Identidades Polinomiais em Álgebras

Sérgio Sardinha de Azevedo

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Sérgio Sardinha de Azevedo e aprovada pela comissão julgadora.

Campinas, 26 de fevereiro de 1999.

Prof. Dr. Plamen Emilov Kochloukov
Orientador

Banca Examinadora:

- Prof. Dr. Plamen Emilov Kochloukov
- Prof. Dr. Paulo Roberto Brumatti
- Prof. Dr. Antonio José Engler

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas, como requisito parcial para obtenção do título de Mestre em Matemática.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

991885

Universidade Estadual de Campinas
1999

UNIDADE BC
N.º CHAMADA:
T UNICAMP
AZ25i
V. Ex.
TOMBO BC/37831
PROC. 229/99
C D X
PREÇO R$ 11,00
DATA 10/06/99
N.º CPD

CM-00124522-6

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP

Azevedo, Sérgio Sardinha de

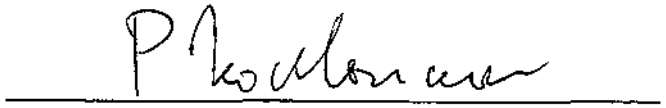
Az25i Identidades polinomiais em álgebras / Sérgio Sardinha de Azevedo -- Campinas, [S.P. :s.n.], 1999.

Orientador : Plamen E. Koshlukov

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

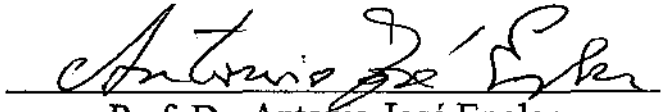
1. Polinômios. 2. Álgebra. I. Koshlukov, Plamen Emilov. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Dissertação de Mestrado defendida e aprovada em 26 de fevereiro de 1999 pela Banca Examinadora composta por



Prof. Dr. Plamen Emilov Kochloukov

Prof. Dr. Paulo Roberto Brumatti



Prof. Dr. Antonio José Engler

Resumo

O tema desta dissertação é a teoria de álgebras com identidades polinomiais. O texto introduz as primeiras definições necessárias para o estudo de tal teoria, bem como alguns dos resultados mais importantes que surgiram nas últimas décadas.

Apresentamos identidades minimais em álgebras matriciais (Teorema de Amitsur–Levitzki). Além disso, provamos que o produto tensorial de duas PI álgebras é uma PI álgebra (Teorema de Regev), o Teorema da Altura com algumas aplicações e o Teorema de Nagata–Higman sobre a nilpotência de nil álgebras. Discutimos também os Problemas de Kurosh e Specht. A maioria destes resultados são aplicações de métodos combinatórios (noção de altura, representações do grupo simétrico) à teoria de álgebras com identidades polinomiais.

Abstract

The subject of this dissertation is the theory of algebras with polynomial identities. The text introduces the first definitions that are needed to study this theory, as well as some of the most important results that appeared in the last decades.

We present minimal identities of matricial algebras (Theorem of Amitsur–Levitzki). Moreover, we prove that the tensor product of any two PI algebras is a PI algebra (Theorem of Regev), the Theorem on Height with some of their applications, and the Theorem of Nagata–Higman about nilpotency of nil algebras. We discuss too the Problems of Kurosh and Specht. Most of these results are applications of combinatorial methods (notion of height, representations of symetric groups) to the theory of algebras with polynomial identities.

Aos meus pais,
Maria José e Sérgio.

"O correr da vida embrulha tudo, a vida é assim: esquenta e esfria, aperta e daí afrouxa, sossega e depois desinqueta. O que ela quer da gente é coragem. O que Deus quer é ver a gente aprendendo a ser capaz de ficar alegre a mais, no meio da alegria, e ainda mais no meio da tristeza."

Grande Sertão: Veredas
João Guimarães Rosa

Índice

Apresentação xi

1 Estruturas algébricas 1
- 1.1 Estruturas simples . . . 1
- 1.2 Grupos . . . 1
- 1.3 Anéis . . . 2
- 1.4 Espaços vetoriais . . . 3
- 1.5 Álgebras . . . 4
- 1.6 Módulos . . . 5

2 Representações de grupos 7
- 2.1 Álgebra do grupo . . . 7
- 2.2 Alguns homomorfismos e isomorfismos . . . 8
- 2.3 Representações de grupos . . . 9
- 2.4 Relação entre representações e módulos . . . 11
- 2.5 Somas diretas em grupos abelianos . . . 13
- 2.6 Lema de Zorn . . . 14
- 2.7 Redutibilidade de módulos . . . 14
- 2.8 Simplicidade de álgebras . . . 16

3 Álgebras com identidade polinomial 19
- 3.1 Alfabeto, letra e palavra . . . 19
- 3.2 Polinômios . . . 20
- 3.3 Álgebras com identidade polinomial . . . 21
- 3.4 Polinômios multilineares . . . 23
- 3.5 Identidades multilineares . . . 24

4 Teorema de Amitsur-Levitzki 29
- 4.1 Identidades polinomiais na álgebra das matrizes . . . 29
- 4.2 Um teorema sobre grafos . . . 30
- 4.3 Teorema de Amitsur-Levitzki . . . 38

5 Teorema de Higman 39
- 5.1 Propriedade da base finita . . . 39
- 5.2 Teorema de Higman . . . 46

5.3 Algumas aplicações combinatórias 49
5.4 Um teorema sobre permutações 51

6 PI álgebras sobre Q **53**
6.1 T-ideais 53
6.2 Codimensões 54
6.3 Tabelas de Young 56
6.4 Representações de S_n 58
6.5 Alguns teoremas de Regev 60
6.6 Relações de ordem em $\mathbf{X}^*$ 61
6.7 Problema de Specht 62

7 Produto tensorial de PI álgebras **65**
7.1 Produto tensorial de espaços vetoriais 65
7.2 Produto tensorial de álgebras 69
7.3 Teorema do Produto Tensorial 70
7.4 Outra prova do Teorema de Amitsur-Levitzki 72

8 Teorema da Altura **76**
8.1 Problema de Kurosh 76
8.2 Lema de Shirshov 78
8.3 Teorema da Altura 81

9 Nilpotência de Álgebras Associativas **85**
9.1 Introdução 85
9.2 Teorema de Nagata-Higman 87

Apresentação

A teoria de álgebras com identidades polinomiais começou a se desenvolver mais intensamente nos últimos cinqüenta anos. Até 1945, havia poucos resultados conhecidos nesta área e um escasso número de artigos publicados. Entre eles, destacam-se os trabalhos de M. Dehn (1922), W. Wagner (1936) e M. Hall (1943), motivados pela geometria. A. G. Kurosh publicou em 1941 um artigo que posteriormente estimulou as pesquisas de alguns matemáticos.

A partir de 1945, aprofundaram-se os conhecimentos sobre as identidades polinomiais em álgebras, devido sobretudo aos trabalhos de Nathan Jacobson e Irving Kaplansky. Em 1950, foi demonstrado um importante resultado desta teoria: o Teorema de Amitsur-Levitzki. Nos anos seguintes, surgiram vários resultados provindos de matemáticos como Higman, Nagata, Shirshov, Regev, Razmyslov, Procesi, Rowen, Cohn, Posner, Vaughan-Lee, Herstein, Formanek, Kostrikin, Kemer e Braun, entre outros.

Quanto ao texto, procurei deixá-lo bastante simples, pensando sempre que poderia servir como um primeiro contato com a teoria de identidades polinomiais para aqueles que se interessassem. Reconheço que em algumas partes, sobretudo dentro de algumas demonstrações, cheguei a escrever o óbvio, porém sempre motivado pela idéia de deixar o texto bastante claro. As provas de proposições, lemas, teoremas e corolários que são muito triviais foram omitidas.

Gostaria de agradecer a todos aqueles que me ajudaram na realização deste trabalho. Entre outros, agradeço especialmente ao Prof. Plamen Kochloukov, meu orientador, pela dedicação e objetividade na condução desta pesquisa. Faço constar também os meus agradecimentos aos amigos Edson Agustini e Fabio Jacon, pelo auxílio na formatação do texto, e ao Antonio Batista de Jesus, por vários favores e sugestões referentes a trâmites acadêmicos.

Campinas, janeiro de 1999.

Capítulo 1

Estruturas algébricas

Neste primeiro capítulo definimos várias estruturas algébricas (grupos, anéis, espaços vetoriais, álgebras etc.) utilizadas posteriormente. O texto é bem simples, tendo como objetivo apenas deixar claras as características e propriedades de cada um destes conjuntos.

1.1 Estruturas simples

Definição 1.1.1 *Um conjunto não-vazio G, no qual está definida uma operação $*$ de $G \times G$ em G, é um* semigrupo *se para quaisquer a, b e c em G a seguinte condição é satisfeita:*
(i) $(a * b) * c = a * (b * c)$.

Definição 1.1.2 *Um conjunto não-vazio G, no qual está definida uma operação $*$ de $G \times G$ em G, é um* monóide *se para quaisquer a, b e c em G as seguintes condições são satisfeitas:*
(i) $(a * b) * c = a * (b * c)$;
(ii) existe em G um elemento denotado por e e chamado identidade *tal que* $a * e = e * a = a$.

1.2 Grupos

Definição 1.2.1 *Um conjunto não-vazio G, no qual está definida uma operação $*$ de $G \times G$ em G, é um* grupo *se para quaisquer a, b e c em G as seguintes condições são satisfeitas:*
(i) $(a * b) * c = a * (b * c)$;
(ii) existe em G um elemento denotado por e e chamado identidade *tal que* $a * e = e * a = a$;
(iii) existe um elemento denotado por a^{-1} e chamado inverso *de a tal que* $a^{-1} * a = a * a^{-1} = e$.

Definição 1.2.2 *Um grupo G no qual $a * b = b * a$ para quaisquer a e b em G é chamado* grupo abeliano *ou* grupo comutativo.

Definição 1.2.3 *Um subconjunto H de um grupo G é um* subgrupo *de G se H é um grupo sob a operação induzida de G. Escrevemos $H \leqslant G$.*

Proposição 1.2.1 *Se H é um subgrupo de um grupo G, então a identidade de H é igual à identidade de G.*

Prova: Sejam e a identidade de G e e' a identidade de H. Como $e' \in H$, então $e' * e = e' \in H$ e, conseqüentemente, $e' * (e' * e) = e' * e$. Escrevendo $(e')^{-1} * (e' * (e' * e)) = (e')^{-1} * (e' * e)$, podemos concluir que $e' * e = e$. Assim, $e' = e' * e = e$ e está demostrado que a identidade de H é igual à identidade de G. ■

Teorema 1.2.1 *Se H é um subconjunto não-vazio de um grupo G, então as seguintes afirmações são equivalentes:*
(i) $H \leqslant G$;
(ii) se a e $b \in H$ então $ab \in H$ e $a^{-1} \in H$;
(iii) se a e $b \in H$ então $a^{-1}b \in H$.

Prova: As demonstrações de que (i) implica (ii) e de que (ii) implica (iii) são triviais. Se a afirmação (iii) é verdadeira então $a^{-1}a = e \in H$ para qualquer $a \in H$. Logo, se $b \in H$ então $b^{-1}e = b^{-1} \in H$. Além disso, se a e b pertencem a H então $a^{-1} \in H$ e $ab = (a^{-1})^{-1}b \in H$. Portanto, H é um subgrupo de G. ■

Definição 1.2.4 *Sejam G e H dois 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 é um isomorfismo de grupos *e escrevemos $G \cong H$.*

1.3 Anéis

Definição 1.3.1 *Um conjunto não-vazio R, no qual estão definidas as operações $+$ e $\cdot$, ambas de $R \times R$ em R e chamadas adição e multiplicação respectivamente, é um* anel *se para quaisquer a, b e c em R as seguintes condições são satisfeitas:*
(i) R é um grupo abeliano com relação à adição;
(ii) $a \cdot (b + c) = a \cdot b + a \cdot c$ e $(a + b) \cdot c = a \cdot c + b \cdot c$.

Definição 1.3.2 *Um anel R no qual $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para quaisquer a, b e c em R é chamado* anel associativo.

Definição 1.3.3 *Um anel R no qual $a \cdot b = b \cdot a$ para quaisquer a e b em R é chamado* anel comutativo.

Definição 1.3.4 *Um anel R no qual existe um elemento 1 tal que $1 \cdot a = a \cdot 1 = a$ para qualquer a em R é chamado* anel com identidade. *O elemento 1 é chamado* identidade multiplicativa *de R.*

Definição 1.3.5 *Um anel com identidade R no qual para todo a em R tal que $a \neq 0$ existe um elemento a^{-1} tal que $a^{-1} \cdot a = a \cdot a^{-1} = 1$ é chamado* anel com divisão. *O elemento a^{-1} é chamado* inverso multiplicativo *de a.*

Definição 1.3.6 *Um anel associativo, comutativo e com divisão é chamado* corpo.

Definição 1.3.7 *Um subconjunto S de um anel R é um* subanel *de R se S é um anel sob as operações induzidas de R.*

Teorema 1.3.1 *Se S é um subconjunto não-vazio de um anel R, então as seguintes afirmações são equivalentes:*
(i) S é um subanel de R;
(ii) S é um subgrupo aditivo de R, se a e $b \in S$ então $ab \in S$;
(iii) se a e $b \in S$ então $a + b \in S$, $-a \in S$ e $ab \in S$;
(iv) se a e $b \in S$ então $-a + b \in S$ e $ab \in S$.

Definição 1.3.8 *Sejam R um anel e I um subgrupo aditivo de R. Se $ax \in I$ para todo $a \in R$ e $x \in I$, dizemos que I é um* ideal à esquerda *de R e escrevemos $I \lhd_e R$. Se $xa \in I$ para todo $a \in R$ e $x \in I$, dizemos que I é um* ideal à direita *de R e escrevemos $I \lhd_d R$. Se $ax \in I$ e $xa \in I$ para todo $a \in R$ e $x \in I$, dizemos que I é um* ideal *(ou* ideal bilateral*) de R e escrevemos $I \lhd R$.*

Definição 1.3.9 *Sejam R e S dois anéis. Uma aplicação $\varphi : R \to S$ é um* homomorfismo de anéis *se para quaisquer a e b em R 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 é um isomorfismo de anéis *e escrevemos $R \cong S$.*

1.4 Espaços vetoriais

Definição 1.4.1 *Seja K um corpo. Um conjunto não-vazio V, no qual estão definidas uma operação $+$ de $V \times V$ em V (chamada* adição*) e uma operação de $K \times V$ em V (chamada* multiplicação por escalar*), é um* espaço vetorial sobre K *se para quaisquer u, v e w em V e α e β em K as seguintes condições são satisfeitas:*
(i) V é um grupo abeliano com relação à adição;
(ii) $1v = v$ onde 1 é a identidade multiplicativa de K;
(iii) $(\alpha\beta)v = \alpha(\beta v)$;
(iv) $\alpha(u + v) = \alpha u + \alpha v$;
(v) $(\alpha + \beta)v = \alpha v + \beta v$.

Observação 1.4.1 *Se na definição anterior ao invés de um corpo K colocássemos um anel associativo e comutativo com identidade R, então V seria chamado um* módulo (à esquerda) sobre R.

Definição 1.4.2 *Um subconjunto W de um espaço vetorial V é um* subespaço *de V se W é um espaço vetorial sob as operações induzidas de V.*

Teorema 1.4.1 *Se W é um subconjunto não-vazio de um espaço vetorial V, então as seguintes afirmações são equivalentes:*
(i) W é um subespaço de V;
(ii) W é fechado com relação à adição e à multiplicação por escalar.

Prova: A demonstração é essencialmente a mesma do **Teorema 1.2.1**, observando que $-v = (-1)v$ para qualquer $v \in V$. ■

Definição 1.4.3 *Sejam V e W dois espaços vetoriais sobre um corpo K. Uma aplicação $\varphi : V \to W$ é uma* transformação linear *se para quaisquer u e v em V e α em K as seguintes condições são satisfeitas:*

$$\varphi(u+v) = \varphi(u) + \varphi(v) \text{ e } \varphi(\alpha v) = \alpha\varphi(v).$$

Se φ é bijetora dizemos que é um isomorfismo de espaços vetoriais *e escrevemos $V \cong W$.*

Definição 1.4.4 *Seja V um espaço vetorial. Uma transformação linear $\varphi : V \to V$ é chamada* operador linear.

1.5 Álgebras

Definição 1.5.1 *Seja K um corpo. Um espaço vetorial A sobre K, munido de uma operação $\cdot$ de $A \times A$ em A (chamada* multiplicação*), é uma* álgebra sobre K *se para quaisquer a, b e c em A e α em K as seguintes condições são satisfeitas:*
(i) $a \cdot (b+c) = a \cdot b + a \cdot c$ *e* $(a+b) \cdot c = a \cdot c + b \cdot c$;
(ii) $\alpha(a \cdot b) = (\alpha a) \cdot b = a \cdot (\alpha b)$.

Observação 1.5.1 *Se na definição anterior ao invés de um espaço vetorial sobre um corpo K colocássemos um módulo sobre um anel associativo e comutativo com identidade R, então A seria uma álgebra sobre R.*

Definição 1.5.2 *Uma álgebra A na qual $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para quaisquer a, b e c em A é chamada* álgebra associativa.

Definição 1.5.3 *Uma álgebra A na qual $a \cdot b = b \cdot a$ para quaisquer a e b em A é chamada* álgebra comutativa.

Definição 1.5.4 *Uma álgebra A na qual existe um elemento $\mathbf{1}$ tal que $\mathbf{1} \cdot a = a \cdot \mathbf{1} = a$ para qualquer a em A é chamada* álgebra com identidade. *O elemento $\mathbf{1}$ é chamado* identidade *de A.*

Definição 1.5.5 *Uma álgebra com identidade A na qual para todo a em A tal que $a \neq 0$ existe um elemento a^{-1} tal que $a^{-1} \cdot a = a \cdot a^{-1} = \mathbf{1}$ é chamada* álgebra com divisão. *O elemento a^{-1} é chamado* inverso multiplicativo *de a.*

Definição 1.5.6 *Um subconjunto B de uma álgebra A é uma* subálgebra *de A se B é uma álgebra sob as operações induzidas de A.*

Teorema 1.5.1 *Se B é um subconjunto não-vazio de uma álgebra A, então as seguintes afirmações são equivalentes:*
(i) B é uma subálgebra de A;
(ii) B é fechado com relação à adição, à multiplicação e à multiplicação por escalar.

Definição 1.5.7 *Sejam A uma álgebra e I um subespaço de A. Se $ax \in I$ para todo $a \in A$ e $x \in I$, dizemos que I é um* ideal à esquerda *de A e escrevemos $I \lhd_e A$. Se $xa \in I$ para todo $a \in A$ e $x \in I$, dizemos que I é um* ideal à direita *de A e escrevemos $I \lhd_d A$. Se $ax \in I$ e $xa \in I$ para todo $a \in A$ e $x \in I$, dizemos que I é um* ideal *(ou* ideal bilateral*) de A e escrevemos $I \lhd A$.*

Definição 1.5.8 *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:*

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

Se φ é bijetora dizemos que é um isomorfismo de álgebras *e escrevemos $A \cong B$.*

Definição 1.5.9 *Seja A uma álgebra. Um homomorfismo $\varphi : A \to A$ é chamado um* endomorfismo da álgebra A.

1.6 Módulos

Definição 1.6.1 *Sejam G um grupo, M um espaço vetorial sobre um corpo K e $\cdot$ uma operação de $G \times M$ em M. O espaço vetorial M é chamado um* G-módulo (à esquerda) sobre K *se para quaisquer g e h em G, m e n em M e α em K as seguintes condições são satisfeitas:*
(i) $e \cdot m = m$ onde e é a identidade de G;
(ii) $(gh) \cdot m = g \cdot (h \cdot m)$;
(iii) $g \cdot (m+n) = g \cdot m + g \cdot n$;
(iv) $g \cdot (\alpha m) = \alpha(g \cdot m)$.

Definição 1.6.2 *Seja G um grupo. Um subconjunto N de um G-módulo M é um* submódulo *de M se N é um G-módulo sob as operações induzidas de M.*

Teorema 1.6.1 *Seja G um grupo. Se N é um subconjunto não-vazio de um G-módulo M, então as seguintes afirmações são equivalentes:*
(i) N é um submódulo de M;
(ii) N é fechado com relação à adição, à multiplicação por escalar e à multiplicação por elementos de G.

Definição 1.6.3 *Sejam G um grupo e M e N dois G-módulos sobre um corpo K. Uma aplicação $\varphi : M \to N$ é um* homomorfismo de módulos *se para quaisquer m e n em M, α em K e g em G as seguintes condições são satisfeitas:*

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

O núcleo *de φ é o conjunto $Ker\varphi = \{m \in M : \varphi(m) = 0\}$ e a* imagem *de φ é o conjunto $\varphi(M) = \{\varphi(m) : m \in M\}$. Se φ é bijetora dizemos que é um* isomorfismo de módulos *e escrevemos $M \cong N$.*

Proposição 1.6.1 *Se G é um grupo e $\varphi : M \to N$ é um homomorfismo de G-módulos, então $Ker\varphi$ é um submódulo de M e $\varphi(M)$ é um submódulo de N. Além disso, se $Ker\varphi = \{0\}$ então φ é injetora.*

Prova: As demonstrações de que $Ker\varphi$ é um submódulo de M e de que $\varphi(M)$ é um submódulo de N são triviais. Se $Ker\varphi = \{0\}$ e $\varphi(m) = \varphi(n)$, com m e n em M, então $\varphi(m - n) = 0$ e, conseqüentemente, $m - n = 0$, ou seja, $m = n$. Portanto, φ é injetora. ■

Definição 1.6.4 *Sejam K um corpo, A uma álgebra com identidade sobre K, M um espaço vetorial sobre K e $\cdot$ uma operação de $A \times M$ em M. O espaço vetorial M é chamado um* A-módulo (à esquerda) sobre K *se para quaisquer a e b em A, m e n em M e α em K as seguintes condições são satisfeitas:*
(i) $\mathbf{1} \cdot m = m$ *onde* $\mathbf{1}$ *é a identidade de A;*
(ii) $(ab) \cdot m = a \cdot (b \cdot m)$;
(iii) $(a + b) \cdot m = a \cdot m + b \cdot m$;
(iv) $a \cdot (m + n) = a \cdot m + a \cdot n$;
(v) $\alpha(a \cdot m) = (\alpha a) \cdot m = a \cdot (\alpha m)$.

Definição 1.6.5 *Seja A uma álgebra com identidade. Um subconjunto N de um A-módulo M é um* submódulo *de M se N é um A-módulo sob as operações induzidas de M.*

Teorema 1.6.2 *Seja A uma álgebra com identidade. Se N é um subconjunto não-vazio de um A-módulo M, então as seguintes afirmações são equivalentes:*
(i) N é um submódulo de M;
(ii) N é fechado com relação à adição, à multiplicação por escalar e à multiplicação por elementos de A.

Definição 1.6.6 *Sejam A uma álgebra com identidade sobre um corpo K e M e N 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:*

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

O núcleo *de φ é o conjunto $Ker\varphi = \{m \in M : \varphi(m) = 0\}$ e a* imagem *de φ é o conjunto $\varphi(M) = \{\varphi(m) : m \in M\}$. Se φ é bijetora dizemos que é um* isomorfismo de módulos *e escrevemos $M \cong N$.*

Proposição 1.6.2 *Se A é uma álgebra com identidade e $\varphi : M \to N$ é um homomorfismo de A-módulos, então $Ker\varphi$ é um submódulo de M e $\varphi(M)$ é um submódulo de N. Além disso, se $Ker\varphi = \{0\}$ então φ é injetora.*

Prova: Semelhante à demonstração da **Proposição 1.6.1.** ■

Capítulo 2

Representações de grupos

Introduzimos os conceitos de representações de grupos, redutibilidade de módulos e simplicidade de álgebras que são utilizados no Capítulo 6. Na última seção é provado o Teorema de Maschke.

2.1 Álgebra do grupo

Definição 2.1.1 *Sejam K um corpo e G um monóide. Indicamos $K(G)$ por o espaço vetorial sobre K cuja base são todos os elementos de G.*

Se K é um corpo e G é um monóide, então

$$K(G) = \{\sum_{g\in G} \alpha_g g : \alpha_g \in K \text{ e } \alpha_g \neq 0 \text{ para um número finito de elementos } g\},$$

a soma de dois elementos de $K(G)$ é dada por

$$(\sum_{g\in G} \alpha_g g) + (\sum_{g\in G} \beta_g g) = \sum_{g\in G} (\alpha_g + \beta_g) g$$

e a multiplicação por um elemento de K é dada por

$$\delta(\sum_{g\in G} \alpha_g g) = \sum_{g\in G} (\delta\alpha_g) g,$$

Teorema 2.1.1 *Sejam K um corpo e G um monóide. Se definirmos a multiplicação de dois elementos de $K(G)$ por*

$$(\sum_{g\in G} \alpha_g g) \cdot (\sum_{g\in G} \beta_g g) = \sum_{g\in G} (\sum_{ab=g} \alpha_a \beta_b) g$$

esta operação está bem definida e $K(G)$ é uma álgebra associativa com identidade sobre K.

Prova: Trata-se de uma verificação rotineira de propriedades, tomando como identidade multiplicativa o elemento $1e$. ■

Definição 2.1.2 *Sejam K um corpo e G um grupo. A álgebra $K(G)$ é chamada* álgebra do grupo *G sobre o corpo K.*

Exemplo 2.1.1 *Consideremos o corpo dos números reais $\mathbf{R}$ e o grupo simétrico $S_3 = \{(1), (12), (13), (23), (123), (132)\}$. Neste caso, a álgebra de S_3 sobre $\mathbf{R}$ é dada por*

$$\mathbf{R}(S_3) = \{\alpha_1(1) + \alpha_2(12) + \alpha_3(13) + \alpha_4(23) + \alpha_5(123) + \alpha_6(132) : \alpha_i \in \mathbf{R}, 1 \leq i \leq 6\}.$$

Talvez, pela definição dada no ***Teorema 2.1.1****, a multiplicação de elementos da álgebra do grupo pode não ter parecido muito intuitiva. No entanto, o que devemos fazer é apenas aplicar a distributividade ao multiplicar dois elementos. Assim, a multiplicação dos elementos $a = 4(1) - 7(13) + 12(123)$ e $b = 3(12) - 8(132)$ é dada por*

$$\begin{aligned} a \cdot b &= (4(1) - 7(13) + 12(123)) \cdot (3(12) - 8(132)) = \\ &= 12(12) - 32(132) - 21(123) + 56(23) + 36(13) - 96(1) = \\ &= -96(1) + 12(12) + 36(13) + 56(23) - 21(123) - 32(132). \end{aligned}$$

2.2 Alguns homomorfismos e isomorfismos

Proposição 2.2.1 *Sejam V e W espaços vetoriais sobre um corpo K e $L(V,W) = \{\varphi : V \to W : \varphi$ é uma transformação linear$\}$. Então $L(V,W)$ é um espaço vetorial sobre K. Além disso, se $dim_K(V) = m$ e $dim_K(W) = n$ então $L(V,W)$ e $M_{n\times m}(K)$ são espaços vetoriais isomorfos.*

Proposição 2.2.2 *Sejam V um espaço vetorial sobre um corpo K e $L(V) = \{\varphi : V \to V : \varphi$ é um operador linear$\}$. Então $L(V)$ é uma álgebra associativa com identidade sobre K. Além disso, se $dim_K(V) = n$ então $dim_K(L(V)) = n^2$ e as álgebras $L(V)$ e $M_n(K)$ são isomorfas.*

Definição 2.2.1 *O grupo das matrizes inversíveis $n \times n$ sobre um corpo K é chamado* grupo linear geral *e denotado por $GL_n(K)$.*

Proposição 2.2.3 *Sejam V um espaço vetorial sobre um corpo K e $GL(V) = \{\varphi : V \to V : \varphi$ é um isomorfismo$\}$. Então $GL(V)$ é um grupo. Além disso, se $dim_K(V) = n$ então $GL(V)$ e $GL_n(K)$ são grupos isomorfos.*

Definição 2.2.2 *Se V é um espaço vetorial sobre um corpo K então $GL(V) = \{\varphi : V \to V : \varphi$ é um isomorfismo$\}$ é chamado* grupo linear geral *de V.*

Proposição 2.2.4 *Sejam G um grupo, K um corpo e $\varphi : G \to GL_n(K)$ um homomorfismo de grupos. Então $\psi : K(G) \to M_n(K)$ dada por*

$$\psi(\sum_{g\in G} \alpha_g g) = \sum_{g\in G} \alpha_g \varphi(g)$$

é um homomorfismo de álgebras.

Prova: Basta manipular algumas equações:

$$\begin{aligned}\psi(\sum_{g\in G}\alpha_g g+\sum_{g\in G}\beta_g g) &= \psi(\sum_{g\in G}(\alpha_g+\beta_g)g)=\sum_{g\in G}(\alpha_g+\beta_g)\varphi(g)= \\ &= \sum_{g\in G}\alpha_g\varphi(g)+\sum_{g\in G}\beta_g\varphi(g)=\psi(\sum_{g\in G}\alpha_g g)+\psi(\sum_{g\in G}\beta_g g).\end{aligned}$$

Portanto, $\psi(a+b)=\psi(a)+\psi(b)$ para quaisquer a e b em $K(G)$.

$$\psi(\lambda\sum_{g\in G}\alpha_g g)=\psi(\sum_{g\in G}(\lambda\alpha_g)g)=\sum_{g\in G}(\lambda\alpha_g)\varphi(g)=\lambda\sum_{g\in G}\alpha_g\varphi(g)=\lambda\psi(\sum_{g\in G}\alpha_g g).$$

Portanto, $\psi(\lambda a)=\lambda\psi(a)$ para quaisquer λ em K e a em $K(G)$.

$$\begin{aligned}\psi(\sum_{g\in G}\alpha_g g\cdot\sum_{g\in G}\beta_g g) &= \psi(\sum_{g\in G}(\sum_{hk=g}\alpha_h\beta_k)g)=\sum_{g\in G}(\sum_{hk=g}\alpha_h\beta_k)\varphi(g)= \\ &= \sum_{h\in G}(\sum_{k\in G}\alpha_h\beta_k\varphi(hk))=\sum_{h\in G}(\sum_{k\in G}\alpha_h\beta_k\varphi(h)\varphi(k))= \\ &= \sum_{h\in G}(\alpha_h\varphi(h)\sum_{k\in G}\beta_k\varphi(k))=(\sum_{h\in G}\alpha_h\varphi(h))\cdot(\sum_{k\in G}\beta_k\varphi(k))= \\ &= \psi(\sum_{g\in G}\alpha_g g)\cdot\psi(\sum_{g\in G}\beta_g g).\end{aligned}$$

Portanto, $\psi(a\cdot b)=\psi(a)\cdot\psi(b)$ para quaisquer a e b em $K(G)$. ■

Proposição 2.2.5 *Sejam G um grupo, K um corpo e $\psi: K(G)\to M_n(K)$ um homomorfismo de álgebras. Então $\psi(1g)\in GL_n(K)$, para qualquer g em G, e $\varphi: G\to GL_n(K)$ dada por*

$$\varphi(g)=\psi(1g)$$

é um homomorfismo de grupos.

Prova: Como $\psi(1g)\cdot\psi(1g^{-1})=\psi(1e)=I$, temos que $\psi(1g)\in GL_n(K)$, para qualquer g em G. Além disso, para quaisquer g e h em G,

$$\varphi(gh)=\psi(1gh)=\psi((1g)(1h))=\psi(1g)\psi(1h)=\varphi(g)\cdot\varphi(h).$$

Portanto, φ é um homomorfismo de grupos. ■

2.3 Representações de grupos

Definição 2.3.1 *Sejam G um grupo e V um espaço vetorial. Uma* representação *de G em V é um homomorfismo ρ do grupo G no grupo $GL(V)$. A dimensão de V é chamada* grau *de ρ.*

Definição 2.3.2 *Seja G um grupo. Uma* representação matricial *de G em $GL_n(K)$ é um homomorfismo ρ do grupo G no grupo $GL_n(K)$, para algum $n \in \mathbf{N}$ e K um corpo.*

Proposição 2.3.1 *Sejam G um grupo e V um espaço vetorial sobre um corpo K de dimensão n. Existe uma representação de G em V se, e somente se, existe uma representação matricial de G em $GL_n(K)$.*

Observação 2.3.1 *No caso de grau finito, não será feita distinção entre representação e representação matricial.*

Exemplo 2.3.1 *Se K é um corpo, uma representação ρ do grupo simétrico S_3 em $GL_3(K)$ de grau 3 é dada por*

$$\rho(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \rho(12) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \rho(13) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\rho(23) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \rho(123) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ e } \rho(132) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Esta representação é chamada de representação natural *de grau 3 de S_3; pois se $\sigma \in S_3$, então*

$$\rho(\sigma) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} \sigma(1) \\ \sigma(2) \\ \sigma(3) \end{pmatrix}.$$

Exemplo 2.3.2 *Se K é um corpo e existe um elemento w em K tal que $w^3 = 1$ e $w \neq 1$, então $\rho : S_3 \to GL_2(K)$, dada por*

$$\rho(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \rho(12) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \rho(13) = \begin{pmatrix} 0 & w \\ w^2 & 0 \end{pmatrix}$$

$$\rho(23) = \begin{pmatrix} 0 & w^2 \\ w & 0 \end{pmatrix}, \rho(123) = \begin{pmatrix} w & 0 \\ 0 & w^2 \end{pmatrix} \text{ e } \rho(132) = \begin{pmatrix} w^2 & 0 \\ 0 & w \end{pmatrix},$$

é uma representação de grau 2 de S_3 em $GL_2(K)$. Isto decorre dos seguintes fatos:
(i) $S_3 = \langle (12), (123) \rangle$, mais explicitamente,

$$1 = (123)^0(12)^0,\ (13) = (123)(12),\ (23) = (123)^2(12) \text{ e } (132) = (123)^2;$$

(ii) a ordem de $\begin{pmatrix} w & 0 \\ 0 & w^2 \end{pmatrix}$ no grupo $GL_n(K)$ é 3;

(iii) $\rho((123)^j(12)^i) = \begin{pmatrix} w & 0 \\ 0 & w^2 \end{pmatrix}^j \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^i$ *para $0 \leqslant i \leqslant 2$ e $0 \leqslant j \leqslant 1$.*

Exemplo 2.3.3 *Se $C_n = \langle g \rangle$ é o grupo cíclico de ordem n, então a aplicação $\rho : C_n \to \mathbf{C}^*$, dada por $\rho(g^k) = (e^{2i\Pi/n})^k$ para $0 \leqslant k \leqslant n-1$, é uma representação de C_n em $GL_1(\mathbf{C})$.*

2.4 Relação entre representações e módulos

Proposição 2.4.1 *Se G é um grupo e V é um G-módulo sobre um corpo K, então as seguintes propriedades são válidas para quaisquer v e w em V e g em G:*
(i) $g \cdot \mathbf{0} = \mathbf{0}$;
(ii) se $g \cdot v = g \cdot w$ *então* $v = w$;
(iii) se $g \cdot v = \mathbf{0}$ *então* $v = \mathbf{0}$.

Proposição 2.4.2 *Se A é uma álgebra com identidade sobre um corpo K e V é um A-módulo sobre K, então $a \cdot \mathbf{0} = \mathbf{0}$ para qualquer a em A.*

Lema 2.4.1 *Sejam G um grupo e V um espaço vetorial sobre um corpo K. Então:*
(i) Se existe uma operação $\cdot$ de $G \times V$ em V que determina uma estrutura de G-módulo em V, então a aplicação $\rho_g : V \to V$ dada por $\rho_g(v) = g \cdot v$ é um isomorfismo de espaços vetoriais para qualquer g em G e a aplicação $\rho : G \to GL(V)$ dada por $\rho(g) = \rho_g$ é uma representação de G em V;
(ii) Se existe uma representação ρ de G em V, a operação $\cdot$ de $G \times V$ em V dada por $g \cdot v = \rho_g(v)$, onde $\rho_g = \rho(g)$, determina uma estrutura de G-módulo em V.

Prova: Para provar (i), seja $g \in G$ e provemos que ρ_g é um isomorfismo. Para v e w em V e λ em K, como

$$\rho_g(v + w) = g \cdot (v + w) = g \cdot v + g \cdot w = \rho_g(v) + \rho_g(w)$$

e

$$\rho_g(\lambda v) = g \cdot (\lambda v) = \lambda(g \cdot v) = \lambda \rho_g(v),$$

segue que ρ_g é um operador linear. Se v e w são elementos de V tais que $\rho_g(v) = \rho_g(w)$, então $g \cdot v = g \cdot w$ e $v = w$; logo, ρ_g é injetora. Se $v \in V$ então $g^{-1} \cdot v \in V$ e $\rho_g(g^{-1} \cdot v) = v$; logo, ρ_g é sobrejetora. Portanto, ρ_g é um isomorfismo de espaços vetoriais.

Se g e h são elementos de G, então

$$\rho_{gh}(v) = (gh) \cdot v = g \cdot (h \cdot v) = g \cdot (\rho_h(v)) = \rho_g(\rho_h(v)) = (\rho_g \circ \rho_h)(v)$$

para qualquer v em V. Portanto, $\rho(gh) = \rho(g) \circ \rho(h)$ e ρ é um homomorfismo e uma representação de G em V.

Para provar (ii), sejam v e w em V, g e h em G e λ em K. Então:
1) $e \cdot v = \rho_e(v) = id(v) = v$, onde e é a identidade de G;
2) $g \cdot (v + w) = \rho_g(v + w) = \rho_g(v) + \rho_g(w) = g \cdot v + g \cdot w$;
3) $g \cdot (\lambda v) = \rho_g(\lambda v) = \lambda \rho_g(v) = \lambda(g \cdot v)$;
4) $(gh) \cdot v = \rho_{gh}(v) = (\rho_g \circ \rho_h)(v) = \rho_g(\rho_h(v)) = g \cdot (h \cdot v)$.

Portanto, a operação $\cdot$ determina uma estrutura de G-módulo em V. ■

Lema 2.4.2 *Sejam G um grupo e V um espaço vetorial sobre um corpo K. Então:*
(i) Se existe uma operação $\cdot$ de $K(G) \times V$ em V que determina uma estrutura de $K(G)$-módulo em V, então a aplicação $\rho_g : V \to V$ dada por $\rho_g(v) = (1g) \cdot v$, onde 1 é a identidade de K, é um isomorfismo de espaços vetoriais para qualquer g em G e a aplicação $\rho : G \to$

$GL(V)$ dada por $\rho(g) = \rho_g$ é uma representação de G em V;
(ii) Se existe uma representação ρ de G em V, a operação $\cdot$ de $K(G) \times V$ em V dada por $(\sum_{g\in G} \alpha_g g) \cdot v = \sum_{g\in G} \alpha_g \rho_g(v)$, onde $\rho_g = \rho(g)$, determina uma estrutura de $K(G)$-módulo em V.

Prova: Para provar (i), seja $g \in G$ e provemos que ρ_g é um isomorfismo. Para v e w em V e λ em K, como

$$\rho_g(v+w) = (1g)\cdot(v+w) = (1g)\cdot v + (1g)\cdot w = \rho_g(v) + \rho_g(w)$$

e

$$\rho_g(\lambda v) = (1g)\cdot(\lambda v) = \lambda((1g)\cdot v) = \lambda\rho_g(v),$$

segue que ρ_g é um operador linear. Se v e w são elementos de V tais que $\rho_g(v) = \rho_g(w)$, como $(1g^{-1})\cdot((1g)\cdot v) = v$ e $(1g^{-1})\cdot((1g)\cdot w) = w$, temos que $v = w$; logo, ρ_g é injetora. Se $v \in V$ então $(1g^{-1})\cdot v \in V$ e $\rho_g((1g^{-1})\cdot v) = v$; logo, ρ_g é sobrejetora. Portanto, ρ_g é um isomorfismo de espaços vetoriais.

