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

Tesis de Licenciatura

El Séptimo Problema de Hilbert. Antecedentes, soluciones y ampliaciones

Emiliano Battaglia

Director: Dr. Ariel Martín Salort

11/2014

1 Introducción

El teorema fundamental del álgebra nos dice en particular que si tenemos un polinomio no nulo con coeficientes enteros (o equivalentemente racionales) entonces éste tendrá una raíz en los números complejos. Uno podría preguntarse si dado un número complejo α, existe un polinomio $P \in \mathbb{Z}[X]$ (o equivalentemente $\mathbb{Q}[X]$) tal que $P(\alpha) = 0$. Si tal polinomio no existe entonces se dice que α es "trascendente" o "trascendental".

Más ampliamente, un conjunto de números $\{\alpha_1, ..., \alpha_n\}$ se dice "algebraicemente independiente sobre un cuerpo K", si no existe un polinomio no nulo $P \in K[X_1, ..., X_n]$ tal que $P(\alpha_1, ..., \alpha_n) = 0$. En caso contrario se dice "algebraicamente dependiente sobre K". De esta manera la trascendencia de un número es un caso particular de independencia algebraica en el que el conjunto considerado consiste sólo en él.

Se plantean entonces dos preguntas : ¿existen los números trascendentes? y si es que existen, ¿cómo reconocerlos?. El primero en intentar responder estas preguntas fue Euler, quien en 1748 conjeturó que si r, $p \in \mathbb{Q} - 0$ y $r^\beta = p$, entonces si β es irracional debe ser trascendente.

La respuesta por primera vez afirmativa a la pregunta de existencia de números trascendentes y al mismo tiempo un criterio para reconocer algunos de ellos, fue dada por Liouville en la década de 1840. Su criterio dice esencialmente que un número algebraico no puede ser aproximado muy bien por racionales, y consecuentemente, que si un número puede ser muy bien aproximado por números racionales debe ser trascendente. En particular lo aplicó para demostrar que dado $a \in \mathbb{N}$, $a \geq 2$, el número $\theta = \sum_{k=0}^{\infty} (-1)^k a^{-k!}$ es trascendente. Posteriormente en 1874 Cantor encontró otra forma de responder la pregunta probando que el conjunto de números algebraicos es numerable, lo cual por la no numerabilidad de los reales implica que los números trascendentes existen.

El interés sobre la cuestión de la trascendencia se centró inicialmente de modo natural en los casos particulares de los números e y π. En 1815 Fourier probó la irracionalidad de e utilizando su representación como serie de potencias para obtener un entero no nulo, y luego mostrar que éste debe ser menor a 1, obteniendo una contradicción. En 1873 Hermite probó que e es trascendente usando una estrategia análoga a la de Fourier, pero en vez de partir de la serie de potencias de e lo hizo de una identidad que involucra la función e^x (con $x \in \mathbb{R}$) y que se deduce integrando por partes sucesivamente. En 1882 Lindemann probó que π es trascendente siguiendo un razonamiento análogo al de Hermite, pero partiendo de que la algebricidad de π implicaría la de $i\pi$, y usando luego la identidad $e^{i\pi} + 1 = 0$. Una consecuencia de la trascendencia de π es la imposibilidad de la cuadratura del círculo. En 1837 Wantzel [L37] mostró que los segmentos de línea que pueden ser construidos solamente con regla y compás son precisamente aquellos cuyas longitudes pueden expresarse en términos de raíces de ecuaciones cuadráticas con coeficientes racionales, raíces de ecuaciones cuadráticas cuyos coeficientes son raíces de ecuaciones cuadráticas con coeficientes racionales, y continuando del mismo modo, pueden ser expresados en términos de números obtenidos solucionando sucesivamente ecuaciones cuadráticas. El conjunto de tales números es un subconjunto del cuerpo de los números algebraicos $\mathbb{A}$. Como se tiene que π es tracendente se sigue que es imposible construir un segmento de longitud $\sqrt{\pi}$ usando regla y compás.

Lindemann aifrmó en 1882 que se podía modificar y extender la prueba de su resultado

anterior para obtener lo que en 1884 Weierstrass lo demostró formalmente que dados $\alpha_1, ..., \alpha_n$ números algebraicos distintos, y $\beta_1, ..., \beta_n$ números algebraicos no nulos, entonces

$$\beta_1 e^{\alpha_1} + ... + \beta_n e^{\alpha_n} \neq 0.$$

En particular esto implica que cualquier número de la forma

$$a_1 e^{\alpha_1} + ... + a_n e^{\alpha_n}$$

cuando los α_j números algebraicos no nulos distintos, y los a_j números algebraicos no todos nulos, es trascendente. En particular lo son los siguientes números debido a sus respectiva expresiones en términos de la función exponencial : π , e^{α} (para α algebraico no nulo) , $sen(\alpha), cos(\alpha), tan(\alpha)$ (para α algebraico no nulo) y $log(\alpha)$ (para α algebraico distinto de 0 y de 1).

Durante el Segundo Congreso Internacional de Matematica de 1900 Hilbert expuso una lista de 23 problemas que consideraba de gran importancia para el desarrollo de la matemática. El séptimo era una conjetura que expande la naturaleza artimética de los números involucrados en la conjetura de Euler. Sean α,$\beta \in \mathbb{A}$ con $\alpha \neq 0$ y $\alpha \neq 1$ y $\beta \notin \mathbb{Q}$, entonces α^{β} es trascendente. Casos particulares son los números $e^{\pi} = i^{-2i}$ y $2^{\sqrt{2}}$.

En 1929 Gelfond prueba que e^{π} es trascendente mostrando que la hipótesis de su algebricidad junto con aproximaciones polinomiales de $e^{\pi z}$ $(z \in \mathbb{C})$ basadas en los enteros gaussianos implicaría una repressentación de función e^z de la que se deduciría que es algebraica. Para obtener esta representación hizo uso de la estrategia anterior, i.e., partir de las aproximaciones polinomiales para obtener un entero no nulo, y luego mostrar que éste debe ser menor a 1, obteniendo una contradicción.

En 1930 Kuzmin prueba que $2^{\sqrt{2}}$ es trascendente, siguiendo una estrategia muy similar a la de Gelfond, pero aproximando 2^z con la fórmula de interpolación de Lagrange, esto es en vez de usar los enteros gaussianos usa los números $\{a + b\sqrt{2} \ : a, b \in \mathbb{Z}\}$ ordenados de manera conveniente.

Finalmente en 1934, en forma independiente, Gelfond y Schneider, prueban la conjetura de Hilbert. Ambas se basan en la observación de Siegel de que todas las pruebas anteriores sobre trascendencia pueden interpretarse con alguna función auxiliar subyacente con muchos ceros (incluyendo multiplicidad) y que "no crece mucho", a partir de las cual se deriva alguna contradicción (existencia de un entero entre cero y uno, o bien cotas contradictorias para algún número involucrado) , y que no es necesario conocerla explícitamente sino que es suficiente con saber que existe y que tiene ciertas propiedades. Para probar la existencia de las funciones auxiliares respectivas, Gelfond y Schneider hacen uso de la hipótesis de algebricidad del número en cuestión y lemas de Siegel o Kronecker que aseguran la existencia de soluciones acotadas para sistemas de ecuaciones o inecuaciones lineales (y que se prueban como consecuencias del criterio de las cajas de Dirichlet).

Complementando al resultado de Gelfond y Schneider, Lang probó que dados $\beta_1, \beta_2 \in \mathbb{C}$ linealmente independientes sobre $\mathbb{Q}$, y $z_v \in \mathbb{C}\,(v = 1, 2, 3)$ también linealmente independientes sobre $\mathbb{Q}$; entonces por lo menos uno de los números

$$e^{\beta_1 z_v},\ e^{\beta_2 z_v} \qquad (v = 1, 2, 3)$$

es trascendente sobre $\mathbb{Q}$.

El resultado de Gelfond y Schneider fue posteriormente ampliado por Lang y por Baker.

En 1966, Lang prueba que dados un cuerpo numérico K y funciones meromorfas $f_1, ..., f_N$ de orden $\leq \rho$, tales que el cuerpo $K(f_1, ..., f_N)$ tiene grado de trascendencia mayor o igual a dos sobre K, y que la derivada $D = \frac{d}{dt}$ mapea el anillo $K[f_1, ..., f_N]$ en sí mismo; y dados $w_1, ..., w_m$ números complejos distintos que no están entre los polos de las f_i, tales que

$$f_i(w_v) \in K$$

para todo $i = 1, ..., N$ y $v = 1, ..., m$. Entonces $m \leq 20\rho[K : Q]$. Este resultado amplia simultáneamente los de Gelfond-Schneider y los de Hermite-Lindemann.

También 1966 Baker prueba (algo ya conjeturado por Gelfond) que dados $\alpha_1, ..., \alpha_n \in \mathbb{A} \setminus \{0\}$, si $\{\log(\alpha_1), ...,\log(\alpha_n)\}$ son linealmente independientes sobre $\mathbb{Q}$ entonces $\{1, \log(\alpha_1), ..., \log(\alpha_n)\}$ es linealmente independiente sobre $\mathbb{A}$. La prueba se basa en la construcción de una función auxiliar de varias variables complejas que amplia la función de una sola variable usada por Gelfond. La mayor dificultad está en relación a las técnicas básicas de interpolación. El trabajo en conexión a esto, hasta el momento de la prueba de Baker siempre involucró una extensión en el orden de las derivadas mientras se dejaban fijos los puntos de interpolación. Sin embargo la prueba de Baker involucra un procedimiento de extrapolación propio de su contexto, en el cual el rango de interpolación es extendido mientras el orden de las derivadas se reduce. Como consecuencias de su resultado Baker deriva la trascendencia de varias clases nuevas de números. Adicionalmente soluciona el problema de número de clase de encontrar todos los cuerpos cuadráticos imaginarios con número de clase 1 y da cotas para soluciones enteras de ciertas ecuaciones polinomiales.

Hay en la actualidad algunas conjeturas que extienden resultados anteriores. Las dos centrales son las conjeturas de Schanuel y de Baker (un caso particular de la anterior). Hay otras conjeturas adicionales relacionados con ellas, casos particulares de las cuales se han probado. La de Conjetura de Schanuel dice que dados $a_1, ..., a_p \in \mathbb{C}$ linealmente independientes sobre $\mathbb{Q}$, entonces hay por lo menos p de los $2p$ números

$$a_1, ..., a_p, e^{a_1}, ..., e^{a_p}$$

que son algebraicamente independienetes sobre $\mathbb{Q}$. La Conjetura de Baker extiende el resultado de Baker diciendo que si $\lambda_1, ..., \lambda_n$ son "logaritmos de números algebraicos no nulos" linealmente independientes sobre $\mathbb{Q}$, entonces son algebraicamente independientes sobre $\mathbb{Q}$.

En este trabajo estudiaremos el desarrollo histórico del séptimo problema de Hilbert. El trabajo está organizado de la siguiente manera: en la sección 2 estudiaremos sus antecedentes, en la sección 3 su formulación, en la secciones 4 y 5 respectivamente las soluciones iniciales de los casos particulares e^π por Gelfond y $2^{\sqrt{2}}$ por Kuzmin, en la secciones 6 y 7 respectivamente las siguientes soluciones plenas independientemente elaboradas por Gelfond y por Schneider, en la sección 8 y 9 respectivamente las posteriores ampliaciones de Baker y de Lang, y finalmente en la sección 10 algunos resultados relacionados y conjeturas abiertas. Adjuntamos dos apéndices. El primero incluye las nociones y resultados sobre números algebraicos necesarios para el desarrollo de la tesis. El segundo consiste en el material de referencia utilizado.

Contenido

2 Antecendentes del séptimo problema de Hilbert

En esta sección hacemos un repaso de los resultados históricos que motivan y contextualizan el séptimo problema de Hilbert.

2.1 Euler

Conjetura 2.1 (1748 Euler logarítmica)**.** *Sean a/b, $b/c \in \mathbb{Q}^+$ con $a/b \neq 1$ entonces el número*

$$\log_{a/b} c/d \ (= \frac{\log c/d}{\log a/b})$$

es racional o trascendente.

Esta conjetura puede formularse equivalentemente en forma exponencial del siguiente modo.

Conjetura 2.2 (1748 Euler exponencial)**.** *Sean r y $p \in \mathbb{Q} \setminus 0$ y $r^\beta = p$, entonces si β es irracional debe ser trascendente.*

Es decir un racional elevado a una potencia irracional algebraica debe ser irracional. Si se compara esta conjetura con la de Hilbert (enunciada en el siguiente capítulo) puede observarse que la última expande la naturaleza artimética de los números involucrados en la primera.

Si bien Euler conjeturó la existencia de números trascendentes, no se conocía ninguno por entonces, ni siquiera se sabía si existían o no. La primera prueba de ello fue dada por Liouville un siglo después.

2.2 Fourier

Teorema 2.3 (1815 Fourier)**.** *e es irracional.*

La importancia de esta prueba de la irracionalidad de e (que no fue la primera, ya que Euler ya lo había demostrado) radica en que su estrategia motivó la de posteriores pruebas de trascendentalidad. La idea de la misma es suponer que e es racional, luego utilizando la representación de e como serie de potencias obtener un entero no nulo, y luego mostrar que éste debe ser menor a 1, obteniendo una contradicción.

Prueba. Supongamos $e = \frac{B}{A}$ con A y B enteros positivos. Sustituyendo e en $Ae - B = 0$ por su representación en serie de potencias se obtiene

$$A\left(\sum_{k=0}^{\infty} \frac{1}{k!}\right) - B = 0 \tag{2.1}$$

Para todo entero $N \geq 1$ es posible separar la serie en un término cabeza M_N y en otro cola T_N,

$$\sum_{k=0}^{\infty} \frac{1}{k!} = \underbrace{\sum_{k=0}^{N} \frac{1}{k!}}_{M_N} + \underbrace{\sum_{k=N+1}^{\infty} \frac{1}{k!}}_{T_N} \tag{2.2}$$

Si reemplazamos esta expresión en (2.1) y sacamos $N!$ de denominador común en M_N tenemos

$$A \left(\frac{1}{N!} \underbrace{\sum_{k=0}^{N} \frac{N!}{k!}}_{M'_N} + \underbrace{\sum_{k=N+1}^{\infty} \frac{1}{k!}}_{T_N} \right) - B = 0. \tag{2.3}$$

Multiplicando (2.3) por $N!$ y reordenando los términos tenemos que para todo $N \in \mathbb{N}$

$$|AM'_N - N!B| = N!AT_N. \tag{2.4}$$

Veamos que el lado de la izquierda de (2.4) es un entero no nulo. Como para cada k tal que $0 \leq k \leq N$, la fracción $\frac{N!}{k!}$ es un entero entonces también lo es M'_N. Si la izquierda de (2.4) fuera cero tendríamos para todo N

$$e = \frac{B}{A} = \frac{M'_N}{N!} = M_N < e \tag{2.5}$$

que es absurdo. Veamos que la derecha de (2.4) debe ser menor a 1. Tenemos

$$\begin{aligned} N!AT_N &= A \sum_{k=N+1}^{\infty} \frac{N!}{k!} \\ &= A \sum_{k=1}^{\infty} \frac{1}{\prod_{h=1}^{k} (N+h)} \end{aligned} \tag{2.6}$$

Tomando $N + 1 = 2A$ obtenemos

$$N!.A.T_N = \sum_{k=1}^{\infty} \frac{A}{\prod_{h=0}^{k-1} (2A+h)} < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 \tag{2.7}$$

Obtenemos así el absurdo buscado. □

2.3 Liouville

Teorema 2.4 (1844-1851 Liouville)**.** *Sea $\alpha \in \mathbb{A}$, $\deg \alpha = n > 0$. Si $\alpha \neq p/q$, entonces*

$$|\alpha - p/q| > c(\alpha) q^{-n}, \qquad \textit{con } c(\alpha) = (1 + |\alpha|)^{1-n} n^{-1} L(\alpha)^{-1}$$

Donde $L(\alpha)$ es la longitud de α .

Prueba. Si $n = 1$. Entonces $\alpha \in \mathbb{Z}$ así que

$$\frac{1}{q} \leq \frac{|\alpha q - p|}{q} = |\alpha - p/q|$$

Si $n > 1$. Si $|\alpha - p/q| \geq 1$ la desigualdad del enunciado es cierta para todo $c \in (0, 1)$. Supongamos que $|\alpha - p/q| < 1$, en cuyo caso $|\alpha| + 1 > |p/q|$. Sea $P(z) = a_n z^n + ... + a_0$ el polinomio minimal de α . Entonces

$$|P(p/q)| = |P(\alpha) - P(p/q)| = |\int_{p/q}^{\alpha} P'(z)\, dz| \leq |\alpha - p/q| \max_{\theta \in [0,1]} |P'(\alpha + \theta(p/q - \alpha))|$$

$$\leq |\alpha - p/q| \sum_{k=1}^{n} k|a_k|(|\alpha| + 1)^{k-1} < |\alpha - p/q| n L(\alpha)(|\alpha| + 1)^{n-1}.$$

Entonces

$$|\alpha - p/q| > c(\alpha)|q^n P(p/q)|q^{-n}.$$

Como $n \geq 2$, todas las raíces de P(z) son irracionales, y entonces $A = q^n P(p/q) \neq 0$. Pero como $A \in \mathbb{Z}$ tenemos que $|A| \geq 1$ y se obtiene la desigualdad buscada. □

Liouville usó este resultado para construir los primeros ejemplos de números trascendentes. Sea $a, m \in \mathbb{N}$, $a \geq 2$, $t \in \mathbb{N}_0$,

$$\theta = \sum_{k=0}^{\infty} (-1)^k a^{-k!}, \qquad \frac{P_{m+t}}{Q_{m+t}} = \sum_{k=0}^{m+t} (-1)^k a^{-k!}, \qquad Q_{m+t} = a^{(m+t)!}$$

Entonces

$$0 < |\theta - \frac{P_{m+t}}{Q_{m+t}}| < Q_{m+t+1}^{-1} = Q_{m+t}^{-m-1-t} \leq Q_{m+t}^{-m-1}.$$

Entonces por (2.4) θ no puede ser algebraico de orden menor o igual a m. De la arbitrariedad de m se deduce la trascendencia de θ.

2.4 Hermite

Teorema 2.5 (1873 Hermite)**.** *e es trascendental.*

Para probarlo se necesitan previamente 2 lemas.

Lema 2.6. *Si $g(x) \in \mathbb{Z}[X]$, entonces para todo $k \in \mathbb{N}$ todos los coeficientes de la k-ésima derivada $g^{(k)}(x)$ son divisibles por $k!$.*

Prueba. Como diferenciar es una operación lineal, alcanza con probar el lema para el polinomio x^s con $s > 0$. Pero la $k-$ésima derivada de x^s es cero si $k > s$, y es igual a $k!\binom{s}{k} x^{s-k}$ si $1 \leq k \leq s$, en donde $\binom{s}{k} \in \mathbb{Z}$. □

Lema 2.7 (Identidad de Hermite)**.** *Sea $f(x)$ un polinomio de grado v con coeficientes reales y*

$$F(x) = f(x) + f'(x) + ... + f^{(v)}(x) \tag{2.8}$$

Entonces

$$e^x \int_0^x f(t)e^{-t}\,dt = F(0)e^x - F(x) \tag{2.9}$$

donde x es un número real o complejo.

Prueba. Integrando por partes obtenemos la relación

$$\int_0^x f(t)e^{-t}\,dt = f(0) - f(x)e^{-x} + \int_0^x f'(t)e^{-t}\,dt \tag{2.10}$$

Si repetimos este proceso $v+1$ veces obtenemos la igualdad

$$\int_0^x f(t)e^{-t}\,dt = F(0) - F(x)e^{-x}$$

de donde se sigue (2.9) □

Probamos a continuación que e es trascendente.

Prueba de 2.5. Supongamos lo contrario, que e es algebraico de orden m. Entonces

$$a_m e^m + ... + a_1 e + a_0 = 0, \qquad a_0 \neq 0, \qquad a_k \in \mathbb{Z}, \qquad k = 0, 1, ..., m \tag{2.11}$$

Si ponemos $x = k$ en (2.9), donde $k = 0, 1, ..., m$, obtenemos

$$e^k \int_0^k f(t)e^{-t}\,dt = F(0)e^k - F(k) \tag{2.12}$$

Multiplicamos ambos lados de (2.12) por a_k, y luego sumamos las ecuaciones resultantes para $k = 0, 1, ..., m$. Usando (2.11) obtenemos que

$$-\sum_{k=0}^{m} a_k F(k) = \sum_{k=0}^{m} a_k e^k \int_0^k f(t)e^{-t}\,dt \tag{2.13}$$

La igualdad (2.13) es cierta para todo polinomio $f(x)$ con coeficientes reales. En particular tomamos

$$f(x) = \frac{1}{(n-1)!} x^{n-1}(x-1)^n ...(x-m)^n, \tag{2.14}$$

donde n es un natural grande que satisface ciertas condiciones que elegiremos luego.

Mostraremos que el lado izquierdo de (2.13) será un entero racional no nulo mientras que el lado derecho tendrá valor abosoluto menor a 1, contradicción que probará el resultado buscado.

El polinomio $f(x)$ en (2.14) tiene 0 como raíz de multiplicidad $n-1$ y a $1,...,m$ como raíces de multiplicidad n. Entonces

$$f^{(l)}(0) = 0, \qquad l = 0, 1, ..., n-2 \tag{2.15}$$

$$f^{(n-1)}(0) = (-1)^{mn}(m!)^n \tag{2.16}$$

$$f^{(l)}(k) = 0, \qquad l = 0, 1, ..., n-1; \qquad k = 1, ..., m \tag{2.17}$$

Por el Lema 2.6, la derivada $l-$ésima de $x^{n-1}(x-1)^n...(x-m)^n$ tiene coeficientes que son divisibles por $l!$. Esto implica que para $l \geq n$ los coeficientes de $f^{(l)}(x)$ son divisibles por n. Entonces de (2.8),(2.15) y (2.16) deducimos que

$$F(0) = \sum_{l=n-1}^{(m+1)n-1} f^{(l)}(0) = (-1)^{mn}(m!)^n + nA, \qquad A \in \mathbb{Z} \tag{2.18}$$

y de (2.17) obtenemos que

$$F(k) = \sum_{l=n}^{(m+1)n-1} f^{(l)}(k) = nB_k, \qquad B_k \in \mathbb{Z}, \qquad k = 1, ...m \tag{2.19}$$

Elegimos n cualquier entero que satisface

$$(n, m!) = 1, \qquad n > |a_0| \tag{2.20}$$

Se sigue de (2.18) y (2.19) que todos los términos en el lado izquierdo de (2.13) son enteros. Aquí las condiciones (2.11) y (2.20) junto con (2.18) implican que $a_0F(0)$ no es divisible por n. Pero todos los otros términos $a_kF(k)$ son divisibles por n, lo que implica que el lado izquierdo de (2.13) es un entenero no nulo, en particular

$$\left| \sum_{k=0}^{m} a_k F(k) \right| \geq 1 \tag{2.21}$$

A continuación encontraremos una cota superior para el lado derecho de (2.13). En el intervalo $0 \leq x \leq m$ cada uno de los factores $x-k$ en (2.14), $0 \leq k \leq m$ está acotado por m. Entonces,

$$|f(x)| \leq \frac{m^{(m+1)n-1}}{(n-1)!}, \qquad 0 \leq x \leq m$$

y

$$\begin{aligned} \left| \sum_{k=0}^{m} a_k \int_0^k f(t)e^{k-t}\,dt \right| &\leq \frac{m^{(m+1)n-1}}{(n-1)!} \sum_{k=0}^{m} |a_k| \int_0^k e^{k-t}\,dt \\ &< \frac{m^{(m+1)n}}{(n-1)!} e^m \sum_{k=0}^{m} |a_k| = C_0 \frac{C^n}{(n-1)!} \end{aligned} \tag{2.22}$$

donde las constantes C_0 y C no dependen de n.

De (2.13), (2.21) y (2.22) obtenemos la desigualdad

$$1 \leq \left| \sum_{k=0}^{m} a_k F(k) \right| < C_0 \frac{C^n}{(n-1)!},$$

la cual es imposible cuando n es grande ya que el lado derecho tiende a cero. □

Notemos que para todo n las fracciones $\frac{F(k)}{F(0)}$ con $k = 0, 1, ..., m$ son aproximaciones simultáneas de los valores e^k con $k = 0, 1, ..., m$, ya que de (2.18) y (2.19) se deduce que $F(k)$ con $k = 0, 1, ..., m$ son enteros, y que si estimamos el lado derecho de (2.12) de la misma forma que lo hicimos para el lado derecho de (2.13), vemos que tiende a cero cuando n tiende a infinito.

Considerando esto podemos ver que la idea de la prueba ha sido similar a la de (2.3), esto es, obtener un entero no nulo y mostrar que debe ser menor a 1, pero aquí mediante la construcción de una secuencia de aproximaciones simultáneas de potencias de e obtenida usando la identidad de Hermite.

2.5 Cantor

Teorema 2.8 (1874 Cantor)**.** *El conjunto de números algebraicos es numerable.*

Como $\mathbb{R}$ es no numerable esto en particular implica que los números trascendentes existen.

Prueba. Sea

$$P(n) = \{f(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{Z}[x] : 1 \leq \sum_{j=0}^{n} |a_j| \leq n\} \tag{2.23}$$

Notar que $P(n)$ es finito, y que todo polinomio no nulo en $\mathbb{Z}[x]$ pertenece a algún P(n). Considerando la sucesión de raíces de $P(1), P(2), ...$ (donde en el posición m se consideran las raíces de $P(h)$ que no son raíces de algún $P(j)$ con $j < h$) se ve que $\mathbb{A}$ es numerable. □

2.6 Lindemann

Teorema 2.9 (1882 Lindemann)**.** *π es trascendente.*

Prueba. La prueba siguiente se basa en la identidad $e^{i\pi} + 1 = 0$ y la Identidad de Hermite (2.7). Supongamos que al afirmación es falsa, es decir que π es algebraico. Entonces $\gamma = \pi i$ también será algebraico. Sea $v = \deg \gamma$ y $\gamma = \gamma_1, ..., \gamma_v$ los conjugados de γ. Como $e^{\gamma} + 1 = 0$ tenemos que

$$\prod_{i=1}^{v} (1 + e^{\gamma_i}) = 0. \tag{2.24}$$

Expandiendo este producto obtenemos

$$\prod_{i=1}^{v}(1+e^{\gamma_i}) = \sum_{\epsilon_1=0}^{1}...\sum_{\epsilon_v=0}^{1} e^{\epsilon_1\gamma_1+...+\epsilon_v\gamma_v} = 0. \tag{2.25}$$

Los exponentes dentro de la suma múltiple en (2.25) incluyen algunos que son no nulos (como $\epsilon_1 = 1, \epsilon_2 = ... = \epsilon_v = 0$) y otros que son cero ($\epsilon_1 = \epsilon_2 = ... = \epsilon_v = 0$). Supongamos que hay exactamente m exponentes no nulos y $a = 2^v - m$ nulos, $a \geq 1$. Luego si notamos $\alpha_1, ..., \alpha_m$ a los exponentes no nulos, podemos reescribir (2.25) de la siguiente forma

$$a + e^{\alpha_1} + ... + e^{\alpha_m} = 0, \qquad a \geq 1. \tag{2.26}$$

Veremos a continuación que los números $\alpha_1, ..., \alpha_m$ son el conjunto de raíces de un polinomio $\psi(x) \in \mathbb{Z}[x]$ de grado m. Para ver esto observemos que el polinomio

$$\varphi(x) = \prod_{\epsilon_1=0}^{1}...\prod_{\epsilon_v=0}^{1}[x - (\epsilon_1\gamma_1 + ... + \epsilon_v\gamma_v)] \tag{2.27}$$

considerado como polinomio en $\gamma_1, ..., \gamma_v$ con coeficientes en $\mathbb{Z}[x]$, es simétrico en $\gamma_1, ..., \gamma_v$. Entonces por el Lema 11.4 $\varphi(x) \in \mathbb{Q}[X]$. Las raíces del polinomio $\varphi(x)$ de grado 2^v son $\alpha_1, ..., \alpha_m$ y 0 con multiplicidad a. Entonces el polinomio $x^{-a}\varphi(x) \in \mathbb{Q}[X]$ de grado m tiene precisamente como raíces $\alpha_1, ..., \alpha_m$. Sea $r \in \mathbb{N}$ el menor denominador común de los coeficientes de este polinomio, entonces el polinomio

$$\psi(x) = \frac{r}{x^a}\varphi(x) = b_m x^m + ... + b_1 x + b_0 \in \mathbb{Z}[x] \qquad b_m > 0\, b_0 \neq 0 \tag{2.28}$$

también tiene precisamente como raíces a $\alpha_1, ..., \alpha_m$.

En la identidad de Hermite (2.7) ponemos sucesivamente $x = \alpha_1, ..., \alpha_m$, sumamos las ecuaciones resultantes y usamos (2.26) para obtener

$$-aF(0) - \sum_{k=1}^{m} F(a_k) = \sum_{k=1}^{m} e^{\alpha_k}\int_0^{\alpha_k} f(t)e^{-t}\,dt. \tag{2.29}$$

El resto de la prueba es similar a la de la trascendencia de e. En (2.29) tomamos la función

$$f(x) = \frac{1}{(n-1)!}b_m^{mn-1}x^{n-1}\psi^n(x) = \frac{1}{(n-1)!}b_m^{(m+1)n-1}x^{n-1}(x-\alpha_1)^n...(x-\alpha_m)^n \tag{2.30}$$

donde $n \in \mathbb{N}$ es suficientemente grande. Mostraremos a continuación que con esta elección de $f(x)$ la igualdad (2.29) lleva a una contradicción.

Como en el caso de la prueba de la trascendencia de e 2.5, obtenemos igualdades similares a (2.15), (2.16) y (2.18)

$$f^{(l)}(0) = 0, \qquad l = 0, ..., n-2 \qquad f^{(n-1)}(0) = b_m^{mn-1}b_0^n \tag{2.31}$$

$$F(0) = \sum_{l=n-1}^{(m+1)n-1} f^{(l)}(0) = b_m^{mn-1}b_0^n + nA \qquad A \in \mathbb{Z}.$$

Como α_k es una raíces de $f(x)$ de multiplicidad n, obtenemos ecuaciones similares a (2.17)

$$f^{(l)}(\alpha_k) = 0, \qquad l = 0, 1, ..., n-1; \qquad k = 1, ..., m \tag{2.32}$$

y por el Lema 2.6 la $l-$ésima derivada de $x^{n-1}\psi^n(x)$ tiene coeficientes enteros todos divisibles por $n!$. Entonces para $l > n$ los coeficientes de $f^{(l)}(x)$ son enteros divisibles por $b_m^{mn-1}n$. De (2.32) obtenemos

$$F(\alpha_k) = \sum_{l=n}^{(m+1)n-1} f^{(l)}(\alpha_k) = nb_m^{mn-1}\Phi(\alpha_k) \qquad k = 1, ..., m \qquad \Phi(z) \in \mathbb{Z}[z]. \tag{2.33}$$

Obtenemos que los números $\beta_k = b_m\alpha_k$, $k = 1, ..., m$ son enteros algebraicos y además todas las raíces de un polinomio de grado m en $\mathbb{Z}[x]$ con coeficiente de mayor grado 1. Más aún

$$b_n^{mn-1}\Phi(\alpha_k) = H(\beta_k) \qquad H(x) \in \mathbb{Z}[X] \tag{2.34}$$

Entonces por el Lema 11.4 se tiene

$$\sum_{k=1}^{m} b_m^{mn-1}\Phi(\alpha_k) = \sum_{k=1}^{m} H(\beta_k) = B \qquad B \in \mathbb{Z} \tag{2.35}$$

De (2.31),(2.33) y (2.35) obtenemos que

$$aF(0) + \sum_{k=1}^{m} F(\alpha_k) = ab_0^n b_m^{mn-1} + n(aA + B) \tag{2.36}$$

Elegimos el valor de n como un número natural que satisface las condiciones

$$(n, b_0 b_m) = 1 \qquad n > a \tag{2.37}$$

Entonces el lado derecho de (2.36) es un entero no divisible por n, y por lo tanto no es cero. En consecuencia

$$\left| aF(0) + \sum_{k=1}^{m} F(\alpha_k) \right| \geq 1. \tag{2.38}$$

Encontraremos a continuación una cota superior para el laado derecho de (2.29). Supongamos que todos los puntos $\alpha_1, ..., \alpha_m$ están contenidos en el círculo $|x| \leq R$. Denotemos

$$C := \max_{|x| \leq R} |b_m^m \psi(x)| \tag{2.39}$$

donde C es independiente de n. Entonces

$$\max_{|x| \leq R} |f(x)| \leq \frac{R^{n-1}C^n}{(n-1)!} \tag{2.40}$$

Obtenemos así que existe un n_0 tal que para todo $n \geq n_0$ satisface (2.37) y que

$$\left| \sum_{k=1}^{m} e^{\alpha_k} \int_0^{\alpha_k} f(x) e^{-x}\, dx \right| \leq \sum_{k=1}^{m} \left| \int_0^{\alpha_k} |f(x)||e^{\alpha_k - x}|\, dx \right| \tag{2.41}$$

$$\leq \frac{R^{n-1} e^R C^n}{(n-1)!} \sum_{k=1}^{m} \left| \int_0^{\alpha_k} dx \right| \leq m e^R \frac{(RC)^n}{(n-1)!} < 1. \tag{2.42}$$

Las desigualdades (2.38) y (2.41), junto con (2.29) llevan a la contradicción $1 < 1$. □

Una consecuencia de la trascendencia de π es la imposibilidad de la cuadratura del círculo. En 1837 Wantzel mostró que los segmentos de línea que pueden ser construidos solamente con regla y compás son precisamente aquellos cuyas longitudes pueden expresarse en términos de raíces de polinomios de grado 2^n con coeficientes en $\mathbb{Q}$. El conjunto de tales números es un subconjunto del cuerpo de los números algebraicos $\mathbb{A}$. Como se tiene que π es tracendente se sigue que es imposible construir un segmento de longitud $\sqrt{\pi}$ usando regla y compás.

