Universidad Nacional Autónoma de México

Facultad de Ciencias

División de la lemniscata con regla y compás.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:
ANTÓN BAUTISTA ROMO

DIRECTOR DE TESIS:
JOSÉ ANTONIO GÓMEZ ORTEGA

2009

Hoja de datos del jurado

1. Datos del alumno

   Bautista

   Romo

   Antón

   57 85 39 92

   Universidad Nacional Autónoma de México

   Facultad de Ciencias

   Matemáticas

   096001888

2. Datos del tutor

   M. en C.

   José Antonio

   Gómez

   Ortega

3. Datos del sinodal 1

   Dr.

   Alberto León

   Kushner

   Schnur

4. Datos del sinodal 2

   Dr.

   Eugenio

   Garnica

   Vigil

5. Datos del sinodal 3

   Dr.

   Rogelio

   Valdez

   Delgado

6. Datos del sinodal 4

   Mat.

   Julio César

   Guevara

   Bravo

7. Datos del trabajo escrito

   División de la lemniscata con regla y compás

   149 p

   2009

Agradecimientos

Un agradecimiento muy especial
al hombre de todas las respuestas
José Antonio Gómez Ortega.

A todos mis sinodales
por el tiempo dedicado.

Y a todos mis amigos.

A mis padres y mi hermana
por siempre estar ahí.

Índice general

Introducción

Esta tesis consistirá en el desarrollo de dos demostraciones del teorema acerca de la lemniscata formulado por Niels Henrik Abel, contenido en la segunda parte de su trabajo *Recherches sur les fonctions elliptiques*.

Teorema. *Es posible dividir la lemniscata en n partes iguales utilizando solamente regla y compás si y sólo si $n = 2^s p_1 \dots p_r$ donde s es un entero positivo y los p_i para $i = \{1, \dots, r\}$ son primos de Fermat distintos.*

Las demostraciones presentadas en este trabajo son contemporáneas y pertenecen a los matemáticos Michael Rosen y David A. Cox, estas demostraciones seguirán las ideas de las demostraciones del teorema presentadas por el mismo Abel y el matemático alemán F. Einsestein, pero con la diferencia de que estas harán uso de herramientas no desarrolladas completamente en el tiempo de Abel, como lo son la multiplicación compleja y la teoría de Galois.

Cabe mencionar que Abel sólo demostró la posibilidad de dividir la lemniscata con regla y compás en n partes iguales si la n tiene la forma descrita, pero el no probó que la construcción para otros valores de n fuera imposible. Pero las pruebas presentadas en este trabajo incluyen demostraciones detalladas de la doble implicación del teorema.

Nota para el lector:

Los primeros dos capítulos de este trabajo correspondientes a la suma de puntos en una cúbica y a las propiedades de la función $\wp(z)$ de Weierstrass, no son utilizados sino hasta la segunda parte del tercer capitulo, así que su lectura no interfiere con la primera demostración proporcionada.

Capítulo 1

Geometría de Curvas Cúbicas

En junio de 1796 la revista *Literature Gazette*, publicada en ese tiempo en Jena, ofreció a sus lectores la siguiente nota (en Alemán):

Nuevos Descubrimientos.

Todo principiante en geometría sabe que es posible construir geométricamente, es decir, con regla y compás, varios polígonos regulares, a saber, un triángulo, un pentágono, un 15-ágono y los polígonos que se pueden obtener de cada uno de estos dividiendo consecutivamente el número de sus lados. Esto ya se conocía en el tiempo de Euclides y, pareciera, que la creencia reinante, a partir de ese tiempo, es que el dominio de la geometría elemental no traspasa estos límites: al menos yo no conozco ningún intento exitoso para expandir la geometría en esta dirección. Por lo tanto, el descubrimiento de que, a parte de estos polígonos regulares, es posible construir geométricamente una multitud de otros polígonos, por ejemplo, un 17-ágono, me parece a mí digno de mención. Este descubrimiento es esencialmente un mero corolario de una teoría de gran envergadura que no ha sido finalmente completada todavía. En el momento que esta teoría sea completada esta será ofrecida al público.

C.F. Gauss de Braunschweig,
estudiante de matemáticas en Göttingen.

La teoría fue completada cinco años más tarde y publicada por Gauss en la séptima edición de las *Disquisitiones Arithmeticae*, las cuales aparecieron en 1801. Gauss probó que si el número n de lados de un polígono regular es de la forma $n = 2^{\alpha} p_1 \cdots p_k$, donde los p_i son primos de Fermat distintos, es decir, números primos de la forma $2^{2^m} + 1$, entonces el polígono podía ser construido con regla y compás. En lenguaje algebraico, esta afirmación significa que para los numeros n indicados, la ecuación $x^n + 1 = 0$ es soluble en radicales cuadrados.

La prueba del teorema de Gauss esta basada en una ingeniosa teoría algebraica la cual sirvió como piedra angular para la teoría de Galois creada treinta años mas tarde de que las *Disquisitiones Arithmeticae* fueran publicadas.

En la séptima sección de las *Disquisitiones Arithmeticae*, aparte de la teoría de la división del circulo, es decir, la teoría algebraica de las funciones circulares, se encuentra una pequeña observación, también hecha por Gauss, que se refiere a que el método que el mismo había desarrollado también es aplicable a ciertas funciones trascendentales; en particular, a las funciones relacionadas con integrales de la forma $\int \frac{dx}{\sqrt{1-x^4}}$.

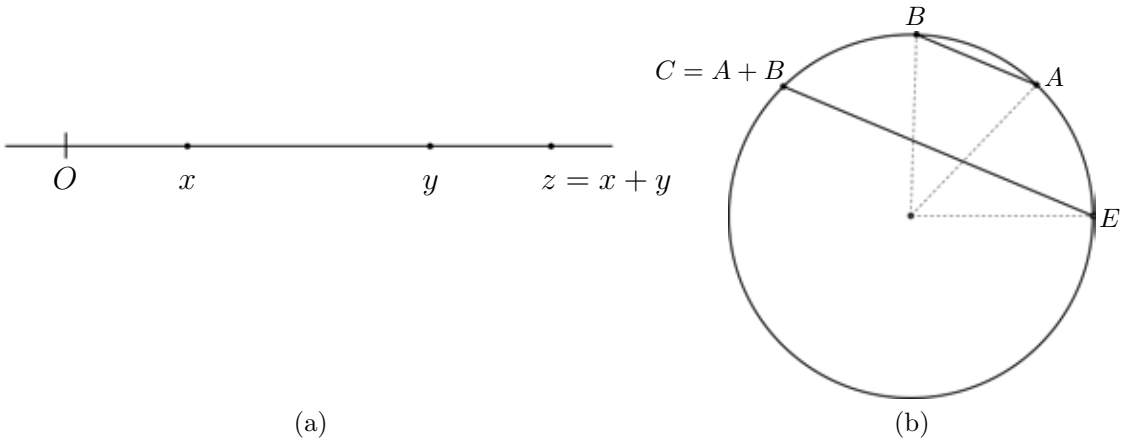
Esta observación se convirtió en el punto de partida para los estudios de Abel, quien en 1827 probó que para los mismos valores de n mencionados por Gauss, es posible dividir la lemniscata de Bernoulli con regla y compás en n partes iguales. Para hacer esto, Abel tuvo que mejorar considerablemente el método de Gauss y, lo que es más importante, crear una nueva disciplina matemática –la ***teoría de las funciones elípticas***.

La teoría de las funciones elípticas y su gemela geométrica –la ***teoría de las curvas elípticas***– ocupan uno de los más importantes lugares en matemáticas unificando varias de sus ramas. A pesar de su avanzadas edades, la teoría de las funciones elípticas y la teoría de curvas elípticas siguen siendo un dominio de las matemáticas vivo y de rápido desarrollo, es decir, estas teorías son una fuente inexhaustible de técnicas, problemas y conjeturas para los investigadores.

1.1. Adición de puntos en una cúbica

Una ***curva algebraica plana*** es el conjunto de puntos $(x, y) \in \mathbb{R}^2$ que satisfacen la ecuación $f(x, y) = 0$, donde $f(x, y)$ es un polinomio en dos variables distinto de cero.

Para algunas curvas planas, existen leyes naturales para la adición de puntos. Por ejemplo, tales leyes existen en cualquier recta y sobre la circunferencia unitaria $x^2 + y^2 = 1$. Para sumar puntos en una recta, debemos fijar un punto O en ésta y entonces la ***suma de los puntos*** X y Y puede ser definida como el punto Z tal que $\overrightarrow{OZ} = \overrightarrow{OX} + \overrightarrow{OY}$.



(a) (b)

Figura 1.1

También es natural definir la suma de los puntos $(\cos\alpha, \operatorname{sen}\alpha)$ y $(\cos\beta, \operatorname{sen}\beta)$ de la circunferencia unitaria (Figura 1.1 (b)) como el punto $(\cos(\alpha+\beta), \operatorname{sen}(\alpha+\beta))$. Esta ley de adición de puntos puede ser expresada de la siguiente manera. Sea E el punto $(1,0)$ y sean $A = (\cos\alpha, \operatorname{sen}\alpha)$ y $B = (\cos\beta, \operatorname{sen}\beta)$ puntos arbitrarios de la circunferencia unitaria. Dibujemos la recta que pasa por E y es paralela a la recta AB; la nueva recta dibujada interseca a la circunferencia en el punto C, definamos la ***suma de los puntos*** A y B como el punto C.

Esta definición sirve para cualquier ***cónica*** (una curva de segundo orden). A saber, fijando un punto E en una cónica y considerando el punto en el cual la recta que pasa por E y es paralela a la recta AB interseca a la cónica por segunda ocasión como la ***suma de los puntos*** A y B. La conmutatividad de la operación obtenida es clara, ya que los papeles de A y B son simétricos; el punto E sirve como elemento neutro o cero. Para encontrar el elemento $-A$, debemos dibujar la recta que pasa por A y es paralela a la tangente en E. Ahora veamos la asociatividad

$$(A+B)+C = A+(B+C)$$

que no es tan clara. Para probar esto, denotemos los puntos $A+B$ y $B+C$ como P y Q, respectivamente. La asociatividad es equivalente a la siguiente proposición: *Si A, B, C, E, P y Q son puntos en la cónica tales que $AB||EP$ y $BC||EQ$, entonces $AQ||CP$.* Pero esta proposición es un caso particular del ***teorema de Pascal*** sobre un hexágono inscrito en una cónica, lo cuál veremos un poco mas adelante.

Ejemplos. a) Para la parabola $y = x^2$ con el punto fijo $E = (0,0)$ la suma de los puntos $(x_1, y_1) = (x_1, x_1^2)$ y $(x_2, y_2) = (x_2, x_2^2)$ es el punto $(x_1+x_2, y_1+y_2+2x_1x_2) = (x_1+x_2, (x_1+x_2)^2)$.

b) Para la hipérbola $x^2 - y^2 = 1$ con el punto fijo $E = (1,0)$ la suma de los puntos (x_1, y_1) y (x_2, y_2) es el punto $(x_1x_2 + y_1y_2, y_1x_2 + y_2x_1)$. Bajo la parametrización de la hipérbola $x = \cosh t$ y $y = \operatorname{senh} t$ esta adición corresponde a la adición del parámetro t, es decir, para $(x_1, y_1) = (\cosh t_1, \operatorname{senh} t_1)$ y $(x_2, y_2) = (\cosh t_2, \operatorname{senh} t_2)$ la suma de estos puntos es $(\cosh(t_1+t_2), \operatorname{senh}(t_1+t_2))$.

Una cúbica es una curva algebraica plana $\sum_{i,j} a_{ij}x^iy^j = 0$, donde el valor máximo de $i+j$ es igual a 3. Para cualquier cúbica no singular, también existe una ley de adición de puntos bastante natural (más adelante discutiremos en detalle que es una ***cúbica no singular***). La ley de adición de puntos distintos en una cúbica no singular puede ser definida de la siguiente manera.

En una cúbica, fijemos un punto arbitrario E (éste resultará ser el elemento cero). Para sumar los puntos A y B, dibujemos la recta AB. Ésta interseca la cúbica en el punto X. El punto de intersección de la recta XE con la cúbica será la suma de A y B (Figura 1.2). En la definición de la adición usamos dos veces la siguiente propiedad de una cúbica:

Si una recta interseca una cúbica en dos puntos, entonces la recta interseca a la cúbica necesariamente en un punto más.

Esta propiedad parece ser más o menos obvia. De hecho, resolviendo la ecuación de la

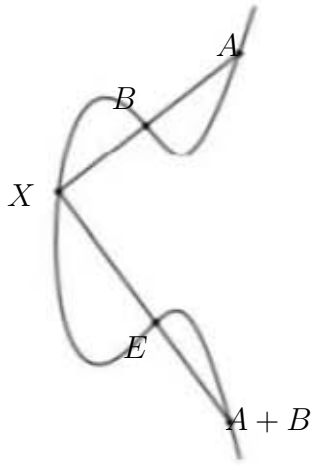


Figura 1.2

recta $ax + by + c = 0$ para x o y y sustituyendo los valores en la ecuación de la cúbica, obtenemos una ecuación de tercer grado. Por hipótesis esta ecuación tiene dos raíces reales y, por lo tanto, debe tener una tercera raíz real.

Pero en realidad todo esto no es tan simple. El problema no es solamente que el polinomio puede tener raíces múltiples sino que el grado del polinomio puede incluso resultar ser menor que 3. En el último caso la operación de adición es degenerada, lo que quiere decir, que no podemos sumar cualquier par de puntos, pero por ahora este caso no es de nuestro interés. Más adelante discutiremos como es posible definir la adición para tales puntos.

La conmutatividad de la operación obtenida es inmediata. También es fácil verificar que E es el elemento cero. La asociatividad de la operación no es obvia en absoluto. La igualdad $(A + B) + C = A + (B + C)$ es equivalente al hecho de que los puntos de intersección de las rectas que conectan a los puntos $(A + B)$ con C y a $(B + C)$ con A se encuentran sobre la cúbica.

Denotemos las rectas descritas como:

$$p_1 = AB, \qquad p_2 = E(B + C), \qquad p_3 = C(A + B),$$

$$q_1 = BC, \qquad q_2 = E(A + B), \qquad q_3 = A(B + C).$$

Asumiendo que todos los puntos de intersección de las rectas p_i y q_j son distintos por pares. Entonces la proposición que garantiza lo que se desea puede escribirse como:

Teorema 1.1.1. *Sea A_{ij} el punto de intersección de las rectas p_i y q_j, donde $1 \leq i, j \leq 3$ y supongamos que los puntos A_{ij} son distintos. Supongamos también que todos los puntos A_{ij}, excepto, quizás, A_{33}, se encuentran sobre una cúbica. Entonces A_{33} también se encuentra sobre la cúbica.*

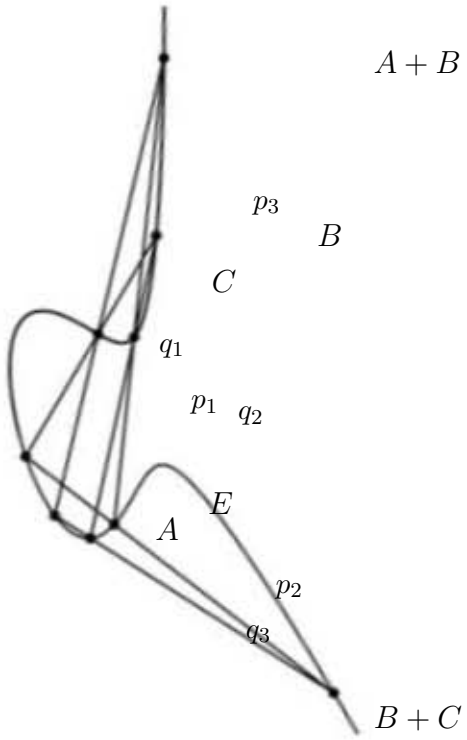


Figura 1.3

Demostración. Sean $p_i(x, y) = 0$ y $q_j(x, y) = 0$ las ecuaciones de las rectas p_i y q_j. Entonces la ecuación de tercer grado $p_1p_2p_3 = 0$ determina las rectas p_1, p_2 y p_3 y la ecuación $q_1q_2q_3 = 0$ determina las rectas q_1, q_2 y q_3. De ahí que, la cúbica $\alpha p_1p_2p_3 + \beta q_1q_2q_3 = 0$ pasa por todos los puntos A_{ij}.

Veamos ahora que de esta forma podemos representar la ecuación de cualquier cúbica que pase por ocho de los nueve puntos A_{ij}. Tomemos las rectas p_1 y q_1 como ejes coordenados, es decir, supongamos que $p_1(x, y) = y$ y $q_1(x, y) = x$. Si la cúbica dada está determinada por la ecuación $P(x, y) = 0$, entonces las funciones $P(0, y)$ y $yp_2(0, y)p_3(0, y)$ se anulan en los tres puntos A_{11}, A_{21} y A_{31} en el eje-y. Además estas funciones son polinomios de grado menor o igual a 3. Por lo tanto, $P(0, y) = \alpha yp_2(0, y)p_3(0, y)$. De forma similar, $P(x, 0) = \beta xq_2(x, 0)q_3(x, 0)$.

Considerando el polinomio

$$Q(x, y) = P(x, y) - \alpha yp_2(x, y)p_3(x, y) - \beta xq_2(x, y)q_3(x, y).$$

Claramente,

$$Q(0, y) = P(0, y) - \alpha yp_2(0, y)p_3(0, y) = 0.$$

El polinomio $a_0(y)+a_1(y)x+a_2(y)x^2+\cdots$ se anula en $x = 0$ si y sólo si $a_0(y)$ es idénticamente igual a cero, es decir, si el polinomio es divisible entre x.

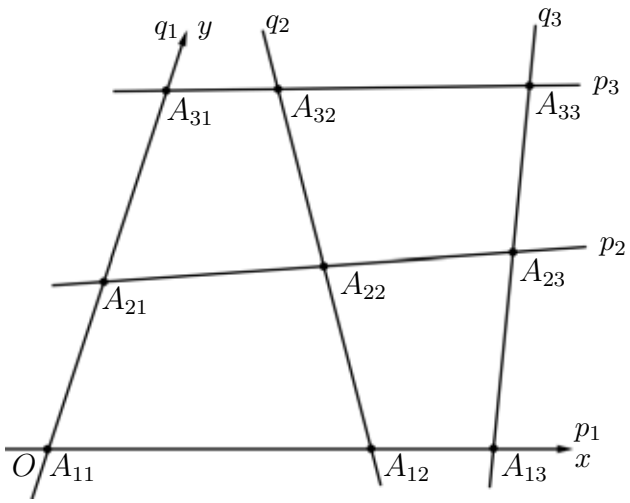


Figura 1.4

Argumentos similares muestran que $Q(x, y)$ también es divisible entre y, es decir, $Q(x, y) = xyQ_1(x, y)$. Como el grado $Q \leq 3$, tenemos que, $Q_1(x, y)$ es o una función lineal o una constante. Ahora, como los polinomios P, p_2p_3 y q_2q_3 se anulan en los puntos A_{22}, A_{23} y A_{32}, se tiene que el polinomio Q también se anula en estos puntos. Como $xy \neq 0$, la función lineal Q_1 tiene que anularse en estos puntos. Pero los puntos A_{22}, A_{23} y A_{32} no se encuentran sobre una recta y como una función lineal distinta de cero $f(x, y) = 0$ determina una recta. De ahí que, $Q_1 = 0$, es decir, $P = \alpha p_1p_2p_3 + \beta q_1q_2q_3$.

En particular, el punto $A_{3,3}$ se encuentra en la curva $P(x, y) = 0$. También hemos probado que cualquier cúbica que pase por los puntos A_{ij} esta dada por la ecuación

$$\alpha p_1p_2p_3 + \beta q_1q_2q_3 = 0.$$

En otras palabras, tales curvas constituyen una familia a un parámetro. La demostración del Teorema 1.1.1 está completa y, junto con esta, la comprobación de la asociatividad de la adición de puntos en una cúbica. □

Del Teorema 1.1.1, podemos obtener una demostración muy simple del siguiente:

Teorema de Pascal. *Los puntos de intersección de lados opuestos de un hexágono inscrito se encuentran sobre una recta.*

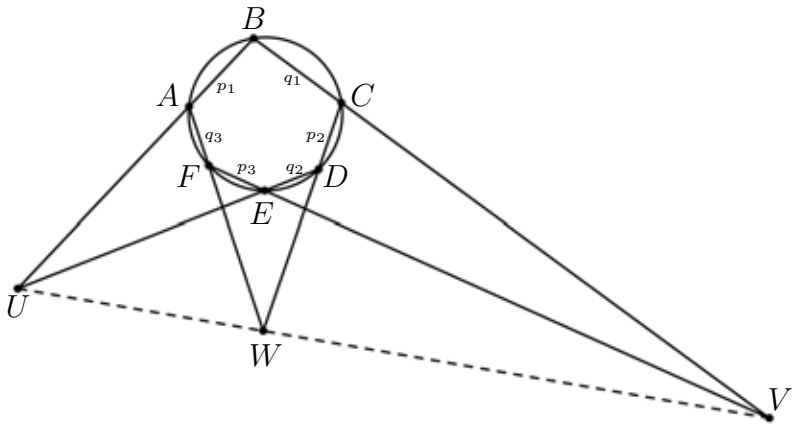


Figura 1.5

Demostración. Sean $p_1 = AB$, $q_1 = BC$, $p_2 = CD$, $q_2 = DE$, $p_3 = EF$ y $q_3 = AF$. Como cúbica tomemos la curva generada por la ecuación $Ql = 0$, donde $Q = 0$ es la ecuación de la circunferencia y $l = 0$ es la ecuación de la recta UV, donde U y V son los puntos de intersección de las rectas p_1 con q_2 y p_3 con q_1, respectivamente. Sea W el punto de intersección de las rectas p_2 y q_3. Ya que sabemos que todos los puntos distintos de W se encuentran en la curva $Ql = 0$, entonces el punto W también se encuentra en esta curva y además como este punto no se encuentra en la cónica tiene que pertenecer a la recta l. □

En lugar de la cónica $Q = 0$ podemos tomar cualquier curva de segundo grado. En particular, podemos suponer que $Q = pq$, donde p y q son funciones lineales. En este caso obtenemos

Teorema de Papus. *Si los puntos A, C y E están sobre una recta p, así como los puntos B, D y F están sobre una recta q, entonces las rectas AB y DE, BC y EF, AF y CD se intersecan en los puntos U, V y W, que se encuentran sobre una recta.* □

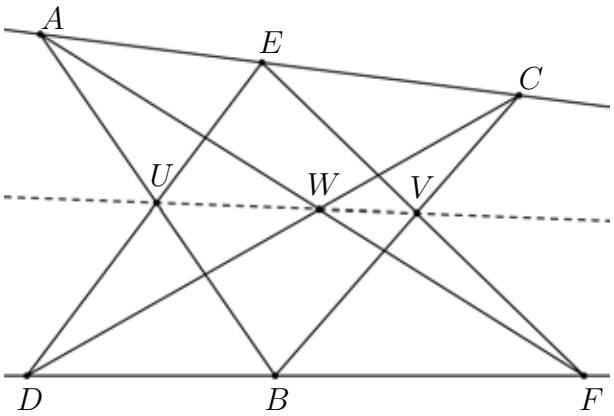


Figura 1.6

Algunas veces tendremos que aplicar el Teorema 1.1.1 cuando varios de los puntos A_{ij} coinciden. Por esta razón, tenemos que entender como debemos reformular el teorema de tal manera que resulte verdadero en tales circunstancias. Durante la demostración del teorema ocupamos dos veces la posibilidad de distinguir los puntos A_{ij}:

1) la función xy es distinta de cero en los puntos A_{22}, A_{23} y A_{32} y, por lo tanto la función lineal Q_1 se anula en ellos;

2) estos puntos no se encuentran en una recta; de ahí que, $Q_1 \equiv 0$. (Aquí usamos el signo $\equiv$ para expresar la noción "idénticamente igual a").

Durante la demostración del Teorema 1.1.1 solo hemos ocupado la restricción del polinomio P a las rectas p_i y q_j. Por esta razón, podemos esperar que en lugar de requerir que los puntos A_{ij} sean distintos, es suficiente suponer que

Si dos (o tres) de los puntos A_{ij} en las rectas p_i o q_j coinciden, entonces la restricción del polinomio P a estas rectas tendrá en el punto de coincidencia una raíz de multiplicidad dos (o tres).

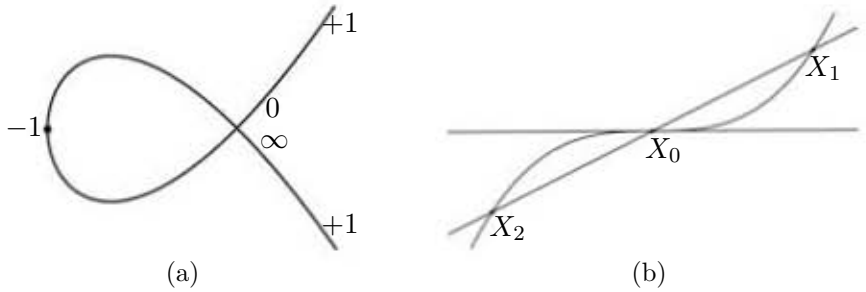
Esta modificación también se aplica al punto A_{33}.

Mostremos que la formulación del Teorema 1.1.1 puede ser modificada de la manera requerida. La demostración del hecho de que el polinomio $Q = P - \alpha p_1 p_2 p_3 - \beta q_1 q_2 q_3$ es divisible entre $xy = p_1 q_1$, funciona sin cambios. Si $A_{ij} = A_{ik} = A$, entonces en el punto A la restricción de P a p_i tiene una raíz de multiplicidad 2, la restricción de $p_1 p_2 p_3$ a p_i es idénticamente cero y la restricción a $q_1 q_2 q_3$ tiene una raíz de multiplicidad 2 ya que $q_j(A_{ij}) = 0$ y $q_k(A_{ik}) = 0$. Por lo tanto, la restricción de Q a p_i tiene una raíz de multiplicidad 2 en A.

Los argumentos para la recta q_j son similares y también en el caso de que los tres puntos coincidan. Por esta razón, es claro que la función lineal Q_1 sigue anulándose en los puntos A_{22}, A_{23} y A_{32}.

Si algunos de estos puntos coinciden, usaremos el hecho de que una función lineal distinta de cero en la recta puede tener una raíz de multiplicidad 2.

En la afirmación del Teorema 1.1.1, es claro que para la restricción de la función $\alpha p_1 p_2 p_3 + \beta q_1 q_2 q_3$ a la recta p_3 la multiplicidad de la raíz en el punto A_{33} es igual al número de rectas q_j que pasan por el punto A_{33}.



(a) (b)

Figura 1.7

Para la recta l tangente a la curva $F(X) = 0$ en el punto X_0 la restricción de F a l es de multiplicidad 2 en el punto X_0. De hecho, si movemos un punto X_1 sobre esta curva acercándonos al punto X_0, la restricción de la función F a la recta X_0X_1 tiene raíces X_0 y X_1. En la posición límite la recta X_0X_1 coincide con l y las raíces X_0 y X_1 se juntan para obtener una sola raíz de multiplicidad 2 (Figura 1.7 (a)). La fusión de tres raíces se da en la tangente de un punto de inflexión. En la Sección 1.3 discutiremos en detalle los puntos de intersección múltiple de una recta con una cúbica.

Para curvas de grado $n \geq 3$ el Teorema 1.1.1 puede ser generalizado de la siguiente manera.

Teorema 1.1.2. *Sean A_{ij} los puntos de intersección de las rectas p_i y q_j, con $1 \leq i, j \leq n$; supongamos que los puntos A_{ij} son distintos. Y supongamos que todos los puntos A_{ij}, donde $i + j \leq n + 3$, se encuentran en una curva de grado n. Entonces los otros puntos A_{ij} también están en la curva.*

Demostración. Tomemos las rectas p_1 y q_1 como ejes coordenados. Supongamos que la curva está dada por la ecuación $P_n(x, y) = 0$. Entonces $P_n(0, y) = \alpha p_1 \cdots p_n$ y $P_n(x, 0) = \beta q_1 \cdots q_n$. Consideremos el polinomio $Q_n = P_n - \alpha p_1 \cdots p_n - \beta q_1 \cdots q_n$. Es suficiente demostrar que $Q_n \equiv 0$. Como en el teorema anterior se tiene que Q_n es divisible entre $xy = p_1q_1$, es decir, $Q_n = p_1q_1Q_{n-2}$. Así que resta probar que el polinomio distinto de cero Q_{n-2} de grado no mayor a $n-2$ no puede anularse en los puntos A_{ij}, donde $i, j \geq 2$ e $i+j < n+3$. (Probaremos esta proposición por inducción sobre n.)

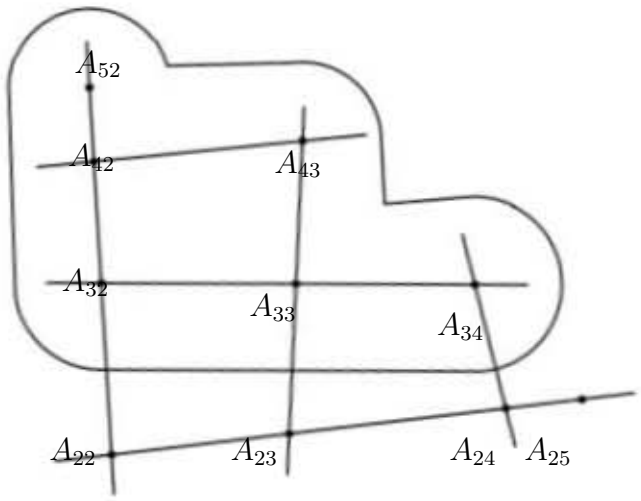


Figura 1.8

Supongamos que tal polinomio distinto de cero Q_{n-2} existe. Su restricción a la recta p_2 se anula en los $n-1$ puntos $A_{22}, A_{23}, \cdots, A_{2n}$. Por lo tanto, la restricción de Q_{n-2} a esta recta es idénticamente cero, es decir, $Q_{n-2} = p_2Q_{n-3}$. El polinomio Q_{n-3} se anula en los puntos formando una configuración semejante de menor tamaño. Estos argumentos ilustran

el paso inductivo. La base de la inducción ($n = 3$) esta ya considerada en la demostración del Teorema 1.1.1. □

El método de probar teoremas geométricamente usando la familia de curvas

$$p_1p_2p_3 + \mu q_1q_2q_3 = 0 \tag{1.1}$$

fue desarrollado por el matemático alemán **Julius Plücker** (1801-1868). La idea de representar una terna de rectas como una cúbica degenerada resulta ser bastante fructifera. Esta representación permitió reducir la demostración de varios complicados teoremas geométricos a la ingeniosa selección del coeficiente μ en (1.1); esta μ comenzó a aparecer regularmente en los papeles de Plücker.

Tal algebraización de la geometría no convenció a todo el mundo. **Jacob Steiner** (1796-1863) – uno de los geómetras más prominentes de estos tiempos – se rehusó a atribuir signos a cantidades geométricas y en vez de eso prefirió considerar distintas variantes de las posiciones de los puntos. A pesar de la complicada manera de tratar los temas que Steiner seleccionó, el consiguió en muchas ocasiones obtener resultados más finos y más profundos que Plücker. Steiner se rehusó a los nuevos métodos algebraicos en geometría.

1.2. Rectas y curvas en el plano proyectivo

En la sección anterior mencionamos que la adicción de puntos en una cúbica no está definida generalmente para todos los puntos. Ilustremos esto usando la curva

$$y^2 = x(x-1)(x-2). \tag{1.2}$$

Sustituyendo la ecuación de la recta $x = \frac{1}{2}$ en (1.2) obtenemos $y^2 = \frac{3}{8}$. El grado de esta ecuación es igual a 2, no 3. De ahí que, la recta $x = \frac{1}{2}$ interseca a la curva (1.2) solamente en dos puntos. Y los puntos de intersección no son multiples. Un intento de sumar estos puntos no será posible.

Si (x, y) es un punto de la curva (1.2), entonces

$$\lim_{x\to\infty} \frac{x^2}{y^2} = \lim_{x\to\infty} \frac{x}{(x-1)(x-2)} = 0.$$

Surge la suposición de que ambas la curva (1.2) y la recta $x = \frac{1}{2}$ pasan por un punto infinito en dirección del *eje y*. El punto de intersección faltante puede resultar estar situado en la recta al infinito. Intentemos aumentar la colección de puntos del plano ordinario con puntos al infinito, pensando en estos puntos como puntos de intersección de rectas paralelas. Otra razón para hacer esto es que de otra manera nuestras formulaciones y demostraciones de los teoremas de Papus y Pascal serían inexactas. De hecho, hemos asumido siempre que las rectas que consideramos se intersecan. Pero estas rectas también pueden ser paralelas. Desde luego, podemos considerar, por separado, los casos cuando ciertas rectas se intersecan

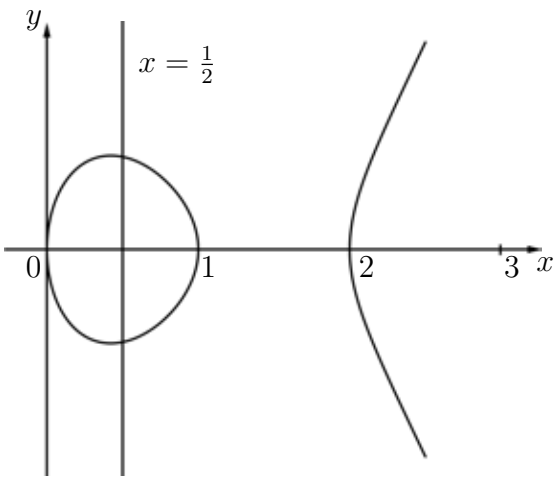


Figura 1.9

y ciertas rectas son paralelas. Pero esto es un poco más fastidioso, pues cada caso requiere no solo una formulación por separado, sino también una demostración tal vez diferente.

Tomemos un plano π en un espacio tridimensional y consideremos un punto O fuera del plano π. A cada punto $A \in \pi$ le asociamos la recta OA. Para una recta $l \in \pi$ no le corresponde asignarle todo el plano Ol sino solamente la recta l' que pasa por O y es paralela a la recta l. Si la recta $l_1 \in \pi$ es paralela a l, entonces los planos Ol y Ol_1 se intersecan a lo largo de la recta l'. También es claro que si el punto A corre a lo largo de la recta l al infinito, entonces la posición límite de la recta OA es la recta l'.

Definiremos ***el plano proyectivo real*** $\mathbb{R}P^2$ de la siguiente manera. Los ***puntos*** en $\mathbb{R}P^2$ son rectas que pasan por O. Las ***rectas*** en $\mathbb{R}P^2$ son los planos que pasan por el punto O descrito anteriormente. Construido así, las rectas paralelas al plano π corresponden a puntos infinitos de π y el plano paralelo a π corresponde a la recta al infinito en π. En el plano proyectivo, dos rectas cualquiera se intersecan en un punto. Las rectas proyectivas correspondientes a rectas paralelas en π se intersecan en un punto de la recta al infinito. Cuando nos olvidamos de π los puntos infinitos, no difieren de los otros puntos.

Para trabajar con curvas algebraicas tenemos que introducir coordenadas en el plano proyectivo. Asumiremos que O es el origen del sistema coordenado en el espacio tridimensional y que el plano π esta dado por la ecuación $z = 1$. Una recta que pasa por O consiste en los puntos de la forma $(\lambda x, \lambda y, \lambda z)$, donde x, y y z están fijos y λ corre sobre $\mathbb{R}$. Por lo tanto, podemos considerar las ternas de números reales distintas del cero (x, y, z) como ***puntos*** en $\mathbb{R}P^2$, aquí las ternas (x, y, z) y $(\lambda x, \lambda y, \lambda z)$, $\lambda \neq 0$, son consideradas ***equivalentes***, y $\mathbb{R}P^2$ es el espacio cociente del conjunto de ternas módulo esta relación de equivalencia. La recta al infinito estará dada por la ecuación $z = 0$.

En la definición de plano proyectivo x, y, z y λ pueden elegirse como números complejos. De esta manera obtenemos una definición del ***plano proyectivo complejo*** $\mathbb{C}P^2$. La geometría de curvas algebraicas en $\mathbb{C}P^2$ es considerable más simple que en $\mathbb{R}P^2$. Este

fenómeno está relacionado con el hecho de que sobre $\mathbb{C}$ cada polinomio de grado n tiene precisamente n raíces (contando multiplicidades).

A la curva

$$y^2 = x(x-1)(x-2) \tag{1.3}$$

podemos asignarle la curva

$$y^2 z = x(x-z)(x-2z) \tag{1.4}$$

en el plano proyectivo. El hecho de que la ecuación (1.4) defina una curva en el plano proyectivo dependerá de que los puntos (x, y, z) y $(\lambda x, \lambda y, \lambda z)$ satisfagan simultáneamente (1.4) o no. Más aún, en el plano π dado por la ecuación $z = 1$, ambas ecuaciones (1.3) y (1.4) coinciden.

De igual forma, a cualquier curva algebraica $\sum a_{ij} x^i y^j = 0$ podemos asignarle la curva

$$\sum a_{ij} x^i y^j z^{n-i-j} = 0, \qquad \text{donde} \quad n = \text{máx}(i+j),$$

en el plano proyectivo. Ahora podemos verificar nuestra hipótesis de que la recta $x = \frac{1}{2}z$ y la curva $y^2 z = x(x-z)(x-2z)$ se encuentran en el punto al infinito en la dirección del *eje* y. Sustituyendo la expresión $x = \frac{z}{2}$ en la ecuación de la curva tenemos que $y^2 x = \frac{3z^3}{8}$. Esta ecuación tiene tres tipos de soluciones, a saber, (1) $z = 0$, donde y es arbitrario; (2) $y = kz$ y (3) $y = -kz$, donde $k = \sqrt{3/8}$ y z es arbitrario. En otras palabras, cada vez que obtenemos una ***familia de soluciones***, la familia corresponde a un punto de $\mathbb{C}P^2$.

Por lo tanto, la recta $x = \frac{1}{2}z$ en el plano proyectivo, interseca a la curva considerada en tres puntos: $(\frac{1}{2}, \sqrt{\frac{3}{8}}, 1), (\frac{1}{2}, -\sqrt{\frac{3}{8}}, 1), (0, 1, 0)$. El tercer punto es el punto al infinito en dirección del *eje* y.

Incluso es posible hacer un esbozo de como se ven la recta $x = \frac{1}{2}z$ y la curva considerada en una vecindad del punto infinito $(0, 1, 0)$. Para esto, en vez del plano $z = 1$ debemos tomar un plano que pase por el punto (0,1,0) y que no pase por el origen. Tomando, por ejemplo, el plano $y = 1$. En este, obtenemos la curva $z = x(x-z)(x-2z)$. Para x y z pequeños nuestra curva luce casi como la curva $z = x^3$ (Figura 1.10).

El paso al plano proyectivo es de ayuda no solamente en el caso anterior. Mostremos, por ejemplo, que cualquier recta en el plano proyectivo pertenece enteramente a una cúbica o interseca a esta (contando multiplicidades) en precisamente tres números complejos; si consideramos solamente puntos reales, entonces ésta interseca a la cúbica en sólo uno o en tres puntos.

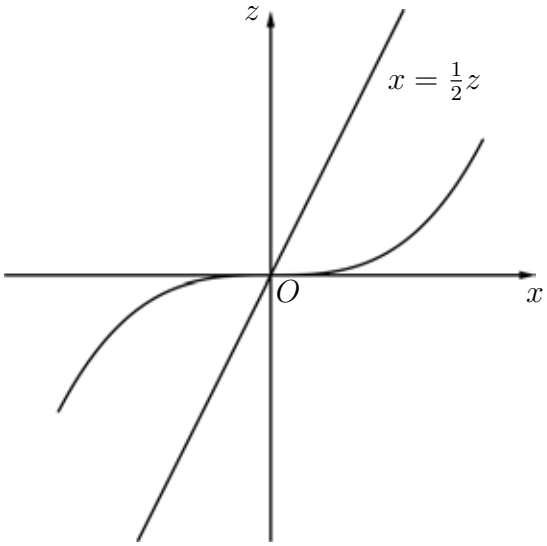


Figura 1.10

Podemos encontrar los puntos de intersección de la recta $ax + by + cz = 0$ y la cúbica $\sum_{i+j+k=3} a_{ij} x^i y^j z^k = 0$ en el plano proyectivo de la siguiente manera. Uno de los números a, b, c es distinto de cero. Sea, por ejemplo, $c \neq 0$. Entonces $z = \alpha x + \beta y$, donde $\alpha = -a/c$ y $\beta = -b/c$ (el caso $\alpha = \beta = 0$ no esta excluido). Insertando esta expresión en la ecuación de la cúbica obtenemos una ecuación de la forma $Q = 0$, donde $Q(x, y) = \sum b_p x^p y^{3-p}$. Los dos casos siguientes son posibles:

1) Todos los coeficientes b_p son cero. Entonces la recta $ax + by + cz = 0$ esta contenida enteramente en la curva, es decir, Q es divisible entre $ax + by + cz$.

2) No todos los coeficientes b_p son cero. Entonces

$$Q(x, y) = bx^r y^s (x - t_1 y) \cdots (x - t_m y),$$

donde $r + s + m = 3$. Para el factor x^r le corresponde el punto de intersección $(0, 1, \beta)$ de multiplicidad r; para el factor y^s le corresponde el punto $(1, 0, \alpha)$ de multiplicidad s; y para el factor $(x - t_i y)$ le corresponde el punto $(t_i, 1, \alpha t_i + \beta)$.

Un polinomio cúbico Q con coeficientes reales puede tener tres raíces reales o una. Por lo tanto, cualquier recta en el plano proyectivo interseca a una cúbica en tres puntos reales o en un punto (contando multiplicidades). Por lo tanto, casi tenemos el manejo claro de como sumar puntos distintos en una cúbica. Solo tendremos problemas con puntos que se intersecan en si mismos o con puntos cúspide. El problema es que cualquier recta que pase por uno de estos puntos tiene una intersección múltiple con la curva. Por lo tanto, tomando la suma de dicho punto con cualquier otro punto nunca obtendremos un nuevos puntos. Más adelante en la Sección 1.5 discutiremos con detalle las cúbicas singulares.

Por ahora, veamos como hacer la operación de adición de puntos que coinciden y entender el significado geométrico de la multiplicidad de estos puntos de intersección.

1.3. Las tangentes y puntos de inflexión

Para sumar puntos A y B en una cúbica tenemos que dibujar la recta AB. ¿Como debemos proceder si los puntos A y B coinciden? Supongamos que el punto A está fijo y que el punto B se mueve hacia A a lo largo de la curva dada. Entonces, bajo ciertas condiciones (por ejemplo, el punto A debe ser no singular), la recta AB tiende a una recta fija, la tangente en A. Por lo tanto, para encontrar la suma $A+A$, en vez de la recta AB, deberemos dibujar la tangente en A (con la condición de que la tangente quede definida de manera única en este punto).

Si la curva que pasa por los puntos A y B está dada por la ecuación $F=0$, entonces la restricción de F a AB tiene raíces en los puntos A y B. En la posición límite, cuando los puntos A y B coinciden, la restricción de F a A tiene una raíz múltiple, por lo tanto, la restricción de F a la tangente tiene una raíz múltiple en el punto de tangencia. Esta propiedad puede ser usada para obtener la ecuación de la tangente.

Sea $P=(p_1,p_2,p_3)$ un punto perteneciente a la curva $F=0$, es decir, $F(P)=0$, y sea $X=(x_1,x_2,x_3)$ un punto arbitrario. Los puntos de la recta proyectiva PX son de la forma $\lambda P+\mu X$. Los puntos de esta recta distintos de X son de la forma $P+tX$. Consideremos la restricción de F a la recta PX como una función de t. En el caso que nos interesa F es un polinomio de grado 3, por esta razón,

$$F(P+tX)=F(P)+at+bt^2+ct^3=Q(t),$$

donde $F(P)=0$, $a=\sum F_i(P)x_i$, y $b=\frac{1}{2}\sum F_{ij}(P)x_ix_j$ (aquí F_i es la derivada parcial con respecto a la i-esima variable). El punto P corresponde al valor $t=0$. El polinomio $Q(t)$ tiene una raíz múltiple en cero si $a=0$, es decir, $\sum F_i(P)x_i=0$.

Un punto P para el cual al menos uno de los números $F_i(P)$ es distinto de cero es llamado un ***punto no singular*** de la curva. Para un punto P no singular la ecuación $\sum F_i(P)x_i=0$ determina de manera única una recta l, la ***recta tangente a la curva*** en P.

La tangente a la curva está definida geométricamente sea cual sea el sistema coordenado. Por ahora no es claro que la definición de tangente y la singularidad de un punto no dependan de la elección del sistema coordenado. Probemos la invariancia de estas definiciones. Veamos que pasa bajo el cambio de coordenadas $(x_1,x_2,x_3)\mapsto(u_1,u_2,u_3)$, donde $x_i=\sum_j a_{ij}u_j$. Sea $G(u_1,u_2,u_3)=F(x_1(u),x_2(u),x_3(u))$. Entonces

$$G_j=\frac{\partial G}{\partial u_j}=\sum_i\frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial u_j}=\sum_i F_ia_{ij}.$$

Como la matriz $J=(a_{ij})$ es no singular, la terna (G_1,G_2,G_3) es distinta de cero si y sólo si la terna (F_1,F_2,F_3) es distinta de cero. Si f y g son los renglones (F_1,F_2,F_3) y (G_1,G_2,G_3), y x y u son las columnas $(x_1,x_2,x_3)^T$ y $(u_1,u_2,u_3)^T$, respectivamente, entonces $x=Ju$ y $g=fJ$. De ahí que $gu=(fJ)(J^{-1}x)=fx$ y las ecuaciones $fx=\sum F_ix_i=0$ y $gu=\sum G_ju_j=0$ determinan la misma recta.

Para pasar de las coordenadas proyectivas (x_1, x_2, x_3) a las coordenadas Cartesianas (x_1, x_2) tenemos que hacer $x_3 = 1$. Para satisfacer la condición $x_3 = 1$ para un punto de la recta P, tenemos que expresar los puntos de la recta PX de la forma $P + t(X - P)$. El desarrollo

$$F(P + t(X - P)) = \sum F_i(P)(x_i - p_i)t + \cdots$$

nos permite expresar la ecuación de la tangente en la forma

$$F_1(P)x_1 + F_2(P)x_2 = F_1(P)p_1 + F_2(P)p_2.$$

En las coordenadas proyectivas, es decir, para una función homogénea F, la expresión $\sum F_i(P)p_i$ es igual a cero. La razón es que para cualquier polinomio homogéneo F de grado n se cumple la ***fórmula de Euler***

$$\sum F_i(X)x_i = nF(X).$$

Por ejemplo para el monomio $M = x_1^{p_1}x_2^{p_2}x_3^{p_3}$, donde $\sum p_i = n$, es claro que para p_i positivo tenemos $x_i\frac{\partial M}{\partial x_i} = p_i M$.

Sea P un punto no singular de la curva $F = 0$, entonces la tangente l en P está bien definida. La restricción de F a l tiene una raíz múltiple en P. Si la multiplicidad de esta raíz es mayor o igual a 3, entonces P es llamado ***punto de inflexión***. En otras palabras, la condición $a = \sum F_i(P)x_i = 0$ debe implicar que $b = \frac{1}{2}\sum F_{ij}(P)x_i x_j = 0$, es decir, la cuadrática $\sum F_{ij}(P)x_i x_j = 0$ debe contener la recta $\sum F_i(P)x_i = 0$.

Recordemos que el polinomio de grado 2, $x^T Ax$ (expresado aquí en su forma matricial) es divisible entre la función lineal $x^T l$ sólo si $x^T Ax = x^T lm^T x$ para alguna m. Estos significa que la matriz $A = lm^T$ es el producto de una columna por un renglón, es decir, es de rango 1 (estando en el caso en el cual l y m no son cero). En particular, $\det A = 0$. Esto es, si P es un punto de inflexión, entonces $\det(F_{ij}(P)) = 0$.

Mostremos que para un punto no singular de la curva el recíproco también es cierto, es decir, si P es un punto no singular y el $\det(F_{ij}(P)) = 0$, entonces P es un punto de inflexión. Consideremos la cuadrática $\sum F_{ij}(P)x_i x_j = 0$. El punto P pertenece a ésta ya que por la fórmula de Euler

$$\sum F_{ij}(P)p_i p_j = 2\sum F_j(P)p_j = 6F(P) = 0.$$

Más aun, la recta $\sum_i F_i(P)x_i = 0$ es la tangente a esta cuadrática en P. De hecho, la ecuación de la tangente a la cuadrática $\sum F_{ij}(P)x_i x_j = 0$ en P es de la forma

$$\sum F_{ij}(P)x_i p_j = 0$$

y por la fórmula de Euler $\sum_{i,j} F_{ij}(P)x_i p_j = 2\sum_i F_i(P)x_i$. Aún no hemos usado que la cuadrática es degenerada; en cualquier caso la tangente a la curva en P es también la tangente a la cuadrática $\sum_{i,j} F_{ij}x_i x_j = 0$. Pero en el caso cuando la cuadrática consiste de un par de rectas ésta contiene enteramente a la tangente.

Resumamos. El conjunto de puntos de intersección de la curvas $F = 0$ y $H = 0$, donde $H(X) = \det(F_{ij}(X))$, contiene a todos los puntos de inflexión de la curva $F = 0$ (de estos puntos de intersección sólo los puntos singulares de esta curva no podrán ser puntos de inflexión). La curva $H = 0$ es llamada la ***curva de Hesse*** o el ***Hessiano*** de $F = 0$. Si F es un polinomio homogéneo de grado n, entonces F_{ij} es un polinomio homogéneo de grado $n - 2$. Por lo tanto, H es un polinomio homogéneo de grado $3(n - 2)$. Para un polinomio cúbico F el polinomio H también es cúbico.

La invarianza de la noción de punto de inflexión y la de la curva de Hesse pueden ser probadas casi de la misma forma en que probamos la invarianza de la tangente.

Sea $G = (u_1, u_2, u_3) = F(x_1(u), x_2(u), x_3(u))$, donde $x_i = \sum a_{ij}u_j$. Entonces

$$G_{pq} = \frac{\partial^2 G}{\partial u_p \partial u_q} = \sum_{i,j} \frac{\partial^2 F}{\partial x_i \partial x_j} \frac{\partial x_i}{\partial u_p} \frac{\partial x_j}{\partial u_q} = \sum_{i,j} a_{ip} F_{ij} a_{jq},$$

es decir, $(G_{pq}) = J^T(F_{ij})J$, donde $J = (a_{ij})$. Por lo tanto, $\det(G_{pq}) = (\det J)^2 \det(F_{ij})$. Por esta razón, la condiciones $\det(F_{ij}) = 0$ y $\det(G_{pq}) = 0$ son equivalentes.

La búsqueda de puntos de inflexión de la curva se reduce a la búsqueda de los puntos de intersección de la curva con el Hessiano. Entonces tenemos que encontrar los puntos de intersección de estas dos curvas. Ya hemos hecho esto en el caso cuando una de estas curvas es una recta. La ecuación de la recta nos permite expresar una variable en términos de la otra. Sustituyendo esta expresión en la ecuación de la curva podemos excluir una de las variables. Para curvas de grado arbitrario también podemos excluir una variable, pero el proceso es más difícil. Para hacer la representación más clara, primero deberemos considerar las curvas en las coordenadas Cartesianas (x, y) para después pasar a las coordenadas proyectivas (x, y, z).

Para simplificar, nos centraremos en las curvas de grado tres. Es posible expresar los polinomios de grado tres $F(x, y)$ y $H(x, y)$ de la forma

$$\begin{aligned} F(x, y) &= a_0 y^3 + a_1(x)y^2 + a_2(x)y + a_3(x), \\ H(x, y) &= b_0 y^3 + b_1(x)y^2 + b_2(x)y + b_3(x), \end{aligned}$$

donde $a_k(x)$ y $b_k(x)$ son polinomios de grado no mayor a k, $0 \leq k \leq 3$. Si (x_0, y_0) es un punto común de las curvas $F(x, y) = 0$ y $H(x, y) = 0$, entonces los polinomios $f(y) = a_0 y^3 + a_1 y^2 + a_2 y + a_3$ y $h(y) = b_0 y^3 + b_1 y^2 + b_2 y + b_3$, donde $a_k = a_k(x_0)$ y $b_k = b_k(x_0)$, tienen una raíz común y_0; el recíproco también es cierto: si los polinomios tienen una raíz común y_0, entonces las curvas tienen un punto en común (x_0, y_0).

Sobre $\mathbb{C}$, dos polinomios tienen una raíz común si y sólo si estos tienen un divisor común que no sea constante (sobre $\mathbb{R}$ el divisor común puede no tener raíces). Si $a_0 b_0 \neq 0$, entonces los polinomios $f(y)$ y $h(y)$ tienen un divisor común d si y sólo si existen polinomios h_1 y f_1 tales que $f = df_1$, $h = dh_1$, por lo que $fh_1 = hf_1$, donde los grados de h_1 y f_1 son menores que los grados de $H(x, y)$ y $F(x, y)$, respectivamente, así grado $f_1 <$ grado f. Entonces todos los divisores primos de f deben encontrarse en la factorización prima de hf_1; además, estos tendrán el mismo grado; pero no todos estos se encontraran en la factorización de f_1.

La restricción $a_0 b_0 \neq 0$ es, sin duda, muy fuerte pero en el caso proyectivo es fácil de satisfacer.

Sean $h_1(y) = u_0 y^2 + u_1 y + u_2$ y $f_1(y) = v_0 y^2 + v_1 y + v_2$. La igualdad $f h_1 = h f_1$ puede ser expresada de la forma

$$\begin{array}{rrrrrrl}
a_0u_0 & & & -b_0v_0 & & & = 0,\\
a_1u_0 & +a_0u_1 & & -b_1v_0 & -b_0v_1 & & = 0,\\
a_2u_0 & +a_1u_1 & +a_0u_2 & -b_2v_0 & -b_1v_1 & -b_0v_2 & = 0,\\
a_3u_0 & +a_2u_1 & +a_1u_2 & -b_3v_0 & -b_2v_1 & -b_1v_2 & = 0,\\
 & a_3u_1 & +a_2u_2 & & -b_3v_1 & -b_2v_2 & = 0,\\
 & & a_3u_2 & & & -b_3v_2 & = 0.
\end{array}$$

Este sistema homogéneo de ecuaciones lineales con respecto a u y v tiene una solución distinta de cero si y sólo si su determinante se anula, equivalentemente si el siguiente determinante se anula

$$\begin{vmatrix}
a_0 & a_1 & a_2 & a_3 & & \\
 & a_0 & a_1 & a_2 & a_3 & \\
 & & a_0 & a_1 & a_2 & a_3\\
b_0 & b_1 & b_2 & b_3 & & \\
 & b_0 & b_1 & b_2 & b_3 & \\
 & & b_0 & b_1 & b_2 & b_3
\end{vmatrix} = 0. \tag{1.5}$$

Este determinante es llamado el ***resultante*** de los polinomios f y h. Los coeficientes a_k y b_k dependen de x y, por lo tanto, el determinante (1.5) es un polinomio R de x, quizás el polinomio cero. Para cada raíz x_0 del polinomio $R(x)$ las curvas $F(x, y) = 0$ y $H(x, y) = 0$ tienen un punto en común (x_0, y_0) (Observemos que en el caso real el hecho de que x_0 sea real no implica necesariamente que y_0 lo sea). Si el polinomio $R(x)$ es idénticamente cero, entonces las curvas tienen una componente común.

Ahora, repitamos los argumentos anteriores para las coordenadas proyectivas. Expresemos primero los polinomios F y H en la forma

$$\begin{aligned}
F(x, y, z) &= a_0 z^3 + a_1(x, y) z^2 + a_2(x, y) z + a_3(x, y),\\
H(x, y, z) &= b_0 z^3 + b_1(x, y) z^2 + b_2(x, y) z + b_3(x, y),
\end{aligned}$$

donde a_k y b_k son polinomios homogéneos de grado k con $0 \leq k \leq 3$, es posible escoger coordenadas tales que las curvas $F = 0$ y $H = 0$ no pasen por el punto $(0, 0, 1)$. Entonces la condición que necesitamos, $a_0 b_0 \neq 0$, se puede satisfacer. En el caso proyectivo el determinante (1.5) es un polinomio en dos variables, $R(x, y)$. Probemos que R es el polinomio cero o un polinomio homogéneo de grado 9 (para curvas de grados m y n el grado de $R(x, y)$

es igual a mn). De hecho,

$$R(\lambda x, \lambda y) = \begin{vmatrix} a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 & & \\ & a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 & \\ & & a_0 & \lambda a_1 & \lambda^2 a_2 & \lambda^3 a_3 \\ b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 & & \\ & b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 & \\ & & b_0 & \lambda b_1 & \lambda^2 b_2 & \lambda^3 b_3 \end{vmatrix}.$$

Multipliquemos el segundo y el quinto renglón por λ y el tercero y sexto por λ^2. Como resultado, tenemos una matriz para $R(x, y)$ en la cual la k-esima columna esta multiplicada por λ^k. Por lo tanto, $\lambda^6 R(\lambda x, \lambda y) = \lambda^{15} R(x, y)$, es decir, $R(\lambda x, \lambda y) = \lambda^9 R(x, y)$, luego R es homogéneo de grado 9.

