Sobre anéis de séries formais: a altura de estrela e uma imersão do anel de grupo livre

Douglas de Araujo Smigly

Relatório apresentado
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
a disciplina
MAT0148 - Introdução ao
Trabalho Científico

Bacharelado em Matemática

Orientador: Prof. Dr. Javier Sánchez Serdà

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq, através do processo CNPq: 157567/2018-4

São Paulo, junho de 2019

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Os contribuintes

O presente trabalho foi desenvolvido por Douglas de Araujo Smigly, aluno do Instituto de Matemática e Estatística da Universidade de São Paulo, regularmente matriculado no curso de Bacharelado em Matemática, no sétimo semestre do curso na presente data, sob orientação do Prof. Dr. Javier Sánchez Serdà, professor doutor do Departamento de Matemática da USP, referente à sua Iniciação Científica em Álgebra, mais especificamente Séries Racionais e Derivações no Anel de Grupo Livre.

Este trabalho é composto de um resumo e apresentação das atividades e projetos desenvolvidos pelo aluno durante o período de Ago/2018 até Jun/2019, que consistiu principalmente em estudos pessoais da teoria presente nas bibliografias pelo aluno toda semana de segunda a sábado, e de apresentações e reuniões semanais entre o aluno e o professor para discussão dos temas abordados, das pesquisas e das elaborações de generalizações feitas por parte do aluno.

O aluno recebeu apoio do CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, através do processo CNPq: 157567/2018-4 durante toda a Iniciação Científica, e recebeu bolsa do PICME - Programa de Iniciação Científica e Mestrado, devido a medalhas obtidas na Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) e Olimpíada Brasileira de Matemática (OBM) obtidas durante o Ensino Fundamental e o Ensino Médio. Com isso, ele participou duas vezes do Simpósio Internacional de Iniciação Científica e Tecnológica da USP - SIICUSP, apresentando pôsteres e os trabalhos desenvolvidos sob os títulos "Localização em Teoria de Anéis" e "Séries Formais Não-Comutativas", nos 25º SIICUSP e 26º SIICUSP, respectivamente.

1.2 Resumo do trabalho

Neste trabalho, estudamos dois resultados relacionados aos anéis de séries formais:

1.2.1 Altura de estrela

Seja $X = \{x_1, \ldots, x_n\}$ um conjunto com n elementos. Denotamos por X^* o monoide livre em X.

Dado um anel comutativo k denotamos por $k\langle X\rangle$ a k-álgebra livre em X e por $k\langle\langle X\rangle\rangle$ o anel de séries formais em X

$$k\langle\langle X\rangle\rangle = \Big\{ \sum_{\omega\in X^*} \alpha_\omega \omega \colon \alpha_\omega \in k \Big\}.$$

Uma série formal $S \in k\langle\langle X\rangle\rangle$ é dita *própria* se o seu termo constante é zero. Pode ser provado que se S é própria, então a expressão $S^* = \sum_{n\geq 0} S^n$ define uma série em $k\langle\langle X\rangle\rangle$.

As operações racionais em $K\langle\langle X\rangle\rangle$ são a soma de séries, o produto de séries, o produto de elementos de k por elementos de $k\langle\langle X\rangle\rangle$ e a operação $*$. Um subconjunto de $K\langle\langle X\rangle\rangle$ é *racionalmente fechado* se for fechado pelas operações racionais. O menor subconjunto contendo um subconjunto

E de $K\langle\langle X\rangle\rangle$ e que seja racionalmente fechado é chamado *fecho racional* de E. Uma série formal é *racional* se estiver no fecho racional de $K\langle X\rangle$.

A *altura de estrela* de uma série racional $S \in K\langle\langle X\rangle\rangle$ é definida como segue. Definimos uma sequência

$$R_0 \subseteq R_1 \subseteq \cdots \subseteq R_n \subseteq \cdots$$

de subconjuntos de $K\langle\langle X\rangle\rangle$ tal que a união de todos os R_n seja o conjunto de todas as séries racionais. O conjunto R_0 é $K\langle X\rangle$, e, para i, o conjunto R_{i+1} é o subanel de $K\langle\langle X\rangle\rangle$ gerado pelo conjunto dos S^* em que $S \in R_i$ e S é própria. A *altura de estrela* de uma série S é o menor inteiro n tal que $S \in R_n$.

Nessa primeira parte, tivemos como objetivo estudar o Capítulo 4 de [28] que trata da altura de estrela. O resultado principal é a caracterização da altura de estrela de uma série racional junto com uma importante consequência, a altura de estrela não está limitada (i.e. se X tem pelo menos n^2 elementos, então existem séries de altura de estrela n). Para chegar nesses resultados, são utilizadas técnicas de teoria dos grafos e autômatos, a partir da construção e definição da representação linear minimal de uma série racional. Esses resultados são estudados nos capítulos 1 até 4 de [28].

1.2.2 Álgebra de grupo livre como subálgebra do anel de séries formais

A segunda parte do trabalho consistiu em estudar as seções 1–4 do artigo [7].

Seja $X = \{x_1, \ldots, x_n\}$ um conjunto com n elementos e K um anel. Denotamos por $F(X)$ o grupo livre em X, e por $K[F(X)]$ a K-álgebra de grupo do grupo $F(X)$ com coeficientes em K. A *função de aumento* é o homomorfismo de grupos

$$\begin{aligned} \varepsilon\colon K[F(X)] &\longrightarrow K \\ \sum_{g\in F(X)} \alpha_g g &\longmapsto \sum_{g\in F(X)} \alpha_g \end{aligned}$$

O *ideal de aumento* é o ideal $I = \ker \varepsilon$.

É fácil provar que as séries da forma $1+x_i \in K\langle\langle X\rangle\rangle$ são inversíveis. Logo a função $X \to K\langle\langle X\rangle\rangle$, $x_i \mapsto 1 + x_i$ pode ser estendida a um homomorfismo de K-álgebras $\Phi\colon K[F(X)] \to K\langle\langle X\rangle\rangle$.

Os resultados que estudamos em [7] foram os seguintes:

1. Φ é um homomorfismo injetor.
2. $\bigcap_{n\geq 1} I^n = \{0\}$.

Esses resultados são obtidos a partir do estudo das derivações do anel de grupo $K[F(X)]$ feito nas seções 2 e 3 do artigo. Na seção 4 de [7], esses resultados são relacionados com o ideal de aumento.

1.3 Estrutura do relatório

O presente relatório está estruturado em três grandes partes:

- Na primeira parte do texto, apresentaremos os resultados básicos para uma melhor compreensão do uso de tais estruturas ao decorrer do trabalho e sua contextualização. Para isso, apresentam-se definições e resultados introdutórios bastante conhecidos em álgebra, como conceitos de anéis, módulos, semianéis, módulos sobre semianéis e relações de ordem.
- A segunda parte deste trabalho, concernente à Altura de Estrela, está apresentada nos capítulos 3-5. No capítulo 3, são tratados conceitos gerais acerca de séries formais e importantes tópicos, como o semianel das séries formais e suas propriedades, séries racionais e representação linear, conforme exposto em [28]. Na seção 3.3, apresentamos a definição de séries

reconhecíveis e suas propriedades, além de explorar outras operações em $A\langle\langle A\rangle\rangle$, e definir subconjuntos estáveis, importantes para a demonstração de um dos principais resultados desse capítulo, o Teorema de Schützenberger 3.3.12, que estabelece a equivalência do conceito de séries racionais e reconhecíveis. Estudamos as representações lineares minimais para uma série, constatando que duas representações lineares minimais quaisquer são semelhantes no teorema 3.3.34. Também são estudados conceitos relacionados à álgebras e ideais sintáticos de séries formais, e demonstramos o teorema de Paz-Carlyle-Fliess 3.3.26, que trata da relação entre o posto de uma série formal, a dimensão de seu ideal sintático e o posto de suas respectiva matriz de Hankel. Para finalizar o capítulo, são apresentados alguns conceitos sobre autômatos ponderados em 3.4, e são estabelecidas relações entre estes e as representações lineares minimais de uma série racional, o que permite criar um autômato ponderado que reconheça a série racional sobre certas condições.

No capítulo 4, definimos o semianel das expressões racionais, e verificamos algumas de suas propriedades importantes, como o morfismo

$$\mathrm{eval}\colon \mathcal{E} \to K\langle\langle A\rangle\rangle$$

conceituado na proposição 4.1.5, e consequências da identidade racional

$$E^* \sim 1 + EE^* \sim 1 + EE^*$$

apresentada em 4.1, bem como sua generalização para matrizes. Estudamos em seguida as expressões racionais sobre um anel, e, a grosso modo, a "trivialidade" das identidades racionais; já no capítulo 5, nos atemos ao resultado principal dessa primeira parte, no qual desenvolvemos a caracterização da altura de estrela de uma série racional, além de mostrarmos no teorema 5.3.2 que a altura de estrela não está limitada (isto é, se X tem pelo menos n^2 elementos, então existem séries de altura de estrela n). Para isso, são introduzidos diversos exemplos introdutórios, e apresentamos alguns conceitos básicos sobre Teoria de Grafos em 5.2 qeu serão de grande utilidade, como a definição de complexidade de ciclo dada em 5.2.27. Com todos esses subsídios, foi possível demonstrar o teorema 5.3.2 sobre a altura de estrela. Ainda, o apêndice A mostra as motivações para dos lemas utilizados na demonstração desse resultado, como o Teorema de Mandel-Simon A.1.4, apresentado em [18] e o crescimento polinomial das séries $\mathbb{Z}$-racionais.

- A terceira parte deste trabalho é sobre derivações no anel de grupo livre, apresentada nos capítulos 6-7. No capítulo 6, são apresentadas generalidades sobre anéis de grupo e derivações. Estudamos várias definições e propriedades dos Anéis de Grupo, e também nos concentramos no estudo dos grupos livres, para a posterior conceituação dos anéis de grupo livres. Em seguida, definimos e estudamos as propriedades das derivações no anel de grupo livre, de acordo com o exposto no artigo [7], demonstrando a existência de uma única derivação tal que $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$ para cada gerador x_i de X, como no teorema 6.1.22 e deduzimos uma fórmula fundamental para esta derivação, obtida no teorema 6.1.23. Em seguida, estudamos as séries centrais descendentes em 6.2 para compreender melhor e poder desenvolver a teoria sobre a estrutura do anel de grupo livre, analisando a série central descendente de $F(X)$ a partir de sua conexão com as potências do ideal fundamental do anel de grupo $R[F(X)]$. Finalmente, com toda a teoria desenvolvida, obtemos o teorema de unicidade da expansão de séries formais 6.2.7, que nos permitiu demonstrar no teorema 6.3.4 a existência do homomorfismo injetor de K-álgebras

$$\Phi\colon K[F(X)] \to K\langle\langle X\rangle\rangle$$

definido em 6.3.2. No capítulo 7, reproduzimos e generalizamos alguns dos resultados obtidos no capítulo 6 para um grupo livre parcialmente comutativo, substituindo o anel de séries $k\langle\langle X\rangle\rangle$ pelo anel de séries formais parcialmente comutativas. Para isso, estudamos os monoides livres parcialmente comutativos e a série central descendente de $F(A, \vartheta)$ seguindo a

abordagem apresentada no artigo [5], estudando brevemente anéis e módulos graduados e álgebras de Lie. Com essas ferramentas, generalizamos a noção de derivação para séries parcialmente comutativas, e conseguimos demonstrar que existe uma única derivação satisfazendo as condições apresentadas em 7.3.7.

Com o objetivo de ser o mais didático e claro, são apresentados sempre que possível e pertinente exemplos que ilustrem a teoria que está sendo apresentada a cada ponto do texto,permitindo assim que o leitor tenha uma compreensão mais clara dos resultados que setão sendo demonstrados e desenvolvidos.

Capítulo 2

Conceitos Introdutórios

Neste capítulo, apresentaremos alguns conceitos básicos para fins de contextualização e que serão utilizados e desenvolvidos no decorrer deste trabalho, como um interlúdio para os estudos principais deste projeto. Para maiores detalhes, recomendamos as referências [28], [1], [15] e [8].

2.1 Anéis

Definição 2.1.1. Um conjunto R, munido com duas operações binárias $+$ e $\cdot$, chamadas de *adição* e *multiplicação*, respectivamente, tais que

1. $(R, +)$ é um grupo abeliano, com elemento neutro 0 e inverso do elemento $r \in R$ denotado por $-r$;
2. A multiplicação é associativa, ou seja,
$$(xy)z = x(yz)$$
3. A multiplicação é distributiva sobre a adição, ou seja,
$$x(y+z) = xy + xz$$
$$(y+z)x = yx + zx$$
4. Existe elemento neutro da multiplicação, ou seja, um elemento $1 \in R$ tal que
$$x1 = 1x = x \qquad x \in R$$

é chamado *anel*. Se além das condições acima, a multiplicação também é comutativa, ou seja:

- $xy = yx \qquad \forall x, y \in R$.

então o anel R é dito *comutativo*.

Definição 2.1.2. Seja R um anel. Um elemento $r \in R$ é *inversível à direita (respectivamente à esquerda)* se existe $a \in R$ tal que $ra = 1$ (respectivamente $ar = 1$). Além disso, r é chamado de *unidade* ou *inversível* se e só se r é simultaneamente inversível à direita e à esquerda. Denotamos por $\mathcal{U}(R)$ o conjunto das unidades de R, ou seja, o conjunto dos elementos inversíveis do anel.

Definição 2.1.3. Um anel R é chamado de *domínio* se $R \neq \{0\}$ e para todo $x, y \in R$,
$$xy = 0 \Rightarrow x = 0 \text{ ou } y = 0$$

Definição 2.1.4. Seja R um anel. Um elemento $r \in R$ é dito

- *nilpotente* se possui alguma potência nula: $r^n = 0$ para algum $n \in \mathbb{N}$.
- *idempotente* se e somente se $r^2 = r$ (e portanto $r^n = r$ para todo $n \in \mathbb{N}$ não nulo).
- *inversível à esquerda (resp. à direita)* se e só se existe $a \in R$ tal que $ar = 1$ (resp. $ra = 1$). Além disso, r é dito *inversível* ou *unidade* se e somente se r é inversível à direita e à esquerda.

O conjunto dos elementos inversíveis de um anel R é denotado por

$$\mathcal{U}(R) = \{u \in R | u \text{ é unidade em } R\}$$

Definição 2.1.5. Num anel A, um elemento $x \in A$ é dito **inversível à direita (resp. à esquerda)** se e só se existe $a \in A$ tal que $xa = 1$ (resp. $ax = 1$). Além disso, x é dito **inversível** se e só se x é inversível à direita e à esquerda. O conjunto dos elementos inversíveis de um anel A é denotado por $U(A)$.

Definição 2.1.6. Um anel R é chamado *anel com divisão* se $R \neq 0$ e $R \setminus \{0\} = \mathcal{U}(R)$. Um *corpo* é um anel comutativo com divisão.

Definição 2.1.7. Um subconjunto S de um anel R é chamado *subanel* de R se :

1. S é um subgrupo abeliano de R.
2. Para todo $r, s \in S$, então $rs \in S$;
3. $1 \in S$.

Definição 2.1.8. Dada uma coleção de anéis $\{B_\lambda\}_{\lambda \in \Lambda}$, o produto cartesiano

$$\prod_{\lambda \in \Lambda} B_\lambda$$

define um anel, em que a soma e multiplicação são efetuadas coordenada a coordenada. O elemento neutro deste anel é a tupla constante com todas entradas iguais a 0 e a identidade é a tupla constante com todas entradas iguais a 1.

2.1.1 Homomorfismo de anéis

Definição 2.1.9. Sejam R e A anéis. Uma função $\varphi \colon R \to A$ tal que:

1. $\varphi(1_R) = 1_A$;
2. $\varphi(x + y) = \varphi(x) + \varphi(y)$;
3. $\varphi(xy) = \varphi(x)\varphi(y)$.

é chamada de *homomorfismo de anéis.*

O *núcleo* ou *kernel* de φ é

$$\text{Ker}(\varphi) = \{r \in R | \varphi(r) = 0\}$$

Dizemos que φ é um *isomorfismo* se for um homomorfismo bijetor.

Definição 2.1.10. Seja $\varphi \colon R \to S$ um homomorfismo de anéis. Dizemos que φ é um *epimorfismo de anéis* se e somente se para qualquer anel A e quaisquer homomorfismos $f, g \colon S \to A$, temos que

$$f \circ \varphi = g \circ \varphi \Rightarrow f = g$$

Definição 2.1.11. Sejam R e S anéis. Então o produto cartesiano $R \times S$, munido de duas operações binárias $+, \cdot$, definida para $r_1, r_2 \in R, s_1, s_2 \in S$:

- $(r_1, s_1) + (r_2, s_2) = (r_1 + r_2, s_1 + s_2)$.
- $(r_1, s_1)(r_2, s_2) = (r_1 r_2, s_1 s_2)$.

é chamado *produto direto* de R e S.

Note que $R \times S$ de fato é um anel, em que o neutro aditivo é $(0, 0)$, o oposto aditivo de (x, y) é $(-x, -y)$, e a unidade é $(1, 1)$.

Proposição 2.1.12. *Sejam R e S anéis, e $F\colon R \to S$, uma função. Então F é um homomorfismo se e somente se F é um subanel de $R \times S$.*

Demonstração. Considere que F é um homomorfismo de R a S. Desse modo:

- Como $F(1_R) = 1_S$, então $(1_R, 1_S) \in F$.
- Para (r_1, s_1) e (r_2, s_2) em F, então

$$F(r_1) = s_1 \quad \text{e} \quad F(r_2) = s_2.$$

  Então,

$$F(r_1 - r_2) = F(r_1) - F(r_2) = s_1 - s_2 \Rightarrow (r_1, s_1) - (r_2, s_2) = (r_1 - r_2, s_1 - s_2) \in F$$

- Para (r_1, s_1) e (r_2, s_2) em F, então novamente usando que

$$F(r_1) = s_1 \quad \text{e} \quad F(r_2) = s_2,$$

  temos

$$F(r_1 r_2) = F(r_1)F(r_2) = s_1 s_2 \Rightarrow (r_1, s_1)(r_2, s_2) = (r_1 r_2, s_1 s_2) \in F.$$

Portanto, concluímos que F é um subanel de $R \times S$. □

2.1.2 Módulo sobre um anel

Definição 2.1.13. Seja R um anel. Um *R-módulo à direita*[1] é um grupo abeliano M com uma função $\cdot\colon M \times R \to M$ tal que, para todos $x, y \in M$ e $r, s \in R$:

- $x(r + s) = xr + xs$;
- $(x + y)r = xr + yr$;
- $x(rs) = xr + yr$;
- $x1 = x$

Definição 2.1.14. Um subconjunto N de um R-módulo à direita M é um *submódulo* de M se N é um subgrupo abeliano de M e para $r \in R$ e $x \in N$, então $xr \in N$.

Definição 2.1.15. Sejam R e S anéis. Um *R, S-bimódulo* é um grupo abeliano M munido de uma estrutura de R-módulo à esquerda e um S-módulo à direita tal que:

- $(rx)s = r(xs)\ \forall\ r \in R, s \in S$ e $x \in M$.

Proposição 2.1.16. *Seja $f : R \to S$ um homomorfismo de anéis. Então podemos considerar S como um R-bimódulo.*

[1] Analogamente se define um R-módulo à esquerda.

Demonstração. Vamos definir as operações em S. Definamos a adição em S como a adição da estrutura de anel de S. A mutliplicação por escalar será dada por:

$$x * r = xf(r) \quad \text{e} \quad r * x = f(r)x$$

Como S é um anel, segue trivialmente que S é um R-bimódulo.

□

Definição 2.1.17. Considere R uma anel. Seja M um R-módulo à esquerda (direita). Se N é um R-submódulo à esquerda (direita) de M, então a relação $\sim$ tal que, para todos $x, y \in M$,

$$x \sim y \Leftrightarrow x - y \in N$$

é uma relação de equivalência. Denotemos por $\overline{x}$ a classe de equivalência de $x \in M$. Tome M/N o conjunto de todas as classes de equivalência. Temos que M/N possui uma estrutura de R-módulo à esquerda (direita) tal que:

- $\overline{x} + \overline{y} = \overline{x + y}$ $\forall x, y \in M$.
- $\overline{x}r = \overline{xr} (r\overline{x} = \overline{rx})$ $\forall r \in R, \forall x \in M$.

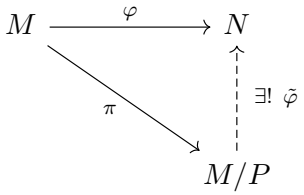
Dizemos que o R-módulo à esquerda (direita) M/N é o *R-módulo à esquerda (direita) quociente.*

Note que a função

$$\begin{array}{rccc} \pi & : & M & \longrightarrow & M/N \\ & & x & \longmapsto & \overline{x} \end{array}$$

é um homomorfismo sobrejetor de R-módulos à esquerda (direita).

Proposição 2.1.18. *Sejam M e N R-módulos à esquerda (direita). Considere um homomorfismo de R-módulos à esquerda (direita) $\varphi\colon M \to N$. Seja ainda $P \subseteq Ker(\varphi)$, um submódulo à esquerda (direita) de N. Então, existe um único homomorfismo $\tilde{\varphi}\colon M/P \to N$ tal que $\tilde{\varphi}(\overline{x}) = \varphi(x)$. Temos o seguinte diagrama:*



Se $P = \ker \varphi$, então M/P é isomorfo a $Im(\varphi)$.

2.1.3 Ideais

Definição 2.1.19. Um *ideal à direita (resp. à esquerda) I* de um anel R é um subconjunto de R que é um subgrupo aditivo tal que

$$xr \in I (\text{resp.} rx \in I), \forall x \in I, \forall r \in R$$

Um *ideal* de R é um subconjunto de R que é um ideal à direita e à esquerda de R. Neste caso, adotamos como notação $I \lhd R$.

Definição 2.1.20. Dada uma família arbitrária $\{b_\lambda\}_{\lambda \in \Lambda}$ de elementos de R, o conjunto de todas as combinações R-lineares (finitas) à esquerda (direita) de elementos nesta família é chamado de *ideal à esquerda (direita) gerado por* $\{b_\lambda\}_{\lambda \in \Lambda}$. Em particular, se R é comutativo, o ideal gerado por uma quantidade finita de elementos $r_1, \ldots, r_n \in R$ será denotado por $\langle r_1, \ldots, r_n \rangle$. Ideais da forma $\langle r \rangle$, isto é, gerados por um único elemento, são chamados de *ideais principais.*

Definição 2.1.21. Considere o conjunto formado por todos os ideias próprios de R, parcialmente ordenado pela inclusão, como dado na definição 2.5.7. Um elemento maximal desse conjunto é dito *ideal maximal*. Em outras palavras, um ideal próprio $\mathfrak{m} \subsetneq R$ é maximal se possui a seguinte propriedade: para qualquer ideal $\mathfrak{a} \subset R$,

$$\mathfrak{a} \supseteq \mathfrak{m} \Rightarrow \mathfrak{a} = \mathfrak{m} \text{ ou } \mathfrak{a} = \langle 1 \rangle = R$$

Definição 2.1.22. Seja R um anel comutativo. Um ideal $\mathfrak{p} \subset R$ é dito *primo* se satisfaz as seguintes condições equivalentes:

- $A/\mathfrak{p}$ é um domínio;
- $\mathfrak{p}$ é um ideal próprio e, dados $a, b \in R$,

$$ab \in \mathfrak{p} \Rightarrow a \in \mathfrak{p} \text{ ou } b \in \mathfrak{p}$$

- $\mathfrak{p}$ é um ideal próprio e, dados ideais $\mathfrak{a}, \mathfrak{b} \subseteq R$,

$$\mathfrak{p} \supseteq \mathfrak{ab} \Rightarrow \mathfrak{p} \subseteq \mathfrak{a} \text{ ou } \mathfrak{p} \subseteq \mathfrak{b}$$

Se I é um ideal de R, podemos considerar a relação de equivalência $\sim$ dada, para todos $r, s \in R$, por

$$r \sim s \Leftrightarrow r - s \in I$$

Considere para $r \in R$ por $\overline{r}$ a classe de equivalência de r. Seja R/I o conjunto das classes de equivalência. Então podemos definir uma estrutura de anel em R/I tal que:

- $\overline{r} + \overline{s} = \overline{r + s}$,
- $\overline{r} \cdot \overline{s} = \overline{r \cdot s}$.

Dizemos que R/I é o *anel quociente* de R por I.

O anel quociente vem equipado com o morfismo quociente, ou morfismo de projeção, que leva um elemento a sua respectiva classe de equivalência:

$$\begin{array}{rccc} \pi & : & R & \longrightarrow & R/I \\ & & r & \longmapsto & \overline{r} \end{array}$$

Claramente π é um homomorfismo sobrejetor de anéis.

Considere $f : R \to S$ é um homomorfismo de anéis. Então para todo ideal I de S, $f^{-1}[I]$ é um ideal de R. Então $\ker(f)$ é um ideal de R.

Proposição 2.1.23. *Dado um homomorfismo de anéis $f : R \to S$, então $Ker(f)$ é um ideal de R e $Im(f)$ é um subanel de S. Além disso, para ideal I, se $I \subseteq Ker(f)$, existe um único homomorfismo*

$$\overline{f} \colon R/I \to S$$

tal que $\overline{f}(\overline{x}) = f(x)$. Se $I = Kerf$, então R/I é isomorfo a $Im(f)$.

Definição 2.1.24. Se R é um anel e I e J são ideais, definimos IJ como o ideal gerado por todos os elementos da forma ab em que $a \in I$ e $b \in J$.

Definição 2.1.25. Um *conjunto multiplicativo* do anel R é um subconjunto $S \subseteq R$ que é fechado por produto, isto é,

$$s, t \in S \Rightarrow st \in S$$

e tal que $1 \in S$.

2.2 Semianéis e monoides

Vamos definir brevemente algumas estruturas algébricas que serão utilizadas de maneira recorrente durante o trabalho.

Definição 2.2.1. Um *monoide* M é um conjunto munido com uma operação que é associativa e possui um elemento neutro, denotado usualmente por 1_M.

Definição 2.2.2. Dizemos que uma função $\varphi\colon M \to N$, onde M e N são monoides, é um *morfismo de monoides* se $\varphi(1_M) = \varphi(1_N)$ e $\varphi(mn) = \varphi(m)\varphi(n), m \in M, n \in N$.

Definição 2.2.3. Um *semianel* é um conjunto R com duas operações $+$, e $\cdot$, que satisfaz os seguintes axiomas:

- $(R, +)$ é um monoide comutativo com elemento neutro denotado por 0.
- $(R, \cdot)$ é um monoide com elemento neutro denotado por 1.
- O produto é distributivo com respeito à soma.
- $\forall\ r \in R, 0r = r0 = 0$.

É interessante notar que o semianel é basicamente um "anel sem subtração", ou seja, sem o elemento inverso para a operação $+$. Note que isso afeta diretamente sua definição, uma vez que exigimos $r = r0 = 0\ \forall\ r \in R$, já que esta propriedade não pode ser demonstrada sem a definição de um inverso para $+$. Exemplos importantes de semianéis são $\mathbb{N}$, $\mathbb{Q}_+$, $\mathbb{R}_+$ e o anel booleano $\mathbb{B} = \{0, 1\}$, onde suas operações estão completamente determinadas por $1 + 1 = 1$.

Definição 2.2.4. Sejam R e S dois semianéis. Então, dizemos que $\varphi\colon R \to S$ é um *morfismo de semianéis* se:

- $\varphi(a + b) = \varphi(a) + \varphi(b), \forall a, b \in R$;
- $\varphi(ab) = \varphi(a)\varphi(b), \forall a, b \in R$;
- $\varphi(0_R) = 0_S$;
- $\varphi(1_R) = 1_S$.

Definição 2.2.5. Sejam R e S dois semianéis. Então, dizemos que $\phi\colon R \to S$ é um *antimorfismo de semianéis* se:

- $\phi(a + b) = \phi(a) + \phi(b), \forall a, b \in R$;
- $\phi(ab) = \phi(b)\phi(a), \forall a, b \in R$;
- $\phi(0_R) = 0_S$;
- $\phi(1_R) = 1_S$.

2.3 Módulos sobre semianéis

Definição 2.3.1. Seja R um semianel. Um *R-módulo à esquerda*[2] é um monoide comutativo $(M, +, 0)$ munido com uma operação externa $R \times M \to M$ denotada por $(k, x) \longmapsto kx$, tal que, $\forall\ r, s \in R$ e $x, y \in M$, são válidas as condições:

[2] Analogamente se define um R-módulo à direita.

- $r(x+y) = rx + ry$
- $(r+s)x = rx + sx$
- $(rs)x = r(sx)$
- $1x = x$
- $0x = 0$
- $r0 = 0$

Um *submódulo* de M é um subconjunto de M contendo 0 e fechado pelas operações de M.[3]

Um R-módulo à esquerda é *finitamente gerado* se existe uma quantidade finita de elementos $\alpha_1, \ldots, \alpha_n \in M$ tais que todo elemento em M pode ser escrito como uma combinação linear da forma $\sum_{i=1}^{n} r_i\alpha_i$, com $r_i \in R$.

A interseção de uma família $(N_i)_{i\in I}$ de submódulos de M é um submódulo de M. Dessa forma, se P é um subconjunto de M, então podemos definir o *submódulo gerado* por P como a interseção de todos os submódulos contendo P.

Definição 2.3.2. Se R é um semianel, dizemos que uma família $(x_i)_{i\in I}$ de elementos de um R-módulo à direita M é *linearmente independente* se, para toda família $(r_i)_{i\in I}$ de elementos de R, em que todos a menos de uma quantidade finita de termos são nulos, ocorre que

$$\sum_i x_i r_i = 0 \Rightarrow r_i = 0 \ \forall \ i \in I$$

Caso contrário, dizemos que a família $(x_i)_{i\in I}$ é *linearmente dependente*.

Definição 2.3.3. Se R é um semianel, dizemos que uma família de elementos $(x_i)_{i\in I}$ de um R-módulo à direita M linearmente independente tal que M é o submódulo gerado pela família é uma *base* de M.

Se M possui uma base com cardinalidade κ, dizemos que M possui posto κ.

Se R é um semianel, um R-módulo M é dito *livre* se ele possui uma base. Por convenção o grupo $\{0\}$ é um R-módulo livre para qualquer anel R, com base $\mathcal{B} = \emptyset$.

Uma das maiores diferenças entre um espaço vetorial (R-módulo, com R corpo) e um R-módulo com R semianel em geral é que nem sempre é possível obter uma base para o segundo, mesmo que R seja anel. Outro fato surpreendente é que as bases de um R-módulo livre podem não ter todas a mesma cardinalidade. Isso só é garantido quando R é um anel com divisão, como registrado no

Teorema 2.3.4. *Seja R um anel com divisão. Então todo R-módulo à direita possui uma base, e quaisquer duas bases possuem a mesma cardinalidade.*

Se R é anel comutativo e M é um R-módulo livre, então esses resultado também vale, como enunciado abaixo:

Teorema 2.3.5. *Seja R um anel comutativo e M um R-módulo livre. Então quaisquer duas bases de M possuem a mesma cardinalidade.*

Observe que esse resultado não é válido para anéis não-comutativos:

Exemplo 2.3.6. Seja K um corpo e denote por $K^{\mathbb{N}}$ o K-espaço vetorial de dimensão contável sobre K. Seja

$$R = \hom_K(K^{\mathbb{N}}, K^{\mathbb{N}})$$

o anel não-comutativo dos endomorfismos K-lineares de $K^{\mathbb{N}}$. A multiplicação em R é dada pela composição. Então, temos isomorfismos de R-módulos à esquerda tais que

$$R \simeq R^2 \simeq R^3 \simeq \cdots$$

Também temos que R é isomorfo ao anel das matrizes com linhas e colunas indexadas por $\mathbb{N}$ e cujas colunas são sequências quase nulas.

[3] Alguns autores referem-se à módulo sobre um semianel como *semimódulo*, e os submódulos sobre semianel como *subsemimódulos*.

Definição 2.3.7. Seja R um semianel. Um *homomorfismo* de R-módulos à esquerda (direita) $\varphi\colon M_1 \to M_2$ é uma aplicação entre R-módulos à esquerda (direita) que satisfaz:

- $\varphi(0) = 0$;
- $\varphi(m_1 + m_2) = \varphi(m_1) + \varphi(m_2),\ \forall\ m_1, m_2 \in M_1$;
- $\varphi(rm) = r\varphi(m)\ (\varphi(mr) = \varphi(m)r),\ \forall\ r \in R,\ \forall\ m \in M_1$.

Se $\varphi\colon M_1 \to M_2$ é um homomorfismo, seu núcleo $\mathrm{Ker}(\varphi)$ e sua imagem $\mathrm{Im}(\varphi)$ são submódulos à esquerda (direita) de M_1 e M_2, respectivamente.

2.4 Álgebras

Definição 2.4.1. Seja R um anel. Uma *R-álgebra* é um anel A com uma multiplicação por escalar

$$\begin{array}{rccc} \cdot & : & R \times A & \longrightarrow & A \\ & & (\lambda, a) & \longmapsto & \lambda a \end{array}$$

satisfazendo as seguintes condições:

(i) $(\lambda\mu)a = \lambda(\mu a)$, para todos $\lambda, \mu \in R$, $a \in A$;

(ii) $(\lambda + \mu)a = \lambda a + \mu a$, para todos $\lambda, \mu \in R$, $a \in A$;

(iii) $\lambda(a + b) = \lambda a + \lambda b$, para todos $\lambda \in K, a, b \in A$;

(iv) $1_R a = a$, para todo $a \in A$;

(v) $\lambda(ab) = (\lambda a)b = a(\lambda b)$, para todos $\lambda, \mu \in R$, $a \in A$.

A tem unidade se existe um elemento $1 \in A$ tal que $1x = x1 = x$ para todo $x \in A$. Tal elemento será chamado de *identidade* em A. Se R é comutativo, A é chamada uma *R-álgebra comutativa.*

É importante notar que, se R é um corpo, então uma R-álgebra também é um espaço vetorial sobre R. Assim, é possível falar sobre a dimensão de uma R-álgebra.

Exemplo 2.4.2. Considere K um corpo. Então, $\mathcal{M}_n(K)$ é uma K-álgebra de dimensão finita $\dim_K(\mathcal{M}_n(K)) = n^2$, tendo como base $\{E_{ij} : 1 \leq i, j \leq n\}$, onde E_{ij} é a matriz que possui 1_K na entra (i, j) e 0 nas demais.

Exemplo 2.4.3. Todo anel A é uma $\mathbb{Z}$-álgebra. Podemos definir a operação de multiplicação por escalar como

$$\begin{array}{rccc} \cdot & : & \mathbb{Z} \times A & \longrightarrow & A \\ & & (\lambda, a) & \longmapsto & \begin{cases} \underbrace{a + a + \ldots + a}_{\lambda \text{ vezes}}, & \text{se } \lambda > 0, \\ 0_A, & \text{se } \lambda = 0, \\ \underbrace{(-a) + (-a) + \ldots + (-a)}_{|\lambda| \text{ vezes}}, & \text{se } \lambda < 0. \end{cases} \end{array}$$

É fácil ver que os axiomas da definição são satisfeitos.

Exemplo 2.4.4. Seja $\mathbb{H} = \mathbb{R}^4$. Denotamos:

$$1 = (1, 0, 0, 0) \quad i = (0, 1, 0, 0) \quad j = (0, 0, 1, 0) \quad k = (0, 0, 0, 1)$$

Definimos uma multiplicação em $\mathbb{H}$ satisfazendo:

- $1\alpha = \alpha 1 = \alpha, \forall \alpha \in \mathbb{H}$;

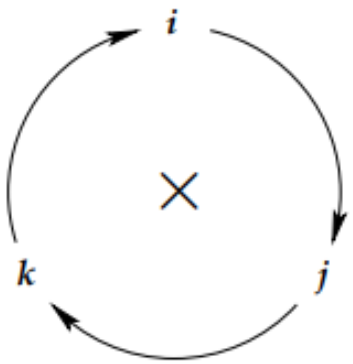


Figura 2.1: *Multiplicação em* $\mathbb{H}$ *e o produto vetorial* $\times$ *em* $\mathbb{R}^3$.

- $i^2 = j^2 = k^2 = -1$;
- $ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j$

Então $\mathbb{H}$ é uma $\mathbb{R}$-álgebra, conhecida como Quatérnios.

Exemplo 2.4.5. Seja $R = \mathbb{C}$, e tome G um grupo topológico Hausdorff com a propriedade de que toda vizinhança da identidade contém um subgrupo aberto e compacto de G. Então G é localmente compacto e possui uma medida de Haar μ, que é uma medida de Borel satisfazendo as seguintes condições:

1. $\mu(xE) = \mu(E)$ para todo $x \in G$ e para todo $E \subseteq G$ mensurável;
2. $\mu(U) > 0$ para todo conjunto aberto não-vazio $U \subseteq G$;
3. $\mu(K) < \infty$ para todo conjunto compacto $K \subseteq G$.

O $\mathbb{C}$-espaço vetorial $\mathcal{C}_c^\infty(G)$ das funções $\varphi\colon G \to \mathbb{C}$ localmente constantes e de suporte compacto pode ser entendido como uma $\mathbb{C}$-álgebra, definindo a multiplicação como a convolução:

$$f * g(x) = \int_G f(y)g(y^{-1}x)d\mu(y)$$

Esta é a chamada *álgebra de Iwahori-Hecke*, e usualmente não é comutativa.

2.5 Distância ultramétrica e relações de ordem

Definição 2.5.1. Uma *distância ultramétrica* num conjunto M é uma função $d\colon M \times M \to \mathbb{R}$ tal que, para todo $x, y, z \in M$:

- $d(x, y) \geq 0$;
- $d(x, y) = 0 \Leftrightarrow x = y$;
- $d(x, y) = d(y, x)$;
- $d(x, z) \leq \max\{d(x, y), d(y, z)\}$.

Exemplo 2.5.2. Um exemplo interessante de distância ultramétrica a qual vale a pena citar é a *distância p-ádica*.

Seja p um número primo. Definimos o *anel dos inteiros p-ádicos* $\mathbb{Z}_p$ como o subanel do anel produto

$$\prod_{n\geq 1} \mathbb{Z}/p^n\mathbb{Z} = \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p^3\mathbb{Z} \times \cdots$$

formado por tuplas coerentes $(\alpha_1, \alpha_2, \ldots)$: se $n \geq m$, exigimos que α_n tenha imagem α_m segundo o morfismo natural $\mathbb{Z}/p^n\mathbb{Z} \to \mathbb{Z}/p^m\mathbb{Z}$. Ou seja:

$$\mathbb{Z}_p = \left\{ (\overline{a}_n) \in \prod_{n\geq 1} \mathbb{Z}/p^n\mathbb{Z} \middle| a_m \equiv a_n \pmod{p^m} \text{ se } n \geq m \right\}$$

Daí, um inteiro p-ádico pode ser representado como uma sequência na forma

$$[\ldots, a_n, a_{n-1}, \ldots, a_2, a_1]_p,$$

onde $0 \leq a_i < p$, e p é um número primo.

Definição 2.5.3. Seja p um primo e seja ∞ um símbolo sujeito às regras $\infty + a = \infty$ e $a \leq \infty$ para todo $a \in \mathbb{N} \cup \{\infty\}$.

- A *valoração p-ádica* em $\mathbb{Z}$ é a função $v_p \colon \mathbb{Z} \to \mathbb{N} \cup \{\infty\}$ dada por

$$v_p(n) = \begin{cases} \text{maior } r \in \mathbb{N} \text{ tal que } p^r \mid n, & \text{se } n \neq 0 \\ \infty, \text{ se } n = 0 \end{cases}$$

- A *norma p-ádica* é a função $\|\bullet\|_p : \mathbb{Z} \to \mathbb{R}$ dada por

$$\|n\|_p = p^{-v_p(n)} \quad (n \in \mathbb{Z})$$

  Aqui interpretamos $\|0\|_p = p^{-\infty} = 0$.

  Analogamente, definimos a valoração e a norma p-ádica em $\mathbb{Z}_p$ (que estendem as respectivas funções em $\mathbb{Z}$).

Diretamente das definições acima temos as seguintes propriedades para quaisquer $a, b \in \mathbb{Z}$ (ou $\mathbb{Z}_p$:)

(i)	$v_p(a) = \infty \Leftrightarrow a = 0$	$\|a\|_p = 0 \Leftrightarrow a = 0$
(ii)	$v_p(ab) = v_p(a) + v_p(b)$	$\|ab\|_p = \|a\|_p \cdot \|b\|_p$
(iii)	$v_p(a+b) \geq \min\{v_p(a), v_p(b)\}$	$\|a+b\|_p \leq \max\{\|a\|_p, \|b\|_p\}$

com igualdade em (iii) se $v_p(a) \neq v_p(b) \Leftrightarrow \|a\|_p \neq \|b\|_p$.

Temos então que

$$d_p(a, b) = \|a - b\|_p$$

define uma métrica em $\mathbb{Z}$ e em $\mathbb{Z}_p$, chamada *métrica p-ádica*. Observe que d_p é uma distância ultramétrica, pois, para todo $a, b, c \in \mathbb{Z}$ (ou $\mathbb{Z}_p$):

- $d_p(a, b) = \|a - b\|_p = p^{-v_p(a+b)} \geq p^{\min\{v_p(a), v_p(b)\}} \geq 0$;
- $d_p(a, b) = 0 \Leftrightarrow \|a - b\|_p = 0 \Leftrightarrow a - b = 0 \Leftrightarrow a = b$;
- $d_p(a, b) = \|a - b\|_p = p^{-v_p(a-b)} = p^{-v_p(b-a)} = \|b - a\|_p = d_p(b, a)$;
- $d_p(a, c) = \|a - c\|_p = \|a - b + b - c\|_p = \big\|(a - b) + (b - c)\big\|_p \leq \max\{\|a - b\|_p, \|b - c\|_p\} = \max\{d_p(a, b), d_p(b, c)\}$.

Portanto, d_p é um exemplo de distância ultramétrica em $\mathbb{Z}$ e no anel de inteiros p-ádicos $\mathbb{Z}_p$.

Definição 2.5.4. Dizemos que uma relação $\preceq$ em um conjunto I é uma *pré-ordem* se ela satisfaz os seguintes dois axiomas:

- (Reflexividade) $i \preceq i$ para todo $i \in I$.

- (Transitividade) se $i \preceq j$ e $j \preceq k$, então $i \preceq k$.

Um conjunto pré-ordenado $(I, \preceq)$ é chamado de *direcionado* se, para quaisquer $i, j \in I$, existe $k \in I$ tal que $i \preceq k$ e $j \preceq k$ (ou seja, existe k majorando dois elementos i e j quaisquer).

Dois exemplos de pré-ordem que serão utilizados neste trabalho são:

Exemplo 2.5.5. Os conjuntos de submódulos finitamente gerados de um módulo M, pré-ordenados pela inclusão. Dados dois submódulos N e P, temos que $N + P$ é finitamente gerado e majora N e P.

Exemplo 2.5.6. Os elementos de um conjunto multiplicativo S de um anel (como definido em 2.1.25, pré-ordenados pela relação de divisibilidade $\mid$ em S (isto é,

$$s \prec t \Leftrightarrow s \mid t \Leftrightarrow \exists u \in S \ : \ t = su$$

Dados dois elementos $t, s \in S$, temos que o produto st majora s e t.

Definição 2.5.7. Uma *ordem parcial* é uma relação binária em um conjunto X se satisfaz as seguintes propriedades:

- Reflexividade: $\forall x \in X(x \leq x)$
- Antissimetria: $\forall x, y \in X((x \leq y \wedge y \leq x) \Rightarrow x = y)$
- Transitividade: $\forall x, y, z \in X((x \leq y \wedge y \leq z) \Rightarrow x \leq z)$

Note que a relação de pré-ordem definida no exemplo 2.5.5 é uma relação de ordem parcial (isto é, vale a propriedade antissimétrica), mas no exemplo 2.5.6, isso não acontece: se S é um conjunto multiplicativo de $\mathbb{Z}$ formado por todos os elementos não nulos, temos que $2 \mid -2$ e $-2 \mid 2$, mas $2 \neq -2$.

Definição 2.5.8. Uma *ordem total* $\leq$ é uma relação binária em um conjunto X se satisfaz as seguintes propriedades, $\forall\ a, b, c \in A$:

- Antissimetria: se $a \leq b$ e $b \leq a$, então $a = b$;
- Transitividade: se $a \leq b$ e $b \leq c$, então $a \leq c$;
- Totalidade: $a \leq b \vee b \leq a$.

Note que a relação de totalidade implica a relação de reflexividade; assim, a ordem total é também uma ordem parcial.

Capítulo 3

Séries Formais

Apresentamos aqui as noções básicas sobre séries formais e seus principais tópicos.

Na seção 3.1, definimos e estudamos as propriedades de elementos básicos da constitutição de séries formais.

Na seção 3, definimos e estudamos o semianel das séries formais, e também definimos o conceito de série racional.

Na seção 3.3, exploramos outras representações para as séries formais, conceitualizando a noção de séries reconhecíveis, e demonstramos o Teorema de Schützenberger 3.3.12, que trata da equivalência dos conceitos de séries reconhecíveis e racionais. São tratados também resultados relacionados a ideias sintáticos de séries formais, que serão utilizados nos próximos capítulos, sobretudo na demonstração dos resultados concernentes à não-limitação da altura de estrela em 5.3. Tratamos também de representações lineares minimais para séries formais, que simplificam demonstrações de diversos resultados que irão se suceder posteriormente.

Continuando em direção ao estudo de representações de séries formais, apresentamos noções básicas sobre autômatos ponderados na seção 3.4, com o fim de utilizá-los para caracterizar séries a partir de suas representações lineares.

3.1 Alfabeto, palavras e linguagens

Definição 3.1.1. Um *alfabeto* é um conjunto finito e não-vazio de elementos, os quais denominamos *símbolos* ou *letras* do alfabeto. Uma *palavra* é uma sequência de letras de um alfabeto. A palavra formada pela combinação de nenhuma letra do alfabeto é chamada *palavra vazia*, e indicada por ε.

O conjunto de todas as palavras que podem ser formadas a partir de certo alfabeto A é denominado *fechamento* de A, e representado como A^*. O *fechamento reflexivo* de A, indicado por A^+, é dado por $A^+ = A^* \setminus \{\varepsilon\}$.

A partir do conjunto definido, vamos introduzir uma operação.

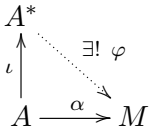
Definição 3.1.2. Sejam $a_1 \cdots a_n, b_1 \cdots b_p \in A^*$. Define-se então uma operação $\star$, dada por

$$(a_1 \cdots a_n) \star (b_1 \cdots b_p) = a_1 \cdots a_n b_1 \cdots b_p,$$

chamada *concatenação*, com o elemento neutro sendo a sequência nula ε. Quando não houver ambiguidade, o símbolo $\star$ será suprimido.

A operação definida acima nos permite observar propriedades algébricas interessantes.

Proposição 3.1.3. *Para qualquer função $\alpha\colon A \to M$, onde M é um monoide, existe um único morfismo de monoides $\varphi\colon A^* \to M$ tal que o diagrama seguinte é comutativo:*



onde ι é a inclusão natural de A em A^*.

Demonstração. Veja [16].

