CENTRO de CIêNCIAS EXATAS e da NATUREZA
DEPARTAMENTO de MATEMáTICA

Em torno do teorema de Roth

por
Tiago Duque Marques

2010

Em torno do teorema de Roth

por
Tiago Duque Marques
sob orientações de Antonio Carlos Monteiro e Aron Simis

dissertação de mestrado
apresentada ao departamento de matemática
como requisito parcial para obtenção do grau
de Mestre em Ciências Matemáticas

Universidade Federal de Pernambuco
Recife, 10 de agosto de 2010

Duque Marques, Tiago
Em torno do teorema de Roth / Tiago Duque Marques. - Recife: O Autor, 2010.
63 folhas

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCEN. Matemática, 2010.

Inclui bibliografia.

1. Teoria dos números. 2. Geometria Diofantina. I. Título.

512.7 CDD (22. ed.) MEI201 – 0119

Dissertação submetida ao Corpo Docente do Programa de Pós-graduação do Departamento de Matemática da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Mestrado em Matemática.

Aprovado: _____

Aron Simis, UFPE

Orientador

_____

Seyed Hamid Hassanzadeh Hafshejani, Univ. Tarbiat Moallem

_____

Hemar Teixeira Godinho, UnB

EM TORNO DO TEOREMA DE ROTH

Por

Tiago Duque Marques

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE MATEMÁTICA
Cidade Universitária – Tels. (081) 2126 - 8414 – Fax: (081) 2126 - 8410
RECIFE – BRASIL

Agosto – 2010

Por que você é Flamengo
e meu pai Botafogo
o que significa
impávido colosso

Por que os ossos doem
enquanto a gente dorme
por que os dentes caem
por onde os filhos saem

Por que os dedos murcham
quando estou no banho
por que as ruas enchem
quando está chovendo

Quanto é mil trilhões
vezes infinito
quem é Jesus Cristo
onde estão meus primos

Por que o fogo queima
por que a lua é branca
por que a terra roda
por que deitar agora

Por que as cobras matam
por que o vidro embaça
por que você se pinta
por que o tempo passa

Por que a gente espirra
por que as unhas crescem
por que o sangue corre
por que a gente morre

Do que é feito a nuvem
do que é feito a neve
como é que se escreve
réveillon

Adriana Partimpim (Oito anos)

Agradecimentos

É um grande prazer poder agradecer aqui a todos aqueles que me ajudaram, motivaram e acompanharam neste período.

Começo por agradecer a Antonio Carlos Monteiro, que me orientou desde a graduação me apresentando uma matemática de muito bom gosto. Obrigado por ter sempre encontrado tempo para discutirmos e sobretudo de ter me escutado com paciência.

Gostaria de estender minha gratidão a Aron Simis, por me encorajar, por seus cursos ao longo desses anos aqui na UFPE que constituem boa parte de minha formação.

Agradeço a Hamid Hassanzadeh e Hemar Godinho por terem aceitado, gentilmente, de fazerem parte da banca examinadora de minha dissertação e por suas valiosas sugestões que me permitiram melhorar o texto.

Agradeço aos professores Sérgio Santa-Cruz, Sóstenes Lins, Hildeberto Cabral e Karl-Otto Sthör por lapidarem através de seus cursos meu modo de ver matemática. Em especial, agradeço a Marc Hindry pelas discussões em Rennes e no IMPA e por sua cuidadosa atenção.

É um honra agradecer pelas interessantes conversas com Hugo, Graciliano, Lucas, Giovana, Karla, Antonio, Eudes, Raphael, Bruno, Zaqueu, André e Abbas; ótimos jovens matemáticos.

Agradeço a Tânia e a Cláudia por me atenderem prontamente a tudo que precisei.

Eu agradeço muito a minha família, hoje maior com Francisco! Meus pais Milton e Iguacy, minha madrinha Iara, meus irmãos Miltinho e Débora pela base indispensável; minha esposa Karla pela força incondicional.

Resumo

Nesta dissertação de mestrado vamos apresentar métodos da aproximação de números algébricos por racionais que são usados para provar resultados de finitude em geometria Diofantina. Faremos isto através do teorema de Roth e de sua generalização a dimensões superiores, o teorema do subespaço de Schmidt; eles permitem demonstrar quase todos os resultados sobre o conjunto de pontos inteiros sobre curvas algébricas, ilustraremos isso com uma nova prova do famoso teorema de Siegel, dada recentemente por P. Corvaja e U. Zannier.

Palavras-chave: aproximação Diofantina, alturas, teorema de Roth, pontos inteiros.

Abstract

In this master thesis we present methods of Diophantine approximation that are used to prove finiteness theorems in Diophantine geometry. We do this through the theorem of Roth and its generalization to higher dimensions, the Schmidt's subspace theorem, they will demonstrate nearly all results on set of integral points on algebraic curves, we will illustrate this with a new proof of the famous theorem of Siegel recently given by P. Corvaja and U. Zannier.

Key words: Diophantine approximation, heights, Roth's theorem, integral points.

Conteúdo

Introdução

Os objetivos desta dissertação é apresentar métodos da aproximação Diofantina que são usados para provar resultados de finitude em geometria Diofantina. Este é um tópico clássico da teoria dos números e suas conexões com geometria algébrica. Faremos isto através do teorema de Roth e de sua generalização, o teorema do subespaço de Schmidt, que permitem demonstrar um famoso teorema de Siegel sobre pontos inteiros sobre curvas algébricas, do final da década de 20.

A inexistência de referências sobre o assunto na literatura em português é uma motivação extra.

O teorema de Roth diz que existe um número finito de soluções racionais da desigualdade

$$(*) \quad \left|\alpha - \frac{p}{q}\right| < q^{-\varepsilon},$$

em que, α é um número algébrico irracional e $\varepsilon > 2$. Graças a esse resultado, em 1958 foi concedida a Roth a medalha Fields! Ele é o resultado central na reta e foi generalizado a dimensões superiores por W.M. Schmidt na década de 70 com o seu teorema do subespaço: sejam $L_1, ..., L_m \in \overline{\mathbf{Q}}[X_1, ..., X_n]$ formas lineares independentes. Então existe uma quantidade finita de subespaços lineares próprios $T_1, ..., T_r \subset \mathbf{C}^n$ tais que

$$\{\underline{x} \in \mathbf{Z}^n || \underline{x}|^{\varepsilon} \cdot |L_1(\underline{x}) \cdots L_m(\underline{x})| \leq 1\} \subset \bigcup_{i=1}^{r} T_i.$$

Para ver que o teorema de Roth segue deste teorema, tomamos $m = n = 2$, $L_1(x, y) = x - \alpha y$ e $L_2(x, y) = y$.

Vamos aplicar esses resultados ao estudo qualitativo dos pontos inteiros sobre curvas algébricas, aqui, o resultado fundamental de finitude é o teorema de Siegel ([17]): suponhamos que uma curva afim C definida sobre um corpo de números K possua uma quantidade infinita de pontos inteiros. Então o seu fecho projetivo $\widetilde{C}$ tem gênero zero e além disso tem no máximo dois pontos no infinito. É claro que para gênero maior ou igual a dois, esse

teorema, é um caso especial do teorema de Faltings—conjectura de Mordell que garante que o próprio $\widetilde{C}(K)$ é finito ([7]). Contudo o tratamento original de Siegel permanece de interesse independente. A prova de Siegel começa por mergulhar $\widetilde{C}$ em sua Jacobiana, assim, essencialmente, os ingredientes são: aproximação Diofantina, teorema de Mordell-Weil e alturas sobre variedades Abelianas. A parte de aproximação Diofantina, hoje, é o teorema de Roth, que já simplifica bastante o argumento original. Recentemente, em [5], P. Corvaja e U. Zannier deram mais uma substancial simplificação da prova do teorema de Siegel invocando o teorema do subespaço de Schmidt ao invés do teorema de Roth; assim, eles evitaram completamente a parte da prova que usa a aritmética das variedades Abelianas, o que deu a possibilidade de estender o resultado a dimensões superiores. Nesse sentido tem-se um bonito teorema de Corvaja-Zannier [6], A. Levin [12] e P. Autissier [1]: uma superfície afim com quatro ou mais divisores amplos intersectando-se propriamente no infinito não pode ter um conjunto Zariski-denso de pontos inteiros; o seminário Bourbaki de Yu. Bilu [3] contém uma prova desse resultado.

Usamos como principais referências os livros *Diophantine Geometry*, de M. Hindry e J. Silverman, e *Heights in Diophantine Geometry*, de E. Bombieri e W. Gubler. Alguns argumentos foram detalhados um pouco mais, lançando mão de outras fontes especializadas.

A dissertação está dividida em quatro capítulos da seguinte forma: no primeiro capítulo expomos a teoria de alturas sobre o espaço projetivo. Definimos a altura de um número algébrico, mostramos algumas relações entre alturas e polinômios e concluímos com a prova da desigualdade de Gelfand, um importante resultado usado na prova do teorema de Roth. O segundo capítulo é dedicado, essencialmente, ao enunciado do teorema de Roth, na versão envolvendo alturas e outros valores absolutos sobre um corpo de números, lá também, fazemos duas reduções do teorema de Roth: mostramos que é suficiente prová-lo para inteiros algébricos e reduzimos as aproximações simultâneas. O terceiro capítulo é o coração da dissertação, é onde damos uma demonstração completa do teorema de Roth, ela consiste basicamente dos seguintes passos:

Primeiro: fixamos um m grande, $d_1, ..., d_m$ inteiros e construimos um polinômio $P \in \mathbf{Z}[X_1, ..., X_m]$ de grau no máximo d_i em X_i que se anula muito em $(\alpha, ..., \alpha)$ (utilizando a noção de índice ou ordem de anulamento ponderada) e cujos coeficientes são de tamanho controlado; a maneira que faremos isso é resolvendo um sistema de equações lineares com coeficientes inteiros (Lema de Siegel).

Segundo: supomos que existem $p_1/q_1, ..., p_m/q_m$ satisfazendo $(*)$, então mostramos que P e um certo número de suas derivadas são pequenas em $\beta =$

$(p_1/q_1, ..., p_m/q_m)$ aplicando a fórmula de Taylor junto com a desigualdade triangular (ou ultramétrica).

Terceiro: empregamos um resultado de não-anulamento (Lema de Roth); esta é a parte difícil da prova, temos que mostrar que uma derivada de ordem muito pequena é não-nula, i.e.,

$$\eta = \partial_i P(\beta) \neq 0 \ \text{ com } \ \partial_i = \frac{1}{i_1!...i_m!}(\partial/\partial X_1)^{i_1} \cdots (\partial/\partial X_m)^{i_m}.$$

Para isso vamos introduzir o determinante Wronskiano generalizado.

Quarto: como η é racional e não-nulo:

$$(\text{denominador}(\eta))^{-1} < |\eta|;$$

esta é a "desigualdade de Liouville", ela fornece uma contradição se m é grande e os q_i são bem escalonados.

Finalmente, temos um quarto capítulo, onde provaremos o teorema de Siegel. Fazemos isso com a demonstração de Corvaja e Zannier que usa o teorema do subespaço de Schmidt. E apenas indicamos, sem riqueza de detalhes, o argumento clássico que usa o teorema de Roth.

Uma ótima leitura!

A Karla
com todo amor que eu possuo

Capítulo 1

Funções alturas

Vamos introduzir uma noção precisa de "tamanho" para os pontos algébricos do espaço projetivo, que chamaremos altura. A versão aqui apresentada é comumente chamada de altura de Weil. Começaremos definindo a altura de um ponto do espaço projetivo a coordenadas racionais, depois a coordenadas algébricas, para enfim, deduzir a noção de altura de um número algébrico. Aproveitaremos para fixar notações, principalmente, no que diz respeito a valores absolutos sobre um corpo. Na segunda seção investigaremos as relações entre alturas e polinômios a fim de demonstrar o teorema de Northcott e a desigualdade de Gelfand.

1.1 Alturas sobre $\mathbf{P}^n$

Definição 1.1.1 *Se P é um ponto de $\mathbf{P}^n(\mathbf{Q})$, podemos escolher suas coordenadas homogêneas $(x_0 : ... : x_n)$ com $x_i \in \mathbf{Z}$ e $\mathrm{mdc}(x_0, ..., x_n) = 1$. Definimos a* **altura** *de P como sendo a quantidade*

$$H_{\mathbf{Q}}(P) = \max\{|x_0|, ..., |x_n|\}.$$

É claro que para qualquer número $B \geq 0$, o conjunto

$$\{P \in \mathbf{P}^n(\mathbf{Q}) | H_{\mathbf{Q}}(P) \leq B\}$$

é finito, pois existe somente uma quantidade finita de inteiros $x \in \mathbf{Z}$ satisfazendo $|x| \leq B$. Essa definição é bastante simples e não se transporta com facilidade as coordenadas algébricas. Será mais cômodo, tecnicamente, reinterpretar as alturas em termo do conjunto dos valores absolutos do corpo. E desejamos manter o importante atributo dessa função altura de que apenas um número finito de pontos tem altura limitada.

O valor absoluto, usual, arquimediano sobre $\mathbf{Q}$ é definido, pondo, para cada número racional x,

$$|x| = |x|_\infty = \max\{x, -x\}.$$

Para definir os valores absolutos, p-ádicos, ultramétricos, escrevemos o número racional x não-nulo na forma $x = \pm p_1^{e_1} \cdots p_r^{e_r}$, com $e_i = \mathrm{ord}_{p_i}(x) \in \mathbf{Z}$ e p_i um número primo. Pomos então, para cada primo p,

$$|x|_p = p^{-\mathrm{ord}_p(x)}.$$

Qualquer valor absoluto v sobre k, um corpo qualquer, define uma distância pondo

$$d(x, y) = |x - y|_v$$

e em particular induz uma topologia sobre k. Dois valores absolutos $|\cdot|_v, |\cdot|_w$ são ditos *equivalentes* se eles induzem a mesma topologia sobre k. Pode ser provado que isso acontece se, e somente se, existe um número real $s > 0$ tal que $|x|_v = |x|_w^s$ para todo $x \in k$. Denotaremos por M_k o conjunto dos *lugares* de k (isto é, classes de equivalência de valores absolutos). Para $M_{\mathbf{Q}}$ tomaremos como representantes o valor absoluto arquimediano $|\cdot|_\infty$, e todos os valores absolutos p-ádicos $|\cdot|_p$ em que p é um número primo.

A fórmula do produto para um número racional x não-nulo se escreve então

$$\prod_{v \in M_{\mathbf{Q}}} |x|_v = 1.$$

Ela simplesmente reflete o fato de que $\mathbf{Z}$ é um domínio de fatoração única.

Corolário 1.1.1 *Sejam $P \in \mathbf{P}^n(\mathbf{Q})$ e $(x_0 : ... : x_n)$ coordenadas homogêneas quaisquer de P. Então*

$$H_{\mathbf{Q}}(P) = \prod_{v \in M_{\mathbf{Q}}} \max\{|x_0|_v, ..., |x_n|_v\}.$$

Demonstração. A fórmula do produto mostra que o lado direito da igualdade do enunciado independe das coordenadas homogêneas. Escolhendo $x_i \in \mathbf{Z}$ primos entre si, temos que para cada p primo $\max\{|x_0|_p, ..., |x_n|_p\} = 1$. Portanto o lado direito será igual a $\max\{|x_0|_\infty, ..., |x_n|_\infty\}$, isto é, a $H_{\mathbf{Q}}(P)$.

■

Para generalizar as alturas aos pontos com coordenadas algébricas, vamos definir os valores absolutos padrões sobre um corpo de números K, i.e., uma extensão finita dos racionais.

Para um corpo de números K fixaremos como representantes de M_K os valores absolutos cuja restrição a $\mathbf{Q}$ seja o valor absoluto arquimediano $|\cdot|_\infty$ ou um dos valores absolutos p-ádicos $|\cdot|_p$. Denotaremos ainda por M_K^∞ o conjunto dos valores absolutos arquimedianos sobre K e por M_K^0 seus valores absolutos ultramétricos.

Sejam $L|K$ uma extensão de corpos de números, $v \in M_K, w \in M_L$ valores absolutos. Diremos que w *divide* v ou w *está acima* de v e escreveremos $w|v$ se a restrição de w a K é v. Diremos que v é p-ádico se ele está acima de algum valor absoluto p-ádico de $\mathbf{Q}$.

Para qualquer valor absoluto $v \in M_K$ denotaremos por K_v o completamento de K com respeito a v, $\mathbf{Q}_v$ o completamento de $\mathbf{Q}$ com respeito a restrição de v a $\mathbf{Q}$ e $n_v = [K_v : \mathbf{Q}_v]$ o grau local de v. O valor absoluto *normalizado* associado a v é

$$\|x\|_v := |x|_v^{n_v}.$$

Lembremos aqui a fórmula do grau cuja demonstração pode ser encontrada em [3] ou [13].

