UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

El Teorema de Roth

Juan Manuel Menconi

Director: Román Sasyk

Marzo 2015

Índice

Introducción

Una de las preguntas básicas de la Teoría de Aproximaciones Diofánticas es investigar las aproximaciones racionales a un número real. Uno de los principales objetivos de esta teoría es comparar, por un lado, la distancia entre un número real α y un número racional p/q, y por el otro, el denominador q de la aproximación. Mientras más chico sea $|\alpha - p/q|$ en comparación con q, se dirá que la aproximación racional será mejor. En 1844, Liouville fue el primero en observar que en el caso de números algebraicos, no se podrán conseguir aproximaciones racionales tan buenas como uno desee; si α es un número algebraico de grado $d \geq 2$, existirá una constante positiva C, que solo depende de α, tal que

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C}{q^d}$$

para todo número racional p/q, con $q > 0$[1].

En 1909, una mejora notable en el Teorema de Liouville fue obtenida por el matemático noruego Axel Thue, durante su investigación sobre la finitud del conjunto de soluciones de ciertas ecuaciones Diofánticas. Thue probó que si α es un número algebraico de grado $d \geq 2$, para todo $k > d/2 + 1$ existe una constante $C(\alpha, k) > 0$, tal que

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C(\alpha,k)}{q^k}$$

para todo número racional p/q, con $q > 0$. Como consecuencia de este resultado, Thue probó que la Ecuación de Thue

$$F(X, Y) = m,$$

con $F \in \mathbb{Z}[X, Y]$ un polinomio homogéneo de grado d con al menos 3 factores lineales sobre $\mathbb{C}$ no proporcionales, posee finitas soluciones enteras x, y, para toda constantes fija no nula m.

Luego del trabajo de matemáticos como Siegel, Dyson, Gelfond, Schneider y Mahler, en 1955 Roth probó que la proposición anterior sigue siendo valida para $k > 2$. Una forma equivalente y quizás mas clásica de enunciar este teorema es la siguiente; si α es un número algebraico de grado $d \geq 2$, para todo $\varepsilon > 0$ la inecuación

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^{d+\varepsilon}} \tag{0.1}$$

posee finitas soluciones racionales p/q, con $q > 0$. A partir de un resultado clásico de Dirichlet[2], el resultado de Roth es esencialmente el mejor posible.

[1] Gracias a este resultado, Liouville fue capaz de probar la existencia de números trascendentes.

[2] Si se toma $\varepsilon = 0$, existen infinitas soluciones racionales

Desde los trabajos de Siegel, este resultado ha pasado por varias mejoras sucesivas. Desde la extensión a aproximaciones en un cuerpo de números K, a la consideración de valores absolutos p-ádicos. En este trabajo estudiaremos el teorema de Roth bajo estas generalizaciones y algunas de sus respectivas aplicaciones, entre otras cosas, a la teoría de ecuaciones diofánticas.

La exposición aquí presentada seguirá la dada en el libro de Hindry y Silverman [**15**], con ligeras modificaciones, que nos permitirán llegar a un resultado más general, como el presentado en los libros de Lang [**18**] y Bombieri y Gubler [**4**].

CAPÍTULO 1

El camino hacia el teorema de Roth

1. Los teoremas de Dirichlet y Liouville

Uno de los problemas fundamentales en la teoría de Aproximaciones Diofánticas consiste en comprender que tan bien se puede aproximar a un número real por medio de números racionales o más generalmente números algebraicos. La completitud de $\mathbb{Q}$ en $\mathbb{R}$ nos asegura que si $\alpha \in \mathbb{R}$, la diferencia $|\alpha - \frac{p}{q}|$ puede hacerse tan pequeña como queramos para algún $\frac{p}{q}$ adecuado. Si bien esta respuesta es válida, no resulta completamente satisfactoria. Reformulemos nuestra pregunta a tratar de entender con qué precisión podemos aproximar a α por medio de racionales, es decir, lograr que esta diferencia sea chica sin que p y q sean muy grandes. Por ejemplo, dado $\alpha \in \mathbb{R}$, podemos intentar responder para que valores $\varepsilon > 0$ la inecuación

$$\left|\alpha - \tfrac{p}{q}\right| \leq \tfrac{1}{q^{\varepsilon}}$$

posee infinitas soluciones racionales $\frac{p}{q} \in \mathbb{Q}$ con $q > 0$. Mientras mayor sea ε, más precisa será la aproximación. Comprender esta pregunta para un número real nos permitirá establecer si dicho número es racional o irracional, o si es algebraico o trascendente. Técnicas de Aproximaciones Diofanticas han sido aplicadas para resolver problemas de Inecuaciones Diofánticas, Ecuaciones Diofánticas, Geometria Diofántica y Teoría de Trascendencia, algunos de los cuales mencionaremos en este trabajo. Nuestro objetivo principal será probar el teorema de Roth, cuya versión original estableces que si α es un número algebraico, dado $\varepsilon > 0$ la desigualdad

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

posee finitas soluciones racionales.

Como punto de partida hacia la prueba del teorema de Roth y el estudio de las aproximaciones diofánticas, probaremos el siguiente teorema, debido a Dirichlet.

TEOREMA 1.1. *Sean α un número irracional y Q un entero > 1. Entonces existen enteros p y q, con $1 \leq q \leq Q$ tal que*

$$\left|\alpha - \tfrac{p}{q}\right| \leq \tfrac{1}{qQ}.$$

DEMOSTRACIÓN. Consideremos los $Q + 1$ números reales

$$q\alpha - [q\alpha] \quad \text{con} \quad q = 0, 1, \ldots, Q.$$

Como α es irracional, estos son $Q + 1$ números distintos en el intervalo $[0, 1]$. Dividiendo el intervalo en Q subintervalos de longitud $1/Q$, el principio del palomar asegura que podremos encontrar dos enteros $0 \leq q_1 < q_2 \leq Q$ tal que

$$|(q_1\alpha - [q_1\alpha]) - (q_2\alpha - [q_2\alpha])| \leq \tfrac{1}{Q}.$$

Por lo tanto

$$\left|\frac{[q_2\alpha]-[q_1\alpha]}{q_2-q_1}-\alpha\right| \leq \frac{1}{(q_2-q_1)Q}.$$

Tomamos entonces, $p=[q_2\alpha]-[q_1\alpha]$ y $q=q_2-q_1\leq Q$. □

COROLARIO 1. [1]*Sea $\alpha\in\mathbb{R}$ irracional. Entonces la inecuación*

$$\left|\alpha-\frac{p}{q}\right| \leq \frac{1}{q^2} \tag{1.1}$$

posee infinitas soluciones $\frac{p}{q}$ con $q>0$.

DEMOSTRACIÓN. Tomemos por ejemplo $Q=2$ en el teorema anterior, entonces existen enteros p y q, con $1\leq q\leq Q$ tal que $|\alpha-p/q|\leq 1/qQ$, en particular $|\alpha-p/q|\leq 1/q^2$. Como α es irracional, $|\alpha-p/q|\neq 0$ con lo cual existe $Q'>1$ tal que $1/Q'<|\alpha-p/q|$. Volviendo a aplicar el teorema para Q' obtenemos un nuevo racional p'/q' que aproxima de la misma forma y es distinto de p/q pues

$$|\alpha-p'/q'|\leq 1/(q'Q')\leq 1/Q'<|\alpha-p/q|.$$

Repitiendo esto sucesivamente, obtenemos el resultado deseado. □

OBSERVACIÓN 1. *Este corolario fue mejorado por Hurwitz.*

TEOREMA 1.2 (Hurwitz 1891). *Sea $\alpha\in\mathbb{R}$ irracional entonces existen infinitos números racionales $\frac{p}{q}$ con $q>0$ que cumplen*

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

Hurwitz además probó que para $\alpha=\frac{1}{2}(\sqrt{5}-1)$, esta constante es optima, es decir, no podemos tomar una constante mayor a $\sqrt{5}$ en el denominador[2]*.*

Observar, que el Corolario 1 resulta falso si α es racional. Pues supongamos que $\alpha=a/b$, si $p/q\neq\alpha$, $aq-pb$ será un número entero no nulo entonces

$$\left|\alpha-\frac{p}{q}\right|=\left|\frac{a}{b}-\frac{p}{q}\right|=\left|\frac{aq-pb}{bq}\right|\geq\frac{1}{bq}. \tag{1.2}$$

Si p/q cumple además (1.1), entonces se tendrá que $q\leq b$. Por lo tanto hay finitos racionales que la cumplen.

Obtenemos así, una forma de distinguir los números racionales de los irracionales en función de cómo pueden ser aproximados. Resumimos este comentario en el siguiente teorema.

TEOREMA 1.3. *Sea $\alpha\in\mathbb{R}$, entonces α es irracional si y solo si*

$$\left|\alpha-\tfrac{p}{q}\right|\leq\tfrac{1}{q^2}$$

posee infinitas soluciones racionales.

La observación hecha en (1.2) puede verse de la siguiente forma. Sea α un número algebraico de grado 1, es decir, $\alpha=a/b$ un número racional. Entonces existe una constante $c(\alpha)=1/b$ tal que

[1]Este resultado ya había sido obtenido por medio de la teoría de Series de Farey y también por medio de la teoría de fracciones continuas, principalmente desarrollada por Euler y Lagrange. La nueva demostración dada por Dirichlet permitió generalizarlo a aproximaciones simultaneas. Ver por ejemplo [**29**] Capitulo II

[2]Una prueba de este resultado aparece en [**29**] Capitulo I.

$$\left|\alpha - \tfrac{p}{q}\right| \geq \tfrac{1}{bq} = \tfrac{c(\alpha)}{q} \quad \text{para todo} \quad \tfrac{p}{q} \neq \alpha.$$

Si consideramos ahora el caso en el que α es un número cuadrático, dado que estos poseen desarrollo en fracción continua periódico, se sabe que existe una constante $c = c(\alpha) > 0$ tal que

$$\left|\alpha - \tfrac{p}{q}\right| \geq \tfrac{c}{q^2}$$

para todo racional $\frac{p}{q}$. Liouville observó que una desigualdad de este estilo era válida para números algebraicos de grado d y por lo tanto existe un límite para la rapidez con la cual podemos aproximar números algebraicos por medio de racionales.

TEOREMA 1.4 (Liouville 1844-51). *Sea $\alpha \in \mathbb{R}$ algebraico de grado $d \geq 2$, entonces existe $c = c(\alpha) > 0$ tal que*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^d} \tag{1.3}$$

para todo racional $\frac{p}{q}$ con $q > 0$.

DEMOSTRACIÓN. Sea $f(x) \in \mathbb{Z}[x], gr(f) = d$ tal que $f(\alpha) = 0$, es decir, f es un múltiplo entero del polinomio minimal de α. Por lo tanto, dado $\frac{p}{q} \in \mathbb{Q}$, como f es irreducible sobre $\mathbb{Q}$ se tiene que f no se anula en $\frac{p}{q}$. Entonces si $f(x) = \sum_{i=0}^{d} a_i x^i$,

$$q^d f(\tfrac{p}{q}) = \textstyle\sum_{i=0}^{d} a_i p^i q^{d-i}$$

es un número entero no nulo y por tanto ≥ 1 lo que implica que

$$|f(\tfrac{p}{q})| \geq \tfrac{1}{q^d}$$

Si $|\alpha - \frac{p}{q}| > 1$, (1.3) se cumple trivialmente tomando $c = 1$. Supongamos entonces que $|\alpha - \frac{p}{q}| \leq 1$. Por medio del desarrollo de Taylor de $f(x)$ alrededor de α se tendrá que

$$\left|f(\frac{p}{q})\right| = \left|\sum_{i=1}^{d} b_i \left(\alpha - \frac{p}{q}\right)^i\right| = \left|\alpha - \frac{p}{q}\right| \left|\sum_{i=1}^{d} b_i \left(\alpha - \frac{p}{q}\right)^{i-1}\right|$$

$$< \left|\alpha - \frac{p}{q}\right| \left(\sum_{i=1}^{d} |b_i|\right) = \left|\alpha - \frac{p}{q}\right| c'$$

Por lo tanto,

$$\tfrac{1}{q^d} < \left|\alpha - \tfrac{p}{q}\right| c'.$$

Tomando $c = min(1/c', 1)$ obtenemos (1.3) para todo p/q.

□

OBSERVACIÓN 2. *El teorema sigue siendo válido si α es un número complejo no real, no necesariamente algebraico, ya que*

$$\left|\alpha - \tfrac{p}{q}\right| \geq |Im(\alpha)| \geq \tfrac{|Im(\alpha)|}{q^d}$$

para cualquier número racional $\frac{p}{q}$; donde $Im(\alpha)$ denota la parte imaginaria de α y d es cualquier entero positivo.

Liouville obtuvo así una condición necesaria que deben cumplir los números algebraicos de grado d. Gracias a esto, logro probar la existencia de los números trascendentes y dar infinitos[3] ejemplos de ellos. Lo que necesitamos es encontrar un número que no cumpla (1.3) para ninguna constante $c > 0$ y ningún entero $d > 0$. Dado que estamos pidiendo que esto no se cumpla para ningún $d > 0$ alcanza con tomar $c = 1$. En definitiva, tenemos el siguiente resultado.

TEOREMA 1.5. *Sea $\alpha \in \mathbb{R}$ y supongamos que para todo entero $d > 1$ existe un número racional $\frac{p}{q}$ tal que*

$$\left|\alpha - \tfrac{p}{q}\right| \leq \tfrac{1}{q^d}.$$

Entonces α resulta trascendente.

En función de lo que queremos, no es difícil construir un número que cumpla esta condición. Consideremos el número $\xi = \sum_{n\geq 1} \frac{1}{10^{n!}}$, conocido como constante de Liouville. Dicho número es trascendente pues sea $d > 1$ y tomamos $p = 10^{d!} \sum_{n\geq 1}^{d} \frac{1}{10^{n!}}$ y $q = 10^{d!}$ entonces

$$\left|\xi - \frac{p}{q}\right| = \sum_{n>d}^{\infty} \frac{1}{10^{n!}} < \frac{1}{10^{(d+1)!}} \left(\frac{1}{1} + \frac{1}{10} + \frac{1}{10^2} \cdots\right) = \frac{9}{10} \frac{1}{10^{(d+1)!}} < \frac{1}{10^{(d+1)!}} < \frac{1}{q^d}.$$

Más en general,

PROPOSICIÓN 1. $\sum_{k\geq 1} \frac{a_k}{b^{k!}}$ *es un número trascendente para todo $b \in \mathbb{Z}$, $b \geq 2$ y $1 \leq a_k \leq b - 1$.*

A partir del teorema de Dirichlet vimos que es posible decidir si un número es racional o irracional si sabemos para que valores $\varepsilon > 0$ se cumple que la inecuación

$$\left|\alpha - \tfrac{p}{q}\right| < \tfrac{1}{q^\varepsilon}$$

posee infinitas soluciones racionales $\frac{p}{q} \in \mathbb{Q}$ con $q > 0$. Resulta entonces natural intentar caracterizar a los números irracionales en función del conocimiento de estos exponente. Se define el exponente de aproximación de un número α, también llamado medida de irracionalidad, como el menor valor $\mu(\alpha)$ para el cual, dado $\varepsilon > 0$, la inecuación

$$\left|\alpha - \tfrac{p}{q}\right| < \tfrac{1}{q^{\mu(\alpha)+\varepsilon}}$$

posee finitas soluciones racionales $\frac{p}{q} \in \mathbb{Q}$ con $q > 0$. Si no existe ningún μ con esta propiedad, diremos que $\mu(\alpha) = \infty$[4]. Como consecuencia del teorema de Dirichlet, tenemos que $\mu(\alpha) \geq 2$ si α es irracional. A partir de (1.2) se tiene que $\mu(a/b) \leq 1$. De hecho se puede ver que $\mu(\frac{a}{b}) = 1$ pues dado $q > 0$ entero, la fracción de denominador q más cercana a a/b cumple que $|a/b - p/q| \leq 1/2q < 1/q$.

El teorema de Liouville establece que si α es un número algebraico de grado d, entonces $\mu(\alpha) \leq d$. Pues si $\varepsilon > 0$ y p/q es solución a $|\alpha - p/q| < 1/q^{d+\varepsilon}$ se tiene que

[3]De hecho exhibió una cantidad no numerables de números trascendentes. La teoría de Cantor sobre cardinalidad y su demostración de la existencia y no numerabilidad de los números trascendentes es 30 años posterior al teorema de Liouville.

[4]Estos números suelen llamarse, números de Liouville. Los números de Proposición 1, y en particular la constantes de Liouville, pertenecen a este conjunto de números.

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

Lo que implica que q debe estar acotado, probando asi que $\mu(\alpha) \leq d$.

La primer mejora en el caso de números algebraicos fue obtenida por Thue [**32**] en 1909, quien probó que $\mu(\alpha) \leq \frac{1}{2}d+1$. El trabajo de Thue es de suma importancia, no solo por sus aplicaciones, sino también porque su procedimiento para establecer este resulto fue la base para los trabajos subsiguientes en este campo. En 1921 Siegel [**25**] probó que

$$\mu(\alpha) \leq \min_{s\in\mathbb{N}} \left(s + \frac{d}{s+1}\right) < 2\sqrt{d},$$ [5]

y conjeturó que para cualquier número algebraico, independientemente del grado, debería valer $\mu(\alpha) = 2$. En 1947 Dyson [**11**] e independientemente Gelfond [**14**] en 1952 probaron que $\mu(\alpha) \leq \sqrt{2d}$. La conjetura de Siegel fue finalmente probada en 1955 por Roth [**23**], trabajo por el cual le fue otorgada la medalla Fields en 1958.

Teorema 1.6 (Roth 1955). *Sea α un número real algebraico de grado $d \geq 2$. Entonces para todo $\varepsilon > 0$, la desigualdad*

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

posee finitas soluciones racionales $\frac{p}{q}$ con $q > 0$.

El teorema de Roth puede formularse de cualquiera de las siguientes formas equivalentes

Proposición 2. *Sea α un número algebraico real de grado $d \geq 2$ entonces son equivalentes:*

(1) *Para todo $\varepsilon > 0$ existe una constante $c(\alpha, \varepsilon) > 0$ tal que*
$$|\alpha - \tfrac{p}{q}| > \tfrac{c(\alpha,\varepsilon)}{q^{2+\varepsilon}}$$
para todo $\frac{p}{q} \in \mathbb{Q}$, con $q > 0$;

(2) *Para todo $\varepsilon > 0$, la desigualdad*
$$|\alpha - \tfrac{p}{q}| < \tfrac{1}{q^{2+\varepsilon}}$$
posee finitas soluciones $\frac{p}{q} \in \mathbb{Q}$ y $q > 0$

(3) *Para todo $\varepsilon > 0$, $C > 0$, la desigualdad*
$$|\alpha - \tfrac{p}{q}| < \tfrac{C}{q^{2+\varepsilon}}$$
posee finitas soluciones $\frac{p}{q} \in \mathbb{Q}$ y $q > 0$.

2. El trabajo de Thue

Si bien intentar caracterizar los números reales por medio de su exponente de aproximación puede ser considerado una forma de justificar los avances antes mencionados o quizas el hecho de que obtener mejoras en el exponente para la cota de Liouville produciría nuevos ejemplos de números trascendentes, se podría decir que las motivaciones que dieron origen a la búsqueda de estas mejoras fueron el estudio de las ecuaciones diofánticas. Thue observó que a una solución entera de una cierta ecuación diofántica se le podía asociar una muy buena aproximación hacia un número algebraico definido por dicha ecuación.

[5] El hecho de que este exponente sea de orden $o(d)$ en lugar de d fue de suma importancia en la prueba de Siegel, del teorema que establece que toda curva de genero $g \geq 1$ posee finitas soluciones enteras(al menos para el caso $g = 1$).

Veamos un ejemplo, supongamos que queremos resolver la siguiente ecuación sobre los enteros

$$X^3 - Y^3 = 5. \tag{1.4}$$

Una posible forma de hacerlo es factorizando el polinomio considerado de la siguiente forma

$$X^3 - Y^3 = (X - Y)(X^2 + XY + Y^2). \tag{1.5}$$

Si (x, y) es una solución entera de (1.4), entonces cada uno de los factores de (1.5) resulta ser un número entero. Dado que en los enteros hay factorización única $x-y$ será igual a alguno de los factores de 5. Por ejemplo, si suponemos que $x - y = 5$, resulta que $x - 5 = y$ con lo cual $x^2 + x(x - 5) + (x - 5)^2 = 1$ obtenemos así una ecuación de grado 2 a la cual le podemos calcular sus raíces y conseguir así las soluciones a nuestra ecuación en el caso $x - y = 5$. Haciendo lo mismo para todos los factores de 5 obtendremos todas las soluciones de la ecuación en cuestión.

Que sucede ahora si modificamos ligeramente nuestra ecuación por

$$X^3 - 2Y^3 = 5. \tag{1.6}$$

Si intentamos proceder de la misma forma que en el caso anterior encontramos que no podremos factorizar al polinomio $X^3 - 2Y^3$ sobre $\mathbb{Z}[X, Y]$. Dado que en el caso anterior estas técnicas funcionaron a la perfección, intentemos hacer lo mismo permitiendo la aparición de coeficientes no necesariamente enteros. Sea ζ una raíz cubica de la unidad entonces

$$X^3 - 2Y^3 = \left(X - \sqrt[3]{2}Y\right)\left(X - \zeta\sqrt[3]{2}Y\right)\left(X - \zeta^2\sqrt[3]{2}Y\right).$$

Sea ahora $(x, y) \in \mathbb{Z}^2$ una solución de (1.6), con $y \neq 0$, entonces

$$\left(\tfrac{x}{y} - \sqrt[3]{2}\right)\left(\tfrac{x}{y} - \zeta\sqrt[3]{2}\right)\left(\tfrac{x}{y} - \zeta^2\sqrt[3]{2}\right) = \tfrac{5}{y^3}.$$

El segundo y tercer factor de este producto están acotados inferiormente pues $\zeta\sqrt[3]{2}$ y $\zeta^2\sqrt[3]{2}$ son números complejos no reales, con lo cual estarán lejos de cualquier número racional x/y. Esto nos dice que podremos encontrar una constante $C > 0$ independiente de (x, y) tal que

$$\left|\frac{x}{y} - \sqrt[3]{2}\right| \leq \frac{C}{|y|^3}. \tag{1.7}$$

Por lo tanto, toda solución (x, y) a nuestra ecuación (1.6), da lugar a una muy buena aproximación racional de $\sqrt[3]{2}$. Luego, si probamos que $\sqrt[3]{2}$ no posee infinitas aproximaciones de este estilo, es decir, $\mu(\sqrt[3]{2}) < 3$, se tendrá que nuestra ecuación posee finitas soluciones enteras.

Por medio de la desigualdad del teorema de Liouville, se obtuvo que $\mu(\sqrt[3]{2}) \leq 3$. Lo cual no nos ayuda en nada en este caso. Como se dijo, Thue logro probar que $\mu(\sqrt[3]{2}) \leq 3/2 + 1 = 2,5 < 3$, lo que implica en este ejemplo que (1.7) posee finitas soluciones racionales y por lo tanto nuestra ecuación (1.6) posee finitas soluciones enteras.

Antes de seguir con la mejora de Thue del teorema de Liouville, veamos la importancia de este tipo de resultado y cómo se utiliza para probar la finitud de soluciones de ciertas ecuaciones diofánticas.

TEOREMA 1.7. *Sea $F(X, Y) \in \mathbb{Z}[X, Y]$ un polinomio homogéneo de grado d con al menos tres factores lineales no proporcionales sobre $\mathbb{C}$. Sea $m \in \mathbb{Z}$ un entero*

fijo no nulo, entonces la ecuación

$$F(X,Y) = m \tag{1.8}$$

posee finitas soluciones con $x, y \in \mathbb{Z}$.

Ecuaciones de este estilo, son llamadas ecuaciones de Thue.

Demostración. Asumamos primero que F es irreducible sobre $Z[X,Y]$ o equivalentemente, $f(X) = F(X,1)$ es irreducible sobre $\mathbb{Q}[X]$. Sea

$$f(X) = a.(X-\alpha_1).(X-\alpha_2)\dots(X-\alpha_d),$$

con $a \in \mathbb{Z}$, $\alpha_i \in \mathbb{C}$. Como $F(X,Y) = Y^d.f(\frac{X}{Y})$ se obtiene la siguiente descomposición

$$a\left(\frac{X}{Y}-\alpha_1\right).\left(\frac{X}{Y}-\alpha_2\right)\dots\left(\frac{X}{Y}-\alpha_d\right) = \frac{m}{Y^d}. \tag{1.9}$$

Supongamos que tenemos infinitas soluciones (x_n, y_n) de (1.8), podemos además suponer $y_n \neq 0$ pues solo hay finitas soluciones bajo esas condiciones. Se tendrá entonces que $|y_n| \to \infty$ y por lo tanto, tomando a partir de (1.9) obtendremos que, $\frac{x_n}{y_n} \to \alpha_j$ para algun cero de $f(X)$. Los factores

$$\left|\frac{x_n}{y_n} - \alpha_i\right|$$

con $i \neq j$, resultaran acotados inferiormente para todo $n \in \mathbb{N}$. Luego tendremos que

$$\left|\frac{x_n}{y_n} - \alpha_i\right| \leq \frac{C}{|y|^d}$$

para todo $n \in \mathbb{N}$, para alguna constante $C > 0$. Por el teorema de Thue, sabemos que esto es imposible pues $d \geq 3$, llegando así a una contradicción.

En general, sea $F(X,Y) = a.F_1(X,Y)^{e_1}\dots F_r(X,Y)^{e_r}$, con $F_i(X,Y)$ los factores no constantes irreducibles en $\in \mathbb{Z}[X,Y]$ de F, que resultan homogéneos, y $e_1, \dots, e_r$ enteros positivos. Podemos suponer que Y no divide a F pues en ese caso, se tendría que si (x,y) es solución de la ecuación de Thue, y deberá ser uno de los finitos divisores de m y como para cada uno de estos hay finitas opciones para x, se tendrá que (1.8) en este caso posee finitas soluciones.

Toda solución entera (x,y) de (1.8) resulta solución de un sistema de la forma

$$F_j(X,Y) = m_j \text{ para } \ j = 1, \dots, r,$$

con m_j divisores de m. Por lo tanto, si asumimos que (1.8) posee infinitas soluciones, dado que hay finitos de estos sistemas, por el principio del palomar alguno de ellos tendrá infinitas soluciones. Sean (x_n, y_n) infinitas soluciones enteras de uno de estos sistemas, que podemos suponer $y_n \neq 0$. Por el mismo argumento mencionado en el caso anterior tenemos que $\frac{x_n}{y_n}$ tiende a algún cero de $F_j(X_j, 1)$ para cada j, y como estos son coprimos entre sí, se tendrá que $r = 1$. Como F posee al menos tres factores lineales distintos, $gr(F_1) \geq 3$. Por el caso anterior, llegamos a un absurdo pues el sistema $F_1(X,Y) = m_1$ posee finitas soluciones enteras. □

Los siguientes ejemplos sencillos muestran que las hipótesis impuestas sobre f no pueden ser excluidas del enunciado. La ecuación $(x+y)^n = 1$ posee infinitas soluciones para todo $n \in \mathbb{N}$; mientras que la ecuación $(x^2 - 2y^2)^n = 1$ se cumple para los pares $(x_m, y_m) \in \mathbb{Z}^2$ definidos por $x_m + \sqrt{2}y_m = (3 + 2\sqrt{2})^m$ con $m \in \mathbb{Z}$.

Recordemos la demostración de $\mu(\alpha) \leq d$, a partir del teorema de Liouville, para α un número algebraico de grado d. Separaremos la demostración en 3 pasos, con el fin de mostrar que dicha estructura se mantiene en las demostraciones de los teoremas de Thue y Roth, como ha sido notado por varios autores.

Supongamos que α es un número algebraico de grado $d > 1$ y sea $k > d$ tal que

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{|q|^k} \tag{1.10}$$

posee infinitas soluciones racionales.

Paso I- Construcción del polinomio auxiliar.

Construimos un polinomio no nulo $f(X)$ con coeficientes enteros que tenga a α como raíz. En el caso del Teorema de Liouville tomamos $f(X)$ como un múltiplo entero del polinomio minimal de α.

Paso II- f se anula en puntos racionales muy cercanos a α.