□

Por causa dessa propriedade (chamada de *propriedade universal*), o conjunto A^* é chamado de *monoide livre* gerado por A, sendo um monoide com a operação de concatenação.

Temos também que A^+ é o semigrupo livre sobre A, e é um semigrupo considerando a operação de concatenação de palavras.

Definição 3.1.4. Dada uma palavra $\omega = a_1 \cdots a_n \in A^*$, $a_i \in A$, denotaremos o *tamanho* ou *comprimento* de ω como a quantidade de letras utilizadas para a formação desta. Indicaremos como

$$|\omega| = |a_1 \cdots a_n| = n$$

o número de ocorrências de determinada letra α do alfabeto A como $|\omega|_\alpha$.

Definição 3.1.5. Seja $\omega \in A^*$. Então, define-se a *potência* de uma palavra como

$$\omega^n = \underbrace{\omega\omega\ldots\omega}_{n \text{ vezes}}, \ \forall n \in \mathbb{N}.$$

Se $n = 0$, então $\omega^0 = \varepsilon$.

Nota-se que $|w \cdot \omega| = |w| + |\omega|$ e, sendo o alfabeto $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, composto por n letras $\alpha_1, \alpha_2, \ldots \alpha_n$ e $\omega \in A^*$, então

$$|\omega| = \sum_{i=1}^{n} |\omega|_{\alpha_i}.$$

Podemos também agrupar conjuntos de palavras formadas a partir de uma certa regra em comum. A esses conjuntos, daremos o nome de linguagens.

Definição 3.1.6. Uma *linguagem* é um subconjunto de A^*, onde A é um alfabeto.

3.2 O semianel das séries formais e séries racionais

3.2.1 Definição e propriedades básicas

Iremos agora definir alguns conceitos fundamentais de séries formais.

Definição 3.2.1. Uma *série formal* S é uma função

$$A^* \to R$$

Onde A^* é o fechamento de um alfabeto A e R é um semianel, como na definição 2.2.3.

Em outras palavras, uma série formal é uma função que tem como objetivo associar a cada palavra do fechamento do alfabeto um elemento do semianel R.

Definição 3.2.2. A imagem por uma série formal S de uma palavra $w \in A^*$ é denominada *coeficiente* de w em S e utilizamos a seguinte notação: (S, w).

O *suporte* de uma série formal S é dado por

$$\text{Supp}(S) = \{w \in A^* | (S, w) \neq 0\}$$

O *núcleo* de uma série formal S é dado por

$$\text{Ker}(S) = \{w \in A^* | (S, w) = 0\}$$

Dizemos que uma série formal S possui *suporte finito* se

$$|\mathrm{Supp}(S)| < \infty.$$

É importante observar que o suporte e o núcleo de uma série formal são linguagens de A. Como o suporte é uma linguagem, este pode ser finito ou infinito. Claramente $\mathrm{Supp}(S) \cup \mathrm{Ker}(S) = A^*$ e $\mathrm{Supp}(S) \cap \mathrm{Ker}(S) = \emptyset$.

Por questão de notação, denotaremos o conjunto de séries formais de A com coeficientes em R de $R\langle\langle A\rangle\rangle$.

Naturalmente, introduziremos operações $+$ e $\cdot$ em $R\langle\langle A\rangle\rangle$ que nos permitirá obter propriedades interessantes sobre a tripla formada $(R\langle\langle A\rangle\rangle, +, \cdot)$.

Proposição 3.2.3. *Seja a tripla $(R\langle\langle A\rangle\rangle, +, \cdot)$, em que definiremos as operações $+$ e $\cdot$ do seguinte modo: dados $S, T \in R\langle\langle A\rangle\rangle$, definimos:*

- ***Adição**: a série $S + T$ como a única tal que, para toda palavra $\omega \in A^*$, tenhamos:*

$$(S + T, \omega) = (S, \omega) + (T, \omega)$$

- ***Multiplicação**: a série ST como a única tal que, para toda palavra $\omega \in A^*$, tenhamos:*

$$(ST, \omega) = \sum_{xy=\omega} (S, x)(T, y)$$

Então, o conjunto $R\langle\langle A\rangle\rangle$, munido das operações $+$ e $\cdot$ definidas acima confere a este uma estrutura de semianel.

Além da adição e multiplicação citadas acima, pode-se definir também outras duas operações externas, que denotaremos por *multiplicação por escalar à direita* e *multiplicação por escalar à esquerda*. Se S é uma série formal e $r \in R$, definimos a série rS como a única que satisfaz:

$$(rS, \omega) = r(S, \omega)$$

e definimos a série Sr, por sua vez, como a única que satisfaz:

$$(Sr, \omega) = (S, \omega)r$$

Note que estamos definindo a multiplicação por escalar à direita e à esquerda pelo fato de não estarmos impondo que a operação de multiplicação é comutativa em R. Naturalmente, se R é um semianel comutativo, podemos omitir o lado em que estamos multiplicando, já que nesse caso

$$(rS, \omega) = r(S, \omega) = (S, \omega)r = (Sr, \omega)$$

Podemos também definir outros tipos de operações em $R\langle\langle A\rangle\rangle$, como *produto shuffle* e o *produto de Hadamard*, importantes para o estudo de algumas estruturas algébricas. Para mais detalhes, veja [16] e [28].

Seguindo a definição 2.3.1, podemos ver que o semianel $R\langle\langle A\rangle\rangle$ das séries formais é um R-módulo à esquerda, onde a multiplicação por escalar $R \times R\langle\langle A\rangle\rangle$ definida em 3.2.3.

Definição 3.2.4. O *termo constante* de uma série formal S é definido como o coeficiente da palavra vazia, ou seja, (S, ε). Uma série formal $S \in R\langle\langle A\rangle\rangle$ é *própria* se esta não possui termo constante.

Se R é um corpo, então o conjunto das séries formais próprias é o único ideal maximal de $R\langle\langle A\rangle\rangle$.

Proposição 3.2.5. *A função*

$$\begin{array}{rccl} \varphi & : & R\langle\langle A\rangle\rangle & \longrightarrow & R \\ & & S & \longmapsto & \varphi(S) = (S, \varepsilon) \end{array}$$

é um morfismo de semianéis.

Demonstração. Temos as seguintes igualdades:

- $\varphi(S+T) = (S+T,\varepsilon) = (S,\varepsilon) + (T,\varepsilon) = \varphi(S) + \varphi(T),\ \forall\ S,T \in R\langle\langle A\rangle\rangle$
- $\varphi(ST) = (ST,\varepsilon) = \sum_{uv=\varepsilon} (S,u)(T,v) = (S,\varepsilon)(T,\varepsilon) = \varphi(S)\varphi(T),\ \forall\ S,T \in R\langle\langle A\rangle\rangle$
- $\varphi(0_{R\langle\langle A\rangle\rangle}) = (0_{R\langle\langle A\rangle\rangle},\varepsilon) = 0_R$
- $\varphi(1_{R\langle\langle A\rangle\rangle}) = (1_{R\langle\langle A\rangle\rangle},\varepsilon) = 1_R$

□

Definição 3.2.6. Uma série formal $S \in R\langle\langle A\rangle\rangle$ com suporte finito é chamada de *polinômio.*

Por questão de notação, o *conjunto de polinômios* será denotado por $R\langle A\rangle$. Note que $R\langle A\rangle$ é um subsemianel de $R\langle\langle A\rangle\rangle$. Veja que $R\langle A\rangle$ é uma R-álgebra livre sobre A, conforme o seguinte teorema:

Teorema 3.2.7 (Propriedade Universal de $R\langle A\rangle$)**.** *Sejam R um anel comutativo, A um conjunto não-vazio e $R\langle A\rangle$ o conjunto de polinômios sobre A. Dada uma R-álgebra K, e uma função $j\colon A \to K$, existe um único homomorfismo de R-álgebras $\varphi\colon R\langle A\rangle \to K$ tal que*

$$\varphi(a) = j(a), \quad \forall\ a \in A,$$

onde identificamos A com sua imagem pela aplicação (injetora) $A \ni a \mapsto 1a \in R\langle A\rangle$.

Definição 3.2.8. O *grau* de um polinômio é o tamanho da maior palavra em seu suporte se este é não-nulo, e é $-\infty$ se o polinômio é nulo (i.e. a série formal nula 0). Ou seja, o grau de um polinômio S é definido como:

$$\deg(S) = \begin{cases} \max\{|\omega| : \omega \in supp(S)\}, & \text{se } S \neq 0 \\ -\infty, & \text{se } S = 0 \end{cases}$$

Quando $A = \{a\}$ possui apenas um elemento, denotamos o conjunto das séries de potências formais como $R\langle\langle a\rangle\rangle = R[[a]]$, e o conjunto de polinômios como $R\langle a\rangle = R[a]$.

Se R for um anel, então $R[[a]]$ possui algumas propriedades importantes.

Proposição 3.2.9. *Seja R um anel. Então*

1. *O grupo das unidades de $R[[a]]$ é*

$$\mathcal{U}(R[[a]]) = \{r_0 + r_1 t + r_2 t^2 + \ldots \in R[[a]] \mid a_0 \in \mathcal{U}(R)\}$$

2. *Se R é um domínio, então $R[[a]]$ também é um domínio.*

3. *Temos um isomorfismo natural*

$$R[[a]] = \operatorname*{proj\,lim}_{n\in\mathbb{N}} \frac{R[a]}{\langle a^n\rangle} = \left\{ (f_n) \in \prod_{n\in\mathbb{N}} \frac{R[a]}{\langle a^n\rangle} \Big| f_m \equiv f_n \pmod{a^m} \text{ para todo } n \geq m \right\},$$

que leva $r_0 + r_1 a + r_2 a^2 + \cdots$ em seus truncamentos $(r_0 \pmod a, r_0 + r_1 a \pmod{a^2}, r_0 + r_1 a + r_2 a^2 \pmod{a^3}, \ldots)$.

Demonstração. 1. Um elemento $r_0 + r_1 a + r_2 a^2 + \ldots \in R[[a]]$ é uma unidade se, e somente se, a seguinte equação nas variáveis $s_i's$

$$(r_0 + r_1 a + r_2 a^2 + \ldots)(s_0 + s_1 a + s_2 b^2 + \ldots) = 1$$

admite solução, ou seja, se e somente se, o seguinte "sistema triangular infinito" possui solução:

$$\begin{aligned} r_0s_0 &= 1 \\ r_1s_0 + r_0s_1 &= 0 \\ r_2s_0 + r_1s_1 + r_0s_2 &= 0 \\ r_3s_0 + r_2s_1 + r_1s_2 + r_0s_3 &= 0 \\ &\vdots \\ \sum_{i+j=n} r_is_j &= 0 \end{aligned}$$

Assim, se $r_0 + r_1a + r_2a^2 + \cdots$ é unidade então $r_0s_0 = 1 \Rightarrow r_0 \in \mathcal{U}(R)$.

Reciprocamente, se $r_0 \in \mathcal{U}(R)$, podemos recursivamente definir

$$\begin{cases} s_0 = r_0^{-1} \\ s_n = -r_0^{-1}(r_ns_0 + r_{n-1}s_1 + \ldots + r_1s_{n-1}) \quad \text{para} \quad n \geq 1, \end{cases}$$

que é solução do sistema

$$\begin{aligned} r_0s_0 &= 1 \\ r_1s_0 + r_0s_1 &= 0 \\ r_2s_0 + r_1s_1 + r_0s_2 &= 0 \\ r_3s_0 + r_2s_1 + r_1s_2 + r_0s_3 &= 0 \\ &\vdots \\ \sum_{i+j=n} r_is_j &= 0 \end{aligned}$$

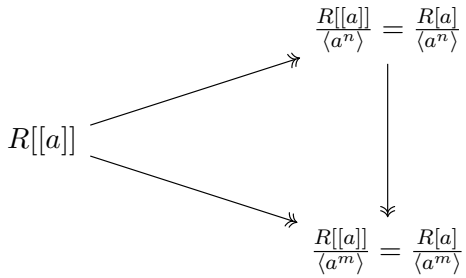
Logo, $r_0 + r_1a + r_2a^2 + \ldots \in \mathcal{U}(R[[a]])$.

2. Dados elementos não nulos $\sum r_na^n$ e $\sum s_na^n$ em $R[[a]]$, sejam i, j mínimos tais que $r_i \neq 0$e $s_j \neq 0$. Então o coeficiente de a^{i+j} no produto

$$\left(\sum r_na^n\right)\left(\sum s_na^n\right)$$

é $r_is_j \neq 0$ (pois R é domínio), o que mostra que este produto de séries formais é não-nulo.

3. Observe que a inclusão $R[a] \hookrightarrow R[[a]]$ induz um isomorfismo de anéis $R[a]/\langle a^n\rangle = R[[a]]/\langle a^n\rangle$: um conjunto de representantes de classe destes quocientes é o conjunto de todos os polinômios de grau menor ou igual a $n-1$. Além disso, para $n \geq m$, temos um diagrama comutativo



Assim, pela propriedade universal do limite projetivo (veja [30] e [23], por exemplo), temos um morfismo de anéis

$$\tau\colon R[[a]] \to \operatorname*{proj\,lim}_{n\in\mathbb{N}} R[a]/\langle a^n\rangle$$

Explicitamente, τ é dado pelo produto dos mapas quocientes

$$R[[a]] \to \prod_{n\in\mathbb{N}} \frac{R[[a]]}{\langle a^n\rangle} = \prod_{n\in\mathbb{N}} \frac{R[a]}{\langle a^n\rangle}$$

cuja imagem consiste em tuplas coerentes pela comutatividade do diagrama.

Vamos mostrar que τ é um isomorfismo.

- Claramente τ é injetor, pois dado elemento não nulo

$$f(a) = \sum r_n a^n \in R[[a]],$$

digamos com $r_i \neq 0$, temos que $f(a) \neq 0 \pmod{a^{i+1}}$.

- τ é sobrejetor: escrevendo um dado elemento $(f_n(a) \pmod a)^n)_{n\in\mathbb{N}} \in \operatorname{proj}\lim_{n\in\mathbb{N}} \frac{R[a]}{\langle a^n\rangle}$ utilizando representantes de classe $f_n(a) \in R[a]$ com $\deg(f_n(a)) < n$, temos que, para $n \geq m$, $f_n(a) \equiv f_n(a) \pmod{a^m}$ implica que $f_m(a)$ é obtido a partir de $f_n(a)$ omitindo-se os monômios de graus maiores ou iguais a m. Assim, os polinômios $f_n(a), n \in \mathbb{N}$, podem ser "colados" em uma série formal $f(a) \in R[[a]]$ com

$$\tau(f(a)) = (f_n(a) \pmod{a^n})_{n\in\mathbb{N}}$$

□

3.2.2 Famílias somáveis

Iremos agora definir o conceito de série racional, a partir da definição de operações racionais sobre o conjunto das séries formais de A com coeficientes em um semianel R e algumas constatações acerca do fecho racional dos subconjuntos de $R\langle\langle A\rangle\rangle$.

Definição 3.2.10. Seja $0 < \sigma < 1$. Uma família $(S_i)_{i\in I}$ é chamada de *somável* em relação a σ se existe uma série formal S tal que, para todo $\varepsilon > 0$, existe um subconjunto finito $I' \subset I$ tal que, para todo subconjunto finito $J \subset I$ contendo I', a seguinte desigualdade é satisfeita:

$$d\left(\sum_{j\in J} S_j, S\right) \leq \varepsilon$$

onde d é a função definida da seguinte maneira:

$$\begin{aligned} d\colon R\langle\langle A\rangle\rangle \times R\langle\langle A\rangle\rangle &\longrightarrow \mathbb{R} \\ (S,T) &\longmapsto d(S,T) = \sigma^{\omega(S,T)} \end{aligned}$$

onde

$$\begin{aligned} \omega\colon R\langle\langle A\rangle\rangle \times R\langle\langle A\rangle\rangle &\longrightarrow \mathbb{N}\cup\{\infty\} \\ (S,T) &\longmapsto \omega(S,T) = \sqcup(S,T) \end{aligned}$$

em que

$$\sqcup(S,T) = \inf\{n \in \mathbb{N} | \exists w \in A^*, |w| = n \ \wedge \ (S,w) \neq (T,w)\}$$

Uma propriedade interessante, cuja demonstração não faremos por ser trabalhosa e entediante, é o seguinte:

Proposição 3.2.11. *Para $0 < \sigma < 1$, então uma família $(S_i)_{i\in I}$ é somável em relação a σ se e somente se é somável em relação a $\frac{1}{2}$.*

Por causa de tal propriedade, definimos o seguinte:

Definição 3.2.12. Uma família $(S_i)_{i\in I}$ é *somável* se e só se é somável em relação a $\frac{1}{2}$.

Definição 3.2.13. Uma família $(S_i)_{i\in I}$ é chamada de *localmente finita* se para toda palavra ω existe apenas uma quantidade finita de índices $i \in I$ tais que $(S_i, \omega) \neq 0$.

Proposição 3.2.14. *Toda família de séries $(S_i)_{i\in I}$ localmente finita é somável.*

Demonstração. Seja $(S_i)_{i\in I}$ uma série localmente finita em $R\langle\langle A\rangle\rangle$.

Para cada inteiro n, seja J_n o conjunto dos $i \in I$ tais que o grau $\deg(S_i)$ é menor ou igual a n. Por hipótese, J_n é finito. Tome

$$T_N = \sum_{i\in J_n} S_i$$

e para todo conjunto finito L que contém J_n, tome

$$T_L = \sum_{i\in L} S_i$$

As séries T_N, T_L e S coincidem para todas as palavras $\omega \in A^*$ de tamanho menor ou igual a n, e temos então que

$$d(T_n, S) \leq 2^{-n} \quad \text{e} \quad d(T_n, T_L) \leq 2^{-n},$$

E daí segue que

$$d(T_L, S) \leq d(T_n, S) + d(T_n, T_L) \leq 2^{-n} + 2^{-n} \leq 2^{-(n-1)}.$$

□

A soma de uma família localmente finita pode ser definida para toda palavra $\omega \in A^*$ pela relação

$$(S, \omega) = \sum_{i\in I} (S_i, \omega).$$

Note que o suporte dessa soma é finito pois a série (S_i) é localmente finita.

A recíproca não é verdadeira, ou seja, nem toda família de séries somável é localmente finita.

Exemplo 3.2.15. Como exemplo de família somável que não é localmente finita, considere $\mathbb{B}$ o semianel booleano e considere, para todo número n natural $S_n = 1$. Então, a família de séries $(S_n)_{n\in\mathbb{N}}$ é somável, mas não é localmente finita. De fato, como

$$(S_i, \omega) = 1 \neq 0 \ \forall i \in \mathbb{N},$$

e sabemos que $\mathbb{N}$ possui infinitos elementos,[1] segue que

$$|\{i \in I | (S_i, \omega) \neq 0\}| = \infty$$

Logo, é claro que $(S_n)_{n\in\mathbb{N}}$ tal que $S_n = 1 \ \forall n \in \mathbb{N}$ não é localmente finita.

Proposição 3.2.16. *Se $S \in R\langle\langle A\rangle\rangle$ é uma série formal própria, então a família $(S^n)_{n\geq 0}$ é localmente finita.*

Demonstração. Seja $S \in R\langle\langle A\rangle\rangle$ é uma série formal própria. Se $n > |\omega|$, então $(S^n, \omega) = 0$. Consequentemente, esta família é localmente finita.

□

Corolário 3.2.17. *Se $S \in R\langle\langle A\rangle\rangle$ é uma série formal própria, então a família $(S^n)_{n\geq 0}$ é somável.*

Demonstração. Da proposição 3.2.16, sabemos que $(S^n)_{n\geq 0}$ é localmente finita. Como toda série localmente finita é somável pela proposição 3.2.14, segue que a família $(S^n)_{n\geq 0}$ é somável, como queríamos demonstrar. □

[1] Aqui não estamos preocupados com a questão de cardinalidade, mas rigorosamente deveríamos dizer que $|\mathbb{N}| = |\{i \in I | (S_i, \omega) \neq 0\}| = \aleph_0$

Definição 3.2.18. Dizemos que soma da família de séries próprias $(S^n)_{n\geq 0}$ é chamada *estrela de Kleene* de S, ou seja:

$$S^* = \sum_{n\geq 0} S^n.$$

Denotamos a série S^+ como sendo

$$S^+ = \sum_{n\geq 1} S^n$$

Exemplo 3.2.19. Seja F_n o n-ésimo número de Fibonacci, definido recursivamente como:

$$\begin{cases} F_0 = 0, F_1 = 1, \\ F_n = F_{n-1} + F_{n-2}, \ \forall n > 1 \end{cases}$$

Considere a série $S \in \mathbb{N}[x]$, definida por[2]

$$S = \sum_{n=0}^{\infty} F_n x^n$$

Então, a série S é própria, pois $(S, \varepsilon) = (S, x^0) = F_0 = 0$. Vamos calcular S^*. Observe que:

$$S^0 = 1, S^1 = x^1 + 1x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + 13x^7 + \ldots,$$

$$S^2 = x^2 + 2x^3 + 5x^4 + 10x^5 + 20x^6 + 38x^7 + 71x^8 + 88x^9 + 125x^{10} + \ldots$$

$$S^3 = x^3 + 3x^4 + 9x^5 + 22x^6 + 51x^7 + 111x^8 + 233x^9 + 411x^{10} + 714x^{11} + \ldots$$

$$S^4 = x^4 + 4x^5 + 14x^6 + 40x^7 + 105x^8 + 256x^9 + 594x^{10} + 1240x^{11} + 2472x^{12} + \ldots$$

$$S^5 = x^5 + 5x^6 + 20x^7 + 65x^8 + 190x^9 + 511x^{10} + \ldots$$

E assim por diante. Note que $(S^n, \omega) = 0$ para $\omega = x^k$, com $k < n$. Então,

$$S^* = S^0 + S^1 + S^2 + S^3 + S^4 + S^5 + S^6 + S^7 + S^8 + \ldots \Rightarrow$$

$$S^* = 1 + x^1 + 2x^2 + 5x^3 + 12x^4 + 29x^5 + 70x^6 + 169x^7 + 408x^8 + 985x^9 + 2378x^{10} + \ldots \Rightarrow$$

$$S^* = 1 + \sum_{n=1}^{\infty} \frac{\left(1+\sqrt{2}\right)^n - \left(1-\sqrt{2}\right)^n}{2\sqrt{2}} x^n = 1 + \sum_{n=1}^{\infty} P_n x^n$$

onde P_n denota o n-ésimo Número de Pell, definido recursivamente por

$$\begin{cases} P_0 = 0, P_1 = 1, \\ P_n = 2P_{n-1} + P_{n-2}, \ \forall n > 1 \end{cases}$$

Desse modo, temos que

$$S^+ = \sum_{n=1}^{\infty} P_n x^n$$

Proposição 3.2.20. *Seja $S \in R\langle\langle A \rangle\rangle$. Então, são válidas as seguintes relações:*

- $S^* = 1 + S^+$
- $S^+ = SS^* = S^*S$

Demonstração. Temos o seguinte:

$$S^* = \sum_{n\geq 0} S^n = S^0 + \sum_{n\geq 1} S^n = 1 + \sum_{n\geq 1} S^n = 1 + S^*.$$

[2]Trata-se da expansão em série de potências da função geradora da Sequência de Fibonacci, $f(x) = \frac{1}{1-x-x^2}$.

E também:

$$S^+ = \sum_{n\geq 1} S^n = \sum_{n\geq 1} SS^{n-1} = S\left(\sum_{n\geq 1} S^{n-1}\right) = S\left(\sum_{k\geq 0} S^k\right) = SS^*$$

Analogamente:

$$S^+ = \sum_{n\geq 1} S^n = \sum_{n\geq 1} S^{n-1}S = \left(\sum_{n\geq 1} S^{n-1}\right) S = \left(\sum_{k\geq 0} S^k\right) S = S^*S$$

□

Se K é um anel, então S^* é o inverso de $1 - S$, pois

$$S^*(1 - S) = S^* - S^*S = S^* - S^+ = 1$$

As relações da proposição acima serão importantes na demonstração do lema que se segue.

Lema 3.2.21 (Lema de Arden)**.** *Sejam T e U séries formais, sendo T uma série formal própria. Então, a única solução S da equação $S = U + TS$ é a série $S = T^*U$.*

Demonstração. Como temos $T^* = 1 + T^+$ de 3.2.20, então $T^* = 1 + TT^*$ e daí $T^*U = U + TT^*$. Logo, tomando $S = T^*U$, segue que $S = U + TS$. Como a série T é própria, temos que

$$\lim_{n\to\infty} T^n = 0 \quad \text{e} \quad \lim_{n\to\infty} \sum_{i=0}^{n} T^i = T^*$$

Como $S = U + TS$, temos

$$S = U + T(U + TS) = U + TU + T^2S$$

Analogamente,

$$S = U + TU + T^2(U + TS) = U + TU + T^2U + T^3S$$

E por indução, obtemos

$$S = \left(\sum_{i=1}^{n} T^i\right) U + T^{n+1}S$$

Consequentemente,

$$\lim_{n\to\infty} S = \lim_{n\to\infty} \left(\left(\sum_{i=1}^{n} T^i\right) U + T^{n+1}S\right) =$$

$$\left(\lim_{n\to\infty} \sum_{i=1}^{n} T^i\right) U + \left(\lim_{n\to\infty} T^{n+1}\right) S = T^*U + 0 \cdot S \Rightarrow S = T^*U$$

□

Agora, iremos definir o conceito de operações racionais de $R\langle\langle A\rangle\rangle$, que será a base para a conceitualização do fecho racional.

Definição 3.2.22. Seja R um semianel e A um alfabeto. Chamaremos as operações:

- Soma e produto, como definidos em 3.2.3;
- Estrela de Kleene, como definido em 3.2.18;

- multiplicação por escalar à esquerda e multiplicação por escalar à direita ($R\langle\langle A\rangle\rangle \times R \to R\langle\langle A\rangle\rangle$)

de *operações racionais* de $R\langle\langle A\rangle\rangle$.

Um subconjunto de $R\langle\langle A\rangle\rangle$ é *racionalmente fechado* se é fechado com respeito às operações racionais. O menor subconjunto contendo um subconjunto D de $R\langle\langle A\rangle\rangle$ de forma que seja racionalmente fechado é chamado *fecho racional* de D.

Definição 3.2.23. Seja $S \in R\langle\langle A\rangle\rangle$ uma série formal. S é dita *racional* se esta pertence ao fecho racional de $R\langle A\rangle$.

Se R for um anel, então o *fecho racional* de $R\langle A\rangle$ é o menor subanel de $R\langle\langle A\rangle\rangle$ contendo $R\langle A\rangle$ e fechado em relação ao inverso.

3.3 Representação linear

3.3.1 Séries reconhecíveis

Definição 3.3.1. Sendo K um semianel, uma série formal $S \in K\langle\langle A\rangle\rangle$ é dita *reconhecível* se existe um inteiro $n \geq 1$ e um morfismo de monoides [3]

$$\mu\colon A^* \to \mathcal{M}_n(K)$$

e duas matrizes $\lambda \in \mathcal{M}_{1\times n}(K)$ e $\gamma \in \mathcal{M}_{n\times 1}(K)$ tais que, para todas as palavras ω,

$$(S, \omega) = \lambda\mu(\omega)\gamma$$

Neste caso, a tripla (λ, μ, γ) é chamada *representação linear* de S, e n é sua dimensão.

Vamos agora definir uma operação de A^* em $K\langle\langle A\rangle\rangle$.

Definição 3.3.2. Para cada palavra x, e para cada série formal S, vamos associar à série $x^{-1}S$ definida por:

$$x^{-1}S = \sum_{\omega\in A^*} (S, x\omega)\omega$$

Em outras palavras, para todas as palavras x e ω, o coeficiente de ω na série $x^{-1}S$ é $(S, x\omega)$, isto é:

$$(x^{-1}S, \omega) = (S, x\omega)$$

A operação $S \to x^{-1}S$ possui algumas propriedades interessantes.

Proposição 3.3.3. *Sejam $k \in K$, $x, y \in A^*$, $S, T, U \in K\langle\langle A\rangle\rangle$, com U própria e a uma letra. Então, é válido que:*

1. $x^{-1}(S + T) = x^{-1}S + x^{-1}T$;
2. $x^{-1}(kS) = k(x^{-1}S), \quad x^{-1}(Sk) = (x^{-1}S)k$;
3. $(xy)^{-1}S = y^{-1}(x^{-1}S)$;
4. $a^{-1}(ST) = (a^{-1}S)T + (S, \varepsilon)(a^{-1}T)$;
5. $a^{-1}U^* = (a^{-1}U)U^*$.

Demonstração. Sendo (S, ω) o coeficiente de ω em S. Então, para qualquer palavra ω:

[3] Aqui estamos considerando $\mathcal{M}_n(K)$ com a estrutura multiplicativa.

1. $(x^{-1}(S+T),\omega) = (S+T, x\omega) = (S,x\omega) + (T,x\omega) = (x^{-1}S,\omega) + (x^{-1}T,\omega)$.
2. $(x^{-1}(kS),\omega) = (kS,x\omega) = k(S,x\omega) = k(x^{-1}S,\omega)$.
3. $((xy)^{-1}S,\omega) = (S,xy\omega) = (x^{-1}S,y\omega) = (y^{-1}x^{-1}S,\omega)$.
4. Temos que:
$$\begin{aligned}(a^{-1}(ST),\omega) &= (ST,a\omega) = \sum_{uv=a\omega}(S,u)(T,v)\\ &= (S,\varepsilon)(T,a\omega) + \sum_{uv=\omega}(S,au)(T,v)\\ &= (S,\varepsilon)(T,a\omega) + \sum_{uv=\omega}(a^{-1}S,u)(T,v)\\ &= (S,\varepsilon)(a^{-1}T,\omega) + ((a^{-1}S)T,\omega).\end{aligned}$$
5. Observando da proposição 3.2.20 que $U^* = 1+UU^*$, e utilizando o item 4, temos que $a^{-1}U^* = (a^{-1}U)U^*$, já que $(U,\varepsilon) = 0$.

□

Exemplo 3.3.4. Seja $S = a^2 + aba^2 + abab + ab^2 + b$. Para esta série, temos que $(S,\omega) = \mathbb{1}_{\mathcal{O}}(\omega) = [\omega \in \mathcal{O}]$[4] para $\mathcal{O} = \{a^2, aba^2, abab, ab^2, b\}$.

Temos que:
$$(ab)^{-1}S = \sum_{\omega\in A^*}(S,ab\omega)\omega = a^2 + ab + b$$
$$b^{-1}(a^{-1}S) = b^{-1}\left(\sum_{\omega\in A^*}(S,a\omega)\omega\right) = b^{-1}(a + ba^2 + bab + b^2) = a^2 + ab + b$$

Além do que já foi exposto, esta operação de A^* em $K\langle\langle A\rangle\rangle$ é associativa no sentido de que $(xy)^{-1}S = y^{-1}(x^{-1}S)$, como pode ser constatado no exemplo 3.3.4 e na proposição 3.3.3.

Outra fórmula muito importante é a apresentada no seguinte lema:

Lema 3.3.5. *Para $S \in K\langle A\rangle$, tem-se*
$$S = (S,\varepsilon) + \sum_{a\in A} a(a^{-1}S)$$

Demonstração. Quando S é uma palavra, é fácil ver que a fórmula é válida. Estendendo de A^* para $K\langle A\rangle$ por linearidade e continuidade, obtém-se o resultado desejado.

Vamos mostrar o resultado por indução na quantidade de termos de S. Suponha inicialmente que $S = \omega \in A^*$. Assim:
$$(\omega,\varepsilon) + \sum_{a\in A} a(a^{-1}\omega) = 0 + \omega = \omega$$

Agora, suponha que para $S_k = \sum_{i=0}^{k}(S,\omega_i)\omega_i$ a fórmula da proposição é válida, ou seja:
$$S_k = (S_k,\varepsilon) + \sum_{a\in A} a(a^{-1}S_k)$$

[4] $\mathbb{1}_X$ ou χ_X indica a função característica do conjunto X e $[P]$ indica o colchete de Iverson para uma proposição P. Nesse caso, ambos possuem o mesmo significado, e atribuem 1 se ω está em $\mathcal{O}$ ou 0, caso contrário.

Então, para $S_{k+1} = \sum_{i=0}^{k+1} (S, \omega_i)\omega_i$, temos:

$$\begin{aligned}
S_{k+1} &= \sum_{i=0}^{k+1} (S, \omega_i)\omega_i \\
&= \sum_{i=0}^{k} (S, \omega_i)\omega_i + \omega_{k+1} \\
&= S_k + \omega_{k+1} \\
&= (S_k, \varepsilon) + \sum_{a\in A} a(a^{-1}S_k) + (\omega_{k+1}, \varepsilon) + \sum_{a\in A} a(a^{-1}\omega_{k+1}) \\
&= (S_k, \varepsilon) + (\omega_{k+1}, \varepsilon) + \sum_{a\in A} a(a^{-1}S_k) + a(a^{-1}\omega_{k+1}) \\
&= ((S_k + \omega_{k+1}), \varepsilon) + \sum_{a\in A} a(a^{-1}(S_k + \omega_{k+1})) \\
&= (S_{k+1}, \varepsilon) + \sum_{a\in A} a(a^{-1}S_{k+1})
\end{aligned}$$

□

Definição 3.3.6. Um subconjunto M de $K\langle\langle A\rangle\rangle$ é dito *estável* se, para todo $S \in M$ e para todo $x \in A^*$, temos que $x^{-1}S \in M$.

Observe que, pela propriedade 3 da proposição 3.3.3, a definição 3.3.6 é equivalente a dizer que $a^{-1}S \in M$ para cada $a \in A$.

A proposição abaixo nos fornece uma maneira de verificar se uma série formal é reconhecível por meio de subconjuntos estáveis.

Proposição 3.3.7. *Uma série formal $S \in K\langle\langle A\rangle\rangle$ é reconhecível se, e somente se, existe um K-submódulo de $K\langle\langle A\rangle\rangle$ à esquerda finitamente gerado estável que contém S.*

Demonstração. Assuma que S é reconhecível e tome (λ, μ, γ) uma representação linear de S com dimensão n. Considere as séries formais $S_1, S_2, \ldots, S_n$, definidas por

$$(S_i, \omega) = (\mu(\omega)\gamma)_i, \quad i = 1, 2, \ldots, n$$

para todas as palavras ω. Considere

$$\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$$

Seja M o K-módulo à esquerda gerado por $\mathcal{S}$. Portanto M é finitamente gerado. M contém S, pois

$$\begin{aligned}
(S, \omega) &= \lambda\mu(\omega)\gamma \\
&= \sum_{i=1}^{n} \lambda_i(\mu(\omega)\gamma)_i \\
&= \sum_{i=1}^{n} \lambda_i(S_i, \omega),
\end{aligned}$$

o que nos fornece

$$S = \sum_{i=1}^{n} \lambda_i S_i$$

Vamos mostrar agora que M é estável, ou seja, que $S \in M$ e para todo $x \in A^*$, temos que

$x^{-1}S \in M$. Sendo $x \in A^*$, temos

$$\begin{aligned}(x^{-1}S_i, \omega) &= (S_i, x\omega)\\ &= (\mu(x\omega)\gamma)_i\\ &= (\mu(x)\mu(\omega)\gamma)_i\\ &= \sum_{j=1}^{n}(\mu(x))_{i,j}(\mu(\omega)\gamma)_j\\ &= \sum_{j=1}^{n}(\mu(x))_{i,j}(S_j, \omega)\end{aligned}$$

Portanto,

$$x^{-1}S_i = \sum_{j=1}^{n}(\mu(x))_{i,j}S_j \in M$$

Consequentmente, M é estável, pois a função $T \mapsto x^{-1}T$ é K-linear, pelas propriedades expostas na proposição 3.3.3, e manda geradores em M.

Reciprocamente, seja M um K-submódulo estável de $K\langle\langle A\rangle\rangle$ gerado por $S_1, S_2, \ldots, S_n$ e que contém S. Então,

$$S = \sum_{i=1}^{n}\lambda_i S_i$$

para certos $\lambda_i \in K$. Além disso, temos também que para toda letra $a \in A$, existe uma matriz $\mu(a) \in \mathcal{M}_n(K)$ tal que, para todo i,

$$a^{-1}S_i = \sum_{j=1}^{n}(\mu(a))_{i,j}S_j$$

Podemos ver essa equação de maneira matricial. Considere $\mathfrak{M}$, dada por

$$\mathfrak{M} = \begin{pmatrix}(\mu(a))_{1,1} & (\mu(a))_{1,2} & \cdots & (\mu(a))_{1,n}\\ (\mu(a))_{2,1} & (\mu(a))_{2,2} & \cdots & (\mu(a))_{2,n}\\ \vdots & \vdots & \ddots & \vdots\\ (\mu(a))_{n,1} & (\mu(a))_{n,2} & \cdots & (\mu(a))_{n,n}\end{pmatrix}\begin{pmatrix}S_1\\ S_2\\ \vdots\\ S_n\end{pmatrix} \Rightarrow$$

$$\mathfrak{M} = \begin{pmatrix}\sum_{j=1}^{n}(\mu(a))_{1,j}S_j & \sum_{j=1}^{n}(\mu(a))_{2,j}S_j & \cdots & \sum_{j=1}^{n}(\mu(a))_{n,j}S_j\end{pmatrix}$$

Ou seja, $a^{-1}S_i$ corresponde à entrada $\mathfrak{M}_{1,i}$ de $\mathfrak{M}$.

Podemos estender a função $\mu\colon A \to \mathcal{M}_n(K)$ para um morfismo de monoides

$$\begin{array}{rcccl}\mu & : & A^* & \longrightarrow & \mathcal{M}_n(K)\\ & & \omega & \longmapsto & \mu(\omega) = \begin{pmatrix}(\mu(\omega))_{1,1} & (\mu(\omega))_{1,2} & \cdots & (\mu(\omega))_{1,n}\\ (\mu(\omega))_{2,1} & (\mu(\omega))_{2,2} & \cdots & (\mu(\omega))_{2,n}\\ \vdots & \vdots & \ddots & \vdots\\ (\mu(\omega))_{n,1} & (\mu(\omega))_{n,2} & \cdots & (\mu(\omega))_{n,n}\end{pmatrix}\end{array}$$

que por conveniência, continuaremos denotando por mu. Nessas condições, temos a seguinte afirmação:

Afirmação 3.3.8. *Para toda palavra $\omega \in A^*$, vale que*

$$\omega^{-1}S_i = \sum_{j=1}^{n}(\mu(\omega))_{i,j}S_j$$

Demonstração. Vamos provar a afirmação por indução no tamanho da palavra.

Para $\omega = \varepsilon$, observando que

$$(\mu(\varepsilon))_{i,j} S_j = 0 \text{ se } i \neq j,$$

temos que

$$\sum_{j=1}^{n} (\mu(\varepsilon))_{i,j} S_j =$$

$$\underbrace{(\mu(\varepsilon))_{i,1} S_1}_{=0} + \underbrace{(\mu(\varepsilon))_{i,2} S_2}_{=0} + \ldots (\mu(\varepsilon))_{i,i} S_i + \ldots + \underbrace{(\mu(\varepsilon))_{i,n} S_n}_{=0} =$$

$$(\mu(\varepsilon))_{i,i} S_i = \varepsilon^{-1} S_i$$

Logo, a relação é válida para $\omega = \varepsilon$, ou seja, para palavras de comprimento 0.

Se a relação funciona para certa palavra ω com comprimento ℓ então por indução

$$\begin{aligned}
(\omega a)^{-1} S_i &= a^{-1}(\omega^{-1} S_i) \\
&= a^{-1} \left(\sum_{k=1}^{n} (\mu(\omega)_{i,k} S_k \right) \\
&= \sum_{k=1}^{n} (\mu(\omega))_{i,k} (a^{-1} S_k) \\
&= \sum_{k=1}^{n} (\mu(\omega))_{i,k} \left(\sum_{j=1}^{n} (\mu(a))_{k,j} S_j \right) \\
&= \sum_{j=1}^{n} \left(\sum_{k=1}^{n} (\mu(\omega))_{i,k} (\mu(a))_{k,j} \right) S_j \\
&= \sum_{j=1}^{n} (\mu(\omega a))_{i,j} S_j
\end{aligned}$$

Logo, a relação vale para toda palavra $\eta = \omega a$ de tamanho $\ell + 1$. Dessa forma, a equação vale para todas as palavras, e a afirmação está provada. □

Agora, considere $\gamma_j = (S_j, \varepsilon)$, e tome a matriz $\gamma \in \mathcal{M}_{n \times 1}(K)$ definida dessa maneira, ou seja:

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \vdots \\ \gamma_{n-1} \\ \gamma_n \end{pmatrix} = \begin{pmatrix} (S_1, \varepsilon) \\ (S_2, \varepsilon) \\ (S_3, \varepsilon) \\ (S_4, \varepsilon) \\ \vdots \\ (S_{n-1}, \varepsilon) \\ (S_n, \varepsilon) \end{pmatrix}$$

Então,

$$\begin{aligned}
(S_i, \omega) &= (\omega^{-1}, S_i, \varepsilon) \\
&= \left(\sum_{j=1}^{n} (\mu(\omega))_{i,j} S_j, \varepsilon \right) \\
&= \sum_{j=1}^{n} (\mu(\omega))_{i,j} (S_j, \varepsilon) \\
&= \sum_{j=1}^{n} (\mu(\omega))_{i,j} \gamma_j \\
&= (\mu(\omega)\gamma)_i
\end{aligned}$$

Assim, temos que

$$\begin{aligned}
\lambda\mu(\omega)\gamma &= \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} (\mu(\omega))_{1,1} & (\mu(\omega))_{1,2} & \cdots & (\mu(\omega))_{1,n} \\ (\mu(\omega))_{2,1} & (\mu(\omega))_{2,2} & \cdots & (\mu(\omega))_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ (\mu(\omega))_{n,1} & (\mu(\omega))_{n,2} & \cdots & (\mu(\omega))_{n,n} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \vdots \\ \gamma_{n-1} \\ \gamma_n \end{pmatrix} \\
&= \sum_{i=1}^{n} \lambda_i(\mu(\omega)\gamma)_i \\
&= \sum_{i=1}^{n} \lambda_i(S_i, \omega) \\
&= (S, \omega)
\end{aligned}$$

Daí, concluímos que existe um morfismo de monoides

$$\mu\colon A^* \to \mathcal{M}_n(K)$$

e duas matrizes $\lambda \in \mathcal{M}_{1\times n}(K)$ e $\gamma \in \mathcal{M}_{n\times 1}(K)$ tais que, para todas as palavras ω,

$$(S, \omega) = \lambda\mu(\omega)\gamma$$

Isso mostra que (λ, μ, γ) é uma representação linear de dimensão n de S.

Daí, da definição 3.3.1, S é reconhecível.

□

A generalização para matrizes é extremamente útil e pode ser feita sem muitas dificuldades.

Para $M \in \mathcal{M}_n(K)$ e $x \in A$, definimos

$$x^{-1}M = \sum_{\omega\in A^*} (M, x\omega)\omega,$$

onde

$$(M, x\omega) = \begin{pmatrix} (M_{11}, x\omega) & (M_{12}, x\omega) & \cdots & (M_{1n}, x\omega) \\ (M_{21}, x\omega) & (M_{22}, x\omega) & \cdots & (M_{2n}, x\omega) \\ \vdots & \vdots & \ddots & vdots \\ (M_{n1}, x\omega) & (M_{n2}, x\omega) & \cdots & (M_{nn}, x\omega) \end{pmatrix}$$

Definição 3.3.9. Uma matriz $M \in \mathcal{M}_n(K\langle\langle A\rangle\rangle)$ é *própria* se, para todos os índices i e j, a série $m_{i,j}$ é própria. Nesse caso, a *estrela* de M pode ser definida como[5]

$$M^* = \sum_{k\geq 0} M^k$$

É fácil ver que a generalização da proposição 3.2.20 é válida para matrizes próprias:

$$M^* = 1 + MM^*, \tag{3.1}$$

onde 1 é a matriz identidade.

O próximo resultado mostra que as operações racionais atuam de modo "racional" nas matrizes. Esse lema é essencial para o estudo de autômatos, pois nos fornece uma versão direta dos algoritmos de McNaughton–Yamada (MNY) e de Brzozowski-McCluskey (BMC), responsáveis por nos municiar de processos para mostrar que uma linguagem reconhecida por um autômato finito pode ser

[5] A existência de M^* pode ser verificada considerando a topologia produto induzida por $K\langle\langle A\rangle\rangle$ em $K\langle\langle A\rangle\rangle^{n\times n}$.

obtida a partir das letras do alfabeto a partir do uso de uniões, produtos e estrelas de linguagens, ou seja, por uma expressão racional, temas que serão tratados na seção 3.4 e no capítulo 4. Para mais detalhes sobre os algoritmos MNY e BMC, veja [29] e [19].

Lema 3.3.10. *Se $M \in \mathcal{M}_n(K\langle\langle A\rangle\rangle)$ é uma matriz própria, então todos os coeficientes de M^* estão no fecho racional dos coeficientes de M.*

Demonstração. A ideia da demonstração será utilizar indução para provar o resultado.

Se $n = 1$, então o resultado é claro, e já temos nosso caso base. Assuma $n > 1$, e considere as decomposições em blocos

$$M = \left(\begin{array}{c|c} a & b \\ \hline c & d \end{array}\right) \quad \text{e} \quad M^* = \left(\begin{array}{c|c} \alpha & \beta \\ \hline \gamma & \delta \end{array}\right)$$

onde a e d são matrizes quadradas, e os blocos de M^* têm as mesmas dimensões correspondentes aos blocos de M. Da equação (3.1), temos que

$$\alpha = 1 + a\alpha + b\gamma \qquad \beta = a\beta + b\delta$$

$$\gamma = c\alpha + d\gamma \qquad \delta = 1 + c\beta + d\delta$$

Observe que o lema 3.2.21 pode ser estendido para matrizes; portanto, temos que

$$\beta = a^*b\delta \quad \text{e} \quad \gamma = d^*c\alpha.$$

Daí, substituindo as equações,

$$\alpha = 1 + a\alpha + b(d^*c\alpha) = 1 + (a + bd^*c)\alpha$$

$$\delta = 1 + c(a^*b\delta) + d\delta = 1 + (ca^*b + d)\delta$$

Novamente, o lema 3.2.21 nos fornece

$$\alpha = (a + bd^*c)^* \quad \text{e} \quad \delta = (ca^*b + d)^*$$

Finalmente,

$$\beta = a^*b(ca^*b + d)^* \quad \text{e} \quad \gamma = d^*c(a + bd^*c)^*$$

Pela hipótese de indução, todos os coeficientes de a^* e d^* estão no fecho racional dos coeficientes de M. O mesmo ocorre para os coeficientes de $ca^*b + d$ e $a + bd^*c$, e novamente pela hipótese de indução, os coeficientes de α, β, γ e δ estão todos no fecho racional de M^*. □

Exemplo 3.3.11. Vamos calcular a estrela da matriz $M = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$. Para isso, vamos decompor M como soma de duas matrizes de modo que uma delas seja diagonal, isto é,

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & f \end{pmatrix} + \begin{pmatrix} 0 & b & c \\ 0 & 0 & e \\ 0 & 0 & 0 \end{pmatrix}$$

utilizando a identidade $(M+N)^* = (M^*N)^*M$, obtemos:

$$\begin{aligned}
\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}^* &= \left(\begin{pmatrix} a^* & 0 & 0 \\ 0 & d^* & 0 \\ 0 & 0 & f^* \end{pmatrix} \begin{pmatrix} 0 & b & c \\ 0 & 0 & e \\ 0 & 0 & 0 \end{pmatrix} \right)^* \begin{pmatrix} a^* & 0 & 0 \\ 0 & d^* & 0 \\ 0 & 0 & f^* \end{pmatrix} = \\
&= \begin{pmatrix} 0 & a^*b & ac \\ 0 & 0 & d^*e \\ 0 & 0 & 0 \end{pmatrix}^* \begin{pmatrix} a^* & 0 & 0 \\ 0 & d^* & 0 \\ 0 & 0 & f^* \end{pmatrix} = \\
&= \begin{pmatrix} 1 & a^*b & a^*c + a^*bd^*e \\ 0 & 1 & d^*e \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^* & 0 & 0 \\ 0 & d^* & 0 \\ 0 & 0 & f^* \end{pmatrix} = \\
&= \begin{pmatrix} a^* & a^*bd & a^*cf^* + a^*bd^*ef^* \\ 0 & d^* & d^*ef^* \\ 0 & 0 & f^* \end{pmatrix}
\end{aligned}$$

Um importante teorema na teoria de séries racionais é o Teorema de Schützenberger, demonstrado pela primeira vez em [9], que estabelece a equivalência do conceito de série reconhecível e série racional.