Proposição 1.1.1 *Seja $L|K$ uma extensão de corpos de números. Seja $v \in M_K$ um valor absoluto sobre K. Então*

$$\sum_{w \in M_L, w|v} [L_w : K_v] = [L : K].$$

Vamos dar agora uma descrição alternativa dos valores absolutos sobre um corpo de números K de grau n. O corpo K admite r_1 mergulhos reais e r_2 pares de mergulhos complexos de modo que $n = r_1 + 2r_2$. Cada mergulho $\sigma : K \longrightarrow \mathbf{R}$ ou $\mathbf{C}$ produz, por composição com o módulo, um valor absoluto. Se σ for um mergulho complexo, ele e seu conjugado dão o mesmo valor absoluto. Dispomos então de $r_1 + r_2$ valores absolutos arquimedianos

$$|x|_\sigma = \begin{cases} |\sigma(x)| & \text{para } \sigma \text{ real} \\ |\sigma(x)|^2 & \text{para } \sigma \text{ complexo.} \end{cases}$$

Seja O_K o anel de inteiros de K. Usaremos agora o fato fundamental de que O_K é um domínio de Dedekind, os ideais fracionários possuem uma fatoração única em um produto de ideais primos. Se p, um primo racional, se fatora em

$$pO_K = \wp_1^{e_1} \cdots \wp_g^{e_g},$$

com $\mathrm{N}(\wp_i) = p^{f_i}$ e $\sum_{i=1}^{g} e_i f_i = n$, podemos definir para cada ideal primo $\wp$ um valor absoluto

$$|x|_\wp = \mathrm{N}(\wp)^{-\mathrm{ord}_\wp(x)},$$

em que $\mathrm{N}(\cdot)$ é a norma de um ideal primo. Esses últimos valores absolutos são ultramétricos.

Decorre dessas escolhas (detalhes em [11] ou [13]) que, para $x \in K$,

$$\prod_{\wp|p} |x|_\wp = |\mathrm{N}_{K|\mathbf{Q}}(x)|_p \quad \text{e} \quad \prod_{v \in M_K^\infty} |x|_v = |\mathrm{N}_{K|\mathbf{Q}}(x)|_\infty.$$

Vamos provar agora a fórmula do produto para K.

Proposição 1.1.2 *Seja $x \in K^*$. Então,*

$$\prod_{v \in M_K} \|x\|_v = 1.$$

Demonstração. Agrupamos os lugares de K que estão acima de um certo lugar de $\mathbf{Q}$ e usamos a fórmula do produto para $\mathbf{Q}$:

$$\prod_{w \in M_K} \|x\|_w = \prod_{v \in M_\mathbf{Q}} \prod_{w|v} \|x\|_w = \prod_{v \in M_\mathbf{Q}} |\mathrm{N}_{K|\mathbf{Q}}(x)|_v = 1.$$

■

Definição 1.1.2 *Seja $P = (x_0 : ... : x_n) \in \mathbf{P}^n(K)$ um ponto cujas coordenadas homogêneas são escolhidas em K. A* **altura** *de P, relativa a K, é a quantidade*

$$H_K(P) = \prod_{v \in M_K} \max\{\|x_o\|_v, ..., \|x_n\|_v\}.$$

Definimos também a **altura** *de um elemento $\alpha \in K$ como sendo a altura do ponto projetivo correspondente $(1 : \alpha) \in \mathbf{P}^1(K)$. Assim,*

$$H_K(\alpha) = \prod_{v \in M_K} \max\{\|\alpha\|_v, 1\}.$$

A fórmula do produto assegura que a altura $H_K(P)$ está bem definida, isto é, independe da escolha das coordenadas homogêneas para P. Notemos ainda que podemos tomar coordenadas homogêneas para P com alguma coordenada igual a 1, então, segue da definição que $H_K(P) \geq 1$. Pode ser conveniente considerarmos a altura *logarítimica* $h_K(P) = \log H_K(P)$.

A altura de um ponto considerado em diversas extensões varia de uma maneira simples, este é o conteúdo da próxima proposição.

Proposição 1.1.3 *Seja L uma extensão finita de K e $P \in \mathbf{P}^n(K)$, então*

$$H_L(P) = H_K(P)^{[L:K]}.$$

Demonstração. Escrevamos $P = (x_0 : ... : x_n)$. Vamos usar a fórmula do grau, Proposição 1.1.1, lembrando que para cada $w \in M_L, w|v, v \in M_K$ temos

$$n_w = [L_w : \mathbf{Q}_w] = [L_w : K_v]n_v.$$

Logo,

$$\begin{aligned}
H_L(P) &= \prod_{w\in M_L} \max\{\|x_0\|_w, ..., \|x_n\|_w\} \\
&= \prod_{v\in M_K} \prod_{w\in M_L, w|v} \max\{\|x_0\|_w, ..., \|x_n\|_w\} \\
&= \prod_{v\in M_K} \prod_{w\in M_L, w|v} \max\{|x_0|_v^{n_w}, ..., |x_n|_v^{n_w}\} \\
&= \prod_{v\in M_K} \prod_{w\in M_L, w|v} \max\{\|x_0\|_v, ..., \|x_n\|_v\}^{[L_w:K_v]} \\
&= \prod_{v\in M_K} \max\{\|x_0\|_v, ..., \|x_n\|_v\}^{[L:K]} \\
&= H_K(P)^{[L:K]}.
\end{aligned}$$

■

Esta última proposição permite definir a altura, *absoluta*, que independe do corpo de números, ela será definida sobre o conjunto de pontos a coordenadas em $\overline{\mathbf{Q}}$, um fecho algébrico de $\mathbf{Q}$.

Definição 1.1.3 *A* **altura** *(absoluta) sobre $\mathbf{P}^n$ é a função*

$$H : \mathbf{P}^n(\overline{\mathbf{Q}}) \longrightarrow [1, \infty), \quad H(P) = H_K(P)^{1/[K:\mathbf{Q}]},$$

onde K é qualquer corpo de números com $P \in \mathbf{P}^n(K)$.

O próximo lema é um resultado chave, junto com a desigualdade de Gelfand (Proposição 1.2.2 abaixo), da teoria de alturas que usaremos na demonstração do teorema de Roth, ele é conhecido como a desigualdade de Liouville.

Lema 1.1.1 *Sejam K um corpo de números, $\alpha \in K^*$ um elemento não-nulo de K e $S \subset M_K$ qualquer subconjunto de valores absolutos sobre K. Então*

$$\prod_{v \in S} \min\{\|\alpha\|_v, 1\} \geq \frac{1}{H_K(\alpha)}.$$

Demonstração. A fórmula do produto, Proposição 1.1.2, diz que

$$\prod_{v \in M_K} \|\alpha\|_v = 1, \quad \text{o que implica,} \quad 1 = \prod_{v \in M_K} \frac{1}{\|\alpha\|_v}.$$

Agora calculamos

$$\begin{aligned} H_K(\alpha) &= \prod_{v \in M_K} \max\{1, \|\alpha\|_v\} = \prod_{v \in M_K} \|\alpha\|_v \cdot \max\left\{1, \frac{1}{\|\alpha\|_v}\right\} \\ &= \prod_{v \in M_K} \max\left\{1, \frac{1}{\|\alpha\|_v}\right\} = \prod_{v \in M_K} \frac{1}{\min\{1, \|\alpha\|_v\}} \\ &\geq \prod_{v \in S} \frac{1}{\min\{1, \|\alpha\|_v\}}. \end{aligned}$$

Portanto, tomando os inversos, obtemos o resultado desejado.

■

1.2 Alturas e polinômios

Definimos a **norma de Gauss** de um polinômio

$$f = \sum_{i=(i_1,\ldots,i_n) \in I} a_i x_1^{i_1} \cdots x_n^{i_n}$$

com coeficientes em um corpo de números K, com respeito a um valor absoluto v como sendo

$$|f|_v = \max_{i \in I} |a_i|_v.$$

Para estabelecer a relação entre a altura de um número algébrico e seu polinômio mínimo, usaremos o seguinte lema clássico, dito, Lema de Gauss.

Lema 1.2.1 [3, Lemma 1.6.3] *Sejam $f, g \in K[X_1, \ldots, X_m]$ polinômios, em que K é um corpo de números. Se v é um valor absoluto ultramétrico. Então,*

$$|fg|_v = |f|_v |g|_v.$$

Lema 1.2.2 *Seja α um número algébrico de grau d e $K = \mathbf{Q}(\alpha)$. Seja $P \in \mathbf{Z}[X]$ o polinômio mínimo de α escrito na forma*

$$P(X) = a_0(X - \alpha_1) \cdots (X - \alpha_d) = a_0 X^d + ...$$

Então,

$$H_K(\alpha) = |a_0| \prod_{i=1}^{d} \max\{1, |\alpha_i|\}.$$

Demonstração. Consideremos o corpo $L = \mathbf{Q}(\alpha_1, ..., \alpha_d)$. Podemos escrever:

$$H_L(\alpha) = H_K(\alpha)^{[L:K]} = \prod_{\wp \in M_L^0} \max\{1, |\alpha|_\wp\} \prod_{w \in M_L^\infty} \max\{1, |\alpha|_w\}.$$

Analizemos o lado direito da segunda igualdade acima:

$$\prod_{w \in M_L^\infty} \max\{1, |\alpha|_w\} = \prod_{v \in M_K^\infty} \max\{1, |\alpha|_v\}^{[L:K]} = \left(\prod_{i=1}^{d} \max\{1, |\alpha_i|\} \right)^{[L:K]}.$$

Agora, o Lema de Gauss aplicado a f e a sua fatoração mostra que, para $\wp \in M_L^0$, temos

$$1 = \|f\|_\wp = |a_0|_\wp \prod_{i=1}^{d} \max\{1, |\alpha_i|_\wp\}.$$

Fazendo o produto sobre $\wp$ e aplicando a fórmula do produto para L (aplicada a a_0), obtemos

$$1 = \prod_{\wp \in M_L^0} |a_0|_\wp \prod_{i=1}^{d} \prod_{\wp \in M_L^0} \max\{1, |\alpha_i|_\wp\} = |a_0|^{-[L:\mathbf{Q}]} \left(\prod_{\wp \in M_L^0} \max\{1, |\alpha|_\wp\} \right)^d.$$

Finalmente, combinando esses resultados, obtemos

$$H_K(\alpha)^{[L:K]} = |a_0|^{[L:\mathbf{Q}]/d} \prod_{i=1}^{d} \max\{1, |\alpha_i|\}^{[L:K]},$$

que dá a igualdade procurada extraindo as raízes $[L : K]$-ésimas.

■

Definimos o **corpo de definição** de um ponto $P = (x_0 : ... : x_n) \in \mathbf{P}^n(\overline{\mathbf{Q}})$ como

$$\mathbf{Q}(P) = \mathbf{Q}(x_0/x_i, x_1/x_i, ..., x_n/x_i)$$

para qualquer i com $x_i \neq 0$. O principal mérito da função altura é o seguinte teorema de finitude.

Teorema 1.2.1 (NORTHCOTT) *Para quaisquer números* $B, D \geq 0$,

$$\{P \in \mathbf{P}^n(\overline{\mathbf{Q}}) | H(P) \leq B \text{ e } [\mathbf{Q}(P) : \mathbf{Q}] \leq D\}$$

é um conjunto finito. Em particular, para qualquer corpo de números K,

$$\{P \in \mathbf{P}^n(K) | H_K(P) \leq B\}$$

é um conjunto finito.

Demonstração. Seja $P = (x_0 : ... : x_n) \in \mathbf{P}^n(\overline{\mathbf{Q}})$, sem perda de generalidade, podemos supor que $x_0 \neq 0$ e escrever $P = (1 : \alpha_1 : ... : \alpha_n)$, com α_i algébrico. Claramente temos que $H(\alpha_i) \leq H(P)$ e $[\mathbf{Q}(\alpha_i) : \mathbf{Q}] \leq [\mathbf{Q}(P) : \mathbf{Q}]$. Portanto, é suficiente mostrarmos que o conjunto $\{\alpha \in \overline{\mathbf{Q}} \mid H(\alpha) \leq C, [\mathbf{Q}(\alpha) : \mathbf{Q}] \leq d\}$ é finito. O limite sobre o grau e a altura dá, pelo Lema 1.2.2, um limite para os coeficientes (que são em número finito) do polinômio mínimo de α, o que demonstra a finitude.

■

A **altura** (projetiva) de um polinômio é definida como a altura dos seus coeficientes tomados como coordenadas homogêneas, assim

$$H_K(f) = \prod_{v \in M_K} |f|_v^{n_v} \text{ e } h(f) = \log H(f) = \frac{1}{[K : \mathbf{Q}]} \sum_{v \in M_K} n_v \log |f|_v.$$

Proposição 1.2.1 [9, Proposition B.7.2] *Sejam* $f_1, ..., f_r \in K[X_1, ..., X_m]$. *Seja* $\mathrm{gr}(f_i)$ *o grau total de* f_i. *Então,*

$$h(f_1 \cdots f_r) \leq \sum_{i=1}^{r} (h(f_i) + (\mathrm{gr}(f_i) + m) \log 2).$$

Proposição 1.2.2 (GELFAND) *Sejam* $d_1, ..., d_r$ *inteiros. Sejam* $f_1, ..., f_r \in \overline{\mathbf{Q}}[X_1, ..., X_m]$ *polinômios cujo produto satisfaz* $\mathrm{gr}_{X_i}(f_1 \cdots f_r) \leq d_i$ *para cada* $1 \leq i \leq r$. *Então*

$$\sum_{i=1}^{r} h(f_i) \leq h(f_1 \cdots f_r) + d_1 + \cdots + d_m.$$

Demonstração. Começamos lembrando o Lema de Gauss

$$|f_1 \cdots f_r|_v = |f_1|_v \cdots |f_r|_v$$

em que v é qualquer valor absoluto ultramétrico. O fato fundamental para concluir a desigualdade de Gelfand é a prova de um resultado análogo, arquimediano,

$$\prod_{i=1}^{r} |f_i| \leq e^{d_1+\cdots+d_m}|f|, \tag{1.1}$$

válido para todos os polinômios $f, f_1, ..., f_r \in \mathbf{C}[X_1, ..., X_m]$. De fato, assumindo por um momento (1.1), temos que

$$\begin{aligned}
\prod_{i=1}^{r} H_K(f_i) &= \prod_{i=1}^{r} \prod_{v\in M_K} |f_i|_v^{n_v} \\
&\leq \prod_{v\in M_K^0} |f_1\cdots f_r|_v^{n_v} \prod_{v\in M_K^\infty} e^{n_v(d_1+\cdots+d_m)}|f_1\cdots f_r|_v^{n_v} \\
&\leq e^{[K:\mathbf{Q}](d_1+\cdots+d_m)} H_K(f_1\cdots f_r),
\end{aligned}$$

logo, tomando raízes $[K : \mathbf{Q}]$-ésimas e aplicando o logaritmo em ambos os membros obtemos a desigualdade de Gelfand.

Provaremos (1.1) introduzindo uma norma multiplicativa e uma norma L^2 sobre o espaço de polinômios.

Definição 1.2.1 *Sejam*

$$I = [0, 1], \quad \mathbf{i} = (i_1, ..., i_m), \quad t = (t_1, ..., t_m), \quad dt = dt_1 \cdots dt_m,$$

e seja $\mathbf{e}(t) = (\exp(2\pi i t_1), ..., \exp(2\pi i t_m))$. *Para qualquer polinômio complexo*

$$f = \sum_{\mathbf{i}} a_{\mathbf{i}} X_1^{i_1} \cdots X_m^{i_m} \in \mathbf{C}[X_1, ..., X_m],$$

a **medida de Mahler** *de f é a quantidade*

$$M(f) = \exp\left(\int_{I^m} \log|f(\mathbf{e}(t))|dt\right),$$

e a L^2**-norma** *de f é a quantidade*

$$L_2(f) = \left(\int_I |f(\mathbf{e}(t))|^2 dt\right)^{1/2} = \left(\sum_{\mathbf{i}} |a_{\mathbf{i}}|^2\right)^{1/2}.$$

Lema 1.2.3 [9, Lemma B.7.3.1] *Sejam* $f, g \in \mathbf{C}[X_1, ..., X_m]$ *polinômios e suponhamos que* $\mathrm{gr}_{X_j}(f) \leq d_j$. *Então*

(i) $L_2(f) \leq [(d_1 + 1) \cdots (d_m + 1)]^{1/2} |f|$.

(ii) $M(fg) = M(f)M(g)$.

(iii) $M(f) \leq L_2(f)$.