Se g e h são elementos de G, então

$$\rho_{gh}(v) = (1gh)\cdot v = ((1g)(1h))\cdot v = (1g)\cdot((1h)\cdot v) = \rho_g(\rho_h(v)) = (\rho_g \circ \rho_h)(v)$$

para qualquer v em V. Portanto, $\rho(gh) = \rho(g)\circ\rho(h)$ e ρ é um homomorfismo e uma representação de G em V.

Para provar (ii), sejam v e w em V, $\sum_{g\in G}\alpha_g g$ e $\sum_{g\in G}\beta_g g$ em $K(G)$ e λ em K. Então:
1) $(1e)\cdot v = 1\rho_e(v) = 1id(v) = 1v = v$, onde e é a identidade de G e 1 é a identidade de K;
2)

$$\begin{aligned}(\sum_{g\in G}\alpha_g g)\cdot(v+w) &= \sum_{g\in G}\alpha_g\rho_g(v+w) = \sum_{g\in G}\alpha_g\rho_g(v) + \sum_{g\in G}\alpha_g\rho_g(w) = \\ &= (\sum_{g\in G}\alpha_g g)\cdot v + (\sum_{g\in G}\alpha_g g)\cdot w;\end{aligned}$$

3)

$$(\sum_{g\in G}\alpha_g g)\cdot(\lambda v) = \sum_{g\in G}\alpha_g\rho_g(\lambda v) = \lambda\sum_{g\in G}\alpha_g\rho_g(v) = \lambda((\sum_{g\in G}\alpha_g g)\cdot v)$$

e

$$(\sum_{g\in G}\alpha_g g)\cdot(\lambda v) = \sum_{g\in G}\alpha_g\rho_g(\lambda v) = \sum_{g\in G}\alpha_g\lambda\rho_g(v) = (\sum_{g\in G}\lambda\alpha_g g)\cdot v = (\lambda(\sum_{g\in G}\alpha_g g))\cdot v$$

4)

$$\begin{aligned}((\sum_{g\in G}\alpha_g g)(\sum_{g\in G}\beta_g g))\cdot v &= (\sum_{g\in G}(\sum_{ab=g}\alpha_a\beta_b)g)\cdot v = \sum_{g\in G}(\sum_{ab=g}\alpha_a\beta_b)\rho_g(v) = \\ &= \sum_{g\in G}(\sum_{h\in G}(\alpha_g\beta_h\rho_{gh}(v))) = \sum_{g\in G}(\alpha_g\sum_{h\in G}(\beta_h\rho_g(\rho_h(v)))) = \\ &= \sum_{g\in G}(\alpha_g\rho_g(\sum_{h\in G}(\beta_h\rho_h(v))) = (\sum_{g\in G}\alpha_g g)\cdot(\sum_{h\in G}\beta_h\rho_h(v)) = \\ &= (\sum_{g\in G}\alpha_g g)\cdot((\sum_{g\in G}\beta_g h)\cdot v).\end{aligned}$$

5)

$$((\sum_{g\in G}\alpha_g g)+(\sum_{g\in G}\beta_g g))\cdot v = (\sum_{g\in G}(\alpha_g+\beta_g)g)\cdot v = \sum_{g\in G}(\alpha_g+\beta_g)\rho_g(v) =$$
$$= \sum_{g\in G}\alpha_g\rho_g(v)+\sum_{g\in G}\beta_g\rho_g(v) = (\sum_{g\in G}\alpha_g g)\cdot v+(\sum_{g\in G}\beta_g g)\cdot v$$

Portanto, a operação $\cdot$ determina uma estrutura de $K(G)$-módulo em V. ■

Corolário 2.4.1 *Sejam G um grupo e V um espaço vetorial sobre um corpo K. Então:*
(i) Se existe uma operação $\cdot$ de $G\times V$ em V que determina uma estrutura de G-módulo em V, então a operação $\odot$ de $K(G)\times V$ em V dada por $(\sum_{g\in G}\alpha_g g)\odot v = \sum_{g\in G}\alpha_g(g\cdot v)$ determina uma estrutura de $K(G)$-módulo em V;
(ii) Se existe uma operação $\odot$ de $K(G)\times V$ em V que determina uma estrutura de $K(G)$-módulo em V, então a operação $\cdot$ de $G\times V$ em V dada por $g\cdot v = (1g)\odot v$ determina uma estrutura de G-módulo em V.

Proposição 2.4.3 *Se G é um grupo e V e W são espaços vetoriais sobre um corpo K, então as seguintes afirmações são equivalentes:*
(i) V e W são G-módulos isomorfos;
(ii) V e W são $K(G)$-módulos isomorfos.

2.5 Somas diretas em grupos abelianos

Definição 2.5.1 *Um subconjunto B de um grupo abeliano $(A,+)$ é* soma *dos subconjuntos da família $F=\{B_i\subseteq A : i\in I\}$ se $B=\{\sum_{i\in I}b_i : b_i\in B_i \text{ e } b_i\neq 0 \text{ para um número finito de índices } i\}$. Neste caso, escrevemos que $B=\sum_{i\in I}B_i$.*

Definição 2.5.2 *Um subconjunto B de um grupo abeliano $(A,+)$ é* soma direta *dos subconjuntos da família $F=\{B_i\subseteq A : i\in I\}$ se*
(i) $B=\sum_{i\in I}B_i$;
(ii) qualquer elemento de B pode ser expresso de maneira única sob a forma de um elemento de $\sum_{i\in I}B_i$.
Neste caso, escrevemos que $B=\oplus_{i\in I}B_i$.

Teorema 2.5.1 *Sejam $(A,+)$ um grupo abeliano, B um subgrupo de A e $F=\{B_i\leqslant A : i\in I\}$ tais que $B=\sum_{i\in I}B_i$. As seguintes afirmações são equivalentes:*
(i) $B=\oplus_{i\in I}B_i$;
(ii) se $\sum_{i\in I}b_i\in\sum_{i\in I}B_i$ e $\sum_{i\in I}b_i=0$, então cada $b_i=0$;
(iii) para cada $j\in I$, $B_j\cap\sum_{i\in I-\{j\}}B_i=\{0\}$.

Prova: É fácil ver que (i) implica (ii), pois caso contrário, o elemento neutro poderia ser expresso de duas maneiras distintas, o que seria um absurdo de acordo com (i).

Também é fácil ver que (ii) implica (iii), pois caso contrário, o elemento neutro poderia ser escrito como uma soma de b_i's os quais não são todos nulos. Assim, teríamos uma contradição com (ii).

Suponhamos que a afirmação (iii) seja verdadeira e que $\sum_{i\in I} a_i = \sum_{i\in I} b_i$ são elementos de $\sum_{i\in I} B_i$. Seja $k \in I$. Logo, $a_k - b_k = \sum_{i\in I-\{k\}} b_i - a_i$ e, conseqüentemente, $a_k - b_k = 0$, ou seja, $a_k = b_k$. Portanto, $a_i = b_i$ para qualquer $i \in I$. Assim, temos que $B = \oplus_{i\in I} B_i$. ■

Observação 2.5.1 *Muitas vezes consideramos somas diretas em estruturas algébricas que são grupos abelianos (tais como anéis, espaços vetorias, álgebras, módulos etc.).*

2.6 Lema de Zorn

Definição 2.6.1 *Uma* relação de ordem parcial *num conjunto P é uma relação simbolizada por $\leqslant$ e que para quaisquer x, y e z em P satisfaz as seguintes propriedades:*
(i) $x \leqslant x$ (reflexiva);
(ii) se $x \leqslant y$ *e* $y \leqslant z$ *então* $x \leqslant z$ (transitiva).
(iii) se $x \leqslant y$ *e* $y \leqslant x$ *então* $x = y$ (anti-simétrica).
Um conjunto P no qual está definida uma relação de ordem parcial é chamado um conjunto parcialmente ordenado.

Definição 2.6.2 *Dizemos que uma seqüência (a_i) num conjunto parcialmente ordenado P é uma* cadeia ascendente *se* $a_1 \leqslant a_2 \leqslant a_3 \leqslant ....$

Definição 2.6.3 *Um elemento x de um conjunto parcialmente ordenado $(P, \leqslant)$ é chamado* maximal *se sempre que* $x \leqslant y$ *então* $x = y$, *para qualquer y em P.*

Definição 2.6.4 *Seja A um subconjunto de um conjunto parcialmente ordenado P. Um elemento x em P é chamado um* limitante superior *de A se* $a \leqslant x$ *para todo* $a \in A$.

Seguem dois enunciados do chamado Lema de Zorn. É fácil ver que o segundo pode ser provado assumindo o primeiro. O Lema de Zorn é uma afirmação equivalente ao Axioma da Escolha, e portanto, não tem uma demonstração propriamente dita.

Lema de Zorn *Se P é um conjunto parcialmente ordenado no qual toda cadeia ascendente tem um limitante superior, então P possui um elemento maximal.*

Lema de Zorn *Seja A um conjunto e P uma propriedade com relação aos subconjuntos de A. Suponha que se P é verdadeira para uma cadeia $A_1 \subseteq A_2 \subseteq ...$ de subconjuntos de A, então P é verdadeira para a união $\cup_{i=1}^{\infty} A_i$. Então existe um subconjunto B de A para o qual P é verdadeira e que é maximal.*

2.7 Redutibilidade de módulos

Definição 2.7.1 *Seja A uma álgebra com identidade. Um A-módulo M é dito* irredutível *se os únicos submódulos de M são $\{0\}$ e M. Caso contrário, M é dito* redutível.

Definição 2.7.2 *Seja A uma álgebra com identidade. Um A-módulo M é dito* completamente redutível *ou* semi-simples *se M é uma soma direta de alguns dos seus submódulos irredutíveis.*

Observação 2.7.1 *A decomposição de um módulo completamente redutível em soma direta de alguns dos seus submódulos irredutíveis não é única. Consideremos, por exemplo,* $\mathbf{R}^2$ *como um espaço vetorial sobre* $\mathbf{R}$, $\mathbf{R}$ *uma álgebra com identidade sobre* $\mathbf{R}$ *e* $\mathbf{R}^2$ *um* $\mathbf{R}$*-módulo sobre* $\mathbf{R}$. *Claramente, se* S *for um subespaço de* $\mathbf{R}^2$ *com dimensão* 1 *então* S *é um submódulo irredutível de* $\mathbf{R}^2$. *Além disso, podemos escrever*

$$\mathbf{R}^2 = \{(x,0) : x \in \mathbf{R}\} \oplus \{(0,x) : x \in \mathbf{R}\}$$

e

$$\mathbf{R}^2 = \{(x,x) : x \in \mathbf{R}\} \oplus \{(x,-x) : x \in \mathbf{R}\}.$$

Lema 2.7.1 *Sejam* A *uma álgebra com identidade e* M *um* A*-módulo. As seguintes afirmações são equivalentes:*
(i) M *é uma soma de alguns dos seus submódulos irredutíveis;*
(ii) M *é completamente redutível;*
(iii) se N *é um submódulo de* M *então* $M = N \oplus N'$ *para algum submódulo* N' *de* M.

Prova: Seja $M = \sum_{i \in I} M_i$, onde cada M_i é um submódulo irredutível de M. Consideremos J como um subconjunto maximal de I para o qual a soma $M' = \sum_{j \in J} M_j$ é direta (a existência é garantida pelo Lema de Zorn). Se $i \in I - J$ então $M_i \cap M' = \{0\}$ ou $M_i \cap M' = M_i$. Portanto, $M' = M$. Assim, M é uma soma direta de alguns dos seus submódulos irredutíveis, ou seja, M é completamente irredutível.

Sejam $M = \oplus_{i \in I} M_i$, onde cada M_i é um submódulo irredutível de M, e N um submódulo de M. Consideremos J como um subconjunto maximal de I para o qual a soma $M' = N + (\oplus_{j \in J} M_j)$ é direta (a existência é garantida pelo Lema de Zorn). Pelo mesmo raciocínio do parágrafo anterior, temos que $M_i \subseteq N + (\oplus_{j \in J} M_j)$ para todo $i \in I$. Portanto, se considerarmos $N' = \oplus_{j \in J} M_j$ então $M = N \oplus N'$.

Se a afirmação (iii) é verdadeira, provemos que todo submódulo de M contém um submódulo irredutível.

Provemos que se P é um submódulo de M então para qualquer submódulo Q de P existe um submódulo Q' de P tal que $P = Q \oplus Q'$. Pela afirmação (iii), existe um submódulo R de M tal que $M = Q \oplus R$. Assim, tomando $Q' = R \cap P$, temos que Q' é um submódulo de P e que $P = Q \oplus Q'$.

Se P é um submódulo de M e $P \neq \{0\}$, escolhemos $v \in P$ tal que $v \neq 0$. Seja Q um submódulo maximal de P que não contém v (a existência é garantida pelo Lema de Zorn). Pelo que vimos no parágrafo anterior, existe um submódulo Q' de P tal que $P = Q \oplus Q'$. Além disso, Q' deve ser irredutível, pois caso contrário poderíamos escrever $Q' = Q'_1 \oplus Q'_2$ e, como $v \notin Q'_1$ ou $v \notin Q'_2$, teríamos que Q não é maximal (pois $Q \subseteq Q \oplus Q'_1$ e $Q \subseteq Q \oplus Q'_2$). Mas isto é uma contradição.

Seja M' a soma de todos os submódulos irredutíveis de M. Logo, $M = M' \oplus N$ para algum submódulo N de M. Se $N \neq \{0\}$ então N contém um submódulo irredutível, o que é uma contradição. Portanto, M é soma de alguns dos seus submódulos irredutíveis. ■

Definição 2.7.3 *Seja* G *um grupo. Um* G*-módulo* M *é dito* irredutível *se os únicos submódulos de* M *são* $\{0\}$ *e* M. *Caso contrário,* M *é dito* redutível.

Definição 2.7.4 *Seja G um grupo. Um G-módulo M é dito* completamente redutível *ou* semi-simples *se M é uma soma direta de alguns dos seus submódulos irredutíveis.*

Proposição 2.7.1 *Se G é um grupo e V é um espaço vetorial sobre um corpo K, então as seguintes afirmações são equivalentes:*
(i) V é redutível (irredutível ou semisimples) como $K(G)$-módulo;
(ii) V é redutível (irredutível ou semisimples) como G-módulo.

Definição 2.7.5 *Uma representação ρ de um grupo G num espaço vetorial V é* irredutível, redutível *ou* completamente redutível *(*semi-simples*) se V é irredutível, redutível ou semi-simples como $K(G)$-módulo (ou como G-módulo) respectivamente.*

2.8 Simplicidade de álgebras

Definição 2.8.1 *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* ideal simples *(ou* minimal*)* à esquerda *(*à direita *ou* bilateral*) de A. Escrevemos $I \lhd_e^s A$ ($I \lhd_d^s A$ ou $I \lhd^s A$).*

Definição 2.8.2 *Uma álgebra associativa com identidade que é uma soma direta de alguns dos seus ideais simples à esquerda é chamada* álgebra semi-simples.

Proposição 2.8.1 *Se A é uma álgebra associativa com identidade então A é um A-módulo.*

Proposição 2.8.2 *Se A é uma álgebra associativa com identidade e I é um subconjunto de A então:*
(i) I é um ideal à esquerda de A se, e somente se, I é um submódulo do A-módulo A.
(ii) I é um ideal simples à esquerda de A se, e somente se, I é um submódulo irredutível do A-módulo A.

Proposição 2.8.3 *Se A é uma álgebra associativa com identidade então as seguintes afirmações são equivalentes:*
(i) A é uma álgebra semi-simples;
(ii) A é um A-módulo completamente redutível.

Definição 2.8.3 *Sejam A uma álgebra e I e J ideais à esquerda de A. Definimos*

$$IJ = \{\sum_{i=1}^{n} x_i y_i : x_i \in I,\ y_i \in J \text{ e } n \in \mathbf{N}\}.$$

Definição 2.8.4 *Sejam A uma álgebra associativa com identidade, I um ideal à esquerda de A, M um A-módulo e m um elemento de M. Definimos*
(i) $Im = \{x \cdot m : x \in I\}$;
(ii) $IM = \{\sum_{i=1}^{n} x_i \cdot m_i : x_i \in I,\ m_i \in M \text{ e } n \in \mathbf{N}\}$;
(iii) $AM = \{\sum_{i=1}^{n} a_i \cdot m_i : a_i \in A,\ m_i \in M \text{ e } n \in \mathbf{N}\}$.

Proposição 2.8.4 *Se A é uma álgebra associativa com identidade, I um ideal à esquerda de A, M um A-módulo e m um elemento de M, então Im, IM e AM são submódulos de M.*

Lema 2.8.1 *Sejam A uma álgebra semi-simples e I um ideal simples à esquerda de A. Se M é um A-módulo irredutível então $IM = \{0\}$ ou $I \cong M$ (como A-módulos).*

Prova: Como IM é um submódulo de M, temos que $IM = \{0\}$ ou $IM = M$. Se $IM = M$, então $Im \neq \{0\}$ para algum $m \in M$. Logo, $Im = M$. Seja $\varphi : I \to M$ tal que $\varphi(x) = x \cdot m$. Provemos que φ é um isomorfismo de módulos. Se x e y pertencem a I, λ pertence a K (onde K é o corpo sobre o qual A está definida) e a pertence a A, então
a) $\varphi(x+y) = (x+y) \cdot m = x \cdot m + y \cdot m = \varphi(x) + \varphi(y)$;
b) $\varphi(\lambda x) = (\lambda x) \cdot m = \lambda(x \cdot m) = \lambda\varphi(x)$;
c) $\varphi(ax) = (ax) \cdot m = a \cdot (x \cdot m) = a \cdot \varphi(x)$;
d) como I é simples temos que $Ker\varphi = \{0\}$ e, portanto, φ é injetora;
e) claramente φ é sobrejetora.
Portanto, $I \cong M$ (como A-módulos). ■

Definição 2.8.5 *Uma álgebra semi-simples A em que todos os ideais simples à esquerda são isomorfos (como A-módulos) é chamada* álgebra simples.

Teorema 2.8.1 *Sejam A uma álgebra semi-simples e $\{I_i\}$ uma classe de ideais simples à esquerda de A tal que*
a) se $i \neq j$ então $I_i \ncong I_j$ (como A-módulos);
b) se $I \lhd_e^s A$ então $I \cong I_i$ (como A-módulos) para algum i.
Se para cada i, A_i é a soma dos ideais simples à esquerda de A isomorfos a I_i, então
(i) $A_iA_j = \{ab : a \in A_i \text{ e } b \in A_j\} = \{0\}$, se $i \neq j$;
(ii) $\{I_i\}$ é uma classe finita, ou seja, podemos escrever $\{I_i\} = \{I_1, ..., I_n\}$;
(iii) A_i é uma álgebra associativa com identidade;
(iv) A_i é um ideal bilateral de A;
(vi) $A = \oplus_{i=1}^n A_i$.
(v) A_i é uma álgebra simples;

Prova: Pelo **Lema 2.8.1**, podemos concluir que $A_iA_j = \{0\}$, se $i \neq j$.

Seja $\mathbf{1} = \sum_i e_i$, onde $e_i \in A_i$ e e_i não é nulo somente para um número finito de índices i. Podemos escrever $\mathbf{1} = \sum_{i=1}^n e_i$, onde cada e_i não é nulo se $1 \leqslant i \leqslant n$. Se $k \notin \{1, ..., n\}$, então para cada $a \in A_k$ obtemos $a = \mathbf{1}a = (e_1 + ... + e_n)a = 0$, ou seja, $A_k = \{0\}$. Portanto, $\{I_i\}$ é uma classe finita e podemos escrever $\{I_i\} = \{I_1, ..., I_n\}$.

Através de uma verificação rotineira de propriedades, podemos concluir que A_i é uma álgebra associativa. Além disso, $a = \mathbf{1}a = (e_1 + ... + e_n)a = e_ia$ e $a = a\mathbf{1} = a(e_1 + ... + e_n) = ae_i$. Portanto A_i é uma álgebra com identidade e e_i é a identidade de A_i.

Claramente, $AA_i = A_i$. Por outro lado, temos que $A_iA = A_i(\sum_{j=1}^n A_j) = A_iA_i = A_i$. Portanto, A_i é um ideal bilateral de A.

Se $a_1 + ... + a_n = 0$, cada $a_i \in A_i$, multiplicando por e_i temos que $a_i = 0$. Portanto, $A = \oplus_{i=1}^n A_i$.

Claramente A_i é uma álgebra semi-simples. Vamos provar que todos os ideais simples à esquerda de A_i são isomorfos (como A-módulos). Para isto basta provarmos que todos os ideais simples à esquerda de A_i são isomorfos (como A-módulos) a I_i. Suponhamos, por absurdo, que I seja um ideal simples à esquerda de A_i que não é isomorfo (como A-módulo) a I_i. Novamente pelo **Lema 2.8.1**, temos que $IA_i = \{0\}$, já que A_i é soma direta de ideais simples à esquerda de A_i isomorfos (como A-módulos) a I_i. Mas isto é um absurdo, pois $IA_i = I$. ■

Definição 2.8.6 *Consideremos a notação do Teorema anterior. As álgebras simples A_i são chamadas* componentes simples *da álgebra semi-simples A. Se I é um ideal simples à esquerda de A e isomorfo (como A-módulo) a I_i, então dizemos que I* está na componente simples A_i.

Teorema 2.8.2 (Maschke) *Se G é um grupo finito e K é um corpo tais que a característica de K não divide a ordem de G, então a álgebra $K(G)$ é semi-simples.*

Prova: Consideremos $K(G)$ como um $K(G)$-módulo, sejam M um submódulo de $K(G)$ e M' um subespaço de $K(G)$ tais que $K(G) = M \oplus M'$.

Seja $\pi : K(G) \to M$ a projeção, isto é, $\pi(u+v) = u$ para quaisquer $u \in M$ e $v \in M'$.

Seja $\pi' : K(G) \to K(G)$ a transformação linear dada por

$$\pi'(x) = \frac{1}{|G|} \sum_{g \in G} (g^{-1} \pi g)(x).$$

Provemos que π' é uma projeção sobre M. Para isto basta observar dois fatos. Primeiro, $\pi'(x) = x$ para todo $x \in M$. Segundo, como $\pi(gx) \in M$ para quaisquer $g \in G$ e $x \in K(G)$, então $\pi'(K(G)) \subseteq M$.

Se provarmos que π' é um homomorfismo de módulos, então, como $Ker\pi' = M'$, temos que M' é um submódulo de $K(G)$. Para isto basta provar que $\pi'(gx) = g\pi'(x)$ para quaisquer $g \in G$ e $x \in K(G)$. Mas isto decorre do seguinte cálculo

$$\begin{aligned} \pi'(gx) &= \frac{1}{|G|} \sum_{h \in G} (h^{-1} \pi h)(gx) = \frac{1}{|G|} g \sum_{h \in G} (g^{-1} h^{-1} \pi hg)(x) = \\ &= g\left(\frac{1}{|G|} \sum_{h \in G} ((hg)^{-1} \pi (hg))(x) = g\pi'(x).\right. \end{aligned}$$

Como M' é um submódulo de $K(G)$, pelo **Lema 2.7.1**, temos que $K(G)$ é um $K(G)$-módulo completamente redutível, ou seja, $K(G)$ é uma álgebra semi-simples. ■

Capítulo 3

Álgebras com identidade polinomial

Definimos álgebra livre, identidades polinomiais e PI álgebras. Desenvolvemos na última seção um resultado importante relacionando polinômios multilineares e identidades polinomiais.

3.1 Alfabeto, letra e palavra

Definição 3.1.1 *Um* alfabeto Σ *é um conjunto não-vazio, cujos elementos são chamados* letras *ou* símbolos. *Uma* palavra *sobre* Σ *é uma seqüência finita de símbolos de* Σ. *O conjunto de todas as palavras sobre* Σ *é denotado por* Σ^*, *ou seja,*

$$\Sigma^* = \{a_1...a_n : a_i \in \Sigma \text{ para } i = 1, ..., n \text{ e } n \text{ um inteiro não-negativo}\}.$$

O número de letras de uma palavra w *é denotado por* $l(w)$ *e é chamado o* comprimento *de* w, *ou seja, se* $w = a_1...a_n$ *é uma palavra então o inteiro não-negativo* n *é o comprimento de* w. *A palavra de comprimento zero é denotada por* λ *e é chamada a* palavra vazia.

Exemplo 3.1.1 *Se considerarmos o alfabeto* $\Sigma = \{a, b, c\}$, *então* $w_1 = abba$, $w_2 = bcc$ *e* $w_3 = bcaca$ *são algumas palavras sobre* Σ *e seus comprimentos são* $l(w_1) = 4$, $l(w_2) = 3$ *e* $l(w_3) = 5$.

Definição 3.1.2 *Seja* Σ *um alfabeto. Dadas duas palavras* $v = a_1...a_m$ *e* $w = b_1...b_n$ *sobre* Σ *definimos a operação de* concatenação *de* $\Sigma^* \times \Sigma^*$ *em* Σ^* *por* $vw = a_1...a_m b_1...b_n$.

Observação 3.1.1 *Observemos que* $l(vw) = l(v) + l(w)$.

Proposição 3.1.1 *Se* Σ *é um alfabeto então* Σ^* *é um monóide com relação à operação de concatenação, cuja identidade é a palavra vazia* λ.

Definição 3.1.3 *A palavra* v *é uma* subpalavra *da palavra* w *se existem palavras* w_1 *e* w_2 *tais que* $w = w_1 v w_2$. *Dizemos que a palavra* w *contém a letra* a *ou que* a *é uma das letras que compõem* w *se* a *é uma subpalavra de* w.

Exemplo 3.1.2 *Seja* $\Sigma = \{a, b, c, d, e\}$ *um alfabeto. A palabra* bd *é uma subpalavra de* $cbdae$. *A palavra vazia* λ *é uma subpalavra de qualquer palavra. A letra* a *ocorre em* $abba$, *mas não ocorre em* bcc.

3.2 Polinômios

Definição 3.2.1 *Definiremos o alfabeto* $\mathbf{X}$ *como o conjunto enumerável* $\{X_1, X_2, ...\}$ *onde cada* X_i *é chamado* indeterminada.

Como $\mathbf{X}^*$ é um monóide, sabemos, pelo **Teorema 2.1.1**, que $K(\mathbf{X}^*)$, onde K é um corpo, é uma álgebra associativa com identidade sobre K. Assim, faz sentido a próxima definição.

Definição 3.2.2 *Seja* K *um corpo. A álgebra* $K\{\mathbf{X}\} = K(\mathbf{X}^*)$ *é chamada* álgebra livre *e os seus elementos são chamados* polinômios. *Um polinômio da forma* αw*, onde* $\alpha \in K$ *e* $w \in \mathbf{X}^*$*, é chamado um* monômio *e o elemento* α *é chamado coeficiente de* αw*. O subconjunto de* $K\{\mathbf{X}\}$ *formado por todos os monômios é denotado por* $\mu(\mathbf{X})$.

A partir de agora, ainda que não seja mencionado, K *denotará sempre o corpo sobre o qual está definida a álgebra dos polinômios.*

Observação 3.2.1 *Se* K *é um corpo então todo polinômio de* $K\{\mathbf{X}\}$ *pode ser escrito como uma soma finita de monômios de* $K\{\mathbf{X}\}$.

Definição 3.2.3 *Dizemos que* X_i ocorre *num polinômio* p *se para algum dos monômios* αw *de* p*, com* $\alpha \neq 0$*, a palavra* w *contém* X_i*. Escrevemos* $p(X_{i_1}, ..., X_{i_k})$ *para denotar que* $X_{i_1}, ..., X_{i_k}$ *são as únicas indeterminadas que ocorrem em* p.

Definição 3.2.4 *Sejam* A *uma álgebra associativa sobre um corpo* K*,* $p(X_{i_1}, ..., X_{i_k}) \in K\{\mathbf{X}\}$ *e* $a_1, ..., a_k$ *elementos de* A*. Denotamos por* $p(a_1, ..., a_k)$ *o elemento de* A *obtido pela substituição de* X_{i_j} *por* a_j *em* $p(X_{i_1}, ..., X_{i_k})$*, para todo* j*, e realizadas as respectivas operações em* A.

Teorema 3.2.1 *Sejam* A *uma álgebra associativa com identidade sobre um corpo* K *e* $\{a_1, a_2, ...\}$ *um subconjunto enumerável de* A*. Então*
(i) existe um único homomorfismo de álgebras $\psi : K\{X\} \to A$ *tal que* $\psi(X_i) = a_i$ *para todo* $i \in \mathbf{N}$*;*
(ii) este homomorfismo ψ *é dado por* $\psi(p(X_{i_1}, ..., X_{i_k})) = p(a_{i_1}, ..., a_{i_k})$.

Prova: Seja $\sigma : \mathbf{X} \to A$ uma aplicação tal que $\sigma(X_i) = a_i$ para todo $i \in \mathbf{N}$. Podemos extender σ unicamente a um homomorfismo de monóides $\sigma' : \mathbf{X}^* \to A$, pondo $\sigma'(\lambda) = \mathbf{1}$ e $\sigma'(X_{i_1}...X_{i_k}) = \sigma(X_{i_1})...\sigma(X_{i_k})$. Podemos agora extender a um homomorfismo de álgebras $\sigma'' : K\{\mathbf{X}\} \to A$ dado por $\sigma''(\sum_{w \in \mathbf{X}^*} \alpha_w w) = \sum_{w \in \mathbf{X}^*} \alpha_w \sigma'(w)$. Observemos que $\sigma'' = \psi$ e portanto é um ψ homomorfismo de álgebras.

Por outro lado, seja $\varphi : K\{\mathbf{X}\} \to A$ um homomorfismo de álgebras tal que $\varphi(X_i) = a_i$ para todo $i \in \mathbf{N}$. Logo,

$$\varphi(X_{i_1}, ..., X_{i_k}) = \varphi(X_{i_1})...\varphi(X_{i_k}) = \sigma(X_{i_1})...\sigma(X_{i_k}) = \sigma'(X_{i_1}, ..., X_{i_k}),$$

e então

$$\varphi(\sum_{w \in \mathbf{X}^*} \alpha_w w) = \sum_{w \in \mathbf{X}^*} \alpha_w \varphi(w) = \sum_{w \in \mathbf{X}^*} \alpha_w \sigma'(w) = \psi(\sum_{w \in \mathbf{X}^*} \alpha_w w),$$

provando que $\varphi = \psi$. ■

Definição 3.2.5 *O polinômio* $s_n(X_1, ..., X_n) = \sum_{\sigma \in S_n} (-1)^\sigma X_{\sigma(1)} ... X_{\sigma(n)}$ *é chamado* polinômio "standard" de grau n.

Definição 3.2.6 *O polinômio* $S_k(X_1, ..., X_n) = \sum_{1 \leqslant i_1 < ... < i_k \leqslant n} X_{i_1} ... X_{i_k}$, *onde* $k \leqslant n$, *é chamado* polinômio simétrico de grau k em n variáveis.

Definição 3.2.7 *O polinômio* $[X_i, X_j] = X_i X_j - X_j X_i$, *com* i *e* j *em* $\mathbf{N}$, *é chamado* polinômio comutador *de* X_i *e* X_j.

Observação 3.2.2 *O polinômio standard de grau 2 é simplesmente o polinômio comutador de* X_1 *e* X_2, *isto é,* $s_2(X_1, X_2) = [X_1, X_2]$.

Definição 3.2.8 *O polinômio* $S_k(X_1, ..., X_n) = \sum_{1 \leqslant i_1 < ... < i_k \leqslant n} X_{i_1} ... X_{i_k}$, *onde* $k \leqslant n$, *é chamado* polinômio simétrico de grau k em n variáveis.

Definição 3.2.9 *O polinômio* $P_k(X_1, ..., X_n) = (X_1)^k + ... + (X_n)^k$ *é chamado* soma de potências de grau k em n variáveis.

3.3 Álgebras com identidade polinomial

Definição 3.3.1 *Sejam* A *uma álgebra associativa sobre um corpo* K *e* $p(X_{i_1}, ..., X_{i_n})$ *um polinômio não-nulo de* $K\{\mathbf{X}\}$. *Se* $p(a_1, ..., a_n) = 0$ *para quaisquer* $a_1, ..., a_n$ *em* A, *então* p *é uma* identidade polinomial *de* A, A *é uma* álgebra com identidade polinomial *(ou abreviadamente uma* PI álgebra*)* e A satisfaz a identidade polinomial p.

Exemplo 3.3.1 *Se* A *é uma álgebra associativa e comutativa, então*

$$s_2(X_1, X_2) = [X_1, X_2] = X_1 X_2 - X_2 X_1$$

é uma identidade polinomial de A. *Portanto, toda álgebra associativa e comutativa é uma PI álgebra.*

Observação 3.3.1 *A classe das álgebras associativas que satisfazem um polinômio standard é na verdade uma generalização da classe das álgebras associativas e comutativas.*

Exemplo 3.3.2 *A álgebra das matrizes de dimensão* 2×2 *sobre um corpo* K, *denotada por* $M_2(K)$, *satisfaz a identidade de Hall*

$$p(X_1, X_2, X_3) = [[X_1, X_2]^2, X_3] = (X_1 X_2 - X_2 X_1)^2 X_3 - X_3 (X_1 X_2 - X_2 X_1)^2.$$

Verifica-se facilmente esta identidade observando dois fatos cuja demonstração é bastante simples. O primeiro é que se A *e* B *pertencem a* $M_n(K)$, $n \in \mathbf{N}$, *então o traço de* $[A, B]$ *é zero; e o segundo é que 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 identidade polinomial é conhecida há bastante tempo, e aparece num artigo de M. Hall [8] de 1943.*

Exemplo 3.3.3 *A álgebra das matrizes triangulares superiores $n \times n$ sobre um corpo K, denotada por $T_n(K)$, é uma PI álgebra pois satisfaz a identidade*

$$p(X_1, ..., X_{2n}) = [X_1, X_2][X_3, X_4]...[X_{2n-1}, X_{2n}].$$

Isto decorre dos fatos de que se A e B são elementos de $T_n(K)$ então $[A, B]$ é um elemento de $T_n(K)$ com a diagonal principal nula; e que o produto de n matrizes triangulares superiores $n \times n$ com a diagonal principal nula é a matriz nula.

Definição 3.3.2 *Uma álgebra associativa A é chamada* nil álgebra *se para cada $a \in A$ existe um número natural n tal que $a^n = 0$. O menor número n com esta propriedade é chamado* índice de nilpotência *do elemento a. Uma álgebra A é chamada* nil álgebra de índice limitado *n se existe um número natural fixo n tal que $a^n = 0$, para cada $a \in A$.*

Exemplo 3.3.4 *Toda nil álgebra associativa de índice limitado n é uma PI álgebra, pois satisfaz a identidade*

$$p(X_1) = X_1^n.$$

Definição 3.3.3 *Uma álgebra associativa A é chamada* nilpotente *se existe um número natural fixo n tal que o produto de quaisquer n elementos de A é igual a zero. O menor número n com esta propriedade é chamado* índice de nilpotência *da álgebra A.*

Exemplo 3.3.5 *Toda álgebra nilpotente é uma PI álgebra, pois satisfaz a identidade*

$$p(X_1, ..., X_n) = X_1...X_n,$$

onde n é o índice de nilpotência da álgebra.

Após vários exemplos de PI álgebras, uma pergunta natural seria: *existem álgebras associativas que não satisfazem nenhuma identidade polinomial?* A resposta para esta pergunta é sim. A álgebra livre $K\{\mathbf{X}\}$, por exemplo, não satisfaz nenhuma identidade polinomial. Isto pode ser compreendido através de um argumento muito simples. Suponhamos, por absurdo, que $p(X_{i_1}, ..., X_{i_n})$ é uma identidade polinomial de $K\{\mathbf{X}\}$. Logo, $p(X_{i_1}, ..., X_{i_n}) = p(p_1(X_{i_1}), ..., p_n(X_{i_n})) = 0$, onde $p_k(X_{i_k}) = X_{i_k}$ para $1 \leqslant k \leqslant n$, o que é um absurdo pois $p(X_{i_1}, ..., X_{i_n}) \neq 0$.

Definição 3.3.4 *Se A é uma PI álgebra sobre um corpo K, então definimos*

$$T(A) = \{p \in K\{\mathbf{X}\} : p \text{ é uma identidade polinomial de } A\}.$$

Proposição 3.3.1 *Se A é uma PI álgebra sobre um corpo K, então $T(A) \lhd K\{\mathbf{X}\}$.*

O próximo lema mostra que toda álgebra de dimensão finita é uma PI álgebra.

Lema 3.3.1 *Se A é uma álgebra associativa de dimensão finita n, então A satisfaz a identidade*

$$s_{n+1}(X_1, ..., X_{n+1}) = \sum_{\sigma \in S_{n+1}} (-1)^\sigma X_{\sigma(1)}...X_{\sigma(n+1)}.$$

Prova: Da definição de polinômio standard é óbvio que ele é igual a zero se dois de seus argumentos são iguais. Sejam $\{e_1, ..., e_n\}$ uma base do espaço vetorial A e $a_1, ..., a_{n+1}$ elementos arbitrários da álgebra A.Podemos representar cada um dos elementos a_i na forma de uma combinação linear dos elementos $e_1, ..., e_n$ com coeficientes em K. É fácil ver que a expressão $s_{n+1}(a_1, ..., a_{n+1})$ é uma combinação linear de termos da forma $s_{n+1}(e_{i_1}, ..., e_{i_{n+1}})$, onde cada e_{i_k} é um elemento do conjunto $\{e_1, ..., e_n\}$. Mas então existem dois argumentos na expressão $s_{n+1}(e_{i_1}, ..., e_{i_{n+1}})$ que necessariamente coincidem, e portanto, ela deve ser igual a zero. Logo, $s_{n+1}(a_1, ..., a_{n+1}) = 0$. ■

3.4 Polinômios multilineares

Definição 3.4.1 *O* grau do monômio αw, *onde* $\alpha \in K$, $\alpha \neq 0$, *e* $w \in \mathbf{X}^*$, *denotado por* $gr(\alpha w)$, *é o comprimento da palavra* w. *O* grau do polinômio $p \in K\{\mathbf{X}\}$ *é definido como o grau máximo de seus monômios.*

Definição 3.4.2 *Um monômio* αw, *onde* $\alpha \in K$ *e* $w \in \mathbf{X}^*$, *tem* tipo $[n_1, n_2, ..., n_k]$ *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. *Monômios que têm o mesmo tipo são chamados* monômios do mesmo-tipo.

Exemplo 3.4.1 *O monômio* $X_1X_2X_2X_1X_4$ *tem tipo* $[2, 2, 0, 1]$ *e* $gr(X_1X_2X_2X_1X_4) = 5$. *O monômio* $X_1X_1X_2X_4X_2$ *tem tipo* $[2, 2, 0, 1]$. *Portanto,* $X_1X_2X_2X_1X_4$ *e* $X_1X_1X_2X_4X_2$ *são do mesmo-tipo.*

Observação 3.4.1 *Se* m *é um monômio de tipo* $[n_1, n_2, ..., n_k]$ *então* $gr(m) = \sum_{i=1}^{k} n_i$.

Definição 3.4.3 *Um polinômio é chamado* polinômio homogêneo em X_i com grau n_i *se todos os seus monômios têm grau* n_i *em* X_i. *Um polinômio é chamado* polinômio homogêneo *se todos os seus monômios são do mesmo-tipo.*

Exemplo 3.4.2 *O polinômio* $p(X_1, X_2, X_4) = 3X_1X_1X_2X_4 + 7X_1X_4X_2X_1 - 5X_2X_1X_1X_4 + 1X_1X_1X_4X_2$ *é homogêneo, pois cada um de seus monômios tem tipo* $[2, 1, 0, 1]$. *O polinômio* $p(X_1, X_2, X_3) = 4X_1X_1X_2 - 9X_1X_3X_1 + 2X_1X_1$ *é homogêneo em* X_1 *com grau 2; no entanto, não é homogêneo.*

Definição 3.4.4 *Seja* p *um polinômio arbitrário. Se agruparmos os monômios do mesmo tipo, o polinômio* p *pode ser representado como uma soma de polinômios homogêneos. Estes polinômios homogêneos são chamados as* componentes homogêneas *do polinômio* p.