Sea $\frac{p}{q}$ un número racional que cumple (1.10). Como f posee coeficiente enteros se cumple que

$$\left|f(\tfrac{p}{q})\right| = \tfrac{N}{|q|^d}$$

con N un número entero positivo. Por otro lado, via la expresión de Taylor de f alrededor de α se obtiene una constante $c > 0$ tal que

$$\left|f(\tfrac{p}{q})\right| < c\left|\alpha - \tfrac{p}{q}\right| < c\tfrac{1}{|q|^k}.$$

Tomando q suficientemente grande, pues suponemos que hay infinitos, y usando que $k > d$ se tendrá que N es un número entero no negativo que cumple

$$N < c\tfrac{1}{|q|^{k-d}} < 1.$$

Por lo tanto, si q es suficientemente grande debe ser $f(\frac{p}{q}) = 0$.

Paso III- El polinomio f, no se anula en un punto racional.

Para el polinomio f construido, podremos elegir $\frac{p}{q}$ de forma tal que q sea grande, para que siga valiendo II, y $\frac{p}{q}$ no sea raíz de f. Dado que f un múltiplo del minimal de α, resulta irreducible sobre $\mathbb{Q}[X]$, por lo tanto, este paso resulta sumamente sencillo en este caso.

A partir del paso II y III obtenemos una contradicción, probando así lo que queríamos.

Un primer intento de mejorar esto sería tomar un polinomio distinto, por ejemplo, que tenga a α como raíz de multiplicidad más grande que 1. Supongamos que tomamos $g(X)$ un polinomio con coeficientes enteros tal que α es raíz de multiplicidad l. Como g posee finitas raíces, el paso III seguirá valiendo si tomamos q suficientemente grandes. En el paso II, por medio de la expresión de Taylor alrededor de p/q obtendríamos una cota del estilo

$$\tfrac{N}{|q|^{gr(g)}} = \left|f(\tfrac{p}{q})\right| < c\left|\alpha - \tfrac{p}{q}\right|^l < c\tfrac{1}{|q|^{kl}},$$

para los p/q que cumplen (1.10), con c una constante positiva que depende de g. Por lo tanto,

$$N < \tfrac{c}{|q|^{kl-gr(g)}}.$$

Podremos asegurar que $N = 0$, si tomamos q suficientemente grande, para los k que cumplan $kl - gr(g) > 0$. Queremos entonces que el cociente $\frac{gr(g)}{l}$ sea lo más chico posible, en particular buscamos que sea menor a d, para que esto sea realmente una mejora. Pero resulta que como α es raíz de multiplicidad l de g y este es un polinomio con coeficientes enteros, entonces f^l divide a g, con f el polinomio minimal de α. Por lo tanto debe ser $gr(g) \geq gr(f^l) = ld$, es decir, $\frac{gr(g)}{l} \geq d$. Vemos de esta forma que el mejor exponente que podemos conseguir vía este método es el obtenido por Liouville.

Thue tuvo la idea de considerar polinomios en dos variables $f(X, Y) \in \mathbb{Z}[X, Y]$ para mejorar el resultado de Liouville, en particular, utilizó polinomios de la forma $f(X, Y) = p(X) + Yq(X)$. A continuación daremos una breve descripción de los pasos para demostrar el teorema de Thue y motivar algunas ideas que usaremos a lo largo de la prueba del teorema de Roth. Sea α algebraico de grado $d \geq 3$(pues el caso 2 ya está comprendido) y supongamos que $|\alpha - p/q| < q^{-k}$ posee infinitas soluciones.

Paso I- Construcción del polinomio.

Thue propuso construir un polinomio $f(X, Y)$ de la forma antes mencionada tal que $h(X) = f(X, \alpha)$ posea una raiz de multiplicidad grande. La construcción es de la siguiente manera. Si el polinomio $f(X, Y)$ posee grado total r[6], habrá que determinar como mucho $2r + 1$ coeficientes. Que $h(X)$ posea un cero de orden s en α es equivalente a que $f(\alpha, \alpha) = \frac{\partial f}{\partial X}(\alpha, \alpha) = \ldots = \frac{\partial^{s-1} f}{\partial X^{s-1}}(\alpha, \alpha) = 0$. Cada una de estas ecuaciones, terminará siendo una ecuación lineal con los coeficientes de f como incóngnitas y con cuerpo de base $\mathbb{Q}[\alpha]$. Escribiendo cada potencia α^n con $n \geq d$ como combinación $\mathbb{Q}$ lineal de las potencias $1, \alpha, \ldots, \alpha^{d-1}$, nuestras ecuaciones se transformaran en una combinación lineal de $1, \alpha, \ldots, \alpha^{d-1}$ cuyos coeficientes son ecuaciones lineales en nuestras incógnitas con coeficientes sobre $\mathbb{Q}$. Por lo tanto, como $1, \alpha, \ldots, \alpha^{d-1}$ son $\mathbb{Q}$ linealmente independientes, cada una de estas s ecuaciones será equivalente a d ecuaciones lineales sobre $\mathbb{Q}$ con incógnitas, los coeficientes de f. En total tendremos sd ecuaciones lineales y $2r + 1$ incógnitas. Podremos obtener una solución no nula si logramos que $sd < 2r + 1$, por ejemplo si tomamos $r/s \approx \frac{1}{2}d$. Esto sugiere la mejora obtenida por Thue.

Paso II- El polinomio se anula en valores racionales cercamos a (α, α).

Al igual que en el Teorema de Liouville, tomemos $\frac{p}{q}$ una muy buena aproximación de α y mediante la fórmula de Taylor tratemos de llegar a que $f(\frac{p}{q}, \frac{p}{q}) = 0$. Se tiene que

$$f(X, Y) = \sum_{i=s}^{r} b_i (X - \alpha)^i + \sum_{j=0}^{r} c_J (X - \alpha)^j (Y - \alpha)$$

Nos gustaría llegar a algo del estilo

$$|f(\tfrac{p}{q}, \tfrac{p}{q})| < B \left|\alpha - \tfrac{p}{q}\right|^s.$$

El problema de esto es que el aporte de los monomios $(X - \alpha)^j (Y - \alpha)$ al evaluarlos en $(\frac{p}{q}, \frac{p}{q})$, no es semejante al aportado por $(X - \alpha)^s$. La forma de Thue para solucionar esto fue tomar dos aproximaciones racionales distintas $(\frac{p_1}{q_1}, \frac{p_2}{q_2})$ de forma tal que $q_2 \approx q_1^s$. Por lo tanto $\left|\frac{p_2}{q_2} - \alpha\right| < \frac{1}{q_2^k} \approx \frac{1}{q_1^{ks}}$ que es similar a lo que aporta $\left|\frac{p_1}{q_1} - \alpha\right|^s$. De esta forma lograremos obtener una desigualdad del estilo

[6]La mayor de las sumas de las potencias de todos los monomios no nulos.

$$\left|f(\tfrac{p_1}{q_1},\tfrac{p_2}{q_2})\right| \leq \tfrac{B}{q_1^{sk}},$$

con B una constante que depende de las coeficientes de f. Mientras que por otro lado tendremos que

$$\left|f(\tfrac{p_1}{q_1},\tfrac{p_2}{q_2})\right| = \tfrac{N}{q_1^r q_2} \approx \tfrac{N}{q_1^{r+s}}.$$

Al igual que en el teorema de Liouville llegaremos a que $N = 0$ y por lo tanto $f(\frac{p_1}{q_1},\frac{p_2}{q_2}) = 0$, tomando q_1 y q_2 suficientemente grande si $sk-(r+s) > 0$ o lo que es lo mismo $r/s + 1 < k$. Como dijimos que tomaríamos $r/s \approx \frac{1}{2}d$ se obtiene el resultado de Thue, $\frac{1}{2}d+1 < k$. Esto será posible si logramos también tener control sobre la constante B. Esta constante depende del valor las derivadas de f en (α,α) que dependerá del valor de los coeficientes de f. El polinomio f, se obtuvo a partir de la existencia de soluciones de un sistema de ecuaciones lineales. Para poder tener control sobre el tamaño de la solución, Thue usó un argumento que se basa en el principio del palomar que permite controlar el tamaño de la solución en función del tamaño de las ecuaciones. Para tener más control sobre los coeficientes del sistema de ecuaciones, por ejemplo, se pueden reemplazar las ecuaciones $\frac{\partial^i f}{\partial X^i}(\alpha,\alpha) = 0$ por las ecuaciones $\frac{1}{i!}\frac{\partial^{s-1} f}{\partial X^{s-1}}(\alpha,\alpha) = 0$ que poseen coeficientes más chicos, ya que se efecturán simplificaciones.

Paso III- El polinomio f no se anula en la aproximación racional $(\frac{p_1}{q_1},\frac{p_2}{q_2})$.

En el teorema de Liouville, este paso resultaba trivial pues un polinomio en una variable posee finitas raíces racionales. Cuando consideramos más variables, este paso se vuelve el más difícil de todos. De hecho no podremos asegurar que f no se anula, sino que una derivada de orden chico no se anula en $(\frac{p_1}{q_1},\frac{p_2}{q_2})$, es decir, f posee orden de anulación chico en dicho punto. El resultado del paso II seguirá siendo válido para las derivadas y se verá que el orden de anulación de f en $(\frac{p_1}{q_1},\frac{p_2}{q_2})$ es grande en comparación con el orden obtenido en este paso. Llegando así a una contradicción.

Supongamos que tanto f como $\frac{\partial f}{\partial X}$ se anulan en $(\frac{p_1}{q_1},\frac{p_2}{q_2})$. Esto significa que

$$P(\tfrac{p_1}{q_1}) + Q(\tfrac{p_1}{q_1})\tfrac{p_2}{q_2} = 0$$
$$P'(\tfrac{p_1}{q_1}) + Q'(\tfrac{p_1}{q_1})\tfrac{p_2}{q_2} = 0.$$

Eliminando p_2/q_2 de estas ecuaciones obtenemos que

$$P(\tfrac{p_1}{q_1})Q'(\tfrac{p_1}{q_1}) - P'(\tfrac{p_1}{q_1})Q(\tfrac{p_1}{q_1}) = 0.$$

Por lo tanto el determinante Wronskiano

$$W(X) = p(X)q'(X) - p'(X)q(X),$$

aparece naturalmente si asumimos que algunas derivadas se anulan. El analisis del tamaño de los coeficientes de W nos permitirá obtener cota para el orden de anulación de f en la aproximación racional considerada.

Siegel trabajó con polinomios en dos variables más generales, y lo mismo hicieron Dyson y Gelfond. Ya era sabido que usar polinomios en más variables permitiría mejorar el exponente de aproximación. Roth fue capaz de superar las dificultades del paso III cuando se consideraban más variables.

3. Teorema de Roth y generalizaciones

Como ya ha sido mencionado, el teorema de Roth es el siguiente

TEOREMA 1.8 (Roth 1955). *Sea α un número real algebraico de grado $d \geq 2$. Entonces para todo $\varepsilon > 0$, la desigualdad*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\varepsilon}} \tag{1.11}$$

posee finitas soluciones racionales $\frac{p}{q}$ con $q > 0$.

Este teorema puede ser generalizado en diferentes direcciones. Ya en 1920, cuando Siegel [**25**] probó que el exponente $d/2 + 1$ probado por Thue podía ser reemplazado por $2\sqrt{d}$, también logro generalizar este resultado para la aproximación de un número algebraico sobre un cuerpo de números K.

TEOREMA 1.9 (Siegel 1920). *Sea K un cuerpo de números y sea α un número algebraico sobre K de grado $d \geq 2$. Entonces para todo $\varepsilon > 0$, la desigualdad*

$$|\alpha - \beta| < \frac{1}{\Lambda(\beta)^{2\sqrt{d}+\varepsilon}}$$

posee finitas soluciones con $\beta \in K$.

Donde $\Lambda(\beta)$ representa el máximo de los valores absolutos de los coeficientes del polinomio irreducible primitivo sobre $\mathbb{Z}$ que anula a β. En realidad Sigel probó que el exponente d/2+1 podía ser reemplazado por

$$s(d) = \min_{0 \leq s \leq d} \left\{ s + \frac{d}{s+1} \right\} < 2\sqrt{d}.$$

Otro tipo de generalización fue hecha por Mahler [**17**] en 1931 cuando introdujo las aproximaciones diophanticas p-adicas.

TEOREMA 1.10 (Mahler 1932). *Sea $f(X)$ un polinomio irreducible de grado $d \geq 3$ con coeficientes racionales. Sean $p_1, \ldots, p_n$ números primos distintos, y sean $\alpha_0, \alpha_1, \ldots, \alpha_n$ respectivamente una raiz real, p_1-adica, ..., p_n-ádica de $f(X)$. Sea $\varepsilon > 0$ entonces*

$$\min\left\{1, \left|\alpha_0 - \frac{p}{q}\right|\right\} \prod_{t=1}^{n} \min\left\{1, |p - q\alpha_t|_{p_t}\right\} \leq \frac{1}{\max\{|p|, |q|\}^{2\sqrt{d}+\varepsilon}}$$

posee finitas soluciones racionales p/q.

De hecho, Mahler también probó que el exponente $2\sqrt{d}$ podía ser reemplazado por el ya mencionado exponente $s(d)$ del teorema de Siegel. A partir de sus trabajos en aproximaciones p-ádicas, Mahler obtuvo como resultado la finitud de la cantidad de soluciones de la ecuación de Thue en S-enteros(enteros cuyos factores primos están restringidos a un conjunto finito de primos S). Mahler también probó la finitud de la llamada ecuación de Thue-Mahler

$$F(x, y) = p_1^{a_1} \ldots p_n^{a_n}$$

donde F es un polinomio homogéneo en $\mathbb{Z}[x, y]$ con al menos 3 factores lineales distintos $x - \alpha_i y$, $p_1, \ldots, p_n$ primos fijos, a ser resuelta en $x, y \in \mathbb{Z}$ y $a_1, \ldots, a_n$ enteros positivos.

La generalización del teorema de Thue sobre un cuerpo de números K y finitos valores absolutos sobre K fue llevada a cabo por Parry [**20**] en 1950. El análogo del teorema de Roth sobre cuerpos de números fue llevado a cabo por LeVeque [**19**] en 1956 y la generalización del teorema de Thue sobre los números p-ádicos con el exponente del teorema de Roth, es debida a Ridout[**22**] en 1958.

La formulación general del teorema de Roth sobre cuerpos de números es la siguiente y será la probada en esta tesis.

TEOREMA 1.11. *Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$. Dado $\varepsilon > 0$, entonces existen finitos $\beta \in K$ que satisfacen la inecuación*

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

Por ejemplo, si tomamos $K = \mathbb{Q}$, $S = \{|\ |_\infty\}$ el valor absoluto usual sobre $\mathbb{Q}$ y α un número algebraico sobre $\mathbb{Q}$ obtenemos que la inecuacion

$$\min\left\{1, |\alpha - \tfrac{p}{q}|\right\} \leq \tfrac{1}{\max\{|x|,|y|\}^{2+\varepsilon}}$$

posee finitas soluciones racionales p/q. Esta versión es equivalente a la versión clásica del Teorema de Roth anteriormente mencionada.

CAPÍTULO 2

El Teorema de Roth

1. Comentarios sobre el enunciado del teorema de Roth

Haremos a continuación algunos comentarios respecto al enunciado del teorema de Roth. Recordemos primero el teorema en cuestion:

TEOREMA 2.1. *Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$. Dado $\varepsilon > 0$, entonces existen finitos $\beta \in K$ que satisfacen la inecuación*

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

(1) Podemos reemplazar el lado derecho de (2.1) por $C/H_K(\beta)^{2+\varepsilon}$, con $C > 0$ una constante fija. Supongamos que β satisface la desigualdad modificada. Entonces si consideramos los β que además cumplen $C/H_K(\beta)^{2+\varepsilon/2} \leq 1$ tendremos que

$$\prod_{v \in S} \min\left\{\|\beta - \alpha_v\|_v, 1\right\} \leq \frac{C}{H_K(\beta)^{2+\varepsilon}} = \frac{C}{H_K(\beta)^{2+\varepsilon/2}} . \frac{1}{H_K(\beta)^{2+\varepsilon/2}} \leq \frac{1}{H_K(\beta)^{2+\varepsilon/2}},$$

los cuales son finitos por el teorema de Roth. Como solo descartamos un conjunto de β con altura acotada, es decir, un conjunto finito, concluimos que este resultado es equivalente al teorema de Roth.

(2) Dado de que toda extensión de un valor absoluto v sobre K a $\bar{K}$, corresponde a un embedding de $\bar{K}$ en la completación $\bar{K}_v$. Podremos ver a los α_v como elementos de K_v algebraicos sobre K. En efecto, tendremos la siguiente versión del teorema de Roth.

TEOREMA 2.2. *Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K. Para cada $v \in S$, sea $\alpha_v \in K_v$ algebraico sobre K. Dado $\varepsilon > 0$, entonces existen finitos $\beta \in K$ que satisfacen la inecuación*

$$\prod_{v \in S} \min\left\{\|\beta - \alpha_v\|_v, 1\right\} \leq \frac{1}{H_K(\beta)^{2+\varepsilon}}. \tag{2.2}$$

Veamos que este teorema se deduce del teorema de Roth.

Sea $\bar{K}$ una clausura algebraica de K, $\bar{K}_v$ una clausura algebraica de K_v para cada $v \in S$ y $|\ |_v$ la extensión de v a K_v que se extiende de forma única a $\bar{K}_v$. Para cada $v \in S$ sea $f_v \in K[X]$ el polinomio minimal de α_v sobre K. Consideremos $\tilde{\alpha}_v$ una raíz de f_v en $\bar{K}$. Como $\bar{K}_v$

es algebraicamente cerrado, el embedding $\sigma_v : K(\tilde{\alpha}_v)/K \to \bar{K}_v/K$ que manda $\tilde{\alpha}_v$ a α_v, puede ser extendido a $\bar{K}$. Consideremos entonces $|\ |'_v$ la extensión de v a $\bar{K}$ dada por $|x|'_v = |\sigma_v(x)|_v$. Si $\beta \in K$, entonces

$$|\beta - \tilde{\alpha}_v|'_v = |\beta - \alpha_v|_v.$$

Por lo tanto, una solución de (2.2) cumple que

$$\prod_{v\in S} \min\left\{\|\beta - \tilde{\alpha}_v\|'_v, 1\right\} \leq \frac{1}{H_K(\beta)^{2+\varepsilon}},$$

las cuales sabemos que son finitas por la validez del teorema de Roth. Queda así probada lo enunciado. De una forma similar puede verse que de hecho este enunciado es equivalente al teorema de Roth.

(3) Si α y α' son dos elementos distintos algebraicos sobre K y β se acerca a α respecto de v, entonces β estará lejos de α'. Por lo tanto, podemos considerar distintos números algebraicos para un mismo valor absoluto.

(4) No hay razón por la cual no podamos permitirle a β tender a infinito, es decir, podemos tomar $\alpha_v = \infty$ y reemplazar el termino sin sentido $\|\infty - \beta\|_v$ por $\|1/\beta\|_v$. Podemos entonces agregarle a la condición de aproximación del lado izquierdo un producto de la forma

$$\prod_{v\in S} \min\left\{\|1/\beta\|_v, 1\right\}.$$

Esta versión del teorema se puede reducir a la anterior tomando una transformación proyectiva $T(\beta) = (a\beta + b)/(c\beta + d)$, $a, b, c, d \in \mathbb{Z}$, tal que $T(\alpha_v)$ es siempre finito y aplicando el teorema con $T(\alpha_v)$ y $T(\beta)$. Dado que $H_K(T(\beta)) >><< H(\beta)$. Esto ultimo se debe a partir de las siguientes propiedades de la altura : $H_K(A) = H_K(1/A)$, $H_K(A + B) \leq 2.H_K(A)H_K(B)$ y $H_K(A.B) \leq H_K(A)H_K(B)$; se concluye que $H_K(T(\beta)) \leq \tilde{C}H_K(\beta)$ para alguna constante $\tilde{C}$. Esto se debe a que T es una composición de traslaciones, rotaciones e inversiones, y en este caso es facil ver la desigualdad. Para probar lo afirmado, consideremos $T(z) = \frac{az+b}{z+d}$ $a, b, d \in \mathbb{Z}$ tal que $T(\alpha_v) \neq \infty$ para todo $v \in S$,es decir, $\alpha_v \neq -d$. Veremos que para todo $v \in S$, existe una constante $C = C_{v,T} > 0$ tal que

$$\min\left\{1, \|T(\alpha_v) - T(\beta)\|_v\right\} \leq C \min\left\{1, \|\alpha_v - \beta\|_v\right\}. \tag{2.3}$$

Para simplificar la notación, lo probaremos para el valor absoluto no normalizado $|\ |_v$, pues solo habrá que elevar a la potencia n_v-esima. Notaremos $|\ | = |\ |_v$ y $\alpha = \alpha_v$.

Caso 1. $\alpha \neq \infty$

Supongamos β cumple que $|\beta + d| \geq |\alpha + d|/2$ entonces

$$|T(\alpha) - T(\beta)| = \frac{|ad-b||\alpha-\beta|}{|\alpha+d||\beta+d|} \leq \frac{2|ad-b||\alpha-\beta|}{|\alpha+d|^2} = C_1|\alpha - \beta|.$$

Luego, (2.3) es válida para estos β tomando $C = C_1$.

Consideremos ahora los β que cumplen $|\beta + d| \leq |\alpha + d|/2$. Si la afirmación es falsa para estos β, entonces para cada $N \in \mathbb{N}$ existe β_N bajo estas condiciones tal que

$$N. \min\left\{1, |\alpha - \beta_N|\right\} \leq \min\left\{1, |T(\alpha) - T(\beta_N)|\right\}.$$

Ademas tenemos que

$$|\alpha - \beta_N| \geq |\alpha + d| - |d + \beta_N| \geq |\alpha + d| - |\alpha + d|/2 = |\alpha + d|/2,$$

luego para todo N,

$$N \min\{1, |\alpha + d|/2\} \leq N \min\{1, |\alpha - \beta_N|\} < \min\{1, |T(\alpha) - T(\beta_N)|\} \leq 1.$$

Tomando N suficientemente grande llegamos a un absurdo. Por lo tanto, existe una constante C_2 tal que (2.3) es válida para estos β tambien. Tomando C como el máximo de estas dos constantes obtenemos lo buscado.

Caso 2. $\alpha = \infty$

Sea β tal que $|\beta + d| \geq 1$ entonces

$$|\beta| \leq |\beta + d| + |d| \leq |\beta + d| + |d||\beta + d| = |\beta + d|(1 + |d|).$$

Por lo tanto,

$$|T(\alpha) - T(\beta)| = |\tfrac{ad-b}{\beta+d}| \leq C'|\tfrac{1}{\beta}| = C'|\alpha - \beta|.$$

Luego (2.3) se cumple para estos β, tomando $C = C'$.

Sea ahora β tal que $|\beta + d| \leq 1$. En particular, tenemos que

$$|\beta| \leq |\beta + d| + |d| \leq 1 + |d|.$$

Nuevamente, si (2.3) es falso en este caso, construimos una sucesión β_N como antes. Por lo tanto,

$$N \min\{1, 1/(1 + |d|)\} \leq N \min\{1, 1/|\beta||\} = N \min\{1, |\alpha - \beta||\} \leq \min\{1, |T(\alpha_v) - T(\beta)|\} \leq 1.$$

Se sigue entonces como en el caso anterior.

2. Reducción a aproximaciones simultaneas de enteros algebraicos

Antes de comenzar con la prueba del teorema de Roth, haremos dos simplificaciones. La primera de ellas, que será de utilidad técnica para simplificar notación y cantidad de índices, establece que es suficiente probar el teorema para el caso de enteros algebraicos.

Afirmación 1. *Si el teorema de Roth es verdadero para enteros algebraicos, entonces es verdadero para números algebraicos arbitrarios.*

Demostración. Lo probaremos por el contrarecíproco. Para cada $v \in S$, sea α_v un número algebraico y supongamos que el teorema de Roth es falso para los elementos $\{\alpha_v\}_{v \in S}$ y algún $\varepsilon > 0$. Tenemos entonces infinitos $\beta \in K$ que satisfacen (2.1). Para cada β, existe al menos un subconjunto de S, digamos S', tal que

$$\prod_{v \in S'} \|\beta - \alpha_v\|_v = \prod_{v \in S} \min\{\|\beta - \alpha_v\|_v, 1\}$$

(descartando todos o algunos de los $v \in S$ que cumplen $\min\{\|\beta - \alpha_v\|_v, 1\} = 1$).

Como S posee finitos subconjuntos y los β son infinitos, quizás reemplazando S por un subconjunto, podremos suponer que existen infinitos $\beta \in K$ para los cuales

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

Para cada α_v, existe $D_v \in \mathbb{N}$ tal que $D_v\alpha_v$ es entero algebraico. Existe entonces $D > 0$ entero, por ejemplo el producto de todos los D_v, tal que $D\alpha_v$ es entero algebraico para todo $v \in S$. Consideremos los $\beta \in K$ solución de (2.1) tal que $H_K(\beta) > H_K(D)^{1+6/\varepsilon}$. Hay infinitos de ellos pues solo descartamos un conjunto de altura acotada. A partir de la definición de altura se tiene que $H_k(D\beta) \leq H_K(D)H_K(\beta)$. Además,

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

Por lo tanto

$$\prod_{v\in S} \min\{\|D\beta - D\alpha_v\|_v, 1\} \leq \prod_{v\in S} \|D\beta - D\alpha_v\|_v = \prod_{v\in S} \|D\|_v \prod_{v\in S} \|\beta - \alpha_v\|_v \leq$$
$$\frac{H_K(D)}{H_K(\beta)^{2+\varepsilon}} = \frac{H_K(D)}{H_K(\beta)^{2+\varepsilon/2}} \cdot \frac{1}{H_K(\beta)^{2+\varepsilon/2}} \leq$$
$$\frac{H_K(D)}{(H_K(D\beta)/H_K(D))^{2+\varepsilon/2}} \cdot \frac{1}{(H_K(D)^{1+6/\varepsilon})^{\varepsilon/2}} = \frac{1}{H_K(D\beta)^{2+\varepsilon/2}}.$$

Se obtiene así que el teorema de Roth es falso para los enteros algebraicos $\{D\alpha_v\}_{v\in S}$, con $\varepsilon' = \varepsilon/2$. □

La segunda simplificación reemplazara la condición de que el producto $\prod \|\alpha_v - \beta\|$ sea pequeño por la más sencilla condición de cada una de las diferencias $\|\alpha_v - \beta\|$ sea pequeña. Esta idea de reducir el trabajo a considerar aproximaciones simultaneas es debida a Mahler, quien fue el primero en considerar aproximaciones sobre varios valores absolutos.

TEOREMA 2.3. *Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$. Sea $\varepsilon > 0$ y sea*

$$\xi : S \to [0, 1] \text{ una función que cumple } \sum_{v\in S} \xi_v = 1.$$

Entonces existen finitos $\beta \in K$ que satisfacen

$$\min\{\|\alpha_v - \beta\|, 1\} \leq \frac{1}{H_K(\beta)^{(2+\varepsilon)\xi_v}} \tag{2.4}$$

para todo $v \in S$.

Veamos que este teorema es equivalente al teorema de Roth.

TEOREMA 2.4. *El teorema 2.3 es equivalente al teorema de de Roth.*

DEMOSTRACIÓN. Supongamos que el teorema de Roth es cierto. Sea $\xi : S \to [0, 1]$ como en el enunciado y sea $\beta \in K$ que cumple (2.4). Multiplicando las desigualdades y usando que $\sum_v \xi_v = 1$ se tiene entonces que β cumple (2.1). Aplicando el teorema de Roth concluimos que hay finitas posibilidades para β, como se quería ver.

Supongamos ahora que el teorema sobre aproximaciones simultaneas es verdadero. Para cada $\beta \in K$ que cumple (2.1) y para cada $v \in S$, se define el número real $\lambda_v(\beta) \geq 0$ mediante

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

Para que $\lambda_v(\beta)$ este unívocamente determinado, supondremos que $H_K(\beta) \neq 1$, pues como mucho estamos eliminando finitas soluciones. Multiplicando sobre $v \in S$ y comparando con (2.1) se tiene que $\sum_{v\in S} \lambda_v(\beta) \geq 1$.