El polinomio distinto de cero $R(x, y)$ puede ser representado en la forma $\prod_{i=1}^{9}(y_i x - x_i y)$, donde x_i y y_i no se anulan simultáneamente. Para cada uno de los nueve pares (x_i, y_i) existe z_i tal que (x_i, y_i, z_i) es un punto de intersección de las curvas $f = 0$ y $h = 0$. El polinomio $R(x, y)$ puede tener raíces múltiples, es decir, ciertos pares (x_i, y_i) pueden ser proporcionales. Por lo tanto, no todos los pares de cúbicas tienen nueve puntos en común distintos. Pero en el plano proyectivo complejo cualesquiera dos cúbicas tienen al menos un punto en común, por esta razón,

cualquier cúbica no singular tiene un punto de inflexión (y entonces, nueve puntos de inflexión, contando multiplicidades).

Ésta es precisamente la propiedad que necesitaremos en la siguiente sección.

1.4. Cúbicas no singulares. Formas normales

Una curva cúbica es llamada ***no singular*** si todos su puntos son no singulares. En esta sección probaremos que sobre $\mathbb{C}$ la ecuación de una cúbica no singular puede ser reducida por cambios lineales de coordenadas homogéneas a cualquiera de las siguientes formas:

1) $y^2 z = x^3 + pxz^2 + qz^3$ (la forma de Weierstrass);

2) $x^3 + y^3 + z^3 = 3\lambda xyz$.

En el primer caso el polinomio $x^3 + px + q$ no tiene raíces múltiples (de no ser así la curva seria singular) y en el segundo caso $\lambda^3 \neq 1$ (de otra forma la curva consistiría de tres rectas, como veremos más adelante).

Consideremos una cúbica no singular $\sum a_{ij} x^i y^j z^{3-i-j} = 0$ sobre $\mathbb{C}$. En la sección anterior mostramos que esta tiene un punto de inflexión. Podemos suponer que las coordenadas del punto de inflexión son $(0, 1, 0)$ y que la tangente en este punto esta dada por la ecuación $z = 0$.

En otras palabras, la restricción de la función $F(x,y,z) = \sum a_{ij}x^i y^j z^{3-i-j}$ a la recta $z=0$ (es decir, el polinomio $a_{30}x^3 + a_{21}x^2 + a_{12}xy^2 + a_{03}y^3$) tiene una raíz $x=0$ de multiplicidad 3. De aquí se sigue que $a_{21} = a_{12} = a_{03} = 0$ pero $a_{30} \neq 0$, ya que de otra manera la curva considerada tendría que contener completamente a la recta $z=0$. La tangente en $(0,1,0)$ está dada por la ecuación

$$F_x(0,1,0)x + F_y(0,1,0)y + F_z(0,1,0)z = 0.$$

Por esta razón, $F_x(0,1,0) = F_y(0,1,0) = 0$ pero $F_z(0,1,0) \neq 0$, ya que de otra manera el punto $(0,1,0)$ debería ser singular. El valor del polinomio homogéneo $F_z(x,y,z)$ de grado 2 en $(0,1,0)$ es igual a a_{02} y podemos suponer que $a_{02} = 1$. En coordenadas Cartesianas la ecuación de la curva toma la forma

$$y^2 - 2(ax+b)y + P_3(x) = 0,$$

donde P_3 es un polinomio de grado tres. Haciendo los cambios de variables $y_1 = y - ax - b$ tenemos que

$$y_1^2 - (ax+b)^2 + P_3(x) = 0,$$

es decir, $y_1^2 = Q_3(x)$, donde $Q_3(x) = (ax+b)^2 - P_3(x)$ es un polinomio de grado tres. Bajo un cambio de la forma $x = \lambda x_1 + \mu$ el polinomio Q_3 puede ser reducido a la forma $x_1^3 + px_1 + q$.

El polinomio Q_3 no tiene raíces múltiples, ya que de otra manera la ecuación de la curva se podría haber reducido a la forma $y^2 = x^2(\alpha x + \beta)$ y para curvas así el origen es un punto singular.

En la sección anterior probamos que cualquier cúbica tiene 9 puntos de inflexión, contando multiplicidades, pero no podemos determinar cuando estos son o no son distintos. Si la ecuación de la cúbica no singular está expresada en la forma $y^2 = Q_3(x) = x^3 + ax^2 + bx + c$, entonces es fácil encontrar los puntos de intersección de la cúbica con el Hessiano y mostrar que todos son distintos.

Teorema 1.4.1. *La cúbica no singular $y^2 = Q_3(x)$ en $\mathbb{C}P^2$ tiene precisamente 9 puntos de inflexión distintos.*

Demostración. Podemos suponer que el polinomio Q_3 tiene una raíz $x=0$, es decir, que la curva considerada está dada por la ecuación $f(x,y)=0$, donde $f(x,y) = y^2 - x^3 - ax^2 - bx$. Como el polinomio Q_3 no tiene raíces múltiples, se sigue que $b \neq 0$ y que $a^2 - 4b \neq 0$. Para obtener una ecuación del Hessiano, pasemos a coordenadas homogéneas: $F(x,y,z) = y^2z - x^3 - ax^2z - bxz^2$. Entonces

$$\begin{aligned} H(x,y,z) &= \begin{vmatrix} -6x - 2az & 0 & -2ax - 2bz \\ 0 & 2z & 2y \\ -2ax - 2bz & 2y & -2bx \end{vmatrix} \\ &= 8[(y^2 + bxz)(3x + az) - (ax + bz)^2 z], \end{aligned}$$

es decir, (dividiendo entre 8 y tomando $z=1$) tenemos

$$h(x,y) = y^2(3x+a) + bx(3x+a) - (ax+b)^2.$$

Es fácil de encontrar los puntos de intersección de las curvas $f = 0$ y $h = 0$. Expresemos la ecuación $f = 0$ en la forma $y^2 = x^3 + ax^2 + bx$ y sustituyamos esta expresión en la ecuación $h = 0$. Como resultado tenemos que

$$(x^3 + ax^2 + bx)(3x + a) + bx(3x + a) - (ax + b)^2 = 0,$$

es decir,

$$q(x) = 3x^4 + 4ax^3 + 6bx^2 - b^2 = 0.$$

Probemos que el polinomio $q(x)$ no tiene raíces múltiples. Supongamos que x_0 es un cero múltiple, entonces $q(x_0) = q'(x_0) = 0$. Notemos que $x_0 \neq 0$ ya que $q(0) = -b^2 \neq 0$. Como $q'(x) = 12(x^3 + ax^2 + bx)$ se tiene que

$$q(x) - \frac{q'(x)}{12}\left(3x + a - \frac{b}{x}\right) = (4b - a^2)x^2.$$

Como $q(x_0) = q'(x_0) = 0$, se tiene que $(4b - a^2)x_0^2 = 0$, luego $4b - a^2 = 0$, que es una contradicción.

Hemos probado que el polinomio $q(x)$ tiene cuatro raíces distintas x_i. A cada raíz x_i le corresponden dos valores distintos de y porque $y^2 = x_i^3 + ax_i^2 + bx_i = \frac{q'(x_i)}{12} \neq 0$. Así que, las curvas $F = 0$ y $H = 0$ tienen 8 puntos de intersección en el dominio finito $z \neq 0$. Como $F(x, y, 0) = -x^3$ y $H(x, y, 0) = 24xy^2$, se sigue que en la recta al infinito $z = 0$ las curvas $F = 0$ y $H = 0$ tienen precisamente un punto en común, $(0, 1, 0)$, con esto hemos completado la demostración del teorema. □

Cualquier recta que pase por dos puntos de inflexión contiene otro punto de inflexión. De hecho, podemos suponer que las coordenadas de uno de los puntos de inflexión son $(0, 1, 0)$. Si (x_0, y_0) es un punto en común de la curva $y^2 = x^3 + ax^2 + bx$ y su Hessiana $y^2(3x + a) + bx(3x + a) = (ax + b)^2$, entonces $(x_0, -y_0)$ también es un punto en común. Los puntos $(0, 1, 0)$ y $(x_0, \pm y_0, 1)$ se encuentran en la recta $x = x_0 z$.

Esquemáticamente la configuración de los nueve puntos de inflexión y las doce rectas que los contienen está representada en la Figura 1.11 (a). Un esquema mas simétrico de esta configuración es mostrada en la Figura 1.11 (b).

Recordemos una noción importante de la geometría proyectiva. Cuatro puntos en $\mathbb{C}P^2$ son llamados ***puntos genéricos*** (o en posición general) si ninguna terna de estos se encuentra sobre una recta; cuatro rectas en $\mathbb{C}P^2$ son llamadas ***rectas genéricas*** (o en posición general) si ninguna terna de ellas pasa por un mismo punto.

Puntos con coordenadas homogéneas $(x_i, y_i, z_i) = e_i$, donde $i = 1, 2, 3, 4$, son genéricos si y sólo si los vectores e_1, e_2, e_3 son linealmente independientes y $e_4 = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3$, donde $\lambda_1\lambda_2\lambda_3 \neq 0$. No es difícil mostrar que *para cualesquiera* 4 *puntos genéricos en* $\mathbb{C}P^2$ *existe una transformación proyectiva (es decir, una transformación lineal de coordenadas homogéneas) que manda estos* 4 *puntos en otros* 4 *puntos genéricos dados.*

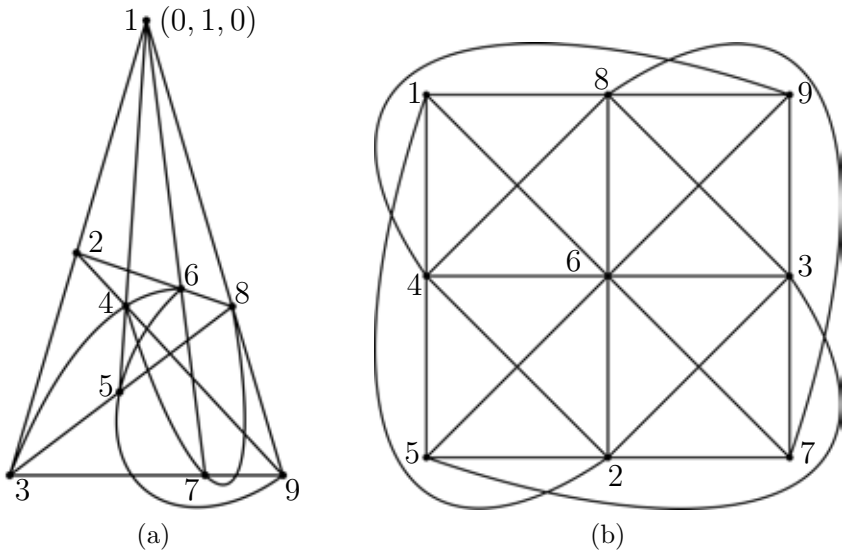


Figura 1.11

De hecho, sean $\{e_1\}_{i=1}^4$ y $\{\varepsilon_i\}_{i=1}^4$ dos conjuntos de cuatro puntos genéricos cada uno de ellos. Entonces $e_4 = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3$ y $\varepsilon_4 = \mu_1\varepsilon_1 + \mu_2\varepsilon_2 + \mu_3\varepsilon_3$, donde $\lambda_1\lambda_2\lambda_3 \neq 0$ y $\mu_1\mu_2\mu_3 \neq 0$. La transformación proyectiva requerida es la siguiente:

$$e_i \mapsto \alpha_i\varepsilon_i, \qquad \text{donde} \quad \alpha_i = \frac{\mu_i}{\lambda_i} \quad \text{para} \quad i = 1, 2, 3.$$

Existe una correspondencia uno-a-uno entre los conjuntos de puntos y los conjuntos de rectas en $\mathbb{C}P^2$: al punto (a, b, c) le corresponde la recta $ax + by + cz = 0$ (***dualidad proyectiva***). Si los puntos A y B se encuentran en la recta l, entonces las rectas a y b duales a A y B se cortan en el punto L dual a la recta l.

A la recta $ax + by + cz = 0$ le corresponde el punto $(a, b, c) = \alpha$. Por lo tanto, si $\mathbb{C}P^2$ esta sujeto a la transformación proyectiva $(x, y, z) \mapsto (x, y, z)A$, donde A es una matriz no singular, entonces la recta $(x, y, z)\alpha^T = 0$ se convierte en la recta $(x, y, z)A\alpha^T = 0$, es decir, $(x, y, z)(\alpha A^T)^T = 0$. Por lo tanto, la ley de transformación de las coordenadas en la recta es $\alpha \mapsto \alpha A^T$. Esta transformación también es proyectiva. Por esta razón, la dualidad proyectiva hace posible probar que para cualesquiera 4 rectas genéricas pueden ser convertidas en otras 4 rectas genéricas por una transformación proyectiva.

Teorema 1.4.2. *Bajo un cambio de coordenadas los 9 puntos de inflexión de una curva cúbica pueden ser transformados en el siguiente conjunto de 9 puntos:*

$$\begin{array}{ccc} (0, 1, -1) & (0, \varepsilon^2, -\varepsilon) & (0, \varepsilon, \varepsilon^2) \\ (-1, 0, 1) & (-\varepsilon^2, 0, 1) & (-\varepsilon, 0, 1) \\ (1, -1, 0) & (-\varepsilon, 1, 0) & (-\varepsilon^2, 1, 0), \end{array}$$

donde $\varepsilon^3 = 1$ *y* $\varepsilon \neq 1$, *es decir,* $\varepsilon^2 + \varepsilon + 1 = 0$.

Demostración. Las rectas 189, 463, 527 (Figura 1.11 (b)) no pueden concurrir a un punto. En efecto, si R es un punto en común de estas rectas, cualesquiera 4 puntos genéricos en $\mathbb{C}P^2$ pueden ser convertidos por una transformación proyectiva en cualesquiera otros 4 puntos genéricos, podemos suponer que los puntos R, 1, 5 y 6 son reales. Entonces también todos los otros puntos de la configuración son reales. Es fácil de ver que esto es imposible.

Tomando la recta 145 junto con la terna de rectas mencionadas tenemos cuatro rectas genéricas. Por lo tanto, podemos suponer que las rectas 145, 189, 463, y 572 están dadas por las ecuaciones $x+y+z=0$, $x=0$, $y=0$ y $z=0$, respectivamente. Entonces las coordenadas de los puntos de inflexión son de la siguiente forma:

$$\begin{array}{lll} (0,1,-1) & (0,a,-b) & (0,c,-d) \\ (-1,0,1) & (-a',0,1) & (-c',0,1) \\ (1,-1,0) & (-b',1,0) & (-d',1,0), \end{array} \tag{1.6}$$

donde todos los números $a, b, \ldots, d'$ son distintos de cero.

Los puntos $(0,a,-b)$, $(-a',0,1)$ y $(-b',1,0)$ se encuentran en una recta, a saber, la recta 862. Si $\alpha x + \beta y + \gamma z = 0$ fuese la ecuación de la recta 862 al sustituir las coordenadas de los puntos que se encuentran en ella, obtenemos las ecuaciones $\beta a - \gamma b = 0$, $-\alpha a' + \gamma = 0$ y $-\alpha b' + \beta = 0$, lo que nos obliga a que $ab' = ba'$, por lo que podemos suponer que $a = a'$ y $b = b'$. De forma análoga, $c = c'$ y $d = d'$.

Utilizando el proceso anterior y considerando las rectas 167 y 123 tenemos que $a = d$ y $b = c$, respectivamente. De igual forma, considerando las rectas 538 y 569, tenemos que $a = bc$ y $c = ad$, respectivamente. De ahí se sigue que $b^3 = 1$ y $a = b^2$. Es claro que $b \neq 1$. Sustituyendo $a = \varepsilon^2$, $b = \varepsilon$, $c = \varepsilon$ y $d = \varepsilon^2$ en (1.6) tenemos los puntos requeridos.

Es fácil verificar que las ocho rectas restantes contienen ternas de ciertos puntos los cuales son dados por ecuaciones de la forma $x + \alpha y + \beta z = 0$, donde α y β toman valores 1, ε, ε^2. □

Con la ayuda del Teorema 1.4.2 es fácil probar que ***cualquier cúbica no singular puede ser reducida a la forma*** $x^3 + y^3 + z^3 - 3\lambda xyz = 0$. En efecto, si los puntos de inflexión de la curva dada tienen las coordenadas indicadas en la formulación del Teorema 1.4.2, estos puntos pertenecen a la terna de rectas $xyz = 0$ y a la terna de rectas

$$(x + y + z)(x + \varepsilon y + \varepsilon^2 z)(x + \varepsilon^2 y + \varepsilon z) = 0.$$

Por lo tanto, cualquier cúbica que pase por estos nueve puntos esta dada por la ecuación

$$\mu xyz + \nu(x + y + z)(x + \varepsilon y + \varepsilon^2 z)(x + \varepsilon^2 y + \varepsilon z) = 0.$$

Pero como,

$$(x + y + z)(x + \varepsilon y + \varepsilon^2 z)(x + \varepsilon^2 y + \varepsilon z) = x^3 + y^3 + z^3 - 3xyz,$$

la cúbica toma la forma

$$(x^3 + y^3 + z^3) + 3\lambda xyz = 0,$$

donde $\lambda = \dfrac{\mu}{3\nu} - 1$.

1.5. Cúbicas singulares

La ecuación de una cúbica no singular puede ser escrita en la forma

$$y^2 = (x - x_1)(x - x_2)(x - x_3),$$

donde los números x_1, x_2 y x_3 son distintos. En el caso real una de tales curvas es mostrada en la Figura 1.12.

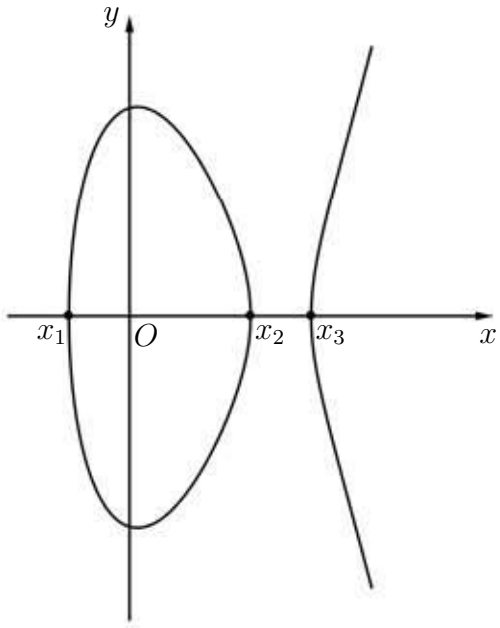


Figura 1.12

Sean $x_1 < x_2 < x_3$. Cuando las raíces x_1 y x_2 se confunden, tenemos una curva de la forma $y^2 = x^2(x - 1)$ (ver Figura 1.13 (a)); cuando las raíces x_2 y x_3 se confunden, tenemos una curva de la forma $y^2 = x^2(x + 1)$ (ver Figura 1.13 (b)). Sobre $\mathbb{R}$ estas curvas son distintas, pero sobre $\mathbb{C}$ la distinción desaparece. Si las tres raíces se confunden obtenemos la curva $y^2 = x^3$ (ver Figura 1.13 (c)). Para estas tres curvas el origen es un punto singular.

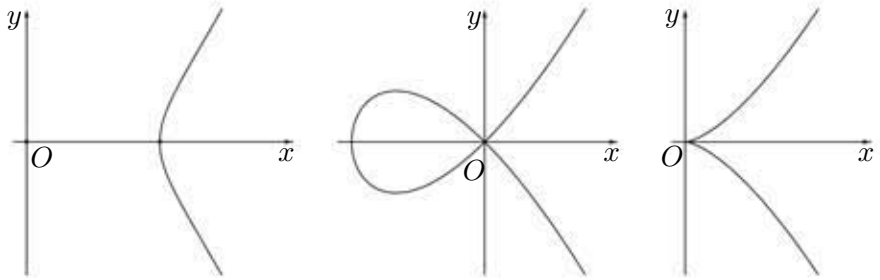


Figura 1.13

Cualquier recta $y = kx$ interseca la curva $y^2 = x^2(x \pm 1)$ y $y^2 = x^3$ en el punto singular con multiplicidad al menos dos. De hecho, para las ecuaciones $k^2x^2 = x^2(x \pm 1)$ y $k^2x^2 = x^3$ la raíz es al menos doble. Por lo tanto, cualquier recta que conecte el punto singular con otro punto de la cúbica, el tercer punto de intersección es nuevamente el punto singular. Por lo tanto, la suma del punto singular con cualquier otro punto siempre debe resultar el punto singular. Por esta razón, no podemos definir la adición para un punto singular. Pero si excluimos el punto singular, entonces para las curvas $y^2 = x^2(x + 1)$ y $y^2 = x^3$ la adición de puntos está bien definida. Si en ambos casos tomamos como elemento cero el punto al infinito, entonces la curva $y^2 = x^2(x + 1)$ sobre $\mathbb{R}$ se convierte en el grupo de los números reales distintos de cero con respecto a la multiplicación y la curva $y^2 = x^3$ se convierte en el grupo de números reales con respecto a la adición (sobre $\mathbb{C}$ obtenemos es grupo $\mathbb{C} \setminus \{0\}$ con respecto a la multiplicación y $\mathbb{C}$ con respecto a la adición, respectivamente).

Comencemos con la curva $y^2 = x^3$. Esta curva admite una parametrización racional $x = t^{-2}$, $y = t^{-3}$. Los puntos de intersección de la curva con la recta $ax + by + c = 0$ están determinados por la relación $ct^3 + at + b = 0$. Si la recta no pasa por el punto singular, entonces $c \neq 0$. En este caso tenemos una ecuación cúbica con coeficiente cero en t^2. La suma de las raíces de tal ecuación es igual a cero: $t_1 + t_2 + t_3 = 0$. Tomemos como elemento cero E el punto al infinito correspondiente al parámetro $t_E = 0$. Si t_A y t_B son los valores de los parámetros correspondientes a los puntos A y B de la curva dada. La recta AB interseca a la cúbica en el punto X; tenemos $t_A + t_B + t_X = 0$. La recta EX interseca a la curva en el punto $A + B$, es decir, $t_E + t_X + t_{A+B} = 0$. Por lo tanto, $t_{A+B} = -t_X = t_A + t_B$. De ahí que para sumar puntos de la curva $y^2 = x^3$, debemos sumar los correspondientes valores del parámetro t. Observemos que al punto singular le corresponde el valor del parámetro $t = \infty$.

La curva $y^2 = x^2(x + 1)$ también admite una parametrización racional. De hecho, sea $y = tx$. Entonces $t^2x^2 = x^2(x+1)$, es decir, $x = t^2-1$ y $y = tx = t^3-t$. La recta $ax+by+c = 0$ interseca la curva $y^2 = x^2(x + 1)$ en el punto cuyo valor del parámetro satisface la relación

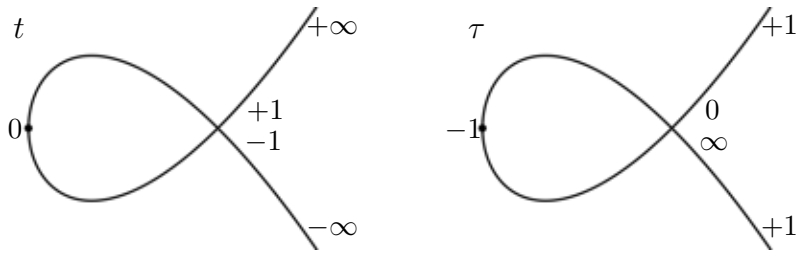
$$a(t^2 - 1) + b(t^3 - t) + c = 0.$$

Si $b \neq 0$, entonces después de la división por b tenemos una ecuación cúbica con coeficiente 1 en t^3 y -1 en t. Las raíces de tal ecuación satisfacen la relación $t_1t_2 + t_2t_3 + t_3t_1 = -1$.

Podemos obtener una relación más simple después de la reparametrización

$$t = (1 + \tau)(1 - \tau)^{-1}.$$

De hecho, es fácil verificar que $\tau_1\tau_2\tau_3 = 1$. Tomemos como elemento cero E el punto al infinito correspondiente al valor del parámetro $\tau_E = 1$. Para encontrar la suma $A + B$, debemos considerar el punto X en el cual la recta AB interseca a la cúbica. Como $\tau_A\tau_B\tau_X = 1$ y $\tau_X\tau_E\tau_{A+B} = 1$, se sigue que $\tau_{A+B} = \tau_A\tau_B$. Cuando sumamos puntos de la curva $y^2 = x^2(x+1)$, multiplicamos los valores correspondientes del parámetro τ. Al punto singular le corresponde no uno sino dos valores de cada parámetro t y τ, a saber, $t = \pm 1$ y $\tau = 0, \infty$ (ver Figura 1.14).



(a) (b)

Figura 1.14

1.6. Una cúbica no singular no admite parametrización racional

Las cúbicas singulares que estudiamos en la sección anterior admiten una parametrización racional. Ahora probaremos que ninguna de las cúbicas no singulares admiten una parametrización racional. Recordemos que una cúbica no singular puede reducirse a la forma $y^2 = x(x-1)(x-\lambda)$, donde $\lambda \neq 0, 1$.

Teorema 1.6.1. *Si $\lambda \neq 0, 1$, entonces no existen polinomios P_1, P_2, Q_1, Q_2 tales que las funciones no constantes $y(t) = P_1(t)/P_2(t)$ y $x(t) = Q_1(t)/Q_2(t)$ satisfagan la relación $y^2 = x(x-1)(x-\lambda)$.*

Demostración. Supongamos que $P_1(t)/P_2(t)$ y $Q_1(t)/Q_2(t)$ no son constantes y que satisfacen,

$$\frac{P_1^2}{P_2^2} = \frac{Q_1}{Q_2} \cdot \frac{Q_1 - Q_2}{Q_2} \cdot \frac{Q_1 - \lambda Q_2}{Q_2}.$$

Podemos suponer que los polinomios P_1 y P_2 son primos relativos y también Q_1 y Q_2. Ya que

$$P_1^2 Q_2^3 = P_2^2 Q_1 (Q_1 - Q_2)(Q_1 - \lambda Q_2),$$

se sigue que el polinomio P_2^2, el cual es primo relativo con P_1^2, es divisible entre Q_2^3 y que el polinomio Q_2^3, el cual es primo relativo con $Q_1, Q_1 - Q_2$ y $Q_1 - \lambda Q_2$, es divisible entre P_2^2. Por esta razón, los polinomios Q_2^3 y P_2^2 son proporcionales entre si. Por lo tanto, reemplazando P_1 con un polinomio proporcional podemos obtener la igualdad

$$P_1^2 = Q_1(Q_1 - Q_2)(Q_1 - \lambda Q_2). \tag{1.7}$$

Además, el polinomio Q_2^3 es el cuadrado de un polinomio; de ahí que, Q_2 también es el cuadrado de un polinomio.

Los polinomios $Q_1, Q_1 - Q_2$ y $Q_1 - \lambda Q_2$ son primos relativos por pares y, por lo tanto, la igualdad (1.7) implica que cada uno de ellos es un cuadrado perfecto. Por esta razón, en la familia de polinomios de la forma $\alpha Q_1 + \beta Q_2$, donde $\alpha, \beta \in \mathbb{C}$, existen 4 cuadrados perfectos,

a saber, $Q_1, Q_2, Q_1 - Q_2$, y $Q_1 - \lambda Q_2$; estos polinomios no son proporcionales y son distintos porque $\lambda \neq 0, 1$.

Para obtener una contradicción mostremos que en la recta proyectiva $\alpha Q_1 + \beta Q_2$, donde Q_1 y Q_2 son primos relativos, no hay más de tres puntos que pueden ser cuadrados perfectos. De hecho, supongamos que en esta recta proyectiva hay 4 cuadrados perfectos:

$$R_1^2, \quad R_2^2, \quad \alpha_1 R_1^2 - \beta_1 R_2^2 \quad \text{y} \quad \alpha_2 R_1^2 - \beta_2 R_2^2.$$

Como los polinomios R_1 y R_2 son primos relativos, se sigue que los polinomios $\sqrt{\alpha_i} R_1 \pm \sqrt{\beta_i} R_2$ deben ser cuadrados perfectos. Como resultado, de la recta proyectiva $\alpha Q_1 + \beta Q_2$ en la cual hay 4 cuadrados perfectos llegamos a la recta proyectiva $\alpha R_1 + \beta R_2$ en la cual también hay 4 cuadrados perfectos. De esta recta proyectiva, podemos llegar a otra recta proyectiva, etc. Pero cada paso decrece el grado máximo de cada polinomio de la forma $\alpha Q_1 + \beta Q_2$ al menos por un factor de 2. Contradicción. □

Capítulo 2

Funciones Elípticas

La adición de puntos en la circunferencia está relacionada con su parametrización dada por las funciones ***seno*** y ***coseno***. De hecho, considerando la función $f : \mathbb{R} \to S^1$ definida por la fórmula $f(t) = (\cos(t), \text{sen}(t))$, resulta que esta función parametriza la circunferencia con números reales de tal forma que la adición de puntos en la circunferencia se corresponde con la adición de números reales.

Una parametrización similar existe para las cúbicas. Ésta se obtiene por medio de funciones elípticas. Bajo esta parametrización la adición de puntos en una cúbica definida en el capítulo 1 se corresponde con la adición de los valores del parámetro.

En este capítulo estudiaremos las principales propiedades de las funciones elípticas y mostraremos como podemos parametrizar una cúbica no singular con su ayuda.

El nombre ***funciones elípticas*** desde luego está ligado con la ***elipse***, pero la relación es bastante indirecta. La relación viene dada por las integrales elípticas que si están directamente relacionadas con la elipse. La longitud de un arco de la elipse es expresada por una integral elíptica de una forma particular. Aquí es precisamente donde se origina el nombre de ***integrales elípticas***. Las funciones elípticas aparecen en el proceso de inversión de las integrales elípticas de una manera especial, no exactamente relacionada con el cálculo de la longitud de arco de una elipse.

Las integrales elípticas aparecen a principios del siglo diecisiete en el cálculo de longitudes de arcos de ciertas curvas, principalmente ***elipses***. Aparte de la elipse un ejemplo interesante es ***la lemniscata de Bernoulli*** cuya longitud de arco se encuentra con una integral de la forma $\int_0^\alpha \frac{dx}{\sqrt{1-x^4}}$. Fue por esta integral que el matemático italiano Conde ***Giulio Fagnano*** (1682-1766) obtuvo una fórmula de duplicación, que después Euler extendió en la primera mitad del siglo dieciocho a un teorema de adición

$$\int_0^\alpha \frac{dx}{\sqrt{1-x^4}} + \int_0^\beta \frac{dx}{\sqrt{1-x^4}} = \int_0^\gamma \frac{dx}{\sqrt{1-x^4}},$$

donde

$$\gamma = \frac{\alpha\sqrt{1-\beta^4} + \beta\sqrt{1-\alpha^4}}{1+\alpha^2\beta^2}.$$

En realidad, Euler consiguió obtener un teorema de adición para integrales de una forma más general. A saber, el probó que

$$\int_0^\alpha \frac{dx}{\sqrt{P(x)}} + \int_0^\beta \frac{dx}{\sqrt{P(x)}} = \int_0^\gamma \frac{dx}{\sqrt{P(x)}},$$

donde

$$\gamma = \frac{\alpha\sqrt{P(\beta)} + \beta\sqrt{P(\alpha)}}{1+n\alpha^2\beta^2} \text{ y } P(x) = 1 + mx^2 + nx^4.$$

Euler incluso obtuvo teoremas de adición para formas más generales.

Después de Euler, ***Legendre*** trabajó incansablemente por muchos años en el desarrollo de la teoría de las integrales elípticas. El resumió los resultados de sus estudios en el libro ***Exercises de calcul integrál*** (Ejercicios de cálculo integral), publicados entre 1811 y 1819. Una edición revisada del libro fue publicada entre 1827-1832 bajo el nombre Traité des fonctions elliptiques et des intégrales eulériennes (Tratado sobre funciones elípticas e integrales eulerianas).

Legendre llamó ***funciones elípticas*** a lo que hoy en día llamamos ***integrales elípticas***. Después de los trabajos de ***Abel*** y ***Jacobi*** la importancia del libro de Legendre disminuyó. Sin embargo, Abel y Jacobi se refirieron al libro de Legendre con gran respeto, como se lo merecía.

La teoría de las funciones elípticas comenzó propiamente con el trabajo de Abel ***Recherches sur les fonctions elliptiques*** (Estudios sobre funciones elípticas). Abel mostró que la inversión de una integral elíptica de primer orden,

$$\alpha = \int \frac{dx}{(1-cx^2)(1+ex^2)},$$

da origen a una función $\varphi(\alpha)$ que tiene dos periodos en el dominio complejo. Abel estudió meticulosamente las ecuaciones que relacionan $\varphi(\alpha)$ con $\varphi(n\alpha)$. Jacobi comenzó a estudiar la teoría de las funciones elípticas casi simultáneamente con Abel. Esto llevó a una tensa, aunque corta, competencia entre ellos. Sin un puesto laboral permanente, casi en la pobreza, Abel finalizó la segunda parte de los ***Recherches***... y continuó sus intensivos estudios. Pero pronto Abel enfermo gravemente y murió en 1829 a los 27 años de edad.

Mucho antes de Abel y Jacobi, ***Gauss*** conocía varios de los descubrimientos de estos personajes. Pero Gauss no publicó sus resultados, se sabe de ellos por sus memorias, su diario y las multiples biografías que existen de él.

Es conveniente considerar a las funciones elípticas como funciones de una variable compleja. Muchas de sus propiedades están desarrolladas sólo en el plano complejo $\mathbb{C}$, y

no en la recta real $\mathbb{R}$. La parametrización de una curva cúbica es mucho más gráfica sobre $\mathbb{C}$. Por lo tanto, comenzaremos con la investigación de la topología de una cúbica no singular en $\mathbb{C}P^2$. Resultará, que desde el punto de vista topológico todas estas curvas son iguales: todas ellas son toros bidimensionales.

2.1. La estructura topológica de las cúbicas no singulares en $\mathbb{C}P^2$

La ecuación de cualquier cúbica no singular en $\mathbb{C}P^2$ puede reducirse a la forma

$$y^2z = (x - a_1z)(x - a_2z)(x - a_3z), \tag{2.1}$$

donde los números a_i son distintos por pares. Esta ecuación determina una curva compleja en $\mathbb{C}P^2$ con dimensión compleja igual a 1 y con dimensión real igual a 2.

Para encontrar la estructura topológica de la curva (2.1) en $\mathbb{C}P^2$, consideremos la proyección

$$p : \mathbb{C}P^2 \setminus \{(0, 1, 0)\} \to \mathbb{C}P^1, \quad (x, y, z) \mapsto (x, z).$$

La recta proyectiva compleja $\mathbb{C}P^1$ (es una compactificación de $\mathbb{C}$ a un punto y por tanto) es homeomorfa a la esfera bidimensional S^2. Para $b \neq 0$ la ecuación $y^2 = b$ tiene exactamente dos soluciones distintas. Por lo tanto, si $z \neq 0$ y $x - a_iz \neq 0$, entonces un punto $(x, z) \in \mathbb{C}P^1$ tiene exactamente dos preimágenes que pertenecen a la curva (2.1). Si $z \neq 0$ pero x/z es igual a uno de los números a_i, entonces solo existe una preimagen. Para $z = 0$ la ecuación (2.1) se convierte en la ecuación $x^3 = 0$. Por lo tanto, el punto $\infty = (0, 1)$ también tiene solamente una preimagen, a saber, $(0, 1, 0)$. Además, la preimagen del punto $(1, z)$ tiende al $(0, 1, 0)$ cuando $z \to 0$.

La proyección bajo p de la curva (2.1) sobre $\mathbb{C}P^1$ funciona de la manera siguiente. Si excluimos de $\mathbb{C}P^1$ los puntos a_1, a_2, a_3, e ∞, entonces todos los puntos tienen exactamente dos preimágenes. La estructura de la función en vecindades de los puntos a_i e ∞ debe ser estudiada con mayor detalle. Para simplificar, supongamos que $a_1 = 0$. Consideremos coordenadas afines, es decir, sea $z = 1$. La proyección de la curva (2.1) en $\mathbb{C}P^1$ con este sistema de coordenadas resulta ser $(x, y) \mapsto x$. Entonces (2.1) toma la forma

$$y^2 = x(x - a_2)(x - a_3),$$

donde $a_2a_3 \neq 0$. Para puntos x cerca del cero la cantidad $(x - a_2)(x - a_3)$ casi es constante, es decir, casi tenemos la ecuación $y^2 = cx$. Esta ecuación tiene soluciones de la forma $x = c\lambda^2e^{2i\varphi}$, $y = c\lambda e^{i\varphi}$. Cuando φ varía de 0 a π, recorremos una vuelta completa alrededor del punto $(0, 1)$ en $\mathbb{C}P^1$. Bajo dicha trayectoria, y cambia de signo. Al nacer esta trayectoria alrededor del $(0, 1)$ en $\mathbb{C}P^1$ de la curva (2.1) no regresamos al punto inicial (ver Figura 2.1). Pero si recorremos una vuelta mas si regresaremos al punto inicial, ya que cambiando el signo de y_0 dos veces obtenemos y_0.

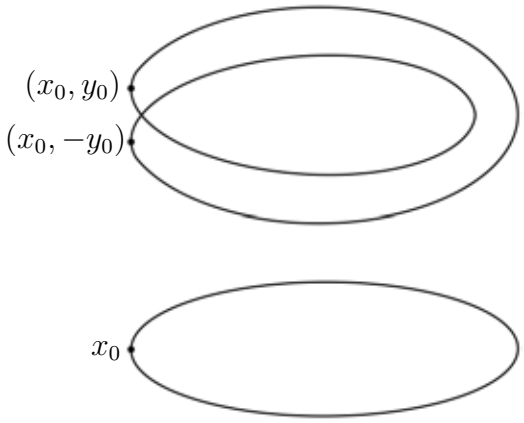


Figura 2.1

La estructura de la proyección de la curva (2.1) en $\mathbb{C}P^1$ en una vecindad de ∞ es la misma que de las vecindades de a_i. De hecho, sea $x = 1$. Entonces en una vecindad de $z = 0$ la ecuación (2.1) aproximadamente se ve como $y^2 = \frac{1}{z}$ y, por esto, el signo de y cambia cuando damos una vuelta completa alrededor del punto $z = 0$.

Cortemos $\mathbb{C}P^1$ de a_1 a a_2 y de a_3 a ∞. Los levantamientos de estos cortes a la curva (2.1) divide esta en dos partes. De hecho, avanzando por cualquier trayectoria cerrada en $\mathbb{C}P^1$ que no interseque los cortes solo podremos encerrar los puntos a_1, a_2, a_3, e ∞ en pares y bajo la trayectoria alrededor de dos puntos el valor de y no cambia. Por lo tanto, es imposible pasar de una preimagen de un punto de $\mathbb{C}P^1$ a su otra preimagen sin intersecar los cortes.

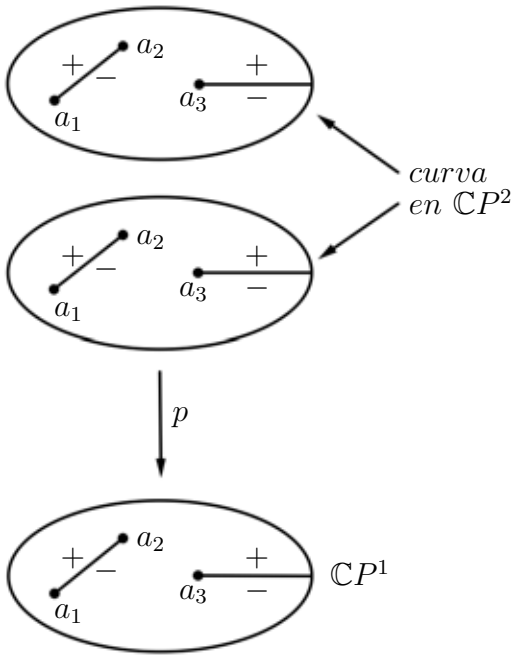


Figura 2.2

Si cortamos $\mathbb{C}P^1$ de a_1 a a_2 y de a_3 a ∞, entonces la parte restante de $\mathbb{C}P^1$ puede ser representada en la forma de un plano con cortes. La parte de la curva (2.1) que se encuentra por encima de este plano consiste en dos piezas. Realizando un corte en $\mathbb{C}P^1$ vamos del límite marcado con un signo más en una pieza de la curva (2.1) al límite marcado con un signo menos de la otra pieza. De ahí que, cuando los límites son pegados, obtenemos un toro (ver Figura 2.3).

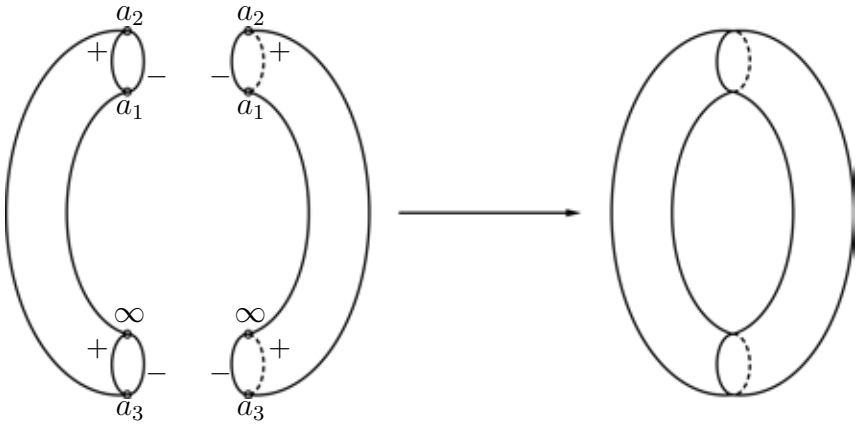


Figura 2.3

Una parametrización de la cúbica en $\mathbb{C}P^2$ puede ser determinada por medio de una función $f : \mathbb{C}^1 \to \mathbb{C}P^2$, donde $f(z) = (F_1(z), F_2(z), 1)$. La imagen de tal función debe ser un toro. La función más simple de $\mathbb{C}^1$ a un toro se obtiene identificando todos los puntos de la forma $z + n\omega_1 + m\omega_2$. En otras palabras, se hace que ω_1 y ω_2 sean periodos de las funciones F_1 y F_2.

2.2. Las funciones elípticas

Una función f se llama ***doblemente periódica*** si existen $\omega_1, \omega_2 \in \mathbb{C}$ distintos del cero tales que $\frac{\omega_1}{\omega_2}$ no pertenece a $\mathbb{R}$ y se cumple que $f(z + n\omega_1 + m\omega_2) = f(z)$ $\forall\, z \in \mathbb{C}$ y $\forall\, n, m \in \mathbb{Z}$. Supondremos que $Im(\frac{\omega_1}{\omega_2}) > 0$. Esto quiere decir que la rotación que se hace en el plano complejo para mandar ω_1 en ω_2 es en el sentido del avance de las manecillas del reloj (Figura 2.4). Al conjunto $\Omega = \{n\omega_1 + m\omega_2\,;\, n, m \in \mathbb{Z}\}$ se le llama ***la latiz generada por*** ω_1 ***y*** ω_2, y resulta que cada punto ω de Ω es un periodo de f.

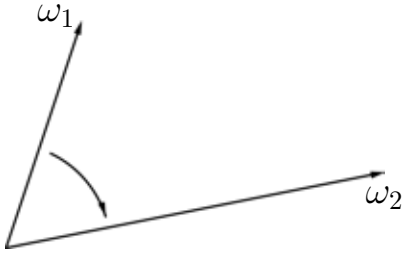


Figura 2.4

En lo que sigue sólo nos interesarán las funciones meromorfas doblemente periódicas. Recordemos que una función es llamada ***meromorfa*** en un dominio de $\mathbb{C}$ si esta función es analítica en el dominio salvo en un conjunto numerable de puntos aislados, y de manera que tales puntos singulares sean polos. En una vecindad de cualquier punto finito a del dominio de la función meromorfa f, esta puede ser desarrollada en una serie

$$f(z) = c_0(z-a)^r + c_1(z-a)^{r+1} + \cdots ,$$

donde $c_0 \neq 0$ y $r \in \mathbb{Z}$ (desde luego si a es un punto donde f es analítica se tendrá que $r \geq 0$ y si a es un polo, $r < 0$). Una función meromorfa doblemente periódica es llamada ***función elíptica***.

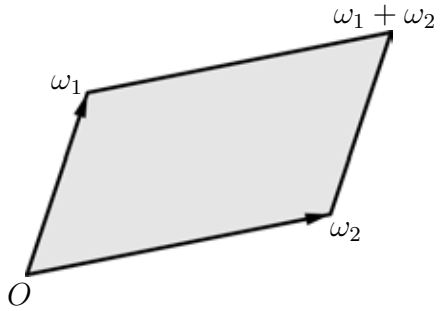


Figura 2.5

Cualquier número complejo z puede ser representado de la forma $z = a_1\omega_1 + a_2\omega_2$, donde $a_i \in \mathbb{R}$. Como el número a_i puede ser representado como la suma de su parte entera y su parte fraccionaria, se tiene que, una función elíptica esta completamente determinada por sus valores en el ***paralelogramo fundamental*** (Figura 2.5).

$$\{\alpha_1\omega_1 + \alpha_2\omega_2 \,:\, 0 \leq \alpha_1, \alpha_2 \leq 1\}.$$

La imagen del paralelogramo fundamental bajo cualquier traslación paralela ($z \mapsto z + b$ con $b \in \mathbb{C}$ fijo) también puede ser considerado como un paralelogramo fundamental.

Teorema 2.2.1 (**de Liouville**). *Una función elíptica sin polos es constante.*

Demostración. Sea f una función elíptica sin polos entonces f es una función analítica en $\mathbb{C}$ y como el paralelogramo fundamental P es compacto entonces para alguna $M \in \mathbb{R}^+$, se tiene que $|f(z)| \leq M \quad \forall\, z \in P$ y por lo tanto en todo $\mathbb{C}$. Entonces por el teorema de Liouville para funciones enteras f es constante. □

Como todo punto singular de una función meromorfa es aislado, el paralelogramo fundamental contiene sólo un número finito de puntos singulares. Por lo tanto, existe una traslación paralela del paralelogramo fundamental de manera que no existen puntos singulares en la frontera del paralelogramo fundamental.

Proposición 2.2.2. *Sea f una función elíptica y P el paralelogramo fundamental con vértices $\alpha, \alpha+\omega_1, \alpha+\omega_1+\omega_2$ y $\alpha+\omega_2$, y sea ∂P su frontera (Figura 2.6), entonces $\int_{\partial P} f(z)dz = 0$.*

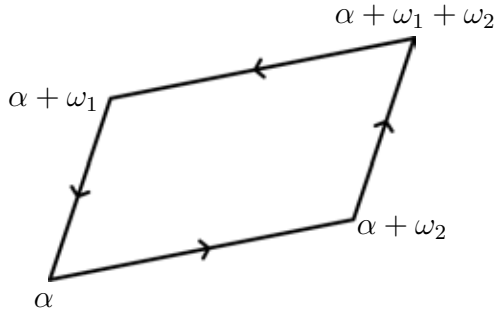


Figura 2.6

Demostración.

$$\int_{\partial P} f(z)dz = \int_{\alpha}^{\alpha+\omega_2} f(z)dz + \int_{\alpha+\omega_2}^{\alpha+\omega_1+\omega_2} f(z)dz + \int_{\alpha+\omega_1+\omega_2}^{\alpha+\omega_1} f(z)dz + \int_{\alpha+\omega_1}^{\alpha} f(z)dz$$

Pero por ser $f(z)$ elíptica tenemos que

$$\int_{\alpha}^{\alpha+\omega_2} f(z)dz = \int_{\alpha+\omega_1}^{\alpha+\omega_1+\omega_2} f(z+\omega_1)d(z+\omega_1) = -\int_{\alpha+\omega_1+\omega_2}^{\alpha+\omega_1} f(z)dz$$

Y análogamente

$$\int_{\alpha+\omega_2}^{\alpha+\omega_1+\omega_2} f(z)dz = -\int_{\alpha+\omega_1}^{\alpha} f(z)dz$$

Por lo tanto

$$\int_{\partial P} f(z)dz = 0.$$

□

Esta proposición nos permitirá obtener información esencial en los ceros y polos de las funciones elípticas, veamos como

Teorema 2.2.3. *a) La suma de los residuos de una función elíptica en los puntos singulares dentro del paralelogramo fundamental es igual a cero.*
b) Para una función elíptica, sean a_i los ceros y polos que se encuentran dentro de un paralelogramo fundamental y sean r_i sus ordenes (positivos para los ceros y negativos para los polos). Entonces $\sum r_i = 0$ y $\sum r_i a_i \equiv 0 \pmod{\Omega}$, es decir, $\sum r_i a_i = m\omega_1 + n\omega_2$, donde m y n son enteros.

Demostración. a) Sea $f(z)$ una función elíptica y sea P un paralelogramo fundamental sin puntos singulares de esta función en la frontera. Entonces por el Teorema del residuo sabemos

que como f es analítica en ∂P y ∂P es homóloga a cero dentro de un rectángulo abierto que contiene a $\overline{P}$ que:

$$\sum_{z\in P} \text{res}(f,z) = \frac{1}{2\pi i}\int_{\partial P} f(z)dz.$$

Pero por la proposición anterior sabemos que $\int_{\partial P} f(z) = 0$. Por lo tanto

$$\sum_{z\in P} \text{res}(f,z) = 0.$$

b) Sea $g(x) = \frac{f'(z)}{f(z)}$, como f es una función elíptica entonces $g(z)$ también es una función elíptica. Si a es una singularidad de g entonces a es un cero o un polo de $f(z)$, si el orden es r entonces $f(z) = (z-a)^r f_1(z)$ con $f_1(z)$ analítica alrededor de a y $f_1(a) \neq 0$, luego

$$g(z) = \frac{f'(z)}{f(z)} = \frac{r}{(z-a)} + \frac{f_1'(z)}{f_1(z)},$$

por lo que el residuo de $g(z)$ en el punto a es igual a r. Por la parte a) aplicada a $g(z)$ se tiene que

$$\sum r_i = 0.$$

Ahora consideremos a $h(z) = \frac{zf'(z)}{f(z)}$ (que es una función meromorfa, pero no necesariamente elíptica), si a es un cero o polo de $f(z)$ de orden r entonces a es un polo de $h(z)$ y como antes

$$h(z) = \frac{zf'(z)}{f(z)} = \frac{rz}{z-a} + \frac{zf_1'(z)}{f_1(z)}$$

por lo que el residuo de $h(z)$ en el punto a es ar. Y entonces

$$\int_{\partial P} \frac{zf'(z)}{f(z)}dz = 2\pi i \sum a_i r_i.$$

Ahora calcularemos la integral anterior de otra manera. Primero veamos la siguiente igualdad

$$\begin{aligned}
\int_\alpha^{\alpha+\omega_2} \frac{zf'(z)}{f(z)}dz &= \int_\alpha^{\alpha+\omega_2} \frac{(z-\omega_1)f'(z)}{f(z)}dz + \omega_1\int_\alpha^{\alpha+\omega_2}\frac{f'(z)}{f(z)}dz \\
&= \int_{\alpha+\omega_1}^{\alpha+\omega_1+\omega 2} \frac{zf'(z+\omega_1)}{f(z+\omega_1)}d(z+\omega_1) + \omega_1\left[\log f(z)|_\alpha^{\alpha+\omega_2}\right] \\
&= -\int_{\alpha+\omega_1+\omega_2}^{\alpha+\omega_1} \frac{zf'(z)}{f(z)}dz + \omega_1\left[\log f(z)|_\alpha^{\alpha+\omega_2}\right].
\end{aligned}$$

Como $f(\alpha+\omega_2) = f(\alpha)$, el logaritmo de $f(z)$ solo puede variar por $2k\pi i$ para alguna $k \in \mathbb{Z}$ cuando z varia de α a $\alpha+\omega_2$. Como resultado tenemos que

$$\int_\alpha^{\alpha+\omega_2} h(z)dz + \int_{\alpha+\omega_1+\omega_2}^{\alpha+\omega_1} h(z)dz = 2\pi i(m\omega_1).$$

Análogamente tenemos que

$$\int_{\alpha+\omega_2}^{\alpha+\omega_1+\omega_2} h(z)dz + \int_{\alpha+\omega_1}^{\alpha} h(z)dz = 2\pi i(n\omega_2).$$

Por lo tanto

$$\int_{\partial P} h(z)dz = 2\pi i[n\omega_1 + m\omega_2],$$

es decir, $\sum a_i r_i = m\omega_1 + n\omega_2$ para algunos $n, m \in \mathbb{Z}$. □

Como hemos mencionado, una función elíptica que no sea constante debe tener al menos un polo dentro del paralelogramo fundamental. Pero como la suma de los residuos en los puntos singulares que se encuentran dentro del paralelogramo fundamental es igual a cero, la función no puede tener exactamente un polo de orden 1. Para una función elíptica el número de polos (contando la multiplicidad) dentro del paralelogramo fundamental es llamado ***el orden de la función elíptica***. El mínimo orden posible es 2; y dos formas de que el orden sea 2 son:
1) un polo de orden 2. (Como sucede en la función ℘ de Weierstrass)
2) dos polos simples. (Como sucede en las funciones elípticas de Jacobi)

Por el Teorema 2.2.3 *b*), para una función elíptica la suma de los ordenes de los ceros dentro del paralelogramo fundamental es igual a la suma de los ordenes de los polos, es decir, es igual al orden de la función. También es claro, que los polos de la función $f(z) - c$ son los mismos que de la función $f(z)$. En consecuencia, una función elíptica de orden r toma cualquier valor finito dentro del paralelogramo fundamental exactamente $r-veces$ (contando la multiplicidad).

2.3. La función ℘ de Weierstrass

Ahora mostraremos que para cualquier latiz Ω la función

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega}{}' \left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right], \tag{2.2}$$

es una función elíptica (donde $\sum'_\omega = \sum_{\omega\in\Omega\setminus\{0\}}$, quiere decir, hacer la suma sobre todos los puntos ω distintos de cero de la latiz Ω). La agrupación de los términos en corchetes es esencial, ya que las series $\sum'_\omega (z-\omega)^{-2}$ y $\sum'_\omega \omega^{-2}$ divergen.

Primero, demostraremos que la serie (2.2) define una función meromorfa. En cualquier compacto K que no contenga los puntos de la latiz, esta serie converge uniforme y absolutamente. Ya que

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} = \frac{2z\omega - z^2}{\omega^2(z-\omega)^2} = \frac{1}{\omega^3}\cdot\frac{2z - z^2\omega^{-1}}{(z\omega^{-1}-1)^2}$$

y como

$$\lim_{\omega\to\infty}\frac{2z-z^2\omega^{-1}}{(z\omega^{-1}-1)^2}=2z$$

tenemos que $\forall\,\omega\in\Omega\backslash 0$ con $|\omega|$ suficientemente grande y $\forall\,z\in K$ existe una constante C tal que

$$\left|\frac{1}{(z-\omega)^2}-\frac{1}{\omega^2}\right|<\frac{C}{|\omega|^3}.$$

Ahora demostraremos que $\sum'|\omega|^{-3}$ converge. Si definimos

$$\Omega_r=\{m\omega_1+n\omega_2\,:\,m,n\in\mathbb{Z}\ y\ \text{máx}(|m|,|n|)=r\}$$

tenemos que para $r>0$ $\ |\Omega_r|=8r$. Ahora sea $d=\text{mín}(|m|,|n|)$ entonces

$$\sum{}'|\omega|^{-3}=\sum_{r=1}^{\infty}\sum_{\omega\in\Omega_r}|\omega|^{-3}\leq\sum_{r=1}^{\infty}8r|rd|^{-3}$$

$$=\frac{8}{d^3}\sum_{r=1}^{\infty}\frac{1}{r^2}$$

por lo tanto $\sum'|\omega|^{-3}$ converge. Así que $\wp(z)$ es una función meromorfa con polos en los puntos de la latiz Ω. Esta función es llamada ***la función $\wp$ de Weierstrass***. Ahora probaremos la periodicidad de $\wp(z)$. Para esto vamos a considerar su derivada

$$\wp'(z)=-2\sum_{\omega\in\Omega}(z-\omega)^{-3}.$$

Claramente ω_1 y ω_2 son periodos de la función $\wp'(z)$. De ahí que $\wp'(z+\omega_i)-\wp'(z)=0$ por lo que, las funciones $\wp(z+\omega_i)$ y $\wp(z)$ pueden diferir solamente por una constante c. Sustituyendo $z=-\frac{\omega_i}{2}$ en la igualdad $\wp(z+\omega_i)=\wp(z)+c$ tenemos que $\wp(\frac{\omega_i}{2})=\wp(-\frac{\omega_i}{2})+c$, pero por la fórmula (2.2) sabemos que la función $\wp(z)$ es par. Por lo que la constante es $c=0$, es decir, ω_1 y ω_2 son periodos de $\wp(z)$.