Exemplo 3.4.3 *Consideremos o polinômio*

$$p(X_1, X_2, X_3) = 4X_1X_1X_2 - 9X_1X_2X_1 + 2X_2X_1X_1 + 7X_3X_1 + 5X_1X_3.$$

Suas componentes homogêneas são $4X_1X_1X_2 - 9X_1X_2X_1 + 2X_2X_1X_1$ *e* $7X_3X_1 + 5X_1X_3$.

Definição 3.4.5 *Um polinômio homogêneo em X_i com grau 1 é chamado* polinômio linear em X_i. *Um polinômio homogêneo de tipo $[n_1, ..., n_k]$, onde $n_i = 0$ ou $n_i = 1$ para cada i, é chamado* polinômio multilinear.

Exemplo 3.4.4 *O polinômio $p(X_1, X_2, X_4) = 3X_1X_2X_4 + 7X_1X_4X_2 - 5X_2X_1X_4$ é multilinear, pois cada um de seus monômios tem tipo $[1, 1, 0, 1]$. O polinômio $p(X_1, X_2, X_3) = 4X_1X_2 - 9X_1X_3X_2 + 2X_1X_3$ é linear em X_1; no entanto, não é multilinear.*

Definição 3.4.6 *Se $A = \{i_1, ..., i_k\}$ então S_A denota o grupo simétrico formado pelas permutações dos elementos de A.*

Proposição 3.4.1 *Seja K um corpo. Se p é um polinômio multilinear de $K\{\mathbf{X}\}$ e $X_{i_1}, ..., X_{i_k}$ são as únicas indeterminadas que ocorrem em p, então*

$$p(X_{i_1}, ..., X_{i_k}) = \sum_{\sigma \in S_A} \alpha_\sigma X_{\sigma(i_1)} ... X_{\sigma(i_k)},$$

onde $A = \{i_1, ..., i_k\}$ e cada $\alpha_\sigma \in K$.

3.5 Identidades multilineares

Conforme vimos no **Exemplo 3.3.4**, toda nil álgebra associativa A de índice limitado n é uma PI álgebra, pois satisfaz a identidade $p(X_1) = X_1^n$. No entanto, o polinômio p não é um polinômio multilinear. Na tentativa de encontrar uma identidade polinomial multilinear para a nil álgebra A, vamos analisar algumas nil álgebras de índices limitados.

Se A é uma nil álgebra de índice limitado 2, afirmamos que

$$p(X_1, X_2) = X_1X_2 + X_2X_1 = \sum_{\sigma \in S_2} X_{\sigma(1)}X_{\sigma(2)}$$

é uma identidade polinomial de A; pois $0 = (a+b)^2 = a^2 + ab + ba + b^2 = ab + ba$ para quaisquer elementos a e b de A.

Se A é uma nil álgebra de índice limitado 3, afirmamos que

$$p(X_1, X_2, X_3) = \sum_{\sigma \in S_3} X_{\sigma(1)}X_{\sigma(2)}X_{\sigma(3)}$$

é uma identidade polinomial de A. Para obter este resultado basta observar dois fatos:

$$0 = (x+y)^3 = x^3 + x^2y + xyx + xy^2 + yx^2 + yxy + y^2x + y^3 = x^2y + xyx + xy^2 + yx^2 + yxy + y^2x$$

para quaisquer elementos x e y de A e

$$\begin{aligned} 0 &= (a+b+c)^3 = ((a+b)+c)^3 \\ &= (a+b)^2c + (a+b)c(a+b) + (a+b)c^2 + c(a+b)^2 + c(a+b)c + c^2(a+b) \\ &= (a^2c + aca + ac^2 + ca^2 + cac + c^2a) + (b^2c + bcb + bc^2 + cb^2 + cbc + c^2b) + \\ &\quad + (abc + acb + bac + bca + cab + cba) \\ &= abc + acb + bac + bca + cab + cba \end{aligned}$$

para quaisquer elementos a, b e c de A.

Estes casos particulares podem servir como motivação para querer encontrar uma identidade polinomial multinilinear para uma nil álgebra de índice limitado n, ou de forma mais genérica, dada uma PI álgebra A qualquer querer encontrar uma identidade polinomial multinilinear para A. Esta seção desenvolve algumas ferramentas necessárias e conclui este resultado no **Corolário 3.5.1.**

Inicialmente introduziremos a seguinte notação: o sinal $\widehat{\ }$ sobrescrito a X_i numa soma $X_1 + ... + \widehat{X}_i + ... + X_n$ denota que X_i aparece na soma com coeficiente zero, ou seja,

$$X_1 + ... + \widehat{X}_i + ... + X_n = (\sum_{k=1}^{n} X_k) - X_i.$$

Por exemplo, $X_1 + \widehat{X}_2 + X_3 + \widehat{X}_4 + X_5 = X_1 + X_3 + X_5$.

Definição 3.5.1 *Sejam $p = p(X_1, ..., X_n)$ um polinômio de $K\{\mathbf{X}\}$ e $Y_1, ..., Y_k \in \mathbf{X} - \{X_1, ..., X_n\}$. Para cada $i = 1, ..., n$ definimos um polinômio $p\mathbf{L}_i^k$ pela fórmula*

$$\begin{aligned} & p\mathbf{L}_i^k(X_1, ..., X_{i-1}, Y_1, ..., Y_k, X_{i+1}, ..., X_n) = \\ = \; & p(X_1, ..., X_{i-1}, Y_1 + ... + Y_k, X_{i+1}, ..., X_n) - \\ & - \sum_{q=1}^{k} p(X_1, ..., X_{i-1}, Y_1 + ... + \widehat{Y}_q + ... + Y_k, X_{i+1}, ..., X_n) + \\ & + \sum_{1 \leqslant q_1 < q_2 \leqslant k}^{k} p(X_1, ..., X_{i-1}, Y_1 + ... + \widehat{Y}_{q_1} + ... + \widehat{Y}_{q_2} + ... + Y_k, X_{i+1}, ..., X_n) + \\ & + ... + (-1)^{k-1} \sum_{q=1}^{k} p(X_1, ..., X_{i-1}, Y_q, X_{i+1}, ..., X_n) \end{aligned}$$

Proposição 3.5.1 *Sejam A uma álgebra associativa sobre um corpo K, P e Q subgrupos do grupo aditivo da álgebra A. Se para qualquer polinômio $p = p(X_1, ..., X_n) \in K\{\mathbf{X}\}$ e quaisquer elementos $a_1, ..., a_n$ em P temos que $p(a_1, ..., a_n) \in Q$, então para quaisquer $a_1, ..., a_{i-1}, b_1, ..., b_k, a_{i+1}, ..., a_n$ em P teremos que $p\mathbf{L}_i^k(a_1, ..., a_{i-1}, b_1, ..., b_k, a_{i+1}, ..., a_n) \in Q$. Em particular, se $p \in T(A)$, então $p\mathbf{L}_i^k \in T(A)$.*

Prova: Verificação direta. ■

Assim o operador $\mathbf{L}_i^k$ transforma uma identidade da álgebra A numa nova identidade da álgebra A. A fim de elucidar a mecânica de sua operação, provaremos primeiro um lema.

Lema 3.5.1 *Sejam A uma álgebra sobre um corpo K e $F : A^n \to A$ uma aplicação em n variáveis que é linear em cada argumento. Se $k \geqslant n$, então para quaisquer $a_1, ..., a_k$ em A, temos que*

$$F(a_1 + ... + a_k, ..., a_1 + ... + a_k) -$$

$$-\sum_{q=1}^{k} F(a_1+...+\widehat{a}_q+...+a_k,...,a_1+...+\widehat{a}_q+...+a_k)+$$

$$+\sum_{1\leqslant q_1<q_2\leqslant k} F(a_1+...+\widehat{a}_{q_1}+...+\widehat{a}_{q_2}+...+a_k,...,a_1+...+\widehat{a}_{q_1}+...+\widehat{a}_{q_2}+...+a_k)+$$

$$+...+(-1)^{k-1}\sum_{q=1}^{k} F(a_q,...,a_q)=$$

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

Prova: Seja $k \geqslant n$. Utilizando a linearidade da função F, podemos remover todas as somas de dentro dos argumentos desta função. Assim o lado esquerdo da igualdade a ser provada fica representada como uma combinação linear de elementos da forma $F(a_{j_1},...,a_{j_n})$ com coeficientes inteiros. Se existem s índices distintos entre $j_1,...,j_n$ e $s<k$, então o coeficiente para $F(a_{j_1},...,a_{j_n})$ é igual à soma alternante

$$1-\binom{k-s}{1}+\binom{k-s}{2}+...+(-1)^{k-s}\binom{k-s}{k-s},$$

que é igual a zero. Observemos que s não excede n, e que k por hipótese não é menor que n. Assim $s \geqslant k$ ocorre apenas se $s=k=n$. Neste caso o coeficiente para $F(a_{j_1},...,a_{j_n})$ é claramente igual a um, e a demonstração do lema está concluída. ■

Proposição 3.5.2 *Para quaisquer p e q em $K\{\mathbf{X}\}$ nos quais o operador $\mathbf{L}_i^k$ está definido, temos que $(p+q)\mathbf{L}_i^k = p\mathbf{L}_i^k + q\mathbf{L}_i^k$. Se $p=p(X_1,...,X_m)$ é um monômio de grau n em X_i e $q=q(X_1,...,X_{i-1},Z_1,...,Z_n,X_{i+1},...,X_m)$ é um monômio linear em $Z_1,...,Z_n \in \mathbf{X}-\{X_1,...,X_m\}$ tais que $p(X_1,...,X_m)=q(X_1,...,X_{i-1},X_i,...,X_i,X_{i+1},...,X_m)$, então*

$$p\mathbf{L}_i^k(X_1,...,X_{i-1},Y_1,...,Y_k,X_{i+1},...,X_m)=$$

$$=\begin{cases}\sum_{\sigma\in S_n} q(X_1,...,X_{i-1},Y_{\sigma(1)},...,Y_{\sigma(n)},X_{i+1},...,X_m) & ,\ se\quad k=n,\\ 0 & ,\ se\quad k>n.\end{cases}$$

Prova: A linearidade de $\mathbf{L}_i^k$ é óbvia. Fixando $X_1,...,X_{i-1},X_{i+1},...,X_n$ podemos considerar o monômio q como uma função das n variáveis $Y_1,...,Y_n$. Aplicando o **Lema 3.5.1**, obtemos a fórmula para $p\mathbf{L}_i^k$. ■

Exemplo 3.5.1 *Vamos calcular $p\mathbf{L}_1^3$ para o monômio $p(X_1,X_2)=X_1X_1X_2X_1$. Considerando o monômio $q(Z_1,Z_2,Z_3,X_2)=Z_1Z_2X_2Z_3$ e aplicando a **Proposição 3.5.2**, temos que $p\mathbf{L}_1^3(Y_1,Y_2,Y_3,X_2)$ é igual a*

$$Y_1Y_2X_2Y_3+Y_1Y_3X_2Y_2+Y_2Y_1X_2Y_3+Y_2Y_3X_2Y_1+Y_3Y_1X_2Y_2+Y_3Y_2X_2Y_1.$$

Exemplo 3.5.2 *Vamos calcular $p\mathbf{L}_1^5$ para o monômio $p(X_1,X_2)=X_1X_1X_2X_1X_1$. Considerando o monômio $q(Z_1,Z_2,Z_3,Z_4,X_2)=Z_1Z_2X_2Z_3Z_4$ e aplicando a **Proposição 3.5.2**, temos que*

$$p\mathbf{L}_1^5(Y_1,Y_2,Y_3,X_2)=0.$$

Exemplo 3.5.3 *Vamos calcular* $p\mathbf{L}_1^2\mathbf{L}_2^2$ *para o monômio* $p(X_1, X_2) = X_1X_1X_2X_2$. *Considerando o monômio* $q(W_1, W_2, X_2) = W_1W_2X_2X_2$ *e aplicando a* ***Proposição 3.5.2****, temos que*

$$p\mathbf{L}_1^2(Y_1, Y_2, X_2) = Y_1Y_2X_2X_2 + Y_2Y_1X_2X_2.$$

Considerando agora os monômios $r(Y_1, Y_2, W_1, W_2) = Y_1Y_2W_1W_2$ *e* $s(Y_1, Y_2, W_1, W_2) = Y_2Y_1W_1W_2$, *temos que*

$$(Y_1Y_2X_2X_2)\mathbf{L}_2^2(Y_1, Y_2, Z_1, Z_2) = Y_1Y_2Z_1Z_2 + Y_1Y_2Z_2Z_1$$

e

$$(Y_2Y_1X_2X_2)\mathbf{L}_2^2(Y_1, Y_2, Z_1, Z_2) = Y_2Y_1Z_1Z_2 + Y_2Y_1Z_2Z_1.$$

Portanto, $p\mathbf{L}_1^2\mathbf{L}_2^2(Y_1, Y_2, Z_1, Z_2) = Y_1Y_2Z_1Z_2 + Y_1Y_2Z_2Z_1 + Y_2Y_1Z_1Z_2 + Y_2Y_1Z_2Z_1$.

Proposição 3.5.3 *Se* $p = p(X_1, ..., X_n)$ *é um polinômio homogêneo de tipo* $[k_1, ..., k_n]$, *então o polinômio* $p\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n}$ *é um polinômio multilinear.*

Prova: A demonstração é bastante simples utilizando a **Proposição 3.5.2.** ■

Definição 3.5.2 *Seja* $p = p(X_1, ..., X_n)$ *um polinômio homogêneo de tipo* $[k_1, ..., k_n]$. *O polinômio multilinear* $p\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n}$ *é chamado a* linearização completa do polinômio p.

Teorema 3.5.1 *Sejam* A *uma álgebra associativa sobre um corpo* K *e* $p = p(X_1, ..., X_n)$ *um polinômio não-nulo de* $K\{\mathbf{X}\}$ *que se anula pela substituição de quaisquer elementos de um subgrupo* P *do grupo aditivo da álgebra* A. *Então a linearização completa de quaisquer das suas componentes homogêneas com grau maximal também se anula em* P.

Prova: Seja $p = p_1 + ... + p_s$ uma decomposição do polinômio p numa soma de componentes homogêneas, onde o grau de p_1 é maximal.

Suponhamos que alguma variável X_j não aparece em p_1 mas aparece em outras componentes homogêneas. Então o polinômio

$$p'(X_1, ..., X_n) = p(X_1, ..., X_{j-1}, 0, X_{j+1}, ..., X_n)$$

também se anula em P, tem p_1 como componente de grau maximal, mas p' não contém X_j.

Por outro lado, se X_j aparece em p_1 mas não aparece em qualquer outra componente homogênea do polinômio p', então podemos considerar o novo polinômio

$$p''(X_1, ..., X_n) = p'(X_1, ..., X_n) - p'(X_1, ..., X_{j-1}, 0, X_{j+1}, ..., X_n).$$

Este polinômio também se anula em P, mas X_j agora aparece em todas as componentes homogêneas do polinômio p'', e, como anteriormente, p_1 é uma componente de grau maximal em p''.

Assim, renumerando as variáveis se necessário, podemos assumir que a componente homogênea p_1 do polinômio p tem tipo $[k_1, ..., k_n]$, onde $k_i \neq 0$ para $i = 1, 2, ..., n$, e demais componentes homogêneas (p_2 por exemplo) têm tipo $[m_1, ..., m_n]$, onde $m_i \neq 0$ para

$i = 1, 2, ..., n$, e onde a desigualdade $m_j < k_j$ mantém-se para algum j. Como $m_j < k_j$, então $p_2\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n} = 0$ pela **Proposição 3.5.2**. Conseqüentemente,

$$p_1\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n} = p\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n}.$$

Pela **Proposição 3.5.1**, concluímos que a linearização completa

$$p_1\mathbf{L}_1^{k_1}\mathbf{L}_2^{k_2}...\mathbf{L}_n^{k_n}$$

da componente p_1 anula-se em P. ■

Deste teorema decorre um corolário que é o resultado mais importante desta seção.

Corolário 3.5.1 *Se uma álgebra associativa A satisfaz uma identidade polinomial de grau n, então A também satisfaz uma identidade polinomial multilinear de grau n.*

Exemplo 3.5.4 *Se A é uma nil álgebra de índice limitado n, então*

$$q(X_1, ..., X_n) = \sum_{\sigma \in S_n} X_{\sigma(1)}...X_{\sigma(n)}$$

é uma identidade polinomial de A, pois é a linearização completa do polinômio $p(X_1) = X_1^n$, o que está de acordo com os casos particulares do início desta seção.

Capítulo 4

Teorema de Amitsur-Levitzki

Apresentamos uma das provas do Teorema de Amitsur-Levitzki. Este teorema afirma que s_{2n} é uma identidade polinomial da álgebra das matrizes de ordem n.

4.1 Identidades polinomiais na álgebra das matrizes

Seja K um corpo e n um número natural. Denotamos por $M_n(K)$ o conjunto das matrizes $n \times n$ com entradas no corpo K. Na verdade, $M_n(K)$ é uma álgebra associativa com identidade sobre K, conhecida como a álgebra das matrizes de ordem n. Será que $M_n(K)$ é uma PI álgebra? Sim, pois $s_{n^2+1}(X_1, X_2, ..., X_{n^2+1})$ é uma identidade polinomial de $M_n(K)$, dado que sua dimensão é n^2. Será que existem outras identidades polinomiais em $M_n(K)$, cujo grau seja menor que n^2+1? O Teorema da Escada, cujo nome torna-se bastante natural ao compreendermos sua demonstração, mostra que não existem identidades polinomiais de grau menor que $2n$ em $M_n(K)$.

Teorema 4.1.1 (da Escada) *Se K é um corpo e n é um número natural, então a álgebra $M_n(K)$ não satisfaz nenhuma identidade polinomial de grau menor que $2n$.*

Prova: Suponhamos, por absurdo, que exista uma identidade polinomial p de grau menor que $2n$. Conforme o **Corolário 3.5.1**, podemos deduzir a existência de uma identidade polinomial multilinear $q(X_1, ..., X_g)$ cujo grau g é menor que $2n$. Podemos assumir também que entre os monômios que aparecem em q, temos um monômio $\alpha X_1 ... X_g$,com $\alpha \neq 0$. Substituamos $X_1, ..., X_g$ por $e_{11}, e_{12}, e_{22}, e_{23}, ..., e_{kk+1}$ se $g = 2k$ é par, ou por $e_{11}, e_{12}, e_{22}, e_{23}, ..., e_{kk}$ se $g = 2k-1$ é ímpar. É claro que o único produto destes elementos que não é zero é $e_{11}e_{12}e_{22}e_{23}...e_{kk+1} = e_{1k+1}$ se $g = 2k$ ou $e_{11}, e_{12}, e_{22}, e_{23}, ..., e_{kk} = e_{1k}$ se $g = 2k-1$. Logo,

$$q(e_{11}, e_{12}, e_{22}, e_{23}, ..., e_{kk+1}) = \alpha e_{1k+1} \neq 0 \text{ se } g = 2k$$

ou

$$q(e_{11}, e_{12}, e_{22}, e_{23}, ..., e_{kk}) = \alpha e_{1k} \neq 0 \text{ se } g = 2k-1,$$

o que é uma contradição. ■

Podemos nos perguntar: existem identidades polinomiais em $M_n(K)$ com grau maior que $2n-1$ e menor que n^2+1? No caso de $n = 1$, isto é óbvio, pois $s_2(X_1, X_2) = X_1X_2 - X_2X_1$ e

a multiplicação no corpo K é comutativa. Na verdade, este resultado pode ser generalizado para um número natural n qualquer. Em 1950, S. A. Amitsur e J. Levitzki [2] provaram que

$$\sum_{\sigma \in S_{2n}} (-1)^{\sigma} A_{\sigma(1)} ... A_{\sigma(2n)} = 0$$

para quaisquer matrizes $A_1, A_2, ..., A_{2n}$ em $M_n(K)$. Após a publicação deste teorema, um marco na história das PI álgebras, surgiram outras três provas diferentes para ele, cada vez mais simples. A primeira foi elaborada em 1958 por B. Kostant [17]. Em 1963, utilizando teoria dos grafos, R. G. Swan [36], [37] apresentou uma nova prova. Por fim, em 1976, S. Rosset [31] publicou um artigo de apenas duas páginas com uma demonstração envolvendo álgebra exterior. Neste capítulo apresentamos a prova de Swan que pode ser encontrada em Procesi [27] e é bastante simples, mas apóia-se num teorema da teoria dos grafos que desenvolveremos na próxima seção. No Capítulo 7, apresentamos a prova de Rosset.

4.2 Um teorema sobre grafos

Definição 4.2.1 *Um* grafo orientado $G = (V, A, \varepsilon_1, \varepsilon_2)$ *é uma quádrupla ordenada na qual* V *é um conjunto finito e não-vazio cujos elementos são chamados* vértices, A *é um conjunto finito cujos elementos são chamados* arestas, *e* ε_1 *e* ε_2 *são funções de* A *em* V. *Se* $a \in A$ *então* $\varepsilon_1(a)$ *é chamado* extremo inicial *de* a *e* $\varepsilon_2(a)$ *é chamado* extremo final *de* a.

Definição 4.2.2 *Um* caminho orientado C *num grafo orientado* $G = (V, A, \varepsilon_1, \varepsilon_2)$ *é uma seqüência finita* $a_1, ..., a_k$ *de arestas tais que* $\varepsilon_2(a_i) = \varepsilon_1(a_{i+1})$ *para* $i = 1, ..., k-1$. *Se* $\varepsilon_1(a_1) = u$ *e* $\varepsilon_2(a_k) = v$, *dizemos que* C *é um caminho orientado de* u *até* v.

Exemplo 4.2.1 *Consideremos o grafo orientado* $G = (V, A, \varepsilon_1, \varepsilon_2)$ *onde* $V = \{u, v, w, x, y\}$, $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$, $\varepsilon_1(a_1) = u$ *e* $\varepsilon_2(a_1) = v$, $\varepsilon_1(a_2) = u$ *e* $\varepsilon_2(a_2) = x$, $\varepsilon_1(a_3) = x$ *e* $\varepsilon_2(a_3) = v$, $\varepsilon_1(a_4) = v$ *e* $\varepsilon_2(a_4) = w$, $\varepsilon_1(a_5) = w$ *e* $\varepsilon_2(a_5) = y$, $\varepsilon_1(a_6) = v$ *e* $\varepsilon_2(a_6) = y$, $\varepsilon_1(a_7) = y$ *e* $\varepsilon_2(a_7) = x$ *(veja a Figura 4.1). Neste grafo orientado, temos, entre outros, o caminho orientado* $C = a_1, a_4, a_5, a_7$ *de* u *até* x.

Definição 4.2.3 *Um* caminho unicursal C *num grafo orientado* G *é um caminho orientado tal que cada aresta de* G *aparece uma e apenas uma vez em* C. *Se a seqüência* $a_1, ..., a_k$ *de arestas é um caminho unicursal então existe uma bijeção entre* $\{1, ..., k\}$ *e* A; *numerando as arestas de* A, *podemos considerar cada caminho unicursal como uma permutação do conjunto* $\{1, ..., k\}$. *Assim, definimos* caminho unicursal par *como um caminho unicursal cuja permutação associada é par e* caminho unicursal ímpar *como um caminho unicursal cuja permutação associada é ímpar.*

Definição 4.2.4 *Se* G *é um grafo orientado e* u *e* v *são vértices de* G, *denotamos por* $n_1(u, v)$ *o número de caminhos unicursais ímpares de* u *até* v *e por* $n_2(u, v)$ *o número de caminhos unicursais pares de* u *até* v.

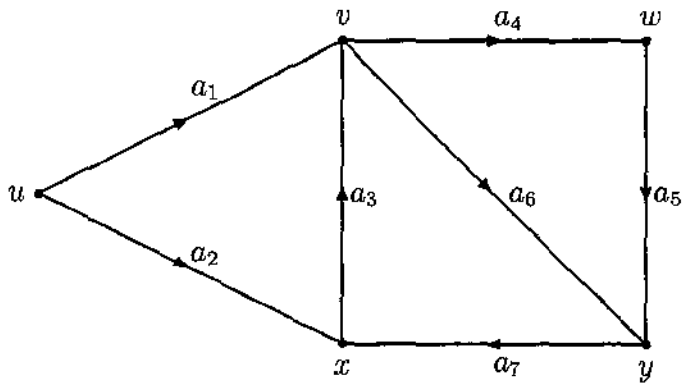


Figura 4.1: Desenho do grafo do Exemplo 4.2.1.

Exemplo 4.2.2 *Consideremos o grafo orientado* $G = (V, A, \varepsilon_1, \varepsilon_2)$ *onde* $V = \{u, v, w, x\}$, $A = \{a_1, a_2, a_3, a_4, a_5\}$, $\varepsilon_1(a_1) = u$ *e* $\varepsilon_2(a_1) = w$, $\varepsilon_1(a_2) = w$ *e* $\varepsilon_2(a_2) = v$, $\varepsilon_1(a_3) = u$ *e* $\varepsilon_2(a_3) = v$, $\varepsilon_1(a_4) = x$ *e* $\varepsilon_2(a_4) = u$, $\varepsilon_1(a_5) = v$ *e* $\varepsilon_2(a_5) = x$ *(veja a Figura 4.2). Os únicos caminhos unicursais de u até v são* $C_1 = a_1, a_2, a_5, a_4, a_3$ *e* $C_2 = a_3, a_5, a_4, a_1, a_2$*; as permutações associadas a eles são respectivamente* $\sigma_1 = (35)$ *e* $\sigma_2 = (134)(25) = (14)(13)(25)$*; portanto, G tem apenas 2 caminhos unicursais de u até v, ambos ímpares.*

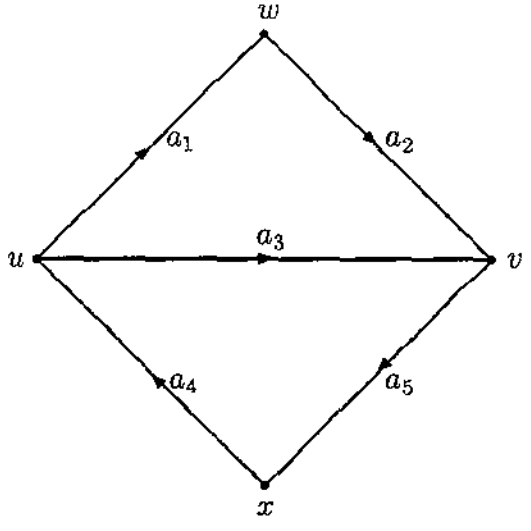


Figura 4.2: Desenho do grafo do Exemplo 4.2.2.

Definição 4.2.5 *Se* $G = (V, A, \varepsilon_1, \varepsilon_2)$ *é um grafo orientado, definimos para um vértice v os números*

$$s(v) = |\{a \in A : \varepsilon_1(a) = v\}|,$$

$$e(v) \;=\; |\{a \in A : \varepsilon_2(a) = v\}|\,.$$

Além disso, definimos o fluxo *e a* ordem *de um vértice como*

$$\begin{aligned} fluxo(v) &= s(v) - e(v), \\ ordem(v) &= s(v) + e(v). \end{aligned}$$

Teorema 4.2.1 *Se $G = (V, A, \varepsilon_1, \varepsilon_2)$ é um grafo orientado tal que $|A| \geqslant 2|V|$, então $n_1(u, v) = n_2(u, v)$ para quaisquer vértices u e v de G.*

Prova: Para provar o teorema vamos primeiro fazer algumas reduções e então prosseguir utilizando indução finita.

Primeira redução: se $|A| > 2|V|$, seja $k = |A| - 2|V|$. Vamos modificar o grafo G conectando ao vértice u uma cadeia de k novas arestas, ou seja, consideramos o novo grafo orientado $G' = (V', A', \varepsilon_1', \varepsilon_2')$ onde $V' = V \cup \{v_1, ..., v_k\}$, $A' = A \cup \{a_1, ..., a_k\}$, $\varepsilon_1'(a) = \varepsilon_1(a)$ e $\varepsilon_2'(a) = \varepsilon_2(a)$ se $a \in A$, $\varepsilon_1'(a_i) = v_i$ e $\varepsilon_2'(a_i) = v_{i+1}$ para $1 \leqslant i \leqslant k-1$, $\varepsilon_1'(a_k) = v_k$ e $\varepsilon_2'(a_k) = u$ (veja a Figura 4.3). No novo grafo G', $|V'| = |V| + k$, $|A'| = |A| + k$ e existe uma bijeção entre os caminhos unicursais de u até v em G e de v_1 até v em G', preservando a paridade. Assim é suficiente provar o teorema para G'. Por sua vez, em G' temos $|A'| = 2|V'|$, e esta é a primeira redução.

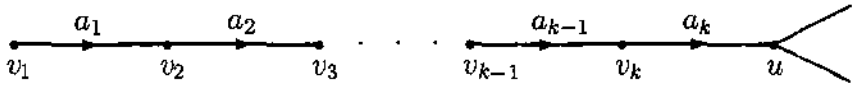


Figura 4.3: Desenho de uma parte do grafo G' da primeira redução.

Segunda redução: Assumamos que em G há pelo menos um caminho unicursal $a_1, ..., a_n$ de u até v. Como $\varepsilon_2(a_i) = \varepsilon_1(a_{i+1})$ para $1 \leqslant i \leqslant n-1$, então
(i) se $u = v$ então $fluxo(x) = 0$ para todo vértice x;
(ii) se $u \neq v$ então $fluxo(x) = 0$ para todo vértice x com exceção de u e v, para os quais $fluxo(u) = 1$ e $fluxo(v) = -1$. Neste caso, adicionamos um novo vértice w e duas novas arestas b_1 e b_2 da seguinte maneira: consideramos o novo grafo orientado $G' = (V', A', \varepsilon_1', \varepsilon_2')$ onde $V' = V \cup \{w\}$, $A' = A \cup \{b_1, b_2\}$, $\varepsilon_1'(a) = \varepsilon_1(a)$ e $\varepsilon_2'(a) = \varepsilon_2(a)$ se $a \in A$, $\varepsilon_1'(b_1) = v$ e $\varepsilon_2'(b_1) = w$, $\varepsilon_1'(b_2) = w$ e $\varepsilon_2'(b_2) = u$ (veja a Figura 4.4). O grafo G' tem $|A'| = |A| + 2$ arestas e $|V'| = |V| + 1$ vértices; novamente, temos $|A'| = 2|V'|$ e existe uma bijeção entre os caminhos unicursais de u até v em G e de w até w em G', preservando a paridade. Além disso, em G' todos os vértices têm fluxo nulo, e esta é a segunda redução.

Resumindo nossas reduções, podemos considerar apenas grafos orientados $G = (V, A, \varepsilon_1, \varepsilon_2)$ tais que
(i) $|A| = 2|V|$;
(ii) $fluxo(x) = 0$ para qualquer vértice x;
(iii) $u = v$ (chamaremos este vértice de vértice base).

Vamos iniciar a prova sob estas hipóteses. Faremos a construção de novos grafos orientados sempre procurando que eles satisfaçam estas hipóteses. Podemos considerar 3 casos:

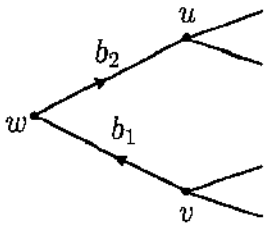


Figura 4.4: Desenho de uma parte do grafo G' da segunda redução.

Primeiro caso: Suponhamos que o grafo orientado G tem uma configuração formada por 3 vértices v_1, v_2, v_3 e por 3 arestas a_1, a_2, a_3 tais que $\varepsilon_1(a_1) = v_1$ e $\varepsilon_2(a_1) = v_2$, $\varepsilon_1(a_2) = v_2$ e $\varepsilon_2(a_2) = v_2$, $\varepsilon_1(a_3) = v_2$ e $\varepsilon_2(a_3) = v_3$, e $ordem(v_2) = 4$. Se v_2 é o vértice base então qualquer caminho unicursal começa ou termina com a_2. Invertendo o papel de a_2 como aresta inicial ou final, temos uma bijeção entre os caminhos unicursais pares e ímpares (já que $|A|$ é um número par). Se o vértice base não é v_2 então todo caminho unicursal contém $a_1a_2a_3$. Podemos transformar as 3 arestas a_1, a_2, a_3 em uma eliminando o vértice v_2, ou seja, consideramos o novo grafo orientado $G' = (V', A', \varepsilon_1', \varepsilon_2')$ onde $V' = V - \{v_2\}$, $A' = (A - \{a_1, a_2, a_3\}) \cup \{b\}$, $\varepsilon_1'(a) = \varepsilon_1(a)$ e $\varepsilon_2'(a) = \varepsilon_2(a)$ se $a \in A - \{a_1, a_2, a_3\}$, $\varepsilon_1'(b) = v_1$ e $\varepsilon_2'(b) = v_3$ (veja a Figura 4.5). Um caminho unicursal no grafo orientado G' é exatamente um caminho unicursal em G com $a_1a_2a_3$ substituídos por b. O grafo orientado G' tem $|V'| = |V| - 1$ vértices e $|A'| = |A| - 2$ arestas, logo $|A'| = 2|V'|$; além disso, existe uma bijeção entre os caminhos unicursais de G e de G' que preserva a paridade. Como G' tem menos vértices que G podemos aplicar indução finita e concluir a prova para este caso.

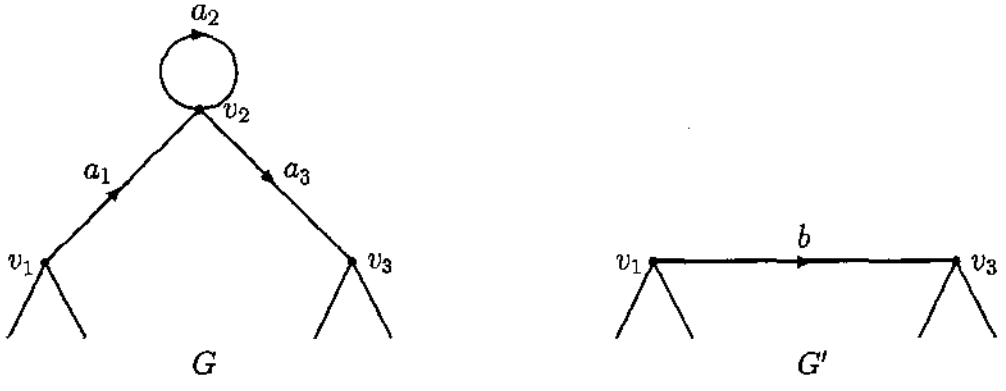


Figura 4.5: Desenhos de partes dos grafos G e G' do primeiro caso.

Segundo caso: Suponhamos que o grafo orientado G tenha um vértice de grau 2. Logo, G tem uma configuração formada por 3 vértices v_1, v_2, v_3 e por $k+2$ arestas $a_1, ..., a_k, a_{k+1}, a_{k+2}$ tais que $\varepsilon_1(a_{k+1}) = v_1$ e $\varepsilon_2(a_{k+1}) = v_2$, $\varepsilon_1(a_{k+2}) = v_2$ e $\varepsilon_2(a_{k+2}) = v_3$, $\varepsilon_2(a_i) = v_1$ para $1 \leqslant i \leqslant k$, $ordem(v_1) > 2$ e $ordem(v_2) = 2$. Para cada $i = 1, ..., k$, construímos um novo

grafo orientado $G_i = (V, A_i, \varepsilon_1^i, \varepsilon_2^i)$ onde $A_i = (A - \{a_i, a_{k+1}\}) \cup \{b_i, b\}$, $\varepsilon_1^i(a) = \varepsilon_1(a)$ e $\varepsilon_2^i(a) = \varepsilon_2(a)$ se $a \in A - \{a_i, a_{k+1}\}$, $\varepsilon_1^i(b_i) = \varepsilon_1(a_i)$ e $\varepsilon_2^i(b_i) = v_2$, $\varepsilon_1^i(b) = v_2$ e $\varepsilon_2^i(b) = v_2$ (veja a Figura 4.6). O grafo orientado G_i satisfaz todas as hipóteses e cai no caso anterior. Cada caminho unicursal em G gera um caminho unicursal em G_i para algum i e vice-versa, preservando a paridade. Assim, caímos no caso anterior.

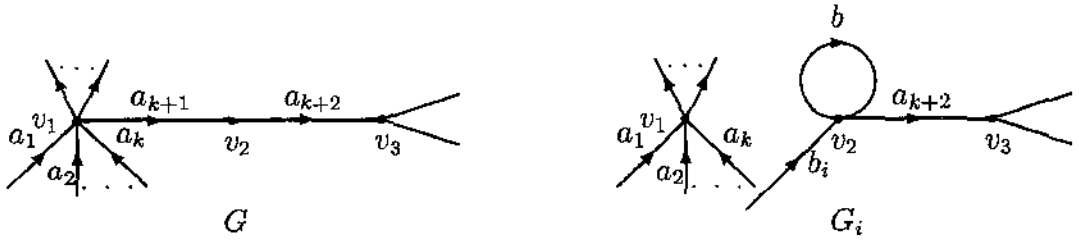


Figura 4.6: Desenhos de partes dos grafos G e G_i do segundo caso.

Terceiro caso: Suponhamos que não ocorra nenhum dos dois casos anteriores. Logo, $ordem(x) \geqslant 4$ para todo vértice x de G. Como $\sum_{x \in V} ordem(x) = 2|A| = 4|V|$ então todo vértice tem ordem 4. Desde que não temos configurações como a do 1° caso, o grafo orientado G deve conter uma configuração formada por 8 vértices $w, t, v_1, v_2, v_3, v_4, v_5, v_6$ e por 7 arestas $a, a_1, a_2, a_3, a_4, a_5, a_6$ tais que $\varepsilon_1(a) = w$ e $\varepsilon_2(a) = t$, $\varepsilon_1(a_1) = v_1$ e $\varepsilon_2(a_1) = w$, $\varepsilon_1(a_2) = v_2$ e $\varepsilon_2(a_2) = w$, $\varepsilon_1(a_3) = t$ e $\varepsilon_2(a_3) = v_3$, $\varepsilon_1(a_4) = t$ e $\varepsilon_2(a_4) = v_4$, $\varepsilon_1(a_5) = w$ e $\varepsilon_2(a_5) = v_5$, $\varepsilon_1(a_6) = v_6$ e $\varepsilon_2(a_6) = t$ (veja a Figura 4.7).

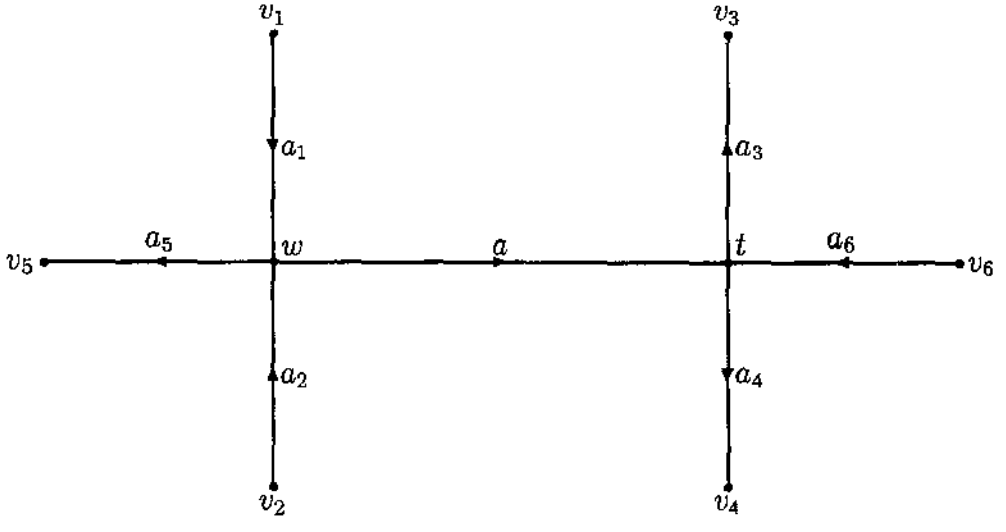


Figura 4.7: Desenho do grafo G do terceiro caso.

