Uma demonstração do Teorema de Thue-Siegel-Dyson-Roth

Luis Fernando Ragognette

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Matemática
Orientador: Prof. Dr. Paulo Agozzini Martin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro do CNPq

São Paulo, março de 2012

Uma demonstração do Teorema de Thue-Siegel-Dyson-Roth

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 11/05/2012. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Paulo Agozzini Martin(Presidente) - IME-USP
- Prof. Dr. Yoshiharu Kohayakawa - IME-USP
- Prof. Dr. Fábio Maia Bertato - CLE-UNICAMP

Agradecimentos

Aos meus pais, Paulo Ragognette e Leonilda Ginete Machiti Ragognette pela paciência, dedicação e amor que recebo e, especialmente, por se dedicarem tanto para que eu alcance meus objetivos.

A minha namorada Mariana Robertti Ambrosio por me aturar, entender e acompanhar durante todo este período. Ela é talvez a maior vítima do meu mau humor, mas, continua, sempre alegre, a me acompanhar.

Ao Paulo Agozzini Martin, meu orientador, pela ajuda e paciência durante todas as exposições e durante a revisão deste texto. Ao Fábio Maia Bertato, membro da Comissão Julgadora, por suas críticas e sugestões que ajudaram a melhorar este trabalho. Ao Yoshiharu Kohayakawa, membro da Comissão Julgadora, professor que teve participação importantíssima na minha formação como matemático.

Aos professores que me influenciaram direta e indiretamente nesta caminhada, em especial, gostaria de agradecer a Roseli Fernandez, ao Paulo Domingos Cordaro e ao Manuel Valentim de Pera Garcia pelas aulas, pelos conselhos e pelo carinho.

Aos meus amigos do IME-USP pela companhia nas aulas, nos seminários, na sala do café e nos corredores do IME. Em especial, agradeço ao Gabriel Cueva Candido Soares de Araújo e ao Max Reinhold Jahnke por serem os amigos com quem mais dividi minhas preocupações e por nunca me deixarem desanimar.

Agradeço ainda ao CNPq pelo auxílio financeiro.

Resumo

Neste trabalho estudamos o célebre Teorema de Klaus F. Roth para aproximações diofantinas, também conhecido como Teorema de Thue-Siegel-Roth. Nossos objetivos consistem em fazer um estudo abrangente da evolução do problema, que se iniciou com um resultado de Liouville em 1844, e chegar à completa compreensão das ideias e das técnicas utilizadas na demonstração do Teorema de Roth.
Palavras-chave: Teorema de Thue-Siegel-Dyson-Roth, aproximações diofantinas, números algébricos.

Abstract

In this work we study the celebrated Klaus F. Roth's Theorem in Diophantine approximations, also known as the Thue-Siegel-Roth Theorem. Our goals are to make a comprehensive study of the evolution of the problem that started with a result of Liouville in 1844 and achieve full understanding of ideas and techniques used in the proof of the Roth's Theorem.
Keywords: Thue-Siegel-Dyson-Roth Theorem, diophantine approximation, algebraic numbers.

Sumário

Introdução

A história do Teorema de Roth começa com o seguinte resultado devido a Joseph Liouville(1809-1882): se α é um irracional algébrico de grau n então existe uma constante positiva $c(\alpha)$ tal que

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^n},$$

para todos p, q inteiros com q positivo.

Uma consequência deste resultado é que, se a desigualdade

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu} \tag{1}$$

possui um número infinito de soluções, então $\mu \leq n$.

Em 1909, Axel Thue(1863-1922) obteve um resultado mais forte do que o obtido por Liouville e assim conseguiu provar a finitude de soluções para um tipo importante de equações diofantinas, que ficaram conhecidas como equações de Thue. Em seu artigo [**5**], Thue provou que, se a equação (1) possui um número infinito de soluções, então $\mu \leq n/2 + 1$.

Em 1921, Carl L. Siegel(1896-1981) [**4**] conseguiu provar que a existência de infinitas de soluções da equação (1) implica que

$$\mu \leq \min_{s \in \{1,\ldots,n-1\}} \left\{ \frac{n}{s+1} + s \right\} < 2\sqrt{n}.$$

Siegel ainda conjecturou que a infinitude de soluções deveria implicar que $\mu \leq 2$. Esta conjectura contrasta com um famoso resultado Johann P. G. L. Dirichlet(1805-1859) que diz que, dado um número irracional, existem infinitos racionais p/q tais que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Em 1947 Freeman J. Dyson(1923) [**2**] conseguiu melhorar o resultado obtido por Siegel e provar que $\mu \leq \sqrt{2n}$. Este Teorema, aparentemente, também foi provado independentemente por Alexander O. Gelfond(1906-1968). Gelfond publicou seu resultado em 1952 em seu livro [**8**].

Na conclusão de seu artigo, Dyson, comenta que o surgimento de uma raiz quadrada está diretamente relacionado com o uso de polinômios em duas variáveis e que acredita que se seus resultados pudessem ser generalizados num contexto de várias variáveis, seria possível então provar a conjectura de Siegel.

Em 1949, Kurt Mahler(1903-1988) [**3**] publicou um artigo sobre o Teorema de Dyson. Mahler conseguiu simplificar algumas ideias presentes no artigo de Dyson. Mais precisamente, em seu artigo, Dyson apresenta um único Lema e a demonstração do Teorema e o resultado provado por Mahler é uma versão simplificada do Lema de Dyson.

Finalmente, em 1955, Klaus F. Roth(1925) [**1**] investigou o comportamento de uma função chamada índice para polinômios em várias variáveis e com isso provou a conjectura de Siegel, isto é, Roth provou que $\mu \leq 2$.

A demonstração do Teorema de Roth é de uma dificuldade técnica muito grande e sua realização fez com que Roth ganhasse a medalha Fields em 1958.

No primeiro capítulo temos uma breve introdução ao problema das aproximações diofantinas e os principais resultados provados são os Teoremas de Dirichlet e de Liouville.

No segundo capítulo apresentamos uma demonstração do Teorema de Thue, comentamos a demonstração do Teorema de Siegel e provamos o finitude de soluções para as equações de Thue.

Iniciamos o terceiro capítulo com uma demonstração de um resultado de Mahler e usamos este resultado para obter o Teorema de Dyson.

Finalmente, no quarto capítulo, provamos o Teorema de Roth. Finalizamos esta dissertação no quinto capítulo onde apresentamos algumas aplicações destes Teoremas e comentamos algumas generalizações do Teorema de Roth.

CAPíTULO 1

Alguns resultados sobre aproximações diofantinas

1. Aproximações diofantinas

A teoria das aproximações diofantinas é uma área da teoria dos números que estuda a aproximação de números irracionais por números racionais.

É bem conhecido que os números racionais são densos dentro do conjunto dos números reais. Nosso interesse, quando pensamos em aproximações diofantinas, é exercer um controle no denominador do número racional. Por exemplo, suponhamos que queremos aproximar um número irracional α por números racionais na forma p/q tais que a diferença entre α e p/q seja menor que 10^{-100}. Se escolhermos q da ordem de 10^{100} é fácil conseguir um p tal que $|\alpha - p/q| < 10^{-100}$. Mas, se escolhermos q da ordem de 10^{60} ou 10^{30}, será que ainda conseguimos uma diferença menor que 10^{-100}? Nossa intuição nos diz que ao aproximar um irracional α não é possível conseguir uma aproximação da forma p/q tal que tanto o denominador q quanto o erro $|\alpha - p/q|$ sejam "pequenos". O quão preciso conseguimos deixar essa afirmação?

Por exemplo, consideremos as seguintes aproximações para o número π: 31415/10000 e 355/113. A primeira aproximação tem o erro da ordem de 10^{-5} e a segunda tem um erro da ordem de 10^{-7}. No segundo exemplo conseguimos com um denominador de ordem 10^2 e um erro de ordem 10^{-7}, portanto, podemos considerar 335/113 uma aproximação melhor que 31415/10000.

Começaremos apresentando alguns resultados simples que servirão para ilustrar as dificuldades de obtermos melhores aproximações. O primeiro resultado é bastante intuitivo. Se quisermos aproximar um irracional por um número inteiro então existe um único inteiro que dista de nosso irracional menos que 1/2.

PROPOSIÇÃO 1. *Para qualquer número irracional α existe um único inteiro p tal que*

$$|\alpha - p| < \frac{1}{2}.$$

DEMONSTRAÇÃO. Seja $\lfloor \alpha \rfloor$ o maior inteiro menor que α. Temos que

$$0 < \alpha - \lfloor \alpha \rfloor < 1.$$

Portanto, se $0 < \alpha - \lfloor \alpha \rfloor < 1/2$ fazemos $p = \lfloor \alpha \rfloor$. Se $1/2 < \alpha - \lfloor \alpha \rfloor < 1$ temos que

$$-1/2 < \alpha - \lfloor \alpha \rfloor - 1 < 0,$$

e fazemos $p = \lfloor \alpha \rfloor + 1$.

Suponhamos agora que m e n são inteiros tais que $|\alpha - m| < 1/2$ e $|\alpha - n| < 1/2$ temos pela desigualdade triangular que

$$|m - n| = |m - \alpha + \alpha - n| \leq |m - \alpha| + |\alpha - n| < 1.$$

Como m e n são inteiros temos que $m = n$. □

Nossa próxima proposição é uma versão para racionais da proposição 1.

PROPOSIÇÃO 2. *Sejam α um número irracional e q um inteiro positivo. Então existe um número inteiro p tal que*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q}.$$

DEMONSTRAÇÃO. Como $q\alpha$ é um número irracional podemos aplicar a proposição 1 que nos diz que existe um único inteiro p tal que

$$|q\alpha - p| < \frac{1}{2}.$$

Disto é claro que

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q}.$$

□

Notação: Durante todo o texto, a menos de menção contraria, p será um número inteiro e q será um inteiro positivo.

Antes de apresentarmos um resultado mais forte que a proposição 2 vamos demonstrar um resultado bastante intuitivo e que usaremos várias vezes durante este texto sem mesmo mencioná-lo.

PROPOSIÇÃO 3. *Seja α um número irracional e suponhamos que*

$$\left| \alpha - \frac{p}{q} \right| < 1 \tag{2}$$

possua infinitas soluções p, q então, dado um inteiro positivo M, temos que o número de soluções tais que $q < M$ é finito. Portanto, a infinitude de soluções implica que sempre podemos escolher soluções p/q com q tão grande quanto desejamos.

DEMONSTRAÇÃO. Fixado q temos que

$$\left|\frac{p}{q}\right| \leq |\alpha| + \left|\alpha - \frac{p}{q}\right| < |\alpha| + 1.$$

Logo

$$|p| \leq q(|\alpha| + 1).$$

Disto concluímos que, dado um denominador q, existe uma quantidade finita de numeradores p tais que p/q é solução. Logo, existe apenas um número finito de soluções p/q tais que $q < M$. □

Uma consequência da proposição anterior é que se α é um irracional dado e queremos aproximar α por um racional na forma p/q de tal forma que α e p/q distem no máximo δ, então, a escolha de um δ "pequeno" implica que q precisa ser um número "grande". O que é expresso de forma mais precisa na seguinte proposição:

PROPOSIÇÃO 4. *Dados um número irracional α e um inteiro positivo n, existe $\delta > 0$ tal que se p/q satisfaz*

$$\left|\alpha - \frac{p}{q}\right| < \delta$$

então $n < q$.

DEMONSTRAÇÃO. Pela proposição anterior temos que o número de soluções de (2) tais que $q \leq n$ é finito. Definimos

$$\delta = \frac{1}{2} \min\left\{\left|\alpha - \frac{p}{q}\right|\right\}$$

onde o mínimo é calculado sobre todas as frações tais que $q \leq n$. Se não existir p/q nestas condições definimos $\delta = 1$. □

2. Teorema da aproximação de Dirichlet

Vamos, agora, provar o Teorema da aproximação de Dirichlet. Este nos diz que, para um irracional dado, existem infinitas aproximações tais que o erro é menor que o inverso do quadrado do denominador. A demonstração do Teorema da aproximação de Dirichlet é uma simples aplicação do princípio da casa dos pombos.

2.1. Teorema de Dirichlet.

TEOREMA 1 (Teorema de Dirichlet, 1842). *Seja α um número irracional. Existem infinitos números racionais da forma p/q, tais que $(p, q) = 1$ e*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^2}. \tag{3}$$

DEMONSTRAÇÃO. Dado α um número irracional e um inteiro $Q \geq 1$, consideremos os seguintes números

$$s_k = k\alpha - \lfloor k\alpha \rfloor,$$

onde $k \in \{0, 1, \ldots, Q\}$.

É claro que $0 \leq s_k < 1$. Dividindo o intervalo $[0, 1]$ em Q partes, isto é, $[0, 1] = [0, 1/Q] \cup [1/Q, 2/Q] \cup \cdots \cup [(Q-1)/Q, 1]$, temos, pelo princípio da casa dos pombos, que existem inteiros $0 \leq k_1 < k_2 \leq Q$ tais que s_{k_1} e s_{k_2} estão em uma mesma parte $[\ell/Q, (\ell+1)/Q]$.

Desta forma é claro que $|s_{k_2} - s_{k_1}| \leq 1/Q$ e, consequentemente,

$$\left|k_2\alpha - \lfloor k_2\alpha \rfloor - k_1\alpha + \lfloor k_1\alpha \rfloor\right| \leq \frac{1}{Q}$$

$$\left|(k_2 - k_1)\alpha - (\lfloor k_2\alpha \rfloor - \lfloor k_1\alpha \rfloor)\right| \leq \frac{1}{Q}.$$

Sejam $q = k_2 - k_1$ e $p = (\lfloor k_2\alpha \rfloor - \lfloor k_1\alpha \rfloor)$. Então:

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{qQ} \leq \frac{1}{q^2}.$$

Suponhamos que o número de racionais satisfazendo (3) seja finito. Isto implica que existe um inteiro positivo n tal que

$$\frac{1}{n} < \left|\alpha - \frac{p}{q}\right|,$$

para todos p, q inteiros com $q \geq 1$ que verificam (3). Repetimos o processo acima com $Q = n$ e obtemos

$$\left|\alpha - \frac{p_n}{q_n}\right| \leq \frac{1}{nq_n} \leq \frac{1}{n}.$$

Consequentemente o número de racionais satisfazendo (3) é infinito. Por fim, para provar que p, q podem ser escolhidos primos entre si observemos primeiro que se p_0, q_0 são tais que p_0/q_0 é solução de (3) e $p_0/q_0 = p/q$, onde p e q são primos entre si, então

$$\left|\alpha - \frac{p}{q}\right| = \left|\alpha - \frac{p_0}{q_0}\right| < \frac{1}{q_0^2} \leq \frac{1}{q^2}.$$

Logo, toda solução de (3) está associada a uma fração reduzida que também é solução. Portanto, basta demonstrar que o número de soluções associadas a uma fração reduzida é finito.

Por absurdo, suponhamos que infinitas soluções de (3) estão associadas a fração reduzida p/q. Logo, existe uma sequência p_n/q_n, com

q_n crescente, tal que $p_n/q_n = p/q$ e assim

$$\left|\alpha - \frac{p}{q}\right| = \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}$$

o que contraria a proposição 4.

□

A próxima proposição nos diz que o Teorema de Dirichlet não pode ser estendido para os racionais.

PROPOSIÇÃO 5. *Sejam α um número real, A um número real positivo e $k > 1$. Denotamos por $S(\alpha)$ o número de frações p/q com p e q primos entre si, tais que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{A}{q^k}.$$

Temos que, se $S(\alpha)$ é infinito, então α é irracional.

DEMONSTRAÇÃO. Suponhamos que $\alpha = a/b$ é racional. Para as soluções vale que

$$\left|\frac{a}{b} - \frac{p}{q}\right| < \frac{A}{q^k},$$

donde concluímos que

$$|aq - pb| < \frac{bA}{q^{k-1}}.$$

Não podemos ter infinitas soluções, pois, o número de soluções tais que $q^{k-1} < bA$ é finito e $|aq - pb|$ é um número inteiro e $|aq - pb|$ só é nulo quando $p/q = a/b$, o que só ocorre uma vez dado que p e q são primos entre si. □

2.2. Teorema de Hurwitz. Em 1891, Adolf Hurwitz(1859-1919) obteve um resultado mais preciso que o de Dirichlet. Hurwitz demonstrou o seguinte Teorema:

TEOREMA 2 (Hurwitz, 1891). *Seja α um número irracional. Então:*

(a) *Existem infinitas soluções racionais p/q tais que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

(b) *Dado um número $C > \sqrt{5}$ existe um α irracional tal que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Cq^2}$$

tem apenas um número finito de soluções.

Não vamos provar o item (a), pois para tanto seria interessante nos desviar um pouco e falar sobre séries de Farey. Uma demonstração deste item pode ser encontrada no livro de LeVeque [**12**].

Provaremos o item (b). Usaremos o seguinte Lema:

LEMA 1. *Suponha que α é um número irracional algébrico de grau 2 que satisfaz $f(x) = ax^2 + bx + c$, onde a, b e c são inteiros e $f(x)$ não é o polinômio nulo. Seja D o discriminante $D = b^2 - 4ac$. Se C é um número real tal que $C > \sqrt{D}$, então a desigualdade*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Cq^2} \tag{4}$$

tem apenas um número finito de soluções.

DEMONSTRAÇÃO. Sejam p e q dois inteiros primos entre si e tais que q é positivo. Temos que $f(p/q) \neq 0$. Pois as únicas raízes de f são α e β, o conjugado de α. Então:

$$\left|f\Big(\frac{p}{q}\Big)\right| \geq \frac{1}{q^2}.$$

Suponha que p/q satisfaz

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Cq^2},$$

para algum $C > \sqrt{D}$.

Como $b = -a(\alpha + \beta)$ e $ac = a^2\alpha\beta$ temos que

$$D = a^2(\alpha + \beta)^2 - 4a^2\alpha\beta = a^2(\alpha - \beta)^2.$$

Donde temos

$$\begin{aligned}
\frac{1}{q^2} &\leq \left|f\Big(\frac{p}{q}\Big)\right| \\
&= \left|\alpha - \frac{p}{q}\right|\left|a\Big(\beta - \frac{p}{q}\Big)\right| \\
&< \frac{1}{Cq^2}\left|a\Big(\beta - \alpha + \alpha - \frac{p}{q}\Big)\right| \\
&\leq \frac{1}{Cq^2}\Big(|a(\beta - \alpha)| + |a|\left|\alpha - \frac{p}{q}\right|\Big) \\
&\leq \frac{1}{Cq^2}\Big(\sqrt{D} + \frac{|a|}{Cq^2}\Big).
\end{aligned}$$

Desta desigualdade conseguimos

$$(C - \sqrt{D})Cq^2 < |a|. \tag{5}$$

Como, por hipótese, $C > \sqrt{D}$ temos que (5) não pode ser válida para q suficientemente grande. Logo, a desigualdade (4) possui um número finito de soluções.

□

Demonstração do item (b) do Teorema 2. Seja $\alpha = \frac{1}{2}(\sqrt{5}-1)$. Então α satisfaz o polinômio $f(x) = x^2 + x - 1$ cujo discriminante D é 5. Logo, pelo Lema 1, temos que se $C > \sqrt{5}$ a desigualdade

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{Cq^2}$$

possui um número finito de soluções. □

3. Teorema de Liouville

O Teorema de Dirichlet nos diz que um número irracional α sempre pode ser aproximado por números racionais p/q tais que o erro desta aproximação é menor que $1/q^2$. Até aqui os resultados que estudamos nos garantem a existência de infinitas aproximações satisfazendo uma certa desigualdade. Estudaremos, a partir de agora, certas desigualdades que só podem ter um número finito de soluções. Para isto, vamos restringir o nosso estudo ao caso em que α é um número algébrico; mais precisamente: dado α algébrico, estudaremos qual é o menor μ tal que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu}$$

tem um número finito de soluções.

3.1. Teorema de Liouville. Nosso primeiro resultado nesta direção foi provado por Liouville em 1844 e ele nos diz que $\mu \leq n$, onde n é o grau do número algébrico.

Teorema 3 (Liouville,1844). *Seja α um número irracional algébrico de grau n, então existe $c(\alpha) > 0$ tal que*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^n},$$

para todo p, q.

Demonstração. Como α é um número algébrico de grau n, existe um polinômio f de grau n e coeficientes inteiros tal que

$$f(\alpha) = a_n\alpha^n + \cdots + a_1\alpha + a_0 = 0.$$

Como f' é contínua, existe M tal que $|f'(x)| < M$ para todo $x \in [\alpha - 1, \alpha + 1]$.

Consideremos as aproximações p/q tais que

$$\left|\alpha - \frac{p}{q}\right| < 1.$$

Como f é irredutível temos que $|f(p/q)| \neq 0$.

Portanto,

$$\left|f\left(\frac{p}{q}\right)\right| = \frac{|a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_0 q^n|}{q^n} \geq \frac{1}{q^n}.$$

Podemos aplicar o Teorema do valor médio para obter

$$f\left(\frac{p}{q}\right) = f\left(\frac{p}{q}\right) - f(\alpha) = f'(a)\left(\frac{p}{q} - \alpha\right)$$

para algum $a \in [\alpha - 1, \alpha + 1]$.

Logo, temos que

$$\left|\frac{p}{q} - \alpha\right| = \frac{|f(p/q)|}{|f'(a)|} \geq \frac{1}{|f'(a)|q^n} > \frac{K}{q^n},$$

onde $K = 1/M$.

Seja $c(\alpha) = \min\{K, 1\}$ temos que

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^n},$$

para todo p, q.

□

Como corolário do Teorema de Liouville temos o seguinte resultado:

Teorema 4. *Seja α um número algébrico de grau $n \geq 2$. Dado $\epsilon > 0$. Então existe apenas uma quantidade finita de racionais $p/q \in \mathbb{Q}$ tais que*

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^{n+\epsilon}}.$$

Os Teoremas de Thue, Siegel, Dyson e Roth são generalizações deste último resultado. É importante deixar claro que, fixado k, temos dois tipos de enunciados:

Tipo 1: *Seja α um irracional algébrico de grau n. Então existe $c(\alpha) > 0$ tal que*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^k}$$

todo p, q.

Tipo 2: *Sejam α um número algébrico de grau n e $\epsilon > 0$. Então existe apenas uma quantidade finita de racionais p/q tais que*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{k+\epsilon}}$$

É um exercício fácil provar que se um resultado do tipo 1 é verdadeiro, então um resultado do tipo 2 é verdadeiro.

Entretanto, não é claro que se um resultado do tipo 2 é verdadeiro, então um resultado do tipo 1 é verdadeiro. Por exemplo, o Teorema de Roth é um resultado do tipo 2 com $k = 2$, mas, não se sabe se um resultado do tipo 1 é verdadeiro quando $k = 2$. Ainda no caso $k = 2$, o Teorema de Liouville nos garante a existência de $c(\alpha)$ quando $n = 2$, entretanto, quando α é um número algébrico de grau $n \geq 3$ conjectura-se que não existe tal $c(\alpha) > 0$.

Vamos introduzir um terceiro tipo de resultado:

Tipo 3: *Sejam α um número algébrico de grau n e $\epsilon > 0$. Então existe uma constante $c(\alpha, \epsilon)$ tal que*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha, \epsilon)}{q^{k+\epsilon}}$$

para todo p, q.

Fixado k, temos que um resultado do tipo 2 é verdadeiro se e somente se um resultado do tipo 3 é verdadeiro.

3.2. Números de Liouville. O resultado de Liouville inspirou a seguinte definição: dizemos que um número real α é um número de Liouville se α é irracional e se para todo $n \geq 2$ existem inteiros p e q, com $q > 1$, tais que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^n}.$$

Denotamos por $\mathbb{L}$ o conjunto dos números de Liouville.

Como consequência do Teorema de Liouville temos que todos os números de Liouville são transcendentais. De fato, suponhamos que exista α um número de Liouville, algébrico de grau n. Temos, pelo Teorema de Liouville, que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^n} \tag{6}$$

tem apenas um número finito de soluções. Logo, existe $\epsilon > 0$ tal que

$$\epsilon < \left|\alpha - \frac{p}{q}\right|,$$

para toda solução de (6).

Escolhemos k tal que $\epsilon > 2^{-n-k}$. Como α é um número de Liouville temos que existem p e q tais que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+k}} \leq \frac{1}{2^{n+k}} < \epsilon.$$

Logo $\alpha \in \mathbb{L}$ não pode ser algébrico.
O número de Liouville mais famoso é

$$\sum_{k=1}^{\infty} \frac{1}{10^{k!}}.$$

Este é, de fato, um número de Liouville, pois

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{1}{10^{k!}} - \sum_{k=1}^{n} \frac{1}{10^{k!}} &= \sum_{k=n+1}^{\infty} \frac{1}{10^{k!}} \\
&\leq \frac{1}{10^{(n+1)!}} \sum_{k=0}^{\infty} \frac{1}{10^{k}} \\
&= \frac{10}{9(10^{(n+1)!})} \\
&< \frac{1}{10^{(n+1)!-1}} \\
&\leq \frac{1}{(10^{n!})^{n}}.
\end{aligned}$$

O número $\sum_1^{\infty} 1/10^{k!}$ foi apresentado por Liouville [**22**] como um exemplo de número transcedental, por isso, este número pode ser considerado o primeiro exemplo de número transcendental. Este número é conhecido como constante de Liouville.

Vamos agora provar que $\mathbb{L}$ é um conjunto não-enumerável.

Consideremos a função $f : [0,1] \setminus \mathbb{Q} \to \mathbb{L}$ definida da seguinte forma: Seja $\alpha \in [0,1]$ um número irracional. Podemos escrever $\alpha = a_1 10^{-1} + a_2 10^{-2} + \ldots + a_k 10^{-k} + \ldots$ onde $a_i \in \{0, 1, \ldots, 9\}$. Definimos

$$f(\alpha) = \sum_{k=1}^{\infty} \frac{a_k}{10^{k!}}.$$

Como α é irracional temos que $f(\alpha)$ também o é ($f(\alpha)$ não pode ser uma dízima periódica!). Além disso,

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{a_k}{10^{k!}}-\sum_{k=1}^{n}\frac{a_k}{10^{k!}} &= \sum_{k=n+1}^{\infty}\frac{a_k}{10^{k!}}\\ &\leq 9\sum_{k=n+1}^{\infty}\frac{1}{10^{k!}}\\ &\leq \frac{9}{10^{(n+1)!}}\sum_{k=0}^{\infty}\frac{1}{10^{k}}\\ &\leq \frac{10}{10^{(n+1)!}}\\ &\leq \frac{1}{(10^{n!})^{n}}.\end{aligned}$$

Além disso, como essa função é injetora, concluímos que $\mathbb{L}$ é não-enumerável.

3.3. Densidade e medida dos números de Liouville. Esta seção segue o livro de John C. Oxtoby [**10**].

Vamos introduzir alguns conceitos de topologia para conseguirmos extrair mais algumas informações sobre $\mathbb{L}$.

Um conjunto $A \subset \mathbb{R}$ é um conjunto denso em nenhuma parte se A satisfaz uma das condições equivalentes:

(a) se todo intervalo I possui um subintervalo em $\mathbb{R} \setminus A$;
(b) se $\mathbb{R} \setminus A$ contém um aberto denso;
(c) o interior do fecho de A é vázio.

Um conjunto é dito de primeira categoria (ou magro) se ele pode ser representado como união enumerável de conjuntos densos em nenhuma parte.

Uma demonstração do próximo Teorema pode ser encontrada em [**10**].

Teorema 5 (Baire). *O complemento de um conjunto de primeira categoria é denso em* $\mathbb{R}$.

Vamos usar este resultado para provar que o conjunto dos números de Liouville é denso na reta real.

Teorema 6. *O conjunto dos números de Liouville é denso em* $\mathbb{R}$.

DEMONSTRAÇÃO. Vamos provar que o complementar do conjunto de Liouville é de primeira categoria. Definimos o seguinte conjunto:

$$G_n = \bigcup_{q=2}^{\infty} \bigcup_{p-\infty}^{\infty} (\frac{p}{q} - \frac{1}{q^n}, \frac{p}{q} + \frac{1}{q^n}).$$

E notemos que

$$\mathbb{L} = (\mathbb{R} \setminus \mathbb{Q}) \cap \bigcap_{n=1}^{\infty} G_n.$$

Como $\mathbb{Q} \subset G_n$ temos, por (b), que $\mathbb{R} \setminus G_n$ é denso em nenhuma parte. Além disso, $\mathbb{Q}$ é enumerável, portanto, é união enumerável de conjuntos densos em nenhuma parte. Portanto, temos que

$$\mathbb{R} \setminus \mathbb{L} = \mathbb{Q} \cup \bigcup_{n=1}^{\infty} (\mathbb{R} \setminus G_n),$$

é união de conjuntos densos em nenhuma parte. Segue que $\mathbb{R} \setminus \mathbb{L}$ é de primeira categoria. Portanto, pelo Teorema 5, $\mathbb{L}$ é denso em $\mathbb{R}$. □

Apesar de não-enumerável e denso em $\mathbb{R}$ temos que $\mathbb{L}$ tem medida nula.

