TRABAJO FIN DE GRADO
FUNCIONES ELÍPTICAS

FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ANÁLISIS MATEMÁTICO

Realizado por: María Vargas Magán
Dirigido por: Juan Carlos García Vázquez

Sevilla, 2023

Mis ideas están basadas en mi asombro
y admiración por las leyes contenidas
en el mundo que nos rodea.
Quien se maravilla de algo,
toma consciencia de algo maravilloso.

M.C ESCHER

Índice

Resumen

La finalidad de este trabajo es introducirnos en la teoría de las funciones elípticas, con una mirada clásica desde la teoría de la variable compleja, para luego establecer su relación con las curvas elípticas y formas modulares.

Aunque estén íntimamente relacionados, hemos dividido el trabajo en tres capítulos diferenciados: funciones elípticas, curvas elípticas y formas modulares. En cada capítulo hemos expuesto las propiedades más destacadas de cada objeto matemático. Para ser más concretos, en el primer capítulo veremos como Jacobi entiende las funciones elípticas como cociente de funciones *casi* elípticas. Weierstrass, por su parte, las entiende como funciones racionales de sólo dos funciones elípticas. En el segundo capítulo estudiaremos como es posible identificar una curva elíptica definida sobre $\mathbb{C}$ con un toro y como toda curva elíptica queda determinada por una única función: la función J-invariante. Además, veremos que es posible definir una estructura de grupo en la curva mediante la suma de números complejos. En el tercer capítulo, estudiaremos la geometría subyacente al grupo modular y veremos como sus elementos pueden entenderse como una generalización de las funciones elípticas. También estudiaremos la importancia de la función J entendida como función modular, obteniendo de nuevo resultados relevantes acerca de esta función. Finalizaremos el trabajo enunciando el teorema de Modularidad, teorema que simplifica la relación entre nuestros tres objetos protagonistas.

Abstract

The purpose of this work is to introduce us to the theory of elliptic functions, with a classical look from the theory of the complex variable, to then establish its relationship with elliptic curves and modular forms.

Although they are closely related, we have divided the work into three differentiated chapters: elliptic functions, elliptic curves and modular forms. In each chapter we have explained the most prominent properties of each mathematical object. To be more concrete, in the first chapter we will see how Jacobi understands elliptic functions as quotient of *almost* elliptical functions. Weierstrass, on the other hand, understands them as rational functions of only two elliptic functions. In the second chapter we will study how it is possible to identify an elliptic curve defined on $\mathbb{C}$ with a torus and we will also see that any elliptic curve is determined by a unique function: the J-invariant function. In addition, we will see that it is possible to define a group structure on the curve by adding complex numbers. In the third chapter, we will study the geometry underlying the modular group and we will see how its elements can be understood as a generalization of elliptic functions. We will also study the importance of the function J understood as a modular function, obtaining relevant results about this function. We will finish the work formulating the Modularity Theorem, theorem that simplifies the relationship between our three main objects.

Introducción histórica

Con el objetivo de entender de la forma más completa el contenido de este trabajo, comprendamos primero el marco histórico en el que se sitúa la teoría de las funciones elípticas. Para ello, nos tenemos que remontar al siglo XVII. Su motivación, impulsada por Wallis (1616-1703) y Newton (1643-1727), la encontramos en algunos problemas diferenciales, como el movimiento que describe un péndulo o en problemas geométricos, como la rectificación de arcos o el estudio de curvas como la elipse o la lemniscata, cuyas longitudes de arco no se pueden expresar por funciones elementales, estas son racionales, trigonométricas, exponenciales o logarítmicas. Por ello, uno de los objetivos era incorporar una nueva clase de funciones, extendiendo así el horizonte del análisis matemático. Es aquí donde interviene Gauss, definiendo las funciones lemniscáticas. Aunque desarrolló el campo de los números complejos y la aritmética modular, priorizó resultados que no eran relativamente novedosos, quedando parte de sus descubrimientos en el campo de las funciones elípticas eclipsado por los avances de Abel y Jacobi.
Por este motivo, históricamente, el origen de dichas funciones se atribuye a Abel (1802-1829) y a Jacobi (1804-1851). Abel, con la ayuda de Legendre, fue el precursor definiendo en *Recherches sur les fonctions elliptiques* (1827) las funciones elípticas como inversas de integrales elípticas. De este documento destacan los teoremas de adición de funciones elípticas y la extensión de tales funciones a números puramente imaginarios.

Tras la muerte de Abel, Jacobi tomó protagonismo en el campo de las funciones elípticas. En 1829 publicó la obra *Fundamenta nova functionum ellipticarum*, donde estudia la doble periodicidad de las funciones elípticas a través de las funciones theta. La obra de Jacobi fue acogida con agrado, convirtiendo a estas funciones en un objeto de gran interés. Sin embargo, el avance del análisis matemático ensombreció las funciones theta de Jacobi. En esta línea, destaca el estudio de Eisenstein (1823-1852) sobre formas modulares definidas por series explícitas, permitiendo a Weierstrass presentar sus funciones $\wp(z)$ y $\wp'(z)$, funciones que nos permiten caracterizar todas las

funciones elípticas.

Estas funciones de Weierstrass son el detonante que marca el rumbo del desarrollo de las funciones elípticas hasta la actualidad. En esta línea, lo más reciente y notable fue el teorema de Taniyama-Shimura (2001), que establece una relación biunívoca entre las formas modulares y las curvas elípticas y como consecuencia, prueba el último teorema de Fermat.

Actualmente, las aplicaciones de las funciones elípticas gozan de una gran diversidad que va desde la demostración del último teorema de Fermat hasta sus usos en sistemas criptográficos, especialmente en aquellos sistemas en los que se apoyan las transacciones financieras. Además, las funciones elípticas, con las curvas elípticas en particular, han permitido el origen de nuevos conceptos matemáticos, como es el caso de la teoría de las formas modulares. Es más, el avance de las formas modulares ha permitido a su vez el desarrollo de la teoría de Galois. En este sentido, las funciones elípticas son un punto de mira de todas las ramas de las matemáticas, pues no sólo abarca la variable compleja, sino también la geometría, el álgebra o la teoría de números.

Capítulo 1

Funciones elípticas

A lo largo de este trabajo nos apoyaremos en resultados dados en la asignatura Funciones de una Variable Compleja. Por destacar algunos de ellos, usaremos el teorema de Liouville, el de los residuos o el principio del argumento, así como sus corolarios más conocidos, entre otros. También vimos una introducción a las funciones elípticas, con algunos resultados que incluimos en este trabajo.

En este capítulo daremos una serie de definiciones necesarias para entender qué es una función elíptica. Asimismo, expondremos resultados relacionados con el número de ceros o de polos de tales funciones. Por seguir el orden cronológico, estudiaremos las funciones elípticas de Jacobi y las contrastaremos con las de Weierstrass, reflejando además el avance del estudio de las funciones elípticas a lo largo de los años.

Para redactar la primera sección de este capítulo nos hemos basado principalmente en [Mar70]. De la misma manera, nos hemos apoyado sobre todo en [WW28] y en [SS03] para la segunda y tercera sección respectivamente.

1.1 Teoría de las funciones elípticas

Definición 1.1. *Decimos que una función f meromorfa en $\mathbb{C}$ es **periódica** si existe $w \in \mathbb{C}$ tal que*

$$f(z) = f(z + w)$$

*para todo $z \in \mathbb{C}$. A w se le llama **período** de f.*

El siguiente lema nos clasifica los distintos y excluyentes tipos de periodicidad que puede tener una función periódica.

Lema 1.1.1. ***(Jacobi)*** *Sea f una función periódica no constante y Λ el conjunto formado por todos sus períodos. Entonces puede ocurrir las siguientes posibilidades*

1. *Existe $w \in \mathbb{C}$ tal que $\Lambda = \{nw : n \in \mathbb{Z}\}$.*
2. *Existen w_1 y w_2 en $\mathbb{C}$ linealmente independientes tal que $\Lambda = \{nw_1 + mw_2 : n, m \in \mathbb{Z}\}$. En este caso, diremos que Λ es el **retículo** en $\mathbb{Z}$ generado por w_1 y w_2.*

*En el primer caso decimos que f es **simplemente periódica** y en el segundo decimos que f es **doblemente periódica o elíptica**. Notemos que, al ser $dim_{\mathbb{R}}\mathbb{C} = 2$, dos períodos es el límite.*

La prueba de este resultado se puede ver en el séptimo capítulo del libro [Mar70].

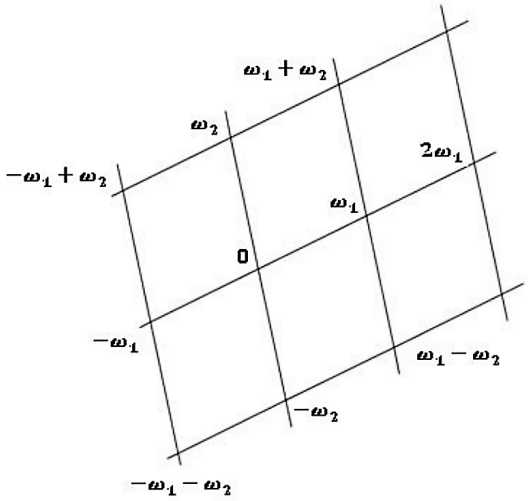

Figura 1.1: Retículo generado por los períodos w_1 y w_2.

Por tanto, siguiendo la definición de Abel, hemos definido las **funciones elípticas** como funciones meromorfas doblemente periódicas. Las llamamos elípticas porque, como veremos al final del capítulo, son las inversas de las funciones definidas por integrales elípticas.

En lo que sigue, supondremos por convenio que para una función elíptica f no constante, los períodos w_1 y w_2 son **mínimos**, en el sentido de que el retículo

$$\Lambda = \{nw_1 + mw_2 : n, m \in \mathbb{Z}\}$$

es el conjunto formado por todos los posibles períodos.

Definición 1.2. *Llamamos* ***paralelogramo fundamental*** ρ_* *asociado a un retículo* $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2$ *al paralelogramo de vértices 0,* w_1, w_2 *y* $w_1 + w_2$*. Esto es*

$$\rho_* = \{tw_1 + sw_2 : 0 \leq t, s < 1\}.$$

A partir de este punto, a cualquier paralelogramo trasladado del fundamental lo llamameros ***paralelogramo período****, o por simplicidad, paralelogramo. Un paralelogramo* ρ *será* ***especial*** *si su frontera* $\partial\rho$ *no contiene ceros ni polos de* f*.*

Por otro lado, a no ser que se diga lo contrario, supondremos que $\Im\left(\frac{w_2}{w_1}\right) > 0$, pues esta condición nos asegura que el recorrido del contorno del paralelogramo se efectúa en sentido contrario a las agujas del reloj.

Debido al carácter periódico, el comportamiento de una función elíptica queda determinado por su comportamiento en cualquier paralelogramo ρ. No obstante, este hecho se enfatizará en el corolario 1.1.7. Además, fijémonos que a partir de la definición de ρ tenemos una partición del plano complejo en el siguiente sentido

$$\mathbb{C} = \cup_{n,m\in\mathbb{Z}}(nw_1 + mw_2 + \rho).$$

Definición 1.3. *Diremos que dos puntos del plano* z *y* z' *son* ***congruentes*** *respecto de* $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2$ *si* $z - z'$ *es cierto período, esto es, si existen* $n, m \in \mathbb{Z}$ *tales que*

$$z - z' = nw_1 + mw_2 \in \Lambda.$$

Es evidente que un paralelogramo no contiene ningún par de puntos congruentes entre sí, mientras que un punto no vértice que esté en un lado del paralelogramo es congruente con otro punto situado en el respectivo lado paralelo. Respecto a los vértices, representan una cuaterna de puntos congruentes. Además, para cada punto z' del plano complejo, siempre hay un único punto congruente a z' en el paralelogramo ρ. Por tanto, una función doblemente periódica suele considerarse como una función definida sobre el toro y recíprocamente, toda función definida sobre el toro puede ser considerada como una función elíptica sobre $\mathbb{C}$ (véase la figura 1.2).

Obsérvese que si f y g son dos funciones elípticas con igual retículo de períodos Λ, entonces también son elípticas y con el mismo retículo de períodos las funciones $\frac{1}{f}$, f', $f + g$, fg, f/g, esta última cuando g no es idénticamente nula.
Demostremos que f' satisface esta propiedad. En efecto, por ser f meromorfa,

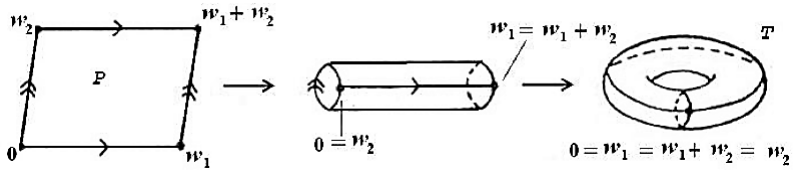

Figura 1.2: Toro bidimensional

f' es también lo es. Sea $z_1 \in \mathbb{C}$ donde f sea analítica y sea z_2 un punto congruente con z_1. Entonces,

$$\frac{f(z_1 + \Delta z) - f(z_1)}{\Delta z} = \frac{f(z_2 + \Delta z) - f(z_2)}{\Delta z}$$

y haciendo $\Delta z \to 0$ obtenemos $f'(z_1) = f'(z_2)$, por lo que f' tiene los mismos períodos que f. Además, f' no puede ser una constante pues en ese caso tendríamos $f(z) = cz + d$, que no es elíptica.

Definición 1.4. *El **orden** de una función elíptica es el número de soluciones de la ecuación $f(z) = \infty$, con $z \in \rho$. Es decir, el orden de f es la suma de los órdenes de todos los polos que pertenecen a un paralelogramo. Se denota por **ord(f)**.*

Desde el principio asumimos que las funciones elípticas son meromorfas en $\mathbb{C}$. Luego, por definición de función meromorfa, las funciones elípticas no constantes tendrán sólo una cantidad finita de ceros y polos en el paralelogramo fundamental. Veamos en esta línea algunos teoremas que manifiestan las propiedades principales de tales funciones.

Teorema 1.1.2. *Una función elíptica no constante no puede ser entera.*

Demostración. Por reducción al absurdo, supongamos que f es una función entera. Entonces, f es continua en el paralelogramo fundamental ρ_* y por tanto, está acotada. Es decir,

$$|f(z)| \leq C \quad \forall z \in \rho_*.$$

Si z' es un punto del plano complejo, entonces existe un punto congruente z del paralelogramo fundamental tal que f toma el mismo valor en ambos puntos. Luego, la desigualdad obtenida ha de cumplirse en todo el plano. Pero en ese caso, del teorema de Liouville llegamos a que f es constante, contradiciendo la hipótesis del teorema. □

Corolario 1.1.3. *Una función elíptica no constante tiene al menos un polo en el paralelogramo fundamental.*

Teorema 1.1.4. *Sea ρ un paralelogramo. Entonces la suma de los residuos en los polos de f contenidos en ρ es igual a cero.*

Demostración. En primer lugar, asumamos por simplicidad que no hay ceros ni polos de f en la frontera de ρ. Si los hubiera, podemos conseguir la condición anterior desplazando ρ al paralelogramo $a + \rho$, siendo $a \in \mathbb{C}$ lo suficientemente pequeño. Notemos que al ser el número de ceros y polos finito, es claro que podemos encontrar un paralelogramo especial. El resultado es consecuencia del teorema de los residuos aplicado a f. Consideraremos como camino de integración los lados del paralelogramo ρ. Por tal teorema, tenemos que

$$\sum_{\substack{a \text{ polo de } f \\ a \in \rho}} \operatorname{res}(f; a) = \frac{1}{2\pi i} \int_{\partial\rho} f(z)dz.$$

Denotemos $\Gamma_1, \Gamma_2, \Gamma_3$ y Γ_4 los lados del paralelogramo ρ de vértices $z_0, z_0 + w_1, z_0 + w_1 + w_2$ y $z_0 + w_2$, donde el sentido de integración a lo largo de cada Γ_i es consistente con la orientación positiva de $\partial\rho$. Entonces la anterior integral es

$$\int_{\partial\rho} f(z)dz = \sum_{i=1}^{4} \int_{\Gamma_i} f(z)dz$$

Demostremos que la suma de la integral en Γ_1 y en Γ_3, así como la suma de la integral en Γ_2 y en Γ_4 es igual a cero.
En efecto, la ecuación del lado Γ_1 es

$$z = z_0 + tw_1, \quad 0 \leq t \leq 1,$$

y entonces

$$\int_{\Gamma_1} f(z) = w_1 \int_0^1 f(z_0 + tw_1)dt.$$

La ecuación del lado que une los vértices $z_0 + w_2$ y $z_0 + w_1 + w_2$ es

$$z_0 + tw_1 + w_2, \quad 0 \leq t \leq 1,$$

por lo que nos queda

$$-\int_{\Gamma_3} f(z) = w_1 \int_0^1 f(z_0 + tw_1 + w_2)dt = w_1 \int_0^1 f(z_0 + tw_1)dt,$$

de donde obtenemos la relación

$$\int_{\Gamma_1} f(z)dz = -\int_{\Gamma_3} f(z)dz.$$

Del mismo modo se demuestra que

$$\int_{\Gamma_2} f(z)dz = -\int_{\Gamma_4} f(z)dz.$$

En conclusión, $\int_{\partial\rho} f(z)dz = 0$, obteniendo así el resultado. □

Corolario 1.1.5. *El orden de una función elíptica f no constante no es menor que 2.*