Para simplificar la notación, llamemos $\mu = \frac{2+\varepsilon}{2+\varepsilon/2}$. Observar que $\mu > 1$. Sea A un entero de forma tal que

$$A\,(\mu - 1) > s.$$

Utilizando repetidas veces el hecho de que $[x + y] \leq [x] + [y] + 1$, se tendrá

$$A + s \leq A\mu \leq \left[\sum_{v\in S} A\mu\lambda_v(\beta)\right] + 1 \leq \sum_{v\in S} [A\mu\lambda_v(\beta)] + s,$$

entonces

$$A \leq \sum_{v\in S} [A\mu\lambda_v(\beta)].$$

Por lo tanto, existen enteros $b_v(\beta)$ tal que

$$0 \leq b_v(\beta) \leq [A\mu\lambda_v(\beta)] \leq A\mu\lambda_v(\beta) \;\text{ y }\; \sum_{v\in S} b_v(\beta) = A.$$

Consideramos el conjunto de funciones

$$\xi : S \to [0, 1] \;\text{ con } \xi_v = \tfrac{a_v}{A},\ a_v \in \mathbb{Z}, a_v \geq 0, \text{ y } \textstyle\sum_{v\in S} a_v = A,$$

este conjunto resulta finito y será denotado por Ω.

Por lo tanto, la función $\xi : S \to [0, 1]$ definida por $\xi_v = b_v(\beta)/A$ pertenece al conjunto Ω. Además, como $(2 + \varepsilon/2)\xi_v \leq (2 + \varepsilon)\lambda_v(\beta)$, se tendrá que

$$\min\left\{\|\alpha_v - \beta\|_v, 1\right\} \leq \tfrac{1}{H(\beta)^{(2+\varepsilon/2)\xi_v}} \;\text{ para todo }\; v \in S.$$

Se concluye así que para cada $\beta \in K$, que cumple (roth) existe al menos una función $\xi \in \Omega$ para la cual cumple (2.4). Por hipótesis, para cada ξ hay finitos β, por lo tanto, como Ω posee finitos elementos, se concluye que * posee finitas soluciones $\beta \in K$. □

3. Preliminares

En esta sección probaremos varios resultados preliminares que será utilizados a lo largo de la prueba del teorema de Roth. Empezaremos con un estudio sobre el tamaño de los coeficientes de polinomios en varias variables.

Sea k un cuerpo y $k[X_1, \ldots, X_n]$ el anillo de polinomios en n variables, abreviaremos $\mathbf{i} = (i_1, \ldots, i_n)$ y $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n_{\geq 0}$ y notaremos

$$\mathbf{x}^{\mathbf{m}} = x_1^{m_1} \ldots x_n^{m_n}$$

y

$$\binom{\mathbf{m}}{\mathbf{i}} = \prod_{j=1}^{n} \binom{m_j}{i_j},$$

donde entendemos que $\binom{m}{i} = 0$ si $m < i$.

Definición 1. *Sea k un cuerpo. Para cada $\boldsymbol{i}$, se define el operador lineal $\partial_{\boldsymbol{i}}$ sobre el anillo de polinomios $k[X_1, \ldots, X_n]$ como*

$$\partial_{\boldsymbol{i}} \boldsymbol{x}^{\boldsymbol{m}} = \binom{\boldsymbol{m}}{\boldsymbol{i}} \boldsymbol{x}^{\boldsymbol{m}-\boldsymbol{i}}.$$

También podremos notar $\partial_{\mathbf{i}} = \partial_{i_1 \dots i_n}$. Observar que este operador está bien definido en cualquier cuerpo pues los binomios son números enteros. Además, si el cuerpo tiene característica cero vale la igualdad

$$\partial_{\mathbf{i}} P = \frac{1}{i_1! i_2! \dots i_n!} \frac{\partial^{i_1+\dots+i_n}}{\partial x_1^{i_1} \dots \partial x_n^{i_n}} P.$$

Los operadores diferenciales están normalizados de esta forma para simplificar lo más posible los factores comunes que aparecen cuando diferenciamos un polinomio.

DEFINICIÓN 2. *Sea $P \in \mathbb{C}[X_1, \dots, X_m]$ un polinomio con coeficientes complejos, definimos la altura de P como*

$$|P| = \textit{máximo de los valores absolutos de los coeficientes de } P.$$

LEMA 1. *Sea $P(X_1, \dots, X_m) \in \mathbb{Z}[X_1, \dots, X_m]$ un polinomio con coeficientes enteros, y sea $\boldsymbol{i} = (i_1, \dots, i_m)$ una m-upla de enteros no negativos. Las siguientes afirmaciones son validas*

(1) $\partial_{\boldsymbol{i}} P \in \mathbb{Z}[X_1, \dots, X_m]$
(2) *Si $gr_{X_h}(P) \leq r_h$ para todo $1 \leq h \leq m$, entonces $|\partial_{\boldsymbol{i}} P| \leq 2^{r_1+\dots+r_m} |P|$*

DEMOSTRACIÓN. Sea

$$P(X_1, \dots, X_m) = \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1 \dots j_m} X_1^{j_1} \dots X_m^{j_m}$$

con $C_{j_1 \dots j_m} \in \mathbb{Z}$. Diferenciando P obtenemos

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

Como los números combinatorios son enteros, queda probado 1).

Para probar 2), recordemos la estimación de los números combinatorios dada por

$$\binom{j}{i} \leq \sum_{k=0}^{j} \binom{j}{k} = (1+1)^j = 2^j$$

Por lo tanto, se tendrá que

$$|\partial_{i_1 \dots i_m} P| = max \left| C_{j_1 \dots j_m} \tbinom{j_1}{i_1} \dots \tbinom{j_m}{i_m} \right| \leq max\, |C_{j_1 \dots j_m}| \,.max\ 2^{j_1+\dots+j_m} \leq |P|.2^{r_1+\dots+r_m}.$$

□

Estos operadores tienen otras ventajas como por ejemplo que la formula de Taylor queda expresada de forma más concisa. Sea $\mathbf{a} = (a_1, \dots, a_n) \in k^n$ entonces,

$$P(X_1, \dots, X_n) = \sum_{\mathbf{i}} \partial_{\mathbf{i}} P(\mathbf{a}) (\mathbf{x} - \mathbf{a})^{\mathbf{i}}$$

La formula de Leibniz de la derivada del producto es también más sencilla

$$\partial_n \left(P_1(X) P_2(X) \dots P_s(X) \right) = \textstyle\sum_{j_1+\dots+j_s=n} \partial_{j_1} P_1(X) \dots \partial_{j_s} P_s(X).$$

El único aspecto negativo de estos operadores es que no satisfacen la formula de composición que si poseen los diferenciales clásicos, por el contrario se tiene

$$\partial_i \partial_j P = \tbinom{j+i}{i} \partial_{i+j} P.$$

A continuación, definiremos una noción de orden de anulación para polinomios en varias variables que será fundamental en la prueba del teorema de Roth.

DEFINICIÓN 3. *Sea k un cuerpo, $P(X_1,\dots,X_m) \in k[X_1,\dots,X_m]$ un polinomio, $(\alpha_1,\dots,\alpha_m) \in k^m$ y sea $(r_1,\dots,r_m) \in \mathbb{Z}^m_{>0}$. Se define el índice de P respecto a $(r_1,\dots,r_m)$ en $(\alpha_1,\dots,\alpha_m)$ como*

$$Ind(P;r_1,\dots,r_m;\alpha_1,\dots,\alpha_m) = \min_{\boldsymbol{i}} \left\{ \frac{i_1}{r_1} + \dots + \frac{i_m}{r_m} | \partial_{\boldsymbol{i}} P(\alpha_1,\dots,\alpha_m) \neq 0 \right\}.$$

Si P es el polinomio nulo definimos su índice como ∞. Cuando $(r_1,\dots,r_m)$ y $(\alpha_1,\dots,\alpha_m)$ estén fijos, o se entiendan sin ambiguedad por contexto, notaremos $Ind\ P$.

Si bien los r_i no tienen a priori ninguna conexión con los grados de P, en la práctica solo aplicaremos esta noción para polinomios con $\text{grado}_{X_i} \leq r_i$. Se cumple además que $Ind\ P \geq 0$, y la igualdad se cumple si y solo si $P(\alpha_1,\dots,\alpha_m) \neq 0$. En el caso de polinomios de una variable, si $gr(f) = r$ entonces

$$Ind(f;r;\alpha) = \frac{mult_\alpha(f)}{r},$$

cociente que ya había aparecido en el primer intento de generalizar el teorema de Liouville. El siguiente lema justificara que el índice es una forma de medir el orden de anulacion, en particular, 2) y 3) dirán que el índice es una valuación en $k[X_1,\dots,X_m]$.

LEMA 2. *Sean P y $P' \in k[X_1,\dots,X_m]$ polinomios, y fijemos $(r_1,\dots,r_m) \in \mathbb{Z}^m_{>0}$ y $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_m) \in k^m$. Entonces, el índice respecto de $(r_1,\dots,r_m)$ en el punto $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_m)$ cumple las siguientes propiedades:*

1. $Ind(\partial_{i_1\dots i_m} P) \geq Ind\ P - \left(\frac{i_1}{r_1} + \dots + \frac{i_m}{r_m}\right)$
2. $Ind(P+P') \geq min\{Ind\ P, Ind\ P'\}$
3. $Ind(PP') = Ind\ P + Ind\ P'$

DEMOSTRACIÓN. 1) Sea $Q = \partial_{\mathbf{i}} P$. Por definición de índice, tomamos $\mathbf{j} = (j_1,\dots,j_m)$ tal que $\partial_{\mathbf{j}} Q(\boldsymbol{\alpha}) \neq 0$ e $Ind\ Q = \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m}$. Entonces $\partial_{\mathbf{j}} Q(\boldsymbol{\alpha}) \neq 0$ implica que $\partial_{\mathbf{i}+\mathbf{j}} P(\alpha) \neq 0$. Por lo tanto, por definición de índice,

$$\frac{j_1+i_1}{r_1} + \dots + \frac{j_m+i_m}{r_m} \geq Ind\ P$$

con lo cual

$$Ind(\partial_{\mathbf{i}} P) = \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq Ind\ P - \left(\frac{i_1}{r_1} + \dots + \frac{i_m}{r_m}\right).$$

2)Sea $\mathbf{j} = (j_1,\dots,j_m)$ tal que $\partial_{\mathbf{j}}(P+P')(\boldsymbol{\alpha}) \neq 0$ e $Ind(P+P') = \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m}$. Por linealidad de la derivada, $\partial_{\mathbf{j}} P(\boldsymbol{\alpha}) \neq 0$ o $\partial_{\mathbf{j}} P'(\boldsymbol{\alpha}) \neq 0$, entonces

$$\frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq Ind(P) \text{ o } \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m} \geq Ind(P').$$

Consecuentemente,

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

3)Sea $\mathbf{j} = (j_1,\dots,j_m)$ tal que $\partial_{\mathbf{j}}(PP')(\boldsymbol{\alpha}) \neq 0$ e $Ind(PP') = \frac{j_1}{r_1} + \dots + \frac{j_m}{r_m}$. Usando la regla del producto, tenemos que

$$\partial_{\mathbf{j}}(PP') = \sum_{\mathbf{i}+\mathbf{i}'=\mathbf{j}} \partial_{\mathbf{i}} P . \partial_{\mathbf{i}'} P'.$$

Luego existen $\mathbf{i} = (i_1, \ldots, i_m)$, $\mathbf{i}' = (i'_1, \ldots, i'_m)$ con $\mathbf{i} + \mathbf{i}' = \mathbf{j}$ tal que $\partial_{\mathbf{i}} P(\boldsymbol{\alpha}) \neq 0$ y $\partial_{\mathbf{i}'} P'(\boldsymbol{\alpha}) \neq 0$. Entonces

$$Ind\ P \leq \tfrac{i_1}{r_1} + \ldots + \tfrac{i_m}{r_m} \ \text{ y } \ Ind\ P' \leq \tfrac{i'_1}{r_1} + \ldots + \tfrac{i'_m}{r_m}.$$

Sumando estas inecuaciones obtenemos que $Ind\ P + Ind\ P' \leq Ind(PP')$.

Para ver la desigualdad opuesta, consideremos todas las m-uplas $\mathbf{i} = (i_1, \ldots, i_m)$ tal que $Ind\ P = \displaystyle\sum_{h=1}^{m} \frac{i_h}{r_h}$ y $\partial_{\mathbf{i}} P(\boldsymbol{\alpha}) \neq 0$. Sea $\bar{\mathbf{i}}$ la menor de ellas respecto del orden lexicográfico. Es decir, si $\mathbf{i}$ es una de estas m-uplas, distinta de $\bar{\mathbf{i}}$ entonces existe $1 \leq k \leq m$ tal que $\bar{i}_h = i_h$ para todo $1 \leq h < k$ e $\bar{i}_h < i_h$. Similarmente elegimos $\bar{\mathbf{i}}'$ la menor m-upla respecto del orden lexicográfico para P' tal que $Ind\ P = \displaystyle\sum_{h=1}^{m} \frac{\bar{i}'_h}{r_h}$ y $\partial_{\bar{\mathbf{i}}'} P(\boldsymbol{\alpha}) \neq 0$, y definimos $\bar{\mathbf{j}} = \bar{\mathbf{i}} + \bar{\mathbf{i}}'$.Entonces

$$\partial_{\bar{\mathbf{j}}}(PP')(\boldsymbol{\alpha}) = \partial_{\bar{\mathbf{i}}} P(\boldsymbol{\alpha}) \partial_{\bar{\mathbf{i}}'} P'(\boldsymbol{\alpha}) \neq 0$$

pues todas las demás son cero por elección de $\bar{\mathbf{i}}$ e $\bar{\mathbf{i}}'$.Luego obtenemos que

$$Ind(PP') \leq \sum_{h=1}^{m} \frac{\bar{j}_h}{r_h} = \sum_{h=1}^{m} \frac{\bar{i}_h + \bar{i}'_h}{r_h} = Ind\ P + Ind\ P',$$

lo que nos da la otra desigualdad, completando la prueba del enunciado.

□

Al igual que en el teorema de Thue, la construcción del polinomio será llevada acabo a partir de la existencia de soluciones de un sistema de ecuaciones lineales sobre $\mathbb{Z}$. Para eso tendremos que contar la cantidad de ecuaciones que obtenemos al imponer condiciones sobre el índice del polinomio. El siguiente lema combinatorio, lidia con este problema. Una versión similar a este lema ya habia sido probada por Schneider [**31**] en 1936. El lema aquí enunciado es el mismo al publicada en el trabajo original de Roth cuya demostración es debida a Davenport.

LEMA 3. *Supongamos que $r_1, \ldots, r_m$ son enteros positivos, y sea $\lambda > 0$, entonces el número de m-uplas $(j_1, \ldots, j_m) \in \mathbb{Z}^m$ tal que*

$$0 \leq j_h \leq r_h \quad (h = 1, \ldots, m) \quad y \quad \tfrac{j_1}{r_1} + \ldots + \tfrac{j_m}{r_m} \leq \tfrac{1}{2}(m - \lambda)$$

es como mucho $2m^{1/2}\lambda^{-1}(r_1 + 1)\ldots(r_m + 1)$

DEMOSTRACIÓN. Probaremos el lema por indución en m. Si $m = 1$ el resultado es valido pues el número de enteros j_1 que satisfacen

$$0 \leq j_1 \leq r_1 \ \text{ y } j_1 \leq 1/2(1 - \lambda) r_1$$

es cero si $\lambda > 1$ y si $\lambda \leq 1$, dado que el conjunto tiene como mucho $r_1 + 1$ elementos se tiene que $r_1 + 1 \leq \frac{1}{\lambda} 2(r_1 + 1)$.

Sea ahora $m > 1$ y supongamos que el resultado es válido para $m - 1$. Podemos suponer además que $\lambda > 2m^{1/2}$ ya que caso contrario, el lema es válido pues hay como mucho $(r_1 + 1)\ldots(r_m + 1)$ m-uplas. Para cada valor de j_m, la condición que se le pide a $j_1, \ldots, j_{m-1}$, es de la misma naturaleza que nuestro problema solo que en $m - 1$ lugares en lugar de m y con λ reemplazado por λ' tal que $1/2(m - 1 - \lambda') = 1/2(m - \lambda) - j_m/r_m$, es decir, $\lambda' = \lambda - 1 + 2j_m/r_m$. Observar

que $\lambda' > 0$ ya que $\lambda > 2m^{1/2} > 1$. La cantidad de m-uplas va a ser menor o igual a la cantidad de $(m-1)$-uplas para cada valor fijo $0 \leq j_m \leq r_m$, y por hipótesis inductiva, el total de m-uplas será menor o igual que

$$\sum_{j_m=0}^{r_m} \frac{2(m-1)^{1/2}}{\lambda - 1 + 2j_m/r_m}(r_1+1)\dots(r_{m-1}+1)$$

Bastara entonces probar que

$$\sum_{j=0}^{r} \frac{1}{\lambda - 1 + 2j/r} < \lambda^{-1}(m-1)^{-1/2}m^{1/2}(r+1)$$

para todo entero positivo r y m, con $\lambda > 2m^{1/2}$.

Supongamos que r es par, y reemplazando j por $\frac{1}{2}r + k$, obtenemos que

$$\sum_{k=-\frac{r}{2}}^{\frac{r}{2}} \frac{1}{\lambda + 2k/r} = \lambda^{-1} + \sum_{k=1}^{\frac{r}{2}} \left(\frac{1}{\lambda + 2k/r} + \frac{1}{\lambda - 2k/r} \right) = \lambda^{-1} + \sum_{k=1}^{\frac{r}{2}} 2\lambda \frac{1}{\lambda^2 - 4k^2/r^2} \leq$$

$$\lambda^{-1} + 2\lambda \sum_{k=1}^{\frac{r}{2}} \frac{1}{\lambda^2 - 1} = \lambda^{-1} + 2\lambda^{-1} \sum_{k=1}^{\frac{r}{2}} \frac{1}{1 - \lambda^{-2}} \leq \frac{\lambda^{-1}(r+1)}{1 - \lambda^{-2}}$$

como $\lambda > 2m^{1/2}$ entonces $1 - \lambda^{-2} > 1 - 1/4m^{-1} > (1 - m^{-1})^{1/2}$ se obtiene lo buscado.

Si ahora r es impar, realizando el cambio $j = (r-1)/2 + k$ se obtiene que

$$\sum_{k=-\frac{r-1}{2}}^{\frac{r+1}{2}} \frac{1}{\lambda + (2k-1)/r} = \sum_{k=1}^{\frac{r+1}{2}} \left(\frac{1}{\lambda + (2k-1)/r} + \frac{1}{\lambda - (2k-1)/r} \right) =$$

$$2\lambda \sum_{k=1}^{\frac{r+1}{2}} \frac{1}{\lambda^2 - (2k-1)^2/r^2} < \frac{\lambda(r+1)}{\lambda^2 - 1} = \frac{\lambda^{-1}(r+1)}{1 - \lambda^{-2}}.$$

□

Es posible dar una interpretación probabilística de este teorema. Si elegimos al azar un m-upla $(j_1, \dots, j_m)$ con $0 \leq j_h \leq r_h$, entonces el valor esperado de j_h/r_h será $1/2$ y por lo tanto el valor esperado de $\sum j_h/r_h$ será de $m/2$.

Si α es un entero algebraico de grado d sobre $\mathbb{Q}$, entonces todo elemento del anillo $\mathbb{Z}[\alpha]$ se puede escribir de manera única como combinación $\mathbb{Z}$ lineal de la base $1, \alpha, \dots, \alpha^{d-1}$. El siguiente lema nos dirá que tan grande serán los coordenadas de las potencias de α en esta base y así obtener cierto control sobre el tamaño de los coeficientes que aparecerán en nuestro sistema de ecuaciones lineales.

LEMA 4. *Sea α un entero algebraico cuyo polinomio minimal es $P(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$. Para cada entero $l \geq 0$ existen enteros $a_0^{(l)}, \dots, a_{n-1}^{(l)}$ tal que*

$$\alpha^l = a_{n-1}^{(l)}\alpha^{n-1} + \dots + a_1^{(l)}\alpha + a_0^{(l)}$$

y vale que

$$|a_j^{(l)}| \leq (1 + |P|)^l \qquad j = 0, \dots, n-1$$

DEMOSTRACIÓN. Procederemos por inducción. Para $l \leq n$ el teorema es trivial. Supongamos entonces que el teorema es cierto para $l - 1$

$$\alpha^l = \alpha^{l-1}\alpha = (a_{n-1}^{(l-1)}\alpha^{n-1} + \ldots + a_1^{(l-1)}\alpha + a_0^{(l-1)})\alpha = a_{n-1}^{(l-1)}\alpha^n + \ldots + a_1^{(l-1)}\alpha^2 + a_0^{(l-1)}\alpha = a_{n-1}^{(l-1)}(-a_{n-1}\alpha^{n-1} - \ldots - a_1\alpha - a_0) + \ldots + a_1^{(l-1)}\alpha^2 + a_0^{(l-1)}\alpha = (a_{n-2}^{(l-1)} - a_{n-1}a_{n-1}^{(l-1)})\alpha^{n-1} + \ldots + (a_0^{(l-1)} - a_1 a_{n-1}^{(l-1)})\alpha - a_0 a_{n-1}^{(l-1)}$$

Se tiene entonces que para $j = n-1, \ldots, 1$

$$|a_j^{(l)}| = |a_{j-1}^{(l-1)} - a_j a_{n-1}^{(l-1)}| \leq (1+|P|)^{l-1}(1+|a_j|) \leq (1+|P|)^l$$

y

$$|a_0^{(l)}| = |a_0 a_{n-1}^{(l-1)}| \leq |P|(1+|P|)^{l-1} \leq (1+|P|)^l.$$

□

El siguiente lema sistematiza el uso de Thue del principio del palomar para dar una cota superior a la solución de un sistema de ecuaciones lineales. Siegel fue el primero en formalizar estas ideas, es por eso que el lema lleva su nombre, a veces también llamando lema de Thue-Siegel.

LEMA 5 (Lema de Siegel). *Sean a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ enteros no todos nulos, acotados por $A > 0$ y supongamos $n > m$. Entonces el sistema homogéneo*

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & 0 \end{array} \tag{2.5}$$

admite una solución entera no trivial $\boldsymbol{x} = (x_1, \ldots, x_n)$ tal que

$$\max_i |x_i| \leq \left[(nA)^{\frac{m}{n-m}}\right] \tag{2.6}$$

DEMOSTRACIÓN. Dado que por hipótesis tenemos más incógnitas que ecuaciones, siempre existe una solución racional no nula y multiplicando por una constante adecuada, obtenemos una solución entera no nula.

Sea H un entero positivo y consideremos el conjunto

$$T := \{\mathbf{x} \in \mathbb{Z}^n \mid 0 \leq x_i \leq H, i = 1, \ldots, n\}.$$

La cantidad de puntos enteros en T es $(H+1)^n$ pues para cada variable tenemos $H+1$ posibilidades.

Sea $F : \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal dada por $F(\mathbf{x}) = (L_1(\mathbf{x}), \ldots, L_n(\mathbf{x}))$ donde $L_j = a_{j1}x_1 + \cdots + a_{jn}x_n$.

Si $\mathbf{x} \in T$, se tiene que $-B_j H \leq L_j(\mathbf{x}) \leq C_j H$ donde $-B_j$ y C_j son la suma de los coeficientes negativos y positivos de L_j, respectivamente. Dado que $B_j + C_j \leq nA$, cada $L_j(\mathbf{x})$ pertenece a un intervalo de longitud $\leq nAH$. Por lo tanto, cada $L_j(\mathbf{x})$ toma como mucho $nAH+1$ valores, lo que implica que $\#F(T) \leq (nAH+1)^m$.

Si H es tal que

$$(nAH+1)^m < (H+1)^n$$

esto nos diria que F no puede ser inyectiva, entonces existirían $\mathbf{x}' \neq \mathbf{x}''$ tal que $F(\mathbf{x}') = F(\mathbf{x}'')$. Tomando $\tilde{\mathbf{x}} = \mathbf{x}' - \mathbf{x}''$, se tiene que $F(\tilde{\mathbf{x}}) = 0$. Por lo tanto, $\tilde{\mathbf{x}}$ resulta solución entera no nula del sistema homogéneo. Observar que $|\tilde{\mathbf{x}}_i| \leq H$ ya que $\mathbf{x}'$ y $\mathbf{x}''$ pertenecen a $T \subset [0, H]^n$).

Basta elegir $H = \left[(nA)^{\frac{m}{n-m}}\right]$ pues entonces

$$(nAH+1)^m < (nA(H+1))^m = (nA)^m(H+1)^m \leq (H+1)^{n-m}(H+1)^m = (H+1)^n.$$

□

El siguiente lema es básicamente el hecho de que no existen enteros entre 0 y 1. Según Mahler, en un cuerpo que posea una formula del producto y una desigualdad de este estilo será posible desarrollar una teoría de aproximaciones diofánticas.

LEMA 6. *Sea K un cuerpo de números, $\alpha \in K^*$, y sea $S \subset M_K$ un subconjunto de valores absolutos sobre K. Entonces*

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

DEMOSTRACIÓN.

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

$$\prod_{v \in M_K} max\left\{1, \frac{1}{\|\alpha\|_v}\right\} = \prod_{v \in M_K} \frac{1}{min\left\{1, \|\alpha\|_v\right\}} \geq \prod_{v \in S} \frac{1}{min\left\{1, \|\alpha\|_v\right\}}.$$

□

4. Construcción del polinomio auxiliar

En esta sección probaremos el análogo al Paso I, descripto en la demostración del teorema de Liouville y el teorema de Thue. Construiremos un polinomio $P(X_1, \ldots, X_m)$ con coeficientes enteros de un tamaño razonablemente chico que tendrá cada m-upla $(\alpha_v, \ldots, \alpha_v)$ como ceros de orden grande, según nuestra noción de orden de anulación. Como ya ha sido mencionado, la construcción será llevada a cabo a partir de una solución no trivial de un sistema de ecuaciones lineales con coeficientes enteros.

TEOREMA 2.5. *Sean $\alpha_1, \ldots, \alpha_s$ enteros algebraico, cada uno de grado d_s sobre $\mathbb{Q}$. Sea $\varepsilon > 0$, y sea m un entero que cumple*

$$m > 16 \left(\sum_{t=1}^{s} d_j\right)^2 \varepsilon^{-2}. \tag{2.7}$$

Sean $r_1, \ldots, r_m$ enteros positivos. Entonces existe un polinomio $P(X_1, \ldots, X_m)$ no nulo con coeficientes enteros tal que

1. *P tiene grado $\leq r_h$ en la variable X_h.*
2. *El indice de P respecto de $(r_1, \ldots, r_m)$ en $(\alpha_t, \ldots, \alpha_t)$ es tal que*
$$Ind(P) \geq \tfrac{m}{2}(1-\varepsilon) \text{ para todo } 1 \leq t \leq s.$$
3. *$|P| \leq B^{r_1+\ldots+r_m}$, con $B = B(\alpha_1, \ldots, \alpha_s)$ una constante que solo depende de $\alpha_1, \ldots, \alpha_s$.*

DEMOSTRACIÓN. La prueba consistirá en considerar los coeficientes de P como incóngnitas que deberán cumplir un sistema de ecuaciones lineales homogéneas provenientes de la condición 2), el cual tendrá más ecuaciones que incóngnitas, probando así la existencia de dicho polinomio.