Teorema 3.3.12 (Schützenberger)**.** *Uma série formal é reconhecível se, e somente se, é racional.*

Demonstração. Para mostrar que toda série racional é reconhecível, basta verificar que todo polinômio P é reconhecível, e para séries reconhecíveis S e T, $S+T, ST$ e S^* (S própria) também são reconhecíveis. De fato:

- P é reconhecível: Vamos nos valer do resultado obtido na proposição 3.3.7. Se P é um polinômio, então $\omega^{-1}P = 0$ para qualquer palavra ω com $|\omega| > \deg(P)$. Consequentemente, o conjunto $\{\omega^{-1}P | \omega \in A^*\}$ é finito. Como este é estável, ele gera um submódulo estável que é finitamente gerado e também contém P, pois $\varepsilon^{-1}P = P$. Assim, P é reconhecível.

- $S + T$ é reconhecível: Se S e T são reconhecíveis, então existem submódulos finitamente gerados estáveis M e N de $K\langle\langle A\rangle\rangle$ com $S \in M$ e $T \in N$. Então $M + N$ contém $S + T$ e é finitamente gerado e estável, mostrando que $S + T$ é reconhecível.

- ST é reconhecível: Seja o submódulo $P = MT + N$. Temos que $ST \in P$, e de acordo com a equação 4 do lema 3.3.3, P é estável. Além disso, é finitamente gerado, pois M e N são finitamente gerados. Consequentemente, ST é reconhecível.

- S^* é reconhecível: Assumindo S própria, seja o submódulo $Q = K + MS^*$. Então, $S^* = \varepsilon + SS^* \in Q$, e Q é estável, já que pelo lema 3.3.3,

$$a^{-1}(S'S^*) = (a^{-1}S')S^* + (S, \varepsilon)(a^{-1}S)S^* \in Q, \quad \forall S' \in M.$$

  Além disso, Q é finitamente gerado. Portanto, S^* é reconhecível.

Dessa forma, toda série racional é reconhecível.

Reciprocamente, seja S uma série reconhecível e tome (λ, μ, γ) a representação linear de S com dimensão n, como definido em 3.3.1. Considere a matriz própria

$$M = \sum_{a \in A} \mu(a)a \in \mathcal{M}_n(K)\langle\langle A\rangle\rangle$$

Utilizando o isomorfismo natural φ entre $\mathcal{M}_n(K)\langle\langle A\rangle\rangle$ e $\mathcal{M}_n(K\langle\langle A\rangle\rangle)$, dado por

$$\varphi\left(\sum_{\omega\in A^*}\begin{pmatrix} {}^{\omega}m_{11} & {}^{\omega}m_{12} & \cdots & {}^{\omega}m_{1n} \\ {}^{\omega}m_{21} & {}^{\omega}m_{22} & \cdots & {}^{\omega}m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{\omega}m_{n1} & {}^{\omega}m_{n2} & \cdots & {}^{\omega}m_{nn} \end{pmatrix}\omega\right) = \begin{pmatrix} \sum\limits_{\omega\in A^*} {}^{\omega}m_{11}\omega & \sum\limits_{\omega\in A^*} {}^{\omega}m_{12}\omega & \cdots & \sum\limits_{\omega\in A^*} {}^{\omega}m_{1n}\omega \\ \sum\limits_{\omega\in A^*} {}^{\omega}m_{21}\omega & \sum\limits_{\omega\in A^*} {}^{\omega}m_{22}\omega & \cdots & \sum\limits_{\omega\in A^*} {}^{\omega}m_{2n}\omega \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{\omega\in A^*} {}^{\omega}m_{n1}\omega & \sum\limits_{\omega\in A^*} {}^{\omega}m_{n2}\omega & \cdots & \sum\limits_{\omega\in A^*} {}^{\omega}m_{nn}\omega \end{pmatrix},$$

temos que

$$M^* = \sum_{k\geq 0} M^k = \sum_{k\geq 0}\left(\sum_{a\in A}\mu(a)a\right)^k = \sum_{k\geq 0}\sum_{\omega\in A^k}\mu(\omega)\omega = \sum_{\omega\in A^*}\mu(\omega)\omega$$

Desse modo, cada entrada $M^*_{i,j}$ da matriz M^* é da forma

$$M^*_{i,j} = \sum_{\omega}(\mu(\omega))_{i,j}\omega$$

Em vista do lema 3.3.10, $M^*_{i,j}$ é racional. Como

$$S = \sum_{i,j}\lambda_i M^*_{i,j}\gamma_j,$$

a série S é racional, pois cada $M^*_{i,j}$ é racional.

Portanto, toda série reconhecível é racional.

□

Como corolário imediato, temos que

Corolário 3.3.13. *Toda série racional S possui representação linear da forma:*

$$(\lambda, \mu, \gamma)$$

Onde $\lambda \in \mathcal{M}_{1\times n}(K)$ é uma matriz-linha, $\mu : A^* \to \mathcal{M}_n(K)$ é um morfismo de monoides e $\gamma \in \mathcal{M}_{n\times 1}(K)$ é uma matriz-coluna.

3.3.2 Ideais sintáticos

Seguindo [27] e [29], vamos assumir que K é um anel comutativo. A álgebra de polinômios $K\langle A\rangle$ é um K-módulo livre que possui como base o monoide livre A^*. Consequentemente, o conjunto $K\langle\langle A\rangle\rangle$ das séries formais pode ser identificado com

$$(K\langle A\rangle)^* = \{f\colon K\langle A\rangle \to K | f \text{ é linear.}\},$$

chamado de *dual* de $K\langle A\rangle$. Registremos tal fato no

Lema 3.3.14. *Seja $(K\langle A\rangle)^*$ o dual de $K\langle A\rangle$). Então, $(K\langle A\rangle)^*$ é isomorfo a $K\langle\langle A\rangle\rangle$.*

Demonstração. Considere

$$\begin{array}{rccl} \varphi & : & (K\langle A\rangle)^* & \longrightarrow & K\langle\langle A\rangle\rangle \\ & & f & \longmapsto & \varphi(f) = \sum\limits_{\omega\in A^*} f(\omega)\omega \end{array}$$

Claramente φ é um monomorfismo. Essa função também é sobrejetora, pois para $S = \sum\limits_{\omega\in A^*}(S,\omega)\omega \in$

$K\langle\langle A\rangle\rangle$, podemos definir g na base canônica de $(K\langle A\rangle)^*$ de modo que $g(\omega) = (S, \omega)$. Assim,

$$\psi(g) = \sum_{\omega \in A^*} g(\omega)\omega = \sum_{\omega \in A^*} (S, \omega)\omega = S$$

Portanto, $(K\langle A\rangle)^*$ é isomorfo a $K\langle\langle A\rangle\rangle$. □

Cada série formal S define uma função linear

$$\begin{array}{rccl} \varphi_S & : & K\langle A\rangle & \longrightarrow & K \\ & & P & \longmapsto & (S, P) = \sum\limits_{\omega \in A^*} (S, \omega)(P, \omega) \end{array}$$

a soma acima possui suporte finito, pois P é um polinômio. Então, podemos considerar o núcleo de φ_S :

$$\mathrm{Ker}(\varphi_S) = \{P \in K\langle A\rangle | (S, P) = 0\}.$$

Assim, $\varphi_S \in (K\langle A\rangle)^*$. Daí, podemos definir

$$\begin{array}{rccl} \psi & : & K\langle\langle A\rangle\rangle & \longrightarrow & (K\langle A\rangle)^* \\ & & S & \longmapsto & \psi(S) = \varphi_S \end{array}$$

Pode-se verificar que ψ é um morfismo de K-módulos, pois

- $\psi(S + T) = \psi(S) + \psi(T),\ \forall\ S, T \in K\langle\langle A\rangle\rangle$:

$$\begin{array}{rcl} \psi(S + T)(P) & = & (\varphi_S + \varphi_T)(P) \\ & = & \varphi_S(P) + \varphi_T(P) = (S, P) + (T, P) \\ & = & \sum\limits_{\omega \in A^*} (S, \omega)(P, \omega) + \sum\limits_{\omega \in A^*} (T, \omega)(P, \omega) \\ & = & \sum\limits_{\omega \in A^*} (S + T, \omega)(P, \omega) \\ & = & (S + T, P) \\ & = & \varphi_{S+T} \end{array}$$

- $\psi(kS) = k\psi(S),\ \forall\ S \in K\langle\langle A\rangle\rangle, \forall k \in K$:

$$\begin{array}{rcl} \psi(kS)(P) & = & (k\varphi_S)(P) \\ & = & (kS, P) \\ & = & \sum\limits_{\omega \in A^*} (kS, \omega)(P, \omega) \\ & = & k \sum\limits_{\omega \in A^*} (S, \omega)(P, \omega) \\ & = & k\varphi_S(P) \\ & = & k\psi(S) \end{array}$$

Qualquer morfismo multiplicativo $\mu\colon A^* \to \mathfrak{M}$, onde $\mathfrak{M}$ é uma K-álgebra, pode ser estendido univocamente a um morfismo de K-álgebras $K\langle A\rangle \to \mathfrak{M}$. Esta extensão também será denotada por μ. Vamos usar esta convenção tacitamente na sequência desta seção. Além disso, temos:

$$\mu(P) = \sum_{\omega \in A^*} (P, \omega)\mu(\omega)$$

Definição 3.3.15. Seja uma série formal $S \in K\langle\langle A\rangle\rangle$. Dizemos que o maior ideal bilateral de $K\langle A\rangle$ contido em $\mathrm{Ker}(S)$ com respeito à inclusão é o *ideal sintático* de S, sendo denotado por I_S.

Este ideal sempre existe, uma vez que é a soma de todos os ideais contidos em $\mathrm{Ker}(S)$, ou seja:

$$I_S = \sum_{I \subset \mathrm{Ker}(S)} I$$

Definição 3.3.16. A *álgebra sintática* de uma série formal $S \in K\langle\langle A\rangle\rangle$, denotada por $\mathfrak{M}_S$, é a álgebra quociente de $K\langle A\rangle$ pelo ideal sintático I_S de S, ou seja:

$$\mathfrak{M}_S = \frac{K\langle A\rangle}{I_S}.$$

O morfismo canônico $K\langle A\rangle \to \mathfrak{M}_S$ será denotado por μ_S. Como $\mathrm{Ker}(\mu_S) = I_S \subset \mathrm{Ker}(S)$, a série S induz em $\mathfrak{M}_S$ uma função linear denotada por φ_S. Dessa forma, temos que

$$S = \varphi_S \circ \mu_S.$$

Lema 3.3.17. *Seja $\mathfrak{M}$ um K-módulo finitamente gerado à direita,* [6] *e tome φ uma função K-linear em $\mathfrak{M}$. Considere m_0 um elemento de $\mathfrak{M}$ e seja $\nu\colon A^* \to End(\mathfrak{M})$. Então, a série formal*

$$S = \sum_{\omega \in A^*} \varphi(\nu\omega(m_0))\omega$$

é reconhecível. Além disso, se $\mathfrak{M}$ possui um sistema gerador com n elementos, então S admite uma representação linear com dimensão n.

Demonstração. Sejam $m_1, m_2, \dots m_n$ os geradores de $\mathfrak{M}$. Então, para cada letra $a \in A$, e para cada $j \in \{1, 2, \dots n\}$, existem coeficientes $\alpha^a_{i,j} \in K$ tais que

$$\nu a(m_j) = \sum_i m_i \alpha^a_{i,j}.$$

As matrizes $(\alpha^a_{i,j})_{i,j} \in \mathcal{M}_n(K)$ definem uma função

$$\begin{array}{rccc} \mu\colon & A & \to & \mathcal{M}_n(K) \\ & a & \longmapsto & (\alpha^a_{i,j})_{i,j} \end{array},$$

que pode ser estendida para um morfismo $\mu\colon A^* \to \mathcal{M}_n(K)$. Uma indução direta mostra que para qualquer palavra ω,

$$\nu\omega(m_j) = \sum_i m_i(\mu\omega)_{i,j}. \tag{3.2}$$

Sejam $\lambda \in \mathcal{M}_{1\times n}(K)$ e $\gamma \in \mathcal{M}_{n\times 1}(K)$ dados por $\lambda_i = \varphi(m_i)$ e $m_0 = \sum_j m_j\gamma_j$. Desse modo, utilizando tal informação e a equação 3.2, temos:

$$\nu\omega(m_0) = \nu\omega\left(\sum_j m_j\gamma_j\right) = \sum_j \nu\omega(m_j)\gamma_j = \sum_j\sum_i m_i(\mu\omega)_{i,j}\gamma_j,$$

e consequentemente,

$$(S,\omega) = \varphi(\nu\omega(m_0)) = \sum_{i,j} \lambda_i(\mu\omega)_{i,j}\gamma_j = \lambda\mu(\omega)\gamma.$$

Da definição 3.3.1, segue que $S = \sum_{\omega\in A^*} \varphi(\nu\omega(m_0))\omega$ é reconhecível. □

Definição 3.3.18. O *ideal sintático à direita* de uma série formal $S \in K\langle\langle A\rangle\rangle$ é o maior ideal à direita contido em $\mathrm{Ker}(S)$. É denotado por I^r_S.

Vamos agora introduzir uma operação de $K\langle A\rangle$ em $K\langle\langle A\rangle\rangle$ à direita. Como visto em 3.3.14, $K\langle\langle A\rangle\rangle$ é identificado com o dual de $K\langle A\rangle$, de modo que cada endomorfismo f do K-módulo $K\langle A\rangle$

[6] Este resultado é verdadeiro para qualquer semianel K, mesmo que esse não seja comutativo.

define um endomorfismo ${}^t f$ do K-módulo $K\langle\langle A\rangle\rangle$, chamado *morfismo adjunto*, pela relação

$$(S, f(P)) = ({}^t f(S), P) \quad \forall S \in K\langle\langle A\rangle\rangle, P \in K\langle A\rangle.$$

A função $f \longmapsto {}^t f$ é um antimorfismo, de acordo com a definição 2.2.5, ou seja:

$${}^t(g \circ f) = {}^t f \circ {}^t g$$

A partir de um polinômio P, podemos considerar o endomorfismo $Q \longmapsto PQ$ de $K\langle A\rangle$ e seu respectivo morfismo adjunto, denotado por $S \longmapsto S \circ P$. Assim,

$$(S, PQ) = (S \circ P, Q).$$

Em particular, para palavras x, y

$$(S, xy) = (S \circ x, y).$$

A equação acima define a operação $\circ$. Dessas observações, segue como consequência o seguinte:

Lema 3.3.19. *Dada uma série formal $S \in K\langle\langle A\rangle\rangle$, polinômios $P, Q \in K\langle A\rangle$ e uma palavra $\omega \in A^*$, temos que*

1. $S \circ \omega = \omega^{-1} S$;
2. $(S \circ P) \circ Q = S \circ (PQ)$.

Demonstração. 1. Lembrando a operação definida em 3.3.2, sendo suficiente estendê-la por linearidade. segue diretamente, para toda palavra w, que

$$(S \circ \omega, w) = (S, \omega w) = (\omega^{-1} S, w)$$

2. Como sabemos que $(S, PQ) = (S \circ P, Q)$, considere um polinômio R. Advém daí que

$$((S \circ P) \circ Q, R) = (S \circ P, QR) = (S, PQR) = (S \circ (PQ), R) \Rightarrow (S \circ P) \circ Q = S \circ (PQ).$$

□

Corolário 3.3.20. *$K\langle\langle A\rangle\rangle$ é um $K\langle A\rangle$- módulo à direita com a operação $\circ$.*

Demonstração. Segue diretamente do item 2 da proposição 3.3.19, já que as demais propriedades da definição 2.3.1 seguem trivialmente. □

Proposição 3.3.21. *O ideal sintático à direita de uma série S é*

$$I_S^r = \{P \in K\langle A\rangle | S \circ P = 0\}$$

Demonstração. Seja $J = \{P \in K\langle A\rangle | S \circ P = 0\}$. Pelo corolário 3.3.20, vê-se que J é um ideal à direita de $K\langle A\rangle$. Além disso, $J \in \mathrm{Ker}(S)$, pois

$$S \circ P = 0 \implies (S, P) = (S \circ P, \varepsilon) = 0.$$

Ademais, J é o maior ideal à direita com esta propriedade, já que, dado um polinômio P, temos que

$$PK\langle A\rangle \subset \mathrm{Ker}(S) \implies (S \circ P, Q) = (S, PQ) = 0 \quad \forall P, Q \ \in K\langle A\rangle$$

Daí, $S \circ P = 0$. Logo, $J = I_S^r$. □

Corolário 3.3.22. *Temos que $K\langle A\rangle / I_S^r$ é isomorfo a $S \circ K\langle A\rangle$ como um $K\langle A\rangle$-módulo à direita.*

Vamos supor a partir de agora que K é um corpo.

Definição 3.3.23. O *posto* de uma série formal S é a dimensão do espaço $S \circ K\langle A\rangle$.

Definição 3.3.24. A *matriz de Hankel* de uma série formal S é a matriz H indexada por $A^* \times A^*$ com entradas em K definida por

$$H(x, y) = (S, xy), \quad \forall x, y \in A^*.$$

Exemplo 3.3.25. Considere o alfabeto $A = \{a, b\}$, e defina recursivamente a função $\varphi\colon A \to \mathbb{N}$ como $\varphi(\varepsilon) = \varphi(a) = 0$, $\varphi(b) = 1$, e $\varphi(\omega x) = 2\varphi(\omega) + \varphi(x)$, $\forall \omega \in A^*, x \in A$. Em essência, φ codifica a palavra no número natural cuja representação em binário coincide com a palavra ao substituir a por 0 e b por 1. Por exemplo,

$$\begin{aligned} \varphi(bab) &= 2\varphi(ba) + \varphi(b) \\ &= 2(2\varphi(b) + \varphi(a)) + \varphi(b) \\ &= 4\varphi(b) + 2\varphi(a) + 1\varphi(b) = 5 \end{aligned}$$

Vamos analisar a matriz de Hankel da série $S = \sum_{\omega \in A^*} \varphi(\omega)\omega$. A tabela 3.1 mostra as entradas da parte superior esquerda da matriz de Hankel H, com linhas e colounas destacando os índices da matriz. As primeiras linha e coluna repetem os índices. Caso não haja confusão com os índices,

	ε	a	b	aa	ab	ba	bb	aaa	$\cdots$
ε	0	0	1	0	1	2	3	0	$\cdots$
a	0	0	1	0	1	2	3	0	$\cdots$
b	1	2	3	4	5	6	7	8	$\cdots$
aa	0	0	1	0	1	2	3	0	$\cdots$
ab	1	2	3	4	5	6	7	8	$\cdots$
ba	2	4	5	8	9	10	11	16	$\cdots$
bb	3	6	7	12	13	14	15	24	$\cdots$
aaa	0	0	1	0	1	2	3	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Tabela 3.1: *Parte superior esquerda da matriz de Hankel.*

podemos simplesmente escrever a matriz H de maneira usual como

$$H = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots \\ 2 & 4 & 5 & 8 & 9 & 10 & 11 & 16 & \cdots \\ 3 & 6 & 7 & 12 & 13 & 14 & 15 & 24 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Teorema 3.3.26 (Teorema de Paz-Carlyle-Fliess)**.** *O posto de uma série formal S é igual à codimensão de seu ideal sintático a direita, e é equivalente ao posto de sua respectiva matriz de Hankel:*

$$posto(S) = codim(I_S^r) = posto(H)$$

Além disso, a série S é racional se e somente se possui posto finito e, nesse caso,

$$posto(S) = \min\{dim(\lambda, \mu, \gamma) | (\lambda, \mu, \gamma) \text{ é uma representação linear de } S\}.$$

Demonstração. Recorde que o posto de uma matriz (mesmo uma matriz infinita) pode ser definido como sendo a maior dimensão de um subdeterminante não-nulo,e este é igual ao posto das linhas e ao posto das colunas.

Lembrando que a codimensão de $W \subset X$ é $\text{codim}(W) = \dim(X/W)$, a primeira identidade, mais especificamente, $\text{posto}(S) = \text{codim}(I_S^r)$, segue como consequência direta do corolário 3.3.22, pois

$$\text{codim}(I_S^r) = \dim(K\langle A\rangle / I_S^r) = \dim(S \circ K\langle A\rangle) = \text{posto}(S).$$

O espaço $S \circ K\langle A\rangle$ possui um conjunto de geradores $\mathcal{T} = \{S \circ \omega | \omega \in A^*\}$. Assim, $\text{posto}(S)$ é igual ao posto de $\mathcal{T}$. Como cada $S \circ \omega$ pode ser identificado com a linha de índice ω na matriz de Hankel de S, pois, a partir da propriedade 1 do lema 3.3.19, temos que, $\forall x \in A^*$:

$$(S \circ \omega, x) = (\omega^{-1}S, x) = (S, \omega x) = H(\omega, x)$$

Portanto, o posto de S é igual ao posto da matriz de Hankel, ou seja:

$$\text{posto}(S) = \text{posto}(H),$$

o que conclui a primeira parte do teorema.

Vamos agora para a segunda parte do teorema. Se S é racional, então pelo teorema de Schützenberger 3.3.12, S é reconhecível e possui uma representação linear (λ, μ, γ) de dimensão n. O ideal à direita

$$J = \{P \in K\langle A\rangle | \lambda\mu(P) = 0\}$$

está contido em $\text{Ker}(S)$, e sua codimensão é menor ou igual a n. Consequentemente, $J \subset I_S^r$, mostrando que

$$\text{posto}(S) = \text{codim}(I_S^r) \leq \text{codim}(J) \leq n.$$

Logo, se S é racional, então $\text{posto}(S) < \infty$.

Reciprocamente, seja $n = \text{posto}(S) = \dim(S \circ K\langle A\rangle)$. Vamos construir uma representação linear para S, mostrando que esta é reconhecível e, portanto, racional. Para lograr nosso objetivo, seja φ, definida por

$$\begin{array}{rccc} \varphi\colon & S \circ K\langle A\rangle & \longrightarrow & K \\ & T & \longmapsto & (T, \varepsilon). \end{array}$$

Então, para qualquer palavra ω,

$$(S, \omega) = (S \circ \omega, \varepsilon) = \varphi(S \circ \omega).$$

Seja $\mu(\omega)$ a matriz singular dos endomorfismos de $S \circ K\langle A\rangle$ que levam uma série T em $T \circ \omega$, para alguma base de $S \circ K\langle A\rangle$. Cada elemento de $S \circ K\langle A\rangle$ é representado por um vetor $\mathcal{M}_{1\times n}(K)$, e cada endomorfismo de $S \circ K\langle A\rangle$ é representado por uma matriz em $\mathcal{M}_n(K)$; então $\mathcal{M}_n(K)$ exerce uma ação à direita em $\mathcal{M}_{1\times n}(K)$. Em vista da equação 2 do lema 3.3.19

$$(S \circ P) \circ Q = S \circ (PQ),$$

temos que $(\mu(x))(\mu(y)) = \mu(xy)$ para quaisquer palavras x e y. Seja λ o vetor linha representando S em uma base escolhida, e tome γ a coluna representando φ. Então, temos que

$$(S, \omega) = \varphi(S \circ \omega) = \lambda\mu(\omega)\gamma,$$

mostrando que S é reconhecível, e pelo Teorema de Schützenberger 3.3.12, racional, com uma representação linear de dimensão n. Logo, se $\text{posto}(S) < \infty$, a série S é racional, como queríamos demonstrar.

□

Exemplo 3.3.27. Considere a série $S = \sum_{\omega \in A^*} \varphi(\omega)\omega = b + ab + 2ba + 3bb + aab + \ldots$ do exemplo 3.3.36.

Esta série possui uma representação linear (λ, μ, γ) de dimensão 2, definida por

$$\lambda = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \mu(a) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \quad \gamma = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Dessa forma, sabemos que $\text{posto}(S) \leq \dim(\lambda, \mu, \gamma) = 2$. Vamos observar a matriz de Hankel de S, onde os índices estão ordenados pela ordem lexicográfica (a mesma adotada na tabela 3.1 do exemplo 3.3.25, reproduzida abaixo por conveniência):

$$H = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots \\ 2 & 4 & 5 & 8 & 9 & 10 & 11 & 16 & \cdots \\ 3 & 6 & 7 & 12 & 13 & 14 & 15 & 24 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 2 & 3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Observe que as linhas indexadas por a e b na matriz de Hankel são linearmente independentes. Assim, o posto de S é no mínimo 2. Portanto, o posto de S é 2.

Exemplo 3.3.28. Tome $A = \{a, b\}$ e a função φ do exemplo 3.3.25, e considere a série $S = \sum_{\omega \in A^*} (\varphi(\omega))^2\omega = b + ab + 4ba + 9bb + aab + \ldots$

Esta série possui uma representação linear (λ, μ, γ) de dimensão 3, definida por

$$\lambda = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \quad \mu(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 3 & -8 & 6 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 3 \\ 6 & -15 & 10 \end{pmatrix} \quad \gamma = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$$

Assim, $\text{posto}(S) \leq \dim(\lambda, \mu, \gamma) = 3$.

A matriz de Hankel de S, onde os índices estão ordenados pela ordem lexicográfica é:

$$H = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & \cdots \\ 4 & 16 & 25 & 64 & 81 & 100 & 121 & 256 & \cdots \\ 9 & 36 & 49 & 144 & 169 & 196 & 225 & 576 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Observe que podemos extrair da matriz de Hankel acima uma submatriz H' formada pelas linhas e colunas indexadas por ε, b e bb, como mostrado abaixo:

$$H = \begin{pmatrix} \boxed{0} & 0 & \boxed{1} & 0 & 1 & 4 & \boxed{9} & 0 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ \boxed{1} & 4 & \boxed{9} & 16 & 25 & 36 & \boxed{49} & 64 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & \cdots \\ 4 & 16 & 25 & 64 & 81 & 100 & 121 & 256 & \cdots \\ \boxed{9} & 36 & \boxed{49} & 144 & 169 & 196 & \boxed{225} & 576 & \cdots \\ 0 & 0 & 1 & 0 & 1 & 4 & 9 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \Rightarrow H' = \begin{pmatrix} 0 & 1 & 9 \\ 1 & 9 & 49 \\ 9 & 49 & 225 \end{pmatrix}$$

Temos que $\det(H') = -72 \neq 0$, pois as linhas são linearmente indepedentes. Logo, o posto de S é ao menos 3. Portanto, o posto de S é 3.

Pode-se utilizar a função φ do exemplo 3.3.25 para mostrar que existe uma série formal com posto n em $\mathbb{N}\langle\langle A\rangle\rangle$ para todo $n \in \mathbb{N}^*$, bastando verificar que a série formal

$$S = \sum_{\omega \in A^*} (\varphi(\omega))^k \omega \in \mathbb{N}\langle\langle A\rangle\rangle$$

possui posto $k+1$.

3.3.3 Representação linear minimal

Sabemos que as representações lineares de uma série S podem ter diferentes dimensões. Trabalhar com a representação linear de menor dimensão pode simplificar o estudo de propriedades das séries racionais.

Para isso, precisamos garantir que podemos utilizar qualquer representação linear com menor dimensão para os cálculos, ou seja, que existe uma relação entre todas as representações lineares "minimais" de uma série racional S, o que será comprovado no teorema 3.3.34.

Definição 3.3.29. Uma representação linear *minimal* de uma série racional S é uma representação linear de S com menor dimensão entre todas as suas representações possíveis.

Proposição 3.3.30. *Uma representação linear (λ, μ, γ) de dimensão n de uma série S é minimal se, e somente se, tomando $\mathfrak{M} = \mu(K\langle A\rangle)$, temos que*

$$\lambda\mathfrak{M} = \mathcal{M}_{1\times n}(K) \quad e \quad \mathfrak{M}\gamma = \mathcal{M}_{n\times 1}(K)$$

Neste caso,

$$I_S^r = \{P \mid \lambda\mu(P) = 0\}$$

onde I_S^r é o ideal sintático à direita de S, como definido em 3.3.18.

Demonstração. Suponha que (λ, μ, γ) é minimal, e tome $\mathcal{J} = \{P | \lambda\mu(P) = 0\}$. Então $\mathcal{J}$ é um ideal à direita de $K\langle A\rangle$, e $\operatorname{codim}(\mathcal{J}) = \dim(K\langle A\rangle/\mathcal{J}) = \dim(\lambda\mathfrak{M}) \leq n$. Como $\mathcal{J} \subset \operatorname{Ker}(S)$, temos $J \subset I_S^r$, e o Teorema de Paz-Carlyle-Fliess 3.3.26 nos garante que $\operatorname{codim}(\mathcal{J}) \geq \operatorname{codim}(I_S^r) = n$. Consequentemente, $\operatorname{codim}(\mathcal{J}) = n, \mathcal{J} = I_S^r$ e $\lambda\mathfrak{M} = \mathcal{M}_{1\times n}(K)$. A igualdade $\mathfrak{M}\gamma = \mathcal{M}_{n\times 1}(K)$ é obtida analogamente.

Reciprocamente, assuma $\lambda\mathfrak{M} = \mathcal{M}_{1\times n}(K)$ e $\mathfrak{M}\gamma = \mathcal{M}_{n\times 1}(K)$. Então, existem palavras $x_1, \ldots, x_n$ e $y_1, \ldots y_n$ tais que $\lambda\mu(x_1), \ldots \lambda\mu(x_n)$ e $\mu(y_1)\gamma, \ldots \mu(y_n)\gamma$ são bases para $\mathcal{M}_{1\times n}(K)$ e $\mathcal{M}_{n\times 1}(K)$, respectivamente. Desse modo,

$$\det(\lambda(\mu(x_i y_j))\gamma)_{1\leq i,j\leq n} \neq 0$$

Como $\lambda(\mu(x_i y_j))\gamma = (S, x_i y_j)$, a matriz de Hankel H de S é tal que $\operatorname{posto}(H) \geq n$. Pelo Teorema de Paz-Carlyle-Fliess 3.3.26, a representação (λ, μ, γ) é minimal. □

Corolário 3.3.31. *Se (λ, μ, γ) é uma representação linear minimal de dimensão n de uma série formal S, então existem polinômios $P_1, \ldots, P_n, Q_1, \ldots, Q_n$ tais que, para toda palavra ω*

$$\mu(\omega) = ((S, P_i \omega Q_j))_{1 \leq i,j \leq n}$$

Demonstração. Pela proposição 3.3.30, existem polinômios $P_1, \ldots, P_n, Q_1, \ldots, Q_n$ tais que

$$(\lambda\mu(P_i))_{1 \leq i \leq n}$$

é a base canônica de $\mathcal{K}_{1\times n}(K)$, e analogamente, $(\mu(Q_j)\gamma)_{1 \leq j \leq n}$ é a base canônica de $\mathcal{K}_{n\times 1}(K)$. Portanto,

$$(\mu(\omega))_{i,j} = \lambda\mu(P_i)\mu(\omega)\mu(Q_j)\gamma = (S, P_i \omega Q_j)$$

□

Proposição 3.3.32. *Seja K um anel comutativo que é um domínio de ideais principais, e seja F um corpo quociente. Seja $S \in K\langle\langle A \rangle\rangle$ uma série racional de posto n sobre F. Então S é racional sobre K e possui uma representação linear em K de dimensão n, ou seja, S possui uma representação minimal com coeficientes em K.*

Demonstração. Seja (λ, μ, γ) uma representação linear minimal de S sobre F. De acordo com o corolário 3.3.31, existem polinômios $P_1, \ldots, P_n, Q_1, \ldots Q_n \in F\langle A \rangle$ tais que para $\omega \in A^*$,

$$\mu(\omega) = ((S, P_i \omega Q_j))_{1 \leq i,j \leq n}.$$

Seja $d \in K \setminus \{0\}$ tal que $dP_i, dQ_j \in K\langle A \rangle$ e $d\lambda \in \mathcal{M}_{1\times n}(K)$. Então, para todo polinômio $P \in K\langle A \rangle$,

$$d^3\lambda\mu(P) = (d\lambda)(S, dP_i P dQ_j)_{i,j} \in \mathcal{M}_{1\times n}(K),$$

pois $(S, R) \in K$ sempre que $R \in K\langle A \rangle$. Daí,

$$\lambda\mu(K\langle A \rangle) \subseteq \frac{1}{d^3}\mathcal{M}_{1\times n}(K)$$

Isso mostra que $\lambda\mu(K\langle A \rangle)$ é um submódulo de um K-módulo livre de posto n. Consequentemente, $\lambda\mu(K\langle A \rangle)$ também é livre, e seu posto é menor ou igual a n. Pelo lema 3.3.17, obtemos uma representação de S sobre K de dimensão menor ou igual a n. Pelo Teorema de Paz-Carlyle-Fliess 3.3.26, segue a dimensão da representação é n. □

No contexto em que estamos trabalhando, é interessante verificar quais as relações existentes entre duas representações lineares. Assim, temos a seguinte definição:

Definição 3.3.33. Duas representações lineares (λ, μ, γ) e $(\lambda', \mu', \gamma')$ são chamadas *semelhantes* se existe uma matriz m tal que $\lambda' = \lambda m, \mu' = m^{-1}\mu(\omega)m, \gamma' = m^{-1}\gamma$.

Vamos agora enunciar um importante teorema, que revela uma conexão entre as representações minimais possíveis para uma série.

Teorema 3.3.34. *Duas representações lineares minimais são semelhantes.*

Demonstração. Seja (λ, μ, γ) uma representação linear minimal de S. Então, pelas proposições 3.3.21 e 3.3.30,

$$I_S^r = \{P \in K\langle A \rangle | \lambda\mu(P) = 0\} = \{P \in K\langle A \rangle | S \circ P = 0\},$$

os dois $K\langle A \rangle$-módulos $S \circ K\langle A \rangle$ e $\mathcal{M}_{1\times n}(K) = \lambda\mu(K\langle A \rangle)$ (com ação em $\mathcal{M}_{1\times n}$ definida por $(v, P) \longmapsto v\mu(P)$) são isomorfos. Consequentemente, existe um K-isomorfismo $f\colon \mathcal{M}_{1\times n}(K) \to S \circ K\langle A \rangle$ tal que, para cada polinômio P, e para $v \in \mathcal{M}_{1\times n}(K)$,

$$f(v\mu(P)) = f(v) \circ P.$$

e, além disso,

$$f(\lambda) = S.$$

Considere a função linear φ em $S \circ K\langle A\rangle$ definida por $\varphi(T) = (T, \varepsilon)$. Então para $v = \lambda\mu(P)$, obtemos

$$\varphi(f(v)) = \varphi(f(\lambda\mu(P))) = \varphi(f(\lambda) \circ P) = \varphi(S \circ P) = (S \circ P, \varepsilon) = (S, P) = \lambda\mu(P)\gamma = v\gamma,$$

nos revelando que $\varphi \circ f = \gamma$, se γ for tomada como a função linear $v \mapsto v\gamma$.

Se $(\lambda', \mu', \gamma')$ é outra representação linear minimal de S, então existe um isomorfismo ϕ, analogamente à f. Desse modo, existe um isomorfismo

$$\psi = \phi^{-1} \circ \phi \colon \mathcal{M}_{1\times n}(K) \to \mathcal{M}_{1\times n}(K)$$

tal que

$$\psi(v\mu'(P)) = \psi(v)\mu(P), \quad \psi(\lambda') = \lambda, \quad \gamma' = \gamma \circ \psi.$$

Basta escrever ψ em sua forma matricial para obter o resultado desejado. □

Corolário 3.3.35. *Sejam (λ, μ, γ) e $(\lambda', \mu', \gamma')$ duas representações lineares de um série S e assuma que a segunda representação é minimal. Então existe uma representação $(\bar\gamma, \bar\mu, \bar\gamma)$ similar a (λ, μ, γ) tendo uma decomposição em blocos da forma*

$$\bar\lambda = \begin{pmatrix} \times & \lambda' & 0 \end{pmatrix}, \ \bar\mu = \begin{pmatrix} \mu_1 & 0 & 0 \\ \times & \mu' & 0 \\ \times & \times & \mu_2 \end{pmatrix}, \ \bar\gamma = \begin{pmatrix} 0 \\ \gamma' \\ \times \end{pmatrix}$$

Demonstração. Veja [27]. □

Exemplo 3.3.36. Considere o alfabeto $A = \{a, b\}$ e a função φ definida em 3.3.25. Utilizemos a mesma série formal $S = \sum_{\omega\in A^*} (\varphi(\omega))^2\omega \in \mathbb{N}\langle\langle A\rangle\rangle$ do exemplo 3.3.28. Esta série admite duas representações lineares minimais (λ, μ, γ) e (η, κ, ξ), ambas de dimensão 3, definidas como:

$$\lambda = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \quad \mu(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{pmatrix} \quad \gamma = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\eta = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \quad \kappa(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \kappa(b) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix} \quad \xi = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Considere $U = \sum_{\omega\in A^*} (\mu(\omega)\gamma)\omega$ e $V = \sum_{\omega\in A^*} (\kappa(\omega)\xi)\omega$. Se tomarmos apenas as colunas de U e V com índice ε, b e bb, então U e V são representados por matrizes

$$U = \begin{pmatrix} 0 & 1 & 9 \\ 0 & 2 & 12 \\ 1 & 4 & 16 \end{pmatrix} \quad \text{e} \quad V = \begin{pmatrix} 0 & 1 & 9 \\ 0 & 4 & 24 \\ 1 & 4 & 16 \end{pmatrix}.$$

É fácil ver que a matriz P tal que $UP = V$ é

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

E podemos verificar que

$$\mu(b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = P^{-1}\kappa(b)P \Rightarrow \mu(b) = P^{-1}\kappa(b)P$$

Trivialmente, vemos que $\lambda = P\eta$, $\mu(a) = P^{-1}\kappa(a)P$ e $\gamma = P^{-1}\xi$. Dessa forma, concluímos que as representações lineares minimais de S são semelhantes.

O lema seguinte será importante para concluir a demonstração do lema 5.3.1, que é imprescindível para a demonstração do teorema 5.3.2, o mais importante dessa primeira parte.

No entanto, sua demonstração utiliza o teorema de Kleene e conceitos de semigrupos de matrizes, que estão brevemente apresentados no apêndice A para preencher as eventuais lacunas. Desse modo, não iremos reproduzí-la aqui, mas sim no apêndice A.

Lema 3.3.37. *Seja K um semianel comutativo. Tome*

$$\mu = \begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix}$$

um morfismo $A^* \to \mathcal{M}_n(K)$, onde μ' e μ'' são morfismos. Toda série reconhecida por μ é uma combinação linear de séries reconhecidas por μ' ou μ'' e de séries da forma $S'aS''$, onde S' é reconhecida por μ', $a \in A$ e S'' é reonhecida por μ''.

3.4 Autômatos ponderados

Vamos agora nos voltar ao estudo dos autômatos ponderados e compreender como utilizá-lo para caracterizar séries a partir de suas representações lineares.

Definição 3.4.1. Um *autômato ponderado*, também chamado de *autômato com pesos*[7] é uma sêxtupla $\mathcal{A} = (A, K, Q, I, E, T)$, onde:

- A é um alfabeto;
- K é um conjunto de pesos (geralmente um semianel);
- Q é um conjunto de um ou mais elementos denominados estados;
- $I\colon Q \to K$ é a função estado inicial;
- $E\colon Q \times A \times Q \to K$ é uma função chamada função de transição;
- $T\colon Q \to K$ é a função de estados finais.

Uma tripla (p, a, q) tal que $E(p, a, q) \neq 0$ é chamada *aresta*, p e q são respectivamente estados inicial e final, a letra a é seu *rótulo* e $E(p, a, q)$ é seu *peso*. Um *caminho* é uma sequência

$$c = (q_0, a_1, q_1)(q_1, a_2, q_2)\dots(q_{n-1}, a_n, q_n)$$

de arestas. O peso de um caminho c é o produto

$$E(c) = \prod_{i=1}^{n} E(q_{i-1}, a_i, q_i) = E(q_0, a_1, q_1) \cdot E(q_1, a_2, q_2) \cdot \ldots \cdot E(q_{n-1}, a_n, q_n)$$

dos pesos de cada aresta. Seu rótulo é a palavra $\omega = a_0 a_1 \dots a_n$.

Cada estado é um vértice, e cada aresta carrega uma expressão ka, onde k é o peso e a é seu rótulo. Sempre que o peso é 1, este é omitido.

[7] em inglês, *weighted automata*

Exemplo 3.4.2. Seja $\mathcal{A} = (A, K, Q, I, E, T)$. Considere o alfabeto $A = \{a, b\}$, e o semianel $K = \mathbb{N}$. Tome $Q = \{1, 2\}$. Defina $I(p) = 2 - p, p \in Q$, $T(q) = 1 + (-1)^q, q \in Q$, e

$$E(p, \alpha, q) = \begin{cases} 1, & \text{se } p = q; \\ 0, & \text{se } p \neq q \text{ e } \alpha = b; \\ 1 + (-1)^p, & \text{se } p \neq q \text{ e } \alpha = a \end{cases}$$

O diagrama de estados de tal $\mathbb{N}$-autômato está retratado na figura 3.1.

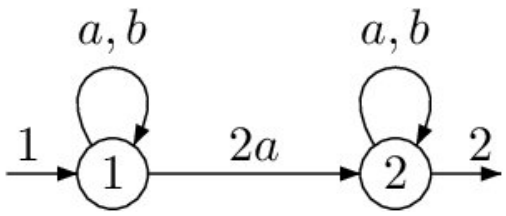


Figura 3.1: *Diagrama de estados de um autômato ponderado.*

Exemplo 3.4.3. Seja o autômato ponderado $\mathcal{A} = (\{a\, b\}, \mathbb{B}, \mathbb{Z}, I, E, T)$,onde $I(x) = T(x) = \delta_{0x}$, e a função E está definida como

$$E(p, \alpha, q) = \begin{cases} 1, & \text{se } \alpha = a \text{ e } q = p \pm 1; \\ 0, & \text{caso contrário.} \end{cases}$$

Este autômato ponderado possui um número infinito de estados, e está representado na figura 3.2.

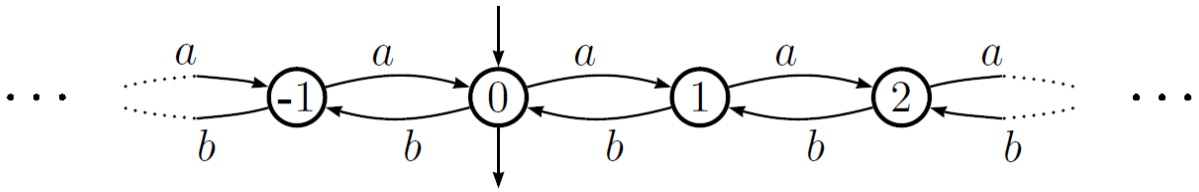


Figura 3.2: *Autômato ponderado com infinitos estados.*

Exemplo 3.4.4. Um caso particular de autômato ponderado é o *autômato finito determinístico*, onde a função E é tratada como uma função característica, e K é o semianel booleano.

Por exemplo, o autômato finito determinístico $\mathcal{D}_6^2 = (A, K, Q, I, E, T)$, onde:

- $A = \{0, 1\}$
- $K = \mathbb{B} = \{\overline{0}, \overline{1}\}$;
- $Q = \{0, 1, 2, 3, 4, 5\}$;
- $I(q) = \begin{cases} 1, & \text{se } q = 0; \\ 0, & \text{caso contrário.} \end{cases}$;
- $E(p, a, q) = \begin{cases} \overline{1}, & \text{se } q = 2p + a \pmod 6; \\ 0, & \text{caso contrário.} \end{cases}$;
- $T(q) = \begin{cases} 1, & \text{se } q = 0; \\ 0, & \text{caso contrário.} \end{cases}$;

O autômato $\mathcal{D}_6^2$ determina se, dada uma sequência de 0's e 1's, o número representado por ele na base 2 é divisível por 6. Podemos construir o seguinte diagrama de estados para este autômato (lembramos aqui que nesse caso específico, o peso dado a cada aresta pode ser suprimido, por se tratar da função característica):

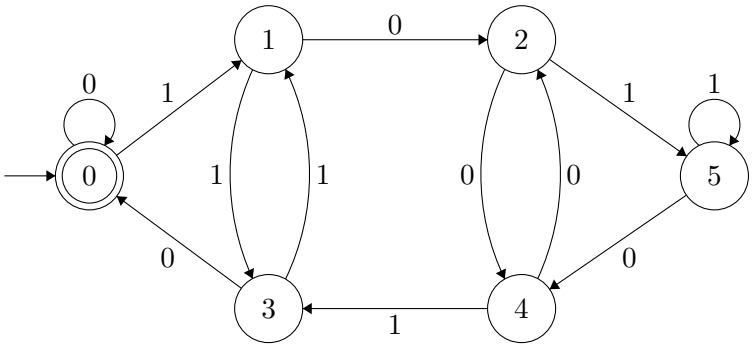


Figura 3.3: *Diagrama de estados para o autômato* $\mathcal{D}_6^2$

Analogamente, pode-se construir um autômato finito determinístico $\mathcal{D}_k^b$ que determina se um número representado na base b é divisível por k. Nessa mesma linha, o diagrama de estados representado em 3.4 verifica se a expansão binária corresponde a um número divisível por 8, correspondendo ao autômato finito determinístico $\mathcal{D}_8^2$.