Lema 1.2.4 *Para qualquer polinômio* $f = \sum_{i=0}^{d} a_i X^i \in \mathbf{C}[X]$,

$$|a_j| \leq \left(\begin{array}{c} d \\ j \end{array} \right) M(f) \leq 2^d M(f).$$

Mais geralmente, consideremos o polinômio complexo

$$f = \sum_{0 \leq j_i \leq d_i} a_{j_1,...,j_m} X_1^{j_1} \cdots X_m^{j_m}, i = 1, ..., m,$$

satisfazendo $\mathrm{gr}_{X_i}(f) \leq d_i$. *Então*

$$|a_{j_1,...,j_m}| \leq \left(\begin{array}{c} d_1 \\ j_1 \end{array} \right) \cdots \left(\begin{array}{c} d_m \\ j_m \end{array} \right) M(f) \leq 2^{d_1 + \cdots + d_m} M(f).$$

Demonstração. A demonstração seguirá por indução sobre o número de variáveis m. Para $m = 1$ fatoramos f

$$f(X) = a_d \prod (X - \alpha_i).$$

Então

$$\begin{aligned} |a_j| = |a_d| \left| \sum_{h_1 < \cdots < h_{d-j}} \alpha_{h_1} \cdots \alpha_{h_{d-j}} \right| &\leq \left(\begin{array}{c} d \\ j \end{array} \right) |a_d| \prod_{i=1}^{d} \max\{1, |\alpha_i|\} \\ &= \left(\begin{array}{c} d \\ j \end{array} \right) M(f). \end{aligned}$$

Antes de prosseguir com a indução vamos introduzir umas notações. Para qualquer $1 \leq n \leq m$, pomos

$$f_{k_1,...,k_n}(X_{n+1}, ..., X_m) = \sum_{h_{n+1}=0}^{d_n+1} \cdots \sum_{h_m=0}^{d_m} a_{k_1,...,k_n,h_{n+1},...,h_m} X_{n+1}^{h_{n+1}} \cdots X_m^{h_m}$$

com $f_{k_1,...,k_m} = a_{k_1,...,k_m}$ no caso $n = m$. Assim, podemos escrever

$$f(X_1, ..., X_m) = \sum_{k_1=0}^{d_1} f_{k_1}(X_2, ..., X_m) X_1^{k_1}, \tag{1.2}$$

e mais geralmente

$$f_{k_1,...,k_{n-1}}(X_n, ..., X_m) = \sum_{k_n=0}^{d_n} f_{k_1,...,k_n}(X_{n+1}, ..., X_m) X_n^{k_n}. \tag{1.3}$$

De (1.2) e da já constatação do lema para uma variável, temos que, para todos $x_2, ..., x_m \in \mathbf{C}$,

$$|f_{k_1}(x_2, ..., x_m)| \leq \begin{pmatrix} d_1 \\ k_1 \end{pmatrix} \exp\left(\int_0^1 \log |f(e^{2\pi i t}, x_2, ..., x_m)| dt \right).$$

Aplicamos agora o logaritmo a ambos os membros, avaliado em $(x_2, ..., x_m) = (e^{2\pi i t_2}, ..., e^{2\pi i t_m})$ e integramos sobre $0 \leq t_2, ..., t_m \leq 1$, donde obtemos

$$\begin{aligned} \log M(f_{k_1}) &= \int_{I^{m-1}} \log |f_{k_1}(e^{2\pi i t_2}, ..., e^{2\pi i t_m})| dt_2 \cdots dt_m \\ &\leq \log \begin{pmatrix} d_1 \\ k_1 \end{pmatrix} + \int_{I^m} \log |f(e^{2\pi i t_1}, ..., e^{2\pi i t_m})| dt_1 \cdots dt_m \\ &\leq \log \begin{pmatrix} d_1 \\ k_1 \end{pmatrix} + \log M(f). \end{aligned}$$

Obtemos assim a desigualdade

$$M(f_{k_1}) \leq \begin{pmatrix} d_1 \\ k_1 \end{pmatrix} M(f).$$

Usando o mesmo argumento para a expressão (1.3) obtemos

$$M(f_{k_1,...,k_n}) \leq \begin{pmatrix} d_n \\ k_n \end{pmatrix} M(f_{k_1,...,k_{n-1}}).$$

Esta dá a cota

$$|a_{k_1,...,k_m}| \leq \begin{pmatrix} d_m \\ k_m \end{pmatrix} M(f_{k_1,...,k_{m-1}})$$

para os coeficientes.

■

Denotemos por $\mu(f)$ o número de variáveis $X_1, ..., X_m$ que de fato aparecem em f. Usando a estimativa básica, válida para $d \geq 1$,

$$\left(\begin{array}{c} d \\ k \end{array} \right) \leq 2^{d-1}$$

obtemos

$$|f| \leq 2^{d_1+\cdots+d_m-\mu(f)} M(f).$$

Finalmente, vamos terminar a demonstração da desigualdade de Gelfand. Sejam $d_{ij} = \mathrm{gr}_{X_j}(f_i)$, $d_j = \mathrm{gr}_{X_j}(f)$, assim

$$d_j = \sum_{i=1}^{r} d_{ij} \quad \text{e} \quad \mu(f) \leq \sum_{i=1}^{r} \mu(f_i).$$

Então

$$\begin{aligned} |f_1| \cdots |f_r| &\leq \prod_{i=1}^{r} \left(2^{d_{i1}+\cdots+d_{im}-\mu(f_i)} M(f_i)\right) \\ &= 2^{d_1+\cdots d_m-\sum_i \mu(f_i)} M(f) \\ &\leq 2^{d_1+\cdots d_m-\mu(f)} ((d_1+1)\cdots(d_m+1))^{1/2} |f|. \end{aligned}$$

Agora, observando que

$$2^d \sqrt{d+1} \leq e^d \quad \text{para} \quad d \geq 2 \quad \text{e para} \quad d = 0,$$

enquanto que se $d_i = 1$, então a variável X_i contribui para $\mu(f)$. Portanto, obtemos

$$\prod_{i=1}^{r} |f_i| \leq e^{d_1+\cdots+d_m} |f|,$$

completando assim a demonstração da Proposição 1.2.2.

■

Capítulo 2

Aproximação racional de números algébricos

A questão clássica em aproximação Diofantina é a de saber quão próxima uma quantidade irracional está de uma quantidade racional. De maneira mais precisa, seja $\alpha \in \mathbf{R}$ e e um expoente dado. Questiona-se: a desigualdade

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^e}$$

pode ter uma quantidade infinita de soluções racionais $p/q \in \mathbf{Q}$? Começamos este capítulo relacionando isto com a construção de números transcedentes feita por Liouville, daremos também a demonstração do resultado de Liouville, pois, embora elementar, contém alguns traços que reaparecerão na prova do teorema de Roth. Nas duas últimas seções fazemos duas reduções do teorema de Roth, primeiro, mostramos que basta prová-lo para inteiros algébricos e depois o reduzimos as aproximações simultâneas.

2.1 Origem transcedente

Teorema 2.1.1 (Liouville 1844) *Se α é um número algébrico com grau $n > 1$ então, para todos os racionais p/q $(p, q \in \mathbf{Z}, q > 0)$, tem-se que*

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

para alguma constante $c = c(\alpha) > 0$ (isto é, c só depende de α).

Este teorema foi usado por Liouville para construir números transcedentes. Por exemplo, o número

$$\xi = \sum_{n=1}^{\infty} 10^{-n!}$$

é transcedente.

De fato, sejam $p_k = 10^{k!} \sum_{n=1}^{k} 10^{-n!}$ e $q_k = 10^{k!}$ para $k = 1, 2, ...$; então p_k e q_k são inteiros racionais relativamente primos e

$$\begin{aligned} \left| \xi - \frac{p_k}{q_k} \right| &= \sum_{n=k+1}^{\infty} 10^{-n!} < 10^{-(k+1)!} \sum_{n=0}^{\infty} 10^{-n} \\ &= \frac{10}{9} q_k^{-(k+1)} < q_k^{-k}. \end{aligned}$$

Como k tende ao infinito não pode existir uma tal constante c, como no teorema de Liouville, dependendo somente de ξ. Portanto ξ é transcedente.

Demonstração do Teorema 2.1.1. Seja $P(x)$ o polinômio mínimo de α, ou seja, irredutível com $P(\alpha) = 0$, com os coeficientes inteiros primos entre si e com coeficiente dominante positivo. Podemos assumir que α é real e que $|\alpha - p/q| < 1$, pois caso contrário o resultado é trivialmente válido. Pelo teorema do valor médio temos que

$$P(\alpha) - P(p/q) = (\alpha - p/q)P'(\xi)$$

para algum ξ entre α e p/q. Assim, $\xi \in (\alpha - 1, \alpha + 1)$ e portanto

$$|P'(\xi)| < 1/c$$

para alguma constante $c = c(\alpha) > 0$. Como $P(\alpha) = 0$ temos que

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

Como P é irredutível de grau n, $P(p/q) \neq 0$ e $|q^n P(p/q)|$ é um inteiro, portanto

$$|P(p/q)| \geq 1/q^n$$

e o resultado segue.

■

Um resultado bem mais profundo nesse contexto do teorema de Liouville foi descoberto por Thue [18], em 1909. Seja α um número algébrico de grau $d > 1$ e consideremos a desigualdade

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^e}$$

para $c = c(\alpha, e) > 0$ e p, q inteiros racionais. Thue mostrou que $c(\alpha, e)$ existe para $e > \frac{1}{2}d+1$. Esse resultado foi melhorado por Siegel para $e > s+d/(s+1)$, com s um inteiro positivo, em particular, para $e > 2\sqrt{d}$, que depois foi melhorado, independentemente, por Dyson e Gelfond, para $e > \sqrt{2d}$. Roth [14], em 1955, colocou um ponto final (ou inicial!) nesta história mostrando que $c(\alpha, e) > 0$ existe para qualquer $e > 2$. Thue foi motivado por seus estudos de certas equações Diofantinas e o melhoramento de Siegel levou ao seu famoso teorema que garante a finitude dos pontos inteiros sobre qualquer curva algébrica de gênero pelo menos 1.

2.2 O enunciado do teorema de Roth

A **aproximação expoente** de um número real α é definida como sendo o maior número $\tau(\alpha)$ com a propriedade que para qualquer expoente $e > \tau(\alpha)$, a desigualdade

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^e}$$

possui somente uma quantidade finita de soluções em números racionais $p/q \in \mathbf{Q}$. O teorema de Roth afirma que todo número algébrico α possui aproximação expoente 2, i.e., para qualquer $\varepsilon > 0$, existe somente uma quantidade finita de números racionais p/q satisfazendo

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

Vejamos que esse expoente do teorema de Roth é o melhor possível. Com este fim, contrua uma sequência de inteiros indutivamente como segue: sejam $a_1 = b_1 = 1$, $a_{n+1} = a_n + 2b_n$, $b_{n+1} = a_n + b_n$. Então, não é difícil ver que,

$$|a_n^2 - 2b_n^2| = 1, \quad \text{para todo } n,$$

$$\left|\frac{a_n}{b_n} - \sqrt{2}\right| < \frac{1}{2b_n^2}, \quad \text{para todo } n.$$

Os números racionais a_n/b_n são todos distintos e eles satisfazem a desigualdade

$$\left|\frac{p}{q} - \sqrt{2}\right| \leq \frac{1}{2q^2},$$

assim, no teorema de Roth, não podemos tomar $\varepsilon = 0$. Um resultado clássico de Dirichlet afirma que se α é qualquer real irracional então $|p/q - \alpha| \leq \frac{1}{q^2}$ possui uma quantidade infinita de soluções racionais.

No próximo capítulo demonstraremos a seguinte formulação do teorema de Roth, dada por Serge Lang, em que a força da aproximação é medida pela altura de Weil dos aproximandos e é introduzido outros valores absolutos definidos sobre um corpo de números (possivelmente não-arquimedianos).

Teorema 2.2.1 (Roth) *Sejam K um corpo de números e $S \subset M_K$ um conjunto finito de valores absolutos sobre K. Assuma que cada valor absoluto em S pode ser estendido, de alguma maneira, a $\overline{K}$. Sejam $\alpha \in \overline{K}$ e $\varepsilon > 0$ dados. Então existe somente uma quantidade finita de números algébricos $\beta \in K$ satisfazendo a desigualdade*

$$\prod_{v \in S} \min\{\|\beta - \alpha\|_v, 1\} \leq \frac{1}{H_K(\beta)^{2+\varepsilon}}. \tag{2.1}$$

A formulação anterior, original, pode ser obtida desta no caso em que $K = \mathbf{Q}$ e $\#S = 1$.

2.3 Redução aos inteiros algébricos

Lema 2.3.1 *Se o Teorema 2.2.1 é verdadeiro para todos os inteiros algébricos, então é verdadeiro para todos os números algébricos.*

Demonstração. Seja α um número algébrico, e suponhamos que o teorema de Roth é falso para α, i.e., existem infinitos números algébricos $\beta \in K$ satisfazendo a desigualdade (2.1). O conjunto S possui uma quantidade finita de subconjuntos, então possivelmente após trocar S por um desses subconjuntos, nós podemos assumir que existem infinitos números algébricos $\beta \in K$ tal que

$$\prod_{v \in S} \|\beta - \alpha\|_v \leq \frac{1}{H_K(\beta)^{2+\varepsilon}}.$$

Escolhamos um inteiro $D > 0$ tal que $D\alpha$ é um inteiro algébrico, e seja $\beta \in K$ uma solução de (2.1) com $H_K(\beta) > H_K(D)^{1+6/\varepsilon}$. Da definição de

altura segue que $H_K(D\beta) \leq H_K(D)H_K(\beta)$. Além disso,

$$\prod_{v\in S} \|D\|_v \leq \prod_{v\in S} \max\{\|D\|_v, 1\} = H_K(D).$$

Portanto

$$\begin{aligned}
\prod_{v\in S} \|D\beta - D\alpha\|_v &\leq \frac{H_K(D)}{H_K(\beta)^{2+\varepsilon}} \\
&= \frac{H_K(D)}{H_K(\beta)^{2+\varepsilon/2}} \cdot \frac{1}{H_K(\beta)^{\varepsilon/2}} \\
&\leq \frac{H_K(D)}{(H_K(D\beta)/H_K(D))^{2+\varepsilon/2}} \cdot \frac{1}{(H_K(D)^{1+6/\varepsilon})^{\varepsilon/2}} \\
&= \frac{1}{H_K(D\beta)^{2+\varepsilon/2}}.
\end{aligned}$$

Assim $D\beta$ é uma aproximação de $D\alpha$ no sentido de que a desigualdade (2.1) é verdadeira quando $\alpha, \beta, \varepsilon$ são trocados por $D\alpha, D\beta, \varepsilon/2$. Portanto a falsidade do teorema de Roth para α implica a falsidade para o inteiro algébrico $D\alpha$.

■

2.4 Redução as aproximações simultâneas

Teorema 2.4.1 *Sejam K um corpo de números, $S \subset M_K$ um conjunto finito de valores absolutos sobre K com cada valor absoluto estendido, de alguma maneira, a $\overline{K}$. Sejam $\alpha \in \overline{K}$ e $\varepsilon > 0$ dados. Suponha que*

$$\xi : S \longrightarrow [0,1] \quad \textit{é uma função satisfazendo} \quad \sum_{v\in S} \xi_v = 1.$$

Então existe somente uma quantidade finita de números algébricos $\beta \in K$ com a propriedade que

$$\|\beta - \alpha\|_v \leq \frac{1}{H_K(\beta)^{(2+\varepsilon)\xi_v}} \quad \textit{para todo} \quad v \in S. \tag{2.2}$$

Este Teorema 2.4.1 lembra bastante o teorema de Roth, mas ele troca a condição que o produto $\prod \|\beta - \alpha\|_v$ é pequeno pela condição que cada diferença $\|\beta - \alpha\|_v$ é pequena. Esta ideia de reduzir a aproximações *simultâneas* é devida a Mahler e a Ridout, que foram os primeiros a estudar aproximação Diofantina para valores absolutos p-ádicos.

Lema 2.4.1 *O Teorema 2.2.1 é verdadeiro se, e somente se, o Teorema 2.4.1 é verdadeiro.*

Demonstração. Primeiro suponhamos que o Teorema 2.2.1 seja verdadeiro. Seja $\xi : S \longrightarrow [0,1]$ uma função como descrita no Teorema 2.4.1, suponhamos que $\beta \in K$ satisfaz (2.2). Então multiplicando a estimativa (2.2) sobre $v \in S$ e usando $\sum_v \xi_v = 1$ mostra que β satisfaz a desigualdade (2.1), portanto o Teorema 2.2.1 garante que existe somente uma quantidade finita de β's.

Agora, reciprocamente, suponhamos que seja verdadeiro o Teorema 2.4.1. Suponhamos, para efeito de contradição, que existem infinitos $\beta \in K$ satisfazendo (2.1).

Seja $s = \#S$. Consideremos a coleção de mapas

$$\xi : S \longrightarrow [0,1] \ \text{ da forma } \ \xi_v = \frac{a_v}{s} \ \text{ com } \ a_v \in \mathbf{Z}, a_v \geq 0 \ \text{ e } \ \sum_{v \in S} a_v = s.$$

Como S é finito existe somente uma quantidade finita de tais mapas. Denotaremos esta coleção de mapas por $\mathcal{Z}$.

Suponhamos que $\beta \in K$ satisfaz (2.1). Mostraremos que β satisfaz (2.2) para um dos mapas em $\mathcal{Z}$. Para cada $v \in S$, definimos um número real $\lambda_v(\beta) \geq 0$ pela fórmula

$$\min\{\|\beta - \alpha\|_v, 1\} = \frac{1}{H_K(\beta)^{(2+\varepsilon)\lambda_v(\beta)}}.$$

Assim, multiplicando sobre $v \in S$ e comparando com (2.1), nós vemos que $\sum_{v \in S} \lambda_v(\beta) \geq 1$, portanto

$$\sum_{v \in S} 2s\lambda_v(\beta) \geq \sum_{v \in S} (2s\lambda_v(\beta) - 1) = 2s \sum_{v \in S} \lambda_v(\beta) - s \geq s.$$

Isso implica que nós podemos encontrar inteiros $b_v(\beta)$ com a propriedade que

$$0 \leq b_v(\beta) \leq 2s\lambda_v(\beta) \ \text{ e } \ \sum_{v \in S} b_v(\beta) = s.$$

Então a função $\xi : S \longrightarrow [0,1]$ definida por $\xi_v = b_v(\beta)/s$ pertence a $\mathcal{Z}$.

Nós provamos que se $\beta \in K$ satisfaz (2.1), então β satisfaz (2.2) para pelo menos uma das funções ξ em $\mathcal{Z}$. Mas, por hipótese, para qualquer ξ existe somente uma quantidade finita de β's satisfazendo (2.2). Portanto (2.1) possui somente uma quantidade finita de soluções.