PROPOSIÇÃO 6. *O conjunto dos números de Liouville* $\mathbb{L}$ *tem medida nula.*

Esta proposição será provada no último capítulo como consequência da proposição 10. Uma outra demonstração pode ser encontrada em [**10**]

CAPíTULO 2

Teoremas de Thue e Siegel

Any improvement of the exponent n in this lower bound is of utmost importance because it allows one to solve diophantine equations $f(x, y) = N$, in integers x, y, where $f(x, y)$ is a binary form of degree n.

(G. V. Chudnovsky [**23**])

1. Introdução

Nosso principal objetivo neste capítulo é provar o seguinte Teorema:

TEOREMA 7 (Thue). *Seja α um número algébrico de grau $n \geq 3$. Se μ é um número real positivo tal que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}} \tag{7}$$

possui infinitas soluções, então

$$\mu \leq \frac{n}{2} + 1.$$

Thue provou esse resultado em 1909, e como consequência, conseguiu provar a finitude de soluções para vários tipos de equações diofantinas. A demonstração que apresentaremos aqui segue o artigo de Shorey [**9**]. Vamos ainda comentar a demonstração do Teorema de Siegel:

TEOREMA 8 (Siegel). *Seja α um número algébrico de grau $n \geq 3$. Se μ é um número real positivo tal que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}} \tag{8}$$

possui infinitas soluções, então

$$\mu \leq n/[s+1] + s,$$

onde $s = 1, 2, \ldots, n-1$.

Terminaremos este capítulo provando a finitude de soluções das chamadas equações de Thue.

2. Teorema de Thue

2.1. Trabalhando com sistemas de equações. Quando provamos o Teorema de Liouville usamos o polinômio minimal de α para obter o resultado desejado. Visando obter um resultado mais forte que o de Liouville parece natural tentar descobrir se existe um polinômio melhor que o minimal para investigar o quão bem α pode ser aproximado por racionais. Esta é a chave para demonstrar os Teoremas de Thue, Siegel, Dyson e Roth.

Thue explorou identidades polinomiais na forma

$$P(x) - \alpha Q(x) = (\alpha - x)^h \big(f_0(x) + \alpha f_1(x) + \cdots + \alpha^s f_s(x)\big).$$

E neste processo surgiu a ideia de utilizar polinômios em duas variáveis (substituindo α por uma variável y). Buscando conseguir estas identidades polinomiais iremos primeiramente provar alguns resultados sobre soluções de sistemas de equações lineares.

LEMA 2. *Sejam $a_{ij} \in \mathbb{Z}$ com $1 \leq i \leq q$, $1 \leq j \leq p$ e $A \geq \max\{|a_{ij}|\}$. Se $q > p$, o sistema*

$$\sum_{i=0}^{q} a_{ij} x_j = 0, \quad 1 \leq j \leq p \tag{9}$$

tem soluções não-triviais em $x_1, \ldots, x_q \in \mathbb{Z}$ satisfazendo

$$|x_i| \leq (2qA)^{p/(q-p)} \tag{10}$$

para todo $1 \leq i \leq q$.

DEMONSTRAÇÃO. Seja X um inteiro positivo. Definimos

$$U = \{(x_1, \ldots, x_q) \in \mathbb{Z}^q : 0 \leq x_i \leq X\}$$

e

$$V = \Big\{(y_1, \ldots, y_p) \in \mathbb{Z}^p : y_j = \sum_{i=1}^{q} a_{ij} x_i \text{ para } 1 \leq j \leq p \text{ e } (x_1, \ldots, x_q) \in U\Big\}.$$

Definimos $F : U \to V$ por

$$F(x_1, \ldots, x_q) = (y_1, \ldots, y_p).$$

Temos que

$$|y_j| \leq qAX$$

para todo $(y_1, \ldots, y_p) \in V$. Denotemos por $|V|$ e $|U|$ as cardinalidades de V e U. Temos que

$$|V| \leq (2qAX + 1)^p$$

e

$$|U| = (X+1)^q.$$

Quando fixamos

$$X = \lfloor (2qA)^{p/(q-p)} \rfloor.$$

temos que

$$X + 1 > (2qA)^{p/(q-p)},$$

e portanto

$$|U| = (X+1)^p (X+1)^{q-p} > (X+1)^p (2qA)^p > (2qAX+1)^p \geq |V|.$$

Logo F não é injetora. Logo, U possui, pelo menos, dois elementos distintos, $(x'_1, \ldots, x'_q)$ e $(x''_1, \ldots, x''_q)$, tais que

$$\sum_{i=1}^{q} a_{ij} x'_i = \sum_{i=1}^{q} a_{ij} x''_i \text{ para todo } 1 \leq j \leq p.$$

Concluímos que

$$x_i = x'_i - x''_i \quad (1 \leq i \leq q)$$

é uma solução inteira não-trivial de (10) satisfazendo (9).

□

Nosso próximo passo será reduzir o problema de aproximação para um α que é um inteiros algébricos.

2.2. Redução para inteiros algébricos. Vamos mostrar que é suficiente considerar o caso em que α é um inteiro algébrico de grau n. Suponhamos que α não seja um inteiro algébrico e que existam infinitas soluções para

$$\left| \alpha - \frac{h}{q} \right| < \frac{1}{q^{\mu+\epsilon}}. \tag{11}$$

Como α é algébrico existem inteiros $a_0, \ldots, a_n$ com $a_n > 0$ tais que

$$a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0 = 0.$$

Multiplicando a identidade acima por a_n^{n-1} concluímos que $a_n \alpha$ é um inteiro algébrico. Mas, a desigualdade (11) implica que

$$\left| a_n \alpha - \frac{h'}{q} \right| < \frac{a_n}{q^{\mu+\epsilon}} < \frac{1}{q^{\mu+\epsilon'}}$$

possui um número infinito de soluções, onde $q^{\epsilon-\epsilon'} > a_n$, $\epsilon' > 0$ e $\epsilon - \epsilon' > 0$.

2.3. Potências de α como combinações lineares. Fixemos um inteiro algébrico α de grau $n \geq 2$. Seja

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$$

o polinômio minimal de α. Fixemos também $H = \max\{1, |a_1|, \ldots, |a_n|\}$.

LEMA 3. *Dado $s \geq 0$ inteiro sempre podemos escrever*

$$\alpha^s = \sum_{j=0}^{n-1} b_{j,s}\alpha^j, \tag{12}$$

onde $b_{j,s} \in \mathbb{Z}$, e vale a seguinte estimativa:

$$\max_{0 \leq j < n} |b_{j,s}| \leq (2H)^s.$$

DEMONSTRAÇÃO. Iremos provar o Lema por indução em s. Se $s = 0$, temos que $b_{0,0} = 1, b_{j,0} = 0$. Suponhamos que (12) seja verdade e que $s > 0$. Desta forma

$$\alpha^{s+1} = \alpha \sum_{j=0}^{n-1} b_{j,s}\alpha^j = \sum_{j=1}^{n} b_{j-1,s}\alpha^j. \tag{13}$$

Além disso, como $f(\alpha) = 0$, temos que

$$\alpha^n = -\sum_{j=0}^{n-1} a_{n-j}\alpha^j$$

e conseguimos

$$b_{n-1,s}\alpha^n = -b_{n-1,s}\sum_{j=0}^{n-1} a_{n-j}\alpha^j. \tag{14}$$

Usando (13) e (14), podemos escrever

$$\alpha^{s+1} = \sum_{j=0}^{n-1} b_{j,s+1}\alpha^j$$

onde

$$b_{0,s+1} = -b_{n-1,s}a_n$$

e

$$b_{j,s+1} = b_{j-1,s} - a_{n-j}b_{n-1,s} \quad (1 \leq j \leq n-1).$$

Logo

$$\max_{0 \leq j < n} |b_{j,s+1}| \leq 2H \max_{0 \leq j < n} |b_{j,s}|.$$

Aplicando agora a hipótese de indução temos

$$\max_{0 \leq j < n} |b_{j,s+1}| \leq (2H)^{s+1}.$$

□

2.4. A construção do polinômio $R(x,y)$. Dado um polinômio $R(x,y)$ definimos

$$R_m(x,y) = \frac{1}{m!}\frac{\partial^m}{\partial x^m}R(x,y).$$

LEMA 4. *Seja $\delta > 0$ e sejam L e k inteiros positivos satisfazendo*

$$2(L+1) > (1+\delta)nk. \tag{15}$$

Então existe um polinômio não-nulo

$$R(x,y) = \sum_{\lambda_1=0}^{L}\sum_{\lambda_2=0}^{1} p(\lambda_1,\lambda_2)x^{\lambda_1}y^{\lambda_2} \in \mathbb{Z}[x,y] \tag{16}$$

satisfazendo

$$R_m(\alpha,\alpha) = 0 \quad (0 \le m < k) \tag{17}$$

e existe um número real positivo u (dependendo apenas de α e δ) para o qual vale

$$\max_{(\lambda_1,\lambda_2)} |p(\lambda_1,\lambda_2)| \le u^L. \tag{18}$$

DEMONSTRAÇÃO. Seja $R(x,y)$ um polinômio qualquer como em (16). Podemos escrever

$$R_m(\alpha,\alpha) = \sum_{\lambda_1=0}^{L}\sum_{\lambda_2=0}^{1} p(\lambda_1,\lambda_2)\binom{\lambda_1}{m}\alpha^{\lambda_1+\lambda_2-m}$$

e usando o Lema 3 conseguimos

$$\begin{aligned} R_m(\alpha,\alpha) &= \sum_{\lambda_1=0}^{L}\sum_{\lambda_2=0}^{1} p(\lambda_1,\lambda_2)\binom{\lambda_1}{m}\sum_{j=0}^{n-1} b_{j,\lambda_1+\lambda_2-m}\alpha^j \\ &= \sum_{j=0}^{n-1}\alpha^j \sum_{\lambda_1=0}^{L}\sum_{\lambda_2=0}^{1} p(\lambda_1,\lambda_2)\binom{\lambda_1}{m} b_{j,\lambda_1+\lambda_2-m}. \end{aligned} \tag{19}$$

Como α é de grau n, segue que $R_m(\alpha,\alpha) = 0$ se, e somente se,

$$\sum_{\lambda_1=0}^{L}\sum_{\lambda_2=0}^{1} p(\lambda_1,\lambda_2)\binom{\lambda_1}{m} b_{j,\lambda_1+\lambda_2-m} = 0 \quad (0 \le j < n). \tag{20}$$

Pelo Lema 3, temos

$$\max_{(\lambda_1,\lambda_2)} \left|\binom{\lambda_1}{m} b_{j,\lambda_1+\lambda_2-m}\right| \le 2^L(2H)^{L+1} \le (4H)^{2L}.$$

Portanto, podemos aplicar o Lema 2 com $q = 2(L+1), p = nk$ e $A = (4H)^{2L}$ para o sistema de equações lineares (20) obtemos

$$0 < \max_{(\lambda_1,\lambda_2)} |p(\lambda_1,\lambda_2)| \le (2qA)^{p/(q-p)} \le \big(4(L+1)(4H)^{2L}\big)^{p/(q-p)},$$

concluímos de (15) que $p/(q-p) < 1/\delta$ e, assim, vale

$$0 < \max_{(\lambda_1,\lambda_2)} |p(\lambda_1,\lambda_2)| \le \Big(4(L+1)(4H)^{2L}\Big)^{1/\delta} \le (4H)^{4L/\delta}.$$

Terminamos escolhendo $u = (4H)^{4/\delta}$. □

2.5. Uma estimativa para $R_m(p_1/q_1, p_2/q_2)$. A partir deste ponto até a demonstração do Teorema de Thue, sempre que escrevermos $R(x,y)$ fica subentendido que $R(x,y)$ é um polinômio com as propriedades do Lema 4. Fica, também, subentendido que $u_1, \ldots, u_8$ são números reais positivos que dependem apenas de α e δ.

Sejam p_1/q_1 e p_2/q_2 duas aproximações de α. Podemos aplicar a expansão em polinômio de Taylor de $R_m(x,y)$ em (α,α) para estimar $|R_m\big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\big)|$.

LEMA 5. *Seja $\delta > 0$. Sejam L e k inteiros positivos satisfazendo* (15), *isto é, tais que*

$$2(L+1) > (1+\delta)nk.$$

Sejam p_1/q_1 e p_2/q_2 frações reduzidas distintas com $q_1 > 0$ e $q_2 > 0$ satisfazendo

$$\Big|\alpha - \frac{p_i}{q_i}\Big| < 1 \text{ para } i = 1, 2. \tag{21}$$

Então existe u_1 tal que

$$\Big|R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big)\Big| \le \Big(\Big|\alpha - \frac{p_1}{q_1}\Big|^{k-m} + \Big|\alpha - \frac{p_2}{q_2}\Big|\Big) u_1^L \text{ para } 0 \le m < k. \tag{22}$$

DEMONSTRAÇÃO. Podemos escrever o polinômio $R(x,y)$ na forma

$$R(x,y) = P(x) - yQ(x) \tag{23}$$

onde

$$P(x) = \sum_{\lambda_1=0}^{L} p(\lambda_1, 0)x^{\lambda_1}, \quad Q(x) = -\sum_{\lambda_1=0}^{L} p(\lambda_1, 1)x^{\lambda_1}. \tag{24}$$

Aplicando o operador

$$\frac{1}{m!}\frac{\partial}{\partial x^m}$$

em (23), conseguimos

$$R_m(x,y) = P_m(x) - yQ_m(x) \tag{25}$$

onde

$$P_m(x) = \frac{1}{m!}P^{(m)}(x) \;\text{ e }\; Q_m(x) = \frac{1}{m!}Q^{(m)}(x) \tag{26}$$

e assim temos que

$$P_m(x) = \sum_{\lambda_1=0}^{L} p(\lambda_1,0)\binom{\lambda_1}{m}x^{\lambda_1-m} \text{ e } Q_m(x) = \sum_{\lambda_1=0}^{L} p(\lambda_1,1)\binom{\lambda_1}{m}x^{\lambda_1-m}. \tag{27}$$

Pelo Lema 4, temos que $|p(\lambda_1,0)| \leq u^L$donde temos

$$|P_m(x)| \leq (L+1)u^L 2^L(1+|x|)^L$$

e, portanto, temos

$$|P_m(x)| \leq \big(u_2(1+|x|)\big)^L, \tag{28}$$

onde u_2 é dado por

$$u_2 = 4u \geq (L+1)^{1/L}2u.$$

Analogamente, conseguimos

$$|Q_m(x)| \leq (u_2(1+|x|))^L. \tag{29}$$

De (25) e das estimativas (28) e (29), conseguimos

$$|R_m(x,y)| \leq |P_m(x)| + |y||Q_m(x)| \leq (u_2(1+|x|))^L(1+|y|). \tag{30}$$

Além disso, podemos escrever (25) na forma

$$R_m(x,y) = P_m(x) - \alpha Q_m(x) - (y-\alpha)Q_m(x).$$

E aplicando a desigualdade triangular e substituindo $x = p_1/q_1$ e $y = p_2/q_2$ temos

$$\Big|R_m\Big(\frac{p_1}{q_1},\frac{p_2}{q_2}\Big)\Big| \leq \Big|R_m\Big(\frac{p_1}{q_1},\alpha\Big)\Big| + \Big|\alpha - \frac{p_2}{q_2}\Big|\Big|Q_m\Big(\frac{p_1}{q_1}\Big)\Big|. \tag{31}$$

Vamos primeiramente estimar

$$\Big|\alpha - \frac{p_2}{q_2}\Big|\Big|Q_m\Big(\frac{p_1}{q_1}\Big)\Big|.$$

Temos, por (29), que

$$\Big|\alpha - \frac{p_2}{q_2}\Big|\Big|Q_m\Big(\frac{p_1}{q_1}\Big)\Big| \leq \Big|\alpha - \frac{p_2}{q_2}\Big|u_2^L\Big(1+\Big|\frac{p_1}{q_1}\Big|\Big)^L$$

e como

$$\Big|\frac{p_1}{q_1}\Big| \leq |\alpha| + \Big|\alpha - \frac{p_1}{q_1}\Big| \leq |\alpha| + 1,$$

basta definir $u_3 = u_2(2 + |\alpha|)$ para obter

$$\Big|\alpha - \frac{p_2}{q_2}\Big|\Big|Q_m\Big(\frac{p_1}{q_1}\Big)\Big| \leq \Big|\alpha - \frac{p_2}{q_2}\Big|u_3^L. \tag{32}$$

Vamos agora estimar

$$\Big|R_m\Big(\frac{p_1}{q_1}, \alpha\Big)\Big|.$$

Fixemos $0 \leq m < k$. Definimos

$$S(x) = R_m(x, \alpha) \ \text{ e } \ S_\nu(\alpha) = \frac{1}{\nu!}S^{(\nu)}(\alpha)$$

e observamos que

$$S_\nu(\alpha) = \binom{m+\nu}{\nu} R_{m+\nu}(\alpha, \alpha). \tag{33}$$

Por (17), temos que $R_{m+\nu}(\alpha, \alpha) = 0$ sempre que $m + \nu < k$ e concluímos que

$$S_\nu(\alpha) = 0 \text{ para } \nu < k - m. \tag{34}$$

Calculando a expansão em polinômio de Taylor em α de S e usando (34) temos

$$S\Big(\frac{p_1}{q_1}\Big) = S\Big(\frac{p_1}{q_1} - \alpha + \alpha\Big) = \sum_{\nu=k-m}^{L-m} S_\nu(\alpha)\Big(\frac{p_1}{q_1} - \alpha\Big)^\nu. \tag{35}$$

Substituindo (33) em (35), conseguimos

$$R_m\Big(\frac{p_1}{q_1}, \alpha\Big) = \sum_{\nu=k-m}^{L-m} \binom{m+\nu}{\nu} R_{m+\nu}(\alpha, \alpha)\Big(\frac{p_1}{q_1} - \alpha\Big)^\nu.$$

Logo,

$$\begin{aligned}\Big|R_m\Big(\frac{p_1}{q_1}, \alpha\Big)\Big| &\leq \sum_{\nu=k-m}^{L-m} \binom{m+\nu}{\nu} |R_{m+\nu}(\alpha, \alpha)|\Big|\frac{p_1}{q_1} - \alpha\Big|^\nu \\ &= \Big|\frac{p_1}{q_1} - \alpha\Big|^{k-m} \sum_{\nu=0}^{L-k} \binom{k+\nu}{k-m+\nu} |R_{k+\nu}(\alpha, \alpha)|\Big|\frac{p_1}{q_1} - \alpha\Big|^\nu.\end{aligned}$$

Por (21), temos que

$$\Big|\frac{p_1}{q_1} - \alpha\Big|^\nu < 1,$$

para todo $\nu \geq 1$. Por (30), temos que

$$|R_{k+\nu}(\alpha, \alpha)| \leq u_2^L(1 + |\alpha|)^{L+1}.$$

Assim conseguimos mostrar que

$$\left|R_m\Big(\frac{p_1}{q_1},\alpha\Big)\right| \leq \left|\frac{p_1}{q_1}-\alpha\right|^{k-m} u_2^L(1+|\alpha|)^{L+1}\sum_{\nu=0}^{L-k}\binom{k+\nu}{k-m+\nu}.$$

Como

$$\sum_{\nu=0}^{L-k}\binom{k+\nu}{k-m+\nu} \leq \sum_{\nu=k}^{L} 2^\nu = 2^k(2^{L-k+1}-1) = 2^{L+1}-2^k.$$

Definimos

$$u_4 = 4u_2(1+|\alpha|)^2.$$

Pelas estimativas acima temos

$$\left|R_m\Big(\frac{p_1}{q_1},\alpha\Big)\right| \leq u_4^L\left|\alpha-\frac{p_1}{q_1}\right|^{k-m}. \tag{36}$$

Encerramos a demonstração observando que $u_3 < u_4$ e, portanto, basta definirmos $u_1 = u_4$ e substituindo as estimativas (32) e (36) em (31), concluímos o Lema.

□

2.6. Conseguindo m tal que $R_m(p_1/q_1, p_2/q_2) \neq 0$. Observamos que, dado um k, a desigualdade (22) é tanto mais precisa quanto menor for o L. Por outro lado, a desigualdade (15) deve ser satisfeita. Deste ponto em diante iremos considerar L e k como inteiros positivos satisfazendo

$$L = \left\lfloor \frac{1}{2}(1+\delta)nk \right\rfloor \tag{37}$$

e assim (15) está satisfeita. Consequentemente podemos usar os Lemas 4 e 5.

Em nosso próximo resultado iremos usar um famoso resultado conhecido como Lema de Gauss.

LEMA 6 (Gauss). *O produto de dois polinômios primitivos é um polinômio primitivo.*

Um polinômio primitivo é um polinômio cujo máximo divisor comum de seus coeficientes é 1.

LEMA 7. *Sejam $0 < \delta < 1$ e L, k inteiros positivos nas condições fixadas. Sejam p_1/q_1 e p_2/q_2 frações reduzidas distintas com $q_1 \geq 1$ e $q_2 \geq 1$. Existem um número real u_5 e um inteiro m satisfazendo*

$$0 \leq m \leq \left\lfloor \frac{u_5 k}{\log q_1} \right\rfloor + 2$$

tais que

$$R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \neq 0.$$

DEMONSTRAÇÃO. Seja t um inteiro positivo. Suponhamos que

$$R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) = 0 \text{ para } 0 \leq m \leq t. \tag{38}$$

E, portanto, temos

$$P^{(m)}\Big(\frac{p_1}{q_1}\Big) - \frac{p_2}{q_2} Q^{(m)}\Big(\frac{p_1}{q_1}\Big) = 0.$$

Donde concluímos que

$$P^{(m)}\Big(\frac{p_1}{q_1}\Big)Q^{(m')}\Big(\frac{p_1}{q_1}\Big) - P^{(m')}\Big(\frac{p_1}{q_1}\Big)Q^{(m)}\Big(\frac{p_1}{q_1}\Big) = 0 \quad (0 \leq m, m' \leq t). \tag{39}$$

Denotamos

$$W(x) = P(x)Q^{(1)}(x) - P^{(1)}(x)Q(x). \tag{40}$$

Então (39) implica que

$$W^{(\mu)}\Big(\frac{p_1}{q_1}\Big) = 0 \text{ para } 0 \leq \mu < t.$$

Disto concluímos que $(x - p_1/q_1)^t$ divide $W(x)$ e, portanto, $(q_1x - p_1)^t$ divide $W(x)$. Assim, podemos escrever

$$W(x) = (q_1x - p_1)^t H(x),$$

onde $H(x) \in \mathbb{Q}[x]$. Escrevemos então

$$H(x) = g\hat{H}(x),$$

onde $\hat{H}(x)$ é um polinômio primitivo e g é um racional. Temos que

$$W(x) = g(q_1x - p_1)^t \hat{H}(x) \tag{41}$$

pelo Lema de Gauss $(q_1x - p_1)^t\hat{H}(x)$ é um polinômio primitivo, logo, como $W(x)$ tem coeficientes inteiros temos que g é inteiro. Portanto, $H(x) \in \mathbb{Z}[x]$.

Definimos w como o máximo dos valores absolutos dos coeficientes de $W(x)$. Da definição de $W(x)$ em (40) e da estimativa de $|p(\lambda_1, \lambda_2)|$ em (18), temos que existe um número u_6, que depende somente de α e ϵ, tal que

$$w \leq u_6^L. \tag{42}$$

Suponhamos que $W(x)$ não seja o polinômio nulo. Vemos que q_1^t divide o coeficiente do monômio de maior grau de $W(x)$ e assim vale que $q_1^t \leq w$. Por (37) e por (42), existe $u_7 > 1$ tal que

$$w \leq u_7^k.$$

Portanto,

$$t \leq \left\lfloor \frac{k \log u_7}{\log q_1} \right\rfloor + 1.$$

E assim, basta escolher $u_5 = \log u_7$.

Precisamos ainda provar que $W(x)$ não é o polinômio nulo. Pela definição de $R(x, y)$ temos que ou $P(x)$ ou $Q(x)$ não é o polinômio nulo. Supondo que $Q(x)$ é o polinômio nulo temos que $P(x)$ é não-nulo e conseguimos de (23) e (17) que

$$P^{(m)}(\alpha) = 0 \text{ para } 0 \leq m < k.$$

Consequentemente

$$P^{(m)}(\alpha^{(j)}) = 0 \text{ para } 0 \leq m < k, 0 \leq j < n,$$

onde

$$\alpha^{(1)}, \ldots, \alpha^{(n)}$$

são todas as raízes de $f(x)$. Então $(x - \alpha^{(i)})^k$ divide P, todo $1 \leq i \leq n$, e, portanto, o grau de P é maior que nk. Mas, o grau de P é, por definição, menor ou igual a L e $L < nk$. Portanto, Q não pode ser identicamente nulo.

Sabemos que para todo x tal que $Q(x) \neq 0$ vale que

$$W(x) = P(x)Q^{(1)}(x) - P^{(1)}(x)Q(x) = -Q(x)^2 \frac{\partial}{\partial x} \frac{P(x)}{Q(x)}.$$

Como $Q(x)^2$ só é nulo em um número finito de pontos $W(x)$ é o polinômio nulo se e somente se $P(x)/Q(x)$ é constante.

Suponhamos que

$$\frac{P(x)}{Q(x)} = \lambda.$$

Temos que λ não pode ser irracional, pois, $P(1)/Q(1)$ é um número racional. Em particular, $\lambda \neq \alpha$. Além disso, observamos de (23) que

$$R(x, \alpha) = (\lambda - \alpha)Q(x).$$

Logo, conseguimos, de (17), que $Q^{(m)}(\alpha^{(j)}) = 0$ para $0 \leq m < k$ e $0 \leq j < n$. Novamente conseguimos concluir que o grau de Q é maior que nk, o que não pode acontecer já que o grau de Q é menor ou igual a L e $L < nk$. Isto prova que $P(x)/Q(x)$ não é constante. Logo

$$\left(P(x)/Q(x)\right)' = -W(x)/Q^2(x)$$

não é identicamente nulo. Isso completa a prova que $W(x)$ não é identicamente nulo. □

2.7. Teorema de Thue.

Demonstração do Teorema 7. Suponhamos que $\mu > (n/2) + 1$.

Seja $0 < \delta < 1/2$. Suponhamos que exista um número infinito de frações p/q com $q > 0$ satisfazendo

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}}. \tag{43}$$

Seja $u_8 = u_1^{(1+\delta)n/2}$. Como (43) possui infinitas soluções podemos escolher inteiros p_1 e q_1 tais que $(p_1, q_1) = 1$, p_1/q_1 é solução de (53) e vale

$$q_1 > \max\{e^{u_5/\delta}, u_8^{1/\delta}\}. \tag{44}$$

Escolhemos também p_2 e q_2 tais que $(p_2, q_2) = 1$, p_2/q_2 é solução de (53) e vale

$$\log q_2 > (\delta^{-1} + 1) \log q_1. \tag{45}$$

Seja

$$k = \left\lfloor \frac{\log q_2}{\log q_1} \right\rfloor. \tag{46}$$

Portanto, temos

$$q_1^k \leq q_2 < q_1^{k+1}. \tag{47}$$

Definimos

$$L = \left\lfloor \frac{1}{2}(1+\delta)nk \right\rfloor$$

como em (37). Estamos, portanto, nas condições dos Lema 5 e 7.

Temos, por (44), que

$$\left\lfloor \frac{u_5 k}{\log q_1} \right\rfloor + 2 \leq k\delta + 2. \tag{48}$$

Pelo Lema 7, achamos m satisfazendo

$$0 \leq m \leq k\delta + 2 \tag{49}$$

e tal que

$$R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \neq 0. \tag{50}$$

Consequentemente, $R_m(p_1/q_1, p_2/q_2)$ é um número racional cujo denominador é no máximo $q_1^L q_2$. Portanto, vale

$$\left|R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big)\right| \geq q_1^{-L} q_2^{-1}. \tag{51}$$

Como $L \leq (1+\delta)nk/2$ e $q_1^{-k-1} \leq q_2^{-1}$ temos que

$$\left|R_m\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big)\right| \geq q_1^{-\frac{1}{2}(1+\delta)nk-k-1}. \tag{52}$$

Usando o Lema 5

$$\begin{aligned}\left|R_m\Big(\frac{p_1}{q_1},\frac{p_2}{q_2}\Big)\right| &\leq \Big(\Big|\alpha-\frac{p_1}{q_1}\Big|^{k-m}+\Big|\alpha-\frac{p_2}{q_2}\Big|\Big)u_1^L\\ &\leq (q_1^{-\mu(k-m)}+q_2^{-\mu})u_1^L\\ &\leq (q_1^{-\mu(k-k\delta-2)}+q_2^{-\mu})u_1^L\\ &\leq \max\{q_1^{-\mu(k(1-\delta)-2)},q_2^{\mu}\}u_8^k.\end{aligned}$$

Por (44), conseguimos

$$\left|R_m\Big(\frac{p_1}{q_1},\frac{p_2}{q_2}\Big)\right| \leq q_1^{-\mu(k(1-\delta)-2)+\delta k}. \tag{53}$$

Combinando (52) e (53) concluímos que

$$\mu k(1-\delta)-2\mu-\delta k \leq \frac{1}{2}(1+\delta)nk+k+1. \tag{54}$$

Por (45), temos que

$$k \geq \delta^{-1}.$$

Logo:

$$\begin{aligned}\mu &\leq \frac{1}{2}(1+\delta)n+1+\frac{1}{k}\delta+\frac{2\mu}{k}+\mu\delta\\ &\leq \frac{n}{2}+1+\delta(2+3\mu+n/2)\\ &\leq \frac{n}{2}+1+\delta(2+3n+n/2).\end{aligned}$$

Escolhemos

$$\delta=\frac{2\epsilon}{4+7n}$$

e assim conseguimos de (52) e (53) que

$$\mu \leq \frac{n}{2}+1+\epsilon.$$

□

3. Teorema de Siegel

3.1. Alguns pré-requisitos. Vamos agora comentar a demonstração do Teorema 8. Os interessados nas demonstrações dos Lemas que iremos enunciar podem encontrá-las no capítulo 22 do livro de Mordell [**13**].

