

**Universidad Autónoma
de Ciudad Juárez**

Instituto de Ingeniería y Tecnología
Departamento de Física y Matemáticas



**Ecuaciones diofantinas en la conjetura
de Collatz para resolver el problema de
los 1-ciclos**

Proyecto de Titulación que presenta:
Kevin Martin Cossio Lozano

Como requisito parcial para obtener el título de:
LICENCIATURA EN MATEMÁTICAS

Asesor:
Isaac Hasse Armengol

Cd. Juárez, Chihuahua,
Mayo de 2024

Declaración de Originalidad

Yo, Kevin Martin Cossio Lozano, declaro que el material contenido en este documento es original y no ha sido copiado de ninguna otra fuente, ni ha sido usado para obtener otro título o reconocimiento en otra institución de educación superior.

Kevin Martin Cossio Lozano

Título del Proyecto

Los miembros del comité evaluador que aprobaron el manuscrito del proyecto de titulación de:

Kevin Martin Cossio Lozano

Isaac Hasse Armengol
Asesor

Gustavo Tapia Sánchez
Sinodal

Francisco Ávila Álvarez
Sinodal

Luis Leobardo Alfaro Avena
Profesor de la materia

Resumen

En el presente escrito se presentan todos los fundamentos de Teoría de números, Álgebra moderna, Teoría de números algebraicos, Látices, Análisis matemático y Variable compleja que son necesarios para realizar un estudio introductorio a la Teoría de números trascendentales a partir de la exposición de los problemas típicos de la materia, desde Aproximación diofantina, Teorema de Lindemann, Teorema de Gelfond-Schneider, el Teorema de Baker, hasta un teorema importante para la solución de ecuaciones diofantinas. Para ello, se asumen conocimientos básicos de Álgebra lineal y Cálculo. Se realiza una introducción intuitiva a la conjetura de Collatz y se demuestran los resultados iniciales de la conjetura de Collatz hasta el problema de los 1-ciclos. Se expone el Teorema de Steiner para demostrar que no existen 1-ciclos utilizando la Teoría de números trascendentales.

Agradecimientos

Agradezco de todo corazón a mi familia por hacer esto posible, a mi asesor Isaac Hasse Armengol por darme la oportunidad de trabajar en el tema y a Gustavo Tapia Sánchez por contribuir ampliamente en mi formación como matemático.

Índice general

1	Introducción	3
1.1	Contexto	3
1.2	Antecedentes	4
1.3	Planteamiento del problema	4
1.4	Solución del problema	5
1.5	Objetivo general	5
1.6	Objetivos específicos	5
1.7	Alcances y delimitaciones	5
2	Fundamentos esenciales	7
2.1	Teoría de números	7
2.1.1	Divisibilidad	7
2.1.2	Congruencias	11
2.1.3	Fracciones continuas	12
2.2	Álgebra	21
2.2.1	Teoría de Grupos	21
2.2.2	Teoría de campos	22
2.3	Teoría de números algebraicos	32
2.4	Látices	38
2.5	Análisis matemático	45
2.6	Variable compleja	50
2.6.1	Las ecuaciones de Cauchy-Riemann	52
3	Metodología	57
4	Teoría trascendental de números	59
4.1	Aproximación diofantina	59
4.2	Teorema de Lindemann	62
4.3	Teorema de Gelfond-Schneider	70
4.4	Teorema de Baker	79

5	Ecuaciones diofantinas	91
5.1	Formas lineales en los logaritmos de números algebraicos . . .	91
6	Conjetura de Collatz	111
6.1	Dificultades de su estudio	112
6.2	Enfoques que han proporcionado mayor visión del problema .	113
6.3	Comportamiento de la función	113
6.4	Resultados relevantes para los ciclos	115
7	Teorema de Steiner	117
7.1	Ecuación diofantina de los 1-ciclos	117
7.2	Borrador de la demostración	118
7.3	Demostración	118

Glosario de siglas

Símbolo	Significado
$<, \leq$	Menor, menor o igual
$>, \geq$	Mayor, mayor o igual
$\forall$	Para todo
$\exists$	Existe
$\therefore$	Por lo tanto
f'	Primera derivada de f
$f^{(j)}$	j -ésima derivada de f
∂	Derivada parcial respecto a una variable
$ x $	Valor absoluto complejo de x
$\sum$	Sumatoria del estilo $a + b + \dots + z$
$\prod$	Producto del estilo $a \cdot b \cdots z$
$\rightarrow$	Denota el próximo valor de una sucesión
$\{a, b, \dots, z\}$	Conjunto de elementos
$\subset$	Subconjunto
$\in$	Pertenece
$\neq$	No es igual
$\emptyset$	Conjunto vacío
max, min	Máximo, mínimo
sup, inf	Supremo, ínfimo

Capítulo 1

Introducción

1.1 Contexto

En términos sencillos, una ecuación diofantina es una ecuación algebraica que tiene como soluciones números enteros. Un caso famoso son las ternas pitagóricas que resuelven la ecuación $x^2 + y^2 = z^2$. Estas ecuaciones pueden resultar increíblemente difíciles de resolver, pues se requieren técnicas pertinentes a las áreas de Teoría de números trascendental, Teoría de números algebraicos y Aproximación diofantina.

Por otro lado, la conjetura de Collatz o el problema $3x+1$ consiste en construir una serie que comienza con cualquier número entero n , si el número es par se divide entre dos, y si es impar lo multiplicamos por tres y le sumamos 1, así, el resultado es el nuevo valor de la sucesión. Por ejemplo:

$$5 \text{ (impar)} \rightarrow 16 \text{ (par)} \rightarrow 8 \text{ (par)} \rightarrow 4 \text{ (par)} \rightarrow 2 \text{ (par)} \rightarrow 1$$

Si la sucesión repite valores sin incluir al número 1, se dice que estos valores forman un ciclo no trivial. La conjetura de Collatz afirma que esta sucesión converge numéricamente a 1 y como consecuencia de ello, se afirma que no hay ciclos no triviales. Por lo cual, determinar que no hay ciclos no triviales es clave para la demostración de la conjetura. Por su parte, la búsqueda de ciclos está estrechamente relacionada con las soluciones de algunas ecuaciones diofantinas, por lo que serán de una herramienta clave para abordar la conjetura de Collatz.

1.2 Antecedentes

La conjetura fue compartida de manera verbal desde 1930 y no hubo documentos formales hasta la década de 1970, sin embargo, se le adjudica principalmente Collatz Lothar. Consiste en construir una serie que comienza con cualquier número n . Si el número es par se divide entre dos, y si es impar lo multiplicamos por tres y le sumamos 1. El resultado es el nuevo valor de la sucesión. La conjetura de Collatz afirma que la sucesión llegará a 1. Sorprendentemente, existen muchas áreas que conectan con estos problemas, desde Teoría de números, Sistemas dinámicos, Teoría de probabilidad, Teoría de la medida, entre otras áreas [15].

El primer avance notable en los ciclos comienza con Steiner [22] en 1977 al demostrar que no existen 1-ciclos no triviales resolviendo una ecuación diofantina utilizando técnicas de teoría de números trascendental, en particular, utilizando resultados del Teorema de Baker. En base al progreso de Steiner, Simons [20] logra demostrar que no existen 2-ciclos no triviales. Refinando esta demostración, Simons y de Weger [23] extienden este resultado hasta 68-ciclos. De esta manera, se ha continuado el método propuesto con Simons y aplicado distintas técnicas hasta 91-ciclos [13, 19]. Por supuesto, no hay evidencia de la existencia de un ciclo no trivial.

1.3 Planteamiento del problema

La principal problemática que se piensa resolver en la conjetura de Collatz, reside en la comprensión de los avances que se han hecho en el área, en particular, en la teoría utilizada para resolver el problema de los 1-ciclos. Esto se debe a que se requieren técnicas matemáticas avanzadas para comprender la demostración de Steiner y su acceso ha sido muy limitado. Aunque Simons [20] lo cita frecuentemente, y su resultado extiende el de Steiner, no se le ha dado el tratamiento adecuado al resultado de Steiner.

En primer lugar, es necesario exhibir la conexión entre la Teoría trascendental de números y las ecuaciones diofantinas para resolver el problema de los 1-ciclos en la conjetura de Collatz. En segundo lugar, se requiere motivar la teoría de manera intuitiva dado que la Teoría trascendental de números utiliza una maquinaria pesada, difícil de comprender, si no se introduce correctamente, lo que puede perder el interés de cualquier persona que quiera participar en la investigación de la conjetura de Collatz. Por último, existe la

necesidad de presentar enfoques didácticos que faciliten el estudio de esta conjetura con todo el contexto de sus herramientas matemáticas. En conjunto, resolver estas problemáticas generarán claridad en la conjetura de Collatz y permitirán que el lector pueda resolver problemas análogos.

1.4 Solución del problema

Se propone un estudio exhaustivo de los tópicos claves para comprender y replicar la demostración de Steiner de 1977. Entre los tópicos, se piensa estudiar a fondo teoremas de la teoría trascendental de números y su conexión con las ecuaciones diofantinas.

Objetivos

1.5 Objetivo general

- Estudiar y presentar las técnicas de Teoría trascendental de números y Teoría de números algebraicos para presentar el problema de los 1-ciclos en la conjetura de Collatz.

1.6 Objetivos específicos

- Presentar los problemas típicos de la Teoría trascendental de números por su orden e importancia histórica: Teorema de Lindemann, Gelfond-Schneider, Teorema de Baker, etc.
- Conectar las ideas de la teoría trascendental de números con las ecuaciones diofantinas con las ideas propuestas por Baker.
- Exponer desde un punto de vista intuitivo, los conocimientos necesarios hasta alcanzar el problema de las ecuaciones diofantinas en los 1-ciclos de la conjetura de Collatz.
- Replicar de manera didáctica la demostración de Steiner de 1977.

1.7 Alcances y delimitaciones

La investigación proporcionará claridad para comprender los avances actuales en los ciclos de la conjetura de Collatz, por supuesto, aunque no se tratará di-

rectamente los n -ciclos, para toda n , se presentarán todas las bases para entender el contexto con el que se trabajan.

No se pretende proporcionar un estudio exhaustivo del área de Teoría trascendental de números o Teoría de números algebraicos, sin embargo, habrá suficiente material para resolver de manera análoga algunas ecuaciones diofánticas el cual será el tema central de este trabajo.

Capítulo 2

Fundamentos esenciales

En este apartado, se tiene como objetivo presentar las definiciones y conceptos importantes para presentar de manera comprensible el capítulo de teoría trascendental de números. Por lo cual, los siguientes temas serán tratados de manera breve y no pretenden presentar un material exhaustivo de cada área, únicamente los resultados indispensables para este trabajo. Para un estudio más elaborado de estas áreas, se recomienda ampliamente los libros de Teoría de números [4], [16], [11], Álgebra [14], Análisis [1] y Variable compleja [8].

2.1 Teoría de números

Principio de Dirichlet: Sean m, n dos números naturales distintos de 0. Si se desean colocar n palomas en m cajas con $n > m$, alguna caja debe contener al menos $n + 1$ palomas.

Principio del buen orden: En cualquier conjunto de números naturales $A \subset \mathbb{N}$ distinto del conjunto vacío, $A \neq \emptyset$, existe un mínimo, es decir, un número $n \in A$ tal que $n \leq a$ para todo $a \in A$.

2.1.1 Divisibilidad

Si bien los conceptos de divisibilidad son comúnmente conocidos en el conjunto de los números enteros, en realidad, las siguientes propiedades son válidas en otros conjuntos. Aunque no se demostrará explícitamente, los siguientes resultados son válidos para polinomios con coeficientes en algún campo. Como referencia, recomendamos [14, Cap. 3, Sec. 9].

Teorema 2.1.1 (Algoritmo de la división). Dados $a, b \in \mathbb{Z}$ tal que $b > 0$, existen enteros únicos q y r que satisfacen

$$a = qb + r \quad 0 \leq r < b$$

Demostración. Primero, definamos el conjunto

$$S = \{a - xb \mid x \in \mathbb{Z}, \quad a - xb \geq 0\}$$

Veamos que el conjunto es no vacío. Como el entero $b \geq 1$, entonces $|a|b > |a|$ y así,

$$u = a - (-|a|)b = a + |a|b \geq a + |a| \geq 0$$

Es decir, $u \in S$. Por el Principio del buen orden, existe el entero más pequeño en S , sea ése entero r . Por definición de S , existe un entero q que satisface

$$r = a - qb \quad 0 \leq r$$

Supongamos que $r \geq b$, entonces $r - b \geq 0$, así,

$$a - (q + 1)b = (a - qb) - b = r - b \geq 0$$

Entonces $a - (q + 1)b \in S$ y además, $a - (q + 1)b = r - b < r$, lo cual contradice que r sea el entero más pequeño contenido en S . Esto implica que $r < b$. Entonces, existen enteros q y r que satisfacen

$$a = qb + r \quad 0 \leq r < b$$

Veamos que q y r son únicos. Supongamos que existe otra representación de la misma forma, es decir,

$$a = qb + r = q'b + r' \quad 0 \leq r < b \quad 0 \leq r' < b$$

Entonces $r' - r = b(q - q')$ y como $b > 0$

$$|r' - r| = |b(q - q')| = |b||q - q'| = b|q - q'|$$

Sumando las desigualdades $-b < -r \leq 0$ y $0 \leq r' < b$ obtenemos que $-b < r' - r < b$, de manera equivalente, $|r' - r| < b$. Así, $b|q - q'| < b$. Por lo cual,

$$0 \leq |q - q'| < 1$$

Entonces $|q - q'|$ es un entero no negativo menor a 1, por lo cual, la única posibilidad es que $|q - q'| = 0$. Por lo tanto, $q = q' \Rightarrow r = r'$, lo cual concluye la demostración. ■

Definición 2.1.1. Un entero b se dice que es divisible por un entero $a \neq 0$, en símbolos se representa como $a|b$, si existe un entero $c \neq 0$ tal que $b = ac$. Escribimos $a \nmid b$ para indicar que b no es divisible por a .

Las siguientes proposiciones son en realidad propiedades de la división, en particular, la más utilizada y de mayor interés es la proposición 2.1.5.

Proposición 2.1.1. Sea a un entero, entonces $a|0$, $1|a$, $a|a$.

Demostración: Claramente $0 = a \cdot 0$, $a = 1 \cdot a$ y $a = a \cdot 1$, por definición $a|0$, $1|a$ y $a|a$ respectivamente. ■

Proposición 2.1.2. Sean a, b, c, d enteros. Si $a|b$ y $c|d$, entonces $ac|bd$.

Demostración: Como $b = am$ para alguna m y $d = cn$ para alguna n , entonces multiplicando ambas expresiones $bd = amcn = (ac)mn$, por lo tanto, $ac|bd$. ■

Proposición 2.1.3. Sean a, b, c enteros. Si $a|b$ y $b|c$, entonces $a|c$.

Demostración: Como $b = am$ para alguna m y $c = bn$ para alguna n , entonces $c = (am)n = (a)mn$, por lo tanto, $a|c$. ■

Proposición 2.1.4. Sean a, b enteros. Si $a|b$ con $b \neq 0$, entonces $|a| \leq |b|$.

Demostración. Como $a|b$, entonces existe un entero $c \neq 0$ tal que $b = ac$. Por la Definición 2.1.1, $a \neq 0$. Claramente $|\frac{b}{a}| = |c| \geq 1$, por lo tanto, $|\frac{b}{a}| \geq 1$, así $|\frac{b}{a}| = \frac{|b|}{|a|} \geq 1$ finalmente $|b| \geq |a|$. ■

Proposición 2.1.5. Sean a, b, c enteros. Si $a|b$ y $a|c$ entonces a divide a cualquier combinación lineal de b y c , es decir, $a|bx + cy$ para cualesquiera enteros x, y .

Demostración. Como $b = am$ para alguna m , entonces $bx = amx$, entonces $a|bx$, análogamente $cy = any$, por lo tanto, $bx + cy = amx + any = a(mx + ny)$, como $mx + ny$ es un entero, entonces $a|bx + cy$. ■

Ahora, veremos un concepto que tiene mucha relevancia no solamente para los números enteros, si no también para los polinomios, aunque de nuevo, no lo demostraremos explícitamente.

Definición 2.1.2. Definimos el máximo común divisor entre dos enteros a, b como el entero d que satisface:

- $d|a, d|b$
- Si $c|a$ y $c|b$, entonces $c \leq d$.

Se denota $d = (a, b)$

Teorema 2.1.2. Sean enteros a, b , con al menos uno no nulo, entonces existen enteros r, s tales que el máximo común divisor se puede escribir como combinación lineal de a y b , es decir:

$$d = (a, b) = ar + bs$$

Demostración. Consideremos el conjunto S de todas las posibles combinaciones lineales de a y b :

$$S = \{au + bv : au + bv > 0 \quad u, v \in \mathbb{Z}\}$$

Claramente este conjunto es no vacío, pues si $u, v = 1$ entonces $a + b \in S$. Por el principio del buen orden, S debe contener un elemento mínimo d . Por lo tanto, por definición de S , $d = ar + bs$. Afirmamos que $d = (a, b)$. Por el algoritmo de la división, $a = dq + r$ con $0 \leq r < d$, entonces

$$r = a - dq = a - (ar + bs)q = a - arq - bsq = a(1 - rq) + b(-sq)$$

Por lo tanto, si r es un entero positivo entonces $r \in S$ con $r < d$, lo cual contradice la elección de d , así, $r = 0$. Con ello concluimos que $a = dq$, es decir, $d|a$, análogamente $d|b$. Supongamos que existe $c|a$ y $c|b$. Por la Proposición 2.1.5 $c|ax + by$ para cualesquiera enteros x, y , en particular, $c|d$. Por la Proposición 2.1.4, $c \leq d$. Por definición, d es el máximo común divisor entre a y b . Lo cual concluye la demostración. ■

Definición 2.1.3. Decimos que a y b son primos relativos si $(a, b) = 1$.

Teorema 2.1.3. Si a y b son primos relativos, entonces existen enteros r, s tales que el número 1 se puede escribir como combinación lineal de a y b , es decir:

$$1 = ar + bs$$

La demostración es una consecuencia directa del Teorema 2.1.2.

Lema 2.1.1 (Lema de Euclides). Si $a|bc$ tal que $(a, b) = 1$, entonces $a|c$.

Demostración. Como $(a, b) = 1$, entonces existen enteros r, s tales que $1 = ar + bs$. Multiplicando por c , $c = arc + bcs$, como $a|bc$, entonces existe algún entero q tal que $bc = aq$, por lo tanto, $c = arc + aqs = a(rc + qs)$, por lo tanto, $a|c$. ■

2.1.2 Congruencias

Definición 2.1.4. Decimos que $a \equiv b \pmod{m}$ si y solo si $m|a - b$.

En palabras, se lee a es congruente con b módulo m si y solo si m divide a $a - b$. Utilizar esta notación tiene muchos beneficios, en primer lugar, se puede demostrar que $\equiv$ es una relación de equivalencia, lo que permite simplificar muchos resultados en teoría de números. Veremos solamente algunas de sus propiedades.

Proposición 2.1.6. Sea $m > 1$ y a un entero. Entonces $a \equiv a \pmod{m}$.

Demostración: Claramente $m|a - a$, ya que $m|0$, por lo tanto, $a \equiv a \pmod{m}$. ■

Proposición 2.1.7. Sea $m > 1$, a y b enteros. Si $a \equiv b \pmod{m}$, entonces $b \equiv a \pmod{m}$.

Demostración: Como $a \equiv b \pmod{m}$, entonces $a - b = m \cdot q$, por lo tanto, $b - a = m \cdot (-q)$, por lo tanto, $m|b - a$. ■

Proposición 2.1.8. Sea $m > 1$, a, b, c enteros. Si $a \equiv b \pmod{m}$ y $b \equiv c \pmod{m}$, entonces $a \equiv c \pmod{m}$.

Demostración: Como $a - b = mq_1$ y $b - c = mq_2$, entonces sumando ambas ecuaciones $a - c = mq_1 + mq_2 = m(q_1 + q_2)$, por lo tanto, $m|a - c$. ■

Por supuesto, estas últimas tres proposiciones son suficientes para concluir que $\equiv$ es una relación de equivalencia. Las siguientes proposiciones son resultados difíciles de obtener sin la estructura que proporcionan las congruencias, es decir, no son evidentes.

Proposición 2.1.9. Sean a, b, c, d enteros y $m > 1$. Si $a \equiv b \pmod{m}$, $c \equiv d \pmod{m}$, entonces $a + c \equiv b + d \pmod{m}$ y $ac \equiv bd \pmod{m}$.

Demostración: Como $a - b = mq_1$ y $c - d = mq_2$, entonces sumando ambas expresiones $a + c - (b + d) = m(q_1 + q_2)$, por lo tanto, $m|(a + c) - (b + d)$. Por otro lado, $ac = (mq_1 + b)(mq_2 + d) = bd + mq_1mq_2 + bmq_2 + dmq_1$, entonces $ac - bd = m(q_1q_2m + b + d)$ con $q_1q_2m + b + d \in \mathbb{Z}$, por lo tanto, $m|ac - bd$. ■

Proposición 2.1.10. Sean a, b, c enteros y $m > 1$. Si $a \equiv b \pmod{m}$, entonces $a + c \equiv b + c \pmod{m}$ y $ac \equiv bc \pmod{m}$.

Demostración: Como $a - b = mq$, entonces $a + c - (b + c) = mq$, es decir, $m|(a + c) - (b + c)$. Además, $c(a - b) = mqc$, por lo tanto, $m|ac - bc$. ■

Proposición 2.1.11. Sean a, b enteros. Si $a \equiv b \pmod{m}$, entonces $a^k \equiv b^k \pmod{m}$ para cualquier entero positivo k .

Demostración: Por la Proposición 2.1.9 y usando $c = a$, $d = b$, entonces $a^2 \equiv b^2 \pmod{m}$ y se puede probar por inducción que $a^k \equiv b^k \pmod{m}$ para $k \in \mathbb{Z}$. ■

2.1.3 Fracciones continuas

Definición 2.1.5. Una fracción continua, es una expresión de la forma:

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}$$

Y se denota como $[a_0, a_1, a_2, \dots]$. A los números a_j se les conoce como los cocientes parciales.

Definición 2.1.6. Definimos a $[a_0, a_1, \dots, a_n]$ como el enésimo **convergente** de $[a_0, a_1, a_2, \dots]$. Es decir, es el resultado de la siguiente fracción continua:

$$\frac{p_n}{q_n} = [a_0, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_{n-1} + \frac{1}{a_n}}}}}$$

Donde $(p_n, q_n) = 1$. Y a los números $x_n = [a_n, a_{n+1}, \dots]$ se les conoce como cocientes completos.

Teorema 2.1.4. Sea $[a_0, a_1, \dots]$ una fracción continua. Entonces, afirmamos que:

- $p_n = a_n p_{n-1} + p_{n-2}$
- $q_n = a_n q_{n-1} + q_{n-2}$

Donde $p_{-1} = 1$ y $q_{-1} = 0$.

Demostración. Procedemos por inducción. Para $n = 1$, donde claramente $p_0 = a_0$ y $q_0 = 1$.

$$\begin{aligned} [a_0, a_1] &= a_0 + \frac{1}{a_1} \\ &= \frac{a_0 a_1 + 1}{a_1} \end{aligned}$$

Como se requería. Supongamos que el teorema es válido para n . Por lo tanto, observemos lo siguiente para utilizar la hipótesis de inducción:

$$\begin{aligned} [a_0, \dots, a_{n+1}] &= [a_0, \dots, a_{n-1}, a_n + \frac{1}{a_{n+1}}] \\ &= \frac{\left(a_n + \frac{1}{a_{n+1}}\right) p_{n-1} + p_{n-2}}{\left(a_n + \frac{1}{a_{n+1}}\right) q_{n-1} + q_{n-2}} \\ &= \frac{a_{n+1} a_n p_{n-1} + p_{n-1} + a_{n+1} p_{n-2}}{a_{n+1} a_n q_{n-1} + q_{n-1} + a_{n+1} q_{n-2}} \\ &= \frac{a_{n+1} (a_n p_{n-1} + p_{n-2}) + p_{n-1}}{a_{n+1} (a_n q_{n-1} + q_{n-2}) + q_{n-1}} \\ &= \frac{a_{n+1} p_n + p_{n-1}}{a_{n+1} q_{n-1} + q_{n-1}} \\ &= \frac{p_{n+1}}{q_{n+1}} \end{aligned}$$

Lo cual finaliza la inducción. ■

Lema 2.1.2. Las siguientes relaciones se cumplen:

1. $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$
2. $p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n$

Demostración. Veamos la primera parte, utilizando el Teorema 2.1.4:

$$\begin{aligned}
 p_n q_{n-1} - p_{n-1} q_n &= (a_n p_{n-1} + p_{n-2}) q_{n-1} - p_{n-1} (a_n q_{n-1} + q_{n-2}) \\
 &= p_{n-2} q_{n-1} + a_n p_{n-1} q_{n-1} - a_n p_{n-1} q_{n-1} - p_{n-1} q_{n-2} \\
 &= (-1)(p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) \\
 &\vdots \\
 &= (-1)^n (p_0 q_{-1} - p_{-1} q_0) \\
 &= (-1)^n (-1) \\
 &= (-1)^{n-1}
 \end{aligned}$$

Análogamente, y utilizando lo anterior,

$$\begin{aligned}
 p_n q_{n-2} - p_{n-2} q_n &= (a_n p_{n-1} + p_{n-2}) q_{n-2} - p_{n-2} (a_n q_{n-1} + q_{n-2}) \\
 &= p_{n-2} q_{n-2} + a_n p_{n-1} q_{n-2} - a_n p_{n-2} q_{n-1} - p_{n-2} q_{n-2} \\
 &= a_n (p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) \\
 &= (-1)^n a_n
 \end{aligned}$$

■

Definición 2.1.7. Sea α un número real. Para $m = 0, 1, 2, \dots$ definimos el siguiente algoritmo recursivo como sigue:

- $\alpha_0 = \alpha$
- $a_m = \lfloor \alpha_m \rfloor$ donde $\lfloor x \rfloor$ representa la función piso.
- $\alpha_m = a_m + \frac{1}{\alpha_m + 1}$

Observemos que dada esta definición, para un entero positivo m , se cumple que $\alpha = [a_0, a_1, \dots, a_{m-1}, \alpha_m]$ (se sigue de expandir $\alpha = \alpha_0$). Por lo cual, se justifica que se pueda aproximar cualquier número real a partir de las fracciones continuas.

Teorema 2.1.5. Sea α un número irracional y $\frac{p_n}{q_n}$ sea un convergente de su fracción continua. Por lo tanto,

$$\left| \alpha - \frac{p_n}{q_n} \right| = \frac{1}{q_n (\alpha_{n+1} q_n + q_{n-1})}$$

Demostración. Sea n un entero positivo. Entonces, $\alpha = [a_0, a_1, \dots, a_n, \alpha_{n+1}]$. Entonces, utilizando el Teorema 2.1.4 y el Lema 2.1.2,

$$\begin{aligned}
 \alpha - \frac{p_n}{q_n} &= \frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} \\
 &= \frac{p_{n-1}q_n + \alpha_{n+1}p_nq_n - \alpha_{n+1}p_nq_n - p_nq_{n-1}}{q_n(\alpha_{n+1}q_n + q_{n-1})} \\
 &= \frac{p_{n-1}q_n - p_nq_{n-1}}{q_n(\alpha_{n+1}q_n + q_{n-1})} \\
 &= \frac{(-1)(p_nq_{n-1} - q_n p_{n-1})}{q_n(\alpha_{n+1}q_n + q_{n-1})} \\
 &= \frac{(-1)^n}{q_n(\alpha_{n+1}q_n + q_{n-1})}
 \end{aligned}$$

■

Corolario 2.1.1. Sea α un número real con $\frac{p_n}{q_n}$ convergente en su fracción continua. Entonces,

$$|q_n\alpha - p_n| < |q_{n-1}\alpha - p_{n-1}|$$

Demostración. Por el Teorema 2.1.5, tenemos que:

$$|q_n\alpha - p_n| = \frac{1}{q_n\alpha_{n+1} + q_{n-1}}$$

Análogamente,

$$|q_{n-1}\alpha - p_{n-1}| = \frac{1}{q_{n-1}\alpha_n + q_{n-2}}$$

Por lo cual, es suficiente probar la siguiente desigualdad,

$$\frac{1}{q_n\alpha_{n+1} + q_{n-1}} < \frac{1}{q_{n-1}\alpha_n + q_{n-2}} \quad (2.1)$$

Notemos que:

$$\begin{aligned}
 \frac{1}{q_{n-1}\alpha_n + q_{n-2}} &= \frac{1}{q_{n-1}(a_n + \frac{1}{\alpha_{n+1}}) + q_{n-2}} \\
 &= \frac{1}{q_{n-1}a_n + \frac{q_{n-1}}{\alpha_{n+1}} + q_{n-2}} \\
 &= \frac{\alpha_{n+1}}{q_{n-1}a_n\alpha_{n+1} + q_{n-1} + q_{n-2}\alpha_{n+1}}
 \end{aligned}$$

Por lo tanto, se cumple la ecuación (2.1) si y solo si,

$$\begin{aligned} q_{n-1}a_n\alpha_{n+1} + q_{n-1} + \alpha_{n+1}q_{n-2} &< q_n\alpha_{n+1}^2 + q_{n-1}\alpha_{n+1} \iff \\ \alpha_{n+1}(q_{n-1}a_n + q_{n-2}) + q_{n-1} &< q_n\alpha_{n+1}^2 + q_{n-1}\alpha_{n+1} \iff \\ q_n\alpha_{n+1} + q_{n-1} &< \alpha_{n+1}(q_n\alpha_{n+1} + q_{n-1}) \iff \\ 1 &< \alpha_{n+1} \end{aligned}$$

Notemos que $\frac{1}{\alpha_{n+1}} = \alpha_n - [\alpha_n]$, por lo cual, $\frac{1}{\alpha_{n+1}} < 1$, por lo tanto, $\alpha_{n+1} > 1$, lo cual concluye la demostración. ■

Corolario 2.1.2. Sea α un número real, $\frac{p_n}{q_n}$ un convergente en la fracción continua de α , entonces,

$$\frac{1}{(a_{n+1} + 2)q_n^2} < \left| \alpha - \frac{p_n}{q_n} \right|$$

Demostración. Argumentamos que

$$a_{n+1} + 2 - \alpha_{n+1} > \frac{q_{n-1}}{q_n}$$

Primero, notemos que $q_n > q_{n-1}$, pues de hecho, $q_n = a_n q_{n-1} + q_{n-2}$, donde todos son enteros positivos. Por lo cual, es suficiente ver que $a_{n+1} + 2 - \alpha_{n+1} > 1$, lo cual es obvio, ya que $\alpha_{n+1} = a_{n+1} + \frac{1}{\alpha_{n+1} + 1} < a_{n+1} + 1$. Por lo tanto,

$$\begin{aligned} q_n(a_{n+1} + 2 - \alpha_{n+1}) &> q_{n-1} \\ (a_{n+1} + 2)q_n - \alpha_{n+1}q_n &> q_{n-1} \\ (a_{n+1} + 2)q_n &> \alpha_{n+1}q_n + q_{n-1} \\ (a_{n+1} + 2)q_n^2 &> q_n(\alpha_{n+1}q_n + q_{n-1}) \\ \frac{1}{q_n(\alpha_{n+1}q_n + q_{n-1})} &> \frac{1}{(a_{n+1} + 2)q_n^2} \end{aligned}$$

Utilizando el Teorema 2.1.5, concluimos que,

$$\left| \alpha - \frac{p_n}{q_n} \right| > \frac{1}{(a_{n+1} + 2)q_n^2}$$
■

De hecho, estamos más interesado en este corolario que en el teorema, pues, con él es posible obtener una cota inferior para a_{n+1} .

Teorema 2.1.6 (Ley de la mejor aproximación). Sea α un número real y $\frac{p_n}{q_n}$ un convergente de su fracción continua y $n \geq 2$. Si p, q son enteros tales que $0 < q \leq q_n$ y $\frac{p}{q} \neq \frac{p_n}{q_n}$, entonces,

$$|q_n \alpha - p_n| < |q \alpha - p|$$

Además, una fracción reducida $\frac{p'}{q'}$ con $q' \geq q_2$ que satisface la desigualdad es convergente.

Demostración. Por el Corolario 2.1.1 hemos probado el caso cuando $\frac{p}{q}$ es convergente en la fracción continua de α .

Caso 1. Supongamos que $q = q_n$, entonces $p \neq p_n$ (por la hipótesis $\frac{p}{q} \neq \frac{p_n}{q_n}$), así,

$$\begin{aligned} \left| \frac{p}{q} - \frac{p_n}{q_n} \right| &= \frac{|p - p_n|}{q_n} \\ &\geq \frac{1}{q_n} \end{aligned}$$

A su vez, notemos que para cada n , $a_n \leq \alpha_n$ (por la propiedad de la función piso). Utilizando el Teorema 2.1.5,

$$\begin{aligned} \left| \alpha - \frac{p_n}{q_n} \right| &= \frac{1}{q_n(\alpha_{n+1}q_n + q_{n-1})} \\ &< \frac{1}{q_n(a_{n+1}q_n + q_{n-1})} \\ &= \frac{1}{q_n q_{n+1}} \\ &< \frac{1}{2q_n} \end{aligned}$$

La última desigualdad porque $q_{n+1} > 3$ cuando $n \geq 2$ (ya que q_n es estrictamente creciente para $n > 1$). Por la desigualdad del triángulo,

$$\begin{aligned} \left| \alpha - \frac{p}{q} \right| &\geq \left| \frac{p}{q} - \frac{p_n}{q_n} \right| - \left| \alpha - \frac{p_n}{q_n} \right| \\ &> \frac{1}{q_n} - \frac{1}{2q_n} \\ &= \frac{1}{2q_n} \\ &> \left| \alpha - \frac{p_n}{q_n} \right| \end{aligned}$$

Por lo tanto, como $q = q_n$, obtenemos:

$$|q\alpha - p| \geq |q_n\alpha - p_n|$$

Como se requería.

Caso 2. Supongamos que $0 < q < q_n$. Consideremos el siguiente sistema de ecuaciones

$$p_n X + p_{n-1} Y = p \quad (2.2)$$

$$q_n X + q_{n-1} Y = q \quad (2.3)$$

La cual induce la siguiente solución (x, y) única:

$$x = \frac{pq_{n-1} - qp_{n-1}}{p_n q_{n-1} - p_{n-1} q_n}$$

$$y = \frac{pq_n - qp_n}{p_n q_{n-1} - p_{n-1} q_n}$$

Por el Lema 2.1.2, los denominadores son ± 1 . Por lo tanto, x, y son enteros, y de hecho podemos considerar que son no nulos, ya que si

$$pq_{n-1} - qp_{n-1} = 0 \Rightarrow \frac{p}{q} = \frac{p_{n-1}}{q_{n-1}}$$

Para lo cual no habría nada que demostrar, por ser $\frac{p}{q}$ convergente en la fracción continua. Ahora, por la ecuación (2.3) como $q_n > q$, necesariamente x, y tienen signos opuestos. Por el Teorema 2.1.5, $\alpha - \frac{p_n}{q_n}$ alterna signos, es decir, $\alpha - \frac{p_n}{q_n}$ y $\alpha - \frac{p_{n-1}}{q_{n-1}}$ tienen signos opuestos. Esto implica que $x(q_n\alpha - p_n)$ y

$y(q_{n-1}\alpha - p_{n-1})$ tienen signos iguales. Por lo cual,

$$\begin{aligned} q\alpha - p &= (q_n x + q_{n-1} y)\alpha - (p_n x + p_{n-1} y) \\ &= x(q_n \alpha - p_n) + y(q_{n-1} \alpha - p_{n-1}) \\ |q\alpha - p| &= |x(q_n \alpha - p_n)| + |y(q_{n-1} \alpha - p_{n-1})| \\ &> |q_{n-1} \alpha - p_{n-1}| \\ &> |q_n \alpha - p_n| \end{aligned}$$

La última desigualdad por el Corolario 2.1.1. Como se requería. Además, como observación, esta última desigualdad se cumple para $q_{n-1} < q < q_n$, lo cual, por el principio de inducción, implica que solo los convergentes satisfacen tal desigualdad. Esto concluye la demostración. ■

Finalmente, tenemos todos los teoremas que necesitamos para determinar un criterio que nos permita encontrar fracciones que sean convergentes en la fracción continua de un número real.

