1, \ldots, x_{n^2}]$. Mas esta última álgebra é um subanel do corpo das funções racionais $K(x_1, \ldots, x_{n^2}) = L$. Se o lema for verdadeiro para matrizes com entradas no corpo L então ele será válido para matrizes com entradas na álgebra B. Assim basta demonstrar o lema somente no caso onde $B = L$ é um corpo de característica 0.

Sejam então $\lambda_1, \ldots, \lambda_n$ as raízes características de a no fecho algébrico de L. Logo, o polinômio característico de a é

$$p(x) = det(xI - a) = x^n - e_1(\lambda_1, \ldots, \lambda_n)x^{n-1} + e_2(\lambda_1, \ldots, \lambda_n)x^{n-2} + \ldots$$

$$\ldots + (-1)^{n-1}e_{n-1}(\lambda_1, \ldots, \lambda_n)x + (-1)^n e_n(\lambda_1, \ldots, \lambda_n).$$

Como as raízes da a^k são $\lambda_1^k, \ldots, \lambda_n^k$ e $tr(a^k) = 0$ para todo $k = 1, 2, \ldots, n$, temos a igualdade $p_k(\lambda_1, \ldots, \lambda_n) = 0$. Pelas fórmulas de Newton e pelo fato da característica de K ser zero, obtemos que $e_i(\lambda_1, \ldots, \lambda_n) = 0$ para todo $i \in \{1, 2, \ldots, n\}$. Assim $p(x) = x^n$ e pelo teorema de Cayley-Hamilton, temos $a^n = 0$. ■

Teorema 6.2 (Amitsur e Levitzki) *Sejam K um corpo e k um número natural. Então a álgebra $M_k(K)$ das matrizes $k \times k$ satisfaz a identidade standard de grau $2k$*

$$s_{2k}(x_1, \ldots, x_{2k}) = 0.$$

Demonstração. Pelo lema (6.0.14) podemos considerar $K = \mathbb{Q}$. Seja E a álgebra de Grassmann de um espaço vetorial sobre K gerado por $e_1, e_2, \ldots, e_{2k}$ e seja $E = E^0 \oplus E^1$ a $\mathbb{Z}_2$-graduação de E. Ou seja, E^0 e E^1 são gerados pelos produtos $e_{i_1} \ldots e_{i_l}$ de comprimentos pares e ímpares respectivamente. Pela prova do lema (6.0.17) temos que E^0 é uma subálgebra comutativa de E. Sejam então $r_1, r_2, \ldots, r_{2k}$ matrizes em $M_k(\mathbb{Q})$ e teremos que

$$b = r_1 e_1 + \ldots + r_{2k} e_{2k}$$

é uma matriz $k \times k$ com entradas da álgebra de Grassmann que é não comutativa. Mas, visto que $e_i \cdot e_j = -e_j \cdot e_i$ temos

$$a = b^2 = \sum_{i=1}^{2k} \sum_{j=1}^{2k} r_i r_j e_i \cdot e_j = \sum_{i<j} (r_i r_j - r_j r_i)(e_i \cdot e_j).$$

Logo, a é uma matriz com entradas da álgebra E^0 que é comutativa. Veja então que

$$a^q = (b^2)^q = \sum_{1 \leq i_1 < \ldots < i_{2q} \leq 2k} s_{2q}(r_{i_1}, \ldots, r_{i_{2q}}) e_{i_1} \ldots e_{i_{2q}}$$

para todo q natural. Assim, $tr(a^q) = \sum_{1 \leq i_1 < \ldots < i_{2q} \leq 2k} tr(s_{2q}(r_{i_1}, \ldots, r_{i_{2q}})) e_{i_1} \ldots e_{i_{2q}}$. Pelo lema (6.0.18), $tr(s_{2q}(r_1, \ldots, r_{2q})) = 0$ para todos $r_1, \ldots, r_{2q}$ em $M_k(\mathbb{Q})$ e pelo lema (6.0.19),

$$a^k = s_{2k}(r_1, \ldots, r_{2k}) e_1 \ldots e_{2k} = 0.$$

Mas $e_1 \ldots e_{2k} \neq 0$ pois é um elemento de uma base de E. Portanto $s_{2k}(r_1, \ldots, r_{2k}) = 0$. ■

Capítulo 7

Polinômio Central e o Teorema de Posner

Não é muito simples exibir polinômios centrais. Que eles existem em abundância foi descoberto somente no começo da década de setenta (1972). O primeiro polinômio central para matrizes $n \times n$ foi encontrado por Formanek e Razmyslov [9] e [22]. Os polinômios centrais permitiram a descoberta de importantes resultados novos, e também deram novas e melhores provas para muitos teoremas anteriores sobre PI-álgebras.

7.1 Polinômio Central

Definição 7.1.1 *Seja R um PI-anel. Seja $f(x_1, \ldots, x_n) \in K\{X\}$, sem termo constante e não sendo identidade polinomial de R. Se para toda escolha de $r_1, \ldots, r_n \in R$, $f(r_1, \ldots, r_n)$ é um elemento do centro de R, então f é chamado um* **polinômio central de** *R.*

Exemplo 7.1.2 *Seja C um anel comutativo. Se x e y são matrizes 2×2 sobre C então $xy - yx$ é uma matriz com traço zero. Cálculos simples mostrarão que o quadrado de uma matriz com traço zero é uma matriz escalar, então estará no centro de $M_2(C)$. Assim o poliômio $p(x, y) = (xy - yx)^2$ é um polinômio central de $M_2(C)$. Este polinômio não tem termo constante e é fácil verificar que não é uma identidade polinomial de $M_2(C)$.*

Teorema 7.1 (Formanek, Razmyslov) *Para cada anel $M_n(C)$ de matrizes $n \times n$ sobre um anel comutativo C, existe um polinômio g_n que é central para $M_n(C)$. Pode-se escolher g_n multi-linear.*

Demonstração. Veja [8] pp. 91–97. ■

Lema 7.1.3 *Um polinômio central para $M_n(C)$ é uma identidade polinomial para $M_k(C)$ quando $k < n$.*

Demonstração. Seja $k < n$ e A o conjunto de todas as matrizes em $M_n(C)$ tais que as últimas $n - k$ linhas e colunas são zero. Nós podemos identificar A com $M_k(C)$. Se g é um polinômio central de $M_n(C)$, então os valores tomados por g em A são matrizes escalares

em $M_k(C)$, e estas matrizes escalares têm que ser nulas visto que as últimas $n-k$ linhas e colunas das matrizes nessa representação de A são formadas por zero. Lembre-se que o termo constante do polinômio central g é zero por definição. Assim g é uma identidade polinomial de $M_k(C)$. ■

Agora é fácil deduzir a existência de polinômios centrais para PI-anéis simples.

Teorema 7.2 *Todo PI-anel simples admite um polinômio central.*

Antes de demonstrarmos o teorema veja que se o grau de um anel R é n (veja 5.0.11), então g_n é um polinômio central de R. Além disso, g_k é uma identidade polinomial de R para todo $k > n$.

Demonstração. Pelo teorema de Kaplansky (3.5), R tem dimensão finita sobre seu centro Z. Seja K o corpo de decomposição de R, que existe pelo teorema (5.2). Então $R \otimes_Z K$ é isomorfo a $M_n(K)$. Pelo teorema (7.1) e pelo lema (7.1.3), g_k é uma identidade para $M_n(K)$ e assim para R se $k > n$. Claramente os valores de g_n obtidos sobre os elementos de R são centrais em R visto que eles são também centrais no anel maior $M_n(K)$. Resta-nos mostrar que g_n não é uma identidade polinomial de R. Suponha o contrário. Visto que g_n é multi-linear, ele é então, pelo lema (4.2.6) também uma identidade polinomial da extensão central $M_n(L)$ isomorfa a $R \otimes_Z L$. Isto é uma contradição pois, visto que g_n é um polinômio central para $M_n(L)$ então, ele não é uma identidade polinomial de $M_n(L)$. ■

7.2 Teorema de Posner

Nós vamos mostrar agora que o centro de um PI-anel semi-primo é "grande".

Teorema 7.3 (Rowen) *Seja R um PI-anel semi-primo. Então todo ideal bilateral, não nulo, I de R, tem interseção não nula com o centro Z de R, ou seja, $I \cap Z \neq 0$.*

Demonstração. Seja t uma indeterminada central sobre R. Considere $R[t]$ o anel dos polinômios com coeficientes em R. Visto que R é um PI-anel semi-primo, o teorema (5.3) nos diz que a interseção de todos os ideais maximais de $R[t]$ é zero. Pelo lema (4.2.6), $R[t]$ também é um PI-anel. Se I é um ideal não nulo de R então $I[t]$, o conjunto dos polinômios com coeficientes em I, é um ideal de $R[t]$. Suponha que $I[t]$ contenha um polinômio pertencente ao centro de $R[t]$. Uma verificação fácil nos mostra que os coeficientes desse polinômio estão no centro de R. Assim I contém elementos não nulos no centro de R. Logo para provar o teorema, nós podemos substituir R por $R[t]$. Então nós podemos assumir que a interseção de todos ideais maximais de R é zero.

Agora seja I um ideal não nulo de R. Se M é um ideal maximal de R, então R/M é simples. Desta forma, a imagem de I, por um homomorfismo, em R/M é zero ou todo o R/M. Como $I \neq 0$, essa imagem é não nula em algum R/M.

Pelo teorema (7.2), para cada ideal maximal M existe um inteiro n_M tal que g_{n_M} é um polinômio central para R/M. O anel R satisfaz uma identidade multi-linear de grau d. Os inteiros n_M são limitados superiormente. Logo concluiremos que R/M também satisfaz esta identidade polinomial, visto que, pelo lema (5.0.13), $d \geq 2n_M$.

Indicamos por n_0 o maior inteiro entre os n_M para os quais a imagem de I em R/M é não nula. Denote o ideal maximal escolhido, por M_0. Agora escolha qualquer elemento em I tal que o valor z obtido quando aplicamos g_{n_0} neste elemento tem imagem não nula em R/M_0. Claramente $z \neq 0$. Visto que o termo constante de g_{n_0} é zero, z pertence a I. Pretendemos mostrar que z está no centro de R.

Seja M um ideal maximal. Mostraremos primeiro que a imagem $\overline{z}$ de z em R/M é central. Se a imagem $\overline{I}$ de I é zero em R/M então $\overline{z} = 0$ e assim $\overline{z}$ é central. Agora assuma que $\overline{I} \neq 0$. Então pela maximalidade de n_0, g_{n_0} é um polinômio central de R/M ou, pelo teorema (7.2), uma identidade polinomial de R/M. Em qualquer caso, $\overline{z}$ é central em R/M.

Finalmente, seja x um elemento qualquer de R. Então, módulo todo ideal maximal M, $xz - zx = 0$ pois z é central em R/M. Assim $xz - zx$ pertence a todos os ideais maximais de R. Como a interseção de todos os ideais maximais de R é zero, segue-se que $xz - zx = 0$ e daí $xz = zx$. Portanto, z é central em R. ■

Antes do teorema de Posner, vamos dar um exemplo que mostra que nem todo PI-anel admite um polinômio central, e que o teorema de Rowen pode falhar para PI-anéis que não são semi-primos.

Exemplo 7.2.1 *Provaremos que existe um PI-anel R que tem as seguintes propriedades:*

1. *R tem um ideal I não nulo cuja interseção com o centro de R é zero e,*

2. *R não admite polinômio central.*

Demonstração. Seja C um anel comutativo. Seja R o anel das matrizes 2×2 triangulares superiores com entradas em C. É fácil verificar que o centro de R consiste de matrizes escalares e é assim isomorfo a C. Seja I o conjunto de todas as matrizes 2×2 tais que a segunda linha é zero. Assim

$$R = \begin{pmatrix} C & C \\ 0 & C \end{pmatrix} \text{ e } I = \begin{pmatrix} C & C \\ 0 & 0 \end{pmatrix}.$$

Não é difícil verificar que I é um ideal de R. Como a matriz nula é a única matriz escalar em I, a interseção de I com o centro de R é zero. Assim R verifica (6.1).

De acordo com o teorema de Rowen, R não pode ser semi-primo. O conjunto de todas as matrizes 2×2 triangulares estritamente superiores

$$J = \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}$$

formam um ideal que é nilpotente, ou seja, $J^2 = 0$.

Agora suponha que f é um polinômio central de R. Considere a aplicação $\varphi : C \to R$ dada por $\varphi(a) = ae_{2,2}$. Esta aplicação é um homomorfismo de anéis que não preserva 1 pois $\varphi(1_C) = e_{2,2} \neq 1_R$. Visto que os valores de f quando aplicado a $\varphi(C)$ são matrizes escalares pertencendo a $\varphi(C)$, segue-se que f anula-se em $\varphi(C)$. Assim f é uma identidade polinomial para C que é isomorfo a $\varphi(C)$. Como R/I também é isomorfo a C, temos que todos os valores de f quando aplicado em R pertencem a I. Logo estes valores são matrizes escalares pertencendo a I, ou seja, elas são zero. Conseqüentemente f é uma identidade

polinomial de R, que contradiz o fato de f ser um polinômio central de R. Isto prova o item (2). ■

Vamos ao teorema de Posner. Para isso definiremos e faremos algumas considerações sobre localização em álgebra comutativa.

Seja S um sub-monóide do monóide multiplicativo do anel comutativo R e seja M um R-módulo. Considere o conjunto $S \times M$ de pares (s, x); $s \in S$, $x \in M$. Defina $(s_1, x_1) \sim (s_2, x_2)$ se existir um $s \in S$ tal que $s(s_2x_1 - s_1x_2) = 0$. Isto é uma relação de equivalência. Denotaremos o conjunto quociente por M_S e a classe de equivalência de (s, x) por $s^{-1}x$ (Compare com (4.3)).

Definição 7.2.2 *Podemos definir em M_S as seguintes operações:*

$$s_1^{-1}x_1 + s_2^{-1}x_2 = (s_1s_2)^{-1}(s_2x_1 + s_1x_2) \quad e \quad r(s^1x) = s^{-1}(rx).$$

O quociente M_S com essas operações é um R-módulo chamado localização de M em S.

Temos naturalmente a função

$$\varphi_S : M \longrightarrow M_S \quad dada \quad por \quad \varphi_S(x) = 1^{-1}x.$$

Podemos verificar que é um homomorfismo.

O núcleo de φ_S é o conjunto dos x para os quais existem $s \in S$ tais que $sx = 0$, que é, $\{r \in R \mid rx = 0\} \cap S \neq \emptyset$.

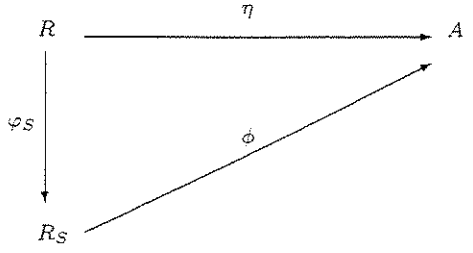
Em particular, podemos formar R_S. Definimos multiplicação em R_S por:

$$(s_1^{-1}r_1)(s_2^{-1}r_2) = (s_1s_2)^{-1}r_1r_2,$$

então R_S é uma R-álgebra comutativa. Neste caso, a função φ_S é um homomorfismo de álgebras e $\varphi_S(s)$ é invertível em R_S para todo $s \in S$ pois

$$(1^{-1}s)(s^{-1}1) = 1^{-1}1 = 1 \quad \text{em} \quad R_S.$$

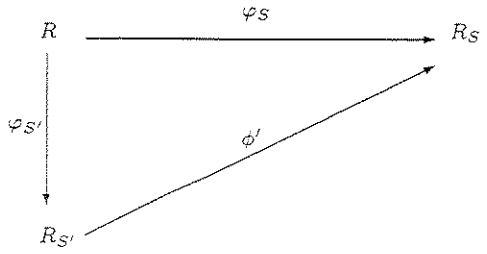
Sejam A uma R-algebra e η: $R \longrightarrow A$ um homomorfismo de álgebras. Entao η é invertível para cada $s \in S$. Logo existe um único homomorfismo ϕ tal que o diagrama



é comutativo. Entao o par (R_S, φ_S) é um objeto universal.

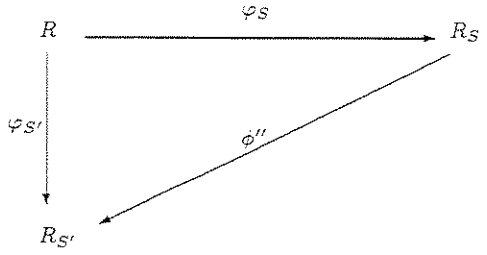
Desta forma, ϕ é dado por $\phi(s^{-1}r) = \eta(s)^{-1}\eta(r)$.

Se S' é um sub-monóide de S, então temos os homomorfismos φ_S de R em R_S e cada $\varphi_S(s')$, $s' \in S$, é invertível em R_S. Logo temos um único homomorfismo ϕ' de $R_{S'}$ em R_S tal que o diagrama



é comutativo.

Agora suponha que temos uma situação em que para cada $s \in S$, $\varphi_{S'}(s)$ é invertível em $R_{S'}$. Então temos um único homomorfismo ϕ'' de R_S em $R_{S'}$ tal que o diagrama



é comutativo.

Segue-se que $\phi''\phi' = I_{R_{S'}}$ e $\phi'\phi'' = I_{R_S}$, ou seja, as composições de ϕ' e ϕ'' são as respectivas identidades em $R_{S'}$ e R_S. Desta forma ϕ' é um isomorfismo de $R_{S'}$ em R_S.

Em vista disso, podemos dar uma definição alternativa para M_S em termos de R_S.

Lema 7.2.3 *A aplicação τ: $M_S \to R_S \otimes_R M$ definida por $\tau(s^{-1}x) = s^{-1}1 \otimes x$ é um isomorfismo.*

Demonstração. Notaremos primeiro que temos $t((st)^{-1}x) = (st)^{-1}(tx) = s^{-1}x$ em M_S e em particular isso vale em R_S com $x \in R$. Agora suponha $s_1^{-1}x_1 = s_2^{-1}x_2$. Então existe $s \in S$ tal que $s(s_2x_1 - s_1x_2) = 0$. Assim, $(ss_1s_2)^{-1}1 \otimes ss_2x_1 = ss_2((s_1s_2s)^{-1}1) \otimes x_1 = s_1^{-1}1 \otimes x_1$. Similarmente $(ss_1s_2)^{-1} \otimes ss_1x_2 = s_2^{-1}1 \otimes x_2$. Logo $s_1^{-1}1 \otimes x_1 = s_2^{-1}1 \otimes x_2$ e temos que τ está bem definida. Uma verificação direta mostra que τ é um homomorfismo. Suponhamos $s_1^{-1}r_1 = s_2^{-1}r_2$ para $s_i \in S$ e $r_i \in R$ com $i \in \{1, 2\}$. Então $s(s_2r_1 - s_1r_2) = 0$ para algum $s \in S$ e isto implica que $s_1^{-1}(r_1x) = s_2^{-1}(r_2x)$ vale em M_S para qualquer x em M. Logo temos a função $f : R_S \times M \to M_S$, dada por $f((s^{-1}r, x)) = s^1(rx)$; ela é aditiva em ambos os fatores do produto $R_S \times M$ e para $r' \in R$ $f((r's^{-1}r, x)) = f((s^{-1}r, r'x))$. Portanto temos um homomorfismo $\widetilde{\tau}$ de $R_S \otimes_R M$ em M_S que associa $s^{-1}1 \otimes rx = s^{-1}r \otimes x$ a $s^{-1}(rx)$. As composições dos dois homomorfismos são identidades. Logo ambos são isomorfismos. ■

É claro que τ é um isomorfismo de R-módulos. Também poderíamos considerar M_S como um R_S-módulo via $(s_1^{-1}r_1)(s^{-1}x) = (ss_1)^{-1}(r_1x)$.