Assim como na demonstração do Teorema de Thue essa demonstração do Teorema de Siegel passa pela construção de um polinômio.

A primeira diferença de abordagem entre essas demonstrações é abrangência da classe de polinômios estudada. Para provar o Teorema de Siegel estudamos polinômios da forma

$$R(x,y) = \sum_{i=0}^{a} \sum_{j=0}^{b} c_{ij} x^i y^j,$$

onde b é um inteiro positivo qualquer. Lembremos que na demonstração do Teorema de Thue o grau de y era no máximo 1.

Iremos admitir os seguintes Lemas:

LEMA 8. *Seja α um inteiro algébrico de grau $n \geq 3$ e sejam a, b, r inteiros não-negativos tais que*

$$(a+1)(b+1) > rn.$$

Então existe um polinômio não-nulo

$$R(x,y) = \sum_{i=0}^{a} \sum_{j=0}^{b} c_{ij} x^i y^j, \tag{55}$$

onde c_{ij} são racionais, tal que

$$R(x,\alpha) = (x-\alpha)^r S(x), \tag{56}$$

e $S(x)$ é um polinômio em x. Além disso, se

$$R_\lambda(x,y) = \frac{1}{\lambda!} \frac{\partial^\lambda}{\partial x^\lambda} R(x,y), \tag{57}$$

então $R_\lambda(\alpha,\alpha) = 0$ para todo $\lambda = 0, 1 \ldots, r-1$.

LEMA 9. *Sejam b um inteiro fixado tal que $0 \leq b < n$, δ um número real tal que $0 < \delta < 1$, e*

$$a = \left\lfloor \left(\frac{n+\delta}{b+1} \right) r \right\rfloor. \tag{58}$$

Então existe um polinômio nas condições do Lema 8, um inteiro $c = c(\alpha,\delta)$ e um polinômio $R(x,y)$ da forma (55) tal que

$$|c_{ij}| \leq c^r, 0 \leq i \leq a, 0 \leq j \leq b. \tag{59}$$

LEMA 10. *Se $(p_1,q_1) = 1$, $(p_2,q_2) = 1$, $q_1 > 0, q_2 > c^r, r \geq 2n^2, \delta < \frac{1}{2}$, então existem $c_1 = c_1(\alpha,\delta)$ e um inteiro $\lambda = \lambda(\alpha,b,\delta,p_1,q_1,p_2,q_2)$ com $0 \leq \lambda < \delta r + n^2$ tais que*

$$c_1^r q_1^a q_2^b \max\left\{ \left| \alpha - \frac{p_1}{q_1} \right|^{r-\lambda}, \left| \alpha - \frac{p_2}{q_2} \right| \right\} > 1. \tag{60}$$

Uma das ferramentas utilizadas nas demonstrações apresentadas no livro de Mordell é o wronskiano. Este terá um papel de destaque nos próximos capítulos.

3.2. Teorema de Siegel. O Teorema provado por Siegel é o seguinte.

TEOREMA 9 (Siegel). *Seja α um número algébrico de grau $n \geq 3$. Se μ é um número real positivo tal que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}} \tag{61}$$

possui infinitas soluções, então

$$\mu \leq \frac{n}{s+1} + s,$$

onde $s = 1, 2, \ldots, n-1$.

DEMONSTRAÇÃO. Começamos fixando um inteiro $0 \leq b < n$, um número real ϵ, $0 < \epsilon < 1$, e $\mu = n/(b+1) + b + \epsilon$. Seja δ um número real tal que $0 < \delta < 1/2$ e tal que $\delta/(b+1) + \delta\mu < \epsilon/2$. Sejam c e c_1 constantes dadas pelos Lemas 9 e 10, respectivamente.

Suponhamos que a equação (61) possui infinitas soluções. Vamos escolher duas soluções de (61), p_1/q_1 e p_2/q_2, que são frações reduzidas, com $q_1 > \max\{c_1^{4/\epsilon}, c\}$ e $q_2 > q_1$.

Definimos r por

$$r = \left\lfloor \frac{\log q_2}{\log q_1} \right\rfloor \text{ e, portanto, temos } q_1^r \leq q_2 < q_1^{r+1}. \tag{62}$$

Notemos que temos liberdade de escolher q_2 grande o suficiente para que $r \geq 2n^2$ e também $4b + 4n^2\mu < \epsilon r$. Observemos que nossas escolhas implicam que $q_2 > c^r$.

Seja

$$a = \left\lfloor \left(\frac{n+\delta}{b+1}\right) r \right\rfloor$$

e seja λ dado pelo Lema 10. Vamos provar que supor a existência de infinitas soluções de (61) contradiz o Lema 10. Denotamos

$$A = c_1^r q_1^a q_2^b \left|\alpha - \frac{p_1}{q_1}\right|^{r-\lambda}$$

e

$$B = c_1^r q_1^a q_2^b \left|\alpha - \frac{p_2}{q_2}\right|.$$

Como

$$Ac_1^{-r} = q_1^a q_2^b \left|\alpha - \frac{p_1}{q_1}\right|^{r-\lambda},$$

usamos (61) e (62) para conseguir

$$\begin{aligned} Ac_1^{-r} &< q_1^{a-\mu(r-\lambda)+(r+1)b} \\ &< q_1^{\frac{n+\delta}{b+1}r+br+b-\mu r+\mu\lambda} \\ &< \left(q_1^{\frac{n+\delta}{b+1}+b+\frac{b}{r}-\mu+\frac{\mu\lambda}{r}}\right)^r. \end{aligned}$$

Usando que $b = \mu - n/(b+1) - \epsilon$ e $\lambda \leq \delta r + n^2$ temos

$$Ac_1^{-r} < \left(q_1^{\frac{\delta}{b+1}+\delta\mu-\epsilon+\frac{b}{r}+\frac{n^2\mu}{r}}\right)^r.$$

Lembrando que escolhemos δ e r tais que $\delta/(b+1) + \delta\mu < \epsilon/2$ e $4b + 4n^2\mu < \epsilon r$ e, portanto, vale

$$Ac_1^{-r} < q_1^{-\epsilon r/4}.$$

Então, $A < (c_1 q_1^{-\epsilon/4})^r$ e pela nossa escolha de q_1 temos $A < 1$.

Também temos que

$$\begin{aligned} Bc_1^{-r} &< q_1^{\frac{n+\delta}{b+1}} q_2^{b-\mu} \\ &< q_1^{\frac{n+\delta}{b+1}r+r(b-\mu)} \\ &< \left(q_1^{\frac{n+\delta}{b+1}+b-\mu}\right)^r = \left(q_1^{\frac{\delta}{b+1}-\epsilon}\right)^r \\ &< (q_1^{-\epsilon/2})^r. \end{aligned}$$

Logo

$$B < (c_1 q_1^{-\epsilon/2})^r < (c_1 q_1^{-\epsilon/4})^r < 1$$

pela escolha de q_1.

Mas, o Lema 10 nos diz que $\max\{A, B\} > 1$ e a contradição vem de supormos que (61) tem infinitas soluções.

□

Uma consequência deste resultado é que $\mu < 2\sqrt{n}$, pois, escolhendo $s = \lfloor\sqrt{n}\rfloor$ temos que

$$\mu \leq \frac{n}{\lfloor\sqrt{n}\rfloor + 1} + \lfloor\sqrt{n}\rfloor < \frac{n}{\sqrt{n}} + \sqrt{n} = 2\sqrt{n}.$$

4. Aplicações do Teorema de Thue-Siegel.

Uma consequência do Teorema de Thue é o seguinte fato: dados α um número algébrico de grau $n \geq 3$ e um número real A temos que

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{A}{q^n}$$

tem um número finito de soluções. Notemos que este resultado é um pouco mais forte que o Teorema de Liouville. A demonstração deste fato é a seguinte. Para todo $q^{\frac{1}{4}} \geq A$ temos que

$$\begin{aligned}\left|\alpha - \frac{p}{q}\right| &\leq \frac{A}{q^n} \\ &\leq \frac{A}{q^{\frac{1}{4}}} \frac{1}{q^{n-\frac{1}{4}}} \\ &\leq \frac{1}{q^{n-\frac{1}{4}}}\end{aligned}$$

como $n \geq 3$ temos que $n - 1/4 > n/2 + 1$ e, pelo Teorema de Thue, a desigualdade tem apenas um número finito de soluções.

Como consequência deste resultado Thue provou o seguinte Teorema:

TEOREMA 10. *Seja m um inteiro não-nulo. Então a equação*

$$f(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n = m \neq 0, \tag{63}$$

onde $n \geq 3$ e $f(x, y)$ é um polinômio irredutível com coeficientes inteiros, então a equação (63) tem um número finito de soluções inteiras.

DEMONSTRAÇÃO. Quando $y = 0$ temos que $f(x, 0) = m$ tem no máximo duas soluções inteiras. Podemos supor sem perda de generalidade $y > 0$. Sejam $\alpha_1, \ldots, \alpha_n$ as n soluções de

$$f(z, 1) = 0.$$

Temos que

$$|f(x, y)| = \left|a_0 \prod_{i=0}^{n} (x - \alpha_i y)\right| = |m|,$$

então, pelo menos para algum i, digamos $i = 1$, temos que

$$|x - \alpha_1 y| \leq |a_0^{-1} m|^{1/n} = B_1.$$

Então, para $i \neq 1$,

$$\begin{aligned}|x - \alpha_i y| &= |x - \alpha_1 y + (\alpha_1 - \alpha_i) y| \\ &\geq |\alpha_1 - \alpha_i| y - |x - \alpha_1 y|.\end{aligned}$$

Seja $B_2 = \min |\alpha_1 - \alpha_i|, i = 2, \ldots, n$ então

$$\begin{aligned}|x - \alpha_i y| &\geq B_2 y - B_1 \\ &> \frac{1}{2} B_2 y \text{ para todo } y > 2B_1/B_2.\end{aligned}$$

Logo

$$\left|a_0 (x - \alpha_1 y)(B_2 y/2)^{n-1}\right| < |m|,$$

isto é,

$$\left|\frac{x}{y}-\alpha_1\right| < \frac{1}{y^n}\left|\frac{m}{a_0}\right|\left(\frac{2}{B_2}\right)^{n-1} \leq \frac{1}{y^{n-\frac{1}{4}}},$$

quando

$$y^{1/4} > \frac{|m|}{|a_0|}\left(\frac{2}{B_2}\right)^{n-1}.$$

Notemos que α_1 é um número algébrico de grau n, pois, f é irredutível e α_1 é raiz de $f(z,1)$ que é um polinômio irredutível de coeficientes inteiros.

Pelo Teorema de Thue,

$$\left|\frac{x}{y}-\alpha_1\right| \leq \frac{1}{y^{n-\frac{1}{4}}}$$

só pode ter um número finito de soluções.

Logo a equação (63) só possui um número finito de soluções.

□

Devido a este resultado as equações do tipo (63) são conhecidas como equações de Thue.

CAPíTULO 3

Uma demonstração do Teorema de Dyson

If a similar lemma could be proved for polynomials in an arbitrarily large number of variables, the full result could be deduced. The present paper probably represents the limit of what can be done with two variables only. (F. J. Dyson [**2**])

1. Introdução

O principal objetivo deste capítulo é apresentar uma demonstração do seguinte Teorema:

TEOREMA 11 (Dyson [**2**]). *Seja α um irracional algébrico de grau n. Se μ é um número real positivo tal que*

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{\mu}} \tag{64}$$

possui infinitas soluções, então $\mu \leq \sqrt{2n}$.

Dyson publicou a demonstração deste Teorema em 1947. Dois anos depois, Mahler publicou um artigo [**3**] com uma simplificação do único Lema do artigo de Dyson. A principal simplificação do artigo de Mahler é a introdução do índice de um polinômio em duas variáveis. A demonstração do Teorema de Dyson apresentada neste texto é uma adaptação da prova original de Dyson usando o resultado de Mahler. Nosso intuito com essa abordagem é apresentar as principais ideias presentes nos artigos de Mahler e Dyson e as principais dificuldades que foram superadas por Roth visando generalizar estas ideias para polinômios em várias variáveis.

2. Teorema de Mahler

Em seu artigo, Dyson provou o seguinte Lema:

LEMA 11. *Seja $R(x, y)$ um polinômio com coeficientes complexos e de grau não excedendo u em x e s em y. Sejam*

$$\{x_0, x_1, \ldots, x_n\} \text{ e } \{y_0, y_1, \ldots, y_n\}$$

dois conjuntos, cada um deles formado por $n+1$ números complexos. Sejam $\delta, \lambda, t_0, t_1, \ldots, t_n$ números reais tais que

$$\begin{cases} 0 < \delta < 1, & \lambda \geq 2\delta^{-1}, & s \leq \frac{1}{2}n\delta(u+1); \\ 0 \leq t_i \leq s, & \lambda\lfloor t_1 + 1 \rfloor \leq u+1, & i = 0, 1, \ldots, n. \end{cases} \tag{65}$$

Suponha que

$$\Big(\frac{\partial}{\partial x}\Big)^l \Big(\frac{\partial}{\partial y}\Big)^m R(x_i, y_i) = 0 \tag{66}$$

para todos inteiros i, l e m tais que

$$0 \leq i \leq n, \quad 0 \leq m \leq t_i, \quad 0 \leq l \leq \lambda(t_i - m). \tag{67}$$

Então

$$\lambda \sum_{i=0}^{n} (1 + \lfloor t_i \rfloor)(t_i - \frac{1}{2}\lfloor t_i \rfloor) \leq \big(1 + \frac{1}{2}n(n+1)\delta\big)(s+1)(u+1). \tag{68}$$

Com este Lema Dyson foi capaz de provar o Teorema 11. O lado esquerdo de (68) é aproximadamente o número de soluções de (66). A estratégia usada por Dyson é seguinte: seja f o polinômio minimal de α, vamos achar um polinômio $R(x, y)$ com duas propriedades importantes. A primeira delas é que o polinômio $f(x_0)$ divide todas as derivadas

$$\Big(\frac{\partial}{\partial x}\Big)^l \Big(\frac{\partial}{\partial y}\Big)^m R(x, y)\Big|_{x=x_0, y=x_0},$$

quando l e m satisfazem as condições em (67). Portanto, as raízes de f anulam essas derivadas. A segunda propriedade é que podemos escolher p_1/q_1 e p_2/q_2 aproximações de α de forma que o Lema 11 nos garante que existem l_0 e m_0 satisfazendo (67) tais que

$$\Big(\frac{\partial}{\partial x}\Big)^{l_0} \Big(\frac{\partial}{\partial y}\Big)^{m_0} R\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \neq 0.$$

E utilizando o polinômio

$$R_*(x, y) = \Big(\frac{\partial}{\partial x}\Big)^{l_0} \Big(\frac{\partial}{\partial y}\Big)^{m_0} R(x, y),$$

conseguimos uma desigualdade para μ que implica que $\mu \leq \sqrt{2n}$.

Optamos por substituir o Lema 11 pelo Teorema 12. A principal vantagem desta escolha é que desta forma podemos introduzir algumas ideias importantes da demonstração de Roth. A principal delas é o índice.

2.1. Wronskianos. Sejam $u_0(x), u_1(x), \ldots, u_{l-1}(x)$ l polinômios com coeficientes racionais. Chamamos de wronskiano dos polinômios $u_0, \ldots, u_{l-1}$ o determinante

$$W(u_0(x), \ldots, u_{l-1}(x)) = \det\left(\frac{1}{\mu!}\frac{d^\mu u_\lambda(x)}{dx^\mu}\right) \quad (\lambda, \mu = 0, 1, \ldots, l-1).$$

Esta definição é um pouco diferente da usual. Geralmente o wronskiano é definido sem que as derivadas de ordem μ sejam multiplicadas por $(\mu!)^{-1}$. Esta multiplicação servirá para amenizar o crescimento dos coeficientes dos polinômios resultantes e pode ser feita de forma que a derivada do polinômio seja um polinômio de coeficientes inteiros. Poderíamos também definir o wronskiano de quaisquer funções satisfazendo a classe de diferenciabilidade necessária, mas, como nossas aplicações serão para polinômios, vamos nos restringir a este caso.

Os wronskianos são geralmente usados em critérios de dependência linear. Usando as propriedades do determinante, é claro que o wronskiano de funções linearmente dependentes é nulo. A recíproca, em geral, não é verdadeira, por exemplo, se $f_0(x) = x^2$ e $f_1(x) = x|x|$ temos que o wronskiano de f_0 e f_1 é nulo, mas, estas funções são linearmente independentes. Todavia, a recíproca é verdadeira para funções analíticas, isto é, se $f_0, \ldots, f_{l-1}$ são funções analíticas linearmente independentes então seu wronskiano é não-nulo. Uma prova deste fato pode ser encontrada em [**14**]. Apresentaremos aqui uma prova desse resultado para polinômios.

LEMA 12. *Sejam $u_0, \ldots, u_{l-1}$ l polinômios com coeficientes racionais. Temos que o wronskiano destes polinômios é identicamente nulo se, e somente se, estes são linearmente dependentes.*

Por simplicidade vamos ignorar os coeficientes da forma $(\mu!)^{-1}$, pois eles não interferem no resultado.

Antes de provar o Lema, vamos provar este mesmo resultado para monômios. Para isto vamos utilizar a matriz de Vandermonde.

Seja

$$D = \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_n \\ \vdots & \ddots & \vdots \\ (d_1)^{n-1} & \cdots & (d_n)^{n-1} \end{pmatrix}$$

o determinante de uma matriz de Vandermonde é:

$$V(d_1, \ldots, d_n) = \det(D) = \prod_{1 \le i < j \le n} (d_j - d_i).$$

LEMA 13. *O wronskiano dos monômios* $a_1x^{d_1}, \ldots, a_nx^{d_n}$ *é igual a*

$$V(d_1, \ldots, d_n)x^{d_1+\ldots+d_n-\binom{n}{2}} \prod_{i=1}^{n} a_i.$$

DEMONSTRAÇÃO. O wronskiano $W(a_1x^{d_1}, \ldots, a_nx^{d_n})$ é igual a

$$\det \begin{pmatrix} a_1x^{d_1} & \cdots & a_nx^{d_n} \\ a_1d_1x^{d_1-1} & \cdots & a_nd_nx^{d_n-1} \\ \vdots & \ddots & \vdots \\ a_1(d_1)_{n-1}x^{d_1-(n-1)} & \cdots & a_n(d_n)_{n-1}x^{d_n-(n-1)} \end{pmatrix},$$

onde $(d)_k = d(d-1)\ldots(d-k+1)$.

Como o determinante é uma função multilinear nas colunas podemos por em evidência $a_kx^{d_k}$ em cada coluna e assim temos que $W(a_1x^{d_1}, \ldots, a_nx^{d_n})$ é igual a

$$x^{d_1+\cdots+d_n} \prod_{i=1}^{n} a_i \det \begin{pmatrix} 1 & \cdots & 1 \\ d_1x^{-1} & \cdots & d_nx^{-1} \\ \vdots & \ddots & \vdots \\ (d_1)_{n-1}x^{-(n-1)} & \cdots & (d_n)_{n-1}x^{-(n-1)} \end{pmatrix}.$$

Como o determinante também é uma função multilinear nas linha temos que $W(a_1x^{d_1}, \ldots, a_nx^{d_n})$ é igual a

$$x^{d_1+\cdots+d_n-\binom{n}{2}} \prod_{i=1}^{n} a_i \det \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_n \\ \vdots & \ddots & \vdots \\ (d_1)_{n-1} & \cdots & (d_n)_{n-1} \end{pmatrix}.$$

Basta provarmos que o determinante acima é igual ao determinante de D. Lembremos que $(d)_k = d(d-1)\ldots(d-k+1) = d^k + p_k(d)$, onde p_k é um polinômio de grau $k-1$. Por fim, observemos que $p_k(d)$ pode ser escrito como combinação linear de $(d)_{k-1}, \ldots, (d)_2, (d)_1$, portanto, executando operações nas linhas podemos reescrever o determinante acima como o determinante de uma matriz de Vandermonde. □

Vamos fazer mais uma simplificação antes de provarmos o Lema 12. Para isto, definimos o grau inferior de um polinômio como o menor expoente de todos os monômios de coeficientes não-nulos. Nosso próximo Lema objetiva justificar a seguinte observação: basta provarmos o Lema 12 para polinômios de graus inferiores distintos.

LEMA 14. *Sejam* $u_0, \ldots, u_{l-1}$ *polinômios linearmente independentes com coeficientes racionais. Então existe uma matriz* A, $n \times n$, *com*

entradas racionais, tal que os polinômios $v_0, \ldots, v_{l-1}$ *definidos por*

$$(v_0, \ldots, v_{l-1}) = (u_0, \ldots, u_{l-1})A$$

são todos não-nulos e têm graus inferiores distintos. Como consequência, segue a seguinte igualdade:

$$W(v_0, \ldots, v_{l-1}) = W(u_0, \ldots, u_{l-1}) \det(A).$$

DEMONSTRAÇÃO. Se dois polinômios u_0 e u_1 são linearmente independentes, então a menos de uma reenumeração, existe uma combinação linear gerando $\hat{u_1}$, um polinômio de grau inferior maior que o de u_0. Podemos repetir esse processo de forma a obter polinômios $v_0, \ldots, v_{l-1}$ com graus inferiores em ordem crescente. Além disso, como todas as operações realizadas com $u_0, \ldots, u_{l-1}$ são operações elementares, provamos a existência da matriz A, que é invertível, pois é produto de matrizes elementares.

Por fim, é claro de

$$(v_0, \ldots, v_{l-1}) = (u_0, \ldots, u_{l-1})A$$

que vale

$$(v_0^{(i)}, \ldots, v_{l-1}^{(i)}) = (u_0^{(i)}, \ldots, u_{l-1}^{(i)})A$$

para todo $i \geq 1$. □

Para um polinômio u não-nulo e de coeficientes racionais denotamos $MI(u)$ o monômio de u de grau inferior, isto é,

$$u = MI(u) + \text{ termos de grau estritamente maior.}$$

Chamamos $MI(u)$ de monômio inferior de u.

DEMONSTRAÇÃO DO LEMA 12. Caso $u_0, \ldots, u_{l-1}$ são monômios, o resultado segue do Lema 13. De fato, o determinante de Vandemonde $V(d_1, \ldots, d_l)$ é não-nulo se, e somente se, $d_1, \ldots, d_l$ são dois a dois distintos.

No caso geral, se $u_0, \ldots, u_{l-1}$ são linearmente independentes temos, pelo Lema 14, que existem $v_0, \ldots, v_{l-1}$ tais que

$$W(v_0, \ldots, v_{l-1}) = W(u_0, \ldots, u_{l-1}) \det(A) \tag{69}$$

e, portanto, o wronskiano de $u_0, \ldots, u_{l-1}$ é nulo se, e somente se, o wronskiano de $v_0, \ldots, v_{l-1}$ é nulo.

Seja $MI(v_j) = a_j x^{d_j}$ o monômio inferior de v_j. Então a entrada (i, j) da matriz do wronskiano é $(w_{i,j}(1+xr_{i,j}))$ onde $w_{i,j} = a_j(d_j)_{i-1}x^{d_j-i+1}$ e $r_{i,j} \in \mathbb{Q}[x]$. Repetindo os mesmos passos do Lema 13 obtemos

$$W(v_0, \ldots, v_{l-1}) = \det D(x) x^{d_1+\cdots+d_n-\binom{n}{2}} \prod_{k=1}^{n} a_k,$$

onde a matriz $D(x) = (d_j)_{i-1}(1 + xr_{i,j})$. Notemos que $D(0) = D$. Como o determinante é uma função contínua e $\det(D) \neq 0$. Existe uma vizinhança V de 0 tal que $\det D(x) \neq 0, \forall x \in V$.

Disto concluímos que $W(v_0, \ldots, v_{l-1}) \neq 0$ e, por (69), concluímos que $W(u_0, \ldots, u_{l-1}) \neq 0$. □

O wronskiano de polinômios de coeficientes racionais é um polinômio de coeficientes racionais. Nosso próximo passo é calcular o grau deste polinômio. Para tanto vamos primeiro simplificar os cálculos provando que basta estimar o grau para polinômios de graus distintos.

Sejam $u_0, \ldots, u_{l-1}$ polinômios linearmente independentes com coeficientes racionais. Podemos achar polinômios $v_0, \ldots, v_{l-1}$ linearmente independentes tais que o grau de v_k é menor que o grau de v_{k-1} para todo $1 \leq k \leq l-1$ e tais que cada v_k é combinação linear dos polinômios $u_0, \ldots, u_{l-1}$.

Podemos supor que o grau de u_0, digamos d_0, é maior ou igual ao grau de todos os outros polinômios. Logo, existem racionais $c_1, \ldots, c_{l-1}$ tais que podemos definir

$$u_\lambda^1(x) = c_\lambda u_0(x) + u_\lambda(x) \quad (\lambda = 1, 2, \ldots, l-1)$$

de forma que u_λ^1 tem grau menor d_0.

É claro que podemos repetir este processo para $u_1^1, \ldots, u_{l-1}^1$ e conseguir polinômios $u_0, u_1^1, u_2^2, \ldots, u_{l-1}^2$, tais que u_1^1 é de grau $d_1 < d_0$ e com grau de u_λ^2 menor que d_1 para todo $\lambda = 2, \ldots, l-1$. E depois de no máximo $l-1$ passos conseguimos os polinômios $v_0, \ldots, v_{l-1}$, onde $v_0 = u_0, \ldots, v_{l-1} = u_{l-1}^{l-1}$, tais que seus graus são, respectivamente, $d_0, \ldots, d_{l-1}$, onde

$$d_0 > d_1 > \ldots > d_{l-1}.$$

Deste processo concluímos que o wronskiano de $u_0, \ldots, u_{l-1}$ é igual ao wronskiano de $v_0, v_1, \ldots, v_{l-1}$.

LEMA 15. *Sejam $u_0, \ldots, u_{l-1}$ polinômios com coeficientes racionais de grau não excedendo d. Então o wronskiano de $u_0, \ldots, u_{l-1}$ é um polinômio de grau não excedendo $l(d-l+1)$.*

DEMONSTRAÇÃO. Se $u_0, \ldots, u_{l-1}$ são linearmente dependentes não há o que fazer. Podemos supor que $u_0, \ldots, u_{l-1}$ são linearmente independentes e, pela observação anterior, que seus graus são $d_0, \ldots, d_{l-1}$, respectivamente, onde $d_0 > \cdots > d_{l-1}$.

Temos, portanto, que

$$d_0 \leq d,\ d_1 \leq d-1,\ \ldots,\ d_{l-1} \leq d-(l-1).$$

O wronskiano de $u_0, \ldots, u_{l-1}$ é soma de $l!$ termos da forma

$$\pm \frac{1}{(l-1)!} \frac{d^{i_0} u_0(x)}{dx^{i_0}} \frac{d^{i_1} u_1(x)}{dx^{i_1}} \cdots \frac{d^{i_{l-1}} u_{l-1}(x)}{dx^{i_{l-1}}}$$

onde $\{i_0, \ldots, i_{l-1}\} = \{0, 1, \ldots, l-1\}$. Cada um destes termos possui grau limitado por

$$\begin{aligned} \sum_{\lambda=0}^{l-1} (d_\lambda - i_\lambda) &= \sum_{\lambda=0}^{l-1} d_\lambda - \sum_{\lambda=0}^{l-1} (l - \lambda - 1) \\ &\leq \sum_{\lambda=0}^{l-1} (d - \lambda) - (l - \lambda - 1) \\ &= l(d - l + 1), \end{aligned}$$

como queríamos. □

2.2. O índice de polinômios em duas variáveis. Agora que já apresentamos as propriedades do wronskiano, nosso próximo passo é definir o índice de polinômios em duas variáveis. O índice é uma função intrinsicamente ligada a expansão de Taylor do polinômio. Por isso é extremamente conviniente fixarmos a notação: sejam $R(x, y) \in \mathbb{R}[x, y]$ e i, j inteiros não-negativos denotamos

$$R_{i,j}(x, y) = \frac{1}{i!j!} \frac{\partial^{i+j}}{\partial x^i \partial y^j} R(x, y).$$

Dada uma dupla (ξ, η) de números complexos e uma dupla r, s de inteiros positivos definimos

$$c_{(i,j)} = R_{i,j}(\xi, \eta).$$

Definimos a função $f_R : \mathbb{N}^2 \to [0, \infty]$ por

$$f_R(i, j) = \begin{cases} \frac{i}{r} + \frac{j}{s} & \text{se } c_{(i,j)} \neq 0 \\ \infty & \text{caso contrário.} \end{cases}$$

Definimos o índice θ de $R(x, y)$ em (ξ, η) relativo à r, s por

$$\theta(R) = \min_{(i,j)} f_R(i, j).$$

Notemos que se R não é o polinômio nulo vale que $\theta(R) < \infty$. Temos que $\theta(R) \geq 0$ e que $\theta(R) = 0$ se e somente se $R(\xi, \eta) \neq 0$.