■

Capítulo 3

Prova do teorema de Roth

3.1 Preliminares

Seja $P(X_1, ..., X_m) \in \mathbf{R}[X_1, ..., X_m]$ um polinômio em várias variáveis com coeficientes em $\mathbf{R}$. Sejam $r_1, ..., r_m$ inteiros não-negativos e suponhamos que o grau de P com respeito a X_j é no máximo r_j. Para qualquer variável t, temos a expansão de Taylor, em torno de t,

$$P(X_1, ..., X_m) = \sum_{i_1=1}^{r_1} \cdots \sum_{i_m=1}^{r_m} \partial_{i_1 \ldots i_m} P(t, ..., t)(X_1 - t)^{i_1} \cdots (X_m - t)^{i_m}.$$

Cada $\partial_{i_1 \ldots i_m} P$ é o polinômio

$$\partial_{i_1 \ldots i_m} P = \frac{1}{i_1! \cdots i_m!} \partial_1^{i_1} \cdots \partial_m^{i_m} P$$

e $\partial_1, ..., \partial_m$ são as derivadas parciais. Observemos que se $P \in \mathbf{Z}[X_1, ..., X_m]$, então os coeficientes de $\partial_{i_1 \ldots i_m} P$ são múltiplos inteiros dos coeficientes de P, pois os fatoriais no denominador dividem os inteiros que ocorrem como resultado da diferenciação parcial repetida. Além disso, tais múltiplos inteiros serão limitados pelo produto dos números combinatórios

$$\left(\begin{array}{c} r_1 \\ i_1 \end{array} \right) \cdots \left(\begin{array}{c} r_m \\ i_m \end{array} \right)$$

que por sua vez é limitado por $2^{r_1 + \cdots + r_m}$. Portanto, se denotamos por $|P|$ o máximo dos valores absolutos dos coeficientes de P, acabamos de observar que $|\partial_{i_1 \ldots i_m} P| \leq 2^{r_1 + \cdots + r_m} |P|$.

É fundamental para a prova do teorema de Roth estimar a ordem de anulamento de um polinômio em várias variáveis em certos pontos. A fim de precisar a medida desta ordem de anulamento temos a seguinte definição.

Definição 3.1.1 *Sejam k um corpo qualquer, $P(X_1, ..., X_m) \in k[X_1, ..., X_m]$ um polinômio, $(\alpha_1, ..., \alpha_m) \in k^m$ um ponto, e sejam $r_1, ..., r_m$ inteiros positivos. O* **índice** *de P com respeito a $(\alpha_1, ..., \alpha_m; r_1, ..., r_m)$, denotado por $\mathrm{Ind}_{(\alpha_1,...,\alpha_m;r_1,...,r_m)}P$ ou, se não houver confusão, $\mathrm{Ind}P$, é o menor valor de*

$$\frac{i_1}{r_1} + \frac{i_2}{r_2} + \cdots + \frac{i_m}{r_m}$$

tal que

$$\partial_{i_1 \ldots i_m} P(\alpha_1, ..., \alpha_m) \neq 0.$$

Se P é o polinômio nulo, definimos o seu índice como sendo ∞.

Lema 3.1.1 *Sejam $P, P' \in k[X_1, ..., X_m]$ polinômios, e sejam $r_1, ..., r_m$ inteiros positivos fixados e um ponto $(\alpha_1, ..., \alpha_m) \in k^m$. O índice com respeito a $(\alpha_1, ..., \alpha_m; r_1, ..., r_m)$ possui as seguintes propriedades:*

(a) $\mathrm{Ind}(\partial_{i_1 \ldots i_m} P) \geq \mathrm{Ind}P - \left(\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m}\right)$.

(b) $\mathrm{Ind}(P + P') \geq \min\{\mathrm{Ind}P, \mathrm{Ind}P'\}$.

(c) $\mathrm{Ind}(PP') = \mathrm{Ind}P + \mathrm{Ind}P'$.

Demonstração. Para facilitar vamos escrever $\alpha = (\alpha_1, ..., \alpha_m)$.
(a) Seja $Q = \partial_{i_1 \ldots i_m} P$. Pela definição de índice podemos escolher uma m-upla $(j_1, ..., j_m)$ tal que $\partial_{j_1 \ldots j_m} Q(\alpha) \neq 0$ e tal que o índice de Q, com respeito a $(\alpha_1, ..., \alpha_m; r_1, ..., r_m)$, é igual a $j_1/r_1 + \cdots + j_m/r_m$. Então

$$\begin{aligned} \partial_{j_1 \ldots j_m} Q(\alpha) \neq 0 &\implies \partial_{i_1+j_1,\ldots,i_m+j_m} P(\alpha) \neq 0 \\ &\implies \frac{i_1 + j_1}{r_1} + \cdots + \frac{i_m + j_m}{r_m} \geq \mathrm{Ind}P \\ &\implies \mathrm{Ind}Q = \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \geq \mathrm{Ind}P - \frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m}. \end{aligned}$$

(b) Escolhamos uma m-upla $(j_1, ..., j_m)$ tal que $\partial_{j_1,\ldots,j_m}(P + P')(\alpha) \neq 0$ e o índice de $P+P'$, com respeito a $(\alpha_1, ..., \alpha_m; r_1, ..., r_m)$, seja $j_1/r_1 + \cdots + j_m/r_m$. Então $\partial_{j_1 \ldots j_m} P(\alpha) \neq 0$ ou $\partial_{j_1 \ldots j_m} P'(\alpha) \neq 0$. Assim, $j_1/r_1 + \cdots + j_m/r_m$ é maior do que ou igual a no mínimo $\mathrm{Ind}P$ ou $\mathrm{Ind}P'$. Logo

$$\mathrm{Ind}(P + P') = \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \geq \min\{\mathrm{Ind}P, \mathrm{Ind}P'\}.$$

(c) Aqui usaremos a fórmula de Leibiniz para a derivada do produto, que pode ser escrita como

$$\partial_{j_1 \ldots j_m}(PP') = \sum_{i_1 + i'_1 = j_1} \cdots \sum_{i_m + i'_m = j_m} (\partial_{i_1 \ldots i_m} P)(\partial_{i'_1 \ldots i'_m} P').$$

Escolhamos uma m-upla $(j_1, ..., j_m)$ tal que $\partial_{j_1 \ldots j_m}(PP')(\alpha) \neq 0$ e tal que

$$\mathrm{Ind}(PP') = \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \text{ em } (\alpha_1, ..., \alpha_m; r_1, ..., r_m).$$

Então existem m-uplas $(i_1, ..., i_m)$ e $(i'_1, ..., i'_m)$ com $\partial_{i_1 \ldots i_m} P(\alpha) \neq 0$ e $\partial_{i'_1 \ldots i'_m} P'(\alpha) \neq 0$. Portanto

$$\mathrm{Ind}P \leq \sum_{h=1}^{m} \frac{i_h}{r_h} \text{ e } \mathrm{Ind}P' \leq \sum_{h=1}^{m} \frac{i'_h}{r_h}.$$

Assim, adicionando essas desigualdades obtemos $\mathrm{Ind}P + \mathrm{Ind}P' \leq \mathrm{Ind}(PP')$.

Para obtermos a desigualdade reversa, vamos considerar o conjunto das m-uplas $(i_1, ..., i_m)$ satisfazendo

$$\partial_{i_1 \ldots i_m} P(\alpha) \neq 0 \text{ e } \mathrm{Ind}P = \sum_{h=1}^{m} \frac{i_h}{r_h}.$$

Vamos ordenar essas m-uplas lexicograficamente e escolhamos a menor delas, digamos, $(\bar{\imath}_1, ..., \bar{\imath}_m)$. Isto significa que se $(i_1, ..., i_m)$ é uma outra m-upla, então existe um $k \geq 1$ tal que

$$i_k = \bar{\imath}_k, \text{ para } 1 \leq h < k \text{ e } i_k > \bar{\imath}_k.$$

Similarmente, escolhamos uma m-upla $(\bar{\imath}'_1, ..., \bar{\imath}'_m)$ para P' e consideremos

$$(\bar{\imath}_1 + \bar{\imath}'_1, ..., \bar{\imath}_m + \bar{\imath}'_m).$$

Então

$$\partial_{\bar{\imath}_1 + \bar{\imath}'_1, ..., \bar{\imath}_m + \bar{\imath}'_m}(PP')(\alpha) = \partial_{\bar{\imath}_1 \ldots \bar{\imath}_m} P(\alpha) \cdot \partial_{\bar{\imath}'_1 \ldots \bar{\imath}'_m} P'(\alpha) \neq 0,$$

pois todas as outras parcelas serão nulas. Portanto

$$\mathrm{Ind}(PP') \leq \sum_{h=1}^{m} \frac{\bar{\imath}_h + \bar{\imath}'_h}{r_h} = \mathrm{Ind}P + \mathrm{Ind}P'.$$

■

3.2 Construção do polinômio auxiliar

Lema 3.2.1 *Seja α um inteiro algébrico de grau d sobre $\mathbf{Q}$, e seja*

$$Q(X) = X^d + a_1X^{d-1} + \cdots + a_{d-1}X + a_d \in \mathbf{Z}[X]$$

o polinômio mínimo de α sobre $\mathbf{Q}$. Então para todo $\ell \geq 0$ podemos escrever

$$\alpha^\ell = a_1^{(\ell)}\alpha^{d-1} + a_2^{(\ell)}\alpha^{d-2} + \cdots + a_{d-1}^{(\ell)}\alpha + a_d^{(\ell)}$$

com inteiros

$$a_i^{(\ell)} \in \mathbf{Z} \quad \textit{satisfazendo} \quad |a_i^{(\ell)}| \leq (|Q| + 1)^\ell.$$

Demonstração. A demonstração seguirá por indução sobre ℓ. Primeiro observemos que a afirmação do lema é clara se $0 \leq \ell \leq d-1$, pois com ℓ neste intervalo podemos tomar todos os $a_i^{(\ell)}$ como sendo 0 ou 1.

Vamos assumir agora que o resultado é válido para α^ℓ. Então

$$\begin{aligned}
\alpha^{\ell+1} &= \alpha \cdot \alpha^\ell = \alpha \sum_{i=1}^{d} a_i^{(\ell)}\alpha^{d-i} \\
&= a_1^{(\ell)}\alpha^d + \sum_{i=2}^{d} a_i^{(\ell)}\alpha^{d+1-i} \\
&= a_1^{(\ell)} \sum_{i=1}^{d} -a_i\alpha^{d-i} + \sum_{i=2}^{d} a_i^{(\ell)}\alpha^{d+1-i} \quad \text{usando que } Q(\alpha) = 0 \\
&= \sum_{i=1}^{d}(-a_1^{(\ell)}a_i + a_{i+1}^{(\ell)})\alpha^{d-i}, \quad \text{com } a_{d+1}^{(\ell)} = 0.
\end{aligned}$$

Donde segue que $a_i^{(\ell+1)} = -a_1^{(\ell)}a_i + a_{i+1}^{(\ell)}$, assim

$$\begin{aligned}
|a_i^{(\ell+1)}| &\leq |a_1^{(\ell)}a_i| + |a_{i+1}^{(\ell)}| \leq \max\{|a_1^{(\ell)}|, |a_{i+1}^{(\ell)}|\} \cdot (|a_i| + 1) \\
&\leq (|Q| + 1)^\ell \cdot (|Q| + 1) = (|Q| + 1)^{\ell+1}.
\end{aligned}$$

■

Lema 3.2.2 *Sejam $r_1, ..., r_m$ inteiros positivos e $0 < \varepsilon < 1$ fixado. Então existem no máximo*

$$(r_1 + 1) \cdots (r_m + 1) \cdot \exp(-\varepsilon^2 m/4)$$

m-uplas de inteiros $(i_1, ..., i_m)$, $0 \leq i_1 \leq r_1$, $0 \leq i_2 \leq r_2$, ..., $0 \leq i_m \leq r_m$ que satisfazem a condição

$$\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} \leq \frac{m}{2} - \varepsilon m.$$

Demonstração. Seja $I(m,\varepsilon)$ o conjunto das m-uplas que vamos tentar contar,

$$I(m,\varepsilon)=\{(i_1,...,i_m)\in\mathbf{Z}^m|0\le i_h\le r_h, 1\le h\le m \ \text{ e } \ \sum_{h=1}^{m}\frac{i_h}{r_h}\le\frac{m}{2}-\varepsilon m\}.$$

Então

$$\begin{aligned}
\#I(m,\varepsilon) &= \sum_{(i_1,...,i_m)\in I(m,\varepsilon)} 1 \\
&\le \sum_{(i_1,...,i_m)\in I(m,\varepsilon)} \exp\left[\frac{\varepsilon}{2}\left(\frac{m}{2}-\varepsilon m-\frac{i_1}{r_1}-\cdots-\frac{i_m}{r_m}\right)\right]
\end{aligned}$$

pois $e^t\ge 1, t\ge 0$

$$\begin{aligned}
&\le \sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}\exp\left[\frac{\varepsilon}{2}\left(\frac{m}{2}-\varepsilon m-\frac{i_1}{r_1}-\cdots-\frac{i_m}{r_m}\right)\right] \\
&= \exp\left(-\frac{\varepsilon^2 m}{2}\right)\sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}\exp\left[\frac{\varepsilon}{2}\left(\frac{m}{2}-\frac{i_1}{r_1}-\cdots-\frac{i_m}{r_m}\right)\right] \\
&= \exp\left(-\frac{\varepsilon^2 m}{2}\right)\prod_{h=1}^{m}\left(\sum_{i_h=0}^{r_h}\exp\left(\frac{\varepsilon}{2}\left(\frac{1}{2}-\frac{i_h}{r_h}\right)\right)\right).
\end{aligned}$$

Usaremos a desigualdade $e^t\le 1+t+t^2$, válida para $|t|\le 1$, para estimar a soma interna como segue

$$\begin{aligned}
\sum_{i=0}^{r}\exp\left(\frac{\varepsilon}{2}\left(\frac{1}{2}-\frac{i}{r}\right)\right) &\le \sum_{i=0}^{r}\left[1+\frac{\varepsilon}{2}\left(\frac{1}{2}-\frac{i}{r}\right)+\frac{\varepsilon^2}{4}\left(\frac{1}{2}-\frac{i}{r}\right)^2\right] \\
&= \sum_{i=0}^{r}\left[\left(1+\frac{\varepsilon}{4}+\frac{\varepsilon^2}{16}\right)-\left(\frac{\varepsilon}{2}+\frac{\varepsilon^2}{4}\right)\frac{i}{r}+\frac{\varepsilon^2}{4}\frac{i^2}{r^2}\right] \\
&= (r+1)\left(1+\frac{\varepsilon^2}{48}+\frac{\varepsilon^2}{24r}\right) \\
&\le (r+1)\left(1+\frac{\varepsilon^2}{4}\right), r\ge 1.
\end{aligned}$$

Finalmente, substituindo essa estimativa acima, obtemos

$$\begin{aligned}
\#I(m,\varepsilon) &\le \exp\left(-\frac{\varepsilon^2 m}{2}\right)\prod_{h=1}^{m}\left((r_h+1)\left(1+\frac{\varepsilon^2}{4}\right)\right) \\
&\le \exp\left(-\frac{\varepsilon^2 m}{2}\right)\prod_{h=1}^{m}\left((r_h+1)\exp\left(\frac{\varepsilon^2}{4}\right)\right), \ \text{ usando } \ 1+t\le e^t \\
&= (r_1+1)\cdots(r_m+1)\exp\left(-\frac{\varepsilon^2 m}{4}\right).
\end{aligned}$$

■

Lema 3.2.3 (SIEGEL) *Seja $A = (a_{ij})$ uma matriz com coeficientes inteiros, M linhas, N colunas e $N > M$. Então, existe um vetor $\underline{t} = (t_1, ..., t_N)$ não-nulo com coordenadas inteiras tal que $A\underline{t} = \underline{0}$ e satisfazendo*

$$\max_{1 \leq i \leq N} |t_i| \leq (N \cdot |A|)^{\frac{M}{N-M}},$$

em que $|A| = \max |a_{ij}|$ com $1 \leq i \leq M$ e $1 \leq j \leq N$.

Demonstração. Consideremos a aplicação

$$L : \mathbf{Z}^N \longrightarrow \mathbf{Z}^M$$

definida por

$$L(\underline{x}) = \left(\sum_{j=1}^{N} a_{1j}x_j, ..., \sum_{j=1}^{N} a_{Mj}x_j \right).$$

Para cada $i = 1, ..., M$ sejam

$$S_i^+ = \sum_{j=1}^{N} \max\{0, a_{ij}\} \quad \text{e} \quad S_i^- = \sum_{j=1}^{N} \min\{0, a_{ij}\}$$

de sorte que

$$S_i^+ - S_i^- = \sum_{j=1}^{N} |a_{ij}| \leq N \cdot |A|.$$

Seja $T = \left[(N \cdot |A|)^{\frac{M}{N-M}} \right] \geq 1$. Seja $I_T = \{0, 1, ..., T\}$ e consideremos os vetores inteiros $\underline{x} = (x_1, ..., x_N) \in I_T^N \subset \mathbf{Z}^N$; notemos que $\#(I_T^N) = (T+1)^N$.

Para $\underline{x} \in I_T^N$

$$S_i^- \cdot T \leq \sum_{j=1}^{N} a_{ij}x_j \leq S_i^+ \cdot T.$$

Como o intervalo $[S_i^- \cdot T, S_i^+ \cdot T]$ contém $S_i^+ \cdot T - S_i^- \cdot T + 1$ inteiros e $S_i^+ \cdot T - S_i^- \cdot T + 1 \leq N \cdot |A| \cdot T + 1$, a imagem de I_T^N por L consiste de no máximo $(N \cdot |A| \cdot T + 1)^M$ pontos.

