UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA.

APROXIMACIÓN DIOFANTINA
TRASCENDENCIA Y ALGEBRAICIDAD DE NÚMEROS

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
GERARDO GONZÁLEZ ROBERT

DIRECTOR DE LA TESIS
DR. TIMOTHY GENDRON
INSTITUTO DE MATEMÁTICAS, CUERNAVACA

MÉXICO, D. F. 25 DE AGOSTO DE 2014

Aproximación Diofantina. Trascendencia y algebraicidad de números

por

Gerardo González Robert

México
2014

Un matemático me ha dicho que el placer no está en hallar la verdad, sino en buscarla.
Anna Karenina, Lev Tolstoi

Presentación

La densidad de los racionales en los reales es uno de los primeros resultados aprendidos al estudiar análisis matemático. Esta densidad significa que para toda $\varepsilon > 0$ y cualquier α real la desigualdad $|\alpha - p/q| < \varepsilon$ tiene soluciones enteras, p y $q > 0$. Por otra parte, la teoría elemental de las fracciones continuadas garantiza que dada una función, φ, de los enteros positivos a los reales positivos, exista $\alpha \in \mathbb{R}$ ($\mathbb{R}$ representa al conjunto de los números reales) tal que

$$\left|\alpha - \frac{p}{q}\right| < \varphi(q)$$

tenga una infinidad de soluciones con p, q enteros y $q > 0$ ([Kh] Teorema 22, p.35). Luego, tiene sentido preguntarse por las soluciones de

$$\left|\alpha - \frac{p}{q}\right| < \psi(q) \tag{1}$$

donde α es un real dado y ψ, una función no negativa con argumento en los enteros positivos; en particular, cuando $\eta > 0$ es fijo y $\psi(q) = q^{-\eta}$.

El problema básico de teoría de la aproximación diofantina es (1). En general, esta teoría estudia de la densidad de los racionales en los reales a través de la solubilidad en los enteros de cierto tipo de desigualdades. Uno de los resultados más importantes es el Teorema de Roth, enunciado con precisión más adelante. El Teorema de Roth dice para qué números reales α existe una cantidad infinita de soluciones de (1) con $\psi(q) = q^{-2-\delta}$ y $\delta > 0$ fijo. Es sorprendente que las condiciones sobre α impuestas por este resultado sean puramente algebraicas.

El Teorema del Subespacio de Schmidt, que es una generalización multidimensional del Teorema de Roth, es otra proposición fundamental de aproximación diofantina. El Teorema de Schmidt también habla sobre las soluciones de una desigualdad; sin embargo, la desigualdad no es precisamente (1) y la conclusión no es sobre si hay una cantidad finita o infinita de soluciones, sino sobre el total de los subespacios lineales de $\mathbb{Q}^n$ que las contienen. Es importante mencionar que ni el Teorema de Roth dice como calcular el total de soluciones, ni el Teorema de Schmidt el total de subespacios.

En 1842 Dirichlet mostró con su célebre Principio del Palomar que para cualesquiera $\alpha, Q \in \mathbb{R}$ con $Q > 1$ existen enteros p, q con $q > 0$ que satisfacen $1 \le q < Q$ y $|\alpha q - p| \le Q^{-1}$. Como una consecuencia se obtiene la siguiente caracterización de los números irracionales.

Teorema. *Un número $\alpha \in \mathbb{R}$ es irracional si y sólo si existe una infinidad de parejas de enteros primos relativos, p, q con $q > 0$, tales que*

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

Apoyándose en la definición de número racional, la cota puede debilitarse.

Teorema. *Sea α un número racional. Para cada $\delta > 0$ existe una cantidad finita de soluciones p, q de*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{1+\delta}}. \tag{3}$$

Además de racionales e irracionales, es posible clasificar a los números reales en otras dos categorías interesantes: algebraicos y trascendentes. Se recuerda que un complejo α es algebraico si satisface polinomio con coeficientes enteros; que el polinomio con coeficientes coprimos, coeficiente principal positivo y grado mínimo que sea satisfecho por α es el polinomio mínimo de α, f_α, y que el grado de α es el grado de f_α. Por ejemplo, todo racional es un algebraico de grado 1. Los reales que no son algebraicos son trascendentes.

Los números reales algebraicos forman un conjunto numerable; en consecuencia, casi todos los reales- en el sentido de Lebesgue- son trascendentes. A pesar de ello, no fue sino hasta el Siglo XIX cuando se exhibió un número real trascendente. El Teorema de Liouville sobre aproximaciones racionales establece que si α es un real algebraico de grado n, entonces existe $C = C(\alpha) > 0$ tal que $|\alpha - pq^{-1}| > Cq^{-n}$ vale para cualquier racional pq^{-1}. De esta proposición se desprende el próximo resultado.

Teorema (Liouville, 1844)**.** *Sean α un real algebraico de grado n y $\delta > 0$. Sólo hay una cantidad finita de soluciones racionales, p/q, de*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{n+\delta}}. \tag{4}$$

La prueba utiliza técnicas del cálculo diferencial aplicadas al polinomio mínimo de α, dando pie a nuevas estrategias para estudiar la trascendencia de ciertos números complejos. Matemáticos como a Dyson, Thue, Siegel, Gelfond y Klaus Roth descubrieron mejores cotas que la propuesta por Liouville con métodos más sofisticados.

Teorema (Roth, 1955)**.** *Sean α un irracional algebraico y $\delta > 0$. Entonces sólo hay una cantidad finita de soluciones racionales, p/q, de*

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

En el Teorema de Roth, a diferencia del Teorema de Liouville, la cota es independiente del grado de α. La prueba presentada por Roth, aunque es complicada y técnica, también se basa en el estudio de un polinomio. Si el grado de α es 2, el Teorema de Liouville implica el de Roth; no obstante, la implicación se invierte cuando el grado es al menos tres. Tiempo después de que Roth publicara su Teorema, Wolfgang Schmidt probó una proposición más profunda. Como es usual, se denota $\mathbb{C}$ al campo de los números complejos y $\mathbb{Q}$, a los números racionales.

Teorema (Schmidt, 1972)**.** *Sean $L_1, \ldots, L_n$, $L_i : \mathbb{C}^n \to \mathbb{C}$, formas lineales independientes con coeficientes algebraicos. Para cualquier $\delta > 0$ existe una cantidad finita de subespacios propios de $\mathbb{Q}^n$, $T_1, \ldots, T_w$, tales que cada punto con entradas enteras $\mathbf{x} \neq \mathbf{0}$ que satisfaga*

$$|L_1(\mathbf{x})L_2(\mathbf{x})\cdots L_n(\mathbf{x})| \leq \frac{1}{\|\mathbf{x}\|_\infty^\delta}. \tag{6}$$

está en alguno de estos subespacios.

La similitud entre las desigualdades (2), (3), (4) y (5) desaparece en (6), que ni siquiera tiene la forma de (1). No es inmediato ver que el Teorema de Schmidt implica el Teorema de Roth. Hay que tomar en cuenta a las transformaciones $L_1(x,y) = \alpha y - x$, $L_2(x,y) = y$. De este modo, soluciones de (5) con denominador suficientemente grande resolverán (6) y, por ende, estarán en una cantidad finita de subespacios propios de $\mathbb{Q}^2$. Tras mostrar que cada uno de estos subespacios tiene una cantidad finita de soluciones, se concluye que (5) únicamente es resuelta por una cantidad finita de racionales.

El Teorema del Subespacio de Schmidt

Wolfgang Schmidt obtuvo su Teorema del Subespacio de otra proposición, el Teorema Fuerte del Subespacio. Su enunciado requiere la noción de mínimos sucesivos. Tómese a $\mathscr{R} \subseteq \mathbb{R}^n$ compacto, convexo ($\mathbf{x}, \mathbf{y} \in \mathscr{R}$ y $\lambda \in [0,1]$ implican $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in \mathscr{R}$), simétrico en el origen ($-\mathbf{x} \in \mathscr{R}$ cada que $\mathbf{x} \in \mathscr{R}$) y con interior no vacío. Para cada $\lambda > 0$ fijo, $\lambda\mathscr{R}$ denota a la imagen de $\mathscr{R}$ bajo la transformación $\mathbf{x} \mapsto \lambda\mathbf{x}$. Si λ es muy chico, $\lambda\mathscr{R}$ carece de puntos enteros distintos del origen. Por otra parte, para $\lambda > 0$ suficientemente grande, $\lambda\mathscr{R}$ contiene a n puntos linealmente independientes con entradas enteras.

El primer mínimo sucesivo de $\mathscr{R}$, λ_1, es el ínfimo de los reales $\lambda > 0$ para los que $\lambda\mathscr{R}$ tiene un punto entero distinto del origen. En general, el j-ésimo mínimo sucesivo de $\mathscr{R}$, λ_j, es el ínfimo de los $\lambda > 0$ tales que $\lambda\mathscr{R}$ tiene j puntos enteros linealmente independientes, entonces $\lambda_1 \leq \ldots \leq \lambda_n$. Además, por la compacidad de $\mathscr{R}$, los ínfimos son mínimos. Con este concepto a la mano ya puede enunciarse el Teorema Fuerte del Subespacio.

Teorema (Teorema Fuerte del Subespacio)**.** *Sean $L_1, \ldots, L_n$ formas lineales, $L_j : \mathbb{R}^n \to \mathbb{R}$, linealmente independientes con coeficientes en $\mathbb{R} \cap \mathbb{A}$ y $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ tal que $c_1 + c_2 + \ldots + c_n = 0$. Para cualquier $Q > 0$ los reales $\lambda_1(Q), \ldots, \lambda_n(Q)$ denotan los mínimos sucesivos de*

$$\Pi(Q) := \{\mathbf{x} \in \mathbb{R}^n : \forall j \in \{1, \ldots, n\} \quad |L_j(\mathbf{x})| \leq Q^{c_j}\}$$

Si existen $\delta > 0$, $d \in [1..n-1]$ y $\mathfrak{Q} \subseteq \mathbb{R}_{>0}$ no acotado que satisfagan

$$\forall Q \in \mathfrak{Q} \qquad \lambda_d(Q) < \frac{\lambda_{d+1}(Q)}{Q^\delta};$$

entonces, hay un subespacio $S^d \subseteq \mathbb{Q}^n$, $\dim S^d = d$, y $\mathfrak{Q}_1 \subseteq \mathfrak{Q}$ no acotado tales que para cualquier $Q \in \mathfrak{Q}_1$ los primeros d mínimos sucesivos de $\Pi(Q)$ se realizan en $\mathbf{g}_1(Q), \ldots, \mathbf{g}_d(Q) \in \mathbb{Z}^n \cap S^d$.

Pasar del Teorema Fuerte del Subespacio al Teorema del Subespacio de Schmidt no es trivial. Es necesario un Lema que conecte a estas dos proposiciones. Con $\mathbf{c} \in \mathbb{R}^n$ y $L_1, \ldots, L_n$ como en el Teorema Fuerte del Subespacio y pensando cierto este resultado, se supone que las soluciones de $|L_1(\mathbf{x})| \leq \|\mathbf{x}\|_\infty^{c_1+\delta}, \ldots, |L_n(\mathbf{x})| \leq \|\mathbf{x}\|_\infty^{c_n+\delta}$ no pertenecen a una cantidad finita de subespacios propios de $\mathbb{Q}^n$. De aquí se obtiene una sucesión de soluciones tal que cualesquiera n son linealmente independientes. Con el Teorema Fuerte del Subespacio se concluye que una infinidad de términos de esta sucesión pertenece a un subespacio propio S de $\mathbb{Q}^n$, llegando a una contradicción. De este modo se obtiene el siguiente resultado.

Lema. *Sean $L_1, \ldots, L_n$ y $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ como en el Teorema Fuerte del Subespacio. Para cada $\delta > 0$ existe una cantidad finita de subespacios propios $T_1, \ldots, T_w \subseteq \mathbb{Q}^n$ tales*

que

$$\forall \mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \quad \left(\left(\forall j \in \{1 \dots n\} \quad |L_j(\mathbf{x})| \leq \|\mathbf{x}\|_\infty^{c_j - \delta} \right) \implies \mathbf{x} \in \bigcup_{k=1}^{w} T_k \right). \tag{7}$$

Para concluir El Teorema del Subespacio se supondrá que los coeficientes de $L_1, \dots, L_n$ son reales, el caso complejo sigue por inducción. La idea es reducir el problema a aplicar una cantidad finita de veces el Lema anterior.

A pesar de las complicaciones y tecnicismos, la prueba del Teorema Fuerte del Subespacio descansa sobre una proposición que sigue de las propiedades elementales de los determinantes y el álgebra exterior de $\mathbb{R}^n$.

Proposición. *Sean p, n enteros con $0 < p < n$ y $\mathbf{x}_1, \dots, \mathbf{x}_p$, $\mathbf{y}_1, \dots, \mathbf{y}_p$ dos colecciones linealmente independientes de vectores en $\mathbb{R}^n$. Entonces, $\mathbf{x}_1, \dots, \mathbf{x}_p$ y $\mathbf{y}_1, \dots, \mathbf{y}_p$ generan el mismo subespacio de $\mathbb{R}^n$ si y sólo si $\mathbf{x}_1 \wedge \dots \wedge \mathbf{x}_p$ y $\mathbf{y}_1 \wedge \dots \wedge \mathbf{y}_p$ son linealmente dependientes.*

Se lleva el problema de $\mathbb{R}^n$ a $\mathbb{R}_p^n$, con $p = n - d$, para buscar un vector fijo, $H^{(p)}$ y una colección no acotada de reales positivos, Q, para los cuales $\mathbf{g}_1(Q) \wedge \dots \wedge \mathbf{g}_n(Q)$ es paralelo a $H^{(p)}$. Aplicando la proposición anterior se concluye el resultado. La existencia de $H^{(p)}$– la parte más pesada de la demostración– se justifica con propiedades de los mínimos sucesivos y con un polinomio multivariado cuya construcción se asemeja mucho a la prueba del Teorema de Roth.

Una familia de números trascendentes

Los Teoremas de Roth y Liouville proveen a través de las fracciones continuadas formas de construir números trascendentes (cfr. [Kh] Teorema 22, p. 35). No obstante, este camino da lugar a números cuyos elementos (en la expansión como fracción continuada simple) no son acotados. Yann Bugeaud demostró que, de hecho, la expansión en fracciones continuadas de un número algebraico no puede ser *sencilla*. Para explicar esta sencillez a cada sucesión de enteros positivos, $\mathbf{a}$, se le asocia la función de complejidad, $p_{\mathbf{a}} : \mathbb{N} \to \mathbb{N}$, con la que puede enunciarse el siguiente resultado.

Teorema (Bugeaud, 2011)**.** *Sea $\mathbf{a} = (a_n)_{n=1}^{\infty}$ una sucesión de naturales que no es periódica tarde o temprano. Si el real $\alpha := [0; a_1, a_2, \dots]$ es algebraico, entonces*

$$\lim_{n \to \infty} \frac{p_{\mathbf{a}}(n)}{n} = +\infty. \tag{8}$$

El Teorema anterior sigue del Teorema del Subespacio de Schmidt y de estudiar la estructura de la expansión en fracciones continuadas. Con esta herramienta puede construirse una familia de sucesiones $(t_n)_{n=1}^{\infty}$ acotadas de enteros positivos para las que $[0; t_1, t_2, \dots]$ es trascendente.

Para fijar las ideas, considérese la sucesión $(t_n)_{n=0}^{\infty}$ en la cual t_n es la suma (mód 3) de los coeficientes de n al escribirlo en su expansión ternaria. Por ejemplo, $t_0 = 0, t_1 = 1, t_2 = 2, t_3 = 1, t_4 = 2$. El último resultado expuesto en esta tesis establece que para cualesquiera tres números enteros positivos distintos- a_0, a_1, a_2- la fracción continuada $[s_0; s_1, s_2, s_3, \dots]$ dada por $s_n = a_0$ si $t_n = 0$, $s_n = a_1$ si $t_n = 1$ y $s_n = a_2$ si t_2 representa a un número trascendente. La sucesión $(t_n)_{n=0}^{\infty}$ se denota TM_3 y, en general, cuando 3 se sustituye por un entero $m > 0$, la sequencia resultante se denota TM_m.

Teorema. *Sean $a_1, a_2, \ldots, a_m$ números positivos distintos por pares. Entonces, el número real α cuya sucesión de cocientes parciales*[1] *es la sucesión TM_m sobre el alfabeto $\{a_1, \ldots, a_m\}$ es trascendente.*

Este modesto y atractivo enunciado, a pesar de ser una aplicación del Teorema de Bugeaud, no está formulado o sugerido en la bibliografía consultada (puede encontrarse, sin embargo, en [ge]). Para probar este Teorema es necesario entrar un poco en la teoría de las sucesiones automáticas. Es importante mencionar que Martine Quéffelec había obtenido ya este resultado para el caso $m = 2$ por otros métodos ([qu]). Poco después, Adamczewski y Bugeaud dieron en [ab] una prueba del resultado de Queffélec con un argumento que también involucra al Teorema del Subespacio de Schmidt, mas su demostración deja de ser válida cuando $m > 2$.

Estructura de la tesis

La tesis consiste de tres capítulos. En el primero, se da una prueba completa del Teorema de Roth. En el segundo, se estudia a detalle la barroca demostración del Teorema del Subespacio de Schmidt. Finalmente, en el tercero se construye una familia de números trascendentes. La tesis es prácticamente autocontenida. Las pruebas que lograron escapar a estas páginas están relacionadas con el álgebra exterior de $\mathbb{R}^n$ y propiedades elementales de las fracciones continuadas.

Las fuentes principales usadas en el Teorema de Roth son [ro], [NS] y [Ca01]. La exposición Teorema de Schmidt sigue a [sch71] y [Sch]. A pesar de la gran calidad de las referencias, por momentos se reduce la claridad considerablemente. La reestructura de las ideas y la disminución de implicaciones complicadas elimina a la mayoría de los pasajes lóbregos del discurso. Con esto en mente, se pretende aportar una lectura más asequible de los Teoremas de Roth y Schmidt además de un resultado original que es consecuencia de ellos a través del Teorema de Bugeaud.

Es pertinente hacer un comentario sobre la enumeración y la notación. La mayoría de los símbolos usados está explicada en el Glosario de la Notación. Los que no están ahí son definidos en el texto y son poco usados. La enumeración de las proposiciones es de la forma **n.m**, el primer término se refiere al capítulo y el segundo dice qué número de elemento es. Las definiciones siguen su propia enumeración mientras que los teoremas, corolarios y lemas siguen otra. En los apéndices la enumeración sólo tiene un elemento.

[1] Los cocientes parciales de α son los elementos de su representación como fracción continuada simple.

Agradecimientos

Agradezco encarecidamente

- A mi familia.
- Al Dr. Tim Gendron por las sugerencias que siempre se tradujeron en mejoras invaluables en la tesis y su apoyo durante la maestría.
- A los sinodales Dr. Aubin Arroyo Camacho, Dr. Florian Luca, Dr. Alberto Verjovsky Sola y Dr. Wilson Zúñiga Galindo.
- A los profesores Dr. Javier Rosenblueth, Dr. Luis Octavio Silva y Dr. Ernesto Rosales.
- A mis amigos en las matemáticas y fuera de ellas.
- A CONACYT.

Índice general

Capítulo 1

El Teorema de Roth

1.1. Introducción

Los números trascendentes forman una abrumadora mayoría en los reales en términos de numerabilidad, también en medibilidad de Lebesgue. No obstante, fue hasta 1844 cuando se probó, gracias a Joseph Liouville, su existencia y se construyó una infinidad de ellos: si $k \geq 2$ es un natural, entonces $[0; k^{1!}, k^{2!}, k^{3!}, \ldots]$ es trascendente (cf. [HW], Sección 11.8).

Teorema 1.1 (Liouville, 1844)**.** *Para cada irracional algebraico α de grado n existe $c = c(\alpha) > 0$ tal que para todos p, q enteros con $q > 0$,*

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

Demostración. Tómese $P \in \mathbb{Z}[x]$ un polinomio de grado mínimo que sea satisfecho por α. Supóngase, primero, que

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

Usando la expansión de Taylor y que $P(p/q) \neq 0$ es una fracción con denominador q^n,

$$\frac{1}{q^n} \leq \left|P\left(\frac{p}{q}\right)\right| = \left|\sum_{j=1}^{n} \left(\alpha - \frac{p}{q}\right)^j \frac{P^{(j)}(\alpha)}{j!}\right| \leq \frac{1}{\tilde{c}(\alpha)} \left|\alpha - \frac{p}{q}\right|.$$

Juntando los dos extremos,

$$\frac{\tilde{c}(\alpha)}{q^n} \leq \left|\alpha - \frac{p}{q}\right|.$$

Por otra parte, si

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

trivialmente se cumple

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

Con $c(\alpha) = \min\{\tilde{c}(\alpha), 1\}$ se tiene la desigualdad para todos los racionales. □

Piénsese por un momento que para un algebraico α de grado n y $\mu > 0$ existe una infinidad de racionales, p/q, tales que

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

Considérese a $P \in \mathbb{Z}[x]$ de grado r tal que α sea una raíz de multiplicidad i y p/q, una solución de la desigualdad anterior. Aplicando el Teorema de Taylor, existe $c = c(\alpha) > 0$ tal que

$$\left|P\left(\frac{p}{q}\right)\right| = \left|\sum_{j=i}^{n}\left(\frac{p}{q} - \alpha\right)^{j} \frac{P^{(j)}(\alpha)}{j!}\right| \leq \frac{c}{q^{\mu i}}.$$

Como P se anula en una cantidad finita de racionales, una infinidad de racionales cumple con

$$\frac{1}{q^{r}} \leq \left|P\left(\frac{p}{q}\right)\right| \leq \frac{c}{q^{\mu i}}.$$

Entonces, para una infinitud de naturales, q, vale $q^{\mu i - r} \leq c$, que implica

$$\mu i - r \leq 0 \implies \mu \leq \frac{r}{i} = \left(\frac{i}{r}\right)^{-1}.$$

El número

$$\operatorname{ind}_r(P;\alpha) := \frac{i}{r},$$

llamado índice de P con respecto a r en α, jugará un papel fundamental en la prueba del Teorema de Roth.

Para acotar a μ se busca maximizar el índice sobre todos los polinomios $P \in \mathbb{Z}[x]$ que satisfagan $P(\alpha) = 0$. La irreducibilidad del polinomio mínimo de α, $f \in \mathbb{Z}[x]$, propicia $f^{i}(x) \mid P(x)$, implicando $ni \leq r$, que es

$$\operatorname{ind}_r(P;\alpha) \leq \operatorname{ind}_n(f;\alpha) = \frac{1}{n},$$

La igualdad se cumple si y sólo si P es una potencia del polinomio mínimo de α. Cuando $\operatorname{ind}_r(P;\alpha)$ asume el mayor valor posible, se concluye $\mu \leq n$, que no mejora la cota de Liouville. Hay que considerar un cambio de estrategia. Se adoptará la usada por Roth, entre otros: trabajar con el índice de polinomios en múltiples variables.

El mundo fue testigo de varios intentos para mejorar la cota de Liouville. Algunos de los resultados más notables fueron obtenidos por Thue (1908), Siegel (1921), Dyson (1947) y Gelfond (1952). El trabajo de Thue establece que si α es un irracional algebraico de grado n y

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^{\kappa}}$$

tiene una infinidad de soluciones racionales h/q, entonces $\kappa \leq 2^{-1}n + 1$; Siegel –quien sugirió el uso de polinomios en dos variables– concluyó $\kappa \leq s + n(s+1)^{-1}$ para $s = 1, 2, \ldots, n-1$ y tanto Dyson como Gelfond probaron, independientemente, que $\kappa \leq \sqrt{2n}$. A pesar de que todas estas cotas mejoran la obtenida por Liouville, siguen dependiendo del grado de α. Siegel conjeturó que, de hecho, $\kappa \leq 2$ y fue Klaus Roth quien lo demostró ganando una Medalla Fields.

Teorema 1.2 (Roth[1], 1955)**.** *Sean α un número algebraico no racional. Si la siguiente desigualdad tiene una infinidad de soluciones con enteros h y q ($q > 0$), entonces $\kappa \leq 2$:*

$$\left|\alpha - \frac{h}{q}\right| < \frac{1}{q^{\kappa}}. \tag{1.1}$$

Piénsese que α es un irracional cuadrático. El Teorema de Liouville garantiza la existencia de $c = c(\alpha) > 0$ tal que $|\alpha - pq^{-1}| > cq^{-2}$ para cualquier racional pq^{-1}. Sea $\delta > 0$. Supóngase que $|\alpha - pq^{-1}| < q^{-2-\delta}$ tiene una infinidad de soluciones: $(p_j/q_j)_{j=1}^{\infty}$. Claramente, $q_j \to \infty$ cuando $j \to \infty$, de donde

$$0 < \frac{c}{q_j^2} < \left|\alpha - \frac{p_j}{q_j}\right| < \frac{1}{q_j^2}\frac{1}{q_j^{\delta}} \implies 0 < c < \frac{1}{q_j^{\delta}} \to 0 \text{ cuando } j \to \infty \implies c = 0.$$

La última contradicción revela que si $n = 2$, el Teorema de Liouville conduce al Teorema de Roth. Sin embargo, la supuesta trivialidad del resultado de Roth desaparece cuando α es un real algebraico de grado mayor o igual a tres. Supóngase que es cierto el Teorema de Roth y que $n \geq 3 > \kappa > 2$. Entonces, (1.1) tiene una cantidad finita de soluciones, $\{p_j/q_j\}_{j=1}^{m}$. Se define

$$\forall j \in \{1, \ldots, m\} \quad \eta_j := \alpha - \frac{p_j}{q_j}, \quad \eta := \min_j \{|\eta_j|\} < 1.$$

Nótese que $\eta > 0$, pues α es irracional. De este modo, se cumple que

$$\begin{aligned} \forall\ j \in \{1, \ldots, m\} \qquad & \left|\alpha - \frac{p_j}{q_j}\right| = |\eta_j| \geq \eta > \frac{\eta}{q^{\kappa}} > \frac{\eta}{q^{n}}, \\ \forall\ \frac{p}{q} \in \mathbb{Q} \setminus \left\{\frac{p_k}{q_k}\right\}_{k=1}^{m} \qquad & \left|\alpha - \frac{p}{q}\right| \geq \frac{1}{q^{\kappa}} > \frac{1}{q^{n}} > \frac{\eta}{q^{n}}. \end{aligned}$$

Estableciendo $c := \eta$ se concluye el Teorema de Liouville.

Por otro lado, si $\alpha := [\alpha_0; \alpha_1, \alpha_2, \ldots]$ es cualquier irracional y $(p_j/q_j)_{j=0}^{\infty}$, la sucesión de sus convergentes, entonces

$$\forall j \in \mathbb{N}_0 \qquad \left|\alpha - \frac{p_j}{q_j}\right| < \frac{1}{q_j^2}$$

mostrando que la cota obtenida por Roth es la mejor posible. En otras palabras, no se puede reemplazar el dos del Teorema de Roth por un número menor.

1.2. Teorema de Roth

La demostración del Teorema de Roth es tan larga como ingeniosa. La prueba del Teorema de Liouville se desarrolla con el polinomio mínimo de α. En el Teorema de Roth se utiliza, en su lugar, polinomios multivariados y se considerará su índice en un punto con respecto a una colección finita de enteros positivos (cf. Definición 1.3). En esta sección se prueba el Teorema de Roth, aunque se posterga la demostración de tres lemas importantes.

[1] Ésta es la formulación original de Roth, es la contrapuesta de la redacción usual.

Bastará con verificar el resultado para enteros algebraicos. Supóngase por un momento que el Teorema de Roth es cierto para enteros algebraicos. Sean[2] $\alpha \in \mathbb{A} \cap \mathbb{R}$, $f(x) = \sum_{i=1}^{n} a_i x^i \in \mathbb{Z}[x]$ su polinomio mínimo y $\beta := a_n\alpha$. Se tiene que

$$\begin{aligned} 0 &= a_n^{n-1} f(\alpha) = (a_n\alpha)^n + a_{n-1}(a_n\alpha)^{n-1} + \ldots + a_n^{n-2}a_1(a_n\alpha) + a_n^{n-1}a_0 \\ &= \beta^n + a_{n-1}\beta^{n-1} + \ldots + a_n^{n-2}a_1\beta + a_n^{n-1}a_0 = g(\beta). \end{aligned}$$

Por la igualdad anterior, β es un entero algebraico (g no es forzosamente irreducible, pero si se escinde en $\mathbb{Q}[x]$, lo hará en $\mathbb{Z}[x]$ y los dos factores son mónicos). Si p/q es una solución de (1.1), satisface

$$\left|\beta - \frac{a_n p}{q}\right| < \frac{|a_n|}{q^\kappa}.$$

Si (1.1) tiene una infinidad de soluciones, $(p_j/q_j)_{n=1}^{\infty}$, entonces $q_j \to \infty$ cuando $j \to \infty$; en consecuencia,

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \left(m \geq N \implies \frac{|a_n|}{q_m^\varepsilon} < 1 \right).$$

Fíjese $\varepsilon > 0$ y para cada $m > N$, $p'_m := a_n p_m$; entonces,

$$\left|\beta - \frac{p'_m}{q_m}\right| < \frac{|a_n|}{q_m^\kappa} = \frac{|a_n|}{q_m^\varepsilon}\frac{1}{q_m^{\kappa-\varepsilon}} < \frac{1}{q_m^{\kappa-\varepsilon}}.$$

Ya que se supuso válido el Teorema de Roth para enteros algebraicos, $\kappa - \varepsilon \leq 2$ y, recordando la abritrariedad de ε, $\kappa \leq 2$. Por lo tanto, sólo hay que probar el resultado para enteros algebraicos.

La demostración del Teorema de Roth tiene la cualidad de no utilizar conceptos muy sofisticados. Los puntos clave del argumento están en los Teoremas 1.4, 1.5 y 1.6, enunciados más adelante. Los lemas que los anteceden ayudan, en la mayoría de las veces, a ver qué ideas yacen detrás de cada una de estas tres proposiciones.

A grandes rasgos, el teorema de Roth se prueba por contradicción. Suponiendo que hay una infinidad de soluciones se fija una cantidad finita de ellas con ciertas propiedades: $p_1/q_1, \ldots, p_m/q_m$. Se construye un polinomio $R \in \mathbb{Z}[x_1, \ldots, x_m]$ y a cada punto en $\mathbb{R}^m$ se le asocia un número llamado índice que dice *qué tanto se anula* R en ese punto. El índice de R en $(p_1/q_1, \ldots, p_m/q_m)$, un punto racional cercano a $(\alpha, \ldots, \alpha)$, estará acotado inferior y superiormente. Finalmente, las cotas conducirán a una contradicción.

Además del índice, Roth utiliza wronskianos generalizados que, salvo una constante multiplicativa, incluyen al wronskiano usual como un caso particular. Esta herramienta dará un criterio para independencia lineal de polinomios multivariados que permitirá probar el Teorema 1.6, el corazón del argumento.

El esquema de la prueba es el siguiente:

I. Suponiendo $\alpha \in \mathbb{A}$ y que hay una infinidad de soluciones de (1.1), se construye con ayuda del Lema de Siegel (Capítulo III, Lema 2.2) un polinomio $R \in \mathbb{Z}[x_1, \ldots, x_m]$ que satisface ciertas propiedades.

II. Se acota inferiormente el índice de R en puntos racionales cercanos a $\boldsymbol{\alpha} := (\alpha, \ldots, \alpha)$. En este punto el argumento sigue el espíritu de la geometría de números.

[2] Se recuerda que el campo de los números algebraicos se denota $\mathbb{A}$.

III. Se acota superiormente el índice de R en puntos racionales cercanos a $\boldsymbol{\alpha}$ (Definición 1.3). Para lograrlo se recurre a los Wronskianos generalizados.

IV. Las cotas del índice conducirán a una contradicción.

La exposición presente se basa, principalmente, en la dada en [Ca01] y [NS], aunque hay elementos de [Sch] y [ro]. Las demostraciones de algunas proposiciones auxiliares difieren de las que pueden encontrarse en las fuentes.

Todos los símbolos cuyo significado no esté aclarado en el texto se explican en el Glosario de Notación. Hay algunas letras cuyo valor será constante a lo largo de esta sección: α será un entero algebraico, f su polinomio mínimo, $n = \operatorname{grad}(f)$, $a := \mathcal{H}(f)$ (cf. Definición 1.1) y $\boldsymbol{\alpha} := (\alpha, \alpha, \ldots, \alpha)$.

Definición 1.1. *Sea* $R(\mathbf{x}) = \sum_{\mathbf{j}} C(\mathbf{j})\mathbf{x}^{\mathbf{j}} \in \mathbb{R}[x_1, \ldots, x_N]$. *La* ***altura*** *de* R, $\mathcal{H}(R)$, *es el máximo de los valores absolutos de los coeficientes de* R; *es decir,*

$$\mathcal{H}(R) := \max\left\{|C(\mathbf{j})| : \mathbf{j} \in \mathbb{N}^N\right\}.$$

1.2.1. Índices de polinomios

A pesar de que se usarán herramientas elementales, es necesario establecer algunas definiciones. En adelante, al hablar de operadores diferenciales se estará pensando en los siguientes.

Definición 1.2. *Para cualquier multi-índice,* $\mathbf{i}$, *se define el operador*

$$(\cdot)_{\mathbf{i}} : \mathbb{R}[x_1, \ldots, x_m] \to \mathbb{R}[x_1, \ldots, x_m], \quad R_{\mathbf{i}}(\mathbf{x}) := \frac{1}{i_1! \ldots i_m!} \frac{\partial^{\mathbf{i}} R}{\partial \mathbf{x}^{\mathbf{i}}}.$$

El ***orden*** *del operador es*[3] $|\mathbf{i}|$.

En la discusión posterior al Teorema de Liouville se observó que el cociente $\operatorname{ind}_r(P; \alpha) = i/r$ (la notación de ese contexto) señala qué tanto se anula P en α. Se generaliza esa noción.

Definición 1.3. *Sean* $R \in \mathbb{Q}[x_1, x_2, \ldots, x_m]$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m) \in \mathbb{R}^m$ *y* $\mathbf{r} \in \mathbb{N}^m$. *El índice de* R *en* $\boldsymbol{\beta}$ *con respecto a* $\mathbf{r}$ *es*

$$\operatorname{ind}_{\mathbf{r}}(R; \boldsymbol{\beta}) = \inf\left\{\mathbf{i} \cdot \mathbf{r}^{-1} : R_{\mathbf{i}}(\boldsymbol{\beta}) \neq 0\right\}.$$

Se adopta la convención $\inf \varnothing = +\infty$.

Cuando el polinomio es univariado la definición anterior se reduce a la dada en la introducción. Además, es claro que el ínfimo es un mínimo excepto cuando el polinomio es idénticamente cero. Nótese que cuando $\mathbf{r} = (r_1, \ldots, r_m)$ y $\mathbf{s} = (r_1, \ldots, r_{m-1})$ son dos multi-índices, $\boldsymbol{\alpha} = (\alpha, \ldots, \alpha) \in \mathbb{R}^m$, $\tilde{\boldsymbol{\alpha}} = (\alpha, \ldots, \alpha) \in \mathbb{R}^{m-1}$ y $R \in \mathbb{R}[x_1, \ldots, x_{m-1}] \subseteq \mathbb{R}[x_1, \ldots, x_{m-1}, x_m]$, se cumple

$$\operatorname{ind}_{\mathbf{r}}(R; \boldsymbol{\alpha}) = \operatorname{ind}_{\mathbf{s}}(R; \tilde{\boldsymbol{\alpha}}).$$

Lema 1.3. *Sean* $R, P \in \mathbb{R}[x_1, \ldots, x_m]$, $\mathbf{r}$ *e* $\mathbf{i}$ *multi-índices con* $\mathbf{r} > \mathbf{0}$ *y* $\boldsymbol{\beta} \in \mathbb{R}^n$, *entonces*

i. $\operatorname{ind}_{\mathbf{r}}(R_{\mathbf{i}}; \boldsymbol{\beta}) \geq \operatorname{ind}_{\mathbf{r}}(R; \boldsymbol{\beta}) - \mathbf{i} \cdot \mathbf{r}^{-1}$,

[3] Ver Glosario de Notación.

ii. $\operatorname{ind}_{\mathbf{r}}(R+P;\boldsymbol{\beta}) \geq \min\{\operatorname{ind}_{\mathbf{r}}(R;\boldsymbol{\beta}), \operatorname{ind}_{\mathbf{r}}(P;\boldsymbol{\beta})\}$,

iii. $\operatorname{ind}_{\mathbf{r}}(RP;\boldsymbol{\beta}) = \operatorname{ind}_{\mathbf{r}}(R;\boldsymbol{\beta}) + \operatorname{ind}_{\mathbf{r}}(P;\boldsymbol{\beta})$.

Demostración. i. Sea $\mathbf{j}$ tal que $(R_{\mathbf{i}})_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0$. Tras definir $\mathbf{s} = \mathbf{i} + \mathbf{j}$,

$$R_{\mathbf{s}}(\boldsymbol{\beta}) = c_{\mathbf{s}} \frac{\partial^{\mathbf{s}} R}{\partial \mathbf{x}^{\mathbf{s}}}(\boldsymbol{\beta}) = \frac{c_{\mathbf{s}}}{c_{\mathbf{i}}} \frac{\partial^{\mathbf{j}}}{\partial \mathbf{x}^{\mathbf{j}}} \left(c_{\mathbf{i}} \frac{\partial^{\mathbf{i}} R}{\partial \mathbf{x}^{\mathbf{i}}} \right)(\boldsymbol{\beta}) = \frac{c_{\mathbf{r}}}{c_{\mathbf{i}} c_{\mathbf{j}}} c_{\mathbf{j}} \frac{\partial^{\mathbf{j}}}{\partial \mathbf{x}^{\mathbf{j}}} R_{\mathbf{i}}(\boldsymbol{\beta}) = c(R_{\mathbf{i}})_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0.$$

con $c_{\mathbf{j}}, c_{\mathbf{i}}, c_{\mathbf{s}}, c$ constantes distintas de cero. Esto prueba que

$$\left\{ \mathbf{j} \cdot \mathbf{r}^{-1} : (R_{\mathbf{i}})_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0 \right\} \subseteq \left\{ \mathbf{j} \cdot \mathbf{r}^{-1} - \mathbf{i} \cdot \mathbf{s}^{-1} : R_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0 \right\}.$$

Sacando ínfimo respecto a $\mathbf{j}$ se obtiene el resultado.

ii. Sea $r = r_1 \cdots r_m$. Para cualquier $S \in \mathbb{Z}[x_1, \ldots, x_m]$ se cumple, tomando la expansión de Taylor alrededor de $\boldsymbol{\beta}$,

$$S(\beta_1 + t^{\frac{r}{r_1}} y_1, \ldots, \beta_1 + t^{\frac{r}{r_m}} y_m) = \sum_{\mathbf{j}} S_{\mathbf{j}}(\boldsymbol{\beta}) t^{r(\mathbf{j}\cdot\mathbf{r}^{-1})} \mathbf{y}^{\mathbf{j}}.$$

El menor exponente de t es justamente $r \operatorname{ind}_{\mathbf{r}}(S;\boldsymbol{\beta})$. La linealidad de los operadores diferenciales da el segundo inciso[4].

iii. Usando la escritura del punto anterior en P y Q, $\operatorname{ind}_{\mathbf{r}}(P;\boldsymbol{\beta})$ es el exponente mínimo de t en P y lo mismo para Q. Igualmente, el exponente mínimo de t en PQ es, por un lado, $\operatorname{ind}_{\mathbf{r}}(PQ;\boldsymbol{\beta})$ y, por otro, $\operatorname{ind}_{\mathbf{r}}(P;\boldsymbol{\beta}) + \operatorname{ind}_{\mathbf{r}}(Q;\boldsymbol{\beta})$.

□

1.2.2. Resultados principales

En adelante, f denotará al polinomio mínimo de α; n, a su grado y a, a su altura. Si $P \in \mathbb{Z}[x_1, \ldots, x_m]$, se define al vector $\operatorname{grad} P \in \mathbb{Z}^m$ como aquél que tiene en la j-ésima entrada el grado de P con respecto a x_j.

El Teorema 1.4 da la existencia del polinomio auxiliar.

Teorema 1.4. *Sean $\varepsilon > 0$, $m \in \mathbb{N}$ con $m > 8n^2\varepsilon^{-2}$, $\mathbf{r}$ un multi-índice. Entonces existe $0 \neq R \in \mathbb{Z}[x_1, \ldots, x_m]$ con $\operatorname{grad} R \leq \mathbf{r}$ tal que*

i. $\operatorname{ind}_{\mathbf{r}}(R;\boldsymbol{\alpha}) \geq m(1-\varepsilon)2^{-1}$,

ii. $\mathcal{H}(R) \leq (4a+4)^{|\mathbf{r}|}$.

Sean δ, ε tales que $0 < \delta < 12^{-1}$ y $0 < \varepsilon < 20^{-1}\delta$. Suponiendo que hay una infinidad de soluciones de (1.10), puede tomarse una cantidad finita de ellas en expresión mínima–con numeradores $(p_k)_{k=1}^m$ y denominadores positivos $(q_k)_{k=1}^m$– que satisfagan

$$\forall k \in [1..m] \quad \eta_k := \alpha - \frac{p_k}{q_k}, \quad |\eta_k| < \frac{1}{q_k^{2+\delta}} \tag{1.2}$$

$$\forall k \in [1..m] \quad 64(a+1) \max\{1, |\alpha|\} < q_k^{\varepsilon}. \tag{1.3}$$

[4] Una prueba más corta se obtiene de la contención

$$\{\mathbf{j} \in \mathbb{N}^m : (R+P)_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0\} \subseteq \{\mathbf{j} \in \mathbb{N}^m : (R)_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0\} \cup \{\mathbf{j} \in \mathbb{N}^m : (P)_{\mathbf{j}}(\boldsymbol{\beta}) \neq 0\}.$$

Teorema 1.5. *Sean* $\mathbf{p} = (p_1, \dots, p_m)$, $\mathbf{q} = (q_1, \dots, q_m)$ *como en el párrafo anterior y* $\mathbf{r}$ *un multi-índice tal que*

$$\forall k \in [1..m] \quad r_1 \log q_1 \leq r_k \log q_k \leq (1+\varepsilon) r_1 \log q_1. \tag{1.4}$$

Entonces, si R *es el polinomio obtenido en el Teorema 1.4,*

$$\operatorname{ind}_{\mathbf{r}}\left(R; \frac{\mathbf{p}}{\mathbf{q}}\right) \geq \frac{\delta m}{8}. \tag{1.5}$$

El Teorema 1.5 da una cota inferior el índice de R con respecto a $\mathbf{r}$ en $\mathbf{p}/\mathbf{q}$, mientras que el el Teorema 1.6 da una superior.

Sean, ahora, $\varepsilon \in \mathbb{R}$ tal que

$$0 < \varepsilon < \frac{1}{12} \tag{1.6}$$

y $\tilde{\omega} : \mathbb{N} \times (0, 12^{-1}) \to \mathbb{R}$ la función dada por

$$\tilde{\omega}(w, \eta) = \frac{24}{2^w} \left(\frac{\eta}{12}\right)^{2^{w-1}}.$$

Para $m \in \mathbb{N}$ fijo se define

$$\omega := \tilde{\omega}(m, \varepsilon) = \frac{24}{2^m} \left(\frac{\varepsilon}{12}\right)^{2^{m-1}}, \tag{1.7}$$

que es a lo más 1 cuando $m \geq 2$.

Teorema 1.6 (Lema fundamental de Roth, 1955)**.** *Sea* $\mathbf{r} = (r_1, \dots, r_m)$ *un multi-índice tal que*

$$\forall j \in [1..m-1] \quad \omega r_j \geq r_{j+1}. \tag{1.8}$$

Considérese a $\gamma \in (0, 1]$ *y a* m *parejas de enteros coprimos,* q_j, p_j*, con* $q_j > 0$ *tales que*

$$\forall j \in [1..m] \quad q_j^{r_j} \geq q_1^{\gamma r_1}, \quad q_j^{\omega\gamma} \geq 2^{3m}. \tag{1.9}$$

Si $0 \neq S \in \mathbb{Z}[x_1, \dots, x_m]$, $\operatorname{grad} S \leq \mathbf{r}$*, cumple con* $\mathcal{H}(S) \leq q_1^{\omega r_1 \gamma}$*, entonces*

$$\operatorname{ind}_{\mathbf{r}}\left(S; \frac{\mathbf{p}}{\mathbf{q}}\right) \leq \varepsilon.$$

1.2.3. Demostración del Teorema de Roth

Al principio de esta sección se presentó la formulación del Teorema de Roth dada por su autor en [ro]. No obstante, es más común encontrarlo actualmente con otra redacción, la equivalencia es obvia.

Teorema 1.7 (Roth, 1955)**.** *Sean* α *un irracional algebraico y* $\delta > 0$*. Existe sólo una cantidad finita de enteros* p, q *con* $q > 0$ *tales que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}. \tag{1.10}$$

Demostración. Sea $0 < \delta < 12^{-1}$. A continuación, se elige una colección de valores para satisfacer las hipótesis de los Teoremas 1.4, 1.5 y 1.6.

i. Sea ε tal que $0 < \varepsilon < \delta/20$, entonces $\varepsilon < 1/12$.

ii. m es un entero tal que $m > 8n^2\varepsilon^{-1}$ y $\delta > 8\varepsilon m^{-1}$, ω está dada por (1.7). Nótese que

$$\frac{m\delta}{8} > \varepsilon.$$

iii. (p_1, q_1) es una solución con q_1 tan grande que satisface (1.3), la segunda condición en (1.9) y

$$q_1^{\omega} > (4a+4)^m.$$

iv. Se elige m soluciones de (1.10), (p_j, q_j), de modo que

$$\forall j \in [1..m] \quad \frac{1}{2}\omega \log q_{j+1} > \log q_j.$$

Esto garantiza $q_m > q_{m-1} > \ldots > q_1$, por lo que la segunda condición en (1.9) y (1.3) serán satisfechas.

v. Se toma $r_1 \in \mathbb{N}$ tal que

$$\varepsilon r_1 \log q_1 \geq \log q_m.$$

vi. Se define[5]

$$\forall j \in [2..m] \quad r_j := \left\lfloor \frac{r_1 \log q_1}{\log q_j} \right\rfloor + 1.$$

Entonces, por vi.,

$$\forall j \in [2..m] \quad r_1 \log q_1 = r_1 \frac{\log q_1}{\log q_j} \log q_j \leq \left(\left\lfloor \frac{r_1 \log q_1}{\log q_j} \right\rfloor + 1 \right) \log q_j = r_j \log q_j,$$

que es la primera condición de (1.9) con $\gamma = 1$; además, v. y la desigualdad anterior dan

$$\forall j \in [2..m] \quad r_1 \log q_1 \leq r_j \log q_j = \left\lfloor \frac{r_1 \log q_1}{\log q_j} \right\rfloor \log q_j + \log q_j \leq r_1 \log q_1 + \log q_j \leq (1+\varepsilon) r_1 \log q_1.$$

que es (1.4). La expresión anterior se puede reescribir como

$$\forall j \in [1..m] \quad \frac{r_1 \log q_1}{\log q_j} \leq r_j \leq (1+\varepsilon) \frac{r_1 \log q_1}{\log q_j}. \tag{1.11}$$

En consecuencia, usando iv., se obtiene (1.8):

$$\omega r_j > 2 \frac{\log q_j}{\log q_{j+1}} r_j = 2 \frac{r_1 \log q_1}{\log q_{j+1}} \left(\frac{r_1 \log q_1}{r_j \log q_j} \right)^{-1} \geq 2 \left(\frac{r_{j+1}}{1+\varepsilon} \right) \cdot 1 > r_{j+1}.$$

En la segunda desigualdad se acota el primer factor con el lado derecho de (1.11) y el segundo, con el lado izquierdo dividiendo entre r_j. Además, ya que $0 < \omega < 1$, se cumple $r_1 > \ldots > r_n$. Luego, si R el polinomio obtenido por el Teorema 1.4 con estos parámetros, iii. conlleva

$$\mathcal{H}(R) \leq (4a+4)^{\mathbf{r}} \leq (4a+4)^{mr_1} < q_1^{\omega r_1}.$$

Esto asegura que se puede aplicar el Teorema 1.6 al polinomio R con $\gamma = 1$. Usando los Teoremas 1.5 y 1.6 en R y $\mathbf{p}/\mathbf{q}$,

$$\frac{\delta m}{8} \leq \operatorname{ind}_{\mathbf{r}} \left(R; \frac{\mathbf{p}}{\mathbf{q}} \right) \leq \varepsilon \implies 0 < \delta \leq \frac{8\varepsilon}{m}.$$

La desigualdad anterior contradice la elección de m en ii.; por lo tanto, (1.10) tiene sólo una cantidad finita de soluciones. □

[5] La función $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{Z}$ está dada por $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$ para toda $x \in \mathbb{R}$.

1.3. Demostración de los resultados principales

1.3.1. Lemas preliminares

El Teorema de Roth, como antes el de Liouville, da propiedades algebraicas de los números reales mediante aproximaciones racionales. Podría esperarse que la prueba requiriera resultados sobre extensiones de campos. Sin embargo, en este tenor sólo se necesitará una proposición terrenal debido a Gauss.

Recuérdese que si A es un dominio de factorización única y $f \in A[x]$, entonces f es **primitivo** cuando cualquier máximo común divisor de los coeficientes de f es una unidad de A.

Lema 1.8 (Gauss)**.** *Sean A un dominio de factorización única, K su campo de fracciones y $h \in A[x]$ primitivo. Si h se factoriza en $K[x]$, entonces h se factoriza en $A[x]$.*

Demostración. Ver [La01] p.181. □

Lema 1.9. *Si $R \in \mathbb{Z}[x_1, \ldots, x_m]$, entonces $R_{\mathbf{i}} \in \mathbb{Z}[x_1, \ldots, x_m]$. Si R tiene grado r_j en x_j, entonces $R_{\mathbf{i}}$ tiene grado a lo más $r_k - i_k$ para toda $k \in [1..m]$. Además,*

$$\mathcal{H}(R_{\mathbf{i}}) \leq 2^{|\mathbf{r}|}\,\mathcal{H}(R).$$

Demostración. Sean $\mathbf{i}, \mathbf{j} \in \mathbb{Z}^m$. Entonces,

$$\begin{aligned}\frac{\partial^{\mathbf{i}}\mathbf{x}^{\mathbf{j}}}{\partial \mathbf{x}^{\mathbf{i}}} &= \prod_{k_1=1}^{i_1-1}(j_1-k_1)\cdots\prod_{k_m=1}^{i_m-1}(j_m-k_m)x_1^{j_1-i_1}\cdots x_m^{j_m-i_m}\\ &= i_1!\binom{j_1}{i_1}\cdots i_m!\binom{j_m}{i_m}x_1^{j_1-i_1}\cdots x_m^{j_m-i_m}\\ &= i_1!\cdots i_m!\binom{\mathbf{j}}{\mathbf{i}}\mathbf{x}^{\mathbf{j}-\mathbf{i}} \implies (\mathbf{x}^{\mathbf{j}})_{\mathbf{i}} = \binom{\mathbf{j}}{\mathbf{i}}\mathbf{x}^{\mathbf{j}-\mathbf{i}}.\end{aligned}$$

Como el operador $(\cdot)_{\mathbf{i}}$ es lineal,

$$R(\mathbf{x}) = \sum_{\mathbf{j}} C(\mathbf{j})\mathbf{x}^{\mathbf{j}} \implies R_{\mathbf{i}}(\mathbf{x}) = \sum_{\mathbf{j}} C(\mathbf{j})\binom{\mathbf{j}}{\mathbf{i}}\mathbf{x}^{\mathbf{j}-\mathbf{i}}.$$

Por lo que $R_{\mathbf{i}} \in \mathbb{Z}[x_1, \ldots, x_m]$. En consecuencia, recordando[6] que $\binom{m}{n} \leq 2^m$ para cualesquiera $m, n \in \mathbb{N}_0$,

$$\mathcal{H}(R_{\mathbf{i}}) = \max_{\mathbf{j}}\left\{|C(\mathbf{j})|\binom{j_1}{i_1}\cdots\binom{j_m}{i_m}\right\} \leq \max_{\mathbf{j}}\left\{|C(\mathbf{j})|\,2^{|\mathbf{j}|}\right\} \leq \max_{\mathbf{j}}\left\{|C(\mathbf{j})|\,2^{|\mathbf{r}|}\right\} = \mathcal{H}(R)2^{|\mathbf{r}|}$$

□

Lema 1.10. *Sean $r_1, \ldots, r_m$ enteros positivos, $\mathbf{r} = (r_1, \ldots, r_m)$ y $\lambda > 0$, entonces*

$$I_m(\lambda, \mathbf{r}) := \left|\left\{\mathbf{i} \in \mathbb{N}_0^m : \mathbf{i}\cdot\mathbf{r}^{-1} \leq \frac{m-\lambda}{2},\ \mathbf{0} \leq \mathbf{i} \leq \mathbf{r}\right\}\right| \leq \frac{\sqrt{2m}}{\lambda}\prod_{k=1}^{m}(r_k+1)$$

[6] Por el Teorema del Binomio $2^n = (1+1)^n = \sum_{j=0}^{n}\binom{m}{n}$.

Demostración. El lema es trivial cuando $\lambda \leq \sqrt{2m}$, pues $I_m(\lambda, \mathbf{r}) \leq |\prod_{k=1}^{m}[0..r_k]| = \prod_{k=1}^{m}(r_k+1)$. Suponiendo lo contrario, $\lambda > \sqrt{2m}$, se prueba por inducción sobre m. Para la base se toma $m = 1$ y $r \in \mathbb{N}$, que se traduce en

$$I_m(\lambda, r) = \left|\left\{i \in \mathbb{N}_0 : \frac{i}{r} \leq \frac{1-\lambda}{2}, 0 \leq i \leq r\right\}\right|.$$

Ya que $\lambda > 1$, $I_m(\lambda, r) = |\varnothing| = 0 \leq r+1$.

Supóngase el lema válido para $m = M-1 > 1$ y sea $1 < \sqrt{2M} < \lambda$. Con $\mathbf{0} < \mathbf{r} \in \mathbb{N}^M$, tómese la partición:

$$\left\{\mathbf{i} \in \mathbb{N}_0^M : \mathbf{i} \cdot \mathbf{r}^{-1} \leq \frac{M-\lambda}{2}, \mathbf{0} \leq \mathbf{i} \leq \mathbf{r}\right\} = \dot{\bigcup_{i=0}^{r_M}} \left\{\mathbf{i} \in \mathbb{N}_0^{M-1} : \mathbf{i} \cdot \mathbf{r}'^{-1} \leq \frac{M-\lambda}{2} - \frac{i}{r_M}, \mathbf{0} \leq \mathbf{i} \leq \mathbf{r}'\right\},$$

donde[7] $\mathbf{r}' = (r_1, \ldots, r_{M-1})$. Por lo anterior, basta con acotar la cardinalidad de cada conjunto en la unión. Primero, si $i \in [0..r_M]$, entonces trivialmente $\lambda - 1 + 2ir_M^{-1} > 0$, porque $\lambda > 1$ y $|-1+2ir_M^{-1}| < 1$. Luego, por la hipótesis de inducción,

$$\forall i \in [0..r_m] \quad I_{M-1}\left(\lambda - 1 + \frac{2i}{r_M}, \mathbf{r}'\right) \leq \frac{\sqrt{2(M-1)}}{\lambda - 1 + \frac{2i}{r_M}} \prod_{j=1}^{M-1}(r_j+1).$$

Lo anterior conduce a

$$I_M(\lambda, \mathbf{r}) = \sum_{i=0}^{r_M} I_{M-1}\left(\lambda - 1 + \frac{2i}{M}, \mathbf{r}'\right) \leq \sqrt{2(M-1)}\left(\sum_{i=0}^{r_m} \frac{1}{\lambda - 1 + \frac{2i}{M}}\right)\prod_{k=1}^{M-1}(r_k+1). \quad (1.12)$$

Ahora se acota el segundo factor del lado derecho de (1.12):

$$\sum_{i=0}^{r_M} \frac{2}{\lambda - 1 + \frac{2i}{r_M}} = \sum_{i=0}^{r_M} \frac{1}{\lambda - 1 + \frac{2i}{r_M}} + \sum_{i=0}^{r_M} \frac{1}{\lambda + 1 - \frac{2i}{r_M}} = \sum_{i=0}^{r_M} \frac{2\lambda}{\lambda^2 - \left(1 - \frac{2i}{r_M}\right)^2} < \frac{2\lambda(r_M+1)}{\lambda^2 - 1}. \quad (1.13)$$

En la primera igualdad los índices se recorren de 1 a r_M y en sentido contrario. Por otra parte, sacando raíces cuadradas y reordenando los términos,

$$1 - \frac{1}{M} < 1 - \frac{1}{M} + \frac{1}{4M^2} = \left(1 - \frac{1}{2M}\right)^2 \implies \sqrt{\frac{2(M-1)}{2M}} = \sqrt{1 - \frac{1}{M}} < 1 - \frac{1}{2M}$$
$$\implies \frac{\sqrt{2(M-1)}}{1 - \frac{1}{2M}} < \sqrt{2M},$$

que, junto con $\lambda > \sqrt{2M}$, da

$$\frac{\lambda\sqrt{2(M-1)}}{\lambda^2 - 1} = \frac{\lambda\sqrt{2(M-1)}}{\lambda^2\left(1 - \frac{1}{\lambda^2}\right)} < \frac{\sqrt{2(M-1)}}{\lambda\left(1 - \frac{1}{2M}\right)} < \frac{\sqrt{2M}}{\lambda}. \quad (1.14)$$

Sustituyendo (1.14) y (1.13) en (1.12), se concluye la desigualdad deseada. □

Aunque la primera parte de la siguiente proposición sea elemental, la segunda da una cota que será útil. Recuérdese que $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ es el polinomio mínimo de α y que $a = \mathcal{H}(f)$.