La función $\wp$ tiene polos dobles en los puntos de la latiz; y ésta no tiene otros puntos singulares. Dentro del paralelogramo fundamental sólo se encuentra un punto de la latiz. Por lo tanto, dentro del paralelogramo fundamental la suma de los polos de $\wp$ es congruente a cero módulo Ω. Por el Teorema 2.2.3, dentro del paralelogramo fundamental se encuentran dos ceros de $\wp$, llamémoslos u y v, tales que $u+v\equiv 0\ (\text{mód}\ \Omega)$. Para cualquier constante c los polos de la función $\wp(z)-c$ coinciden con los polos de la función $\wp(z)$ y, por lo tanto, dentro del paralelogramo fundamental se encuentran exactamente dos puntos, u y v, para los cuales $\wp(u)+\wp(v)=c$ y $u+v\equiv 0\ (\text{mód}\ \Omega)$. Si $u\equiv -u\ (\text{mód}\ \Omega)$, entonces estos dos puntos coinciden, es decir, el valor correspondiente de $\wp$ es alcanzado dos veces. En los puntos donde los ceros de $\wp(z)-c$ se confunden, la derivada $\wp'(z)$ se anula. Es posible seleccionar el paralelogramo fundamental de tal forma que éste contenga exactamente cuatro puntos para los cuales $u\equiv -u\ (\text{mód}\ \Omega)$, más precisamente, los puntos

$$0,\frac{\omega_1}{2},\frac{\omega_2}{2},\frac{\omega_1+\omega_2}{2}$$

(Figura 2.7). El primero de estos puntos es el polo de $\wp$ y los otros tres puntos son ceros de $\wp'$.

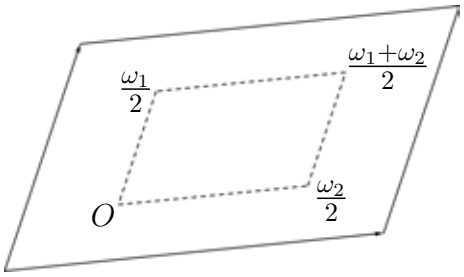


Figura 2.7

Así, los valores

$$e_1 = \wp\left(\frac{\omega_1}{2}\right),\ e_2 = \wp\left(\frac{\omega_2}{2}\right) \text{ y } e_3 = \wp\left(\frac{\omega_1+\omega_2}{2}\right)$$

de $\wp$ son de multiplicidad 2 y no hay otros valores de multiplicidad 2. Los valores de multiplicidad 2 corresponden a los ceros de la derivada, por eso, $\wp'(z) = 0$ si y sólo si

$$z \equiv \frac{1}{2}\omega_1,\ \frac{1}{2}\omega_2,\ \frac{1}{2}(\omega_1+\omega_2) \qquad (\text{mód } \Omega).$$

Observemos que los números e_1, e_2 y e_3 son distintos. Sea $f_j(z) = \wp(z) - e_j$ para $j = 1, 2, 3$. Como los polos de f_j son los mismos que los de $\wp$, entonces f_j es una función elíptica de orden 2 y por lo tanto tiene dos ceros (contando multiplicidades). Ya que

$$f_j(\tfrac{1}{2}\omega_j) = f'(\tfrac{1}{2}\omega_j) = 0,$$

f_j tiene ceros dobles en $\frac{1}{2}\omega_j$ y por esta razón no tiene otros ceros. En particular, $f_j(\frac{1}{2}\omega_k) \neq 0$ para $j \neq k$. Como

$$\begin{aligned} f_j(\tfrac{1}{2}\omega_k) &= \wp(\tfrac{1}{2}\omega_k) - e_j \\ &= e_k - e_j, \end{aligned}$$

de ahí se sigue que $e_j \neq e_k$ si $j \neq k$.

La función de Weierstrass no solamente nos da un ejemplo de una función elíptica, sino que veremos, que nos permite describir la estructura de todas las funciones elípticas con respecto a la misma latiz Ω.

Teorema 2.3.1. *Sea $f(z)$ una función elíptica arbitraria y $\wp(z)$ la función de Weierstrass con los mismos periodos. Entonces existen funciones racionales R y R_1 tal que*

$$f = R(\wp) + R_1(\wp)\wp'$$

Demostración. Sabemos que es posible representar a $f(z)$ como la suma de la función par $g(z) = \frac{1}{2}(f(z) + f(-z))$ y la función impar $h(z) = \frac{1}{2}(f(z) - f(-z))$. Puesto que la función $\wp'(z)$ es una función impar, $h_1(z) = \frac{h(z)}{\wp'(z)}$ es una función par, entonces las funciones pares g y h_1 son pares y tenemos que

$$f(z) = g(z) + h_1(z)\wp'(z)$$

Por lo tanto, es suficiente probar que cualquier función elíptica par puede ser representada como una función racional de $\wp$.

Primero resaltamos dos propiedades de los ceros y polos de una función elíptica.

1°. Si f es una función par, y u es uno de sus ceros (o polos) de orden m, entonces $-u$ también es un cero (o polo) de orden m. De hecho, en el caso de los ceros es suficiente observar que para f una función par tenemos:

$$f^{(k)}(-z) = (-1)^k f^{(k)}(z) \quad \forall\, k \geq 0.$$

En el caso de los polos tenemos que considerar $\frac{1}{f}$ en lugar de f.

2°. Si f es una función elíptica par y $u \equiv -u \pmod{\Omega}$; entonces el orden de un cero o un polo de f en u es par. Probaremos esto solamente para los ceros (puesto que para los polos podemos considerar $\frac{1}{f}$ en vez de f). La condición $u \equiv -u \pmod{\Omega}$ es equivalente al hecho de que

$$u \equiv 0, \frac{\omega_1}{2}, \frac{\omega_2}{2}, \frac{\omega_1 + \omega_2}{2} \qquad (\text{mód } \Omega).$$

Además, la periodicidad de f' implica que $f'(u) = f'(-u)$. Pero la derivada de una función par es impar; por esta razón, $f'(u) = 0$. Por lo tanto, si la función f tiene un cero en u, entonces la multiplicidad de este cero es al menos dos. Para cualquiera de los casos

$$u \equiv \frac{\omega_1}{2}, \frac{\omega_2}{2}, \frac{\omega_1 + \omega_2}{2} \qquad (\text{mód } \Omega)$$

la función $F(z) = \wp(z) - \wp(u)$ tiene un cero de orden dos en u, pero si $u \equiv 0 \pmod{\Omega}$ entonces la función $F(z) = \frac{1}{\wp(z)}$ tiene tal propiedad. Usando F, podemos construir una función elíptica par $f_1(z) = \frac{f(z)}{F(z)}$ para la cual el orden del cero en u es el orden de u en f menos dos. De ahí que, si $f_1(u) \neq 0$, entonces el orden del cero en u es igual a 2 y si $f_1 = 0$, entonces podemos aplicar los mismos argumentos a f_1 en vez de a f, etcetera.

Por la proposiciones anteriores de los ceros y polos de la función elíptica par f estos pueden ser divididos en pares de la forma $\{x, -x\}$. Seleccionando un representante de cada par, de tal manera que $a_1, ..., a_k$ serán los representantes de los ceros y $b_1, ..., b_k$ los representantes de los polos. Consideremos la función elíptica

$$Q(z) = R(\wp(z)) = \frac{\prod(\wp(z) - \wp(a_i))}{\prod(\wp(z) - \wp(b_i))}$$

donde solo tomaremos los a_i y b_i distintos de los puntos de la latiz (puesto que en tales puntos latiz la función $\wp$ toma valores infinitos). Si no tomamos en cuenta los puntos de la latiz, entonces el sistema completo de ceros y polos de Q es el mismo que el de f, puesto que $\wp(z) = \wp(a)$ si y sólo si $z \equiv \pm a$ (mód Ω). Pero por el Teorema 2.2.2 b) para una función elíptica la suma de los ordenes de sus ceros y polos dentro del paralelogramo fundamental es igual a cero; por esta razón, el orden de un cero ó un polo en un punto de la latiz es únicamente determinado por los ordenes de los otros ceros y polos. Por lo tanto $\frac{f(z)}{Q(z)}$ es una función elíptica sin polos, es decir, una constante (digamos c). Como resultado tenemos que $f(z) = cR(\wp(z))$ como se deseaba. □

2.4. Una ecuación diferencial para la función $\wp$ de Weierstrass

En la sección anterior probamos que una función elíptica puede ser expresada racionalmente en términos de $\wp(z)$ y $\wp'(z)$ y la expresión fue descrita explícitamente. Esto puede ser aplicado a la función par $(\wp'(z))^2$ ésta tiene ceros de multiplicidad 2 en $\frac{\omega_1}{2}, \frac{\omega_2}{2}, \frac{\omega_1+\omega_2}{2}$ y un polo de multiplicidad 6 en un punto de la latiz. Por esta razón,

$$(\wp'(z))^2 = c(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3), \tag{2.3}$$

donde $e_1 = \wp(\frac{\omega_1}{2}), e_2 = \wp(\frac{\omega_2}{2}), e_3 = \wp(\frac{\omega_1+\omega_2}{2})$ y puesto que $\wp(z) = z^2 + \cdots$ y $\wp'(z) = -2z^{-3} + \cdots$, se deduce que $c = 4$.

Existe otra manera de obtener esta ecuación diferencial para $\wp(z)$. La siguiente manera no solamente nos ofrece un nuevo método de deducir la ecuación, si no también nos provee otra forma de expresarla. Para deducirla, usaremos el hecho de que si los coeficientes de las potencias negativas en el desarrollo de Laurent de las funciones $(\wp'(z))^2$ y $a\wp^3(z) + b\wp^2(z) + c\wp(z) + d$ coinciden, entonces estas funciones son iguales. De hecho bastará ver que su diferencia es una función elíptica sin polos que se anula en el origen. Por esta razón, su diferencia es constante e igual a 0.

Como

$$\left(\frac{1}{1-z}\right) = 1 + z^2 + z^3 + \cdots,$$

entonces

$$\left(\frac{1}{1-z}\right)^2 = \frac{d}{dz}\left(\frac{1}{1-z}\right) = 1 + 2z + 3z^2 + \cdots,$$

de ahí se sigue que

$$\begin{aligned}
\wp(z) &= \tfrac{1}{z^2} + \sum{}' \left(\tfrac{1}{(z-\omega)^2} - \tfrac{1}{\omega^2}\right) \\
&= \tfrac{1}{z^2} + \sum{}' \left(\tfrac{1}{\omega^2}\left(1 + 2\tfrac{z}{\omega} + 3\left(\tfrac{z}{\omega}\right)^2 + \cdots\right) - \tfrac{1}{\omega^2}\right) \\
&= \tfrac{1}{z^2} + \sum{}' \left(2\tfrac{z}{\omega^3} + 3\tfrac{z^2}{\omega^4} + 4\tfrac{z^3}{\omega^5} + 5\tfrac{z^4}{\omega^6} + \cdots\right)
\end{aligned}$$

Pero si $\omega \in \Omega$ también $-\omega \in \Omega$, y entonces tenemos que para k impar

$$\sum{}' \omega^{-k} = 0.$$

Por lo tanto

$$\wp(z) = z^{-2} + 3G_4 z^2 + 5G_6 z^4 + \cdots$$

donde $G_k = \sum{}' \omega^{-k}$. De ahí que,

$$\begin{aligned} \wp^2(z) &= z^{-4} + 6G_4 + 10G_6 z^2 + \cdots, \\ \wp^3(z) &= z^{-6} + 9G_4 z^{-2} + 15G_6 + \cdots, \\ (\wp'(z))^2 &= 4z^{-6} - 24G_4 z^{-2} - 80G_6 + \cdots. \end{aligned}$$

De esta manera,

$$a\wp^3(z) + b\wp^2(z) + c\wp(z) + d = az^{-6} + bz^{-4} + (9aG_4 + c)z^{-2} + (15aG_6 + 6bG_4 + d) + \cdots.$$

Por lo tanto, $a\wp^3 + b\wp^2 + c\wp + d = (\wp')^2$ si

$$a = 4\,,\ b = 0\,,\ 9aG_4 + c = -24G_4 \ \text{ y } \ 15aG_6 + 6bG_4 + 6bG_4 + d = -80G_6,$$

es decir,

$$a = 4\,,\ b = 0\,,\ c = -60G_4 \ \text{ y } \ d = -140G_6.$$

Si definimos $g_2 = 60G_4$ y $g_3 = 140G_6$. Entonces tenemos la ecuación diferencial,

$$(\wp'(z))^2 = 4\wp^3(z) - g_2\wp(z) - g_3. \tag{2.4}$$

Comparando (2.3) y (2.4) observamos que

$$\begin{aligned} (\wp'(z))^2 &= 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3) \\ &= 4\wp^3(z) - 4(e_1 + e_2 + e_3)\wp^2(z) + 4(e_1e_2 + e_2e_3 + e_3e_1)\wp(z) - 4e_1e_2e_3 \end{aligned}$$

que son validas las relaciones:

$$e_1 + e_2 + e_3 = 0\,,\ e_1e_2 + e_2e_3 + e_3e_1 = -\frac{g_2}{4} \ \text{ y } \ e_1e_2e_3 = \frac{g_3}{4}.$$

De aquí, podemos verificar la siguiente igualdad

$$g_2^3 - 27g_3^2 = 16(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2.$$

Observemos primero que de las igualdades anteriores podemos deducir que

$$e_1^2 + e_2^2 + e_3^2 = (e_1 + e_2 + e_3)^2 - 2(e_1e_2 + e_2e_3 + e_3e_1) = \frac{g_2}{2}$$

y que

$$e_1^2e_2^2 + e_2^2e_3^2 + e_3^2e_1^2 = (e_1e_2 + e_2e_3 + e_3e_1)^2 - 2e_1e_2e_3(e_1 + e_2 + e_3) = \frac{g_2^2}{16}.$$

Ahora diferenciando $(\wp'(z))^2 = 4(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3)$ y $(\wp'(z))^2 = 4\wp^3(z) - g_2\wp(z) - g_3$ con respecto a $\wp$ y evaluando en $z = \frac{\omega_1}{2}$, tenemos

$$\begin{aligned} 4(e_1-e_2)(e_1-e_3) &= \left[(\wp'(\tfrac{\omega_1}{2}))^2\right]' \\ &= 12e_1^2 - g_2, \end{aligned}$$

con expresiones similares para $z = \frac{\omega_2}{2}$ y $z = \frac{\omega_1+\omega_2}{2}$. De ahí que si definimos $\Delta_\wp = 16(e_1-e_2)^2(e_2-e_3)^2(e_3-e_1)^2$ tenemos que

$$\begin{aligned} \Delta_\wp &= -\tfrac{1}{4}\prod_{i=1}^{3}(12e_i^2 - g_2) \\ &= -\tfrac{1}{4}(1728(e_1e_2e_3)^2 - 144g_2(e_1^2e_2^2 + e_2^2e_3^2 + e_3^2e_1^2) + 12g_2^2(e_1^2+e_2^2+e_3^2) - g_2^3) \\ &= -\tfrac{1}{4}(108g_3^2 - 9g_2^3 + 6g_2^3 - g_2^3) \\ &= g_2^3 - 27g_3^2. \end{aligned}$$

En la sección anterior se mostró que los números e_1, e_2 y e_3 son distintos. Por lo tanto, $g_2^3 - 27g_3^2 \neq 0$.

De aquí surge una pregunta natural: ¿Dados números g_2 y g_3 tales que $g_2^3 \neq 27g_3^2$, existirá una latiz para la cual $g_2 = 60\sum' \omega^{-4}$ y $g_3 = 140\sum' \omega^{-6}$?

La respuesta a esta pregunta es afirmativa, ver el capítulo 6 de [5].

2.5. Una parametrización de la cúbica via la función $\wp$ de Weierstrass

La ecuación diferencial para $\wp$ nos permite aclarar la naturaleza de la adición de puntos en una cúbica. Para esto tendremos que ocupar el hecho que dejamos sin demostración la sección anterior:

Para cualquier par de números g_2 y g_3 tal que $g_2^3 \neq 27g_3^2$ existe una latiz para la cual la función de Weierstrass asociada a la latiz satisface la ecuación

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3.$$

La curva cúbica $y^2 = 4x_3 - g_2x - g_3$ puede ser parametrizada con la ayuda de $\wp$ definiendo $x = \wp(z)$ y $y = \wp'(z)$. Pasando a las coordenadas homogéneas en $\mathbb{C}P^2$ la función $f : \mathbb{C}/\Omega \to \mathbb{C}P^2$ puede ser definida como

$$f(x) \mapsto \begin{cases} (\wp(z), \wp'(z), 1) & para \quad z \neq 0, \\ (z^3\wp(z), z^3\wp'(z), z^3) = (0,1,0) & para \quad z = 0. \end{cases}$$

Obviamente, esta función es analítica en todos los puntos distintos de los puntos de la latiz. Expresando esta función de la forma

$$z \mapsto \left(\frac{\wp(z)}{\wp'(z)}, 1, \frac{1}{\wp'(z)}\right)$$

podemos verificar que también es analítica en una vecindad de un punto de la latiz. La función f es una función inyectiva del toro $\mathbb{C}/\Omega$ en la cúbica $y^2z = 4x^3 - g_2xz^2 - g_3z^3$ en $\mathbb{C}P^2$.

De hecho, en la recta al infinito $z = 0$ solamente se encuentra el punto $(0, 1, 0)$ de la curva; los puntos de la latiz que corresponden a un solo punto en el toro son enviados por la función f precisamente a este punto. Para todos los otros puntos podemos considerar la curva afín $y^2 = 4x^3 - g_2x - g_3$ y la función $z \mapsto (\wp(z), \wp'(z))$.

La ecuación $\wp(z) = c$ puede tener una o dos soluciones (si es una, es de multiplicidad 2). Esta ecuación tiene dos soluciones si $\wp'(z) \neq 0$, además de las dos soluciones una es la negativa de la otra. Las imágenes de estos dos puntos bajo la función bajo la función $z \mapsto (\wp(z), \wp'(z))$ no coinciden, puesto que los números $\wp'(z)$ y $\wp'(-z) = -\wp'(z)$ difieren por un signo.

La suma de números complejos induce una adición de puntos en el toro la cual, en consecuencia, con la ayuda de la función f induce una adición de los puntos en una cúbica. Resulta que esta suma es precisamente la adición de puntos en una cúbica definida en la Sección 1.1 si tomamos como elemento neutro el punto infinito $(0, 1, 0)$, veamos como.

Sean P_1 y P_2 los puntos en la cúbica correspondientes a los puntos z_1 y z_2 en $\mathbb{C}$, es decir, $P_i = (\wp(z_i), \wp'(z_i))$. Dibujemos la recta $y = ax + b$ que pasa por P_1 y P_2. Entonces $\wp'(z_i) = a\wp(z_i) + b$, donde $i = 1, 2$.

En el punto $z = 0$ la función elíptica $\wp'(z) - a\wp(z) - b$ tiene un polo de multiplicidad 3 y no tiene ningún otro polo en el paralelogramo fundamental. Por lo tanto, el orden de esta función es igual a 3, es decir, está tiene precisamente 3 ceros, a saber, los que ya conocemos z_1 y z_2 y un tercer cero z_3. Puesto que la suma de los polos y ceros es igual a cero, tenemos que $z_1 + z_2 + z_3 \equiv 0 \pmod{\Omega}$, luego, $z_3 \equiv -z_1 - z_2 \pmod{\Omega}$. Así, el tercer punto de intersección de la recta P_1P_2 con la cúbica es el punto

$$\begin{aligned} P_3' &= (\wp(z_3), \wp'(z_3)) = (\wp(-z_1 - z_2), \wp'(-z_1 - z_2)) \\ &= (\wp(z_1 + z_2), -\wp'(z_1 + z_2)). \end{aligned}$$

Por lo tanto, el punto $P_3 = (\wp(z_1 + z_2), \wp'(z_1 + z_2))$ correspondiente a la suma de z_1 y z_2 es simétrico a P_3' con respecto al eje x. En otras palabras, P_3 es el punto de intersección de la cúbica con la recta $P_3'E$, donde $E = (0, 1, 0)$ es el punto infinito en la cúbica. Esto es precisamente lo que queríamos establecer.

Extendiendo un poco más los argumentos anteriores, podemos mostrar que existe un teorema algebraico de adición para $\wp$, esto es, $\wp(z_1+z_2)$ puede ser expresado algebraicamente en términos de $\wp(z_1)$ y $\wp(z_2)$. De hecho la recta $y = ax + b$ que pasa por los puntos P_1 y P_2 interseca la cúbica $y^2 = 4x^3 - g_2x - g_3$ en los tres puntos (x_i, y_i), donde $x_1 = \wp(z_1)$, $x_2 = \wp(z_2)$ y $x_3 = \wp(z_1 + z_2)$. Por lo tanto la ecuación cúbica

$$(ax + b)^2 = 4x^3 - g_2x - g_3$$

tiene las raíces indicadas x_1 , x_2 y x_3. Expresando el coeficiente de x^2 en términos de estas

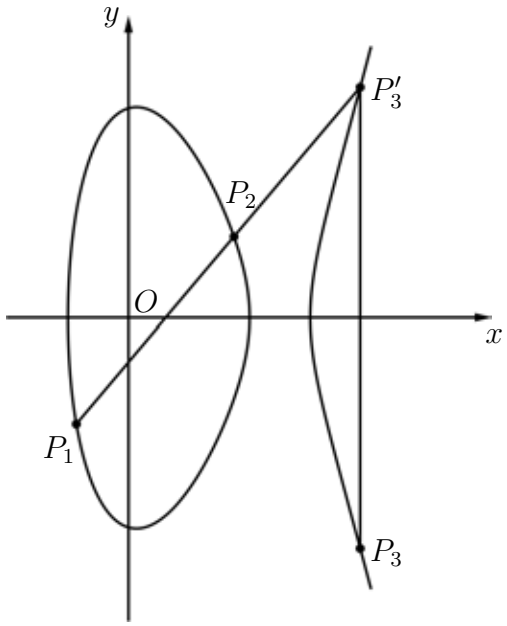


Figura 2.8

raíces tenemos que

$$\wp(z_1) + \wp(z_2) + \wp(z_1 + z_2) = \frac{a^2}{4}.$$

Como $\wp'(z_1) = a\wp(z_1) + b$ y $\wp'(z_2) = a\wp(z_2) + b$, se sigue que $a = \left(\frac{\wp'(z_1)-\wp'(z_2)}{\wp(z_1)-\wp(z_2)}\right)$. Por lo tanto

$$\wp(z_1 + z_2) = -\wp(z_1) - \wp(z_2) + \frac{1}{4}\left(\frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)}\right)^2. \tag{2.5}$$

De esta manera, $\wp(z_1 + z_2)$ puede ser expresado racionalmente en términos de $\wp(z_i)$ y $\wp'(z_i)$, con $i = 1, 2$. Es preciso recordar que $\wp'(z_i)$ puede ser expresado algebraicamente en términos de $\wp(z_i)$, como vimos anteriormente

$$\wp'(z_i) = \sqrt{4\wp^3(z_i) - g_2\wp(z_i) - g_3}.$$

Notemos también que si en la fórmula de adición hacemos $z_2 \to z_1 = z$ se obtiene la fórmula de duplicación

$$\wp(2z) = -2\wp(z) + \frac{1}{4}\left(\frac{\wp''(z)}{\wp'(z)}\right)^2.$$

Con la ayuda de la función de Weierstrass podemos también parametrizar la curva

$$y^2 = G_4(x),$$

donde $G_4(x) = a_0x^4 + a_1x^3 + a_2^2 + a_3x + a_4$ es un polinomio de grado cuatro sin raíces múltiples. Para este fin hagamos los cambios de variable $x = x_1^{-1} + \alpha$ y $y = y_1x_1^{-2}$. Obteniendo

$$y_1^2x_1^{-4} = b_4x_1^{-4} + b_3x_1^{-3} + b_2x_1^{-2} + b_1x_1^{-1} + G_4(\alpha).$$

Si α es una raíz de G_4, entonces $y_1^2 = b_1x_1^3 + b_2x_1^2 + b_3x_1 + b_4$. Esta cúbica pude ser parametrizada con la ayuda de la función de Weierstrass. Observe que los cambios de variable

$$x = x_1^{-1} + \alpha \quad \text{y} \quad y = y_1x_1^{-n}$$

nos permiten de una forma análoga pasar de la curva $y^2 = G_{2n}(x)$ a la curva $y^2 = G_{2n-1}(x)$.

2.6. Las integrales elípticas

La función de Weierstrass $\wp(z)$ satisface, como hemos visto, la ecuación diferencial

$$\left(\frac{d\wp}{dz}\right)^2 = 4\wp^3 - g_2\wp - g_3.$$

Por lo tanto,

$$dz = \frac{d\wp}{\sqrt{4\wp^3 - g_2\wp - g_3}},$$

es decir,

$$z = \int \frac{du}{\sqrt{4u^3 - g_2u - g_3}},$$

donde $u = \wp(z)$. Así que, $z = \wp^{-1}(u)$, luego, la inversión de la integral

$$\int \frac{du}{\sqrt{4u^3 - g_2u - g_3}}$$

da lugar a la función de Weierstrass.

Una ***integral elíptica*** es una integral de la forma

$$\int R(x, \sqrt{G(x)})dx,$$

donde $G(x)$ es un polinomio de grado 3 ó 4 sin raíces múltiples y $R(x, y)$ es una función racional de dos variables. En un principio, estas integrales aparecieron en los cálculos de la longitud de arco de varias curvas, por ejemplo, elipses. Solamente más tarde se notó que para ciertas integrales elípticas las funciones inversas poseían propiedades más interesantes, principalmente la doble periodicidad.

Las integrales elípticas se pueden reducir a ciertas integrales simples. Iniciaremos considerarendo integrales más simples que las integrales elípticas. Primero, probaremos que si $R(x)$ es una función racional, entonces $\int R(x)dx$ es la suma de una función racional y un cierto número de sumandos de la forma $c_i \log(x - a_i)$. Es suficiente probar que una función racional $R(x)$ puede ser representada en la forma

$$A(x) + \sum_{i,k} \frac{c_{i,k}}{(x - a_i)^k},$$

donde $A(x)$ es un polinomio. Sea $R(x) = P(x)/Q(x)$, donde P y Q son polinomios. Dividiendo P entre Q tenemos que $P(x) = A(x)Q(x) + S(x)$, con $A(x)$ y $S(x)$ polinomios, donde $S(x)$ es el residuo y grado $S <$ grado Q. Sea $Q = Q_1Q_2$, donde Q_1 y Q_2 son polinomios primos relativos. Entonces, existen polinomios a y b tal que $a(x)Q_1(x) + b(x)Q_2(x) = 1$. Por lo tanto,

$$\frac{S}{Q_1Q_2} = \frac{aSQ_1 + bSQ_2}{Q_1Q_2} = \frac{aS}{Q_2} + \frac{bS}{Q_1}$$

En las fracciones obtenidas podemos dividir el numerador y el denominador con un residuo. Repitiendo este procedimiento llegamos a la suma de un polinomio $A(x)$ y varias fracciones de la forma $p(x)(x-a)^{-n}$, donde grado $p(x) < n$. La demostración se completa expresando el polinomio $p(x)$ de la forma

$$p(x) = b_1(x-a)^{n-1} + b_2(x-a)^{n-2} + \cdots + b_n.$$

Leibniz fue el primero en estudiar la integración de funciones racionales. El considero la factorización de polinomios en factores con coeficientes reales y, por lo tanto, el se cuestiono: *¿Sera cierto ó no que cualquier polinomio real puede ser factorizado en factores de grado 1 y 2 con coeficientes reales?*. En 1702 Leibniz publicó un articulo en el cual el decía que era imposible factorizar el polinomio $x^4 + a^4$ en la forma requerida puesto que

$$\begin{aligned} x^4 + a^4 &= (x^2 + a^2 i)(x^2 - a^2 i) \\ &= (x + a\sqrt{i})(x - a\sqrt{i})(x + a\sqrt{-i})(x - a\sqrt{-i}) \end{aligned}$$

y el producto de cualesquiera 2 de estos factores no puede ser, como el creyó, una cuadrática con coeficientes reales. Solo 17 años después fue que Nicolás Bernoulli (1687- 1759) indicara que

$$x^4 + a^4 = (x^2 + a^2)^2 - 2a^2x^2 = (x^2 + \sqrt{2}ax + a^2)(x^2 - \sqrt{2}ax + a^2).$$

En su correspondencia Leibniz y Jacob Bernoulli también discutieron integrales de expresiones irracionales que aparecían en varios problemas físicos y matemáticos. Muchas de estas integrales son elípticas.

Pasemos ahora de las funciones racionales a las irracionales mas simples. Para calcular la integral

$$\int R(x, \sqrt{G(x)})dx,$$

donde $G(x) = ax+b$ es una función lineal, primero haremos el cambio de variables $u = ax+b$. Como resultado tenemos una integral de la forma

$$\int R_1(u, \sqrt{u})du,$$

donde R_1 es nuevamente una función racional. Ahora, sea $t = \sqrt{u}$. Entonces $du = d(t^2) = 2tdt$ y, por lo tanto,

$$\int R_1(u, \sqrt{u})du = \int R_1(t^2, t)2tdt = \int R_2(t)dt,$$

donde R_2 es una función racional, por lo que el estudio de estas segundas integrales se reduce las primeras.

Ahora sea $G(x) = ax^2 + bx + c$. Como hemos dicho en la sección anterior, con la ayuda de los cambios de variable $x = x_1^{-1} + \alpha$ y $y = y_1 x_1^{-1}$ podemos pasar de la curva $y^2 = G(x)$ a la curva $y_1^2 = G_1(x_1)$ donde G_1 es una función lineal. Apliquemos estos cambios de variable para calcular la integral $\int R(x, y)dx$, donde $y^2 = G(x)$. Sea $x = x_1^{-1} + \alpha$ y $y = y_1 x_1^{-1}$, donde $G(\alpha) = 0$. Entonces $dx = -x_1^{-2}dx_1$ y

$$\int R(x, y)dx = -\int R(x_1^{-1} + \alpha, y_1 x_1^{-1})x_1^{-2}dx_1 = \int R_1(x_1, y_1)dx_1,$$

donde $y_1^2 = Ax_1 + B$.

De esta manera, las integrales de la forma $\int R(x, y)dx$, donde R es una función racional y $y = \sqrt{G(x)}$, pueden ser expresadas en términos de funciones elementales si el grado $G < 2$.

En el caso donde grado $G = 3$ pueden aparecer funciones las cuales son inversas de funciones elípticas. La integral $\int R(x, y)dx$, donde $y = \sqrt{G_4(x)}$, puede ser reducida a la integral $\int Q(x, y)dx$, donde $y = \sqrt{4x^3 - g_2x - g_3}$. En efecto, usando los cambios de variable $x = x_1^{-1} + \alpha$ y $y = y_1 x_1^{-2}$ podemos pasar del polinomio de grado cuatro G_4 a un polinomio de grado tres y de cualquier polinomio de grado tres podemos pasar con la ayuda de un cambio lineal a un polinomio de la forma $4x^3 - g_2x - g_3$.

Habríamos podido confinarnos al cálculo de integrales de la forma $\int R(x, y)dx$, donde $y^2 = 4x^3 - g_2x - g_3$, pero en muchos otros casos son convenientes algunas otras formas de integrales elípticas. Por lo tanto, primero calcularemos las integrales elípticas en su forma general y después estudiaremos algunas formas especiales.

Sea $I = \int R(x, y)dx$, donde R es una función racional, y

$$y^2 = a_0x^4 + 4a_1x^3 + 6a_2x^2 + 4a_3x + a_4,$$

donde al menos uno de los coeficientes a_0 y a_1 es distinto de cero.

Teorema 2.6.1 (**Legendre**). *La integral elíptica I puede ser representada como una combinación lineal de una función racional en x y y; la integral de una función racional en x; y de las integrales*

$$\int \frac{dx}{y}, \int \frac{xdx}{y}, \int \frac{x^2dx}{y} \quad y \quad \int \frac{dx}{(x - c)y}.$$

Demostración. Como y^2 puede ser expresado en forma de polinomio en términos de x, podemos suponer que la función racional R no contiene y^k para $k \geq 2$. Por otra parte,

$$\frac{a + by}{c + dy} = \frac{(a + by)(c - dy)y}{(c + dy)(c - dy)y} = \frac{A}{y} + B,$$

donde A y B son funciones racionales de x. Por lo tanto, el calculo de la integral $\int R(x, y)ds$ de reduce al cálculo de las integrales $\int B(x)dx$ y $\int \frac{A(x)dx}{y}$. La función racional $A(x)$ puede

ser representada de la forma

$$A(x) = \sum a_n x^n + \sum \frac{a_{r,m}}{(x-c_r)^m}.$$

De ahí que, solamente es preciso considerar las integrales de la forma

$$J_n = \int \frac{x^n dx}{y} \ (n \geq 0) \ \text{ y } \ H_m = \int \frac{dx}{(x-c)^m y} \ (m \geq 1).$$

Puesto que

$$\begin{aligned} \tfrac{d}{dx}(x^m y) = & \ m x^{m-1} y + x^m \tfrac{dy}{dx} = \tfrac{1}{y}\left[m x^{m-1} y^2 + \tfrac{1}{2} x^m \tfrac{d(y^2)}{dx}\right] \\ = & \ (m+2)a_0 \tfrac{x^{m+3}}{y} + 2(2m+3)a_1 \tfrac{x^{m+2}}{y} + 6(m+1)a_2 \tfrac{x^{m+1}}{y} + 2(2m+1)a_3 \tfrac{x^m}{y} + m a_4 \tfrac{x^{m-1}}{y}, \end{aligned}$$

de ahí se sigue que

$$x^m y = \ (m+2)a_0 J_{m+3} + 2(2m+3)a_1 J_{m+2} + 6(m+1)a_2 J_{m+1} + 2(2m+1)a_3 J_m + m a_4 J_{m-1}.$$

Usando estas identidades para $m = 0, 1, 2, \ldots$, consecutivamente podemos expresar J_3 en términos de J_0, J_1, J_2 (y una función racional de x y y), después J_4 en términos de J_0, J_1, J_2, J_3, etcetera. (En el caso cuando $a_0 = 0$ podemos expresar J_2 en términos de J_0 y J_1; después J_3 en términos de J_0 y J_1, etcetera.)

Para calcular la integral $H_m = \int \frac{dx}{(x-c)^m y}$, escribamos el polinomio $G(x)$ en la forma

$$G(x) = b_0(x-c)^4 + 4b_1(x-c)^3 + 6b_2(x-c)^2 + 4b_3(x-c) + b_4,$$

donde $b_0 = a_0$. Como en el caso anterior, tenemos la identidad

$$\begin{aligned} \frac{d}{dx}\left[(x-c)^m y\right] = (m+2)b_0 \frac{(x-c)^{m+3}}{y} + 2(2m+3)b_1 \frac{(x-c)^{m+2}}{y} \\ + 6(m+1)b_2 \frac{(x-c)^{m+1}}{y} + 2(2m+1)b_3 \frac{(x-c)^m}{y} + m b_4 \frac{(x-c)^{m-1}}{y}. \end{aligned}$$

Integrando estas identidades para $m = -1, -2, -3, \ldots$ tenemos que

$$\begin{array}{rcccccc} \dfrac{y}{x-c} & = & b_0 \displaystyle\int \frac{(x-c)^2}{y} dx & +2b_1 \displaystyle\int \frac{x-c}{y} dx & & -2b_3 H_1 & -b_4 H_2, \\ \dfrac{y}{(x-c)^2} & = & & -2b_1 J_0 & -6b_2 H_1 & -6b_3 H_2 & -2b_4 H_3, \\ \dfrac{y}{(x-c)^3} & = & -b_0 J_0 & -6b_1 J_1 & -12b_2 H_2 & -10b_3 H_3 & -3b_4 H_4, \\ \cdots & & \cdots & \cdots & \cdots & \cdots & \cdots \end{array}$$

Estas identidades nos permiten expresar $H_2, H_3, H_4, \ldots$ en términos de J_0, J_1, J_2, H_1 y funciones racionales de x y y. □

Como hemos mencionado, cualquier integral elíptica puede reducirse a la integral $\int R(x,y)dx$, donde

$$y^2 = 4x^3 - g_2x - g_3.$$

Esta forma de las integrales elípticas es llamada la ***forma de Weierstrass***. Puesto que en este caso $a_0 = 0$, se sigue que J_2 puede ser expresado en términos de J_0 y J_1; por lo tanto, existen tres tipos de integrales con las cuales el resto pueden ser calculadas:

$$\int \frac{dx}{\sqrt{4x^3 - g_2x - g_3}}, \int \frac{xdx}{\sqrt{4x^3 - g_2x - g_3}} \text{ y } \int \frac{dx}{(x-c)\sqrt{4x^3 - g_2x - g_3}}.$$

Otro forma muy utilizada de las integrales elípticas es la ***forma de Legendre***. Para esto es usada la ecuación

$$y^2 = (1-x^2)(1-k^2x^2).$$

Nosotros podemos pasar de la forma de Weierstrass a la forma de Legendre de la siguiente manera. Usando el cambio lineal de variable $x = ax_1+b$ pasamos de la expresión $4x^3-g_2x-g_3$ a $x_1(x_1-1)(x_1-k^2)$. En seguida, aplicando los cambios de variable $\xi^2 = x_1^{-1}$ y $\eta^2 = y^2x_1^{-3}$. Tenemos que $\eta^2 = (1-\xi^2)(1-k^2\xi^2)$.

Para la forma de Legendre, aparecen cuatro tipos de integrales:

$$\int \frac{dx}{\sqrt{G(x)}}, \int \frac{xdx}{\sqrt{G(x)}}, \int \frac{x^2dx}{\sqrt{G(x)}}, \int \frac{dx}{(x-c)\sqrt{G(x)}},$$

donde $G(x) = (1-x^2)(1-k^2x^2)$. Pero, como

$$\int \frac{xdx}{(1-x^2)(1-k^2x^2)} = \frac{1}{2}\int \frac{du}{(1-u)(1-k^2u)},$$

donde $u = x^2$, esta integral puede ser expresada en términos de funciones elementales.

Para simplificar bastante la forma de las integrales en la forma de Legendre, hagamos el cambio de variable $x = \operatorname{sen}\varphi$. Entonces

$$dx = \cos\varphi d\varphi, \qquad \sqrt{1-x^2} = \cos\varphi, \qquad \sqrt{1-k^2x^2} = \sqrt{1-k^2\operatorname{sen}^2\varphi}.$$

Por lo tanto, las integrales (no elementales) mencionadas anteriormente toman la forma

$$\int \frac{d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}}, \int \frac{\operatorname{sen}^2\varphi d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}}, \int \frac{d\varphi}{(\operatorname{sen}\varphi - c)\sqrt{1-k^2\operatorname{sen}^2\varphi}}.$$

Estas integrales son llamadas ***integrales elípticas de primer, segundo*** y ***tercer orden***, respectivamente.

Observemos que en lugar de la integral $\int \frac{\operatorname{sen}^2\varphi d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}}$ podemos tomar la integral $\int \sqrt{1-k^2\operatorname{sen}^2\varphi}d\varphi$ porque

$$k^2\int \frac{\operatorname{sen}^2\varphi d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}} = \int \frac{d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}} - \int \sqrt{1-k^2\operatorname{sen}^2\varphi}d\varphi.$$

Para las integrales elípticas de primer y segundo orden es usada la notación de Legendre:

$$F(\varphi) = \int_0^{\varphi} \frac{d\varphi}{\sqrt{1 - k^2 \operatorname{sen}^2 \varphi}} \quad y \quad E(\varphi) = \int_0^{\varphi} \sqrt{1 - k^2 \operatorname{sen}^2 \varphi} d\varphi.$$

2.7. Teorema de adición para las integrales elípticas $F(\varphi)$ y $E(\varphi)$

Primero consideraremos a

$$F(\varphi) = \int_0^{\varphi} \frac{d\varphi}{\Delta(\varphi)}$$

donde $\Delta(\varphi) = \sqrt{1 - k^2 \operatorname{sen}^2 \varphi}$. Si $F(\varphi) + F(\psi) = F(\mu)$, entonces $\operatorname{sen} \mu$ puede ser expresado algebraicamente in términos de $\operatorname{sen} \varphi$ y $\operatorname{sen} \psi$. Para probar esto, consideremos la ecuación diferencial

$$(*) \qquad \frac{d\varphi}{\sqrt{1 - k^2 \operatorname{sen}^2 \varphi}} + \frac{d\psi}{\sqrt{1 - k^2 \operatorname{sen}^2 \psi}} = 0.$$

Su integral $F(\varphi) + F(\psi)$ es constante y podemos suponer que es de la forma $F(\mu)$ donde μ es una constante. Considerando que F es una función impar, la integral puede ser expresada de la forma $F(\varphi) + F(\psi) + F(-\mu) = 0$. Mostremos ahora que la integral de la ecuación diferencial satisface la relación

$$\cos \varphi \cos \psi - \operatorname{sen} \varphi \operatorname{sen} \psi \sqrt{1 - k^2 \operatorname{sen}^2 \mu} = \cos \mu \tag{2.6}$$

la cual puede ser reescrita en una forma mas simétrica:

$$\cos^2 \varphi + \cos^2 \psi + \cos^2 \mu - 2 \cos \varphi \cos \psi \cos \mu + k^2 \operatorname{sen}^2 \varphi \operatorname{sen}^2 \psi \operatorname{sen}^2 \mu = 1. \tag{2.7}$$

Esto se obtiene de reordenar (2.6) así, $\cos \varphi \cos \psi - \cos \mu = \operatorname{sen} \varphi \operatorname{sen} \psi \sqrt{1 - k^2 \operatorname{sen}^2 \mu}$, elevar al cuadrado y reducir. El termino "simétrico" significa que no solo se satisface (2.6) sino también las siguientes dos relaciones

$$\cos \mu cos\varphi + \operatorname{sen} \mu \operatorname{sen} \varphi \sqrt{1 - k^2 \operatorname{sen}^2 \psi} = \cos \psi, \tag{2.8}$$

$$\cos \mu cos\psi + \operatorname{sen} \mu \operatorname{sen} \psi \sqrt{1 - k^2 \operatorname{sen}^2 \varphi} = \cos \varphi, \tag{2.9}$$

ya que los argumentos φ, ψ y $-\mu$ entran en (2.7) simétricamente.

Dividiendo ambos lados de (2.6) entre $\operatorname{sen} \varphi \operatorname{sen} \psi$ y derivando la expresión obtenida. El resultado puede ser expresado de la forma

$$d\varphi \left(\frac{\cos \psi - \cos \mu \cos \varphi}{\operatorname{sen} \varphi} \right) + d\psi \left(\frac{\cos \varphi - \cos \mu \cos \psi}{\operatorname{sen} \psi} \right) = 0.$$

Ocupando las fórmulas (2.8) y (2.9) tenemos que

$$\frac{d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}}+\frac{d\psi}{\sqrt{1-k^2\operatorname{sen}^2\psi}}=0.$$

Por esta razón, (2.6) es de hecho una integral de la ecuación diferencial $(*)$. Pero está no puede tener dos integrales independientes y, por lo tanto, la igualdad $F(\varphi)+F(\psi)=F(\mu)$ implica que

$$\cos\varphi\cos\psi-\operatorname{sen}\varphi\operatorname{sen}\psi\sqrt{1-k^2\operatorname{sen}^2\mu}=\cos\mu.$$

Esto nos provee de una expresión implícita para $\cos\mu$. No es difícil obtener también una expresión explícita. Indiquemos como, sea $x=\cos\mu$; entonces $\operatorname{sen}^2\mu=1-x^2$ y la relación (2.7) puede ser considerada como una expresión cuadrática en x. Resolviendo esta para x tenemos que

$$\cos\mu=\frac{\cos\varphi\cos\psi-\operatorname{sen}\varphi\operatorname{sen}\psi\Delta(\varphi)\Delta(\psi)}{1-k^2\operatorname{sen}^2\varphi\operatorname{sen}^2\psi}. \tag{2.10}$$

(El signo es tomado de tal forma que las fórmulas (2.10) y (2.6) son compatibles para φ y ψ pequeñas.)

Mediante transformaciones algebraicas podemos obtener las siguientes expresiones para $\operatorname{sen}\mu$ y $\Delta(\mu)$:

$$\operatorname{sen}\mu=\frac{\operatorname{sen}\varphi\cos\psi\Delta(\psi)+\operatorname{sen}\psi\cos\varphi\Delta(\varphi)}{1-k^2\operatorname{sen}^2\varphi\operatorname{sen}^2\psi}, \tag{2.11}$$

$$\Delta(\mu)=\frac{\Delta(\varphi)\Delta(\psi)-k^2\operatorname{sen}\varphi\operatorname{sen}\psi\cos\varphi\cos\psi}{1-k^2\operatorname{sen}^2\varphi\operatorname{sen}^2\psi}. \tag{2.12}$$

Dividiendo (2.11) entre (2.10) obtenemos

$$\tan\mu=\frac{\tan\varphi\Delta(\varphi)+\tan\psi\Delta(\psi)}{1-\tan\varphi\tan\psi\Delta(\varphi)\Delta(\psi)}. \tag{2.13}$$

La última fórmula puede ser interpretada de la manera siguiente. Sean los ángulos φ' y ψ' de tales que $\tan\varphi'=\tan\varphi\Delta(\varphi)$ y $\tan\psi'=\tan\psi\Delta(\psi)$. Entonces $\mu=\varphi'+\psi'$.

En aplicaciones el caso cuando $\mu=\frac{1}{2}\pi$ es bastante importante. En este caso $\cos\mu=0$ y $\operatorname{sen}\mu=1$. De (2.8) y (2.9) tenemos que $\operatorname{sen}\varphi=\cos\psi/\Delta(\psi)$ y $\operatorname{sen}\psi=\cos\varphi/\Delta(\varphi)$, y de (2.6) tenemos que $\cos\varphi\cos\psi=b\operatorname{sen}\varphi\operatorname{sen}\psi$, es decir, $b\tan\varphi\tan\psi=1$, donde $b=\sqrt{1-k^2}$. Entonces de (2.10) se sigue que

$$\cos\varphi\cos\psi=\Delta(\varphi)\Delta(\psi)\operatorname{sen}\varphi\operatorname{sen}\psi$$

y, por lo tanto, $\Delta(\varphi)\Delta(\psi)=b$. De ahí que,

$$\cos\varphi=\operatorname{sen}\psi\Delta(\varphi)=\frac{b\operatorname{sen}\psi}{\Delta\psi}\quad\text{y}\quad\cos\psi=\frac{b\operatorname{sen}\varphi}{\Delta\varphi}.$$

Las fórmulas (2.10) y (2.11) se parecen, hasta cierto punto, a las fórmulas para coseno y seno de la suma de dos ángulos. Con su ayuda podemos obtener expresiones similares para $\cos n\varphi$ y $\operatorname{sen} n\varphi$ en términos de $\cos\varphi$ y $\operatorname{sen}\varphi$. Sea $F(\varphi_n) = nF(\varphi)$. Entonces

$$\operatorname{sen}\varphi_2 = \frac{2\operatorname{sen}\varphi\cos\varphi\Delta(\varphi)}{1-k^2\operatorname{sen}^4\varphi}, \qquad \cos\varphi_2 = \frac{1-2\operatorname{sen}^2\varphi+k^2\operatorname{sen}^4\varphi}{1-k^2\operatorname{sen}^4\varphi},$$

$$\Delta(\varphi_2) = \frac{1-2k^2\operatorname{sen}^2\varphi+k^2\operatorname{sen}^4\varphi}{1-k^2\operatorname{sen}^4\varphi}, \quad \tan\varphi_2 = \frac{2\tan\varphi\Delta(\varphi)}{1-(\tan\varphi\Delta(\varphi))^2}.$$

Para encontrar φ de un φ_2 dado, podemos usar el hecho de que $\tan(\frac{\varphi_2}{2}) = \tan\varphi\Delta(\varphi)$ y también podemos resolver la ecuación

$$\cos\varphi_2 = \frac{1-2x^2-k^2x^4}{1-k^2x^4},$$

donde $x = \operatorname{sen}\varphi$. Esta ecuación corresponde a la división de $F(\varphi_2)$ en mitades.

Para dividir $F(\psi)$ en tres partes iguales, debemos resolver la ecuación

$$\operatorname{sen}\psi = \frac{3x-4(1-k^2)x^3+6k^2x^5-k^4x^9}{1-6k^2x^4+4k^2(1+k^2)x^6-3k^4x^8},$$

donde $x = \operatorname{sen}\varphi$. Para $\psi = \frac{\pi}{2}$ tenemos la ecuación

$$(1+x)(1-2x+2k^2x^3-k^2x^4)^2 = 0,$$

es decir, la división de $F(\pi/2)$ en tres partes iguales se reduce a la solución de la ecuación

$$1-2\operatorname{sen}\varphi+2k^2\operatorname{sen}^3\varphi+k^2\operatorname{sen}^4\varphi = 0.$$

Para la integral elíptica de segundo orden $E(\varphi)$, solo hay un teorema de adición, con un término algebraico extra. Este teorema está directamente relacionado con el teorema de adición para la integral elíptica de primer orden, $F(\varphi)$.

Teorema 2.7.1. *Si* $F(\varphi)+F(\psi)-F(\mu) = 0$ *entonces*

$$E(\varphi)+E(\psi)-E(\mu) = k^2\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu.$$

Demostración. Sea $E(\varphi)+E(\psi)-E(\mu) = P(\varphi,\psi,\mu)$. Tomemos la diferencial de esta igualdad con un valor constante μ.

Como resultado tenemos que

$$\Delta(\varphi)d\varphi+\Delta(\psi)d\psi = dP.$$

Pero por (2.8) y (2.9)

$$\Delta(\varphi) = \frac{\cos\varphi-\cos\psi\cos\mu}{\operatorname{sen}\psi\operatorname{sen}\mu} \quad \text{y} \quad \Delta(\psi) = \frac{\cos\psi-\cos\varphi\cos\mu}{\operatorname{sen}\varphi\operatorname{sen}\mu}.$$

Por esta razón,

$$\begin{aligned} dP &= \left(\frac{\cos\varphi - \cos\psi\cos\mu}{\operatorname{sen}\psi\operatorname{sen}\mu}\right)d\varphi + \left(\frac{\cos\psi - \cos\varphi\cos\mu}{\operatorname{sen}\varphi\operatorname{sen}\mu}\right)d\psi \\ &= \frac{d(\operatorname{sen}^2\varphi + \operatorname{sen}^2\psi + 2\cos\varphi\cos\psi\cos\mu)}{2\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu}. \end{aligned}$$

Por (2.7) tenemos que

$$1 - \operatorname{sen}^2\varphi + 1 - \operatorname{sen}^2\psi - 2\cos\varphi\cos\psi\cos\mu = 1 - \cos^2\mu - k^2\operatorname{sen}^2\varphi\operatorname{sen}\psi\operatorname{sen}\mu.$$

De ahí que,

$$dP = \frac{d(k\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu)^2}{2\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu} = k^2 d(\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu).$$

Ya que P y $\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu$ se anulan en $\varphi = 0$, tenemos que $P = k^2\operatorname{sen}\varphi\operatorname{sen}\psi\operatorname{sen}\mu$. □

Este teorema de adición nos permite obtener las siguientes expresiones para $nE(\varphi) - E(\varphi_n)$, donde $F(\varphi_n) = nF(\varphi)$:

$$\begin{aligned} 2E(\varphi) - E(\varphi_2) &= k^2\operatorname{sen}^2\varphi\operatorname{sen}\varphi_2, \\ 3E(\varphi) - E(\varphi_3) &= (2E(\varphi) - E(\varphi_2)) + (E(\varphi_2) + E(\varphi) - E(\varphi_3)) \\ &= k^2\operatorname{sen}\varphi\operatorname{sen}\varphi_2(\operatorname{sen}\varphi + \operatorname{sen}\varphi_3), \end{aligned}$$

etcetera.

2.8. Las funciones elípticas de Jacobi

La función considerada en la sección anterior

$$F(\varphi) = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \int_0^\varphi \frac{d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}},$$

donde $x = \operatorname{sen}\varphi$, tiene mucha analogía con la función

$$\operatorname{arc\,sen} x = \int_0^x \frac{dx}{\sqrt{1-x^2}} = \int_0^\varphi d\varphi,$$

donde $x = \operatorname{sen}\varphi$. Por ejemplo si $F(\varphi) + F(\psi) = F(\mu)$, entonces $\operatorname{sen}\mu$ es expresado en términos de $\operatorname{sen}\varphi$ y $\operatorname{sen}\psi$ por la fórmula

$$\operatorname{sen}\mu = \frac{\operatorname{sen}\varphi\cos\psi\Delta(\psi) + \operatorname{sen}\psi\cos\varphi\Delta\varphi}{1 - k^2\operatorname{sen}^2\varphi\operatorname{sen}^2\psi}, \tag{2.14}$$

de igual manera si $\operatorname{arc\,sen} x + \operatorname{arc\,sen} y = \operatorname{arc\,sen} z$, donde $x = \operatorname{sen}\varphi$, $y = \operatorname{sen}\psi$, y $z = \operatorname{sen}\mu$, entonces

$$\operatorname{sen}\mu = sen\varphi\cos\psi + \operatorname{sen}\psi\cos\varphi. \tag{2.15}$$

Esta analogía no es accidental, para $k = 0$ la función $F(\varphi)$ se vuelve la función arc sen $x = \varphi$ y la fórmula (2.14) se convierte en (2.15).

Por muchas razones es mas conveniente considerar la función $\varphi = \operatorname{arc\,sen} x$ en lugar de la función inversa $x = \operatorname{sen}\varphi$. La función inversa de la función $F(\varphi)$ es incluso en varios sentidos más conveniente que la función $F(\varphi)$ misma. Reemplacemos ahora la notación de Legendre $F(\varphi)$ por la notación de Jacobi, para ello sea $u(\varphi) = F(\varphi)$. La función $\varphi(u)$ inversa de $u(\varphi)$ es llamada la ***amplitud*** de u y es denotada por $\varphi = \operatorname{am} u$. Obtuvimos las fórmulas (2.10)-(2.12) para las funciones $\operatorname{sen}\varphi$, $\cos\varphi$ y $\Delta(\varphi) = \sqrt{1 - k^2 \operatorname{sen}^2\varphi}$. Introduciendo las funciones

$$\operatorname{sn} u = \operatorname{sen}\operatorname{am} u,\ \operatorname{cn} u = \cos\operatorname{am} u \quad \text{y} \quad \operatorname{dn} u = \Delta(\operatorname{am} u)$$

podemos expresar las fórmulas (2.10)-(2.12) de la siguiente manera:

$$\operatorname{cn}(u+v) = \frac{\operatorname{cn} u \operatorname{cn} v - \operatorname{sn} u \operatorname{sn} v \operatorname{dn} u \operatorname{dn} v}{1 - k^2 \operatorname{sn}^2 u \operatorname{sn}^2 v}, \tag{2.16}$$

$$\operatorname{sn}(u+v) = \frac{\operatorname{sn} u \operatorname{cn} v \operatorname{dn} v + \operatorname{sn} v \operatorname{cn} u \operatorname{dn} u}{1 - k^2 \operatorname{sn}^2 u \operatorname{sn}^2 v}, \tag{2.17}$$

$$\operatorname{dn}(u+v) = \frac{\operatorname{dn} u \operatorname{dn} v - k^2 \operatorname{sn} u \operatorname{sn} v \operatorname{cn} u \operatorname{cn} v}{1 - k^2 \operatorname{sn}^2 u \operatorname{sn}^2 v}. \tag{2.18}$$

Las funciones $\operatorname{sn} u$, $\operatorname{cn} u$ y $\operatorname{dn} u$ usualmente son llamadas funciones ***elípticas de Jacobi*** aunque muchas propiedades de estas funciones fueron establecidas, hasta cierto punto, por Legendre y Abel antes que Jacobi.