Escribamos a P como

$$P(X_1, \ldots, X_m) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} C_{j_1 \ldots j_m} X_1^{j_1} \ldots X_m^{j_m},$$

donde $C_{j_1\dots j_m}$ son enteros a determinar para que se cumpla 2). El número total de coeficientes es

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

Para cada m-upla $(i_1, \dots, i_m)$ se tiene que

$$\partial_{i_1\dots i_m} P = \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1\dots j_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} X_1^{j_1-i_1} \dots X_m^{j_m-i_m}.$$

Para que valga 2), para todo α_t debe cumplirse que

$$\partial_{i_1\dots i_m} P(\alpha_t, \dots, \alpha_t) = 0 \text{ para todo } (i_1, \dots, i_m) \text{ con } \sum_{h=0}^{m} \frac{i_h}{r_h} < \frac{m}{2}(1-\varepsilon).$$

Para disminuir la cantidad de índices en las cuentas siguientes, llamemos α, a cualquiera de estos enteros algebraicos y sea d a su grado. Evaluando en $(\alpha, \dots, \alpha)$ y usando el Lema 4 para escribir las potencias de α como combinación lineal de $1, \alpha, \dots, \alpha^{d-1}$, obtenemos que

$$\begin{aligned}\partial_{i_1\dots i_m} P(\alpha, \dots, \alpha) &= \sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1\dots j_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} \alpha^{j_1-i_1+\dots+j_m-i_m} = \\ &\sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1\dots j_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} \left(\sum_{k=0}^{d-1} a_k^{j_1-i_1+\dots+j_m-i_m} \alpha^k \right) = \\ &\sum_{k=0}^{d-1} \left(\sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1\dots j_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} a_k^{j_1-i_1+\dots+j_m-i_m} \right) \alpha^k\end{aligned}$$

Como $1, \alpha, \dots, \alpha^{d-1}$ son linealmente independientes, $\partial_{i_1\dots i_m} P(\alpha, \dots, \alpha)$ será cero si y solo si

$$\sum_{j_1=0}^{r_1} \dots \sum_{j_m=0}^{r_m} C_{j_1\dots j_m} \binom{j_1}{i_1} \dots \binom{j_m}{i_m} a_k^{j_1-i_1+\dots+j_m-i_m} = 0$$

para $0 \le k \le d-1$. Estas son d ecuaciones lineales homogéneas con coeficientes enteros para cada m-upla $(i_1, \dots, i_m)$ tal que

$$\sum_{h=0}^{m} \frac{i_h}{r_h} < \frac{m}{2}(1-\varepsilon) = \frac{1}{2}(m - \varepsilon m)$$

Según Lema 3 con $\lambda = \varepsilon m$, la cantidad total de ecuaciones M' será

$$M' \le 2dm^{-1/2}\varepsilon^{-1}(r_1 + 1)\dots(r_m + 1) = 2dm^{-1/2}\varepsilon^{-1}N.$$

Dado que queremos que esto suceda para todo α_t, la cantidad total de ecuaciones M será

$$M \le 2\left(\sum_{t=1}^{s} d_t\right) m^{-1/2}\varepsilon^{-1}N \le \frac{N}{2},$$

donde la ultima desigualdad es válida por la hipótesis hecha sobre m.

Por lo tanto, tenemos M ecuaciones lineales homogéneas con coeficientes enteros en N incóngnitas. Aplicaremos ahora el Lema de Siegel para así obtener una cota sobre la altura de nuestro polinomio. Para esto, necesitamos tener una estimación de los coeficientes de las ecuaciones. Por Lema 4 sabemos que $|a_k^{(l)}| \leq (1+|Q|)^l$, con Q el polinomio minimal sobre $\mathbb{Q}$ del α considerado. Consecuentemente podemos estimar los coeficientes del sistema lineal por

$$\left|\binom{j_1}{i_1}\dots\binom{j_m}{i_m}a_k^{j_1-i_1+\dots+j_m-i_m}\right| \leq 2^{j_1+\dots+j_m}(|Q|+1)^{j_1+\dots+j_m} \leq (2|Q|+2)^{r_1+\dots+r_m}.$$

Esta cota, nos sirve para las ecuaciones que provienen de considerar un solo α. Con lo cual, la cota que se obtiene de considerar todos los α_t será el máximo de todas estas. Digamos $A = \max|Q_t|$ donde Q_t es el polinomio minimal de α_t, entonces los coeficientes de nuestro sistema estarán acotados por $(2A+2)^{r_1+\dots+r_m}$. Aplicando el lema de Siegel a nuestro sistema obtenemos que los coeficientes de P están acotados por

$$\begin{gathered}|P| \leq (N(2A+2)^{r_1+\dots+r_m})^{\frac{M}{N-M}} \leq N(2A+2)^{r_1+\dots+r_m} \leq \\ 2^{r_1+\dots+r_m}(2A+2)^{r_1+\dots+r_m} = B(\alpha_1,\dots,\alpha_s)^{r_1+\dots+r_m}\end{gathered}$$

con $B(\alpha_1,\dots,\alpha_s) = 4A+4$. La segunda desigualdad se debe a que $\frac{M}{N-M} \leq 1$ pues $M \leq \frac{N}{2}$. □

5. El indice es grande

Una vez que tenemos el polinomio con coeficientes de tamaño "chico" y orden grande en $(\alpha_v,\dots,\alpha_v)$, debemos mostrar ahora que si las aproximaciones $\beta_1,\dots,\beta_m$ son lo suficientemente cercanas a los α_v, entonces el polinomio deberá de anularse con orden alto en $(\beta_1,\dots,\beta_m)$.

Ilustraremos este fenómeno con polinomios de una variable. Sea entonces $\beta = p/q$ un número racional, $P(X) \in \mathbb{Z}[X]$ un polinomio con coeficientes enteros, de los cuales tenemos alguna especie de control y P se anula con orden grande en un número α. Pidamos por ejemplo

$$grP(X) = r, \quad Ind(P;r;\alpha) = \tfrac{multi(P,\alpha)}{r} \geq \tfrac{1}{2}, \quad |P| \leq B(\alpha)^r,$$

y supongamos que p/q es una buena aproximación de α, es decir, $|p/q-\alpha| \leq q^{-(2+\varepsilon)}$. La expresión de Taylor de P alrededor de α será

$$P(X) = \textstyle\sum_{i=0}^{r}\partial_i P(\alpha)(X-\alpha)^i = \sum_{i\geq r/2}\partial_i P(\alpha)(X-\alpha)^i.$$

Evaluando en p/q obtenemos que

$$\left|P(\tfrac{p}{q})\right| \leq \textstyle\sum_{i\geq r/2}\partial_i P(\alpha)\left|\frac{p}{q}-\alpha\right|^i.$$

Recordando el lema 1 y usando que $|P| \leq B(\alpha)^r$ podremos estimar el tamaño de las derivadas en α. Como $\partial_i P$ es un polinomio de grado como mucho r y coeficientes acotados por $|\partial_i P|$ se tiene que

$$|\partial_i P(\alpha)| \leq r|\partial_i P|\max\{1,|\alpha|\}^r \leq 2^r.2^r.B(\alpha)^r\max\{1,|\alpha|\}^r = C(\alpha)^r$$

Por lo tanto, tendremos que

$$\left|P(\tfrac{p}{q})\right| \leq rC^r\left|\frac{p}{q}-\alpha\right|^{r/2} \leq (2C)^r\left(\tfrac{1}{q^{2+\varepsilon}}\right)^{r/2} = \left(\tfrac{2C}{q^{1+\varepsilon/2}}\right)^r.$$

Por otro lado, si $P(X)$ no se anula en p/q, $P(p/q)$ es una fracción no nula con denominador q^r y por lo tanto $P(p/q) \geq 1/q^r$. Entonces tendremos que

$$\frac{1}{q^r} \leq \left|P(\tfrac{p}{q})\right| \leq \left(\tfrac{2C}{q^{1+\varepsilon/2}}\right)^r,$$

lo que implica que $q \leq (2C)^{2/\varepsilon}$. Esto nos da una cota para q, si asumimos que p/q no es raíz de $P(X)$. Por lo tanto, si nuestras aproximaciones son de altura suficientemente grande $P(X)$ se anulara en ellas. Un argumento similar funcionara para las derivadas $\partial_j P$. Por lo tanto, con las hipótesis correctas impuestas sobres las alturas de las aproximaciones, podremos probar que $P(X)$ se anula con orden grande en p/q.

Por lo tanto, para generalizar este argumento y poder probar lo deseado necesitaremos tener cierto control sobre la altura de las derivadas del polinomio construido, evaluado en números algebraicos.

LEMA 7. *Sea $P \in \mathbb{Z}[X_1, \ldots, X_m]$ tal que $gr_{X_h}(P) \leq r_h$, y sea $\beta = (\beta_1, \ldots, \beta_m)$ una m-upla de números algebraicos en un cuerpo de números K. Entonces para toda m-upla $j = (j_1, \ldots, j_m)$ de enteros no negativos se cumple que*

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

DEMOSTRACIÓN. Sea $(j_1, \ldots, j_m)$ una m-upla de enteros no negativos. Sea $T(X_1, \ldots, X_m) = \partial_{j_1 \ldots j_m} P(X_1, \ldots, X_m)$. Por lema 1 sabemos que T posee coeficientes enteros acotados por

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

Consideremos $v \in M_K^\infty$ cualquier valor absoluto arquimediano. Usando la desigualdad triangular y el hecho de que T posee como mucho $(r_1+1)\ldots(r_m+1) \leq 2^{r_1+\ldots+r_m}$ coeficientes, cada uno de ellos acotados por $|T|$ se tiene que

$$|T(\beta_1, \ldots, \beta_m)|_v \leq (r_1+1)\ldots(r_m+1)\ |T| max\left\{|\beta_1|_v, 1\right\}^{r_1} \ldots\ max\left\{|\beta_m|_v, 1\right\}^{r_m} \leq 4^{r_1+\ldots+r_m}\ |P|\ max\left\{|\beta_1|_v, 1\right\}^{r_1}\ \ldots max\left\{|\beta_m|_v, 1\right\}^{r_m}.$$

Similarmente, si $v \in M_K^0$ es algun valor absoluto noarquimediano, usando la desigualdad triangular noarquimediana, el hecho de que T posee coeficientes enteros y recordando que $|n|_v \leq 1\ \forall n \in \mathbb{Z}$ se tiene que

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

Elevando estas desigualdades a la potencia n_v-esima($n_v = [K_v : \mathbb{Q}_v]$), multiplicandolas sobre todos los $v \in M_K$, se obtiene que

$$H_K(T(\beta_1, \ldots, \beta_m)) \leq \prod_{v \in M_K^\infty} 4^{(r_1+\ldots+r_m)n_v}\ |P|^{n_v}\ max\left\{|\beta_1|_v, 1\right\}^{r_1 n_v} \ldots\ max\left\{|\beta_m|_v, 1\right\}^{r_m n_v} \times$$

$$\prod_{v \in M_K^0} max\left\{|\beta_1|_v, 1\right\}^{r_1 n_v} \ldots\ max\left\{|\beta_m|_v, 1\right\}^{r_m n_v} = 4^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ H_K(P)\ H_K(\beta_1)^{r_1} \ldots\ H_K(\beta_m)^{r_m}$$

(Observar que como P posee coeficientes enteros, $H_K(P) = |P|^{[K:\mathbb{Q}]}$). □

LEMA 8. *Sean $r_1, \ldots, r_m$ enteros positivos y $P \in \mathbb{Z}[X_1, \ldots, X_m]$ un polinomio con coeficientes enteros tal que $gr_{X_h} \leq r_h$.Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$ y sea $\theta_v = Ind(P; r_1, \ldots, r_m; \alpha_v, \ldots, \alpha_v)$. Sea $\theta = \min_{v \in S} \{\theta_v\}$. Sea $0 < \delta < 1$ una constante y elijamos $0 < \theta_0 < \theta$.*

Sea

$$\xi : S \to [0,1] \text{ una función tal que } \sum_{\xi \in S} \xi_v = 1.$$

Supongamos que $\beta_1, \ldots, \beta_m \in K$ cumplen que

$$\|\beta_v - \alpha_v\|_v \leq \frac{1}{H_K(\beta_v)^{(2+\delta)\xi_v}} \text{ para todo } v \in S \text{ y para todo } 1 \leq h \leq m.$$

Definamos $D := min\{H_K(\beta_h)^{r_h}\}$, y sea $j = (j, \ldots, j_m)$ una m-upla de enteros no negativos tal que $\sum_{h=1}^m \frac{j_h}{r_h} \leq \theta_0$. Entonces

$$\prod_{v \in S} \|T(\beta_1, \ldots, \beta_m)\|_v \leq$$
$$\left(16 \prod_{v \in S} max\{|\alpha_v|_v, 1\}\right)^{(r_1+\ldots+r_m)[K:\mathbb{Q}]} H_K(P)\ D^{-(2+\delta)(\theta-\theta_0)}.$$

DEMOSTRACIÓN. Sea $v \in S$ y $j = (j_1, \ldots, j_m)$ como en el enunciado y $T = \partial_j P$. Estimaremos $\|\partial_j P(\beta_1, \ldots, \beta_m)\|_v$ mediante la expresion de Taylor de T alrededor de $(\alpha_v, \ldots, \alpha_v)$. Sea $i = (i_1, \ldots, i_m)$ una m-upla de enteros no negativos.

De la misma forma en la que se obtuvieron las cotas para $|T(\beta_1, \ldots, \beta_m)|_v$ en el lema anterior, se tiene que, si v es arquimediano entonces

$$|\partial_{i_1 \ldots i_m} T(\alpha_v, \ldots, \alpha_v)|_v \leq |T|(4max\{|\alpha_v|_v, 1\})^{r_1+\ldots+r_m} \leq$$
$$|P|(8max\{|\alpha_v|_v, 1\})^{r_1+\ldots+r_m},$$

mientras que si v es noarquimediano se tendrá que

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

Por Lema 2, podemos acotar inferiormente el índice de T respecto de $(r_1, \ldots, r_m)$ en $(\alpha_v, \ldots, \alpha_v)$, se tiene entonces que

$$Ind(T; r_1, \ldots, r_m; \alpha_v, \ldots, \alpha_v) = Ind\partial_{j_1 \ldots j_m} P \geq IndP - \sum_{h=1}^m \frac{j_h}{r_h} \geq \theta_v - \theta_0.$$

Por lo tanto, si escribimos la expresión de Taylor de T alrededor de $(\alpha_v, \ldots, \alpha_v)$, muchos de los términos serán ceros. Entonces

$$T(X_1, \ldots, X_m) = \sum{}^* \partial_{i_1 \ldots i_m} T(\alpha_v, \ldots, \alpha_v)(X_1 - \alpha_v)^{i_1} \ldots (X_m - \alpha_v)^{i_m},$$

$\sum^*$ quiere decir que se suma sobre las m-uplas $(i_1, \ldots, i_m)$ tal que $\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m} \geq \theta_v - \theta_0$. Sea $(i_1, \ldots, i_m)$ una de estas m-uplas entonces

$$|\beta_1 - \alpha_v|_v^{i_1} \ldots |\beta_m - \alpha_v|_v^{i_m} \leq \frac{1}{(H_K(\beta_1)^{i_1} \ldots H_K(\beta_m)^{i_m})^{(2+\delta)\xi_v/n_v}},$$

el denominador podrá ser acotado por

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

Por lo tanto, para toda m-uplas $(i_1, \ldots, i_m)$ tal que $\frac{i_1}{r_1} + \ldots + \frac{i_m}{r_m} \geq \theta_v - \theta_0$ vale que

$$|\beta_1 - \alpha_v|_v^{i_1} \dots |\beta_m - \alpha_v|_v^{i_m} \leq \frac{1}{D^{(2+\delta)(\theta_v - \theta_0)\xi_v/n_v}}.$$

Sea $v \in S$, arquimediano, si evaluamos la expresión de Taylor en $(\beta_1, \dots, \beta_m)$, usando que los β_h son próximos a α_v y la desigualdad triangular, obtenemos que

$$|T(\beta_1, \dots, \beta_m)|_v \leq \sum^{*} |\partial_{i_1 \dots i_m} T(\alpha_v, \dots, \alpha_v)|_v |\beta_1 - \alpha_v|_v^{i_1} \dots |\beta_m - \alpha_v|_v^{i_m} \leq$$
$$(r_1 + 1) \dots (r_m + 1). \max_{i_1 \dots i_m} \{|\partial_{i_1 \dots i_m} T(\alpha_v, \dots, \alpha_v)|_v\} . \max_{\frac{i_1}{r_1} + \dots + \frac{i_m}{r_m} \geq \theta_v - \theta_0} \left\{|\beta_1 - \alpha|_{v_v}^{i_1} \dots |\beta_m - \alpha_v|_v^{i_m}\right\} \leq$$
$$2^{r_1 + \dots + r_m} \max_{i_1 \dots i_m} \{|\partial_{i_1 \dots i_m} T(\alpha_v, \dots, \alpha_v)|_v\} \frac{1}{D^{(2+\delta)(\theta_v - \theta_0)\xi_v/n_v}} \leq$$
$$|P|(16 max\{|\alpha_v|_v, 1\})^{r_1 + \dots + r_m} \frac{1}{D^{(2+\delta)(\theta_v - \theta_0)\xi_v/n_v}} \leq$$
$$|P|(16 max\{|\alpha_v|_v, 1\})^{r_1 + \dots + r_m} \frac{1}{D^{(2+\delta)(\theta - \theta_0)\xi_v/n_v}}.$$

Mientras que si $v \in S$, es noarquimediano, por la desigualdad triangular noarquimediana se tendrá

$$|T(\beta_1, \dots, \beta_m)|_v \leq max\{|\alpha_v|_v, 1\}^{r_1 + \dots + r_m} \frac{1}{D^{(2+\delta)(\theta_v - \theta_0)\xi_v/n_v}} \leq$$
$$max\{|\alpha_v|_v, 1\}^{r_1 + \dots + r_m} \frac{1}{D^{(2+\delta)(\theta - \theta_0)\xi_v/n_v}}.$$

Elevando estas desigualdades a la potencia n_v-esima, multiplicando, recordando que $n_v \leq [K : \mathbb{Q}]$ y que $\sum \xi_v = 1$, conseguimos

$$\prod_{v \in S} \|T(\beta_1, \dots, \beta_m)\|_v \leq$$
$$\left(16 \prod_{v \in S} max\{|\alpha_v|_v, 1\}\right)^{(r_1 + \dots + r_m)[K:\mathbb{Q}]} H_K(P)\ D^{-(2+\delta)(\theta - \theta_0)}.$$

□

Veremos ahora que con las hipótesis adecuadas sobre las alturas de las aproximaciones consideradas, podremos concluir que el orden de anulación de nuestro polinomio sobre ellas será elevado. En efecto, tenemos la siguiente proposición.

PROPOSICIÓN 3. *Sea $0 < \delta < 1$ una constante dada, y sea ε tal que*

$$0 < \varepsilon < \frac{\delta}{22}. \tag{2.8}$$

Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$ de grado d_v, sea m un entero tal que $m > 16 \left(\sum_{t=1}^{s} d_j\right)^2 \varepsilon^{-2}$, sean $r_1, \dots, r_m$ enteros positivos dados y sea $P(X_1, \dots, X_m) \in \mathbb{Z}[X_1, \dots, X_m]$ el polinomio del Teorema 2.5.

Sea

$$\xi : S \to [0, 1] \text{ una función tal que } \sum_{\xi \in S} \xi_v = 1.$$

Supongamos que $\beta_1, \dots, \beta_m \in K$ cumplen que

$$\|\alpha_v - \beta_v\|_v \leq \frac{1}{H_K(\beta_v)^{(2+\delta)\xi_v}} \tag{2.9}$$

para todo $v \in S$ y para todo $1 \leq h \leq m$. Supongamos además que

$$\max_{1\leq h\leq m} \{H_K(\beta_h)^{r_h}\} \leq \min_{1\leq h\leq m} \{H_K(\beta_h)^{r_h}\}^{1+\varepsilon} \tag{2.10}$$

y que existe una constante $C = C(\{\alpha_v\}_{v\in S}, \delta, K)$ tal que

$$C \leq H_K(\beta_h) \tag{2.11}$$

para todo $1 \leq h \leq m$. Entonces el índice de P respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$ cumple

$$Ind\ P \geq \varepsilon m.$$

Demostración. Sea $j = (j_1, \ldots, j_m)$ una m-upla de enteros no negativos tal que $\sum_{h=1}^{m} j_h/r_h \leq \varepsilon m$, y sea $B = B\left(\{\alpha_v\}_{v\in S}\right)$ la cota para $|P|$ que proviene del Teorema 2.5 y notemos $M = \prod_{v\in S} max\{|\alpha_v|_v, 1\}$. Veremos que $\partial_j\ P(\beta_1, \ldots, \beta_m) = 0$. Aplicaremos el lema 8 sobre el polinomio P, tomando $\theta_0 = \varepsilon m$. Observar que $\theta \geq \frac{m}{2}(1-\varepsilon)$ y dado que $\varepsilon < 1/22$, se cumple que $\theta_0 < \theta$, por lo tanto estamos bajo las hipótesis del Lema 8 obteniendo así que

$$\prod_{v\in S} \|\partial_j\ P(\beta_1, \ldots, \beta_m)\|_v \leq \frac{(16M)^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ H_K(P)}{D^{(2+\delta)(\theta-\theta_0)}} \leq \frac{(16BM)^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}}{D^{(2+\delta)(\frac{m}{2}(1-\varepsilon)-\varepsilon m)}}.$$

Por otro lado, por lema 7 se tiene que

$$H_K(\partial_j P(\beta_1, \ldots, \beta_m)) \leq 4^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ H_K(P)\ \prod_{h=1}^{m} H_K(\beta_h)^{r_h} \leq (4B)^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ D^{m(1+\varepsilon)} \leq (16BM)^{(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ D^{m(1+\varepsilon)}.$$

Por la desigualdad de Liouville(Lema 6) se tiene que o bien $\partial_j\ P(\beta_1, \ldots, \beta_m) = 0$ o bien

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

Veamos que esto último contradice nuestras hipótesis. Asumiendo entonces que $\partial_j\ P(\beta_1, \ldots, \beta_m) \neq 0$, la desigualdad de Liouville implica que

$$D^{m((1+\delta/2)(1-3\varepsilon)-(1+\varepsilon))} \leq (16BM)^{2(r_1+\ldots+r_m)[K:\mathbb{Q}]}\ .$$

Como asumimos que $\delta < 1$ y $\varepsilon < \delta/22$, se tiene

$$(1+\tfrac{\delta}{2})(1-3\varepsilon) - (1+\varepsilon) = \tfrac{\delta}{2} - 4\varepsilon - \tfrac{3}{2}\varepsilon\delta > \tfrac{\delta}{2} - \tfrac{4}{22}\delta - \tfrac{3}{44}\delta = \tfrac{\delta}{4}$$

y entonces

$$\max_{1\leq h\leq m} \{H_K(\beta_h)^{r_h}\} \leq D^{1+\varepsilon} \leq (16BM)^{8(r_1+\ldots+r_m)[K:\mathbb{Q}](1+\varepsilon)/(\delta m)}.$$

Sea j tal que $r_j = \max_{h=1}^{m} r_h$, se deduce que

$$H_K(\beta_h) \leq (16BM)^{8[K:\mathbb{Q}](1+\varepsilon)/\delta}.$$

Eligiendo $C = C(\{\alpha_v\}_{v\in S}, \delta, K)$ suficientemente grande,por ejemplo $C = (16BM)^{8[K:\mathbb{Q}]((8/\delta)+4/11)}$, se consigue la contradicción deseada. □

6. Lema de Roth

Ya fue mencionado que ni bien consideramos más de una variable, probar la no anulación del polinomio P o alguna de sus derivadas en el punto de aproximación $(\beta_1, \ldots, \beta_m)$ resulta ser el paso más difícil de todos. En esta sección explicaremos el argumento propuesto por Roth para solucionar las dificultades del Paso III. En la sección anterior probamos que si P se anula con orden grande en $(\alpha_\upsilon, \ldots, \alpha_\upsilon)$ entonces también se anulara con orden alto en $(\beta, \ldots, \beta)$. Aquí mostraremos que de hecho, no es posible para P anularse con orden grande en $(\beta, \ldots, \beta)$.

Consideremos el caso de polinomios en una variable para ilustrar una de las dificultades que aparecen.

Sea $P(X) \in \mathbb{Z}[X]$ un polinomio de grado a lo sumo r,con coeficientes enteros acotados por $|P| \leq B^r$. Notemos por $I = Ind(P; r; p/q) = multi(P; p/q)/r$ el índice de P en algun número racional p/q. Tendremos entonces que $(X - p/q)^{rI}$ divide a P. Como P posee coeficientes enteros, el Lema de Gauss[1] asegura que $(qX - p)^{rI}$ divide a P. Dado que q^{rI} divide al coeficiente principal de P y p^{rI} divide al termino constantes de P, se tiene que

$$\max\{|p|, |q|\}^{rI} \leq |P| \leq B^r$$

por lo tanto

$$IndP = I \leq \tfrac{logB}{logH(p/q)}.$$

Esto muestra que $IndP$ será chico si $H(p/q)$ es relativamente grande comparado con B.

Cuando el polinomio considerado P posee más de una variable, este argumento de divisibilidad ya no funciona. Veamos a continuación algunas de las ideas utilizadas por Thue, quien considero polinomios de dos variables de la forma $P(X, Y) = f(X) + g(X)Y$. Notemos por I al índice de P respecto de $(r, 1)$ en (β_1, β_2). Tenemos entonces, por definición de índice que

$$\partial_{i,0} P(\beta_1, \beta_2) = \partial_i f(\beta_1) + \partial_i g(\beta_1)\beta_2 = 0 \text{ para todo } i/r \leq I.$$

Como ha sido previamente comentado, el determinante Wronskiano, que es un polinomio de una variable, aparece naturalmente cuando tratamos de probar que P no puede anularse con orden grande en (β_1, β_2). Sea entonces

$$W(X) = f(X)\partial_1 g(X) - g(X)\partial_1 f(X).$$

Diferenciando W k veces, se ve que vale

$$\partial_k W = \textstyle\sum_{i+j=k} (\partial_i f \partial_{j+1} g - \partial_{j+1} f \partial_i g).$$

Por lo mencionado más arriba tenemos que si $i \leq rI$ y $j + 1 \leq rI$, entonces

$$\partial_{i,0} P(\beta_1, \beta_2) = \partial_i f(\beta_1) + \partial_i g(\beta_1)\beta_2 = 0$$

y

$$\partial_{j+1,0} P(\beta_1, \beta_2) = \partial_{j+1} f(\beta_1) + \partial_{j+1} g(\beta_1)\beta_2 = 0.$$

Si eliminamos β_2 de estas ecuaciones obtenemos

$$\partial_i f(\beta_1)\partial_{j+1} g(\beta_1) - \partial_{j+1} f(\beta_1)\partial_i g(\beta_1) = 0 \text{ para todo } i \leq rI \text{ y } j \leq rI - 1.$$

Se sigue que $\partial_k W(\beta_1) = 0$ para todo $k \leq rI - 1$, lo que significa que el índice de W respecto de r en β_1 cumple

$$IndW \geq IndP - \tfrac{1}{r}.$$

[1]El lema de Gauss asegura que si un polinomio con coeficientes enteros se factoriza en $\mathbb{Q}[X]$, entonces se factoriza en $\mathbb{Z}[X]$.

Si estimamos el tamaño de $|W|$, mediante el mismo razonamiento utilizado en el caso de una variable, conseguiríamos una cota superior para $IndW$ y por lo tanto una cota superior para $IndP$. Al igual que el caso de una variable, si imponemos las hipótesis necesarias sobre las alturas de las aproximaciones, tendremos que $IndP$ no podrá ser muy grande.

Roth utiliza un argumento inductivo sobre la cantidad de variables y mediante el uso de Wronskianos generalizados[2] logra factorizar uno de estos Wronskianos de m variables en un producto de la forma $v(X_1, \ldots, X_{m-1})u(X_m)$ para poder aplicar la hipótesis inductiva sobre estos.

DEFINICIÓN 4. *Sea $\boldsymbol{i} = (i_1, \ldots, i_m)$, definimos el orden del operador diferencial*

$$\Delta_{\boldsymbol{i}} = \frac{\partial^{i_1+\ldots+i_m}}{\partial X_1^{i_1} \ldots \partial X_m^{i_m}}$$

como orden$(\Delta_{\boldsymbol{i}}) = i_1 + \ldots + i_m$.

DEFINICIÓN 5. *Sea k un cuerpo de característica cero y $\phi_1, \ldots, \phi_r \in k(X)$ funciones racionales. Un Wronskiano generalizado de $\phi_1, \ldots, \phi_r$ es un determinante de la forma*

$$det((\Delta_i \phi_j)_{1 \leq i,j \leq r}),$$

donde $\Delta_1, \ldots, \Delta_r$ son operadores diferenciales, con orden$(\Delta_i) \leq i - 1$

Si $m = 1$ y $\Delta_i = \frac{\partial^{i-1}}{\partial X^{i-1}}$ se recupera el Wronskiano clásico de las funciones racionales $\phi_1, \ldots, \phi_r \in k(X)$

$$W(\phi_1, \ldots, \phi_r) = \det\Big(\tfrac{\partial^{i-1}}{\partial X^{i-1}} \phi_j\Big).$$

Un Teorema clásico establece que las funciones $\phi_1, \ldots, \phi_r$ son linealmente independientes sobre k si y solo si $W(\phi_1, \ldots, \phi_r) \neq 0$. Veremos que un resultado de esta índole es válido en más variables.