[7] Ver Notación

Lema 1.11. *Para cualquier $l \in \mathbb{N}$ existen $a_{j,l} \in \mathbb{Z}$ con $j \in \{0, \dots, n-1\}$ tales que*

$$\alpha^l = \sum_{j=0}^{n-1} a_{j,l}\alpha^j, \quad \max_j \{|a_{j,l}|\} \le (a+1)^l.$$

Demostración. Notemos que cuand $l \le n$ el lema es trivial. Para $l = n$ se tiene

$$\alpha^n = -\sum_{j=0}^{n-1} a_j\alpha^j, \quad \max_j \{|a_{j,l}|\} = \max_j \{|a_j|\} = a \le (a+1)^n.$$

Supóngase válido para $l - 1 \ge n$, entonces

$$\begin{aligned}
\alpha^l = \alpha\alpha^{l-1} &\overset{\text{H.I.}}{=} \alpha \sum_{j=0}^{n-1} a_{j,l-1}\alpha^j = \sum_{j=0}^{n-1} a_{j,l-1}\alpha^{j+1} = \left[\sum_{j=1}^{n-1} a_{j-1,l-1}\alpha^j\right] + a_{n-1,l-1}\alpha^n \\
&= \left[\sum_{j=1}^{n-1} a_{j-1,l-1}\alpha^j\right] - a_{n-1,l-1}\sum_{j=0}^{n-1} a_j\alpha^j \\
&= \sum_{j=0}^{n-1} a_{j,l}\alpha^j,
\end{aligned}$$

con $a_{0,l} := -a_{n-1,l-1}a_0$ y $a_{j,l} := a_{j-1,l-1} - a_{n-1,l-1}a_j$ para $j \in [1..n-1]$. La desigualdad buscada es obvia para $a_{0,l}$:

$$|a_{0,l}| = |-a_{n-1,l-1}a_0| = |-a_{n-1,l-1}||a_0| \overset{\text{H.I.}}{\le} (a+1)^{l-1}a \le (a+1)^l.$$

Y cuando $j \in [1..n-1]$,

$$|a_{j,l}| = |a_{j-1,l-1} - a_{n-1,l-1}a_j| \le |a_{j-1,l-1}| + |a_{n-1,l-1}a_j| \overset{\text{H.I.}}{\le} (a+1)^{l-1} + a(a+1)^{l-1} = (a+1)^l. \ \square$$

1.3.2. Construcción del polinomio

Juntando los Lemas 1.11 y 1.10 se construye el polinomio que probará el Teorema de Roth.

Demostración del Teorema 1.4. Para construir a R se considerarán a los coeficientes, $C(\mathbf{i})$, como incógnitas. La segunda condición da lugar a un conjunto de multi-índices, Ω, y como $[\mathbb{Q}(\alpha) : \mathbb{Q}] = n$, cada $\mathbf{j} \in \Omega$ conduce a n formas lineales igualadas a cero. Se acotan los coeficientes del sistema lineal homogéneo y, con el Lema de Siegel (Capítulo III, Lema 2.2), se concluirá la existencia de una solución entera no trivial del sistema y, por lo tanto, la existencia del polinomio buscado.

Se determinarán los coeficientes apropiados, $C(\mathbf{j})$, del siguiente polinomio

$$R(\mathbf{x}) := \sum_{j \le \mathbf{r}} C(\mathbf{j})\mathbf{x}^{\mathbf{j}}$$

de modo que las condiciones del teorema se satisfagan.

Sean

$$\Omega := \left\{\mathbf{i} : \mathbf{i}\cdot\mathbf{r}^{-1} \le \frac{m(1-\varepsilon)}{2},\ \mathbf{0} \le \mathbf{i} \le \mathbf{r}\right\},\ M := |\Omega|,\ \lambda := m\varepsilon,\ N := \prod_{k=1}^{m}(r_k+1).$$

Hay, entonces, N incógnitas, éstas se acomodan en un vector C siguiendo un orden lexicográfico[8]:

$$\mathcal{C} := (C(0,0,\dots,0), C(0,0,\dots,1), \dots, C(\mathbf{r})).$$

Por otra parte, por la definición de índice y como se está forzando la segunda condición, cualquier $\mathbf{i} \in \Omega$ tendrá que cumplir

$$0 = R_{\mathbf{i}}(\boldsymbol{\alpha}) = \sum_{\mathbf{j} \le \mathbf{r}} C(\mathbf{j}) \binom{\mathbf{j}}{\mathbf{i}} \boldsymbol{\alpha}^{\mathbf{j}-\mathbf{i}} = \sum_{\mathbf{j} \le \mathbf{r}} C(\mathbf{j}) \binom{\mathbf{j}}{\mathbf{i}} \alpha^{|\mathbf{j}-\mathbf{i}|} = \sum_{k=0}^{n-1} L_{\mathbf{i},k}(\mathcal{C}) \alpha^k.$$

Las funciones $L_{\mathbf{i},k}$ son las formas lineales en $\mathcal{C}$ que se obtienen al aplicar la primera parte del Lema 1.11 a cada potencia de α que aparezca en la segunda sumatoria. Los vectores que determinan a cada una de estas formas dependen de los coeficientes binomiales y de los números $a_{j,l}$ que devuelve el Lema 1.11. Hay, entonces, a lo más nM ecuaciones lineales en los coeficientes igualadas a cero.

Usando el Lema 1.11 y tomando en cuenta los coeficientes binomiales, el valor absoluto de los coeficientes de cada $L_{i,k}$ está acotado por $2^{|\mathbf{r}|}(a+1)^{|\mathbf{r}|}$. Además, por el Lema 1.10 y por $\varepsilon^2 > (8n^2)/m$,

$$M \le \frac{\sqrt{2m}}{m\varepsilon} N < \frac{\sqrt{2m}}{m} \frac{\sqrt{m}}{2\sqrt{2}n} N = \frac{N}{2n} \iff nM < \frac{N}{2} \iff \frac{N}{2} < N - nM \iff \frac{nM}{N-nM} < 1.$$

Finalmente, existe, por el Lema de Siegel (Apéndice C, Lema 2.2) ($N \mapsto N, M \to nM, A \mapsto (2a+1)^{|\mathbf{r}|}$) un vector $\mathbf{z} \in \mathbb{Z}^N$, $\mathbf{z} \neq \mathbf{0}$, tal que $L_{\mathbf{i},k}(\mathbf{z}) = 0$ para todas $\mathbf{i} \in \Omega$, $k \in [0..n-1]$ y

$$\forall j \in [1..n] \quad |z_j| \le \left\lfloor (NA)^{\frac{nM}{N-nM}} \right\rfloor \le NA = A \prod_{k=1}^{m} (r_k + 1) \le (2a+2)^{|\mathbf{r}|} 2^{|\mathbf{r}|} = (4a+4)^{\mathbf{r}}.$$

La primera desigualdad es cierta por el Lema de Siegel y la tercera, por $1 + r_k \le 2^{r_k}$. Tomando a $\mathbf{z}$ como los coeficientes de R se obtiene el resultado: la primera condición se cumple porque $\mathbf{z} \neq \mathbf{0}$; la segunda, porque para toda $\mathbf{i} \in \Omega$ se satisface $R_{\mathbf{i}}(\boldsymbol{\alpha}) = 0$ y la tercera, por la desigualdad anterior. □

1.3.3. Cota inferior para los índices de R

Los Teoremas 1.5 y 1.6 dan propiedades del índice de R en ciertos vectores racionales que aproximan a $\boldsymbol{\alpha}$. Antes, es necesario probar una pequeña proposición técnica.

Proposición 1.12. *Sean $0 < \delta < 10^{-1}$ y $0 < \varepsilon < 20^{-1}\delta$, entonces*

$$\left(1 + \frac{\delta}{3}\right)\left(1 - \frac{\delta}{2}\right) > (1+\varepsilon)^2. \tag{1.15}$$

Demostración.

$$1 + \frac{\delta}{40} < \frac{3}{2} \implies \frac{\delta}{10}\left(1 + \frac{\delta}{40}\right) < \frac{\delta}{6}\frac{9}{10} < \frac{\delta}{6}(1-\delta)$$
$$\implies (1+\varepsilon)^2 < \left(1 + \frac{\delta}{20}\right)^2 = 1 + \frac{\delta}{10}\left(1 + \frac{\delta}{40}\right) < 1 + \frac{\delta}{6} - \frac{\delta^2}{6} = \left(1 + \frac{\delta}{3}\right)\left(1 - \frac{\delta}{2}\right)$$

□

[8] El orden elegido es irrelevante.

Demostración del Teorema 1.5. Sean $\mathbf{i} \in \mathbb{N}^m$ tal que $\mathbf{i}\cdot\mathbf{r}^{-1} < \delta/8$, $T := R_{\mathbf{i}}$. Se mostrará que

$$T\left(\frac{p_1}{q_1},\dots,\frac{p_m}{q_m}\right) = 0$$

que implicará el teorema. Para ello bastará con establecer $|\mathbf{q}^{\mathbf{r}}T((p_1/q_m,\dots,p_m/q_m)| < 1$, pues el término del lado izquierdo es un entero. Para lograrlo, se acotará burda y apropiadamente cada término en la expansión de Taylor de T alrededor de $\boldsymbol{\alpha}$.

Utilizando el teorema de Taylor, recordando que $\boldsymbol{\eta} = \boldsymbol{\alpha} - \mathbf{p}/\mathbf{q}$,

$$T\left(\frac{\mathbf{p}}{\mathbf{q}}\right) = \sum_{\mathbf{j}\le\mathbf{r}} T_{\mathbf{j}}(\boldsymbol{\alpha})\boldsymbol{\eta}^{\mathbf{j}} \tag{1.16}$$

Por el Lema 1.9 y por el Teorema 1.4, $\mathcal{H}(T) \le 2^{|\mathbf{r}|}\,\mathcal{H}(R) \le 2^{|\mathbf{r}|}(4a+4)^{|\mathbf{r}|}$ y T tiene a lo más $(r_1+1)\cdots(r_m+1)$ sumandos, entonces

$$\begin{aligned}
|T_{\mathbf{i}}(\boldsymbol{\alpha})| &= \left|\sum_{\mathbf{j}\le\mathbf{r}} C'(\mathbf{j})\binom{\mathbf{j}}{\mathbf{i}}\boldsymbol{\alpha}^{\mathbf{j}}\right| \\
&\le 2^{|\mathbf{r}|}(4a+4)^{|\mathbf{r}|}2^{|\mathbf{r}|}(\max\{|\alpha|,1\})^{|\mathbf{r}|}\prod_{k=1}^{m}(r_k+1) \\
&\le 2^{|\mathbf{r}|}(4a+4)^{|\mathbf{r}|}2^{|\mathbf{r}|}(\max\{|\alpha|,1\})^{|\mathbf{r}|}2^{|\mathbf{r}|} \\
&= [32(a+1)\max\{|\alpha|,1\}]^{|\mathbf{r}|} =: \gamma^{|\mathbf{r}|}.
\end{aligned} \tag{1.17}$$

El Lema 1.3 y la hipótesis sobre $\mathbf{i}$ conllevan $\operatorname{ind}_{\mathbf{r}}(T;\boldsymbol{\alpha}) > m2^{-1}(1-3^{-1}\delta)$, pues

$$\begin{aligned}
\operatorname{ind}_{\mathbf{r}}(T;\boldsymbol{\alpha}) &= \operatorname{ind}_{\mathbf{r}}(R_{\mathbf{i}};\boldsymbol{\alpha}) \ge \operatorname{ind}_{\mathbf{r}}(R;\boldsymbol{\alpha}) - \mathbf{i}\cdot\mathbf{r}^{-1} \\
&\ge \frac{m(1-\varepsilon)}{2} - \mathbf{i}\cdot\mathbf{r}^{-1} > \frac{m(1-\varepsilon)}{2} - \frac{m\delta}{8} \\
&= \frac{m}{2}\left(1-\varepsilon-\frac{\delta}{4}\right) > \frac{m}{2}\left(1-\frac{\delta}{20}-\frac{\delta}{4}\right) = \frac{m}{2}\left(1-\frac{3}{10}\delta\right) \\
&> \frac{m}{2}\left(1-\frac{\delta}{3}\right).
\end{aligned}$$

En consecuencia, los sumandos de (1.16) que no se anulan son aquellos cuyos índices, $\mathbf{j}$, satisfacen

$$\mathbf{j}\cdot\mathbf{r}^{-1} > \frac{m}{2}\left(1-\frac{\delta}{3}\right) \tag{1.18}$$

y cuando esto se cumple, se sigue que

$$\begin{aligned}
-\log|\boldsymbol{\eta}^{\mathbf{j}}| &= -\sum_{k=1}^{m} j_k \log|\eta_k| \geq \sum_{k=1}^{m} j_k(2+\delta)\log q_k && \text{(por (1.2))}\\
&\geq (2+\delta)\sum_{k=1}^{m}(r_1\log q_1)\frac{j_k}{r_k} && \text{(por la primera desigualdad en (1.4))}\\
&= (2+\delta)(r_1\log q_1)\mathbf{j}\cdot\mathbf{r}^{-1}\\
&> (2+\delta)(r_1\log q_1)\frac{m}{2}\left(1-\frac{\delta}{3}\right) && \text{(por (1.18))}\\
&= \left(1+\frac{\delta}{2}\right)\left(1-\frac{\delta}{3}\right)m(r_1\log q_1)\\
&\geq \left(1+\frac{\delta}{2}\right)\left(1-\frac{\delta}{3}\right)\frac{\sum_{k=1}^{m} r_k\log q_k}{1+\varepsilon} && \text{(sumando la segunda desigualdad de(1.4))}\\
&> \frac{(1+\varepsilon)^2}{1+\varepsilon}\sum_{k=1}^{m} r_k\log q_k = (1+\varepsilon)\sum_{k=1}^{m} r_k\log q_k && \text{(por (1.15)).}
\end{aligned}$$

Evaluando a la exponencial en los extremos,

$$\boldsymbol{\eta}^{\mathbf{j}} < \frac{1}{(\mathbf{q}^{\mathbf{r}})^{1+\varepsilon}}.$$

Finalmente, utilizando la ecuación anterior, (1.17) y la hipótesis (1.3)

$$\begin{aligned}
\left|\mathbf{q}^{\mathbf{r}}T\left(\frac{\mathbf{p}}{\mathbf{q}}\right)\right| &= \left|\mathbf{q}^{\mathbf{r}}\sum_{\mathbf{j}\leq\mathbf{r}} T_{\mathbf{j}}(\alpha)\boldsymbol{\eta}^{\mathbf{r}}\right| \leq |\mathbf{q}^{\mathbf{r}}|\,\frac{\sum_{k=1}^{m}|T_{\mathbf{j}}(\alpha)|}{|\mathbf{q}^{\mathbf{r}}|^{1+\varepsilon}} \leq \frac{\gamma^{|\mathbf{r}|}\prod_{k=1}^{m}(r_k+1)}{|\mathbf{q}^{\mathbf{r}}|^{\varepsilon}}\\
&\leq \frac{[64(a+1)\max\{|\alpha|,1\}]^{|\mathbf{r}|}}{|\mathbf{q}^{\mathbf{r}}|^{\varepsilon}} = \prod_{k=1}^{m}\left[\frac{64(a+1)\max\{|\alpha|,1\}}{q_k^{\varepsilon}}\right]^{r_k} < 1. \qquad \square
\end{aligned}$$

1.3.4. Cota superior para los índices de R

La prueba del Teorema 1.6 es por inducción; no obstante, ésta no es directa. Los obstáculos son sorteados generalizando el wronskiano a polinomios en varias variables. Un caso particular de los wronskianos generalizados será el wronskiano usual. Con estos objetos se concluirán criterios para la dependencia lineal de polinomios en varias variables.

Wronskianos generalizados

Las pruebas detalladas de las proposiciones de esta breve sección están en el Apéndice D. A pesar de ser parte importante de la prueba, estudiarlas en este momento distraerían del argumento principal.

Definición 1.4. *Sean Δ_0, Δ_1, …, Δ_{l-1} operadores diferenciales de orden a lo más $0,1,\ldots,l-1$, respectivamente. Sean $R_k \in \mathcal{C}^{l-1}(\mathbb{R}^m;\mathbb{R})$, $k\in[0..l-1]$, funciones de $\mathbb{R}^m$ a $\mathbb{R}$ diferenciables $l-1$ veces. Un* ***wronskiano generalizado*** *de $R_1, R_2, \ldots, R_l$ es*

$$G(R_1,R_2,\ldots,R_m;\mathbf{x}) := \det(\Delta_j R_k(\mathbf{x})) \quad j,k\in[0..m].$$

Cuando $f_1, f_2, \ldots, f_n$ fuesen polinomios en una variable, el único wronskiano generalizado que no se anula completamente, G, es proporcional al wronskiano usual:

$$G(f_1, f_2, \ldots, f_n; x) = \frac{1}{0!1!\cdots(n-1)!}\mathfrak{W}(f_1, f_2, \ldots, f_n)(x)$$

con $\mathfrak{W}(f_1, f_2, \ldots, f_n)(x)$ el wronskiano usual de $f_1, f_2, \ldots, f_m$.

En general, la anulación del wronskiano usual no es un criterio para determinar la independencia lineal de las funciones involucradas. Sin embargo, la situación es diferente en el caso de polinomios.

Lema 1.13. *Sean $f_1, f_2, \ldots, f_n$ polinomios de $\mathbb{R}$ en $\mathbb{R}$, éstos son linealmente dependientes si y sólo si su wronskiano es idénticamente cero.*

Demostración. Apéndice D, Lema 2. □

Si $R_1, R_2, \ldots, R_l \in \mathcal{C}^{l-1}(\mathbb{R}^m; \mathbb{R})$ son linealmente dependientes, todos los wronskianos generalizados de estas funciones serán idénticamente cero. El siguiente lema, análogo al lema anterior, establece el recíproco.

Lema 1.14. *Sean $R_1, R_2, \ldots, R_l \in \mathbb{R}[x_1, \ldots, x_m]$ linealmente independientes, entonces al menos un wronskiano generalizado no es idénticamente cero.*

Demostración. Apéndice D, Lema 3. □

Lema Fundamental de Roth

A lo largo de la prueba habrá que usar ciertas desigualdades cuya verificación, aunque sencilla, distraería la atención. Son enunciados para poder referirlas cuando se requieran.

- Claramente, $0 < \omega < 1$, que junto con (1.8) da

$$r_m < r_{m-1} < \ldots < r_1. \tag{1.19}$$

- Es fácil probar la desigualdad $\log_2(x) \leq x - 1$ para cualquier $x > 0$. Entonces,

$$\forall n \in \mathbb{N} \left(\sum_{j=1}^{n} \log_2 j \leq \frac{n(n-1)}{2} \leq n(n-1) \right) \implies \forall n \in \mathbb{N} \left(n! \leq 2^{n(n-1)} \right). \tag{1.20}$$

- También, cuando $r_1, r_2, \ldots, r_h$ son enteros no negativos,

$$\sum_{j=1}^{h} \log_2(1 + r_j) \leq \sum_{j=1}^{h} r_j \implies \prod_{j=1}^{h} (1 + r_j) \leq 2^{r_1 + \ldots + r_n}. \tag{1.21}$$

Demostración del Teorema 1.6. La prueba es por inducción sobre m. La base no tiene mayor dificultad. Una idea natural para el paso inductivo es factorizar el polinomio en m variables en dos cuyos coeficientes sean enteros y tales que un factor dependa de a lo más $m - 1$ variables y el otro, del resto. Ya que esto no es siempre posible, se construyen dos familias de polinomios linealmente independientes. Con la ayuda de los wronskianos generalizados apropiados se obtiene un polinomio en el que aplicará la hipótesis de inducción y el teorema se concluirá.

Caso $m = 1$. Supóngase que $\operatorname{ind}_r(S, p/q) = n_1/r$, entonces

$$S(t) =: \left(t - \frac{p}{q}\right)^{n_1} S_1(t) = (qt - p)^{n_1} \left(\frac{1}{q^{n_1}} S_1(t)\right).$$

Tanto los coeficientes de S como los de $(qt-p)^{n_1}$ son enteros y por el Lema de Gauss, los de $(qt-p)^{n_1}$ son coprimos. Se afirma que los coeficientes de $q^{-n_1}S_1$ son enteros, por lo que q^{n_1} divide a todos los coeficientes de S_1. En efecto, si los coeficientes de S_1 no fueran enteros, el máximo común divisor de los coeficientes de $(qt-p)^{n_1}$ cancelaría a todos los denominadores en $q^{n_1}S_1$, sin embargo $(qt-p)^{n_1}$ es primitivo. Esto da

$$q^{n_1} \leq \mathcal{H}(S) \leq q^{\omega r \gamma} \implies \operatorname{ind}_r\left(P; \frac{p}{q}\right) = \frac{n_1}{r} \leq \omega\gamma \leq \omega = \varepsilon.$$

Caso $m > 1$.

I. Sean $\{\psi_j\}_{j=1}^h$, $\{\phi_j\}_{j=1}^h$ polinomios en una y $m-1$ variables, respectivamente, con coeficientes racionales tales que

$$S = \sum_{j=1}^{h} \phi_j(x_1, \ldots, x_{m-1})\psi_j(x_m)$$

donde h es mínimo con respecto a la propiedad de escribir de ese modo a S. Considerando a la posibilidad $\psi_j(x_m) = x_m^j$, se tiene que $h - 1 < h \leq r_m \leq |\mathbf{r}|$. Supóngase que $\{\phi_j\}_{j=1}^h$ son linealmente dependientes y que $c_1 \neq 0$ en

$$c_1\phi_1 + c_2\phi_2 + \ldots + c_h\phi_h \equiv 0 \implies \phi_1 = \sum_{j=2}^{h} \frac{c_j}{c_1}\phi_j.$$

Sustituyendo

$$S = \sum_{j=1}^{h} \phi_j\psi_j = \sum_{j=2}^{h} \phi_j \left(\psi_j - \frac{c_j}{c_1}\psi_1\right)$$

contradiciendo la minimalidad de h. Por lo tanto, $\{\phi_j\}_{j=1}^h$ es un conjunto linealmente independiente. Análogamente, las funciones $\{\psi_j\}_{j=1}^h$ son linealmente independientes.

II. Por el Lema 1.14, existe un wronskiano generalizado de $\{\phi_j\}_{j=1}^h$ que no se anula idénticamente y el wronskiano de $\{\psi_j\}_{j=1}^h$ no se anula idénticamente. Sean $\Delta^{(i)}$, $i = 1, 2, \ldots, h$, operadores diferenciales de orden a lo más i tales que

$$0 \not\equiv U(x_1, x_2, \ldots, x_{m-1}) := \det(\Delta^{(i-1)}\phi_j) \quad i, j \in [1..h].$$

Tómese

$$V(x_m) = \frac{1}{0!1!\ldots(h-1)!}\mathfrak{W}(\psi_1(x_m), \psi_2(x_m), \ldots, \psi_h(x_m)) \not\equiv 0.$$

Adquiriendo la notación $\partial^k := (k!)^{-1}\partial^k/\partial x_m^k$ se tiene

$$\begin{aligned}
W(\mathbf{x}) &:= U(x_1, x_2, \ldots, x_{m-1})V(x_m) \\
&= \det\left[\begin{pmatrix} \triangle^{(0)}\phi_1 & \ldots & \triangle^{(0)}\phi_h \\ \vdots & & \vdots \\ \triangle^{(h-1)}\phi_1 & \ldots & \triangle^{(h-1)}\phi_h \end{pmatrix}\begin{pmatrix} \psi_1 & \ldots & \partial^{h-1}\psi_1 \\ \vdots & & \vdots \\ \psi_h & \ldots & \partial^{h-1}\psi_h \end{pmatrix}\right] \\
&= \det\begin{pmatrix} \sum_{j=1}^{h} \triangle^{(0)}\phi_j\psi_j & \ldots & \sum_{j=1}^{h} \triangle^{(0)}\phi_j\partial^{h-1}\psi_j \\ \vdots & & \vdots \\ \sum_{j=1}^{h} \triangle^{(h-1)}\phi_j\psi_j & \ldots & \sum_{j=1}^{h} \triangle^{(h-1)}\phi_j\partial^{h-1}\psi_j \end{pmatrix} \\
&= \det\begin{pmatrix} \triangle^{(0)}\sum_{j=1}^{h} \phi_j\psi_j & \ldots & \triangle^{(0)}\partial^{h-1}\sum_{j=1}^{h} \phi_j\psi_j \\ \vdots & & \vdots \\ \triangle^{(h-1)}\sum_{j=1}^{h} \phi_j\psi_j & \ldots & \triangle^{(h-1)}\partial^{h-1}\sum_{j=1}^{h} \phi_j\psi_j \end{pmatrix} \\
&= \det\begin{pmatrix} \triangle^{(0)}S & \ldots & \triangle^{(0)}\partial^{h-1}S \\ \vdots & & \vdots \\ \triangle^{(h-1)}S & \ldots & \triangle^{(h-1)}\partial^{h-1}S \end{pmatrix} \\
&= \det(S_{(\mathbf{i},k-1)}) \quad k \in [1..h],\ \mathbf{i} \in J. \qquad (1.22)
\end{aligned}$$

en donde J es el conjunto de los multi-índices que definen a los operadores $\Delta^{(i)}$. El paso del primer al segundo renglón es válido, porque el determinante de un producto es el producto de determinantes, cada ϕ_j es constante con respecto a x_m y cada ϕ_j, con respecto a $x_1, \ldots, x_{m-1}$. Así, W tiene coeficientes enteros.

Nótese que U y V tienen coeficientes racionales, existe $k \in \mathbb{Q}$ tal que $kU \in \mathbb{Z}[x_1, \ldots x_{m-1}]$ es primitivo. Entonces, la evidente igualdad $W = (kU)(k^{-1}V)$ y el corolario al Lema de Gauss implican que los coeficientes de $u := kU$ y $v := k^{-1}V$ son enteros y $\mathcal{H}(u), \mathcal{H}(v) \leq \mathcal{H}(W)$.

III. La hipótesis de inducción será aplicada a u. Se verificará que las condiciones requeridas son satisfechas por u. Definiendo

$$\omega' := \omega\left(m-1, \frac{\varepsilon^2}{12}\right), \quad \forall\ j \in [1..m-1]\ r'_j := hr_j$$

se buscará corroborar la validez de

$$\forall\ j \in [1..m-2] \quad \omega' r'_j \geq r'_{j+1} \qquad (1.23)$$

$$\forall\ j \in [1..m-1] \quad \omega' q_j^{r'_j} \geq q_1^{\gamma r'_1},\ q_j^{\omega'\gamma} \geq 2^{3(m-1)} \qquad (1.24)$$

$$\mathcal{H}(u) \leq q_1^{\omega' r'_1 \gamma}, \qquad (1.25)$$

para poder concluir

$$\mathrm{ind}_{h\mathbf{r}}\left(u; \frac{\mathbf{p}}{\mathbf{q}}\right) < \frac{\varepsilon^2}{12}. \qquad (1.26)$$

Bajo las hipótesis del Teorema y la luz de la ecuación

$$\omega' = \left(m-1, \frac{\varepsilon^2}{12}\right) = \frac{24}{2^{m-1}}\left(\frac{\varepsilon^2}{12^2}\right)^{(2^{m-2})} = \frac{24}{2^{m-1}}\left(\frac{\varepsilon}{12}\right)^{2^{m-1}} = 2\omega(m,\varepsilon) = 2\omega$$

(1.23) y (1.24) son claras, no así la concerniente a la altura de u.

Primero, por (1.22), se puede escribir, con $\mathbf{i} \in J$ y $k \in [1..h]$,

$$W(\mathbf{x}) = \det\left(S_{(\mathbf{i},k-1)}\right) = \sum_{\sigma\in\mathbb{S}_h} (-1)^{\operatorname{sgn}\sigma} s_{\sigma(1),0}\cdots s_{\sigma(h),h-1}. \tag{1.27}$$

$\mathbb{S}_h$ es el grupo simétrico de orden h y $s_{j,k}$ son las entradas de la matriz que determina a $W(\mathbf{x})$. Como la última suma tiene $h!$ términos,

$$\mathcal{H}(W) \le \sum_{\sigma\in\mathbb{S}_h} \mathcal{H}(s_{\sigma(1),0}\cdots s_{\sigma(h),h-1}) \le h!\max_{\sigma\in\mathbb{S}_h}\{\mathcal{H}(s_{\sigma(1),0}\cdots s_{\sigma(h),h-1})\}. \tag{1.28}$$

Segundo, se toma $\sigma \in \mathbb{S}_h$. Al calcular $s_{\sigma(1),0}\cdots s_{\sigma(h),h-1}$ se obtiene a lo más $[\prod_{j=1}^{h}(r_j+1)]^h$ sumandos antes de simplificar, por lo que

$$\forall\sigma\in\mathbb{S}_h \quad \mathcal{H}(s_{\sigma(1),0}\cdots s_{\sigma(h),h-1}) \le \left[\prod_{j=1}^{h}(1+r_j)\right]^h \left[\max_{j\in[1..h]}\{\mathcal{H}(s_{\sigma(j),j}\}\right]^h. \tag{1.29}$$

Tercero, por (1.9) y la hipótesis sobre la altura de S,

$$\forall\ i,j\in[0..h-1] \quad \mathcal{H}(\Delta^{(i)}\partial^j S) \le 2^{|\mathbf{r}|}\,\mathcal{H}(S) \le 2^{|\mathbf{r}|}q_1^{\omega r_1\gamma}. \tag{1.30}$$

Finalmente, juntando las desigualdades,

$$\begin{aligned}
\mathcal{H}(W) &\le h!\max_{\sigma\in\mathcal{S}_h}\{\mathcal{H}(s_{\sigma(1),0}\cdots s_{\sigma(h),h-1})\} && \text{(por (1.28))}\\
&\le h!\left[\prod_{j=1}^{h}(1+r_j)\right]^h\left[\max_{j\in[1..h]}\{\mathcal{H}(s_{\sigma(j),j})\}\right]^h && \text{(por (1.29))}\\
&\le h!\left[\prod_{j=1}^{h}(1+r_j)\right]^h 2^{h|\mathbf{r}|}q_1^{\omega h r_1\gamma} && \text{(por 1.30)}\\
&\le 2^{h(h-1)}2^{h|\mathbf{r}|}2^{h|\mathbf{r}|}q_1^{\omega r_1'\gamma} && \text{(por (1.20) y (1.21))}\\
&\le 2^{3h|\mathbf{r}|}q_1^{\omega r_1'\gamma} && \text{(porque } h-1<h\le|\mathbf{r}|\text{)}\\
&\le 2^{3hmr_1}q_1^{\omega r_1'\gamma} && \text{(por (1.19))}\\
&\le q_1^{\omega h r_1\gamma}q_1^{\omega r_1'\gamma} = q_1^{\omega r_1'\gamma}q_1^{\omega r_1'\gamma} = q_1^{\omega' r_1'\gamma} && \text{(por (1.9)).}
\end{aligned}$$

En particular, $\mathcal{H}(u) \le q_1^{\omega' r_1'\gamma}$ y se puede aplicar la hipótesis de inducción en u. Además, con el caso $m=1$ bajo estos parámetros, se concluye, con $\mathbf{p}' = (p_1,\ldots,p_{m-1})$ y $\mathbf{q}' = (q_1,\ldots,q_{m-1})$,

$$\operatorname{ind}_{h\mathbf{r}}\left(u;\frac{\mathbf{p}'}{\mathbf{q}'}\right) < \frac{\varepsilon^2}{12}, \quad \operatorname{ind}_{h\mathbf{r}}\left(v;\frac{p_m}{q_m}\right) < \frac{\varepsilon^2}{12}.$$

IV. Es sencillo ver, con la ayuda de la definición de índice y el Lema 1.3, que

$$\operatorname{ind}_{\mathbf{r}}\left(W;\frac{\mathbf{p}}{\mathbf{q}}\right)=\operatorname{ind}_{\mathbf{r}}\left(u;\frac{\mathbf{p}}{\mathbf{q}}\right)+\operatorname{ind}_{\mathbf{r}}\left(v;\frac{\mathbf{p}}{\mathbf{q}}\right)=h\operatorname{ind}_{h\mathbf{r}}\left(u;\frac{\mathbf{p}}{\mathbf{q}}\right)+h\operatorname{ind}_{h\mathbf{r}}\left(v;\frac{\mathbf{p}}{\mathbf{q}}\right)<\frac{h\varepsilon^2}{6}.$$

V. Sea $\theta:=\operatorname{ind}_{\mathbf{r}}(S;\mathbf{p}/\mathbf{q})$. Para enteros no negativos $i_1+i_2+\ldots+i_{m-1}\le h-1\le r_m$ se tiene, por el Lema 1.3. *i*,

$$\begin{aligned}\operatorname{ind}_{\mathbf{r}}\left(S_{(i_1,\ldots,i_{m-1},j)};\frac{\mathbf{p}}{\mathbf{q}}\right)&\ge\theta-\sum_{k=1}^{m-1}\frac{i_k}{r_k}-\frac{j}{r_m}\ge\theta-\frac{1}{r_{m-1}}\sum_{k=1}^{m-1}i_k-\frac{j}{r_m}\ge\theta-\frac{r_m}{r_{m-1}}-\frac{j}{r_m}\\&\ge\theta-\omega-\frac{j}{r_m}\ge\theta-\frac{\varepsilon^2}{24}-\frac{j}{r_m}\end{aligned}\tag{1.31}$$

La expansión de W como determinante está dada en (1.27). Así, por el segundo inciso del Lema 1.3 e inducción, existen multi-índices $\mathbf{i}_1,\ldots,\mathbf{i}_h$ que verifican $|\mathbf{i}_j|\le h-1$ para $j\in[1..h]$ y también

$$\operatorname{ind}_{\mathbf{r}}\left(W;\frac{\mathbf{p}}{\mathbf{q}}\right)\ge\operatorname{ind}_{\mathbf{r}}\left(S_{(\mathbf{i}_1,0)}\cdots S_{(\mathbf{i}_h,h-1)};\frac{\mathbf{p}}{\mathbf{q}}\right).$$

Entonces, el tercer inciso del Lema 1.3, (1.31) y la no negatividad de los índices traen consigo la certeza de

$$\frac{h\varepsilon^2}{6}>\operatorname{ind}_{\mathbf{r}}\left(W;\frac{\mathbf{p}}{\mathbf{q}}\right)\ge\sum_{j=0}^{h-1}\operatorname{máx}\left\{\theta-\frac{\varepsilon^2}{24}-\frac{j}{r_m},0\right\}\ge-\frac{h\varepsilon^2}{24}+\sum_{j=0}^{h-1}\operatorname{máx}\left\{\theta-\frac{j}{m},0\right\},$$

de donde sigue

$$0\le\frac{1}{h}\sum_{j=1}^{h-1}\operatorname{máx}\left\{\theta-\frac{j}{r_m},0\right\}\le\frac{\varepsilon^2}{6}+\frac{\varepsilon^2}{24}<\frac{\varepsilon^2}{4}.$$

Ahora, hay que considerar dos casos: $\theta\ge(h-1)r_m^{-1}$ y $\theta<(h-1)r_m^{-1}$.
Caso I. Si $\theta\ge(h-1)r_m^{-1}$, entonces

$$\frac{\varepsilon}{2}>\frac{\varepsilon^2}{4}>\frac{1}{h}\sum_{j=0}^{h-1}\theta-\frac{j}{r_m}=\theta-\frac{1}{r_m}\frac{h-1}{2}\ge\frac{\theta}{2}\implies\varepsilon>\theta.$$

Caso II. Si $\theta<(h-1)r_m^{-1}$, entonces

$$\begin{aligned}\frac{\varepsilon^2}{4}&>\frac{1}{h}\sum_{j=0}^{\lfloor r_m\theta\rfloor}\theta-\frac{j}{r_m}\\&=\frac{1}{h}\left[\theta(\lfloor r_m\theta\rfloor+1)-\frac{\lfloor r_m\theta\rfloor(\lfloor r_m\theta\rfloor+1)}{2r_m}\right]\\&=\frac{\lfloor r_m\theta\rfloor+1}{2r_mh}\left[r_m\theta+(r_m\theta-\lfloor r_m\theta\rfloor)\right]\\&\ge\frac{\lfloor r_m\theta\rfloor+1}{2hr_m\theta}r_m\theta>\frac{r_m\theta^2}{2h}>\frac{h}{2}\frac{\theta^2}{2h}=\frac{\theta}{4}\implies\varepsilon>\theta=\operatorname{ind}_{\mathbf{r}}\left(S;\frac{\mathbf{p}}{\mathbf{q}}\right).\end{aligned}$$

□

Capítulo 2

El Teorema del Subespacio de Schmidt

2.1. Introducción

El Teorema del Subespacio de Schmidt, que emana del Teorema Fuerte del Subespacio de Schmidt, es una generalización en varias dimensiones del Teorema de Roth. Las prolijas demostraciones de los resultados de Schmidt son mayormente comprensibles, mas no fáciles. Hay en ellas aspectos importantes que exigen un estudio meticuloso. En este capítulo se expone detalladamente la prueba del Teorema Fuerte del Subespacio y, naturalmente, la del Teorema del Subespacio de Schmidt.

Es necesario establecer ciertas convenciones. Recuérdese que las normas $\|\cdot\|_\infty$ y $\|\cdot\|_1$ en $\mathbb{C}^n$ están dadas por

$$\forall \mathbf{z} \in \mathbb{C} \quad \|\mathbf{z}\|_\infty = \max\{|z_1|, \ldots, |z_n|\}, \quad \|\mathbf{z}\|_1 = \sum_{j=1}^{n} |z_j|,$$

en donde $\mathbf{z} = (z_1, \ldots, z_n)$. Estas reglas de correspondencia también dan lugar a normas en $\mathbb{R}^n$ y los mismos símbolos serán utilizados para denotarlas. Si $\varnothing \neq X \subseteq \mathbb{R}$ y $f, g : \mathbb{R} \to \mathbb{R}_{\geq 0}$ son funciones, se escribe $f \ll g$ cuando exista $K > 0$ tal que $f(x) \leq Kg(x)$ para $x \in X$ suficientemente grande y $g \gg f$ significa $f \ll g$. El signo $\ll$ se llama **símbolo de Vinogradov**.

Teorema 2.1 (Teorema del Subespacio de Schmidt, 1972)**.** *Sean $L_1, L_2, \ldots, L_n$, $L_j : \mathbb{C}^n \to \mathbb{C}$, formas lineales en n variables linealmente independientes sobre $\mathbb{C}$ con coeficientes complejos algebraicos. Para toda $\delta > 0$ existe sólo una cantidad finita de subespacios propios de $\mathbb{Q}^n$, $T_1, T_2, \ldots, T_w$, tales que cualquier $\mathbf{0} \neq \mathbf{x} \in \mathbb{Z}^n$ que satisfaga*

$$|L_1(\mathbf{x})L_2(\mathbf{x})\cdots L_n(\mathbf{x})| < \frac{1}{\|\mathbf{x}\|_\infty^\delta} \tag{2.1}$$

está en $T_1 \cup \ldots \cup T_w$.

Con el Teorema del Subespacio de Schmidt se descubre una prueba más accesible del Teorema de Roth.

Teorema (Roth, 1955)**.** *Sean α un irracional algebraico y $\delta > 0$. Existe sólo una cantidad finita de enteros p, q con $q > 0$ tales que*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}. \tag{2.2}$$

Demostración. Tómese (x, y), una pareja de enteros coprimos con $y > 0$ que resuelva (2.2), entonces

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{y^{2+\delta}} \implies \left|\frac{x}{y}\right| \leq \frac{1}{y^{2+\delta}} + |\alpha| \leq 1 + |\alpha| \implies |x| \leq (1 + |\alpha|)y \implies \|(x, y)\|_\infty \ll |y|.$$

Considerando las formas linealmente independientes $L_1 : \mathbb{C}^2 \to \mathbb{C}$, $L_2 : \mathbb{C}^2 \to \mathbb{C}$ dadas por

$$L_1(X, Y) = Y, \quad L_2(X, Y) = X - \alpha Y$$

se tiene

$$0 < |L_1(x, y)L_2(x, y)| = |y(x - \alpha y)| = y^2 \left|\frac{x}{y} - \alpha\right| < \frac{1}{y^\delta} \ll \frac{1}{\|(x, y)\|_\infty^\delta}.$$

Como L_1 y L_2 son $\mathbb{C}$-linealmente independientes, por el Teorema del Subespacio de Schmidt, todas las parejas que satisfagan (2.2) yacen en una cantidad finita de subespacios propios de $\mathbb{Q}^2$: $T_1, \ldots, T_r$.

Finalmente, hay que ver que cada subespacio contiene sólo una cantidad finita de soluciones. Fíjese $i \in [1..r]$. Como $T_i \subseteq \mathbb{Q}^2$ es un subespacio propio y distinto del cero, es unidimensional; luego, existe $(x_i, y_i) \in \mathbb{Z}$ con entradas coprimas y $y_i > 0$ tal que

$$T_i = \{t(x_i, y_i) : t \in \mathbb{Q}\}.$$

Cuando una pareja de enteros $(x, y) \in T_i$ satisfaga (2.2), existirá $t \in \mathbb{Z}$ tal que $(x, y) = t(x_i, y_i)$, por lo que

$$0 < \left|\alpha - \frac{x_i}{y_i}\right| = \left|\alpha - \frac{x}{y}\right| \leq \frac{1}{|t|^2 y_i^2} \implies t^2 \leq \left|\alpha - \frac{x_i}{y_i}\right|^{-1} \frac{1}{y_i^2}.$$

Así, t puede asumir una cantidad finita de valores y, por lo tanto, T_i contiene sólo un número finito de soluciones. □

Estos párrafos podrían hacer creer que los esfuerzos del capítulo fueron vanos. La larga prueba del Teorema de Roth no sólo fungirá como modelo para la imponente cadena de argumentos que soporta al Teorema Fuerte del Subespacio, también brinda algunas ideas como el Lema Fundamental de Roth.

2.2. Elementos de la Geometría de Números

La formulación del Teorema Fuerte del Subespacio obliga a una pequeña digresión sobre la Geometría de los Números. Además de asentar conceptos y terminología, esta sección permite exhibir algunas ideas usadas para apuntalar al Teorema Fuerte del Subespacio. Muchas demostraciones, sin embargo, son relegadas a los Apéndices B y C.

2.2.1. El Lema de Siegel

Se recuerda que $X \subseteq \mathbb{R}^n$ es **convexo** cuando $\mathbf{x}, \mathbf{y} \in X$ y $t \in [0,1]$ implican $t\mathbf{x}+(1-t)\mathbf{y} \in X$ y que X es **simétrico con respecto a 0** si $-\mathbf{x} \in X$ cada que $\mathbf{x} \in X$. Se denota $\mathfrak{m}$ a la medida de Lebesgue en $\mathbb{R}^n$ y $\mathfrak{M}$, a la σ-álgebra de los subconjuntos Lebesgue Medibles de $\mathbb{R}^n$.

Con operaciones elementales de matrices es fácil probar que para cualquier matriz $A \in \mathcal{M}(\mathbb{Z}, m, n)$ con $m > n$ existe $\mathbf{z} \in \mathbb{Z}^n$ tal que $\mathbf{z} \neq \mathbf{0}$ y $A\mathbf{z} = \mathbf{0}$. El famoso Lema de Siegel dice qué tan grande debe ser un cubo centrado en el origen para asegurar que tenga un elemento de $\mathbb{Z}^n \cap \ker A$.

Lema 2.2 (Siegel, 1929)**.** *Sea A una matriz de $m \times n$ con entradas en $\mathbb{Z}$, $A = (a_{i,j})$, con $m > n$ y $K \geq 0$ tal que $|a_{i,j}| \leq K$ para cualesquiera i, j. Entonces, existe $\mathbf{x} \in \mathbb{Z}^n$ que satisface*

$$A\mathbf{x} = \mathbf{0}, \quad 1 \leq \|\mathbf{x}\|_\infty \leq \lfloor (nK)^{\frac{m}{n-m}} \rfloor.$$

Demostración. Ver Apéndice C, Lema 6. □

2.2.2. Mínimos sucesivos

La idea detrás de los mínimos sucesivos es sencilla. Piénsese en un elipsoide, $P \subseteq \mathbb{R}^3$, centrado en el origen. Es posible que P no contenga puntos enteros; pero al multiplicarlo por $\lambda > 1$ éste se *infla* y cuando λ sea muy grande habrá uno, dos o tres puntos enteros linealmente independientes sobre $\mathbb{R}$ contenidos en λP. El menor $\lambda > 0$ que garantice la existencia de j puntos enteros linealmente independientes es el j-ésimo mínimo sucesivo de P.

Definición 2.1. *Sea $\mathscr{R} \subseteq \mathbb{R}^n$ convexo, compacto, simétrico con respecto al origen y tal que $0 < \mathfrak{m}(\mathscr{R}) < \infty$. Para cada $j \in [1..n]$ el j-ésimo* ***mínimo sucesivo*** *de $\mathscr{R}$ es*

$$\lambda_j := \inf\{\lambda > 0 : \lambda\mathscr{R} \text{ tiene } j \text{ puntos enteros } \mathbb{R}\text{-linealmente independientes}\}.$$

Definición 2.2. *Sean $i \in [1..n]$ y $\mathbf{g}_1, \ldots, \mathbf{g}_i \in \mathbb{Z}^n$ linealmente independientes. Se dice que los primeros i mínimos sucesivos de $\mathscr{R}$ se* ***alcanzan*** *en $\mathbf{g}_1, \ldots, \mathbf{g}_i$ si para toda $j \in [1..i]$ se tiene $\mathbf{g}_j \in \lambda_j \Pi$.*

Es importante notar que los vectores $\mathbf{g}_1, \ldots, \mathbf{g}_n$ no están únicamente determinados.

Segundo Teorema de Minkowski

En general, establecer cotas para los mínimos sucesivos es una tarea muy complicada. El Segundo Teorema de Minkowski, sin embargo, dice que el comportamiento de estos números no es tan caótico. Específicamente, cuando $\mathscr{R}$ se ha fijado y sus mínimos sucesivos se llaman $\lambda_1 \leq \ldots \leq \lambda_n$, el Segundo Teorema de Minkowski establece $1 \ll \lambda_1 \cdots \lambda_n \ll 1$, donde las constantes implícitas en $\ll$ dependen de n y $\mathscr{R}$.

Teorema 2.3 (Segundo Teorema de Minkowski)**.** *Sea $\mathscr{R} \subseteq \mathbb{R}^n$ compacto, convexo, simétrico con respecto al origen y de medida positiva. Los mínimos sucesivos de $\mathscr{R}$ satisfacen*

$$\frac{2^n}{n!} \leq \lambda_1 \cdots \lambda_n \mathfrak{m}(\mathscr{R}) \leq 2^n.$$

Demostración. Apéndice C, Teorema 10 □

Los ejemplos $\mathscr{R}_1 = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_1 \leq 1\}$ y $\mathscr{R}_2 = \{\mathbf{x} \in \mathbb{R}^2 : |x_1| \leq 10, |x_2| \leq 1\}$ atestiguan que las cotas del Segundo Teorema de Minkowski son las mejores posibles, ya que $\lambda_1(\mathscr{R}_1) = \lambda_2(\mathscr{R}_1) = 1$, $\mathfrak{m}(\mathscr{R}_1) = 2$ y $\lambda_1(\mathscr{R}_2) = 10^{-1}$, $\lambda_2(\mathscr{R}_2) = 1$, $\mathfrak{m}(\mathscr{R}_2) = 40$.

Ahora se considera un caso particular, sea $\Pi = \{\mathbf{x} \in \mathbb{R}^n : \forall i \in [1..n] \quad |\langle \mathbf{a}_i, \mathbf{x}\rangle| \leq 1\}$ con $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ linealmente independientes y fijos. Obviamente, Π satisface las cuatro condiciones pedidas en el Teorema 2.3. Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \Pi$ y

$$\Omega := \left\{ t_1\mathbf{x}_1 + \ldots + t_n\mathbf{x}_n : \sum_{j=1}^{n} |t_j| = 1 \right\}.$$

Integrando con respecto a $t_1, \ldots, t_n$ se obtiene

$$\mathfrak{m}(\Omega) = \frac{2^n}{n!} |\det(\mathbf{x}_1 | \ldots | \mathbf{x}_n)|.$$

Además, como $\Omega \subseteq \Pi$,

$$\frac{2^n}{n!} |\det(\mathbf{x}_1 | \ldots | \mathbf{x}_n)| = \mathfrak{m}(\Omega) \leq \mathfrak{m}(\Pi) \implies |\det(\mathbf{x}_1 | \ldots | \mathbf{x}_n)| \leq \frac{n!\mathfrak{m}(\Pi)}{2^n}.$$

Si $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$ son linealmente independientes y $\mathbf{g}_1 \in \lambda_1\Pi, \ldots, \mathbf{g}_n \in \lambda_n\Pi$, existen $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \Pi$ independientes tales que $\mathbf{g}_j = \lambda_j\mathbf{y}_j$ para $j \in [1..n]$. La desigualdad anterior, las propiedades elementales de los determinantes y el Segundo Teorema de Minkowski conllevan

$$|\det(\mathbf{g}_1 | \ldots | \mathbf{g}_n)| = \lambda_1 \cdots \lambda_n |\det(\mathbf{y}_1 | \ldots | \mathbf{y}_n)| \leq \lambda_1 \cdots \lambda_n \frac{n!\mathfrak{m}(\Pi)}{2^n} \leq n! \ll 1. \tag{2.3}$$

El Lema de Davenport

Las instancias más importantes de subconjuntos convexos de $\mathbb{R}^n$ y simétricos con respecto al origen en este contexto son los paralelepípedos. Sea $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subseteq \mathbb{R}^n$ una base. El **paralelepípedo generado** por $\mathbf{a}_1, \ldots, \mathbf{a}_n$ es

$$\Pi = \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |\langle \mathbf{x}, \mathbf{a}_j\rangle| \leq 1\}. \tag{2.4}$$

Los 1 en las desigualdades que determinan a Π pueden ser cambiados por valores positivos y el conjunto resultante vuelve a ser un paralelepípedo, pero es generado por otros vectores. Por ejemplo, si $Q_1, \ldots, Q_n > 0$, el conjunto

$$\Pi = \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |\langle \mathbf{x}, \mathbf{a}_j\rangle| \leq Q_j\}$$

es el paralelepípedo generado por $Q_1^{-1}\mathbf{a}_1, \ldots, Q_n^{-1}\mathbf{a}_n$.

Nota. Esta noción de paralelepípedo generado por una base de $\mathbb{R}^n$ **NO** es la usual. En otras palabras, si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ es una base de $\mathbb{R}^n$, Π definido como arriba no es

$$\left\{ \sum_{j=1}^{n} t_j\mathbf{a}_j : \forall j \in [1..n] \quad 0 \leq t_j \leq 1 \right\}.$$

Sin embargo, Π es la imagen bajo una transformación lineal del cubo $[-1, 1]^n$; por lo tanto, Π es la traslación de un paralelepípedo en el sentido usual. En efecto, sea $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ dada por $\Phi(\mathbf{x}) = A\mathbf{x}$, con $A = (\mathbf{a}_1 | \ldots | \mathbf{a}_n)^t$. Entonces, Φ es invertible y tanto ella como su inversa son continuas. De aquí sigue que $\Pi = \Phi^{-1}([-1, 1]^n)$ es compacto, además

$$\mathfrak{m}(\Pi) = |\det(A^{-1})| \, \mathfrak{m}([-1, 1]^n) = \frac{2^n}{|\det A|} \tag{2.5}$$

Es evidente que Π es convexo y simétrico en el origen.

Ya que toda transformación lineal $L : \mathbb{R}^n \to \mathbb{R}$ puede escribirse de manera única como $L(\mathbf{x}) = \langle \mathbf{x}, \mathbf{a} \rangle$, una colección de n funcionales lineales independientes, $L_1, \ldots, L_n : \mathbb{R}^n \to \mathbb{R}$, también determina un paralelepípedo:

$$\Pi = \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \ |L_j(\mathbf{x})| \leq 1\} = \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \ |\langle \mathbf{x}, \mathbf{a}_j \rangle| \leq 1\},$$

si $L_j(\mathbf{x}) = \langle \mathbf{a}_j, \mathbf{x} \rangle$ para cada $j \in [1..n]$. Es pertinente observar que

$$\forall \lambda > 0 \quad \lambda\Pi := \{\lambda \mathbf{x} : \mathbf{x} \in \Pi\} = \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |L_j(\mathbf{x})| \leq \lambda\}.$$

Más adelante se apelará a la igualdad anterior sin mecionarla de manera explícita.

Definición 2.3. *Sea $\mathbf{a}_1, \ldots, \mathbf{a}_n$ una base de $\mathbb{R}^n$. La **base recíproca**[1] de $\mathbf{a}_1, \ldots, \mathbf{a}_n$, denotada $\mathbf{a}_1^*, \ldots, \mathbf{a}_n^*$, es la base de $\mathbb{R}^n$ que cumple con*

$$\forall i, j \in [1..n] \quad \left\langle \mathbf{a}_i, \mathbf{a}_j^* \right\rangle = \delta_{i,j},$$

*$\delta_{i,j}$ el símbolo de Kronecker. Si Π es el paralelepípedo generado por $\mathbf{a}_1, \ldots, \mathbf{a}_n$, el **paralelepípedo recíproco de Π** es*

$$\Pi^* = \left\{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad \left|\left\langle \mathbf{a}_j^*, \mathbf{x} \right\rangle\right| \leq 1\right\}.$$

Por un argumento análogo al usado en (2.5), $\mathfrak{m}(\Pi^*) = 2^n |\det A|$ y $A = (\mathbf{a}_1 | \cdots | \mathbf{a}_n)^t$, por lo que $\mathfrak{m}(\Pi^*)\mathfrak{m}(\Pi) = 4^n$.

En [Sch] y [sch71] puede encontrarse una propiedad mucho más elusiva descubierta por Davenport.

Teorema 2.4 (Davenport)**.** *Sean $L_1, \ldots, L_n : \mathbb{R}^n \to \mathbb{R}$, $\det(L_1, \ldots, L_n) = 1$, formas lineales y $\lambda_1 \leq \ldots \leq \lambda_n$ los mínimos sucesivos de Π, el paralelepípedo generado por $L_1, \ldots, L_n$. Sean $\rho_1, \ldots, \rho_n$ reales tales que*

$$\rho_1 \geq \rho_2 \geq \ldots \geq \rho_n > 0, \tag{2.6}$$

$$\rho_1\lambda_1 \leq \rho_2\lambda_2 \leq \ldots \leq \rho_n\lambda_n, \tag{2.7}$$

$$\rho_1\rho_2 \cdots \rho_n = 1. \tag{2.8}$$

Entonces, existe una permutación de $L_1, \ldots, L_n$ –dígase $L_1', \ldots, L_n'$– tal que los mínimos sucesivos del paralelepípedo generado por $\rho_1 L_1', \ldots, \rho_n L_n'$, llamados $\lambda_1', \ldots, \lambda_n'$, verifican

$$\forall i \in [1..n] \quad \frac{\lambda_i \rho_i}{2^n} \leq \lambda_i' \leq 2^{n^2} (n!)^2 \lambda_i \rho_i. \tag{2.9}$$

Además, si $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$ son linealmente independientes y $\mathbf{g}_j \in \lambda_j \Pi$ para $j \in [1..n]$; entonces,

$$\forall i \in [2..n] \quad \forall \mathbf{x} \in \mathbb{Z}^n \setminus \operatorname{span}\{\mathbf{g}_1, \ldots, \mathbf{g}_i\} \quad \frac{\rho_i \lambda_i}{2^n} \leq \max\{|\rho_1 L_1'(\mathbf{x})|, |\rho_2 L_2'(\mathbf{x})|, \ldots, |\rho_n L_n'(\mathbf{x})|\}. \tag{2.10}$$

[1] En algunos textos como [Ca02] en lugar de recíproca se usa el término **polar**.

2.2.3. Dos teoremas de Mahler

Los últimos dos resultados de esta sección muestran cierta proporcionalidad entre los mínimos sucesivos de paralelepípedos vinculados de manera particular. El primero relaciona a los mínimos de un paralelepípedo con los de su recíproco y el segundo, con los del p-ésimo pseudocompuesto, un conjunto contenido en el álgebra exterior de $\mathbb{R}^n$ definido más adelante.

Teorema 2.5. *Sean $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ una base de $\mathbb{R}^n$; $\mathbf{a}_1^*, \ldots, \mathbf{a}_n^*$ la base recíproca y Π, Π^* los paralelepípedos generados por cada una. Si $\lambda_1, \ldots, \lambda_n$ y $\lambda_1^*, \ldots, \lambda_n^*$ son, respectivamente, los mínimos sucesivos de Π y Π^*, entonces*

$$\forall i \in [1..n] \quad \lambda_i^* \ll \frac{1}{\lambda_{n+1-i}} \ll \lambda_i^*, \tag{2.11}$$

Además, cuando $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$ son linealmente independientes y $\mathbf{g}_j \in \lambda_j \Pi$ para cualquier $j \in [1..n]$– es decir, $|\langle \mathbf{a}_i, \mathbf{g}_j \rangle| \leq \lambda_j$ para $i, j \in [1..n]$– y $\mathbf{g}_1^*, \ldots, \mathbf{g}_n^*$ es la base recíproca de $\mathbf{g}_1, \ldots, \mathbf{g}_n$, sucede

$$\forall i, j \in [1..n] \qquad \left|\left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle\right| \ll \frac{1}{\lambda_j}. \tag{2.12}$$

La constantes en (2.11) y (2.12) dependen de n.

Aunque se deduzca directamente del Segundo Teorema de Minkowski, el segundo resultado de Mahler requiere algunos comentarios sobre el Álgebra de Grassman en $\mathbb{R}^n$. En aras de facilitar la lectura, los detalles se relegan al Apéndice B.

Cuando n sea un natural dado, se define

$$\forall p \in [1..n] \quad C(n,p) \coloneqq \{X \subseteq [1..n] : |X| = p\}.$$

Además, si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ es una base de $\mathbb{R}^n$, escribiendo $\sigma = \{i_1 < \ldots < i_p\} \in C(n,p)$,

$$\forall \sigma \in C(n,p) \quad A_\sigma \coloneqq \mathbf{a}_{i_1} \wedge \ldots \wedge \mathbf{a}_{i_n}.$$

El variar a σ a lo largo de $C(n,p)$ da $l \coloneqq \binom{n}{p}$ vectores, A_σ , linealmente independientes (cfr. Apéndice B, Lema 9) pertenecientes al espacio $\mathbb{R}_p^n$, cuya dimensión es justamente l (cfr. Apéndice B, Definición 1). Por lo tanto, $\{A_\sigma : \sigma \in C(n,p)\}$ es una base de $\mathbb{R}_n^p$ y genera a un paralelepípedo tan importante que posee su propio nombre.

Definición 2.4. *Sea $\mathbf{a}_1, \ldots, \mathbf{a}_n$ una base de $\mathbb{R}^n$. El paralelepípedo en $\mathbb{R}_p^n$ generado por los vectores A_σ,*

$$\Pi^{(p)} = \left\{X \in \mathbb{R}_p^n : \forall \sigma \in C(n,p) \;\; |\langle A_\sigma, X \rangle| \leq 1\right\},$$

*se llama el **p-ésimo pseudocompuesto de Π**.*

Una vez que n y p han sido fijados, los elementos de $C(n,p)$ son acomodados de forma conveniente. Primero, si $\lambda_1, \ldots, \lambda_n$ son los mínimos sucesivos de Π, se define

$$\forall \tau \in C(n,p) \quad \lambda_\tau \coloneqq \prod_{i \in \tau} \lambda_i.$$

Y después, a cada $\tau \in C(n,p)$ se le asigna una etiqueta en $[1..l]$ de modo que $\lambda_{\tau_1} \leq \lambda_{\tau_2} \leq \ldots \leq \lambda_{\tau_n}$.

Supóngase que los mínimos sucesivos de Π se alcanzan en $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$. Escribiendo $\tau = \{j_1 < \ldots < j_p\} \in C(n,p)$,

$$\forall \tau \in C(n,p) \quad G_\tau \coloneqq \mathbf{g}_{j_1} \wedge \ldots \wedge \mathbf{g}_{j_p}.$$

No se está que los mínimos sucesivos de $\Pi^{(p)}$ se alcancen en los vectores G_τ.