Por definição de T, $T + 1 > (N \cdot |A|)^{\frac{M}{N-M}}$. Donde

$$\begin{aligned} \#(I_T^N) = (T+1)^N &= (T+1)^{N-M}(T+1)^M \\ &> (N \cdot |A| \cdot (T+1))^M \\ &\geq (N \cdot |A| \cdot T + 1)^M \\ &\geq \#(L(I_T^N)). \end{aligned}$$

Portanto, pelo princípio da casa dos pombos, existem dois vetores distintos $\underline{x}', \underline{x}'' \in I_T^N$ tais que $L(\underline{x}') = L(\underline{x}'')$. Tomemos então, $\underline{t} = \underline{x}' - \underline{x}'' \neq \underline{0}$. Então, $L(\underline{t}) = \underline{0}$ e

$$|t_i| \leq T \leq (N \cdot |A|)^{\frac{M}{N-M}} \quad \text{para} \quad i = 1, ..., N.$$

■

Proposição 3.2.1 *Sejam α um inteiro algébrico de grau d sobre $\mathbf{Q}$, $\varepsilon > 0$ uma constante fixada e m um inteiro satisfazendo*

$$\exp(\varepsilon^2 m/4) > 2d. \tag{3.1}$$

Sejam $r_1, ..., r_m$ inteiros positivos dados. Então existe um polinômio

$$P(X_1, ..., X_m) \in \mathbf{Z}[X_1, ..., X_m]$$

satisfazendo as seguintes condições:

(i) *P possui grau no máximo r_h na variável X_h.*

(ii) *O índice de P com respeito a $(\alpha, \alpha, ..., \alpha; r_1, ..., r_m)$ satisfaz*

$$\mathrm{Ind}P \geq \frac{m}{2}(1 - \varepsilon). \tag{3.2}$$

(iii) *O maior coeficiente de P satisfaz*

$$|P| \leq B^{r_1 + \cdots + r_m}, \tag{3.3}$$

onde $B = B(\alpha)$ é uma constante que depende apenas de α.

Demonstração. Escrevamos o polinômio P como

$$P(X_1, ..., X_n) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} p_{j_1,...,j_m} X_1^{j_1} \cdots X_m^{j_m},$$

onde os inteiros $p_{j_1,...,j_m}$ ainda serão determinados. O número de coeficientes $p_{j_1,...,j_m}$ é

$$N = (r_1 + 1) \cdots (r_m + 1).$$

Para qualquer m-upla $(i_1, ..., i_m)$ tem-se que

$$\partial_{i_1 ... i_m} P = P_{i_1 ... i_m} = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} p_{j_1,...,j_m} \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} X_1^{j_1 - i_1} \cdots X_m^{j_m - i_m}.$$

Avaliando esta identidade em $(\alpha, ..., \alpha)$ e usando o Lema 3.2.1 para expressar as potências de α como combinação linear de $1, \alpha, ..., \alpha^{d-1}$, obtemos

$P_{i_1\ldots i_m}(\alpha,\ldots,\alpha)$

$$
\begin{aligned}
&= \sum_{j_1=0}^{r_1}\cdots\sum_{j_m=0}^{r_m} p_{j_1,\ldots,j_m}\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}\alpha^{(j_1-i_1+\ldots+j_m-i_m)}\\
&= \sum_{j_1=0}^{r_1}\cdots\sum_{j_m=0}^{r_m} p_{j_1,\ldots,j_m}\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}\sum_{k=1}^{d} a_k^{(j_1+\ldots+j_m-i_1-\ldots-i_m)}\alpha^{d-k}\\
&= \sum_{k=1}^{d}\left\{\sum_{j_1=0}^{r_1}\cdots\sum_{j_m=0}^{r_m}\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}a_k^{(j_1+\ldots+j_m-i_1-\ldots-i_m)}p_{j_1,\ldots,j_m}\right\}\alpha^{d-k}.
\end{aligned}
$$

Teremos então $P_{i_1\ldots i_m}(\alpha,\ldots,\alpha)=0$ se nós escolhemos $p_{j_1,\ldots,j_m}$ para satisfazer as d equações lineares

$$\sum_{j_1=0}^{r_1}\cdots\sum_{j_m=0}^{r_m}\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}a_k^{(j_1+\ldots+j_m-i_1-\ldots-i_m)}p_{j_1,\ldots,j_m}=0,$$

$1\le k\le d$.

A fim de satisfazer a condição (ii), necessitamos $P_{i_1\ldots i_m}=0$ para todas as m-uplas $(i_1,\ldots,i_m)$ satisfazendo

$$\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\le\frac{m}{2}(1-\varepsilon)=\frac{m}{2}-\frac{\varepsilon}{2}m.$$

De acordo com o Lema 3.2.2, existem no máximo

$$(r_1+1)\cdots(r_m+1)\cdot\exp(-\varepsilon^2 m/4)$$

tais m-uplas. Portanto nós podemos encontrar um P que satisfaz (ii) escolhendo $p_{j_1,\ldots,j_m}$ para satisfazer um sistema de M equações lineares com coeficientes inteiros, onde, pela escolha de m na equação (3.1),

$$M\le d\cdot(r_1+1)\cdots(r_m+1)\cdot\exp(-\varepsilon^2 m/4)=dN\exp(-\varepsilon^2 m/4)\le\frac{1}{2}N.$$

Agora, nós temos M equações lineares para as N variáveis $p_{j_1,\ldots,j_m}$. Lembremos que do Lema 3.2.1 as quantidades $a_k^{(\ell)}$ satisfazem $|a_k^{(\ell)}|\le(|Q|+1)^\ell$, onde Q é o polinômio mínimo de α sobre $\mathbf{Q}$. Logo, nós podemos estimar os coeficientes dessas equações lineares por

$$
\begin{aligned}
\left|\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}a_k^{(j_1+\ldots+j_m-i_1-\ldots-i_m)}\right| &\le 2^{j_1+\ldots+j_m}(|Q|+1)^{j_1+\ldots+j_m}\\
&\le (2|Q|+2)^{r_1+\ldots+r_m}.
\end{aligned}
$$

Aplicando o Lema de Siegel, temos que existe um polinômio P satisfazendo (i) e (ii) cujo os coeficientes $p_{j_1,\ldots,j_m}$ são limitados por

$$\begin{aligned} |P| &\leq \left(N(2|Q|+2)^{r_1+\cdots+r_m}\right)^{M/(N-M)} \\ &\leq N(2|Q|+2)^{r_1+\cdots+r_m}, \quad \text{pois} \quad M \leq \frac{1}{2}N \\ &\leq 2^{r_1+\cdots+r_m}(2|Q|+2)^{r_1+\cdots+r_m} \\ &\leq B^{r_1+\cdots+r_m}, \quad B = 4|Q|+4. \end{aligned}$$

Portanto P também satisfaz (iii), completando a demonstração.

■

3.3 O índice é *grande*

Proposição 3.3.1 *Sejam $0 < \delta < 1$ uma constante dada e ε satisfazendo*

$$0 < \varepsilon < \frac{\delta}{22}. \tag{3.4}$$

Sejam α um inteiro algébrico de grau d sobre $\mathbf{Q}$, m um inteiro satisfazendo $e^{\varepsilon^2 m/4} > 2d$ e $r_1, \ldots, r_m$ inteiros positivos. Usemos a Proposição 3.2.1 para escolher um polinômio $P \in \mathbf{Z}[X_1, \ldots, X_m]$ satisfazendo as condições (3.2), (3.3) e (3.4).

Seja $S \subset M_K$ um conjunto finito de valores absolutos sobre K com cada valor absoluto estendido, de alguma maneira, a $\overline{K}$. Suponhamos que para cada $v \in S$ um número real $\xi_v \geq 0$ é dado tal que $\sum_{v\in S} \xi_v = 1$. Suponhamos que $\beta_1, \ldots, \beta_m \in K$ satisfazem

$$\|\beta_h - \alpha\|_v \leq \frac{1}{H_K(\beta_h)^{(2+\delta)\xi_v}} \tag{3.5}$$

para todo $v \in S$ e para todo $1 \leq h \leq m$. Suponhamos ainda que

$$\max_{1\leq h\leq m} \{H_K(\beta_h)^{r_h}\} \leq \min_{1\leq h\leq m} \{H_K(\beta_h)^{r_h}\}^{1+\varepsilon} \tag{3.6}$$

e que existe uma constante $C = C(\alpha, \delta)$ tal que

$$C \leq H(\beta_h), \quad 1 \leq h \leq m. \tag{3.7}$$

Então, o índice de P com respeito a $(\beta_1, \ldots, \beta_m; r_1, \ldots, r_m)$ satisfaz

$$\mathrm{Ind}_{(\beta_1,\ldots,\beta_m;r_1,\ldots,r_m)} P \geq \varepsilon m.$$

Lema 3.3.1 *Sejam $P \in \mathbf{Z}[X_1, ..., X_m]$ com $\mathrm{gr}_{X_h}(P) \leq r_h$ e $\beta = (\beta_1, ..., \beta_m)$ uma m-upla de números algébricos em um corpo de números K. Então para todas as m-uplas de inteiros não-negativos $j = (j_1, ..., j_m)$ tem-se que*

$$H_K(\partial_j P(\beta)) \leq 4^{(r_1+\cdots+r_m)[K:\mathbf{Q}]} H_K(P) \prod_{h=1}^{m} H_K(\beta_h)^{r_h}.$$

Demonstração. Seja $(j_1, ..., j_m)$ qualquer m-upla de inteiros não-negativos. Seja

$$T(X_1, ..., X_m) = \partial_{j_1 \cdots j_m} P(X_1, ..., X_m).$$

Já observamos que $T(X_1, ..., X_m)$ tem coeficientes inteiros e além disso

$$|T| \leq 2^{r_1+\cdots+r_m} |P|.$$

E observemos que como P possui coeficientes inteiros, $H_K(P) = |P|^{[K:\mathbf{Q}]}$.

Vamos usar as desigualdades triangular e ultramétrica para obter uma cota superior para a altura de $T(\beta_1, ..., \beta_m)$. Assim, seja $v \in M_K^\infty$ qualquer valor absoluto arquimediano, então
$|T(\beta_1, ..., \beta_m)|_v$

$$\begin{aligned} &\leq (r_1 + 1) \cdots (r_m + 1) \cdot |T| \cdot \max\{|\beta_1|_v, 1\}^{r_1} \cdots \max\{|\beta_m|_v, 1\}^{r_m} \\ &\leq 4^{r_1+\cdots+r_m} |P| \cdot \max\{|\beta_1|_v, 1\}^{r_1} \cdots \max\{|\beta_m|_v, 1\}^{r_m}. \end{aligned}$$

Se $v \in M_K^0$, a desigualdade ultramétrica mais o fato de que T possui coeficientes inteiros dão uma cota mais forte

$$|T(\beta_1, ..., \beta_m)|_v \leq \max\{|\beta_1|_v, 1\}^{r_1} \cdots \max\{|\beta_m|_v, 1\}^{r_m}.$$

Agora, elevamos essas desigualdades a $n_v = [K_v : \mathbf{Q}_v]$ e multiplicamos elas, juntas, para todo $v \in M_K$, isso nos leva a estimativa desejada

$$H_K(T(\beta_1, ..., \beta_m)) \leq 4^{(r_1\cdots+r_m)[K:\mathbf{Q}]} H_K(P) H_K(\beta_1)_1^r \cdots H_K(\beta_m)^{r_m}.$$

■

Lema 3.3.2 *Sejam $r_1, ..., r_m$ inteiros positivos dados e $P \in \mathbf{Z}[X_1, ..., X_m]$ um polinômio tal que $\mathrm{gr}_{X_h}(P) \leq r_h$. Denotemos por $\theta = \mathrm{Ind}P$ o índice de P com respeito a $(\alpha, ..., \alpha; r_1, ..., r_m)$. Seja $0 < \delta < 1$ uma constante e escolhamos $0 < \theta_0 < \theta$.*

Seja $S \subset M_K$ um conjunto finito de valores absolutos sobre K com cada valor absoluto estendido, de alguma maneira, a $\overline{K}$. Suponhamos que para cada $v \in S$ um número real $\xi_v \geq 0$ é dado tal que $\sum_{v \in S} \xi_v = 1$. Suponhamos que $\beta_1, ..., \beta_m \in K$ satisfazem

$$\|\beta_h - \alpha\|_v \leq \frac{1}{H_K(\beta_h)^{(2+\delta)\xi_v}}$$

para todo $v \in S$ e para todo $1 \leq h \leq m$. Denotemos por $D = \min\{H_K(\beta_h)^{r_h}\}$ e seja $j = (j_1, ..., j_m)$ qualquer m-upla de inteiros não-negativos satisfazendo

$$\sum_{h=1}^{m} \frac{j_h}{r_h} \leq \theta_0.$$

Então

$$\prod_{v \in S} \|\partial_j P(\beta_1, ..., \beta_m)\|_v \leq \frac{(8H(\alpha))^{[K:\mathbf{Q}](r_1+\cdots+r_m)} H_K(P)}{D^{(2+\delta)(\theta-\theta_0)}}.$$

Demonstração. Seja $j = (j_1, ..., j_m)$ como no enunciado acima e seja $T = \partial_j P$. Desejamos usar a expansão de Taylor de T em torno de $(\alpha, ..., \alpha)$, para isso, vamos estimar primeiramente os coeficientes da expansão. Seja $v \in M_K$ um valor absoluto sobre K (que pode ser estendido a $K(\alpha)$). Temos que $|\partial_{i_1 ... i_m} T(\alpha, ..., \alpha)|_v$ é uma soma de no máximo $(r_1 + 1) \cdots (r_m + 1)$ parcelas, cada uma limitada superiormente por

$$|T| \max\{|\alpha|_v, 1\}^{r_1+\cdots+r_m} \leq |P| (2\max\{|\alpha|_v, 1\})^{r_1+\cdots+r_m}.$$

Portanto

$$|\partial_{i_1 ... i_m} T(\alpha, ..., \alpha)|_v \leq (4\max\{|\alpha|_v, 1\})^{r_1+\cdots+r_m} |P|.$$

Observemos agora que T se anula com alta ordem em $(\alpha, ..., \alpha)$, mais precisamente, pelo item (a) do Lema 3.1.1

$$\mathrm{Ind}T = \mathrm{Ind}\partial_{j_1,...,j_m} P \geq \mathrm{Ind}P - \sum_{h=1}^{m} \frac{j_h}{r_h} \geq \theta - \theta_0.$$

Logo, a expansão de Taylor de T em torno de $(\alpha, ..., \alpha)$ terá muitos dos termos iniciais nulos, ou seja,

$$
\begin{aligned}
&T(X_1,...,X_m)\\
&\quad = \sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}\partial_{i_1,...,i_m}T(\alpha,...,\alpha)(X_1-\alpha)^{i_1}\cdots(X_m-\alpha)^{i_m}
\end{aligned}
$$

com $\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\geq\theta-\theta_0$.

Dessa forma, ponhamos $X_h=\beta_h$ e usemos o fato de que β_h está próximo de α. Então, para cada valor absoluto $v\in S$,

$$
\begin{aligned}
&|T(\beta_1,...,\beta_m)|_v\\
&\quad\leq \sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}|\partial_{i_1,...,i_m}T(\alpha,...,\alpha)|_v|\beta_1-\alpha|_v^{i_1}\cdots|\beta_m-\alpha|_v^{i_m}\\
&\quad\leq (r_1+1)\cdots(r_m+1)\max_{i_1,...,i_m}|\partial_{i_1,...,i_m}T(\alpha,...,\alpha)|_v\\
&\qquad\qquad\times\max_{\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\geq\theta-\theta_0}|\beta_1-\alpha|_v^{i_1}\cdots|\beta_m-\alpha|_v^{i_m}
\end{aligned}
$$

$$
\leq (8\max\{|\alpha|_v,1\})^{r_1+\cdots+r_m}|P|\max_{\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\geq\theta-\theta_0}\frac{1}{(H_K(\beta_1)^{i_1}\cdots H_K(\beta_m)^{i_m})^{(2+\delta)\xi_v}}.
$$

Podemos estimar o denominador dessa última desigualdade como segue:

$$
H_K(\beta_1)^{i_1}\cdots H_K(\beta_m)^{i_m}=(H_K(\beta_1)^{r_1})^{\frac{i_1}{r_1}}\cdots(H_K(\beta_m)^{r_m})^{\frac{i_m}{r_m}}\geq D^{\theta-\theta_0}.
$$

Logo, segue dessas estimativas acima que

$$
|T(\beta_1,...,\beta_m)|_v\leq\frac{(8\max\{|\alpha|_v,1\})^{r_1+\cdots+r_m}|P|}{D^{(\theta-\theta_0)(2+\delta)\xi_v}}.
$$

Finalmente, para chegarmos a estimativa do enunciado, elevamos a n_v, multiplicamos para todo $v\in S$ e usamos o fato de que $\sum_{v\in S}n_v\xi_v\geq\sum_{v\in S}\xi_v=1$, assim,

$$
\begin{aligned}
\prod_{v\in S}\|T(\beta_1,...,\beta_m)\|_v &\leq \frac{\prod_{v\in S}(8^{n_v}\max\{\|\alpha\|_v,1\})^{r_1+\cdots+r_m}|P|^{n_v}}{D^{(\theta-\theta_0)(2+\delta)}}\\
&\leq \frac{\prod_{v\in M_K}(8^{n_v}\max\{\|\alpha\|_v,1\})^{r_1+\cdots+r_m}|P|^{n_v}}{D^{(\theta-\theta_0)(2+\delta)}}\\
&= \frac{(8H(\alpha))^{[K:\mathbf{Q}](r_1+\cdots+r_m)}H_K(P)}{D^{(\theta-\theta_0)(2+\delta)}}.
\end{aligned}
$$

■

Demonstração da Proposição 3.3.1. Seja $j = (j_1, ..., j_m)$ uma m-upla de inteiros não-negativos satisfazendo $\sum_{h=1}^{m} j_h/r_h \leq \varepsilon m$. Queremos mostrar que $\partial_j P(\beta_1, ..., \beta_m) = 0$. Do resultado anterior, Lema 3.3.2, temos que

$$\begin{aligned}\prod_{v\in S} \|\partial_j P(\beta_1, ..., \beta_m)\|_v &\leq \frac{(8H(\alpha))^{[K:\mathbf{Q}](r_1+\cdots+r_m)} H_K(P)}{D^{(\theta-\theta_0)(2+\delta)}} \\ &\leq \frac{(8H(\alpha)B)^{[K:\mathbf{Q}](r_1+\cdots+r_m)}}{D^{(\frac{m}{2}(1-\varepsilon)-\varepsilon m)(2+\delta)}}.\end{aligned}$$

Para essa última desigualdade usamos as propriedades (3.3) e (3.4). Por outro lado, pelo Lema 3.3.1, temos que

$$\begin{aligned}H_K(\partial_j P(\beta_1, ..., \beta_m)) &\leq 4^{[K:\mathbf{Q}](r_1+\cdots+r_m)} H_K(P) \prod_{h=1}^{m} H_K(\beta_h)^{r_h} \\ &\leq (4B)^{[K:\mathbf{Q}](r_1+\cdots+r_m)} D^{m(1+\varepsilon)}.\end{aligned}$$

Aqui, para a última desigualdade, usamos mais uma vez (3.4) e a hipótese (3.7). Agora, a desigualdade de Liouville, Lema 1.1.1, diz que

$$\partial_j P(\beta_1, ..., \beta_m) = 0$$

ou

$$\prod_{v\in S} \|\partial_j P(\beta_1, ..., \beta_m)\|_v \geq \frac{1}{H_K(\partial_j P(\beta_1, ..., \beta_m))}.$$

Vamos mostrar que nossas hipóteses contradizem essa última, possível, desigualdade.