2.7 Lindemann-Weierstrass

Lindemann aifrmó que se podía modificar y extender la prueba del resultado anterior para obtener el siguiente resultado. Weierstrass lo demostró formalmente.

Teorema 2.10 (1884 Lindemann-Weierstrass)**.** *Sean $\alpha_1, ..., \alpha_n$ números algebraicos distintos, y $\beta_1, ..., \beta_n$ números algebraicos no nulos. Entonces*

$$\beta_1 e^{\alpha_1} + ... + \beta_n e^{\alpha_n} \neq 0$$

.

En particular esto implica que cualquier número de la forma

$$a_1 e^{\alpha_1} + ... + a_h e^{\alpha_h}$$

con los α_j números algebraicos no nulos distintos, y los a_j algebraicos no todos nulos, es trascendente. En particular lo son los siguientes números debido a sus respectiva expresiones en términos de la función exponencial:
1) π
2) e^α para α algebraico no nulo
3) $sen(\alpha), cos(\alpha), tan(\alpha)$ para α algebraico no nulo
4) $log(\alpha)$ para α algebraico distinto de 0 y de 1

Prueba de 2.10. Supongamos que el enunciado es falso, es decir que

$$\beta_1 e^{\alpha_1} + ... + \beta_n e^{\alpha_n} = 0. \tag{2.43}$$

Reordenando podemos suponer que

$$|\alpha_1| = \max_{1 \leq j \leq n} |\alpha_j|. \tag{2.44}$$

Sea

$$F_1(u_1, ..., u_n, v_1, ..., v_n) = u_1v_1 + u_2v_2 + ... + u_nv_n \tag{2.45}$$

y sea

$$F_2(u_2, ..., u_n, v_1, ..., v_n) = \prod_{u_1} F_1(u_1, ..., u_n, v_1, ..., v_n), \tag{2.46}$$

donde u_1 recorre todos los conjugados de β_1. En particular por el Lema 11.4, F_2 tiene coeficientes racionales, el coeficiente de la potencia más alta de v_1 es un número racional no nulo y

$$F_2(\beta_2, ..., \beta_n, e^{\alpha_1}, ..., e^{\alpha_n}) = 0. \tag{2.47}$$

Observemos también que F_1 y F_2 son polinomios homogéneos en las variables $v_1, ..., v_n$. Sea

$$F_3(u_3, ..., u_n, v_1, ..., v_n) = \prod_{u_2} F_2(u_2, ..., u_n, v_1, ..., v_n), \tag{2.48}$$

donde u_2 recorre todos los conjugados de β_2. En particular por el Lema 11.4, F_3 tiene coeficientes racionales, el coeficiente de la potencia más alta de v_1 es un número racional no nulo, y

$$F_3(\beta_3, ..., \beta_n, e^{\alpha_1}, ..., e^{\alpha_n}) = 0. \tag{2.49}$$

También tenemos que F_3 es un polinomio homogéneo en $v_1, ..., v_n$. Continuando de esta forma y limpiando denominadores (y combinando términos), terminamos con una expresión de la forma (2.43) con posiblemente un nuevo valor de n y con cada β_j siendo un entero racional. Veamos también que los β_j no son todos 0. Observemos que hay un N tal que F_{n+1} es un polinomio homogéneo en $v_1, ..., v_n$ de grado N. Cada uno de los términos de $F_{n+1}(e^{\alpha_1}, ..., e^{\alpha_n})$ es de la forma $ce^{\sum_{j=1}^n a_j\alpha_j}$ donde los a_j son naturales cuya suma da N. El coeficiente de $e^{N\alpha_1}$ es, en particular, no nulo. Aún más, la condición (2.44) implica que con los a_j como antes $\sum_{j=1}^n a_j\alpha_j = N\alpha_1$ si y solo si $a_1 = N$ y $a_j = 0$ en otro caso (como puede mostrarse a partir de la desigualdad triangular). En conclusión podemos suponer desde el comienzo que los β_j son enteros racionales.

A continuación haremos una observación similar sobre los α_j. Existe un polinomio no nulo con coeficientes enteros que tiene a los $\alpha_1, ..., \alpha_n$ como raíces. Sea $\alpha_1, ..., \alpha_N$ (N posiblemente distinto del anterior) el conjunto de todas las raíces de este polinomio. Observemos que todos los conjugados de cada α_j aparecen entre los $\alpha_1, ..., \alpha_N$. Sea

$$\beta_{n+1} = \beta_{n+2} = ... = \beta_N = 0. \tag{2.50}$$

Consideremos el producto

$$\prod (\beta_1 e^{\alpha_{k_1}} + \beta_2 e^{\alpha_{k_2}} + ... + \beta_N e^{\alpha_{k_N}}) \tag{2.51}$$

donde en el producto, $k_1, k_2, ..., k_N$ recorre todas las $N!$ permutaciones de los números $1, 2, ..., N$. Notemos que al expandirse el producto se tendría algo de la forma (2.43) y

sería igual a 0. Si βe^{α} denota un término en este producto expandido, entonces $\alpha = a_1\alpha_1 + ... + a_N\alpha_N$ para algunos enteros no negativos $a_1, ..., a_N$ cuya suma es $N!$. Aquí β sería un entero racional (ya que los β_j lo son). También estarían en el producto expandido todos los términos de la forma $\beta e^{\alpha'}$ con $\alpha' = a_1\alpha_{k_1} + ... + a_N\alpha_{k_N}$ donde $k_1, k_2, ..., k_N$ es cualquiera de las $N!$ permutaciones de los números $1, 2, ..., N$. Notemos que tal α' recorrerá un conjunto de todos los conjugados de α (y quizás más).

Finalmente, observemos que algún término βe^{α} en el producto expandido es no nulo, para ver esto consideremos el término no nulo βe^{α} en cada factor del producto no expandido con α teniendo la mayor parte real, y de éstos el que tiene α con la mayor parte imaginaria, el producto de éstos será un término βe^{α} no nulo. Entonces en (2.43), podemos suponer que hay enteros $n_0 = 0 < n_1 < ... < n_r = n$ tales que para cada $t \in \{0, 1, ..., r-1\}$ los números $\alpha_{n_t+1}, ..., \alpha_{n_{t+1}}$ forman un conjunto de conjugados completo y

$$\beta_{n_t+1} = ... = \beta_{n_{t+1}}. \tag{2.52}$$

Sea b un entero positivo para el cual $b\alpha_j$ son enteros algebraicos para cada j. Sea p un primo grande que elegiremos luego. Para $i \in \{1, ..., n\}$ sea

$$f_i(x) = b^{np}(x-\alpha_1)^p ... (x-\alpha_n)^p/(x-\alpha_i). \tag{2.53}$$

Finalmente, para $i \in \{1, ..., n\}$ sea

$$J_i = \beta_1 I_i(\alpha_1) + ... + \beta_n I_i(\alpha_n) \tag{2.54}$$

donde

$$I_i(t) = \int_0^t e^{t-u} f_i(u)\, du. \tag{2.55}$$

Integrando por partes obtenemos que

$$I_i(t) = e^t \sum_{j=0}^{\infty} f_i^{(j)}(0) - \sum_{j=0}^{\infty} f_i^{(j)}(t) = e^t \sum_{j=0}^{n} f_i^{(j)}(0) - \sum_{j=0}^{n} f_i^{(j)}(t) \tag{2.56}$$

donde n es el grado de $f_i(x)$. Si $f_i(x) = \sum_{j=0}^{n} a_j x^j$, definimos

$$\overline{f}_i(x) = \sum_{j=0}^{n} |a_j| x^j. \tag{2.57}$$

Entonces

$$|I(t)| \le |\int_0^t |e^{t-u} f_i(u)|\, du| \le |t| \max\{|e^{t-u}|\} \max\{|f(u)|\} \le |t| e^{|t|} \overline{f}(|t|). \tag{2.58}$$

Obtendremos una contradicción entre cotas superior e inferior para $|J_1 J_2 ... J_n|$. De (2.55) y (2.43) obtenemos que

$$J_i = -\sum_{j=0}^{m} \sum_{k=1}^{n} \beta_k f_i^{(j)}(\alpha_k) \tag{2.59}$$

donde $m = np-1$. Notemos que $f_i^{(j)}(\alpha_k)$ es $p!$ por un entero algebraico a menos que $j = p-1$ y $k = i$. También

$$f_i^{(p-1)}(\alpha_i) = b^{np}(p-1)! \prod_{1\leq k\leq n, k\neq i} (\alpha_i - \alpha_k)^p = (p-1)![F'(\alpha_i)]^p \tag{2.60}$$

donde

$$F(x) = \prod_{k=1}^{n} (bx - b\alpha_k) \in \mathbb{Z}[x]. \tag{2.61}$$

Consideremos el producto $J_1...J_n$. Observemos que uno de sus términos será

$$\prod_{i=1}^{n} \{(p-1)![F'(\alpha_i)]^p\}, \tag{2.62}$$

que es $(p-1)!$ (hasta $[(p-1)!]^n$) veces un entero racional no nulo. Si p es suficientemente grande, este entero no será divisible por p. Todo otro término puede escribirse como $p!$ por un entero algebraico. Mostraremos que la suma de estos enteros algebraicos es racional, y entonces por el Lema 11.1 que dice que "si α es un entero algebraico y α es racional, entonces α es un entero racional" los términos restantes sumados dan un entero racional divisible por $p!$. Observemos que para obtener esto alcanzaría con que mostremos que $J_1...J_n$ es un racional, pues entonces la suma que buscamos es simplemente

$$\{J_1...J_n - [(p-1)!]^n \prod_{i=1}^{n} [F'(\alpha_i)]^p\}/p! \tag{2.63}$$

que como se ve es racional. Consideremos entonces

$$J_1...J_n = (-1)^n \prod_{i=1}^{n} \left[\sum_{j=0}^{m}\sum_{k=1}^{n} \beta_k f_i^{(j)}(\alpha_k)\right]. \tag{2.64}$$

Se obtendría usando el Lema 11.3 que $J_1...J_n$ es un número racional si podemos mostrar que para cada $t \in \{0, 1, ..., r-1\}$, el lado derecho es simétrico en $\alpha_{n_t+1}, ..., \alpha_{n_{t+1}}$. Sea σ cualquier permutación de $n_t + 1, ..., n_{t+1}$ y extendamos σ para que fije los otros elementos de $\{1, 2, ..., n\}$. Entonces queremos mostrar que

$$\prod_{i=1}^{n} \left[\sum_{j=0}^{m}\sum_{k=1}^{n} \beta_k f_{\sigma(i)}^{(j)}(\alpha_{\sigma(k)})\right] = \prod_{i=1}^{n} \left[\sum_{j=0}^{m}\sum_{k=1}^{n} \beta_k f_i^{(j)}(\alpha_k)\right]. \tag{2.65}$$

Ya que $\sigma(i)$ recorre todos los números 1,...,n, tal como lo hace i, entonces obtenemos

$$\prod_{i=1}^{n} \left[\sum_{j=0}^{m}\sum_{k=1}^{n} \beta_k f_{\sigma(i)}^{(j)}(\alpha_{\sigma(k)})\right] = \prod_{i=1}^{n} \left[\sum_{j=0}^{m}\sum_{k=1}^{n} \beta_k f_i^{(j)}(\alpha_{\sigma(k)})\right]. \tag{2.66}$$

Poniendo $l = \sigma(k)$ vemos que

$$\prod_{i=1}^{n}\left[\sum_{j=0}^{m}\sum_{k=1}^{n}\beta_k f_i^{(j)}(\alpha_{\sigma(k)})\right] = \prod_{i=1}^{n}\left[\sum_{j=0}^{m}\sum_{l=1}^{n}\beta_{\sigma^{-1}(l)} f_i^{(j)}(\alpha_l)\right] \tag{2.67}$$

La ecuación $\beta_{n_t+1} = ... = \beta_{n_{t+1}}$ y la definición de σ dan

$$\prod_{i=1}^{n}\left[\sum_{j=0}^{m}\sum_{l=1}^{n}\beta_{\sigma^{-1}(l)} f_i^{(j)}(\alpha_l)\right] = \prod_{i=1}^{n}\left[\sum_{j=0}^{m}\sum_{l=1}^{n}\beta_l f_i^{(j)}(\alpha_l)\right] \tag{2.68}$$

y deducimos entonces que $J_1...J_n$ es un número racional como queríamos. Entonces $J_1...J_n$ es un entero racional divisible por $(p-1)!$ y no divisible por p, lo que implica que

$$(p-1)! \leq |J_1...J_n|. \tag{2.69}$$

Por otro lado podemos obtener una cota superior usando la cota superior para $|I(t)|$ de (2.58):

$$|J_1...J_n| \leq \prod_{i=1}^{n}[\sum_{k=1}^{n}|\beta_k||\alpha_k|e^{|\alpha_k|}\overline{f}_i(|\alpha_k|)] \leq (c_1 c_2^p)^n \leq c_3 c_4^p \tag{2.70}$$

para algunas constantes c_1, c_2, c_3, c_4 independientes de p. Obtenemos una contradicción cuando p es suficientemente grande, completando así la prueba. □

3 Formulación del séptimo problema de Hilbert

Durante el Segundo Congreso Internacional de Matematica de 1900 Hilbert expresó que "los problemas definidos tienen un profundo valor para el progreso de la ciencia matemática" y expusó una lista de 23 problemas cuyo estudio consideraba "podría estimular fuertemente el desarrollo de nuestra ciencia", en sus palabras

> *...el cierre de una gran epóca no sólo nos invita mirar hacia antrás en el pasado sino también dirige nuestros pensamientos hacia el futuro desconocido. El profundo significado de ciertos problemas para el avance de la ciencia matemática y el importante rol que juegan en el trabajo del investigador individual no deben negarse. Mientras que una rama de la ciencia ofrezca un gran número de problemas está viva, una falta de ellos amenaza con la extinción o el cese del desarrollo independiente. Así como cualquier actividad humana busca ciertos objetos, del mismo modo la investigación matemática requiere sus problemas...*

El séptimo problema de la lista se titulaba "irracionalidad y trascendencia de ciertos números", en sus palabras

> *Sabemos que ciertas funciones trascendentales especiales que juegan un rol importante en el análisis toman valores algebraicos para ciertos valores algebraicos dados del argumento, es esta una circunstancia que nos impacta como especialmente sorprendente y merecedora de mayor estudio. Siempre esperamos que las funciones trascendentales suelan tomar valor trascendentales en valores algebraicos del argumento. Pero aunque sabemos que inclusive existen funciones trascendentales enteras que toman valores racionales en todos los números algebraicos, de todos modos consideramos altamente probable que funciones como $e^{\pi iz}$, que obviamente toma valores algebraicos para todo racional z, tomen sin embargo valores trascendentales para todo algebraico irracional z. A esta afirmación se le puede dar aspecto geométrico de la siguiente forma. Si el cociente entre el ángulo del vertice y el ángulo de la base en un triángulo isósceles es un número algebraico pero irracional, entonces el cociente entre la base y uno de los lados laterales es un número trascendental.*

Para ver la equivalencia de las 2 formas de la conjetura de Hilbert, sea α el ángulo de la base del triángulo isósceles, de modo tal que el ángulo del vértice es $\pi - 2\alpha$, entonces el cociente entre los ángulos es $(\pi/\alpha) - 2$, que es racional o algebraico si y sólo si $z = \alpha/\pi$. Pero el cociente entre los lados del triángulo es $2cos\alpha = e^{i\alpha} - e^{-i\alpha} = e^{\pi iz} - (e^{\pi iz})^{-1}$. Este número es trascendental si y solo si $e^{\pi iz}$ lo es. Continuando con las palabras de Hilbert

> *A pesar de la simplicidad de esta afirmación, y su similitud con los problemas resueltos por Hermite y Lindemann, me parece que es extraordinariamente difícil de probarlo, tal como parece ser probar que α^β es siempre trascendental (o por lo menos irracional) si α es algebraico y β es irracional y algebraico, por ejemplo $2^{\sqrt{2}}$ o $e^\pi = i^{-2i}$. Puede ser cierto que las soluciones de éste problema y de problemas similares requiera desarrollar nuevos métodos y nuevos puntos de vista sobre la naturaleza esencial de ciertos números irracionales y trascendentales.*

Es interesante señalar que, según Siegel, Hilbert solía decir que la prueba de la irracionalidad de $2^{\sqrt{2}}$ pertenecía a un futuro mas distante que la prueba de la Hipótesis de Riemann o la Última Conjetura de Fermat. Aunque Hilbert no acertó en esta estimación sí lo hizo en que su séptimo problema sería un gran estímulo para la investigación. Su solución implicó el desarrollo de poderosos nuevos métodos analíticos que se siguieron aplicando y desarrollando hasta el presente, haciendo posible solucionar muchos problemas que no habían podido ser atacados mediante métodos anteriores. A continuacíon enunciaremos la conjetura de Hilbert y algunas formulaciones equivalentes

Conjetura 3.1 (Hilbert 1)**.** *Sean* α,$\beta \in \mathbb{A}$ *con* $\alpha \neq 0$ *y* $\alpha \neq 1$ *y* $\beta \notin \mathbb{Q}$,*entonces* α^{β} *es trascendente.*

Conjetura 3.2 (Hilbert 2)**.** *Sean* γ, $\beta \in \mathbb{C}$ *con* $l \neq 0$ *y* $\beta \notin \mathbb{Q}$, *entonces al menos uno de los tres números* e^{γ}, β, $e^{\beta\gamma}$ *es trascendente.*

Conjetura 3.3 (Hilbert 3)**.** *Sean* α, $\beta \in \mathbb{A} - 0$ *con* $\log\alpha$ *y* $\log\beta$ *linealmente independientes sobre* $\mathbb{Q}$, *entonces* $\log\alpha$ *y* $\log\beta$ *son linealmente independientes sobre* $\mathbb{A}$.

Conjetura 3.4 (Hilbert 4)**.** *Sean* α, $\beta \in \mathbb{A} - 0, 1$, *si* $\eta = \log\alpha/\log\beta$ *es irracional, entonces es trascendente.*

Prueba de las equivalencias. Notar que 3.3 es equivalente 3.4 ya que $\mathbb{A}$ es cerrado para división.

Para ver que 3.1 implica 3.2, tomamos $\alpha = e^{\gamma}$, entonces α no es ni 0 ni 1, entonces 3.1 implica que α y β son algebraicos, entonces $\alpha^{\beta} = e^{\gamma\beta}$ lo cual implica 3.1.

Para ver que 3.2 implica 3.3, observamos que la condicion de que $\log\alpha$ y $\log\beta$ son linealmente independientes sobre $\mathbb{Q}$ implica que ni α ni β es 1, y además que $\log\alpha/\log\beta \notin \mathbb{Q}$. Sea $\gamma = \log\beta$ y $\beta' = \log\alpha/\log\beta$. Entonces 3.2 implica que β' es trascendente, lo cual implica 3.3.

Para ver que 3.3 implica 3.1, consideramos $\beta' = e^{\beta\log\alpha}$. Entonces $\log\alpha$ y $\log\beta'$ son linealmente independientes sobre $\mathbb{A}$. Así que por 3.3 $\log\alpha$ y $\log\beta'$ son linealmente independientes sobre $\mathbb{Q}$. Esto contradice que $\beta \notin \mathbb{Q}$. □

Como mencionamos antes, si se compara la conjetura 3.1 con la de Euler en (2.2), se ve que la de Hilbert expande la naturaleza artimética de los números involucrados en la primera.

Veamos que si se remueve la condición de que α o β sea algebraico, la afirmación de la Conjetura 3.1 no es universalmente válida. Para ello usaremos el hecho que probaremos más adelante de que la conjetura es cierta.

Caso $\alpha \notin \mathbb{A}$. Sea $\alpha = (\sqrt{2}^{\sqrt{2}})$ y $\beta = \sqrt{2}$. Por la Conjetura 3.1 α es trascendente. Pero se tiene que $\alpha^{\beta} = 2$ que es algebraico.

Caso $\beta \notin \mathbb{A}$. Sea $\alpha = 3$ y $\beta = \frac{\log 2}{\log 3}$. Si β es irracional por la Conjetura 3.2 (poniendo en él $\gamma = \log 3$ y $\beta = \frac{\log 2}{\log 3}$) se obtiene que β es trascendente. Pero se tiene que $\alpha^{\beta} = e^{\log 3\beta} = 2$ que es algebraico.

4 Caso particular de Gelfond: e^π es trascendente

En esta sección expondremos este resultado obtenido en 1929 por Gelfond.

Teorema 4.1 (1929 Gelfond)**.** *e^π es trascendente.*

4.1 Preliminares

La prueba de Gelfond no se basa en la representación de e como serie de potencias, o más bien en la de $e^{\pi z}$, sino en otras aproximaciones polinomiales de $e^{\pi z}$. Estas aproximaciones polinomiales están basadas en los enteros gaussianos.

La colección de enteros gaussianos es el conjunto $\{a + bi : b \in \mathbb{Z}\}$. Lo esencial de ellos es que si e^π se supone algebraico, entoces la función $f(z) = e^{\pi z}$ tomará un valor algebraico en cada uno de los enteros gaussianos. Especificamente si $a + bi$ es un entero gaussiano, entonces

$$f(a + bi) = e^{\pi(a+bi)} = e^{\pi a}e^{\pi bi} = (-1)^b(e^\pi)^a \tag{4.1}$$

es algebraico. Para describir más adelante como Gelfond usó a los enteros gaussianos para dar aproximaciones polinomiales de la función $e^{\pi z}$ necesitamos primero una forma de ordenarlos. Gelfond los ordenó por su valor absoluto, y para enteros gaussianos con mismo valor absoluto por sus argumentos. Esto da el siguiente ordenamiento:

$$z_=0, z_1 = 1, z_2 = i, z_3 = -1, z_4 = -i, z_5 = i + 1, z_6 = -1 + i, z_7 = -1 - iz_8 = 1 - i, ... \tag{4.2}$$

Por otro lado, es posible aproximar la función e^z por un serie infinita cada uno de cuyos términos es un polinomial con todos sus ceros dentro los enteros gaussianos ordenados como antes (esto lleva a la llamada serie de Newton para la función e^z). Sea $\{z_0 = 0, z_1, z_2, ...\}$ los enteros gaussianos (ordenados), y consideremos los polinomios

$$P_0(z) = 1, P_1(z) = z = z - z_0, P_2(z) = z(z - z_1), ..., P_h(z) = z(z - z_1)...(z - z_{h-1}) \tag{4.3}$$

Queremos una interpolación de la función $f(z) = e^{\pi z}$ por estos polinomios:

$$e^{\pi z} = A_0P_0(z) + A_1P_1(z) + A_2P_2(z) + ... + A_nP_n(z) + R_n(z), \tag{4.4}$$

donde los coeficientes numéricos $A_0, A_1, ..., A_n$ y el término resto polinomial $R_n(z)$ tengan representaciones integrales. En particular si γ_n y γ'_n son cualesquiera par de líneas cerradas simples que encierran a los puntos de interpolación $z_0, z_1, z_2, ..., z_n$ entonces

$$A_n = \frac{1}{2\pi i}\int_{\gamma_n} \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)}\, d\sigma \tag{4.5}$$

y

$$R_n(z) = \frac{P_{n+1}(z)}{2\pi i}\int_{\gamma'_n} \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)}\, d\sigma. \tag{4.6}$$

Más adelante volveremos a la cuestión de cómo elegir estos contornos de las integrales.

Haremos uso además del siguiente lema:

Lema 4.2. *La función $f(z) = e^z$ es trascendente.*

Prueba. Alcanza con probarlo para el caso $z \in \mathbb{R}$, pues entonces e^z con $z \in \mathbb{C}$ también debe ser trascendente porque sino restringiendo z a los números reales se contradice lo primero.

Supongamos que e^x es una función algebraica. Entonces existe un conjunto de polinomios $p_n(x), p_{n-1}(x), p_{n-2}(x), ..., p_1(x), p_0(x)$, donde $n \geq 1$ y $p_n(x)$ no es idénticamente nulo, tales que la ecuación

$$p_n(x)(e^x)^n + p_{n-1}(x)(e^x)^{n-1} + p_{n-2}(x)(e^x)^{n-2} + ... + p_1(x)e^x + p_0(x) = 0, \tag{4.7}$$

o

$$p_n(x)(e^{nx}) + p_{n-1}(x)e^{(n-1)x} + p_{n-2}(x)e^{(n-2)x} + ... + p_1(x)e^x + p_0(x) = 0, \tag{4.8}$$

se satisface $\forall x \in \mathbb{R}$. Entre todos los posibles polinomios anteriores, elegimos aquel para el cual el grado en e^x es mínimo. Sea m el grado de $p_0(x)$. Entonces

$$\frac{d^{m+1}}{dx^{m+1}} p_0(x) = 0 \tag{4.9}$$

y para $h = n, n-1, n-2, ..., 1$ se tiene:

$$\begin{aligned}
\frac{d}{dx}p_h(x)e^{hx} &= p_h'(x)e^{hx} + hp_h(x)e^{hx} = (p_h'(x) + hp_h(x))e^{hx} \\
\frac{d^2}{dx^2}p_h(x)e^{hx} &= (p_h''(x) + hp_h'(x))e^{hx} + h(p_h'(x) + hp_h(x))e^{hx} = (p_h''(x) + 2hp_h'(x) + h^2p_h(x))e^{hx} \\
\frac{d^3}{dx^3}p_h(x)e^{hx} &= (p_h'''(x) + 2hp_h''(x) + h^2p_h'(x))e^{hx} + h(p_h''(x) + 2hp_h'(x) + h^2p_h(x))e^{hx} \\
&\vdots \\
\frac{d^{m+1}}{dx^{m+1}}p_h(x)e^{hx} &= q_h(x)e^{hx}
\end{aligned}$$

donde $q_h(x)$ es una suma de múltiplos constantes de $p_h(x)$ y de sus derivadas hasta de orden $m+1$, y por lo tanto es un polinomio y no es idénticamente nulo si $p_h(x)$ no lo es, ya que en la suma los polinomios son de distintos grados.

En consecuencia $q_n(x)$ no es idénticamente nulo. Entonces tomando $m+1$ derivadas sucesivas en ambos lados de (4.8) tenemos que:

$$q_n(x)(e^{nx}) + q_{n-1}(x)e^{(n-1)x} + q_{n-2}(x)e^{(n-2)x} + ... + q_2(x)e^{2x} + q_1(x)e^x = 0 \tag{4.10}$$

para todo x, donde $q_n(x)$ no es idénticamente nulo. Dividiendo ambos lados de (4.10) por e^x tenemos

$$q_n(x)(e^{(n-1)x}) + q_{n-1}(x)e^{(n-2)x} + q_{n-2}(x)e^{(n-3)x} + ... + q_2(x)e^x + q_1(x) = 0 \tag{4.11}$$

que puede reescribirse como

$$q_n(x)(e^x)^{(n-1)} + q_{n-1}(x)(e^x)^{(n-2)} + q_{n-2}(x)(e^x)^{(n-3)} + ... + q_2(x)e^x + q_1(x) = 0 \tag{4.12}$$

para todo x, donde $q_n(x)$ no es idénticamente nulo y $n-1 \geq 0$. Si $n-1=0$ entonces $n=1$ así $q_1(x)$ no es idénticamente 0, y la ecuación (4.12) se reduce a $q_1(x)=0\ \forall x$, es decir, $q_1(x)$ idénticamente nulo, una contradicción. Entonces se deduce que debemos tener $n-1>0$. En consecuencia, (4.12) con grado $n-1$ en e^x satisface la definición de función algebraica aplicada a e^x contradiciendo la minimalidad de n. Como esta contradicción provino de suponer que la función e^x era algebraica, queda probado el lema. □

4.2 Idea de la prueba

Paso 1) Suponer que e^π es algebraico (entoncs la función $e^{\pi z}$ toma una valor algebraico en cada entero gaussiano).
Paso 2) Mostrar que para n suficientemente grande $A_n = 0$. Esto significa que existe un entero positivo N' tal que si $n > N'$ entonces $A_n = 0$. Esto nos dice que para todo $n > N'$ tenemos la representación de $e^{\pi z}$ siguiente:

$$e^{\pi z} = A_0P_0(z) + A_1P_1(z) + A_2P_2(z) + ... + A_{N'}P_{N'}(z) + R_n(z). \tag{4.13}$$

Paso 3) Tomando γ_n' igual a un círculo de radio n y haciendo $n \to \infty$ se obtiene que $R_n(z) \to 0$ para todo z. Esto implica que la función $e^{\pi z}$ puede representarse por un polinomio.
Paso 4) Concluir que e^z no es una función trascendente.
Esta última conclusión contradice la trascendencia de la función e^z y por lo tanto la suposición inicial de que e^π es algebraico no puede ser cierta, entonces es trascendente.

4.3 Prueba: Paso 2

La prueba de que para n suficientemente grande $A_n = 0$ tiene dos partes. Primero se utilizan herramientas analíticas para encontrar una cota superior de $|A_n|$ que depende en parte de elegir un contorno de integración razonablemente corto γ_n. Segundo usando herramientas puramente algebraicas, a partir de la suposición de que e^π es algebraico (es la única parte de la prueba en la que esta hipótesis se utiliza) se obtiene que cada una de la expresiones A_n es un número algebraico, y luego usando estimaciones sutiles Gelfond encuentra un denominador chico para A_n. Si $A_n \neq 0$, multiplicando A_n por un denominador y tomando norma algebraica se contradice la cota superior obtenida en la primer parte. Entonces $A_n = 0$.

Veamos la primer parte: encontrar una cota superior para $|A_n|$.

Usando la respresentación (4.5) tenemos que

$$|A_n| = \left| \frac{1}{2\pi i} \int_{\gamma_n} \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)}\, d\sigma \right| \leq \frac{1}{2\pi}(\text{long de } \gamma_n) \max_{\sigma\in\gamma_n} \frac{|e^{\pi\sigma}|}{|\sigma||(\sigma - z_1)|...|(\sigma - z_n)|} \tag{4.14}$$

$$\leq \frac{1}{2\pi}(\text{long de } \gamma_n) \frac{\max_{\sigma\in\gamma_n} |e^{\pi\sigma}|}{\min_{\sigma\in\gamma_n} |\sigma||(\sigma - z_1)|...|(\sigma - z_n)|}.$$

Para dar una cota superior razonablemente chica para $|A_n|$ Gelfond necesitba entender los posibles contornos γ_n que encerrarían a los enteros gaussianos que aparecen en la representación integral de A_n. Necesitaba que la longitud del contorno fuese lo más chica posible,

pero que también cada uno de los

$$\max_{\sigma\in\gamma_n}|e^{\pi\sigma}| \qquad \text{y} \qquad \frac{1}{\min_{\sigma\in\gamma_n}|\sigma||(\sigma-z_1)|...|(\sigma-z_n)|} \tag{4.15}$$

fuera chico. Cualquier estimación para cualquiera de estos dos números dependerá del contorno de integración elegido.