Teorema 2.6. *Sean* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *una base,* Π *el paralelepípedo generado por ella,* $\Pi^{(p)}$ *el p-ésimo pseudocompuesto de* Π*,* $\lambda_1, \ldots, \lambda_n$ *y* $\nu_1, \ldots, \nu_l$ *los mínimos sucesivos de* Π *y* $\Pi^{(p)}$*, respectivamente. Entonces, tomando a* A_σ *y* G_τ *como arriba,*

$$\forall \sigma, \tau \in C(n,p) \qquad |\langle A_\sigma, G_\tau \rangle| \leq p! \lambda_\tau. \tag{2.13}$$

Además, existen constantes que sólo dependen de n para las que

$$\forall i \in [1..l] \quad \lambda_{\tau_i} \ll \nu_i \ll \lambda_{\tau_i}. \tag{2.14}$$

Cuando $p = 1$, (2.13) se reduce a la ya conocida $|\langle \mathbf{a}_i, \mathbf{g}_j \rangle| \leq \lambda_j$ para $i, j \in [1..n]$.

Una pequeña y aparentemente inofensiva proposición es parte medular de la prueba del Teorema Fuerte del Subespacio. Este lema participa directamente en la existencia del subespacio fijo en el Teorema Fuerte del Subespacio.

Lema 2.7. *Sean* $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^n$ *y* $\mathbf{y}_1, \ldots, \mathbf{y}_p \in \mathbb{R}^n$ *dos colecciones de vectores linealmente independientes. Los vectores* $\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p$ *y* $\mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p$ *son proporcionales si y sólo si* $\operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_p\} = \operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$.

Demostración. Ver [Sch], Lema 6C, p. 105. □

2.3. El Teorema Fuerte del Subespacio

El Teorema del Subespacio de Schmidt, igual que antes el Teorema de Roth, relaciona aspectos analíticos y algebraicos a través de la geometría y el álgebra lineal. El Teorema Fuerte del Subespacio recalca desde su enunciado la vena geométrica.

Teorema 2.8 (Teorema Fuerte del Subespacio)**.** *Sean* $L_1, \ldots, L_n$ *formas lineales,* $L_j : \mathbb{R}^n \to \mathbb{R}$*, linealmente independientes con coeficientes en* $\mathbb{R} \cap \mathbb{A}$ *y* $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ *tal que* $c_1 + c_2 + \ldots + c_n = 0$*. Para cualquier* $Q > 0$ *los reales* $\lambda_1(Q), \lambda_2(Q), \ldots, \lambda_n(Q)$ *denotan los mínimos sucesivos de*

$$\Pi(Q) := \{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |L_j(\mathbf{x})| \leq Q^{c_j}\}$$

Si existen $\delta > 0$*,* $d \in [1..n-1]$ *y* $\mathfrak{Q} \subseteq \mathbb{R}_{>0}$ *no acotado que satisfagan*

$$\forall Q \in \mathfrak{Q} \qquad \lambda_d(Q) < \frac{\lambda_{d+1}(Q)}{Q^\delta};$$

entonces, hay un subespacio $S^d \subseteq \mathbb{Q}^n$*,* $\dim S^d = d$*, y* $\mathfrak{Q}_1 \subseteq \mathfrak{Q}$ *no acotado tales que para cualquier* $Q \in \mathfrak{Q}_1$ *los primeros d mínimos sucesivos de* $\Pi(Q)$ *se alcanzan en* $\mathbf{g}_1(Q), \ldots, \mathbf{g}_d(Q) \in \mathbb{Z}^n \cap S^d$*; en otras palabras,* $\operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1(Q), \ldots, \mathbf{g}_d(Q)\} = S^d$.

Antes de trabajar en el Teorema Fuerte del Subespacio se cerciora que éste implica el Teorema del Subespacio. Para lograrlo es necesario un lema que traduce la conclusión del Teorema 2.8 sobre un subespacio fijo y mínimos sucesivos en términos de una cantidad finita de subespacios.

Lema 2.9. *Sean* $L_1, \ldots, L_n$ *y* $\mathbf{c} \in \mathbb{R}^n$ *como en el Teorema 2.8. Para cada* $\delta > 0$ *existe una cantidad finita de subespacios propios* $T_1, \ldots, T_w \subsetneq \mathbb{Q}^n$ *tales que*

$$\forall \mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \quad \left(\left(\forall j \in [1..n] \quad |L_j(\mathbf{x})| \leq \|\mathbf{x}\|_\infty^{c_j - \delta} \right) \implies \mathbf{x} \in \bigcup_{k=1}^{w} T_k \right). \tag{2.15}$$

Demostración. La prueba es por contradicción. En breve, se supone que las soluciones de (2.15) no están en una cantidad finita de subespacios propios de $\mathbb{Q}^n$. Se toma una sucesión de soluciones, $(\mathbf{x}_j)_{j=1}^{\infty}$, de modo que cualesquiera n términos sean linealmente independientes. Después de mostrar que las hipótesis del Teorema 2.8 son satisfechas, se fija un subespacio, S^d, de dimensión menor que n y se extrae una subsucesión de $(\mathbf{x}_j)_{j=1}^{\infty}$, $(\mathbf{y}_j)_{j=1}^{\infty}$, cuyos elementos estén contenidos en S^d. Así, cualesquiera n términos de $(\mathbf{y}_j)_{j=1}^{\infty}$ serán linealmente dependientes, contradiciendo la elección de $(\mathbf{x}_j)_{j=1}^{\infty}$.

I. Algunos nombres son establecidos antes de empezar. Se define

$$\forall Q > 0 \quad \Pi(Q) := \{\mathbf{y} \in \mathbb{R}^n : \forall j \in [1..n] \quad |L_i(\mathbf{y})| \leq Q^{c_i}\}.$$

Para $i \in [1..n]$ el i-ésimo mínimo sucesivo de Π es $\lambda_i := \lambda_i(Q)$; $\mathbf{g}_1(Q), \ldots, \mathbf{g}_n(Q) \in \mathbb{Z}^n$, $\mathbf{g}_i := \mathbf{g}_i(Q)$, son puntos en los que $\lambda_1, \ldots, \lambda_n$ se alcanzan y $S_i := S_i(Q) := \operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1, \ldots, \mathbf{g}_i\}$.

II. Supóngase que la conclusión del Lema 2.9 es falsa; esto es, que no existe una colección finita de subespacios propios de $\mathbb{Q}^n$ cuya unión contenga a todas las soluciones de (2.15). Sea $(\mathbf{z}_k)_{k=1}^{\infty}$ una enumeración de estas soluciones. Se extraerá una subsucesión $(\mathbf{z}_{k_j})_{j=1}^{\infty}$ con la particularidad de que cualesquiera n términos son linealmente independientes.

Sean $\mathbf{z}_{k_1} := \mathbf{z}_1$ y

$$\forall j \in [2..n] \quad k_j := \min\left\{i \in \mathbb{N}_{\geq 2} : \mathbf{z}_{k_1}, \ldots, \mathbf{z}_{k_{j-1}}, \mathbf{z}_i \text{ son } \mathbb{Q}\text{-linealmente independientes}\right\}.$$

No es complicado ver que existen $\mathbf{z}_{k_1}, \ldots, \mathbf{z}_{k_n}$. En efecto, si para alguna $i \in [2..n]$ se tuviera a $\mathbf{z}_{k_1}, \ldots, \mathbf{z}_{k_{i-1}}$ pero no a $\mathbf{z}_{k_i}$, todos los términos de $(\mathbf{z}_j)_{j=1}^{\infty}$ pertenecerían a $\operatorname{span}_{\mathbb{Q}}\{\mathbf{z}_{k_1}, \ldots, \mathbf{z}_{k_{i-1}}\}$, un subespacio propio de $\mathbb{Q}^n$, que contraviene a la suposición inicial.

Recuérdese que si $j \in \mathbb{N}_{>n}$, $C(j, n-1)$ es la colección de subconjuntos de $[1..j]$ con $n-1$ elementos. Escribiendo $\boldsymbol{l} = \{l_1 < \ldots < l_{n-1}\} \in C(j, n-1)$,

$$\forall j \in \mathbb{N}_{\geq n} \quad k_{j+1} := \min\left\{i \in \mathbb{N} : \forall \boldsymbol{l} \in C(j, n-1) \quad \mathbf{z}_{k_{l_1}}, \ldots, \mathbf{z}_{k_{l_{n-1}}}, \mathbf{z}_i \text{ son } \mathbb{Q}\text{-lin. indep.}\right\}.$$

Si no existiera k_{n+1}, para cualquier $j \in \mathbb{N}_{>k_n}$ habría una subcolección de $\{\mathbf{z}_{k_1}, \ldots, \mathbf{z}_{k_n}\}$ con $n-1$ vectores– $\mathbf{z}_{k'_1}, \ldots, \mathbf{z}_{k'_{n-1}}$– tal que $\mathbf{z}_{k'_1}, \ldots, \mathbf{z}_{k'_{n-1}}, \mathbf{z}_j$ son linealmente dependientes. En una combinación lineal no trivial de $\mathbf{z}_{k'_1}, \ldots, \mathbf{z}_{k'_{n-1}}, \mathbf{z}_j$ igualada a $\mathbf{0}$, la independencia de $\mathbf{z}_{k'_1}, \ldots, \mathbf{z}_{k'_{n-1}}$ obliga al coeficiente de $\mathbf{z}_j$ a no anularse, por lo que $\mathbf{z}_j \in \operatorname{span}_{\mathbb{Q}}\{\mathbf{z}_{k'_1}, \ldots, \mathbf{z}_{k'_{n-1}}\}$. De este modo, los términos de $(\mathbf{z}_j)_{j=1}^{\infty}$ estarían en a lo más n subespacios propios de $\mathbb{Q}^n$. En general, si para $i \geq n+1$ existiesen $k_1, \ldots, k_{i-1}$ pero no k_i, sería posible dar $\binom{i-1}{n-1}$ subespacios propios de $\mathbb{Q}^n$ que contuvieran a todas las soluciones de (2.15). Por lo tanto, $(\mathbf{z}_{k_j})_{j=1}^{\infty}$ está bien definida. Adoptando una escritura sencilla, $(\mathbf{x}_j)_{j=1}^{\infty}$ es $(\mathbf{z}_{k_j})_{j=1}^{\infty}$.

III. Como los términos de $(\mathbf{x}_j)_{j=1}^{\infty}$ son enteros, $\|\mathbf{x}_j\|_{\infty} \to \infty$ cuando $j \to \infty$. Fíjese por un momento a $j \in \mathbb{N}$. Sean $\mathbf{x} = \mathbf{x}_j$ y $Q = Q_j = \|\mathbf{x}\|_{\infty}$. De (2.15) se obtiene

$$\left(\forall i \in [1..n] \quad \left|L_i(Q^{\delta}\mathbf{x})\right| \leq Q^{\delta}Q^{c_i-\delta} = Q^{c_i}\right) \implies Q^{\delta}\mathbf{x} \in \Pi(Q) \implies \mathbf{x} \in Q^{-\delta}\Pi(Q).$$

La última expresión implica $\lambda_1 \leq Q^{-\delta}$. Luego, $\lambda_n > 1$ cuando j es suficientemente grande, porque el segundo Teorema de Minkowski dice $\lambda_1 \cdots \lambda_n \gg 1$. Como $\lambda_1 \leq \lambda_2 \leq$

$\ldots \le \lambda_n$ y $\lambda_1 < \lambda_n$, existe $i \in [1..n-1]$ tal que $\lambda_i < \lambda_{i+1}$. Nótese que $\mathbf{x} \in \lambda_i \Pi \supseteq \lambda_1 \Pi$. Si $\mathbf{x} \notin \operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1, \ldots, \mathbf{g}_i\}$, el conjunto $\lambda_i \Pi$ tendría $i+1$ puntos en $\mathbb{Z}^n$ linealmente independientes, algo imposible por $\lambda_i < \lambda_{i+1}$. Entonces, $\mathbf{x} \in S_i$ y, por $S_1 \subseteq S_2 \subseteq \ldots \subseteq S_{n-1} \subseteq S_n = \mathbb{Q}^n$, se tiene $\mathbf{x} \in S_{n-1}$.

Llámese k al mínimo entero tal que $\mathbf{x} \in S_k$. Se afirma que $\lambda_k \le Q^{-\delta}$. Si no fuese así, se tendría $\lambda_1 \le Q^{-\delta} < \lambda_k$ y habría un $i \in [2..k]$ tal que $\lambda_{i-1} \le Q^{-\delta} < \lambda_i$. De estas desigualdades y $\mathbf{x} \in Q^{-\delta}\Pi$ se concluye que $\mathbf{x} \in \operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1, \ldots, \mathbf{g}_{i-1}\} = S_{i-1}$, contrario a la definición de k. En consecuencia,

$$\exists\, d \in [k..n-1] \qquad \lambda_d \le \frac{\lambda_{d+1}}{Q^{\frac{\delta}{n}}}.$$

Si así no fuera, para toda $d \in [k..n-1]$ valdría $Q^{-\delta/n} > \lambda_{d+1}/\lambda_d$ que, junto con $1 < \lambda_n$ y $\lambda_1 \le Q^{-\delta}$, da

$$Q^{-\delta} > \frac{Q^{-\delta}}{\lambda_n} \ge \frac{\lambda_k}{\lambda_n} = \frac{\lambda_k}{\lambda_{k_j+1}} \cdots \frac{\lambda_{n-1}}{\lambda_n} \ge Q^{-\frac{\delta(n-k)}{n}} \implies -\delta > -\delta\left(\frac{n-k}{n}\right) \implies 0 > k,$$

una aseveración ridícula. En resumen, $\mathbf{x} \in S_k \subseteq S_i$.

IV. Variando $Q = Q_j = \|\mathbf{x}_j\|_\infty$ en los índices suficientemente grandes, al menos una $d \in [1..n-1]$ se repetirá una infinidad de veces. Sean $(\mathbf{y}_j)_{j=1}^\infty$ una subsucesión de $(\mathbf{x}_j)_{j=1}^\infty$ en la que todos los términos tengan asociada la misma d y

$$\mathfrak{Q} := \{\|\mathbf{y}_j\|_\infty : j \in \mathbb{N}\}.$$

Por el Teorema Fuerte del Subespacio, existen un subespacio $S \subseteq \mathbb{R}^n$ con $\dim S = d$ y $\mathfrak{Q}' \subseteq \mathfrak{Q}$ no acotado tales que $Q \in \mathfrak{Q}'$ implica $\mathbf{g}_1, \ldots, \mathbf{g}_d \in S$ o, en otras palabras, $S_d = S$. Tómese una subsucesión $(\mathbf{y}'_j)_{j=1}^\infty$ de $(\mathbf{y}_j)_{j=1}^\infty$ que cumpla con $\|\mathbf{y}'_j\|_\infty \in \mathfrak{Q}'$ para toda $j \in \mathbb{N}$. Estableciendo $Q'_j := \|\mathbf{y}'_j\|_\infty$ se tiene

$$\forall\ j \in \mathbb{N} \quad \mathbf{y}'_j \in S_d(Q'_j) = S \ne \mathbb{Q}^n$$

y cualesquiera n términos de $(\mathbf{y}'_j)_{j=1}^\infty$ son dependientes, contradiciendo la elección de $(\mathbf{x}_j)_{j=1}^\infty$. □

Demostración del Teorema del Subespacio.

I. Supóngase que los coeficientes de $L_1, \ldots, L_n$ son reales algebraicos. Ya que las formas lineales no son iguales a cero, su núcleo es de dimensión $n-1$ y todas las soluciones que satisfagan $L_1(\mathbf{x})\cdots L_n(\mathbf{x}) = 0$ pertenecen a una cantidad finita de subespacios propios de $\mathbb{Q}^n$: $\bigcup_{j=1}^n \ker L_j$. Se trabaja, pues, con

$$\mathcal{S} := \left\{\mathbf{x} \in \mathbb{Z}^n : 0 < |L_1(\mathbf{x})\cdots L_n(\mathbf{x})| \le \frac{1}{\|\mathbf{x}\|_\infty^\delta}\right\}.$$

El Lema 2 del Apéndice E avala la existencia de $A = A(L_1, \ldots, L_n) > 0$ tal que

$$\forall j \in [1..n] \quad \frac{1}{\|\mathbf{x}\|_\infty^A} \ll |L_j(\mathbf{x})| \ll \|\mathbf{x}\|_\infty,$$

donde las constantes implicadas en los símbolos de Vinogradov dependen exlusivamente de los coeficientes $L_1, \dots, L_n$. Entonces, cuando $\|\mathbf{x}\|_\infty$ es grande[2], satisface

$$\frac{1}{\|\mathbf{x}\|_\infty^{2A}} \le |L_j(\mathbf{x})| < \|\mathbf{x}\|_\infty^2.$$

II. Divídase al intervalo $[-2A, 2)$ en subintervalos disjuntos $[c_i', c_i'')$ con $c_i'' - c_i' < \delta/(2n)$. A partir de esta partición de $[-2A, 2)$ se obtiene una para $\mathcal{S}$:

$$\bigcup \left\{ \mathbf{x} \in \mathbb{Z}^n : (\mathbf{x} \in \mathcal{S}) \ \& \ \left(\forall j \in [1..n] \quad \|\mathbf{x}\|_\infty^{c_{i_j}'} \le |L_j(\mathbf{x})| < \|\mathbf{x}\|_\infty^{c_{i_j}''} \right) \right\},$$

la unión corre sobre las formas de elegir a los intervalos $[c_{i_1}', c_{i_1}''), \dots, [c_{i_n}', c_{i_n}'')$. Fíjese algún término en la unión que no sea vacío, B. Para mantener manejable la notación, $C_j' := c_{i_j}'$, $C_j'' := c_{i_j}''$ y

$$B = \left\{ \mathbf{x} \in \mathbb{Z}^n : (\mathbf{x} \in \mathcal{S}) \ \& \ \left(\forall j \in [1..n] \quad \|\mathbf{x}\|_\infty^{C_j'} \le |L_j(\mathbf{x})| < \|\mathbf{x}\|_\infty^{C_j''} \right) \right\} \neq \varnothing.$$

Las desigualdades que determinan a B dan

$$\sum_{j=1}^n C_j' < -\delta \implies \sum_{j=1}^n C_j'' \le \sum_{j=1}^n \left[C_j' + \frac{\delta}{2n} \right] < -\delta + \frac{\delta}{2} = -\frac{\delta}{2}.$$

Al definir $\mathbf{C} = (C_1, \dots, C_n)$ con

$$\forall j \in [1..n] \quad C_j := C_j'' - \frac{1}{n} \sum_{k=1}^n C_k''$$

se tiene $C_1 + \ldots + C_n = 0$ y, como para cualesquiera $\mathbf{x} \in B$ y $j \in [1..n]$ valen $|L_j(\mathbf{x})| \le \|\mathbf{x}\|_\infty^{C_j''}$ y $C_j'' < C_n + \delta/(2n)$,

$$\forall \mathbf{x} \in B \quad \left(\forall j \in [1..n] \quad \left(|L_j(\mathbf{x})| < \|\mathbf{x}\|_\infty^{C_j - \frac{\delta}{2n}} \right) \right).$$

Las hipótesis del Lema 2.9 se cumplen para $\mathbf{C}$ y $\delta/(2n)$ en lugar de δ, por lo que todos los puntos de B pertenecen a una cantidad finita de subespacios propios de $\mathbb{Q}^n$. En consecuencia, las soluciones de (2.1) están en una cantidad finita de subespacios propios de $\mathbb{Q}^n$.

III. Cuando los coeficientes de $L_1, \dots, L_n$ no estén en $\mathbb{R}$, el Teorema del Subespacio de Schmidt se verifica por inducción. La base– todas tienen coeficientes reales– ya se ha verificado. Supóngase cierto el resultado cuando hay a lo más j formas cuyos coeficientes no son reales y que hay $j+1$ formas con esa característica: $L_{n-j-1}, L_{n-j}, \dots, L_n$.

Escríbase $L_n(\mathbf{x}) = R(\mathbf{x}) + iI(\mathbf{x})$ donde R, I tienen coeficientes en $\mathbb{R}$. Al menos para una de las formas R o I, llámese L_n', el conjunto $L_1, \dots, L_{n-1}, L_n'$ es linealmente independiente [3]. Entonces, si $\mathbf{x}$ resuelve (2.1),

$$\frac{1}{\|\mathbf{x}\|^{2\delta}} > |L_1 \cdots L_{n-1} L_n(\mathbf{x})|^2 = |L_1 \cdots L_{n-1}(R + iI)(\mathbf{x})|^2 = |L_1 \cdots L_{n-1}(\mathbf{x})|^2 \left(|R(\mathbf{x})|^2 + |I(\mathbf{x})|^2 \right)$$
$$\ge |L_1 \cdots L_{n-1}(\mathbf{x})|^2 |L_n'(\mathbf{x})|^2.$$

[2] Si K_1, K_2 son constantes para las que $K_1 \|\mathbf{x}\|^{-A} \le |L_j(\mathbf{x})| \le K_2 \|\mathbf{x}\|_\infty$ basta con $\|\mathbf{x}\|_\infty \ge \max\{K_1^{1/A}, K_2\}$.

[3] De lo contrario, I y R estarían en $\operatorname{span}_{\mathbb{Q}}\{L_1, \dots, L_{n-1}\}$, que daría $L_n = R + iI \in \operatorname{span}_{\mathbb{Q}}\{L_1, \dots, L_n\}$.

En consecuencia, la Hipótesis de Inducción es aplicable a $L_1, \ldots, L_{n-1}, L'_n$ y el resultado sigue. □

2.3.1. Prueba del Teorema Fuerte del Subespacio

La teoría sobre los mínimos sucesivos y la concerniente al álgebra de Grassman son las dos columnas que sostienen la demostración del Teorema Fuerte del Subespacio. A primera vista, la pruebas de este resultado y la del Teorema de Roth difieren significativamente. Las similitudes, no obstante, serán evidentes al justificar la existencia de un vector mediante la construcción de un polinomio auxiliar. Para no perder de vista la idea general, las largas demostraciones de algunos resultados son diferidas.

Demostración del Teorema 2.8. En esta demostración las constantes implícitas en $\ll$ dependen únicamente de n y $L_1, \ldots, L_n$.

I. Supóngase, adicionalmente, que $\|\mathbf{c}\|_\infty \leq n^{-1}$. Con $p := n - d$ y $\sigma = \{i_1 < \ldots < i_p\}$ para $\sigma \in C(n,p)$, defínase

$$\forall \sigma \in C(n,p) \quad A_\sigma := \mathbf{a}_{i_1} \wedge \ldots \wedge \mathbf{a}_{i_p},\ c_\sigma := \sum_{i \in \sigma} c_i,$$

en donde $L_i(\mathbf{x}) = \langle \mathbf{a}_i, \mathbf{x} \rangle$, $i \in [1..n]$. Llámese $l = \binom{n}{p}$.

II. Se buscará una $Q_1 > 0$ para aplicar el Lema 2.17[4] a $\mathbf{c} \leftarrow \{c_\sigma : \sigma \in C(n,p)\}$, $\delta \leftarrow \delta/2 > 0$, $\mathfrak{F} \leftarrow \mathfrak{Q}_{>Q_1}$ y a $\Pi(Q) \leftarrow \Pi^{(p)}(Q)$ para $Q \in \mathfrak{Q}_{>Q_1}$. Recuérdese que $\Pi^{(p)}(Q)$, el p-ésimo pseudocopuesto de $\Pi(Q)$, es

$$\Pi^{(p)}(Q) := \{X \in \mathbb{R}^n_p : \forall \sigma \in C(n,p) \quad |\langle A_\sigma, X \rangle| \leq Q^{c_\sigma}\}.$$

Los mínimos sucesivos de $\Pi^{(p)}(Q)$ se denotan $\nu_1, \ldots, \nu_l$.

Hay dos suposiciones tácitas en el Lema 2.17, son fácilmente verificables:

$$\sum_{\sigma \in C(n,p)} c_\sigma = \binom{n-1}{p-1} \sum_{j=1}^{n} c_j = 0, \quad \forall \sigma \in C(n,p) \quad |c_\sigma| \leq 1. \tag{2.16}$$

Para cualquier $Q \in \mathfrak{Q}$, con $\lambda_j = \lambda_j(Q)$, $\nu_i = \nu_i(Q)$ ($j \in [1..n]$, $i \in [1..l]$) y recordando el orden de $C(n,p)$, se tiene

$$\lambda_{\tau_l} = \lambda_{n-p}\lambda_{n-p+1}\cdots\lambda_n = \lambda_{d+1}\lambda_{d+2}\cdots\lambda_n, \quad \lambda_{\tau_{l-1}} = \lambda_{n-p-1}\lambda_{n-p+1}\cdots\lambda_n = \lambda_d\lambda_{d+2}\cdots\lambda_n,$$

Entonces, el Teorema 2.6 da $\lambda_{d+1}\lambda_{d+2}\cdots\lambda_n \ll \nu_l$ y $\nu_{l-1} \ll \lambda_d\lambda_{d+2}\cdots\lambda_n$. En consecuencia, de $\lambda_d < \lambda_{d+1}Q^{-\delta}$ se deduce $\nu_{l-1} \ll \nu_l Q^{-\delta}$ cuando $Q \in \mathfrak{Q}$. Luego, existe[5] $Q_1 > 0$ tal que

$$\forall Q \in \mathfrak{Q}_{>Q_1} \quad \nu_{l-1}(Q) < \frac{\nu_l(Q)}{Q^{\frac{\delta}{2}}}. \tag{2.17}$$

A cada $Q > 0$ se le asocia un par de bases de $\mathbb{R}^n_p$. La primera, $\left\{V_1^{(p)}, \ldots, V_l^{(p)}\right\}$, tiene coeficientes en $\mathbb{Z}$ con respecto a la base canónica y $V_j^{(p)} \in \nu_j \Pi^{(p)}(Q)$, $j \in [1..l]$;

[4] Algunos parámetros no aparecen en el enunciado del Lema 2.17, están definidos al principio de esa sección.

[5] Cualquiera que satisfaga $Q_1 > K^{\delta/2}$ con K la constante implícita en $\nu_{l-1} \ll \nu_l Q^{-\delta}$

es decir, en esta base se alcanza a los mínimos sucesivos de $\Pi^{(p)}(Q)$. La segunda, $\left\{V_1^{(p)*},\dots,V_l^{(p)*}\right\}$, es la recíproca de $\left\{V_1^{(p)},\dots,V_l^{(p)}\right\}$. Entonces, por el Lema 2.17, existen $\mathfrak{Q}' \subseteq \mathfrak{Q}_{>Q_1}$ no acotado y $H^{(p)} \in \mathbb{R}_p^n$ tales que

$$\forall Q \in \mathfrak{Q}' \quad V_l^{(p)*} = H^{(p)}.$$

III. Se asigna a cada $Q > 0$ el conjunto $\{G_\tau : \tau \in C(n,p)\}$, igual que en la discusión precedente al Teorema 2.6. Debido a que $\Pi^{(p)}$ está generado por los vectores $Q^{-c_\sigma}A_\sigma = Q^{-c_{i_1}}\mathbf{a}_{i_1} \wedge \dots Q^{-c_{i_p}}\mathbf{a}_{i_p}$ donde $\sigma = \{i_1 < \dots < i_p\} \in C(n,p)$, el Teorema 2.6 dice

$$\forall Q > 0 \quad \forall \sigma \in C(n,p) \quad \forall \tau \in C(n,p) \smallsetminus \{\tau_l\} \quad |\langle A_\sigma, G_\tau\rangle| \ll \nu_{l-1} Q^{c_\sigma}.$$

En consecuencia, como $\mathfrak{Q}' \subseteq \mathfrak{Q}_{>Q_1}$, de (2.17) se obtiene

$$\forall Q \in \mathfrak{Q}' \quad \forall \sigma \in C(n,p) \quad \forall \tau \in C(n,p) \smallsetminus \{\tau_l\} \quad |\langle A_\sigma, G_\tau\rangle| \ll \nu_l Q^{c_\sigma - \frac{\delta}{2}}.$$

Luego, existe $Q_2 > 0$ tal que $Q \in \mathfrak{Q}'_{>Q_2}$ implica $|\langle A_\sigma, G_\tau\rangle| < \nu_l Q^{c_\sigma}$ para $\sigma, \tau \in C(n,p)$, $\tau \neq \tau_l$. Así, cuando $Q \in \mathfrak{Q}'_{>Q_2}$, el conjunto $\{G_\tau : \tau \in C(n,p), \tau \neq \tau_l\}$ está contenido en el interior de $\nu_l \Pi^{(p)}(Q)$, por lo que $\operatorname{span}_{\mathbb{Q}}\{G_{\tau_1},\dots,G_{\tau_{l-1}}\} = \operatorname{span}_{\mathbb{Q}}\{V_1^{(p)},\dots,V_{l-1}^{(p)}\}$. Tras bautizar $G^*_{\tau_1},\dots,G^*_{\tau_l}$ a la base recíproca de $\{G_\tau : \tau \in C(n,p)\}$ se tiene $G^*_{\tau_l} \parallel V_l^{(p)*}$, pues

$$\operatorname{span}_{\mathbb{Q}}\{G^*_{\tau_l}\} = \{G_{\tau_1},\dots,G_{\tau_{l-1}}\}^\perp = \left\{V_1^{(p)},\dots,V_{l-1}^{(p)}\right\}^\perp = \operatorname{span}_{\mathbb{Q}}\left\{V_l^{(p)*}\right\}.$$

IV. Para $Q \in \mathfrak{Q}'_{>Q_2}$ sea $G := G(Q) := \mathbf{g}^*_{d+1} \wedge \dots \wedge \mathbf{g}^*_n$. Por la Identidad de Laplace (Apéndice B, Lema 6),

$$\forall \tau \in C(n,p) \quad \left\langle G_\tau, G^*_{\tau_l}\right\rangle = \langle G_\tau, G\rangle.$$

Debido a que $\{G_\tau : \tau \in C(n,p)\}$ es una base de $\mathbb{R}_p^n$, $G = G^*_{\tau_l}$. Esto conlleva $G \parallel V_l^{(p)*} = H^{(p)}$, implicando que $S^* = \operatorname{span}_{\mathbb{Q}}\{\mathbf{g}^*_{d+1}(Q),\dots,\mathbf{g}^*_n(Q)\}$ sea el mismo subespacio para cualquier $Q \in \mathfrak{Q}'_{>Q_2}$ (cfr. Lema 2.7). Por lo tanto, con $\mathfrak{Q}_1 := \mathfrak{Q}'_{>Q_2}$,

$$\forall Q \in \mathfrak{Q}_1 \quad \operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1(Q),\dots,\mathbf{g}_d(Q)\} = \left(\operatorname{span}_{\mathbb{Q}}\{\mathbf{g}^*_{d+1}(Q),\dots,\mathbf{g}^*_n(Q)\}\right)^\perp = (S^*)^\perp =: S^d.$$

V. Para poder concluir hay que erradicar la hipótesis $\|\mathbf{c}\|_\infty \leq n^{-1}$. Supóngase que todas las hipótesis son satisfechas, pero que $\|\mathbf{c}\| > n^{-1}$ y considérese el significado de $\Pi(Q)$, $\mathfrak{Q}$ y $\lambda_1(Q),\dots,\lambda_n(Q)$ como en el enunciado del Teorema. Con $\tilde{\mathbf{c}} := (n\|\mathbf{c}\|_\infty)^{-1}\mathbf{c}$ es claro que

$$\begin{aligned}\forall Q > 0 \qquad \widetilde{\Pi}(Q) &:= \left\{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |L_j(\mathbf{x})| \leq Q^{\tilde{c}_j}\right\} \\ &= \left\{\mathbf{x} \in \mathbb{R}^n : \forall j \in [1..n] \quad |L_j(\mathbf{x})| \leq Q^{\frac{c_j}{n\|\mathbf{c}\|_\infty}}\right\} = \Pi\left(Q^{\frac{1}{n\|\mathbf{c}\|_\infty}}\right).\end{aligned}$$

De este modo, si $\tilde{\lambda}_1(Q) < \dots < \tilde{\lambda}_n(Q)$ son los mínimos sucesivos de $\widetilde{\Pi}(Q)$, la igualdad anterior da

$$\forall Q > 0 \qquad \forall j \in [1..n] \qquad \tilde{\lambda}_j(Q) = \lambda\left(Q^{\frac{1}{n\|\mathbf{c}\|_\infty}}\right). \tag{2.18}$$

Entonces, con $\widetilde{\mathfrak{Q}} := \{Q^{n\|\mathbf{c}\|_\infty} : Q \in \mathfrak{Q}\}$ y $\tilde{\delta} := (n\|\mathbf{c}\|_\infty)^{-1}\delta > 0$,

$$\forall Q \in \widetilde{\mathfrak{Q}} \qquad \tilde{\lambda}_d(Q) = \lambda_d\left(Q^{\frac{1}{n\|\mathbf{c}\|_\infty}}\right) < \frac{\lambda_{d+1}\left(Q^{\frac{1}{n\|\mathbf{c}\|_\infty}}\right)}{\left(Q^{\frac{1}{n\|\mathbf{c}\|_\infty}}\right)^\delta} = \frac{\tilde{\lambda}_{d+1}(Q)}{Q^{\tilde{\delta}}}.$$

Así, por los cuatro puntos anteriores, hay un subespacio propio $S \subseteq \mathbb{Q}^n$ y $\widetilde{\mathfrak{Q}}_1 \subseteq \widetilde{\mathfrak{Q}}$ no acotado tales que para $Q \in \widetilde{\mathfrak{Q}}_1$ los primeros d mínimos sucesivos de $\widetilde{\Pi}(Q)$ se alcanzan en $\tilde{\mathbf{g}}_1(Q), \ldots, \tilde{\mathbf{g}}_d(Q)$ y $S = \operatorname{span}_{\mathbb{Q}}\{\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_d\}$. Definiendo $\mathfrak{Q}_1 := \{Q^{\frac{1}{n\|\mathbf{c}\|_\infty}} : Q \in \widetilde{\mathfrak{Q}}_1\} \subseteq \mathfrak{Q}$ que no es acotado, pues $\widetilde{\mathfrak{Q}}_1$ no lo es, y, en vista de (2.18), con $\mathfrak{Q}_1$ y S se obtiene la conclusión deseada. □

2.4. Construcción del Teorema Fuerte del Subespacio

En términos de la prueba del Teorema Fuerte del Subespacio, no es trivial ver cómo se usarán los métodos expuestos en el Teorema de Roth para justificar la existencia de $H^{(p)}$. Esto se probará asignando a cada $Q > 1$ un paralelepípedo y a éste, un subespacio unidimensional relacionado con el último mínimo sucesivo y se verá que para un subconjunto no acotado de $\mathbb{R}_{>1}$, este subespacio permanece fijo.

A pesar de que se busca un objeto en $\mathbb{R}_p^n$, bastará trabajar en $\mathbb{R}^n$. El Teorema 2.14 es parte fundamental de la prueba. Esta proposición traduce un problema de mínimos sucesivos a uno de polinomios en varias variables. Es en la construcción de este polinomio en la que las ideas inspiradas en el trabajo de Roth resolverán varias vicisitudes.

La longitud y los recovecos del largo razonamiento pueden ocultar las ideas que lo conforman. Es pertinente ver una guía general sobre el edificio que está por estudiarse. La prueba se divide en dos grandes partes. Después de definir un nuevo índice de polinomios, ahora con respecto a formas lineales y un multi-índice, se trabaja a grandes rasgos bajo el siguiente esquema:

I. Construcciones geométricas

1. $\Pi(Q)$ es el paralelepípedo generado por $Q^{-c_1}L_1, \ldots, Q^{-c_n}L_n$ para cada $Q > 1$, sus mínimos sucesivos son $\lambda_1, \ldots, \lambda_n$ y $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$ es una base obtenida de la definición de mínimos sucesivos. El objetivo es encontrar $\mathbf{h} \in \mathbb{R}^n$ tal que $\mathbf{g}_n = \mathbf{h}$ para una infinidad no acotada Q suponiendo $\lambda_{n-1} < \lambda_n Q^{-\delta}$. Es fundamental notar que $\mathfrak{m}(\Pi(Q)) = \mathfrak{m}(\Pi(1)) = 2^n|\det(L_1, \ldots, L_n)|^{-1}$ para cualquier $Q > 0$ por propiedades del determinante y $c_1 + \ldots + c_n = 0$.
2. Se demuestra el resultado bajo la suposición $\lambda_{n-1} < Q^{-\delta}$. Esto se prueba por contradicción con un polinomio auxiliar y un argumento similar al del Teorema de Roth.
3. Usando el Lema de Davenport se concluye la existencia de $\mathbf{h}$ suponiendo $\lambda_{n-1} < \lambda_n Q^{-\delta}$.

II. El polinomio auxiliar

1. A partir de un sistema de ecuaciones lineales y el Lema de Siegel, se construye un polinomio, P, que satisface propiedades de homogeneidad y con altura acotada. Este resultado, el Teorema 2.18, corresponde al Teorema 1.4.
2. El Teorema 1.5 tiene su análogo en el Teorema 2.19. Con respecto a ciertas formas lineales $M_1, \ldots, M_n$ con coeficientes en $\mathbb{Z}$ coprimos y un $\mathbf{0} \leq \mathbf{r} \in \mathbb{Z}^m$ el índice de P es acotado inferiormente. Aquí se echa mano de las mallas, conjuntos finitos que determinarán si los polinomios se anulan en ciertos subespacios lineales.
3. Finalmente, se acota superiormente el índice de P con respecto a $L_1, \ldots, L_n$ con el Teorema 2.25, análogo al Lema Fundamental de Roth.

2.4.1. Índices de polinomios

En el capítulo anterior se construyó un polinomio auxiliar cuyo índice se sujetaba a ciertas desigualdades que, suponiendo falso el Teorema de Roth, desembocaron en una contradicción. En esta sección se definirá otro índice para polinomios con respecto a formas lineales y puntos. Para distinguir entre estos dos objetos se llamará al ya conocido **índice de Roth**.

En adelante $n, m \in \mathbb{N}$ serán fijos, $\mathbf{r} \in \mathbb{N}^m$,

$$\mathbf{X} = X_{1,1}, \ldots, X_{1,n}; X_{2,1}, \ldots, X_{2,n}; \ldots; X_{m,1}, \ldots, X_{m,n}$$

y $\mathbb{R}[\mathbf{X}]$ será el anillo

$$\mathbb{R}[\mathbf{X}] := \mathbb{R}[X_{1,1}, \ldots, X_{1,n}; X_{2,1}, \ldots, X_{2,n}; \ldots; X_{m,1}, \ldots, X_{m,n}].$$

Dada una colección de formas lineales $L_1, \ldots, L_m : \mathbb{R}^n \to \mathbb{R}$ con coeficientes algebraicos se define para cada $c \in \mathbb{R}$ el ideal[6]

$$I(c) := \left\langle \left\{ L_1^{i_1} L_2^{i_2} \cdots L_m^{i_m} : i_1, i_2, \ldots, i_m \in \mathbb{N}_0, \sum_{j=1}^{m} \frac{i_j}{r_j} \geq c \right\} \right\rangle.$$

Nótese que $c_1 > c_2$ implica $I(c_1) \subseteq I(c_2)$.

Definición 2.5. *Sean $P \in \mathbb{R}[\mathbf{X}]$ y $\mathbf{L} = (L_1, \ldots, L_m)$. Si $P \neq 0$, el **índice de P con respecto a $\mathbf{L}$ y $\mathbf{r}$** es*

$$\operatorname{ind} P := \operatorname{ind}_{\mathbf{r}}(P; \mathbf{L}) := \sup\{c \geq 0 : P \in I(c)\}$$

y si $P = 0$, $\operatorname{ind} P := +\infty$.

Las formas $L_1, \ldots, L_n : \mathbb{R}^n \to \mathbb{R}$ dadas por

$$\forall h \in [1..m] \quad L_h(X_{h,1}, \ldots, X_{h,n}) = \alpha_{h,1} X_{h,1} + \ldots + \alpha_{1,n} X_{h,n},$$

se mantendrán fijas por lo que resta de la sección. No se pierde generalidad al suponer $\alpha_{h,1} \neq 0$ para $h \in [1..m]$, por lo que cada $X_{h,1}$ se expresa en términos de $L_h, X_{h,2}, \ldots, X_{h,n}$ y

$$P = \sum c(j_1, a_{1,2}, \ldots, a_{1,n}; \ldots; j_m, a_{m,2}, \ldots, a_{m,n}) L_1^{j_1} X_{1,2}^{a_{1,2}} \cdots X_{1,n}^{a_{1,n}} \cdots L_m^{j_m} X_{m,2}^{a_{m,2}} \cdots X_{m,n}^{a_{m,n}}. \tag{2.19}$$

Escribiendo $\mathbf{j} = (j_1, \ldots, j_m)$ para $\mathbf{j} \in \mathbb{N}_0^m$ se define

$$\mu := \min\left\{ \mathbf{j} \cdot \mathbf{r}^{-1} | : \mathbf{j} \in \mathbb{N}_0^m, \; c(j_1, a_{1,2}, \ldots; \ldots; j_n, \ldots a_{m,n}) \neq 0 \right\}.$$

Entonces, se tiene $P \in I(\mu)$ y $\mu \leq \operatorname{ind} P$. Por otra parte, como $P \in I(\operatorname{ind} P)$, cada término en (2.19) con coeficiente no nulo, indicado por el multi-índice $\mathbf{j}$, satisface

$$\mathbf{j} \cdot \mathbf{r}^{-1} = \sum_{i=1}^{n} \frac{j_i}{r_i} \geq \operatorname{ind} P \implies \mu \geq \operatorname{ind} P \qquad \therefore \mu = \operatorname{ind} P. \tag{2.20}$$

Para cualesquiera $P, Q \in \mathbb{R}[\mathbf{X}]$ vale $P, Q \in I(\min\{\operatorname{ind} P, \operatorname{ind} Q\})$ y, por ser un ideal, $P + Q \in I(\min\{\operatorname{ind} P, \operatorname{ind} Q\})$, por lo que $\operatorname{ind}(P + Q) \geq \min\{\operatorname{ind} P, \operatorname{ind} Q\}$. Escríbase $P =$

[6] $\langle L \rangle$ significa el ideal generado por L.

$P_1 + P_2$ en donde P_1 comprende a los términos para los que $\sum_{h=1}^{m} j_h/r_h = \operatorname{ind} P$ y P_2, a los que tienen $\sum_{h=1}^{m} j_h/r_h > \operatorname{ind} P$ y, similarmente, $Q = Q_1 + Q_2$; entonces,

$$PQ = (P_1 + P_2)(Q_1 + Q_2) = P_1Q_1 + P_1Q_2 + P_2Q_1 + P_2Q_2.$$

En consecuencia, por la caracterización del índice dada en (2.20), $\operatorname{ind}(PQ) = \operatorname{ind}(P) + \operatorname{ind}(Q)$.

Lema 2.10. *Para cualesquiera $P, Q \in \mathbb{R}[\mathbf{X}]$ se cumple, con respecto a $L_1, \ldots, L_m, \mathbf{r}$,*

1. $\operatorname{ind}(P+Q) \geq \min\{\operatorname{ind} P, \operatorname{ind} Q\}$,
2. $\operatorname{ind}(PQ) = \operatorname{ind} P + \operatorname{ind} Q$.

El Lema 2.10 debe parecer familiar, corresponde al Lema 1.3. No obstante, la proposición del capítulo anterior consta de tres incisos. La propiedad faltante también tiene su contraparte, mas ella exige nueva notación. Igual que antes, cuando

$$\mathfrak{T} = (i_{1,1}, \ldots, i_{1,n}; \ldots; i_{m,1}, \ldots, i_{m,n}) \in \mathbb{Z}^{mn}$$

tenga entradas no negativas, el operador diferencial $(\cdot)_{\mathfrak{T}} : \mathbb{R}[\mathbf{X}] \to \mathbb{R}[\mathbf{X}]$ es

$$\forall P \in \mathbb{R}[\mathbf{X}] \quad P_{\mathfrak{T}} := \frac{1}{i_{1,1}! \cdots i_{m,n}!} \frac{\partial^{|\mathfrak{T}|} P}{\partial X_{1,1}^{i_{1,1}} \cdots \partial X_{m,n}^{i_{m,n}}},$$

donde $|\mathfrak{T}| = i_{1,1} + \ldots + i_{m,n}$. Recordando que $\mathbf{r} \in \mathbb{N}^m$, no existe problema alguno en definir

$$\frac{\mathfrak{T}}{\mathbf{r}} := \sum_{j=1}^{m} \frac{i_{j,1} + \ldots + i_{j,n}}{r_j}.$$

A pesar de chocar con el cociente de multi-índices cuando $m = 1$, no hay riesgo de confusión con esta escritura. El contexto y la tipografía dejarán sin ambigüedades el significado de la división.

Para cada $h \in [1..m]$ la forma $L_h : \mathbb{R}^n \to \mathbb{R}$ se extiende a $\mathbb{R}^{nm}$ con la misma regla de correspondencia:

$$L_h(X_{1,1}, \ldots, X_{1,n}; \ldots; X_{m,1}, \ldots, X_{m,n}) = \alpha_{h,1}X_{h,1} + \ldots + \alpha_{h,n}X_{h,n}.$$

En los Lemas 2.11 y 2.12 se considera a las extensiones en $\mathbb{R}^{nm}$ y a $T \subseteq \mathbb{R}^{nm}$ como

$$T = \bigcap_{h=1}^{m} \ker L_h.$$

Nótese que $\dim T = m(n-1)$.

Lema 2.11. *Con respecto a $L_1, \ldots, L_m$ y $\mathbf{r}$ se cumple*

$$\forall P \in \mathbb{R}[\mathbf{X}] \quad \forall \mathfrak{T} \in \mathbb{N}_0^{mn} \quad \operatorname{ind} P_{\mathfrak{T}} \geq \operatorname{ind} P - \frac{\mathfrak{T}}{\mathbf{r}}.$$

Además, $\operatorname{ind} P > \frac{\mathfrak{T}}{\mathbf{r}}$ *implica* $P_{\mathfrak{T}}\big|_T = 0$.

Demostración. Con $\mathbf{a} = (a_{1,2}, \dots, a_{1,n}; \dots; a_{m,2}, \dots, a_{m,n})$ se redacta la fórmula (2.19) de manera más breve:

$$P = \sum_{\mathbf{j},\mathbf{a}} c(\mathbf{j},\mathbf{a}) L_1^{j_1} \cdots L_m^{j_m} X_{1,2}^{a_{1,2}} \cdots X_{1,n}^{a_{1,n}} \cdots X_{m,2}^{a_{m,2}} \cdots X_{m,n}^{a_{m,n}}.$$

Aplicando $(\cdot)_{\mathfrak{T}}$ a P con $\mathfrak{T} = (i_{1,1}, \dots, i_{1,n}; \dots; i_{m,1}, \dots, i_{m,n})$ se llega a

$$\begin{aligned} P_{\mathfrak{T}} &= \sum_{\mathbf{j},\mathbf{a}} c(\mathbf{j},\mathbf{a}) \left[(X_{1,2}^{a_{1,2}} \cdots X_{m,n}^{a_{m,n}})_{\mathfrak{T}} L_1^{j_1} \cdots L_m^{j_m} + X_{1,2}^{a_{1,2}} \cdots X_{m,n}^{a_{m,n}} (L_1^{j_1} \cdots L_m^{j_m})_{\mathfrak{T}} \right] \\ &= \sum_{\mathbf{j},\mathbf{a}} c(\mathbf{j},\mathbf{a}) \left[(X_{1,2}^{a_{1,2}} \cdots X_{m,n}^{a_{m,n}})_{\mathfrak{T}} L_1^{j_1} \cdots L_m^{j_m} + K_{\mathbf{j}} X_{1,2}^{a_{1,2}} \cdots X_{m,n}^{a_{m,n}} L_1^{j_1 - (i_{1,1}+\dots+i_{1,n})} \cdots L_m^{j_m - (i_{m,1}+\dots+i_{m,n})} \right], \end{aligned}$$

con $K_{\mathbf{j}}$ constantes adecuadas. Si en algún $\mathbf{j}$ los exponentes de las formas son negativos, $K_{\mathbf{j}} = 0$ y la desigualdad buscada sigue de la caracterización 2.20.

Cuando $\operatorname{ind} P > \frac{\mathfrak{T}}{\mathbf{r}}$ vale $\operatorname{ind} P_{\mathfrak{T}} > 0$. Representando a $P_{\mathfrak{T}}$ como en (2.19) y usando (2.20) se ve que en cada sumando al menos una forma tendrá exponente es positivo. En otras palabras, P como polinomio en $L_1, \dots, L_m$ no tiene términos constantes. Por lo tanto, cualquier $\mathbf{x} \in T$ satisface $P_{\mathfrak{T}}(\mathbf{x}) = 0$. □

Lema 2.12. *Con respecto a $L_1, \dots, L_m, \mathbf{r}$ se tiene*

$$\forall P \in \mathbb{R}[\mathbf{X}] \setminus \{0\} \quad \exists\, \mathfrak{T} \in \mathbb{N}_0^{mn} \quad \left(\operatorname{ind} P = \frac{\mathfrak{T}}{\mathbf{r}} \right) \ \&\ \left(P_{\mathfrak{T}}\big|_T \neq 0 \right).$$

Además, cuando $\alpha_{h,1} \neq 0$ para $h \in [1..m]$, puede elegirse $\mathfrak{T}$ de la forma

$$\mathfrak{T} = (i_1, 0, \dots, 0; i_2, 0, \dots, 0; \dots; i_m, 0, \dots, 0).$$

Demostración. Agrupando los términos de (2.19) se producen polinomios en

$$X_{1,2} \dots, X_{1,n}, \dots, X_{m,2}, \dots, X_{m,n}$$

que dependen de los multi-índices, esto da la representación

$$P = \sum_{\mathbf{j}} R(\mathbf{j}; X_{1,2} \dots, X_{1,n}; \dots; X_{m,2}, \dots, X_{m,n}) L_1^{j_1} \cdots L_m^{j_m}.$$

Por la definición del índice y la elaboración de los polinomios R, cuando $\mathbf{j} \cdot \mathbf{r}^{-1} < \operatorname{ind} P$, $R = 0$. Entonces, la caracterización $\mu = \operatorname{ind} P$ da la existencia de $0 \leq \mathbf{j}' \in \mathbb{Z}^m$ tal que $R(\mathbf{j}'; X_{1,2} \dots, X_{1,n}; \dots; X_{m,2}, \dots, X_{m,n}) \neq 0$ y $\mathbf{j}' \cdot \mathbf{r}^{-1} = \operatorname{ind} P$.

Suponiendo que $\alpha_{h,1} \neq 0$ para toda $h \in [1..m]$ (si no fuese así, simplemente se elige en cada forma una variable con coeficiente no 0) y se define

$$\mathfrak{T} = (j_1', 0, \dots, 0; j_2', 0, \dots, 0; \dots; j_n', 0, \dots, 0).$$

Entonces, como $\dim T = m(n-1)$ y $P_{\mathfrak{T}} = KR(\mathbf{j}'; X_{1,2}, \dots, X_{m,n})$ con $0 \neq K \in \mathbb{R}$ es un polinomio en $n(m-1)$ variables distinto de cero, se cumple $P_{\mathfrak{T}}|_T = KR(\mathbf{j}'; X_{1,2}, \dots, X_{m,n}) \neq 0$. □

La discusión previa revela similitudes imposibles de ignorar entre el índice de Roth y el recién definido. Los parecidos se explican porque el índice de Roth puede entenderse como un caso particular del actual ([Ba01], p. 70).

2.4.2. Construcciones geométricas

En esta sección, para cada $h \in [1..m]$ las funciones $L_h(\mathbf{X}) = \langle \boldsymbol{\alpha}_h, \mathbf{X} \rangle$ serán linealmente independientes con coeficientes en los enteros algebraicos. Los vectores $\boldsymbol{\alpha}_1^*, \ldots, \boldsymbol{\alpha}_n^*$ denotarán a la base recíproca de $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$; $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ cumplirá $\|\mathbf{c}\|_\infty \leq 1$, $c_1 + \ldots + c_n = 0$ y $\delta > 0$. Nótese que las entradas de $\boldsymbol{\alpha}_1^*, \ldots, \boldsymbol{\alpha}_n^*$ también son números algebraicos.

Ya establecida la simbología, nuevos conjuntos son construidos. Si todas las entradas de $\mathbf{c}$ fuesen menores que $-\delta/2$, su suma sería estrictamente negativa; luego,

$$\mathfrak{S} := \left\{ i \in [1..n] : c_i + \frac{\delta}{2} \geq 0 \right\} \neq \varnothing. \tag{2.21}$$

Para cada $Q > 1$ el paralelepípedo $\Pi(Q)$ es el generado por $\{Q^{-c_1}\boldsymbol{\alpha}_1, \ldots, Q^{-c_n}\boldsymbol{\alpha}_n\}$ y $\Pi^*(Q)$, el recíproco:

$$\begin{aligned} \forall Q > 1 \quad \Pi(Q) &:= \{ \mathbf{x} \in \mathbb{R} : \forall i \in [1..n] \,|\, |\langle \boldsymbol{\alpha}_i, \mathbf{x} \rangle| \leq Q^{c_i} \}, \\ \forall Q > 1 \quad \Pi^*(Q) &:= \{ \mathbf{x} \in \mathbb{R} : \forall i \in [1..n] \,|\, |\langle \boldsymbol{\alpha}_i^*, \mathbf{x} \rangle| \leq Q^{-c_i} \}. \end{aligned}$$

Los mínimos sucesivos de $\Pi(Q)$ son $\lambda_1(Q), \ldots, \lambda_n(Q)$ y $\mathbf{g}_1(Q), \ldots, \mathbf{g}_n(Q) \in \mathbb{Z}^n$ son una base de $\mathbb{R}^n$ tal que

$$\forall i \in [1..n] \quad \mathbf{g}_i(Q) \in \lambda_i(Q)\Pi(Q).$$

A la base recíproca de $\mathbf{g}_1(Q), \ldots, \mathbf{g}_n(Q)$ se le denota $\mathbf{g}_1^*(Q), \ldots, \mathbf{g}_n^*(Q)$. Expresar todo el tiempo que cada uno de estos objetos depende de Q dificulta la escritura, a veces la dependencia será omitida.

Resultados principales

Lo que se busca es probar la existencia de $H^{(p)}$ en la prueba del Teorema Fuerte del Subespacio. Trabajar en $\mathbb{R}_p^n$ daría complicaciones innecesarias; de este modo, se piensa el problema simplemente en $\mathbb{R}^n$ con el producto interior usual. La notación precedente permite una formulación más clara de la meta: el Lema 2.17.

Bajo $\lambda_{n-1} < Q^{-\delta}$, el Teorema 2.14 da un conjunto no acotado cuyos elementos, Q, satisfacen $\mathbf{g}_n^*(Q) \in \{\alpha_i^* : i \in \mathfrak{S}\}^\perp$. Con un real $\overline{Q}$ que cumpla esto, se fija un múltiplo entero adecuado de $\mathbf{g}_n^*(\overline{Q})$, $\mathbf{h} \in \mathbb{Z}^n$, y se concluye que $\mathbf{h} \in \Pi^*(Q)$ para Q en un subconjunto no acotado de $\mathbb{R}_{>0}$. Si se varía a $\mathbf{c}$, $\mathbf{c} = \mathbf{c}(Q)$, también puede encontrarse un vector que se comporte de manera similar a $\mathbf{h}$ para nuevos paralelepípedos $\widetilde{\Pi}(Q)$. Esta variación permite modificar la hipótesis del Lema 2.15 para llegar al Lema 2.17.

Lema 2.13. *Sea $\mathfrak{B} \subseteq \mathbb{R}_{>1}$ no acotado. Si existe $i \in \mathfrak{S}$ tal que $\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle \neq 0$ para toda $Q \in \mathfrak{B}$ y toda $Q \in \mathfrak{B}$ satisface $\lambda_{n-1} < Q^{-\delta}$; entonces, existen constantes $C_1, C_2, C_3 > 0$ dependientes de $\delta, L_1, \ldots, L_n, \mathbf{c}$ tales que*

$$\forall Q > C_3 \qquad Q^{C_1} \leq \|\mathbf{g}_n^*(Q)\|_\infty \leq Q^{C_2}. \tag{2.22}$$

Una desigualdad sigue de manera más o menos directa del segundo Teorema de Minkowski. La otra requiere de algunas ideas algebraicas, específicamente, de campos numéricos y enteros algebraicos.

Teorema 2.14 (El Teorema del Penúltimo Mínimo)**.** *Sea $\mathfrak{B} \subseteq \mathbb{R}_{>1}$ no acotado. Si para toda $Q \in \mathfrak{B}$ se cumple $\lambda_{n-1} < Q^{-\delta}$, entonces existe $Q_1 = Q_1(\delta; L_1, \ldots, L_n; c_1, \ldots, c_n) > 0$ tal que*

$$\forall\, Q > Q_1 \qquad \forall i \in \mathfrak{S} \quad \langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle = 0. \tag{2.23}$$

A primera vista, el Teorema 2.14 va en un sentido contrario del objetivo, dice que el promisorio Lema 2.13 no es aplicable para Q suficientemente grande. Lejos de desilusionar, esta observación sugiere que la prueba del Teorema 2.14 será por contradicción. Una larga construcción que emula a la prueba del Teorema de Roth se utiliza en este resultado.

Lema 2.15. *Sea* $\mathfrak{D} \subseteq \left\{Q \in \mathbb{R}_{>1} : \lambda_{n-1} < Q^{-\delta}\right\}$ *no acotado. Existen* $\mathbf{h} \in \mathbb{R}^n$ *fijo y* $\mathfrak{D}' \subseteq \mathfrak{D}$ *no acotado tales que*

$$\forall Q \in \mathfrak{D}' \quad \mathbf{g}_n^* = \mathbf{h}.$$

Nótese que este lema dice que, aunque los vectores $\mathbf{g}_1, \ldots, \mathbf{g}_n$ no son únicamente determinados forzosamente, $\mathbf{g}_n^*$ sí lo es.

Mientras que el Lema 2.15 el vector $\mathbf{c}$ permanece fijo, el Lema 2.16 afirma (después de aplicar log_Q a los números A_i) que si $\mathbf{c}$ depende de Q, $c_1 + \ldots + c_n = 0$ y $\|\mathbf{c}\|_\infty \leq 1$, la conclusión del Lema 2.15 sigue siendo válida.