Construímos novos grafos orientados $G_1 = (V, A_1, \varepsilon_1^1, \varepsilon_2^1)$ onde $A_1 = (A - \{a, a_1\}) \cup \{b, b_1\}$, $\varepsilon_1^1(a) = \varepsilon_1(a)$ e $\varepsilon_2^1(a) = \varepsilon_2(a)$ se $a \in A - \{a, a_1\}$, $\varepsilon_1^1(b) = t$ e $\varepsilon_2^1(b) = t$, $\varepsilon_1^1(b_1) = v_1$ e $\varepsilon_2^1(b_1) = t$

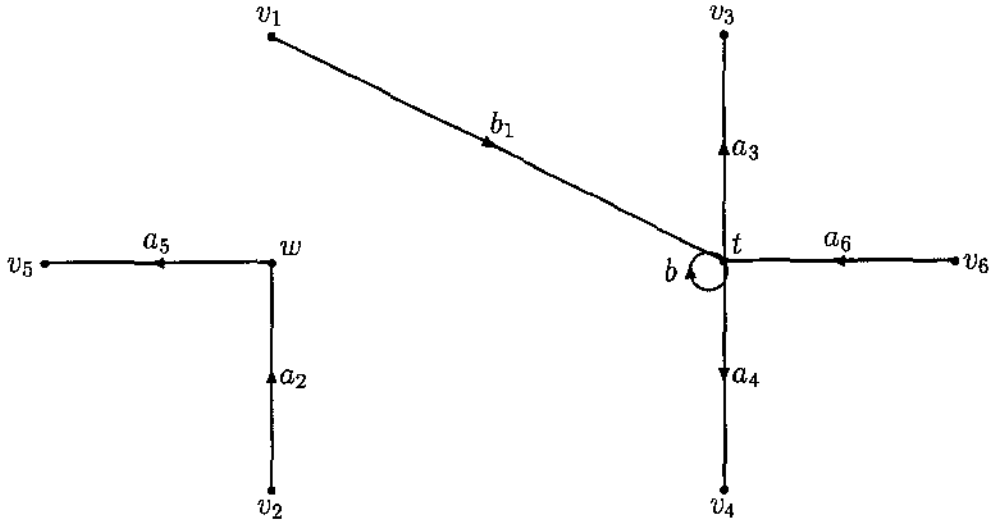


Figura 4.8: Desenho do grafo G_1 do terceiro caso.

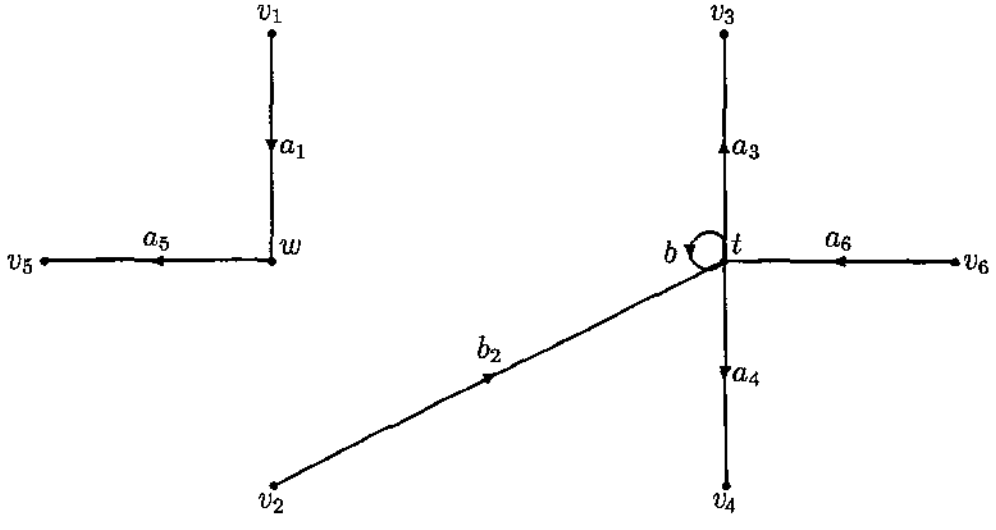


Figura 4.9: Desenho do grafo G_2 do terceiro caso.

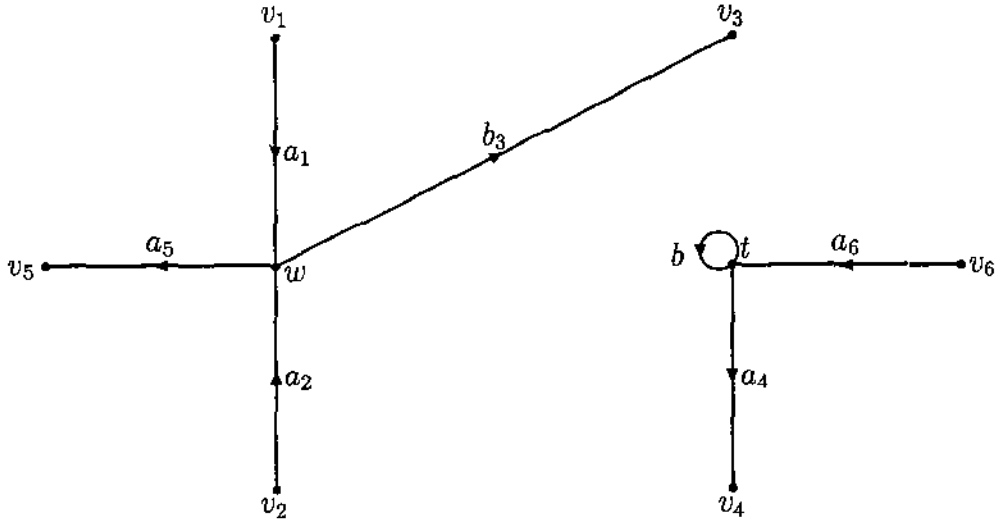


Figura 4.10: Desenho do grafo G_3 do terceiro caso.

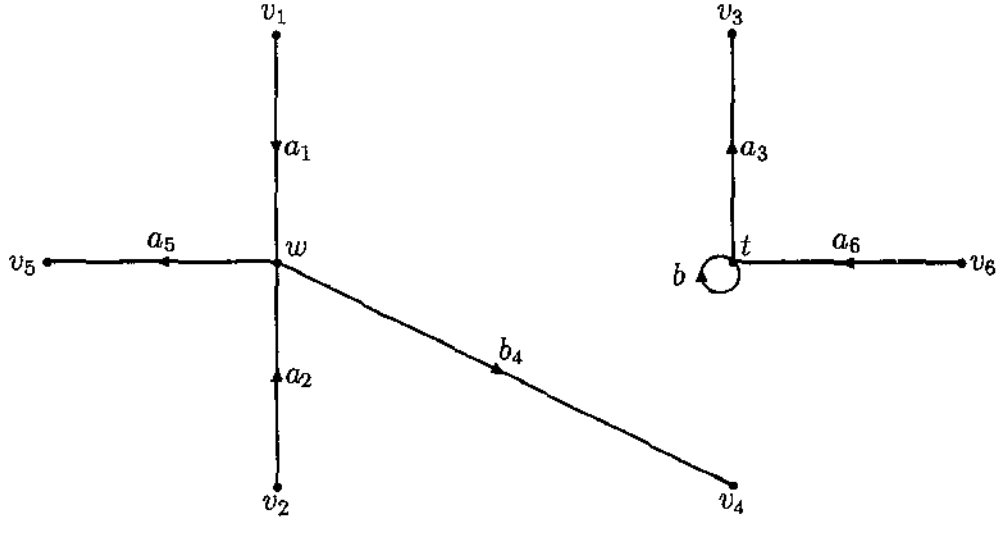


Figura 4.11: Desenho do grafo G_4 do terceiro caso.

e $G_2 = (V, A_2, \varepsilon_1^2, \varepsilon_2^2)$ onde $A_2 = (A - \{a, a_2\}) \cup \{b, b_2\}$, $\varepsilon_1^2(a) = \varepsilon_1(a)$ e $\varepsilon_2^2(a) = \varepsilon_2(a)$ se $a \in A - \{a, a_2\}$, $\varepsilon_1^2(b) = t$ e $\varepsilon_2^1(b) = t$, $\varepsilon_1^1(b_2) = v_2$ e $\varepsilon_2^1(b_2) = t$ (veja a Figuras 4.8 e a Figura 4.9). Qualquer caminho unicursal de G contendo $a_1 a$ gera um caminho unicursal em G_1, e um contendo $a_2 a$ gera um caminho unicursal em G_2; além disso, a paridade é preservada.

No entanto, há caminhos unicursais em G_1 e em G_2 que não provém de caminhos unicursais em G, explicitamente são aqueles contendo $b_1 a_3$ ou $b_1 a_4$ em G_1 e aqueles contendo $b_2 a_3$ ou $b_2 a_4$ em G_2; eles são exatamente os caminhos unicursais nos grafos $G_3 = (V, A_3, \varepsilon_1^3, \varepsilon_2^3)$ onde $A_3 = (A - \{a, a_3\}) \cup \{b, b_3\}$, $\varepsilon_1^3(a) = \varepsilon_1(a)$ e $\varepsilon_2^3(a) = \varepsilon_2(a)$ se $a \in A - \{a, a_3\}$, $\varepsilon_1^3(b) = t$ e $\varepsilon_2^3(b) = t$, $\varepsilon_1^3(b_3) = w$ e $\varepsilon_2^3(b_3) = v_3$ e $G_4 = (V, A_4, \varepsilon_1^4, \varepsilon_2^4)$ onde $A_4 = (A - \{a, a_4\}) \cup \{b, b_4\}$, $\varepsilon_1^4(a) = \varepsilon_1(a)$ e $\varepsilon_2^4(a) = \varepsilon_2(a)$ se $a \in A - \{a, a_4\}$, $\varepsilon_1^4(b) = t$ e $\varepsilon_2^4(b) = t$, $\varepsilon_1^4(b_4) = w$ e $\varepsilon_2^4(b_4) = v_4$ (veja a Figura 4.10 e a Figura 4.11). As paridades são preservadas.

Como o teorema é verdadeiro para G_1, G_2, G_3 e G_4 pelos casos anteriores, é verdadeiro para G. ■

Observação 4.2.1 *A condição $|A| \geqslant 2|V|$ é realmente necessária para o* ***Teorema 4.1.1.*** *Consideremos o grafo $G = (V, A, \varepsilon_1, \varepsilon_2)$ onde $V = \{v_1, ..., v_n\}$, $A = \{a_1, ..., a_{2n-1}\}$, $\varepsilon_1(a_i) = v_i$ e $\varepsilon_2(a_i) = v_{i+1}$ se $1 \leqslant i \leqslant n-1$, $\varepsilon_1(a_n) = v_n$ e $\varepsilon_2(a_n) = v_n$, $\varepsilon_1(a_i) = v_{i-n+1}$ e $\varepsilon_2(a_i) = v_{i-n}$ se $n+1 \leqslant i \leqslant 2n-1$ (veja a Figura 4.12). Há apenas um caminho unicursal de v_1 até v_1.*

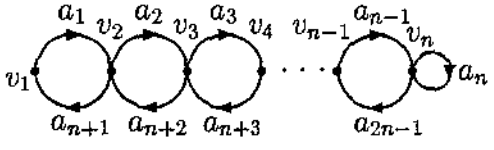


Figura 4.12: Desenho do grafo da Observação 4.2.1.

Exemplo 4.2.3 *No* ***Exemplo 4.1.2,*** *$n_1(u, v) = 2 \neq 0 = n_2(u, v)$ pois $|A| = 5 < 8 = 2|V|$ e o* ***Teorema 4.1.1*** *não pode ser aplicado.*

Exemplo 4.2.4 *Consideremos o grafo orientado $G = (V, A, \varepsilon_1, \varepsilon_2)$ onde $V = \{u, v, w\}$, $A = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, $\varepsilon_1(a_1) = u$ e $\varepsilon_2(a_1) = w$, $\varepsilon_1(a_2) = w$ e $\varepsilon_2(a_2) = v$, $\varepsilon_1(a_3) = u$ e $\varepsilon_2(a_3) = v$, $\varepsilon_1(a_4) = v$ e $\varepsilon_2(a_4) = u$, $\varepsilon_1(a_5) = u$ e $\varepsilon_2(a_5) = u$, $\varepsilon_1(a_6) = w$ e $\varepsilon_2(a_6) = w$ (veja a Figura 4.13). Como $|A| = 6 \leqslant 6 = 2|V|$, o* ***Teorema 4.1.1*** *é válido para G. Escolhendo os vértices u e v, temos que os únicos caminhos unicursais de u até v são $C_1 = a_1, a_6, a_2, a_4, a_5, a_3$, $C_2 = a_3, a_4, a_5, a_1, a_6, a_2$, $C_3 = a_5, a_1, a_6, a_2, a_4, a_3$ e $C_4 = a_5, a_3, a_4, a_1, a_6, a_2$; as permutações associadas a eles são respectivamente $\sigma_1 = (236) = (23)(26)$, $\sigma_2 = (135624) = (14)(12)(16)(15)(13)$, $\sigma_3 = (1542)(36) = (12)(14)(15)(36)$ e $\sigma_4 = (156234) = (14)(13)(12)(16)(15)$; portanto, G tem apenas 4 caminhos unicursais de u até v, 2 pares e 2 ímpares.*

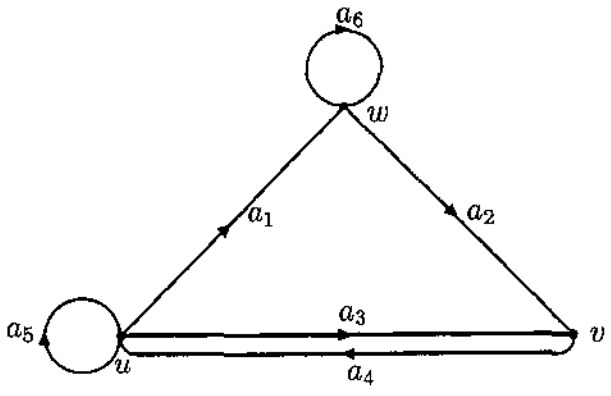


Figura 4.13: Desenho do grafo do Exemplo 4.2.4.

4.3 Teorema de Amitsur-Levitzki

A seguir, provamos o Teorema de Amitsur-Levitzki pela prova de Swan.

Teorema 4.3.1 (Amitsur-Levitzki) *Se K é um corpo e n um número natural, então s_{2n} é uma identidade polinomial de $M_n(K)$.*

Prova: Como s_{2n} é multilinear, é suficiente mostrar que este polinômio se anula quando calculado em elementos de uma base de $M_n(K)$. Escolhemos a base usual, formada pelas matrizes e_{ij}, nas quais a posição ij é igual a 1 e as demais posições são nulas. Para simplificar escreveremos (i,j) ao invés de e_{ij}. Tomemos $2n$ matrizes desta base

$$A_1 = (i_1, j_1), A_2 = (i_2, j_2), ..., A_{2n} = (i_{2n}, j_{2n})$$

e calculemos $s_{2n}(A_1, ..., A_{2n})$. Se $A_i = A_j$ para i e j distintos, então $s_{2n}(A_1, ..., A_{2n}) = 0$. Assumimos então que as matrizes $A_1, ..., A_{2n}$ são distintas. Construímos um grafo orientado cujos vértices são os números $1, 2, ..., n$ e cujas arestas são os $2n$ pares $(i_1, j_1), (i_2, j_2), ..., (i_{2n}, j_{2n})$; definimos também $\varepsilon_1(i,j) = i$ e $\varepsilon_2(i,j) = j$. Um produto $A_{\sigma(1)}A_{\sigma(2)}...A_{\sigma(2n)}$ é não-nulo se e somente se representa um caminho unicursal neste grafo, e ele produz o valor (i,j) se e somente se começa em i e termina em j. Pelo **Teorema 4.2.1**, (i,j) é obtido o mesmo número de vezes através de permutações pares e ímpares. Logo, na soma $\sum_{\sigma \in S_{2n}} (-1)^{\sigma} A_{\sigma(1)}...A_{\sigma(2n)}$ as duas contribuições são canceladas. Portanto, $s_{2n}(A_1, ..., A_{2n}) = 0$. ■

Capítulo 5

Teorema de Higman

Em 1952, G. Higman [10] publicou um artigo provando um teorema, hoje conhecido como Teorema de Higman, que é uma ferramenta importante utilizada no Capítulo 6. Espomos este resultado seguindo o artigo original de Higman.

5.1 Propriedade da base finita

Definição 5.1.1 *Uma* relação de quasi-ordem *num conjunto P é uma relação simbolizada por $\leqslant$ e que para quaisquer x, y e z em P satisfaz as seguintes propriedades:*
(i) $x \leqslant x$ (reflexiva);
(ii) se $x \leqslant y$ *e* $y \leqslant z$ *então* $x \leqslant z$ (transitiva).
Um conjunto P no qual está definida uma relação de quasi-ordem é chamado um conjunto quasi-ordenado.

Definição 5.1.2 *Uma relação de quasi-ordem num conjunto P é chamada uma* relação de ordem parcial *se para quaisquer x e y em P satisfaz a seguinte propriedade:*
(iii) se $x \leqslant y$ *e* $y \leqslant x$ *então* $x = y$ (anti-simétrica).
Um conjunto P no qual está definida uma relação de ordem parcial é chamado um conjunto parcialmente ordenado.

Definição 5.1.3 *Uma relação de ordem parcial num conjunto P é chamada uma* relação de ordem total (*ou* linear) *se para quaisquer x e y em P satisfaz a seguinte propriedade:*
(iv) $x \leqslant y$ *ou* $y \leqslant x$ *(x e y são chamados* comparáveis*).*
Um conjunto P no qual está definida uma relação de ordem total é chamado um conjunto totalmente ordenado (*ou um* conjunto linearmente ordenado).

Exemplo 5.1.1 *Seja* $A = \{a = (a_1, a_2, ...) : a_i \in \mathbf{N}$ *para* $i \in \mathbf{N}\}$. *Para* $a = (a_1, a_2, ...)$ *e* $b = (b_1, b_2, ...) \in A$, *definimos uma relação de quasi-ordem escrevendo que* $a \leqslant b$ *se existe uma aplicação* $\varphi : \mathbf{N} \longrightarrow \mathbf{N}$ *tal que* $a_i \leqslant b_{\varphi(i)}$ *para cada* $i \in \mathbf{N}$. *É fácil ver que esta relação de quasi-ordem não é uma relação de ordem parcial. Sejam* $a = (1, 2, 3, ...)$ *e* $b = (2, 4, 6, ...)$. *Considerando as aplicações* $\varphi : \mathbf{N} \longrightarrow \mathbf{N}$ *dada por* $\varphi(n) = 2n$ *e* $id : \mathbf{N} \longrightarrow \mathbf{N}$ *a aplicação identidade, temos que* $b \leqslant a$ *e* $a \leqslant b$ *respectivamente. No entanto,* $a \neq b$, *mostrando que não é uma relação de ordem parcial.*

Exemplo 5.1.2 *Consideremos o conjunto* **N** *dos números naturais e definamos* $m \leqslant n$ *se* m *divide* n. *Esta é uma relação de ordem parcial, porém não é uma relação de ordem total, porque, por exemplo,* 4 *não divide* 7 *e* 7 *não divide* 4.

Exemplo 5.1.3 *Consideremos* A *como o conjunto das letras do alfabeto latino. O conjunto* P *formado por todos os subconjuntos de* A *onde* $M \leqslant N$ *se* $M \subseteq N$ *é parcialmente ordenado. Porém,* P *não é totalmente ordenado pois* $\{a, b\}$ *não contém* $\{c\}$ *e* $\{c\}$ *não contém* $\{a, b\}$.

Exemplo 5.1.4 *Se tomarmos o conjunto dos números inteiros* **Z** *e* $x \leqslant y$ *com o seu significado habitual, teremos um conjunto totalmente ordenado.*

Proposição 5.1.1 *Todo subconjunto de um conjunto quasi-ordenado também é um conjunto quasi-ordenado.*

Proposição 5.1.2 *Todo subconjunto de um conjunto parcialmente ordenado também é um conjunto parcialmente ordenado.*

Proposição 5.1.3 *Todo subconjunto de um conjunto totalmente ordenado também é um conjunto totalmente ordenado.*

Definição 5.1.4 *O* fecho *de um subconjunto* S *de um conjunto quasi-ordenado* P, *denotado por* $cl(S)$, *é o conjunto de todos os elementos* b *em* P *tais que para algum* a *em* S, $a \leqslant b$.

Definição 5.1.5 *Um subconjunto de um conjunto quasi-ordenado é* fechado *se ele é igual ao seu fecho. Um subconjunto de um conjunto quasi-ordenado é* aberto *se seu complemento é fechado.*

Proposição 5.1.4 *Em qualquer conjunto quasi-ordenado* P, *o conjunto vazio* $\emptyset$ *e o conjunto todo* P *são conjuntos fechados.*

Prova: Para mostrar que $\emptyset$ é fechado, precisamos mostrar que para todo elemento a em $\emptyset$ não existe um elemento b fora de $\emptyset$ tal que $a \leqslant b$; mas desde que não existem elementos em $\emptyset$, este requisito está automaticamente satisfeito. O conjunto P é fechado, pois se $a \in P$, então todo elemento b tal que $a \leqslant b$ é um elemento de P. ■

Corolário 5.1.1 *Em qualquer conjunto quasi-ordenado* P, *o conjunto vazio* $\emptyset$ *e o conjunto todo* P *são conjuntos abertos.*

Proposição 5.1.5 *Se* S *é um subconjunto de um conjunto quasi-ordenado* P *e* F *é um subconjunto fechado de* P *tal que* $S \subseteq F$, *então* $cl(S) \subseteq F$.

Prova: Seja $a \in cl(S)$. Logo, existe b em S tal que $b \leqslant a$. Como $b \in F$ e F é fechado, temos que $a \in F$. ■

Observação 5.1.1 *Expressamos também a proposição anterior dizendo que* $cl(S)$ *é o menor conjunto fechado que contém* S.

Proposição 5.1.6 *Seja P um conjunto quasi-ordenado. Então*
(i) qualquer união de conjuntos fechados em P é fechada;
(ii) qualquer interseção de conjuntos fechados em P é fechada.

Prova: Seja $\{F_i\}$ uma classe arbitrária de conjuntos fechados em P.

Para provar (i), devemos mostrar que $F = \cup_i F_i$ é fechado. Se $\{F_i\}$ é vazio, então F é vazio, e pela **Proposicão 5.1.4**, F é fechado. Suponhamos que $\{F_i\}$ não seja vazio. Logo, se $a \in cl(F)$ então existe $b \in F$ tal que $b \leqslant a$. Como $b \in F_i$, para algum i, temos que $a \in F_i$. Logo, $cl(F) \subseteq F$, ou seja, $F = cl(F)$ e F é fechado.

Para provar (ii), devemos mostrar que $F = \cap_i F_i$ é fechado. Se $\{F_i\}$ é vazio, então $F = P$, e pela **Proposição 5.1.4**, F é fechado. Suponhamos que $\{F_i\}$ não seja vazio. Logo, se $a \in cl(F)$ então existe $b \in F$ tal $b \leqslant a$. Como $b \in F_i$, para todo i, temos que $a \in F_i$, para todo i. Logo, $cl(F) \subseteq F$, ou seja, $F = cl(F)$ e F é fechado. ■

Corolário 5.1.2 *Seja P um conjunto quasi-ordenado. Então*
(i) qualquer união de conjuntos abertos em P é aberta;
(ii) qualquer interseção de conjuntos abertos em P é aberta.

Proposição 5.1.7 *Sejam P um conjunto quasi-ordenado, S um subconjunto de P e F um subconjunto fechado de P. Então $F - S$ é um subconjunto fechado de $P - S$.*

Prova: Seja $a \in P - S$ tal que existe $b \in F - S$ onde $b \leqslant a$. Logo $a \in F$ e $a \notin S$. Portanto, $a \in F - S$. Assim, $F - S$ é fechado em $P - S$. ■

Proposição 5.1.8 *Sejam P um conjunto quasi-ordenado e A e B subconjuntos de P. Então*
(i) $cl(A \cup B) = cl(A) \cup cl(B)$;
(ii) $cl(A \cap B) \subseteq cl(A) \cap cl(B)$.

Prova: Para provar (i), consideremos $a \in cl(A \cup B)$. Logo, existe $b \in A$ ou $b \in B$ tal que $b \leqslant a$. Assim, $a \in cl(A)$ ou $b \in cl(B)$. Portanto, $cl(A \cup B) \subseteq cl(A) \cup cl(B)$. Consideremos agora $a \in cl(A) \cup cl(B)$. Logo, existe $b \in A$ ou $b \in B$ tal que $b \leqslant a$. Assim, $a \in cl(A \cup B)$. Portanto, $cl(A) \cup cl(B) \subseteq cl(A \cup B)$.

Para provar (ii), consideremos $a \in cl(A \cap B)$. Logo, existe $b \in A$ e $b \in B$ tal que $b \leqslant a$. Assim, $a \in cl(A)$ e $a \in cl(B)$. Portanto, $cl(A \cap B) \subseteq cl(A) \cap cl(B)$. ■

Observação 5.1.2 *Não necessariamente $cl(A \cap B) = cl(A) \cap cl(B)$. Se considerarmos o conjunto quasi-ordenado $\mathbf{Z}$ onde $a \leqslant b$ se a divide b, temos que $cl(\{2\} \cap \{3\}) = cl(\emptyset) = \emptyset$ e $cl(\{2\}) \cap cl(\{3\}) = 6\mathbf{Z}$.*

Definição 5.1.6 *Um conjunto quasi-ordenado P tem a* propriedade da base finita *se todo subconjunto fechado de P é fecho de um conjunto finito. Um conjunto quasi-ordenado com a propriedade da base finita é chamado* conjunto parcialmente bem ordenado.

Definição 5.1.7 *Dizemos que um subconjunto S de um conjunto quasi-ordenado P satisfaz a* condição da cadeia ascendente *se para toda seqüência infinita $a_1 \leqslant a_2 \leqslant a_3 \leqslant \ldots$ em S existe $n \in \mathbf{N}$ tal que $a_n = a_{n+1} = a_{n+2} = \ldots$.*

Proposição 5.1.9 *Seja P um conjunto quasi-ordenado tal que não existe uma infinidade de elementos mutuamente incomparáveis em P. Se $a_1, a_2, ...$ é uma seqüência infinita de elementos de P e $\Omega_i = \{j \in \mathbf{N} : a_j \leqslant a_i \text{ ou } a_i \leqslant a_j\}$ para cada i em $\mathbf{N}$, então existe $i_0 \in \mathbf{N}$ tal que Ω_{i_0} é infinito.*

Prova: Suponhamos, por absurdo, que para qualquer i em $\mathbf{N}$, Ω_i é finito. Seja $b_1, b_2, ...$ a subseqüência infinita definida por
$b_1 = a_{m_1}$ onde $m_1 = 1$ e
$b_i = a_{m_i}$ onde $m_i = max(\Omega_{m_{i-1}}) + 1$, para $i \geqslant 2$.

Logo, os elementos de (b_i) constituem uma infinidade de elementos mutuamente incomparáveis em P, o que é uma contradição. ■

Proposição 5.1.10 *Seja P um conjunto quasi-ordenado tal que não existe uma infinidade de elementos mutuamente incomparáveis em P. Então toda seqüência infinita de elementos de P tem uma subseqüência infinita com todos os elementos comparáveis.*

Prova: Sejam $a_1, a_2, ...$ uma seqüência infinita de elementos de P e $\Omega_i = \{j \in \mathbf{N} : a_j \leqslant a_i$ ou $a_i \leqslant a_j\}$ para cada i em $\mathbf{N}$. Seja $b_1, b_2, ...$ a subseqüência definida por
$b_1 = a_{m_1}$ onde $m_1 = \min(\{j \in \mathbf{N} : \Omega_j \text{ é infinito}\})$ e
$b_i = a_{m_i}$ onde $m_i = \min(\{j \in \Omega_{m_{i-1}} : \Omega_j \text{ é infinito}\})$, para $i \geqslant 2$.

Pela **Proposição 5.1.9**, esta subseqüência é infinita. Além disso, todos os seus elementos são comparáveis. ■

O próximo teorema caracteriza bem os conjuntos parcialmente bem ordenados, apresentando uma várias das suas propriedades.

Teorema 5.1.1 *As seguintes afirmações são equivalentes num conjunto quasi-ordenado P:*
(i) P tem a propriedade da base finita;
(ii) a condição da cadeia ascendente é mantida para subconjuntos fechados de P;
(iii) se S é um subconjunto qualquer de P, existe um conjunto finito S_0 tal que $S_0 \subseteq S \subseteq cl(S_0)$;
(iv) toda seqüência infinita de elementos de P tem uma subseqüência infinita ascendente;
(v) se $a_1, a_2, ...$ é uma seqüência infinita de elementos de P, existem inteiros positivos i e j tais que $i < j$ e $a_i \leqslant a_j$;
(vi) não existe uma seqüência infinita estritamente descendente em P, nem uma infinidade de elementos mutuamente incomparáveis em P.

Prova: A prova deste teorema será feita em três etapas. Primeiro, provaremos a equivalência das três primeiras afirmações; depois, provaremos a equivalência das três últimas afirmações; por fim, provaremos que (iii) implica (v) e que (v) implica (ii).

Se P tem a propriedade da base finita, suponhamos, por absurdo, que exista uma seqüência infinita $F_1 \subseteq F_2 \subseteq ...$ de subconjuntos fechados de P com a propriedade de que para todo inteiro positivo i existe um inteiro positivo j tal que F_i está contido propriamente em F_j. Pela **Proposição 5.1.6**, $F = \cup_{i=1}^{\infty} Fi$ é fechado. Sejam G um conjunto finito tal que $F = cl(G)$ e $n \in \mathbf{N}$ tal que $G \subseteq F_n$. Como F_n é fechado, temos que $F = cl(G) \subseteq F_n$. Logo, $F = F_n = F_{n+1} = ...$, o que é uma contradição. Portanto, (i) implica (ii).

Se a condição da cadeia ascendente é mantida para subconjuntos fechados de P, vamos supor, por absurdo, que exista um subconjunto S de P tal que para qualquer subconjunto finito S_0 de S temos que $S \nsubseteq cl(S_0)$. Sejam $a_1 \in S$ e $A_1 = \{a_1\}$. Logo, existe $a_2 \in S$ tal que $a_2 \notin cl(A_1)$; seja $A_2 = \{a_1, a_2\}$. Logo, existe $a_3 \in S$ tal que $a_3 \notin cl(A_2)$; seja $A_3 = \{a_1, a_2, a_3\}$. Continuando este processo sucessivamente, teremos uma cadeia de conjuntos fechados $cl(A_1) \subseteq cl(A_2) \subseteq ... \subseteq cl(An) \subseteq ...$ tal que a condição da cadeia ascendente não se mantém, o que é uma contradição. Portanto, (ii) implica (iii).

Se a afirmação (iii) é verdadeira e F é um subconjunto fechado de P, então existe um conjunto finito F_0 tal que $F_0 \subseteq F \subseteq cl(F_0)$. Como $cl(F_0) \subseteq F$ (**Proposição 5.1.5**), temos que $F = cl(F_0)$ e F é fecho de um subconjunto finito. Portanto, (iii) implica (i) e fica demonstrada a equivalência das três primeiras afirmações.

Se a afirmação (iv) é verdadeira e $a_1, a_2, ...$ é uma seqüência infinita de elementos de P, tomemos uma subseqüência infinita ascendente $a_i, a_j, ...$ de $a_1, a_2, ...$ Logo, $i < j$ e $a_i \leqslant a_j$. Portanto, (iv) implica (v).

Consideremos que a afirmação (v) é verdadeira. Suponhamos, por absurdo, que exista uma seqüência infinita estritamente descendente $a_1, a_2, ...$ em P. Logo, existem inteiros positivos i e j tais que $i < j$ e $a_i \leqslant a_j$; como $a_j \leqslant a_{i+1}$, temos que $a_i \leqslant a_{i+1}$, o que é uma contradição. Se A é um subconjunto infinito de P, então tomando uma seqüência infinita $a_1, a_2, ...$ de elementos distintos em A temos que existem inteiros positivos i e j tais que $i < j$ e $a_i \leqslant a_j$. Portanto, (v) implica (vi).

Consideremos que a afirmação (vi) é verdadeira. Seja $a_1, a_2, ...$ uma seqüência infinita de elementos de A. Pela **Proposição 5.1.10**, existe uma subseqüência $b_1, b_2, ...$ de (a_i) tal que $b_i \leqslant b_j$ ou $b_j \leqslant b_i$ para quaisquer i e j em $\mathbf{N}$. Para cada b_i, seja $\Omega_i = \{j > i : b_i \leqslant b_j\}$. Seja $A = \{i \in \mathbf{N} : \Omega_i \text{ é finito}\}$ e consideremos os dois possíveis casos:

Primeiro caso: A é finito.

Seja m_1 um índice suficientemente grande tal que $i \notin A$ se $i \geqslant m_1$. Logo, Ω_i é infinito se $i \geqslant m_1$. Seja $c_1, c_2, ...$ a subseqüência infinita definida por

$c_1 = b_{m_1}$ e

$c_i = b_{m_i}$ onde $m_i = min(\Omega_{m_{i-1}})$, para $i \geqslant 2$.

Como $c_i \leqslant c_{i+1}$ para todo $i \geqslant 1$, temos que (c_i) é uma subseqüência infinita ascendente de (a_i).

Segundo caso: A é infinito.

Seja $c_1, c_2, ...$ a subseqüência infinita definida por

$c_1 = b_{m_1}$ onde $m_1 = min(A)$ e

$c_i = b_{m_i}$ onde $m_i = min(\{j > m_{i-1} : j \in A - \Omega_{m_{i-1}}\})$, para $i \geqslant 2$.

Como $c_{i+1} \geqslant c_i$ para todo $i \geqslant 1$, temos que (c_i) é uma subseqüência infinita estritamente descendente de (a_i), o que é uma contradição.

Portanto (vi) implica (iv) e fica demonstrada a equivalência das três últimas afirmações.

Consideremos que a afirmação (iii) é verdadeira. Sejam $a_1, a_2, ...$ uma seqüência infinita de elementos de P e B o conjunto dos elementos desta seqüência. Logo, existe um conjunto finito B_0 tal que $B_0 \subseteq B \subseteq cl(B_0)$. Como B_0 é finito, existe um inteiro j tal que todo elemento de B_0 é um a_i com $i < j$; e como $B \subseteq cl(B_0)$, para algum a_i em B_0 temos que $a_i \leqslant a_j$. Portanto (iii) implica (v).

Consideremos que a afirmação (v) é verdadeira. Afirmamos que não pode existir uma seqüência infinita propriamente ascendente $F_1, F_2, ...$ de conjuntos fechados em P. Pois

supondo o contrário, se construirmos a seqüência (a_i) onde $a_i \in F_{i+1} - F_i$, teremos um absurdo. Portanto, (v) implica (ii) e a prova está concluída. ∎

Definição 5.1.8 *Seja $\varphi : A \to B$ uma aplicação entre dois conjuntos quasi-ordenados A e B tais que a relação de quasi-ordem de A é denotada por $\leqslant$ e a relação de quasi-ordem de B é denotada por $\preccurlyeq$. Se $a \leqslant b$ implicar $\varphi(a) \preccurlyeq \varphi(b)$ para quaisquer a e b em A, dizemos que φ é um* homomorfismo.

Proposição 5.1.11 *Se $\varphi : A \to B$ é um homomorfismo entre dois conjuntos quasi-ordenados A e B e F é um subconjunto fechado de B, então $\varphi^{-1}(F)$ é um subconjunto fechado de A.*

Prova: Denotemos as relações de quasi-ordem de A e B por $\leqslant$ e $\preccurlyeq$ respectivamente. Se $a \in cl(\varphi^{-1}(F))$ então existe $b \in \varphi^{-1}(F)$ tal que $b \leqslant a$. Como $\varphi(b) \in F$ e $\varphi(b) \preccurlyeq \varphi(a)$, temos que $\varphi(a) \in F$. Logo, $a \in \varphi^{-1}(F)$ e portanto $\varphi^{-1}(F)$ é um subconjunto fechado de A. ∎

Teorema 5.1.2 *Todo subconjunto de um conjunto parcialmente bem ordenado também é parcialmente bem ordenado. Toda imagem homomórfica de um conjunto parcialmente bem ordenado é também parcialmente bem ordenada.*

Prova: Seja S um subconjunto de um conjunto parcialmente bem ordenado P. Pela **Proposição 5.1.1**, S é quasi-ordenado. Seja $a_1, a_2, ...$ uma seqüência infinita de elementos de S. Como $a_1, a_2, ...$ também é uma seqüência infinita de elementos de P, pelo **Teorema 5.1.1**, temos que existem inteiros positivos i e j tais que $i < j$ e $a_i \leqslant a_j$. Novamente pelo **Teorema 5.1.1**, temos que S tem a propriedade da base finita. Portanto, S é um conjunto parcialmente bem ordenado.

Sejam A um conjunto parcialmente bem ordenado, B um conjunto quasi-ordenado, $\varphi : A \to B$ um homomorfismo e $F_1 \subseteq F_2 \subseteq ...$ uma seqüência infinita de conjuntos fechados de $\varphi(A)$. Pela **Proposição 5.1.11**, temos que $\varphi^{-1}(F_1), \varphi^{-1}(F_2), ...$ é uma seqüência infinita de conjuntos fechados de A e, pelo **Teorema 5.1.1**, temos que existe $n \in \mathbf{N}$ tal que $\varphi^{-1}(F_n) = \varphi^{-1}(F_{n+1}) = ...$ Como $\varphi(\varphi^{-1}(S)) = S$ para qualquer subconjunto S de B, temos que $F_n = F_{n+1} = ...$ Assim, novamente pelo **Teorema 5.1.1**, $\varphi(A)$ tem a propriedade da base finita. Portanto, $\varphi(A)$ é um conjunto parcialmente bem ordenado. ∎

Definição 5.1.9 *O produto cartesiano de dois conjuntos quasi-ordenados P e Q é o conjunto de pares ordenados (a, b) com $a \in A$ e $b \in B$, munido da relação $\leqslant$ onde $(a_1, b_1) \leqslant (a_2, b_2)$ se, e somente se, $a_1 \leqslant a_2$ e $b_1 \leqslant b_2$.*

Proposição 5.1.12 *O produto cartesiano de dois conjuntos quasi-ordenados também é um conjunto quasi-ordenado.*

Teorema 5.1.3 *O produto cardinal de dois conjuntos parcialmente bem ordenados também é um conjunto parcialmente bem ordenado.*

Prova: Sejam P e Q dois conjuntos parciamente bem ordenados e (a_i, b_i), $i \in \mathbf{N}$, uma seqüência infinita do produto cardinal de P e Q. Pelo **Teorema 5.1.1**, temos que a seqüencia (a_i) possui uma subseqüência infinita ascendente; chamemo-la (α_j). Tomemos a subseqüência infinita (α_j, b_j) de (a_i, b_i). Novamente, pelo **Teorema 5.1.1**, temos que a seqüência (b_j)

possui uma subseqüência infinita ascendente; chamemo-la (β_k). Assim, (α_k, β_k) é uma subseqüência infinita ascendente de (a_i, b_i) e, pelo **Teorema 5.1.1**, o produto cardinal de P e Q é um conjunto parcialmente bem ordenado. ■

Teorema 5.1.4 *Seja $P(M)$ uma proposição com relação a um conjunto quasi-ordenado M. A proposição $P(M)$ é verdadeira se $P(A)$ é verdadeira para qualquer subconjunto aberto próprio A de M. Se M é um conjunto parcialmente bem ordenado então $P(M)$ é verdadeira.*

Prova: Suponhamos, por absurdo, que $P(M)$ seja falsa. Logo, $P(A_1)$ é falsa para algum subconjunto aberto próprio de M, bem como $P(A_2)$ é falsa para algum subconjunto aberto próprio de A_1, e assim por diante. Os complementos em M de $A_1, A_2, ...$ formam uma cadeia infinita propriamente ascendente de subconjuntos fechados de M. Pelo **Teorema 5.1.1**, M não tem a propriedade da base finita, o que é uma contradição. ■

Teorema 5.1.5 *Se M é um conjunto quasi-ordenado tal que para todo a em M, $M - cl(a)$ é um conjunto parcialmente bem ordenado, então M é um conjunto parcialmente bem ordenado.*

Prova: Sejam F um subconjunto fechado não-vazio de M e a um elemento de F. Pela **Proposição 5.1.7**, $F - cl(a)$ é um subconjunto fechado de $M - cl(a)$. Logo, $F - cl(a)$ é o fecho em $M - cl(a)$ de um conjunto finito F_0. Pela **Proposição 5.1.8**, o fecho de $F_0 \cup \{a\}$ em M é F. Portanto, M é um conjunto parcialmente bem ordenado. ■

Definição 5.1.10 *Se $(M_i) = (M_0, M_1, ..., M_n)$ é uma seqüência de conjuntos parcialmente bem ordenados, então uma segunda seqüência (M_i') é dita obtida da primeira* por descendência *se para algum inteiro r, $0 \leqslant r \leqslant n$, temos que*
(i) para $i > r$, $M_i' = M_i$;
(ii) M_r' é um subconjunto aberto próprio de M_r;
(iii) para $i < r$, M_i' é qualquer conjunto parcialmente bem ordenado.