A grosso modo, dizer que um polinômio possui um índice grande significa que a expansão de Taylor deste polinômio possui termos iniciais nulos. E conseguir um polinômio $R_*(x, y)$ tal que a expansão em polinômio de Taylor em torno de (α, α) tem termos iniciais nulos é

uma das chaves para conseguirmos desigualdades para μ o expoente que aparece em (11).

Buscamos uma relação entre o índice de um polinômio e o índice de suas derivadas. Seja $P(x,y)$ um polinômio e fixemos i_0 e j_0 inteiros não-negativos. Definimos $P_0(x,y)$ como

$$P_0(x,y) = \Big(\frac{\partial}{\partial x}\Big)^{i_0}\Big(\frac{\partial}{\partial y}\Big)^{j_0} P(x,y).$$

Dados dois números reais (ξ,η) temos que os coeficientes da expansão de Taylor de P_0 em torno de (ξ,η) são

$$\begin{aligned} c_{(i,j)} &= \frac{1}{i!j!}\Big(\frac{\partial}{\partial x}\Big)^{i}\Big(\frac{\partial}{\partial y}\Big)^{j} P_0(x,y)\Big|_{x=\xi,y=\eta} \\ &= \frac{1}{i!j!}\Big(\frac{\partial}{\partial x}\Big)^{i+i_0}\Big(\frac{\partial}{\partial y}\Big)^{j+j_0} P(x,y)\Big|_{x=\xi,y=\eta} \\ &= \frac{(i_0+i)!}{i!}\frac{(j_0+j)!}{j!}c_{(i+i_0,j+j_0)}. \end{aligned}$$

Donde concluímos que

$$\theta(P_0) \geq \theta(P) - \Big(\frac{i_0}{r} + \frac{j_0}{s}\Big). \tag{70}$$

O índice tem um bom comportamento em relação ao produto e a soma de polinômios.

LEMA 16. *Sejam P,Q polinômios. Então, se calculamos todos os índices na mesma dupla (ξ,η) relativos à r,s, temos que*

$$\theta(P+Q) \geq \min\{\theta(P),\theta(Q)\}, \tag{71}$$

$$\theta(PQ) = \theta(P) + \theta(Q). \tag{72}$$

DEMONSTRAÇÃO. Vamos primeiro provar (71).

Temos que $c_{P+Q}(i,j) = c_P(i,j) + c_Q(i,j)$. Se

$$\frac{i}{r} + \frac{j}{s} < \min\{\theta(P),\theta(Q)\},$$

temos que $c_{P+Q}(i,j) = 0$.

Provaremos agora (72).

Existem inteiros não-negativos k_0 e l_0 tais que

$$\frac{k_0}{r} + \frac{l_0}{s} = \theta(P),$$

o termo $P_{k_0,l_0}(\xi,\eta)$ é não-nulo e ainda k_0 é o menor inteiro não-negativo com essa propriedade. Também existem inteiros não-negativos k_1 e l_1

tais que

$$\frac{k_1}{r} + \frac{l_1}{s} = \theta(Q)$$

o termo $Q_{k_1,l_1}(\xi,\eta)$ é não-nulo e k_1 é o menor inteiro não-negativo com essa propriedade.

Escrevemos a expansão em polinômio de Taylor de $P(x,y)$ e $Q(x,y)$ em torno de (ξ,η):

$$P(x,y) = \sum_{i_0=k_0}^{r} \sum_{j_0=0}^{s} P_{i_0,j_0}(\xi,\eta)(x-\xi)^{i_0}(y-\eta)^{j_0}$$

e

$$Q(x,y) = \sum_{i_1=k_1}^{r} \sum_{j_1=0}^{s} Q_{i_1,j_1}(\xi,\eta)(x-\xi)^{i_1}(y-\eta)^{j_1}.$$

Multiplicando temos

$$P(x,y)Q(x,y) = \sum_{i_0=k_0}^{r} \sum_{j_0=0}^{s} \sum_{i_1=k_1}^{r} \sum_{j_1=0}^{s} P_{i_0,j_0}(\xi,\eta)Q_{i_1,j_1}(\xi,\eta)(x-\xi)^{i_0+i_1}(y-\eta)^{j_0+j_1}.$$

Assim temos que o coeficiente de $(x-\xi)^{k_0+k_1}(y-\eta)^{l_0+l_1}$ é

$$P_{k_0,l_0}(\xi,\eta)Q_{k_1,l_1}(\xi,\eta) \neq 0.$$

Logo $\theta(PQ) \leq \theta(P) + \theta(Q)$.

Mas $\theta(PQ) < \theta(P) + \theta(Q)$ não pode ocorrer, pois os coeficientes de PQ são somas de produtos $P_{i_0,j_0}(\xi,\eta)Q_{i_1,j_1}(\xi,\eta)$ e estes são todos nulos se

$$\frac{i_0}{r} + \frac{j_0}{s} < \theta(P) \text{ ou se } \frac{i_1}{r} + \frac{j_1}{s} < \theta(Q).$$

□

2.3. Obtendo um polinômio a partir do Wronskiano. Fixemos o polinômio

$$R(x,y) = \sum_{h=0}^{r} \sum_{k=0}^{s} c_{(h,k)} x^h y^k$$

onde os coeficientes $c_{(h,k)}$ são racionais não todos nulos. Sejam

$$\theta_0, \theta_1, \ldots, \theta_n$$

números reais satisfazendo

$$0 < \theta_f \leq 1 \quad (0 \leq f \leq n),$$

e sejam

$$\{\xi_0, \xi_1, \ldots, \xi_n\} \text{ e } \{\eta_0, \eta_1, \ldots, \eta_n\}$$

dois conjuntos de números complexos cada um destes com $n+1$ elementos.

Supondo que $R(x,y)$ é de índice θ_f em (ξ_f, η_f) temos que

$$R_{i,j}(\xi_f, \eta_f) = 0 \text{ se } i \geq 0,\ j \geq 0,\ \frac{i}{r} + \frac{j}{s} < \theta_f, \tag{73}$$

para todo $0 \leq f \leq n$.

Consideremos as representações do polinômio $R(x,y)$ da seguinte forma

$$R(x,y) = \sum_{\lambda=0}^{l-1} u_\lambda(x) v_\lambda(y),$$

onde u_λ e v_λ são polinômios em uma variável tais que u_λ é de grau no máximo r e v_λ é de grau no máximo s. É claro, tomando $l-1 = s$ e $v_\lambda(y) = y^\lambda$, que existe pelo menos uma dessas representações. Escolhemos uma destas representações que minimize l e, consequentemente, obtemos que $1 \leq l \leq \min\{r,s\} + 1$. A minimalidade de l nos garante que

$$u_0, \ldots, u_{l-1}$$

são linearmente independentes. De fato, caso $u_0, \ldots, u_{l-1}$ fossem linearmente dependentes, poderíamos escrever, reordenando se necessário, que

$$u_{l-1}(x) = \sum_{\lambda=0}^{l-2} a_\lambda u_\lambda(x),$$

onde os coeficientes a_λ são racionais, e assim teríamos

$$R(x,y) = \sum_{\lambda=0}^{l-2} u_\lambda(x)\Big(v_\lambda(y) + a_\lambda v_{l-1}(y)\Big)$$

contrariando a minimalidade de l.

Analogamente, temos que

$$v_0, \ldots, v_{l-1}$$

são linearmente independentes.

Seja $U(x)$ o wronskiano de $u_0, \ldots, u_{l-1}$ e seja $V(x)$ o wronskiano de $v_0, \ldots, v_{l-1}$. Pelo Lema 12, nenhum destes wronskianos é o polinômio nulo. E, pelo Lema 15, o grau de U é no máximo $l(r-l+1)$ e o grau de V é no máximo $l(s-l+1)$.

Para todo f tal que $0 \leq f \leq n$ definimos r_f sendo a maior potência tal que $(x-\xi_f)^{r_f}$ divide $U(x)$. Definimos s_f como a maior potência tal que $(y-\eta_f)^{s_f}$ divide $V(y)$. Concluímos que $U(x)$ é divisível por

$$\prod_{f=0}^{n} (x-\xi_f)^{r_f},$$

e que $V(y)$ é divisível por

$$\prod_{f=0}^{n}(y-\eta_f)^{s_f}.$$

E assim obtemos

$$\left.\begin{array}{l} r_0+r_1+\cdots+r_n \le l(r-l+1), \\ s_0+s_1+\cdots+s_n \le l(s-l+1). \end{array}\right\} \tag{74}$$

Agora seja

$$W(x,y)=\det\Big(R_{\chi,\mu}(x,y)\Big), \quad (\chi,\mu=0,1,\ldots,l-1),$$

onde

$$R_{\chi,\mu}(x,y)=\frac{1}{\chi!\mu!}\sum_{\lambda=0}^{l-1}u_\lambda^{(\chi)}(x)v_\lambda^{(\mu)}(y).$$

Queremos provar que $W(x,y)$ não é o polinômio nulo. Para isso basta provar que $U(x)V(y)=W(x,y)$ e desta forma concluir que $W(x,y)$ é produto de dois polinômios não-nulos. E este resultado é claro se escrevermos os wronskianos

$$U(x)=\det\begin{pmatrix} u_0(x) & u_1(x) & \ldots & u_{l-1}(x) \\ u_0^{(1)}(x) & u_1^{(1)}(x) & \ldots & u_{l-1}^{(1)}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(l-1)!}u_0^{(l-1)}(x) & \frac{1}{(l-1)!}u_1^{(l-1)}(x) & \ldots & \frac{1}{(l-1)!}u_{l-1}^{(l-1)}(x) \end{pmatrix}$$

e

$$V(y)=\det\begin{pmatrix} v_0(y) & v_0^{(1)}(y) & \ldots & \frac{1}{(l-1)!}v_0^{(l-1)}(y) \\ v_1(y) & v_1^{(1)}(y) & \ldots & \frac{1}{(l-1)!}v_1^{(l-1)}(y) \\ \vdots & \vdots & \ddots & \vdots \\ v_{l-1}(y) & v_{l-1}^{(1)}(y) & \ldots & \frac{1}{(l-1)!}v_{l-1}^{(l-1)}(y) \end{pmatrix},$$

e usarmos que o produto dos determinantes é o determinante dos produtos.

Fixemos $f\in\{0,1,\ldots,n\}$. Então, o índice de $R(x,y)$ em (ξ_f,η_f) relativo à r,s é θ_f; por (70) o índice de $R_{\chi,\mu}(x,y)$ em (ξ_f,η_f) é pelo menos

$$\max\Big\{0,\theta_f-\frac{\chi}{r}-\frac{\mu}{s}\Big\}.$$

Lembremos que $W(x,y)$ é soma de $l!$ termos da forma

$$\pm R_{i_0,0}(x,y)R_{i_1,1}(x,y)\ldots R_{i_{l-1},l-1}(x,y),$$

onde $\{i_0, \ldots, i_{l-1}\} = \{0, 1, \ldots l-1\}$. Por (72), o índice de cada um dos $l!$ termos é pelo menos

$$\sum_{\lambda=0}^{l-1} \max\Big\{0, \theta_f - \frac{i_\lambda}{r} - \frac{\lambda}{s}\Big\} \geq \sum_{\lambda=0}^{l-1} \max\Big\{ -\frac{i_\lambda}{r}, \theta_f - \frac{i_\lambda}{r} - \frac{\lambda}{s}\Big\}$$
$$= \sum_{\lambda=0}^{l-1} \max\Big\{0, \theta_f - \frac{\lambda}{s}\Big\} - \sum_{\lambda=0}^{l-1} \frac{i_\lambda}{r}.$$

Usando que

$$\sum_{\lambda=0}^{l-1} \frac{i_\lambda}{r} = \sum_{\lambda=0}^{l-1} \frac{\lambda}{r} = \frac{l(l-1)}{2r},$$

e que, por (71), o índice da soma é maior ou igual ao menor dos índices, concluímos que o índice de $W(x, y)$ em (ξ_f, η_f) é pelo menos

$$\sum_{\lambda=0}^{l-1} \max\Big\{0, \theta_f - \frac{\lambda}{s}\Big\} - \frac{l(l-1)}{2r}.$$

Mas $U(x)V(y)$ é divisível por

$$(x - \xi_f)^{r_f} (y - \eta_f)^{s_f}$$

e, portanto,

$$\frac{1}{i!j!} \frac{\partial^{i+j}}{\partial x^i \partial y^j} \Big(U(x)V(y)\Big)\Big|_{x=\xi_f, y=\eta_f}$$

é nulo se

$$i \geq 0,\ j \geq 0,\ \frac{i}{r} + \frac{j}{s} < \frac{r_f}{r} + \frac{s_f}{s}$$

e é não-nulo se $i = r_f, j = s_f$.

De

$$U(x)V(y) = W(x, y)$$

temos que o índice de W em (ξ_f, η_f) relativo à r, s é

$$\frac{r_f}{r} + \frac{s_f}{s}.$$

Desta forma obtemos a seguinte relação

$$\sum_{\lambda=0}^{l-1} \max\Big\{0, \theta_f - \frac{\lambda}{s}\Big\} - \frac{l(l-1)}{2r} \leq \frac{r_f}{r} + \frac{s_f}{s} \quad (f = 0, 1, \ldots, n). \quad (75)$$

2.4. Melhorando algumas estimativas. Vamos agora trabalhar com as desigualdades em (75). O Teorema de Mahler sairá de manipulações algébricas com essas desigualdades.

Somando estas $n+1$ desigualdades e usando (74) obtemos

$$\sum_{f=0}^{n}\sum_{\lambda=0}^{l-1}\max\Big\{0,\theta_f-\frac{\lambda}{s}\Big\}\le(n+1)\frac{l(l-1)}{2r} + \frac{l(r-l+1)}{r}+\frac{l(s-l+1)}{s}. \tag{76}$$

Nosso próximo passo é simplificar a dupla soma à esquerda em (76). Definimos

$$\Lambda_f=\min\{\lfloor\theta_f s\rfloor+1,l\}\quad(f=0,1,\dots,n), \tag{77}$$

e assim temos

$$\max\Big\{0,\theta_f-\frac{\lambda}{s}\Big\}=\begin{cases}\theta_f-\frac{\lambda}{s} & \text{se } 0\le\lambda\le\Lambda_f-1,\\ 0 & \text{se } \lambda\ge\Lambda_f.\end{cases}$$

Portanto,

$$\sum_{\lambda=0}^{l-1}\max\Big\{0,\theta_f-\frac{\lambda}{s}\Big\}=\sum_{\lambda=0}^{\Lambda_f-1}\Big(\theta_f-\frac{\lambda}{s}\Big)=\frac{1}{2}\Lambda_f\Big(2\theta_f-\frac{\Lambda_f-1}{s}\Big), \tag{78}$$

e assim conseguimos uma primeira simplificação. Sejam

$$X=\frac{l}{s}\ \text{ e }\ X_f=\min\{\theta_f,X\}\quad(f=0,1,\dots,n). \tag{79}$$

Então

$$sX_f=\min\{s\theta_f,sX\}=\min\{s\theta_f,l\}$$

e por (77) vale que

$$\Lambda_f-1\le sX_f\le\Lambda_f,\ \text{logo, }\Lambda_f\Big(2\theta_f-\frac{\Lambda_f-1}{s}\Big)\ge sX_f(2\theta_f-X_f).$$

Juntando essa expressão com (76) e com (78) conseguimos provar que

$$\frac{s}{2}\sum_{f=0}^{n}X_f(2\theta_f-X_f)\le(n+1)\frac{l(l-1)}{2r}+\frac{l(r-l+1)}{r}+\frac{l(s-l+1)}{s}. \tag{80}$$

Agora vamos reescrever o lado direito da desigualdade com X no lugar de l/s

$$\begin{aligned}(n+1)\frac{l(l-1)}{2r}+\frac{l(r-l+1)}{r}+\frac{l(s-l+1)}{s}&\\ =\Big(2l-\frac{l^2}{s}\Big)+\Big(\frac{l}{s}+\frac{(n-1)l(l-1)}{2r}\Big)&\\ =s(2X-X^2)\bigg(1+\frac{1}{2-X}\Big(\frac{1}{s}+\frac{(n-1)(l-1)}{2r}\Big)\bigg)&\\ =s(1-(1-X)^2)\bigg(1+\frac{1}{2-X}\Big(\frac{1}{s}+\frac{(n-1)(l-1)}{2r}\Big)\bigg),&\end{aligned}$$

Usando que

$$l\leq \min\{r,s\}+1\leq s+1$$

conseguimos a seguinte desigualdade

$$\sum_{f=0}^{n}X_f(2\theta_f-X_f)\leq 2(1-(1-X)^2)\bigg(1+\frac{1}{2-X}\Big(\frac{1}{s}+\frac{(n-1)s}{2r}\Big)\bigg). \quad (81)$$

Agora seja δ um número real satisfazendo

$$0<\delta\leq 1,$$

e escolhemos r e s satisfazendo

$$s\geq\frac{5}{\delta},\quad r\geq\frac{5(n-1)s}{2\delta}.$$

Então

$$X=\frac{l}{s}\leq\frac{s+1}{s}\leq 1+\frac{1}{5},\quad 2-X\geq\frac{4}{5},\quad \frac{1}{s}\leq\frac{\delta}{5},\quad \frac{(n-1)s}{2r}\leq\frac{\delta}{5},$$

e assim a temos a seguinte estimativa

$$\frac{1}{2-X}\Big(\frac{1}{s}+\frac{(n-1)s}{2r}\Big)\leq\frac{5}{4}\Big(\frac{\delta}{5}+\frac{\delta}{5}\Big)=\frac{\delta}{2},$$

e desta forma conseguimos estimar o lado direito de (81)

$$\sum_{f=0}^{n}X_f(2\theta_f-X_f)\leq(2+\delta)(1-(1-X)^2). \quad (82)$$

Voltemos a trabalhar com o lado esquerdo da desigualdade em (81).

Fixando f temos que

$$X_f(2\theta_f-X_f)-\theta_f^2(1-(1-X)^2)=\theta_f^2(1-X)^2-(\theta_f-X_f)^2$$

é não-negativo, pois, ou $X\geq\theta_f$ e, por (79), $X_f=\theta_f$ donde obtemos

$$\theta_f^2(1-X)^2-(\theta_f-X_f)^2=\theta_f^2(1-X)^2\geq 0;$$

ou $X < \theta_f$, e, por (79), $X_f = X$ e $X < \theta_f \leq 1$ e portanto, pela diferença de quadrados, temos que

$$\theta_f^2(1-X)^2 - (\theta_f - X)^2 = X(1-\theta_f)\big(\theta_f(1-X) + (\theta_f - X)\big) \geq 0.$$

e desta forma conseguimos provar que

$$(1-(1-X)^2)\sum_{f=0}^{n}\theta_f^2 \leq \sum_{j=0}^{n} X_f(2\theta_f - X_f) \leq (2+\delta)\big(1-(1-X)^2\big).$$

Como $(1-X)^2 < 1$, finalmente obtemos

$$\sum_{f=0}^{n}\theta_f^2 \leq 2+\delta.$$

Acabamos de provar o seguinte Teorema:

TEOREMA 12 (Mahler). *Sejam $\delta, \theta_0, \theta_1, \ldots, \theta_n$, $n+2$ números reais satisfazendo*

$$0 < \delta \leq 1, \quad 0 < \theta_0 \leq 1, \ldots, 0 < \theta_n \leq 1,$$

e sejam r e s inteiros satisfazendo

$$s \geq \frac{5}{\delta}, \quad r \geq \frac{5(n-1)s}{2\delta}. \tag{83}$$

Seja $R(x,y)$ um polinômio não-nulo de grau não excedendo r em x e não excedendo s em y, com coeficientes racionais e denotamos

$$R_{i,j}(x,y) = \frac{1}{i!j!}\frac{\partial^{i+j}}{\partial x^i \partial y^j}R(x,y).$$

Além disso, sejam

$$\{\xi_0, \xi_1, \ldots, \xi_n\} \text{ e } \{\eta_0, \eta_1, \ldots, \eta_n\}$$

dois conjuntos com $n+1$ elementos cada. Se

$$R_{i,j}(\xi_f, \eta_f) = 0 \text{ para todo } i \geq 0,\ j \geq 0,\ \frac{i}{r} + \frac{j}{s} < \theta_f,\ f = 0, 1, \ldots, n,$$

então

$$\theta_0^2 + \theta_1^2 + \cdots + \theta_n^2 \leq 2 + \delta.$$

3. Teorema de Dyson

Agora que provamos o Teorema de Mahler estamos prontos para iniciar a prova do Teorema de Dyson. Nosso primeiro passo será fixar algumas condições iniciais.

3.1. Condições iniciais. Nesta demonstração vamos escolher algumas constantes satisfazendo diversas condições.

Seja α um inteiro algébrico de grau $n > 1$ e seja $f(x)$ o polinômio minimal de α. Temos que $f(x)$ possui n raízes distintas α_i e que podemos escrever

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$$

e assim definimos $A = \max\{1, |a_1|, \ldots, |a_n|\}$.

Escolhemos $0 < \delta \leq 1/(16n)$. Como estamos supondo infinitas soluções de (64) podemos escolher p_1, q_1 tais que p_1/q_1 é solução e satisfaz

$$\delta^2 \log q_1 > 4 \log(2 + 2A) + 4 \log(1 + |\alpha|). \tag{84}$$

Também podemos escolher p_2, q_2 tais que p_2/q_2 é solução de (64) e satisfaz

$$\log q_2 > \frac{5(n-1)}{\delta} \log q_1. \tag{85}$$

Escolhemos r inteiro satisfazendo

$$r > \frac{16n \log q_2}{\delta \log q_1} > \frac{\log q_2}{\delta \log q_1}. \tag{86}$$

E também s inteiro satisfazendo

$$\frac{r \log q_1}{\log q_2} \leq s < 1 + \frac{r \log q_1}{\log q_2}$$

e consequentemente

$$q_1^{r/s} \leq q_2. \tag{87}$$

Destas escolhas temos que

$$\frac{5}{\delta} < \frac{16n}{\delta} < \frac{r \log q_1}{\log q_2} \leq s. \tag{88}$$

Também temos que

$$\frac{s \log q_2}{r \log q_1} < 1 + \frac{\log q_2}{r \log q_1} < 1 + \delta. \tag{89}$$

donde conseguimos

$$q_2^s < q_1^{r(1+\delta)}. \tag{90}$$

Obtemos usando (89) e (85) que

$$\frac{r}{s} > \frac{\log q_2}{\log q_1} \Big(1 + \delta\Big)^{-1} > \frac{5(n-1)}{2\delta}. \tag{91}$$

Notemos que, por (86) e (91), já garantimos que r e s satisfazem as condições (83) do Teorema de Mahler.

Por fim, escolhemos $B = \lfloor q_1^{\delta r} \rfloor$.

3.2. Uma contagem em um conjunto de polinômios. Consideremos o conjunto P de todos os polinômios

$$V(x,y) = \sum_{i=0}^{r} \sum_{j=0}^{s} v_{i,j} x^i y^j, \tag{92}$$

com coeficientes inteiros satisfazendo $0 \leq v_{i,j} \leq B$. Logo P tem

$$N = (B+1)^{(s+1)(r+1)}$$

polinômios.

Fixado $V(x,y) \in P$ vamos considerar todas as derivadas

$$\frac{1}{i!j!}\Big(\frac{\partial}{\partial x}\Big)^i \Big(\frac{\partial}{\partial y}\Big)^j V(x,y)$$

tais que

$$\frac{i}{r} + \frac{j}{s} < k \tag{93}$$

onde k é dado por

$$k = \sqrt{\frac{2+2\delta}{n+\epsilon^2}}$$

e onde $\epsilon = \sqrt{2n\delta(2+\delta)}$. Notemos que ϵ fica pequeno com δ, ϵ é maior que δ e que

$$\epsilon \leq \frac{1}{2}\sqrt{1+\frac{1}{32n}} < 1.$$

Seja N_0 o número de soluções de (93). E seja

$$\phi : \{1, \ldots, N_0\} \longrightarrow \Big\{(i,j) : 0 \leq i \leq r,\ 0 \leq j \leq s,\ \frac{i}{r} + \frac{j}{s} < k\Big\}$$

uma bijeção. Denotamos por D a função que a cada polinômio V em P associa a N_0-upla $D(V)$ tal que a m-ésima coordenada de $D(V)$ é a derivada

$$V_m(x,y) = \Big(\frac{\partial}{\partial x}\Big)^{\phi_1(m)} \Big(\frac{\partial}{\partial y}\Big)^{\phi_2(m)} V(x,y),$$

onde $\phi_1(m)$ e $\phi_2(m)$ são as coordenadas de $\phi(m)$.

Vamos estimar N_0. Para isto vamos primeiramente fixar um j, portanto, as soluções de

$$\frac{i}{r} + \frac{j}{s} < k$$

são todos inteiros $0 \leq i \leq (k - j/s)r$. Logo

$$N_0 = \sum_{j=0}^{\lfloor sk \rfloor} \Big(1 + \Big\lfloor \Big(k - \frac{j}{s}\Big) r \Big\rfloor\Big).$$

Vamos, então, obter uma cota superior para N_0:

$$\begin{aligned} N_0 &= \sum_{j=0}^{\lfloor sk \rfloor} \Big(1 + \Big\lfloor \Big(k - \frac{j}{s}\Big) r \Big\rfloor\Big) \\ &\leq (1 + \lfloor sk \rfloor)(1 + kr) - \frac{(\lfloor sk \rfloor + 1)\lfloor sk \rfloor r}{2s} \\ &\leq (1 + sk)\Big(1 + kr - \frac{\lfloor sk \rfloor r}{2s}\Big) \\ &\leq (1 + sk)\Big(1 + kr - \frac{(sk - 1) r}{2s}\Big) \\ &\leq (1 + sk)\Big(1 + \frac{kr}{2} + \frac{r}{2s}\Big) \\ &= 1 + kr + \frac{r}{2s} + sk + \frac{k^2 rs}{2}. \end{aligned} \tag{94}$$

Depois desta estimativa voltemos a trabalhar com o conjunto $D(V)$. Vamos associar a $D(V)$ uma nova N_0-upla de polinômios $T(V)$, onde a m-ésima coordenada de $T(V)$ é o resto da divisão de $V_m(x, x)$ por $f(x)$.

Se olharmos para $T(V)$ como um conjunto de coeficientes vemos que possui nN_0-uplas de coeficientes. Vamos estimar os coeficientes dos polinômios em $T(V)$. Quando calculamos V em (x, x) temos a soma de todos os monômios de mesmo grau e V tem no máximo $s + 1$ monômios de mesmo grau, portanto, o valor absoluto dos coeficientes de $V(x, x)$ é no máximo $(1 + s)B$. Os coeficientes de $V_m(x, x)$ são produtos de coeficientes de $V(x, x)$ por binômios e disto concluímos que os coeficientes de $V_m(x, x)$ são limitados por

$$2^{r+s}(1 + s)B.$$

Vamos iniciar a divisão $V_m(x, x)$ por $f(x)$. O primeiro passo consiste em multiplicar o polinômio mônico $f(x)$ pelo o coeficiente líder de $V_m(x, x)$ e subtrair de $V_m(x, x)$. E assim obtemos que os coeficientes do polinômio resultante desta diferença são ou da forma $u_1 - au_2$ ou da forma u, onde u, u_1, u_2 são coeficientes de $V_m(x, x)$ e a é coeficiente de $f(x)$. Portanto, temos a estimativa $|u_1 - au_2| \leq |u_1| + A|u_2| \leq (1 + A)2^{r+s}(1 + s)B$. Seja t o número de vezes que executamos esse processo até o grau do polinômio resultante se torne menor que n. Temos que os coeficientes do polinômio resultante são limitados por $(1 + A)^t 2^{r+s}(1 + s)B$.