Lema 2.16. *Sea* $\mathfrak{B}$ *un conjunto no acotado de reales y supóngase que para cada* $Q \in \mathfrak{B}$ *existen reales positivos* $A_1 = A_1(Q), \ldots, A_n = A_n(Q)$ *que satisfacen*

$$A_1 A_2 \cdots A_n = 1, \quad \max\left\{A_1, \ldots, A_n, \frac{1}{A_1}, \ldots, \frac{1}{A_n}\right\} \leq Q. \tag{2.24}$$

Sean $\widetilde{\Pi} = \widetilde{\Pi}(Q)$ *el paralelepípedo generado por* $\{A_1^{-1}\boldsymbol{\alpha}_1, \ldots, A_n^{-1}\boldsymbol{\alpha}_n\}$, $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ *los mínimos sucesivos de* $\widetilde{\Pi}$, $\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_n \in \mathbb{Z}$ *una base de* $\mathbb{R}^n$ *tal que* $\tilde{\mathbf{g}}_i \in \lambda_i \Pi$ *para* $i \in [1..n]$ *y* $\tilde{\mathbf{g}}_1^*, \ldots, \tilde{\mathbf{g}}_n^*$ *su base recíproca.*

Si para cada $Q \in \mathfrak{B}$ *se cumple* $\tilde{\lambda}_{n-1} < Q^{-\delta}$*; entonces, existen* $\mathbf{h} \in \mathbb{R}^n$ *fijo y* $\mathfrak{B}' \subseteq \mathfrak{B}$ *no acotado tales que*

$$\forall Q \in \mathfrak{B}' \quad \tilde{\mathbf{g}}_n^* = \mathbf{h}$$

Finalmente, el Lema de Davenport asocia a cada $\Pi(Q)$ un nuevo paralelepípedo, Π', al que se le aplica el Lema 2.16. Con un argumento de perpendicularidad y el Principio del Palomar de Dirichlet se alcanza la conclusión del Lema 2.17.

Lema 2.17. *Sea* $\mathfrak{F} := \left\{Q \in \mathbb{R}_{>1} : \lambda_{n-1} < \lambda_n Q^{-\delta}\right\}$*. Existen* $\mathbf{h} \in \mathbb{R}^n$ *fijo y* $\mathfrak{F}' \subseteq \mathfrak{F}$ *no acotado tales que*

$$\forall Q \in \mathfrak{F}' \quad \mathbf{g}_n^* = \mathbf{h}.$$

Demostración del Lema 2.13

Demostración. Cota superior. Sea $Q \in \mathfrak{B}$. El Segundo Teorema de Minkowski y $\lambda_{n-1} < Q^{-\delta}$ dan

$$\frac{1}{\lambda_n} \ll \lambda_1 \cdots \lambda_{n-1} \leq \lambda_{n-1}^{n-1} < \frac{1}{Q^{\delta(n-1)}}.$$

donde se usa el hecho de que $\mathfrak{m}\Pi(Q)$ es constante (ver I.1. del esquema de la prueba al inicio de esta sección). Del Teorema 2.5, con $\mathbf{a}_i = \boldsymbol{\alpha}_i Q^{-c_i}$ y $\mathbf{a}_i^* = \boldsymbol{\alpha}_i^* Q^{c_i}$, para constantes que dependen de $L_1, \ldots, L_n$,

$$\forall i \in [1..n] \qquad |\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle| \ll \frac{1}{\lambda_n Q^{c_i}} \ll \frac{1}{Q^{\delta(n-1)+c_i}}, \tag{2.25}$$

Debido a que $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n \in \mathbb{R}^n$ es una base, para cada $\mathbf{x} \in \mathbb{R}^n$ hay exactamente una colección de reales- $\tilde{x}_1, \ldots, \tilde{x}_n$- tal que $\mathbf{x} = \sum_{j=1}^n \tilde{x}_j \boldsymbol{\alpha}_j$; en particular, $\mathbf{g}_n^* = \sum_{j=1}^n g_j \boldsymbol{\alpha}_j$.

La linealidad del producto interior y la definición de base recíproca aseguran que $g_i = \langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n \rangle$ para cualquier $i \in [1..n]$. En efecto,

$$\forall i \in [1..n] \quad \langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle = \left\langle \boldsymbol{\alpha}_i^*, \sum_{j=1}^{n} g_j \boldsymbol{\alpha}_j \right\rangle = \sum_{j=1}^{n} g_j \langle \boldsymbol{\alpha}_i^*, \boldsymbol{\alpha}_j \rangle = \sum_{j=1}^{n} g_j \delta_{ij} = g_i.$$

Tras notar que la función que $\|\cdot\|_A : \mathbb{R}^n \to \mathbb{R}$ dada por $\|\mathbf{x}\|_A = \max\{\tilde{x}_1, \ldots, \tilde{x}_n\}$ es una norma, se deduce de (2.25) y $\|\mathbf{c}\|_\infty \leq 1$ que

$$\forall j \in [1..n] \quad |g_j| = \left|\left\langle \boldsymbol{\alpha}_j^*, \mathbf{g}_n^* \right\rangle\right| \ll \frac{1}{Q^{\delta(n-1)+c_j}} = \frac{Q^{-c_j}}{Q^{\delta(n-1)}} \leq \frac{Q}{Q^{\delta(n-1)}}.$$

Luego,

$$\|\mathbf{g}_n^*\|_\infty \ll \|\mathbf{g}_n^*\|_A \ll \frac{Q}{Q^{\delta(n-1)}},$$

en donde las constantes dependen de $L_1, \ldots, L_n$ (por (2.25) y la equivalencia de las normas en $\mathbb{R}^n$). Así, como $\delta(n-1) > 0$, tomando[7] $C_2 = 2 - 2^{-1}(\delta(n-1))$ y a $Q \in \mathfrak{B}$ suficientemente grande, se tendrá $\|\mathbf{g}_n\|_\infty \leq Q^{C_2}$.

Cota inferior. Sea $Q \in \mathfrak{B}$. La definición de $\mathfrak{S}$ y (2.25) dan

$$\forall j \in \mathfrak{S} \quad \left|\left\langle \boldsymbol{\alpha}_j^*, \mathbf{g}_n^* \right\rangle\right| \ll \frac{1}{Q^{\delta(n-1)-\frac{\delta}{2}}} \leq \frac{1}{Q^{\frac{\delta}{2}}};$$

Por otra parte, para cada $j \in [1..n]$ escríbase $\boldsymbol{\alpha}_j^* = \left(\alpha_{j,1}^*, \ldots, \alpha_{j,n}^*\right)$, cuyas entradas son algebraicas, $\Delta_j^* := [\mathbb{Q}\left(\alpha_{j,1}^*, \ldots, \alpha_{j,n}^*\right) : \mathbb{Q}]$ y $\Delta^* := \max\{\Delta_1^*, \ldots, \Delta_n^*\}$. Por la Regla de Cramer, $\mathbf{g}_n^* \in \mathbb{Q}^n$ y por (2.3), los denominadores de sus entradas son a lo más $n!$. Entonces, con $i \in \mathfrak{S}$ el de la hipótesis, la cota para los denominadores de las entradas de $\mathbf{g}_n^*$ y el que haya sólo una cantidad finita de $\mathbb{Q}$-automorfismos de $\mathbb{Q}(\alpha_{1,1}^*, \ldots, \alpha_{n,n}^*)$ dan la existencia de constantes dependientes de $(\alpha_{i,j})_{i,j}$ y n tales que[8]

$$|N(\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle)| \gg 1.$$

Por otra parte, $\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n \rangle \neq 0$ tiene a lo más Δ_i conjugados y cada uno de ellos es en valor absoluto $\ll \|\mathbf{g}_n^*\|_\infty$, la cota se obtiene con un argumento que apela a la cantidad finita de $\mathbb{Q}$-automorfismos de $\mathbb{Q}(\alpha_{1,1}^*, \ldots, \alpha_{n,n}^*)$ y la desigualdad de Cauchy-Bunyakowski-Schwarz. Por consiguiente, retomando $|N(\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle)| \gg 1$, se tiene

$$\frac{1}{Q^{\frac{\delta}{2}}} \geq |\langle \boldsymbol{\alpha}_i^*, \mathbf{g}_n^* \rangle| \gg \frac{1}{\|\mathbf{g}_n\|_\infty^{\Delta_i^*-1}} \geq \frac{1}{\|\mathbf{g}_n\|_\infty^{\Delta^*-1}} \implies \|\mathbf{g}_n^*\|_\infty^{\Delta^*-1} \gg Q^{\frac{\delta}{2}} \implies \Delta^* > 1.$$

La última implicación es porque $\Delta^* \in \mathbb{N}$ y si fuese 1, $\mathfrak{B}$ estaría acotado (Q es cualquier elemento de $\mathfrak{B}$). Entonces, para $C_1 := \delta/(3(\Delta^* - 1))$ y $Q \in \mathfrak{B}$ suficientemente grande[9] vale

$$\|\mathbf{g}_n^*\|_\infty^{\Delta^*-1} \geq Q^{C_1}. \qquad \square$$

[7] Si K es la constante en $\|\mathbf{g}_n\|_\infty \ll Q^{1-\delta(n-1)}$, basta con que $Q^{\delta(n-1)/2} > K$ para tener $\|\mathbf{g}_n\|_\infty \leq Q^{C_2}$ con la C_2 propuesta.

[8] Cuando α sea un número algebraico, $N(\alpha)$ denota al producto de α y sus conjugados.

[9] Si K es la constante implícita, basta con $Q \geq K^{\frac{\delta}{6}(\Delta^*-1)}$.

Con las hipótesis y la notación del Lema 2.13, se considera a $M : \mathbb{R}^n \to \mathbb{R}$ dada por $M(\mathbf{x}) = \langle \mathbf{m}, \mathbf{x} \rangle$, donde $\mathbf{m} \in \{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\}^{\perp} \cap \mathbb{Z}^n$ tiene entradas coprimas. Por la definición de base recíproca, $\mathbf{m} \parallel \mathbf{g}_n^*$. Además, de la Regla de Cramer aplicada para calcular $\mathbf{g}_n^*$ con $E = \det(\mathbf{g}_1, \ldots, \mathbf{g}_n)$, se tiene $\mathbf{g}_n^* = (t/E)\mathbf{m}$ para algún $t \in \mathbb{Z}$. De aquí se sigue que $t \mid E$, pues

$$\langle \mathbf{g}_n, \mathbf{g}_n^* \rangle = \frac{t}{E} \langle \mathbf{g}_n, \mathbf{m} \rangle = 1 \implies \frac{E}{t} = \langle \mathbf{g}_n, \mathbf{m} \rangle \in \mathbb{Z}.$$

Llamando $F = E/t$ se tiene $1 \leq |F| \leq n!$ y $\mathbf{g}_n^* = F^{-1}\mathbf{m}$; en consecuencia, $\|\mathbf{m}\|_\infty \ll \|\mathbf{g}_n^*\|_\infty \leq \|\mathbf{m}\|_\infty$, que implica $\|\mathbf{g}_n^*\|_\infty \ll \mathcal{H}(M) \ll \|\mathbf{g}_n^*\|_\infty$. Luego, el Teorema 2.13 da constantes positivas C_4, C_5, C_6 sujetas a $\delta, L_1, \ldots, L_n, \mathbf{c}$ tales que

$$\forall Q > \mathfrak{B}_{>C_6} \qquad Q^{C_4} \leq \mathcal{H}(M) \leq Q^{C_5}. \tag{2.26}$$

Demostración del Teorema 2.14

En breve, la prueba es por contradicción. Primero, se supone que el conjunto de reales que no cumplen con la conclusión no es acotado y se especifican algunos parámetros. A partir de estos objetos se obtiene un polinomio auxiliar, P. El índice de P con respecto a ciertas formas lineales y un multi-índice estará acotado superior e inferiormente. Las dos cotas implicarán una desigualdad ridícula. Durante la prueba se hará referencia a un teorema posterior, pero la prueba de éste no depende de la demostración actual.

Bajo este resultado- en el polinomio auxiliar- subyace una larga construcción muy parecida a la hecha por Klaus Roth. Las pruebas de las proposiciones concernientes a las propiedades del polinomio construido son postergadas buscando enfatizar al edificio geométrico.

Demostración. I. Supóngase que no es acotado el conjunto

$$\mathfrak{A} := \left\{ Q > 1 : \left(\lambda_{n-1} < \frac{1}{Q^{\delta}} \right) \,\&\, \left(\exists\, i \in \mathfrak{S} \;\; \langle \mathbf{a}_i^*, \mathbf{g}_n^* \rangle \neq 0 \right) \right\}.$$

1. Se elige δ_1 tal que $0 < \delta_1 < \min\{1, \delta\}$.
2. Se toma $\varepsilon > 0$ tal que
$$16n^2\varepsilon < \delta_1. \tag{2.27}$$
3. Con $\Delta = \max\{\Delta_1, \ldots, \Delta_n\}$, $\Delta_j := [\mathbb{Q}(a_{j,1}, \ldots, a_{j,n}) : \mathbb{Q}]$ para $j \in [1..n]$, sea $m \in \mathbb{N}$ tal que
$$m > \frac{4\log(2n\Delta)}{\varepsilon^2}.$$
También, se considera
$$\omega := \tilde{\omega}(m, \varepsilon) = \frac{24}{2^m} \left(\frac{\varepsilon}{12} \right)^{2^{m-1}}.$$
4. Sean E y D como en el Teorema 2.18 (dependen únicamente de los los coeficientes de $L_1, \ldots, L_m$. Con C_4, C_5, C_6 como en (2.26) se toma $Q_1 \in \mathfrak{A}$ tal que
$$Q_1^{\frac{\omega C_4^2}{C_5}} \geq 2^{3mq^2}, \quad Q_1^{\frac{\omega C_4^2}{C_5}} \geq D^{mq^2}, \quad Q_1^{\varepsilon} > 2^n E, \quad Q_1^{\varepsilon} > n\left(\frac{1}{\varepsilon} + 1 \right).$$

5. Sean $Q_2, \ldots, Q_m \in \mathfrak{A}$ tales que

$$\begin{aligned} \forall h \in [1..m-1] \quad & \omega \log Q_{h+1} \geq 2 \log Q_h, \\ \forall h \in [1..m] \quad & Q_h^{\varepsilon} > 2^n E, \quad Q_h^{\varepsilon} > n\left(\frac{1}{\varepsilon} + 1\right) \\ \forall h \in [1..m] \quad & Q_h \geq C_6, \quad Q_h^{\frac{\omega C_4^2}{C_5}} \geq 2^{3mq^2} \end{aligned} \tag{2.28}$$

El primer juego de desigualdades implica $Q_m \geq Q_{m-1} \geq \ldots \geq Q_1$.

6. Se fija a $\mathbf{r} \in \mathbb{N}^m$ de modo que

$$\begin{aligned} & \varepsilon r_1 \log Q_1 \geq \log Q_m, \\ \forall h \in [2..m] \quad & r_h := \left\lfloor \frac{r_1 \log Q_1}{\log Q_h} \right\rfloor + 1. \end{aligned}$$

Nótese que las entradas de $\mathbf{r}$ satisfacen (2.37), porque

$$\begin{aligned} \forall h \in [2..m] \quad r_1 \log Q_1 &= \frac{r_1 \log Q_1}{\log Q_h} \log Q_h \\ &\leq r_h \log Q_h \\ &= \left\lfloor \frac{r_1 \log Q_1}{\log Q_h} \right\rfloor \log Q_h + \log Q_h \leq r_1 \log Q_1 + \log Q_h \\ &\leq r_1(1+\varepsilon) \log Q_1. \end{aligned} \tag{2.29}$$

II. La elección de ε, m y $\mathbf{r}$ permite usar el Teorema 2.18 para obtener un polinomio

$$P = P(X_{1,1}, \ldots, X_{1,n}; \ldots; X_{m,1}, \ldots, X_{m,n}).$$

En P se aplicará primero el Teorema 2.19. Para cada $h \in [1..m]$, $\{\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,n}\} \subseteq \mathbb{Z}^n$ es una base de $\mathbb{R}^n$ que satisface

$$\forall h \in [1..m] \quad \forall i \in [1..n] \quad \mathbf{g}_{h,i} \in \lambda_i(Q_h)\Pi(Q_h).$$

Con esto y la hipótesis sobre el penúltimo mínimo, cuando $h \in [1..m], i \in [1..n], t \in [1..n-1]$ se obtiene

$$|L_i(\mathbf{g}_{h,t})| \leq \lambda_t Q^{c_i} \leq \lambda_{n-1} Q^{c_i} \leq Q_h^{c_i - \delta}.$$

Entonces, por el Teorema 2.19 ($\delta \leftarrow \delta_1$, $L^{(i)} \leftarrow L_i$ para $i \in [1..n]$),

$$m\varepsilon \leq \operatorname{ind} P(M_1, \ldots, M_m; r_1, \ldots, r_m),$$

donde $M_h = M_h(\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,n-1})$ es la forma de la discusión posterior al Teorema 2.13.

III. Ahora, el índice anterior se acota inferiormente con el Teorema 2.20. Para verificar sus hipótesis, se ve que (2.29) conduce a

$$\forall h \in [1..m-1] \quad r_{h+1} \log Q_{h+1} \leq (1+\varepsilon) r_h \log Q_h,$$

que al reordenar y apelar a la primera parte de (2.28) se convierte en (2.40):

$$\forall h \in [1..m-1] \quad \omega r_h \geq \omega \frac{r_{h+1} \log Q_{h+1}}{(1+\varepsilon) \log Q_h} \geq \frac{2 \log Q_h}{\log Q_{h+1}} \frac{r_{h+1} \log Q_{h+1}}{(1+\varepsilon) \log Q_h} = \frac{2 r_{h+1}}{1+\varepsilon} \geq r_{h+1}.$$

La suposición sobre $\mathfrak{A}$ permite usar el Lema 2.13 y las líneas subsecuentes para llegar a

$$\forall h \in [1..m] \quad Q_h^{C_4} \le \mathcal{H}(M_h) \le Q_h^{C_5}.$$

Aplicando varias veces estas desigualdades y (2.29),

$$\forall h \in [1..m] \quad \mathcal{H}(M_h)^{r_h} \ge Q_h^{r_h C_4} \ge Q_1^{r_1 C_4} \ge \mathcal{H}(M_1)^{\frac{r_1 C_4}{C_5}}. \tag{2.30}$$

Llamando $\xi = C_4/C_5$, esto se reescribirse justo como (2.41):

$$\forall h \in [1..m] \quad \mathcal{H}(M_h)^{r_h} \ge \mathcal{H}(M_1)^{r_1 \xi}.$$

Entonces, usando la desigualdad de arriba y (2.28) en (2.30), se obtiene (2.42):

$$\forall h \in [1..m] \quad \mathcal{H}(M_h)^{\omega\xi} \ge Q_h^{\omega\xi C_4} \ge 2^{3mq^2}.$$

Finalmente, el Teorema 2.18 asegura $\mathcal{H}(P) \le D^{|\mathbf{r}|} \le D^{mr_1}$ que, dada la elección de Q_1, conduce a (2.43):

$$\mathcal{H}(P)^{q^2} \le D^{q^2 m r_1} \le Q_1^{\frac{\omega r_1 C_4^2}{C_5}} = Q_1^{\omega r_1 C_4 \xi} \le \mathcal{H}(M_1)^{\omega r_1 \xi}.$$

Así, el Teorema 2.25 es aplicable y su conclusión es

$$\operatorname{ind}(P; M_1, \ldots, M_m; r_1, \ldots, r_m) \le \varepsilon.$$

IV. Las dos desigualdades que involucran al índice conllevan $0 < m\varepsilon \le \varepsilon$ o $m \le 1$. Sin embargo, la elección de m–tomando en cuenta los valores posibles de Δ, n y ε– fuerza a que m exceda a 1. La constradicción proviene de suponer que $\mathfrak{A}$ no es acotado. □

Demostración del Lema 2.15

Demostración. El Lema 2.14 avala la existencia de $\overline{Q} \in \mathfrak{D}$ que satisface $\mathbf{g}_n^*(\overline{Q}) \in \{\boldsymbol{\alpha}_i^* : i \in \mathfrak{S}\}^\perp$. Multiplicando por un entero racional adecuado, se fija a $\widetilde{\mathbf{h}} \in \mathbb{Z}^n$, $\widetilde{\mathbf{h}} \neq \mathbf{0}$, con entradas coprimas y paralelo a $\mathbf{g}_n^*(\overline{Q})$. Así, existe $Q_1 > 0$ tal que $Q \in \mathfrak{D}_{>Q_1}$ implica $\widetilde{\mathbf{h}} \in \Pi^*(Q)$, pues

$$\forall Q \in \mathfrak{D}_{>Q_1} \quad \left(\forall i \in \mathfrak{S} \quad \left|\left\langle \boldsymbol{\alpha}_i^*, \widetilde{\mathbf{h}} \right\rangle\right| = 0\right) \;\&\; \left(\forall i \in [1..n] \setminus \mathfrak{S} \quad \left|\left\langle \boldsymbol{\alpha}_i^*, \widetilde{\mathbf{h}} \right\rangle\right| \ll 1 \le \frac{1}{Q^{c_i + \frac{\delta}{2}}} < \frac{1}{Q^{c_i}}\right).$$

Del Teorema de los Paralelepípedos Recíprocos de Mahler (Teorema 2.12) y la definición de $\mathfrak{D}$ sigue que

$$\forall Q \in \mathfrak{D} \quad \frac{1}{\lambda_2^*} \ll \lambda_{n-1} < \frac{1}{Q^\delta} \implies \forall Q \in \mathfrak{D} \quad Q^\delta \ll \lambda_2^*.$$

Luego, existe $Q_2 > 0$ tal que $Q \in \mathfrak{D}_{>Q_2}$ conlleva $\lambda_2^* > 1$; por ende, si $Q \in \mathfrak{D}$ y $Q > \max\{Q_1, Q_2\}$, todos los puntos enteros de $\Pi^*(Q)$ son múltiplos de $\tilde{\mathbf{h}}$.

Por (2.25) existe $Q_3 > 0$ tal que $\mathbf{g}_n^*(Q) \in \Pi^*(Q)$ cuando $Q \in \mathfrak{D}_{>Q_3}$. Tómense $Q \in \mathfrak{D}_{>Q_4}$, $Q_4 := \max\{Q_1, Q_2, Q_3\}$, y $\mathbf{m} \in \{\mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\}^\perp \cap \mathbb{Z}^n$ con entradas coprimas, entonces, con F como en el párrafo anterior a (2.26),

$$\mathbf{g}_n^* = \frac{1}{F}\mathbf{m} \quad \text{con } 1 \le |F| \le n!$$

A pesar de que $\mathbf{g}_n^*$ y F dependen de Q, se cumple $\mathbf{g}_n^* \parallel \widetilde{\mathbf{h}}$ porque $\mathbf{g}_n^* \in \Pi^*(Q)$. Entonces $\mathbf{m} \parallel \widetilde{\mathbf{h}}$ y, ya que $\mathbf{m}, \mathbf{h} \in \mathbb{Z}$ tienen entradas coprimas, $\mathbf{m} = \mathbf{h}$ o $\mathbf{m} = -\mathbf{h}$. En consecuencia, si $J_n = \{1, 2, \ldots, n!, -1, -2, \ldots, -n!\}$,

$$\forall Q \in \mathfrak{D}_{>Q_4} \quad \exists F \in J_n \quad \mathbf{g}_n^*(Q) = \pm \frac{1}{F} \widetilde{\mathbf{h}}.$$

Al variar a $Q \in \mathfrak{D}_{>Q_4}$, $\mathbf{g}_n^*(Q)$ asume una cantidad finita de valores. Por lo tanto, existen $\mathfrak{D}' \subseteq \mathfrak{D}$ no acotado y $\mathbf{h} \in \mathbb{R}^n$ tales que $\mathbf{g}_n^*(Q) = \mathbf{h}$ para $Q \in \mathfrak{D}'$. $\square$

Demostración del Lema 2.16

Para cada $Q \in \mathfrak{B}$ sea $\mathbf{c} = \mathbf{c}(Q) \in \mathbb{R}^n$ el vector dado por

$$\forall Q \in \mathfrak{B} \quad \forall i \in [1..n] \quad c_i(Q) := \frac{\log(A_i(Q))}{\log(Q)}.$$

Este vector satisface $c_1 + \ldots + c_n = 0$, $\|\mathbf{c}\|_\infty \leq 1$ y $A_j = Q^{c_j}$.

Sea $C = [-1, 1]^n$. No es complicado ver que una partición $\{P_i\}_{i \in I}$ de $[-1, 1]$ induce una en C, $\mathcal{P} = \{P_{i_1} \times \ldots \times P_{i_n} : i_1, \ldots, i_n \in I\}$. En particular, se considera a $\{P_i\}_{i \in I}$ integrada por intervalos de la forma $[c, c')$ y $[c, 1]$ tan fina que si $\mathbf{x}, \mathbf{y} \in C$ están en el mismo elemento de $\mathcal{P}$, entonces

$$|\mathbf{s}(\mathbf{x}) - \mathbf{s}(\mathbf{y})| < \frac{\delta}{2n},$$

donde $\mathbf{s} : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{s}(\mathbf{z}) = z_1 + \ldots + z_n$.

Sea $\tilde{C} = [d_1, d_1') \times \ldots \times [d_n, d_n') \in \mathcal{P}$[10], tal que $\mathbf{c}(Q) \in \tilde{C}$ para una infinidad no acotada de $Q \in \mathfrak{B}$, llámese $\mathfrak{B}_1$. Se define la colección

$$\forall j \in [1..n] \qquad \tilde{d}_j := d_j' - \frac{1}{n} \sum_{j=1}^{n} d_j'.$$

Nótese que $\tilde{d}_1 + \ldots + \tilde{d}_n = 0$. Ya que $c_j \leq d_j'$ y $s(d_1', \ldots, d_n') n^{-1} < 2^{-1}\delta$, valen $c_j \leq \tilde{d}_j + 2^{-1}\delta$ y

$$\forall j \in [1..n] \qquad c_j - \delta \leq \tilde{d}_j - \frac{\delta}{2}.$$

Sean $Q \in \mathfrak{B}$ con $\mathbf{c}(Q) \in \tilde{C}$ y $\tilde{\mathbf{g}}_1 = \tilde{\mathbf{g}}_1(Q), \ldots, \tilde{\mathbf{g}}_n = \tilde{\mathbf{g}}_n(Q)$ puntos que realizan los mínimos sucesivos de $\widetilde{\Pi}(Q)$, $\tilde{\lambda}_1 \leq \ldots \leq \tilde{\lambda}_n$. Retomando la hipótesis $\tilde{\lambda}_{n-1} < Q^{-\delta}$

$$\forall j \in [1..n-1] \quad \forall i \in [1..n] \quad |L_i(\mathbf{g}_j)| \leq \tilde{\lambda}_j Q^{c_j} \leq Q^{c_j - \delta} \leq Q^{\tilde{d}_j - \frac{\delta}{2}} = \frac{1}{Q^{\frac{\delta}{2}}} Q^{\tilde{d}_j}. \tag{2.31}$$

La desigualdad anterior implica que el penúltimo mínimo sucesivo de $\Pi(Q)' = \{\mathbf{x} \in \mathbb{R}^n : |L_j(\mathbf{x})| \leq Q^{\tilde{d}_j}\}$, λ_{n-1}', satisface $\lambda_{n-1}' \leq Q^{\delta/2}$. Entonces, por el Lema 2.15, existen $\mathfrak{B}_2 \subseteq \mathfrak{B}_1$ no acotado y $\mathbf{h}' \in \mathbb{R}^n$ tal que $\mathbf{g}_n'^*(Q) = \mathbf{h}$ para cualquier $Q \in \mathfrak{B}_1$, en donde $\mathbf{g}_1'(Q), \ldots, \mathbf{g}_n'(Q)$ realiza a los mínimos sucesivos de $\Pi(Q)'$ y $\mathbf{g}_1'^*(Q), \ldots, \mathbf{g}_n'^*(Q)$ es su base recíproca.

Por el Segundo Teorema de Minkowski y $\lambda_{n-1}' < Q^{-\delta/2}$, para $Q \in \mathfrak{B}_2$ suficientemente grande $\lambda_n' > 1 > Q^{-\delta/2}$. De (2.31) se obtiene que $\{\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_{n-1}\}, \{\mathbf{g}_1', \ldots, \mathbf{g}_{n-1}'\} \subseteq Q^{-\delta/2}\Pi'(Q)$ que está propiamente contenido en $\lambda_n'\Pi'(Q)$; luego, $\operatorname{span}_{\mathbb{Q}}\{\mathbf{g}_1', \ldots, \mathbf{g}_{n-1}'\} = \operatorname{span}_{\mathbb{Q}}\{\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_{n-1}\}$ y $\mathbf{g}_n^* \parallel \mathbf{g}_n' = \mathbf{h}'$. Argumentando como en el Lema 2.15, existe $\mathfrak{B}' \subseteq \mathfrak{B}_2$ no acotado tal que $\tilde{\mathbf{g}}_n = \mathbf{h}'$.

[10] Los intervalos pueden ser cerrados.

Demostración del Lema 2.17

Demostración. En esta prueba las constantes involucradas en los símbolos de Vinogradov dependenrán de $L_1, \dots, L_n$ y n.

I. Las condiciones (2.6), (2.7) y (10) del Lema de Davenport se satisfacen con

$$\forall Q \in \mathfrak{F} \qquad \rho_0 := \left(\lambda_1\lambda_2\cdots\lambda_{n-2}\lambda_{n-1}^2\right)^{\frac{1}{n}}, \qquad \forall i \in [1..n-1] \quad \rho_i := \frac{\rho_0}{\lambda_i}, \qquad \rho_n := \frac{\rho_0}{\lambda_{n-1}}.$$

En consecuencia, hay una permutación de $[1..n]$– $t_1, \dots, t_n$– tal que los mínimos sucesivos de

$$\Pi'(Q) = \left\{\mathbf{x} \in \mathbb{R}^n : \forall i \in [1..n] \quad \frac{|L_i(\mathbf{x})|}{Q^{c_i}} \le \frac{1}{\rho_{t_i}}\right\},$$

llamados $\lambda'_1, \dots, \lambda'_n$, satisfacen $\lambda_j\rho_j \ll \lambda'_j \ll \lambda_j\rho_j$.

II. Por el Segundo Teorema de Minkowski y la definición de $\mathfrak{F}$,

$$\forall Q \in \mathfrak{F} \quad \lambda'_{n-1} \ll \rho_{n-1}\lambda_{n-1} = \rho_0 = (\lambda_1\lambda_2\cdots\lambda_{n-1}\lambda_n)^{\frac{1}{n}}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{\frac{1}{n}} \ll \left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{\frac{1}{n}} \ll \frac{1}{Q^{\frac{\delta}{n}}}. \tag{2.32}$$

Sean $Q \in \mathfrak{F}$ y $\mathbf{g}_1(Q)', \dots, \mathbf{g}'_n(Q)$ puntos que realicen a $\lambda'_1, \dots, \lambda'_n$. Nótese que

$$\forall \mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \quad \max_{1\le j\le n}\left\{\frac{|L_j(\mathbf{x})|}{Q^{c_j}}\right\} \ge \frac{1}{Q}\max_{1\le j\le n}\{|L_j(\mathbf{x})|\} \gg \frac{\|\mathbf{x}\|_\infty}{Q} \ge \frac{1}{Q},$$

que con $\mathbf{x} = \mathbf{g}'_1$ conduce a $\lambda_1 \gg Q^{-1}$. Por otra parte, si $\mathbf{e}_1, \dots, \mathbf{e}_n$ la base canónica,

$$\left(\forall i, j \in [1..n] \qquad \frac{|L_i(\mathbf{e}_j)|}{Q^{c_i}} \ll \frac{1}{Q^{c_i}} \le Q\right) \Longrightarrow \left(\forall i \in [1..n] \qquad \frac{|L_i(\mathbf{g}'_n)|}{Q^{c_i}} \ll Q\right),$$

de donde $\lambda_n \ll Q$, que se traduce por (2.32) y $\lambda_n^{\frac{n-1}{n}} \ge \lambda_{n-1}^{\frac{n-1}{n}}$ en

$$\rho_n = \frac{\rho_0}{\lambda_{n-1}} \gg \frac{1}{\lambda_{n-1}}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{\frac{1}{n}} \ge \frac{1}{\lambda_n} \gg \frac{1}{Q}, \quad \rho_1 = \frac{\rho_0}{\lambda_1} \ll \frac{1}{\lambda_{n-1}}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{\frac{1}{n}} \le \frac{1}{\lambda_1} \ll Q. \tag{2.33}$$

Juntando a (2.33) con $\rho_1 \ge \ldots \ge \rho_n > 0$ se deduce $Q^{-1} \ll \rho_n \le \rho_{n-1} \le \ldots \le \rho_1 \ll Q$. Luego, si $A_i := Q^{c_i}/\rho_{t_i}$,

$$\forall i \in [1..n] \quad \frac{1}{Q^2} \ll \frac{Q^{c_i}}{Q} \ll \frac{Q^{c_i}}{\rho_i} = A_i \ll Q^{c_i+1} \le Q^2.$$

Entonces, existe $Q_1 > 0$ que cumple con

$$\forall Q \in \mathfrak{F}_{>Q_1} \quad A_1\cdots A_n = \frac{Q^{c_1+\ldots+c_n}}{\rho_{t_1}\cdots\rho_{t_n}} = 1, \quad \max\left\{A_1, \dots, A_n, \frac{1}{A_1}, \dots, \frac{1}{A_n}\right\} \le Q^3.$$

Además, si $\delta_2 = \delta/(4n)$, (2.32) da $\lambda'_{n-1} < Q^{-3\delta_2}$ cuando $Q \in \mathfrak{F}_{>Q_1}$. Aplicando el Lema 2.16 ($\delta_2 \leftarrow \delta$, $\mathfrak{B} \leftarrow \{Q^3 : Q \in \mathfrak{B}_{>Q_1}\}$) se obtiene $\mathfrak{F}' \subseteq \mathfrak{F}$ no acotado y $\mathbf{h}' \in \mathbb{R}^n$ tales que

$$\forall Q^3 \in \mathfrak{B}'' \quad \mathbf{g}'^*_n(Q^3) = \mathbf{h}'. \tag{2.34}$$

III. Tómese $Q \in \mathfrak{F}'$. Si $T_{n-1} = \operatorname{span}\{\mathbf{g}_1, \dots, \mathbf{g}_{n-1}\}$, de (12) y (2.33) sigue que

$$\forall \mathbf{g} \in \mathbb{Z}^n \setminus T_{n-1} \qquad \max_{1\le i\le n} \frac{|L_i(\mathbf{g})|}{A_i} = \max_{1\le i\le n} \frac{|L_i(\mathbf{g})|\rho_{t_i}}{Q^{c_i}} \ge \frac{\lambda'_n \rho_n}{2^n} \gg \frac{\lambda'_n}{Q} \ge \lambda'_n \gg \lambda_n \rho_n \gg 1.$$

Por otra parte, (2.32) dice $\lambda'_{n-1} \ll Q^{-\frac{\delta}{n}}$ de donde $\lambda_{n-1} < 1$ si $Q \in \mathfrak{F}'$ es grande. En consecuencia, $\{\mathbf{g}'_1, \dots, \mathbf{g}'_{n-1}\} \subseteq T_{n-1}$ que da $T_{n-1} = \operatorname{span}\{\mathbf{g}'_1, \dots, \mathbf{g}'_{n-1}\}$. Llámese $\mathbf{m} \in \mathbb{Z}^n \cap T_{n-1}^{\perp}$ con entradas coprimas. Un argumento similar al de la demostración del Lema 2.15 indica que

$$\forall Q \in \mathfrak{F}' \qquad \mathbf{g}_n^*(Q) = \frac{1}{F}\mathbf{m}, \quad \mathbf{g}_n'^*(Q) = \frac{1}{F'}\mathbf{m} = \mathbf{h}' \quad \text{con } 1 \le |F|, |F'| \le n!.$$

Luego, $\mathbf{g}_n^*(Q) = (F'/F)\mathbf{h}'$ toma una cantidad finita de valores cuando $Q^3 \in \mathfrak{F}'$. Por lo ranto, existen $\mathfrak{F}'' \subseteq \mathfrak{F}$ no acotado y $\mathbf{h}$ para los que $\mathbf{g}_n^* = \mathbf{h}$ para $Q \in \mathfrak{F}''$. □

2.4.3. El polinomio auxiliar

La cuenta pendiente está en el Teorema 2.14. La idea que guía a este argumento se asemeja a la del Teorema de Roth. No obstante, el costo de la generalización está en el uso de nociones matemáticas más sofisticadas y, sobre todo, más rincones detallados. Buscando facilitar la comprensión, en la medida de lo posible se imita la exposición del capítulo pasado.

Resultados principales

Para construir el polinomio auxiliar se recurre a un argumento que debe sonar familiar. Sean $L^{(1)}, \dots, L^{(n)} : \mathbb{R}^n \to \mathbb{R}$ transformaciones lineales independientes[11] fijas dadas por

$$\forall i \in [1..n] \quad L^{(i)}(X_1, \dots, X_n) = \sum_{j=1}^{n} \alpha_{i,j} X_j,$$

con todos los números $\alpha_{i,h}$ enteros algebraicos reales. A cada una se le asocia una colección de formas lineales definidas por

$$\forall i \in [1..n] \quad \forall h \in [1..m] \quad L_h^{(i)}(X_{h,1}, \dots, X_{h,n}) := \sum_{j=1}^{n} \alpha_{i,j} X_{h,j}.$$

Asimismo, cuando $\mathbf{r} \in \mathbb{N}^m$ y $P \in \mathbb{R}[\mathbf{X}]$, se denota al índice de P con respecto a $L_1, \dots, L_m$ y $\mathbf{r}$, $\operatorname{ind}(P; L_1, \dots, L_m, \mathbf{r})$, simplemente $\operatorname{ind}(P; L, \mathbf{r})$. Además, retomando la notación de arriba,

$$\forall i \in \{1, 2, \dots, n\} \quad \Delta_i := [\mathbb{Q}(\alpha_{i,1}, \dots, \alpha_{i,n}) : \mathbb{Q}], \qquad \Delta := \max\{\Delta_1, \dots, \Delta_n\}.$$

Para construir el polinomio auxiliar se recurre a un resultado parecido al Lema 1.4, cuya prueba también sigue un camino semejante. En breve, los coeficientes de P se determinan a partir de las condiciones impuestas. Las restricciones sobre el polinomio conducen a un sistema homogéneo de ecuaciones lineales en el que las incógnitas serán los coeficientes de P. Finalmente, el Lema de Siegel proporcionará una solución no trivial.

[11] Para las primeras dos conclusiones del Teorema 2.18 puede tomarse t formas lineales con $t \in \mathbb{N}$ arbitrario, sin importar la pérdida de independencia lineal.

Teorema 2.18 (Teorema del Polinomio)**.** *Sean $\varepsilon > 0$, $m \in \mathbb{N}$ con $m \geq 4\log(2t\Delta)\varepsilon^{-2}$ y $\mathbf{r} \in \mathbb{N}^m$. Entonces, existe $P \in \mathbb{R}[\mathbf{X}]$ con coeficientes en $\mathbb{Z}$ y homogéneo en $X_{h,1}, \ldots, X_{h,n}$ de grado r_h para $h \in [1..m]$ tal que*

I. *Para cada $i \in [1..n]$ vale $\operatorname{ind}\left(P; L^{(i)}, \mathbf{r}\right) \geq \left(n^{-1} - \varepsilon\right) m$.*

II. *Existe una constante $D = D(\alpha_{i,j}) > 0$ tal que $\mathcal{H}(P) \leq D^{|\mathbf{r}|}$.*

III. *Escribiendo para cada $\mathfrak{T} = (k_{1,1}, \ldots, k_{m,n}) \in \mathbb{N}_0^{mn}$ – de manera única –*

$$P_{\mathfrak{T}} = \sum_{\mathfrak{J}} d_{\mathfrak{T}}(\mathfrak{J}) \left(L_1^{(1)}\right)^{k_{1,1}} \cdots \left(L_1^{(n)}\right)^{k_{1,n}} \cdots \left(L_m^{(1)}\right)^{k_{m,1}} \cdots \left(L_m^{(1)}\right)^{k_{m,n}}, \tag{2.35}$$

existe $E = E(\alpha_{i,h}) > 0$ tal que $|d_{\mathfrak{T}}(\mathfrak{J})| \leq E^{|\mathbf{r}|}$ para cualquier $\mathfrak{T} \in \mathbb{N}_0^{mn}$.

IV. *Si $\mathfrak{J} = (j_{1,1}, \ldots, j_{m,n}) \in \mathbb{N}_0^{mn}$ cumple $\mathfrak{J}/\mathbf{r} \leq 2\varepsilon m$, entonces $d^{\mathfrak{J}}(\mathfrak{J}) = 0$ en (2.35), excepto cuando*

$$\forall k \in [1..n] \quad \left| \left(\sum_{h=1}^{m} \frac{j_{h,k}}{r_h} \right) - \frac{m}{n} \right| \leq 3nm\varepsilon.$$

Tómese $\delta \in (0,1)$ y supóngase, adicionalmente, que ε satisface $16n^2\varepsilon < \delta$. Además, manténgase el significado de $L_h^{(i)}$, $\mathbf{r}$, m, D y E. El paso siguiente es acotar superiormente el índice de P con respecto a ciertas formas lineales y $\mathbf{r}$. Esta meta se alcanza con una proposición similar al Teorema 1.5 y, como en este resultado, se fija una colección de reales $Q_1, \ldots, Q_m$ suficientemente grandes, concretamente

$$\forall h \in [1..m] \qquad Q_n^{\varepsilon} > 2^n E, \quad Q_h^{\varepsilon} > n\left(\frac{1}{\varepsilon} + 1\right) \tag{2.36}$$

La prueba de este lema requiere trabajar un tipo de conjuntos finitos llamados mallas.

Teorema 2.19. *Tómese $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ tal que $\|\mathbf{c}\|_\infty \leq 1$ y $c_1 + \ldots + c_n = 0$. Supóngase que $Q_1, \ldots, Q_m$ satisfacen*

$$\forall h \in [1..m] \qquad r_1 \log Q_1 \leq r_h \log Q_h \leq (1+\varepsilon) r_1 \log Q_1. \tag{2.37}$$

Para cada $h \in [1..m]$ sea $\{\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,n-1}\} \subseteq \mathbb{Z}^n$ un conjunto linealmente independiente tal que

$$\forall k \in [1..n] \quad \forall h \in [1..m] \quad \forall t \in [1..q] \qquad \left|L^{(k)}(\mathbf{g}_{h,t})\right| \leq Q_h^{c_k - \delta} \tag{2.38}$$

Sea $M_h = M_h(\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,q})$ una forma lineal en $X_{h,1}, \ldots, X_{h,n}$ dada por $M_h(\mathbf{x}) = \langle \mathbf{m}_h, \mathbf{x} \rangle$ donde $\mathbf{m}_h \in \{\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,q}\}^{\perp} \cap \mathbb{Z}^n$ tiene entradas coprimas (como en la derivación de (2.26)). Entonces,

$$m\varepsilon \leq \operatorname{ind}(P; M_1, \ldots, M_m; \mathbf{r}). \tag{2.39}$$

Si bien los vectores $\mathbf{m}_h$ del enunciado anterior no son únicos, sólo hay dos posibilidades para elegirlos y éstas son inversos aditivos. Afortunadamente, por la definición de índice de un polinomio, el valor de $\operatorname{ind}(P; M_1, \ldots, M_m, \mathbf{r})$ es el mismo en las dos opciones.

Apoyándose en el Lema Fundamental de Roth (Capítulo I, Teorema 1.6) se concluye su homólogo. Como en aquella ocasión, se considera $\varepsilon \in (0, 12^{-1})$, $m \in \mathbb{N}$ y

$$\omega = 24 \frac{1}{2^m} \left(\frac{\varepsilon}{12}\right)^{2^{m-1}}.$$

Teorema 2.20. *Sean* $\mathbf{r} = (r_1, \ldots, r_m) \in \mathbb{N}^m$ *tal que*

$$\forall h \in [1..m] \quad \omega r_h \geq r_{h+1}, \tag{2.40}$$

$n \in \mathbb{N}_{\geq 2}$, $q = n - 1$ *y* $0 \neq M_h : \mathbb{R}^n \to \mathbb{R}$ *lineal con coeficientes en* $\mathbb{Z}$ *coprimos para cada* $h \in [1..m]$. *Tómese* $0 < \xi \leq q$ *que satisfaga*

$$\forall h \in [1..m] \qquad \mathcal{H}(M_h)^{r_h} \geq \mathcal{H}(M_1)^{r_1 \xi} \tag{2.41}$$

$$\forall h \in [1..m] \qquad \mathcal{H}(M_h)^{\omega \xi} \geq 2^{3mq^2} \tag{2.42}$$

Si $0 \neq P \in \mathbb{R}[\mathbf{X}]$ *tiene coeficientes en* $\mathbb{Z}$, *es homogéneo en* $X_{h,1}, \ldots, X_{h,n}$ *de grado* r_h *para* $h \in [1..m]$ *y*

$$\mathcal{H}(P)^{q^2} \leq \mathcal{H}\, {M_1}^{\omega r_1 \xi}; \tag{2.43}$$

entonces, $\mathrm{ind}(P, M_1, \ldots, M_m; \mathbf{r}) \leq \varepsilon$

Lemas preliminares

Las demostraciones de los Resultados Principales exigen resultados previos. Probarlas representaría un fuerte corte en el flujo de de las ideas. Así que por el momento bastará con enunciarlas. Las pruebas están en el Apéndice E.

En el Teorema 2.20 aparece un símbolo que estuvo presente en el capítulo anterior, pero se recuerda su significado. La **altura** de un polinomio $P \in \mathbb{R}[\mathbf{X}]$, $\mathcal{H}(P)$, es el máximo de los valores absolutos de sus coeficientes. Con fines de referencia se reformula el Lema 1.9 en el contexto actual.

Lema 2.21. *Si* $P \in \mathbb{R}[\mathbf{X}]$ *tiene coeficientes en* $\mathbb{Z}$, *también los tendrá* $P_{\mathbf{i}}$. *Además, cuando* P *sea homogéneo en* $X_{h,1}, X_{h,2}, \ldots, X_{h,n}$ *de grado* r_h *para* $h \in [1..m]$, *entonces* $\mathcal{H}(P_{\mathbf{i}}) \leq 2^{|\mathbf{r}|}\, \mathcal{H}(P)$.

Dados $\alpha_1, \ldots, \alpha_n$ enteros algebraicos, se considera a $K = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ y a $\Delta = [K : \mathbb{Q}] < \infty$. Recuérdese que, en general, si $L|\mathbb{Q}$ es una extensión finita y B es la cerradura entera de $\mathbb{Z}$ en L, una **base entera** de B es un conjunto $\{\omega_1, \ldots, \omega_{[L:\mathbb{Q}]}\} \subseteq B$ tal que para toda $b \in B$ existen $a_1, \ldots, a_{[L:\mathbb{Q}]} \in \mathbb{Z}$ únicos que verifican $b = a_1\omega_1 + \ldots + a_n\omega_{[L:\mathbb{Q}]}$. La Teoría Algebraica de los Números establece la existencia de una base entera de K (cfr. [Ne], Teorema 2.10, p. 12). Fíjese una base entera de K, $\{\beta_1, \ldots, \beta_\Delta\}$. Como el producto de enteros algebraicos es entero algebraico, existen $c_k(i,j) \in \mathbb{Z}$ únicos con $i, j, k \in [1..\Delta]$ tales que

$$\beta_i \beta_j = \sum_{k=1}^{\Delta} c_k(i,j)\beta_k.$$

Con esta información se construye una norma en K denotada $\|\cdot\|_K$. Ya que $\{\beta_1, \ldots, \beta_\Delta\}$ son $\mathbb{Q}$ linealmente independientes, forman una base de K. En consecuencia, para cada $\gamma \in K$ existe un único vector $\mathbf{r} \in \mathbb{Q}^\Delta$ tal que

$$\gamma = \sum_{j=1}^{\Delta} r_j \beta_j.$$

Con $A := \max_{i,j,k}\{|c_k(i,j)|\}$, la fórmula $\|\gamma\|_K := A\Delta^2 \|\mathbf{r}\|_\infty$ da una norma submultiplicativa. En símbolos,

$$\forall \gamma, \delta \in K \qquad (\|\gamma + \delta\|_K \leq \|\gamma\|_K + \|\delta\|_K) \quad \& \quad (\|\gamma\delta\|_K \leq \|\gamma\|_K \|\delta\|_K)\,.$$

La desigualdad del triángulo no necesita mayor explicación. Para ver la segunda, tómese a $\gamma \in K$ como arriba y a $\delta \in K$ con vector de coeficientes $\mathbf{s}$. Se tiene, pues,

$$\gamma\delta = \sum_{i,j} r_i s_j \beta_i \beta_j = \sum_{i,j,k} r_i s_j c_k(i,j)\beta_k.$$

Ahora, los coeficientes de cada β_k son números raciones y satisfacen

$$\forall k \in [1..\Delta] \qquad \left|\sum_{i,j} r_i s_j c_k(i,j)\right| \le A\Delta^2 \|\mathbf{r}\|_\infty \|\mathbf{s}\|_\infty = \frac{1}{A\Delta^2}\|\gamma\|\|\delta\|,$$

que implica $\|\gamma\delta\|_K \le \|\gamma\|_K \|\delta\|_K$.

Lema 2.22. *Sean $\alpha_1, \dots, \alpha_n$ enteros algebraicos y $\mathfrak{J} \in \mathbb{Z}^{mn}$ con entradas no negativas y de la forma $\mathfrak{J} = (j_1, 0, \dots, 0; j_2, 0, \dots, 0; \dots; j_m, 0, \dots, 0)$. Para cada $P \in \mathbb{R}[\mathbf{X}]$ homogéneo en $X_{h,1}, \dots, X_{h,n}$ con $h \in [1..m]$,*

$$P(X_{1,1}, \dots, X_{1,n}; \dots; X_{m,1}, \dots, X_{m,n}) = \sum c(j_{1,1}, \dots, j_{m,n}) X_{1,1}^{j_{1,1}}, \dots, X_{1,n}^{j_{m,n}} \in \mathbb{R}[\mathbf{X}],$$

se define el polinomio $P^* = P^*(X_{1,2}, \dots, X_{1,n}; \dots; X_{m,2}, \dots, X_{m,n})$ en $nm - m$ variables mediante

$$P^* = P_{\mathfrak{J}}(-\alpha_2 X_{1,2} - \dots - \alpha_n X_{1,n}, \alpha_1 X_{1,2}, \dots, X_{1,n}; \dots; -\alpha_2 X_{m,2} - \dots - \alpha_n X_{m,n}, \alpha_1 X_{m,2}, \dots, X_{m,n}).$$

Entonces, los coeficientes de P^* son una combinación lineal de los coeficientes de P y los coeficientes de estas combinaciones, γ, son enteros en $K := \mathbb{Q}(\alpha_1, \dots, \alpha_n)$ que cumplen con

$$\|\gamma\|_K \le (4^n B)^{|\mathbf{r}|}, \quad B := \max\{\|\alpha_1\|_K, \dots, \|\alpha_n\|_K\}.$$

2.4.4. Construcción de polinomio

Demostración del Teorema 2.18. ***I.*** *y* ***II.*** Supopiendo que $P \in \mathbb{R}[\mathbf{X}]$ resuelve el problema, el objetivo es determinar a sus coeficientes.

I. Con la abreviatura $\mathbf{j}_h = (j_{h,1}, \dots, j_{h,m})$,

$$P(X_{1,1}, \cdots, X_{m,n}) = \sum_{h=1}^{m} \sum_{|\mathbf{j}_h| = r_h} c(j_{1,1}, \dots, j_{m,n}) X_{1,1}^{j_{1,1}} \cdots X_{m,n}^{j_{m,n}}.$$

Puede suponerse $a_{1,1} \neq 0$, pues ninguna forma es idénticamente cero. A cada vector

$$\mathfrak{T} = (i_1, 0, \dots, 0; \dots; i_m, 0, \dots, 0) \in \mathbb{Z}^{mn}$$

con entradas no negativas se le asocia el multi-índice $\mathbf{i}_{\mathfrak{T}} = (i_1, \dots, i_m)$. Tras llamar I al conjunto de vectores en $\mathbb{Z}^{mn}$ con esta forma, se obtiene trivialmente del Lema 2.12 que si vale[12]

$$\forall \mathfrak{T} \in I \qquad \left(\mathbf{i}_{\mathfrak{T}} \cdot \mathbf{r}^{-1} < \left(\frac{1}{n} - \varepsilon\right) m \quad \& \quad \left(\forall h \in [1..m] \quad 0 \le i_h \le r_h\right)\right) \implies P_{\mathfrak{T}}|_{T_1} = 0. \tag{2.44}$$

con $T_1 := \bigcap_{j=1}^{m} \ker L_j^{(1)}$, entonces $\operatorname{ind}(P; L^{(1)}, \mathbf{r}) \ge \left(n^{-1} - \varepsilon\right) m$.

Fíjese $\mathfrak{T} \in I$. Cuando $\mathbf{X} = (X_{1,1}, \dots, X_{1,n}; \dots; X_{m,1}, \dots, X_{m,n})$ pertenece a T, puede escribirse a $X_{h,1}$ en términos de $X_{h,2}, \dots, X_{h,n}$ para cada $h \in [1..m]$. Así, la anulación de $P_{\mathfrak{T}}$ en T_1 se traduce en $P^* = 0$, donde P^* es el polinomio definido en el Lema 2.22 con $\mathfrak{J} = \mathfrak{T}$.

[12] Recuérdese que $\mathbf{i} \cdot \mathbf{r}^{-1} := \sum_j i_j / r_j$.

II. Por construcción, P^* es homogéneo en $X_{h,2},\dots,X_{h,n}$ de grado $r_h - i_h$ para $h \in [1..m]$. Entonces, la cantidad de coeficientes que tiene es a lo más[13]

$$\binom{r_1 - i_1 + n - 2}{n-2} \cdots \binom{r_m - i_m + n - 2}{n-2}.$$

Por la levedad en la escritura se adopta

$$\forall h \in [1..m] \quad \nu_h(i_h) := \binom{r_h - i_h + n - 2}{n-2}.$$

Por el Lema 2.22, cada coeficiente de P^* da pie a una ecuación lineal en los coeficientes de P igualada a cero. De este modo, los coeficientes de P satisfacen un sistema lineal homogéneo de $\nu(i_1)\cdots\nu(i_m)$ ecuaciones con coeficientes enteros en $K_1 = \mathbb{Q}(\alpha_{1,1},\dots,\alpha_{1,n})$, γ, que satisfacen

$$\|\gamma\|_K \le (4^n B_1)^{|\mathbf{r}|}, \quad B_1 = \max\{\|\alpha_{1,1}\|_{K_1},\dots,\|\alpha_{1,n}\|_{K_1}\}.$$

Escribiendo a cada γ en términos de la base entera de K_1 fijada, cada una de estas ecuaciones da Δ_1 ecuaciones homogéneas con coeficientes en $\mathbb{Z}$, k, tales que

$$|k| \le (4^n B_1)^{|\mathbf{r}|}. \tag{2.45}$$

Por lo tanto, para cumplir la primera conclusión hay que satisfacer a lo más

$$\Delta_1 \sum_{\mathfrak{T}} \nu_1(i_1)\cdots\nu_m(i_m) \tag{2.46}$$

ecuaciones, donde la suma corre sobre todos los $\mathfrak{T}$ que verifiquen el antecedente en la expresión (2.44). La suma tiene una cantidad finita de términos porque es sobre un subconjunto acotado de puntos enteros.

III. Aprovechando las hipótesis y las desigualdades combinatorias se llega a

$$\begin{aligned}
\Delta_1 \sum_{\mathfrak{T}} \nu_1(i_1)\cdots\nu_m(i_m) &\le \Delta_1 \binom{r_1 + n - 1}{r_1} \cdots \binom{r_m + n - 1}{r_m} \frac{1}{\exp\left(\frac{\varepsilon^2 m}{4}\right)} \\
&\le \Delta_1 \binom{r_1 + n - 1}{r_1} \cdots \binom{r_m + n - 1}{r_m} \frac{1}{2n\Delta} \\
&\le \frac{N}{2n} \qquad \text{con} \qquad N = \binom{r_1 + n - 1}{r_1} \cdots \binom{r_m + n - 1}{r_m}.
\end{aligned}$$

La primera línea vale por el Lema (4) del Apéndice E y la segunda, por $m \ge 4\log(2n\Delta)\varepsilon^{-2}$. Nótese que todo este argumento no es exclusivo de $L^{(1)}$. En consecuencia, al sumar sobre las n formas lineales se tiene que el total de ecuaciones a resolver para alcanzar **II.**, M, satisface

$$M \le \frac{N}{2}.$$

IV. La ecuación (2.45) aplicada a cada $i \in [1..n]$, designando $B = \max\{B_1,\dots,B_n\}$, implica que el valor absoluto de los coeficientes de estas M ecuaciones es a lo más

[13] Para $r, n \in \mathbb{N}$ la ecuación $y_1 + \ldots + y_n = r$ tiene $\binom{n+r-1}{r-1}$ soluciones enteras no negativas.

$A := (4^n B)^{|\mathbf{r}|}$. Una aplicación del Lema de Siegel (Lema 2.2) con las asignaciones $m \leftarrow M$, $n \leftarrow N$, $K \leftarrow A$, da una solución $\mathbf{c} \neq \mathbf{0}$ con entradas en $\mathbb{Z}$ que cumple con

$$\|\mathbf{c}\|_\infty \leq (NA)^{\frac{M}{N-M}} \leq NA = \binom{r_1+n-1}{r_1} \cdots \binom{r_m+n-1}{r_m} A \leq 2^{n|\mathbf{r}|} A = (8^n B)^{|\mathbf{r}|} =: D^{|\mathbf{r}|}.$$

El polinomio P con coeficientes $\mathbf{c}$ satisface las primeras dos conclusiones, pues $\mathcal{H}(P) = \|\mathbf{c}\|_\infty$.