Una de las mas importantes propiedades de las funciones $\operatorname{sn} u$, $\operatorname{cn} u$ u $\operatorname{dn} u$ es su doble periodicidad. La existencia de un periodo para estas funciones es bastante obvio. De hecho, las funciones $\operatorname{sen}\varphi$ y $\cos\varphi$ tienen periodo $2\pi = 4(\frac{\pi}{2})$ y la función $\operatorname{sen}^2\varphi$ tiene periodo $\pi = 2(\frac{\pi}{2})$. Por lo tanto, las funciones $\operatorname{sn} u$ y $\operatorname{cn} u$ tienen periodo $4K$, donde

$$K = \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2 \operatorname{sen}^2\varphi}};$$

y la función $\operatorname{dn} u = \sqrt{1 - k^2 \operatorname{sn}^2 u}$ tiene periodo 2K.

Con la ayuda de las fórmulas de adición (2.16)-(2.18) veamos como cambian las funciones $\operatorname{sn} u$, $\operatorname{cn} u$ y $\operatorname{dn} u$ si el argumento se incrementa en un cuarto del periodo, K, y en un medio del periodo, 2K. Sustituyendo

$$\operatorname{sn} K = 1, \quad \operatorname{cn} K = 0 \quad \text{y} \quad \operatorname{dn} K = \sqrt{1 - k^2}$$

en (2.16)-(2.18) tenemos que

$$\operatorname{sn}(u+K) = \frac{\operatorname{cn} u}{\operatorname{dn} u},\ \operatorname{cn}(u+K) = -\frac{\sqrt{1-k^2}\operatorname{sn} u}{\operatorname{dn} u} \quad \text{y} \quad \operatorname{dn}(u+K) = \frac{\sqrt{1-k^2}}{\operatorname{dn} u}.$$

Puesto que $\operatorname{sn} 2K = 0$, $\operatorname{cn} 2K = -1$ y $\operatorname{dn} 2K = 1$, se sigue que

$$\operatorname{sn}(u+2K) = -\operatorname{sn} u,\ \operatorname{cn}(u+K) = -\operatorname{cn} u \quad \text{y} \quad \operatorname{dn}(u+2K) = \operatorname{dn} u.$$

Es bastante mas difícil imaginar cual será el otro periodo de las funciones elípticas de Jacobi. Primero recordemos que la integral

$$\int \frac{d\varphi}{\sqrt{1-k^2\operatorname{sen}^2\varphi}}$$

se obtuvo de la integral

$$\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

con la ayuda del cambio de variable $x = \operatorname{sen}\varphi$. Ahora será mas conveniente trabajar con la integral original. En el plano complejo $\mathbb{C}$, la función

$$u(x) = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

no esta definida generalmente porque el valor de $u(x)$ depende de la trayectoria de integración que se une para si de 0 a x. Los valores de la función $u(x)$ en el mismo punto pueden diferir por un números de la forma

$$L = \int_C \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}},$$

donde la integral está tomada a lo largo de una trayectoria cerrada C. Aquí cada uno de los números L es un periodo de la función inversa $x(u)$. El integrando tiene puntos singulares en $\pm 1, \pm k^{-1}$.

Veamos ahora como la elección de la trayectoria alrededor de estos puntos afecta el valor de las funciones $\operatorname{sn} u = x$, $\operatorname{cn} u = \sqrt{1-x^2}$ y $\operatorname{dn} u = \sqrt{1-k^2x^2}$.

Sean

$$\int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = K \quad \text{y} \quad \int_0^X \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = \alpha.$$

La trayectoria mostrada en la Figura 2.9 muestra que los valores de la integral en el punto X son iguales a α y a $K+(K-\alpha) = 2K-\alpha$. De hecho, en la última parte de la trayectoria el signo de la función $\sqrt{1-x^2}$ y la dirección del segmento de integración cambian, como resultado de estos dos cambios el signo de la integral no cambia. Por lo tanto, $\operatorname{sn}\alpha = \operatorname{sn}(K-\alpha)$. Además, en esta trayectoria el signo de la función $\sqrt{1-x^2}$ cambia pero el signo de $\sqrt{1-k^2x^2}$ no. De ahí que, $\operatorname{cn}\alpha = -\operatorname{cn}(2K-\alpha)$ y $\operatorname{dn}\alpha = \operatorname{dn}(2K-\alpha)$. Reemplazando α por $-\alpha$ tenemos que

$$\begin{aligned}
&\operatorname{sn}(\alpha+2K) = \operatorname{sn}(-\alpha) = -\operatorname{sn}\alpha,\\
&\operatorname{cn}(\alpha+2K) = -\operatorname{cn}(-\alpha) = -\operatorname{cn}\alpha,\\
&\operatorname{dn}(\alpha+2K) = \operatorname{dn}(-\alpha) = \operatorname{dn}\alpha.
\end{aligned}$$

Hemos obtenido estas fórmulas por otro método.

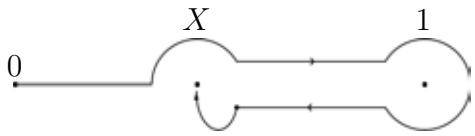


Figura 2.9

Ahora, consideremos la trayectoria a lo largo de la curva mostrada en la Figura 2.10. Sea

$$\int_1^{k^{-1}} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} = iK'.$$

(El número K' es real puesto que el número $\sqrt{1-x^2}$ es imaginario puro para $x \in (1, k^{-1})$). Por lo tanto, los valores de la integral en el punto X son iguales a α y a $K+iK'+iK'-(K-\alpha)$. Aquí sólo el signo del último sumando necesita aclaración.

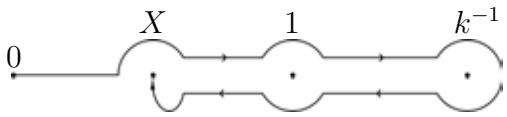


Figura 2.10

Observe que ahí ocurren tres cambios de signo: la dirección del contorno de integración a cambiado, y por esto los signos de las funciones $\sqrt{1-x^2}$ y $\sqrt{1-k^2x^2}$ también. Como resultado, tenemos que

$$\begin{aligned} \operatorname{sn}\alpha &= \operatorname{sn}(\alpha + 2iK'), \\ \operatorname{cn}\alpha &= -\operatorname{cn}(\alpha + 2iK'), \\ \operatorname{dn}\alpha &= -\operatorname{dn}(\alpha + 2iK'). \end{aligned}$$

Por lo tanto, las funciones $\operatorname{sn} u$, $\operatorname{cn} u$ y $\operatorname{dn} u$ tienen periodos $2iK'$, $4iK'$, y $4iK'$, respectivamente. Además, la función $\operatorname{cn} u$ tiene periodo $2K + 2iK'$. De hecho,

$$\operatorname{cn}(\alpha + 2iK' + 2K) = cn(\alpha + 2iK') = \operatorname{cn}\alpha.$$

Por lo tanto, la función $\operatorname{sn} u$ tiene periodos $4K$ y $2Ki$; la función $\operatorname{cn} u$ tiene periodos $4K$ y $2K + 2iK'$; y la función $\operatorname{dn} u$ tiene periodos $2K$ y $4iK'$.

2.9. El teorema de Weierstrass sobre funciones que poseen un teorema algebraico de adición

Diremos que una función meromorfa $\varphi(z)$ ***posee un teorema algebraico de adición*** si existe un polinomio F en tres variables distinto de cero tal que

$$F(\varphi(z_1 + z_2), \varphi(z_1), \varphi(z_2)) = 0;$$

esta identidad significa que $\varphi(z_1 + z_2)$ puede ser expresado algebraicamente en términos de $\varphi(z_1)$ y $\varphi(z_2)$.

Por ejemplo, la función $\operatorname{sn} z$ posee un teorema algebraico de adición (que se sigue de (2.17)). De hecho, si $a = \operatorname{sen}(z_1 + z_2)$, $b = \operatorname{sn} z_1$ y $c = \operatorname{sn} z_2$, entonces

$$a = \frac{b\sqrt{1-c^2}\sqrt{1-k^2c^2} + c\sqrt{1-b^2}\sqrt{1-k^2b^2}}{1-k^2b^2c^2}.$$

Elevando al cuadrado dos veces podemos librarnos de los radicales y obtener una relación polinomial

$$F(a, b, c) = 0.$$

La función de Weierstrass $\wp(z)$ también posee un teorema algebraico de adición. De hecho, si $a = \wp(z_1 + z_2)$, $b = \wp(z_1)$ y $c = \wp(z_2)$, entonces

$$a = -b - c + \frac{1}{4}\left(\frac{\sqrt{4b^3 - g_2b - g_3} - \sqrt{4c^3 - g_2c - g_3}}{b-c}\right)^2.$$

Es posible simplificar esta fórmula para obtener una relación de la forma $F(a, b, c) = 0$, donde F es un polinomio.

Como vimos en la Sección 2.3, cualquier función elíptica puede ser representada de la forma

$$f = R(\wp) + R_1(\wp)\wp',$$

donde R y R_1 son funciones racionales. Esta representación nos permite obtener un teorema algebraico de adición para una función elíptica arbitraria si tomamos en cuenta que

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3.$$

De hecho, si $R_1 \neq 0$, entonces $\wp' = \frac{f - R(\wp)}{R_1(\wp)}$. Por esta razón,

$$\left(\frac{f - R(\wp)}{R_1(\wp)}\right)^2 = 4\wp^3 - g_2\wp - g_3.$$

Y ésta última relación puede ser reescrita de la forma $P(f, \wp) = 0$, donde P es un polinomio (para el cual su grado con respecto a f es igual a 2). Si, $R_1 = 0$, la relación $f = R(\wp)$ también puede ser representada en la forma expresada. Sean $A = f(z_1 + z_2)$, $B = f(z_1)$ y $C = f(z_2)$. De las relaciones $F(a, b, c) = 0$ y $P(A, a) = 0$ podemos obtener una relación $F_1(A, b, c) = 0$ calculando el resultante de los polinomios $f(a) = F(a, b, c)$ y $g(a) = P(A, a)$.

Después, de las relaciones $F_1(A, b, c) = 0$ y $P(B, b) = 0$ obtenemos $F_2(A, B, c) = 0$ y de las relaciones $F_2(A, B, c) = 0$ y $P(C, c) = 0$ obtenemos la relación $G(A, B, C) = 0$, como requeríamos.

No solamente las funciones elípticas poseen un teorema algebraico de la adición. Por ejemplo, la función ***exponencial*** e^z posee un teorema algebraico de adición, ya que $e^{z_1+z_2} = e^{z_1}e^{z_2}$. Aún más si $u = R(f)$, donde R es una función racional, entonces $P(u, f) = 0$, para algún polinomio P (lineal en u). Para ver esto, definamos $F(a, b, c)$ así

$$F(a,b,c) = \begin{cases} a-(b+c) & \text{si } f = z, \\ a - bc & \text{si } f = e^z, \end{cases}$$

podemos obtener, como antes, una relación $G(A, B, C) = 0$, donde $A = R(a)$, $B = R(b)$ y $C = R(c)$. Por lo tanto, cualquier función racional y también cualquier función racional en $e^{\lambda z}$ posee un teorema algebraico de adición. Resulta que los ejemplos descritos anteriormente agotan todas las funciones meromorfas que poseen un teorema algebraico de adición.

Teorema 2.9.1 (Weierstrass)**.** *Toda función meromorfa $\varphi(z)$ que posea un teorema algebraico de adición o es una función elíptica o es de la forma $R(z)$ o $R(e^{\lambda z})$, donde $R(z)$ es una función racional.*

Demostración. (W. S. Osgood). En un dominio finito una función meromorfa tiene como puntos singulares solamente a sus polos. Si los límites

$$\lim_{z\to\infty} \varphi(z) \quad \text{ó} \quad \lim_{z\to\infty} \frac{1}{\varphi(z)}$$

existen, entonces la función $\varphi(z)$ es racional. De hecho, restando a φ la suma de sus partes principales en todos los polos (si el punto ∞ es un polo, entonces la parte principal de la función en este punto es de la forma $a_r z^r + a_{r+1}z^{r+1} + \cdots$, donde $r > 0$), dan como resultado, una función f sin puntos singulares; el punto ∞ también es un punto no singular. Por lo tanto, f es una constante y la función inicial φ es racional.

En lo que sigue supondremos que la función φ no es racional, es decir, el punto ∞ es una singularidad esencial. Para probar el teorema de Weierstrass, necesitaremos el siguiente teorema sobre el comportamiento de una función en una vecindad de un punto singular esencial.

Gran Teorema de Picard. *Cualquier función analítica $\varphi(z)$ toma en una vecindad arbitraria de un punto singular esencial cualquier valor finito excepto, quizás, un valor.*

La demostración de este teorema puede se encontrada en [9], pág. 240.

Sea $F(\varphi(z_1 + z_2), \varphi(z_1), \varphi(z_2)) = 0$, donde F es un polinomio de su grado con respeto al primer argumento igual a n. Tenemos que probar que φ es una función periódica y si no es doblemente periódica, entonces $\varphi(z) = R(e^{\lambda})$. El gran teorema de Picard implica que en una vecindad de un punto singular esencial la función φ toma un cierto valor c infinitas veces.

Sean $a_1, \ldots, a_{n+1}$ puntos en los cuales φ toma el valor c. Los puntos singulares de φ, junto con los puntos z para los cuales los puntos $z + a_i$ son singulares, forman un conjunto de medida cero. Por lo tanto, existe un punto no singular z_0 de la función φ para el cual todos

los puntos $a_i + z_0$ son también no singulares. Entonces los puntos z y $a_i + z$ para valores z suficientemente cerca de z_0 son no singulares. Para tales puntos z considera la ecuación

$$F(x, \varphi(z), c) = 0. \tag{2.19}$$

Esta tiene $n+1$ raíces $x_i = \varphi(z + a_i)$, porque

$$F(\varphi(z + a_i), \varphi(z), c) = F(\varphi(z + a_i), \varphi(z), \varphi(a_i)) = 0.$$

La ecuación (2.19) es un polinomio en x distinto de cero de grado n y, por lo tanto, esta tiene a lo mas n raíces distintas. De ahí se sigue que $\varphi(z + a_p) = \varphi(z + a_q)$ para ciertos p y q distintos. Tal relación se satisface por cualquier punto z de una vecindad de z_0, pero los pares (p, q) pueden diferir.

Sin embargo, solo hay un número finito de pares (p, q) y, por lo tanto, alguna relación $\varphi(z + a_p) = \varphi(z + a_q)$ se cumple para un conjunto infinito de puntos z. Estos punto tienen un punto limite z_1 y la función φ es regular en z_1. Por el teorema de unicidad vemos que las funciones $\varphi(z + z_p)$ y $\varphi(z + a_q)$ coinciden, es decir, $a_p - a_q$ es un periodo de φ.

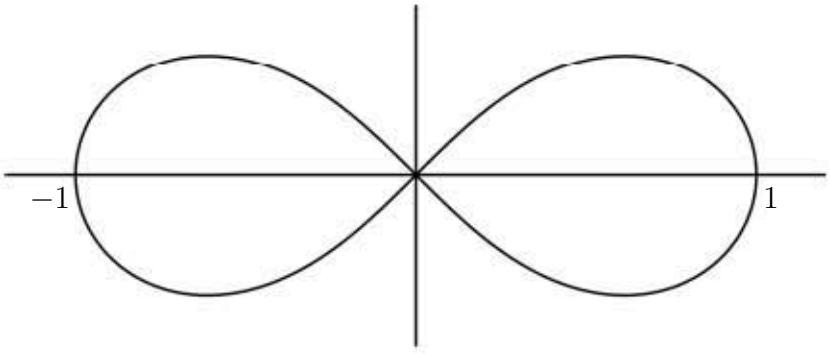
Tenemos probado que φ es una función periódica; por definición, podemos suponer que el periodo mínimo de φ es igual a 2π. Supongamos que φ no tiene otros periodos y entonces mostremos que $\varphi(z) = R(w)$, donde $w = e^{iz}$ y R es una función racional. La función $z \mapsto w = e^{iz}$ manda la franja $0 \leq \operatorname{Re} z < 2\pi$ en el plano con el corte de 0 a ∞. La función $\psi(w) = \varphi(z)$ es meromorfa en $\widehat{\mathbb{C}}\backslash\{0, \infty\}$. Si estos puntos no son singularidades esenciales, entonces la función ψ es racional.

Supongamos que al menos uno de estos puntos es una singularidad esencial. Entonces por el gran teorema de Picard existen puntos $b_1, \ldots, b_{n+1}$ tales que $\psi(b_i) = c$. Las preimágenes $\beta_1, \ldots, \beta_{n+1}$ de estos puntos con respecto a la función $z \mapsto w$ son distintos y pertenecen a la franja $0 \leq \operatorname{Re} z < 2\pi$. La ecuación $F(x, \varphi(z), c) = 0$ tiene raíces $x_i = \varphi(\beta_i + z)$. Repitiendo los mismos argumentos anteriores, vemos que la función φ tiene un periodo $\beta_p - \beta_q$, donde $0 \leq \operatorname{Re} \beta_p, \operatorname{Re} \beta_q < 2\pi$. Por lo tanto, φ tiene otro periodo aparte del periodo puramente real 2π o tiene un periodo real mas pequeño que 2π. El último caso es imposible, ya que supusimos que el periodo 2π era el más pequeño. □

Capítulo 3

Lemniscata

La lemniscata es una curva plana definida por la ecuación $(x^2 + y^2)^2 = x^2 - y^2$ y su gráfica es:



El astrónomo **Cassini** fue el primero en estudiar la lemniscata. El incluso consideró curvas más generales para las cuales el producto de las distancias de uno de sus puntos a dos puntos fijos es una constante. En las bases de las observaciones astronómicas Cassini creyó que con las ayuda de tales curvas se podría describir más precisamente el movimiento de los planetas que con la ayuda de las elipses. Ahora estas curvas con llamadas ***ovalos de Cassini***.

Pero el libro de Cassini *Eléments d'astronomie*, en el cual fueron estudiadas, fue publicado en 1749, muchos años después de su muerte. Para la comunidad matemática la lemniscata fue conocida por los artículos de **Jacob Bernoulli** y **Johann Bernoulli** publicados en 1694 y, por lo tanto, usualmente es llamada la ***lemniscata de Bernoulli***.

Las propiedades más importantes de la lemniscata fueron señaladas por el matemático italiano Conde **Fagnano** (1682-1766). Fagnano descubrió que la longitud de arco de la lemniscata puede ser expresada en términos de una integral elíptica de primer orden. Por cierto, él fue quien creo el término ***integrales elípticas***. El obtuvo un teorema de adición para esta integral y, por lo tanto, demostró que la división de arcos de la lemniscata en n

partes iguales es un problema algebraico. En 1750 Fagnano publicó una colección de sus artículos bajo el nombre *Produzioni matematiche*.

En ese momento Fagnano sabía que la división de la lemniscata podía reducirse a la solución de una ecuación algebraica. Sin embargo, los métodos para investigar la solubilidad de ecuaciones en cuadraturas (es decir, via raíces cuadradas) no estaban desarrollados en ese tiempo. El primero en lograr avances esenciales en este campo fue a sus 19 años **Gauss** (en 1796). El descubrió que un 17-agono regular podía ser construido con regla y compás, es decir, que la ecuación $x^{17}-1=0$ es soluble en raíces cuadradas. Más tarde, Gauss mostró que con regla y compás uno podía construir un n-agono regular para cualquier n de la forma $2^{\alpha}p_1\cdots p_k$, donde los p_i son ***primos de Fermat*** distintos, es decir, primos de la forma $2^{2^m}+1$. Gauss escribió que para cualquier otro n era imposible construir un n-agono con regla y compás pero no existen evidencias suficientes de que en verdad él sabía demostrar esto.

Gauss también estaba interesado en la ecuación para la división de la lemniscata. Por ejemplo, el mostró que una ecuación de grado 25 relacionada con la división de la lemniscata en 5 partes iguales se podía resolver en raíces cuadradas. Sus argumentos estaban basados en el hecho de que el numero 5 puede ser representado como el producto de $2+i$ por $2-i$. Gauss no publicó esta investigación pero en su libro *Disquisitiones Arithmeticae*, el cual fue publicado en 1801, mencionó que los métodos que el había desarrollado eran aplicables no solamente a las funciones trigonométricas sino también a las funciones relacionadas a la integrales de la forma $\int \frac{dx}{\sqrt{1-x^4}}$.

Esta afirmación intrigó a **N. Abel**. Abel investigó en detalle la ecuación para la división de la lemniscata y probó que la lemniscata podía ser dividida en n partes iguales para todo numero n de la forma $2^{\alpha}p_1\cdots p_k$, donde los p_i son primos de Fermat distintos (es decir, los mismos numeros descritos por Gauss). Abel consideró su teorema como uno de sus mas importantes resultados. Este resultado esta contenido en la segunda parte de su trabajo más grande *Recherches sur les fonctions elliptiques*. La prueba de Abel es algo larga y complicada.

Más tarde, **F. Eisenstein** (1823-1852) obtuvo una prueba mas simple. Haciendo esto Eisenstein descubrió interesantes propiedades de los polinomios relacionados con la división de la lemniscata. La prueba de Eisenstein sigue siendo un poco complicada.

En tiempos relativamente recientes Rosen encontró en 1981 una elegante prueba del teorema de Abel. Su prueba es bastante más simple y, mas aún, en su demostración resalta el papel decisivo de la invariancia de la latiz de periodos de la función lemniscática con respecto a la multiplicación por i, ésta la presentaremos aquí.

Además de la prueba encontrada por Rosen, mostraremos una prueba un poco más simple propuesta por Cox en [3].

Abel solamente probó la posibilidad de dividir la lemniscata en n partes iguales con regla y compás para los valores indicados de n. El no probó que para los otros valores de n la construcción es imposible. En la prueba de Rosen se muestra que para otros valores de n es imposible construir las coordenadas de los puntos que dividan la lemniscata en n

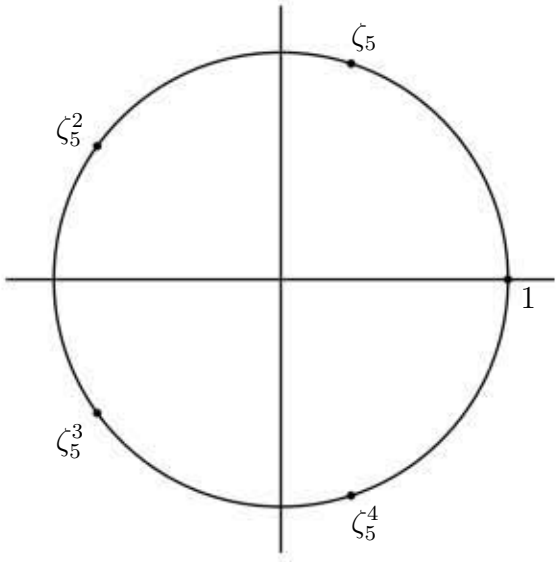
partes iguales con regla y compás. Esto, sin embargo, no significa que para otros valores de n es imposible dividir la lemniscata en n partes iguales con regla y compás si la lemniscata esta *ya dibujada*. De hecho, si la lemniscata esta dada, hay posibilidades adicionales para nuevas construcciones. Considerando los puntos de intersección de rectas y circunferencias con la lemniscata uno puede, en general, construir algo más que solamente raíces de funciones cuadráticas con coeficientes racionales.

3.1. Puntos de división y longitud de arco

Para formular cuidadosamente el teorema de Abel en la lemniscata, necesitamos definir los *n-puntos de división* de la lemniscata y estudiar la longitud de arco de esta curva.

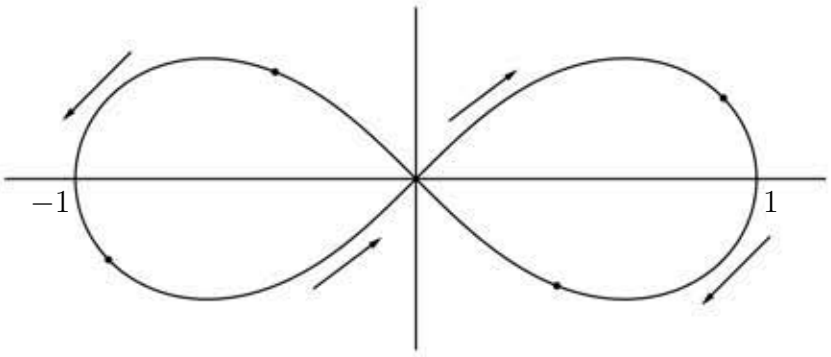
3.1.1. Puntos de división de la lemniscata

Las n raíces de la unidad ayudan a determinan cuando un n-agono regular puede ser construido con regla y compás (ver Apéndice A.3). En términos de la circunferencia unitaria centrada en el origen, las n raíces de la unidad dividen la circunferencia en n segmentos de la misma longitud, empezando en el punto $(1, 0)$. Para $n = 5$, las cinco raíces de la unidad 1, ζ_5, ζ_5^2, ζ_5^3, ζ_5^4 dividen la circunferencia de la siguiente forma

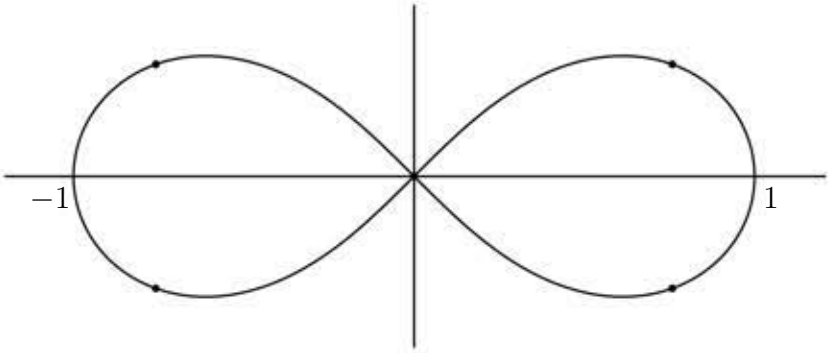


En general, las n raíces de la unidad $\{\zeta_n^i\}_{i=0,\ldots,n-1}$, son los ***n-puntos de división*** de la circunferencia unitaria. Entonces el teorema de Gauss sobre la circunferencia puede ser reformulado de la siguiente manera: los *n-puntos de división* de la circunferencia unitaria pueden construirse con regla y compás si y sólo si n es una potencia de 2 por un producto de distintos primos de Fermat.

Abel, siguiendo la idea de Gauss, se formuló la misma pregunta para la lemniscata. Aquí, los ***n*-puntos de división de la lemniscata** se obtienen de la siguiente manera. Comenzando en el origen y siguiendo la curva por el primer cuadrante, bajando por el cuarto cuadrante, volviendo al origen para seguir la curva por el segundo cuadrante, continuando por el tercer cuadrante, y finalmente regresar al origen. Este recorrido completa la longitud de arco total de la lemniscata. Para $n = 5$, los 5-puntos de división dividen la lemniscata de la siguiente forma



Así los n-puntos de división, dividen la lemniscata en n segmentos de la misma longitud. Cuando n es impar, como en la figura anterior, el segmento de en medio cruza el origen. Cuando n es par, los n-puntos de división son simétricos con respecto al eje x y al eje y, con el punto de división de en medio en el origen. Para $n = 6$, los 6-puntos de división de la lemniscata son de la siguiente forma, note que el origen es un punto doble.



Los n-puntos de división de la lemniscata nos llevarán a algunos importantes polinomios análogos a los polinomios ciclotómicos. La teoría de Galois de estos polinomios nos permitirá entender cuando los n-puntos de división pueden construirse con regla y compás.

Al principio, definimos la lemniscata usando la ecuación Cartesiana $(x^2+y^2)^2 = x^2-y^2$. Al expresar la ecuación en su forma polar, ésta toma la siguiente forma

$$r^2 = \cos(2\theta). \tag{3.1}$$

En efecto si $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$, se tiene que $x^2 + y^2 = r^2(\cos^2\theta - \mathrm{sen}^2\,\theta) = r^2\cos 2\theta$ y $(x^2 + y^2)^2 = r^4$, al igualar y cancelar se obtiene (3.1).

La coordenada polar r juega un papel central en la división de la lemniscata. Una razón es que para construir un punto de la lemniscata, sólo necesitamos r. Esto puede parecer obvio ya que obtendremos el punto deseado (y sus imágenes reflejadas con respecto a los ejes x y y) intersecando la lemniscata con la circunferencia de radio r. Pero de hecho la lemniscata no es necesaria. En otras palabras, si $0 < r < 1$ es un número que se puede construir, entonces también lo son las coordenadas x y y de los cuatro puntos de la lemniscata a distancia r del origen. Para ver esto, usaremos que $(x^2 + y^2)^2 = x^2 - y^2$ y que $r^2 = x^2 + y^2$. Esto nos proporcionan las ecuaciones

$$r^4 = x^2 - y^2 \quad \text{y} \quad r^2 = x^2 + y^2.$$

Resolviendo para x y y en términos de r, obtenemos

$$x = \pm\sqrt{\tfrac{1}{2}(r^2 + r^4)} \quad \text{y} \quad y = \pm\sqrt{\tfrac{1}{2}(r^2 - r^4)}.$$

Ya que por el Teorema A.2.2 del Apéndice los números construibles forman un subcampo de $\mathbb{C}$ cerrado bajo raíces cuadradas, observamos que x y y son números construibles cuando r lo es. De esta manera, para probar si un punto dado sobre la lemniscata es construible con regla y compás, es suficiente mostrar que la correspondiente coordenada polar r es construible. Notemos también que el inverso se cumple, es decir, si x y y son construibles entonces también lo es $r = \sqrt{x^2 + y^2}$. Hemos probado el siguiente resultado.

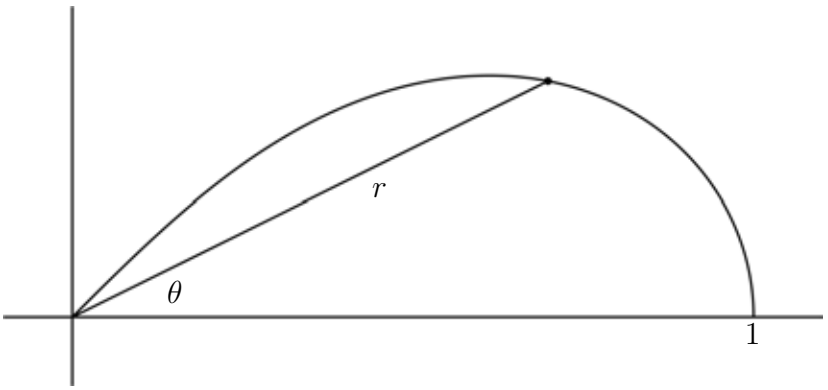
Proposición 3.1.1. *Sea P un punto sobre la lemniscata, y sea r la distancia de P al origen. Entonces P puede construirse con regla y compás si y sólo si r es un número construible.* □

3.1.2. Longitud de arco de la lemniscata

Los n-puntos de división de la lemniscata están definidos en términos de la longitud de arco. Por esta razón necesitamos estudiar la longitud de arco de la lemniscata, partamos de la ecuación polar de la lemniscata,

$$r^2 = \cos(2\theta).$$

Si nos fijamos en el primer cuadrante, entonces obtenemos la figura



Resolviendo la ecuación anterior para θ obtenemos $\theta = \frac{1}{2}\cos^{-1}(r^2)$. Esto convierte a θ en una función con respecto a r. Notemos que θ decrece de $\frac{\pi}{4}$ a 0 cuando r se incrementa de 0 a 1.

Recordemos que la longitud de arco en coordenadas cartesianas y polares está dada por

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{dr^2 + r^2 d\theta^2}.$$

De aquí se sigue que ***la longitud de arco de la lemniscata*** desde el origen a un punto en el primer cuadrante con coordenadas polares (r_0, θ_0) es

$$\text{longitud de arco} = \int_0^{r_0} \sqrt{1 + r^2 \left(\frac{d\theta}{dr}\right)^2}\, dr.$$

Diferenciando $r^2 = \cos(2\theta)$ con respecto a r tenemos $2r = -\operatorname{sen}(2\theta) \cdot 2\frac{d\theta}{dr}$, por lo tanto

$$1 + r^2(\frac{d\theta}{dr})^2 = 1 + r^2(-\frac{r}{\operatorname{sen}(2\theta)})^2 = 1 + \frac{r^4}{\operatorname{sen}^2(2\theta)}.$$

Como $\operatorname{sen}^2(2\theta) = 1 - \cos^2(2\theta) = 1 - r^4$, obtenemos

$$1 + r^2 \left(\frac{d\theta}{dr}\right)^2 = 1 + \frac{r^4}{1 - r^4} = \frac{1}{1 - r^4}.$$

Por lo tanto nuestra fórmula para la longitud de arco se transforma en

$$\text{longitud de arco} = \int_0^{r_0} \frac{1}{\sqrt{1 - r^4}} dr. \tag{3.2}$$

La integral (3.2) es impropia cuando $r = 1$, sin embargo esta converge por lo que, $\int_0^1 (1 - r^4)^{-1/2} dr$ es la longitud de arco de la porción de la lemniscata en el primer cuadrante. En el siglo dieciocho, este número fue denotado como $\frac{\varpi}{2}$, donde ϖ es una variante de la letra griega π. Por lo tanto

$$\varpi = 2\int_0^1 \frac{1}{\sqrt{1 - r^4}} dr \approx 2{,}62206.$$

De aquí se sigue que la longitud de arco de la lemniscata es 2ϖ y que la longitud de arco entre dos n-puntos de división consecutivos es $\frac{2\varpi}{n}$.

Escribiremos (3.2) como

$$s = \int_0^r \frac{1}{\sqrt{1 - t^4}} dt, \tag{3.3}$$

donde s representa la longitud de arco sobre la lemniscata desde el origen a un punto en el primer cuadrante con coordenadas polares (r, θ). Entonces (3.3) expresa a s como una función de r. Siguiendo a Abel la función inversa será escrita como $r = \varphi(s)$, por lo tanto

$$r = \varphi(s) \Longleftrightarrow s = \int_0^r \frac{1}{\sqrt{1 - t^4}} dt. \tag{3.4}$$

Como

$$0 \leq r \leq 1 \qquad \text{corresponde a} \qquad 0 \leq s \leq \frac{\varpi}{2},$$

observamos que φ está definida en el intervalo $[0, \frac{\varpi}{2}]$, la llamaremos ***función lemniscática o de Abel***. En la Sección 3.2 extenderemos φ a una función periódica en $\mathbb{R}$, y en la Sección 3.3 extenderemos φ a una función meromorfa doblemente periódica en $\mathbb{C}$.

En particular, cuando $n \geq 4$, el primer n-punto de división de la lemniscata se encuentran en el primer cuadrante. Como su longitud de arco desde el origen es $\frac{2\varpi}{n}$, la Proposición 3.1.1 implica que el primer n-punto de división es construible con regla y compás si y sólo si

$$r_0 = \varphi\left(\frac{2\varpi}{n}\right)$$

es un número construible. En la siguiente sección desarrollaremos fórmulas de multiplicación para $\varphi(ns)$, $n \in \mathbb{Z}$ y usaremos éstas para mostrar que:

- $\varphi(\frac{2\varpi}{n})$ es la raíz de un polinomio con coeficientes en $\mathbb{Z}$.
- $\varphi(\frac{2\varpi}{n})$ es construible si y sólo si todos los n-puntos de división son construibles con regla y compás.

En la Sección 3.5 consideraremos los grupos de Galois de la extensión

$$\mathbb{Q}(i) \subset \mathbb{Q}\left(i, \varphi\left(\frac{2\varpi}{n}\right)\right).$$

La aparición de $i = \sqrt{-1}$ es inesperada pero tendrá un sentido perfecto una vez que hallamos estudiado las fórmulas de la multiplicación compleja para $\varphi((n+im)s)$, $n+im \in \mathbb{Z}[i]$, en la Sección 3.4. Usando esto y algunas astutas ideas de Eisenstein, se podrá probar el teorema de Abel sobre la división de la lemniscata.

Notas Matemáticas

• **Integrales y Funciones Inversas.** La definición de la función de Abel $r = \varphi(s)$ envuelve una integral definida s en términos de r y entonces una función inversa para obtener r en términos de s.

La idea de una función inversa de una integral es más común de lo que se podría esperar. Por ejemplo, una definición estándar de e^x es primero definir el logaritmo natural via la integral

$$\ln(x) = \int_1^x \frac{1}{t} dt, \quad x > 0,$$

y después definir e^x como la función inversa de $\ln(x)$. Por lo tanto e^x es la función inversa de una integral.

Otro ejemplo del cálculo es la integral

$$\operatorname{sen}^{-1}(x) = \int_0^x \frac{1}{\sqrt{1-t^2}} dt, \quad -1 \leq x \leq 1.$$

La función inversa de sen^{-1} es obviamente $\text{sen}(x)$. Por lo tanto $\text{sen}(x)$ es la función inversa de una integral.

Ahora surge una intrigante idea. Supongamos que no conocemos la función $\text{sen}(x)$ ni el $\text{sen}^{-1}(x)$. ¿Como podemos entender la integral $\int(1-x^2)^{-1/2}dx$? Una forma podría ser *definir* $\text{sen}(x)$ como la función inversa de

$$x \mapsto \int_0^x \frac{1}{\sqrt{1-t^2}}dt.$$

Además, si *definimos* $\cos(x) = \frac{d}{dx}\text{sen}(x)$ y $\pi = 2\int_0^1(1-t^2)^{-1/2}dt$, entonces todas las propiedades estándar de $\text{sen}(x)$ y $\cos(x)$ pueden ser derivadas de estas definiciones.

Una manera de entender (3.4) es que la función de Abel $\varphi(s)$ se obtiene aplicando la misma idea a la integral $\int(1-x^4)^{-1/2}dx$. Existen muchas analogías entre $\text{sen}(x)$ y $\varphi(s)$, y es posible desarrollar paralelamente las propiedades de estas funciones, una presentación de esto se puede ver en [1] y [6].

3.2. La función lemniscática

En (3.4) definimos la función de Abel $\varphi(s)$ por

$$r = \varphi(s) \Longleftrightarrow s = \int_0^r \frac{1}{\sqrt{1-t^4}}dt. \tag{3.5}$$

Puesto que s representa la longitud de arco del origen a algún punto de la lemniscata en el primer cuadrante, observamos que $\varphi(s)$ esta definida en $[0, \frac{\varpi}{2}]$, donde

$$\varpi = 2\int_0^1 \frac{1}{\sqrt{1-t^4}}dt.$$

En esta sección, extenderemos $\varpi(s)$ a una función en $\mathbb{R}$ de periodo 2ϖ y mostraremos que la extensión satisface algunas importantes fórmulas de adición y multiplicación. También aplicaremos estas fórmulas a las construcciones con regla y compás sobre la lemniscata.

3.2.1. Una función periódica

Nuestra primera tarea será extender $\varphi(s)$ a una función en $\mathbb{R}$ de periodo 2ϖ. Haremos esto extendiendo la interpretación de longitud de arco $\varphi(s)$ dada en la Sección 3.1.

La parametrización de la lemniscata que obtenemos se su longitud de arco se define mandando los números reales s al punto P en la lemniscata de tal forma que:

- Si $s = 0$, entonces P es el origen.

- Si $s > 0$, entonces nos desplazaremos comenzando por el primer cuadrante y continuaremos recorriendo la curva hasta llegar al punto P de manera que la longitud de arco acumulada desde el origen sea s.
- Si $s < 0$, entonces nos desplazaremos comenzando por el tercer cuadrante y continuaremos recorriendo la curva hasta llegar al punto P para el cual su longitud de arco acumulada desde el origen sea $-s$.

Llamaremos a s la ***variable de la longitud de arco con signo*** de la lemniscata. Cuando $|s|$ es grande, podemos necesitar dar varias vueltas a la lemniscata completa antes de alcanzar el punto P. Puesto que el total de la longitud de arco de la lemniscata es 2ϖ, observamos que s y $s \pm 2\varpi$ nos proporcionan el mismo punto sobre la lemniscata para cualquier $s \in \mathbb{R}$. Esto es parecido a medir ángulos en la circunferencia unitaria, donde s y $s \pm 2\pi$ denotan el mismo punto sobre la circunferencia.

La lemniscata en coordenadas polares se expresa como $r^2 = \cos(2\theta)$. Recordemos que está permitido que r sea negativo así como positivo o cero. Restringiremos θ a $[-\frac{\pi}{4}, \frac{\pi}{4}]$, de tal forma que $0 \leq r \leq 1$ denotará la mitad derecha de la lemniscata y $-1 \leq r \leq 0$ denotará la mitad izquierda. Llamaremos a r la ***distancia polar*** del punto correspondiente sobre la lemniscata. Estrictamente hablando, r es en realidad la ***distancia polar con signo***, puesto que r es negativo en la mitad izquierda de la lemniscata. Usaremos el término más corto "distancia polar" para simplificar.

Ahora consideremos (3.5). Es fácil observar que la longitud de arco con signo s satisface

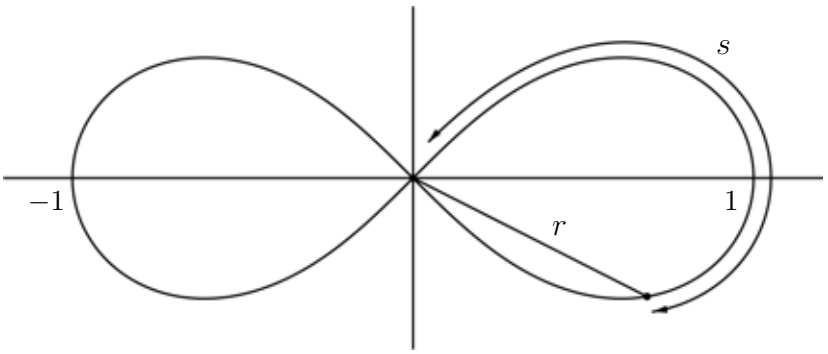
$$s = \int_0^r \frac{1}{\sqrt{1-t^4}} dt$$

para $-\frac{\varpi}{2} \leq s \leq \frac{\varpi}{2}$ y $-1 \leq r \leq 1$. Esto implica que (3.5) puede usarse para definir $\varphi(s)$ para $-\frac{\varpi}{2} \leq s \leq \frac{\varpi}{2}$. En otras palabras, para s en este rango, la función de Abel es simplemente la distancia polar (con la convención anterior sobre r) del punto sobre la lemniscata con longitud de arco con signo s.

Ahora es fácil extender φ a todo $\mathbb{R}$: dada $s \in \mathbb{R}$, $\varphi(s)$ es la distancia polar del punto sobre la lemniscata cuya longitud de arco es s. De esta manera

$$\varphi(s) = r,$$

donde s y r están relacionadas de acuerdo al diagrama



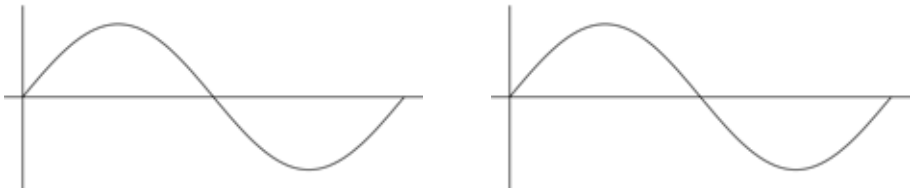
Notemos que $\varphi(s)$ tiene periodo 2ϖ, puesto que s y $s \pm 2\varpi$ proporcionan el mismo punto sobre la lemniscata. Además, $\varphi(s)$ también satisface las siguientes identidades:

$$\begin{aligned} \varphi(-s) &= -\varphi(s), \\ \varphi(\varpi - s) &= \varphi(s). \end{aligned} \tag{3.6}$$

Lo primero se sigue de que s y $-s$ corresponden a puntos sobre la lemniscata simétricos con respecto al origen, y lo segundo se sigue de que s y $\varpi - s$ corresponden a puntos simétricos con respecto al eje x (recordemos que cada mitad de la lemniscata tiene longitud ϖ). Usando la interpretación de la longitud de arco de $\varphi(s)$, podemos probar que $\varphi(s)$ es infinitamente diferenciable para todo $s \in \mathbb{R}$, pero omitiremos esta prueba.

La función $\text{sen}(\frac{\pi}{\varpi}s)$ tiene el mismo periodo y la amplitud que $\varphi(s)$. Si trazamos las gráficas de $\varphi(s)$ y $\text{sen}(\frac{\pi}{\varpi}s)$ para $0 \leq s \leq 2\varpi$, entonces obtenemos:

La función de Abel $\varphi(s)$ La función sen $\left(\frac{\pi}{\omega}s\right)$



La función $\text{sen}(x)$ satisface identidades similares a (3.6), pero la teoría completa de $\text{sen}(x)$ requiere $\text{sen}'(x) = \cos(x)$. Lo mismo es cierto para $\varphi(s)$, donde usaremos $\varphi'(s)$. También necesitaremos la siguiente identidad que es esencial.

Proposición 3.2.1. *Sea $\varphi(s)$ definida como antes. Entonces*

$$\varphi'^2(s) = 1 - \varphi^4(s).$$

Demostración. Primero observemos que por la forma en que definimos $\varphi(s)$ en (3.6) tenemos que

$$\varphi'^2(-s) = \varphi'^2(s) = \varphi'^2(\varpi - s)$$

por lo tanto es suficiente probar que

$$\varphi'(s) = \sqrt{1 - \varphi^4(s)}, \quad 0 \leq s \leq \frac{\varpi}{2},$$

para demostrar que la igualdad se cumple para todo $s \in \mathbb{R}$.

Para demostrar esta ecuación, primero observemos que por (3.5) obtenemos la identidad

$$s = \int_0^{\varphi(s)} \frac{1}{\sqrt{1 - t^4}} dt, \quad 0 \leq s \leq \frac{\varpi}{2}.$$

Derivando ambos lados con respecto a s y utilizando la Regla de la Cadena junto con el Teorema Fundamental del Cálculo, obtenemos que

$$1 = \frac{1}{\sqrt{1 - \varphi^4(s)}} \varphi'(s), \quad 0 \leq s < \frac{\varpi}{2}$$

(observemos que hemos excluido el valor $s = \frac{\varpi}{2}$ ya que en este valor la integral esta indeterminada). De aquí se sigue que

$$\varphi'(s) = \sqrt{1 - \varphi^4(s)}, \quad 0 \leq s < \frac{\varpi}{2}.$$

Para $s = \frac{\varpi}{2}$, notemos que $\sqrt{1 - \varphi^4(s)}$ se anula en $\frac{\varpi}{2}$ debido a que $\varphi(\frac{\varpi}{2}) = 1$. Por la forma en la que hemos definido a $\varphi(s)$ sabemos que 1 es un valor máximo para esta función por lo que $\varphi'(\frac{\varpi}{2}) = 0$. Por lo tanto la ecuación se cumple para todo $0 \leq s \leq \frac{\varpi}{2}$. □

Ahora que probamos la identidad podemos junto con (3.6) obtener algunas otras propiedades de $\varphi'(s)$:

- $\varphi'(\varpi - s) = -\varphi'(s)$.
- Puesto que $\varphi(s)$ es una función impar tenemos que

  $$\varphi'(-s) = \sqrt{1 - \varphi^4(-s)} = \sqrt{1 - \varphi^4(s)} = \varphi'(s),$$

  por lo que $\varphi'(s)$ es una función par.
- Puesto que $\varphi(s) = \varphi(s + n \cdot 2\varpi)$ se cumple para todo $n \in \mathbb{Z}$ tenemos que

  $$\varphi'(s + n \cdot 2\varpi) = \sqrt{1 - \varphi^4(s + n \cdot 2\varpi)} = \sqrt{1 - \varphi^4(s)} = \varphi'(s),$$

  por lo que $\varphi'(s)$ también tiene periodo 2ϖ.
- Derivando la ecuación de la proposición anterior tenemos que

  $$\varphi''(s) = -2\varphi^3(s).$$

3.2.2. Leyes de adición

La ley de adición para el $\operatorname{sen}(x)$ afirma que

$$\operatorname{sen}(x + y) = \operatorname{sen}(x)\cos(y) + \cos(x)\operatorname{sen}(y).$$

Para $\varphi(x)$, su ley de adición se remonta a Euler, quien en 1753 probó la identidad

$$\int_0^\alpha \frac{1}{\sqrt{1 - t^4}} dt + \int_0^\beta \frac{1}{\sqrt{1 - t^4}} dt = \int_0^\gamma \frac{1}{\sqrt{1 - t^4}} dt \tag{3.7}$$

$$\text{cuando} \quad \alpha, \beta \in [0, 1] \quad \text{y} \quad \gamma = \frac{\alpha\sqrt{1 - \beta^4} + \beta\sqrt{1 - \alpha^4}}{1 + \alpha^2\beta^2}.$$

Para afirmar esto en términos de φ, representemos como x, y y z a las tres integrales en (3.7), de tal forma que $\varphi(x) = \alpha$, $\varphi(y) = \beta$, y $\varphi(z) = \gamma$. Entonces $x + y = z$ implica que

$$\varphi(x+y) = \varphi(z) = \gamma = \frac{\alpha\sqrt{1-\beta^4} + \beta\sqrt{1-\alpha^4}}{1+\alpha^2\beta^2},$$

utilizando este hecho en combinación con $\varphi(x) = \alpha$ y $\varphi(y) = \beta$ obtenemos

$$\varphi(x+y) = \frac{\varphi(x)\sqrt{1-\varphi^4(y)} + \varphi(y)\sqrt{1-\varphi^4(x)}}{1+\varphi^2(x)\varphi^2(y)}. \tag{3.8}$$

Además, por la Proposición 3.2.1 $\sqrt{1-\varphi^4(x)} = \varphi'(x)$ para $0 \leq x \leq \frac{\varpi}{2}$. Por esta razón

$$\varphi(x+y) = \frac{\varphi(x)\varphi'(y) + \varphi(y)\varphi'(x)}{1+\varphi^2(x)\varphi^2(y)}. \tag{3.9}$$

En vez de ocupar el resultado de Euler, Abel proporcionó una prueba diferente de (3.9) que se cumple para todo x, $y \in \mathbb{R}$. Para esto Abel definió una función $g(x, y)$ diferenciable en $\mathbb{R}^2$ y definió la función $h(u, v)$ de la forma: $h(u, v) = g(\frac{1}{2}(u+v), \frac{1}{2}(u-v))$, derivando esta función tenemos con la ayuda de la regla de la cadena que

$$\frac{\partial h}{\partial v}(u,v) = \frac{1}{2}\cdot\frac{\partial g}{\partial x}\left(\frac{1}{2}(u+v), \frac{1}{2}(u-v)\right) - \frac{1}{2}\cdot\frac{\partial g}{\partial y}\left(\frac{1}{2}(u+v), \frac{1}{2}(u-v)\right).$$

Utilizando esta ecuación podemos mostrar que $g(x, y) = g(x+y, 0)$ en $\mathbb{R}^2$ si y sólo si $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y}$ en $\mathbb{R}^2$.

Primero supongamos que $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y}$, sustituyendo este hecho en la parcial de h que obtuvimos tenemos que

$$\frac{\partial h}{\partial v}(u,v) = 0, \quad \text{para todo} \quad (u,v) \in \mathbb{R}^2,$$

por lo que observamos que $h(u, v) = h(u, 0)$ para toda $u, v \in \mathbb{R}$. De aquí que si tomamos $x + y = \frac{1}{2}(u + v)$ y a $y = \frac{1}{2}(u - v)$ obtenemos

$$g(x,y) = h(x+y, x-y) = h(x+y, 0) = h(x+y, x+y) = g(x+y, 0).$$

Finalmente Abel tomó la función $g(x, y) = \frac{\varphi(x)\varphi'(y)+\varphi(y)\varphi'(x)}{1+\varphi^2(x)\varphi^2(y)}$ y puesto que esta función cumple la propiedad de que $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y}$ aplicando el criterio anterior tenemos ***la ley de adición***:

$$\varphi(x+y) = g(x+y, 0) = g(x, y) = \frac{\varphi(x)\varphi'(y) + \varphi(y)\varphi'(x)}{1+\varphi^2(x)\varphi^2(y)}.$$

Usando $\varphi(-x) = -\varphi(x)$ podemos mostrar que (3.9) implica la ***ley de sustracción***

$$\varphi(x-y) = \frac{\varphi(x)\varphi'(y) - \varphi(y)\varphi'(x)}{1+\varphi^2(x)\varphi^2(y)}.$$

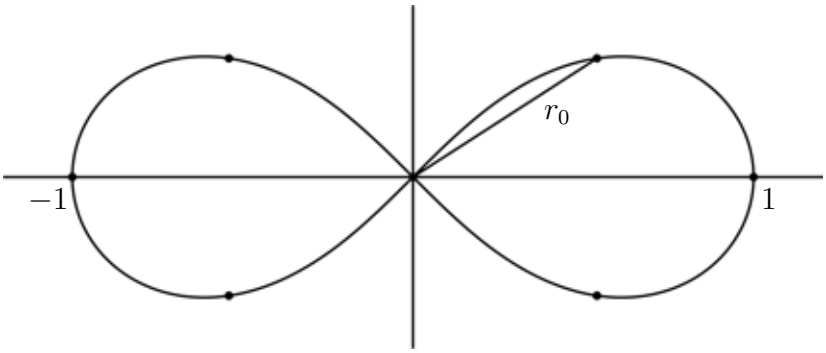
Cuando combinamos esto con (3.9), podemos fácilmente obtener la identidad

$$\varphi(x+y)+\varphi(x-y)=\frac{2\varphi(x)\varphi'(y)}{1+\varphi^2(x)\varphi^2(y)}. \tag{3.10}$$

Esto será de utilidad en la siguiente sección.

La leyes de adición proporcionan algunas construcciones con regla y compás.

Ejemplo 3.2.2. Dividamos la lemniscata en ocho partes de longitud $\frac{2\varpi}{8}=\frac{\varpi}{4}$. En la figura se muestran $r_0=\varphi(\frac{\varpi}{4})$ y los 8-puntos de división:



La Proposición 3.1.1 muestra que para construir los 8-puntos de división, solamente es necesario construir r_0. Puesto que $\varphi(\frac{\varpi}{2})=1$, la ley de adición (3.8) implica que

$$1=\varphi\left(\frac{\varpi}{2}\right)=\varphi\left(\frac{\varpi}{4}+\frac{\varpi}{4}\right)=\frac{2\varphi(\frac{\varpi}{4})\sqrt{1-\varphi^4(\frac{\varpi}{4})}}{1+\varphi^4(\frac{\varpi}{4})}=\frac{2r_0\sqrt{1-r_0^4}}{1+r_0^4}.$$

Para resolver esta ecuación necesitamos obtener las raíces del polinomio

$$r_0^8+4r_0^6+2r_0^4-4r_0^2+1=0,$$

para obtener estas raíces sea $x=r_0^2$, entonces tenemos que

$$x^4+4x^3+2x^2-4x+1=(x^2+2x-1)^2,$$

de ahí que la única raíz positiva de este polinomio sea $x=\sqrt{2}-1$, por lo tanto la única solución real positiva del primer polinomio será

$$r_0=\sqrt{\sqrt{2}-1}.$$

Pero este número es construible y por lo tanto proporciona la construcción deseada.

El razonamiento detrás del Ejemplo 3.2.2 puede ser generalizado de la siguiente manera.

Proposición 3.2.3. *Si $\varphi(x_0)$ es construible, entonces $\varphi(\frac{x_0}{2})$ también lo es.*

Demostración. Tomando $x = y$ en (3.9) obtenemos la ***fórmula de duplicación***

$$\varphi(2x) = \frac{2\varphi(x)\varphi'(x)}{1+\varphi^4(x)}. \tag{3.11}$$

Sean $r_0 = \varphi(\frac{x_0}{2})$ y $a = \varphi(x_0)$. Entonces (3.11) y $\varphi'^2(\frac{x_0}{2}) = 1 - \varphi^4(\frac{x_0}{2})$ implican que

$$a^2 = \left(\frac{2r_0\varphi'(\frac{x_0}{2})}{1+r_0^4}\right)^2 = \frac{4r_0^2(1-r_0^4)}{(1+r_0^4)^2}.$$

Para resolver esta ecuación para r_0, sea $t \in \mathbb{C}$ tal que satisface

$$t^2 = \frac{2ir_0^2}{1-r_0^4}, \quad i = \sqrt{-1} \tag{3.12}$$

y observemos que

$$\frac{-2it^2}{1-t^4} = \frac{-2i\frac{2ir_0^2}{1-r_0^4}}{1-\left(\frac{2ir_0^2}{1-r_0^4}\right)^2} = \frac{4r_0^2(1-r_0^4)}{(1+r_0^4)^2} = a^2. \tag{3.13}$$

Resolviendo (3.13) para t^2 y (3.12) para r_0 por la fórmula cuadrática tenemos que

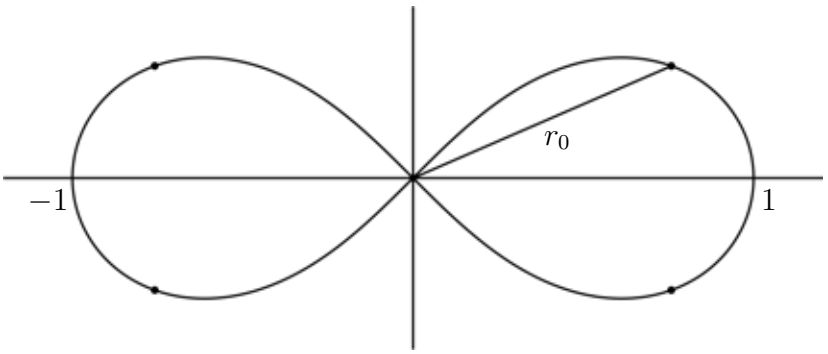
$$t^2 = \frac{i \pm \sqrt{a^4-1}}{a^2} \quad \text{y que} \quad r_0 = \sqrt{\frac{-i \pm \sqrt{t^4-1}}{t^2}},$$

entonces t^2 es construible por que a es construible por lo tanto r_0 es construible. □

Las fórmulas (3.12) y (3.13) en la prueba anterior al parecer aparecen de la nada. En la Sección 3.4 utilizaremos la ***multiplicación compleja*** y la factorización $2 = (1+i)(1-i)$ para dar la fórmula de duplicación para $\varphi(2x)$ en (3.12) y (3.13). Entonces estas fórmulas tendrán más sentido.