Como los valores absolutos de los enteros gaussianos ordenados es no decreciente, antes de especificar γ_n Gelfond necesitaba estimar $|z_n|$ y entonces saber qué tan grande sería el contorno a usar, para lo cual usó ciertas estimaciones hechas por E. Landau. Si temporalmente denotamos por $G(r)$ a los enteros gaussianos x_k+y_ki con $x_k^2+y_k^2\leq r^2$ Landau había probado que para $r>\sqrt{2}$, se tiene que

$$\pi(r-\sqrt{2})^r\leq G(r)\leq\pi(r+\sqrt{2})^2 \tag{4.16}$$

y que a partir de ello se tiene

$$|z_n|=\sqrt{\frac{n}{\pi}}+o(\sqrt{n}), \quad \text{con} \quad \lim_{n\to\infty}\frac{o(\sqrt{n})}{\sqrt{n}}=0. \tag{4.17}$$

Gelfond comprendió que esta estimación quería decir que podía tomar como contorno de integración un círculo de radio mayor a una constante por $\sqrt{\frac{n}{\pi}}$, Gelfond decidió usar el radio relativamente grande de n. Con este contorno estimó la primer expresión en (4.15).

$$\max_{\sigma\in\gamma_n}|e^{\pi\sigma}|\leq e^{\pi\max Re(\sigma):\sigma\in\gamma_n}=e^{\pi n} \tag{4.18}$$

Para estimar la segunda expresión en (4.15) necesitamos una estimación para la mínima distancia de cada uno de los primeros n enteros gaussianos $z_1,...,z_n$ y los puntos del círculo γ_n y queremos que este mínimo sea chico. La necesidad de tal estimación puede ser una de las razones por las que Gelfond tomó un contorno de integración con un radio mayor al que hubiera sido necesario para solamente contener los primeros n enteros gaussianos. De las estimaciones para $|z_n|$ anteriores, vemos que para n suficientemente grande:

$$|z_n|\leq\sqrt{\pi}\sqrt{\frac{n}{\pi}}=\sqrt{n}, \tag{4.19}$$

así que para $1\leq i\leq n, \min\{|\sigma-z_i|:\sigma\in\gamma_n\}\geq n-\sqrt{n}\geq\frac{1}{2}n$, para n suficientemente grande. Entonces tenemos

$$\frac{1}{\min_{\sigma\in\gamma_n}|\sigma||(\sigma-z_1)|...|(\sigma-z_n)|}\leq\left(\frac{2}{n}\right)^{n+1} \tag{4.20}$$

Juntando todas las estimaciones anteriores obtenemos que para n suficientemente grande

$$|A_n|=\left|\frac{1}{2\pi i}\int_{\gamma_n}\frac{e^{\pi\sigma}}{\sigma(\sigma-z_1)...(\sigma-z_n)}\,d\sigma\right|\leq\frac{1}{2\pi}(\text{long de }\gamma_n)\max_{\sigma\in\gamma_n}\frac{|e^{\pi\sigma|}}{|\sigma||(\sigma-z_1)|...|(\sigma-z_n)|} \tag{4.21}$$

$$\leq \frac{1}{2\pi} 2\pi n e^{\pi n} \left(\frac{2}{n}\right)^{n+1} \leq e^{\log n + n - (n+1)\log(n/2)}.$$

Veamos la segunda parte: encontrar una cota inferior para $|A_n|$ para aquellos n para los que $A_n \neq 0$.

Comenzamos aplicando residuos (6.3) para expresar A_n como un número algebraico

$$A_n = \frac{1}{2\pi i} \int_{\gamma_n} \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)}\, d\sigma$$

$$= \sum_{k=0}^{n} \text{res de } \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)} \text{en } z = z_k$$

$$= \sum_{k=0}^{n} \frac{e^{\pi z_k}}{\prod_{j=0, j\neq k}^{n} (z_k - z_j)}$$

.

Y si para cada uno de los enteros gaussianos ordenados usamos la notación $z_k = x_k + y_k i$ tenemos

$$A_n = \sum_{k=0}^{n} \frac{e^{\pi z_k}}{\prod_{j=0, j\neq k}^{n} (z_k - z_j)}$$

$$= \sum_{k=0}^{n} \frac{e^{\pi(x_k + y_k i)}}{\prod_{j=0, j\neq k}^{n} (z_k - z_j)}$$

$$= \sum_{k=0}^{n} \frac{(e^{\pi})^{x_k}(-1)^{y_k}}{\prod_{j=0, j\neq k}^{n} (z_k - z_j)}$$

Esta ecuación muestra que para cada n, A_n es un número algebraico porque cada uno de los sumandos

$$\frac{(e^{\pi})^{x_k}(-1)^{y_k}}{\prod_{j=0, j\neq k}^{n} (z_k - z_j)} \tag{4.22}$$

es un cociente de dos números algebraicos.

La norma algebraica de un entero algebraico no nulo es un entero no nulo, pero la norma algebraica de un número algebraico que no es un entero algebraico es simplemente un número racional. Para obtener un entero a partir de A_n necesitamos primero multiplicar por su denominador. El denominador de cada uno de los sumandos de la respresentación anterior de A_n es un producto de diferencias de enteros gaussianos. Como $\mathbb{Z}[i]$ es un anillo, estos denominadores son a su vez enteros gaussianos. Necesitamos comprender mejor los numeradores y denominadores con el fin de encontrar enteros apropiados Ω_n por los cuales multiplicar A_n para obtener un entero algebraico. Como hizo Gelfond, introducimos para simplificar la notación

$$\omega_{n,k} := \prod_{j=0, j\neq k}^{n} (z_k - z_j). \tag{4.23}$$

Entonces tenemos

$$A_n = \frac{(e^\pi)^{x_0}(-1)^{y_0}}{\omega_{n,0}} + \frac{(e^\pi)^{x_1}(-1)^{y_1}}{\omega_{n,1}} + ... + \frac{(e^\pi)^{x_n}(-1)^{y_n}}{\omega_{n,n}}. \tag{4.24}$$

El anillo $\mathbb{Z}[i]$ es un domino de factorización única, así que la noción de mímino común múltiplo de una colección de enteros gaussianos tiene sentido. En otro paper, también publicado en 1929, Gelfond estudió la distribución de los elementos irreducibles en $\mathbb{Z}[i]$ y concluyó que para

$$\Omega_n := \text{Mínimo común múltiplo}\{(z_1 - z_2)(z_1 - z_3)...(z_1 - z_n),$$

$$(z_2 - z_1)(z_2 - z_3)...(z_2 - z_n), ..., (z_n - z_1)(z_n - z_2)...(z_n - z_{n-1})\}$$

se tiene la cota

$$|\Omega_n| \leq e^{\frac{1}{2}n\log n + 163n + O(\sqrt{n})}. \tag{4.25}$$

Recordando que $\omega_{n,k} := \prod_{j=0, j\neq k}^{n}(z_k - z_j)$ tenemos la respresentación para $\Omega_n A_n$

$$\Omega_n A_n = \frac{\Omega_n}{\omega_{n,0}}(e^\pi)^{x_0}(-1)^{y_0} + \frac{\Omega_n}{\omega_{n,1}}(e^\pi)^{x_1}(-1)^{y_1} + ... + \frac{\Omega_n}{\omega_{n,n}}(e^\pi)^{x_n}(-1)^{y_n}. \tag{4.26}$$

Observemos por un lado que en la expresión anterior cada $\frac{\Omega_n}{\omega_{n,k}}$ es un elemento de $\mathbb{Z}[i]$, así que es un entero algebraico, y por otro que en el ordenamiento de Gelfond de los elementos $z_k = x_k + y_k i \in \mathbb{Z}[i]$, x_k puede ser positivo, negativo o cero, entonces cada uno de los numeradores de la expresión anterior involucra e^π o $e^{-\pi}$.

En conclusión, estas dos observaciones nos dicen que $\Omega_n A_n$ es una expresión polinomial con coeficientes enteros en e^π, $e^{-\pi}$ y i. Es posible simplificar multiplicando todo por una potencia alta apropiada de e^π. Especificamente, si ponemos $r_n = \max_{0\leq k\leq n}|x_k|$, notando para uso posterior que $r_n \leq \sqrt{n}$, entonces $(e^\pi)^{r_n}\Omega_n A_n$ es una expresión polinomial con coeficientes enteros en e^π y i:

$$(e^\pi)^{r_n}\Omega_n A_n = \frac{\Omega_n}{\omega_{n,0}}(e^\pi)^{r_n+x_0}(-1)^{y_0} + \frac{\Omega_n}{\omega_{n,1}}(e^\pi)^{r_n+x_1}(-1)^{y_1} + ... + \frac{\Omega_n}{\omega_{n,n}}(e^\pi)^{r_n+x_n}(-1)^{y_n}. \tag{4.27}$$

Finalmente, sea δ un denominador para e^π. Tenemos el entero algebraico:

$$P_n(i, e^\pi) = (\delta)^{2\sqrt{n}}(e^\pi)^{r_n}\Omega_n A_n \tag{4.28}$$

$$= (\delta)^{2\sqrt{n}-(r_n+x_0)}\frac{\Omega_n}{\omega_{n,0}}(\delta e^\pi))^{r_n+x_0}(-1)^{y_0}$$

$$+(\delta)^{2\sqrt{n}-(r_n+x_1)}\frac{\Omega_n}{\omega_{n,1}}(\delta e^\pi))^{r_n+x_1}(-1)^{y_1}$$

$$+... + (\delta)^{2\sqrt{n}-(r_n+x_n)}\frac{\Omega_n}{\omega_{n,n}}(\delta e^\pi))^{r_n+x_n}(-1)^{y_n},$$

donde $P_n(x,y)$ es el polinomio con coeficientes enteros subyacente en la expresión anterior. Nuestro objectivo es calcular la norma algebraica del entero algebraico anterior, es decir, $N(P_n(i,e^\pi))$. Como $P_n(i,e^\pi) \neq 0$, tenemos que $N(P_n(i,e^\pi))$ es un entero no nulo. Denotamos a los conjugados de e^π por $\theta_1(=e^\pi), \theta_2, ..., \theta_d$. Entonces una potencia entera de $N(P_n(i,e^\pi))$ está dada por el producto

$$N = P_n(i,\theta_1)P_n(-i,\theta_1)[\prod_{j=2}^{d} P_n(i,\theta_j)][\prod_{j=2}^{d} P_n(-i,\theta_j)]. \tag{4.29}$$

De lo anterior sigue que si $A_n \neq 0$ entonces $N \neq 0$.

Usaremos nuevamente el trabajo analítico de la parte anterior para acotar superiormente el primer factor $P_n(i,\theta_1)$, e información algebraica sobre los enteros gaussianos para estimar el valor absoluto de los demás factores.

Nuestra cota superior anterior para $|A_n|$ combinada con la cota de Gelfond para $|\Omega_n|$ y las cotas $r_n \leq \sqrt{n}$ y $|x_n| \leq \sqrt{n}$ nos dice que

$$|P_n(i,\theta_1)| \leq e^{-1/2n\log n+170n}, \qquad n \text{ suficientemente grande} \tag{4.30}$$

Cada uno de los otros $2d-1$ factores en (4.29) se acota con la desigualdad triangular.

Los términos más difíciles de acotar son los cocientes $\frac{\Omega_n}{\omega_{n,k}}$. Ya vimos que en otro paper Gelfond probó una cota superior para el numerador $|\Omega_n|$. Para encontrar una cota inferior para el denominador, Gelfond usa una cota dada por otro matemático, Seigo Fukasawa, quien en 1926 probó que para n suficientemente grande

$$|\omega_{n,k}| > e^{1/2n\log n-10n}. \tag{4.31}$$

Entonces para n suficientemente grande tenemos que

$$\left|\frac{\Omega_n}{\omega_{n,k}}\right| \leq e^{\frac{1}{2}n\log n+164n-(\frac{1}{2}n\log n-10n)} \leq e^{174n}. \tag{4.32}$$

Juntando todas estas cotas tenemos que para cada uno de los factores en (4.29) se cumple que

$$|P_n(\pm i,\theta_j)| \leq (n+1)e^{174n}(\delta e^\pi)^{2\sqrt{n}} \leq e^{175n} \tag{4.33}$$

para n suficientemente grande.

Finalmente, tenemos un entero no nulo N, tal que

$$0 < |N| < e^{-\frac{1}{2}n\log n+(2d-1)175n}. \tag{4.34}$$

Pero para n suficientemente grande la expresión de la derecha tiende a 0 y tenemos entonces una contradicción, con lo cual nuestra suposición inicial de que $A_n \neq 0$ debe ser falsa, es decir que para n suficientemente grande, digamos $n > N'$, se tiene que $A_n = 0$.

4.4 Prueba: Pasos 3 y 4

Lo anterior nos dice que para $n > N'$ tenemos la representación de la función $e^{\pi z}$:

$$e^{\pi z} = A_0P_0(z) + A_1P_1(z) + A_2P_2(z) + ... + A_{N'}P_{N'}(z) + R_n(z). \tag{4.35}$$

Tomamos por contorno de integración en la representación

$$R_n(z) = \frac{P_{n+1}(z)}{2\pi i} \int_{\gamma_n'} \frac{e^{\pi\sigma}}{\sigma(\sigma - z_1)...(\sigma - z_n)}\, d\sigma \tag{4.36}$$

un círculo de radio n. Como para cualquier z, $R_n(z) \to 0$ cuando $n \to \infty$ se tiene que $e^{\pi z}$ es igual al polinomio $A_0P_0(z) + A_1P_1(z) + A_2P_2(z) + ... + A_{N'}P_{N'}(z)$, y no es entonces una función trascendente, lo que implica que e^z tampoco lo es, pero en verdad lo es, así que tenemos una contradicción proveniente de la suposición de que e^π era algebraico, con lo cual debe ser trascendente.

5 Caso particular de Kuzmin : $2^{\sqrt{2}}$ es trascendente

En esta sección comentaremos el siguiente resultado.

Teorema 5.1 (1930 Kuzmin)**.** $2^{\sqrt{2}}$ *es trascendente*

Tanto la estructura como los detalles de la prueba de Kuzmin son muy similares a los de la de Gelfond sobre la trascendencia de e^{π}. Kuzmin comienza suponiendo que $2^{\sqrt{2}}$ es algebraico y considera la función a valores reales $2^z = e^{\log 2z}$. Luego aproxima 2^z usando la fórmula de interpolación de Lagrange, esto es, en vez de usar los enteros gaussianos usa los números $\{a + b\sqrt{2} \ : a, b \in \mathbb{Z}\}$ ordenados de cierta forma. Luego usando la notación $P_h = (z - z_1)(z - z_2)...(z - z_h)$ Kuzmin tiene que para cada n

$$2^z = \sum_{j=1}^{n} \frac{P_n(z)}{z - z_j} \frac{2^{z_j}}{P_n'(z_j)} + \frac{P_n(z)}{n!} 2^{\sigma} (\log 2)^n, \tag{5.1}$$

donde σ está entre el menor y el mayor de los z_j.

Pero $P_n'(z) = \sum_{r=1}^{n} \prod_{l \neq r} z - z_l$, lo cual implica que $P_n'(z_j) = \prod_{l \neq j} z_j - z_l$.

Entonces para $z_0 \notin z_1, ..., z_n$ se tiene

$$2^{z_0} = \sum_{j=1}^{n} \frac{P_n(z_0) 2^{z_j}}{(z_0 - z_j) \prod_{l \neq j} z_j - z_l} + \frac{P_n(z_0)}{n!} 2^{\sigma} (\log 2)^n. \tag{5.2}$$

Dividiendo por $P_n(z_0)$ y reescribiendo obtenemos que

$$\sum_{j=0}^{n} \frac{2^{z_j}}{\prod_{l \neq j} z_j - z_l} = \frac{2^{\sigma} (\log 2)^n}{n!} \tag{5.3}$$

Luego

1) Multiplicando el lado izquierdo de esta igualdad por el mínimo común múltiplo de los denominadores y luego multiplicando por un denominador algebraico se obtiene un entero algebraico no nulo.

2) Tomando n suficientemente grande el lado derecho toma valor absoluto chico.

3) Tomando la norma algebraica del lado izquierdo se obtiene un entero no nulo de valor absoluto menor a 1, lo cual es absurdo.

Esta contradicción proviene de suponer que el término de error en la interpolación de Lagrange es no nulo, entonces debe serlo, pero esto implica que la función trascendental $f(z) = 2^z$ es una función polinomial, absurdo que proviene de la suposición de que $2^{\sqrt{2}}$ es algebraico, así que debe ser trascendente.

6 Solución de Gelfond

6.1 Preliminares

Recordamos en esta sección ciertos resultados (algunos supondremos conocidos y otros no) de los cuales haremos uso en la demostración de Gelfond.

Teorema 6.1 (Matriz de potencias)**.** *Sea $m_1, ..., m_h$ pertenecientes a un domino íntegro. Sea la matriz $M \in \mathbb{C}^{h\times n}$ tal que $M_{i,j} = m_i^{j-1}$. Se tiene que* $\det M = 0$ *si y solo si tiene 2 columnas iguales.*

Teorema 6.2 (Representación integral de holomorfa)**.** *Sea D disco cerrado de $\mathbb{C}$, C su frontera y $f(z) : D \to \mathbb{C}$ holomorfa, entonces para todo $z \in int(D)$ se tiene*

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_C \frac{f(w)}{(w-z)^{n+1}}\,dw$$

Teorema 6.3 (Fórmula de residuos)**.** *Sea $f(z)$ holomorfa en un abierto conexo $\Omega \subset \mathbb{C}$, excepto por eventuales singularidades aisladas a_j $1 \le j \le n$. Entonces*

$$\frac{1}{2\pi i}\int_\gamma f(z)\,dz = \sum_{j=1}^{n} R_{z=a_j} f(z) \tag{6.1}$$

para cualquier ciclo γ homólogo a cero en Ω que no pasa por ninguno de los a_j.

Recordemos que el residuo $R_{z=a}$ es el único número complejo R, que hace que $f(z) - \frac{R}{z-a}$ sea la derivada de una función holomorfa en $0 < |z-a| < \delta$. Equivalentemente R es el único número complejo para el cual

$$\int_C f(z) - \frac{R}{z-a}\,dz = 0 \tag{6.2}$$

donde C es una circunferencia tal que $f(z)$ es holomorfa en ella y su interior. Es decir

$$R_{z=a} = \frac{1}{2\pi i}\int_C f(z)\,dz, \tag{6.3}$$

lo que es igual al coeficiente con subíndice -1 en la expansión en serie de Laurent de $f(z)$ entorno de $z = a$.

Teorema 6.4 (Módulo max)**.** *Sea $R \subset \mathbb{C}$ un abierto conexo y $f(z) : R \to \mathbb{C}$ holomorfa, entonces $|f(z)|$ alcanza su máximo en la frontera de R.*

Recordemos que si α es un número algebraico de orden m sobre $\mathbb{Q}$, se tiene para cada $\beta \in \mathbb{Q}(\alpha)$ con gr $\beta \le m-1$, que β es algebraico de orden d sobre $\mathbb{Q}$ con $d|m$, y además los conjugados de β se obtienen reemplzando en el polinomio que lo define α por sus conjugados.

Lema 6.5. *Si γ es algebraico de grado m sobre $\mathbb{Q}$ y si A es un entero racional tal que $A\gamma$ es un entero algebraico, entonces o bien $\gamma = 0$ o bien $|\gamma| \ge A^{-m}[\mu(\gamma)]^{1-m}$*

Este lema esencialmente expresa el hecho de que $|N(A\gamma)| \geq 1$.

Lema 6.6 (Kronecker). *Dada $A = (a_{ij}) \in \mathbb{C}^{m\times n}, n > 2m$una matriz (a_{ij}) y $|a_{ij}| \leq A$. Sea P un número positivo dado. Entonces existen n enteros racionales $N_1, N_2, ..., N_n$ tales que*

$$|\sum_{i=1}^{n} N_i a_{ij}| \leq \frac{1}{P} \qquad j = 1, ..., m \tag{6.4}$$

$$\sum_{i=1}^{n} |N_i| \geq 1 \tag{6.5}$$

$$|N_i| \leq [2^{3/2} nAP]^{\frac{2m}{n-2m}} \qquad i = 1, ..., n. \tag{6.6}$$

Prueba. Sea $w_1, w_2, ..., w_n$ un conjunto de enteros no negativos cada uno menor o igual a un entero B que elegiremos luego. Hay $(B+1)^n$ posibles distintas elecciones de tales conjuntos, colección de elecciones que llamamos M_1.

Pongamos

$$\sum_{i=1}^{n} w_i a_{ij} = c_j + id_j \qquad j = 1, 2, ..., m. \tag{6.7}$$

En un espacio de $2m$ dimensiones marquemos los puntos $(c_1, d_1, c_2, d_2..., c_m, d_m)$. Llamamos M_2 a la colección de puntos $2m-$dimensionales obtenida de esta forma a partir de cada elemento de M_1.

Estos puntos están contenidos en un cubo $2m-$dimensional de lado $\leq 2nAB$ ya que de (6.7) se tiene que $|c_j|, |d_j| \leq |c_j + id_j| \leq nAB$. Si subdividimos el lado en intervalos de longitud $2^{-1/2}P^{-1}$, obtenemos como mucho $2^{3/2}nABP + 1$ (su parte entera) intervalos ya que puede sobrar un pedacito de longitud menor a $2^{-1/2}P^{-1}$. Éstos a su vez determinan una partición del cubo grande en como mucho $(2^{3/2}nABP + 1)^{2m}$ subceldas (paralelepípedos) ya que cada uno de los 2m ejes lo tengo partido en $(2^{3/2}nABP + 1)$ intervalos, entonces el total de posibilidades de elegir un intervalo de cada eje es el número anterior. Por construcción cada lado de estas subceldas tiene longitud menor o igual a $2^{-1/2}P^{-1}$.

Si elegimos B de forma tal que

$$(B + 1)^n > (2^{3/2}nABP + 1)^{2m} \tag{6.8}$$

entonces habrá por lo menos una de las subceldas $2m-$dimensionales conteniendo 2 puntos de nuestra colección M_2, digamos $P = (c_1, d_1, c_2, d_2..., c_m, d_m)$ y $P' = (c'_1, d'_1, c'_2, d'_2..., c'_m, d'_m)$. Supongamos que los correspondientes conjuntos de enteros de los que provienen son $(v_1, v_2, ..., v_n)$ y $(v'_1, v'_2, ..., v'_n)$.

Elegimos $B := [2^{3/2}nAP]^{\frac{2m}{n-2m}}$ y luego $N_i := v_i - v'_i$ para $i = 1, ..., n$; nombramos $C := 2^{3/2}nABP$. Veamos que con estas elecciones de N_i y B se satisfacen todas las cualidades que queremos.

Veamos que se satisface (6.8). Si se tiene $B^n \geq C^{2m}$ esto implica que $(B+1)^n \geq (C+1)^{2m}$ (es decir (6.8)) ya que $(B+1)^n = \sum_{i=0}^{n} \binom{n}{i} B^n$ y $(C+1)^{2m} = \sum_{i=0}^{2m} \binom{2m}{i} C^{2m}$ y $n > 2m$, alcanza con probar que $B^n \geq C^{2m}$. Por definición de C esto equivale a $B^n \geq 2^{3m}(nAP)^{2m}B^{2m}$ que dividiendo por B^{2m} equivale a $B^{n-2m} \geq 2^{3m}(nAP)^{2m}$. Reemplazando B por su definición se obtiene que el lado izquiero es igual a la expresión de la derecha.

Veamos que se satisface (6.4). Dado j fijo entre 1 y m, tenemos por definición de N_i que

$$\begin{aligned}\sum_{i=1}^{n} N_i a_{ij} &= \sum_{i=1}^{n} (v_i - v_i')a_{ij} = \sum_{i=1}^{n} v_i a_{ij} - \sum_{i=1}^{n} v_i' a_{ij} \\ &= (c_j + id_j) - (c_j' + id_j') = (c_j - c_j') + i(d_j - d_j')\end{aligned}$$

Como P y P' pertenecen a la misma subcelda, la diferencia entre coordenadas homólogas ($|d_j - d_j'|$ o $|c_j + id_j|$ con $j = 1, .., m$) es como mucho $2^{-1/2}P^{-1}$, entonces el módulo del complejo anterior es menor o igual a

$$[2(2^{-1/2}P^{-1})^2]^{1/2} = P^{-1}$$

Veamos que se satisface (6.5). Como P y P' son distintos, también deben serlo los conjuntos de enteros de los que provienen, así que para algún $i = 1, ..., n$ se tiene $1 < v_i - v_i' =: N_i$

Veamos que se satisface (6.6). Como cada elemento de M_1 está formado por enteros no negativos $\leq B$ (en particular v_i y v_i' para cada $i = 1, .., n$) se tiene que $|N_i| = |v_i - v_i'| \leq B$. □

Lema 6.7 (Jensen)**.** *Sea $G(z)$ holomorfa en $|z| \leq R$ con ceros $z_1, z_2, ..., z_n$ en el interior de la bola de radio R, donde cada cero se incluye su multiplicidad de veces. Entonces*

$$|G(0)| \leq R^{-n}|z_1, z_2, ..., z_n| \max_{0 \leq \theta < 2\pi} |G(Re^{i\theta})|$$

Prueba. Observemos que la desigualdad es cierta si $G(0) = 0$, así que podemos suponer $G(0) \neq 0$. Para cada cero z_j consideramos la función

$$B_j(z) = \frac{R(z - z_j)}{R^2 - z\overline{z_j}}$$

que se anula en $z = z_j$, tiene un polo simple en $z = \frac{R^2}{\overline{z_j}}$ fuera de $|z| = R$, y tiene módulo 1 en $|z| = R$. Sea

$$F(z) = G(z) \left[\prod_{j=1}^{n} B_j(z) \right]^{-1}$$

Esta función es holomorfa en $|z| \leq R$ y

$$\max |F(z)| = \max G(z), \qquad |z| = R.$$

Pero por (6.2)

$$F(0) = \frac{1}{2\pi i} \int_{|t|=R} \frac{F(t)}{t}\, dt$$

entonces

$$|F(0)| \leq \max |F(z)| = \max |G(z)|, \qquad |z| = R$$

y

$$G(0) = (-R)^{-n} \left[\prod_{j=1}^{n} z_j \right] F(0).$$

Combinando las últimas 2 expresiones se obtiene la desigualdad del enunciado. □

Obervemos que este resultado nos dice que a mayor número de ceros podamos agregarle a la función $G(z)$ entonces $|G(0)|$ será más chico.

6.2 Prueba de la Conjetura 3.4

Gelfond probó en 1934 la Conjetura 3.4.

6.2.1 Idea de la prueba

Por el enunciado α y β , son algebraicos y definen un cuerpo algebraico $\mathbb{Q}(\alpha, \beta)$. Supongamos que también η es algebraico. Podemos construir entonces el cuerpo de extensión $\mathbb{Q}(\alpha, \beta, \eta)$. Luego definimos una función auxiliar $f(x)$, con valores en $\mathbb{Q}(\alpha, \beta)$ para cada entero racional x. La más simple de tales funcoines sería un polinomio en x con coeficientes en $\mathbb{Q}(\alpha, \beta)$, pero tal función no nos permitiría agregar η. En cambio Gelfond define

$$f(x) = \sum_{k=-q}^{q} \sum_{l=-q}^{q} C_{kl} \alpha^{kx} \beta^{lx} \tag{6.9}$$

donde los C_{kl} son enteros racionales y q es un entero positivo, todos éstos a elegir más adelante. Aquí $\alpha^{kx} = e^{kx \log \alpha}$ con alguna elección fija del logaritmo, y lo mismo para β^{lx}. Esta función toma valores en $\mathbb{Q}(\alpha, \beta)$ cuando x es un entero racional. Además se tiene que

$$f_s(x) := f^{(s)}(x)(\log \beta)^{-s} = \sum_{k=-q}^{q} \sum_{l=-q}^{q} C_{kl} (k\eta + l)^s \alpha^{kx} \beta^{lx} \tag{6.10}$$

lo que implica que $f_s(x)$ toma valores en $\mathbb{Q}(\alpha, \beta, \eta)$ para cada entero x y $s = 0, 1, 2, ...$

La función $f(x)$ tiene $(2q + 1)^2$ parámetros C_{kl}, que junto con el q podemos elegir. En particular podemos tomar q tan grande como queramos. Por el Lema (6.6), los C_{kl} podemos elegirlos de forma tal que un conjunto bastante grande de números $\{|f_s(j)|\}$ sean muy chicos, tan chicos que la cota del Lema (6.5) sea contradicha y entonces deban ser cero.

Dicho de otra forma, podemos imponer sobre $f(x)$ un cierto número de ceros sobre puntos enteros que elijamos, a la vez que mantenemos los coeficientes como enteros acotados. Gelfond distribuyó los ceros entre varios puntos, pero es más simple ubicarlos todos en el origen.

Esta acumulación de ceros en un entorno del origen hará que por el Lema (6.7), $|f(x)|$ tenga que ser chica, digamos para $|x| < q^{2/3}$.

Por la representación como integral de las derivadas de $f(x)$ dada en el Lema (6.2), esta acotación se extenderá también a $|f^{(s)}(x)|$ para valores de s no demasiado grandes. Pero

esto a su vez agrandará el conjunto de de números algebraicos $f_s(j)$, tan chicos que por contradecir la cota del Lema (6.5) deban ser cero.

Entonces tenemos un número adicional de ceros de multiplicidad alta en un entorno del 0, lo que nuevamente hará que por el Lema (6.7), $|f(x)|$ tenga que ser todavía más chica en un entorno más grande.

Usando alternadamente de esta forma los Lemas (6.5) y (6.7) se puede mostrar que $f(x)$ y todas sus derivadas se anulan en todos los enteros, lo cual implica que $f(x) \equiv 0$.

De hecho dos aplicaciones del Lema (6.7) dan suficientes ecuaciones sobre los C_{kl} como para concluir que o bien son todos cero o bien η es racional. La prmera posibilidad está excluida porque por la elección de los C_{kl} mediante el Lema (6.6) se tiene que $\sum\sum |C_{kl}| \geq 1$, mientras que la segunda contradice la hipótesis de que η es irracional, y por ende concluimos que η no puede ser, como inicialmente supusimos, algebraico.

6.2.2 Primer parte: elección de los coeficientes de $f(x)$

Queremos elegir, para poner en (6.6), los números m, n, A, P y obtener como consecuencias del lema la existencia de los C_{kl} y las cotas: (6.5) para usar más adelante, y que (6.4) y (6.6) sean tales que permitan contradecir la cota de (6.5) para $f_s(0)$ para cada $s = 0, 1, .., m-1$ (también debemos elegir este m), y que entonces los $f_s(0)$ deban ser cero.

Intentemos de encontrar alguna motivación de esta estrategia. De aplicar el (6.6) podemos obtener la siguiente cota superior

$$|f_s(0)| = \sum_{k=-q}^{q} \sum_{l=-q}^{q} C_{kl}(k\eta + l)^s \leq \frac{1}{P} \qquad s = 0, 1, ..., m-1 \tag{6.11}$$

lo cual podría usarse para contradecir la cota inferior de (6.5).

Supongamos que η es algebraico de grado h sobre $\mathbb{Q}$ y que E es un denominador para η (es decir E es un entero racional y $E\eta$ es un entero algebraico), entonces se tiene que E^s es un denominador para $f_s(0)$ (es decir $E^s f_s(0)$ es un entero algebraico, en particular de grado $\leq h$ sobre Q), si es que los C^{kl} son enteros racionales.

Como $f_s(0) \in R[\eta]$ sus conjugados se obtienen según (11.14), y entonces usando que entre $-q$ y q hay $(2q+1)$ números, $k, l \leq q$ y $\eta \leq \mu(\eta)$ y suponiendo $|C_{kl}| \leq B$ se obtiene

$$\mu[f_s(0)] \leq B(2q+1)^2\{q[1+\mu(\eta)]\}^s. \tag{6.12}$$

Por otro lado aplicando (6.5) a $f_s(0)$ con E^s se obtiene

$$|f_s(0)| \geq (E^s)^{-y}[\mu(f_s(0))]^{1-y} \qquad y \leq h. \tag{6.13}$$

Reemplazando aquí con (6.12) se tiene

$$|f_s(0)| \geq (E^s)^{-y}[B(2q+1)^2\{q[1+\mu(\eta)]\}^s]^{1-y} \qquad y \leq h. \tag{6.14}$$

Si se tuviera un P tal que

$$P > (E^s)^h[B(2q+1)^2\{q[1+\mu(\eta)]\}^s]^{h-1}, \tag{6.15}$$

reemplazando esa cota en (6.11) se obtiene que

$$|f_s(0)| \leq (E^s)^{-h}[B(2q+1)^2\{q[1+\mu(\eta)]\}^s]^{1-h}, \tag{6.16}$$

Lo cual contradice la cota de (6.14) y obtenemos así la contradicción de la cota de (6.5), lo que implicaría que $f_s(0) = 0$.