III. Por la independencia lineal de $L^{(1)}, \ldots, L^{(n)}$, la escritura de $X_{h,1}, \ldots, X_{h,n}$ como combinación lineal de $L_h^{(1)}, \ldots, L_h^{(n)}$ es única y viceversa ($h \in [1..m]$). En consecuencia, cualquier polinomio en $X_{h,1}, \ldots, X_{h,n}$ se escribe de manera única como un polinomio en $L_h^{(1)}, \ldots, L_h^{(n)}$. Recordando que $L^{(h)}(X_1, \ldots, X_n) = \alpha_{h,1}X_1 + \ldots + \alpha_{h,n}X_n$ para $h \in [1..m]$, se llama $A := (\alpha_{ij})$ y $B := (\beta_{ij})$ a la inversa de A. Entonces,

$$\forall i \in [1..n] \quad \left(X_i = \sum_{k=1}^{n} \beta_{ik} L^{(k)} \right) \quad \& \quad \left(\forall h \in [1..m] \quad X_{hi} = \sum_{k=1}^{n} \beta_{ik} L_h^{(k)} \right). \tag{2.47}$$

Tómese $\mathfrak{T}$ un multi-índice de mn entradas. Escribiendo $\mathbf{X} = (X_{1,1}, \ldots, X_{m,n})$ y $\mathbf{X}^{\mathfrak{J}} = X_{1,1}^{j_{1,1}} \cdots X_{m,n}^{j_{m,n}}$,

$$P_{\mathfrak{T}} = \sum c_{\mathfrak{T}}(\mathfrak{J}) \mathbf{X}^{\mathfrak{J}}. \tag{2.48}$$

La suma corre sobre el conjunto de los multi-índices; salvo una cantidad finita, todos tienen un coeficiente igual a cero. Por la segunda parte de (2.47), cada sumando en (2.48) es de la forma

$$X_{1,1}^{j_{1,1}} \cdots X_{m,n}^{j_{m,n}} = \left(\sum_{k=1}^{n} \beta_{1,k} L_1^{(k)} \right)^{j_{1,1}} \cdots \left(\sum_{k=1}^{n} \beta_{n,k} L_m^{(k)} \right)^{j_{m,n}}. \tag{2.49}$$

A continuación, se acotará a cada $d_{\mathfrak{J}}$, $\mathfrak{J} \in \mathbb{N}_0^{mn}$. Primero, expandiendo el producto (2.49) se llega a un polinomio en $L_h^{(i)}$ cuyos coeficientes tienen un valor absoluto menor o igual que[14]

$$(nG)^{j_{1,1}} \cdots (nG)^{j_{m,n}} = (nG)^{|\mathfrak{J}|} \leq (nG)^{|\mathbf{r}|}.$$

donde $G := \max\{1, |\beta_{1,1}|, \ldots, |\beta_{m,n}|\}$. Segundo, por el Lema 2.21 y la segunda conclusión del presente teorema,

$$\mathcal{H}(P_{\mathfrak{T}}) \leq 2^{|\mathbf{r}|} \mathcal{H}(P) \leq (2D)^{|\mathbf{r}|}.$$

Luego, $c_{\mathfrak{T}}(\mathfrak{J})\mathbf{X}^{\mathfrak{J}}$ como polinomio en las formas $L_h^{(i)}$ tiene coeficientes con valor absoluto a lo más $(2DnG)^{|\mathbf{r}|}$. Tercero, ya que P es homogéneo de grado r_h en $X_{h,1}, \ldots, X_{h.n}$ para $h \in [1..m]$, la expresión en (2.48) cuenta a lo más con

$$\binom{r_1+n-1}{r_1} \cdots \binom{r_m+n-1}{r_m} \leq 2^{r_1+n-1} \cdots 2^{r_m+n-1} \leq 2^{nr_1} \cdots 2^{nr_m} = 2^{n|\mathbf{r}|}.$$

sumandos diferentes de cero. En consecuencia, los coeficientes de $P_{\mathfrak{J}}$ como polinomio en $L_h^{(i)}$ tienen valor absoluto a lo más

$$2^{n|\mathbf{r}|}(2DnG)^{|\mathbf{r}|} = (2^n 2DnG)^{|\mathbf{r}|} =: E^{|\mathbf{r}|}.$$

[14] El Teorema Multinomial dice que $(x_1 + \ldots + x_n)^J = \sum_{\mathbf{j}} \binom{J}{j_1, \ldots, j_n} x_1^{j_1} \cdots x_n^{j_n}$ donde $\mathbf{0} \leq \mathbf{j}$ corre en el conjunto de los multi-índices tales que $j_1 + \ldots + j_n = J$. En particular, si $x_1 = \ldots = x_n = 1$, entonces $n^J = \sum_{\mathbf{j}} \binom{J}{j_1, \ldots, j_n}$.

IV. Sea $\mathfrak{J} = (j_{1,1}, \ldots, j_{m,n}) \in \mathbb{Z}^{mn}$ con entradas no negativas. Para cada $k \in [1..n]$, la construcción de P garantiza $\operatorname{ind}(P; L^{(k)}, \mathbf{r}) \geq m(1/n - \varepsilon)$; entonces, si $\mathfrak{J}/\mathbf{r} \leq 2\varepsilon m$, el Lema 2.11 conlleva

$$\operatorname{ind}\left(P_{\mathfrak{J}}; L^{(k)}, \mathbf{r}\right) \geq \operatorname{ind}\left(P; L^{(k)}, \mathbf{r}\right) - \frac{\mathfrak{J}}{\mathbf{r}} \geq \left(\frac{1}{n} - \varepsilon\right) m - 2\varepsilon m = \left(\frac{1}{n} - 3\varepsilon\right) m.$$

Por la definición de índice, $d_{\mathfrak{T}}(\mathfrak{J}) \neq 0$ implica $\mathfrak{J}/\mathbf{r} \geq \operatorname{ind}(P; L^{(h)}, \mathbf{r})$ que, junto con la desigualdad anterior, da

$$\forall k \in [1..n] \quad \sum_{h=1}^{m} \frac{j_{h,k}}{r_h} - \frac{m}{n} \geq -3\varepsilon m. \tag{2.50}$$

Se concluye una de las desigualdades buscadas.

Para la otra, obsérvese que el grado total de $P_{\mathfrak{J}}$ en $X_{h,1}, \ldots, X_{h,n}$ es menor o igual que r_h para $h \in [1..n]$. Esto produce

$$\left(\forall h \in [1..m] \quad \sum_{k=1}^{n} \frac{j_{h,k}}{r_h} \leq 1\right) \Longrightarrow \sum_{k=1}^{n} \sum_{h=1}^{m} \frac{j_{h,k}}{r_h} \leq m \Longrightarrow \sum_{k=1}^{n} \left[\sum_{h=1}^{m} \left(\frac{j_{h,k}}{r_h}\right) - \frac{m}{n}\right] \leq 0.$$

Finalmente, la desigualdad anterior y (2.50) dan, argumentando por contradicción,

$$\forall k \in [1..n] \quad \sum_{h=1}^{m} \frac{j_{h,k}}{r_h} - \frac{m}{n} \leq 3m\varepsilon(n-1) < 3\varepsilon mn. \qquad \square$$

2.4.5. Cota inferior del índice

Mallas

Las mallas son una idea sencilla de poderosas consecuencias. El resultado principal de esta breve sección asegura que basta con estudiar a un polinomio en múltiples variables sobre una malla- un conjunto finito- para concluir globalmente su comportamiento. El caso univariado no será sino una consecuencia de una propiedad elemental los polinomios con coeficientes reales. Por ahora, si $n \in \mathbb{N}_{\geq 2}$, $q = n - 1$.

Definición 2.6. *Sean $H^q \subseteq \mathbb{R}^n$ un subespacio lineal de $\mathbb{R}^n$, $\mathbf{w}_1, \ldots, \mathbf{w}_q \in H^q$ una base fija y $s \in \mathbb{N}$. Una **malla** en H^q de **tamaño** s es*

$$\Gamma := \Gamma(s; \mathbf{w}_1, \ldots, \mathbf{w}_q) := \{\mathbf{w} = \mathbf{w}_1 h_1 + \ldots + \mathbf{w}_q h_q : \forall j \in [1..q] \quad h_j \in [1..s]\}.$$

Lema 2.23. *Sean $P \in \mathbb{R}[X_1, \ldots, X_n]$ de grado total r y $s, t \in \mathbb{N}$ tales que $s(t+1) > r$. Si para cualesquiera $\mathbf{j} \in \mathbb{N}_0^n$ con $|\mathbf{j}| \leq t$ y $\mathbf{x} \in [1..s]^n$ se cumple $P_{\mathbf{j}}(\mathbf{x}) = 0$, entonces $P \equiv 0$.*

Demostración. La prueba es por inducción sobre n. La base, $n = 2$, es evidente, porque las condiciones sobre las derivadas dan $s(t+1)$ raíces, contando multiplicidades, de un polinomio de grado $r < s(t+1)$. Por lo que P es idénticamente cero.

La Hipótesis de Inducción es el Lema para $n = N - 1 \geq 1$ y se probará para $n = N$. Supóngase que $P \in \mathbb{R}[X_1, \ldots, X_n]$ y $s, t, r \in \mathbb{N}$ satisfacen las hipótesis del enunciado, pero que la conclusión es falsa: $P \neq 0$. Para cada $h \in [1..n]$ se representa al máximo natural tal que $(X_1 - h) \mid P$ mediante α_h; además, $\alpha := \min\{\alpha_1, \ldots, \alpha_n\}$ y para facilitar la escritura, se supone $\alpha = \alpha_1$. Puede escribirse, entonces $P(X_1, \ldots, X_N) = (X_1 - 1)^{\alpha}(X_1 - 2)^{\alpha_2} \cdots (X_1 - S)^{\alpha_s} \tilde{P}(X_1, \ldots, X_N)$ donde $\tilde{P} \neq 0$. Sea R el polinomio

$$R(X_2, \ldots, X_N) := \frac{P(1, X_2, \ldots, X_N)}{(X_1 - 2)^{\alpha} \cdots (X_1 - s)^{\alpha}} = (X_1 - 2)^{\alpha_2 - \alpha} \cdots (X_1 - S)^{\alpha_s - \alpha} \tilde{P}(1, X_2, \ldots, X_N)$$

Nótese que $\tilde{P}(1, X_2, \ldots, X_N)$ es un polinomio en $N-1$ variables, cuyo grado total, $\tilde{r}$, satisface

$$r - (\alpha_1 + \ldots + \alpha_s) \leq r - \alpha s < s(t - \alpha + 1).$$

Para cada multi-índice $\mathbf{j} \in \mathbb{N}_0^{N-1}$ con $|\mathbf{j}| \leq t - s\alpha$ el polinomio $\tilde{P}$ se anula en $[1..s]^{N-1}$ ($P_{\mathbf{i}}(\mathbf{x}) = 0$ si $\mathbf{i} \in \mathbb{N}_0^N$, $|\mathbf{i}| \leq t$ y $\mathbf{x} \in [1..s]^N$). Entonces, por la Hipótesis de Inducción, $\tilde{P} \equiv 0$, una contradicción. Por lo tanto, $P \equiv 0$. □

Manteniendo la notación, sea Γ es una malla en H^q. Aplicando un isomorfismo lineal, no se pierde generalidad al pensar- con $\mathbf{e}_1, \ldots, \mathbf{e}_n$ la base canónica de $\mathbb{R}^n$- que

$$H^q = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n = 0\}, \quad \Gamma = \Gamma(s; \mathbf{e}_1, \ldots, \mathbf{e}_{n-1})$$

Tómese $P \in \mathbb{R}[X_1, \ldots, X_n]$ de grado total r. Piénsese que para cada $\mathbf{j} \in \mathbb{N}_0^n$ con $|\mathbf{j}| \leq t$ vale $P_{\mathbf{j}}|_\Gamma \equiv 0$, donde t es un natural que satisface $s(t+1) > r$. Sea $Q \in \mathbb{R}[X_1, X_2, \ldots, X_{q-1}]$ dada por

$$Q(X_1, X_2, \ldots, X_{q-1}) = P(X_1, X_2, \ldots, X_{q-1}, 0).$$

Las condiciones del Lema 2.23 están presentes, por lo que $Q \equiv 0$, implicando $P|_{H^q} \equiv 0$. Esto prueba el caso base del siguiente Lema, con inducción sobre m se concluye su veracidad.

Lema 2.24. *Sean $P \in \mathbb{R}[\mathbf{X}]$, $\mathbf{r}, \mathbf{s} \in \mathbb{N}^m$. Supóngase que para cada $h \in [1..m]$ el polinomio homogéneo P tiene grado total en $X_{h,1}, \ldots, X_{h,n}$ menor o igual que r_h, que $H_h^q \subseteq \mathbb{R}^n$ es un subespacio de dimensión q, que Γ_h es una malla en H_h^q de tamaño s_h y que $t_h \in \mathbb{N}$ satisface $s_h(t_h + 1) > r_h$. Además, sean $T := H_1^q \times \ldots \times H_m^q$ y $\Gamma^* := \Gamma_1 \times \Gamma_2 \times \ldots \times \Gamma_m$.*

Si para cada $\mathfrak{T} = (t_{1,1}, \ldots, t_{m,n}) \in \mathbb{N}_0^{mn}$ que satisfaga

$$\forall h \in [1..m] \quad t_{h,1} + t_{h,2} + \ldots + t_{h,n} \leq t_h$$

se tiene $P_{\mathfrak{T}}|_{\Gamma^*} \equiv 0$; entonces, $P|_T \equiv 0$.

Cota inferior

Demostración del Teorema 2.19 Las hipótesis del Teorema 2.19 fuerzan $0 < \varepsilon < 1$ que da $n^{-1} + 3n\varepsilon < 4n$, de donde

$$(1+\varepsilon)\left(\frac{1}{n} + 3n\varepsilon\right)m = \left(\frac{1}{n} + 3n\varepsilon\right)m + \varepsilon\left(\frac{1}{n} + 3n\varepsilon\right)m < \left(\frac{1}{n} + 3n\varepsilon\right)m + 4\varepsilon nm = \left(\frac{1}{n} + 7n\varepsilon\right)m.$$

Demostración del Teorema 2.19. Por los Lemas 2.12 y 2.24, basta con probar que si $\mathfrak{J}/\mathbf{r} < 2\varepsilon m$, entonces $P_{\mathfrak{J}}|_{\Gamma^*} \equiv 0$ con Γ^* un producto de mallas adecuadas. Con esto en mente, ya que para cada $\mathbf{v} \in \Gamma^*$ se cumple $P_{\mathfrak{J}}(\mathbf{v}) \in \mathbb{Z}$, se prueba $|P_{\mathfrak{J}}(\mathbf{v})| < 1$ para concluir $P_{\mathfrak{J}}(\mathbf{v}) = 0$.

1. Se define al subespacio[15] $T := \ker M_1 \times \cdots \times \ker M_m \subseteq \mathbb{R}^{mn}$. De acuerdo con el Lema 2.12, la desigualdad 2.39 es válida si para cada $\mathfrak{J} \in \mathbb{N}_0^{mn}$ con $\frac{\mathfrak{J}}{\mathbf{r}} < \varepsilon m$ se cumple $P_{\mathfrak{J}}\big|_T = 0$. La forma de T y la conclusión sugieren fuertemente el empleo del Lema 2.24. Para ello, considérese a los siguientes objetos:

$$\forall h \in [1..m] \qquad s_h := \left\lfloor \frac{1}{\varepsilon} \right\rfloor + 1, \quad t_h := \lfloor r_h \varepsilon \rfloor, \quad \Gamma_h := \Gamma\left(\left\lfloor \frac{1}{\varepsilon} \right\rfloor + 1; \mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,q}\right)$$

[15] También se puede tomar a $T = \bigcap_h \ker M_h$ siempre y cuando $M_1, \ldots, M_n$ hayan sido extendidas a $\mathbb{R}^{mn}$.

y $\Gamma^* = \Gamma_1 \times \cdots \times \Gamma_n$. Usando el Lema 2.24, la prueba habrá concluido si se muestra que

$$\forall \mathbf{v} \in \Gamma \quad \left(\frac{\mathfrak{J}}{\mathbf{r}} < 2\varepsilon m \implies P_{\mathfrak{J}}(\mathbf{v}) = 0 \right). \tag{2.51}$$

En efecto, supóngase verdadera la expresión anterior. Al fijar a $\mathfrak{J} \in \mathbb{N}_0^{mn}$, $\mathfrak{J}/\mathbf{r} < \varepsilon m$, todas las hipótesis del Lema 2.24 son ciertas para $P_{\mathfrak{J}}$. Primero, de la definición de los parámetros,

$$\forall h \in [1..m] \qquad s_h(t_h + 1) = \left(\left\lfloor \frac{1}{\varepsilon} \right\rfloor + 1 \right) \lfloor r_h \varepsilon \rfloor > \frac{r_h \varepsilon}{\varepsilon} = r_h.$$

Sea $\mathfrak{T} = (t_{1,1}, \ldots, t_{m,n})$ un multi-índice que satisfaga

$$(\forall h \in [1..m] \quad t_{h,1} + \ldots + t_{h,n} \le t_h) \implies \big(\forall \mathbf{v} \in \Gamma \quad (P_{\mathfrak{J}})_{\mathfrak{T}}(\mathbf{v}) = 0 \big).$$

Las desigualdades y la definición de $t_1, \ldots, t_m$ se convierten en

$$\frac{\mathfrak{T}}{\mathbf{r}} \le \frac{\lfloor r_1 \varepsilon \rfloor}{r_1} + \ldots + \frac{\lfloor r_m \varepsilon \rfloor}{r_m} \le \varepsilon m.$$

Y, por $\mathfrak{J}/\mathbf{r} < \varepsilon m$, se tiene $(\mathfrak{T}+\mathfrak{J})/\mathbf{r} < 2\varepsilon m$. Entonces, (2.51) implica que $(P_{\mathfrak{J}})_{\mathfrak{T}}(\mathbf{v}) = 0$, pues $P_{\mathfrak{J}+\mathfrak{T}}$ y $(P_{\mathfrak{J}})_{\mathfrak{T}}$ son proporcionales. Finalmente, por el Lema 2.24, $P_{\mathfrak{J}}|_T = 0$. Así, las fuerzas estarán centradas en verificar 2.51.

II. Supóngase que $\mathfrak{J}$ satisface $\mathfrak{J}/\mathbf{r} < 2\varepsilon m$ y $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_m) \in \Gamma^*$. Claramente, $P_{\mathfrak{J}}(\mathbf{v}) \in \mathbb{Z}$, los coeficientes de P están en $\mathbb{Z}$. Luego, $P_{\mathfrak{J}}(\mathbf{v}) = 0$ sigue de $|P_{\mathfrak{J}}(\mathbf{v})| < 1$, la meta ahora es esta segunda desigualdad. La tercera conclusión del Teorema 2.18 permitirá acotar a $|P_{\mathfrak{J}}(\mathbf{v})|$. Tómese

$$P_{\mathfrak{J}}(\mathbf{v}) = \sum_{\mathfrak{T}} d_{\mathfrak{J}}(\mathfrak{T}) \big(L^{(1)}(\mathbf{v}_1)\big)^{t_{1,1}} \cdots \big(L^{(n)}(\mathbf{v}_1)\big)^{t_{1,n}} \cdots \big(L^{(1)}(\mathbf{v}_m)\big)^{t_{m,1}} \cdots \big(L^{(n)}(\mathbf{v}_m)\big)^{t_{m,n}}, \tag{2.52}$$

donde $\mathfrak{T} = (t_{1,1}, \ldots, t_{1,n}; \ldots; t_{m,1}, \ldots, t_{m,n})$.

Sean $h \in [1..m]$, $k \in [1..n]$. Como $\mathbf{v}_h$ es una combinación lineal de $\mathbf{g}_{h,1}, \ldots, \mathbf{g}_{h,q}$ con coeficientes en $[1..s_h]$,

$$\begin{aligned}
\big|L^{(k)}(\mathbf{v}_h)\big| &\le \sum_{j=1}^{m} \left(1 + \frac{1}{\varepsilon}\right) \big|L^{(k)}(\mathbf{g}_{h,j})\big| && \text{(linealidad de } L^{(k)} \text{ y la definición de } s_h) \\
&< n \left(\frac{1}{\varepsilon} + 1\right) Q_h^{c_k - \delta} && \text{(desigualdad (2.38))} \\
&< Q_h^{c_k - \delta + \varepsilon} && \text{(desigualdad (2.36))} \\
&\le Q_h^{c_k - 15n^2\varepsilon}. && \text{(hipótesis sobre } \delta \text{ y } \varepsilon)
\end{aligned}$$

Esto conlleva $\big|L^{(k)}(\mathbf{v}_h)^{t_{h,k}}\big| < Q_h^{t_{h,k}(c_k - 15n^2\varepsilon)}$. Por la arbitrariedad de k y h,

$$\forall k \in [1..n] \qquad \big|L^{(k)}(\mathbf{v}_1)^{t_{1,k}} \cdots L^{(k)}(\mathbf{v}_m)^{t_{m,k}}\big| \le \exp\left((c_k - 15n^2\varepsilon) \sum_{h=1}^{m} t_{h,k} \log Q_h \right). \tag{2.53}$$

Ahora, si $d_{\mathfrak{J}}(t_{1,1}, \ldots, t_{m,n}) \neq 0$, del cuarto punto del Teorema 2.18 se deduce

$$-3nm\varepsilon \le \sum_{h=1}^{m} \frac{t_{h,k}}{r_h} - \frac{m}{n} \le 3nm\varepsilon.$$

Entonces, por la primera desigualdad en (2.37),

$$\sum_{h=1}^{m} t_{h,k} \log Q_h \geq r_1 \log Q_1 \sum_{h=1}^{m} \frac{t_{h,k}}{r_h} \geq r_1 \log Q_1 \left(\frac{1}{n} - 3n\varepsilon\right) m$$

y, por la segunda desigualdad de (2.37) y la observación al principio de la prueba,

$$\begin{aligned}\sum_{h=1}^{m} t_{h,k} \log Q_h &\leq (1+\varepsilon) r_1 \log Q_1 \sum_{h=1}^{m} \frac{t_{h,k}}{r_h} \leq r_1 \log Q_1 (1+\varepsilon) \left(\frac{1}{n} + 3n\varepsilon\right) m \\ &\leq r_1 \log Q_1 \left(\frac{1}{n} + 7n\varepsilon\right) m.\end{aligned}$$

Juntando las expresiones anteriores,

$$\left| \sum_{h=1}^{m} t_{h,k} \log Q_h - (r_1 \log Q_1) \frac{m}{n} \right| \leq (r_1 \log Q_1) 7nm\varepsilon.$$

De las restricciones sobre $\mathbf{c}$, δ y ε es fácil obtener $|c_k - 15n^2\varepsilon| \leq 2$ para cada $k \in [1..n]$; en consecuencia,

$$\forall k \in [1..n] \qquad \left|L^{(k)}(\mathbf{v}_1)^{j_{1,k}} \cdots L^{(k)}(\mathbf{v}_m)^{j_{m,k}}\right| \leq Q_1^{r_1 \frac{m}{n}(c_k - 15n^2\varepsilon) + 2r_1 7nm\varepsilon} = Q^{r_1\left(\frac{m}{n}c_k - nm\varepsilon\right)}.$$

La tercera conclusión del Teorema 2.18 establece $|d_{\mathfrak{J}}(\mathfrak{T})| < E^{|\mathbf{r}|}$. Por $c_1 + \ldots + c_n = 0$ y (2.53),

$$\left|d_{\mathfrak{J}}(\mathfrak{T}) \left(L^{(1)}(\mathbf{v}_1)\right)^{t_{1,1}} \cdots \left(L^{(n)}(\mathbf{v}_m)\right)^{t_{m,n}}\right| \leq \frac{E^{|\mathbf{r}|}}{Q_1^{r_1 n^2 m\varepsilon}} \leq \frac{E^{|\mathbf{r}|}}{(Q_1^{r_1\varepsilon} \cdots Q_m^{r_m\varepsilon})^{\frac{n^2}{1+\varepsilon}}} \leq \frac{E^{|\mathbf{r}|}}{Q_1^{r_1\varepsilon} \cdots Q_m^{r_m\varepsilon}}.$$

Finalmente, ya que (2.52) tiene menos de $2^{|\mathbf{r}|}$ sumandos, aplicando (2.36) se concluye

$$|P_{\mathfrak{J}}(\mathbf{v}_1, \ldots, \mathbf{v}_n)| \leq \prod_{h=1}^{m} \left(\frac{2^n E}{Q_h^{\varepsilon}}\right)^{r_h} < 1. \qquad \square$$

2.4.6. Cota superior del índice

Lema 2.25. *Sean $n \in \mathbb{N}_{\geq 2}$ y $m_1, \ldots, m_n$ enteros racionales coprimos con $m_1 \neq 0$. Entonces, existe un entero j con $2 \leq j \leq n$ tal que*

$$\text{m. c. d.}(m_1, m_j) \leq |m_1|^{\frac{n-2}{n-1}}.$$

Demostración. Para cada $j \in [2..n]$ se define $d_j = mcd(m_1, m_j)$, entonces $(d_2, d_3, \ldots, d_n) = 1$, de donde $[d_1, \ldots, d_n] = d_1 \cdots d_n$. Por definición, $d_j \mid m_1$ para toda j, por lo que

$$d_2 d_3 \cdots d_n \mid m_1 \implies d_2 d_3 \cdots d_n \mid m_1^{n-2} \implies d_2 d_3 \cdots d_n \leq |m_1|^{n-2}.$$

Luego, al menos para alguna j se cumple la desigualdad busada. $\square$

Demostración del Lema 2.20. La idea general de la prueba es obtener sucesivamente a partir de P nuevos polinomios en menos variables cuyo índice con respecto a ciertas funciones lineales excederá a $\text{ind}(P; M_1, \ldots, M_m; \mathbf{r})$. En el último polinomio se aplica el Lema Fundamental de Roth y su índice quedará acotado superiormente por $\varepsilon > 0$. De aquí se concluirá el resultado.

I. Si $M_h(X_{h,1},\dots,X_{h,n}) = \sum_{j=1}^{n} u_{h,j}X_{h,j}$ para cada $h \in [1..m]$, puede pensarse que $\mathcal{H}(M_h) = |u_{h,1}|$ y, por el Lema 2.25, que

$$\forall h \in [1..m] \quad \text{m. c. d.}(m_{h,1}, m_{h,2}) \leq |m_{h,1}|^{\frac{n-2}{n-1}} = |m_{h,1}|^{\frac{q-1}{q}}. \tag{2.54}$$

Sean $\eta = \text{ind}(P; M_1,\dots,M_m; r_1,\dots,r_m)$ e $I(\eta)$ al ideal generado por $M_1^{i_1}\cdots M_m^{i_m}$ con $\mathbf{i}\cdot\mathbf{r}^{-1} \geq \eta$. Así, $P \in I(\eta)$.

II. Se construirá un polinomio $\widehat{P} = \widehat{P}(X_{1,1}, X_{1,2}; X_{2,1}, X_{2,2}; \dots; X_{m,1}, X_{m,2})$. Si $n = 2$, entonces $\hat{P} = P$ y cuando $n > 2$, se eliminan metódicamente las $p = m(n-2)$ variables $X_{1,3},\dots,X_{1,n},\dots,X_{m,3},\dots,X_{m,n}$.

1. Se escribe, igual que antes, a P como un polinomio en

$$M_1, X_{1,2},\dots,X_{1,n}, M_2, X_{2,2},\dots,X_{2,n},\dots,M_m,X_{m,2},\dots,X_{m,n}.$$

2. Sea $a \in \mathbb{N}_0$ tal que $X_{1,3}^a \| P$, se define $\overline{P} = X_{1,3}^{-a}P$. De este modo, $\overline{P} \in \mathbb{R}[\mathbf{X}]$, $\mathcal{H}(\overline{P}) = \mathcal{H}(P)$, $\overline{P} \in I(\eta)$ y tiene grado a lo más r_h para cada $h \in [1..m]$.
3. Se define al polinomio

$$P^{(1)} = \overline{P}(X_{1,1}, X_{1,2}, 0, X_{1,4},\dots,X_{1,n}; X_{2,1},\dots,X_{2,n};\dots;X_{m,1},\dots,X_{m,n}) \neq 0.$$

Para convencerse de $P^{(1)} \neq 0$, hay que observar cómo se construyó $\overline{P}$. Escribiendo a P como

$$P(X_{1,1},\dots,X_{m,n}) = \sum c(j_{1,1},\dots,j_{m,n})X_{1,1}^{j_{1,1}}\cdots X_{1,n}^{j_{1,n}},$$

a es justo el menor exponente positivo con el que aparece $X_{1,3}$. Entonces, $\overline{P}$ tiene al menos un sumando en el que no aparece $X_{1,3}$ y, por lo tanto, $P^{(1)}$ tiene al menos un sumando distinto de cero. Por construcción, $\mathcal{H}(P^{(1)}) \leq \mathcal{H}(\overline{P})$, $P^{(1)}$ es homogéneo de grado a lo más r_1 en $X_{1,1}, X_{1,2}, X_{1,4},\dots,X_{1,n}$, de grado r_h en $X_{h,1},\dots,X_{h,n}$ para $h \in [2..m]$ y $P^{(1)} \in I^{(n)}(\eta)$, donde $I^{(n)}(\eta)$ es el ideal del anillo de polinomios en las $nm-1$ variables restantes generado por

$$M_1(X_{1,1}, X_{1,2}, 0, X_{1,4},\dots,X_{1,n})^{i_1} M_2(X_{2,1},\dots,X_{2,n})^{i_2}\cdots M_m(X_{m,1},\dots,X_{m,n})^{i_m}$$

con $\mathbf{i} = (i_1,\dots,i_m)$ cumpliendo con $\mathbf{i}\cdot\mathbf{r}^{-1} \geq \eta$.

4. Repitiendo este procedimiento p veces para que queden sólo $X_{h,1}, X_{h,2}$, $h \in [1..m]$, se obtiene un polinomio $P^{(p)}$ y se define $\widehat{P} = P^{(p)}$. $\widehat{P}$ es homogéneo en $X_{h,1}, X_{h,2}$ de grado a lo más r_h para $h \in [1..m]$ y $\widehat{P} \in \widehat{I}(\eta)$, el ideal de $\mathbb{Q}[X_{1,1}, X_{1,2},\dots,X_{m,1}, X_{m,2}]$ generado por los polinomios

$$(u_{1,1}X_{1,1} + u_{1,2}X_{1,2})^{i_1}(u_{2,1}X_{1,1} + u_{2,2}X_{1,2})^{i_2}\cdots(u_{m,1}X_{m,1} + u_{m,2}X_{m,2})^{i_m}$$

con el multi-índice $\mathbf{i}$ satisfaciendo $\mathbf{i}\cdot\mathbf{r}^{-1} \geq \eta$.

III. Por la homogeneidad de $\widehat{P} \neq 0$ cada pareja $X_{h,1}, X_{h,2}$ con $h \in [1..n]$,

$$\tilde{P}(X_1,\dots,X_m) \coloneqq \widehat{P}(X_1, 1, X_2, 1,\dots,X_m, 1) \neq 0.$$

También, $\mathcal{H}(\tilde{P}) = \mathcal{H}(\widehat{P}) \leq \mathcal{H}(P)$ y $\tilde{P} \in \tilde{I}(\eta)$, el ideal en $\mathbb{Q}[X_1,\dots,X_m]$ generado por

$$\left(X_1 + \frac{u_{1,2}}{u_{1,1}}\right)^{i_1}\left(X_2 + \frac{u_{2,2}}{u_{2,1}}\right)^{i_1}\cdots\left(X_m + \frac{u_{m,2}}{u_{m,1}}\right)^{i_m}$$

con $\mathbf{i}\cdot\mathbf{r}^{-1}\geq\eta$. Definiendo

$$\forall h\in[1..m]\quad q_h:=\frac{|u_{h,1}|}{\mathrm{m.c.d.}(u_{h,1},u_{h,2})},\quad p_h:=-\frac{u_{h,2}}{\mathrm{m.c.d.}(u_{h,1},u_{h,2})}\frac{u_{h,1}}{|u_{h,1}|}$$

se cumple $q_h>0$ y $\mathrm{m.c.d.}(q_h,p_h)=1$ para $h\in[1..m]$. Además, si se reescribe a los polinomios que generan a $I(\eta)$ como

$$\left(X_1-\frac{p_1}{q_1}\right)^{i_1}\cdots\left(X_m-\frac{p_m}{q_m}\right)^{i_m}.$$

Entonces,

$$\forall\mathbf{j}\in\mathbb{N}_0^n\quad\left(\mathbf{j}\cdot\mathbf{r}^{-1}\geq\eta\implies P_{\mathbf{j}}\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right)=0\right).$$

Luego, con $\mathbf{p}=(p_1,\ldots,p_m)$ y $\mathbf{q}=(q_1,\ldots,q_m)$,

$$\mathrm{ind}_{\mathbf{r}}\left(\tilde{P};\frac{\mathbf{p}}{\mathbf{q}}\right)\geq\eta.$$

IV. Ahora, simplemente se verá que $\tilde{P}$ satisface las hipótesis del Lema Fundamental de Roth (Teorema 1.6) para arribar a

$$\eta\leq\mathrm{ind}_{\mathbf{r}}\left(\tilde{P};\frac{\mathbf{p}}{\mathbf{q}}\right)\leq\varepsilon.$$

Tres condiciones del Lema Fundamental de Roth– (1.6), (1.7) y (1.8)– valen por hipótesis. Nótese que

$$\forall h\in[1..m]\quad\mathcal{H}(M_h)^{\frac{1}{q}}=|u_{h,1}|^{\frac{1}{q}}\leq\frac{|u_{h,1}|}{\mathrm{m.c.d.}(u_{h,1},u_{h,2})}=q_h\leq\mathcal{H}(M_h).$$

Entonces, llamando $\gamma=\xi/q\in(0,1]$, se obtiene la primera condición en (1.9):

$$\forall h\in[1..m]\quad q_h^{r_h}\geq\mathcal{H}(M_h)^{\frac{r_h}{q}}\geq\mathcal{H}(M_1)^{\frac{r_1\xi}{q}}=\mathcal{H}(M_1)^{r_1\gamma}\geq q_1^{r_1\gamma}.$$

Con lo anterior y (2.42) se llega a la segunda desigualdad en (1.9):

$$\forall h\in[1..m]\quad q_h^{\omega}\geq q_h^{\omega\gamma}=q_h^{\frac{\omega\xi}{q}}\geq\mathcal{H}(M_h)^{\frac{\omega\xi}{q^2}}\geq 2^{3m}.$$

Finalmente, con la desigualdad previa y (2.41) se verifica la hipótesis sobre la altura,

$$\mathcal{H}(\tilde{P})\leq\mathcal{H}(P)\leq\mathcal{H}(M_1)^{\frac{\omega r_1\xi}{q^2}}=\mathcal{H}(M_1)^{\frac{\omega r_1\gamma}{q}}\leq q_1^{\omega\gamma r_1}\leq q_1^{\omega r_1}.$$

Por lo tanto, son ciertas las hipótesis del Lema Fundamental de Roth y $\eta\leq\varepsilon$. □

Capítulo 3

Una infinidad de números trascendentes mal aproximables

A finales de los años noventa, Martine Queffélec dio un método para construir números trascendentes a través de la sucesión de Thue-Morse. Poco tiempo después, parándose sobre el Teorema del Subespacio de Schmidt y resultados propios sobre fracciones continuadas palindrómicas, Yann Bugeaud y Boris Adamczewski llegaron a una proposición que incluía el resultado de Queffélec. Aprovechando un teorema reciente de Bugeaud (2012), se dará otra generalización del trabajo de Queffélec.

En este último capítulo se construye una familia infinita de números trascendentes mal aproximables. En primer lugar, se establecen algunos términos de la teoría de las sucesiones automáticas y se presenta la notación. Después, se expone de manera detallada las demostraciones de los resultados de Bugeaud. Finalmente, juntando todos los elementos, se concluye el teorema que contiene como caso particular al resultado de Queffélec sobre sucesiones de Thue-Morse. La formulación del teorema principal requiere un poco de terminología, se pospone su enunciado.

Las condiciones de trascendencia aquí estudiadas giran en torno a la aproximación mediante números racionales o irracionales cuadráticos. La maquinaria que se empleará para construir estas aproximaciones es la teoría de las fracciones continuadas. Si bien todo lo requerido sobre fracciones continuadas es elemental, hay lemas no tan famosos. Para evitar dificultades, todos los resultados son enunciados en el Apéndice A.

3.1. Definiciones de la teoría de sucesiones automáticas

Los resultados principales de este capítulo se apoyan en la estructura de la expansión en fracción continuada de los números irracionales para obtener propiedades algebraicas. Este acercamiento conduce a la famosa caracterización de los irracionales cuadráticos debida a Lagrange. No será sorprendente que este teorema se utilice varias veces.

Definición 3.1. *Si $\mathcal{A}$ es un conjunto no vacío, una palabra en $\mathcal{A}$ es una sucesión finita*

o infinita en $\mathcal{A}$. Si $\mathbf{a} = (a_i)_{i=1}^{n}$ es una palabra en $\mathcal{A}$, se escribe

$$\mathbf{a} = a_1 a_2 a_3 \dots a_n,$$

*los términos a_j se llaman **elementos**. La palabra **vacía**, ϵ, es aquélla que carece de elementos. Al conjunto de las palabras finitas en $\mathcal{A}$ se denotará $\mathcal{A}^*$; al de las palabras infinitas, $\mathcal{A}^\omega$; al de las palabras no vacías, $\mathcal{A}^+$ y al conjunto de todas las palabras, $\mathcal{A}^\infty$.*

La infinidad de la representación como fracción continuada de los irracionales sugiere que las palabras infinitas serán los objetos con los que se trabajará. En varias instancias se aproximará la palabra infinita de interés mediante otras que serán construidas a partir de ciertas palabras finitas.

Definición 3.2. *Si $U \in \mathcal{A}^\infty$, entonces una palabra finita V es un **prefijo** de U si existe $W \in \mathcal{A}^\infty$ tal que $U = VW$.*

Definición 3.3. *Si $\mathbf{a} \in \mathcal{A}^*$, entonces la **longitud** de $\mathbf{a} = a_1 a_2 \dots a_n$, $|\mathbf{a}|$, es n.*

Si $\mathbf{a} = a_1 \dots a_n$, $\mathbf{b} = b_1 \dots b_m \in \mathcal{A}^*$, la concatenación de $\mathbf{a}$ y $\mathbf{b}$ es la palabra $\mathbf{ab} = a_1 \dots a_n b_1 \dots b_m$. Las potencias n-ésimas se definen de manera natural:

$$\forall \mathbf{a} \in \mathcal{A}^* \quad \mathbf{a}^1 := \mathbf{a}, \quad \forall n \in \mathbb{N}\ \mathbf{a}^{n+1} = \mathbf{a}\mathbf{a}^n.$$

También se consideran potencias racionales positivas. Cuando $r \in \mathbb{Q}_{>0}$, si $\mathbf{a}'$ es el prefijo de $\mathbf{a}$ de longitud $\lceil (q - \lfloor q \rfloor)|\mathbf{a}| \rceil$,

$$\forall \mathbf{a} \in \mathcal{A}^* \quad \mathbf{a}^r := \mathbf{a}^{\lfloor r \rfloor}\mathbf{a}',$$

Definición 3.4. *Una palabra $\mathbf{a} \in \mathcal{A}^\omega$, $\mathbf{a} = a_1 a_2 a_3 \dots$, es **puramente periódica** si existe un natural m tal que*

$$\forall n \in \mathbb{N} \quad a_n = a_{m+n}.$$

*Una palabra $\mathbf{a} \in \mathcal{A}^\omega$ es **tarde o temprano periódica** si existen $\mathbf{b} \in \mathcal{A}^*$ y $\mathbf{c} \in \mathcal{A}^\omega$ puramente periódica tales que $\mathbf{a} = \mathbf{bc}$.*

En estos términos, la caracterización de Lagrange dice que un número irracional $\alpha = [a_0; a_1, a_2, \dots]$ es irracional si y sólo si la palabra $\mathbf{a} = a_0 a_1 a_2 \dots$ es tarde o temprano periódica.

A grandes rasgos, el Teorema de Bugeaud que se utilizará establece que un real algebraico de grado al menos tres no puede tener una representación *sencilla* como fracción continuada. La última definición ayudará a aclarar qué significa esta sencillez.

Definición 3.5. *Tómese $\mathbf{a} \in \mathcal{A}^\omega$. Se define la **función de complejidad**, $p_{\mathbf{a}} : \mathbb{N} \to \mathbb{N} \cup \{+\infty\}$, como*

$$p_{\mathbf{a}}(n) := |\{\mathbf{b} \in \mathcal{A}^* : \mathbf{b} \text{ es subpalabra de } \mathbf{a},\ |\mathbf{b}| = n\}|.$$

Claramente, si $\mathbf{a}$ es una palabra sobre un alfabeto finito con k elementos, entonces $p_{\mathbf{a}}(n) \leq k^n$ y cuando se constituye por una infinidad de letras, $p_{\mathbf{a}}(n) = +\infty$ para cualquier $n \in \mathbb{N}$.

3.2. Teorema de Bugeaud

La base sobre la cual se construirá una infinidad de reales trascendentes es el Teorema 3.1. Hay que tener un poco de cuidado con esta proposición. No establece que los irracionales algebraicos tengan elementos acotados. Mucho menos lo afirma para los trascendentes, el Teorema de Liouville da trascendentes con elementos arbitrariamente grandes y con argumentos de numerabilidad se concluye la existencia trascendentes con elementos acotados.

Teorema 3.1. *Sea* $\mathbf{a} = (a_n)_{n=1}^{\infty}$ *una sucesión de naturales que no es periódica tarde o temprano. Si el real* $\alpha := [0; a_1, a_2, \ldots]$ *es algebraico, entonces*

$$\lim_{n\to\infty} \frac{p_{\mathbf{a}}(n)}{n} = +\infty. \tag{3.1}$$

El Teorema 3.1 es consecuencia del más profundo Teorema 3.2, que se apoya fuertemente en el Teorema del Subespacio de Schmidt. De hecho, para concluir el Teorema 3.1 se apela a la formulación contrapuesta y simplemente se verifica que las hipótesis del Teorema 3.2 son satisfechas.

Definición 3.6. *Sea* $\mathbf{a} = (a_n)_{n=1}^{\infty}$ *una sucesión en* $\mathcal{A}$*. Se dice que* $\mathbf{a}$ *satisface la condición* (♠) *si no es tarde o temprano periódica y existen sucesiones en* $\mathcal{A}^*$*,* $(U_n)_{n=1}^{\infty}$*,* $(V_n)_{n=1}^{\infty}$ *y* $(W_n)_{n=1}^{\infty}$*, tales que*

I. $\forall n \in \mathbb{N}$ *la palabra* $W_n U_n V_n U_n$ *es un prefijo de* $\mathbf{a}$*,*

II. $(|V_n||U_n|^{-1})_{n=1}^{\infty}$ *es acotada,*

III. $(|W_n||U_n|^{-1})_{n=1}^{\infty}$ *es acotada,*

IV. $(|U_n|)_{n=1}^{\infty}$ *es creciente.*

Teorema 3.2. *Sean* $\mathbf{a} = (a_n)_{n=1}^{\infty}$ *una sucesión en* $\mathbb{N}$*,* $\alpha := [0; a_1, a_2, \ldots]$ *y* $(p_n/q_n)_{n=1}^{\infty}$ *los convergentes de* α*. Si* $(a_n)_{n=1}^{\infty}$ *satisface la condición* (♠) *y* $\limsup \sqrt[n]{q_n} < +\infty$*, entonces* α *es trascendente.*

Para probar el Teorema 3.2 se supondrá que $\alpha \in \mathbb{R}$ es algebraico. Como la sucesión de sus elementos es infinita y no es periódica, el grado de α es al menos tres. Tras aproximar mediante irracionales cuadráticos, se construyen transformaciones lineales y sucesiones de puntos racionales que satisfagan las hipótesis del Teorema del Subespacio de Schmidt. Con la asistencia de un vector ortogonal a uno de los subespacios propios se construye, tras un proceso iterativo, un polinomio cuadrático en $\mathbb{Z}[x]$ satisfecho por α, contradiciendo $[\mathbb{Q}(\alpha) : \mathbb{Q}] \geq 3$.

En las pruebas de los dos teoremas anteriores se adopta la notación tradicional, los convergentes de α serán $(p_k/q_k)_{k=1}^{\infty}$.

Demostración del Teorema 3.1. Supóngase cierto el Teorema 3.2. Si el límite en (3.1) no se cumple, existe $C \geq 2$ tal que $\mathcal{N} = \{n \in \mathbb{N} : p_{\mathbf{a}}(n) \leq nC\}$ es infinito, por lo que $\mathbf{a}$ es una palabra sobre un alfabeto finito y $\alpha := [0; a_1, a_2, \ldots]$ satisface (cfr. Apéndice A, Teorema 11)

$$\limsup_{n\to\infty} \sqrt[n]{q_n} < +\infty.$$

Para construir las sucesiones exigidas por el Teorema 3.2 se fija a $n \in \mathcal{N}$. Por el Principio del Palomar, hay una subpalabra X_n de $\mathbf{a}$, $|X_n| = n$, que aparece al menos un

par de veces en $a_1a_2\dots a_{C(n+1)}$. Nótese que existen palabras W_n, W'_n, B_n, B'_n para las que

$$a_1a_2\dots a_{C(n+1)} = W_nX_nB_n = W'_nX_nB'_n, \quad |W_n| < |W'_n|.$$

La longitud de W_n y la de W'_n dan dos casos a considerar, éstos son ilustrados en la Figura 1 y en la Figura 2.

Caso i. Si $|W_nX_n| \le |W'_n|$, V_n es la palabra para la cual $W_nX_nV_n = W'_n$, entonces

$$a_1a_2\dots a_{C(n+1)} = W_nX_nV_nX_nB'_n.$$

Así, $C(n+1) = |W_n| + |X_n| + |V_n| + |X_n| + |B'_n|$, de donde

$$|W_n| + |V_n| \le (C+1)n - 2n = (C-1)n \implies \frac{|W_n| + |V_n|}{|X_n|} \le C - 1 \ll 1. \tag{3.2}$$

Se define $U_n := X_n$.

Caso ii. Si $|W_nX_n| > |W'_n|$, X'_n es la palabra que cumple con $W'_n = W_nX'_n$; entonces, $X_nB_n = X'_nX_nB'_n$ y

$$|W_n| + |X'_n| = |W'_n| < |W_n| + |X_n| \implies |X'_n| < |X_n|.$$

En consecuencia, X_n es una potencia racional de X'_n (Figura 2). Sean $x_n \in \mathbb{N}$ y $y_n \in \mathbb{Q} \cap [0,2)$ tales que

$$X'_nX_n = X_n^{\prime 1+\frac{|X_n|}{|X'_n|}} = X_n^{\prime 2x_n+y_n} = (X_n^{\prime x_n})^2 X_n^{\prime y_n}.$$

Calculando longitudes,

$$\begin{aligned} |X'_nX_n| = 2x_n|X'_n| + y|X'_n| \le 2x_n|X'_n| + 2|X'_n| &\implies n = |X_n| \le (2x_n+1)|X'_n| \le 3x_n|X'_n| \\ &\implies |X_n^{\prime x_n}| \ge \frac{n}{3}. \end{aligned}$$

Luego, $W_nX_n^{\prime x_n}X_n^{\prime x_n}$ es un prefijo de $\mathbf{a}$ tal que

$$\frac{|W_n|}{|X_n^{\prime x_n}|} \le \frac{3}{n}\left[(C+1)n - 2|X_n^{\prime x_n}|\right] \le 3(C+1).$$

Con $U_n := X_n^{\prime x_n}$ y $V_n := \epsilon$, $W_nU_nV_nU_n$ es un prefijo de $\mathbf{a}$ que satisface $|W_n| + |V_n| \le (3C+3)|U_n|$, que es

$$\frac{|W_n| + |V_n|}{|U_n|} \ll 1. \tag{3.3}$$

Las sucesiones $(U_n)_{n\in\mathcal{N}}$, $(V_n)_{n\in\mathcal{N}}$, $(W_n)_{n\in\mathcal{N}}$ se reetiquetan para que estén indicadas en los naturales y, tomando una subsucesión de ser necesario, puede pensarse que $(|U_n|)_{n=1}^{\infty}$ es estrictamente creciente. Las desigualdades (3.2) y (3.3) garantizan las condiciones II. y III. del Teorema 3.2; por lo tanto, α es trascendente. □

Demotración del Teorema 3.2. I. Por el Teorema de Lagrange y (♠), $[\mathbb{Q}(\alpha) : \mathbb{Q}] \ge 3$. El objetivo es llegar a que α es un irracional cuadrático, una contradicción. Sean las sucesiones $(u_n)_{n=1}^{\infty}$, $(v_n)_{n=1}^{\infty}$, $(w_n)_{n=1}^{\infty}$ dadas por

$$\forall n \in \mathbb{N} \qquad u_n := |U_n|, \quad v_n := |V_n|, \quad w_n := |W_n|.$$

W_n	X_n	B_n		
W'_n			X_n	B'_n
W_n	X_n	V_n	X_n	B'_n

Figura 3.1: Configuración en el Caso I.

W_n	X_n	B_n	
W'_n		X_n	B'_n
W_n	X'_n	X_n	B'_n

Figura 3.2: Configuración en el Caso II.

Ahora, se aproxima a α mediante irracionales cuadráticos $(\alpha_n)_{n=1}^{\infty}$ definidos a través de su expansión en fracción continuada. Denotando A_n a la palabra formada por los elementos de α_n,

$$\forall n \in \mathbb{N} \quad A_n := W_n U_n V_n U_n V_n U_n V_n \dots.$$

La coincidencia de los primeros $w_n + u_n + v_n + u_n$ elementos de α y α_n conlleva (cfr. Apéndice A, Teorema 10)

$$\forall n \in \mathbb{N} \quad |\alpha - \alpha_n| \leq \frac{2}{q_{w_n+2u_n+v_n}}.$$

II. Los reales α_n dan lugar a aparatosos polinomios que producen cotas, con ellas se verá que ciertas transformaciones lineales satisfacen las hipótesis del Teorema del Subespacio. Unas cuentas directas y muy largas junto con

$$\forall n \in \mathbb{N} \quad \alpha_n = \frac{p_{w_n} r_n + p_{w_n-1}}{q_{w_n} r_n + q_{w_n-1}} = \frac{p_{w_n+u_n+v_n} r_n + p_{w_n+u_n+v_n-1}}{q_{w_n+u_n+v_n} r_n + q_{w_n+u_n+v_n-1}}, \quad r_n = \frac{q_{w_n-1} r_n - p_{w_n-1}}{-q_{w_n} r_n + p_{w_n}},$$

donde los elementos de r_n forman $U_n V_n U_n V_n \dots$, muestran que α_n anula a

$$P_n(X) := \begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix} X^2 - \left(\begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} + \begin{vmatrix} p_{w_n-1} & q_{w_n+u_n+v_n-1} \\ p_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix} \right) X + \\ + \begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1} \\ p_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}$$

La precisión con la que los convergentes se acercan a α (cfr. Apéndice A, Teorema 7) conduce a

$$\left| \begin{vmatrix} q_{w_n-1} & q_{w_+u_n+v_n-1} \\ q_{w_n} & q_{w_+u_n+v_n} \end{vmatrix} \alpha - \begin{vmatrix} p_{w_n-1} & q_{w_+u_n+v_n-1} \\ p_{w_n} & q_{w_+u_n+v_n} \end{vmatrix} \right| = \left| \det \begin{pmatrix} q_{w_n-1}\alpha - p_{w_n-1} & q_{w_+u_n+v_n-1} \\ q_{w_n}\alpha - p_{w_n} & q_{w_+u_n+v_n} \end{pmatrix} \right| \\ \leq q_{w_n+u_n+v_n} |q_{w_n-1}\alpha - p_{w_n-1}| + \\ + q_{w_n+u_n+v_n-1} |q_{w_n}\alpha - p_{w_n}| \\ \leq \frac{q_{w_n+u_n+v_n}}{q_{w_n}} + \frac{q_{w_n+u_n+v_n-1}}{q_{w_n}} \\ \leq 2 \frac{q_{w_n+u_n+v_n}}{q_{w_n}}. \tag{3.4}$$

Similarmente,

$$\left|\begin{vmatrix} q_{w_n-1} & q_{w_+u_n+v_n-1} \\ q_{w_n} & q_{w_+u_n+v_n} \end{vmatrix}\alpha - \begin{vmatrix} q_{w_n-1} & p_{w_+u_n+v_n-1} \\ q_{w_n} & p_{w_+u_n+v_n} \end{vmatrix}\right| \leq \frac{2q_{w_n}}{q_{w_n+v_n+u_n}}. \tag{3.5}$$

Como los primeros $w_n + 2u_n + v_n$ elementos de α y α_n coinciden, se obtiene una cota parecida a la anterior:

$$\begin{aligned}\left|\begin{vmatrix} q_{w_n-1} & q_{w_+u_n+v_n-1} \\ q_{w_n} & q_{w_+u_n+v_n-1} \end{vmatrix}\alpha_n - \begin{vmatrix} q_{w_n-1} & p_{w_+u_n+v_n-1} \\ q_{w_n} & p_{w_+u_n+v_n} \end{vmatrix}\right| &= \left|\det\begin{pmatrix} q_{w_n-1} & q_{w_+u_n+v_n-1}\alpha_n - p_{w_+u_n+v_n-1} \\ q_{w_n} & q_{w_+u_n+v_n}\alpha_n - p_{w_+u_n+v_n} \end{pmatrix}\right| \\ &\leq q_{w_n-1}|q_{w_n+u_n+v_n}\alpha_n - p_{w_n+u_n+v_n}| + \\ &\quad + q_{w_n}|q_{w_n+u_n+v_n-1}\alpha_n - p_{w_n+u_n+v_n-1}| \\ &\leq \frac{q_{w_n-1}}{q_{w_n+u_n+v_n}} + \frac{q_{w_n}}{q_{w_n+u_n+v_n}} \\ &\leq 2\frac{q_{w_n}}{q_{w_n+u_n+v_n}}. \end{aligned} \tag{3.6}$$

Usando (3.6) y (3.4),

$$\begin{aligned}|P_n(\alpha)| &= |P_n(\alpha) - P_n(\alpha_n)| \\ &= |\alpha - \alpha_n|\left|\begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}(\alpha + \alpha_n) - \begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} - \begin{vmatrix} p_{w_n-1} & q_{w_n+u_n+v_n-1} \\ p_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}\right| \\ &= |\alpha - \alpha_n|\left|\begin{vmatrix} q_{w_n-1}\alpha - p_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n}\alpha - p_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix} + \begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1}\alpha_n - p_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n}\alpha_n - p_{w_n+u_n+v_n} \end{vmatrix}\right| \\ &\ll |\alpha - \alpha_n|\left(\frac{q_{w_n+u_n+v_n}}{q_{w_n}} + \frac{q_{w_n}}{q_{w_n+u_n+v_n}}\right) \\ &\ll \frac{1}{q^2_{w_n+2u_n+v_n}}\left(\frac{q_{w_n+u_n+v_n}}{q_{w_n}} + \frac{q_{w_n}}{q_{w_n+u_n+v_n}}\right) \\ &\ll \frac{q_{w_n+u_n+v_n}}{q_{w_n}q^2_{w_n+2u_n+v_n}} \end{aligned} \tag{3.7}$$

III. Sea $\mathscr{L}^{(1)} : \mathbb{R}^4 \to \mathbb{R}^4$, $\mathbf{X} = (X_1, X_2, X_3, X_4)$, dada por

$$\mathscr{L}^{(1)}(\mathbf{X}) := \begin{pmatrix} \mathscr{L}^{(1)}_1(\mathbf{X}) \\ \mathscr{L}^{(1)}_2(\mathbf{X}) \\ \mathscr{L}^{(1)}_3(\mathbf{X}) \\ \mathscr{L}^{(1)}_4(\mathbf{X}) \end{pmatrix} := \begin{pmatrix} \alpha^2 & -\alpha & -\alpha & 1 \\ \alpha & -1 & 0 & 0 \\ \alpha & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}.$$

Llámese $(\mathbf{r}_n)_{n=1}^{\infty}$ a la sucesión en $\mathbb{R}^4$, $\mathbf{r}_n = (r_{n,1}, r_{n,2}, r_{n,3}, r_{n,4})$ dada por

$$\begin{aligned} r_{n,1} &:= \begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}, \quad r_{n,2} := \begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}, \\ r_{n,3} &:= \begin{vmatrix} p_{w_n-1} & q_{w_n+u_n+v_n-1} \\ p_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}, \quad r_{n,4} := \begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1} \\ p_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}. \end{aligned}$$

Como $\alpha \in (0,1)$ y $(q_k)_{k=1}^{\infty}$ es estrictamente creciente, $\|\mathbf{r}_n\|_{\infty} \ll q_{w_n}q_{w_n+u_n+v_n}$; por ejemplo,

$$\begin{aligned}|r_{n,1}| = |q_{w_n-1}q_{w_n+u_n+v_n} - q_{w_n+u_n+v_n-1}q_{w_n}| &= q_{w_n}q_{w_n+u_n+v_n}\left|\frac{q_{w_n-1}}{q_{w_n}} - \frac{q_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}}\right| \\ &\leq 2q_{w_n}q_{w_n+u_n+v_n}. \end{aligned} \tag{3.8}$$

Se evalúa $\mathscr{L}^{(1)}$ en $\mathbf{r}_n$ y de (3.7), (3.5), (3.4) y (3.8) se obtiene para $n \in \mathbb{N}$ suficientemente grande

$$\begin{aligned}\prod_{j=1}^{4}\left|\mathscr{L}_j^{(1)}(\mathbf{r}_n)\right| &= |P_n(\alpha)|\,|\alpha r_{n,1}-r_{n,2}|\,|\alpha r_{n,1}-r_{n,3}|\,|r_{n,1}|\\ &\ll \frac{q_{w_n+u_n+v_n}}{q_{w_n}q_{w_n+2u_n+v_n}^2}\,\frac{q_{w_n}}{q_{w_n+u_n+v_n}}\,\frac{q_{w_n+u_n+v_n}}{q_{w_n}}\,q_{w_n}q_{w_n+u_n+v_n}\\ &= \frac{q_{w_n+u_n+v_n}^2}{q_{w_n+2u_n+v_n}^2}\\ &\ll \frac{1}{2^{u_n}} \ll \left(q_{w_n}q_{w_n+u_n+v_n}\right)^{-\delta u_n s_n},\end{aligned}$$

en donde $s_n = (2w_n+u_n+v_n)^{-1}$, $\delta = \log 2/\log M$ y $M = 1+\limsup_{k\to\infty} q_k^{\frac{1}{k}}$. La última desigualdad vale porque para $n \in \mathbb{N}$ suficientemente grande

$$\begin{aligned}2 = \left(M^{w_n}M^{w_n+u_n+v_n}\right)^{\frac{\log 2}{\log M}(2w_n+u_n+v_n)^{-1}} &\geq \left(q_{w_n}q_{w_n+u_n+v_n}\right)^{\frac{\log 2}{\log M}(2w_n+u_n+v_n)^{-1}}\\ &= \left(q_{w_n}q_{w_n+u_n+v_n}\right)^{\delta s_n}.\end{aligned}$$

La condición (♠) da la existencia de constantes $M_1, M_2 > 0$ tales que

$$\liminf_{n\to\infty} u_n s_n = \liminf_{n\to\infty}\frac{u_n}{2w_n+u_n+v_n} = \liminf_{n\to\infty}\frac{1}{\frac{2w_n}{u_n}+1+\frac{v_n}{u_n}} \geq \liminf_{n\to\infty}\frac{1}{M_1+1+M_2} > 0.$$

Entonces, puede fijarse $\varepsilon > 0$ tal que para $n \in \mathbb{N}$ grande

$$\prod_{j=1}^{4}\left|\mathscr{L}_j^{(1)}(\mathbf{r}_n)\right| \ll \frac{1}{\left(q_{w_n}q_{w_n+u_n+v_n}\right)^{\varepsilon}} \ll \frac{1}{\|\mathbf{r}_n\|_\infty^{\varepsilon}}. \tag{3.9}$$

Por el Teorema del Subespacio de Schmidt, los $\mathbf{r}_n$ yacen en una cantidad finita de subespacios propios de $\mathbb{Q}^4$; por ende, existe uno, T_1, con una infinidad de puntos $\mathbf{r}_n$. Se define $\mathcal{N}_1 = \{n \in \mathbb{N} : \mathbf{r}_n \in T_1\} \subseteq \mathbb{N}$. Fíjese $\mathbf{x} \in T_1^{\perp} \cap \mathbb{Z}^4$, $\mathbf{x} = (x_1, x_2, x_3, x_4)$. Luego,

$$\forall n \in \mathcal{N}_1 \quad \langle \mathbf{x}, \mathbf{r}_n\rangle = 0.$$

Expandiendo la ecuación anterior se ve que cualquier $n \in \mathcal{N}_1$ cumple con

$$\begin{aligned}0 = x_1\begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1}\\ q_{w_n} & q_{w_n+u_n+v_n}\end{vmatrix} + x_2\begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1}\\ q_{w_n} & p_{w_n+u_n+v_n}\end{vmatrix} + x_3\begin{vmatrix} p_{w_n-1} & q_{w_n+u_n+v_n-1}\\ p_{w_n} & q_{w_n+u_n+v_n}\end{vmatrix} +\\ + x_4\begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1}\\ p_{w_n} & p_{w_n+u_n+v_n}\end{vmatrix}.\end{aligned} \tag{3.10}$$

IV. Ahora, se considera dos casos según si $(w_n)_{n=1}^{\infty}$ tiene o no una subsucesión constante.

IV.1.1. *Caso I.* $(w_n)_{n=1}^{\infty}$ tiene una subsucesión constante. Considerando una subsucesión y un residuo adecuados[1] puede suponerse que $w_n = 0$ para toda $n \in \mathbb{N}$. Entonces, como $q_{-1} = p_0 = 0$ y $p_{-1} = q_0 = 1$,

$$\forall n \in \mathcal{N}_1 \quad \mathbf{r}_n = (-q_{u_n+v_n-1}, -p_{u_n+v_n-1}, q_{u_n+v_n}, p_{u_n+v_n}) \implies$$

[1] El Teorema 6 del Apéndice A implica $[\mathbb{Q}(\alpha^{(m)}), \mathbb{Q})] = [\mathbb{Q}(\alpha), \mathbb{Q})]$ para cada residuo $\alpha^{(m)} = [0; a_{m+1}, a_{m+2}, \ldots]$.

que implica

$$\forall n \in \mathcal{N}_1 \quad x_1 q_{u_n+v_n-1} + x_2 p_{u_n+v_n-1} - x_3 q_{u_n+v_n} - x_4 p_{u_n+v_n} = 0. \tag{3.11}$$

De esta ecuación se obtiene $|x_1| + |x_2| > 0$, pues $|x_1| + |x_2| = 0$ conlleva $\mathbf{x} = \mathbf{0}$:

$$\mathbb{Q} \ni x_3 = \lim_{\substack{n\to\infty \\ n\in\mathcal{N}_1}} x_3 = \lim_{\substack{n\to\infty \\ n\in\mathcal{N}_1}} \frac{p_{u_n+v_n}}{q_{u_n+v_n}} = x_4\alpha \implies x_3 = x_4 = 0.$$

Dividiendo (3.11) entre $q_{u_n+v_n}$ se obtiene

$$\begin{aligned} &\forall n \in \mathcal{N}_1 \quad x_1 \frac{q_{u_n+v_n-1}}{q_{u_n+v_n}} + x_2 \frac{p_{u_n+v_n-1}}{q_{u_n+v_n-1}} \frac{q_{u_n+v_n-1}}{q_{u_n+v_n}} - x_3 - x_4 \frac{p_{u_n+v_n}}{q_{u_n+v_n}} = 0 \\ \implies &\forall n \in \mathcal{N}_1 \quad \frac{q_{u_n+v_n-1}}{q_{u_n+v_n}} \left(x_1 + x_2 \frac{p_{u_n+v_n-1}}{q_{u_n+v_n-1}} \right) = x_3 + x_4 \frac{p_{u_n+v_n}}{q_{u_n+v_n}} \\ \implies &\lim_{\substack{n\to\infty \\ n\in\mathcal{N}_1}} \frac{q_{u_n+v_n-1}}{q_{u_n+v_n}} = \frac{x_3 + x_4\alpha}{x_1 + x_2\alpha} =: \beta. \end{aligned}$$

IV.1.2. Se prueba ahora la irracionalidad de β. Primero, como α es irracional, existe $K_1 > 0$ tal que $K_1 < |x_1 + x_2 p_m/q_m|$ para m suficientemente grande. Entonces, hay constantes dependientes de α y $\mathbf{x} = \mathbf{x}(\alpha)$ que garantizan

$$\begin{aligned} \left| \beta - \frac{q_{u_n+v_n-1}}{q_{u_n+v_n}} \right| &= \left| \frac{x_3 + x_4\alpha}{x_1 + x_2\alpha} - \frac{x_3 + x_4 \dfrac{p_{u_n+v_n}}{q_{u_n+v_n}}}{x_1 + x_2 \dfrac{p_{u_n+v_n-1}}{q_{u_n+v_n-1}}} \right| \\ &\ll \left| \frac{1}{x_1 + x_2\alpha} - \frac{1}{x_1 + x_2 \dfrac{p_{u_n+v_n}}{q_{u_n+v_n}}} \right| + \left| \frac{\alpha}{x_1 + x_2\alpha} - \frac{\dfrac{p_{u_n+v_n}}{q_{u_n+v_n}}}{x_1 + x_2 \dfrac{p_{u_n+v_n-1}}{q_{u_n+v_n-1}}} \right| \\ &\ll \left| \frac{p_{u_n+v_n}}{q_{u_n+v_n}} - \alpha \right| + \left| \alpha - \frac{p_{u_n+v_n}}{q_{u_n+v_n}} \right| + \left| \frac{p_{u_n+v_n}}{q_{u_n+v_n}} - \frac{p_{u_n+v_n-1}}{q_{u_n+v_n-1}} \right| \\ &\ll \frac{2}{q_{u_n+v_n}^2} + \frac{1}{q_{u_n+v_n} q_{u_n+v_n-1}} \ll \frac{1}{q_{u_n+v_n} q_{u_n+v_n-1}}. \end{aligned} \tag{3.12}$$

La recurrencia de $(q_k)_{k=0}^{\infty}$ (cf. Apéndice A, Teorema 5) y las propiedades elementales del máximo común divisor implican que q_m y q_{m+1} son coprimos para cualquier natural m; en consecuencia, $q_{u_n+v_n-1}/q_{v_n+u_n}$ está en expresión mínima y todos estos racionales son distintos por pares. Entonces, puede pensarse que $q_{u_n+v_n-1}/q_{v_n+u_n} \neq \beta$ para toda $n \in \mathbb{N}$. Si $\beta = a/b$ fuese racional en expresión mínima, entonces $|\beta - q_{u_n+v_n-1} q_{v_n+u_n}^{-1}| \geq (b q_{v_n+u_n})^{-1}$. Juntando esta desigualdad con (3.12) se concluiría $q_{v_n+u_n} \ll b \ll 1$, contraviniendo el carácter estrictamente creciente de $(q_m)_{m=1}^{\infty}$.