Lema 5.1.1 *Se $(M_i) = (M_0, M_1, ..., M_n)$ é uma seqüência de conjuntos parcialmente bem ordenados, então não existe nenhuma cadeia infinita de seqüências $(M_i^{(r)})$, $r = 1, 2, ...$, tal que $(M_i^{(1)}) = (M_i)$ e $(M_i^{(r+1)})$ é obtida de $(M_i^{(r)})$ por descendência para todo r.*

Prova: Provemos por indução em n.

Para $n = 0$, suponhamos, por absurdo, que exista uma cadeia infinita de seqüências $(M_0^{(r)})$, $r = 1, 2, ...$, tal que para todo r, $(M_0^{(r+1)})$ é obtida de $(M_0^{(r)})$ por descendência, ou seja, $M_0^{(r+1)}$ é um subconjunto aberto próprio de $M_0^{(r)}$. Os complementos em $M_0^{(1)}$ de $(M_0^{(r)})$ formam uma cadeia propriamente ascendente infinita de conjuntos fechados em $M_0^{(1)}$, o que é uma contradição, pois $M_0^{(1)}$ é um conjunto parcialmente bem ordenado.

Seja $n > 0$. Suponhamos, por absurdo, que exista uma cadeia infinita de seqüências $(M_0^{(r)}, M_1^{(r)}, ..., M_n^{(r)})$, $r = 1, 2, ...$, tal que para todo r, $(M_i^{(r+1)})$ é obtida de$(M_i^{(r)})$ por descendência. Como para cada r, $M_n^{(r+1)} = M_n^{(r)}$ ou $M_n^{(r+1)}$ é um subconjunto aberto próprio de $M_n^{(r)}$, temos que $M_n^{(k)} = M_n^{(k+1)} = ...$ para um k suficientemente grande. Assim se consideramos a cadeia $(M_0^{(r)}, M_1^{(r)}, ..., M_{n-1}^{(r)})$, $r = k, k+1, ...$, temos que ela é uma cadeia infinita tal que para todo r, $(M_0^{(r+1)}, M_1^{(r+1)}, ..., M_{n-1}^{(r+1)})$ é obtida de $(M_0^{(r)}, M_1^{(r)}, ..., M_{n-1}^{(r)})$ por descendência, o que é uma contradição segundo a hipótese de indução. ■

Teorema 5.1.6 *Sejam* $(M_i) = (M_0, M_1, ..., M_n)$ *uma seqüência de conjuntos parcialmente bem ordenados e* $P(M_i)$ *uma proposição com relação à seqüência* (M_i). *Se a proposição* $P(M_i)$ *é falsa então* $P(M_i')$ *é falsa para alguma seqüência* (M_i') *obtida de* (M_i) *por descendência. Então* $P(M_i')$ *é verdadeira.*

Prova: Suponhamos, por absurdo, que $P(M_i)$ seja falsa. Logo, $P(M_i^{(1)})$ é falsa para alguma seqüência $(M_i^{(1)})$ obtida de (M_i) por descendência, bem como $P(M_i^{(2)})$ é falsa para alguma seqüência $(M_i^{(2)})$ obtida de $(M_i^{(1)})$ por descendência, e assim por diante. Estas seqüências $(M_i^{(r)})$, $r = 1, 2, ...$, formam uma cadeia infinita que para todo r, $(M_i^{(r+1)})$ é obtido de $(M_i^{(r)})$ por descendência, o que é uma contradição com relação ao **Lema 5.1.1**. ■

5.2 Teorema de Higman

Definição 5.2.1 *Uma* álgebra abstrata (A, M) *(ou simplesmente uma* álgebra*) é um par ordenado de conjuntos* A *e* M *tal que cada* μ *em* M *é um operador* r*-ário, para algum inteiro não-negativo* r*, que leva cada seqüência* $a_1, a_2, ..., a_r$ *de* r *elementos de* A *num único elemento* $\mu(a_1, a_2, ..., a_r)$ *de* A. *O conjunto* A *é chamado de* conjunto de elementos *e o conjunto* M *de* conjunto de operações.

Notação 5.2.1 *Numa álgebra* (A, M), *denotaremos:*
(i) M_r *como o subconjunto de* M *que contém todas as operações* r*-árias;*
(ii) $\mu a_1 a_2 ... a_r = \mu(a_1, a_2, ..., a_r)$ *para* $\mu \in M_r$ *e* $a_i \in A$, $1 \leqslant i \leqslant r$;
(iii) as primeiras letras do alfabeto, $a, b, ...$, *como elementos de* A;
(iv) as últimas letras do alfabeto, $x, y, ...$, *como seqüências finitas (possivelmente vazias) de elementos de* A;
(v) os comprimentos das seqüências de elementos de A *sempre fazendo sentido; por exemplo, se* $\mu \in M_r$, *na expressão* μxay *os comprimentos* s *e* t *de* x *e* y *satisfazem* $s + t + 1 = r$.

Definição 5.2.2 *Uma* subálgebra (B, M) *de uma álgebra* (A, M) *é uma álgebra em que* B *é um subconjunto de* A.

Observação 5.2.1 *Numa subálgebra abstrata* (B, M) *de uma álgebra abstrata* (A, M), $\mu x \in B$ *para qualquer operador* r*-ário* μ *de* M *e qualquer seqüência* x *de* r *elementos de* B.

Definição 5.2.3 *Uma álgebra abstrata é chamada* minimal *se não possui nenhuma subálgebra distinta dela própria.*

Proposição 5.2.1 *Se* (A, M) *é uma álgebra e* F *uma família de subálgebras de* (A, M), *então* (B, M) *é uma subálgebra de* (A, M), *onde* $B = \cap_{(A', M) \in F} A'$.

Prova: Sejam μ um operador r-ário de M e $a_1, ..., a_r$ uma seqüência de r elementos de B. Como os elementos $a_1, ..., a_r$ pertencem a cada A' em F, temos que $\mu(a_1, ..., a_r)$ pertence a cada A'. Logo, $\mu(a_1, ..., a_r) \in B$. ■

Definição 5.2.4 *Sejam (A, M) uma álgebra, S um subconjunto de A e $B = \cap_{(A',M)\in F} A'$ onde $F = \{(A', M)$ subálgebra de $(A, M) : S \subseteq A'\}$. A subálgebra (B, M) é chamada* álgebra gerada *por S, e o conjunto S é chamado* conjunto gerador *da subálgebra (B, M).*

Proposição 5.2.2 *Sejam (A, M) uma álgebra e $A_0 = \{a \in A : \mu x = a$ para algum $\mu \in M_0$ e x a seqüência vazia$\}$. A álgebra (A, M) é minimal se, e somente se, A_0 é um conjunto gerador da álgebra $(A, M - M_0)$.*

Prova: Consideremos que a álgebra (A, M) seja minimal. Suponhamos, por absurdo, que A_0 não é um conjunto gerador da álgebra $(A, M - M_0)$. Logo, existe $a \in A$ tal que $a \notin A_0$. Assim, a álgebra gerada por A_0, com relação a M, é uma subálgebra de (A, M). Portanto, (A, M) não é minimal, o que é uma contradição.

Consideremos agora que A_0 é um conjunto gerador da álgebra $(A, M - M_0)$. Seja (B, M) uma subálgebra de (A, M). Provemos que $(A, M) = (B, M)$, ou seja, $A = B$. Claramente, $A_0 \subseteq B$. Como A_0 é um conjunto gerador da álgebra $(A, M - M_0)$, temos que $(A, M) = (B, M)$. Portanto, (A, M) é minimal. ■

Proposição 5.2.3 *Toda álgebra (A, M) tem uma única subálgebra minimal.*

Prova: Sejam $A_0 = \{a \in A : \mu x = a$ para algum $\mu \in M_0$ e x a seqüência vazia$\}$ e (B, M) a subálgebra gerada por A_0. Como (B, M) não contém nenhuma subálgebra distinta dela própria e está contida em toda subálgebra de (A, M), ela é a única subálgebra minimal de (A, M). ■

Definição 5.2.5 *Uma álgebra (A, M), tal que A é um conjunto quasi-ordenado, é uma* álgebra ordenada *se satisfaz a seguinte propriedade para todo μ em M e para quaisquer a, b, x, y:*
(i) se $a \leqslant b$ então $\mu xay \leqslant \mu xby$.

Definição 5.2.6 *Numa álgebra ordenada (A, M), a relação de quasi-ordem é chamada* ordem de divisibilidade *se satisfaz a seguinte propriedade para todo μ em M e para quaisquer a, x, y:*
(ii) $a \leqslant \mu xay$.
Uma álgebra ordenada na qual a relação de quasi-ordem é uma ordem de divisibilidade é chamada uma álgebra com ordem de divisibilidade.

Definição 5.2.7 *Numa álgebra com ordem de divisibilidade (A, M), tal que cada M_r é um conjunto quasi-ordenado, dizemos que a relação de quasi-ordem em A é* compatível *com as relações de quasi-ordem de cada M_r se satisfaz a seguinte propriedade para quaisquer μ e λ em M_r e para qualquer x:*
(iii) se $\lambda \leqslant \mu$ então $\lambda x \leqslant \mu x$.

Lema 5.2.1 *Seja (A, M) uma álgebra com ordem de divisibilidade. Se X é um subconjunto fechado de A e (B, M) é uma subálgebra de (A, M) então $(B \cup X, M)$ é uma subálgebra de (A, M).*

Prova: Sejam μ um operador r-ário em M, $a_1, ..., a_r$ elementos de $B \cup X$ e $b = \mu a_1 ... a_r$. Se $a_1, ..., a_r$ pertencem todos a B, então b pertence a B, pois (B, M) é uma subálgebra. Caso contrário, existe um a_i que pertence a X e, neste caso, b pertence a X, desde que $a_i \leqslant b$ em qualquer ordem de divisibilidade e X é fechado. Portanto, b pertence a $B \cup X$. ■

Teorema 5.2.1 (Higman) *Suponha que (A, M) seja uma álgebra abstrata minimal, e que M_r, o conjunto de operações r-árias em M, seja um conjunto parcialmente bem ordenado para $r = 1, 2, ..., n$, e seja vazio para $r > n$. Se (A, M) tem uma ordem de divisibilidade compatível com as quasi-ordens de cada M_r então A tem a propriedade da base finita.*

Prova: Suponhamos, por absurdo, que A não tem a propriedade da base finita. Vamos deduzir a existência de uma álgebra abstrata (A', M') com as mesmas propriedades de (A, M) e que a seqüência (M_i') é obtida de (M_i) por descendência. Assim teremos uma contradição com relação ao Lema 1.

Seja $X = \{a \in A : A - cl(a)$ tem a propriedade da base finita$\}$. Pelo **Teorema 5.1.5**, $X \neq A$ e X não é fechado sob M pois (A, M) é minimal. Assim, existe um μ em M_r para algum r em $0 \leqslant r \leqslant n$, e uma seqüência $z = a_1 ... a_r$ de elementos de X tais que $\mu z = b \notin X$. Vamos definir M' como
$M_i' = M_i$ se $i \neq r-1$ e $i \neq r$;
$M_r' = M_r - cl(\mu)$;
M_{r-1}' é um caso um pouco mais complexo. No entanto, basta considerar $r \geqslant 1$. Se λ é uma operação r-ária, a um elemento de A e $s \in \{1, ..., r\}$, denotamos por $\lambda_{(a,s)}$ a operação $(r-1)$-ária definida como $\lambda_{(a,s)} xy = \lambda xay$, onde x tem comprimento $s-1$. Seja $M_{r-1}^{(s)}$ o conjunto das operações $\lambda_{(a,s)}$ com $\lambda \in cl(\mu)$ e $a \in A - cl(a_s)$, ordenado como o produto cartesiano de $cl(\mu)$ e $A - cl(a_s)$. Assim, definimos
$M_{r-1}' = M_{r-1} \cup (\cup_{s=1}^{r} M_{r-1}^{(s)})$,
sem relações de ordem entre termos diferentes desta união.

Seja (A', M') a única subálgebra minimal de (A, M), conforme a **Proposição 5.2.3**.

Provemos que (A', M') tem as mesmas propriedades de (A, M).

Se considerarmos A' um conjunto quasi-ordenado como um subconjunto de A, não é difícil ver que (A', M') é uma álgebra com ordem de divisibilidade e que a relação de quasi-ordem de A é compatível com as relações de quasi-ordem de (M_i').

Se $i \neq r-1$, não é difícil ver que M_i' tem a propriedade da base finita. Provemos que M_{r-1}' tem a propriedade da base finita. Como M_{r-1} tem a propriedade da base finita por hipótese e cada $M_{r-1}^{(s)}$ tem a propriedade da base finita pelo **Teorema 5.1.3**, obviamente a união finita de conjuntos com a propriedade da base finita também tem esta propriedade.

Provemos agora que A' não tem a propriedade da base finita. Podemos provar isto mostrando que A' contém $A - cl(b)$, que por hipótese não tem a propriedade da base finita. Nosso método será provar que $A' \cup cl(b) = A$, mostrando que $(A' \cup cl(b), M')$ é uma subálgebra de (A, M).

Pelo **Lema 5.2.1**, $A' \cup cl(b)$ é fechado sob as operações de M', e então temos apenas que provar que é fechado sob as operações de M que não estão em M', as quais são as operações de $cl(\mu)$.

Suponha então que os elementos $c_1, ..., c_r$ pertencem a $A' \cup cl(b)$ e que $d = \lambda c_1 ... c_r$ para um λ em $cl(\mu)$. Se $a_i \leqslant c_i$ para todo i, então $b = \mu a_1 ... a_r \leqslant \lambda c_1 ... c_r = d$ e $d \in cl(b)$. Mas se

(o que ocorre apenas se $r \geqslant 1$) para algum s, $c_s \in A - cl(a_s)$, então $d = \lambda_{(c_s,s)}x$, onde x é uma seqüência de elementos de $A' \cup cl(b)$. Como e $A' \cup cl(b)$ é fechado sob M', novamente temos que $d \in A' \cup cl(b)$. Logo $A' \cup cl(b)$ é fechado sob M' e $(A' \cup cl(b), M)$ é uma subálgebra de (A, M).

Portanto, (A', M') tem as mesmas propriedades de (A, M), e a prova está concluída. ■

Em 1960, J. B. Kruskal [20] melhorou o Teorema de Higman provando que a condição $M_r = \emptyset$ se $r > n$ pode ser substituída por $M_n = M_{n+1} = ...$

Corolário 5.2.1 *Seja (A, M) uma álgebra com um número finito de operadores. Se (A, M) tem uma ordem de divisibilidade e um conjunto gerador com a propriedade da base finita, então A tem a propriedade da base finita.*

Prova: Sejam B um conjunto gerador de (A, M) com a propriedade da base finita e $N = \{\mu_a$ um operador 0-ário: $\mu_a x = a$ onde x é a seqüência vazia e $a \in B\}$. Seja $M' = M \cup N$.

Pela **Proposição 5.2.2**, a álgebra (A, M') é minimal.

Além disso, como cada M_r é finito, para $r \geqslant 1$, se considerarmos a ordem trivial ($\mu \leqslant \lambda$ somente se $\mu = \lambda$), cada M_r é um conjunto parcialmente bem ordenado. Como $M'_0 = M_0 \cup N$, M_0 é finito e N é um conjunto parcialmente bem ordenado, temos que M'_0 é um conjunto parcialmente bem ordenado. Como M é finito, temos que existe um inteiro não-negativo n tal que se $r > n$ então M_r é vazio.

Não é difícil ver que a quasi-ordem de A é uma ordem de divisibilidade em (A, M') compatível com as quasi-ordens de cada M'_r.

Portanto, A tem a propriedade da base finita. ■

5.3 Algumas aplicações combinatórias

Mostraremos a seguir algumas aplicações do Teorema de Higman em diversas áreas da Combinatória.

Teorema 5.3.1 *Se X é um conjunto de inteiros positivos com a propriedade de que em qualquer subconjunto infinito de X existem elementos distintos a e b tais que a divide b, então a mesma propriedade é válida para $P(X)$, o conjunto de inteiros que podem ser escritos como produtos de elementos de X.*

Prova: Os conjuntos X e $P(X)$ são quasi-ordenados se considerarmos $a \leqslant b$ como a divide b. Pelo item (vi) do **Teorema 5.1.1**, temos que X é um conjunto parcialmente bem ordenado. Consideremos o operador binário μ que associa cada par de elementos de $P(X)$ ao seu produto. Claramente $(P(X), \{\mu\})$ é uma álgebra com ordem de divisibilidade e X é um conjunto gerador para esta álgebra. Pelo **Corolário 5.2.1**, $P(X)$ tem a propriedade da base finita. Novamente pelo item (vi) do **Teorema 5.1.1**, em qualquer subconjunto infinito de $P(X)$ existem elementos distintos a e b tais que a divide b. ■

Definição 5.3.1 *Se A é um conjunto, definimos*

$$S(A) = \{(a_1, ..., a_n) : a_i \in A \text{ para } i = 1, ..., n, \text{ e } n \text{ um inteiro não-negativo}\}.$$

Definição 5.3.2 *Seja A um conjunto quasi-ordenado. Para $x = (a_1, ..., a_r)$ e $y = (b_1, ..., b_s)$ em $S(A)$, definimos uma relação de quasi-ordem em $S(A)$ escrevendo que $x \leqslant y$ se existe uma aplicação $f : \{1, ..., r\} \longrightarrow \{1, ..., s\}$ tal que*
(i) se $i < j$ então $f(i) < f(j)$, para quaisquer i e j em $\{1, ..., r\}$;
(ii) $a_i \leqslant b_{f(i)}$ para todo i em $\{1, ..., r\}$.
A relação de ordem $\leqslant$ em $S(A)$ é chamada ordem natural induzida pela ordem de A.

Exemplo 5.3.1 *Se $A = \mathbf{Z}$ com o ordem usual, temos que $(1, 2, 4) \leqslant (5, 2, 3, 1, 7)$, utilizando a função $f(1) = 2, f(2) = 3$ e $f(3) = 5$. Por outro lado, $(1, 2, 4) \not\leqslant (6, 5, 1, 2, 3, 3, 3, 1)$.*

Teorema 5.3.2 *Se A é um conjunto parcialmente bem ordenado, então $S(A)$ também é um conjunto parcialmente bem ordenado.*

Prova: Consideremos o operador binário μ que associa cada par de elementos de $S(A)$ à sua concatenação (justaposição). Claramente $(S(A), \{\mu\})$ é uma álgebra com ordem de divisibilidade e A é um conjunto gerador para esta álgebra. Pelo **Corolário 5.2.1**, $S(A)$ é um conjunto parcialmente bem ordenado.. ■

Corolário 5.3.1 *Se X é um conjunto qualquer de palavras sobre um alfabeto finito, é possível encontrar um subconjunto finito X_0 de X tal que, dada uma palavra w em X, é possível encontrar w_0 em X_0 tal que as letras de w_0 ocorrem em w na mesma ordem, embora não necessariamente consecutivamente.*

Definição 5.3.3 *Se A é um conjunto, definimos $P(A)$ como o conjunto dos subconjuntos finitos de A.*

Definição 5.3.4 *Seja A um conjunto quasi-ordenado. Para R e S em $P(A)$, definimos uma relação de quasi-ordem em $P(A)$ escrevendo que $R \leqslant S$ se existe uma aplicação $f : R \longrightarrow S$ tal que*
(i) f é injetora;
(ii) $a \leqslant f(a)$ para todo a em R.

Exemplo 5.3.2 *Se $A = \mathbf{Z}$ com o ordem usual, temos que $\{1, 2, 4\} \leqslant \{2, 3, 5, 7\}$, utilizando a função $f(1) = 2, f(2) = 3$ e $f(3) = 5$. Por outro lado, $\{1, 2, 5\} \not\leqslant \{1, 2, 3, 4\}$.*

Teorema 5.3.3 *Se A é um conjunto parcialmente bem ordenado, então $P(A)$ também é um conjunto parcialmente bem ordenado.*

Prova: Consideremos o operador binário μ que associa cada par de elementos de $P(A)$ à sua união. Claramente $(P(A), \{\mu\})$ é uma álgebra com ordem de divisibilidade e A é um conjunto gerador para esta álgebra. Pelo **Corolário 5.2.1**, $P(A)$ é um conjunto parcialmente bem ordenado. ■

Teorema 5.3.4 *Se X é um conjunto qualquer de palavras formadas a partir de um alfabeto finito, é possível encontrar um subconjunto finito X_0 de X tal que, dada uma palavra w em X, é possível encontrar w_0 em X_0 tal que as letras de w_0 ocorrem em w na mesma ordem, embora não necessariamente consecutivamente.*

Corolário 5.3.2 *Se $w_1, w_2, ...$ é uma seqüência infinita de palavras formadas a partir de um alfabeto finito, então existem w_i e w_j, $i < j$ tais que w_i é obtida a partir de w_j retirando-se algumas das letras de w_j.*

5.4 Um teorema sobre permutações

Definição 5.4.1 *A seqüência de números naturais* $(i_1, ..., i_k)$ *tem* altura t *se pode ser particionada em* t *seqüências ascendentes consecutivas (com* t *tomado o menor possível).*

Exemplo 5.4.1 *A seqüência* $(1, 4, 7, 2, 5, 9, 3, 6, 8, 11, 10)$ *pode ser particionada em* $(1, 4, 7)$, $(2, 5, 9)$, $(3, 6, 8, 11)$, (10), *e assim tem altura* 4. *A seqüência* $(1, 2, 5, 3)$ *pode ser particionada em* $(1, 2, 5)$, (3), *e assim tem altura* 2.

Definição 5.4.2 *Se* s *é uma seqüência de altura* t *cujos elementos formam o conjunto* $\{1, ..., k\}$, *então* s *pode ser escrita como uma matriz* $t \times k$ *da seguinte maneira. Particione* s *em* t *subseqüências ascendentes consecutivas* $s_1, ..., s_t$; *se* j *aparece em* s_i, *escreva* 1 *na posição* ij; *caso contrário escreva* 0 *na posição* ij. *Esta matriz é chamada* matriz representação *de* s.

Exemplo 5.4.2 *A matriz representação de* $(1, 4, 2, 3, 5, 6)$ *é* $\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$.

Observação 5.4.1 *Claramente a matriz representação está bem definida. Além disso, se* σ *e* τ *são permutações distintas de* S_n *e as seqüências* $(\sigma(1), ..., \sigma(n))$ *e* $(\tau(1), ..., \tau(n))$ *têm altura* t, *então obviamente suas matrizes representação são distintas.*

Definição 5.4.3 *Seja* $P_{nt} = \{\sigma \in S_n : a\ seqüência\ (\sigma(1), ..., \sigma(n))\ tem\ altura \leqslant t\}$, *e seja* $P_t = \cup_{i=1}^{\infty} P_{it}$.

O próximo teorema é importante, pois será usado no Capítulo 6.

Teorema 5.4.1 *Se* $(\sigma_1, \sigma_2, ...)$ *é uma seqüência infinita de membros de* P_t, *então existem* $\sigma_i \in S_{n_i}$ *e* $\sigma_j \in S_{n_j}$, *com* $i < j$, *e existe uma aplicação* $\psi : \{1, ..., n_i\} \to \{1, ..., n_j\}$ *com as seguintes propriedades*
(i) se $r < s$ *então* $\psi(r) < \psi(s)$ *para quaisquer* r *e* s *em* $\{1, ..., n_i\}$;
(ii) $\sigma_j(1)\sigma_j(2)...\sigma_j(n_j) = w_0\psi(\sigma_i(1))w_1\psi(\sigma_i(2))w_2...w_{n_i-1}\psi(\sigma_i(n_i))w_{n_i}$ *para palavras* w_u *adequadas.*

Prova: A matriz representação de qualquer elemento de P_{nt} é uma matriz $t \times n$ (se a altura for menor que t, basta completar a matriz com linhas nulas), que pode ser vista como uma seqüência de n colunas, cada coluna tomada de $\{0, 1\}^t$. Considerando a ordem trivial ($a \leqslant b$ somente se $a = b$) em $\{0, 1\}^t$, como $\{0, 1\}^t$ é finito, temos que $\{0, 1\}^t$ é um conjunto parcialmente bem ordenado. Pelo **Teorema 5.3.2**, temos que $S(\{0, 1\}^t)$ é um conjunto parcialmente bem ordenado. Pelo **Teorema 5.1.1**, existem inteiros positivos i e j tais que $i < j$ e a matriz representação de σ_j é obtida da matriz representação de σ_i inserindo colunas apropriadas em lugares apropriados. Assim, se $\sigma_i \in S_{n_i}$ e $\sigma_j \in S_{n_j}$, existe uma aplicação $\psi : \{1, ..., n_i\} \to \{1, ..., n_j\}$ tal que se $r < s$ então $\psi(r) < \psi(s)$ para quaisquer r e s em $\{1, ..., n_i\}$ e $\sigma_j(1)\sigma_j(2)...\sigma_j(n_j) = w_0\psi(\sigma_i(1))w_1\psi(\sigma_i(2))w_2...w_{n_i-1}\psi(\sigma_i(n_i))w_{n_i}$ para palavras w_u adequadas. ■

Exemplo 5.4.3 *Na demonstração do teorema anterior, se*

$$\sigma_i = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 3 & 5 & 6 \end{pmatrix} \text{ e } \sigma_j = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 7 & 2 & 3 & 4 & 6 & 8 \end{pmatrix},$$

então as matrizes representação de σ_i e σ_j são respectivamente

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

A segunda é obtida a partir da primeira inserindo $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ *depois da coluna* 1 *e* $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *depois da coluna* 5. *Para obter ψ devemos ignorar as colunas inseridas (no caso a segunda e a sétima) e observar que ψ envia* 142356 *em* 153468. *Assim temos que* $15723468 = \psi(1)\psi(4)72\psi(2)\psi(3)\psi(5)\psi(6)$.

Capítulo 6

PI álgebras sobre Q

Apresentamos algumas propriedades das PI álgebras sobre **Q**. Embora este capítulo tenha sido baseado em Rowen [32], indicamos os artigos originais dos resultados.

6.1 T-ideais

Definição 6.1.1 *Seja K um corpo. Para cada $n \in \mathbf{N}$, definimos V_n como o subespaço de $K\{\mathbf{X}\}$ gerado por $\{X_{\sigma(1)}X_{\sigma(2)}...X_{\sigma(n)} : \sigma \in S_n\}$ e $V_0 = K$.*

Observação 6.1.1 *V_n é o conjunto dos polinômios multilineares de grau n em $X_1, ..., X_n$ com coeficientes em K, além disso $dim_K(V_n) = n!$.*

Definição 6.1.2 *Seja A uma álgebra e $I \lhd A$. O ideal I é um T-*ideal *de A se $\psi(I) \subseteq I$ para todo endomorfismo de álgebras ψ de A. Escrevemos $I \lhd_T A$.*

Agora mostramos que o ideal de todas as identidades polinomiais de uma álgebra (cf. **Definição 3.3.4**) é um T-ideal de $K\{\mathbf{X}\}$.

Proposição 6.1.1 *Se A é uma PI álgebra com identidade sobre um corpo K, então $T(A) \lhd_T K\{\mathbf{X}\}$.*

Prova: Sejam ψ um endomorfismo de álgebras de $K\{\mathbf{X}\}$ e $f(X_1, ..., X_n) \in T(A)$. Pelo **Teorema 3.2.1**, podemos escrever que

$$\psi(f(X_1, ..., X_n)) = f(\psi(X_1), ..., \psi(X_n)).$$

Como $f(a_1, ..., a_n) = 0$ para quaisquer $a_1, ..., a_n$ em A é fácil ver que $\psi(f(X_1, ..., X_n) \in T(A)$. Portanto, $\psi(T(A)) \subseteq T(A)$, ou seja, $T(A)$ é um T-ideal de $K\{\mathbf{X}\}$. ■

Lema 6.1.1 *Sejam K um corpo e I um T-ideal de $K\{\mathbf{X}\}$. Se $p(X_{i_1}, ..., X_{i_n}) \in I$ então $p(q_1, ..., q_n) \in I$, para quaisquer $q_1, ..., q_n$ em $K\{\mathbf{X}\}$.*

Prova: Pelo **Teorema 3.2.1**, temos que $\psi : K\{\mathbf{X}\} \to K\{\mathbf{X}\}$, dada por

$$\psi(X_i) = \begin{cases} q_k & \text{, se } i = i_k \text{ para algum } k \in \{1, ..., n\} \\ X_i & \text{, caso contrário} \end{cases}$$

é um endomorfismo de $K\{\mathbf{X}\}$. Logo, $\psi(p(X_{i_1}, ..., X_{i_n})) = p(q_1, ..., q_n) \in I$. ■

Lema 6.1.2 *Seja K um corpo. Se $I \lhd_T K\{\mathbf{X}\}$ então $I = T(K\{\mathbf{X}\}/I)$.*

Prova: Sejam $p(X_{i_1}, ..., X_{i_k}) \in I$ e $p_1 + I$,..., $p_k + I$ elementos de $K\{\mathbf{X}\}/I$. Logo

$$p(p_1 + I, ..., p_k + I) = p(p_1, ..., p_k) + I.$$

Pelo **Lema 6.1.1**, $p(p_1, ..., p_k) \in I$. Portanto, $I \subseteq T(K\{\mathbf{X}\}/I)$.
Seja $p(X_{i_1}, ..., X_{i_k}) \in T(K\{\mathbf{X}\}/I)$. Como

$$I = p(X_{i_1} + I, ..., X_{i_k} + I) = p(X_{i_1}, ..., X_{i_k}) + I,$$

temos que $p(X_{i_1}, ..., X_{i_k}) \in I$. Portanto, $T(K\{\mathbf{X}\}/I) \subseteq I$. ■

6.2 Codimensões

Definição 6.2.1 *Se I é um T-ideal de $K\{\mathbf{X}\}$, definimos $I_n = I \cap V_n$ e $c_n(I) = dim_K(V_n/I_n)$; $c_n(I)$ é chamada a n-*codimensão *de I.*

Obviamente I_n é um subespaço de V_n e $c_n(I) \leqslant n!$.

Qual a importância das codimensões? Veremos adiante que a seqüência $\{c_n(I)\}$ fornece informações sobre I quando $c_n(I)$ é bem menor que $n!$. Em 1972, A. Regev [29] provou que se I satisfaz uma identidade polinomial de grau g, então $c_n(I) \leqslant (3 \cdot 4^{g-3})^n$. No mesmo ano, V. N. Latyshev [21] melhorou este resultado provando que $c_n(I) < (g-1)^{2n}$. Desenvolveremos a seguir o método de Latyshev, que se apóia no Teorema de Dilworth.

Definição 6.2.2 *Para $\sigma \in S_n$, definimos $\rho(\sigma)$ como o maior k para o qual existem $i_1 < i_2 < ... < i_k \in \{1, ..., n\}$ tais que $\sigma(i_1) > \sigma(i_2) > ... > \sigma(i_k)$.*

Exemplo 6.2.1 *Se $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 4 & 1 \end{pmatrix}$, então $\rho(\sigma) = 4$, porque $1 < 2 < 3 < 5$ e $\sigma(1) > \sigma(2) > \sigma(3) > \sigma(5)$.*

Definição 6.2.3 *Para $\sigma \in S_n$, definimos as* tabelas de Amitsur *$T_1(\sigma) = (t_{ij})$ e $T_2(\sigma) = (u_{ij})$ da seguinte maneira: $t_{11} = 1$ e $u_{11} = \sigma(1)$. Indutivamente, t_{1j} (se existir) é o menor k tal que $t_{1,j-1} < k \leqslant n$ e $\sigma(k) > u_{1,j-1}$; $u_{1j} = \sigma(k)$. Continuamos ao longo da primeira linha até onde for possível e então começamos a segunda linha. Construímos a segunda linha escolhendo o menor k que não aparece na primeira linha e definindo $t_{21} = k$ e $u_{21} = \sigma(k)$. Continuamos como na primeira linha mas ignorando todo k que já aparece em T_1. Assim, quando terminadas, T_1 e T_2 têm cada uma n entradas.*

Exemplo 6.2.2 *Se* $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 5 & 3 & 2 & 6 \end{pmatrix}$, *então* $\rho(\sigma) = 3$ *e*

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

Teorema 6.2.1 (Dilworth) $\rho(\sigma)$ *é o número de linhas em* $T_1(\sigma)$.

Prova: Seja g o número de linhas em $T_1(\sigma)$. Se $i_1 < ... < i_k$ e $\sigma(i_1) > ... > \sigma(i_k)$, então $i_1, ..., i_k$ estão em linhas distintas, implicando $\rho(\sigma) \leq g$. Por outro lado, podemos construir uma seqüência $i_g, ..., i_1$ da seguinte maneira: $i_g = t_{g1}$ e, indutivamente, dado i_{m+1} da linha $m+1$, definimos $i_m = t_{mj}$ como j maximal tal que $t_{mj} < i_{m+1}$. Observe que $\sigma(i_m) = u_{mj} > \sigma(i_{m+1})$, pois caso contrário ao construirmos as tabelas teríamos colocado i_{m+1} na linha m de $T_1(\sigma)$. Assim, $i_1 < ... < i_g$ e $\sigma(i_1) > ... > \sigma(i_g)$, implicando $\rho(\sigma) \geqslant g$. Portanto, $\rho(\sigma) = g$. ■

Teorema 6.2.2 (Latyshev) *Se* I *é um* T*-ideal de* $K\{\mathbf{X}\}$ *tal que* $K\{\mathbf{X}\}/I$ *satisfaz uma identidade polinomial* p *de grau* g, *então* $c_n(I) \leqslant |\{\sigma \in S_n : \rho(\sigma) < g\}|$.

Prova: Sejam $S = \{X_{\sigma(1)}...X_{\sigma(n)} : \sigma \in S_n \text{ e } \rho(\sigma) < g\}$ e A o K-subespaço de $K\{\mathbf{X}\}$ gerado por S. Basta provarmos que $A + I_n = V_n$, pois

$$c_n(I) = dim(V_n/I_n) = dim((A + I_n)/I_n) \leqslant dim(A) = |S|.$$

Claramente, $A + I_n \subseteq V_n$. Suponhamos, por absurdo, que exista algum $\sigma \in S_n$ tal que $m = X_{\sigma(1)}...X_{\sigma(n)} \notin A + I_n$; escolha σ tal que m seja minimal segundo a ordem parcial definida para monômios. Como $\rho(\sigma) \geqslant g$, existem $i_1 < ... < i_g$ com $\sigma(i_1) > ... > \sigma(i_g)$. Podemos escrever $m = m_1 X_{\sigma(i_1)} m_2 X_{\sigma(i_2)} ... m_g X_{\sigma(i_g)} m_{g+1}$ escolhendo apropriadamente cada $m_i \in \mu(\mathbf{X})$. Então $m - m_1 p(X_{\sigma(i_1)} m_2, X_{\sigma(i_2)} m_3, ..., X_{\sigma(i_g)} m_{g+1})$ é uma soma de monômios de ordem menor, e que portanto estão em $A + I_n$. Mas, pelo **Lema 6.1.1**,

$$m_1 p(X_{\sigma(i_1)} m_2, X_{\sigma(i_2)} m_3, ..., X_{\sigma(i_g)} m_{g+1}) \in I \cap V_n,$$

implicando que $m \in A + I_n$, o que é uma contradição. ■

Corolário 6.2.1 *Se* I *é um* T*-ideal de* $K\{\mathbf{X}\}$ *tal que* $K\{\mathbf{X}\}/I$ *satisfaz uma identidade polinomial* p *de grau* g, *então* $c_n(I) < (g-1)^{2n}$.

Prova: Para cada $\sigma \in S_n$ tal que $\rho(\sigma) < g$, podemos ver $T_1(\sigma)$ e $T_2(\sigma)$ como funções de $\{1, ..., n\}$ em $\{1, ..., g-1\}$. Como $T_1(\sigma)$ e $T_2(\sigma)$ determinam σ, então

$$|\{\sigma \in S_n : \rho(\sigma) < g\}| < ((g-1)^n)^2 = (g-1)^{2n}.$$

Pelo **Teorema 6.2.2**, temos que $c_n(I) < (g-1)^{2n}$. ■

6.3 Tabelas de Young

Nesta seção, definimos as tabelas de Young e desenvolvemos algumas de suas propriedades. Elas são instrumentos úteis para encontrar identidades polinomiais em álgebras sobre **Q** como veremos mais adiante.

Definição 6.3.1 *Uma* partição $\lambda = (\lambda_1, ..., \lambda_u)$ *de um número natural* n *é uma seqüência* $\lambda_1 \geqslant ... \geqslant \lambda_u > 0$ *de números naturais tal que* $\sum_{i=1}^{u} \lambda_i = n$. *Escrevemos* $\lambda \vdash n$.

Observação 6.3.1 *Podemos estabelecer de forma bastante natural uma bijeção entre as partições de* n *e as estruturas cíclicas de* S_n.

Definição 6.3.2 *Seja* $\lambda = (\lambda_1, ..., \lambda_u)$ *uma partição de* n. *O* diagrama de Young *desta partição consiste de* u *linhas de células tais que a linha* i *tem* λ_i *células. Uma* tabela de Young *para esta partição é um diagrama no qual estão escritos os números* $1, ..., n$, *um por célula. Dada uma tabela* t, *escrevemos* t_{ij} *para indicar o número que aparece na célula* j *da linha* i; *os elementos* t_{ij} *são chamados as* entradas *de* t.

Exemplo 6.3.1 *As tabelas correspondentes à partição* $(3,1,1)$ *de* 5 *são da forma*

t_{11}	t_{12}	t_{13}
t_{21}		
t_{31}		

Claramente existem 5! *tabelas correspondentes a esta partição.*

Definição 6.3.3 *Uma tabela é* "standard" *se* $t_{i_1 j_1} \leqslant t_{i_2 j_2}$, *sempre que* $i_1 \leqslant i_2$ *e* $j_1 \leqslant j_2$.

Em outras palavras, uma tabela é standard se os números nas linhas e nas colunas formam seqüências crescentes.