Veamos que para que P satisfaga (6.15) alcanza con que P satisfaga

$$P > [B(E_1 q)^m]^h \tag{6.17}$$

donde E_1 es un entero fijo $\geq 2E[1+\mu(\eta)]$, si m es suficientemente grande (en particular es suficiente con que $m \geq h-1$ y $2q \geq \mu(\eta)$).

Estas desigualdades todavía dejan bastante libertad para elegir los parámetros involucrados. Muestran que P debe ser más grande que B y que q^m, pero no muestran cual debe ser mayor entre éstas 2 últimas. Sin embargo la aplicación más adelante de (6.7) requiere que q^m sea mayor.

Las elecciones de Gelfond son

$$B = 3^{q^2}, \qquad m = \lfloor q^2 \frac{\log\log q}{\log q} \rfloor, \qquad P = e^{\gamma q^2 \frac{\log q}{\log\log q}}, \tag{6.18}$$

$$n = (2q+1)^2, \qquad a_{k,l;s} = (k\eta + l)^s, -q \leq k, l \leq q \qquad A = q[1+\mu(\eta)]^{m-1} \tag{6.19}$$

donde γ depende solo de $\mathbb{Q}(\alpha, \beta, \eta)$ y q es un número grande, y donde notamos $\lfloor x \rfloor$ a la parte entera de x.

Veamos que con éstas elecciones se tienen las cualidades necesarias : las hipótesis de (6.6), $|C_{k,l}| \leq B$ y la cota (6.17) para P. Éstas dos ltimas son necesarias para poder hacer el razonamiento anteriormente realizado que permite contradecir la cota obtenida de (6.5) para $f_s(0)$ para cada $s = 0, 1, ..., m-1$.

Para usar en estas pruebas notemos

$$u := \frac{2m}{(n-2m)}, \qquad w := \frac{\log\log q}{\log q} \tag{6.20}$$

y observemos que poniendo $q = e^{e^y}$ y las elecciones se tiene que

$$w = \frac{y}{e^y} \to 0, \qquad w^2 \log q = \frac{y^2}{e^y} \to 0, \qquad uw \log q = \frac{w^2 \log q}{2 + \frac{1+4q}{4q^2} - w} \to 0, \qquad q \to \infty \tag{6.21}$$

y también por (6.21) se tiene que

$$wu = \frac{w^2}{2 + \frac{1+4q}{4q^2} - w} \to 0, \qquad u = \frac{1}{\frac{4q^2+1+4q}{4q^2 w} - 1} \to 0 \qquad q \to \infty \tag{6.22}$$

Veamos que se satisfacen las hipótesis de (6.6).

Calculemos la cota para los elementos de la matriz. Usando que $l, k \leq q$, $\eta \leq \mu(\eta)$ y $s \leq m-1$ se tiene

$$|a_{k,l;s}| = |(k\eta + l)^s| \leq (q\mu(\eta) + q)^{m-1} = A \tag{6.23}$$

tenemos $m < 2n$, ya que por definiciones de m y n esto equivale a

$$\left[q^2 \underbrace{\frac{\log\log q}{\log q}}_{=m<1} \right] < q^2 < 2 \underbrace{(2q+1)^2}_{=n}. \tag{6.24}$$

Veamos que se cumple (6.17). Poniendo las definiciones de B y P y que $m = q^2 w$ ella equivale a

$$e^{\gamma q^2 w^{-1}} > 3^{q^2 h} (E_1 q)^{q^2 w h} = e^{[q^2 h \log 3 + q^2 w h \log(E_1 q)]} \tag{6.25}$$

que equivale a

$$\gamma q^2 w^{-1} > 3^{q^2 h} (E_1 q)^{q^2 w h} = [q^2 h] \log 3 + q^2 w h \log(E_1 q). \tag{6.26}$$

Reordenando, por (6.21) se tiene

$$\gamma > h[w log 3 + w^2 log(E_1 q)] \to 0, \qquad q \to \infty. \tag{6.27}$$

Veamos que se satisface $|C_{k,l}| \leq B$. La definición de B junto a la consecuencia (6.6) del Lema 6.6 ella que equivale a

$$|C_{k,l}| \leq [2^{3/2} \underbrace{(2q+1)^2}_{n} \underbrace{q[1+\mu(\eta)]^{q^2 w - 1}}_{A} \underbrace{e^{\gamma q^2 \frac{\log q}{\log\log q}}}_{P}] = e^{[3/2 \log 2 + 2\log(2q+1) + (m-1)\log q + \gamma q^2 w]u} \leq 3^{q^2}. \tag{6.28}$$

en particular lo anterior será cierto si

$$[3/2 \log 2 + 2\log(2q+1) + (q^2 w - 1)\log q + \gamma q^2 w] u \leq q^2 \tag{6.29}$$

que equivale a

$$u q^{-2} [3/2 \log 2 + 2\log(2q+1) - \log q] + [uw \log q + \gamma u w] \leq 1. \tag{6.30}$$

Con (6.21) y (6.22) se ve que la expresión de la izquierda tiende a cero cuando q tiende a infinito.

Verificadas las cualidades requeridas, entonces tenemos que es posible elegir C_{kl}, no todos nulos, tales que $|C_{kl}| \leq 3^{q^2}$, y $s = 0$ siendo un cero de $f(x)$ definida por (6.9) con multiplicidad por lo menos m, y que además

$$|f(x)| \leq \sum_{k=-q}^{q} \sum_{l=-q}^{q} |C_{kl}| |e^{kx \log\alpha + lx \log\beta}| < 3^{q^2} (2q+1)^2 e^{\delta q |x|} = e^{q^2 log 3 + 2\log(2q+1) + \delta q |x|} \leq e^{2q^2 + \delta q |x|} \tag{6.31}$$

para $g \geq 2$, donde $\delta = |\log\alpha| + |\log\beta|$ depende sólo de α y β .

6.2.3 Segunda parte: Reconocimiento de ceros adicionales de $f(x)$

A continuación aplicaremos por primera vez el Lema 6.7. Sea $G(z) = f(a+z)$ donde a es un punto cualquiera de la circunsferencia $|x| = q^{2/3}$ (puede usarse cualquier potencia de q menor a 1). $G(z)$ tiene un cero de multiplicidad m en $z = -a$. Consideramos un disco con centro en $z = 0$ de radio q. Aplicando (6.7) obtenemos

$$|f(a)| = |G(0)| \leq q^{-m}(q^{2/3})^m \max_{0\leq\theta<2\pi} |G(qe^{i\theta})|, \tag{6.32}$$

luego, poniendo la cota (6.31) para $f(a+z)$ en la circunsferencia $|z| = q$, y usando la definición de m obtenemos

$$|f(a)| = |G(0)| \leq q^{-m/3} e^{[2q^2+\delta q(q+q^{2/3})]} = e^{[-q^2/3\log\log q]} e^{[(1+\delta)q^2+q^2+\delta q^{5/3}]} < e^{-q^2/3\log\log q+2(1+\delta)q^2} \tag{6.33}$$

para q suficientemente grande. Escribiendo $1/3 = 1/6 + 1/6$, como δ no depende de q

$$= \exp\left(-q^2/6\log\log q + \underbrace{q^2[2(1+\delta) - 1/6\log\log q]}_{\to-\infty, q\to\infty}\right) < e^{-q^2/6\log\log q} \tag{6.34}$$

para q suficientemente grande. Por 6.4 se tiene que

$$|f(x)| < e^{-1/6q^2\log\log q}, \qquad |x| \leq q^{2/3} \tag{6.35}$$

para q suficientemente grande. Por (6.2) se tiene

$$f^{(s)}(x) = \frac{s!}{2\pi i}\int_C \frac{f(t)}{(t-z)^{n+1}}\,dt \tag{6.36}$$

sobre el disco $|t| = q^{2/3}$. Considerando $|x| \leq 1/2q^{2/3}$, acotando $s!$ por s^s y f por (6.35) obtenemos

$$|f^{(s)}(x)| \leq s^s \frac{1}{2\pi i} 2\pi i q^{2/3} \frac{1}{(q^{2/3} - 1/2q^{2/3})^{s+1}} e^{-1/6q^2\log\log q} = 2(2sq^{-2/3})^s e^{-1/6q^2\log\log q} \tag{6.37}$$

Si $0 \leq s \leq \frac{q^2}{\log q}$, usando esta cota para s, distribuyendo el logaritmo en los productos y potencias, escribiendo $1/6 = 1/12 + 1/12$ tenemos que

$$|f^{(s)}(x)| \leq e^{-1/12q^2\log\log q + w} \tag{6.38}$$

donde

$$w = \left[\log 2 + \frac{q^2}{\log q}(\log 2 + 4/3\log q - \log\log q) - 1/12q^2\log\log q\right]. \tag{6.39}$$

Sacando factor común q, escribiendo $q = e^{e^y}$ y haciendo tender q a infinito (y tiende a infinito) se observa que w tiende a $-\infty$, con lo cual tenemos que

$$|f^{(s)}(x)| \leq e^{-1/12q^2\log\log q}, \qquad |x| \leq 1/2q^{2/3}, \qquad 0 \leq s \leq \frac{q^2}{\log q} \tag{6.40}$$

para q suficientemente grande. Entonces la expresión (6.10) implica lo siguiente.

Si $|\log\beta| \geq 1$, como $s \in \mathbb{N}$ se tiene $|\log\beta|^{-s} \leq 1$, y en consecuencia usando la cota de (6.40), y luego que e^x es creciente

$$|f_s(j)| \leq |f^{(s)}(j)| \leq e^{-1/12q^2 \log\log q} \leq e^{-1/24q^2 \log\log q}. \tag{6.41}$$

Si $|\log\beta| < 1$ se tiene $|\log\beta|^{-1} \geq 1$ entonces usando las cotas para s y para $f^s(x)$ dadas en (6.40)

$$|f_s(j)| \leq f^{(s)}(j)(|\log\beta|^{-1})^{\frac{q^2}{\log q}} \leq \exp\left(-1/24q^2\log\log q + \underbrace{[-1/24q^2\log\log q - \frac{q^2}{\log q}\log|\log\beta|]}_{\to-\infty, q\to+\infty}\right). \tag{6.42}$$

Entonces de (6.41) y (6.42) se tiene que

$$|f_s(j)| < e^{-1/24q^2\log\log q} \tag{6.43}$$

para $j = 0, \pm1, \pm2, ...\lfloor\pm q^{2/3}\rfloor$, $s = 0, 1, 2, ..., \lfloor\frac{q^2}{\log q}\rfloor$, y q suficientemente grande.

Veamos que (6.43) implica que $f_s(j) = 0$ por el Lema 6.5. Por definición tenemos

$$f_s(j) = \sum_{k=-q}^{q}\sum_{l=-q}^{q} C_{kl}(k\eta + l)^s \alpha^{kj}\beta^{lj}. \tag{6.44}$$

Eligiendo enteros racionales A, B, C, D, E tales que $A\alpha, B\beta, C/\alpha, D/\beta, E\eta$ sean enteros algebraicos, tenemos que $(ABCD)^{jq}E^s f_s(j)$ es un entero en el cuerpo $\mathbb{Q}(\alpha, \beta, \eta)$. Supongamos que el grado de éste es H, se tiene que H es como mucho igual al producto de los grados de α, β, η sobre $\mathbb{R}$. Por la forma en la que se obtienen los conjugados de $f_s(j)$, enunciada en (11.14), tenemos a partir de (6.44) que

$$\mu[f_s(j)] \leq 3^{q^2}(2q+1)^2\{q[1+\mu(\eta)]\}^s\lambda^{2jq} \tag{6.45}$$

donde $\lambda = \max\{\mu(\alpha), \mu(\beta), \mu(1/\alpha), \mu(1/\beta)\}$. Entonces si $f_s(j) \neq 0$, por (6.5) se tiene que

$$|f_s(j)| \geq [(ABCD)^{jq}E^s]^{-H}\mu[f_s(j)]^{1-H} = e^p \tag{6.46}$$

donde

$$p = (-H)[jq\log(ABCD) + s\log E] + (1-H)\log\mu[f_s(j)] \tag{6.47}$$

Poniendo la cota (6.40) de s y la (6.45) de $\mu[f_s(j)]$ se tiene

$$p \leq (-H)[jq\log(ABCD) + \frac{q^2}{logq}\log E]$$

$$+(1-H)[q^2\log 3 + 2\log(2q+1) + s[\log q + \log(1+\mu(\eta))] + 2jq\log\lambda$$

Poniendo nuevamente la cota (6.40) de s y que $j \leq q$ tenemos que

$$p \leq (-H)[q^2 \log(ABCD) + \frac{q^2}{logq} \log E]$$

$$+(1-H)[q^2 \log 3 + 2\log(2q+1) + \frac{q^2}{\log q}[\log q + \log(1+\mu(\eta))] + 2q^2 \log \lambda$$

Como en esta expresión todos los términos tienen orden q^2 se tiene que

$$p \leq -\sigma q^2 \tag{6.48}$$

para q suficientemente grande, donde σ no depende de q. Poniendo esto en (6.46) se obtiene que

$$|f_s(j)| \geq e^{-\sigma q^2}. \tag{6.49}$$

Pero esto contradice (6.43), y contrario a como se supuso debe tenerse $f_s(j) = 0$.

Entonces $f(x)$ tiene un cero en cada entero j con $|j| \leq 1/2q^{2/3}$ y cada cero tiene multiplicidad por lo menos $q^2/logq$. Recurriremos nuevamente al Lema 6.7 poniendo $G(z) = f(a+z), |a| = q^{4/3}, R = q^{3/2}$ y usando que entonces todo cero conocido dentro del disco $|z| = R$ satisface $|z| < 2q^{4/3}$, y que se tienen por lo menos (con multiplicidad) $v := \frac{q^2}{\log q} 1/2q^{2/3} = 1/2q^{8/3}/logq$ ceros , y que por (6.31) $\max_{|z|=R} G(z+a) \leq e^{2q^2+\delta q+q^{4/3}+q^{3/2} \leq e^{2(1+\delta)q^{5/2}}}$ para q grande se tiene que

$$|f(a)| \leq |G(0)| \leq q^{-3/2v}(2q^{4/3})^v e^{2(1+\delta)q^{5/2}} = e^y \tag{6.50}$$

con

$$y = -1/6v \log q + v \log 2 + 2(1+\delta)q^{5/2} \tag{6.51}$$

Poniendo la definición de v

$$y = -1/12q^{8/3} + \underbrace{1/2q^{8/3}/logq \log 2}_{<0} + 2(1+\delta)q^{5/2} \leq -1/12q^{8/3} + 2(1+\delta)q^{5/2} = \tag{6.52}$$

$$-1/24q^{8/3} + \underbrace{[2(1+\delta)q^{5/2} - 1/24q^{8/3}]}_{\to -\infty, q \to +\infty} \leq -1/24q^{8/3}$$

para q grande. Poniendo esta cota en (6.50) obtenemos

$$|f(a)| \leq e^{-1/24q^{8/3}}, \qquad |a| = q^{4/3} \tag{6.53}$$

para q suficientemente grande. Y por 6.4

$$|f(x)| \leq e^{-1/24q^{8/3}}, \qquad |x| \leq q^{4/3}. \tag{6.54}$$

6.2.4 Tercer parte: racionalidad de η

Observemos que usando la fórmula (6.2) para $f^{(s)}(0)$ (como se hizo antes en (6.35)) con $|t| = q^{2/3}s! \leq s^s$, y la cota (6.54) con $x = 0$, se tiene

$$|f^{(s)}(0)| = s^s \frac{1}{2\pi} 2\pi q^{2/3} e^{-1/24q^{8/3}} \cdot \frac{1}{[q^{2/3}]^{s+1}} = \frac{s^s}{q^{\frac{2}{3}s}} e^{-1/24q^{8/3}}. \tag{6.55}$$

Utilizando que $0 \leq s \leq q^{5/2}$, acotamos (6.55) por

$$q^{\frac{11}{6}q^{5/2}} e^{-1/24q^{8/3}} = \exp\left(-1/48q^{8/3} + \underbrace{5/2 \log(\frac{11}{6}q) - 1/48q^{8/3}}_{\to -\infty, q \to +\infty} \right) \leq e^{-1/48q^{8/3}} \tag{6.56}$$

para q suficientemente grande. Es decir

$$|f^{(s)}(0)| \leq e^{-1/48q^{8/3}} \tag{6.57}$$

para $0 \leq s \leq q^{5/2}$ y q suficientemente grande. Con esta cota se obtiene, de igual modo como se obtuvo antes (6.43) de (6.40), lo siguiente

$$|f_s(0)| = |f^s(0)(\log b)^{-s}| \leq e^{-1/96q^{8/3}} \qquad 0 \leq s \leq q^{5/2}. \tag{6.58}$$

Si tomamos en (6.11) P^{-1} igual al lado derecho de (6.58) veremos que (6.15) se satisface y que entonces la cota de (6.5) se contradice, con lo que debe ser $f^{(s)}(0) = 0$ para $s \leq q^{5/2}$.

Verifiquemos esto. Poniendo en (6.15) $P^{-1} = e^{-1/96q^{8/3}}$ y luego que $s \leq q^{5/2}$ se tiene que la acotación es equivalente a

$$1 > P^{-1} E^{sh} [B(2q+1)^2 (q(1+\mu(\eta)))^s]^{h-1} = e^y \tag{6.59}$$

con

$$y \leq -1/96q^{8/3} + q^{5/2} h \log E + (h-1)[\log B + 2\log(2q+1) + q^{5/2} \log q + q^{5/2} \log(1+\mu(\eta))] \tag{6.60}$$

Vemos que aquí el término de mayor peso es $-1/96q^{8/3}$, con lo que $y \to -\infty$ cuando q tiende a infinito. Entonces se verifica (6.59) para q suficientemente grande.

El hecho de que $f^{(s)}(0) = 0$ para $s \leq q^{5/2}$ da un sistema de $\lfloor q^{5/2} \rfloor$ ecuaciones lineales

$$\sum_{k=-q}^{q} \sum_{l=-q}^{q} (k\eta + l) C^{kl} = 0 \tag{6.61}$$

en los $(2q+1)^2$ coeficientes C_{kl}. Como se sabe que éstos no son todos nulos debido a la construcción de $f(x)$ por la consecuencia (6.5), cualquier determinante de $(2q+1)^2$ filas debe ser cero. En particular

$$\det |(k\eta + l)^s| = 0 \qquad -q \leq k, l \leq q, 0 \leq s \leq 4q(q+1). \tag{6.62}$$

Pero como esta es una matriz del tipo (6.1) que se anula si y solo si 2 columnas son iguales, esto implica que existen $k, k', l, l' \in \mathbb{Z}$ tales que $k\eta + l = k'\eta + l'$, pero esto dice que η es racional, contradiciendo la suposición inicial de que no lo era, con lo cual concluye la prueba de Gelfond.

7 Solución de Schneider

7.1 Preliminares

7.1.1 Un lema de Siegel para los cuerpos de números

Para la prueba de Schneider se necesiterá encontrar una solución de un sistema homogéneo de ecuaciones lineales con coeficientes en un cuerpo K, que no sea demasiado grande. Para ello se utilizará el lema de Siegel que exponemos a continuación, cuya demostración se basa en el siguiente lema de Dirichlet

Lema 7.1 (Dirichlet). *Si $\varphi : E \to F$ es una función de un conjunto E de n elementos a un conjunto $F = \bigcup_{1\leq j\leq m} F_j$ y si $m < n$, entonces por lo menos uno de los conjuntos $F_1, ..., F_m$ contiene las imágenes por φ de dos elementos distintos de E. O dicho más brevemente, que una función de un conjunto de n elementos en otro de m elementos, no es inyectiva si $m < n$.*

Enunciamos el lema de Siegel que queremos probar.

Lema 7.2 (Siegel). *Sea K un cuerpo de números, de grado δ sobre $\mathbb{Q}$. Sean $a_{i,j}$ $(1 \leq i \leq n, 1 \leq j \leq m)$ elementos de K enteros sobre $\mathbb{Z}$. Sean $\sigma_1, ..., \sigma_\delta$ los diferentes isomorfismos de K sobre $\mathbb{C}$, y sa A un entero racional que verifica*

$$A \geq \max_{1\leq j\leq m, 1\leq h\leq \delta} \sum_{i=1}^{n} |\sigma_h(a_{i,j})|. \tag{7.1}$$

Si $n > \delta m$ entoces el sistema

$$\sum_{i=1}^{n} a_{i,j}x_i = 0 \qquad 1 \leq j \leq m \tag{7.2}$$

admite una solución no nula $(x_1, ..., x_n) \in \mathbb{Z}^n$ que verifica

$$\max_{1\leq i\leq n} |x_i| \leq (\sqrt{2}A)^{\frac{m\delta}{n-m\delta}}. \tag{7.3}$$

Observemos que para solucionar un sistema

$$\sum_{i=1}^{n} a_{i,j}x_i = 0 \qquad 1 \leq j \leq m$$

con coeficientes en K, podemos reducir al caso en el que los $a_{i,j}$ son enteros sobre $\mathbb{Z}$ multiplicando la ecuación $j-$ésima por un denominador común d_j de $a_{1,j}, ..., a_{n,j}$. Alcanza entonces con que reemplacemos A por $A \max_{1\leq j\leq m} d_j$.

Para demostrar (7.2) usaremos el siguiente lema que permite solucionar un sistema de inecuaciones lineales.

Lema 7.3. *Sean $u_{i,j}\,(1 \leq i \leq v, 1 \leq j \leq \mu)$ números reales; sea U un úmero entero que verifica*

$$\max_{1\leq j\leq \mu} \sum_{i=1}^{v} |u_{i,j}| \leq U \tag{7.4}$$

y sean X y l dos números enteros positivos tales que

$$l^{\mu} < (X+1)^{v}. \tag{7.5}$$

Entonces existen elementos $\epsilon_1, ..., \epsilon_v$ de $\mathbb{Z}$ no todos nulos tales que

$$\max_{1\leq i\leq v} |\epsilon_i| \leq X \tag{7.6}$$

y

$$\max_{1\leq j\leq \mu} |\sum_{i=1}^{v} u_{i,j}\epsilon_i| \leq \frac{UX}{l} \tag{7.7}$$

Prueba. Consideremos la función φ del conjunto

$$N(v,X) := \{(\epsilon_1, ..., \epsilon_v) \in \mathbb{Z}^v : 0 \leq \epsilon_i \leq X (1 \leq i \leq v)\}$$

en $\mathbb{R}^{\mu}$, que manda $(\epsilon_1, ..., \epsilon_v)$ a $(\eta_1, ..., \eta_v)$ con

$$\eta_j = \sum_{i=1}^{v} u_{i,j}\epsilon_i \qquad 1 \leq j \leq \mu. \tag{7.8}$$

Para $1 \leq j \leq \mu$ notamos $-V_j$ (respectivamente W_j) a la suma de elementos negativos (respectivamente positivos) del conjunto

$$u_{1,j}, ..., u_{v,j}. \tag{7.9}$$

Tendremos entonces

$$V_j + W_j \leq U \qquad j = 1, ..., \mu. \tag{7.10}$$

Observemos que si $(\epsilon_1, ..., \epsilon_v) \in N(v,X)$ entonces la imagen $(\eta_1, ..., \eta_\mu) = \varphi(\epsilon_1, ..., \epsilon_v)$ pertenece al conjunto

$$E := \{(\eta_1, ..., \eta_\mu) \in \mathbb{R}^{\mu}; -V_j X \leq \eta_j \leq W_j X\}. \tag{7.11}$$

Partimos cada uno de los intervalos $[-V_j X, W_j X]$ en l intervalos (de longitud $\leq \frac{UX}{l}$), lo cual parte a E en l^{μ} subconjuntos $E_k\,(1 \leq k \leq l^{\mu})$. La condición

$$l^{\mu} < (X+1)^{v} = |N(v,X)| \tag{7.12}$$

permite aplicar (7.1): existen dos elementos distintos ϵ' y ϵ'' de $N(v,X)$ tales que sus imágenes por φ pertenecen al mismo suconjunto E_k de E. Notemos ϵ la diferencia $\epsilon'-\epsilon''$ y η a $\varphi(\epsilon)$. Tendremos

$$\epsilon=(\epsilon_1,...,\epsilon_v)\neq 0, \qquad \max_{1\leq i\leq v}|\epsilon_i|\leq X \tag{7.13}$$

y

$$\eta=(\eta_1,...,\eta_\mu), \qquad \max_{1\leq j\leq \mu}|\eta_j|\leq \frac{UX}{l} \tag{7.14}$$

de donde se obtiene el Lema 7.3 □

Estamos entonces en condiciones de demostrar el Lema 7.2.
Numeremos los diferentes morfismos $\sigma_1,...,\sigma_\delta$ de K en $\mathbb{C}$ de tal forma que tenemos

$$\sigma_h(K)\subset\mathbb{R} \qquad 1\leq h\leq r \tag{7.15}$$

y

$$\sigma_{r+s+k}=\overline{\sigma_{r+k}} \qquad 1\leq k\leq s \tag{7.16}$$

(donde la barra significa conjugado de complejo de δ_{r+k}) y r y s son dos enteros que verifican $\delta=r+2s$. Defininos las funciones $t_1,...,t_\delta$ de K a $\mathbb{R}$ por

$$t_h=\begin{cases}\sigma_h & 1\leq h\leq r\\ \mathrm{Re}\sigma_h & r+1\leq h\leq r+s\\ \mathrm{Im}\sigma_h & r+s+1\leq h\leq \delta=r+2s\end{cases}$$

Elegimos dos enteros X y l

$$X=\lfloor(\sqrt{2}A)^{\frac{m\delta}{n-m\delta}}\rfloor \qquad l=1+=\lfloor\sqrt{2}AX\rfloor \tag{7.17}$$

de forma tal que

$$X\leq(\sqrt{2}A)^{\frac{m\delta}{n-m\delta}} \tag{7.18}$$

y

$$(1+X)^{n-m\delta}>(\sqrt{2}A)^{m\delta}. \tag{7.19}$$

Entonces, como $A\geq 1$

$$(1+X)^n\geq(1+\sqrt{2}AX)^{m\delta}\geq l^{m\delta}. \tag{7.20}$$

El Lema 7.3 (con $v=n\,,\mu=m\delta\,,U=A$) muestra que existen enteros racionales $x_1,...,x_n$ no todos nulos que verifican

$$\max_{1\leq i\leq n}|x_i|\leq(\sqrt{2}A)^{\frac{m\delta}{n-m\delta}} \tag{7.21}$$

y

$$\max_{1\le h\le\delta,1\le j\le m}|\sum_{i=1}^{n}t_h(a_{i,j})x_i|\le\frac{AX}{1+\lfloor\sqrt{2}AX\rfloor}. \tag{7.22}$$

Deducimos que

$$\max_{1\le h\le r,1\le j\le m}|\sum_{i=1}^{n}\sigma_h(a_{i,j})x_i|\le\frac{AX}{1+\lfloor\sqrt{2}AX\rfloor}, \tag{7.23}$$

y

$$\max_{r+1\le h\le\delta,1\le j\le m}|\sum_{i=1}^{n}\sigma_h(a_{i,j})x_i|\le\frac{\sqrt{2}AX}{1+\lfloor\sqrt{2}AX\rfloor} \tag{7.24}$$

de donde

$$|N_{K/Q}(\sum_{i=1}^{n}a_{i,j}x_i)|\le 2^s(\frac{AX}{1+\lfloor\sqrt{2}AX\rfloor})^{\delta}. \tag{7.25}$$

En esta última desigualdad, el miembro izquierdo es un entero racional y el derecho está acotado superiormente (ya que $s\le\frac{\delta}{2}$) por

$$(\frac{\sqrt{2}AX}{1+\lfloor\sqrt{2}AX\rfloor})^{\delta}<1. \tag{7.26}$$

De donde

$$\sum_{i=1}^{n}a_{i,j}x_i=0\qquad 1\le j\le m, \tag{7.27}$$

con lo que queda probado el Lema 7.2 □

7.1.2 Funciones complejas

Sean f y g dos funciones reales de variable real, notamos "$f<<g$" si existen dos números reales positivos A y C tales que $x>A$ implica que $f(x)\le Cg(x)$.

Sea ρ un real positivo y f una función entera (es decir holomorfa de $\mathbb{C}$ en $\mathbb{C}$). Diremos que f es "de orden (estricto) inferior o igual a ρ" si

$$\log|f|_R=\log\sup_{|z|=R}|f(z)|<<R^{\rho}\qquad R\to\infty. \tag{7.28}$$

Una función meromorfa se dice "de orden inferior o igual a ρ" si ella es el cociente de 2 funciones enteras de orden inferior o igual a ρ.

Ejemplos. Una fracción racional es de orden menor o igual a ρ para cualquier $\rho>0$. Las funciones seno, coseno, exponenciales son de orden inferior o igual a 1. Si $n\in\mathbb{Z},n>0$ la función e^{z^n} es de orden inferior o igual a n. Si f es una función par ($f(-z)=f(z)\forall z\in\mathbb{C}$) de orden inferior o igual a ρ, entonces $f(\sqrt(z))$ es de orden inferior o igual a $\rho/2$. Por otro lado la función e^{e^z} no es de orden finito.

Para obtener información de cierta precisión sobre los ceros de funciones enteras, consideraremos la siguiente fórmula.

Teorema 7.4 (Jensen)**.** *Sea f una función holmorfa no nula sobre un disco abierto de centro 0 y radio $R > 0$. Sea r un número real $0 < r < R$, sean $a_1, ..., a_p$ los ceros no nulos de f (repetidos según sus multiplicidades) en el disco $|z| \leq r$ y sea $c_k z^k$ ($k \geq 0$ entero) el primer término del desarrollo en potencias de f en cero. Entonces tenemos que*

$$\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})|\, d\theta = \log|c_k| + k\log r + \sum_{h=1}^{p} \log\frac{r}{|a_h|}. \tag{7.29}$$

La fórmula de Jensen muestra que si f es una función entero no nula de orden inferior o igual a ρ entonces el número $n(f, R)$ de ceros de f en el disco $|z| < R$ verifica

$$n(f, R) << R^{\rho} \qquad R \to +\infty. \tag{7.30}$$

Una consecuencia de esto que utilizaremos será la siguiente:

Lema 7.5. *Si f es una función meromorfa no nula de orden inferior o igual a ρ, y si $x_1, ..., x_n$ son números complejos $\mathbb{Q}-$linealmente independientes, con $n > \rho$, entonces por lo menos uno de los números*

$$f(k_1x_1 + ... + k_nx_n), \qquad k_i \in \mathbb{Z}, \qquad k_i \geq 1 \qquad (1 \leq i \leq n) \tag{7.31}$$

es no nulo.

Sea $f_0 : U \to \mathbb{C}$ la función identidad sobre U, y sea $K = \mathbb{C}(f_0)$ (que lo escribimos tambén $K\mathbb{C}(z)$). Decimos que una función meromorfa $f : U \to \mathbb{C}$ es "algebraica"(respectivamente "trascendente") si f es algebraico sobre K (respectivamente trascendente sobre K), es decir si existe (respectivamente si no existe) un polinomio no nulo $P \in \mathbb{C}[X_1, X_2]$ tal que

$$P(z, f(z)) = 0 \qquad \forall z \in U \tag{7.32}$$

Tenemos el siguiente resultado.

Lema 7.6. *Sean $b_1, ..., b_h \in \mathbb{C}$. Las funciones enteras*

$$z, e^{b_1 z}, ..., e^{b_h z} \tag{7.33}$$

son algebraicamente independientes sobre $\mathbb{C}$ si y solo si los números $b_1, ..., b_h$ son $\mathbb{Q}$ linealmente independientes.

Prueba. La relación

$$\lambda_1 b_1 + ... + \lambda_h b_h = 0 \qquad \lambda_j \in \mathbb{Z} \qquad 1 \leq j \leq h$$

implica que

$$e^{b_1 z\lambda_1}...e^{b_h z\lambda_h} = 1 \qquad \forall z \in \mathbb{C}.$$

Supongamos que los números $b_1, ..., b_h$ son $\mathbb{Q}$ -linealmente independientes, y sea

$$P \in \mathbb{C}[X_0, ..., X_h] \tag{7.34}$$

un polinomio no nulo. Alcanza con probar que la función entera

$$F : z \to P(z, e^{b_1 z}, ..., e^{b_h z}) \tag{7.35}$$

no es la función nula.