LEMA 9. *Sea k un cuerpo de caracteristica cero y $\phi_1, \ldots, \phi_r \in k(X)$ funciones racionales. Si $\phi_1, \ldots, \phi_r$ son linealmente independientes sobre k entonces existe un Wronskiano generalizado de $\phi_1, \ldots, \phi_r$ no nulo.*

DEMOSTRACIÓN. La prueba será por inducción en r. Si $r = 1$, el único Wronskiano generalizado es $\Delta_1 \phi_1 = \phi_1$ ya que el unico operador diferencial de orden 0 es la identidad. Por lo tanto, en el caso $r = 1$, el lema dice que ϕ_1 es linealmente independiente sobre k si y solo si ϕ_1 es no nula.

Supongamos que el lema es verdadero para cualquier conjunto de $r - 1$ funciones racionales linealmente independiente sobre k y sean $\phi_1, \ldots, \phi_r$ k-linealmente independientes. Sea $\lambda \in k(X_1, \ldots, X_m)$ una función racional no nula, las funciones $\lambda\phi_1, \ldots, \lambda\phi_r$ siguen siendo linealmente independientes sobre k. Observar que cualquier Wronskiano generalizado $det(\Delta_i(\lambda\phi_j))$ de $\lambda\phi_1, \ldots, \lambda\phi_r$ es una $k(X_1, \ldots, X_m)$-combinacion lineal de Wronskianos generalizados de $\phi_1, \ldots, \phi_r$ (donde los coeficientes serán funciones racionales en las derivadas parciales de λ). Para probar el lema, será suficiente probar que existe un wronskiano generalizado no nulo de $\lambda\phi_1, \ldots, \lambda\phi_r$ para un λ apropiado. Si tomamos $\lambda = \phi_1^{-1}$, podremos suponer que $\phi_1 = 1$.

Como $r > 1$, $\phi_1 = 1$ y ϕ_2 son linealmente independiente, con lo cual ϕ_2 no es una constante. Entonces $\frac{\partial \phi_2}{\partial X_J} \neq 0$ para algun j. Sin pérdida de generalidad, podemos suponer que $\frac{\partial \phi_2}{\partial X_1} \neq 0$.

[2]Herramienta utilizada con anterioridad por Siegel.

Sea V el k-espacio vectorial generado por $\phi_1, \ldots, \phi_r$ y definamos el k-subespacio vectorial de V

$$W = \left\{ \phi \in V | \ \frac{\partial \phi}{\partial X_1} = 0 \right\}, \text{ y sea } \ t = dim W.$$

Observar que $\phi_1 \in W$ y $\phi_2 \notin W$, entonces $1 \leq t \leq r-1$. Sean $\psi_1, \ldots, \psi_r$ funciones racionales tal que $\psi_1, \ldots, \psi_t$ es base de W y $\psi_1, \ldots, \psi_r$ es base de V. Por hipótesis inductiva, existen operadores $\Delta_1^*, \ldots, \Delta_t^*$ operadores diferenciales tal que

$$\det(\Delta_i^* \psi_j) \neq 0 \quad \text{con} \ \ 1 \leq i, j \leq t \ \ \text{y} \ \ orden(\Delta_i^*) \leq i-1.$$

Afirmo que las funciones racionales $\frac{\partial \psi_{t+1}}{\partial X_1}, \ldots, \frac{\partial \psi_r}{\partial X_1}$ son k-linealmente independientes, pues si $c_{t+1}, \ldots, c_r \in k$ tal que

$$0 = c_{t+1} \frac{\partial \psi_{t+1}}{\partial X_1} + \ldots + c_r \frac{\partial \psi_r}{\partial X_1} = \frac{\partial}{\partial X_1}(c_{t+1}\psi_{t+1} + \ldots + c_r \psi),$$

entonces $c_{t+1} = \ldots = c_r = 0$ pues el subespacio generado por $\psi_{t+1}, \ldots, \psi_r$ posee intersección nula con W. Aplicando hipótesis inductiva nuevamente, obtenemos que existen operadores $\Delta_{t+1}^*, \ldots, \Delta_r^*$ operadores diferenciales tal que

$$\det(\Delta_i^* \frac{\partial \psi_j}{\partial X_1}) \neq 0 \quad \text{con} \ \ t+1 \leq i, j \leq r \ \ \text{y} \ \ orden(\Delta_i^*) \leq i-r-1.$$

Definimos operadores Δ_i para $1 \leq i \leq r$ como

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

Notar que $orden(\Delta_i) \leq i-1$ y ademas

$$\Delta_i \psi_j = \Delta_i^* \frac{\partial \psi_j}{\partial X_1} = \Delta_i^* 0 = 0 \quad \text{para} \quad 1 \leq j \leq t \ \ \text{y} \ \ t+1 \leq i \leq r.$$

Se tiene entonces que

$$det((\Delta_i \psi_j)_{1 \leq i,j \leq r}) = det \begin{pmatrix} \Delta_i^* \psi_j & \Delta_i^* \psi_j \\ \\ 0 & \Delta_i^* \frac{\partial \psi_j}{\partial X_1} \end{pmatrix} = det(\Delta_i^* \psi_j) det(\Delta_i^* \frac{\partial \psi_j}{\partial X_1})$$

Por lo tanto, existe un Wronskiano generalizado no nulo de $\psi_1, \ldots, \psi_t$.Pero como las funciones $\phi_1, \ldots, \phi_r$ y $\psi_1, \ldots, \psi_t$ son base del mismo k-espacio vectorial, se tiene que $\psi_j = \sum_l a_{jl} \phi_l$ con (a_{jl}) matriz sobre k, inversible. Se tiene entonces

$$0 \neq \det(\Delta_i \psi_j) = \det(\Delta_i(\textstyle\sum_l a_{jl} \phi_l)) = \det(\sum_l a_{jl} \Delta_i(\phi_l)) = \det(a_{jl}) \ \det(\Delta_i \phi_l) \neq 0,$$

y consecuentemente $\det(\Delta_i \phi_l) \neq 0$, lo que concluye la prueba del lema. □

Antes de empezar con la demostración del lema de Roth, recordemos la definición de altura de un polinomio que usaremos en el lema. Sea $f = \sum_{i \in I} a_i x^i$ un polinomio en varias variables sobre un cuerpo de números K, sea $n_v = [K_v : \mathbb{Q}_v]$ y $|f|_v = \max_{i \in I} |a_i|_v$, entonces

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

LEMA 10. *Sea m un entero positivo y $P \in \bar{\mathbb{Q}}[X_1, \ldots, X_m]$ un polinomio con coeficientes algebraicos y $gr_{X_h} \leq r_h$. Sea $\beta = (\beta_1, \ldots, \beta_m)$ una m-upla de números algebraicos. Sea $\eta > 0$ un número real fijo tal que*

$$\frac{r_{h+1}}{r_h} \leq \eta^{2^m - 1} \text{ para todo } 1 \leq h \leq m-1, \tag{2.12}$$

y

$$\eta^{2^m-1} \min_{1 \leq h \leq m} \{r_h log\ H(\beta_h)\} \geq logH(P) + 4mr_1. \tag{2.13}$$

Entonces el índice de P respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$ cumple

$$Ind\ P \leq 2m\eta.$$

Dado que siempre se tiene que $IndP \leq m$, podremos suponer que $\eta < 1/2$, pues de otra forma el resultado se cumplirá trivialmente.

La prueba será por inducción en m, el número de variables. Para simplificar la notación, llamaremos K a un cuerpo de números que contenga a todos los β_h y los coeficientes de P. Y sea $d = [K, \mathbb{Q}]$.

Probemos el caso $m = 1$. Para simplificar notación, escribamos $\beta = \beta_1$ y $r = r_1$. Sea l el orden de anulación de $P(X)$ en β, entonces $P(X) = (X - \beta)^l Q(X)$, con $Q(\beta) \neq 0$. Además $Ind\ P$ respecto de r en β es igual a l/r. Usando la desigualdad de Gelfand, tenemos que

$$H(\beta)^{r Ind\ P} = H(\beta)^l = H(X-\beta)^l \leq H(X-\beta)^l H(Q) \leq H(P)2^r,$$

lo que implica

$$Ind\ P \leq \frac{log\ H(P) + rlog2}{rlog\ H(\beta)} \leq \eta \quad \text{usando hipótesis (2.13).}$$

Lo que demuestra el caso $m = 1$.

OBSERVACIÓN 3. *La cota aquí obtenida es mejor que la enunciada. Conseguimos una cota igual a η, en lugar de 2η, y solo usamos que $\eta r log\ H(\beta) \geq log\ H(P) + rlog2$ en lugar de $\geq log\ H(P) + 4r$. Esta observación será usada en el paso inductivo con $m = 1$.*

Supongamos ahora, que el lema de Roth es cierto para polinomios de grado menor a m y lo probaremos para $P(X_1, \ldots, X_m)$ un polinomio en m variables.

Sea

$$P(X_1, \ldots, X_m) = \sum_{j_1=0}^{r_1} \sum_{j_2=0}^{r_2} \cdots \sum_{j_m=0}^{r_m} C_{j_1 \ldots j_m} X_1^{j_1} \ldots X_m^{j_m}$$

con $C_{j_1 \ldots j_m} \in \bar{\mathbb{Q}}$. Dado que

$$P(X_1, \ldots, X_m) = \sum_{j_m=0}^{r_m} \left(\sum_{j_1=0}^{r_1} \cdots \sum_{j_{m-1}=0}^{r_{m-1}} C'_{j_1, \ldots, j_{m-1}} X_1^{j_1} \ldots X^{jm-1} \right) X_m^{j_m},$$

con $C'_{j_1 \ldots j_{m-1}} \in \bar{\mathbb{Q}}$, siempre será posible escribir al polinomio como una suma del estilo

$$P(X_1, \ldots, X_m) = \sum_{j=1}^{k} \varphi_j(X_1, \ldots, X_{m-1}) \psi_j(X_m),$$

donde $\varphi_1, \ldots, \varphi_k$ y $\psi_1, \ldots, \psi_k$ son polinomios con coeficientes en $\bar{\mathbb{Q}}$. Desde ahora, tomemos una descomposición con k mínimo. En particular, $k \leq r_m + 1$.

La minimalidad de k, implica que tanto $\varphi_1, \ldots, \varphi_k$ como $\psi_1, \ldots, \psi_k$, son $\bar{\mathbb{Q}}$-linealmente independientes. Pues de no serlo, existirán $c_1, \ldots, c_k \in \bar{\mathbb{Q}}$, no todos

nulos, tal que $c_1\varphi_1 + \cdots + c_k\varphi_k = 0$. Supongamos sin pérdida de generalidad que $c_k \neq 0$, entonces

$$P(X_1, \ldots, X_m) = \sum_{j=1}^{k-1} \varphi_j . \left(\psi_j - \frac{c_j}{c_k}\psi_k\right),$$

lo que contradice la minimalidad de k. Similarmente, se prueba que $\psi_1, \ldots, \psi_k$ son linealmente independientes sobre $\bar{\mathbb{Q}}$.

Por Lema 9, se tiene que existe un determinante Wronskiano generalizado no nulo, digamos

$$U(X_m) = det\left(\tfrac{1}{(i-1)!}\tfrac{\partial^{i-1}}{\partial X_m^{i-1}}\psi_j(X_m)\right) \neq 0 \quad \text{con} \quad 1 \leq i, j \leq k,$$

y también existen operadores

$$\Delta_i' = \tfrac{1}{i_1!\ldots i_{m-1}!}\tfrac{\partial^{i_1+\ldots+i_{m-1}}}{\partial X_1^{i_1}\ldots\partial X_{m-1}^{i_{m-1}}} \quad \text{para} \quad 1 \leq i \leq k,$$

con $orden(\Delta_i') = i_1 + \ldots + i_{m-1} \leq i - 1 \leq k - 1 \leq r_m$. Tal que

$$V(X_1, \ldots, X_{m-1}) := det(\Delta_i'\varphi_j) \neq 0 \quad \text{con} \quad 1 \leq i, j \leq k.$$

Definimos al polinomio $W(X_1, \ldots, X_m)$ por

$$W(X_1, \ldots, X_m) := det\left(\tfrac{1}{(j-1)!}\tfrac{\partial^{j-1}}{\partial X_m^{j-1}}\Delta_i' P\right) =$$

$$det\left(\sum_{r=1}^{k}(\Delta_i'\varphi_r)\left(\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial X_m^{j-1}}\psi_r\right)\right) = V(X_1, \ldots, X_{m-1})U(X_m) \neq 0.$$

Por lo tanto, el uso de los determinantes Wronskianos nos permite obtener un polinomio W fuertemente relacionado con P y factorizarlo como el producto de polinomios en menos variables. Por hipótesis inductiva, lograremos obtener cotas superiores para los índices de U y V, lo que derivara en una cota para W. Dado que W es un polinomio en P y sus derivadas, podremos conseguir una cota inferior para el índice de W en función del índice de P.

Observar que $W \in K[X_1, \ldots, X_m]$ y dado que U y V no comparten variables, a partir de la definición de altura se tendrá que $h(U) + h(V) = h(W)$.

Para poder aplicar el lema de Roth, necesitaremos cotas para los grados y las alturas de U y V.

Afirmación 2. *Las siguientes estimaciones son validas*

(1) $gr_{X_m}(U) \leq kr_m$ *y* $gr_{X_j}(V) \leq kr_j$ *para todo* $1 \leq j \leq m-1$.
(2) $h(U) + h(V) = h(W) \leq k(h(P) + 4r_1)$

Demostración. (1) En efecto, los determinantes que definen a U y V son de tamaño k y cada entrada posee grado a lo sumo r_j respecto de X_j.

(2) Desarrollando el determinante que define a W por columnas tenemos que

$$W = \sum_{\pi}(-1)^{sg(\pi)}\prod_{i=1}^{k}\Delta_i'\partial_{\pi(i)}P.$$

Por definición de altura,

$$h(W) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in M_K} n_v log\left(\left|\sum_{\pi}(-1)^{sg(\pi)}\prod_{i=1}^{k}\Delta_i'\partial_{\pi(i)}P\right|_v\right).$$

Si υ es arquimediano se tiene que $|f_1 + \ldots + f_r|_\upsilon \leq r \max |f_j|_\upsilon$, mientras que si υ es no-arquimediano $|f_1 + \ldots + f_r|_\upsilon \leq \max |f_j|_\upsilon$. Aplicando estas propiedades a la suma sobre las permutaciones obtenemos que

$$h(W) \leq \frac{1}{[K:\mathbb{Q}]} \sum_{\upsilon \in M_K} n_\upsilon \max_\pi \log \left| \prod_{i=1}^{k} \Delta_i' \partial_{\pi(i)} P \right|_\upsilon + log(k!).$$

El lema de Gauss establece que si υ es no-arquimediano entonces $|f.g|_\upsilon = |f|_\upsilon |g|_\upsilon$, mientras que si υ es arquimediano por la desigualdad de Gelfand, $|f_1 \ldots f_r|_\upsilon \leq 2^d |f_1|_\upsilon \ldots |f_r|_\upsilon$ con d la suma de los grados parciales de $f_1 \ldots f_r$. Aplicándolos a nuestra desigualdad obtenemos que

$$h(W) = \frac{1}{[K:\mathbb{Q}]} \sum_{\upsilon \in M_K} n_\upsilon \max_\pi \sum_{i=1}^{k} \log \left| \Delta_i' \partial_{\pi(i)} P \right|_\upsilon + k(r_1 + \ldots + r_m) \log 2 + log(k!).$$

Si υ es no-arquimediano entonces $\left| \Delta_i' \partial_{\pi(i)} P \right|_\upsilon \leq |P|_\upsilon$ mientras que en el caso arquimediano $\left| \Delta_i' \partial_{\pi(i)} P \right|_\upsilon \leq 2^{r_1 + \ldots + r_m} |P|_\upsilon$. Esto implica que

$$h(W) \leq \sum_{i=1}^{k} (h(P) + (r_1 + \ldots + r_m) \log 2) + k(r_1 + \ldots + r_m) \log 2 + log(k!) = \\ kh(P) + k(r_1 + \ldots + r_m) 2 \log 2 + log(k!).$$

Por hipótesis,

$$r_1 + \ldots + r_m \leq r_1 (1 + \eta' + \ldots + \eta'^{m-1}) \quad \text{con} \quad \eta' = \eta^{2^{m-1}}.$$

Dado que $\eta \leq 1/2$ y que $m \geq 2$, se tiene que $\eta' \leq 1/4$ y $r_1 + \ldots + r_m \leq 4/3\ r_1$. Por otro lado,

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

y por lo tanto,

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

Observar que la constante 4 puede ser reemplazada por la constante más chica $\frac{4}{3}\ 2\ log 2 + \frac{1}{2} \approx 2,35$. Esta observación será utilizada más adelante. □

Ahora usaremos inducción para acotar el índice de U,V y W.

Afirmación 3. *Si el lema de Roth es verdadero para menos de m variables, entonces*

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

y por lo tanto el índice de W respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$ cumple que

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

DEMOSTRACIÓN. Queremos aplicar el lema de Roth para V, que es un polinomio en $m' = m-1$ variables, con $r'_j = kr_j$ y $\eta' = \eta^2$. Verifiquemos que se cumplen las hipótesis del lema. Se tiene que $gr_{X_j}(V) \leq r'_j$ por Afirmación 2. Condición 2.12 se cumple pues

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

Veamos que vale la condición (2.13),

$$r'_j h(\beta_j) = kr_j h(\beta_j) \geq k\eta^{-2^{m-1}}(h(P) + 4mr_1) = k\eta'^{-2^{m'-1}}(h(P) + 4mr_1).$$

Dado que $h(V) \leq h(W) \leq k(h(P) + 4r_1)$ obtenemos que $k(h(P) + 4r_1) \geq h(V) + 4m'kr_1$. Por lo tanto, se verifican todas las hipótesis del lema. Entonces la hipótesis inductiva asegura que

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

Para aplicar el lema de Roth al polinomio de una variable a U, con $\eta'' = \eta^{2^{m-1}}$ y $r'' = kr_m$. Tenemos que $gr_{X_m}(U) \leq r'_m$, por Afirmación 2. Como $m = 1$, la primer condición 2.12 es vacia, con lo cual solo hay que verificar (2.13). Para ver esto, usaremos la versión mejorada del lema de Roth en una variable, Comentario 3. Entonces

$$h(U) + r''\ log2 \leq k(h(P) + c_1 r_1) + kr_m\ log2 \leq k(h(P) + c_1 r_1) + k\tfrac{1}{2} r_1\ log2 \leq k(h(P) + 4r_1) \leq k\eta^{2^{m-1}} r_m h(\beta_m) = \eta'' r'' h(\beta_m).$$

(donde $c_1 = 4/3\ 2log2 + 1/2 \approx 2,35$ y $c_1 + 1/2\ log2 < 4$). Como las hipotesis se cumplen, aplicamos lema de Roth para una variable y obtenemos

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

□

El siguiente paso será relacionar el índice de W con el índice de P. A partir de la definición de W, se sigue que si P se anula con orden alto en $(\beta_1, \ldots, \beta_m)$, lo mismo pasara con los coeficientes de la matriz que define a W y entonces W también poseerá orden alto. A continuación, cuantificaremos esta observación.

AFIRMACIÓN 4. *Se tiene que*

$$IndW \geq \tfrac{k}{2} \min\left\{IndP, (IndP)^2\right\} - k\tfrac{r_m}{r_{m-1}}.$$

DEMOSTRACIÓN. Estimaremos primero el índice de los coeficientes de la matriz que define W respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$. Recordar que $\text{orden}(\Delta'_i) = i_1 + \ldots + i_{m-1} \leq i - 1 \leq k - 1 \leq r_m$ y $r_1 \geq r_2 \geq \ldots$; aplicando Lema 2 tendremos que

$$Ind\left(\Delta'_i\left(\tfrac{1}{(j-1)!}\tfrac{\partial^{j-1}}{\partial X_m^{j-1}} P\right)\right) = Ind\ \partial_{i_1,\ldots,i_{m-1},j-1} P \geq IndP - \tfrac{i_1}{r_1} - \cdots - \tfrac{i_{m-1}}{r_{m-1}} - \tfrac{j-1}{r_m} \geq IndP - \tfrac{i_1+\ldots+i_{m-1}}{r_{m-1}} - \tfrac{j-1}{r_m} \geq Ind(P) - \tfrac{r_m}{r_{m-1}} - \tfrac{j-1}{r_m}.$$

Desarrollando el determinante que define a W por columnas se tiene que

$$W = \sum_{\pi} (-1)^{sg(\pi)} \prod_{i=1}^{k} \Delta'_i \partial_{\pi(i)} P.$$

Observemos que W es suma de productos de k elementos, uno de cada columna, los cuales son de la forma $\partial_{i_1\ldots,i_{m-1},j-1}P$. Por Lema 2 se tiene que el indice de W es mayor o igual al mínimo de los índices de cada uno de los sumandos, es decir,

$$IndW \geq \min_{\pi}\left\{Ind\left(\prod_{i=1}^{k}\Delta_i'\partial_{\pi(i)}P\right)\right\}.$$

Aplicando nuevamente Lema 2, se tiene que el índice de un producto es igual a la suma de los índices de cada factor, y dado que cada sumando del desarrollo del determinante tiene un representante de cada columna se obtiene así que

$$IndW \geq \sum_{j=1}^{k}\min_{i_1,\ldots,i_{m-1}} Ind\ \partial_{i_1,\ldots,i_{m-1},j-1}P.$$

Sustituyendo en esta desigualdad la cota obtenida más arriba para $Ind\,\partial_{i_1\ldots,i_{m-1},j-1}P$, en los casos en los que sea positivo, obtenemos que

$$IndW \geq \sum_{j=1}^{k}\max\left\{Ind(P)-\tfrac{r_m}{r_{m-1}}-\tfrac{j-1}{r_m},0\right\}\geq \sum_{j=1}^{k}\max\left\{Ind(P)-\tfrac{j-1}{r_m},0\right\}-\tfrac{kr_m}{r_{m-1}}.$$

Por lo tanto, para probar nuestra afirmación, será suficiente ver que

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

Separaremos en dos casos.

Caso 3. $IndP \geq \frac{k-1}{r_m}$

En este caso tendremos que

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

Caso 4. $IndP \leq \frac{k-1}{r_m}$

Sea $N=[r_mIndP]$, con lo cual nuestra suposición implica que $N\leq k-1$. Entonces nuestra inecuación en cuestión resulta

$$\sum_{j=1}^{N+1}\left(Ind(P)-\frac{j-1}{r_m}\right)=(N+1).IndP-\frac{N(N+1)}{2r_m}=$$
$$(N+1).\left(IndP-\frac{[r_mIndP]}{2r_m}\right)\geq(N+1).\frac{1}{2}IndP\geq r_mIndP.\frac{1}{2}IndP\geq\frac{k}{2}(IndP)^2$$

si tenemos que $k\leq r_m$.

Si llegara a pasar que $k=r_m+1$, la cantidad que queremos acotar será

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

Observar que $q(N)$ es cuadrática en N. Por definición de N, es este caso tendremos

$$(k-1)IndP-1\leq N\leq(k-1)IndP.$$

Calculando el valor de q en los extremos de la desigualdad se puede ver a partir de un calculo directo que

$$q((k-1)IndP-1) = q((k-1)IndP) = \tfrac{(k-1)(IndP)^2+IndP}{2}.$$

Por lo tanto, como q como función cuadratica posee coeficiente principal negativo,

$$q(N) \geq \tfrac{(k-1)(IndP)^2+IndP}{2} \geq \tfrac{k(IndP)^2}{2},$$

donde la última desigualdad se debe a que bajo nuestras suposiciones, $IndP \leq 1$. Se completa así la afirmación hecha. □

Estamos ahora en condiciones de terminar la prueba del lema de Roth. Dado que $IndP \leq m$ por definición de índice, Afirmación 4 implica que

$$IndW + \tfrac{kr_m}{r_{m-1}} \geq \tfrac{k}{2} \min\left\{IndP, (IndP)^2\right\} \geq \tfrac{k(IndP)^2}{2m},$$

mientras que por Afirmación 3 se tiene que

$$IndW + \tfrac{kr_m}{r_{m-1}} \leq 2k(m-1)\eta^2 + k\eta^{2^{m-1}} + \tfrac{kr_m}{r_{m-1}} \leq k(2(m-1)\eta^2 + 2\eta^{2^{m-1}}).$$

Por lo tanto, comparando estas desigualdades obtenemos que $(IndP)^2 \leq 4\eta^2 m^2$, lo que implica $IndP \leq 2m\eta$.

7. Prueba del teorema de Roth

Probaremos a continuación el teorema de Roth. La versión que demostraremos será la obtenida luego de aplicar las reducciones probadas al principio del capitulo. Observar que se ha cambiado δ en lugar de ε y que se reemplazara la condición de aproximación $\min\{\|\alpha_v - \beta\|_v, 1\}$ por la condición más fuerte $\|\alpha_v - \beta\|_v$.

TEOREMA 2.6. *Sea K un cuerpo de números, $S \subset M_K$ un conjunto finito de valores absolutos sobre K y asumamos que cada valor absoluto en S ha sido extendido de alguna forma a $\bar{K}$. Para cada $v \in S$, sea $\alpha_v \in \bar{K}$ un entero algebraico. Sea $\delta > 0$ y sea*

$$\xi : S \to [0,1] \text{ una función que cumple } \sum_{v\in S} \xi_v = 1.$$

Entonces existen finitos $\beta \in K$ que satisfacen

$$\|\alpha_v - \beta\|_v \leq \frac{1}{H_K(\beta)^{(2+\delta)\xi_v}} \tag{2.14}$$

para todo $v \in S$.

DEMOSTRACIÓN. Para probar el teorema, supondremos que hay infinitas soluciones y llegaremos a una contradicción. Probarlo para δ más chico que alguna constante solo hará el resultado más fuerte, asumiremos entonces $0 < \delta < 1$. Dado que necesitaremos hacer referencia a las condiciones del Teorema 2.5, Proposición 3 y Lema 10, haremos una lista de dichas condiciones para que la lectura de la demostración sea más sencilla. La constante $B = B(\alpha_v; v \in S)$ esta definida en Teorema 2.5 y la constante $C = C(\{\alpha_v\}_{v\in S}, \delta, K)$ es la definida en Proposición 3:

Teorema 2.5 (2.7) $m > 16\left(\sum_{t=1}^{s} d_v\right)^2 \varepsilon^{-2}$.
Teorema 2.5-(3) $|P| \leq B^{r_1+\ldots+r_m}$.
Proposicion 3 (2.8) $0 < \varepsilon < \frac{\delta}{22}$.
Proposicion 3 (2.9) $\|\alpha_v - \beta_v\|_v \leq \frac{1}{H_K(\beta_v)^{(2+\delta)\xi_v}}$.

Proposicion 3 (2.10) $D := \min_{1\leq h\leq m}\{H_K(\beta_h)^{r_h}\} \leq \max_{1\leq h\leq m}\{H_K(\beta_h)^{r_h}\} \leq D^{1+\varepsilon}$.
Proposicion 3 (2.11) $C \leq H_K(\beta_h)$.
Lema 10 (2.12) $r_{h+1} \leq \omega r_h$.
Lema 10 (2.13) $logH(P) + 4mr_1 \leq \omega\ log\ D$.