Veamos un ejemplo un poco más complicado.

Ejemplo 3.2.4. Dividiendo la lemniscata en seis partes de la misma longitud obtenemos la siguiente figura:



Para calcular $r_0 = \varphi(\frac{2\varpi}{6}) = \varphi(\frac{\varpi}{3})$, primero observemos que si tomamos (3.10) con $2x$ y x en lugar de x y y obtenemos

$$\varphi(3x) + \varphi(x) = \varphi(2x + x) + \varphi(2x - x) = \frac{2\varphi(2x)\varphi'(x)}{1 + \varphi^2(2x)\varphi^2(x)}$$

Utilizando (3.11) y la igualdad $\varphi'^2(x) = 1 - \varphi^4(x)$ junto con un poco de álgebra obtenemos la fórmula de triplicación

$$\varphi(3x) = -\varphi(x)\frac{\varphi^8(x) + 6\varphi^4(x) - 3}{1 + 6\varphi^4(x) - 3\varphi^8(x)}. \tag{3.14}$$

Ya que $\varphi(\varpi) = 0$ sustituyendo $x = \frac{\varpi}{3}$ en (3.14) obtenemos que $r_0 = \varphi(\frac{\varpi}{3})$ satisface $r_0^8 + 6r_0^4 - 3 = 0$, fácilmente podemos ver que para esta ecuación la única solución real positiva es $r_0 = \sqrt[4]{2\sqrt{3} - 3}$. Este valor claramente es construible y por lo tanto obtenemos la construcción con regla y compás deseada.

3.2.3. Multiplicación por enteros

Las fórmulas de duplicación y triplicación

$$\begin{aligned} \varphi(2x) &= \frac{2\varphi(x)\varphi'(x)}{1 + \varphi^4(x)}, \\ \varphi(3x) &= -\varphi(x)\frac{\varphi^8(x) + 6\varphi^4(x) - 3}{1 + 6\varphi^4(x) - 3\varphi^8(x)} \end{aligned}$$

de (3.11) y (3.14) pueden ser generalizadas a fórmulas para expresar $\varphi(nx)$ en términos de $\varphi(x)$ y $\varphi'(x)$ para cualquier entero positivo n.

Teorema 3.2.5. *Dado un entero $n > 0$, existen polinomios primos relativos $P_n(u)$, $Q_n(u) \in \mathbb{Z}[u]$ tales que si n es impar, entonces*

$$\varphi(nx) = \varphi(x)\frac{P_n\left(\varphi^4(x)\right)}{Q_n\left(\varphi^4(x)\right)},$$

y si n es par, entonces

$$\varphi(nx) = \varphi(x)\frac{P_n\left(\varphi^4(x)\right)}{Q_n\left(\varphi^4(x)\right)}\varphi'(x).$$

Además, $Q_n(0) = 1$.

Demostración. Probaremos el teorema por inducción sobre n. Definiendo $P_1(u) = Q_1(u) = 1$ obtenemos la fórmula deseada para $n = 1$. Para $n = 2$, notemos que (3.11) se puede reescribir como

$$\varphi(2x) = \varphi(x)\frac{2}{1 + \varphi^4(x)}\varphi'(x).$$

Por lo tanto el teorema se cumple para $n = 2$ con $P_2(u) = 2$, $Q_2(u) = 1 + u$. Ahora supongamos que se cumple para $n-1$ y para n. Usando (3.10) con nx y x en lugar de x y y, obtenemos

$$\varphi\left((n+1)x\right) = -\varphi\left((n-1)x\right) + \frac{2\varphi(nx)\varphi'(x)}{1+\varphi^2(nx)\varphi^2(x)}.$$

Si n es par, entonces $n-1$ es impar, entonces tenemos que nuestra hipótesis de inducción implica que

$$\varphi\left((n+1)x\right) = -\left(\varphi(x)\frac{P_{n-1}\left(\varphi^4(x)\right)}{Q_{n-1}\left(\varphi^4(x)\right)}\right) + \frac{2\left(\varphi(x)\frac{P_n\left(\varphi^4(x)\right)}{Q_n(\varphi^4(x))}\varphi'(x)\right)\varphi'(x)}{1+\left(\varphi(x)\frac{P_n(\varphi^4(x))}{Q_n(\varphi^4(x))}\varphi'(x)\right)^2\varphi^2(x)}.$$

Utilizando $\varphi'^2(x) = 1-\varphi^4(x)$ y simplificando obtenemos

$$\varphi\left((n+1)x\right) = \varphi(x)\frac{P_{n+1}(\varphi^4(x))}{Q_{n+1}(\varphi^4(x))},$$

donde

$$P_{n+1}(u) = 2P_n(u)Q_n(u)Q_{n-1}(u)(1-u) - P_{n-1}(u)[Q_n^2(u) + uP_n^2(u)(1-u)] \tag{3.15}$$

y

$$Q_{n+1}(u) = Q_{n-1}(u)[Q_n^2(u) + uP_n^2(u)(1-u)].$$

Para el caso cuando n es impar, tenemos que $n-1$ es par, entonces nuestra hipótesis de inducción implica que

$$\varphi\left((n+1)x\right) = -\left(\varphi(x)\frac{P_{n-1}\left(\varphi^4(x)\right)}{Q_{n-1}\left(\varphi^4(x)\right)}\varphi'(x)\right) + \frac{2\left(\varphi(x)\frac{P_n\left(\varphi^4(x)\right)}{Q_n(\varphi^4(x))}\right)\varphi'(x)}{1+\left(\varphi(x)\frac{P_n(\varphi^4(x))}{Q_n(\varphi^4(x))}\right)^2\varphi^2(x)}.$$

De donde eliminando los denominadores y factorizando obtenemos que

$$\varphi\left((n+1)x\right) = \varphi(x)\frac{P_{n+1}(\varphi^4(x))}{Q_{n+1}(\varphi^4(x))}\varphi'(x),$$

donde

$$P_{n+1}(u) = 2P_n(u)Q_n(u)Q_{n-1}(u) - P_{n-1}(u)[Q_n^2(u) + uP_n^2(u)]$$

y

$$Q_{n+1}(u) = Q_{n-1}(u)[Q_n^2(u) + uP_n^2(u)]$$

De aquí se sigue por nuestra hipótesis de inducción que en ambos casos $P_{n+1}(u)$, $Q_{n+1}(u) \in \mathbb{Z}[u]$, y puesto que $\mathbb{Z}[u]$ es un dominio de factorización única entonces podemos escribir $P_{n+1}(u) = C_{n+1}\widetilde{P}_{n+1}(u)$ y $Q_{n+1}(u) = C_{n+1}\widetilde{Q}_{n+1}(u)$, donde C_{n+1}, $\widetilde{P}_{n+1}$, $\widetilde{Q}_{n+1} \in \mathbb{Z}[u]$ y $\widetilde{P}_{n+1}$ y $\widetilde{Q}_{n+1}$ son polinomios primos relativos. Sustituyendo estos polinomios en las fórmulas

obtenidas para $\varphi((n+1)x)$ tenemos que $\widetilde{P}_{n+1}$ y $\widetilde{Q}_{n+1}$ son polinomios que cumplen todas las condiciones del teorema. Finalmente falta demostrar que $\widetilde{Q}_n(0) = 1$, para esto observemos que en ambos casos por las fórmulas recursivas para $Q_{n+1}(u)$ tenemos que

$$Q_{n+1}(0) = Q_{n-1}(0)Q_n^2(0),$$

recordemos que $Q_1(u) = 1$ y que $Q_2(u) = 1 + u$, entonces $Q_1(0) = 1$ y $Q_2(0) = 1$, ahora tomando estos valores como hipótesis de inducción y aplicando inducción sobre n para la fórmula recursiva tenemos que $Q_{n+1}(0) = 1$. Puesto que $Q_{n+1}(0) = C_{n+1}(0)\widetilde{Q}_{n+1}(0) = 1$ y que $C_{n+1}(0)$, $\widetilde{Q}_{n+1}(0) \in \mathbb{Z}$ tenemos que $Q_{n+1}(0)$ puede tomar el valor 1 ó -1, en el caso en el cual toma el valor -1 podemos multiplicar los polinomios $\widetilde{P}_{n+1}$ y $\widetilde{Q}_{n+1}$ por -1 sin alterar las fórmulas y seguir afirmando que $\widetilde{Q}_{n+1}(0) = 1$. □

El Teorema 3.2.5 tiene algunas importantes consecuencias en lo concerniente a los puntos de división de la lemniscata. Las distancias polares de los n-puntos de división son

$$\varphi\left(m\frac{2\varpi}{n}\right), \quad m = 0, 1, \ldots, n-1.$$

Cuando n es impar, la periodicidad de φ y el Teorema 3.2.5 implican que

$$0 = \varphi(m \cdot 2\varpi) = \varphi\left(n \cdot m\frac{2\varpi}{n}\right) = \varphi\left(m\frac{2\varpi}{n}\right)\frac{P_n\left(\varphi^4\left(m\frac{2\varpi}{n}\right)\right)}{Q_n\left(\varphi^4\left(m\frac{2\varpi}{n}\right)\right)},$$

por lo tanto la distancia polar $\varphi\left(m\frac{2\varpi}{n}\right)$ es una raíz de $uP_n(u^4)$ cuando n es impar. En el caso cuando n es par, por la periodicidad de φ y el Teorema 3.2.5 tenemos que

$$0 = \varphi(m \cdot 2\varpi) = \varphi\left(n \cdot m\frac{2\varpi}{n}\right) = \varphi\left(m\frac{2\varpi}{n}\right)\frac{P_n\left(\varphi^4\left(m\frac{2\varpi}{n}\right)\right)}{Q_n\left(\varphi^4\left(m\frac{2\varpi}{n}\right)\right)}\varphi'\left(m\frac{2\varpi}{n}\right),$$

y recordando que $\varphi'(x) = \sqrt{1-\varphi^4(x)}$ podemos observar que

$$uP_n(u^4)\sqrt{1-u^4} = 0 \Leftrightarrow uP_n(u^4)(1-u^2) = 0,$$

por lo tanto las distancias polares serán las raíces del polinomio $uP_n(u^4)(1-u^2)$. Llamaremos a estos polinomios los ***n-polinomios de división***. Hemos probado entonces el siguiente corolario del Teorema 3.2.5.

Corolario 3.2.6. *Sea $n \in \mathbb{Z}^+$, entonces las distancias polares de los n-puntos de división de la lemniscata son las raíces de los n-polinomios de división.* □

También tenemos el siguiente resultado referente a las construcciones con regla y compás.

Corolario 3.2.7. *Sea $n \in \mathbb{Z}^+$ tal que $\varphi\left(\frac{2\varpi}{n}\right)$ es construible, entonces*

(a) *$\varphi\left(m\frac{2\varpi}{n}\right)$ es construible para cada $m \in \mathbb{Z}$.*

(*b*) *Los n-puntos de división de la lemniscata son construibles con regla y compás.*

(*c*) *Si además* $\varphi\left(\frac{2\varpi}{m}\right)$ *es construible para un entero positivo* m*, entonces también lo es* $\varphi\left(\frac{2\varpi}{N}\right)$*, donde* $N = \text{mcm}(n, m)$.

Demostración. Si $\varphi\left(\frac{2\varpi}{n}\right)$ es construible, entonces $\varphi'\left(\frac{2\varpi}{n}\right)$ también lo es puesto que $\varphi'(x) = \pm\sqrt{1-\varphi^4(x)}$. La parte (a) es obvia para $n = 1$ y 2, por lo que supondremos que $n > 2$.

Cuando $m > 0$, el Teorema 3.2.5 implica que $\varphi(m\frac{2\varpi}{n})$ es una función racional en términos de $\varphi(\frac{2\varpi}{n})$ y $\varphi'(\frac{2\varpi}{n})$ con coeficientes en $\mathbb{Z}$. Para demostrar la existencia de este valor probaremos que el denominador $Q_m(u^4)$ de esta función racional no se anula cuando $u = \varphi(\frac{2\varpi}{n})$, para esto utilizaremos que $n > 2$ y que del Teorema 3.2.5 tenemos que los polinomios $P_m(u)$ y $Q_m(u)$ son primos relativos.

Primero observemos que si tomamos polinomios primos relativos $P(u), Q(u) \in \mathbb{Z}[u]$ tales que $Q(0) = 1$ podemos mostrar que los polinomios $uP(u^4)$ y $uP(u^4)(1-u^2)$ no comparten raíces con $Q(u^4)$ en ninguna extensión de $\mathbb{Q}$. Como $u = 0$ es una raíz del polinomio $uP(u^4)$ pero $Q(0) = 1$ tenemos que $u = 0$ no es raíz común, ahora si suponemos que u_0 es una raíz común de estos polinomios tendríamos que

$$u - u_0 \,\big|\, P(u^4) \quad \text{y que} \quad u - u_0 \big|\, Q(u^4),$$

pero esto es imposible dado que estos polinomios son primos relativos, por lo tanto $uP(u^4)$ y $Q(u^4)$ no tienen raíces en común, tomando este hecho, pero ahora tomando en cuenta los polinomios $uP(u^4)(1-u^2)$ y $Q(u^4)$ tenemos que su única raíz común posible es $u^2 = 1$, pero tomando las fórmulas recursivas del Teorema 3.2.5 podemos observar que 1 no es raíz de $Q(u^4)$, por lo tanto estos polinomios no comparten raíces.

Ahora fijemos $x \in \mathbb{R}$ y $m > 0$ impar en $\mathbb{Z}$ y sean $P_m(u), Q_m(u) \in \mathbb{Z}[u]$ como en el Teorema 3.2.5. Así $\varphi(mx)Q_m(\varphi^4(x)) = \varphi(x)P_m(\varphi^4(x))$. Si suponemos que $Q_m(\varphi^4(x)) = 0$, entonces por la igualdad anterior tenemos que $\varphi(x)P_m(\varphi^4(x)) = 0$, pero puesto que los polinomios P_m y Q_m son primos relativos, tenemos que $\varphi(x) = 0$, lo cual implica que $Q_m(\varphi^4(x)) \neq 0$ cuando $\varphi(x) \neq 0$. El caso $m > 0$ par es similar.

Finalmente como tenemos que $\varphi(\frac{2\varpi}{n}) \neq 0$ para todo $n > 2$ en $\mathbb{Z}$ por lo anterior tenemos que $Q_m(\varphi^4(\frac{2\varpi}{n})) \neq 0$. Por lo tanto el denominador no se anula para estos valores por lo que $\varphi(m\frac{2\varpi}{n})$ es un numero construible, puesto que los números construibles forman un subcampo de $\mathbb{C}$.

El caso $m = 0$ es obvio, y $m < 0$ se sigue de $m > 0$ puesto que φ es una función impar. Esto completa la prueba de (a)

La parte (b) se sigue inmediatamente de la parte (a) y la Proposición 3.1.1.

Para la parte (c), sea $d = \text{mcd}(n, m)$, Entonces $N = \text{mcm}(n, m) = \frac{nm}{d}$. De aquí se sigue que si los enteros μ y ν satisfacen $\mu n + \nu m = d$, entonces

$$\mu\frac{2\varpi}{m} + \nu\frac{2\varpi}{n} = (\mu n + \nu m)\frac{2\varpi}{nm} = d\frac{2\varpi}{nm} = \frac{2\varpi}{N}.$$

Por la parte (a), $\varphi\left(\mu\frac{2\varpi}{m}\right)$ y $\varphi\left(\nu\frac{2\varpi}{n}\right)$ son construibles, y lo mismo que $\varphi'\left(\mu\frac{2\varpi}{m}\right)$ y $\varphi'\left(\nu\frac{2\varpi}{n}\right)$. Entonces la ley de adición (3.9) define a $\varphi\left(\frac{2\varpi}{N}\right)$ como una expresión racional con coeficientes en $\mathbb{Z}$ en los números construibles dados por los valores de φ y φ' en $\mu\frac{2\varpi}{m}$ y $\nu\frac{2\varpi}{n}$. Como el denominador de esta expresión racional es

$$1+\varphi^2\left(\mu\frac{2\varpi}{m}\right)\varphi^2\left(\nu\frac{2\varpi}{n}\right)\neq 0,$$

se sigue que $\varphi\left(\frac{2\varpi}{N}\right)$ es construible. □

Las partes (b) y (c) del Corolario 3.2.7 implican que si los n-puntos de división y los m-puntos de división de la lemniscata son construibles con regla y compás, entonces también son construibles los N-puntos de división para $N=\text{mcm}(n,m)$. Este hecho será útil más adelante.

Ahora mostremos algunas aplicaciones del Corolario 3.2.7.

Ejemplo 3.2.8. Como $\varphi(2\varpi)=0$, la Proposición 3.2.3 implica que $\varphi\left(\frac{2\varpi}{2^n}\right)$ es construible para $n\geq 0$, entonces la parte (b) del Corolario 3.2.7 muestra que los 2^n-puntos de división de la lemniscata se pueden construir con regla y compás.

Ejemplo 3.2.9. Cuando $n=5$, podemos mostrar que

$$\begin{aligned}&\varphi(5x)=\varphi(x)\frac{P_5(\varphi^4(x))}{Q_5(\varphi^4(x))},\quad \text{donde}\\ &P_5(u)=u^6+50u^5-125u^4+300u^3-105u^2-62u+5,\\ &Q_5(u)=1+50u-125u^2+300u^3-105u^4-62u^5+5u^6.\end{aligned}\tag{3.16}$$

Notemos la "simetría inversa" de los coeficientes de $P_5(u)$ y $Q_5(u)$. Para los 5-puntos de división de la lemniscata, la discusión que precede al Corolario 3.2.6 implica que $r_0=\varphi\left(\frac{2\varpi}{5}\right)$ es una raíz del 5-polinomio de división $uP_5(u^4)$. Así que tenemos que obtener las raíces del polinomio

$$0=r_0P_5(r_0^4)=r_0(r_0^{24}+50r_0^{20}-125r_0^{16}+300r_0^{12}-105r_0^8-62r_0^4+5).$$

esta ecuación puede ser factorizada como

$$0=r_0(r_0^8-2r_0^4+5)(r_0^{16}+52r_0^{12}-26r_0^8-12r_0^4+1)$$

y las únicas soluciones reales positivas son

$$\sqrt[4]{-13+6\sqrt{5}\pm 2\sqrt{85-38\sqrt{5}}},\text{que son construibles.}$$

Por lo tanto $\varphi\left(\frac{2\varpi}{5}\right)$ es construible, esto junto con el Corolario 3.2.7 muestra que los 5-puntos de división de la lemniscata son construibles con regla y compás.

Esta discusión aclara la relación de los n-puntos de división de la lemniscata con la fórmula de multiplicación para $\varphi(nx)$. Pero para utilizar todo el poder de estas fórmulas, necesitaremos extender φ a una función de variable compleja. Haremos esto en la siguiente sección.

3.3. La función lemniscática compleja

El Corolario 3.2.6 implica que las distancias polares $r = \varphi\left(m\frac{2\varpi}{n}\right)$ de los n-puntos de división de la lemniscata son raíces de los n-polinomios de división. Para probar el teorema de Abel sobre la lemniscata, necesitamos representar *todas* las raíces de estos polinomios usando φ. Puesto que muchas de estas raíces son complejas se requiere que sigamos a Gauss y a Abel y extendamos φ a una función definida en $\mathbb{C}$.

Abel comenzó considerando $\varphi(iy)$ para $y \in \mathbb{R}$. Nosotros sabemos que $r = \varphi(y)$ es la función inversa de $y = \int_0^r (1-t^4)^{-1/2}dt$. El cambio de variable $t = ui$ muestra que

$$\int_0^{ir} \frac{1}{\sqrt{1-t^4}}dt = i\int_0^r \frac{1}{\sqrt{1-u^4}}du = iy.$$

Esto sugiere definir $\varphi(iy)$ como $\varphi(iy) = ir = i\varphi(y)$. Entonces Abel usó la ley de adición para definir $\varphi(x+iy)$ como

$$\varphi(x+iy) = \frac{\varphi(x)\varphi'(iy) + \varphi(iy)\varphi'(x)}{1+\varphi^2(x)\varphi^2(iy)}.$$

Como de $\varphi(iy) = i\varphi(y)$ se implica fácilmente que $\varphi'(iy) = \varphi'(y)$, entonces la fórmula para $\varphi(x+iy)$ se simplifica a

$$\varphi(z) = \varphi(x+iy) = \frac{\varphi(x)\varphi'(y) + i\varphi(y)\varphi'(x)}{1-\varphi^2(x)\varphi^2(y)}. \tag{3.17}$$

Para hacer la extensión de Abel rigurosa, debemos ***definir*** $\varphi(z)$ usando (3.17). Sobre $\mathbb{R}$, $\varphi(x)$ es periódica y definida en todas partes; sobre $\mathbb{C}$, debemos ver que $\varphi(z)$ es doblemente periódica y tiene polos. Las propiedades de $\varphi(z)$ jugaran un papel crucial en las Secciones 3.4 y 3.5.

3.3.1. Una función doblemente periódica

Como antes, definamos $\varphi(z) = \varphi(x+iy)$ usando la ecuación (3.17) y veamos algunas propiedades básicas de esta función.

Proposición 3.3.1. *La función $\varphi(z)$ satisface lo siguiente:*

(a) *$\varphi(z)$ es analítica para toda $z \neq (m+in)\frac{\varpi}{2}$, donde m, n son impares.*

(b) *La ley de adición*

$$\varphi(z+w) = \frac{\varphi(z)\varphi'(w) + \varphi(w)\varphi'(z)}{1+\varphi^2(z)\varphi^2(w)}$$

se cumple para todos z, $w \in \mathbb{C}$ tales que ambos lados estén definidos.

(c) *Para $z \in \mathbb{C}$ y m, $n \in \mathbb{Z}$, tenemos que*

$$\varphi(z + m\varpi + n\varpi i) = (-1)^{m+n}\varphi(z).$$

Demostración. Primero observemos que $\varphi(z)$ esta definida siempre y cuando el denominador $1-\varphi^2(x)\varphi^2(y)$ en (3.17) no se anule. La interpretación de $\varphi(x)$ como distancia polar muestra que $\varphi^2(x) \leq 1$ para toda $x \in \mathbb{R}$, donde la igualad se cumple si y sólo si x es un múltiplo impar de $\frac{\varpi}{2}$. De ahí que $\varphi(z)$ esta definida en el conjunto abierto $G = \{z \in \mathbb{C} \,|\, z \neq (m+in)\frac{\varpi}{2},$ donde m, $n \in \mathbb{Z}$ y son impares$\}$.

Escribiendo $\varphi(z) = \varphi(x+iy) = u(x,y) + iv(x,y)$, donde $u(x,y)$ y $v(x,y)$ son las partes real e imaginaria del lado derecho de (3.17). Es fácil ver que $u(x,y)$ y $v(x,y)$ como funciones de (x,y) son diferenciables sobre G, puesto que $\varphi(x)$ es infinitamente diferenciable en $\mathbb{R}$. Además, usando la identidad $\varphi'^2 = 1 - \varphi^4(x)$ para $x \in \mathbb{R}$, es sencillo verificar que $u(x,y)$ y $v(x,y)$ cumplen que

$$\frac{\partial u}{\partial x} = \frac{\varphi'(x)\varphi'(y)[1+\varphi^2(x)\varphi^2(y)]}{[1-\varphi^2(x)\varphi^2(y)]^2} = \frac{\partial v}{\partial y}$$

y

$$\frac{\partial u}{\partial y} = \frac{2\varphi(x)\varphi(y)[\varphi^2(x)-\varphi^2(y)]}{[1-\varphi^2(x)\varphi^2(y)]^2} = -\frac{\partial v}{\partial x},$$

es decir, cumplen las ecuaciones de Cauchy-Riemann, por esta razón $\varphi(z)$ es analítica en G.

Para la parte (b), sean z y w variables complejas, y definamos

$$g(z,w) = \frac{\varphi(z)\varphi'(w) + \varphi(w)\varphi'(z)}{1+\varphi^2(z)\varphi^2(w)}.$$

Cuando $x_0 \in \mathbb{R}$ está fijo, $\varphi(x_0+w)$ y $g(x_0,w)$ son analíticas en $w \in G$ y cuando $w \in \mathbb{R}$ coinciden, entonces por el Teorema de identidad[1], $\varphi(x_0+w) = g(x_0,w)$ para toda $w \in G$. De aquí se sigue que cuando $w_0 \in \mathbb{C}$ esta fijo, $\varphi(z+w_0) = g(z,w_0)$ son analíticas en z y coinciden cuando $z \in \mathbb{R}$. Usando nuevamente el Teorema de identidad, tendremos que $\varphi(z+w_0) = g(z,w_0)$ para toda $z \in \mathbb{C}$ tal que ambas funciones están definidas. Puesto que w_0 es arbitrario, esto prueba la ley de adición.

Para probar la parte (c) necesitaremos una serie de propiedades de $\varphi(z)$ y $\varphi'(z)$. Comenzaremos con la siguiente tabla de valores:

(3.18)

x	$\varphi(x)$	$\varphi'(x) = \sqrt{1-\varphi^4(x)}$
0	0	1
$\frac{\varpi}{2}$	1	0
ϖ	0	-1
$\frac{3\varpi}{2}$	-1	0

También necesitaremos las siguiente identidades válidas para $z \in \mathbb{C}$:

(3.19)

$$\varphi(iz) = i\varphi(z),$$
$$\varphi'(iz) = \varphi'(z).$$

[1] **Teorema de identidad.**
Si f y g son funciones analíticas en una región $G \in \mathbb{C}$ y el conjunto donde coinciden tiene un punto de acumulación en G entonces coinciden en todo G.

Un poco antes, estas identidades fueron probadas para $x \in \mathbb{R}$, por lo que nuevamente por el Teorema de identidad tenemos que estas identidades se cumplen para $z \in \mathbb{C}$.

Puesto que φ y φ' tienen periodo 2ϖ en $\mathbb{R}$, (3.18) y (3.19) fácilmente implican que

$$\begin{aligned} \varphi(m\varpi) &= \varphi(m\varpi i) = 0, \\ \varphi'(m\varpi) &= \varphi'(m\varpi i) = (-1)^m \end{aligned} \tag{3.20}$$

para toda $m \in \mathbb{Z}$. Usando la ley de adición junto con las identidades anteriores es sencillo mostrar que:

$$\varphi(z + m\varpi) = (-1)^m \varphi(z) \quad \text{y} \quad \varphi(z + n\varpi i) = (-1)^n \varphi(z) \tag{3.21}$$

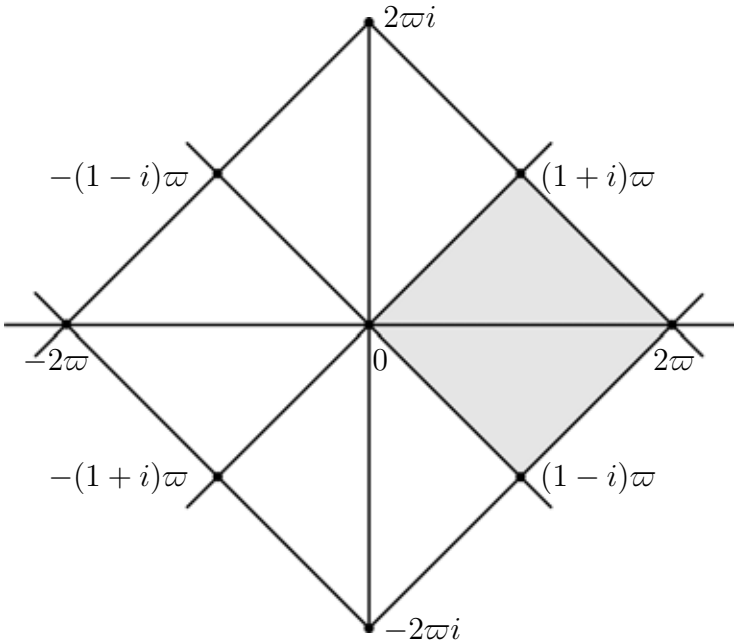
para $m, n \in \mathbb{Z}$. Por lo tanto la identidad $\varphi(z + m\varpi + n\varpi i) = (-1)^{m+n}\varphi(z)$ se sigue inmediatamente. □

La parte (c) de la Proposición 3.3.1 implica que φ es *doblemente periódica*:

$$\varphi(z) = \varphi(z + (1+i)\varpi) = \varphi(z + (1-i)\varpi). \tag{3.22}$$

Notemos que los periodos $(1+i)\varpi$ y $(1-i)\varpi$ son linealmente independientes en $\mathbb{R}$.

(3.23)



Los puntos señalados en la figura anterior son los números complejos del siguiente conjunto

$$\Omega_1 = \{(m+ni)\varpi \,|\, m+n \equiv 0 \pmod 2\} = \{m(1+i)\varpi + n(1-i)\varpi \,|\, m, n \in \mathbb{Z}\}.$$

Esta es la ***latiz de periodos*** de φ. La periodicidad doble implica que una vez que conocemos los valores de $\varphi(z)$ para toda z en uno de los rombos pequeños, conocemos sus valores para toda $z \in \mathbb{C}$.

3.3.2. Ceros y polos

Nuestra siguiente tarea será estudiar los ceros y polos de la función $\varphi(z)$. Recordemos que $z_0 \in \mathbb{C}$ es un ***cero simple*** de una función analítica $g(z)$ si $g(z_0) = 0$ y $g'(z_0) \neq 0$. Esto es equivalente a decir que la expansión de $g(z)$ en una serie de potencias en z_0 es de la forma

$$g(z) = a_1(z - z_0) + \sum_{n=2}^{\infty} a_n(z - z_0)^n, \quad a_1 \neq 0.$$

Por otro lado, z_0 es un ***polo simple*** de una función meromorfa $g(z)$ si la expansión de Laurent de $g(z)$ en z_0 es

$$g(z) = \frac{a_{-1}}{z - z_0} + \sum_{n=0}^{\infty} a_n(z - z_0)^n, \quad a_{-1} \neq 0.$$

Teorema 3.3.2. *$\varphi(z)$ es meromorfa en $\mathbb{C}$ con los siguientes ceros y polos:*

(*a*) *Todos los ceros son simples y ocurren cuando $z = (m + ni)\varpi$ para $m, n \in \mathbb{Z}$.*

(*b*) *Todos los polos son simples y ocurren cuando $z = (m + ni)\frac{\varpi}{2}$ para m, n impares.*

Demostración. Puesto que $\varphi(0) = 0$ y $\varphi'(0) = 1$, la parte (c) de la Proposición 3.3.1 fácilmente implica que φ tiene un cero simple en $(m + ni)\varpi$ para toda $m, n \in \mathbb{Z}$.

Usando la ley de adición junto con (3.18) y (3.19), observamos que

$$\varphi\left(z + \frac{\varpi}{2}\right) = \frac{\varphi(z)\varphi'(\frac{\varpi}{2}) + \varphi(\frac{\varpi}{2})\varphi'(z)}{1 + \varphi^2(z)\varphi^2(\frac{\varpi}{2})} = \frac{\varphi'(z)}{1 + \varphi^2(z)}.$$

De forma análoga,

$$\varphi\left(z \pm \frac{\varpi}{2}i\right) = \pm i\frac{\varphi'(z)}{1 - \varphi^2(z)}.$$

Multiplicando las dos ecuaciones anteriores obtenemos la importante identidad

$$\varphi\left(z + \frac{\varpi}{2}\right)\varphi\left(z \pm \frac{\varpi}{2}i\right) = \left(\frac{\varphi'(z)}{1 + \varphi^2(z)}\right)\left(\pm i\frac{\varphi'(z)}{1 - \varphi^2(z)}\right) = \pm i,$$

puesto que $\varphi'^2(z) = 1 - \varphi^4(z)$.

Reemplazando z con $z + \frac{\varpi}{2}$ y utilizando $\varphi(z + \varpi) = -\varphi(z)$, obtenemos que

$$\varphi(z)\varphi\left(z + (1 \pm i)\frac{\varpi}{2}\right) = \mp i. \tag{3.24}$$

Si $\varphi(z_0) = 0$, entonces $\varphi\left(z_0 + (1 + i)\frac{\varpi}{2}\right)$ esta indefinida por (3.24). De ahí que

$$z_0 + (1 + i)\frac{\varpi}{2} = (m + ni)\frac{\varpi}{2}, \quad m, n \quad \text{impares}$$

por la Proposición 3.3.1. Se sigue fácilmente que z_0 es uno de los ceros simples conocidos.

Para analizar los polos de φ, escribiremos (3.24) como

$$\varphi\left(z + (1 \pm i)\frac{\varpi}{2}\right) = \frac{\mp i}{\varphi(z)}.$$

Puesto que $\varphi(z)$ tiene un cero simple en $z = 0$, observamos que φ tiene polos simples en $z = (1 \pm i)\frac{\varpi}{2}$. Usando la doble periodicidad de φ, concluimos que φ tiene polos simples en $(m + ni)\frac{\varpi}{2}$ para m, n impares. Entonces hemos terminado, puesto que estos son las únicas singularidades posibles de φ por la Proposición 3.3.1. □

Nuestro siguiente resultado jugará un papel importante en la siguiente sección.

Teorema 3.3.3. *Fijando un número complejo w_0. Entonces la ecuación $\varphi(z) = w_0$ tiene una solución $z_0 \in \mathbb{C}$. Además, si z_0 es una solución entonces todas las soluciones están dadas por*

$$z = (-1)^{m+n} z_0 + (m + ni)\varpi, \quad m, n \in \mathbb{Z}.$$

Demostración. Recordemos que si $g(z)$ es analítica en una región $G \subset \mathbb{C}$, y sea $C \subset \mathbb{C}$ una curva cerrada simple, orientada en el sentido contrario al avance de las manecillas del reloj, ***el principio del argumento*** (ver [7] capítulo 6) asegura que si $g(z)$ no tiene ceros o polos en C, entonces

$$\frac{1}{2\pi i}\int_C \frac{g'(z)}{g(z)}dz = Z - P,$$

donde Z es el numero de ceros de $g(z)$ dentro de C, contando las multiplicidades, y P es el número de polos de $g(z)$ dentro de C, contando también las multiplicidades.

La función $g(z) = \varphi(z) - w_0$ tiene los mismos polos que φ, los cuales son $(m + ni)\frac{\varpi}{2}$, con m, n impares, por el Teorema 3.3.2. Esto significa que no podemos usar el rombo de (3.23). Sin embargo, puesto que los ceros de $g(z)$ están aislados, podemos desplazar uno de los rombos hacia la izquierda como en la siguiente figura para obtener una curva C tal que $g(z)$ no tenga ni ceros ni polos en C (Figura (3.25)).

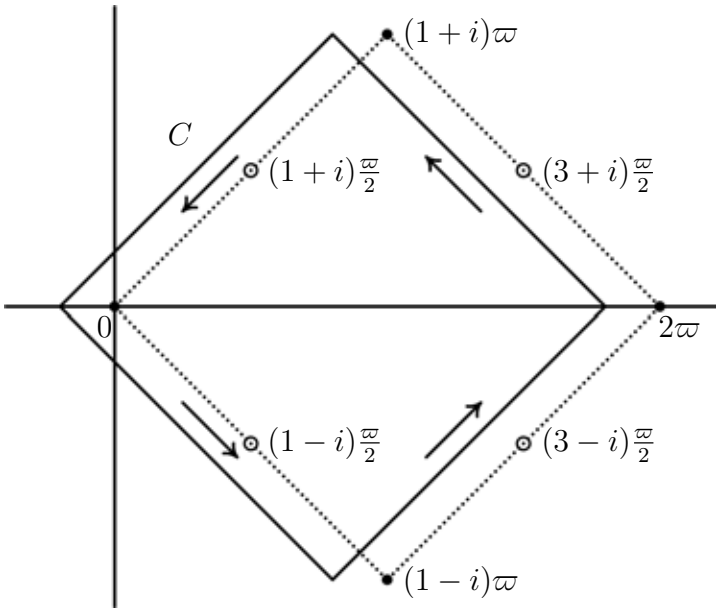
Los puntos dentro de los círculos pequeños son polos de $g(z)$ y son simples por el Teorema 3.3.2. Exactamente dos de ellos se encuentran en el interior de C, por lo tanto $P = 2$.

Puesto que $g(z) = \varphi(z) - w_0$ tiene periodos $(1 \pm i)\varpi$, lo mismo es cierto para $g'(z)$ y $g'(z)/g(z)$. Los bordes opuestos de C difieren por $(1 \pm i)\varpi$, por lo que $g'(z)/g(z)$ toma los mismos valores en los bordes opuestos. Por esta razón las integrales sobre bordes opuestos se cancelan, puesto que estos bordes tienen orientación opuesta. De aquí obtenemos que

$$Z - 2 = Z - P = \frac{1}{2\pi i}\int_C \frac{g'(z)}{g(z)}dz = 0.$$

Concluimos que dentro de C, $g(z)-w_0$ tiene dos ceros simples o un cero doble. En particular, $g(z)=w_0$ debe tener una solución z_0 dentro de C.

(3.25)



Conocido z_0, la Proposición 3.3.1 proporciona las soluciones adicionales, primero veamos que las que se proponen lo son

$$\varphi\left((-1)^{m+n}z_0+(m+ni)\varpi\right)=(-1)^{m+n}\varphi\left((-1)^{m+n}z_0\right)=\varphi(z_0)=w_0,$$

donde la segunda igualdad se sigue puesto que φ es impar. Debemos mostrar que no existen otras soluciones. Sea D la región encerrada por C (incluyendo la frontera). Trasladando D por elementos de la latiz de periodos Ω cubrimos completamente el plano complejo. En particular, $-z_0+\varpi$ tiene un trasladado por Ω que se encuentra en el interior de D, esto es, existen $m,n\in\mathbb{Z}$ con $m+n$ par tal que

$$-z_0+\varpi+(m+ni)\varpi=(-1)^{n+m+1}z_0+((m+1)+ni)\,\varpi \tag{3.26}$$

se encuentra dentro de la curva C. Puesto que cualquier otro cero tiene un trasladado por Ω que se encuentra dentro de C, se sigue que todas las soluciones de $\varphi(z)=w_0$ tienen la forma deseada. Finalmente, si (3.26) coincide con z_0, entonces es fácil de ver que

$$z_0=(a+bi)\frac{\varpi}{2},\quad a,b\in\mathbb{Z},\ a+b \text{ impar}.$$

De aquí podemos observar que por (3.17) tenemos que

$$\varphi(z_0)=\varphi\left(a\frac{\varpi}{2}+ib\frac{\varpi}{2}\right)=\frac{\varphi(a\frac{\varpi}{2})\varphi'(b\frac{\varpi}{2})+i\varphi(b\frac{\varpi}{2})\varphi'(a\frac{\varpi}{2})}{1-\varphi^2(a\frac{\varpi}{2})\varphi^2(b\frac{\varpi}{2})}.$$

Puesto que $a+b$ es impar solamente si uno de los numeros es par y el otro impar, por (3.18) y la periodicidad de φ y φ' tenemos que

$$\varphi(z_0)=\pm i\quad \text{ó}\quad \varphi(z_0)=\pm 1,$$

en cualquiera de los cuatro caso tenemos que $\varphi^4(z_0) = 1$, recordando que $\varphi'^2(z) = 1 - \varphi^4(z)$ hemos mostrado que $\varphi'(z_0) = 0$ cuando z_0 tiene la forma anteriormente mencionada. Por lo que probamos antes, se sigue que z_0 es el único cero de $g(z)$ dentro de C. Luego, concluimos que las soluciones tienen la forma deseada. □

Notas Matemáticas

Dos ideas implícitas en esta sección requieren ser comentadas.

• **Funciones Elípticas.** Por la proposición 3.3.1, φ es una función meromorfa en $\mathbb{C}$ con periodos $(1+i)\varpi$, $(1-i)\varpi$ que son linealmente independientes en $\mathbb{R}$. Como hemos visto en el capítulo anterior, una ***función elíptica*** es una función meromorfa en $\mathbb{C}$ con periodos ω_1, ω_2 que son linealmente independientes en $\mathbb{R}$. Mientras las ideas básicas de las funciones elípticas remontan a Abel y Jacobi, ahora los textos siguen la presentación de Weierstrass, quien definió la ***función*** $\wp$ ***de Weierstrass*** como

$$\wp(z;\omega_1,\omega_2) = \frac{1}{z^2} + \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \left(\frac{1}{(z-(n\omega_1+m\omega_2))^2} - \frac{1}{(n\omega_1+m\omega_2)^2} \right).$$

Por ejemplo, si denotamos como $\wp_1(z)$ a la función $\wp$ con periodos $(1+i)\varpi$, $(1-i)\varpi$, entonces podemos mostrar que

$$\varphi(z) = -2\frac{\wp_1(z)}{\wp_1'(z)} \quad \text{y} \quad \varphi'(z) = \frac{4\wp_1^2(z)-1}{4\wp_1^2(z)+1}. \tag{3.27}$$

Además, la relación

$$\varphi'^2(z) = 1 - \varphi^4(z)$$

se traduce en la relación

$$\wp_1'^2(z) = 4\wp_1^3(z) + \wp_1(z). \tag{3.28}$$

En general, la $\wp(z) = \wp(z;\omega_1,\omega_2)$ satisface

$$\wp'^2(z) = 4\wp^3(z) - g_2\wp(z) - g_3, \tag{3.29}$$

donde g_2 y g_3 son constantes determinadas por los periodos ω_1 , ω_2. Existe también una ley de adición para $\wp(z+w)$. Como lo hemos señalado en el capítulo 2.

• **Curvas Elípticas.** El objeto geométrico primario es la lemniscata, la cual es la curva definida por la ecuación $(x^2+y^2)^2 = x^2 - y^2$. Sin embargo, las funciones elípticas que hemos estado estudiando darán otras curvas de interés. Por ejemplo, la relación

$$\varphi'^2(z) = 1 - \varphi^4(z)$$

muestra que la función $z \mapsto (\varphi(z), \varphi'(z))$ parametriza la curva $y^2 = 1 - x^4$. De forma similar, la relación (3.28) para la función de Weierstrass $\wp_1(z)$ muestra que $z \mapsto (\wp_1(z), \wp_1'(z))$ parametriza la curva

$$y^2 = 4x^3 + x,$$

y en general para una función $\wp$, la ecuación (3.29) muestra que $z \mapsto (\wp(z), \wp'(z))$ parametriza

$$y^2 = 4x^3 - g_2x - g_3.$$

Estas son las ***curvas elípticas***. Y tienen una ley de adición intrínseca con la ley de adición de la función $\wp$. Algunos de los más importantes teoremas y conjeturas de la teoría de números moderna está relacionada con las curvas elípticas.

3.4. Multiplicación compleja

Utilizando el teorema de identidad (ver pie de página de la sección 3.3.1) podemos observar que las fórmulas de multiplicación para $\varphi(nx)$, $x \in \mathbb{R}$, se pueden extender a fórmulas $\varphi(nz)$, $z \in \mathbb{C}$. Sobre $\mathbb{C}$ también tenemos la fórmula (3.19) dada por

$$\varphi(iz) = i\varphi(z), \quad i = \sqrt{-1}. \tag{3.30}$$

Por lo tanto además de multiplicar por $n \in \mathbb{Z}$, también podemos multiplicar por i. Combinando esto con las leyes de adición obtenemos fórmulas para $\varphi((n+mi)x)$, donde $n + mi \in \mathbb{Z}[i]$ es un entero Gaussiano. En otras palabras, $\varphi(z)$ tiene ***multiplicación compleja*** por elementos de $\mathbb{Z}[i]$.

Antes de desarrollar la teoría general, daremos un ejemplo para ilustrar el poder de la multiplicación compleja.

Ejemplo 3.4.1. Como hemos mencionado utilizando las leyes de adición junto con las identidades $\varphi(iz) = i\varphi(z)$ y $\varphi(-z) = -\varphi(z)$ obtenemos las siguientes fórmulas

$$\begin{aligned} \varphi((1+i)z) &= \frac{(1+i)\varphi(z)\varphi'(z)}{1-\varphi^4(z)}, \\ \varphi((1-i)z) &= \frac{(1-i)\varphi(z)\varphi'(z)}{1-\varphi^4(z)}. \end{aligned} \tag{3.31}$$

Estas fórmulas son ejemplos simples de la multiplicación compleja.

Para observar la relevancia de (3.31), elevemos al cuadrado cada lado y apliquemos $\varphi'^2(z) = 1 - \varphi^4(z)$. Obtenemos que

$$\begin{aligned} \varphi^2((1+i)z) &= \frac{2i\varphi^2(z)}{1-\varphi^4(z)}, \\ \varphi^2((1-i)z) &= \frac{-2i\varphi^2(z)}{1-\varphi^4(z)}. \end{aligned} \tag{3.32}$$

La sorpresa es que hemos visto versiones similares de estas fórmulas en la prueba de la Proposición 3.2.3. Para explicar por que, sean $r_0 = \varphi\left(\frac{x_0}{2}\right)$ y $a = \varphi(x_0)$ como en la prueba

de la proposición requerida. Entonces definamos $t = \varphi\left((1+i)\frac{x_0}{2}\right)$ y apliquemos la primera fórmula de (3.32) para obtener

$$t^2 = \frac{2ir_0^2}{1-r_0^4}.$$

Puesto que $2 = (1-i)(1+i)$, la segunda fórmula de (3.32) implica que

$$a^2 = \varphi^2(x_0) = \varphi^2\left((1-i)(1+i)\frac{x_0}{2}\right) = \frac{-2i\varphi^2((1+i)\frac{x_0}{2})}{1-\varphi^4((1+i)\frac{x_0}{2})} = \frac{-2it^2}{1-t^4}.$$

Las dos fórmulas anteriores son justamente las fórmulas (3.12) y (3.13) de la Proposición 3.2.3. Antes, estas al parecer surgieron de la nada, pero ahora que conocemos la multiplicación compleja, ya no son tan misteriosas.

La prueba de la Proposición 3.2.3 usó la fórmula de duplicación para $\varphi(2x)$. El Ejemplo 3.4.1 muestra que la factorizar 2 en $\mathbb{Z}[i]$ nos permite factorizar la fórmula de duplicación en ecuaciones que son mas simples de entender. Usaremos factorizaciones similares en la Sección 3.5 cuando probemos el Teorema de Abel sobre la lemniscata.

La teoría de la *multiplicación compleja* proporciona fórmulas para $\varphi(\beta z)$, donde $z \in \mathbb{C}$ y $\beta = n + mi \in \mathbb{Z}[i]$ es un entero Gaussiano. En esta sección primero revisaremos algunos hechos básicos de $\mathbb{Z}[i]$ y después derivaremos fórmulas para $\varphi(\beta z)$, poniendo especial atención al caso cuando β es primo en $\mathbb{Z}[i]$.

3.4.1. Los enteros Gaussianos

El anillo de los ***enteros Gaussianos*** esta definido por

$$\mathbb{Z}[i] = \{a + ib \,|\, a, b \in \mathbb{Z}\}.$$

Las ***unidades*** de $\mathbb{Z}[i]$ forman el grupo $\mathbb{Z}[i]^* = \{\pm 1, \pm i\} = \{i^\varepsilon \,|\, \varepsilon = 0, 1, 2, 3\}$, y dos enteros Gaussianos α y β están ***asociados*** si $\alpha = i^\varepsilon\beta$ para algún $i^\varepsilon \in \mathbb{Z}[i]^*$. Además, $\mathbb{Z}[i]$ es un dominio de factorización única con los siguientes primos (más los asociados):

- $2 = (1+i)(1-i)$, donde $1+i$ y $1-i$ son primos asociados en $\mathbb{Z}[i]$.
- Cuando $p \equiv 3 \mod 4$ es un primo en $\mathbb{Z}$, p también es un primo en $\mathbb{Z}[i]$.
- Cuando $p \equiv 1 \mod 4$ es un primo en $\mathbb{Z}$, existen $a, b \in \mathbb{Z}$ tal que $p = a^2 + b^2 = (a+bi)(a-bi)$, donde $a+bi$ y $a-bi$ son primos no asociados en $\mathbb{Z}[i]$.

También, $\mathbb{Z}[i]$ es un dominio de ideales principales, por lo tanto todo ideal es de la forma $\beta\mathbb{Z}[i]$ para algún $\beta \in \mathbb{Z}[i]$. Estos hechos están probados en la Sección 3.8. de [4].

Dada $\alpha \in \mathbb{Z}[i]$, decimos que $\beta \equiv \gamma \mod \alpha$ si α divide a $\beta - \gamma$ en $\mathbb{Z}[i]$. Para entender el anillo cociente $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$, necesitamos definir para $\alpha = a + ib \in \mathbb{Z}[i]$ su ***norma***

$$N(\alpha) = \alpha\overline{\alpha} = |\alpha|^2 = a^2 + b^2 \in \mathbb{Z}.$$

La norma cumple que $N(\alpha\beta) = N(\alpha)N(\beta)$. Entonces tenemos el siguiente resultado.

Lema 3.4.2. *Sea α un elemento de $\mathbb{Z}[i]$ distinto del cero. Entonces:*

(a) *$\mathbb{Z}[i]/\alpha\mathbb{Z}[i]$ es un anillo finito con $N(\alpha)$ elementos.*

(b) *Si α es primo, entonces $\mathbb{Z}[i]/\alpha\mathbb{Z}[i]$ es el campo finito*

$$\mathbb{Z}[i]/\alpha\mathbb{Z}[i] \simeq \mathbb{F}_{N(\alpha)}.$$

Demostración. Para probar (a) sea $\alpha \in \mathbb{Z}[i]$ y sea m el máximo común divisor de las partes real e imaginaria de α, de tal forma que $\alpha = m(a+bi)$, donde el $\mathrm{mcd}(a,b) = 1$ y tomemos $c, d \in \mathbb{Z}$ tales que $ad - bc = 1$. Entonces observando que la función de $\mathbb{Z}[i] \to \mathbb{Z} \oplus \mathbb{Z}$ definida por

$$\mu + i\nu \mapsto \mu(d,-b) + \nu(-c,a) = (\mu d - \nu c, -\mu b + \nu a)$$

es un isomorfismo de grupos bajo la adición, y puesto que, las imágenes bajo la función de α e $i\alpha$ están dadas por

$$\begin{array}{rcl} \alpha = ma + imb & \mapsto & (m,0) \\ i\alpha = -mb + ima & \mapsto & (-m(ac+bd), m(a^2+b^2)) \end{array}$$

tenemos que la función manda el subgrupo $\alpha\mathbb{Z}[i] \subset \mathbb{Z}[i]$ al subgrupo

$$m\mathbb{Z} \oplus m(a^2+b^2)\mathbb{Z} \subset \mathbb{Z} \oplus \mathbb{Z}.$$

Por lo tanto $|\mathbb{Z}[i]/\alpha\mathbb{Z}[i]| = m \cdot m(a^2+b^2) = N(\alpha)$.

Para la parte (b) sea $\alpha \in \mathbb{Z}[i]$ primo, entonces $\alpha = p$ donde p es primo en $\mathbb{Z}$ ó $\alpha = a + ib$ donde $a^2 + b^2 = p$ y p es un primo en $\mathbb{Z}$. Entonces por la parte (a) tenemos que $\mathbb{Z}[i]/\alpha\mathbb{Z}[i]$ es un anillo finito con cardinalidad p o p^2 al cual sabemos que se le puede asociar un isomorfismo con $\mathbb{F}_p$. □

Diremos que un entero Gaussiano $a + ib \in \mathbb{Z}[i]$ es ***impar*** si $a + b$ es impar y es ***par*** si $a + b$ es par. Si $\alpha, \beta \in \mathbb{Z}[i]$, son inmediatas las afirmaciones:

$$\begin{array}{rcl} \alpha\beta \text{ es impar} & \Leftrightarrow & \alpha \text{ y } \beta \text{ son impares,} \\ \alpha + \beta \text{ es par} & \Leftrightarrow & \alpha, \beta \text{ son ambos pares o ambos impares,} \\ \alpha \text{ es par} & \Leftrightarrow & 1 + i \text{ divide a } \alpha. \end{array} \tag{3.33}$$

Puesto que $1 + i$ es primo en $\mathbb{Z}[i]$, la última línea de (3.33) puede ser reformulada como

$$\alpha \text{ es impar} \Leftrightarrow 1 + i \text{ y } \alpha \text{ son primos relativos.}$$

Mas adelante utilizaremos el hecho de que el criterio de irreducibilidad de Schönemann-Eisenstein para polinomios en $\mathbb{Z}[u]$ y primos en $\mathbb{Z}$, también es aplicable para polinomios en $\mathbb{Z}[i][u]$ y primos en $\mathbb{Z}[i]$. Mostremos esto.

Teorema 3.4.3. *Sea $\beta \in \mathbb{Z}[i]$ primo, y sea $f = a_0u^d + a_1u^{d-1} + \cdots + a_d \in \mathbb{Z}[i][u]$ de grado $d > 0$. Si β es tal que $\beta \nmid a_0$, $\beta | a_1, \ldots, \beta | a_d$ y $\beta^2 \nmid a_d$, entonces f es irreducible sobre $\mathbb{Q}(i)$.*

Demostración. Supongamos que f es reducible en $\mathbb{Q}(i)$ entonces existen polinomios $g, h \in \mathbb{Z}[i][u]$ de grado $< d$ tales que $f = gh$. Entonces considerando el homomorfismo de anillos $[Z][i][u] \to \mathbb{F}_{N(\beta)}[u]$ que se define enviando $q = b_0u^k + b_1u^{k-1} + \cdots + b_k \in \mathbb{Z}[i][u]$ a $\overline{q} = [b_0]u^k + [b_1]u^{k-1} + \cdots + [b_k] \in \mathbb{F}_{N(\beta)}[u]$ donde $[b] \in \mathbb{F}_{N(\beta)}[u]$ es la clase de congruencia módulo β de $b \in \mathbb{Z}[i]$.

Entonces $f = gh$ implica que $[a_0]u^d = \overline{g}\overline{h}$, puesto que $\beta | a_1, \ldots, \beta | a_d$. Sin embargo, $\mathbb{F}_{N(\beta)}$ es un campo, lo cual significa que la factorización única se cumple en $\mathbb{F}_{N(\beta)}[u]$. Dada que $\beta \nmid a_0$, se sigue que $\overline{g} = [a]x^r$ y $\overline{h} = [b]x^s$, donde $[a][b] = [a_0]$ y $r + s = d$.

Ahora, si $r = 0$, entonces $\overline{g} = [a]$ y grado$(g) > 0$ implicaran que el primer término de g es divisible entre β, entonces $f = gh$ implicara que lo mismo es cierto para el primer término a_0 de f. De esta manera $\beta \nmid a_0$ implica que $r > 0$, y para $s > 0$ se procede de forma similar.

Pero entonces $\overline{g} = [a]x^r$ para $r > 0$ implica que β divide al término constante de g, y lo mismo pasa para el término constante de h, puesto que $s > 0$. Dado que el término constante a_d de f es el producto de los términos constantes de g y h, se sigue que $\beta^2 | a_d$. Esto contradice $\beta^2 \nmid a_d$. □

3.4.2. Multiplicación por enteros Gaussianos

Cuando $n \in \mathbb{Z}$, el Teorema 3.2.5 expresa a $\varphi(nz)$ en términos de $\varphi(z)$ cuando n es impar y en términos de $\varphi(z)$ y $\varphi'(z)$ cuando n es par. Aquí, generalizaremos el caso anterior proporcionando fórmulas para $\varphi(\beta z)$ en términos de $\varphi(z)$ cuando $\beta \in \mathbb{Z}[i]$ sea impar.

En cierto sentido, las fórmulas son fáciles – la prueba del Teorema 3.4.5 que proporcionaremos mas adelante muestra que estas son simples consecuencias de la ley de adición, las fórmulas de la multiplicación para $\varphi(nz)$ del Teorema 3.2.5, y la identidad $\varphi(iz) = i\varphi(z)$. Sin embargo, para probar el teorema de Abel sobre la lemniscata, necesitamos entender la fina estructura de estas fórmulas.