Teorema 2.1.7. Sea α un número real, una fracción reducida de la forma $\frac{p}{q}$, tal que se satisface:

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

Entonces $\frac{p}{q}$ es un convergente de la fracción continua de α .

Demostración. Supongamos que $\frac{p}{q}$ satisface tal desigualdad. Por el Ley de la mejor aproximación, es suficiente demostrar que $\frac{p}{q}$ es la mejor aproximación de α .

Sea $\frac{P}{Q}$ tales que $\frac{P}{Q} \neq \frac{p}{q}$ y $|Q\alpha - P| \leq |q\alpha - p| = q \left| \alpha - \frac{p}{q} \right| < \frac{1}{2q}$. Entonces,

$$\begin{aligned} \frac{1}{qQ} &\leq \frac{|pQ - Pq|}{qQ} \\ &= \left| \frac{p}{q} - \frac{P}{Q} \right| \\ &\leq \left| \alpha - \frac{p}{q} \right| + \left| \frac{P}{Q} - \alpha \right| \\ &< \frac{1}{2q^2} + \frac{1}{2qQ} \\ &= \frac{q + Q}{2q^2Q} \end{aligned}$$

Por lo tanto,

$$1 < \frac{q + Q}{2q}$$
$$q < Q$$

Lo cual completa la demostración.



2.2 Álgebra

Si bien no se incluyen todos los resultados pertinentes de cada una de las áreas de Álgebra, nuestra misión es tener el mínimo conocimiento indispensable para entender los elementos de la Teoría de Galois, la cual se intenta abordar en la Teoría de campos. Como veremos, esta teoría tiene muchas aplicaciones en Teoría de números algebraicos, Látices, Teoría de números trascendentales y seguirá siendo una constante durante todo el trabajo. Se espera que con esto el lector tenga una noción exacta de lo que se puede lograr con esta teoría, es decir, nos importa mucho más sus aplicaciones y entender bajo qué condiciones se puede aplicar.

2.2.1 Teoría de Grupos

Definición 2.2.1. Una operación binaria definida en un conjunto G es una función $G \times G \rightarrow G$ cerrada que asocia un par de elementos de G a un único elemento en G .

Definición 2.2.2. Un conjunto no vacío de elementos G se dice que forma un grupo si en G existe una operación binaria definida por un producto $\cdot$, tal que para todo $a, b, c \in G$:

- (a). El producto es asociativo, es decir, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
- (b). Existe un elemento neutro $e \in G$ tal que $a \cdot e = e \cdot a = a$.
- (c). Existe un elemento $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$.

Por ejemplo, si n es un entero, podemos considerar el conjunto de todas las matrices invertibles de $n \times n$ con entradas en los reales. Por construcción, ése conjunto forma un grupo con la operación usual de matrices.

Definición 2.2.3. Decimos que $H \subset G$ es un subgrupo de G si es un grupo bajo la misma operación de G .

Es claro que $\mathbb{Z}$ es un grupo con la operación de suma usual. Es posible demostrar que los números pares forman un subgrupo de $\mathbb{Z}$.

Definición 2.2.4. Si G es un conjunto que contiene n elementos, definimos el orden de G como el cardinal de G , es decir,

$$o(G) = n$$

Proposición 2.2.1. Sea G un grupo y $g \in G$. Definimos $gG = \{gh : h \in G\}$. Afirmamos que $gG = G$.

Demostración. Sea $x \in gG$, entonces $x = gh$ para algún $h \in G$. Como la operación binaria es cerrada, entonces $gh \in G$, por lo tanto, $x \in G$, es decir, $gG \subset G$. Ahora, sea $y \in G$, como $g \in G$, existe $g^{-1} \in G$ tal que $g \cdot g^{-1} = e$. En particular, $g^{-1}y \in G$ porque la operación binaria es cerrada. Por lo tanto, $g \cdot (g^{-1} \cdot y) \in gG$, notemos que, $g \cdot (g^{-1} \cdot y) = (g \cdot g^{-1}) \cdot y = e \cdot y = y$ (por la propiedad (a) de la definición de grupo). Por lo tanto, $y \in gG$, es decir, $G \subset gG$. Lo cual demuestra que $gG = G$ como se quería. ■

Aunque este resultado podrá aparentar no ser importante, es realmente clave para entender algunos conceptos en la Teoría de Galois.

2.2.2 Teoría de campos

Definición 2.2.5. Decimos que un conjunto K es un campo si tiene dos operaciones binarias en K las cuales se conocen como *suma* y *multiplicación* con las que se cumplen las siguientes propiedades para todo $a, b, c \in K$:

(a). Asociatividad en la suma y multiplicación:

$$a + (b + c) = (a + b) + c \quad \text{y} \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

(b). Conmutatividad en la suma y multiplicación:

$$a + b = b + a \quad \text{y} \quad a \cdot b = b \cdot a$$

(c). Distributividad de la multiplicación sobre la suma:

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

(d). Identidad en la suma y multiplicación:

$$a + 0 = a \quad \text{y} \quad a \cdot 1 = a$$

(e). Inversos en la suma y multiplicación (exceptuando $a = 0$):

$$a + (-a) = 0 \quad \text{y} \quad a \cdot (a^{-1}) = 1$$

Claramente, bajo esta definición K es un espacio vectorial sobre F , lo que nos permite utilizar las herramientas de Álgebra lineal.

Un ejemplo de un campo son los números racionales bajo las operaciones de suma y multiplicación usual. Por supuesto, veremos otros ejemplos más adelante.

Definición 2.2.6. Decimos que F es un subcampo de K si es un campo bajo las mismas operaciones de K y está contenido en K . De manera equivalente, se dice que K es una extensión de F .

Es claro que K es un subcampo de sí mismo, si F es un subcampo de K distinto de K , entonces se dice que es un subcampo propio. A lo largo de esta sección, asumiremos que K siempre es una extensión de algún subcampo F .

Definición 2.2.7. El grado de K sobre F es la dimensión de K como espacio vectorial sobre F . Se denota como $[K : F]$. Además, decimos que K es una extensión finita de F si $[K : F]$ es finito.

La demostración del siguiente teorema puede ser encontrada en [14, pp. 208]. La omitimos por ser una demostración propia de Álgebra lineal.

Teorema 2.2.1. Sea L una extensión finita de K y K una extensión finita de F , entonces L es una extensión finita de F y además:

$$[L : F] = [L : K] \cdot [K : F]$$

Ahora, bien, notemos que dado un campo K , $1 \in K$. Ya que para cualquier $a \in K \Rightarrow a \cdot a^{-1} \in K$, entonces $1 \in K$. Además, $1 + 1 \in K \Rightarrow 2 \in K$. A su vez, el inverso aditivo de 2 está en K , es decir, $-2 \in K$, utilizando inducción, podemos construir una copia de $\mathbb{Z}$ contenida en K . Lo cual, motiva la siguiente definición.

Definición 2.2.8. Sea K un campo y $m \in K$, $a \in K$ un entero. Se dice que K tiene característica 0 si $ma \neq 0$ para $a \neq 0 \in K$ y $m > 0$ un entero. En caso contrario, si $ma = 0$ para alguna $m > 0$ y algún $a \neq 0 \in K$ entonces K se dice que tiene característica finita. La característica de K en este caso es el menor entero positivo p tal que $pa = 0$ para todo $a \in K$.

Ejemplo: Se puede demostrar que $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ con las operaciones de suma y producto mod 5 es un campo. Notemos que $5 \cdot x \equiv 0 \pmod{5}$, ya que $5|5x$. Por lo tanto, $5(x) \equiv 0$ para todo $x \in \mathbb{Z}_5$. Por lo tanto, $\mathbb{Z}_5$ tiene característica 5.

Definición 2.2.9. Decimos que $a \in K$ es algebraico sobre F si existe un polinomio no nulo $p(x) \in F[x]$, es decir, un polinomio donde sus coeficientes están en F , tal que $p(a) = 0$.

Definición 2.2.10. Sea M una colección de todos los subcampos de K que contienen a F y un elemento $a \in K$. Claramente M no es vacío, pues K es un elemento de M . Definimos la intersección de todos los subcampos de K que son miembros de M como $F(a)$.

Teorema 2.2.2. $F(a)$ es el mínimo subcampo de K que contiene a F y a .

Demostración: Primero, veamos que la intersección de subcampos de K es un subcampo de K . Claramente $0, 1$ están en la intersección porque están en todos los subcampos de K en M . Dados dos elementos a, b en la intersección, entonces a, b están en todos los subcampos de K en M , por lo cual, $a \pm b, ab$ están en cada uno de los subcampos y por tanto están en la intersección. Por el mismo argumento los inversos multiplicativos y aditivos están en la intersección. Las demás propiedades se heredan de K . Por lo tanto, $F(a)$ es un subcampo de K . Supongamos que hay otro subcampo que contiene a F y a , entonces están en M y claramente por construcción, $F(a)$ está contenido en él (por ser la intersección de este y otros subcampos de K). Lo que completa el teorema. ■

Observación. La definición anterior se puede generalizar si consideramos $a, b \in K$ y $F(a)(b) = F(a, b)$ como la intersección de todos los subcampos de K , que contienen a $F(a)$ y b . Es decir, es el mínimo subcampo que contiene a $F(a)$ y b , por lo tanto, es el mínimo subcampo que contiene a F , a y b . Análogamente para más variables.

Definición 2.2.11. La extensión K de F es una extensión simple de F si $K = F(\alpha)$ para algún $\alpha \in K$.

El siguiente teorema, forma parte de una discusión del libro [14, pp. 210–212]. Por su extensión, no incluiremos su demostración.

Teorema 2.2.3. Sea $a \in K$ algebraico de grado n sobre F si y solo si $[F(a) : F] = n$ y además, una base linealmente independiente sobre F para $F(a)$ es $\{1, a, a^2, \dots, a^{n-1}\}$

Lema 2.2.1. El polinomio $f(x) \in F[x]$ tiene una raíz múltiple, es decir, repetida, si y solo si $f(x)$ y $f'(x)$ tienen un factor común no trivial (es decir, de grado positivo).

Demostración. Supongamos que $f(x)$ y $g(x)$ en $F[x]$ tienen un factor común no trivial en $K[x]$, para alguna extensión K sobre F . Entonces $f(x)$ y $g(x)$ tienen un factor no trivial en $F[x]$. De lo contrario, si fueran primos relativos como elementos en $F[x]$, podríamos encontrar polinomios $a(x), b(x) \in F[x]$

tal que $a(x)f(x) + b(x)g(x) = 1$ por el Teorema 2.1.3. Entonces, el máximo común divisor entre $f(x)$ y $g(x)$ es 1 (incluso como polinomios en $K[x]$), por lo tanto, en $K[x]$ serían primos relativos, lo cual es una contradicción. (Nótese que este argumento asume que se comprende la Teoría de números para funciones definidas en un campo).

Por lo tanto, sin pérdida de generalidad, podemos asumir que todas las raíces de $f(x)$ están en F , de lo contrario, extendemos F a K , el campo de descomposición de F , si podemos ver que $f(x)$ y $g(x) = f'(x)$ tiene un factor no trivial en común en $K[x]$, entonces este factor está en $F[x]$. Supongamos que $f(x)$ tiene una raíz múltiple α , entonces $f(x) = (x - \alpha)^m q(x)$ con $m > 1$. Además, $f'(x) = (x - \alpha)^m q'(x) + m(x - \alpha)^{m-1}$. Pero esto implica que $f(x)$ y $f'(x)$ tienen el factor común $x - \alpha$. Lo cual demuestra la primera dirección del lema.

Ahora, supongamos que $f(x)$ y $f'(x)$ tienen un factor común no trivial. Y supongamos que $f(x)$ no tienen ninguna raíz múltiple. Entonces $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$ donde cada α_i son todos distintos. Entonces:

$$f'(x) = \sum_{i=1}^n (x - \alpha_1) \cdots \overline{(x - \alpha_i)} \cdots (x - \alpha_n)$$

Donde $\overline{(x - \alpha_i)}$ en este caso específico significa que el término es omitido. Afirmamos que ninguna raíz de $f(x)$ es una raíz de $f'(x)$, notemos que

$$f'(a_i) = \prod_{j \neq i} (a_i - a_j) \neq 0$$

Ya que todas las raíces son distintas. Lo cual contradice la hipótesis. Por lo tanto, $f(x)$ tiene una raíz múltiple. ■

Teorema 2.2.4. Sea F un campo de característica 0 y sea $f(x) \in F[x]$ un polinomio irreducible. Entonces $f(x)$ no tiene raíces repetidas.

Demostración. Si $f(x)$ tiene una raíz repetida, entonces $f(x)$ y $f'(x)$ tienen un factor repetido por el lema anterior, entonces $f(x) | f'(x)$. Sin embargo, como $\deg f'(x) < \deg f(x)$, entonces $f'(x) = 0$. En un campo de característica 0, esto implica que f es una función constante. ■

Observación: En un campo de característica p se tiene que $f(x) = x^p \Rightarrow f'(x) = px^{p-1} = 0$. Por lo cual, la conclusión sería que $f(x) = g(x^p)$, así, es una condición necesaria que el campo sea de característica 0. En parte por ello y por simplicidad, en este trabajo solamente necesitaremos campos de característica 0.

Teorema 2.2.5. Sea F un campo de característica 0 y a, b algebraicos sobre F . Entonces existe un elemento $c \in F(a, b)$ tal que $F(a, b) = F(c)$.

Demostración. Sean $f(x)$ y $g(x)$ los polinomios irreducibles de grado m y n sobre F tal que $f(a) = g(b) = 0$. Sea K una extensión de F donde $f(x)$ y $g(x)$ se factoricen completamente. Como la característica de F es 0, entonces todas las raíces de $f(x)$ son distintas, de la misma manera para $g(x)$ por el teorema anterior. Sean las raíces de $f(x)$ $a = a_1, a_2, \dots, a_m$ y las de $g(x)$ $b = b_1, b_2, \dots, b_n$. Si $j \neq 1$, entonces $b_j \neq b_1 = b$, por lo tanto la ecuación $a_i + \lambda b_j = a_1 + \lambda b_1 = a + \lambda b$ solamente tiene una solución $\lambda \in K$, la cual es,

$$\lambda = \frac{a_i - a}{b - b_j}$$

Como F es de característica 0, tiene un número infinito de elementos, por lo cual, podemos encontrar un elemento $\gamma \in F$ tal que $a_i + \gamma b_j \neq a + \gamma b$ para todo i y para todo $j \neq 1$, ya que por lo anterior solo tienen un número finito de soluciones. Sea $c = a + \gamma b$, queremos demostrar que $F(c) = F(a, b)$. Como $c \in F(a, b)$, claramente $F(c) \subset F(a, b)$. Ahora mostraremos que $a, b \in F(c)$, lo que concluye $F(a, b) \subset F(c)$.

Definamos $h(x) = f(c - \gamma x)$, entonces $h(x) \in F(c)[x]$ y $h(b) = f(c - \gamma b) = f(a) = 0$, ya que $a = c - \gamma b$. Así, en alguna extensión de $F[c]$, $h(x)$ y $g(x)$ tienen $x - b$ como factor común. Por el primer argumento del Lema 2.2.1, $x - b \in F[c][x]$, por lo tanto, $b \in F[c]$. Como $a = c - \gamma b$, obtenemos que $a \in F[c]$. Lo cual concluye la demostración. ■

En realidad, hemos demostrado más, como vemos en el siguiente teorema.

Teorema 2.2.6. Cualquier extensión finita de un campo de característica 0 es una extensión simple.

Demostración. Por ser una extensión finita, $K = F(a, b, \dots, z)$, aplicando el teorema anterior una cantidad finita de veces por un argumento inductivo se tiene que $K = F(c)$. ■

Este resultado es útil para algunos de los teoremas que veremos, pues es más sencillo trabajar bajo la suposición de que un campo es una extensión simple.

Definición 2.2.12. Si $f(x) \in F[x]$ entonces una extensión finita K sobre F se conoce como el campo de descomposición de F para $f(x)$ si $f(x)$ se puede factorizar como un producto de factores lineales pero no sobre un subcampo propio de K .

Nótese que si K es el campo de descomposición para algún polinomio $f(x) \in F[x]$, entonces K es el campo de descomposición para cualquier L tal que $F \subset L \subset K$. Este argumento en ocasiones es útil, pues $f(x) \in L[x]$.

Definición 2.2.13. Un automorfismo de un campo K es una función biyectiva $\sigma : K \rightarrow K$ tal que $\forall a, b \in K$:

- $\sigma(a + b) = \sigma(a) + \sigma(b)$
- $\sigma(ab) = \sigma(a)\sigma(b)$

Decimos que dos automorfismos σ y τ de K son distintos si $\sigma(a) \neq \tau(a)$ para algún $a \in K$. Y cotidianamente, decimos que los automorfismos abren sumas y productos.

Teorema 2.2.7. El conjunto de todos los automorfismos de K forman un grupo bajo la composición.

Demostración. Primero, sean σ, τ dos automorfismos, notemos lo siguiente $\sigma(\tau(K)) = \sigma(K) = K$ por ser automorfismos, por lo tanto, $\sigma\tau$ es un automorfismo, es decir, la operación es cerrada. Claramente es una operación asociativa porque las funciones lo son. Existe un automorfismo neutro, el cual se define como $1_K : K \rightarrow K$ tal que $1_K(x) = x$ para todo $x \in K$, ya que $\sigma(1_K(x)) = \sigma(x) = 1_K(\sigma(x))$. Por último, dado que todo automorfismo es biyectivo entonces todos tienen un automorfismo inverso. Por lo tanto, el conjunto de todos los automorfismos forma un grupo. ■

Definición 2.2.14. Sea G un grupo de automorfismos de K , entonces el campo fijo de G es el conjunto de todos los elementos $\alpha \in K$ tales que $\sigma(\alpha) = \alpha$ para todo $\sigma \in G$. Se denota K_G .

De manera coloquial, decimos que los elementos de K_G son invariantes bajo cualquier automorfismo de G .

Teorema 2.2.8. El campo fijo de G es un subcampo de K .

Demostración. Primero, es no nulo ya que $\sigma(0) = \sigma(0+0) = \sigma(0)+\sigma(0) \Rightarrow \sigma(0) = 0$ y además si $\sigma(1) = 0$, entonces $\sigma(a) = \sigma(1 \cdot a) = \sigma(1)\sigma(a) = 0$, para todo $a \in K$ (lo cual contradice que sea una función biyectiva), por lo tanto, $\sigma(1) \neq 0$, así, $\sigma(1) = \sigma(1 \cdot 1) = \sigma(1) \cdot \sigma(1)$, como hay inversos para todo elemento de K exceptuando el 0, entonces $\sigma(1) \cdot \sigma(1)^{-1} = \sigma(1) \cdot \sigma(1) \cdot \sigma(1)^{-1}$, por lo tanto, $1 = \sigma(1)$. Es decir, 0 y 1 están en el campo fijo de G . Ahora, sean a, b en el campo fijo, entonces dado $\sigma \in G$, $\sigma(a \pm b) = \sigma(a) \pm \sigma(b) = a \pm b$, análogamente para el producto. Entonces $a \pm b, ab$ en el campo fijo. Si $b \neq 0$,

entonces $\sigma(b^{-1}) \cdot \sigma(b) = \sigma(1) = 1$, por lo tanto, $\sigma(b^{-1}) = \sigma(b)^{-1} = b^{-1}$, por lo tanto, b^{-1} está en el campo fijo. Análogamente para la suma, y las demás propiedades se heredan de K . Por lo tanto, el campo fijo es un subcampo de K . ■

En particular, el campo fijo de G , K_G , es un campo.

Definición 2.2.15. Sea K una extensión de F . Entonces el grupo de automorfismos de K relativo a F , denotado como $G(K, F)$, es el conjunto de todos los automorfismos de K tal que cada elemento de F queda fijo, es decir, un automorfismo de K está en $G(K, F)$ si y solo si $\sigma(\alpha) = \alpha$ para todo $\alpha \in F$.

Teorema 2.2.9. $G(K, F)$ es un subgrupo del grupo de todos los automorfismos de K .

Demostración. Claramente $1_K \in G(K, F)$, ya que por definición deja fijo a todos los elementos de K , en particular a los de F . La asociatividad se hereda de las funciones. Dado $\sigma, \tau \in G(K, F)$ y $x \in F$, entonces $\sigma(\tau(x)) = \sigma(x) = x$, por lo tanto, $\sigma\tau \in G(K, F)$. Por último, dado $\sigma \in G(K, F)$, veamos que $\sigma^{-1} \in G(K, F)$. Sea $x \in F$, entonces $\sigma\sigma^{-1}(x) = 1_K(x) = x$. Entonces $\sigma(\sigma^{-1}(x)) = x$, por inyectividad $\sigma^{-1}(x) = x$. Por lo tanto, $G(K, F)$ es un grupo contenido en el grupo de todos los automorfismos de K , por lo tanto, es un subgrupo.

Definición 2.2.16. Sea K una extensión finita de F , decimos que K es una extensión normal de F si F es el campo fijo de $G(K, F)$. Es decir, $F = K_{G(K, F)}$.

Cuando K sea una extensión normal de F y sea el campo de descomposición de F para algún polinomio $f(x) \in F[x]$, diremos que $G(K, F)$ es el grupo de Galois. Además, veremos que ser una extensión normal y un campo de descomposición son de hecho conceptos que coinciden.

Por su extensión, remitimos la demostración del siguiente lema al libro [14, pp. 245].

Lema 2.2.2. Sea K el campo de descomposición de $f(x)$ en $F[x]$ y sea $p(x)$ un factor irreducible de $f(x) \in F[x]$. Si las raíces de $p(x)$ son $\alpha_1, \dots, \alpha_r$, entonces para cada i existe un automorfismo $\sigma_i \in G(K, F)$ tal que $\sigma_i(\alpha_1) = \alpha_i$.

Observación. En esencia, este lema nos da información acerca de lo que hacen los automorfismos de $G(K, F)$, en pocas palabras, siempre podemos encontrar un automorfismo entre dos raíces de un polinomio irreducible. Es decir, los automorfismos de $G(K, F)$ permutan las raíces de un polinomio irreducible. Lo cual motiva la siguiente definición.

Definición 2.2.17. Sea K una extensión normal de F y $\alpha \in K$. Definimos los **conjugados** de α como $\sigma(\alpha)$ para cada $\sigma \in G(K, F)$.

Nuevamente, omitiremos la demostración del siguiente teorema, la cual puede ser consultada en [14, pp. 244].

Teorema 2.2.10. Sea K una extensión normal de F , entonces

$$o(G(K, F)) = [K : F]$$

Este teorema en particular también nos da información acerca del comportamiento de los automorfismos de $G(K, F)$. Si la extensión es normal, entonces $o(G(K, F))$ es finito, este hecho será usado constantemente para el último teorema sin mayor mención.

Teorema 2.2.11. K es una extensión normal de F si y solo si K es el campo de descomposición de algún polinomio con coeficientes en F .

Demostración. Supongamos que K es una extensión normal de F , por la Definición 2.2.16, es una extensión finita. Por el Teorema 2.2.6, $K = F(a)$. Consideremos el polinomio

$$p(x) = (x - \sigma_1(a))(x - \sigma_2(a)) \cdots (x - \sigma_n(a))$$

El cual es un polinomio sobre K , donde $\sigma_1, \dots, \sigma_n$ son todos los elementos de $G(K, F)$. Se puede verificar por inducción que

$$p(x) = x^n - \alpha_1 x^{n-1} + \alpha_2 x^{n-2} + \dots + (-1)^n \alpha_n \quad (2.4)$$

donde los coeficientes equivalen a:

$$\begin{aligned} \alpha_1 &= \sigma_1(a) + \sigma_2(a) + \dots + \sigma_n(a) \\ \alpha_2 &= \sum_{i < j} \sigma_i(a) \sigma_j(a) \\ &\vdots \\ \alpha_n &= \sigma_1(a) \sigma_2(a) \cdots \sigma_n(a) \end{aligned}$$

Afirmamos que cada α_i es invariante sobre $\sigma \in G(K, F)$, es decir, $\sigma(\alpha_i) = \alpha_i$ para toda i . Supongamos lo contrario, entonces existe α_i tal que $\sigma(\alpha_i) \neq \alpha_i$. Por lo tanto, notemos lo siguiente

$$\begin{aligned} \sigma(\alpha_i) &= \sigma\left(\sum_{i < j < \dots < k} \sigma_i(a) \sigma_j(a) \cdots \sigma_k(a)\right) \\ &= \sum_{i < j < \dots < k} \sigma(\sigma_i(a)) \sigma(\sigma_j(a)) \cdots \sigma(\sigma_k(a)) \end{aligned}$$

Por la Proposición 2.2.1, $\sigma(G(K, F)) = G(K, F)$, por lo tanto, la sumatoria se permuta en algún orden distinto, entonces, la única manera de que $\sigma(\alpha_i) \neq \alpha_i$, ocurre cuando al menos dos subíndices de la lista $\{i, j, \dots, k\}$ son iguales, sin pérdida de generalidad, supongamos que son m y n . Por lo tanto, $\sigma(\sigma_m) = \sigma(\sigma_n) = \sigma_h$ para alguna h , pero esto contradice que $\sigma(G(K, F)) = G(K, F)$, porque existe una biyección entre los elementos de $\sigma(G(K, F))$ y $G(K, F)$ y aquí hay dos elementos de $\sigma(G(K, F))$ que le corresponde a un elemento de $G(K, F)$. En conclusión, $\sigma(\alpha_i) = \alpha_i$ para toda i .

Por lo tanto, $\alpha_1, \dots, \alpha_n$ pertenecen al campo fijo de $G(K, F)$, como K es una extensión normal de F , entonces F es el campo fijo. Como a es una raíz de $p(x)$, pues al menos algún $\sigma_i = 1_{G(K, F)}$, y $K = F(a)$ es la mínima extensión que contiene a a y F , entonces K es el campo de descomposición de $p(x) \in F[x]$.

Para la otra dirección, asumamos que K es el campo de descomposición del polinomio $f(x) \in F[x]$. Queremos mostrar que K es una extensión normal de F . Procederemos por inducción sobre la dimensión $[K : F]$, asumiendo que para cualquier par de campos K_1, F_1 tal que $[K_1, F_1] \leq [K, F]$ donde K_1 es el campo de descomposición sobre F_1 de un polinomio en $F_1[x]$, entonces K_1 es una extensión normal de F_1 .

Sea $[K : F] = 1$. Afirmo que $K = F$, de lo contrario, existe $a \in K$ y $a \notin F$. Pero una base de K sobre F es $\{1\}$, por lo cual, $a = 1 \cdot x$, con $x \in F$, claramente $x = a$, lo cual es una contradicción. Claramente K es una extensión normal de F , ya que $K = K_{G(K, F)} = K_{G(K, K)}$.

Supongamos que el teorema es válido para $[K : F] < n$. Ahora, asumamos que $f(x)$ tiene un factor irreducible $p(x) \in F[x]$ de grado $r > 1$. Entonces, hay r distintas raíces $\alpha_1, \dots, \alpha_r$ de $p(x)$ que pertenecen a K y K es el campo de descomposición de $f(x)$ visto como un polinomio sobre $F(\alpha_1)$. Por el teorema 2.2.1:

$$[K : F(\alpha_1)] = \frac{[K : F]}{[F(\alpha_1) : F]} = \frac{n}{r} < n$$

Por la hipótesis de inducción, K es una extensión normal de $F(\alpha_1)$. Ahora, sea $\theta \in K_{G(K, F)}$, queremos ver que $\theta \in F$. Ahora, cualquier $\tau \in G(K, F(\alpha_1))$ deja a F fijo, por lo tanto, deja a θ fijo. Por la normalidad de K sobre $F(\alpha_1)$, esto implica que $\theta \in F(\alpha_1)$, así, puede expresarse como un elemento de la base, por el teorema 2.2.3.

$$\theta = \lambda_0 + \lambda_1 \alpha_1 + \dots + \lambda_{r-1} \alpha_1^{r-1} \quad \text{donde } \lambda_0, \dots, \lambda_{r-1} \in F$$

Por el lema 2.2.2, existe $\sigma_i \in G(K, F)$ tal que $\sigma_i(\alpha_1) = \alpha_i$, como σ_i deja fijo a θ y cada λ_j queda fijo por estar en F , entonces, para toda i existe σ_i tal que:

$$\sigma_i(\theta) = \theta = \lambda_0 + \lambda_1 \alpha_i + \dots + \lambda_{r-1} \alpha_i^{r-1}$$

Consideremos

$$q(x) = \lambda_{r-1} x^{r-1} + \lambda_{r-2} x^{r-2} + \dots + \lambda_1 x + (\lambda_0 - \theta) \in K[x]$$

$$\therefore q(\alpha_i) = 0 \quad \forall i$$

Entonces $q(x)$ tiene grado a lo más $r - 1$ y tiene las r distintas raíces $\alpha_1, \dots, \alpha_r$. Esto solo puede pasar si todos sus coeficientes son 0, en particular, $\lambda_0 - \theta = 0$, pero $\theta = \lambda_0 \in F$. Por lo tanto, $F = K_{G(K, F)}$, por lo tanto, K es una extensión normal de F . Esto completa la inducción y el teorema. ■

Observación. Este teorema es el más importante de todos y conecta todas las ideas que necesitamos de esta sección. Como ser una extensión normal y ser el campo de descomposición de F son equivalentes, entonces solo necesitamos garantizar que K es el campo de descomposición de algún polinomio en $F[x]$ para garantizar la existencia del grupo de Galois $G(K, F)$ y poder utilizar el hecho de que $\sigma(\alpha)$ para todo $\sigma \in G(K, F)$ son los conjugados de α .

Más aún, cualquier polinomio en $F[x]$ debe tener algún campo de descomposición, pues basta considerar $F(\alpha_1, \dots, \alpha_r) = F(c)$ la mínima extensión de F que contiene todas las raíces de $f(x)$. Entonces, bajo estas ideas, siempre podremos hablar del grupo de Galois en cualquier campo F , independientemente de si describimos explícitamente su campo de descomposición.

Ejemplo: Supongamos que queremos una extensión normal sobre $\mathbb{Q}$ para los números algebraicos $\sqrt{2}$ y $\sqrt{3}$. Consideremos

$$f(x) = (x^2 - 2)(x^3 - 3) \in \mathbb{Q}[x]$$

Notemos que $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ contiene todas las raíces de $f(x)$ y ningún subcampo propio de K que es una extensión de $\mathbb{Q}$ contiene todas las raíces (pues de hacerlo contendría a $\sqrt{2}, \sqrt{3}$ y $\mathbb{Q}$, por lo tanto a la mínima extensión K), por lo tanto, K es el campo de descomposición de F para el polinomio $f(x) \in \mathbb{Q}[x]$, por el teorema anterior, K es una extensión normal sobre $\mathbb{Q}$.

2.3 Teoría de números algebraicos

A partir de ahora, haremos una distinción entre ser algebraico, número algebraico y entero algebraico. Aunque esta terminología no es estándar, aclarará con qué se está trabajando sin lugar a ambigüedades. Sin más, veamos los siguientes resultados.

Definición 2.3.1. Sea ζ un número complejo. Decimos que ζ es un número algebraico si existe un polinomio no nulo $p(x) \in \mathbb{Q}[x]$ tal que $p(\zeta) = 0$.

Denotamos el conjunto de todos los números algebraicos como $\mathbb{A}$, aunque también es común verlos como $\overline{\mathbb{Q}}$.

Proposición 2.3.1. Existe un polinomio no nulo $p(x) \in \mathbb{Z}[x]$ tal que $p(\zeta) = 0$ si y solo si ζ es un número algebraico.

Demostración: Supongamos que ζ es un número algebraico, entonces existe $p(x) \in \mathbb{Q}[x]$ tal que $p(\zeta) = 0$. Por lo tanto, los coeficientes del polinomio están en $\mathbb{Q}$, multiplicando el polinomio por el producto de los denominadores se obtiene un polinomio con coeficientes en los enteros. Entonces, existe $q(x) \in \mathbb{Z}[x]$ tal que $q(\zeta) = p(\zeta)(\prod q_i) = 0$ como se requiere. Por otro lado, supongamos que existe un polinomio no nulo $p(x) \in \mathbb{Z}[x]$ tal que $p(\zeta) = 0$, claramente $p(x) \in \mathbb{Q}[x]$, por lo tanto, ζ es un número algebraico. ■

Definición 2.3.2. Definimos el grado de un número algebraico ζ como el grado del mínimo polinomio no nulo $p(x) \in \mathbb{Z}[x]$ tal que $p(\zeta) = 0$.

Es decir, no existe $q(x) \in \mathbb{Z}[x]$ tal que $q(\zeta) = 0$ y $\deg q(x) < \deg p(x)$.

Definición 2.3.3. Decimos que un polinomio $p(x) \in F[x]$ es reducible si existen polinomios $q(x), h(x) \in F[x]$ con $\deg q(x), \deg h(x) < \deg p(x)$ tales que $p(x) = q(x)h(x)$. En caso contrario se dice que el polinomio es irreducible.

Es importante que los polinomios $q(x)$ y $h(x)$ estén en $F[x]$.

Definición 2.3.4. Decimos que un polinomio $p(x)$ de grado n de la forma $p(x) = a_0 + a_1x + \dots + a_nx^n$ es mónico si $a_n = 1$.

Notemos que estas son las definiciones formales, pero queremos demostrar que los números algebraicos comparten las propiedades que estamos acostumbrados para realizar operaciones. Para ello, veamos el siguiente teorema.

Teorema 2.3.1. El conjunto de números algebraicos $\mathbb{A}$ forma un campo.