IV.1.3. Sea $\mathscr{L}^{(2)} : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$\mathscr{L}^{(2)}(\mathbf{Y}) := \begin{pmatrix} \mathscr{L}_1^{(2)}(\mathbf{Y}) \\ \mathscr{L}_2^{(2)}(\mathbf{Y}) \\ \mathscr{L}_3^{(2)}(\mathbf{Y}) \end{pmatrix} := \begin{pmatrix} \beta & -1 & 0 \\ \alpha & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$

Como $\alpha \in (0,1)$, sus convergentes yacen en $[0,1]$ (cf. Apéndice A, Teorema 5) y

$$\forall\, n \in \mathcal{N}_1 \quad \|(q_{u_n+v_n}, q_{u_n+v_n-1}, p_{u_n+v_n})\|_\infty = \|(q_{u_n+v_n}, q_{u_n+v_n-1}, p_{u_n+v_n-1})\|_\infty = q_{u_n+v_n}.$$

Si $(\mathbf{w}_n)_{n\in\mathcal{N}_1}$ es la sucesión recién definida, (3.12) y las aproximaciones de los convergentes dicen que toda $n \in \mathcal{N}_1$ cumple

$$\begin{aligned}\prod_{j=1}^{3}\left|\mathscr{L}_j^{(2)}(\mathbf{w}_n)\right| &= |\beta q_{u_n+v_n} - q_{u_n+v_n-1}|\,|\alpha q_{u_n+v_n} - p_{u_n+v_n}|\,|q_{u_n+v_n-1}| \\ &\ll \frac{1}{q_{u_n+v_n-1}}\frac{1}{q_{u_n+v_n}} q_{u_n+v_n-1} \\ &= \frac{1}{q_{u_n+v_n}}.\end{aligned}$$

Por el Teorema del Subespacio de Schmidt, todos los $\mathbf{w}_n$ están en una cantidad finita de subespacios propios de $\mathbb{Q}^3$; por lo tanto, hay uno, T_2, con una infinidad de $\mathbf{w}_n$. Sean $\mathcal{N}_2 = \{n \in \mathcal{N}_1 : \mathbf{w}_n \in T_2\}$ y $\mathbf{0} \neq \mathbf{y} = (y_1, y_2, y_3) \in T_2^{\perp} \cap \mathbb{Z}^3$, entonces

$$\forall\, n \in \mathcal{N}_2 \quad y_1 q_{u_n+v_n} + y_2 q_{u_n+v_n-1} + y_3 p_{u_n+v_n} = 0.$$

Dividiendo entre $q_{u_n+v_n}$ y tomando el límite cuando $n \to \infty$ a lo largo de $\mathcal{N}_2$ se concluye

$$y_1 + y_2\beta + y_3\alpha = 0. \tag{3.13}$$

Como β no es racional, $y_3 \neq 0$. Similarmente, al evaluar $\mathscr{L}^{(2)}$ en $(q_{u_n+v_n}, q_{u_n+v_n-1}, p_{u_n+v_n-1})$ $(n \in \mathcal{N}_1)$ se obtiene $\mathbf{0} \neq \mathbf{z} \in \mathbb{Z}^3$ y $\mathcal{N}_3 \subseteq \mathcal{N}_1$ infinito tales que

$$\forall\, n \in \mathcal{N}_3 \quad z_1 q_{u_n+v_n} + z_2 q_{u_n+v_n-1} + z_3 p_{u_n+v_n-1} = 0.$$

Dividiendo entre $q_{u_n+v_n-1}$ y tomando el límite a lo largo de $\mathcal{N}_3$,

$$\frac{z_1}{\beta} + z_2 + z_3\alpha = 0. \tag{3.14}$$

La irracionalidad de β implica $z_3 \neq 0$. Aislando y multiplicando los términos con β en (3.13) y (3.14) se llega a

$$(z_3\alpha + z_2)(y_3\alpha + y_1) = y_2 z_1.$$

Ya que $y_3 z_3 \neq 0$, α satisface un polinomio cuadrático en $\mathbb{Z}[X]$, contradiciendo $[\mathbb{Q}(\alpha) : \mathbb{Q}] \geq 3$.

IV.2. *Caso II.* Ahora se supondrá que $(w_n)_{n=1}^{\infty}$ no tiene una subsucesión constante. Igual que en el caso anterior, el objetivo es encontrar un polinomio cuadrático con coeficientes en los enteros racionales que tenga a α por raíz para llegar a una contradicción.

IV.2.1. Tómese un subconjunto $\mathcal{N}_4 \subseteq \mathcal{N}_1$ tal que $(w_n)_{n\in\mathcal{N}_4}$ sea estrictamente creciente, esto da

$$\lim_{\substack{n\to\infty \\ n\in\mathcal{N}_4}} \frac{p_{w_n}}{q_{w_n}} = \lim_{\substack{n\to\infty \\ n\in\mathcal{N}_4}} \frac{p_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}} = \alpha.$$

Sin pérdida de generalidad, se puede suponer que

$$\forall n \in \mathcal{N}_4 \quad a_{w_n} \neq a_{w_n+u_n+v_n}. \tag{3.15}$$

Para verlo, piénsese que $\mathbf{a}$ comienza con $WUVU$ y $a_{|W|} = a_{|W|+|U|+|V|}$. Si $V = \epsilon$, se definen W' y V' mediante

$$W = W'a, \quad U = U'a \implies WUVU = W'aU'aU'a = W'(aU')(aU')a.$$

Por otra parte, si $V \neq \epsilon$, se definen W' y V' con

$$W = W'a, \quad V = V'a \implies WUVU = W'aUV'aU = W'(aU)V'(aU).$$

Iterando, si es que se requiere, se llega a (3.15).

IV.1.2. Para evitar que la notación se vuelva pesada, se adopta

$$\forall n \in \mathbb{N} \quad Q_n := \frac{q_{w_n-1}q_{w_n+u_n+v_n}}{q_{w_n}q_{w_n+u_n+v_n-1}}, \quad R_n := \alpha - \frac{p_n}{q_n}.$$

Por la teoría de fracciones continuadas (cf. Apéndice A, Teorema 1), $R_k \searrow 0$ cuando $n \to \infty$; además,

$$\forall n \in \mathbb{N} \quad |R_n| < \frac{1}{q_n q_{n+1}}. \tag{3.16}$$

Dividiendo (3.10) entre $q_{w_n}q_{w_n+u_n+v_n-1}$ se obtiene

$$0 = x_1(Q_n - 1) + x_2\left(Q_n\frac{p_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}} - \frac{p_{w_n+u_n+v_n-1}}{q_{w_n+u_n+v_n-1}}\right) + x_3\left(Q_n\frac{p_{w_n-1}}{q_{w_n-1}} - \frac{p_{w_n}}{q_{w_n}}\right) +$$
$$+ x_4\left(Q_n\frac{p_{w_n-1}}{q_{w_n-1}}\frac{p_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}} - \frac{p_{w_n}}{q_{w_n}}\frac{p_{w_n+u_n+v_n-1}}{q_{w_n+u_n+v_n-1}}\right),$$

que conlleva, sumando y restando α varias veces y reacomodando,

$$(Q_n - 1)(x_1 + (x_2 + x_3)\alpha + x_2\alpha^2) = x_2(Q_nR_{w_n+u_n+v_n} - R_{w_n+w_n+v_n-1}) + x_3(Q_nR_{w_n-1} - R_{w_n})$$
$$+ \alpha x_4(Q_nR_{w_n-1} + Q_nR_{w_n+u_n+v_n} - R_{w_n} - R_{w_n+u_n+v_n-1}) \tag{3.17}$$

IV.2.3. Más adelante se verá que de (3.17) se deduce:

$$x_1 + (x_2 + x_3)\alpha + x_4\alpha^2 = 0. \tag{3.18}$$

Como $\mathbf{x} \neq \mathbf{0}$ y $[\mathbb{Q}(\alpha) : \mathbb{Q}] \geq 3$, $x_4 = x_1 = 0$ y $x_2 = -x_3 \neq 0$; entonces, (3.10) se reduce a

$$\begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} = \begin{vmatrix} p_{w_n-1} & q_{w_n+u_n+v_n-1} \\ p_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}.$$

Además, los polinomios P_n se convierten en

$$P_n(X) = \begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix} X^2 - 2\begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} X + \begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1} \\ p_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}.$$

Sea $\mathscr{L}^{(3)} : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$\mathscr{L}^{(3)}(Y_1, Y_2, Y_3) := \begin{pmatrix} \mathscr{L}_1^{(3)}(Y_1, Y_2, Y_3) \\ \mathscr{L}_2^{(3)}(Y_1, Y_2, Y_3) \\ \mathscr{L}_3^{(3)}(Y_1, Y_2, Y_3) \end{pmatrix} := \begin{pmatrix} \alpha^2 & -2\alpha & 1 \\ \alpha & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}.$$

Para cada $n \in \mathcal{N}_4$ se considera la terna

$$\mathbf{v}_n := \left(\begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix}, \begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}, \begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1} \\ p_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix}\right).$$

La nueva expresión de P_n, (3.7) y (3.5) dan para $n \in \mathcal{N}_4$ grande y $\varepsilon > 0$ de (3.9)

$$\prod_{j=1}^{3} \left| \mathscr{L}_j^{(3)}(\mathbf{v}_n) \right| = |P_n(\alpha)(\alpha v_{n,1} - v_{n,2})v_{n,1}| \ll \frac{q_{w_n} q_{w_n+u_n+v_m}}{q_{w_n+2u_n+v_n}^2} \ll \frac{1}{(q_{w_n} q_{w_n+u_n+v_n})^{\varepsilon}}.$$

IV.2.4. Por el Teorema del Subespacio de Schmidt a la expresión previa, los puntos $(\mathbf{v}_n)_{n\in\mathcal{N}_4}$ están en una cantidad finita de subespacios propios de $\mathbb{Q}^3$. En consecuencia, hay uno, T_3, con una infinidad de vectores $\mathbf{v}_n$. Nombrando $\mathcal{N}_5 = \{n \in \mathcal{N}_4 : \mathbf{v}_n \in T_3\}$ y $\mathbf{0} \neq (t_1, t_2, t_3) \in \mathbb{Z}^3 \cap T_3^{\perp}$ se tiene

$$t_1 \begin{vmatrix} q_{w_n-1} & q_{w_n+u_n+v_n-1} \\ q_{w_n} & q_{w_n+u_n+v_n} \end{vmatrix} + t_2 \begin{vmatrix} q_{w_n-1} & p_{w_n+u_n+v_n-1} \\ q_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} + t_3 \begin{vmatrix} p_{w_n-1} & p_{w_n+u_n+v_n-1} \\ p_{w_n} & p_{w_n+u_n+v_n} \end{vmatrix} = 0.$$

Dividiendo entre $q_{w_n} q_{w_n+u_n+v_n-1}$ y con manipulaciones similares a las hechas para llegar a la ecuación anterior a (3.17),

$$t_1(Q_n - 1) + t_2 \left(Q_n \frac{p_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}} - \frac{p_{w_n+u_n+v_n-1}}{q_{w_n+u_n+v_n-1}} \right) + t_3 \left(Q_n \frac{p_{w_n-1}}{q_{w_n-1}} \frac{p_{w_n+u_n+v_n}}{q_{w_n+u_n+v_n}} - \frac{p_{w_n}}{q_{w_n}} \frac{p_{w_n+u_n+v_n-1}}{q_{w_n+u_n+v_n-1}} \right) = 0.$$

Un argumento análogo al empleado para concluir (3.18) lleva a

$$t_3\alpha^2 + t_2\alpha + t_1 = 0.$$

Como $(t_1, t_2, t_3) \neq 0$, α no puede ser un número algebraico de grado al menos tres, una contradicción.

IV.1.5. La única deuda que no permite concluir el resultado es (3.18).

Caso i. Para una infinidad de $n \in \mathcal{N}_4$ vale $Q_n \geq 2$ o $Q_n \leq 1/2$. Si fuese la primera, $Q_n \geq 2$, se tendría $|Q_n/(Q_n - 1)| \leq 2$ y $(Q_n - 1) \geq 1$ y dividiendo (3.17) entre $(Q_n - 1)$,

$$\begin{aligned} |x_1 + (x_2 + x_3)\alpha + x_4\alpha^2| \leq\ & |x_2| \left(2|R_{w_n+u_n+v_n}| + |R_{w_n+u_n+v_n-1}|\right) + \quad + |x_3| \left(2|R_{w_n-1}| + |R_{w_n}|\right) + \\ & + |x_3| \left(2|R_{w_n-1}| + |R_{w_n+u_n+v_n-1}|\right) + \\ & + |x_4| \left(2|R_{w_n-1}||R_{w_n+u_n+v_n}| + |R_{w_n-1}||R_{w_n+u_n+v_n}|\right) + \\ & + |\alpha x_4| \left(2|R_{w_n-1}| + 2|R_{w_n+u_n+v_n}| + |R_{w_n}| + |R_{w_n+u_n+v_n-1}|\right) \\ \ll\ & |R_{w_n-1}| \ll \frac{1}{q_{w_n-1} q_{w_n}}. \end{aligned}$$

Como el lado derecho tiende a cero cuando $n \in \mathcal{N}_4$ se va a infinito, el lado izquierdo es cero. Por otra parte, si $Q_n \leq 1/2$ para una infinidad de $n \in \mathcal{N}_4$ el procedimiento es muy parecido observando que $|Q_n/(1 - Q_n)| \leq 1$.

Caso ii. Para toda $n \in \mathcal{N}_4$ grande vale $1/2 \leq Q_n \leq 2$, entonces

$$|(Q_n - 1)(x_1 + (x_2 + x_3)\alpha + x_4\alpha^2| \ll |R_{w_n-1}| \leq \frac{1}{q_{w_n-1} q_{w_n}}.$$

Si $x_1 + (x_2 + x_3)\alpha + x_4\alpha^2 \neq 0$, entonces

$$|Q_n - 1| \ll \frac{1}{q_{w_n-1} q_{w_n}}. \tag{3.19}$$

Por definición, Q_n puede expresarse como un cociente de irracionales cuyo comportamiento es conocido (cf. Apéndice A, Teorema 6):

$$Q_n = \frac{q_{w_n-1} q_{w_n+u_n+v_n}}{q_{w_n} q_{w_n+u_n+v_n-1}} = \frac{[a_{w_n+u_n+v_n}; a_{w_n+u_n+v_n-1}, \ldots, a_1]}{[a_{w_n}; a_{w_n-1}, \ldots, a_1]}.$$

La suposición $a_{w_n} \neq a_{w_n+u_n+v_n}$ da pie a dos subcasos: $a_{w_n} - 1 \leq a_{w_n+u_n+v_n}$ y $a_{w_n} - 1 \geq a_{w_n+u_n+v_n}$. En el primero, tras algunos cálculos elementales, se llega a

$$Q_n \geq \frac{a_{w_n+u_n+v_n}}{a_{w_n} + \dfrac{1}{1 + \dfrac{1}{a_{w_n-2}+1}}} \geq \frac{a_{w_n}+1}{a_{w_n} + \dfrac{a_{w_n-2}+1}{a_{w_n-2}+2}} \geq 1 + \frac{1}{(a_{w_n}+1)(a_{w_n-2}+2)} \geq 1 + \frac{1}{4a_{w_n-2}a_{w_n}}.$$

Mientras que en el segundo,

$$\frac{1}{Q_n} \geq \frac{a_{w_n} + \dfrac{1}{a_{w_n-1}+1}}{a_{w_n+u_n+v_n}+1} \geq 1 + \frac{1}{(a_{w_{n-1}}+1)(a_{w_n+u_n+v_n}+1)} \geq 1 + \frac{1}{(a_{w_n-1}+1)a_{w_n}} \geq 1 + \frac{1}{2a_{w_n-1}a_{w_n}}.$$

En ambos subcasos, recordando que $Q_n \in [1/2, 2]$, se concluye

$$|Q_n - 1| \gg \frac{1}{a_{w_n} \operatorname{máx}\{a_{w_n-2}, a_{w_n-1}\}} \gg \frac{1}{a_{w_n} q_{w_n-1}}.$$

Juntando la expresión previa con (3.19) se concluye

$$\frac{1}{q_{w_n-1}q_{w_n}} \gg \frac{1}{a_{w_n}q_{w_n-1}} \implies a_{w_n} \gg q_{w_n} \geq a_{w_n}q_{w_n-1} \implies 1 \gg q_{w_n-1}.$$

La última desigualdad significa que los números q_{w_n-1} son acotados que, por la naturaleza estrictamente creciente de $(q_j)_{j=0}^{\infty}$, se contrapone a la infinidad de $\mathcal{N}_4$. □

3.3. Construcción de números trascendentes

Un número irracional, θ, es **mal aproximable** si satisface $\liminf n[\![n\theta]\!] > 0$, donde $[\![\cdot]\!]$ es la distancia la entero más cercano y al conjunto de los reales mal aproximables se le denota **Bad**. Un conocido resultado asegura que $\theta \in$ **Bad** si y sólo si los elementos de la expansión en fracción continuada están acotados (cf. [Bu], p. 11, Teorema 1.9 o [Sch], p. 22, Teorema 5F). En particular, **Bad** es no numerable, más aún, si a, b son dos enteros positivos, el conjunto de los irracionales cuyos elementos toman valores en $\{a, b\}$ es no numerable. Hay, entonces, una infinidad de reales trascendentes cuyos elementos en la expansión como fracción continuada asumen únicamente dos valores. Martine Queffélec dio en [qu] una forma de construir algunos.

Teorema 3.3 (Queffélec)**.** *Sean a, b dos enteros positivos distintos. Entonces, el real α cuya sucesión de cocientes parciales es la sucesión de Thue-Morse sobre el alfabeto $\{a, b\}$ (cf. Definición 3.8) es trascendente.*

Adamczewski y Bugeaud obtuvieron en [ab], a partir de la estructura de la sucesión de elementos, el resultado de Queffélec's como un corolario del Teorema 3.4.

Definición 3.7. *Sea $W \in \mathcal{A}^+$. Si $W \in \mathcal{A}^*$, $W = w_0w_1 \dots w_{n-1}$, se dice que W es un **palíndromo** si $w_j = w_{n-1-j}$ para toda $j \in [0..n-1]$. Se dice que W es **puramente palindrómica** si existe una sucesión de palíndromos $(W_n)_{n=1}^{\infty}$ tal que W_n es prefijo de W y $(|W_n|)_{n=1}^{\infty}$ es estrictamente creciente. Se dice que W es **tarde o temprano palindrómica** si existen $V \in \mathcal{A}^*$ y $W' \in \mathcal{A}^{\omega}$ tal que $W = VW'$ y W' es puramente palindrómica.*

Teorema 3.4 (Adamczewski-Bugeaud)**.** *Sea* $\mathbf{a} = (a_n)_{n=1}^{\infty}$ *una sucesión de enteros positivos. Si la palabra* $\mathbf{a}$ *no es tarde o temprano palindrómica, entonces* $[0; a_1, a_2, a_3, \ldots]$ *es trascendente.*

Demostración. Ver [ab], Teorema 1. □

El enunciado original del Teorema 3.4 contempla únicamente el caso en el que $\mathbf{a}$ es puramente periódica. La redación aquí presentada es una consecuencia de que, si $\alpha := [0; a_1, a_2, \ldots]$ y $\alpha_n := [a_n; a_{n+1}, a_{n+2}, \ldots]$, entonces $\mathbb{Q}(\alpha) = \mathbb{Q}(\alpha_n)$ para toda $n \in \mathbb{N}$ (cf. Apéndice A, Teorema 6).

En esta última sección se generaliza la sucesión de Thue-Morse a las sucesiones TM_m, las cuales toman valores en un alfabeto de m letras y satisfacen ciertas condiciones. Se muestra que ellas dan pie a números trascendentes como en el Teorema 3.3; sin embargo, no será posible apelar al Teorema 3.4, pues estas secuecnias no serán palindrómicas salvo cuando $m = 2$. Explícitamente, se probará el siguiente resultado.

Teorema 3.5. *Sean* $a_1, a_2, \ldots, a_m$ *números positivos distintos por pares. Entonces, el número real* α *cuya sucesión de cocientes parciales es la sucesión* TM_m *sobre el alfabeto* $\{a_1, \ldots, a_m\}$ *es trascendente.*

3.3.1. Sucesiones TM_m

Hay varias maneras de establecer y generalizar la sucesión de Thue-Morse (TM_2). Se definirá TM_2 en términos de la expansión binaria de un número natural y su generalización, en términos de la expansión m-aria finita de un número natural. Siguiendo la notación de [AS]: para enteros no tegativos n, m, $m \geq 2$, $(n)_m$ es la palabra formada por los coeficientes de la expansión m-aria finita de n; esto es,

$$n = \sum_{i}^{r} c_i m^i \;\Rightarrow\quad (n)_m := c_0 c_1 \ldots c_r.$$

Para cualquier palabra finita sobre el alfabeto $\{0, 1, \ldots, m-1\}$, dígase $\mathbf{w} = w_0 w_1 \ldots w_r$, se utilizará $[\mathbf{w}]_m = \sum_{i=0}^{r} w_i m^i$ y $\mathbf{w}_{[j_1:j_2]} = w_{j_1} w_{j_1+1} \ldots w_{j_2-1}$. También, se denota $\Sigma_m := \{1, 2, \ldots, m\}$.

Definición 3.8. *Sea* $m \geq 2$ *un entero. La sucesión* $(t_n(m))_{n=0}^{\infty}$ *se construye en dos pasos. Primero, para cualquier entero no negativo,* n*, se considera su expansión* m*-aria finita*

$$n = \sum_{j=0}^{k_n} c_j(n) m^j.$$

Segundo, se define

$$t_n(m) := \sum_{j=0}^{k_n} c_j(n) \mod m.$$

Se le conoce a $(t_n(m))_{n=0}^{\infty}$ *como la* ***sucesión*** $\boldsymbol{TM_m}$*. Sean* $\Sigma \subset \mathbb{N}$, $|\Sigma| = m$ *y* $f : \Sigma_m \to \Sigma$ *una biyección, entonces* $(f(t_n(m)))_{n=0}^{\infty}$ *es la* ***sucesión*** $\boldsymbol{TM_m}$ ***sobre*** $\boldsymbol{\Sigma}$*.*

Cuando $m = 2$ se obtiene la sucesión clásica de Thue - Morse (cf. [qu]). Para aligerar la notación, cuando m sea un entero positivo fijo, se escribirá $(t_n)_{n=0}^{\infty}$ en lugar de $(t_n(m))_{n=0}^{\infty}$. La definición conduce a algunas propiedades inmediatas.

Proposición 3.6. *Fíjese un entero $m \geq 2$ y sea $(t_n)_{n=0}^{\infty}$ la sucesión TM_m asociada. Entonces,*

1. $t_n = t_{nm}$ *para toda* $n \in \mathbb{N}$,

2. $t_{n+1} - t_n \not\equiv 1 \pmod{m}$ *implica* $n \equiv m-1 \pmod{m}$,

3. *Para toda* $r \in \Sigma_m$ *y cualquier* $n \in \mathbb{N}_0$, $t_{nm+r} \equiv t_{nm} + r \pmod{m}$.

Demostración. **1.** Si $n = [c_0\, c_1\, \ldots\, c_r]_m$, se tiene

$$n = \sum_{i=0}^{r} c_i m^i \;\Rightarrow nm = \sum_{i=0}^{r} c_i m^{i+1} \;\Rightarrow\; nm = [0,\, c_0,\, c_1,\, \ldots,\, c_r]_m.$$

2. Supóngase que $n \not\equiv m-1 \pmod{m}$, entonces $n = [c_0\, c_1\, \ldots\, c_r]_m$ con $c_0 \in \{0, 1, \ldots, m-2\}$, por lo que

$$n+1 = \sum_{i=0}^{r} c_i m^i + 1 = (c_0+1) + c_1 m + c_2 m^2 + \ldots + c_r m^r \;\Rightarrow\; n+1 = [c_0+1,\, c_1,\, \ldots,\, c_r]_m.$$

Luego,

$$t_{n+1} \equiv (c_0+1) + c_1 + \ldots + c_r \equiv 1 + \sum_{i=0}^{r} c_i \equiv 1 + t_n \pmod{m}.$$

3. La tercera parte sigue de la contrapuesta de la segunda. De $nm \equiv 0 \pmod{m}$ se tiene que $t_{nm+1} - t_{nm} \equiv 1 \pmod{m}$, …, $t_{nm+(m-1)} - t_{nm+(m-2)} \equiv 1 \pmod{m}$. Sumando adecuadamente estas ecuaciones se llega a la conclusión deseada. □

Una consecuencia inmediata de la Proposición 3.6 es que $t_j = t_{j+1} = t_{j+2}$ es imposible. Si no lo fuera, entonces $j \equiv -1 \pmod{m}$ y $j+1 \equiv -1 \mod m$, que implicaría $1 \equiv 0 \pmod{m}$.

Sea $f : \Sigma_m \to \mathbb{N}$ inyectiva. Si $(x_n)_{n=0}^{\infty}$ es una sucesión en Σ_m, obviamente $(f(x_n))_{n=0}^{\infty}$ es tarde o temprano peroódica o tarde o temprano palindrómica si y sólo si la sucesión original es tarde o temprano periódica o tarde o temprano palindrómica, respectivamente. Los números reales considerados en el Teorema 3.5 tienen precisamente la forma $\alpha = [0; f(x_0), f(x_1), f(x_2), \ldots]$ con $x_n = t_n(m)$ y f inyectiva. Será muy importante verificar que TM_m no es a la larga periódica. Si este fuera el caso, el Teorema de Lagrange garantizaría que todos los números generados de este modo sean irracionales cuadráticos.

Teorema 3.7. *Para cualquier entero $m \geq 2$ la sucesión TM_m no es tarde o temprano periódica.*

Demostración. Sean $m \in \mathbb{N}$, $m \geq 2$, y $(t_n)_{n=0}^{\infty}$ la sucesión TM_m. Se mostrará que no existen enteros positivos a, b tales que $t_{a+n} = t_{a+n+b}$ para cualquier $n \in \mathbb{N}$. Supóngase que no es así y tómese a b mínimo.

Caso $b < m$. Supóngase que $t_{a+n} = t_{a+b+n}$ para todo $n \in \mathbb{N}_0$ y alguna $a \in \mathbb{N}_0$. En particular, si $a + a' \equiv 0 \mod m$, entonces la Proposición 3.6.3 implica

$$t_{a+a'} \equiv t_{a+a'+b} \equiv t_{a+a'} + b \mod m$$

que conlleva $b = 0$, una contradicción.

Caso $b \equiv 0 \mod m$ *y* $b \geq m$. La hipótesis implica que para algún entero positivo b' se verifica $b = mb'$. Entonces, por la Proposición 3.6.1., para cualquier $n \in \mathbb{N}$ se tiene

$$t_{a+n} = t_{m(a+n)} = t_{m(a+n)+b} = t_{m(a+n+b')} = t_{a+n+b'},$$

que contradice la minimalidad de b.

Caso $b \not\equiv 0 \mod m$ *y* $b > m$. El intervalo cerrado $[a, a+2b]$ contiene al menos $2m+3$ enteros consecutivos, por lo que contiene 2 múltiplos de m estrictamente mayores que a, llámense $j = mq$ y $j+m = m(q+1)$. Supóngase que $t_{j+m} - t_{j+m-1} \equiv 1 \mod m$ y $t_j - t_{j-1} \equiv 1 \mod m$, entonces $t_{j-m} = t_j = t_{j+m}$: de hecho, por la Proposición 3.6,

$$t_{(j-m)+r} = t_{j-m} + r,\ t_{j+r} = t_j + r,\ t_{j+m+r} = t_{j+m} + r \quad \forall r \in \Sigma_m,$$

por lo que

$$\begin{aligned} t_j &\equiv 1 + t_{j-1} \equiv 1 + t_{(j-m)+(m-1)} \equiv 1 + t_{j-m} + m - 1 \equiv t_{j-m} \mod m, \\ t_{j+m} &\equiv 1 + t_{j+m-1} \equiv 1 + t_{j+(m-1)} \equiv 1 + t_j + m - 1 \equiv t_j \mod m. \end{aligned}$$

Luego, $t_{q-1} = t_q = t_{q+1}$, que es imposible por la Proposición 3.6.2. En consecuencia, al menos una de las expresiones $t_{j+m} - t_{j+m-1} \not\equiv 1 \mod m$, $t_j - t_{j-1} \not\equiv 1 \mod m$ se satisface. Sin perder generalidad, se supone que $t_j - t_{j-1} \not\equiv 1 \mod m$ es cierto. Entonces, $t_{j+b} - t_{j-1+b} \equiv t_j - t_{j-1} \not\equiv 1 \mod m$ que da $j + b - 1 \equiv -1 \mod m$ and $j - 1 \equiv -1 \mod m$, implicando que $b \equiv 0 \mod m$, contrario a la hipótesis. □

Al dotar a Σ^* con la operación binaria de la concatenación se obtiene un monoide cuyo neutro es ϵ. Los morfismos constituyen un tipo importante de función de Σ^* en Σ^*, éstos son simplemente homomorfismos de monoides.

Definición 3.9. *Un* ***morfismo*** *es un homomorfismo de monoides de* Σ^* *en sí mismo, i.e.* $\varphi : \Sigma^* \to \Sigma^*$ *cumple con*

$$\mathbf{w} = w_0 w_1 \dots w_k \in \Sigma^* \ \Rightarrow\ \varphi(\mathbf{w}) = \varphi(w_0)\varphi(w_1)\dots\varphi(w_k).$$

Cualquier morfismo φ *puede extenderse a una función* $\overline{\varphi} : \Sigma^\infty \to \Sigma^\infty$ *mediante*

$$\overline{\varphi}(\mathbf{w}) = \begin{cases} \varphi(w_0)\varphi(w_1)\varphi(w_2)\dots & si\ \ \mathbf{w} = w_0 w_1 w_2 \dots \in \Sigma^\infty \setminus \Sigma^*, \\ \varphi(\mathbf{w}) & si\ \ \mathbf{w} \in \Sigma^*. \end{cases}$$

En este caso, se dice que $\overline{\varphi}$ *es* ***generada*** *por* φ.

Evidentemente, cualquier morfismo $\varphi : \Sigma^* \to \Sigma^*$, y en consecuencia el morfismo generado por él, está completamente determinado por $\varphi(j)$ con j variando en Σ. Para evitar recargar la notación, no se distinguirá entre φ y $\overline{\varphi}$. También, se define $\varphi^1 = \varphi$ y para toda $n \in \mathbb{N}$, $\varphi^{n+1} := \varphi^n \circ \varphi$. Algunos morfismos ayudarán conseguir nuevas palabras.

Definición 3.10. *Un morfismo* $\varphi : \Sigma^* \to \Sigma^*$ *es* ***prolongable en*** $j \in \Sigma$ *si* $\varphi(j) = j\mathbf{x}$ *para algún* $\mathbf{x} \in \Sigma^*$. *Si* φ *es prolongable en* j, *el* ***punto fijo de*** $\overline{\varphi}$ ***anclado en*** j *es la palabra* $j\mathbf{x}\varphi(\mathbf{x})\varphi^2(\mathbf{x})\varphi^3(\mathbf{x})\dots$.

Definición 3.11. *Sean* $\varphi : \Sigma^* \to \Sigma^*$ *un morfismo y* $k \in \mathbb{N}$. *Se dice que* φ *es* ***k-uniforme*** *si* $|\varphi(j)| = k$ *para cualquier* $j \in \Sigma$.

La corta demostración de Adamczewski y Bugeaud dada en [ab] del resultado de Quéffelec se basa en que la sucesión clásica de Thue-Morse es palindrómica. Desgraciadamente, esta característica deja de cumplirse para las demás sucesiones TM_m.

Teorema 3.8. *Sean $m \geq 2$ un entero y $(t_n)_{n=0}^{\infty}$ la sucesión TM_m asociada. Si $m = 2$, entonces $(t_n)_{n=0}^{\infty}$ es palindrómica. Si $m > 2$, entonces $(t_n)_{n=0}^{\infty}$ no es tarde o temprano palindrómica.*

Demostración. Se exhibe la prueba del caso $m = 2$ encontrada en [AS] porque motivará las ideas que se usarán en adelante. La Proposición 3.6 aplicada a la sucesión de Thue-Morse ($m = 2$) implica que $(t_n)_{n=0}^{\infty}$ es palindrómica, donde $(n_j)_{j=1}^{\infty}$ está dada por $n_j = 4^{j-1}$. Para verlo, sea $\varphi : \{0,1\}^{\infty} \to \{0,1\}^{\infty}$ el morfismo generado por la función $\varphi(0) = 01$, $\varphi(1) = 10$. Por definición, $\varphi^2(0) = 0110$, $\varphi^2(1) = 1001$, por lo que el conjunto de palabras palindrómicas sobre el alfabeto $\{0,1\}$ es invariante bajo φ^2. De hecho, si $\mathbf{w} \in \Sigma^*$ un palíndromo, $\mathbf{w} = w_0 w_1 \dots w_{k-1}$, entonces, si $\overline{w}_j := 1 - w_j$, se tiene

$$\varphi^2(\mathbf{w}) = \varphi^2(w_0)\varphi^2(w_1)\dots\varphi^2(w_{k-1}) = w_0\,\overline{w}_0\,\overline{w}_0 w_0\, w_1\,\overline{w}_1\,\overline{w}_1\, w_1 \dots w_{k-1}\,\overline{w}_{k-1}\,\overline{w}_{k-1}\, w_{k-1}.$$

Nótese que $|\varphi(\mathbf{w})| = 4k$. Sean $j \in [0..4k-1]$ y $q, r \in \mathbb{N}_0$ tales que $j = 4q + r$ y $0 \leq r < 4$. Claramente,

$$\varphi(\mathbf{w})_j = \varphi(\mathbf{w})_{4q+r} = \begin{cases} w_q \text{ si } r \in \{0,3\}, \\ \overline{w}_q \text{ si } r \in \{1,2\}. \end{cases}$$

En consecuencia, como $r \in \{0,3\}$ si y sólo si $3-r \in \{0,3\}$ y $r \in \{1,2\}$ si y sólo si $3-r \in \{1,2\}$,

$$\varphi(\mathbf{w})_{4k-1-j} = \varphi(\mathbf{w})_{4k-1-4q-r} = \varphi(\mathbf{w})_{4(k-1-q)+(3-r)} = \begin{cases} w_{k-1-q} \text{ si } r \in \{0,3\}, \\ \overline{w}_{k-1-q} \text{ si } r \in \{1,2\}. \end{cases}$$

Entonces, si $\mathbf{w}$ es un palíndromo, para cualquier $l \in [0..k-1]$ la igualdad $w_l = w_{k-1-l}$ se cumple implicando $\varphi(\mathbf{w})_j = \varphi(\mathbf{w})_{4k-1-j}$ y, consecuentemente, que $\varphi(\mathbf{w})$ es un palíndromo. Por definición, $\varphi^2(0) = 0110$ es un palíndromo y, por inducción, $\varphi^{2n}(0)$ es un palíndromo para cualquier $n \in \mathbb{N}$. Obsérvese que φ es prolongable en 0. Además, se tiene $\varphi(0) = 01$, $\varphi^2(0) = 0110 = 01\varphi(1)$, si $\varphi^{2n}(0) = 01\varphi(1)\dots\varphi^{2n-1}(0)$, entonces

$$\begin{aligned}\varphi^{2(n+1)}(0) &= \varphi^2(\varphi^n(0)) = \varphi^2(01\varphi(1)\dots\varphi^{2n-1}(0)) = \varphi^2(0)\varphi^2(1)\dots\varphi^{2n+1}(0) \\ &= 01\varphi(1)\dots\varphi^{2n+1}(0).\end{aligned}$$

Esto significa que $\varphi^{2n}(0)$ es un segmento inicial de un punto fijo de φ anclado en 0, $\mathbf{v}$. En particular, $\mathbf{v}$ es palindrómica con los segmentos palindrómicos iniciales dados por $\varphi^{2n}(0)$. Por el Teorema 3.9 de la siguiente sección (cuya prueba es independiente de ésta), $\mathbf{v}$ es precisamente TM_2. Luego, TM_2 es una sucesión palindrómica.

Ahora se considera el caso $m \geq 3$. En $(t_n)_{n=0}^{\infty}$ la palabra 011 aparece una infinidad de veces, porque

$$j := \big[m-2, \underbrace{m-1,\dots,m-1}_{\text{m-2 veces}}\big]_m, j+1 = \big[\underbrace{m-1,\dots,m-1,m-1}_{\text{m-1 veces}}\big]_m, j+2 = \big[\underbrace{0,\dots,0}_{\text{m-1 veces}},1\big]_m$$

da $t_j = 0$, $t_{j+1} = 1$, $t_{j+2} = 1$. Además, $t_{j'} = 0$, $t_{j'+1} = 1$, $t_{j'+2} = 1$ cuando j'_m está definida por

$$j' := \big[m-2, \underbrace{m-1,\dots,m-1}_{\text{m-2 veces}}, 0, m-2, 2\big]_m = j + (m-2)m^m + 2m^{m+1}.$$

En general, $j(k) := j + (m-2)m^k + 2m^{k+1}$ para $k \geq m$ da lugar a una ocurrencia de 011; esto es, $t_{j(k)} = 0$, $t_{j(k)+1} = 1$, $t_{j(k)+2} = 1$ donde

$$j(k) := \big[m-2, \underbrace{m-1, \ldots, m-1}_{\text{m-2 veces}}, \underbrace{0, \ldots, 0}_{\text{k+1-m veces}}, m-2, 2\big]_m = j + (m-2)m^k + 2m^{k+1}.$$

Sin embargo, la configuración 110 es imposible. Supóngase que $t_k = 1$, $t_{k+1} = 1$, $t_{k+2} = 0$ para algún $k \in \mathbb{N}_0$. Como $m \geq 3$, $-1 \not\equiv 1 \pmod m$, por lo que $t_{k+2} - t_{k+1} \not\equiv 1 \pmod m$. También se tiene $t_{k+1} - t_k \not\equiv 1 \pmod m$. Por la Proposición 3.6, $k+1 \equiv m-1 \pmod m$ y $k \equiv m-1 \pmod m$. Las últimas congruencias conducen a $1 \equiv 0 \pmod m$, que es absurdo. Por lo tanto, las sucesiones TM_m no son tarde o temprano palindrómicas. □

3.3.2. Una definición alternativa de las sucesiones TM_m

La sucesión de Thue-Morse clásica (TM_2) puede definirse como el punto fijo anclado en 0 de un morfismo (cf. [AS], Corolario 1.7.7, p.23). Con este espíritu, para cualquier $m \in \mathbb{N}$ fijo, $m \geq 2$, se considera el morfismo $\varphi : \Sigma_m^* \to \Sigma_m^*$ dado por

$$\forall j \in \Sigma_m \quad \varphi(j) := j,\ j+1,\ \ldots,\ j+m-1.$$

La suma en la expresión anterior debe entenderse módulo m. Por ejemplo, si $m = 4$ y $j = 2$, entonces $\varphi(j) = 2\,3\,0\,1$ (como se notó antes, φ está completamente determinado por su acción en Σ_m). Por definición, φ es prolongable en 0 y, de hecho, en cualquier elemento de Σ_m. También, φ^k es m^k-uniforme para cualquier $k \in \mathbb{N}$. Esto da otra perspectiva para estudiar las sucesiones TM_m.

Definición 3.12. *Para cada $m \geq 2$ sea φ el morfismo definido como arriba. La **sucesión TM_m** es el punto fijo de φ anclado en 0.*

El siguiente Teorema afirma que en verdad no se le está dando el mismo nombre a dos objetos distintos. En adelante, se supondrá que m ha sido fijado.

Teorema 3.9. *Las Definiciones 3.8 y 3.12 son equivalentes.*

Lema 3.10. *Sean $c_0, \ldots, c_n, c_{n+1} \in \Sigma_m$ y $\mathbf{c} := c_0 \ldots c_{n+1}$. Entonces, se cumple la siguiente igualdad*

$$\varphi^{n+1}(c_{n+1})_{[\mathbf{c}]_m} = \sum_{i=0}^{n+1} c_i \pmod m. \tag{3.20}$$

Demostración. En esta prueba las llaves carecen de significado, son sólo para facilitar la lectura. Considérese, primero, el caso $n = 0$. Sean c_0, $c_1 \in \Sigma_m$, entonces

$$\varphi(c_1)_{c_0} = \{c_1,\ c_1+1,\ \ldots,\ c_1+m-1\}_{c_0} = c_1 + c_0.$$

Ahora considérese el caso $n = 1$. De que φ sea un morfismo m-uniforme y de la definición de φ se sigue que

$$\varphi^2(c_2)_{c_1 m + c_0} = \{\varphi(c_2)\varphi(c_2+1)\ldots\varphi(c_2+m-1)\}_{c_1 m + c_0} = \varphi(c_2 + c_1)_{c_0} = c_2 + c_1 + c_0.$$

Supóngase que (3.20) se cumple para $n-1$. Tomando $c_0, c_1, \ldots, c_{n+1} \in \Sigma_m$ y $\mathbf{c} := c_0 c_1 \ldots c_n$,

$$\begin{aligned}
\varphi^{n+1}(c_{n+1})_{[\mathbf{c}]_m} &= \{\varphi^n(c_{n+1})\varphi^n(c_{n+1}+1)\cdots\varphi^n(c_{n+1}+m-1)\}_{[\mathbf{c}]_m} \\
&= \varphi^n(c_{n+1} + c_n)_{\sum_{i=0}^{n-1} c_i m^i} && (\varphi^n \text{ es } m^n\text{-uniforme}) \\
&= \sum_{i=0}^{n+1} c_i. && (\text{Hipótesis de Inducción})
\end{aligned}$$

El lema sigue del Principio de Inducción Matemática. □

Demostración del Teorema 3.9. Sea $\mathbf{v} \in \Sigma_m^\infty$ un punto fijo de φ anclado en 0, entonces

$$\begin{aligned}\mathbf{v} &= 0, 1, \ldots, m-1, \varphi(1, \ldots, m-1), \varphi^2(1, \ldots, m-1), \ldots \\ &= \varphi(0)\,\varphi(1)\ldots\varphi(m-1)\varphi^2(1, \ldots, m-1)\varphi^3(1, \ldots, m-1)\ldots \\ &= \varphi^2(0)\varphi^2(1)\ldots\varphi^2(m-1)\varphi^3(1, \ldots, m-1)\varphi^4(1, \ldots, m-1)\ldots \\ &\vdots \\ &= \varphi^n(0)\varphi^n(1)\ldots\varphi^n(m-1)\varphi^{n+1}(1, \ldots, m-1)\varphi^{n+2}(1, \ldots, m-1)\ldots.\end{aligned}$$

Sea j un entero no negativo, $j = \sum_{i=0}^{n} c_i m^i$, y defínase $\mathbf{c} := (j)_m$. Tomando un segmento inicial suficientemente largo y aplicando el Lema 3.10 se consigue

$$v_j = \varphi^{n+1}(0)_j = \varphi^{n+1}(0)_{[\mathbf{c}]_m} = \sum_{i=0}^{n} c_i \pmod m = t_j.$$

□

3.3.3. Prueba del Teorema 3.5.

La definición 3.12 hace que la teoría de los morfismos de palabras sea accesible. En particular, algunos aspectos combinatorios extremadamente útiles serán empleados.

Lema 3.11. *Sean n, m enteros positivos con $m \geq 2$ y $(t_j)_{j=1}^\infty$ la sucesión TM_m. Si $\mathbf{t} = t_0 t_1 t_2 \ldots$, entonces $p_{\mathbf{t}}(n) \leq m^3 n$.*

Demostración. Supóngase que n es un entero positivo fijo y que $r \in \mathbb{N}$ es tal que $m^{r-1} \leq n < m^r$. Llámese $\mathbf{t}[k]$ al conjunto de subpalabras de $\mathbf{t}$ de longitud k. Sea $f : \Sigma_{m^r} \times \mathbf{t}[2] \to \mathbf{t}[n]$ la función dada por

$$f(s, \mathbf{v}) := \varphi^r(\mathbf{v})_{[s:s+n]}. \tag{3.21}$$

Como $\mathbf{t}$ es un punto fijo de φ y φ es un morfismo m-uniforme, $\mathbf{t}$ es un punto fijo de φ^p y φ^p es un morfismo m^p-uniforme para cualquier $p \in \mathbb{N}$. En consecuencia,

$$\forall\, p \in \mathbb{N} \quad \mathbf{t} = t_0 t_1 t_2 \ldots = \varphi^p(t_0)\varphi^p(t_1)\varphi^p(t_2)\ldots = \varphi^p(\mathbf{t}).$$

Tómese $\mathbf{v} = t_j t_{j+1}$ con $j \in \mathbb{N}$ fija. La ecuación de arriba y las propiedades recién enunciadas de φ implican

$$\varphi^r(\mathbf{v}) = \varphi^r(t_j)\varphi^r(t_{j+1}) = t_{jm^r} t_{jm^r+1} \ldots t_{(j+1)m^r-1} t_{(j+1)m^r} t_{(j+1)m^r+1} \ldots t_{(j+2)m^r-1},$$

y se obtiene

$$f(s, \mathbf{v}) = \varphi^r(\mathbf{v})_{[s:s+n]} = t_{jm^r+s} t_{jm^r+s+1} \ldots t_{jm^r+s+n-1}.$$

Sea $\mathbf{x}$ una subpalabra de $\mathbf{t}$ de longitud n, entonces $\mathbf{x} = t_i \ldots t_{i+n-1}$ para alguna i. Por el Algoritmo de Euclides, hay dos enteros no negativos q', s' tales que $i = q'm^r + s'$ y $0 \leq s' < m^r$. Obviamente, $\mathbf{x}$ es una subpalabra de

$$\varphi^r(t_{q'} t_{q'+1}) = t_{q'm^r} \ldots t_{(q'+1)m^r-1} t_{(q'+1)m^r} \ldots t_{(q'+2)m^r-1};$$

de hecho, $\mathbf{x} = f(s', \mathbf{v})$ con $\mathbf{v} = t_{q'} t_{q'+1}$, probando que f es sobre. En consecuencia,

$$p_{\mathbf{t}}(n) = \left|\mathbf{t}[n]\right| \leq \left|\Sigma_{m^r} \times \mathbf{t}[2]\right| = m^r m^2 = m^{r-1} m^3 \leq n m^3$$

que implica

$$\limsup_{n\to+\infty} \frac{p_{\mathbf{t}}(n)}{n} \leq m^3. \tag{3.22}$$

□

El argumento anterior es sólo un caso particular de un teorema más general (cf. [AS], Teorema 10.3.1, p. 304).

Prueba del Teorema 3.5.

Sean $m \geq 3$ un entero, $(t_j)_{j=0}^{\infty}$ la sucesión TM_m, $\Sigma \subset \mathbb{N}$, $\psi : \Sigma_m \to \Sigma$ biyectiva, $\mathbf{w} := \psi(t_0)\psi(t_1)\psi(t_2)\ldots$ y $\alpha = [0; \psi(t_0), \psi(t_1), \psi(t_2), \ldots]$. La conclusión del Lema 3.11 sigue siendo válida–ψ es una biyección– cuando se aplica a $\mathbf{w}$, por lo que (3.22) es válida y, por el Teorema 3.7, $\mathbf{w}$ no es tarde o temprano periódica. Como

$$\limsup_{n\to+\infty} \frac{p_n(\mathbf{w})}{n} < \infty,$$

el Teorema 3.1 implica que α es trascendente.

Apéndice A
Elementos de fracciones continuadas

Fracciones continuadas

Además de una vistosa notación, las fracciones continuadas brindan valiosos teoremas en la teoría de aproximación. Este breve apéndice es para recordar algunas definiciones y propiedades básicas de los objetos mencionados. Varias pruebas son omitidas, muchas son sólo ejercicios elementales de inducción matemática y pueden encontrarse en referencias clásicas como [Kh] o [La02]

Definición 1. *Sean $N \in \mathbb{N}$, $(a_n)_{n=1}^{N}$ una sucesión tal que $a_0 \in \mathbb{Z}$ y $a_j \in \mathbb{N}$ para $j \geq 1$. Una* ***fracción continuada finita simple de orden N*** *es una expresión de la forma*

$$[a_0; a_1, a_2, \ldots, a_N] := a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots + \cfrac{1}{a_N}}}}}.$$

Sean $(a_n)_{n=1}^{\infty}$ es una sucesión de enteros positivos y $a_0 \in \mathbb{Z}$. Una expresión de la forma

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

se llama ***fracción continuada infinita simple****. Tanto en el caso finito como en el infinito, los términos $(a_n)_{n=0}^{\infty}$ se llaman* ***elementos*** *o* ***cocientes parciales****.*

Definición 2. *La* ***representación canónica*** *de la fracción $[a_0]$ es $a_0/1$. Si se han definido la representación canónica de las fracciones de orden $k-1$, entonces si la de $[a_1; a_2, a_3, \ldots, a_k]$ es p'/q', se define la de $[a_0; a_1, a_2, \ldots, a_k]$ como $\frac{a_0 p' + q'}{p'}$.*

Manteniendo la notación, se observa que

$$[a_0; a_1, a_2, \ldots, a_k] = a_0 + \frac{1}{[a_1; a_2, a_3, \ldots, a_k]} = a_0 + \frac{q'}{p'} = \frac{a_0 p' + q'}{p'}.$$

Una de las principales cualidades de las fracciones continuadas es que permiten aproximar rápida y recursivamente a los números reales. Por el resto de esta sección, $(a_n)_{n=0}^{\infty}$ y $(b_n)_{n=0}^{\infty}$ representarán sucesiones de números naturales[2] salvo, posiblemente, a_0 y b_0 que son enteros.

Teorema 3. *1. Existe el límite*

$$[a_0; a_1, a_2, \ldots] := \lim_{n\to\infty} [a_0; a_1, a_2, \ldots, a_n].$$

2. Todo número real admite una representación como fracción continuada.

3. Si $\alpha \in \mathbb{R}$, entonces α es irracional si y sólo si su representación como fracción continuada es infinita. En este caso, la representación es única.

4. Si $\alpha \in \mathbb{Q}$, entonces hay sólo dos representaciones y ambas son finitas; de hecho, éstas serían

$$\alpha = [a_0; a_1, \ldots, a_n] = [a_0; a_1, \ldots, a_{n-1} - 1, 1].$$

Demostración. [La02] p.7, [Kh] p.14. □

Definición 4. *Si $(a_n)_{n=0}^{\infty}$ es una sucesión como antes, la sucesión de* ***convergentes*** *de la fracción $[a_0; a_1, a_2, \ldots]$, $(p_n/q_n)_{n=0}^{\infty}$, es la sucesión cuyo término n-ésimo es la representación canónica de $[a_0; a_1, a_2, \ldots, a_n]$.*

Teorema 5. *Los convergentes de una fracción continuada están dados por $q_0 = 1$, $p_0 = a_0$, $q_1 = a_1$, $p_1 = a_0 a_1 + 1$ y*

$$\forall n \geq 2 \qquad p_{n+1} = a_{n+1} p_n + p_{n-1}, \quad q_{n+1} = a_{n+1} q_n + q_{n-1}.$$

Si se definen $q_{-1} = 0$, $p_{-1} = 1$, la recurrencia es válida desde $n = 1$.

Demostración. [La02] p. 2., Teorema 1; [Kh] p.4, Teorema 1. □

El carácter recursivo de los convergentes permite entender mejor el comportamiento de estas aproximaciones. Por ejemplo, los convergentes forman una sucesión alternante; los errores, una estrictamente decreciente. El descenso de la secuencia de errores no es descontrolado, es posible encontrar una cota inferior.

Teorema 6. *Sean $\alpha \in \mathbb{R}$ y $(p_n/q_n)_{n=1}^{\infty}$ la sucesión de convergentes de α. Entonces se cumplen las siguientes propiedades.*

1. Para toda $k \in \mathbb{N}$ o para aquéllos naturales, k, para los que p_n/q_n esté definido

$$q_k p_{k-1} - p_k q_{k-1} = (-1)^k, \quad q_k p_{k-2} - p_k q_{k-2} = (-1)^{k-1} a_k.$$

2. Para toda $k \in \mathbb{N}$ o para aquéllos naturales, k, para los que p_n/q_n esté definido

$$\alpha = \frac{p_{k-1}[a_k; a_{k+1}, a_{k+2}, \ldots] + p_{k-2}}{q_{k-1}[a_k; a_{k+1}, a_{k+2}, \ldots] + q_{k-2}}.$$

[2]Se llama *naturales* al conjunto de los enteros positivos

3. Para toda $k \in \mathbb{N}$ o para aquéllos naturales, k, para los que p_n/q_n esté definido

$$\frac{q_{k+1}}{q_k} = [a_{k+1}; a_k, \ldots, a_1].$$

Demostración. 1. Ver [La02] p.4, Teorema 2 o [Kh] p. 5, Teorema 2 y p.6, Teorema 3.