Exemplo 6.3.2 *Considerando as tabelas*

$$t_1 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 6 \\ \cline{1-2} 5 \\ \cline{1-1} \end{array} \quad e \; t_2 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 5 & 4 \\ \cline{1-2} 6 \\ \cline{1-1} \end{array}$$

verificamos que t_1 *é standard e* t_2 *não é standard.*

Por inspeção, vamos contar o número de tabelas standard para o diagrama do **Exemplo 6.3.1**. Obviamente t_{11} deve ser 1, e podemos colocar qualquer par de números ascendentes nas outras duas células da primeira linha, determinando uma tabela standard; assim há $\binom{4}{2} = 6$ tabelas standard correspondentes à partição $(3,1,1)$.

Vamos desenvolver agora uma fórmula para calcular o número de tabelas standard correspondentes a uma dada partição.

Definição 6.3.4 *Suponhamos que temos um diagrama de Young correspondente à partição* $(\lambda_1, ..., \lambda_u)$ *de* n. *Para a célula na linha* i *e coluna* j *definimos o número de gancho* g_{ij} *como* $g_{ij} = |\{(i', j') : i \leqslant i' \leqslant u,\ j \leqslant j' \leqslant m_i,$ *satisfazendo a condição adicional* $(i' = i$ *ou* $j' = j)\}|$.

O número de gancho é igual ao número de células "puxadas com um gancho" para a direita e para baixo a partir da célula dada.

Exemplo 6.3.3 *No* ***Exemplo 6.3.1****, temos os números de gancho* $g_{11} = 5$, $g_{12} = 2$, $g_{13} = 1$, $g_{21} = 2$ *e* $g_{31} = 1$. *No diagrama correspondente à partição* $(3, 2)$ *de* 5, *temos os números de gancho* $g_{11} = 4$, $g_{12} = 3$, $g_{13} = 1$, $g_{21} = 2$ *e* $g_{22} = 1$.

O próximo teorema determina exatamente o número de tabelas standard para uma dada partição. Sua demonstração envolve sobretudo métodos combinatórios e aparece em [32] como uma série de exercícios.

Teorema 6.3.1 (Fórmula do gancho) *O número de tabelas standard correspondentes à partição* $(\lambda_1, ..., \lambda_u)$ *de* n *é igual a*

$$\frac{n!}{\prod\{g_{ij} : 1 \leqslant i \leqslant u \text{ e } 1 \leqslant j \leqslant m_i\}}.$$

Exemplo 6.3.4 *O número de tabelas standard correspondentes à partição* $(3, 1, 1)$ *é*

$$\frac{5!}{5 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 3 \cdot 2 = 6,$$

o que está de acordo com o resultado obtido anteriormente. O número de tabelas standard correspondentes à partição $(3, 2)$ *é*

$$\frac{5!}{4 \cdot 3 \cdot 1 \cdot 2 \cdot 1} = 5.$$

Definição 6.3.5 *Sejam* $\sigma \in S_n$ *e* t *uma tabela correspondente à partição* $\lambda = (\lambda_1, ..., \lambda_u)$ *de* n. *Definimos* $\sigma(t)$ *como a tabela obtida de* t *trocando cada* t_{ij} *por* $\sigma(t_{ij})$. *Dizemos que* σ *é uma* permutação de linha *de* t *se* $\sigma(t_{ij}) \in \{t_{i1}, ..., t_{i\lambda_i}\}$ *para quaisquer* i *e* j, *ou seja,* σ *permuta apenas elementos dentro da mesma linha. Analogamente dizemos que* σ *é uma* permutação de coluna *de* t *se* $\sigma(t_{ij}) \in \{t_{1i}, t_{2i}, ...\}$ *para quaisquer* i *e* j. *Denotamos por* L_t *o conjunto de todas as permutações de linha de* t *e por* C_t *o conjunto de todas as permutações de coluna de* t. *O conjunto* L_t *é chamado* estabilizador das linhas *de* t *e* C_t *é chamado* estabilizador das colunas *de* t.

Proposição 6.3.1 *Se* t *é uma tabela correspondente à partição* $\lambda = (\lambda_1, ..., \lambda_u)$ *de* n, *então*

$$L_t \leqslant S_n,\ C_t \leqslant S_n \text{ e } L_t \cap C_t = \{1\}.$$

Além disso,

$$L_t = S_{L_1} \cdot S_{L_2} \cdot ... \cdot S_{L_u} \text{ e } C_t = S_{C_1} \cdot S_{C_2} \cdot ... \cdot S_{C_v},$$

onde L_i *é o conjunto dos elementos da linha* i, C_j *é o conjunto dos elementos da coluna* j *e* $v = \lambda_1$ *é o número de colunas de* t.

Exemplo 6.3.5 *Para a tabela*

$$t = \begin{array}{|c|c|c|}\hline 4 & 1 & 2 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$$

temos $R_t = S_{\{1,2,4\}} \cdot S_{\{3,5\}}$ *e* $C_t = S_{\{3,4\}} \cdot S_{\{1,5\}} \cdot S_{\{2\}}$.

6.4 Representações de S_n

Nesta seção, estenderemos a teoria da seção 6.1 quando $K = \mathbf{Q}$, o corpo dos números racionais, por motivos de simplicidade. Assim, o corpo em questão será sempre $\mathbf{Q}$ e a álgebra livre será $\mathbf{Q}\{\mathbf{X}\}$.

É possível introduzir uma estrutura de S_n-módulo no subespaço V_n de $\mathbf{Q}\{\mathbf{X}\}$, conforme o próximo lema.

Lema 6.4.1 *Sejam $\sigma \in S_n$ e $X_{\tau(1)}...X_{\tau(n)} \in V_n$, definimos $\sigma \cdot (X_{\tau(1)}...X_{\tau(n)}) = X_{\tau(\sigma(1))}...X_{\tau(\sigma(n))}$. Então V_n é um S_n-módulo sobre $\mathbf{Q}$.*

Recordemos que $S_n(Q)$ é a álgebra do grupo S_n sobre o corpo dos números racionais.

Lema 6.4.2 *A aplicação $\varphi : \mathbf{Q}(S_n) \to V_n$ dada por*

$$\varphi(\sum_{\sigma \in S_n} \alpha_\sigma \sigma) = \sum_{\sigma \in S_n} \alpha X_{\sigma(1)}...X_{\sigma(n)}$$

é um isomorfismo de álgebras.

Proposição 6.4.1 *Se I é um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$, então $I_n = I \cap V_n$ é um submódulo do S_n-módulo V_n.*

Prova: Se $p(X_1, ..., X_n) = \sum_{\sigma \in S_n} \alpha_\sigma X_{\sigma(1)}...X_{\sigma(n)} \in V_n$ e $\tau \in S_n$, então

$$\tau \cdot p(X_1, ..., X_n) = \sum_{\sigma \in S_n} \alpha_\sigma X_{\sigma(\tau(1))}...X_{\sigma(\tau(n))} = p(X_{\tau(1)}, ..., X_{\tau(n)}).$$

Pelo **Lema 6.1.1**, é fácil ver que I_n é um submódulo de V_n. ■

Teorema 6.4.1 *$\mathbf{Q}(S_n)$ é uma álgebra semisimples.*

Prova: Aplicação imediata do Teorema de Maschke, pois $|S_n| = n! < \infty$ e a característica de $\mathbf{Q}$ é zero. ■

Definição 6.4.1 *Seja A uma álgebra sobre um corpo K. Um elemento $a \in A$ é* idempotente *se $a^2 = a$. Um elemento $a \in A$ é* semi-idempotente *se $a^2 = \alpha a$ para algum $\alpha \in K$ e $\alpha \neq 0$.*

Proposição 6.4.2 *Sejam A uma álgebra sobre um corpo K, a um elemento de A e α um elemento não-nulo de K. Então $a^2 = \alpha a$ se, e somente se, $\alpha^{-1}a$ é idempotente.*

Definição 6.4.2 *Dada uma tabela t, definimos o elemento*

$$e^t = \sum_{\sigma \in L_t, \tau \in C_t} (-1)^\tau (\tau \circ \sigma).$$

de $\mathbf{Q}(S_n)$.

Proposição 6.4.3 *Se A é uma álgebra e a é um elemento de A, então $Aa = \{b \cdot a : b \in A\}$ é um ideal à esquerda de A.*

Os próximos teoremas descrevem quase completamente a estrutura de $\mathbf{Q}(S_n)$ em termos de tabelas de Young. Suas provas podem ser encotradas em Miller [24].

Teorema 6.4.2 *Para qualquer tabela t, e^t é um elemento semi-idempotente de $\mathbf{Q}(S_n)$ e $[\mathbf{Q}(S_n)]e^t$ é um ideal simples à esquerda de $\mathbf{Q}(S_n)$. Além disso, as seguintes afirmações são equivalentes:*
(i) As tabelas t_1 e t_2 correspondem ao mesmo diagrama;
(ii) $[\mathbf{Q}(S_n)]e^{t_1}$ e $[\mathbf{Q}(S_n)]e^{t_2}$ estão na mesma componente simples de $\mathbf{Q}(S_n)$;
(iii) $e^{t_2} = \sigma^{-1}e^{t_1}\sigma$ para algum σ em S_n.

Teorema 6.4.3 *Todo ideal simples à esquerda de $\mathbf{Q}(S_n)$ é isomorfo (como $\mathbf{Q}(S_n)$-módulo) a algum ideal à esquerda $[\mathbf{Q}(S_n)]e^t$ para alguma tabela standard t apropriada.*

Teorema 6.4.4 *Se t é uma tabela de um diagrama d, então $dim_{\mathbf{Q}}([\mathbf{Q}(S_n)]e^t)$ é o número de tabelas standard correspondentes a d.*

Definição 6.4.3 *Dada uma tabela t, definimos*

$$p_t = \sum_{\sigma \in I_t, \tau \in C_t} (-1)^\tau X_{\tau\circ\sigma(1)} \ldots X_{\tau\circ\sigma(n)}$$

e

$$q_t = \sum_{\tau \in C_t} (-1)^\tau X_{\tau(1)} \ldots X_{\tau(n)}.$$

Observação 6.4.1 *Se considerarmos o isomorfismo $\varphi : \mathbf{Q}(S_n) \to V_n$ como no **Lema 6.4.2**, então $\varphi(e^t) = p_t$.*

Proposição 6.4.4 *Se t é uma tabela então*

$$p_t = \sum_{\sigma \in I_t} \sigma(q_t).$$

Prova: Manipulando algumas fórmulas

$$\begin{aligned} p_t &= \sum_{\sigma \in I_t, \tau \in C_t} (-1)^\tau X_{\tau\circ\sigma(1)} \ldots X_{\tau\circ\sigma(n)} = \sum_{\sigma \in I_t, \tau \in C_t} (-1)^\tau \sigma(X_{\tau(1)} \ldots X_{\tau(n)}) = \\ &= \sum_{\sigma \in I_t} \sigma \cdot \sum_{\tau \in C_t} (-1)^\tau X_{\tau(1)} \ldots X_{\tau(n)}) = \sum_{\sigma \in I_t} \sigma(q_t). \end{aligned}$$

concluímos o resultado. ■

Proposição 6.4.5 *Seja t uma tabela correspondente à partição $\lambda = (\lambda_1, \ldots, \lambda_u)$ de n e I um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ tal que $p_t \in I$. Se ψ é um endomorfismo de $\mathbf{Q}\{\mathbf{X}\}$ tal que $\psi(X_{t_{ij}}) = X_i$ para todo $1 \leqslant i \leqslant u$ e $1 \leqslant j \leqslant \lambda_i$, então $\psi(p_t) \in I$.*

Prova: Para todo $\sigma \in L_t$, $\psi(X_{\sigma(t_{ij})}) = X_i = \psi(X_{t_{ij}})$. Logo,

$$\begin{aligned}\psi(p_t) &= \psi(\sum_{\sigma \in L_t}((X_{\sigma(1)}...X_{\sigma(n)})q_t)) = (\lambda_1!...\lambda_u!)\psi(X_1...X_n)\psi(q_t) = \\ &= (\lambda_1!...\lambda_u!)\psi((X_1...X_n)q_t) = (\lambda_1!...\lambda_u!)\psi(q_t).\end{aligned}$$

Como $\psi(p_t) \in I$ então $\psi(q_t) \in I$. ■

Proposição 6.4.6 *Seja t uma tabela. tal que a coluna j de t tem comprimento q_j. Se t tem entradas $t_{ij} = i + \sum_{p=1}^{j-1} q_p$ para quaisquer i e j, então*

$$p_t = s_{q_1}(X_1, ..., X_{q_1})s_{q_2}(X_{q_1+1}, ..., X_{q_1+q_2})...,$$

um produto de polinômios standard.

Proposição 6.4.7 *Suponha que I seja um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ e que para algum diagrama d temos $p_t \in I$ para qualquer tabela t correspondente a d. Suponha também que d tem v colunas e que a coluna j tem comprimento q_j, $1 \leqslant j \leqslant v$. Então*

$$s_{q_1}(X_1, ..., X_{q_1})...s_{q_v}(X_1, ..., X_{qv}) \in I.$$

Prova: Basta combinar as duas proposições anteriores. ■

6.5 Alguns teoremas de Regev

Os próximos resultados foram publicados em 1978 por Regev [30].

Teorema 6.5.1 (Regev) *Seja I um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$. Se existe uma partição $(\lambda_1, ..., \lambda_u)$ de n cujo diagrama d tem mais que $c_n(I)$ tabelas standard, então I_n contém todos os S_n-módulos irredutíveis que correspondem a d.*

Prova: Seja s o número de tabelas standard correspondentes a d, e seja B o submódulo de V_n correspondente a d. Então B é uma soma de submódulos irredutíveis, assim é suficiente mostrar que $L \subseteq I_n$ para todo submódulo irredutível L de V_n contido em B. Por outro lado, se $L \cap I_n = \{0\}$ então

$$dim_{\mathbf{Q}}(V_n) \geqslant dim_{\mathbf{Q}}(L) + dim_{\mathbf{Q}}(I_n) \text{ e } c_n(I) \geqslant dim_{\mathbf{Q}}(L) = s,$$

contrariando a hipótese. ■

Lema 6.5.1 *Seja d o diagrama correspondente à partição $(\lambda, ..., \lambda)$ de n com u linhas. Seja e a base dos logaritmos naturais. Se I é um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ com $c_n(I) \leqslant \alpha^n$, onde $(\lambda u)/(\lambda + u) \geqslant e\alpha/2$, então d tem mais que $c_n(I)$ tabelas standard.*

Prova: Sejam g_{ij} os números de gancho de d. Então, desde que a média aritmética é maior que a média geométrica, temos

$$\begin{aligned}(\prod_{i,j} g_{ij})^{1/n} &\leqslant (\sum_{i,j} g_{ij})/n = (\sum_{i=1}^{u}\sum_{j=1}^{\lambda}(i+j-1))/n = (\sum_{i=1}^{u}(i\lambda+\lambda(\lambda-1)/2))/n = \\ &= (\lambda u(u+1)+\lambda(\lambda-1)u)/2n = (u+\lambda)/2.\end{aligned}$$

É fácil ver (através de integração por partes) que

$$n(\log(n)-1)+1 = \int_1^n \log(x)dx < \sum_{i=1}^{n}\log(i) = \log(n!),$$

onde log denota o logaritmo natural, então $(\lambda u/e)^n < n!$. Por hipótese,

$$\alpha \leqslant (\lambda u/e)(2/(\lambda+u)) < (n!/\prod_{i,j} g_{ij})^{1/n}$$

e então, pela fórmula do gancho, o número de tabelas standard de d é maior que $\alpha^n \geqslant c_n(I)$.■

Teorema 6.5.2 (Regev) *Seja I um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ que contém um polinômio multilinear de grau g. Se m e u são números naturais tais que $(mu)/(m+u) \geqslant (g-1)^2e/2$, onde e é a base dos logaritmos naturais, então I_n contém o ideal bilateral de V_n correspondente ao diagrama retangular de u linhas e m colunas; em particular, $(s_u(X_1,...,X_u))^m \in I$.*

Prova: Pelo **Corolário 6.1.1**, podemos aplicar o **Lema 6.3.3**, com $\alpha = (g-1)^2$ ao **Teorema 6.5.1** e ao **Lema 6.1.2**. Assim, obtemos o resultado desejado. ■

Corolário 6.5.1 (Regev) *Sejam A uma álgebra associativa sobre $\mathbf{Q}$ que satisfaz uma identidade polinomial p de grau g e $\beta = (g-1)^2e/2$. Se m e u são números naturais tais que $m > u\beta/(u-\beta)$, então $(s_u(X_1,...,X_u))^m$ é uma identidade polinomial de A.*

Prova: Como

$$m > \frac{u\beta}{u-\beta} \text{ implica } \frac{mu}{m+u} > \beta,$$

pelo **Teorema 6.5.2**, temos que $p \in T(A)$, ou seja, p é uma identidade polinomial de A. ■

Em 1971, Amitsur [1] provou que se A é uma PI álgebra sobre $\mathbf{Q}$, então A satisfaz uma identidade polinomial da forma $(s_u(X_1,...,X_u))^m$ para números naturais m e u apropriados. O resultados obtido por Regev é melhor, pois nele m é explicitamente computado.

6.6 Relações de ordem em $\mathbf{X}^*$

Definição 6.6.1 *A relação $\leqslant$ para o conjunto das palavras sobre $\mathbf{X}$ é definida da seguinte forma: escrevemos $X_{i_1}...X_{i_k} \leqslant X_{j_1}...X_{j_l}$ se um dos dois casos ocorre:*
(i) $X_{i_1}...X_{i_k} = X_{j_1}...X_{j_l}$, isto é, $k=l$ e $i_t = j_t$ para $1 \leqslant t \leqslant k$;
(ii) para algum $r \in \{1,...,k\}$ temos que $i_t = j_t$ se $1 \leqslant t \leqslant r-1$ e $i_r < j_r$.
Escrevemos $X_{i_1}...X_{i_k} < X_{j_1}...X_{j_l}$ se $X_{i_1}...X_{i_k} \leqslant X_{j_1}...X_{j_l}$ e $X_{i_1}...X_{i_k} \neq X_{j_1}...X_{j_l}$.

Exemplo 6.6.1 *Para as palavras* X_1X_2, X_1X_3, $X_2X_3X_4$, X_3, *temos que* $X_1X_2 \leqslant X_1X_3$, $X_1X_3 \leqslant X_2X_3X_4$ *e* $X_2X_3X_4 \leqslant X_3$.

Observação 6.6.1 *A relação* $\leqslant$ *é uma relação de ordem parcil no conjunto das palavras sobre* $\mathbf{X}$, *mas não é uma relação de ordem total. As palavras* X_1X_2 *e* $X_1X_2X_3$, *por exemplo, não são comparáveis.*

Definição 6.6.2 *A ordem* lexicográfica *(denotada por* $\preccurlyeq$*) para o conjunto das palavras sobre* $\mathbf{X}$ *é definida da seguinte forma: escrevemos* $X_{i_1}...X_{i_k} \preccurlyeq X_{j_1}...X_{j_l}$ *se um dos três casos ocorre:*
(i) $X_{i_1}...X_{i_k} = X_{j_1}...X_{j_l}$, *isto é,* $k = l$ *e* $i_t = j_t$ *para* $1 \leqslant t \leqslant k$;
(ii) para algum $r \in \{1, ..., k\}$ *temos que* $i_t = j_t$ *se* $1 \leqslant t \leqslant r-1$ *e* $i_r < j_r$,
(iii) $k < l$ *e* $i_t = j_t$ *para* $1 \leqslant t \leqslant k$.
Escrevemos $X_{i_1}...X_{i_k} \prec X_{j_1}...X_{j_l}$ *se* $X_{i_1}...X_{i_k} \preccurlyeq X_{j_1}...X_{j_l}$ *e* $X_{i_1}...X_{i_k} \neq X_{j_1}...X_{j_l}$.

Observação 6.6.2 *A ordem lexicográfica, embora diferencie-se da ordem* $\leqslant$ *apenas pelo acréscimo do caso (iii) na sua definição, é uma relação de ordem total no conjunto das palavras sobre* $\mathbf{X}$. *As palavras* X_1X_2 *e* $X_1X_2X_3$, *por exemplo, são comparáveis,* $X_1X_2 \preccurlyeq X_1X_2X_3$. *É interessante observar que a ordem lexicográfica é a relação que ordena as palavras, por exemplo, num dicionário, daí o nome* lexicográfica.

Consideremos um subconjunto S de $\mu(\mathbf{X})$ tal que $gr(m) = k$ para todo monômio m de S. É fácil ver que a ordem $\leqslant$ coincide com a ordem lexicográfica para as palavras que compõem os monômios de S. Portanto, faz sentido a próxima definição.

Definição 6.6.3 *Se* K *é um corpo e* p *é um polinômio multilinear de* $K\{\mathbf{X}\}$, *denotamos por* $L(p)$ *o máximo do conjunto das palavras que compõem os monômios de* p, *com relação à ordem* $\leqslant$.

Exemplo 6.6.2 *Para o polinômio* $p(X_1, X_2, X_3) = 14X_1X_2X_3 + 8X_2X_1X_3 + 5X_3X_1X_2 \in \mathbf{R}\{\mathbf{X}\}$, *temos que* $L(p) = X_3X_1X_2$.

6.7 Problema de Specht

Definição 6.7.1 *Seja* S *um subconjunto de uma álgebra* A. *O* T-ideal gerado por S *(denotado por* $\langle S\rangle_T$*) é definido como a interseção de todos os* T*-ideais de* A *que contém* S. *Se* S *é um conjunto finito dizemos que o* T*-ideal gerado por* S *é um* T-ideal finitamente gerado, *e denotamos* $\langle S\rangle_T = \langle a_1, ..., a_n\rangle_T$ *se* $S = \{a_1, ..., a_n\}$.

Em 1950, W. Specht [35] levantou a seguinte questão:

Problema de Specht: *Se* A *é uma álgebra associativa sobre um corpo de característica zero, então* $T(A)$ *é finitamente gerado como* T*-ideal?*

Alguns dos primeiros exemplos não-triviais de T-ideais finitamente gerados foram dados por Yu. P. Razmyslov. Entre outros resultados, em 1973 ele provou que $T(M_2(K))$ é gerado

por 9 identidades polinomiais, quando a característica de K é zero (ver [28]). Em 1981, V. Drensky [4] provou que $T(M_2(K))$ é gerado por

$$\{[[X_1, X_2]^2, X_3], s_4(X_1, X_2, X_3, X_4)\},$$

quando a característica de K é zero.

Em 1976, Latyshev [22] obteve uma resposta positiva ao Problema de Specht para álgebras associativas sobre $\mathbf{Q}$ que satisfazem a identidade polinomial $[X_1, X_2]...[X_{2n-1}, X_{2n}]$, para algum número natural n. Provaremos nesta seção este resultado de Latyshev que se apóia fortemente no Teorema de Higman.

Definição 6.7.2 *Seja $w = X_{i_1}...X_{i_k}$ uma palavra de $\mathbf{X}^*$. A* altura da seqüência $(i_1, ..., i_k)$ *é chamada* altura da palavra w.

Definição 6.7.3 *Sejam K um corpo, p um polinômio multilinear de $K\{\mathbf{X}\}$ e $J = \langle p \rangle_T$. Dizemos que p é um* n-polinômio *se para todo T-ideal I que contém J, para cada número natural k e para todo q em I_k temos que $q \in J_k + \{r \in I_k : L(r) \leqslant L(q)$ e $L(r)$ tem altura $\leqslant n\}$.*

Proposição 6.7.1 *O polinômio $p(X_1, ..., X_{2n}) = [X_1, X_2]...[X_{2n-1}, X_{2n}]$ é um $2n$-polinômio.*

Prova: Sejam $J = \langle [X_1, X_2]...[X_{2n-1}, X_{2n}] \rangle_T$, $k \in \mathbf{N}$ e I um T-ideal que contém J. Para g em I_k, escrevemos $V(g) = J_k + \{r \in I_k : L(r) \leqslant L(p)$ e $L(r)$ tem altura $\leqslant n\}$. Suponhamos, por absurdo, que exista um g em I_k tal que $g \notin V(g)$; escolhemos g com $L(g)$ minimal. Podemos assumir que o coeficiente de $L(g)$ é 1. Obviamente, $L(g)$ tem altura maior que $2n$ então, escrevendo $L(g) = X_{\sigma(1)}...X_{\sigma(k)}$, temos que existem inteiros $1 \leqslant i_1 < ... < i_{2n} < k$ tais que $\sigma(i_u) > \sigma(i_u + 1)$ para $1 \leqslant u \leqslant 2n$. Consideremos $i_{2n+1} = k + 1$ e vamos construir os monômios $y_0, y_1, y_2, ..., y_{2n}$ da seguinte maneira:

1) $y_0 = X_{\sigma(1)}...X_{\sigma(i_1 - 1)}$ (se $i_1 = 1$ então tome $y_0 = 1$);

2) $y_{2j-1} = X_{\sigma(i_{2j-1})}$ para $1 \leqslant j \leqslant n$;

3) $y_{2j} = X_{\sigma(i_{2j-1}+1)}...X_{\sigma(i_{2j+1}-1)}$ para $1 \leqslant j \leqslant n$.

Seja $g' = g - y_0[y_1, y_2]...[y_{2n-1}, y_{2n}]$. Pela construção dos monômios $y_0, y_1, y_2, ..., y_{2n}$, temos que

$$L(y_0[y_1, y_2]...[y_{2n-1}, y_{2n}]) = L(g) = y_0 y_1 y_2 ... y_{2n-1} y_{2n};$$

logo, $L(g') < L(g)$, e, por hipótese, $g' \in V(g') \subseteq V(g)$. Pelo **Lema 6.1.1** temos que

$$[y_1, y_2]...[y_{2n-1}, y_{2n}] \in J;$$

logo,

$$y_0[y_1, y_2]...[y_{2n-1}, y_{2n}] \in J_k.$$

Assim $g \in V(g) + J_k = V(g)$, o que é uma contradição. ■

Para o próximo teorema precisaremos da seguinte proposição:

Proposição 6.7.2 *Seja K um corpo de característica zero. Se $I \lhd_T K\{\mathbf{X}\}$, então I é gerado como T-ideal por todos os seus polinômios multilineares.*

Teorema 6.7.1 *Todo T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ contendo um $2n$-polinômio, para algum n, é finitamente gerado como T-ideal.*

Prova: Seja I um T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ contendo um $2n$-polinômio p. Suponhamos, por absurdo, que I não é finitamente gerado. Vamos formar uma seqüência infinita de polinômio multilineares $p_i \in I$, $1 \leqslant i < \infty$, escolhidos indutivamente como mostramos a seguir. Seja $p_1 = p$; dados $p_1, ..., p_j$ tomamos $p_{j+1} \in I - \langle p_1, ..., p_j \rangle_T$ tal que $L(p_{j+1})$ é minimal e tem altura $\leqslant 2n$. Podemos assumir que o coeficiente de $L(p_{j+1})$ é 1. Aplicando o **Teorema 5.4.1** às matrizes representação de $\{L(p_i) : 1 \leqslant i < \infty\}$, para inteiros $i < j$ apropriados, escrevendo $L(p_i) = X_{\sigma(1)}...X_{\sigma(k)}$ e $L(p_j) = X_{\tau(1)}...X_{\tau(l)}$, onde $\sigma \in S_k$ e $\tau \in S_l$, existe alguma aplicação $\psi : \{1, ..., k\} \to \{1, ..., l\}$ tal que $\psi(r) < \psi(s)$ se $r < s$ e

$$L(p_j) = w_0 X_{\psi(\sigma(1))} w_1 X_{\psi(\sigma(2))} w_2 ... w_{k-1} X_{\psi(\sigma(k))} w_k$$

para palavras w_u adequadas. Então

$$L(p_j - w_0 p_i(X_{\psi(1)} w_{\sigma^{-1}(1)}, X_{\psi(2)} w_{\sigma^{-1}(2)}, ..., X_{\psi(k)} w_{\sigma^{-1}(k)})) < L(p_j),$$

pois o monômio líder foi eliminado. Por hipótese,

$$p_j - w_0 p_i(X_{\psi(1)} w_{\sigma^{-1}(1)}, X_{\psi(2)} w_{\sigma^{-1}(2)}, ..., X_{\psi(k)} w_{\sigma^{-1}(k)}) \in \langle p_1, ..., p_{j-1} \rangle_T,$$

implicando $p_j \in \langle p_1, ..., p_{j-1} \rangle_T$, o que é uma contradição. ■

Corolário 6.7.1 *Todo T-ideal de $\mathbf{Q}\{\mathbf{X}\}$ que contém o polinômio $p(X_1, ..., X_{2n}) = [X_1, X_2]...[X_{2n-1}, X_$ para algum n, é finitamente gerado como T-ideal.*

Exemplo 6.7.1 *Seja A a álgebra das matrizes triangulares superiores $n \times n$ sobre $\mathbf{Q}$. Como*

$$p(X_1, ..., X_{2n}) = [X_1, X_2][X_3, X_4]...[X_{2n-1}, X_{2n}] \in T(A)$$

*(ver **Exemplo 3.3.3**), temos que $T(A)$ é finitamente gerado.*

Em 1987, A. R. Kemer provou que se A é uma álgebra associativa sobre um corpo de característica zero, então $T(A)$ é finitamente gerado como T-ideal (ver [15] ou [16]). Dessa forma, ele conseguiu uma resposta positiva para o Problema de Specht. O resultado de Latyshev é um caso particular do Teorema de Kemer, e foi incluído pois historicamente é o primeiro resultado não-trivial nessa direção. Além disso, a prova do Teorema de Kemer é de tal extensão que não poderia ser incluso neste trabalho.

Capítulo 7

Produto tensorial de PI álgebras

Definimos produto tensorial de espaços vetoriais e de álgebras, e provamos algumas de suas propriedades. Provamos que o produto tensorial de duas PI álgebras é uma PI álgebra e apresentamos a prova de Rosset para o Teorema de Amitsur e Levitzki.

7.1 Produto tensorial de espaços vetoriais

Sejam A e B espaços vetoriais sobre um corpo K. O produto cartesiano $A \times B$ é também um espaço vetorial sobre K, considerando a soma como

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

e a multiplicação por escalar como

$$\lambda(a, b) = (\lambda a, \lambda b).$$

Este espaço é chamado soma direta externa de A e B.

Definição 7.1.1 *Sejam A e B espaços vetoriais sobre um corpo K. Indicamos por $(A \times B)^+$ o conjunto*

$$\{ \sum_{(a,b) \in A \times B} n_{ab}(a, b) : n_{ab} \in \mathbf{Z} \text{ e } n_{ab} \neq 0 \text{ para um número finito de pares ordenados } (a, b) \}.$$

Observação 7.1.1 *Se definirmos a soma de dois elementos de $(A \times B)^+$ por*

$$\sum_{(a,b) \in A \times B} m_{ab}(a, b) + \sum_{(a,b) \in A \times B} n_{ab}(a, b) = \sum_{(a,b) \in A \times B} (m_{ab} + n_{ab})(a, b)$$

então $(A \times B)^+$ é um grupo abeliano.

Observação 7.1.2 *Cada elemento $\sum_{(a,b) \in A \times B} n_{ab}(a, b)$ pode ser escrito como $\sum_{i=1}^{n} (a_i, b_i)$.*

Definição 7.1.2 *Sejam A e B espaços vetoriais sobre um corpo K. Definimos $H_{A\times B}$ como o subgrupo de $(A \times B)^+$ gerado por todos os elementos dos tipos:*
(i) $(a_1 + a_2, b_1) - (a_1, b_1) - (a_2, b_1)$
(ii) $(a_1, b_1 + b_2) - (a_1, b_1) - (a_1, b_2)$
(iii) $(\lambda a_1, b_1) - (a_1, \lambda b_1)$
para quaisquer a_1 e a_2 em A, b_1 e b_2 em B e λ em K. Definimos $A \otimes B = (A \times B)^+ / H_{A\times B}$ como o produto tensorial *de A e B. Denotamos por $a \otimes b$ cada elemento $(a, b) + H_{A\times B}$ de $A \otimes B$.*

Observação 7.1.3 *Obviamente $A \otimes B$ é um grupo abeliano. Além disso, cada elemento $(\sum_{(a,b)\in A\times B} n_{ab}(a, b)) + H_{A\times B}$ pode ser escrito como $\sum_{i=1}^{n} a_i \otimes b_i$.*

Definição 7.1.3 *Sejam A e B espaços vetoriais sobre um corpo K e G um grupo abeliano. Dizemos que a aplicação $\psi : A \times B \to G$ é* bilinear *se para quaisquer a_1 e a_2 em A, b_1 e b_2 em B e λ em K as seguintes condições são satisfeitas:*
(i) $\psi(a_1 + a_2, b_1) = \psi(a_1, b_1) + \psi(a_2, b_1)$;
(ii) $\psi(a_1, b_1 + b_2) = \psi(a_1, b_1) + \psi(a_1, b_2)$;
(iii) $\psi(\lambda a_1, b_1) = \psi(a_1, \lambda b_1)$.

Proposição 7.1.1 *Sejam A e B espaços vetoriais e G um grupo abeliano. Se $\psi : A \times B \to G$ é uma aplicação bilinear, então existe um único homomorfismo de grupos $\psi' : A \otimes B \to G$ tal que $\psi'(a \otimes b) = \psi(a, b)$ para todo a em A e b em B.*

Prova: A aplicação ψ pode ser extendida a um homomorfismo de grupos $\psi : (A \times B)^+ \to G$ definindo $\psi(\sum_{(a,b)\in A\times B} n_{ab}(a, b)) = \sum_{(a,b)\in A\times B} \psi(n_{ab}(a, b))$. Observemos que o subgrupo $H_{A\times B}$ está contido no núcleo de ψ. Consideremos a aplicação $\psi' : A \otimes B \to G$, dada por $\psi'(x + H_{A\times B}) = \psi(x)$.

Vamos provar que ψ' está bem definida. Se $x + H_{A\times B} = y + H_{A\times B}$ então $x - y \in H_{A\times B}$. Logo, $\psi(x - y) = 0$ implica $\psi(x) = \psi(y)$, ou seja, $\psi'(x + H_{A\times B}) = \psi'(y + H_{A\times B})$.

É fácil ver que ψ' é um homomorfismo de grupos e que $\psi'(a \otimes b) = \psi(a, b)$ para todo a em A e b em B.

Claramente, ψ' é único. ■

Lema 7.1.1 *Sejam A, B, A' e B' espaços vetoriais. Se $\psi_1 : A \to A'$ e $\psi_2 : B \to B'$ são transformações lineares, então existe um único 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 em A e b em B.*

Prova: Vamos definir $\psi : A \times B \to A' \otimes B'$ por $\psi(a, b) = \psi_1(a) \otimes \psi_2(b)$. Obviamente, ψ é bilinear. Logo, pela **Proposição 7.1.3**, existe um homomorfismo de grupos de $A \otimes B$ em $A' \otimes B'$ possuindo a propriedade estabelecida. Claramente, este homomorfismo é único. ■

Teorema 7.1.1 *Se A e B são espaços vetoriais sobre um corpo K, então $A \otimes B$ é um espaço vetorial sobre K, com multiplicação por escalar definida por $\lambda \sum_{i=1}^{n} (a_i \otimes b_i) = \sum_{i=1}^{n} ((\lambda a_i) \otimes b_i)$.*

Prova: Dado λ em K, definimos $\psi_\lambda : A \to A$ por $\psi_\lambda(a) = \lambda a$. Como ψ_λ é uma transformação linear, definimos $\lambda \sum_{i=1}^n (a_i \otimes b_i) = (\psi_\lambda \otimes id_B)(\sum_{i=1}^n (a_i \otimes b_i))$. Logo, $\lambda \sum_{i=1}^n (a_i \otimes b_i) = \sum_{i=1}^n ((\lambda a_i) \otimes b_i)$. Como $\psi_\lambda \otimes id_B$ está bem definida, temos que o produto por escalar está bem definido e $A \otimes B$ é um espaço vetorial sobre K. ■

Teorema 7.1.2 *Sejam A, B, A' e B' espaços vetoriais. Se $\psi_1 : A \to A'$ e $\psi_2 : B \to B'$ são transformações lineares, então existe uma única transformação linear, denotada 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 em A e b em B.*

Prova: Vamos definir $\psi : A \times B \to A' \otimes B'$ por $\psi(a, b) = \psi_1(a) \otimes \psi_2(b)$. Obviamente, ψ é bilinear. Logo, pela **Proposição 7.1.3**, existe um único homomorfismo de grupos $\psi' : A \otimes B \to A' \otimes B'$ tal que $\psi'(a \otimes b) = \psi(a, b) = \psi_1(a) \otimes \psi_2(b)$ para todo a em A e b em B. Além disso,

$$\begin{aligned}
\psi'(\lambda \sum_{i=1}^n (a_i \otimes b_i)) &= \psi'(\sum_{i=1}^n ((\lambda a_i) \otimes b_i)) = \sum_{i=1}^n \psi'((\lambda a_i) \otimes b_i)) = \\
&= \sum_{i=1}^n \psi_1(\lambda a_i) \otimes \psi_2(b_i) = \sum_{i=1}^n (\lambda \psi_1(a_i)) \otimes \psi_2(b_i) = \\
&= \sum_{i=1}^n (\lambda(\psi_1(a_i) \otimes \psi_2(b_i))) = \sum_{i=1}^n (\lambda \psi'(a_i \otimes b_i)) = \\
&= \lambda \sum_{i=1}^n \psi'(a_i \otimes b_i) = \lambda \psi'(\sum_{i=1}^n a_i \otimes b_i).
\end{aligned}$$

Portanto, ψ' é uma transformação linear. Definindo $\psi_1 \otimes \psi_2 = \psi'$ temos a transformação linear desejada. Claramente, $\psi_1 \otimes \psi_2$ é única. ■

Teorema 7.1.3 *Se A e B são espaços vetoriais sobre um corpo K, então*
(i) se $a_1, ..., a_m$ são vetores linearmente independentes de A e $b_1, ..., b_m$ são vetores de B tais que $\sum_{i=1}^m a_i \otimes b_i = 0$, então $b_i = 0$ para todo $i \in \{1, ..., m\}$;
(ii) se $a_1, ..., a_m$ são vetores de A e $b_1, ..., b_m$ são vetores linearmente independentes de B tais que $\sum_{i=1}^m a_i \otimes b_i = 0$, então $a_i = 0$ para todo $i \in \{1, ..., m\}$;
(iii) se $\{e_1, ..., e_m\}$ é uma base de A e $\{f_1, ..., f_n\}$ é uma base de B, então $\{e_i \otimes f_j : 1 \leqslant i \leqslant m$ e $1 \leqslant j \leqslant n\}$ é uma base de $A \otimes B$. Além disso, se $a = \sum_{i=1}^m \alpha_i e_i$ e $b = \sum_{j=1}^n \beta_j f_j$ então $a \otimes b = \sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j (e_i \otimes f_j)$.

Prova: Para provar a parte (i), consideremos **A** um conjunto de vetores linearmente independentes de A tais que $\{a_1, ..., a_m\} \cup \mathbf{A}$ é uma base de A. Seja $\psi_k : A \times B \to B$ a aplicação dada por $\psi_k(a, b) = \alpha_k b$, onde α_k é o coeficiente de a_k, quando escrevemos a na base $\{a_1, ..., a_m\} \cup \mathbf{A}$. Claramente, ψ_k é bilinear. Logo, pela **Proposição 7.1.3**, existe um homomorfismo de grupos $\psi'_k : A \times B \to B$ tal que $\psi'_k(a \otimes b) = \psi_k(a, b)$. Como $\sum_{i=1}^m a_i \otimes b_i = 0$, temos que $b_k = \psi'_k(\sum_{i=1}^m a_i \otimes b_i) = \psi'_k(0) = 0$. Portanto, $b_i = 0$ para todo $i \in \{1, ..., m\}$.

Analogamente provamos a parte (ii).