En el siguiente ejemplo se ilustran algunas de las relaciones implicadas.

Ejemplo 3.4.4. Deduzcamos una fórmula para $\varphi((2+i)z)$. Sean $x = z$ y $y = (1+i)z$ por (3.10) tenemos

$$\varphi(z + (1+i)z) + \varphi(z - (1+i)z) = \frac{2\varphi(z)\varphi'((1+i)z)}{1 + \varphi^2(z)\varphi^2((1+i)z)}$$

y ahora utilizando (3.19) tenemos que

$$\varphi((2+i)z) = \varphi(z)\left(\frac{2\varphi'((1+i)z)}{1 + \varphi^2(z)\varphi^2((1+i)z)} + i\right),$$

para obtener el valor de $\varphi'((1+i)z)$, recordemos que por (3.32) tenemos

$$\varphi^4\left((1+i)z\right) = -\frac{4\varphi^4(z)}{(1-\varphi^4(z))^2},$$

de este hecho junto con la fórmula $\varphi'^2 = 1-\varphi^4(z)$ obtenemos que

$$\varphi'\left((1+i)z\right) = \frac{1+\varphi^4(z)}{1-\varphi^4(z)}.$$

Sustituyendo los valores de $\varphi'((1+i)z)$ y $\varphi^2((1+i)z)$ en la primera fórmula obtenemos

$$\varphi\left((2+i)z\right) = -i\,\varphi(z)\frac{\varphi^4(z)+(-1+2i)}{(-1+2i)\varphi^4(z)+1}. \tag{3.34}$$

Factorizamos el valor $-i$ para asegurarnos de que el numerador es mónico y de que el denominador tiene termino constante 1. Notemos también la "simetría inversa" de los coeficientes del numerador y el denominador. Este hecho será importante más adelante.

El siguiente Teorema generaliza la fórmula (3.34).

Teorema 3.4.5. *Sea $\beta \in \mathbb{Z}[i]$ impar. Entonces existen polinomios primos relativos $P_\beta(u), Q_\beta(u)$ en el anillo de polinomios $\mathbb{Z}[i][u]$ y existe $\varepsilon \in \{0,1,2,3\}$ tales que*

(a) *Para toda $z \in \mathbb{C}$, tenemos*

$$\varphi(\beta z) = i^\varepsilon \varphi(z)\frac{P_\beta\left(\varphi^4(z)\right)}{Q_\beta\left(\varphi^4(z)\right)}.$$

(b) $\beta \equiv i^\varepsilon \mod 2(1+i)$.

(c) *$P_\beta(u)$ y $Q_\beta(u)$ tienen* grado $d = (N(\beta)-1)/4$, *donde $N(\beta)$ es la norma de β.*

(d) *Las raíces del β-***polinomio de división** *$uP_\beta(u^4)$ son los números complejos $\varphi\left(\alpha\frac{\varpi}{\beta}\right)$ para $\alpha \in \mathbb{Z}[i]$ impar.*

(e) *$P_\beta(u)$ es mónico, $Q_\beta(0) = 1$, y $Q_\beta(u) = u^d P_\beta(1/u)$, donde d es el de la parte (c).*

Antes de comenzar la prueba, expliquemos que dice el teorema acerca de $\varphi(\beta z)$ cuando $\beta \in \mathbb{Z}[i]$ es impar. Sean $P_\beta(u), Q_\beta(u) \in \mathbb{Z}[i][u]$ los polinomios dados en el teorema. Las partes (c) y (e) implican que $P_\beta(u)$ y $Q_\beta(u)$ pueden ser escritos de la forma

$$\begin{aligned} P_\beta(u) &= u^d + a_1 u^{d-1} + \cdots + a_d \\ Q_\beta(u) &= u^d P_\beta(1/u) \\ &= u^d\left((1/u)^d + a_1(1/u)^{d-1} + \cdots + a_d\right) \\ &= 1 + a_1 u + \cdots + a_d u^d, \end{aligned}$$

donde $d = (N(\beta) - 1)/4$ y $a_1, \ldots, a_d \in \mathbb{Z}[i]$. Esta es la "simetría inversa" mencionada anteriormente. Entonces la fórmula de la multiplicación compleja para $\varphi(\beta z)$ puede ser escrita como

$$\varphi(\beta z) = i^{\varepsilon}\varphi(z)\frac{\varphi^{4d}(z) + a_1\varphi^{4d-4}(z) + \cdots + a_d}{1 + a_1\varphi^4(z) + \cdots + a_d\varphi^{4d}(z)},$$

donde $\beta \equiv i^{\varepsilon} \mod 2(1+i)$ por la parte (b). Demos otro ejemplo.

Ejemplo 3.4.6. Supongamos que $\beta = 2+i$. Puesto que $d = (N(\beta)-1)/4 = (5-1)/4 = 1$ y $\beta \equiv -i \mod 2(1+i)$, la fórmula anterior se reduce a

$$\varphi\left((2+i)z\right) = -i\varphi(z)\frac{\varphi^4(z) + a_1}{a_1\varphi^4(z) + 1},$$

donde $a_1 \in \mathbb{Z}[i]$. Comparando esto con (3.34), deberá suceder que $a_1 = -1 + 2i$.

El siguiente lema será útil en la prueba del Teorema 3.4.5.

Lema 3.4.7. *Sea $\beta \in \mathbb{Z}[i]$ impar. Entonces el conjunto*

$$R_\beta = \{\varphi(z) \,|\, z \in \mathbb{C},\, \varphi(\beta z) = 0\}$$

tiene precisamente $N(\beta)$ elementos y consiste en todos los números complejos de la forma

$$\varphi\left(\alpha\frac{\varpi}{\beta}\right), \quad \textit{con} \quad \alpha \in \mathbb{Z}[i] \textit{ impar.}$$

Demostración. Primero observemos que si $\alpha \in \mathbb{Z}[i]$ es impar, entonces $\varphi\left(\alpha\frac{\varpi}{\beta}\right) \in R_\beta$, puesto que $\varphi\left(\beta \cdot \alpha\frac{\varpi}{\beta}\right) = \varphi(\alpha\varpi) = 0$, donde la última igualdad esta dada por el Teorema 3.3.2. Tomando otro camino, supongamos que $\varphi(\beta z) = 0$. Entonces el Teorema 3.3.2 implica que

$$\beta z = (a + ib)\varpi, \text{ con } a, b \in \mathbb{Z}.$$

Sea $\alpha = a + ib \in \mathbb{Z}[i]$. Entonces $z = \alpha\frac{\varpi}{\beta}$, por lo que $\varphi(z) = \varphi\left(\alpha\frac{\varpi}{\beta}\right)$. Si α es impar, entonces hemos terminado. Por otra parte, si α es par, entonces $\beta - \alpha$ es impar. Usando la identidad $\varphi(\varpi - z) = \varphi(z)$ de (3.6), obtenemos

$$\varphi\left((\beta-\alpha)\frac{\varpi}{\beta}\right) = \varphi\left(\varpi - \alpha\frac{\varpi}{\beta}\right) = \varphi\left(\alpha\frac{\varpi}{\beta}\right).$$

Esto muestra que los elementos de R_β, tienen la forma deseada.

Para determinar el tamaño de R_β, fijemos $\varphi\left(\alpha\frac{\varpi}{\beta}\right) \in R_\beta$, donde $\alpha \in \mathbb{Z}[i]$ es impar. Afirmamos que α es única módulo $\beta\mathbb{Z}[i]$. Para ver esto, supongamos que

$$\varphi\left(\alpha\frac{\varpi}{\beta}\right) = \varphi\left(\widetilde{\alpha}\frac{\varpi}{\beta}\right), \quad \alpha, \widetilde{\alpha} \in \mathbb{Z}[i] \text{ impares.}$$

Por el Teorema 3.3.3, existe $a+ib \in \mathbb{Z}[i]$ tal que

$$\widetilde{\alpha}\frac{\varpi}{\beta} = (-1)^{a+b}\alpha\frac{\varpi}{\beta} + (a+ib)\varpi.$$

Esto implica que

$$\widetilde{\alpha} = (-1)^{a+b}\alpha + (a+ib)\beta.$$

Puesto que $\alpha, \widetilde{\alpha}$ y β son impares, $a+ib$ es par por (3.33). Por esta razón $(-1)^{a+b} = 1$ y de ahí que

$$\widetilde{\alpha} = \alpha + (a+ib)\beta,$$

por lo tanto α y $\widetilde{\alpha}$ son congruentes módulo $\beta\mathbb{Z}[i]$. Puesto que todo elemento de $\mathbb{Z}[i]/\beta\mathbb{Z}[i]$ puede ser representado por un entero Gaussiano impar (dado cualquier α se tiene que α ó $\alpha+\beta$ es impar), se sigue que

$$|R_\beta| = |\mathbb{Z}[i]/\beta\mathbb{Z}[i]| = N(\beta),$$

donde la última igualdad se obtiene del Lema 3.4.2. □

Comencemos ahora la prueba del teorema.

***Demostración del Teorema* 3.4.5**. Probaremos el teorema en cinco pasos.

Paso 1: Existencia de $P_\beta(u)$ y $Q_\beta(u)$ para toda β. Dada $\beta \in \mathbb{Z}[i]$, afirmamos que existen $P_\beta(u), Q_\beta(u) \in \mathbb{Z}[i][u]$ tales que $Q_\beta(0) = 1$ y

$$\varphi(\beta z) = \varphi(z)\frac{P_\beta(\varphi^4(z))}{Q_\beta(\varphi^4(z))} \tag{3.35}$$

cuando β es impar, y

$$\varphi(\beta z) = \varphi(z)\frac{P_\beta(\varphi^4(z))}{Q_\beta(\varphi^4(z))}\varphi'(z) \tag{3.36}$$

cuando β es par. Probaremos (3.35) y (3.36) usando las fórmulas de multiplicación del Teorema 3.2.5 junto con las identidades

$$\begin{aligned}
\varphi(iz) &= i\varphi(z), \\
\varphi((1+i)z) &= \frac{(1+i)\varphi(z)\varphi'(z)}{1-\varphi^4(z)}, \\
\varphi((\beta+1)z) &= -\varphi((\beta-1)z) + \frac{2\varphi(\beta z)\varphi'(z)}{1+\varphi^2(\beta z)\varphi^2(z)}, \\
\varphi((\beta+i)z) &= -\varphi((\beta-i)z) + \frac{2i\varphi(\beta z)\varphi'(z)}{1-\varphi^2(\beta z)\varphi^2(z)}.
\end{aligned} \tag{3.37}$$

Nosotros solamente hemos demostrado la primera y la segunda líneas, la tercera y la cuarta se siguen de la primera y de (3.10) sustituyendo los valores $x = \beta z, y = z$ y $x = \beta z, y = iz$ en cada caso.

Primero obtendremos fórmulas para $\varphi((n+i)z)$ para todo entero $n \geq 0$. Para esto bastará tomar $\beta = n+i$, sustituir este valor en la tercera línea y repetir los argumentos de la prueba del Teorema 3.2.5, tomando como hipótesis de inducción las fórmulas para $\varphi(iz)$ y $\varphi((1+i)z)$. En particular cuando n es impar, siguiendo este procedimiento obtenemos la fórmula recursiva

$$Q_{n+1+i}(u) = Q_{n-1+i}(u)\left[Q_{n+i}^2(u) + uP_{n+i}^2(u)(1-u)\right]$$

similar a (3.15). Esto facilita mostrar que $Q_{n+i}(0) = 1$ para toda $n \geq 0$ impar, el argumento para probar que $Q_{n+1}(0) = 1$ para toda $n \geq 0$ par es similar.

Ahora obtendremos fórmulas para $\varphi((n+im)z)$ para $n+im \in \mathbb{Z}[i]$, $n, m \geq 0$. Para esto fijemos un entero $n \geq 0$ y tomemos $\beta = n+im$, sustituyendo este valor en la cuarta linea y repitiendo los argumentos de la prueba del Teorema 3.2.5, tomando en este caso como hipótesis de inducción las fórmulas para $\varphi(nz)$ y $\varphi((n+i)z)$. De forma similar al caso anterior, en particular cuando $n+im$ es impar, siguiendo el procedimiento descrito obtenemos la fórmula recursiva

$$Q_{n+(m+1)i}(u) = Q_{n+(m-1)i}(u)\left[Q_{n+im}^2(u) + uP_{n+im}^2(u)\right].$$

Con esto podemos mostrar que $Q_{n+im}(0) = 1$ para toda $n+im$, $n, m \geq 0$ par, el argumento para probar que $Q_{n+im}(0) = 1$ para toda $n+im$, $n, m \geq 0$ impar es similar. De esta manera hemos demostrado que $Q_{n+im}(0) = 1$ para todo par de enteros positivos n, m.

Por lo tanto hemos obtenido fórmulas para $\varphi((n+im)z)$ para todo par de enteros $n, m \geq 0$. Entonces como

$$\begin{array}{rcccl} \varphi((-m+in)z) & = & \varphi(i(n+im)z) & = & i\varphi((n+im)z), \\ \varphi((-n-im)z) & = & \varphi(-(n+im)z) & = & -\varphi((n+im)z), \\ \varphi((m-in)z) & = & \varphi(-i(n+im)z) & = & -i\varphi((n+im)z) \end{array} \tag{3.38}$$

se facilitrá la construcción de los polinomios deseados $P_\beta(u), Q_\beta(u) \in \mathbb{Z}[i][u]$ para toda $\beta \in \mathbb{Z}[i]$.

Paso 2: Remover los factores comunes. Para el resto de la demostración, asumiremos que β es impar. Los polinomios P_β, Q_β construidos en el Paso 1 pueden tener un factor común. Puesto que $\mathbb{Z}[i]$ es un dominio de factorización única entonces también esto es cierto para $\mathbb{Z}[i][u]$ por el Teorema A.1.2 del ápendice. De ahí que

$$P_\beta(u) = C_\beta(u)\widetilde{P}_\beta(u) \quad \text{y} \quad Q_\beta(u) = C_\beta(u)\widetilde{Q}_\beta(u),$$

donde $C_\beta(u), \widetilde{P}_\beta(u), \widetilde{Q}_\beta(u) \in \mathbb{Z}[i][u]$ y $\widetilde{P}_\beta(u), \widetilde{Q}_\beta(u)$ son primos relativos. Como $Q_\beta(0) = 1$, podemos multiplicar $C_\beta(u), \widetilde{P}_\beta(u), \widetilde{Q}_\beta(u)$ por unidades convenientes en $\mathbb{Z}[i]^* = \{\pm 1, \pm i\}$ de tal forma que $\widetilde{Q}_\beta(0) = 1$. Al ser β es impar tenemos que

$$\varphi(\beta z) = \varphi(z)\frac{P_\beta(\varphi^4(z))}{Q_\beta(\varphi^4(z))} = \varphi(z)\frac{C_\beta(\varphi^4(z))\widetilde{P}_\beta(\varphi^4(z))}{C_\beta(\varphi^4(z))\widetilde{Q}_\beta(\varphi^4(z))} = \varphi(z)\frac{\widetilde{P}_\beta(\varphi^4(z))}{\widetilde{Q}_\beta(\varphi^4(z))}.$$

Por lo tanto podemos suponer que $P_\beta(u)$ y $Q_\beta(u)$ son primos relativos en $\mathbb{Z}[i][u]$.

Paso 3: La Constante i^ε. Un ejercicio de rutina muestra que $(\mathbb{Z}[i]/2(1+i)\mathbb{Z}[i])^* = \{\pm[1], \pm[i]\}$. Ahora si β es impar entonces β no es múltiplo de $2(1+i)$ y entonces la clase $[\beta]$ de β es alguna de $\{\pm[1], \pm[i]\}$ por lo que $\beta \equiv i^\varepsilon \mod 2(1+i)$ para alguna $\varepsilon \in \{0,1,2,3\}$. Multiplicando $P_\beta(u)$ por una unidad apropiada de $\mathbb{Z}[i]^*$, obtenemos

$$\varphi(\beta z) = i^\varepsilon \varphi(z) \frac{P_\beta(\varphi^4(z))}{Q_\beta(\varphi^4(z))}. \tag{3.39}$$

Si $\beta \equiv i^\varepsilon \mod 2(1+i)$ entonces $\beta - i^\varepsilon = 2(1+i)(m+in)$, ahora usando que φ es periódica o bien la proposición 3.3.1 (c) junto con (3.18) y (3.19) se puede ver que

$$\varphi\left(\beta\frac{\varpi}{2}\right) = i^\varepsilon. \tag{3.40}$$

Esto último se usará después. Es inmediato que los polinomios relativos $P_\beta(u)\,, Q_\beta(u) \in \mathbb{Z}[u]$ satisfacen las partes (a) y (b) del teorema, junto con la condición $Q_\beta(0) = 1$ de la parte (e). Los pasos siguientes serviran para demostrar que $P_\beta(u)\,, Q_\beta(u)$ cumplen las condiciones restantes del teorema.

Paso 4: Las Raíces de $uP_\beta(u^4)$. Usaremos el Lema 3.4.7 para determinar las raíces del β-polinomio de división $A_\beta(u) = uP_\beta(u^4)$. También sea $B_\beta(u) = Q_\beta(u^4)$. Puesto que β es impar, (3.39) implica que

$$\varphi(\beta z) = i^\varepsilon \frac{A_\beta(\varphi(z))}{B_\beta(\varphi(z))}. \tag{3.41}$$

Puesto que $P_\beta(u^4)$ y $Q_\beta(u^4)$ son primos relativos en $\mathbb{Z}[i][u]$, estos polinomios no tiene raíces en común por lo tanto la única raíz común que pueden tener $A_\beta(u)$ y $B_\beta(u)$ es el cero, pero ya que $B_\beta(0) = Q_\beta(0) = 1$ se observa que $A_\beta(u)$ y $B_\beta(u)$ también son primos relativos en $\mathbb{Z}[i][u]$. Utilizando este hecho y (3.41), se sigue que

$$A_\beta(\varphi(z)) = 0 \Leftrightarrow \varphi(\beta z) = 0.$$

Puesto que cualquier raíz de $A_\beta(u)$ es de la forma $\varphi(z)$ para alguna $z \in \mathbb{C}$ por el Teorema 3.3.3, concluimos que las raíces de $A_\beta(u)$ forman el conjunto

$$R_\beta = \{\varphi(z) \,|\, z \in \mathbb{Z}\,,\, \varphi(\beta z) = 0\}$$

del Lema 3.4.7. Entonces el lema implica que la raíz puede ser escrita en la forma que afirma la parte (d) del teorema.

Ahora mostraremos que todas las raíces tienen multiplicidad 1. Para esto supongamos que $u_0 = \varphi(z_0)$ es una raíz múltiple. Entonces $A_\beta(u_0) = A'_\beta(u_0) = 0$, y de ahí que $B_\beta(u_0) \neq 0$ por el párrafo anterior. Diferenciando (3.41) con respecto a z y evaluando en $z = z_0$ tenemos que

$$\varphi'(\beta z_0)\beta = i^\varepsilon \frac{B_\beta(u_0)A'_\beta(u_0)\varphi'(z_0) - B'_\beta(u_0)A_\beta(u_0)\varphi'(z_0)}{B^2_\beta(u_0)} = 0$$

(notemos que $\varphi'(z_0)$ está definida por que $\varphi(z_0)$ también esta definida). Puesto que $\varphi(\beta z_0) = 0$, observamos que φ tiene un cero múltiple en βz_0. Pero esto es imposible por el Teorema 3.3.2.

Concluimos que el grado de $A_\beta(u)$ es el número de elementos de R_β. Por el Lema 3.4.7, se sigue que $A_\beta(u) = uP_\beta(u^4)$ tiene grado $N(\beta)$, por lo que $P_\beta(u)$ tiene grado $d = (N(\beta)-1)/4$. Esto prueba la parte (c) para $P_\beta(u)$.

Paso 5: Relacionar P_β y Q_β. Una vez que probemos que

$$Q_\beta(u) = u^d P_\beta(1/u), \text{ con } d = (N(\beta) - 1)/4, \tag{3.42}$$

se seguirá inmediatamente que $Q_\beta(u)$ tiene grado d y que $P_\beta(u)$ es mónico (puesto que $Q_\beta(u)$ tiene término constante 1). Por esta razón solo necesitamos probar (3.42) para completar la prueba.

La identidad (3.24) asegura que

$$\varphi(z)\varphi\left(z + (1+i)\frac{\varpi}{2}\right) = -i = i^3.$$

Definiendo $w = z + (1+i)\frac{\varpi}{2}$, obtenemos que

$$\varphi(z)\varphi(\omega) = i^3. \tag{3.43}$$

De nuevo como en 3.40 se llega a que

$$\varphi(\beta z)\varphi(\beta w) = i^{3+2\varepsilon}. \tag{3.44}$$

Entonces

$$\frac{\varphi(\beta z)}{i^\varepsilon \varphi(z)} = \frac{i^\varepsilon \varphi(w)}{\varphi(\beta w)} = \frac{Q_\beta(\varphi^4(w))}{P_\beta(\varphi^4(w))} = \frac{Q_\beta(1/\varphi^4(z))}{P_\beta(1/\varphi^4(z))}, \tag{3.45}$$

donde la primer igualdad utiliza (3.43) y (3.44), la segunda utiliza (3.39) con w en lugar de z, y la tercera se sigue de elevar (3.43) a la cuarta potencia para obtener $\varphi^4(w) = 1/\varphi^4(z)$. Comparando (3.45) con (3.39), concluimos que

$$\frac{Q_\beta(1/u^4)}{P_\beta(1/u^4)} = \frac{P_\beta(u^4)}{Q_\beta(u^4)}$$

como funciones racionales en u con coeficientes en $\mathbb{Q}(i)$. Por esta razón

$$\frac{Q_\beta(1/u)}{P_\beta(1/u)} = \frac{P_\beta(u)}{Q_\beta(u)}.$$

Recordemos del Paso 4 que el grado$(P_\beta(u)) = d$, donde $d = (N(\beta)-1)/4$. Ahora supongamos que el grado$(Q_\beta(u)) = m$ tal que $m \geq d$ entonces por la ecuación anterior tenemos que $Q_\beta(u)Q_\beta(1/u) = P_\beta(u)P(1/u)$, si suponemos

$$P(u) = a_d u^d + \cdots + a_0 \quad \text{y} \quad Q(u) = b_m u^m + \cdots + b_0$$

entonces operando y factorizando $1/u^d$ y $1/u^m$ respectivamente tenemos que

$$b_0 b_m u^{2m} + \cdots + b_0 b_m = \frac{u^m}{u^d}(a_o a_d u^{2d} + \cdots + a_0 a_d)$$

por lo cual esta igualdad se cumple si y solo si los grados de estos dos polinomios son iguales, y esto se cumple cuando $2m = m - d + 2d$, es decir, cuando $m = d$. El caso $m \leq d$ es análogo al anterior, por lo tanto P_β y Q_β tienen el mismo grado.

Ahora, utilizando nuevamente la igualdad anterior tenemos que $u^d Q_\beta(u) Q_\beta(1/u) = u^d P_\beta(u) P_\beta(1/u)$ entonces si suponemos que u_0 es raíz de $u^d P_\beta(1/u)$ entonces por ser $P_\beta(1/u)$ y $Q_\beta(1/u)$ primos relativos u_0 es raíz de $Q_\beta(u)$. El caso para cual u_0 es raíz de $Q_\beta(u)$ es análogo. Por lo cual se demuestra que

$$u^d P_\beta(1/u) = \lambda Q_\beta(u) \tag{3.46}$$

para alguna constante distinta de cero $\lambda \in \mathbb{Q}(i)$. Sin embargo, si evaluamos (3.39) en $z = \frac{\varpi}{2}$ y utilizamos (3.40) y $\varphi(\frac{\varpi}{2}) = 1$, entonces obtenemos

$$i^\varepsilon = i^\varepsilon \frac{P_\beta(1)}{Q_\beta(1)}.$$

Por esta razón $P_\beta(1) = Q_\beta(1) \neq 0$. Entonces sustituyendo $u = 1$ en (3.46) implica que $\lambda = 1$, por lo tanto $Q_\beta(u) = u^d P_\beta(1/u)$. Esto completa la prueba. □

Mostraremos ahora dos ejemplos del Teorema 3.4.5 del principio del capítulo.

Ejemplo 3.4.8. Cuando $\beta = 3$, la ecuación (3.14) proporciona

$$\varphi(3z) = -\varphi(z)\frac{\varphi^8(z) + 6\varphi^4(z) - 3}{1 + 6\varphi^4(z) - 3\varphi^8(z)}.$$

En la notación del Teorema 3.4.5, esto significa

$$P_3(u) = u^2 + 6u - 3 \quad \text{y} \quad Q_3(u) = u^2 P_3(1/u) = 1 + 6u - 3u^2.$$

Estos polinomios tienen grado $(N(3) - 1)/4 = 2$. Notemos también que $i^\varepsilon = -1$, puesto que $3 \equiv -1 \mod 2(1 + i)$.

Cuando $\beta = 5$, la ecuación (3.16) proporciona

$$\begin{aligned} \varphi(5z) &= \varphi(z)\frac{P_5(\varphi^4(z))}{Q_5(\varphi^4(z))}, \quad \text{donde} \\ P_5(u) &= u^6 + 50u^5 - 125u^4 + 300u^3 - 105u^2 - 62u + 5, \\ Q_5(u) &= 1 + 50u - 125u^2 + 300u^3 - 105u^4 - 62u^5 + 5u^6. \end{aligned}$$

Estos polinomios tienen grado$(N(5) - 1)/4 = 6$ y satisfacen $Q_5(u) = u^6 P_5(u)$. Además, tenemos $i^\varepsilon = 1$, puesto que $5 \equiv 1 \mod 2(1 + i)$.

En general, se puede mostrar que para un entero $n > 0$, los polinomios $P_n(u)$ y $Q_n(u)$ del Teorema 3.4.5 se encuentran en $\mathbb{Z}[u]$.

Para calcular $\varphi(5\alpha)$, podemos usar el hecho de que

$$\varphi(3\alpha + 2\alpha) + \varphi(3\alpha - 2\alpha) = \frac{2\varphi(3\alpha)\varphi'(2\alpha)}{1 + \varphi^2(3\alpha)\varphi^2(2\alpha)}.$$

Simples pero laboriosos cálculos muestran que

$$\varphi(5\alpha) = \varphi \cdot \frac{\varphi^{24} + 50\varphi^{20} - 125\varphi^{16} + 300\varphi^{12} - 105\varphi^8 - 62\varphi^4 + 5}{1 + 50\varphi^4 - 125\varphi^8 + 300\varphi^{12} - 105\varphi^{16} - 62\varphi^{20} + 5\varphi^{24}}. \tag{3.47}$$

Para resolver la ecuación $F_5(\varphi) = 0$ en raíces cuadradas, Gauss usó el hecho de que sobre $\mathbb{Z}[i]$ el número 5 se factoriza en el producto de $2+i$ y $2-i$. Sean $\beta = (2+i)\alpha$ y $\overline{\beta} = (2-i)\alpha$. Entonces

$$\varphi(\beta) = \varphi \cdot \frac{-i\varphi^4 + (2+i)}{1 - (1-2i)\varphi^4} = \psi, \quad \varphi(\overline{\beta}) = \varphi \cdot \frac{i\varphi^4 + (2-i)}{1 - (1+2i)\varphi^4} = \overline{\psi},$$

$$\varphi(5\alpha) = \psi \cdot \frac{i\psi^4 + (2-i)}{1 - (1+2i)\psi^4}, \qquad \varphi(5\alpha) = \overline{\psi} \cdot \frac{-i\overline{\psi}^4 + (2+i)}{1 - (1-2i)\overline{\psi}^4}.$$

Observemos que en (3.47) el numerador es divisible entre los numeradores de las fracciones ψ y $\overline{\psi}$. Dividiendo el numerador de (3.47) por

$$(-i\varphi^4 + (2+i))(i\varphi^4 + (2-i)) = \varphi^8 - 2\varphi^4 + 5,$$

obtenemos el polinomio

$$\varphi^{16} + 52\varphi^{12} - 26\varphi^8 - 12\varphi^4 + 1.$$

La solución de $\varphi(5\alpha) = 0$ (si hacemos caso omiso del caso obvio $\psi = 0$) puede obtenerse, primero, resolviendo la ecuación $i\psi^4 + (2-i) = 0$ y, después, resolviendo la ecuación

$$\varphi \cdot \frac{-i\varphi^4 + (2+i)}{1 - (1-2i)\varphi^4} = \psi = \sqrt[4]{1+2i}. \tag{3.48}$$

Podemos, alternativamente, resolver la ecuación $-i\overline{\psi}^4 + (2+i) = 0$ y entonces la ecuación

$$\varphi \cdot \frac{i\varphi^4 + (2-i)}{1 - (1+2i)\varphi^4} = \overline{\psi} = \sqrt[4]{1-2i}. \tag{3.49}$$

Dividiendo (3.48) entre (3.49) obtenemos una ecuación cuadrática para φ^4.

3.4.3. Multiplicación por primos Gaussianos

Cuando β es un primo impar en $\mathbb{Z}[i]$, el Teorema 3.4.5 tiene el siguiente importante refinamiento debido a Eisenstein. Este resultado jugará un papel central en la prueba del teorema de Abel.

Teorema 3.4.9. *Sea $\beta \in \mathbb{Z}[i]$ un primo impar, y sea*

$$P_\beta(u) = u^d + a_1 u^{d-1} + \cdots + a_d \in \mathbb{Z}[i][u], \quad d = (N(\beta) - 1)/4,$$

el polinomio correspondiente del Teorema 3.4.5. *Entonces:*

(a) *$a_1, \ldots, a_d$ son divisibles entre β y $a_d = i^\varepsilon \beta$, donde $\beta \equiv i^\varepsilon \mod 2(1+i)$.*

(b) *$P_\beta(u)$ es irreducible sobre $\mathbb{Q}(i)$.*

Demostración. Primero observemos que la parte (a) y el Teorema 3.4.3 implican que

$$P_\beta(u^4) = u^{4d} + a_1 u^{4(d-1)} + \cdots + a_d \in \mathbb{Z}[i][u]$$

es un polinomio irreducible sobre $\mathbb{Q}(i)$. Por lo tanto la parte (b) del teorema se sigue de la parte (a).

Probar la parte (a) será más difícil. Puesto que β es impar, el Teorema 3.4.5 implica que

$$\varphi(\beta z) = i^\varepsilon \varphi(z) \frac{\varphi^{4d}(z) + a_1 \varphi^{4(d-1)}(z) + \cdots + a_d}{1 + a_1 \varphi(z) + \cdots + a_d \varphi^{4d}(z)}, \tag{3.50}$$

donde los coeficientes $a_1, \ldots, a_d \in \mathbb{Z}[i]$ dependen de β. Para probar la parte (a), analizaremos la relación entre $a_1, \ldots, a_d$ y β desarrollando cada lado de (3.50) como una serie de potencias en z.

Durante la prueba aparecerán varias series de potencias. La primera proviene de

$$i^\varepsilon \frac{u^d + a_1 u^{d-1} + \cdots + a_d}{1 + a_1 u + \cdots + a_d u^d},$$

la cual escribiremos como

$$i^\varepsilon \frac{u^d + a_1(\beta) u^{d-1} + \cdots + a_d(\beta)}{1 + a_1(\beta) u + \cdots + a_d(\beta) u^d}$$

para enfatizar la dependencia de β. Esta función racional es analítica en $u = 0$ (el denominador no se anula en 0) y de ahí que tenga un desarrollo en serie de potencias

$$\begin{aligned} i^\varepsilon \frac{u^d + a_1(\beta) u^{d-1} + \cdots + a_d(\beta)}{1 + a_1(\beta) u + \cdots + a_d(\beta) u^d} &= \sum_{k=0}^{\infty} b_k(\beta) u^k \\ &= b_0(\beta) + b_1(\beta) u + b_2(\beta) u^2 + \cdots . \end{aligned} \tag{3.51}$$

Notemos que $b_k(\beta) \in \mathbb{Z}[i]$ para toda k ya que $b_k(\beta)$ es la k-ésima derivada de la parte izquierda evaluada en 0 y dividida entre $k!$, La derivación dará una función racional que al evaluar en $u = 0$, tendrá denominador igual a 1 y el denominador de la parte izquierda en $u = 0$ es 1. Utilizando la serie de potencias (3.51), la fórmula de multiplicación (3.50) puede ser escrita como

$$\begin{aligned} \varphi(\beta z) &= \varphi(z)(b_0(\beta) + b_1(\beta)\varphi^4(z) + b_2(\beta)\varphi^8(z) + \cdots) \\ &= b_0(\beta)\varphi(z) + b_1(\beta)\varphi^5(z) + b_2(\beta)\varphi^9(z) + \cdots . \end{aligned} \tag{3.52}$$

Puesto que $\varphi(z)$ es analítica en $z = 0$, esta puede ser desarrollada en una serie de potencias en z alrededor de $z = 0$. Para observar como es esta serie, tomamos la serie de Taylor al rededor del cero dada por

$$\varphi(z) = b_0 + b_1 z^1 + b_2 z^2 + b_3 z^3 + b_4 z^4 + b_5 z5 + \cdots$$

donde $b_k = \varphi^{(k)}(0)/k!$ y utilizamos la igualdad $\varphi(iz) = i\varphi(z)$, de donde podemos observar que esta se cumple solamente si $b_k = 0$ para toda $k \neq 4j + 1$ para toda $j \in \mathbb{N}$. Puesto que $b_0 = 0$ y $b_1 = 1$ podemos reescribir a la serie de potencias como

$$\varphi(z) = z + \sum_{j=1}^{\infty} c_j z^{4j+1} = z + c_1 z^5 + c_2 z^9 + \cdots , \quad c_j \in \mathbb{Q}. \tag{3.53}$$

Tomando en cuenta que $\varphi'^2(z) = 1 - \varphi^4$ y que $\varphi''(z) = -2\varphi^3(z)$, derivando reiteradamente la función $\varphi(z)$ tenemos que $c_1 = -\frac{1}{10}$ y que $c_2 = \frac{1}{120}$. Entonces reemplazando z por βz en (3.53) obtenemos la tercera serie de potencias.

$$\varphi(\beta z) = \sum_{j=0}^{\infty} c_j \beta^{4j+1} z^{4j+1} = \beta z + c_1 \beta^5 z^5 + c_2 \beta 9 z^9 + \cdots . \tag{3.54}$$

De aquí, la prueba procederá en tres pasos. Antes escribiremos una perspectiva general de lo que haremos en cada paso:

- **Paso 1.** Deduciremos una fórmula para $b_k(\beta)$ en términos de β, válida para todo impar $\beta \in \mathbb{Z}[i]$. Esto se seguirá de sustituir las series para $\varphi(z)$ y $\varphi(\beta z)$ en (3.52).

- **Paso 2.** Probaremos que β divide $b_0(\beta), \ldots, b_{d-1}(\beta)$ cuando β es un primo impar. Esto se hará analizando la fórmula del Paso 1 utilizando una inteligente idea de Eisenstein.

- **Paso 3.** Relacionaremos $a_1(\beta), \ldots, a_d(\beta)$ con $b_0(\beta), \ldots, b_{d-1}(\beta)$ y concluiremos que β divide a $a_1(\beta), \ldots, a_d(\beta)$. Esto se seguirá fácilmente de (3.51).

Comencemos ahora con el primer paso.

Paso 1: Expresar $b_k(\beta)$ en términos de β. Si sustituimos (3.53) y (3.54) en la identidad (3.52), entonces obtenemos

$$\begin{aligned} \beta z + c_1\beta^5 z^5 + c_2\beta^9 z^9 + \cdots &= b_0(\beta)(z + c_1 z^5 + c_2 z^9 + \cdots) + \\ &\quad b_1(\beta)(z + c_1 z^5 + c_2 z^9 + \cdots)^5 + \\ &\quad b_2(\beta)(z + c_1 z^5 + c_2 z^9 + \cdots)^9 + \cdots . \end{aligned} \tag{3.55}$$

Cuando desarrollamos el lado derecho de (3.55), una potencia dada de z aparece solamente un número finito de veces, puesto que para toda j la siguiente igualad

$$(z + c_1 z^5 + c_2 z^9 + \cdots)^{4j+1} = z^{4j+1}(1 + c_1 z^4 + c_2 z^8 + \cdots)^{4j+1}$$

garantiza que el menor grado de las potencias de z es mayor o igual a $4j+1$. El lado derecho de (3.55) comienza así:

$$b_0(\beta)z + (b_0(\beta)c_1 + b_1(\beta))z^5 + (b_0(\beta)c_2 + 5b_1(\beta)c_1 + b_2(\beta))z^9 + \cdots . \tag{3.56}$$

Puesto que esto es igual a $\beta z + c_1\beta^5 z^5 + c_2\beta^9 z^9 + \cdots$, comparando coeficientes obtenemos

$$\begin{aligned} \beta &= b_0(\beta), \\ c_1\beta^5 &= b_0(\beta)c_1 + b_1(\beta), \\ c_2\beta^9 &= b_0(\beta)c_2 + 5b_1(\beta)c_1 + b_2(\beta), \ldots \end{aligned}$$

y entonces resolviendo para $b_0(\beta), b_1(\beta), b_2(\beta)$ tenemos

$$\begin{aligned} b_0(\beta) &= \beta, \\ b_1(\beta) &= \beta(c_1\beta^4 - c_1), \\ b_2(\beta) &= \beta(c_2\beta^8 - 5c_1^2\beta^4 + 5c_1^2 - c_2), \ldots \end{aligned}$$

Estas ecuaciones se cumplen para todo impar $\beta \in \mathbb{Z}[i]$. Mas adelante observaremos que $b_0(\beta) = \beta$ es muy importante.

En general, para cualquier k, existe un polinomio $S_k(u) \in \mathbb{Q}[u]$ de grado $4k$ tal que

$$b_k(\beta) = \beta S_k(\beta), \quad \beta \in \mathbb{Z}[i] \text{ impar}. \tag{3.57}$$

Esto se sigue por que todos los términos c_j se encuentran en $\mathbb{Q}$. El punto crucial aquí es que el mismo polinomio $S_k(u)$ funciona para todo β impar. Por ejemplo, como $c_1 = -\frac{1}{10}$, las ecuaciones anteriores implican que

$$b_1(\beta) = \beta S_1(\beta), \quad S_1(u) = -\frac{1}{10}u^4 + \frac{1}{10}.$$

Paso 2: Probar que β divide a $b_0(\beta), \ldots, b_{d-1}(\beta)$ cuando β es un primo impar. La ecuación (3.57) parece implicar que $b_k(\beta)$ es un múltiplo de β para toda $k \geq 0$. El problema es que $S_k(u) \in \mathbb{Q}[u]$ necesariamente no tiene coeficientes enteros, como mostramos para $S_1(u)$. Por esta razón necesitamos estudiar los denominadores de los coeficientes de $S_k(u)$.

Sea s_k el mínimo común múltiplo de estos denominadores. Entonces

$$S_k(u) = \frac{1}{s_k} T_k(u),$$

donde $s_k \in \mathbb{Z}\backslash\{0\}$, $T_k(u) \in \mathbb{Z}[u]$, y ± 1 son los únicos enteros enteros que dividen a s_k y a todos los coeficientes de $T_k(u)$. Eisenstein observó que si $\alpha \in \mathbb{Z}[i]$ es un primo impar, entonces

$$\alpha | s_k \Rightarrow N(\alpha) \leq 4k + 1. \tag{3.58}$$

Para probar esto, primero observo que (3.57) implica que

$$s_k b_k(\beta) = \beta T_k(\beta), \quad \beta \in \mathbb{Z}[i] \text{ impar}. \tag{3.59}$$

Arriba notamos que $b_k(\beta)$ siempre se encuentra en $\mathbb{Z}[i]$. Esto significa que si un primo Gaussiano impar α divide a s_k, entonces α también divide a $\beta T_k(\beta)$. De ahí se sigue que

$$\beta T_k(\beta) \equiv 0 \mod \alpha, \quad \beta \in \mathbb{Z}[i] \text{ impar}. \tag{3.60}$$

Entonces consideremos lo siguiente:

- Puesto que α es impar, la prueba del Lema 3.4.7 muestra que los elementos de $\mathbb{Z}[i]/\alpha\mathbb{Z}[i]$ son de la forma $[\beta]$, β impar. Así que (3.60) implica que la reducción de $uT_k(u)$ módulo α es un polinomio con al menos $|\mathbb{Z}[i]/\alpha\mathbb{Z}[i]|$ raíces.
- Puesto que α divide a s_k, la definición de s_k muestra que la reducción de $uT_k(u)$ modulo α es un polinomio distinto del cero de grado a lo mas $4k+1$. De ahí que que la reducción tenga a lo más $4k + 1$ raíces puesto que $\mathbb{Z}[i]/\alpha\mathbb{Z}[i]$ es un campo por el Lema 3.4.2.

Estos puntos implican que $|\mathbb{Z}[i]/\alpha\mathbb{Z}[i]| \leq 4k + 1$. Sin embargo, $|\mathbb{Z}[i]/\alpha\mathbb{Z}[i]| = N(\alpha)$ por el Lema 3.4.2. Por esta razón $N(\alpha) \leq 4k + 1$, y por lo tanto se cumple (3.58).

Ahora fijemos un primo Gaussiano impar β. Entonces (3.58), aplicado a β, nos dice que

$$N(\beta) > 4k + 1 \Rightarrow \beta \nmid s_k.$$

Notemos que $N(\beta) > 4k + 1$ si y sólo si $k < d = (N(\beta) - 1)/4$. De aquí se sigue $\beta \nmid s_k$ para $k = 0, \ldots, d - 1$. Como β es primo, 3.59 implica que β divide a $b_k(\beta)$ para $k = 0, \ldots, d - 1$. Esto es lo que necesitamos probar.

Paso 3: Relacionar $a_1(\beta), \ldots, a_d(\beta)$ con $b_0(\beta), \ldots, b_{d-1}(\beta)$. Esto es fácil, si escribimos (3.51) de la forma

$$i^{\varepsilon}(u^d + a_1(\beta)u^{d-1} + \cdots + a_d(\beta)) = (1 + a_1(\beta)u + \cdots + a_d(\beta)u^d) \left(\sum_{k=0}^{\infty} b_k(\beta)u^k \right)$$

y al hacer la multiplicación del lado derecho, y comparando los coeficientes de las potencias de u obtenemos las ecuaciones

$$\begin{aligned} i^{\varepsilon}a_d(\beta) &= b_0(\beta), \\ i^{\varepsilon}a_{d-1}(\beta) &= a_1(\beta)b_0(\beta) + b_1(\beta), \\ i^{\varepsilon}a_{d-2}(\beta) &= a_2(\beta)b_0(\beta) + a_1(\beta)b_1(\beta) + b_2(\beta), \\ &\vdots \\ i^{\varepsilon}a_1(\beta) &= a_{d-1}(\beta)b_0(\beta) + a_{d-2}(\beta)b_1(\beta) + \cdots + b_{d-1}(\beta). \end{aligned}$$

Los $a_j(\beta)$ se encuentran en $\mathbb{Z}[i]$, y $b_0(\beta), \cdots, b_{d-1}(\beta)$ son divisibles entre β por el Paso 2. De aquí se sigue que en las ecuaciones anteriores, el lado derecho siempre es divisible entre β. Esto muestra que β divide a $a_1(\beta), \ldots, a_d(\beta)$, puesto que i^{ε} es una unidad. Además, antes probamos que $b_0(\beta) = \beta$, por lo tanto la primera ecuación implica que $a_d(\beta) = i^{-\varepsilon}\beta$. Esto completa la prueba de la parte (a). □

Nota Matemática

Mencionemos ahora algunos comentarios más acerca de la multiplicación compleja.

• **Multiplicación Compleja.** En nuestra discusión de las funciones elípticas en la Sección 3.3, mencionamos que la función $\wp$ de Weierstrass $\wp(z;\omega_1,\omega_2)$ para periodos ω_1, ω_2 tiene una ley de adición. De aquí se sigue fácilmente que esta función también satisface las fórmulas de multiplicación para $n \in \mathbb{Z}$ que generaliza el Teorema 3.2.5. Sin embargo, la función $\wp$ pocas veces tiene multiplicación compleja. Más preciso, $\wp(z;\omega_1,\omega_2)$ tiene multiplicación compleja para algunos $\beta \in \mathbb{C}\backslash\mathbb{S}$ si y sólo si ω_2/ω_1 es una raíz de un polinomio cuadrático con coeficientes enteros. Esto significa que ω_2/ω_1 se encuentra en un ***campo imaginario cuadrático***, el cual es un campo de la forma $\mathbb{Q}(\sqrt{-m})$ para alguna $m > 0$ en $\mathbb{Z}$. Por ejemplo, para los periodos $\omega_1 = (1-i)\varpi$, $\omega_2 = (1+i)\varpi$ de la función de Abel $\varphi(z)$ se tiene razón

$$\frac{\omega_2}{\omega_1} = \frac{(1+i)\varpi}{(1-i)\varpi} = i,$$

la cual es raíz de $x^2 + 1 = 0$. Por lo tanto el campo imaginario cuadrático asociado es $\mathbb{Q}(i)$. En general, las funciones elípticas con multiplicación compleja tienen una profunda relación con los campos imaginarios cuadráticos.

3.5. Teorema de Abel

En esta sección, probaremos el teorema de Abel acerca de las construcciones con regla y compás de los puntos que dividen a la lemniscata en n partes iguales. Las herramientas que utilizaremos incluirán teoría de Galois y la teoría de la multiplicación compleja desarrollada en la Sección 3.4.

3.5.1. El grupo de Galois lemniscático

Sea n un entero positivo impar y consideremos

$$L = \mathbb{Q}\left(i, \varphi\left(\frac{\varpi}{n}\right)\right).$$

Observaremos que el grupo de Galois de $\mathbb{Q}(i) \subset L$ involucra al grupo

$$(\mathbb{Z}[i]/n\mathbb{Z}[i])^*$$

de las unidades en $\mathbb{Z}[i]/n\mathbb{Z}[i]$. Puesto que $\mathbb{Z}[i]$ es un dominio de ideales principales, tenemos que α es primo relativo a n en $\mathbb{Z}[i]$

$$\begin{array}{ll} \iff & \text{existen } a, b \in \mathbb{Z}[i] \text{ tales que } a\alpha + bn = 1 \\ \iff & a\alpha \equiv 1 \mod n\mathbb{Z}[i] \\ \iff & [\alpha] \in (\mathbb{Z}[i]/n\mathbb{Z}[i])^* \end{array}$$

Teorema 3.5.1. *$\mathbb{Q}(i) \subset L$ es una extensión de Galois y existe un homomorfismo inyectivo de grupos*

$$\mathrm{Gal}(L/\mathbb{Q}(i)) \to (\mathbb{Z}[i]/n\mathbb{Z}[i])^*.$$

En particular, $\mathrm{Gal}(L/\mathbb{Q}(i))$ *es Abeliano.*

Demostración. Sea $A_n(u) = uP_n(u^4)$ el n-polinomio de división definido en la parte (d) del Teorema 3.4.5. El teorema nos dice que las raíces de $A_n(u)$ están dadas por

$$\varphi\left(\alpha\frac{\varpi}{n}\right), \quad \alpha \in \mathbb{Z}[i] \text{ impar} \tag{3.61}$$

y la prueba del Lema 3.4.7 muestra que para cada raíz, la $\alpha \in \mathbb{Z}[i]$ asociada es única módulo $n\mathbb{Z}[i]$.

Puesto que cada α en (3.61) es impar, la fórmula de la multiplicación compleja para $\varphi(\alpha z)$ dada por el Teorema 3.4.5 muestra que $\varphi(\alpha\frac{\varpi}{n})$ es una función racional en $\varphi(\frac{\varpi}{n})$ con coeficientes en $\mathbb{Q}(i)$, es decir, $\varphi(\alpha\frac{\varpi}{n}) \in \mathbb{Q}\left(i, \varphi(\frac{\varpi}{n})\right) = L$. De aquí se sigue que $A_n(u)$ se descompone completamente en $L = \mathbb{Q}(i, \varphi(\frac{\varpi}{n}))$. Puesto que una de la raíces es $\varphi(\frac{\varpi}{n})$, se sigue inmediatamente que L es el campo de descomposición de $A_n(u)$ sobre $\mathbb{Q}(i)$. Por lo tanto $\mathbb{Q}(i) \subset L$ es una extensión de Galois.

Ahora tomemos $\sigma \in \mathrm{Gal}(L/\mathbb{Q}(i))$. Puesto que $0 \in \mathbb{Q}(i)$ y que $A_n(u) \in \mathbb{Z}[i][u]$ tenemos que

$$0 = \sigma(0) = \sigma\left(A_n\left(\varphi\left(\frac{\varpi}{n}\right)\right)\right) = A_n\left(\sigma\left(\varphi\left(\frac{\varpi}{n}\right)\right)\right)$$

Entonces $\sigma(\varphi(\frac{\varpi}{n}))$ es una raíz de $A_n(u)$ y por lo cual es uno de los números (3.61). Por lo tanto existe $\alpha \in \mathbb{Z}[i]$ impar tal que

$$\sigma\left(\varphi\left(\frac{\varpi}{n}\right)\right) = \varphi\left(\alpha\frac{\varpi}{n}\right). \tag{3.62}$$

Como notamos antes, α es única modulo $n\mathbb{Z}[i]$.

Utilizando el hecho anterior y el Teorema 3.4.5 podemos mostrar que si $\beta \in \mathbb{Z}[i]$ es impar, entonces

$$\begin{aligned} \sigma\left(\varphi\left(\beta\tfrac{\varpi}{n}\right)\right) &= \sigma(i^{\varepsilon})\sigma\left(\varphi\left(\frac{\varpi}{n}\right)\right)\frac{P_\beta(\sigma(\varphi^4(\frac{\varpi}{n})))}{Q_\beta(\sigma(\varphi^4(\frac{\varpi}{n})))} \\ &= i^{\varepsilon}\varphi\left(\alpha\frac{\varpi}{n}\right)\frac{P_\beta(\varphi^4(\alpha\frac{\varpi}{n}))}{P_\beta(\varphi^4(\alpha\frac{\varpi}{n}))} \\ &= \varphi\left(\alpha\beta\tfrac{\varpi}{n}\right). \end{aligned} \tag{3.63}$$

Ahora probaremos que α es primo relativo con n. Sea m el orden de σ en $\mathrm{Gal}(L/\mathbb{Q}(i))$, luego es σ^m es la identidad. Entonces aplicando repetidamente (3.63) obtenemos

$$\varphi\left(\frac{\varpi}{n}\right) = \sigma^m\left(\varphi\left(\frac{\varpi}{n}\right)\right) = \varphi\left(\alpha^m\frac{\varpi}{n}\right).$$

Por unicidad, concluimos que

$$1 \equiv \alpha^m \mod n.$$

De ahí que α sea primo relativo con n en $\mathbb{Z}[i]$, por lo que $[\alpha]$ es unidad de $\mathbb{Z}[i]/n\mathbb{Z}[i]$, de tal forma que $\sigma \mapsto [\alpha]$ proporciona una función bien definida

$$\mathrm{Gal}(L/\mathbb{Q}(i)) \to (\mathbb{Z}[i]/n\mathbb{Z}[i])^*. \tag{3.64}$$

Si σ y τ son mandados a α y β respectivamente, entonces (3.63) fácilmente implica que $\sigma\tau(\varphi(\frac{\varpi}{n})) = \varphi(\alpha\beta\frac{\varpi}{n})$. Por lo tanto $\sigma\tau$ es mandado a $\alpha\beta$, lo cual muestra que la función es un homomorfismo de grupos. Además, si $[\alpha] = [\beta]$ en $(\mathbb{Z}[i]/n\mathbb{Z}[i])^*$, entonces

$$\alpha = \beta + (a + ib)n$$

donde $a + ib$ es par por que α, β y n son impares. Entonces la Proposición 3.3.1 implica que

$$\sigma\left(\varphi\left(\frac{\varpi}{n}\right)\right) = \varphi\left(\alpha\frac{\varpi}{n}\right) = \varphi\left(\beta\frac{\varpi}{n}\right) = \tau\left(\varphi\left(\frac{\varpi}{n}\right)\right),$$

de lo cual concluimos que la función es inyectiva puesto que $\varphi(\frac{\varpi}{n})$ genera L sobre $\mathbb{Q}(i)$. Esto completa la prueba. □

Ahora, gracias al Teorema 3.5.1 podemos mostrar que los n-puntos de división de la lemniscata se pueden expresar por radicales sobre $\mathbb{Q}$. Veamos esto, por el teorema sabemos que la extensión $\mathbb{Q}(i) \subset L$ es de Galois y que el grupo de Galois $\mathrm{Gal}(L/\mathbb{Q}(i))$ es Abeliano, por estos hechos junto con los Teoremas A.1.16 y A.1.17 sabemos que por ser el grupo Abeliano es soluble y que esto implica que la extensión $\mathbb{Q}(i) \subset L$ es soluble. Por la definición de extensión soluble sabemos que existe una extensión de campo $L \subset M$ tal que la extensión $\mathbb{Q}(i) \subset M$ es radical, puesto que la extensión de campo $\mathbb{Q} \subset \mathbb{Q}(i)$ también es radical, por el Teorema A.1.20 sabemos que la extensión $\mathbb{Q} \subset M$ también es radical. Finalmente como

sabemos que los n-puntos de división se encuentran en L y por lo tanto en M y que son raíces de los n-polinomios de división que se encuentran en $\mathbb{Q}[x]$ tenemos que los n-puntos de división son expresables por radicales.

El homomorfismo (3.64) construido en el Teorema 3.5.1 es el análogo lemniscático del homomorfismo

$$\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^*.$$

3.5.2. Construcciones con regla y compás

Teorema 3.5.2. *Sea n un entero positivo. Entonces los siguientes enunciados son equivalentes:*

(a) *Los n-puntos de división de la lemniscata pueden ser construidos usando regla y compás.*

(b) *$\varphi(\frac{2\varpi}{n})$ es construible.*

(c) *n es un entero de la forma*

$$n = 2^s p_1 \cdots p_r,$$

donde $r, s \geq 0$ son enteros y $p_1 \cdots p_r$ son primos de Fermat distintos.

Demostración. La implicación $(a) \Rightarrow (b)$ es fácil, puesto que $\varphi(\frac{2\varpi}{n})$ es la distancia polar de un n-punto de división. El recíproco $(b) \Rightarrow (a)$ se sigue de la parte (b) del Corolario 3.2.7.

La prueba de $(c) \Rightarrow (b)$ será una aplicación del Teorema 3.5.1 junto con algunos resultados de la Sección 3.2. Primero observemos por la parte (c) del Corolario 3.2.7 que bastará ver, para que $\varphi(\frac{2\varpi}{n})$ sea construible, que

$$\varphi\left(\frac{2\varpi}{2^s}\right), \varphi\left(\frac{2\varpi}{p_1}\right), \ldots, \varphi\left(\frac{2\varpi}{p_r}\right) \text{ son construibles.}$$

Pero como $\varphi(\frac{2\varpi}{2^s})$ es construible por la Proposición 3.2.3, así solamente necesitamos mostrar que $\varphi(\frac{2\varpi}{p})$ es construible cuando p es un primo de Fermat.