Assim sendo, τ será também um isomorfismo de R_S-módulo de $R_S \otimes_R M$ em M_S.

Teorema 7.4 (Posner) *Seja R um PI-anel primo. Então o centro Z de R é um domínio. Denote por S o conjunto de elementos não nulos de Z. Então R_S é um PI-anel Artiniano simples. Além disso, R_S é isomorfo a $M_t(D)$ para algum anel com divisão D que tem dimensão finita sobre seu centro.*

O anel R_S é chamado anel total de frações de R (Compare com 4.3). Note que o centro de D é igual ao centro de R_S que é apenas Z_S, o corpo de frações do centro Z de R.

Demonstração. Como para localização em álgebra comutativa, verifica-se que os ideais de R_S estão em correspondência biunívoca com os ideais de R que tenham interseção nula com S. Mas, pelo teorema de Rowen (7.3), todos os ideais não nulos de R tem interseções não nulas com Z. Assim R_S é um anel simples. Novamente, como para localização em álgebra comutativa, R_S é isomorfo a $R \otimes_Z Z_S$. Então, pelo lema (4.2.6), R_S é um PI-anel. Logo, pelo teorema de Kaplansky (5.1), R_S é isomorfo a $M_t(D)$ para algum anel com divisão D que tem dimensão finita sobre seu centro. Conseqüentemente, R_S é espaço vetorial de dimensão finita sobre seu centro e, desta forma, Artiniano. Isto conclui a demonstração. ■

7.3 Exemplos de PI-anéis Primos

Vamos ver aqui alguns exemplos de PI-anéis primos e rever os resultados da seção anterior em algumas situações concretas.

Veja que o lema, a seguir, é um critério para verificar se um anel é primo.

Lema 7.3.1 *Seja S um subanel de um anel primo R. Se S contém um ideal bilateral não nulo I de R, então S é primo.*

Demonstração. Digamos que A e B são ideais de S tais que $AB = 0$. Visto que $I = RIR \subseteq S$ e $AI \subseteq AS = A$, segue-se que $ARIRB = 0$. Então um destes três ideais RAR, I e RBR do anel primo R tem que ser zero. Como, por hipótese, I é não nulo, temos que A ou B têm que ser zero. ■

Exemplo 7.3.2 *Ideais maximais, primos e semi-primos de $M_n(C)$.*
Seja C um anel comutativo e seja $R = M_n(C)$ o anel das matrizes $n \times n$ sobre C. Nós vemos C como subanel de R via imersão por matrizes escalares. Note que C é o centro de R. Os ideais de R são todos da forma $M_n(I')$ onde I' é um ideal de C. Para verificar isso, veja que $I' = I \cap C$. Assim seja I_0 o subconjunto de elementos de C que são entradas de matrizes pertencentes a I. Claramente $I \subseteq M_n(I_0)$. Usando a multiplicação de matrizes unidades (veja a demonstração do teorema (3.4)) verificamos a igualdade. Deduzimos que $I_0 = I \cap C$ e temos o que queríamos.

Segue-se que R é simples se seu centro é simples, ou seja, se C é um corpo. Se I e J são ideais de R, então $IJ = M_n(I \cap C).M_n(J \cap C) = M_n(I \cap C)(J \cap C)$. Assim R é primo se C é primo, ou seja, se C é um domínio. Igualmente, R é semi-primo se C também é, ou seja, se C é fatorial.

Transladando essas observações para ideais, nós vemos que o ideal I de R é maximal, primo ou semi-primo se o ideal $I \cap C$ de C também é. Finalmente, note que se C é um corpo, então R é simples, logo primo, que contém divisores de zero se $n \geq 2$.

Sabemos que todo PI-anel primo S é um subanel de um anel de matrizes $R = M_n(C)$ onde C é um anel comutativo. Nos concentraremos agora numa classe especial de PI-anéis. Para exemplificar, considere a matriz unidade $e_{i,j} \in R$. Se $s \in S$ é uma matriz com entradas $s_{i,j}$ então $e_{i,i}se_{j,j} = s_{i,j}e_{i,j}$ é a matriz com as mesmas entradas que s na i,j-ésima posição

e zero nas demais. Chamaremos S de componente-discreta se para todo elemento $s \in S$, S também contém todos os $e_{i,i}se_{j,j} = s_{i,j}e_{i,j}$. Nós já vimos um anel componente-discreta no exemplo (7.2.1). Vejamos mais alguns exemplos.

Exemplo 7.3.3 *Seja A um anel comutativo, e seja $A[t]$ o anel comutativo dos polinômios em uma variável sobre A. Denote por (t) o ideal de $A[t]$ gerado por t. Sejam A_1 e A_2 subanéis de A. Seja $R = M_2(A[t])$ o conjunto das matrizes 2×2 sobre $A[t]$, e seja*

$$S = \begin{pmatrix} A_1 + (t) & A[t] \\ (t) & A_2 + (t) \end{pmatrix}.$$

Aqui, S é um anel componente-discreta.

Demonstração. Uma matriz de R pertence a S se a entrada $(1,1)$ pertence a $A_1 + (t)$, a entrada $(1,2)$ pertence a $A[t]$ e assim por diante. Verifica-se facilmente que S é fechado para a multiplicação de matrizes e que S é um subgrupo aditivo de R. Assim S é um anel. Como um subanel de um PI-anel R, S também é um PI-anel. Considere $I = M_2((t)) = tR \neq 0$. Do exemplo (7.3.2) temos que I é um ideal de R e R é primo. Visto que I está contido em S e é um ideal bilateral não nulo de S, pelo lema (7.3.1), temos que S é primo. Assim S é um PI-anel primo.

Vamos ver quem é o centro Z de S. Se uma matriz em S comuta com todas as outras matrizes em S então ela é uma matriz escalar. Assim, Z é o conjunto de todas as matrizes escalares em S, ou seja, $Z = (A_1 \cap A_2) + (t) = (A_1 \cap A_2) + tA[t]$. ■

Exemplo 7.3.4 *Vamos verificar o teorema de Rowen e Posner neste exemplo.*
Seja $x = (a_{i,j})$ um elemento não nulo de um ideal I de S, onde S é o mesmo anel do exemplo anterior. Digamos que $a_{1,2}$, a entrada na posição $(1,2)$ de x é não nula. Então I contém a matriz escalar não nula $e_{1,1}x(te_{2,1}) + (te_{2,1})xe_{2,2} = ta_{1,2}I_2$, onde I_2 denota a matriz identidade 2×2. Neste caso I contém um elemento central não nulo. O argumento é similar se alguma outra entrada de x é não nula. Assim, $I \cap Z \neq 0$.

Como S é um PI-anel primo, seu centro $Z = (A_1 \cap A_2) + tA[t]$ é um domínio. O corpo de frações do centro Z de R é apenas o corpo de frações de $A[t]$. Denote este corpo por K. Visto que $A_1K = A_2K = (t)K = K$, temos que $RK = M_2(K)$. Logo, o anel de frações centrais de R é um anel de matrizes sobre um corpo. Em particular, ele é simples (veja o exemplo (7.3.2)) e tem dimensão finita sobre seu centro, que é K.

Exemplo 7.3.5 *Vamos verificar que o anel $R[c^{-1}]$ é um anel de matrizes, para algum valor c de um polinômio central de R. Recordando do exemplo (7.1.2), que $p(x,y) = (xy - yx)^2$ é um polinômio central para matrizes 2×2, façamos $x = te_{2,1}$ e $y = e_{1,2}$ vemos que $c = t^2 = t^2I_2$ é um valor de um polinômio central de R. Note que $R[c^{-1}] = R[t^{-2}] = R[t^{-1}] = M_2(A[t, t^{-1}])$ é o anel das matrizes 2×2 sobre $A[t, t^{-1}]$.*

Definição 7.3.6 *1. Um elemento r de um anel R é dito inteiro (ou algébrico) de grau t sobre um subanel S de R se $r^t = \sum_{i=0}^{t-1} s_i r^i$ para apropriados $s_0, \ldots, s_{t-1}$ em S.*

2. Um anel R é dito inteiro se todos os seus elementos são inteiros.

3. Um anel R é inteiro de grau limitado t, se cada elemento de R é inteiro de grau menor ou igual a t.

Exemplo 7.3.7 *Exibiremos um PI-anel primo S tal que*

1. *S não é um módulo finito sobre seu centro. Ainda mais, S não é inteiro sobre seu centro.*
2. *S é uma álgebra finitamente gerada sobre um corpo, ou seja S é afim.*
3. *S não é Noetheriano à esquerda nem à direita.*

Demonstração. Seja S como no exemplo (7.3.3) e denote por Z seu centro. Como um Z-módulo, S é a soma direta de quatro Z-módulos $A_1+(t)$, $A[t]$, $A_2+(t)$ e (t). Assim S será um Z-módulo finito, apenas se $A[t]$ for um Z-módulo finito. Visto que $Z=(A_1\cap A_2)+tA(t)$, isto será verdade apenas se A for um $(A_1\cap A_2)$-módulo finito. Mas, é fácil encontrar um exemplo em que isso falha. Seja $A=K[x,y]$ o anel de polinômios sobre um corpo K e seja $A_1=K[x]$ e $A_2=K[y]$. Então $A=K[x,y]$ não é finito sobre $A_1\cap A_2=K$. Explicitamente, o anel S é da forma

$$S=\begin{pmatrix} K[x]+tK[x,y,t] & K[x,y,t] \\ tK[x,y,t] & K[y]+tK[x,y,t] \end{pmatrix}.$$

Note que o subanel das matrizes diagonais não é inteiro sobre $Z=K+tK[x,y,t]$ visto que ambos $A_1=K[x]$ e $A_2=K[y]$, não são inteiros sobre $A_1\cap A_2=K$. Assim S não é inteiro sobre Z.

Para o segundo ítem, podemos verificar que S é gerada como K-álgebra pelas seis seguintes matrizes

$$e_{1,1},\quad xe_{1,1},\quad e_{1,2},\quad te_{1,2},\quad e_{2,2}\quad e\quad ye_{2,2}.$$

Finalmente, vamos verificar que S não é Noetheriano. Seja I o ideal $M_2((t))$. Então

$$S/I=\begin{pmatrix} K[x] & K[x,y] \\ 0 & K[y] \end{pmatrix}.$$

Nós pretendemos mostrar que S/I não é Noetheriano, que implica que S também não é Noetheriano. Para isso, seja V qualquer $K[x]$-submódulo de $K[x,y]$. Então $Ve_{1,2}$ é um ideal à esquerda de R/I. Mas $K[x,y]$ não é um $K[x]$-submódulo Noetheriano, porque $K[x,y]$ não é finitamente gerado como $K[x]$-módulo, então S/I não satisfaz a condição de cadeias ascendentes para ideais à esquerda, ou seja, S/I não é Noetheriano. Similarmente, usando $K[y]$-submódulo de $K[x,y]$ mostra-se que S/I também não é Noetheriano à direita. ■

Note que (2) e (3) mostram que o teorema da base de Hilbert [2], que diz que se uma álgebra comutativa R é Noetheriana então a álgebra polinomial $R[X]$ também é, não se generaliza para as PI-álgebras que são finitamente geradas sobre um subcorpo central.

Exemplo 7.3.8 *As noções, condições de cadeias à esquerda e à direita para anéis Noetherianos, coincidem para "bons" PI-anéis. O teorema de Cauchon* (veja [27] pp. 226)*, afirma que se um PI-anel semi-primo satisfaz a condição de cadeias ascendentes em ideais bilaterais, ele também satisfará para ambos, ideais à esquerda e à direita. Visto que a condição de cadeias ascendentes em ideais à esquerda ou à direita, implicam sempre a mesma condição para ideais bilaterais, temos que um PI-anel semi-primo é Noetheriano à esquerda se é também à direita.*

Capítulo 8

O Teorema sobre a Altura

Neste capítulo desenvolveremos tópicos relacionados com o bem conhecido problema de Kurosh formulado em 1941 [19]: Toda álgebra algébrica é localmente finita? Este problema foi solucionado afirmativamente para casos específicos por Jacobson [15], Levitzki [20] e Kaplansky [17]. Mais ainda, Shirshov ([29] e [30]) em 1957 demonstrou o teorema sobre a altura de uma álgebra dando um tratamento combinatorial ao problema de Kurosh. Aqui seguiremos o texto em [37]. Trabalharemos primeiro o resultado de Shirshov e na seção seguinte voltaremos ao problema de Kurosh.

8.1 Lema de Shirshov

Nós consideraremos palavras associativas formadas por elementos de algum conjunto ordenado finito $X_k = \{x_1, \ldots, x_k\}$ que é um subconjunto de um alfabeto $X = \{x_1, x_2, \ldots\}$ (veja a definição (1.3.13)).

Definição 8.1.1 1. *Diremos que uma palavra w é x_k-indecomponível se ela tem a forma $w = x_k x_k \cdots x_k x_{i_1} x_{i_2} \cdots x_{i_s}$, onde $s \geq 1$ e $i_t \neq k$ para $t = 1, 2, \ldots, s$.*

2. *Seja v uma palavra tal que $v = v_1 \cdots v_m$ onde cada v_i é uma palavra x_k-indecomponível. Então diremos que v é uma palavra x_k-fatorável e $v_1 \cdots v_m$ é a x_k-fatoração de v.*

Note que a x_k-fatoração de uma palavra v existe e é única se, e somente se ela começa com x_k e termina com um símbolo diferente de x_k.

Lembre-se da definição da ordem lexicográfica (1.3.13) dada sobre um alfabeto X. Consideremos a mesma ordem sobre X_k^* que é, como na definição (1.3.13), o conjunto das palavras sobre o alfabeto X_k.

Será útil agora estabelecer uma notação para um importante conjunto:

Definição 8.1.2 *Usaremos T_k para o conjunto das palavras que são x_k-indecomponíveis.*

Palavras x_k-fatoráveis podem ser consideradas como formadas por elementos de T_k, ou seja, podemos olhar para T_k como um alfabeto e assim as palavras sobre T_k serão as palavras x_k-fatoráveis. Mais a frente faremos isso e o conjunto dessas palavras será T_k^*.

Nós agora pretendemos introduzir uma ordem no conjunto T_k. Dados w e v em T_k, diremos que $v \prec w$ se v é lexicograficamente menor do que w ou se w é um segmento inicial de v. Formalmente diríamos:

Definição 8.1.3 *Sejam* $v = x_k \cdots x_k x_{i_1} \cdots x_{i_l}$ *e* $w = x_k \cdots x_k x_{j_1} \cdots x_{j_p}$ *palavras em* T_k, *então diremos que* $v \preceq w$ *se um dos três seguintes casos ocorre:*

(i) $v = w$, *ou seja,* $l = p$ *e* $i_t = j_t$ *para* $t = 1, 2, \ldots, l$;

(ii) existe $r \in \{1, 2, \ldots, l\}$ *tal que* $i_t = j_t$ *para* $1 \leq t \leq r-1$ *e* $i_r < j_r$;

(iii) $p < l$ *e* $i_t = j_t$ *para* $1 \leq t \leq p$.

Diremos também que $v \prec w$ *se* $v \preceq w$ *e* $v \neq w$.

Veja que o item *(iii)* inverte a ordem lexicográfica em T_k. Além disso, diz-se que w é um segmento inicial de v se ocorre *(iii)*.

Introduziremos agora uma ordenação no conjunto T_k^* citado anteriormente.

Definição 8.1.4 *Sejam* $v = v_1 \cdots v_l$ *e* $w = w_1 \cdots w_p$ *palavras em* T_k^* *onde* $v_1 \cdots v_l$ *e* $w_1 \cdots w_p$ *são suas* x_k*-fatorações, então diremos que* $v \trianglelefteq w$ *se um dos três seguintes casos ocorre:*

(i) $v = w$, *ou seja,* $l = p$ *e* $v_i = w_j$ *para* $t = 1, 2, \ldots, l$;

(ii) existe $r \in \{1, 2, \ldots, l\}$ *tal que* $v_t = w_t$ *para* $1 \leq t \leq r-1$ *e* $v_r \prec w_r$;

(iii) $l < p$ *e* $v_t = w_t$ *para* $1 \leq t \leq l$.

Diremos também que $v \triangleleft w$ *se* $v \trianglelefteq w$ *e* $v \neq w$.

Veja que o que temos é, na verdade, a ordem lexicográfica em T_k^*, sendo que a comparação feita entre as palavras v_i e w_j é determinada pela ordem $\preceq$.

Mais uma noção importante será dada pela seguinte definição:

Definição 8.1.5 *Diremos que uma palavra* y *sobre um alfabeto* X *é* n*-particionável se ela pode ser representada por um produto de* n *sub-palavras sobre* X, $y = y_1 \cdots y_n$, *tais que* $y_{\sigma(1)} \cdots y_{\sigma(n)} < y$ *para toda permutação* $\sigma \in S_n$ *diferente da identidade (A ordem aqui é a lexicográfica). Além disso, diremos que* $y_1 \cdots y_n$ *é uma* n*-partição de* y.

Exemplo 8.1.6 *A palavra* $w = x_3x_1x_2x_2x_1x_1x_2x_1x_1x_1$ *é 3-particionável pois temos*

$$w = (x_3x_1)(x_2x_2x_1x_1)(x_2x_1x_1x_1), \quad w = (x_3x_1x_2)(x_2x_1x_1)(x_2x_1x_1x_1),$$

$$w = (x_3)(x_1x_2x_2x_1x_1x_2)(x_1x_1x_1),$$

entre outras. Veja porém que $x_1x_2x_1x_3x_2x_1x_2x_3x_2$ *não é 2-particionável.*

Vamos agora usar a definição de n-paticionável pensando em $X = T_k$, ou seja, y é uma palavra sobre T_k, e estabeleceremos o que significa y ser n_T-particionável. Formalmente:

Definição 8.1.7 *Diremos que uma palavra y sobre T_k (em T_k^*) é n_T-particionável se ela pode ser representada por um produto de n sub-palavras sobre T_k, $y = y_1 \cdots y_n$, tais que $y_{\sigma(1)} \cdots y_{\sigma(n)} < y$ para toda permutação $\sigma \in S_n$ diferente da identidade. Além disso, diremos que $y_1 \cdots y_n$ é uma n_T-partição de y.*

Exemplo 8.1.8 *Seja $y = x_3x_3x_2x_3x_1x_3x_2x_1 \in T_3^*$, y é 3_T-particionável com a partição*

$$(x_3x_3x_2)(x_3x_1)(x_3x_2x_1)$$

e y é 2_T-particionável admitindo duas 2_T-partições,

$$(x_3x_3x_2)(x_3x_1x_3x_2x_1) \quad e \quad (x_3x_3x_2x_3x_1)(x_3x_2x_1).$$

Veja porém que a palavra $w = x_3x_3x_1x_3x_3x_3x_2x_3x_3x_2$ não é 2_T-particionável.