Para provar a parte (iii), vamos mostrar que $\mathbf{B} = \{e_i \otimes f_j : 1 \leqslant i \leqslant m \text{ e } 1 \leqslant j \leqslant n\}$ é uma base de $A \otimes B$. Suponhamos, por absurdo, que existam coeficientes γ_{ij}, $1 \leqslant i \leqslant m$ e $1 \leqslant j \leqslant n$, tais que $\sum_{i=1}^{m}(\sum_{j=1}^{n} \gamma_{ij}(e_i \otimes f_j)) = 0$. Logo,

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\gamma_{ij}(e_i \otimes f_j) = \sum_{i=1}^{m}\sum_{j=1}^{n}(e_i \otimes \gamma_{ij}f_j) = \sum_{i=1}^{m} e_i \otimes (\sum_{j=1}^{n}\gamma_{ij}f_j).$$

Portanto, para cada i, temos que $\sum_{j=1}^{n}\gamma_{ij}f_j = 0$, de onde concluímos que $\gamma_{ij} = 0$ para $1 \leqslant i \leqslant m$ e $1 \leqslant j \leqslant n$. Portanto, $\mathbf{B}$ é linearmente independente.

Vamos provar que $\mathbf{B}$ gera $A \otimes B$. Sejam $a = \sum_{i=1}^{m}\alpha_i e_i \in A$ e $b = \sum_{j=1}^{n}\beta_j f_j \in B$. Logo,

$$\begin{aligned} a \otimes b &= \sum_{i=1}^{m}\alpha_i e_i \otimes \sum_{j=1}^{n}\beta_j f_j = \sum_{i=1}^{m}(\alpha_i e_i \otimes \sum_{j=1}^{n}\beta_j f_j) = \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha_i e_i \otimes \beta_j f_j) = \sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_i\beta_j(e_i \otimes f_j). \end{aligned}$$

Assim, é fácil ver que qualquer elemento $\sum_{i=1}^{p} a_i \otimes b_i$ de $A \otimes B$ pode ser escrito como uma combinação linear de elementos de $\mathbf{B}$. Portanto, $\mathbf{B}$ gera $A \otimes B$ e é uma base de $A \otimes B$. ■

Quando consideramos aplicações bilineares que têm como contra-domínio um espaço vetorial (ao invés de apenas um grupo), acrescentando mais uma propriedade na condição (iii) da **Definição 7.1.3**, podemos obter novos resultados.

Definição 7.1.4 *Sejam A, B e V espaços vetoriais sobre um corpo K. Dizemos que a aplicação $\psi : A \times B \to V$ é* bilinear *se para quaisquer a_1 e a_2 em A, b_1 e b_2 em B e λ em K as seguintes condições são satisfeitas:*
(i) $\psi(a_1 + a_2, b_1) = \psi(a_1, b_1) + \psi(a_2, b_1)$;
(ii) $\psi(a_1, b_1 + b_2) = \psi(a_1, b_1) + \psi(a_1, b_2)$;
(iii) $\psi(\lambda a_1, b_1) = \psi(a_1, \lambda b_1) = \lambda\psi(a_1, b_1)$.

Teorema 7.1.4 *Sejam A, B e V espaços vetoriais sobre uma corpo K. Se $\psi : A \times B \to V$ é uma aplicação bilinear, então existe uma única transformação linear $\psi' : A \otimes B \to V$ tal que $\psi'(a \otimes b) = \psi(a, b)$ para todo a em A e b em B.*

Prova: Pela **Proposição 7.1.3**, temos que existe um homomorfismo de grupos $\psi' : A \otimes B \to V$ tal que $\psi'(a \otimes b) = \psi(a, b)$ para todo a em A e b em B. Se $\sum_{i=1}^{n} a_i \otimes b_i \in A \otimes B$ e $\lambda \in K$, então

$$\begin{aligned} \psi'(\lambda\sum_{i=1}^{n} a_i \otimes b_i) &= \psi'(\sum_{i=1}^{n}(\lambda a_i) \otimes b_i) = \sum_{i=1}^{n}\psi'((\lambda a_i) \otimes b_i)) = \sum_{i=1}^{n}\psi(\lambda a_i, b_i) = \\ &= \sum_{i=1}^{n}\lambda\psi(a_i, b_i) = \lambda\sum_{i=1}^{n}\psi(a_i, b_i) = \lambda\sum_{i=1}^{n}\psi'(a_i \otimes b_i) = \\ &= \lambda\psi'(\sum_{i=1}^{n} a_i \otimes b_i). \end{aligned}$$

Portanto, ψ' é uma transformação linear.
Claramente, ψ' é única. ■

Teorema 7.1.5 *Se A e B são espaços vetoriais sobre um corpo K e $\tau : A \times B \to A \otimes B$ é a aplicação dada por $\tau(a,b) = a \otimes b$, então τ é uma aplicação bilinear e o subespaço gerado por $\tau(A \times B)$ em $A \otimes B$ é o próprio $A \otimes B$.*

Prova: O fato de que τ é uma aplicação bilinear decorre facilmente de alguns cálculos. É fácil verificar também que o subespaço gerado por $\tau(A \times B)$ em $A \otimes B$ é o próprio $A \otimes B$.■

7.2 Produto tensorial de álgebras

Sejam A e B álgebras associativas com identidade sobre um corpo K. O produto cartesiano $A \times B$ é também uma álgebra associativa com identidade sobre K, considerando a soma como

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$

a multiplicação como

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 b_2)$$

e a multiplicação por escalar como

$$\lambda(a, b) = (\lambda a, \lambda b).$$

Teorema 7.2.1 *Se A e B são álgebras associativas com identidade sobre um corpo K, então $A \otimes B$ é uma álgebra associativa com identidade sobre K, com multiplicação induzida por $(a_1 \otimes b_1)(a_2 \otimes b_2) = (a_1 a_2 \otimes b_1 b_2)$.*

Prova: Fixando a_2 e b_2, definimos ψ_1 e ψ_2 por $\psi_1(a_1) = a_1 a_2$ e $\psi_2(b_1) = b_1 b_2$ para quaisquer a_1 em A e b_1 em B. Claramente, ψ_1 e ψ_2 são transformações lineares. Logo, $\psi_1 \otimes \psi_2 : A \otimes B \to A \otimes B$ define multiplicação à direita por $a_2 \otimes b_2$.

Fazendo isto para todo a_2 em A e b_2 em B, revertemos o procedimento, fixando x em $A \otimes B$ e definindo $\psi_x : A \times B \to A \otimes B$ por $\psi_x(a, b) = x(a \otimes b)$ para qualquer (a, b) em $A \times B$. Como ψ_x é bilinear, pela **Proposição 7.1.1**, temos que existe um homomorfismo de grupos $\psi'_x : A \otimes B \to A \otimes B$ tal que $\psi'_x(a \otimes b) = \psi(a, b)$ para todo a em A e b em B. A aplicação ψ'_x corresponde à multiplicação à esquerda por x.

É fácil ver agora que $A \otimes B$ é uma álgebra associativa com identidade. ■

Teorema 7.2.2 *Sejam A, B, A' e B' álgebras associativas com identidade. Se $\psi_1 : A \to A'$ e $\psi_2 : B \to B'$ são homomorfismos de álgebras, então existe um único homomorfismo de álgebras $\psi_1 \otimes \psi_2 : A \otimes B \to A' \otimes B'$ tal que $(\psi_1 \otimes \psi_2)(a \otimes b) = \psi_1(a) \otimes \psi_2(b)$ para todo a em A e b em B.*

Prova: Vamos definir $\psi : A \times B \to A' \otimes B'$ por $\psi(a, b) = \psi_1(a) \otimes \psi_2(b)$. Obviamente, ψ é bilinear. Logo, pela **Proposição 7.1.2**, existe uma única transformação linear $\psi' : A \otimes B \to A' \otimes B'$ tal que $\psi'(a \otimes b) = \psi(a, b) = \psi_1(a) \otimes \psi_2(b)$ para todo a em A e b em B.

Além disso,

$$\begin{aligned}\psi'((a_1\otimes b_1)(a_2\otimes b_2)) &= (\psi_1\otimes\psi_2)((a_1\otimes b_1)(a_2\otimes b_2)) = (\psi_1\otimes\psi_2)(a_1a_2\otimes b_1b_2) = \\ &= \psi_1(a_1a_2)\otimes\psi_2(b_1b_2) = \psi_1(a_1)\psi_1(a_2)\otimes\psi_2(b_1)\psi_2(b_2) = \\ &= (\psi_1(a_1)\otimes\psi_1(a_2))(\psi_2(b_1)\otimes\psi_2(b_2)) = \\ &= (\psi_1\otimes\psi_2(a_1\otimes b_1))(\psi_1\otimes\psi_2(a_2\otimes b_2)) = \\ &= \psi'(a_1\otimes b_1)\psi'(a_2\otimes b_2)).\end{aligned}$$

Assim, é fácil ver que ψ' é um homomorfismo de álgebras. Definindo $\psi_1\otimes\psi_2=\psi'$ temos o homomorfismo desejado. Claramente, $\psi_1\otimes\psi_2$ é único. ■

7.3 Teorema do Produto Tensorial

Provaremos nesta seção o Teorema do Produto Tensorial, o qual afirma que o produto tensorial de duas PI álgebras é uma PI álgebra. A prova deste teorema, feita em 1972 por Regev [29], utiliza codimensões e pode ser encontrada em Rowen [32].

Lema 7.3.1 (Regev) *Se K é um corpo e I e J são T-ideais de $K\{\mathbf{X}\}$ tais que $c_n(I)\cdot c_n(J) < n!$ para algum n, então $(K\{X\}/I)\otimes(K\{X\}/J)$ satisfaz uma identidade polinomial de grau n.*

Prova: Consideremos que o espaço V_n/I_n é gerado por $\{p_i(X_1,...,X_n) : 1\leqslant i\leqslant c_n(I)\}$ e que o espaço V_n/J_n é gerado por $\{q_j(X_1,...,X_n) : 1\leqslant j\leqslant c_n(J)\}$. Para cada $\sigma\in S_n$, podemos encontrar elementos α_i^σ e β_j^σ em K, $1\leqslant i\leqslant c_n(I)$ e $1\leqslant j\leqslant c_n(J)$, tais que

$$X_{\sigma(1)}...X_{\sigma(n)} = \sum_{i=1}^{c_n(I)}\alpha_i^\sigma p_i(X_1,...,X_n) \pmod{I_n}$$

e

$$X_{\sigma(1)}...X_{\sigma(n)} = \sum_{j=1}^{c_n(J)}\beta_j^\sigma q_j(X_1,...,X_n) \pmod{J_n}.$$

Queremos encontrar elementos t_σ em K, $\sigma\in S_n$, tais que o polinômio $f(X_1,...,X_n) = \sum_{\sigma\in S_n} t_\sigma X_{\sigma(1)}...X_{\sigma(n)}$ seja uma identidade polinomial de $(K\{X\}/I)\otimes(K\{X\}/J)$. Tomemos elementos arbitrários $a_1,...,a_n$ de $K\{X\}/I$ e $b_1,...,b_n$ de $K\{X\}/J$. Logo,

$$a_{\sigma(1)}...a_{\sigma(n)} = \sum_{i=1}^{c_n(I)}\alpha_i^\sigma p_i(a_1,...,a_n)$$

e

$$b_{\sigma(1)}...b_{\sigma(n)} = \sum_{j=1}^{c_n(J)}\beta_j^\sigma q_j(b_1,...,b_n).$$

Então

$$\begin{aligned} f(a_1 \otimes b_1, ..., a_n \otimes b_n) &= \sum_{\sigma \in S_n} t_\sigma (a_{\sigma(1)} \otimes b_{\sigma(1)}) ... (a_{\sigma(n)} \otimes b_{\sigma(n)}) = \\ &= \sum_{\sigma \in S_n} t_\sigma (a_{\sigma(1)} ... a_{\sigma(n)}) \otimes (b_{\sigma(1)} ... b_{\sigma(n)}) = \\ &= \sum_{\sigma \in S_n} t_\sigma (\sum_{i=1}^{c_n(I)} \alpha_i^\sigma p_i(a_1, ..., a_n)) \otimes (\sum_{j=1}^{c_n(J)} \beta_j^\sigma q_j(b_1, ..., b_n)) = \\ &= \sum_{\sigma \in S_n} \sum_{i=1}^{c_n(I)} \sum_{j=1}^{c_n(J)} t_\sigma \alpha_i^\sigma p_i(a_1, ..., a_n) \otimes \beta_j^\sigma q_j(b_1, ..., b_n) = \\ &= \sum_{i=1}^{c_n(I)} \sum_{j=1}^{c_n(J)} (\sum_{\sigma \in S_n} t_\sigma \alpha_i^\sigma \beta_j^\sigma) p_i(a_1, ..., a_n) \otimes q_j(b_1, ..., b_n). \end{aligned}$$

Assim, precisamos apenas encontrar elementos t_σ não todos nulos tais que $\sum_{\sigma \in S_n} t_\sigma \alpha_{i,\sigma} \beta_{j,\sigma} = 0$ para $1 \leqslant i \leqslant c_n(I)$ e $1 \leqslant j \leqslant c_n(J)$. Observando que isto é um sistema linear e homogêneo com $c_n(I) \cdot c_n(J)$ equações e $n!$ incógnitas, ele tem solução não trivial pois $c_n(I) \cdot c_n(J) < n!$.■

Teorema 7.3.1 (do Produto Tensorial) *O produto tensorial de duas PI álgebras é uma PI álgebra.*

Prova: Sejam A e B duas PI álgebras sobre um corpo K. Consideremos que A possui uma identidade polinomial de grau g e que B possui uma identidade polinomial de grau h. Notemos que para qualquer k em $\mathbf{N}$, existe $n \in \mathbf{N}$ com $n! > k^n$. Assim pelo **Corolário 6.2.1**, existe n em $\mathbf{N}$ tal que

$$c_n(T(A)) \cdot c_n(T(B)) < ((g-1)^2 \cdot (h-1)^2)^n < n!.$$

Pelo **Lema 7.3.1**, a álgebra $(K\{X\}/T(A)) \otimes (K\{X\}/T(B))$ satisfaz uma identidade polinomial de grau n; consideremos que esta identidade polinomial seja

$$p(X_1, ..., X_n) = \sum_{\sigma \in S_n} \alpha_\sigma X_{\sigma(1)} ... X_{\sigma(n)}.$$

Tomemos elementos arbitrários $a_1, ..., a_n$ de A; definamos $a_i = 0$ para $i > n$. Pelo **Teorema 3.2.1**, $\psi_1 : K\{X\} \to A$ dada por $\psi_1(f(X_{i_1}, ..., X_{i_k})) = f(a_{i_1}, ..., a_{i_k})$ é um homomorfismo de álgebras tal que $\psi_1(X_i) = a_i$ para todo $i \in \mathbf{N}$. Consideremos a aplicação $\varphi_1 : K\{X\}/T(A) \to A$ dada por $\varphi_1(f + T(A)) = \psi_1(f)$. Esta aplicação está bem definida, pois

$$\begin{aligned} f + T(A) &= g + T(A) \Rightarrow f - g \in T(A) \Rightarrow \psi_1(f-g) = 0 \Rightarrow \\ &\Rightarrow \psi_1(f) = \psi_1(g) \Rightarrow \varphi_1(f + T(A)) = \varphi_1(g + T(A)). \end{aligned}$$

É fácil ver que φ_1 é um homomorfismo de álgebras.

Tomando elementos arbitrários $b_1, ..., b_n$ de B, definindo $b_i = 0$ para $i > n$ e procedendo de modo análogo, temos que existe um homomorfismo de álgebras $\varphi_2 : K\{X\}/T(B) \to B$ dado por $\varphi_2(f(X_{i_1}, ..., X_{i_k}) + T(B)) = f(b_{i_1}, ..., b_{i_k})$.

Pelo **Teorema 7.2.1**, existe um homomorfismo de álgebras $\varphi_1 \otimes \varphi_2 : (K\{X\}/T(A)) \otimes (K\{X\}/T(B)) \to A \otimes B$ tal que $(\varphi_1 \otimes \varphi_2)(a \otimes b) = \varphi_1(a) \otimes \varphi_2(b)$ para todo a em A e b em B. Assim, obtemos

$$\begin{aligned}
p(a_1 \otimes b_1, ..., a_n \otimes b_n) &= \sum_{\sigma \in S_n} \alpha_\sigma (a_{\sigma(1)} \otimes b_{\sigma(1)}) ... (a_{\sigma(n)} \otimes b_{\sigma(n)}) = \\
&= \sum_{\sigma \in S_n} \alpha_\sigma (a_{\sigma(1)} ... a_{\sigma(n)} \otimes b_{\sigma(1)} ... b_{\sigma(n)}) = \\
&= \sum_{\sigma \in S_n} \alpha_\sigma (\varphi_1(X_{\sigma(1)} ... X_{\sigma(n)} + T(A)) \otimes \varphi_2(X_{\sigma(1)} ... X_{\sigma(n)} + T(B))) = \\
&= \sum_{\sigma \in S_n} \alpha_\sigma ((\varphi_1 \otimes \varphi_2)((X_{\sigma(1)} ... X_{\sigma(n)} + T(A)) \otimes (X_{\sigma(1)} ... X_{\sigma(n)} + T(B)))) = \\
&= (\varphi_1 \otimes \varphi_2) \sum_{\sigma \in S_n} \alpha_\sigma ((X_{\sigma(1)} ... X_{\sigma(n)} + T(A)) \otimes (X_{\sigma(1)} ... X_{\sigma(n)} + T(B)) = \\
&= (\varphi_1 \otimes \varphi_2)(p(X_1 + T(A) \otimes X_1 + T(B), ..., X_n + T(A) \otimes X_n + T(B))) = \\
&= (\varphi_1 \otimes \varphi_2)(0) = 0
\end{aligned}$$

Portanto, p é uma identidade polinomial de $A \otimes B$. ■

7.4 Outra prova do Teorema de Amitsur-Levitzki

Voltemos a considerar o Teorema de Amitsur-Levitzki. Utilizando o **Teorema 4.2.1**, cuja demonstração é um pouco trabalhosa, a prova de Swan torna-se simples e curta. A prova que Rosset publicou em 1976 também é interessante pela sua simplicidade, principalmente quando comparada com a longa demonstração original proposta por Amitsur e Levitzki em 1950. Apresentaremos a seguir a prova de Rosset, que aparece somente neste capítulo pois utiliza, entre outras ferramentas matemáticas, o produto tensorial. Ela pode ser encontrada também em Pierce [26]. Para podermos desenvolvê-la necessitamos de alguns resultados auxiliares.

Lema 7.4.1 *Sejam K um corpo, n e r dois números naturais e $A_1, ..., A_r$ matrizes de $M_n(K)$. Se r é par, então o traço de $s_r(A_1, ..., A_r)$ é zero.*

Prova: Se A e B são matrizes de $M_n(K)$, é fácil ver que o traço de $AB - BA$ é zero. Seja $\sigma \in S_r$ e consideremos a permutação $\tau = \sigma \circ (123...r) \in S_r$. As paridades de σ e τ são distintas, pois $(123...r)$ é uma permutação ímpar. Assim, se definirmos $A = B_{\sigma(1)}$ e $B = B_{\sigma(2)} ... B_{\sigma(r)}$, teremos $BA = B_{\sigma(2)} ... B_{\sigma(r)} B_{\sigma(1)} = B_{\tau(1)} ... B_{\tau(r)}$. Logo, o traço de $(-1)^\sigma B_{\sigma(1)} ... B_{\sigma(r)} + (-1)^\tau B_{\tau(1)} ... B_{\tau(r)}$ é igual ao traço de $(-1)^\sigma AB + (-1)^\tau BA$ que é zero. Portanto, o traço de $s_r(B_1, ..., B_r)$ é zero. ■

O próximo teorema apresenta um resultado clássico: as Fórmulas de Newton. Sua demonstração não é complicada e pode ser encontrada em Kostrikin [18].

Teorema 7.4.1 (Fórmulas de Newton) *Sejam K um corpo e n e k dois números naturais. Se $k \leqslant n$ então*

$$P_k - P_{k-1}S_1 + P_{k-2}S_2 + ... + (-1)^{k-1}P_1S_{k-1} + (-1)^k kS_k = 0,$$

e se $k > n$ então

$$P_k - P_{k-1}S_1 + P_{k-2}S_2 + ... + (-1)^{n-1}P_{k-n+1}S_{n-1} + (-1)^n P_{k-n}S_n = 0,$$

onde, para cada $i \in N$, P_i é a soma de potências de grau i em n variáveis e S_i é o polinômio simétrico de grau i em n variáveis.

Lema 7.4.2 *Sejam K um corpo de característica zero e E uma álgebra associativa e comutativa sobre K. Se $A \in M_n(E)$ e o traço de A^k é zero para $1 \leqslant k \leqslant n$, então $A^n = 0$.*

Prova: Sejam $\lambda_1, ..., \lambda_n$ as raízes características de A no fecho algébrico do corpo K. Logo, o polinômio característico p de A é dado por

$$\begin{aligned} p(x) &= \det(xI - A) = \\ &= x^n - S_1(\lambda_1, ..., \lambda_n)x^{n-1} + S_2(\lambda_1, ..., \lambda_n)x^{n-2} + ... + (-1)^{n-1}S_{n-1}(\lambda_1, ..., \lambda_n)x + (-1)^n S_n(\lambda_1, ..., \end{aligned}$$

Seja $k \in \{1, ..., n\}$. Como as raízes características de A^k são $\lambda_1^k, ..., \lambda_n^k$ e o traço de A^k é zero, então $p(\lambda_1, ..., \lambda_n) = 0$. Pelas Fórmulas de Newton, já que a característica de K é zero, obtemos que $S_i(\lambda_1, ..., \lambda_n) = 0$ para $1 \leqslant i \leqslant n$. Logo, $p(x) = x^n$, e pelo Teorema de Cayley-Hamilton, temos que $A^n = 0$. ■

Definição 7.4.1 *Sejam R um anel associativo com identidade, a um elemento de R, n um número natural, i e j elementos de $\{1, ..., n\}$. Denotamos por $[a]_{ij}$ a matriz de $M_n(R)$ que tem a na posição (i, j) e zeros nas demais posições.*

Proposição 7.4.1 *Sejam A uma álgebra associativa com identidade de dimensãao n sobre um corpo K e $\{e_1, ..., e_n\}$ uma base de A. A transformação linear $\phi : M_n(A) \to M_n(K) \otimes A$, dada por $\phi([e_k]_{ij}) = [1]_{ij} \otimes e_k$, é um isomorfismo de álgebras.*

Prova: Prova-se com alguns cálculos que ϕ é um homomorfismo de álgebras. Como ϕ é sobrejetora e

$$dim(M_n(A)) = n^2 \cdot dim(A) = dim(M_n(K) \otimes A),$$

temos que ϕ é um isomorfismo de álgebras. ■

Outro instrumento utilizado na prova de Rosset é o conceito de álgebra exterior, que definimos a seguir.

Definição 7.4.2 *Sejam K um corpo e n um número natural. A* álgebra exterior *(ou* álgebra alternante *ou* álgebra de Grassman*) E de um espaço vetorial n dimensional sobre K é definida como*

$$E = K\{\mathbf{X}\}/I,$$

onde I é o ideal de $K\{\mathbf{X}_k\}$ gerado por todos os elementos da forma $(\alpha_i X_1 + ... + \alpha_n X_n)^2$, onde cada $\alpha_i \in K$. Denotamos $1 = 1 + I$, $e_i = X_i + I$ para $1 \leqslant i \leqslant n$ e a multiplicação de dois elementos a e b em E por $a \wedge b$.

A próxima proposição descreve uma série de propriedades das álgebras exteriores.

Proposição 7.4.2 *Sejam K um corpo, n um número natural e E a álgebra exterior de um espaço vetorial V de dimensão n sobre K. Então valem as seguintes propriedades:*
(i) $e_i \wedge e_i = 0$ *para qualquer* i *em* $\{1, ..., n\}$;
(ii) $e_i \wedge e_j = -e_j \wedge e_i$ *para quaisquer* i *e* j *em* $\{1, ..., n\}$;
(iii) $B = \{e_{i_1} \wedge ... \wedge e_{i_k} : 0 \leqslant k \leqslant n \text{ e } 1 \leqslant i_1 < ... < i_k \leqslant n\}$ *é uma base de* E *(por* $k = 0$ *entendemos* $e_{i_1} \wedge ... \wedge e_{i_k} = \mathbf{1}$*)*;
(iv) para cada k *em* $\{0, 1, ..., n\}$, *definimos* E_k *como o subespaço de* E *gerado por* $\{e_{i_1} \wedge ... \wedge e_{i_k} : 1 \leqslant i_1 < ... < i_k \leqslant n\}$. *Então* $E = E_0 \oplus E_1 \oplus ... \oplus E_n$;
(v) para cada k *em* $\{0, 1, ..., n\}$, $\dim E_k = \binom{n}{k}$ *e* $\dim E = 2^n$;
(vi) $E_0 \cong K$ *(como álgebras) e* $E_1 \cong V$ *(como espaços vetoriais);*
(vii) O conjunto $Z(E) = \{a \in E : a \cdot b = b \cdot a \text{ para todo } b \in E\}$, *chamado centro de* E, *pode ser escrito como* $Z(E) = E_0 \oplus E_2 \oplus E_4 \oplus ... \oplus E_m$, *onde* $m = n$ *(se* n *é par) ou* $m = n - 1$ *(se* n *é ímpar).*

Demonstramos a seguir o Teorema de Amitsur e Levitzki pela prova de Rosset.

Prova de Rosset para o Teorema de Amitsur-Levitzki: Como s_{2n} é multilinear, é suficiente mostrar que este polinômio se anula quando calculado em elementos de uma base de $M_n(K)$. Se escolhermos a base usual, formada pelas matrizes e_{ij}, nas quais a posição ij é igual a 1 e as demais posições são nulas, vemos que basta provar o teorema para o subcorpo primo de K.Por sua vez, se a característica de K for p então o subcorpo primo de K é isomorfo a $\mathbf{Q}$ (caso $p = 0$) ou a $\mathbf{Z}_p$ (caso p seja um número primo). Como $\mathbf{Z}_p$ é isomorfo a um subanel de $\mathbf{Q}$, precisamos provar o teorema apenas quando o subcorpo primo é isomorfo a $\mathbf{Q}$. Assim, podemos considerar que a característica de K é zero.

Sejam $A_1, ..., A_{2n}$ matrizes de $M_n(K)$ e E a álgebra exterior de um espaço vetorial $2n$ dimensional sobre K.

Se

$$a = A_1 \otimes e_1 + A_2 \otimes e_2 + ... + A_{2n} \otimes e_{2n} \in M_n(K) \otimes E,$$

então para todo $k \geqslant 1$,

$$a^k = \sum_{1 \leqslant i_1 < ... < i_k \leqslant 2n} s_k(A_{i_1}, ..., A_{i_k}) \otimes e_{i_1} \wedge ... \wedge e_{i_k},$$

e em particular,

$$a^{2n} = s_{2n}(A_1, ..., A_{2n}) \otimes e_1 \wedge ... \wedge e_{2n}.$$

Se provarmos que $a^{2n} = 0$, a prova está concluída, pois $e_1 \wedge ... \wedge e_{2n}$ é um elemento de uma base de E e, pelo **Teorema 7.1.3**, temos que $s_{2n}(A_1, ..., A_{2n}) = 0$.

Consideremos a transformação linear $\phi : M_n(E) \to M_n(K) \otimes E$ dada por $\phi([e_{i_1} \wedge ... \wedge e_{i_k}]_{ij}) = [1]_{ij} \otimes e_{i_1} \wedge ... \wedge e_{i_k}$ para quaisquer i e j em $\{1, ..., n\}$, $0 \leqslant k \leqslant n$ e $1 \leqslant i_1 < ... < i_k \leqslant 2n$. Pela **Proposição 7.4.1**, ϕ é um isomorfismo de álgebras. Sejam $k \in \{1, ..., 2n\}$ e $A \in M_n(E)$ tal que $\phi(A) = a$. Assim,

$$tr(A^{2k}) = \sum_{1 \leqslant i_1 < ... < i_k \leqslant 2n} tr(s_{2k}(A_{i_1}, ..., A_{i_k}))e_{i_1} \wedge ... \wedge e_{i_k}.$$

Mas pelo **Lema 7.4.1**, temos que o traço de $s_{2k}(A_{i_1}, ..., A_{i_k})$ é zero para quaisquer $1 \leqslant i_1 < ... < i_k \leqslant 2n$. Portanto, o traço de A^{2k} é zero.

Observando que $A^2 \in M_n(Z(E))$ e que o traço de $(A^2)^k$ é zero para $1 \leqslant k \leqslant n$, concluímos, pelo **Lema 7.4.2**, que $A^{2n} = 0$. Portanto, $a^{2n} = 0$. ■

É interessante observarmos que toda álgebra exterior é uma PI álgebra.

Proposição 7.4.3 *Toda álgebra exterior satisfaz a identidade polinomial* $[[X_1, X_2], X_3]$.

Prova: Sejam K um corpo, n um número natural e E a álgebra exterior de um espaço vetorial n dimensional sobre K. Como $[[X_1, X_2], X_3]$ é multilinear, é suficiente mostrar que este polinômio se anula quando calculado em elementos de uma base de E. Consideremos a base $B = \{e_{i_1} \wedge ... \wedge e_{i_k} : 0 \leqslant k \leqslant n \text{ e } 1 \leqslant i_1 < ... < i_k \leqslant n\}$ de E. Sejam a e b elementos de B. Se provarmos que $[a, b] \in Z(E)$, a prova está concluída. Sejam a' e b' elementos de $B \cap Z(E)$ tais que $a = a' \wedge e_a$ e $b = b' \wedge e_b$ onde e_a e e_b pertencem a $\{1, e_1, ..., e_n\}$. Observemos que $[e_a, e_b] \in Z(E)$. Assim,

$$\begin{aligned}
[a, b] &= a \wedge b - b \wedge a = (a' \wedge e_a) \wedge (b' \wedge e_b) - (b' \wedge e_b) \wedge (a' \wedge e_a) = \\
&= a' \wedge b' \wedge e_a \wedge e_b - a' \wedge b' \wedge e_b \wedge e_a = \\
&= a' \wedge b' \wedge (e_a \wedge e_b - e_b \wedge e_a) = a' \wedge b' \wedge [e_a, e_b],
\end{aligned}$$

e portanto, $[a, b] \in Z(E)$. ■

Capítulo 8

Teorema da Altura

Os primeiros resultados deste capítulo apareceram pela primeira vez em dois artigos de A. I. Shirshov de 1957 (ver [33] e [34]). Eles também podem ser encontrados em [39]. O Teorema da Altura é um dos principais resultados na teoria combinatória das álgebras com identidades polinomiais. Este teorema abrange e generaliza vários resultados obtidos por Jacobson, Kaplansky e Levitzki, entre outros. O Teorema possui aplicações no caso de álgebras não-associativas (ver [39] para mais detalhes).

8.1 Problema de Kurosh

Definição 8.1.1 *Seja A uma álgebra associativa sobre um corpo K. Um elemento a de A é chamado* algébrico *se existe um número natural n tal que $a^n = \sum_{i=1}^{n-1} \alpha_i a^i$, onde $\alpha_i \in K$. O menor número n para o qual isso ocorre é chamado o* índice algébrico *do elemento a. A álgebra A é chamada* algébrica *se todos os seus elementos são algébricos. Além disso, se os índices algébricos dos elementos da álgebra A são limitados, então a álgebra A é chamada uma* álgebra algébrica de índice limitado.

Exemplo 8.1.1 *Seja A uma álgebra associativa n dimensional sobre um corpo K. Para cada elemento a de A existem elementos $\alpha_1, ..., \alpha_n, \alpha_{n+1}$ em K tais que $\sum_{i=1}^{n+1} \alpha_i a^i = 0$. Assim, é fácil ver que A é uma álgebra algébrica de índice limitado por $n+1$.*

Proposição 8.1.1 *Toda álgebra algébrica de índice limitado é uma PI álgebra.*

Prova: 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 $[a^{n-1}, b], ..., [a, b]$ com coeficientes em K, pois $a^n = \sum_{i=1}^{n-1} \alpha_i a^i$, onde $\alpha_i \in K$.Logo, os elementos $[a^n, b], ..., [a, b]$ reduzem o polinômio standard $s_n(X_1, ..., X_n)$ a zero. Conseqüentemente

$$s_n([X_1^n, X_2], [X_1^{n-1}, X_2], ..., [X_1, X_2])$$

é uma identidade polinomial da álgebra A. ■

Proposição 8.1.2 *Seja A uma álgebra. A interseção de uma coleção arbitrária de subálgebras de A é uma subálgebra de A.*

Definição 8.1.2 *Seja S um subconjunto de uma álgebra A. A* subálgebra gerada por S *é definida como sendo a interseção de todas as subálgebras de A que contém S. Se S é um conjunto finito dizemos que a álgebra gerada por S é uma* álgebra finitamente gerada.

Observação 8.1.1 *Toda nil álgebra é uma álgebra algébrica. Toda nil álgebra de índice limitado é uma álgebra algébrica de índice limitado.*

Definição 8.1.3 *Uma álgebra A é* localmente finita *se cada subálgebra finitamente gerada de A tem dimensão finita.*

Em 1941, A. G. Kurosh [19] formulou a seguinte pergunta:

Problema de Kurosh: *Toda álgebra algébrica é localmente finita?*

Este problema contribuiu muito para o desenvolvimento da teoria das PI álgebras. Em 1945, N. Jacobson [12] encontrou uma resposta positiva para álgebras algébricas de índice limitado. Em 1946, J. Levitzki [23] provou que toda PI álgebra finitamente gerada e que é nil álgebra é nilpotente. Finalmente, em 1948 I. Kaplansky [14] ampliou o resultado obtido por Jacobson resolvendo o Problema de Kurosh para PI álgebras.

Os métodos utilizados por Kaplansky e Levitzki são completamente diferentes. No entanto, A. I. Shirshov [34], em 1957, abordando o Problema de Kurosh do ponto de vista combinatório, provou o Teorema da Altura e mostrou que os teoremas de Kaplansky e Levitzki são alguns de seus corolários.

Em 1964, E. S. Golod e I. R. Shafarevich (ver [7] e [6], ou [9]) provaram que existem nil álgebras finitamente geradas que não são nilpotentes. Pelo **Teorema 8.1.1**, vemos que isto é uma resposta negativa para o Problema de Kurosh.

Teorema 8.1.1 *Toda nil álgebra de dimensão finita é nilpotente.*

Prova: Observemos, em primeiro lugar, 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, $1 \leqslant i \leqslant 2$.

Seja A uma nil álgebra de dimensão finita n. Se $a \in A$, é fácil ver que $Aa = \{xa : x \in A\}$ é um ideal à esquerda de A. Além disso, para a transformação linear $T_a : A \to Aa$ dada por $T_a(x) = xa$, temos que $KerT_a \neq \{0\}$, pois $(a^{k-1}) \cdot a = 0$ e $a^{k-1} \neq 0$ para algum número natural k maior que 1.

Provemos o teorema por indução finita em n. Se $n = 1$, então $KerT_a = A$ para todo $a \in A$, ou seja, $Aa = \{0\}$ para todo $a \in A$; logo, A é nilpotente. Se $n > 1$, então $Aa \neq A$ (pois $KerT_a \neq \{0\}$) e, conseqüentemente, $\dim(Aa) < n$; por hipótese de indução, Aa é nilpotente para todo $a \in A$. Basta provarmos agora que existem elementos $a_1, ..., a_m$ em A tais que $Aa \subseteq Aa_1 + ... + Aa_m$ para todo $a \in A$; pois neste caso teremos $A^2 \subseteq Aa_1 + ... + Aa_m$ e, além disso, como $Aa_1 + ... + Aa_m$ é nilpotente, as álgebras A^2 e A serão nilpotentes. 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 $Aa_3 \subseteq Aa_1 + Aa_2 + Aa_3$. Procedendo desta forma, obteremos elementos $a_1, ..., a_m$ em A tais que $Aa \subseteq Aa_1 + ... + Aa_m$ para todo $a \in A$, pois $\dim(Aa_1 + ... + Aa_i) < \dim(Aa_1 + ... + Aa_{i+1})$. ■

8.2 Lema de Shirshov

Definição 8.2.1 *Definiremos o alfabeto* $\mathbf{X}_n$ *como o subconjunto* $\{X_1, ..., X_n\}$ *de* $\mathbf{X}$, *para cada* $n \in \mathbf{N}$.

Definição 8.2.2 *Uma palavra* w *sobre* $\mathbf{X}_n$. *é chamada:*
(i) X_n-indecomponível *se tem a forma* $w = X_n...X_nX_{i_1}...X_{i_s}$, *onde* $s \geqslant 1$ *e* $i_t \neq n$ *para* $t = 1, ..., s$;
(ii) X_n-fatorável *se tem a forma* $w = w_1...w_n$, *onde cada* w_i *é uma palavra* X_n*-indecomponível; dizemos que* $w_1...w_n$ *é uma* X_n-fatoração *de* w.

Exemplo 8.2.1 *As palavras* $X_3X_3X_3X_1X_2$, X_3X_2 *e* $X_3X_3X_1$ *são* X_3*-indecomponíveis. Por outro lado, as palavras* $X_2X_3X_3X_1X_2$, X_3 *e* $X_3X_2X_3$ *não são* X_3*-indecomponíveis. A palavra* $w = X_3X_3X_2X_3X_1$ *é* X_3*-fatorável e* $(X_3X_3X_2)(X_3X_1)$ *é uma* X_3*-fatoração para* w.

Observação 8.2.1 *Uma palavra sobre* $\mathbf{X}_n$ *tem uma* X_n*-fatoração se, e somente se, ela começa com o símbolo* X_n *e termina com um símbolo diferente de* X_n. *Além disso, é fácil ver que uma palavra* X_n*-fatorável possui uma única* X_n*-fatoração.*

Definição 8.2.3 *O conjunto das palavras sobre* $\mathbf{X}_n$ *que são* X_n*-indecomponíveis.é denotado por* T_n.

Observação 8.2.2 *Se considerarmos* T_n *como um novo alfabeto, então o conjunto das palavras sobre* T_n *é o conjunto de todas as palavras* X_n*-fatoráveis.*

Na próxima definição utilizaremos a relação de ordem $\leqslant$ da **Definição 6.6.1**.

Definição 8.2.4 *Uma palavra* w *sobre* $\mathbf{X}$ *é* k-particionável *se* $w = w_1...w_k$, *onde cada* w_i *é uma palavra sobre* $\mathbf{X}$, *e* $w_{\sigma(1)}...w_{\sigma(k)} < w$, *para cada* $\sigma \in S_k - \{1\}$. *Dizemos que* $w_1...w_k$ *é uma* k-partição *de* w.

Exemplo 8.2.2 *A palavra* $X_3X_1X_2X_2X_1X_1X_2X_1X_1X_1$ *é 3-particionável e permite várias 3-partições como, por exemplo,*

$$(X_3X_1)(X_2X_2X_1X_1)(X_2X_1X_1X_1),$$
$$(X_3X_1X_2)(X_2X_1X_1)(X_2X_1X_1X_1),$$
$$(X_3)(X_1X_2X_2X_1X_1X_2)(X_1X_1X_1).$$

Mas a palavra $X_1X_2X_1X_3X_2X_1X_2X_3X_2$ *não é 2-particionável.*

Definição 8.2.5 *Definimos uma relação de ordem* $\trianglelefteq$ *no conjunto* T_n *como sendo:*
escrevemos $X_nX_{i_1}...X_{i_l} \trianglelefteq X_nX_{j_1}...X_{j_p}$ *se um dos três casos ocorre:*
(i) $X_nX_{i_1}...X_{i_l} = X_nX_{j_1}...X_{j_p}$, *isto é,* $l = p$ *e* $i_t = j_t$ *para* $1 \leqslant t \leqslant l$;
(ii) para algum $r \in \{1, ..., l\}$ *temos que* $i_t = j_t$ *se* $1 \leqslant t \leqslant r-1$ *e* $i_r < j_r$;
(iii) $p < l$ *e* $i_t = j_t$ *para* $1 \leqslant t \leqslant p$.
Escrevemos $X_nX_{i_1}...X_{i_l} \triangleleft X_nX_{j_1}...X_{j_p}$ *se* $X_nX_{i_1}...X_{i_l} \trianglelefteq X_nX_{j_1}...X_{j_p}$ *e* $X_nX_{i_1}...X_{i_l} \neq X_nX_{j_1}...X_{j_p}$.