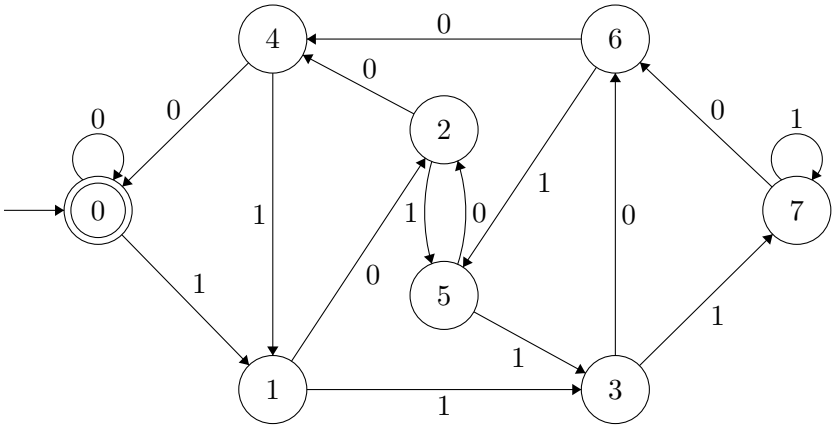


Figura 3.4: *Diagrama de estados para o autômato* $\mathcal{D}_8^2$

Definição 3.4.5. Seja $S \in K\langle\langle A \rangle\rangle$. Dizemos que S é *reconhecida* pelo autômato ponderado $\mathcal{A} = (A, K, Q, I, E, T)$ se para $\omega = a_1 a_2 \dots a_n$, temos

$$(S, \omega) = \sum_{q_0, \dots, q_n \in Q} I(q_0) E(q_0, a_1, q_1) \cdot \dots \cdot E(q_{n-1}, a_n, q_n) T(q_n)$$

É útil chamar um estado q de inicial (final) se $I(q) \neq 0$ ($T(q) \neq 0$). O coeficiente (S, ω) é a soma dos pesos de todos os caminhos c que percorrem ω de u estado inicial p para um estado final q, cada peso sendo mutiplicado à esquerda pelo coeficiente do estado inicial e à direita pelo coeficiente do estado final.

Exemplo 3.4.6. Considere a sequência de Fibonacci e sua série $S = \sum_{n=0}^{\infty} F_n x^n \in \mathbb{N}\langle\langle x \rangle\rangle$. O diagrama de estados do autômato ponderado que reconhece S está apresentado na figura 3.5 abaixo:

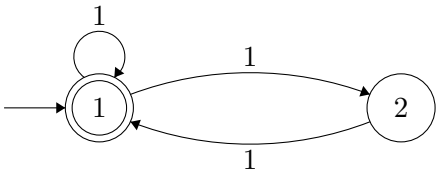


Figura 3.5: *Autômato reconhecendo a série de Fibonacci.*

Exemplo 3.4.7. Considere $A = \{a, b\}$ e a série $S \in \mathbb{Z}\langle\langle A \rangle\rangle$, definida por

$$(S, \omega) = \begin{cases} 2^n, & \text{se } \omega = a^n, n \geq 1 \\ -3 \cdot 2^n, & \text{se } \omega = a^n b, n \geq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

em outras palavras,

$$S = \sum_{n \geq 1} 2^n a^n - 3 \sum_{n \geq 0} 2^n a^n b$$

O suporte de S é a linguagem $a^+ \cup a^* b$. A série é reconhecida pelo seguinte $\mathbb{Z}$-autômato ponderado:

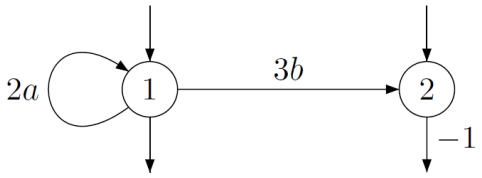


Figura 3.6: *Diagrama de estados do autômato ponderado que reconhece* $S = \sum_{n \geq 1} 2^n a^n - 3 \sum_{n \geq 0} 2^n a^n b$.

Os autômatos ponderados permitem reconhecer se uma linguagem é finita ou não com relativa facilidade, como nos revela o teorema 3.4.8 abaixo, cuja demonstração pode ser encontrada em [32]:

Teorema 3.4.8. *São equivalentes:*

(i) *Uma linguagem L é finita se, e somente se, existe algum autômato ponderado finito que a reconhece cujo diagrama de estados não tem ciclos.*

(ii) *Uma linguagem L é infinita se, e somente se, não existe autômato ponderado finito que reconhece L ou o diagrama de estados de qualquer autômato ponderado finito que a reconhece tem ciclo.*

O diagrama de estados de um autômato ponderado é capaz de trazer à tona diversas características da respectiva série reconhecida por ele. Veremos agora que as séries que são reconhecíveis por autômatos são justamente as séries reconhecíveis, definidas em 3.3.1. Consequentemente, o Teorema de Schützenberger 3.3.12 nos garantirá que toda série racional admite um autômato ponderado que a reconhece.

Proposição 3.4.9. *Seja $S \in K\langle\langle A \rangle\rangle$. Então S é reconhecível por um autômato ponderado finito que não contém ciclos se, e somente se, S é um polinômio.*

Demonstração. Se o autômato ponderado finito que reconhece S não possui ciclos, então pelo teorema 3.4.8, que pode ser facilmente estendido para autômatos ponderados, a linguagem que este autômato reconhece é finita. Como esta linguagem corresponde ao suporte de S, segue pela definição 3.2.4 que S é um polinômio.

Reciprocamente, se S é um polinômio, então este possui suporte finito. Dessa forma, a linguagem que o autômato que reconhece S aceita é finita, e pelo teorema 3.4.8, segue que o autômato ponderado finito que reconhece S não contém ciclos. □

Proposição 3.4.10. *Uma série é reconhecida por um autômato ponderado finito se, e somente se, é reconhecível.*

Demonstração. Assuma que $S \in K\langle\langle A\rangle\rangle$ é reconhecida pelo autômato $\mathcal{A} = (A, K, Q, I, E, T)$. Podemos supor sem perda de generalidade que $Q = \{1, \ldots, n\}$. Então S é reconhecida por uma representação linear (λ, μ, γ), onde $\lambda \in \mathcal{M}_{1\times n}(K)$ é um vetor-linha, $\mu\colon A^* \to \mathcal{M}_n(K)$ é um morfismo e $\gamma \in \mathcal{M}_{n\times 1}(K)$ é um vetor-coluna. Cada uma das entradas de λ e γ estão definidos respectivamente como $\lambda_{1p} = I(p)$ e $\gamma_{q1} = T(q)$, enquanto para μ, temos $\mu(a)_{pq} = E(p, a, q)$ para cada $a \in A$, e $1 \leq p, q \leq n$.

De fato, para $\omega = \alpha_1\alpha_2 \ldots \alpha_n$,

$$(\mu(\omega))_{pq} = \sum_{p_1,\ldots,p_{m-1}} E(p, \alpha_1, p_1)E(p_1, \alpha_2, p_2) \cdot \ldots \cdot E(p_{m-1}, \alpha_m, q)$$

é a soma dos pesos dos caminhos de p a q formando ω. Portanto (S, ω), que é dado pela equação da definição 3.4.5, é igual a $\lambda\mu(\omega)\gamma$.

Reciprocamente, seja (λ, μ, γ) uma representação linear reconhecendo $S \in K\langle\langle A\rangle\rangle$, e defina um autômato ponderado $\mathcal{A} = (A, K, Q, I, E, T)$, tomando $I(p) = \lambda_{1p}$, $E(p, a, q) = (\mu(a))_{pq}$ e $T(q) = \gamma_{q1}$. Então $\mathcal{A}$ reconhece S, pois para $\omega = a_1a_2 \ldots a_m$, temos

$$\begin{aligned}
\sum_{p_1,\ldots,p_{m-1}} I(p_1)E(p_1, a_1, p_2) \cdot \ldots \cdot E(q_{m-1}, a_m, q)T(q) &= \\
&= \sum_{q_0,\ldots,q_n \in Q} \lambda_{1p_1}(\mu(a_1))_{p_1p_2} \cdot \ldots \cdot \ldots \cdot (\mu(a_m))_{p_{m-1}q}\gamma_{q1} \\
&= \sum_{k=1}^{m} \lambda\mu(a_k)\gamma \\
&= \lambda\mu(\omega)\gamma \\
&= (S, \omega)
\end{aligned}$$

□

A demonstração mostra que existe uma clara equivalência entre a noção de autômato ponderado e a representação linear: elas estão associadas entre si.

Corolário 3.4.11. *Toda série racional é reconhecida por um autômato ponderado finito.*

Demonstração. Pelo Teorema de Schützenberger 3.3.12, toda série racional é reconhecível. Pela proposição 3.4.10, toda série reconhecível é reconhecida por um autômato ponderado finito. Segue imediatamente o resultado. □

Exemplo 3.4.12. Considerando o mesmo autômato ponderado do exemplo 3.4.2, este corresponde à representação linear

$$\lambda = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \mu(a) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \gamma = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Exemplo 3.4.13. O autômato do exemplo 3.4.7 corresponde à seguinte representação:

$$\lambda = \begin{pmatrix} 1 & 1 \end{pmatrix} \quad \mu(a) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \mu(b) = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \quad \gamma = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Observe que em particular

$$\mu(a^n) = \begin{pmatrix} 2^n & 0 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad \mu(a^nb) = \begin{pmatrix} 0 & 3 \cdot 2^n \\ 0 & 0 \end{pmatrix}$$

Capítulo 4

Expressões racionais

Neste capítulo, estudaremos as expressões racionais. Na seção 4.1, abordaremos o semianel das expressões racionais $\mathcal{E}$, definindo-o e estudando algumas de suas principais propriedades. São estudadas relações de congruência em $\mathcal{E}$, as consequências da identidade racional $E \equiv 1 + EE^* \equiv 1 + E^*E$, e propriedades da operação da operação $a^{-1}E$, que generaliza a operação definida em 3.3.2 para séries formais.

A seção seguinte trata sobre identidades racionais sobre um anel e suas propriedades, bem como sua "trivialidade".

Durante este capítulo, iremos sempre considerar K como sendo um semianel comutativo e A um alfabeto, a menos que seja explicitado o contrário.

4.1 O semianel das expressões racionais

Vamos definir o semianel das *expressões racionais* em A sobre K. Esse semianel, que denotaremos por $\mathcal{E}$, será construído como a união de uma sequência crescente de semianéis $\mathcal{E}_n$, para $n \geq 0$, ou seja:

$$\mathcal{E} = \bigcup_{n \geq 0} \mathcal{E}_n$$

Cada um desses semianéis será da forma $\mathcal{E}_n = K\langle A_n \rangle$ para algum alfabeto A_n (podendo esse alfabeto ser infinito).

Sem mais delongas, vamos definir o semianel $\mathcal{E}$.

Definição 4.1.1. Considere K um semianel comutativo e A um alfabeto. Seja $\mathcal{A} = \{A_0, A_1, A_2, \ldots\}$ um conjunto de alfabetos, no qual

$$A_n = \begin{cases} A & \text{se } n = 0 \\ A_{n-1} \cup \{E^* | E \in K\langle A_{n-1} \rangle \wedge (E, \varepsilon) = 0\} & \text{se } n \geq 1 \end{cases}$$

Definimos $\mathcal{E}_n = K\langle A_n \rangle \ \forall n \geq 0$ e $A_n \in \mathcal{A}$.

O *semianel* $\mathcal{E}$ *das expressões racionais* em A sobre K é dado por:

$$\mathcal{E} = \bigcup_{n=0}^{\infty} \mathcal{E}_n$$

É importante notar que E^* é uma expressão formal, obtida a partir de E colocando $*$ como expoente. Nota-se que a definição dada acima está incompleta, uma vez que não sabemos a partir dela como calcular (E, ε) para $E \in \mathcal{E}$ em geral. Vamos construir uma maneira de preencher tal lacuna na definição.

Em $\mathcal{E}_0 = K\langle A \rangle$, é claro que o termo constante de $S \in K\langle A \rangle$ pode ser calculado, uma vez que já está definido em 3.2.4, pois estamos tratando simplesmente do anel de polinômios. Mais ainda,

podemos considerar este como o morfismo de termos constantes:

$$\begin{array}{rccc} \varphi_0 & : & \mathcal{E}_0 = K\langle A_0\rangle & \longrightarrow & K \\ & & S & \longmapsto & \varphi_0(S) = (S, \varepsilon) \end{array}$$

Desse modo, podemos definir o termo constante indutivamente da seguinte forma: para ε_{n-1}, já temos definida a função φ_{n-1}, dada por

$$\begin{array}{rccc} \varphi_{n-1} & : & \mathcal{E}_{n-1} = K\langle A_{n-1}\rangle & \longrightarrow & K \\ & & E & \longmapsto & \varphi_{n-1}(E) = (E, \varepsilon) \end{array}$$

Podemos então estender tal função para $K\langle A_n\rangle$ pel propriedade universal 3.2.7, tomando a função φ_n,

$$\varphi_n \colon \mathcal{E}_n = K\langle A_n\rangle \longrightarrow K \ ,$$

como o único morfismo de semianéis tal que $\varphi_n(E^*) = 1$ e $\varphi \restriction_{K\langle A_{n-1}\rangle} = \varphi_{n-1}$

Podemos então definir o morfismo de termos constantes:

Definição 4.1.2. A função

$$\begin{array}{rccc} \varphi & : & \mathcal{E} & \longrightarrow & K \\ & & E & \longmapsto & \varphi(E) = \varphi_k(E) \end{array} \ ,$$

onde $k = \min\{n \in \mathbb{N} : E \in \mathcal{E}_n\}$, é chamada de *morfismo de termos constantes.*

Definição 4.1.3. Um elemento de $\mathcal{E}_n \setminus \mathcal{E}_{n-1}$ é chamado *expressão racional com altura de estrela n.*

Exemplo 4.1.4. Seja $A = \{a, b\}$.

Então $ab \in \mathcal{E}_0$, $(ab)^* \in A_1$ e $\varepsilon + b(ab)^*a \in \mathcal{E}_1$. Como $a \in A_0$, $a^* \in \mathcal{E}_1, (a^*b)^* \in A_2, (a^*b)^*a^* \in \mathcal{E}_2$. O termo constante de $\varepsilon + b(ab)^*a$ é 1, e também é o de $(a^*b)^*a^*$.

A seguir, a proposição abaixo, cuja demonstração pode ser encontrada em [12], asserta sobre a existência de um morfismo, univocamente definido, que generaliza o morfismo de termos constantes para qualquer expressão racional em $\mathcal{E}$.

Proposição 4.1.5. *Seja K um semianel, e φ o morfismo de termos constantes definido em 4.1.2. Então, existe um único morfismo*

$$\mathit{eval}\colon \mathcal{E} \to K\langle\langle A\rangle\rangle$$

estendendo a identidade em $K \cup A$ tal que a operação estrela é preservada, ou seja, $\mathit{eval}(k) = k\ \forall k \in K$, $\mathit{eval}(a) = a\ \forall a \in A$, e $\mathit{eval}(E^*) = (\mathit{eval}(E))^*$. Em outras palavras, eval é o único morfismo que satisfaz as condições dadas e torna comutativo o diagrama seguinte:

$$\begin{array}{ccc} \mathcal{E} & \xhookrightarrow{\mathit{eval}} & K\langle\langle A\rangle\rangle \\ \downarrow{\scriptstyle \varphi} & \nearrow{\scriptstyle \iota} & \\ K & & \end{array}$$

onde ι designa a inclusão natural.

Além das propriedades expostas na proposição 4.1.5, eval também preserva termos constantes, isto é, $(\mathrm{eval}(E), \varepsilon) = (E, \varepsilon)$ para toda expressão racional E.

Definição 4.1.6. Sejam E e F duas expressões racionais. Escrevemos $E \equiv F$ quando $\mathrm{eval}(E) = \mathrm{eval}(F)$. Dizemos que $E \equiv F$ ou (E, F) é uma *identidade K-racional.*

Todas as identidades dadas nos exemplos a seguir são identidades K-racionais. Elas foram colocadas por Conway [3] no caso do semianel booleano:

Exemplo 4.1.7.

- Identidades aperiódicas:

$$(ab)^* \equiv \varepsilon + a(ba)^*b \quad \text{e} \quad (a+b)^* \equiv (a^*b)^*a^*$$

- Identidades cíclicas:

$$\forall n \in \mathbb{N}, \quad (P(n))a^* \equiv (1 + a + \ldots + a^{n-1})(a^n)^*$$

- Identidades de estrela:

$$a^* \equiv \varepsilon + a^*a \quad \text{e} \quad a^* \equiv \varepsilon + aa^*$$

Proposição 4.1.8. *A relação $\equiv$, definida em 4.1.6 por*

$$E \equiv F \Leftrightarrow \mathit{eval}(E) = \mathit{eval}(F),$$

é uma congruência no semianel $\mathcal{E}$,

Demonstração. Basta verificar que a relação é reflexiva, transitiva e simétrica e compatível com as operações do semianel $\mathcal{E}$. Façamos então a verificação rotineira:

- Reflexividade: $E \equiv E$, pois $\text{eval}(E) = \text{eval}(E)$.
- Simetria: $E \equiv F \Rightarrow F \equiv E$, pois $\text{eval}(E) = \text{eval}(F)$.
- Transitividade: $E \equiv F \wedge F \equiv G \Rightarrow E \equiv G$, pois

$$\text{eval}(E) = \text{eval}(F) \quad \text{e} \quad \text{eval}(F) = \text{eval}(G)$$

  Implicam $\text{eval}(E) = \text{eval}(G)$. Além disso, se $E \equiv F$ e $G \equiv H$, então

$$E + G \equiv F + H \text{ e} EG \equiv FH$$

  Pois eval é um morfismo de semianéis.

□

Podemos definir outra congruência em $\mathcal{E}$, a qual denotaremos $\sim$ para fins de diferenciação. Esta é a menor congruência de $\mathcal{E}$ tal que para qualquer $E \in \mathcal{E}$, com $(E, \varepsilon) = 0$, tenhamos

$$E^* \sim 1 + EE^* \sim 1 + EE^* \tag{4.1}$$

A ideia é encontrar a menor congruência que satisfaça a propriedade acima.Vamos mostrar que a congruência $\sim$ que satisfaz tal condição é justamente a congruência $\equiv$ dada em 4.1.6.

Proposição 4.1.9. *Se $E \sim F$, então $E \equiv F$ e $(E, \varepsilon) = (F, \varepsilon)$.*

Demonstração. De fato, como $\equiv$ é uma congruência satisfazendo

$$E \equiv 1 + EE^* \equiv 1 + E^*E$$

para qualquer $E \in \mathcal{E}$ com $(E, \varepsilon) = 0$. Isso ocorre pois, para $S = \text{eval}(E)$, temos $S = 1 + SS^* = 1 + S^*S$ pela proposição 3.2.20. Assim,

$$E \sim F \Rightarrow E \equiv F \Rightarrow \text{eval}(E) = \text{eval}(F) \Rightarrow (E, \varepsilon) = (F, \varepsilon).$$

□

Vamos agora tentar "expandir" o domínio da função eval. Uma boa ideia é verificar se podemos definir algo parecido para matrizes. O morfismo de termos constantes

$$\begin{array}{rccc} \varphi & : & \mathcal{E} & \longrightarrow & K \\ & & E & \longmapsto & (E, \varepsilon) \end{array}$$

pode ser estendido naturalmente para matrizes:

$$\begin{array}{rccc} \varphi & : & \mathcal{M}_n(\mathcal{E}) & \longrightarrow & \mathcal{M}_n(K) \\ & & M & \longmapsto & (M,\varepsilon) \end{array}$$

Onde, para

$$M = \begin{pmatrix} E_{11} & E_{12} & \cdots & E_{1n} \\ E_{21} & E_{22} & \cdots & E_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ E_{n1} & E_{n2} & \cdots & E_{nn} \end{pmatrix},$$

temos que

$$(M,\varepsilon) = \begin{pmatrix} (E_{11},\varepsilon) & (E_{12},\varepsilon) & \cdots & (E_{1n},\varepsilon) \\ (E_{21},\varepsilon) & (E_{22},\varepsilon) & \cdots & (E_{2n},\varepsilon) \\ \vdots & \vdots & \ddots & \vdots \\ (E_{n1},\varepsilon) & (E_{n2},\varepsilon) & \cdots & (E_{nn},\varepsilon) \end{pmatrix}$$

Tal qual a definição 3.3.9, chamamos a matriz $M \in \mathcal{M}_n(\mathcal{E})$ de própria se $(M,\varepsilon) = 0$. A congruência $\sim$ também é estendida naturalmente para matrizes $A, B \in \mathcal{M}_n(\mathcal{E})$, considerando $A \sim B$ se e somente se $a_{ij} \sim b_{ij}$ para $1 \leq i,j \leq n$.

Proposição 4.1.10. *Dada uma matriz quadrada própria M sobre $\mathcal{E}$, existem matrizes M_1 e M_2 de mesmo tamanho de M sobre $\mathcal{E}$ tais que $M_1 \sim 1 + MM_1$ e $M_2 \sim 1 + M_2M$.*

Demonstração. Vamos utilizar indução na dimensão da matriz para demonstrar o resultado. Se a matriz M for de tamanho 1×1, o resultado é claro, pois basta notar que $M_1 = M_2 = M^*$. Daí, temos que

$$M \equiv 1 + MM^* \equiv 1 + M^*M,$$

o que corresponde à identidade racional 4.1.

Considere então M com dimensão maior do que 1. Escreva

$$M = \begin{pmatrix} I & J \\ N & L \end{pmatrix}$$

com blocos não-triviais, e I e L quadradas. Por indução, como I e L são próprias, existem matrizes I_1, L_1 de mesmas dimensões de I e L tais que

$$I_1 \sim 1 + II_1 \text{ e } L_1 \sim 1 + LL_1.$$

Seja $I' = I + JL_1N$ e $L' = L + NI_1J$. Utilizando o argumento indutivo novamente, como L' e I' são matrizes próprias, existem I_1' e L_1' tais que

$$I_1' \sim 1 + I'I_1' \text{ e } L_1' \sim 1 + L'L_1'$$

Considere então

$$M_1 = \begin{pmatrix} I_1' & I_1JL_1' \\ L_1NI_1' & L_1' \end{pmatrix}$$

Basta agora apenas verificar que $M_1 \sim 1 + MM_1$. Vamos então calcular cada um dos coeficientes de $1 + MM_1$:

$$\begin{aligned}
1+MM_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} I & J \\ N & L \end{pmatrix} \begin{pmatrix} I_1' & I_1JL_1' \\ L_1NI_1' & L_1' \end{pmatrix} \\
&= \begin{pmatrix} 1+II_1'+JL_1NI_1' & II_1JL_1'+JL_1' \\ NI_1'+LL_1NI_1' & 1+NI_1JL_1'+LL_1' \end{pmatrix} \\
&= \begin{pmatrix} 1+(I+JL_1N)I_1' & (II_1+1)JL_1' \\ (1+LL_1)NI_1' & 1+(NI_1J+L)L_1' \end{pmatrix} \\
&\sim \begin{pmatrix} 1+I'I_1' & I_1JL_1' \\ L_1NI_1' & 1+L'L_1' \end{pmatrix} \sim \begin{pmatrix} I_1' & I_1JL_1' \\ L_1NI_1' & L_1' \end{pmatrix} \\
&= M_1
\end{aligned}$$

O que termina a demonstração. A existência de M_2 é provada simetricamente. □

Corolário 4.1.11. *Se K é um anel e M é própria, $1-M$ é invertível módulo $\sim$.*

Demonstração. Se K é um anel, então também são $\mathcal{E}$ e $\mathcal{E}/\sim$. Da proposição 4.1.10, temos que:

$$M_1 \sim 1+MM_1 \Rightarrow M_1 - MM_1 \sim 1 \Rightarrow M_2(1-M) \sim 1 \quad \textbf{(ii)}$$

Juntando **(i)** e **(ii)**, obtemos

$$M_2 \sim 1+M_2M_{\Rightarrow}M_2 - M_2M \sim 1 \Rightarrow (1-M)M_1 \sim 1 \quad \textbf{(i)}$$

$$(1-M)M_1 \sim 1 \sim M_2(1-M)$$

Desse modo,

$$1 \sim M_2(1-M) \Rightarrow M_1 \sim M_2(1-M)M_1$$

$$(1-M)M_1 \sim 1 \Rightarrow M_2(1-M)M_1 \sim M_2$$

Daí, obtemos

$$M_1 \sim M_2(1-M)M_1 \sim M_2$$

Assim, $1-M$ é invertível em $\mathcal{E}/\sim$. □

Nossa ideia agora será definir um operador K-linear para cada letra $a \in A$. Procurando generalizar o que foi feito em 3.3.2, queremos um operador

$$\begin{array}{rccc} \psi & : & \mathcal{E} & \longrightarrow & \mathcal{E} \\ & & E & \longmapsto & a^{-1}E \end{array}$$

Vamos fazer isso de forma recursiva nos subsemianéis $\mathcal{E}_n$. Para $n=0$, este equivale ao operador em $\mathcal{E}_0 = K\langle A\rangle$ já definido em 3.3.2.

Suponha que o operador já esteja definido em $\mathcal{E}_{n-1}$. Vamos definir $a^{-1}E$ primeiramente para $E \in A_n$; se $E \in A_{n-1}$, então $a^{-1}E$ já está definido. Caso contrário, $E = F^*$ para algum $F \in \mathcal{E}_{n-1}$ com $(F,\varepsilon) = 0$; então $a^{-1}F$ está definida. Logo, podemos considerar

$$a^{-1}E = (a^{-1}F)F^*,$$

como será demonstrado em 4.1.12.

Agora $a^{-1}E$ está definido para $E \in A_n$, e podemos tomar a função

$$\begin{array}{rccc} \mu & : & A_n & \longrightarrow & \mathcal{M}_2(\mathcal{E}_n) \\ & & E & \longmapsto & \mu(E) = \begin{pmatrix} E & 0 \\ a^{-1}E & (E,\varepsilon) \end{pmatrix} \end{array}$$

A função μ pode ser estendida para um morfismo de monoides $A_n^* \to \mathcal{M}_2(\mathcal{E}_n)$, o último com uma estrutura multiplicativa.

Note que A_n^* é uma base para o K-módulo $\mathcal{E}_n$. A partir dessa observação, podemos estender pela K-linearidade para

$$\mathcal{E}_n = K\langle A_n\rangle \to \mathcal{M}_2(\mathcal{E}_n)$$

Podemos então calcular $a^{-1}E$ para qualquer E em $\mathcal{E}_n$, por $a^{-1}E = \mu(E)_{21}$.

Então, o operador está definido em $\mathcal{E}_n$, recursivamente para todo $\mathcal{E}_n$. Como μ é um morfismo multiplicativo, temos para todo E e F em $\mathcal{E}$:

$$\begin{pmatrix} EF & 0 \\ a^{-1}(EF) & (EF,\varepsilon) \end{pmatrix} = \begin{pmatrix} E & 0 \\ a^{-1}E & (E,\varepsilon) \end{pmatrix}\begin{pmatrix} F & 0 \\ a^{-1}F & (F,\varepsilon) \end{pmatrix}$$

A igualdade acima implica uma generalização dos itens 4 e 5 da proposição 3.3.3:

Proposição 4.1.12. *Sejam E e F expressões racionais e a uma letra. Então*

$$a^{-1}(EF) = (a^{-1}E)F + (E,\varepsilon)a^{-1}F$$

Além disso, $a^{-1}E^* = (a^{-1}E)E^*$ se $(E,\varepsilon) = 0$.

Demonstração. Considere o morfismo multiplicativo:

$$\begin{array}{rcccl} \mu & : & \mathcal{E} & \longrightarrow & \mathcal{M}_2(\mathcal{E}) \\ & & E & \longmapsto & \mu(E) = \begin{pmatrix} E & 0 \\ a^{-1}E & (E,\varepsilon) \end{pmatrix} \end{array}$$

Então

$$\mu(EF) = \mu(E)\mu(F) \Rightarrow \begin{pmatrix} EF & 0 \\ a^{-1}(EF) & (EF,\varepsilon) \end{pmatrix} = \begin{pmatrix} E & 0 \\ a^{-1}E & (E,\varepsilon) \end{pmatrix}\begin{pmatrix} F & 0 \\ a^{-1}F & (F,\varepsilon) \end{pmatrix}$$

Mas também temos:

$$\begin{pmatrix} E & 0 \\ a^{-1}E & (E,\varepsilon) \end{pmatrix}\begin{pmatrix} F & 0 \\ a^{-1}F & (F,\varepsilon) \end{pmatrix} = \begin{pmatrix} EF & 0 \\ (a^{-1}E)F + (E,\varepsilon)a^{-1}F & (E,\varepsilon)(F,\varepsilon) \end{pmatrix}$$

Segue portanto que

$$\begin{pmatrix} EF & 0 \\ a^{-1}(EF) & (EF,\varepsilon) \end{pmatrix} = \begin{pmatrix} EF & 0 \\ (a^{-1}E)F + (E,\varepsilon)a^{-1}F & (E,\varepsilon)(F,\varepsilon) \end{pmatrix} \Rightarrow$$

$$a^{-1}(EF) = (a^{-1}E)F + (E,\varepsilon)a^{-1}F$$

□

Lema 4.1.13. *Se $a \in A$ e E é uma expressão racional, então $eval(a^{-1}E) = a^{-1}eval(E)$.*

Demonstração. Por definição, a fórmula é verdadeira para $E \in \mathcal{E}_0$. Vamos demonstrar o lema por indução. Vamos supor que a afirmação é válida para $E \in \mathcal{E}_{n-1}$, $n \geq 1$. Defina o morfismo de semianéis

$$\begin{array}{rcccl} \mu' : & K\langle\langle A\rangle\rangle & \to & \mathcal{M}_2(K\langle\langle A\rangle\rangle) \\ & S & \longmapsto & \begin{pmatrix} S & 0 \\ a^{-1}S & (S,\varepsilon) \end{pmatrix} \end{array}$$

Observe que μ' é multiplicativa como consequência da proposição 3.3.3, pois

$$\begin{aligned}\mu'(S)\cdot\mu'(T) &= \begin{pmatrix} S & 0 \\ a^{-1}S & (S,\varepsilon)\end{pmatrix}\cdot\begin{pmatrix} T & 0 \\ a^{-1}T & (T,\varepsilon)\end{pmatrix} \\ &= \begin{pmatrix} ST & 0 \\ a^{-1}ST + a^{-1}T(S,\varepsilon) & (ST,\varepsilon)\end{pmatrix} \\ &= \begin{pmatrix} ST & 0 \\ a^{-1}(ST) & (ST,\varepsilon)\end{pmatrix} = \mu'(ST)\end{aligned}$$

Temos para $E \in \mathcal{E}_n$:

$$\mu' \circ \mathrm{eval}(E) = \begin{pmatrix} \mathrm{eval}(E) & 0 \\ a^{-1}\mathrm{eval}(E) & (\mathrm{eval}(E),\varepsilon)\end{pmatrix}$$

$$\mathrm{eval} \circ \mu(E) = \begin{pmatrix} \mathrm{eval}(E) & 0 \\ \mathrm{eval}(a^{-1}E) & (E,\varepsilon)\end{pmatrix}$$

Assim, é suficiente mostrar que, para $E \in \mathcal{E}_n$,

$$\mu' \circ \mathrm{eval}(E) = \mathrm{eval} \circ \mu(E)$$

Uma vez que $\mu' \circ \mathrm{eval}$ e $\mathrm{eval} \circ \mu$ são homomorfismos de semianéis K-lineares e como $\mathcal{E}_n = K\langle A_n\rangle$ é suficiente verificar isso para $E \in A_n$.

Vamos mostrar que as entradas 21 das matrizes $\mu' \circ \mathrm{eval}$ e $\mathrm{eval} \circ \mu$ coincidem. Então, ou $E \in A_{n-1} \subset \mathcal{E}_{n-1}$ e isso funciona por indução, ou $E = F^*$ para algum $F \in \mathcal{E}_{n-1}$ com $(F,\varepsilon) = 0$. Assim, sabemos que

$$a^{-1}E = (a^{-1}F)F^*$$

e

$$\mathrm{eval}(a^{-1}E) = \mathrm{eval}(a^{-1}F)\mathrm{eval}(F^*) = (a^{-1}\mathrm{eval}(F))\mathrm{eval}(F)^* = a^{-1}(\mathrm{eval}(F)^*) =$$
$$a^{-1}(\mathrm{eval}(F^*)) = a^{-1}\mathrm{eval}(E).$$

Então, segue que $\mathrm{eval}(a^{-1}F) = a^{-1}\mathrm{eval}(F)$. □

Lema 4.1.14. *Se E é uma expressão racional, então*

$$E \sim (E,\varepsilon) + \sum_{a\in A} a(a^{-1}E).$$

Demonstração. A fórmula é válida por definição quando $E \in \mathcal{E}_0$, como mostrado pelo lema 3.3.5.

Novamente, vamos utilizar um argumento indutivo para mostrar nosso resultado. Vamos supor então que a fórmula é válida para $E \in \mathcal{E}_{n-1}, n \geq 1$, e provar para $E \in \mathcal{E}_n$. Primeiramente, tome $E \in A_n$. Se $E \in A_{n-1}$, e o resultado está válido por indução. Caso contrário, $E = F^*$ para algum $F \in \mathcal{E}_{n-1}$, com $(F,\varepsilon) = 0$. Então, por indução,

$$F \sim \sum_{a\in A} a(a^{-1}F)$$

Então

$$E = F^* \sim \varepsilon + FF^* \sim \varepsilon + \sum_{a\in A} a(a^{-1}F)F^* = \varepsilon + \sum_{a\in A} a(a^{-1}F^*) = \varepsilon + \sum_{a\in A} a(a^{-1}E)$$

Como $\mathcal{E}_n$ é um K- módulo livre com base A_n^*, é suficiente provar que o produto preserva a

fórmula. Assim, suponha que

$$E \sim (E,\varepsilon) + \sum_{a\in A} a(a^{-1}E) \quad \text{e} \quad F \sim (F,\varepsilon) + \sum_{a\in A} a(a^{-1}F)$$

Assim, temos que

$$\begin{aligned}
(EF,\varepsilon) + \sum_{a\in A} a(a^{-1}(EF)) &= (EF,\varepsilon) + \sum_{a\in A} a((a^{-1}E)F + (E,\varepsilon)(a^{-1}F)) \\
&= (E,\varepsilon)(F,\varepsilon) + \sum_{a\in A} a(a^{-1}E)F + (E,\varepsilon)\sum_{a\in A} a(a^{-1}F) \\
&= (E,\varepsilon)\left((F,\varepsilon) + \sum_{a\in A} a(a^{-1}F)\right) + \sum_{a\in A} a(a^{-1}E)F \\
&\sim (E,\varepsilon)F + \sum_{a\in A} a(a^{-1}E)F \\
&= \left((E,\varepsilon) + \sum_{a\in A} a(a^{-1}E)\right)F \\
&\sim EF.
\end{aligned}$$

□

4.2 Identidades racionais sobre um anel

O objetivo dessa seção é mostrar que, se K é um anel comutativo, então todas as identidades racionais sobre K, como definidas em 4.1.6, são a grosso modo trivias. Isso significa que todas as identidades racionais são consequências do fato de que S^* é o inverso de $\varepsilon - S$, para qualquer série própria S. Como K é um anel, $\mathcal{E}$ também é um anel, e podemos considerar equivalentemente $\mathrm{Ker}(eval)$, chamado *ideal das identidades racionais.*

Vamos primeiramente analisar alguns exemplos:

Exemplo 4.2.1. Seja $A = \{a, b\}$.

Considere a igualdade

$$(ab)^* = \varepsilon + a(ba)^*b.$$

Combinatoriamente, significa que cada palavra em $(ab)^*$ é vazia ou é da forma $a\omega b$, onde $\omega \in (ba)^*$. Vamos mostrar que essa identidade pode ser deduzida algebricamente das identidades $(\varepsilon - S)S^* = \varepsilon = S^*(\varepsilon - S)$. De fato,

$$\begin{aligned}
\varepsilon &= \varepsilon - ab + ab = \varepsilon - ab + a(\varepsilon - ba)(ba)^*b = \varepsilon + a(ba)^*b - ab - aba(ba)^*b \\
&= (\varepsilon - ab)(\varepsilon + a(ba)^*b)
\end{aligned}$$

Aqui usamos $(\varepsilon - ba)(ba)^* = \varepsilon$ na segunda igualdade. Como $(ab)^*$ é o inverso de $\varepsilon - ab$, obtemos:

$$(ab)^* \cdot \varepsilon = \underbrace{(ab)^*(\varepsilon - ab)}_{=\varepsilon}(\varepsilon + a(ba)^*b) \Rightarrow (ab)^* = \varepsilon + a(ba)^*b$$

Exemplo 4.2.2. Seja novamente $A = \{a, b\}$, e considere

$$(a + b)^* = (a^*b)^*a^*$$

Combinatoriamente, isso significa que cada palavra em $\{a, b\}^*$ possui uma única fatoração $a^{i_0}ba^{i_1}b\cdots ba^{i_n}$ com $n \geq 0$ e $i_0, \ldots, i_n \geq 0$. Essa identidade também pode ser deduzida algebricamente das

identidades $(\varepsilon - S)S^* = \varepsilon = S^*(\varepsilon - S)$. Temos:

$$\begin{aligned}
\varepsilon &= (a^*b)^*(\varepsilon - a^*b) \\
&= (a^*b)^* - (a^*b)^*a^*b \\
&= (a^*b)^*\varepsilon - (a^*b)^*a^*b \\
&= (a^*b)^*(a^*(1-a)) - (a^*b)^*a^*b \\
&= (a^*b)^*(a^* - a^*a) - (a^*b)^*a^*b \\
&= (a^*b)^*a^* - (a^*b)^*a^*a - (a^*b)^*a^*b \\
&= (a^*b)^*a^*(\varepsilon - a - b)
\end{aligned}$$

Aqui usamos o fato de que $(a^*b)^*$ é o inverso de $\varepsilon - a^*b$ e a^* é o inverso de $\varepsilon - a^*$. Então

$$\varepsilon \cdot (a+b)^* = (a^*b)^*a^* \underbrace{(\varepsilon - a - b)\cdot(a+b)^*}_{=\varepsilon} \Rightarrow (a+b)^* = (a^*b)^*a^*$$

Já que $(a+b)^*$ é o inverso de $\varepsilon - a - b$.

Os dois exemplos 4.2.1 e 4.2.2 mostram que as respectivas identidades podem ser deduzidas algebricamente a partir de $(\varepsilon - S)S^* = \varepsilon = S^*(\varepsilon - S)$. De fato, veremos que isto sempre é possível, nos revelando que todas as identidades racionais na verdade podem ser "geradas"tendo como ponto de partida uma única identidade racional. Enunciamos formalmente esta constatação no

Teorema 4.2.3. *Se K é um anel, o ideal das identidades racionais é gerado pelas expressões racionais $(\varepsilon - E)E^* - \varepsilon$ e $E^*(\varepsilon - E) - \varepsilon$, com $E \in \mathcal{E}$ e $(E, \varepsilon) = 0$.*

Demonstração. Vamos dividir a prova do teorema em 3 etapas:

1. Iniciamos lembrando que se um anel comutativo K é noetheriano, então cada submódulo de um K-módulo (à esquerda ou à direita) finitamente gerado também é um módulo finitamente gerado. [1] Uma vez que uma identidade racional envolve apenas uma quantidade finita de coeficientes de K, é suficiente provar o teorema quando K é um anel finitamente gerado. Nesse caso, K é um anel Noetheriano, pois todo anel finitamente gerado é Noetheriano (veja [14]). Consequentemente, cada submódulo de um módulo finitamente gerado sobre K é finitamente gerado.

2. Vamos associar agora para cada expressão racional um K-módulo finitamente gerado de $\mathcal{E}$ o qual é estável, ou seja, fechado sobre os operadores $a^{-1}E$, e que contém E, como definido em 3.3.6. Para lograr o objetivo, basta repetir analogamente o que já foi feito com séries racionais na primeira parte da demonstração do Teorema de Schützenberger 3.3.12: Se $E \in \mathcal{E}_0 = K\langle A\rangle$, a existência do módulo é clara: podemos tomar o K-submódulo gerado pelas palavras que aparecem em E. Para o passo indutivo, note que, tomando o resultado garantido para $E \in \mathcal{E}_{n-1}$, o resultado será satisfeito se $E \in A_{n-1}$. Agora, tome $E \in A_n \setminus A_{n-1}$. Assim, $E = F^*$ para alguma $F \in \mathcal{E}_{n-1}$, com $(F, \varepsilon) = 1$. Por indução, existe um K-submódulo M estável, que é finitamente gerado (pois K é noetheriano) de $\mathcal{E}$ o qual contém F. Defina

   $$T = ME + KE.$$

   Então T é um K-submódulo de $\mathcal{E}$ finitamente gerado contendo E. Além disso, T é estável, já que, como $F \in M$ e M é estável, então $a^{-1}F \in M$, e temos que

   $$a^{-1}E = (a^{-1}F)E \in ME$$

   e, para $G \in M$,

   $$a^{-1}(GE) = (a^{-1}G)E + (G, \varepsilon)(a^{-1}E) \in ME$$

[1] Dentre outras definições equivalentes. Para mais detalhes, veja por exemplo [15].

pois $a^{-1}G \in M$. Vamos provar a existência de um submódulo para todos os elementos de $\mathcal{E}_n$ mostrando que se E, F possuem tal submódulo, então também $E+F$ e EF possuem. Denote os correspondentes submódulos porM_E e M_F. É fácil ver que $M_E + M_F$ e $M_E F + M_F$ são os submódulos procurados. Observe que apenas utilizamos o fato de que K é um semianel comutativo.

3. Tome $E \equiv 0$ uma identidade racional. Seja M o menor K-submódulo estável de $\mathcal{E}$ contendo E. As etapas 1 e 2 de nossa demonstração garantem que M é finitamente gerado. Já que M é finitamente gerado, considere $\mathfrak{E} = \{E_1, \ldots, E_n\}$ um conjunto gerador.

   Então, é suficiente provar que $\mathfrak{E} \subset \mathcal{J}$, onde $\mathcal{J}$ é o ideal de $\mathcal{E}$ gerado pelos elementos indicados no teorema. Note que $\sim$ é a igualdade módulo $\mathcal{J}$. Consequentemente, temos que mostrar que $E \sim 0 \ \forall E \in \mathfrak{E}$.

   Pelo lema 4.1.13, cada elemento de M é em si uma identidade racional, já que M é gerado pelo menor subconjunto de $\mathcal{E}$ contendo E e fechado para a operação $F \longmapsto a^{-1}F, a \in A$. Em particular, $E_i, \varepsilon) = 0$. Dessa forma, pelo lema 4.1.14, temos que

$$E \sim \sum_{a \in A} a(a^{-1}E), \forall \ E \in \mathscr{E}.$$

   Como M é estável, $a^{-1}E_i$ é uma combinação K-linear de elementos $E_j \in \mathfrak{E}$. Assim, podemos encontrar polinômios homogêneos $M_{i,j}$ de grau 1 tais que $E_i \sim \sum_j M_{i,j}E_j$. De outro modo, se tomarmos $M = (M_{i,j})$, obtemos

$$(1 - M) \cdot \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_{n-1} \\ E_n \end{pmatrix} \sim 0.$$

   Pela proposição 4.1.10, $1 - M$ é invertível módulo $\mathcal{J}$. Assim, $E_i \in \mathcal{J}$ para qualquer i. Consequentemente, $\mathfrak{E} \subset \mathcal{J}$.

□

Capítulo 5

Altura de estrela

A partir dos pré-requisitos vistos nos capítulos anterior, estamos prontos para apresentar as definições mais importantes da primeira parte deste trabalho.

Na seção 5.1, apresentamos a caracterização da altura de estrela de uma série racional e exploramos alguns exemplos, com o objetivo de analisar o comportamento da altra de estrela de certas séries, servindo como motivação para a questão natural que se apresenta sobre a limitação da altura de estrela, ou seja, a existência de séries com altura de estrela arbitrariarmente grande.

Na seção 5.2, são apresentados conceitos básicos da Teoria de Grafos que serão utilizados para a conceitualização de caminhos e complexidade de ciclo em 5.2.4, que serão de suma importância para prover uma maneira de se calcular a altura de estrela de uma série racional, o que será explorado na seção seguinte.

A seção 5.3 apresenta o principal resultado sobre a altura de estrela, no qual mostrarmos que a altura de estrela não está limitada.

5.1 Definição e primeiros exemplos

Agora, finalmente estamos guarnecidos da maquinaria básica necessária para apresentar a definição mais importante desta primeira parte:

Definição 5.1.1. Defina uma sequência

$$\overline{\mathcal{O}}_0 \subseteq \overline{\mathcal{O}}_1 \subseteq \ldots \subseteq \overline{\mathcal{O}}_n \subseteq \ldots$$

de semianéis de $K\langle\langle A\rangle\rangle$ tal que a união de todos os $\overline{\mathcal{O}}_n$ seja o conjunto de todas as séries racionais. Temos que $\overline{\mathcal{O}}_0 = K\langle A\rangle$, (seguindo a notação da definição 3.2.4) e para $S, T \in \overline{\mathcal{O}}_i$ ambas $S + T$ e ST estão em $\overline{\mathcal{O}}_i$; se S_i for própria, então $S^* \in \overline{\mathcal{O}}_{i+1}$.

A *altura de estrela* de uma série racional S é o menor inteiro n tal que $S \in \overline{\mathcal{O}}_n$.

Apresentaremos a seguir alguns exemplos do cômputo da altura de estrela em diversos ambientes, para nos acostumarmos com este conceito:

Exemplo 5.1.2. Seja $S \in \mathbb{R}_+\langle\langle\{a, b, c\}\rangle\rangle$, dada por

$$(S, \omega) = \sqrt{\left\lfloor \frac{5}{\sqrt{|\omega| + 11} - 3} \right\rfloor}$$

Note que $(S, \omega) = 0 \Leftrightarrow \frac{5}{\sqrt{|\omega|+11}-3} < 1 \Leftrightarrow |\omega| > 53$. Logo, $\mathrm{Supp}(S) = \{\omega \in A^* || \omega \leq 53\}$ é finito. Assim, S é um polinômio, e sua altura de estrela é 0.

Exemplo 5.1.3. Como mostrado no exemplo 3.2.19, vamos considerar a série $S_{\mathcal{F}} = \sum_{n=0}^{\infty} F_n x^n \in \mathbb{N}\langle\langle A\rangle\rangle$, para $A = \{x\}$. Sabemos que a função geradora da sequência de Fibonacci é

$$\frac{1}{1-(x+x^2)} = \sum_{n=0}^{\infty} F_n x^n$$

O autômato ponderado que reconhece a série $S_{\mathcal{F}}$ está representado na figura 3.5.

Das propriedades da proposição 3.2.20, sabemos que $S_{\mathcal{F}}^*$ é o inverso de $1 - S_{\mathcal{F}}$. podemos então escrever que

$$(x+x^2)^* = \frac{1}{1-(x+x^2)}$$

Tomando $T = x + x^2$, é fácil ver que $T \in \overline{\mathcal{O}}_0$, pois T é um polinômio. Assim, $S_{\mathcal{F}} = T^*$, implicando $S_{\mathcal{F}} \in \overline{\mathcal{O}}_1$.

Logo, a série $S_{\mathcal{F}} = \sum_{n=0}^{\infty} F_n x^n$ possui altura de estrela 1.

Exemplo 5.1.4. Ainda sobre a sequência de Fibonacci, considere a série formada pelos seus termos pares, ou seja , $S_{\mathcal{F}_e} = \sum_{n=0}^{\infty} F_{2n} x^n \in \mathbb{N}\langle\langle A\rangle\rangle$. O autômato que a reconhece encontra-se na figura 5.1.