Assim concluímos que o valor absoluto dos coeficientes de elementos de $T(V)$ não excede

$$M = (2(1+A))^{r+s}(1+s)B.$$

Podemos estimar $2M+1$ por

$$2M+1 < 2(2(1+A))^{r+s}(1+s)(B+1).$$

Sejam

$$T = \bigcup_{V \in P} T(V)$$

e N' o número elementos em T. Usando nossas estimativas concluímos que

$$\begin{aligned} N' &\leq (2M+1)^{nN_0} \\ &\leq (B+1)^{nN_0} 2^{nN_0} (1+s)^{nN_0} (2(1+A))^{(r+s)nN_0} \end{aligned}$$

Por (84), temos que

$$\delta^2 \log q_1 > 4 \log(2(1+A))$$

donde obtemos

$$(2(1+A))^{r+s} < (2(1+A))^{2r} < q_1^{\frac{\delta^2 r}{2}}.$$

Também vale que $2(1+s) < 2^{s+1} < 2^r < q_1^{\frac{\delta^2 r}{4}}$ e como $B = \lfloor q_1^{\delta r} \rfloor$ temos que $B+1 > q_1^{\delta r}$ e assim conseguimos mostrar que

$$2(2(1+A))^{r+s}(1+s) < (B+1)^{\delta}.$$

Logo nossa nova estimativa para N' é:

$$N' < (B+1)^{(1+\delta)nN_0}.$$

Vamos provar que $nN_0(1+\delta) < (s+1)(r+1)$. Por (94) temos que

$$(1+\delta)nN_0 \leq (1+\delta)n\Big(1 + rk + \frac{r}{2s} + sk + \frac{k^2 rs}{2}\Big).$$

Como

$$\frac{k^2}{2} rs(1+\delta)n = rs - \frac{\epsilon^2 rs}{2(n+\epsilon^2)},$$

vale que

$$nN_0(1+\delta) \leq n(1+\delta)\Big(1 + \Big(k + \frac{1}{2s}\Big)r + ks\Big) + rs - \frac{\epsilon^2 rs}{2(n+\epsilon^2)}.$$

Vamos agora provar que

$$n(1+\delta)\Big(1 + \Big(k + \frac{1}{2s}\Big)r + ks\Big) < \frac{\epsilon^2 rs}{2(n+\epsilon^2)}.$$

Usando que $k < 1$ e $s < r$ obtemos que

$$n(1+\delta)\Big(1+\Big(k+\frac{1}{2s}\Big)r+ks\Big) < 8nr.$$

Portanto, basta mostrar que

$$8nr < \frac{\epsilon^2 rs}{2(n+\epsilon^2)} = \frac{\delta(2+\delta)rs}{(1+2\delta(2+\delta))}.$$

Mas isso é verdade, pois, por (88), vale

$$\frac{16n}{\delta} < s.$$

Concluímos que

$$nN_0(1+\delta) < rs < rs+s+r+1 = (s+1)(r+1).$$

Segue que o números de elementos de T é estritamente menor que o número de elementos de P. Logo, existem $V_1(x,y)$ e $V_2(x,y)$, dois polinômios de P, tais que para todos i,j satisfazendo

$$\frac{i}{r}+\frac{j}{s} < k$$

vale que

$$\frac{\partial^{i+j}V_1}{\partial x^i \partial y^j}(x,y)\Big|_{y=x} \quad \text{e} \quad \frac{\partial^{i+j}V_2}{\partial x^i \partial y^j}(x,y)\Big|_{y=x}$$

têm os mesmos restos na divisão por $f(x)$.

3.3. Aplicando o Teorema de Mahler. Vamos aplicar o Teorema de Mahler para $R(x,y) = V_1(x,y) - V_2(x,y)$, $\theta_0 = \epsilon k$, $\xi_0 = p_1/q_1$, $\eta_0 = p_2/q_2$, $\theta_\ell = k$ e $\xi_\ell = \eta_\ell = \alpha_\ell$ onde $1 \le \ell \le n$.

Vamos calcular a soma dos quadrados

$$\begin{aligned}\sum_{\ell=0}^{n}\theta_\ell^2 &= \Big(n+\epsilon^2\Big)k^2 \\ &= \Big(n+\epsilon^2\Big)\Big(\frac{2+2\delta}{n+\epsilon^2}\Big) \\ &= 2+2\delta.\end{aligned}$$

Como $R(x,y) = V_1(x,y) - V_2(x,y)$ temos que

$$R_{i,j}(\alpha_\ell,\alpha_\ell) = 0$$

para todo i,j satisfazendo

$$\frac{i}{r}+\frac{j}{s} < k.$$

Portanto, o Teorema de Mahler implica que existem i_0 e j_0 tais que

$$\frac{i_0}{r} + \frac{j_0}{s} < \epsilon k$$

e tais que

$$R_{i_0,j_0}\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \neq 0.$$

Definimos

$$R_*(x,y) = \frac{1}{i_0! j_0!}\Big(\frac{\partial}{\partial x}\Big)^{i_0}\Big(\frac{\partial}{\partial y}\Big)^{j_0} R(x,y).$$

Temos que

$$\Big(\frac{\partial}{\partial x}\Big)^{i}\Big(\frac{\partial}{\partial y}\Big)^{j} R_*(x,y)\Big|_{x=\alpha,y=\alpha} = 0$$

todo i, j satisfazendo

$$0 \leq \frac{i+i_0}{r} + \frac{j+j_0}{s} < k.$$

Logo

$$\Big(\frac{\partial}{\partial x}\Big)^{i}\Big(\frac{\partial}{\partial y}\Big)^{j} R_*(x,y)\Big|_{x=\alpha,y=\alpha} = 0$$

para todo i, j tais que

$$0 \leq \frac{i}{r} + \frac{j}{s} \leq (1-\epsilon)k.$$

3.4. Finalizando a demonstração. A expansão de $R_*(x,y)$ em polinômio de Taylor nos diz que

$$R_*(x,y) = \sum_{i=0}^{r}\sum_{j=0}^{s} c_{i,j}(x-\alpha)^i(y-\alpha)^j.$$

Os coeficientes são da forma

$$\begin{aligned} c_{i,j} &= \frac{1}{i!j!}\Big(\frac{\partial}{\partial x}\Big)^{i}\Big(\frac{\partial}{\partial y}\Big)^{j} R_*(x,y)\Big|_{x=\alpha,y=\alpha} \\ &= \sum_{\rho=0}^{r}\sum_{\sigma=0}^{s} a_{\rho,\sigma}\Big(\frac{\rho!}{i!i_0!(\rho-i-i_0)!}\Big)\Big(\frac{\sigma!}{j!j_0!(\sigma-j-j_0)!}\Big)\alpha^{\rho-i-i_0+\sigma-j-j_0}, \end{aligned}$$

onde os coeficientes $a_{\rho,\sigma}$ são coeficientes do polinômio $R(x,y)$. Por (87), temos que $q_1^{r/s} \leq q_2$ e assim obtemos

$$\Big|R_*\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big)\Big| \leq \sum_{i=0}^{r}\sum_{j=0}^{s} |c_{i,j}| q_1^{-\mu i} q_2^{-\mu j} \leq \sum_{i=0}^{r}\sum_{j=0}^{s} |c_{i,j}| q_1^{-\mu r(\frac{i}{r}+\frac{j}{s})}.$$

Lembremos que $c_{i,j} = 0$ se

$$\frac{i}{r} + \frac{j}{s} < (1-\epsilon)k$$

e assim

$$\left| R_*\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \right| \leq \sum_{i=0}^{r} \sum_{j=0}^{s} |c_{i,j}| q_1^{-\mu r(1-\epsilon)k}.$$

Como $a_{\rho,\sigma}$ é coeficiente de R temos que

$$|a_{\rho,\sigma}| \leq B.$$

O multinômio

$$\Big(\frac{\rho!}{i!i_0!(\rho - i - i_0)!}\Big)$$

é termo da expansão do multinomial

$$(1+1+1)^\rho,$$

donde concluímos que os coeficientes $c_{i,j}$ são limitados por

$$B(r+1)(s+1)3^{r+s}m^{r+s},$$

onde $m = \max\{1, |\alpha|\}$. Desta forma obtemos

$$\left| R_*\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \right| \leq (r+1)^2(s+1)^2 B(3m)^{r+s} q_1^{-r(1-\epsilon)k\mu}.$$

Vamos majorar essa desigualdade por uma potência de q_1. Por (84) temos

$$s + 1 < r + 1 \leq 2^r \leq q_1^{\delta^2 r/4}.$$

Agora, por (84), também vale que

$$(3m)^{r+s} \leq (3m)^{2r} \leq \big((2+2A)(1+|\alpha|)\big)^{2r} < q^{\delta^2 r}.$$

Nossa estimativa superior é

$$\left| R_*\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big) \right| \leq q_1^{-r(1-\epsilon)k\mu + r(\delta + 2\delta^2)}. \tag{95}$$

Vamos agora conseguir uma cota inferior. Sabemos que $R_*(x, y)$ é um polinômio de grau não excedendo r em x e não excedendo s em y e todos os seus coeficientes são inteiros. Portanto,

$$R_*\Big(\frac{p_1}{q_1}, \frac{p_2}{q_2}\Big)$$

é um racional não nulo e sua fração reduzida possui denominador não excedendo

$$q_1^r q_2^s.$$

Por (90), vale que $q_1^{r(1+\delta)} \geq q_2^s$ e disto temos

$$\left|R_*\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\right| \geq q_1^{-r} q_2^{-s} \geq q_1^{-r(2+\delta)}. \tag{96}$$

Logo, por (95) e (96), temos que

$$q_1^{-r(2+\delta)} < q_1^{-r(1-\epsilon)k\mu + r(\delta+2\delta^2)}.$$

donde vale que

$$-r(2+\delta) \leq r\delta(1+2\delta) - r(1-\epsilon)k\mu.$$

Assim conseguimos

$$(1-\epsilon)k\mu \leq 2 + 2(\delta + \delta^2)$$

e isolando μ concluímos que

$$\begin{aligned}
\mu &\leq \frac{2}{k(1-\epsilon)}(1+\delta+\delta^2) \\
&\leq \frac{2\sqrt{n+\epsilon^2}}{(1-\epsilon)\sqrt{2+2\delta}}(1+\delta+\delta^2) \\
&= \sqrt{2}\sqrt{n+\epsilon^2}\left(\frac{1+\delta+\delta^2}{(1-\epsilon)(\sqrt{1+\delta})}\right).
\end{aligned}$$

Como δ pode ser escolhido suficientemente pequeno concluímos que $\mu \leq \sqrt{2n}$.

Dyson conjecturou que se alguém conseguisse provar um Lema similar ao dele para várias variáveis então seria possível obter $\mu \leq 2$.

O Lema usado por Dyson foi aqui substituído pelo Teorema de Mahler. Notemos que nossa estimativa para μ depende diretamente de k e que, para aplicar o Teorema de Mahler, provamos que a soma dos quadrados é maior que $2+\delta$ e por esse motivo aparece a raiz quadrada na escolha de k.

Para conseguir um resultado melhor que o obtido por Dyson, utilizando estas mesmas técnicas, parece necessário conseguir um resultado mais forte que o obtido por Mahler.

CAPíTULO 4

O Teorema de Roth

My impression of his proof is that it is a struture, inevitably of some complexity, every part of which fits into its proper place and carries its proper share of the total load. (H. Davenport [**24**])

1. Introdução

Depois do artigo de Dyson, parecia evidente a necessidade de utilizar polinômios em várias variáveis para se obter um resultado mais forte. Entretanto, resultados análogos para polinômios com várias variáveis são muito mais complicados.

Vamos agora apresentar uma segunda demonstração do Teorema de Liouville e discutiremos o roteiro da prova do Teorema de Roth.

TEOREMA 13 (Liouville, versão 2). *Seja α um número algébrico de grau $n \geq 2$. Dado $\epsilon > 0$. Então existe apenas uma quantidade finita de racionais $p, q \in \mathbb{Z}$ tais que*

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^{n+\epsilon}}. \tag{97}$$

DEMONSTRAÇÃO. *Passo 1: Construção de um polinômio que anule α.*

O primeiro passo é encontrar um polinômio $P \in \mathbb{Z}[x]$ que anule α. Podemos utilizar o polinômio minimal de α sobre $\mathbb{Q}$ para obter um polinômio nestas condições. Fixemos $P(x) \in \mathbb{Z}[x]$ um polinômio de grau n que anule α.

Passo 2: A existência de soluções de (97) *com denominador "grande" implica na existência de raízes racionais para P.*

Usando o fato que $P(x)$ tem grau n e tem coeficientes inteiros temos que

$$P\Big(\frac{p}{q}\Big) = \frac{N}{q^n} \text{ para algum inteiro } N.$$

Expandindo $P(x)$ em torno de α temos

$$P(x) = \sum_{i=1}^{n} \frac{1}{i!} \frac{d^i P}{dx^i}(\alpha)(x-\alpha)^i.$$

As soluções de (97) satisfazem $|\alpha - p/q| \leq 1$, logo temos

$$\begin{aligned}\frac{|N|}{q^n} &= \Big|P\Big(\frac{p}{q}\Big)\Big| \\ &\leq \Big|\frac{p}{q} - \alpha\Big|\Big(\sum_{i=1}^{n}\Big|\frac{1}{i!}\frac{d^iP}{dx^i}(\alpha)\Big|\Big|\frac{p}{q} - \alpha\Big|^{i-1}\Big) \\ &\leq \Big|\frac{p}{q} - \alpha\Big| n \max_{1\leq i\leq n}\Big|\frac{1}{i!}\frac{d^iP}{dx^i}(\alpha)\Big| \\ &= B(\alpha)\Big|\frac{p}{q} - \alpha\Big| \leq \frac{B(\alpha)}{q^{n+\epsilon}}.\end{aligned}$$

onde

$$B(\alpha) = n \max_{1\leq i\leq n}\Big|\frac{1}{i!}\frac{d^iP}{dx^i}(\alpha)\Big|.$$

Temos que $B(\alpha)$ é uma constante positiva que depende somente de α e P.

Segue que $|N| \leq B(\alpha)/q^{\epsilon}$. Mas N é um inteiro, portanto:

$$\text{Se } \frac{p}{q} \text{ satisfaz (97) e } q > B(\alpha)^{1/\epsilon}, \text{ então } P\Big(\frac{p}{q}\Big) = 0.$$

Passo 3: O polinômio P não pode ter raízes racionais

O polinômio P é um polinômio de grau n que anula α, portanto, é irredutível e não pode ter raízes racionais.

Passo 4: Completando a prova

Suponhamos que a desigualdade (97) tenha infinitas soluções $p/q \in \mathbb{Q}$. Isto implica na existência de uma solução p/q tal que $q > B(\alpha)^{1/\epsilon}$. O passo 2 nos diz que $P(p/q) = 0$, mas, o passo 3 nos diz que P não pode se anular em um racional. Concluímos que a desigualdade (97) não pode ter infinitas soluções. □

Na demonstração do Teorema de Roth, provaremos que a infinitude de soluções de

$$\Big|\alpha - \frac{h}{q}\Big| < \frac{1}{q^k}, \tag{98}$$

onde $k > 2$, nos permite escolher soluções distintas $h_1/q_1, \ldots, h_m/q_m$ de forma a garantir a existência de um polinômio $Q(x_1, \ldots, x_m)$ que, entre outras propriedades, não se anula em $h_1/q_1, \ldots, h_m/q_m$. Este é nosso passo 1.

O nosso passo 2 é mostrar que a expansão em polinômio de Taylor de Q nos dá uma cota superior para $Q(h_1/q_1, \ldots, h_m/q_m)$, por exemplo:

$$|Q(h_1/q_1, \ldots, h_m/q_m)| \leq S.$$

O nosso passo 3 é concluir que como $Q(h_1/q_1, \ldots, h_m/q_m) \neq 0$, então existe um limitante inferior:

$$0 < s \leq |Q(h_1/q_1, \ldots, h_m/q_m)|.$$

Por fim, nosso passo 4 essencialmente é concluir que $S < s$ e, portanto, supor que (98) possui infinitas soluções, quando $k > 2$, nos leva a uma contradição.

A parte mais trabalhosa da demonstração será provar a existência deste polinômio Q. Para provarmos esse resultado precisamos de estimativas para o índice de polinômios em várias variáveis e, para tanto, iniciaremos com o estudo de wronskianos generalizados.

2. Teorema de Roth

2.1. Wronskianos no Teorema de Roth. Em seu trabalho [**1**], Roth utilizou polinômios em várias variáveis para generalizar resultados obtidos por Dyson. Em seu artigo, Dyson relata que a raiz quadrada que aparece em sua estimativa está diretamente ligada ao fato dele ter usado polinômios em duas variáveis e completou conjecturando que a utilização de várias variáveis poderia culminar no resultado esperado. O wronskiano é fundamental na demonstração do Teorema de Dyson e sua versão para polinômios em várias variáveis, o wronskiano generalizado, é fundamental para o Teorema de Roth.

O wronskiano generalizado foi introduzido por Ostrowski e Roth estava ciente que Siegel utilizou o wronskiano generalizado em um contexto semelhante ao seu. Definimos, para polinômios em várias variáveis, o wronskiano generalizado do seguinte modo: sejam

$$\phi_0(x_1, \ldots, x_p), \cdots, \phi_{l-1}(x_1, \ldots, x_p)$$

l polinômios em p variáveis. Consideremos operadores diferenciais da seguinte forma

$$\Delta = \frac{1}{i_1! \ldots i_p!} \Big(\frac{\partial}{\partial x_1}\Big)^{i_1} \ldots \Big(\frac{\partial}{\partial x_p}\Big)^{i_p} \tag{99}$$

e chamamos $i_1 + \cdots + i_p$ a ordem do operador Δ. Sejam

$$\Delta_0, \ldots, \Delta_{l-1}$$

quaisquer operadores diferenciais da forma (99) cujas ordens são no máximo $0, 1, \ldots, l-1$, respectivamente. Chamamos o determinante

$$G_1(x_1, \ldots, x_p) = \det\Big(\Delta_\mu \phi_\nu(x_1, \ldots, x_p)\Big) \quad (\mu, \nu = 0, \ldots, l-1)$$

de um wronskiano generalizado de $\phi_0, \ldots, \phi_{l-1}$. Se $p = 1$ e $\Delta_0, \ldots, \Delta_{l-1}$ são de ordens distintas, então recuperamos o Wronskiano original. Caso

$p > 1$ e $l > 1$ existem mais do que um wronskiano generalizado. Os operadores possuem coeficientes da forma

$$\frac{1}{i_1! \dots i_p!}$$

para amenizar o crescimento dos coeficientes dos polinômios que surgem depois que aplicamos tal operador. É claro que se $\phi_0, \dots, \phi_{l-1}$ são linearmente dependentes então todos os wronskianos generalizados são nulos.

O Lema 17 nos diz que se todos os wronskianos generalizados são identicamente nulos então os polinômios são linearmente dependentes.

LEMA 17. *Se $\phi_0(x_1, \dots, x_p), \dots, \phi_{l-1}(x_1, \dots, x_p)$ são l polinômios linearmente independentes em p variáveis, com coeficientes racionais, então pelo menos um dos wronskianos generalizados não é identicamente nulo.*

DEMONSTRAÇÃO. Fixemos k um inteiro estritamente maior que o grau de todos os polinômios $\phi_0, \dots, \phi_{l-1}$ em cada uma das variáveis $x_1, \dots, x_p$. Consideremos os l polinômios

$$f_\nu(t) = \phi_\nu(t, t^k, \dots, t^{k^{p-1}}) \quad (\nu = 0, \dots, l-1) \tag{100}$$

em uma única variável t. Vamos provar que esses polinômios são linearmente independentes. Escrevemos

$$\phi_\nu(x_1, \dots, x_p) = \sum_{s_1=0}^{k-1} \cdots \sum_{s_p=0}^{k-1} b^{(\nu)}(s_1, \dots, s_p) x_1^{s_1} \dots x_p^{s_p}.$$

Suponhamos que os polinômios em (100) sejam linearmente dependentes. Logo existem racionais $c_0, \dots, c_{l-1}$, não todos nulos, tais que

$$\sum_{\nu=0}^{l-1} c_\nu f_\nu(t) = \sum_{\nu=0}^{l-1} c_\nu \sum_{s_1=0}^{k-1} \cdots \sum_{s_p=0}^{k-1} b^{(\nu)}(s_1, \dots, s_p) t^{s_1 + k s_2 + \cdots + k^{p-1} s_p} = 0.$$

Como a representação de inteiros na forma

$$s_1 + k s_2 + \cdots + k^{p-1} s_p \quad (0 \le s_1 \le k-1, \dots, 0 \le s_p \le k-1)$$

é única (pois a representação de um número da base k é única), temos, para todo ν tal que $c_\nu \neq 0$, que $b^{(\nu)}(s_1, \dots, s_p) = 0$ para todo $(s_1, \dots, s_p)$ e isso implica que todos os coeficientes de ϕ_ν são nulos o que não pode acontecer, pois $\phi_0, \dots, \phi_{l-1}$ são linearmente independentes. Segue do Lema 12 que o wronskiano dos l polinômios em (100), a saber

$$W(t) = \det\Big(\frac{1}{\mu!}\Big(\frac{d}{dt}\Big)^\mu \phi_\nu(t, t^k, \dots, t^{k^{p-1}})\Big) \quad (\mu, \nu = 0, \dots, l-1),$$

não é identicamente nulo.

Resta-nos escrever o wronskiano $W(t)$ a partir dos wronskianos generalizados e com isso provar que existe um wronskiano generalizado não-nulo.

Seja $f : \mathbb{R}^p \to \mathbb{R}$ um polinômio em p variáveis e seja $g : \mathbb{R} \to \mathbb{R}^p$ a aplicação $g(t) = (t, t^k, \ldots, t^{k^{p-1}})$ temos que

$$\frac{d(f \circ g)}{dt}(t) = \sum_{i=1}^{p} \Big(\frac{\partial f}{\partial x_i} \circ g(t)\Big) \frac{dg_i}{dt}(t).$$

Como

$$\frac{dg_i}{dt}(t) = k^i t^{k^i - 1},$$

podemos escrever o operador d/dt na forma

$$\frac{d}{dt} = \frac{\partial}{\partial x_1} + kt^{k-1}\frac{\partial}{\partial x_2} + \cdots + k^{p-1}t^{k^p-1}\frac{\partial}{\partial x_p},$$

onde os operadores à direita são aplicados em um polinômio em $x_1, \ldots, x_p$ e essas variáveis são substituídas por $t, t^k, \ldots, t^{k^{p-1}}$. Visando generalizar esta expressão, vamos calcular $(d/dt)^2$:

$$\frac{d^2(f \circ g)}{dt^2}(t) = \sum_{i=1}^{p}\sum_{j=1}^{p} \Big(\frac{\partial^2 f}{\partial x_i \partial x_j} \circ g(t)\Big) \frac{dg_i}{dt}(t)\frac{dg_j}{dt}(t) + \sum_{i=1}^{p} \Big(\frac{\partial f}{\partial x_i} \circ g_i(t)\Big) \frac{d^2 g_i}{dt^2}(t).$$

Por indução em μ, vemos que o operador $(d/dt)^\mu$ é expressado como combinação linear de operadores diferenciais em $x_1, \ldots, x_p$ da forma (99), de ordem não excedendo μ,

$$\Big(\frac{d}{dt}\Big)^\mu = f_{\mu,1}(t)\Delta_\mu^{(1)} + \cdots + f_{\mu,r}(t)\Delta_\mu^{(r)},$$

onde r depende de μ, $\Delta_\mu^{(1)}, \ldots, \Delta_\mu^{(r)}$ são operadores de ordens não excedendo μ e $f_{\mu,1}(t), \ldots, f_{\mu,r}(t)$ são polinômios com coeficientes racionais.

Concluímos assim que $W(t)$ pode ser escrito na seguinte forma:

$$\det \begin{pmatrix} \phi_0 & \phi_1 & \cdots & \phi_{l-1} \\ \sum_i f_{1,i}(t)\Delta_1^{(i)}\phi_0 & \sum_i f_{1,i}(t)\Delta_1^{(i)}\phi_1 & \cdots & \sum_i f_{1,i}(t)\Delta_1^{(i)}\phi_{l-1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_i f_{l-1,i}(t)\Delta_{l-1}^{(i)}\phi_0 & \sum_i f_{l-1,i}(t)\Delta_{l-1}^{(i)}\phi_1 & \cdots & \sum_i f_{l-1,i}(t)\Delta_{l-1}^{(i)}\phi_{l-1} \end{pmatrix}.$$

Utilizando a linearidade do determinante, conseguimos a seguinte expressão:

$$W(t) = g_1(t)G^{(1)}(t, t^k, \ldots, t^{k^{p-1}}) + \cdots + g_s(t)G^{(s)}(t, t^k, \ldots, t^{k^{p-1}}),$$

onde $G^{(1)}, \ldots, G^{(s)}$ são wronskianos generalizados de $\phi_0, \ldots, \phi_{l-1}$ e $g_1(t)$, $\ldots$, $g_s(t)$ são polinômios em t. Como $W(t)$ não é identicamente nulo,

existe pelos menos um i para o qual $G^{(i)}(t, t^k, \ldots, t^{k^{p-1}})$ não é identicamente nulo e, portanto, $G^{(i)}(x_1, \ldots, x_p)$ não é identicamente nulo.
□

2.2. Conseguindo um polinômio a partir do wronskiano. Nosso objetivo é conseguir associar a um polinômio $R(x_1, \ldots, x_p)$ não-nulo com coeficientes inteiros um polinômio $F(x_1, \ldots, x_p)$ não-nulo com coeficientes inteiros tal que temos a decomposição $F = UV$ onde U é um polinômio em $x_1, \ldots, x_{p-1}$ e V é um polinômio em x_p.

LEMA 18. *Seja $R(x_1, \ldots, x_p)$ um polinômio em $p \geq 2$ variáveis, com coeficientes inteiros, que não é identicamente nulo. Seja R de grau no máximo r_j em x_j para $j = 1, \ldots, p$. Então existe um inteiro l satisfazendo*

$$1 \leq l \leq r_p + 1, \tag{101}$$

e existem operadores diferenciais $\Delta_0, \ldots, \Delta_{l-1}$ nas variáveis $x_1, \ldots, x_{p-1}$, de ordem no máximo $0, \ldots, l-1$, respectivamente, tais que se

$$F(x_1, \ldots, x_p) = \det\Big(\Delta_\mu \frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^\nu R\Big) \quad (\mu, \nu = 0, \ldots, l-1) \tag{102}$$

então

(i) *F tem coeficientes inteiros e não é identicamente nulo;*

(ii) *temos*

$$F(x_1, \ldots, x_p) = U(x_1, \ldots, x_{p-1})V(x_p), \tag{103}$$

onde U e V têm coeficientes inteiros e o grau de U é no máximo lr_j em x_j para $j = 1, \ldots, p-1$ e o grau de V é no máximo lr_p em x_p.

DEMONSTRAÇÃO. Consideremos todas as representações de R na forma

$$R(x_1, \ldots, x_p) = \phi_0(x_p)\psi_0(x_1, \ldots, x_{p-1}) + \ldots + \phi_{l-1}(x_p)\psi_{l-1}(x_1, \ldots, x_{p-1}),$$

onde ϕ_ν e ψ_ν são polinômios com coeficientes racionais, o grau de ϕ_ν é no máximo r_p e o grau de ψ_ν é no máximo r_j em x_j para $j = 1, \ldots, p-1$. É fácil achar uma representação desta forma tomando $l - 1 = r_p$ e $\phi_\nu = x_p^\nu$. Logo, tais representações existem. Escolhemos uma destas representações com l minimal. Segue que $1 \leq l \leq r_p + 1$. Assim temos que

$$\phi_0(x_p), \ldots, \phi_{l-1}(x_p)$$

são linearmente independentes. Suponhamos que não, isto é, temos, por exemplo, que

$$\phi_{l-1} = d_0\phi_0 + \cdots + d_{l-2}\phi_{l-2}$$

com coeficientes racionais $d_0, \ldots, d_{l-2}$ e desta forma concluímos que

$$R = \phi_0(\psi_0 + d_0\psi_{l-1}) + \cdots + \phi_{l-2}(\psi_{l-2} + d_{l-2}\psi_{l-1}),$$

contrariando a minimalidade de l. Analogamente, temos que

$$\psi_0(x_1, \ldots, x_{p-1}), \ldots, \psi_{l-1}(x_1, \ldots, x_{p-1})$$

são linearmente independentes.

Seja $W(x_p)$ o wronskiano de $\phi_0(x_p), \ldots, \phi_{l-1}(x_p)$. Sabemos, portanto, que W é um polinômio com coeficientes racionais, não identicamente nulo e que W se escreve na forma

$$W(x_p) = \det\Big(\frac{1}{\mu!}\Big(\frac{\partial}{\partial x_p}\Big)^{\mu}\phi_\nu(x_p)\Big)(\mu, \nu = 0, \ldots, l-1).$$

Seja $G(x_1, \ldots, x_{p-1})$ um wronskiano generalizado de

$$\psi_0(x_1, \ldots, x_{p-1}), \ldots, \psi_{l-1}(x_1, \ldots, x_{p-1})$$

que não é identicamente nulo, a existência de tal wronskiano é garantida pelo Lema 17. Então

$$G(x_1, \ldots, x_{p-1}) = \det\big(\Delta_\mu\psi_\nu(x_1, \ldots, x_{p-1})\big) \quad (\mu, \nu = 0, \ldots, l-1),$$

onde $\Delta_0, \ldots, \Delta_{l-1}$ são certos operadores diferenciais da forma (99), mas atuando sobre polinômios em $p-1$ variáveis. Temos também que G é um polinômio não-nulo em $x_1, \ldots, x_{p-1}$ com coeficientes racionais.