Por la parte (a) del Corolario 3.2.7, $\varphi(\frac{2\varpi}{p})$ es construible cuando $\varphi(\frac{\varpi}{p})$ lo es. Para esto último, tenemos que el Teorema 3.5.1 proporciona una extensión de Galois $\mathbb{Q}(i) \subset L = \mathbb{Q}(i, \varphi(\frac{\varpi}{p}))$ con

$$\mathrm{Gal}(L/\mathbb{Q}(i)) \simeq \text{un subgrupo de } (\mathbb{Z}[i]/p\mathbb{Z}[i])^*.$$

Primero mostraremos que si

$$|(\mathbb{Z}[i]/p\mathbb{Z}[i])^*| = \text{una potencia de 2}, \tag{3.65}$$

entonces $\varphi(\frac{\varpi}{p})$ es construible. Para esto recurriremos a algunos teoremas del apéndice. Sea $H = \mathrm{Gal}(L/\mathbb{Q}(i)) \subset (\mathbb{Z}[i]/p\mathbb{Z}[i])^*$ el subgrupo del Teorema 3.5.1, entonces por el Lema 3.4.2

(b) sabemos que $(\mathbb{Z}[i]/p\mathbb{Z}[i])^*$ es finito, entonces por el Teorema A.1.21 sabemos que $|H|$ divide a $|(\mathbb{Z}[i]/p\mathbb{Z}[i])^*|$, por lo cual el $|H|$ también es una potencia del 2, sea $s \geq 0$ tal que $|H| = 2^s$. Ahora, por el isomorfismo dado en el Teorema 3.5.1 y puesto que la extensión $\mathbb{Q}(i) \subset L$ es de Galois, por el Teorema A.1.22 tenemos que

$$2^s = |H| = |\operatorname{Gal}(L/\mathbb{Q}(i))| = [L : \mathbb{Q}(i)].$$

Como el grupo $\operatorname{Gal}(L/\mathbb{Q}(i))$ es Abeliano sabemos que es soluble entonces por definición tenemos que existen subgrupos

$$\{e\} = G_n \subset G_{n-1} \subset \cdots \subset G_1 \subset G_0 = \operatorname{Gal}(L/\mathbb{Q}(i)),$$

tal que $\{e\}$ es la identidad en el grupo de Galois, G_i es normal en G_{i-1} y $[G_{i-1} : G_i]$ es primo para toda $i = 1, \ldots, n$. Dado que estamos suponiendo que el orden del grupo de Galois es una potencia de 2, entonces tenemos que $[G_{i-1} : G_i] = 2$ para toda $i = 1, \ldots, n$. Por otra parte sabemos que $L_{G_n} = L_{\{e\}} = L$ y que por ser $\mathbb{Q}(i) \subset L$ una extensión de Galois por definición tenemos que L es un campo de descomposición de un polinomio separable en $\mathbb{Q}(i)[u]$ lo cual por el Teorema A.1.23 es equivalente a que $L_{\operatorname{Gal}(L/\mathbb{Q}(i))} = \mathbb{Q}(i)$. Por esto último, tomando los campos fijos de los subgrupos del grupo de Galois obtenemos los subcampos

$$\mathbb{Q}(i) = L_{G_0} \subset L_{G_1} \subset \cdots \subset L_{G_{n-1}} \subset L_{G_n} = L,$$

dado que la extensión $\mathbb{Q}(i) \subset L$ es de Galois por la Proposición A.1.24 tenemos que $L_{G_i} \subset L$ es una extensión de Galois para toda $i = 0, \ldots, n-1$, este resultado junto con el hecho de que $G_i \subset G_0 = \operatorname{Gal}(L/\mathbb{Q}(i))$ es subgrupo para toda $i = 1, \ldots, n$ nos proporciona por el Teorema A.1.25 (b) que

$$\operatorname{Gal}(L/L_{G_i}) = G_i \text{ para toda } i = 0, \ldots, n$$

y que

$$[L_{G_i} : L_{G_{i-1}}] = [\operatorname{Gal}(L/L_{G_{i-1}}) : G_i] = [G_{i-1} : G_i] = 2 \text{ para } i = 1, \ldots, n.$$

Finalmente, puesto que $\varphi(\frac{\varpi}{p}) \in L$ y que sabemos que $\mathbb{Q}$ es un subcampo de $\mathbb{Q}(i)$ tal que $[\mathbb{Q}(i) : \mathbb{Q}] = 2$, agregando este subcampo a los subcampos anteriores tenemos por el Teorema A.2.3 que $\varphi(\frac{\varpi}{p})$ es un número construible.

Mostraremos ahora que (3.65) se cumple cuando $p = 2^{2^m} + 1$ es un primo de Fermat. Para el caso cuando $p = 3$ por el Lema 3.4.2 (b) sabemos que

$$\mathbb{Z}[i]/3\mathbb{Z}[i] \simeq F_{N(3)} = F_9,$$

por la definición de campo sabemos que todo elemento distinto del cero en F_9 tiene inverso multiplicativo, y por último sabemos que los isomorfismo mandan la identidad en la identidad, por lo tanto

$$|(\mathbb{Z}[i]/3\mathbb{Z}[i])^*| = |F_9^*| = 8.$$

Ahora, si $p > 3$, entonces $m \geq 1$, por lo tanto

$$p = 2^{2^m} + 1 = \left(2^{2^{m-1}} + i\right)\left(2^{2^{m-1}} - i\right) = \beta\overline{\beta},$$

donde $\beta, \overline{\beta}$ son primos no asociados en $\mathbb{Z}[i]$ y tales que $N(\beta) = N(\overline{\beta}) = \beta\overline{\beta} = p$, en este caso, el Teorema chino del residuo para $\mathbb{Z}[i]$ y el Lema 3.4.2 proporcionan los isomorfismos

$$\mathbb{Z}[i]/p\mathbb{Z}[i] = \mathbb{Z}[i]/\beta\overline{\beta}\mathbb{Z}[i] \simeq \mathbb{Z}[i]/\beta\mathbb{Z}[i] \times \mathbb{Z}[i]/\overline{\beta}\mathbb{Z}[i] \simeq \mathbb{F}_p \times \mathbb{F}_p.$$

De ahí que

$$|(\mathbb{Z}[i]/p\mathbb{Z}[i])^*| = |\mathbb{F}_p^* \times \mathbb{F}_p^*| = (p-1)^2 = 2^{2^m} \cdot 2^{2^m} = 2^{2^{m+1}}.$$

Esto muestra (3.65) para todos los primos de Fermat y completa la demostración de la afirmación $(c) \Rightarrow (b)$.

Falta probar $(b) \Rightarrow (c)$. Aquí es donde usaremos el resultado de irreducibilidad probado en el Teorema 3.4.9. Sea n un entero tal que $\varphi(\frac{2\varpi}{n})$ es construible. Podemos suponer que $n > 1$ puesto que el teorema es trivial para cuando $n = 1$. Además, la fórmula de duplicación (3.11) implica que podemos asumir que n es impar, ya que duplicando repetidas veces podemos eliminar el factor 2^s en n, y la Proposición 3.2.3 muestra que $\varphi(\frac{\varpi}{n})$ es construible.

Sea p un primo que divida a n. Entonces p es impar puesto que n lo es. Sea β un primo complejo tal que $p = \beta$ si $p \equiv 3 \mod 4$ y $p = \beta\overline{\beta}$ si $p \equiv 1 \mod 4$. Entonces $n/\beta \in \mathbb{Z}[i]$ es impar (puesto que n y β lo son), por lo tanto $\frac{\varpi}{\beta} = \frac{n}{\beta}\frac{\varpi}{n}$ es un múltiplo impar de $\frac{\varpi}{n}$ y entonces el Teorema 3.4.5 garantiza que,

$$\varphi\left(\frac{\varpi}{\beta}\right) \in \mathbb{Q}\left(i, \varphi\left(\frac{\varpi}{n}\right)\right).$$

De aquí se sigue que $\varphi(\frac{\varpi}{\beta})$ es construible, puesto que i y $\varphi(\frac{\varpi}{n})$ lo son. Por lo tanto por el corolario A.2.4 sabemos que el polinomio mínimo de $\varphi(\frac{\varpi}{\beta})$ sobre $\mathbb{Q}$ tiene grado igual a una potencia de 2. Entonces el Teorema de la torre[2] muestra que el polinomio mínimo de $\varphi(\frac{\varpi}{\beta})$ sobre $\mathbb{Q}(i)$ también tiene grado igual a una potencia de 2.

El Teorema 3.4.5 implica que $\varphi(\frac{\varpi}{\beta})$ es una raíz del polinomio de $uP_\beta(u^4)$. Es fácil mostrar que $\varphi(\frac{\varpi}{\beta}) \neq 0$, primero en el caso cuando $\beta = p$ para algún primo de la forma $4k+3$ sabemos por la definición de $\varphi(z)$ que para todo $n \in \mathbb{N}$ se tiene que $\varphi(\frac{\varpi}{n}) \neq 0$ en particular para p, para el caso cuando β es tal que $\beta\overline{\beta} = p$ para algún primo de la forma $4k+1$ tenemos que $\frac{\varpi}{\beta} = \overline{\beta}\frac{\varpi}{\beta\overline{\beta}} = \overline{\beta}\frac{\varpi}{p}$ es un múltiplo impar, entonces por el Teorema 3.4.5 sabemos que

$$\varphi\left(\frac{\varpi}{\beta}\right) = \varphi\left(\overline{\beta}\frac{\varpi}{p}\right) = i^\varepsilon \varphi\left(\frac{\varpi}{p}\right) \frac{P_{\overline{\beta}}(\varphi^4(\frac{\varpi}{p}))}{Q_{\overline{\beta}}(\varphi^4(\frac{\varpi}{p}))},$$

y por el mismo teorema sabemos $\varphi(\frac{\varpi}{\beta})$ es cero cuando $\frac{\varphi}{p} = \alpha\frac{\varpi}{\beta}$, es decir, cuando $\alpha\beta = 1$, pero esto no puede pasar puesto que β es primo, por lo tanto $\varphi(\frac{\varpi}{\beta})$ es una raíz de $P_\beta(u^4)$.

[2] **Teorema de la torre.**
Suponiendo que tenemos campos $F \subset K \subset L$.

(a) Si $[K : F] = \infty$ ó $[L : K] = \infty$, entonces $[L : F] = \infty$.

(b) Si $[K : F] \leq \infty$ y $[L : K] \leq \infty$, entonces $[L : F] = [L : K][K : F]$.

Puesto que β es un primo impar, por el Teorema 3.4.5 $P_\beta(u^4)$ tiene grado $N(\beta)-1$ y por el Teorema 3.4.9 este polinomio es irreducible sobre $\mathbb{Q}(i)$. Esto prueba que el polinomio mínimo de $\varphi(\frac{\varpi}{\beta})$ sobre $\mathbb{Q}(i)$ tiene grado $N(\beta)-1$.

Cuando $p=\beta$ para $p\equiv 3 \mod 4$, tenemos que $N(\beta)-1=p^2-1=(p+1)(p-1)$. Observamos que este producto es una potencia de 2 si y sólo si $p=3$. Por otra parte, cuando $p=\beta\overline{\beta}$ para $p\equiv 1 \mod 4$, tenemos que $N(\beta)-1=p-1$, lo cual es una potencia de 2 si y sólo si p es un primo de Fermat.

Por esta razón los únicos primos que dividen a n son los primos de Fermat. Para completar la prueba del teorema, necesitamos mostrar que no puede ocurrir que $p^2|n$. Por lo tanto asumamos que $p^2|n$, donde p es primo. Entonces existe un primo complejo impar β tal que $\beta^2|n$, igual que antes puesto que β^2 es impar y n también lo es tenemos que $n/\beta^2\in\mathbb{Z}[i]$ también es impar, de ahí se tiene que $\frac{\varpi}{\beta^2}=\frac{n}{\beta^2}\frac{\varpi}{n}$ es un múltiplo impar por lo tanto nuevamente el Teorema 3.4.5 nos garantiza que,

$$u_0=\varphi\left(\frac{\varpi}{\beta^2}\right)\in L=\mathbb{Q}\left(i,\varphi\left(\frac{\varpi}{n}\right)\right),$$

lo cual implica como antes que u_0 es construible. De ahí que el grado del polinomio mínimo sobre $\mathbb{Q}(i)$ sea una potencia de 2. Probaremos que el polinomio mínimo tiene grado $N(\beta)(N(\beta)-1)$. Observemos que este producto no puede ser una potencia de 2 puesto que $N(\beta)=p$ ó p^2.

Puesto que β es impar, el Teorema 3.4.5 implica que

$$\varphi\left(\frac{\varpi}{\beta}\right)=\varphi\left(\beta\frac{\varpi}{\beta^2}\right)=i^\varepsilon\varphi\left(\frac{\varpi}{\beta^2}\right)\frac{P_\beta\left(\varphi^4\left(\frac{\varpi}{\beta^2}\right)\right)}{Q_\beta\left(\varphi^4\left(\frac{\varpi}{\beta^2}\right)\right)}=i^\varepsilon u_0\frac{P_\beta(u_0^4)}{Q_\beta(u_0^4)}.$$

Puesto que $\varphi(\frac{\varpi}{\beta})$ es una raíz de $P_\beta(u^4)$, esta fórmula para $\varphi(\frac{\varpi}{\beta})$ proporciona la ecuación

$$0=P_\beta\left(\varphi^4\left(\frac{\varpi}{\beta}\right)\right)=P_\beta\left(u_0^4\frac{P_\beta(u_0^4)^4}{Q_\beta(u_0^4)^4}\right)=0.$$

Si escribimos $P_\beta(u)=u^d+a_1u^{d-1}+\cdots+a_d$, $d=(N(\beta)-1)/4$, entonces despejando el denominador en la ecuación anterior mostramos que u_0 es una raíz de

$$P(u)=u^{4d}P_\beta(u^4)^{4d}+a_1u^{4d-4}P_\beta(u^4)^{4d-4}Q_\beta(u^4)^4+\cdots+a_dQ_\beta(u^4)^{4d}.$$

Este polinomio tiene coeficientes en $\mathbb{Z}[i]$ y grado $4d(4d-1)=N(\beta)(N(\beta)-1)$, puesto que $P_\beta(u)$, $Q_\beta(u)\in\mathbb{Z}[i][u]$ tienen grado d. Además, el Teorema 3.4.9 implica que

$$\beta \quad \text{divide a} \quad a_1,\ldots,a_d. \tag{3.66}$$

Por esta razón $P_\beta(u)\equiv u^d \mod \beta$. Utilizando esto y (3.66), observamos que

$$P(u)\equiv u^{N(\beta)(N(\beta)-1)} \mod \beta,$$

puesto que $4d(4d+1) = N(\beta)(N(\beta)-1)$. Además, $Q_\beta(0) = 1$ por el Teorema 3.4.5, por lo tanto el termino constante de $P(u)$ es

$$P(0) = 0 + \cdots + 0 + a_d Q_\beta(0)^{4d} = a_d.$$

El teorema 3.4.9 muestra que a_d no es divisible entre β^2, por lo tanto por el criterio de Schönemann-Eisenstein (Teorema 3.4.3) sobre $\mathbb{Q}(i)$, $P(u)$ es irreducible sobre $\mathbb{Q}(i)$. Por esta razón el polinomio mínimo de u_0 sobre $\mathbb{Q}(i)$ tiene grado $N(\beta)(N(\beta)-1)$. La prueba esta completada. □

3.5.3. Otra demostración del teorema.

En esta sección daremos otra demostración del teorema de Abel propuesta por ***Rosen***. La piedra angular de esta demostración es, al igual que en la demostración original de Abel, el hecho de que la latiz de periodos de la función $\varphi(\alpha)$ es invariante con respecto a la multiplicación por la unidad compleja i.

En la demostración de Rosen, una función de Weierstrass $\wp(z)$ es usada en lugar de la función lemniscática $\varphi(\alpha)$. La función de Weierstrass usada corresponde a la latiz $\Omega = \{2a\varpi + 2bi\varpi \mid a, b \in \mathbb{Z}\}$. Observemos que esta latiz está contenida en la latiz de periodos de φ pero no coincide con esta. Al final de esta sección mostraremos que para Ω tenemos $g_2 = \frac{1}{4}$ y $g_3 = 0$, y entonces tal $\wp$ satisface:

$$\wp'^2(z) = 4\wp^3(z) - \frac{1}{4}\wp(z).$$

La justificación de utilizar la función $\wp(z)$ está relacionada con la siguiente proposición.

Proposición 3.5.3. *Si un segmento de longitud $\wp(\alpha)$ puede ser construido con regla y compás, entonces el segmento de longitud $\varphi(\alpha)$ también puede ser construido con regla y compás.*

Demostración. Módulo Ω, los ceros de φ son de la forma 0, ϖ, $i\varpi$, $(1+i)\varpi$ y sus polos son de la forma $\frac{(1+i)\varpi}{2}$, $\frac{(3+i)\varpi}{2}$, $\frac{(1+3i)\varpi}{2}$, $\frac{(3+3i)\varpi}{2}$. La función $\wp'(z)$ también tiene ceros ϖ, $i\varpi$, $(1+i)\varpi$, recuerde que estos son los puntos críticos de $\wp$, mientras que 0 es un polo de $\wp'(z)$. Además,

$$\wp\left(\frac{1+i}{2}\varpi\right) = \wp\left(\frac{3+3i}{2}\varpi\right) \quad \text{y} \quad \wp\left(\frac{3+i}{2}\varpi\right) = \wp\left(\frac{1+3i}{2}\varpi\right).$$

Consideremos la función

$$g(z) = \frac{\wp'(z)}{(\wp(z) - \wp(\frac{1+i}{2}\varpi))(\wp(z) - \wp(\frac{3+i}{2}\varpi))}.$$

En una vecindad de cero tenemos $\wp(z) = z^{-2} + \cdots$ y $\wp'(z) = -2z^{-3} + \cdots$; por esta razón, $g(0) = 0$. De aquí se sigue que g tiene los mismos ceros y polos que φ. Por lo que, $\varphi(z) = Cg(z)$.

Ahora probaremos que es posible construir segmentos de longitud $\wp(\frac{\varpi}{2}), \wp(\frac{1+i}{2}\varpi)$ y $\wp(\frac{3+i}{2}\varpi)$. Primero que todo, observemos que los segmentos de longitud $\wp(\varpi)$, $\wp(i\varpi)$ y $\wp((1+i)\varpi)$ pueden ser construidos. Puesto que ϖ, $i\varpi$, $(1+i)\varpi$ son ceros de $\wp'$ y como $\wp'^2 = 4\wp^3 - \frac{1}{4}\wp$ entonces $\wp(\varpi)$, $\wp(i\varpi)$ y $\wp((1+i)\varpi)$ son raíces de la ecuación $4x^3 - \frac{1}{4}x = 0$, por lo tanto son construibles.

Observemos que de la definición de la función $\wp$ (2.2) del Capítulo 2 y de la invariancia de la latiz Ω al multiplicarse por i (es decir, $\Omega = i\Omega$) tenemos

$$\wp(-z) = \wp(z) \quad \text{y} \quad \wp(iz) = -\wp(z).$$

Derivando estas fórmulas obtenemos

$$\wp'(-z) = -\wp'(z) \quad \text{y} \quad \wp'(iz) = i\wp'(z).$$

Utilizando estas propiedades de la función $\wp$ junto con la fórmula de adición (2.5) del Capítulo 2 podemos verificar que

$$\begin{aligned}
\wp(z \pm iz) &= -\wp(z) - \wp(\pm iz) + \frac{1}{4}\left(\frac{\wp'(z) - \wp'(\pm iz)}{\wp(z) - \wp(iz)}\right)^2 \\
&= \frac{1}{4}\left(\frac{\wp'(z) \mp i\wp'(z)}{\wp(z) + \wp(z)}\right)^2 \\
&= \frac{1}{4}\frac{(1 \mp i)^2\wp'^2(z)}{4\wp^2(z)} \\
&= \mp\frac{i}{8}\frac{4\wp^2 - \frac{1}{4}}{\wp} \\
&= \mp\frac{i}{8}\frac{4\wp^2(z) - \frac{1}{4}}{\wp(z)}.
\end{aligned}$$

Evaluando las fórmulas anteriores en $z = \frac{\alpha}{(1+i)}$ para el caso positivo y $z = \frac{x}{(1-i)} = \frac{\alpha}{2}$ para el caso negativo tenemos que

$$\wp(\alpha) = \wp((1+i)x) = -\frac{i}{8}\frac{4\wp^2(x) - \frac{1}{4}}{\wp(x)}$$

y que

$$\wp(x) = \wp\left((1-i)\frac{\alpha}{2}\right) = \frac{i}{8}\frac{4\wp^2(\frac{\alpha}{2}) - \frac{1}{4}}{\wp(\frac{\alpha}{2})}$$

de esto último tenemos que si $\wp(\alpha)$ es construible entonces también lo es $\wp(x)$, pero entonces por la segunda igualdad también lo es $\wp(\frac{\alpha}{2})$. Notemos que hemos demostrado: $\wp(\alpha)$ implica $\wp(\frac{\alpha}{2})$ construible.

Por lo tanto puesto que los segmentos de longitud $\wp(\varpi)$, $\wp(i\varpi)$ y $\wp((1+i)\varpi)$ son construibles entonces también lo son los segmentos $\wp(\frac{\varpi}{2}), \wp(\frac{1+i}{2}\varpi)$ y $\wp(\frac{3+i}{2}\varpi)$. Con esto

y por el hecho de que si es posible construir $\wp(\alpha)$, entonces también es posible construir $\wp'(\alpha) = \sqrt{4\wp^3(\alpha) - \frac{1}{4}\wp(\alpha)}$ podemos asegurar que $g(\frac{\varpi}{2})$ es un número construible y puesto que $\varphi(\frac{\varpi}{2}) = 1$, tenemos que la constante C también es construible. Como resultado, de la relación $\varphi(z) = Cg(z)$ tenemos que si es posible construir $g(\alpha)$ también es posible construir $\varphi(\alpha)$. □

Por lo tanto, para probar el teorema de Abel, debemos verificar que si $n = 2^\alpha p_1 \cdots p_m$, donde los p_i son primos de Fermat distintos, entonces los segmentos de longitud $\wp(\frac{k\varpi}{n})$, donde $k = 1, \ldots, n-1$, pueden construirse.

La función $z \mapsto (\wp(z), \wp'(z))$ puede ser considerada como un homeomorfismo del toro $\mathbb{C}/\Omega$ a la curva E definida por la ecuación $y^2 = 4x^3 - \frac{1}{4}x$. La adición de puntos en el toro induce bajo esta función una adición de puntos en E. Los elementos del grupo E cuyo orden divide a n forman un subgrupo

$$E_n = \left\{ \left(\wp\left(\frac{2a\varpi + 2bi\varpi}{n} \right), \wp'\left(\frac{2a\varpi + 2bi\varpi}{n} \right) \right) \middle| \, 0 \leq a, b < n \right\}.$$

El elemento cero corresponde a $a = b = 0$; éste es el punto al infinito.

El grupo E_n es análogo al grupo C_n para la circunferencia (para la definición de C_n ver la sección A.3.2 del apéndice). Extendiendo la analogía, mostremos que si (a, b) y (c, d) son puntos de E, entonces

$$(a, b) + (c, d) = (f(a, b, c, d), g(a, b, c, d)),$$

donde $\sigma f(u) = f(\sigma u)$ y $\sigma g(u) = g(\sigma u)$ para cualquier automorfismo σ del campo $\mathbb{C}$. El teorema de adición de la función $\wp$,

$$\wp(z_1 + z_2) = -\wp(z_1) - \wp(z_2) + \frac{1}{4}\left(\frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)} \right)^2 \tag{3.67}$$

muestra que $f(a, b, c, d) = -a - c - \frac{1}{4}(\frac{b-d}{a-c})^2$ para $a \neq c$. Diferenciando la ecuación (3.67) y tomando en cuenta que $\wp''(z) = 6\wp^2(z) - \frac{1}{2}g_2 = 6\wp^2(z) - \frac{1}{8}$ podemos representar a g como una función racional de a, b, c, d con coeficientes racionales. En el caso $z_1 \equiv z_2 \pmod{\Omega}$ podemos usar la fórmula

$$\wp(2z) = -2\wp(z) + \frac{1}{4}\left(\frac{\wp''(z)}{\wp'(z)} \right)^2.$$

Si $z_1 \equiv -z_2 \pmod{\Omega}$, entonces la fórmula (3.67) se sigue cumpliendo; solamente se tiene que considerar a ambas expresiones como infinitas.

Con la ayuda de las funciones f y g podemos obtener funciones f_n y g_n para las cuales $(f_n(x, y), g_n(x, y)) = n \cdot (x, y)$. Los puntos de E_n están dados en esta forma por las ecuaciones $f_n(x, y) = \infty$ y $g_n(x, y) = \infty$. Para los puntos finitos esto significa que los denominadores de las fracciones f_n y g_n se anulan.

Ahora podemos considerar el campo K_n generado sobre $\mathbb{Q}$ por las coordenadas de los puntos finitos de E_n. Repitiendo para E_n los mismos argumentos que para C_n, vemos que el grupo G_n de automorfismos de K_n sobre $\mathbb{Q}$ es isomorfo a un subgrupo del grupo $\mathrm{Aut}(E_n)$. Puesto que

$$E_n \cong \Omega/n\Omega \cong \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z},$$

se sigue que $\mathrm{Aut}(E_n) \cong GL_2(\mathbb{Z}/n\mathbb{Z})$. En el caso de un primo n el orden de $GL_2(\mathbb{Z}/n\mathbb{Z})$ es igual al número de bases de $(\mathbb{Z}/n\mathbb{Z})^2$, en otras palabras, es igual a $(n^2-1)(n^2-n)$. Este número es divisible entre $n(n+1)$ y, por lo tanto, para $n \geq 2$ este no puede ser una potencia de 2. Nuestros argumentos han llegado a un callejón sin salida!

Solamente nos podemos librar de esto con un truco el cual hemos usado repetidas veces en el estudio de polinomios para la división de la lemniscata, a saber, la invariancia de la latiz de periodos bajo la multiplicación por i. En nuestro caso, esto significa que $\wp(iz)$ y $\wp'(iz)$ pueden expresarse en términos de $\wp(z)$ y $\wp'(z)$. Probemos que $\wp(iz) = -\wp(z)$ y que $\wp'(iz) = i\wp'(z)$. De hecho, como

$$\wp(z) = z^{-2} + \sum_{\omega\in\Omega}{}'((z-\omega)^{-2} - \omega^{-2}).$$

La invariancia de Ω con respecto a la multiplicación por i (es decir, $i\Omega = \Omega$) implica que

$$\wp(iz) = \frac{1}{iz} + \sum_{\omega\in\Omega}{}'\frac{1}{(iz-\omega)^2} - \frac{1}{\omega^2} = -\frac{1}{z^2} - \sum_{\omega\in\Omega}{}'\frac{1}{(z-i\omega)^2} + \frac{1}{(i\omega)^2} = -\wp(z).$$

Diferenciando esta ecuación tenemos $i\wp'(iz) = -\wp'(z)$. Por lo tanto, la acción de i sobre el toro $\mathbb{C}/\Omega$ induce la i-acción sobre E dada por la fórmula $i(x,y) = (-x, iy)$. Sobre el grupo $\Omega/n\Omega$ isomorfo a E_n la acción de $k+il \in \mathbb{Z}[i]$ esta dada por la fórmula

$$(2a\varpi + 2bi\varpi) \pmod n \mapsto (k+il)(2a\varpi + 2bi\Omega) \pmod n.$$

Esta acción puede trasladarse al grupo E_n.

Sean $F = \mathbb{Q}(i)$, F_n el campo generado por las coordenadas de los puntos de E_n sobre F, y G_n el grupo de automorfismos de F_n sobre F. Si $\sigma \in G_n$, entonces $\sigma(i) = i$; por esta razón, $\sigma(i(x,y)) = \sigma(-x, iy) = (\sigma(-x), \sigma(iy) = (-\sigma(x), i\sigma(y)) = i(\sigma(x), \sigma(y))$. Además, $\sigma((a+b)+(c+d)) = \sigma(a,b) + \sigma(c,d)$. De ahí se sigue que G_n es un subgrupo del grupo de automorfismos del $\mathbb{Z}[i]$-módulo $\Omega/n\Omega$.

La inversa de la función $a+ib \mapsto (k+il)(a+ib)$, donde a y b están tomados módulo n, es la función

$$a+ib \mapsto \frac{k-il}{k^2+l^2}(a+ib).$$

Esta función esta definida si y sólo si el número k^2+l^2 es primo relativo a n. Para obtener distintas funciones, tenemos que suponer que $0 \leq k, l \leq n-1$. Por esta razón, el orden del grupo de automorfismos del $\mathbb{Z}[i]$-módulo $\Omega/n\Omega$ es igual al número de pares (k,l), donde $0 \leq k, l \leq n-1$ y k^2+l^2 es primo relativo a n. Sea $\Phi(n)$ el numero total de dichos pares. Nosotros dividiremos el cálculo de $\Phi(n)$ en varios lemas.

Lema 3.5.4. *Si p y q son primos relativos, entonces $\Phi(pq) = \Phi(p)\Phi(q)$.*

Demostración. Sean $0 \leq a_1, b_1 \leq p-1$ y $0 \leq a_2, b_2 \leq q-1$. Entonces los pares $(a,b) = (a_1q + a_2p, b_1q + b_2p)$ constituyen un sistema completo de pares de residuos modulo pq. Además $a^2 + b^2 = (a_1^2 + b_1^2)q^2 + (a_2^2 + b_2^2)p^2$. Por lo tanto, si (u,v) representa el máximo común divisor de u y v,

$$(a^2 + b^2, pq) = 1 \Leftrightarrow \{(a_1^2 + b_1^2, p) = 1 \text{ y } (a_2^2 + b_2^2, q) = 1\}.$$

□

Lema 3.5.5. *Si p es un primo de la forma $4k+3$, entonces $\Phi(p) = p^2 - 1$.*

Demostración. Debemos probar que si $a^2 + b^2$ es divisible entre p, entonces ambos números a^2 y b^2 son divisibles por p. Supongamos que $a^2 + b^2$ es divisible entre p pero que al menos uno de los números a y b no es divisible entre p. Entonces ambos números a y b no son divisibles entre p; de ahí que, por el pequeño teorema de Fermat $a^{p-1} \equiv 1$ (mód p) y $b^{p-1} \equiv 1$ (mód p), consecuentemente, $a^{p-1} + b^{p-1} \equiv 2$ (mód p). Por otra parte, $a^{p-1} + b^{p-1} = a^{4k+2} + b^{4k+2} = (a^2)^{2k+1} + (b^2)^{2k+1}$ es divisible entre $a^2 + b^2$; por lo tanto, es divisible entre p. □

Lema 3.5.6. *Si p es un primo de la forma $4k+1$, entonces $\Phi(p) = (p-1)^2$.*

Demostración. Antes que de comenzar la demostración, recordemos que cualquier primo p de la forma $4k+1$ puede ser representado como la suma de dos cuadrados. La demostración conocida más simple de esta proposición fue sugerida por ***D. Zagier*** y es como sigue:

Se considera al conjunto de todas las soluciones de la ecuación $x^2 + 4yz = p$ en números naturales. Es suficiente probar que esta ecuación tiene una solución para la cual $y = z$. En otras palabras, la involución $\sigma(x,y,z) = (x,z,y)$ definida sobre el conjunto de soluciones tiene un punto fijo. (Recordemos que una ***involución*** es una función f tal que $f(f(x)) = x$ para toda x; si $f(x_0) = x_0$, entonces x_0 es llamado ***punto fijo***.) En un conjunto que consista de un número impar de elementos cualquier involución tiene un punto fijo. Por lo tanto, es suficiente probar que el número total de soluciones de la ecuación dada es impar. Para este fin, es suficiente construir otra involución τ de este conjunto, que tenga exactamente un punto fijo. A saber, definiremos:

$$\tau(x,y,z) = \begin{cases} (x+2z, z, y-x-z) & \text{para} \quad x < y-z, & \text{(A)} \\ (2y-x, y, x-y+z) & \text{para} \quad y-z < x < 2y, & \text{(B)} \\ (x-2y, x-y+z, y) & \text{para} \quad 2y < x. & \text{(C)} \end{cases}$$

Es fácil verifica que $x \neq 2y$ y $x \neq y - z$; junto, cualquier solución es de hecho transformada por τ en una solución.

Dividamos las soluciones en tres tipos (A)-(C) de acuerdo con cual de las siguientes tres desigualdades se satisface:

$$x < y - z, \quad y - z < x < 2y, \quad 2y < x.$$

La función τ manda soluciones del tipo (A) en soluciones del tipo (C), soluciones del tipo (B) en soluciones del tipo (B), y soluciones del tipo (C) en soluciones del tipo (A). Ahora es fácil mostrar que τ es una involución. Sólo un punto del tipo (B) puede ser fijo. La ecuación $(x, y, z) = (2y - x, y, x - y + z)$ implica que $y = x$. Por lo tanto, $p = x(x + 4z)$, es decir, $x = y = 1$ (aquí tenemos que ocupar el hecho de que p es primo). Por esta razón, existe exactamente un punto fijo; a sabe, el punto $(1, 1, k)$ (y aquí haremos uso del hecho de que p es de la forma $4k + 1$).

Ahora probemos que para un $a \neq 0$ fijo la ecuación $x^2 + a^2 \equiv 0 \pmod{p}$ tiene precisamente dos soluciones. De hecho, existen numeros distintos de cero b y c tal que $b^2 + c^2 \equiv 0 \pmod{p}$. Multiplicando esta desigualdad por $(ac^{-1})^2$ vemos que $b_1^2 + a^2 \equiv 0 \pmod{p}$, donde $b_1 = abc^{-1}$. Por lo tanto, $x^2 \equiv b_1^2 \pmod{p}$; las soluciones de esta ecuación son $x = \pm b_1$. Por esta razón, solamente $2(p - 1)$ pares con a y b distintos de cero y el par $(0, 0)$ no entran en $\Phi(p)$. De ahí se sigue que

$$\Phi(p) = p^2 - 1 - 2(p - 1) = (p - 1)^2.$$

□

También es obvio que $\Phi(2) = 2$.

Lema 3.5.7. *Sean p un primo, $k \geq 1$. Entonces $\Phi(p^k) = (p^{k-1})^2\Phi(p)$.*

Demostración. Los números $a + a_1p$, donde $0 \leq a \leq p - 1$ y $0 \leq a_1 \leq p^{k-1} - 1$, forman un sistema completo de residuos módulos p^k. El número $(a + a_1p)^2 + (b + b_1p)^2$ no es primo relativo a p^k si y sólo si este no es primo relativo a p, es decir, $a^2 + b^2 \equiv 0 \pmod{p}$. Falta observar que a todo par (a, b) que entre en $\Phi(p)$ le corresponden $(p^{k-1})^2$ pares que entran en $\Phi(p^k)$. □

Ahora es fácil verificar que $\Phi(n)$ es una potencia de 2 si y sólo si $n = 2\alpha p_1 \cdots p_m$, donde los p_i son primos de Fermat distintos. De hecho, $\Phi(2^\alpha p_1^{k_1} \cdots p_m^{k_m})$ puede ser una potencia de 2 si y sólo si $k_1 = \cdots = k_m = 1$. Si $p = 4k+1$, entonces $\Phi(p) = (p-1)^2$, y este número es una potencia de 2 si y sólo si $p = 1+2^c$. Si $p = 4k+3$ entonces como $\Phi(p) = p^2-1 = (p-1)(p+1)$, este sera una potencia de 2, si los números pares consecutivos $p-1$ y $p+1$ son una potencia de 2, pero esto sera verdadero si y sólo si $p = 3$.

Para completar la demostración del teorema de Abel falta verificar que para la latiz considerada, es decir, $\Omega = \{(2a\varpi + 2ib\varpi) \,|\, a, b \in \mathbb{Z}\}$, tenemos

$$g_2 = 60 \sum_{\omega} \in \Omega'(2a\varpi + 2ib\varpi)^{-4} = \frac{1}{4} \quad \text{y} \quad g_3 = 140 \sum_{\omega} \in \Omega'(2a\varpi + 2ib\varpi)^{-6} = 0.$$

La segunda igualdad es obvia ya que $g_3(\Omega) = -g_3(i\Omega)$ y en nuestro caso $i\Omega = \Omega$. La principal dificultad es probar que $\sum'(a\varpi + bi\Omega)^{-4} = \frac{1}{15}$.

Consideremos las tres latices

$$\begin{aligned} L_0 &= \{a\varpi + bi\varpi\}, \\ L_1 &= \left\{ \frac{a\varpi + bi\varpi}{2} \middle| \, a \text{ y } b \text{ son impares} \right\}, \\ L_2 &= \left\{ \frac{a\varpi + bi\varpi}{2} \middle| \, a - b \text{ es impar} \right\}. \end{aligned}$$

Entonces $\frac{1}{2}L_0 = L_0 \cup L_1 \cup L_2$ y $L_2 = \frac{1+i}{2}L_1$. Definamos $|L| = \sum{}'_{l\in L} l^{-4}$. Entonces

$$16|L_0| = \left|\frac{1}{2}L_0\right| = |L_0| + |L_1| + |L_2| \quad \text{y} \quad |L_2| = \left(\frac{2}{1+i}\right)^4 |L_1| = -4|L_1|;$$

de ahí que, $|L_1| = -5|L_0|$. Para probar la desigualdad deseada $|L_0| = \frac{1}{15}$ necesitamos una relación mas entre $|L_1|$ y $|L_0|$. Para esto usemos el hecho de que L_0 es la latiz de ceros de $\varphi(z)$ y L_1 es la latiz de sus polos. Tomando en cuenta que $\varphi'(0) = 1$ tenemos que

$$\varphi(z) = z \prod_{\alpha\in L_0}{}' \left(1 - \frac{z}{\alpha}\right) \prod_{\beta\in L_1}{}' \left(1 - \frac{z}{\beta}\right)^{-1},$$

donde los productos infinitos deben entenderse como límites de productos finitos sobre $|\alpha|, |\beta| \le N$ cuando $N \to \infty$. Los elementos distintos de cero de las latices L_0 y L_1 pueden dividirse en cuartetas de la forma $\{\pm\gamma, \pm i\gamma\}$; por esta razón,

$$\varphi(z) = z \prod{}' \left(1 - \frac{z^4}{\alpha^4}\right) \prod \left(1 - \frac{z^4}{\beta^4}\right)^{-1},$$

donde $0 \le \arg\alpha, \arg\beta < \frac{\pi}{2}$. Por lo tanto,

$$z\frac{\varphi'(z)}{\varphi(z)} = z\frac{d}{dz}\ln\varphi(z) = 1 + (|L_1| - |L_0|)z^4 + \cdots. \tag{3.68}$$

La función $z^{-1}\varphi(z)$ no cambia bajo la transformación de z a $-z$ ó a $\pm iz$; de ahí que $\varphi(z) = z(1 + cz^4 + \cdots)$. Además como, $(\varphi'(z))^2 = 1 - \varphi^4(z)$, se tiene que

$$(1 + 5cz^4 + \cdots)^2 = 1 - z^4(1 + cz^4 + \cdots)^4$$

y, por lo tanto, $c = -\frac{1}{10}$. De ahí se sigue que

$$z\frac{\varphi'(z)}{\varphi(z)} = 1 + 4cz^4 + \cdots = 1 - \frac{2}{5}z^4 + \cdots.$$

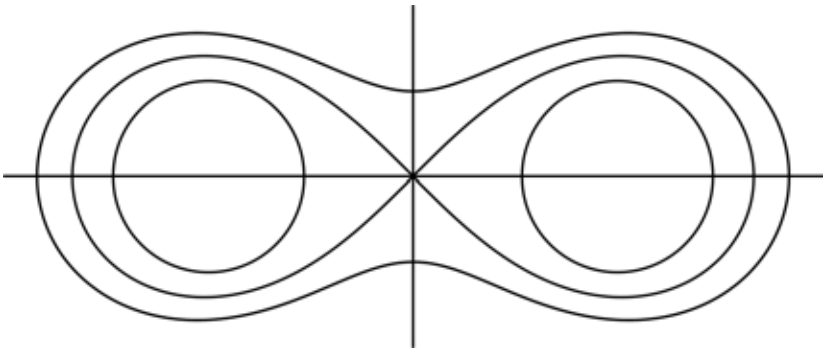
Comparando (3.68) con (3.5.3) deducimos que $|L_1| - |L_0| = -\frac{2}{5}$. Como $|L_1| = -5|L_0|$, se sigue que $|L_0| = \frac{1}{15}$. $\square$

3.6. Notas históricas

La primera aparición de la lemniscata en la literatura matemática fue como parte de los ***óvalos de Cassini***, descritos en 1680 por el astrónomo frances de origen italiano Jean-Dominique (Giovanni Dominico) Cassini (1625-1712). En coordenadas Cartesianas, los óvalos son la familia de curvas definidas por la ecuación

$$((x-a^2)+y^2)((x+a^2)^2+y^2)=b^4. \tag{3.69}$$

La lemniscata que hemos estado estudiando corresponde a $a=b=1/\sqrt{2}$. En general, $a<b$ dibuja una especie de burbuja y $a>b$ dibuja dos óvalos, como en la siguiente imagen:



Sin conocer el trabajo de Cassini, en 1694 Jacob Bernoulli expresó la ecuación de la lemniscata como

$$xx+yy=a\sqrt{xx-yy}.$$

El describió la curva como "la figura en forma de 8 acostada, también como una cinta amarrada como nudo, o como una lemniscus, o como un nudo de un listón Frances". Aquí, "lemniscus" es una palabra en Latin (tomada del Griego) que se refiere al listón que llevaba colgando la guirnalda que era portada por los ganadores de una competencia atlética.

Bernoulli fue llevado a esta curva por una ruta indirecta. En 1691 el encontró la integral $2\int_0^1(1-t^4)^{-1/2}dt$ (esto debe parecer familiar) en su estudio de la *curva elástica*. Para representar esto geométricamente, el buscaba una curva definida por una ecuación algebraica cuya longitud de arco fuese igual a $2\int_0^1(1-t^4)^{-1/2}dt$. En 1694 el mostró que la lemniscata tenia la longitud de arco que él deseaba, usando la descripción polar de la lemniscata como hicimos nosotros antes.

El uso de Bernoulli de las coordenadas polares para calcular la longitud de arco de la lemniscata representa el primer uso de la longitud de arco con coordenadas polares. Es irónico que los estudiantes de cálculo estudien la lemniscata en una parte del curso y la longitud de arco en coordenadas polares en otra, pero que nunca junten estas ideas, aún cuando la integral que resulta no pueda ser evaluada por los métodos usuales del cálculo.

Nuestra discusión muestra que la lemniscata y su longitud de arco fueron bien conocidas a principios del siglo dieciocho. Por esta razón el teorema de Abel sobre la división de

la lemniscata en arcos de la misma longitud tiene relación con un tópico de moda en la comunidad matemática de esa época.

Aunque la relación entre la longitud de arco y la lemniscata se remonta a Bernoulli, la primera persona en realizar un progreso sustancial en esta área fue el matemático italiano Conde Giulio Carlo Fagnano (1682-1766). En 1718 el probó el caso $\alpha = \beta$ de la ley de adición de Euler (3.7), a saber

$$2\int_0^{\alpha} \frac{1}{\sqrt{1-t^4}}dt = \int_0^{\gamma} \frac{1}{\sqrt{1-t^4}}dt \quad \text{cuando} \quad \gamma = \frac{2\sqrt{1-\alpha^4}}{1+\alpha^4}.$$

Usando éste y otros resultados, Fagnano fue capaz de dividir un arco de la lemniscata en dos, tres, y cinco segmentos de la misma longitud con regla y compás. Los resultado y los métodos de Fagnano son discutidos en [1].

Los descubrimientos se tornaron interesantes cuando los trabajos de Fagnano fueron enviados a la Academia de Berlin como parte de su solicitud a ser miembro. En Diciembre de 1751 se le pidió a Euler que leyera estos escritos. En enero del siguiente año da un dictamen favorable y para 1753 Euler fue capaz de mostrar que la fórmula de duplicación de Fagnano era un caso especial de (3.7). Más importante, el también se dio cuenta de que $\sqrt{1-t^4}$ podia ser reemplazada por $\sqrt{P(t)}$, donde $P(t)$ puede ser cualquier polinomio separable de grado 4 con coeficientes reales. Esto preparó el camino para la teoría de las *integrales elípticas*, la cual fue desarrollada en mayor proporción por Lagrange y Legendre. Eventualmente estas integrales fueron expresadas en su forma estándar

$$\int \frac{1}{\sqrt{1-k^2\operatorname{sen}^2\theta}}d\theta, \tag{3.70}$$

la cual después de la sustitución $t = \operatorname{sen}\theta$ se expresa como

$$\int \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}}dt. \tag{3.71}$$

Llamaremos a k el módulo de la integral, nuestra integral $\int(1-t^4)^{-1/2}dt$ corresponde al módulo $k = i$.

La primera persona en considerar la función inversa de $\int(1-t^4)^{-1/2}dt$ fue Gauss en 1797, pero su trabajo no fue publicado hasta después de su muerte en 1855. En el artículo de Abel *Recherches sur les fonctions elliptiques*, él considera la función inversa de una integral elíptica de la forma

$$\int \frac{1}{\sqrt{(1-c^2t^2)(1+e^2t^2)}}dt, \tag{3.72}$$

de tal manera que $c = e = 1$ proporciona la función lemniscática que hemos estado estudiando. Jacobi, por otra parte usó la integral (3.70) y expresó su función inversa como $\theta = \operatorname{am} u$. Por esta razón $\operatorname{sen}\theta = \operatorname{sen}\operatorname{am} u$ es la función inversa de (3.71). Hoy en día, se

escribe $\operatorname{sen}\operatorname{am} u$ como $\operatorname{sn}(u,k)$ o simplemente $\operatorname{sn}(u)$ si tenemos definido el módulo. De esta forma tenemos que la función lemniscática es $\varphi(u) = \operatorname{sn}(u,i)$.

Uno de los descubrimientos críticos de Gauss, Abel y Jacobi es que las funciones inversas de integrales elípticas son funciones doblemente periódicas de una variable compleja. Más de la historia de las integrales elípticas puede encontrarse en [1] o en [11].

Abel denotó a la función inversa de la integral (3.72) como $\varphi(x)$ y entonces definió $f(x) = \sqrt{(1-c^2\varphi^2(x))}$ y $F(x) = \sqrt{(1+e^2\varphi^2(x))}$. Estas funciones están relacionadas via $\varphi'(x) = f(x)F(x)$. Abel proporcionó leyes de adición para $\varphi(x)\,,\ f(x)\,,\ F(x)$ y fórmulas de multiplicación para $\varphi(nx)\,,\ f(nx)\,,\ F(nx)$ en un espíritu similar al Teorema 3.2.5. El también extendió estas funciones para números complejos y determinó sus periodos, ceros y polos.

Jacobi desarrolló una teoría similar basada en la integral (3.71). El definió las funciones $\operatorname{sen}\operatorname{am} x, \cos\operatorname{am} x$ y $\Delta\operatorname{am} x$, para las cuales su notación se simplificó a $\operatorname{sn}(x), \operatorname{cn}(x), \operatorname{dn}(x)$. Su versión de la teoría llegó a ser muy influyente, aunque eventualmente fue reemplazada por la función $\wp$ introducida por Weierstrass en 1822. Un importante resultado de Weierstrass es que toda función elíptica con los mismos periodos de $\wp$ es una función racional de $\wp(z)$ y $\wp'(z)$. Por lo tanto una vez que la latiz de periodos esta fija, bastará conocer a éstas dos funciones elípticas para obtener todas las demás.

Gauss anticipó mucho del trabajo de Abel y Jacobi sobre funciones elípticas pero nunca publicó sus resultados. Como el escribió en 1828,

> Muy probablemente no prepararé todavía mi investigación sobre funciones trascendentales que he tenido por muchos años –desde 1798– por que tengo muchos otros temas que deben ser esclarecidos. El distinguido Abel me ha anticipado y relevado de la carga con respecto a un tercio de estos temas, particularmente desde que el realizó todos estos desarrollos en forma concisa y con gran elegancia.

¿Que llevó a Gauss y a Abel a trabajar sobre los números complejos? Al parecer ellos estuvieron inspirados en definir $\varphi(z)$ para $z \in \mathbb{C}$ para representar *todas* las raíces de los n-polinomios de división de la lemniscata. Además para estos polinomios conocían que las raíces no pueden ser todas reales.

En adición a la teoría general de las funciones elípticas, Abel también consideró la función lemniscática $\varphi(z)$ que hemos estado estudiando. Sea $m + \mu i \in \mathbb{Z}[i]$ impar, y definamos $x = \varphi(\delta)$. Entonces Abel estableció la multiplicación compleja por $m + \mu i$ como "una que cumple

$$\varphi(m + \mu i)\delta = x \cdot T,$$

donde T es una función racional de x^4". Como un ejemplo, el escribió la fórmula para la multiplicación compleja por $2 + i$ como

$$\varphi(2+i)\delta = x\frac{2 - 2x^8 + i(1 - 6x^4 + x^8)}{1 - 2x^4 + 5x^8} = xi\frac{1 - 2i - x^4}{1 - (1-2i)x^4}.$$

Esto esta muy relacionado con lo que hicimos en el Ejemplo 3.4.6.

Eisenstein también juega un papel importante en esta historia puesto que el fue el primero en probar los Teoremas 3.4.5 y 3.4.9. Transcribiremos un fragmento de una carta que el escribió a Gauss en 1847:

> Si $m = a + ib$ es un entero complejo impar, p es su norma y $y = \frac{U}{V} = \frac{A_0x+A_1x^5+\cdots+A_{(p-1)/4}x^9}{1+B_1x^4+\cdots+B_{(p-1)/4}x^p}$ es la integral algebraica de la ecuación
>
> $$\int dy/\sqrt{1-y^4} = m\int dx/\sqrt{1-x^4},$$
>
> he mostrado que para un número complejo primo de *dos términos* m los coeficientes del numerador, excepto por el último el cual es una unidad compleja, y los coeficientes del denominador, excepto por el primero el cual es $= 1$, son divisibles por m. Yo conjeturo que el teorema también es correcto, cuando m es un número primo de *un solo término* ...

Aquí, un primo complejo impar de "dos términos" es $m = a + bi$ tal que $p = a^2 + b^2$ es un primo en $\mathbb{Z}$ con $p \equiv 1 \mod 4$, y un primo complejo de "un término" es un primo en $\mathbb{Z}$ tal que $p \equiv 3 \mod 4$. En esta carta, Eisenstein puede probara la parte (a) del Teorema 3.4.9 solo en el caso de los "dos términos", pero mas tarde él obtendrá una prueba general. También, si pensamos en $\varphi(z)$ como la función inversa de la integral elíptica $\int(1-t^4)^{-1/2}dt$, entonces debe ser claro que la ecuación mostrada en la carta de Eisenstein se refiere a la fórmula de multiplicación para $\varphi(mz)$ en términos de $\varphi(z)$.

La afirmación mas clara del criterio de irreducibilidad de Eisenstein aparece en un articulo que escrito en 1850, donde encontramos el siguiente teorema:

> Si en un polinomio $F(x)$ de s de grado arbitrario cuyo coeficiente del mayor termino es $= 1$, y todos los siguientes coeficientes son enteros (reales o complejos), en los cuales un cierto numero primo (real o resp. complejo) m aparece, y además el último coeficiente es $= \varepsilon m$, donde ε representa un número no divisible por m; entonces es imposible llevar a $F(x)$ a la forma
>
> $$(x^\mu + a_1x^{\mu-1} + \cdots + a_\mu)(x^\nu + b_1x^{\nu-1} + \cdots + b_\nu),$$
>
> donde μ y $\nu \geq 1$, $\mu+\nu =$ a el grado de $F(x)$, y todas las a y b son enteros (reales o resp. complejos); y le ecuación $F(x) = 0$ es por lo tanto irreducible.

La razón por la cual Eisenstein afirmó el teorema para $\mathbb{Z}$ y $\mathbb{Z}[i]$ es que probablemente el primero descubrió esto para $\mathbb{Z}[i]$ en sus estudios acerca de la multiplicación compleja en la lemniscata y entonces se dio cuenta de que esto también se aplica en $\mathbb{Z}$.

Apéndice A

Apéndice de Álgebra

A.1. Extensiones de campos y teoría de Galois

Definición A.1.1. *Un dominio integral R es un* ***dominio de factorización única*** *o* ***(DFU)****, si las siguientes dos condiciones se cumplen:*

(a) *Cualquier elemento de R distinto del cero es o una unidad o un producto de irreducibles.*

(b) *Si $r_1 \cdots r_k = s_1 \cdots s_l$, donde $r_1, \cdots, r_k, s_1, \cdots, s_l \in R$ son irreducibles, entonces $k = l$, y existe una permutación $\sigma \in S_k$ tal que para cada $1 \leq i \leq k$ existe una unidad $a_i \in R^*$ tal que $r_i = a_i s_{\sigma(i)}$.*

Teorema A.1.2. *Sea R un dominio de factorización única, y sea $R[x]$ el anillo de polinomios en una variable x con coeficientes en R. Entonces $R[x]$ también es un dominio de factorización única.*

Definición A.1.3. *Dado un homomorfismo de campos inyectivo $\varphi : F \to L$, decimos que L es una* ***una extensión de campo de*** *F* ***via*** *φ. Usualmente se identifica a F con su imagen*

$$\varphi(F) = \{\varphi(a) \,|\, a \in F\} \subset L$$

y se escribe $F \subset L$.

Definición A.1.4. *Sea $F \subset L$ una extensión de campo.*

(a) *L es una* ***extensión finita*** *de F si L es un espacio vectorial de dimensión finita sobre F.*

(b) *El* ***grado*** *de L sobre F, denotado por $[L : F]$, esta definido de la siguiente manera:*

$$[L : F] = \begin{cases} \dim_F L & si\ L\ es\ una\ extensión\ finita\ de\ F, \\ \infty & de\ otra\ forma, \end{cases}$$

Definición A.1.5. *Sea $f \in F[x]$ de grado $n > 0$. Entonces una extensión $F \subset L$ es un* ***campo de descomposición*** *de f sobre F si*

(a) $f = c(x - \alpha_1) \cdots (x - \alpha_n)$, *donde* $c \in F$ *y* $\alpha_i \in L$, *y*

(b) $L = F(\alpha_1, \ldots, \alpha_n)$.

Definición A.1.6. *Sea $F \subset L$ una extensión finita. El* ***grupo de Galois de la extensión*** *es* $\mathrm{Gal}(L/F)$ *es el conjunto*

$$\{\sigma : L \to L \,|\, \sigma \text{ es un automorfismo y } \sigma(a) = a \text{ para toda } a \in F\}$$

En otras palabras, $\mathrm{Gal}(L/F)$ *consiste en todos los automorfismos de L que fijan a todos los elementos de F.*

Definición A.1.7. *Sea $f \in F[x]$.* ***El grupo de Galois de f sobre*** *F es* $\mathrm{Gal}(L/F)$, *donde L es un campo de descomposición de f sobre F.*

Definición A.1.8. *Un polinomio $f \in F[x]$ es* ***separable*** *si es no constante y sus raíces en un campo de descomposición son todas simples*

Definición A.1.9. *Una extensión $F \subset L$ es llamada una* ***extensión de Galois*** *si L es el campo de descomposición de un polinomio separable en $F[x]$.*

Definición A.1.10. *Sea L una extensión de F, y sea $\alpha \in L$. Entonces α es un elemento* ***algebraico*** *sobre F si existe un polinomio no constante $f \in F[x]$ tal que $f(\alpha) = 0$. Si α no es algebraico sobre F, entonces se dice que α es* ***trascendente*** *sobre F.*

Definición A.1.11. *Una extensión de campo $F \subset L$ es* ***algebraica*** *si cada elemento de L es algebraico sobre F.*

Definición A.1.12. *Una extensión algebraica $F \subset L$ es* ***normal*** *si cada polinomio irreducible en $F[x]$ que tiene una raíz en L se descompone completamente sobre L.*

Definición A.1.13. *Un* ***grupo finito*** *G* ***es soluble*** *si existen subgrupos*

$$\{e\} = G_n \subset G_{n-1} \subset \cdots \subset G_1 \subset G_0 = G$$

tales que para cada $i = 1, \ldots, n$ se tiene que:

(a) G_i *es normal en* G_{i-1}.

(b) $[G_{i-1} : G_i]$ *es primo.*

Definición A.1.14. *Una extensión de campo $F \subset L$ es* ***radical*** *si existen campos*

$$F = F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = L$$

donde para cada $i = 1, \ldots, n$, existe $\gamma_i \in F_i$ con $F_i = F_{i-1}(\gamma_i)$ y $\gamma_i^{m_i} \in F_{i-1}$, con $m_i > 0$.

Notemos que si tomamos $b_i = \gamma_i^{m_i} \in F_{i-1}$, entonces γ_i es una m_i-esima raíz de b_i. Esto permitirá escribir $\gamma_i = \sqrt[m_i]{b_i}$, por lo cual $F_i = F_{i-1}\left(\sqrt[m_i]{b_i}\right)$, $b_i \in F_{i-1}$. Esto muestra que una extensión radical se obtiene adjuntando radicales sucesivamente.