Agora vamos aos primeiros resultados desta seção.

Lema 8.1.9 *Seja w uma palavra associativa sobre T_k. Se w é n_T-particionável então w é n-particionável.*

Demonstração. Sejam $w \in T_k$ uma palavra associativa e $w = w_1 \cdots w_n$ uma n_T-partição da palavra w. Então $w, w_1, \ldots, w_n$ são x_k-fatoráveis. Segue-se da definição de n_T-particionável que $w_{\sigma(1)} \cdots w_{\sigma(n)} \lhd w$ para qualquer permutação $\sigma \in S_n$ diferente da identidade. É fácil ver que $w_{\sigma(1)} \cdots w_{\sigma(n)} < w$ (A ordem aqui é a lexicográfica). Desta forma dada uma n_T-partição teremos também uma n-partição, provando o que queríamos. ■

Lema 8.1.10 *Se temos uma palavra w $(n-1)_T$-particionável então a palavra wx_k é n-paticionável.*

Demonstração. Pelo lema (8.1.9) a palavra w é $(n-1)$-particionável:

$$w = (x_k x_{i_1} \cdots x'_{i_1})(x_k x_{i_2} \cdots x'_{i_2}) \cdots (x_k x_{i_{n-1}} \cdots x'_{i_{n-1}})$$

onde $x_i, x'_i \in X_k$, com $x'_{i_t} \neq x_k$ para $t = 1, 2, \ldots, n-1$.

Mostraremos que a palavra wx_k admite a seguinte n-partição:

$$wx_k = (x_k)(x_{i_1} \cdots x'_{i_1} x_k)(x_{i_2} \cdots x'_{i_2} x_k) \cdots (x_{i_{n-1}} \cdots x'_{i_{n-1}} x_k).$$

Qualquer permutação de sub-palavras da palavra wx_k que mantenha (x_k) na primeira posição transforma a palavra wx_k em uma nova palavra $w'x_k$, onde w' é obtida por meio de alguma permutação de sub-palavras na $(n-1)$-partição da palavra w, dada acima. Por esta razão $w' < w$ e $w'x_k < wx_k$. Por outro lado, se consideramos permutações que retiram símbolos x_k da primeira posição, então teremos que palavras obtidas desta maneira irão começar com um número menor de símbolos x_k em comparação com a palavra wx_k. Logo, elas serão estritamente menores do que a palavra wx_k. Isto prova o lema. ■

Lema 8.1.11 (Shirshov) *Sejam k, s, n números naturais. Então existe um número natural $N(k, s, n)$ tal que em qualquer palavra w associativa sobre k símbolos ordenados com comprimento $N(k, s, n)$ ocorrem s sub-palavras consecutivas iguais ou podemos determinar uma sub-palavra n-particionável.*

Demonstração. Podemos considerar palavras associativas sobre X_k. É fácil ver que o número $N(k,s,1) = 1$ para quaisquer k e s visto que toda palavra é 1-particionável. Desta forma temos uma base para indução em n. Assumiremos que existe um número $N(k,s,n-1)$ para todos k e s. Analogamente, $N(1,s,n) = s$ pois 1^s é uma sub-palavra da palavra w, e isso nos dá uma base para uma indução por k. Assim, suponhamos que o número $N(k-1,s,n)$ existe.

Consideremos uma palavra associativa w com comprimento

$$[s + N(k-1,s,n)][N(k^{N(k-1,s,n)+s}, s, n-1) + 1].$$

Se em algum segmento inicial w' de w, o número de entradas $x_i \neq x_k$ for maior ou igual a $N(k-1,s,n)$, então a hipótese de indução é satisfeita com respeito a sub-palavra w' que é um segmento inicial da palavra w e depende somente de $k-1$ geradores. Desta forma, nós podemos assumir que o comprimento da palavra w' é menor do que $N(k-1,s,n)$. No fim da palavra w podemos determinar uma sub-palavra $w'' = x_k x_k \cdots x_k$. Assim, podemos assumir que o comprimento da palavra w'' é menor que s, visto que, caso contrário, o lema estaria provado. Descartando, se existirem, as palavras w' e w'', obtemos uma sub-palavra w_1 com comprimento maior do que o número:

$$[s + N(k-1,s,n)].N(k^{N(k-1,s,n)+s}, s, n-1).$$

Para a palavra w_1 existe uma x_k-fatoração $w_1 = w_{11} \cdots w_{1m}$. Nós podemos assumir que o comprimento de cada palavra w_{1i} x_k-indecomponível é menor do que $s + N(k-1,s,n)$ pois, caso contrário, poderíamos determinar s símbolos consecutivos x_k ou uma sub-palavra de comprimento $N(k-1,s,n)$ que não conteria o símbolo x_k. Considerando esse limite para o comprimento, vemos que não existem mais do que $k^{N(k-1,s,n)+s}$ palavras x_k-indecomponíveis diferentes. Podemos considerar w_1 como uma palavra sobre T_k. Visto que seu comprimento m é estritamente maior do que $N(k^{N(k-1,s,n)+s}, s, n-1)$, podemos determinar na palavra w_1 s sub-palavras consecutivas iguais ou uma sub-palavra v $(n-1)_T$-particionável. Se a segunda hipótese ocorre; então, como a palavra w_1, considerada como uma palavra sobre T_k, tem comprimento maior do que $N(k^{N(k-1,s,n)+s}, s, n-1)$, podemos considerar que o símbolo x_k segue a sub-palavra v. Pelo lema (8.1.10), a sub-palavra vx_k é n-particionável. Assim, em qualquer caso, o lema se verifica. Desta forma nós verificamos que o número

$$N(k,s,n) = [s + N(k-1,s,n)][N(k^{N(k-1,s,n)+s}, s, n-1) + 1],$$

satisfaz o lema. ■

Lema 8.1.12 *Seja w uma palavra associativa com comprimento m que não pode ser representada na forma v^t, onde v é uma sub-palavra própria de w. Então, para qualquer número natural $n < m$, a palavra w^{2n} contém uma sub-palavra n-particionável.*

Demonstração. É possível obter por meio de permutações cíclicas das letras da palavra w, m palavras $w = w_0, w_1, \ldots, w_{m-1}$. Visto que a palavra w não é representável na forma de potência de uma de suas sub-palavras próprias, podemos escrever $w = x_{i_1}^{t_{i_1}} x_{i_2}^{t_{i_2}} \ldots x_{i_l}^{t_{i_l}}$ com $2 \leq l \leq m$; $x_{i_j} \neq x_{i_{j+1}}$ e $\sum_1^l t_{i_j} = m$. Como cada permutação cíclica retirará uma letra da

primeira potência $x_{i_1}^{t_{i_1}}$ acrescentando-a ao final da palavra w, e são m letras em w, todas as palavras $w_0, w_1, \ldots, w_{m-1}$ são diferentes.

Nós assumiremos que $w_{i_0} > w_{i_1} > \cdots > w_{i_{m-1}}$ no sentido lexicográfico. Claro que cada palavra w_i pode ser representada por $w_i = u_i v_i$ onde $v_i u_i = w$. Agora, consideraremos a palavra

$$w^{2n} = v_{i_0}u_{i_0}v_{i_0}u_{i_0}v_{i_1}u_{i_1}v_{i_1}u_{i_1}\cdots v_{i_{n-1}}u_{i_{n-1}}v_{i_{n-1}}u_{i_{n-1}}.$$

Façamos assim $w'_{i_k} = u_{i_k}v_{i_k}u_{i_k}v_{i_{k+1}}$ para $k = 0, 1, \ldots, n-2$, e $w'_{i_{n-1}} = u_{i_{n-1}}v_{i_{n-1}}u_{i_{n-1}}$, e ainda $y = v_{i_0}$. Então a palavra w^{2n} é representada na forma $w^{2n} = yw'_{i_0}w'_{i_1}\cdots w'_{i_{n-1}}$. Visto que o segmento inicial da palavra w'_j coincide com o segmento inicial da palavra w_j para $j = i_0, i_1, \ldots, i_{n-1}$ e $w_{i_0} > w_{i_1} > \cdots > w_{i_{m-1}}$, é fácil ver que a palavra $w'_{i_0}w'_{i_1}\cdots w'_{i_{n-1}}$ é n-particionável. Isto prova o lema. ■

Desta forma estamos prontos para demonstrar o teorema sobre a altura de uma álgebra, que é o resultado mais importante dessa seção.

Analogamente à seção (1.3) nós consideraremos a álgebra associativa com identidade $K\{X_k\}$ dos polinômios sobre um corpo K com indeterminadas em X_k. Ou seja, de forma natural pensaremos em X_k no lugar de X.

Definição 8.1.13 *Seja X um alfabeto e seja Y um conjunto de palavras sobre X. Seja w uma palavra sobre X tal que $w = w_1^{n_1}\cdots w_h^{n_h}$ onde cada $w_i \in Y$ para $i = 1, \ldots, h$ e $w_i \neq w_{i+1}$ para $i = 1, \ldots, h-1$. Assim, o menor número h com essa propriedade será chamado a altura de w em relação ao conjunto Y.*

Exemplo 8.1.14 *Seja $w = x_1x_2x_2x_1x_2x_2x_1x_2x_2$. Veja que w tém altura 6 em relação a $Y_1 = \{x_1, x_2\}$ e altura 1 em relação a $Y_2 = \{x_1x_2x_2\}$.*

Exemplo 8.1.15 *Seja $w = x_1x_1x_2x_3x_1x_2x_3x_1$. Veja que w tém altura 7 em relação a $Y_1 = \{x_1, x_2, x_3\}$ e altura 3 em relação a $Y_2 = \{x_1, x_2x_3x_1\}$.*

Definição 8.1.16 *Sejam A uma álgebra associativa sobre um corpo K, gerada por um conjunto finito $G = \{a_1, \ldots, a_k\}$ e $P = \{p_1, \ldots, p_l\}$ um conjunto finito de polinômios homogêneos de $K\{X_k\}$. Para cada $p \in P$, denotaremos $\overline{p} = p(a_1, \ldots, a_k)$, a imagem de $p \in K\{X_k\}$ na álgebra A por meio do homomorfismo que associa x_i a a_i. Então A tem altura limitada em relação a (G, P) se existe um número h tal que para cada monômio m em $K\{X_k\}$ existem monômios $u_1, \ldots, u_n$ em $K\{X_k\}$ para os quais valem*

1. $\overline{m} = \sum_{i=1}^{n} u_i(\overline{p_1}, \overline{p_2}, \ldots, \overline{p_l})$;
2. *a altura de cada u_i em relação a X_k é menor ou igual a h;*
3. *cada $u_i(p_1, \ldots, p_l)$ tem o mesmo tipo que m.*

O menor número h com essa propriedade é chamado altura da álgebra A em relação a (G, P).

Exemplo 8.1.17 *Veja que se A é uma álgebra associativa e comutativa gerada por $G = \{a_1, \ldots, a_k\}$ e m é um monômio de $K\{X_k\}$ do tipo $[i_1, \ldots, i_k]$, então $\overline{m} = \alpha a_1^{i_1}\ldots a_k^{i_k}$, para algum $\alpha \in K$. Assim, a altura de todo monômio de $K\{X_k\}$ em relação a (G, X_k) não excede k. Portanto, toda álgebra associativa e comutativa finitamente gerada é um exemplo de álgebra com altura limitada.*

Teorema 8.1 (Teorema sobre a Altura) *Seja A uma K-álgebra associativa, com conjunto de geradores $G = \{a_1, \ldots, a_k\}$, satisfazendo uma identidade polinomial de grau n. Se Y é o conjunto de todas as palavras sobre X_k com comprimento menor do que n e P é o conjunto de todos os monômios da forma $1y$ onde $y \in Y$ então a álgebra A tem altura limitada com respeito a (G, P).*

Demonstração. Pelo que foi provado na seção (3.1), se uma álgebra associativa satisfaz uma identidade polinomial de grau n, então ela também satisfaz uma identidade polinomial multi-linear de grau n e ainda mais, vimos que podemos assumir que A satisfaz uma identidade da forma

$$x_1 x_2 \cdots x_n - \sum_{\sigma \in S_n} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(n)} = 0.$$

Para provarmos o teorema, basta mostrarmos que existe um número $M = M(n,k)$ tal que cada palavra associativa s sobre X_k de altura, com respeito a Y, maior ou igual a M, contém uma sub-palavra n-particionável. Suponha que existe tal M e, veja que na álgebra A nós teremos $\overline{s} = \sum_i \alpha_i \overline{s}_i$, onde $\alpha_i \in K$ e as s_i são palavras sobre X_k com as mesmas letras que s mas estritamente menores que s. Se alguma das palavras s_i tem altura maior ou igual a M, então nós continuamos o processo sobre essa s_i. Em vista desta estrita monotonicidade, nós obteremos que $\overline{s} = \sum_i \alpha_i \overline{s}_i$, onde cada uma das palavras s_i tem altura, com respeito a Y, menor do que M.

Pelo lema (8.1.11), existe um número $N = N(k, 2n, n)$ tal que cada palavra associativa s de comprimento N sobre X_k que não contém sub-palavras n-particionáveis, contém uma sub-palavra da forma v^{2n} onde v é uma sub-palavra própria de s. Mais ainda, podemos assumir que a palavra v não pode ser representada na forma $(v_1)^k$ com $k > 1$. Então, pelo lema (8.1.12), o comprimento $d(v)$ da palavra v é menor que n.

Agora vamos mostrar que para cada palavra sobre X_k com altura, com respeito a Y, maior ou igual a $N + 2$, que não contém sub-palavras n-particionáveis, contém uma sub-palavra v_1 da forma $v_1 = v^n v'$ onde $n > d(v) \geq d(v') > 0$ e a palavra v' não é segmento inicial da palavra v. Se w tem altura, em relação a Y, maior do que $N+2$, então w tem comprimento maior do que $N + 2$. O segmento inicial de comprimento N contém uma sub-palavra u^n, onde $d(u) < n$; logo, $w = w_0 u^n w_0'$, onde w_0' tem altura maior do que 2. Podemos escrever $w_0' = u^k u' w_0''$ e $u = u'u''$ onde u' pode ser vazia mas u'' não; além disso, w_0'' não começa com um segmento inicial de u'' e tem altura positiva, ou seja, não é vazia. Então

$$w = w_0 u^k u^n u' w_0'' = w_0 u^k (u'u'')^n u' w_0'' = w_0 u^k u' (u''u')^n w_0'' = w_0 u^k u' v^n w_0'',$$

onde $v = u''u'$ tem o mesmo comprimento que u e w_0'' não começa com um segmento inicial de v. Tomando v' como o segmento inicial de w_0'' de comprimento não nulo e menor ou igual a $d(v)$, teremos a palavra $v_1 = v^n v'$ procurada.

Como o conjunto destas sub-palavras é finito, para um número natural M suficientemente grande, cada palavra com altura, com respeito a Y, maior ou igual a M que não contém sub-palavras n-particionáveis, conterá n sub-palavras iguais da forma $v_1 = v^n v'$ que não serão necessariamente consecutivas. Para cada palavra assim, existirá uma sub-palavra que é n-particionável em uma das seguintes formas:

$$(v^n v' u_1 v)(v^{n-1} v' u_2 v^2) \cdots (v v' u_n), \text{ caso } v < v',$$

ou

$$(v'u_1v^{n-1})(vv'u_2v^{n-2})\cdots(v^{n-1}v'u_n),\ \text{caso } v' < v.$$

Conseqüentemente cada palavra que tem altura, com respeito a Y, maior do que M contém uma sub-palavra n-particionável. Isto prova o teorema. ■

Na próxima seção demonstraremos os resultados devidos a Levitzki e Kaplansky como corolários do teorema sobre a altura de uma álgebra.

8.2 O Problema de Kurosh

Vamos descrever o problema de Kurosh, que motivou a demonstração do teorema sobre a altura de uma álgebra tratado na seção anterior.

Para uma definição de álgebra algébrica de índice limitado, veja a definição análoga em (7.3.6).

Exemplo 8.2.1 *Considere uma álgebra associativa A de dimensão n sobre um corpo K. Devido à dimensão n de A, para cada elemento $a \in A$ sempre existirão $\alpha_1, \alpha_2, \ldots, \alpha_{n+1} \in K$ tais que $\sum_{i=1}^{n+1} \alpha_i a^i = 0$ pois a, a^2, $\ldots$, a^{n+1} são linearmente dependentes sobre K. Assim A é uma álgebra algébrica de índice limitado por $n+1$.*

Definição 8.2.2 *Uma álgebra A é localmente finita se cada subálgebra finitamente gerada de A tem dimensão finita.*

Proposição 8.2.3 *Toda álgebra algébrica de índice limitado sobre um corpo K é uma PI-álgebra.*

Demonstração. Sejam A uma álgebra algébrica de índice limitado sobre um corpo K e n o limite dos índices algébricos dos elementos de A. Então para quaisquer a e b em A, o elemento $[a^n, b]$ é uma combinação linear dos elementos $[a^{n-1}, b], \ldots, [a, b]$ com coeficientes em K, pois $a^n = \sum_{i=1}^{n-1} \alpha_i a^i$, onde $\alpha_i \in K$. Logo os elementos $[a^n, b], \ldots, [a, b]$ anulam o polinômio standard $s_n(x_1, \ldots, x_n)$. Conseqüentemente

$$s_n([x_1^n, x_2], [x_1^{n-1}, x_2], \ldots, [x_1, x_2]) = 0$$

é uma identidade polinomial de A ■

O problema de Kurosh: se A é uma álgebra algébrica sobre um corpo K, então uma subálgebra de A, finitamente gerada, tem sempre dimensão finita sobre K? Ou, em outras palavras, toda álgebra algébrica é localmente finita?

Com esta generalidade, a resposta é negativa devido aos resultados de E. S. Golod e I. R. Shafarevich (Voltaremos a comentá-los à frente). Mas, em 1945, Jacobson encontrou uma resposta positiva para álgebras algébricas de índice limitado. No ano seguinte, Levitzki provou que toda PI-álgebra finitamente gerada e que é nil álgebra é nilpotente. Em 1948, Kaplansky ampliou o resultado de Jacobson, resolvendo o problema de Kurosh para PI-álgebras. Apesar das abordagens de Levitzki e Kaplansky serem diferentes, Shirshov, abordando o problema de Kurosh do ponto de vista combinatório, demonstrou o teorema sobre a altura de uma álgebra e obteve seus resultados como corolários.