Demostración. En efecto, si el orden fuese 0, f no tendría polos y por tanto sería entera, contradiciendo el teorema 1.1.2.
Si el orden fuera 1 entonces existiría un polo simple β y del desarrollo de Laurent en un entorno de β, $f(z)$ sería de la forma $\dfrac{B}{z-\beta} + h(z)$, siendo B el residuo de f en β y $h(z)$ función holomorfa. Pero por el teorema anterior, $B = 0$, de donde se deduce que f no tiene polos en ρ y según el teorema 1.1.2, f sería constante. □

Teorema 1.1.6. *Una función elíptica f en un paralelogramo ρ tiene igual cantidad de polos que de ceros, contando ambos según sus multiplicidades.*

Demostración. Al igual que hicimos en el teorema 1.1.4, podemos suponer que la frontera de ρ no contiene ceros ni polos de f. Por ser f doblemente periódica con períodos en Λ, el cociente $\dfrac{f'}{f}$ es también doblemente periódico con igual retículo de períodos. En virtud del principio del argumento aplicado a $\partial\rho$ tenemos que

$$\frac{1}{2\pi i}\int_{\partial\rho} \frac{f'(z)}{f(z)}dz = \sum_{\substack{a \text{ cero de } f \\ a\in\rho}} m_a - \sum_{\substack{b \text{ polo de } f \\ b\in\rho}} m_b,$$

siendo m_a y m_b la multiplicidad de a o b como cero o polo de f respectivamente. Procediendo como en la prueba del teorema 1.1.4, la integral $\displaystyle\int_{\partial\rho} \frac{f'(z)}{f(z)}dz$ es nula por la simetría propia de $\partial\rho$. □

Corolario 1.1.7. *Si una función elíptica f tiene orden $N > 0$, entonces f toma cada valor complejo en un paralelogramo exactamente N veces.*

Demostración. Si c es un complejo cualquiera, la función $f(z) - c$ también es elíptica y de orden N. Por el teorema anterior, $f(z) - c$ tiene exactamente N ceros para todo $c \in \mathbb{C}$. □

Teorema 1.1.8. *La suma de los ceros y la suma de los polos en un paralelogramo contando su multiplicidad son números congruentes.*

Demostración. Por el principio del argumento se tiene que

$$\frac{1}{2\pi i}\int_{\partial\rho} z\frac{f'(z)}{f(z)}dz = \sum_{\substack{a \text{ cero de } f \\ a\in\rho}} m_a a - \sum_{\substack{b \text{ polo de } f \\ b\in\rho}} m_b b.$$

Por definición de congruencia, es suficiente demostrar que la integral es igual a algún período de la función f. Por la simetría de $\partial\rho$, tenemos que

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial\rho} z\frac{f'(z)}{f(z)}dz = &\frac{1}{2\pi i}\int_{\Gamma_1} z\frac{f'(z)}{f(z)}dz + \frac{1}{2\pi i}\int_{\Gamma_2} z\frac{f'(z)}{f(z)}dz \\ &- \frac{1}{2\pi i}\int_{\Gamma_3} z\frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{\Gamma_4} z\frac{f'(z)}{f(z)}dz,\end{aligned}$$

donde hemos representado los lados de $\partial\rho$ por $\Gamma_1, \Gamma_2, \Gamma_3$ y Γ_4, de modo que sus puntos iniciales y finales son: $z_0, z_0 + w_1$ para Γ_1, $z_0 + w_1, z_0 + w_1 + w_2$ para Γ_2, $z_0 + w_2, z_0 + w_1 + w_2$ para Γ_3 y $z_0, z_0 + w_2$ para Γ_4.

Veamos que la suma de la primera y tercera, así como la suma de la segunda y cuarta integral del segundo miembro de la igualdad son unos períodos de f.

En efecto, ya hemos visto que el lado Γ_1 tiene por ecuación $z = z_0 + w_1 t$, con $0 \le t \le 1$. Luego,

$$\frac{1}{2\pi i}\int_{\Gamma_1} z\frac{f'(z)}{f(z)}dz = \frac{1}{2\pi i}\int_0^1 (z_0 + w_1 t)\frac{w_1 f'(z_0 + w_1 t)}{f(z_0 + w_1 t)}dt.$$

La ecuación de Γ_3 se puede expresar como $z = z_0 + w_2 + w_1 t$, con $0 \le t \le 1$. Por tanto,

$$\begin{aligned}-\frac{1}{2\pi i}\int_{\Gamma_3} z\frac{f'(z)}{f(z)}dz &= -\frac{1}{2\pi i}\int_0^1 (z_0 + w_2 + w_1 t)\frac{w_1 f'(z_0 + w_2 + w_1 t)}{f(z_0 + w_2 + w_1 t)}dt \\ &= -\frac{1}{2\pi i}\int_0^1 (z_0 + w_2 + w_1 t)\frac{w_1 f'(z_0 + w_1 t)}{f(z_0 + w_1 t)}dt.\end{aligned}$$

Si denotamos por $\gamma(t)$ a la curva cerrada dada por $\gamma(t) = f(z_0 + w_1 t)$, con $0 \le t \le 1$, se tiene que

$$\frac{1}{2\pi i}\left(\int_{\Gamma_1} z\frac{f'(z)}{f(z)}dz - \int_{\Gamma_3} z\frac{f'(z)}{f(z)}dz\right)$$

$$\begin{aligned}
&= \frac{1}{2\pi i}\int_0^1 \frac{w_1 f'(z_0+w_1t)}{f(z_0+w_1t)}(z_0+w_1t-z_0-w_2-w_1t)dt\\
&= -\frac{w_2}{2\pi i}\int_0^1 \frac{w_1 f'(z_0+w_1t)}{f(z_0+w_1t)}dt\\
&= -\frac{w_2}{2\pi i}\int_0^1 \frac{\gamma'(t)}{\gamma(t)}dt\\
&= -kw_2, \text{ con } k\in\mathbb{Z},
\end{aligned}$$

pues la última integral es el índice de la curva $\gamma(t)$ respecto del origen, por lo que ha de ser un número entero.

Análogamente, hallamos que

$$\frac{1}{2\pi i}\int_{\Gamma_2} z\frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{\Gamma_4} z\frac{f'(z)}{f(z)}dz = lw_1, \text{ con } l\in\mathbb{Z}.$$

Así pues,

$$\frac{1}{2\pi i}\int_{\partial\rho} z\frac{f'(z)}{f(z)}dz = lw_1 - kw_2 = \Omega,$$

siendo Ω cierto período de la función f. □

1.2 Funciones elípticas de Jacobi

Aunque Weierstrass fue verdaderamente quien marcó la dirección del estudio de las funciones elípticas hasta la actualidad, no por ello debemos desestimar el avance anterior de Jacobi. En esta sección seguiremos los pasos dados en *Fundamenta nova functionum ellipticarum* (1829), obra en la que Jacobi representa las funciones elípticas como cociente de sus funciones theta.

De la *Integral Elíptica de Lengendre* con módulo k, $0 \le k \le 1$,

$$F(x,k) = \int_0^x \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

Jacobi define la función **sn(u,k)** dada por la inversa de la anterior integral con $x = \mathrm{sn}(u,k)$. Es decir,

$$u = \int_0^{\mathrm{sn}(u,k)} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}.$$

A partir de la definición de $\mathrm{sn}(u,k)$, Jacobi da nombre a las funciones **cn(u,k)** y **dn(u,k)** dadas por las soluciones positivas de las ecuaciones

$$\mathrm{sn}^2(u,k) + \mathrm{cn}^2(u,k) = 1$$
$$k^2\mathrm{sn}^2(u,k) + \mathrm{dn}^2(u,k) = 1.$$

La notación de estas funciones no es por mero azar, pues cuando $k = 0$, las funciones $\mathrm{sn}(u)$ y $\mathrm{cn}(u)$ son los análogos a las funciones $\sin(u)$ y $\cos(u)$ respectivamente, mientras que $\mathrm{dn}(u)$ degenera a la función constante 1.

El objetivo de este capítulo, así como el de Jacobi en su obra, es demostrar que estas tres funciones son en efecto elípticas. Para ello, los cálculos se reducen gracias a la ayuda de ciertas funciones auxiliares, conocidas como las funciones theta de Jacobi.

Siguiendo la notación clásica de Gauss, si τ es un número complejo fijado con parte imaginaria positiva, escribimos $q = e^{i\pi\tau}$, de forma que $|q| < 1$.

La **función theta** se define entonces como

$$\Theta(z,q) = \sum_{n=-\infty}^{+\infty} (-1)^n q^{n^2} e^{2niz}.$$

$\Theta(z,q)$ resulta ser una función entera ya que el rápido decaimiento de q^{n^2} nos asegura la convergencia uniforme de la serie en cada disco $D(0,R)$, con $R > 0$. Aunque Θ no sea doblemente periódica, esta función tiene un carácter similar al de una función elíptica.

Es sencillo ver que

$$\Theta(z+\pi,q) = \Theta(z,q)$$

y además,

$$\begin{aligned}\Theta(z+\pi\tau,q) &= \sum_{n=-\infty}^{+\infty} (-1)^n q^{n^2} e^{2niz} q^{2n} \\ &= -q^{-1}e^{-2iz} \sum_{n=-\infty}^{+\infty} (-1)^{n+1} q^{(n+1)^2} e^{2(n+1)iz} \\ &= -q^{-1}e^{-2iz}\Theta(z,q).\end{aligned} \tag{1.1}$$

Como consecuencia de estos resultados, decimos que Θ es una función *casi* doblemente periódica. En este sentido, 1 y $-q^{-1}e^{-2iz}$ se denominan **multiplicadores** o **factores de periodicidad** asociados a los períodos π y $\pi\tau$ respectivamente.

Las cuatro **funciones theta de Jacobi** se definen como

$$\begin{aligned}\Theta_1(z,q) &= -iM\Theta(z+\frac{1}{2}\pi\tau, q)\\ \Theta_2(z,q) &= \Theta_1(z+\frac{1}{2}\pi, q)\\ \Theta_3(z,q) &= \Theta(z+\frac{1}{2}\pi, q)\\ \Theta_4(z,q) &= \Theta(z,q).\end{aligned}$$

siendo $M = q^{\frac{1}{4}}e^{iz}$.

Por brevedad, el parámetro q generalmente no se especificará, quedando $\Theta_j(z) = \Theta_j(z,q)$.

De la definición de cada función theta, obtenemos las siguientes relaciones

$$\Theta_2(z) = M\Theta_3(z+\frac{1}{2}\pi\tau) \tag{1.2}$$

$$\Theta_3(z) = M\Theta_2(z+\frac{1}{2}\pi\tau) \tag{1.3}$$

$$\Theta_4(z) = -iM\Theta_1(z+\frac{1}{2}\pi\tau). \tag{1.4}$$

Demostremos la última relación.

Observemos que

$$\begin{aligned}\Theta_1(z) &= -iq^{\frac{1}{4}}e^{iz}\Theta(z+\frac{1}{2}\pi\tau)\\ &= -iq^{\frac{1}{4}}e^{iz}\sum_{n=-\infty}^{+\infty}(-1)^n q^{n^2}e^{2ni(z+\frac{1}{2}\pi\tau)}\\ &= -iq^{\frac{1}{4}}e^{iz}\sum_{n=-\infty}^{+\infty}(-1)^n q^{n^2+n}e^{2niz\pm ni\pi\tau}\\ &= -i\sum_{n=-\infty}^{+\infty}(-1)^n q^{(n+\frac{1}{2})^2}e^{i(2n+1)z},\end{aligned}$$

por lo que $\Theta_1(z)$ se puede escribir como

$$\Theta_1(z) = -i \sum_{n=-\infty}^{+\infty} (-1)^n q^{(n+\frac{1}{2})^2} e^{i(2n+1)z}. \tag{1.5}$$

Entonces,

$$\begin{aligned}
\Theta_1(z+\frac{1}{2}\pi\tau) &= -i \sum_{n=-\infty}^{+\infty} (-1)^n q^{(n+\frac{1}{2})^2} e^{i(2n+1)(z+\frac{1}{2}\pi\tau)} \\
&= i \sum_{n=-\infty}^{+\infty} (-1)^{n+1} q^{(n+\frac{1}{2})^2} q^{n+\frac{1}{2}} e^{2niz+iz\pm iz} \\
&= iq^{-\frac{1}{4}} e^{-iz} \sum_{n=-\infty}^{+\infty} (-1)^{n+1} q^{(n+\frac{1}{2})^2} q^{n+\frac{1}{2}} e^{2iz(n+1)} \\
&= iq^{-\frac{1}{4}} e^{-iz} \Theta_4(z),
\end{aligned}$$

obteniendo así el resultado.
Para las otras dos relaciones, basta tener en cuenta que podemos escribir $\Theta_2(z)$ como

$$\Theta_2(z) = \sum_{n=-\infty}^{+\infty} q^{(n+\frac{1}{2})^2} e^{i(2n+1)z}$$

y procedemos de manera parecida como hicimos con la relación anterior.

Como decíamos al principio de la sección, vamos a ver que $\mathrm{sn}(u)$ (y por tanto $\mathrm{cn}(u)$ y $\mathrm{dn}(u)$) se puede escribir como cociente de las cuatro funciones theta. La idea es que al hacer cociente, los factores extras que aparecen en los casi-períodos se cancelan, lo que da lugar a una función doblemente periódica. Recogemos una serie de resultados que nos serán de utilidad para nuestro objetivo.

Proposición 1.2.1. *Denotemos $N = q^{-1}e^{-2iz}$. Entonces se satisfacen las siguientes relaciones*

$$\begin{aligned}
&\Theta_1(z+\pi) = -\Theta_1(z) && \Theta_1(z+\pi\tau) = -N\Theta_1(z) \\
&\Theta_2(z+\pi) = -\Theta_2(z) && \Theta_2(z+\pi\tau) = N\Theta_2(z) \\
&\Theta_3(z+\pi) = \Theta_3(z) && \Theta_3(z+\pi\tau) = N\Theta_3(z) \\
&\Theta_4(z+\pi) = \Theta_4(z) && \Theta_4(z+\pi\tau) = -N\Theta_4(z).
\end{aligned}$$

Demostración. Las igualdades al incrementar z por π son directas ya que $e^{2ni\pi} = 1$ y $e^{ni\pi} = (-1)^n$ para todo $n \in \mathbb{N}$.

La igualdad $\Theta_4(z+\pi\tau) = -N\Theta_4(z)$ ya la tenemos demostrada en (1.1). Demostremos la igualdad para $\Theta_1(z+\pi\tau) = -N\Theta_1(z)$, el resto no la haremos para no alargar innecesariamente la sección.

Usando la igualdad (1.5), se tiene que

$$\begin{aligned}\Theta_1(z+\pi\tau) &= -i\sum_{n=-\infty}^{+\infty}(-1)^n q^{(n+\frac{1}{2})^2}e^{i(2n+1)z}\\ &= iq^{-1}e^{-2iz}\sum_{n=-\infty}^{+\infty}(-1)^{n+1}q^{(n+1+\frac{1}{2})^2}e^{i((2(n+1)+1)z}\\ &= -N\Theta_1(z).\end{aligned}$$

□

Proposición 1.2.2. *La función $\Theta_j(z)$ tiene exactamente un cero en cada paralelogramo ρ_t, siendo ρ_t el paralelogramo con vértices en los puntos $t, t+\pi, t+\pi+\pi\tau$ y $t+\pi\tau$ y $t \in \mathbb{C}$, con $j \in \{1,2,3,4\}$.*

Demostración. Haremos la prueba para Θ_4, pues para el resto se hace de forma análoga.
Al igual que hicimos en la sección anterior, podemos suponer sin pérdida de generalidad que la frontera de ρ_t no contiene ceros ni polos de $\Theta_4(z)$. Sabemos que $\Theta_4(z)$ tiene período π y casi período $\pi\tau$. Por ser $\Theta_4(z)$ entera, no tiene polos y en virtud del principio del argumento, el número de ceros contando multiplicidades en cada ρ_t es

$$\frac{1}{2\pi i}\int_{\partial\rho_t}\frac{\Theta_4'(z)}{\Theta_4(z)}dz.$$

Esta integral se puede escribir como

$$\frac{1}{2\pi i}\left(\int_t^{t+\pi}\frac{\Theta_4'(z)}{\Theta_4(z)}dz + \int_{t+\pi}^{t+\pi+\pi\tau}\frac{\Theta_4'(z)}{\Theta_4(z)}dz + \int_{t+\pi+\pi\tau}^{t+\pi\tau}\frac{\Theta_4'(z)}{\Theta_4(z)}dz + \int_{t+\pi\tau}^{t}\frac{\Theta_4'(z)}{\Theta_4(z)}dz\right).$$

Por ser $\Theta_4(z+\pi) = \Theta_4(z)$, entonces tenemos que

$$\int_{t+\pi}^{t+\pi+\pi\tau}\frac{\Theta_4'(z)}{\Theta_4(z)}dz = \int_t^{t+\pi\tau}\frac{\Theta_4'(z)}{\Theta_4(z)}dz = -\int_{t+\pi\tau}^{t}\frac{\Theta_4'(z)}{\Theta_4(z)}dz,$$

y por tanto la integral en ρ_t se reduce a

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial P_t}\frac{\Theta_4'(z)}{\Theta_4(z)}dz &= \frac{1}{2\pi i}\left(\int_t^{t+\pi}\frac{\Theta_4'(z)}{\Theta_4(z)}dz - \int_{t+\pi\tau}^{t+\pi+\pi\tau}\frac{\Theta_4'(z)}{\Theta_4(z)}dz\right)\\ &= \frac{1}{2\pi i}\left(\int_t^{t+\pi}\frac{\Theta_4'(z)}{\Theta_4(z)}dz - \int_t^{t+\pi}\frac{\Theta_4'(z+\pi\tau)}{\Theta_4(z+\pi\tau)}dz\right)\\ &= \frac{1}{2\pi i}\int_t^{t+\pi}\left(\frac{\Theta_4'(z)}{\Theta_4(z)} - \frac{\Theta_4'(z+\pi\tau)}{\Theta_4(z+\pi\tau)}\right)dz.\end{aligned}$$

Por otro lado, sabemos que $\Theta_4(z+\pi\tau) = -q^{-1}e^{-2iz}\Theta_4(z)$. Derivando a ambos lados de la igualdad nos queda

$$\Theta_4'(z+\pi\tau) = -q^{-1}e^{-2iz}\Theta_4'(z) + 2iq^{-1}e^{-2iz}\Theta_4(z).$$

Por tanto,

$$\frac{\Theta_4'(z+\pi\tau)}{\Theta_4(z+\pi\tau)} = \frac{\Theta_4'(z)}{\Theta_4(z)} - 2i.$$

Así podemos concluir que

$$\frac{1}{2\pi i}\int_{\partial\rho_t}\frac{\Theta_4'(z)}{\Theta_4(z)}dz = \frac{1}{2\pi i}\int_t^{t+\pi}2idz = 1,$$

por lo que $\Theta_4(z)$ tiene exactamente un cero simple en cada ρ_t. □

Proposición 1.2.3. *Las siguientes relaciones manifiestan que es posible expresar cualquier función theta en términos de cualquier otro par de funciones theta:*

1. $\Theta_2(z)\Theta_4^2(0) = \Theta_4^2(z)\Theta_2^2(0) - \Theta_1^2(z)\Theta_3^2(0)$
2. $\Theta_3^2(z)\Theta_4^2(0) = \Theta_4^2(z)\Theta_3^2(0) - \Theta_1^2(z)\Theta_2^2(0)$.
3. $\Theta_4^2(z)\Theta_4^2(0) = \Theta_3^2(z)\Theta_3^2(0) - \Theta_2^2(z)\Theta_2^2(0)$.