Demostración. Claramente $\mathbb{A} \subset \mathbb{C}$, por lo cual, muchas propiedades se heredan. Es suficiente ver que $-\alpha, \alpha^{-1}, \alpha + \beta, \alpha \cdot \beta \in \mathbb{A}$ cuando $\alpha, \beta \in \mathbb{A}$. Sea $\alpha \in \mathbb{A}$, entonces existe un polinomio irreducible $p(x) \in \mathbb{Q}[x]$ tal que $p(\alpha) = 0$. Supongamos que $p(x)$ es de la forma

$$p(x) = a_n x^n + \dots + a_1 x + a_0$$

Definimos $q(x) = (-1)^n a_n x^n + \dots + (-1)a_1 x + a_0$, entonces:

$$q(-\alpha) = a_n \alpha^n + \dots + a_1 \alpha + a_0 = 0$$

Por lo tanto, $-\alpha \in \mathbb{A}$. De manera análoga α^{-1} satisface el polinomio $\alpha^n x^n f(\frac{1}{x})$ (donde n es el grado del polinomio f y $f(\alpha) = 0$ con $\alpha \neq 0$). Por lo tanto, $\alpha^{-1} \in \mathbb{A}$. Ahora, sea K una extensión normal sobre $\mathbb{Q}$ que contiene los n conjugados de α y m conjugados de β , definimos el polinomio:

$$h(x) = \prod_{i=1}^n \prod_{j=1}^m (x - \sigma_i(\alpha) \sigma_j(\beta))$$

$$t(x) = \prod_{i=1}^n \prod_{j=1}^m (x - (\sigma_i(\alpha) + \sigma_j(\beta)))$$

Utilizando los argumentos vistos en la demostración del Teorema 2.2.11, es claro que $h(x)$ y $t(x)$ son polinomios en $\mathbb{Q}[x]$ tal que $\alpha\beta$ y $\alpha + \beta$ son raíces de estos polinomios. Por supuesto, esto no garantiza que sean irreducibles, sin embargo, supongamos que $h(x) = g(x)r(x) \in \mathbb{Q}[x]$, entonces $h(\alpha\beta) = g(\alpha\beta)r(\alpha\beta) = 0$, por lo tanto, $\alpha\beta$ es raíz de $g(x)$ o $r(x)$, como es suficiente repetir este argumento una cantidad finita de veces (pues $\deg g(x), \deg r(x) < \deg h(x)$), en algún punto obtendremos un polinomio irreducible. Esto garantiza que $\alpha\beta$ y $\alpha + \beta$ son raíces de un polinomio irreducible en $\mathbb{Q}[x]$. Por lo tanto, $\alpha\beta, \alpha + \beta \in \mathbb{A}$. ■

Definición 2.3.5. Sea ζ un número complejo. Decimos que ζ es un entero algebraico, si existe un polinomio mónico no nulo $p(x) \in \mathbb{Z}[x]$ con $p(\zeta) = 0$.

La distinción entre ser un número algebraico y un entero algebraico es muy sutil, pero no debe ignorarse. Los argumentos del Teorema 2.3.1 son suficientes para ver que $-\alpha, \alpha + \beta$ y $\alpha\beta$ son enteros algebraicos si α y β lo son, pero no podemos garantizar que α^{-1} sea un entero algebraico. De manera equivalente, podemos garantizar que el polinomio irreducible de un número algebraico está en $\mathbb{Z}[x]$, pero no podemos garantizar que sea mónico.

Proposición 2.3.2. Sea $\alpha \in \mathbb{Q}$ un entero algebraico, entonces $\alpha \in \mathbb{Z}$.

Demostración. Supongamos que α es un entero racional de la forma $\frac{p}{q}$ una fracción irreducible, es decir, $(p, q) = 1$, por ser un entero algebraico existe un polinomio mónico no nulo $p(x) \in \mathbb{Z}[x]$ tal que $p(\alpha) = 0$. Se tiene que,

$$\frac{p^n}{q^n} + \dots + a \frac{p}{q} + a_0 = 0$$

Multiplcando por q^n :

$$p^n + \dots + a(q^n)p + a_0q^n = 0 \Rightarrow p^n \equiv 0 \pmod{q}$$

Así, $q|p^n$, entonces $q|p \cdot p^{n-1}$, por el lema 2.1.1 $q|p^{n-1}$, por un argumento inductivo, $q|p$, además, $q|q$, como 1 es el máximo común divisor entre p y q , por definición, $q \leq 1$, por ser un entero, $q = 1$. Por lo tanto, $\alpha \in \mathbb{Z}$. ■

Lema 2.3.1. Sea α y β números algebraicos en un campo K tal que $h = [K : \mathbb{Q}]$. Si los conjugados de α son $\alpha_1, \dots, \alpha_h$ y para β son $\beta_1, \dots, \beta_h$, entonces los conjugados para $\alpha\beta$ y $\alpha + \beta$ son $\alpha_1\beta_1, \dots, \alpha_h\beta_h$ y $\alpha_1 + \beta_1, \dots, \alpha_h + \beta_h$ respectivamente.

Demostración. Recordemos que los conjugados de α y β son $\sigma(\alpha), \sigma(\beta)$ para todo $\sigma \in G(K, \mathbb{Q})$. Como σ es un automorfismo de K , entonces abre sumas y productos, es decir, $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$ y $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$. ■

Lema 2.3.2. Sea α un número algebraico, entonces existe un entero r tal que $r\alpha$ es un entero algebraico.

Demostración. Por la Proposición 2.3.1 existe un polinomio no nulo $p(x) \in \mathbb{Z}[x]$ tal que $p(\alpha) = 0$. Entonces, el polinomio tiene la forma:

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 : a_i \in \mathbb{Z} \quad \forall i$$

Definamos el siguiente polinomio:

$$q(x) = x^n + a_{n-1} x^{n-1} + a_{n-2} (a_n) x^{n-2} + \dots + a (a_n^{n-2}) x + a_0 (a_n^{n-1})$$

Claramente $q(x) \in \mathbb{Z}[x]$ y además:

$$\begin{aligned} q(a_n \alpha) &= a_n^n \alpha^n + a_{n-1} (a_n^{n-1}) \alpha^{n-1} + \dots + a (a_n^{n-1}) \alpha + (a_n^{n-1}) a_0 \\ &= a_n^{n-1} (a_n \alpha^n + a_{n-1} \alpha^{n-1} + \dots + a \alpha + a_0) = a_n^{n-1} (p(\alpha)) = 0 \end{aligned}$$

Por lo tanto, $q(x)$ es un polinomio mónico no nulo con coeficientes en los enteros tal que $q(a_n \alpha) = 0$, por lo tanto, $a_n \alpha$ es un entero algebraico. ■

Definición 2.3.6. Sea K es una extensión de grado h sobre los racionales. Decimos que los enteros algebraicos $\beta_1, \dots, \beta_h$ forman una base integral para K , si cada entero algebraico $\alpha \in K$ puede ser expresado de manera única de la manera:

$$\alpha = b_1 \beta_1 + \dots + b_h \beta_h : b_i \in \mathbb{Z} \quad \forall i$$

Ejemplo: Sea $K = \mathbb{Z}(\sqrt{3})$, es decir, una extensión finita de los racionales. Notemos que $\alpha = \sqrt{3}$ es raíz del polinomio irreducible $p(x) = x^2 - 3$, con raíces $\pm\sqrt{3}$, entonces K es el campo de descomposición de $p(x) \in \mathbb{Z}[x]$, por el Teorema 2.2.11 K es el campo de descomposición de $\mathbb{Z}$, por el Teorema 2.2.3 $[K : \mathbb{Z}] = 2$, ya que $\alpha \in K$ es algebraico de grado 2 sobre $\mathbb{Z}$. Además, por el mismo teorema, una base para K es $\{1, \alpha\}$. Por lo tanto, una base integral para K es:

$$K = \{b_1(1) + b_2(\sqrt{3}) : b_1, b_2 \in \mathbb{Z}\}$$

El siguiente lema puede ser encontrado en el libro [17, pp. 63].

Lema 2.3.3. Si K es una extensión de grado h sobre los racionales. Entonces existen enteros algebraicos $\beta_1, \dots, \beta_h$ que forman una base integral y el determinante de dicha base es un entero distinto de 0.

Definición 2.3.7. Sea α un número algebraico del campo K de grado n sobre los racionales, donde K es una extensión normal. Entonces la norma $N(\alpha)$ se define como el producto de α y sus conjugados, es decir:

$$N(\alpha) = \prod_{\sigma \in G(K, \mathbb{Q})} \sigma(\alpha)$$

Observación: Como cada σ abre productos, entonces también N .

Lema 2.3.4. Sea α un entero algebraico de grado n sobre los racionales. Entonces $N(\alpha) \in \mathbb{Z}$.

Demostración. Sea $p(x) \in \mathbb{Q}[x]$ el polinomio irreducible de α , y sea K el campo de descomposición de $p(x)$, por lo tanto, K es una extensión normal de $\mathbb{Q}$, notemos que para todo $\tau \in G(K, \mathbb{Q})$, por ser automorfismos, abren productos, así,

$$\tau(N(\alpha)) = \tau \left(\prod_{\sigma \in G(K, F)} \sigma(\alpha) \right) = \prod_{\sigma \in G(K, \mathbb{Q})} \tau\sigma(\alpha)$$

Por el Teorema 2.2.1, $\tau G(K, F) = G(K, F)$, como en el producto se encuentran todos los $\sigma \in G(K, F)$, entonces esencialmente τ solo permuta a los σ 's, por lo cual, $\tau(N(\alpha)) = N(\alpha)$. Es decir, $N(\alpha)$ es invariante bajo automorfismos de $G(K, F)$. Entonces, $N(\alpha) \in K_{G(K, \mathbb{Q})}$, por ser una extensión normal $K_{G(K, \mathbb{Q})} = \mathbb{Q}$, así, $N(\alpha) \in \mathbb{Q}$ y notemos que $N(\alpha)$ es un producto de enteros algebraicos (pues cada $\sigma(\alpha)$ es una raíz del mismo polinomio irreducible), por lo tanto es un entero algebraico, tal que pertenece a los racionales. Por la Proposición 2.3.2 $N(\alpha) \in \mathbb{Z}$. ■

Ejemplo: Consideremos w una raíz de la unidad tal que $w^5 = 1$. Consideremos el polinomio $p(x) = x^5 - 1 = (x - 1)(x^4 + x^3 + x^2 + x + 1)$. Entonces el polinomio $q(x) = x^4 + x^3 + x^2 + x + 1$ es el polinomio irreducible tal que $q(w) = 0$ ya que $p(w) = 0$, entonces el campo de descomposición de $p(x)$ es de grado 4 y por lo tanto, $[K : \mathbb{Q}] = o(G(K, \mathbb{Q})) = 4$. Utilizando un software, podemos conocer las raíces de este polinomio, las cuales son:

$$\begin{aligned} \bullet w_1 &= -\frac{1}{4} - \frac{\sqrt{5}}{4} - i\sqrt{\frac{5}{8} - \frac{\sqrt{5}}{8}} \\ \bullet w_2 &= -\frac{1}{4} + \frac{\sqrt{5}}{4} + i\sqrt{\frac{5}{8} + \frac{\sqrt{5}}{8}} \\ \bullet w_3 &= -\frac{1}{4} + \frac{\sqrt{5}}{4} - i\sqrt{\frac{5}{8} + \frac{\sqrt{5}}{8}} \\ \bullet w_4 &= -\frac{1}{4} - \frac{\sqrt{5}}{4} + i\sqrt{\frac{5}{8} - \frac{\sqrt{5}}{8}} \end{aligned}$$

Por el lema 2.2.2, existe $\sigma_i \in G(K, \mathbb{Q})$ tal que $\sigma_i(w_1) = w_i$. Por lo tanto,

$$N(w_1) = \prod_{\sigma \in G(K, \mathbb{Q})} \sigma(w_1) = w_1 \cdot w_2 \cdot w_3 \cdot w_4$$

Por el lema anterior, sorprendentemente $N(w_1) \in \mathbb{Z}$. Auxiliados de un software, podemos confirmar que $N(w_1) = 1$ y efectivamente está en los enteros. De manera alternativa, como $w^5 = 1$, entonces $|N(w_1)|^5 = |w_1 \cdots w_4|^5 = |w_1^5 \cdots w_4^5| = 1$, en los enteros, esto implica que $N(w_1) = 1$.

Lema 2.3.5. Sea α un número algebraico, entonces

$$N(\alpha) = 0 \iff \alpha = 0$$

Demostración. Si $N(\alpha) = 0$, entonces $\sigma_i(\alpha) = 0$ para alguna i , por el Lema 2.2.2, 0 es una raíz del polinomio irreducible de a . Supongamos que

$$h(x) = a_n x^n + \dots + a_1 x + a_0 \in \mathbb{Q}[x]$$

$$h(0) = a_0 = 0$$

Por lo tanto, $h(x) = x(a_n x^{n-1} + \dots + a_1)$, necesariamente $a_j = 0$ para $j > 1$, de lo contrario $h(x)$ es un polinomio irreducible. Por lo tanto, $h(x) = a_1 x$, como $h(\alpha) = a_1(\alpha) = 0$, entonces $\alpha = 0$, ya que $h(x)$ no es el polinomio nulo por definición. Por otro lado, si $\alpha = 0$, entonces su polinomio irreducible es $h(x) = c(x)$ con $c \neq 0$. Entonces $N(\alpha)$ es el producto de las raíces de $h(x)$, pero solo α es una raíz, entonces $N(\alpha) = 0$.

Definición 2.3.8. Sean α, β números algebraicos y $m \in \mathbb{Z}^+$, escribimos $\alpha \equiv \beta \pmod{m}$ si $\frac{\alpha - \beta}{m}$ es un entero algebraico.

En particular, nos interesan las propiedades de las congruencias. Sin embargo, estas se cumplen trivialmente siguiendo las demostraciones del capítulo anterior considerando que $\gamma = \frac{\alpha - \beta}{m}$ es un número algebraico y las sumas o productos de números algebraicos son nuevamente un número algebraico.

Definición 2.3.9. Decimos que un número α es **trascendental** si no es un número algebraico.

Desafortunadamente, con las herramientas vistas no podemos afirmar o negar casi nada acerca de los números trascendentales. Resulta, de hecho, que su estudio tiene muchos beneficios, como veremos más adelante. Por último, veamos la siguiente definición.

Definición 2.3.10. Definimos la **altura** de un número algebraico α como el máximo de los valores absolutos de los coeficientes del polinomio mínimo de α . Es decir, si $p(x) = a_n x^n + \dots + a_1 x + a_0$ es el polinomio mínimo de α , entonces la altura de α es el $\max |a_i|$ para $0 \leq i \leq n$.

2.4 Látices

Definición 2.4.1. Un retículo o *látice*, representado como Λ , en un espacio euclideo $\mathbb{R}^n$ es el conjunto de combinaciones lineales enteras de n vectores linealmente independientes en $\mathbb{R}^n$. En símbolos:

$$\Lambda = \sum_{i=1}^n b_i v_i \quad b_i \in \mathbb{Z}$$

Además los puntos de Λ son conocidos como puntos *látice*.

Ejemplo: Sea $\Lambda = \mathbb{Z}^n$, es decir, los puntos del espacio $\mathbb{R}^n$ cuyas coordenadas son todas enteras. Este *látice* se conoce como *cúbico*, *canónico* o *estándar*, y es el único en el que se trabajará.

Lema 2.4.1. Consideremos las m ecuaciones en las n incógnitas:

$$\begin{pmatrix} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,n)} \\ a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,n)} \\ \vdots & \vdots & & \vdots \\ a_{(m,1)} & a_{(m,2)} & \cdots & a_{(m,n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Donde $a_{(i,j)} \in \mathbb{Z}$ y $0 < m < n$. Sea $A \geq |a_{(i,j)}|$ una cota superior para toda i, j . Entonces existe una solución no trivial $x_1, \dots, x_n \in \mathbb{Z}$ tales que:

$$|x_j| < 1 + (nA)^{\frac{m}{n-m}} \quad j = 1, 2, \dots, n.$$

Demostración: Sea $y_k = a_{(k,1)}x_1 + \dots + a_{(k,n)}x_n$ de tal manera que para cada punto $x = (x_1, x_2, \dots, x_n)$ se corresponde a un punto $y = (y_1, y_2, \dots, y_m)$, donde cada coordenada de y_i es la suma de una fila de la matriz anterior para esa específica solución x . Claramente si x es un punto *látice*, y también (pues sus coordenadas son una combinación lineal de enteros). Sea q cualquier entero positivo. Recordemos que estamos trabajando en $\mathbb{Z}^n$, sin embargo, lo podemos restringir al n -cubo dimensional $(2q+1)^n$ donde $|x_j| \leq q$ para cualquier subíndice j . Es decir, vamos a considerar todos los posibles valores de x que cumplan $|x_j| < q$. Por lo tanto, los valores y_k satisfacen:

$$|y_k| = \left| \sum_{j=1}^n a_{(k,j)} x_j \right| \leq \sum_{j=1}^n |a_{(k,j)}| \cdot |x_j| \leq \sum_{j=1}^n Aq = nAq$$

Así, conforme x toma sus valores en los $(2q+1)^n$ puntos *látice*, las correspondientes coordenadas y_k son enteros entre $2nAq+1$ valores $0, \pm 1, \pm 2, \dots, \pm nAq$.

Así, las coordenadas y tienen a lo más $(2nAq + 1)^m$ posibles ubicaciones. Aún con tantas posibles ubicaciones, demostraremos que para alguna q existen más puntos que posibles ubicaciones (lo cual es posible porque $n > m$), es decir, que al menos dos puntos láctice y coinciden. Es decir:

$$(2q + 1)^n > (2nAq + 1)^m$$

Ahora, definamos el valor de $2q$ como el único entero par en el intervalo de tamaño 2 definido por:

$$(nA)^{\frac{m}{n-m}} - 1 \leq 2q < (nA)^{\frac{m}{n-m}} + 1$$

Como $(nA)^{\frac{m}{n-m}} - 1 \leq 2q$, entonces $(nA)^m \leq (2q + 1)^{n-m}$, por lo tanto,

$$\begin{aligned} (2nAq + 1)^m &= (nA)^m \left(2q + \frac{1}{nA}\right)^m < (nA)^m (2q + 1)^m \\ &\leq (2q + 1)^{n-m} (2q + 1)^m = (2q + 1)^n \end{aligned}$$

Lo cual demuestra que $(2q + 1)^n > (2nAq + 1)^m$ como se quería. Por lo tanto, conforme x varía en los $(2q + 1)^n$ puntos láctice definidos por $|x_j| < q$, los correspondientes puntos láctice y no son todos distintos. Digamos que tenemos un punto láctice y que corresponden a dos soluciones distintas $x = (x'_1, \dots, x'_n)$ y $x = (x''_1, \dots, x''_n)$. Veamos que la suma $x = (x'_1 - x''_1, \dots, x'_n - x''_n)$ es una solución, su correspondiente punto y' tiene como coordenadas:

$$\begin{aligned} y'_k &= a_{(k,1)}(x'_1 - x''_1) + \dots + a_{(k,n)}(x'_n - x''_n) \\ &= (a_{(k,1)}x'_1 + \dots + a_{(k,n)}x'_n) - (a_{(k,1)}x''_1 + \dots + a_{(k,n)}x''_n) = y_k - y_k = 0 \end{aligned}$$

Por lo tanto, la suma de cada fila de la matriz es 0. Por lo tanto, x es una solución del sistema. Además, esta solución es la buscada, pues notemos que:

$$|x'_j - x''_j| \leq |x'_j| + |x''_j| \leq q + q < (nA)^{\frac{m}{n-m}} + 1$$

■

Definición 2.4.2. Sea α un algebraico en un campo K , definimos $\|\alpha\|$ como el máximo de los valores absolutos de α y sus conjugados.

Lema 2.4.2 (Thue-Siegel). Consideremos las p ecuaciones en las q incógnitas:

$$\begin{pmatrix} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,n)} \\ a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,n)} \\ \vdots & \vdots & & \vdots \\ a_{(p,1)} & a_{(p,2)} & \cdots & a_{(p,n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (2.5)$$

Con coeficientes $a_{(i,j)}$ enteros algebraicos en un campo K de grado finito. Asumamos que $0 < p < q$. Sea $A \geq 1$ una cota superior para los coeficientes y conjugados en K , así, $A \geq \|a_{(i,j)}\|$ para toda i, j . Entonces, existe una constante positiva c que depende del campo K e independiente de $a_{(i,j)}$, p, q , tal que el sistema de ecuaciones tiene una solución no trivial $x = (x_1, \dots, x_p)$ en los enteros algebraicos de K donde:

$$\|x_k\| < c + c(cqA)^{\frac{p}{q-p}} \quad k = 1, \dots, q$$

Demostración: Sea K una extensión normal sobre los racionales tal que $[K : \mathbb{Q}] = h$, $\beta_1, \dots, \beta_h$ una base integral para el campo (la cual existe por el Lema 2.3.3, si $\alpha \in K$ es cualquier entero algebraico, entonces podemos expresar α como una combinación lineal única de la base integral,

$$\alpha = g_1\beta_1 + g_2\beta_2 + \dots + g_h\beta_h$$

con enteros g_j . Como K es una extensión normal, entonces $o(G(K, F)) = [K : \mathbb{Q}] = h$, es decir, existen h conjugados para cada $a \in K$. Denotemos los conjugados de α como $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(h)}$, análogamente para β_j . Por el lema 2.3.1,

$$\alpha^{(i)} = g_1\beta_1^{(i)} + g_2\beta_2^{(i)} + \dots + g_h\beta_h^{(i)} \quad i = 1, 2, \dots, h$$

Por lo tanto, viéndolo como un sistema de Álgebra lineal de la forma $Ax = b$,

$$\begin{pmatrix} \beta_1^{(1)} & \beta_2^{(1)} & \dots & \beta_h^{(1)} \\ \beta_1^{(2)} & \beta_2^{(2)} & \dots & \beta_h^{(2)} \\ \vdots & \vdots & \dots & \vdots \\ \beta_1^{(h)} & \beta_2^{(h)} & \dots & \beta_h^{(h)} \end{pmatrix} \cdot \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_h \end{pmatrix} = \begin{pmatrix} \alpha^{(1)} \\ \alpha^{(2)} \\ \vdots \\ \alpha^{(h)} \end{pmatrix}$$

Utilizando la regla de Cramer de Álgebra lineal [10, pp. 221], podemos resolver este sistema de ecuaciones para las g_j con coeficientes que dependen solamente de la base y sus conjugados como sigue

$$g_j = \frac{\det(A_i)}{\det(A)} \quad \text{tal que } A_i = \begin{pmatrix} \beta_1^{(1)} & \dots & \alpha^{(1)} & \dots & \beta_h^{(1)} \\ \beta_1^{(2)} & \dots & \alpha^{(2)} & \dots & \beta_h^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_1^{(h)} & \dots & \alpha^{(h)} & \dots & \beta_h^{(h)} \end{pmatrix}$$

Es decir, se reemplaza la columna i por el vector b . Por simplicidad, se puede intercambiar la columna i para facilitar el cálculo del $\det(A_i)$, haciendo esto,

es claro que podemos encontrar una constante c_1 que depende únicamente del campo K (e independiente de α) tal que

$$|g_j| < c_1 \|\alpha\| \quad (2.6)$$

Ahora, regresamos al sistema original (2.5) y expresamos x_i como un elemento de la base, por lo tanto,

$$x_i = \sum_{j=1}^h \epsilon_{(i,j)} \beta_j \quad \forall i = 1, \dots, h \quad (2.7)$$

Por lo tanto, el sistema de ecuaciones (2.5) puede escribirse como:

$$\sum_{i=1}^q \sum_{j=1}^h a_{(k,i)} \beta_j \epsilon_{(i,j)} = 0 \quad k = 1, 2, \dots, p \quad (2.8)$$

Ahora, los enteros algebraicos $a_{(k,i)} \beta_j$ pueden escribirse en términos de la base integral,

$$a_{(k,i)} \beta_j = \sum_{r=1}^h m_{(k,i,j,r)} \beta_r \quad k = 1, \dots, p \quad i = 1, \dots, q \quad j = 1, \dots, h$$

Con enteros $m_{(k,i,j,r)}$ (por ser una base integral). Entonces, el sistema de ecuaciones modificado (2.8) puede escribirse como

$$\sum_{i=1}^q \sum_{j=1}^h \sum_{r=1}^h m_{(k,i,j,r)} \epsilon_{(i,j)} \beta_r = 0 \quad k = 1, 2, \dots, p$$

Ahora, como los β_r son linealmente independiente sobre los racionales, la única manera que se cumpla la ecuación anterior es que los coeficientes de cada β_r sean 0. Esto nos lleva al siguiente sistema de ecuaciones

$$\sum_{i=1}^q \sum_{j=1}^h m_{(k,i,j,r)} \epsilon_{(i,j)} = 0 \quad k = 1, \dots, p \quad r = 1, \dots, h \quad (2.9)$$

Ahora, hemos obtenido un sistema de ecuaciones final, donde tenemos ph ecuaciones en las qh incógnitas $\epsilon_{(i,j)}$, y queremos aplicar el Lema 2.4.1. Para ello, necesitamos una cota superior de los valores absolutos de los coeficientes $m_{(k,i,j,r)}$. Aplicamos el resultado que obtuvimos en (2.6) con α reemplazado por $\alpha_{(k,i)} \beta_j$ y g_j sustituido por $m_{(k,i,j,r)}$, así

$$|m_{(k,i,j,r)}| < c_1 \|a_{k,i} \beta_j\| \leq c_1 \|\alpha_{(k,i)}\| \cdot \|\beta_j\| \leq c_1 A \|\beta_j\| \leq c_2 A$$

Donde A fue especificado en el presente Lema 2.4.2 y c_2 es una constante positiva seleccionada tal que $c_2 \geq c_1 \|\beta_j\|$ y $c_2 A$ es un entero, lo cual es posible por el Lema 2.3.4. Por lo tanto, por el Lema 2.4.1, con m, n y A sustituidos por ph, qh y $c_2 A$ respectivamente, existe una solución no trivial para el sistema (2.9) con enteros $\epsilon_{(i,j)}$ que satisface

$$|\epsilon_{(i,j)}| < 1 + (qhc_2 A)^{\frac{ph}{qh-ph}} = 1 + (hc_2 qA)^{\frac{p}{q-p}}$$

Sustituyendo estos estimados en (2.7), obtenemos una solución para el sistema (2.5) que satisface las siguientes desigualdades

$$\|x_i\| < h \cdot \max_j \|\beta_j\| (1 + hc_2 qA)^{\frac{p}{q-p}}$$

$$\Rightarrow \|x_i\| < c + c(cqA)^{\frac{p}{q-p}}$$

Donde se escoge $c > hc_2$ y $c > h \cdot \|\beta_j\|$ para toda j . Así, c depende únicamente del campo K y es independiente de $\alpha_{(i,j)}$, p y q . Aún más, como al menos algún $\epsilon_{(i,j)}$ es distinto de 0, entonces al menos algún $x_{(i,j)}$ no es cero, porque los β_j son linealmente independientes sobre los racionales. Lo cual finaliza la demostración. ■

El siguiente Lema es análogo al Lema 2.4.1, aunque es mi propia adaptación de las ideas encontradas en [18] para simplificar el Teorema de Baker.

Lema 2.4.3. Sea K una extensión normal sobre los racionales de grado $[K : \mathbb{Q}] = d > 1$, consideremos las w ecuaciones en las v incógnitas:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_w \end{pmatrix} = \begin{pmatrix} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,v)} \\ a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,v)} \\ \vdots & \vdots & \cdots & \vdots \\ a_{(w,1)} & a_{(w,2)} & \cdots & a_{(w,v)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Donde $a_{(i,j)}$ son enteros algebraicos del campo K y $1 \leq w < v$. Sea A una cota superior entera y positiva para los coeficientes y conjugados en K , así $A \geq \|a_{(i,j)}\|$ para toda i, j . Entonces existe una solución no trivial $x_1, \dots, x_v \in \mathbb{Z}$ tales que:

$$|x_j| < 1 + (vA)^{\frac{wd(d+1)}{v-wd(d+1)}} \quad j = 1, 2, \dots, v.$$

Cuando $v > wd(d+1)$.

Demostración: Sea $X \geq 2$, Y números naturales. Definamos:

$$I_X = \{(x_1, \dots, x_v) : |x_i| \leq X\}$$

$$J_Y = \{(y_1, \dots, y_w) : \|y_j\| \leq Y\}$$

Entonces, seguiremos el argumento del Lema 2.4.1, por lo que tenemos,

$$I_X \leq (2X + 1)^v \quad y \quad \|y_j\| \leq vAX$$

Una distinción importante, es que y_j es un entero algebraico, por lo tanto, satisface una ecuación de la forma

$$y_j^d + b_1 y_j^{d-1} + \dots + b_v = 0 \quad b_i \in \mathbb{Z}$$

Ahora, utilizando las fórmulas para los coeficientes que se obtienen a partir de la ecuación (2.4), podemos concluir que:

$$|b_i| \leq \binom{d}{i} (vAX)^i$$

Por lo tanto, el número de ecuaciones que se satisfacen por y_j es a lo más

$$\prod_{i=1}^d \left(2 \binom{d}{i} (vAX)^i + 1 \right)$$

Y cada ecuación tiene a lo más d posibles valores para y_j (pues es el máximo de conjugados que puede tener por el grado del campo sobre los racionales).

Así,

$$|J_{vAX}| \leq \left(d \prod_{i=1}^d \left(2 \binom{d}{i} (vAX)^i + 1 \right) \right)^w = d^w 2^{wd} (vAX)^{w \frac{d(d+1)}{2}} \prod_{i=1}^d \left(\binom{d}{i} + \frac{1}{2(vAX)^i} \right)^w$$

Se puede demostrar que para $n \geq 2$

$$\frac{a_1 + a_2 + \dots + a_n}{n} \geq \sqrt[n]{a_1 \cdot a_2 \cdots a_n}$$

Lo cual se conoce como desigualdad entre media aritmética y geométrica. Así, utilizando lo anterior y $vAX \geq 2$ obtenemos

$$\begin{aligned} \prod_{i=1}^d \left(\binom{d}{i} + \frac{1}{2^{i+1}} \right)^w &\leq \left(\frac{\binom{d}{1} + \dots + \binom{d}{d} + \frac{1}{2^2} + \dots + \frac{1}{2^{v+1}}}{d} \right)^{vw} \\ &\leq \left(\frac{2^d - \frac{1}{2}}{d} \right)^{dw} \end{aligned}$$

Entonces

$$|J_{vAX}| \leq 2^{wd(d+1)}(vAX)^{w\frac{d(d+1)}{2}} \leq (2vAX + 1)^{wd(d+1)}$$

Buscamos lo siguiente

$$(2X + 1)^v > (2vAX + 1)^{wd(d+1)}$$

Cuando $v > wd(d + 1)$. Repitiendo exactamente los mismos argumentos del Lema 2.4.1, se obtiene el resultado de este lema. ■

2.5 Análisis matemático

Definición 2.5.1. Decimos que l es el límite de $f(x)$ cuando x se aproxima a a , denotado por:

$$\lim_{x \rightarrow a} f(x) = l$$

si y solo si para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $0 < |x - a| < \delta$, entonces $|f(x) - l| < \epsilon$.

Definición 2.5.2. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función, decimos que f es continua en a si $\lim_{x \rightarrow a} f(x)$ existe y $\lim_{x \rightarrow a} f(x) = f(a)$.

Proposición 2.5.1. La función $f(x) = x$ es continua para todo $x \in \mathbb{R}$.

Demostración. Sea $\epsilon > 0$ cualquier número real. Definamos $\delta = \epsilon$. Entonces si

$$0 < |x - a| < \delta \Rightarrow |f(x) - f(a)| = |x - a| < \delta = \epsilon$$

Por lo tanto, f es continua en a para toda $a \in \mathbb{R}$. ■

Teorema 2.5.1. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función polinomial. Entonces f es continua para todo $x \in \mathbb{R}$.

Demostración. Notemos que cualquier función polinomial se puede escribir como sigue

$$f(x) = a_n x^n + \dots + a_1 x + a_0$$

Acabamos de demostrar que la función $g(x) = x$ es continua, y por lo tanto,

$$\lim_{x \rightarrow a} g(x) = g(a)$$

Por propiedades de límites, y utilizando el hecho de que $f(x)$ es una combinación lineal de sumas y productos de $g(x)$

$$\lim_{x \rightarrow a} f(x) = a_n f(a)^n + \dots + a_1 (f(a)) + a_0$$

Por lo tanto, f es continua para toda $x \in \mathbb{R}$. ■

Definición 2.5.3. Sea A un conjunto. Decimos que α es una cota superior de A si y solo si $\alpha \geq a$ para toda $a \in A$. Decimos que β es cota inferior de A si $\beta \leq a$ para toda $a \in A$.

Por ejemplo, si consideramos el intervalo $(5, 10)$, entonces todo x tal que $x > 10$ es una cota superior de $(5, 10)$. Análogamente para la cota inferior.

Definición 2.5.4. Sea A un conjunto, decimos que α es el máximo de A si α es cota superior y $\alpha \in A$. Decimos que β es el mínimo si es cota inferior y $\beta \in A$.

No necesariamente existe el máximo y el mínimo. Por ejemplo, en el intervalo $(5, 10)$, intuitivamente, el máximo de A es 10, sin embargo, $10 \notin (5, 10)$, análogamente el mínimo de A no puede ser 5.

Definición 2.5.5. Sea A un conjunto, decimos que α es el supremo de A si es la mínima cota superior de A , y se denota como $\sup A = \alpha$. Decimos que β es el ínfimo de A si es la máxima cota inferior de A , y se denota por $\inf A = \beta$.

Es posible demostrar que el supremo e ínfimo siempre existen, pues ya no es necesario que pertenezcan al conjunto A .

Lema 2.5.1. Sea f una función continua en a . Entonces existe $\delta > 0$, tal que f es acotada en $(a - \delta, a + \delta)$.

Demostración. Como f es continua en a para $\epsilon = 1$, existe $\delta > 0$ tal que si $|x - a| < \delta$, entonces $|f(x) - f(a)| < 1$, en consecuencia, si $x \in (a - \delta, a + \delta)$, entonces $f(x) \in (f(a) - 1, f(a) + 1)$. Lo cual demuestra el lema. ■

A continuación, veremos que cualquier función continua definida en un intervalo cerrado hacia los reales está acotada.

Proposición 2.5.2. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función continua. Entonces existe $M > 0$ tal que $|f(x)| < M$ para toda $x \in [a, b]$.

Demostración. Consideremos el siguiente conjunto:

$$A = \{x \in [a, b] : \text{tales que } f \text{ está acotada en } [a, x]\}$$

Es claro que $A \neq \emptyset$, dado que por el lema anterior $a \in A$. Veamos que $b = \max A$, sea $\alpha = \sup A$. Supongamos que $\alpha < b$, por lo tanto, f es continua en α y por el lema anterior existe $\delta > 0$ tal que f está acotado en $(\alpha - \delta, \alpha + \delta)$ lo cual contradice que sea la mínima cota superior de A . Por lo tanto, $b = \alpha$. Ahora, veamos que $b \in A$, por lo anterior, sabemos que f es acotada en $[a, x]$ para todo $x \in [a, b)$. Como f es continua en b , existe $\delta' > 0$ tal que f está acotada en $(b - \delta', b]$. Escojamos $y \in (b - \delta', b]$, entonces f está acotada en $[a, y]$, tomando el máximo entre las dos cotas, entonces f está acotada en $[a, b]$. ■

Definición 2.5.6. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función, decimos que f es diferenciable (o derivable) en un punto $x_0 \in [a, b]$ si

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existe. Si f es derivable en cada punto de $[a, b]$ decimos que f es diferenciable en $[a, b]$.

Proposición 2.5.3. La identidad $f(x) = x$ es diferenciable.

Demostración. Consideremos

$$\lim_{x \rightarrow 0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = 1$$

■

Análogamente se puede demostrar para una función constante.