2. Ver [La02] p.3, Corolario 1 o [Kh] p.7, Teorema 5.

3. Ver [La02] p.6, Teorema 4 o [Kh] p.7, Teorema 6.

□

Lema 7. *Sean $\alpha \in (0,1)$ y $n \in \mathbb{N}$ tal que el $(n+1)$-ésimo convergente está definido; entonces,*

$$\left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right| < \left|\alpha - \frac{p_n}{q_n}\right|. \tag{1}$$

Demostración. [La02] p.9, Corolario. □

Teorema 8. *Sea $\alpha \in (0,1)$, entonces para cualquier $n \in \mathbb{N}$ se tiene*

$$\frac{1}{q_n(q_n + q_{n+1})} < \left|\alpha - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}}. \tag{2}$$

Demostración. [La02] p.8, Teorema 5; [Kh] p.9, Teorema 9 y p.15, Teorema 13. □

La expansión en fracciones continuadas da un criterio sencillo para decidir si un número es racional o irracional. La estructura de la sucesión de cocientes parciales también permite caracterizar a los irracionales cuadráticos.

Teorema 9 (Lagrange)**.** *Sea $\alpha = [a_0; a_1, a_2, a_3, \ldots]$ un real. Entonces, α es un irracional cuadrático si y sólo si la sucesión $(a_n)_{n=0}^{\infty}$ tarde o temprano es periódica.*

Demostración. Ver [La02] p.54, Teorema 3 o [Kh] p.48, Teorema 28. □

Teorema 10. *Sean $\alpha = [a_0; a_1, a_2, a_3, \ldots]$ y $\beta = [b_0; b_1, b_2, b_3, \ldots]$ dos reales. Si existe $n \in \mathbb{N}$ tal que $a_j = b_j$ para $j \in \{0, 1, \ldots, n\}$, entonces*

$$|\alpha - \beta| \leq \frac{1}{q_{n-1} q_n}$$

en donde $(p_n/q_n)_{n=1}^{\infty}$ son los convergentes de α.

Demostración. Dependiendo de la paridad de n alguna de las dos desigualdades es cierta:

$$\frac{p_{n-1}}{q_{n-1}} < \alpha, \beta < \frac{p_n}{q_n}, \quad \frac{p_n}{q_n} < \alpha, \beta < \frac{p_{n-1}}{q_{n-1}}.$$

Restando los extremos se concluye el resultado. □

Si bien las próximas dos proposiciones no están en las referencias que se han citado, siguen fácilmente de ellas. Son enunciadas en [bu].

Teorema 11. *Sean* $(a_n)_{n=1}^{\infty}$ *una sucesión de naturales,* $\alpha = [0; a_1, a_2, \ldots]$ *y* $(p_n/q_n)_{n=1}^{\infty}$ *los convergentes de* α*. Si los cocientes parciales de* $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ *están acotados por* M*, entonces*

$$\limsup_{n\to\infty} \sqrt[n]{q_n} < +\infty.$$

Demostración. Supóngase que $a_n < M$ para toda $n \in \mathbb{N}$. Por el Teorema 5, $q_0 = 1 \leq (M+1)^0$ y $q_1 = a_1 \leq (M+1)^1$. Suponiendo válido $q_n \leq (M+1)^n$ para $n \in \{1, 2, \ldots, N\}$, el Teorema 5 da

$$q_{N+1} = a_{N+1}q_N + q_{N-1} \leq M(M+1)^N + (M+1)^{N-1} = (M+1)^{N-1}(M(M+1)+1) \leq (M+1)^{N+1}.$$

Así, para cualquier $n \in \mathbb{N}$ se cumple $q_n^{1/n} < M+1$ y el resultado se sigue. □

Teorema 12. *Si* $[0; a_1, a_2, a_3, \ldots]$ *es un irracional y* $(q_n)_{n=0}^{\infty}$ *es la sucesión de denominadores, entonces*

$$\forall h, l \in \mathbb{N} \quad q_l(\sqrt{2})^{h-1} \leq q_{l+h}.$$

Demostración. Como $(q_n)_{n=1}^{\infty}$ es creciente, 2 se traduce en que para cualquier $n \in \mathbb{N}$ se cumple

$$\frac{1}{2q_{l+1}^2} < \left|\alpha - \frac{p_l}{q_l}\right| \leq \frac{1}{q_{l+1}^2} \implies \forall n \in \mathbb{N} \quad q_l \leq q_{l+1}\sqrt{2}.$$

Entonces, fijando a $l \in \mathbb{N}$ y usando repetidamente la desigualdad anterior se concluye el teorema. □

Dado un irracional α, la teoría de fracciones continuadas provee una sucesión de números racionales convergente a α; en términos de aproximación, ésta es la mejor que existe.

Números mal aproximables

El Teorema de Lioville asegura que los reales que se aproximan muy bien son trascendentes. Los números trascendentes, sin embargo, como se mostrará en el último capítulo, no se agotan con este criterio.

Definición 13. *Los reales mal aproximables,* **Bad***, son el siguiente conjunto*

$$\mathbf{Bad} := \left\{\alpha \in \mathbb{R} : \exists c > 0 \left(\forall \frac{p}{q} \in \mathbb{Q} \quad \left|\alpha - \frac{p}{q}\right| > \frac{c}{q^2}\right)\right\}.$$

Es un resultado sorprendente que se pueda dar un criterio tan simple para caracterizar a los números mal aproximables.

Teorema 14. *Si* $\alpha \in \mathbb{R} \setminus \mathbb{Q}$*, entonces* $\alpha \in \mathbf{Bad}$ *si y sólo si* α *tiene elementos acotados. Un corolario inmediato es que los irracionales cuadráticos son mal aproximables.*

Demostración. Ver [La02] p.23, Teorema 6 o [Kh] p.36, Teorema 23. □

De la caracterización anterior, se tiene que **Bad** no es numerable. Por otra parte, un resultado de Khinchin asegura que este conjunto tiene medida de Lebuesgue–en adelante denotada $\mathfrak{m}$– cero.

Teorema 15. *El conjunto* **Bad** *es no numerable y satisface* $\mathfrak{m}(\mathbf{Bad}) = 0$*.*

Demostración. Ver [La02] p.23, Teorema 6 o [Kh] p.60, Teorema 29. □

Apéndice B
El álgebra de Grassman

El álgebra de Grassman

El álgebra exterior en $\mathbb{R}^n$, aun siendo una noción conocida, se desarrolla brevemente en este apéndice. El objetivo principal es presentar resultados usados en el texto y establecer la notación. La mayor parte de las demostraciones sigue de las propiedades elementales de los determinantes, son omitidas. Todas pueden encontrarse en la referencia principal de esta sección, [Sch], o, bajo un contexto más general, en [La01].

Construcción del Álgebra de Grassman

Primero, se dotará de significado a la expresión $\mathbb{R}_p^n$ donde n es un entero positivo y p es un entero no negativo menor o igual que n. En adelante, n será un número natural fijo. Apegándose a la tradición, la base canónica de $\mathbb{R}^n$ se representa mediante $e_1, \ldots, e_n$.

El símbolo $\mathbb{R}_0^n$ es simplemente el conjunto de los números reales. Tómese a p en $\{1, \ldots, n\}$. Se denota por $C(n,p)$ al conjunto cuyos elementos son los subconjuntos de $[1..n]$ de cardinalidad p:

$$C(n,p) := \{X \subseteq [1..n] : |X| = p\}.$$

Para cada $\sigma \in C(n,p)$, $\sigma = \{i_1 < \ldots < i_p\}$, se consideran los productos cuña, expresiones de la forma

$$e_{\sigma_1} \wedge \cdots \wedge e_{\sigma_p}.$$

El espacio $\mathbb{R}_p^n$ es el $\mathbb{R}$-espacio vectorial de dimensión $\binom{n}{p} =: l$ generado por estos objetos. Este espacio admite un producto interior muy natural. Primero, si $\sigma \in C(n,p)$ es $\sigma = \{i_1 < \ldots < i_p\}$, se escribe

$$E_\sigma := e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p}.$$

Después, con $\delta_{\tau,\sigma}$ la delta de Kronecker, se define

$$\forall \sigma, \tau \in C(n,p) \quad \langle E_\sigma, E_\tau \rangle = \delta_{\tau,\sigma}.$$

Para calcular el producto en cualquier pareja de vectores en $\mathbb{R}_p^n$ se extiende la igualdad anterior por linealidad.

La formación de los productos cunã exige, hasta ahora, que los subíndices sean estrictamente crecientes. Para eliminar esta restricción se consideran dos casos: hay dos subíndices repetidos y todos son distintos. Cuando un vector aparezca dos veces en el producto, éste valdrá cero. Por otra parte, si los índices $i_1, i_2, \ldots, i_p$ son distintos pero

no forman una sucesión creciente, se toma $\pi \in \mathcal{S}_p$ (el grupo simétrico de orden p) para la que $i_{\pi(1)} < i_{\pi(2)} < \ldots < i_{\pi(n)}$ y se establece

$$e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p} = \operatorname{sgn}(\pi)\, e_{i_{\pi(1)}} \wedge e_{i_{\pi(2)}} \wedge \ldots \wedge e_{i_{\pi(p)}}.$$

Ahora, sea G_n el $\mathbb{R}$ espacio vectorial de dimensión 2^n dado por

$$G_n := R_0^n \oplus R_1^n \oplus \ldots \oplus R_n^n.$$

El producto cuña, hasta este momento tiene sentido dentro de cada $\mathbb{R}_q^n$ con $q \in [0..n]$. Las siguientes fórmulas lo llevan hasta los vectores de G_n obtenidos como el producto de vectores canónicos de $\mathbb{R}^n$:

$$1 \wedge 1 = 1$$
$$1 \wedge (e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p}) = (e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p}) \wedge 1 = e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p}$$
$$(e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p}) \wedge (e_{j_1} \wedge e_{j_2} \wedge \ldots \wedge e_{j_q}) = e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_p} \wedge e_{j_1} \wedge e_{j_2} \wedge \ldots \wedge e_{j_q}$$

Nótese que, por el Principio del Palomar de Dirichlet, cualquier producto que cuente con más de n factores es igual al cero en G_n.

Definición 1. *El álgebra obtenida al extender linealmente las fórmulas anteriores a G_n es el* ***Álgebra de Grassman****.*

Resultados útiles

Si $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(p)}$ son vectores en $\mathbb{R}^n$, se escribirá

$$X^{(p)} := \mathbf{x}^{(1)} \wedge \mathbf{x}^{(2)} \wedge \ldots \wedge \mathbf{x}^{(p)}. \tag{1}$$

Lema 2. *Sean $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^n$ con $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,n})$ para toda $i \in [1..n]$. Entonces*

$$\mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p = \sum_{\sigma \in C(n,p)} X_\sigma E_\sigma \quad \text{con} \quad X_\sigma = \det(x_{i,j}),$$

en donde $i \in [1..p]$ y $j \in \sigma$.

Demostración. Ver [Sch], Lema 6A, p. 104. □

Para ejemplificar el lema anterior, supóngase que $p = 2$, $n = 4$ y $\mathbf{x}_1 = (x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4})$ para $i \in \{1, 2\}$. Calculando,

$$\begin{aligned}
\mathbf{x}_1 \wedge \mathbf{x}_2 &= \left(\sum_{i=1}^{4} x_{1,j}\mathbf{e}_j\right) \wedge \left(\sum_{j=1}^{4} x_{2,j}\mathbf{e}_j\right) \\
&= \sum_{i,j=1}^{4} x_{1,i}x_{2,j}\ \mathbf{e}_i \wedge \mathbf{e}_j \\
&= x_{1,1}\left[x_{2,2}\ \mathbf{e}_1 \wedge \mathbf{e}_2 + x_{2,3}\ \mathbf{e}_1 \wedge \mathbf{e}_3 + x_{2,4}\ \mathbf{e}_1 \wedge \mathbf{e}_4\right] + \\
&\quad + x_{1,2}\left[x_{2,1}\left(\mathbf{e}_2 \wedge \mathbf{e}_1\right) + x_{2,3}\left(\mathbf{e}_2 \wedge \mathbf{e}_3\right) + x_{2,4}\left(\mathbf{e}_2 \wedge \mathbf{e}_4\right)\right] + \\
&\quad + x_{1,3}\left[x_{2,1}\left(\mathbf{e}_3 \wedge \mathbf{e}_1\right) + x_{2,2}\left(\mathbf{e}_3 \wedge \mathbf{e}_3\right) + x_{2,4}\left(\mathbf{e}_3 \wedge \mathbf{e}_4\right)\right] \\
&\quad + x_{1,4}\left[x_{2,1}\left(\mathbf{e}_4 \wedge \mathbf{e}_1\right) + x_{2,2}\left(\mathbf{e}_4 \wedge \mathbf{e}_3\right) + x_{2,3}\left(\mathbf{e}_4 \wedge \mathbf{e}_3\right)\right] \\
&= \begin{vmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{vmatrix} \mathbf{e}_1 \wedge \mathbf{e}_2 + \begin{vmatrix} x_{1,1} & x_{1,3} \\ x_{2,1} & x_{2,3} \end{vmatrix} \mathbf{e}_1 \wedge \mathbf{e}_3 + \begin{vmatrix} x_{1,1} & x_{1,4} \\ x_{2,1} & x_{2,4} \end{vmatrix} \mathbf{e}_1 \wedge \mathbf{e}_4 + \\
&\quad + \begin{vmatrix} x_{1,2} & x_{1,3} \\ x_{2,2} & x_{2,3} \end{vmatrix} \mathbf{e}_2 \wedge \mathbf{e}_3 \begin{vmatrix} x_{1,2} & x_{1,4} \\ x_{2,2} & x_{2,4} \end{vmatrix} \mathbf{e}_2 \wedge \mathbf{e}_4 + \begin{vmatrix} x_{1,3} & x_{1,4} \\ x_{2,3} & x_{2,4} \end{vmatrix} \mathbf{e}_3 \wedge \mathbf{e}_4.
\end{aligned}$$

La existencia de los subespacios en el Teorema de Schmidt la dará precisamente estas ideas, después de una larga y complicada construcción de los elementos a los que se les aplicará.

Lema 3. *Sean $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p \in \mathbb{R}^n$ con $p \leq n$. Estos puntos son linealmente dependientes si y sólo si para cualquier $\sigma \in C(n,p)$ se cumple $\det(x_{i,j}) = 0$ con $i \in [1..p]$ y $j \in \sigma$.*

Demostración. La necesidad es trivial. Para la suficiencia, se fija $p \in \mathbb{N}$. El resultado es evidente para $n = p + 0$. Con inducción matemática se verifica el resultado para $n \in \mathbb{N}$ con $n \geq p$. □

Lema 4. *Sean $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^n$. Entonces, $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n = 0$ si y sólo si $\mathbf{x}_1, \ldots, \mathbf{x}_p$ son linealmente dependientes.*

Demostración. Ver [Sch], Lema 6B, p. 104. □

Lema 5. *Sean $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$ dos colecciones de puntos de $\mathbb{R}^n$ linealmente independientes. Los vectores $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n$ y $\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_n$ son linealmente dependientes si y sólo si $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$ generan el mismo subespacio de $\mathbb{R}^n$.*

Demostración. Ver [Sch], Lema 6C, p. 105. □

Lema 6 (Identidad de Laplace)**.** *Para $\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{y}_1, \ldots, \mathbf{y}_p \in \mathbb{R}^n$ se cumple*

$$\langle \mathbf{x}_1 \wedge \ldots \wedge \mathbf{x}_p, \mathbf{y}_1 \wedge \ldots \wedge \mathbf{y}_p \rangle_{\mathbb{R}^n_p} = \begin{vmatrix} \langle \mathbf{x}_1, \mathbf{y}_1 \rangle & \langle \mathbf{x}_1, \mathbf{y}_2 \rangle & \ldots & \langle \mathbf{x}_1, \mathbf{y}_p \rangle \\ \langle \mathbf{x}_2, \mathbf{y}_1 \rangle & \langle \mathbf{x}_2, \mathbf{y}_2 \rangle & \ldots & \langle \mathbf{x}_2, \mathbf{y}_p \rangle \\ \vdots & & & \vdots \\ \langle \mathbf{x}_p, \mathbf{y}_1 \rangle & \langle \mathbf{x}_p, \mathbf{y}_2 \rangle & \ldots & \langle \mathbf{x}_p, \mathbf{y}_p \rangle \end{vmatrix}. \tag{2}$$

Demostración. Ver [Sch], Lema 6D, p. 105. □

Cabe mencionar que si $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ y $\{\mathbf{y}_1, \ldots, \mathbf{y}_p\}$ son colecciones linealmente independientes, el lema anterior implica que el producto interior del lado izquierdo es distinto de cero. Tras aplicar una transformación lineal adecuada, el lado derecho es un determinante de una matriz de $p \times p$ cuyas columnas son linealmente independientes (serían imagen de un isomorfismo lineal).

El lema siguiente requiere una construcción llamada el p-ésima compuesto de una matriz. Se supondrá que los elementos de $C(n,p)$ han sido enlistados de acuerdo con algún orden fijo. En el texto, el orden que se elige depende de los mínimos sucesivos del paralelepípedo formado por una base fija de $\mathbb{R}^n$.

Definición 7. *Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^n$ y $p \in \{1, 2, \ldots, n\}$. Además, supóngase que $C(n_p) = \{\sigma_1, \ldots, \sigma_l\}$. El p-ésimo compuesto de $(\mathbf{x}_1 | \ldots | \mathbf{x}_n)^t$ es la matriz de tamaño $\binom{n}{p} \times \binom{n}{p}$ cuya entrada (r,j) está dada por*

$$X_{\sigma,\tau} = \begin{vmatrix} x_{i_1,j_1} & x_{i_2,j_1} & \ldots & x_{i_p,j_1} \\ x_{i_1,j_2} & x_{i_2,j_2} & \ldots & x_{i_p,j_2} \\ \vdots & & & \vdots \\ x_{i_1,j_p} & x_{i_2,j_p} & \ldots & x_{i_p,j_p} \end{vmatrix}$$

donde $\sigma = \sigma_r = \{i_1 < \ldots < i_p\}$ y $\tau = \sigma_s = \{j_1 < \ldots < j_p\}$.

Por ejemplo, sean $n = 4, p = 3$ y $\mathbf{x}_i = (x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4})$ para $i \in \{1,2,3\}$. Se considera el orden

$$C(4,3) = \{\{1,2,3\},\{1,2,4\},\{2,3,4\}\}.$$

Entonces, el tercer compuesto de la matriz

$$\begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} \end{pmatrix}$$

es la matriz de 3×3 es

$$\begin{pmatrix} \begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{vmatrix} & \begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,4} \end{vmatrix} & \begin{vmatrix} x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,2} & x_{3,3} & x_{3,4} \end{vmatrix} \\ \begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{4,1} & x_{4,2} & x_{4,3} \end{vmatrix} & \begin{vmatrix} x_{1,1} & x_{1,2} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,4} \\ x_{4,1} & x_{4,2} & x_{4,4} \end{vmatrix} & \begin{vmatrix} x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,2} & x_{2,3} & x_{2,4} \\ x_{4,2} & x_{4,3} & x_{4,4} \end{vmatrix} \\ \begin{vmatrix} x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \\ x_{4,1} & x_{4,2} & x_{4,3} \end{vmatrix} & \begin{vmatrix} x_{2,1} & x_{2,2} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,4} \\ x_{4,1} & x_{4,2} & x_{4,4} \end{vmatrix} & \begin{vmatrix} x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,2} & x_{3,3} & x_{3,4} \\ x_{4,2} & x_{4,3} & x_{4,4} \end{vmatrix} \end{pmatrix}$$

Lema 8. *Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ puntos en $\mathbb{R}^n$. Entonces,*

$$\det\left(X_{\sigma,\tau}\right)_{\sigma,\tau \in C(n,p)} = \left(\det(\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n)\right)^{\binom{n-1}{p-1}}.$$

Demostración. Ver [Sch], Lema 6E, p. 105. □

Con las proposiciones hasta ahora dadas es sencillo concluir el último resultado de esta sección. Ésta se utilizará fuertemente en la prueba del Teorema del Subespacio.

Lema 9. *Sea $\mathbf{a}_1, \ldots, \mathbf{a}_n$ una base de $\mathbb{R}^n$. Para cada $\sigma \in C(n,p)$, $\sigma = \{i_1 < \ldots < i_p\}$ se escribe*

$$A_\sigma := \mathbf{a}_{i_1} \wedge \mathbf{a}_{i_2} \wedge \ldots \wedge \mathbf{a}_{i_p}.$$

Entonces, son ciertas los siguientes enunciados.

I. *Los $l = \binom{n}{p}$ elementos A_σ forman una base de $\mathbb{R}^n_p$.*

II. *Si $\det(\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n) = 1$, entonces el determinante de los A_σ también es 1.*

III. *Sea $\mathbf{a}_1^*, \ldots, \mathbf{a}_n^*$ la base recíproca de $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Para cada $\sigma \in C(n,p)$ con $\sigma = \{i_1 < \ldots < i_p\}$ se escribe*

$$A_\sigma^* = a_{i_1}^* \wedge a_{i_2}^* \wedge \ldots \wedge a_{i_p}^*.$$

Entonces, los vectores $\{A_\sigma^*\}_{\sigma \in C(n,p)}$ forman la base recíproca de $\{A_\sigma\}_{\sigma \in C(n,p)}$; esto es, si $\delta_{\sigma,\tau}$ es el símbolo de Kronecker,

$$\forall \sigma, \tau \in C(n,p) \quad \langle A_\sigma, A_\tau^* \rangle_{\mathbb{R}^n_p} = \delta_{\sigma\tau}. \tag{3}$$

Demostración. Ver [Sch], Lema 6F, p. 108. □

Apéndice C
Resultados de la Geometría de Números

Demostraciones de la geometría de números

Las fronteras que separan diversas áreas de las matemáticas se difuminan hasta desaparecer cuando se atacan algunos problemas. Los Teoremas de Roth y de Schmidt son buenas instancias de ello. Las pruebas se basan, fundamentalmente, en conceptos geométricas para concluir aspectos algebraica.

Este apéndice expone las ideas sobre la Geometría de Números. La estructura es la misma que en el segundo capítulo. Todas las proposiciones y las definiciones son reformuladas para evitar las referencias al cuerpo del texto con la única excepción de la expresión 2.3. Algunos detalles que han aparecido en la discusión del segundo capítulo son omitidos.

Primeras definiciones

Los objetos centrales en la geometría de los números son las retículas[3], siendo $\mathbb{Z}^n$ el ejemplo más importante. Minkowski observó en el Siglo XIX que la geometría de $\mathbb{R}^n$ ayuda a deducir propiedades de los enteros racionales. En esta ocasión, los teoremas presentados serán de existencia y sobre proporciones.

Definición 1. *Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\} \subset \mathbb{R}^n$ vectores linealmente independientes. La **retícula** generada por $\mathcal{A}$, Λ, es el conjunto*

$$\Lambda := \left\{ \sum_{j=1}^{n} \alpha_j \mathbf{a}_j : \alpha_j \in \mathbb{Z} \right\}.$$

*En este caso, se dice que $\mathcal{A}$ forma una **base** de Λ. Equivalentemente, si $A = (\mathbf{a}_1|\mathbf{a}_2|\ldots|\mathbf{a}_n)$, y $T_A : \mathbb{R}^n \to \mathbb{R}^n$ está dada por $T_A(\mathbf{x}) = A\mathbf{x}$, entonces $\Lambda = T(\mathbb{Z}^n)$. Si Λ, Γ son retículas y $\Lambda \subseteq \Gamma$, entonces se dice que Λ es una **subretícula** de Γ.*

El ejemplo más importante es la retícula, Λ_0, generada por la base canónica. En ella y, en general, es claro que puede existir una infinidad de bases. Además, es fácil caracterizarlas. Antes, se recuerda que las matrices con entradas enteras con determinante 1 o -1 se llaman **unimodulares**.

[3]Se traduce *retícula* al término *lattice*.

Tómense $\Lambda \subseteq \mathbb{R}^n$una retícula y $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ y $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ bases de Λ con $A = (\mathbf{a}_1|\ldots|\mathbf{a}_n)$, $B = (\mathbf{b}_1|\ldots|\mathbf{b}_n)$. Existen una matrices con coeficientes enteros, V y W, tales que

$$AV = B, \qquad BW = A.$$

La invertibilidad de las matrices involucradas implica $W = B^{-1}A = (A^{-1}B)^{-1} = V^{-1}$; en consecuencia,

$$\det(V),\ \frac{1}{\det(V)} = \det(W) \in \mathbb{Z} \implies |\det(V)| = 1.$$

Por otra parte, si $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ es una base de Λ, $A = (\mathbf{a}_1|\mathbf{a}_2|\ldots|\mathbf{a}_n)$ y V es una matriz unimodular, las columnas de $B = AV$ forman una base de Λ. En efecto $A = BV^{-1}$ y V^{-1} tiene entradas enteras (por la Regla de Cramer); luego, la retícula generada por las columnas de B contiene a Λ (la contención opuesta se da por cómo fue definida B).

Teorema 2. *Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ una base de Λ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} \subset \Lambda$ un conjunto linealmente independiente. Si $A = (\mathbf{a}_1|\mathbf{a}_2|\ldots|\mathbf{a}_n)$ y $B = (\mathbf{b}_1|\mathbf{b}_2|\ldots|\mathbf{b}_n)$, entonces $\mathcal{B}$ es base de Λ si y sólo si existe una matriz unimodular V tal que $B = AV$.*

Manteniendo la notación, se ve que el valor absoluto del determinante de una base de Λ sólo depende de la retícula:

$$|\det(A)| = |\det(BV)| = |\det(B)||\det(V)| = |\det(B)|.$$

Definición 3. *Sean Λ una retícula, $\mathcal{A}$ una base de Λ y $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. El **determinante** de Λ, $\mathrm{d}(\Lambda)$, es*

$$\mathrm{d}(\Lambda) := |\det(A)|.$$

*Si $\Lambda \subseteq M$ son retículas, el **índice** de M en Λ es*

$$D := \frac{d(\Lambda)}{d(M)}.$$

Cuando Λ, con base $\{a_1, \ldots, a_n\}$, es subretícula de M, con base $\{b_1, \ldots, b_n\}$ existe una matriz con entradas enteras, V, tal que $A = BV$ con A y B definidas como antes. Utilizando, por ejemplo, la Regla de Cramer, se deduce la existencia de una matriz W con entradas enteras tal que $AW = BD$. Esto da

$$DM \subseteq \Lambda \subseteq M. \tag{1}$$

Estudiar la estructura de las matrices no sólo rinde frutos en un entorno de cómputo científico. Las implicaciones teóricas tampoco son despreciables, facilitan los cálculos y hacen evidente. En algunos casos, la invertibilidad de ciertas transformaciones vía la independencia lineal resulta evidente.

Lema 4. *Sean Λ, M retículas tales que $\Lambda \subseteq M$. Entonces, para cada base $\mathbf{b}_1, \ldots, \mathbf{b}_n$ de M existen una base $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de Λ y $V \in \mathcal{M}(\mathbb{Z}, n)$ triangular superior tales que, con $B = (\mathbf{b}_1|\ldots|\mathbf{b}_n)$ y $A = (\mathbf{a}_1|\ldots|\mathbf{a}_n)$,*

$$A = BV.$$

Demostración. Como $DM \subseteq \Lambda$, para cada $i \in [1..n]$ existen puntos de la forma $\alpha_1 b_1 + \ldots + \alpha_i b_i \in \Lambda$. Defínanse $a_1, \ldots, a_n \in \Lambda$ de la forma

$$\forall i \in [1..n] \quad \sum_{j=1}^{i} v_{i,j} b_j$$

con cada $v_{i,j} \in \mathbb{Z}$ y $|v_{i,i}| > 0$ mínimo. Claramente, $\{a_1, \ldots, a_n\} \subseteq \Lambda$ es un conjunto linealmente independiente. Supóngase que no es una base de Λ. Tómese $c \in \Lambda$ fuera de la retícula generada por $a_1, \ldots, a_m$ con

$$c = \sum_{i=1}^{k} t_i b_i,\ t_i \in \mathbb{Z}$$

y k mínimo. Por el Algoritmo de Euclides, existe $s \in \mathbb{Z}$ tal que $|t_k - s v_{k,k}| < |v_{k,k}|$. Entonces, el vector

$$c - s a_k = \sum_{i=1}^{k} (t_i - v_{k,i} s) b_i \in \Lambda$$

no está en la retícula generada por $\{a_1, \ldots, a_n\}$ y, como $|v_{k,k}| > 0$ es mínimo, $t_k - v_{k,k} s \neq 0$ que contradice la minimalidad de $v_{k,k}$. Por lo tanto, no puede existir tal c y $\{a_1, \ldots, a_n\}$ genera a Λ. □

Un caso particular importante del lema anterior es cuando Λ y M son la misma retícula. Obedeciendo el dictamen del pragmatismo, En el resto de este apéndice se considerará únicamente la retícula generada por la base canónica. Si bien los primeros resultados no exigen mucho trabajo adicional para presentarlos en retículas arbitrarias, el resto aumentaría significativamente la longitud del texto.

Teorema de Minkowski y Lema de Siegel

La existencia de los polinomios en los Teoremas de Schmidt y Roth se apoya fuertemente en el Lema de Siegel. Los resultados de Minkowski y Siegel dan condiciones bajo las que los subconjuntos convexos de $\mathbb{R}^n$ y simétricos con respecto al origen contienen puntos con entradas en $\mathbb{Z}$. La medida de Lebesgue en $\mathbb{R}^n$ se denota $\mathfrak{m}$ y la σ-álgebra de los Lebesgue Medibles de $\mathbb{R}^n$, $\mathfrak{M}$.

Teorema 5 (Minkowski, 1896)**.** *Sea $\mathscr{S} \in \mathfrak{M}$ convexo y simétrico con respecto al origen. Supóngase que $\mathfrak{m}(\mathscr{S}) > 2^n$ o que $\mathscr{S}$ es compacto y $\mathfrak{m}(\mathscr{S}) = 2^n$. Entonces, $\mathscr{S}$ contiene una pareja de puntos enteros, $\mathbf{u}$ y $-\mathbf{u}$, distintos de cero.*

Demostración. I. Se afirma que[4] si $\mathscr{R} \subseteq \mathbb{R}^n$ es un un covexo, acotado y $\mathfrak{m}(\mathscr{R}) > 1$, entonces existen $\mathbf{x}, \mathbf{y} \in \mathscr{R}$ tales que $\mathbf{x} - \mathbf{y} \in \mathbb{Z}^n$. En efecto, escribiendo $\mathbf{v} = (v_1, \ldots, v_n)$ para cada $\mathbf{v} \in \mathbb{Z}^n$

$$\left(\forall \mathbf{v} \in \mathbb{Z}^n \ \ \mathscr{R}(\mathbf{v}) := \left\{ \mathbf{r} \in \mathscr{R} : \mathbf{r} \in \prod_{j=1}^{n} [v_n, v_{n+1}) \right\} \right) \implies \left(\forall \mathbf{v} \in \mathbb{Z}^n \ \ \mathfrak{m}(\mathscr{R}(\mathbf{v})) = \mathfrak{m}(\mathscr{R}(\mathbf{v}) - \mathbf{v}) \leq 1 \right).$$

Para cada $\mathbf{v} \in \mathbb{Z}^n$ vale $\mathscr{R}(\mathbf{v}) - \mathbf{v} \subseteq [0,1]^n$, por lo que $\mathfrak{m}(\mathscr{R}(\mathbf{v})) = \mathfrak{m}(\mathscr{R}(\mathbf{v}) - \mathbf{v}) \leq 1$. Entonces,

$$\mathfrak{m}(\mathscr{R}) = \sum_{\mathbf{v} \in \mathbb{Z}^n} \mathfrak{m}(\mathscr{R}(\mathbf{v})) > 1 \geq \mathfrak{m}\left(\bigcup_{\mathbf{v} \in \mathbb{Z}^n} \mathscr{R}(\mathbf{v}) - \mathbf{v} \right) \implies \exists \mathbf{u}, \mathbf{v} \in \mathbb{Z}^n \quad [\mathscr{R}(\mathbf{u}) - \mathbf{u}] \cap [\mathscr{R}(\mathbf{v}) - \mathbf{v}] \neq \varnothing \tag{2}$$

Fijando a dos puntos $\mathbf{u}, \mathbf{v}$ distintos que cumplen con lo anterior se concluye

$$\left(\exists\, \mathbf{x}, \mathbf{y} \in \mathscr{R} \quad \mathbf{x} - \mathbf{u} = \mathbf{y} - \mathbf{v} \right) \implies \left(\exists\, \mathbf{x}, \mathbf{y} \in \mathscr{R} \quad \mathbf{0} \neq \mathbf{x} - \mathbf{y} = \mathbf{u} - \mathbf{v} \in \mathbb{Z}^n \right).$$

[4] Este resultado se conoce como el Lema de Blichfeldt

II. Supóngase que $\mathfrak{m}(\mathscr{S}) > 2^n$. Sea $\mathscr{R} := 2^{-1}\mathscr{S}$, por lo que $\mathfrak{m}(\mathscr{R}) > 1$ y, por las ideas de arriba, existen $\mathbf{x}, \mathbf{y} \in \mathscr{R}$ tales que $\mathbf{0} \neq \mathbf{x} - \mathbf{y} \in \mathscr{R} \cap \mathbb{Z}^n$. Por la simetría de $\mathscr{S}$, $2\mathbf{y} \in \mathscr{S}$ da $-2\mathbf{y} \in \mathscr{S}$. Entonces, recordando la convexidad de $\mathscr{S}$,

$$\mathbf{x}, \mathbf{y} \in \frac{1}{2}\mathscr{S} \implies 2\mathbf{x}, 2\mathbf{y} \in \mathscr{S} \implies \mathbf{x} - \mathbf{y} = \frac{1}{2}(2\mathbf{x}) - \frac{1}{2}(2\mathbf{y}) \in \mathbb{Z}^n \cap \mathscr{S}.$$

III. Finalmente, supóngase que $\mathscr{S}$ es compacto y $\mathfrak{m}(\mathscr{S}) = 2^n$. Sea $(\varepsilon_n)_{n=1}^{\infty}$ una sucesión en $\mathbb{R}$ tal que $\varepsilon_n \searrow 0$ cuando $n \to \infty$. Por lo anterior, hay una sucesión, $(\mathbf{z}_n)_{n=1}^{\infty}$, con

$$\forall n \in \mathbb{N} \qquad \mathbf{z}_n \in \mathscr{S} + \overline{B}(0; \varepsilon_n) \quad \& \quad 0 \neq \mathbf{z}_n \in \mathbb{Z}^n.$$

Como $\mathscr{S} + \overline{B}(0; \varepsilon_1)$ es compacto y contiene a todos los términos de $(\mathbf{z}_n)_{n=1}^{\infty}$, hay una subsucesión $(\mathbf{z}_{n_j})_{j=1}^{\infty}$ convergente. Ya que $\mathbb{Z}^n$ es discreto, a la larga $(\mathbf{z}_{n_j})_{j=1}^{\infty}$ es constante. Por lo tanto, llamando $\mathbf{z}$ al límite,

$$\mathbf{0} \neq \mathbf{z} \in \bigcap_{j=1}^{n} \left\{\mathscr{S} + \overline{B}(0; \varepsilon_{n_j})\right\} = \overline{\mathscr{S}} = \mathscr{S}. \qquad \square$$

Lema 6 (Siegel, 1929)**.** *Sea $A \in \mathcal{M}(\mathbb{Z}, m, n)$, $A = (a_{i,j})$, con $m > n$ y $K \geq 0$ tal que $|a_{i,j}| \leq K$ para cualesquiera i, j. Entonces, existe $\mathbf{x} \in \mathbb{Z}^n$ tal que*

$$A\mathbf{x} = \mathbf{0}, \quad 1 \leq \|\mathbf{x}\|_{\infty} \leq \lfloor (nK)^{\frac{m}{n-m}} \rfloor.$$

Demostración. Tomando $Z = \lfloor (nK)^{\frac{m}{n-m}} \rfloor$ se obtiene $Z \leq (nK)^{\frac{m}{n-m}} < Z + 1$ y $(nK) < (Z+1)^{\frac{m}{m-n}}$ que, como $n, K \geq 1$, implica

$$nKZ + 1 \leq nK(Z+1) \leq (Z+1)^{\frac{n}{m}}. \tag{3}$$

Por otra parte, para cualquier $\mathbf{z} \in \mathbb{Z}^n$ tal que $0 \leq \mathbf{z} \leq \mathbf{Z} = (Z, \ldots, Z)$ se cumple

$$\forall j \in \{1, \ldots, m\} \quad B_j Z \leq (A\mathbf{z})_j \leq C_j Z,$$

donde, llamando χ_X a la función característica de $X \subseteq \mathbb{R}^n$,

$$\forall\ j \in \{1, \ldots, m\} \quad B_j := \sum_{k=1}^{n} a_{j,k}\chi_{(-\infty,0)}(a_{j,k}), \quad C_j := \sum_{k=1}^{n} a_{j,k}\chi_{(0,+\infty)}(a_{j,k}).$$

Entonces, $(A\mathbf{z})_j$ puede asumir a lo más $Z(B_j + C_j) + 1 \leq nZK + 1$ valores y el vector $A\mathbf{z}$, a lo más $(nKZ+1)^m$. Como hay $(Z+1)^n$ valores posibles para $\mathbf{z}$, (3) y el Principio del Palomar de Dirichlet, dan la existencia $\tilde{\mathbf{z}}, \mathbf{z}'$ distintos tales que $\mathbf{0} \leq \tilde{\mathbf{z}}, \mathbf{z}' \leq \mathbf{Z}$ y $A\tilde{\mathbf{z}} = A\mathbf{z}'$. El vector $\mathbf{z} = \tilde{\mathbf{z}} - \mathbf{z}'$ realiza la conclusión. $\square$

Mínimos sucesivos

En esta sección se supondrá que $\mathscr{R} \subseteq \mathbb{R}^n$ es convexo, compacto, simétrico en el origen y $0 < \mathfrak{m}(\mathscr{R}) < \infty$.

Definición 7. *La función distancia $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ con respecto a $\mathscr{R}$ es*

$$F(\mathbf{x}) = \inf\{\lambda > 0 : \lambda^{-1}\mathbf{x} \in \mathscr{R}\}$$

La compacidad de $\mathscr{R}$ implica $\mathbf{x} \in F(\mathbf{x})^{-1}\mathscr{R}$ para cualquier $\mathbf{x} \in \mathbb{R}^n$, $F(\mathbf{x}) \leq 1$ si y sólo si $\mathbf{x} \in \mathscr{R}$, además para cualquier $\lambda > 0$ se tiene que $\lambda \geq F(x)$ es equivalente a $\mathbf{x} \in \lambda^{-1}\mathscr{R}$. Si bien F no es la función dada por $d_{\mathscr{R}}(\mathbf{x}) = \inf_{\mathbf{r}\in\mathscr{R}} \|\mathbf{x} - \mathbf{r}\|$, sí es una norma.

Lema 8. *Sean* $\mathbf{x}$ *y* $\mathbf{y}$ *elementos de* $\mathbb{R}^n$.

1. $F(\mathbf{x}) = 0$ *si y sólo si* $\mathbf{x} = \mathbf{0}$.
2. *Para cualquier* $\lambda \in \mathbb{R}$ *se cumple* $F(\lambda\mathbf{x}) = |\lambda|F(\mathbf{x})$.
3. $F(\mathbf{x} + \mathbf{y}) \leq F(\mathbf{x}) + F(\mathbf{y})$.

Demostración. Los primeros dos puntos son evidentes: el primero por la compacidad de $\mathscr{R}$ y el segundo por las propiedades elementales de los ínfimos y la simetría de $\mathscr{R}$. El tercero no ofrece resistencia por la convexidad de $\mathscr{R}$ y

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \frac{1}{F(\mathbf{x}) + F(\mathbf{y})} = \frac{F(\mathbf{x})}{F(\mathbf{x}) + F(\mathbf{y})}\frac{\mathbf{x}}{F(\mathbf{x})} + \frac{F(\mathbf{y})}{F(\mathbf{x}) + F(\mathbf{y})}\frac{\mathbf{y}}{F(\mathbf{y})}. \qquad \square$$

Dado que $\mathfrak{m}(\mathscr{R}) > 0$, hay n vectores linealmente independientes en $\mathscr{R}$, $\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(n)}\}$. Si $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$ satisface $\|\boldsymbol{\mu}\|_1 \leq 1$, el tercer inciso del Lema 8 e inducción garantizan[5]

$$F\left(\sum_{j=1}^{n} \mu_j \mathbf{y}^{(j)}\right) \leq \sum_{j=1}^{n} |\mu_j| F(\mathbf{y}^{(j)}) \leq \sum_{j=1}^{n} |\mu_j| < 1 \implies \overline{\mathrm{co}}\left\{\pm\mathbf{y}^{(1)}, \ldots, \pm\mathbf{y}^{(n)}\right\} \subseteq \mathscr{R}.$$

Claramente, $\overline{\mathrm{co}}\left\{\pm\mathbf{y}^{(1)}, \ldots, \pm\mathbf{y}^{(n)}\right\}$ es una vecindad del origen y, por ende, dado $\mathbf{x} \in \mathbb{R}^n$, existe $\lambda > 0$ tal que $\mathbf{x} \in \lambda\mathscr{R}$. En particular, para alguna $\lambda > 0$ el conjunto $\lambda\mathscr{R}$ tiene k puntos enteros linealmente independientes ($k \in [1..n]$). Esto da pie a la siguiente concepto.

Definición 9. *Para cada* $j \in \{1, 2, \ldots, n\}$ *el* j*-ésimo* ***mínimo sucesivo*** *de* $\mathscr{R}$ *es*

$$\lambda_j := \inf\{\lambda > 0 : \lambda\mathscr{R} \text{ tiene } j \text{ puntos enteros linealmente independientes}\}.$$

Segundo Teorema de Minkowski

Llamando $\lambda_1 \leq \ldots \leq \lambda_n$ a los mínimos sucesivos de $\mathscr{R}$, el Segundo Teorema de Minkowski establece $1 \ll \lambda_1 \cdots \lambda_n \ll 1$, las constantes en $\ll$ dependen de n. La desigualdad izquierda se prueba a detalle; pero en la segunda, nada más se da un esquema.

Teorema 10 (Segundo Teorema de Minkowski)**.** *Los mínimos sucesivos de* $\mathscr{R}$ *satisfacen*

$$\frac{2^n}{n!} \leq \lambda_1 \cdots \lambda_n \mathfrak{m}\mathscr{R} \leq 2^n.$$

Demostración. **Lado izquierdo.** Elíjanse n puntos linealmente independientes con coordenadas enteras, $\mathbf{g}_1, \ldots, \mathbf{g}_n$ tales que $\mathbf{g}_k \in \lambda_k\mathscr{R}$. Es claro que la matriz $X = (\mathbf{g}^{(1)}| \ldots |\mathbf{g}^{(n)})$ es invertible y cumple con $|\det(X)| \geq 1$. Las observaciones hechas tras la definición de F se traducen en las desigualdad $F(\mathbf{g}_j) \leq \lambda_j$ para cada $j \in [1..n]$. Cuando $\mu_1, \ldots, \mu_n \in \mathbb{R}$ satisfacen $|\mu_1|\lambda_1 + \ldots + |\mu_n|\lambda_n \leq 1$ el tercer punto del Lema 8 implica

$$F\left(\sum_{j=1}^{n} \mu_j \mathbf{g}_j\right) \leq \sum_{j=1}^{n} |\mu_j| F(\mathbf{g}_j) \leq \sum_{j=1}^{n} |\mu_j|\lambda_j \leq 1 \implies \mu_1\mathbf{x}^{(1)} + \ldots + \mu_n\mathbf{x}^{(n)} \in \mathscr{R}.$$

[5] $\overline{\mathrm{co}}\left\{\pm\mathbf{y}^{(1)}, \ldots, \pm\mathbf{y}^{(n)}\right\}$ es la cerradura de la envolvente convexa de $\{\pm\mathbf{y}^{(1)}, \ldots, \pm\mathbf{y}^{(n)}\}$.

En consecuencia,

$$\{\mu_1\mathbf{g}_1 + \ldots + \mu_n\mathbf{g}_n : |\mu_1|\lambda_1 + \ldots + |\mu_n|\,\lambda_n \le 1\} \subseteq \mathscr{R};$$

por lo que, integrando con respecto a $\mu_1, \ldots, \mu_n$,

$$\mathfrak{m}(\mathscr{R}) \ge \mathfrak{m}\,\{\mu_1\mathbf{g}_1 + \ldots + \mu_n\mathbf{g}_n : |\mu_1|\lambda_1 + \ldots + |\mu_n|\,\lambda_n \le 1\} = \frac{2^n|\det(X)|}{n!\lambda_1\cdots\lambda_n} \ge \frac{2^n}{n!\lambda_1\cdots\lambda_n},$$

que da la primera desigualdad.

Lado derecho.

I. Haciendo un cambio de base adecuado (cfr. Lema 4) se puede suponer que $(\mathbf{g}_1|\ldots|\mathbf{g}_n)$ es una matriz triangular superior. Bajo esta suposición se tienen dos propiedades inmediatas:

   I. Si $\mathbf{x} \in \mathbb{Z}^n$ y $F(\mathbf{x}) < \lambda_k$, entonces $x_{k+1} = \ldots = x_n = 0$.
   II. Si $\mathbf{x} - \mathbf{y} \in \mathbb{Z}^n$, $F(\mathbf{x}) < \frac{1}{2}\lambda_j$ y $F(\mathbf{y}) < \frac{1}{2}\lambda_j$, entonces $x_{k+1} = y_{k+1}, \ldots, x_n = y_n$.

II. Para cada $\lambda \ge 0$ se definen $\mathcal{W}_0(\lambda) := \lambda\mathscr{R}$ y

$$\forall j \in [1..n] \quad \mathcal{W}_j(\lambda) := \{(\{x_1\}, \ldots, \{x_j\}, x_{j+1}, \ldots, x_n) : \mathbf{x} \in \lambda\mathscr{R}\}.$$

Para cada $j \in [0..n]$ la función $\lambda \mapsto \mathfrak{m}\mathcal{W}_j(\lambda)$ es continua. Se define para cualquier $j \in [0..n]$, $V_j(\lambda) := \mathfrak{m}\mathcal{W}_j(\lambda)$. Se verifica rápidamente que para cualquier $\lambda \ge 0$ vale $V_n(\lambda) \le 1$.

III. Para cada $j \in [0..n-1]$ se tiene que $0 \le \lambda \le \frac{1}{2}\lambda_{j+1}$ implica $V_n(\lambda) = V_j(\lambda)$; en particular,

$$0 < \lambda \le \frac{\lambda_1}{2} \implies V_n(\lambda) = V_0(\lambda) = \lambda^n\mathfrak{m}\mathscr{R}. \tag{4}$$

IV. Integrando secciones de $\mathscr{R}$ en subespacios de dimensión $j \in [1..n]$ se concluye

$$\forall j \in [1..n] \qquad 0 < \mu \le \lambda \implies \left(\frac{\lambda}{\mu}\right)^{n-j} V_j(\mu) \le V_j(\lambda). \tag{5}$$

V. De (4) y (5) siguen las expresiones

$$\begin{aligned}
\left(\frac{1}{2}\lambda_1\right)^n \mathfrak{m}\mathscr{R} &= V_n\left(\frac{1}{2}\lambda_1\right),\\
\left(\frac{\lambda_2}{\lambda_1}\right)^{n-1} V_n\left(\frac{1}{2}\lambda_1\right) &\le V_n\left(\frac{1}{2}\lambda_2\right),\\
\left(\frac{\lambda_3}{\lambda_2}\right)^{n-2} V_n\left(\frac{1}{2}\lambda_2\right) &\le V_n\left(\frac{1}{2}\lambda_3\right),\\
&\vdots\\
\left(\frac{\lambda_n}{\lambda_{n-1}}\right)^{1} V_n\left(\frac{1}{2}\lambda_{n-1}\right) &\le V_n\left(\frac{1}{2}\lambda_n\right).
\end{aligned}$$

Sustituyendo de abajo y usando que $\mathcal{W}_n(2^{-1}\lambda_n) \subseteq [0,1)^n$ se concluye lo que se quería probar:

$$\frac{1}{2^n}\lambda_1\cdots\lambda_n\mathfrak{m}\mathscr{R} \le V_n\left(\frac{1}{2}\lambda_n\right) \le 1 \implies \lambda_1\cdots\lambda_n\mathfrak{m}\mathscr{R} \le 2^n. \qquad \square$$

El Lema de Davenport

Se recuerda algunos conceptos discutidos y utilizados ampliacmente en el texto. Si $\{\mathbf{a}_1,\dots,\mathbf{a}_n\}\subseteq\mathbb{R}^n$ es una base, el paralelepípedo generado por $\mathbf{a}_1,\dots,\mathbf{a}_n$ es

$$\Pi=\{\mathbf{x}\in\mathbb{R}^n:\forall j\in[1..n]\ \ |\langle\mathbf{x},\mathbf{a}_j\rangle|\le 1\}\,. \tag{6}$$

Una colección de n funcionales lineales independientes, $L_1,\dots,L_n:\mathbb{R}^n\to\mathbb{R}$, también determina un paralelepípedo:

$$\Pi=\{\mathbf{x}\in\mathbb{R}^n:\forall j\in[1..n]\ \ |L_j(\mathbf{x})|\le 1\}\,.$$

Es claro que estos conjuntos son convexos, simétricos con respecto al cero y tienen medida finita y positiva. La base **recíproca** de $\mathbf{a}_1,\dots,\mathbf{a}_n$, denotada $\mathbf{a}_1^*,\dots,\mathbf{a}_n^*$ es aquella base que cumple

$$\forall i,j\in[1..n]\quad \left\langle\mathbf{a}_i,\mathbf{a}_j^*\right\rangle=\delta_{i,j},$$

donde $\delta_{i,j}$ es el símbolo de Kronecker. Si Π es el paralelepípedo generado por $\mathbf{a}_1,\dots,\mathbf{a}_n$, el **paralelepípedo recíproco** es

$$\Pi^*=\left\{\mathbf{x}\in\mathbb{R}^n:\forall j\in[1..n]\ \ \left|\left\langle\mathbf{a}_j^*,\mathbf{x}\right\rangle\right|\le 1\right\}.$$

La teoría básica de integración facilita una importante propiedad sobre los volúmenes:

$$\mathfrak{m}(\Pi)=\frac{2^n}{D},\ \mathfrak{m}(\Pi^*)=2^nD\quad\text{ con }\quad D=\det(\mathbf{a}_1|\dots|\mathbf{a}_n). \tag{7}$$

Teorema 11 (Davenport)**.** *Sean $L_1,\dots,L_n:\mathbb{R}^n\to\mathbb{R}$, $\det(L_1,\dots,L_n)=1$, formas lineales y $\lambda_1\le\dots\le\lambda_n$ los mínimos sucesivos del paralelepípedo, Π, generado por $L_1,\dots,L_n$. Sean $\rho_1,\dots,\rho_n$ reales tales que*

$$\rho_1\ge\rho_2\ge\dots\ge\rho_n>0, \tag{8}$$

$$\rho_1\lambda_1\le\rho_2\lambda_2\le\dots\le\rho_n\lambda_n, \tag{9}$$

$$\rho_1\rho_2\cdots\rho_n=1. \tag{10}$$

Entonces, existe una permutación de $L_1,\dots,L_n$ –dígase $L_1',\dots,L_n'$– tal que los mínimos sucesivos del paralelepípedo generado por $\rho_1L_1',\dots,\rho_nL_n'$, llamados $\lambda_1',\dots,\lambda_n'$, verifican

$$\forall i\in[1..n]\quad \frac{\lambda_i\rho_i}{2^n}\le\lambda_i'\le 2^{n^2}(n!)^2\lambda_i\rho_i. \tag{11}$$

Además, si $\mathbf{g}_1,\dots,\mathbf{g}_n\in\mathbb{Z}^n$ son linealmente independientes y $\mathbf{g}_j\in\lambda_j\Pi$ para $j\in[1..n]$; entonces,

$$\forall i\in[2..n]\ \ \left(\forall\mathbf{x}\in\mathbb{Z}^n\setminus\operatorname{span}\{\mathbf{g}_1,\dots,\mathbf{g}_i\}\ \ \left(\frac{\rho_i\lambda_i}{2^n}\le\max\{|\rho_1L_1'(\mathbf{x})|,|\rho_2L_2'(\mathbf{x})|,\dots,|\rho_nL_n'(\mathbf{x})|\}\right)\right). \tag{12}$$

Demostración. I. Sea $N:\mathbb{R}^n\to\mathbb{R}$ la función dada por

$$N(\mathbf{x}):=\max_{1\le i\le n}\{|L_1(\mathbf{x})|,|L_2(\mathbf{x})|,\dots,|L_n(\mathbf{x})|\}\,.$$

II. Se definen $S_0 \coloneqq \{\mathbf{0}\}$ y $\forall i \in [1..n] \qquad S_i \coloneqq \operatorname{span}\{\mathbf{g}_1, \ldots, \mathbf{g}_i\}$. Debido a que S_{n-1} es isomorfo como $\mathbb{R}$-espacio vectorial a $\mathbb{R}^{n-1}$, $S_{n-1}^* \cong \mathbb{R}^{n-1}$ y los funcionales $L_1|_{S_{n-1}}, \ldots, L_n|_{S_{n-1}}$ son dependientes. Existen, pues, $\alpha_1^{(1)}, \ldots, \alpha_{n-1}^{(1)} \in \mathbb{R}$ no todos cero tales que

$$\alpha_1^{(1)} L_1|_{S_{n-1}} + \ldots + \alpha_n^{(1)} L_n|_{S_{n-1}} \equiv 0.$$

Se ordenan las funciones $L_1, \ldots, L_n$ de modo que

$$\left|\alpha_n^{(1)}\right| = \max_{1 \le j \le n} \left|\alpha_j^{(1)}\right|.$$

Supóngase que la permutación de las formas es $L_1^1, L_2^1, \ldots, L_{n-1}^1, L_n'$. Repitiendo el argumento anterior en $L_1^1, \ldots, L_{n-1}^1$, las $n-1$ formas lineales restringidas a S_{n-2} son linealmente dependientes, por lo que existen reales $\alpha_1^{(2)}, \ldots, \alpha_{n-1}^{(2)}$ no todos cero tales que

$$\alpha_1^{(2)} L_1|_{S_{n-2}} + \ldots + \alpha_{n-1}^{(2)} L_{n-1}|_{S_{n-2}} \equiv 0.$$

Se ordenan las formas $L_1^1, \ldots, L_{n-1}^1$ de forma que se verifique

$$\left|\alpha_n^{(2)}\right| = \max_{1 \le j \le n-1} \left|\alpha_j^{(2)}\right|.$$

Se llama $L_1^2, \ldots, L_{n-2}^2, L_{n-1}'$ a la nueva permutación. Continuando de este modo, se llega a una permutación $L_1', \ldots, L_n'$.