Podemos escrever

$$W = \det\begin{pmatrix} \phi_0 & \phi_1 & \cdots & \phi_{l-1} \\ \frac{\partial}{\partial x_p}\phi_0 & \frac{\partial}{\partial x_p}\phi_1 & \cdots & \frac{\partial}{\partial x_p}\phi_{l-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(l-1)!}\frac{\partial^{l-1}}{\partial x_p^{l-1}}\phi_0 & \frac{1}{(l-1)!}\frac{\partial^{l-1}}{\partial x_p^{l-1}}\phi_1 & \cdots & \frac{1}{(l-1)!}\frac{\partial^{l-1}}{\partial x_p^{l-1}}\phi_{l-1} \end{pmatrix}$$

e

$$G = \det\begin{pmatrix} \psi_0 & \Delta_1\psi_0 & \cdots & \Delta_{l-1}\psi_0 \\ \psi_1 & \Delta_1\psi_1 & \cdots & \Delta_{l-1}\psi_1 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{l-1} & \Delta_1\psi_{l-1} & \cdots & \Delta_{l-1}\psi_{l-1} \end{pmatrix}.$$

Multiplicando os dois determinantes e usando que o produto dos determinantes é o determinante do produto obtemos

$$\begin{aligned} WG &= \det\Big(\sum_{\rho=0}^{l-1}\Delta_\mu\frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^{\nu}\phi_\rho(x_p)\psi_\rho(x_1, \ldots, x_{p-1})\Big) \\ &= \det\Big(\Delta_\mu\frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^{\nu}R\Big) \quad (\mu, \nu = 0, \ldots, l-1). \end{aligned}$$

Conseguimos construir $F(x_1, \ldots, x_p) = W(x_p)G(x_1, \ldots, x_{p-1})$ como em (102). Sabemos que F tem coeficientes inteiros já que é o determinante de uma matriz na qual as entradas são polinômios com coeficientes inteiros e que W, G têm coeficientes racionais.

Suponhamos que $h \neq 0, q > 1$ e que a fração reduzida h/q seja coeficiente de $x_1^{i_1} \ldots x_{p-1}^{i_{p-1}}$ em W. Temos que todos os monômios de G são múltiplos de q, caso contrário, existe j tal que $x_1^{i_1} \ldots x_{p-1}^{i_{p-1}} x_p^j$ é um monômio de F com coeficiente não-inteiro, o que não pode acontecer. O mesmo ocorre para os coeficientes de G, donde segue que existe um racional g tal que os polinômios $U(x_1, \ldots, x_{p-1}) = gG(x_1, \ldots, x_{p-1})$ e $V(x_p) = g^{-1}W(x_p)$ têm coeficientes inteiros.

Finalmente, como W é um determinante de uma matriz $l \times l$ cujos elementos são polinômios em x_p de grau no máximo r_p, segue que W, e portanto V, é de grau no máximo lr_p. Analogamente, G, e portanto U, é de grau no máximo lr_j em x_j para $j = 1, \ldots, p-1$. □

Nosso próximo Lema nos dá um limitante superior para o valor absoluto dos coeficientes de F a partir de um limitante superior para os coeficientes de R.

LEMA 19. *Seja R um polinômio satisfazendo as condições do Lema 18 e tal que seus coeficientes têm valor absoluto não excedendo B. Então todos os coeficientes de $F(x_1, \ldots, x_p)$, definidos em* (102)*, têm valor absoluto não excedendo*

$$\Big((r_1 + 1) \ldots (r_p + 1)\Big)^l l! B^l 2^{(r_1 + \cdots + r_p)l}.$$

DEMONSTRAÇÃO. Lembrando a definição de F temos que

$$F = \det\Big(\Delta_\mu \frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^\nu R\Big) \quad (\mu, \nu = 0, 1, \ldots, l-1).$$

Usando a notação de multi-índice podemos escrever

$$R(x_1, \ldots, x_p) = \sum_\alpha b_\alpha x^\alpha,$$

onde α é um multi-índice da forma $\alpha = (i_1, \ldots, i_p)$ tal que $0 \leq i_j \leq r_j$,

$$x^\alpha = x_1^{i_1} \ldots x_p^{i_p} \quad \text{e} \quad b_\alpha = b(i_1, \ldots, i_p)$$

assim F é o determinante da matriz

$$\begin{pmatrix} \sum_\alpha b_\alpha x^\alpha & \frac{\partial}{\partial x_p}\sum_\alpha b_\alpha x^\alpha & \dots & \frac{1}{(l-1)!}\Big(\frac{\partial}{\partial x_p}\Big)^{l-1}\sum_\alpha b_\alpha x^\alpha \\ \Delta_1\sum_\alpha b_\alpha x^\alpha & \Delta_1\frac{\partial}{\partial x_p}\sum_\alpha b_\alpha x^\alpha & \dots & \Delta_1\frac{1}{(l-1)!}\Big(\frac{\partial}{\partial x_p}\Big)^{l-1}\sum_\alpha b_\alpha x^\alpha \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{l-1}\sum_\alpha b_\alpha x^\alpha & \Delta_{l-1}\frac{\partial}{\partial x_p}\sum_\alpha b_\alpha x^\alpha & \dots & \Delta_{l-1}\frac{1}{(l-1)!}\Big(\frac{\partial}{\partial x_p}\Big)^{l-1}\sum_\alpha b_\alpha x^\alpha \end{pmatrix}.$$

Notemos que temos $(r_1+1)\dots(r_p+1)$ possíveis multi-índices distintos.

Vamos agora utilizar as propriedades do determinante para escrever F como a soma de determinantes de matrizes formadas apenas por monômios.

Olhando apenas para a primeira coluna e usando a linearidade do determinante para a soma, conseguimos escrever F como a soma de no máximo $(r_1+1)\dots(r_p+1)$ determinantes de matrizes onde a primeira coluna é formada apenas por monômios. Prosseguimos com esse processo para as próximas colunas. Obtemos assim, uma soma de determinantes de matrizes formadas por monômios. Desta forma conseguimos uma representação de F como soma de no máximo $\Big((r_1+1)\dots(r_p+1)\Big)^l$ termos.

Agora cada um destes determinantes têm entradas da forma

$$b_{(i_1,\dots,i_p)}\Delta_\mu\frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^\nu x_1^{i_1}\dots x_p^{i_p},$$

e já sabemos que $|b_{(i_1,\dots,i_p)}|\le B$ e como

$$\Delta_\mu\frac{1}{\nu!}\Big(\frac{\partial}{\partial x_p}\Big)^\nu x_1^{i_1}\dots x_p^{i_p} = Ax_1^{t_1}\dots x_p^{t_p},$$

para todos $t_1\le i_1,\dots,t_p\le i_p$ e o coeficiente A, se não é nulo, é dado por

$$A=\binom{i_1}{t_1}\dots\binom{i_p}{t_p},$$

que tem como limitante

$$A\le 2^{i_1+\dots+i_p}\le 2^{r_1+\dots+r_p}.$$

Desta forma cada um dos $l!$ termos da expansão de cada um dos determinantes têm valor absoluto não excedendo

$$(B2^{r_1+\dots+r_p})^l,$$

portanto, cada um dos $\Big((r_1+1)\dots(r_p+1)\Big)^l$ determinantes têm $l!$ termos de valor absoluto não excedendo $(B2^{r_1+\dots r_p})^l$ donde concluímos

que os coeficientes de F têm valores absolutos limitados por

$$\Big((r_1+1)\dots(r_p+1)\Big)^l l! B^l 2^{(r_1+\dots+r_p)l}.$$

□

2.3. Definindo o Índice. Dados um polinômio em p varíaveis $R(x_1,\dots,x_p)\in\mathbb{R}[x_1,\dots,x_p]$, uma p-upla de números reais $(\alpha_1,\dots,\alpha_p)$ e uma p-upla de inteiros positivos $r_1,\dots,r_p$ definimos a função $f_R:\mathbb{N}^p\to[0,\infty]$ por

$$f_R(i_1,\dots,i_p)=\begin{cases}\frac{i_1}{r_1}+\dots+\frac{i_p}{r_p} & \text{se } c_{(i_1,\dots,i_p)}\neq 0,\\ \infty & \text{caso contrário,}\end{cases}$$

onde $c_{(i_1,\dots,i_p)}$ é dado pela expansão de Taylor de R em $\alpha_1,\dots,\alpha_p$, isto é,

$$R(\alpha_1+y_1,\dots,\alpha_p+y_p)=\sum_{(i_1,\dots,i_p)} c_{(i_1,\dots,i_p)} y_1^{i_1}\dots y_p^{i_p}.$$

Pela expansão em polinômio de Taylor fica claro que

$$c_{(i_1,\dots,i_p)}=\frac{1}{i_1!\dots i_p!}\Big(\frac{\partial}{\partial x_1}\Big)^{i_1}\dots\Big(\frac{\partial}{\partial x_p}\Big)^{i_p}R(\alpha_1,\dots,\alpha_p).$$

Definimos o índice θ de um polinômio $R(x_1,\dots,x_p)$ em $(\alpha_1,\dots,\alpha_p)$ relativo à $r_1,\dots,r_p$ por

$$\theta=\min_{(i_1,\dots,i_p)} f_R(i_1,\dots,i_p).$$

Notemos que se considerarmos apenas polinômios não-nulos temos que $\theta(R)<\infty$ para todo R. Além disso, $\theta(R)\geq 0$ e $\theta(R)=0$ se, e somente se, $R(\alpha_1,\dots,\alpha_p)\neq 0$.

Vamos relacionar o índice de

$$R'(x_1,\dots,x_p)=\Big(\frac{\partial}{\partial x_1}\Big)^{k_1}\dots\Big(\frac{\partial}{\partial x_p}\Big)^{k_p}R(x_1,\dots,x_p)$$

em $(\alpha_1,\dots,\alpha_p)$ relativo à $r_1,\dots,r_p$ com o índice de $R(x_1,\dots,x_p)$ calculado nos mesmos pontos.

Temos que

$$\Big(\frac{\partial}{\partial x_1}\Big)^{k_1}\dots\Big(\frac{\partial}{\partial x_p}\Big)^{k_p}R(x_1,\dots,x_p)=$$
$$\sum_{\substack{(i_1,\dots,i_p)\\ i_j\geq k_j}}\frac{i_1!}{(i_1-k_1)!}\dots\frac{i_p!}{(i_p-k_p)!}c_{(i_1,\dots,i_p)}x_1^{i_1-k_1}\dots x_p^{i_p-k_p}.$$

Portanto, $f_{R'}(i_1 - k_1, \ldots, i_p - k_p) = f_R(i_1, \ldots, i_p) - (k_1/r_1 + \cdots + k_p/r_p)$ sempre que $i_j - k_j \geq 0$ e assim temos que

$$\begin{aligned}\theta(R') &= \min_{\substack{(i_1,\ldots,i_p)\\ i_j \geq k_j}} f_{R'}(i_1 - k_1, \ldots, i_p - k_p) \\ &= \min_{\substack{(i_1,\ldots,i_p)\\ i_j \geq k_j}} f_R(i_1, \ldots, i_p) - \Big(\frac{k_1}{r_1} + \cdots + \frac{k_p}{r_p}\Big) \\ &\geq \quad \theta(R) - \Big(\frac{k_1}{r_1} + \cdots + \frac{k_p}{r_p}\Big).\end{aligned}$$

O próximo Lema nos apresenta outras duas propriedades do índice.

LEMA 20. *Sejam P, Q polinômios não identicamente nulos. Então, se calcularmos todos os índices no mesmo ponto $(\alpha_1, \ldots, \alpha_p)$ relativos à $r_1, \ldots, r_p$, temos que*

$$\theta(P + Q) \geq \min\{\theta(P), \theta(Q)\}, \tag{104}$$

$$\theta(PQ) = \theta(P) + \theta(Q), \tag{105}$$

onde (105) *permanece verdadeiro se P é um polinômio em $x_1, \ldots, x_{p-1}$ com índice calculado em $(\alpha_1, \ldots, \alpha_{p-1})$ relativo à $(r_1, \ldots, r_{p-1})$ e Q é um polinômio em x_p com índice calculado em α_p relativo à r_p.*

DEMONSTRAÇÃO. Provaremos primeiro (104).

Temos que $c_{P+Q}(j_1, \ldots, j_p) = c_P(j_1, \ldots, j_p) + c_Q(j_1, \ldots, j_p)$. Logo, é claro que se

$$\frac{j_1}{r_1} + \cdots + \frac{j_p}{r_p} < \min\{\theta(P), \theta(Q)\}$$

temos que $c_{P+Q}(j_1, \ldots, j_p) = 0$.

Provaremos agora (105).

Seja P_{ind} o polinômio formado pela soma de todos os monômios de P cujo índice é igual ao índice de P. Analogamente, definimos Q_{ind}.

Temos que o polinômio

$$P_{ind}Q_{ind} = \sum_{(j_1,\ldots,j_p)} \sum_{(k_1,\ldots,k_p)} c_P(j_1, \ldots, j_p) c_Q(k_1, \ldots, k_p) x_1^{j_1+k_1} \ldots x_p^{j_p+k_p},$$

é não-nulo e tem índice $\theta(P_{ind}Q_{ind}) = \theta(P_{ind}) + \theta(Q_{ind}) = \theta(P) + \theta(Q)$.

Resta-nos provar que na multiplicação de P por Q todos os monômios que não aparecem no produto $P_{ind}Q_{ind}$ possuem índice maior que $\theta(P) + \theta(Q)$.

Seja $(i_1, \ldots, i_p)$ tal que existem $(j_1, \ldots, j_p)$ e $(k_1, \ldots, k_p)$ com a decomposição

$$(i_1, \ldots, i_p) = (j_1, \ldots, j_p) + (k_1, \ldots, k_p)$$

e $c_P(j_1,\ldots,j_p)\neq 0$ e $c_Q(k_1,\ldots,k_p)\neq 0$. Se

$$\frac{i_1}{r_1}+\cdots+\frac{i_p}{r_p}\leq \theta(P)+\theta(Q)$$

temos obrigatoriamente que

$$\frac{j_1}{r_1}+\cdots+\frac{j_p}{r_p}=\theta(P) \text{ e } \frac{k_1}{r_1}+\cdots+\frac{k_p}{r_p}=\theta(Q),$$

caso contrário, teríamos ou

$$\frac{j_1}{r_1}+\cdots+\frac{j_p}{r_p}<\theta(P)\Rightarrow c_P(j_1,\ldots,j_p)=0$$

ou

$$\frac{k_1}{r_1}+\cdots+\frac{k_p}{r_p}<\theta(P)\Rightarrow c_P(k_1,\ldots,k_p)=0.$$

Assim temos que $\theta(PQ)=\theta(P)+\theta(Q)$ como queríamos. □

2.4. Estimando o índice de um conjunto de polinômios. Vamos fixar um conjunto de polinômios com algumas propriedades e passaremos a investigar estimativas para o maior índice que um polinômio neste conjunto pode atingir.

Dados um número real $B\geq 1$ e uma m-upla de inteiros positivos $r_1,\ldots,r_m$ denotaremos por $\mathcal{R}_m=\mathcal{R}_m(B;r_1,\ldots,r_m)$ o conjunto dos polinômios R tais que

(a) R tem coeficientes inteiros e é não-nulo;
(b) o grau de x_j em R é no máximo r_j;
(c) os coeficientes de R têm valores absolutos menores ou iguais a B.

Sejam $q_1,\ldots,q_m$ inteiros positivos e sejam $h_1,\ldots,h_m$ inteiros satisfazendo $(h_j,q_j)=1$ para $j=1,\ldots,m$. Seja $\theta(R)$ o índice de $R(x_1,\ldots,x_m)$ no ponto $(h_1/q_1,\ldots,h_m/q_m)$ relativo à $r_1,\ldots,r_m$. Definimos

$$\Theta_m(B;q_1,\ldots,q_m;r_1,\ldots,r_m)=\sup\theta(R)$$

onde o sup é calculado em todos os polinômios $R\in\mathcal{R}_m$ e para todos os inteiros $h_1,\ldots,h_m$ que são, respectivamente, primos com $q_1,\ldots,q_m$, isto é, $(h_i,q_i)=1$.

Observemos o duplo significado de $r_1,\ldots,r_m$: eles são ao mesmo tempo limitantes para o grau das variáveis de R no item (b) e também aparecem na definição de $\theta(R)$.

Notemos que estamos calculando o índice para todos as m-uplas $h_1,\ldots,h_m$ de inteiros respectivamente primos com $q_1,\ldots,q_m$ e uma outra forma de dizermos isto é dizer que estamos calculando os índices para todas as m-uplas de racionais com denominador $q_1,\ldots,q_m$.

Nosso objetivo é estimar $\Theta_m(B; q_1, \ldots, q_m; r_1, \ldots, r_m)$ para tanto obteremos uma fórmula indutiva no Lema 22. Para o cálculo de Θ_m quando $m \geq 2$ adicionaremos algumas hipóteses que não são necessárias para a estimativa de $\Theta_1(B; q_1, r_1)$.

LEMA 21. *Temos que*

$$\Theta_1(B; q; r) \leq \frac{\log B}{r \log q}. \tag{106}$$

DEMONSTRAÇÃO. Dados $R \in \mathcal{R}_1$ e h primo com q. Seja $\theta(R)$ o índice de R em h/q relativo à r. Se $\theta(R) = 0$ é claro que $\theta(R) \leq \frac{\log B}{r \log q}$. Suponhamos que $\theta(R) \neq 0$. Seja $k = \theta(R)r$. Pela definição de índice temos que $(x - h/q)^k$ divide $R(x) = \sum_{j=k}^{r} c_j (x - h/q)^j$. Portanto, $(qx-h)^k$ divide $R(x)$. Logo, existe um polinômio $\hat{Q}(x)$ com coeficientes racionais tal que vale $R(x) = (qx-h)^k \hat{Q}(x)$. Podemos escrever $\hat{Q}(x) = gQ(x)$ onde g é um racional e $Q(x)$ é um polinômio primitivo, pelo Lema de Gauss o produto $(qx - h)^k Q(x)$ é um polinômio primitivo donde temos que g é um número inteiro. Disto concluímos que o coeficiente de maior grau de $R(x)$ é múltiplo de $q^k = q^{\theta(R)r}$ desta forma $q^{\theta(R)r} \leq B$ donde $\theta(R) \leq \log B/(r \log q)$. □

Nos próximos Lemas estimaremos Θ_m. Para isto precisamos de algumas hipóteses adicionais sobre $r_1, \ldots, r_m$. A ideia principal do Lema é: dado $R \in \mathcal{R}_m$ conseguir um polinômio "bom o suficiente" para a estimativa do índice e usar essa estimativa para o cálculo do índice de R. Por "bom o suficiente" queremos dizer que este novo polinômio tem propriedades que facilitam o cálculo de seu índice, por exemplo, sabemos que se $F(x_1, \ldots, x_p) = U(x_1, \ldots, x_{p-1})V(x_p)$ vale que $\theta(F) = \theta(U) + \theta(V)$. Ficará claro, ao final da demonstração do Lema 23, a importância do Lema 18.

LEMA 22. *Seja $p \geq 2$ um inteiro positivo e sejam $r_1, \ldots, r_p$ inteiros positivos satisfazendo*

$$r_p > 10\delta^{-1}, r_{j-1}/r_j > \delta^{-1} \text{ para } j = 2, \ldots, p, \tag{107}$$

onde $0 < \delta < 1$. Sejam $B \geq 1$ e $q_1, \ldots, q_p$ inteiros positivos. Então

$$\Theta_p(B; q_1, \ldots, q_p; r_1, \ldots, r_p) \leq 2 \max_l \{\Phi + \Phi^{1/2} + \delta^{1/2}\}, \tag{108}$$

onde o máximo é calculado sobre todos os inteiros l tais que

$$1 \leq l \leq r_p + 1, \tag{109}$$

e onde

$$\Phi = \Theta_1(M; q_p; lr_p) + \Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}) \tag{110}$$

e

$$M = (r_1 + 1)^{pl} l! B^l 2^{plr_1}. \tag{111}$$

DEMONSTRAÇÃO. Sejam $R \in \mathcal{R}_p(B; r_1, \ldots, r_p)$ e $h_1, \ldots, h_p$ inteiros respectivamente primos com $q_1, \ldots, q_p$. Temos que mostrar que o índice $\theta(R)$ é menor que $2 \max_l(\Phi + \Phi^{1/2} + \delta^{1/2})$.

Sabemos pelo Lema 18 que existem um inteiro l satisfazendo (109) e um polinômio $F(x_1, \ldots, x_p)$ com as propriedades (i) e (ii) do Lema 18. Pelo Lema 19, os coeficientes de F têm valores absolutos menores que

$$\Big((r_1 + 1) \ldots (r_p + 1)\Big)^l l! B^l 2^{(r_1 + \cdots + r_p)l}.$$

De (107), temos que $r_1 > r_2 > \ldots > r_p$ e assim concluímos que os coeficientes de F são menores que M definido em (111). Como

$$F = U(x_1, \ldots, x_{p-1}) V(x_p),$$

U, V têm coeficientes inteiros e no produto entre U e V não ocorre nenhuma soma entre os coeficientes dos monômios resultantes, temos que os valores absolutos dos coeficientes de U e V são menores que M.

O polinômio U tem grau no máximo lr_j em x_j para $j = 1, \ldots, p-1$, e, portanto,

$$U \in \mathcal{R}_{p-1}(M; lr_1, \ldots, lr_{p-1}).$$

Segue que o índice de U em $(h_1/q_1, \ldots, h_{p-1}/q_{p-1})$ relativo à $lr_1, \ldots, lr_{p-1}$ não excede

$$\Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}).$$

Donde obtemos que o índice de U relativo à $r_1, \ldots, r_{p-1}$ não excede

$$l\Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}).$$

Analogamente, $V \in \mathcal{R}_1(M; lr_p)$ e seu índice em h_p/q_p relativo à r_p não excede

$$l\Theta_1(M; q_p; lr_p).$$

Pelo Lema 20, o índice de $F = UV$ em $(h_1/q_1, \ldots, h_p/q_p)$ relativo à $r_1, \ldots, r_p$ é a soma dos índices de U e V donde temos

$$\theta(F) \leq l\Phi. \tag{112}$$

Como F foi construído no Lema 18 a partir do determinante de algumas derivadas de R, usaremos esta relação para obtermos um limitante superior para o índice $\theta(R)$ utilizando para isto o índice $\theta(F)$.

Pelo Lema 18, F é o determinante de um polinômio com entradas da forma

$$\Delta \frac{1}{\nu!} \Big(\frac{\partial}{\partial x_p}\Big)^\nu R(x_1, \ldots, x_p) \tag{113}$$

onde o operador

$$\Delta = \frac{1}{i_1! \dots i_{p-1}!} \Big(\frac{\partial}{\partial x_1}\Big)^{i_1} \dots \Big(\frac{\partial}{\partial x_{p-1}}\Big)^{i_{p-1}}$$

é de ordem $w = i_1 + \dots + i_{p-1} \leq l-1$. Portanto, índice do polinômio (113) em $(h_1/q_1, \dots, h_p/q_p)$ relativo à $r_1, \dots, r_p$ é pelo menos

$$\theta(R) - \Big(\frac{i_1}{r_1} + \dots + \frac{i_{p-1}}{r_{p-1}}\Big) - \frac{\nu}{r_p} \geq \theta(R) - \frac{w}{r_{p-1}} - \frac{\nu}{r_p} \geq \theta(R) - \frac{\nu}{r_p} - \delta,$$

onde usamos $w \leq (l-1) \leq r_p < \delta r_{p-1}$, por (109) e (107). Usando que o índice nunca é negativo temos que o índice do polinômio (113) é pelo menos

$$\max\{0, \theta(R) - \nu/r_p - \delta\}.$$

Se expandirmos o determinante da definição de F obtemos uma soma de $l!$ termos cada um da forma

$$\pm(\Delta_{\mu_0} R)\Big(\Delta_{\mu_1} \frac{1}{1!} \frac{\partial}{\partial x_p} R\Big) \dots \Big(\Delta_{\mu_{l-1}} \frac{1}{(l-1)!} \Big(\frac{\partial}{\partial x_p}\Big)^{l-1} R\Big), \qquad (114)$$

onde $\Delta_{\mu_0}, \dots, \Delta_{\mu_{l-1}}$ são operadores diferenciais em $x_1, \dots, x_{p-1}$ e cujas ordens destes operadores são no máximo $l-1$. Pelo item (ii) do Lema 20, o índice de cada termo não-nulo é, pelo menos,

$$\sum_{\nu=0}^{l-1} \max\{0, \theta(R) - \nu/r_p - \delta\} \geq \Big(\sum_{\nu=0}^{l-1} \max\{0, \theta(R) - \nu/r_p\}\Big) - l\delta.$$

Como F é soma de polinômios da forma (114) o item (i) do Lema 20 nos diz que

$$\theta(F) \geq \sum_{\nu=0}^{l-1} \max\{0, \theta(R) - \nu/r_p\} - l\delta.$$

Se $\theta(R) r_p < l$, temos que

$$\begin{aligned}
\sum_{\nu=0}^{l-1} \max\{0, \theta(R) - \nu/r_p\} &= r_p^{-1} \sum_{0 \leq \nu \leq \theta(R) r_p} (\theta(R) r_p - \nu) \\
&= r_p^{-1} \Big(\theta(R) r_p (\lfloor \theta(R) r_p \rfloor + 1) - \frac{\lfloor \theta(R) r_p \rfloor}{2} (\lfloor \theta(R) r_p \rfloor + 1)\Big) \\
&= r_p^{-1} (\lfloor \theta(R) r_p \rfloor + 1) \Big(\theta(R) r_p - \frac{\lfloor \theta(R) r_p \rfloor}{2}\Big) \\
&\geq r_p^{-1} (\lfloor \theta(R) r_p \rfloor + 1) \frac{\theta(R) r_p}{2} \\
&\geq \frac{1}{2} \theta(R)^2 r_p.
\end{aligned}$$

Se $\theta(R)r_p \geq l$, temos

$$\begin{aligned}\sum_{\nu=0}^{l-1} \max\{0, \theta(R) - \nu/r_p\} &= \sum_{\nu=0}^{l-1} (\theta(R) - \nu/r_p) \\ &= \theta(R)l - \frac{l(l-1)}{2r_p} \\ &\geq \theta(R)l - \frac{l\theta(R)r_p}{2r_p} \\ &= \frac{1}{2}l\theta(R).\end{aligned}$$

Logo

$$\theta(F) \geq \min\left\{\frac{1}{2}l\theta(R), \frac{1}{2}\theta(R)^2 r_p\right\} - l\delta. \tag{115}$$

Combinando as desigualdades (112) e (115), obtemos

$$\min\left\{\frac{1}{2}l\theta(R), \frac{1}{2}\theta(R)^2 r_p\right\} \leq l(\Phi + \delta).$$

No primeiro caso $\frac{1}{2}l\theta(R) \leq l(\Phi + \delta)$ temos que $\theta(R) \leq 2(\Phi + \delta)$ e a desigualdade (108) está satisfeita.

No segundo caso temos que

$$\frac{1}{2}r_p\theta(R)^2 \leq l(\Phi + \delta) \leq (r_p + 1)(\Phi + \delta).$$

Como $r_p + 1 < 2r_p$ por (107), donde temos

$$\theta(R) \leq 2(\Phi + \delta)^{1/2} \leq 2(\Phi^{1/2} + \delta^{1/2})$$

e, também neste caso, a desigualdade (108) está satisfeita.

□

Notemos que poderíamos substituir a hipótese $r_{j-1}/r_j > \delta^{-1}$ no Lema 22 por $r_1 > \ldots > r_{p-1}$ e $r_{p-1}/r_p > \delta^{-1}$. Usaremos que $r_{j-1}/r_j > \delta^{-1}$ apenas no Lema 23 quando aplicarmos a hipótese de indução.

LEMA 23. *Sejam m um inteiro positivo e δ satisfazendo*

$$0 < \delta < m^{-1}. \tag{116}$$

Sejam $r_1, \ldots, r_m$ de inteiros positivos satisfazendo

$$r_m > 10\delta^{-1}, r_{j-1}/r_j > \delta^{-1} \text{ para } j = 2, \ldots, m. \tag{117}$$

Sejam $q_1, \ldots, q_m$ inteiros positivos satisfazendo

$$\log q_1 > \delta^{-1}m(2m+1), \tag{118}$$

$$r_j \log q_j \geq r_1 \log q_1 \text{ para } j = 2, \ldots, m. \tag{119}$$

Então

$$\Theta_m(q_1^{\delta r_1}; q_1, \ldots, q_m; r_1, \ldots, r_m) < 10^m \delta^{(1/2)^m}. \tag{120}$$

DEMONSTRAÇÃO. Provaremos o Lema por indução em m. Caso $m = 1$, temos, pelo Lema 21, que

$$\Theta_1(q_1^{\delta r_1}; q_1; r_1) \leq \frac{\log q_1^{\delta r_1}}{r_1 \log q_1} = \delta \leq 10\delta^{1/2},$$

onde usamos que $\delta < 1$.