Proposición 2.5.4. Sean $f, g : [a, b] \rightarrow \mathbb{R}$ funciones diferenciables en x_0 . Entonces:

1. $f + g$ es diferenciable en x_0
2. $f \cdot g$ es diferenciable en x_0

Demostración. Sea $h(x) = f(x) + g(x)$, queremos ver que h es diferenciable en x_0 . Por definición,

$$\begin{aligned} h'(x_0) &= \lim_{x \rightarrow x_0} \frac{h(x) - h(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) + g(x) - f(x_0) - g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = f'(x_0) + g'(x_0) \end{aligned}$$

Pues f y g son diferenciables en x_0 . Lo cual concluye la demostración del inciso (1). Ahora, sea $h(x) = f(x)g(x)$. Por definición:

$$\begin{aligned} h'(x_0) &= \lim_{x \rightarrow x_0} \frac{h(x) - h(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{[f(x)g(x) + f(x)g(x_0)] - [f(x)g(x_0) + f(x_0)g(x_0)]}{x - x_0} \end{aligned}$$

$$\begin{aligned}
&= \lim_{x \rightarrow x_0} \frac{f(x)[g(x) - g(x_0)] + g(x_0)[f(x) - f(x_0)]}{x - x_0} \\
&= \lim_{x \rightarrow x_0} f(x) \frac{g(x) - g(x_0)}{x - x_0} + \lim_{x \rightarrow x_0} g(x_0) \frac{f(x) - f(x_0)}{x - x_0} \\
&= \lim_{x \rightarrow x_0} f(x) \cdot g'(x_0) + \lim_{x \rightarrow x_0} g(x_0) \cdot f'(x_0) \\
&= f(x_0)g'(x_0) + g(x_0) \cdot f'(x_0)
\end{aligned}$$

Lo cual concluye la demostración del inciso (2). ■

Corolario 2.5.1. Cualquier función polinomial es diferenciable

$$a_0 + a_1x + \dots + a_nx^n$$

Demostración. Como la identidad x y cualquier función constante es diferenciable por la Proposición 2.5.3 y cualquier función polinomial se puede obtener a través de sumas y productos de la identidad y la función constante, entonces es la suma y producto de funciones diferenciables, por lo tanto, por el teorema anterior, cualquier función polinomial es diferenciable. ■

Proposición 2.5.5. Si $f : [a, b] \rightarrow \mathbb{R}$ tiene un máximo en $x_0 \in (a, b)$ y $f'(x_0)$ existe, entonces $f'(x_0) = 0$.

Demostración. Supongamos que f alcanza su máximo en x_0 , por hipótesis,

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existe, entonces los límites laterales existen y son iguales, pero

$$\lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

$$\lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

Entonces

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = 0$$

Análogamente si f alcanza su mínimo en x_0 entonces $f'(x_0) = 0$. ■

Teorema 2.5.2. Sean $f, g : [a, b] \rightarrow \mathbb{R}$ son funciones continuas en $[a, b]$ y diferenciables en (a, b) . Entonces existe $x \in (a, b)$ tal que:

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$$

Demostración. Sea $h(t) = (f(b) - f(a))g(t) - (g(b) - g(a))f(t)$, entonces h es continua en $[a, b]$ y diferenciable en (a, b) , más aún,

$$h'(t) = (f(b) - f(a))g'(t) - (g(b) - g(a))f'(t)$$

Por otro lado, $h(a) = f(b)g(a) - g(b)f(a) = h(b)$. Si h es constante, entonces $h'(x) = 0$ para toda $x \in (a, b)$, si h no es constante por la continuidad existe $x \in (a, b)$ tal que h alcanza su máximo y por el teorema anterior $h'(x_0) = 0$ lo cual concluye la demostración. ■

Teorema 2.5.3 (Teorema del valor intermedio). Sea $f : [a, b] \rightarrow \mathbb{R}$ una función continua en el intervalo cerrado $[a, b]$ y diferenciable en el intervalo abierto (a, b) con $a < b$ entonces existe al menos algún punto $c \in (a, b)$ tal que:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Demostración. Sea $g(x) = x$, aplicando el teorema anterior se sigue que:

$$f(b) - f(a) = (b - a)f'(c)$$

Donde c es el máximo de $h'(t)$, lo cual concluye la demostración. ■

Aunque claramente es un corolario del teorema anterior, por su importancia, lo nombramos como un teorema.

2.6 Variable compleja

Este apartado tiene como intención facilitar herramientas para poder estudiar funciones $f(z)$ definidas en variables complejas. Nuestro objetivo será presentar conceptos básicos y teoremas que permitan derivar e integrar $f(z)$ como estamos acostumbrados en Cálculo.

Definición 2.6.1. Un número complejo en $\mathbb{C}$ es una expresión de la forma $a + bi$ donde a y b son números reales. Decimos que dos números complejos $a + bi$ y $c + di$ son iguales si y solo si $a = c$ y $b = d$.

Sea w la imagen de la función $f : \mathbb{C} \rightarrow \mathbb{C}$ en el punto z , como $f(z) = w$, entonces como $z = x + iy$ entonces $w = u(x, y) + i v(x, y)$. Por lo tanto, en esencia, una función de variable compleja es un par de funciones reales de dos variables reales donde $u(x, y)$ se le conoce como la parte real de f , o $\text{Re } f$ y a $v(x, y)$ se le conoce como la parte imaginaria de f .

En cálculo, estamos acostumbrados que los resultados estén típicamente enunciados para funciones que están definidas en algún intervalo. Para las funciones de variable compleja, los resultados se formulan para funciones que están definidas en regiones cerradas de 2 dimensiones. Por lo cual, definiremos de manera explícita el significado de estos conjuntos.

Definición 2.6.2. El conjunto de todos los puntos $z \in \mathbb{C}$ que satisfacen la desigualdad

$$|z - z_0| < p$$

Donde p es un número real positivo, se le conoce como un disco abierto o vecindades de z_0

Este conjunto, consiste de todo los puntos que están contenidos en el círculo de radio p y centro z_0 . En particular, los conjuntos que son una solución para las siguientes desigualdades:

$$|z - i| < 2, \quad |z + 2| < 5, \quad |z| < 10$$

Son las vecindades para los puntos i , -2 y 0 respectivamente.

Definición 2.6.3. Un punto z_0 que está contenido en un conjunto S se le llama punto interior de S , si existe una vecindad de z_0 que está completamente contenido en S . Si cada punto del conjunto S es un punto interior de S , decimos que S es un conjunto abierto.

A partir de esta definición, es claro que cada vecindad de algún punto z_0 es un conjunto abierto.

Definición 2.6.4. Sea f una función de variable compleja definida en una vecindad de z_0 . Entonces f es diferenciable (o derivable) en z_0 si:

$$\frac{df}{dz}(z_0) = f'(z_0) := \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

Es decir, f es derivable en z_0 si el límite existe.

La distinción que se hace, respecto a ser derivable en los números reales, es que Δz es un número complejo. Así que puede aproximarse hacia el cero en muchas direcciones, desde la derecha, desde abajo, a través de un espiral, etc. Es decir, $f'(z_0)$ debe tener un único límite independiente de cómo Δz se aproxima a 0.

Ejemplo: Sea $f(z) = \bar{z}$ (el conjugado complejo de z). Entonces, si aproximamos $\Delta z \rightarrow 0$ a través de los reales positivos y luego por los imaginarios positivos, obtenemos que:

$$\begin{aligned} f'(z_0) &= \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{\overline{z_0 + \Delta z} - \bar{z}_0}{\Delta z} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\overline{\Delta x}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta x}{\Delta x} = 1 \\ &= \lim_{\Delta y \rightarrow 0} \frac{\overline{\Delta y}}{\Delta y} = \lim_{\Delta y \rightarrow 0} -\frac{\Delta y}{\Delta y} = -1 \end{aligned}$$

Por lo tanto, concluimos que $f(z)$ no es analítica.

Proposición 2.6.1. Si f, g son diferenciables en z_0 , entonces:

- $f + g$ es diferenciable en z_0
- $f \cdot g$ es diferenciable en z_0
- $\frac{f}{g}$ es diferenciable en z_0 si $g(z_0) \neq 0$.

La demostración es una copia fiel del resultado para funciones reales, el cual vimos en la Proposición 2.5.4.

Definición 2.6.5. Una función de variable compleja $f(z)$ se dice que es analítica en un conjunto abierto G si $f(z)$ es derivable en cada punto de G . Si f es analítica en todo el plano complejo, entonces se dice que f es una función entera.

Proposición 2.6.2. La función identidad $f(z) = z$ es una función entera.

Demostración. La demostración, nuevamente, es una copia fiel del resultado para funciones reales, el cual vimos en la Proposición 2.5.3. ■

Análogamente, este resultado es válido para las funciones constantes, es decir, $f(z) = z_0$ para cualquier número complejo z_0 .

Proposición 2.6.3. Sea $f(z) = z_0 + z_1 z + \dots + z_n z^n$ una función de variable compleja con z_i números complejos. Entonces $f(z)$ es una función entera.

Demostración. Como la identidad es analítica y las funciones constantes son analíticas, por la Proposición 2.6.1 es claro que $f(z)$ es analítica, pues es la suma y producto de funciones analíticas. ■

2.6.1 Las ecuaciones de Cauchy-Riemann

La propiedad de ser una función analítica, claramente depende de la función $f(z) = u(x, y) + i v(x, y)$. Por lo cual, resulta útil demostrar que la función es analítica a partir del análisis de las funciones reales $u(x, y)$ y $v(x, y)$. Sin embargo, estas funciones aún dependen de dos variables. Intuitivamente, solamente necesitamos que Δz se aproxime a cero desde los valores reales positivos y desde los valores imaginarios positivos. Más adelante, demostraremos esto formalmente.

Si la función $f(z) = u(x, y) + i v(x, y)$ es diferenciable en el punto $z_0 = x_0 + i y_0$, entonces el límite:

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

Puede ser calculado permitiendo que $\Delta z = \Delta x + i \Delta y$ y se aproxime desde cualquier dirección conveniente en el plano complejo, ya que el límite es único en cualquier dirección. Si Δz se aproxima horizontalmente (a través de los reales), entonces $\Delta z = \Delta x$ y obtenemos:

$$f'(z_0) = \lim_{\Delta x \rightarrow 0} \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0) + i v(x_0 + \Delta x, y_0) - i v(x_0, y_0)}{\Delta x}$$

$$f'(z_0) = \lim_{\Delta x \rightarrow 0} \left[\frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} \right] + i \lim_{\Delta x \rightarrow 0} \left[\frac{v(x_0 + \Delta x, y_0) - v(x_0, y_0)}{\Delta x} \right]$$

Como los límites son las derivadas parciales de u y v respecto a x , entonces:

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0)$$

Análogamente, si Δz se aproxima verticalmente (a través de la parte imaginaria), entonces $\Delta z = i \Delta y$ y además, como u, v no dependen de i , entonces:

$$f'(z_0) = \lim_{\Delta y \rightarrow 0} \left[\frac{u(x_0, y_0 + \Delta y) - u(x_0, y_0)}{i \Delta y} \right] + i \lim_{\Delta y \rightarrow 0} \left[\frac{v(x_0, y_0 + \Delta y) - v(x_0, y_0)}{\Delta y} \right]$$

$$f'(z_0) = \frac{i}{i} \cdot \frac{1}{i} \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} = -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

Por lo tanto, como los límites coinciden, entonces concluimos que:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

Las cuales se conocen como las ecuaciones de Cauchy-Riemann. Por lo tanto, hemos demostrado el siguiente teorema.

Teorema 2.6.1. Una condición necesaria para que una función de variable compleja $f(z) = u(x, y) + i v(x, y)$ sea diferenciable en el punto z_0 es que las ecuaciones de Cauchy-Riemann se satisfagan en el punto z_0 . Consecuentemente, si f es analítica en un conjunto abierto G , entonces las ecuaciones de Cauchy-Riemann se satisfacen para cualquier punto de G .

Teorema 2.6.2. Sea $f(z)$ una función de variable compleja de la forma $f(z) = u(x, y) + i v(x, y)$ definida en algún conjunto abierto G que contiene al punto z_0 . Si las derivadas parciales de u y v existen en G , son continuas en z_0 , y satisfacen las ecuaciones de Cauchy-Riemann en z_0 , entonces f es diferenciable en z_0 . Consecuentemente, si las derivadas parciales son continuas y satisfacen las ecuaciones de Cauchy-Riemann para todos los puntos de G , entonces f es analítica en G .

Demostración. Notemos que el cociente de f en el punto z_0 puede escribirse como sigue:

$$\frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0) + i[v(x_0 + \Delta x, y_0 + \Delta y) - v(x_0, y_0)]}{\Delta x + i \Delta y}$$

Donde $z_0 = x_0 + iy_0$ y $\Delta z = \Delta x + i \Delta y$. Ahora, notemos que:

$$\begin{aligned} &u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0) = \\ &u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0 + \Delta y) + u(x_0, y_0 + \Delta y) - u(x_0, y_0) \end{aligned}$$

Como las derivadas parciales existen en G , por el teorema del valor intermedio 2.5.3, existe un número x^* entre x_0 y $x_0 + \Delta x$ tal que:

$$u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0 + \Delta y) = \Delta x \frac{\partial u}{\partial x}(x^*, y_0 + \Delta y)$$

Aún más, como las derivadas parciales son continuas en el punto (x_0, y_0) , podemos escribir:

$$\frac{\partial u}{\partial x}(x^*, y_0 + \Delta y) = \frac{\partial u}{\partial x}(x_0, y_0) + \epsilon_1$$

Donde $\epsilon \rightarrow 0$ conforme $x^* \rightarrow x_0$ y $\Delta y \rightarrow 0$ (en particular, cuando $\Delta z \rightarrow 0$). Así, obtenemos:

$$u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0 + \Delta y) = \Delta x \left[\frac{\partial u}{\partial x}(x_0, y_0) + \epsilon_1 \right]$$

Repitiendo este procedimiento, concluimos que:

$$\frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{\Delta x \left[\frac{\partial u}{\partial x} + \epsilon_1 + i \frac{\partial v}{\partial x} + i \epsilon_3 \right] + \Delta y \left[\frac{\partial u}{\partial y} + \epsilon_2 + i \frac{\partial v}{\partial y} + i \epsilon_4 \right]}{\Delta x + i \Delta y}$$

Donde cada derivada parcial es evaluada en el punto (x_0, y_0) y donde cada $\epsilon_i \rightarrow 0$ conforme $\Delta z \rightarrow 0$. Ahora, por hipótesis, se satisfacen las ecuaciones de Cauchy-Riemann, lo que nos permite escribir el cociente de manera equivalente como:

$$\frac{\Delta x \left[\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right] + i \Delta y \left[\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right]}{\Delta x + i \Delta y} + \frac{\lambda}{\Delta x + i \Delta y}$$

Donde $\lambda := \Delta x(\epsilon_1 + i \epsilon_3) + \Delta y(\epsilon_2 + i \epsilon_4)$. Ahora, notemos lo siguiente:

$$\begin{aligned} \left| \frac{\lambda}{\Delta x + i \Delta y} \right| &\leq \left| \frac{\Delta x}{\Delta x + i \Delta y} \right| |\epsilon_1 + i \epsilon_3| + \left| \frac{\Delta y}{\Delta x + i \Delta y} \right| |\epsilon_2 + i \epsilon_4| \\ &\leq |\epsilon_1 + i \epsilon_3| + |\epsilon_2 + i \epsilon_4| \end{aligned}$$

Así,

$$\lim_{\Delta z \rightarrow 0} \frac{\lambda}{\Delta x + i \Delta y} \leq \lim_{\Delta z \rightarrow 0} |\epsilon_1 + i \epsilon_3| + |\epsilon_2 + i \epsilon_4| = 0$$

Por lo tanto concluimos que,

$$\lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0)$$

Es decir, $f'(z_0)$ existe, es decir, f es diferenciable en el punto z_0 , lo cual concluye la demostración. ■

Proposición 2.6.4. La función $f(z) = e^{pz}$ es una función entera.

Demostración. Notemos que $e^z = e^x(\cos y + i \sin y)$, por lo tanto, e^z es la suma y producto de funciones continuas. Por el Teorema anterior, es suficiente ver que $f(z)$ satisface las ecuaciones de Cauchy-Riemann. Así,

$$f(z) = e^{px+ipy} = e^{px}(\cos py + i \sin py)$$

$$u(x) = e^{px} \cos py \quad y \quad v(x) = e^{px} \sin py$$

$$\frac{\partial u}{\partial x} = p e^{px} \cos py = \frac{\partial v}{\partial y}$$

$$\frac{\partial u}{\partial y} = -p e^{px} \sin py = -\frac{\partial v}{\partial x}$$

Por lo tanto, se satisfacen las ecuaciones de Cauchy-Riemann. ■

Los siguientes teoremas por motivos de extensión los referimos directamente a [8, pp. 218, 402].

Teorema 2.6.3 (Principio de módulo máximo-mínimo). Una función analítica en un dominio acotado y continua (incluyendo la frontera) alcanza su módulo máximo-mínimo en la frontera.

Teorema 2.6.4 (Fórmula integral de Cauchy). Supongamos que C es una curva cerrada simple y la función $f(z)$ es analítica en una región que contiene a C y su interior. Asumimos que C está orientada en sentido antihorario, entonces para cualquier z_0 dentro de C :

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz$$

Capítulo 3

Metodología

En primer lugar, se presentará de manera selectiva resultados de cada una de las áreas planteadas, incluyendo la motivación y explicación de las demostraciones (que pueden ser bastante extensas). Por lo tanto, se estima que la mayor parte de la investigación se dedicará en abordar estos resultados en la parte introductoria del trabajo:

1. Aproximación diofantina.
2. Teorema de Lindmann.
3. Teorema de Gelfond-Schneider.
4. Teorema de Baker.
5. Ecuaciones diofantinas.

En segundo lugar, se expondrá de forma efectiva todo lo relevante a la conjetura de Collatz para el problema de los ciclos en el siguiente orden:

1. Enunciado de la conjetura.
2. Dificultades de su estudio.
3. Enfoques que han proporcionado mayor visión del problema.
4. Estudio de los ciclos.
5. Resultados preliminares relevantes en los ciclos.

Por último, se presentará con el detalle suficiente el Teorema de Steiner de 1977. Para ello, se expondrá de manera sistemática:

1. La eucación diofantina de los 1-ciclos de la conjetura.
2. Un borrador sencillo con los pasos para completar la demostración.
3. La demostración explicada en términos comprensibles con la teoría estudiada.

Capítulo 4

Teoría trascendental de números

Hasta el momento, hemos desarrollado las herramientas esenciales que nos van a permitir estudiar los números trascendentales. Lo natural, es encontrar propiedades que sean ciertas para todos los números algebraicos, así, si alguna de estas propiedades no se cumplen para algún número, es porque éste no puede ser un número algebraico, es decir, debe ser un número trascendental. Es decir, intentamos determinar la esencia de un número algebraico. Por lo tanto, no es sorpresa que estas propiedades tengan aplicaciones en otras áreas, como en las ecuaciones diofantinas.

4.1 Aproximación diofantina

Históricamente, la primera propiedad encontrada que separa a los números algebraicos de los números trascendentales, es la precisión con la que se pueden aproximar por un número racional.

Teorema 4.1.1 (Teorema de Dirichlet). Dado $\zeta \in \mathbb{R}$, existen enteros p, q tales que:

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

Demostración. Sea $\{x\}$ la parte fraccionaria, $[x\zeta]$ la parte entera y Q un entero suficientemente grande para que $q < Q$. Y consideremos los $Q + 1$ números $0, \{\zeta\}, \{2\zeta\}, \dots, \{(Q - 1)\zeta\}, 1$ que pertenecen a los $\frac{1}{Q}$ intervalos. Por el principio de Dirichlet, existen dos de estos que pertenecen al mismo intervalo. Afirmando que la diferencia tiene la forma $q\zeta - p$, veamos:

$$\{q_2\zeta\} - \{q_1\zeta\} = q_2\zeta - [q_2\zeta] - q_1\zeta + [q_1\zeta] = (q_2 - q_1)\zeta - [q_2\zeta] + [q_1\zeta]$$

Definiendo $q = q_2 - q_1$ y $p = \lfloor q_2 \zeta \rfloor - \lfloor q_1 \zeta \rfloor$ entonces $\{q_2 \zeta\} - \{q_1 \zeta\} = q\zeta - p$. Por lo tanto,

$$|q\zeta - p| < \frac{1}{Q} \Rightarrow \left| \zeta - \frac{p}{q} \right| < \frac{1}{qQ} < \frac{1}{q^2}$$

Para cualquier $Q > q$, lo cual completa la demostración. ■

Lo más natural es preguntarse si es posible obtener una mejor aproximación para ciertos números o hay un límite para ello. Liouville encontró una cota para los números algebraicos que impide obtener una mejor aproximación por los racionales. Y con ello, fabricó un número que se aproximara mejor que cualquier número algebraico, es decir, un número trascendental.

Teorema 4.1.2 (Teorema de Liouville). Sea ζ cualquier número algebraico de grado n , y cualesquiera enteros p, q se tiene que:

$$\left| \zeta - \frac{p}{q} \right| > \frac{K}{q^n}$$

Demostración. Dado que ζ es un número algebraico, entonces existe el mínimo polinomio irreducible $f \in \mathbb{Z}[x]$ tal que $f(\zeta) = 0$ por la Proposición 2.3.1. Ahora, sea $x \in [\zeta - 1, \zeta + 1]$, utilicemos el Teorema 2.5.2 para $f'(x)$ (dado que la derivada de cualquier función polinomial es polinomial y por lo tanto continua y diferenciable por el Teorema 2.5.1 y Corolario 2.5.1) se tiene que:

$$f'(x) < M$$

Observemos que M depende únicamente de ζ , ya que $M = f'(c)$ para alguna $c \in [\zeta - 1, \zeta + 1]$ y f depende de ζ por ser su mínimo polinomio irreducible.

Por otro lado, es claro que $f\left(\frac{p}{q}\right) \neq 0$, pues es suficiente garantizar que $\frac{p}{q}$ no es una raíz. En caso de serlo, podemos buscar una aproximación más cercana de ζ con $\zeta \neq \frac{p_2}{q_2}$ tal que $\frac{p_2}{q_2}$ no sea raíz, como el grado del polinomio es finito, no puede contener infinitas raíces, por lo que es un proceso finito y $\frac{p_2}{q_2} \approx \frac{p}{q}$ por lo que el teorema vale para p, q originales. Así,

$$\left| f\left(\frac{p}{q}\right) \right| = \sum_{i=0}^n \left| a_i \left(\frac{p}{q}\right)^i \right| = \frac{1}{q^n} \left| \sum_{i=0}^n a_i p^i q^{n-i} \right|$$

Afirmo que $\left| \sum_{i=0}^n a_i p^i q^{n-i} \right| > 1$, como es la suma de enteros, es suficiente ver que

la suma es distinta de 0. Supongamos lo contrario, entonces $\left| f\left(\frac{p}{q}\right) \right| = 0$ lo

cual es una contradicción por lo anterior. Por lo tanto,

$$\left| f\left(\frac{p}{q}\right) \right| = \frac{1}{q^n} \left| \sum_{i=0}^n a_i p^i q^{n-i} \right| > \frac{1}{q^n}$$

Ahora, consideremos el intervalo $(\frac{p}{q}, \zeta)$, por el Teorema 2.5.3:

$$f(\zeta) - f\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \zeta\right) f'(x)$$

Donde $x \in (\zeta, \frac{p}{q}) \subset [\zeta - 1, \zeta + 1]$, por lo tanto,

$$\left| \frac{p}{q} - \zeta \right| = \frac{\left| f\left(\frac{p}{q}\right) \right|}{|f'(x)|} > \frac{1}{Mq^n} = \frac{K}{q^n}$$

■

Observación. Baker propone $M = n^2(1 + |\zeta|)^{n-1}H$ donde $H = \max\{|a_i|\}$, es decir, el máximo de los valores absolutos de los coeficientes de f . Lo cual, es muy claro si consideramos que M solo debe satisfacer $f'(x) < M$ para $x \in (\zeta - 1, \zeta + 1)$. Como $\{|a_i|\} \leq H$, $i \leq n$ y $x < 1 + |\zeta|$ entonces:

$$f'(x) = \sum_{i=1}^n i(a_i)x^{i-1} < \sum_{i=1}^n nH(1 + |\zeta|)^{n-1} = n^2(1 + |\zeta|)^{n-1}H$$

Por lo cual esta cota no es particularmente buena, por mencionar una mejor, directamente de lo anterior:

$$M = \frac{n(n+1)(1 + |\zeta|)^{n-1}H}{2}$$

Aunque es posible reemplazar 1 por $\frac{p}{q}$ para mejorar la cota, la intención es proporcionar una constante independiente de p y q .

Trascendencia a partir del Teorema de Louville: A partir de lo anterior, Louville se dio cuenta que esta propiedad no es universal para los números trascendentes, puesto que se pueden generar números con una mejor aproximación por racionales. Consideremos

$$\zeta = \sum_{n=1}^{\infty} \frac{1}{10^{n!}}$$

Definamos:

$$p_j = 10^{j!} \sum_{n=1}^j \frac{1}{10^{n!}}$$

Entonces,

$$\begin{aligned} \left| \zeta - \frac{p_j}{10^{j!}} \right| &= \left| \sum_{n=1}^{\infty} \frac{1}{10^{n!}} - \sum_{n=1}^j \frac{1}{10^{n!}} \right| = \left| \sum_{n=j+1}^{\infty} \frac{1}{10^{n!}} \right| \\ &= \frac{1}{10^{(j+1)!}} + \frac{1}{10^{(j+2)!}} + \dots = \frac{1}{10^{(j+1)!}} \left(1 + \frac{1}{10^{(j+2)}} + \frac{1}{10^{(j+3)}} + \dots \right) \\ &< \frac{1}{10^{(j+1)!}} \left(1 + \frac{1}{10} + \frac{1}{10^2} + \dots \right) = \frac{1}{10^{(j+1)!}} \left(\frac{10}{9} \right) = \left(\frac{10}{9} \right) \frac{1}{10^{(j+1)(j!)}} < \frac{1}{(10^{j!})^j} \end{aligned}$$

Tomando $j > n$ donde n es el grado de ζ , contradice el teorema de Louville, por lo tanto, ζ no es algebraico, por lo tanto, es trascendental.

4.2 Teorema de Lindemann

Teorema 4.2.1 (Teorema de Lindemann). Para cualesquiera números algebraicos distintos $\alpha_1, \dots, \alpha_n$ y cualesquiera números algebraicos no nulos $\beta_1, \dots, \beta_n$ se tiene que:

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

Demostración. Supongamos lo contrario, entonces existe una combinación lineal tal que:

$$\beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n} = 0$$

Sea K la extensión normal que contiene las raíces $\beta_1, \dots, \beta_n, \alpha_1, \dots, \alpha_n$. Consideremos

$$\prod_{\sigma \in G(K, \mathbb{Q})} [\sigma(\beta_1) e^{\sigma(\alpha_1)} + \dots + \sigma(\beta_n) e^{\sigma(\alpha_n)}] = 0$$

El producto tiene como resultado 0, porque $1_K \in G(K, \mathbb{Q})$, es decir, $(\beta_1 e^{\alpha_1}) + \dots + \beta_n e^{\alpha_n}$ está en el producto. Supongamos que $G(K, \mathbb{Q}) = \{\sigma_1, \dots, \sigma_d\}$. Expandiendo el producto se obtiene:

$$\sum \sigma_1(\beta_{1*}) \cdots \sigma_d(\beta_{n*}) e^{\sigma_1(\alpha_{1*}) + \dots + \sigma_d(\alpha_{n*})} = 0$$

Donde $1*, \dots, n*$ son en realidad todas las posibles combinaciones de $1, \dots, n$. Por ejemplo, $\sigma_1(\beta_1) \cdots \sigma_d(\beta_1)$ está en la suma. Es decir, es la suma de n^n ecuaciones. Formalmente, lo escribimos de la siguiente manera:

$$\sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \sigma_1(\beta_{i_1}) \cdots \sigma_d(\beta_{i_n}) \cdot e^{\sigma_1(\alpha_{i_1}) + \dots + \sigma_d(\alpha_{i_n})} = 0$$

Los exponentes $\sigma_1(a_{i_1}) + \dots + \sigma_d(a_{i_n})$ pueden tener el mismo valor. Agrupamos los términos con exponentes iguales. Sean $\gamma_1, \dots, \gamma_s$ los distintos valores de $\sigma_1(a_{i_1}) + \dots + \sigma_d(a_{i_n})$ y sea J_k el conjunto de vectores $(i_1, \dots, i_n)$ tal que

$$\sigma(a_{i_1}) + \dots + \sigma(a_{i_n}) = \gamma_k$$

Por lo tanto, la sumatoria se convierte en:

$$\sum_{k=1}^s \delta_k e^{\gamma_k} = 0 \quad \text{donde } \delta_k = \sum_{(i_1, \dots, i_n) \in J_k} \sigma_1(\beta_{i_1}) \cdots \sigma_d(\beta_{i_n})$$

Notemos que cada $\sigma \in G(K, \mathbb{Q})$ permuta los pares $(\gamma_1, \delta_1), \dots, (\gamma_s, \delta_s)$. Veremos que es imposible que todos los δ_k sean 0 mostrando que existe un vector tal que $\sigma(\alpha_{i_1}) + \dots + \sigma(\alpha_{i_n})$ es distinto de todos los demás exponentes. Así, para alguna k , el conjunto J_k tiene cardinalidad 1 y $\delta_k \neq 0$.

Definamos un orden total en $\mathbb{C}$ definiendo $\zeta < \omega$ si $\text{Re } \zeta < \text{Re } \omega$ o bien, $\text{Re } \zeta = \text{Re } \omega$ y $\text{Im } \zeta < \text{Im } \omega$. Este orden total tiene la propiedad que si ζ_i, ω_i son números complejos con $\zeta_i < \omega_i$ para $i = 1, \dots, r$, entonces

$$\sum_{j=1}^r \zeta_j < \sum_{j=1}^r \omega_j$$

Como $\alpha_1, \dots, \alpha_n$ son distintos por hipótesis, entonces $\sigma(\alpha_1), \dots, \sigma(\alpha_n)$ son distintos, de lo contrario, $\sigma(\alpha_i) = \sigma(\alpha_j)$, por ser un grupo, existe σ^{-1} tal que

$$\sigma^{-1}\sigma(\alpha_i) = \sigma^{-1}\sigma(\alpha_j) \Rightarrow \alpha_i = \alpha_j$$

Lo cual es una contradicción. Entonces para cada $k \in \{1, \dots, d\}$ existe un índice i_k tal que $\sigma_k(\alpha_{i_k}) > \sigma_k(\alpha_j)$ para $j \neq i_k$. Esto implica que

$$\sigma_1(\alpha_{i_1}) + \dots + \sigma_d(\alpha_{i_n}) > \sigma_1(\alpha_{j_1}) + \dots + \sigma_d(\alpha_{j_n})$$

Para todos los vectores $(j_1, \dots, j_n) \neq (i_1, \dots, i_n)$ y así, $\sigma_1(\alpha_{i_1}) + \dots + \sigma_d(\alpha_{i_n})$ es distinto del resto de los exponentes. Y claramente $\delta_k \neq 0$, de lo contrario, algún $\sigma_j(\beta_i) = 0$ y repetimos la demostración del Lema 2.3.5. En conclusión, asumamos sin pérdida de generalidad que $\delta_1, \dots, \delta_t$ son los números no nulos entre $\delta_1, \dots, \delta_s$. Entonces,

$$\sum_{k=1}^t \delta_k e^{\gamma_k} = 0$$

Donde cada δ_k es distinto de 0 y cada γ_k son distintos números algebraicos tales que σ permuta los pares $(\gamma_1, \delta_1), \dots, (\gamma_t, \delta_t)$. Utilizando el Lema 2.3.2, podemos

encontrar un entero m tal que $m\delta_k$ es un entero algebraico para toda k . Por lo tanto, definamos $\mu_k = m\delta_k$, así,

$$\sum_{k=1}^t \mu_k e^{\gamma_k}$$

Donde cada μ_k es un entero algebraico distinto de 0 y cada γ_k son números algebraicos distintos tales que σ permuta los pares $(\gamma_1, \mu_1), \dots, (\gamma_t, \mu_t)$. Sea p un número primo y utilizando el Lema 2.3.2, podemos encontrar un entero l tal que $l\gamma_k$ es un entero algebraico para toda k . Para $k = 1, \dots, t$ definimos

$$f_k(x) := \frac{1}{(p-1)!} \cdot l^{tp}(x - \gamma_k)^{p-1} \prod_{\substack{j=1 \\ j \neq k}}^t (x - \gamma_j)^p$$

$$f_k(z) := \int_0^z e^{z-u} f_k(u) du$$

$$M_k := \mu_1 F_k(\gamma_1) + \dots + \mu_t F_k(\gamma_t)$$

Mostraremos que se cumple:

1. Para cada $\sigma \in G(K, \mathbb{Q})$, se tiene que $\sigma(M_1) \in \{M_1, \dots, M_t\}$
2. Si p es suficientemente grande, entonces M_1 es un entero algebraico distinto de 0.
3. $|M_k| < 1$ para $k = 1, \dots, t$ para p suficientemente grande.

Como veremos, estos tres puntos serán suficientes para demostrar el Teorema de Lindemann de manera eficaz. Para ello, necesitaremos los siguientes lemas.

Lema 4.2.1. Supongamos que f tiene grado $m - 1$. Entonces

$$F_k(z) = e^z \left(\sum_{j=0}^m f_k^{(j)}(0) \right) - \sum_{j=0}^m f_k^{(j)}(z)$$

Demostración. Definimos $t = uz^{-1}$, por lo tanto, $u = zt$ y así, $\frac{du}{dt} = z$, entonces $du = zdt$. Notemos que cuando $u = 0$ entonces $t = 0$ y cuando $u = z$ entonces $t = 1$. Sustituyendo:

$$F_k(z) = \int_0^1 e^{z(1-t)} f_k(zt) z dt$$

Integramos por partes utilizando la fórmula

$$\int u^{(0)} v^{(n)} dt = \sum_{k=0}^{n-1} (-1)^k u^{(k)} v^{(n-1-k)} + (-1)^n \int u^{(n)} v^{(0)} dt$$

Definimos $u^{(0)} = f_k(tz)$, $v^{(m)} = ze^{z(1-t)} dt$, notemos que

$$v^{(n-1)} = \int e^{z(1-t)} z dt = -e^{z(1-t)} \Rightarrow v^{n-1-j} = (-1)^{j+1} e^{z(1-t)}$$

$$u^{(j)} = f_k^{(j)}(tz) z^j \quad y \quad u^{(m)} = 0$$

Por lo tanto,

$$\begin{aligned} \int_0^1 e^{z(1-t)} f_k(tz) z dt &= \sum_{j=0}^m (-1)^j f_k^{(j)}(tz) z^j (-1)^{j+1} e^{z(1-t)} \Big|_0^1 \\ &= e^z \sum_{j=0}^m f_k^{(j)}(0) - \sum_{j=0}^m f_k^{(j)}(z) \end{aligned}$$

■

Lema 4.2.2. Se cumple:

$$(i) \quad M_k = - \sum_{j=1}^t \sum_{m=0}^{p(t+1)} \mu_j f_k^{(m)}(\gamma_j) \quad \text{para } k = 1, \dots, t.$$

(ii) Para cada $\sigma \in G(K, \mathbb{Q})$, tenemos que $\sigma(M_1) \in \{M_1, \dots, M_t\}$

Demostración. (i) Como f_k tiene grado $p(t+1) - 1$, entonces por el lema anterior

$$M_k = \sum_{j=1}^t \mu_j \left[e^{\gamma_j} \sum_{m=0}^{p(t+1)} f_k^{(m)}(0) - \sum_{m=0}^{p(t+1)} f_k^{(m)}(\gamma_j) \right]$$

Sin embargo, por hipótesis, $\sum_{j=1}^t \mu_j e^{\gamma_j} = 0$, así,

$$M_k = - \sum_{j=1}^t \sum_{m=0}^{p(t+1)} \mu_j f_k^{(m)}(\gamma_j)$$

(ii) Por el lema 2.2.2, existe $\sigma_i \in G(K, \mathbb{Q})$ tal que

$$(\sigma_i(\gamma_k), \sigma_i(\mu_k)) = (\gamma_i, \mu_i) \quad \text{para } i = 1, \dots, t$$

Así, aplicando σ_i a los coeficientes de f_1 obtenemos

$$\begin{aligned} f_{\sigma(1)}(x) &= \frac{1}{(p-1)!} l^{tp} (x - \sigma_i(\gamma_1))^{p-1} \prod_{\substack{j=1 \\ j \neq \sigma(1)}}^t (x - \sigma_i(\gamma_j))^p \\ &= \frac{1}{(p-1)!} l^{tp} (x - \gamma_i)^{p-1} \prod_{\substack{j=1 \\ j \neq i}}^t (x - \gamma_j)^p = f_i(x) \end{aligned}$$

Con lo cual, concluimos que

$$\begin{aligned} \sigma_i(M_1) &= - \sum_{j=1}^t \sum_{m=0}^{p(t+1)} \sigma(\mu_j) f_{\sigma(1)}^{(m)}(\sigma(\gamma_j)) \\ &= - \sum_{j=1}^t \sum_{m=0}^{p(t+1)} \sigma(\mu_j) f_i^{(m)}(\sigma(\gamma_j)) = M_i \end{aligned}$$

■

Lema 4.2.3. Sea $k \in \{1, \dots, t\}$. Entonces se cumple:

- (i) $f_1^{(p-1)}(\gamma_1) = l^{tp} \left(\prod_{k=2}^t (\gamma_1 - \gamma_k) \right)^p$
- (ii) $f_1^{(j)}(\gamma_m) = 0$ para $m = 1, \dots, t$, $j = 0, \dots, p-1$ y $(m, j) \neq (1, p-1)$.
- (iii) $f_1^{(j)}(\gamma_m) \equiv 0 \pmod{p}$ para $m = 1, \dots, t$ y $j \geq p$.