Elegiremos los parámetros

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

en el siguiente orden

(1) Tomemos ε tal que $0 < \varepsilon < \delta/22$. En particular ε cumple (2.8); notar ademas que $\varepsilon < 1/22 < 1$.
(2) Elijamos m un entero tal que $m > 16\left(\sum_{t=1}^{s} d_v\right)^2 \varepsilon^{-2}$ donde d_v es el grade de α_v. Entonces (2.7) se cumple. Definimos $\omega = \omega(m,\varepsilon) = (\varepsilon/4)^{2^{m-1}}$. Lo que implica que $2\omega^{2^{-m+1}} = \varepsilon/2 < \varepsilon$.
(3) Como estamos asumiendo que (2.14) posee infinitas soluciones en K, dado que K posee finitos elementos de altura acotada, podemos elegir $\beta_1 \in K$ cuya altura cumple
$$H(\beta_1) \geq C \text{ y } log\ H(\beta_1) \geq \tfrac{m(logB+4)}{\omega}.$$
(4) Luego elegimos sucesivamente $\beta_2, \ldots, \beta_m$ soluciones de (2.14) que satisfagan
$$H_K(\beta_{h+1})^{\omega} \geq H_K(\beta_h)^2 \text{ para todo } 1 \leq h < m.$$
En particular, como $\omega < 1$, se tendrá que $H_K(\beta_h) \geq H_K(\beta_1)$. Es decir, la sucesión $H_K(\beta_h)$ será creciente, y (2.11) se cumplirá por la elección hecha en (3).
(5) Sea r_1 un entero que cumpla $H_K(\beta_1)^{\omega r_1} \geq H_K(\beta_m)^2$.
(6) Queremos elegir enteros $r_2, \ldots, r_m$ de forma tal que todos los $H_K(\beta_h)^{r_h}$ sean aproximadamente iguales. Definimos $r_2, \ldots, r_m$ como los enteros
$$r_h = \left\lceil \tfrac{r_1 logH_K(\beta_1)}{logH_K(\beta_h)} \right\rceil = \left\lceil \tfrac{r_1 logH(\beta_1)}{logH(\beta_h)} \right\rceil.$$
Para verificar la condición (2.10), calculamos

$$\begin{aligned}
&r_1 logH(\beta_1)\\
&\leq r_h logH(\beta_h) \ \text{ por definicion de } r_h \text{ y } \lceil t\rceil \geq t\\
&\leq r_1 logH(\beta_1) + logH(\beta_h) \ \text{ definicion de } r_h \text{ y } \lceil t\rceil \leq t+1\\
&\leq r_1 logH(\beta_1) + logH(\beta_m) \ \text{ puse } H_K(\beta_h) \text{ es creciente a partir de (4)}\\
&\leq (1+\varepsilon) r_1 logH(\beta_1) \ \text{ por la eleccion de } r_1 \text{ en (5)}.
\end{aligned}$$

Exponenciando obtenemos (2.10). Verificamos ahora la condición (2.12)

$$\begin{gathered}
\frac{r_{h+1}}{r_h} = \left\lceil \frac{r_1 logH(\beta_1)}{logH(\beta_{h+1})} \right\rceil \Big/ \left\lceil \frac{r_1 logH(\beta_1)}{logH(\beta_h)} \right\rceil \leq \\
\left(\frac{r_1 logH(\beta_1)}{logH(\beta_{h+1})} + 1\right) \Big/ \ \frac{r_1 logH(\beta_1)}{logH(\beta_h)} = \frac{logH(\beta_h)}{logH(\beta_{h+1})} + \frac{logH(\beta_1)}{r_1 logH(\beta_1)} \leq \\
\frac{\omega}{2} + \frac{\omega}{2} = \omega.
\end{gathered}$$

(7) Dado que m fue elegido para cumplir (2.7), aplicamos Teorema 2.5 para producir un polinomio $P(X_1, \ldots, X_m)$ con coeficientes enteros tal que $gr_{X_h} P \leq r_h$ y cumpla $|P| \leq B^{r_1+\ldots+r_m}$.

(8) Ya hemos verificado que las condiciones (2.8),(2.9), (2.10) y (2.11) se cumplen, entonces aplicando Proposicion 3 obtenemos que el índice de P respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$ cumple
$$IndP \geq m\varepsilon.$$

(9) Para poder aplicar el lema de Roth, nos falta verificar (2.13). En efecto, como
$$logD = \min_{1\leq h\leq m} \{r_h logH(\beta_h)\} = r_1 logH(\beta_1)$$
se tiene que
$$\frac{log|P|+4mr_1}{logD} \leq \frac{(r_1+\ldots+r_m)logB(\alpha)+4mr_1}{logD} \leq \frac{m(logB+4)}{logH(\beta_1)} \leq \omega.$$
Esto completa la verificación de las hipótesis del lema de Roth tomando $\eta = \omega^{2^{-m+1}} = \varepsilon/4$, entonces se concluye que el índice de P respecto de $(r_1, \ldots, r_m)$ en $(\beta_1, \ldots, \beta_m)$ satisface
$$IndP \leq 2m\eta = m\varepsilon/2.$$

Obtenemos así una cota superior e inferior para el índice de P que se contradicen. Así se completa la demostración del teorema de Roth.

□

CAPÍTULO 3

Aplicaciones

1. Sobre el desarrollo decimal de números algebraicos

A continuación veremos un ejemplo sencillo sobre cómo utilizar el Teorema de Roth para estudiar el desarrollo decimal de números algebraicos, lo que derivara en un criterio de trascendencia. El Teorema de Liouville visto en el Capitulo 1, implica que si $\{a_n\}$ es una sucesión creciente de enteros positivos que cumple $liminf \frac{a_{n+1}}{a_n} = \infty$, entonces el número

$$\sum_n \frac{1}{10^{a_n}}$$

resulta trascendente. Veamos que podemos reemplazar la condición $= \infty$ por > 1.

Sea S' un conjunto finitos de primos. Diremos que un número racional $\beta = a/b$ con $(a, b) = 1$, es un S'-entero, si el denominador es divisible solamente por primos de S'. Recordando la fórmula del producto

$$H(\beta) = \prod_p max\left\{1, |\beta|_p\right\} = \left(\prod_p min\left\{1, |\beta|_p\right\}\right)^{-1},$$

y usando que $H(\beta) = H(1/\beta)$ se tiene que en particular si β es un S'-entero, entonces

$$H(\beta) = \left(\prod_{p\in S'} min\left\{1, 1/|\beta|_p\right\}\right)^{-1}.$$

Sea α un número algebraico. Si aplicamos el teorema de Roth (segun el comentario (4) hecho en la sección 1 del capitulo 2) en el caso $K = \mathbb{Q}$, $S = S' \cup \infty$, $\alpha_\infty = \alpha$, $\alpha_p = \infty$ para todo $p \in S'$, se obtiene que dado $\varepsilon > 0$, existen finitos $\beta \in \mathbb{Q}$ para los cuales

$$min\{1, |\alpha - \beta|\} \prod_{p\in S'} min\left\{1, 1/|\beta|_p\right\} < \frac{1}{H(\beta)^{2+\varepsilon}}.$$

Usando el comentario anterior, se concluye que la desigualdad

$$|\alpha - \beta| < \frac{1}{H(\beta)^{1+\varepsilon}} \tag{3.1}$$

posee finitas soluciones en S'-enteros.

Una consecuencia de este hecho es que la expresión decimal de un número algebraico no puede tener bloques de ceros "muy grandes". Más precisamente, sea $0.a_1a_2\ldots$ la expresión decimal de un número algebraico, y para cada n definamos $l(n)$ como el mínimo $l \geq 0$ tal que $a_{n+l} \neq 0$; entonces $l(n) = o(n)$ con $n \to \infty$. Pues consideremos $\beta = 0, a_1 \ldots a_n$ con $a_n \neq 0$, entonces

$$\alpha - \beta = \sum_{i \geq n+l(n)} \frac{a_i}{10^i} \leq \frac{9}{10^{n+l(n)}} \sum_{i \geq 0} \frac{1}{10^i} = \frac{1}{10^{n+l(n)}}.$$

Por lo tanto, para no entrar en contradicción con (3.1), para todo $\varepsilon > 0$, debe valer que $l(n) < \varepsilon n$, para n suficientemente grande.

En particular, si $\{a_n\}$ es una sucesión creciente de enteros positivos que no cumpla $a_{n+1} - a_n = o(n)$, por ejemplo si $liminf\ \frac{a_{n+1}}{a_n} > 1$, el número

$$\sum_n \frac{1}{10^{a_n}}$$

resulta trascendente. Por ejemplo, podemos tomar $a_n = 2^n$. Esto es solo un caso sencillo de un resultado obtenido por Ferenczi y Mauduit [**13**], que esencialmente establece que si la expresión en base $b \geq 2$ de un número contiene infinitas potencias $(2+\varepsilon)$ de bloques (esto es un bloque seguido de sigo mismo y luego por su comienzo, de tamaño relativo al menos ε), a distancia del origen no mucho más grande que la longitud del bloque considerado, el número resulta trascendente. Veremos en el capitulo siguiente un resultado más fuerte que este que se obtiene a partir de Teorema del Subespacio, una generalización multidimensional del Teorema de Roth.

2. El problema de Waring

En 1770, Lagrange probó un problema de Diofanto, el cual asegura que todo entero positivo puede ser escrito como suma de 4 cuadrados. El mismo año, Waring afirmo, sin demostración, que todo entero positivo se podía escribir como suma de 9 cubos, 19 potencias cuartas y en general como s potencias k-esimas para algún s. Se define el número $g(k)$ como el menor entero tal que todo entero positivo se puede escribir como la suma de como mucho $g(k)$ potencias k-esimas positivas. En 1909 Hilbert probó que en general $g(k) < \infty$. Como fue observado por J.A.Euler(hijo de Leonhard Euler), el número $\left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor 2^k - 1$ requiere $\left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor - 1$ potencias 2^k y $2^k - 1$ potencias 1^k para su representacion, probando asi que $g(k) \geq 2^k + \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor - 2$.

Luego del tabajo de Hardy, Littlewood, y Vinogradov en el problema de Waring, se probó que para algun entero $c(k)$, que podía ser calculado en función de k, todo entero podría ser escrito como suma de menos de $g(k)$ potencias k-esimas. Por lo tanto, ąue realizar finitos cálculos para probar que todo entero menor a $c(k)$ podría ser escrito con a lo sumo $g(k)$ potencias k-esimas. Este trabajo fue llevado a cabo en una serie de papers por Dickson, Pillai, Rubugunday, y Niven, entre otros matemáticos. Todos estos trabajos se reducen en el siguiente teorema:

Teorema 3.1. *Sea $k \geq 6$, si la siguiente inecuación se verifica*

$$2^k . \left(\left(\frac{3}{2}\right)^k - \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor \right) + \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor \leq 2^k, \tag{3.2}$$

entonces $g(k) = 2^k + \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor - 2$. Mientras que si,

$$2^k . \left(\left(\tfrac{3}{2}\right)^k - \left\lfloor \left(\tfrac{3}{2}\right)^k \right\rfloor \right) + \left\lfloor \left(\tfrac{3}{2}\right)^k \right\rfloor > 2^k,$$

si definimos $N(k) = \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor . \left\lfloor \left(\frac{4}{3}\right)^k \right\rfloor + \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor + \left\lfloor \left(\frac{4}{3}\right)^k \right\rfloor$ se tiene que

$$g(k) = \left\lfloor \left(\tfrac{3}{2}\right)^k \right\rfloor + \left\lfloor \left(\tfrac{4}{3}\right)^k \right\rfloor + 2^k - 3 \quad si \quad 2^k < N(k),$$

o

$$g(k) = \left\lfloor \left(\tfrac{3}{2}\right)^k \right\rfloor + \left\lfloor \left(\tfrac{4}{3}\right)^k \right\rfloor + 2^k - 2 \quad si \quad 2^k = N(k).$$

Mahler probó usando la versión p-adica del teorema de Roth dada por Ridiout que $g(k) = 2^k + \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor - 2$ para todo k suficientemente grande.

Apliquemos el teorema tomando $K = \mathbb{Q}$, $S = \{\infty, 2, 3\}$, $\alpha_\infty = 1$, $\alpha_2 = \infty$, $\alpha_3 = 0$, y sea $\beta = 3^k/(n.2^k)$ con $n = \left\lceil \left(\frac{3}{2}\right) \right\rceil$. Se tiene que $|\alpha_2 - \beta|_2 = 2^{-k}|n|_2$, $|\alpha_3 - \beta| = 3^{-k}|n|_3^{-1}$. Por lo tanto, por el teorema de Roth, la inecuación

$$|\alpha_\infty - \beta|_\infty |\alpha_2 - \beta|_2 |\alpha_3 - \beta|_3 = |1 - 3^k/(n.2^k)|.2^{-k}|n|_2.3^{-k}|n|_3^{-1} < H(\beta)^{-2-\varepsilon}$$

posee finitas soluciones sobre k. En particular, usando que $|n|_2 \leq 1$ y que por definición de altura, $3^k|n|_3 \leq H(\beta)$. Se tendrá que la desigualdad más fina

$$|1 - 3^k/(n.2^k)|.2^{-k}.3^{-k}|n|_3^{-1} < (3^k|n|_3)^{-2-\varepsilon}$$

también posee finitas soluciones sobre k. Despejando, multiplicando por n y usando que $(2/3)^k.n \geq 1$ y que $1 \geq |n|_3$, se tendrá que la inecuación

$$\left\lceil \left(\tfrac{3}{2}\right)^k \right\rceil - \left(\tfrac{3}{2}\right)^k \geq 3^{-\varepsilon k}$$

es válida para todos salvo un números finitos de enteros positivos k, para cualquier $\varepsilon > 0$ fijo. Si tomamos $\varepsilon = log(4/3)/log\ 3$, y usando que esta desigualdad implica (3.2), deducimos el teorema de Mahler, que asegura que $g(k) = 2^k + \left\lfloor \left(\frac{3}{2}^k\right) \right\rfloor - 2$ para todo entero k suficientemente grande. Debido a la inefectividad del teorema de Roth, sigue siendo un problema abierto determinar efectivamente un k_0 para el cual el resultado sea válido para $k \geq k_0$.

3. Ecuaciones Diofánticas

Ya hemos mencionado que gran parte de esta teoría se ha desarrollado con el fin de resolver ciertas ecuaciones diofánticas o al menos poder determinar la finitud de su conjunto de soluciones. Como fue mostrado en el Capítulo 1, con su resultado sobre aproximaciones diofánticas, Thue fue capaz de probar que que llamada Ecuación de Thue,

$$F(X, Y) = m,$$

posee finitas soluciones enteras, si F es un polinomio homogéneo irreducible sobre $\mathbb{Q}[X, Y]$ con coeficientes enteros y m un entero no negativo.

Probaremos ahora un resultado más general utilizando el teorema de Roth sobre $\mathbb{Q}$ y el valor absoluto usual. Recordemos que en este caso el teorema de Roth es equivalente a que si α es algebraico de grado $d \geq 3$, y $k > 2$, entonces existe una constante $c(\alpha, k) > 0$ tal que

$$|\alpha - \xi| \geq c(\alpha, k)H(\xi)^k,$$

para todo $\xi \in \mathbb{Q}$. Para $d = 1$ o 2, el teorema de Liouville asegura que una desigualdad de este estilo sigue valiendo.

Un polinomio homogéneo $F(X, Y) \in \mathbb{Z}[X, Y]$ se dirá libre de cuadrados si no es divisible por $G(X, Y)^2$ para algun polinomio homogéneo $G \in \mathbb{Z}[X, Y]$.

Teorema 3.2. *Sea $F(X, Y) \in \mathbb{Z}[X, Y]$ un polinomio homogéneo libre de cuadrados de grado $d \geq 3$. Entonces para todo $k > 2$ existe una constante $c(F, k) > 0$ tal que para par de enteros (x, y) con $F(x, y) \neq 0$ se tiene que*

$$|F(x, y)| \geq c(F, k) \max\{|x|, |y|\}^{d-k}. \tag{3.3}$$

Si F posee grado $d \leq 2$, el teorema es verdadero trivialmente ya que $|F(x,y)| \in \mathbb{Z}$, luego ≥ 1.

DEMOSTRACIÓN. Se probara la desigualdad solo para los pares de enteros (x,y) con $|y| \geq |x|$. Intercambiando los roles de x e y, con el mismo argumento se prueba la desigualdad para los pares (x,y) con $|x| > |y|$.

Luego, nos restringiremos al caso en el que $|y| \geq |x|$ y F no es divisible por Y. Si F es divisible por Y, tendremos que $F = Y.F_1$, con $F_1 \in \mathbb{Z}[X,Y]$ un polinomio homogéneo libre de cuadrados de grado $d-1 \geq 2$, el cual no es divisible por Y. Por lo tanto, si la desigualdad se cumple para F_1 con $d-1$ en lugar de d, se cumple automáticamente para F.

Supongamos entonces que F es un polinomio homogéneo de grado $d \geq 2$, libre de cuadrados, que no es divisible por Y. Luego $F(X,Y) = a_0X^d + a_1X^{d-1}Y + \cdots + a_dY^d$, $a_0 \neq 0$, y entonces,

$$F(X,Y) = a_0\,(X - \alpha_1 Y)\ldots(X - \alpha_d Y)$$

con $\alpha_1, \ldots, \alpha_d$ distintos, pues F es libre de cuadrados. Sea (x,y) un par de enteros con $F(x,y) \neq 0$ y $|y| \geq |x|$. Como $y \neq 0$, sea $\xi := x/y$. Observar que $|y| = \max\{|x|,|y|\} \geq H(\xi)$(con igualdad si $mcd(x,y) = 1$). Sea i el índice para el cual

$$|\xi - \alpha_i| = \min_{j=1,\ldots,d} |\xi - \alpha_j|.$$

Por el teorema de Roth, tendremos que existe una constante $c(\alpha_i, k)$ tal que

$$|\xi - \alpha_i| \geq c(\alpha_i,k)H(\xi)^{-k} \geq c(\alpha_i,k)\max\{|x|,|y|\}^{-k}.$$

Para $j \neq i$, se tiene que

$$|\alpha_i - \alpha_j| \leq |\alpha_i - \xi| + |\xi - \alpha_j| \leq 2|\xi - \alpha_j|$$

lo que implica que

$$|\xi - \alpha_j| \geq \tfrac{1}{2}|\alpha_i - \alpha_j|.$$

Por lo tanto,

$$|F(x,y)| = |y|^d.|a_0|\prod_{j=1}^d |\xi - \alpha_j| = \max\{|x|,|y|\}^d\,|a_0|\prod_{j=1}^d |\xi - \alpha_j| \geq$$
$$c(\alpha_i,k)|a_0|\prod_{j\neq i}\left(\tfrac{1}{2}|\alpha_i - \alpha_j|\right).\max\{|x|,|y|\}^{d-k} = c'(\alpha_i,k)\max\{|x|,|y|\}^{d-k}.$$

Tomando el mínimo sobre las constantes $c'(\alpha_i,k)$ para $i = 1,\ldots,d$, obtenemos el resultado deseado. □

COROLARIO 2. *Sea $F(X,Y)$ un polinomio homogéneo en $\mathbb{Z}[X,Y]$ libre de cuadrados de grado $d \geq 3$, y sea $G(X,Y) \in \mathbb{Z}[X,Y]$ un polinomio de grado total $\leq d-3$. Entonces existen finitas pares $(x,y) \in \mathbb{Z}^2$ tal que $F(x,y) = G(x,y)$ y $F(x,y) \neq 0$.*

DEMOSTRACIÓN. Sea $G(X,Y) = \sum_{i+j\leq d-3} a_{ij}X^iY^j$, entonces si $(x,y) \in \mathbb{Z}$ se tendrá que

$$|G(x,y)| \leq \textstyle\sum_{(i+j\leq d-3)} |a_{ij}|\max\{|x|,|y|\}^{i+j} \leq C_1.\max\{|x|,|y|\}^{d-3}.$$

Sea $(x,y) \in \mathbb{Z}$ tal que $F(x,y) \neq 0$ y $F(x,y) = G(x,y)$ entonces

$$|F(x,y)| = |G(x,y)| \leq C_1\max\{|x|,|y|\}^{d-3}.$$

Sea $2 < k < 3$, entonces por el teorema anterior, existe $C_2 > 0$ tal que para todo (x, y) con $F(x, y) \neq 0$,

$$|F(x, y)| \geq C_2 \max\{|x|, |y|\}^{d-k}.$$

Por lo tanto,

$$\max\{|x|, |y|\}^{3-k} \leq C_3,$$

como $3 - k > 0$, concluimos lo deseado.

□

Si tomamos $G(x, y) = m \neq 0$, obtenemos nuevamente la Ecuación de Thue. La condición $F(x, y) \neq 0$ es necesaria pues si F y G posee algún factores lineal en común, con coeficientes racionales, habrá infinitas soluciones tal que $F(x, y) = G(x, y) = 0$.

Siegel fue el primero en probar un teorema de aproximaciones racionales sobre cuerpos de números algebraicos, lo que le permitío llegar a resultados más generales y probar la finitud de soluciones de ciertas ecuaciones sobre el anillo de enteros de un cuerpo de números. Podemos generalizar la ecuación de Thue en este contexto. De hecho, podremos considerar otro tipo de ecuaciones diofánticas que también posee finitas soluciones. En lo que sigue K será un cuerpo de números y R_K su anillo de enteros.

Proposición 4. *Las siguientes afirmaciones son equivalentes*

- *(M) Para todo cuerpo de números K y todo elemento no nulo $k \in K$, la ecuación de Mordell*
$$Y^2 = X^3 + k$$
posee finitas soluciones $(x, y) \in R_K \times R_K$.
- *(E) Para todo cuerpo de números K y todo polinomio $f \in K[X]$ de grado 3 con tres raíces complejas distintas, la ecuación elíptica*
$$Y^2 = f(X)$$
posee finitas soluciones $(x, y) \in R_K \times R_K$.
- *(HE) Para todo cuerpo de números K y todo polinomio $f \in K[X]$ de grado al menos 3 con raíces complejas simples, la ecuación hiperelíptica*
$$Y^2 = f(X)$$
posee finitas soluciones $(x, y) \in R_K \times R_K$.
- *(SE) Para todo cuerpo de números K, todo entero $m \geq 3$ y todo polinomio $f \in K[X]$ con al menos dos raíces complejas distintas cuyos ordenes de multiplicidad son coprimos con m, la ecuación superelíptica*
$$Y^m = f(X)$$
posee finitas soluciones $(x, y) \in R_K \times R_K$.
- *(T) Para todo cuerpo de números K, para todo elemento no nulo $k \in K$ y elementos $\alpha_1, \ldots, \alpha_n$ en K con $\#\{\alpha_1, \ldots, \alpha_n\} \geq 3$, la ecuación de Thue*
$$(X - \alpha_1 Y) \ldots (X - \alpha_n Y) = k$$
posee finitas soluciones $(x, y) \in R_K \times R_K$.
- *(S) Para todo cuerpo de números K y elementos $a_1, a_2 \in K$ no nulos, la ecuación de Siegel*
$$a_1 U + a_2 V = 1$$

posee finitas soluciones $(u, v) \in R_K^* \times R_K^*$.

Cada una de estas afirmaciones es un teorema: Mordell probó (M) en el caso de números enteros, y Thue lo hizo con (T), las primeras cuatro resultados son un caso particular del Teorema de Siegel sobre finitud de soluciones en el anillo de enteros de un cuerpo de números para curvas $F(X, Y) = 0$ que poseen genero al menos uno.

Veremos cómo se utiliza el Teorema de Roth para probar la finitud de la ecuación de Siegel. De hecho probaremos un resultado más general debido a Mahler quien fue el primero en trabajar con aproximaciones en cuerpos p-ádicos. Luego probaremos como deducir la finitud de soluciones enteras en caso de curvas hiperelípticas. Para una prueba completa de la equivalencia entre estos enunciados el lector se puede dirigir a [**34**].

En lo que sigue, $S \subset M_K$ será un subconjunto finito de valores absolutos sobre K que contiene a todos los valores absolutos arquimedianos y notaremos por R_S al anillo de S-enteros

$$R_S = \{x \in K |\ |x|_\upsilon \leq 1\ \forall\ \upsilon \notin S\}.$$

El Teorema de Dirichlet sobre las unidades de R_K establece que el rango del grupo de unidades del anillo de enteros de un cuerpo de números es igual a $r_1 + r_2 - 1$, con r_1 la cantidad de embedding reales de K en $\mathbb{C}$ y $2.r_2$ la cantidad de embeddings no reales. En el caso de S-enteros, este teorema sigue valiendo y el rango R_S será igual a $r_1 + r_2 - 1 + |S|$.

Teorema 3.3 (Siegel-Mahler). *Sea $K/\mathbb{Q}$ un cuerpo de números, sea $S \subset M_K$ un subconjunto finito de valores absolutos sobre K que contiene a todos los valores absolutos arquimedianos, sea R_S el anillo de S-enteros de K y sean $a, b \in K^*$. Entonces la ecuación*

$$aU + bV = 1 \tag{3.4}$$

tiene solo finitas soluciones en S-unidades $U, V \in R_S$

Demostración. Sea m un entero fijo, suficientemente grande(alcanzara con tomar $m = 2\#S + 1$). El teorema de S-unidades de Dirichlet implica que el grupo cociente $R_S^*/\left(R_S^*\right)^m$ es finito. Sea $c_1, \ldots, c_r$ un conjunto de representantes. Entonces toda solución (U, V) de (3.4) se puede escribir de la forma

$$U = c_i X^m \quad V = c_j Y^m$$

para algun $X, Y \in R_S^*$ y algun c_i, c_j, y por lo tanto (X, Y) es solución de la ecuación

$$ac_i X^m + bc_j Y^m = 1.$$

Como hay finitas elecciones posibles para c_i y c_j, será suficiente probar que para α, $\beta \in K^*$, la ecuación

$$\alpha X^m + \beta Y^m = 1$$

tiene finitas soluciones con $X, Y \in R_S^*$.

Supongamos por el contrario, que poseemos infinitas soluciones para una ecuación de este estilo. Como el conjunto S es finito, por el principio del palomar podremos encontrar un valor absoluto $\omega \in S$, para el cual la ecuación $\alpha X^m + \beta Y^m = 1$ posee infinitas soluciones (X, Y) y además

$$\|Y\|_\omega = \max\{\|Y\|_\upsilon \,|\upsilon \in S\}.$$

Sea γ una raíz m-esima de -β/α, que luego será determinada apropiadamente. Entonces

$$\frac{1}{\alpha Y^m} = \frac{X^m}{Y^m} + \frac{\beta}{\alpha} = \frac{X^m}{Y^m} - \gamma^m = \prod_{\zeta\in\mu_m}\left(\frac{X}{Y} - \zeta\gamma\right),$$

donde el producto está tomado sobre todas las raíces m-esimas de la unidad. Dado que hay infinitas soluciones, Y tiene valor absoluto no acotado con lo cual alguno de los factores del producto será chico. De hecho, solo uno de ellos podrá ser chico. En efecto, si ζ, $\zeta' \in \mu_m$ son raíces m-ésimas de la unidad distintas, usando la desigualdad triangular tendremos que

$$\left|\tfrac{X}{Y} - \zeta\gamma\right|_\omega + \left|\tfrac{X}{Y} - \zeta'\gamma\right|_\omega \geq |\zeta'\gamma - \zeta\gamma| \geq C_1.$$

Donde la constante $C_1 = C_1(\alpha, \beta, S, m)$ es independiente de X e Y. Por esta desigualdad, en el producto, todos salvo quizás uno de los factores deben ser mayores a $C_1/2$. Entonces

$$\frac{1}{|\alpha Y^m|_\omega} = \prod_{\zeta\in\mu_m}\left|\frac{X}{Y} - \zeta\gamma\right|_\omega \geq \left(\min_{\zeta\in\mu_m}\left|\frac{X}{Y} - \zeta\gamma\right|_\omega\right)\left(\frac{C_1}{2}\right)^{m-1}.$$

Por lo tanto, existe una constante $C_2 = C_2(\alpha, \beta, S, m)$ tal que

$$\frac{1}{\|Y\|_\omega^m} \geq C_2 \min_{\zeta\in\mu_m}\left\|\frac{X}{Y} - \zeta\gamma\right\|_\omega.$$

Para cada solucion (X, Y), existirá al menos una raíz m-ésima de la unidad ζ, que alcanza el mínimo del lado derecho de la inecuación anterior. Como estamos suponiendo que tenemos infinitas soluciones y hay finitas raíces m-esimas, entonces podemos suponer que hay infinitas soluciones $X, Y \in R_S^*$ de la ecuación $\alpha X^m + \beta Y^m = 1$ que cumplen

$$\frac{1}{\|Y\|_\omega^m} \geq C_2\left\|\frac{X}{Y} - \zeta\gamma\right\|_\omega.$$

Esto nos dice que X/Y es una buena aproximación de $\zeta\gamma$. Para poder aplicar el Teorema de Roth, necesitamos relacionar $\|Y\|_\omega^m$ con la altura de X/Y.