Suponhamos que $p \geq 2$ é um inteiro e que o Lema 23 seja válido quando $m = p - 1$. Notemos (116) e (117) nos garantem que podemos utilizar o Lema 22. E assim $\Theta_m(q_1^{\delta r_1}; q_1, \ldots, q_m; r_1, \ldots, r_m) \leq 2\max\{\Phi + \Phi^{1/2} + \delta^{1/2}\}$. Prosseguiremos com estimativas para M e Φ.

Queremos escrever M da forma $q^{\delta l r}$ para podermos proceder por indução

$$M = (r_1 + 1)^{pl} l! 2^{plr_1} q_1^{\delta l r_1} \leq \left((r_1 + 1)^p l 2^{pr_1} q_1^{\delta r_1}\right)^l.$$

Usando que $l \leq r_p + 1 < r_1 + 1 \leq 2^{r_1}$ temos que

$$M < (2^{(2p+1)r_1} q_1^{\delta r_1})^l < (e^{(2p+1)r_1} q_1^{\delta r_1})^l.$$

Por (118) com $m = p$, temos que $2p + 1 < \delta p^{-1} \log q_1$, portanto vale

$$M < q_1^{\delta_1 l r_1},$$

onde

$$\delta_1 = \delta(1 + p^{-1}) \tag{121}$$

Desta forma é claro que

$$\Theta_1(M; q_p; lr_p) \leq \Theta_1(q_1^{\delta_1 l r_1}; q_p; lr_p) \tag{122}$$

e

$$\begin{aligned} &\Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}) \\ &\qquad \leq \Theta_{p-1}(q_1^{\delta_1 l r_1}; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}). \end{aligned} \tag{123}$$

Temos, pelo Lema 21, que o termo à direita em (122) é limitado por

$$\frac{\log(q_1^{\delta_1 l r_1})}{lr_p \log q_p} \leq \frac{\delta_1 l r_1 \log q_1}{lr_1 \log q_1} = \delta_1,$$

onde usamos (119).

Queremos utilizar a hipótese de indução em (123). Vamos verificar que $\delta_1 < (p-1)^{-1}$, dado que $\delta_1 > \delta$, esta é a única hipótese cuja verificação não é trivial.

Dado que $\delta_1 = \delta(1+p^{-1})$ temos que

$$\delta\Big(1+\frac{1}{p}\Big) \leq \frac{1}{p}\Big(1+\frac{1}{p}\Big) = \frac{p+1}{p^2} \leq \frac{1}{p-1}.$$

Segue que

$$\Theta_{p-1}(q_1^{\delta_1 l r_1}; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}) < 10^{p-1}\delta_1^{(1/2)^{p-1}}.$$

Como $\delta_1 < 2\delta$, temos que

$$\begin{aligned}\Phi &< 2\delta + 2^{(1/2)^{p-1}} 10^{p-1}\delta^{(1/2)^{p-1}} \\ &\leq 2\delta + 2(10^{p-1}\delta^{(1/2)^{p-1}}) \\ &< 3(10^{p-1}\delta^{(1/2)^{p-1}}).\end{aligned}$$

Aplicando o Lema 22 obtemos

$$\Theta_p(q_1^{\delta r_1}; q_1, \ldots, q_p; r_1, \ldots, r_p) < 2\Big(3(10^{p-1}\delta^{(1/2)^{p-1}}) + 3^{1/2}10^{(p-1)/2}\delta^{(1/2)^p} + \delta^{1/2}\Big).$$

Vamos simplificar os três somandos à direita visando obter $10^p(\delta^{(1/2)^p})$. Utilizando que $0 < \delta < 1$ concluímos que

$$3(10^{p-1}\delta^{(1/2)^{p-1}}) \leq \frac{3}{10}10^p\delta^{(1/2)^p}$$

e que

$$\delta^{1/2} \leq \frac{1}{100}10^p\delta^{(1/2)^p}.$$

Por fim, como $p \geq 2$ temos que $10^{(p+1)/2} \geq 10^{3/2}$ e assim

$$3^{1/2}10^{(p-1)/2}\delta^{(1/2)^p} \leq \frac{3^{1/2}}{10^{3/2}}10^p\delta^{(1/2)^p}.$$

Juntando as últimas quatro desigualdades e usando que

$$\frac{3}{10} + \Big(\frac{3}{1000}\Big)^{1/2} + \frac{1}{100} < \frac{3}{10} + \Big(\frac{1}{100}\Big)^{1/2} + \frac{1}{100} < \frac{1}{2},$$

concluímos que

$$\begin{aligned}\Theta_p(q_1^{\delta r_1}; q_1, \ldots, q_p; r_1, \ldots, r_p) &< 2\Big(\frac{3}{10} + \frac{3^{1/2}}{10^{3/2}} + \frac{1}{10^2}\Big)10^p\delta^{(1/2)^p} \\ &< 10^p\delta^{(1/2)^p}.\end{aligned}$$

Como queríamos. □

2.5. Um Lema combinatório. Nosso próximo Lema nos diz o número máximo de soluções inteiras satisfazendo algumas desigualdades.

LEMA 24. *Se $r_1, \ldots, r_m$ são inteiros positivos e $\lambda > 0$, então o número de soluções inteiras $j_1, \ldots, j_m$ das desigualdades*

$$0 \le j_1 \le r_1, \ldots, 0 \le j_m \le r_m, \quad \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \le \frac{1}{2}(m - \lambda)$$

não excede

$$2m^{1/2}\lambda^{-1}(r_1 + 1)\ldots(r_m + 1).$$

DEMONSTRAÇÃO. Provaremos este resultado por indução. Para o caso $m = 1$ temos que o número de inteiros j_1 satisfazendo

$$0 \le j_1 \le r_1, \quad j_1 \le \frac{1}{2}(1 - \lambda)r_1$$

é no máximo $r_1 + 1$ se $\lambda \le 1$ e 0 se $\lambda > 1$.

Suponhamos que $m > 1$ e que já sabemos que o resultado é válido para $m-1$. Se $\lambda \le 2m^{1/2}$ não há o que fazer pois o número de soluções das primeiras desigualdades é $(r_1 + 1)\ldots(r_m + 1)$.

Suponhamos que $\lambda > 2m^{1/2}$. Fixando j_m temos que o número de soluções $j_1, \ldots, j_{m-1}$ de

$$0 \le j_1 \le r_1, \ldots, 0 \le j_{m-1} \le r_{m-1}, \quad \frac{j_1}{r_1} + \cdots + \frac{j_{m-1}}{r_{m-1}} \le \frac{1}{2}(m - 1 - \lambda')$$

não excede

$$2(m-1)^{1/2}\lambda'^{-1}(r_1 + 1)\ldots(r_{m-1} + 1)$$

onde $\lambda' = \lambda - 1 + 2j_m/r_m$.

Logo temos que o total de soluções $j_1, \ldots, j_m$ não excede

$$\sum_{j=0}^{r_m} 2(m-1)^{1/2}(\lambda - 1 + 2j_m/r_m)^{-1}(r_1 + 1)\ldots(r_{m-1} + 1).$$

Portanto, para completarmos o Lema, basta provar que

$$\sum_{j=0}^{r}(\lambda - 1 + 2j/r)^{-1} < \lambda^{-1}(m-1)^{-1/2}m^{1/2}(r + 1)$$

para inteiros positivos r e m com $m > 1$ e $\lambda > 2m^{1/2}$.

Caso r seja um número par temos

$$\begin{aligned}
\sum_{j=0}^{r}\Big(\lambda-\Big(\frac{r-2j}{r}\Big)\Big)^{-1} &= \sum_{j=-\frac{r}{2}}^{\frac{r}{2}}\Big(\lambda-\frac{2j}{r}\Big)^{-1}\\
&= \lambda^{-1}+\sum_{j=1}^{\frac{r}{2}}\Big(\Big(\lambda-\frac{2j}{r}\Big)^{-1}+\Big(\lambda+\frac{2j}{r}\Big)^{-1}\Big)\\
&= \lambda^{-1}+2\lambda\sum_{j=1}^{\frac{r}{2}}\Big(\lambda^2-\Big(\frac{2j}{r}\Big)^2\Big)^{-1}\\
&\leq \lambda^{-1}+2\lambda\sum_{j=1}^{\frac{r}{2}}(\lambda^2-1)^{-1}\\
&\leq \lambda^{-1}+r\lambda(\lambda^2-1)^{-1}\\
&\leq (r+1)\lambda^{-1}(1-\lambda^{-2})^{-1},
\end{aligned}$$

onde usamos $\lambda(\lambda^2-1)^{-1}=\lambda^{-1}(1-\lambda^{-2})^{-1}$.

Usando que $\lambda>2m^{1/2}$ temos que $1-\lambda^{-2}>1-\frac{1}{4}m^{-1}>(1-m^{-1})^{1/2}$, como queríamos.

Caso r seja um ímpar, temos

$$\begin{aligned}
\sum_{j=0}^{r}\Big(\lambda-\frac{r-2j}{r}\Big)^{-1} &= \sum_{j=0}^{\lfloor\frac{r}{2}\rfloor}\Big(\lambda-\frac{r-2j}{r}\Big)^{-1}+\sum_{j=\lceil\frac{r}{2}\rceil}^{r}\Big(\lambda-\frac{r-2j}{r}\Big)^{-1}\\
&= \sum_{j=0}^{\lfloor\frac{r}{2}\rfloor}\Big(\Big(\lambda-\frac{r-2j}{r}\Big)^{-1}+\Big(\lambda+\frac{r-2j}{r}\Big)^{-1}\Big)\\
&= \sum_{j=0}^{\lfloor\frac{r}{2}\rfloor}\frac{2\lambda}{\lambda^2-\Big(\frac{r-2j}{r}\Big)^2}\\
&= \Big(\Big\lfloor\frac{r}{2}\Big\rfloor+1\Big)2\lambda(\lambda^2-1)^{-1}\\
&= \Big(\frac{r}{2}-\frac{1}{2}+1\Big)2\lambda(\lambda^2-1)^{-1}\\
&\leq (r+1)\lambda(\lambda^2-1)^{-1}\\
&= (r+1)\lambda^{-1}(1-\lambda^{-2})^{-1}
\end{aligned}$$

e o resultado segue pelos mesmos cálculos usados no caso r par.

□

2.6. A existência do polinômio desejado. Tudo o que foi feito até aqui objetiva a demonstração do Lema 25. Este é o último Lema e o único que aparecerá na demonstração do Teorema de Roth. Precisamos aqui, da mesma forma que precisamos para o Teorema de Dyson, fixar várias constantes.

Seja α um número irracional algébrico e suponha que a desigualdade

$$\left|\alpha - \frac{h}{q}\right| \leq \frac{1}{q^k} \tag{124}$$

é satisfeita por uma infinidade de pares de inteiros h, q com $q > 0$ e $k > 2$.

Vimos no terceiro capítulo que basta provarmos o Teorema para inteiros algébricos. Visando simplificar nossos cálculos faremos esta redução.

Se α é um inteiro algébrico existe um polinômio

$$f(x) = x^n + a_1 x^{n-1} + \ldots + a_n, \tag{125}$$

com coeficientes inteiros, tal que $f(\alpha) = 0$. Seja

$$A = \max\{1, |a_1|, \ldots, |a_n|\}. \tag{126}$$

Vamos estabelecer algumas condições que devem ser satisfeitas por $m, \delta, q_1, h_1, \ldots, q_m, h_m, r_1, \ldots, r_m$ (que serão escolhidos nesta exata ordem). Nos preocuparemos em provar que, de fato, o conjunto de números que satisfazem estas condições é não-vazio apenas durante a demonstração do Teorema. Estas desigualdades serão hipóteses do Lema 25:

$$0 < \delta < m^{-1}, \tag{127}$$

$$10^m \delta^{(1/2)^m} + 2(1+3\delta)nm^{1/2} < \frac{1}{2}m, \tag{128}$$

$$r_m > 10\delta^{-1}, \quad r_{j-1}/r_j > \delta^{-1} \text{ para } j = 2, \ldots, m, \tag{129}$$

$$\delta^2 \log q_1 > 2m + 1 + 2m\log(1+A) + 2m\log(1+|\alpha|), \tag{130}$$

$$r_j \log q_j \geq r_1 \log q_1. \tag{131}$$

Notemos que essas desigualdades implicam as condições do Lema 23, pois (127) e (130) implicam que $\delta \log q_1 > m(2m+1)$. Sempre que formos usar (130) estaremos, na verdade, usando uma das seguintes desigualdades

$$\begin{cases} \delta^2 \log q_1 > 2m + 1; \\ \delta^2 \log q_1 > 2m \log(1+A); \\ \delta^2 \log q_1 > 2m \log(1+|\alpha|). \end{cases}$$

Definimos $\lambda, \gamma, \eta, B_1$ por

$$\lambda = 4(1+3\delta)nm^{1/2}, \tag{132}$$

$$\gamma = \frac{1}{2}(m-\lambda) \tag{133}$$

$$\eta = 10^m \delta^{(1/2)^m}, \tag{134}$$

$$B_1 = \lfloor q_1^{\delta r_1} \rfloor. \tag{135}$$

Notemos que (128) é equivalente à

$$\eta < \gamma. \tag{136}$$

Notemos que B_1 é necessariamente grande, pois $r_1 > 10$ e $q_1^{\delta^2} > e^{2m+1} \geq e^3$. Logo vale $q_1^{\frac{1}{2}\delta r_1} < B_1$ que implica

$$q_1^{\frac{1}{2}\delta^2 r_1} < B_1^{\delta.} \tag{137}$$

LEMA 25. *Suponhamos que as condições 127 - 131 estão satisfeitas e também que $h_1, \ldots, h_m$ são inteiros relativamente primos com $q_1, \ldots, q_m$, respectivamente. Então existe um polinômio $Q(x_1, \ldots, x_m)$ com coeficientes inteiros, de grau no máximo r_j em x_j para $j = 1, \ldots, m$, tal que*

(i) *o índice de Q no ponto $(\alpha, \ldots, \alpha)$ relativo à $r_1, \ldots, r_m$ é pelo menos $\gamma - \eta$;*[1]

(ii) $Q(h_1/q_1, \ldots, h_m/q_m) \neq 0$;

(iii) *para todas as derivadas*

$$Q_{i_1,\ldots,i_m}(x_1, \ldots, x_m) = \frac{1}{i_1! \ldots i_m!} \Big(\frac{\partial}{\partial x_1}\Big)^{i_1} \ldots \Big(\frac{\partial}{\partial x_m}\Big)^{i_m} Q,$$

temos que

$$|Q_{i_1,\ldots,i_m}(\alpha, \ldots, \alpha)| < B_1^{1+3\delta}. \tag{138}$$

DEMONSTRAÇÃO. Consideremos P o conjunto de todos os polinômios $W(x_1, \ldots, x_m)$ da forma

$$W(x_1, \ldots, x_m) = \sum_{s_1=0}^{r_1} \ldots \sum_{s_m=0}^{r_m} c(s_1, \ldots, s_m) x_1^{s_1} \ldots x_m^{s_m}, \tag{139}$$

onde os coeficientes $c(s_1, \ldots, s_m)$ são inteiros positivos satisfazendo

$$0 \leq c(s_1, \ldots, s_m) \leq B_1. \tag{140}$$

O número de polinômios W em P é

$$N = (B_1 + 1)^r, \tag{141}$$

[1]Notemos que $\gamma - \eta > 0$ por (136), caso contrário, este item não nos daria informação alguma.

onde

$$r = (r_1 + 1) \dots (r_m + 1). \tag{142}$$

Consideremos todas as m-uplas $j_1, \dots, j_m$ satisfazendo

$$0 \leq j_1 \leq r_1, \dots, 0 \leq j_m \leq r_m, \quad \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \leq \gamma. \tag{143}$$

Denotamos por N_0 o número de $j_1, \dots, j_m$ satisfazendo (143). Pelo Lema 24 e por (133), concluímos que N_0 é limitado por

$$2m^{1/2}\lambda^{-1}r, \tag{144}$$

onde r foi definido em (142).

Seja

$$\phi : \{1, \dots, N_0\} \longrightarrow \Big\{(j_1, \dots, j_m), 0 \leq j_\ell \leq r_\ell, \forall\ 1 \leq \ell \leq m,\ \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \leq \gamma\Big\}$$

uma bijeção.

Para cada polinômio W associamos uma N_0-upla de polinômios $D(W)$ onde a ℓ-ésima coordenada é dada por

$$W_{\phi(\ell)}(x_1, \dots, x_m) = \frac{1}{\phi_1(\ell)! \dots \phi_m(\ell)!} \Big(\frac{\partial}{\partial x_1}\Big)^{\phi_1(\ell)} \cdots \Big(\frac{\partial}{\partial x_m}\Big)^{\phi_m(\ell)} W, \tag{145}$$

onde $\phi_i(\ell)$ é a i-ésima coordenada da função ϕ.

Logo, as coordenadas de $D(W)$ são as derivadas

$$W_{j_1,\dots,j_m}(x_1, \dots, x_m),$$

onde $j_1, \dots, j_m$ satisfazem (143). Para cada N_0-upla de polinômios $D(W)$ associamos uma nova N_0-upla de polinômios $T(W)$ onde a ℓ-ésima coordenada de $T(W)$ é dada pelo resto da divisão de $W_{\phi(\ell)}(x, \dots, x)$ por $f(x)$, onde $W_{\phi(\ell)}$ é como em (145).

Denotamos por

$$T_{j_1,\dots,j_m}(W; x),$$

o resto da divisão de $W_{j_1,\dots,j_m}(x, \dots, x)$ por $f(x)$.

Nosso interesse agora é estimar o maior valor absoluto dos coeficientes de todos os restos $T_{j_1,\dots,j_m}(W; x)$. Os coeficientes de cada uma das derivadas $W_{j_1,\dots,j_m}(x_1, \dots, x_m)$ têm valores absolutos não excedendo

$$2^{r_1+\dots+r_m} B_1,$$

pois cada coeficiente de $W_{j_1,\dots,j_m}(x_1, \dots, x_m)$ é da forma

$$Bc(i_1, \dots, i_m),$$

onde

$$B = \begin{cases} \binom{i_1}{j_1} \dots \binom{i_m}{j_m} & \text{se } i_\ell \geq j_\ell, \quad \forall\ 1 \leq \ell \leq m; \\ 0 & \text{caso contrário.} \end{cases}$$

Quando $x_1, \ldots, x_m$ são substituídos por x temos a soma dos coeficientes associados a monômios de mesmo grau. Assim, os coeficientes em $W_{j_1,\ldots,j_m}(x,\ldots,x)$ têm valores absolutos menores que $r2^{r_1+\cdots+r_m}B_1$. Nos resta considerar a operação de divisão do polinômio

$$W_{j_1,\ldots,j_m}(x,\ldots,x) = w_s x^s + w_{s-1}x^{s-1} + \cdots + w_0,$$

por $f(x)$, o polinômio minimal de α dado em (125). Supondo que $s \geq n$ temos que nossa primeira operação é subtrair $w_s x^{s-n} f(x)$ (note que aqui estamos usando a redução para o caso em $f(x)$ é um polinômio mônico) e obtemos um novo polinômio cujos coeficientes são ou da forma $w' - aw''$ ou da forma w, onde w, w', w'' são coeficientes de $W_{j_1,\ldots,j_m}$ e a um dos coeficientes de f. Logo, os coeficientes do novo polinômio têm valores absolutos menores que $|w'| + |aw''| < r2^{r_1+\cdots+r_m}(1+A)B_1$, com A definido como em (126). Repetindo o mesmo raciocínio temos que os coeficientes do resto $T_{j_1,\ldots,j_m}(W;x)$ têm valores absolutos menores que

$$M = r2^{r_1+\cdots+r_m}(1+A)^{s-n+1}B_1.$$

Seja

$$T = \bigcup_{V \in P} T(V).$$

Temos que T tem no máximo $(2M+1)^{nN_0}$ elementos.

Podemos estimar $2M+1$ por:

$$2M + 1 \leq r2^{r_1+\cdots+r_m+1}(1+A)^{s-n+1}(B_1+1).$$

Vamos estimar o lado direito de tal forma a obter uma potência de $B_1 + 1$.

Usando que $r_1 > r_2 > \cdots > r_m$ temos

$$2^{r_1+\cdots+r_m+1} \leq 2^{mr_1},$$

mas também temos, por (130), que $mr_1 \log 2 < mr_1 \log(1+A) < \delta^2 r_1 \log q_1$ e isto implica que $2^{mr_1} < q_1^{\delta^2 r_1} < (B_1+1)^{\delta}$.

Também temos que

$$r = (r_1+1)\ldots(r_m+1) \leq 2^{r_1+\ldots r_m} \leq 2^{mr_1} < (B_1+1)^{\delta}. \qquad (146)$$

Como $s \leq r_1 + \cdots + r_m \leq mr_1$, obtemos que

$$(1+A)^{s-n+1} \leq (1+A)^{mr_1}$$

por (130) temos que

$$m \log(1+A) < \delta^2 \log q_1$$

donde temos que $(1+A)^{mr_1} < q_1^{\delta^2 r_1} < (B_1+1)^{\delta}$.

Concluímos que

$$2M + 1 < (B_1+1)^{1+3\delta}.$$

Portanto T tem no máximo

$$(B_1+1)^{(1+3\delta)nN_0}$$

elementos.

Por (144) e da definição de λ em (132), temos

$$(1+3\delta)nN_0 \leq 2(1+3\delta)nm^{1/2}\lambda^{-1}r = \frac{1}{2}r < r,$$

portanto

$$(1+2M)^{nN_0} < (1+B_1)^r = N.$$

Vemos que o número de elementos em T é estritamente menor que o número de polinômios W em P. Logo, temos que existem polinômios W^1 e W^2 da forma (138) tais que

$$W^1_{j_1,\ldots,j_m}(x,\ldots,x) - W^2_{j_1,\ldots,j_m}(x,\ldots,x)$$

é divisível por $f(x)$ para todo $j_1,\ldots,j_m$ satisfazendo (143). Denotando $W^* = W^1 - W^2$, temos que todas as derivadas de

$$W^*_{j_1,\ldots,j_m}(x_1,\ldots,x_m)$$

são nulas quando $x_1 = \cdots = x_m = \alpha$. Pela definição de T o índice de W^* no ponto $(\alpha,\ldots,\alpha)$ relativo à $r_1,\ldots,r_m$ é pelo menos γ. Também os coeficientes de W^* são inteiros, não todos nulos e têm valores absolutos não excedendo B_1.

Vamos aplicar o Lema 23. O polinômio $W^*(x_1,\ldots,x_m)$ pertence à

$$\mathcal{R}_m(q_1^{\delta r_1}; r_1,\ldots,r_m).$$

Pelo Lema 23, seu índice em $(h_1/q_1,\ldots,h_m/q_m)$ relativo à $r_1,\ldots,r_m$ é menor que η, definido em (134). Logo W^* possui alguma derivada

$$Q(x_1,\ldots,x_m) = \frac{1}{k_1!\ldots k_m!}\Big(\frac{\partial}{\partial x_1}\Big)^{k_1}\ldots\Big(\frac{\partial}{\partial x_m}\Big)^{k_m}W^*(x_1,\ldots,x_m),$$

com

$$\frac{k_1}{r_1}+\cdots+\frac{k_m}{r_m} < \eta,$$

tal que

$$Q(h_1/q_1,\ldots,h_m/q_m) \neq 0.$$

O índice de Q no ponto $(\alpha,\ldots,\alpha)$ relativo à $r_1,\ldots,r_m$ é pelo menos $\gamma-\eta$. Por (136) temos que $\gamma-\eta>0$ e, portanto, Q possui índice positivo em $(\alpha,\ldots,\alpha)$ relativo à $r_1,\ldots,r_m$. Logo, Q tem as propriedades (i) e (ii) do enunciado.

Como os coeficientes de W^* têm valores absolutos no máximo B_1 segue que os coeficientes de Q têm valores absolutos no máximo

$$2^{r_1+\cdots+r_m}B_1 \leq 2^{mr_1}B_1 < B_1^{1+\delta},$$

onde usamos que

$$2m \log 2 \leq 2m \log(1+A) \leq \delta^2 \log q_1 \tag{147}$$

e a observação (137).

Portanto, os coeficientes de qualquer uma das derivadas de

$$Q_{i_1,\ldots,i_m}(x_1,\ldots,x_m)$$

têm valores absolutos menores que $2^{mr_1} B_1^{1+\delta} < B_1^{1+2\delta}$. Assim segue que

$$\begin{aligned} |Q_{i_1,\ldots,i_m}(x_1,\ldots,x_m)| &\leq B_1^{1+2\delta} \sum_{s_1=0}^{r_1} \cdots \sum_{s_m=0}^{r_m} |x_1^{s_1} \ldots x_m^{s_m}| \\ &= B_1^{1+2\delta} (1+|x_1|)^{r_1} \ldots (1+|x_m|)^{r_m} \end{aligned}$$

donde concluímos que

$$|Q_{i_1,\ldots,i_m}(\alpha,\ldots,\alpha)| < B_1^{1+2\delta}(1+|\alpha|)^{r_1+\cdots+r_m},$$

e isso implica (iii) já que, por (130), vale

$$2m \log(1+|\alpha|) < \delta^2 \log q_1,$$

portanto, por (137), vale

$$(1+|\alpha|)^{mr_1} < q_1^{\delta^2 r_1/2} < B_1^{\delta}.$$

□

2.7. Teorema de Roth.

TEOREMA 14 (Roth). *Seja α um número irracional algébrico. Temos que*

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^k},$$

tem infinitas soluções somente se $k \leq 2$.

DEMONSTRAÇÃO. Suponhamos que α é um inteiro algébrico, que $k > 2$ e que

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^k} \tag{148}$$

tem infinita soluções. Nosso objetivo é provar que as infinitas soluções de (148) nos permitem escolher $m, \delta, q_1,\ h_1, \ldots, q_m, h_m, r_1, \ldots, r_m$ nas condições do Lema 25.

Seja n o grau de α. Vamos escolher m satisfazendo $m > 4nm^{1/2}$ e também

$$\frac{2m}{m - 4nm^{1/2}} = \frac{2}{1 - 4nm^{-1/2}} < k, \tag{149}$$

o que só é possível pois $k > 2$. Para um δ suficientemente pequeno temos que

$$m - 4(1+3\delta)nm^{1/2} - 2\eta > 0, \tag{150}$$

onde η é dado por (134). Essa condição é exatamente (128). Escolhemos δ satisfazendo (150), (127) e também satisfazendo

$$\frac{2m(1+4\delta)}{m - 4(1+3\delta)nm^{1/2} - 2\eta} < k, \tag{151}$$

o que é possível graças a (149) e pelo fato que η é pequeno quando δ é pequeno.

A desigualdade (151) é equivalente a

$$\frac{m(1+4\delta)}{\gamma - \eta} < k, \tag{152}$$

por (132) e (133). Vale notar que até aqui não usamos que (148) tem infinitas soluções. Nesta demonstração usaremos as infinitas soluções para conseguir a desigualdade contrária a (152).

Escolhemos h_1/q_1 solução de (148) com $(h_1, q_1) = 1$ e com q_1 suficientemente grande para satisfazer (130) e prosseguimos escolhendo soluções $h_2/q_2, \ldots, h_m/q_m$ com $(h_i, q_i) = 1$ satisfazendo

$$\frac{\log q_j}{\log q_{j-1}} > \frac{2}{\delta} \quad (j = 2, \ldots, m). \tag{153}$$

Agora escolhemos um inteiro r_1 satisfazendo

$$r_1 > \frac{10 \log q_m}{\delta \log q_1}, \tag{154}$$

e também $r_2, \ldots, r_m$ satisfazendo

$$\frac{r_1 \log q_1}{\log q_j} \leq r_j < 1 + \frac{r_1 \log q_1}{\log q_j} \quad (j = 2, \ldots, m). \tag{155}$$

Assim a condição (131) está satisfeita.