Demostración. En general, si $g(x) = (x-a)^r h$, entonces $g^{(1)}(x) = r(x-a)^{r-1} h + h'(x-a)^r$ y $g^{(1)}(a) = 0$, es claro que $g^{(m)}(a) = 0$ para $m = 0, \dots, r-1$ y $g^{(r)} = r!h(a)$, esto demuestra (i) y (ii). Ahora, sea $m \in \{1, \dots, t\}$ y $j \geq p$, definimos:

$$g(x) := f_1\left(\frac{x}{l}\right) = \frac{1}{(p-1)!} \cdot l(x - l\gamma_1)^{p-1} \prod_{k=2}^t (x - l\gamma_k)^p$$

Entonces $(p-1)!g$ es un polinomio con coeficientes enteros algebraicos, pues escogimos l tal que $l\gamma_j$ sea un entero algebraico. Notemos que si $g = c_r x^r + \dots + c_0 \in \mathbb{C}[x]$, entonces para todo $j \geq 0$:

$$\frac{1}{j!} g^{(j)}(x) = c_r \binom{r}{j} x^{r-j} + c_{r-1} \binom{r-1}{j} + \dots + c_j$$

Por lo tanto, $(p-1)! \frac{g^{(j)}(x)}{j!}$ es un polinomio con coeficientes enteros algebraicos.

En particular $\frac{g^{(j)}(l\gamma_j)}{p}$ es un entero algebraico. Además, por la regla de la cadena $g^{(1)}(x) = f_1^{(1)}\left(\frac{x}{l}\right) \cdot \frac{1}{l}$ y por lo tanto,

$$\frac{f_1^{(j)}(\gamma_m)}{p} = \frac{l^j g^{(j)}(l\gamma_j)}{p}$$

es un entero algebraico. Esto demuestra (iii). ■

Lema 4.2.4. Para p suficientemente grande, M_1 es un entero algebraico no nulo.

Demostración. Aplicando los Lemas 4.2.2 y 4.2.4, tenemos que:

$$M_1 \equiv -\mu_1 A^p \pmod{p} \quad \text{con} \quad A := l^t \left(\prod_{k=2}^t (\gamma_1 - \gamma_k) \right)$$

Por lo tanto, $\frac{M_1 + \mu_1 A^p}{p}$ es un entero algebraico, por lo tanto, M_1 tiene que ser un entero algebraico y veremos que no puede ser nulo para p suficientemente grande. Asumamos que $\frac{\mu_1 A^p}{p}$ (tomando $M_1 = 0$) es un entero algebraico, y $b = N(\mu_1)$, $B = N(A)$, entonces $b, B \in \mathbb{Z}$ por el Lema 2.3.2, utilizando el mismo lema:

$$N\left(\frac{\mu_1 A^p}{p}\right) = \frac{b B^p}{p^d} \in \mathbb{Z}$$

Lo cual es imposible si $p > |bB|$. Por lo tanto, $M_1 \neq 0$. ■

Lema 4.2.5. Para p suficientemente grande $|M_k| < 1$ para $k = 1, \dots, t$.

Demostración. Notemos que:

$$\begin{aligned} F_k(z) &= \int_0^z e^{z-t} f(tz) z dt \leq \int_0^1 e^{|z|} |f(tz)| |z| dt \\ &\leq |z| e^{|z|} \cdot \sup_{\substack{u \in \mathbb{C} \\ |u| \leq |z|}} |f(u)| \end{aligned}$$

Pues f está acotada por una constante M por el Lema 2.5.2, es decir, se puede acotar por la integral de una constante (que es fácil de calcular). Así, sea $\beta =$

$\max \prod_{\substack{j=1 \\ j \neq k}}^t (|\gamma_k - \gamma_j|)$ para toda k , $\omega = \max (|\gamma_i - \gamma_k|)$ para toda i, k , y $\mu = \max \mu_i$

para toda i , entonces, se cumple que para toda k, i :

$$F_k(\gamma_i) \leq |\gamma_i| e^{|\gamma_i|} \cdot \frac{\omega^{p-1} \beta^p l^{tp}}{(p-1)!}$$

Por lo tanto,

$$|M_k| < \mu \cdot t |\gamma_i| e^{|\gamma_i|} \cdot \frac{\omega^p \beta^p l^{tp}}{(p-1)!} = \frac{c^p}{(p-1)!}$$

Donde c es una constante que depende de $\mu, t, \gamma_i, \omega, \beta$ y l , pero independiente de p . Dado que para todo $c > 1$, $\frac{c^p}{(p-1)!} \rightarrow 0$ conforme $p \rightarrow \infty$, entonces $|M_k| < 1$ para toda k para p suficientemente grande. ■

Finalmente, hemos completado los tres puntos propuestos. Por lo tanto, por los Lemas 4.2.2, 4.2.4, 4.2.5

$$|N(M_1)| = \left| \prod_{\sigma \in G(K, \mathbb{Q})} \sigma(M_1) \right| = |M_1 \cdots M_t| = |M_1| \cdots |M_t| < 1$$

Por el Lema 2.3.4, tenemos que $N(M_1)$ es un entero con la propiedad de $N(M_1) < 1$. Por lo tanto, $N(M_1) = 0$, por el Lema 2.3.5, $M_1 = 0$, lo cual contradice el Lema 4.2.4. Así, hemos producido un entero $N(M_1)$ que depende de $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ tal que $0 < N(M_1) < 1$, lo cual es imposible. Con esto finaliza la demostración del Teorema de Lindemann. ■

Como veremos, este Teorema es en realidad una herramienta útil para generar números trascendentales. Veamos algunas consecuencias.

Consecuencias del Teorema de Lindemann

Corolario 4.2.1. Sea α un número algebraico no nulo. Entonces $\beta = e^\alpha$ es un número trascendental.

Demostración: Supongamos lo contrario, entonces β es algebraico, por lo tanto, $e^\alpha - \beta e^0 = 0$, lo cual contradice el Teorema de Lindemann. ■

Corolario 4.2.2. π es trascendental.

Demostración: Supongamos lo contrario, entonces πi es algebraico. Como $e^{\pi i} = -1$ entonces $e^{\pi i} + e^0 = 0$, lo cual contradice el Teorema de Lindemann, por lo tanto, π es trascendental. (Es claro que este argumento se puede extender para demostrar que πi es trascendental.) ■

Corolario 4.2.3. Sea α un número algebraico distinto de 0, entonces $\sin \alpha$, $\cos \alpha$ y $\tan \alpha$ son trascendentales.

Demostración: Supongamos que $\sin \alpha$ es algebraico. Como $\cos \alpha = \sqrt{1 - \sin^2 \alpha}$ entonces $\cos \alpha$ es algebraico. Como $i\alpha$ es algebraico distinto de 0, entonces $e^{i\alpha} = \cos \alpha + i \sin \alpha$ es trascendental por el Corolario 4.2.1. Lo cual es una contradicción por ser una suma y producto de números algebraicos. Análogamente para $\cos \alpha$ y $\tan \alpha$, ya que pueden probarse con la identidad $1 + \tan^2 \alpha = \sec^2 \alpha$. ■

Corolario 4.2.4. Sea α un número algebraico distinto de 0, 1, entonces $\ln \alpha$ es trascendental.

Demostración: Supongamos que $\ln \alpha$ es algebraico. Entonces $e^{\ln \alpha} = \alpha$ es trascendental, lo cual es una contradicción por el Corolario 4.2.1. ■

Definición. Decimos que los números $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ son algebraicamente independientes si no existe un polinomio (no nulo) $f \in \overline{\mathbb{Q}}[x_1, \dots, x_n]$ tal que $f(\alpha_1, \dots, \alpha_n) = 0$.

Observación. Supongamos que α_1 es algebraicamente independiente, entonces es trascendental.

Definición. Dado un subconjunto S de $\mathbb{C}$, definimos el **grado de trascendencia** de S como $\text{trdeg } S$ como el máximo número t tal que S contiene t elementos algebraicamente independientes.

Corolario 5: Supongamos que $\alpha_1, \dots, \alpha_n$ son linealmente independientes sobre $\mathbb{Q}$. Entonces $e^{\alpha_1}, \dots, e^{\alpha_n}$ son algebraicamente independientes.

Demostración: Sea f cualquier polinomio no nulo en $\overline{\mathbb{Q}}[x_1, \dots, x_n]$. Entonces:

$$f(x_1, \dots, x_n) = \sum_{i_j \in I} \beta_{i_1} \cdots \beta_{i_n} x_1^{i_1} \cdots x_n^{i_n}$$

para $\beta_{i_1} \cdots \beta_{i_n} \in \overline{\mathbb{Q}}[x_1, \dots, x_n]$ y I un conjunto no negativo. Por lo tanto,

$$f(e^{\alpha_1}, \dots, e^{\alpha_n}) = \sum_{i_j \in I} \beta_{i_1} \cdots \beta_{i_n} e^{\alpha_1 i_1 + \dots + \alpha_n i_n}$$

Como $\alpha_1, \dots, \alpha_n$ son linealmente independientes sobre $\mathbb{Q}$, entonces los exponentes $\alpha_1 i_1 + \dots + \alpha_n i_n$ son distintos. Supongamos lo contrario, entonces $\alpha_1 i_1 + \dots + \alpha_n i_n = \alpha_1 j_1 + \dots + \alpha_n j_n$, entonces $\alpha_1(i_1 - j_1) + \dots + \alpha_n(i_n - j_n) = 0$ lo cual contradice que $\alpha_1, \dots, \alpha_n$ sean linealmente independientes sobre $\mathbb{Q}$ o bien $i_k = j_k$. Por lo tanto, por el Teorema de Lindemann $f(e^{\alpha_1}, \dots, e^{\alpha_n}) \neq 0$, por lo tanto, $e^{\alpha_1}, \dots, e^{\alpha_n}$ son algebraicamente independientes o bien, tiene grado de trascendencia n . ■

El siguiente corolario no fue encontrado en ningún artículo relacionado al tema. Reconozco que el resultado podría ser de la autoría de alguien más, pero lo incluyo como resultado de mis observaciones personales.

Corolario 4.2.5. Dadas las condiciones del Teorema de Lindemann, dado $\zeta = \beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n}$, entonces ζ es un número trascendental.

Demostración: Supongamos lo contrario, entonces ζ es algebraico. Podemos suponer que $\alpha_i \neq 0 \forall i$, de lo contrario, consideremos $\zeta_2 = \zeta - \beta_i e^0$ como nuestro nuevo $\zeta = \beta_1 e^{\alpha_1} + \dots + \beta_n e^{\alpha_n}$, el cual es la suma de números algebraicos, por lo tanto, es un número algebraico tal que ninguno de sus $\alpha_j = 0$ para toda j . Así,

$$\zeta e^0 - \beta_1 e^{\alpha_1} - \dots - \beta_n e^{\alpha_n} = 0$$

En cualquier caso, hemos generado un número ζ que contradice el Teorema de Lindemann. Por lo tanto, ζ es un número trascendental. ■

4.3 Teorema de Gelfond-Schneider

En el siglo 20, había un gran interés por expandir nuestro conocimiento acerca de los números trascendentales. Motivado por la lista de 23 problemas de David Hilbert [6], el 7mo problema fue resuelto primero por Gelfond, seguido por Schneider de manera independiente, por lo cual, ahora se conoce como el Teorema de Gelfond-Schneider. Como veremos, es en realidad, una máquina para producir números trascendentales.

Teorema 4.3.1 (Teorema de Gelfond-Schneider). Sean α y β números algebraicos tal que $a \neq 0$, $\alpha \neq 1$ y β no es un número racional real, entonces cualquier valor α^β es trascendental.

Demostración. En general, $\alpha^\beta = e^{\beta \log(\alpha)}$ donde $\log(\alpha)$ es en realidad una función multivaluada en los números complejos. Asumamos que α y β satisfacen las hipótesis del problema y definimos $\gamma = e^{\beta \log(\alpha)}$. Sea K una extensión

normal sobre los racionales tal que $[K : \mathbb{Q}] = h$ para α, β y γ . Establecemos las siguientes relaciones:

$$m = 2h + 3, \quad q > 4m^2, \quad n = \frac{q^2}{2m}$$

$$t = q^2 = 2mn, \quad n > q$$

Por supuesto, estas variables serán complementadas por requisitos futuros, pues buscamos que excedan ciertas constantes. Definimos $p_1, \dots, p_t$ como los números

$$(r + k\beta) \log(\alpha) \quad \text{para } r = 1, \dots, q \quad k = 1, \dots, q \quad (4.1)$$

Definimos la función entera (la demostración es análoga a la realizada en la Proposición 2.6.4)

$$F(z) = \sum_{j=1}^t \mu_j e^{(z p_j)} \quad (4.2)$$

Donde los μ_j son enteros algebraicos en el campo K que serán especificados ahora. Por el Lema 2.3.2, podemos encontrar un entero c_1 tal que $c_1\alpha, c_1\beta$ y $c_1\gamma$ sean enteros algebraicos. Consideremos las mn ecuaciones en las $2mn$ incógnitas μ_j :

$$c_1^{n+2mq} (\log(\alpha))^{-a} F^{(a)}(b) = 0 \quad (4.3)$$

$$a = 0, \dots, n-1 \quad \text{y} \quad b = 1, 2, \dots, m$$

Veamos que se cumplen las condiciones del Lema 2.4.2. Los coeficientes que acompañan a μ_j son:

$$\begin{aligned} c_1^{n+2mq} (\log \alpha)^{-a} p_j^a e^{b p_j} &= c_1^{n+2mq} (r + k\beta)^a e^{b(r+k\beta) \log(\alpha)} \\ &= c_1^{n+2mq} (r + k\beta)^a e^{br \log(\alpha)} e^{k\beta b \log(\alpha)} \\ &= c_1^{n+2mq} (r + k\beta)^a \alpha^{br} \gamma^{kb} \end{aligned} \quad (4.4)$$

Notemos que $a + rb + kb \leq n-1 + qm + qm = n-1 + 2mq$, por lo tanto, podemos garantizar que cada coeficiente es un entero algebraico al multiplicarse por c_1 . Para encontrar una cota para los coeficientes y sus conjugados observemos que:

$$\begin{aligned} |r + k\beta| &\leq \|r\| + \|k\| \cdot \|\beta\| \\ q + q\|\beta\| &= q(1 + \|\beta\|) \end{aligned}$$

Definimos c_2 como el máximo de $\|\alpha\|, \|\gamma\|, 1 + \|\beta\|$ y utilizando $q^2 = 2mn$, podemos afirmar que una cota superior para cualquier coeficiente (o conjugado) de la ecuación (4.4) es:

$$c_1^{n+2mq} (qc_2)^n c_2^{2mq} = (c_1 c_2)^n (c_1 c_2)^{2mq} (\sqrt{2m})^n n^{\frac{n}{2}}$$

Definimos $c_3 = (c_1 c_2)^{2m+1} \sqrt{2m}$ y como $q < n$ entonces puede simplificarse como:

$$\begin{aligned} &\leq (c_1 c_2)^n (c_1 c_2)^{2mn} (\sqrt{2m})^n n^{\frac{n}{2}} \\ &= (c_1 c_2)^{n(2m+1)} (\sqrt{2m})^n n^{\frac{n}{2}} \\ &= [(c_1 c_2)^{2m+1} \sqrt{2m}]^n n^{\frac{n}{2}} \\ &= c_3^n n^{\frac{n}{2}} \end{aligned}$$

Como observación, notemos que c_3 , al igual que c_1 , c_2 y m , son independientes de n . Así, podemos aplicar el Lema 2.4.1 para el sistema de ecuaciones (4.3) con A reemplazado por $c_3^n n^{\frac{n}{2}}$ y concluimos que estas ecuaciones tienen una solución no trivial μ_j , tal que, para cada j ,

$$\begin{aligned} \|\mu_j\| &< c + c[c(2mn)c_3^n n^{\frac{n}{2}}]^{\frac{mn}{2mn-mn}} \\ &= c + 2c^2 mnc_3^n n^{\frac{n}{2}} < 3c^2 mnc_3^n n^{\frac{n}{2}} \end{aligned}$$

Donde la constante c es dependiente del campo K e independiente de n . Como $2^n > n > q > m$, por lo cual podemos reemplazar $4^n = 2^n 2^n > mn$. Por lo tanto, definiendo $c_4 = \sqrt[n]{3c^2 4c_3}$

$$\|\mu_j\| < c_4^n n^{\frac{n}{2}} \quad (4.5)$$

Donde c_4 es independiente de n . De tal manera, la función $F(z)$ en (4.2) queda completamente definida. Antes de continuar, enunciamos el siguiente lema.

Lema 4.3.1. Existen enteros $p \geq n$ y B en el rango $1 \leq B \leq m$ tal que $F^{(a)}(b) = 0$ para $a = 0, 1, \dots, p-1$ y $b = 1, 2, \dots, m$ y $F^{(p)}(B) \neq 0$.

Demostración. Si existe tal entero p , ciertamente satisface $p \geq n$, debido a la ecuación (4.3) (de lo contrario $F^{(p)}(B) = 0$). Además, podemos garantizar que $F^{(a)}(b) = 0$ porque concluimos que el sistema (4.3) sí tiene solución. Por lo tanto, es suficiente demostrar que $F^{(a)}(1)$ no es cero para todos los valores $a = 0, 1, \dots, t-1$. Asumamos lo contrario, entonces:

$$\sum_{j=1}^t \mu_j p_j^a e^{p_j} = 0, \quad 0 \leq a \leq t-1$$

Sin embargo, μ_j no es cero para toda j por el Lema 2.4.2. Por lo tanto, obtenemos un determinante igual a cero.

$$0 = \det(p_j^a e^{p_j}) = \det(p_j^a) \cdot \prod_{j=1}^t e^{p_j} \Rightarrow 0 = \det(p_j^a)$$

Obtenemos esto porque $e^{p_j} \neq 0$. Notemos que $\det(p_j^a)$ es un determinante de Vandermonde, el cual referimos nuevamente a [10, pp. 217], el cual solo equivale a cero si dos p 's son iguales, digamos, $p_j = p_k$. Entonces

$$(r_i + k_j\beta) \log(\alpha) = (r_a + k_b\beta) \log(\alpha)$$

Por lo tanto, β es un número racional, lo cual contradice las hipótesis del Teorema de Gelfond-Schneider. Lo cual demuestra el lema. ■

Ahora, utilizamos el Lema anterior para definir el valor no nulo ζ

$$\begin{aligned} \zeta &= (\log(\alpha))^{-p} F^{(p)}(B) = \sum_{j=1}^t \mu_j (\log(\alpha))^{-p} p_j^p e^{B p_j} \\ &= \sum_{j=1}^t \mu_j (r + k\beta) \alpha^{B r} \gamma^{B k} \end{aligned} \quad (4.6)$$

Notemos que $c_1^{p+2mq} \zeta$ es un entero algebraico en K por la elección de c_1 y la construcción de ζ . Ahora, veamos algunos lemas.

Lema 4.3.2. Existen constantes positivas C , independientes de n y p tales que:

$$|N(\zeta)| \geq C^{-p}$$

Demostración. Notemos que $q < n \leq p$, por lo tanto,

$$c_1^{p+2mq} < c_1^{p+2mp} = (c_1^{1+2m})^p = c_5^p$$

Al definir $c_5 = c_1^{1+2m}$. A su vez, c_5 es como c_1 y m , por ser independiente de p y n donde $c_5^p \zeta$ es un entero algebraico, por lo tanto, utilizando la observación de la Definición 2.3.7, tenemos que:

$$1 \leq |N(c_5^p \zeta)| = |N(c_5^p) N(\zeta)| = c_5^{ph} |N(\zeta)|$$

Ya que $c_5^p \in K_{G(K, \mathbb{Q})}$ (por estar en $\mathbb{Q}$). Entonces,

$$|N(\zeta)| \geq (c_5^h)^{-p}$$

Lo cual demuestra el lema, ya que h es dependiente del campo e independiente de n y p . ■

Lema 4.3.3. Existe una constante positiva c , independiente de n y p , tales que:

$$\|\zeta\| < c^p p^p$$

Demostración. Por la ecuación (4.6) tenemos que:

$$\|\zeta\| \leq t \max_j \|\mu_j\| \cdot \|r + k\beta\|^p \|\alpha\|^{\beta r} \gamma^{\beta k}$$

Ahora, $q < n \leq p$ y $t = 2mn < 2^n$, para alguna n suficientemente grande. Recordemos que r, k, β está acotados por q, q, m respectivamente. Usando también la ecuación (4.5) obtenemos:

$$\begin{aligned} \|\zeta\| &< 2^n c_4^n n^{\frac{n}{2}} \cdot \|q + q\beta\|^p (\|\alpha\|^m \|\gamma\|^m)^q \\ &\leq [2c_4(1 + \|\beta\|) \|\alpha\|^m \cdot \|\gamma\|^m]^p n^{\frac{n}{2}} q^p \end{aligned}$$

Por definición $n2m = q^2$, entonces:

$$q^p = (\sqrt{2m})^p n^{\frac{p}{2}} \leq (\sqrt{2m})^p p^{\frac{p}{2}}$$

Así, obtenemos que:

$$\|\zeta\| < (2c_4(1 + \|\beta\|) \|\alpha\|^m \cdot \|\gamma\|^m \sqrt{2m})^p p^p = c^p p^p$$

Como se quería. ■

Ahora, por el Lema 4.3.1, la función entera $F(z)$ y $F'(z)$ tienen como factor $(z - b)$ para $b = 1, \dots, m$ (aunque estamos asumiendo que una función entera se escribe como el producto de sus raíces, para ello, véase [8, pp. 278]). Por el Lema 2.2.1, $F(z)$ tiene una raíz múltiple $(z - b)^k$, argumentamos que $k \geq p$. Supongamos lo contrario, entonces $F(z) = (z - b)^k h(z)$ con $k < p$ y $(z - b) \nmid h(z)$, por lo tanto, $F^{(k)}(b) = k!h(b) = 0$, entonces $(z - b) \mid h(z)$, lo cual es una contradicción. Por lo tanto, $f(z)$ tiene raíces en $z = 1, \dots, m$ de multiplicidad al menos p . Consideremos la siguiente función:

$$S(z) = p!F(z) \prod_{b=1}^m (z - b)^{-p} \prod_{\substack{b=1 \\ b \neq B}}^m (B - b)^p \quad (4.7)$$

Por lo tanto, $S(z)$ es un cociente de funciones enteras de la forma $\frac{p(z)}{q(z)}$ con $q(z) \neq 0$ (pues los ceros son cancelables), por lo tanto $S(z)$ es una función entera. Expandiendo $F(z)$ en una serie de potencias por la serie de Taylor en $z - B$, obtenemos que:

$$F(z) = \frac{(z - B)^p F^{(p)}(B)}{p!} + \frac{(z - B)^{p+1} F^{(p+1)}(B)}{(p+1)!} + \dots$$

Sustituyendo en $S(z)$, y tomando $z = B$ obtenemos:

$$S(B) = F^{(p)}(B) \quad \text{y} \quad \zeta = (\log(\alpha))^{-p} F^{(p)}(B) \quad (4.8)$$

Con la última expresión cumpliéndose por (4.6). Por la Fórmula integral de Cauchy, tenemos que:

$$S(B) = \frac{1}{2\pi i} \int_C \frac{S(z)}{z - B} dz$$

Para cualquier curva cerrada C alrededor del punto $z = B$. Tomaremos C el círculo $|z| = \frac{p}{q}$. Esto rodea el punto $z = B$ porque

$$\frac{p}{q} > \frac{p}{2q} \geq \frac{n}{2q} = \frac{q}{4m} > \frac{4m^2}{4m} = m \geq B$$

Ahora, si u es cualquier número complejo, entonces $|e^u| < e^{|u|}$, entonces, para toda z en el círculo $|z| = \frac{p}{q}$ y utilizando la ecuación (4.1) se cumple:

$$\begin{aligned} |e^{zp_j}| &\leq e^{|zp_j|} \\ &\leq e^{\frac{p}{q}(q+q|\beta|)|\log(\alpha)|} = c_6^p \end{aligned}$$

Así, hemos definido $c_6 = e^{(1+|\beta|)\log(\alpha)}$ como todas las demás constantes, independientes de n y p . Uniendo (4.2) y (4.5), para toda z en el círculo $|z| = \frac{p}{q}$ obtenemos:

$$|F(z)| \leq t c_4^n n^{\frac{n}{2}} c_6^p < (2c_4 c_6)^p n^{\frac{n}{2}} \leq c_7^p p^{\frac{p}{2}} \quad (4.9)$$

Donde hemos utilizado $t = 2mn < 2^n \leq 2^p$ para n suficientemente grande y definido $c_7 = 2c_4 c_6$ independiente de n y p . Ahora, para cualquier $b = 1, \dots, m$ obtenemos:

$$\begin{aligned} |z - b| &\geq |z| - |b| \geq \frac{p}{q} - m \geq \frac{p}{q} - \frac{p}{2q} \\ &\geq \frac{2qp - qp}{q2q} = \frac{q(2p - p)}{q2q} = \frac{p}{2q} \Rightarrow \\ |z - b|^{-k} &\leq \left(\frac{2q}{p}\right)^k \quad \forall k \end{aligned} \quad (4.10)$$

Donde hemos aplicado $\frac{p}{2q} > m$. Ahora aplicamos las ecuaciones (4.9), (4.7) y recordamos $q^2 = 2mn$, por lo tanto, concluimos que en el círculo $|z| = \frac{p}{q}$:

$$\begin{aligned} |S(z)| &< p! c_7^p p^{\frac{p}{2}} \left(\frac{2q}{p} \right)^{mp} \prod_{\substack{b=1 \\ b \neq B}}^m |B - b|^p \\ &= (c_7 2^m (2m)^{\frac{m}{2}} \prod_{\substack{b=1 \\ b \neq B}}^m |B - b|)^p p! p^{\frac{p}{2}} \left(\frac{\sqrt{n}}{p} \right)^{mp} \\ &= c_8^p p! p^{\frac{p}{2}} \left(\frac{\sqrt{n}}{p} \right)^{mp} \end{aligned}$$

Como $p! < p^p$ y $\frac{\sqrt{n}}{p} \leq \frac{1}{\sqrt{p}}$ (ya que $n < p$), entonces

$$\begin{aligned} |S(z)| &< c_8^p p^p p^{\frac{p}{2}} p^{-\frac{mp}{2}} \\ &= c_8^p p^{\frac{p(3-m)}{2}} \end{aligned}$$

Para toda z en el círculo $|z| = \frac{p}{q}$. A partir de la ecuación (4.8) obtenemos:

$$|\zeta| \leq |\log(\alpha)|^{-p} |S(B)| = \frac{1}{2\pi} |\log(\alpha)|^{-p} \int_C \frac{S(z)}{z - B} dz$$

Porque $|i^{-1}| = 1$. Y la longitud de la circunferencia de $|z| = \frac{p}{q}$ es $2\pi \frac{p}{q}$. Por lo anterior y la ecuación (4.10), obtenemos:

$$\begin{aligned} |\zeta| &< \frac{1}{2\pi} |\log(\alpha)|^{-p} c_8^p p^{\frac{p(3-m)}{2}} \left(\frac{2q}{p} \right)^1 \int_C |dz| \\ &= \frac{1}{2\pi} |\log(\alpha)|^{-p} c_8^p p^{\frac{p(3-m)}{2}} \left(\frac{2q}{p} \right)^1 2\pi \frac{p}{q} \\ &= |\log(\alpha)|^{-p} \frac{p}{q} c_8^p p^{\frac{p(3-m)}{2}} \frac{2q}{p} \\ &< (2c_8 \log(\alpha)^{-1})^p p^{\frac{p(3-m)}{2}} \\ &= c_9^p p^{\frac{p(3-m)}{2}} \end{aligned}$$

Con este estimado y utilizando el Lema 4.3.3 para sus conjugados obtenemos:

$$|N(\zeta)| < c_9^p p^{\frac{p(3-m)}{2}} (c^p p^p)^{h-1} = (c_9 c^{h-1})^p p^{-p} = c_0^p p^{-p}$$

Donde hemos utilizado que $m = 2h + 3$ y definido $c_0 = c_9 c^{h-1}$ (nuevamente, una constante independiente de n y p). Finalmente, utilizando el Lema 4.3.2, concluimos que:

$$c_0^p p^{-p} > |N(\zeta)| \geq C^{-p} \Rightarrow \\ C_0 c > p$$

Para constantes independientes de n y p . Sin embargo, esto es una contradicción, porque $p \geq n$ y podemos escoger a n arbitrariamente grande. Con esto finaliza la demostración del Teorema de Gelfond-Schneider. ■

Consecuencias del Teorema de Gelfond-Schneider

Corolario 4.3.1. El número $2^{\sqrt{2}}$ es un número trascendental.

Demostración. Claramente $2 \neq 0, 1$ y $\sqrt{2}$ no es un número racional real, con $2, \sqrt{2}$ números algebraicos. Por lo tanto, $2^{\sqrt{2}}$ es un número trascendental por el Teorema de Gelfond-Schneider. ■

Corolario 4.3.2. El número e^π es trascendental.

Demostración. Notemos que:

$$e^\pi = (e^{i\pi})^{-i} = (-1)^{-i}$$

Donde $-1, -i$ son algebraicos con $-1 \neq 0, 1$ y como $-i$ no es un número racional real, por el Teorema de Gelfond-Schneider, $(-1)^{-i} = e^\pi$ es un número trascendental. ■

Corolario 4.3.3. Sea α, β números algebraicos diferentes de 0, 1 tales que $\log(\alpha), \log(\beta)$ son linealmente independientes sobre $\mathbb{Q}$. Entonces para cualesquiera números algebraicos no nulos γ, δ , se cumple:

$$\gamma \log(\alpha) + \delta \log(\beta) \neq 0$$

Demostración. Supongamos lo contrario, entonces:

$$\gamma \log(\alpha) + \delta \log(\beta) = 0$$

Entonces $\log(\alpha) = -\left(\frac{\delta}{\gamma}\right) \log(\beta)$. Entonces $\alpha = \beta^{-\frac{\delta}{\gamma}}$. Por el Teorema de Gelfond-Schneider, esto solo es posible si $a := \frac{\delta}{\gamma} \in \mathbb{Q}$. Pero entonces, $\log(\alpha) + a \log(\beta) = 0$, lo cual contradice nuestra hipótesis original. ■

Proposición 4.3.1. El Corolario 4.3.3 implica el Teorema de Gelfond-Schneider.

Demostración. Supongamos que α^β no es trascendental. Como $\beta \log(\alpha) = \log(\alpha^\beta)$, por lo tanto, α y α^β son linealmente dependientes sobre los números algebraicos, por lo tanto, α y α^β son linealmente dependientes sobre $\mathbb{Q}$, necesariamente $\beta \in \mathbb{Q}$. Sin embargo, esto contradice las hipótesis. ■

A partir de estos últimos dos corolarios, es que Baker intenta extender el Teorema de Gelfond-Schneider. Como veremos más adelante, el Teorema de Gelfond-Schneider es un caso particular del Teorema de Baker.

4.4 Teorema de Baker

Las ideas de Baker revolucionaron la Teoría trascendental de números. A partir de ese momento, fue evidente la necesidad de encontrar una función con ciertas características, para su teorema, es necesaria la construcción de una función compleja multivariable. Su resultado no solamente generaliza los resultados de Lindemann y Gelfond-Schneider, también son clave para las ecuaciones diofantinas.

Mi interpretación de los resultados de Baker, solamente facilitan su lectura. Algunas ideas mías, permiten desarrollar los resultados con mayor facilidad. Por su puesto, el libro original [3] de Baker, tiene todo el crédito al respecto, pues no se mejora ni altera los resultados propuestos.

Teorema 4.4.1 (Teorema de Baker). Sean $\alpha_1, \dots, \alpha_n$ números algebraicos no nulos tales que $\log(\alpha_1), \dots, \log(\alpha_n)$ son linealmente independientes sobre los racionales, entonces $1, \log(\alpha_1), \dots, \log(\alpha_n)$ son linealmente independientes sobre el campo de todos los números algebraicos. Es decir, para cualquier $\beta_0, \dots, \beta_n$ números algebraicos (no todos cero), entonces

$$\beta_0 + \beta_1 \log(\alpha_1) + \dots + \beta_n \log(\alpha_n) \neq 0$$

Demostración. Sea K una extensión normal sobre $\mathbb{Q}$ de grado $d = [K : \mathbb{Q}]$. Procedamos por contradicción, sin pérdida de generalidad, $\beta_n \neq 0$. Por lo tanto, podemos dividir entre $-\beta_n$ y redefinimos los β_j para obtener:

$$\beta_0 + \beta_1 \log(\alpha_1) + \dots + \beta_{n-1} \log(\alpha_{n-1}) = \log(\alpha_n)$$

$$\alpha_n = e^{\beta_0} e^{\beta_1 \log(\alpha_1)} \dots e^{\beta_{n-1} \log(\alpha_{n-1})}$$

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

Ahora, utilizando esta igualdad, produciremos los siguientes lemas.