Definición A.1.15. *Una* ***extensión de campo*** $F \subset L$ ***es soluble*** *(o* ***soluble por radicales****) si existe una extensión de campo* $L \subset M$ *tal que* $F \subset M$ *es radical.*

Teorema A.1.16. *Todo grupo Abeliano finito* G *es soluble.*

Teorema A.1.17. *Sea* $F \subset L$ *una extensión de Galois. Entonces las siguientes afirmaciones son equivalentes:*

(*a*) $F \subset L$ *es una extensión soluble.*

(*b*) $\mathrm{Gal}(L/F)$ *es un grupo soluble.*

Definición A.1.18. *Sea* $f \in F[x]$ *no constante con campo de descomposición* $F \subset L$*.*

(*a*) *Una raíz* $\alpha \in L$ *de* f *es* ***expresable por radicales sobre*** F *si* α *se encuentra en alguna extensión radical de* F*.*

(*b*) *El polinomio* f *es* ***soluble por radicales sobre*** F *si* $F \subset L$ *es una extensión soluble.*

Definición A.1.19. *Sea* $F \subset L$ *un extensión finita con grupo de Galois* $\mathrm{Gal}(L/F)$*. Dado un subgrupo* $H \subset \mathrm{Gal}(L/F)$*, llamaremos al campo*

$$L_H = \{\alpha \in L \,|\, \sigma(\alpha) = \alpha \text{ para toda } \sigma \in H\}$$

el ***campo fijo*** *de* H*. Donde* $F \subset L_H \subset L$ *son subcampos.*

Lema A.1.20. *Si* $F \subset L$ *y* $L \subset M$ *son extensiones radicales, entonces* $F \subset M$ *también es extensión radical.*

Teorema A.1.21. *Si* H *es un subgrupo de un grupo finito* G*, entonces* $|H|$ *divide a* $|G|$*.*

Teorema A.1.22. *Sea* $F \subset L$ *una extensión finita. Entonces:*

(*a*) $|\mathrm{Gal}(L/F)|$ *divide a* $[L : F]$*.*

(*b*) $F \subset L$ *es una extensión de Galois si y sólo si* $|\mathrm{Gal}(L/F)| = [L : F]$*.*

Teorema A.1.23. *Sea* $F \subset L$ *es una extensión finita. Entonces las siguientes afirmaciones son equivalentes:*

(*a*) L *es el campo de descomposición de un polinomio separable en* $F[x]$*.*

(*b*) $L_{\mathrm{Gal}(L/F)} = F$*.*

(*c*) $F \subset L$ *es una extensión normal y separable.*

Proposición A.1.24. *Suponiendo que* $F \subset L$ *es una extensión de Galois y si tenemos un campo intermedio* $F \subset K \subset L$*. Entonces* $K \subset L$ *es una extensión de Galois.*

Teorema A.1.25. *Sea $F \subset L$ una extensión de Galois.*

(*a*) *Para un campo intermedio $F \subset K \subset L$, su grupo de Galois* $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$ *tiene campo fijo*

$$L_{\mathrm{Gal}(L/K)} = K.$$

Además, $|\mathrm{Gal}(L/K)| = [L : K]$ *y* $[\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)] = [K, F]$.

(*b*) *Para un subgrupo $H \subset \mathrm{Gal}(L, F)$, su campo fijo $F \subset L_H \subset L$ tiene grupo de Galois*

$$\mathrm{Gal}(L/L_H) = H.$$

Además, $[L : L_H] = |H|$ *y* $[L_H : F] = [\mathrm{Gal}(L/F) : H]$.

Las demostraciones de estas afirmaciones pueden encontrarse en [3] y [4].

A.2. Números construibles

Para definir las construcciones con regla y compás, sean α, β y γ puntos en un plano. Entonces, una ***regla*** es un instrumento que nos permite construir la recta que pasa por dos puntos dados y un ***compás*** es un instrumento que nos permite construir una circunferencia con centro en un punto dado y que a su vez pase por otro punto dado (equivalentemente nos permite construir una circunferencia con un punto dado y un radio dado). De esta manera tenemos nuestras primeras construcciones:

C1. De $\alpha \neq \beta$, podemos dibujar una recta ℓ que vaya de α a β.

C2. De $\alpha \neq \beta$ y γ, podemos dibujar una circunferencia $\mathcal{C}$ con centro en γ cuyo radio sea la distancia de α a β.

Finalmente, los puntos que utilizaremos para nuestras construcciones serán las intersecciones de las figuras construidas de las dos formas descritas antes.

P1. El punto de intersección de distintas rectas ℓ_1 y ℓ_2 construidas como antes.

P2. Los puntos de intersección de una recta ℓ y una circunferencia $\mathcal{C}$ construidas como antes.

P3. Los puntos de intersección de circunferencias distintas $\mathcal{C}_1$ y $\mathcal{C}_2$ construidas como antes.

Por lo tanto, con estas construcciones podemos ahora definir que es un número construible.

Definición A.2.1. *Un numero complejo α es* ***construible*** *si existe una sucesión finita de construcciones con regla y compás utilizando* **C1, C2, P1, P2, P3** *que comience con 0 y 1 y termine con α.*

Teorema A.2.2. *El conjunto de* $\mathcal{C} = \{\alpha \in \mathbb{C} \,|\, \alpha$ *es un numero construible*$\}$ *es un subcampo de* $\mathbb{C}$*. Además:*

(a) *Sea* $\alpha = a + ib \in \mathbb{C}$*, donde* $a, b \in \mathbb{R}$*. Entonces* $\alpha \in \mathcal{C}$ *si y solo si* $a, b \in \mathcal{C}$*.*

(b) $\alpha \in \mathcal{C}$ *si y sólo si* $\sqrt{\alpha} \in \mathcal{C}$*.*

Teorema A.2.3. *Sea* α *un numero complejo. Entonces* α *es un numero construible si y sólo si existen subcampos*

$$\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n \subset \mathbb{C}$$

tales que $\alpha \in F_n$ *y* $[F_i : F_{i-1}] = 2$ *para* $1 \leq i \leq n$

Corolario A.2.4. *Si* $\alpha \in \mathcal{C}$*, entonces* $[\mathbb{Q}(\alpha), \mathbb{Q}] = 2^m$ *para alguna* $m \geq 0$*. Por lo tanto todo número construible es algebraico sobre* $\mathbb{Q}$*, y el* grado *de su polinomio mínimo sobre* $\mathbb{Q}$ *es una potencia del 2.*

Teorema A.2.5. *Sea* $\alpha \in \mathbb{C}$ *algebraico sobre* $\mathbb{Q}$*, y se a* $\mathbb{Q} \subset L$ *el campo de descomposición del polinomio mínimo de* α *sobre* $\mathbb{Q}$*. Entonces* α *es construible si y sólo si* $[L : \mathbb{Q}]$ *es una potencia del 2.*

Las demostraciones de estas afirmaciones las podemos encontrar en el capítulo 10 de [3].

A.3. Construcción de polígonos regulares

Debido a la estrecha relación existente entre la división de la lemniscata y la división de la circunferencia, en este apartado estudiaremos este último caso detenidamente. Primero en A.3.1 considerando una ecuación más simple para la división de la circunferencia, mostraremos como resolver en raíces cuadradas la ecuación $x^{17} - 1 = 0$ por un método más o menos elemental, a pesar de que este método no puede ser generalizado para la ecuación de la división de la lemniscata. Después, en la sección A.3.2 discutiremos como el estudio de la solubilidad de la ecuación $x^n - 1 = 0$ en raíces cuadradas puede ser generalizado para la ecuación de la división de la lemniscata.

A.3.1. Construcción de un 17-ágono regular. Una aproximación elemental

Las raíces de la ecuación $x^n - 1 = 0$ son los vértices de un n-ágono regular. De hecho, si $\varepsilon = \exp(2\pi i/n)$, entonces $\varepsilon, \varepsilon^2, \ldots, \varepsilon^n = 1$ son raíces de esta ecuación. Dividiendo el polinomio $x^n - 1$ por $x - 1$ obtenemos el polinomio $x^{n-1} + x^{n-2} + \cdots + x + 1$. De esta manera si la ecuación

$$x^{n-1} + x^{n-2} + \cdots + x + 1 = 0 \tag{A.1}$$

se puede resolver en raíces cuadradas, entonces es posible construir un n-ágono regular con regla y compás.

Para $n = 3$ no hay problema, ya que sin duda, la cuadrática $x^2 + x + 1 = 0$ se puede resolver en raíces cuadradas. Para $n = 5$ la ecuación (A.1) también es fácil de resolver. De hecho, la substitución $u = x + x^{-1}$ transforma tal ecuación en $u^2 + u - 1 = 0$.

Para $n = 17$ no es tan fácil resolver la ecuación (A.1) en raíces cuadradas. Para hacer esto, Gauss usó una partición especial de los números $\varepsilon, \varepsilon^2, \varepsilon^3, \ldots, \varepsilon^{16}$ en grupos, donde $\varepsilon = \exp(2\pi i/17)$. Para obtener dicha división, enumeraremos los números dados de tal forma que para l fijo la raíz ε_{k+l} se obtiene de ε_k de la misma forma, a saber, elevándolas a una potencia fija: $\varepsilon_{k+l} = (\varepsilon_k)^c$:

$$\varepsilon_k \varepsilon_l = \varepsilon_{k+l}.$$

Dicha enumeración se puede obtener definiendo $\varepsilon_k = \varepsilon^{g^k}$, donde los residuos de los números $1, g, g^2, \ldots, g^{15}$ después de dividirlos por 17 toman todos los valores de 1 a 16. Es fácil ver que $g = 3$ posee tal propiedad. Para $g = 3$ los números $\varepsilon_0, \ldots, \varepsilon_{15}$ y sus respectivos valores son escritos uno bajo el otro en la siguiente tabla:

ε	ε^3	ε^9	ε^{10}	ε^{13}	ε^5	ε^{15}	ε^{11}	ε^{16}	ε^{14}	ε^8	ε^7	ε^4	ε^{12}	ε^2	ε^6
ε_0	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8	ε_9	ε_{10}	ε_{11}	ε_{12}	ε_{13}	ε_{14}	ε_{15}

Sea x_1 la suma de los números ε_k con índice par, y x_2 la suma de los números ε_k con índice impar, es decir,

$$\begin{aligned} x_1 &= \varepsilon + \varepsilon^9 + \varepsilon^{13} + \varepsilon^{15} + \varepsilon^{16} + \varepsilon^8 + \varepsilon^4 + \varepsilon^2, \\ x_2 &= \varepsilon^3 + \varepsilon^{10} + \varepsilon^5 + \varepsilon^{11} + \varepsilon^{14} + \varepsilon^7 + \varepsilon^{12} + \varepsilon^6. \end{aligned}$$

La suma de todas las raíces de la ecuación $x^{17} - 1 = 0$ (incluida la raíz $x = 1$) es igual a cero, de ahí que, $x_1 + x_2 = -1$. Cálculos simples muestran que $x_1 x_2 = -4$. De hecho, si $\alpha = 2\pi/17$, entonces $\varepsilon^k = \cos k\alpha + i \operatorname{sen} k\alpha$, por esta razón,

$$\begin{aligned} &\varepsilon + \varepsilon^{16} = 2\cos\alpha, \qquad \varepsilon^9 + \varepsilon^8 = 2\cos 8\alpha, \\ &\varepsilon^{13} + \varepsilon^4 = 2\cos 4\alpha, \quad \varepsilon^{15} + \varepsilon^2 = 2\cos 2\alpha, \end{aligned}$$

es decir,

$$x_1 = 2(\cos\alpha + \cos 8\alpha + \cos 4\alpha + \cos 2\alpha).$$

De forma análoga,

$$x_2 = 2(\cos 3\alpha + \cos 7\alpha + \cos 5\alpha + \cos 6\alpha).$$

Usando la fórmula

$$2\cos p\alpha \cos q\alpha = \cos(p+q)\alpha + \cos(p-q)\alpha$$

obtenemos

$$x_1 x_2 = 8(\cos\alpha + \cos 2\alpha + \cos 3\alpha + \cdots + \cos 8\alpha) = 4(x_1 + x_2) = -4.$$

Así que, podemos encontrar x_1 y x_2 de la ecuación cuadrática

$$x^2 - x - 4 = 0. \tag{A.2}$$

Ya que

$$\cos\alpha + \cos 2\alpha > 2\cos\frac{\pi}{4} = \sqrt{2} > -\cos 8\alpha$$

y $\cos 4\alpha > 0$, se sigue que $x_1 > 0$. Por esto, $x_2 = -\frac{4}{x_1} < 0$, es decir, x_1 es la raíz positiva de la ecuación (A.2) y x_2 es la raíz negativa.

Denotando a y_1, y_3, y_2 y y_4 como la suma de los números ε_k con indices tales que su residuo modulo 4 sea igual a 0,2,1 y 3, respectivamente, tenemos que

$$\begin{array}{rcccl} y_1 & = & \varepsilon + \varepsilon^{13} + \varepsilon^{16} + \varepsilon^4 & = & 2(\cos\alpha + \cos 4\alpha), \\ y_2 & = & \varepsilon^9 + \varepsilon^{15} + \varepsilon^8 + \varepsilon^2 & = & 2(\cos 8\alpha + \cos 2\alpha), \\ y_3 & = & \varepsilon^3 + \varepsilon^5 + \varepsilon^{14} + \varepsilon^{12} & = & 2(\cos 3\alpha + \cos 5\alpha), \\ y_4 & = & \varepsilon^{10} + \varepsilon^{11} + \varepsilon^7 + \varepsilon^6 & = & 2(\cos 7\alpha + \cos 6\alpha). \end{array}$$

Es claro que $y_1 + y_2 = x_1$ y $y_1 > y_2$, por que $\cos\alpha > \cos 2\alpha$ y $\cos 4\alpha > \cos 8\alpha$. Además,

$$y_1 y_2 = 4(\cos\alpha + \cos 4\alpha)(\cos 8\alpha + \cos 2\alpha) = 2(\cos\alpha + \cdots + \cos 8\alpha) = -1.$$

Por lo tanto, y_1 y y_2 satisfacen la ecuación $y^2 - x_1 y - 1 = 0$. Es fácil verificar que y_3 y y_4 satisfacen la ecuación $y^2 - x_2 y - 1 = 0$; además, $y_3 > y_4$.

Finalmente, consideremos $z_1 = \varepsilon + \varepsilon^{16} = 2\cos\alpha$ y $z_2 = \varepsilon^{13} + \varepsilon^4 = 2\cos 4\alpha$, es decir, la suma de los números ε_k con indices tales que su residuo modulo 8 sea igual a 0 y 4, respectivamente. Entonces $z_1 > z_2, z_1 + z_2 = y_1$ y

$$z_1 z_2 = 4\cos\alpha\cos 4\alpha = 2(\cos 5\alpha + \cos 3\alpha) = y_3.$$

Por lo tanto, z_1 es la raíz mas grande de la ecuación $z^2 - y_1 z + y_3 = 0$. Por esta razón, la longitud del segmento $z_1 = 2\cos(2\pi/17)$ puede ser construido con regla y compás. Ahora es claro como construir un 17-ágono regular.

Para finalizar esta sección mostraremos un construcción simplificada del polígono regular de 17 lados preparada por H. W. Richmond en 1983. En la página siguiente se mostrara la ilustración paso a paso de la construcción.

La construcción procede de la siguiente manera:

(1) Se traza una circunferencia de centro O y radio OP_0 de longitud arbitraria. Se traza la recta OB, perpendicular a OP_0. Se determina un punto J a la cuarta parte del recorrido OB. Se halla el punto E tal que el ángulo OJE sea una cuarta parte del ángulo OJP_0 (lo cual puede hacerse mediante doble bisección de este ángulo). Se determina un punto F tal que el ángulo FJE mida 45 grados (lo cual puede obtenerse por bisección de un ángulo recto).

(2) Se construye una circunferencia de diámetro FP_0. Esta circunferencia corta a OB en el punto K.

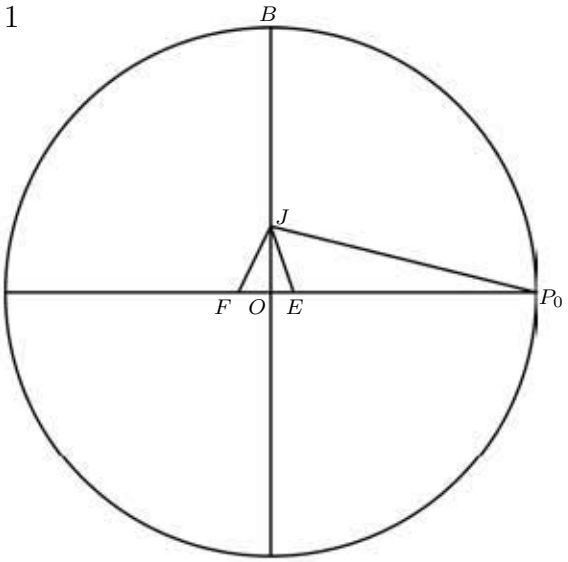
(3) Se traza otra circunferencia de centro E y radio EK. Esta circunferencia define los puntos N_5 y N_3.

(4) Se trazan las rectas N_3P_3 y N_5P_5, perpendiculares a OP_0.

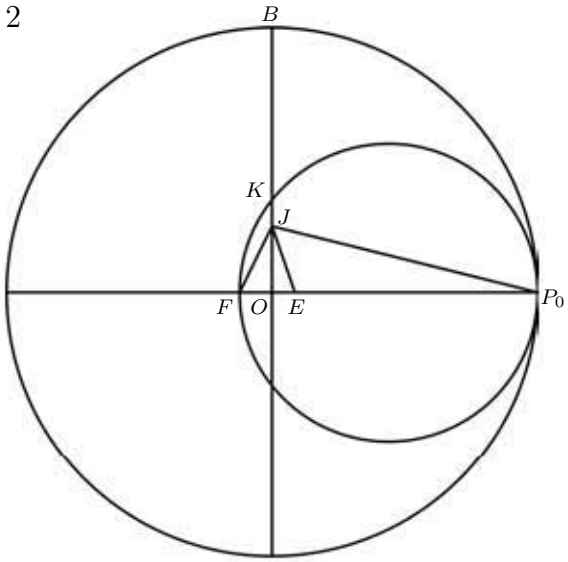
(5) Se traza la bisectriz del arco P_3P_5, a fin de obtener el punto P_4.

(6) Se lleva sucesivamente la cuerda P_4P_5 sobre la circunferencia, a partir de P_0. Finalmente, los puntos obtenidos se unen por segmentos rectilíneos para formar el polígono deseado.

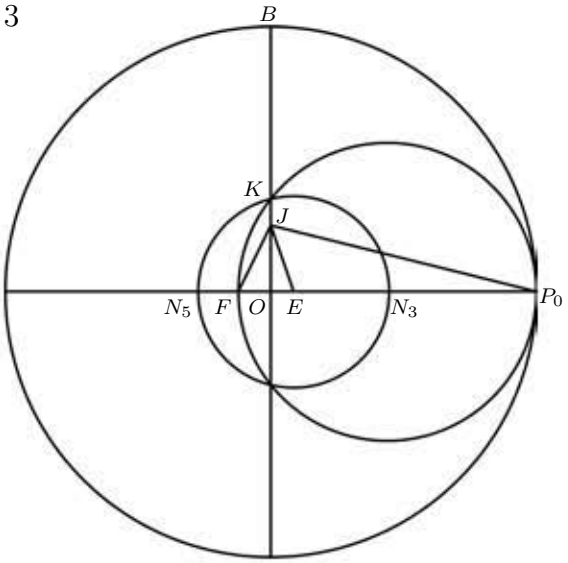
1


B
J
F
O
E
P0

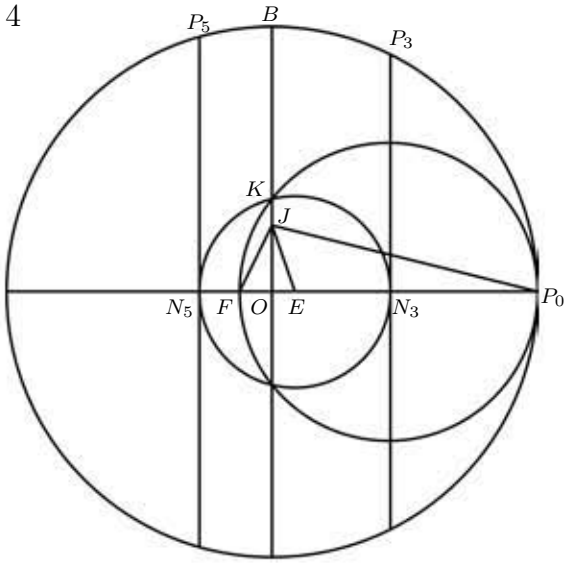
2


B
K
J
F
O
E
P0

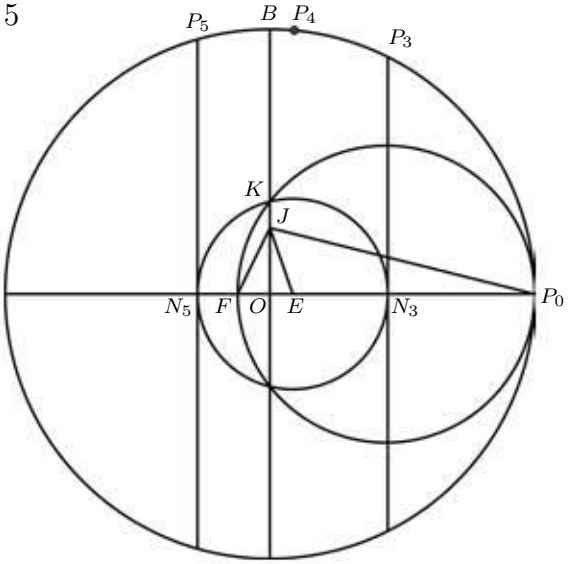
3


B
K
J
N5
F
O
E
N3
P0

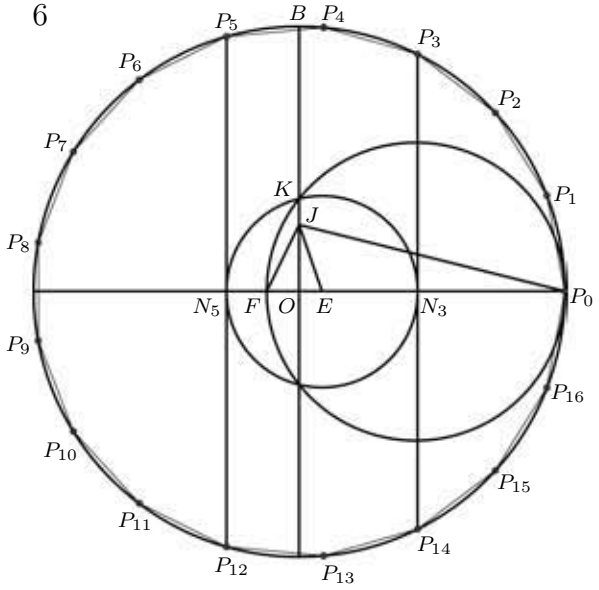
4


P5
B
P3
K
J
N5
F
O
E
N3
P0

5


P5
B
P4
P3
K
J
N5
F
O
E
N3
P0

6


P5
B
P4
P3
P6
P2
P7
K
P1
J
P8
N5
F
O
E
N3
P0
P9
P16
P10
P15
P11
P14
P12
P13

A.3.2. Construcción de polígonos regulares. Elementos de la teoría de Galois

En la sección anterior mostramos como resolver en raíces cuadradas la ecuación $x^{17}-1=0$. Ahora probaremos que para todos los números de la forma $2^n p_1 \cdots p_k$, donde los p_i son primos de Fermat distintos, la ecuación $x^n - 1 = 0$ también se puede resolver en raíces cuadradas. Nuestra exposición será de tal forma que un buen manejo de esta, pueda ser utilizada en el caso de la lemniscata casi sin cambios.

Asignando a cada número real t el punto con coordenadas $(\cos t, \operatorname{sen} t)$, obtendremos una parametrización de la circunferencia unitaria C mediante números reales. Como resultado, C resultará un grupo abeliano con elemento unitario $(1,0)$ con parámetro real t.

Ya que

$$\cos(t+s) = \cos t \cos s - \operatorname{sen} t \operatorname{sen} s \quad \text{y} \quad \operatorname{sen}(t+s) = \operatorname{sen} t \cos s + \cos t \operatorname{sen} s,$$

La ley de adición de puntos en la circunferencia puede ser expresada de la siguiente manera:

$$(a,b)+(c,d) = (ac-bd, ad+bc) = (f(a,b,c,d), g(a,b,c,d)).$$

Es fácil verificar que

$$2(x,y) = (x,y)+(x,y) = (x^2-y^2, 2xy)$$

y que

$$3(x,y) = (x^3 - 3xy^2, 3x^2y - y^3).$$

De forma análoga,

$$n(x,y) = (f_n(x,y), g_n(x,y)),$$

donde f_n y g_n son polinomios con coeficientes enteros. Ahora, usando la relación de Moivre: $\cos n\varphi + i \operatorname{sen} n\varphi = (\cos\varphi + i \operatorname{sen}\varphi)^n$ tenemos que

$$f_n(x,y) = \frac{(x+iy)^n + (x-iy)^n}{2}, \quad g_n(x,y) = \frac{(x+iy)^n - (x-iy)^n}{2i}. \tag{A.3}$$

Sea C_n el conjunto de puntos $(x,y) \in C$ tal que $n(x,y) = (1,0)$, es decir, $f_n(x,y) = 1$ y $g_n(x,y) = 0$. Estos puntos pueden servir como vértices de un n-ágono regular. Es claro que C_n es un subgrupo de C isomorfo a $\mathbb{Z}/n\mathbb{Z}$, el grupo aditivo de residuos modulo n.

Sobre $\mathbb{C}$, además de los puntos de C_n existen otras soluciones del sistema

$$f_n(x,y) = 1, \quad g_n(x,y) = 0.$$

Encontremos estas soluciones. Usando las fórmulas (A.3) podemos pasar a un sistema de ecuaciones equivalentes

$$(x+iy)^n = 1, \quad (x-iy)^n = 1.$$

Por lo tanto, $x + iy = \varepsilon^p$ y $x - iy = \varepsilon^q$ para $\varepsilon = \exp(2\pi i/n)$. En particular $x^2 + y^2 = \varepsilon^p\varepsilon^q$; de ahí que, la igualdad $x^2 + y^2 = 1$ se cumple si y sólo si $\varepsilon^p = \varepsilon^{-q}$. Por esta razón, C_n puede ser caracterizado como el conjunto de todas las soluciones del sistema de ecuaciones

$$f_n(x, y) = 1, \quad g_n(x, y) = 0, \quad x^2 + y^2 = 1. \tag{A.4}$$

Consideremos el campo K_n generado sobre $\mathbb{Q}$ por las coordenadas Cartesianas de todos los puntos de C_n. Por ejemplo, C_3 consiste en los puntos $(1, 0)$ y $(-\frac{1}{2}, \pm\frac{\sqrt{3}}{2})$; y, por lo tanto, $K_3 = \mathbb{Q}(\sqrt{3})$; C_4 consiste en los puntos $(\pm 1, 0)$ y $(0, \pm 1)$; de ahí que, $K_4 = \mathbb{Q}$.

Sea σ un automorfismo de K_n que sea la identidad sobre $\mathbb{Q}$. Puesto que los coeficientes de los polinomios f_n y g_n son enteros, σ manda cualquier solución del sistema (A.4) en otra solución del sistema, es decir, σ determina una permutación de los puntos de C_n. Cualquier automorfismo σ de estos puede ser recuperado de forma única de esta permutación por que las coordenadas de los puntos de C_n generan el campo K_n. También es claro que

$$\sigma((a, b) + (c, d)) = \sigma(ac - bd, ad + bc) = \sigma(a, b) + \sigma(c, d),$$

es decir, esta no es una permutación arbitraria de los puntos de C_n que corresponde a σ sino induce un automorfismo del grupo $\mathbb{Z}/n\mathbb{Z}$. Por lo tanto, el grupo G_n de automorfismos de K_n sobre $\mathbb{Q}$ es isomorfo a un subgrupo del grupo $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z})$.

La idea de la parte restante de la demostración es la siguiente. Primero, deberemos mostrar que si $n = 2^a p_1 \cdots p_k$, donde los p_i son primos de Fermat distintos, entonces el grupo $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z})$ es una potencia de 2. En particular, el orden de G_n también será una potencia de 2.

Después, deberemos probar que si el orden de un grupo G es igual a 2^k, entonces existe una sucesión de subgrupos

$$G = G^0 \supset G^1 \supset \cdots \supset G^k = \{e\}$$

de manera que G^i es un subgrupo de G^{i-1} de índice 2 para $i = 1, \ldots, k$.

Finalmente, usando esta sucesión de subgrupos construiremos una sucesión de extensiones cuadráticas de campo comenzando con $\mathbb{Q}$ y terminando con K_n (en particular, las coordenadas de los puntos C_n son irracionales cuadráticos), es decir, pueden ser construidos usando una sucesión de raíces cuadradas.

Lema A.3.1. *El orden del grupo* $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z})$ *es igual a una potencia de* 2 *si y sólo si* $n = 2^a p_1 \cdots p_k$*, donde los* p_i *son primos de Fermat distintos.*

Demostración. Como se sabe, todos los automorfismos del grupo aditivo $\mathbb{Z}/n\mathbb{Z}$ son de la forma $x \mapsto mx$, donde m es un entero que es primo relativo a n. Por lo tanto, el orden de $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z})$ es igual a $\varphi(n)$, donde $\varphi(n)$ es el número de enteros positivos menores a n y que son primos relativos con n. Si los números p y q son primos relativos, entonces $\varphi(pq) = \varphi(p)\varphi(q)$. En efecto, sean $0 \leq a \leq p - 1$ y $0 \leq b \leq q - 1$. Entonces *los residuos de*

la división entre n de los $n = pq$ números de la forma $aq + bp$, forman un sistema completo de residuos modulo n.

Para probar esto, es suficiente observar que $a_1q + b_1p \equiv a_2q + b_2p \pmod{pq}$, si y sólo si $(a_1 - a_2)q \equiv (b_2 - b_1)p \pmod{pq}$, es decir, $a_1 \equiv a_2 \pmod{p}$ y $b_1 \equiv b_2 \pmod{q}$. También es claro que los números $aq + bp$ y pq son primos relativos si y sólo si a y p y también b y q son primos relativos.

Si $n = p^k$, donde p es un primo, entonces $\varphi(n) = p^{k-1}(p-1)$. En efecto, de los números que no exceden a n solo los números $p, 2p, \ldots, p^{k-1}p$ tienen un común divisor con n. El número total de dichos números es p^{k-1}.

Sea $n = p_1^{k_1} \cdots p_m^{k_m}$. Entonces $\varphi(n)$ es el producto de los números $p_i^{k_i-1}(p_i-1)$. El número $p^{k-1}(p-1)$ puede ser una potencia de 2 solamente en los siguientes dos casos:

a) $p = 2$ y k es un entero positivo arbitrario;

b) $p - 1 = 2^c$ y $k = 1$.

En el segundo caso el número c no puede tener divisores impares. De hecho, si d es un divisor impar de c, entonces $2^c + 1 = (2^a)^d + 1$ es divisible entre $2^a + 1$. Por lo tanto, p es un primo de la forma $2^{2^l} + 1$, es decir, un primo de Fermat. □

Lema A.3.2. *Si el orden de un grupo G es igual a 2^k, entonces existe una sucesión de subgrupos*

$$G = G^0 \supset G^1 \supset \cdots \supset G^k = \{e\}$$

tal que G^i es un subgrupo de G^{i-1} de índice 2 para $i = 1, \ldots, k$.

Demostración. Apliquemos inducción sobre k. Para $k = 1$ la afirmación es obvia. Supongamos que la afirmación es verdadera para todos los grupos de orden 2^{k-1}. Para cada elemento $x \in G$ consideremos la clase de elementos conjugados a este, es decir, el conjunto $[gxg^{-1}]$ de elementos de la form gxg^{-1}, donde $g \in G$. Cualesquiera dos de estas clases o coinciden o son disjuntas, es decir, G esta dividido en la union de clase que no se intersecan de elementos conjugados. La igualdad $g_1xg_1^{-1} = g_2xg_2^{-1}$ es equivalente a la igualdad $xh = hx$, donde $h = g_2^{-1}g_1$. Consideremos el subgrupo

$$G_x = \{h \in G \mid xh = hx\}.$$

Los elementos $g_1xg_1^{-1}$ y $g_2xg_2^{-1}$ son iguales si y sólo si $g_1 \in g_2G_x$. Por lo tanto, el número de elementos en la clase $[gxg^{-1}]$ es igual al índice del subgrupo G_x en G; de ahí que este sea de la forma 2^s.

La clase $[gxg^{-1}]$ contiene exactamente un elemento sólo si x conmuta con todos los elementos de G, es decir, x es un elemento del centro de G. Supongamos que el centro de G consiste solamente del elemento unitario. Entonces la suma de cardinalidades de todas las clases conjugadas es igual a $1 + 2^{s_1} + \cdots + 2^{s_p}$, donde $s_i \geq 1$. Por esta razón, esta suma es

un número impar. Por otra parte, esta suma es igual al orden de G, es decir, es igual a 2^k. Esta contradicción implica que el centro de G tiene un elemento $a \neq e$.

El elemento a genera un subgrupo cíclico de orden 2^r. Consideremos el elemento $b = a^m$, donde $m = 2^{r-1}$. El elemento b genera un subgrupo H de orden 2 que pertenece al centro de G; en particular, H es un subgrupo normal. Por la hipótesis de inducción para el grupo $F = G/H$ de orden 2^{k-1}, existe una sucesión

$$F = F^0 \supset F^1 \supset \cdots \supset F^{k-1} = \{e\},$$

donde F^i es un subgrupo de F^{i-1} de índice 2 para $i = 1, \ldots, k-1$. Para obtener la sucesión de subgrupos requerida, definamos $G^k = e$ y $G^i = F^i \cup bF^i$ para $i = 0, \ldots, k-1$. □

Sea $n = 2^a p_1 \cdots p_m$, donde p_i son primos de Fermat distintos. Entonces el grupo G_n de automorfismo del campo K_n sobre $\mathbb{Q}$ tiene una sucesión de subgrupos

$$G_n = G^0 \supset G^1 \supset \cdots \supset G^k = \{e\},$$

donde G^i es un subgrupo de G^{i-1} de índice 2 para $i = 1, \ldots, k$. Para el subgrupo G^i asignaremos el conjunto L^i que consista de los elementos de K_n que estén fijos bajo la acción de todos los automorfismos en G^i. Como la suma, resta, producto y cociente de elementos de L^i pertenece a L^i, se sigue que L^i es un campo. Es claro que $L^k = K_n$ y que $L^i \supset L^{i-1}$.

Podemos mostrar que $L^0 = \mathbb{Q}$, es decir, para cualquier $x \in K_n/\mathbb{Q}$ existe un automorfismo del campo K_n sobre $\mathbb{Q}$ que mueve x. En general, no toda extensión de $\mathbb{Q}$ posee tal propiedad. Por ejemplo, el campo

$$\{p + q\sqrt[3]{2} + r\sqrt[3]{4} \mid p, q, r \in \mathbb{Q}\}$$

no tiene automorfismos distintos de la identidad. En efecto, el elemento $\sqrt[3]{2}$ solo puede obtenerse como raíz de la ecuación $x^3 - 2 = 0$, pero sólo una raíz de esta ecuación pertenece al campo considerado.

La razón por la cual $L^0 = \mathbb{Q}$ será dada un poco más tarde. Por el momento asumamos esto sin una demostración.

Todo automorfismo en G^i preserva los elementos de L^i. Además, $G^{i-1} = G^i \cup \sigma G^i$, donde $\sigma \in G^{i-1}$. De ahí que, $\sigma^2 \in G^i$ por que σ^2 no puede pertenecer a σG^i. Por lo tanto, el automorfismo σ de K_n es tal que si $x \in L^i$, entonces $\sigma^2 x = x$. Por otra parte, $\sigma x = x$ si y sólo si $x \in L^{i-1}$. Cualquier elemento $x \in L^i$ puede ser representado como la suma de los elementos $x_1 = \frac{1}{2}(x + \sigma x)$ y $x_2 = \frac{1}{2}(x - \sigma x)$, donde $\sigma x_1 = x_1$ y $\sigma x_2 = -x_2$. Por lo tanto, $x_1 \in L^{i-1} \subset L^i$ y $x_2 = x - x_1 \in L^i$.

Supongamos que $L^i \neq L^{i-1}$. Sean $a \in L^i \setminus L^{i-1}$ y $\alpha = \sigma a - a$. Entonces, $\sigma\alpha = -\alpha$, donde $\alpha \neq 0$. Además, $\sigma(\alpha x_2) = (-\alpha)(-x_2) = \alpha x_2$, es decir, $\alpha x_2 \in L^{i-1}$ y $x_2 \in \alpha^{-1} L^{i-1}$. Por lo tanto, $L^i = L^{i-1} + \alpha^{-1} L^{i-1}$. De ahí que, si $x \in L^i$, entonces los elementos $1, x, x^2$ son linealmente dependientes sobre L^{i-1}; por lo tanto, $x^2 + px + q = 0$, donde $p, q \in L^{i-1}$. De esto se sigue que cualquier elemento de K_n es un irracional cuadrático sobre L^0.

Ahora probemos que $L^0 = \mathbb{Q}$. Primero, observemos que K_n es una extensión algebraica de $\mathbb{Q}$, es decir, las coordenadas de los puntos de C_n son números algebraicos. De hecho, las soluciones del sistema de ecuaciones

$$f_n(x, y) = 1, \quad g_n(x, y) = 0$$

son algebraicas; para probar esto, es suficiente considerar $f_n(x, y) - 1$ y $g_n(x, y)$ como polinomios de y y examinar sus resultantes.

Sea $\overline{\mathbb{Q}}$ el conjunto de todos los números algebraicos (sobre $\mathbb{Q}$). Es fácil verificar que $\overline{\mathbb{Q}}$ es un campo. De hecho, sea $\alpha, \beta \in \overline{\mathbb{Q}}$. Entonces $\alpha^p = a_0 + a_1\alpha + \cdots + a_{p-1}\alpha^{p-1}$ y $\beta^q = b_0 + b_1\beta + \cdots + b_{q-1}\beta^{q-1}$, donde $a_i, b_j \in \mathbb{Q}$ y $a_0 b_0 \neq 0$. Por lo tanto, cualquier elemento de un anillo generado sobre $\mathbb{Q}$ por elementos α y β puede ser representado en la forma de una combinación lineal con coeficientes racionales de elementos $\alpha^i\beta^j$, donde $0 \leq i < p$ y $0 \leq j < q$. En particular, cada uno de los elementos $1, \alpha + \beta, \ldots, (\alpha + \beta)^{pq}$ puede ser representado como una combinación lineal de los pq elementos indicados; por esta razón, estos elementos son linealmente dependientes sobre $\mathbb{Q}$, es decir, $\alpha + \beta \in \overline{\mathbb{Q}}$.

De manera similar se demuestra que $\alpha\beta \in \overline{\mathbb{Q}}$. Además, $a_0\alpha^{-1} = \alpha^{p-1} - a_1 - \cdots - a_{p-1}\alpha^{p-2}$; de ahí que, $\alpha^{-1} \in \overline{\mathbb{Q}}$.

Cualquier automorfismo τ del campo $\overline{\mathbb{Q}}$ sobre $\mathbb{Q}$ manda K_n en si mismo. De hecho, el campo K_n esta generado por coordenadas de puntos de C_n y C_n coincide con el conjunto de todas las soluciones (sobre $\mathbb{Q}$) del sistema de ecuaciones

$$f_n(x, y) = 1, \quad g_n(x, y) = 0, \quad x^2 + y^2 = 1.$$

La igualdad $L^0 = \mathbb{Q}$ significa que

si $a \in K_n \setminus \mathbb{Q}$, entonces existe un automorfismo $\sigma : K_n \to K_n$ para el cual $\sigma(a) \neq a$.

Para probar esto, es suficiente exhibir un automorfismo $\tau : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$ para el cual $\tau(a) \neq a$.

Teorema A.3.3. *Sean a y b dos raíces de un polinomio irreducible sobre $\mathbb{Q}$. Entonces existe un automorfismo $\tau : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$ tal que $\tau(a) = b$.*

Demostración. Sean K un campo, α una raíz de un polinomio irreducible P sobre K, y $K(\alpha)$ el campo generado por α sobre K. Entonces un isomorfismo arbitrario $f : K \to K'$ puede ser extendido a un isomorfismo $g : K(\alpha) \to K'(\beta)$, donde β es una raíz del polinomio $f(P)$. En efecto, el campo $K(\alpha)$ consiste en los elementos de la forma $\sum k_j\alpha^j$, donde $j \geq 0$ y $k_j \in K$. Definamos $g(\sum k_j\alpha^j) = \sum f(k_j)\beta^j$. Esta función queda bien definida por que la igualdad $\sum k_j\alpha^j = 0$ es equivalente al hecho de que el polinomio $F = \sum k_j x^j$ es divisible entre P.

Primero, con lo anterior, construyamos un isomorfismo $\tau_1 : \mathbb{Q}(a) \to \mathbb{Q}(b)$. Entonces seleccionando un elemento $t_2 \in \overline{\mathbb{Q}} \setminus \mathbb{Q}(a)$. Este elemento es una raíz de un polinomio irreducible P_2 sobre $\mathbb{Q}(a)$. Sea t_2' una raíz del polinomio $\tau_1(P_2)$.

Después, nuevamente por lo descrito en el primer párrafo podemos construir un isomorfismo $\tau_2 : \mathbb{Q}(a, t_2) \to \mathbb{Q}(b, t_2')$. Seleccionando un elemento $t_3 \in \overline{\mathbb{Q}} \setminus \mathbb{Q}(a, t_2)$ y construyendo un isomorfismo $\tau_3 : \mathbb{Q}(a, t_2, t_3) \to \mathbb{Q}(b, t_2', t_3')$, etcetera. Puesto que la dimensión de $\overline{\mathbb{Q}}$ sobre $\mathbb{Q}$ es numerable, podemos construir una base $\{1, \varepsilon_1 = a, \varepsilon_2, \varepsilon_3, \ldots\}$ de $\overline{\mathbb{Q}}$ sobre $\mathbb{Q}$. Los elementos $t_2, t_3, \ldots$ pueden escogerse tal que el campo $\mathbb{Q}(a, t_2, \ldots, t_k)$ contenga al subespacio generado por los elementos $1, \varepsilon_1, \ldots, \varepsilon_k$.

Como resultado, obtenemos un monomorfismo $\tau : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$ tal que $\tau(a) = b$. Falta verificar que τ es un epimorfismo. Sea $\gamma_1 \in \overline{\mathbb{Q}}$ una raíz de un polinomio irreducible R sobre $\mathbb{Q}$ y sean $\gamma_1, \ldots, \gamma_n$ todas las raíces de este polinomio. Entonces $\tau(\gamma_i) \in \{\gamma_1, \ldots, \gamma_n\}$, donde todos los números $\tau(\gamma_1), \ldots, \tau(\gamma_n)$ son distintos. Por esta razón, los conjuntos $\{\gamma_1, \ldots, \gamma_n\}$ y $\{\tau(\gamma_1), \ldots, \tau(\gamma_n)\}$ coinciden. En particular, $\gamma_1 = \tau(\gamma_j)$ para alguna j. □

Es posible generalizar significativamente el Teorema A.3.3. Primero observemos que $\overline{\mathbb{Q}}$ es cerrado algebraicamente. De hecho, sea x_0 una raíz del polinomio $\alpha + \beta x + \cdots + \omega x^n = 0$, donde $\alpha, \beta, \ldots, \omega \in \overline{\mathbb{Q}}$. Consideremos un polinomio R irreducible sobre $\mathbb{Q}$ y con una raíz α. Sean $\alpha_1, \ldots, \alpha_p$ todas las raíces de R. Entonces todo polinomio simétrico elemental de $\alpha_1, \ldots, \alpha_p$ puede ser expresado en términos de los coeficientes de R y, por lo tanto, son racionales. De forma similar definamos $\beta_1, \ldots, \beta_q; \ldots; \omega_1, \ldots, \omega_r$ y considerando el polinomio

$$P(x) = \prod_{i,j,\ldots,k} (\alpha_i + \beta_j x + \cdots + \omega_k x^n).$$

Este polinomio es distinto de cero y todos sus coeficientes pueden ser expresados en términos de los polinomios simétricos elementales $\sigma_s(\alpha_1, \ldots, \alpha_p), \ldots, \sigma_t(\omega_1, \ldots, \omega_r)$; de ahí que, sus coeficientes sean racionales. Como $P(x_0) = 0$, se sigue que $x_0 \in \overline{\mathbb{Q}}$.

En el caso general la siguiente proposición se cumple para los automorfismos de un campo algebraicamente cerrado Ω sobre su subcampo K.

Teorema A.3.4. *Si los elementos $x, y \in \Omega$ son trascendentes sobre K, entonces existe un automorfismo de ω sobre K que manda a x en y. Si los elementos $x, y \in \Omega$ son raíces de un mismo polinomio irreducible sobre K, entonces existe un automorfismo de Ω sobre K que manda a x en y.*

Nosotros no probaremos este teorema para campos arbitrarios, pero si para el caso más interesante –los automorfismos de $\mathbb{C}$ sobre $\mathbb{Q}$– pero aún más no probaremos solamente este teorema, sino muchas de sus generalizaciones. Por ejemplo, probaremos que la cardinalidad del conjunto de automorfismos de $\mathbb{C}$ coincide con la cardinalidad del conjunto de todos los mapeos de $\mathbb{C} \to \mathbb{C}$ (es decir, esta cardinalidad es más grande que la cardinalidad de $\mathbb{C}$).

Primero, observemos que un isomorfismo de campos $\varphi : F \to G$ puede ser extendido a un isomorfismo $\varphi' : F(\alpha) \to G(\beta)$ si y sólo si las siguientes condiciones se cumplen:

1. Si un elemento α es algebraico sobre F y P es un polinomio irreducible sobre F con un raíz α, entonces β es una raíz del polinomio $\varphi(P)$;

2. Si α es trascendente sobre F, entonces β es trascendente sobre G.

Para nuestros argumentos necesitaremos el lema de Zorn. El punto es que hacer pruebas por inducción solo es aplicable a conjuntos numerables mientras que la dimensión de $\mathbb{C}$ sobre $\mathbb{Q}$ no es numerable. Por lo tanto, para trabajar con automorfismos de $\mathbb{C}$ sobre $\mathbb{Q}$ necesitaremos otra técnica y el lema de Zorn es suficientemente conveniente para este propósito.

Antes de formular el lema de Zorn, daremos algunas definiciones. Sea g un conjunto. Denotemos por 2^g el conjunto de todos los subconjuntos de g. Un conjunto $A \subset 2^g$ es llamada una ***cadena*** si para cualquier par de sus elementos $a, b \in A$ tenemos que $a \subset b$ o que $b \subset a$ (recordemos que a y b son subconjuntos de un mismos conjunto g). Un conjunto $B \subset 2^g$ es llamado ***Zorn cerrado*** si para cualquier cadena $A \subset B$ el conjunto B también contiene la union de todos los elementos de A. Un elemento $m \in B$ es llamado ***maximal*** si el conjunto $m \subset g$ no esta contenido en cualquier otro subconjunto $a \subset g$ el cual es un elemento de B (es decir, $a \in B$).

Lema de Zorn. *Cualquier conjunto $B \subset 2^g$ que sea Zorn cerrado y distinto del vacío contiene al menos un elemento maximal m.*

Con la ayuda del lema de Zorn podemos extender cualquier automorfismo φ de un subcampo de $\mathbb{C}$ a un automorfismo de todo el campo $\mathbb{C}$. Para esto, tenemos que aplicar el lema de Zorn a la familia de automorfismos que extiende φ. Pero aquí hay una dificultad. Puede pasar que una extensión de un automorfismo del campo F al campo F' que contenga a F no deje F' invariante, tal extensión del automorfismo no resultará un automorfismo de F' sino un isomorfismo de F' con algún otro campo. Por ejemplo, el automorfismo del campo $\mathbb{Q}(\sqrt{2})$ dado por la fórmula $a + b\sqrt{2} \mapsto a - b\sqrt{2}$. Su extensión a $\mathbb{Q}(\sqrt[4]{2})$ es como sigue:

$$a + b\sqrt[4]{2} + c\sqrt{2} + d\sqrt[4]{8} \mapsto a + ib\sqrt[4]{2} - c\sqrt{2} - id\sqrt[4]{8}.$$

Este mapeo es un isomorfismo de $\mathbb{Q}(\sqrt[4]{2})$ a $\mathbb{Q}(i\sqrt[4]{2})$ pero no un automorfismo de $\mathbb{Q}(\sqrt[4]{2})$.

Para superar esta dificultad primero probemos la siguiente proposición.

Teorema A.3.5. *Cualquier isomorfismo de campos $\varphi : F \to G$ puede ser extendida a un isomorfismo de cerraduras algebraicas $\overline{F} \to \overline{G}$.*

Demostración. Consideremos todas las posibles extensiones del isomorfismo φ a un isomorfismo $\varphi_\alpha : F_\alpha \to G_\alpha$, donde $F_\alpha \subset \overline{F}$ y, por lo tanto, $G_\alpha \subset \overline{G}$. Mostremos que el conjunto

$$S = \{\text{los subconjuntos de } \overline{F} \times \overline{G} \text{ de la forma } \{(a, \varphi_\alpha(a)) \,|\, a \in F_\alpha\}\}$$

es Zorn cerrado. Considerando una cadena arbitraria en S. Por la definición de una cadena, para cualesquiera dos de sus elementos los isomorfismos correspondientes φ_α y φ_β son tales que uno de estos isomorfismos es una extensión del otro. Esto significa que para la unión de todos los elementos de la cadena corresponde un isomorfismo, es decir, su unión pertenece al conjunto considerado.

Por el lema de Zorn el conjunto S tiene un elemento maximal. A este elemento le corresponde un isomorfismo $\psi : F' \to G'$, entonces debemos probar que $F' = \overline{F}$ y que $G' = \overline{G}$. Supongamos que un elemento $a \in F$ no pertenece a F'. Pero a es algebraico sobre F' y $\overline{G}$ es cerrado algebraicamente. Por lo tanto, $\overline{G}$ contiene un elemento b que es la raíz de la imagen (bajo ψ) del polinomio minimal de a sobre F'. Por esta razón, es posible extender este isomorfismo $\psi : F' \to G'$ a un isomorfismo $F'(a) \to G'(b)$, pero este contradice la maximalidad del elemento correspondiente a ψ.

Por lo tanto, $F' = \overline{F}$. Falta probar que $G' = \overline{G}$. El campo G' es isomorfo a $\overline{F}$; de ahí que, G' es en si mismo cerrado algebraicamente. En adición, G' contienes a G. Por lo tanto, $G' = \overline{G}$. □

Ahora podemos probar el teorema sobre extensiones de automorfismos de subcampos de $\mathbb{C}$.

Teorema A.3.6. *Cualquier automorfismo φ de un subcampo de $\mathbb{C}$ puede extenderse a un automorfismo de $\mathbb{C}$.*

Demostración. Consideremos todas las posibles extensiones del automorfismo dado $\varphi : F \to F$ al automorfismo $\varphi_\alpha : F_\alpha \to F_\alpha$, donde $F_\alpha \subset \mathbb{C}$. Como en la demostración del Teorema A.3.5, vemos que el conjunto consistente de conjuntos de la forma $\{(a, \varphi_\alpha(a)) \,|\, a \in F_\alpha\}$ tiene un elemento maximal. A este elemento le corresponde un automorfismo $\varphi' : F' \to F'$. Debemos probar que $F' = \mathbb{C}$.

Supongamos que un número complejo a no pertenece a F'. Si a es algebraico sobre F', entonces por el Teorema A.3.5 existe una extensión de φ' a un automorfismo de la cerradura algebraica de F' y esta cerradura es estrictamente mayor que F'. Si a es trascendental sobre F', entonces existe una extensión de φ' a un isomorfismo $F'(a) \to F'(a)$ que manda a en a. En ambos casos obtenemos una contradicción con la maximalidad de F'. Por lo tanto, $F' = \mathbb{C}$. □

Observemos que no siempre es posible extender un isomorfismo de dos subcampos de $\mathbb{C}$ a un automorfismo en $\mathbb{C}$. Por ejemplo, existe un isomorfismo $\mathbb{C} \to F \subset \mathbb{C}$, donde $F \neq \mathbb{C}$. Tal isomorfismo es construido de la siguiente manera. Sea $a_1, a_2, \ldots$ un conjunto numerable de números complejos algebraicamente independientes sobre $\mathbb{Q}$. El mapeo $a_i \mapsto a_{i+1}$ determina un isomorfismo

$$\mathbb{Q}(a_1, a_2, \ldots) \to \mathbb{Q}(a_2, a_3, \ldots) \subset \mathbb{Q}(a_1, a_2, \ldots).$$

Considerando todas las posibles extensiones de este isomorfismo al isomorfismo $\theta_\alpha : F_\alpha \to G_\alpha (F_\alpha, G_\alpha \subset \mathbb{C})$ tal que a_1 es trascendental sobre G_α.

Por el lema de Zorn el conjunto $\{F_\alpha\}$ tiene un elemento maximal el cual, como es fácil mostrar, coincide con $\mathbb{C}$. Por lo tanto, obtenemos un isomorfismo $\mathbb{C} \to F \subset \mathbb{C}$, donde el campo F no contiene a a_1 y, por lo tanto, $F \neq \mathbb{C}$.

La demostración del Teorema A.3.4 para el caso de automorfismos de $\mathbb{C}$ sobre $\mathbb{Q}$ ya no es un problema. De hecho, si los números complejos x y y son trascendentes, podemos considerar

un automorfismo del campo $\mathbb{Q}(x, y)$ que intercambie x con y. Por el Teorema A.3.6 es posible extender este isomorfismo a un automorfismo del campo $\mathbb{C}$. Si x y y son raíces del mismo polinomio irreducible sobre $\mathbb{Q}$, entonces existe un isomorfismo de campos $\mathbb{Q}(x) \to \mathbb{Q}(y)$ que manda x en y. Por el teorema A.3.5 este isomorfismo puede extenderse a un isomorfismo de las cerraduras algebraicas de $\mathbb{Q}(x)$ y $\mathbb{Q}(y)$. Pero las cerraduras algebraicas de estos campos coinciden y, por lo tanto, obtenemos no solo un isomorfismo, sino un automorfismo. Este automorfismo puede extenderse a un automorfismo de $\mathbb{C}$.

Ahora es claro que la cardinalidad del conjunto de automorfismos de $\mathbb{C}$ no es menor que la cardinalidad del continuo. Resulta que la cardinalidad del conjunto de todos los automorfismos de $\mathbb{C}$ es, de hecho, mayor que la cardinalidad del continuo.

Teorema A.3.7. *La cardinalidad del conjunto de todos los automorfismos de* $\mathbb{C}$ *coincide con la cardinalidad de todos los mapeos de* $\mathbb{C} \to \mathbb{C}$.

Demostración. Tenemos que probar que la cardinalidad de todos los automorfismos de $\mathbb{C}$ coincide con la cardinalidad del continuo. Basta probar que la cardinalidad de todos los automorfismos de $\mathbb{C}$ no es menor que la cardinalidad del conjunto de todos los subconjuntos del continuo. Un conjunto $B \subset \mathbb{C}$ es llamada una ***base de trascendentalidad*** sobre $\mathbb{Q}$ si B es algebraicamente independiente sobre $\mathbb{Q}$ y B no esta contenido en cualquier otro conjunto de números complejos algebraicamente independientes sobre $\mathbb{Q}$. La maximalidad de B implica que $\mathbb{C}$ es algebraico sobre $\mathbb{Q}(B)$. Por lo tanto, en particular, la cardinalidad de B es igual que la del continuo.

Mostremos que a cualquier subconjunto $S \subset B$ le podemos asignar un automorfismo φ_S de $\mathbb{C}$ tal que a conjuntos diferentes les corresponden automorfismos diferente.

Consideremos un automorfismo de $\mathbb{Q}(B)$ sobre $\mathbb{Q}$ que manda $x \in B$ a x si $x \in S$ y a $-x$ si $x \notin S$. Por el Teorema A.3.6 este automorfismo puede extenderse a un automorfismo φ_S de todo $\mathbb{C}$. Claramente, si x pertenece a uno de los conjuntos S o T y no pertenece al otro conjunto, entonces $\varphi_S(x) = -\varphi_T(x)$ y, por lo tanto $\varphi_S \neq \varphi_T$. □

Bibliografía

[1] R. Ayoub, *The Lemniscate and Fagnano's contributions to elliptic integrals,* Arch. Hist. Exact Sci. **29**, pág. 131-149, 1984.

[2] C. Sánchez y T. Noriega, *ABEL, El romántico nórdico,* Nivola, 2005.

[3] D. A. Cox, *Galois Theory,* Wiley Intercience, 2004.

[4] I. N. Herstein, *Topics in Algebra,* Second Edition, Wiley, 1975.

[5] G. A. Jones y D. Singerman, *Complex Functions: An Algebraic and Geometric Viewpoint,* Cambridge U. P., 1987.

[6] A. I Markushévich, *Funciones Maravillosas,* Mir, 1977.

[7] J. E. Marsden y M. J. Hoffman, *Basic Complex Analisis,* Third Edition, W. H. Freeman, 1999.

[8] V. Prosolov y Y. Solovyev, *Elliptic Functions and Elliptic Integrals,* Translations of Mathematical Monographs, Volume 170, AMS, 1997.

[9] R. Remmert, *Theory of Complex Functions,* Graduate Texts in Mathematics, Readings in Mathematics, Volume 122, 1999.

[10] M. Rosen, *Abel's theorem on the lemniscate,* Amer. Math. Monthly **88**, pág. 387-395, 1981.

[11] C. L. Siegel, *Topics in Complex Function Theory,* Vol I, Wiley, 1969.

[12] I. Stewart, *Gauss,* Scientific American **237**, pág. 122-131, 1977.