A primeira desigualdade em (155) e a desigualdade em (154) implicam que

$$r_m \geq \frac{r_1 \log q_1}{\log q_m} > 10\delta^{-1}$$

e por (155) temos

$$\frac{r_{j-1}}{r_j} > \frac{r_1 \log q_1}{\log q_{j-1}} \Big(1 + \frac{r_1 \log q_1}{\log q_j}\Big)^{-1} = \frac{\log q_j}{\log q_{j-1}} \Big(\frac{\log q_j}{r_1 \log q_1} + 1\Big)^{-1}.$$

Usando novamente (155) conseguimos

$$\begin{aligned}\frac{r_j \log q_j}{r_1 \log q_1} &< \Big(1 + \frac{r_1 \log q_1}{\log q_j}\Big)\frac{\log q_j}{r_1 \log q_1} \\ &= \frac{\log q_j}{r_1 \log q_1} + 1 \leq \frac{\log q_m}{r_1 \log q_1} + 1 < \frac{1}{10}\delta + 1. \qquad (156)\end{aligned}$$

Por (156), temos que

$$\Big(\frac{\log q_j}{r_1 \log q_1} + 1\Big)^{-1} > \Big(1 + \frac{1}{10}\delta\Big)^{-1} > \frac{1}{2},$$

e por (153) concluímos que

$$\frac{r_{j-1}}{r_j} > \frac{\log q_j}{\log q_{j-1}}\Big(\frac{\log q_j}{r_1 \log q_1} + 1\Big)^{-1} > \frac{2}{\delta}\frac{1}{2} = \delta^{-1}.$$

Notemos que todas as condições do Lema 25 estão satisfeitas, e, portanto, existe um polinômio $Q(x_1, \ldots, x_m)$ com as propriedades (i), (ii) e (iii) do Lema 25. Como Q tem coeficientes inteiros e não se anula em $(h_1/q_1, \ldots, h_m/q_m)$ temos que

$$\begin{aligned}|Q(h_1/q_1, \ldots, h_m/q_m)| &= \Big|\sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} c(j_1, \ldots, j_m)\frac{h_1^{j_1}}{q_1^{j_1}} \ldots \frac{h_m^{j_m}}{q_m^{j_m}}\Big| \\ &> \frac{1}{q_1^{r_1} \ldots q_m^{r_m}}.\end{aligned}$$

Usando a desigualdade (156),

$$\frac{r_j \log q_j}{r_1 \log q_1} < \frac{1}{10}\delta + 1,$$

conseguimos

$$q_j^{r_j} < q_1^{r_1(1+\delta/10)} < q_1^{r_1(1+\delta)},$$

para todo $j = 1, \ldots, m$. Disto concluímos que

$$|Q(h_1/q_1, \ldots, h_m/q_m)| > q_1^{-mr_1(1+\delta)}. \qquad (157)$$

Entretanto, expandindo Q em polinômio de Taylor em torno de $(\alpha, \ldots, \alpha)$ temos

$$Q\Big(\frac{h_1}{q_1}, \ldots, \frac{h_m}{q_m}\Big) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} Q_{j_1,\ldots,j_m}(\alpha, \ldots, \alpha)\Big(\frac{h_1}{q_1} - \alpha\Big)^{i_1} \ldots \Big(\frac{h_m}{q_m} - \alpha\Big)^{i_m},$$

e sabemos pela propriedade (i) do Lema 25 que

$$Q_{j_1,\ldots,j_m}(\alpha, \ldots, \alpha) = 0$$

para todo $j_1, \dots, j_m$ satisfazendo

$$\frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} < \gamma - \eta.$$

Como (155) implica que $q_j \geq q_1^{r_1/r_j}$ temos que

$$\frac{1}{q_1^{j_1} \dots q_m^{j_m}} \leq q_1^{-r_1(\frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m})},$$

logo para todo $j_1, \dots, j_m$ tal que

$$\frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \geq \gamma - \eta$$

vale que

$$\left|\Big(\frac{h_1}{q_1} - \alpha\Big)^{j_1} \dots \Big(\frac{h_m}{q_m} - \alpha\Big)^{j_m}\right| < \frac{1}{(q_1^{j_1} \dots q_m^{j_m})^k} \leq q_1^{-r_1(\gamma-\eta)k}.$$

Agora usando (iii) do Lema 25 temos

$$\begin{aligned}\left|Q\Big(\frac{h_1}{q_1}, \dots, \frac{h_m}{q_m}\Big)\right| &= \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} |Q_{j_1,\dots,j_m}(\alpha, \dots, \alpha)| \left|\frac{h_1}{q_1} - \alpha\right|^{i_1} \dots \left|\frac{h_m}{q_m} - \alpha\right|^{i_m} \\ &< (r_1 + 1) \dots (r_m + 1) B_1^{1+3\delta} q_1^{-r_1(\gamma-\eta)k}.\end{aligned}$$

Lembrando de (137) e de (147) temos que

$$(r_1 + 1) \dots (r_m + 1) \leq 2^{mr_1} \leq B_1^{\delta},$$

portanto, vale que

$$\begin{aligned}\left|Q\Big(\frac{h_1}{q_1}, \dots, \frac{h_m}{q_m}\Big)\right| &< B_1^{1+4\delta} q_1^{-r_1(\gamma-\eta)k} \\ &< q_1^{(1+4\delta)\delta r_1 - r_1(\gamma-\eta)k}.,\end{aligned}$$

e esta desigualdade junto com (157) nos garante que

$$-mr_1(1+\delta) < (1+4\delta)\delta r_1 - r_1(\gamma-\eta)k,$$

ou, equivalentemente,

$$k < \frac{m(1+\delta) + \delta(1+4\delta)}{\gamma - \eta}.$$

Como $\delta(1+4\delta) < 3\delta m$ temos que

$$k < \frac{m(1+4\delta)}{\gamma - \eta},$$

contrariando a desigualdade (152). Logo (148) não pode ter infinitas soluções quando $k > 2$.

□

CAPíTULO 5

Aplicações e comentários.

1. Aplicações

Já apresentamos algumas aplicações quando falamos de números de Liouville e quando provamos a finitude de soluções para equações de Thue.

Aplicação 1. O nosso próximo resultado é uma aplicação do Teorema de Roth para polinômios homogêneos, isto é, polinômios tais que seus monômios de coeficientes não-nulos têm o mesmo grau. Encontramos essa aplicação no livro de Lequain [**11**].

PROPOSIÇÃO 7. *Seja $f(x,y) \in \mathbb{Q}[x,y]$ um polinômio homogêneo de grau d sem fatores múltiplos.*

(a) *Sejam $\epsilon > 0$ e $A > 0$. Se $d \geq 2$, então o sistema*

$$0 < |f(x,y)| < A \max\{|x|,|y|\}^{d-2-\epsilon}$$

tem somente um número finito de soluções em $\mathbb{Z}^2$.

(b) *Se $d \geq 3$ e se $g(x,y) \in \mathbb{C}[x,y]$ tem grau menor ou igual a $d-3$. Então, o sistema*

$$\begin{cases} f(x,y) = g(x,y) \\ f(x,y) \neq 0 \end{cases}$$

tem somente um número finito de soluções em $\mathbb{Z}^2$.

DEMONSTRAÇÃO. (a) Suponhamos que o conjunto

$$S = \left\{(x,y) \in \mathbb{Z}^2 : 0 < |f(x,y)| < A \max\{|x|,|y|\}^{d-2-\epsilon}\right\}$$

seja infinito. Podemos supor, sem perda de generalidade, que S possui um número infinito de elementos tais que $|x| \leq |y|$. Assim o conjunto

$$S_1 = \{(x,y) \in \mathbb{Z}^2 : \ 0 < |f(x,y)| < A|y|^{d-2-\epsilon}\} \tag{158}$$

é infinito.

Como $f(x,y)$ é homogêneo de grau d e não tem fatores múltiplos concluímos que ou $f(x,y)$ é da forma

$$f(x,y) = \beta(x - \alpha_1 y) \cdots (x - \alpha_d y)$$

ou da forma

$$f(x,y) = \beta y(x-\alpha_1 y)\cdots(x-\alpha_{d-1}y)$$

onde $\beta, \alpha_1, \ldots, \alpha_d \in \mathbb{C}, \beta \neq 0, \alpha_i \neq \alpha_j$ se $i \neq j$.

No último caso temos que

$$f_1(x,y) = f(x,y)/y = \beta(x-\alpha_1 y)\cdots(x-\alpha_{d-1}y).$$

Neste caso segue que

$$S_1 = \{(x,y) \in \mathbb{Z}^2 : 0 < |f_1(x,y)| < A|y|^{d-3-\epsilon}\}.$$

Portanto, em qualquer um dos casos, podemos supor que

$$f(x,y) = \beta(x-\alpha_1 y)\cdots(x-\alpha_r y)$$

onde $r = d$ ou $r = d-1$.

Como S_1 é infinito, então existe α_i, digamos $i = 1$, tal que o conjunto

$$S_2 = \{(x,y) \in S_1; |x-\alpha_1 y| \leq |x-\alpha_i y|, i = 1, \ldots, r\}$$

é infinito.

Para todo $i = 2, \ldots, r$ e todo $(x,y) \in S_2$, temos:

$$\begin{aligned} |x-\alpha_i y| &\geq \frac{1}{2}(|x-\alpha_i y| + |x-\alpha_1 y|) \\ &\geq \frac{1}{2}|\alpha_1 y - \alpha_i y| \\ &= \frac{1}{2}|y||\alpha_1 - \alpha_i|. \end{aligned}$$

Visando estimar o valor absoluto de f calculado em elementos de S_2 escrevemos

$$|f(x,y)| = |\beta||x-\alpha_1 y| \prod_{i=2}^{r} |x-\alpha_i y|,$$

e com isso obtemos

$$|f(x,y)| \geq b|x-\alpha_1 y||y|^{r-1} \ \forall (x,y) \in S_2,$$

onde $b = |\beta/2^{r-1}| \prod_{i=2}^{r} |\alpha_i - \alpha_i| > 0$.

Lembramos que $S_2 \subset S_1$ e, portanto, temos

$$A|y|^{r-2-\epsilon} > |x-\alpha_1 y| b|y|^{r-1} > 0 \quad \forall (x,y) \in S_2.$$

Como b e y são diferentes de zero podemos dividir por $b|y|^r$ para obtermos

$$\frac{A/b}{|y|^{2+\epsilon}} > \left|\frac{x}{y} - \alpha_1\right| > 0,$$

para todo $(x,y) \in S_2$. Como S_2 é infinito concluímos pela proposição 5 que α_1 é um irracional. Pelo Teorema de Roth, temos que a desigualdade

$$\frac{A/b}{|y|^{2+\epsilon}} > \left|\frac{x}{y} - \alpha_1\right|$$

possui um número finito de soluções. Contrariando a hipótese de S ser infinito.

(b) Agora seja $g(x,y) \in \mathbb{C}[x,y]$ um polinômio de grau s. Podemos escrever g na forma

$$g(x,y) = \sum_{i+j\leq s} b_{ij}x^i y^j.$$

Temos, para todo $(x,y) \in \mathbb{Z}^2$, que

$$|g(x,y)| \leq \sum_{i+j\leq s} |b_{ij}x^i y^j| \leq A\max\{|x|,|y|\}^s$$

com $A = \sum |b_{ij}|$.

O conjunto dos pares $(x,y) \in \mathbb{Z}^2$ tais que $f(x,y) = g(x,y)$ e tais que $f(x,y) \neq 0$ está contido no conjunto dos pares $(x,y) \in \mathbb{Z}^2$ tais que $0 < |f(x,y)| = |g(x,y)| < A\max\{|x|,|y|\}^s$. Portanto, temos, por (a), que quando $s \leq d-3$, existe apenas um número finito de soluções tais que $f(x,y) = g(x,y)$. □

Aplicação 2. Nossa próxima aplicação está presente no livro de A.N. Parshin e I.R. Shafarevich [**19**].

PROPOSIÇÃO 8. *Seja $\{p_1,\ldots,p_n\}$ um conjunto formado por n números primos. Seja $P = \{p_1^{i_1}p_2^{i_2}\ldots p_n^{i_n}$ para todo $(i_1,\ldots,i_n) \in \mathbb{N}^n\}$. Então a equação*

$$x - y = c,\ c \in \mathbb{Z}, c \neq 0, \tag{159}$$

tem apenas um número finito de soluções com $x,y \in P$.

DEMONSTRAÇÃO. Suponhamos que x,y seja uma solução de (159), com

$$x = p_1^{i_1}\ldots p_n^{i_n}, \quad y = p_1^{j_1}\ldots p_n^{j_n}, \quad i_k, j_k \in \mathbb{N}.$$

Escrevemos

$$i_k = 3a_k + u_k, \quad j_k = 3b_j + w_k, \quad a_k,u_k,b_k,w_k \in \mathbb{N}, u_k,w_k \leq 2,$$

$$z = p_1^{a_1}\cdots p_m^{a_m}, \quad t = p_1^{b_1}\cdots p_m^{b_m},$$

então

$$Az^3 - Bt^3 = c,$$

onde $A = p_1^{u_1}\ldots p_m^{u_m}$ e $B = p_1^{w_1}\ldots p_m^{w_m}$. Seja

$$S = \{p_1^{i_1}\ldots p_n^{i_n},\ \text{onde } i_1,\ldots,i_n \in \{0,1,2\}\}.$$

Logo cada solução de $x - y = c$ é solução de uma equação do tipo $Az^3 - Bt^3 = c$, onde $A, B \in S$. Como S é um conjunto finito, temos que se $x - y = c$ tem um número infinito de soluções, então, uma equação da forma $Az^3 - Bt^3 = c$ tem infinitas soluções. Como vimos no Teorema 10, estas equações têm um número finito de soluções. Portanto, as equações do tipo (159) têm um número finito de soluções em P. □

Aplicação 3. Nossa terceira aplicação foi encontrada em **[20]**.
Fixemos $r \in \mathbb{R}$, com $r > 2$. Definimos uma função f_r por

$$f_r(x) = \begin{cases} \frac{1}{q^r} & \text{se } x \text{ é uma fração irredutível } \frac{p}{q} \text{ com } p \neq 0. \\ 0 & \text{se } x = 0 \text{ ou } x \text{ é irracional.} \end{cases}$$

Nosso interesse será analisar em quais pontos essa função é diferenciável. Vamos verificar inicialmente quais são os pontos em que essa função é contínua.

PROPOSIÇÃO 9. *A função f_r é contínua no zero e em todos os irracionais, mas é descontínua em todos os racionais não-nulos.*

DEMONSTRAÇÃO. Dado um racional não-nulo p/q, existe uma sequência x_n formada por irracionais que converge para p/q, mas $f(p/q) = 1/q^r$ e $f(x_n) = 0$ provando que f_r não é contínua em p/q. Agora suponhamos que α é 0 ou um irracional. Dado $\epsilon > 0$ escolhemos n tal que $0 < 1/n^r < \epsilon$. Pela proposição 4 existe $\delta > 0$ tal que

$$\left|\alpha - \frac{p}{q}\right| < \delta$$

somente se $n < q$. Assim, para toda sequência $(x_n)_n$ convergindo para α temos que se

$$|\alpha - x_n| < \delta$$

então

$$|f(\alpha) - f(x_n)| < \frac{1}{q^r} < \frac{1}{n^r} < \epsilon.$$

□

Vamos definir agora uma função conhecida como medida de irracionalidade. Dado $\alpha \in \mathbb{R}$, definimos $\mu(\alpha)$ como o menor número tal que se $\delta > \mu(\alpha)$ então

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^\delta}$$

tem apenas um número finito de soluções, onde p/q é uma fração reduzida. Notemos que o conjunto dos números de Liouville é precisamente o conjunto

$$\{\alpha \in \mathbb{R} : \mu(\alpha) = \infty\}.$$

Já vimos na proposição 5 que $\mu(\alpha) = 1$ se α é racional e pelo Teorema de Roth vimos que $\mu(\alpha) = 2$ para α irracional algébrico.

Vamos provar que o número a maioria dos números reais tem medida de irracionalidade 2.

PROPOSIÇÃO 10. *Seja*

$$T = \{\alpha \in \mathbb{R} : \mu(\alpha) > 2\}.$$

Então T tem medida nula.

DEMONSTRAÇÃO. Seja $T_0 = T \cap [0,1]$. Como a medida de T_0 é igual a medida de $T \cap [n, n+1]$ para todo $n \in \mathbb{Z}$. Temos que basta provar que T_0 tem medida nula.

Definimos os seguintes conjuntos

$$T_n = \Big\{\alpha \in [0,1] : \mu(\alpha) > 2 + \frac{1}{n}\Big\}.$$

Temos que

$$T_0 = \bigcup_{n=1}^{\infty} T_n.$$

Definimos também os seguintes conjuntos

$$I_{p,q} = \Big(\frac{p}{q} - \frac{1}{q^{2+\frac{1}{n}}}, \frac{p}{q} + \frac{1}{q^{2+\frac{1}{n}}}\Big)$$

onde p, q são inteiros tais que $0 < p < q$ e também

$$I_{0,q} = \Big(0, \frac{1}{q^{2+\frac{1}{n}}}\Big) \text{ e } I_{q,q} = \Big(1 - \frac{1}{q^{2+\frac{1}{n}}}, 1\Big).$$

Dado $N \in \mathbb{N}$ temos que

$$T_n \subset \bigcup_{q=N}^{\infty} \bigcup_{p=0}^{q} I_{p,q},$$

pois se $\alpha \in T_n$ então

$$\Big|\alpha - \frac{p}{q}\Big| < \frac{1}{q^{2+\frac{1}{n}}}$$

tem solução para $q \geq N$. Mas sabemos que

$$\begin{aligned}\mu\Big(\bigcup_{q=N}^{\infty} \bigcup_{p=0}^{q} I_{p,q}\Big) &\leq \sum_{q=N}^{\infty} \sum_{p=0}^{q} \mu(I_{p,q}) \\ &= \sum_{q=N}^{\infty} \frac{2}{q^{1+\frac{1}{n}}}.\end{aligned}$$

Sabemos que $\sum q^{-1-\frac{1}{n}}$ converge. Logo, existe N tal que

$$\sum_{q=N}^{\infty} \frac{2}{q^{1+\frac{1}{n}}} < \epsilon.$$

Consequentemente, temos que T_n tem medida nula. Como T_0 é união enumerável de conjuntos de medida nula temos que a medida de T_0 é nula. Como

$$T = \bigcup_{n\in\mathbb{Z}} \Big(T \cap [n, n+1]\Big)$$

temos que T tem medida nula. □

Uma consequência disto é que o conjunto dos número de Liouville tem medida nula. Como enuncíamos na proposição 6.

PROPOSIÇÃO 11. *Seja $r > 2$. Definimos*

$$S = \{\alpha \in \mathbb{R} : \alpha \text{ é irracional e } f_r \text{ não é diferenciável em } \alpha\}$$

e

$$T = \{\alpha \in \mathbb{R} : \mu(\alpha) > 2\}.$$

Então $S \subset T$.

DEMONSTRAÇÃO. Seja S_n o conjunto dos α irracionais tais que, para todo $\delta > 0$, existe $x \in (\alpha - \delta, \alpha + \delta)$ tal que

$$\Big|\frac{f_r(x) - f_r(\alpha)}{x - \alpha}\Big| = \Big|\frac{f_r(x)}{x - \alpha}\Big| > \frac{1}{n}.$$

Temos que

$$S = \bigcup_{n=1}^{\infty} S_n.$$

Pelo definição de S_n dado δ existem infinitos racionais $p/q \in (\alpha - \delta, \alpha + \delta)$ satisfazendo

$$\left|\frac{f_r\left(\frac{p}{q}\right) - f_r(\alpha)}{\frac{p}{q} - \alpha}\right| = \frac{1}{q^r}\frac{1}{\left|\alpha - \frac{p}{q}\right|} > \frac{1}{n}.$$

Desta forma, conseguimos provar que

$$\left|\alpha - \frac{p}{q}\right| < \frac{n}{q^r}$$

tem infinitas soluções. Portanto, $\mu(\alpha) > 2$ e $S_n \subset T$, por fim, $S \subset T$. □

Temos, pela proposição 10, que S tem medida nula. Temos pela proposição 11 e pelo Teorema de Roth a seguinte proposição:

PROPOSIÇÃO 12. *Temos que f_r é diferenciável em todos os números algébricos de grau $n \geq 2$ e descontínua em todos os racionais não-nulos.*

2. Comentários

A demonstração do Teorema de Roth que apresentamos segue a demonstração original apresentada no artigo de Roth [**1**]. Uma outra demonstração foi obtida por Esnault e Viehweg, em 1984. Esta demonstração aparece no artigo [**16**], cujo título é "Dyson's Lemma for polynomials in several variables (and the Theorem of Roth).". Como o próprio título já diz eles generalizaram o Lema de Dyson para várias variáveis e com isso provaram o Teorema de Roth. Uma outra demonstração foi obtida por Corvaja [**17**] em 1992.

Existem também algumas generalizações do Teorema de Roth para contextos diferentes. Ridout publicou, em 1958, uma generalização do Teorema de Roth para números p-ádicos. O artigo de Ridout [**15**] segue a estrutura do artigo de Roth, ele cita alguns Lemas do artigo de Roth com as devidas modificações e assim consegue uma versão do Teorema de Roth usando a métrica p-ádica.

LeVeque apresenta, em seu livro [**12**], o seguinte resultado:

TEOREMA 15. *Seja K um corpo de números algébricos de grau N e seja α algébrico de grau $n \geq 2$ sobre K. Então para cada $\mu > 2$ a desigualdade*

$$|\alpha - \zeta| < \frac{1}{(H(\zeta))^{\mu}}$$

tem apenas um número finito de soluções $\zeta \in K$.

Onde $H(\zeta)$ é a altura de ζ. Esta demonstração também segue os passos da demonstração de Roth. Os interessados nesta generalização podem se interessar pelo trabalho de Daniel Ishak [**21**]. Neste trabalho, Ishak segue as demonstrações do livro de LeVeque corrigindo alguns problemas que surgem na generalização para corpos de números algébricos.

O Teorema de Roth é também corolário do Teorema dos subespaços de Schmidt provado em 1972.

TEOREMA 16 (Schmidt). *Seja $x = (x_1, \ldots, x_m)$. Suponha que as formas lineares*

$$L_i(x) = a_{i,1}x_1 + \cdots + a_{i,m}x_m, \quad i = 1, \ldots, m,$$

têm coeficientes números algébricos e são linearmente independentes. Seja $\delta > 0$. Então, existe um número finito de subespaços próprios de

$\mathbb{Q}^n$ *tais que, todo* $x \in \mathbb{Z}^m$ *não-nulo que satisfaz*

$$|L_1(x) \dots L_m(x)| < |x|^{-\delta},$$

pertence um destes subespaços.

Uma demonstração deste Teorema pode ser encontrada no livro de Schmidt [**18**].

Vamos mostrar como o Teorema do subespaço implica no Teorema de Roth. Sejam $m = 2$, α um número algébrico, $L_1(x, y) = \alpha y - x$ e $L_2(x, y) = y$. Para $\delta > 0$ o Teorema do subespaço de Schmidt nos diz que todos os pontos da forma $(p, q) \in \mathbb{Z}^2$ tais que

$$|\alpha q - p||q| < \sqrt{p^2 + q^2} \leq \max\{|p|, |q|\}^{-\delta}$$

pertence a um conjunto finito de retas.

Se

$$|\alpha y - x||y| < \max\{|x|, |y|\}^{-\delta} \tag{160}$$

possuir infinitas soluções inteiras, então um número infinitos de soluções estão na mesma reta.

Seja $x = ky$, onde $k \in \mathbb{Q}$, uma reta com infinitas soluções de (160). Seja $(p_0, q_0) \in \mathbb{Z}^2$ satisfazendo $p_0 = kq_0$ e p_0 o menor inteiro positivo com esta propriedade. Os pontos de coordenadas inteiras de $x = ky$ são da forma (tp_0, tq_0), onde $t \in \mathbb{Z}$. E as soluções são da forma de (160) satisfazem:

$$|tq_0\alpha - tp_0||tq_0| < \max\{|tq_0|, |tp_0|\}^{-\delta} < |tq_0|^{-\delta} \max\{1, |k|^{-\delta}\}.$$

Da desigualdade acima conseguimos a desigualdade:

$$|t|^{2+\delta} < |q_0|^{1-\delta}|q_0\alpha - p_0|^{-1} \max\{1, |k|^{-\delta}\}.$$

Portanto, o valor absoluto de t é limitado e (160) só pode ter um número finito de soluções.

Como (160) tem um número finito de soluções concluímos que existe $c(\alpha, \delta) > 0$ tal que

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha, \delta)}{q^{2+\delta}}$$

tem um número finito de soluções e este é o Teorema de Roth.

3. Conclusão

A demonstração original do Teorema de Roth é bastante complicada e é necessário trabalhar com várias variáveis restritas a várias condições. Nessa dissertação optamos por demonstrar alguns dos principais Teoremas que antecedem o Teorema de Roth e mostrar como as ideias deles influenciaram a demonstração deste Teorema.

Tentamos explorar algumas questões que surgem naturalmente ao estudar este problema:

- Como é feita a transição de uma variável (Teorema de Liouville) para duas variáveis e depois para várias variáveis?
- Como surgiu a ideia de utilizar o wronskiano generalizado?
- O que motivou Roth a definir o índice para polinômios em várias variáveis?

As duas últimas questões são respondidas quando estudamos o trabalho de Dyson e Mahler e vemos como o wronskiano e o índice já aparecem nestes trabalhos. Já a questão sobre o número de variáveis é mais interessante. Em uma variável é clara a associação do número algébrico ao seu polinômio minimal. A passagem para duas variáveis ocorreu devido à tentativa de achar um polinômio melhor que o polinômio minimal para estudar as aproximações por números racionais. Enquanto parecia claro que as técnicas usadas para estudar esse problema precisavam ser aperfeiçoadas para várias variáveis, a abordagem utilizada por Roth não nos permite fixar o número de variáveis e estudar o quão bem o método funciona neste caso. Uma das dificuldades da demonstração é que o número de variáveis precisa satisfazer uma desigualdade que depende do grau do número algébrico e de k, o expoente de q na desigualdade (148). Se uma das principais dificuldades das técnicas usadas para atacar esse problema é o grande número de condições que devem ser satisfeitas, a passagem para o contexto de várias variáveis fazendo com que o número de variáveis ainda dependa de certas condições faz com que o Teorema de Roth seja talvez inevitavelmente difícil, mas, Roth foi muito bem sucedido, conseguindo uma demonstração muito clara e bem estruturada.

Índice Remissivo

Referências Bibliográficas

[1] K. F. Roth (1955). *Rational approximations to algebraic numbers*, Mathematika, **2**, 1-20.

[2] F. J. Dyson (1947). *The approximation to algebraic numbers by rationals*, Acta Math, **79**, 225-240.

[3] K. Mahler (1949). *On Dyson's improvement of the Thue-Siegel theorem*, Proc. Kon. Nederlandsche Akad. v . Wetenschappen **52**, 449-458.

[4] C.L. Siegel(1921). *Approximation algebraischer Zahlen*, Mathematische Zeitschrift **10** (3), 173-213.

[5] A. Thue (1909). *Über Annäherungswerte algebraischer Zahlen*, Journal für die reine und angewandte Mathematik **135**, 284-305.

[6] G.H. Hardy, E. M. Wright (1983). *An Introduction to the Theory of Numbers*, 5th edition, Oxford Science Publications.

[7] I. M. Niven (1961). *Numbers: Rational and Irrational*. New York: Random House

[8] A.O. Gelfond, (1960) [1952]. *Transcendental and algebraic numbers*, Dover Phoenix editions. New York: Dover Publications.

[9] T.N. Shorey (2003). *Approximations of algebraic numbers by rationals: A theorem of Thue*, Elliptic Curves, Modular Forms and Cryptography ed. By A.K. Bhandari, D.S. Nagraj, B. Ramakrishnan, T.N. Venkataramana, Hinustan Book Agency, 119-137.

[10] J. C. Oxtoby (1980). *Measure and Category*, Graduate Texts in Mathematics, (2nd ed.), Springer-Verlag.

[11] Y. Lequain (1993). *Aproximação de um número real por números racionais*, 19º Colóquio. Brasileiro de Matemática, IMPA.

[12] W. J. LeVeque (1956). *Topics in Number Theory, Volumes I and II*, New York: Dover Publications.

[13] L. J. Mordell (1969). *Diophantine Equations*, Academic Press, London.

[14] A. Bostan and P. Dumas (2010). *Wronskians and Linear Independence*, The American Mathematical Monthly **117** (8), 722-727.

[15] D. Ridout (1958) *The p-adic generalization of the Thue-Siegel-Roth theorem*, Mathematika, **5**, 40-48.

[16] H. Esnault, E. Viehweg (1984). *Dyson's Lemma for polynomials in several variables (and the Theorem of Roth)*, Inventiones Mathematicae **78** (3), 445-490.

[17] P. Corvaja (1992). *Roth's theorem via an interpolation determinant*, C.R. Acad. Sci. Paris, Sér. A **315**, 517-521.

[18] W. M. Schmidt.(1980) *Diophantine approximation*, Lecture Notes in Mathematics **785**. Springer.

[19] A. N. Parshin, I. R. Shafarevich (1997) *Number Theory IV: Transcendental Numbers*, Encyclopaedia of Mathematical Sciences **44**. Springer.

[20] J. Sally, P. Sally (2007) *Roots to Research: A vertical development of mathematical problems*, AMS.

[21] D. Ishak (2008) *The Thue-Siegel-Roth Theorem*, Examensarbete i matematik Uppsala Universitet. Atualmente, este trabalho pode ser encontrado em: http://www2.math.uu.se/research/pub/Ishak1.pdf.

[22] J. Liouville (1851) *Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques*, Journal de mathématiques pures et appliquées 1re série, **16**, p. 133-142.

[23] G. V. Chudnovsky (1983) *On the Method of Thue-Siegel: Dedicated to the Memory of Carl Ludwig Siegel*, The Annals of Mathematics, Second Series, Vol. **117**, No. **2**, pp. 325-382

[24] H. Davenport (1958) *The Work of K F Roth*, Fields Medallists' Lectures, 2nd Edition, World Scientific Series in 20th Century Mathematics - Vol. **9**