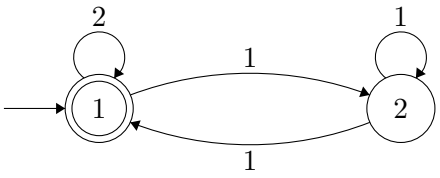


Figura 5.1: *Autômato reconhecendo os termos de índice par na série de Fibonacci.*

Da teoria de funções geradoras, [1], sabemos que $F_{\text{pares}}(x) = \frac{F(x)+F(-x)}{2}$. Aplicando esta equação, vem:

$$F_{\text{pares}}(x) = \frac{1}{2}\left(\frac{1}{1-x-x^2} - \frac{1}{1+x-x^2}\right) = \frac{1}{1-3x-x^2}$$

Agora, vamos ajustar o denominador. Temos:

$$\frac{1}{1-3x-x^2} = \frac{1}{1-(2x+x+x^2)} = \frac{1}{1-\left(2x+x^2\left(\frac{1}{x}+1\right)\right)} = \frac{1}{1-\left(2x+x^2\left(\frac{x+1}{x}\right)\right)} =$$

$$\frac{1}{1-\left(2x+x^2\left(\frac{\frac{1}{x}}{\frac{x}{x+1}}\right)\right)} = \frac{1}{1-\left(2x+x^2\left(\frac{1}{1-\frac{1}{1-x}}\right)\right)} = \frac{1}{1-\left(2x+x^2x^*\right)}$$

Então, temos que $S_{\mathcal{F}_e} = (2x + x^2x^*)^*$. Sendo $T = 2x + x^2x^*$, vemos que $T \in \overline{\mathcal{O}}_1$. Logo $S_{\mathcal{F}_e} = T^* \in \overline{\mathcal{O}}_2$. Portanto, a altura de estrela de $S_{\mathcal{F}_e}$ é 2.[2]

Exemplo 5.1.5. Considere agora a série de Pell em $\mathbb{N}\langle\langle x\rangle\rangle$, que já foi usada no exemplo 3.2.19:

$$S_{\mathcal{P}} = \sum_{n\geq 0} P_n x^n = 1 + 2x + 5x^2 + 12x^3 + 29x^4 + 70x^5 + 169x^6 + \ldots$$

Temos que $\frac{1}{1-2x-x^2}$ é uma função geradora da série de Pell. Com os mesmos argumentos de 5.1.3, temos que $S_{\mathcal{P}} = (2x + x^2)^*$. Assim, a altura de estrela de $S_{\mathcal{P}}$ é 1.

[1] Para mais detalhes, consulte [33]

[2] Rigorosamente, isto mostra na verdade que a altura de estrela de $S_{\mathcal{F}_e}$ é *menor ou igual* a 2. Para mostrar que esta é *exatamente* 2, precisaremos do teorema **??**, que está apresentado no apêndice.

Cabe observar que a série $S_{\mathcal{F}_e}$ com os termos de índice par da sequência de Fibonacci do exemplo 5.1.4 possui altura de estrela 2 levando em conta $\mathbb{N}\langle\langle\{x\}\rangle\rangle$. Considerando $S_{\mathcal{F}_e} \in \mathbb{Z}\langle\langle\{x\}\rangle\rangle$, sua altura de estrela será 1, pois $S_{\mathcal{F}_e} = (3x + x^2)^*$, e $3x + x^2 \in \mathbb{Z}[x]$. Isso mostra que a altura de estrela de uma mesma série pode ser diferente dependendo do semianel K considerado.

Para compreender melhor as propriedades da altura de estrela e como computá-la, precisaremos de alguns conceitos básicos sobre grafos. Outrossim, a série do exemplo 5.1.2 é um polinômio, tendo assim altura de estrela 0 e o autômato que a representa não contém ciclos, em consonância com a proposição 3.4.9.

Veremos mais adiante que tais constatações não são coincidências, e que existe uma correlação direta entre a altura de estrela de uma série racional e a quantidade de ciclos do autômato poderado finito que a reconhece.

5.2 Grafos

A fim de introduzir as definições e resultados sobre altura de estrela da maneira mais clara possível, apresentaremos noções básicas sobre a teoria dos grafos que serão requeridas nas próximas seções. Entre outros trabalhos, esta seção foi inspirada em [24], [6] e [28].

5.2.1 Conceitos básicos

Definição 5.2.1. Um *grafo direcionado* (também chamado de *dirigido* ou *orientado*) consiste de

- um conjunto V de vértices,
- um conjunto E de arestas,
- mapas $s, t\colon E \to V$, onde $s(e)$ é a fonte e $t(e)$ é o alvo da aresta direcionada e.

Adotaremos como notação $G = (V, E)$, pois em geral as funções s, t ficam subentendidas. Em alguns momentos, para deixar claro a que grafo o conjunto de vértices ou arestas pertence, escreveremos $V = V(G)$ e $E = E(G)$.

Um grafo é dito apenas *grafo não-direcionado* se satisfaz apenas as duas primeiras condições.

Os diagramas de estado dos autômatos podem ser interpretados como grafos dirigidos, onde o conjunto de estados Q é o conjunto de vértices e a função de transição E determina as arestas.

Para as próximas definições, denotaremos por $V^{(2)}$ o conjunto de todos os pares não-ordenados de elementos de V, para um conjunto qualquer V. Se V tem n elementos então $V^{(2)}$ tem $\binom{n}{2} = \frac{n(n-1)}{2}$. Os elementos de $V^{(2)}$ serão identificados com os subconjuntos de V que têm cardinalidade 2. Assim, cada elemento de $V^{(2)}$ terá a forma $\{v, w\}$, sendo v e w dois elementos distintos de V.

Definição 5.2.2. O *complemento* de um grafo (V, E) é o grafo $(V, V^{(2)} \setminus E)$. O complemento de um grafo G será denotado por $\overline{G}$.

Definição 5.2.3. O *oposto* de um grafo dirigido (V, E) é o grafo obtido invertendo a direção das arestas de G. O oposto de um grafo G será denotado por $\tilde{G}$.

Definição 5.2.4. Um grafo G é *completo* se $E(G) = V(G)^{(2)}$. A expressão "G é um K_n"é uma abreviatura de "G é um grafo completo com n vértices".

Dois exemplos primordiais de grafos bem conhecidos e importantes são os grafos de Clebsch e de Petersen. Para mais detalhes sobre suas características e importância, consulte [24].

Exemplo 5.2.5 (Grafo de Clebsch)**.** O Grafo de Clebsch, representado na figura, é um grafo não-orientado que possui 16 vértices e 40 arestas, sendo muito importante na Teoria de Ramsey.

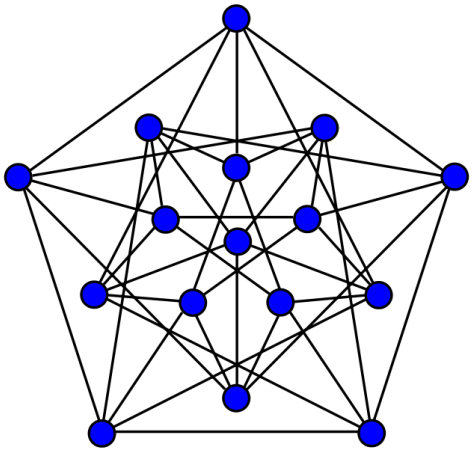


Figura 5.2: *Grafo de Clebsch.*

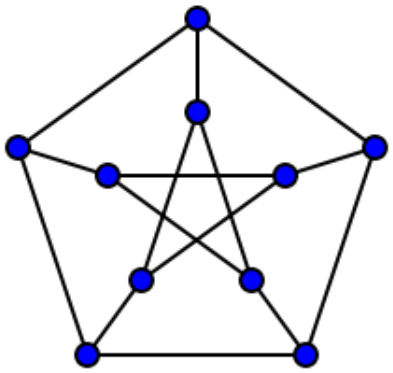


Figura 5.3: *Grafo de Petersen.*

Exemplo 5.2.6 (Grafo de Petersen)**.** Seja V o conjunto de todos os subconjuntos de $\{1, 2, 3, 4, 5\}$ que têm exatamente 2 elementos. Digamos que dois elementos v e w de V são adjacentes se $v \cap w = \emptyset$. Por exemplo, os elementos $v = \{1, 2\}$ e $w = \{3, 4\}$ são adjacentes. Essa relação de adjacência sobre V define o *grafo de Petersen,* como mostra a figura abaixo:

Este é um grafo não-orientado com 10 vértices e 15 arestas. É um pequeno grafo que serve como um exemplo útil e contra-exemplo para muitos problemas.

5.2.2 Isomorfismos de grafos

Definição 5.2.7. Sejam $G = (V_G, E_G)$ e $H = (V_H, E_H)$ grafos dirigidos. Considere f uma função bijetora de V_G em V_H e g uma função bijetora de E_G em E_H. Denote $\theta = (f, g)$. Dizemos então que θ é um *isomorfismo entre os grafos* G e H se a seguinte condição é válida: o vértice x é incidente na aresta A em G se, e somente se, o vértice $f(x)$ é incidente na aresta $g(A)$ em H, ou seja, se para toda aresta e,

$$\begin{cases} s(g(e)) = f(s(e)) \\ t(g(e)) = f(t(e)) \end{cases}$$

Como notação, escrevemos $\theta(G) = H$.

Se tal isomorfismo existe, dizemos que G e H são *isomorfos*. Claramente, temos:

$$|V_G| = |V_H|$$

$$|E_G| = |E_H|$$

Um isomorfismo de G em si mesmo será denotado *automorfismo* do grafo G.

Exemplo 5.2.8. Os dois grafos direcionados abaixo são isomorfos:

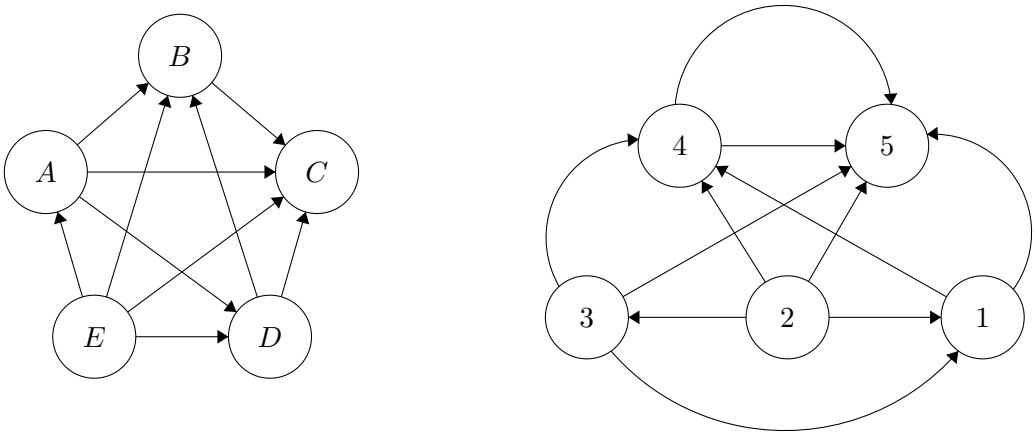


Figura 5.4: *Grafos direcionados isomorfos.*

Note que o isomorfismo θ é composto na verdade por duas funções bijetoras. É natural perguntar então em que condições a existência de uma das funções implica na existência da outra. Em vista disso, tem-se o seguinte teorema, cuja demonstração pode ser encontrada em [31].

Teorema 5.2.9. *Sejam G e H grafos. Seja f uma função bijetora de V_G em V_H que possui a seguinte propriedade: dois vértices distintos x e y de G são adjacentes em G se, e somente se, os correspondentes vértices $f(x)$ e $f(y)$ são adjacentes em H. Então, existe uma única função $g\colon E_G \to E_H$ tal que (f, g) é um isomorfismo entre G e H.*

Definição 5.2.10. Dado um grafo G, seu conjunto de automorfismos $\Phi(G)$ é um grupo sob a operação de composição. Dizemos que $\Phi(G)$ é o *grupo dos automorfismos* de G.

Exemplo 5.2.11. O grupo de automorfismo de um grafo e de seu complemento é sempre o mesmo, ou seja, $\Phi(G) = \Phi(\overline{G})$. Para um grafo completo K_n, temos que $\Phi(K_n) = S_n$.

Exemplo 5.2.12. Seja P o grafo de Petersen, introduzido no exemplo 5.2.6. Então, $\Phi(P) = S_5$.

Exemplo 5.2.13. Seja G o grafo de Nauru, que possui 24 vértices e 36 arestas. Então, $\Phi(G) = S_4 \times S_3$.

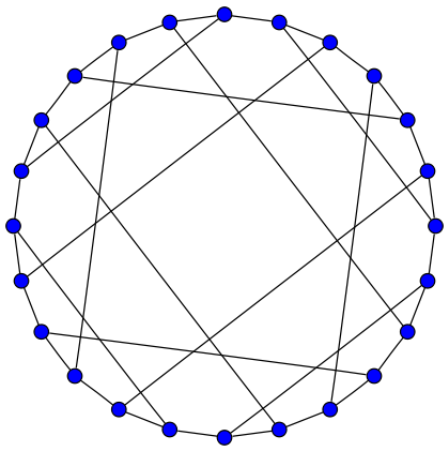


Figura 5.5: *Grafo de Nauru.*

Exemplo 5.2.14. O grafo de Desargues, ilustrado na figura 5.6, possui 20 vértices e 30 arestas, e surge a partir de diferentes construções combinatórias, e tem diversas propriedades de simetrias interessantes, além de possuir aplicações práticas em Química. Seu grupo de automorfismos é $\Phi(G) = S_5 \times \mathbb{Z}/2\mathbb{Z}$.

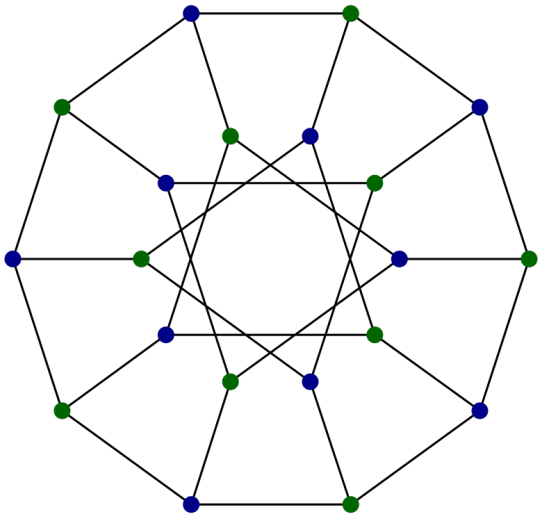


Figura 5.6: *Grafo de Desargues.*

Teorema 5.2.15. *Grafos isomorfos possuem o mesmo grupo de automorfismo.*

Demonstração. Seja θ um isomorfismo entre os grafos G e H. Assim, se $\xi \in \Phi(H)$, o isomorfismo $\theta^{-1}\xi\theta$ está em $\Phi(G)$. Analogamente, se $\eta \in \Phi(G)$, então $\theta\eta\theta^{-1}$ está em $\Phi(H)$. Assim, $|\Phi(G)| = |\Phi(H)|$. Considere

$$\begin{aligned} \varphi\colon \Phi(G) &\longrightarrow \Phi(H) \\ \eta \quad &\longmapsto \theta\eta\theta^{-1} \end{aligned}$$

Claramente φ é bijetora. Vamos mostrar que φ preserva produtos. De fato, $\forall \alpha, \beta \in \Phi(G)$,

$$\varphi(\alpha)\varphi(\beta) = (\theta\alpha\theta^{-1})(\theta\beta\theta^{-1}) = \theta\alpha\beta\theta^{-1} = \varphi(\alpha\beta)$$

Logo $\Phi(G) = \Phi(H)$.

□

5.2.3 Subgrafos

É possível notar que alguns grafos contém outros grafos em seu interior. Diremos que H é um subgrafo de G se está contido em G. Mais precisamente:

Definição 5.2.16. Um grafo H é dito *subgrafo* de um grafo G se

$$V(H) \subseteq V(G) \quad \wedge \quad E(H) \subseteq E(G)$$

e cada aresta de H as mesmas direções de cada aresta de G, caso este seja dirigido. Ou seja, as funções s, t em H são dadas pela restrição de s e t a $E(H)$:

$$\begin{cases} s' = s \restriction_{E(H)}\colon E(H) \to V(H) \\ t' = t \restriction_{E(H)}\colon E(H) \to V(H) \end{cases}$$

Por questão de notação, escrevemos $H \subseteq G$.

Note que, por esta definição, G é um subgrafo de si mesmo.

Definição 5.2.17. Seja G um grafo dirigido. A remoção de um subconjunto S de vértices de G é o subgrafo contendo os vértices de G que não estão em S e as arestas de G que não incidem em vértices de S. Este subgrafo é denotado por $G - S$ e $G - S = G[V_G - S]$, e é chamado *subgrafo próprio* de G.

Se $S = \{w\}$, indicamos $G - w = (V \setminus \{w\}, E')$, onde

$$E' = \{(u, v) \in E | u \neq w \text{ e } v \neq w\}$$

Definição 5.2.18. Seja $G = (V, E)$ um grafo. Um subgrafo $H \subset G$ é maximal se não está propriamente contido em nenhum subgrafo de G, diferente de G. Ou seja, se existe um subgrafo K de G tal que $H \subseteq K \subset G$, então $H = K$.

5.2.4 Caminhos e complexidade de ciclo

Definição 5.2.19. Um *caminho* em um grafo é uma sequência

$$(v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n)$$

onde $s(e_i) = v_i$, $t_{e_i} = v_{i+1}$, para todo $1 \leq i \leq n$.

Os vértices v_0 e v_n são os *extremos* do caminho.

Definição 5.2.20. Um *ciclo* é um caminho com $v_1 = v_n$, com $n \geq 2$.

Um grafo dirigido é *acíclico* se não possui nenhum ciclo; isto é, para qualquer vértice v, não há nenhuma ligação dirigida começando e acabando em v. O tamanho do ciclo é a quantidade de vértices no caminho que o forma.

Definição 5.2.21. Dizemos que um grafo é *conexo* ou *conectado* se e somente se existir um caminho entre quaisquer dois vértices.

Exemplo 5.2.22. O grafo de Markström, representado na figura 5.7, é um grafo conexo, pois é possível obter um caminho entre quaisquer dois vértices. Além disso, este grafo possui ciclos com todos os comprimentos possíveis, a exceção de ciclos de tamanhos $1, 2, 4$ e 8.

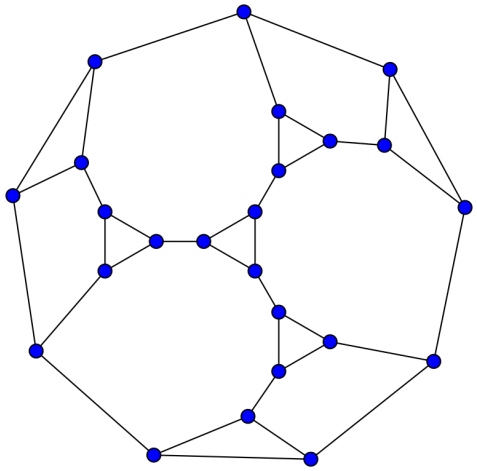


Figura 5.7: *Grafo de Markström.*

Definição 5.2.23. Um grafo dirigido é *fortemente conexo* ou *fortemente conectado* se existe um caminho entre quaisquer dois vértices.

É importante notar as sutilezas existentes entre as definições 5.2.21 e 5.2.23. Referimo-nos a conexo para um grafo não-dirigido, enquanto utilizamos o termo fortemente conexo para grafos dirigidos. Note que se um grafo dirigido é fortemente conexo, então o grafo não-dirigido equivalente (a grosso modo, o grafo obtido ao "ignorar"as direções das setas) sempre será conexo; em contrapartida, a recíproca pode não ocorrer, ou seja, um grafo não-dirigido pode ser conexo, mas existem grafos dirigidos com o mesmo conjunto de vértices que não são fortemente conexos.

Definição 5.2.24. Um grafo dirigido $G = (V, E)$ é chamado *bola* se é fortemente conexo e possui pelo menos uma aresta.

Definição 5.2.25. Uma *componente fortemente conexa* de um grafo dirigido G é um subgrafo maximal que é fortemente conexo. O conjunto das componentes fortementes conexas de um grafo G é denotado por $\mathfrak{S}^{\emptyset}(G)$.

Lema 5.2.26. *Seja $G = (V, E)$ um grafo finito dirigido, $v \in V$ e $H \in \mathfrak{S}^{\emptyset}(G)$. Se $v \in H$, então $\mathfrak{S}^{\emptyset}(H - v) \subset \mathfrak{S}^{\emptyset}(G - v)$. Se $v \notin H$, então $H \in \mathfrak{S}^{\emptyset}(G - v)$.*

Demonstração. Suponha que $v \in H$. Então, tome $K \in \mathfrak{S}^{\emptyset}(H-v)$, e suponha que exista um subgrafo $L \in G - v$ tal que $K \subset L \subseteq G - v$. Como $H \in \mathfrak{S}^{\emptyset}(G)$, então $\nexists P \subset G : H \subset P \subseteq G$. Logo, $L = K$, e $K \in \mathfrak{S}^{\emptyset}(G - v), \forall K \in \mathfrak{S}^{\emptyset}(H - v)$. Consequentemente, $\mathfrak{S}^{\emptyset}(H - v) \subset \mathfrak{S}^{\emptyset}(G - v)$.

Se $v \notin H$, por um raciocínio análogo, obtemos $H \in \mathfrak{S}^{\emptyset}(G - v)$. □

Definição 5.2.27. A *complexidade de ciclo* de um grafo dirigido G, denotada por $\mathcal{C}(G)$, é definida da seguinte maneira:

- Se G não possui nenhum ciclo[3], então sua complexidade de ciclo é 0;
- Se G é fortemente conectado, então será 1 adicionado ao menor valor entre as complexidades de ciclo dos grafos $G - v$, para todo vértice v, ou seja:

$$\mathcal{C}(G) = 1 + \min\{\mathcal{C}(G - v) | v \in G\}$$

- Se G não é fortemente conectado, é o máximo entre as complexidades de ciclo das componentes fortemente conectadas de G, ou seja:

$$\mathcal{C}(G) = \max\{\mathcal{C}(H) | H \in \mathfrak{S}^{\emptyset}(G)\}$$

Note que a definição 5.2.27 cobre todas as possibilidades para o grafo G, e portanto a complexidade de ciclo está bem-definida.

Exemplo 5.2.28. Considere o grafo $G = (V, E)$ abaixo, com:

$$V = \{1, 2, 3\} \quad \text{e} \quad E = \{(1, 1), (1, 2), (1, 3), (2, 1), (3, 1), (3, 2)\}$$

Vamos verificar que sua complexidade de ciclo é 1.

Claramente, G possui bolas. Logo, $\mathcal{C}(G) > 0$. O grafo G é fortemente conexo. Como $\mathcal{C}(G-1) = \mathcal{C}((\{2, 3\}, \{3, 2\})) = 0$, pois $G - 1$ não contém bolas, obtemos $\min\{\mathcal{C}(G - v) | v \in G\} = 0$, e segue que

$$\mathcal{C}(G) = 1 + min\{\mathcal{C}(G - v) | v \in G\} \Rightarrow \mathcal{C}(G) = 1.$$

Exemplo 5.2.29. Considere o grafo $G = (V, E)$, com $V = \{1, 2\}$ e $E = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$, correspondente à representação da figura:

O grafo G abaixo possui complexidade de ciclo 2, pois G é fortemente conexo, e $\min\{\mathcal{C}(G-v) | v \in G\} = 1$, e assim

$$\mathcal{C}(G) = 1 + \min\{\mathcal{C}(G - v) | v \in G\} \Rightarrow \mathcal{C}(G) = 2.$$

[3]Alguns autores, como [28], definem a complexidade de ciclo nula para grafos sem caminhos infinitos ou bolas.

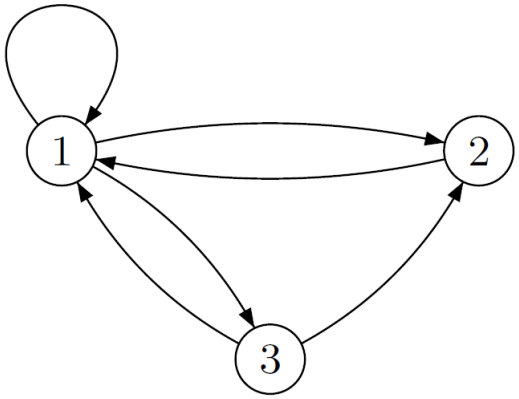


Figura 5.8: *Grafo com complexidade de ciclo* 1.

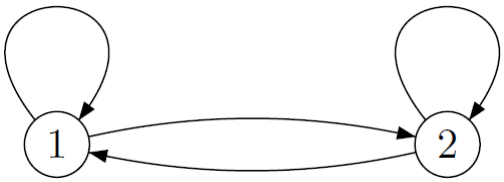


Figura 5.9: *Grafo com complexidade de ciclo* 2.

Note que a quantidade de vértices nem sempre indica que a complexidade de ciclo é grande. Como visto, o grafo da figura 5.8 tem 3 vértices, mas sua complexidade de ciclo é 1, enquanto o grafo da figura 5.9 tem 2 vértices e complexidade de ciclo 2.

Vamos verificar algumas propriedades da complexidade de ciclo.

Lema 5.2.30. *Seja K_n um grafo completo, como definido em 5.2.4. Então, temos que $\mathcal{C}(K_n) = n$.*

Demonstração. Vamos demonstrar o resultado por indução. Claramente, $\mathcal{C}(K_1) = 1$ e $\mathcal{C}(K_2) = 2$. Suponhamos agora que $\mathcal{C}(K_n) = n$ para algum $n \in \{2, 3, \ldots\}$. Vamos mostrar que ela também vale para $n + 1$.

Observe que, como K_{n+1} é um grafo completo, existe um caminho ligando quasiquer dois vértices, sendo portanto fortemente conexo. Logo, pela definição 5.2.27,

$$\mathcal{C}(K_{n+1}) = 1 + \min\{\mathcal{C}(K_{n+1} - v) | v \in K_{n+1}\}$$

Mas $K_{n+1} - v = K_n,\ \forall v \in K_{n+1}$. Pela hipótese de indução, sabemos que $\mathcal{C}(K_n) = n$. Daí,

$$\mathcal{C}(K_{n+1}) = 1 + \mathcal{C}(K_n) \Rightarrow \mathcal{C}(K_{n+1}) = n + 1$$

□

Exemplo 5.2.31. O grafo da figura 5.9 é um exemplo de grafo completo com dois vétices, ou seja, é um K_2. Como calculado no exemplo 5.2.29, temos que $\mathcal{C}(K_2) = 2$.

Lema 5.2.32. *Seja G um grafo dirigido finito. Então G e $\tilde{G}$ têm a mesma complexidade de ciclo.*

Demonstração. Claramente, G e $\tilde{G}$ simultaneamente possuem bolas ou não. Além disso, os componentes fortemente conectados de G e $\tilde{G}$ formam grafos opostos. Outrossim, se v é um vértice, então $\widetilde{G - v} = \tilde{G} - v$. De posse desses fatos, estamos preparados para demonstrar o lema.

Vamos utilizar indução na quantidade de vértices de $G = (V, E)$. Seja $|V| = 1$. Então claramente $\mathcal{C}(G) = \tilde{\mathcal{G}}$.

Suponha agora que para $|V| = k-1$, e vale $\mathcal{C}(G) = \tilde{\mathcal{G}}$. Podemos supor sem perda de generalidade que G possui bolas e é fortemente conectado. Então, temos que

$$\mathcal{C}(G) = 1 + min\{\mathcal{C}(G - v) | v \in G\}$$

Da mesma forma,

$$\mathcal{C}(\tilde{G}) = 1 + min\{\mathcal{C}(\widetilde{G} - v)|v \in \tilde{G}\}$$

Mas, pela hipótese de indução,

$$\mathcal{C}(\tilde{G}) - 1 = min\{\mathcal{C}(\tilde{G} - v)|v \in \tilde{G}\} = min\{\mathcal{C}(\widetilde{G - v})|v \in \widetilde{G}\} = min\{\mathcal{C}(G - v)|v \in G\}$$

Daí,

$$\mathcal{C}(G) = 1 + \mathcal{C}(\tilde{G}) - 1 \Rightarrow \mathcal{C}(G) = \mathcal{C}(\tilde{G})$$

como desejado. □

Para compreender melhor a complexidade de ciclo, vamos observar como esta se comporta, procurando funções que possam auxiliar em sua quantificação, como a *função próximo* e *função altura*, que serão o alvo de nossas próximas definições.

Definição 5.2.33. Seja V um conjunto finito totalmente ordenado e $h\colon V \to \mathbb{N}$ uma função. A partir dela, defina outra função

$$n\colon V \to V \cup \{\infty\},$$

onde $\infty \notin V$ e $v < \infty$ para qualquer $v \in V$. Chamaremos n de *função próximo:* $n(v)$ é o menor $v' > v$ tal que $h(v') \geq h(v)$ e $n(v) = \infty$ caso contrário. Sucintamente, se

$$A = \{v' \in V | v' > v \ \wedge \ h(v') \geq h(v)\}$$

então:

$$n(v) = \begin{cases} \min(A), \text{ se } A \neq \emptyset, \\ \infty, \text{ caso contrário.} \end{cases}$$

Exemplo 5.2.34. Seja $V = \{a, b, c\}$, em que $a < b < c$, e considere a função

$$\begin{aligned} h\colon V &\longrightarrow \mathbb{N} \\ v \quad &\longmapsto h(v) = \begin{cases} 2, \text{ se } v = a, \\ 3, \text{ se } v = b, \\ 5, \text{ se } v = c. \end{cases} \end{aligned}$$

Vamos definir a função n. Temos que:

- $n(a) = b$, pois
$$A = \{v' \in V | v' > v \text{ e } h(v') \geq h(v)\} = \{b, c\}$$
e $\min(A) = b$.
- $n(b) = c$, pois
$$A = \{v' \in V | v' > v \text{ e } h(v') \geq h(v)\} = \{c\}$$
e $\min(A) = c$.
- $n(c) = \infty$, pois
$$A = \{v' \in V | v' > v \text{ e } h(v') \geq h(v)\} = \emptyset$$

Assim, a função próximo para este caso é:

$$\begin{aligned} v\colon V &\longrightarrow V \cup \{\infty\} \\ v \quad &\longmapsto n(v) = \begin{cases} b, \text{ se } v = a, \\ c, \text{ se } v = b, \\ \infty, \text{ se } v = c. \end{cases} \end{aligned}$$

Vamos observar agora como a função próximo se comporta "localmente", ou seja, para subconjuntos de V.

Lema 5.2.35. *Seja W um intervalo em V e considere $g = h\restriction_W$, e seja n' a função próximo de g, como definido em 5.2.33. Então, para qualquer $w \in W$,*

$$n'(v) = \begin{cases} n(w), & \text{se } n(w) \in W, \\ \infty, & \text{caso contrário.} \end{cases}$$

Em particular, $n'(w) \geq n(w)$.

Demonstração. Uma vez que

$$n'(w) = \min\{u \in W | u > w \ \wedge \ h(u) \geq h(w)\}$$

e

$$n(w) = \min\{u \in V | u > w \ \wedge \ h(u) \geq h(w)\}$$

é possível notar que se $n(w) \in w$, então $n(w) = n'(w)$. Se $n(w) \notin W$, então, como $w < n(w)$ e W é um intervalo, $n(w)$ é maior que qualquer elemento de V; assim, para $v > w$, temos $h(v) < h(w)$ e $n'(w) = \infty$. □

Definição 5.2.36. Seja $G = (V, E)$ um grafo direcionado finito. Dizemos que $h\colon V \to \mathbb{N}$ é uma *função altura* para G se existe uma ordem total em V tal que, sendo n a função próximo de h, a seguinte condição é satisfeita:

Para qualquer $v \in V$ se $h(v) = 0$ (respectivamente $h(v) \geq 1$), então para cada aresta $v \to v'$, temos que $v' < v$. (respectivamente $v' < n(v)$)

Note que no caso $h(v) \geq 1$, se $n(v) = \infty$, então a conclusão é sempre válida.

Exemplo 5.2.37. Considere o grafo do exemplo 5.2.28, ilustrado na figura 5.8. Dada a ordem natural dos vértices, podemos levar em conta a função h tal que $h(1) = 1$, $h(2) = h(3) = 0$, ou seja:

$$\begin{aligned} h\colon V &\longrightarrow \mathbb{N} \\ v &\longmapsto h(v) = \delta_{1v} \end{aligned}$$

onde δ_{ij} representa o delta de Kronecker.

A respectiva função próximo associada é $n(1) = n(3) = \infty$ e $n(2) = 3$.

Exemplo 5.2.38. Considere o grafo do exemplo 5.2.29, ilustrado na figura 5.9. Dada a ordem natural dos vértices, podemos levar em conta a função h tal que $h(1) = 2$ $h(2) = 1$, ou seja:

$$\begin{aligned} h\colon V &\longrightarrow \mathbb{N} \\ v &\longmapsto h(v) = 3 - v \end{aligned}$$

A respectiva função próximo associada é $n(1) = n(2) = \infty$.

Lema 5.2.39. *Se $G = (V_G, E_G)$ e $H = (V_H, E_H)$ são dois grafos dirigidos isomorfos, e existe uma função altura para G, então também existe uma função altura para H.*

Demonstração. De acordo com a definição 5.2.7, seja $\varphi = (\varphi_V, \varphi_E)\colon G \to H$ um isomorfismo entre G e H, e $h\colon V_G \to \mathbb{N}$ uma função altura para G. Então, é fácil ver que $h \circ \varphi_V^{-1}\colon V_H \to \mathbb{N}$ é uma função altura para H. □

Lema 5.2.40. *Seja $G = (V, E)$ um grafo dirigido finito com função altura h, seja V' um intervalo em V e tome G' o grafo obtido pela restrição de G a V'. Então, $h\restriction_{V'}$ é uma função altura para G'.*

Demonstração. Seja n' a função próximo de h'. Vamos verificar o que ocorre quando $h = 0$, e quando $h \geq 1$.

Seja $v \in V'$ com $h'(v) = 0$. Pela definição 5.2.36, $h(v) = 0$, e assim para cada aresta $v \to v'$ com $v' \in V'$ temos $v' < v$.

Agora, considere $v \in V'$ com $h'(v) \geq 1$. Pela definição 5.2.36, $h(v') \geq 1$, e para cada aresta $v \to v'$, com $v' \in V'$, temos $v' < n(v)$. Temos pelo lema 5.2.35 que $n'(v) \geq n(v)$. Então $v' < n'(v)$. Isso prova que h' é uma função altura para G'.

□

O teorema 5.2.41 abaixo é crucial para relacionar a função altura com a complexidade de ciclo de um grafo.

Teorema 5.2.41. *Um grafo $G = (V, E)$ possui complexidade de ciclo com valor no máximo m se, e somente se, G possui uma função altura h com a propriedade de que* $\max(h) \leq m$, *ou seja*

$$\mathcal{C}(G) \leq m \Leftrightarrow \exists h \colon V \to \mathbb{N} : \max(h) \leq m$$

Demonstração. Considere $\mathcal{C}(G) \leq m$. Podemos assumir que $\mathcal{C}(G) = m$. Vamos provar que existe uma função altura satisfazendo as condições do teorema lançando mão da indução. Se $m = 0$, então G não possui bolas, e de acordo com 2.5.7, podemos considerar uma ordem total em V de modo que $v \to v'$ implica $v > v'$. Consequentemente, podemos tomar $h(v) = 0 \forall\, v \in V$.

Suponha agora que $m \geq 1$. Se G é fortemente conexo, existe um vértice ν tal que $\mathcal{C}(G - \nu) = m - 1$. Por indução, existe uma função altura $\hbar \colon V - \nu \to \mathbb{N}$ com $\max(h) \leq m - 1$. Assim, podemos estender h para o domínio V, definindo:

$$\begin{aligned} h \colon V &\longrightarrow \mathbb{N} \\ v \quad &\longmapsto h(v) = \begin{cases} \hbar(v), \text{ se } v \neq \nu, \\ m, \text{ se } v = \nu. \end{cases} \end{aligned}$$

e estendendo a em $V - \nu$ por $\nu \neq v'$ para todo $v' \in V - \nu$. Pode ser prontamente verificado que a função próximo de h satisfaz $n(\nu) = \infty$, e segue do lema 5.2.35 que $n(v') = n_{G-\nu}(v')$ para $v' \neq \nu$.

Daí, segue que h é uma função altura para G.

Suponha agora que G não é fortemente conexo. Vamos instituir uma ordem total no conjunto $\mathfrak{S}^{\emptyset}(G)$, respeitando 2.5.7 de modo que $\forall\, H, H' \in \mathfrak{S}^{\emptyset}(G)$. Se $H < H'$ então não existem arestas de H para H'. Em cada componente fortemente conexa H, existe, por indução uma ordem total de seu conjunto de vértices e uma função altura h_H com $\max(h_H) \leq m$. Definimos h em V estendendo essas funções naturalmente a V, e definimos a ordem total em V colando juntos todas as ordens de maneira compatível com a ordem total de $\mathfrak{G}^{\emptyset}(G)$ e tal que cada $I \in \mathfrak{S}^{\emptyset}(G)$ é um intervalo de V.

Note que se v, v' estão em elementos diferentes de $\mathfrak{S}^{\emptyset}(G)$, então $v \to v'$ implica $v' < v$.

Seja $v \in H$. Vamos supor primeiro que $h(v) = 0$. Então $v \to v'$ implica em duas possibilidades: ou $v' \in H$, e por 5.2.36 $v' < v$ ou $v' \notin H$ e $v' < v$, pois v, v' estarão em elementos diferentes de $\mathfrak{S}^{\emptyset}(G)$. Suponha agora que $h(v) \geq 1$. Então $n_H(v) \in H$ ou $n_H(v) = \infty$.

Se $n_H(v) \in H$, então $n_H(v) = n(v)$ pelo lema 5.2.35; suponha que $v \to v'$: então temos duas opções: ou $v' \in H$, e pela definição 5.2.36 $v' < n_H(v) = n(v)$, ou então $v' \notin H$ e portanto $v' < v < n(v)$.

Se $n_H(v) = \infty$, então $n(v) \notin H$, e $v < n(v)$. Suponha que $v \to v'$: então temos duas alternativas: $v' \in H$ e $v' < n(v)$ (de fato, $v, v' \in H, v < n(v), n(v) \notin H$ e H é um intervalo implicando $v' < n(v)$), ou $v' \notin H$ e nesse caso v, v' estarão em elementos diferentes de $\mathfrak{G}^{\emptyset}(G)$, o que acarreta $v' < v < n(v)$. Dessa forma, concluímos que h é uma função altura para G.

Vamos agora mostrar a recíproca, ou seja, provar que se G possui uma função altura h com $\max(h) \leq m$, então $\mathcal{C}(G) \leq m$.

Seja G possui uma função altura h com $\max(h) \leq m$. Podemos assumir que $\max(h) = m$. Se $m = 0$, o lema 5.2.35 implica que não existem bolas em G, e consequentemente de acordo com a

definição 5.2.27, $\mathcal{C}(G) = 0$. Assuma que $m \geq 1$. Suponha primeiramente que $\nu = \min(V)$ é o único vértice tal que $h(\nu) = m$.

Considere a restrição $h' \colon h \restriction_{V-\nu}$. Uma vez que $\nu = \min(V)$, pelo lema 5.2.40, h' é uma função altura para $G - \nu$ e seu $\max(h') \leq m-1$. Por indução, $\mathcal{C}(G-\nu) \leq m-1$. Seja $H \in \mathfrak{S}^{\emptyset}(G)$ contendo ν. Então pelo lema 5.2.26, $H - \nu = \bigcup_{\mathcal{H} \in \mathscr{H} \subseteq \mathfrak{S}^{\emptyset}(G-\nu)} \mathcal{H}$, e consequentemente $\mathcal{C}(H-\nu) \leq m-1$. Portanto, $\mathcal{C}(H) \leq m$. Se $h' \in \mathfrak{S}^{\emptyset}(G)$ é outra componente fortemente conexa de G, pelo lema 5.2.26, temos também que $H' \in \mathfrak{S}^{\emptyset}(G-\nu)$ e dessa forma $\mathcal{C}(H') \leq m-1$. Concluímos assim que $\mathcal{C}(G) \leq m$.

Suponha agora que $h(\min(V)) \neq m$ ou que $\min(V)$ não é o único vértice para o qual h toma o valor m, e seja v o maior vértice tal que $h(v) = m$ na ordem total de V. Então $V_1 = \{v' \in V \mid v' < v\}$ é diferente de $\emptyset$ e de V. Seja $V_2 = V \setminus V_1$. Então, não existem arestas de V_1 para V_2, porque $v = \min(V_2)$ e $n(v_1) \leq v \ \forall v_1 \in V_1$. Seja $G_i = G \restriction_{V_i}$. Então pelo lema 5.2.40, os grafos G_i herdam uma função altura pela restrição de h, e concluímos por indução que $\mathcal{C}(G_i) \leq m$. Agora, cada commponente fortemente conexa de G está contida numa componente fortemente conexa de G_1 ou G_2, culminando na constatação de que G possui complexidade de ciclo no máximo m, ou seja, $\mathcal{C}(G) \leq m$, como queríamos. □

Definição 5.2.42. Seja K um corpo, e considere E um espaço vetorial de dimensão finita sobre K, com base $\mathcal{B}$ e tome $\Phi \subseteq \mathrm{End}(E)$. Vamos associar à tripla $\mathfrak{V} = (E, \mathcal{B}, \Phi)$ um grafo direcionado com conjunto de vértices $\mathcal{B}$ e arestas $b \to b'$ sempre que existir um certo $\varphi \in \Phi$ tal que $\varphi(b)$ envolve b' quando expandido na base $\mathcal{B}$.

A noção de complexidade de ciclo e função altura para $\mathfrak{V} = (E, \mathcal{B}, \Phi)$ são definidas de maneira correspondente às definições 5.2.27 e 5.2.36, e indicamos sua complexidade de ciclo como $\mathcal{C}(E, \mathcal{B}, \Phi)$.

Exemplo 5.2.43. Seja $E = \mathbb{R}^3$ e $\mathcal{B} = \{e_1, e_2, e_3\}$ a base canônica. Considere o conjunto $\Phi = \{\varphi_1, \varphi_2, \varphi_3\}$, onde

$$\varphi_1(x, y, z) = (x - y, y - z, z)$$
$$\varphi_2(x, y, z) = (x + z, 2y, y - z)$$
$$\varphi_3(x, y, z) = (x, x + 8y + 64z, y - x - z)$$

Calculando os valores dos elementos de Φ nas componentes da base canônica, podemos preencher a tabela abaixo:

Endomorfismo	e_1	e_2	e_3
φ_1	$(1,0,0) = e_1$	$(-1,1,0) = e_2 - e_1$	$(0,-1,1) = e_3 - e_2$
φ_2	$(1,0,0) = e_1$	$(0,2,1) = 2e_2 + e_3$	$(1,0,-1) = e_1 - e_3$
φ_3	$(1,1,-1) = e_1 + e_2 - e_3$	$(0,8,1) = 8e_2 - e_3$	$(0,64,-1) = 64e_2 - e_3$

Tabela 5.1: *Endomorfismos de Φ calculados na base canônica.*

Assim, para os conjuntos $\mathfrak{V}_1 = (\mathbb{R}^3, \mathcal{B}, \{\varphi_1\})$, $\mathfrak{V}_2 = (\mathbb{R}^3, \mathcal{B}, \{\varphi_1, \varphi_3\})$, $\mathfrak{V}_3 = (\mathbb{R}^3, \mathcal{B}, \{\varphi_2, \varphi_3\})$ e $\mathfrak{V}_4 = (\mathbb{R}^3, \mathcal{B}, \Phi)$, a partir dos valores obtidos na tabela 5.1, obtemos os seguintes grafos:

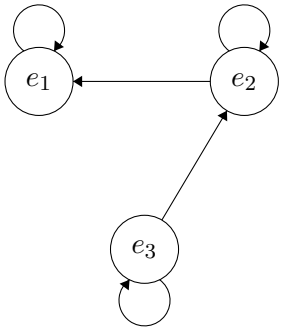
Definição 5.2.44. A complexidade de ciclo de (E, Φ) é o menor valor dentre as complexidades de ciclo das triplas $\mathfrak{V}_B = (E, B, \Phi)$ entre todas as bases B de E, ou seja:

$$\mathcal{C}(E, \Phi) = \min\{\mathcal{C}(E, \mathcal{B}, \Phi) | \mathcal{B} \text{ é uma base de } E.\}$$

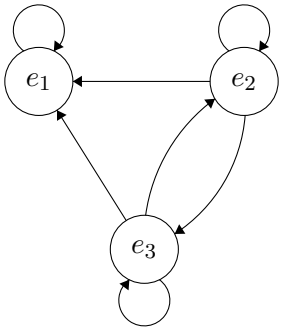
Denotaremos por E' o espaço dual de E,[4] $\mathcal{B}'$ a base dual de $\mathcal{B}$ e por Φ' o conjunto de operadores adjuntos φ' para $\varphi \in \Phi$.

Proposição 5.2.45. *Os conjuntos $\mathfrak{V} = (E, \mathcal{B}, \Phi)$ e $\mathfrak{V}' = (E', \mathcal{B}', \Phi')$ possuem a mesma complexidade de ciclo.*

[4] normalmente, o espaço dual é denotado por E^*, mas utilizaremos esta notação para evitar confusão com a altura de estrela.

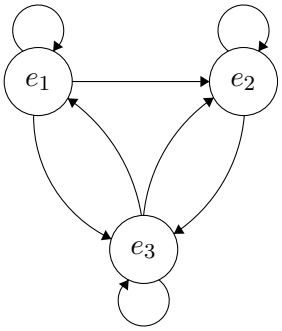


(a) *Grafo associado a* $\mathfrak{V}_1$.

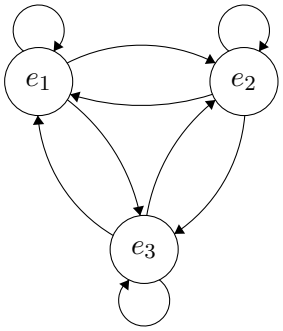


(b) *Grafo associado a* $\mathfrak{V}_2$.

Figura 5.10: *Grafos associados a* $\mathfrak{V}_1$ *e* $\mathfrak{V}_2$.



(a) *Grafo associado a* $\mathfrak{V}_3$.



(b) *Grafo associado a* $\mathfrak{V}_4$.

Figura 5.11: *Grafos associados a* $\mathfrak{V}_3$ *e* $\mathfrak{V}_4$.

Demonstração. É claro que b_j aparece na expansão de $\varphi(b_i)$ na base $\mathcal{B}$ se e somente se b'_i aparece na expansão de $\varphi'(b'_j)$ na base $\mathcal{B}'$. Portanto, os grafos associados a $\mathfrak{V}$ e $\mathfrak{V}'$ são opostos. Como grafos opostos têm a mesma complexidade de ciclo pelo lema 5.2.32, segue que $\mathfrak{V} = (E, \mathcal{B}, \Phi)$ e $\mathfrak{V}' = (E', \mathcal{B}', \Phi')$ possuem a mesma complexidade de ciclo. □

Corolário 5.2.46. (E, Φ) *e* (E', Φ') *possuem a mesma complexidade de ciclo.*

Demonstração. Basta tomar a base $\mathcal{B}$ que torna o valor de $\mathcal{C}(E, \mathcal{B}, \Phi)$ mínimo, e o resultado segue naturalmente da proposição 5.2.45. □

Observe que $h\colon \mathcal{B} \to \mathbb{N}$ é uma função altura para $\mathfrak{V}$ se e somente se a seguinte condição é válida:

Se $h(b)$ (resp. $h(b) \geq 1$), então para qualquer $\varphi \in \Phi$, a imagem $\varphi(b)$ é uma combinação linear de $b' < b$ (resp. de $b' < n(b)$)

Note que $\mathcal{B}$ deve ser totalmente ordenado.