Vamos mostrar os resultados de Levitzki e Kaplansky como corolários do teorema sobre a altura de uma álgebra demonstrado na seção anterior.

Teorema 8.2 (Levitzki – 1946) *Toda PI-álgebra finitamente gerada e que é nil álgebra é nilpotente.*

Demonstração. Seja A uma nil álgebra associativa sobre um corpo K com um conjunto de geradores $G = \{a_1, \ldots, a_k\}$, satisfazendo uma identidade polinomial de grau n. Sejam Y o conjunto de todas as palavras sobre $X_k = \{x_1, \ldots, x_k\}$ com comprimento menor do que n e P o conjunto de todos os monômios da forma $1y$ onde $y \in Y$. Estas são as hipóteses do teorema sobre a altura de uma álgebra, então a álgebra A tem altura limitada com respeito a (G, P).

Seja h essa altura; sejam V o conjunto de todos os produtos com menos que n geradores da álgebra A e m o máximo dos índices de nilpotência dos elementos do conjunto V. Então a álgebra A é gerada como um espaço vetorial por um conjunto finito formado por todos os produtos com menos do que $(m-1).n.h$ geradores.

Seja $M = (m-1).n.h + 1$. Se a é um elemento de A que é produto de M elementos de G, então podemos escrever

$$a = \sum_i \alpha_i a_{i_1}^{m_{i_1}} \ldots a_{i_l}^{m_{i_l}},$$

onde cada $i_l < h$, cada $a_{i_l} \in V$ e $\alpha \in K$. Como $(m-1).n.h < M$ e $M < (m_{i_1} + \cdots + m_{i_l}).n$, podemos concluir que existe um j em $\{1, \ldots, l\}$ tal que $m_{i_j} > m-1$. Logo, $a = 0$. Desta forma, podemos ver que todo produto de quaisquer M elementos de A é igual a zero. Portanto, A é uma álgebra nilpotente. ■

Teorema 8.3 (Kaplansky) *Se A é uma álgebra algébrica finitamente gerada que satisfaz uma identidade polinomial então A tem dimensão finita.*

Demonstração. Seja A uma álgebra associativa e algébrica sobre um corpo K com um conjunto de geradores $G = \{a_1, \ldots, a_k\}$, satisfazendo uma identidade polinomial de grau n. Sejam Y o conjunto de todas as palavras sobre $X_k = \{x_1, \ldots, x_k\}$ com comprimento menor do que n e P o conjunto de todos os monômios da forma $1y$ onde $y \in Y$. Estas são as hipóteses do teorema sobre a altura de uma álgebra, então a álgebra A tem altura limitada com respeito a (G, P).

Seja h essa altura; sejam V o conjunto de todos os produtos com menos que n geradores da álgebra A onde n é o grau da identidade polinomial de A e m o máximo dos graus algébricos dos elementos do conjunto V. Então a álgebra A é gerada como um espaço vetorial por um conjunto finito formado por todos os produtos com menos do que $(m-1).n.h$ geradores. ■

Depois de considerarmos as duas próximas definições, poderemos enunciar mais dois corolários, entre os vários resultados obtidos a partir do teorema sobre a altura de uma álgebra.

Definição 8.2.4 *Uma álgebra A é localmente nilpotente se cada subálgebra finitamente gerada de A é nilpotente.*

Definição 8.2.5 *Uma álgebra A tem altura localmente limitada se cada subálgebra finitamente gerada de A tem altura limitada.*

Corolário 8.2.6 *Toda nil álgebra de indice limitado é localmente nilpotente.*

Corolário 8.2.7 *Toda PI-álgebra tem altura localmente limitada.*

Vamos exemplificar porque o problema de Kurosh não tem resposta afirmativa em todos os casos.

Em 1964, E. S. Golod e I. R. Shafarevich (veja [10] e [11] ou [13] pp. 188) provaram que existem nil álgebras finitamente geradas que não são nilpotentes. Agora, considere o seguinte teorema:

Teorema 8.4 *Toda nil álgebra de dimensão finita é nilpotente.*

Demonstração. Antes da demonstração propriamente dita, veja que, fazendo cálculos simples podemos mostrar que se I_1 e I_2 são ideais nilpotentes à esquerda de uma álgebra associativa, então $(I_1+I_2)^{n_1+n_2-1} = \{0\}$, onde n_i é o índice de nilpotência de I_i com $1 \leq i \leq 2$.

Seja A uma nil álgebra de dimensão finita n. Se $a \in A$, então $Aa = \{xa : x \in A\}$ é um ideal à esquerda de A. Seja k o índice de nilpotência de A. Vejamos que para a transformação linear $T_a : A \to Aa$ dada por $T_a(x) = xa$, se $a \neq 0$, $\ker T_a \neq \{0\}$. Isto é verdade pois $T_a(a^{k-1}) = a^{k-1}.a = a^k = 0$ e $a^{k-1} \neq 0$ para algum $k > 1$.

Usaremos indução sobre n para mostrarmos que Aa é nilpotente para todo $a \in A$, e isto nos levará a demonstração do teorema. Se $n = 1$, a dimensão da imagem de T_a é zero pois já vimos que $\ker T_a \neq \{0\}$. Assim, $Aa = \{0\}$ para todo $a \in A$; logo, A é nilpotente. Se $n > 1$, então $Aa \neq A$ pois $\dim(Aa) < n$ visto que $\ker T_a \neq \{0\}$. Desta forma, por hipótese de indução, Aa é nilpotente para todo $a \in A$.

Sendo assim, basta mostrarmos que existem $a_1, \ldots, a_r \in$ tais que $Aa \subseteq Aa_1 + \cdots + Aa_m$ para todo $a \in A$ pois, neste caso, $A^2 \subseteq Aa_1 + \cdots + Aa_r$ e, além disso, como $Aa_1 + \cdots + Aa_m$ é nilpotente, A^2 é nilpotente e conseqüentemente A também é.

Para provarmos tal existência, seja $a_1 \in A$; se existe $a_2 \in A$ tal que $Aa_2 \nsubseteq Aa_1$, então $Aa_2 \subseteq Aa_1 + Aa_2$. Se existe $a_3 \in A$ tal que $Aa_3 \nsubseteq Aa_1 + Aa_2$, então $Aa_3 \subseteq Aa_1 + Aa_2 + Aa_3$. seguindo desta forma, obteremos os elementos $a_1, \ldots, a_r \in A$ tais que $Aa \subseteq Aa_1 + \cdots + Aa_r$ para todo $a \in A$ pois sempre teremos $\dim(Aa_1 + \cdots + Aa_i) < \dim(Aa_1 + \cdots + Aa_{i+1})$. ■

A partir do que provaram E. S. Golod e I. R. Shafarevich, vemos que este teorema nos dá uma resposta negativa para o Problema de Kurosh.

Capítulo 9

Álgebras Concretas

Neste capítulo pretendemos estudar dois resultados sobre álgebras específicas: uma base de identidades para a álgebra de Lie $sl_2(K)$ das matrizes de traço zero de ordem 2 sobre um corpo K de característica zero (veja o exemplo 1.3.16), e uma base de identidades para a álgebra $M_2(K)$ das matrizes de ordem dois. Em outras palavras, veremos respostas positivas nos casos de $sl_2(K)$ e $M_2(K)$ para um caso particular do problema de W. Specht [31] em 1950: se A é uma álgebra associativa sobre um corpo de característica zero, então $T(A)$ (o ideal das identidades polinomiais de A) é finitamente gerado como T-ideal?

9.1 As álgebras $sl_2(K)$ e $gl_2(K)$

Buscaremos aqui um pouco mais de familiaridade com as álgebras de Lie com as quais pretendemos trabalhar nesse capítulo. Vamos começar relembrando o conceito de álgebra de Lie dado na definição (1.3.8).

Definição 9.1.1 *Diremos que o espaço vetorial $\boldsymbol{L}$ é uma álgebra de Lie se para todo a, b e c em $\boldsymbol{L}$ temos $aa = 0$ que é a anti-comutatividade e $a(bc)+b(ca)+c(ab) = 0$ que é a identidade de Jacobi.*

Lema 9.1.2 *Seja A uma K-álgebra associativa. O espaço vetorial de A pode ser dotado com uma estrutura de álgebra de Lie por introduzirmos nele uma nova operação binária multilinear $[,] : A \otimes_K A \to A$ por meio da fórmula $[a, b] = ab - ba$ para quaisquer elementos a e b do espaço vetorial de A. A álgebra de Lie obtida dessa maneira, da álgebra associativa A, será denotada por $A^{(-)}$.*

Demonstração. Mostraremos primeiro que a operação é distributiva; dados a, b, $c \in A^{(-)}$ temos:

$$[a, (b + c)] = a(b + c) - (b + c)a = ab + ac - ba - ca = (ab - ba) + (ac - ca) = [a, b] + [a, c].$$

De maneira similar mostra-se que $[(a + b), c] = [a, c] + [b, c]$. Vejamos agora que se $a \in A^{(-)}$ então $[a, a] = aa - aa = 0$. Logo $A^{(-)}$ é anti-comutativa. Para verificarmos a identidade de Jacobi façamos:

$$[a, [b, c]] = [a, bc - cb] = abc - acb - bca + cba.$$

Então, fazendo uma permutação cíclica das letras temos:

$$[b,[c,a]] = [b, ca - ac] = bca - bac - cab + acb \quad \text{e}$$

$$[c,[a,b]] = [c, ab - ba] = cab - cba - abc + bac.$$

Somando as três últimas igualdades temos:

$$[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0 \in A^{(-)},$$

que é a identidade desejada. Finalmente, se $\lambda \in K$ e $a, b \in A^{(-)}$ então

$$\lambda[a,b] = \lambda(ab - ba) = (\lambda a)b - b(\lambda a) = [\lambda a, b].$$

Similarmente, mostra-se que $\lambda[a,b] = [a, \lambda b]$. Mostramos então que $A^{(-)}$ satisfaz os axiomas de uma álgebra de Lie sobre um corpo K. ■

Recordemos nossa notação $End_K M$ para a álgebra associativa dos endomorfismos de um K-módulo M, não esquecendo que K-módulo, quando K é corpo, é um espaço vetorial. Se φ e ψ pertencem a $End_K M$, $x \in M$ e $\lambda \in K$, então, por definição

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x), \quad (\varphi\psi)(x) = \varphi(\psi(x)) \text{ e } (\lambda\varphi)(x) = \lambda\varphi(x).$$

Sejam M um espaço vetorial com dimensão finita sobre um corpo K e $B = \{b_1, b_2, \ldots, b_n\}$ uma base de M. Então, a cada endomorfismo φ de M nós associamos a matriz A_φ do operador linear φ na base B. Essa associação $\varphi \to A_\varphi$ é um isomorfismo entre a álgebra associativa $End_K(M)$ e a álgebra $M_n(K)$ das matrizes $n \times n$ com entradas de K (essa última com as operações usuais). A álegbra de Lie $M_n(K)^{(-)}$ será denotada por $gl_n(K)$.

Para nós, um importante exemplo de álgebra de Lie será a subalgebra $sl_n(K)$ de $gl_n(K)$ das matrizes X com traço zero. Denotando traço de X por $tr(X)$, lembre-se que, se $X = (x_{ij})$, $1 \leq i, j \leq n$, então $trX = \sum_{i=1}^{n} x_{ii}$. Além disso, $tr(XY) = tr(YX)$ e o fato de que tr é uma função linear nos mostram que $sl_n(K)$ é fechada para comutadores.

Outro exemplo é a álgebra de Lie $O_n(K)$ das matrizes anti-simétricas de ordem n sobre um corpo K (exemplo 1.3.16). Veja que $[X,Y]^t = (XY - YX)^t = (XY)^t - (YX)^t = Y^tX^t - X^tY^t = YX - XY = -[X,Y]$.

9.2 Derivação Interna

Vamos obter aqui as primeiras noções sobre derivação interna que faremos uso nas próximas seções.

Definição 9.2.1 *Seja A uma K-álgebra. Diremos que um endomorfismo ζ do espaço vetorial A é uma derivação se para quaisquer x, $y \in A$, temos*

$$\zeta(xy) = \zeta(x)y + x\zeta(y).$$

Definição 9.2.2 *Sejam L uma álgebra de Lie sobre um corpo K e $x \in L$. Por $ad\ x$ denotaremos o endomorfismo do espaço vetorial L dado por*

$$ad\ x(y) = xy \quad \textit{para todo } y \in L.$$

Vamos verificar que ad x é uma derivação. Visto que L é anti-comutativa e satisfaz a identidade de Jacobi, $x(yz) + z(xy) + y(zx) = 0$, temos

$$ad\ x(yz) = x(yz) = -y(zx) - z(xy) = y(xz) + (xy)z = ad\ x(y)z + y\ ad\ x(z).$$

Qualquer derivação da forma ad x, $x \in L$, é chamada uma derivação interna determinada pelo elemento x. Assim podemos considerar formalmente as seguintes notações; o conjunto de todas as derivações de L, $Der_K(L)$, e o subconjunto de $Der_K(L)$, $ad\ L$, de todas as derivações internas de L.

Lema 9.2.3 *Seja A uma álgebra associativa sobre um corpo K. Então $Der_K(A)$ é uma subálgebra da álgebra $Der_K(A)^{(-)}$.*

Demonstração. Sejam $\delta_1, \delta_2 \in Der_K(A)$ e $x_1, x_2 \in A$. Então temos

$$[\delta_1, \delta_2](x_1x_2) = (\delta_1\delta_2 - \delta_2\delta_1)(x_1x_2)$$

$$= \delta_1(\delta_2(x_1)x_2 + x_1\delta_2(x_2)) - \delta_2(\delta_1(x_1)x_2 + x_1\delta_1(x_2))$$

$$= \delta_1\delta_2(x_1)x_2 + \delta_2(x_1)\delta_1(x_2) + \delta_1(x_1)\delta_2(x_2) + x_1\delta_1\delta_2x_2$$

$$-\delta_2\delta_1(x_1)x_2 - \delta_1(x_1)\delta_2(x_2) - \delta_2(x_1)\delta_1(x_2) - x_1\delta_2\delta_1(x_2).$$

Agrupando os termos semelhantes,

$$[\delta_1, \delta_2](x_1x_2) = \delta_1\delta_2(x_1)x_2 - \delta_2\delta_1(x_1)x_2 + x_1\delta_1\delta_2x_2 - x_1\delta_2\delta_1(x_2).$$

Usando a distributividade e a definição do comutador, obtemos

$$[\delta_1, \delta_2](x_1x_2) = (\delta_1\delta_2(x_1)\delta_2\delta_1(x_1))x_2 + x_1(\delta_1\delta_2x_2 - \delta_2\delta_1(x_2))$$

$$= [\delta_1, \delta_2](x_1)x_2 + x_1[\delta_1, \delta_2](x_2).$$

Assim, $[\delta_1, \delta_2] \in Der_K(A)$, ou seja, $Der_K(A)$ é fechada para o comutador. Basta agora ver que $Der_K(A)$ é uma álgebra de Lie. Esta álgebra é chamada a álgebra de Lie das derivações de uma K-álgebra A. ■.

Veja que existe uma aplicação natural $ad : L \longrightarrow Der_K(L)$ que a cada $x \in L$ associa a derivação interna $ad\ x$ determinada por x. Vamos então ao resultado seguinte.

Teorema 9.1 *A aplicação ad, acima mencionada, é um homomorfismo de álgebras de Lie. A imagem ad L deste homomorfismo é um ideal de $Der_K(L)$.*

Demonstração. A prova de que ad é um homomorfismo de espaços vetoriais não é difícil. Sejam então $x, y \in L$. Logo,

$$ad\ (xy)(z) = (xy)z = -z(xy) = x(yz) + y(zx)$$

$$= x(yz) - y(xz) = (ad\ x\ ad\ y)(z) - (ad\ y\ ad\ x)(z)$$

$$= [ad\ x, ad\ y](z).$$

Portanto $ad\ (xy) = [ad\ x, ad\ y]$, provando a primeira parte do teorema.

Agora podemos mostrar que $ad\ L$ é um ideal de $Der_K(L)$. Para isso computaremos $[\delta, ad\ x]$, onde $\delta \in Der_K(L)$ e $x \in L$, em um dado $y \in L$. Então teremos

$$[\delta, ad\ x](y) = (\delta ad\ x - ad\ x\delta)(y)$$

$$= \delta(xy) - ad\ x(\delta(y))$$

$$= (\delta x)(y) + x(\delta(y)) - x(\delta(y))$$

$$= (\delta x)(y) = ad\ \delta(x)(y).$$

Assim $[\delta, ad\ x] = ad\ \delta(x) \in ad\ L$. Logo $ad\ L$ é um ideal de $Der_K(L)$. ■

9.3 Identidades Polinomiais e Diagramas de Young

Começaremos definindo as tabelas de Young e desenvolvendo algumas de suas propriedades.

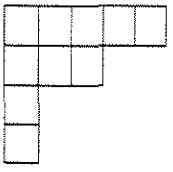
Definição 9.3.1 *Seja n um número natural, uma decomposição de n é uma seqüência $d = (n_1, n_2, \ldots, n_t)$ de números naturais tais que*

$$n = n_1 + n_2 + \cdots + n_t \quad onde \quad n_1 \geq n_2 \geq \cdots \geq n_t.$$

Definição 9.3.2 *Seja $d = (n_1, n_2, \ldots, n_t)$, uma decomposição de n. O conjunto $[d] = \{(i,j) | i, j \in \mathbb{N}, i \leq t, j \leq n_i\}$ é chamado diagrama de Young correspondendo a decomposição d. Os elementos (i,j) são os quadrados, ou as células de $[d]$.*

Exemplo 9.3.3 *Se $d = (5,3,1,1) = (5,3,1^2)$ é uma decomposição de 10 então o diagrama de Young $[d]$ associado à d tem a seguinte representação gráfica*

$[d] =$



Se $\tau \in S_n$ então τd designará à tabela de Young que é o diagrama de Young preenchido com os números $1, 2, \ldots, n$ da seguinte maneira: nós consideramos a seqüência $\tau(1), \ldots, \tau(n)$ e, preservando sua ordem, inserimos no diagrama de Young de cima para baixo e da esquerda para a direita.

Exemplo 9.3.4 *Seja* $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 1 & 10 & 7 & 6 & 4 & 5 & 3 & 9 & 8 \end{pmatrix} \in S_{10}$ *então*

$\tau d =$

2	6	5	9	8
1	4	3		
10				
7				

Sejam $d = (n_1, n_2, \cdots, n_t)$ uma decomposição de n e τd a tabela de Young com a permutação $\tau = \{d_{ij}\}$ de {1,...,n}, o número d_{ij} estando situado na (i,j)-ésima célula de τd.