Demostración. Demostremos las dos primeras igualdades. La tercera veremos que es consecuencia de la segunda. Cada una de las cuatro funciones $\Theta_1^2(z), \Theta_2^2(z), \Theta_3^2(z)$ y $\Theta_4^2(z)$ es analítica para todos los valores de z y tiene factores de periodicidad 1 y $q^{-2}e^{-4iz}$ asociado a los períodos π y $\pi\tau$ respectivamente. Además, cada una tiene un cero doble (y no más) en cualquier paralelogramo ρ_t definido como en la proposición anterior. Por tanto, es obvio que si a, b, a' y b' son elegidos de forma adecuada, las funciones

$$\frac{a\Theta_1^2(z) + b\Theta_4^2(z)}{\Theta_2^2(z)} \quad \text{y} \quad \frac{a'\Theta_1^2(z) + b'\Theta_4^2(z)}{\Theta_3^2(z)}$$

son elípticas (de período π y $\pi\tau$) y tienen a lo sumo sólo un polo simple en cada ρ_t. Luego han de ser funciones constantes, dando lugar a relaciones de la forma

$$\Theta_2^2(z) = A\Theta_1^2(z) + B\Theta_4^2(z), \quad \Theta_3^2(z) = A'\Theta_1^2(z) + B'\Theta_4^2(z),$$

para ciertas constantes A, B, A' Y B'.

De la definición y de la casi periodicidad de cada una de las funciones theta obtenemos las relaciones

$$\Theta_2\left(\frac{1}{2}\pi\tau\right) = q^{\frac{-1}{4}}\Theta_3, \quad \Theta_4\left(\frac{1}{2}\pi\tau\right) = 0, \quad \Theta_1\left(\frac{1}{2}\pi\tau\right) = iq^{\frac{-1}{4}}\Theta_4$$

y dando a z los valores $\frac{1}{2}\pi\tau$ y 0 en las relaciones anteriores, las constantes A, B, A' y B' satisfacen

$$\Theta_3^2(0) = -A\Theta_4^2(0), \quad \Theta_2^2(0) = B\Theta_4^2(0); \quad \Theta_2^2(0) = -A'\Theta_4^2(0), \quad \Theta_3^2(0) = B'\Theta_4^2(0).$$

Despejando las constantes A, B, A' y B' de lo anterior y sustituyendo en las relaciones que teníamos, obtenemos las ecuaciones 1) y 2).

Probemos la tercera igualdad. La segunda igualdad evaluada en $z + \frac{1}{2}\pi$ es

$$\Theta_3^2(z + \frac{1}{2}\pi)\Theta_4^2(0) = \Theta_4^2(z + \frac{1}{2}\pi)\Theta_3^2(0) - \Theta_1^2(z + \frac{1}{2}\pi)\Theta_2^2(0).$$

Ahora bien, por ser $\Theta_4(z)$ periódica de período π, tenemos que

$$\Theta_3^2(z + \frac{1}{2}\pi) \coloneqq \Theta_4^2(z + \pi) = \Theta_4^2(z)$$

y por definición de $\Theta_2(z)$ y $\Theta_3(z)$,

$$\Theta_4^2(z + \frac{1}{2}\pi) = \Theta_3^2(z), \quad \Theta_1^2(z + \frac{1}{2}\pi) = \Theta_2^2(z),$$

obteniendo la igualdad.

□

Veamos una expresión de $\Theta_4(z)$ como producto infinito que nos será de ayuda más adelante.

Proposición 1.2.4. *$\Theta_4(z)$ puede ser expresada como*

$$\Theta_4(z) = G\prod_{n=1}^{\infty}(1-2q^{2n-1}cos(2z)+q^{4n-2})$$

donde G es cierta constante que depende únicamente de q[1].

Demostración. Sea $f(z) \coloneqq \prod_{n=1}^{\infty}(1-q^{2n-1}e^{2iz})\prod_{n=1}^{\infty}(1-q^{2n-1}e^{-2iz})$. Debido a la convergencia absoluta de $\sum_{n=1}^{\infty} q^{2n-1}$, cada uno de los dos productos converge absoluta y uniformemente en cualquier región[2] acotada. Luego, $f(z)$ es analítica en $\mathbb{C}$ y por tanto es una función integrable.

Los ceros de $f(z)$ son simples en puntos tales que $e^{2iz} = e^{(2n+1)\pi i\tau}$, con $n \in \mathbb{Z}$, esto es, tales que $2iz = (2n+1)\pi i\tau + 2m\pi i$. Luego, $f(z)$ y $\Theta_4(z)$ tienen los mismos ceros y por tanto, el cociente $\dfrac{\Theta_4(z)}{f(z)}$ no tiene ceros ni polos en $\mathbb{C}$.

Por otro lado, es evidente que $f(z+\pi) = f(z)$ y de la siguiente igualdad

$$\begin{aligned} f(z+\pi\tau) &= \prod_{n=1}^{\infty}(1-q^{2n+1}e^{2iz})\prod_{n=1}^{\infty}(1-q^{2n-3}e^{-2iz}) \\ &= f(z)\frac{1-q^{-1}e^{-2iz}}{1-qe^{2iz}} \\ &= -q^{-1}e^{-2iz}f(z), \end{aligned}$$

vemos que $f(z)$ y $\Theta_4(z)$ tiene los mismos factores de periodicidad asociado a π y a $\pi\tau$. Luego, la función $\dfrac{\Theta_4(z)}{f(z)}$ es elíptica sin ceros ni polos, por lo que ha de ser constante y dicha constante ha de ser G, obteniendo el resultado. □

Usando de nuevo la expresión (1.5) y la conocida igualdad $\sin z = \dfrac{e^{iz}-e^{-iz}}{2i}$ $\forall z \in \mathbb{C}$, podemos expresar $\Theta_1(z)$ como

$$\Theta_1(z) = 2\sum_{n=0}^{+\infty}(-1)^n q^{(n+\frac{1}{2})^2}\sin(2n+1)z.$$

Luego, $z=0$ es un cero de $\Theta_1(z)$ y en virtud de la proposición 1.2.2, ya tenemos estudiados todos los ceros de $\Theta_1(z)$. De las definiciones del resto de

[1] La expresión de la constante G puede consultarse en *§21.42* de [WW28].
[2] Entendemos por región a cualquier conjunto de $\mathbb{C}$ abierto y conexo.

funciones theta, se sigue que los ceros de $\Theta_2(z), \Theta_3(z)$ y $\Theta_4(z)$ son los puntos congruentes a $-\frac{1}{2}\pi, -\frac{1}{2}\pi + \frac{1}{2}\pi\tau$ y $\frac{1}{2}\pi\tau$ respectivamente. Por tanto, los ceros en el paralelogramo ρ_t de $\Theta_1(z), \Theta_2(z), \Theta_3(z)$ y $\Theta_4(z)$ son congruentes a los puntos $0, \frac{1}{2}\pi, \frac{1}{2}\pi + \frac{1}{2}\pi\tau$ y $\frac{1}{2}\pi\tau$ respectivamente.

Veamos ahora la ecuación que satisface $\mathrm{sn}(z)$.

De la proposición 1.2.1, es obvio que la función $f(z) = \dfrac{\Theta_1(z)}{\Theta_4(z)}$ tiene factores de periodicidad -1 y 1 asociados a los períodos π y $\pi\tau$ respectivamente. Por tanto, su derivada $f'(z) = \dfrac{\Theta_1'(z)\Theta_4(z) - \Theta_1(z)\Theta_4'(z)}{\Theta_4^2(z)}$ tiene los mismos factores de periodicidad.

Por otro lado, es fácil verificar que $g(z) = \dfrac{\Theta_2(z)\Theta_3(z)}{\Theta_4^2(z)}$ tiene -1 y 1 como factores de periodicidad. Por tanto, la función

$$\phi(z) = \frac{f'(z)}{g(z)} = \frac{\Theta_1'(z)\Theta_4(z) - \Theta_1(z)\Theta_4'(z)}{\Theta_2(z)\Theta_3(z)}$$

es elíptica con períodos π y $\pi\tau$ y los únicos posibles polos de $\phi(z)$ son simples en puntos congruentes a $\frac{1}{2}\pi$ y $\frac{1}{2}\pi + \frac{1}{2}\pi\tau$. Ahora bien, de las relaciones (1.2), (1.3) y (1.4) vemos que

$$\phi(z + \frac{1}{2}\pi\tau) = \frac{-\Theta_4'(z)\Theta_1(z) + \Theta_1'(z)\Theta_4(z)}{\Theta_3(z)\Theta_2(z)} = \phi(z),$$

por lo que $\phi(z)$ es elíptica con períodos π y $\frac{1}{2}\pi\tau$. Ahora bien, los únicos posibles polos de ϕ son simples en puntos congruentes a $\frac{1}{2}\pi$ y $\frac{1}{2}\pi + \frac{1}{2}\pi\tau$, pero por ser π y $\frac{1}{2}\pi\tau$ los períodos, tenemos que los puntos anteriores son congruentes y por tanto tenemos un único polo. La figura 1.3 resume gráficamente esta idea.

Por el teorema 1.1.2, ϕ es una constante y haciendo $z \to 0$, el valor de esta constante es $\dfrac{\Theta_1'(0)\Theta_4(0)}{\Theta_2(0)\Theta_3(0)} = \Theta_4^2(0)$. Notemos que hemos usado la igualdad $\Theta_1'(0) = \Theta_2(0)\Theta_3(0)\Theta_4(0)$, igualdad probada en *§21.41* de [WW28]. No hemos expuesto la prueba para no cargar con demasiadas cuentas la sección.

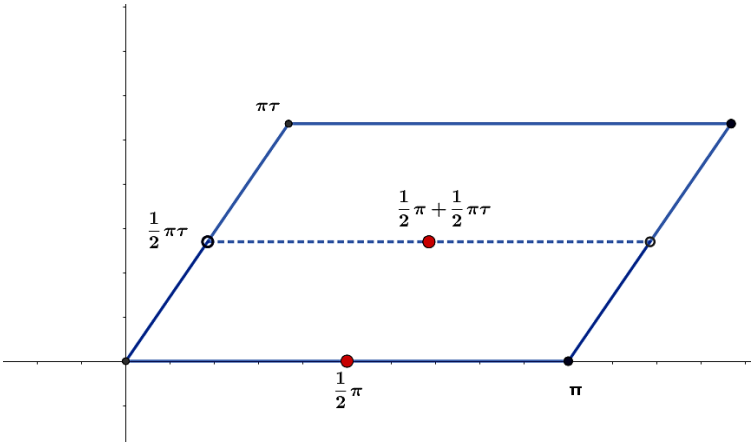

Figura 1.3

Por tanto, podemos escribir $f'(z)$ como

$$f'(z) = \phi(z)g(z) = \Theta_4^2(0)\frac{\Theta_2(z)\Theta_3(z)}{\Theta_4^2(z)} = \frac{\partial}{\partial z}\left(\frac{\Theta_1(z)}{\Theta_4(z)}\right).$$

Con el cambio $\xi(z) = \dfrac{\Theta_1(z)}{\Theta_4(z)}$, lo anterior es

$$\frac{d}{dz}(\xi(z)) = \Theta_4^2(0)\frac{\Theta_2(z)\Theta_3(z)}{\Theta_4^2(z)}$$

y usando las relaciones de la proposición 1.2.3, obtenemos la ecuación diferencial

$$\left(\frac{d\xi}{dz}\right)^2 = \left(\Theta_2^2(0) - \xi^2\Theta_3^2(0)\right)\left(\Theta_3^2(0) - \xi^2\Theta_2^2(0)\right),$$

cuya solución es $\xi(z) = \dfrac{\Theta_1(z)}{\Theta_4(z)}$.

Esta ecuación diferencial puede ser llevada a la forma canónica con el cambio $y = \xi\dfrac{\Theta_3(0)}{\Theta_2(0)}$ y $u = z\Theta_3^2(0)$. Si probamos que existe τ (con $\Im(\tau) > 0$) tal que

$$k^{\frac{1}{2}} = \frac{\Theta_2(0)}{\Theta_3(0)}, \tag{1.6}$$

la ecuación que determina y en función de u es

$$\left(\frac{dy}{du}\right)^2 = (1-y^2)(1-k^2y^2)$$

y esta ecuación tiene la solución particular

$$y = \frac{\Theta_3(0)}{\Theta_2(0)} \frac{\Theta_1(u\Theta_3^{-2}(0))}{\Theta_4(u\Theta_3^{-2}(0))}.$$

Ahora bien, por definición de $\mathrm{sn}(u,k)$, ha de ser $y = \mathrm{sn}(u,k)$, con $0 \leq k \leq 1$ anteriormente fijado. Es decir, $\mathrm{sn}(u)$ satisface

$$\mathrm{sn}(u) = \frac{\Theta_3(0)}{\Theta_2(0)} \frac{\Theta_1(u\Theta_3^{-2}(0))}{\Theta_4(u\Theta_3^{-2}(0))}$$

y la función de la derecha tiene multiplicadores -1 y 1 asociado a los períodos $\pi\Theta_3^2(0)$ y $\pi\tau\Theta_3^2(0)$. Por tanto, **sn(u,k) es elíptica** con períodos $2\pi\Theta_3^2(0)$ y $\pi\tau\Theta_3^2(0)$. Si somos fieles a la notación original de Jacobi, el casi período $\pi\Theta_3^2(0)$ y el período $\pi\tau\Theta_3^2(0)$ es denotado por $2K$ y $2iK$, de forma que $\mathrm{sn}(u,k)$ tiene períodos $4K$ y $2iK$.

Por tanto, todo se reduce a probar que existe τ con $\Im(\tau) > 0$ tal que se cumpla (1.6). De forma equivalente, si escribimos $k^2 = c$ con $0 < c < 1$, tenemos que ver que la ecuación

$$c = \frac{\Theta_2^4(0)}{\Theta_3^4(0)}$$

tiene solución. En efecto, haciendo $z = 0$ en la igualdad 3) de la proposición 1.2.3, la anterior ecuación equivale a

$$1 - c = \frac{\Theta_4^4(0)}{\Theta_3^4(0)}.$$

En virtud de la proposición 1.2.4, tenemos que

$$\Theta_4^4(0) = G^4 \prod_{n=1}^{\infty} (1 - q^{2n-1})^8 \quad \text{y} \quad \Theta_3^4(0) = G^4 \prod_{n=1}^{\infty} (1 + q^{2n-1})^8,$$

por lo que la ecuación anterior puede ser escrita como

$$1 - c = \prod_{n=1}^{\infty} \left(\frac{1 - q^{2n-1}}{1 + q^{2n-1}} \right)^8.$$

Como q varía de 0 a 1, el producto de la derecha es continuo y decrece estrictamente de 1 a 0, por lo que toma el valor $1 - c$ una sóla vez. Por tanto, el problema de la inversión ha sido resuelto y la función $\mathrm{sn}(u,k)$ puede ser

escrita como cociente de funciones theta, por lo que es elíptica.

Finalmente, otras nueve funciones elípticas se pueden definir tomando inversas y cocientes. Jacobi las definió como

$$\begin{aligned}
\operatorname{ns}(u) &= \frac{1}{\operatorname{sn}(u)} & \operatorname{nc}(u) &= \frac{1}{\operatorname{cn}(u)} & \operatorname{nd}(u) &= \frac{1}{\operatorname{dn}(u)} \\
\operatorname{sc}(u) &= \frac{\operatorname{sn}(u)}{\operatorname{cn}(u)} & \operatorname{cd}(u) &= \frac{\operatorname{cn}(u)}{\operatorname{dn}(u)} & \operatorname{ds}(u) &= \frac{\operatorname{dn}(u)}{\operatorname{sn}(u)} \\
\operatorname{cs}(u) &= \frac{\operatorname{cn}(u)}{\operatorname{sn}(u)} & \operatorname{dc}(u) &= \frac{\operatorname{dn}(u)}{\operatorname{cn}(u)} & \operatorname{sd}(u) &= \frac{\operatorname{sn}(u)}{\operatorname{dn}(u)}.
\end{aligned}$$

1.3 La función $\wp$ de Weierstrass

En 1899, Henri Poincaré dijo:

> La forma que Jacobi había dado a la teoría de las funciones elípticas está lejos de la perfección; encontramos sólo tres funciones fundamentales, muy específicas: sn, cn y dn... En el sistema de Weierstrass, en lugar de tres funciones, sólo hay una, $\wp(z)$, de forma que su definición no cambia cuando se reemplaza un sistema de períodos por otro sistema equivalente.