Vamos generalizar levemente esta definição, considerando qualquer conjunto gerador. Sejam E, Φ como antes, e considere uma família finita totalmente ordenada$(b_i)_{i\in I}$ a qual gera E como um espaço vetorial, com a função $h\colon I \to \mathbb{N}$ (também chamada função altura) tal que a seguinte condição é satisfeita:

Se $h(j)$ (resp. $h(j) \geq 1$), então para qualquer $\varphi \in \Phi$, a imagem $\varphi(b_j)$ é uma combinação linear de b_i com $i < j$ (resp. de $i < n(j)$).

Como notação, adotaremos $\mathfrak{V}((b_i)_{i\in I}) = (E, (b_i)_{i\in I}, \Phi)$.

Veja que estamos considerando dessa forma as triplas $\mathfrak{V}((b_i)_{i\in I})$ em relação a um conjunto gerador, ao invés de nos restringirmos apenas às bases. Tal generalização será útil para verificar propriedades sobre a complexidade de ciclo de (E, Φ), como mostrado nos lemas seguintes.

Lema 5.2.47. *Sejam $E, \Phi, (b_i)_{i\in I}$ e $h\colon I \to \mathbb{N}$ como definidos anteriormente. Então,*

$$\mathcal{C}(E, \Phi) \leq \max(h)$$

Demonstração. Vamos remover sucessivamente elementos da família até obter uma base. Para fazer isso, vamos seguir a seguinte receita: Se $(b_i)_{i\in I}$ não é uma base, então para algum $k \in I$ temos a relação

$$b_k = \sum_{j<k} \alpha_j b_j$$

para $\alpha_j \in K$. Cada combinação linear de elementos b_i com $i < p$, onde $p \in I \cup \infty$ é também uma combinação linear de elementos b_i com $i < p$, e com $i \neq k$. Considere a família $(b_i)_{i\in I\setminus k}$ e a restrição $h' = h \restriction_{I\setminus k}$. Isso implica, pela observação acima, que para $j \in I \setminus k$ tal que $h(j) = 0$ (resp. $h(j) \geq 1$) a imagem $\varphi(b_j)$ é uma combinação linear de elementos b_i com $i \in I \setminus k$ e $i < j$ (resp. $i < n'(j)$). Então obtemos uma família menor e concluímos o resultado por indução. □

Lema 5.2.48. *Sejam E, Φ com $\mathcal{C}(E, \Phi) = m$. Seja F um subespaço de E que é invariante sob a ação dos elementos de Φ. Então E/F e F, com o conjunto de endomorfismos induzidos, possuem complexidade de ciclo menor ou igual a m.*

Demonstração. Sabemos que E possui uma base $\mathcal{B}$ com função altura h satisfazendo a seguinte condição: se $h(b)$ (resp. $h(b) \geq 1$), então para qualquer $\varphi \in \Phi$, a imagem $\varphi(b)$ é uma combinação linear de $b' < b$ (resp. de $b' < n(b)$), e com $\max(h) = m$. Consequentemente, E/F possui uma família geradora com função altura h satisfazendo tal condição com $\max(h) = m$. Pelo lema 5.2.47, a complexidade de ciclo E/F com o conjunto de endormorfismos induzidos é no máximo m.

Sabemos que para uma certa base $\mathcal{B}$ de E, a tripla $\mathfrak{V}(\mathcal{B}) = (E, \mathcal{B}, \Phi)$ possui complexidade de ciclo m, e pela proposição 5.2.45, $\mathcal{C}(\mathfrak{V}'(\mathcal{B}')) = m$. Seja $F^\perp$ o conjunto das funções lineares de E' que se anulam em F. Então, sabemos que

$$F' \sim \frac{E'}{F^\perp}$$

Note que cada endomorfismo $\varphi' \in \Phi'$ satisfaz

$$\varphi'(f) \in F^\perp,\ \forall f \in F^\perp.$$

Dessa forma, $\mathcal{C}(F', \Phi') \leq m$. Concluímos pelo corolário 5.2.46 que $\mathcal{C}(F, \Phi) \leq m$. □

Para um conjunto $\mathcal{M}$ de matrizes quadradas de ordem n, vamos associar o grafo G com conjunto de vértices $\{1, \ldots, n\}$ e arestas $i \to j$ se $M_{i,j} \neq 0$ para alguma matriz $M \in \mathcal{M}$. Dizemos que a complexidade de ciclo de $\mathcal{M}$ é a complexidade de ciclo do grafo G. De forma similar, dizemos que a complexidade de ciclo de uma representação (λ, μ, γ) é a complexidade de ciclo do conjunto de matrizes $\mu(a)$, com $a \in A$.

Definição 5.2.49. Uma matriz é chamada *genérica* se seus coeficientes são variáveis distintas não-comutativas.

Definição 5.2.50. A complexidade de ciclo de uma série formal $S \in K\langle\langle A\rangle\rangle$ é a complexidade de ciclo da representação linear minimal (λ, μ, γ) de S.

5.3 Cômputo da altura de estrela de uma série racional

A partir dos conceitos vistos na seção 5.2.4, estamos prontos para obter uma maneira de quantificar a altura de estrela de uma série racional.

Lema 5.3.1. *Seja (λ, μ, γ) uma representação linear de uma série S com $\mathcal{C}(\lambda, \mu, \gamma) \leq m$. Então a altura de estrela de S é no máximo m.*

Demonstração. Se $m = 0$, então o grafo associado à representação linear não possui bolas. Consequentemente, S é um polinômio (como visto no exemplo 5.1.2) e possui altura de estrela 0.

Podemos assumir que $m \geq 1$. Suponha que o grafo associado G seja fortemente conexo, com $\mathcal{C}(G) \leq m$. Então, $\mathcal{C}(G \setminus 1) \leq m - 1$. a matriz $M = \sum_{a \in A} a\mu(a)$ pode ser escrita como

$$M = \left(\begin{array}{c|c} M_1 & M_2 \\ \hline M_3 & M_4 \end{array}\right)$$

onde M_1 é de tamanho 1×1. Então a complexidade de ciclo de M_4 é no máximo $m - 1$, e por indução, cada entrada de M_4^* é uma série com altura de estrela no máximo $m-1$, Agora, tomando $N = M_1 + M_2 M_4^* M_3$, temos

$$M^* = \left(\begin{array}{c|c} N^* & N^* M_2 M_4^* \\ \hline M_4^* M_3 N^* & M_4^* + M_4^* M_3 N^* M_2 M_4^* \end{array}\right)$$

por conta do lema 3.3.10, tomando $\alpha = M_1, \beta = M_2, \gamma = M_3$ e $\delta = M_4$. Note que N é uma série com altura de estrela menor ou igual a $m - 1$, e consequentemente N^* tem altura de estrela no máximo m. Segue que cada entrada de M^* possui altura de estrela no máximo m, e S também.

Suponha agora que G não é fortemente conexo. Então a representação μ possui um bloco na forma triangular e cada bloco diagonal tem complexidade de ciclo no máximo m. Utilizando iterativamente o lema 3.3.37, concluímos o resultado. □

O próximo teorema é o principal dessa parte, e estabelece a não-limitação da altura de estrela de uma série racional.

Teorema 5.3.2. *Uma série racional em $K\langle\langle A \rangle\rangle$ tem altura de estrela no máximo m se, e somente se, possui uma representação minimal com complexidade de ciclo com valor máximo m.*

Demonstração. Seja $S \in K\langle\langle A \rangle\rangle$. Pelo lema 5.3.1, sabemos que se S possui uma representação linear minimal com complexidade de ciclo menor ou igual a m, então a altura de estrela de S é no máximo m.

Falta mostrar que, se S tem altura de estrela no máximo m, então possui uma representação linear minimal com complexidade de ciclo menor ou igual a m.

Primeiramente, vamos mostrar que, à luz das hipóteses do teorema, existe uma subespaço estável E de $K\langle\langle A \rangle\rangle$ contendo S de acordo com 3.3.6, e tal que o conjunto de endomorfismos

$$\Phi = \{\varphi \colon T \to a^{-1}T | a \in A\}$$

de E com complexidade de ciclo no máximo m.

Em vista do lema 5.2.47, é suficiente mostrar que E possui uma família geradora $(S_i)_{i \in I}$ com uma função altura $h \colon I \to \mathbb{N}$ satisfazendo a condição:

Se $h(j)$ (resp. $h(j) \geq 1$), então para qualquer $\varphi \in \Phi$, a imagem $\varphi(b_j)$ é uma combinação linear de b_i com $i < j$ (resp. de $i < n(j)$).

e com $\max(h) \leq m$. Para provar isso, vamos argumentar por indução no tamanho da expressão racional para S. Pela definição de altura de estrela 5.1.1, é suficiente mostrar o resultado quando

1. S é um polinômio e $m = 0$;
2. $S = T + U$ ou $S = UT$, com subespaços estáveis F, G (para U e T, respectivamente) e famílias $(T_i)_{i \in I}$ e $(U_j)_{j \in J}$ e funções altura k, ℓ, com $\max(k), \max(\ell) \leq m$;
3. $S = T^*$, com T própria e subespaço estável F, família $(T_i)_{i \in I}$ e função altura k com $\max(k) \leq m - 1$.

De fato:

1. Considere $\mathcal{F}$ uma família de palavras que aparecem em S, com uma ordem compatível com o tamanho, com $h = 0$. Note que $a^{-1}\omega$ tem tamanho menor que ω ou é 0 para qualquer palavra ω. Lembramos aqui que $a^{-1}\omega = \sum_{\omega \in A^*} (S, a\omega)\omega$, conforme definição 3.3.2.

2. Se $S = T + U$, assumindo que I e J são disjuntos, considere o subespaço estável $F + G$ gerado pela união $\bigcup_{\substack{i\in I\\ j\in J}} (T_i) \cup (U_j)$ das famílias, com uma ordem total estendendo as de I e J, e além disso, $i < j$ para $i \in I, j \in J$. Tome h que estende k e ℓ.

   Se $S = UT$, tome o subespaço estável $GT + F$, gerado pela família $\bigcup_{\substack{i\in I\\ j\in J}} (T_i) \cup (U_jT)$ com a mesma ordem e função altura como acima. Pelo item 4 da proposição 3.3.3, temos:

   $$a^{-1}(U_jT) = (a^{-1}U_j)T + (U_j, \varepsilon)(a^{-1}T) \in GT + F$$

   e como $a^{-1}U_j$ (respectivamente $a^{-1}T$) é uma combinação linear de $U_{j'}$ (respectivamente T_i), temos que a condição é satisfeita.

3. Se $S = T^*$, tome $E = KS + FS$, com $J = I \cup \{\omega\}$, com $\omega < i$, para todo $i \in I$, e considere $S_i = T_iS$ para $i \in I$, e $S_\omega = S$.

   Seja k estendendo h por $h(\omega) = m$. Pelo item 5 da proposição 3.3.3, segue:

   $$a^{-1}T^* = (a^{-1}T)T^* \Rightarrow a^{-1}S = (a^{-1}T)S$$

   Além disso, para $i \in I$,

   $$a^{-1}(T_iS) = (a^{-1}T_i)S + (T_i, \varepsilon)S$$

   Como $a^{-1}T_i$ é uma combinação linear de elementos de $T_{i'}$, a condição é satisfeita.

Pelo procedimento acima e com auxílio do lema 5.2.48, vemos que $S \circ K\langle A\rangle$ possui complexidade de ciclo no máximo m com respeito à Φ, uma vez que $S \circ K\langle A\rangle$ é um subespaço de E invariante pelos endomorfismos de Φ. Isso mostra, pela construção do lema 3.3.17, que S possui uma representação com complexidade de ciclo no máximo m e dimensão $\dim(S \circ K\langle A\rangle$. Como este último é o posto de S, concluímos que a representação é minimal, pelo corolário 3.3.22 e pelo teorema 3.3.26.

□

A força do resultado contido no teorema 5.3.2 habita na condição de minimalidade imposta. Isto é bem diferente do que ocorre com linguagens e autômatos.

O corolário seguinte estabelece o resultado almejado nesse capítulo. Com ele, garante-se que existem séries racionais com altura de estrela n para todo $n \in \mathbb{N}$, ou seja, que existem séries racionais com alturas de estrela arbitrariarmente grandes.

Corolário 5.3.3. *Seja $M \in \mathcal{M}_n(K)$ uma matriz genérica, tal como na definição 5.2.49. Então, cada entrada de M é uma série racional com altura de estrela n.*

Demonstração. Considere a série $S_{u,v} = (M^*)_{u,v}$. Pela segunda parte do teorema de Schützenberger 3.3.12, esta série possui uma representação linear (e_u, μ, e_u^T), onde μ leva elementos $a_{i,j}$ na matriz elementar $E_{i,j}$. Esta representação é minimal pela proposição 3.3.30. Posto isto, segue assim que $S_{u,v}$ tem altura de estrela menor ou igual a n, uma vez que um grafo com n vértices possui complexidade de ciclo no máximo n. Pelo lema 5.2.30, se $S_{u,v}$ possuir altura de estrela menor do que n, o teorema mostra que paraa alguma representação linear minimal $(\lambda', \mu', \gamma')$ de $S_{u,v}$ e para alguns i, j temos $(\mu'(a))_{i,j} = 0$ para cada letra a. Agora, temos que $\mu'(a) = P\mu(a)P^{-1}$ para algum $P \in \mathrm{GL}_n(K)$, pois duas representações lineares minimais são semelhantes, de acordo com 3.3.34. Assim, $(PE_{k,\ell}P^{-1})_{i,j} = 0$ para cada matriz elementar $E_{k,\ell}$. Isso é um absurdo, pois implicaria $(PNP^{-1})_{i,j} = 0$ para qualquer matriz N, e então teríamos $N_{i,j} = 0$ para qualquer N. □

Capítulo 6

Derivações no anel de grupo livre

6.1 Anel de Grupo

Apresentaremos nesta seção alguns conceitos básicos sobre anéis de grupos. Para um estudo mais aprofundado, veja por exemplo [25] e [26].

6.1.1 Definição e propriedades básicas

Definição 6.1.1. Seja G um grupo e R um anel. Definimos o *anel de grupo* RG como sendo o conjunto de todas as combinações lineares

$$\alpha = \sum_{g \in G} a_g g$$

onde $a_g \in R$ e a_g é nulo a menos de uma quantidade finita de termos. Definimos a soma em RG por:

$$\alpha + \beta = \left(\sum_{g \in G} a_g g \right) + \left(\sum_{g \in G} b_g g \right) = \sum_{g \in G} (a_g + b_g) g.$$

Definimos o produto em RG por:

$$\alpha\beta = \left(\sum_{g \in G} a_g g \right) \left(\sum_{h \in G} b_h h \right) = \sum_{g,h \in G} a_g b_h gh$$

Algumas observações importantes sobre a definição feita acima:

- Podemos também escrever o produto $\alpha\beta$ como $\sum_{u \in G} C_u u$, onde $C_u = \sum_{gh=u} a_g b_h$.
- RG é um anel com a adição e multiplicação definidos em 6.1.1.
- Dado um $\alpha \in RG$ e $\lambda \in R$, podemos definir a multiplicação por escalar à esquerda por:

  $$\lambda \cdot \alpha = \lambda \cdot \sum_{g \in G} a_g g = \sum_{g \in G} (\lambda a_g) g.$$

  A multiplicação por escalar à direita é definida analogamente.

  RG é um R-módulo à esquerda (à direita) com a multiplicação por escalar à esquerda (à direita) definida acima.

- Se R é um corpo, RG pode ser considerado como a R-álgebra de grupo de G.

- O elemento $a \in R$ pode ser identificado com o elemento $a \cdot 1_G \in RG$, e o elemento $g \in G$ pode ser identificado com o elemento $1 \cdot g \in RG$, e então R e G são subconjuntos de RG.
- RG é um R-módulo livre com base G.

Os anéis de grupo são caracterizados pela seguinte propriedade universal:

Teorema 6.1.2 (Propriedade universal dos anéis de grupo)**.** *Seja R um anel e G um grupo, e RG o correspondente anel de grupo. Seja T um anel, e considere $f\colon R \to T$ um homomorfismo de anéis e $\varphi\colon G \to \mathcal{U}(T)$ um homomorfismo de grupos, tais que $f(r)\varphi(g) = \varphi(g)f(r)$. Então, existe um único homomorfismo de anéis $\psi\colon RG \to T$ tal que $\psi \restriction_R = f$ e $\psi \restriction_G = \psi$.*

Para uma demonstração, veja o [25].

Definição 6.1.3. Seja RG um anel de grupo. Dizemos que um elemento $\alpha \in RG$ é *invertível* em RG se $\exists \beta \in RG$ tal que $\alpha\beta = \beta\alpha = 1$. Escrevemos $\beta^{-1} = \alpha$, e dizemos que α é uma *unidade* de RG. Dizemos que

$$\mathcal{U}(RG) = \{\alpha \in RG | \alpha \text{ é uma unidade de } RG\}$$

de *grupo das unidades* de RG.

Observe que $\mathcal{U}(RG)$ é um grupo multiplicativo.

Definição 6.1.4. Seja RG um anel de grupo. Dizemos que o *centro* de RG é

$$\mathcal{Z}(RG) = \{\zeta \in RG | \zeta\alpha = \alpha\zeta \ \forall \alpha \in RG\}$$

Definição 6.1.5. Dizemos que $\alpha \in RG$ é *central* se $\alpha \in \mathcal{Z}(RG)$. Dizemos que $\alpha \in RG$ é *idempotente* se $\alpha^2 = \alpha$. Dizemos que $\alpha \in RG$ é *idempotente-central* se $\alpha^2 = \alpha$ e $\alpha \in \mathcal{Z}(RG)$.

Exemplo 6.1.6. Seja $R = \mathbb{Z}_2$ e $G = C_2$. Podemos escrever os elementos de $\mathbb{Z}_2C_2$ a partir dos elementos de $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ e $C_2 = \{1, x\} = \langle x\rangle = \langle x | x^2 = 1\rangle$. Temos que:

$$\begin{aligned}
\mathbb{Z}_2C_2 &= \left\{ \sum_{g \in C_2} a_g g \,\middle|\, a_g \in \mathbb{Z}_2 \right\} \\
&= \{\overline{0}\cdot 1 + \overline{0}\cdot x, \overline{1}\cdot 1 + \overline{0}\cdot x, \overline{0}\cdot 1 + \overline{1}\cdot x, \overline{1}\cdot 1 + \overline{1}\cdot x\} \\
&= \{\overline{0}, \overline{1}, x, \overline{1} + x\} \\
&= \{0, 1, x, 1 + x\}
\end{aligned}$$

$+$	0	1	x	$1+x$
0	0	1	x	$1+x$
1	1	0	$1+x$	x
x	x	$1+x$	0	1
$1+x$	$1+x$	x	1	0

Tabela 6.1: *Adição em $\mathbb{Z}_2C_2$*

$\cdot$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	1	$1+x$
$1+x$	0	$1+x$	$1+x$	0

Tabela 6.2: *Multiplicação em $\mathbb{Z}_2C_2$*

Note que $(\mathbb{Z}_2C_2, +)$ é um grupo, mas $(\mathbb{Z}_2C_2, \cdot)$ não é um grupo, pois $0 \cdot \alpha = 0 \ \forall \alpha \in \mathbb{Z}_2C_2$. Além disso, $(\mathbb{Z}_2C_2 \setminus \{0\}, \cdot)$ também não é um grupo, pois $(1+x)^2 = 0$, e $0 \notin \mathbb{Z}_2C_2 \setminus \{0\}$.

O grupo das unidades de $\mathbb{Z}_2C_2$ é $\mathcal{U}(\mathbb{Z}_2C_2) = \{1, x\} \cong C_2$.

Observe pela 6.2 que todos os elementos comutam entre si. Logo, o centro de $\mathbb{Z}_2C_2$ é $\mathcal{Z}(\mathbb{Z}_2C_2) = \mathbb{Z}_2C_2$. Em geral, se R e G são comutativos, então $\mathcal{Z}(RG) = RG$.

6.1.2 Ideais e homomorfismos em RG

Seja R um anel e G um grupo. Como visto em 6.1.1, RG é um anel de grupo, e como RG é um anel, podemos tratar dos ideais e dos homomorfismos de anéis de RG.

Definição 6.1.7. Dizemos que a função

$$\begin{aligned} \varepsilon \colon RG &\longrightarrow R \\ \sum_{g\in G} a_g g &\longmapsto \sum_{g\in G} a_g \end{aligned}$$

é a *função de aumento.*

Observe que, para $r \in R$, temos que $\varepsilon(r \cdot 1) = r$. Então, para $rg, rh \in RG$, segue $\varepsilon(rg) = \varepsilon(rh) = r$. Porém, $rg \neq rh$, e ε não é injetora. Uma verificação direta comprova que

Proposição 6.1.8. *ε é um homomorfismo de anéis de RG em R. Em particular, é um homomorfismo de grupos.*

Demonstração. Sejam $\alpha = \sum_{g\in G} a_g g$ e $\beta = \sum_{g\in G} b_g g$. Então

$$\varepsilon(\alpha+\beta) = \varepsilon\left(\sum_{g\in G}(a_g+b_g)g\right) = \sum_{g\in G}(a_g+b_g) = \sum_{g\in G} a_g + \sum_{g\in G} b_g = \varepsilon(\alpha)+\varepsilon(\beta)$$

Agora, tomando $\alpha = \sum_{g\in G} a_g g$ e $\beta = \sum_{h\in G} b_h h$:

$$\varepsilon(\alpha\beta) = \varepsilon\left(\sum_{g,h\in G} a_g b_h gh\right) = \sum_{g,h\in G} a_g b_h = \left(\sum_{g\in G} a_g\right)\left(\sum_{h\in G} b_h\right) =$$

$$\varepsilon\left(\sum_{g\in G} a_g g\right)\varepsilon\left(\sum_{h\in G} b_h h\right) = \varepsilon(\alpha)\varepsilon(\beta).$$

Logo, ε é um homomorfismo de anéis.[1]

□

Sendo um homomorfismo de anéis, então podemos considerar o núcleo de ε, que é um ideal de RG. Desse modo, pelo Primeiro Teorema do Isomorfismo para anéis, temos que $\frac{RG}{\mathrm{Ker}(\varepsilon)} \cong R$.

Definição 6.1.9. Considere $\varepsilon \colon RG \to R$ a função de aumento definida em 6.1.7. Seja

$$\mathrm{Ker}(\varepsilon) = \left\{ \alpha = \sum_{g\in G} a_g g \,\middle|\, \varepsilon(\alpha) = \sum_{g\in G} a_g = 0 \right\}$$

o núcleo de ε. Então, $\mathrm{Ker}(\varepsilon)$ é chamado *ideal de aumento* ou *ideal fundamental* de RG, e é denotado usualmente por $\mathrm{Ker}(\varepsilon) = \Delta(RG) = \mathfrak{X}$. [2]

[1] Na verdade, ε é um epimorfismo, como dado na definição 2.1.10

[2] Seguindo as notações de [7] e [25].

Proposição 6.1.10. *O conjunto* $\mathfrak{G} = \{g - 1 \mid g \in G, g \neq 1\}$ *é uma base para* $\Delta(G)$ *sobre* R *como* R*-módulo. Ou seja,*

$$\Delta(G) = \left\{ \sum_{g \in G} a_g(g-1) \middle| g \in G, g \neq 1 \right\}$$

e os elementos de $\mathfrak{G}$ *são linearmente independentes sobre* R*.*

Demonstração. Seja $\alpha = \sum\limits_{g \in G} a_g g \in \Delta(G)$. Então $\sum\limits_{g \in G} a_g = 0$. Logo,

$$\alpha = \sum_{g \in G} a_g g - 0 = \sum_{g \in G} a_g g - \sum_{g \in G} a_g = \sum_{g \in G} a_g(g-1)$$

e portanto $\mathfrak{G}$ é um conjunto gerador para $\Delta(G)$. Vamos mostrar a independência linear: seja $\alpha = \sum\limits_{g \in G} a_g(g-1) = 0$. Então

$$0 = \sum_{g \in G \setminus 1} a_g g - \sum_{g \in G} a_g = \sum_{g \in G} a_g(g-1) = 0 \Leftrightarrow a_g = 0 \quad \forall g \in G.$$

Como G é linearmente independente sobre R, pela definição de RG dada em 6.1.1, segue o resultado. □

Proposição 6.1.11. *Seja* R *um anel comutativo. A função*

$$\begin{aligned} *\colon RG &\longrightarrow RG \\ \sum_{g \in G} a_g g &\longmapsto \sum_{g \in G} a_g g^{-1} \end{aligned}$$

é uma involução*, possuindo assim as seguintes propriedades:*

- $(\alpha + \beta)^* = \alpha^* + \beta^*$
- $(\alpha\beta)^* = \beta^*\alpha^*$
- $(\alpha^*)^* = \alpha$

Proposição 6.1.12. *Seja* $I \lhd R$ *e* G *um grupo. Então*

$$IG = \left\{ \sum_{g \in G} a_g g \middle| a_g \in I \right\} \lhd RG$$

Além disso,

$$\frac{RG}{IG} \cong \left(\frac{R}{I}\right) G$$

Demonstração. Observe que $(IG, +)$ é um grupo comutativo. Seja $\alpha = \sum\limits_{g \in G} a_g g \in IG$ e $\beta = \sum\limits_{h \in G} b_h h \in RG$. Veja que $a_g \in I$ e $b_h \in R$ para todo $g, h \in G$. Então:

$$\alpha\beta = \left(\sum_{g \in G} a_g g \right) \left(\sum_{h \in G} b_h h \right) = \sum_{g,h \in G} \underbrace{a_g b_h}_{\in I} gh \in IG.$$

Como $\gamma + \delta \in IG \ \forall \gamma, \delta \in IG$ e $\alpha\beta \in IG \ \forall \alpha \in IG$ e $\beta \in RG$, segue que IG é um ideal à esquerda de RG. Analogamente, usando a involução definida em 6.1.11, prova-se que IG é um ideal à direita de RG.

Para mostrar a segunda parte, basta considerar os homomorfismos de anéis

$$\begin{aligned} \pi\colon RG &\longrightarrow \tfrac{R}{I} \\ a &\longmapsto \overline{a} \end{aligned}$$

e

$$\begin{aligned} \varphi\colon RG &\longrightarrow \left(\tfrac{R}{I}\right) G \\ \sum_{g\in G} a_g g &\longmapsto \sum_{g\in G} \overline{a_g} g \end{aligned}$$

Segue então pela propriedade universal 6.1.2 que $Ker(\varphi) = IG$, e obtém-se diretamente que $\frac{RG}{IG} \cong \left(\frac{R}{I}\right) G$.

□

6.1.3 Derivações

Definição 6.1.13. Seja RG um anel de grupo. Dizemos que uma função R-linear à esquerda $D\colon RG \to RG$ é uma *derivação* se satisfaz

- $D(\alpha + \beta) = D(\alpha) + D(\beta)$
- $D(\alpha \cdot \beta) = D(\alpha)\varepsilon(\beta) + \alpha \cdot D(\beta)$

$\forall\ \alpha, \beta \in RG$.

Proposição 6.1.14. *Considere uma derivação em RG definida em 6.1.13. Então:*

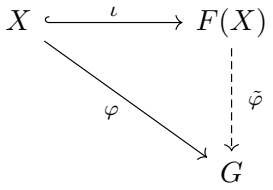
1. *Se $a \in R$, então $D(a) = 0$.*
2. *Se $\alpha = \sum_{g\in G} a_g g$, então*
$$D(\alpha) = D\left(\sum_{g\in G} a_g g\right) = \sum_{g\in G} a_g D(g)$$
3. *Se $\alpha_1, \alpha_2, \ldots, \alpha_\ell \in RG$, então*
$$D\left(\prod_{k=1}^{\ell} \alpha_k\right) = \sum_{i=1}^{\ell} \left(\left(\prod_{k=1}^{i-1} \alpha_k\right) \cdot D(\alpha_i) \cdot \left(\prod_{k=i+1}^{\ell} \varepsilon(\alpha_k)\right)\right)$$
4. *Se $g \in G$, então $D(g^{-1}) = -g^{-1}D(g)$.*

6.1.4 Grupos livres

Vamos apresentar alguns resultados e definições básicos sobre grupos livres que serão utilizadas ao longo desde trabalho. Para uma exposição mais detalhada, veja [17].

Por questão de conveniência, apresentamos aqui a definição de grupo livre por meio da propriedade universal, o que será justificado em 6.1.20, e também a algébrica de grupo livre, dada em 6.1.16, pois ambas serão úteis no decorrer desse capítulo.

Definição 6.1.15 (Via propriedade universal)**.** Seja X um conjunto. O *grupo livre* em X é o grupo $F(X)$ com uma função $\iota\colon X \to F(X)$, tal que, para G grupo e para $\varphi\colon X \to G$, existe um único homomorfismo de grupo $\tilde{\varphi}\colon F(X) \to G$ tal que $\tilde{\varphi} \circ \iota$ estende φ.



Seja X um conjunto arbitrário. Vamos definir o *grupo livre* gerado por X, denotado por $F(X)$.

Definimos uma palavra em X usualmente como em 3.1.1. Dizemos que uma palavra é *reduzida* se não contém subpalavras do tipo ss^{-1} ou $s^{-1}s$ para todo $s \in X$.

Considere

$$X^{-1} = \{x^{-1} \mid x \in X\}$$

Denotamos

$$X^{\pm 1} = X \cup X^{-1}$$

Para $y \in X^{\pm 1}$, definimos y^{-1} por

$$y^{-1} = \begin{cases} x^{-1}, & \text{se } y = x \in X, \\ x, & \text{se } y = x^{-1} \in X^{-1} \end{cases}$$

Uma expressão do tipo

$$\omega = \prod_{k=1}^{n} x_{i_k}^{\varepsilon_k} = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}, \quad x_{i_j} \in X, \varepsilon_j \in \{-1, 1\}$$

será chamada *palavra de grupo* de X.

Definição 6.1.16 (Algébrica)**.** Dizemos que um grupo G é *livre* se existe um conjunto[3] gerador $X = \{x_1, x_2, \ldots\}$ em G tal que toda palavra de grupo não-vazia reduzida em X define um elemento não trivial de G. Nesse caso, dizemos que G é livremente gerado por X (ou que G é livre por X) e X é chamado de base livre para G.

Um elemento de $F(X)$ é representado pela única palavra reduzida $\prod_{k=1}^{\ell} x_{j_k}^{\varepsilon_k}$, com $\varepsilon_k = \pm 1$ e $\varepsilon_k + \varepsilon_{k+1} \neq 0$ se $j_k = j_{k+1}$. O tamanho da palavra u é o comprimento ℓ da palavra reduzida que a representa. O elemento identidade 1 é representado pela palavra vazia, e possui tamanho 0. O inverso u^{-1} de u é representado pela palavra reduzida $\prod_{k=\ell}^{1} x_{j_k}^{-\varepsilon_k}$.

Dado um conjunto arbitrário X, podemos construir um grupo livre canônico com base X.

Uma redução elementar de uma palavra de grupo ω consiste em deletar uma subpalavra do tipo yy^{-1} de ω.

A *redução* de ω (ou processo de redução começando em ω) consiste numa sequência de aplicações de reduções elementares começando em ω e terminando numa palavra reduzida:

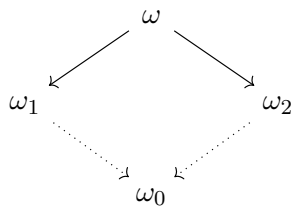
$$\omega \mapsto \omega_1 \mapsto \ldots \mapsto \omega_n \quad (\omega_n \text{ é reduzida})$$

ω_n é chamada *forma reduzida* de ω.

Em geral,poderiam existir reduções diferentes para ω. No entanto, verifica-se que todas as reduções possíveis de w acabam com a mesma forma reduzida. Para ver isso, precisaremos do lema seguinte:

Lema 6.1.17. *Para quaisquer duas reduções elementares $\omega \mapsto \omega_1$ e $\omega \mapsto \omega_2$ de uma palavra de grupo $\omega \in X$ existem reduções elementares $\omega_1 \mapsto \omega_0$ e $\omega_2 \mapsto \omega_0$, tais que o diagrama seguinte comuta:*

[3] O conjunto de geradores não precisa ser enumerável, mas a notação por meio de sequências é mais conveniente para $\langle x \rangle$ de qualquer maneira.



Demonstração. Sejam $\omega \xrightarrow{\lambda_1} \omega_1$ e $\omega \xrightarrow{\lambda_2} \omega_2$ reduções elementares de uma palavra ω. Há duas maneiras possíveis de se realizar as reduções λ_1 e λ_2 :

- **Reduções disjuntas:** Nesse caso,

  $$\omega = u_1 y_1 y_1^{-1} u_2 y_2 y_2^{-1} u_3, \quad y_i \in X^{\pm 1},$$

  e λ_i deleta a subpalavra $y_i y_i^{-1}$, para $i = 1, 2$. Então

  $$\omega \xrightarrow{\lambda_1} u_1 u_2 y_2 y_2^{-1} u_3 \xrightarrow{\lambda_2} u_1 u_2 u_3$$

  $$\omega \xrightarrow{\lambda_2} u_1 y_1 y_1^{-1} u_2 u_3 \xrightarrow{\lambda_1} u_1 u_2 u_3$$

  Consequentemente, o lema é válido.

- **Reduções sobrepostas:** Nesse caso, $y_1 = y_2$ e ω toma a forma

  $$\omega = u_1 y y^{-1} y u_2$$

  Então

  $$\omega = u_1 y (y^{-1} y) u_2 \xrightarrow{\lambda_2} u_1 y u_2,$$
  $$\omega = u_1 (y y^{-1}) y u_2 \xrightarrow{\lambda_1} u_1 y u_2;$$

  e o lema é válido.

□

Proposição 6.1.18. *Seja ω uma palavra de grupo em X. Então quaisquer duas reduções de ω :*

$$\omega \mapsto \omega_1' \mapsto \omega_2' \mapsto \ldots \mapsto \omega_n'$$

$$\omega \mapsto \omega_1'' \mapsto \omega_2'' \mapsto \ldots \mapsto \omega_n''$$

resultam na mesma forma reduzida, isto é, $\omega_n' = \omega_n''$.

Demonstração. Vamos provar o resultado por indução no tamanho da palavra. Se $|\omega| = 0$, então ω é reduzida e não há nada a se provar. Tome agora $|\omega| \geq 1$, e

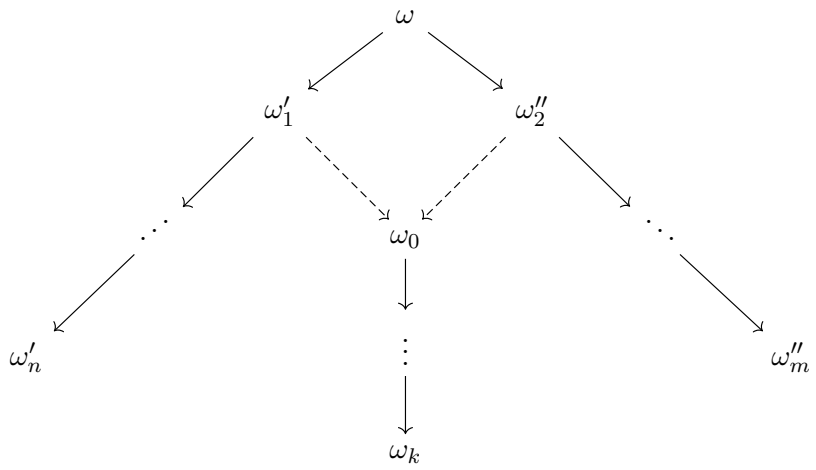
$$\omega \mapsto \omega_1' \mapsto \omega_2' \mapsto \ldots \mapsto \omega_n'$$

$$\omega \mapsto \omega_1'' \mapsto \omega_2'' \mapsto \ldots \mapsto \omega_m''$$

duas reduções de ω. Então, pelo lema 6.1.17, existem reduções elementares $\omega_1' \mapsto \omega_0$ e $\omega_2'' \mapsto \omega_0$. Considere o processo de redução para ω_0 :

$$\omega_0 \mapsto \omega_1 \mapsto \omega_2 \mapsto \ldots \mapsto \omega_k$$

Isso corresponde visualmente ao seguinte diagrama:



Por indução, todas as formas reduzidas da palavra ω_1' são iguais umas às outras, bem como as formas reduzidas de ω_1''. Como ω_k é uma forma reduzida tanto de ω_1' e de ω_1'', então

$$\omega_n' = \omega_k = \omega_m''$$

como desejado. Isso prova a proposição. □

Para uma palavra de grupo ω, denote por $\overline{\omega}$ sua respectiva forma reduzida. Seja $F(X)$ o conjunto de todas as palavras reduzidas em $X^{\pm 1}$. Para $u, v \in F(X)$, definimos a multiplicação como segue:

$$u \cdot v = \overline{uv}$$

Teorema 6.1.19. *O conjunto $F(X)$ forma um grupo com respeito à multiplicação $\cdot$. Este grupo é livre em X.*

Demonstração. A multiplicação definida acima é associativa:

$$u \cdot (v \cdot w) = (u \cdot v) \cdot w$$

para quaisquer $u, v, w \in F(X)$. Para ver isso, é suficiente mostrar que

$$\overline{\overline{(uv)}w} = \overline{u\overline{(vw)}}$$

para u, v, w dados. Observe que cada uma das palavras reduzidas $\overline{\overline{(uv)}w}, \overline{u\overline{(vw)}}$ pode ser obtida de uma palavra uvw por uma sequência de reduções elementares. Consequentemente, pela proposição 6.1.18:

$$\overline{\overline{uv}w} = \overline{uvw} = \overline{u\overline{vw}}.$$

Claramente, a palavra vazia ε é a identidade em $F(X)$ com respeito à multiplicação, ou seja,

$$w \cdot \varepsilon = \varepsilon \cdot w \quad \forall\ w \in F(X)$$

Seja $\omega = y_1 \cdots y_n \in X^{\pm 1}$. Então a palavra

$$\omega^{-1} = y_n^{-1} \cdots y_1^{-1}$$

é também reduzida e

$$\omega \cdot \omega^{-1} = \overline{y_1 \cdots y_n y_n^{-1} \cdots y_1^{-1}} = \varepsilon$$

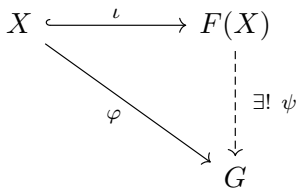
Consequentemente ω^{-1} é o inverso de ω. Isso mostra que $F(X)$ é um grupo. Note que X é um conjunto gerador para $F(X)$, e toda palavra não-vazia reduzida

$$\omega = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}$$

em $x^{\pm 1}$ define um elemento não trivial em $F(X)$ (a palavra ω em si). Consequentmente X é uma base livre de $F(X)$, portanto $F(X)$ é livre em X.

□

Teorema 6.1.20 (Propriedade Universal dos Grupos Livres)**.** *Seja X um conjunto e G um grupo, e considere a função $\varphi\colon X \to G$. Seja $F(X)$ o grupo livre em X. Então, existe um único homomorfismo de grupos $\psi\colon F(X) \to G$, tal que o seguinte diagrama é comutativo:*



onde ι representa a inclusão natural.

Demonstração. Seja $F(X)$ o grupo livre gerado por X e $\varphi\colon X \to G$ uma função de X no grupo G. Como $F(X)$ é livre em X então todo elemento $g \in F(X)$ é definido por uma única palavra reduzida

$$g = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n},$$

com $x_{i_j} \in X, \varepsilon_i \in \{-1, 1\}$. Considere

$$\begin{array}{rccl} \psi & : & F(X) & \longrightarrow & G \\ & & g & \longmapsto & \psi(g) = (\varphi(x_{i_1}))^{\varepsilon_1} \cdots (\varphi(x_{i_n}))^{\varepsilon_n} \end{array}$$

Vamos provar que ψ é um homomorfismo. De fato, tome $g, h \in F(X)$, e

$$g = y_1 \cdots y_n z_1 \cdots z_m,$$

$$h = z_m^{-1} \cdots z_1^{-1} y_{n+1} \cdots y_k$$

são as correspondentes palavras reduzidas em $X^{\pm 1}$, onde $y_i, z_j \in X^{\pm 1}$ e $y_n \neq y_{n+1}^{-1}$ (as subpalavras $y_1 \cdots y_n, z_1 \cdots z_m$ e $y_{n+1} \cdots y_k$ podem ser vazias). Então

$$gh = y_1 \cdots y_n z_1 \cdots z_m z_m^{-1} \cdots z_1^{-1} y_{n+1} \cdots y_k \Rightarrow$$

$$gh = y_1 \cdots y_n y_{n+1} \cdots y_k$$

é uma palavra reduzida em $X^{\pm 1}$ representando gh. Agora

$$\begin{array}{rcl} \psi(gh) & = & \psi(y_1 \cdots y_n y_{n+1} \cdot y_k) \\ & = & \psi(y_1) \cdots \psi(y_n)\psi(y_{n+1}) \cdots \psi(y_k) \\ & = & \psi(y_1) \cdots \psi(y_n)\psi(z_1) \cdots \psi(z_m)(\psi(z_m))^{-1} \cdots (\psi(z_1))^{-1}\psi(y_{n+1}) \cdots \psi(y_k) \\ & = & \psi(y_1 \cdots y_n z_1 \cdots z_m)\psi(z_m^{-1} \cdots z_1^{-1} y_{n+1} \cdots y_k) \\ & = & \psi(g)\psi(h) \end{array}$$

Consequentemente ψ é um homomorfismo. Claramente, ψ estende φ e o diagrama correspondente comuta. Observe que qualquer homomorfismo $\psi\colon F(X) \to G$ que faz o diagrama comutar deve satisfazer

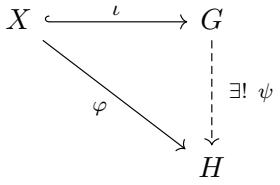
$$\psi(g) = (\varphi(x_{i_1}))^{\varepsilon_1} \cdots (\varphi(x_{i_n}))^{\varepsilon_n}$$

Isso mostra que $F(X)$ satisfaz a propriedade universal.

□

O resultado seguinte mostra que a propriedade universal 6.1.20 acima caracteriza os grupos livres.

Teorema 6.1.21. *Seja G um grupo com conjunto gerador $X \subseteq G$. Então G é livre em X se e somente se satisfaz a seguinte propriedade universal: toda função $\varphi\colon X \to H$ de X num grupo H pode ser estendida univocamente a um homomorfismo $\psi\colon G \to H$, de modo que o diagrama abaixo comute:*



onde ι representa a inclusão natural.

Demonstração. A ida já está provada em 6.1.20. Reciprocamente, suponha que o grupo G com conjunto gerador X satisfaz a propriedade universal. Tome $G = F(X)$, e defina a função

$$\begin{array}{rccc} \varphi & : & X & \longrightarrow & G \\ & & x & \longmapsto & x \end{array}$$

Então pela propriedade universal φ se estende a um único homomorfismo $\psi\colon G \to F(X)$. Seja ω uma palavra de grupo nreduzida nao-vazia em X. Então ω define um elemento $g \in G$ para o qual

$$\psi(g) = \omega \in F(X).$$

Consequentemente, $\psi(g) \neq \varepsilon$, e $g \neq 1$ em G. Isso mostra que G é um grupo livre em X. □

Observação: Os teoremas 6.1.20 e 6.1.21 foram enunciados aqui separadamente por motivos puramente didáticos.

6.1.5 Derivações no anel de grupo livre

Um elemento do anel de grupo livre $R[F(X)]$ é escrito como

$$f = \sum_{u \in F(X)} \alpha_u u, \quad \alpha_u \in R.$$

Assim como definido em 6.1.7, o homomorfismo $\varepsilon\colon RX \to R$ é responsável por enviar f em $\varepsilon(f) = \sum_{u \in F(X)} \alpha_u \varepsilon(u) = \sum_{u \in F(X)} \alpha_u$. O ideal fundamental $\mathrm{Ker}(\varepsilon) = \Delta(RX)$, que denotaremos por $\mathfrak{X}$, portanto, pode ser escrito como

$$\mathfrak{X} = \{f \in R[F(X)] | \varepsilon(f) = 0\}$$

Teorema 6.1.22. *Para cada gerador x_j de X, existe uma derivação correspondente $f \mapsto D_j f = f_{x_j} = \frac{\partial f}{\partial x_j}$, chamada* derivação com respeito a x_j, *que possui a propriedade*

$$\frac{\partial x_k}{\partial x_j} = \delta_{jk}$$

Além disso, existe uma e somente uma derivação $f \mapsto f'$ mapeando $x_1, x_2, \ldots$ em $h_1, h_2, \ldots$ de $R[F(X)]$, tais que os coeficientes de h_i estão em $\mathcal{Z}(R)$. É dado pela fórmula

$$f' = \sum \frac{\partial f}{\partial x_j} \cdot h_j(x)$$

Demonstração. Para mostrar que tal derivação existe, basta mostrar que

$$\mathcal{D}\colon R[F(X)] \longrightarrow \mathcal{M}_2(R[F(X)])$$
$$f \longmapsto \begin{pmatrix} f & D(f) \\ 0 & \varepsilon(f) \end{pmatrix}$$

é um homomorfismo. De fato,

$$\begin{aligned} \mathcal{D}(f)+\mathcal{D}(g) &= \begin{pmatrix} f & D(f) \\ 0 & \varepsilon(f) \end{pmatrix} + \begin{pmatrix} g & D(g) \\ 0 & \varepsilon(g) \end{pmatrix} \\ &= \begin{pmatrix} f+g & D(f)+D(g) \\ 0 & \varepsilon(f)+\varepsilon(g) \end{pmatrix} \\ &= \begin{pmatrix} f+g & D(f+g) \\ 0 & \varepsilon(f+g) \end{pmatrix} \\ &= \mathcal{D}(f+g) \end{aligned}$$

e também:

$$\mathcal{D}(f)\mathcal{D}(g) = \begin{pmatrix} f & D(f) \\ 0 & \varepsilon(f) \end{pmatrix}\begin{pmatrix} g & D(g) \\ 0 & \varepsilon(g) \end{pmatrix} =$$

$$\begin{pmatrix} fg & fD(g)+\varepsilon(g)D(f) \\ 0 & \varepsilon(f)\varepsilon(g) \end{pmatrix} = \begin{pmatrix} fg & D(fg) \\ 0 & \varepsilon(fg) \end{pmatrix} = \mathcal{D}(fg)$$

No nosso caso, temos a função

$$D_j\colon F(X) \longrightarrow \mathcal{M}_2(R[F(X)])$$
$$x_k \longmapsto \begin{cases} \begin{pmatrix} x_j & 1 \\ 0 & 1 \end{pmatrix}, & \text{se } k=j; \\ \begin{pmatrix} x_k & 0 \\ 0 & 1 \end{pmatrix}, & \text{caso contrário.} \end{cases}$$

Observe que as imagens dos x_k são inversíveis. Pela propriedade universal do grupos livres 6.1.20, segue que esta é única.

Além disso, $\mathcal{D}$ tal que

$$x_i \longmapsto \begin{pmatrix} x_i & h_i \\ 0 & \varepsilon(x_i) \end{pmatrix}$$

é uma derivação.