Definição 9.3.5 *A tabela τd é chamada standard se $d_{ij} \leq d_{pq}$ para todo $i \leq p$ e $j \leq q$. Em outras palavras, as entradas nas linhas formam uma seqüência crescente da esquerda para a direita e nas colunas de cima para baixo.*

Exemplo 9.3.6 *Consideremos* $d = (3,2,1)$, $\tau_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 2 & 6 & 3 \end{pmatrix}$ *e* $\tau_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 6 & 2 & 4 & 3 \end{pmatrix}$ *então*

$$\tau d_1 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 6 \\ \cline{1-2} 5 \\ \cline{1-1} \end{array} \quad e \quad \tau d_2 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 5 & 4 \\ \cline{1-2} 6 \\ \cline{1-1} \end{array} .$$

Veja que τd_1 é standard porém τd_2 não é.

Definição 9.3.7 *Seja d uma decomposição de n com entradas (i,j). O número h_{ij} de células abaixo da (i,j)-ésima célula (e na mesma j-ésima coluna) e à direita da mesma (i,j)-ésima célula (e na mesma i-ésima linha) mais a própria célula (i,j) é chamado número do gancho da célula (i,j).*

Exemplo 9.3.8 *Na decomposição $d = (3,2,1)$ do exemplo 9.3.6 temos $h_{11} = 5$, $h_{12} = 3$, $h_{13} = 1$, $h_{21} = 3$, $h_{22} = 1$, $h_{31} = 1$.*

Recordemos que P_n (veja seção 4.4) denota o K-subespaço da álgebra livre de polinômios multi-lineares de grau n em $x_1, \ldots, x_n$. Além disso, P_n tem uma estrutura natural de S_n-módulo definida por $\sigma(x_{i_1} \cdots x_{i_n}) = x_{\sigma(i_1)} \cdots x_{\sigma(i_n)}$, $\sigma \in S_n$ e $x_{i_1} \cdots x_{i_n} \in P_n$. Todos os módulos nós estamos considerando à esquerda. Em outras palavras, os elementos de S_n agem como transformações lineares não singulares em P_n. Recordamos também que KS_n é a álgebra do grupo S_n (veja o exemplo 1.3.16). Para o nosso caso, seja $X = \{x_1, \ldots, x_n\}$ um conjunto de geradores livres da álgebra de Lie livre $L = L\{X\}$. Então PL_n denotará o K-espaço vetorial dos polinômios multi-lineares de grau n em L.

Definição 9.3.9 *Sejam $x_1, x_2, \ldots, x_n \in X$. O produto definido por*

$$x_1 x_2 \cdots x_n = x_1(x_2 \cdots x_n), \text{ se } n > 1,$$

é chamado produto normado à direita. Denotamos por $x^m y$, o produto $\underbrace{x \cdots x}_{m \text{ vezes}} y$.

Lema 9.3.10 *Uma base do K-espaço PL_n pode ser escolhida na forma*

$$x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(n-1)} x_n \quad \sigma \in S_{n-1}. \tag{9.1}$$

Demonstração. Usando indução sobre n podemos assumir que um monômio $w \in PL_n$ tem a forma uvx_n onde vx_n é um monômio normado à direita. Se o grau d de u é igual a 1, w é como queríamos. Caso contrário $u = u_1u_2$. Então $w = (u_1u_2)vx_n = u_1u_2vx_n - u_2u_1vx_n$. Usando indução sobre u_1v_n e u_2v_n diminuiremos o grau de u. Por indução sobre o grau de u, mostramos que w pode ser representado da forma enunciada.

Se $\sum_{\sigma \in S_{n-1}} \alpha_\sigma x_{\sigma(1)}x_{\sigma(2)} \cdots x_{\sigma(n-1)}x_n = 0$ e o coeficiente α_1 do monômio $x_1x_2 \cdots x_{n-1}x_n$ for não nulo, substituiremos os x_i por elementos apropriados da álgebra de Lie $gl_{n+1}(K)$. Seja $x_{n+1} = e_{11}$, $x_n = e_{21}$, $x_{n-1} = e_{32}$, ..., $x_1 = e_{n+1,n}$ onde e_{ij} são a base usual da álgebra $M_n(K)$. A única parcela da somatória que não se anula após tal substituição, é $\alpha_1x_1x_2 \cdots x_{n-1}x_n$ cujo valor será $\alpha e_{n+1,1}$. Logo a somatória não se anula em $gl_{n+1}(K)$, um absurdo. Portanto os elementos do enunciado do lema são linearmente independentes. ■

Do lema temos que $\dim PL_n = (n-1)!$. Veja também que se $f, g \in PL_n$ então $g = 0$ é uma conseqüência de $f = 0$ (veja definição (2.3.9)) se, e somente se $g \in KS_nf$. Desta forma, temos tantos sistemas distintos de identidades multi-lineares de grau n quantos forem os KS_n-submódulos distintos em PL_n. Pelo lema (9.3.10) qualquer elemento em P_n é uma combinação linear de monômios da forma (9.1). Então qualquer um deles, digamos $x_1x_2 \ldots x_n$, pode ser escolhido como gerador de PL_n que é um KS_n-módulo cíclico. Portanto PL_n é um módulo quociente do S_n-módulo KS_n sob um homomorfismo que associa 1 a $x_1x_2 \ldots x_n$.

Como estamos considerando K de característica 0, a descrição dos S_n-módulos irredutíveis pode ser baseada em decomposições e diagramas de Young. Para estudarmos tal descrição, vamos precisar de alguns resultados da teoria de representações do grupo S_n.

Definição 9.3.11 *Sejam d uma decomposição de n e τd uma tabela de Young correspondente. Definiremos uma ação de S_n no conjunto das tabelas de Young τd da forma seguinte. Se as entradas de τd são d_{ij} (i.é., d_{ij} está na (ij)-ésima célula de τd), e $\sigma \in S_n$ então $\sigma(\tau d)$ é uma tabela de Young tendo entradas $\sigma(d_{ij})$ na (i,j)-ésima célula. Denotaremos então $L_{\tau d} = \{\sigma \in S_n \mid \sigma(d_{ij}) = d_{ij'}\}$ e $C_{\tau d} = \{\sigma \in S_n \mid \sigma(d_{ij}) = d_{i'j}\}$ os subgrupos de S_n que preservam os conjuntos de elementos nas linhas e, respectivamente, nas colunas de τd. Em outras palavras, as permutações em $L_{\tau d}$ fixam o primeiro índice e as de $C_{\tau d}$ fixam o segundo índice de d_{ij}.*

Exemplo 9.3.12 *Sejam $d = (2,1)$ uma decomposição de 3 e $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12)$, então $\tau d = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1} \end{array}$. Logo, $L_{\tau d} = \{1, (23)\}$ e $C_{\tau d} = \{1, (12)\}$ onde 1 é a permutação idêntica.*

Teorema 9.2 *KS_n é uma álgebra semi-simples.*

Demonstração. O teorema de Maschke ([13] pp. 26) garante que, como a ordem de S_n é $n! < \infty$ e a característica de K, que é zero, não divide a ordem de S_n, KS_n é uma álgebra semi-simples. ■

Claro que submódulos e quocientes de módulos semi-simples também são semi-simples.

Definição 9.3.13 *Dada uma tabela de Young τd, definimos os elementos*

$$e_{\tau d} = \sum_{\substack{\rho \in L_{\tau d} \\ \sigma \in C_{\tau d}}} (-1)^{\sigma} \rho\sigma \in S_n \quad e \quad f_{\tau d} = \sum_{\substack{\rho \in L_{\tau d} \\ \sigma \in C_{\tau d}}} (-1)^{\sigma} x_{\rho\sigma(1)} \cdots x_{\rho\sigma(n)} \in P_n.$$

Quando $\tau = 1$ escreveremos somente e_d e f_d.

Teorema 9.3 *Seja $charK = 0$. Os S_n-sub-módulos $KS_n e_{\tau d}$ de KS_n e $KS_n f_{\tau d}$ de P_n são isomorfos e irredutíveis. Se M é um S_n-módulo irredutível, $\dim M < \infty$, então M é isomorfo a algum $KS_n e_{\tau d}$.*

Demonstração. A demonstração das duas afirmações do teorema pode ser encontrada por exemplo no livro [5]. ■

Exemplo 9.3.14 *Consideremos d e τd como no exemplo* (9.3.12) e teremos $e_{\tau d} = 1 - (12) + (23) - (132)$ e $f_{\tau d}(x_1, x_2, x_3) = x_1x_2x_3 - x_2x_1x_3 + x_1x_3x_2 - x_3x_2x_1$.

Recordamos que os S_n-módulos KS_n e P_n são isomorfos via $\varphi\colon KS_n \to P_n$, $\varphi(\sigma) = x_{\sigma(1)} \ldots x_{\sigma(n)}$. Observe que $\varphi(e_{\tau d}) = f_{\tau d}$. Acrescentando à nossa observação o teorema (9.2), vemos que a álgebra de grupo $A = KS_n$, como S_n-módulo, pode ser representada na forma

$$A = \sum_{\tau d} Ae_{\tau d}$$

onde $Ae_{\tau d}$ são sub-módulos simples de A. Do isomorfismo φ acima, temos

$$P_n = \sum_{\tau d} Af_{\tau d}.$$

Resumiremos então estas afirmações e acrescentaremos mais alguns fatos da representação do grupo S_n, que poderão ser verificadas em [5], no seguinte teorema:

Teorema 9.4 *Se K é um corpo de característica zero, então:*

1. *O S_n-módulo P_n decompõe-se como uma soma direta de sub-módulos irredutíveis: $P_n = \oplus k_d KS_n f_d$. Aqui d percorre todas as decomposições de n, e a multiplicidade k_d de $KS_n f_d$ é igual a $\dim KS_n f_d$.*

2. *$Af_{\tau d} \cong Af_{\tau' d'}$ se, e somente se $d = d'$.*

3. *Qualquer sistema de identidades multi-lineares de grau n é equivalente (veja definição (2.3.7)) a um sistema de identidades da forma $f_{\tau d} = 0$, $\tau d \in T$ onde T é um conjunto formado por algumas tabelas de Young correspondentes a n.*

4. *Se $n \neq 4, 6$ então P_n é isomorfo a um S_n-sub-módulo de KS_n correspondendo às tabelas de Young, exceto as associadas a decomposições da forma $n = 1 + 1 + \cdots + 1$ e $n = n$. Na verdade, esses diagramas devem ser descartados para $n > 2$.*

Teorema 9.5 (Fórmula do Gancho) [16] Seja d uma decomposição de n. Então a dimensão b_d de um S_n-submódulo irredutível em P_n é $b_d = \dfrac{n!}{\prod_{i,j} hij}$ que é igual ao número de tabelas standard τd com entradas dos números $\{1, \ldots, n\}$.

Se identificarmos em $f_{\tau d}$ as variáveis cujos índices correspondem as entradas de uma mesma linha em τd como a mesma variável, obteremos um polinômio $g_{\tau d}$ dependendo de m variáveis, onde m é o número de linhas em d. Claro que $g_{\tau d}$ é uma conseqüência de $f_{\tau d}$ e vice-versa pois $f_{\tau d}$ é a linearização completa de $g_{\tau d}$ (veja a seção (3.1)).

Exemplo 9.3.15 *Novamente considere d e τd como no exemplo* (9.3.12) já vimos que $f_{\tau d}(x_1, x_2, x_3) = x_1x_2x_3 - x_2x_1x_3 + x_1x_3x_2 - x_3x_2x_1$. Identificaremos $x_1 = x$ e $x_2 = x_3 = y$ e obteremos o polinômio $g_{\tau d} = g(x, y) = yx^2 - xyx + yx^2 - xyx = 2(yx^2 - xyx)$.

Seja $\tau = 1$ e sejam $l_1, l_2, \ldots, l_t$ os comprimentos das colunas respectivas de d. Então $g_{\tau d}$ é obtido, pelo rearranjo dos parênteses da forma indicada na definição do produto normado à direita, como um produto de polinômios standard

$$S_{l_1}(x_1, \ldots, x_{l_1})S_{l_2}(x_1, \ldots, x_{l_2}) \cdots S_{l_t}(x_1, \ldots, x_{l_t}).$$

Vamos agora descrever as estruturas de PL_n para alguns valores pequenos de n.

Vejamos para $n = 3$. O único diagrama possível é: $[d] =$  . Logo, pela fórmula do gancho, $b_d = \dim P_3 = \dfrac{3!}{3.1.1} = 2$ é a dimensão do correspondente S_3-módulo. Tomando então $\tau = 1$ vemos que $L_{\tau d} = L_d = \{1, (13)\}$, $C_{\tau d} = C_d = \{1, (12)\}$, $e_d = 1 - (12) + (13) - (123)$ e $f_d = x_1x_2x_3 - x_2x_1x_3 + x_3x_2x_1 - x_2x_3x_1$. Identificando $x_3 = x_1$ temos $g_d = x_1x_2x_1 - x_2x_1^2 + x_1x_2x_1 - x_2x_1^2 = 2(x_1x_2x_1 - x_2x_1^2) = 2(x_1x_2 - x_2x_1)x_1 \neq 0$. Consideremos agora $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (23)$. Então teremos $L_{\tau d} = \{1, (12)\}$, $C_{\tau d} = \{1, (13)\}$, $e_{\tau d} = 1 - (13) + (12) - (132)$ e $f_{\tau d} = x_1x_2x_3 - x_3x_2x_1 + x_2x_1x_3 - x_3x_1x_2$. Identificando x_2 por x_1 e x_3 por x_2 temos $g_{\tau d} = x_1^2x_2 - x_2x_1^2 + x_1^2x_2 - x_2x_1^2 = 2(x_1^2x_2 - x_2x_1^2) \neq 0$. Assim $x_1^2x_2 = 0$ é a única identidade de grau 3. Claro que $x_1x_2x_3 = 0$ é uma identidade equivalente a esta última.

Para $n = 4$, $\dim P_4 = 3! = 6$. Os diagramas possíveis têm as formas

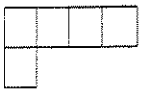
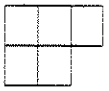
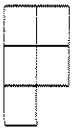
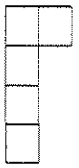
$$[d_1] = \text{  }, \quad [d_2] = \text{  } \quad \text{e} \quad [d_3] = \text{  }.$$

As dimensões dos módulos correspondentes são $b_{d_1} = \dfrac{4!}{4.2.1.1} = 3$, $b_{d_3} = \dfrac{4!}{4.1.2.1} = 3$ e $b_{d_2} = \dfrac{4!}{3.2.2.1} = 2$. Tomando, por exemplo, as tabelas

$$\tau d_1 = \begin{array}{|c|c|c|}\hline 3 & 1 & 2 \\ \hline 4 \\ \cline{1-1} \end{array} \quad \text{e} \quad d_3 = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 \\ \cline{1-1} 3 \\ \cline{1-1} \end{array}$$

obteremos identidades da forma $g_{\tau d_1} = 2x_1^3 x_2 \neq 0$ e $g_{d_3} = s_3(ad\ x_1, x_2, ad\ x_3)(x_1) \neq 0$. Então d_2 não corresponde a uma componente simples em P_4. Assim podemos escrever $P_4 = Af_{\tau d_1} \oplus Af_{d_3}$ com $Af_{\tau d_1}$ não isomorfo a Af_{d_3}.

Para $n = 5$, $\dim P_5 = 4! = 24$. Os diagramas possíveis têm as formas

$[d_1] =$  , $[d_2] =$  , $[d_3] =$  , $[d_4] =$  e $[d_5] =$  .

As dimensões dos módulos correspondentes são $b_{d_1} = b_{d_5} = 4$, $b_{d_2} = b_{d_4} = 5$ e $b_{d_3} = 6$.

Pelo teorema (9.4) todos os módulos correspondentes entram na decomposição de P_5. Agora nós temos $24 = 4 + 5 + 6 + 5 + 4$. Obteremos identidades associadas usando as seguintes tabelas:

$$\tau_1 d_1 = \begin{array}{|c|c|c|c|}\hline 4 & 1 & 2 & 3 \\ \hline 5 \\ \cline{1-1}\end{array}\ ,\ g_{\tau_1 d_1} = 2x_1^4 x_2;$$

$$\tau_2 d_2 = \begin{array}{|c|c|c|}\hline 1 & 4 & 3 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}\ ,\ g_{\tau_2 d_2} = 2(x_1 x_2)x_1^2 x_2;$$

$$\tau_3 d_3 = \begin{array}{|c|c|c|}\hline 2 & 1 & 5 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array}\ ,\ g_{\tau_3 d_3} = x_1 s_3(ad\ x_1, ad\ x_2, ad\ x_3)x_1;$$

$$\tau_4 d_4 = d_4 = \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \cline{1-2} 3 \\ \cline{1-1}\end{array}\ ,\ g_{d_4} = 2s_3(ad\ x_1, ad\ x_2, ad\ x_3)x_1 x_2 \quad e$$

$$\tau_5 d_5 = d_5 = \begin{array}{|c|c|}\hline 1 & 5 \\ \hline 2 \\ \cline{1-1} 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array}\ ,\ g_{d_5} = s_4(ad\ x_1, ad\ x_2, ad\ x_3, ad\ x_4)x_1.$$

Cada um desses polinômios é não nulo em $L\{X\}$ pois, por exemplo, eles não se anulam quando calculados em sl_3. Desta forma, podemos escrever

$$P_5 = Af_{\tau_1 d_1} \oplus Af_{\tau_2 d_2} \oplus Af_{\tau_3 d_3} \oplus Af_{d_4} \oplus Af_{d_5}.$$

Como todos estes módulos são não isomorfos entre si, qualquer submódulo é igual a soma direta de alguns destes submódulos. Logo, qualquer sistema de identidades homogêneas de grau 5 é equivalente a precisamente um sistema da forma $\{g_{\tau d} = 0 \mid \tau d \in T\}$ onde $T \subseteq \{\tau_1 d_1, \tau_2 d_2, \tau_3 d_3, d_4, d_5\}$. O número total de tais sistemas (variedades) é 32.

9.4 Uma Base de Identidades para a Álgebra de Lie $sl_2(K)$

Nesta seção pretendemos provar a existência de uma base finita para as identidades da álgebra de Lie $gl_2(K)$ onde K é um corpo de característica zero. Este teorema falha no

caso em que K tem característica 2 (veja [35]). Como no caso de característica diferente de 2, $var\ gl_2(K) = var\ sl_2(K)$, trataremos predominantemente de $sl_2(K)$. O fato sobre as variedades geradas por gl_2 e por sl_2 ocorre pois sobre um corpo com característica distinta de dois, qualquer matriz pode ser representada como uma soma de uma matriz escalar e uma matriz com traço zero. As identidades que nós estamos considerando são polinômios de Lie (i.é. comutadores), logo elas se transformam em igualdades verdadeiras quando substituídas por 1 em lugar de qualquer uma das suas variáveis.