Escribamos el polinomio P en la forma

$$P(X_0, ..., X_h) = \sum_{\lambda_0=0}^{\delta_0} ... \sum_{\lambda_h=0}^{\delta_h} p_{\lambda_0,...,\lambda_h} X_0^{\lambda_0} ... X_h^{\lambda_h}, \tag{7.36}$$

entonces tenemos que

$$F(z) = \sum_{(\lambda)} p(\lambda) z^{\lambda_0} e^{(\lambda_1 b_1 + ... + \lambda_h b_h) z} \tag{7.37}$$

donde notamos $(\lambda) = (\lambda_0, ..., \lambda_n)$.

Los números

$$\lambda_1 b_1 + ... + \lambda_h b_h, \qquad 0 \leq \lambda_j \leq \delta_j \qquad 1 \leq j \leq h \tag{7.38}$$

son dos a dos distintos, notémoslos

$$w_1, ..., w_q \qquad q = (\delta_1 + 1)...(\delta_h + 1). \tag{7.39}$$

Entonces podemos escribir la función F en la forma

$$F(z) = \sum_{i=1}^{p} \sum_{j=1}^{q} a_{i,j} z^{i-1} e^{w_j z} \tag{7.40}$$

donde $p = \delta_0 + 1$ y $a_{i,j} \quad (1 \leq i \leq p, 1 \leq j \leq q)$ son números complejos no todos nulos, ya que P no es identicamente nulo. De donde se sigue el resultado aplicando el Lema 7.7 siguiente. □

Lema 7.7. *Sean $P_1, ..., P_q$ polinomios no nulos de $\mathbb{C}[X]$, sean $w_1, ..., w_q$ números complejos dos a dos distintos. Entonces la función entera*

$$F : z \to \sum_{k=1}^{q} P_k(z) e^{w_k z} \tag{7.41}$$

no es idénticamente nula.

Prueba. Lo probaremos por inducción sobre q. El caso $q = 1$ es inmediato. Supongamos que $q > 1$ y notemos p_i al grado del polinomio $P_i \qquad (1 \leq i \leq q)$. Observemos que existen polinomios $Q_1, ..., Q_{q-1}$ de $\mathbb{C}[X]$ de grados respectivos $p_1, ..., p_{q-1}$ tales que

$$\frac{d^{p_q+1}}{dz^{p_q+1}} e^{-wq^z} F(z) = \sum_{j=1}^{q-1} Q_j(z) e^{(w_j - w_q) z}. \tag{7.42}$$

Por la hipótesis inductiva el lado derecho no es idénticamente nulo, y entonces F tampoco. □

7.2 Prueba de 3.2

La demostración que exponemos a continuación es una versión simplificada por Lang, de aquella dada por Schneider. La diferencia esencial reside en la construcción de un número $\gamma_N \neq 0$ que luego se acota superior e inferiormente de modo contradictorio. Mientras que en la primer prueba de Schneider éste se obtiene como el determinante de una matriz, en la de Lang se obtiene como un valor de una función F_N cuya existencia se prueba usando el Lema de Siegel 7.2.

Demostraremos 3.2, es decir

Teorema 7.8 (1934 Schneider)**.** *Sean $l \neq 0, b \notin \mathbb{Q}$ números complejos. Entonces por lo menos uno de los tres números*

$$a := e^l, \qquad b, \qquad a^b = e^{bl} \tag{7.43}$$

es trascendente.

Prueba. La demostración es por el absurdo, supongamos que los tres números complejos

$$b, e^l, e^{bl}$$

son algebraicos, con $l \neq 0$ y b irracional. Observemos que las dos funciones

$$z, e^{lz}$$

que son algebraicamente independientes debido a la condición $l \neq 0$ y el Lema 7.6, y toman valores en el cuerpo

$$K = \mathbb{Q}(b, e^l, e^{bl})$$

para $z = i + jb$, $(i, j) \in \mathbb{Z}^2$.

Sea $\delta = [\mathbb{K} : \mathbb{Q}]$ y sea d un denominador común de los tres números

$$b, e^l, e^{bl}.$$

Tomemos un entero N suficientemente grande (es decir acotado inferiormente por un número finito de desigualdades que escribiremos). Podremos suponer que $N^{1/2}$ es entero.

Mostraremos a continuación que existe un polinomio no nulo

$$P_N \in \mathbb{Z}[X_1, X_2]$$

de grado menor a $N^{3/2}$ respecto a X_1, de grado menor a $2\delta N^{1/2}$ respecto a X_2, y de tamaño menor o igual a $2N^{3/2} \log N$, tal que la función

$$F_N(z) = P_N(z, e^{lz})$$

verifica que

$$F_N(i + jb) = 0 \qquad i = 1, ..., N\,; j = 1, ..., N.$$

Para obtener este resultado escribimos al polinomio desconocido P_N en la forma

$$P_N(X_1, X_2) = \sum_{\lambda=0}^{N^{3/2}-1} \sum_{\mu=0}^{2\delta N^{1/2}-1} p_{\lambda,\mu}(N) X_1^{\lambda} X_2^{\mu}$$

con $p_{\lambda,\mu}(N) \in \mathbb{Z}$ y consideramos el sistema de ecuaciones en $p_{\lambda,\mu}(N)$

$$d^{(4\delta+1)N^{3/2}} F_N(i+jb) = 0, \qquad 1 \le i \le N\,; 1 \le j \le N \tag{7.44}$$

es decir

$$\sum_{\lambda=0}^{N^{3/2}-1} \sum_{\mu=0}^{2\delta N^{1/2}-1} p_{\lambda,\mu}(N)(di+djb)^{\lambda}(de^{l})^{i\mu}(de^{bl})^{j\mu} d^{(4\delta+1)N^{3/2}-\lambda-i\mu-j\mu} = 0, \qquad 1 \le i \le N; 1 \le j \le N. \tag{7.45}$$

Obtenemos así un sistema de N^2 ecuaciones en $2\delta N^2$ incógnitas, a coeficientes en K enteros sobre $\mathbb{Z}$, estos coeficientes tienen un tamaño mayorado por

$$N^{3/2}\log N + N^{3/2}[(8\delta+2)\log d + log(1+|\bar{b}|) + 2\delta \log|e^{\overline{l+bl}}|] \le \frac{3}{2} N^{3/2}\log N \tag{7.46}$$

$$1 \le i \le N, 1 \le j \le N$$

debido al Lema 11.20. El Lema 7.2 permite encontrar números enteros racionales

$$P_{\lambda,\mu}(N) \qquad 0 \le \lambda \le N^{3/2}-1 \qquad 0 \le \mu \le 2\delta N^{1/2}-1 \tag{7.47}$$

no todos nulos que verifican

$$\log \max_{\lambda,\mu} |P_{\lambda,\mu}(N)| \le 2N^{3/2}\log N \tag{7.48}$$

(notar que el exponente $\frac{m\delta}{n-m\delta}$ del Lema 7.2 es aquí igual a 1), y tal que la función F_N verifica

$$F_N(i+jb) = 0 \qquad (1 \le i \le N\,; 1 \le j \le N) \tag{7.49}$$

Las condiciones $P_N \neq 0$ y $l \neq 0$ muestran que la función F_N no es nula, así F_N es una función entera de orden menor o igual a 1 ya que

$$\log|F_N|_R \le 2\delta N^{1/2}|l|R + N^{3/2}\log R + 2N^{3/2}\log N + \log(2\delta N^2) << R \qquad R \to \infty. \tag{7.50}$$

Como hemos visto en el Lema 7.5, esto implica que uno de los números

$$F_N(k_1+k_2b), \qquad (k_1,k_2) \in \mathbb{Z}^2 \qquad k_i \ge 1$$

es no nulo (usamos la hipótesis $b \notin \mathbb{Q}$). En consecuencia existe un entero $M \ge N$ tal que

$$F_N(i+jb) = 0, \qquad 1 \le i \le M\,; 1 \le j \le M \tag{7.51}$$

y existe

$$(i_0,j_0) \in \mathbb{Z}^2, \qquad 1 \le i_0 \le M+1; 1 \le j_0 \le M+1$$

con

$$\gamma_N = F_N(i_0+j_0b) \neq 0 \tag{7.52}$$

Para concluir la demostración se encontrarán cotas de γ_N superior e inferior contradictorias entre sí.

Veamos la cota superior

$$\log|\gamma_N| \leq -1/5M^2 \log M. \tag{7.53}$$

Notemos para ello que la función

$$F_N(z)\prod_{i=1}^{M}\prod_{j=1}^{M}(z-i-jb)^{-1} \tag{7.54}$$

es entera a causa de (7.51). Aplicaremos módulo máximo para funciones holomorfas (Lema 6.4) sobre el disco $|z| \leq R = (1+|b|)M^{5/4}$. Obtenemos

$$|\gamma_N| = |F_N(i_0+j_0b)| \leq |F_N|_R \sup_{|z|=R}\prod_{i=1}^{M}\prod_{j=1}^{M}\left|\frac{(i_0-i)+(j_0-j)b}{(z-i-jb)}\right|. \tag{7.55}$$

Mayoramos en $|z| = R$

$$\left|\frac{(i_0-i)+(j_0-j)b}{(z-i-jb)}\right| \leq \frac{(M+2)(1+|b|)}{R-M(1+|b|)} \leq \frac{M+2}{M^{5/4}-M} \leq 2M^{-1/4} \tag{7.56}$$

para N (y por ende M) suficientemente grande.

Por otro lado, debido a (7.50) tenemos

$$\log|F_N|_R \leq (2\delta|l|+1)RN^{1/2} \leq (2\delta|l|+1)(1+|b|)M^{7/4} \leq M^2 \tag{7.57}$$

donde N es suficientemente grande.

Obtenemos entonces

$$\log|\gamma_N| \leq 2M^2 - \frac{1}{4}M^2\log M \leq -1/5M^2\log M \tag{7.58}$$

lo que prueba (7.53)

Veamos la cota inferior. Para acotar inferiormente γ_N alcanza con acotar superiormente su tamaño $s(\gamma_N)$ y luego usar la relación del Lema 11.19

$$-2\delta s(\gamma_N) \leq log|\gamma_N| \tag{7.59}$$

ya que $\gamma_N \in K$ con $[K:\mathbb{Q}] = \delta$, y $\gamma_N \neq 0$ debido a (7.52).

Mostraremos que tenemos que

$$s(\gamma_N) \leq 4M^{3/2}logM \tag{7.60}$$

así entonces

$$\log|\gamma_N| \geq -8\delta M^{3/2}\log M \tag{7.61}$$

lo que contradecirá (7.53).

Para calcular el tamaño, usando el Lema 11.20 : obtenemos que

$$d^{N^{3/2}+4\delta N^{1/2}(M+1)} \tag{7.62}$$

es un denominador de γ_N, y que

$$\overline{|\gamma_N|} \leq N^{2N^{3/2}} M^{2N^{3/2}} \leq M^{4M^{3/2}} \tag{7.63}$$

lo que demuestra (7.60) y concluye la prueba de (7.8) □

Vista la prueba, podemos explicar las razones de la elección de las dos funciones $R_1(N) = N^{3/2}$ y $R_2(N) = N^{1/2}$ expresantes del grado del polinomio P_N respecto a X_1 y X_2.

Para aplicar el Lema de Siegel (7.2), hemos utilizado la desigualdad

$$R_1(N)R_2(N) \geq 2\delta N^2.$$

La acotación superior de γ_N involucra únicamente la cantidad

$$R_1(N)\log M + R_2(N)M.$$

Si las dos funciones R_1 y R_2 son monótonas crecientes, tendremos que

$$s(\gamma_N) << \max\{R_1(M)\log M, R_2(M)M\}.$$

Es entonces natural elegir R_1, R_2 de forma tal que las dos cantidades

$$R_1(N)\log N \text{ y } R_2(N)N$$

tengan el mismo orden. La elección óptima (dada la desigualdad $R_1(N)R_2(N) \geq 2\delta N^2$) será

$$R_1(N) = \lfloor (2\delta)^{1/2} N^{3/2} (\log N)^{-1/2} \rfloor + 1 \tag{7.64}$$

y

$$R_2(N) = \lfloor (2\delta)^{1/2} N^{1/2} (\log N)^{1/2} \rfloor + 1 \tag{7.65}$$

La elección que hemos hecho no es esencialmente distinta y hace que las funciones sean más sencillas.

Una vez elegidas R_1 y R_2, queda darle un valor al parámetro R, radio del disco sobre el cual utilizamos el Lema de módulo máximo de holomorfas 6.4 para acotar superiormente γ_N. Elegiremos R mucho más grande que M y acotamos superiormente

$$\sup_{|z|=R} \prod_{i=1}^{M}\prod_{j=1}^{M} \left| \frac{(i_0-i)+(j_0-j)b}{(z-i-jb)} \right| \leq -M^2 \log \frac{R}{M} + M^2 \log 2(1+|b|).$$

Si satisficiéramos la desigualdad

$$M^2 \log 2(1+|b|) + \log |F_N|_R \leq 1/5 M^2 \log \frac{R}{M} \tag{7.66}$$

obtendríamos

$$\log|\gamma_N| \leq -4/5M^2 \log \frac{R}{M} \tag{7.67}$$

(el 4/5 no tiene ninguna importancia). En la acotación superior (7.50) de $\log|F_N|_R$ el término de mayor grado es

$$2\delta N^{1/2}|l|R. \tag{7.68}$$

Para obtener el resultado, es suficiente que elijamos $R \leq M^{3/2}$, una elección posible es la que hemos hecho

$$R = (1 + |b|)M^{5/4} \tag{7.69}$$

8 Ampliación de Baker

Esta sección trata sobre el siguiente resultado demostrado por Baker en 1966.

Teorema 8.1 (1966 Baker)**.** *Sean* $\alpha_1, ..., \alpha_n \in \mathbb{A}\setminus\{0\}$*. Si* $\log(\alpha_1), ..., \log(\alpha_n)$ *son linealmente independientes sobre* $\mathbb{Q}$ *entonces* $1, \log(\alpha_1), ..., \log(\alpha_n)$ *son linealmente independientes sobre* $\mathbb{A}$*.*

Este resultado amplia el de Gelfond-Schneider en su formulación 3.3. La prueba se basa en la construcción de una función auxiliar de varias variables complejas que amplía la función de una sola variable usada por Gelfond. La mayor dificultad es en relación a las técnicas básicas de interpolación. El trabajo en conexión a esto hasta el momento de la prueba de Baker siempre involucró una extensión en el orden de las derivadas mientras se dejaban fijos los puntos de interpolación. Pero la prueba de 8.1 involucra un procedimiento de extrapolación propio de su contexto, en el cual el rango de interpolación es extendido mientras el orden de las derivadas se reduce.

8.1 Algunas consecuencias

Corolario 8.2. *Cualquier combinación lineal no nula de logaritmos de números algebraicos con coeficientes algebraicos es trascendente.*

En otras palabras, para cualquier conjunto de números algebraicos no nulos $\alpha_1, ..., \alpha_n$ y cualquier conjunto de números algebraicos $\beta_0, ..., \beta_n$ con $\beta_0 \neq 0$ se tiene que

$$\beta_0 + \beta_1 \log(\alpha_1) + ... + \beta_n \log(\alpha_n) \neq 0. \tag{8.1}$$

Prueba. Esto es válido para $n = 0$. Supongamos que es cierto para $n < m$ y probémoslo para $n = m$. Si $\log(\alpha_1), ..., \log(\alpha_m)$ son linealmente independientes sobre los racionales, el resultado se obtiene de 8.1. Entonces podemos suponer que existen racionales $\rho_1, ..., \rho_m$ con digamos $\rho_r \neq 0$ tales que

$$\rho_1 \log(\alpha_1) + ... + \rho_m \log(\alpha_m) = 0. \tag{8.2}$$

Tenemos entonces

$$\rho_r(\beta_0 + \beta_1 \log(\alpha_1) + ... + \beta_m \log(\alpha_m)) = \beta'_0 + \beta'_1 \log(\alpha_1) + ... + \beta'_m \log(\alpha_m), \tag{8.3}$$

donde

$$\beta'_0 = \rho_r\beta_0, \qquad \beta'_j = \rho_r\beta_j - \rho_j\beta_r \qquad 1 \leq j \leq m, \tag{8.4}$$

y también $\beta'_0 \neq 0, \beta'_r = 0$; así el resultado buscado se obtiene por inducción. □

Corolario 8.3. $e^{\beta_0}\alpha_1^{\beta_1}...\alpha_n^{\beta_n}$ *es trascendente para cualquier conjunto de números algebraicos no nulos* $\alpha_1, ..., \alpha_n, \beta_0, ..., \beta_n$*.*

Prueba. Si $\alpha_{n+1} = e^{\beta_0}\alpha_1^{\beta_1}...\alpha_n^{\beta_n}$ fuese algebraico, entonces

$$\beta_1 \log(\alpha_1) + ... + \beta_n \log(\alpha_n) - \log(\alpha_{n+1}) \qquad (= -\beta_0) \tag{8.5}$$

sería algebraico y no nulo, contradiciendo (8.2). □

Se tiene un análogo al Corolario 8.3 en el caso $\beta_0 = 0$:

Corolario 8.4. *$\alpha_1^{\beta_1}, ..., \alpha_n^{\beta_n}$ son trascendentes para cualquier conjunto de números algebraicos $\alpha_1, ..., \alpha_n$ distintos de 0 y 1, y cualesquiera números algebraicos $\beta_1, ..., \beta_n$, con $1, \beta_1, ..., \beta_n$ linealmente independientes sobre los racionales.*

Prueba. Alcanza con mostrar que para cualquier conjunto de números algebraicos $\alpha_1, ..., \alpha_n$ distintos de 0 y 1, y cualquier conjunto de números algebraicos $\beta_1, ..., \beta_n$ linealmente independientes sobre los racionales, tenemos

$$\beta_1 \log(\alpha_1) + ... + \beta_n \log(\alpha_n) \neq 0; \tag{8.6}$$

de hecho el resultado se sigue aplicando esto con n reemplazado por $n+1$ y $\beta_{n+1} = -1$. El resultado es cierto para $n = 1$; supongamos que es cierto para $n < m$ donde m es natural y probemos que es cierto para $n = m$. El resultado es una consecuencia inmediata de 8.1 si $\log \alpha_1, ..., \log \alpha_n$ son linealmente independientes sobre los racionales; así que podemos suponer que existen racionales $\rho_1, ..., \rho_m$ y números β'_j como en la prueba del Corolario 8.2 pero aquí con $\beta_0 = \beta'_0 = 0$. Si los $\beta_1, ..., \beta_m$ son linealmente independientes sobre los racionales, entonces también lo son los β'_j con j distinta de 0 y r, y el resultado buscado se sigue por inducción. □

Finalmente de casos particulares de los resultados anteriores, se tiene que $\pi + \log \alpha$ es trascendente para cualquier número algebraico $\alpha \neq 0$ (lo cual implica la trascendencia de π) y que $e^{\alpha\pi+\beta}$ es trascendente para cualquier par de números algebraicos α, β con $\beta \neq 0$ (lo cual implica la trascendencia de e).

Además de las consecuencias sobre trascendencia, Baker usó su resultado para derivar cotas efectivas para las soluciones de ciertas ecuaciones, y para solucionar el problema de número de clase de encontrar todos los cuerpos cuadráticos imaginarios con número de clase 1. Estos desarrollos pueden encontrarse en el libro de Baker indicado en la sección de referencias.

8.2 Preliminares

Supondremos que 8.1 ese falso, es decir que existen números algebraicos $\beta_0, ..., \beta_n$ no todo nulos tales que

$$\beta_0 + \beta_1 \log(\alpha_1) + ... + \beta_n \log(\alpha_n) = 0, \tag{8.7}$$

y derivaremos al final una contradicción. Por lo menos uno de los $\beta_1, ..., \beta_n$ es no nulo, y podemos suponer que $\beta_n \neq 0$. Entonces la ecuación anterior es cierta con $\beta'_j = -\beta_j/\beta_n$ en lugar de β_j, podemos suponer aún más, que $\beta_n = -1$, tenemos entonces

$$e^{\beta_0}\alpha_1^{\beta_1}...\alpha_{n-1}^{\beta_{n-1}} = \alpha_n. \tag{8.8}$$

Denotamos con $c, c_1, c_2, ...$ números positivos que dependen solamente de los α_j, β_j y las elecciones iniciales de los logaritmos. Por h denotamos un entero positivo que es mayor que un número suficientemente grande c como los anteriores.

Observamos para referencia posterior, que si α es un número algebraico cualquiera que satisface

$$A_0\alpha^d + A_1\alpha^{d-1} + ... + A_d = 0, \tag{8.9}$$

donde $A_0, ..., A_d$ son enteros racionales con valor absoluto menor o igual a A, entonces para cada entenero no negativo j, tenemos

$$(A_0\alpha)^j = A_0^{(j)} + A_1^{(j)}\alpha + ... + A_{d-1}^{(j)}\alpha^{d-1} \tag{8.10}$$

para algunos enteros racionales $A_m^{(j)}$ con valor absoluto menor o igual a $(2A)^j$; esto es una consecuencia de las relaciones recursivas

$$A_m^{(j)} = A_0 A_{m-1}^{(j-1)} - A_{d-m}A_{d-1}^{j-1} \qquad 0 \le m < d, j \ge d, \tag{8.11}$$

donde $A_{-1}^{(j-1)} = 0$. Se sigue que si d es el máximo de los grados de $\alpha_1, ..., \alpha_n, \beta_0, ..., \beta_n$ y que si $a_1, ..., a_n, b_0, ..., b_{n-1}$ son los coeficientes de mayor grado de sus respectivos polinomios minimales, entonces

$$(a_r\alpha_r)^j = \sum_{s=0}^{d-1} a_{rs}^{(j)}\alpha_r^s, \qquad (b_r\beta_r)^j = \sum_{t=0}^{d-1} b_{ts}^{(j)}\beta_r^t, \tag{8.12}$$

donde $a_{rs}^{(j)}, b_{ts}^{(j)}$ son enteros racionales con valor absoluto como mucho c_1^j.

Por brevedad notaremos

$$f_{m_0,...,m_{n-1}}(z_0, ..., z_{n-1}) = (d/dz_0)^{m_0}...(d/dz_{n-1})^{m_{n-1}} f(z_0, ..., z_{n-1}), \tag{8.13}$$

donde f denota una función entera y $m_0, ..., m_{n-1}$ son enteros no negativos.

Usaremos también el siguiente resultado

Teorema 8.5 (Fórmula de Leibnitz)**.** *Sean $u(z)$ y $v(z)$ funciones de clase C^n. Entonces su producto también es de clase C^n y*

$$\frac{d^n}{dz^n}[u(z)v(z)] = \sum_{r=0}^{n} \binom{n}{r} \frac{d^r}{dz^r}[u(z)] \frac{d^{n-r}}{dz^{n-r}}[v(z)], \tag{8.14}$$

donde

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{8.15}$$

es el coeficiente binomial usual.

Prueba. Por inducción sobre n. El resultado es trivial para $n = 0$, y para $n = 1$ es la regla conocida para la derivada del producto. Supongamos entonces que

$$\frac{d^{n-1}}{dz^{n-1}}[uv] = \sum_{r=0}^{n-1} \binom{n-1}{r} u^{(r)} v^{(n-1-r)} \tag{8.16}$$

para algún $n \geq 1$, donde $f^{(y)} = \frac{d^y}{dz^y} f$. Entonces

$$\begin{aligned}
\frac{d^n}{dz^n}(uv) &= \sum_{r=0}^{n-1} \binom{n-1}{r} [u^{(r)} v^{(n-r)} + u^{(r+1)} v^{(n-1-r)}] \\
&= uv^{(n)} + \sum_{r=1}^{n-1} \binom{n-1}{r} u^{(r)} v^{(n-r)} + \sum_{r=0}^{n-2} \binom{n-1}{r} u^{(r+1)} v^{(n-1-r)} + u^{(n)} v \\
&= uv^{(n)} + \sum_{r=1}^{n-1} \binom{n-1}{r} u^{(r)} v^{(n-r)} + \sum_{r=1}^{n-1} \binom{n-1}{r-1} u^{(r)} v^{(n-r)} + u^{(n)} v \\
&= uv^{(n)} + \sum_{r=1}^{n-1} [\binom{n-1}{r} + \binom{n-1}{r-1}] u^{(r)} v^{(n-r)} + u^{(n)} v,
\end{aligned}$$

donde

$$\left[\binom{n-1}{r} + \binom{n-1}{r-1}\right] = \frac{(n-1)!}{r!(n-1-r)!} + \frac{(n-1)!}{(r-1)!(n-r)!} = \frac{(n-1)!}{r!(n-r)!}[(n-r)+r] = \binom{n}{r}, \tag{8.17}$$

entonces

$$\frac{d^n}{dz^n}(uv) = uv^{(n)} + \sum_{r=1}^{n-1} \binom{n}{r} u^{(r)} v^{(n-r)} + u^{(n)} v = \sum_{r=0}^{n} \binom{n}{r} u^{(r)} v^{(n-r)}. \tag{8.18}$$

Lo que completa la prueba. □

8.3 La función auxiliar

Nuestro propósito en esta sección es describir la función auxiliar Φ esencial para la prueba de 8.8; es construida en el Lema 8.7, luego de un resultado preliminar sobre ecuaciones lineales. Estimaciones básicas en relación a Φ son establecidas en el Lema 8.8, y éstas son usadas luego para el algoritmo de extrapolación. Adicionalmente se incluyen los lemas 8.11 y 8.12, el primero da una cota inferior para formas lineales en logaritmos, y el segundo establece un polinomio aumentativo especial. Se verá que la inclusión del 1 en el enunciado de 8.8 introduce complejidad significativa en la prueba; en particular el lema final se necesita esencialmente para tratarla.

Lema 8.6. *Sean M, N enteros con $N > M > 0$ y sean*

$$u_{ij}, \qquad 1 \leq i \leq M, \quad 1 \leq j \leq N$$

enteros con valor absoluto como mucho $U \geq 1$. *Entonces existen enteros* $x_1, ..., x_N$ *no todos nulos, con valor absoluto como mucho* $(NU)^{\frac{M}{N-M}}$ *tales que*

$$\sum_{j=1}^{N} u_{ij}x_j = 0 \qquad 1 \leq i \leq M. \tag{8.19}$$

Prueba. Ponemos $B = \lfloor (NU)^{\frac{M}{N-M}} \rfloor$. Hay $(B+1)^N$ diferentes conjuntos de enteros $x_1, ..., x_N$ con $0 \leq x_j \leq B$, $1 \leq j \leq N$, y para cada uno de tales conjuntos tenemos

$$-V_iB \leq y_i \leq W_iB, \qquad 1 \leq i \leq M,$$

donde y_i denota el lado izquierdo de (8.19) y $-V_i, W_i$ denotan la suma de los negativos y positivos u_{ij}, $1 \leq j \leq N$ respectivamente. Como $V_i + W_i \leq NU$, entonces hay como mucho $(NUB+1)^M$ diferentes conjuntos $y_1, ..., y_M$. Como $(B+1)^{N-M} > (NU)^M$ entonces $(B+1)^N > (NUB+1)^M$. Entonces hay 2 conjuntos distintos $x_1, ..., x_N$ que corresponden al mismo conjunto $y_1, ..., y_M$, y su diferencia da la solución buscada de (8.19). □

Lema 8.7. *Existen enteros* $p(\lambda_0, ..., \lambda_n)$ *no todos nulos, con valor absoluto como mucho* e^{h^3}, *tales que la función*

$$\Phi(z_0, ..., z_{n-1}) = \sum_{\lambda_0=0}^{L} ... \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n) z_0^{\lambda_0} e^{\lambda_n\beta_0 z_0} \alpha_1^{\gamma_1 z_1} ... \alpha_{n-1}^{\gamma_{n-1} z_{n-1}}, \tag{8.20}$$

donde $\gamma_r = \lambda_r + \lambda_n\beta_r$ $(1 \leq r < n)$ *y* $L = \lfloor h^{2-\frac{1}{4n}} \rfloor$, *satisface*

$$\Phi_{m_0,...,m_{n-1}}(l, ..., l) = 0 \tag{8.21}$$

para todos los enteros l *con* $1 \leq l \leq h$ *y todos los enteros no negativos* $m_0, ..., m_{n-1}$ *con* $m_0 + ... + m_{n-1} \leq h^2$.

Prueba. Alcanza, en vista de (8.8), determinar los $p(\lambda_0, ..., \lambda_n)$ tales que

$$\sum_{\lambda_0=0}^{L} ... \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} ... \alpha_n^{\lambda_n l} \gamma_1^{m_1} ... \gamma_{n-1}^{m_{n-1}} = 0 \tag{8.22}$$

para los rangos anteriores de $l, m_0, ..., m_{n-1}$, donde

$$q(\lambda_0, \lambda_n, z) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1)...(\lambda_0 - \mu_0 + 1)(\lambda_n\beta_0)^{m_0-\mu_0} z^{\lambda_0-\mu_0}. \tag{8.23}$$

Multiplicando (8.22) por

$$P' = (a_1...a_n)^{Ll} b_0^{m_0} ... b_{n-1}^{m_{n-1}}, \tag{8.24}$$

escribiendo

$$\gamma_r^{m_r} = \sum_{\mu_r=0}^{m_r} \binom{m_r}{\mu_r} \lambda_r^{m_r-\mu_r} (\lambda_n\beta_r)^{\mu_r}, \tag{8.25}$$

y substituyendo con (8.12) para las potencias de $a_r\alpha_r$ y $b_r\beta_r$ que resultan, obtenemos

$$\sum_{s_1=0}^{d-1}...\sum_{s_n=0}^{d-1}\sum_{t_0=0}^{d-1}...\sum_{t_{n-1}=0}^{d-1}A(s,t)\alpha_1^{s_1}...\alpha_n^{s_n}\beta_0^{t_0}...\beta_{n-1}^{t_{n-1}}=0, \tag{8.26}$$

donde

$$A(s,t)=\sum_{\lambda_0=0}^{L}...\sum_{\lambda_n=0}^{L}\sum_{\mu_0=0}^{m_0}...\sum_{\mu_{n-1=0}}^{m_{n-1}}p(\lambda_0,...,\lambda_n)q'q''q''', \tag{8.27}$$

y q', q'', q''' están dados por

$$q'=\prod_{r=1}^{n}[\alpha_r^{(L-\lambda_r)l}\alpha_{r,s_r}^{(\lambda_r l)}], \tag{8.28}$$

$$q''=\prod_{r=1}^{n-1}\left[\binom{m_r}{\mu_r}(b_r\lambda_r)^{m_r-\mu_r}\lambda_n^{\mu_r}b_{r,t_r}^{(\mu r)}\right], \tag{8.29}$$

$$q'''=\binom{m_0}{\mu_0}\lambda_0(\lambda_0-1)...(\lambda_0-\mu_0+1)\lambda_n^{m_0-\mu_0}b_n^{\mu_0}l^{\lambda_0-\mu_0}b_{0,t_0}^{(m_0-\mu_0)}. \tag{8.30}$$

Entonces (8.21) será satisfecha si las d^{2n} ecuaciones $A(s,t)=0$ son ciertas. Éstas representan ecuaciones lineales en los $p(\lambda_0,...,\lambda_n)$ con coeficientes enteros.