La importancia del estudio de Weierstrass radica en el hecho de que todas las funciones elípticas se pueden expresar en términos de $\wp(z)$ y su derivada $\wp'(z)$. Por ello, el objetivo de esta sección será demostrar tal resultado, señalando además algunas consecuencias que nos serán de utilidad para futuros capítulos.

Describamos una normalización de períodos que usaremos constantemente en esta sección. Esta normalización tendrá más sentido cuando en el tercer capítulo introduzcamos el concepto de grupo modular. Sea $\tau = \dfrac{w_2}{w_1}$ con w_1 y w_2 los períodos linealmente independientes sobre $\mathbb{R}$ de una función elíptica, digamos f. Ya que τ no es real (Lema 1.1.1) y τ y $\dfrac{1}{\tau}$ tienen partes imaginarias de distinto signo, asumamos sin pérdida de generalidad que $\Im(\tau) > 0$. Obsérvese que la función f tiene períodos w_1 y w_2 si y sólo si $F(z) \coloneqq f(w_1 z)$ tiene períodos 1 y τ y además f es elíptica si y sólo si F lo es. Como las propiedades de f se deducen de las de F, asumiremos por comodidad que f es meromorfa en $\mathbb{C}$ con períodos 1 y τ, siendo $\Im(\tau) > 0$. De estas condiciones se tiene

$$f(z + n + m\tau) = f(z), \quad \text{para todo } n, m \in \mathbb{Z} \text{ y } z \in \mathbb{C},$$

por lo que resulta natural considerar el retículo en $\mathbb{C}$ definido por

$$\Lambda = \{n + m\tau : n, m \in \mathbb{Z}\}$$

es decir, el retículo generado por 1 y τ.

El estudio que realizó Weierstrass sobre las funciones elípticas está enfocado desde el estudio de sus polos. Como ya hemos visto, cualquier función elíptica debe tener al menos dos polos. Construiremos por tanto una función cuya única singularidad sea un polo doble en los puntos del retículo Λ.

Antes de estudiar el caso de las funciones doblemente periódicas, consideremos funciones con un sólo período. Si uno desea construir una función con período 1 y polos todos enteros, una sencilla elección es

$$F(z) = \sum_{n=-\infty}^{\infty} \frac{1}{z+n}.$$

No obstante, la serie que define F no es absolutamente convergente. Para solucionar este problema podemos expresar $F(z)$ como

$$F(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z+n} + \frac{1}{z-n}\right) = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z+n} - \frac{1}{n}\right) + \left(\frac{1}{z-n} - \frac{1}{-n}\right),$$

es decir,

$$F(z) = \frac{1}{z} + \sum_{n\neq 0} \left(\frac{1}{z+n} - \frac{1}{n}\right).$$

Como la nueva expresión cumple $\dfrac{1}{z+n} - \dfrac{1}{n} = \mathcal{O}\left(\dfrac{1}{n^2}\right)$[3], la serie es absolutamente convergente en $\mathbb{C} \setminus \mathbb{Z}$ y uniformemente convergente para cada disco $D(0, R)$ con $R > 0$. Luego, F es meromorfa con polos precisamente en los enteros.

Veamos ahora cómo imitar el desarrollo anterior para construir una función elíptica. Para ello, utilizamos una suma análoga sobre el conjunto Λ anteriormente definido. En concreto, estudiamos la suma

$$\sum_{w\in\Lambda} \frac{1}{(z+w)^2} = \frac{1}{z^2} + \sum_{w\in\Lambda^*} \left(\frac{1}{(z+w)^2} - \frac{1}{w^2}\right),$$

[3]Decimos que $f(n) = \mathcal{O}(g(n))$ si se cumple que $|f(n)| \leq A|g(n)|$ para cierta constante $A > 0$.

donde $\Lambda^* = \Lambda - \{(0,0)\}$.

Igual que antes, por ser $\dfrac{1}{(z+w)^2} - \dfrac{1}{w^2} = \dfrac{-z^2-2zw}{(z+w)^2w^2}$, el término del sumatorio es $\mathcal{O}\left(\dfrac{1}{w^3}\right)$ cuando $|w| \to \infty$ y en virtud del siguiente lema, la serie es absolutamente para todo $z \in \mathbb{C} \setminus \Lambda$. Es más, la serie es uniformemente convergente en cada disco $D(0,R)$, con $R > 0$ ya que $\left|\dfrac{1}{(w-z)^2} - \dfrac{1}{w^2}\right| \leq C\dfrac{|z|}{|w|^3}$, definiendo así una función meromorfa con polos en Λ de orden 2.

Lema 1.3.1. *Las dos series*

$$\sum_{(n,m)\neq(0,0)} \frac{1}{(|n|+|m|)^r} \quad y \quad \sum_{n+m\tau\in\Lambda^*} \frac{1}{|n+m\tau|^r}$$

convergen si $r > 2$.

Demostración. Probemos el resultado para la primera serie.
Para cada $n \neq 0$,

$$\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r} = \frac{1}{|n|^r} + 2\sum_{m\geq 1} \frac{1}{(|n|+|m|)^r} \leq \frac{1}{|n|^r} + 2\sum_{k\geq|n|+1} \frac{1}{k^r}$$

$$\leq \frac{1}{|n|^r} + 2\int_{|n|}^{\infty} \frac{dx}{x^r} \leq \frac{1}{|n|^r} + C\frac{1}{|n|^{r-1}}.$$

Luego, si $r > 2$,

$$\sum_{(n,m)\neq(0,0)} \frac{1}{(|n|+|m|)^r} = \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0}\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r}$$

$$\leq \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0} \left(\frac{1}{|n|^r} + C\frac{1}{|n|^{r-1}}\right) < \infty.$$

Por tanto, para probar que la segunda serie converge, es suficiente ver que

$$|n| + |m| \lesssim |n + m\tau|, \quad \forall n, m \in \mathbb{Z},$$

donde $x \lesssim y$ si existe una constante positiva a tal que $x \leq ay$. Si se cumple también $y \lesssim x$, escribiremos $x \approx y$.
Para ello veamos antes que para cualesquiera dos números positivos A y B se satisface la siguiente relación

$$(A^2 + B^2)^{\frac{1}{2}} \approx A + B.$$

Por un lado,

$$A \leq (A^2 + B^2)^{\frac{1}{2}}$$
$$B \leq (A^2 + B^2)^{\frac{1}{2}}$$

y sumando ambas desigualdades llegamos a que

$$A + B \leq 2(A^2 + B^2)^{\frac{1}{2}}.$$

Por otro lado, al ser A y B positivos,

$$(A + B)^2 = A^2 + B^2 + 2AB \geq A^2 + B^2$$

y tomando raíz cuadrada obtenemos que

$$(A^2 + B^2)^{\frac{1}{2}} \leq A + B,$$

por lo que se tiene la relación anterior.

Probemos ahora la siguiente relación:

$$|n| + |m| \approx |n + m\tau|, \quad \forall \tau \in \mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\}.$$

En efecto, si $\tau = s + it$, con $t > 0$, por la observación previa,

$$|n + m\tau| = |(n + ms)^2 + (mt)^2|^{\frac{1}{2}} \approx |n + ms| + |mt| \approx |n + ms| + |m|.$$

En conclusión, $|n| + |m| \approx |n + m\tau|$, por lo que la segunda serie converge si $r > 2$.

□

Con esta idea de Weierstrass, a simple vista sencilla pero no por ello menos brillante, podemos ya sí definir la función $\wp(z)$ dada por la serie

$$\begin{aligned}\wp(z) &= \frac{1}{z^2} + \sum_{w \in \Lambda^*} \left(\frac{1}{(z+w)^2} - \frac{1}{w^2} \right) \\ &= \frac{1}{z^2} + \sum_{(n,m) \neq (0,0)} \left(\frac{1}{(z + n + m\tau)^2} - \frac{1}{(n + m\tau)^2} \right)\end{aligned}$$

y en virtud del lema anterior, tenemos que $\wp(z)$ es una función meromorfa. Además, por la propia definición, $\wp(z)$ tiene polos dobles en los puntos del retículo Λ. Notar que, debido a la inserción de los términos $-\dfrac{1}{w^2}$, ya no es tan obvio que $\wp(z)$ sea doblemente periódica. No obstante, vamos a demostrar que sí lo es. Para ello, observemos que $\wp$ es claramente par y por tanto, la derivada $\wp'$, impar.

Teorema 1.3.2. *$\wp(z)$ es una función elíptica con períodos 1 y τ.*

Demostración. Veamos primero que la derivada de $\wp$ es doblemente periódica. Derivando término a término, la expresión de $\wp'$ es

$$\wp'(z) = -2 \sum_{n,m\in\mathbb{Z}} \frac{1}{(z+n+m\tau)^3}.$$

Por el lema 1.3.1, $\wp'$ converge absolutamente cuando z no es un punto del retículo. Además, la diferenciación elimina el término $-\frac{1}{w^2}$ y por tanto la serie para $\wp'$ es claramente periódica con períodos 1 y τ, pues permanece sin cambios después de reemplazar z por $z+1$ o por $z+\tau$. Luego existen dos constantes a y b tales que

$$\wp(z+1) = \wp(z) + a \quad \text{y} \quad \wp(z+\tau) = \wp(z) + b, \quad \forall z \in \mathbb{C}.$$

Por ser $\wp$ par, $\wp\left(-\frac{1}{2}\right) = \wp\left(\frac{1}{2}\right)$ y $\wp\left(-\frac{\tau}{2}\right) = \wp\left(\frac{\tau}{2}\right)$ y tomando $z = -\frac{1}{2}$ y $z = -\frac{\tau}{2}$ en las expresiones anteriores, llegamos a que $a = b = 0$. Por lo tanto, $\wp$ es elíptica con períodos 1 y τ.

□

Demostremos ya sí que todas las funciones elípticas se pueden expresar en función de $\wp(z)$ y $\wp'(z)$.

Teorema 1.3.3. *Toda función elíptica f con períodos 1 y τ es una función racional de $\wp$ y $\wp'$.*

Demostración. Haremos la prueba en dos pasos. Primero demostraremos el resultado para una función F elíptica par con los mismos períodos que f y luego demostraremos el caso general.

En efecto, por ser F par, si tiene un cero o polo en el origen será de orden par. Luego existe un entero de orden m de modo que la función $F(z)(\wp(z))^m$ no tiene ceros o polos en los puntos del retículo. Por ser $\wp$ también par con períodos 1 y τ, podemos asumir que F no tiene polos o ceros en Λ.

Sea $\{a_1, a_{-1}, \ldots, a_m, a_{-m}\}$ el conjunto de los ceros de F contado con multiplicidades y $\{b_1, b_{-1}, \ldots, b_m, b_{-m}\}$ el conjunto de sus polos contado con multiplicidades. Vamos a usar $\wp$ para construir una función elíptica G con los mismos ceros y polos que F. Para ello, recordemos que la función $\wp(z) - \wp(a)$

tiene un cero de orden 2 si a es medio período y dos ceros distintos, a y $-a$, en otro caso. Entonces, la función

$$(\wp(z) - \wp(a_1)) \cdot \ \ldots \cdot (\wp(z) - \wp(a_m))$$

tiene exactamente los mismos ceros que F.

Para los polos, razonando al igual que para los ceros, tenemos que la función

$$\frac{1}{(\wp(z) - \wp(b_1)) \cdot \ \ldots \cdot (\wp(z) - \wp(b_m))}$$

tiene los mismos polos que F y entonces, la función G dada por

$$G(z) \coloneqq \frac{(\wp(z) - \wp(a_1)) \cdot \ \ldots \cdot (\wp(z) - \wp(a_m))}{(\wp(z) - \wp(b_1)) \cdot \ \ldots \cdot (\wp(z) - \wp(b_m))}$$

tiene los mismos ceros y polos que F. Además, por ser $\wp$ elíptica, G también lo es. Así pues, $\dfrac{F}{G}$ es entera y doblemente periódica. Por el teorema de Liouville, $\dfrac{F}{G}$ ha de ser constante, finalizando en virtud del resultado para el caso par.

Demostremos el caso general.

Podemos escribir la función f como

$$f(z) = f_{par}(z) + f_{impar}(z),$$

siendo

$$f_{par}(z) = \frac{f(z) + f(-z)}{2} \quad \text{y} \quad f_{impar}(z) = \frac{f(z) - f(-z)}{2}.$$

Recordando que $\wp$ es par y su derivada $\wp'$ impar, el cociente $\dfrac{f_{impar}}{\wp'}$ es par. Aplicando el caso anterior a las funciones f_{par} y $\dfrac{f_{impar}}{\wp'}$, concluimos que f es una función racional de $\wp$ y $\wp'$.

□

Recordemos la normalización que hicimos al principio de la sección. Ya que la construcción de $\wp$ depende de τ, podríamos escribir $\wp_\tau$. Esto nos lleva a cambiar nuestro punto de vista y pensar en $\wp_\tau(z)$ principalmente como una función de τ. Este enfoque produce nuevas ideas interesantes que nos lleva a describir el carácter modular de las funciones elípticas y su conexión con las series de Eisenstein. Además, esta consideración está motivada por las siguientes observaciones.

- Ya que 1 y τ generan los períodos de $\wp_\tau(z)$ y 1 y $\tau+1$ generan los mismos períodos, cabe esperar una relación estrecha entre $\wp_\tau(z)$ y $\wp_{\tau+1}(z)$. De hecho, es fácil ver que son idénticas.
- Si $\tau = \dfrac{w_2}{w_1}$ con $\Im(\tau) > 0$ entonces intercambiando los dos períodos tenemos $-\dfrac{1}{\tau} = -\dfrac{w_1}{w_2}$ y además $\Im\left(-\dfrac{1}{\tau}\right) > 0$. De nuevo, podemos esperar una conexión entre $\wp_\tau$ y $\wp_{-\frac{1}{\tau}}$. En particular, $\wp_{\frac{-1}{\tau}}(z) = \tau^2\wp_\tau(\tau z)$.

Por tanto, nos vemos impulsados a considerar el grupo de transformaciones del semiplano superior $\Im(\tau) > 0$, generado por las dos transformaciones

$$\tau \to \tau + 1$$
$$\tau \to -\frac{1}{\tau}.$$

Tal grupo recibe el nombre de **grupo modular**, que estudiaremos con más detalle en el tercer capítulo. Para estudiar las relaciones anteriores con más claridad, recurrimos a las series de Eisenstein.

Definición 1.5. *Las **series de Eisenstein de orden k** se definen como*

$$E_k(\tau) = \sum_{\substack{n,m\in\mathbb{Z}\\(n,m)\neq(0,0)}} \frac{1}{(n+m\tau)^k},$$

donde $k \geq 3$ es un entero y τ es un número complejo con $\Im(\tau) > 0$.

Teorema 1.3.4. *Las series de Eisenstein satisfacen las siguientes propiedades*

1. *$E_k(\tau)$ converge si $k \geq 3$ y es holomorfa en el semiplano superior.*
2. *$E_k(\tau) = 0$ si k es impar.*
3. *$E_k(\tau)$ satisface las siguientes relaciones de transformación*

$$E_k(\tau+1) = E_k(\tau) \quad y \quad E_k(\tau) = \tau^{-k}E_k\left(-\frac{1}{\tau}\right).$$ [4]

[4] En esta propiedad apreciamos el carácter modular de las series de Eisenstein.

Demostración. En virtud del lema 1.3.1, la serie $E_k(\tau)$ converge absolutamente en el semiplano superior $\Im(z) > 0$ siempre que $k \geq 3$. Además, si $k \geq 3$, la serie converge uniformemente en el semiplano $\Im(z) \geq \tau$ para todo $\tau > 0$ y por la arbitrariedad de τ, podemos concluir que $E_k(\tau)$ es holomorfa en todo el semiplano superior.

La segunda propiedad es consecuencia de la definición de $E_k(\tau)$, pues por la simetría, al sustituir n y m por $-n$ y $-m$ cuando k es impar, los términos opuestos se suprimen dos a dos y por tanto la serie de Eisenstein suma cero.

Por último, veamos las dos igualdades dadas.

El hecho de que $E_k(\tau)$ sea periódica de período 1 es evidente ya que $n + m(\tau + 1) = n + m\tau + m$ y de la primera propiedad, podemos reordenar los términos de la suma reemplanzando $n + m$ por n.

La segunda igualdad se deduce por ser $\left(n + m(-\frac{1}{\tau})\right)^k = \tau^{-k}(n\tau - m)^k$ y tras reemplazar esta vez $(-m, n)$ por (n, m). □

La conexión de las series de Eisenstein con la función $\wp$ surge cuando estudiamos la expansión de la serie que define $\wp$ cerca de 0.

Teorema 1.3.5. *Para z cerca de 0, se tiene*

$$\wp(z) = \frac{1}{z^2} + 3E_4 z^2 + 5E_6 z^4 + \cdots = \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) E_{2k+2} z^{2k}.$$

Demostración. De la definición de $\wp$, si reemplazamos w por $-w$, el valor de la suma no cambia. Esto es

$$\wp(z) = \frac{1}{z^2} + \sum_{w \in \Lambda^*} \left(\frac{1}{(z+w)^2} - \frac{1}{w^2} \right) = \frac{1}{z^2} + \sum_{w \in \Lambda^*} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right), \quad (1.7)$$

donde $w = n + m\tau$.