Definição 9.4.1 1. *Uma K-álgebra A é dita uma álgebra envelopante para uma K-álgebra de Lie L se L é uma subálgebra de Lie em $A^{(-)}$ e A, considerada como uma K-álgebra associativa, é gerada pelo subespaço L.*

2. *O par de objetos formado por uma álgebra de Lie L e sua álgebra envelopante associativa A é chamado um par de Lie (álgebra associativa, álgebra de Lie); ou simplesmente, um par, e é denotado por (A, L).*

Definição 9.4.2 *Seja (A, L) um par de Lie. Qualquer relação da forma $f(x_1, \ldots, x_n) = 0$, $f \in K\{X\}$ com f não nulo, é dita uma identidade fraca do par (A, L) se para qualquer escolha de $l_1, \ldots, l_n \in L$ tivermos $f(l_1, \ldots, l_n) = 0$.*

Definição 9.4.3 *Se A é uma álgebra qualquer e $x_1, \ldots, x_n \in A$ então definiremos os comutadores longos por indução:*

$$[x_1, x_2] = x_1x_2 - x_2x_1, \quad [x_1, x_2, x_3] = [x_1, [x_2, x_3]] \quad e$$
$$[x_1, x_2, \ldots, x_n] = [x_1, [x_2, \ldots, x_n]], n \geq 3$$

Exemplo 9.4.4 *Definiremos $a \circ b = ab + ba$. Veremos que*

$$[x \circ y, z] = 0 \tag{9.2}$$

é uma identidade fraca de $(M_2(K), sl_2(K))$. Considere

$$x = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \quad e \quad y = \begin{pmatrix} d & e \\ f & -d \end{pmatrix}$$

pertencentes a $sl_2(K)$. Então

$$x \circ y = \begin{pmatrix} 2ad + bf + ce & 0 \\ 0 & 2ad + bf + ce \end{pmatrix} = (2ad + bf + ce) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Portanto $x \circ y$ é uma matriz escalar (quer dizer, múltiplo da matriz identidade), logo comuta com qualquer matriz $z \in sl_2(K)$.

Lema 9.4.5 *As seguintes identidades são conseqüências da identidade* (9.2)*. Isto é, elas valem para quaisquer pares que satisfaçam* (9.2)*.*

$$[x, y, z] = 2(x \circ y)z - 2(x \circ z)y \tag{9.3}$$

$$x(y \circ [z, u]) = (x \circ y)[z, u] - (x \circ z)[y, u] + (x \circ u)[y, z] \tag{9.4}$$

$$4(x \circ y)[z, u] = [z, x, y, u] + [z, y, x, u] - [x, u, y, z] - [y, u, x, z] \tag{9.5}$$

Demonstração. Para a primeira destas identidades observe que para qualquer álgebra associativa temos

$$[x, y, z] = (x \circ y) \circ z - (x \circ z) \circ y.$$

Logo,

$$[x, y, z] = (x \circ y)z + z(x \circ y) - (x \circ z)y - y(x \circ z) = 2(x \circ y)z - 2(x \circ z)y.$$

Para a identidade (9.4), se multiplicarmos por 2 ambos os lados teremos

$$2x(y \circ [z, u]) = 2(x \circ y)[z, u] - 2(x \circ z)[y, u] + 2(x \circ u)[y, z].$$

Trocando o primeiro termo do lado direito para o esquerdo, teremos

$$2x(y \circ [z, u]) - 2(x \circ y)[z, u] = -2(x \circ z)[y, u] + 2(x \circ u)[y, z].$$

Usando a equação (9.2), vemos que a expressão acima equivale a

$$2(y \circ [z, u])x - 2(x \circ y)[z, u] = -2[y, (x \circ z), u] + 2[y, (x \circ u), z].$$

Agora, usando a equação (9.3) para o lado esquerdo, temos

$$2(y \circ [z, u])x - 2(x \circ y)[z, u] = [y, [z, u], x] = [y, x, u, z].$$

Para o lado direito, usando as equações (9.2) e (9.3), temos

$$-2[y, (x \circ z)u] + 2[y, (x \circ u)z] = [y, 2(x \circ u)z - 2(x \circ z)u] = [y, [x, u, z]].$$

Mostramos assim a igualdade entre os lados, e a identidade (9.4) está demonstrada.

Para a identidade (9.5) reescrevemos cada termo do lado direito, usando a equação (9.3), da seguinte forma

$$[z, x, y, u] = [z, [x, y, u]] = [z, 2(x \circ y)u - 2(x \circ u)y] = 2[z, (x \circ y), u] - 2[z, (x \circ u), y]$$

$$[z, y, x, u] = [z, [y, x, u]] = [z, 2(y \circ x)u - 2(y \circ u)x] = 2[z, (y \circ x), u] - 2[z, (y \circ u), x]$$

$$-[x, u, y, z] = [-x, [u, y, z]] = [-x, 2(u \circ y)z - 2(u \circ z)y] = 2[x, (u \circ z), y] - 2[x, (u \circ y), z]$$

$$-[y, u, x, z] = [-y, [u, x, z]] = [-y, 2(u \circ x)z - 2(u \circ z)x] = 2[y, (u \circ z), x] - 2[y, (u \circ x), z]$$

Somando estas expressões e organizando em termos semelhantes, lembrando sempre da anti comutatividade e que $x \circ y$ é central, obteremos

$$2[z, (x \circ y), u] - 2[z, (x \circ u), y] + 2[z, (y \circ x), u] - 2[z, (y \circ u), x] +$$

$$+2[x, (u \circ z), y] - 2[x, (u \circ y), z] + 2[y, (u \circ z), x] - 2[y, (u \circ x), z] =$$

$$= 2(x \circ y)[z, u] - 2(x \circ u)[z, y] + 2(x \circ y)[z, u] - 2(y \circ u)[z, x] +$$

$$+2(u \circ z)[x, y] + 2(y \circ u)[z, x] - 2(u \circ z)[x, y] + 2(x \circ u)[z, y] = 4(x \circ y)[z, u],$$

provando a identidade (9.5). ■

Teorema 9.6 *Qualquer identidade fraca do par* $(M_2(K), sl_2(K))$ *sobre um corpo de característica zero é uma conseqüência da identidade* (9.2).

Demonstração. Já vimos no exemplo (9.4.4) que a equação (9.2) é uma identidade fraca deste par.

Sejam $f(x_1, x_2, x_3, \ldots, x_l) \in K\{X\}$ um polinômio multi-linear e $f' = f(x_2, x_1, x_3, \ldots, x_l)$. Usando indução sobre l, veremos que existe um polinômio multi-linear $v(y_0, y_1, y_2, \ldots, y_{l-2})$ tal que a identidade

$$f - f' = v([x_1, x_2], x_3, \ldots, x_l) \tag{9.6}$$

é uma conseqüência da equação (9.2). Seja C a subálgebra de $K\{X\}$ gerada pelos elementos $x_i \circ x_j$ e $x_i \circ [x_j, x_m]$ com $1 \leq i, j, m \leq l$ cujas imagens em $M_2(K)$ são centrais. Observe que, como f é multi-linear, usando as equações (9.2) à (9.4), podemos escrever f como uma combinação C-linear de comutadores de grau 1 e/ou 2 em $x_1, x_2, \ldots, x_l$. Isso é verdade pois de acordo com as equações (9.2) e (9.3), qualquer comutador é igual a uma combinação linear de comutadores de grau 1 e 2 (módulo a identidade (9.2)). Portanto, é necessário mostrar que o mesmo vale para o produto da forma $x_i[x_j, x_m]$. Mas, veja que podemos escrever

$$x_i[x_j, x_m] = \frac{1}{2}x_i \circ [x_j, x_m] + \frac{1}{2}[x_i, [x_j, x_m]] = \frac{1}{2}x_i \circ [x_j, x_m] + (x_i \circ x_j)x_m - (x_i \circ x_m)x_j$$

pois

$$\frac{1}{2}x_i \circ [x_j, x_m] + \frac{1}{2}[x_i, [x_j, x_m]] = \frac{1}{2}x_i[x_j, x_m] + \frac{1}{2}[x_j, x_m]x_i + \frac{1}{2}x_i[x_j, x_m] - \frac{1}{2}[x_j, x_m]x_i$$

e para a segunda igualdade, pela identidade (9.3), temos

$$(x_i \circ x_j)x_m - (x_i \circ x_m)x_j = \frac{1}{2}[x_i, x_j, x_m] = \frac{1}{2}[x_i, [x_j, x_m]].$$

Esse argumento e a equação (9.4) mostram que para demonstrar a existência da conseqüência da forma (9.6) basta considerar f com uma das seguintes formas:

1. $f = [x_1, x_2]$,
2. $f = x_1 \circ x_2$,
3. $f = (x_1 \circ x_3)x_2$,
4. $f = [x_1, x_3] \circ x_2$.

No caso (1), $f - f' = [x_1, x_2] - [x_2, x_1] = 2[x_1, x_2]$, então basta tomar $v = v(y_0) = 2y_0$. No caso (2), $f - f' = (x_1 \circ x_2) - (x_2 \circ x_1) = 0$, então tomaremos $v = 0$. No caso (3), $f - f' = (x_1 \circ x_3)x_2 - (x_2 \circ x_3)x_1 = (x_3 \circ x_1)x_2 - (x_3 \circ x_2)x_1 = \frac{1}{2}[x_3, x_1, x_2]$, e assim temos $v = v(y_0, y_1) = \frac{1}{2}[y_1, y_0]$. No caso (4), observamos que $0 = [x, y \circ z] = [x, y] \circ z + x \circ [y, z]$ de onde concluímos que $f = [x_1, x_3] \circ x_2 = x_3 \circ [x_1, x_2]$ e assim

$$f - f' = x_3 \circ [x_1, x_2] - x_3 \circ [x_2, x_1] = x_3[x_1, x_2] + [x_1, x_2]x_3 - x_3[x_2, x_1] - [x_2, x_1]x_3 =$$

$$= 2x_3[x_1, x_2] + 2[x_1, x_2]x_3 = 2(x_3[x_1, x_2] + [x_1, x_2]x_3) = 2(x_3 \circ [x_1, x_2]).$$

Então, basta tomar $v = v(y_0, y_1) = 2(y_1 \circ y_0)$.

Agora, suponhamos que $f = 0$ é uma identidade fraca do nosso par. Então o mesmo acontece para f'=0 e, conseqüentemente, $v([x_1, x_2], x_3, \ldots, x_l) = 0$ também é uma identidade fraca de $(M_2(K), sl_2(K))$. Visto que $sl_2(K)$ é igual a sua subálgebra derivada então $v(y_1, y_2, \ldots, y_{l-1}) = 0$ é uma identidade fraca também. Por uma indução, com base óbvia, esta identidade é uma conseqüência da identidade (9.2). Assim, determinamos que $f - f' = 0$ é uma conseqüência da identidade (9.2). Como a escolha do par (x_1, x_2) é arbitrária, ela pode ser feita para qualquer permutação $\sigma \in S_l$ e portanto

$$f(x_1, x_2, \ldots, x_l) - f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(l)}) = 0$$

é uma conseqüência da identidade (9.2). Somando todas essas equações obtemos que

$$l!f(x_1, x_2, \ldots, x_l) - \lambda \sum_{\sigma \in S_l} x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(l)} = 0,$$

com um apropriado $\lambda \in K$, é uma conseqüência da identidade (9.2). Substituindo

$$x_1 = x_2 = \cdots = x_l = h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

deveremos ter $\lambda l! h^l = 0$ apenas no caso em que $\lambda = 0$. Então $f(x_1, x_2, \ldots, x_l) = 0$ é uma conseqüência da identidade (9.2). ■

Vamos observar agora algumas identidades da álgebra de Lie sl_2. Veja que $G = sl_2$ tem dimensão 3. Logo ela satisfaz a identidade standard

$$[\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_{n-1}, x_n] = 0, \quad n > 4, \tag{9.7}$$

onde as barras sobre $x_1, x_2, \ldots, x_{n-1}$ querem dizer que o lado esquerdo da equação (9.7) é uma soma alternada, em relação às variáevis com barras, de tais comutadores.

Mostraremos que para qualquer s, sl_2 satisfaz

$$(ad\ x)^{2s+1}([y, z]) = [(ad\ x)^{2s+1}(y), z] + [y, (ad\ x)^{2s+1}(z)]. \tag{9.8}$$

Realmente, pela equação 9.5 temos

$$4x^2[x, y] = (ad\ x)^3(y), \tag{9.9}$$

pois

$4(x \circ x)[x, y] = [x, x, x, y] + [x, x, x, y] - [x, y, x, x] - [x, y, x, x]$ e assim

$4(2x^2)[x, y] = 2[x, x, x, y]$ de onde segue o que queríamos.

Então obtemos que

$$(ad\ x)^{2s+1}([y, z]) = 2^{2s}x^{2s}[x, [y, z]] =$$
$$= [2^{2s}x^{2s}[x, y], z] + [y, 2^{2s}x^{2s}[x, z]]$$

$$= [(ad\ x)^{2s+1}(y), z] + [y, (ad\ x)^{2s+1}(z)].$$

Sabemos do teorema (3.4) que a álgebra $M_2(K)$ não satisfaz identidades de graus menores que 4. A única identidade de grau 4, a menos de múltiplo, é s_4. Mas esta última não é polinômio de Lie. Portanto sl_2 não satisfaz identidades de graus menores que 5.

Observaremos que, de acordo com a seção (9.3), o espaço PL_5 de todos os polinômios comutadores multi-lineares (i.é., polinômios de Lie com respeito a colchetes) é igual a soma direta de 5 S_5-módulos irredutíveis:

$$PL_5 = Af_{\tau_1 d_1} \oplus Af_{\tau_2 d_2} \oplus Af_{\tau_3 d_3} \oplus Af_{\tau_4 d_4} \oplus Af_{\tau_5 d_5}.$$

Denotando por $P'(V)$, $V = var(G)$, o conjunto de todas as identidades multi-lineares de grau 5 válidas em V, mostraremos que

$$P'_5(V) = Af_{\tau_3 d_3} \oplus Af_{d_5}.$$

Realmente, de acordo com a seção (9.3) cada S_5-sub-módulo em P_5 é igual a uma soma de alguns dos $Af_{\tau d}$, $\tau d \in T$. Neste caso é suficiente mostrar a validade das identidades da forma $g_{t_i d_i} = 0$ em G. Veja então que $g_{d_5} = [\overline{y}_1, \overline{y}_2, \overline{y}_3, \overline{y}_4, y_1] = 0$ é uma conseqüência da identidade (9.7) com $n = 5$. Além disso,

$$g_{\tau_3 d_3} = 2[y_1, [y_1, y_3], y_1, y_2] = \frac{1}{3}(2(ad\ y_1)^3([y_3, y_2]) - 2[(ad\ y_1)^3(y_3), y_2] - 2[y_3, (ad\ y_1)^3(y_2)]).$$

Então $g_{\tau_3 d_3}$ é uma conseqüência da identidade (9.8) com $s = 1$. Agora, temos

$$g_{\tau_1 d_1} = 2(ad\ x)^4(y), \quad g_{\tau_4 d_4} = 2[\overline{y}_1, \overline{y}_2, \overline{y}_3, y_1, y_2], \quad g_{\tau_2 d_2} = 2[[y_1, y_2], y_1, y_1, y_2].$$

Escolhendo em $sl_2(K)$ uma base da forma

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

com multiplicação $[h, e] = 2e$, $[h, f] = -2f$ e $[e, f] = h$, vemos que $g_{\tau_1 d_1} = 32e$ para $x = h$ e $y = e$. Além disso, $g_{\tau_4 d_4} = 8h$ para $y_1 = e$, $y_2 = f$ e $y_3 = h$ e $g_{\tau_2 d_2} = -8e$ para $y_1 = e$ e $y_2 = f$.

Precisaremos também das seguintes identidades de grau 6:

$$[x, x, x, y, z, u] = [x, y, z, x, x, u] - [x, y, x, u, x, z] \tag{9.10}$$

e

$$[z, u, x, x, x, y] = [z, x, x, u, x, y] - [x, [u, x, y], x, z]. \tag{9.11}$$

Elas seguem da identidade (9.9) e de $4x^2[y, z] = [y, x, x, z] - [x, z, x, y]$, que são conseqüências da identidade (9.5). Vejamos:

$$[x, x, x, y, z, u] = 4x^2[x, y, z, u] = [x, y, 4x^2[z, u]] = [x, y, z, x, x, y] - [x, y, x, u, x, z]$$

e

$$[z, u, x, x, x, y] = 4x^2[z, u, x, y] = [z, x, x, u, x, y] - [x, [u, x, y], x, z].$$

O resultado principal dessa seção é o próximo teorema, que afirmará que a identidade (9.7) com $n = 5$ e a identidade (9.8) com $s = 2$ formam uma base para as identidades de $sl_2(K)$, onde K é um corpo arbitrário de característica zero. A prova deste teorema usa o mesmo método empregado para o teorema (9.6). Uma parte importante desta desmonstração será dada em separado como um lema. Neste lema usaremos a notação $[\ldots, \widetilde{x}_{i_1}, \ldots, \widetilde{x}_{i_2}, \ldots, \widetilde{x}_{i_t}, \ldots]$ para abreviar a soma $\sum_{\sigma \in S_l} [\ldots, x_{i_{\sigma(1)}}, \ldots, x_{i_{\sigma(2)}}, \ldots, x_{i_{\sigma(t)}}, \ldots]$. Pretendemos também escrever $\sigma_{ij} f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_l) = f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_l)$. Finalmente, usaremos $\widehat{\ }$ para dizer que a variável que está sob esse sinal será omitida assim como fizemos na seção (3.1).

Lema 9.4.6 *Seja V a variedade das álgebras de Lie determinada pelas identidades*

$$[\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, x_5] = 0 \quad e \quad (ad\ x)^3([y,z]) = [(ad\ x)^3(y), z] + [y, (ad\ x)^3(z)].$$

Seja $f \in PL_l$ um polinômio multi-linear de grau $l > 4$. Então para quaisquer $1 \leq i, j \leq n$, existe $v \in PL_{l-1}$ de grau $l-1$ tal que

$$f - \sigma_{ij} f = c([x_i, \widetilde{x}_1, \ldots, \widehat{x}_i, \ldots, \widehat{x}_j, \ldots, \widetilde{x}_l, x_j] - \tag{9.12}$$
$$-[x_j, \widetilde{x}_1, \ldots, \widehat{x}_i, \ldots, \widehat{x}_j, \ldots, \widetilde{x}_l, x_i]) + v(x_1, \ldots, [x_i, x_j], \ldots, \widehat{x}_j, \ldots, x_l])$$

módulo as identidades de V. No caso de l ímpar podemos tomar $c = 0$.

Demonstração. Usaremos indução por l. Sem perda de generalidade podemos assumir que $i = 1$ e $j = l$. Visto que, na seção (9.3), vimos que uma base para PL_l pode ser composta por monômios da forma $[x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(l-1)}, x_l]$, podemos nos restringir ao caso em que f é um monômio da forma $f = [x_{\sigma(1)}, \ldots, x_1, \ldots, x_{\sigma(l-1)}, x_l]$.