Como $l\leq h$ y $\binom{m_r}{\mu_r}\leq 2^{m_r}$, tenemos que

$$|q'|\leq\prod_{r=1}^{n}[a_r^{(L-\lambda_r)l}c_1^{\lambda_r l}]\leq c_2^{Lh}, \tag{8.31}$$

$$|q''|\leq\prod_{r=1}^{n-1}(c_3L)^{m_r} \tag{8.32}$$

$$|q'''|\leq 2^{m_0}(\lambda_0 b_n)^{\mu_0}(c_1\lambda_n)^{m_0-\mu_0}l^{\lambda_0-\mu_0}\leq(c_3L)^{m_0}h^L, \tag{8.33}$$

y por las desigualdades

$$(m_0+1)...(m_{n-1}+1)\leq 2^{m_0+...+m_{n-1}}\leq 2^{h^2}, \tag{8.34}$$

se obtiene que el coeficiente de $p(\lambda_0,...,\lambda_n)$ en la forma lineal $A(s,t)$, es decir

$$\sum_{\mu_0=0}^{m_0}...\sum_{\mu_{n-1}=0}^{m_{n-1}}q'q''q''', \tag{8.35}$$

tiene valor absoluto como mucho $U = (2c_3L)^{h^2 c_4^{Lh}}$. Además, hay como mucho $h(h^2+1)^n$ distintos conjuntos de enteros $l, m_0, ..., m_{n-1}$, y entonces hay $M \leq d^{2n}h(h^2+1)^n$ ecuaciones $A(s,t) = 0$ correspondientes a ellos. Aún más, hay $N = (L+1)^{n+1}$ incógnitas $p(\lambda_0, ..., \lambda_n)$ y tenemos

$$N > h^{[2-\frac{1}{4n}](n+1)} \geq h^{2n+\frac{3}{2}} > 2d^{2n}h(h^2+1)^n \geq 2M. \tag{8.36}$$

Entonces por el Lema 8.6, las ecuaciones pueden solucionarse en forma no trivial y los enteros $p(\lambda_0, ..., \lambda_n)$ pueden elegirse con valor absoluto como mucho

$$NU \leq h^{2n+2}(2c_3L)^{h^2}c_4^{Lh} \leq e^{h^3} \tag{8.37}$$

si h es suficientemente grande, como se quería probar. □

Lema 8.8. *Sean $m_0, ..., m_{n-1}$ enteros no negativos cualesquiera con*

$$m_0 + ... + m_{n-1} \leq h^2, \tag{8.38}$$

y sea

$$f(z) = \Phi_{m_0,...,m_{n-1}}(z, ..., z) \tag{8.39}$$

donde Φ es la función definida en el Lema 8.7. Entonces para cualquier número z, tenemos $|f(z)| \leq c_5^{h^3+L|z|}$. Además para cualquier entero positivo l, o bien $f(l) = 0$ o bien $|f(l)| > c_6^{-h^3-Ll}$.

Prueba. La función $f(z)$ está dada por

$$P\sum_{\lambda_0=0}^{L} ... \sum_{\lambda_n=0}^{L} p(\lambda_0, ..., \lambda_n)q(\lambda_0, \lambda_n, z)\alpha_1^{\lambda_1 z}...\alpha_n^{\lambda_n z}\gamma_1^{m_1}...\gamma_{n-1}^{m_{n-1}}, \tag{8.40}$$

donde $q(\lambda_0, \lambda_n, z)$ está definida en el Lema 8.7 y

$$P = (\log \alpha_1)^{m_1}...(\log \alpha_{n-1})^{m_{n-1}}. \tag{8.41}$$

Tenemos

$$|q(\lambda_0, \lambda_n, z)| \leq (c_7L)^{m_0}|z|^L \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} = (2c_7L)^{m_0}|z|^L, \tag{8.42}$$

$$|\alpha_1^{\lambda_1 z}...\alpha_n^{\lambda_n z}| \leq c_8^{L|z|}, \tag{8.43}$$

$$|P\gamma_1^{m_1}...\gamma_{n-1}^{m_{n-1}}| \leq (c_9L)^{m_1+...+m_{n-1}}, \tag{8.44}$$

y el número de términos en la suma múltiple anterior es como mucho h^{2n+2}; la estimación buscada para $|f(z)|$ se obtiene entonces a partir de las desigualdades

$$L \leq h^2 \tag{8.45}$$

$$m_0 + ... + m_{n-1} \leq h^2 \tag{8.46}$$

$$|p(\lambda_0, ..., \lambda_n)| \leq e^{h^3}. \tag{8.47}$$

Para probar la segunda afirmación, comenzamos observando que el número $f' = (P'/P)f(l)$ donde P' está definido por (8.24), es un entero algebraico con grado como mucho d^{2n}. Además por estimaciones como las anteriores, vemos que cualquier conjugado de f', obtenido sustituyendo conjugados arbitrarios para los α_r, β_r, tiene valor absoluto como mucho $c_{10}^{h^3+Ll}$; y la misma cota se obtiene para P'/P. Pero si $f' \neq 0$, entonces la norma de f' (es decir, el producto de sus conjugados) tiene valor absoluto por lo menos 1, y entonces

$$|f'| \geq c_{10}^{-(h^3+Ll)d^{2n}}. \tag{8.48}$$

Esto da el resultado buscado. □

Lema 8.9. *Sea J cualquier enetero que satisface $0 \leq J \leq (8n)^2$. Entonces (8.21) es cierta para todo entero l con $1 \leq l \leq h^{1+\frac{J}{8n}}$ y todos enteros no.negativos $m_0, ..., m_{n-1}$ con $m_0 + ... + m_{n-1} \leq \frac{h^2}{2^J}$.*

Prueba. El resultado es cierto para $J = 0$ por el Lema 8.7. Sea K un entero con $0 \leq K < (8n)^2$ y supongamos que el resultado buscado es válido para

$$J = 0, 1, ..., K.$$

Probaremos el resultado para $J = K + 1$.

Alcanza con mostrar que para cualquier entero l con $R_K < l \leq R_{K+1}$ y cualquier conjunto de enteros no negativos $m_0, ..., m_{n-1}$ con

$$m_0 + ... + m_{n-1} \leq S_{K+1},$$

tenemos $f(l) = 0$ donde $f(z)$ está definida por (8.39) y

$$R_J = \lfloor h^{1+\frac{J}{8n}} \rfloor \qquad J = 0, 1, ...$$

$$S_J = \lfloor \frac{h^2}{2^J} \rfloor \qquad J = 0, 1, ...$$

Por la hipótesis inductiva vemos que $f_m(r) = 0$ para todos los enteros r, m con $1 \leq r \leq R_K$, $0 \leq m \leq S_{K+1}$; ya que $f_m(r)$ está dada por

$$(d/dz_0 + ... + d/dz_{n-1})^m \Phi_{m_0,...,m_{n-1}}(z_0, ..., z_{n-1}), \tag{8.49}$$

calculada en el punto $z_0 = ... = z_{n-1} = r$, eso es por

$$\sum m!(j_0!...j_{n-1}!)^{-1} \Phi_{m_0+j_0,...,m_{n-1}+j_{n-1}}(r, ..., r), \tag{8.50}$$

donde la suma es sobre todos los enteros no negativos $j_0, ..., j_{n-1}$ con $j_0 + ... + j_{n-1} = m$, y las derivadas aquí son 0, ya que

$$m_0 + ... + m_{n-1} + j_0 + ... + j_{n-1} \leq 2S_{K+1} \leq S_K. \tag{8.51}$$

Entonces $f(z)/F(z)$, donde

$$F(z) = [(z-1)...(z-R_K)]^{S_{K+1}}, \tag{8.52}$$

es holomorfa dentro y sobre la circunsferencia C con centro el punto cero y radio $R_{K+1}h^{\frac{1}{8n}}$, y por lo tanto por Módulo Máximo 6.4, se tiene

$$\theta|F(l)| \geq \Theta|f(l)|, \tag{8.53}$$

donde θ, Θ denotan respectivamente la cota superior de $|f(z)|$ y la cota inferior de $|F(z)|$ con z sobre C. Se tiene $\Theta \geq (\frac{1}{2}R)^{R_K S_{K+1}}$, y por el Lema 8.8, $\theta \leq c_5^{h^3+LR}$. Además tenemos $|F(l)| \leq R_{K+1}^{R_K S_{K+1}}$ y, nuevamente por el Lema 8.8, o bien $f(l) = 0$ o bien $|f(l)| > c_6^{-h^3-LR}$. Pero en vista de (8.53), esta última posibilidad da

$$(c_5 c_6)^{h^3+LR} \geq (\frac{1}{2}h^{\frac{1}{8n}})^{R_K S_{K+1}}, \tag{8.54}$$

y como $K < (8n)^2$ y

$$LR \leq h^{3+\frac{K}{8n}} \leq 2^{K+3} R_K S_{K+1}, \tag{8.55}$$

la desigualdad es imposible si h es suficientemente grande. Así que debe tenerse $f(l) = 0$, y el lema que se quería probar se obtiene por inducción. □

Lema 8.10. *Escribiendo $\phi(z) = \Phi(z, ..., z)$ (donde Φ es la función definida en el Lema 8.7) tenemos*

$$|\phi_j(0)| < e^{-h^{8n}} \qquad 0 \leq j \leq h^{8n}. \tag{8.56}$$

Prueba. Del Lema 8.9 vemos que (8.21) es cierta para todos los enteros l y enteros no negativos $m_0, ..., m_{n-1}$ que satisfacen $1 \leq l \leq X$ y

$$m_0 + ... + m_{n-1} \leq Y, \tag{8.57}$$

donde $X = h^{8n}$ y $Y = \lfloor \frac{h^2}{2^{(8n)^2}} \rfloor$. Entonces como en la prueba del Lema 8.9, obtenemos $\phi_m(r) = 0$, para todos los enteros r, m con $1 \leq r \leq X, 0 \leq m \leq Y$.

Definiendo $E(z) = [(z-1)...(z-X)]^Y$, se obtiene que $\phi(z)/E(z)$ es holomorfa dentro y sobre de la circunsferencia Γ con centro en el punto cero y radio $R = Xh^{\frac{1}{8n}}$, y por lo tanto por Módulo Máximo 6.4, se tiene, para cualquier w con $|w| < X$,

$$|\phi(w)| \leq \xi \Xi^{-1} |E(w)|, \tag{8.58}$$

donde ξ y Ξ denotan respectivamente la cota superior de $|\phi(z)|$ y la cota inferior de $|E(z)|$ con z sobre Γ. Tenemos que

$$|E(w)| \leq (2X)^{XY}, \qquad |\Xi| \geq (\frac{1}{2}R)^{XY}, \tag{8.59}$$

y por, el Lema 8.8, $\xi \leq c_5^{h^3+LR}$. Entonces obtenemos

$$|\phi(w)| \leq c_5^{h^3+LR}(\frac{1}{2}h^{\frac{1}{8n}})^{-XY}, \tag{8.60}$$

y como

$$LR \leq h^{8n+2} \leq 2^{(8n)^2+1}XY, \tag{8.61}$$

se obitene que $|\phi(w)| < e^{-XY}$. Además por la representación integral de holomorfa (6.2), tenemos

$$\phi_j(0) = \frac{j!}{2\pi i}\int_\Lambda \frac{\phi(w)}{w^{j+1}}\,dw, \tag{8.62}$$

donde Λ denota a la circunsferencia $|w| = 1$, y la expresión de la derecha tiene valor absoluto como mucho $j^j e^{-XY}$. La estimación buscada (8.56) se obtiene entonces de esto. □

Lema 8.11. *Para cualesquiera enteros $t_1, ..., t_n$ no todos nulos, y con valores absolutos como mucho T, tenemos que*

$$|t_1 \log \alpha_1 + ... + t_n \log \alpha_n| > c_{11}^{-T}. \tag{8.63}$$

Prueba. Sea a_j, $1 \leq j \leq n$ el coeficiente de mayor grado del polinomio minimal de α_j o α_j^{-1} según si $t_j \geq 0$ o $t_j < 0$. Entonces

$$\omega = a_1^{|t_1|}...a_n^{|t_n|}(\alpha_1^{t_1}...\alpha_n^{t_n} - 1) \tag{8.64}$$

es un entero algebraico de grado como mucho d^n, y cualquier conjugado de ω, obtenido sustituyendo conjugados arbitrarios por $\alpha_1, ..., \alpha_n$, tiene valor absoluto como mucho c_{12}^T. Si $\omega = 0$ entonces

$$\Omega = t_1 \log \alpha_1 + ... + t_n \log \alpha_n \tag{8.65}$$

es un múltiplo de $2\pi i$, y de hecho un múltiplo no nulo, ya que $\log \alpha_1, ..., \log \alpha_n$ son por hipótesis, linealmente independientes sobre los racionales; entonces en este caso, el lema es trivialmente válido. En otro caso la norma de ω tiene valor absoluto por lo menos 1 y entonces $|\omega| \geq c_{12}^{-Td^n}$. Pero como, para todo z,

$$|e^z - 1| \leq |z|e^{|z|},$$

obtenemos que

$$|\omega| \leq |\Omega|e^{|\Omega|}c_{13}^T$$

y entonces tomando $|\Omega| < 1$, se obtiene el lema buscado. □

Lema 8.12. *Sean R, S enteros positios y sean $\sigma_0, ..., \sigma_{R-1}$ distintos números complejos.*

Definamos $\sigma = \max_{0\leq i\leq R-1}\{1, |\sigma_i|\}$ y $\rho = \min_{0\leq i<j\leq R}\{1, |\sigma_i - \sigma_j|\}$.

Entonces, para cualesquiera enteros r, s con $0 \leq r < R$, $0 \leq s < S$, existen números complejos w_i, $0 \leq i < RS$ con valor absoluto como mucho $(\frac{8\sigma}{\rho})^{RS}$ tales que el polinomio

$$W(z) = \sum_{j=0}^{RS-1} w_j z^j \tag{8.66}$$

satisface $W_j(\sigma_i) = 0$ para todos i, j con $0 \leq i < R$, $0 \leq j < S$ distintos de $i = r$, $j = s$, y $W_s(\sigma_r) = 1$.

Prueba. El polinomio buscado está dado por

$$W(z) = \frac{-1}{s!}\frac{1}{2\pi i}\int_{C_r} \frac{(\zeta-\sigma_r)^S U(z)}{(\zeta - z)U(\zeta)}\, d\zeta, \tag{8.67}$$

donde

$$U(z) = [(z-\sigma_0)...(z-\sigma_{R-1})]^S \tag{8.68}$$

y C_r denota a la circunsferencia con centro σ_r y radio suficientemente chico, menor a, digamos, ρ y $|z-\sigma_r|$ para $z \neq \sigma_r$. La prueba depende de dos expresiones alternativas para $W(z)$.

Primero como el valor absoluto del integrando multiplicado por $|\zeta|$ tiende a 0 cuando $|\zeta| \to \infty$, tenemos por la fórmula de residuos (6.3),

$$W(z) = \frac{(z-\sigma_r)^S}{s!} + \frac{U(z)}{s!}\frac{1}{2\pi i}\sum_{j=0, j\neq r}^{R-1}\int_{C_j} \frac{(\zeta-\sigma_r)^S}{(\zeta-z)U(\zeta)} d\zeta, \tag{8.69}$$

donde tanto C_j, como el C_r anterior, es una circunsferencia de centro σ_j con radio suficientemente chico. La suma sobre j es una función racional de z, holmorfa en $z = \sigma_r$ y, como $U(z)$ tiene un cero de orden S en $z = \sigma_r$, se obtiene que $W_j(\sigma_r) = 1$ si $j = s$ y 0 en otro caso.

Por otro lado, de la fórmula integral para funciones holomorfas (6.2), obtenemos

$$W(z) = \frac{-1}{s!t!}\left[\frac{d^t}{d\zeta^t}\frac{(\zeta-\sigma_r)^S U(z)}{(\zeta-z)U(\zeta)}\right]_{\zeta=\sigma_r}, \tag{8.70}$$

donde $t = S - s - 1$, y entonces

$$W(z) = (-1)^{t-1}(s!)^{-1}U(z)\sum v(j_0,...,j_{R-1})(\sigma_r - z)^{-j_r-1}, \tag{8.71}$$

donde la suma es sobre todos los enteros no negativos $j_0,...,j_{R-1}$ con $j_0 + ... + j_{R-1} = t$, y

$$v(j_0,...,j_{R-1}) = \prod_{i=0, i\neq r}^{R-1}\binom{S+j_i-1}{j_i}(\sigma_r - \sigma_i)^{-S-j_i}. \tag{8.72}$$

Tenemos que $j_r + 1$ está entre 1 y S inclusive, y entonces $W(z)$ es un polinomio de grado como mucho $RS - 1$. Además, vemos que tanto $W(z)$, como $U(z)$ tienen un cero de orden S en $z = \sigma_i$ $(i \neq r)$, y entonces $W_j(\sigma_i) = 0$ para todo $j < S$. Aún más, los factores en el producto que definen v tienen valor absoluto como mucho $2^{S+j_i-1}\rho^{-S-j_i}$, y entonces

$$|v(j_0,...,j_{R-1})| \leq \left(\frac{2}{\rho}\right)^{(R-1)S+j_0+...+j_{R-1}} \leq \left(\frac{2}{\rho}\right)^{RS}. \tag{8.73}$$

Observando que los coeficientes de $(\sigma_r - z)^{-j_r-1}U(z)$ tienen valor absoluto como mucho $(\sigma+1)^{RS}$, y que además, el número de términos en la suma anterior no excede S^R, se obtiene que los coeficientes de $W(z)$ tienen valor absoluto como mucho

$$S^R(\sigma+1)^{RS}\left(\frac{2}{\rho}\right)^{RS} \leq \left(\frac{8\sigma}{\rho}\right)^{RS}, \tag{8.74}$$

y esto completa la prueba del lema. □

8.4 Prueba de Baker 8.1

Mostraremos a continuación que las desigualdades (8.56) obtenidas en el Lema 8.10 no pueden ser todas ciertas, y esta contradicción establecerá el resultado de Baker 8.1.

Prueba. Comenzamoes escribiendo $S = L + 1$, $R = S^n$, y observando que cualquier entero i con $0 \leq i < RS$ puede expresarse de modo único en la forma

$$i = \lambda_0 + \lambda_1 S + ... + \lambda_n S^n, \tag{8.75}$$

donde $\lambda_0, ..., \lambda_n$ denota enteros entre 0 y L inclusive. Para cada tal i definimos

$$v_i = \lambda_0, \qquad p_i = p(\lambda_0, ..., \lambda_n), \tag{8.76}$$

y ponemos

$$\psi_i = \lambda_1 \log \alpha_1 + ... + \lambda_n \log \alpha_n. \tag{8.77}$$

Entonces tenemos que

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{v_i} e^{\psi_i z}. \tag{8.78}$$

Además, por el Lema 8.11, dos cualesquieras ψ_i correspondientes a distintos conjuntos $\lambda_1, ..., \lambda_n$ difieren por lo menos en c_{11}^{-L} ; en particular, exactamente R de los ψ_i son distintos, y denotamos los diferentes valores, en algún orden, por $\sigma_0, ..., \sigma_{R-1}$. Si σ, ρ son definidos como en el Lema 8.12, tenemos entonces

$$\sigma \leq c_{14} L \qquad \text{y} \qquad \rho \geq c_{15}^{-L}. \tag{8.79}$$

Sea t cualquier subíndice tal que $p_t \neq 0$, sea $s = v_t$, sea r el subíndice para el cual $\psi_t = \sigma_r$, y sea $W(z)$ el polinomio definido en el Lema 8.12. Entonces por las propiedades de $W(z)$ especificadas en el Lema, se obtiene que

$$p_t = \sum_{i=0}^{RS-1} p_i W_{v_i}(\psi_i). \tag{8.80}$$

Además, por la fórmula de Leibnitz dada en el Lema 8.5, tenemos que

$$W_{v_i}(\psi_i) = \sum_{j=0}^{RS-1} j(j-1)...(j-v_i+1) w_j \psi_i^{j-v_i} = \sum_{j=0}^{RS-1} w_j \left[\frac{d^j}{dz^j} (z^{v_i} e^{\psi_i z}) \right]_{z=0}, \tag{8.81}$$

y entonces de (8.78) obtenemos

$$p_t = \sum_{j=0}^{RS-1} w_j \phi_j(0). \tag{8.82}$$

Tenemos que $RS \leq h^{2n+2}$, y entonces, del Lema 8.10, se deduce que (8.56) es cierta para todo j con $0 \leq j \leq RS$. Además por el Lema 8.12, tenemos que

$$|w_j| \leq \left(\frac{8\sigma}{\rho}\right)^{RS} \leq (8c_{14}Lc_{15}^{L})^{RS} \leq c_{16}^{h^{2n+4}}. \tag{8.83}$$

Entonces, como $|p_t| \geq 1$, concluimos que

$$0 \leq \log RS + c_{17}h^{2n+4} - h^{8n}. \tag{8.84}$$

Esta desigualdad es imposible si h es suficientemente grande, y la contradicción prueba finalmente el resultado de Baker 8.1. □

9 Ampliación de Lang

Teorema 9.1 (1966 Lang). *Sea K un cuerpo numérico. Sean $f_1, ..., f_N$ funciones meromorfas de orden menor o igual a ρ. Supongamos que el cuerpo $K(f_1, ..., f_N)$ tiene grado de trascendencia mayor o igual a 2 sobre K, y que la derivada $D = \frac{d}{dt}$ mapea el anillo $K[f_1, ..., f_N]$ en sí mismo. Sean $w_1, ..., w_m$ números complejos distintos que no están entre los polos de las f_i, tales que*

$$f_i(w_v) \in K$$

para todo $i = 1, ..., N$ y $v = 1, ..., m$. Entonces $m \leq 20\rho[K : Q]$.

Veamos que este resultado amplia otros ya conocidos.

Corolario 9.2 (Hermite-Lindemann). *Sea α un número algebraico no nulo. Entonces e^α es trascendente.*

Prueba. Supongamos que e^α es algebraico. Sea K el cuerpo construido por α y e^α sobre $\mathbb{Q}$. Seanf, g las funciones$f(t) = t$ y $g(t) = e^{\alpha t}$. Entonces el anillo $K[f, g]$ es mapeado en sí mismo por $D = \frac{d}{dt}$, y f y g son algebraicamente independientes. Observemos que f y g toman valores en K en todos los números $\alpha, 2\alpha, ..., m\alpha$ para m arbitrariamente grande. Esto contradice 9.1. □

Corolario 9.3 (Gelfond-Schneider). *Sean α, β algebraicos, con $\alpha \neq 0, 1$ y β irracional. Entonces α^β es trascendente.*

Prueba. Supongamos que α^β es algebraico, sea entonces K el cuerpo construido sobre $\mathbb{Q}$ por $\alpha, \beta, \alpha^\beta$. Si aplicamos 9.1 a las funciones e^t y $e^{\beta t}$, y al conjunto de múltiplos enteros de $\log \alpha$ obtenemos una contradicción como en el caso anterior. □

9.1 Preliminares

Sea

$$P(T_1, ..., T_N) = \sum \alpha_{(v)} M_{(v)}(T) \tag{9.1}$$

un polinomio con coeficientes complejos $\alpha_{(v)}$, en N variables $T_1, ..., T_N$, y sea

$$Q(T_1, ..., T_N) = \sum \beta_{(v)} M_{(v)}(T) \tag{9.2}$$

un polinomio con coeficientes reales no negativos. Decimos que Q "domina" P, si $|\alpha_{(v)}| \leq \beta_{(v)}$ para todo (v), y lo notamos $P \prec Q$. Se obtienen de forma inmediata las siguientes propiedades.

Si Q_1 domina P_1 y Q_2 domina P_2 entonces

$$P_1 + P_2 \prec Q_1 + Q_2 \qquad \text{y} \qquad P_1 P_2 \prec Q_1 Q_2. \tag{9.3}$$

Además si D_i, $i = 1, ..., N$ es la derivada $i-$ésima respecto T_i, y si Q domina P, entonces

$$D_i P \prec D_i Q. \tag{9.4}$$

Sea P un polinomio con coeficientes complejos como el anterior. Ponemos

$$|P| = \max |\alpha_{(v)}| \tag{9.5}$$

el mayor de los valores absolutos de los coeficientes. Si el grado total de P es menor o igual a r, entonces

$$P \prec |P|(1 + T_1 + ... + T_N)^r. \tag{9.6}$$

Como es sencillo tomar la derivada del lado derecho respecto a cualquier variable T_i, tenemos una forma sencilla de estimar las sucesivas derivadas formales de polinomios.

Sea K un cuerpo numérico.

Sea S un conjunto de elementos de K. Denotamos por $||S||$ al mayor de los valores absolutos de todos los conjugados de elementos de S. Por "denominador" de S nos referiremos a un entero positivo que es un denominador común de todos los elementos de S. Si P es un polinomio con coeficientes en K, definimos $||P||$ y un denominador para P en términos del conjunto de coeficientes del polinomio. Abreviamos "denominador" con "den".

Lema 9.4. *Sea K un cuerpo numérico. Sean $f_1, ..., f_N$ funciones holomorfas en un entorno de un punto $w \in \mathbb{C}$, y supongamos que $D = \frac{d}{dt}$ mapea el anillo $K[f_1, ..., f_N]$ en sí mismo. Supongamos que $f_i(w) \in K$ para todo i. Entonces existe un número C_1 con la siguiente propiedad. Sea $P(T_1, ..., T_N)$ un polinomio con coeficientes en K, de grado $\leq r$. Si ponemos $f = P(f_1, ..., f_N)$, entonces tenemos, para todos los k enteros positivos,*

$$||D^k f(w)|| \leq ||P|| r^k k! C_1^{k+r}. \tag{9.7}$$

Además, existe un denominador para $D^k f(w)$ acotado por $den(P)C_1^{k+r}$.

Prueba. Existen polinomios $P_i(T_1, ..., T_N)$ con coeficientes en K tales que

$$Df_i = P_i(f_1, ..., f_N). \tag{9.8}$$

Sea δ el mayor de sus grados. Existe una única derivación $\overline{D}$ en $K[T_1, ..., T_N]$ tal que $\overline{D}T_i = P_i(T_1, ..., T_N)$. Para cualquier polinomio P en $K[T]$, tenemos

$$\overline{D}(P(T_1, ..., T_N)) = \sum_{i=1}^{N} (D_i P)(T_1, ..., T_N).P_i(T_1, ..., T_N) \tag{9.9}$$

donde $D_1, ..., D_N$ son las derivadas parciales formales. El polinomio P está dominado por

$$||P||(1 + T_1 + ... + T_N)^r, \tag{9.10}$$

y cada P_i está dominado por

$$||P_i||(1 + T_1 + ... + T_N)^\delta. \tag{9.11}$$

Entonces $\overline{D}P$ está dominado por

$$||P||C_2 r(1 + T_1 + ... + T_N)^{r+\delta}. \tag{9.12}$$

Procediento inductivamente, se obtiene que

$$\overline{D}^k P \prec ||P|| C_3^k r^k k! (1 + T_1 + ... + T_N)^{r+k\delta}. \tag{9.13}$$

Substituyendo los valores $f_i(w)$ por T_i, tenemos

$$D^k f(w) = \overline{D}^k P(f_1(w), ..., f_N(w)), \tag{9.14}$$

y obtenemos así la cota buscada para $D^k f(w)$. La segunda afirmación sobre denominadores se prueba también por inducción inmediata. □

Tenemos también las siguientes formas de Lemas de Siegel (cuyas demostraciones son similares a las ya vistas en secciones anteriores).

Lema 9.5. *Sea*

$$a_{11}x_1 + ... + a_{1n}x_n = 0$$

$$...$$

$$a_{r1}x_1 + ... + a_{rn}x_n = 0$$

un sistema de ecuaciones lineales con coeficientes enteros a_{ij}, y $n > r$. Sea A un número ≥ 1 tal que $|a_{ij}| \leq A$ para todos i, j. Entonces existe un solución entera no trivial con

$$|x_j| \leq 2(2nA)^{\frac{r}{(n-r)}}.$$

Lema 9.6. *Sea K una extensión finita de $\mathbb{Q}$. Sea*

$$a_{11}x_1 + ... + a_{1n}x_n = 0$$

$$...$$

$$a_{r1}x_1 + ... + a_{rn}x_n = 0$$

un sistema de ecuaciones lineales con coeficientes en $\mathbb{Z}_K$, y $n > r$. Sea A un número tal que $||a_{ij}|| \leq A$ para todos i, j. Entonces existe un solución no trivial X en $\mathbb{Z}_K$ tal que

$$||X|| \leq C_1(C_2 nA)^{\frac{r}{(n-r)}} + C_1$$

donde $||X||$ es el mayor de los valores absolutos de todos los conjugados de los x_i, y C_1, C_2 son constantes que dependen solo de K.

9.2 Idea de la prueba

Lang toma dos funciones algebraicamente independientes de entre $f_1, ..., f_n$, digamos f y g, y construye una función auxiliar que es un polinomio en ellas. No expresa esta función explícitamente, sino que usando el Lema de Siegel prueba que existe y es tal que se anula hasta un orden alto en los m números complejos $w_1, ..., w_m$. Debido a su alto orden de anulación puede mostrarse que una derivada de alto orden de F toma un valor de tamaño (size algebraico) muy chico en uno de los w_i. Usando el criterio de Módulo Máximo 6.4, encuentra otra forma de acotar las derivadas de F, y usando resultados que comparan el size de un número y su valor absoluto muestra que estas estimaciones se contradicen a menos que la cota del enunciado sobre m sea cierta.

9.3 Prueba de 9.1

Probamos a continuación 9.1.

Prueba. Sean f, g dos funciones entre $f_1, ..., f_N$, que son algebraicamente independientes sobre K. Sea r un entero positivo divisible por $2m$. Al final de la prueba haremos tender r a infinito.

Sea

$$F = \sum_{i,j=1}^{r} a_{ij} f^i g^j \tag{9.15}$$

con coeficientes a_{ij} en Z_K (el anillo de enteros algebraicos en K). Sea $n = \frac{r^2}{2m}$. Podemos elegir los a_{ij} no todos igual a 0, y tales que

$$D^k F(w_v) = 0 \tag{9.16}$$

para $0 \leq k < n$ y $v = 1, ..., m$. Tenemos que solucionar un sistema de mn ecuaciones lineales en $r^2 = 2mn$ incógnitas. Observemos que

$$\frac{mn}{(2mn - mn)} = 1. \tag{9.17}$$

El tamaño de los coeficientes, y un denominador para ellos, son obtenidos como consecuencia del Lema 9.4. Por el Lema de Siegel 9.6, vemos que podemos tomar los a_{ij} tales que

$$size(a_{ij}) \leq n \log r + n \log n + (n + r)C_1 \leq 2n \log n \tag{9.18}$$

para n suficientemente grande.

Como f y g son algebraicamente independientes sobre K, nuestra función F no es idénticamente-nula. Sea s el menor entero tal que todas las derivadas de F hasta el orden $s - 1$ se anulan en todos los puntos $w_1, ..., w_m$ pero $D^s F$ no se anula en alguno de los w_v, digamos w. Entonces $s \geq n$. Además usando el Lema 9.4 tenemos que

$$size(D^s F(w)) \leq 5s \log s \tag{9.19}$$

para n (y entonces s) suficientemente grande.

Por otro lado, sea θ una función entera de orden menor o igual a ρ, tal que θf y θg son enteras, y $\theta(w) \neq 0$. Entonces $\theta^{2r} F$ es entera. Consideremos la función entera

$$E(t) = \frac{\theta(t)^{2r} F(t)}{\prod_{v=1}^{m} (t - w_v)^s}. \tag{9.20}$$

Entonces $E(w)$ difiere de $D^s F(w)$ por ciertos factores, acotados por $C_4^s s!$ para alguna constante C_4. Usaremos Módulo Máximo 6.4 para estimar E en un círculo de radio $R = s^{1/2\rho}$. En este círculo $|t - w_v|$ tiene aproximadamente el mismo valor absoluto que R y en consecuencia

$$|E(w)| \leq |E|_R \leq \frac{s^{3s} C_5^{2rR}}{R^{ms}}. \tag{9.21}$$

Obtenemos entonces la estimación

$$\log |D^s F(w)| \leq 4s \log s - \frac{1}{2\rho} ms \log s. \tag{9.22}$$

Comparando esto con la estimación para $size(D^s F(w))$, obtenemos $m \leq 20\rho[K : \mathbb{Q}]$, como queríamos.

□

10 Algunos resultados relacionados y conjeturas

En esta sección veremos un resultado adicional de Lang que puede verse como un complemento del de Gelfond-Schneider, y también algunas conjeturas abiertas relacionadas con los resultados previos.

10.1 Gelfond

El séptimo problema de Hilbert pregunta por la trascendencia del valor e^z en el punto trascendente $z = \beta \log \alpha$, donde $\alpha, \beta \in \mathbb{A}$ con $\alpha \neq 0, 1$ y β irracional. Parece ser que Gelfond fue el primero en estudiar la independencia algebraica [1] de valores de la función exponencial en puntos que no son necesariamente algebraicos. En relación a ello formuló las siguientes dos conjeturas.

Conjetura 10.1. *Sea α un número algebraico distinto de 0 y 1, y sean $\beta_1, ..., \beta_m$ números algebraicos tales que $1, \beta_1, ..., \beta_m$ son linealmente independientes sobre $\mathbb{Q}$. Entonces los números*

$$\alpha^{\beta_1}, ..., \alpha^{\beta_m}$$

son algebraicamente independientes sobre $\mathbb{Q}$.