Suponhamos que $\partial_j P(\beta_1, ..., \beta_m) \neq 0$, logo, a desigualdade de Liouville garante que

$$\frac{(8H(\alpha)B)^{[K:\mathbf{Q}](r_1+\cdots+r_m)}}{D^{(\frac{m}{2}(1-\varepsilon)-\varepsilon m)(2+\delta)}} \geq \frac{1}{(4B)^{[K:\mathbf{Q}](r_1+\cdots+r_m)} D^{m(1+\varepsilon)}}$$

logo

$$D^{m((1+\delta/2)(1-3\varepsilon)-(1+\varepsilon))} \leq (32H(\alpha)B^2)^{[K:\mathbf{Q}](r_1+\cdots+r_m)}.$$

Como assumimos $\delta < 1$ e $\varepsilon < \delta/22$ temos que

$$(1+\delta/2)(1-3\varepsilon) - (1+\varepsilon) < \delta/2 - 11\varepsilon/2 < \delta/4 < 1,$$

portanto

$$D^{m\delta/4} \leq (32H(\alpha)B^2)^{[K:\mathbf{Q}](r_1+\cdots+r_m)}$$

e

$$\max_{1\leq h\leq m}\{H_K(\beta_h)^{r_h}\} \leq D^{1+\varepsilon} \leq (32H(\alpha)B^2)^{4[K:\mathbf{Q}](r_1+\cdots+r_m)(1+\varepsilon)/m\delta}.$$

Selecionando j tal que $r_j = \max_{1\leq h\leq m} r_h$, obtemos que

$$H_K(\beta_j) \leq (32H(\alpha)B^2)^{4[K:\mathbf{Q}](1+\varepsilon)/\delta}.$$

Finalmente, escolhendo a constante C da hipótese (3.8) suficientemente grande temos a contradição desejada e com isso concluímos a demonstração.

■

3.4 O índice é *pequeno*

Nesta seção vamos mostrar que não é possível que P se anule com alta ordem em $(\beta_1, ..., \beta_m)$, o que dará uma contradição com a existência de infinitas aproximações de α.

Sejam $f_1(X), ..., f_n(X) \in K(X)$ funções racionais de uma variável. O determinante Wronskiano *clássico* de $f_1(X), ..., f_n(X)$ é a função

$$W(f_1, ..., f_n) = \det \begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ \frac{df_1}{dX} & \frac{df_2}{dX} & \cdots & \frac{df_n}{dX} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{d^{n-1}f_1}{dX^{n-1}} & \frac{d^{n-1}f_2}{dX^{n-1}} & \cdots & \frac{d^{n-1}f_n}{dX^{n-1}} \end{pmatrix}.$$

É bem conhecido que as funções $f_1, ..., f_n$ são linearmente independentes sobre K se, e somente se, $W(f_1, ..., f_n) \neq 0$. Vamos precisar de uma versão que se aplique a polinômios de várias variáveis.

Definição 3.4.1 *A* **ordem** *de um operador diferencial*

$$\Delta = \frac{1}{i_1! \cdots i_m!} \frac{\partial^{i_1+\cdots+i_m}}{\partial X_1^{i_1} \cdots \partial X_m^{i_m}}$$

é a quantidade $\text{ordem}(\Delta) = i_1 + \cdots + i_m$. *Sejam* $\phi_1, ..., \phi_k \in K(X_1, ..., X_m)$ *funções racionais. Um determinante Wronskiano* **generalizado** *de* $\phi_1, ..., \phi_k$ *é qualquer determinante da forma*

$$\det\left((\Delta_i\phi_j)_{1\leq i,j\leq k}\right),$$

em que o operador diferencial Δ_i *satisfaz* $\text{ordem}(\Delta_i) \leq i - 1$.

O próximo lema generaliza o resultado referente a independência linear.

Lema 3.4.1 *As funções $\phi_1, ..., \phi_k$ são linearmente independentes sobre K se, e somente se, existe um determinante Wronskiano generalizado não-nulo de $\phi_1, ..., \phi_k$.*

Demonstração. A implicação que precisamos para a prova do Lema de Roth abaixo é que se $\phi_1, ..., \phi_k$ são linearmente independentes, então existe um Wronskiano generalizado não-nulo para elas.

A demonstração seguirá por indução sobre k, o número de funções. Suponhamos que $\phi_1, ..., \phi_k$ são K-linearmente independentes. Para $k = 1$ o único determinante Wronskiano generalizado é um múltiplo constante de $\Delta_1\phi_1 = \phi_1$, pois Δ_1 tem ordem nula. Portanto para $k = 1$ este Lema 3.4.1 diz que ϕ_1 é linearmente independente sobre K se, e somente se, $\phi_1 \neq 0$.

Suponhamos agora que o resultado seja válido para qualquer conjunto de $k - 1$ funções. Precisamos encontrar um determinante Wronskiano generalizado não-nulo para $\phi_1, ..., \phi_k$. Se conseguirmos mostrar que algum Wronskiano generalizado de $\lambda\phi_1, ..., \lambda\phi_k$ em que $0 \neq \lambda \in K(X_1, ..., X_m)$, é não-nulo, então seguirá que algum Wronskiano generalizado de $\phi_1, ..., \phi_k$ é não-nulo. Então tomando $\lambda = 1/\phi_1$, podemos nos reduzir ao caso em que $\phi_1 = 1$.

Seja

$$V = K\phi_1 + \cdots + K\phi_k \subset K(X_1, ..., X_m)$$

o K-subespaço vetorial gerado por $\phi_1, ..., \phi_k$. Por hipótese a dimensão de V é igual a k. Em particular, ϕ_2 não é constante pois $\phi_1 = 1$, portanto podemos assumir que a variável X_1 aparece em ϕ_2, ou, podemos assumir que $\partial\phi_2/\partial X_1 \neq 0$.

Consideremos agora um K-subespaço vetorial de V definido por

$$W = \{\phi \in V | \partial\phi/\partial X_1 = 0\}$$

e seja t a sua dimensão. Como $\phi_1 \in W$ e ϕ_2 não está em W, $1 \leq t \leq k - 1$. Escolhamos $\psi_1, ..., \psi_t$ uma base de W e estendamos a uma de V $\psi_1, ..., \psi_k$. Por hipótese de indução, existem operadores diferenciais $\Delta_1^*, ..., \Delta_t^*$ satisfazendo

$$\det(\Delta_i^*\psi_j)_{1\leq i,j\leq t} \neq 0, \quad \text{com} \quad \text{ordem}(\Delta_i^*) \leq i - 1.$$

Observemos que as funções $\partial\psi_{t+1}/\partial X_1, ..., \partial\psi_k/\partial X_1$ são K-linearmente independentes pois $\{\psi_{t+1}, ..., \psi_k\}$ é uma base de V/W. Portanto podemos aplicar novamente a hipótese indutiva e encontrar operadores diferenciais $\Delta_{t+1}^*, ..., \Delta_k^*$ satisfazendo

$$\det\left(\Delta_i^*\frac{\partial\psi_j}{\partial X_1}\right)_{t+1\leq i,j\leq k} \neq 0, \quad \text{com} \quad \text{ordem}(\Delta_i^*) \leq i - t - 1.$$

Juntemos esses dois determinantes para definir os seguintes operadores diferenciais

$$\Delta_i = \begin{cases} \Delta_i^* & \text{se } 1 \leq i \leq t \\ \Delta_i^* \frac{\partial}{\partial X_1} & \text{se } t+1 \leq i \leq k. \end{cases}$$

Temos que ordem$(\Delta_i) \leq i-1$, para $1 \leq i \leq k$. Além disso, e é crucial, $\psi_j \in W$, para $1 \leq j \leq t$, logo

$$\Delta_i \psi_j = \Delta_i^* \frac{\partial \psi_j}{\partial X_1} = 0,$$

para $t+1 \leq i \leq k$. Portanto

$$\det(\Delta_i \psi_j)_{1 \leq i,j \leq k} = \det(\Delta_i^* \psi_j)_{1 \leq i,j \leq t} \cdot \det\left(\Delta_i^* \frac{\partial \psi_j}{\partial X_1}\right)_{t+1 \leq i,j \leq k} \neq 0.$$

Com isso mostramos que $\psi_1, ..., \psi_k$ possuem um Wronskiano generalizado não-nulo. Mas as funções $\phi_1, ..., \phi_k$ e $\psi_1, ..., \psi_k$ geram o mesmo K-espaço vetorial, logo $\psi_j = \sum_\ell a_{j\ell} \psi_\ell$ para alguma matriz inversível $(a_{j\ell})$ com entradas em K. Donde segue que

$$0 \neq \det(\Delta_i \psi_j) = \det\left(\sum_\ell a_{j\ell} \Delta_i \phi_\ell\right) = \det(a_{j\ell}) \det(\Delta_i \phi_\ell),$$

e portanto $\det(\Delta_i \phi_\ell) \neq 0$.

A outra implicação é mais fácil: se as funções $\phi_1, ..., \phi_k$ são linearmente dependentes sobre K, então todos os determinantes Wronskianos generalizados se anulam!

■

Antes de enunciar o próximo lema, lembremos que se $\beta \in K$ é um número algébrico,

$$h(\beta) = \log H(\beta) = \frac{1}{[K : \mathbf{Q}]} \log H_K(\beta)$$

e para um polinômio P,

$$h(P) = \log H(P) = \frac{1}{[K : \mathbf{Q}]} \sum_{v \in M_K} n_v \log |P|_v.$$

Lema 3.4.2 (ROTH) *Sejam m um inteiro positivo e $P \in \overline{\mathbf{Q}}[X_1, ..., X_m]$ um polinômio com coeficientes algébricos com $\mathrm{gr}_{X_j}(P) \leq r_j$. Seja $(\beta_1, ..., \beta_m)$ uma m-upla de números algébricos e seja $\eta > 0$ um número real fixado tal que*

$$\frac{r_{j+1}}{r_j} \leq \eta^{2^{m-1}} \quad \textit{para todo} \quad 1 \leq j \leq m-1, \tag{3.8}$$

e

$$\eta^{2^{m-1}} \min_{1 \leq j \leq m} \{r_j h(\beta_j)\} \geq h(P) + 2mr_1. \tag{3.9}$$

Então o índice de P com respeito a $(\beta_1, ..., \beta_m; r_1, ..., r_m)$ satisfaz

$$\mathrm{Ind}_{(\beta_1,...,\beta_m;r_1,...,r_m)}(P) \leq 2m\eta.$$

Demonstração. A demonstração seguirá por indução sobre m, o número de variáveis. Consideremos o caso em que $m = 1$. Para facilitar a notação, escrevamos aqui, $\beta = \beta_1$ e $r = r_1$. Seja ℓ a ordem exata de anulamento de $P(X)$ em $X = \beta$, assim $P(X) = (X - \beta)^\ell Q(X)$ com $Q(\beta) \neq 0$. Neste caso o índice de P com respeito a $(\beta; r)$ é $\mathrm{Ind}_{(\beta;r)}(P) = \ell/r$. Usando a desigualdade de Gelfand, Proposição 1.2.2, obtemos a seguinte estimativa

$$H(\beta)^{r\mathrm{Ind}(P)} = H(\beta)^\ell = H(X - \beta)^\ell \leq H(X - \beta)^\ell H(Q) \leq H(P) \cdot e^r,$$

que aplicando o logaritmo obtemos

$$\begin{aligned} \mathrm{Ind}(P) &\leq \frac{h(P) + r}{rh(\beta)} \\ &\leq \eta \ \text{ usando } \ (3.9). \end{aligned}$$

O que completa a demonstração no caso de polinômios de uma variável.

A fim de continuar a demonstração escrevamos P na forma

$$P(X_1, ..., X_m) = \sum_{j=1}^{k} \phi_j(X_1, ..., X_{m-1}) \psi_j(X_m), \tag{3.10}$$

onde as funções ϕ_j e ψ_j são polinômios com coeficientes em $\overline{\mathbf{Q}}$. Entre as muitas maneiras de decompor P, escolhemos entre elas, aquela em que k é o menor possível, isto é, escolhemos uma decomposição (3.10) com o menor número de parcelas, mínima. Observemos agora que $\psi_1 = 1, \psi_2 = X_m$, $\psi_3 = X_m^2, ..., \psi_k = X_m^{r_m}$ é uma possível decomposição, donde obtemos que

$$k \leq r_m + 1.$$

A minimalidade de k garante o seguinte.

Afirmação 1. *Os polinômios $\phi_1, ..., \phi_k$ aparecendo na decomposição, mínima, (3.10) de P são linearmente independentes sobre $\overline{\mathbf{Q}}$. Similarmente, $\psi_1, ..., \psi_k$ são $\overline{\mathbf{Q}}$-linearmente independentes.*

Prova (da afirmação 1). De fato, suponhamos que os ϕ_j's são linearmente dependentes, logo, existe uma relação linear não-trivial $\sum c_j\phi_j = 0$, digamos, com $c_k \neq 0$. Então

$$\phi_k = -\sum_{j=1}^{k-1} \frac{c_j}{c_k}\phi_j,$$

e portanto

$$P = \sum_{j=1}^{k} \phi_j\psi_j = \sum_{j=1}^{k-1} \phi_j\psi_j - \sum_{j=1}^{k-1} \frac{c_j}{c_k}\phi_j\psi_j = \sum_{j=1}^{k-1} \phi_j \left(\psi_j - \frac{c_j}{c_k}\psi_k\right),$$

o que contradiz a minimalidade de k. A prova para $\psi_1, ..., \psi_k$ é a mesma.

■

Definamos um polinômio $U(X_m)$ por

$$U(X_m) := \det\left(\frac{1}{(i-1)!}\frac{\partial^{i-1}}{\partial X_m^{i-1}}\psi_j(X_m)\right)_{1\leq i,j\leq k}.$$

Este é o determinante Wronskiano de $\psi_1, ..., \psi_k$, portanto, pelo Lema 3.3.1 e pela Afirmação 1 temos que $U(X_m) \neq 0$. Também, o Lema 3.3.1 e a $\overline{\mathbf{Q}}$-independência linear de $\phi_1, ..., \phi_k$ implicam que existe operador diferencial

$$\Delta_i' = \frac{1}{i_1! \cdots i_{m-1}!}\frac{\partial^{i_1+\cdots+i_{m-1}}}{\partial X_1^{i_1} \cdots \partial X_{m-1}^{i_{m-1}}}$$

com

$$\text{ordem}(\Delta_i') = i_1 + \cdots + i_{m-1} \leq i - 1 \leq k - 1 \leq r_m$$

tal que o determinante Wronskiano generalizado, $\det(\Delta_i'\phi_j)_{1\leq i,j\leq k}$, é não-nulo. Definimos

$$V(X_1, ..., X_{m-1}) := \det(\Delta_i'\phi_j)_{1\leq i,j\leq k} \neq 0.$$

Vamos agora definir um polinômio com m variáveis $W(X_1, ..., X_m)$.

$$W(X_1, ..., X_m) := \det\left(\Delta_i'\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial X_m^{j-1}}P(X_1, ..., X_m)\right)_{1\leq i,j\leq k}$$

$$
\begin{aligned}
&= \det\left(\Delta_i' \frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial X_m^{j-1}}\sum_{r=1}^{k}\phi_r(X_1,...,X_{m-1})\psi_r(X_m)\right)_{1\le i,j\le k}\\
&= \det\left(\sum_{r=1}^{k}\Delta_i'\phi_r\cdot\frac{1}{(j-1)!}\frac{\partial^{j-1}\psi_r}{\partial X_m^{j-1}}\right)_{1\le 1,j\le k}\\
&= \det(\Delta_i'\phi_r)_{1\le i,r\le k}\cdot\det\left(\frac{1}{(j-1)!}\frac{\partial^{j-1}\psi_r}{\partial X_m^{j-1}}\right)_{1\le j,r\le k}\\
&= V(X_1,...,X_{m-1})U(X_m).
\end{aligned}
$$

A penúltima igualdade acima foi obitida por multiplicação matricial! Assim, o uso dos determinantes Wronskianos serviu para criar um polinômio W intimamente relacionado com P e que se fatora em um produto de dois polinômios, cada um, envolvendo menos variáveis do que P. A continuação da demonstração consiste basicamente de dois passos fundamentais, primeiro, obter uma cota superior para os índices de U e V, depois, obter uma cota inferior para o índice de W em termos do índice de P.