Para w con $|w| < 1$, si derivamos la serie geométrica

$$\frac{1}{1-w} = \sum_{n=0}^{\infty} w^n,$$

obtenemos que

$$\frac{1}{(1-w)^2} = \sum_{n=0}^{\infty} n w^{n-1} = \sum_{n=0}^{\infty} (n+1) w^n.$$

Si aplicamos la igualdad anterior a $\dfrac{z}{w}$ con z cercano a cero se sigue que

$$\frac{1}{\left(1-\dfrac{z}{w}\right)^2}=\frac{w^2}{(z-w)^2}=\sum_{n=0}^{\infty}(n+1)\left(\frac{z}{w}\right)^n=1+\sum_{n=1}^{\infty}(n+1)\left(\frac{z}{w}\right)^n.$$

Es decir,

$$\frac{1}{(z-w)^2}=\frac{1}{w^2}+\frac{1}{w^2}\sum_{n=1}^{\infty}(n+1)\left(\frac{z}{w}\right)^n$$

y sustituyendo en (1.7),

$$\wp(z)=\frac{1}{z^2}+\sum_{w\in\Lambda^*}\sum_{n=1}^{\infty}(n+1)\frac{z^n}{w^{n+2}}=\frac{1}{z^2}+\sum_{n=1}^{\infty}(n+1)\left(\sum_{w\in\Lambda^*}\frac{1}{w^{n+2}}\right)z^n$$

$$=\frac{1}{z^2}+\sum_{n=1}^{\infty}(n+1)E_{n+2}z^n=\frac{1}{z^2}+\sum_{k=1}^{\infty}(2k+1)E_{2k+2}z^{2k},$$

donde hemos usado que la serie de Eisenstein suma cero para subíndices impares. □

De este teorema, para z cerca de 0 tenemos las expansiones

$$\begin{aligned}\wp'(z)&=-\frac{2}{z^3}+6E_4z+20E_6z^3+\dots\\(\wp'(z))^2&=\frac{4}{z^6}-24\frac{E_4}{z^2}-80E_6+\dots\\(\wp'(z))^3&=\frac{1}{z^6}+9\frac{E_4}{z^2}+15E_6+\dots,\end{aligned}$$

por lo que la diferencia $(\wp'(z))^2-4(\wp(z))^3+60E_4\wp(z)+140E_6$ es holomorfa cerca de 0. De hecho, es igual a 0 en el origen. Ya que esta diferencia es doblemente periódica, podemos concluir que es constante y por tanto, igual a 0. Esto prueba el siguiente corolario.

Corolario 1.3.6. *Si $g_2=g_2(\tau)\coloneqq 60E_4$ y $g_3=g_3(\tau)\coloneqq 140E_6$, entonces se satisface*

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

Luego, por ser $\wp(z)$ solución de la ecuación diferencial $(y')^2=4y^3-g_2y-g_3$, g_2 y g_3 determinan $\wp$. Recíprocamente, si conocemos $\wp$, por su propia definición, conoceremos el retículo. Por ello, decimos que g_2 y g_3 son los **invariantes del retículo**. En lo que sigue del trabajo, escribiremos $g_2(\Lambda)$ o $g_3(\Lambda)$ si queremos enfatizar la dependencia del retículo Λ generado por 1 y

τ. No obstante, más adelante veremos cómo definir estos invariantes para un retículo Λ de períodos generales w_1 y w_2 tales que $\Im\left(\frac{w_2}{w_1}\right) > 0$.

Demostramos ahora un resultado que nos servirá para estudiar la conexión del polinomio cúbico $4x^3 - g_2x - g_3$ con las curvas elípticas.

Teorema 1.3.7. *El polinomio $4x^3 - g_2x - g_3$ tiene tres raíces distintas. De forma equivalente, el discriminante del anterior polinomio $\triangle = g_2^3 - 27g_3^2$ no se anula.*

Demostración. Sea ρ un paralelogramo especial tal que en su interior el único polo de $\wp$ es $z = 0$. Notar que esto es posible desplazando si es necesario el paralelogramo fundamental ρ_* al paralelogramo ρ_a de vértices $a, a + w_1, a + w_1 + w_2$ y $a + w_2$, con $a = -\epsilon w_1 - \delta w_2$, siendo ϵ y δ tales que $0 < \epsilon, \delta < \frac{1}{2}$.

El paralelogramo ρ_a es entonces

$$\rho_a = \{a + tw_1 + sw_2 : 0 \leq t, s < 1\}$$

y por tanto los números $\frac{w_1}{2}$, $\frac{w_2}{2}$ y $\frac{w_1 + w_2}{2}$ no pertenecen al retículo, pues recordemos que dos puntos del retículo se obtienen multiplicando por números enteros.

Sean los números complejos

$$e_1 \coloneqq \wp\left(\frac{w_1}{2}\right), \quad e_2 \coloneqq \wp\left(\frac{w_2}{2}\right), \quad e_3 \coloneqq \wp\left(\frac{w_1 + w_2}{2}\right).$$

Si consideramos por ejemplo la función

$$\wp(z) - e_1,$$

tal función tiene un cero en $z = \frac{w_1}{2}$. Además, tiene un polo doble en $z = 0$. Luego, necesariamente ha de tener otro cero en ρ_a, digamos u. Por ser $u \in \rho_a$, u es de la forma

$$u = a + t_0w_1 + s_0w_2 = (-\epsilon + t_0)w_1 + (-\delta + s_0)w_2.$$

Del teorema 1.1.8 se tiene que

$$(u + \frac{w_1}{2} - 2 \cdot 0) \in \Lambda$$

y entonces

$$u = (-\epsilon + \frac{1}{2} + t_0)w_1 + (-\delta + s_0)w_2 \in \Lambda,$$

por lo que $-\epsilon + \frac{1}{2} + t_0$ y $-\delta + s_0$ son números enteros y como cumplen

$$0 < -\epsilon + \frac{1}{2} + t_0 < \frac{3}{2}, \qquad -\frac{1}{2} < -\delta + s_0 < 1,$$

se tiene que

$$-\epsilon + \frac{1}{2} + t_0 = 1 \quad \text{y} \quad -\delta + s_0 = 0.$$

Volviendo a la expresión de u, se obtiene que $u = \dfrac{w_1}{2}$, por lo que $\dfrac{w_1}{2}$ es un cero doble de $\wp(z) - e_1$.

Análogamente se prueba que $\dfrac{w_2}{2}$ y $\dfrac{w_1 + w_2}{2}$ son los únicos ceros dobles de $\wp(z) - e_2$ y $\wp(z) - e_3$ respectivamente. Además, una prueba similar demuestra que los e_i son distintos entre sí.

La función

$$g(z) \coloneqq 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3)$$

es de orden 6. También es de orden 6 la función $(\wp')^2$, pues recordemos que tal función tiene en $z = 0$ un polo de orden 3.

El cociente

$$\frac{(\wp'(z))^2}{g(z)} = \frac{4\wp^3(z) - g_2\wp(z) - g_3}{4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3)}$$

es una función elíptica que no tiene polos, por lo que ha de ser constante. De hecho, calculando su límite en 0, vemos que la constante ha de ser 1, obteniendo la siguiente igualdad

$$4\wp^3(z) - g_2\wp(z) - g_3 = 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3)$$

y como $e_i \neq e_j$ si $i \neq j$, haciendo $x = \wp(z)$, el polinomio

$$4x^3 - g_2x - g_3$$

tiene tres raíces distintas.

Por último, veamos que esto equivale a decir que el discriminante $\triangle = g_2^3 - 27g_3^2$ del anterior polinomio no se anula. En efecto, hemos visto que $P(x) \coloneqq 4x^3 - g_2x - g_3$ factoriza como

$$P(x) = 4(x - e_1)(x - e_2)(x - e_3)$$

y una cuenta directa muestra que

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

□

Para finalizar este capítulo, vamos a ver que las funciones elípticas son el resultado de la inversión de integrales elípticas.

Como vimos con las funciones de Jacobi, la inversa de la función $\wp$ de Weierstrass es también una integral elíptica.

En general, una integral elíptica de primera clase se define como una del tipo

$$\int_a^z \frac{dt}{\sqrt{P(t)}}$$

donde P es un polinomio de grado 3 o 4 sin raíces múltiples. En principio, el valor de tal integral depende tanto de la elección de la raíz cuadrada como de la curva que conecta a con z.

El siguiente teorema nos dice que la función inversa de una integral elíptica de primera clase es una función elíptica. En este trabajo nos centraremos en el caso en el que $P(t)$ es un polinomio de grado 3.

Teorema 1.3.8. *Si $P(t)$ es un polinomio de grado 3 con raíces simples, existe una función elíptica f no constante tal que si $\Omega \subset \mathbb{C}$ es un abierto donde f es invertible y si $g : f(\Omega) \longrightarrow \Omega$ es la inversa de la restricción $f : \Omega \longrightarrow f(\Omega)$, entonces*

$$g'(z) = \frac{1}{\sqrt{P(z)}}.$$

Demostración. Observemos que para cualquier punto $a \in \mathbb{C}$ en el que $f(a) \neq \infty$ y $f'(a) \neq 0$, podemos encontrar un entorno Ω de a en las condiciones del teorema.

Si $P(t)$ es un polinomio de grado 3, el método de reducción de Tartaglia-Cardano para ecuaciones cúbicas permite que nos podamos restringir al caso de un polinomio cúbico sin término cuadrático, es decir, podemos restringirnos al caso de un polinomio P de la forma $P(t) = at^3 + bt + c$. Es más, podemos normalizar el coeficiente líder y poner $a = 4$. Si llamamos $b = -g_2$ y $c = -g_3$ con $g_2 = g_2(\Lambda)$ y $g_3 = g_3(\Lambda)$ para cierto retículo $\Lambda \subset \mathbb{C}$, obtenemos la llamada forma normal de Weierstrass

$$P(t) = 4t^3 - g_2 t - g_3.$$ [5]

Ahora bien, la función f definida como $f(z) \coloneqq \wp(z)$ satisface las condiciones del teorema. En efecto, del corolario 1.3.6, tenemos que $\wp$ satisface la ecuación diferencial

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

Luego, para una inversa local de $f = \wp$, digamos g, se tiene que

$$g'(t)^2 = \frac{1}{\wp'(g(t))^2} = \frac{1}{4\wp^3(g(t)) - g_2\wp(g(t)) - g_3} = \frac{1}{P(t)}.$$

□

La teoría de las integrales elípticas fue al principio una teoría puramente real. Dejemos de lado momentáneamente la variable compleja y centrémonos en el marco real. Sea $P(x)$ un polinomio de grado 3 (o 4) con coeficientes reales y supongamos que P no tiene raíces múltiples. Supongamos también que el coeficiente líder es positivo. Entonces, $P(x) > 0$ para todo x suficientemente grande, digamos $x > x_0$. En ese caso podemos considerar la raíz cuadrada positiva $\sqrt{P(x)} > 0$, de forma que la siguiente integral impropia

$$E(x) = -\int_x^\infty \frac{dt}{\sqrt{P(t)}}$$

está bien definida para todo $x > x_0$. La integral es convergente ya que podemos compararla con la integral $\int_1^\infty t^{-s}dt$, integral que converge para todo $s > 1$ (en nuestro caso $s = \frac{3}{2}$). Como la función $E(x)$ es estrictamente creciente, ya que el integrando es positivo, podemos considerar la inversa de $E(x)$, definida sobre un intervalo real adecuado.

Por el teorema anterior, tenemos el siguiente resultado.

[5] La elección $b = -g_2$ y $c = -g_3$ siempre va a ser posible gracias al problema de la inversión de Abel, problema que se planteará y se resolverá en el próximo capítulo.

Teorema 1.3.9. *La función inversa de la integral elíptica real*

$$E(x) = -\int_x^\infty \frac{dt}{\sqrt{P(t)}}, \quad x > x_0$$

con $P(t) = 4t^3 - g_2 t - g_3, g_2 = g_2(\Lambda), g_3 = g_3(\Lambda)$ *con* Λ *cierto retículo de* $\mathbb{C}$*, se puede extender de forma meromorfa a todo* $\mathbb{C}$*, siendo tal inversa la función elíptica* $\wp$ *de Weierstrass asociada al retículo* Λ*.*

En concreto, el teorema anterior nos dice que

$$-\int_{\wp(u)}^\infty \frac{dt}{\sqrt{P(t)}} = u$$

donde $\wp(u)$ varía en un intervalo real (t_0, ∞) y u varía en un correspondiente intervalo real adecuado.

Capítulo 2

Curvas elípticas

Se puede decir sin temor a equivocarnos que las curvas elípticas son uno de los objetos más sorprendentes de las matemáticas. Como comentamos en la introducción histórica, conectan muchas áreas de investigación diferentes, como la teoría de números, la geometría algebraica y el análisis complejo. Al mismo tiempo, sus propiedades aritméticas están estrechamente relacionadas con la teoría de las formas modulares y han visto aplicaciones espectaculares en teoría de números como la prueba de Andrew Wiles del último teorema de Fermat. Son objetos de conjeturas abiertas desde hace mucho tiempo, como la de Birch y Swinnerton-Dyer, conjetura que describe el conjunto de soluciones racionales a las ecuaciones que definen una curva elíptica.

En este capítulo vamos a exponer las propiedades más destacadas de las curvas elípticas. En concreto, vamos a ver cómo la función de Weierstrass $\wp$ transforma cada elemento de $\mathbb{C}$ en un punto de una curva elíptica. Probaremos de hecho que es una biyección si uno se restringe al paralelogramo fundamental ρ identificando los lados opuestos. Además, vamos a ver que a través de $\wp$, podemos inducir una estructura de grupo en la curva elíptica que proviene de la suma de números complejos. Para desarrollar este capítulo nos hemos basado principalmente en [Sil09] y en [Kob12].

Definamos ya sí qué es una curva elíptica y veamos cómo la función de Weierstrass $\wp$ y su derivada $\wp'$ parametriza tales curvas.

Definición 2.1. *Sea $\mathbb{K}$ un cuerpo. Una **curva elíptica sobre** $\mathbb{K}$ es el conjunto de puntos $(x, y) \in \mathbb{K} \times \mathbb{K}$ tales que*

$$E(\mathbb{K}): \quad y^2 + axy + by = x^3 + cx^2 + dx + e, \quad a, b, c, d, e \in \mathbb{K}, \tag{2.1}$$

unido con el punto en el infinito, que denotaremos por $\mathcal{O}$.

A la ecuación definida por $E(\mathbb{K})$ en (2.1) la llamaremos la **forma normal de Weierstrass**.

Un cambio de variables será admisible si es de la forma

$$x = u^2 x'$$
$$y = u^3 y',$$

siendo $u \in \mathbb{K}$ elemento no nulo.

Para ciertos $A, B \in \mathbb{K}$, es posible reducir mediante cambios de variables admisibles cada curva elíptica dada en la forma normal de Weierstrass a una única forma simplificada del tipo

$$y^2 = 4x^3 - Ax - B, \tag{2.2}$$

por lo que es costumbre llamar a la ecuación dada en (2.2) la **forma simplificada de Weierstrass**. La demostración de esta reducción la podemos encontrar en [Sil09] (capítulo III, *§1*). En esta línea, dadas dos curvas elípticas, diremos que son **isomorfas** si se puede pasar de una a otra mediante cambios de variable admisibles.

La definición de curva elíptica requiere que la curva no sea singular. Geométricamente, esto significa que el grafo no tiene vértices, autointersecciones ni puntos aislados. Analíticamente, esto se cumple si el polinomio $4x^3 - Ax - B$ tiene tres raíces distintas, o de forma equivalente, si el discriminante del polinomio anterior no se anula, en virtud del teorema 1.3.7. Notemos que estamos identificando implícitamente los coeficientes A y B como los invariantes g_2 y g_3, pero más adelante veremos que el problema de la inversión de Abel justifica esta identificación. No obstante, esta interpretación geométrica se verá más clara cuando al final del capítulo tratemos el caso en el que $\mathbb{K} = \mathbb{R}$.

En el capítulo anterior vimos que la función elíptica de Weierstrass $\wp(z)$ satisface la ecuación diferencial

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

donde insistimos en que las funciones $g_2 = g_2(\tau)$ y $g_3 = g_3(\tau)$ son funciones modulares definidas en términos de las series de Eisenstein en el retículo generado por los períodos 1 y τ. Por tanto, las funciones $\wp(z)$ y $\wp'(z)$ son ecuaciones paramétricas para la curva

$$y^2 = 4x^3 - g_2 x - g_3,$$

con $x = \wp(z)$ e $y = \wp'(z)$. De nuevo, el teorema 1.3.7 asegura que el discriminante del polinomio $4x^3 - g_2x - g_3$ no se anula, lo que garantiza que la curva definida por tal polinomio es no singular. Viendo la analogía entre esta curva y la ecuación que define la forma simplificada de Weierstrass, concluimos que en efecto $\wp(z)$ y $\wp'(z)$ permiten parametrizar las curvas elípticas.

2.1 Curvas elípticas complejas

Tal y como hemos visto, la definición de una curva elíptica depende del cuerpo $\mathbb{K}$ que consideremos. Por tanto, la estructura de la curva elíptica $E(\mathbb{K})$ dependerá de la naturaleza de $\mathbb{K}$. Muchas de las preguntas más interesantes se plantean cuando consideramos el cuerpo $\mathbb{Z}/\mathbb{Z}_p$, para p primo, pues aunque parezca ser de los casos más abstractos, tienen una importante aplicación en criptografía, especialmente en el cifrado de claves. En esta sección nos ocuparemos de las curvas elípticas definidas sobre el cuerpo $\mathbb{C}$ de los números complejos, en cuyo estudio podemos emplear la teoría de la variable compleja.