Vamos mostrar agora que

$$D(f) = \sum_j D_j(f)h_j,$$

onde $\frac{\partial f}{\partial x_j} = D_j$ também é uma derivação. Claramente, a soma de derivações é uma derivação:

$$\begin{aligned} D(f)+D(g) &= \sum_j D_j(f)h_j + \sum_j D_j(g)h_j \\ &= \sum_j (D_j(f)+D_j(g))h_j \\ &= \sum_j D_j(f+g)h_j \\ &= D(f+g) \end{aligned}$$

para $f, g \in R[F(X)]$.

Resta apenas verificar o produto para $D_j h_j$. De fato, para $a, b \in R[F(X)]$:

$$\begin{aligned} D_j(ab)h_j &= (D_j(a)\varepsilon(b) + aD_j(b))h_j \\ &= D_j(a)\varepsilon(b)h_j + aD_j(b)h_j \\ &= (D_j(a)h_j)\varepsilon(b) + a(D_j(b)h_j), \end{aligned}$$

o que é válido pois, como $h_i \in \mathcal{Z}(R)$, temos que $\varepsilon(b)h_j = h_j\varepsilon(b)$.

□

Teorema 6.1.23. *Seja $f \in R[F(X)]$, e D_j como dado em 6.1.22. Então, é válido o seguinte resultado, chamado de* fórmula fundamental:

$$f = \varepsilon(f) + \sum_j \frac{\partial f}{\partial x_j}(x_j - 1) \tag{6.1}$$

Demonstração. Podemos verificar facilmente que $D = Id - \varepsilon$ é uma derivação, pois

$$\begin{aligned} D(a+b) &= a + b - \varepsilon(a) - \varepsilon(b) \\ &= a - \varepsilon(a) + b - \varepsilon(b) \\ &= D(a) + D(b) \end{aligned}$$

e

$$\begin{aligned} D(ab) &= ab - \varepsilon(a)\varepsilon(b) \\ &= ab - \varepsilon(a)\varepsilon(b) + a\varepsilon(b) - a\varepsilon(b) \\ &= (a - \varepsilon(a))\varepsilon(b) + a(b - \varepsilon(b)) \\ &= D(a)\varepsilon(b) + aD(b) \end{aligned}$$

Além disso, temos que para $x_j \in X$,

$$D(x_j) = x_j - \varepsilon(x_j) = x_j - 1$$

Como visto em 6.1.22, tomando $h_j = x_j - 1$, sabemos que $\mathcal{D}(f) = \sum_j D_j(f)(x_j - 1)$ é uma derivação. Além disso, $\mathcal{D}(x_j) = x_j - 1$. Como $\mathcal{D} = D$ para todo $x_i \in X$ (ou seja, coincidem nos geradores), temos pela propriedade universal que se tratam da mesma derivação. Assim,

$$Id - \varepsilon = \sum_j D_j(x_j - 1)$$

Avaliando esta equação em $f \in R[F(X)]$, obtemos a fórmula fundamental desejada.

□

A derivada de potências de um elemento gerador pode ser facilmente obtida a partir da fórmula acima.

Proposição 6.1.24. *Seja x_j um gerador de $F(X)$. Então,*

$$D_j(x_j^p) = \begin{cases} \sum_{\ell=0}^{p-1} x_j^\ell, & \text{se } p \geq 1, \\ 0, & \text{se } p = 0, \\ -\sum_{\ell=p}^{-1} x_j^\ell, & \text{se } p \leq -1. \end{cases}$$

Demonstração. Seja $p \geq 1$. Então, utilizando a fórmula definida em 6.1.23, segue que

$$x_j^p = \varepsilon(x_j^p) + \sum_j D_j(x_j^p)(x_j - 1) \Rightarrow D_j(x_j^p)(x_j - 1) = x_j^p - 1 \Rightarrow D_j(x_j^p) = \sum_{\ell=0}^{p-1} x_j^\ell.$$

A demonstração para $p = 0$ e $p \leq -1$ é análoga.

□

Também podemos definir derivadas em $R[F(X)]$ de ordem superior.
As derivadas de ordem superior são definidas indutivamente

$$\frac{\partial^n f}{\partial x_{j_n}\partial x_{j_{n-1}}\cdots\partial x_{j_1}} = \frac{\partial}{\partial x_{j_n}}\left(\frac{\partial^{n-1} f}{\partial x_{j_{n-1}}\cdots\partial x_{j_1}}\right)$$

Uma notação alternativa para $\frac{\partial^n f}{\partial x_{j_n}\partial x_{j_{n-1}}\cdots\partial x_{j_1}}$ é $D_{j_n\cdots j_2}(f)$, que será utilizada conforme a conveniência.

A partir da definição 6.1.13, obtemos as propriedades básicas listadas na

Proposição 6.1.25. *Sejam $f, g \in R[F(X)]$. Então:*

$$D_{j_n\cdots j_1}(f+g) = D_{j_n\cdots j_1}f + D_{j_n\cdots j_1}g$$

$$D_{j_n\cdots j_1}(fg) = \sum_{p=1}^{n} fD_{j_n\cdots j_p}fD_{j_{p-1}\cdots j_1}\varepsilon(g) + fD_{j_n\cdots j_1}g$$

A demonstração será omitida por consistir essencialmente de verificação direta das igualdades.

Proposição 6.1.26. *Seja $f \in R[F(X)]$. Então*

$$f = \varepsilon(f) + \sum_{k=1}^{n-1}\left(\sum_{j_k, j_{k-1},\ldots,j_1} \varepsilon(D_{j_k,j_{k-1},\ldots,j_1}(f))\prod_{\ell=1}^{k}(x_{j_\ell}-1)\right) + \pitchfork(f),$$

onde

$$\pitchfork(f) = \sum_{j_n,\ldots,j_1}(D_{j_n\ldots j_1}f(1))(x_{j_n}-1)\cdots(x_{j_1}-1)$$

Demonstração. Aplicando a fórmula fundamental 6.1 para $f_{x_{j_1}}(x)$ etc. obtemos

$$D_{j_1}f(x) = \varepsilon(D_{j_1}(f)) + \sum_j (D_{jj_1}(f))(x_j - 1),$$

$$D_{j_2j_1}f(x) = \varepsilon(D_{j_2j_1}(f)) + \sum_j (D_{jj_2j_1}(f))(x_j - 1),$$

etc., e portanto segue para cada inteiro positivo n que

$$\begin{aligned} f(x) &= \varepsilon(f) + \sum_{j_1}(\varepsilon(D_{j_1}(f))(x_{j_1}-1) + \sum_{j_2j_1}(\varepsilon(D_{j_2j_1}(f)))(x_{j_2}-1)(x_{j_1}-1) + \ldots \\ &+ \sum_{j_{n-1},\ldots,j_1}\varepsilon(D_{j_{n-1}\ldots j_1}(f))(x_{j_{n-1}}-1)\cdots(x_{j_1}-1) + \sum_{j_n,\ldots,j_1}(D_{j_n\ldots j_1}f(1))(x_{j_n}-1)\cdots(x_{j_1}-1) \\ &= \varepsilon(f) + \sum_{k=1}^{n}\left(\sum_{j_k,j_{k-1},\ldots,j_1}\varepsilon(D_{j_k,j_{k-1},\ldots,j_1}(f))\prod_{\ell=1}^{k}(x_{j_\ell}-1)\right) + \pitchfork(f), \end{aligned}$$

onde

$$\pitchfork(f) = \sum_{j_n,\ldots,j_1}(D_{j_n\ldots j_1}f(1))(x_{j_n}-1)\cdots(x_{j_1}-1)$$

□

Diversas identidades relacionando as derivadas parciais podem ser deduzidas. Para uma lista detalhada, veja [10].

6.2 Série Central Descendente

6.2.1 Definição e exemplos

Definição 6.2.1. Seja G um grupo, e $g, h \in G$. Então

$$[g,h] = ghg^{-1}h^{-1}$$

é chamado *comutador* de g e h. Para $H, K \subseteq G$, $[H,K]$ denota o subgrupo de G gerado por todos os comutadores $[h,k]$, para $h \in H$ e $k \in K$.

Definição 6.2.2. A *série central descendente* $\{\gamma_\alpha(G)\}$ é uma série descendente de subgrupos indexada por um ordinal α, definida por

$$\gamma_\alpha(G) = \begin{cases} G, & \text{se } \alpha = 1; \\ \left[G, \gamma_{\alpha-1}(G)\right], & \text{se } \alpha \text{ é um ordinal sucessor;} \\ \bigcap\limits_{\beta<\alpha} \gamma_\beta(G), & \text{se } \alpha \text{ é um ordinal limite.} \end{cases}$$

Temos uma cadeia

$$G = \gamma_1(G) \supset \gamma_2(G) \supset \gamma_3(G) \supset \dots$$

A tabela 6.3 abaixo resume as séries centrais descendentes para alguns grupos conhecidos:

Grupo G	$\gamma_i(G), i \geq 2$
D_m, $m \geq 3$ ímpar	$\langle\sigma\rangle$
$D_{2^k m}$, $k \geq 1$, $m \geq 3$ ímpar	$\langle\sigma^{2^k}\rangle, i \geq k$
A_n, $n \geq 5$	A_n
A_4	V^4
S_n, $n \geq 3$	A_n
$\mathrm{GL}_2(K)$, $\lvert K\rvert > 2$	$\mathrm{SL}_2(K)$

Tabela 6.3: *Exemplos de séries centrais descendentes*

Proposição 6.2.3. *$\gamma_n(X)/\gamma_{n-1}(X)$ é um grupo abeliano livre com*

$$\psi_n = \frac{1}{n}\sum_{d|n} \mu(d) q^{\frac{n}{d}}$$

geradores, onde μ é a função de Möbius, dada por

$$\mu(n) = \begin{cases} 1, & \textit{se } n \textit{ é livre de quadrados e tem uma quantidade par de fatores primos;} \\ -1, & \textit{se } n \textit{ é livre de quadrados e tem uma quantidade ímpar de fatores primos;} \\ 0 & \textit{caso contrário.} \end{cases} \tag{6.2}$$

Demonstração. Veja [34].[5] □

Para mais detalhes sobre séries centrais descendentes, veja [22].

6.2.2 Estrutura do anel de grupo livre

As potências do ideal fundamental $\mathfrak{X}$ de um anel de grupo $R[F(X)]$ formam uma cadeia descendente de ideais

$$\mathfrak{X} \supset \mathfrak{X}^2 \supset \mathfrak{X}^3 \dots$$

[5] Uma prova alternativa também é dada em [10].

que está conectado com a série central descendente de $F(X)$.

Vamos analisar essas relações, e relacioná-las com as derivações em $R[F(X)]$.

Proposição 6.2.4. *Um elemento $f \in R[F(X)]$ pertence a $\mathfrak{X}^n$ se, e somente se a função de aumento avalidas em todas as suas derivadas de ordem $0, 1, \ldots, n-1$ se anulam.*

Demonstração. Se $f \in \mathfrak{X}^n$, então

$$f = \sum_k f_{1k} f_{2k} \cdots f_{nk},$$

onde $\varepsilon(f_{ik}) = 0$. Das propriedades básicas da derivação contempladas na proposição 6.1.25, segue que

$$\varepsilon(f) = \varepsilon(D_j(f)) = \cdots = \varepsilon(D_{j_{n-1}\cdots j_1}(f)) = 0$$

Reciprocamente, se a função de aumento se anula em todas as derivadas de ordem $0, 1, \ldots, n-1$, segue de 6.1.26 que

$$\begin{aligned} f &= \varepsilon(f) + \sum_{k=1}^{n} \left(\sum_{j_k, j_{k-1}, \ldots, j_1} \varepsilon(D_{j_k, j_{k-1}, \ldots, j_1}(f)) \prod_{\ell=1}^{k} (x_{j_\ell} - 1) \right) + \Updownarrow(f) \\ &= \Updownarrow(f) = \sum_{j_{n-1}, \ldots, j_1} (D_{j_{n-1}\ldots j_1}(f))(x_{j_{n-1}} - 1) \cdots (x_{j_1} - 1) \in \mathfrak{X}^n \end{aligned}$$

□

Definição 6.2.5. Definimos o *comprimento ou tamanho* de $f = a_1 u_1 + \ldots + a_m u_m$ não-nulo como sendo

$$\ell(f) = \max_{i=1,\cdots m} \{|u_i|\}$$

onde $|u_i|$ representa o tamanho da palavra u_i, como definido em 3.1.4, assumindo que $u_i \neq u_k$ para $i \neq k$ e todos os coeficientes são não-nulos. O tamanho de 0 é $\ell(0) = 0$.

Lema 6.2.6. *O tamanho de um elemento não-nulo $f \in \mathfrak{X}^n$ não é menor que $\frac{n}{2}$.*

Demonstração. Iremos mostrar o resultado por indução sobre n. A veracidade da afirmação para $n = 1, 2$ é óbvia.

Suponha $n \geq 3$ e tome f um elemento de $\mathfrak{X}^n$ tal que $\ell = \ell(f) < \frac{n}{2}$. Cada palavra reduzida que aparece em f deve terminar em um gerador ou o inverso de um gerador. Consequentemente,

$$f = \sum_j (g^{(j)} \cdot x_j + h^{(j)} \cdot x_j^{-1}),$$

onde $g^{(j)}$ e $h^{(j)}$ são polinômios livres de tamanho menor que ℓ. Assim, $D_j(f) = a^{(j)} - h^{(j)} \cdot x_j^{-1}$, onde $a^{(j)}$ possui comprimento menor que ℓ.

Portanto, se $i \neq j$, o comprimento de $D_{ij}(f)$ é menor que ℓ, e o tamanho de

$$D_j((D_j(f)) \cdot x_j) = D_{jj}(f) + D_j(f)$$

é menor que ℓ. Uma vez que $f \in \mathfrak{X}^n$, segue da proposição 6.2.4 que $D_{ij}(f)$, para $i \neq j$, e $D_j((D_j(f)) \cdot x_j)$ pertencem a $\mathfrak{X}^{n-2}$.

De nossa hipótese de indução, temos que $D_{ij}(f) = 0$, para $i \neq j$, e $D_{jj} f = -D_j f$. Desse modo, para qualquer j, segue da equação fundamental 6.1 que

$$D_j f = \varepsilon(D_j(f)) + \sum_i D_{ij}(f) \cdot (x_j - 1) = -(D_j(f)) \cdot (x_j - 1)$$

Da equação acima, segue que $D_j(f) = (D_j(f) + (D_j(f))(x_j - 1)x_j^{-1} = 0$. Assim, aplicando novamente 6.1, chegamos a $f = \varepsilon(f) + \sum_j (D_j(f)) \cdot (x_j - 1) = 0$. Concluímos portanto que, se f tem tamanho

$\ell < \frac{n}{2}$, então $f = 0$. Desse modo, segue que seu tamanho é $\ell \geq \frac{n}{2}$, ou seja, não é menor que $\frac{n}{2}$, como desejado. □

Teorema 6.2.7 (Unicidade da expansão de séries formais)**.** *Sejam $f, g \in R[F(X)]$. Se*

$$\varepsilon(f) = \varepsilon(g),$$

$$\varepsilon(D_j(f)) = \varepsilon(D_j(g)),$$

$$\varepsilon(D_{ij}(f)) = \varepsilon(D_{ij}(g)), \ldots,$$

então $f = g$.

Demonstração. Pelas condições apresentadas no enunciado, $\varepsilon(f-g) = \varepsilon(D_j(f)-D_j(g)) = \varepsilon(D_{ij}(f)-D_{ij}(g)) = \ldots = 0$, o que significa que a função de aumento avaliada em todas as derivadas de $f-g$ se anula. Consequentemente, pela proposição 6.2.4, $f - g \in \mathfrak{X}^n$ para todo n.

Assim, pelo lema 6.2.6, $f - g$ não possui tamanho positivo. Portanto, $f - g = 0 \Rightarrow f = g$. □

Corolário 6.2.8.

$$\bigcap_{n=1}^{\infty} \mathfrak{X}^n = \{0\}.$$

Podemos generalizar 6.2.4, com uma prova similar, para a seguinte proposição, onde J é qualquer ideal de $R[F(X)]$ que está contido em $\mathfrak{X}$:

Proposição 6.2.9. *f pertence a $J\mathfrak{X}^n$ se, e somente se, $f \in \mathfrak{X}$ e todas as suas derivadas de ordem n pertencem a J (nesse caso, suas derivadas de ordem i pertencem a $J\mathfrak{X}^{n-1}$, para $i = 0, 1, 2, \ldots, n$).*

A partir do exposto em [34], sabemos que um elemento $u \in F(X)$ pertence à n-ésima série central descendente $\gamma_n(X)$ se, e somente se, a função de aumento avaliada em todas as derivadas de u de ordem $1, \ldots, n-1$ se anula. Isso significa que $u \in \gamma_n(X)$ se e somente se $u - 1 \in \mathfrak{X}^n$. Assim:

Proposição 6.2.10. *O ideal $\mathfrak{X}^n$ determina o n-ésimo grupo central descendente $\gamma_n(X)$.*

Claro que $\gamma_n(X)$ é também determinado pelo ideal $\mathfrak{X}_n$ que corresponde a $\gamma_n(X)$.

Definição 6.2.11. $\mathfrak{X}_n$ é chamado n-ésimo *ideal central descendente*; $\mathfrak{X}_1 = \mathfrak{X}$ e $\mathfrak{X}_n$ é gerado pelos comutadores de anéis $gh - hg$, $g \in \mathfrak{X}_{n-1}, h \in \mathfrak{X}$.

Em vista de 6.2.10, o teorema 6.2.8 é equivalente a dizer que $\bigcap_{n\geq 1} \gamma_n(X) = \{1\}$. (para mais detalhes, veja [7])

Além disso, um homomorfismo φ de X em um grupo G manda $\gamma_n(X)$ em $\gamma_n(G)$ e $\mathfrak{X}^n$ em $\mathfrak{G}^n$. Dessa forma,

Proposição 6.2.12. *O ideal $\mathfrak{G}^n$ determina o n-ésimo grupo central descendente $\gamma_n(G)$.*

Consequentemente, $\bigcap_{n\geq 1} \mathfrak{G}^n$ determina $\bigcap_{n\geq 1} \gamma_n(G)$. Isso mostra que qualquer invariante de G formada a partir de sua série central descendente $G \supset \gamma_2(G) \supset \gamma_3(G) \supset \ldots$ deve ser calculável a partir da sequência de ideais $\mathfrak{G} \supset \mathfrak{G}^2 \supset \mathfrak{G}^3 \supset \ldots$. Vale a pena notar que esta última seqüência pode ser mais fácil de lidar, porque o anel quociente $\mathfrak{X}^n/\mathfrak{X}^{n-1}$ possui uma base explícita $(x_{j_1} - 1)(x_{j_2} - 1) \cdot \ldots \cdot (x_{j_n} - 1), j_1, j_2, \ldots, j_n = 1, \ldots, q$ de q^n elementos, enquanto pela proposição 6.2.3 sabemos que $\gamma_n(X)/\gamma_{n-1}(X)$ possui $\psi_n = \frac{1}{n}\sum_{d|n} \mu(d) q^{\frac{n}{d}}$ geradores.

6.3 Derivações e homomorfismos

Seja $X = \{x_1, \ldots, x_n\}$ um conjunto com n elementos. Denotamos por $F(X)$ o *grupo livre* em X, e por $K[F(X)]$ a K-álgebra de grupo de $F(X)$ com coeficientes no corpo K.

De forma análoga ao feito na definição 6.1.7, definimos a *função de aumento* como o homomorfismo de grupos

$$\begin{aligned} \varepsilon\colon K[F(X)] &\longrightarrow K \\ \sum_{g\in F(X)} \alpha_g g &\longmapsto \sum_{g\in F(X)} \alpha_g \end{aligned}$$

Com auxílio dos resultados obtidos nas seções 6.1 e 6.2, vamos estabelecer resultados sobre a K-álgebra $K[F(X)]$, como caso particular do estudo geral feito para anéis de grupo.

Tal como definido em 6.1.9, dizemos que o *ideal fundamental* ou *ideal de aumento* de ε é $I = \mathrm{Ker}(\varepsilon)$.

Proposição 6.3.1. *Considere a série central descendente*

$$I \supset I^2 \supset I^3 \supset \ldots$$

onde $I = Ker(\varepsilon)$. Então

$$\bigcap_{n\geq 1} I^n = \{0\}$$

Demonstração. □

Como explorado a partir da proposição 3.2.20, pode-se ver que as séries da forma $1-x_i \in K\langle\langle X\rangle\rangle$ são inversíveis, com inverso x_i^*. Sendo K um corpo, podemos mostrar que $1+x_i$ é inversível, e com isso concluir o seguinte resultado:

Proposição 6.3.2. *Considere a função φ, definida por:*

$$\begin{aligned} \varphi\colon X &\longrightarrow K\langle\langle X\rangle\rangle \\ x_i &\longmapsto 1 + x_i \end{aligned}$$

Então, φ pode ser estendida a um homomorfismo de K-álgebras $\Phi\colon K[F(X)] \to K\langle\langle X\rangle\rangle$.

Demonstração. Observe que a série $1 + x_i$ possui inversa $\sum_{n=0}^{\infty}(-1)^n x^n$ em $K\langle\langle X\rangle\rangle$. Logo, segue o resultado. □

O homomorfismo Φ pode ser caracterizado da seguinte forma:

$$\begin{aligned} \Phi\colon K[F(X)] &\longrightarrow K\langle\langle X\rangle\rangle \\ \sum_{g\in F(X)} \alpha_g g &\longmapsto \sum_{g\in F(X)} \alpha_g \varphi(g) \end{aligned}$$

Exemplo 6.3.3. Seja $X = \{a, b, c\}$, e considere o grupo livre $F(X)$. Tome o elemento $S = 6ab + 2a^2b + 3c^{-1} \in \mathbb{R}[F(X)]$. Vamos calcular $\Phi(S)$:

$$\Phi(S) = 6\varphi(ab) + 2\varphi(a^2b) + 3\varphi(c^{-1}) = 6(a+1)(b+1) + 2((a+1)^2(b+1)) + 3((c+1)^{-1}) =$$

$$6ab + 6a + 6b + 6 + 2a^2b + 2a^2 + 4ab + 4a + 2b + 2 + 3\left(\sum_{n=0}^{\infty}(-1)^n c^n\right) =$$

$$2a^2b + 2a^2 + 10ab + 10a + 8b + 11 + \sum_{n=1}^{\infty} 3(-1)^n c^n = \sum_{\omega\in A^*}(S,\omega)\omega,$$

Teorema 6.3.4. *A função* Φ*, definida em 6.3.2, é um homomofismo injetor.*

Demonstração. Note que, para $x_i \in X$, temos que

$$\varphi(x_i) = 1 + x_i = D_i(x_i^2)$$

Logo, temos que

$$\begin{aligned} \Phi\colon K[F(X)] &\longrightarrow K\langle\langle X\rangle\rangle \\ f &\longmapsto \sum \varepsilon(D_{i_r\ldots i_1}(f))x_{i_1}\ldots x_{i_r} \end{aligned}$$

Se $\sum \varepsilon(D_{i_r\ldots i_1}(f))x_{i_1}\ldots x_{i_r} = \sum \varepsilon(D_{i_r\ldots i_1}(g))x_{i_1}\ldots x_{i_r}$, então temos que

$$\sum \varepsilon(D_{i_r\ldots i_1}(f))x_{i_1}\ldots x_{i_r} - \sum \varepsilon(D_{i_r\ldots i_1}(g))x_{i_1}\ldots x_{i_r} = 0 \Rightarrow$$

$$\sum \varepsilon(D_{i_r\ldots i_1}(f-g))x_{i_1}\ldots x_{i_r} = 0 \Rightarrow \varepsilon(D_{i_r\ldots i_1}(f-g)) = 0$$

Isso significa que $\varepsilon(f) = \varepsilon(g), \varepsilon(D_j(f)) = \varepsilon(D_j(g)), \varepsilon(D_{ij}(f)) = \varepsilon(D_{ij}(g)), \ldots$ Assim, pelo teorema de unicidade das séries formais 6.2.7, temos que $f = g$. □

Capítulo 7

Séries formais parcialmente comutativas

7.1 Monoides livres parcialmente comutativos

Definição 7.1.1. Seja A um alfabeto. Uma *relação de comutação parcial* ϑ em A é um subconjunto simétrico e irreflexivo de $A \times A$, ou seja:

- $\forall a, b \in A$, $(b, a) \in \vartheta$ sempre que $(a, b) \in \vartheta$;
- $\forall a \in A$, $(a, a) \notin \vartheta$.

Podemos representar ϑ pelo seu *grafo de comutação* , que é o grafo não direcionado construído sobre A onde duas letras $(a, b) \in A$ estão relacionadas por uma aresta se, e somente se, $(a, b) \in \vartheta$.

Definição 7.1.2. O *monoide livre parcialmente comutativo* em A associado a ϑ é o monoide denotado por $M(A, \vartheta)$ o qual é definido pela presentação monoidal

$$M(A, \vartheta) = \langle A | [a, b] = 0, (a, b) \in \vartheta \rangle$$

onde $[a, b] := ab = ba$ denota o *Colchete de Lie*. Um elemento $\omega \in M(A, \vartheta)$ é chamado *palavra parcialmente comutativa.*

Definição 7.1.3. Denotamos por $F(A, \vartheta)$ o *grupo livre parcialmente comutativo*, definido pela presentação de grupo

$$F(A, \vartheta) = \langle A | [a, b] = 0, (a, b) \in \vartheta \rangle$$

Agora, $[a, b] = aba^{-1}b^{-1}$. Denotamos por π_ϑ a *projeção canônica* de A^* em $M(A, \vartheta)$, e por $\overline{\omega}$ a imagem em $M(A, \vartheta)$ de uma palavra $\omega \in A^*$ por π_ϑ. Como o tamanho de qualquer $u, v \in A^*$ tal que $\overline{u} = \overline{v}$ é o mesmo, (de acordo com a definição 3.1.4) podemos definir o *tamanho* (ou *grau*) $|\omega|$ de uma palavra $\overline{\omega} \in M(A, \vartheta)$ como o tamanho de qualquer representante $\omega \in A^*$ de $\overline{\omega}$. Denotamos por $M_n(A, \vartheta)$ o conjunto de elementos de tamanho n em $M(A, \vartheta)$.

Podemos definir uma congruência em A^* denotada por $\cong_\vartheta$, da seguinte forma:

$$\forall (a, b) \in \vartheta, ab \cong_\vartheta ba.$$

Pode-se mostrar então que $M(A, \vartheta) = A^* / \cong_\vartheta$. Para mais detalhes, veja [13] e [5].

Definição 7.1.4. Seja $\omega \in M(A, \vartheta)$. O *alfabeto inicial* de ω é o subconjunto de A denotado por $\mathcal{I}_A(\omega)$ e definido por

$$\mathcal{I}_A(\omega) = \{a \in A | \exists u \in M(A, \vartheta) : \omega = au\}$$

Exemplo 7.1.5. Seja o alfabeto $A = \{a, b, c, d\}$. Então, o conjunto

$$\vartheta = \{(a, b), (b, a), (b, c), (c, b), (c, d), (d, c)\}$$

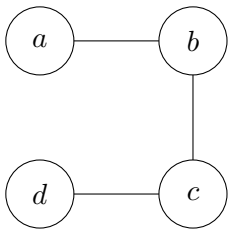


Figura 7.1: *Grafo de comutação G_ϑ.*

é uma relação de comutação parcial, pois satisfaz as condições da definição 7.1.1. O grafo de comutação para ϑ é $G_\vartheta = (A, \{(a,b),(b,c),(c,d)\}$, mostrado na figura 7.1.

Vamos determinar todas as palavras em $M_2(A,\vartheta)$. Como $\overline{ab} = \overline{ba}$, $\overline{bc} = \overline{cb}$ e $\overline{cd} = \overline{dc}$, temos que

$$M_2(A,\vartheta) = \{\overline{aa}, \overline{ab}, \overline{ac}, \overline{ad}, \overline{bb}, \overline{bc}, \overline{bd}, \overline{ca}, \overline{cc}, \overline{cd}, \overline{da}, \overline{db}, \overline{dd}\}$$

A classe de congruência da palavra $\omega = cdab$ é por exemplo $\{cdba, dcab\}$ e seu alfabeto inicial é $\mathcal{I}_A(\omega) = \{a, c\}$

Podemos também instituir uma ordem total em $M(A,\vartheta)$, respeitando a definição 2.5.7. Para isso, vamos assumir que A é totalmente ordenado por certa ordem total $\prec$. Associamos $\omega \in M(A,\vartheta)$ à palavra $\mathrm{std}(\omega) \in A^*$ que é a palavra maximal (para a ordem lexicográfica em A^*) na classe $\pi_\vartheta^{-1}(\omega)$. Isso nos permite equipar $M(A,\vartheta)$ com uma ordem total definida por

$$\omega \prec \eta \Leftrightarrow \mathrm{std}(\omega) \prec \mathrm{std}(\eta), \quad \forall \omega, \eta \in M(A,\vartheta)$$

Note que se $u, v \in M(a,\vartheta)$, então claramente temos $\mathrm{std}(u)\mathrm{std}(v) \in \pi_\vartheta^{-1}(uv)$. Segue imediatamente que $\mathrm{std}(uv) \succcurlyeq \mathrm{std}(u)\mathrm{std}(v) \succcurlyeq \mathrm{std}(u)$. Assim, qualquer prefixo de uma palavra parcialmente comutativa ω é menor que ω.

Exemplo 7.1.6. Vamos nos guarnecer novamente do alfabeto $A = \{a,b,c,d\}$ do exemplo 7.1.5 anterior. Considere uma relação parcialmente comutativa

$$\vartheta = \{(a,c),(c,a),(a,d),(d,a),(b,d),(d,b)\}$$

Vamos ordenar A tomando $a < b < c < d$. Então, temos por exemplo que

$$\pi_\vartheta^{-1}(acbd) = \{acbd, acdb, cabd, cadb, cdab\}, \quad \text{com } acbd < acdb < cabd < cadb < cdab$$

Assim sendo, $\mathrm{std}(acbd) = cdab$. Desfrutando de um raciocínio análogo, temos que

$$\pi_\vartheta^{-1}(bacd) = \{bacd, bcad, bcda\}, \quad \text{com } bacd < bcad < bcda$$

e portanto $\mathrm{std}(bacd) = bcda$. Observe que $bacd \cong_\vartheta bcda < cdab \cong_\vartheta acbd$. Logo $bacd < acbd$.

7.2 A série Central Descendente de $F(A,\vartheta)$

Estudamos a série central descendente de $F(A,\vartheta)$, conforme o exposto em [5], para obter resultados análogos aos desenvolvidos em 6.2, e usar essas propriedades na tentativa de generalizar os resultados do Capítulo 6 para séries parcialmente comutativas.

Precisaremos de algumas definições preliminares básicas de anéis e módulos graduados.

Definição 7.2.1. Seja G um grupo. Então dizemos que um *Anel graduado* é uma soma direta de grupos abelianos

$$R = \sum_{g \in G} R_g,$$

indexada por G, munida de uma estrutura de anel em R, de maneira que

$$\text{Se } x \in R_g \text{ e } y \in R_h, \text{ então } xy \in R_{gh} \ \forall g, h \in G.$$

Proposição 7.2.2. *Num anel graduado R, algumas propriedades importantes são:*

- $1 \in R_e$.
- *R_e é um subanel de R;*
- *R_g é um R_e-bimódulo para todo $g \in G$;*
- *Se $r \in R$ possui inverso à direita (esquerda), então r possui inverso à direita (esquerda) em $R_{g^{-1}}$.*

Definição 7.2.3. Dado um anel graduado R, um **subanel graduado** de R é um subanel S de R tal que para $x = (x_g)_{g \in G}$ em S e para todo $g \in G$, então $x_g \in S$.

Proposição 7.2.4. *Dado um anel graduado R e um conjunto $\mathcal{S}$ de subanéis graduados de R, então $\bigcap \mathcal{S}$ é um subanel graduado.*

Demonstração. Sabemos que $\bigcap \mathcal{S}$ é um subanel de R. Além disso, para $x \in \bigcap \mathcal{S}$, então para $g \in G$, então para $S \in \mathcal{S}$, então $x \in S$, aí $x_g \in S$; logo $x_g \in \bigcap \mathcal{S}$. □

Definição 7.2.5. Dado um anel graduado R, dizemos que um *ideal graduado* de R é um ideal I de R tal que, para $x = (x_g)_{g \in G} \in I$ e para todo $g \in G$, então $x_g \in I$. Como notação, adotamos $I \lhd R$.

Proposição 7.2.6. *Seja R um anel graduado e um conjunto $\mathcal{I} = \{I_j | j \in J\}$ de ideais graduados de R. Então,*

$$\bigcap_{I_j \in \mathcal{I}} I_j$$

é um ideal graduado, ou seja, a intersecção de ideias graduados também é um ideal graduado.

Demonstração. Sabemos que $\bigcap \mathcal{I} \lhd R$. Para $x \in \bigcap \bigcap_{I_j \in \mathcal{I}} I_j$ e para $g \in G$, para $I \in \mathcal{I}$, temos $x \in I$, e então $x_g \in I$.

Logo,

$$x_g \in \bigcap \mathcal{I}$$

□

Definição 7.2.7. Seja R um anel graduado. Um *R-módulo à direita graduado* é uma soma direta de grupos abelianos

$$X = \bigoplus_{g \in G} X_g$$

munida de uma estrutura de R-módulo à direita tal que

$$X_g R_h \subseteq X_{gh} \ \forall g, h \in G$$

O R-módulo à esquerda graduado é definido de maneira análoga.

Definição 7.2.8. Se R e S são anéis graduados, um *R, S-bimódulo graduado* é um R, S-bimódulo X tal que

$$R_a X_b S_c \subseteq X_{abc} \quad \forall a, b, c \in G$$

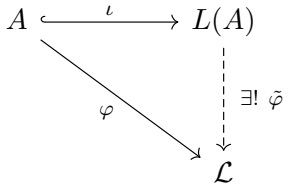
Exemplo 7.2.9. Seja A um anel qualquer. O exemplo canônica de anel graduado é $A[x_1, \ldots, x_n]$, que admite a graduação

$$A[x_1, \ldots x_n] = \bigoplus_{d \geq 0} A[x_1, \ldots, x_n]_d$$

em que $A[x_1, \ldots, x_n]_d$ é o A-módulo livre de posto $\binom{n+d-1}{d}$ com base dada pelos mônomios $x_1^{e_1} x_2^{e_2} \ldots x_n^{e_n}$ de grau

$$d = \sum_{i=1}^{n} e_i = e_1 + \ldots + e_n$$

Definição 7.2.10. Seja $(L(A), \iota)$ um par onde $L(A)$ é uma álgebra de Lie e $\iota\colon A \to L(A)$ é uma aplicação tal que, se existe $\varphi\colon A \to \mathcal{L}$, com $\mathcal{L}$ uma álgebra de Lie, então existe um único homomorfismo de álgebras de Lie $\tilde{\varphi}$ de $L(A)$ a $\mathcal{L}$ tal que o diagrama abaixo seja comutativo, isto é, existe único homomorfismo de álgebras de Lie tal que $\varphi = \tilde{\varphi} \circ \iota$.



Dizemos que $L(A)$ é a *álgebra de Lie livre gerada por A*.

Definição 7.2.11. Uma *Álgebra de Lie* é um espaço vetorial L sobre um corpo K com uma operação $[\cdot,\cdot]\colon L \times L \to L$, chamada *Colchete de Lie*, que satisfaz as seguintes condições para todos $x, y, z \in L$:

- $[\cdot,\cdot]$ é bilinear;
- $[x, x] = 0$ (o que implica $[x, y] = -[y, x]$)
- $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ (Identidade de Jacobi)

Definição 7.2.12. Um *homomorfismo de Álgebras de Lie* é uma aplicação linear $h \in \hom(L, \mathcal{L})$ entre aĺgebras de Lie L e $\mathcal{L}$ que é compatível com o colchete de Lie, ou seja:

$$h([x, y]) = [h(x), h(y)]$$

h é um *isomorfismo de Álgebras de Lie* se for um homomorfismo de Álgebras de Lie bijetor.

Definição 7.2.13. Uma *Álgebra de Lie graduada parcialmente comutativa* $L(A, \vartheta)$ é a Álgebra de Lie definida pela presentação

$$L(A, \vartheta) = \langle A : [a, b] = 0, (a, b) \in \vartheta \rangle$$

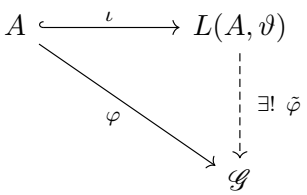
Pode-se mostrar que $L(A, \vartheta)$ é isomorfo à subálgebra de Lie de $K\langle\langle A, \vartheta \rangle\rangle$ gerada pelas letras de A (veja [4] para mais detalhes).

A Álgebra de Lie livre parcialmente comutativa é caracterizada pela seguinte propriedade universal:

Teorema 7.2.14 (Propriedade Universal da Álgebra de Lie livre parcialmente comutativa)**.** *Seja A um alfabeto munido de uma relação de comutação parcial ϑ. Se $\mathscr{G}$ é uma Álgebra de Lie, e se $\varphi\colon A \to \mathcal{G}$ é tal que*

$$[\varphi(a), \varphi(b)] = 0 \quad \forall (a, b) \in \vartheta,$$

então existe um único morfismo de álgebras de Lie $\tilde{\varphi}$ de $L(A, \vartheta)$ em $\mathscr{G}$ que estende φ, isto é, tal que o seguinte diagrama é comutativo:



Demonstração. Veja [4]. □

Teorema 7.2.15. *$L(A,\vartheta)$ e $L_n(A,\vartheta)$ são K-módulos livres.*

Exemplo 7.2.16. Seja $A=\{a,b,c\}$ e $\vartheta=\{(a,b),(b,a)\}$. Então podemos construir uma $\mathbb{Z}$-base da Álgebra de Lie $L(A,\vartheta)$. Temos que:

$$L(A,\vartheta)=L(\tau)\oplus L(a,c),\ \text{onde } \tau=\{b,[c,b],[c,[c,b]],\ldots\}$$

Onde $L(X)$ representa a álgebra de Lie livre gerada por X, como dado na definição 7.2.10.

Seja $(F_n(A,\vartheta))_{n\geq 1}$ a série central descendente de $F(A,\vartheta)$ Então, temos que os grupos quocientes $F_n(A,\vartheta)/F_{n+1}(A,\vartheta)$ são todos grupos abelianos, como constatado em 6.2.3, ou seja, $\mathbb{Z}$-módulos.

Portanto, podemos considerar de acordo com a definição 7.2.7 o $\mathbb{Z}$-módulo graduado $\mathcal{G}$ definido por

$$\mathcal{G}=\bigoplus_{n\in\mathbb{N}} F_n(A,\vartheta)/F_{n+1}(A,\vartheta)$$

É possível munir $\mathcal{G}$ com uma estrutura de $\mathbb{Z}$-álgebra de Lie, seguindo a definição 7.2.11. Para isso, sejam $g\in F_p(A,\vartheta)$ e $h\in F_q(A,\vartheta)$. Então

$$(g,h)=ghg^{-1}h^{-1}\in F_{p+a}(A,\vartheta),$$

pelas propriedades vistas em 6.2. Logo, podemos definir um colchete de Lie em $\mathcal{G}$ para todo $p,q\in\mathbb{N}^*$, pela relação

$$[g,h]=(g,h)\in F_{p+q}(A,\vartheta)/F_{(p+q)+1}(A,\vartheta),$$

para todo $g\in F_p(A,\vartheta)/F_{p+1}(A,\vartheta)$ e $h\in F_q(A,\vartheta)/F_{q+1}(A,\vartheta)$.

Também temos a seguinte propriedade:

$$(a,b)\in\vartheta\Rightarrow[aF_2(A,\vartheta),bF_2(A,\vartheta)]=(a,b)F_3(A,\vartheta)=F_3(A,\vartheta)=0$$

Então, de acordo com a propriedade universal 7.2.14 para $L_{\mathbb{Z}}(A,\vartheta)$, existe um único $\mathbb{Z}$-morfismo $\alpha\colon L_{\mathbb{Z}}(A,\vartheta)\to\mathcal{G}$ tal que

$$\alpha(a)=aF_2(A,\vartheta)\quad\forall\ a\in A$$

Teorema 7.2.17. *O $\mathbb{Z}$-morfismo α definido acima é um isomorfismo de Álgebras de Lie graduadas de $L_{\mathbb{Z}}(A,\vartheta)$ graduado por $(L_n(A,\vartheta))_{n\geq 1}$ em $\mathcal{G}$.*

Demonstração. Veja [2]. □

Como consequência imediata do teorema 7.2.17, obtemos que os dois grupos abelianos abaixo são isomorfos para todo $n\geq 1$:

$$F_p(A,\vartheta)/F_{p+1}(A,\vartheta)\simeq L_n(A,\vartheta)$$

Utilizando os resultados vistos, pode-se provar que

Teorema 7.2.18. *O grupo livre parcialmente comutativo $F(A,\vartheta)$ é um grupo residualmente nilpotente, ou seja, a intersecção de todos os termos da série central descendente é o grupo trivial:*

$$\bigcup_{n\geq 1} F_n(A,\vartheta)=\{1\}$$

Demonstração. Veja [5]. □

O teorema permite desenvolver uma versão análoga dos resultados das proposições 6.2.10 e 6.2.12 para o caso parcialmente comutativo, generalizando esses resultados do capítulo anterior.

7.3 Derivações de séries parcialmente comutativas

O objetivo será generalizar o que fizemos na seção 6 no contexto de séries parcialmente comutativos e inspirados na abordagem feita em [7].

Vamos procurar definir uma derivação no anel de grupo livre parcialmente comutativo $R[F(A,\vartheta)]$.

Definição 7.3.1. Seja $R[F(A,\vartheta)]$. Dizemos que uma função $R[F(A,\vartheta)] \to R[F(A,\vartheta)]$ é uma *derivação* se é R-linear e

- $D(\alpha\cdot\beta) = D(\alpha)\varepsilon(\beta) + \alpha\cdot D(\beta)$

$\forall \alpha,\beta \in R[F(A,\vartheta)]$,

Para verificar que está bem-definida para $R[F(A,\vartheta)]$, precisamos verificar que, dado $(x_i,x_j)\in\vartheta$, temos $D(\overline{x_ix_j}) = D(\overline{x_jx_i})$, pois são representantes de uma mesma classe de equivalência. De fato:

$$D(\overline{x_ix_j}) = D(\overline{x_i})\varepsilon(x_j) + \overline{x_i}D(\overline{x_j}) = D(\overline{x_i}) + \overline{x_i}D(\overline{x_j})$$

As propriedades da derivação definida acima são análogas às dadas em 6.1.14, pois estas essencialmente não dependem da comutatividade das variáveis envolvidas.

Proposição 7.3.2. *Considere a derivação em $R[F(A,\vartheta)]$, definida em 7.3.1. Então:*

1. *Se $a\in R$, então $D(a)=0$.*
2. *Se $\alpha = \sum\limits_{g\in F(A,\vartheta)} a_g g$, então*

$$D(\alpha) = D\left(\sum_{g\in F(A,\vartheta)} a_g g\right) = \sum_{g\in F(A,\vartheta)} a_g D(g)$$

3. *Se $\alpha_1,\alpha_2,\ldots,\alpha_\ell \in R[F(A,\vartheta)]$, então*

$$D\left(\prod_{k=1}^{\ell}\alpha_k\right) = \sum_{i=1}^{\ell}\left(\left(\prod_{k=1}^{i-1}\alpha_k\right)\cdot D(\alpha_i)\cdot\left(\prod_{k=i+1}^{\ell}\varepsilon(\alpha_k)\right)\right)$$

4. *Se $g\in F(A,\vartheta)$, então $D(g^{-1}) = -g^{-1}D(g)$.*

Analogamente como 6.1.7, podemos definir a função de aumento e o ideal de aumento para $R[F(A,\vartheta)]$.

Definição 7.3.3. Dizemos que o homomorfismo de grupos

$$\begin{aligned}\varepsilon\colon R[F(A,\vartheta)] &\longrightarrow R\\ \sum_{g\in R[F(A,\vartheta)]} a_g g &\longmapsto \sum_{g\in R[F(A,\vartheta)]} a_g\end{aligned}$$

é a *função de aumento* de $R[F(A,\vartheta)]$.

Proposição 7.3.4. *ε é um homomorfismo de anéis de $R[F(A,\vartheta)]$ em R.*

Demonstração. Sejam $\alpha = \sum\limits_{g\in G} a_g g$ e $\beta = \sum\limits_{g\in G} b_g g$. Então

$$\varepsilon(\alpha+\beta) = \varepsilon\left(\sum_{g\in G}(a_g+b_g)g\right) = \sum_{g\in G}(a_g+b_g) = \sum_{g\in G}a_g + \sum_{g\in G}b_g = \varepsilon(\alpha)+\varepsilon(\beta)$$

Agora, tomando $\alpha = \sum_{g\in G} a_g g$ e $\beta = \sum_{h\in G} b_h h$:

$$\varepsilon(\alpha\beta) = \varepsilon\left(\sum_{g,h\in G} a_g b_h gh\right) = \sum_{g,h\in G} a_g b_h = \left(\sum_{g\in G} a_g\right)\left(\sum_{h\in G} b_h\right) =$$

$$\varepsilon\left(\sum_{g\in G} a_g g\right)\varepsilon\left(\sum_{h\in G} b_h h\right) = \varepsilon(\alpha)\varepsilon(\beta).$$

Logo, ε é um homomorfismo de anéis.

□

Sendo um homomorfismo de anéis, então podemos considerar o núcleo de ε, que é um ideal de $R[F(A,\vartheta)]$.

Definição 7.3.5. Considere $\varepsilon\colon R[F(A,\vartheta)] \to R$ a função de aumento definida em 6.1.7. Seja

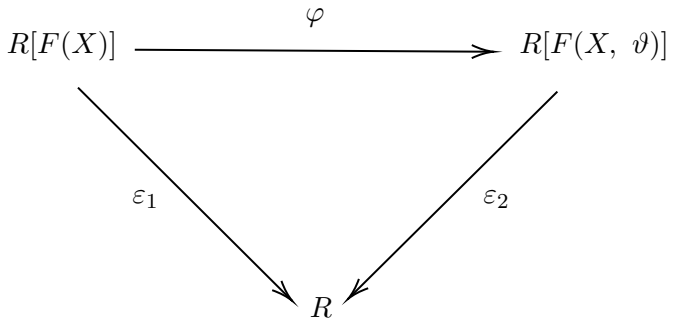
$$\mathrm{Ker}(\varepsilon) = \left\{ \alpha = \sum_{g\in F(A,\vartheta)} a_g g \,\middle|\, \varepsilon(\alpha) = \sum_{g\in F(A,\vartheta)} a_g = 0 \right\}$$

o núcleo de ε. Então, $\mathrm{Ker}(\varepsilon)$ é chamado *ideal de aumento* ou *ideal fundamental* de $R[F(A,\vartheta)]$.