El valor absoluto ω fue elegido para maximizar $\|Y\|_\upsilon$. Como ademas $\|Y\|_\upsilon = 1$ para todo $\upsilon \notin S$, se tiene que

$$\|Y\|_\omega = \max_{\upsilon\in S}\|Y\|_\upsilon \geq \left(\prod_{\upsilon\in S}\|Y\|_\upsilon\right)^{1/\#S} = \left(\prod_{\upsilon\in M_K}\|Y\|_\upsilon\right)^{1/\#S} \geq H_K(Y)^{1/\#S}.$$

Recordando propiedades elementales de la altura, $H(x+y) \leq 2H(x)H(y)$, $H(xy) \leq H(x)H(y)$, $H_K(x) = H(x)^{[K:\mathbb{Q}]}$, y usando que (X, Y) es solución de $\alpha X^m + \beta Y^m = 1$

$$H_K\left(\tfrac{X^m}{Y^m}\right) = H_K\left(\tfrac{1}{\alpha Y^m} - \tfrac{\beta}{\alpha}\right) \leq 2^{[K:\mathbb{Q}]}H_K(\tfrac{1}{Y^m})H_K(\tfrac{1}{\alpha})H_K(\tfrac{\beta}{\alpha}).$$

Tomando raíz m-esima y usando que $H_K(T^m) = H_K(T)^m$, se tiene que existe una constante $C_3 = C_3(\alpha, \beta, S, m)$ tal que

$$H_K(X/Y) \leq C_3 H_K(1/Y) = C_3 H_K(Y).$$

Utilizando la cota obtenida anteriormente para $H_K(Y)$, se tendrá que si llamamos $C_4 = C_3^{1/\#S}$, entonces

$$\|Y\|_\omega \geq C_4 H_K(X/Y)^{1/\#S}.$$

Se concluye que existe una constante $C_5 = C_5(\alpha, \beta, S, m, K) = 1/C_2 C_4^m$ para la cual

$$\frac{C_5}{H_K(X/Y)^{m/\#S}} \geq \left\| \frac{X}{Y} - \zeta\gamma \right\|_\omega,$$

por suposición, posee infinitas soluciones $X, Y \in R_S^*$. Recordando que podemos tomar $m = 2\#S + 1$, el teorema de Roth nos dice que esto es un absurdo, llegando así a la contradicción deseada. □

Teorema 3.4. *Sea $K/\mathbb{Q}$ un cuerpo de números, sea $S \subset M_K$ un subconjunto finito de valores absolutos sobre K que contiene a todos los valores absolutos arquimedianos y sea R_S el anillo de S-enteros de K. Sea $f(X) \in K[X]$ un polinomio de grado al menos 3 con raíces distintas(sobre $\overline{K}$). Entonces la ecuación*

$$Y^2 = f(X)$$

posee finitas soluciones en S-enteros $X, Y \in R_S$.

Demostración. Primero observemos que es posible reemplazar K por una extensión finita L, en la cual $f(X)$ se factorice linealmente. En efecto, supongamos que el teorema es válido bajo estas condiciones y sea $f(X)$ un polinomio en $K[X]$. Sea L una extensión finita de K en la cual $f(X)$ se factoriza linealmente. Consideremos S' el conjunto de valores absolutos sobre L que consiste de todas las extensiones a L de los valores absolutos en S, es decir, todos los valores absolutos en L que estan arriba de algun valor absoluto de S. Como L/K es una extensión finita de cuerpos de números, cada valor absoluto de S posee finitas extensiones. Luego S' resulta finito y posee todos los valores absolutos arquimedianos de L. Dado que $R_S^K \subset R_{S'}^L$ se concluye que la validez del teorema para L y S', implica la validez para K y S.

Notar que mientras más grande sea S, más fuerte será el resultado del teorema. Por lo tanto, probar el teorema para un conjunto de valores absolutos más grande no afectara la conclusión buscada. Ya vimos que podemos suponer que $f(X)$ se factoriza linealmente en K, entonces tenemos que

$$f(X) = a(X - \alpha_1) \dots (X - \alpha_n) \text{ con } \alpha_1, \dots, \alpha_n \in K.$$

Por hipótesis, $n \geq 3$ y los α_i son todos distintos.

Incrementaremos el tamaño de S para que se satisfagan las siguientes condiciones

(1) $a \in R_S^*$.
(2) $\alpha_i - \alpha_j \in R_S^*$ para todo $i \neq j$.
(3) R_S es un anillo de ideales principales.

Recordemos que para todo $x \in K$, $|x|_v = 1$ salvo finitos $v \in M_K$. Por lo tanto, agregando finitos valores absolutos a S, la condicion (1) y (2) serán válidas. Para ver (3), recordar que el grupo de clases de R_S es finito, por lo tanto es suficiente agregar a S un primo de cada clase de ideales.

Sea $(X,Y) \in R_S^2$ una solución de la ecuación $Y^2 = f(X)$. Veamos que el ideal $(X-\alpha_i)R_S$ es el cuadrado de un ideal en R_S. Sea $\mathfrak{p}$ un ideal primo de R_S, entonces p solo puede dividir como mucho a uno de los ideales $(X-\alpha_i)R_S$ pues sino divide a dos de ellos, dividirá a $(\alpha_i-\alpha_j)R_S$ lo cual contradice (2). A partir de (1), tenemos que $\mathfrak{p}$ no puede dividir a aR_S. Por lo tanto, se sigue de la la factorización de $f(X)$, que si $\mathfrak{p}$ divide a uno de los ideales $(X-\alpha_i)R_S$, el orden de $\mathfrak{p}$ debe ser par. Entonces existen ideales $\mathfrak{a}_i \subset R_S$ tal que

$$(X-\alpha_i)R_S = \mathfrak{a}_i^2 \text{ para todo } 1 \leq i \leq n.$$

Por propiedad (3) tenemos que R_S es un anillo de ideales principales, entonces $\mathfrak{a}_i = (Z_i)R_S$, para algun $Z_i \in R_S$. Por lo tanto existe una unidad $U_i \in R_S^*$ tal que $X-\alpha_i = U_i Z_i^2$, para todo $1 \leq i \leq n$.

Extendemos K a un cuerpo L, adjuntándole todas las raíces cuadradas de elementos de R_S^*. Por el teorema de S-unidades de Dirichlet, el grupo R_S^* es finitamente generado, por lo tanto $R_S^*/(R_S^*)^2$ es finito, entonces L/K será una extensión finta pues estara generado por las raíces cuadradas de un conjunto de representantes de $R_S^*/(R_S^*)^2$. Sea $T \subset M_L$ el conjunto de lugares que están arriba de los elementos de S. Como L/K es finita, T resulta finito. En L, cada U_i resulta un cuadrado, digamos $U_i = V_i^2$. Entonces se tiene que $X-\alpha_i = (V_iZ_i)^2 = W_i^2$ para algun $W_i \in R_T$. Considerando la diferencia de dos de estas igualdades tenemos que

$$\alpha_i - \alpha_j = W_i^2 - W_j^2 = (W_i - W_j)(W_i + W_j).$$

Por (2), el lado izquierdo de esta ecuación esta en R_T^*, mientras que cada uno de los factores de la derecha están en R_T. Por lo tanto, cada uno de estos factores debe ser una unidad,

$$W_i - W_j, W_i + W_j \in R_T^* \text{ para todo } i \neq j.$$

Como $f(X)$ posee grado al menos 3, podemos escribir la siguiente identidad(a veces llamada identidad de Siegel)

$$\tfrac{W_1-W_2}{W_1-W_3} + \tfrac{W_2-W_3}{W_1-W_2} = 1.$$

Cada uno de los términos del lado derecho posee finitas opciones por teorema*(S-unit equation). Similarmente, la identidad

$$\tfrac{W_1+W_2}{W_1-W_3} + \tfrac{W_3+W_2}{W_3-W_1} = 1$$

y el teorema* nos dice que cada uno de los términos en esta ecuación posee finitas opciones. Se sigue que hay solo finitos valores posibles para la cantidad

$$\tfrac{W_1-W_2}{W_1-W_3}.\tfrac{W_1+W_2}{W_1-W_3} = \tfrac{W_1^2-W_2^2}{(W_1-W_3)^2} = \tfrac{\alpha_2-\alpha_1}{(W_1-W_3)^2}.$$

Por lo tanto, también habrá finitos valores posibles para $W_1 - W_3$, lo que implica finitos valores posibles para

$$\tfrac{1}{2}\left((W_1-W_3) + \tfrac{\alpha_3-\alpha_1}{W_1-W_3}\right) = \tfrac{1}{2}\left((W_1-W_3)+(W_1+W_3)\right) = W_1,$$

y entonces finitos valores posibles para $\alpha_1 + W_1^2 = X$. Finalmente, para cada valor de X, hay como mucho 2 valores posibles para Y. Lo que completa la prueba de la finitud de las soluciones de $Y^2 = f(X)$ con $X, Y \in R_S$. □

CAPÍTULO 4

El Teorema del Subespacio

1. El Teorema del Subespacio

El Teorema del Subespacio, probado por Schmidt en 1972, es una generalización multidimensional del Teorema de Roth y fue originalmente desarrollado para el estudio de las aproximaciones de números algebraicos por medio de números algebraicos de grado acotado y la norm form equation(una clase de ecuaciones diofánticas que incluyen a las ecuaciones de Thue). Este teorema, involucra sistema de inecuaciones en formas lineales. Es importante mencionar que esta no es una generalización rutinaria, sino que nuevas dificultades aparecen a lo largo de prueba, las cuales fueron resueltas por Schmidt al introducir nuevas ideas de la Geometría de Números. Puede verse una prueba del teorema en [**29**], [**12**] se basa principalmente en la referencia anterior pero posee varias simplificaciones. Una prueba del teorema en su versión más general puede encontrarse en [**4**].

Sea n un entero positivo y $r \leq n$. Diremos que las formas lineales $L_1 = \sum_{j=1}^{n} a_{1j}x_j, \ldots, L_r = \sum_{j=1}^{n} a_{rj}x_j$ con $a_{ij} \in \mathbb{C}$ son linealmente dependientes si existen $c_1, \ldots, c_r \in \mathbb{C}$, no todos nulos, tal que $c_1L_1 + \ldots + c_rL_r \equiv 0$. Caso contrario, diremos que son linealmente independientes. Es facil ver que, si $r = n$, $L_1, \ldots, L_n$ son linealmente independientes si y solo si $det(L_1, \ldots, L_n) = det(a_{ij}) \neq 0$.

La norma de $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ esta dada por $\|\mathbf{x}\| := \max\{|x_1|, \ldots, |x_n|\}$.

Teorema 4.1 (Schmidt, 1972). *Sea $n \geq 2$, y sean $L_1(\boldsymbol{X}), \ldots, L_n(\boldsymbol{X})$, n formas lineales en las variables $\boldsymbol{X} = (X_1, \ldots, X_n)$, linealmente independientes, con coeficientes en $\overline{\mathbb{Q}}$. Sean $C > 0$ y $\varepsilon > 0$, entonces existe un número finito de subespacios lineales propios $T_1, \ldots, T_h$ de $\mathbb{Q}^n$ tal que el conjunto de soluciones de*

$$|L_1(\boldsymbol{x}) \cdots L_n(\boldsymbol{x})| < C \|\boldsymbol{x}\|^{-\varepsilon} \qquad \boldsymbol{x} \in \mathbb{Z}^n, \tag{4.1}$$

está contenido en $T_1 \cup \cdots \cup T_h$.

El Teorema del Subespacio puede reformularse en una forma más geométrica de la siguiente manera. La topología del Subespacio (noción introducida por Schmidt) en $\mathbb{Q}^n$ es la topología cuyos conjuntos cerrados son uniones finitas de subespacioes lineales de $\mathbb{Q}^n$. Entonces, el Teorema 4.1 establece que para todo $C > 0$ y $\varepsilon > 0$, el conjunto de soluciones de (4.1) no es denso en $\mathbb{Q}^n$ con respecto a la topología del Subespacio.

Veamos que el Teorema del Subepacio implica el Teorema de Roth. Recordemos que la altura de $\xi \in \mathbb{Q}$ es $H(\xi) = \max\{|x|, |y|\}$, con $\xi = x/y$, x,$y \in \mathbb{Z}$, $mcd(x, y) = 1$.

Corolario 3 (Teorema de Roth). *Sea $\alpha \in \overline{\mathbb{Q}}$ y $C > 0$, $k > 2$. Entonces la inecuación*

$$|\alpha - \xi| \leq CH(\xi)^{-k} \qquad \xi \in \mathbb{Q} \tag{4.2}$$

posee finitas soluciones.

DEMOSTRACIÓN. Sea $\xi = x/y$ una solución de (4.2), con x,$y \in \mathbb{Z}$, $\text{mcd}(x, y) = 1$. Sea $\varepsilon > 0$ tal que $k = 2 + \varepsilon$. Multiplicando (4.2) por y^2 obtenemos

$$|y(x - \alpha y)| \leq Cy^2 \max\{|x|, |y|\}^{-2-\varepsilon} \leq C. \max\{|x|, |y|\}^{-\varepsilon}$$

Dado que las formas lineales Y y $X - \alpha Y$ son linealmente independientes, podemos aplicar el Teorema del Subespacio. Se sigue entonces que los pares de enteros $(x, y) \in \mathbb{Z}^2$, con $\text{mcd}(x, y) = 1$ tal que $\xi = x/y$ es solución de (4.2), están contenidos en una unión finita de subespacios propios de $\mathbb{Q}^2$, es decir, subespacios lineales de dimensión uno. Los subespacios uno dimensionales de $\mathbb{Q}^2$ consisten en todos los puntos de la forma $\lambda(x_0, y_0)$ con $\lambda \in \mathbb{Q}$ y (x_0, y_0) puede tomarse en $\mathbb{Z}^2$ con $\text{mcd}(x_0, y_0) = 1$. Por lo tanto, $\xi = x_0/y_0$, esta unívocamente determinado por el subespacio. □

El Teorema del Subespacio establece que el conjunto de soluciones de (4.1) esta contenido en finitos subespacios lineales propios de $\mathbb{Q}^n$. En este caso, logramos ver que de hecho existen finitas soluciones. Algo similar ocurrirá siempre que consideremos $n = 2$. Si tenemos una solución no nula $x_0 \in \mathbb{Z}^n$ tal que $L_1(x_0) = 0$, entonces $\lambda.x_0$ será solución de (4.1) para todo $\lambda \in \mathbb{Z}$. Por lo tanto, tendremos así infinitas soluciones de (4.1). Para evitar este tipo de construcción, consideremos

$$0 < |L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| < C|\mathbf{x}|^{-\varepsilon} \qquad \mathbf{x} \in \mathbb{Z}^n, \tag{4.3}$$

En el caso $n = 2$, el conjunto de soluciones de (4.3) será de hecho finito.

LEMA 11. *Sea $L_i = a_{i1}X + \alpha_{i2}$, para $i = 1, 2$, dos formas lineales linealmente independientes con coeficientes en $\overline{\mathbb{Q}}$. Sea $C > 0$, $\varepsilon > 0$. Entonces la inecuación*

$$0 < |L_1(\boldsymbol{x})L_2(\boldsymbol{x})| \leq C \|\boldsymbol{x}\|^{-\varepsilon} \text{ en } x \in \mathbb{Z}^2 \tag{4.4}$$

posee finitas soluciones.

DEMOSTRACIÓN. Por el Teorema del Subespacio, las soluciones de (4.4) pertenecen a un número finito de subespacios lineales de dimensión 1 de $\mathbb{Q}^2$. Cada uno de estos subespacios T se puede representar de la forma $T = \{\lambda(x_0, y_0) | \lambda \in \mathbb{Q}\}$ donde $(x_0, y_0) \in \mathbb{Z}^2$ se puede elegir de forma tal que $mcd(x_0, y_0) = 1$. Observar que $\lambda(x_0, y_0) \in \mathbb{Z}^2$ si y solo si $\lambda \in \mathbb{Z}$. Si $L_1(x_0, y_0)L_2(x_0, y_0) = 0$, entonces (4.4) no posee soluciones en T. Supongamos que $L_1(x_0, y_0)L_2(x_0, y_0) \neq 0$. Entonces $\lambda(x_0, y_0)$es solución de (4.4) si y solo si

$$0 < \lambda^2 |L_1(x_0, y_0)L_2(x_0, y_0)| \leq C|\lambda|^{-\varepsilon} \|(x_0, y_0)\|^{-\varepsilon}.$$

Por lo tanto, $|\lambda|^{2+\varepsilon} \leq C|L_1(x_0, y_0)L_2(x_0, y_0)|^{-1} \|(x_0, y_0)\|^{-\varepsilon}$. Es decir, λ esta acotado. □

Si $n \geq 3$, (4.3) puede tener infinitas soluciones, como muestra el siguiente ejemplo. Sea $0 < \varepsilon < 1$ y consideremos la inecuación

$$0 < |(x_1 + \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 - \sqrt{3}x_3)| \leq \|\mathbf{x}\|^{-\varepsilon} \tag{4.5}$$

en $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$. Observar que las 3 formas lineales consideradas son linealmente independientes.

Consideremos las uplas de enteros $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{Z}^3$ con $x_3 = 0$, $x_1.x_2 \neq 0$. Para estos $\mathbf{x}$, $\|\mathbf{x}\| = \max\{|x_1|, |x_2|\}$. Por el teorema de Dirichlet, existen infinitos racionales x_1/x_2 tal que

$$\left|\sqrt{2} - \tfrac{x_1}{x_2}\right| \leq |x_2|^2.$$

Para estas soluciones $\mathbf{x} = (x_1, x_2, 0)$, dado que

$$|x_1/x_2| \leq |x_2|^2 + \sqrt{2} \leq 1 + \sqrt{2},$$

entonces $\|\mathbf{x}\| = \max\{|x_1|, |x_2|\} \leq (1+\sqrt{2})|x_2|$. Por lo tanto, para los puntos considerados,

$$0 < |(x_1 + \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 - \sqrt{3}x_3)| =$$
$$|(x_1 + \sqrt{2}x_2)(x_1 - \sqrt{2}x_2)^2| \leq (1+\sqrt{2})\,\|\mathbf{x}\|\,.(x_2^{-1})^2 \leq (1+\sqrt{2})^3\,\|\mathbf{x}\|^{-1} \leq \|\mathbf{x}\|^{-\varepsilon},$$

si $\|\mathbf{x}\|$ es suficientemente grande. Por lo tanto, (4.5) posee infinitas soluciones en el subespacio $\{x_3 = 0\}$. De la misma forma se prueba que hay infinitas soluciones en los subespacios $\{x_1 = 0\}$, $\{x_2 = 0\}$ y con un poco más de trabajo ver que (4.5) posee finitas soluciones con $x_1 x_2 x_3 \neq 0$.

Mencionaremos a continuación una versión mejorada del Teorema del Subespacio, probada por Vojta [**33**] en 1989.

TEOREMA 4.2. *Sean $L_1, \ldots, L_n$ formas lineales linealmente independientes en n variables, con coeficientes en $\overline{\mathbb{Q}}$. Entonces existe un conjunto finito de subespacios lineales de $\mathbb{Q}^n$ efectivamente determinables $T_1, \ldots, T_l$. Tal que para todo $\varepsilon > 0$, el conjunto de soluciones de*

$$|L_1(\boldsymbol{x}) \ldots L_n(\boldsymbol{x})| \leq \|\boldsymbol{x}\|^{-\varepsilon} \quad en \quad \boldsymbol{x} \in \mathbb{Z}^n$$

está contenido en $T_1, \ldots, T_n$, salvo por un conjunto finito de soluciones que pueden depender de ε.

2. Aproximaciones Simultaneas y Algebraicas de grado acotado

El siguiente teorema de Dirichlet de 1842, generaliza al probado en el capitulo 1. Su prueba es similar y puede leerse en [**29**].

TEOREMA 4.3. *Sean $\alpha_1, \ldots, \alpha_n$ números reales $\mathbb{Q}$-linealmente independientes. Entonces para alguna constantes $C > 0$, la inecuación*

$$|\alpha_1 x_1 + \ldots + \alpha_n x_n| \leq C\,\|\boldsymbol{x}\|^{-n+1} \quad en \quad \boldsymbol{x} \in \mathbb{Z}^n$$

posee infinitas soluciones.

En 1970, Schmidt probó que el exponente -n+1 no puede ser reemplazado por ningún número más chico si $\alpha_1, \ldots, \alpha_n$ son algebraicos.

TEOREMA 4.4. *Sean $\alpha_1, \ldots, \alpha_n$ números algebraicos. Para toda $C > 0$ y $\varepsilon > 0$, la inecuacion*

$$0 < |\alpha_1 x_1 + \ldots + \alpha_n x_n| \leq C\,\|\boldsymbol{x}\|^{-n+1-\varepsilon} \quad en \quad \boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n \tag{4.6}$$

posee finitas soluciones.

DEMOSTRACIÓN. Procederemos por inducción en n. El caso $n = 1$ resulta trivial. Sea $n > 1$ y supongamos que el teorema es cierto para $n-1$. Podemos asumir que al menos uno de los coeficientes $\alpha_1, \ldots, \alpha_n$ es no nulo pues sino, no habría soluciones. Asumamos entonces que $\alpha_1 \neq 0$. Para cada solución $\mathbf{x}$ de (4.6) se tiene que

$$|(\alpha_1 x_1 + \ldots + \alpha_n x_n) x_2 \ldots x_n| \leq C\,\|\mathbf{x}\|^{-\varepsilon}. \tag{4.7}$$

Dado que el conjunto de formas lineales $\{\alpha_1 X_1 + \ldots + \alpha_n X_n, X_2, \ldots, X_n\}$ es linealmente independiente, el teorema del subespacio asegura que existen finitos subespacios lineales propios de $\mathbb{Q}^n$, $T_1, \ldots, T_t$, que contienen las soluciones de (4.7).

Sea T uno de estos subespacios, y consideremos $a_1X_1 + \ldots + a_nX_n$ una forma lineal que se anula completamente en T. Sin perdida de generalidad, podemos suponer que $a_n \neq 0$. Podemos entonces expresar a X_n como combinacion lineal de $X_1, \ldots, X_n$. Se tendrá que existen $\beta_1, \ldots, \beta_{n-1}$ números algebraicos para los cuales $\alpha_1X_1 + \ldots + \alpha_nX_n = \beta_1X_1 + \ldots + \beta_{n-1}X_{n-1}$ sobre T. Por lo tanto, toda solución $\mathbf{x}$ de (4.6) con $\mathbf{x} \in T$,cumple que

$$0 < |\beta_1x_1 + \ldots + \beta_{n-1}x_{n-1}| \leq C \|\mathbf{x}\|^{-n+1-\varepsilon} \leq max(|x_1|, \ldots, |x_{n-1}|)^{-n+2-\varepsilon} = \|\mathbf{x}'\|^{-n+1-\varepsilon},$$

con $\mathbf{x}' = (x_1, \ldots, x_{n-1})$. Por hipótesis inductiva, esta desigualdad posee finitas soluciones en $\mathbf{x}' \in \mathbb{Z}^{n-1}$. Por lo tanto, (4.6) tiene finitas soluciones con $\mathbf{x}$ en T. Como hay finitos subespacios posibles, completamos así la prueba del teorema. □

En el trabajo sobre aproximaciones diofánticas de Siegel [**25**], además de considerar las aproximaciones de números algebraicos por medio de números algebraicos en un cuerpo de números fijo, tambien estudio la aproximación por medio de números algebraico de grado acotado. Si bien los argumentos de Roth se generalizaron bien en el primer caso, no admitieron una extensión similar para el segundo problema. El siguiente corolario se debe a Schmidt [**30**]. En este corolario, si ξ es un número algebraico, $H(\xi)$ denotara al máximo de los valores absolutos de los coeficientes del polinomio minimal de ξ en $\mathbb{Z}[X]$.

Corolario 4. *Sea α un número algebraico. Para todo entero $d \geq 1$ y $\varepsilon > 0$, existen finitos números algebraicos ξ de grado d tal que*

$$|\alpha - \xi| < H(\xi)^{-d-1-\varepsilon}. \tag{4.8}$$

Demostración. Sea ξ un número algebraico de grado d que satisface (4.8). Dado que α posee solo finitos conjugados, podemos asumir que ξ no es uno de ellos. Sea $f(X) = x_{d+1}X^d + \ldots + x_1 \in \mathbb{Z}[X]$ el polinomio minimal de ξ sobre $\mathbb{Z}$. Se tiene que si $\mathbf{x} = (x_1, \ldots, x_{d+1})$ entonces $H(\xi) = \|\mathbf{x}\|$, ademas $f(\alpha) \neq 0$. A partir del desarrollo de Taylor o el teorema del valor medio, obtenemos que $|f(\alpha)| \leq C(\alpha, d)|\alpha - \xi|H(\xi)$. Entonces

$$0 < |x_1 + x_2\alpha + \ldots + x_{d+1}\alpha^d| = |f(\alpha)| \leq C(\alpha, d)|\alpha - \xi|H(\xi) < C(\alpha, d)H(\xi)^{-d-\varepsilon} = C(\alpha, d)\|\mathbf{x}\|^{-d-\varepsilon}.$$

Por el teorema anterior, existen solo finitos $\mathbf{x} \in \mathbb{Z}^{d+1}$ que cumplen esta desigualdad. Esto implica que hay solo finitos ξ algebraicos de grado d que cumple (4.8). □

En particular, dados $d \geq 1$ y $\varepsilon > 0$ existirán finitos números algebraicos ξ de grado a lo sumo d para los cuales valga (4.8). Pues hay finitos para cada $d' \leq d$ ya que

$$|\alpha - \xi| < H(\xi)^{-d-1-\varepsilon} \leq H(\xi)^{-d'-1-\varepsilon}.$$

Si $d = 1$, este resultado se reduce al Teorema de Roth. Un resultado más débil, con $d + 1 + \varepsilon$ reemplazado por $2d + \varepsilon$ fue probado por Wirsing [**35**] por medio de un método distinto.

Un problema abierto es encontrar el mejor exponente κ_d para el cual la inecuación

$$|\alpha - \xi| < H(\xi)^{-\kappa_d+\varepsilon}$$

posee infinitas soluciones reales algebraicas ξ de grado como mucho d, para todo $\varepsilon > 0$ fijo y todo α real que no es algebraico de grado como mucho d, para alguna constante c dependiente de α y ε. Si $d = 1$ y α es irracional, el teorema de Dirichlet prueba que $\kappa_1 = 2$(incluso con $\varepsilon = 0$). Si $d = 2$ y α no es racional ni cuadrático irracional, Davenport y Schmidt [**9**] probaron que $\kappa_2 = 3$(incluso con $\varepsilon = 0$). Para $d \geq 3$ este resultado esta abierto y no se sabe cual podría ser una respuesta correcta.

Respecto al problema de aproximanciones por medio de enteros algebraicos, Davenport y Schmidt [**10**] probaron que la inecuación

$$|\alpha - \xi| \leq cH(\xi)^{-2}$$

posee infinitas soluciones en enteros algebraicos ξ de grado como mucho 2 para alguna constante $c = c(\alpha)$, si α es irracional. Además, el exponente es óptimo si α es irracional cuadratico. Para el caso $d = 3$, probaron que si α no es algebraico de grado como mucho 2, entonces la inecuación

$$|\alpha - \xi| \leq cH(\xi)^{-(3+\sqrt{5})/2}$$

posee infinitas soluciones con ξ entero algebraico de grado como mucho 3. Pasaron más de 30 años hasta que en 2003 Roy [**24**] construyo un numero trascendente real α y una constantes $c > 0$ tal que

$$|\alpha - \xi| \geq cH(xi)^{-(3+\sqrt{5})/2}$$

para todo entero algebraico ξ de grado como mucho 3.

3. La version p-adica del Teorema del Subespacio

Al igual que para el Teorema de Roth, existe una version p-ádica del Teorema del Subespacio. Esta fue probada por Schlickewei [**27**], [**28**].

Teorema 4.5 (Schlickewei). *Sea $n \geq 2$,$\boldsymbol{X} = (X_1, \ldots, X_n)$, $p_1, \ldots, p_s$ finitos números primos distintos y tomemos una extensión de cada valor absoluto p-adico a $\overline{\mathbb{Q}}$. Sean $L_{1,\infty}(\boldsymbol{X}), \ldots, L_{n,\infty}(\boldsymbol{X})$ formas lineales linealmente independientes, con coeficientes algebraicos. Para cada primo p_j sean $L_{1,p_j}(\boldsymbol{X}), \ldots, L_{n,p_j}(\boldsymbol{X})$ formas lineales linealmente independientes, con coeficientes algebraicos. Dado $\varepsilon > 0$, entonces existe un número finito de subespacios lineales propios $T_1, \ldots, T_h$ de $\mathbb{Q}^n$ tal que el conjunto de soluciones de*

$$|L_{1,\infty}(\boldsymbol{x}) \cdots L_{n,\infty}(\boldsymbol{x})| \prod_{j=1}^{s} |L_{1,p_j}(\boldsymbol{x}) \cdots L_{n,p_j}(\boldsymbol{x})|_{p_j} < \|\boldsymbol{x}\|^{-\varepsilon} \quad x \in \mathbb{Z}^n, \tag{4.9}$$

esta contenido en $T_1 \cup \cdots \cup T_h$.