Lema 4.4.1. Existen enteros $p(\lambda_0, \dots, \lambda_n)$, no todos cero, con valores absolutos a lo más e^{h^3} , tales que la función

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

Donde $\gamma_r = \lambda_r + \lambda_n \beta_r$ ($1 \leq r < n$) y $L = h^{2-\frac{1}{4n}}$ satisface:

$$\Phi^{(m_0, \dots, m_{n-1})}(l, \dots, l) = 0$$

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

Demostración. Claramente, es necesario determinar los enteros $p(\lambda_0, \dots, \lambda_n)$ tales que:

$$\Phi^{(m_0, \dots, m_{n-1})}(l, \dots, l) = 0$$

Para calcular las derivadas parciales, es suficiente utilizar:

$$\frac{d^n(uv)}{dx^n} = \sum_{\mu=0}^{m_0} \binom{m_0}{\mu} \frac{d^{n-k}(u)}{dx^{n-k}} \frac{d^k(v)}{dx^k}$$

$$\frac{d(a^{kx})}{dx} = \frac{d(e^{kx \ln(a)})}{dx} = k \ln(a) e^{kx \ln(a)} = k \ln(a) a^{kx}$$

Por lo tanto,

$$\Phi^{(m_0, \dots, m_{n-1})}(l, \dots, l) = \sum_{\lambda_0=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_0, \dots, \lambda_n) q(\lambda_0, \lambda_n, l) e^{\lambda_n \beta_0 l} \alpha_1^{\gamma_1 l} \cdots \alpha_{n-1}^{\gamma_{n-1} l} h(\gamma_1, \dots, \gamma_{n-1}) = 0$$

Donde

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

$$h(\gamma_1, \dots, \gamma_{n-1}) = \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} \log(\alpha_1)^{m_1} \cdots \log(\alpha_{n-1})^{m_{n-1}}$$

Por supuesto, los logaritmos son constantes por lo cual pueden cancelarse al dividirse (utilizando el hecho de que todo está igualado a cero). Además, notemos que:

$$\begin{aligned} e^{\lambda_n \beta_0 l} \alpha_1^{\gamma_1 l} \cdots \alpha_{n-1}^{\gamma_{n-1} l} &= \\ &= e^{\lambda_n \beta_0 l} \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} \alpha_1^{\lambda_n \beta_1 l} \cdots \alpha_{n-1}^{\lambda_n \beta_{n-1} l} \\ &= \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} \left(e^{\beta_0} \alpha_1^{\beta_1} \cdots \alpha_{n-1}^{\beta_{n-1}} \right)^{\lambda_n l} \\ &= \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} \cdot \alpha_n^{\lambda_n l} \end{aligned} \quad (4.11)$$

Por lo tanto, concluimos que:

$$\Phi^{(m_0, \dots, m_{n-1})}(l, \dots, l) = \sum_{\lambda_0=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_0, \dots, \lambda_n) q(\lambda_0, \lambda_n, l) \alpha_1^{\lambda_1 l} \cdots \alpha_{n-1}^{\lambda_{n-1} l} \cdot \alpha_n^{\lambda_n l} \cdot \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} = 0$$

Para todos los enteros l con $1 \leq l \leq h$ y todos los enteros no negativos $m_0, \dots, m_{n-1}$ con $m_0 + \dots, m_{n-1} \leq h^2$, lógicamente podemos preguntar, ¿cuántos conjuntos

de enteros $(m_0, \dots, m_{n-1})$ hay? En este caso no es necesario una respuesta precisa, tenemos la libertad de tomar la cota máxima h^{2n} . Por lo tanto, se tienen a lo más $p := h(h^{2n})$ ecuaciones con $q := (L+1)^{n+1}$ variables. Veamos que hay más variables que ecuaciones (pues de hecho es requisito para utilizar el Lema 2.4.2, ya que los coeficientes son enteros algebraicos multiplicando por alguna r por el Lema 2.3.2).

$$(L+1)^{n+1} > h^{2-\frac{1}{4n}(n+1)} = h^{2(n+1)-\frac{n+1}{4n}}$$

No es difícil ver que para n entero positivo se cumple que $\frac{n+1}{4n} \leq \frac{1}{2}$, por lo cual,

$$(L+1)^{n+1} > h^{2(n+1)-\frac{1}{2}} = h^{2n+\frac{3}{2}} = h^{2n} h \sqrt{h} > 2h(h^{2n})d(d+1)$$

Para ello, solo necesitamos $h > 4d^2(d+1)^2$. Como se requería. Además,

$$q > 2pd(d+1) \Rightarrow pd(d+1) < q - pd(d+1) \Rightarrow \frac{pd(d+1)}{q - pd(d+1)} < 1 \quad (4.12)$$

Todos los coeficientes que acompañan a los enteros $p(\lambda_0, \dots, \lambda_n)$ no dependen de h con excepción de los $l^{\lambda_0 - \mu_0}$ que lógicamente son menores a h^L . Entonces, podemos afirmar que existe $A = kh^L$ una cota superior para los coeficientes y conjugados en algún campo K . Por el Lema 2.4.3 y la ecuación (4.12), existe una solución no trivial de enteros $p(\lambda_0, \dots, \lambda_n)$ tal que tienen valores absolutos a lo más:

$$((L+1)^{n+1} kh^L)^{\frac{pd(d+1)}{q-pd(d+1)}} < ((L+1)^{n+1} kh^L) < e^{h^3}$$

Pues el lado izquierdo toma un valor menor a h^{kh^2} , con $h = e^k$, entonces $h^{kh^2} = e^{kh^2} < e^{h^3}$ para una h lo suficientemente grande. ■

Lema 4.4.2. Sea $m_0, \dots, m_{n-1}$ enteros no negativos tales que

$$m_0 + \dots + m_{n-1} \leq h^2$$

Y sea $f(z) = \Phi^{(m_0, \dots, m_{n-1})}(z, \dots, z)$. Entonces para cualquier número z , obtenemos:

$$|f(z)| < c^{h^3 + L|z|}$$

Además, para cualquier entero positivo l , se cumple $f(l) = 0$ o bien $|f(l)| > c_2^{-(h^3 + Ll)}$.

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

$$P \sum_{\lambda_0=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_0, \dots, \lambda_n) q(\lambda_0, \lambda_n, z) \alpha_1^{\lambda_1 z} \cdots \alpha_{n-1}^{\lambda_{n-1} z} \cdot \alpha_n^{\lambda_n z} \cdot \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$$

Donde $q(\lambda_0, \lambda_n, z)$ está definida en el Lema 4.4.1 y

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

Así, tenemos que:

$$|q(\lambda_0, \lambda_n, z)| \leq (c_3 L)^{m_0} (|z| + 1)^L \sum_{\mu_0=0}^{m_0} = (2c_3 L)^{m_0} (|z| + 1)^L,$$

$$|\alpha_1^{\lambda_1 z} \cdots \alpha_{n-1}^{\lambda_{n-1} z} \cdot \alpha_n^{\lambda_n z}| \leq (c_4)^{L|z|}, \quad |P \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}| \leq (c_5 L)^{m_1 + \dots + m_{n-1}}$$

Por lo tanto,

$$|f(z)| \leq (c_5 L)^{h^2} (c_4)^{L|z|} (c_3 L)^{m_0} (|z| + 1)^L (e^{h^3}) \leq c^{h^3 + L|z|}$$

Por otro lado, por el Lema 2.3.2, $rf(l)$ es un entero algebraico y de hecho, distinto de cero. Así, como $N(rf(l)) \in \mathbb{Z}$ por el Lema 2.3.2,

$$|N(rf(l))| = |r^d| \prod_{\sigma \in G(K, \mathbb{Q})} |\sigma(f(l))| \geq 1$$

Es fácil ver que los conjugados pueden acotarse por $c_6^{h^3 + Ll}$, por lo tanto,

$$f(l) \geq \frac{1}{r^d c_6^{h^3 + Ll}} \geq c_2^{-(h^3 + Ll)}$$

■

Lema 4.4.3. Sea J cualquier entero que satisface $0 \leq J \leq (8n)^2$. Entonces $f(z) = 0$ para todos los enteros l que satisfacen $1 \leq l \leq h^{1 + \frac{J}{8n}}$ y todos los enteros no negativos $m_0 + \dots + m_{n-1}$ con $m_0 + \dots + m_{n-1} \leq \frac{h^2}{2^J}$.

Demostración. Probaremos el lema por inducción sobre J . Cuando $J = 0$ entonces $h^{1 + \frac{J}{8n}} = h$ y $\frac{h^2}{2^J} = h^2$. Por el Lema 4.4.1, se cumple el caso base. Supongamos que el resultado es válido para $J = k$ y probaremos que es cierto para $J = k + 1$.

Primero, definamos los enteros $m_0, \dots, m_{n-1}$ los enteros no negativos tales que:

$$m_0 + \dots + m_{n-1} \leq \frac{h^2}{2^{k+1}} \leq \frac{h^2}{2^k}$$

Si l es cualquier entero tal que $1 \leq l \leq h^{1+\frac{k}{8n}}$, por hipótesis de inducción satisface $f(l) = 0$. Por lo tanto, podemos asumir que

$$h^{1+\frac{k}{8n}} < l \leq h^{1+\frac{k+1}{8n}}$$

Y procedamos suponiendo que $f(l) \neq 0$ y busquemos una contradicción. Para ello, crearemos una cota inferior y superior de $|f(l)|$ que sea contradictoria. Por el Lema 4.4.2 y utilizando $L \geq h^2$, es claro que

$$|f(l)| \geq c^{-(h^3+Ll)} \geq c_9^{-(h^3+\frac{k}{8n})} \quad (4.13)$$

Tomando $c_9 = c^2$. Para la cota superior, necesitamos construir una función entera. Definamos:

$$R_k = [h^{1+\frac{k}{8n}}], \quad S_k = \left[\frac{h^2}{2^k} \right]$$

Donde $[x]$ denota la parte entera. Por lo tanto, l satisface $R_k < l \leq R_{k+1}$ y $m_0 + \dots + m_{n-1} \leq S_{k+1}$. Sea cualquier entero no negativo $m \leq S_{k+1}$ y escribiendo $m = j_0 + \dots + j_{n-1}$ con j_i no negativo, entonces:

$$(m_0 + j_0) + \dots + (m_{n-1} + j_{n-1}) \leq 2S_{k+1} \leq \frac{h^2}{2^k}$$

Entonces, por hipótesis de inducción, para cualquier entero r con $1 \leq r \leq R_k$ tenemos que:

$$\Phi_{m_0+j_0, \dots, m_{n-1}+j_{n-1}}(r) = 0$$

Con $j_0 + \dots + j_{n-1} = m \leq S_{k+1}$. Ahora, por la fórmula de Leibnitz, tenemos que:

$$f^{(m)}(r) = \sum_{\substack{j_0, \dots, j_{n-1} \\ j_0 + \dots + j_{n-1} = m}} \binom{m}{j_0, \dots, j_{n-1}} \Phi_{m_0+j_0, \dots, m_{n-1}+j_{n-1}}(r)$$

Por lo tanto, la función $f(z)$ tiene ceros en $z = 1, 2, \dots, R_k$ de orden $\geq S_{k+1}$. De manera análoga a la demostración del Teorema de Gelfond-Schneider sea:

$$F(z) = [(z-1)(z-2) \cdots (z-R_k)]^{S_{k+1}}$$

Entonces $\frac{f(z)}{F(z)}$ es una función entera. Así, definimos $R = R_{k+1} h^{\frac{1}{8n}}$, entonces $R > 2R_k$ (basta escoger $h > 2^{8n}$ para verificarlo). Sea D_R el disco cerrado $|z| \leq R$. Por el Teorema 2.6.3 obtenemos:

$$\left| \frac{f(z)}{F(z)} \right| \leq \frac{\max_{|z|=R} |f(z)|}{\min_{|z|=R} |F(z)|} \quad \text{para } z \in D_R$$

Así, para cada $z \in D_r$ obtenemos:

$$|f(z)| \leq \frac{|F(z)|(\max_{|z|=R} |f(z)|)}{\min_{|z|=R} |F(z)|} \quad (4.14)$$

Como $l > R_k$, $F(l) \neq 0$. Además,

$$|F(l)| = (|l-1| \cdots |l-R_k|)^{S_{k+1}} \leq R_{k+1}^{R_k S_{k+1}}$$

Utilizando $|l-j| \leq l \leq R_{k+1}$. Dado que $R > 2R_k \Rightarrow 2R - 2R_k > R \Rightarrow |R - R_k| > \frac{R}{2}$, entonces para $1 \leq j \leq R_k$ se obtiene $|z-j| > \frac{R}{2}$. Por lo tanto,

$$\min_{|z|=R} |F(z)| \geq \left(\frac{R}{2}\right)^{R_k S_{k+1}}$$

Regresando a 4.14, concluimos que:

$$|f(l)| \leq c^{h^3+LR} \left(\frac{2R_{k+1}}{R}\right)^{R_k S_{k+1}}$$

Observemos que

$$R_k S_{k+1} \leq \frac{h^{3+\frac{k}{8n}}}{2^{k+1}}$$

$$c^{h^3+LR} \leq c^{h^3+h^3+\frac{k}{8n}} \leq c_7^{h^3+\frac{k}{8n}}$$

Lo cual es claro si tomamos $c_7 = c^2$. Así, finalmente,

$$|f(l)| \leq c_7^{h^3+\frac{k}{8n}} \left(\frac{2}{h^{\frac{1}{8n}}}\right)^{\frac{h^{3+\frac{k}{8n}}}{2^{k+1}}} \quad (4.15)$$

$$\leq c_8^{h^3+\frac{k}{8n}} (h)^{-\frac{h^{3+\frac{k}{8n}}}{2^{k+4n}}}$$

Con $c_8 = 2c_7$. Así, utilizando las ecuaciones (4.13) y (4.15) obtenemos:

$$\begin{aligned} c_8^{h^3 + \frac{k}{8n}} (h)^{-\frac{h^{3+\frac{k}{8n}}}{2^{k+4n}}} &\geq |f(l)| \geq c_9^{-(h^3 + \frac{k}{8n})} \\ \left(h^{3+\frac{k}{8n}}\right) \ln(c_8) - \frac{h^{3+\frac{k}{8n}}}{n2^{k+4}} \ln(h) &\geq -\left(h^{3+\frac{k}{8n}}\right) \ln(c_9) \\ \left(h^{3+\frac{k}{8n}}\right) [\ln(c_8) - \frac{\ln(h)}{n2^{k+4}}] &\geq \left(h^{3+\frac{k}{8n}}\right) (-\ln(c_9)) \\ n2^{k+4} [\ln(c_8) + \ln(c_9)] &\geq \ln(h) \end{aligned}$$

Lo cual es una contradicción para una h lo suficientemente grande (pues el lado izquierdo contiene constantes independientes de h). La contradicción surge de asumir que $f(l) \neq 0$, así, $f(l) = 0$, lo cual, completa la inducción sobre J . ■

Lema 4.4.4. Escribiendo $\phi(z) = \Phi(z, \dots, z)$, se obtiene:

$$|\phi^{(j)}(0)| < e^{-h^{8n}} \quad (0 \leq j \leq h^{8n}) \quad (4.16)$$

Demostración. Definamos los números $X = h^{8n}$ y $Y = \lfloor \frac{h^2}{2^{(8n)^2}} \rfloor$, por el Lema 4.4.3 tenemos que:

$$\Phi_{m_0, \dots, m_{n-1}}(l) = 0 \quad \text{para } 1 \leq l \leq X$$

Y para todos los enteros no negativos $m_0, \dots, m_{n-1}$ tales que $m_0 + \dots + m_{n-1} \leq Y$. De aquí, se concluye que Φ tiene ceros en $z = 1, 2, \dots, X$ de orden al menos Y . Así, definamos $R = Xh^{\frac{1}{8n}}$, entonces $R > 2X$ ya que $h > 2^{8n}$. Ahora, sea C el círculo $|z| = R$ y D_R el disco cerrado $|z| \leq R$, nuevamente construimos la función

$$F(z) = \frac{\phi(z)}{[(z-1) \cdots (z-X)]^Y}$$

La cual ya hemos visto varias veces que es una función entera. Por el Lema 2.6.3 se tiene que para toda $z \in D_R$ obtenemos:

$$|F(z)| \leq \frac{\max_{|z|=R} |\phi(z)|}{\min_{|z|=R} (|z-1| \cdots |z-X|)^Y}$$

Por lo tanto,

$$\phi(z) \leq \frac{(|z-1| \cdots |z-X|)^Y \max_{|z|=R} |\phi(z)|}{\min_{|z|=R} (|z-1| \cdots |z-X|)^Y}$$

Primero, obtenemos una cota para $\max_{|z|=R} |\phi(z)|$, por el Lema 4.4.2 y utilizando $LR \leq h^{8n+2}$ obtenemos que

$$\max_{|z|=R} |\phi(z)| \leq c^{h^3+LR} \leq c^{h^2(h+h^{8n})} \leq c^{h^2(2h^{8n})} = c^{2h^{8n+2}}$$

Como $XY = \frac{h^{8n+2}}{2^{(8n)^2}}$, entonces

$$\max_{|z|=R} |\phi(z)| \leq c_{10}^{XY} \quad \text{con} \quad c_{10} = c^{\frac{1}{2^{(8n)^2}+1}}$$

Ahora, como $X < \frac{R}{2}$, obtenemos:

$$\min_{|z|=R} (|z-1| \cdots |z-X|)^Y > (R-X)^{XY} > \left(\frac{R}{2}\right)^{XY}$$

Notemos que para todo $|z| \leq X$ tenemos que:

$$(|z-1| \cdots |z-X|)^Y \leq (2X)^{XY}$$

Combinando estas desigualdades se obtiene:

$$|\phi(z)| \leq c_{10}^{XY} \left(\frac{4X}{R}\right)^{XY} = (4c_{10} h^{-\frac{1}{8n}})^{XY} < e^{-XY}$$

Lo cual se puede lograr definiendo $h > (4ec_{10})^{8n}$. Finalmente, para cualquier entero $0 \leq j \leq h^{8n}$ dentro del círculo $C' : |z| = X$, por el Teorema 2.6.4 y lo anterior,

$$\begin{aligned} |\phi^{(j)}(0)| &\leq \frac{j!}{2\pi} \int_{C'} \frac{|\phi(z)|}{|z|^{j+1}} dz \\ &\leq j^j \frac{e^{-XY}}{X^j} \\ &\leq e^{-XY} < e^{-h^{8n}} \end{aligned}$$

Donde $XY = \frac{h^{8n+2}}{2^{(8n)^2}} > h^{8n}$ si $h > 2^{\frac{(8n)^2}{2}}$. Lo cual concluye la demostración. ■

Lema 4.4.5. Para cualesquiera enteros $t_1, \dots, t_n$ (no todos cero) y valores absolutos a lo más T , obtenemos

$$|t_1 \log(\alpha_1) + \dots + t_n \log(\alpha_n)| > c_{11}^{-T}$$

Demostración. Como $\log(\alpha_1), \dots, \log(\alpha_n)$ son $\mathbb{Q}$ -linealmente independientes, entonces

$$t_1 \log(\alpha_1) + \dots + t_n \log(\alpha_n) \neq 0$$

Por lo tanto, definamos

$$\Omega = t_1 \log(\alpha_1) + \dots + t_n \log(\alpha_n)$$

Por el Lema 2.3.2,

$$\omega = r(e^\Omega - 1) = r(\alpha_1^{t_1} \dots \alpha_n^{t_n} - 1) \in \mathbb{A}$$

Por ser el producto de números algebraicos. Ahora, por la demostración del mismo lema, $r \geq 1$. Por lo tanto, si $\omega = 0$, entonces $e^\Omega = 1$, por lo tanto, Ω es un múltiplo de $2\pi i$, y de hecho, uno no nulo, pues $\log(\alpha_1), \dots, \log(\alpha_n)$ son por hipótesis, linealmente independientes sobre los racionales, en este caso, el lema es válido trivialmente. De lo contrario, $N(\omega) \geq 1$, entonces

$$N(\omega) = \prod_{\sigma \in G(K, F)} \sigma(\omega) \geq 1 \Rightarrow |\omega| \geq c_{12}^{-Td^n}$$

Además, para cualquier z , $|e^z - 1| \leq ze^{|z|}$, por lo tanto,

$$c_{12}^{-Td^n} \leq |\omega| \leq |r\Omega|e^{|\Omega|} \leq |r||\Omega|e$$

Cuando $|\Omega| < 1$ (pues el lema es trivialmente válido para $|\Omega| > 1$), por lo tanto, $|\Omega| \geq \frac{c_{12}^{-Td^n}}{|r|e} = c_{11}^{-T}$, como se quería. ■

Lema 4.4.6. Sean R, S , enteros positivos y sean $\gamma_0, \dots, \gamma_{R-1}$ distintos números complejos. Definamos como Γ al máximo entre $1, |\gamma_0|, \dots, |\gamma_{R-1}|$ y ρ como el mínimo entre 1 y $|\gamma_i - \gamma_j|$ con $0 \leq i < j < R$. Entonces, para cualesquiera enteros r, s con $0 \leq r < R$ y $0 \leq s < S$, existen números complejos ω_i ($0 \leq i < RS$) con valores absolutos a lo más $\left(\frac{2\Gamma}{\rho}\right)^{RS}$ tal que el polinomio

$$W(z) = \sum_{j=0}^{RS-1} \omega_j z^j$$

satisface $W^{(j)}(\gamma_i) = 0$ para toda i, j con $0 \leq i < R$ y $0 \leq j < S$ exceptuando $i = r, j = s$ y $W_s(\gamma_r) = 1$.

Demostración. Sea el par r, s fijo. Consideremos

$$W(z) = a_0(z - \gamma_r)^s \prod_{\substack{m=1 \\ m \neq r}}^{R-1} (z - \gamma_m)^S$$

Donde

$$a_0 = \left(s! \prod_{\substack{m=1 \\ m \neq r}}^{R-1} (\gamma_r - \gamma_m)^S \right)^{-1}$$

Entonces, claramente $W(z)$ es un polinomio con coeficientes complejos de grado a lo más $RS - 1$. Ahora, utilicemos la fórmula de Leibnitz,

$$W^{(j)}(z) = a_0 \sum_{k=0}^j \binom{j}{k} [(z - \gamma_r)^S]^{(j-k)} \left[\prod_{\substack{m=1 \\ m \neq r}}^{R-1} (z - \gamma_m)^S \right]^{(k)}$$

Es claro que si $j < S$, entonces el término $(z - \gamma_m)$ para $m \neq r$ sobrevive, pues es de grado al menos S . Por lo tanto, $W^{(j)}(\gamma_i) = 0$. Por otro lado, cuando $j = s, i = r$ se obtiene:

$$W^{(s)}(\gamma_r) = a_0 s! \prod_{\substack{m=1 \\ m \neq r}}^{R-1} (z - \gamma_m)^S = 1$$

Por definición de a_0 . Ahora, vamos a estimar los coeficientes de $W(z)$. Escribiendo a $W(z)$ como sigue:

$$W(z) = c_0 + c_1 z + \dots + c_{RS-1} z^{RS-1}$$

Notemos que $W(z) \leq |a_0|(z + \Gamma)^{RS}$ por definición de Γ , sea u_m el coeficiente de z^m en $(z + \Gamma)^{RS}$, $0 \leq m < RS$, entonces, $|c_m| \leq |a_0|u_m$. Además,

$$u = \binom{RS}{m} \Gamma^m < (\Gamma + 1)^{RS}$$

Esto es obvio porque $(\Gamma + 1)^{RS} = \sum_{k=0}^{RS} \binom{RS}{k} \sigma^{RS-k}$ (es decir, el lado izquierdo es uno de los sumados del lado derecho). Por otro lado,

$$|a_0| = \left(s! \prod_{\substack{m=1 \\ m \neq r}}^{R-1} |\gamma_r - \gamma_m|^S \right)^{-1} < \rho^{-S(R-1)}$$

Por definición de ρ . Así, obtenemos,

$$|c_m| < \frac{(\Gamma + 1)^{RS}}{\rho^{S(R-1)}} = \left(\frac{\Gamma + 1}{\rho} \right)^{RS} \rho^S \leq \left(\frac{2\Gamma}{\rho} \right)^{RS}$$

Porque $\rho \leq 1$. ■

Finalmente, podemos proceder a la demostración del Teorema 4.4.1. Veremos que las desigualdades obtenidas en el Lema 4.4.4 no pueden ser todas válidas y la contradicción demostrará el teorema. Empecemos escribiendo $S = L + 1$ y $R = S^n$. Es fácil demostrar por inducción que para cualquier entero i con $0 \leq i < RS$ se puede representar de manera única de la forma

$$i = \lambda_0 + \lambda_1 S + \dots + \lambda_n S^n \quad 0 \leq \lambda_j \leq L \quad \forall j$$

Para cada i , definimos,

$$v_i = \lambda_0, \quad p_i = p(\lambda_0, \dots, \lambda_n)$$

Y definimos $\varphi_i = \lambda_1 \log(\alpha_1) + \dots + \lambda_n \log(\alpha_n)$. Entonces utilizando la ecuación (4.11) se obtiene:

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{v_i} e^{\varphi_i z}$$

por el Lema 4.4.5, cualesquiera dos φ_i , quienes corresponden a dos conjuntos $\lambda_1, \dots, \lambda_n$ distintos difieren en al menos c_{11}^{-T} . En particular, exactamente $(L + 1)^n = R$ de los φ_i son distintos, denotamos estos valores como $\gamma_0, \dots, \gamma_{R-1}$. Si definimos Γ y ρ como en el Lema 4.4.6, entonces $\Gamma \leq c_{14} L$ y $\rho \geq c_{15}^{-L}$.

Sea t tal que $p_t \neq 0$, $s = v_t$ y r tal que $\varphi_t = \gamma_r$. Definimos $W(z)$ como en el Lema 4.4.6, por las propiedades de $W(z)$ especificadas en el Lema, vemos que

$$p_t = \sum_{i=0}^{RS-1} p_i W^{(v_i)}(\varphi_i)$$

Además, por el Teorema de Leibnitz, tenemos que:

$$W^{(v_i)}(\varphi_i) = \sum_{j=0}^{RS-1} j(j-1) \cdots (j-v_i+1) \omega_j \varphi_i^{j-v_i} = \sum_{j=0}^{RS-1} \omega_j \left[\frac{d^j}{dz^j} (z^{v_i} e^{\varphi_i z}) \right]_{z=0}$$

Entonces

$$\begin{aligned}
 p_i &= \sum_{k=0}^{RS-1} p_k W^{(v_k)}(\varphi_k) \\
 &= \sum_{k=0}^{RS-1} p_k \sum_{j=0}^{RS-1} \omega_j \left[\frac{d^j}{dz^j} (z^{v_k} e^{\varphi_k z}) \Big|_{z=0} \right] \\
 &= \sum_{j=0}^{RS-1} \omega_j \left[\frac{d^j}{dz^j} \left(\sum_{k=0}^{RS-1} p_k z^{v_k} e^{\varphi_k z} \right) \Big|_{z=0} \right] \\
 &= \sum_{j=0}^{RS-1} \omega_j \phi^{(j)}(0)
 \end{aligned}$$

Como $RS \leq h^{2n+2}$, por el Lema 4.4.4, se sigue que se cumple (4.16) para toda j con $0 \leq j \leq RS$. Además, por el Lema 4.4.6 tenemos que:

$$|\omega_j| \leq \left(\frac{2\Gamma}{\rho} \right)^{RS} \leq (2c_{14} L c_{15}^L)^{RS} \leq L^{h^{2n+2}} c_{16}^{h^{2n+4}}$$

Como $|p_t| \geq 1$, concluimos que

$$\begin{aligned}
 1 \leq |p_t| &\leq \sum_{j=0}^{RS-1} |\omega_j| |\phi^{(j)}(0)| \\
 &\leq RSL^{h^{2n+2}} c_{16}^{h^{2n+4}} e^{-h^{8n}} \\
 &\leq h^{2n+2} h^{2h^{2n+2}} c_{16}^{h^{2n+4}} e^{-h^{8n}}
 \end{aligned}$$

Aplicando logaritmos

$$\begin{aligned}
 &[2n+2] \log(h) + [2h^{2n+2}] \log(h) + [h^{2n+4}] \log(c_{16}) \\
 &\leq h^{2n+4} (\log(c_{16}) + \log(h)) \\
 &\leq h^{2n+4} (\log(c_{16}h)) \\
 &\leq h^{2n+5} c_{16} \\
 &\leq h^{8n}
 \end{aligned}$$

Lo cual podemos garantizar cuando $h > c_{16}^{(6n+5)^{-1}}$. Por lo tanto,

$$1 \leq h^{2n+2} h^{2h^{2n+2}} c_{16}^{h^{2n+4}} e^{-h^{8n}} < 1$$

Lo cual es una contradicción. Esto finaliza el Teorema de Baker. ■

Capítulo 5

Ecuaciones diofantinas

Una de las aplicaciones más sorprendentes de la teoría que desarrolló Baker fue permitir encontrar una cota para el tamaño de todas las soluciones de algunas ecuaciones diofantinas. Por supuesto, el teorema que vamos a presentar es de una naturaleza práctica, que sigue en gran medida los lemas presentados en el Teorema de Baker, pero como veremos, los detalles son distintos.

El propósito de esta demostración, no es repetir los mismos detalles que estuvieron presentes en el Teorema de Baker, si no, centrarse en aquellos detalles que realmente son propios de esta demostración y requieren mayor cuidado.

5.1 Formas lineales en los logaritmos de números algebraicos

A continuación, veremos uno de los Teoremas que Baker utiliza para

Teorema 5.1.1. Sea K una extensión normal sobre los racionales de grado $[K : \mathbb{Q}] = d \geq 4$ para los $\alpha_1, \dots, \alpha_n$ números algebraicos no nulos con $n \geq 2$, definimos $0 < \delta \leq 1$ y $A \geq 4$ la altura de los números algebraicos $\alpha_1, \dots, \alpha_n$. Sean $b_1, \dots, b_n$ enteros, con valor absoluto a lo más H que satisfacen

$$0 < |b_1 \log(\alpha_1) + \dots + b_n \log(\alpha_n)| < e^{-\delta H} \quad (5.1)$$

Entonces

$$H < (4^{n^2} \delta^{-1} d^2 \log(A))^{(2n+1)^2} \quad (5.2)$$

Demostración. Comenzaremos por observar que si α es un número algebraico distinto de 0, 1, de grado a lo más d y altura a lo más $A \geq 4$, entonces,

$$(2dA)^{-d-4} \leq |\log(\alpha)| \leq 4 \log(dA) \quad (5.3)$$

Para ello, veamos dos casos. Si $|\alpha| < 1$, entonces $|\alpha| < dA$, y si $|\alpha| > 1$, consideremos el polinomio mínimo e irreducible f de α tal que $f(\alpha) = 0$, por lo cual,

$$a_d \alpha^d + \dots + a_1 \alpha + a_0 = 0 \quad a_i \in \mathbb{Z} \Rightarrow \\ |a_d \alpha^d| = |a_{d-1} \alpha^{d-1} + \dots + a_1 \alpha + a_0|$$

Entonces

$$|\alpha| \leq |a_d \alpha| = |a_{d-1} + \dots + a_1 \alpha^{1-(d-1)} + a_0 \alpha^{-(d-1)}| \leq Ad$$

De manera análoga, si $|\alpha| > 1$, entonces $Ad > 1 > |\alpha|^{-1}$. Y si $|\alpha| < 1$, entonces, notemos que $|a_0| > 1$, de lo contrario $a_0 = 0$, por lo tanto f es un polinomio reducible, lo cual es una contradicción, entonces

$$|A\alpha d| > |a_d \alpha^d + \dots + a_1 \alpha| = |a_0| \geq 1 \Rightarrow$$

$$Ad \geq |\alpha|^{-1} \text{ además considerando } Ad > |\alpha| \Rightarrow |\log(|\alpha|)| < \log(Ad)$$

Como $\sqrt{a^2 + b^2} \leq |a| + |b|$ para $|a|, |b| \geq 2$, entonces

$$|\log(\alpha)| \leq ((\log(|\alpha|))^2 + \pi^2)^{\frac{1}{2}} \leq |\log(|\alpha|)|\pi| < 4 \log(Ad)$$

Ahora notemos que $|N(\alpha - 1)| \geq 1 \geq A^{-1}$, y cada conjugado de $\alpha - 1$ tiene un valor a lo más $2dA$, pues $\sigma(\alpha - 1) = \sigma(\alpha) - \sigma(1) = \sigma(\alpha) - 1 \leq \sigma(\alpha) \leq Ad < 2Ad$ por los mismos argumentos anteriores, entonces $|\alpha - 1| \geq (2dA)^{-d}$. Utilizando

$$|e^z - 1| \leq |z|e^{|z|}$$

Concluimos que

$$|\alpha - 1| = |e^{\log(\alpha)} - 1| \leq |\log(\alpha)|e^{|\log(\alpha)|} \leq |\log(\alpha)|(dA)^4 \\ (2dA)^{-d} \leq |\log(\alpha)|(dA)^4 \Rightarrow (2dA)^{-d-4} \leq |\log(\alpha)| \leq 4 \log(dA)$$

Lema 5.1.1. Se cumple la ecuación (5.2) o bien, existe un entero k con $2 \leq k \leq n$ tal que se cumplen las siguientes condiciones:

1. Existen enteros $b_1'', \dots, b_n''$ con valores absolutos a lo mas $(2H)^{n-k+1}$ tales que

$$0 < |b_1'' \log(\alpha_1) + \dots + b_n'' \log(\alpha_n)| < H^{n-k} e^{-\delta H}$$

2. Al menos $n - k$ de los enteros $b_1'', \dots, b_n''$ equivalen a 0

3. Los únicos enteros $b_1', \dots, b_n'$ con valores absolutos a lo más H tales que

$$b_1' \log(\alpha_1) + \dots + b_n' \log(\alpha_n) = 0 \quad (5.4)$$

Y $b_j' = 0$ cuando $b_j'' = 0$ están dados por $b_1' = \dots = b_n' = 0$.