En el capítulo anterior definimos la función de Weierstrass $\wp$, los invariantes g_2 y g_3 y el discriminante $\triangle$ para los períodos 1 y τ, siendo τ un número complejo con parte imaginaria positiva. Podemos deshacer la normalización que hicimos en la sección 1.2 y definir $\wp$, g_2, g_3 y $\triangle$ de forma general para unos períodos cualesquiera w_1 y w_2 linealmente independientes satisfaciendo $\Im\left(\frac{w_2}{w_1}\right) > 0$. Si $\mathbb{H} \subset \mathbb{C}$ es el semiplano superior abierto, podemos encontrar w_1 y w_2 en las condiciones anteriores de forma que $\tau = \frac{w_2}{w_1}$. Luego, si $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$ es el retículo generado por 1 y τ, multiplicando por w_1 podemos pasar al retículo $\Lambda' = \mathbb{Z}w_1 + \mathbb{Z}w_2$. Es natural esperar una relación entre las funciones $\wp(1,\tau)$, $g_2(1,\tau)$, $g_3(1,\tau)$ y $\triangle(1,\tau)$ y las funciones $\wp(w_1,w_2)$, $g_2(w_1,w_2)$, $g_3(w_1,w_2)$ y $\triangle(w_1,w_2)$ respectivamente.

Definición 2.2. *Diremos que dos* ***retículos*** Λ *y* Λ' *son* ***equivalentes*** *si existe un número complejo* λ *no nulo tal que* $\lambda\Lambda = \Lambda'$*. Geométricamente, esto significa que un retículo puede obtenerse a partir del otro mediante una rotación y una homotecia.*

Proposición 2.1.1. *Para dos retículos equivalentes $\Lambda' = \lambda\Lambda$, se tienen las siguientes relaciones*

$$\begin{aligned}
&\wp(z;\Lambda) = \lambda^2\wp(\lambda z;\lambda\Lambda), \quad \wp'(z;\Lambda) = \lambda^3\wp'(\lambda z;\lambda\Lambda),\\
&G_{2k}(\Lambda) = \lambda^{2k}G_{2k}(\lambda\Lambda),\\
&g_2(\Lambda) = \lambda^4 g_2(\lambda\Lambda) \quad y \quad g_3(\Lambda) = \lambda^6 g_3(\lambda\Lambda).
\end{aligned}$$

Demostración. Las fórmulas anteriores se verifican fácilmente a partir de sus definiciones. Demostramos la igualdad definida por $\wp'$, por ejemplo.

$$\begin{aligned}
\wp'(\lambda z;\lambda\Lambda) &= -2\sum_{\lambda w\in\lambda\Lambda-\{0\}}\frac{1}{(\lambda z-\lambda w)^3} = \lambda^{-3}(-2)\sum_{w\in\Lambda-\{0\}}\frac{1}{(z-w)^3}\\
&= \lambda^{-3}\wp'(z;\Lambda).
\end{aligned}$$

□

En lo que sigue del trabajo, usaremos las funciones $\wp, g_2, g_3$ y Λ asociadas al retículo $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$ o al retículo $\Lambda' = \mathbb{Z}w_1 + \mathbb{Z}w_2$ libremente según nos convenga.

Al principio del primer capítulo adelantábamos cómo una función elíptica $f(z)$ se puede considerar como una función definida sobre el toro. Es decir, si $\Lambda = \mathbb{Z}w_1 + \mathbb{Z}w_2 \subset \mathbb{C}$ es el retículo de f generado por los períodos w_1 y w_2, el conjunto Λ es claramente un sugbrupo de $\mathbb{C}$, por lo que el cociente $\mathbb{C}/\Lambda$, identificado como el toro, es un grupo, que además es abeliano. La operación de grupo que consideramos es la suma habitual y notemos que es fácil ver que en efecto se satisfacen los axiomas de grupo ya que cualquier retículo $\Lambda \subset \mathbb{C}$ satisface $-\Lambda = \Lambda$ y $w+\Lambda = \Lambda$ para cualquier w período de Λ.

En el siguiente teorema vamos a ver que existe una biyección entre todo toro complejo $T = \mathbb{C}/\Lambda$ y la curva elíptica $E(\mathbb{C})$ definida naturalmente a partir de los invariantes del retículo Λ.

Teorema 2.1.2. *Sea Λ un retículo y sea $\wp$ la función de Weierstrass asociada al retículo Λ. Si $E(\mathbb{C}) \subset \mathbb{C}_\infty \times \mathbb{C}_\infty$ es la curva elíptica dada por $E(\mathbb{C}) : y^2 = 4x^3 - g_2x - g_3$ con $\triangle = g_2^3 - 27g_3^2 \neq 0$, entonces la función*

$$\begin{aligned}
&\phi : \mathbb{C}/\Lambda \to E(\mathbb{C})\\
&z \mapsto (\wp(z), \wp'(z))\\
&0 \mapsto \infty
\end{aligned}$$

es una biyección.

Demostración. Observemos primero que la imagen de ϕ está contenida en $E(\mathbb{C})$ en virtud del corolario 1.3.6, luego ϕ toma valores en $E(\mathbb{C})$.

Para ver que ϕ es sobreyectiva, sea $(x, y) \in E(\mathbb{C})$. Entonces, $\wp(z) - x$ es una función elíptica no constante y de los teoremas 1.1.2 y 1.1.6 necesariamente ha de tener un cero, digamos $z = a$. Por tanto, $(\wp'(a))^2 = y^2$ y reemplazando a por $-a$ si fuera necesario, obtenemos $\wp'(a) = y$, por lo que $\phi(a) = (x, y)$.

Probemos que ϕ es inyectiva. Por reducción al absurdo, supongamos que z_1 y z_2 son puntos distintos de $\mathbb{C}/\Lambda$ tales que $\phi(z_1) = \phi(z_2)$. Entonces,

$$\begin{aligned} \wp(z_1) &= \wp(z_2) = \alpha \\ \wp'(z_1) &= \wp'(z_2). \end{aligned}$$

Supongamos primero que $2z_1 \notin \Lambda$. Luego, la función $\wp(z) - \alpha$ es una función elíptica de orden 2 que se anula en $z_1, -z_1$ y en z_2, por lo que necesariamente dos de esos valores han de ser congruentes módulo Λ. Como $2z_1 \notin \Lambda$, entonces $z_2 \equiv \pm z_1 \ (mod\ \Lambda)$ para alguna elección de signo. Pero entonces, la igualdad

$$\wp'(z_1) = \wp'(z_2) = \wp'(\pm z_1) = \pm \wp'(z_1)$$

implica que $z_2 \equiv \pm z_1 \ (mod\ \Lambda)$. (Notar que el corolario 1.3.6 asegura $\wp'(z_1) \neq 0$).

De forma similar, si $2z_1 \in \Lambda$, entonces $\wp(z) - \alpha$ tiene un cero doble en z_1 y en z_2, por lo que de nuevo concluimos que $z_2 \equiv \pm z_1 \bmod(\Lambda)$.

□

En la próxima sección, el resultado de que ϕ es una biyección se verá con más claridad cuando demostremos que la suma en números complejos induce a través de $\wp(z)$ una ley de grupo en la curva elíptica.

En realidad, el anterior teorema posee un recíproco. Es decir, dada una curva elíptica compleja $E(\mathbb{C})$ vamos a ver que existe un retículo de períodos Λ de forma que es posible establecer una biyección entre el toro $\mathbb{C}/\Lambda$ y la curva $E(\mathbb{C})$ a través de una función ϕ definida de forma análoga. Para ello, planteamos primero el siguiente problema.

Fijado un retículo $\Lambda \subset \mathbb{C}$ generado por los períodos w_1 y w_2, hemos visto que los invariantes $g_2(w_1, w_2)$ y $g_3(w_1, w_2)$ determinan $\wp$. Ahora bien, si A y

B son números complejos satisfaciendo $4A^3 - 27B^2 \neq 0$, ¿existirá un retículo $\Lambda \subset \mathbb{C}$ tal que $g_2(\Lambda) = A$ y $g_3(\Lambda) = B$?

El anterior problema planteado es lo que se conoce como el **problema de la inversión de Abel** (1827) y el siguiente teorema muestra que en efecto tiene solución. Antes de enunciar y demostrar tal resultado, definamos una función que es de gran importancia.

Definición 2.3. *Si w_1 y w_2 son números complejos tales $\Im\left(\frac{w_2}{w_1}\right) > 0$, se define la función* ***J-invariante*** *como*

$$J(w_1, w_2) = \frac{g_2^3(w_1, w_2)}{\triangle(w_1, w_2)}.$$

La función anterior está bien definida, pues el teorema 1.3.7 muestra que $\triangle(1, \tau) \neq 0$ para todo τ del semiplano superior $\mathbb{H}$, y por tanto, $\triangle(w_1, w_2) \neq 0$ para w_1 y $w_2 \in \mathbb{C}$ tales que $\Im\left(\frac{w_2}{w_1}\right)$. Por simplicidad, escribiremos $J(\tau)$ para $J(1, \tau)$. Además, por ser las funciones $g_2(\tau)$, $g_3(\tau)$ y $\triangle(\tau)$ analíticas en $\mathbb{H}$, $J(\tau)$ también lo es.

Teorema 2.1.3. ***(Inversión de Abel).*** *Dados dos números complejos a_2 y a_3 tales que $a_2^3 - 27a_3^2 \neq 0$, existen números complejos w_1 y w_2 cuyo cociente no es real tales que*

$$g_2(w_1, w_2) = a_2 \quad y \quad g_3(w_1, w_2) = a_3.$$

Demostración. Consideramos tres casos excluyentes:

$$(1)\ a_2 = 0; \quad (2)\ a_3 = 0; \quad (3)\ a_2 a_3 \neq 0.$$

Caso 1. Si $a_2 = 0$ entonces $a_3 \neq 0$ ya que $a_2^3 - 27a_3^2 \neq 0$. Sea w_1 cualquier número complejo tal que

$$w_1^6 = \frac{g_3(1, \tau)}{a_3}$$

y tomemos $w_2 = \tau w_1$ donde $\tau = e^{\frac{2\pi i}{3}}$. Sabemos que $g_3(1, \tau) \neq 0$ ya que $g_2(1, \tau) = 0$ y $\triangle(1, \tau) = g_2^3 - 27g_3^2 \neq 0$. Luego,

$$g_2(w_1, w_2) = g_2(w_1, w_1\tau) = \frac{1}{w_1^4} g_2(1, \tau) = 0 = a_2$$

y

$$g_3(w_1, w_2) = g_3(w_1, w_1\tau) = \frac{1}{w_1^6} g_3(1, \tau) = a_3.$$

Caso 2. Si $a_3 = 0$ entonces $a_2 \neq 0$. Sea w_1 tal que

$$w_1^4 = \frac{g_2(1,i)}{a_2}$$

y consideremos ahora $w_2 = iw_1$. Entonces,

$$g_2(w_1, w_2) = g_2(w_1, iw_1) = \frac{1}{w_1^4} g_2(1,i) = a_2$$

y

$$g_3(w_1, w_2) = g_3(w_1, iw_1) = \frac{1}{w_1^6} g_3(1,i) = 0 = a_3.$$

Caso 3. Asumamos ahora $a_2 \neq 0$ y $a_3 \neq 0$. La sobreyectividad de la función J en el semiplano superior complejo[1] nos permite tomar $\tau^* \in \mathbb{C}$ con $\Im(\tau^*) > 0$ tal que

$$J(\tau^*) = \frac{a_2^3}{a_2^3 - 27a_3^2}.$$

Notemos que $J(\tau^*) \neq 0$ ya que $a_2 \neq 0$ y que además

$$\frac{J(\tau^*) - 1}{J(\tau^*)} = \frac{27a_3^2}{a_2^3}. \tag{2.3}$$

Para este τ^* elegimos w_1 tal que verifique

$$w_1^2 = \frac{a_2}{a_3} \frac{g_3(1,\tau^*)}{g_2(1,\tau^*)}$$

y tomamos $w_2 = \tau^* w_1$. Entonces,

$$\frac{g_2(w_1,w_2)}{g_3(w_1,w_2)} = \frac{w_1^{-4} g_2(1,\tau^*)}{w_1^{-6} g_3(1,\tau^*)} = w_1^2 \frac{g_2(1,\tau^*)}{g_3(1,\tau^*)} = \frac{a_2}{a_3}$$

y por tanto

$$g_3(w_1,w_2) = \frac{a_3}{a_2} g_2(w_1,w_2). \tag{2.4}$$

Pero de lo anterior,

$$\frac{J(\tau^*) - 1}{J(\tau^*)} = \frac{27g_3^2(w_1,w_2)}{g_2^3(w_1,w_2)} = \frac{27(\frac{a_3}{a_2})^2 g_2^2(w_1,w_2)}{g_2^3(w_1,w_2)} = \frac{27a_3^2}{a_2^2 g_2(w_1,w_2)}$$

y comparando con (2.3) concluimos que $g_2(w_1,w_2) = a_2$ y por (2.4) tenemos $g_3(w_1,w_2) = a_3$. □

[1] La sobreyectividad de J será demostrada en la proposición 3.1.3 del tercer capítulo cuando pensemos en dicha función como función modular.

En la prueba anterior nos hemos apoyado en la función J-invariante. Es natural preguntarse el por qué del calificativo *invariante*. Esto es porque dos curvas elípticas complejas son isomorfas si y sólo si tienen el mismo J-invariante. Recogemos este resultado en el siguiente teorema.

Teorema 2.1.4. *Sean E y E' dos curvas elípticas complejas no singulares definidas por las ecuaciones $y^2 = 4x^3 - g_2x - g_3$ y $y'^2 = 4x'^3 - g'_2x' - g'_3$ respectivamente.*

Entonces, ambas curvas son isomorfas sobre $\mathbb{C}$ si y sólo si:

$$J_E = \frac{g_2^3}{g_2^3 - 27g_3^2} = \frac{g_2'^3}{g_2'^3 - 27g_3'^2} = J_{E'}$$

Demostración. Supongamos que E y E' son dos curvas elípticas complejas isomorfas. Entonces, existe un cambio admisible de la forma

$$\begin{aligned} x &= u^2x' \\ y &= u^3y', \end{aligned}$$

siendo u un elemento no nulo.

Elevando al cuadrado la segunda igualdad, obtenemos

$$4x^3 - g_2x - g_3 = y^2 = u^6y'^2 = 4u^6x'^3 - u^6g'_2x' - u^6g'_3$$

y de la primera igualdad, se tiene que

$$4x^3 - g_2x - g_3 = 4x^3 - u^4g'_2x - u^6g'_3,$$

de donde se deduce que

$$\begin{aligned} g_2 &= u^4g'_2 \\ g_3 &= u^6g'_3. \end{aligned}$$

Luego,

$$J_E = \frac{g_2^3}{g_2^3 - 27g_3^2} = \frac{u^{12}g_2'^3}{u^{12}g_2'^3 - 27u^{12}g_3'^2} = \frac{g_2'^3}{g_2'^3 - 27g_3'^2} = J_{E'}.$$

.

Para probar el recíproco, probemos primero que si los J-invariantes coinciden, los retículos han de ser equivalentes. En ese caso, existiría $\lambda \in$

$\mathbb{C}-\{0\}$ tal que si Λ y Λ' son los retículos asociados a E y E' respectivamente, entonces $\Lambda' = \lambda\Lambda$ y por la proposición 2.1.1, tenemos que

$$\wp(z;\Lambda) = \lambda^2\wp(\lambda z;\Lambda') \quad \text{y} \quad \wp'(z;\Lambda) = \lambda^3\wp'(\lambda z;\Lambda').$$

Ahora bien, ya que $\wp(z)$ y $\wp'(z)$ parametrizan las curvas elípticas, podemos tomar

$$\begin{aligned} x &= \wp(z;\Lambda), \quad y = \wp'(z;\Lambda) \\ x' &= \wp(\lambda z;\Lambda'), \quad y' = \wp'(\lambda z;\Lambda') \end{aligned}$$

y lo anterior equivale a

$$x = \lambda^2 x' \quad \text{y} \quad y = \lambda^3 y',$$

por lo que ambas curvas son isomorfas.

Probemos por tanto que la igualdad de las funciones J implica la equivalencia de los retículos. Para ello, nos apoyamos en el siguiente lema.

Lema 2.1.5. *Si Λ y Λ' son retículos tales que $J(\Lambda) = J(\Lambda')$, entonces existe $\lambda \in \mathbb{C} - \{0\}$ tal que*

$$g_2(\Lambda) = \lambda^{-4} g_2(\Lambda') \quad y \quad g_3(\Lambda) = \lambda^{-6} g_3(\Lambda').$$

Para demostrar el lema, distingamos tres casos.

Supongamos primero que $g_2(\Lambda')$ y $g_3(\Lambda')$ son distintos de cero. Sea $\lambda \in \mathbb{C} - \{0\}$ tal que $g_2(\Lambda) = \lambda^{-4} g_2(\Lambda')$ y veamos que cumple también $g_3(\Lambda) = \lambda^{-6} g_3(\Lambda')$. Como $J(\Lambda) = J(\Lambda')$, se tiene que

$$\begin{aligned} \frac{g_2(\Lambda')^3}{g_2(\Lambda')^3 - 27g_3(\Lambda')^2} &= \frac{g_2(\Lambda)^3}{g_2(\Lambda)^3 - 27g_3(\Lambda)^2} \\ &= \frac{\lambda^{-12} g_2(\Lambda')^3}{\lambda^{-12} g_2(\Lambda')^3 - 27g_3(\Lambda)^2} \\ &= \frac{g_2(\Lambda')^3}{g_2(\Lambda')^3 - 27\lambda^{12} g_3(\Lambda)^2} \end{aligned}$$

y si despejamos λ^{-12} tenemos que

$$\lambda^{-12} = \left(\frac{g_3(\Lambda)}{g_3(\Lambda')}\right)^2.$$

Luego, reemplanzando en caso de ser necesario λ por $i\lambda$, se tiene que $g_3(\Lambda) = \lambda^{-6} g_3(\Lambda')$, lo que demuestra el lema en este caso.