Se $l = 5$ podemos considerar f em uma das formas

$$[x_2, x_3, x_4, x_1, x_5],\ [x_2, x_3, x_1, x_4, x_5],\ [x_2, x_1, x_3, x_4, x_5],\ [x_1, x_2, x_3, x_4, x_5].$$

Em ambos, o primeiro e segundo casos, a desejada representação com $c = 0$ resulta da aplicação da anti-comutatividade e da identidade de Jacobi. No primeiro caso

$$[x_2, x_3, x_4, x_1, x_5] - [x_2, x_3, x_4, x_5, x_1] = [x_2, x_3, x_4, [x_1, x_5]] + [x_2, x_3, x_4, [x_1, x_5]] =$$

$$= 2[x_2, x_3, x_4, [x_1, x_5]].$$

No segundo caso

$$[x_2, x_3, x_1, x_4, x_5] - [x_2, x_3, x_5, x_4, x_1] = [x_2, x_3, [x_1, x_4, x_5] - [x_5, x_4, x_1]] =$$

$$= [x_2, x_3, [x_1, x_4, x_5] + [x_5, x_1, x_4]].$$

Usando a identidade de Jacobi, $[x_1, x_4, x_5] + [x_5, x_1, x_4] = -[x_4, x_5, x_1]$, temos

$$[x_2, x_3, x_4, x_1, x_5] - [x_2, x_3, x_4, x_5, x_1] = -[x_2, x_3, [x_4, x_5, x_1]] = [x_2, x_3, x_4, x_1, x_5]$$

que tem a forma desejada.

Antes de passarmos aos casos restantes, observaremos que qualquer polinômio de Lie de grau 5 que é uma conseqüência da identidade (9.2) é uma identidade de $sl_2(K)$ de grau 5. Então, pela decomposição de $P'(V)$ vista anteriormente, isso vale para as identidades no enunciado desse lema. Isto quer dizer que determinando uma representação como na equação (9.12) no caso de grau 5 poderemos usar a identidade fraca (9.2) e suas conseqüências.

No terceiro caso faremos

$$[x_2,x_1,x_3,x_4,x_5]-[x_2,x_5,x_3,x_4,x_1=2[x_2,x_1,(x_3\circ x_4)x_5]-2[x_2,x_1,(x_3\circ x_5)x_4]-$$

$$-2[x_2,x_5,(x_3\circ x_4)x_1]+2[x_2,x_5,(x_3\circ x_1)x_4]$$

$$=4(x_3\circ x_4)[x_2,x_1,x_5]-[x_2,2(x_3\circ x_5)x_1-2(x_3\circ x_1)x_5,x_4]$$

$$=4(x_3\circ x_4)[x_2,x_1,x_5]-[x_2,[x_3,x_5,x_1],x_4].$$

Aplicando a equação (9.5) com $x=x_3$, $y=x_4$, $z=x_2$ e $u=[x_1,x_5]$, obteremos uma soma de comutadores que contém $[x_1,x_5]$ como um dos seus fatores, obtendo a desejada decomposição. Finalmente, no quarto caso, aplicando a equação (9.3), podemos escrever

$$[x_1,x_2,x_3,x_4,x_5]-[x_5,x_2,x_3,x_4,x_1]=2(x_3\circ x_4)([x_1,x_2,x_5]-[x_5,x_2,x_1])-$$

$$-2(x_3\circ x_5)[x_1,x_2,x_4]+2(x_3\circ x_1)[x_5,x_2,x_4]=$$

$$=2(x_3\circ x_4)[x_2,x_1,x_5]+[2(x_3\circ x_1)x_5-2(x_3\circ x_5)x_1,x_2,x_4]$$

$$=2(x_3\circ x_4)[x_2,x_1,x_5]+[[x_3,x_1,x_5],[x_2,x_4]].$$

Aplicando a equação (9.5) como acima obtemos uma representação da forma (9.12) com $c=0$.

Agora passamos ao caso $l\geq 6$. Notamos primeiro que um comutador da forma $f=[x_1,x_2,\ldots,x_{l-1},x_l]$ é representável como uma soma de $f_0=[x_1,\widetilde{x}_2,\ldots,\widetilde{x}_{l-1},x_l]$ mais uma soma de comutadores da forma $f_i=[x_1,\ldots,[x_i,x_{i+1}],\ldots,x_{l-1},x_l]$ (a menos da reenumeração de $x_2,\ldots,x_{l-1}$). Se l é ímpar então para $f=f_0$ nós temos a equação (9.12) com $c=1$ e $v=0$. Para $f=f_i$ a hipótese de indução aplica-se e então $c=0$.

Antecipando o teorema 9.7, observamos que ele deverá mostrar no decorrer de sua prova que qualquer identidade de $sl_2(K)$ da forma $f=0$ com uma representação da forma (9.12) é uma conseqüência de identidades de grau inferior. Assim, determinando uma representação da forma (9.12) para identidades de graus ≥ 7, podemos aplicar as identidades (9.10) e (9.11) de grau 6. Voltando à nossa demonstração, lembramos que estamos considerando o caso onde l é um número ímpar com $l\geq 7$. Aqui a hipótese de indução aplica-se aos comutadores da forma $[x_2,x_3,[\ldots,x_1,\ldots,x_{l-1},x_l]]$, basta considerar um fator da forma $[\ldots,x_1,\ldots,x_{l-1},x_l]$ de grau $l-2$. No caso de $[x_2,x_1,x_3,\ldots,x_{l-1},x_l]$, pela hipótese de indução, escrevemos

$$f-\sigma_{1l}f=c([x_2,x_1,\widetilde{x}_3,\ldots,\widetilde{x}_{l-1},x_l]-[x_2,x_l,\widetilde{x}_3,\ldots,\widetilde{x}_{l-1},x_1])+[x_2,v'([x_1,x_l],x_3,\ldots,x_{l-1})].$$

Temos assim a forma desejada. Será proveitoso observar que isto é a linearização do polinômio da forma

$$c([x_2,x_1,\underbrace{y,\ldots,y}_{l-3\text{ vezes}},x_l]-[x_2,x_l,\underbrace{y,\ldots,y}_{l-3\text{ vezes}},x_1]).$$

Ambos os termos podem ser transformados usando a seguinte conseqüência da identidade (9.11):

$$[z,u,x,\underbrace{x,\ldots,x}_{2s \text{ vezes}},y] = [z,x,x,u,\underbrace{x,\ldots,x}_{2s-1 \text{ vezes}},y] - [x,[u,\underbrace{x,\ldots,x}_{2s-1 \text{ vezes}},y],x,z]$$

de onde obtemos

$$[x_2,x_1,\widetilde{x}_3,\ldots,\widetilde{x}_{l-1},x_l] = [x_2,\widetilde{x}_3,\widetilde{x}_4,x_1,\widetilde{x}_5,\ldots,\widetilde{x}_{l-1},x_l] - [\widetilde{x}_4,[x_1,\widetilde{x}_5,\ldots,\widetilde{x}_{l-2},x_l],\widetilde{x}_{l-1},x_2].$$

Agora a hipótese de indução aplica-se a $[\widetilde{x}_3,\widetilde{x}_4,x_1,\widetilde{x}_5,\ldots,\widetilde{x}_{l-1},x_l]$ bem como a apropriados fatores de grau $l-1$ na última soma simétrica de comutadores (use a identiadde de Jacobi!). O que nos resta é considerar os seguintes casos:

1. $f = [x_1,\widetilde{x}_2,\ldots,\widetilde{x}_{l-1},x_l]$. Aqui, podemos aplicar a linearização completa da identidade (9.8) com $2s+1 = l-2$ (que é uma conseqüência de (9.8) com $s=1$). Ou seja, a linearização completa da identidade

$$(ad\ x)^{l-2}([y,z]) = [(ad\ x)^{l-2}(y),z] + [y,(ad\ x)^{l-2}(z)]$$

que é igual à linearização completa de

$$[\underbrace{x,\ldots,x}_{l-2\ vezes},[y,z]] = [[\underbrace{x,\ldots,x}_{l-2\ vezes},y],z] + [y,\underbrace{x,\ldots,x}_{l-2\ vezes},z],$$

ou ainda,

$$[y,\underbrace{x,\ldots,x}_{l-2\ vezes},z] = -[[\underbrace{x,\ldots,x}_{l-2\ vezes},y],z] + [\underbrace{x,\ldots,x}_{l-2\ vezes},[y,z]].$$

2. $f = [x_1,x_2,\ldots,[x_i,x_{i+1}],\ldots,x_{l-1},x_l]$. Neste caso, pela hipótese de indução, escrevemos

$$\begin{aligned} f - \sigma_{1l}f = {} & c'([x_1,\widetilde{x}_2,\ldots,[x_i,x_{i+1}],\ldots,\widetilde{x}_{l-1},x_l] - \\ & -[x_l,\widetilde{x}_2,\ldots,[x_i,x_{i+1}],\ldots,\widetilde{x}_{l-1},x_1]) + \\ & +v'([x_1,x_l],x_2,\ldots,[x_i,x_{i+1}],\ldots,x_{l-1}) \end{aligned} \tag{9.13}$$

Temos assim a forma desejada.

Para transformarmos o anterior podemos usar uma conseqüência de (9.10) que tem a forma

$$[\underbrace{x,\ldots,x}_{t \text{ vezes}},y,z,u] = [\underbrace{x,\ldots,x}_{t-2 \text{ vezes}},y,z,x,x,u] - [\underbrace{x,\ldots,x}_{t-2 \text{ vezes}},y,x,u,x,z].$$

Aplicando esta identidade repetidamente obteremos o seguinte:

$$[\underbrace{x,\ldots,x}_{t \text{ vezes}},y,z,u] = \sum_i [x,x,y,w_i'] + \sum_j [x,y,u_j',v_j'] \tag{9.14}$$

onde os u_j', v_j' e w_i' são alguns comutadores em x,z,u. Reescrevemos

$$g = [x_1,\widetilde{x}_2,\ldots,[x_i,x_{i+1}],\ldots,\widetilde{x}_{l-1},x_l]$$

como uma combinação linear de comutadores da forma $[\ldots,[x_i,x_{i+1}]]$. Agora g é simétrico em $x_2,\ldots,x_{i-1},x_{i+2},\ldots,x_{l-1}$. Aplicando a linearização da equação (9.14) vamos obter

$$g=\sum_i \pm[x_1,x_p,x_q,x_l,w_i]+\sum_j \pm[x_1,x_p,x_l,u_j,v_j]+\sum_s \pm[x_1,x_l,a_s,b_s,c_s],$$

a_s,b_s,c_s,u_j,w_i, comutadores nas variáveis distintas de x_1,x_l. Então g pode ser manuseado como um comutador de grau 5, provando assim a existência da representação (9.12) também neste caso. ■

Teorema 9.7 *As identidades da álgebra de Lie $G=sl_2(K)$ sobre um corpo K de característica zero, admitem uma base da forma*

$$[\overline{x}_1,\overline{x}_2,\overline{x}_3,\overline{x}_4,x_5]=0 \tag{9.15}$$

$$(ad\ x)^3([y,z])=[(ad\ x)^3(y),z]+[y,(ad\ x)^3(z)]. \tag{9.16}$$

Demonstração. Começaremos mostrando que se $f=0$ é uma identidade multi-linear de grau l em G, então na equação (9.12) poderemos sempre tomar $c=0$. Para isso é suficiente substituir $x_i=e$, $x_j=f$ e $x_t=h$ (para todo $t\neq i,j$). Visto que l é par e $f=0$ é uma identidade, o lado esquerdo anula-se e nós temos

$$0=c(l-2)!2^{l-1}h+v(h,\ldots,[e,f],\ldots,h).$$

Porém $[e,f]=h$. Como v é um polinômio comutador $v(h,h\ldots,h)=0$, de onde obtemos que $c(l-2)!2^{l-1}=0$ e concluímos que $c=0$. Portanto para qualquer identidade polinomial $f=0$ de grau $l\geq 6$ e para qualquer par i,j; $1\leq i,j\leq l$, existe um polinômio comutador multi-linear v de grau $l-1$ tal que

$$f-\sigma_{i,j}f=v(x_1,\ldots,[x_i,x_j],\ldots,\widehat{x}_j,\ldots,x_l)$$

é uma conseqüência das identidades (9.15) e (9.16). Visto que $f-\sigma_{i,j}f=0$ é uma identidade de G e $G=G^2$, então $v=0$ é uma identidade de G de grau $l-1$. Procedendo por indução, $v=0$ segue-se das identidades (9.15) e (9.16). Então também $f-\sigma_{i,j}f=0$ é uma conseqüência das identidades (9.15) e (9.16), logo, deduzimos que para qualquer permutação $\tau\in S_l$, $f-\tau f=0$ é uma conseqüência das identidades (9.15) e (9.16). Tomando a soma de todas essas identidades obtemos que

$$n!f=\sum_{\tau\in S_l}\tau f$$

é uma conseqüência das identidades (9.15) e (9.16). Agora, o lado direito é um múltiplo da soma $\sum_{\tau\in S_l}[x_{\tau(1)},x_{\tau(2)},\ldots,x_{\tau(l)}]$. Basta então mostrar que essa soma é nula. Consideremos $l\geq 2$ e as duas últimas entradas no comutador em cada parcela desta somatória, ou seja, vamos escrever para cada parcela, $[\ldots,x_{\tau(i)},x_{\tau(j)}]$. Veja que é possível rearrumar as parcelas em pares tal que cada par tenha a seguinte forma:

$$[x_{\tau(i_1)},x_{\tau(i_2)},\ldots,x_{\tau(i_{l-2})},x_{\tau(i)},x_{\tau(j)}]+[x_{\tau(i_1)},x_{\tau(i_2)},\ldots,x_{\tau(i_{l-2})},x_{\tau(j)},x_{\tau(i)}].$$

Nestes pares, as duas últimas entradas nos comutadores estão invertidas e as restantes são iguais. A soma de cada par desses, por causa da anti-comutatividade, é identicamente nula em G visto que teremos

$$[x_{\tau(i_1)}, x_{\tau(i_2)}, \ldots, x_{\tau(i_{l-2})}, x_{\tau(i)}, x_{\tau(j)}] + [x_{\tau(i_1)}, x_{\tau(i_2)}, \ldots, x_{\tau(i_{l-2})}, x_{\tau(j)}, x_{\tau(i)}] =$$

$$= [x_{\tau(i_1)}, x_{\tau(i_2)}, \ldots, x_{\tau(i_{l-2})}, [x_{\tau(i)}, x_{\tau(j)}] + [x_{\tau(j)}, x_{\tau(i)}]] = 0.$$

Assim, $\sum_{\tau \in S_l} [x_{\tau(1)}, x_{\tau(2)}, \ldots, x_{\tau(l)}] = 0$. ■

9.5 Uma Base de Identidades para a Álgebra das Matrizes de Ordem Dois

O objetivo principal dessa seção é a demonstração do teorema seguinte.

Teorema 9.8 *Se o corpo base K é de característica zero, então o conjunto de todas as identidades da álgebra M_2 das matrizes de ordem 2 é equivalente às identidades de grau 4, 5 e 6 que são válidas em M_2.*

Demonstraremos que se a característica do corpo K é diferente de dois e todas as identidades multi-lineares de Lie da álgebra $sl_2(K)$ são equivalentes a um conjunto finito de identidades de Lie, então todas as identidades da álgebra associativa M_2 são equivalentes a um conjunto finito de identidades de M_2. Portanto, o teorema deverá seguir do teorema (9.7) e do fato que sobre um corpo de característica zero, qualquer conjunto de identidades de M_2 é equivalente a algum conjunto de identidades multi-lineares de M_2.

Seja $K\{X\}$ a álgebra associativa livre com geradores $x_1, x_2, \ldots$. Nosso objetivo é determinar um conjunto finito de identidades P satisfeitas em M_2 tais que todas as identidades multi-lineares da álgebra M_2 são conseqüências de P. Este conjunto deverá ser completamente determinado no decorrer da prova do teorema (9.8). Mais a frente, notaremos que P consistirá das identidades

$$[\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, x_5] = 0,$$

e

$$(ad\ x)^3([y, z]) = [(ad\ x)^3(y), z] + [y, (ad\ x)^3(z)],$$

que são, respectivamente, as identidades (9.15) e (9.16). Das identidades multi-lineares

$$\sum_{\sigma \in S_4} (-1)^\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} x_{\sigma(4)} = 0, \tag{9.17}$$

$$[x \circ y, z] = 0,$$

e

$$4(x \circ y)[z, u] = [z, x, y, u] + [z, y, x, u] - [x, u, y, z] - [y, u, x, z],$$

onde x e y são comutadores de comprimento 2 e as duas últimas são, respectivamente, as identidades (9.2) e (9.5). Além dessas, P consistirá de outras duas identidades multi-lineares de grau 6, que são duais, em um apropriado sentido, as identidades (9.17) e (9.2).

Lema 9.5.1 *As identidades*

$$[x \circ y, z] = 0,$$
$$[\overline{x}_1, \overline{x}_2, \overline{x}_3, \overline{x}_4, x_5] = 0,$$
$$(ad\ x)^3([y,z]) = [(ad\ x)^3(y), z] + [y, (ad\ x)^3(z)] \quad e$$
$$\sum_{\sigma \in S_4} (-1)^\sigma x_{\sigma(1)} x_{\sigma(2)} x_{\sigma(3)} x_{\sigma(4)} = 0,$$

que são, respectivamente (9.2), (9.15), (9.16) e (9.17) valem na álgebra M_2.

Demonstração. Sobre um corpo com característica distinta de dois, qualquer matriz de segunda ordem representa-se como uma soma de uma matriz escalar e uma matriz com traço zero. As identidades que nós estamos considerando são próprias (i.é. produtos de comutadores), temos que elas se transformam em igualdades verdadeiras quando substituídas por uma unidade em lugar de qualquer uma de suas variáveis. Assim, é suficiente verificar que todas essas identidades são satisfeitas no par $(M_2, sl_2(K))$. Mas, (9.2) já foi vista no exemplo (9.4.4). Temos também que (9.15) é uma conseqüência da identidade (9.7) com $n = 5$ e (9.16) é uma conseqüência da identidade (9.8) com $s = 2$. Finalmente para (9.17), veja que sl_2 tem dimensão 3, e essa é a identidade standard de grau 4. O lema está demonstrado. ■

Lema 9.5.2 *Módulo as identidades (9.2) e (9.5), qualquer polinômio a da forma* $a = u_1 \dots u_n$, *onde os* u_i *são comutadores nos geradores* $x_1, x_2, \dots$ *de comprimento maior ou igual do que* 2, é representável na forma

$$a = c + v, \tag{9.18}$$

onde v é um polinômio de Lie, $c = \sum_i \beta_i(u_i' \circ [x_i, x_t])$, u_i' são comutadores nos geradores $x_1, x_2, \dots$ de comprimentos ≥ 2 e podemos tomar para x_t qualquer gerador que participa em a.