Conjetura 10.2. *Sean $\alpha_1, ..., \alpha_m$ números algebraicos no nulos, y $\log \alpha_1, ..., \log \alpha_m$ linealmente independientes sobre $\mathbb{Q}$. Entonces $\log \alpha_1, ..., \log \alpha_m$ son algebraicamente independientes sobre $\mathbb{Q}$.*

La primer conjectura, que para $m = 1$ coincide con el seṕtimo problema de Hilbert, es un análogo natural de Lindemann-Weierstrass. Aún no fue probada para $m \geq 2$. Es sencillo ver que es equivalente a la siguiente conjetura.

Conjetura 10.3. *Sea α un número algebraico distinto de 0 y 1, y sea β un número algebraico de grado $d \geq 2$. Entonces los números*

$$\alpha^{\beta}, \alpha^{\beta^2}, ..., \alpha^{\beta^{d-1}}$$

son algebraicamente independientes sobre $\mathbb{Q}$.

Gelfond logró demostrar esta para el caso $d = 3$.

En relación a la segunda conjetura, todo lo que se sabe hasta el momento es que $1, \log \alpha_1, ..., \log \alpha_m$ son linealmente independientes sobre el cuerpo $\mathbb{A}$ de los números algebraicos, lo cual fue probado por Baker.

[1] Dados L, K cuerpos con $L \subset K$, y elementos $b_1, ..., b_n \in K$, se dice que estos elementos son "algebraicamente independientes sobre L" si para cada $p \in L[X_1, ..., X_n]$ no idénticamente-nulo, se tiene que $p(b_1, .., b_n) \neq 0$. De lo contrario se dicen "algebraicamente dependientes".

10.2 Schanuel

Hay otra conjetura sobre los valores de e^z en puntos trascendentes que amplía simultáneamente el resultado de Lindemann-Weierstrass 2.10 y las conjeturas de Gelfond.

Conjetura 10.4 (Schanuel)**.** *Sean $a_1, ..., a_p \in \mathbb{C}$ linealmente independientes sobre $\mathbb{Q}$. Entonces hay por lo menos p de los $2p$ números*

$$a_1, ..., a_p, e^{a_1}, ..., e^{a_p}$$

que son algebraicamente independienetes sobre $\mathbb{Q}$. En otras palabras, el grado de trascendencia sobre $\mathbb{Q}$ de

$$\mathbb{Q}(a_1, ..., a_p, e^{a_1}, ..., e^{a_p})$$

es por lo menos p.

Cuando todos los a_j son algebraicos la conjetura de Schanuel se reduce a Lindemann-Weierstrass2.10. Si tomammos $p = 2, a_1 = 1, a_2 = 2\pi i$ entonces el cuerpo $\mathbb{Q}(a_1, ..., a_p, e^{a_1}, ..., e^{a_p})$ es una extensión algebraica de $\mathbb{Q}(e, \pi)$, y entoncs la conjetura dice que e y π son algebraicamente independientes sobre $\mathbb{Q}$ (otro conocido problema abierto sobre trascendencia). Esta conjetura implica resultados aún más amplios que la tercer conjetura de Gelfond. Supongamos que ponemos $p = d$ y $\alpha_i = \beta^{i-1} \log \alpha$, $i = 1, ..., d$. Entonces el cuerpo $\mathbb{Q}(a_1, ..., a_p, e^{a_1}, ..., e^{a_p})$ es una extensión algebraica de $\mathbb{Q}(\log \alpha, \alpha^{\beta}, \alpha^{\beta^2}, ..., \alpha^{\beta^{d-1}})$. La conjetura de Schanuel implica que los números

$$\log \alpha, \alpha^{\beta}, \alpha^{\beta^2}, ..., \alpha^{\beta^{d-1}}$$

son algebraicamente independientes sobre $\mathbb{Q}$. La segunda conjetura de Gelfond es también una consecuencia de la conjetura de Schanuel.

10.3 Baker

El resultado de Baker asegura la independencia lineal sobre los números algebraicos, de logaritmos de números algebraicos. Esto es menos fuerte que probar su independencia algebraica, sobre lo cual no hay progresos hasta el momento. Se conjeturó la siguiente "Extensión de Baker": si $\lambda_1, ..., \lambda_n$ son "logaritmos de números algebraicos no nulos" linealmente independientes sobre los racionales, entonces son algebraicamente independientes sobre los racionales. Éste es un caso especial de la conjetura de Schanuel. Pero aún debe probarse que siquiera existen dos números algebraicos cuyos logaritmos son algebraicamente independientes.

10.4 Lang

Lang probó en 1966 el siguiente resultado que puede verse como un complemento del resultado de Gelfond.

Teorema 10.5 (6 exponenciales (1966 Lang))**.** *Sean β_1, β_2 números complejos linealmente independientes sobre $\mathbb{Q}$, y z_v $(v = 1, 2, 3)$ números complejos también linealmente independientes sobre $\mathbb{Q}$. Entonces por lo menos uno de los números*

$$e^{\beta_1 z_v},\ e^{\beta_2 z_v} \qquad (v = 1, 2, 3)$$

es trascendente sobre $\mathbb{Q}$.

Veamos a la función e^t como una función entera y observemos que

$$\max_{|t|=R} |e^t| \leq e^R. \tag{10.1}$$

Dadas dos funciones a valores reales f, g escribimos $f << g$ si existe una constante $C > 0$ tal que para todo x suficientemente grande se tiene $f(x) \leq Cg(x)$. Y escribimos $f = O(g)$ si $|f| << g$. Recordamos que una función entera F se dice "de orden $\leq \rho$ " si existe una constante $C > 0$ tal que

$$|F|_R = \max_{|t|=R} |F(t)| \leq C^{R^\rho} \tag{10.2}$$

para todo R suficientemente grande. Con la notación definida antes podemos escribir

$$\log |F|_R = O(R^\rho) \tag{10.3}$$

o también

$$\log |F|_R << R^\rho \tag{10.4}$$

para R suficientemente grande. La constante implícita depende de F. Es un hecho conocido del análisis complejo que el número de ceros de una tal función F en un círculo de radio R es $O(R^\rho)$ si F no es idénticamente nula.

Prueba. Supongamos que la conclusión de 10.5 es falsa, y sea K una extensión finita de $\mathbb{Q}$ que contiene

$$e^{\beta_1 z_v},\ e^{\beta_2 z_v} \qquad (v = 1, 2, 3).$$

Sea n un entero grande, que suponemos cuadrado, y que haremos tender a infinito luego. Sea $r = (4n)^{\frac{3}{2}}$. Podemos encontrar núeros algebraicos a_{ij} no todos nulos en K tales que la función

$$F(t) = \sum_{i,j=1}^{r} a_{ij} e^{i\beta_1 t} e^{j\beta_2 t} \tag{10.5}$$

tenga un cero en cada punto $k.z = k_1 z_1 + k_2 z_2 + k_3 z_3$ donde $k = (k_1, k_2, k_3)$ es una $3-$upla de enteros tales que $1 \leq k_v \leq n$. Esto puede hacerse solucionando un sistema de ecuaciones lineales en r^2 incógnitas, con $r^2 = (4n)^3$ y n^3 ecuaciones.

Los coeficientes de las ecuaciones son los valores

$$e^{i\beta_1(k.z)} e^{j\beta_2(k.z)}$$

que son elementos de K. Para cada ecuacin (correspondiente a algún k), tenemos

$$\text{tamaño de los coeficientes} << nr << n^{\frac{5}{2}},$$

con la constante implícita dependiendo de los valores $e^{\beta_\mu z_v}$. Para cada k, podemos encontrar un denominador común d para los coeficientes de la ecuación $k-$ésima, satisfaciendo la cota

$$\log d << nr, \tag{10.6}$$

porque los coeficientes son potencias de los números algebraicos fijos $e^{\beta_\mu z_v}$, y estas potencias están esencialmente acotadas por nr. Podemos entonces aplicar el Lema de Siegel, y encontrar los a_{ij} tales que

$$size\,(a_{ij}) << n^{\frac{5}{2}}$$

para n suficientemente grande. Como β_1, β_2 son linealmente independientes sobre $\mathbb{Q}$, se sigue que F no es idénticamente nula, y toma valores en K para todas las combinaciones lineales de z_1, z_2, z_3 con coeficientes enteros positivos. Por otro lado, F no puede anularse en todas estas combinaciones lineales, porque ellas forman un conjunto no discreto, o alternativamente porque F es entera de orden 1, y en un círculo de radio grande R hay muchas más de tales combinaciones lineales que la cota $O(R)$ para el número de posibles ceros de F. Sea s el mayor entero para el que $F(k.z) = 0\ \forall k_v$ con $1 \leq k_v \leq s$. Entonces $s \geq n$. Sea

$$2 = k_1 z_1 + k_2 z_2 + k_3 z_3$$

con algún $k_v = s + 1$, y $1 \leq k_v \leq s + 1$ para todo v, y $F(w) \neq 0$. Entonces

$$size\, F(w) << s^{\frac{5}{2}}.$$

Estimamos a continuación $|F(w)|$, y usamos la expresión

$$F(w) = \frac{F(t)}{\prod (t - k.z)} \prod w - k.z|_{t=w}, \tag{10.7}$$

con los productos tomados sobre todos los k_v con $1 \leq k_v \leq s$. Hay s^3 términos en el producto. La función en la derecha de esta última igualdad es entera, y aplicamos Módulo Máximo 6.4, en un círculo de radio $R = s^{\frac{3}{2}}$. Notemos que para $|t| = R$, tenemos $|t - k.z| \geq \frac{R}{2}$ (para s grande), y también

$$\frac{|w - k.z|}{|t - k.z|} \leq \frac{C_1 s}{R} \leq \frac{C_1}{s^{\frac{1}{2}}} \tag{10.8}$$

para alguna constante C_1 y s grande. Entonces

$$\log |F(w)| << \log |F|_R + s^3 - \frac{1}{2} s^3 \log s. \tag{10.9}$$

Además podemos estimar

$$|F|_R \leq r^2 C_2^{n^{\frac{5}{2}}} C_3^{rR} \leq C_4^{s^3}, \tag{10.10}$$

y entonces

$$\log |F(w)| << s^3 - s^3 \log s. \tag{10.11}$$

Esto contradice la cota inferior

$$-\text{size } F(w) << \log |F(w)| \tag{10.12}$$

si hacemos tender n, y entonces s, a infinito lo que concluye la prueba. □

10.5 Otras conjeturas y resultados

A continuación listamos conjeturas y resultados cuya relación con "6 exponenciales" y la "Extensión de Baker" es la siguiente

Extensión de Baker		
4 exponenciales strong	4 exponenciales sharp	4 exponenciales
5 exponenciales strong	5 exponenciales sharp	5 exponenciales (P)
6 exponenciales strong (P)	6 exponenciales sharp (P)	6 exponenciales (P)

donde cada enunciado implica a los que están a su derecha y debajo de él, los que tienen a su lado "(P)" están probados y los demás son conjeturas.

Teorema 10.6 (1988 Waldschmidt, 5 exponenciales)**.** *Sean x_1, x_2 y y_1, y_2 dos pares de números complejos, cada uno linealmente independiente sobre $\mathbb{Q}$, y sea $\gamma \in \mathbb{A}\setminus\{0\}$. Entonces por lo menos uno de los siguiente números es trascendente*

$$e^{x_1y_1}, e^{x_1y_2}, e^{x_2y_1}, e^{x_2y_2}, e^{\gamma\frac{x_2}{x_1}}.$$

Otra conjetura más fuerte que la de las "5 exponenciales" llamada la de las "4 exponenciales" dice que de hecho uno de los primeros cuatro números de la lista debe ser trascendente.

Teorema 10.7 (6 exponenciales sharp)**.** *Sean $x_1, x_2, x_3 \in$ C linealmente independientes sobre $\mathbb{Q}$, y $y_1, y_2 \in \mathbb{C}$ linealmente indpendientes sore $\mathbb{Q}$, y $\beta_{ij} \in \mathbb{A}$ $(1 \le i \le 3,\, 1 \le j \le 2)$ tales que los siguientes seis números son algebraicos*

$$e^{x_1y_1-\beta_{11}}, e^{x_1y_2-\beta_{12}}, e^{x_2y_1-\beta_{21}}, e^{x_2y_2-\beta_{22}}, e^{x_3y_1-\beta_{31}}, e^{x_3y_2-\beta_{32}}.$$

Entonces $x_iy_j = \beta_{ij}$ $(1 \le i \le 3,\, 1 \le j \le 2)$.

Se deduce el resultado de las "6 exponenciales" poniendo $\beta_{ij} = 0$ $\forall i$, $\forall j$, y también se deduce el de las "5 exponenciales" poniendo $x_3 = \frac{\gamma}{x_1}$ y usando el resultado de Baker para asegurar que los x_i sean lineamente independientes.

Conjetura 10.8 (5 exponenciales sharp)**.** *Sean $x_1, x_2 \in$ C linealmente independientes sobre $\mathbb{Q}$, y $y_1, y_2 \in \mathbb{C}$ linealmente indpendientes sore $\mathbb{Q}$, y β_{ij} $(1 \le i \le 2,\, 1 \le j \le 2), \alpha, \gamma \in \mathbb{A}$ con $\gamma \neq 0$ tales que los siguientes cinco números son algebraicos*

$$e^{x_1y_1-\beta_{11}}, e^{x_1y_2-\beta_{12}}, e^{x_2y_1-\beta_{21}}, e^{x_2y_2-\beta_{22}}, e^{(\gamma\frac{x_2}{x_1})-\alpha}.$$

Entonces $x_iy_j = \beta_{ij}$ $(1 \le i \le 2,\, 1 \le j \le 2)$ y $\gamma x_2 = \alpha x_1$.

Una consecuencia de esta conjetura que todavía no se sabe sería la trascendencia de e^{π^2}, poniendo $x_1 = y_1 = \beta^{11} = 1$, $x_2 = y_2 = i\pi$, y todos los demás valores cero.

Teorema 10.9 (1992 Roy, 6 exponenciales strong)**.** *Sea $L' := \{\beta_0 + \sum_{i=1}^{n} \beta_i \log \alpha_i : n \ge 0,\, \beta_i, \alpha_i \in \mathbb{A}\}$. Si $x_1, x_2, x_3 \in \mathbb{C}$ linealmente independientes sobre $\mathbb{A}$, y $y_1, y_2 \in C$ linealmente independientes sobre $\mathbb{A}$, entonces alguno de los seis números siguientes no está en L'*

$$x_iy_j, \quad 1 \le i \le 3,\, 1 \le j \le 2.$$

Este resultado es más fuerte que "6 exponenciales" que dice que uno de estos seis números no es el logartimo de un número algebraico.

Hay también una versión strong de "5 exponenciales" formulada por Waldschmidt

Conjetura 10.10 (5 exponenciales strong)**.** *Sean x_1, x_2 y y_1, y_2 dos pares de números complejos, cada uno linealmente independiente sobre $\mathbb{A}$, y sea $\gamma \in \mathbb{A} \setminus \{0\}$. Entonces por lo menos uno de los siguiente números no está en L'*

$$x_1y_1,\ x_1y_2,\ x_2y_1,\ x_2y_2,\ \frac{x_1}{x_2}.$$

Conjetura 10.11 (4 exponenciales sharp)**.** *Sean $x_1, x_2 \in$ C linealmente independientes sobre $\mathbb{Q}$, y $y_1, y_2 \in \mathbb{C}$ linealmente indpendientes sore $\mathbb{Q}$, y $\beta_{ij}\,(1 \le i \le 2,\ 1 \le j \le 2), \alpha, \gamma \in \mathbb{A}$ con $\gamma \neq 0$ tales que los siguientes cuatro números son algebraicos*

$$e^{x_1y_1-\beta_{11}}, e^{x_1y_2-\beta_{12}}, e^{x_2y_1-\beta_{21}}, e^{x_2y_2-\beta_{22}}.$$

Entonces $x_iy_j = \beta_{ij}$, $1 \le i \le 2$, $1 \le j \le 2$, así que las cuatro exponenciales son de hecho una.

Conjetura 10.12 (4 exponenciales strong)**.** *Sean x_1, x_2 y y_1, y_2 dos pares de números complejos, cada uno linealmente independiente sobre $\mathbb{A}$. Entonces por lo menos uno de los siguiente números no está en L'*

$$x_iy_j, \quad 1 \le i \le 2,\ 1 \le j \le 2.$$

11 Apéndice: Números algebraicos

Incluimos en estas sección los resultados centrales sobre números algebraicos de lo cuales hacemos uso. Notaremos respectivamente: $\mathbb{N}$ al conjunto de números naturales, $\mathbb{Z}$ al anillo de enteros racionales, $\mathbb{Z}^+$ al conjunto de enteros racionales no negativos, $\mathbb{Q}$ al cuerpo de los números racionales, $\mathbb{R}$ al cuerpo de números reales, y $\mathbb{C}$ al cuerpo de números complejos. Y $\mathbb{A}$ o $\overline{\mathbb{Q}}$ al conjunto de los complejos algebraicos sobre $\mathbb{Q}$.

11.1 Número algebraico, número trascendente

Sean K y L dos cuerpos; si $K \subset L$ decimos que L es una "extension" de K. En ese caso L es un K-espacio vectorial, y decimos que L es una "extensión finita" de K si L es un K-espacio vectorial de dimensión finita y notamos a esa dimensión

$$[L : K].$$

Un elemento $\alpha \in L$ se dice "algebraico sobre" K si existe un polinomio no nulo $P \in K[X]$ tal que $P(\alpha) = 0$, es decir si el morfismo canónico $\beta : K[X] \to L$ que deja invariantes los elementos de K y manda X a α, tiene núcleo no nulo (si no existe un polinomio tal decimos que α es "trascendente" o "trascendental" sobre K). En este caso, este núcleo es construido por un polinomio irreducible $p \in K[X]$, y la imagen de β, es decir el subanillo $K[\alpha]$ de L construido sobre K por α, es isomorfo al cuerpo $K[X]/p(X)$. Si tomamos a $p(X)$ unitario (o mónico, es decir el coeficiente de mayor grado es igual a 1) entonces $p(X)$ es único, y decimos que $p(X)$ es el "polinomio irreduclbie de α sobre K" .

Inversamente si el morfismo β asociado a un elemento α de L es inyectivo, entonces diremos que α es "trascendente sobre" K.

Una extension L de K se dice "algebraica" (sobre K) si todo elemento de L es algebraico sobre K. Por ejemplo una extensión finita es algebraica.

Si E es una subconjunto de una extensión L de un cuerpo K, notamos $K(E)$ al subcuerpo de L construido por E sobre K, es decir la intersección de los subcuerpos de L que contienen a K y E. De forma análoga notamos $K[E]$ al subanillo L construido por E sobre K.

Si α es algebraico sobre K tenemos que $K(\alpha) = K[\alpha]$ y el grado del polinomio minimal de α sobre K es igual a $[K(\alpha) : K]$, llamamos a tal número "grado de α sobre K".

Un cuerpo K se dice "algebraicamente cerrado" si todo polinomio no constante (es decir de grado mayor o igual a 1) de $K[X]$ tiene por lo menos una raíz en K. El cuerpo $\mathbb{C}$ de números complejos es un ejemplo. Si K es un cuerpo, entonces existen exstensiones algebraicas de K que son algebraicamente cerradas; si Q es un cuerpo algebraicamente cerrado que contiene a K, el conjunto de elementos de Q algebraicos sobre K se llama "cerramiento algebraico de K sobre Q, y lo notamos $\overline{K}$". Entonces notaremos $\overline{\mathbb{Q}}$ al cerramiento algebraico de $\mathbb{Q}$ en $\mathbb{C}$, es decir el subcuerpo de $\mathbb{C}$ formado por los números complejos algebraicos sobre $\mathbb{Q}$.

Un número complejo se dice "algebraico" (respectivamente "trascendente") si es algebraico sobre $\mathbb{Q}$ (respectivamente trascendente sobre $\mathbb{Q}$). Aplicar operaciones artiméticas a números algebraicos nos da nuevamente números algebraicos, entonces el conjunto de todos los números algebraicos forman un cuerpo que denotamos $\mathbb{A}$.

Sea $\alpha \in \overline{\mathbb{Q}}$ y sea p el polinomio irreducible de α sobre $\mathbb{Q}$. Podemos escribir p en la forma

$$p(X) = X^n + \frac{a_{n-1}}{b_{n-1}}X^{n-1} + \frac{a_0}{b_0} \tag{11.1}$$

donde para todo $i = 0, ..., n-1$, a_i y b_i son números enteros racionales coprimos entre sí con $b_i > 0$. Sea c_n el mínimo común múltiplo de $b_0, ..., b_{n-1}$; notamos

$$c_j = \frac{c_n}{b_j}a_j \qquad 0 \leq j \leq n-1. \tag{11.2}$$

El polinomio $c_n p(X) = c_n X^n + c_{n-1}X^n + c_{n-1}X^{n-1} + ... + c_0 \in \mathbb{Z}[X]$ se llama "polinomio minimal de α sobre $\mathbb{Z}$". Llamamos "grado" de α al grado de p.

Para un número algebraico α , las dos propiedades siguientes son equivalentes:
1) El polinomio minimal de α sobre $\mathbb{Z}$ es unitario (o "mónico", es decir su coeficiente de mayor grado igual a 1), lo que equivale a decir que el polinomio irreducible de α sobre $\mathbb{Q}$ tiene coeficientes enteros racionales.
2) Existe un polinomio unitario no nulo $Q \in \mathbb{Z}[X]$ tal que $Q(\alpha) = 0$.
Decimos entonces que α es "entero algebraico" (sobre $\mathbb{Z}$). El conjunto de enteros algebraicos forma un subanillo de $\overline{\mathbb{Q}}$ que denotamos $\mathbb{Z}_A$. Puede probarse que si α es raíz de un polinomio $\varphi(x) \in \mathbb{Z}[x]$ con coeficiente de mayor grado 1 (no necesariamente irreducible) , entonces α es un entero algebraico.

Teorema 11.1. *Si α es un entero algebraico y también es un número racional, entonces α es un entero racional (es decir $\alpha \in \mathbb{Z}$).*

Prueba. Supongamos que $f(a/b) = 0$ donde $f(x) = \sum_{j=0}^{n} a_j x^j$ con $a_n = 1$ y donde a y b son coprimos con $b > 0$. Alcanza con probar que $b = 1$. De $f(a/b) = 0$ se obtiene que

$$a^n + a_{n-1}a^{n-1}b + ... + a_1 ab^{n-1} + a_0 b^n = 0. \tag{11.3}$$

Se sigue que a^n tiene a b como factor. Ya que $mcd(a, b) = 1$ y $b > 0$, deducimos que $b = 1$. □

Para probar los resultados siguientes necesitamos algunos sobre funciones simétricas.

Definición 11.2. *Un polinomio se llama "polinomio simétrico en $\alpha_1, ..., \alpha_n$ " si sigue siendo el mismo polinomio cuando se permutan de cualquier forma los $\alpha_1, ..., \alpha_n$. Denotamos*

$$\sigma_1 := \alpha_1 + ... + \alpha_n$$

$$\sigma_2 := \alpha_1\alpha_2 + \alpha_1\alpha_3 + ... + \alpha_{n-1}\alpha_n$$

y siguiendo así hasta

$$\sigma_n := \alpha_1 ... \alpha_n$$

Éstos se llaman "polinomios simétricos elementales en $\alpha_1, ..., \alpha_n$", son (excepto por el signo) los coeficientes del polinomio $(x - \alpha_1)...(x - \alpha_n)$.

Teorema 11.3. *Sea R un anillo conmutativo con unidad. Entonces todo polinomio simétrico en $\alpha_1, ..., \alpha_n$ con coeficientes en R puede expresarse como un polinomio en $\sigma_1, ...\sigma_n$ con coeficientes en R.*

Prueba. Para un polinomio simétrico $h(\alpha_1, ..., \alpha_n) \in R[\alpha_1, ..., \alpha_n]$, ponemos $T = T_h$ el conjunto de las n-uplas $(l_1, ..., l_n)$ con el coeficiente de $\alpha_1^{l_1}...\alpha_n^{l_n}$ en $h(\alpha_1, ..., \alpha_n)$ no nulo. Definimos el "tamaño" de h como $(k_1, ..., k_n)$ donde $(k_1, ..., k_n)$ es el elemento de T con k_1 tan grande como sea posible, k_2 tan grande como sea posible dado k_1, y siguiendo así.

Como $h(\alpha_1, ..., \alpha_n)$ es simétrico, se sigue que $(l_1, ..., l_n) \in T$ si y solo si cada permutación de $(l_1, ..., l_n)$ está en T. Esto implica que $k_1 \geq k_2 \geq ... \geq k_n$. Observemos que podemos usar la noción de tamaño para ordenar los elementos de $R[\alpha_1, ..., \alpha_n]$ en el sentido de que si h_1 tiene tamaño $(k_1, ..., k_n)$ y h_2 tiene tamaño $(k_1', ..., k_n')$, entonces $h_1 > h_2$ si existe un $i \in \{0, 1, ..., n-1\}$ tal que $k_1 = k_1', ..., k_i = k_i'$, y $k_{i+1} > k_{i+1}'$. Observemos que los elementos de $R[\alpha_1, ..., \alpha_n]$ que tienen tamaño $(0, 0, ..., 0)$ son precisamente las constantes (los elementos de R).

Supongamos que $(k_1, ..., k_n)$ representa el tamaño de algún simétrico $g \in R[\alpha_1, ..., \alpha_n]$ con $g \notin R$. Para enteros no negativos $d_1, ..., d_n$, el tamanño de $h = \sigma_1^{d_1}...\sigma_n^{d_n}$ es $(d_1 + d_2 + ... + d_n, d_2 + ... + d_n, ..., d_{n-1} + d_n, d_n)$. Tomando $d_1 = k_1 - k_2, d_2 = k_2 - k_3, ..., d_{n-1} = k_{n-1} - k_n$ y $d_n = k_n$, tenemos que el tamanõ de h es $(k_1, ..., k_n)$. El coeficiente de $h = \alpha_1^{k_1}...\alpha_n^{k_n}$ en h es 1. Se sigue que existe un $a \in R$ tal que $g - ah$ es de tamaño menor que g.

Lo anterior implica que para cualquier simétrico $f \in R[\alpha_1, ..., \alpha_n]$, existen $a_1, ..., a_m \in R$ y $h_1, ..., h_m \; in R[\sigma_1, ...\sigma_n]$ tales que $f - a_1h_1 - ... - a_mh_m$ tiene tamaño $(0, 0, ..., 0)$. Esto implica el resultado buscado. □

A partir de (11.2) y (11.3) es sencillo ver que dado $f(x) = \sum_{j=0}^{n} a_j x^j$ un polinomio no nulo en $\mathbb{C}[X]$ con grado n y raíces no necesariamente distintas $\alpha_1, ..., \alpha_n$ se tiene que

$$f(x) = a_n \prod_{j=1}^{n} (x - \alpha_j) = a_n x^n - a_n \sigma_1 x^{n-1} + a_n \sigma_2 x^{n-2} + ... + (-1) a_n \sigma_n, \tag{11.4}$$

donde viendo a los σ_j como funciones simétricas elementales en los números $\alpha_1, ..., \alpha_n$ se sigue que

$$\sigma_1 = -\frac{a_{n-1}}{a_n}, \qquad \sigma_2 = \frac{a_{n-2}}{a_n}, ..., \sigma_n = (-1)^n \frac{a_0}{a_n}. \tag{11.5}$$

De (11.3) se deduce también el siguiente resultado.

Lema 11.4. *Sea $\alpha \in \mathbb{A}, \deg \alpha = n$ y $\alpha = \alpha_1, ..., \alpha_n$ los conjugados de α . Supongamos además que*

$$F(x_1, ..., x_k; \alpha_1, ..., \alpha_n) \in \mathbb{Q}[x_1, ..., x_k; \alpha_1, ..., \alpha_n] \qquad k \geq 0 \tag{11.6}$$

y que como polinomio en $\alpha_1, ..., \alpha_n$ con coeficientes en $\mathbb{Q}[x_1, ..., x_k]$, F es un polinomio simétrico en $\alpha_1, ..., \alpha_n$. Entonces

$$F = F(x_1, ..., x_k) \in \mathbb{Q}[x_1, ..., x_k] \qquad k > 0 \tag{11.7}$$

y $F \in \mathbb{Q}$ en el caso $k = 0$. Además si tenemos $\alpha \in \mathbb{Z}_{\mathbb{A}}$ y

$$F = F(x_1, ..., x_k; \alpha_1, ..., \alpha_n) \in \mathbb{Z}[x_1, ..., x_k; \alpha_1, ..., \alpha_n] \qquad k \geq 0 \tag{11.8}$$

entonces se tiene que

$$F = F(x_1, ..., x_k) \in \mathbb{Z}[x_1, ..., x_k]; \qquad k > 0 \tag{11.9}$$

y $F \in \mathbb{Z}$ *en el caso* $k = 0$.

Continuemos con las pruebas sobre números algebraicos.

Teorema 11.5. *Los números algebraicos forman un cuerpo. Los enteros algebraicos forman un anillo.*

Prueba. Para probar estas 2 afirmaciones, supondremos que α y β son números algebraicos o enteros algebraicos, y provaremos que $-\alpha, \alpha + \beta$ y $\alpha\beta$ también lo son (respectivamente). En el caso de que α sea un número algebraico no nulo, mostraremos que $1/\alpha$ también lo es.

• El caso para $-\alpha$. Si $f(x)$ es un polinomio con coeficientes enteros con α como raíz, entonces consideramos $\pm f(-x)$. Si $f(x)$ es mónico entonces uno de éstos también lo será. Entonces, si α es un número algebraico también lo es $-\alpha$; y si α es un entero algebraico, entonces también lo es $-\alpha$.

• El caso para $\alpha+\beta$. Supongamos que α es una raíz de $f(x) \in \mathbb{Z}[X]$ y que β es una raíz de $g(x) \in \mathbb{Z}[X]$. Denotemos $\alpha_1 = \alpha, \alpha_2, ..., \alpha_n$ el conjunto completo de raíces de $f(x)$ (repetidas según su multiplicidad, de forma que el grado de $f(x)$ se n), y denotemos $\beta_1 = \beta, \beta_2, ..., \beta_n$ el conjunto completo de raíces de $g(x)$. Consideremos el polinomio

$$F(x) \prod_{i=1}^{n} \prod_{j=1}^{m} [x - (\alpha_i + \beta_j)]. \tag{11.10}$$

Tomando $R = \mathbb{Z}[\beta_1, ..., \beta_m]$ en (11.3), vemos que los coeficientes de $F(x)$ son polinomios simétricos en $\alpha_1, ..., \alpha_n$. Entonces si $\sigma_1, ..., \sigma_n$ corresponden a las funciones simétricas elementales en $\alpha_1, ..., \alpha_n$ y A es algún coeficiente (de x^k) en $F(x)$, entonces $A = B(\sigma_1, ..., \sigma_n, \beta_1, ..., \beta_m)$ para algún polinomio B con coeficientes enteros.

Por otro lado, los coeficientes de $F(x)$ también son simétricos en $\beta_1, ..., \beta_m$. Tomando $R = [\sigma_1, ..., \sigma_n]$ en (11.3), y siendo $\sigma'_1, ..., \sigma'_m$ las funciones simétricas elementales en $\beta_1, ..., \beta_m$, obtenemos que $A = B'(\sigma_1, ..., \sigma_n, \sigma'_1, ..., \sigma'_m)$ para algún polinomio B' con coeficientes enteros.

Además (11.5) implica que $\sigma_1, ..., \sigma_n, \sigma'_1, ..., \sigma'_m$ son todos racionales, entonces $A \in \mathbb{Q}$. Entonces $F(x) \in \mathbb{Q}[x]$ y $m'F(x) \in \mathbb{Z}[x]$ para algún entero m'.

Como $\alpha + \beta$ es una raíz de $m'F(x)$, deducimos que $\alpha + \beta$ es un número algebraico.

Si α y β son enteros algebraicos, entonces podemos tomar los coeficientes de mayor grado de $f(x)$ y $g(x)$ iguales a 1 y así (11.5) implica que cada uno de $\sigma_1, ..., \sigma_n, \sigma'_1, ..., \sigma'_m$ está en $\mathbb{Z}$, y entonces $F(x) \in \mathbb{Z}[X]$. Como $F(x)$ es mónico, obtenemos que en este caso $\alpha + \beta$ es un entero algebraico.