La siguiente generalización del Teorema del Subespacio es debida a Vojta y será útil en algunas aplicaciones. Sean $L_1, \ldots, L_r$ formas lineales con coeficientes en $\mathbb{C}$ en variables $X_1, \ldots, X_n$, con $r \geq n$. Diremos que $L_1, \ldots, L_r$, estan en posición general si cada subconjunto de cardinal $\leq n$ de $\{L_1, \ldots, L_r\}$ es linealmente independiente.

Teorema 4.6. *Ses S un conjunto finito de primos, incluyendo $p = \infty$ y tomemos una extensión de cada valor absoluto p-adico a $\overline{\mathbb{Q}}$. Para cada $p \in S$ sea $r_p \geq n$ y sean $L_{1,p}(\boldsymbol{X}), \ldots, L_{r_p,p}(\boldsymbol{X})$ formas lineales en las variables $\boldsymbol{X} = (X_1, \ldots, X_n)$, en posición general, con coeficientes algebraicos. Sea $C > 0$ y $\varepsilon > 0$, entonces existe un número finito de subespacios lineales propios $T_1, \ldots, T_h$ de $\mathbb{Q}^n$*

tal que el conjunto de soluciones de

(4.10)
$$\prod_{p\in S} |L_{1,p}(\boldsymbol{x})\cdots L_{r_p,p}(\boldsymbol{x})|_p < C\,\|\boldsymbol{x}\|^{r_\infty - n - \varepsilon} \quad x\in\mathbb{Z}^n \text{ con } mcd(x_1,\ldots,x_n)=1,$$

está contenido en $T_1\cup\cdots\cup T_h$.

Para probar esta versión del teorema, usaremos el siguiente lema el cual se deduce a partir de la equivalencia de normas en $\mathbb{C}^n$.

LEMA 12. *Sean* $M_1,\ldots,M_n$ *formas lineales linealmente independientes en las variables* $X_1,\ldots,X_n$ *con coeficientes complejos. Entonces existe una constante* $C>0$ *tal que*

$$\|\boldsymbol{x}\| \le C\max\{|M_1(\boldsymbol{x})|,\ldots,|M_n(\boldsymbol{x})|\} \text{ para todo } x\in\mathbb{C}^n.$$

DEMOSTRACIÓN DEL TEOREMA. Partimos al conjunto de soluciones $\mathbf{x}\in\mathbb{Z}^n$ de (4.10) en un número finito de clases dependiendo en cuales de las n cantidades sobre $|L_{1,p}(\mathbf{x})|_p\cdots|L_{r_p,p}(\mathbf{x})|_p$ para cada $p\in S$. Será suficiente probar que las soluciones en cada una de estas clases pertenece a un número finito de subespacios lineales propios de $\mathbb{Q}^n$.

Sin pérdida de generalidad, consideremos las soluciones $\mathbf{x}\in\mathbb{Z}^n$ de (4.10) para las cuales $|L_{1,p}(\mathbf{x})|_p\cdots|L_{n,p}(\mathbf{x})|_p$ son los menores entre $|L_{1,p}(\mathbf{x})|_p\cdots|L_{r_p,p}(\mathbf{x})|_p$ para cada $p\in S$.

Por el lema anterior, para $i=n+1,\ldots,r_\infty$, dado que $L_{1,\infty},\ldots,L_{n-1,\infty},L_{i,\infty}$ son linealmente independientes, existe una constante $C_i>0$ tal que para todas las soluciones x consideradas,

(4.11)
$$\|\mathbf{x}\| \le C_i\max\{|L_{1,\infty}(\mathbf{x})|_\infty,\ldots,|L_{n-1,\infty}(\mathbf{x})|_\infty,|L_{i,\infty}(\mathbf{x})|_\infty\} = C_i|L_{i,\infty}(\mathbf{x})|_\infty.$$

Sea ahora $p\in S$, $p\neq\infty$. Dado que solo estamos considerando soluciones para las cuales el mcd es igual a 1, para cada solución $\mathbf{x}=(x_1,\ldots,x_n)$ en consideración, existe un índice k para el cual $|x_k|_p=1$. Para cada $i=n+1,\ldots,r_p$, como $L_{1,p},\ldots,L_{n-1,p},L_{i,p}$ son linealmente independientes, general al espacio de formas lineales en $\overline{\mathbb{Q}}$, se tiene que existen constantes $\alpha_1,\ldots,\alpha_n$ tal que

$$X_k = \alpha_1 L_{1,p} + \ldots + \alpha_{n-1}L_{n-1,p} + \alpha_n L_{i,p}.$$

Por la desigualdad triangular fuerte, obtenemos

$$1 = |x_k|_p \le \max_{j=1,\ldots,n} |\alpha_j|_p |L_{j,p}(\mathbf{x})|_p \le C_{i,p}|L_{i,p}(\mathbf{x})|_p \tag{4.12}$$

para alguna constante $C_{i,p}$. Juntando (4.11), (4.12) con (4.10) obtenemos

$$|L_{1,\infty}(\mathbf{x})\cdots L_{n,\infty}(\mathbf{x})|_\infty \prod_{p\in S,p\neq\infty} |L_{1,p}(\mathbf{x})\cdots L_{n,p}(\mathbf{x})|_p \le$$

$$C'|L_{1,\infty}(\mathbf{x})\cdots L_{n,\infty}(\mathbf{x})|_\infty.\frac{|L_{n+1,\infty}(\mathbf{x})|_\infty}{\|\mathbf{x}\|}\cdots\frac{|L_{r_\infty,\infty}(\mathbf{x})|_\infty}{\|\mathbf{x}\|}\prod_{p\in S,p\neq\infty}|L_{1,p}(\mathbf{x})\cdots L_{r_p,p}(\mathbf{x})|_p$$

$$\le C.C'\,\|\mathbf{x}\|^{r_\infty-n-\varepsilon}.\,\|\mathbf{x}\|^{-(r_\infty-n)} = \widetilde{C}\,\|\mathbf{x}\|^{-\varepsilon}.$$

Por lo tanto, por el Teorema del Subespacio p-ádico, los $\mathbf{x}$ en consideración pertenecen a finitos subespacios lineales propios de $\mathbb{Q}^n$.

□

Utilizaremos esta versión del Teorema para probar la finitud de las soluciones de la ecuación de Thue-Mahler.

TEOREMA 4.7 (Mahler, 1933). *Sea $F(X,Y) \in \mathbb{Z}[X,Y]$ un polinomio homogéneo libre de cuadrados de grado $n \geq 3$ y sean $p_1, \ldots, p_s$. Entonces la ecuación*

$$|F(x,y)| = p_1^{z_1} \ldots p_s^{z_s} \quad en\ x, y, z_1, \ldots, z_s \in \mathbb{Z} \quad con\ mcd(x,y) = 1, \tag{4.13}$$

posee finitas soluciones.

En la prueba del teorema utilizaremos el siguiente lema, que es una consecuencias sencilla de la formula del producto en $\mathbb{Q}$.

LEMA 13. *Sea $u \in \mathbb{Q}$. Entonces $u = \pm p_1^{z_1} \ldots p_s^{z_s}$ para ciertos enteros $z_1, \ldots, z_s$ si y solo si $|u|.|u|_{p_1} \ldots |u|_{p_s} = 1$.*

DEMOSTRACIÓN. Si $F(1,0) \neq 0$ entonces F se factoriza de la forma $a_0 (X - \alpha_1 Y) \ldots (X - \alpha_n Y)$ con $\alpha_1, \ldots, \alpha_n$ distintos, mientras que si $F(1,0) = 0$, F se factoriza como $a_0 Y (X - \alpha_1 Y) \ldots (X - \alpha_{n-1} Y)$ con $\alpha_1, \ldots, \alpha_{n-1}$ distintos. En ambos casos, F es el producto de n formas lineales en dos variables en posición general.

Sea ε tal que $0 < \varepsilon < n - 2$. Entonces por el lema anterior, para cada solución $(x, y, z_1, \ldots, z_s)$ de (4.13) se tendrá que

$$|F(x,y)| \cdot \prod_{j=1}^{s} |F(x,y)|_{p_j} = 1 \leq \max\{|x|, |y|\}^{n-2-\varepsilon}.$$

Por teorema (subespa p adico), el conjunto de soluciones $(x,y) \in \mathbb{Z}^2$ de está desigualdad esta contenido en la unión de finitos subespacios uno dimensionales de $\mathbb{Q}^2$. Cada uno de estos subespacios contiene solo dos soluciones con $\text{mcd}(x,y) = 1$. Por lo tanto, (4.13) posee finitas soluciones. □

4. Sobre la complejidad de números algebraicos

Vimos en el capítulo anterior, como utilizar el teorema de Roth para obtener información sobre el desarrollo decimal de números algebraicos y resultados de trascendencia. En 2004, Adamczewski, Bugeaud y F.Luca [**2**], aplicaron la version p-adica del Teorema del Subespacio para obtener un criterio combinatorio de trascendencia. Luego Adamczewski y Bugeaud aplicaron este criterio para obtener información sobre la complejidad de números algebraicos.

TEOREMA 4.8. *Sea b un número entero $b \geq 2$. Sea $\alpha \in \mathbb{R}$ y sea $0, a_1 a_2 a_3 \ldots$ el desarrollo de α en base b. Supongamos que existe un número real positivo ε y infinitas 3-uplas de números enteros positivos (j,k,l) tal que*

$$a_{j+i} = a_{j+k+i}, \quad i = 1, \ldots, l, \tag{4.14}$$

y

$$l \geq \varepsilon (j+k), \quad l \leq k \tag{4.15}$$

entonces α es o bien racional o trascendente.

Lo que nos están diciendo las hipótesis del teorema es que existen bloques de tamaño l arbitrariamente grande, que ocurren dos veces muy cerca del inicio del desarrollo b-ario de α.

DEMOSTRACIÓN. Para hacer la notación más sencilla, asumamos que $\alpha \in (0,1)$. Observemos que dentro de todas las uplas (j,k,l) que cumplen (4.14), (4.15), los k's son no acotados. Si lo fueran, los l's también lo serian por (4.15) y por (4.14) los j's también, contradiciendo el hecho de que hay infinitas uplas. Sea entonces $(j_1,k_1,l_1),(j_2,k_2,l_2),\ldots$ una sucesión infinita de 3-uplas de enteros positivos que satisfacen (4.14), (4.15) y $k_1 < k_2 < \ldots$, $k_m \to \infty$.

Supondremos que α es algebraico y veamos que de hecho es racional. Definamos el número racional

$$\alpha_m = a_1a_2\ldots a_{j_m}a_{j_{m+1}}\ldots a_{j_m+k_m}a_{j_{m+1}}\ldots a_{j_m+k_m}a_{j_{m+1}}\ldots a_{j_m+k_m}\ldots$$

con pre-periodo $a_1a_2\ldots a_{j_m}$ y periodo $a_{j_{m+1}}\ldots a_{j_m+k_m}$. A partir de un cálculo directo con la serie que define la expresión b-aria de α_m, obtenemos que existe un entero p_m tal que

$$\alpha_m = \tfrac{p_m}{b^{j_m}(b^{k_m}-1)}.$$

Por (4.14) tenemos que

$$\alpha = a_1a_2\ldots a_{j_m}a_{j_{m+1}}\ldots a_{j_m+k_m}a_{j_{m+1}}\ldots a_{j_m+l_m}\ldots,$$

con lo cual α y α_m tienen, al menos los primeros $j_m+k_m+l_m$ digitos en común. Luego

$$|\alpha-\alpha_m| = |\alpha - \frac{p_m}{b^{j_m}(b^{k_m}-1)}| < \frac{1}{b^{j_m+k_m+l_m}}. \tag{4.16}$$

Multiplicando por $b^{j_m}(b^{k_m}-1)$ y usando (4.15), obtenemos que

$$|b^{j_m+k_m}\alpha - b^{j_m}\alpha - p_m| < \left(b^{j_m+k_m}\right)^{-l_m/(j_m+k_m)} \le \left(b^{j_m+k_m}\right)^{-\varepsilon}. \tag{4.17}$$

Definamos ahora la información para aplicar el teorema del subespacio. Consideremos las formas lineales linealmente independientes con coeficientes reales algebraicos:

$$L_{1,\infty}(\mathbf{X}) = X_1, \quad L_{1,\infty}(\mathbf{X}) = X_2, \quad L_{1,\infty}(\mathbf{X}) = \alpha X_1 - \alpha X_2 - X_3.$$

Sea S el conjunto de primos que dividen a b. Para cada primo $p \in S$, consideremos la formas lineales, linealmente independientes con coeficientes enteros

$$L_{1,p}(\mathbf{X}) = X_1, \quad L_{1,p}(\mathbf{X}) = X_2, \quad L_{1,p}(\mathbf{X}) = X_3$$

A partir de la fórmula del producto y (4.17) obtenemos que para $\mathbf{x} = \left(b^{j_m+k_m}, b^{j_m}, p_m\right)$,

$$\prod_{i=1}^{3}|L_{i,\infty}(\mathbf{x})|\cdot\prod_{p\in S}\prod_{i=1}^{3}|L_{i,p}(\mathbf{x})|_p \le |b^{j_m+k_m}\alpha - b^{j_m}\alpha - p_m| \le \left(b^{j_m+k_m}\right)^{-\varepsilon} \le \|x\|^{-\varepsilon}$$

Por lo tanto, por el Teorema del Subespacio, el conjunto de uplas $\left(b^{j_m+k_m}, b^{j_m}, p_m\right)$, $m \ge 1$, esta contenido en una unión finita de subespacios lineales propios de $\mathbb{Q}^3$. Luego infinitas uplas pertenecen a alguno de estos subespacios. Llamemos a ese subespacio T. Sean z_1, z_2, z_3 enteros tal que la forma lineal $z_1X_1 + z_2X_2 + z_3X_3$ se anula íntegramente en T, entonces

$$z_1b^{j_m+k_m} + z_2b^{j_m} + z_3p_m = 0,$$

para infinitos m. Dividiendo por $b^{j_m}(b^{k_m}-1)$ obtenemos que

$$z_1 \frac{b^{k_m}}{(b^{k_m}-1)} + z_2 \frac{1}{b^{k_m}-1} + z_3 \alpha_m = 0.$$

Si hacemos tender m a infinito, dado que teníamos $k_m \to \infty$ y $\alpha_m \to \alpha$ por (4.16) obtenemos

$$z_1 + z_3 \alpha = 0,$$

es decir, $\alpha \in \mathbb{Q}$ como queríamos ver. □

Sea $b \geq 2$ un número entero, entonces la expresión b-adica de cualquier número racional es eventualmente periódica. Si α es un número algebraico de grado > 1, poco se sabe sobre la expresión b-adica de dicho número. Desde los trabajos de Borel [**8**], se cree que los números algebraicos comparten la misma propiedad que casi todos los números reales. Recordemos que un número real α se dice normal en base b, si para todo entero positivo n, cualquier combinación de longitud n a partir de los dígitos $\{0, 1, \ldots, b-1\}$ ocurre en el desarrollo b-ario de α con igual frecuencia, es decir $1/b^n$. Muy poco se sabe sobre este tema, por ejemplo, ni siquiera se sabe si el digito 7 aparece infinitas veces en el desarrollo decimal de $\sqrt{2}$.

Una posible forma de medir la complejidad de la expresión de un número irracional en base b es la complejidad por bloques. Sea α un número real, $b \geq 2$ un entero, y n un entero positivo. Notaremos por $p(n, \alpha, b)$ la cantidad de bloques distintos de longitud n que aparecen en la expresión b-adica de α, es decir,

$$p(n, \alpha, b) = |\{a_k a_{k+1} \ldots a_{k+n-1} | k = 1, 2, \ldots\}|.$$

Obviamente se tiene que $1 \leq p(n, \alpha, b) \leq b^n$

En [**1**] se prueba una versión similar al siguiente lema combinatorio

LEMA 14. *Sea α un número real y $b \geq 2$ un entero. Supongamos que la función de complejidad $p(n, \alpha, b)$ cumple que*

$$\liminf_{n\to\infty} \frac{p(n, \alpha, b)}{n} < \infty.$$

Entonces existe $\varepsilon > 0$ e infinitas uplas (j, k, l) que cumplen (4.14) *y* (4.15).

Combinando esto con el resultado anterior, se obtiene el siguiente teorema

TEOREMA 4.9 (Adamczewski y Bugeaud, 2007). *Si α es un número algebraico irracional real, entonces*

$$\lim_{n\to\infty} \frac{p(n, \alpha, b)}{n} = \infty$$

De hecho el resultado es más fuerte que el recién enunciado. En el mismo trabajo, también se prueban resultado similares sobre complejidad del desarrollo de Hensel de números p-ádicos y un criterio de trascendencia para números p-ádicos.

APPENDIX A

Alturas

Recordar que un valor absoluto sobre un cuerpo K es una función real

$$|\ |: K \to [0, \infty)$$

que cumple las siguientes 3 propiedades

1: $|x| = 0 \Leftrightarrow x = 0$.
2: $|xy| = |x||y|$.
3: $|x + y| \leq |x| + |y|$.

El valor absoluto se dirá no-arquimediano si cumple la condición más fuerte

3': $|x + y| \leq max\,\{|x|, |y|\}$.

Si $3'$ falla para algún valor de x,$y \in K$ se dice que el valor absoluto es Arquimediano.

El valor absoluto usual sobre $\mathbb{Q}$ es un valor absoluto arquimediano. Para cada número primo p hay un valor absoluto no-arquimediano asociado. Para cada $x \in \mathbb{Q}$, sea $ord_p(x)$ el único entero tal que x se escribe de la forma

$$x = p^{ord_p(x)} \cdot \frac{a}{b} \quad \text{con } a, b \in \mathbb{Z} \quad \text{y } p \not| ab.$$

(Si $x = 0$, definimos $ord_p(x) = \infty$ por convención.) Se define el valor absoluto p-adico como

$$|x|_p = p^{-ord_p(x)}.$$

El conjunto de valores absolutos estándar sobre $\mathbb{Q}$ será notado por $M_{\mathbb{Q}}$, el cual consistirá en el valor absoluto arquimediano $|\ |_\infty$ y los valores absolutos p-adicos $|\ |_p$ para cada número primo p.

El conjunto de valores absolutos estándar sobre un cuerpo de números K será el conjunto M_K que consiste de los valores absolutos sobre K cuya restricción a $\mathbb{Q}$ pertenece a $M_{\mathbb{Q}}$. Notaremos por M_K^∞ al conjunto de valores absolutos arquimedianos en M_K y similarmente M_K^0 al conjunto de valores absolutos no-arquimedianos en K.

Para simplificar la notación, notaremos al valor absoluto $\upsilon \in M_K$ como $|\ |_\upsilon$.

Sea K'/K una estensión de cuerpos de números y sean $\upsilon \in M_K$, $\omega \in M_{K'}$ valores absolutos. Diremos que ω divide a υ(o que ω esta sobre υ) y notaremos $\omega|\upsilon$ si la restriccion de ω a K es υ. Para cada valor absoluto υ sobre K, notaremos K_υ a la completacion de K respecto de υ. Sobre $\mathbb{Q}$ tenemos que $\mathbb{Q}_\upsilon = \mathbb{R}$ si $\upsilon = \infty$ es el valor absoluto arquimediano y $\mathbb{Q}_\upsilon = \mathbb{Q}_p$ si υ es el valor absoluto p-adico.

Proposición 5. *Sea K'/K una estensión de cuerpos de números y sean $\upsilon \in M_K$. Entonces*

$$\sum_{\omega \in M_{K'}, \omega|\upsilon} [K'_\omega : K_\upsilon] = [K' : K]$$

Definición 6. *Sea $\upsilon \in M_K$ un valor absoluto sobre un cuerpo de números K. Se define el grado local de υ como el número*

$$n_\upsilon = [K_\upsilon : \mathbb{Q}_\upsilon].$$

El valor absoluto normalizado asociado a υ será

$$\|x\|_\upsilon = |x|_\upsilon^{n_\upsilon}.$$

PROPOSICIÓN 6. *Sea K un cuerpo de números y sea $x \in K^*$. Entonces*

$$\prod_{\upsilon \in M_K} \|x\|_\upsilon = 1.$$

DEFINICIÓN 7. *Sea $\alpha \in K$, definiremos la altura (multiplicativa) relativa a K como*

$$H_K(\alpha) = \prod_{\upsilon \in M_K} max\left\{\|x\|_\upsilon, 1\right\}.$$

La altura (logaritmica) relativa a K como

$$h_K(\alpha) = log\ H_K(\alpha)$$

DEFINICIÓN 8. *Sea $\alpha \in \overline{\mathbb{Q}}$, se define la altura absoluta (multiplicativa) de α como*

$$H(\alpha) = H_K(\alpha)^{1/[K:\mathbb{Q}]},$$

donde K es un cuerpo de números que contiene a α. La altura absoluta(logaritmica) será

$$h(\alpha) = log\ H(\alpha) = \tfrac{1}{[K:\mathbb{Q}]} h_K(\alpha).$$

DEFINICIÓN 9. *Sea $f = \sum_{i \in I} a_i x^i$ un polinomio en varias variables sobre un cuerpo de números K. Se define la norma de Gauss de f respecto al valor absoluto υ como*

$$|f|_\upsilon = \max_{i \in I} |a_i|_\upsilon$$

Definimos la altura (proyectiva) relativa a K y absoluta de f como

$$H_K(f) = \prod_{\upsilon \in M_K} |f|_\upsilon^{n_\upsilon} \quad y \quad h_K(f) = log\ H_K(f)$$

y

$$H(f) = \prod_{\upsilon \in M_K} |f|_\upsilon^{n_\upsilon/[K:\mathbb{Q}]} \quad y \quad h(f) = log\ H(f) = \frac{1}{[K:\mathbb{Q}]} \sum_{\upsilon \in M_K} n_\upsilon log\ |f|_\upsilon.$$

LEMA 15 (Lema de Gauss). *Sea $f_1, \ldots, f_r$ polinomios sobre K, entonces*

$$|f_1 \ldots f_r|_\upsilon = |f_1|_\upsilon \ldots |f_r|_\upsilon \quad \textit{para todo valor absoluto no-arquimediano } \upsilon$$

PROPOSICIÓN 7 (Desigualdad de Gelfand). *Sean $f_1, \ldots, f_r \in \mathbb{C}[X_1, \ldots, X_m]$. Entonces*

$$\textstyle\prod_{i=1}^r |f_i| \leq 2^{d_1+\ldots+d_m} |f_1 \ldots . f_r|,$$

para todo valor absoluto arquimediano, donde $gr_{X_j}(f_1 \ldots . f_r) \leq d_j$. En particular se tiene que

$$\prod_{i=1}^r H(f_i) \leq 2^{d_1+\ldots+d_m} . H(f_1 \ldots . f_r).$$

Bibliografía

[1] B. Adamczewski and Y. Bugeaud, On the complexity of algebraic numbers. I. Expansions in integer bases, Ann. of Math. (2) **165** (2007), no. 2, 547–565. MR2299740 (2008a:11130)

[2] B. Adamczewski, Y. Bugeaud and F. Luca, Sur la complexité des nombres algébriques, C. R. Math. Acad. Sci. Paris **339** (2004), no. 1, 11–14. MR2075225 (2005g:11033)

[3] A. Baker, *Transcendental number theory*, Cambridge Univ. Press, London, 1975. MR0422171 (54 #10163)

[4] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, 4, Cambridge Univ. Press, Cambridge, 2006. MR2216774 (2007a:11092)

[5] E. Bombieri and A. J. van der Poorten, Some quantitative results related to Roth's theorem, J. Austral. Math. Soc. Ser. A **45** (1988), no. 2, 233–248. MR0951583 (89i:11075)

[6] E. Bombieri, On the Thue-Siegel-Dyson theorem, Acta Math. **148** (1982), 255–296. MR0666113 (83m:10052)

[7] E. Bombieri, D. C. Hunt and A. J. van der Poorten, Determinants in the study of Thue's method and curves with prescribed singularities, Experiment. Math. **4** (1995), no. 2, 87–96. MR1377411 (97b:11092)

[8] É. Borel, Sur les chiffres décimaux de $\sqrt{2}$ et divers problèmes de probabilités en chaîne, C. R. Acad. Sci. Paris **230** (1950), 591–593. MR0034544 (11,605d)

[9] H. Davenport and W. M. Schmidt, Approximation to real numbers by quadratic irrationals, Acta Arith. **13** (1967/1968), 169–176. MR0219476 (36 #2558)

[10] H. Davenport and W. M. Schmidt, Approximation to real numbers by algebraic integers, Acta Arith. **15** (1968/1969), 393–416. MR0246822 (40 #91)

[11] F. J. Dyson, The approximation to algebraic numbers by rationals, Acta Math. **79** (1947), 225–240. MR0023854 (9,412h)

[12] B. Edixhoven and J.-H. Evertse (eds), *Diophantine approximation and abelian varieties*, Lecture Notes in Mathematics, 1566, Springer, Berlin, 1993. MR1288998 (95g:11061)

[13] S. Ferenczi and C. Mauduit, Transcendence of numbers with a low complexity expansion, J. Number Theory **67** (1997), no. 2, 146–161. MR1486494 (98m:11079)

[14] A. O. Gelfond, *Transcendental and algebraic numbers*, Translated from the first Russian edition by Leo F. Boron, Dover, New York, 1960. MR0111736 (22 #2598)

[15] M. Hindry and J. H. Silverman, *Diophantine geometry*, Graduate Texts in Mathematics, 201, Springer, New York, 2000. MR1745599 (2001e:11058)

[16] K. Mahler, *Lectures on diophantine approximations. Part I*, Prepared from the notes by R. P. Bambah of my lectures given at the University of Notre Dame in the Fall of 1957, Univ. Notre Dame Press, Notre Dame, Ind, 1961. MR0142509 (26 #78)

[17] K. Mahler, Zur Approximation algebraischer Zahlen. I, Math. Ann. **107** (1933), no. 1, 691–730. MR1512822

[18] S. Lang, *Fundamentals of Diophantine geometry*, Springer, New York, 1983. MR0715605 (85j:11005)

[19] W. J. LeVeque, *Topics in number theory. Vols. 1 and 2*, Addison-Wesley Publishing Co., Inc., Reading, MA, 1956. MR0080682 (18,283d)

[20] C. J. Parry, The p-adic generalisation of the Thue-Siegel theorem, Acta Math. **83** (1950), 1–100. MR0037875 (12,320d)

[21] D. Ridout, Rational approximations to algebraic numbers, Mathematika **4** (1957), 125–131. MR0093508 (20 #32)

[22] D. Ridout, The p-adic generalization of the Thue-Siegel-Roth theorem, Mathematika **5** (1958), 40–48. MR0097382 (20 #3851)

[23] K. F. Roth, Rational approximations to algebraic numbers, Mathematika **2** (1955), 1–20; corrigendum, 168. MR0072182 (17,242d)

[24] D. Roy, Approximation to real numbers by cubic algebraic integers. II, Ann. of Math. (2) **158** (2003), no. 3, 1081–1087. MR2031862 (2004k:11110)

[25] C. Siegel, Approximation algebraischer Zahlen, Math. Z. **10** (1921), no. 3-4, 173–213. MR1544471

[26] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, 106, Springer, New York, 1986. MR0817210 (87g:11070)

[27] H. P. Schlickewei, Die p-adische Verallgemeinerung des Satzes von Thue-Siegel-Roth-Schmidt, J. Reine Angew. Math. **288** (1976), 86–105. MR0422166 (54 #10158)

[28] H. P. Schlickewei, On products of special linear forms with algebraic coefficients, Acta Arith. **31** (1976), no. 4, 389–398. MR0429784 (55 #2794)

[29] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, 785, Springer, Berlin, 1980. MR0568710 (81j:10038)

[30] W. M. Schmidt, Simultaneous approximation to algebraic numbers by rationals, Acta Math. **125** (1970), 189–201. MR0268129 (42 #3028)

[31] T. Schneider, Über die Approximation algebraischer Zahlen, Journal für die reine und angewandte Mathematik **175** (1936): 182–192.

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

[33] P. Vojta, A refinement of Schmidt's subspace theorem, Amer. J. Math. **111** (1989), no. 3, 489–518. MR1002010 (90f:11054)

[34] M. Waldschmidt, Diophantine equations and transcendental methods(written by Noriko Hirata). In Transcendental numbers and related topics, RIMS Kokyuroku, Kyoto, **599** (1986), no. 8,82–94.

[35] E. A. Wirsing, On approximations of algebraic numbers by algebraic numbers of bounded degree, in *1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969)*, 213–247, Amer. Math. Soc., Providence, RI. MR0319929 (47 #8470)