Sean $j \in [2..n]$ y $\mathbf{x} \in S_{j-1}$, entonces

$$L_j'(\mathbf{x}) = \frac{\alpha_1^{(j)}}{\alpha_j^{(j)}} L_1'(\mathbf{x}) + \ldots + \frac{\alpha_{j-1}^{(j)}}{\alpha_j^{(j)}} L_{j-1}'(\mathbf{x}) \implies \left|L_j'(\mathbf{x})\right| \le \sum_{i=1}^{j-1} |L_i'(\mathbf{x})|$$

$$\implies \sum_{i=1}^{j} |L_i'(\mathbf{x})| \le 2 \sum_{i=1}^{j-1} |L_i'(\mathbf{x})|.$$

Un uso recursivo de la desigualdad anterior desemboca en

$$\forall j \in [2..n] \quad \forall \mathbf{x} \in S_{j-1} \quad \sum_{i=1}^{j} |L_i'(\mathbf{x})| \ge \frac{1}{2} \sum_{i=1}^{j+1} |L_i'(\mathbf{x})| \ge \ldots \ge \frac{1}{2^{n-j}} \sum_{i=1}^{n} |L_i'(\mathbf{x})|. \tag{13}$$

III. Sea $\mathbf{x} \in \mathbb{R}^n$ un punto entero distinto de cero y supóngase que para $i \in [2..n]$ fija $\mathbf{x} \notin S_{i-1}$. Existe, entonces, $j \in [1..n]$ tal que $\mathbf{x} \in S_j \setminus S_{j-1}$; de hecho, $j \ge i$. Esto implica que $N(\mathbf{x}) \ge \lambda_j$. La desigualdad no es inmediata, pero no es difícil. Primero, en el caso extremo, $N(\mathbf{x}) < \lambda_1$ y $\mathbf{x} = \mathbf{0} \in S_{j-1}$ (por la definición de λ_1 y por $\mathbf{x} \in \mathbb{Z}^n$), contradicción. La imposibilidad de $N(\mathbf{x}) < \lambda_1$ fuerza a la existencia de algún $J \in [1..n]$ para el cual

$$\lambda_{J-1} \le N(\mathbf{x}) < \lambda_J \le \lambda_j.$$

Esto significa que $\mathbf{g}_1, \ldots, \mathbf{g}_{J-1}, \mathbf{x} \in N(\mathbf{x})\Pi$ y, debido a que $N(x)$ es estrictamente menor que λ_J, existen $\beta_1, \ldots, \beta_J \in \mathbb{R}$ tales que

$$\beta_1 \mathbf{g}_1 + \ldots + \beta_{J-1} \mathbf{g}_{J-1} + \beta_J \mathbf{x} = \mathbf{0}.$$

La independencia lineal de $\mathbf{g}_1, \ldots, \mathbf{g}_{J-1}$ se traduce en $\beta_J \neq 0$ conduciendo al absurdo $\mathbf{x} \in S_{J-1} \subseteq S_{j-1}$. Por lo tanto, $N(x) \geq \lambda_j$; luego,

$$\begin{aligned}
\max\{\rho_1|L_1'(\mathbf{x})|, \ldots, \rho_n|L_n'(\mathbf{x})|\} &\geq \max\{\rho_1|L_1'(\mathbf{x})|, \ldots, \rho_n|L_j'(\mathbf{x})|\} \\
&\geq \rho_j \max_{1\leq k\leq j} |L_k'(\mathbf{x})| && \text{(por (8))} \\
&= \frac{\rho_j}{j} \sum_{k=1}^{j} |L_k'(\mathbf{x})| \\
&\geq \frac{\rho_j}{2^n}\frac{2^j}{j} \sum_{k=1}^{n} |L_k'(\mathbf{x})| && \text{(por (13))} \\
&\geq \frac{\rho_j}{2^n} \sum_{k=1}^{n} |L_k'(\mathbf{x})| && (\forall\, j \in \mathbb{N} \quad j \leq 2^j) \\
&\geq \frac{\rho_j}{2^n} N(\mathbf{x}) \geq \frac{\rho_j \lambda_j}{2^n} \geq \frac{\rho_i \lambda_i}{2^n} && \text{(por (9))}.
\end{aligned}$$

Los extremos son, precisamente, (12). La construcción de Π' conlleva la primera parte de (11):

$$\left(\forall i \in [1..n] \quad \lambda_i' \geq \frac{\rho_i \lambda_i}{2^n}\right) \Longrightarrow \left(\forall i \in [1..n] \quad \frac{2^n}{\rho_i \lambda_i} \geq \frac{1}{\lambda_i'}\right) \tag{14}$$

IV. Tras acudir $\rho_1 \cdots \rho_n = 1$ y al Teorema de Cambio de Variable para integrales múltiples se observa que $\mathfrak{m}(\Pi') = \mathfrak{m}(\Pi) = 2^n$; en consecuencia, por el Segundo Teorema de Minkowski,

$$\frac{1}{n!} \leq \lambda_1 \cdots \lambda_n \leq n!, \quad \frac{1}{n!} \leq \lambda_1' \cdots \lambda_n' \leq n!.$$

Juntando todas las desigualdades se concluye la segunda parte de (11):

$$\begin{aligned}
\forall i \in [1..n] \qquad \lambda_i' &\leq \frac{n!}{\lambda_1' \cdots \lambda_{i-1}' \lambda_{i+1}' \cdots \lambda_n'} && \text{(Segundo Teorema de Minkowski)} \\
&\leq \frac{2^{n(n-1)} n!}{\rho_1 \lambda_1 \cdots \rho_{i-1} \lambda_{i-1} \rho_{i+1} \lambda_{i+1} \cdots \rho_n \lambda_n} && \text{(por (14))} \\
&= \frac{2^{n(n-1)} n! \rho_i \lambda_i}{\lambda_1 \cdots \lambda_n} && \text{(por } \rho_1 \cdots \rho_n = 1) \\
&< \frac{2^{n^2} n! \rho_i \lambda_i}{\lambda_1 \cdots \lambda_n} \\
&\leq 2^{n^2} (n!)^2 \rho_i \lambda_i. && \text{(Segundo Teorema de Minkowski).}
\end{aligned}$$

□

Pruebas de los dos teoremas de Mahler

Teorema 12. *Sean $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ una base de $\mathbb{R}^n$; $\mathbf{a}_1^*, \ldots, \mathbf{a}_n^*$ la base recíproca y Π y Π^* los paralelepípedos generados por cada base. Si $\lambda_1, \ldots, \lambda_n$ y $\lambda_1^*, \ldots, \lambda_n^*$ son, respectivamente, los mínimos sucesivos de Π y Π^*, entonces*

$$\forall i \in [1..n] \quad \lambda_i^* \ll \frac{1}{\lambda_{n+1-i}} \ll \lambda_i^* \tag{15}$$

con las constantes implicadas dependientes tan sólo de n. Además, si $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{R}^n$ tienen entradas en los enteros racionales, son linealmente independientes y $\mathbf{g}_j \in \lambda_j \Pi$ para cualquier $j \in [1..n]$, entonces

$$\forall i, j \in [1..n] \qquad |\left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle| \ll \frac{1}{\lambda_j}. \tag{16}$$

en donde $\mathbf{g}_1^*, \ldots, \mathbf{g}_n^*$ es la base recíproca de $\mathbf{g}_1, \ldots, \mathbf{g}_n$.

Demostración. I. Juntando las definiciones de los objetos en juego es claro que

$$\forall i, j \in [1..n] \quad |\langle \mathbf{a}_i, \mathbf{g}_j \rangle| \leq \lambda_j. \tag{17}$$

Al tomar $E = |\det(\mathbf{g}_1 | \ldots | \mathbf{g}_n)|$ se llega, gracias a (2.3) y a que $\mathbf{g}_1, \ldots, \mathbf{g}_n$ son enteros, a la existencia de una constante dependiente sólo de n que garantiza

$$1 \leq E \ll 1. \tag{18}$$

II. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}$. Existen reales $\alpha_1, \ldots, \alpha_n$ y $\beta_1, \ldots, \beta_n$ únicos tales que

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i, \quad \mathbf{y} = \sum_{j=1}^{n} \beta_j \mathbf{a}_j^*.$$

No significa dificultad alguna la determinación de los coeficientes; por ejemplo,

$$\forall j \in [1..n] \quad \left\langle \mathbf{x}, \mathbf{a}_j^* \right\rangle = \sum_{i=1}^{n} \alpha_i \left\langle \mathbf{a}_i^*, \mathbf{a}_j \right\rangle = \sum_{i=1}^{n} \alpha_i \delta_{i,j} = \alpha_i.$$

Similarmente,

$$\forall j \in [1..n] \quad \beta_j = \left\langle \mathbf{y}, \mathbf{a}_j \right\rangle.$$

De este modo,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{a}_i^*, \sum_{j=1}^{n} \beta_j \mathbf{a}_j \right\rangle = \sum_{i,j} \alpha_i \beta_j \left\langle \mathbf{a}_i, \mathbf{a}_j^* \right\rangle = \sum_{i,j} \alpha_i \beta_j \delta_{i,j} = \sum_{i=1}^{n} \alpha_i \beta_i = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{a}_i \rangle \left\langle \mathbf{y}, \mathbf{a}_i^* \right\rangle.$$

En particular,

$$\forall k, j \in [1..n] \quad \sum_{i=1}^{n} \langle \mathbf{a}_i, \mathbf{g}_k \rangle \left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle = \left\langle \mathbf{g}_k, \mathbf{g}_j^* \right\rangle = \delta_{k,j}.$$

III. En términos matriciales, si $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n), G^t = (\mathbf{g}_1 | \ldots | \mathbf{g}_n)$, $M = AG$, entonces

$$M^{-1} = \left(\left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle \right)_{i,j=1}^{n}$$

que da, representando al cofactor (i, j) de M mediante $M_{i,j}$,

$$\forall i, j \in [1..n] \quad \left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle = \frac{M_{i,j}}{M}$$

También, $\det M = ED$ con $D = \det(\mathbf{g}_1 | \ldots \mathbf{g}_n)$. De la desigualdad (17) y la definición de cofactor se sigue que

$$\forall i, j \in [1..n] \quad |M_{i,j}| \ll \lambda_1 \cdot \lambda_{j-1} \lambda_{j+1} \cdots \lambda_n.$$

La desigualdad anterior, $\det(M) = ED$ y las ecuaciones concernientes al volumen de Π y Π^* (por (7)) llevan a

$$\forall i, j \in [1..n] \quad \left| \left\langle \mathbf{a}_i^*, \mathbf{g}_j^* \right\rangle \right| = \left| \frac{M_{i,j}}{M} \right| \ll \frac{\lambda_1 \cdots \lambda_n}{\lambda_j E D} = \frac{\lambda_1 \cdots \lambda_n}{2^n} \mathfrak{m}(\Pi) \frac{1}{\lambda_j} \frac{1}{E} \leq \frac{n!}{\lambda_j} \ll \frac{1}{\lambda_j}. \tag{19}$$

IV. Acudiendo de nuevo a la Regla de Cramer, es posible asegurar que los puntos $E\mathbf{g}_1^*$ tiene entradas distintas de cero en los enteros racionales, de donde $E\mathbf{g}_1^* \in \lambda\Pi$ si y sólo si $\lambda \geq \lambda_1^*$; así,

$$\forall i \in [1..n] \quad \lambda_1^* \leq |\langle \mathbf{a}_i, E\mathbf{g}_n^* \rangle| \ll \frac{1}{\lambda_n}.$$

Igualmente, $E\mathbf{g}_{n-1}^*$ y $E\mathbf{g}_n^*$ son vectores linealmente independientes con entradas en $\mathbb{Z}$. Por esto, para $E\mathbf{g}_{n-1}^*, E\mathbf{g}_n^* \in \lambda\Pi$ es necesario y suficiente $\lambda \geq \lambda_2^*$, que se traduce en

$$\forall i \in [1..n] \quad \lambda_2^* \leq \max\left\{|\langle \mathbf{a}_i, \mathbf{g}_{n-1}^* \rangle|, |\langle \mathbf{a}_i, \mathbf{g}_n^* \rangle|\right\} \ll \frac{1}{\lambda_{n-1}}.$$

En general, para $j \in [1..n]$ los puntos son $E\mathbf{g}_n^*, \ldots, E\mathbf{g}_{n+1-j}^*$ son linealmente independientes y sus entradas están en $\mathbb{Z}$. Por lo que $E\mathbf{g}_n^*, \ldots, E\mathbf{g}_{n+1-j}^*$ están en $\lambda\Pi$ si y sólo si $\lambda \geq \lambda_j^*$ que deviene en el lado izquierdo de (15):

$$\forall i \in [1..n] \quad \lambda_j^* \leq \max\left\{|\langle \mathbf{a}_1, E\mathbf{g}_n^* \rangle|, |\langle \mathbf{a}_1, E\mathbf{g}_{n-1}^* \rangle|, \ldots, \left|\left\langle \mathbf{a}_1, E\mathbf{g}_{n+1-j}^* \right\rangle\right|\right\} \ll \frac{1}{\lambda_{n+1-j}}.$$

Se reescribe esta desigualdad como

$$\forall j \in [1..n] \quad \lambda_j \lambda_{n+1-j}^* \ll 1.$$

Aplicando $n-1$ veces esta expresión, las relaciones sobre los volúmenes de Π y Π^* (por (7)) y el Segundo Teorema de Minkowski, en ese orden, se concluye el lado derecho de (15):

$$\forall j \in [1..n] \quad \lambda_j \lambda_{n-j+1}^* \gg \lambda_1 \cdots \lambda_n \lambda_1^* \cdots \lambda_n^* = \frac{\lambda_1 \cdots \lambda_n}{2^n} \mathfrak{m}(\Pi) \frac{\lambda_1^* \cdots \lambda_n^*}{2^n} \mathfrak{m}(\Pi^*) \gg 1.$$

□

Antes de atacar al segundo teorema de Mahler se recuerda a las convenciones en la escritura. Algunas nociones del Apéndice B son usadas. Para cualquier $p \in [1..n]$, $C(n,p)$ es

$$C(n,p) = \{X \subseteq [1..n] : |X| = p\}.$$

Si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ es una base de $\mathbb{R}^n$, se escribe para cualquier $\sigma \in C(n,p)$ con $\sigma = \{i_1 < \ldots < i_p\}$

$$A_\sigma := \mathbf{a}_{i_1} \wedge \ldots \wedge \mathbf{a}_{i_n}.$$

Esto da una base en $\mathfrak{R}_n^p$, $\{A_\sigma : \sigma \in C(n,p)\}$. El paralelepípedo generado por esta base de $\mathbb{R}_p^n$ se llama **p-ésimo pseudocompuesto** de Π y se denota $\Pi^{(p)}$. Se ordena $C(n,p)$ como sigue: primero, si $\lambda_1, \ldots, \lambda_n$ son los mínimos sucesivos de Π,

$$\forall \tau \in C(n,p) \quad \lambda_\tau := \prod_{i \in \tau} \lambda_i.$$

Posteriormente, a cada elemento de $C(n.p)$ se le asigna una etiqueta en $[1..l]$ tal que $\lambda_{\tau_1} \leq \lambda_{\tau_2} \leq \ldots \leq \lambda_{\tau_n}$. Finalmente, si los mínimos sucesivos de Π se realizan en $\mathbf{g}_1, \ldots, \mathbf{g}_n \in \mathbb{Z}^n$ se define, con $\tau = \{j_1 < \ldots < j_p\}$ para $\tau \in C(n,p)$,

$$\forall \tau \in C(n,p) \quad G_\tau := \mathbf{g}_{j_1} \wedge \ldots \wedge \mathbf{g}_{j_p}.$$

Teorema 13. *Sean* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *una base de* $\mathbb{R}^n$ *con determinante* 1, Π *el paralelepípedo generado por ella,* $\Pi^{(p)}$ *el p-ésimo pseudocompuesto de* Π, $\lambda_1, \ldots, \lambda_n$ *y* $\nu_1, \ldots, \nu_l$ *los mínimos sucesivos de* Π *y* $\Pi^{(p)}$ *respectivamente. Entonces, tomando a* A_σ *y* G_τ *como antes,*

$$\forall \sigma, \tau \in C(n,p) \qquad |\langle A_\sigma, G_\tau \rangle| \leq p! \lambda_\tau. \tag{20}$$

Además, existen constantes que sólo dependen de n que dan

$$\forall i \in \{1, 2, \ldots, l\} \quad \lambda_{\tau_i} \ll v_i \ll \lambda_{\tau_i}. \tag{21}$$

Demostración. La expresión (20) sigue de la Identidad de Laplace (Apéndice B, Lema 6) y la definición de determinante. Para la desigualdad derecha se observa que los vectores G_τ, $\tau \in C(n,p)$, generan a $\mathbb{R}^n_p$ y, como éste es de dimensión ℓ, son linealmente independientes. Además, cada elemento de esta colección tiene coeficientes enteros con respecto a la base canónica. Entonces, se sigue de $\lambda_{\tau_1} \leq \ldots \leq \lambda_{\tau_n}$ y (20) que para $i \in [1..l]$ arbitrario es cierto que

$$\big(\forall j \in [0..i] \quad \forall \sigma \in C(n,p) \qquad |\langle A_\sigma, G_\tau \rangle| \leq p! \lambda_{\tau_j}\big) \implies \big(\forall j \in [0..i] \quad G_{\tau_j} \in p! \lambda_{\tau_i} \Pi\big).$$

Estas anotaciones y la definición de mínimo sucesivo se traducen en $\nu_i \leq p! \lambda_{\tau_i}$, o bien, $\nu_i \ll \lambda_{\tau_i}$.

Por otro lado, para la desigualdad izquierda, nótese que, por el Segundo Teorema de Minkowski aplicado a Π y el significado de λ_{τ_i}, hay constantes dependientes de n que garantizan

$$\prod_{i=1}^{l} \lambda_{\tau_i} = \left(\prod_{i=1}^{n} \lambda_i \right)^t \ll 1 \quad \text{con} \quad t = \binom{n-1}{p-1}.$$

Aplicando ahora el Segundo Teorema de Minkowski a $\Pi^{(p)}$ se llega, usando constantes que están sujetas a n, a

$$1 \ll \nu_1 \cdots \nu_l \ll 1.$$

Sea $i \in [1..l]$. Se deduce

$$1 \ll \prod_{j=1}^{l} \nu_j = \nu_i \prod_{\substack{j=1 \\ j \neq i}}^{l} \nu_i \ll \nu_i \frac{\prod_{j=1}^{l} \lambda_{\tau_j}}{\lambda_{\tau_i}} \ll \frac{\nu_i}{\lambda_{\tau_i}} \implies \lambda_{\tau_i} \ll \nu_i. \qquad \square$$

Apéndice D
Wronskianos Generalizados

Wronskiano Generalizado

El wronskiano de n funciones suaves de los reales en los reales es utilizado en la teoría de ecuaciones diferenciales ordinarias. Axel Thue introdujo en su trabajo el uso de wronskianos clásicos y Klaus Roth, el de wronskianos generalizados. El resultado principal da condiciones necesarias y suficientes para la independencia lineal de polinomios en varias variables con coeficientes reales.

Definición 1. *Sean Δ_0, Δ_1, ..., Δ_{l-1} operadores diferenciales de orden a lo más $0, 1, \ldots, l-1$, respectivamente. Si $R^{(k)} \in \mathcal{C}^{l-1}(\mathbb{R}^m; \mathbb{R})$, $k \in [0..l-1]$, un **wronskiano generalizado** de $R^{(1)}, , \ldots, R^{(l)}$ es*

$$G(R^{(1)}, R^{(2)}, \ldots, R^{(m)}; \mathbf{x}) := \det(\Delta_j R^{(k)}(\mathbf{x}))\ j, k = 0, 1, \ldots, m.$$

Supóngase que $f_1, \ldots, f_n$ son polinomios en una variable. Entonces, sólo existe un wronskiano generalizado que no es trivialmente cero, G, para el cual

$$G(f_1, f_2, \ldots, f_n; x) = \frac{1}{0!1!\cdots(n-1)!}\mathfrak{W}(f_1, f_2, \ldots, f_n)(x)$$

con $\mathfrak{W}(f_1, f_2, \ldots, f_n)(x)$ el wronskiano usual de $f_1, f_2, \ldots, f_m$. La igualdad se sigue inmediatamente de las propiedades básicas de los determinantes. Los wronskianos generalizados se usarán para determinar independencia lineal y para ello se verá si se anulan o no idénticamente. De este modo, la constante multiplicativa en el caso univariado es irrelevante y se puede identificar $G(x)$ con $\mathfrak{W}(x)$.

Sean $f_1, \ldots, f_n \in \mathcal{C}^{n-1}(\mathbb{R}; \mathbb{R})$ linealmente dependientes, entonces $\mathfrak{W}(f_1, \ldots, f_n)(x) \equiv 0$. El recíproco en general es falso. Por ejemplo, sean $f, g : \mathbb{R} \to \mathbb{R}$ las funciones linealmente independientes dadas por

$$f(x) := \begin{cases} 0 & \text{si } x \leq 0 \\ x^2 & \text{si } x > 0 \end{cases},\ g(x) := \begin{cases} x^2 & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases} \implies \mathfrak{W}(f, g)(x) \equiv 0.$$

Sin embargo, al restringirse a funciones polinomiales el resultado es cierto.

Lema 2. *Sean $f_1, f_2, \ldots, f_n$ polinomios de $\mathbb{R}$ en $\mathbb{R}$, éstos son linealmente dependientes si y sólo si su wronskiano es idénticamente cero.*

Demostración. Necesidad. Es trivial.

Suficiencia. Se trata primero el caso $n = 2$. Sea $L \in \mathbb{R}$ tal que $t \geq L$ implica $f_1(t)f_2(t) \neq 0$. El que el wronskiano sea idénticamente cero en $\mathbb{R}$ asegura la existencia de una función $c_1 : \mathbb{R} \to \mathbb{R}$ tal que

$$c(t)f_1(t) = f_2(t),\ c(t)f_1'(t) = f_2'(t) \implies c(t) = \frac{f_2(t)}{f_1(t)} \in \mathcal{C}^\infty((L, +\infty)).$$

Derivando la primera igualdad

$$f_2'(t) = c'(t)f_1(t) + c(t)f_1'(t) \implies c'(t) = 0 \quad \forall t \in (L, +\infty).$$

Existen, entonces, una constante, c_1, y un intervalo abierto para los cuales $c_1 f_2 = f_2$. Como las funciones son polinomiales, esta igualdad local fuerza a la igualdad global. Por lo tanto, f_1 y f_2 son linealmente dependientes.

Supóngase que el lema ha sido demostrado para $k \leq n-1$ y sean $f_1, f_2, \ldots, f_n \in \mathbb{R}[x]$ tales que

$$\forall x \in \mathbb{R} \quad \mathfrak{W}(f_1, f_2, \ldots, f_n)(x) = 0.$$

Se supondrá que $f_1, f_2, \ldots, f_{n-1}$ son linealmente independientes y $f_n \not\equiv 0$, de otro modo, ya se habría terminado. Por la hipótesis de inducción, $\mathfrak{W}(f_1, \ldots, f_{n-1}) \not\equiv 0$ y, por ser un polinomio, existe $L \in \mathbb{R}$ tal que $t \geq L$ implica $\mathfrak{W}(f_1, \ldots, f_{n-1})(t) \neq 0$. Por la Regla de Cramer, es claro que las soluciones $c_1(t), c_2(t), \ldots, c_{n-1}(t)$ del siguiente sistema están en $\mathcal{C}^\infty((L, +\infty); \mathbb{R})$:

$$\begin{pmatrix} f_1(t) & f_2(t) & \cdots & f_{n-1}(t) \\ f_1'(t) & f_2'(t) & \cdots & f_{n-1}'(t) \\ \vdots & & & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_{n-1}^{(n-1)}(t) \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_{n-1}(t) \end{pmatrix} = \begin{pmatrix} f_n(t) \\ f_n'(t) \\ \vdots \\ f_n^{(n-1)}(t) \end{pmatrix} \tag{22}$$

Llámese $\mathbf{W}(t)$ a la matriz de coeficientes. Para cualquier $t > L$ se cumple

$$\begin{aligned} 0 &= \mathfrak{W}(f_1, f_2, \ldots, f_n)(t) \\ &= \begin{vmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & & \vdots & \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{vmatrix} \\ &= \begin{vmatrix} f_1(t) & f_2(t) & \cdots & f_{n-1}(t) & f_n(t) - \sum_{j=1}^{n-1} c_j(t) f_j(t) \\ f_1'(t) & f_2'(t) & \cdots & f_{n-1}'(t) & f_n'(t) - \sum_{j=1}^{n-1} c_j(t) f_j'(t) \\ \vdots & & & & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_{n-1}^{(n-1)}(t) & f_n^{(n-1)}(t) - \sum_{j=1}^{n-1} c_j(t) f_j^{(n-1)}(t) \end{vmatrix} \\ &= (-1)^n \det(\mathbf{W}(t)) \left[f_n^{(n-1)}(t) - \sum_{j=1}^{n-1} c_j(t) f_j^{(n-1)}(t) \right]. \end{aligned}$$

La última igualdad implica

$$\sum_{j=1}^{n-1} c_j(t) f_j^{(n-1)}(t) = f_n^{(n-1)}(t).$$

Derivando la k-ésima, $k \in [1..n]$, ecuación de (22) y tomando en cuenta la igualdad anterior si $k = n$ se obtiene

$$\sum_{j=1}^{n-1} c_j'(t) f_j^{(k-1)}(t) + c_j(t) f_j^{(k)}(t) = f_n^{(k)}(t) \implies \sum_{j=1}^{n-1} c_j'(t) f_j^{(k-1)}(t) = 0 \quad \forall k \in [1..n].$$

Luego, para cada $t > L$ el vector $(c_1'(t), c_2'(t), \ldots, c_{n-1}'(t))^T$ satisface un sistema lineal homogéneo cuya matriz de coeficientes es $\mathbf{W}(t)$, que implica $c_j'(t) = 0$ para toda $j \in [1..n]$. En consecuencia, las funciones c_j son constantes en $(L, +\infty)$ y, como $f_n \neq 0$, al menos una es diferente de cero. Finalmente, existe un intervalo abierto, (ξ_1, ξ_2) tal que $\sum_{j=1}^{n-1} c_j f_j(t) = f_n(t)$ para toda $t \in (\xi_1, \xi_2)$ y la igualdad se satisface en $\mathbb{R}$ y $f_n \in \operatorname{span}\{f_1, f_2, \ldots, f_{n-1}\}$. □

Supóngase ahora que $R^{(1)}, R^{(2)}, \ldots, R^{(l)} \in \mathcal{C}^{l-1}(\mathbb{R}^m; \mathbb{R})$ son linealmente dependientes. Entonces, todos los wronskianos generalizados de estas funciones serán idénticamente cero. El siguiente lema establece el recíproco.

Lema 3. *Sean $R^{(1)}, R^{(2)}, \ldots, R^{(l)} \in \mathbb{R}[x_1, \ldots, x_m]$ linealmente independientes, entonces al menos un wronskiano generalizado no es idénticamente cero.*

Demostración. La prueba consiste en convertir mediante una transformación lineal inyectiva a cada $R^{(i)}$ en un polinomio univariado. La conclusión se sigue de la independencia lineal de la imagen y el Lema 2.

Sean $R^{(1)}, R^{(2)}, \ldots, R^{(l)}$ como en la proposición y

$$k := \max_j \max_i \operatorname{grad}_{x_i}(R^{(j)}) + 1.$$

Llámense X al subespacio lineal generado por los polinomios en $x_1, x_2, \ldots, x_m$ variables de grado relativo a lo más k y T al subespacio de los polinomios en t de grado menor o igual que k^m. Sea $\Psi : X \to T$ la transformación obtenida al extender linearmente la asociación

$$\mathbf{x}^{\mathbf{i}} \mapsto t^{i_1 + i_2 k + \ldots + i_m k^{m-1}} \quad \mathbf{0} \leq \mathbf{i} \leq (k-1, \ldots, k-1).$$

Como la expansión k-aria finita de cualquier entero no negativo es única, $\mathbf{i} \neq \mathbf{j}$ implica $\Psi(\mathbf{x}^{\mathbf{i}}) \neq \Psi(\mathbf{x}^{\mathbf{j}})$. En consecuencia, el conjunto de los coeficientes de $R \in X$, $R = \sum_{\mathbf{i}} C(\mathbf{i}) x^{\mathbf{i}}$ y el formado por los de $\Psi(R) \in \mathbb{R}[t]$ coinciden; luego, $\ker \Psi = \{0\}$ y Ψ es inyectiva. Por lo tanto, $\left\{\Psi\left(R^{(j)}\right)\right\}_{j=1}^{l}$ es un conjunto linealmente independiente de polinomios.

Por la independencia lineal de $\left\{\Psi(R^{(j)})\right\}_{j=1}^{l}$ su wronskiano no es idénticamente cero:

$$\begin{aligned} 0 \not\equiv W(t) &:= \mathfrak{W}\left(\Psi(R^{(1)}), \Psi(R^{(2)}), \ldots, \Psi(R^{(l-1)}, \Psi(R^{(l)})\right) \prod_{j=0}^{l-1} \frac{1}{j!} \\ &= \det\left(\frac{1}{j!} \frac{\partial^j}{\partial t^j} R^{(s)}\left(t, t^k, \ldots, t^{k^{m-1}-1}\right)\right) \; s = 1, 2, \ldots, l; \; j = 0, 1, \ldots, l-1. \end{aligned} \tag{23}$$

Por otra parte, sean $R \in \mathbb{R}[x_1, \ldots, x_m]$ y $\phi : \mathbb{R} \to \mathbb{R}^m$, $\phi(t) := \left(t, t^k, \ldots, t^{k^{m-1}}\right)$, entonces

$$\begin{aligned} \frac{\partial (R \circ \phi)}{\partial t}(t) &= \left\langle \nabla R(\phi(t)), \left(1, k t^{k-1}, \ldots, k^{m-1} t^{k^{m-1}-1}\right)\right\rangle \\ &= \frac{\partial R}{\partial x_1}(\phi(t)) + k t^{k-1} \frac{\partial R}{\partial x_2}(\phi(t)) + \ldots + k^{m-1} t^{k^{m-1}-1} \frac{\partial R}{\partial x_m}(\phi(t)). \end{aligned}$$

O, de manera más breve,

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial x_1} + kt^{k-1}\frac{\partial}{\partial x_2} + \ldots + k^{m-1}t^{k^{m-1}-1}\frac{\partial}{\partial x_m}.$$

Aplicando recursivamente la fórmula anterior, se concluye

$$\forall j \in \mathbb{N} \quad \frac{\partial^j}{\partial t^j} = f_{1,j}\Delta_{1,j} + \ldots + f_{r_j,j}\Delta_{r_j,j}$$

en donde cada $f_{i,j}$ es un polinomio en t, $r_j = r_j(j,m) \in \mathbb{N}$ y cada operador $\Delta_{i,j}$ es un operador diferencial de orden $j-1$. Sustituyendo lo anterior en (23) y expandiendo $W(t)$ aprovechando las propiedades de los determinantes[6] se llega a

$$W(t) = g_1(t)G^{(1)}\left(t, \ldots, t^{k^{m-1}}\right) + \ldots + g_s(t)G^{(s)}\left(t, \ldots, t^{k^{m-1}}\right),$$

en donde cada $G^{(i)}\left(t, \ldots, t^{k^{m-1}}\right)$ es un wronskiano generalizado de $R^{(1)}, \ldots, R^{(l)}$. Como existe t_0 tal que $W(t_0) \neq 0$, para alguna j se satisface $G^{(j)}\left(t, \ldots, t^{k^{m-1}}\right)(t_0) \neq 0$. □

[6] Si $\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{w} \in \mathbb{R}^m$, entonces

$$\det(\mathbf{u}_1|\ldots|\mathbf{u}_i + \mathbf{w}|\ldots|\mathbf{u}_n) = \det(\mathbf{u}_1|\ldots|\mathbf{u}_i + \mathbf{w}|\ldots|\mathbf{u}_n) + \det(\mathbf{u}_1|\ldots|\mathbf{u}_i + \mathbf{w}|\ldots|\mathbf{u}_n)$$

y el que para cualquier matriz cuadrada A se cumple $\det(A) = \det(A^t)$.

Apéndice E
Lemas auxiliares para el Teorema de Schmidt

El Teorema del Subespacio de Schmidt se apoya en un sinnúmero de proposiciones auxiliares. Algunas de ellas dan ideas fundamentales sobre la construcción y otras simplemetne ayudan a sortear los problemas técnicos. En este apéndice se demuestran lemas requeridos en el Teorema de Schmidt que no entran en los otros apéndices. No necesariamente hay relación entre uno y otro.

Cuando $\alpha \in \mathbb{C}$ sea un número algebraico, $\mathfrak{N}(\alpha)$ es el producto de todos los conjugados de α en $\mathbb{C}$.

Lema 1. *Sea $1, \alpha_1, \ldots, \alpha_m$ es una base de un campo numérico y $L : \mathbb{R}^n \to \mathbb{R}$ dada por $L(\mathbf{X}) = \alpha_1 X_1 + \ldots + \alpha_m X_m$. Entonces, para cualesquiera $\mathbf{0} \neq \mathbf{q} \in \mathbb{Z}^n$ y $p \in \mathbb{Z}$ se cumple $\|\mathbf{q}\|_\infty^m |L(\mathbf{q}) - p| \gg 1$.*

Demostración. Sean $\mathbf{0} \neq \mathbf{q} \in \mathbb{Z}^n$ y $p \in \mathbb{Z}$. La atención se circunscribe cuando $|L(\mathbf{q}) - p| < 1$; de otro modo, el resultado es evidente. Por la desigualdad del triángulo se tiene $|p| \ll \|\mathbf{q}\|_\infty$; en consecuencia, cada conjugado de $L(\mathbf{q}) - p$ tiene valor absoluto $\ll \|\mathbf{q}\|_\infty$, de donde

$$|N(L(\mathbf{q}) - p)| \ll \|\mathbf{q}\|_\infty^m |L(\mathbf{q}) - p| .$$

Sea $h \in \mathbb{Z}$ tal que $h\alpha_1, \ldots, h\alpha_m$ son enteros algebraicos. Se tiene que $|N(hL(\mathbf{q}) - hp)| \geq 1$ y

$$|N(L(\mathbf{q}) - p)| \geq \frac{1}{h^{m+1}} \gg 1.$$

Por lo tanto, $|L(\mathbf{q}) - p| \, \|\mathbf{q}\|_\infty^m \gg 1$. □

Supóngase que $\alpha_1, \ldots, \alpha_n$ son reales algebraicos linealmente independientes sobre $\mathbb{Q}$ y que $L : \mathbb{R}^n \to \mathbb{R}$ está dada por $L(\mathbf{x}) = \langle (\alpha_1, \ldots, \alpha_n), \mathbf{x} \rangle$. Por el Lema 1, poniendo $d = [Q(\alpha_1, \ldots, \alpha_n) : Q]$, se tiene

$$\forall \mathbf{q} \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \quad \forall p \in \mathbb{Z} \quad \|\mathbf{q}\|_\infty^d |L(\mathbf{q}) - p| \gg 1.$$

Ahora, considérese a una transformación lineal $L : \mathbb{R}^n \to \mathbb{R}$ con coeficientes algebraicos, $L(\mathbf{x}) = \langle (\alpha_1, \ldots, \alpha_n), \mathbf{x} \rangle$. Supóngase, reordenando de ser necesario, que $\alpha_1, \ldots, \alpha_k$ son $\mathbb{Q}$-linealmente independientes y que el resto de los coeficientes está en $\mathrm{span}_{\mathbb{Q}}\{\alpha_1, \ldots, \alpha_k\}$. Defínase una colección de racionales $r_1^{(j)}, \ldots, r_k^{(j)} \in \mathbb{Q}$ tales que

$$\forall j \in [k+1..n] \quad \alpha_j = \sum_{i=1}^{k} r_i^{(j)} \alpha_i.$$

De este modo, para cada $\mathbf{x} \in \mathbb{Z}^n$,

$$L(\mathbf{x}) = \left(1 + \sum_{i=k+1}^{n} r_1^{(i)} x_i\right)\alpha_1 x_1 + \ldots + \left(1 + \sum_{i=k+1}^{n} r_k^{(i)} x_i\right)\alpha_k x_k.$$

El Lema 1 se aplica en una infinidad de formas lineales en $x_1, \ldots, x_k$, cuyos coeficientes son $\mathbb{Q}$-linealmente independientes. En principio, la constante en $\gg$ depende de cada forma; sin embargo, rastreando el origen de estas constantes (h en la demostración del Lema 1), observando que el racional que acompaña a cada α_j tiene un denominador fijo y recordando que el producto de enteros algebraicos es un entero algebraico, puede obtenerse una constante que sirva para todas las formas y así, se concluye la siguiente afirmación.

Lema 2. *Si $L : \mathbb{R}^n \to \mathbb{R}$ es lineal y tiene coeficientes algebraicos, entonces existe $d > 0$ tal que $|L(\mathbf{x})| \gg \|\mathbf{x}\|_\infty^d$ para toda $\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$.*

Lema 3. *Sean $\alpha_1, \ldots, \alpha_n$ enteros algebraicos. Para cada $P \in \mathfrak{R}$ homogéneo en $X_{h,1}, \ldots, X_{h,n}$ con $h \in [1..m]$,*

$$P(X_{1,1}, \ldots, X_{1,n}; \ldots; X_{m,1}, \ldots, X_{m,n}) = \sum c(j_{1,1}, \ldots, j_{m,n}) X_{1,1}^{j_{1,1}}, \ldots, X_{1,n}^{j_{m,n}} \in \mathfrak{R},$$

el polinomio $P^* = P^*(X_{1,2}, \ldots, X_{1,n}; \ldots; X_{m,2}, \ldots, X_{m,n})$ en $nm - m$ variables es

$$P^* = P_{\mathfrak{J}}(-\alpha_2 X_{1,2} - \ldots - \alpha_n X_{1,n}, \alpha_1 X_{1,2}, \ldots, X_{1,n}; \ldots; -\alpha_2 X_{m,2} - \ldots - \alpha_n X_{m,n}, \alpha_1 X_{m,2}, \ldots, X_{m,n}),$$

donde $\mathfrak{J} \in \mathbb{Z}^{mn}$ de entradas no negativas tiene la forma

$$\mathfrak{J} = (j_1, 0, \ldots, 0; j_2, 0, \ldots, 0; \ldots; j_m, 0, \ldots, 0).$$

Entonces, los coeficientes de P^* son una transformación lineal en los coeficientes de P y los coeficientes de estas formas, γ, son enteros en $K := \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ que cumplen con

$$\|\gamma\|_K \leq (4^n B)^{|\mathbf{r}|}, \quad B := \max\{\|\alpha_1\|_K, \ldots, \|\alpha_n\|_K\}.$$

Demostración. La única aseveración que no es suficientemente clara es la cota, las demás siguen de la forma de fabricar a P^*. La idea es disectar al polinomio P^* en partes manejables, acotar cada una y, después, unir todas.

Recordando que para $w \in \mathbb{N}$ con $w \geq n$ la ecuación $y_1 + \ldots + y_n = w$ tiene $\binom{w+n-1}{n-1}$ soluciones enteras no negativas, se afirma que P^* tiene a lo más

$$\binom{r_1 + n - 1}{n-1} \cdots \binom{r_m + n - 1}{n-1} \leq 2^{r_1+n-1} \cdots 2^{r_m+n-1} = 2^{n|\mathbf{r}|}$$

sumandos de la forma

$$\begin{aligned} S &= \pm c(\mathfrak{J}) \binom{j_{1,1}}{j_1} \cdots \binom{j_{m,1}}{j_m} \prod_{k=1}^{n} \left[(\alpha_2 X_{k,2} + \ldots + \alpha_n X_{k,n})^{j_{k,1}-j_k} (\alpha_1 X_{k,2})^{j_{k,2}} \cdots (\alpha_1 X_{k,n})^{j_{k,n}}\right], \\ &= \pm c(\mathfrak{J}) \binom{j_{1,1}}{j_1} \cdots \binom{j_{m,1}}{j_m} S_1 S_2 \cdots S_m \end{aligned}$$

donde para cada $h \in [1..m]$, llamando $j'_h = j_{h,1} - j_h$ y aplicando el Teorema del Multinomio,

$$\begin{aligned} S_h &= (\alpha_2 X_{h,2} + \ldots + \alpha_n X_{h,n})^{j_{h,1}-j_1} (\alpha_1 X_{h,2})^{j_{h,2}} \cdots (\alpha_1 X_{h,n})^{j_{h,n}} \\ &= \sum_{c_2 + \ldots + c_n = j'_h} \binom{j_{h,1} - j_h}{c_1, \ldots, c_n} \alpha_2^{c_2} \cdots \alpha_n^{c_n} \alpha_1^{j_{h,2}+\ldots+j_{h,n}} X_2^{c_2+j_{h,2}} \cdots X_n^{c_n+j_{h,n}} \end{aligned}$$

Para cualquier $h \in [1..m]$ se aplica la desigualdad submultiplicativa de $\|\cdot\|_K$ y se concluye

$$\begin{aligned}\|\alpha_2^{c_2}\cdots\alpha_n^{c_n}\alpha_1^{j_{h,2}+\ldots+j_{h,n}}\|_K &\le \|\alpha_2\|_K^{c_2}\cdots\|\alpha_n\|_K^{c_n}\|\alpha_1\|_K^{j_{h,2}+\ldots+j_{h,n}}\\ &\le B^{c_2+\ldots+c_n+j_{h,2}+\ldots+j_{h,n}} = B^{r_h-j_h} \le B^{r_h}.\end{aligned}$$

Además, por el Teorema del Multinomio,

$$\sum_{c_2+\ldots+c_n=j'_h}\binom{j_{h,1}-j_h}{c_1,\ldots,c_n} = (1+\ldots+1)^{j_{h,1}-j_h} < n^{r_h}.$$

En consecuencia, los coeficientes de S_h, α, satisfacen $\|\alpha\|_K \le (nB)^{r_h}$. Por lo que los coeficientes, β, de cada S son de la forma $\pm c(j_{1,1},\ldots,j_{m,n}\beta$ con

$$\|\beta\|_K \le \binom{j_{1,1}}{j_1}\cdots\binom{j_{m,1}}{j_m}(nB)^{|\mathbf{r}|} \le (2nB)^{|\mathbf{r}|}.$$

Volviendo a P^*, cada coeficiente es una combinación lineal de los originales, $c(j_{1,1},\ldots,j_{m,n})$, y los escalares, γ, que fungen como coeficientes en esta transformación lineal satisfacen

$$\|\gamma\|_K \le 2^{n|\mathbf{r}|}(2nB)^{|\mathbf{r}|} \le (2^{2n}nB)^{|\mathbf{r}|} = (4^n nB)^{|\mathbf{r}|}.$$

□

Lema 4. *Sean $\mathbf{r} \in \mathbb{N}^m$, $\varepsilon > 0$ y $n \in \mathbb{N}_{\ge 2}$. Entonces, el total de elementos del conjunto*

$$\left\{\mathfrak{I} \in \mathbb{N}_0^{nm} : \left(\forall h \in \{1,2,\ldots,m\} \quad \sum_{k=1}^{n} i_{h,k} = r_h\right) \;\&\; \left(\left|\sum_{h=1}^{m}\frac{i_{h,1}}{r_h} - \frac{m}{n}\right| \ge \varepsilon m\right)\right\} \tag{1}$$

con $\mathfrak{I} = (i_{1,1},\ldots,i_{1,n};i_{2,1},,\ldots,i_{2,n};\ldots;i_{m,1},\ldots,i_{m,n})$ es a lo más

$$2\binom{r_1+n-1}{r_1}\cdots\binom{r_m+n-1}{r_m}\exp\left(-\frac{\varepsilon^2 m}{4}\right).$$

Demostración. I. Son probadas dos identidades que involucran al coeficiente binomial. Para $r, n \in \mathbb{N}$ valen

$$\forall r,n \in \mathbb{N} \quad \sum_{c=0}^{r}\binom{r-c+n-2}{r-c} = \binom{r+n-1}{r}, \quad \sum_{c=0}^{r} c\binom{r-c+n-2}{r-c} = \frac{r}{n}\binom{r+n-1}{r}, \tag{2}$$

Con $R \coloneqq [0..r]^n$ se definen los conjuntos

$$\mathcal{S} \coloneqq \left\{\mathbf{x} \in R : \left(\sum_{i=1}^{n} x_i = r\right) \;\&\; (\forall i \in [1..n] \quad x_i \ge 0)\right\},$$
$$\forall i \in [1..n]\ \forall c \in [1..r] \quad \mathcal{S}_{i,c} \coloneqq \{\mathbf{x} \in \mathcal{S} : x_i = c\}.$$

Para cada $i \in [1..n]$, la colección $\{S_{i,c}\}_{c=0}^{r}$ es una partición de $\mathcal{S}$, que implica la primera igualdad

$$\sum_{c=0}^{r}\binom{r-c+n-2}{r-c} = \sum_{c=0}^{r}|\mathcal{S}_{i,c}| = \left|\bigcup_{c=0}^{r}\mathcal{S}_{i,c}\right| = |\mathcal{S}| = \binom{r+n-1}{r}.$$

También, para $c \in [0..r]$ se cumple $|\mathcal{S}_{i,c}| = |\mathcal{S}_{j,c}|$ para $i, j \in [1..n]$. Entonces, la segunda igualdad puede escribirse en términos de $|\mathcal{S}|$ y $|\mathcal{S}_{i,c}|$ para $i \in [1..n]$. Calculando

$$\begin{aligned}\sum_{c=0}^{r} c|\mathcal{S}_{1,c}| &= \sum_{c=0}^{r} c\frac{1}{n}\left(|\mathcal{S}_{1,c}| + \ldots + |\mathcal{S}_{n,c}|\right)\\ &= \frac{1}{n}\left(\sum_{c=0}^{r}\sum_{\mathbf{x}\in\mathcal{S}_1} x_1 + \ldots + \sum_{c=0}^{r}\sum_{\mathbf{x}\in\mathcal{S}_n} x_n\right)\\ &= \frac{1}{n}\left(\sum_{\mathbf{x}\in\mathcal{S}} x_1 + \ldots + \sum_{\mathbf{x}\in\mathcal{S}} x_n\right) = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{S}}(x_1 + \ldots + x_n) = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{S}} r = \frac{r}{n}|\mathcal{S}|.\end{aligned}$$

II. Por el Teorema de Taylor, cuando $x \in [-1, 1]$ se cumple $e^x \leq 1 + x + x^2$ y $1 + x \leq e^x$ para cualquier $x \in \mathbb{R}$.

III. El conjunto en (1) puede partirse naturalmente en dos subconjuntos cuya cardinalidad será M_+ y M_-. De manera más precisa,

$$M_+ := \left|\left\{\mathfrak{I} \in \mathbb{N}_0^{nm} : \left(\forall h \in \{1, 2, \ldots, m\} \quad \sum_{k=1}^{n} i_{h,k} = r_h\right) \;\&\; \left(\sum_{h=1}^{m} \frac{i_{h,1}}{r_h} - \frac{m}{n} \geq \varepsilon m\right)\right\}\right|$$

y para M_- se sustituye a $\geq \varepsilon m$ por $\leq -\varepsilon m$. Escribiendo $f_i(c) = \binom{r_i - c + n - 2}{n-2}$ se tiene

$$\begin{aligned}M_\pm \exp\left(\frac{\varepsilon^2 m}{2}\right) &\leq \sum_{\mathbf{0}\leq\mathbf{c}\leq\mathbf{r}} f_1(c_1)\cdots f_m(c_m)\exp\left(\frac{\varepsilon^2 m}{2}\right)\\ &\leq \sum_{\mathbf{0}\leq\mathbf{c}\leq\mathbf{r}} f_1(c_1)\cdots f_m(c_m)\exp\left(\pm\frac{\varepsilon}{2}\left(\left[\sum_{h=1}^{m}\frac{c_h}{r_h}\right] - \frac{m}{n}\right)\right)\\ &= \prod_{j=1}^{m}\sum_{c_j=0}^{r_j} f_j(c_j)\exp\left(\pm\frac{\varepsilon}{2}\left(\frac{c_j}{r_j} - \frac{1}{n}\right)\right).\end{aligned}$$

Ahora, para cualquier $j \in [1..n]$ se cumple

$$\begin{aligned}\sum_{c=0}^{r_j} f_j(c)\exp\left(\pm\frac{\varepsilon}{2}\left(\frac{c_j}{r_j} - \frac{1}{n}\right)\right) &\leq \sum_{c=0}^{r_j} f_j(c)\left(1 + \pm\frac{\varepsilon}{2}\left(\frac{c_j}{r_j} - \frac{1}{n}\right) + \frac{\varepsilon^2}{4}\left(\frac{c_j}{r_j} - \frac{1}{n}\right)^2\right)\\ &\leq \sum_{c=0}^{r_j} f_j(c)\left(1 + \frac{\varepsilon^2}{4}\right) \pm \frac{\varepsilon}{2r_j}\left(\sum_{c=0}^{r_j} c f_j(c) - \frac{r_j}{n}\sum_{c=0}^{r_j} f_j(c)\right)\\ &= \binom{r_j + n - 1}{r_j}\left(1 + \frac{\varepsilon^2}{4}\right)\end{aligned}$$

La primera desigualdad la da el punto II. y I. propicia la igualdad. Sustituyendo y usando la segunda desigualdad del segundo punto se concluye

$$\begin{aligned}M_\pm \exp\left(\frac{\varepsilon^2 m}{2}\right) &\leq \binom{r_1 + n - 1}{r_1}\cdots\binom{r_m + n - 1}{r_m}\left(1 + \frac{\varepsilon^2}{4}\right)^m\\ &\leq \binom{r_1 + n - 1}{r_1}\cdots\binom{r_m + n - 1}{r_m}\exp\left(\frac{\varepsilon^2 m}{4}\right).\end{aligned}$$

Dividiendo entre $\exp\left(\frac{\varepsilon^2 m}{2}\right)$, el renglón anterior da la cota buscada.

□

Epílogo

La prueba del Teorema del Subespacio de Schmidt invita a encontrar nuevos argumentos más ligeros. Enrico Bombieri y Walter Gubler exponen en su libro *Heights in Diophantine Geometry* una demostración que sigue la línea de Schmidt incorporando cambios debidos a Schlickewei y Evertse. Las modificaciones permiten expandir la fuerza del Teorema del Subespacio; no obstante, no la simplifican de un modo considerable.

Los Teoremas de Roth y Schmidt no son resultados efectivos, en el sentido de que no dan formas de calcular la cantidad de soluciones o de subespacios que existen. En esta dirección hay trabajos de Evertse, Schmidt y también del afamado matemático inglés Alan Baker.

El tercer capítulo deja varias lecciones. Una de ellas es que la aproximación diofantina es un campo vivo que sigue generando nuevos resultados. Prueba de ello es el trabajo de notables matemáticos como Adamczewski, Bugeaud, Luca, Evertse, Kristensen, entre otros. Incluso existen atractivos problemas abiertos como la famosa conjetura de Littlewood, según la cual para cualesquiera $\alpha, \beta \in \mathbb{R}$ se tiene

$$\liminf_{n\to\infty} n\|\alpha n\|\|\beta n\| = 0,$$

donde $\|\cdot\|$ es la distancia al entero más cercano. En esta línea está el trabajo de Manfred Einsiedler, Anatole Katok y Elon Lindenstrauss.

Definitivamente, el uso de nociones más poderosas podría ayudar a reducir la longitud y la pesadez de estos resultados tan importantes. Por ejemplo, podría aprovecharse que el índice de un polinomio univariado es una valuación. Queda como meta del autor e interesante ejercicio al lector investigar la utilidad de nuevas técnicas de optimización, en particular el Análisis No Suave (*Non-Smooth Analysis*) en este tipo de problemas.

Glosario de notación

La notación usada en la tesis es la usual. Si bien el significado de la mayoría de los símbolos se define a lo largo del texto, algunos son sólo son explicados en este glosario.

- $\mathbb{Z}$ es el conjunto de los enteros racionales.
- $\mathbb{R}$ es el conjunto de los números reales.
- $\mathbb{C}$ es el conjunto de los números complejos.
- $\mathbb{A}$ es el conjunto de los complejos algebraicos.
- $\mathbb{N}$ es el conjunto de los números naturales (enteros positivos).
- $\mathbb{Z}[x]$ es el conjunto de los polinomios en x con coeficientes en $\mathbb{Z}$.
- $\mathbb{Z}[x_1, \ldots, x_m]$ es el conjunto de los polinomios en $x_1, \ldots, x_n$ con coeficientes en $\mathbb{Z}$.
- En general, cuando A es un anillo, $A[x_1, \ldots, x_n]$ es anillo de polinomios en $x_1, \ldots, x_n$ con coeficientes en A.
- $\mathcal{H}(P)$ con $P \in \mathbb{R}[x_1, \ldots, x_m]$ es el máximo de los valores absolutos de los coeficientes de P, se llama la altura de P.
- $\mathbb{N}_0$ es el conjunto de los números enteros no negativos (enteros no negativos).
- $L|K$ denota a la extensión de campos $L \supseteq K$.
- $\mathbb{Q}(L)$ con $L \subseteq \mathbb{C}$ es el mínimo campo que contiene a $\mathbb{Q}$ y a L. Si $L = \{\alpha_1, \ldots, \alpha_n\}$, se escribe $\mathbb{Q}(\alpha_1, \ldots, \alpha_n)$.
- $[L : K]$ es el grado de la extensión $L|K$, que es la dimensión de L como K espacio vectorial.
- $[m..n] := \{m, m+1, m+2, \ldots, n\}$ si $m, n \in \mathbb{Z}$ y $m < n$
- $[x]$ es la parte entera de x (el máximo entero menor o igual que x).
- $\{x\}$ es la parte fraccionaria de x, $\{x\} = x - [x]$.
- Para $\mathbf{x} \in \mathbb{R}^n$ o $\mathbb{C}^n$ se escribe $\mathbf{x} = (x_1, \ldots, x_n)$.
- $\|\mathbf{x}\|_2 := \sqrt{x_1^2 + \ldots + x_n^2}$ para toda $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ o $\mathbb{C}^n$.
- $\|\mathbf{x}\|_\infty := \max\limits_{1 \le i \le n} |x_i|$ para toda $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ o $\mathbb{C}^n$.

- $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{j=1}^{n} x_j y_j$ para cualesquiera $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, es el producto interior usual en $\mathbb{R}^n$.
- $[\![x]\!]$ La distancia al entero más cercano.
- $\bigcup_{\alpha \in \mathcal{I}}^{\cdot} A_\alpha$ es la unión de conjuntos disjuntos por pares.
- $\mathfrak{M}$ es la σ-álgebra de conjuntos Lebesgue medibles.
- $\mathfrak{m}$ es la medida de Lebesgue en $\mathbb{R}^n$.
- $\mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R})$ es el conjunto de las funciones suaves de $\mathbb{R}^m$ en $\mathbb{R}$.
- $|\mathbf{r}| := \sum_{k=1}^{m} r_k$ para todo $\mathbf{r}$ multi-índice.
- $\dfrac{\partial^{\mathbf{r}} f}{\partial \mathbf{x}^{\mathbf{r}}} := \dfrac{\partial^{|\mathbf{r}|} f}{\partial x_1^{r_1} \cdots \partial x_m^{r_m}}$ para cualesquiera $\mathbf{r}$ multi-índice y $f \in \mathcal{C}^\infty(\mathbb{R}^m; \mathbb{R})$.
- $R_{\mathbf{i}} := \dfrac{1}{i_1! \cdots i_n!} \dfrac{\partial^{\mathbf{r}} R}{\partial \mathbf{x}^{\mathbf{i}}}$ para un multi-índice $\mathbf{i}$ y $R \in \mathbb{Z}[x_1, \ldots, x_n]$.
- $\mathbf{x}^{\mathbf{j}} := x_1^{j_1} \ldots x_m^{j_m}$ para cualquier $\mathbf{j}$ multi-índice y $x_1, \ldots, x_m$ variables o puntos en $\mathbb{R}^m$.
- $\mathbf{i} \leq \mathbf{r}$ si y sólo si $i_k \leq r_k$ para toda $k \in [1..m]$.
- $\dfrac{\mathbf{i}}{\mathbf{r}} := \left(\dfrac{i_1}{r_1}, \ldots, \dfrac{i_m}{r_m} \right)$ si $\mathbf{i}$ y $\mathbf{r}$ son multi-índices y $r_1 \cdots r_m \neq 0$.
- $\mathbf{i} \cdot \mathbf{r}^{-1} = \sum_{j=1}^{m} \dfrac{i_j}{r_j}$ si $\mathbf{r}$ tiene entradas positivas.
- $\mathbf{0}$ es el origen en $\mathbb{R}^m$.
- Para $n, k \in \mathbb{N}_0$ el se tendrá

$$\binom{n}{k} := \begin{cases} \frac{n!}{k!(n-k)!} & \text{si } k \leq n, \\ 0 & \text{si } n < k. \end{cases}$$

- $\binom{\mathbf{j}}{\mathbf{i}} := \binom{j_1}{i_1} \cdots \binom{j_m}{i_m}$ con $\mathbf{i}$ y $\mathbf{j}$ multi-índices.
- $\operatorname{ind}_{\mathbf{r}}(P; \boldsymbol{\beta})$ es el índice de un polinomio $P \in \mathbb{Z}[x_1, \ldots, x_n]$ en $\boldsymbol{\beta} \in \mathbb{R}^n$ con respecto al multi-índice $\mathbf{r} \in \mathbb{N}^n$. (Definición 1.3).
- $\mathbf{e}_j$ El j-ésimo vector canónico en $\mathbb{R}^n$.
- $A_{>c} := \{x \in A : x > c\}$ donde $A \subseteq \mathbb{R}$ y $c \in \mathbb{R}$.
- $C(n, p) = \{A \subseteq [1..n] : |A| = p\}$ donde $0 < p < n$ son enteros.
- $\operatorname{ind} P = \operatorname{ind}_{\mathbf{r}}(P; \mathbf{L})$ es el índice de un polinomio P con respecto al multi-índice $\mathbf{r} = (r_1, \ldots, r_n)$ y las formas lineales $\mathbf{L} = (L_1, \ldots, L_n)$. (Definición 2.5).

- $\mathbb{R}[\mathbf{X}] \coloneqq \mathbb{R}[X_{1,1}, \ldots, X_{1,n}; X_{2,1}, \ldots, X_{2,n}; \ldots; X_{m,1}, \ldots, X_{m,n}]$ es el anillo de polinomios en las variables $X_{1,1}, \ldots, X_{m,n}$ con coeficientes en $\mathbb{R}$.
- $\dfrac{\mathfrak{T}}{\mathbf{r}} \coloneqq \displaystyle\sum_{j=1}^{m} \frac{i_{j,1} + \ldots + i_{j,n}}{r_j}$, donde $\mathfrak{T} = (i_{1,1}, \ldots, i_{1,n}; \ldots; i_{m,1}, \ldots, i_{m,n}) \in \mathbb{N}^{mn}$.
- $[a_0; a_1, a_2, \ldots]$ donde $a_1, a_2, \ldots \in \mathbb{N}$ y $a_0 \in \mathbb{Z}$ es

$$[a_0; a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}.$$

Bibliografía

Libros

[AS] Allouche, J., Shallit, J. (2003). *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge: Cambridge University Press.

[Ba01] Baker, A. (1975). *Transcendental Number Theory*. Cambridge: Cambridge University Press.

[Ba02] Baker, A. (2012). *A Comprehensive Course in Number Theory*. Cambridge: Cambridge University Press.

[Bu] Bugeaud, Y. (2007). *Approximation by Algebraic Numbers*. Cambridge: Cambridge University Press.

[Ca01] Cassels, J.W.S. (2008/1965) *An Introduction to Diophantine Approximation*. Cambridge: Cambridge University Press.

[Ca02] Cassels, J.W.S. (1997/1971) *An Introduction to the Geometry of Numbers*. Berlin: Springer Verlag.

[HW] Hardy, G.H., Wright, E.M. (2008/1938) *An Introduction to the Theory of Numbers*. Oxford: Oxford University Press.

[Kh] Khinchin, A. (1961/1935) *Continued Fractions*. Chicago: The University of Chicago Press.

[La01] Lang, S. (2005) *Algebra*. New York: Springer Verlag.

[La02] Lang, S. (1995) *Introduction to Diophantine Approximation*. New York: Springer Verlag.

[Mo] Morandi, P. (1991) *Field and Galois Theory*. Berlin: Springer Verlag.

[Ne] Neukirch, J. (1999) *Algebraic Number Theory*. Berlin: Springer Verlag.

[NS] Nikishin, E. M., Sorokin, V.N. (1991) *Rational Approximations and Orthogonality*. Rhode Island: American Mathematical Society.

[Ro] Roman, S. (2005) *Field Theory*. Berlin: Springer Verlag.

[Sch] Schmidt, W.M. (1996) *Diophantine Approximation*. Berlin: Springer Verlag.

[St] Steuding, J. (2005) *Diophantine Analysis*. New York: Chapman & Hall/CRC.

[Zo] Zorich, V.(2006) *Mathematical Analysis Vol. 1*. Berlin : Springer Verlag.

Artículos

[ab] Adamczewski, B., Bugeaud, Y.(2007). *A Short Proof of the Transcendence of Thue-Morse Continued Fractions*. The American Mathematical Monthly, Vol.114 (No.6), pp. 536-540.

[bu] Bugeaud, Y. (2011). *Automatic Continued Fractions are trascendental or quadratic*. arXiv:1012.1709 [math.NT].

[ge] González Robert, G. (2013). *Generalized Thue-Morse Continued Fractions*. arxiv:1302.1900 [math.NT].

[ro] Roth, K.F. (1955). *Rational Approximations to Algebraic Numbers*. Mathemakika, Vol.2 Part 1. (No.3), pp. 1-20

[sch67] Schmidt, W. (1967). *On Simultaneous Approximations of Two Algebraic Numbers by Rationals*. Acta Mathematica (Vol. 119, No.1), pp. 27-50

[sch71] Schmidt, W. (1971). *Linear Forms with Algebraic Coefficients. I*. Journal of Number Theory (No.3), pp. 253-277

[qu] Queffélec, M. (1998). *Transcendance des fractions continues de Thue-Morse*. Journal of Number Theory, Vol.73, pp. 201-211

Notas

[ev] Evertse, J.H. (2011) *The Subspace Theorem*. arXiv:1012.1709 [math.NT].

[kr] Kristensen, S. (2007) *Diophantine Analysis*. Encontrado en http://home.imf.au.dk/sik/oldcourses.html el 13 de septiembre de 2014.