Supongamos ahora que $g_2(\Lambda') = 0$. Entonces, $g_3(\Lambda') \neq 0$ y por ser $J(\Lambda) = J(\Lambda') = 0$, se tiene que $g_2(\Lambda) = 0$. Así, podemos elegir $\lambda \in \mathbb{C} - \{0\}$ tal que $g_3(\Lambda) = \lambda^{-6} g_3(\Lambda')$.

De forma parecida se razona si $g_3(\Lambda') = 0$. En este caso $J(\Lambda') = J(\Lambda) = 1$, y podemos elegir $\lambda \in \mathbb{C} - \{0\}$ tal que $g_3(\Lambda) = \lambda^{-6} g_3(\Lambda')$, por lo que queda probado el lema.

Consideremos entonces un $\lambda \in \mathbb{C} - \{0\}$ tal que se tiene las igualdades del lema anterior y consideremos los retículos Λ y $\lambda\Lambda'$. De la proposición 2.1.1, tenemos que

$$\begin{aligned} g_2(\lambda\Lambda') &= \lambda^{-4} g_2(\Lambda') = g_2(\Lambda) \\ g_3(\lambda\Lambda') &= \lambda^{-6} g_3(\Lambda') = g_3(\Lambda). \end{aligned}$$

y como los coeficientes g_2 y g_3 determinan $\wp$, $\lambda\Lambda'$ y Λ tienen la misma función de Weierstrass, por lo que necesariamente ha de ser $\lambda\Lambda' = \Lambda$. En otras palabras, Λ y Λ' son equivalentes.

□

En consecuencia, se tiene el recíproco del teorema 2.1.2 que buscábamos.

Corolario 2.1.6. *Sea $E(\mathbb{C})$ una curva elíptica. Entonces existe un retículo de períodos $\Lambda \subset \mathbb{C}$ tal que*

$$\begin{aligned} \phi : \mathbb{C}/\Lambda &\to E(\mathbb{C}) \\ \phi(z) &= (\wp(z), \wp'(z)) \\ \phi(0) &= \infty \end{aligned}$$

es una biyección.

Demostración. La existencia es inmediata de los teoremas 2.1.2 y 2.1.3 y la unicidad se tiene salvo homotecia y giro. □

Hemos visto que para cada toro $\mathbb{C}/\Lambda$, podemos calcular su J-invariante $J(\tau)$ siendo el retículo $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$. Como decíamos, se dice *invariante* porque una curva elíptica compleja $E(\mathbb{C})$ se caracteriza por dicho valor J_E. En este sentido, para cada función J-invariante tenemos una clase de isomorfía.

En el siguiente capítulo vamos a pensar en la función J-invariante como función modular y de nuevo veremos su gran importancia. Luego, tiene sentido preguntarse si existe alguna relación entre las curvas elípticas y las formas modulares. Esta cuestión será el motivo del tercer capítulo.

Por último y como curiosidad, el J-invariante también nos permite contar el número de automorfismos sobre una curva elíptica dada, dependiendo si la característica del cuerpo donde la curva está definida es 2 ó 3. Una prueba de este resultado se puede encontrar en el capítulo III de [Sil09].

2.2 La ley de grupo

La correspondencia entre $\mathbb{C}/\Lambda$ y la curva elíptica $E(\mathbb{C})$ nos lleva a definir la ley de la suma para los puntos de la curva elíptica. Si Q_1 y Q_2 son dos puntos de la curva, hemos visto que existen unos únicos u_1 y $u_2 \in \mathbb{C}$ tales que

$$Q_1 = (\wp(u_1), \wp'(u_1))$$
$$Q_2 = (\wp(u_2), \wp'(u_2)).$$

Definimos entonces la suma de dos puntos de la curva como

$$Q_1 \oplus Q_2 = (\wp(u_1 + u_2), \wp'(u_1 + u_2)).$$

Notemos que esto es sólo un caso del principio general: siempre que tengamos una correspondencia biunívoca entre elementos de un grupo conmutativo y elementos de otro conjunto, podemos usar esta correspondencia para dotar a ese conjunto de estructura de **grupo abeliano**. En nuestro caso, ese conjunto es precisamente la curva $E(\mathbb{C})$. Es más, este hecho no sólo ocurre cuando consideramos el cuerpo $\mathbb{C}$, sino para cualquier otro cuerpo finito $\mathbb{K}$.

Estamos listos para deducir el procedimiento geométrico para sumar dos puntos de una curva elíptica definida sobre cualquier cuerpo $\mathbb{K}$.

Sea ρ un paralelogramo especial de forma que el único polo de $\wp$ en el interior de ρ sea $z = 0$ y consideremos los puntos Q_1 y Q_2 de la curva anteriormente parametrizados, con u_1 y $u_2 \in \mathbb{K}$.

Supongamos en primer lugar que Q_1 y Q_2 son distintos y que además no son puntos simétricos respecto al eje OX. Denotamos este caso por $Q_1 \neq \ominus Q_2$. El caso donde se da la igualdad lo estudiaremos luego.

Sea $y = mx + b$ la recta que une Q_1 con Q_2. Entonces, se cumple

$$\wp'(u_1) = m\wp(u_1) + b$$
$$\wp'(u_1) = m\wp(u_1) + b,$$

por lo que u_1 y u_2 son ceros de la función

$$\wp'(z) - (m\wp(z) + b).$$

Por ser $\wp'$ de orden 3, ha de existir $u_3 \in \rho$ cero de la función anterior. En virtud del teorema 1.1.8,

$$(u_1 + u_2 + u_3 - 3 \cdot 0) \in \Lambda,$$

y por tanto,

$$\wp(u_1 + u_2) = \wp(u_1 + u_2 - u_1 - u_2 - u_3) = \wp(-u_3) = \wp(u_3)$$
$$\wp'(u_1 + u_2) = \wp'(u_1 + u_2 - u_1 - u_2 - u_3) = \wp'(-u_3) = -\wp'(u_3).$$

Si denotamos $Q_3 = (\wp(u_3), \wp'(u_3))$ al tercer punto de corte de la recta que une Q_1 con Q_2 con la curva elíptica, hemos obtenido la siguiente relación

$$Q_1 \oplus Q_2 = (\wp(u_3), -\wp'(u_3)).$$

Tratemos el caso que mencionamos antes: Q_1 y Q_2 son simétricos respecto al eje OX, esto es $Q_1 = \ominus Q_2$.
Recordando la paridad de $\wp$ y la imparidad de $\wp'$, se tienen las siguientes relaciones

$$Q_1 = (\wp(u_1), \wp'(u_1))$$
$$Q_2 = (\wp(u_2), \wp'(u_2)) = (\wp(u_1), \wp'(-u_1)) = (\wp(-u_1), \wp'(-u_1)).$$

Por ser $x = \wp(u_1)$ la recta que une Q_1 con Q_2, la función $\wp(z) - \wp(u_1)$ tiene ceros en u_1 y u_2, siendo cada uno un cero simple. Luego,

$$u_1 + u_2 \in \Lambda$$

y además se tiene que

$$\wp(u_1 + u_2) = \wp(u_1 + u_2 - u_1 - u_2) = \wp(0)$$
$$\wp'(u_1 + u_2) = \wp(u_1 + u_2 - u_1 - u_2) = \wp'(0).$$

Luego,

$$Q_1 \oplus Q_2 = (\wp(0), \wp'(0)) = \infty,$$

pues recordemos que la función de Weierstrass tiene un polo en $z = 0$. Por ello, consideraremos al punto del infinito, denotado como $\mathcal{O}$, elemento neutro del grupo.

En definitiva, si P y Q son dos puntos de la curva $E(\mathbb{K})$, denotando $(P * Q)$ al punto resultante de intersectar E con la recta que une P con Q, hemos obtenido que

$$P \oplus Q = \ominus(P * Q)$$

y además

$$P \oplus \mathcal{O} = \mathcal{O} \oplus P = P.$$

El elemento inverso de P es su simétrico respecto al eje OX, que como ya hemos dicho lo denotamos por $\ominus P$.

No obstante, probar la asociatividad ya no es tan sencillo. Una prueba geométrica la muestra el libro [Ful08].

Para visualizar la operación de grupo, asumamos por el momento que $\mathbb{K} = \mathbb{R}$. Los cálculos anteriores muestran que para sumar los puntos P y Q de $E(\mathbb{R})$, dibujamos la línea que los une, buscamos el tercer punto de intersección de esa línea con la curva y luego tomamos el punto simétrico del otro lado del eje OX (véase la figura 2.1). De hecho, hubiera sido posible definir la ley de grupo de esta manera geométrica en primer lugar y probar directamente que se satisfacen los axiomas de grupo.

Finalmente, si el discriminante se anula, tenemos raíces múltiples, de acuerdo con el teorema 1.3.7. Luego, en este caso estamos ante una curva singular. Como hay tres raíces en total, pueden ocurrir dos casos: las tres raíces son iguales o dos de ellas lo son. La figura 2.2 muestra dos ejemplos de curvas singulares.

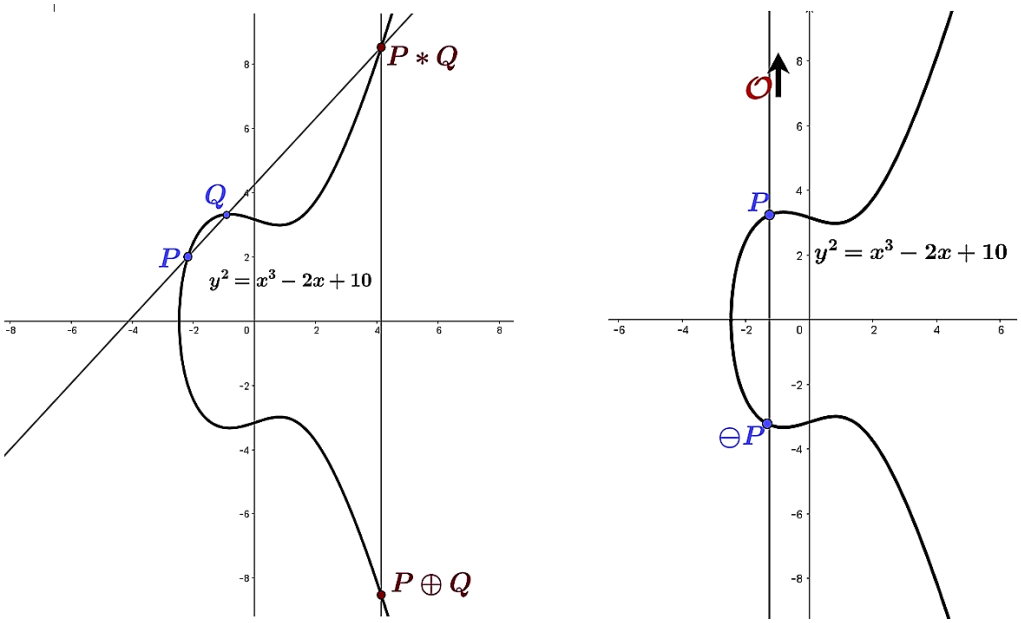

Figura 2.1: Suma de dos puntos en una curva elíptica real.

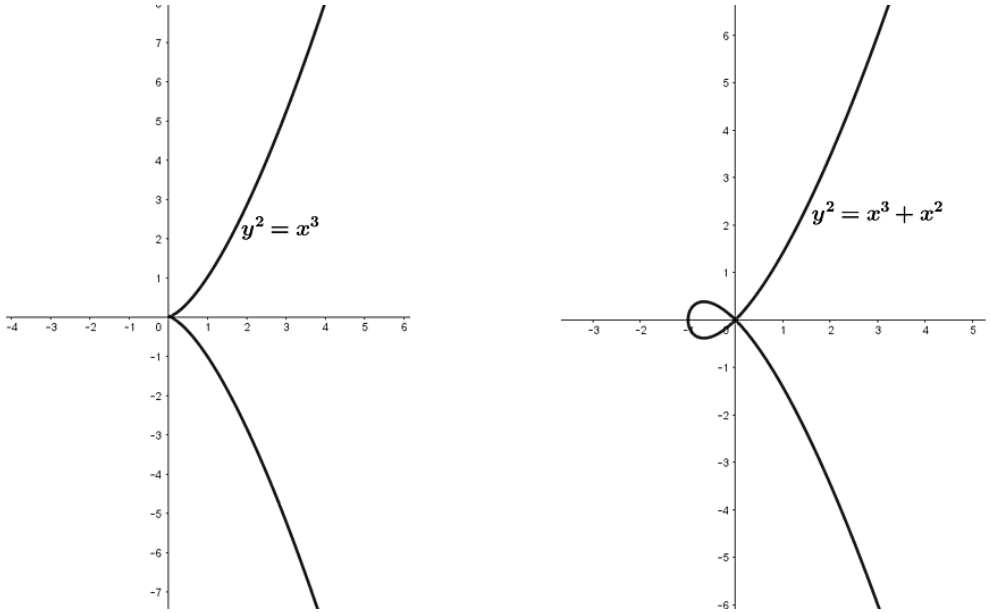

Figura 2.2: Curvas singulares.

Capítulo 3

Formas modulares

Se atribuye a Eichler (1912) la siguiente cita:

> Sólo hay cinco operaciones aritméticas elementales: suma, resta, multiplicación, división y formas modulares.

La teoría de las formas modulares es una de las herramientas más poderosas de la teoría de números. Una de las aplicaciones más sencillas la encontramos en la representación de todo número natural como suma de cuatro cuadrados. Aunque la prueba clásica, dada por Lagrange en 1770, se basa en el método del descenso infinito, no fue hasta 1834 cuando Jacobi demostró el resultado usando formas modulares, dando además una fórmula exacta para el número de representaciones según la paridad del número natural. Por otro lado, gracias a las formas modulares, se han desarrollado conceptos importantes en la teoría algebraica de números, la geometría aritmética y la teoría de representación de Galois.

Durante 358 años, el *Último Teorema de Fermat* fue un problema abierto muy atractivo reconocido en matemáticas. En él se establecía que para los naturales $n \geq 3$, la ecuación $a^n + b^n = c^n$, siendo a, b y c números enteros positivos, no tiene solución. El interés por demostrar esta conjetura propició el auge de la teoría de las formas modulares, que en resumen, son funciones definidas en el semiplano superior complejo que satisfacen ciertas condiciones de transformación y holomorfía.

En este capítulo vamos a definir las formas modulares exponiendo algunas de sus propiedades más relevantes. Como complemento, expondremos cómo se relacionan las formas modulares con las curvas elípticas definidas sobre $\mathbb{Q}$ a través de las funciones L. Tal relación se materializa con el teorema de Modularidad, finalmente probado en 1999, que en definitiva, nos dice que las

curvas elípticas racionales provienen de formas modulares.

Para escribir este capítulo nos hemos apoyado principalmente en [Apo89], así como en [Kob12].

3.1 El grupo modular

Sea $\mathbb{C}$ el plano complejo y denotemos por $\mathbb{H}$ el semiplano superior, es decir, $\mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\}$. Empecemos recordando las transformaciones de Möbius, estudiadas con detalle en la asignatura Variable Compleja.

Una función f se dice que es una **transformación de Möbius** si es de la forma

$$f(z) = \frac{az+b}{cz+d}, \quad \text{donde } a,b,c,d \in \mathbb{C} \text{ y } ad-bc \neq 0.$$

Podemos extender la ecuación que define $f(z)$ a $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$ y obtener un automorfismo $f : \mathbb{C}_\infty \longrightarrow \mathbb{C}_\infty$ si definimos

$$f\left(\frac{-d}{c}\right) \coloneqq \infty \quad \text{y} \quad f(\infty) \coloneqq \frac{a}{c} \quad \text{si } c \neq 0$$

y

$$f(\infty) \coloneqq \infty \quad \text{si } c = 0.$$

Las transformaciones de Möbius son analíticas en $\mathbb{C}_\infty$ excepto en el polo simple $z = \dfrac{-d}{c}$. Además, son funciones conformes y envían circunferencias generalizadas en circunferencias generalizadas, entendiendo tal circunferencia como una en el plano complejo $\mathbb{C}$ o bien como la unión de una recta en $\mathbb{C}$ con el punto del infinito.

Para cada transformación de Möbius, asociamos la matriz

$$A = \begin{pmatrix} a & b \\ c & d, \end{pmatrix}$$

de forma que $det(A) \neq 0$.

Si A y B son matrices asociadas a las transformaciones f y g respectivamente, es fácil verificar que el producto matricial AB está asociado a la composición $f \circ g$.

La matriz identidad $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ está asociada a la transformación identidad $f(z) = z = \dfrac{1z+0}{0z+1}$ y la matriz inversa $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ está asociada a la inversa de f, $f^{-1}(z) = \dfrac{dz-b}{-cz+a}$. Por tanto, las transformaciones de Möbius forman un grupo bajo composición.

Estamos ya listos para definir el grupo modular.