Afirmação 2. *As seguintes estimativas são válidas:*

(a) $\mathrm{gr}_{X_m}(U) \le kr_m$ *e* $\mathrm{gr}_{X_j}(V) \le kr_j$ *para todo* $1 \le j \le m-1$.

(b) $h(U) + h(V) = h(W) \le k(h(P) + 2r_1)$.

Prova (da afirmação 2). (a) Cada determinante é de tamanho k e as entradas de V (respectivamente U) possuem grau no máximo r_j em X_j (respectivamente no máximo r_m em X_m).
(b) Primeiro observamos que como U e V usam conjuntos de variáveis disjuntos, segue da definição da altura que

$$h(U) + h(V) = h(W).$$

Agora, o determinante é uma soma de $k!$ parcelas, cada uma sendo o produto de k polinômios de grau no máximo r_j com respeito a variável X_j e satisfazendo

$$H(\Delta_i'\partial_j P(X_1,...,X_m)) \le 2^{r_1+\cdots+r_m}H(P).$$

Donde, pela Proposição 1.2.1, obtemos a cota superior

$$h(W) \le k(h(P) + (r_1+\cdots+r_m)\log 2) + \log(k!).$$

Temos também que

$$r_1+\cdots+r_m \le r_1(1+\eta'+\cdots+\eta'^{m-1}), \quad \text{com} \quad \eta' = \eta^{2^{m-1}}.$$

Como $\eta \leq \frac{1}{2}$ e $m \geq 2$, temos que $\eta' \leq \frac{1}{4}$ e $r_1 + \cdots + r_m \leq \frac{4}{3} r_1$. Por outro lado,

$$\frac{\log(k!)}{k} \leq \log k \leq k - 1 \leq r_m \leq \frac{1}{2} r_1,$$

portanto

$$h(W) \leq k \left(h(P) + \left(\frac{4}{3} \log 2 + \frac{1}{2} \right) r_1 \right) \leq k(h(P) + 2r_1).$$

■

Agora usaremos a indução para obter um cota para o índice de U, V e W.

Afirmação 3. *Se o Lema de Roth é válido para polinômios com $m-1$ ou menos variáveis, então*

$$\mathrm{Ind}_{(\beta_m, r_m)}(U) \leq k\eta^{2^{m-1}} \quad \text{e} \quad \mathrm{Ind}_{(\beta_1, \ldots, \beta_{m-1}; r_1, \ldots, r_{m-1})}(V) \leq 2k(m-1)\eta^2.$$

Portanto o índice de W com respeito a $(\beta_1, \ldots, \beta_m; r_1, \ldots, r_m)$ satisfaz

$$\begin{aligned} \mathrm{Ind}W &= \mathrm{Ind}_{(\beta_m, r_m)}(U) + \mathrm{Ind}_{(\beta_1, \ldots, \beta_{m-1}; r_1, \ldots, r_{m-1})}(V) \\ &\leq 2k(m-1)\eta^2 + k\eta^{2^{m-1}}. \end{aligned}$$

Prova (da afirmação 3). Vamos primeiro aplicar o Lema de Roth a V polinômio com $m' = m-1$ variáveis, com $r'_j = kr_j$ e $\eta' = \eta^2$. Já sabemos, da afirmação anterior, que $\mathrm{gr}_{X_j}(V) \leq r'_j$. Temos que checar as condições (3.9) e (3.10). Verifiquemos a condição (3.9),

$$\frac{r'_{j+1}}{r'_j} = \frac{r_{j+1}}{r_j} \leq \eta^{2^{m-1}} = \eta'^{2^{m'-1}}.$$

Condição (3.10):

$$r'_j h(\beta_j) = kr_j h(\beta_j) \geq k\eta^{-2^{m-1}}(h(P) + 2mr_1) = \eta'^{-2^{m'-1}} k(h(P) + 2mr_1).$$

Observamos agora que

$$h(V) + 2m'r'_1 = h(V) + 2(m-1)kr_1 \leq kh(P) + k2r_1 + 2(m-1)kr_1 = k(h(P) + 2mr_1),$$

a desigualdade vem do item (b) da Afirmação 2, assim, completamos a verificação de (3.10). Portanto

$$\mathrm{Ind}_{(\beta_1, \ldots, \beta_{m-1}; r_1, \ldots, r_{m-1})}(V) = k\mathrm{Ind}_{(\beta_1, \ldots, \beta_{m-1}; r'_1, \ldots, r'_{m-1})}(V) \leq k(2m'\eta') = 2k(m-1)\eta^2,$$

em que a primeira igualdade foi obtida indutivamente.

Do mesmo modo, só que agora, apliquemos o Lema de Roth em uma variável a U com $\eta'' = \eta^{2^{m-1}}$ e $r'' = kr_m$. Temos que $\mathrm{gr}_{X_m}(U) \leq r''$ pela Afirmação 2. Neste caso, como $m = 1$, precisamos checar apenas a condição (3.10). Assim

$$h(U) + r'' \leq k(h(P) + 2r_1) \leq \eta^{2^{m-1}} k r_m h(\beta_m) = \eta'' r'' h(\beta_m).$$

Donde concluímos que

$$\mathrm{Ind}_{(\beta_m; r_m)}(U) = k\mathrm{Ind}_{(\beta_m; r'')}(U) \leq k\eta'' = k\eta^{2^{m-1}}.$$

■

Afirmação 4.

$$\mathrm{Ind}W \geq \frac{k}{2}\min\{\mathrm{Ind}P, (\mathrm{Ind}P)^2\} - k\frac{r_m}{r_{m-1}}.$$

Prova (da afirmação 4). Vamos começar estimando o índice, com respeito a $(\beta_1, ..., \beta_m; r_1, ..., r_m)$, de uma entrada típica da matriz que define W

$$\begin{aligned}
&\mathrm{Ind}\left(\Delta_i'\left(\tfrac{1}{(j-1)!}\tfrac{\partial^{j-1}}{\partial X_m^{j-1}}P\right)\right) \\
&= \mathrm{Ind}\partial_{i_1,...,i_m,j-1}P \\
&\geq \mathrm{Ind}P - \frac{i_1}{r_1} - \cdots - \frac{i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m}, \text{ pelo item (a) do Lema 3.1.1} \\
&\geq \mathrm{Ind}P - \frac{i_1 + \cdots + i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m} \text{ pois } r_1 \geq r_2 \geq \cdots \text{ de (3.9)} \\
&\geq \mathrm{Ind}P - \frac{r_m}{r_{m-1}} - \frac{j-1}{r_m}
\end{aligned}$$

pois $\mathrm{ordem}(\Delta_i') = i_1 + \cdots + i_{m-1} \leq i - 1 \leq k - 1 \leq r_m$.

Como cada entrada da $j-$ésima coluna da matriz que define W tem a forma $\partial_{i_1,...,i_m,j-1}P$ e W pode ser visto como uma soma de $k!$ parcelas, em que cada parcela é o produto de k polinômios, um de cada coluna da matriz que define W; a estimativa acima dá uma cota inferior para o índice de cada entrada da matriz que define W. Portanto o índice de W com respeito a $(\beta_1, ..., \beta_m; r_1, ..., r_m)$ satisfaz

$$\begin{aligned}
\mathrm{Ind}W &\geq \min_{k!}\left\{\mathrm{Ind}\left(\prod_{j=1}^{k}\partial_{i_1,...,i_{m-1},j-1}P\right)\right\} \\
&\geq \sum_{j=1}^{k}\min_{i_1,...,i_{m-1}}\mathrm{Ind}\partial_{i_1,...,i_{m-1},j-1}P.
\end{aligned}$$

Para essas duas desigualdades usamos os itens (b) e (c) do Lema 3.1.1.

Substituindo a cota inferior, obtida antes, para $\mathrm{Ind}\partial_{i_1,\ldots,i_m,j-1}P$, naqueles casos em que é positiva, obtemos que

$$\begin{aligned}\mathrm{Ind}W &\geq \sum_{j=1}^{k}\max\left\{\mathrm{Ind}P-\frac{r_m}{r_{m-1}}-\frac{j-1}{r_m},0\right\}\\ &\geq \sum_{j=1}^{k}\max\left\{\mathrm{Ind}P-\frac{j-1}{r_m},0\right\}-\frac{kr_m}{r_{m-1}}.\end{aligned}$$

Para terminar vamos mostrar que

$$\sum_{j=1}^{k}\left(\mathrm{Ind}P-\frac{j-1}{r_m}\right)\geq\frac{k}{2}\min\{\mathrm{Ind}P,(\mathrm{Ind}P)^2\}.$$

Para fazer isto vamos separar em dois casos:

$$\textbf{Caso 1}.\ \ \mathrm{Ind}P\geq\frac{k-1}{r_m}.$$

Neste caso

$$\sum_{j=1}^{k}\left(\mathrm{Ind}P-\frac{j-1}{r_m}\right)=k\mathrm{Ind}P-\frac{(k-1)k}{2r_m}\geq\frac{k}{2}\mathrm{Ind}P.$$

$$\textbf{Caso 2}.\ \ \mathrm{Ind}P\leq\frac{k-1}{r_m}.$$

Seja $N=[r_m\mathrm{Ind}P]$, assim, nossa hipótese implica que $N\leq k-1$. Então

$$\begin{aligned}&\sum_{j=1}^{N+1}\left(\mathrm{Ind}P-\frac{j-1}{r_m}\right)\\ &\qquad=\ (N+1)\mathrm{Ind}P-\frac{N(N+1)}{2r_m}\\ &\qquad=\ (N+1)\left(\mathrm{Ind}P-\frac{[r_m\mathrm{Ind}P]}{2r_m}\right)\\ &\qquad\geq\ (N+1)\cdot\frac{1}{2}\mathrm{Ind}P\\ &\qquad\geq\ r_m\mathrm{Ind}P\cdot\frac{1}{2}\mathrm{Ind}P\\ &\qquad\geq\ \frac{k}{2}(\mathrm{Ind}P)^2,\ \ \text{se}\ \ k\leq r_m.\end{aligned}$$

Resta apenas a possibilidade $k = r_m + 1$. Estimemos então a quantidade

$$q(N) = \sum_{j=1}^{N+1} \left(\mathrm{Ind}P - \frac{j-1}{r_m} \right) = (N+1)\mathrm{Ind}P - \frac{N(N+1)}{2(k-1)}.$$

Observemos que $q(N)$ é uma função quadrática de N e que

$$(k-1)\mathrm{Ind}P - 1 \leq N \leq (k-1)\mathrm{Ind}P.$$

Agora, um simples cálculo, dá

$$q((k-1)\mathrm{Ind}P - 1) = q((k-1)\mathrm{Ind}P) = \frac{(k-1)(\mathrm{Ind}P)^2 + \mathrm{Ind}P}{2}.$$

Portanto, como neste caso $\mathrm{Ind}P \leq 1$, obtemos

$$q(N) \geq \frac{(k-1)(\mathrm{Ind}P)^2 + \mathrm{Ind}P}{2} \geq \frac{k}{2}(\mathrm{Ind}P)^2.$$

■

Podemos agora concluir a prova do Lema de Roth usando as cotas superior e inferior para $\mathrm{Ind}W$ obtidas, respectivamente, nas Afirmações 3 e 4. Como $\mathrm{Ind}P \leq m$, usando a Afirmação 4 podemos escrever

$$\mathrm{Ind}W + \frac{kr_m}{r_{m-1}} \geq \frac{k}{2}\min\{\mathrm{Ind}P, (\mathrm{Ind}P)^2\} \geq \frac{k(\mathrm{Ind}P)^2}{2m},$$

por outro lado, a Afirmação 3 dá

$$\begin{aligned} \mathrm{Ind}W + \frac{kr_m}{r_{m-1}} &\leq 2k(m-1)\eta^2 + k\eta^{2^{m-1}} + \frac{kr_m}{r_{m-1}} \\ &\leq k\left(2(m-1)\eta^2 + 2\eta^{2^{m-1}}\right) \leq k(2\eta^2 m). \end{aligned}$$

Donde obtemos que $(\mathrm{Ind}P)^2 \leq 4\eta^2 m^2$. Portanto $\mathrm{Ind}P \leq 2m\eta$.

■

3.5 Conclusão da prova

Temos agora todos os pedaços necessários para completar a prova do teorema de Roth, mais precisamente, para provar o Teorema 2.4.1, que provamos ser equivalente ao teorema de Roth. Vamos reenunciá-lo aqui.

Teorema 2.4.1 *Sejam K um corpo de números, $S \subset M_K$ um conjunto finito de valores absolutos sobre K com cada valor absoluto estendido, de alguma maneira, a $\overline{K}$. Sejam $\alpha \in \overline{K}$ e δ dados. Suponhamos que para cada $v \in S$ um número real $\xi_v \geq 0$ é dado tal que $\sum_{v \in S} \xi_v = 1$. Então existe somente uma quantidade finita de números algébricos $\beta \in K$ com a propriedade que*

$$\|\beta - \alpha\|_v \leq \frac{1}{H_K(\beta)^{(2+\delta)\xi_v}} \quad \textit{para todo} \quad v \in S. \tag{3.11}$$

Demonstração. Vamos assumir que existe uma quantidade infinita de soluções de (3.12) e vamos obter uma contradição. Precisaremos nos referir as várias condições descritas nas Proposições 3.2.1 e 3.3.1 e no Lema de Roth (Lema 3.4.2) por isso listaremos elas aqui. A constante $B = B(\alpha)$ foi definida na Proposição 3.2.1 e a constante $C = C(\alpha, \delta)$ na Proposição 3.3.1:

(3.2) $e^{\varepsilon^2 m/4} > 2[\mathbf{Q}(\alpha) : \mathbf{Q}]$.

(3.3) $\mathrm{Ind}_{(\alpha,\ldots,\alpha;r_1,\ldots,r_m)} P \geq \frac{m}{2}(1-\varepsilon)$.

(3.4) $|P| \leq B^{r_1+\cdots+r_m}$.

(3.5) $0 < \varepsilon < \frac{\delta}{22}$.

(3.6) $\|\beta_h - \alpha\|_v \leq \frac{1}{H_K(\beta_h)^{(2+\delta)\xi_v}}$.

(3.7) $D = \min\limits_{1\leq h\leq m} \{H(\beta_h)^{r_h}\} \leq \max\limits_{1\leq h\leq m} \{H(\beta_h)^{r_h}\} \leq D^{1+\varepsilon}$.

(3.8) $H(\beta_h) \geq C, \quad 1 \leq h \leq m$.

(3.9) $r_{h+1} \leq \omega r_h, \quad 1 \leq h \leq m-1$.

(3.10) $\log|P| + 2mr_1 \leq \omega \log D$.

Agora suponhamos que existem infinitas soluções para a desigualdade (3.12). Vamos escolher as quantidades

$$\varepsilon, m, \omega, \beta_1, \ldots, \beta_m, r_1, \ldots, r_m, P(X_1, \ldots, X_m)$$

como segue:

(1) Escolhamos um ε com $0 < \varepsilon < \delta/22$. Então ε satisfaz (3.5) e $\varepsilon < 1/22 < 1$.

(2) Escolhamos um inteiro m com $e^{\varepsilon^2 m/4} > 2[\mathbf{Q}(\alpha) : \mathbf{Q}]$. Então (3.2) vale. Definimos $\omega = \omega(m, \varepsilon) = (\varepsilon/4)^{2^{m-1}}$ o que implica $2\omega^{2^{-m+1}} = \varepsilon/2 < \varepsilon$.

(3) Como estamos assumindo que (3.12) tem uma quantidade infinita de soluções em K e como K possui apenas uma quantidade finita de elementos de altura limitada (Northcott, Teorema 1.2.1), podemos encontrar um solução β_1 satisfazendo

$$H(\beta_1) \geq C \quad \text{e} \quad \log H(\beta_1) \geq \frac{m(\log B + 2)}{\omega}.$$

(4) Escolhemos então sucessivamente $\beta_2, ..., \beta_m$ soluções de (3.12) satisfazendo

$$H_K(\beta_{h+1})^{\omega} \geq H_K(\beta_h)^2, \quad 1 \leq h < m,$$

donde verificamos (3.6). Como $\omega < 1$, teremos que $H_K(\beta_h) \geq H_K(\beta_1)$. Da escolha feita em (3) temos que (3.8) é satisfeita. Aqui está a razão da natureza inefetiva do método!

(5) Escolhamos um inteiro r_1 satisfazendo $H_K(\beta_1)^{\omega r_1} \geq H_K(\beta_m)^2$.

(6) Vamos escolher $r_2, ..., r_m$ de modo que todas as alturas $H_K(\beta_h)^{r_h}$ sejam aproximadamente as mesmas. Assim, definimos $r_2, ..., r_m$ como sendo os inteiros

$$r_h = \left[\frac{r_1 \log H_K(\beta_1)}{\log H_K(\beta_h)}\right] = \left[\frac{r_1 \log H(\beta_1)}{\log H(\beta_h)}\right],$$

em que $[\cdot]$ denota o menor inteiro maior do que ou igual. Chequemos agora as condições que aparecem em (3.7):
$r_1 \log H_K(\beta_1)$

$\leq$ $r_h \log H_K(\beta_h)$, $[t] \geq t$
$\leq$ $r_1 \log H_K(\beta_1) + \log H_K(\beta_h)$, $[t] \leq t + 1$
$\leq$ $r_1 \log H_K(\beta_1) + \log H_K(\beta_m)$, pois de (4) as $H_K(\beta_h)$ crescem
$\leq$ $(1 + \varepsilon) r_1 \log H_K(\beta_1)$, da escolha de r_1 em (5).