Demostración. Claramente (1) y (2) son válidos para $k = n$ ya que se reduce al caso inicial (5.1) con $b_j'' = b_j$. Si también se cumple (3) no habría nada que probar, por lo tanto, supongamos que (3) no es válido para $k = n$, entonces, existe el mínimo valor $k \geq 2$ para el cual (1) y (2) se cumplen pero (3) no. Sean $b_1'', \dots, b_n''$ los enteros que satisfacen (1), (2) y sean $b_1', \dots, b_n'$ los enteros con las propiedades especificadas en (3) distintos de $b_1' = \dots = b_n' = 0$. Supongamos que $b_m' \neq 0$ y definamos

$$b_j''' = b_j'' b_m' - b_j' b_m'' \quad (1 \leq j \leq n)$$

Notemos que

$$0 < |b_1''' \log(\alpha_1) + \dots + b_n''' \log(\alpha_n)| < |b_m'| H^{n-k} e^{-\delta H} \leq H^{n-k+1} e^{-\delta H}$$

Además, tenemos que $b_j''' = 0$ cuando $b_j'' = 0$ y también $b_m''' = 0$ con $b_m'' \neq 0$. Por lo tanto, al menos $n - k + 1$ de los enteros $b_1''', \dots, b_n'''$ son 0 y los b_j''' tienen valores absolutos a lo más $(2H)^{n-k+2}$. Entonces (1) y (2) se satisfacen con $k - 1$. Por la selección de k , se sigue que se cumple (3) para $k - 1$ con las condiciones del lema o bien no se cumple porque $k - 1 = 1$. De ser lo segundo, entonces al menos $n - (k - 1) = n - 1$ de los enteros $b_1'', \dots, b_n''$ son cero, por lo cual,

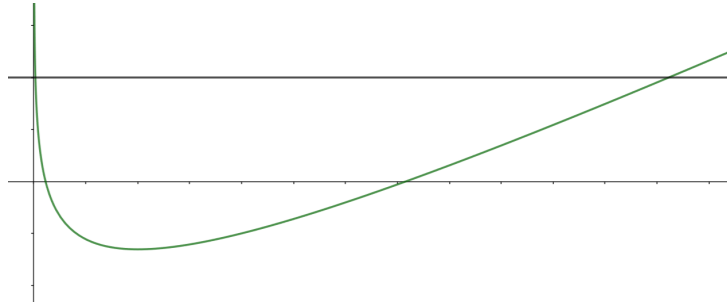
$$0 < |b_j'' \log(\alpha_j)| < H^{n-1} e^{-\delta H}$$

Para algún entero b_j'' y así por (5.3) con $\alpha = \alpha_j$ obtenemos:

$$(2dA)^{-d-4} < H^{n-1} e^{-\delta H}$$

$$\delta H - (n - 1) \log H < (d + 4) \log(2dA)$$

Ahora, nuestro objetivo es acotar H . Baker no proporciona argumentos para esto, por lo cual, probablemente exista alguna idea obvia que se está omitiendo, aún así, presentaremos nuestro propio análisis para llegar al resultado. Para ello, notemos que la función tiene el siguiente comportamiento:



Función $\delta H - (n - 1) \log(H) < (d + 4) \log(2dA)$

Así, para hallar el punto máximo que puede tomar H definimos:

$$f(H) = \delta H - (n-1)\log(H) - (d+4)\log(2dA) = 0$$

Es decir, el punto máximo se obtiene cuando $\delta H - (n-1)\log(H) = (d+4)\log(2dA)$ (notando que hay de hecho dos valores de H que satisfacen esta ecuación), de lo cual se obtiene:

$$\begin{aligned} H^{n-1} &= e^{\delta H} (2dA)^{-(d+4)} \\ H &= e^{\frac{\delta H}{n-1}} (2dA)^{-\frac{(d+4)}{n-1}} \\ -\frac{\delta H}{(n-1)} e^{-\frac{\delta H}{n-1}} &= -\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}} \end{aligned}$$

Este tipo de ecuaciones, se resuelven utilizando la función W de Lambert, la cual satisface la ecuación $ye^y = x \Rightarrow y = W(x)$. Por lo tanto,

$$\begin{aligned} -\frac{\delta H}{n-1} &= W\left(-\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}}\right) \\ H &= -(n-1)W\left(-\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}}\right) \delta^{-1} \end{aligned}$$

Ahora, de hecho, para especificar cuál de los valores de H se obtiene, definimos $W = W_{-1}$ (el cual admite una única solución y de hecho, la que buscamos), para acotar $W_1(x)$ referimos a [5], el cual obtiene para $u > 0$:

$$-1 - \sqrt{2u} - u < W_{-1}(-e^{-u-1}) < -1 - \sqrt{2u} - \frac{2}{3}u \quad (5.5)$$

Entonces, ajustamos como sigue

$$\begin{aligned} -\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}} &= -e^{\ln\left(\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}}\right)} \\ &= -e^{-\left[\frac{(d+4)}{n-1} \ln\left(\frac{\delta}{n-1} (2dA)\right) - 1\right] - 1} \end{aligned} \quad (5.6)$$

Por supuesto, es necesario justificar que

$$k = \left\lceil \frac{(d+4)}{n-1} \ln\left(\frac{\delta}{n-1} (2dA)\right) - 1 \right\rceil > 0$$

Notemos que cuando $u \leq 0$, entonces $-e^{-u-1} \leq -e^{-1}$. Supongamos que $k \leq 0$, por lo tanto, por (5.6):

$$-\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}} \leq -e^{-1} \quad (5.7)$$

5.1. FORMAS LINEALES EN LOS LOGARITMOS DE NÚMEROS ALGEBRAICOS 95

Veamos que es suficiente demostrar lo siguiente para llegar a una contradicción:

$$(n-1)(2dA)^{\frac{1}{n-1}} > e \quad (5.8)$$

Notemos la siguiente serie de desigualdades que dependen de la anterior:

$$\begin{aligned} (n-1)(2dA)^{\frac{(d+4)}{n-1}} &> e\delta \quad \text{ya que } (0 < \delta < 1) \\ (2dA)^{\frac{(d+4)}{n-1}} &> e \frac{\delta}{n-1} \\ e^{-1} &> \frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}} \\ -\frac{\delta}{n-1} (2dA)^{-\frac{(d+4)}{n-1}} &> -e^{-1} \end{aligned}$$

Lo cual contradice (5.7), por lo cual $k > 0$ como se requería. Ahora verificaremos (5.8). Por simplicidad, definamos $m = n-1$ y como $d \geq 4, A \geq 4$:

$$m(2dA)^{\frac{1}{m}} > m(e)^{\frac{1}{m}}$$

Definamos la función $f(m) = me^{\frac{1}{m}}$ y veamos cual es su punto mínimo.

$$f'(m) = e^{\frac{1}{m}} - \frac{e^{\frac{1}{m}}}{m} = 0$$

$$\Rightarrow 1 - \frac{1}{m} = 0 \Rightarrow m = 1$$

Donde $f(1) = e$. Por lo tanto, concluimos que:

$$m(2dA)^{\frac{1}{m}} > me^{\frac{1}{m}} \geq e$$

Para toda m . Con la consideración que la derivada de f no es válida cuando $m = 0$, en cuyo caso, $n = 1$ y sabemos que $n \geq 2$. Ahora, ya que hemos visto que se cumplen todas las condiciones de (5.8), concluimos que:

$$\begin{aligned} W_{-1}(-e^{-k-1}) &> -1 - \sqrt{2k} - k \\ \Rightarrow -(n-1)\delta^{-1}W_{-1}(-e^{-k-1}) &< -(n-1)\delta^{-1}(-1 - \sqrt{2k} - k) \\ \Rightarrow H &< (n-1)\delta^{-1}(1 + \sqrt{2k} + k) \end{aligned}$$

Donde $k = \left\lceil \frac{(d+4)}{n-1} \ln \left(\frac{\delta}{n-1} (2dA) \right) - 1 \right\rceil$. Lo cual concluye el lema. ■

Ahora, tenemos los siguientes dos casos. Si por un lado, no existe k que satisfice las condiciones del Lema 5.1.1, entonces el teorema se cumple trivialmente. Por lo tanto, de ahora en adelante asumiremos que existe un entero k con $2 \leq k \leq n$ tal que se satisfacen las siguientes condiciones:

1. Existen enteros $b_1, \dots, b_n$ no nulos con valores absolutos a lo mas $(2H)^{n-k+1}$ tales que

$$0 < |b_1 \log(\alpha_1) + \dots + b_n \log(\alpha_n)| < H^{n-k} e^{-\delta H}$$

2. No existen enteros $b'_1, \dots, b'_n$ con valores absolutos a lo más H tales que

$$b'_1 \log(\alpha_1) + \dots + b'_n \log(\alpha_n) = 0 \quad (5.9)$$

Notemos que $H^{n-k} e^{-\delta H} < e^{-\frac{1}{2}\delta H}$ para $k \geq 2$. Se asumirá que

$$H \geq (4^{n^2 - \frac{1}{2}} \delta^{-1} d^{2n'} \log(A))^{(2n+1)^2} \quad (5.10)$$

Y se mostrará que esta última suposición generará una contradicción. La contradicción será suficiente para demostrar el teorema, pues si (2) no se satisface, entonces para $k \geq 2$ tendremos que $(2H)^{n-k+1} \leq H^n$ y $H^{n-k} e^{-\delta H} \leq e^{-\frac{1}{2}\delta H}$.

Lemas. Supondremos, sin pérdida de generalidad, que $b_n \neq 0$ y escribiremos

$$\beta_j = -\frac{b_j}{b_n} \quad 1 \leq j \leq n$$

Entonces $\beta_1, \dots, \beta_{n-1}$ satisfacen

$$0 < |\beta_1 \log(\alpha_1) + \dots + \beta_{n-1} \log(\alpha_{n-1}) - \log(\alpha_n)| < e^{-\delta H} \quad (5.11)$$

Utilizando $|e^z - 1| \leq |z|e^{|z|}$ y $z = \beta_1 \log(\alpha_1) + \dots + \beta_{n-1} \log(\alpha_{n-1}) - \log(\alpha_n)$, obtenemos:

$$|\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}} - \alpha_n| < |\alpha_n| e^{-\delta H + 1} \quad (5.12)$$

Definimos

$$k = \left[H^{\frac{2}{2n+1}} \right] \quad h = \left[k^{\frac{1}{4n+2}} \right] \quad L = \left[k^{\frac{2n-1}{2n+1}} \right]$$

Donde $[x]$ denota la parte entera de x . Por virtud de (5.10) (ser suficientemente grande), se cumple que

$$k > H^{\frac{2}{2n+1}} - 1 > \frac{3}{4} H^{\frac{2}{2n+1}} \quad (5.13)$$

$$h > k^{\frac{1}{4n+2}} - 1 > \frac{3}{4} k^{\frac{1}{4n+2}} \quad (5.14)$$

De donde

$$h > 4^{n^2 - \frac{1}{2}} \delta^{-1} d^{2n} \log(A) \quad (5.15)$$

Lema 5.1.2. Existen enteros $p(\lambda_1, \dots, \lambda_n)$, no todos cero, con valores absolutos a lo más e^{hk} tales que la función

$$\Phi(z_1, \dots, z_{n-1}) = \sum_{\lambda_1=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_1, \dots, \lambda_n) \alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}}$$

Donde $\gamma_r = \lambda_r + \lambda_n \beta_r$ ($1 \leq r < n$) satisface

$$|\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)| < e^{-\frac{1}{2} \delta H} \quad (5.16)$$

Para todos los enteros l con $1 \leq l \leq Dh$ y todos los enteros no negativos $m_1, \dots, m_{n-1}$ que satisfacen $m_1 + \dots + m_{n-1} \leq k$. Donde $D = [K : \mathbb{Q}]^n = d^n$

Demostración. Primero determinaremos los enteros $p(\lambda_1, \dots, \lambda_n)$ tales que:

$$\sum_{\lambda_1=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_1, \dots, \lambda_n) \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}} = 0 \quad (5.17)$$

Para todos los enteros l con $1 \leq l \leq Dh$ y todos los enteros no negativos $m_1, \dots, m_{n-1}$ que satisfacen $m_1 + \dots + m_{n-1} \leq k$. Primero, ¿cuántos conjuntos $(m_1, \dots, m_{n-1})$ hay? Nuevamente, podemos tomar la cota máxima k^{n-1} . Por lo tanto, se tienen a lo más, $p := Dh(k^{n-1})$ ecuaciones con $q := (L+1)^n$ variables. Veamos que hay más variables que ecuaciones, lo cual es requisito para utilizar el Lema 2.4.3.

$$q = (L+1)^{n+1} > k^{\frac{(2n-1)n}{2n+1}} = k^{n-1 + \frac{1}{2n+1}} \geq h^2 k^{n-1} > Dh(k^{n-1}) = p$$

De hecho, $q > 2pd(d+1)$, ya que $h > 2Dd(d+1)$. Esto directamente implica que:

$$\frac{pd(d+1)}{q - pd(d+1)} < 1 \quad (5.18)$$

Ahora, visto como un sistema de ecuaciones, los únicos coeficientes $\alpha_j^{\lambda_j l}$ lógicamente son menores a A^{LDh} con $L \leq kh^{-4}$ y los $\gamma_j^{m_j}$ son menores a $(L(A+1))^k$. En cualquier caso, son menores a la cota superior $(kh^{-4}(A+1))^{kh^{-2}}$. Utilizando

el Lema 2.4.3 y la ecuación (5.18), existe una solución no trivial (que de hecho son los enteros $p(\lambda_1, \dots, \lambda_n)$) con valor menor a

$$1 + (qkh^{-4}(A+1)^{kh^{-2}})^{\frac{pd(d+1)}{q-pd(d+1)}} \leq qkh^{-4}(A+1)^{kh^{-2}} \leq e^{hk}$$

Por lo cual, es suficiente demostrar que (5.17) implica (5.16). Ahora, claramente (5.17) se obtiene de $\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)$ exceptuando por el factor faltante

$$P = (\log(\alpha_1))^{m_1} \dots (\log(\alpha_{n-1}))^{m_{n-1}} \quad (5.19)$$

Y el hecho de que en (5.17) tenemos α_n y en (5.16) se tiene $\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}}$. Por lo tanto, restando ambas ecuaciones, recordando que una de las ecuaciones equivale a cero, obtenemos

$$|\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l) - P \cdot 0| \leq (L+1)^n |P \alpha_1^{\lambda_1 l} \cdot \alpha_{n-1}^{\lambda_{n-1} l} \gamma_1^{m_1} \dots \gamma_{n-1}^{m_{n-1}}| |(\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}})^{\lambda_n l} - \alpha_n^{\lambda_n l}|$$

Ahora, se puede probar por inducción que para $n \geq 0$ y $x > y > 0$ se cumple:

$$nx^{n-1}(x-y) \geq x^n - y^n$$

Por lo tanto, utilizando (5.12) obtenemos que:

$$\begin{aligned} |(\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}})^{\lambda_n l} - \alpha_n^{\lambda_n l}| &\leq \lambda_n l (|\alpha_n| + 1)^{\lambda_n l} |(\alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}}) - \alpha_n| \\ &\leq \lambda_n l (|\alpha_n| + 1)^{\lambda_n l} |\alpha_n| e^{-\delta H + 1} \\ &\leq (4A)^{2Ll} e^{-\delta H} \end{aligned}$$

Y a partir de (5.3) obtenemos que:

$$|P \alpha_1^{\lambda_1 l} \cdot \alpha_{n-1}^{\lambda_{n-1} l} \gamma_1^{m_1} \dots \gamma_{n-1}^{m_{n-1}}| \leq (4dA)^{n(k+Ll)} (2LH^n)^k$$

Claramente $Ll < k$ y a partir de (5.15) se obtiene $(4dA)^{2nk} < e^{hk}$. Por lo cual concluimos que

$$|\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)| < e^{5hk - \delta H}$$

Lo cual es suficiente para demostrar el resultado, pues $h \leq k^{\frac{1}{10}} \leq H^{\frac{1}{25}}$ y utilizando (5.10) $5hk < 5H^{\frac{1}{2}} < \frac{1}{2}\delta H$. Lo cual concluye la demostración. ■

Lema 5.1.3. Para cualesquiera enteros no negativos $m_1, \dots, m_{n-1}$ que satisfacen $m_1 + \dots + m_{n-1} \leq k$ y cualquier número complejo z se cumple

$$|\Phi_{m_1, \dots, m_{n-1}}(z, \dots, z)| < e^{2hk} (dA)^{5nL|z|} \quad (5.20)$$

Además, para cualquier entero l con $1 \leq l \leq k^{\frac{(2n^2+1)}{2n+1}}$ se cumple (5.16) o bien

$$|\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l)| > (e^{2hk} (dA)^{3nLl})^{-D} \quad (5.21)$$

Demostración. Tenemos

$$\Phi_{m_1, \dots, m_{n-1}}(z, \dots, z) = P \sum_{\lambda_1=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_1, \dots, \lambda_n) q(\lambda, z)$$

Donde P está dada por (5.19) y

$$q(\lambda, z) = \alpha_1^{\gamma_1 z} \cdots \alpha_n^{\gamma_n z} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$$

Ahora, utilizando (5.3) y (5.11) obtenemos

$$|\alpha_r^{\lambda_r z}| \leq e^{\lambda_r |\log \alpha_r| |z|} \leq (dA)^{4L|z|} \quad (1 \leq r < n)$$

$$\left| \left(\alpha_1^{\beta_1 z} \cdots \alpha_{n-1}^{\beta_{n-1} z} \right)^{\lambda_n} \right| \leq e^{\lambda_n (|\log(\alpha_n)| + 1) |z|} \leq (dA)^{5L|z|}$$

Y así finalmente obtenemos

$$|Pq(\lambda, z)| < (dA)^{(4n+1)L|z|} \left(4e^{\frac{19}{20}h} \log(dA) \right)^{m_1 + \dots + m_{n-1}}$$

Notemos que hay a lo más $k^n < e^{\frac{1}{40}hk}$ términos en la suma, los términos $p(\lambda_1, \dots, \lambda_n)$ tienen valores absolutos a lo más e^{hk} y por (5.15) se tiene $4 \log(dA) < e^{\frac{1}{40}h}$, combinando estas ideas, se obtiene (5.20).

Ahora, para probar la segunda parte del lema (5.21) comenzamos definiendo $a_1, \dots, a_n$ como los coeficientes líderes de los polinomios irreducibles mínimos de $\alpha_1, \dots, \alpha_n$ respectivamente.

$$Q = P' \sum_{\lambda_1=0}^L \cdots \sum_{\lambda_n=0}^L p(\lambda_1, \dots, \lambda_n) q'(\lambda, z)$$

Donde

$$P' = (a_1, \dots, a_n)^{Ll} b_n^{m_1 + \dots + m_{n-1}}, \quad q'(\lambda, l) = \alpha_1^{\lambda_1 l} \cdots \alpha_n^{\lambda_n l} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}$$

Claramente Q es un entero algebraico de grado a lo más $D = d^n$ (recordando que $[K : \mathbb{Q}] = d$). Además, se tiene que,

$$|b_n^{m_1 + \dots + m_{n-1}} \gamma_1^{m_1} \cdots \gamma_{n-1}^{m_{n-1}}| < e^{\frac{19}{20}hk}$$

Y así, para cada conjugado de Q (que de hecho se obtiene sustituyendo por los conjugados de α_r) tiene un valor absoluto a lo más

$$(L+1)^n (dA)^{2nLl} e^{\left(\frac{19}{20}+1\right)hk} < (dA)^{2nLl} e^{2hk}$$

Así, si $Q \neq 0$, sabemos que $|N(Q)| > 1$ y por lo tanto,

$$|Q| \geq (e^{2hk}(dA)^{2nLl})^{-D+1}$$

Nuevamente utilizando (5.12) (y de manera análoga al procedimiento del lema anterior) se obtiene

$$|q(\lambda, l) - q'(\lambda, l)| < (4dA)^{(n+2)Ll} e^{hk-\delta h}$$

Notemos que además,

$$Ll \leq k^{\frac{2n(n+1)}{2n+1}} \leq H^{1-\frac{1}{(2n+1)^2}}, \quad hk < H^{\frac{1}{2}} \quad (5.22)$$

Por lo cual podemos obtener lo siguiente,

$$|P^{-1}\Phi_{m_1, \dots, m_{n-1}}(l, \dots, l) - P'^{-1}Q| \leq (L+1)^n e^{hk-\frac{3}{4}\delta H} \leq e^{-\frac{5}{8}\delta H}$$

Pero, como se tiene que

$$|P'| \leq A^{nLl} H^{nk} < A^{nLl} e^{hk}$$

Y utilizando (5.3), (5.15)

$$|P| \geq (2dA)^{-(d+4)k} > e^{-hk}$$

Se concluye que $Q = 0$ o bien

$$|PP'^{-1}Q| > (e^{2hk}(dA)^{2nLl})^{-D}$$

Pero es claro que e^{2Dhk} y $(dA)^{2nDLl}$ no pueden exceder $e^{\frac{1}{8}\delta H}$, por lo cual, el lado derecho de esta desigualdad es al menos $2e^{-\frac{1}{2}\delta H}$. Y utilizando (5.3) obtenemos

$$|P| \leq (4\log(dA))^k < e^{\frac{1}{8}\delta H}$$

Lo cual implica directamente la segunda parte del lema (5.21) y las alternativas corresponden a los casos $Q = 0$ y $Q \neq 0$. ■

Lema 5.1.4. Sea J cualquier entero que satisface $0 < J < \frac{1}{2}n(4n-1)$. Entonces la ecuación (5.16) se cumple para todos los enteros l con $1 \leq l \leq k^{\frac{(2J+1)}{(4n+2)}}$ y cada conjunto de enteros no negativos $m_1, \dots, m_{n-1}$ con $m_1 + \dots + m_{n-1} \leq \frac{k}{2^J}$

5.1. FORMAS LINEALES EN LOS LOGARITMOS DE NÚMEROS ALGEBRAICOS 101

Demostración. Supongamos que K es un entero que satisface $0 \leq K \leq \frac{1}{2}n(4n-1)-1$, y asumimos que el lema es cierto para los enteros $J = 0, 1, \dots, K$ (por supuesto, dándose cuenta que el caso base $J = 0$ es de hecho la conclusión del Lema 2). Por lo tanto, procederemos a demostrar la validez del lema para $J = K + 1$. Definimos $R_0 = Dh$ y $S_0 = k$ y

$$R_J = \left[k^{\frac{(2J+1)}{(4n+2)}} \right], \quad S_J = \left[\frac{k}{2^J} \right] \quad (J = 1, 2, \dots)$$

Por lo tanto, es suficiente probar que para cada entero l con $R_k < l \leq R_{k+1}$ y cualquier conjunto de enteros no negativos $m_1, \dots, m_{n-1}$ con $m_1 + \dots + m_{n-1} \leq S_{k+1}$ se tiene

$$|f(l)| < e^{-\frac{1}{2}\delta H} \quad (5.23)$$

Por hipótesis de inducción, vemos que para cada entero r con $1 \leq r \leq R_K$ y cada entero m satisface $0 \leq m \leq S_{K+1}$ tenemos que

$$|f^{(m)}(r)| < n^k e^{-\frac{1}{2}\delta H} \quad (5.24)$$

Donde

$$f(z) = \Phi_{m_1, \dots, m_{n-1}}(z, \dots, z)$$

Nuevamente, por la fórmula de Leibnitz, tenemos que:

$$f^{(m)}(r) = \sum_{\substack{j_1, \dots, j_{n-1} \\ j_1 + \dots + j_{n-1} = m}} \binom{m}{j_1, \dots, j_{n-1}} \Phi_{m_1+j_1, \dots, m_{n-1}+j_{n-1}}(r)$$

Con el valor absoluto de las derivadas a lo más $e^{-\frac{1}{2}\delta H}$ porque

$$m_1 + \dots + m_{n-1} + j_1 + \dots + j_{n-1} \leq \frac{k}{2^K}$$

Por lo tanto, la función $f(z)$ tiene ceros en $z = 1, 2, \dots, R_k$ de orden $\geq S_{k+1}$. De manera análoga a la demostración del Teorema de Gelfond-Schneider sea:

$$F(z) = [(z-1)(z-2)\dots(z-R_k)]^{S_{k+1}}$$

Entonces $\frac{f(z)}{F(z)}$ es una función entera. Así, definimos $R = R_{k+1}h$, entonces $R > 2R_k$. Sea D_R el disco cerrado $|z| \leq R$. Por el Teorema 2.6.3 obtenemos:

$$\left| \frac{f(z)}{F(z)} \right| \leq \frac{\max_{|z|=R} |f(z)|}{\min_{|z|=R} |F(z)|} \quad \text{para } z \in D_R$$

Así, para cada $z \in D_r$ obtenemos:

$$|f(z)| \leq \frac{|F(z)|(\max_{|z|=R} |f(z)|)}{\min_{|z|=R} |F(z)|} \quad (5.25)$$

Como $l > R_k$, $F(l) \neq 0$. Además,

$$|F(l)| = (|l-1| \cdots |l-R_k|)^{S_{k+1}} \leq R_{k+1}^{R_k S_{k+1}} \quad (5.26)$$

Ahora, se deduce de (5.24) y de las desigualdades

$$R_k(S_{k+1} + 1) \leq k^{1+\frac{(2K+1)}{(4n+2)}} \leq k^{\frac{4n^2+3n+1}{4n+2}} \leq H^{1-\frac{n}{(2n+1)^2}} \quad (5.27)$$

Además, como $R_{K+1} \leq k^{\frac{(4n^2-n+1)}{(4n+2)}}$, vemos que l satisface las condiciones del lema anterior, y por lo tanto, se cumple la ecuación (5.23) como se quería, o bien, por 5.21,

$$|f(l)| > (e^{2hk}(dA)^{3nLl})^{-D} > 2e^{-\frac{1}{8}\delta H} \quad (5.28)$$

Veremos que esta última ecuación implica una contradicción, completando el lema. Utilizando $R_{k+1} < H$ llegamos a que

$$|F(l)| < H^{R_k(S_{K+1}+1)} < e^{\frac{1}{8}\delta H}$$

Y para la última desigualdad solo se requiere que

$$H^{\frac{n}{(2n+1)^2}} > 8\delta^{-1} \log(H)$$

Pero es una consecuencia directa de (5.15), las definiciones de h, k y la desigualdad $H < e^h$, por lo cual se tiene $h > 8\delta^{-1}$ y

$$H^{\frac{1}{(2n+1)^2}} \geq h > \log(H)$$

Utilizando $|l-j| \leq l \leq R_{k+1}$. Dado que $R > 2R_k \Rightarrow 2R - 2R_k > R \Rightarrow |R - R_k| > \frac{R}{2}$, entonces para $1 \leq j \leq R_k$ se obtiene $|z-j| > \frac{R}{2}$. Por lo tanto,

$$\min_{|z|=R} |F(z)| \geq \left(\frac{R}{2}\right)^{R_k S_{k+1}} \geq \left(\frac{R_{k+1} h}{2}\right)^{R_k S_{k+1}}$$

Y a partir de la ecuación (5.26) se obtiene:

$$\min_{|z|=R} |F(z)| |F(l)|^{-1} \geq \left(\frac{1}{2}h\right)^{R_k(S_{K+1}+1)}$$

5.1. FORMAS LINEALES EN LOS LOGARITMOS DE NÚMEROS ALGEBRAICOS 103

Ahora bien, utilizando la ecuación (5.20) del lema anterior tenemos que:

$$\max_{|z|=R} |f(z)| \leq e^{2hk} (dA)^{5nLR_{k+1}h}$$

Y combinando con la ecuación (5.28) obtenemos:

$$\max_{|z|=R} |f(z)| |f(l)|^{-1} \leq (e^{2hk} (dA)^{5nLR_{k+1}h})^{D+1}$$

Por lo cual, podemos utilizar (5.25) para afirmar que:

$$\max_{|z|=R} |f(z)| |f(l)|^{-1} > \min_{|z|=R} |F(z)| |F(l)|^{-1}$$

Es decir,

$$(e^{2hk} (dA)^{5nLR_{k+1}h})^{D+1} > \left(\frac{1}{2}h\right)^{R_k(S_{K+1}+1)}$$

Aplicando logaritmos en ambos lados cuando $K = 0$

$$(D+1)(2hk + 5nk \log(dA)) \geq \frac{1}{2} Dhk \log\left(\frac{1}{2}h\right)$$

Sin embargo, el número de la izquierda es a lo más $4Dhk$, mientras que por otro lado,

$$\log\left(\frac{1}{2}h\right) > \log(2^5 d^4) > 9$$

Lo cual es incompatible con el número de la derecha. Ahora, cuando $K > 0$, obtenemos

$$(D+1)(2hk + 5nk^{1+\frac{K}{2n+1}} \log(dA)) \geq 2^{-(K+1)} k R_K \log\left(\frac{1}{2}h\right) \quad (5.29)$$

Pero $R_K \geq \frac{8}{9} k^{\frac{2K+1}{4n+2}}$ y tenemos $k^{\frac{1}{4n+2}} \geq h$, $\log\left(\frac{1}{2}h\right) > 9$ y $K \leq 2(n^2 - 1)$. Por lo tanto, el número de la derecha es al menos

$$2^{-K+2} h k^{1+\frac{K}{2n+1}} > 4D^2 k^{1+\frac{K}{2n+1}} \log(A)$$

Por otro lado, como $D \geq 16$ y $k^{\frac{1}{2n+1}} \geq h^2$, tenemos que

$$(D+1)2hk < 4Dhk \leq \frac{1}{4} D^2 k^{1+\frac{K}{2n+1}}$$

Y como claramente $D \geq d^{n-1} \log(d) > 2n \log(d)$ (pues $D = d^n$) y $D \geq 4^n \geq 8n$, vemos que

$$\begin{aligned} 5n(D+1) \log(dA) &\leq 5 \cdot \frac{17}{16} nD (\log(d) + \log(A)) \\ &\leq 6D^2 \left(\frac{1}{2} + \frac{1}{8} \log(A) \right) \\ &\leq \frac{15}{4} D^2 \log(A) \end{aligned}$$

Claramente con todas estas desigualdades, la ecuación (5.29) es inconsistente. Por lo cual, la ecuación (5.28) implica una contradicción, como se quería, completando el lema. ■

Lema 5.1.5. Para cada entero j con $0 \leq j \leq k^n$ tenemos que

$$\log |\phi_j(0)| < -2^{-\frac{1}{2}n(4n-1)} k^{\frac{n(4n+3)}{(4n+2)}} \quad (5.30)$$

donde

$$\phi(z) = \Phi(z, \dots, z)$$

Demostración. Primero definimos

$$X = \left\lfloor k^{\frac{4n^2-n-2}{4n+2}} \right\rfloor, \quad Y = \left\lfloor 2^{-\frac{1}{2}n(4n-1)} k \right\rfloor$$

Entonces $\left\lfloor k^{\frac{2J+1}{4n+2}} \right\rfloor \geq X$ y $\left\lfloor \frac{k}{2^J} \right\rfloor \geq Y$ donde J denota el entero más grande menor a $\frac{1}{2}n(4n-1)$ y así, por el lema anterior, se satisface la ecuación (5.16) para cada entero l , con $1 \leq l \leq X$ y cada conjunto de enteros no negativos $m_1, \dots, m_{n-1}$ que satisfacen $m_1 + \dots + m_{n-1} \leq Y$. Entonces,

$$|\phi^{(m)}(r)| \leq n^k e^{-\frac{1}{2}\delta H} \quad (5.31)$$

Para cada entero r con $1 \leq r \leq X$ y cada entero m que satisface $0 \leq m \leq Y$. Definamos $R = Xh$, entonces $R > 2X$ ya que $h > 2^{8n}$. Ahora, sea C el círculo $|z| = R$ y D_R el disco cerrado $|z| \leq R$, nuevamente construimos la función

$$F(z) = \frac{\phi(z)}{[(z-1) \cdots (z-X)]^{Y+1}}$$

La cual ya hemos visto varias veces que es una función entera. Por el Lema 2.6.3 se tiene que para toda $z \in D_R$ obtenemos:

$$|F(z)| \leq \frac{\max_{|z|=R} |\phi(z)|}{\min_{|z|=R} (|z-1| \cdots |z-X|)^{Y+1}}$$

Por lo tanto,

$$\phi(z) \leq \frac{(|z-1| \cdots |z-X|)^{Y+1} \max_{|z|=R} |\phi(z)|}{\min_{|z|=R} (|z-1| \cdots |z-X|)^{Y+1}}$$

Es claro que

$$|z-1| \cdots |z-X|^{Y+1} \leq (X+1)^{X(Y+1)}, \quad \min_{|z|=R} (|z-1| \cdots |z-X|)^{Y+1} \geq \left(\frac{1}{2}Xh\right)^{X(Y+1)}$$

Y por la ecuación (5.20) del Lema 5.1.3 tenemos que:

$$\max_{|z|=R} |\phi(z)| \leq e^{2hk} (dA)^{5nLXh}$$

Por lo cual, obtenemos

$$|\phi(z)| \leq 2e^{2hk} (dA)^{5nLXh} \left(\frac{1}{4}h\right)^{-X(Y+1)} + (2X)^{2XY} e^{-\frac{1}{4}\delta H}$$

Donde $(2X)^{2XY} e^{-\frac{1}{4}\delta H}$ es menor a $e^{-\frac{1}{8}\delta H}$ ya que

$$2XY \log(2X) < 2H^{1-\frac{(n+1)}{(2n+1)^2}} \log(H) < \frac{1}{8}\delta H$$

Y a partir de las desigualdades

$$L \leq kh^{-4}, \quad X > k, \quad Y+1 > 2^{-\frac{1}{2}n(4n-1)} h^{4n+2}$$

Determinamos que $2e^{2hk}$ y $(dA)^{5nLXh}$ no pueden exceder $e^{\frac{1}{2}X(Y+1)}$. Por todo lo anterior,

$$|\phi(z)| \leq \left(\frac{1}{16}h\right)^{-X(Y+1)} + e^{-\frac{1}{8}\delta H} \leq h^{-\frac{1}{2}X(Y+1)}$$

Finalmente, para cualquier entero $0 \leq j \leq k^n$ dentro del círculo $C' : |z| = X$, por el Teorema 2.6.4 y lo anterior,

$$\begin{aligned} |\phi^{(j)}(0)| &\leq \frac{j!}{2\pi} \int_{C'} \frac{|\phi(z)|}{|z|^{j+1}} dz \\ &\leq j! 4^j h^{-\frac{1}{2}X(Y+1)} \\ &\leq (4j)^j h^{-\frac{1}{2}X(Y+1)} \\ &\leq (4k)^{nk^n} h^{-\frac{1}{2}X(Y+1)} \\ &\leq e^{2nk^n \log(k)} h^{-\frac{1}{2}X(Y+1)} \end{aligned}$$

Notemos que utilizando (5.14) es claro que

$$2nk^n \log(k) \leq 8n(2n+1)k^n \log(h)$$

Pero tenemos que

$$X(Y+1) > 2^{-\frac{1}{2}n(4n-1)-1} k^{\frac{n(4n+3)}{(4n+2)}}$$

Y como $k^{\frac{1}{4n+2}} \geq h$, concluimos que $X(Y+1) > 8hk^n$. Además, $h > (2n+1)^3 \log(h)$ y por lo tanto,

$$e^{2nk^n \log(k)} < e^{\frac{1}{4}X(Y+1)}$$

Y así,

$$|\phi^{(j)}(0)| < h^{-\frac{1}{4}X(Y+1)}$$

Por lo cual, la ecuación (5.30) se cumple (cuando $\log(h) > 9$, lo cual es trivialmente cierto). Lo cual concluye la demostración. ■

Lema 5.1.6. Sean $t_1, \dots, t_n$ enteros con valores absolutos a lo más T y sea

$$W = t_1 \log(\alpha_1) + \dots + t_n \log(\alpha_n)$$

Entonces $W = 0$ o bien $|W| > (dA)^{-4nDT}$

Demostración. Por el Lema 2.3.2,

$$\omega = r(e^W - 1) = r(\alpha_1^{t_1} \dots \alpha_n^{t_n} - 1) \in \mathbb{A}$$

Por ser el producto de números algebraicos, donde r es exactamente el coeficiente líder del polinomio mínimo de $\alpha_1^{t_1} \dots \alpha_n^{t_n} - 1$ (por el Lema 2.3.2). Ahora, por la demostración del mismo lema, $r \geq 1$. Por lo tanto, si $\omega = 0$, entonces $e^\Omega = 1$, por lo tanto, Ω es un múltiplo de $2\pi i$, en este caso, el lema es válido trivialmente ($W = 0$ o bien mayor a 1). De lo contrario, $N(w) \geq 1$, entonces

$$N(\omega) = \prod_{\sigma \in G(K, F)} \sigma(\omega) \geq 1 \Rightarrow |w| \geq (dA)^{-4nT(D-1)}$$