Verifica-se facilmente que $(T_n, \leqslant)$ é totalmente ordenado.

Observação 8.2.3 *Pela definição anterior, se w_1 e w_2 são elementos de T_n e w_1 é um segmento inicial de w_2 então $w_2 \trianglelefteq w_1$. O caso (iii) da definição da relação $\trianglelefteq$ é justamente o contrário do caso (iii) da definição da relação $\preccurlyeq$.*

Definição 8.2.6 *Definimos uma relação de ordem $\lll$ no conjunto das palavras sobre T_n como sendo:*
se v e w são palavras sobre T_n e $v = v_1...v_l$ e $w = w_1...w_p$ são suas respectivas X_n-fatorações, escrevemos $v \lll w$ se um dos três casos ocorre:
(i) $v = w$;
(ii) para algum $r \in \{1, ..., l\}$ temos que $v_t = w_t$ se $1 \leqslant t \leqslant r-1$ e $v_r \triangleleft w_r$;
(iii) $l < p$ e $v_t = w_t$ para $1 \leqslant t \leqslant l$.
Escrevemos $v \ll w$ se $v \lll w$ e $v \neq w$.

Verifica-se facilmente que $(T_n^*, \lll)$ é totalmente ordenado.

Observação 8.2.4 *A relação $\lll$ é na verdade a ordem lexicográfica para as palavras sobre o alfabeto T_n com a relação de ordem $\trianglelefteq$.*

Exemplo 8.2.3 *Para as palavras X_3X_1, $X_3X_1X_3X_1$, $X_3X_1X_3X_2$, $X_3X_1X_3X_2X_1$, temos que $X_3X_1X_3X_2X_1 \lll X_3X_1X_3X_2$, $X_3X_1X_3X_1 \lll X_3X_1X_3X_2X_1$ e $X_3X_1 \lll X_3X_1X_3X_2$.*

Definição 8.2.7 *Uma palavra w sobre T_n é k_T-particionável se $w = w_1...w_k$, onde cada w_i é uma palavra sobre T_n, e $w_{\sigma(1)}...w_{\sigma(k)} \ll w$, para cada $\sigma \in S_k - \{1\}$. Dizemos que $w_1...w_k$ é uma k_T-partição de w.*

Exemplo 8.2.4 *A palavra $w = X_3X_3X_2X_3X_1X_3X_2X_1$ é 3_T-particionável com a 3_T-partição*

$$(X_3X_3X_2)(X_3X_1)(X_3X_2X_1).$$

Além disso, w é 2_T-particionável, permitindo duas 2_T-partições:

$$(X_3X_3X_2)(X_3X_1X_3X_2X_1)$$

e

$$(X_3X_3X_2X_3X_1)(X_3X_2X_1).$$

Mas a palavra $X_3X_3X_1X_3X_3X_3X_2X_3X_3X_2$ não é 2_T-particionável.

Lema 8.2.1 *Se uma palavra w sobre T_n é k_T-particionável, então w é k-particionável.*

Prova: Seja $w = w_1...w_k$ uma k_T-partição da palavra w. Segue da definição de k_T-partição que $w_{\sigma(1)}...w_{\sigma(k)} \ll w$ para qualquer σ em S_k distinto da identidade. É fácil ver que $w_{\sigma(1)}...w_{\sigma(k)} < w$. Portanto, $w_1...w_k$ é uma k-partição de w. ■

Lema 8.2.2 *Se uma palavra w sobre T_n é $(k-1)_T$-particionável, então a palavra wX_n é k-particionável.*

Prova: Do **Lema 8.2.1**, segue a existência da seguinte $(k-1)$-partição da palavra w:

$$w = (X_n X_{i_1} ... X_{j_1})(X_n X_{i_2} ... X_{j_2}) ... (X_n X_{i_{k-1}} ... X_{j_{k-1}}),$$

onde $j_t \neq n$ para $t = 1, 2, ..., k-1$.

Vamos provar que a seguinte k-partição é válida para a palavra wX_n:

$$w = (X_n)(X_{i_1} ... X_{j_1} X_n)(X_{i_2} ... X_{j_2} X_n) ... (X_{i_{k-1}} ... X_{j_{k-1}} X_n).$$

Qualquer permutação de subpalavras da palavra wX_n que retém o símbolo X_n na primeira posição transforma a palavra wX_n numa palavra $w'X_n$, onde w' é obtida por meio de alguma permutação das subpalavras da $(k-1)$-partição referida anteriormente. Logo $w' < w$, e $w'X_n < wX_n$.

Se considerarmos permutações que retiram o símbolo X_n da primeira posição, então as palavras obtidas desta maneira começarão com um número estritamente menor de símbolos X_n em comparação com a palavra wX_n. Portanto, elas serão estritamente menores que a palavra wX_n. ■

O próximo lema foi provado em 1957 por Shirshov [33] e é de grande utilidade para demonstrarmos o Teorema da Altura.

Lema 8.2.3 (Shirshov) *Para quaisquer três números naturais n, s, k, pode-se encontrar um número natural $N(n, s, k)$ tal que qualquer palavra sobre $\mathbf{X}_n$ com comprimento $N(n, s, k)$ contém s subpalavras iguais consecutivas ou uma subpalavra k-particionável.*

Prova: A prova será feita por indução finita em k e em n. Se $k = 1$, é fácil ver que o número $N(n, s, 1)$ existe para quaisquer n e s. Suponhamos que o número $N(n, s, k-1)$ exista para quaisquer n e s, e vamos provar que $N(n, s, k)$ existe.

Notemos que o número $N(1, s, k)$ também existe, e isto nos dá uma base para uma indução em n. Assim, supomos também que o número $N(n-1, s, k)$ existe e vamos mostrar que $N(n, s, k)$ existe.

Consideremos uma palavra qualquer w de comprimento

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

Se no início da palavra w há algum número de símbolos X_i diferentes do símbolo X_n, e esse número não é menor que o número $N(n-1, s, k)$, então a hipótese de indução é satisfeita com relação à subpalavra w' que está no início da palavra w e contém apenas símbolos de $\mathbf{X}_{n-1}$. Logo, podemos assumir que o comprimento da palavra w' é menor que $N(n-1, s, k)$.

Se no final da palavra w há uma subpalavra $w'' = X_n X_n ... X_n$, podemos assumir que o comprimento de w'' é menor que s, pois caso contrário o lema está provado.

Desconsiderando, caso existam, as palavras w' e w'', obtemos uma subpalavra w_1 cujo comprimento é maior que o número

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

Para a palavra w_1 existe uma X_n-fatoração $w_1 = w_{11} ... w_{1m}$. Podemos assumir que o comprimento de cada palavra X_n-indecomponível w_{1i} é menor que o número $s + N(n-1, s, k)$,

pois caso contrário tal palavra conteria s símbolos X_n consecutivos ou uma subpalavra de comprimento $N(n-1,s,k)$ que não contém o símbolo X_n. Considerando esse limite para o comprimento, é fácil ver que não existem mais que $n^{N(n-1,s,k)+s}$ palavras X_n-indecomponíveis diferentes. Como $m > N(n^{N(n-1,s,k)+s}, s, k-1)$, temos que a palavra w_1, considerada como uma palavra sobre T_n, tem comprimento maior que $N(n^{N(n-1,s,k)+s}, s, k-1)$, e aplicando a hipótese de indução concluímos que existem s subpalavras iguais consecutivas ou uma subpalavra v $(k-1)_T$-particionável. Se a segunda conclusão ocorre, então como a palavra w_1, considerada como uma palavra sobre T_n, tem comprimento estritamente maior que $N(n^{N(n-1,s,k)+s}, s, k-1)$, podemos considerar que o símbolo X_n segue a subpalavra v. Pelo **Lema 8.2.2**, a subpalavra vX_n é n-particionável. Assim, o número

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

prova o lema. ■

Lema 8.2.4 *Seja w uma palavra sobre $\mathbf{X}$ de comprimento m que não é representável na forma v^t, onde v é uma subpalavra própria da palavra w. Então para qualquer número natural $k \leqslant m$, a palavra w^{2k} contém uma subpalavra k-particionável.*

Prova: É possível obter a partir da palavra w, por meio de permutações cíclicas das letras, m palavras $w = w_0, w_1, ..., w_{m-1}$. Desde que a palavra w não é representável na forma de uma potência de uma subpalavra, é fácil ver que todas as palavras obtidas $w_0, w_1, ..., w_{m-1}$ são diferentes. Vamos assumir que $w_{i_0} > w_{i_1} > ... > w_{i_{m-1}}$. É óbvio que cada uma das palavras w_i pode ser representada na forma $w_i = u_i v_i$, onde $v_i u_i = w$. Consideremos agora a palavra

$$w^{2k} = v_{i_0}u_{i_0}v_{i_0}u_{i_0}v_{i_1}u_{i_1}v_{i_1}u_{i_1}...v_{i_{k-1}}u_{i_{k-1}}v_{i_{k-1}}u_{i_{k-1}}.$$

Vamos definir $w'_{i_t} = u_{i_t}v_{i_t}u_{i_t}v_{i_{t+1}}$ para $t = 0, 1, ..., k-2$, e $w'_{i_{k-1}} = u_{i_{k-1}}v_{i_{k-1}}u_{i_{k-1}}$. Então a palavra w^{2k} pode ser representada na forma $w^{2k} = v_{i_0}w'_{i_0}w'_{i_1}...w'_{i_{k-1}}$. Como o início de cada uma das palavras w'_j coincide com a palavra w_j e $w_{i_0} > w_{i_1} > ... > w_{i_{m-1}}$, é fácil ver que a palavra $w'_{i_0}w'_{i_1}...w'_{i_{k-1}}$ é k-particionável. ■

8.3 Teorema da Altura

Seja K um corpo. Assim como $K\{\mathbf{X}\}$ é o conjunto de todos os polinômios com coeficientes no corpo K e indeterminadas em $\mathbf{X}$, vamos definir $K\{\mathbf{X}_n\}$ como o subconjunto de $K\{\mathbf{X}\}$ formado pelos polinômios com indeterminadas apenas no conjunto $\mathbf{X}_n$. A fim de definir $K\{\mathbf{X}_n\}$ com mais precisão e destacar a sua estrutura de álgebra, lembremos que $\mathbf{X}_n^*$, o conjunto das palavras sobre $\mathbf{X}_n$, é um monóide e que pelo **Teorema 2.1.1**, $K(\mathbf{X}_n^*)$ é uma álgebra associativa com identidade sobre K. Assim faz sentido definirmos como se segue.

Definição 8.3.1 *Se K é um corpo, definimos a álgebra associativa com identidade $K\{\mathbf{X}_n\} = K(\mathbf{X}_n^*)$.*

Definição 8.3.2 *Sejam Σ um alfabeto e A um conjunto de palavras sobre Σ. Seja w uma palavra sobre Σ tal que $w = w_1^{n_1}...w_h^{n_h}$, onde cada $w_i \in A$ para $i \in \{1, ..., h\}$ e $w_i \neq w_{i+1}$ para $i \in \{1, ..., h-1\}$. O menor número h com esta propriedade é chamado a* altura de w em relação ao conjunto A.

Exemplo 8.3.1 *Consideremos a palavra* $w = X_1X_1X_2X_1X_1X_2X_1X_1X_2$ *sobre* $\mathbf{X}$. *É fácil ver que* w *tem altura* 6 *em relação a* $\{X_1, X_2\}$ *e altura* 1 *em relação a* $\{X_1X_1X_2\}$.

Definição 8.3.3 *Sejam* A *uma álgebra associativa sobre um corpo* K *gerada por um conjunto finito* $G = \{a_1, ..., a_k\}$ *e* $P = \{p_1, ..., p_l\}$ *um conjunto finito de polinômios homogêneos de* $K\{\mathbf{X}_k\}$. *Para cada polinômio* $p = p(X_1, ..., X_k)$ *de* $K\{\mathbf{X}_k\}$, *denotamos* $\overline{p} = p(a_1, ..., a_k)$. *Se existe um número* h *com a seguinte propriedade: para cada monômio* m *de* $K\{\mathbf{X}_k\}$ *existem monômios* $u_1, ..., u_n$ *em* $K\{\mathbf{X}_k\}$ *tais que*
(i) $\overline{m} = \sum_{i=1}^{n} u_i(\overline{p}_1, \overline{p}_2, ..., \overline{p}_l)$,
(ii) a altura de cada u_i *em relação a* $\mathbf{X}_k$ *é menor ou igual a* h,
(iii) cada $u_i(p_1, ..., p_l)$ *tem o mesmo tipo que* m,
dizemos que A *é uma* álgebra de altura limitada em relação a (G, P). *O menor número* h *com esta propriedade é chamado a* altura da álgebra A em relação a (G, P).

Recordemos que o tipo de um monômio foi definido na **Definição 3.4.2**.

Exemplo 8.3.2 *Seja* A *uma álgebra associativa e comutativa gerada pelo conjunto* $G = \{a_1, ..., a_k\}$. *Se* m *é um monômio de* $K\{\mathbf{X}_k\}$ *do tipo* $[i_1, ..., i_k]$, *então* $\overline{m} = \alpha a_1^{i_1}...a_k^{i_k}$, *para algum* α *em* K. *Portanto, a altura de qualquer monômio de* $K\{\mathbf{X}_k\}$ *em relação a* $(G, \mathbf{X}_k)$ *não excede* k. *Assim, toda álgebra associativa e comutativa finitamente gerada é uma álgebra de altura limitada.*

Apresentamos a seguir o importante Teorema da Altura, que foi demonstrado em 1957 por Shirshov [34].

Teorema 8.3.1 (da Altura) *Seja* A *uma álgebra associativa sobre um corpo* K *gerada por um conjunto finito* $G = \{a_1, ..., a_k\}$ *e que satisfaz uma identidade polinomial de grau* n. *Se* W *é o conjunto de todas as palavras sobre* $\mathbf{X}_k$ *de comprimento menor que* n *e* P *é o conjunto de todos os monômios da forma* $1w$ *onde* $w \in W$, *então a álgebra* A *tem altura limitada em relação a* (G, P).

Prova: Pelo **Corolário 3.5.1**, podemos assumir que A satisfaz uma identidade polinomial da forma

$$X_1...X_n = \sum_{\substack{\sigma \in S_n \\ \sigma \neq id}} \alpha_\sigma X_{\sigma(1)}...X_{\sigma(n)}.$$

Suponhamos que exista um número $M = M(n, k)$ tal que cada palavra w sobre $\mathbf{X}_k$ cuja altura em relação a W é maior ou igual a M contém uma subpalavra n-particionável. Se isso ocorre a prova está concluída, conforme argumentaremos. Na álgebra A temos que

$$\overline{w} = \sum_i \alpha_i \overline{w}_i,$$

onde $\alpha_i \in K$ e w_i é uma palavra sobre $\mathbf{X}_k$ tal que cada indeterminada de $\mathbf{X}_k$ ocorre em w e em w_i o mesmo número de vezes e, além disso, $w_i < w$. Se alguma das palavras w_i

tem altura maior ou igual a M em relação a W, então continuamos o processo. Devido à monotonicidade estrita, teremos por fim que

$$\overline{w} = \sum_i \alpha_i \overline{w}_i,$$

onde cada uma das palavras w_i tem altura menor que M em relação a W.

De acordo com o **Lema 8.2.3**, existe um número $N = N(k, 2n, n)$ tal que cada palavra w sobre $\mathbf{X}_k$ com comprimento N, que não contém subpalavras n-particionáveis, contém uma subpalavra da forma v^{2n}. Além disso, é possível assumir que a palavra v não pode ser representada na forma $(v_1)^k$ com $k > 1$. Então pelo **Lema 8.2.4**, o comprimento $l(v)$ da palavra v é menor que n.

Vamos provar que cada palavra w sobre $\mathbf{X}_k$ com altura maior que $N + 2$ em relação a W, e que não contém subpalabvras n-particionáveis, contém uma subpalavra v_1 da forma $v_1 = v^n v'$, onde $0 < l(v') \leqslant l(v) < n$ e a palavra v' não é um segmento inicial da palavra v. Se w tem altura maior que $N + 2$ em relação a W, então w tem comprimento maior que $N + 2$. O segmento inicial de comprimento N contém uma subpalavra u^n, onde $l(u) < n$; logo, $w = w_o u^n w_0'$, onde w_0' tem altura maior que 2. Podemos escrever $w_o' = u^k u' w_0''$ e $u = u'u''$ (onde u'' não é a palavra vazia, embora u' possa sê-la), onde w_0'' não começa com um segmento inicial de u'' e tem altura maior que 0 (ou seja, não é a palavra vazia). Então

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

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

Desde que o conjunto de subpalavras da forma indicada é finito, para um número natural M suficientemente grande cada palavra com altura maior ou igual a M em relação a W, que não contém subpalavras n-particionáveis, contém n subpalavras iguais da forma $v_1 = v^n v'$ (não necessariamente consecutivas). Em cada palavra podemos encontrar uma subpalavra que é n-particionável procedendo da seguinte maneira:

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

ou

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

Conseqüentemente, cada palavra cuja altura em relação a W é maior que M contém uma subpalavra n-particionável. ■

Utilizando o Teorema da Altura, podemos provar que os teoremas de Levitzki [23] e Kaplansky [14] são simples corolários de um mesmo resultado.

Corolário 8.3.1 (Levitzki) *Toda PI álgebra finitamente gerada e que é nil álgebra é nilpotente.*

Prova: Sejam A uma nil álgebra associativa sobre um corpo K gerada por um conjunto finito $G = \{a_1, ..., a_k\}$ e que satisfaz uma identidade polinomial de grau n, W o conjunto de todas as palavras sobre $\mathbf{X}_k$ de comprimento menor que n e P o conjunto de todos os monômios da forma $1w$ onde $w \in W$. Pelo Teorema da Altura, a álgebra A tem altura limitada em relação a (G, P). Seja h esta 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 que $(m-1) \cdot n \cdot h$ geradores.

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

$$a = \sum_{i=1}^{p} \alpha_i a_{i_1}^{m_{i_1}} ... a_{i_l}^{m_{i_l}},$$

onde cada $i_l < h$, cada $a_{i_j} \in V$ e cada $\alpha_i \in K$. Como $(m-1) \cdot n \cdot h < M$ e se $i \in \{1, ..., p\}$ então $M < (m_{i_1} + ... + m_{i_l}) \cdot n$, podemos concluir que existe um j em $\{1, ..., i_l\}$ tal que $m_{i_j} > m - 1$. Logo, $a = 0$. É fácil ver agora que todo produto de quaisquer M elementos de A é igual a zero. Portanto, A é uma álgebra nilpotente. ■

Corolário 8.3.2 (Kaplansky) *Toda PI álgebra finitamente gerada e algébrica tem dimensão finita.*

Prova: Sejam A uma álgebra associativa e algébrica sobre um corpo K gerada por um conjunto finito $G = \{a_1, ..., a_k\}$ e que satisfaz uma identidade polinomial de grau n, W o conjunto de todas as palavras sobre $\mathbf{X}_k$ de comprimento menor que n e P o conjunto de todos os monômios da forma $1w$ onde $w \in W$. Pelo Teorema da Altura, a álgebra A tem altura limitada em relação a (G, P). Seja h esta altura.

Sejam V o conjunto de todos os produtos com menos que n geradores da álgebra 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 que $(m-1) \cdot n \cdot h$ geradores. ■

Definição 8.3.4 *Uma álgebra A é* localmente nilpotente *se cada subálgebra finitamente gerada de A é nilpotente.*

Corolário 8.3.3 *Toda nil álgebra de índice limitado é locamente nilpotente.*

Definição 8.3.5 *Uma álgebra A tem* altura localmente limitada *se cada subálgebra finitamente gerada de A tem altura limitada.*

Corolário 8.3.4 *Toda PI álgebra tem altura localmete limitada.*

Capítulo 9

Nilpotência de Álgebras Associativas

Estudamos a questão da nilpotência de nil álgebras de índice limitado. Apresentamos o Teorema de Nagata-Higman que é um dos resultados mais importantes neste estudo.

9.1 Introdução

Ao provarmos, como um dos corolários do Teorema da Altura, que *toda nil álgebra de índice limitado é localmente nilpotente*, surge uma pergunta: podemos conseguir um resultado mais forte provando que *toda nil álgebra de índice limitado é nilpotente*? Estudando esta questão, M. Nagata [25] publicou um artigo em 1952. Posteriormente, num artigo de 1956, G. Higman [11] aprofundou os resultados obtidos por Nagata. A síntese destes dois artigos ficou conhecida como Teorema de Nagata-Higman. Um de seus corolários afirma que *toda nil álgebra de índice limitado n e sem elementos de ordem aditiva menor ou igual a n é nilpotente.* Assim, impondo uma condição sobre o grupo aditivo destas nil álgebras, obtemos uma resposta positiva para a nossa pergunta inicial. Estes resultados podem ser encontrados também em [39].

Uma curiosidade histórica é que o resultado de Nagata e Higman foi obtido em 1943, de forma diferente, pelos russos J. Dubnov e V. Ivanov [5]. No entanto, devido à Segunda Guerra Mundial, o trabalho deles ficou desconhecido até por volta de 1970.

Definição 9.1.1 *Se A é uma álgebra sobre um corpo K e n é um número natural, definimos os seguintes conjuntos*

$$I_n(A) = \{\sum_{i=1}^{k} \alpha_i (a_i)^n : k \in \mathbf{N},\ \alpha_i \in K \text{ e } a_i \in A \text{ para } i = 1, ..., k\}$$

e

$$J_n(A) = \{a \in A : (n!)^k a \in I_n(A) \text{ para algum inteiro não-negativo } k\}.$$

Observação 9.1.1 *Os conjuntos $I_n(A)$ e $J_n(A)$ são subespaços do espaço vetorial A.*

Lema 9.1.1 *Se A é uma álgebra associativa e n é um número natural então $S_n(a_1, ..., a_n) \in I_n(A)$ para quaisquer $a_1, ..., a_n$ em A.*

Prova: Considerando o monômio $m(X_1, ..., X_n) = X_1...X_n$, vale a igualdade

$$S_n(X_1, ..., X_n) = \sum_{\sigma \in S_n} m(X_{\sigma(1)}, ..., X_{\sigma(n)}).$$

Pelo **Lema 3.5.1**, temos que $S_n(a_1, ..., a_n)$ é igual a

$$\begin{aligned}
&m(a_1 + ... + a_n, ..., a_1 + ... + a_n) - \\
&- \sum_{q=1}^{n} m(a_1 + ... + \widehat{a}_q + ... + a_n, ..., a_1 + ... + \widehat{a}_q + ... + a_n) + \\
&+ \sum_{1 \leqslant q_1 \leqslant q_2 \leqslant n}^{n} m(a_1 + ... + \widehat{a}_{q_1} + ... + \widehat{a}_{q_2} + ... + a_n, ..., a_1 + ... + \widehat{a}_{q_1} + ... + \widehat{a}_{q_2} + ... + a_n) + \\
&+ ... + (-1)^{n-1} \sum_{q=1}^{n} m(a_q, ..., a_q).
\end{aligned}$$

Logo,

$$\begin{aligned}
S_n(a_1, ..., a_n) \;=\; & (a_1 + ... + a_n)^n - \sum_{q=1}^{n} (a_1 + ... + \widehat{a}_q + ... + a_n)^n + \\
& + \sum_{1 \leqslant q_1 \leqslant q_2 \leqslant n}^{n} (a_1 + ... + \widehat{a}_{q_1} + ... + \widehat{a}_{q_2} + ... + a_n)^n + ... + (-1)^{n-1} \sum_{q=1}^{n} (a_q)^n.
\end{aligned}$$

Portanto, $S_n(a_1, ..., a_n) \in I_n(A)$ para quaisquer $a_1, ..., a_n$ em A. ■

Lema 9.1.2 *Se A é uma álgebra associativa e n é um número natural então $(n!)^2 I_n(A)A \subseteq I_n(A)$ e $(n!)^2 A I_n(A) \subseteq I_n(A)$.*

Prova: Se a e b são elementos da álgebra A, então

$$S_n(ab, a, ..., a) = (n-1)! \sum_{i=0}^{n-1} a^i (ab) a^{n-1-i} = a(n-1)! \sum_{i=0}^{n-1} a^i b a^{n-1-i} = a S_n(b, a, ..., a),$$

e, pelo **Lema 9.1.1**, temos que

$$a S_n(b, a, ..., a) \in I_n(A).$$

Considerando $F(X_1, ..., X_n) = X_1 S_n(b, X_2, ..., X_n)$, pelo **Lema 3.5.1** e pela conclusão anterior, temos que

$$\sum_{\sigma \in S_n} F(a_{\sigma(1)}, ..., a_{\sigma(n)}) = \sum_{\sigma \in S_n} a_{\sigma(1)} S_n(b, a_{\sigma(2)}, ..., a_{\sigma(n)}) \in I_n(A).$$

Logo,

$$(n-1)! \sum_{i=1}^{n} a_i S_n(b, a_1, ..., a_{i-1}, a_{i+1}, ..., a_n) \in I_n(A).$$

Substituindo $a_1 = a$ e $a_2 = a_3 = ... = a_n = b$ na última fórmula, obtemos

$$(n-1)!aS_n(b,b,...,b) + (n-1)(n-1)!bS_n(a,b,...,b) \in I_n(A),$$

e

$$(n!)^2 ab^n \in I_n(A).$$

Analogamente provamos que para quaisquer a e b em A

$$(n!)^2 b^n a \in I_n(A).$$

Portanto, $(n!)^2 I_n(A)A \subseteq I_n(A)$ e $(n!)^2 AI_n(A) \subseteq I_n(A)$. ■

Lema 9.1.3 *Sejam A uma álgebra associativa, n um número natural e a um elemento de A. Se $(n_1...n_r)a \in I_n(A)$, onde $r \in \mathbf{N}$ e $0 < n_i \leqslant n$ para cada i em $\{1,...,r\}$, então $a \in J_n(A)$.*

Prova: Seja k um inteiro não-negativo tal que $(n!)^k[(n_1...n_r)a] \in I_n(A)$. Logo, $(n!)^{k+r}a \in I_n(A)$ e $a \in J_n(A)$. ■

Corolário 9.1.1 *Se A é uma álgebra associativa e n é um número natural então $J_n(A)$ é um ideal da álgebra A, e o quociente $A/J_n(A)$ não contém elementos não-nulos de ordem aditiva menor ou igual a n.*

Prova: Como já foi observado, $I_n(A)$ é um subespaço do espaço vetorial A. Sejam $a \in J_n(A)$, $b \in A$ e k um inteiro não-negativo para o qual $(n!)^k a \in I_n(A)$. Pelo **Lema 9.1.2**, temos que $(n!)^2[(n!)^k a]b \in I_n(A)$ e $(n!)^2 b[(n!)^k a] \in I_n(A)$, e conseqüentemente $(n!)^{k+2}ab \in I_n(A)$ e $(n!)^{k+2}ba \in I_n(A)$. Portanto, $J_n(A)$ é um ideal da álgebra A.

Suponhamos, por absurdo, que exista um elemento não-nulo $a + J_n(A)$ de $A/J_n(A)$ cuja ordem aditiva é um número natural $k \leqslant n$. Logo, $k(a + J_n(A)) = 0 + J_n(A)$ e $ka \in J_n(A)$. Assim, $a \in J_n(A)$ e $a + J_n(A) = 0 + J_n(A)$, o que é uma contradição. Portanto, $A/J_n(A)$ não contém elementos não-nulos de ordem aditiva menor ou igual a n. ■

9.2 Teorema de Nagata-Higman

Teorema 9.2.1 (Nagata-Higman) *Se A é uma álgebra associativa então para qualquer número natural n*

$$A^{2^n-1} \subseteq J_n(A).$$

Prova: Provaremos o teorema por indução finita em n. Se $n = 1$, então o resultado é óbvio; suponhamos que o teorema seja válido para $n-1$ e vamos prová-lo para $n > 1$. Devido ao fato de que $S_n(b,a,...,a) \in I_n(A)$, para quaisquer a e b em A temos que

$$\sum_{i=0}^{n-1} a^i b a^{n-1-i} \in J_n(A).$$

Sejam a, b, c elementos arbitrários da álgebra A. Consideremos a soma

$$\begin{aligned}
\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} a^{n-1-i}cb^j a^i b^{n-1-j} &= \sum_{j=1}^{n-1}\sum_{i=1}^{n-1} a^{n-1-i}cb^j a^i b^{n-1-j} = \sum_{j=1}^{n-1}(\sum_{i=1}^{n-1} a^{n-1-i}(cb^j)a^i)b^{n-1-j} = \\
&= \sum_{j=1}^{n-1}(\sum_{i=0}^{n-2} a^i(cb^j)a^{n-1-i})b^{n-1-j} = \sum_{j=1}^{n-1}(\sum_{i=0}^{n-2} a^i(cb^j)a^{n-1-i})b^{n-1-j} = \\
&= \sum_{j=1}^{n-1}((\sum_{i=0}^{n-1} a^i(cb^j)a^{n-1-i}) - a^{n-1}(cb^j))b^{n-1-j} = \\
&= k - \sum_{j=1}^{n-1} a^{n-1}cb^j b^{n-1-j} = \\
&= k - \sum_{j=1}^{n-1} a^{n-1}cb^{n-1} = k - (n-1)a^{n-1}cb^{n-1},
\end{aligned}$$

onde $k \in J_n(A)$. Por outro lado, manipulando a mesma soma obtemos

$$\begin{aligned}
\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} a^{n-1-i}cb^j a^i b^{n-1-j} &= \sum_{i=1}^{n-1} a^{n-1-i}c\sum_{j=1}^{n-1} b^j a^i b^{n-1-j} = \\
&= \sum_{i=1}^{n-1} a^{n-1-i}c(\sum_{j=0}^{n-1} b^j a^i b^{n-1-j} - a^i b^{n-1}) = \\
&= k_1 - \sum_{i=1}^{n-1} a^{n-1-i}ca^i b^{n-1} = k_1 - \sum_{i=0}^{n-1} a^{n-1-i}ca^i b^{n-1} + a^{n-1}cb^{n-1} = \\
&= k_1 + k_2 + a^{n-1}cb^{n-1},
\end{aligned}$$

onde k_1 e k_2 pertencem a $J_n(A)$. Combinando os dois resultados temos que

$$na^{n-1}cb^{n-1} \in J_n(A).$$

Como os elementos a, b, c são arbitrários, segue que

$$nI_{n-1}(A)AI_{n-1}(A) \subseteq J_n(A),$$

e além disso

$$J_{n-1}(A)AJ_{n-1}(A) \subseteq J_n(A).$$

Pela hipótese de indução $A^{2^{n-1}-1} \subseteq J_{n-1}(A)$, e conseqüentemente

$$A^{2^n-1} = A^{2^{n-1}-1}AA^{2^{n-1}-1} \subseteq J_{n-1}(A)AJ_{n-1}(A) \subseteq J_n(A),$$

provando o teorema. ■

Agora estamos em condições para provarmos o resultado que motivou o desenvolvimento do presente capítulo.

Corolário 9.2.1 *Toda nil álgebra de índice limitado n, e sem elementos de ordem aditiva menor ou igual a n, é nilpotente, com índice de nilpotência menor ou igual a $2^n - 1$.*

Prova: Seja A uma nil álgebra de índice limitado n. Logo, $I_n(A) = \{0\}$.

Vamos provar que $J_n(A) = \{0\}$. Seja a um elemento não-nulo de A. Se $n > 2$, então $2a \neq 0$. Se $n > 3$, então $3(2a) = (3!)a \neq 0$. Repetindo este processo até n, obtemos que $(n!)a \neq 0$. Por indução finita em k, provamos facilmente que $(n!)^k a \neq 0$ para todo inteiro não-negativo k. Portanto, $a \notin J_n(A)$. Assim, $I_n(A) = \{0\}$.

Logo, $A^{2^n-1} \subseteq \{0\}$ e A é uma álgebra nilpotente com índice de nilpotência menor ou igual a $2^n - 1$. ■

Corolário 9.2.2 *Se A é uma álgebra associativa e n é um número natural então existe um número natural k tal que para qualquer $a \in A^{2^n-1}$, temos que*

$$(n!)^k a = \sum_{i=1}^{r} \alpha_i (a_i)^n,$$

para algum $r \in \mathbf{N}$, $\alpha_i \in K$ e $a_i \in A$ onde $i = 1, ..., r$.

Prova: Considerando a álgebra $K\{\mathbf{X}\}$, pelo Teorema de Nagata-Higman, temos que

$$X_1 X_2 ... X_{2^n-1} \in J_n(K\{\mathbf{X}\})$$

e conseqüentemente, para algum k,

$$(n!)^k X_1 X_2 ... X_{2^n-1} = \sum_{i=1}^{r} \alpha_i (p_i)^n,$$

para algum $r \in \mathbf{N}$, $\alpha_i \in K$ e $p_i \in K\{\mathbf{X}\}$ onde $i = 1, ..., r$. Assim, é fácil ver que k é o número procurado. ■

Consideremos $f : \mathbf{N} \to \mathbf{N}$ a aplicação onde, para cada $n \in \mathbf{N}$, $f(n)$ é o menor número natural tal que $A^{f(n)} \subseteq J_n(A)$. Pelo Teorema de Nagata-Higman, sabemos que $f(n) \leqslant 2^n - 1$. Será que $f(n) = 2^n - 1$? A resposta dessa pergunta é negativa, pois um resultado de Higman mostra que $A^6 \subseteq J_3(A)$. Assim $2^n - 1$ não é exatamente $f(n)$, mas apenas um limite superior. Razmyslov provou que $f(n) \leqslant n^2$ e Kuz'min que $f(n) \geqslant n(n+1)/2$. Portanto, o que sabemos sobre a aplicação f é que $n(n+1)/2 \leqslant f(n) \leqslant n^2$.

Bibliografia

[1] S. A. Amitsur, *A note on PI-rings*, Israel J. Math. **10** (1971), 210-211.

[2] S. A. Amitsur e J. Levitzki, *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1** (1950), 449-463.

[3] M. Dehn, *Über die Grundlangen der projektiven Geometrie und allgemeine Zahlsysteme*, Math. Ann. **85** (1922), 184-193.

[4] V. Drensky, *A minimal basis for the identities of a second order matrix algebra over a field of characteristic* 0, **43** (1981), 188-194.

[5] J. Dubnov e V. Ivanov, *Sur l'abaissement du degré des polynômes affineurs*, C. R. Acad. Sci. URSS **41** (1943), 95-98.

[6] E. S. Golod, On nil-algebras and residually finite p-groups, Izv. Akad. Nauk SSSR Ser. Mat. **28** (1964), 273-276 (original em russo); Amer. Math. Soc. Transl. (2) **48** (1965) (tradução para o inglês)

[7] E. S. Golod e I. R. Shafarevich, *On class field towers*, Izv. Akad. Nauk SSSR Ser. Mat. **28** (1964), 261-272 (original em russo); Amer. Math. Soc. Transl. (2) **48** (1965) (tradução para o inglês)

[8] M. Hall, *Projective planes*, Trans. Amer. Math. Soc. **54** (1943), 229-277.

[9] I. N. Herstein, *Noncommutative rings*, The Carus Math. Monographs **15**, Math. Assoc. Amer., New York, 1968.

[10] G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. **2** (1952), 326-336.

[11] G. Higman, *On a conjecture of Nagata*, Proc. Cambridge Philos. Soc. **52** (1956), 1-4.

[12] N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. Math. **46** (1945), 695-707.

[13] I. Kaplansky, *On a problem of Kurosh and Jacobson*, Bull. Amer. Math. Soc. **52** (1946), 496-500.

[14] I. Kaplansky, *Rings with a polynomial identity*, Bull. Amer. Math. Soc. **54** (1948), 575-580.

[15] A. R. Kemer, *Finite bases for identities of associative algebras*, Algebra i Logika **27** (1987), 597-641.

[16] A. R. Kemer, *Ideals of identities of associative algebras*, Trans. Math. Monographs **87**, Amer. Math. Soc., 1991.

[17] B. Kostant, *A theorem of Frobenius, a theorem of Amitsur-Levitzki, and cohomology theory*, Indiana J. Math. and Mech. **7** (1958), 237-264.

[18] A. I. Kostrikin, *Introducción al álgebra*, Mir, Moscú, 1983.

[19] A. G. Kurosh, *Ringtheorestische Probleme, die mit dem Burnsideschen Probleme über periodiche Gruppen in Zusammenghang stehen*, Bull. Acad. Sci. URSS Ser. Math. **5** (1941), 223-240.

[20] J. B. Kruskal, *Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture*, Trans. Amer. Math. Soc. **95** (1960), 210-225.

[21] V. N. Latyshev, *On Regev's theorem on identities in a tensor product of PI-algebras*, Ups. Mat. Nauk. **27** (1972), 213-214 (em russo).

[22] V. N. Latyshev, *Partially ordered sets and nonmatrix identities of associative algebras*, Algebra i Logika **15** (1976), 34-44.

[23] J. Levitzki, *On a problem of Kurosh*, Bull. Amer. Math. Soc. **52** (1946), 1033-1035.

[24] W. Miller Jr., *Symetric groups and their applications*, Pure and Appl. Math. **50**, Academic Press, 1972.

[25] M. Nagata, *On the nilpoteny of nil-algebras*, J. Math. Soc. Japan **4** (1952), 296-301.

[26] R. S. Pierce, *Associative algebras*, GTM **88**, Springer, New York, 1982.

[27] C. Procesi, *Rings with polynomial identities*, Pure and Appl. Math. **17**, Marcel Dekker, New York, 1973.

[28] Yu. P. Razmyslov, *Finite basing of the identities of a matrix algebra of second order over a field of characteristic* 0, Algebra i Logika **12** (1973), 47-63.

[29] A. Regev, *Existence of identities in* $A \otimes B$, Israel J. Math. **11** (1972), 131-152.

[30] A. Regev, *The representations of* S_n *and explicit identities for PI-algebras*, J. Algebra **51** (1978), 25-40.

[31] S. Rosset, *A new proof of the Amitsur-Levitzki identity*, Israel J. Math. **23** (1976), 187-188.

[32] L. H. Rowen, *Polynomial identities in Ring Theory*, Pure and Appl. Math. **84**, Academic Press, 1980.

[33] A. I. Shirshov, *On certain nonassociative nil rings and algebraic algebras*, Mat. Sb. **41** (83) (1957), 381-394 (original em russo); Amer. Math. Soc. Transl. **119** (2) (1983), 119-132 (tradução para o inglês).

[34] A. I. Shirshov, *On rings with polynomial identities*, Mat. Sb. **43** (85) (1957), 277-283 (original em russo); Amer. Math. Soc. Transl. **119** (2) (1983), 133-139 (tradução para o inglês).

[35] W. Specht, *Gesetze in Ringen I*, Math. Z. **52** (1950), 557-589.

[36] R. G. Swan, *An application of graph theory to algebra*, Proc. Amer. Math. Soc. **14** (1963), 363-373.

[37] R. G. Swan, *Correction to "An application of graph theory to algebra"*, Proc. Amer. Math. Soc. **21** (1969), 379-380.

[38] W. Wagner, *Über die Grundlangen der projektiven Geometrie und allgemeine Zahlsysteme*, Math. Ann. **113** (1936), 528-567.

[39] K. A. Zhevlakov, A. M. Slin'ko, I. P. Shestakov, A. I. Shirshov, *Rings that are nearly associative*, Pure and Appl. Math. **104**, Academic Press, 1982.