Exponenciando obtemos (3.7). Verificamos aqui ainda (3.9):

$$\frac{r_{h+1}}{r_h} = \frac{\left[\frac{r_1 \log H(\beta_1)}{\log H(\beta_{h+1})}\right]}{\left[\frac{r_1 \log H(\beta_1)}{\log H(\beta_h)}\right]} \text{ das escolhas dos } r_h$$

$$
\begin{aligned}
&\leq \left(\frac{r_1 \log H(\beta_1)}{\log H(\beta_{h+1})} + 1\right) / \left(\frac{r_1 \log H(\beta_1)}{\log H(\beta_h)}\right) \\
&= \frac{\log H(\beta_h)}{\log H(\beta_{h+1})} + \frac{\log H(\beta_h)}{r_1 \log H(\beta_1)} \\
&\leq \frac{\omega}{2} + \frac{\omega}{2} = \omega \quad \text{das escolhas feitas em (4) e (5).}
\end{aligned}
$$

(7) Como m foi escolhido para verificar (3.2), podemos usar a Proposição 3.2.1 para construir um polinômio $P(X_1, ..., X_m)$ com $\mathrm{gr}_{X_h} P \leq r_h$ satisfazendo (3.3) e (3.4).

(8) Verificamos acima que as quantidades escolhidas satisfazem (3.5), (3.6), (3.7) e (3.8). Portanto podemos aplicar a Proposição 3.3.1 para concluir que

$$\mathrm{Ind}_{(\beta_1,...,\beta_m;r_1,...,r_m)} P \geq m\varepsilon.$$

(9) Gostaríamos agora de aplicar o Lema de Roth, Lema 3.4.2. Já verificamos a condição (3.9) com $\eta^{2^{m-1}} = \omega$ em (6), logo, resta apenas checar a condição (3.10). Usando que

$$\log D = \min_{1 \leq h \leq m} r_h \log H(\beta_h) = r_1 \log H(\beta_1) \quad \text{e} \quad r_1 = \max_h r_h$$

obtemos

$$
\begin{aligned}
\frac{\log |P| + 2mr_1}{\log D} &\leq \frac{(r_1 + \cdots + r_m) \log B + 2mr_1}{\log D} \quad \text{de (3.4)} \\
&\leq \frac{m(\log B + 2)}{\log H(\beta_1)} \quad \text{pois } r_1 > r_2 > \cdots \\
&\leq \omega \quad \text{da escolha de } \beta_1 \text{ em (3).}
\end{aligned}
$$

Isto completa a verificação de todas as condições necessárias para aplicar o Lema de Roth com $\eta = \omega^{2^{-m+1}} = \varepsilon/4$, portanto

$$\mathrm{Ind}_{(\beta_1,...,\beta_m;r_1,...,r_m)} P \leq 2m\eta = m\varepsilon/2.$$

Observamos que as cotas inferior e superior para o índice de P obtidas em (8) e (9), respectivamente, se contradizem! Isto completa a prova de que (3.12) possui apenas uma quantidade finita de soluções. Assim, usando o Lema 2.4.1, concluímos a prova do teorema de Roth.

■

Capítulo 4

Pontos inteiros sobre curvas

No capítulo anterior completamos a prova do teorema de Roth: um número algébrico possui uma quantidade finita de aproximações de ordem $2 + \varepsilon$. Neste capítulo daremos a prova de mais um fundamental teorema de finitude em geometria Diofantina, o teorema de Siegel. Usaremos a notação $f \ll g$ significando que $|f(x)| \leq \kappa |g(x)|$ para alguma constante positiva não especificada κ e aproveitamos para lembrar que o anel de inteiros de K pode ser caracterizado usando valores absolutos como

$$O_K = \{x \in K || x|_v \leq 1, \forall v \in M_K^0\}.$$

Mais geralmente, se $S \subset M_K$ é qualquer subconjunto de valores absolutos contendo os valores absolutos arquimedianos M_K^∞, definimos o **anel de S-inteiros** de K como

$$O_{K,S} = \{x \in K || x|_v \leq 1, \forall v \in M_K, v \notin S\}.$$

Assumiremos que o leitor tenha tido contato com a teoria das curvas algébricas. Denotaremos por $C \subset \mathbf{A}_K^m$ uma curva afim irredutível sobre $\overline{K}$ e chamaremos de **ponto S-inteiro** de C um ponto de C com coordenadas em $O_{K,S}$. Os pontos de $\widetilde{C} \backslash \widetilde{C}_{af}$ serão chamados **pontos de C no infinito**, em que $\widetilde{C}_{af}$ é uma normalização de C e $\widetilde{C}$ é uma extensão de $\widetilde{C}_{af}$ a uma curva projetiva lisa.

4.1 Aproximação Diofantina

Praticamente todos os resultados sobre os conjuntos de pontos inteiros sobre variedades algébricas definidas sobre corpos de números são demonstrados por métodos de aproximação Diofantina. No caso de curvas, o teorema de Roth, permite demonstrar o teorema de Siegel:

Teorema 4.1.1 (SIEGEL) *Suponhamos que C possui uma quantidade infinita de pontos em $\mathbf{A}^m_K(O_K)$. Então $\widetilde{C}$ possui gênero 0 e $\#(\widetilde{C}\backslash C) \leq 2$.*

O argumento clássico é o seguinte: considera-se $\{P_n\}_{n\in\mathbf{N}}$ uma sequência de pontos inteiros sobre $\widetilde{C}$, convergindo a um ponto $Q \in \widetilde{C}\backslash C$ com respeito a algum valor absoluto v de K. Então, para uma função não-constante $\varphi \in K(C)$, pode-se assumir que

$$|\varphi(P_n) - \varphi(Q)|_v \ll H(\varphi(P_n))^{-\delta}$$

para algum $\delta > 0$ dependendo apenas da geometria de $\widetilde{C}$. Como $\varphi(Q)$ deve ser um número algébrico e $\varphi(P_n) \in K$, se tivermos $\delta > 2$ teremos uma contradição, pelo teorema de Roth; mas a princípio, este não é o caso! Contudo, trocando $\widetilde{C}$ por seu *pull-back* pela isogenia multiplicação por m sobre a Jacobiana $J(\widetilde{C})$, concluímos o argumento invocando o teorema de Mordell-Weil fraco ([9]).

Recentemente, em [5], Corvaja e Zannier deram uma nova prova do teorema de Siegel que não usa o teorema de Roth, mas baseia-se no teorema do subespaço de Schmidt que é uma generalização em dimensão superior do teorema de Roth. Esse emprego do teorema do subespaço por Corvaja e Zannier deu a possibilidade de obter resultados sobre o conjunto de pontos inteiros sobre superfícies ([1],[6],[12]). Enunciamos agora a versão que utilizaremos do teorema do subespaço.

Teorema 4.1.2 (***Teorema do subespaço de Schmidt***) *Seja K um corpo de números e seja S um conjunto finito de lugares de K que contém os lugares arquimedianos. Sejam $L_{1,v}, ..., L_{m,v}$ formas lineares linearmente independentes em m variáveis com coeficientes algébricos. Então para qualquer $\varepsilon > 0$ as soluções $\underline{x} \in O^m_{K,S}$ da desigualdade*

$$\prod_{v\in S}\prod_{i=1}^{m} |L_{i,v}(\underline{x})|_v \leq H(\underline{x})^{-\varepsilon}$$

estão contidas em uma quantidade finita de subespaços lineares próprios de K^m.

A prova do teorema do subespaço de Schmidt se assemelha a prova do teorema de Roth, embora apareçam novas dificuldades; para uma demonstração remetemos o leitor a [3].

4.2 Uma nova prova do teorema de Siegel

Provaremos a seguinte versão:

Teorema 4.2.1 (Siegel) *Se C possui pelo menos três pontos no infinito, então C possui apenas uma quantidade finita de pontos S-inteiros.*

A versão usual do teorema de Siegel, Teorema 4.1.1, é mais forte do que essa, pois requer a condição de pelo menos três pontos distintos no infinito somente se $\widetilde{C}$ possui gênero 0, contudo a versão do Teorema 4.1.1 pode ser obtida desta (ver [3] p.184).

Demonstração do Teorema 4.2.1 (de acordo com Corvaja-Zannier). Primeiro, sem perda de generalidade, podemos assumir que C é lisa. Seja $\widetilde{C}$ a curva projetiva lisa associada, $C \subset \widetilde{C}$. Sejam $Q_1, ..., Q_r$, $r \geq 3$, pontos distintos de C no infinito, que podemos assumir definidos sobre K.

Seja N um inteiro positivo, a ser especificado depois. Consideremos o espaço vetorial $V = V_N$ sobre K definido por

$$V = V_N = \{\varphi \in K(\widetilde{C}) | \mathrm{div}(\varphi) \geq -N(Q_1 + \cdots + Q_r)\}.$$

Pelo teorema de Riemann-Roch ([8], p.108) temos que

$$d = d_N := \dim_K V_N \geq Nr + 1 - g,$$

onde g é o gênero de $\widetilde{C}$.

Seja $\{\varphi_1, ..., \varphi_d\}$ uma base para V. As φ_i são funções regulares sobre C, multiplicando cada uma por um denominador adequado, podemos assumir que todas elas possuem coeficientes que estão em $O_{K,S}$. Portanto, se $\{P_n\}_{n \in \mathbf{N}}$ é uma sequência infinita de pontos distintos em $C \cap \mathbf{A}^m(O_{K,S})$, então $\varphi_i(P_n) \in O_{K,S}$ para $i = 1, ..., d$ e todo $n \in \mathbf{N}$.

Como $\widetilde{C}$ é uma curva projetiva, o conjunto $\widetilde{C}(K_v)$ é compacto, com respeito a topologia v-ádica, para todo v. Assim, podemos trocar $\{P_n\}$ por um subsequência adequada, que podemos assumir que esta converge, para todo $v \in S$, v-adicamente a um ponto $P^v \in \widetilde{C}(K_v)$. Agora, vamos repartir o conjunto S da seguinte forma: $S = S_1 \cup S_2$, onde S_1 é o conjunto dos lugares $v \in S$ tais que $P^v \in \widetilde{C} \backslash C$ e $S_2 = S \backslash S_1$.

Nós agora desejamos estimar $|\varphi_i(P_n)|_v$ para $n = 1, 2, ...$ e $v \in S$. Para $v \in S_2$ isto é simples, pois os valores $|\varphi_i(P_n)|_v$ são uniformemente limitados.

Fixemos agora $v \in S_1$ e para $j \geq 1$ consideremos o subespaço vetorial de V definido por

$$W_j = W_{j,v} = \{\varphi \in V | \mathrm{ord}_{P^v} \varphi \geq j - 1 - N\}.$$

Notemos que $V = W_1 \supset W_2 \supset \cdots$ e $\dim(W_j/W_{j+1}) \leq 1$. Em particular,

$$\dim W_j \geq \dim V - j + 1 = d - j + 1.$$

Assim, escolhemos uma base de $W_d \neq \{0\}$ que sucessivamente completamos a uma base de $W_{d-1}, W_{d-2}, ..., W_1$, obtendo vetores $w_j \in W_j$ para $1 \leq j \leq d$. Expressando esses vetores como combinações lineares das φ_i, obtemos formas lineares linearmente independentes $L_{dv}, ..., L_{1v}$ em $\varphi_1, ..., \varphi_d$, definidas sobre K tais que

$$\mathrm{ord}_{P^v} L_{jv} \geq j - 1 - N, \quad j = 1, ..., d.$$

Definimos também tais formas lineares para $v \in S_2$, simplesmente pondo $L_{jv} = \varphi_j$, para $j = 1, ..., d$, aqui, também temos que $|L_{jv}(P_n)|_v \ll 1$.

Voltando para $v \in S_1$, escolhamos um parâmetro local $t_v \in K(\widetilde{C})$ em P^v. Então, para uma função $\psi \in K(\widetilde{C})$ tendo ordem q em P^v, temos que $t_v^{-q}\psi$ é regular em P^v, portanto $|t_v^{-q}(P_n)\psi(P_n)|_v$ é limitado, pois $P_n \to P^v$, ou seja, $|\psi(P_n)|_v \ll |t_v(P_n)|_v^q$, em que a constante implícita independe de n. Como $|t_v(P_n)|_v \leq 1$ para n grande, obtemos que

$$|L_{jv}(P_n)|_v \ll |t_v(P_n)|_v^{j-1-N}, \quad j = 1, ..., d,$$

portanto

$$\prod_{j=1}^{d} |L_{jv}(P_n)|_v \ll |t_v(P_n)|_v^{d(d-1)/2-dN} = |t_v(P_n)|_v^{(d/2)(d-2N-1)}.$$

Portanto

$$\begin{aligned} \prod_{v \in S} \prod_{j=1}^{d} |L_{jv}(P_n)|_v &\ll \left(\prod_{v \in S_1} |t_v(P_n)|_v \right)^{(d/2)(d-2N-1)} \\ &\leq \left(\prod_{v \in S_1} |t_v(P_n)|_v \right)^{d((r-2)N-g)/2} . \quad (4.1) \end{aligned}$$

Por outro lado, $\varphi_j(P_n)$ são pontos S-inteiros, portanto, $\max_j |\varphi_j(P_n)|_v \leq 1$ quando v não está em S, $\max_j |\varphi_j(P_n)|_v \ll 1$ para $v \in S_2$ e $\max_j |\varphi_j(P_n)|_v \ll$

$|t_v(P_n)|_v^{-N}$ para $v \in S_1$, de novo, as constantes implícitas não dependem de n. Portanto,

$$H(\varphi_1(P_n) : ... : \varphi_d(P_n)) \ll \left(\prod_{v \in S_1} |t_v(P_n)|_v \right)^{-N}$$

que comparando com (4.1) obtemos que

$$\prod_{v \in S} \prod_{j=1}^{d} |L_{jv}(P_n)|_v \ll H(\varphi_1(P_n) : ... : \varphi_d(P_n))^{-d((r-2)N-g)/(2N)}.$$

Para $N \geq g+1$, temos que $d \geq 2$, logo as φ_j não são todas proporcionais. Portanto, podemos assumir que $H(\varphi_1(P_n) : ... : \varphi_d(P_n)) \to \infty$ quando $n \to \infty$, pois caso contrário todas as razões $\varphi_j(P_n)/\varphi_1(P_n)$ devem pertencer a um conjunto finito, independentemente de n, pelo teorema de Northcott, o mesmo valendo para os P_n, que é o que desejamos concluir. Podemos então aplicar o teorema do subespaço de Schmidt e concluir que todos os pontos $(\varphi_1(P_n), ..., \varphi_d(P_n)), n \in \mathbf{N}$, estão numa união finita de subespaços lineares próprios de K^d. Como $\varphi_1, ..., \varphi_d$ são linearmente independentes sobre $\widetilde{C}$, isto implica que, de novo, os P_n pertencem a um conjunto finito, independetemente de n.

■

"Ai! Que preguiça!..."
Mário de Andrade (Macunaíma)

Bibliografia

[1] Autissier, P., Géométrie des surfaces algébriques et points entiers, Prépublication de l'IRMAR, 2006.

[2] Bilu, Yu., The many faces of the subspace theorem, Séminaire Bourbaki, 59ème année, 2006-2007, nº 967.

[3] Bombieri, E., Gubler, W., *Heights in Diophantine Geometry*, New Mathematical Monographs, 4, Cambridge, 2007.

[4] Cassels, J.W.S., *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Mathematical Physics, 45, Cambridge University Press, 1957.

[5] Corvaja, P., Zannier, U., Points entiers sur les courbes et théorème des sous-espaces, C. R. Acad. Sci. Paris Ser. I, 334 (2002), 267-271.

[6] Corvaja, P., Zannier, U., On integral points on surfaces, Ann. of Math., (2) 160 (2004), 705-726.

[7] Faltings, G., Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, Invent. Math., 73 (1983), 349-366.

[8] Fulton, W., *Algebraic Curves*, Benjamin, 1969.

[9] Hindry, M., Silverman, J., *Diophantine Geometry*, Graduate Texts in Mathematics, 201, Springer-Verlag, 2000.

[10] Lang, S., *Fundamentals of Diophantine Geometry*, Springer-Verlag, 1983.

[11] Lang, S., *Algebraic Number Theory*, Graduate Texts in Mathematics, 110, Springer-Verlag, 1986.

[12] Levin, A., Generalizations of Siegel's and Picard's Theorems, Ann. of Math., 170 (2009), 609-655.

[13] Neukirch, J., *Algebraic Number Theory*, Grundlehren der mathematischen Wissenschaften, 322, Springer-Verlag, 1999.

[14] Roth, K.F., Rational approximations to algebraic numbers, Mathematika, 2 (1955), 1-20.

[15] Schmidt, W.M., *Diophantine Approximation*, Lectures Notes in Mathematics, 785, Springer-Verlag, 1980.

[16] Serre, J.-P., *Lectures on the Mordell-Weil Theorem*, Friedr. Vieweg & Sohn, 1989.

[17] Siegel, C.L., Über einege Anwendungen Diophantischer Approximationen, Abh. Preuss. Akad. Wiss. Phys. Math. Kl., 1 (1929), 41-69.

[18] Thue, A., Über Annäherungswerte algebraischer Zahlen, J. reine angew. Math., 135 (1909), 284-305.

[19] Zannier, U., *Lectures Notes on Diophantine Analysis*, Edizioni della Normale, 8, 2009.

[20] Zannier, U., Roth's theorem, integral points and certain ramified covers of $\mathbb{P}^1$, Analytic Number Theory: Essays in Honour of Klaus Roth, Cambridge, 2009.