• El caso para $\alpha\beta$. La idea del caso anterior funciona también para mostrar que $\alpha\beta$ es un número algebraico (o un entero algebraico), definiendo

$$F(x) = \prod_{i=1}^{n} \int_{j=1}^{m} (x - \alpha_i \beta_j). \tag{11.11}$$

• El caso para $\frac{1}{\alpha}$. Supongamos que $\alpha \neq 0$ y α es una raíz de $\sum_{j=0}^{n} a_j x^j \in \mathbb{Z}[x]$. Entonces es sencillo probar que $\frac{1}{\alpha}$ es una raíz de $\sum_{j=0}^{n} a_{n-j} x^j \in \mathbb{Z}[x]$. Entonces $\frac{1}{\alpha}$ es un entero algebraico.

Completamos así la prueba del resultado buscado. □

Supongamos que α es un entero algebraico no nulo. Observemos que $\frac{1}{\alpha}$ es un entero algebraico si y solo si α es raíz de un polinomio mónico en $\mathbb{Z}[x]$ con término constante ± 1. □

Teorema 11.6. *Si α es un número algebraico, entonces existe un entero racional positivo d tal que $d\alpha$ es un entero algebraico.*

Prueba. Supongamos que α es una raíz de $f(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{Z}[x]$ con $a_n \neq 0$. Considerando $-f(x)$ si es necesario, podemos suponer que $a_n > 0$. Como α es una raíz de $a_n^{n-1} f(x) = \sum_{j=0}^{n} a_j a_n^{n-j-1}$, se sigue que $a_n \alpha$ es una raíz de un polinomio mónico. El resultado se obtiene tomadno $d = a_n$. □

Es conveniente recordar el argumento del resultado anterior además del resultado mismo. Así si uno sabe que un polinomio $f(x)$ tiene por raíz a α, entonces uno podrá saber el valor de d en el resultado.

Llamamos a un tal d "denominador" de α . Además se tiene que si E es un denominador para η y $g(\eta) \in Q[\eta]$ de grado s, entonces E^s es un denominador para $g(\eta)$.

Recordemos que dado un número algebraico α llamamos "polinomio minimal para α" (en $\mathbb{Q}[x]$) al polinomio mónico en $\mathbb{Q}[x]$ de grado minimal que tiene a α como raíz.

Queremos probar que

Teorema 11.7. *El polinomio minimal para un número algebraico α está en $\mathbb{Z}[x]$ si y solo si α es un entero algebraico.*

Definición 11.8. *Sea $f(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{Z}[x]$ no idénticamente-nulo. Entonces el "contenido" de $f(x)$ es $mcd(a_n, a_{n-1}, ..., a_1, a_0)$. Si el contenido de $f(x)$ es 1, entonces llamamos a $f(x)$ "primitivo".*

Lema 11.9. *Si $u(x)$ y $v(x)$ son polinomios primitivos, entonces $u(x)v(x)$ también lo es.*

Prueba. Alcanza con probar que el contenido de $u(x)v(x)$ no es divisible por cada primo. Sea p un primo. Escribamos

$$u(x) = \sum_{j=0}^{n} a_j x^j \qquad \text{y} \qquad v(x) = \sum_{j=0}^{m} b_j x^j.$$

Sean k y l enteros no negativos tan chicos como sea posible tales que p no divide ni a a_k ni a b_l; éstos existen ya ue $u(x)$ y $v(x)$ son primitivos. Se verifica que el coeficiente de x^{k+l} no es divisible por p, y de esto se deduce que el contenido de $u(x)v(x)$ no puede ser divisible por p, completando la prueba. □

Teorema 11.10 (Gauss)**.** *Sea $f(x) \in \mathbb{Z}[x]$. Supongamos que existen $u_1(x)$ y $v_1(x)$ en $\mathbb{Q}[x]$ tales que $f(x) = u_1(x)v_1(x)$. Entonces existen $u_2(x)$ y $v_2(x)$ en $\mathbb{Z}[x]$ tales que $f(x) = u_2(x)v_2(x)$ y $\deg u_2 = \deg u_1$ y $\deg v_2 = \deg v_1$.*

El resultado implica que si $f(x) \in \mathbb{Z}[x]$ tiene contenido 1, entonces una condición necesaria y suficiente para que $f(x)$ sea irreducible sobre los racionales, es que sea irreducible sobre los enteros. Además, observemos que la prueba mostrará que además uno puede tomar a $u_2(x)$ y $v_2(x)$ iguales a números racionales multiplicados por $u_1(x)$ y $v_1(x)$ respectivamente.

Prueba. de 11.10 Sea d el contenido de $f(x)$. Entonces existen enteros racionales positivos a y b y polinomios primitivos $u(x)$ y $v(x)$ en $\mathbb{Z}[x]$ con $\deg u = \deg u_1$ y $\deg v = \deg v_1$ tales que $u_1(x)v_1(x) = \frac{a}{b}u(x)v(x)$. Entonces existe un primitivo $g(x) \in \mathbb{Z}[X]$ para el cual $f(x) = dg(x)$ y $bdg(x) = bf(x) = au(x)v(x)$. Por el Lema 11.9 $u(x)v(x)$ es primitivo. Se sigue que el contenido de $au(x)v(x)$ es a. Como $g(x)$ es primitivo, el contenido de $bdg(x)$ es bd. Entonces $a = bd$. Ponemos $u_2(x) = du(x)$ y $v_2(x) = v(x)$. Entonces $f(x) = u_1(x)v_1(x) = du(x)v(x) = u_2(x)v_2(x)$, y obtenemos el resultado buscado. □

Probemos a continuación el resultado (11.7).

Prueba. de 11.7 Sabemos que si el polinomio minimal para α está en $\mathbb{Z}[x]$, entoncs α es un entero algebraico.

Consideremos un entero algebraico α y sea $f(x) \in \mathbb{Z}[x]$ mónico con $f(\alpha) = 0$. Sea $u_1(x)$ el polinomio minimal de α. Queremos probrar que $u_1(x) \in \mathbb{Z}$.

Por el algoritmo de división para polinomios en $\mathbb{Q}[X]$, existen $v_1(x)$ y $r(x) \in \mathbb{Q}[X]$ tales que $f(x) = u_1(x)v_1(x) + r(x)$ y o bien $r = 0$ o bien $0 \leq \deg r(x) < \deg u_1(x)$. Observemos que $r(\alpha) = f(\alpha) - u_1(\alpha)v_1(\alpha) = 0$. Ya que $u_1(x)$ es el polinomio mónico de grado mínimo que tiene a α como raíz, se sigue que $r(x) \equiv 0$ (de otra forma existiría un $k \in \mathbb{Z}$ para el cual $(1/k)r(x) \in \mathbb{Q}[x]$ es mónico, de menor grado que $u_1(x)$ y con α como raíz). Entonces $f(x) = u_1(x)v_1(x)$ es una factorización de $f(x)$ en $\mathbb{Q}[x]$.

Por el Lema de Gauss y el comentario a su continuación, existen $u_2(x)$ y $v_2(x \in \mathbb{Z}[x])$ con $f(x) = u_2(x)v_2(x)$ y con $u_2(x) = mu_1(x)$ para algún número racional no nulo m. Considerando $f(x) = [-u_2(x)][-v_2(x)]$ si es necesario, podemos suponer que el coeficiente de mayor grado de $u_2(x)$ es positivo. Como $f(x)$ es mónico, deducimos que $u_2(x)$ es mónico. Comparando los coeficiente de mayor grado en $u_2(x) = mu_1(x)$, vemos que $m = 1$ y entonces $u_1(x) = u_2(x) \in \mathbb{Z}[X]$ como queríamos. □

11.2 Cuerpo de números algebraicos

Dado α número algebraico, entonces $\mathbb{Q}(\alpha)$ se define como el menor cuerpo que contiene simultáneamente a los racionales y a α.

Sea $f(x) \in \mathbb{Q}[X]$ el polinomio minimal para α. Considerando cada entero $j \geq 0$ sucesivamente y $\alpha^j f(\alpha) = 0$, se prueba que α^{n+j} puede expresarse como un polinomio en α con coeficientes en $\mathbb{Q}$ de grado $\leq n-1$. Se sigue que $\mathbb{Q}(\alpha)$ es el conjunto de todos los números de la forma $\frac{g(\alpha)}{h(\alpha)}$, donde $g(x)$ y $h(x) \in \mathbb{Z}[x]$, $\deg g(x) \leq n-1$,$\deg h(x) \leq n-1$, y $h(\alpha) \neq 0$. Por el resultado (11.5), cada elemento de $\mathbb{Q}(\alpha)$ es un número algebraico. Por este motivo, nos referimos a $\mathbb{Q}(\alpha)$ como un "cuerpo de número algebraico".

Lema 11.11. *El conjunto de enteros algebraicos contenido en $\mathbb{Q}(\alpha)$ forma un anillo.*

Prueba. Si α y β están en $\mathbb{Q}(\alpha)$, entonces también están $\alpha\beta$ y $\alpha - \beta$ ya que $\mathbb{Q}(\alpha)$ es un cuerpo. Si también α y β son enteros algebraicos, entonces el resultado 11.5implica que $\alpha\beta$ y $\alpha - \beta$ son enteros algebraicos. De lo cual se sigue el resultado buscado. □

Lema 11.12. *Sea α un número algebraico y $f(x) \in \mathbb{Q}[x]$, su polinomio minimal. Sea $g(x) \in \mathbb{Q}[x]$ tal que $g(\alpha) = 0$. Si β es tal que $f(\beta) = 0$, entonces $g(\beta) = 0$. Aún más, $g(x)$ es divisible por $f(x)$ en $\mathbb{Q}[X]$.*

Prueba. Consideremos $q(x), r(x) \in \mathbb{Q}[X]$ tales que $g(x) = f(x)q(x) + r(x)$ donde o bien $r = 0$ o bien $\deg r(x) < \deg f(x)$. Como $f(x)$ es el polinomio minimal de α, se deduce que $r = 0$, y se obtiene el resultado buscado. □

Si α es un número algebraico, su polinomio minimal $f(x)$ es irreducible. Si α' es otra raíz de $f(x)$, entonces por (11.12) se tiene que el polinomio minimal de α' divide $f(x)$. Se sigue que $f(x)$ es también el polinomio minimal para α'. En otras palabras tenemos:

Teorema 11.13. *Sea $f(x)$ el polinomio minimal para α y sean $\alpha_2, ..., \alpha_n$ las otras raíces de $f(x)$. Entonces $f(x)$ es irreducible y es también el polinomio minimal para $\alpha_2, ..., \alpha_n$.*

Teorema 11.14 (Cuerpo algebraico)**.** *Con el número α construye el "cuerpo numérico algebraico" $\mathbb{Q}(\alpha)$ de grado m sobre Q consistente en todos los polinomios en α de grado $\leq m-1$ con coeficientes en Q. Todo elemento γ de $\mathbb{Q}(\alpha)$ es un número algebraico sobre Q de grado d, donde d es un divisor de m, y los conjugados de γ se obtienen reemplzando α por sus conjugados en el polinomio definiente. Si β es un número algebraico sobre $\mathbb{Q}(\alpha)$ de grado n, es decir la raíz de un ecuación de grado n con coeficientes en $\mathbb{Q}(\alpha)$ e irreducible sobre $\mathbb{Q}(\alpha)$, entonces el "cuerpo de extensión" $\mathbb{Q}(\alpha, \beta)$ consiste en todos los polinomios en β degrado $\leq n - 1$ con coeficientes en $\mathbb{Q}(\alpha)$, y sus elementos son de grado $\leq mn$ sobre R. De forma similar para cuerpos de extensión subsiguientes.*

11.3 Conjugados

Sea β un número algebraico con polinomio minimal $g(x)$. A las raíces de $g(x)$ las llamamos "conjugados" de β . Supongamos que $\beta \in \mathbb{Q}(\alpha)$ donde α es un núnmero algebraico con polinomio minimal $f(x)$. Si $\deg f = n$, entonces podemos encontrar un $h(x) \in \mathbb{Q}[x]$, con $h = 0$ o bien $\deg h \leq n - 1$, tal que $\beta = h(\alpha)$. Sea $\alpha_1 = \alpha, \alpha_2, ..., \alpha_n$ las n raíces de $f(x)$, y sean $\beta_1 = \beta, \beta_2, ..., \beta_m$ las raíces de $g(x)$. Entonces $\beta_1, \beta_2, ..., \beta_m$ son los conjugados de β . Los números $h(\alpha_1), h(\alpha_2), ..., h(\alpha_n)$ los llamamos "conjugados de cuerpo" de β en $\mathbb{Q}(\alpha)$.

Teorema 11.15. *Con la notación anterior, $m|n$ y $h(\alpha_1), h(\alpha_2), ..., h(\alpha_n)$ es algún ordenamiento de $\frac{n}{m}$ copias de $\beta_1, \beta_2, ..., \beta_m$. En otras palabras si $F(x) = \prod_{j=1}^{n} [x - h(\alpha_j)]$, entonces $F(x) = g(x)^{\frac{n}{m}}$. También, si $F(x) = g(x)$ (esto es si $n = m$), entonces $\mathbb{Q}(\alpha) = \mathbb{Q}(\beta)$.*

Prueba. Como $F(x)$ es simétrico en $\alpha_1, \alpha_2, ..., \alpha_n$, deducimos que $F(x) \in \mathbb{Q}[X]$. Escribimos

$$F(x) = f_1(x)f_2(x)...f_r(x)$$

donde cada $f_j(x)$ es un polinomio mónico irreducible en $\mathbb{Q}[x]$. Además elegimos $f_1(x)$ tal que $f_1(h(\alpha)) = 0$. Entonces $f_1(\beta) = 0$, Como $g(x)$ y $f_1(x)$ son polinomios mónicos irreducibles con $g(\beta) = f_1(\beta) = 0$, obtenemos que $f(x) \equiv g(x)$.

Cada $f_j(x)$ restante tiene algún (no necesariamente el mismo) $h(\alpha_i)$ como raíz. Observemos que $f_j(h(\alpha_i)) = 0$ implica que α_i es raíz de $f_j(h(x))$. Pero esto implica que $f(x)$ divide $f_j(h(x))$, y entonces $f_j(h(\alpha)) = 0$. Como hicimos con $f_1(x)$, deducimos que $f_j(x) \equiv g(x)$. Entonces, obtenemos $F(x) = g(x)^r$. Comparando grados vemos que $r = \frac{n}{m}$.

Supongamos que $n = m$ de forma que $F(x) = g(x)$. Como sabemos que $\beta \in \mathbb{Q}(\alpha)$, alcanza con mostrar que $\alpha \in \mathbb{Q}(\beta)$ para establecer que $\mathbb{Q}(\alpha) = \mathbb{Q}(\beta)$. Obervemos que

$$G(x) = F(x) \sum_{j=1}^{n} \frac{\alpha_j}{x - h(\alpha_j)}$$

es un polinomio simétrico en $\alpha_1, ..., \alpha_n$, y entonces $G(x) \in \mathbb{Q}[x]$. Observemos que

$$G(\beta) = G(h(\alpha)) = \alpha F'(h(\alpha)) = \alpha F'(\beta).$$

Como β es raíz de $g(x)$ con multiplicidad uno, tenemos que $F'(\beta) = g'(\beta) \neq 0$. Deducimos que $\alpha = \frac{G(\beta)}{F'(\beta)} \in \mathbb{Q}(\beta)$, y obtenemos así el resultado buscado. □

Al polinomio $F(x)$ de (11.15) lo llamamos "polinomio de cuerpo" para β .

Lema 11.16. *Con la notación anterior, sea $w(x) \in \mathbb{Q}[X]$ con $\beta = w(\alpha)$ (no requerimos $\deg w \leq n - 1$). Entonces para cada $j \in \{1, 2, ..., n\}$ el conjugado de cuerpo $\beta_j = h(\alpha_j)$ satisface $\beta_j = w(\alpha_j)$.*

Prueba. Dividimos $w(x)$ por $f(x)$ (el polinomio minimal para α), para obtener $w(x) = f(x)q(x) + r(x)$ donde $q(x)$ y $r(x)$ están en $\mathbb{Q}[x]$ con o bien $r(x) \equiv 0$ o bien $\deg r \leq n - 1$. Entonces $\beta = w(\alpha) = r(\alpha)$, entonces $r(x) \equiv h(x)$. Obtenemos el resultado buscado ya que

$$w(\alpha_j) = f(\alpha_j)q(\alpha_j) + r(\alpha_j) = r(\alpha_j) = h(\alpha_j)$$

para cada $j \in \{1, 2, ..., n\}$. □

11.4 Norma, tamaño, longitud, altura

Sea $\beta \in \mathbb{Q}(\alpha)$ donde α es un número algebraico, sean $\beta_1, ..., \beta_n$ los conjugados de cuerpo de β . La norma de β está definida por

$$N(\beta) = N_{\mathbb{Q}(\alpha)}(\beta) = N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\beta) = \beta_1 ... \beta_n.$$

Notemos que si $F(x) = \sum_{j=0}^{n} a_j x^j$ es el polinomio de cuerpo para β (de modo que $a_n = 1$), entonces

$$N(\beta) = (-1)^n a_0.$$

También si $g(x) = \sum_{j=0}^{m} b_j x^j$ es el polinomio minimal para β como en (11.15) con raíces $\beta_1, ..., \beta_m$ entonces

$$N(\beta) = (\beta_1 ... \beta_m)^{\frac{n}{m}} = (-1)^n b_0^{\frac{n}{m}}.$$

Teorema 11.17. *Sean β y $\gamma \in \mathbb{Q}(\alpha)$. Entonces $N(\beta\gamma) = N(\beta)N(\gamma)$.*

Prueba. Sea n el grado de la extensión $\mathbb{Q}(\alpha)$ sobre $\mathbb{Q}$ (definido como elgrado del polinomio minimal para α). Entonces existen únicos números racionales $b_0, ..., b_{n-1}$ y $c_0, ..., c_{n-1}$ tales que

$$\beta = \sum_{j=0}^{n-1} b_j \alpha^j \qquad \text{y} \qquad \gamma == \sum_{j=0}^{n-1} c_j \alpha^j.$$

Pongamos

$$g(x) = \sum_{j=0}^{n-1} b_j x^j \qquad \text{y} \qquad h(x) = \sum_{j=0}^{n-1} c_j x^j.$$

Entonces $g(\alpha_1), ..., g(\alpha_n)$ son los conjugados de cuerpo de β y $h(\alpha_1), ..., h(\alpha_n)$ son los conjugados de cuerpo de γ. Sea $w(x) = g(x)h(x) \in \mathbb{Q}[x]$, de modo que $\beta\gamma = w(\alpha)$. Entoces el lema anterior (11.16) implica que

$$N(\beta\gamma) = w(\alpha_1)...w(\alpha_n) = g(\alpha_1)h(\alpha_1)...g(\alpha_n)h(\alpha_n) = N(\beta)N(\gamma),$$

completando así la prueba. □

De lo anterior deducimos que

Lema 11.18. *Sea $\beta \in \mathbb{Q}(\alpha)$. Entonces $N(\beta) \in \mathbb{Q}$. Si β es un entero algebraico, entonces $N(\beta) \in \mathbb{Z}$.*

En el caso de $\alpha \in \mathbb{A}$ notaremos a su norma $N(\alpha) = \alpha_1...\alpha_n$ donde α_i son los conjugados de α sobre $\mathbb{Q}$. Se tienen para ella las siguientes propiedades: 1) $N(\alpha) \in \mathbb{Q} \qquad \forall \alpha \in \mathbb{A}$
2) $N(\alpha) \in \mathbb{Z} \qquad \forall \alpha \in \mathbb{Z}_{\mathbb{A}}$
3) $N(\alpha) = 0$ si y solo si $\alpha = 0$
4) $N(\alpha) = \alpha^h$ si $\alpha \in \mathbb{Q}$
5) $N(\alpha\beta) = N(\alpha)N(\beta) \qquad \forall \alpha, \beta \in \mathbb{A}$
6) $N(a\alpha) = a^h N(\alpha)$ para $a \in \mathbb{Q}, \alpha \in \mathbb{A}$

Definimos la "longitud" L de $\alpha \in \mathbb{C}$ como

$$L(\alpha) = \sum_{i=0}^{n} |c_i|$$

donde c_i son los coeficientes del polinomio minimal de α sobre $\mathbb{Z}$.

Cuando $K = \mathbb{Q}$ y $\alpha \in \overline{\mathbb{Q}}$, notamos

$$\overline{|\alpha|} = \max_{1 \le j \le n} |\alpha_j| \tag{11.12}$$

o también $\mu(\alpha)$. Definimos el "tamaño" $s(\alpha)$ o bien $size(\alpha)$ de α por

$$s(\alpha) = \max(\log \overline{|\alpha|}, \log d(\alpha)) \tag{11.13}$$

Recordemos que $d(\alpha)$ designa al denominador de α , es decir el menor de los enteros racionales positivos tales que $d\alpha$ es un entero algebraico.

La propiedad esencial del tamaño es la siguiente

Lema 11.19. *Si α es un número algebraico de grado menor o igual a n, tenemos que*

$$-2ns(\alpha) \leq \log|\alpha|$$

Para ver esto notamos que la norma

$$N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(d(\alpha)\alpha) = \prod_{j=1}^{n} d(\alpha)\alpha_j \tag{11.14}$$

sobre $\mathbb{Q}$ el $d(\alpha)\alpha$ es un entero racional no nulo, entonces

$$\prod_{j=1}^{n} d(\alpha)|\alpha_j| \geq 1. \tag{11.15}$$

Deducimos que

$$-n\log d(\alpha) - (n-1)\log\overline{|\alpha|} \leq \log|\alpha| \tag{11.16}$$

de donde se obtiene la relación (11.19) □

Para calcular el tamaño de ciertos números algebraicos utilizaremos las propiedades siguientes:

$$d(\alpha\beta) \leq d(\alpha)d(\beta), \qquad d(\alpha+\beta) \leq d(\alpha)d(\beta) \qquad d(a\alpha) \leq ad(\alpha) \qquad d(\alpha^m) \leq (d(\alpha))^m \tag{11.17}$$

$$\overline{|\alpha\beta|} \leq \overline{|\alpha|}\overline{|\beta|}, \qquad \overline{|\alpha+\beta|} \leq \overline{|\alpha|} + \overline{|\beta|} \qquad \overline{|a\alpha|} \leq a\overline{|\alpha|} \qquad \overline{|\alpha^m|} \leq \overline{|\alpha|}^m \tag{11.18}$$

para $\alpha, \beta \in \overline{\mathbb{Q}}$ y $a, m \in \mathbb{N}$.

Deducimos, para $\alpha_1, ..., \alpha_m \in \overline{\mathbb{Q}}$, que

$$s(\alpha_1...\alpha_m) \leq s(\alpha_1) + ... + s(\alpha_m) \qquad s(\alpha_1 + ... + \alpha_m) \leq s(\alpha_1) + ... + s(\alpha_m) + \log m \tag{11.19}$$

Si además $\alpha_1, ..., \alpha_m$ son enteros algebraicos, tenemos que

$$s(\alpha_1 + ... + \alpha_m) \leq \max_{1\leq h\leq m} s(\alpha_h) + \log m. \tag{11.20}$$

Los números algebraicos a los cuales les calcularemos su tamaño, estarán dados como valores de polinomios con coeficientes enteros racionales en puntos algebraicos. Por esta razón introducimos en lo siguiente las nociones de altura y tamaño para polinomios.

Sea $P \in \mathbb{C}[X_1, ..., X_q]$ un polinomio no nulo en q variables con coeficientes complejos. Notamos

$$deg_{X_i} P \tag{11.21}$$

al grado de P respecto a X_i y

$$H(P) \tag{11.22}$$

a la "altura de P", es decir al máximo de los valroes absolutos de los coeficientes de P. Definimos la "altura" de un $\alpha \in \mathbb{C}$ como $H(\alpha) := H(P)$ donde P es su polinomio minimal.

Por otro lado, si $P \in \overline{\mathbb{Q}}[X_1, ..., X_q]$ con coeficientes enteros algebraicos, notamos

$$\overline{|P|} \tag{11.23}$$

al máximo de los valores absolutos de los conjugados de los coeficientes de P, y definimos el "tamaño de P" por

$$t(P) = \max\{\log \overline{|P|}, \max_{1 \le i \le q} (1 + deg_{X_i} P)\}. \tag{11.24}$$

Observemos que , para $P \in \mathbb{Z}[X_1, ..., X_q]$,tenemos que

$$H(P) = \overline{|P|} \tag{11.25}$$

Deducimos entonces de las propiedades de la función s (11.19) el resultado siguiente.

Lema 11.20. *Sean $\alpha_1, ..., \alpha_q$ números algebraicos, y sea $P \in \mathbb{Z}[X_1, ..., X_q]$ un polinomio de grado menor o igual a r_i respecto a X_i $\quad 1 \le i \le q$. Entonces*

$$P(\alpha_1, ..., \alpha_q) = \beta$$

es un número algebraico,

$$d(\alpha_1)^{r_1}, ..., d(\alpha_q)^{r_q}$$

es un denominador de β y tenemos que

$$s(\beta) \le \log H(P) + \sum_{i=1}^{q} [r_i s(\alpha_i) + \log(r_i + 1)]$$

Para refinar algo las desigualdades definimos para

$$P(X_1, ..., X_q) = \sum_{\lambda_1=0}^{r_1} ... \sum_{\lambda_q}^{r_q} p(\lambda_1, ..., \lambda_q) X_1^{\lambda_1} ... X_q^{\lambda_q} \tag{11.26}$$

definimos

$$||P|| = (\sum_{\lambda_1=0}^{r_1} ... \sum_{\lambda_q}^{r_q} |p(\lambda_1, ..., \lambda_q)|^2)^{1/2} \tag{11.27}$$

Tenemos entonces (por la relación de Parseval):

$$||P|| = (\int_{H_q} |P(e^{2i\pi y_1}, ..., e^{2i\pi y_q})|^2 \, dy_1 ... dy_q)^{1/2} \qquad H_q = \{(y_1, ..., y_q) \in \mathbb{R}^q, 0 \le y_j \le 1, 1 \le j \le q\} \tag{11.28}$$

Obtenemos inmediatamente

$$H(P) \le ||P|| \le H(P) \prod_{k=1}^{q} (1 + deg^{X_k} P)^{1/2} \tag{11.29}$$

y

$$||P|| \le \max_{|x_1|=1, ..., |x_q|=1} |P(x_1, ..., x_q)| \tag{11.30}$$

Referencias

[Baka] Alan Baker, *"linear forms in the logarithms of algebraic numbers. i*, Mathematika **13**, 204–216.

[Bakb] _____, *Linear forms in the logarithms of algebraic numbers. ii*, Mathematika **14**, 102–107.

[Bakc] _____, *Linear forms in the logarithms of algebraic numbers. iii*, Mathematika **14**, 220–228.

[Bak90] _____, *Transcendental number theory*, second ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1990. MR 1074572 (91f:11049)

[BBR90] F. Beukers, J.-P. Bézivin, and P. Robba, *An alternative proof of the Lindemann-Weierstrass theorem*, Amer. Math. Monthly **97** (1990), no. 3, 193–197. MR 1048429 (91d:11077)

[But08] Butler, *Transcendence and irrationality proofs*, 2008, http://www.maths.bris.ac.uk/~malab/PDFs/MA469.pdf.

[Fel81] N. Fel'dman, *Approximation of algebraic numbers*, Moskov. Gos. Univ., Moscow, 1981. MR 646065 (83e:10045)

[Fel00] _____, *Algebraic and transcendental numbers*, Quantum **10** (2000), no. 6, 22–26. MR 1763861 (2001c:11080)

[Fil10] Filaseta, *Algebraic number theory, math 784 class notes*, 2010, http://people.math.sc.edu/filaseta/gradcourses/TheMath784Notes.pdf.

[Fis07] Fischler, *Algebraic and transcendental numbers.*, 2007, http://www.math.u-psud.fr/~fischler/inde/inde_fischlerpondichery.pdf.

[FN98] N. I. Fel'dman and Yu. V. Nesterenko, *Transcendental numbers*, Number theory, IV, Encyclopaedia Math. Sci., vol. 44, Springer, Berlin, 1998, pp. 1–345. MR 1603608 (99a:11088b)

[FS67] N. Fel'dman and A. B. Shidlovskii, *The development and present state of the theory of transcendental numbers*, Uspehi Mat. Nauk **22** (1967), no. 3 (135), 3–81. MR 0214551 (35 #5400)

[Gel29] A. O. Gel'fond, *Sur les proprietes arithmetiques des fonctions entieres*, 280285.

[Gel34] _____, *Sur le septieme probleme de hilbert*, 623–630.

[Gel52] _____, *The approximation of algebraic numbers by algebraic numbers and the theory of transcendental numbers*, Amer. Math. Soc. Translation **1952** (1952), no. 65, 45. MR 0046392 (13,727b)

[Gel60] ———, *Transcendental and algebraic numbers*, Translated from the first Russian edition by Leo F. Boron, Dover Publications, Inc., New York, 1960. MR 0111736 (22 #2598)

[Hil42] Einar Hille, *Gelfond's solution of Hilbert's seventh problem*, Amer. Math. Monthly **49** (1942), 654–661. MR 0007770 (4,191b)

[Kuz29] R. O. Kuzmin, *On a new class of transcendental numbers*, I.A.N. 3 A930 (1929), 583–597.

[L37] Wantzel P. L, *Recherches sur les moyens de reconnaitre si un probleme de geometrie peut se resoudre avec la regle et compos*, J. Math. Pures Appl. **2** (1837), 366–372. 2

[Lan64] Serge Lang, *Algebraic numbers*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-Palo Alto-London, 1964. MR 0160763 (28 #3974)

[Lan66] ———, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966. MR 0214547 (35 #5397)

[Lan94] ———, *Algebraic number theory*, second ed., Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994. MR 1282723 (95f:11085)

[Lin82] F. Lindemann, *Ueber die Zahl π)*, Math. Ann. **20** (1882), no. 2, 213–225. MR 1510165

[Lio40] A. Liouville, *Sur la irrationalite du nombre e*, J. Math. Pures Appl. **5** (1840), no. 2, 192.

[Lio50] ———, *Sur des classes tres-etendues de quantites dont la valeur n'est ni algebrique, ni meme rductible a des irrationelles algebriques*, J. Math. pures appl. (1850), no. 15, 133–142.

[Nes06] Yu. V. Nesterenko, *Hilbert's seventh problem*, Mathematical events of the twentieth century, Springer, Berlin, 2006, pp. 269–282. MR 2182788 (2006j:11105)

[Per90] Veselin Perić, *Hilbert's seventh problem and transcendence of the numbers π and e*, Matematika **19** (1990), no. 1, 5–22. MR 1090250 (92d:11072)

[Roy92] Damien Roy, *Matrices whose coefficients are linear forms in logarithms*, J. Number Theory **41** (1992), no. 1, 22–47. MR 1161143 (93d:11077)

[Sch57] Theodor Schneider, *Einführung in die transzendenten Zahlen*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957. MR 0086842 (19,252f)

[Shi82] A. B. Shidlovskii, *iophantine approximation and transcendental numbers*, Moskov. Gos. Univ., Moscow, 1982. MR 687225 (84f:10042)

[Shi89a] ———, *Hilbert's seventh problem*, Proceedings of the Congress on Number Theory (Spanish) (Zarauz, 1984), Univ. País Vasco-Euskal Herriko Unib., Bilbao, 1989, pp. 176–180. MR 1203327 (93m:11062)

[Shi89b] ———, *Transcendental numbers*, de Gruyter Studies in Mathematics, vol. 12, Walter de Gruyter & Co., Berlin, 1989, Translated from the Russian by Neal Koblitz, With a foreword by W. Dale Brownawell. MR 1033015 (90j:11066)

[Sie67] Carl Ludwig Siegel, *Transcendental numbers*, Bibliographisches Institut, Mannheim, 1967. MR 0209231 (35 #133)

[Tij76] R. Tijdeman, *Hilbert's seventh problem: on the Gel′fond-Baker method and its applications*, Mathematical developments arising from Hilbert problems (Proc. Sympos. Pure Math., Northern Illinois Univ., De Kalb, Ill., 1974), Amer. Math. Soc., Providence, R. I., 1976, pp. 241–268. MR 0434974 (55 #7936)

[Wal74] Michel Waldschmidt, *Nombres transcendants*, Lecture Notes in Mathematics, Vol. 402, Springer-Verlag, Berlin-New York, 1974. MR 0360483 (50 #12931)

[Wal79] ———, *Transcendence methods*, Queen's Papers in Pure and Applied Mathematics, vol. 52, Queen's University, Kingston, Ont., 1979. MR 633068 (83a:10068)

[Wal88] ———, *On the transcendence methods of Gel′fond and Schneider in several variables*, New advances in transcendence theory (Durham, 1986), Cambridge Univ. Press, Cambridge, 1988, pp. 375–398. MR 972013 (90d:11089)