Existe um homomorfismo natural entre $R[F(X)]$ em $R[F(X,\vartheta)]$, como definido abaixo:

$$\begin{array}{rccl} \varphi \;:\; & R[F(X)] & \longrightarrow & R[F(X,\vartheta)] \\ f = \sum_{g\in F(X)} \alpha_g g & \longmapsto & \varphi\left(\sum_{g\in F(X)} \alpha_g g\right) = \sum_{h\in F(X,\vartheta)} \left(\sum_{g\in\varphi^{-1}(h)} \alpha'_g\right) g \end{array}$$

Proposição 7.3.6. *Sejam ε_1 e ε_2 as funções de aumento de $R[F(X)]$ e de $R[F(X,\vartheta)]$, respectivamente. Então, $\varepsilon_1 = \varepsilon_2 \circ \varphi$, ou seja, o seguinte diagrama é comutativo:*



Demonstração. Claramente

$$\begin{aligned}
\varepsilon_1 \circ \varphi(f) &= \varepsilon_1(\varphi(f)) \\
&= \varepsilon\left(\varphi\left(\sum_{g\in F(X)} \alpha_g g\right)\right) \\
&= \varepsilon\left(\sum_{h\in F(X,\vartheta)} \left(\sum_{g\in\varphi^{-1}(h)} \alpha'_g\right) g\right) \\
&= \sum_{h\in F(X,\vartheta)} \left(\sum_{g\in\varphi^{-1}(h)} \alpha_g\right) \\
&= \sum_{g\in F(X)} \alpha_g \\
&= \varepsilon_2(f)
\end{aligned}$$

□

Definindo

$$\begin{array}{rccc} \pi & : & F(X) & \longrightarrow & F(X,\vartheta) \\ & & \omega & \longmapsto & \overline{\omega} \end{array},$$

temos por definição que

$$\varphi\left(\sum_{g\in F(X)} \alpha_g g\right) = \sum_{g\in F(X)} \alpha_g \pi(g)$$

Vamos agora definir uma derivação em $R[F(X,\vartheta)]$, com auxílio do que já fi feito em 6.1.22.

Proposição 7.3.7. *Considere*

$$\begin{array}{rccc} D & : & F(X,\vartheta) & \longrightarrow & F(X,\vartheta) \\ & & \overline{x_i} & \longmapsto & \overline{x_i} - 1 \end{array}$$

Então, D é uma derivação.

Demonstração. Basta mostrar que

$$\begin{aligned} \mathcal{D}\colon F(X,\vartheta) &\longrightarrow \mathcal{M}_2(RF(X,\vartheta)) \\ f &\longmapsto \begin{pmatrix} f & D(f) \\ 0 & \varepsilon(f) \end{pmatrix} \end{aligned}$$

é um homomorfismo. De fato, para $(x_i, x_j) \in \vartheta$:

$$\begin{aligned}
\mathcal{D}(x_i) + \mathcal{D}(x_j) &= \begin{pmatrix} x_i & D(x_i) \\ 0 & \varepsilon(x_i) \end{pmatrix} + \begin{pmatrix} x_j & D(x_j) \\ 0 & \varepsilon(x_j) \end{pmatrix} \\
&= \begin{pmatrix} x_i & x_i + 1 \\ 0 & \varepsilon(x_i) \end{pmatrix} + \begin{pmatrix} x_j & D(x_j) \\ 0 & \varepsilon(x_j) \end{pmatrix} \\
&= \begin{pmatrix} x_i + x_j & D(x_i) + D(x_j) \\ 0 & \varepsilon(f) + \varepsilon(g) \end{pmatrix} \\
&= \begin{pmatrix} f + g & D(f+g) \\ 0 & \varepsilon(f+g) \end{pmatrix} \\
&= \mathcal{D}(f+g)
\end{aligned}$$

e também:

$$\begin{aligned}\mathcal{D}(x_i)\mathcal{D}(x_j) &= \begin{pmatrix} x_i & x_i-1 \\ 0 & \varepsilon(x_i)\end{pmatrix}\begin{pmatrix} x_j & x_j-1 \\ 0 & \varepsilon(x_j)\end{pmatrix} \\ &= \begin{pmatrix} x_ix_j & x_i(x_j-1)+(x_i-1)\varepsilon x_j \\ 0 & \varepsilon(x_i)\varepsilon(x_j)\end{pmatrix} \\ &= \begin{pmatrix} x_ix_j & x_ix_j-1 \\ 0 & \varepsilon(x_ix_j)\end{pmatrix} \\ &= \mathcal{D}(x_ix_j)\end{aligned}$$

Como $(x_i, x_j) \in \vartheta$, precisamos verificar que $D(x_ix_j) = D(x_jx_i)$. Trivialmente:

$$\begin{aligned}\mathcal{D}(x_ix_j) &= \begin{pmatrix} x_ix_j & x_ix_j-1 \\ 0 & \varepsilon(x_ix_j)\end{pmatrix} \\ &= \begin{pmatrix} x_jx_i & x_jx_i-1 \\ 0 & \varepsilon(x_jx_i)\end{pmatrix} \\ &= \mathcal{D}(x_jx_i)\end{aligned}$$

Logo, D é uma derivação. □

Proposição 7.3.8. *Seja $\partial\colon R[F(X)] \to R[F(X)]$ a derivação*

$$\partial(f) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_j - 1),$$

onde $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$.

Seja D como definido na proposição 7.3.7, e $\varphi\colon R[F(X)] \to R[F(X,\vartheta)]$, dada por

$$\varphi\left(\sum_{g\in F(X)} \alpha_g g\right) = \sum_{g\in F(X)} \alpha_g \pi(g)$$

Então, o seguinte diagrama é comutativo:

$$\begin{array}{ccc} R[F(X)] & \xrightarrow{\ \varphi\ } & R[F(X,\vartheta)] \\ {\scriptstyle\partial}\downarrow & & \downarrow{\scriptstyle D} \\ R[F(X)] & \xrightarrow[\ \varphi\ ]{} & R[F(X,\vartheta)] \end{array}$$

Demonstração. Uma condição necessária é mostrar que, para $x \in \mathrm{Ker}\varphi$, então $\partial(x) \in \mathrm{Ker}\varphi$, ou seja, $\mathrm{Ker}\varphi$ é ∂-invariante.

Sendo $N = \langle \mu_{i_1j_1}\mu_{i_2j_2}, \ldots \mu_{i_kj_k} | ((i_n, j_n))_{1\le n\le k} \in \vartheta\rangle$, temos que $\mathrm{Ker}\varphi = \langle\{n-1 : n \in N\}\rangle$. De fato, para $x_ix_jx_i^{-1}x_j^{-1} \in N$, temos que

$$\begin{aligned}\partial(x_ix_jx_i^{-1}x_j^{-1} - 1) &= \partial(x_ix_jx_i^{-1}x_j^{-1}) \\ &= \sum_{k=1}^{n} \frac{\partial(x_ix_jx_i^{-1}x_j^{-1})}{\partial x_k}(x_k - 1) \\ &= \frac{\partial(x_ix_jx_i^{-1}x_j^{-1})}{\partial x_i}(x_i - 1) + \frac{\partial(x_ix_jx_i^{-1}x_j^{-1})}{\partial x_j}(x_j - 1)\end{aligned}$$

Temos que:

$$\begin{aligned}
\partial_i(x_ix_jx_i^{-1}x_j^{-1}) &= \partial_i(x_ix_jx_i^{-1}) + x_ix_jx_i^{-1}\underbrace{\partial_i(x_j^{-1})}_{=0} \\
&= \partial_i(x_ix_jx_i^{-1}) \\
&= \partial_i(x_ix_j) + x_ix_j\partial_i(x_i^{-1}) \\
&= \partial_i(x_ix_j) - x_ix_jx_i^{-1} \\
&= \partial_i(x_i) + x_i\underbrace{\partial_i(x_j)}_{=0} - x_ix_jx_i^{-1} \\
&= 1 - x_ix_jx_i^{-1}
\end{aligned}$$

$$\begin{aligned}
\partial_j(x_ix_jx_i^{-1}x_j^{-1}) &= \partial_j(x_ix_jx_i^{-1}) + x_ix_jx_i^{-1}\partial_j(x_j^{-1}) \\
&= \partial_j(x_ix_jx_i^{-1}) - x_ix_jx_i^{-1}x_j^{-1} \\
&= \partial_j(x_ix_j) + x_ix_j\underbrace{\partial_j(x_i^{-1})}_{=0} - x_ix_jx_i^{-1}x_j^{-1} \\
&= \partial_j(x_ix_j) - x_ix_jx_i^{-1}x_j^{-1} \\
&= \underbrace{\partial_j(x_i)}_{=0} + x_i\partial_j(x_j) - x_ix_jx_i^{-1}x_j^{-1} \\
&= x_i - x_ix_jx_i^{-1}x_j^{-1}
\end{aligned}$$

Assim:

$$\begin{aligned}
\partial(x_ix_jx_i^{-1}x_j^{-1} - 1) &= \tfrac{\partial(x_ix_jx_i^{-1}x_j^{-1})}{\partial x_i}(x_i - 1) + \tfrac{\partial(x_ix_jx_i^{-1}x_j^{-1})}{\partial x_j}(x_j - 1) \\
&= (1 - x_ix_jx_i^{-1})(x_i - 1) + (x_i - x_ix_jx_i^{-1}x_j^{-1})(x_j - 1) \\
&= x_i - x_ix_j - 1 + x_ix_jx_i^{-1} + x_ix_j - x_i + x_ix_jx_i^{-1} - x_ix_jx_i^{-1}x_j^{-1} \\
&= x_ix_jx_i^{-1}x_j^{-1} - 1
\end{aligned}$$

Seja $u_{ij} = x_ix_jx_i^{-1}x_j^{-1}$. Vamos mostrar agora que $\partial(x) \in \mathrm{Ker}\varphi$ para todo $x \in \mathrm{Ker}\varphi$. Tome

$$x = f_1^{-1}u_{i_1j_1}f_1f_2^{-1}u_{i_2j_2}f_2f_3^{-1}u_{i_3j_3}f_3 \ldots f_k^{-1}u_{i_kj_k}f_k$$

Usaremos indução. Temos como caso base que:

$$\begin{aligned}
\partial f_1^{-1}u_{i_1j_1}f_1) &= \partial(f_1^{-1}u_{i_1j_1}) + f_1^{-1}u_{i_1j_1}\partial(f_1) \\
&= \partial(f_1^{-1}) + f_1^{-1}\partial(u_{i_1j_1}) + f_1^{-1}u_{i_1j_1}\partial(f_1) \\
&= \underbrace{\left(f_1^{-1}(u_{i_1j_1} - 1)\right)}_{\in \mathrm{Ker}\varphi}\partial(f_1) + \underbrace{\partial(u_{i_1j_1})}_{\in \mathrm{Ker}\varphi} \in \mathrm{Ker}\varphi
\end{aligned}$$

Suponha $\partial(f_1^{-1}u_{i_1j_1}f_1f_2^{-1}u_{i_2j_2}f_2 \ldots f_k^{-1}u_{i_kj_k}f_k) \in \mathrm{Ker}\varphi$. Então, temos que

$$\partial(f_1^{-1}u_{i_1j_1}f_1f_2^{-1}u_{i_2j_2}f_2 \ldots f_k^{-1}u_{i_kj_k}f_kf_{k+1}^{-1}u_{i_{k+1}j_{k+1}}f_{k+1}) =$$

$$\partial(f_1^{-1}u_{i_1j_1}f_1f_2^{-1}u_{i_2j_2}f_2 \ldots f_k^{-1}u_{i_kj_k}f_k) + f_1^{-1}u_{i_1j_1}f_1f_2^{-1}u_{i_2j_2}f_2 \ldots f_k^{-1}u_{i_kj_k}f_k\partial(f_{k+1}^{-1}u_{i_{k+1}j_{k+1}}f_{k+1}) \in \mathrm{Ker}(\varphi)$$

Se $\sum_{n\in N} z(n-1) \in \mathrm{Ker}(\varphi)$, com $z \in K[F(X)]$, então

$$\partial\left(\sum_{n\in N} z(n-1)\right) = \left(\sum_{n\in N}\partial(z)\right)\underbrace{\varepsilon(n-1)}_{=0} + \sum_{n\in N} z\partial(n-1) =$$

$$\sum_{n\in N} z\partial(n-1) \in \mathrm{Ker}(\varphi)$$

Sejam ω_1 e ω_2 tais que $\varphi(\omega_1) = \varphi(\omega_2) = \overline{\omega}$. Então,

$$\pi(\omega_1) = \pi(\omega_2) \Leftrightarrow \omega_1\omega_2^{-1} \in N.$$

Precisamos verificar que $\varphi(\partial(\omega_1)) = \varphi(\partial(\omega_2))$. Observe que $\partial(\omega_1) - \partial(\omega_2) \in \mathrm{Ker}(\varphi)$.

$$\partial(\omega_1 - \omega_2) = \partial((\omega_1\omega_2^{-1} - 1)\omega_2) \in \mathrm{Ker}(\varphi)$$

Logo,

$$\varphi(\partial(\omega_1 - \omega_2)) = 0 \Rightarrow \varphi(\partial(\omega_1)) - \varphi(\omega_2)) = 0 \Rightarrow \varphi(\partial(\omega_1)) = \varphi(\omega_2))$$

Concluímos então que os resultados valem independentemente do representante escolhido para $\overline{\omega}$. Isso nos permite simplificar as contas. Considere então $\overline{x_i}$, e sua escolha natural $\overline{x_i} = \varphi(x_i)$.

Mostraremos que $D \circ \varphi = \varphi \circ \partial$. Veja que:

$$D(\varphi(x_i)) = D(\overline{x_i}) = \overline{x_i} - 1$$

$$\varphi(\partial(x_i)) = \varphi(x_i - 1) = \overline{x_i} - 1$$

Além disso, para $\omega = x_1x_2$, temos

$$D(\varphi(x_1x_2)) = D(\overline{x_1x_2}) = \overline{x_1x_2} - 1$$

$$\varphi(\partial(x_1x_2)) = \varphi(\partial(x_1) + x_1\partial(x_2)) =$$

$$\varphi((x_1 - 1) + x_1(x_2 - 1)) = \varphi(x_1x_2 - 1) = \overline{x_1x_2} - 1$$

Por indução, seja $D(\varphi(\omega)) = \varphi(\partial(\omega))$ para $\omega = x_1x_2\ldots x_k$. Então, para $\omega' = x_1x_2\ldots x_kx_{k+1}$, segue que

$$\begin{aligned} D(\varphi(\omega')) &= D(\varphi(\omega x_{k+1})) \\ &= D(\overline{\omega x_{k+1}}) \\ &= D(\overline{\omega}) + \overline{\omega}D(\overline{x_{k+1}}) \\ &= D(\overline{\omega}) + \omega(\overline{x_{k+1}} - 1) \end{aligned}$$

e também:

$$\begin{aligned} \varphi(\partial(\omega')) &= \varphi(\partial(\omega x_{k+1})) \\ &= \varphi(\partial(\omega)\varepsilon(x_{k+1}) + \omega\partial(x_{k+1})) \\ &= \varphi(\partial(\omega) + \omega(x_{k+1} - 1)) \\ &= \varphi(\partial(\omega)) + \varphi(\omega(x_{k+1} - 1)) \\ &= \varphi(\partial(\omega)) + \varphi(\omega)\varphi(x_{k+1} - 1) \\ &= D(\overline{\omega}) + \omega(\overline{x_{k+1}} - 1) \end{aligned}$$

Portanto, para todo $\omega \in F(X)$, temos que $D(\varphi(\omega)) = \varphi(\partial(\omega))$.

Logo, o diagrama é comutativo. □

Utilizaremos a comutatividade do diagrama para estender nossa derivação definida em 7.3.7 para $R[F(X, \vartheta)]$. De fato, podemos tomar $\forall \overline{f} \in R[F(X, \vartheta)]$:

$$D(\overline{f}) = \sum_i \varphi\left(\frac{\partial\varphi^{-1}(\overline{f})}{\partial x_i}\right)(\overline{x_i} - 1) = \sum_i \varphi\left(\frac{\partial f}{\partial x_i}\right)(\overline{x_i} - 1)$$

Ou seja, obtemos uma fórmula para calcular $D(\overline{f})$ para qualquer elemento de $R[F(X, \vartheta)]$.

Capítulo 8

Conclusões

A partir dos estudos e pesquisas realizados, conseguimos compreender diversos aspectos sobre a teoria de Séries Racionais Não-Comutativas, demonstrando importantes resultados, como o Teorema de Schützenberger 3.3.12, que trata da equivalência dos conceitos de séries reconhecíveis e racionais. Além disso, foi desenvolvido o conceito de representação linear de uma série racional, e estudamos as representações lineares minimais para uma série, constatando que duas representações lineares minimais quaisquer são semelhantes no teorema 3.3.34. Também estudamos conceitos relacionados à álgebras e ideais sintáticos de séries formais, e fizemso a demonstração do teorema de Paz-Carlyle-Fliess 3.3.26, que trata da relação entre o posto de uma série formal, a dimensão de seu ideal sintático e o posto de suas respectiva matriz de Hankel.

Neste trabalho, demonstramos a "trivialidade" das identidades racionais, ou seja, se K é um anel comutativo, então todas as identidades racionais sobre K, ou seja, $E \equiv F$ quando $\text{eval}(E) = \text{eval}(F)$, na qual dizemos que $E \equiv F$ ou (E, F) é uma identidade K-racional, são a grosso modo trivias. Isso significa que mostramos que todas as identidades racionais são consequências do fato de que S^* é o inverso de $\varepsilon - S$, para qualquer série própria S. Além disso, utilizamos conceitos de Teoria dos Grafos e Autômatos para obter formas de computar a altura de estrela de uma série racional, e demonstramos o importante teorema 5.3.2 sobre a não-limitação da altura de estrela.

O estudo dos anéis de grupos e dos grupos livres permitiu compreender melhor as derivações nos anéis de grupo livre, estabelecendo uma fórmula fundamental para a derivação *canônica* tal que $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$ para cada gerador x_i de X, o que nos permitiu explorar relações entre o anel de séries formais e a álgebra de grupo livre, demonstrando a existência de um homomorfismo injetor de $K[F(X)]$ em $K\langle\langle X\rangle\rangle$. O estudos das séries central descendentes possibilitou uma melhor abordagem para o entendimento da estrutura do anel de grupo livre, relacionando á série central descendente do grupo às potências do ideal fundamental de $R[F(X)]$.

Além disso, conseguimos fazer a genealização de alguns dos resultados vistos para o caso parcialmente comutativo, com algumas adaptações, como a noção de derivações no anel de grupo livre parcialmente comutativo e a obtenção de uma fórmula geral para a generalização em $R[F(X, \vartheta)]$ da derivação dada por $D(x_i) = x_i - 1$ para $x_i \in F(X, \vartheta)$ em termos da derivação estudada no artigo [7].

Apêndice A

Semigrupos de matrizes

Neste apêndice, apresentamos brevemente os conceitos relacionados a semigrupos de matrizes, como subsídio ao melhor entendimento do lema 3.3.37, e o contexto no qual este se insere. Para mais detalhes, recomendamos a consulta de [27] e [18].

A.1 O Teorema de Mandel-Simon

Lema A.1.1. *Seja α um morfismos de A^* em um monoide finito M. Então, para cada palavra ω de tamanho maior ou igual a $|M|^2$, então existe uma fatoração $\omega = x'zx''$, com $z \neq 1$, $\alpha x' = \alpha(x'z)$ e $\alpha(zx'') = \alpha x''$.*

Demonstração. O conjunto

$$\{(x, y) \in (A^*)^2 | \omega = xy\}$$

tem no mínimo $1 + |M|^2$ elementos, e portanto existem duas fatorações distintas

$$\omega = x'y' = y''x''$$

tais que

$$\alpha x' = \alpha y'' \quad \text{e} \quad \alpha y' = \alpha x''$$

Podemos assumir $|x'| < |y'|$. Então existe uma palavra $z \neq 1$ tal que $y'' = x'z$ e $y' = zx''$. Então $\omega = x'zx''$ com as propriedades desejadas. □

Lema A.1.2. *Seja $\mu\colon A^* \to \mathcal{M}_n(\mathbb{Q})$ um morfismo multiplicativo da forma*

$$\begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix}$$

e seja $\omega = x'zx''$ de forma que $\mu'x' = \mu'(x'z)$ e $\mu''(zx'') = \mu''x''$, que existem pelo lema A.1.1. Então para todo $n \in \mathbb{N}$,

$$\mu'x'\nu z^n\mu''x'' = n\mu'x'\nu z\mu''x''$$

$$\nu(x'z^nx'') = \nu(x'x'') + n\mu'x'\nu z\mu''x''$$

Demonstração. Sabemos que

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}^n = \begin{pmatrix} a^n & \sum\limits_{k+\ell=n-1} a^k b c^\ell \\ 0 & c^n \end{pmatrix}$$

Temos então que

$$\begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix}^n = \begin{pmatrix} \mu'^n & \sum\limits_{k+\ell=n-1} \mu'^k\ \nu\mu''^\ell \\ 0 & \mu''^n \end{pmatrix} \Rightarrow$$

$$\nu(z^n) = \sum_{k+\ell=n-1} \mu'(z^k)\nu(z)\mu''(z^\ell)$$

Temos portanto que

$$\mu' x' (\nu(z^n))\mu'' x'' = \mu' x' \left(\sum_{k+\ell=n-1} \mu'(z^k)\nu(z)\mu''(z^\ell) \right) \mu'' x'' =$$

$$\sum_{k+\ell=n-1} \mu' x' \mu'(z^k)\nu(z)\mu''(z^\ell)\mu'' x'' =$$

$$\sum_{k+\ell=n-1} \mu'(x' z^k)\nu(z)\mu''(z^\ell x'') = n\mu' x' \mu z \mu'' x''$$

Utilizando o produto $\mu(x' z^n x'') = \mu x' \mu z^n \mu'' x''$, chegamos a

$$\begin{aligned} \nu(x' z^n x'') &= \nu x' \mu''(z^n x'') + \mu' x' \nu(z^n)\mu'' x'' + \mu'(x' z^n)\nu x'' \\ &= \nu x' \mu'' x'' + n\mu' x' \nu z \mu'' x'' + \mu' x' \nu x'' \\ &= \nu(x' x'') + n\mu' \nu z \mu'' x'' \end{aligned}$$

□

Corolário A.1.3. *Seja $\mu\colon A^* \to \mathbb{M}_n(\mathbb{Q})$ em um mmonoide de matrizes triangulares por blocos, com*

$$\mu = \begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix}$$

Assuma que $\mu'(A^*)$ e $\mu''(A^*)$ são finitos, com $H_1 = \left|\mu'(A^*)\right|$ e $H_2 = \mu''(A^*)$, e que $\mu(\omega^*)$ é finito para toda palavra ω. Então

$$\left|\nu(A^*)\right| \leq \sum_{i=0}^{(H_1 H_2)^2 - 1} \left|\nu(A^i)\right|.$$

Demonstração. Pelo lema A.1.1, tomando $\alpha = (\mu', \mu'')$, então cada palavra ω de tamanho maior ou igual a $(H_1 H_2)^2$ tem uma fatoração $\omega = x' z x''$, com $z \neq 1$, e satisfazendo para todo $n \in \mathbb{N}$,

$$\mu' x' \nu z^n \mu'' x''$$

$$\nu(x' z^n x'') = \nu(x' x'') + n\mu' x' \nu z \mu'' x'',$$

como consequência do lema A.1.2.

Assim, como $\mu(z^*)$ é finito, $\nu(x' z^* x'')$ também é finito e devemos ter $\mu' x' \nu z \mu'' x'' = 0$ e $\nu(\omega) = \nu(x' x'')$. Como $\left|x' x''\right| < |\omega|$, segue o resultado. □

Teorema A.1.4 (Mandel e Simon)**.** *Seja $\mu\colon A^* \to \mathcal{M}_n(\mathbb{Q})$ um morfismo de monoide tal que, para todo $\omega \in A^*$, o monoide $\mu(\omega^*)$ é finito. Então existe um número inteiro N efetivamente computável, dependendo apenas de $|A|$ e de n, tal que*

$$\left|\mu(A^*)\right| \leq N$$

Demonstração. Vamos demonstrar o resultado verificando o que ocorre com $\mu(A^*)$ quando este é ou não irredutível.

- $\mu(A^*)$ **é irredutível:** Vamos assumir que o monoide $\mu(A^*)$ é irredutível. e considere qualquer matriz $\mu(\omega) \in \mu(A^*)$. Como $\mu(z^*)$ é finito, existem inteiros $0 \leq i < j$ com $\mu(\omega^i) = \mu(\omega^j)$. Isso implica que os autovalores de ω são 0 ou raízes da unidade. Pelo coroláio A.1.3, segue o resultado do teorema.

- $\mu(A^*)$ **não é irredutível:** Se $\mu(A^*)$ não é irredutível, existe um certo subespaço V de $\mathcal{M}_{1\times n}(\mathbb{Q})$ que é invariante sob $\mu(A^*)$. Considere o espaço suplementar W de V. Em uma base a qual é adapatada à decomposição

$$\mathcal{M}_{1\times n}(\mathbb{Q}) = W \oplus V,$$

o morfismo μ admite a forma

$$\begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix},$$

que é a mesma encontrada no lema A.1.2. Por esse lema, para todo $n \in \mathbb{N}$,

$$\mu' x' \nu z^n \mu'' x'' = n\mu' x' \nu z \mu'' x''$$

$$\nu(x' z^n x'') = \nu(x' x'') + n\mu' x' \nu z \mu'' x''$$

Logo, argumentando por indução na dimensão da representação, o resultado segue do corolário A.1.3.

□

Como podemos ver, a função $(|A|, n) \to N$ cresce extremamente rápido. Apesar disso, existem casos nos quais há um limitante razoável que não dependem de $|A|$, descritos no lema A.1.7. Precisaremos de algumas definições antes:

Definição A.1.5. Dizemos que um conjunto de matrizes $E \subset \mathcal{M}_n(\mathbb{Q})$ é *irredutível* se não existe um subespaço de $\mathcal{M}_{1\times n}(\mathbb{Q})$ próprio invariante por todas as matrizes de E (as matrizes agem à direita em $\mathcal{M}_{1\times n}(\mathbb{Q})$.

Definição A.1.6 (Identidades de Newton)**.** Sejam $x_1, \ldots, x_n$ variáveis, e denote para $k \geq 1$ por $p_k(x_1, \ldots, x_n)$ a k-ésima soma de potências, ou seja:

$$p_k(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^k = x_1^k + \ldots + x_n^k,$$

e para $k \geq 0$, denote $e_k(x_1, \ldots, x_n)$ o *polinômio simétrico elementar* (isto é, a soma de todos os produtos distintos de k variáveis distintas), ou seja:

$$e_0(x_1, \ldots, x_n) = 1,$$

$$e_1(x_1, \ldots, x_n) = \sum_{1\leq j\leq n} x_j = x_1 + x_2 + \cdots + x_n,$$

$$e_2(x_1, \ldots, x_n) = \sum_{1\leq i<j\leq n} x_i x_j,$$

e assim por diante, terminando em

$$e_n(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n$$

Além disso,

$$e_k(x_1, \ldots, x_n) = 0, \quad \text{para } k > n.$$

Em geral, para $k \geq 0$, definimos

$$e_k(x_1, \ldots, x_n) = \sum_{1\leq j_1<j_2<\cdots<j_k\leq n} x_{j_1} \cdots x_{j_k}$$

As identidades de Newton podem ser apresentadas como

$$ke_k(x_1,\ldots,x_n) = \sum_{i=1}^{k}(-1)^{i-1}e_{k-i}(x_1,\ldots,x_n)p_i(x_1,\ldots,x_n),$$

válidas para todo $n \geq 1$ e $n \geq k \geq 1$. Também temos

$$0 = \sum_{i=k-n}^{k}(-1)^{i-1}e_{k-i}(x_1,\ldots,x_n)p_i(x_1,\ldots,x_n),$$

para todo $k > n \geq 1$.

Concretamente, temos abaixo as identidades de Newton para os primeiros valores de k :

$$\begin{aligned}
e_1(x_1,\ldots,x_n) &= p_1(x_1,\ldots,x_n),\\
2e_2(x_1,\ldots,x_n) &= e_1(x_1,\ldots,x_n)p_1(x_1,\ldots,x_n) - p_2(x_1,\ldots,x_n),\\
3e_3(x_1,\ldots,x_n) &= e_2(x_1,\ldots,x_n)p_1(x_1,\ldots,x_n) - e_1(x_1,\ldots,x_n)p_2(x_1,\ldots,x_n) + p_3(x_1,\ldots,x_n).
\end{aligned}$$

A forma e a validade dessas equações não dependem do número n de variáveis, [1] o que torna possível declará-las como identidades no anel de funções simétricas. Temos então

$$\begin{aligned}
e_1 &= p_1,\\
2e_2 &= e_1p_1 - p_2,\\
3e_3 &= e_2p_1 - e_1p_2 + p_3,\\
4e_4 &= e_3p_1 - e_2p_2 + e_1p_3 - p_4,
\end{aligned}$$

e assim por diante; aqui os lados esquerdos da igualdade nunca se anulam. Essas equações permitem expressar recursivamente e_i em termos de p_k; para poder fazer o inverso, pode-se reescrevê-los colocando a soma de k-ésimas potências em evidência. A tabela A.1 mostra as identidades de newton para alguns valores de k.

k	p_k
1	e_1
2	$e_1p_1 - p_2$
3	$e_2p_1 - e_1p_2 + p_3$
4	$e_3p_1 - e_2p_2 + e_1p_3 - p_4$

Tabela A.1: *Identidades de Newton para* $1 \leq k \leq 4$

Em geral, temos que

$$p_k(x_1,\ldots,x_n) = (-1)^{k-1}ke_k(x_1,\ldots,x_n) + \sum_{i=1}^{k-1}(-1)^{k-1+i}e_{k-i}(x_1,\ldots,x_n)p_i(x_1,\ldots,x_n),$$

válido para todo $n \geq 1$ e $n \geq k \geq 1$. Também temos

$$p_k(x_1,\ldots,x_n) = \sum_{i=k-n}^{k-1}(-1)^{k-1+i}e_{k-i}(x_1,\ldots,x_n)p_i(x_1,\ldots,x_n), \quad \forall\ k > n \geq 1$$

[1] embora o ponto em que o lado esquerdo se torna 0, ou seja, após a n-ésima identidade

O polinômio com raízes $x_1, \ldots, x_n$ pode ser expandido como

$$\prod_{i=1}^{n}(x - x_i) = \sum_{k=0}^{n}(-1)^k e_k x^{n-k} \tag{A.1}$$

onde os coeficientes $e_k(x_1, \ldots, x_n)$ são os polinômios simétricos definidos em A.1.6.

Dadas as somas de potências de raízes

$$p_k(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^k,$$

os coeficientes do polinômio com raízes $x_1, \ldots, x_n$ pode ser expresso recursivamente em termos de soma de potências como

$$\begin{aligned}
e_0 &= 1, \\
-e_1 &= -p_1, \\
e_2 &= \frac{1}{2}(e_1 p_1 - p_2), \\
-e_3 &= -\frac{1}{3}(e_2 p_1 - e_1 p_2 + p_3), \\
e_4 &= \frac{1}{4}(e_3 p_1 - e_2 p_2 + e_1 p_3 - p_4), \\
&\vdots
\end{aligned}$$

Quando o polinômio A.1 é o polinômio característico de uma matriz A (em particular quando A é a matriz companheira do polinômio), as raízes x_i são os autovalores da matriz, contados com suas respectivas multiplicidades algébricas. Para qualquer inteiro positivo k, a matriz A^k tem como autovalores as potências x_i^k, e cada autovalor x_i de A tem a mesma multiplicidade que o autovalor x_i^k de A^k. Então os coeficientes do polinômio característico de A^k são dados pelos polinômios simétricos elementares nas potências x_i^k.. Em particular, a soma dos x_i^k,, que é a soma das k-ésimas potências das raízes do polinômio característico de A, é dada pelo seu traço:

$$p_k = \operatorname{tr}(A^k) \tag{A.2}$$

As identidades de Newton agora relacionam os traços das potências A^k com os coeficientes do polinômio característico de A. Usando-as em sentido inverso para expressar os polinômios simétricos elementares em termos de somas de potência, elas podem ser usadas para encontrar o polinômio característico computando apenas as potências A^k e seus traços. Para mais detalhes, consulte [11] e [21].

Lema A.1.7. *Seja $M \subset \mathcal{M}_n(\mathbb{Q})$ um monoide irredutível de matrizes tal que todo autovalor não-nulo de matrizes em M são raízes da unidade. Então*

$$|M| \leq (2n+1)^{n^2}$$

Demonstração. Considere $m \in M$. Os autovalores não-nulos de m são raízes da unidade, e portanto são inteiros algébricos sobre $\mathbb{Z}$. Como o traço de m é a soma de seus autovalores, então $\operatorname{tr}(m)$ é um inteiro algébrico. Como $\operatorname{tr}(m) \in \mathbb{Q}$ e $\mathbb{Z}$ é integralmente fechado, então $\operatorname{tr}(m) \in \mathbb{Z}$. A norma de cada autovalor é 0 ou 1. Então $\left|\operatorname{tr}(m)\right| \leq n$. Isso mostra que $\operatorname{tr}(m)$ assume no máximo $2n+1$ valores distintos para $m \in M$.

Seja $m_1, \ldots m_k \in M$ uma base do subespaço N de $\mathcal{M}_n(\mathbb{Q})$ gerado por M. Defina uma relação

de equivalência em M tomando

$$m \sim \mathfrak{m} \Leftrightarrow \mathrm{tr}(mm_i) = \mathrm{tr}(\mathfrak{m}m_i) \quad \forall i \in \{1, \ldots, k\}$$

De fato, temos que:

- $\sim$ é reflexiva, pois trivialmente

$$m \sim m \Leftrightarrow \mathrm{tr}(mm_i) = \mathrm{tr}(mm_i) \quad \forall i \in \{1, \ldots, k\}$$

- $\sim$ é simétrica, pois

$$m \sim \mathfrak{m} \Leftrightarrow \mathrm{tr}(mm_i) = \mathrm{tr}(\mathfrak{m}m_i) \Leftrightarrow \mathrm{tr}(\mathfrak{m}m_i) = \mathrm{tr}(mm_i) \Leftrightarrow \mathfrak{m} \sim m$$

- $\sim$ é transitiva, pois se

$$m \sim \mathfrak{m} \Leftrightarrow \mathrm{tr}(mm_i) = \mathrm{tr}(\mathfrak{m}m_i)$$

e

$$\mathfrak{m} \sim \ \Leftrightarrow \mathrm{tr}(\mathfrak{m}m_i) = \mathrm{tr}(m_i)$$

, então

$$\mathrm{tr}(mm_i) = \mathrm{tr}(\mathfrak{m}m_i) = \mathrm{tr}(m_i) \Leftrightarrow m \sim$$

A quantidade de classes de equivalência dessa relação é no máximo $(2n+1)^k$. Então, é suficiente provar que $m \sim m'$ implica $m = m'$.

Sejam $m, \mathfrak{m} \in M$ tais que $m \sim \mathfrak{m}$. Tome $q = m - \mathfrak{m}$, e suponha por absurdo que $q \neq 0$. Então existe um vetor $v \in \mathcal{M}_{1\times n}(\mathbb{Q})$ tal que $vq \neq 0$. Considere o subespaço

$$vqN = \left\{ \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix} \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1n} \\ n_{21} & n_{22} & \cdots & n_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ n_{n1} & n_{n2} & \cdots & n_{nn} \end{pmatrix} : \right.$$

$$\left. \begin{pmatrix} n_{11} & n_{12} & \cdots & n_{1n} \\ n_{21} & n_{22} & \cdots & n_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ n_{n1} & n_{n2} & \cdots & n_{nn} \end{pmatrix} \in N \right\}$$

O subespaço vqN de $\mathcal{M}_{1\times n}(\mathbb{Q})$ não é o espaço nulo, pois $vq \neq 0$. Vamos mostrar que ele é invariante sobre M. Sendo vqN invariante sobre M e M irredutível, temos que $vqN = \mathcal{M}_{1\times n}(\mathbb{Q})$. Consequentemente, existe um $p \in \mathbb{N}$ tal que $vqp = v$. Isso mostra que qp tem autovalor 1.

Agora, para $j \geq 1$, temos que

$$\mathrm{tr}((qp)^j) = \mathrm{tr}((qp)(qp^{j-1})) = 0,$$

pois $p(qp)^{j-1}$ é uma combinação linear de matrizes $m_1, \ldots, m_k$, e por hipótese, $\mathrm{tr}(qr) = 0$ para $r \in M$. Basta mostrar agora que isso implica que todos os autovalores de qp se anulam. Para isso, iremos utilizar as identidades de Newton, definidas em A.1.6. Pela equação A.2, temos que $p_k = 0$, o que implica que todos os autovalores de pq se anulam.

Temos portanto uma contradição. □

Definição A.1.8. Dizemos que um elemento s de um semigrupo S é de torsão se s gera um

subsemigrupo finito de S; equivalentemente, $s^k = s^\ell$ para algum $1 \leq k < \ell$. Dizemos que S é um semigrupo de torsão se cada elemento de S é de torsão.

Corolário A.1.9 (McNaughton-Zalcstein)**.** *Todo semigrupo de torsão finitamente gerado de matrizes quadradas sobre $\mathbb{Q}$ é finito.*

Demonstração. Veja [20]. □

Dizemos que o *raio* é um subconjunto de A^* da forma uv^*w, com $u, v, w \in A^*$.

Corolário A.1.10. *Seja $S \in \mathbb{Q}\langle\langle A \rangle\rangle$ uma série racional tal que, para todo raio R, o conjunto*

$$\{(S, \omega) | \omega \in R\}$$

é finito. Então o conjunto de coeficientes de S é finito.

Demonstração. Seja (λ, μ, γ) uma representação linear minimal de S, como definido em 3.3.29. Pelo corolário 3.3.31, existem polinômios $P_1, \ldots, P_n, Q_1, \ldots, Q_n$ tais que para todas as palavras ω,

$$\mu(\omega) = ((S, P_i \omega Q_j))_{1 \leq i,j \leq n}$$

Por hipótese, o conjunto

$$\{(S, u\omega^m v) | m \in \mathbb{N}\}$$

é finito para todas as palavras u, v, ω. O mesmo ocorre para o conjunto

$$\{(S, P\omega^m Q | m \in \mathbb{N}\},$$

onde P e Q são polinômios. Isso mostra que $\mu(\omega^*)$ é finito para qualquer palavra ω. Pelo corolário A.1.10, o monoide $\mu(A^*)$ é finito, e em particular

$$\{(S, \omega) | \omega \in A^*\}$$

é finito, uma vez que $(S, \omega) = \lambda\mu(\omega)\gamma$. □

A.2 Crescimento polinomial de séries ℤ-racionais

Vamos abordar agora questões relacionadas ao crescimento de séries racionais sobre $\mathbb{Z}$.

Definição A.2.1. Dizemos que uma série $S \in \mathbb{Z}\langle\langle A \rangle\rangle$ tem *crescimento polinomial* ou é dita *polinomialmente limitada* se existe um número rela $q \geq 0$ e um número real C tal que

$$\big|(S, \omega)\big| \leq C|\omega|^q$$

para todas as palavras não-vazias ω. O menor q que satisfaz a desigualdade é chamado *grau de crescimento* de S.

Observe que as séries com grau de crescimento 0 são exatamente as séries com imagem finita.

Na sequência, iremos considerar morfismos $\mu\colon A^* \to \mathcal{M}_n(\mathbb{Q})$ que possui uma forma triangular em blocos

$$\mu = \begin{pmatrix} \mu_0 & \nu_1 & \star & \cdots & \star \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \star \\ \vdots & \ddots & \ddots & \ddots & \nu_q \\ 0 & \cdots & \cdots & \cdots & \mu_q \end{pmatrix}$$

onde cada μ_i é quadrada e portanto é em si um morfismo.

Teorema A.2.2. *Seja $S \in \mathbb{Z}\langle\langle A\rangle\rangle$ uma série racional e tome (λ, μ, γ) uma representação linear minimal de S. Então, se S possui crescimento polinomial,*

$$\left|\{tr(\mu(\omega))|\omega \in A^*\}\right| < \infty$$

Demonstração. Suponha que S tem crescimento polinomial. Então, pelo corolário 3.3.31, existem números reais C, q tais que para todo i, j,

$$\left|(\mu(\omega))_{i,j}\right| \leq C|\omega|^q \quad \forall \omega \neq \varepsilon, \omega \in A^*$$

Segue que, para qualquer $r \in \mathbb{N}^*$, temos que

$$\left|(\mu(\omega^r))_{i,j}\right| \leq Cr^q|\omega|^q \quad \forall \omega \neq \varepsilon, \omega \in A^*$$

Logo, para todo autovalor ρ de $\mu(\omega)$, temos

$$|\rho|^r \leq Dr^q,$$

para alguma constante D. Daí, $|\rho| \leq 1$. Isso implica que

$$-n \leq \mathrm{tr}(\mu(\omega)) \leq n,$$

onde n é a dimensão de μ. Como S é $\mathbb{Z}$-racional, existe uma representação linear minimal com coeficientes em $\mathbb{Z}$, de acordo com a proposição 3.3.32. Essa representação é semelhante a (λ, μ, γ) pelo teorema 3.3.34, e consequentemente, o traço de qualquer matriz $\mu(\omega)$ é inteiro. Logo, para cada $\mathrm{tr}(\mu(\omega))$, temos que

$$\mathrm{tr}(\mu(\omega)) \in \{-n, \ldots, n\}.$$

□

Agora que já vimos na demonstração do teorema de Mandel e Simon A.1.4 a importância de utilizar um morfismo multiplicativo $\mu\colon A^* \to \mathcal{M}_n(\mathbb{Q})$ da forma

$$\begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix},$$

tal como apresentado no lema A.1.2, vamos agora demonstrar o lema 3.3.37:

Lema A.2.3. *Seja K um semianel comutativo. Tome*

$$\mu = \begin{pmatrix} \mu' & \nu \\ 0 & \mu'' \end{pmatrix}$$

um morfismo $A^* \to \mathcal{M}_n(K)$, onde μ' e μ'' são morfismos. Toda série reconhecida por μ é uma combinação linear de séries reconhecidas por μ' ou μ'' e de séries da forma $S'aS''$, onde S' é reconhecida por μ', $a \in A$ e S'' é reonhecida por μ''.

Demonstração. Uma série reconhecida por μ é uma combinação linear de séries da forma

$$\sum_{\omega \in A^*} (\mu(\omega))_{i,j}\omega$$

com $0 \leq i, j \leq n$. É suficiente mostrar que, quando i, j são coordenadas equivalentes a ν, a série apresentada acima é uma combinação linear de séries da forma $S'aS''$.

Isso por sua vez é uma consequência da fórmula

$$\nu(\omega) = \sum_{\omega=xay} \mu'(x)\nu(a)\mu''(y).$$

□

Referências Bibliográficas

[1] Atiyah, M. F., and MacDonald, I. G. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969. 5

[2] Bourbaki, N. *Groupes et Algébres de Lie.* Hermann, Paris, 1968. 99

[3] Conway, J. H. Regular algebras and finite machines. *Chapman & Hall* (1971). 50

[4] Duchamp, G., and Krob, D. The free partially commutative lie algebra: Bases and ranks. *Advances in Mathematics 95* (1992), 92–126. 98, 99

[5] Duchamp, G., and Krob, D. The lower central series of the free partially commutative group. *Semigroup Forum 45* (1995), 385–394. 4, 95, 96, 99

[6] Eggan, L. C. Transition graphs and the star-height of regular events. *Michigan Mathematical Journal 10* (1963), 385–397. 61

[7] Fox, R. Free diferential Calculus. I: Derivations in the Free Group Rings,. *Annals of Mathematics, Secondi Series 57*, 3 (1953), 547–560. 2, 3, 79, 92, 100, 107

[8] Goodearl, K. R., and Warfield, R. B. *An Introduction to Noncommutative Noetherian Rings*, vol. 16. London Mathematical Society, Cambridge University Press, 1989. 5

[9] Green, J. A. On the definition of a family of automata. *Information and Control* (1961). 33

[10] K. T. Chen, R. H. F., and Lyndon, R. C. Free differential calculus, iv. the quotient groups of the lower central series. *Annals of Mathematics, Secondi Series 68* (1958). 89, 90

[11] Kalman, D. A Matrix Proof of Newton's Identities. *Mathematics Magazine 73*, 4 (2000), 313–315. 113

[12] Krob, D. Expressions rationnelles sur un anneau. *Topics in invariant theory (Paris, 1989/1990), Lecture Notes in Mathematics, 1478* (1991), 215–243. 50

[13] Krob, D., and Lalonde, P. Partially commutative lyndon words. 10. 95

[14] Lang, S. *Algebra*, segunda ed. Addison-Wesley, 1984. 57

[15] Lequain, Y., and Garcia, A. *Elementos de Álgebra.* Projeto Euclides. Instituto de Matematica Pura e Aplicada, 2005. 5, 57

[16] Lothaire, M. *Combinatorics on words.* Encyclopedia of Mathematics and Its Applications, Cambridge University Press, 1997. 18, 19

[17] Magnus, W., Karrass, A., and Solitar, D. *Combinatorial Group Theory.* Nova York. Wiley, 1966. 81

[18] Mandel, A., and Simon, I. On finite semigroups of matrices. *Theoretical Computer Science 5*, 2 (1977), 101 – 111. 3, 109

[19] McNaughton, R., and Yamada, H. Regular expressions and state graphs for automata. *Trans. IRS EC-9* (1960). 32

[20] McNaughton, R., and Zalcstein, Y. The burnside problem for semigroups. *Journal of Algebra 34* (1975), 292–299. 115

[21] Mead, D. G. Netwon's Identities. *The American Mathematical Monthly 99*, 8 (1992), 749–751. 113

[22] Mikhailov, R., and Passi, I. B. S. *Lower Central and Dimension Series of Groups*. Lecture Notes in Mathematics, Springer, 1952. 90

[23] Năstăsescu, C., and Oystaeyen, F. V. *Methods of Graded Rings*. Springer-Verlag, 2004. 21

[24] P. Feofiloff, Y. Kohayakawa, Y. W. Uma introdução sucinta à teoria dos grafos. 61

[25] Passman, D. S. What is a Group Ring? *American Mathematical Monthly 83* (1976). 77, 78, 79

[26] Passman, D. S. *The algebraic structure of group rings*. John Wiley & Sons, 1977. 77

[27] Reutenauer, C., and Berstel, J. *Rational Series and Their Languages*. Springer-Verlag, 2008. 34, 43, 109

[28] Reutenauer, C., and Berstel, J. *Noncommutative Rational Series With Applications*. Encyclopedia of Mathematics and its Applications, Cambridge University Press., 2010. 2, 5, 19, 61, 66

[29] Sakarovitch, J. *Elements of Automata Theory*. Cambridge University Press, 2009. 32, 34

[30] Tengan, E., and Borges, H. *Álgebra Comutativa em quatro movimentos*. Projeto Euclides, IMPA, 2015. 21

[31] Tutte, W. T. *Graph Theory*, vol. 21. Cambridge University Press, 1984. 63

[32] Vieira, N. J. *Introdução aos fundamentos da computação: linguagens e máquinas*. Cengage Learning, 2015. 47

[33] Wilf, H. S. *Generatingfunctionology*, primeira ed. Academic Press, Jan. 1990. 60

[34] Witt, E. Treue darstellung liescher ringe. *Journal für die reine und angewandte Mathematik 177* (1937). 90, 92

Índice Remissivo