Además, para cualquier z , $|e^z - 1| \leq ze^{|z|}$, por lo tanto,

$$(dA)^{-4nT(D-1)} \leq |\omega| \leq |rW|e^{|W|} \leq |r||W|e$$

Cuando $|W| < 1$ (pues el lema es trivialmente válido para $|W| > 1$). Notemos que de hecho, podemos tomar $r = a_1^{|t_1|} \dots a_n^{|t_n|}$ (el producto de los coeficientes líderes de cada α_i , pues estos garantizan que cada $\alpha_i^{t_i}$ sea un entero algebraico). Por lo cual, $|r| \leq A^{nT}$. Por lo tanto,

$$|W| \geq \frac{(dA)^{-4nT(D-1)}}{|r|e} \geq (dA)^{-4nDT}$$

■

5.1. FORMAS LINEALES EN LOS LOGARITMOS DE NÚMEROS ALGEBRAICOS 107

Ahora, procederemos a demostrar que las desigualdades (5.30) no pueden ser todas válidas, lo cual es suficiente para demostrar el Teorema. Definimos $R = (L + 1)^n - 1$, por lo cual, cualquier entero con $0 \leq r \leq R$ puede ser expresado de forma única como

$$r = \lambda_1 + \lambda_2(L + 1) + \dots + \lambda_n(L + 1)^{n-1}$$

Donde $\lambda_1, \dots, \lambda_n$ denota los enteros que satisfacen $0 \leq \lambda_j \leq L$ ($1 \leq j \leq n$). Para cada r definimos,

$$p_r = p(\lambda_1, \dots, \lambda_n) \quad \varphi_r = \lambda_1 \log(\alpha_1) + \dots + \lambda_n \log(\alpha_n)$$

Y definimos,

$$\Psi_j = \sum_{r=0}^R p_r \varphi_r^j \quad (0 \leq j \leq R)$$

Primero demostraremos que Ψ difiere de $\phi^{(j)}(0)$ por una cantidad a lo más $e^{-\frac{1}{2}\delta H}$. Notemos que,

$$\phi^{(j)}(0) = \sum_{\lambda_1}^L \dots \sum_{\lambda_n}^L p(\lambda_1, \dots, \lambda_n) (\gamma_1 \log(\alpha_1) + \dots + \gamma_{n-1} \log(\alpha_{n-1}))^j$$

A partir de (5.3), obtenemos que $|\varphi_r| < 4nL \log(dA)$ y por (5.11), se sigue que para cualquier $j \leq R$,

$$|(\gamma_1 \log(\alpha_1) + \dots + \gamma_{n-1} \log(\alpha_{n-1}))^j - \varphi_r^j| < (16nL \log(dA))^R e^{-\delta H}$$

Como $L < kh^{-4}$ y por definición de h , es fácil ver que,

$$(16nL \log(dA))^R e^{-\delta H} < k^R e^{-\delta H}$$

De esta manera,

$$|\varphi^{(j)}(0) - \Psi_j| < (R + 1)k^R e^{hk - \delta H}$$

A partir de las desigualdades

$$(R + 1)k^R \leq (2k)^R \leq e^{2k^n \log(k)}, \quad H \geq k^{n+\frac{1}{2}}, \quad k^{\frac{1}{2}} \geq h^{2n+1}$$

Obtenemos,

$$|\varphi^{(j)}(0) - \Psi_j| < e^{-\frac{1}{2}\delta H}$$

Ahora, utilizando el Lema 5.1.5 (el cual es válido pues $j \leq k^n$), obtenemos que,

$$\log |\Psi_j| < -2^{-\frac{1}{2}n(4n-1)-1} k^{\frac{n(4n+3)}{4n+2}}$$

Baker argumenta que cada entero p_r ($0 \leq r \leq R$) satisface una ecuación de la forma

$$p_r \Delta_r(\varphi_r) = \sum_{j=0}^R \sigma_{(r,j)} \Psi_j$$

Donde $\Delta_r(x)$ y $\sigma_{(r,0)}, \dots, \sigma_{(r,R)}$ están definidos por las identidades,

$$\Delta_r(x) = \sum_{\substack{s=0 \\ s \neq r}}^R (x - \varphi_s) = \sigma_{(r,0)} + \sigma_{(r,1)}x + \dots + \sigma_{(r,R)}x^R$$

Claramente $\varphi_r - \varphi_s$ con $r \neq s$, es una combinación lineal de los logaritmos $\log(\alpha_1), \dots, \log(\alpha_n)$ con coeficientes enteros, con cada coeficiente con un valor absoluto a lo más $L \leq k^{1-\epsilon} < H$ y no todos los coeficientes cero. Entonces, por nuestra suposición inicial, los únicos enteros $b'_1, \dots, b'_n$ con valores absolutos a lo más H tales que se satisface (5.9) están dados por $b'_1 = \dots = b'_n = 0$, por lo cual, $\varphi_r - \varphi_s \neq 0$ y así, por el Lema (5.1.6),

$$|\varphi_r - \varphi_s| > (dA)^{-4nDL}$$

Notemos que hay $R < (2L)^n$ factores en el producto definido por Δ_r y usando la desigualdad $L \leq k^{\frac{2n-1}{2n+1}}$ se sigue que,

$$\log |\Delta_r(\varphi_r)| > -2^{2n+2} D k^{\frac{(n+1)(2n-1)}{(2n+1)}} \log(dA) \quad (5.32)$$

Procederemos a calcular una cota superior para $\log(\Delta_r(\varphi_r))$ para cualquier r tal que $p_r \neq 0$ (y de hecho hay al meno una r con esta propiedad, por construcción). Tenemos que,

$$|\sigma_{(r,j)}| \leq \prod_{\substack{s=0 \\ s \neq r}}^R (1 + |\varphi_s|) \quad (0 \leq j \leq R) \leq (2k)^R$$

Ya que $|\varphi_j| \leq k$. Entonces,

$$\log(|p_r \Delta_r(\varphi_r)|) \leq \log(R+1) + R \log(2k) - 2^{-\frac{1}{2}n(4n-1)-1} k^{\frac{n(4n+3)}{4n+2}}$$

Como $R \leq k^n$ y $k^{\frac{1}{4n+2}} \geq h$, entonces

$$\log(R+1) + R \log(2k) \leq \frac{1}{2} \left| -2^{-\frac{1}{2}n(4n-1)-1} k^{\frac{n(4n+3)}{4n+2}} \right|$$

Por lo cual,

$$\log(|\Delta_r(\varphi_r)|) \leq -2^{-\frac{1}{2}n(4n-1)-2} k^{\frac{n(4n+3)}{4n+2}}$$

5.1. FORMAS LINEALES EN LOS LOGARITMOS DE NÚMEROS ALGEBRAICOS 109

Combinando esta desigualdad con la obtenida en (5.32), esto implica que,

$$k^{\frac{(n+2)}{(4n+2)}} < 2^{2n^2 + \frac{3}{2}n+4} D \log(dA)$$

Y como $k^{\frac{1}{4n+2}} \geq h$, esto es inconsistente con la ecuación (5.15). Y esta demostración finalmente concluye el teorema. ■

Para ver una de las aplicaciones de este teorema en las ecuaciones diofánticas, contextualizaremos el problema de los 1-ciclos en la conjetura de Collatz.

Capítulo 6

Conjetura de Collatz

Construyamos una sucesión de números que puede comenzar con cualquier número. Si nuestro número es par, el siguiente será el anterior dividido entre dos. Si nuestro número es impar, el siguiente será el anterior multiplicado por 3 y sumado con 1. Por ejemplo,

$$5 \text{ (impar)} \rightarrow 16 \text{ (par)} \rightarrow 8 \text{ (par)} \rightarrow 4 \text{ par} \rightarrow 2 \text{ par} \rightarrow 1$$

La conjetura de Collatz afirma que esta sucesión de números siempre alcanzará el valor 1. Curiosamente, si repetimos este proceso una vez más $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$ se obtiene un pequeño ciclo. Nuestra tarea es el estudio de los ciclos. Problemas como estos pueden generarnos una falsa ilusión, pues aparentan ser sencillos y uno puede sorprenderse al darse cuenta que conecta con áreas tan grandes como sistemas dinámicos, teoría de números trascendentales, teoría de la medida, sistemas computacionales, ecauciones diofanti-
nas, entre muchas otras.

Dada la altura de estas áreas, no es difícil convencerse que proporcionar cualquier avance significativo para demostrar la conjetura requiere de una maquinaria pesada, difícil de comprender e inaccesible para la mayoría de los matemáticos. Es por ello, que se intentará presentar el material con la mayor naturalidad posible para que los problemas surjan a partir de pequeñas observaciones. Esta, es una invitación para cualquier matemático que busque entender las raíces que originan los misterios de la conjetura.

Conjetura de Collatz: Sea $f : \mathbb{N} \rightarrow \mathbb{N}$ definida como:

$$f(n) = \begin{cases} \frac{n}{2} & \text{si } n \equiv 0 \pmod{2} \\ 3n + 1 & \text{si } n \equiv 1 \pmod{2} \end{cases}$$

Consideremos la sucesión definida como:

$$a_i = \begin{cases} n & \text{para } i=0 \\ f(a_{i-1}) & \text{para } i > 1 \end{cases}$$

Entonces la sucesión eventualmente toma el valor de 1 para todo $n \in \mathbb{N}$. Es decir, $a_j = f^j(n) = 1$ para algún $j \in \mathbb{N}$.

En términos simples, estamos definiendo una función que hace lo siguiente:

- Si n es par, lo dividimos entre 2.
- Si n es impar, lo multiplicamos por 3 y le sumamos 1.
- Con el nuevo valor, repetimos el proceso.

6.1 Dificultades de su estudio

Dada la simplicidad de esta conjetura, no hay muchas posibilidades para cualquier sucesión. Si $a_j = 1$, para alguna j , entonces $a_{j+1} = 4$, $a_{j+2} = 2$, $a_{j+3} = 1$. Es decir, eventualmente se obtiene lo que se considera un ciclo que se repite infinitamente. Si $a_i \neq 1 \forall k$, entonces la sucesión crece indefinidamente o bien, se obtiene algún otro ciclo. Por lo tanto, es suficiente ver que se cumplen dos condiciones para demostrar la veracidad de la conjetura:

- La sucesión a_i no crece indefinidamente.
- No existen ciclos además del trivial $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$.

Muchos matemáticos han intentado abordar de manera directa o indirecta cualquiera de estas dos condiciones, sin éxito. A diferencia de otras sucesiones, no se conoce alguna "fórmula" para obtener el valor de la sucesión en la k -ésima iteración. Esto es particularmente problemático, porque la sucesión puede alcanzar valores muy altos y bajos, sin ningún orden aparente. No obstante, existen acercamientos a dicha fórmula como veremos.

Además, otra de las dificultades es que no tenemos los conocimientos para predecir la descomposición en números primos al sumar 1.

6.2 Enfoques que han proporcionado mayor visión del problema

Las áreas más fructíferas han sido aquellas que trabajan en sistemas dinámicos. Las ecuaciones parciales tienen un comportamiento similar a la función de interés. El resultado más importante hasta el momento fue proporcionado por Terence Tao, al demostrar que la conjetura es casi verdadera para casi todos los números. Por supuesto, visto desde la perspectiva de teoría de la medida, esto significa que el conjunto de números que no cumplen con la conjetura de Collatz tiene medida 0.

6.3 Comportamiento de la función

Para entender acerca de los ciclos, uno primero tiene que estar interesado en comprender el comportamiento de la función. Principalmente, nos interesa conocer por qué la función crece o decrece. Claramente, si n es par, la función decrecerá hasta que deje de serlo. Por lo cual es más interesante conocer hasta qué punto la función crecerá si n es impar. Para ello, analicemos un ejemplo específico:

n	$f(n)$	$f^2(n)$
31	94	47
47	142	71
71	214	107
107	322	161
161	484	242

Podemos notar un comportamiento interesante, primeramente, $f(n)$ siempre es par y $f^2(n)$ es siempre impar (con una excepción). Veamos qué información se puede demostrar a partir de estas observaciones.

Proposición: Si n es impar, entonces $f(n)$ debe ser par.

Demostración Si n es impar, entonces $f(n) = 3n + 1$. Como n es impar, $n = 2k + 1$ para alguna $k \in \mathbb{N}$, por lo tanto, $f(n) = 3(2k + 1) + 1 = 6k + 3 + 1 = 6k + 4 = 2(3k + 2)$. Por lo tanto, $f(n)$ es par y además, $f^2(n) = 3k + 2$.

Observación. Notemos que la diferencia entre 47 y 31 es de 16. La diferencia entre 71 y 47 es de 24. La diferencia entre 107 y 71 es de 36. Y la diferencia entre 161 y 107 es de 54. Para estos valores, la siguiente diferencia será 1.5

veces mayor, es decir, $36 = 24(1.5)$, etc. Sin embargo, la diferencia entre 242 y 161 es de 81. Y $81(1.5)$ no es un entero. Así que algo falló al obtener 242. La clave para entender los ciclos está en comprender por qué este valor rompió con el comportamiento anterior.

Supongamos que el patrón continuara indefinidamente. Es decir, dado n impar $f^2(n)$ sea impar. Entonces por definición:

Iteración k -ésima	Valor obtenido
1	$\frac{3n+1}{2}$
2	$\frac{9n+5}{4}$
3	$\frac{27n+19}{8}$
4	$\frac{81n+65}{16}$
5	$\frac{243n+211}{32}$

Como sospechábamos, estas iteraciones tienen un patrón bien definido:

Proposición: Sea $g(n)$ la n -ésima iteración del patrón anterior, es decir, $g(n+1) = f^2(g(n))$ tal que $g(k)$ es impar para todo $k \in \mathbb{N}$. Entonces, el k -ésimo valor obtenido de $g(k)$ está dado por:

$$g(k) = \frac{3^k n + (3^k - 2^k)}{2^k}$$

Demostración: Veamos por inducción, sea $n = 1$, entonces:

$$g(1) = \frac{3^1(n) + 3^1 - 2^1}{2^1} = \frac{3n + 1}{2}$$

Supongamos que el resultado es válido para $n = k$ y veamos que es cierto para $n = k + 1$. Por hipótesis de inducción:

$$g(k) = \frac{3^k(n) + (3^k - 2^k)}{2^k}$$

Como $g(k)$ sigue el patrón, por definición $g(k+1) = f^2(g(k))$ donde $g(k)$ es impar, por lo tanto,

$$g(k+1) = \frac{3 \left(\frac{3^k(n) + (3^k - 2^k)}{2^k} \right) + 1}{2} = \frac{3^{k+1}(n) + 3^{k+1} - 3(2^k) + 2^k}{2}$$

$$g(k+1) = \frac{3^{k+1}(n) + (3^{k+1} - 2 \cdot 2^k)}{2 \cdot 2^k} = \frac{3^{k+1}(n) + (3^{k+1} - 2^{k+1})}{2^{k+1}}$$

Lo cual completa la inducción.

Corolario:

$$g(k) = \frac{3^k(n) + (3^k - 2^k)}{2^k} = \frac{3^k(n+1) - 2^k}{2^k} = \frac{3^k(n+1)}{2^k} - 1$$

Por supuesto, este análisis finalmente nos permite saber qué es lo que falló con 242. Específicamente, ¿hasta qué punto podemos garantizar que $g(k)$ es impar?

Proposición: Afirmamos que si $n+1 = 2^k \cdot m$ con m impar, entonces las siguientes iteraciones $g(1), g(2), \dots, g(k-1)$ serán impares y $g(k)$ será par.

Demostración: Como $n+1 = 2^k \cdot m$, entonces $g(j)$ para $j < k$ satisface

$$g(j) = \frac{3^j(2^k \cdot m)}{2^j} - 1 = 3^j(2^{k-j} \cdot m) - 1$$

Como $k > j$, entonces $g(j)$ es impar. Por otro lado,

$$g(k) = \frac{3^k(2^k \cdot m)}{2^k} - 1 = 3^k \cdot m - 1$$

Como el producto de impares es impar, entonces $3^k \cdot m - 1$ debe ser par.

Conclusión: En resumen, hemos obtenido dos informaciones valiosas. Dado n (par o impar), podemos determinar hasta qué punto crecerá o decrecerá hasta cambiar su comportamiento de la siguiente manera:

Si n es par	Tomará el valor:
$n = 2^k \cdot m$	$f^k(n) = m$
Si n es impar	Tomará el valor:
$n = 2^k \cdot m - 1$	$f^{2k}(n) = g(k) = 3^k \cdot m - 1$

6.4 Resultados relevantes para los ciclos

A partir de lo anterior, es claro que podemos dar ciertos saltos en la función. Esto nos permite definir lo que se conoce como k -ciclos.

Definición: Sea n un número impar de la forma $n = 2^k \cdot m - 1$, decimos que se forma un 1-ciclo si existen enteros (k, m, p) tales que $n = \frac{3^k \cdot m - 1}{2^p}$, es decir, si $n = \frac{f^{2^k}(n)}{2^p}$. Análogamente, decimos se forma un 2-ciclo si existen enteros (k, m, p) y (k_2, m_2, p_2) tales que $n \rightarrow \frac{3^k \cdot m - 1}{2^p} \rightarrow \frac{3^{k_2} \cdot m_2 - 1}{2^{p_2}}$ y $n = \frac{3^{k_2} m_2 - 1}{2^{p_2}}$. Así sucesivamente se definen los k -ciclos.

Proposición: Supongamos que no existen k -ciclos, entonces no existen ciclos en la conjetura de Collatz.

Demostración: Sin pérdida de generalidad, sea n impar, entonces $n \rightarrow 3^k \cdot m - 1 \rightarrow \frac{3^k \cdot m - 1}{2^p}$, como no existen 1-ciclos, entonces $n \neq \frac{3^k \cdot m - 1}{2^p}$, ahora, $\frac{3^k \cdot m - 1}{2^p}$ es un número impar, por lo tanto, tomará el valor $n \rightarrow \frac{3^k \cdot m - 1}{2^p} \rightarrow \frac{3^{k_2} \cdot m_2 - 1}{2^{p_2}}$ y claramente $n \neq \frac{3^{k_2} \cdot m_2 - 1}{2^{p_2}}$ porque no existen 2-ciclos. Por un argumento inductivo, no existen ciclos en la conjetura de Collatz.

Proposición: Supongamos que existe un k -ciclo, entonces existe un ciclo en la conjetura de Collatz.

Demostración: Por hipótesis, supongamos que existe un k -ciclo, entonces $n = 2^k \cdot m - 1$ tal que $n \rightarrow \dots \rightarrow \frac{3^{k_r} \cdot m_r - 1}{2^{p_r}}$ tal que $n = \frac{3^{k_r} \cdot m_r - 1}{2^{p_r}}$, por definición, n forma un ciclo.

Corolario: Existe un k -ciclo si y solo si existe un ciclo en la conjetura de Collatz. (Consecuencia de las dos proposiciones anteriores)

Conclusión: Los k -ciclos son la herramienta más cercana que tenemos para encontrar ciclos en la conjetura de Collatz, principalmente porque podemos utilizar una fórmula explícita para intentar encontrar los k -ciclos, mientras que no existe una fórmula explícita para los ciclos. Es por ello que todos los esfuerzos se han centrado en demostrar que no existen los k -ciclos refinando los argumentos de Steiner.

Capítulo 7

Teorema de Steiner

Es impresionante la cantidad de material necesario para intentar comprender en su totalidad la demostración de Steiner. Este capítulo, representa la conclusión y combinación de todos los resultados relevantes de los capítulos anteriores. En él, se puede apreciar la importancia de las formas lineales de los logaritmos, y la manera en que se utiliza para resolver problemas diofantinos.

Si bien es cierto que se puede leer la demostración de Steiner sin comprender la demostración de los teoremas que utiliza, seguir los pasos históricos que motivaron cada demostración tiene mucho valor para extender los resultados.

Para desarrollar esta demostración, utilizo la notación con la que desarrollé estas ideas de manera independiente, mejoro algunos detalles que Steiner no termina de aclarar y presento de manera convincente una cota inferior para las soluciones de la ecuación diofantina de los 1-ciclos utilizando Python.

7.1 Ecuación diofantina de los 1-ciclos

Sea $n = 2^k \cdot m - 1$, y supongamos que existe un 1-ciclo. Por lo tanto, existen enteros (k, m, p) tales que:

$$n = \frac{3^k \cdot m - 1}{2^p} \Rightarrow 2^k \cdot m - 1 = \frac{3^k \cdot m - 1}{2^p}$$

Por lo tanto, podemos resolver para m como sigue:

$$2^{k+p} \cdot m - 2^p = 3^k \cdot m - 1 \Rightarrow 2^{k+p} \cdot m - 3^k \cdot m = 2^p - 1$$

Finalmente, obtenemos que:

$$m = \frac{2^p - 1}{2^{k+p} - 3^k}$$

Recordemos que m es un entero. Por lo tanto, es suficiente encontrar todas las parejas (k, p) tales que la ecuación diofantina tiene solución, es decir:

$$h(k, p) = \frac{2^p - 1}{2^{k+p} - 3^k} \in \mathbb{Z} \quad (7.1)$$

7.2 Borrador de la demostración

Dividiremos la demostración en tres partes:

1. Mostraremos que la ecuación no es válida para $k = 1, \dots, 100$ utilizando algunas observaciones de la misma ecuación diofantina y un programa de Python.
2. Mostraremos que para que la ecuación sea válida, necesariamente $k < 4^{225}$. Para ello, emplearemos el Teorema 5.1.1 de Baker.
3. Usaremos la teoría de fracciones continuas para demostrar que k es en realidad, convergente en la fracción continua de $\log_2(3)$ y utilizaremos esto para determinar que $k > 4^{225}$.

7.3 Demostración

Primeramente, afirmamos que la ecuación diofantina (7.1) tiene a lo más una posible solución para cada valor de k . Para ello, notemos que si

$$\frac{2^p - 1}{2^{k+p} - 3^k} = z \in \mathbb{Z}^+$$

Entonces $2^{k+p} - 3^k > 0$. Por lo tanto,

$$k + p > k \log_2(3) \quad (7.2)$$

Por otro lado, $2^p - 1 \equiv 0 \pmod{2^{k+p} - 3^k}$, entonces,

$$\begin{aligned} 2^{k+p} &\equiv 2^k \pmod{2^{k+p} - 3^k} \\ 3^k &\equiv 2^k \pmod{2^{k+p} - 3^k} \end{aligned}$$

Concluimos que $3^k - 2^k = (2^{k+p} - 3^k)x$ para alguna $x \in \mathbb{Z}^+$, ahora notemos lo siguiente:

$$2^{k+p} - 3^k < (2^{k+p} - 3^k)x = 3^k - 2^k < 3^k \quad (7.3)$$

Por lo tanto, $2^{k+p} < 2(3^k)$, así,

$$k + p < k \log_2(3) + 1 \quad (7.4)$$

Combinando las ecuaciones (7.2) y (7.4) concluimos que:

$$k \log_2(3) < k + p < k \log_2(3) + 1$$

Por lo tanto, necesariamente $k + p = \lceil k \log_2(3) \rceil$, donde $\lceil x \rceil$ representa la función techo. En particular observemos que:

$$p = \lceil k \log_2(3) \rceil - k \quad (7.5)$$

Por lo cual concluimos que para cada valor de k , existe un único valor para p definido por la ecuación anterior. Es decir, para cada valor de k , hay a lo más una posible solución que resuelve la ecuación (7.1).

Procederemos a mostrar que los primeros valores $k = 1, \dots, 100$ no resuelven la ecuación (7.1). Para ello es suficiente utilizar una computadora que lo verifique. En este caso, se utilizó el siguiente código de Python:

```
import math

k = 1
while k <= 100:
    p = math.ceil(k*(math.log2(3))) - k
    if 2**(k+p) - 3**k < 2**p - 1:
        print("Valor de k =", k)
    k += 1
```

En pocas palabras, determina todos los valores de k que satisfacen la ecuación $2^{k+p} - 3^k < 2^p - 1$, la cual es necesaria para que se cumpla (7.1). Concluimos entonces que $k > 100$.

Ahora bien, consideremos la misma ecuación $2^{k+p} - 3^k < 2^p - 1$ dividiendo ambos lados por 2^{k+p} .

$$0 < 1 - \frac{3^k}{2^{k+p}} \leq \frac{2^p - 1}{2^{k+p}} \quad (7.6)$$

Por lo cual, obtenemos la siguiente serie de desigualdades:

$$\begin{aligned} 1 &> \frac{3^k}{2^{k+p}} > \frac{2^{k+p} + 1 - 2^p}{2^{k+p}} \\ 0 &> k \ln(3) - (k + p) \ln(2) > \ln\left(\frac{2^{k+p} + 1 - 2^p}{2^{k+p}}\right) \\ 0 &< (k + p) \ln(2) - k \ln(3) < -\ln\left(\frac{2^{k+p} + 1 - 2^p}{2^{k+p}}\right) \end{aligned}$$

Utilizando la siguiente desigualdad válida para toda $x > 0$,

$$\frac{x-1}{x} \leq \ln(x) \Rightarrow -\ln(x) \leq \frac{1-x}{x}$$

Por lo tanto,

$$\begin{aligned} 0 < (k+p)\ln(2) - k\ln(3) &< \frac{1 - \frac{2^{k+p} + 1 - 2^p}{2^{k+p}}}{\frac{2^{k+p} + 1 - 2^p}{2^{k+p}}} \\ &= \frac{2^{k+p} - 2^{k+p} - 1 + 2^p}{2^{k+p} + 1 - 2^p} \\ &= \frac{2^p - 1}{2^{k+p} + 1 - 2^p} \end{aligned}$$

Notemos que como $k, p > 1$, entonces

$$\begin{aligned} 2^{k+p} - 2^k - 2^p + 1 &\leq 2^{k+p} + 1 - 2^p \\ (2^k - 1)(2^p - 1) &\leq 2^{k+p} + 1 - 2^p \\ \frac{2^p - 1}{2^{k+p} + 1 - 2^p} &\leq \frac{1}{2^k - 1} \end{aligned}$$

Concluimos que

$$0 < (k+p)\ln(2) - k\ln(3) < \frac{1}{2^k - 1} \quad (7.7)$$

Ahora bien, notemos que por (7.5) es obvio que

$$p < 1.6k - k = 0.6k \Rightarrow k + p < 1.6k$$

Por lo tanto, $e^{0.64k} > e^{0.4(k+p)}$. Afirmamos que $2^k > e^{0.69k}$. De lo contrario,

$$k\ln(2) \leq 0.69k \Rightarrow \ln(2) \leq 0.69$$

Lo cual es una contradicción. Entonces $2^k - 1 > e^{0.69k} - 1$. Ahora afirmamos que $e^{0.69k} - 1 > e^{0.64k}$, ya que $f(x) = e^{0.69x} - e^{0.64x}$ es una función estrictamente creciente y de hecho mayor a 1 para $x \geq 3$.

Combinando estas ideas concluimos que

$$2^k - 1 > e^{0.64k} > e^{0.4(k+p)}$$

Por lo tanto, uniendo esta idea con la ecuación (7.7) se obtiene:

$$0 < (k + p) \ln(2) - k \ln(3) < \frac{1}{2^k - 1} < e^{-0.4(k+p)}$$

Ahora, aplicamos el Teorema 5.1.1 con $H = k + p$, $n = 2$, $d = 4$, $A = 4$ y $\delta = 0.4$. Entonces

$$k + p < \left(4^4 \cdot \frac{10}{4} \cdot 4^4 \cdot \ln(4)\right)^{25} < (4^9)^{25}$$

En particular, esto implica que $k < 4^{225}$, por lo cual, hemos calculado una cota superior para las soluciones de esta ecuación diofantina. Como comentario, esta cota es excesivamente grande, y fuera del alcance de cualquier ordenador. Por lo cual, no podemos verificar cada valor de k , lo que sí podemos hacer, es una serie de observaciones pertinentes en el área de fracciones continuas.

Consideremos la ecuación (7.7) y dividamos todo por $k \ln(2)$, entonces,

$$\begin{aligned} 1 + \frac{p}{k} - \frac{\ln(3)}{\ln(2)} &\leq \frac{1}{k \ln(2)(2^k - 1)} \\ \frac{p}{k} - \frac{\ln(3)}{\ln(2)} &< \frac{1}{k \ln(2)(2^k - 1)} \end{aligned}$$

Claramente $\ln(2)(2^k - 1) > 2k$ para $k > 4$. Por lo tanto,

$$\frac{p}{k} - \log_2(3) < \frac{1}{k \ln(2)(2^k - 1)} < \frac{1}{2k^2}$$

Notemos que por la ecuación (7.5), sabemos que:

$$\frac{p}{k} = \frac{[k \log_2(3)]}{k} - 1 > \log_2(3)$$

Por lo cual, claramente,

$$0 < \frac{p}{k} - \log_2(3) < \frac{1}{2k^2}$$

Y así, finalmente,

$$\left| \log_2(3) - \frac{p}{k} \right| < \frac{1}{2k^2}$$

Por el Teorema 2.1.7, $\frac{p}{k}$ debe ser un convergente en la fracción continua de $\log_2(3)$. Lo cual, además de sorprendente, es más fácil de calcular. Para ello, notemos que no es necesario calcular tanto, pues, utilizando el Corolario 2.1.2,

$$\frac{1}{(a_{n+1} + 2)q_n^2} < \left| \log_2(3) - \frac{p}{k} \right| < \frac{1}{k \ln(2)(2^k - 1)}$$

En particular, $k > 100$, implica necesariamente que $q_n > 100$ (pues de hecho $k = q_j$, para alguna j y los q_j son estrictamente crecientes). Por lo tanto,

$$\frac{1}{(a_{n+1} + 2)100^2} < \frac{1}{100 \ln(2)(2^{100} - 1)}$$

$$2^{99} < \ln(2)(2^{100} - 1) - 2 < a_{n+1}$$

Entonces necesariamente a_{n+1} tiene que exceder 2^{99} , veamos que esto no sucede, para ello, necesitamos determinar la fracción continua de $\log_2(3)$ hasta que su convergente $\frac{p_n}{q_n}$ cumpla con la condición $q_n > 4^{225}$. Sin más, utilizaremos WolframAlpha para determinar los cocientes parciales a_j y utilizaremos Python, motivado por el Teorema 2.1.4 para calcular los q_j , información que se encuentra disponible en la siguiente página. Como todos los cocientes parciales $a_j < 2^{99}$, y el denominador de los convergentes $q_n = k > 4^{225}$, entonces, no existen soluciones para la ecuación diofantina. ■

```

ListaDeCocientesParciales = [1, 1, 1, 2, 2, 3, 1, 5, 2, 23, 2
                             , 2, 1, 1, 55, 1, 4, 3, 1, 1,
                             15, 1, 9, 2, 5, 7, 1, 1, 4, 8,
                             1, 11, 1, 20, 2, 1, 10, 1, 4,
                             1, 1, 1, 1, 1, 37, 4, 55, 1,
                             1, 49, 1, 1, 1, 4, 1, 3, 2, 3,
                             3, 1, 5, 16, 2, 3, 1, 1, 1, 1,
                             , 1, 5, 2, 1, 2, 8, 7, 1, 1, 2,
                             , 1, 1, 3, 3, 1, 1, 1, 1, 5, 4,
                             , 2, 2, 2, 16, 8, 10, 1, 25, 2,
                             , 1, 1, 1, 2, 18, 10, 1, 1, 1,
                             1, 9, 1, 5, 6, 2, 1, 1, 12, 1,
                             , 1, 1, 6, 2, 12, 1, 1, 12, 1,
                             1, 2, 12, 1, 12, 3, 1, 5, 1,
                             14, 1, 1, 14, 2, 3, 1, 2, 2, 1,
                             , 4, 1, 4, 8, 1, 1, 1, 3, 5, 1,
                             , 1, 1, 1, 2, 1, 4, 3, 7, 5, 3,
                             , 1, 32, 1, 1, 1, 18, 1, 3, 2,
                             5, 2, 1, 3, 1, 8, 1, 1, 1, 2,
                             6, 6, 5, 33, 2, 2, 3, 1, 1, 1,
                             , 1, 29, 1, 3, 2, 1, 21, 1, 6,
                             52, 1, 8, 1, 4, 14, 9, 7, 1,
                             4, 18, 2, 2, 1, 1, 2, 100, 39,
                             1, 2, 1, 1, 19, 1, 5, 9, 1, 3,
                             , 964, 5, 1, 1, 1, 39, 1, 1, 1,
                             , 1, 5, 3, 1, 88, 1, 2, 1, 3,
                             1, 11, 1, 23, 11, 1, 1, 1, 2,
                             1, 1, 4, 3, 1, 5, 1, 4, 2, 1,
                             75, 1, 2, 5]

Q = 1
k = len(ListaDeCocientesParciales)
ListaDeConvergentesDenominador = [1,1,2]
a = False
i=3
while i<k:
    Q = ListaDeCocientesParciales[i]*
        ListaDeConvergentesDenominador
        [i-1]+
        ListaDeConvergentesDenominador
        [i-2]
    ListaDeConvergentesDenominador.append(Q)
    if Q>4**(225):
        print(i)
        a = True
        break
    i += 1
print(ListaDeConvergentesDenominador)
print('Q excede el valor 4**(225):', a)

```


Bibliografía

- [1] Stephen Abbott. *Understanding analysis*. Springer, 2001.
- [2] Alan Baker. Linear forms in the logarithms of algebraic numbers (iv). 1968.
- [3] Alan Baker. *Transcendental Number Theory*. Cambridge Mathematical Library, 2022.
- [4] David M. Burton. *Elementary number theory*. International series in pure and applied mathematics. McGraw-Hill, 5th ed edition, 2002.
- [5] Ioannis Chatzigeorgiou. Bounds on the lambert function and their application to the outage analysis of user cooperation. *CoRR*, abs/1601.04895, 2016.
- [6] Hilbert D. *Mathematical problems*. ICM Paris, lectures edition, 1900.
- [7] Samit Dasgupta. *Ranks of matrices of logarithms of algebraic numbers I*. 2023.
- [8] Arthur David Snider Edward B. Saff. *Fundamentals of complex analysis with applications to engineering and science*. Prentice Hall, 3rd ed edition, 2003.
- [9] Jan-Hendrik Evertse. Linear forms in logarithms. 2011.
- [10] Spence. Friedberg, Insel. *Algebra lineal*.
- [11] E.M. Wright G.H. Hardy. *An introduction to the Theory of numbers*. Oxford University Press, 4 edition, 1960.
- [12] Adam J. Harper. *A version of Baker's theorem on linear forms in logarithms*. 2010.
- [13] Christian Hercher. There are no collatz-m-cycles with $m \leq 91$, 2023.

- [14] I. N. Herstein. *Topics in Algebra 2nd Edition*. Wiley, 2 edition, 1975.
- [15] Jeffrey Lagarias. The $3x+1$ problem: An overview. 01 2010.
- [16] Ivan Niven. *Irrational numbers*. Carus Mathematical Monographs. The Mathematical Association of America, maa edition, 2005.
- [17] Harry Pollard. *The Theory of Algebraic Numbers*. The Mathematical Association of America, 2nd printing edition, 1961.
- [18] Natarajan Saradha. *Pillars of Transcendental Number Theory*. Springer, 2020.
- [19] John Simons. On the (non)-existence of m -cycles for generalized syracuse sequences. *Acta Arithmetica*, 131, 12 2007.
- [20] John L Simons. A simple (inductive) proof for the non-existence of 2-cycles of the $3x+1$ problem. *Journal of Number Theory*, 123(1):10–17, 2007.
- [21] Kannan Soundararajan. *Transcendental Number Theory*. 2011.
- [22] Ray. P. Steiner. *A theorem on the Syracuse problem*. Proceedings of the Seventh Manitoba Conference on Numerical Mathematics and Computing, 1977.
- [23] Weger. Theoretical and computational bounds for m -cycles of the $3n+1$ problem. *Acta Arithmetica*, 117(1), 01 2005.
- [24] Seung Hyun Yang. Continued fractions and pell's equation. 2008.