Demonstração. Faremos por indução sobre n; sendo $n = 1$ é óbvio. Seja $n > 1$, então

$$a = u_1 \dots u_{n-1} u_n = u_1 \dots u_{n-2} \{\frac{1}{2}(u_{n-1} \circ u_n) + \frac{1}{2}[u_{n-1}, u_n]\}$$

Se $n > 2$, então a segunda parcela reduz-se à forma que nós necessitamos para a hipótese de indução, e a identidade (9.5) implica que $u_{n-2}(u_{n-1} \circ u_n)$ é um polinômio de Lie e o primeiro termo é também reduzível a forma (9.18). Seja $n = 2$. Então

$$a = \frac{1}{2}\{(u_1 \circ u_2) + [u_1, u_2]\}.$$

De (9.2) obtemos a identidade

$$([x,z] \circ y) = (x \circ [z,y]). \tag{9.19}$$

Como qualquer comutador dependendo de x_t pode ser representado como uma combinação linear de comutadores normados à direita da forma $[\dots, x_t]$, vemos que $(u_1 \circ u_2)$(e portanto a também) pode ser reduzido a forma (9.18) com o auxílio de (9.19). ■

Se $f(x_i, \dots, x_l)$ é um polinômio, então $f|_{x_i = a}$ denotará o polinômio obtido de f pela substituição de x_i por a, por exemplo, $x_3 x_1 x_2|_{x_1 = a} = x_3 a x_2$.

Lema 9.5.3 *Seja A uma álgebra associativa com unidade. O conjunto de todas as identidades multi-lineares satisfeitas por A é equivalente ao subconjunto de identidades multi-lineares da forma $f(x_1, \ldots, x_l) = 0$, tais que $f|_{x_1=1}, \ldots, f|_{x_l=1}$ são polinômios nulos.*

Demonstração. Seja $g(x_1, \ldots, x_l) = 0$ uma identidade multi-linear da álgebra A, onde $g|_{x_1=1}, \ldots, g|_{x_i=1}$ são polinômios nulos, $t \leq l$. Suponha $g|_{x_{t+1}=1} \neq 0$. Então $f = g - g|_{x_{t+1}=1} x_{t+1} = 0$ é uma conseqüência de g e também é uma identidade da álgebra A. Claro que f e g são equivalentes como identidades. Mas $f|_{x_i=1} = 0$ para $1 \leq i \leq t+1$. Procedemos por indução e obtemos a afirmação do Lema. ■

Seja $f(x_1, \ldots, x_l) \in P_l$ e suponha que $f|_{x_1=1}, \ldots, f|_{x_l=1}$ são os polinômios nulos. Então f é uma combinação linear de elementos da forma $u_1 \ldots u_t$ onde os $u_1, \ldots, u_t$ são comutadores de comprimento maior do que 1. Os lemas (9.5.2) e (9.5.3) implicam que, módulo as identidades (9.2) e (9.5), uma base de identidades da álgebra M_2 pode ser escolhida entre as identidades da forma $c+v = 0$ nas quais $v(x_1, \ldots, x_l)$ é um polinômio de Lie multi-linear e $c = \sum_i \beta_i(u_i' \circ [x_i, x_l])$, onde cada u_i' é um polinômio de Lie multi-linear, $u' = u'(x_1, \ldots, \widehat{x_i}, \ldots, x_{l-1})$ e $\beta_i \in K$. Da identidade (9.2) segue-se que os valores do polinômio c em M_2 são matrizes escalares. É claro que os valores do polinômio de Lie v têm traço zero. Portanto, se a característica de K é diferente de dois, então uma identidade da forma $c+v = 0$ vale em M_2 se, e somente se as identidades $c = 0$ e $v = 0$ valem em M_2. Pelo teorema (9.7), a identidade $v = 0$ é uma conseqüência das identidades (9.15) e (9.16), então, módulo P, nós podemos escolher uma base de identidades para a álgebra M_2 entre as identidades $c = 0$.

Lema 9.5.4 *Seja $c = \sum_i \beta_i(u_i \circ [v_i, x_l])$ um polinômio multi-linear onde $\beta_i \in K$, u_i e v_i são comutadores em $x_1, \ldots, x_{l-1}$, e comprimento de u_i é ≥ 2. Então a identidade $c = 0$ vale na álgebra M_2 se, e somente se $\sum_i \beta_i[u_i, v_i] = 0$ é uma identidade de Lie para a álgebra de Lie $sl_2(K)$.*

Demonstração. Se $c = 0$ em M_2, então $\sum_i \beta_i(u_i \circ [v_i, x_l]) = (\sum_i \beta_i[u_i, v_i]) \circ x_l$ como uma identidade fraca. Sabemos que $(\sum_i \beta_i[u_i, v_i]) \circ x_l = 0$ qualquer que seja x_l, ocorre se, e somente se $\sum_i \beta_i[u_i, v_i] = 0$. Logo $\sum_i \beta_i[u_i, v_i] = 0$ é uma identidade de Lie para sl_2.

Reciprocamente, seja $\sum_i \beta_i[u_i, v_i] = 0$ uma identidade de Lie para a álgebra de Lie $sl_2(K)$. Então, usando o argumento, provado anteriormente, na direção oposta, nós determinaremos que $c = 0$ é uma identidade para o par $(M_2, sl_2(K))$. Visto que o polinômio c é multi-linear e ele anula-se sob a substituição por uma unidade de qualquer uma de suas variáveis, $c = 0$ é uma identidade para a álgebra M_2. ■

Na variedade de todas álgebras de Lie, vale a identidade de Jacobi $[x_1, x_2, x_3]+[x_2, x_3, x_1]+[x_3, x_1, x_2] = 0$. Portanto, pelo lema (9.5.4), em M_2 temos a identidade

$$([x_2, x_3] \circ [x_1, x_l]) + ([x_3, x_1] \circ [x_2, x_l]) + ([x_1, x_2] \circ [x_3, x_l]) = 0 \qquad (9.20)$$

que é, como pode facilmente ser verificado, a identidade standard de grau 4, veja (9.17).

Recordaremos o que é a base de Hall na álgebra livre de Lie $L\{X\}$. Definimos uma ordem total sobre os comutadores, nos geradores $x_1, \ldots, x_{l-1}$, que é compatível com o comprimento i.é., se $\deg u < \deg v$ então u precede v. Os comutadores básicos são os seguintes:

Definição 9.5.5 1. *os comutadores básicos de comprimento* 1 *são* $x_1, \ldots, x_{l-1}$;

2. *se os comutadores básicos de comprimento menor do que n estão definidos, então o comutador* $[u, v]$ *de comprimento* n *é um comutador básico se, e somente se*

   (a) u *e* v *são comutadores básicos e* $u < v$,

   (b) se $v = [v_1, v_2]$ *então* $u \geq v_1$.

É conhecido que os comutadores básicos formam uma base da álgebra de Lie livre, ver por exemplo [24, Theorem 52.2].

Lema 9.5.6 *Sejam* u *e* v *polinômios de Lie nos geradores* $x_1, \ldots, x_{l-1}$ *com comprimento de* $v \geq 2$ *e além disso, seja* $[u, v] = \sum_i \delta_i [u_i, v_i]$, *onde* $[u_i, v_i]$ *são comutadores básicos em* $x_1, \ldots, x_{l-1}$. *Então a igualdade*

$$(u \circ [v, x_l]) = \sum\nolimits_i \delta_i (u_i \circ [v_i, x_l]) \tag{9.21}$$

é uma conseqüência das identidades (9.2) e (9.20)

Demonstração. Seguiremos um dos possíveis métodos usados para demonstrar que todo polinômio de Lie apresenta-se por meio de comutadores básicos. Usaremos a identidade (9.20) no lugar da identidade de Jacobi e a identidade (9.19) na forma

$$(z_1 \circ [x_1, z_2]) = -(z_2 \circ [x_1, z_1]), \tag{9.22}$$

onde z_1 e z_2 são comutadores com comprimento maior do que 1, em lugar da identidade $[x, x] = 0$. (Como foi mencionado na prova do lema (9.5.2) a identidade (9.19) é uma conseqüência da identidade (9.2)).

Qualquer comutador representa-se como uma combinação linear de comutadores básicos do mesmo comprimento. Usamos (9.22), então basta mostrar o lema quando u e v são comutadores básicos, $u < v$ e o comprimento de v é ≥ 2.

Demonstraremos a igualdade (9.21) por indução sobre o comutador básico v. A base de indução é o caso em que $[u, v]$ é um comutador básico.

Suponha que para comutadores da forma $[w_1, w_2]$ onde o comprimento é igual ao comprimento de $[u, v]$, o comprimento de w_2 é ≥ 2, $w_2 > w_1 > u$ e, w_1 e w_2 são comutadores básicos, a afirmação do lema já está demonstrada. Vamos provar essa afirmação para o comutador $[u, v]$. Visto que o comprimento de v é ≥ 2, temos $v = [v_1, v_2]$ e pela definição de comutadores básicos, v_1 e v_2 são comutadores básicos e $v_1 < v_2$. Se $v_1 \leq u$, então $[u, v]$ é um comutador básico e assim está demonstrado. Suponha então que $v_1 > u$. Logo $[u, [v_1, v_2]] = [v_2, [v_1, u]] - [v_1, [v_2, u]]$ e a identidade (9.20) implicará que

$$([v_1, v_2] \circ [u, x_l]) = ([v_1, u] \circ [v_2, x_l]) - ([v_2, u] \circ [v_1, x_l])$$

portanto é suficiente provar a afirmação do lema para os comutadores $[v_2, [v_1, u]]$ e $[v_1, [v_2, u]]$.

Consideraremos cada um desses casos. Seja $[v_1, u] = \sum_i \delta_i w_i$, onde os w_i são comutadores básicos. Então, de $v_1 < v_2$ vemos que o comprimento de w_i é igual ao comprimento de $[v_1, u]$, que é maior do que o comprimento de v_2. Logo $w_i > v_2$ para todo i. Por indução, usando que $v_2 > u$, o lema é válido para comutadores $[w_i, v_2]$ e, assim também será para $[v_2, [v_1, u]]$.

Seja $[v_2, u] = \sum_i \varepsilon_i w_i'$, onde os w_i' são comutadores básicos. Se $w_i' = v_1$, então $[w_i', v_1] = 0$ e a afirmação do lema é válida nesse caso por (9.22). Se $w_i' > v_1$, então da relação $v_2 > v_1 > u$ segue-se que o lema é válido, por indução, para o comutador $[w_i', v_1] = 0$. Se $w_i' < v_1$, então o comprimento de w_i' é igual ao comprimento de $[v_2, u]$, que é maior do que o comprimento de u e portanto, $w_i' > u$ e o lema vale para o comutador $[v_1, w_i']$. Além disso, por (9.22), também vale para $[w_i', v_1]$. Desta forma, a afirmação do lema esta demonstrada também para o comutador $[v_1, [v_2, u]]$. O lema está demonstrado. ■

Corolário 9.5.7 *Suponha que na álgebra de Lie livre com geradores livres $x_1 \ldots, x_{l-1}$ a igualdade $\sum_i \delta_i [u_i, v_i] = 0$ vale, onde u_i e v_i são comutadores e o comprimento dos u_i é ≥ 2. Então a identidade $\sum_i \delta_i (u_i \circ [v_i, x_l]) = 0$ segue das identidades* (9.2) e (9.20).

Demonstração. Como os comutadores básicos formam uma base na álgebra de Lie livre, o elemento $\sum_i \delta_i [u_i, v_i] = 0$ pode ser representado como uma combinação linear de comutadores básicos com coeficientes zero. Assim, neste caso a nossa afirmação segue-se do lema (9.5.6). O corolário está demonstrado. ■

Vamos transferir todos os termos das identidades de Lie (9.15), (9.16) para o lado esquerdo escrevendo-as então na forma $f_1 = 0$ e $f_2 = 0$ respectivamente. Vamos escrever os mesmos polinômios de Lie na forma $f_i = \sum_j \beta_{ij} [u_{ij}, v_{ij}]$, onde os u_{ij} e v_{ij} são comutadores e o comprimento dos u_{ij} é ≥ 2. Então, pelo lema (9.5.4), as identidades $f_i' = \sum_j \beta_{ij} (u_{ij} \circ [v_{ij}, x_6]) = 0$ para $i = 1, 2$ são identidades para a álgebra M_2.

Foi mostrado anteriormente que, módulo as identidades (9.2), (9.5), (9.15) e (9.16) uma base de identidades da álgebra M_2 pode ser escolhida entre as identidades da forma $c = 0$, onde $c = \sum_i \beta_i (u_i \circ [v_i, x_l])$ e u_i são comutadores com comprimento ≥ 2. Do lema (9.5.4), segue-se que $\sum_i \beta_i [u_i, v_i] = 0$ é uma identidade de Lie em $sl_2(K)$. Do teorema 9.7 concluímos que na álgebra de Lie livre temos a igualdade

$$\sum_i \beta_i [u_i, x_i] = \sum_{ij} \delta_{ij} f_{ij},$$

onde o polinômio de Lie f_{ij} é obtido do f_i pela substituição de alguns polinômios de Lie no lugar dos seus argumentos. Se f_{ij}' é o polinômio associativo obtido de $f_i'|_{x_6 = x_l}$ pela mesma substituição como em f_i, então, do corolário do lema (9.5.6), vemos que a igualdade

$$\sum_i \beta_i (u_i \circ [x_i, x_l]) = \sum_{ij} \delta_{ij} f_{ij}'$$

é uma conseqüência das identidades (9.17) e (9.2). As identidades $f_{ij}' = 0$ são conseqüências das identidades $f_1' = 0$ e $f_2' = 0$. Portanto, provamos que todas as identidades para a álgebra M_2 segue-se do conjunto de identidades

$$P = \{(9.2), (9.5), (9.15), (9.16), (9.17), f_i = 0 (i = 1, 2)\}.$$

O teorema está demonstrado.

Ressaltamos que mais tarde em [7] foi demonstrado que as identidades de $M_2(K)$ seguem de duas, a saber a identidade standard s_4 e a identidade de Hall (veja exemplo (2.1.6)). A demonstração em [7] foi obtida analisando-se a estrutura dos polinômios multi-lineares.

Bibliografia

[1] S. A. Amtsur, J. Levitzki, *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1**, 449–463 (1950).

[2] M. F. Atiyah, I. G. Macdonald, *Introduction to Conmutative Algebra*, Addison-Wesley, Reading, Mass. (1969), 81–83, 90.

[3] S. S. de Azevedo, *Identidades Polinomiais em Álgebras*, Universidade Estadual de Campinas-Unicamp, (1999).

[4] Yu. A. Bahturin, *Identical Relations in Lie Algebras*, VNU Science Press, Utrecht, (1987).

[5] C. Curtis e I. Reiner, *The representation theory of finite groups and associative algebras*, Wiley–Interscience, (1962).

[6] R. P. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. Math. **51**, 161–166 (1950).

[7] V. Drensky, *A minimal basis for the identities of a second-order matrix algebra over a field of characteristic 0*, Algebra and Logic **20**, 188–194 (1980).

[8] V. Drensky, *Free Algebras and PI-Algebras*, Springer, (2000).

[9] E. Formanek, *Central polynomials for matrix rings*, J. Algebra, **23**, 129–133 (1972).

[10] E. S. Golod, *On nil-algebras and residually finite p-groups*, Amer. Math. Soc. Transl. **48** (2) (1965).

[11] E. S. Golod e I. R. Shafarevich, *On class field towers*, Amer. Math. Soc. Trans. **48** (2) (1965).

[12] M. Hall, *Projective Planes*, Trans. Amer. Math. Soc. **54**, 229–277 (1943).

[13] I. N. Herstein, *Noncommutative Rings*, Carus Math. Monographs **15**, Wiley and Sons, Inc., New York, (1968), 39–40, 46–47, 188, 94–96.

[14] N. Jacobson, *PI-Algebras: An Introduction*, Lecture Notes in Math. **441**, Springer, Berlin-New York, (1975).

[15] N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. Math. **46**, 695–707 (1945).

[16] G. James e A. Kerber, *The representation theory of the symmetric group*, Encyclopaedia Math. and its Applications **16**, Addison–Wesley (1981).

[17] I. Kaplansky, *On a problem of Kurosh and Jacobson*, Bull. Amer. Math. Soc. **52**, 496–500 (1946).

[18] B. Kostant, *A theorem of Frobenius, a theorem of Amitsur–Levitzki and cohomolohy theory*, J. Math. Mech. **7** (1958).

[19] A. G. Kurosh, *Ringtheorestische Probleme, die mit dem Burnsideschen Probleme über periodiche Gruppen in Zusammenghang stehen*, Bull. Acad. Sci. URSS Ser. Math. **5**, 223–240 (1941).

[20] J. Levitzki, *On a problem of Kurosh*, Bull. Amer. Math. Soc. **52**, 1033–1035 (1946).

[21] C. Processi, *Rings With Polynomial Identities*, Marcel Dekker, New York, (1973).

[22] Yu. P. Razmyslov, *On a problem of Kaplansky*, Math. USSR, Izv. **7**, 479–496 (1973).

[23] Yu. P. Razmyslov, *Trace identities of full matrix algebras over a field of characteristic zero*, Math. USSR, Izv. **8**, 727–760 (1974).

[24] Yu. P. Razmyslov, *Identities of Algebra and Their Representations*, Translation: Translations of Math. Monographs **138**, AMS, Providence, R. I., (1994).

[25] A. Regev, *Existence of identities in* $A \otimes B$, Israel J. Math. **11**, 131–152 (1972).

[26] S. Rosset, *A new proof of the Amitsur–Levitzki identity*, Israel J. Math. **23**, 187–188 (1976).

[27] L. H. Rowen, *Polynomial Identities of Ring Theory*, Acad. Press, (1980).

[28] L. A. B. San Martin, *Álgebras de Lie*, Editora da Unicamp (1999).

[29] A. I. Shirshov, *On certain nonassociative nil rings and algebraic algebras*, Amer. Math. Soc. Trans. **119** (2), 119–132 (1983).

[30] A. I. Shirshov, *On rings with polynomial identities*, Amer. Math. Soc. Transl. **119** (2), 133–139 (1983).

[31] W. Specht, *Gesetze in Ringen I*, Math. Z. **52**, 557–589 (1950).

[32] R. G. Swan, *An application of graph theory to algebra*, Proc. Amer. Math. Soc. **14**, 363–373 (1963).

[33] R. G. Swan, *Correction to "An application of graph theory to algebra"*, Proc. Amer. Math. Soc. **21**, 379–380 (1969).

[34] J. Szigeti, Z. Tuza, G. Révész, *Eulerian Polynomial identities on matrix rings*, J. Algebra **161**, 90–101 (1993).

[35] M. R. Vaughan-Lie, *Varieties of Lie algebras*, D. Phil. Thesis, Oxford (1970).

[36] N. Vonessen, *Rings with polynomial identities an elementary introduction*, Matemática Contemporânea **7**, 199–231 (1994).

[37] K.A. Zhevlakov, A.M. Slin'ko, I.P. Shestakov, A.I. Shirshov, *Rings That Are Nearly Associative*, Translation: Academic Press, New York, (1982).