Definición 3.1. *El conjunto de transformaciones de Möbius*

$$f(z) = \frac{az+b}{cz+d}$$

con a, b, c y d enteros satisfaciendo $ad - bc = 1$ se llama ***grupo modular*** *y lo denotaremos indistintamente por Γ o $SL_2(\mathbb{Z})$.*

Como decíamos, el grupo modular puede ser representado por las matrices de coeficientes enteros

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

con $det(A) = 1$. En esta línea, A y $-A$ representan la misma transformación en el grupo modular. Por este motivo, es frecuente denotar al grupo modular como el cociente $SL_2(\mathbb{Z})/\pm I$, siendo I la matriz identidad de orden 2.
De aquí en adelante, no haremos distinción entre una matriz y su transformación asociada, es decir, si la matriz A viene dada por $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, escribiremos

$$Az = \frac{az+b}{cz+d}.$$

Una propiedad importante es la invarianza del semiplano superior complejo $\mathbb{H}$ mediante cualquier función f del grupo modular. Esto es, si $\tau \in \mathbb{H}$, entonces $f(\tau) \in \mathbb{H}$, en virtud del siguiente cálculo

$$\Im(f(\tau)) = \Im\left(\frac{a\tau+b}{c\tau+d}\right) = \frac{ad-bc}{|c\tau+d|^2}\Im(z) = \frac{1}{|c\tau+d|^2}\Im(\tau). \tag{3.1}$$

Como adelantábamos en el primer capítulo cuando definimos las series de Eisenstein, el siguiente teorema muestra que Γ es generado por la traslación T e inversión S dadas por las ecuaciones

$$T\tau = \tau + 1 \quad \text{y} \quad S\tau = -\frac{1}{\tau}.$$

Teorema 3.1.1. *El grupo modular $SL_2(\mathbb{Z})$ está generado por las matrices*

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad y \quad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Esto es, cada matriz A de $SL_2(\mathbb{Z})$ puede ser expresada como

$$A = T^{n_1} S T^{n_1} S T^{n_2} ... S T^{n_k}$$

con n_i enteros.

Demostración. Empecemos observando que

$$T^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \quad \text{para todo } k \text{ entero.}$$

Es suficiente considerar las matrices de la forma

$$A = (a_{i,j})_{i,j=1,2} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

de Γ con $c \geq 0$.

Si $c = 0$, entonces $ad = 1$ y $a = d = \pm 1$. En este caso, A es una potencia de T, pues

$$A = \begin{pmatrix} \pm 1 & b \\ 0 & \pm 1 \end{pmatrix} = \begin{pmatrix} 1 & \pm b \\ 0 & 1 \end{pmatrix} = T^{\pm b}.$$

Si $c = 1$, tenemos $b = ad - 1$, por lo que A es

$$A = \begin{pmatrix} a & ad-1 \\ 1 & d \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} = T^a S T^d.$$

Ahora, supongamos cierto el teorema para todas las matrices A con el coeficiente $a_{21} < c$ con $c \geq 1$. Ya que $ad - bc = 1$, por la identidad de Bézout tenemos $mcd(c, d) = 1$. Dividiendo d por c, obtenemos que

$$d = cq + r, \quad \text{con } 0 < r < c.$$

Luego,

$$AT^{-q} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -q \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & -aq + b \\ c & r \end{pmatrix}$$

y

$$AT^{-q}S = \begin{pmatrix} a & -aq + b \\ c & r \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -aq + b & -a \\ r & -c \end{pmatrix}.$$

Por hipótesis de inducción, la última matriz es un producto de potencias de T y S, por lo que A también. □

A continuación, vamos a ver cómo es posible formar una teselación de $\mathbb{H}$ mediante el grupo modular Γ, para entender de alguna manera la geometría subyacente a los elementos del grupo modular.

Definición 3.2. *Sea G un subgrupo del grupo modular Γ. Diremos que dos puntos τ y τ' del semiplano superior $\mathbb{H}$ son **G-equivalentes** si $\tau' = A\tau$ para algún automorfismo $A \in G$.*

La relación anterior es de equivalencia ya que G es un grupo. Además, esta relación de equivalencia divide el semiplano $\mathbb{H}$ en una colección disjunta de clases de equivalencia llamadas *órbitas*, siendo la órbita $G\tau$ el conjunto de números complejos de la forma $A\tau$ donde $A \in G$.

Definición 3.3. *Diremos que F es un **dominio fundamental para el sugbrupo G** de Γ si satisface las siguientes propiedades*

1. *Cada $z \in \mathbb{H}$ es G-equivalente a un punto de F.*
2. *No hay dos puntos en el interior de F distintos que sean G-equivalentes.*

El ejemplo más famoso de dominio fundamental para el grupo modular Γ se define como

$$F \coloneqq \left\{ z \in \mathbb{H} : \frac{-1}{2} \leq \Re(z) \leq \frac{1}{2}, |z| \geq 1 \right\}$$

Recogemos el resultado anterior en la siguiente proposición.

Proposición 3.1.2. *La región F definida anteriormente es un dominio fundamental para Γ.*

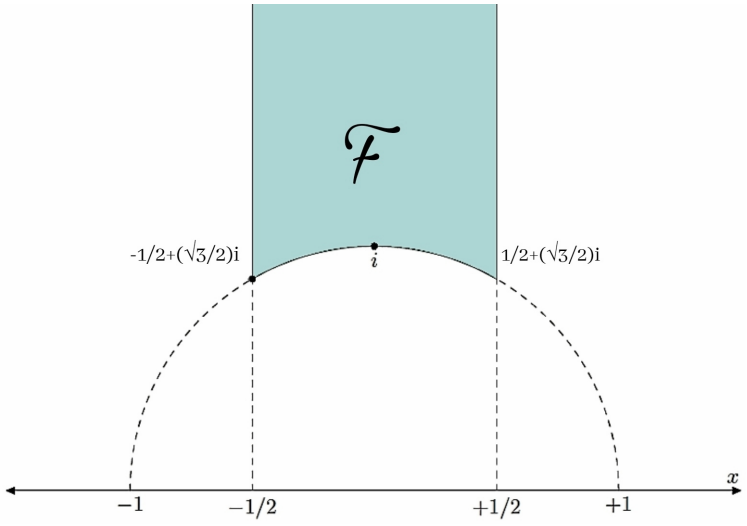

Figura 3.1: Dominio fundamental F para el grupo modular $SL_2(\mathbb{Z})$

Demostración. Probemos primero que cada $z \in \mathbb{H}$ es Γ-equivalente a un punto de F. Sea z un punto de $\mathbb{H}$ fijado. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, entonces $\Im(Az) = \dfrac{1}{|cz+d|^2}\Im(z)$, por lo obtenido en (3.1). Como c y d varían a lo largo de los enteros, el número complejo $cz+d$ toma los valores en la red generada por 1 y z. Por ser c y d no nulos a la vez, $|cz+d|$ está acotado inferiormente por un valor $k > 0$. Por tanto, existe cierto $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ tal que $\Im(Az)$ es máximo. Sustituyendo A por T^jA para algún j adecuado, podemos suponer sin pérdida de generalidad que Az está en la banda $\dfrac{-1}{2} \leq \Re(Az) \leq \dfrac{1}{2}$. Ahora bien, si Az no estuviera en F, es decir, si tuviéramos $|Az| < 1$, entonces de nuevo por (3.1) tendríamos

$$\Im(SAz) = \frac{1}{|Az|^2}\Im(Az) > \Im(Az),$$

que contradice nuestra elección de $A \in \Gamma$, pues $\Im(Az)$ es máximo. Por tanto, existe $A \in \Gamma$ tal que $Az \in F$.

Probemos ahora que no hay dos puntos distintos en el interior de F que sean Γ-equivalentes. Por reducción al absurdo, supongamos que $z_1 \in F$ y $z_2 \in F$ son Γ-equivalentes. Notemos que no suponemos que z_1 y z_2 sean necesariamente distintos o que ambos estén en el interior de F. Supongamos

por ejemplo que $\Im(z_2) \geq \Im(z_1)$. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ tal que $z_2 = Az_1$. Ya que $\Im(z_2) \geq \Im(z_1)$, por (3.1) tenemos $|cz_1 + d| \leq 1$. Como z_1 está en F y d es real, fijándonos en la figura 3.1, es fácil ver que si $|c| \geq 2$, la desigualdad anterior $|cz_1 + d| \leq 1$ no puede darse. Esto lleva a estudiar los casos:

1. $c = 0, d = \pm 1$.
2. $c = \pm 1, d = 0$ y z_1 en la circunferencia unidad.
3. $c = d = \pm 1$ y $z_1 = \dfrac{-1}{2} + \dfrac{\sqrt{3}}{2}i$
4. $c = -d = \pm 1$ y $z_1 = \dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i$

En el primer caso, A (o $-A$) es una traslación T^j, pero tal A puede llevar un punto de F a otro punto de F si y sólo si es la identidad o $j = \pm 1$ y los puntos z_1 y z_2 están en las rectas verticales $\Re(z) = \pm\dfrac{1}{2}$.
En el segundo caso es fácil ver que $A = \pm T^a S$ con $a = 0$ o $a = \pm 1$. Si $a = 0$, z_1 y z_2 están en el círculo unidad simétricamente posicionados respecto al eje imaginario y si $a = \pm 1$, $z_1 = z_2 = \pm\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i$.
Para el tercer caso, A puede ser escrita como $\pm T^a \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, y si tal automorfismo lleva $z_1 \in F$ a $z_2 \in F$ ha de ser $a = 0$ y $z_2 = z_1 = \dfrac{-1}{2} + \dfrac{\sqrt{3}}{2}i$ o bien $a = 1$ y $z_2 = z_1 = \dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i$.
El último caso es tratado de forma análoga al anterior.

Concluimos de este modo que en ningún caso z_1 y z_2 pueden pertenecer al interior de F, a no ser que $\pm A$ sea la identidad y $z_1 = z_2$.

□

De hecho, en la prueba anterior hemos demostrado un resultado más preciso. En concreto, hemos visto que dos puntos distintos z_1 y z_2 de la frontera de F son Γ-equivalentes si y sólo si se da alguno de los dos casos siguientes: $\Re(z_1) = \pm\dfrac{1}{2}$ y $z_2 = z_1 = \pm 1$ o bien si z_1 está en la circunferencia unidad y $z_2 = \dfrac{-1}{z_1}$.

La siguiente figura muestra el dominio fundamental F y algunas de sus imágenes bajo transformaciones del grupo modular, que como sabemos, se pueden escribir en términos de S y T.

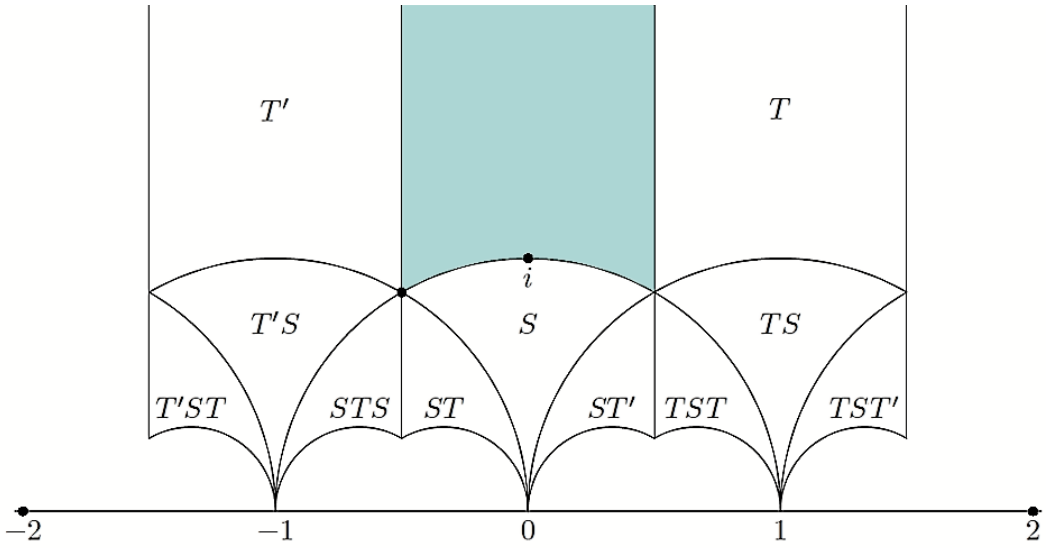

Figura 3.2: Teselación del plano complejo por elementos de $SL_2(\mathbb{Z})$

En cierto sentido, las órbitas de los elementos del grupo modular Γ son un análogo a los paralelogramos períodos de las funciones elípticas. En este caso, el grupo era el retículo Λ y la acción de un período $w \in \Lambda$ en un punto $z \in \mathbb{C}$ era simplemente $w(z) = w + z$. El dominio fundamental es el paralelogramo $\rho \subset \mathbb{C}$ para el retículo Λ asociado: cada $z \in \mathbb{C}$ es Λ-equivalente a un punto de ρ y no hay dos puntos en el interior de ρ que sean Λ-equivalentes. Por tanto, en un sentido geométrico, las transformaciones de un subgrupo de Γ y las funciones elípticas tienen un comportamiento semejante.

La geometría hiperbólica está íntimamente relacionada con estas teselaciones. De hecho, este tipo de teselaciones se han representado en el arte del siglo XX, gracias a la obra del artista neerlandés Escher, mostrando gran interés en la métrica desarrollada por Poincaré y en las obras del geómetra H. Coxeter. Escher, a través de su obra, pretendía simbolizar el infinito usando esta geometría hiperbólica. En el siguiente mosaico, identificado como el disco unidad $\mathbb{D}$, Escher relaciona las teselaciones de $\mathbb{H}$ y $\mathbb{D}$ y en términos matemáticos, podemos entender esta relación como la imagen de cierta transformación de Möbius.

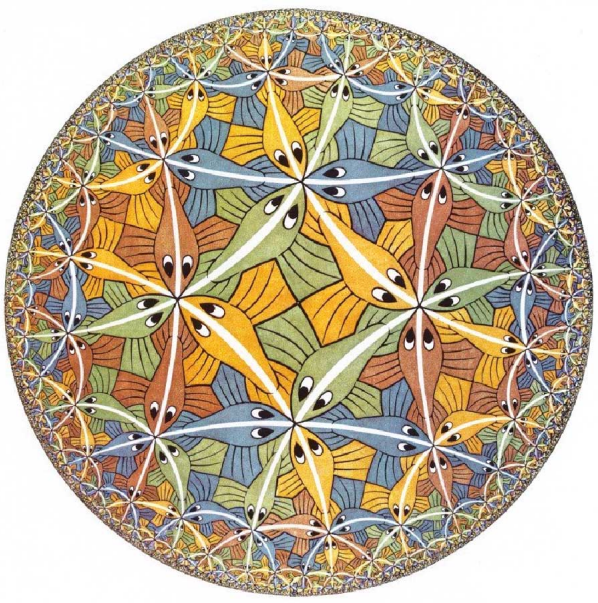

Figura 3.3: El disco de Poincaré ilustrado por Escher.

Antes de definir las formas y funciones modulares, vamos a extender la topología usual en $\mathbb{H}$ a $\mathbb{H}_\infty = \mathbb{H} \cup \{\infty\}$. [1]

Un entorno del $\infty \in \mathbb{H}_\infty$ es $N_C = \{z \in \mathbb{H} : \Im(z) > c\} \cup \{\infty\}$. Si llevamos $\mathbb{H}$ al disco unidad perforado mediante la aplicación

$$z \longrightarrow q := e^{2\pi i z} \tag{3.2}$$

y si aceptamos llevar $\infty \in \mathbb{H}_\infty$ al origen bajo esta aplicación, entonces N_C es la imagen inversa del disco abierto de radio $e^{-2\pi C}$ centrado en el origen. Hemos definido nuestra topología en $\mathbb{H}_\infty$ para hacer la aplicación anterior continua.

El cambio de variable dado en (3.2) de z a q juega un papel importante en la teoría de las funciones modulares. Usaremos este cambio para definir una estructura analítica en $\mathbb{H}_\infty$. En otras palabras, dada una función f en $\mathbb{H}$ de período 1, diremos que es **meromorfa** en ∞ si puede ser expresada como una serie de potencias en la variable q con un número finito de términos negativos no nulos. Es decir, si la función definida por

$$\hat{f}(q) := f(z)$$

[1] Es frecuente denotar al infinito como $i\infty$, ya que se suele pensar como el punto en el infinito del semieje positivo y.

para z tal que $q = e^{2\pi iz}$, admite en un entorno de 0 un desarrollo de la forma

$$\sum_{n\in\mathbb{Z}} a_n q^n$$

con $a_n = 0$ para todo n menor que cierto entero negativo n_0.

Diremos que es **holomorfa** en el ∞ si $a_n = 0$ para todo $n < 0$ y diremos que $f(z)$ **se anula** en ∞ si $f(z)$ es holomorfa en ∞ y $a_0 = 0$.

Definición 3.4. *Dado un entero k, decimos que $f(z)$ es una **función modular de peso k** para $\Gamma = SL_2(\mathbb{Z})$ si $f(z)$ es una función meromorfa en el semiplano superior $\mathbb{H}$ tal que $f(z)$ satisface la relación*

$$f(\gamma z) = (cz + d)^k f(z) \tag{3.3}$$

para todo $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ y además $f(z)$ es meromorfa en el infinito, es decir, la serie de Fourier

$$f(z) = \sum_{n\in\mathbb{Z}} a_n q^n, \quad \text{con } q = e^{2\pi iz},$$

tiene como máximo un número finito de coeficientes a_n no nulos con $n < 0$.

*En tales condiciones, si $f(z)$ es holomorfa en $\mathbb{H}$ y en el infinito, $f(z)$ es una **forma modular de peso k** para $SL_2(\mathbb{Z})$. Denotaremos al conjunto de las formas modulares de peso k por $M_k(\Gamma)$.*

*Finalmente, si la forma modular se anula en el infinito, entonces $f(z)$ se denomina **forma cúspide de peso k** para Γ. Denotaremos al conjunto de estas funciones por $S_k(\Gamma)$.*

En particular, para los elementos $\gamma = T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ y $\gamma = S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, la relación (3.3) es

$$f(z + 1) = f(z) \tag{3.4}$$

$$f\left(\frac{-1}{z}\right) = (-z)^k f(z). \tag{3.5}$$

Exponemos ahora una serie de observaciones fáciles de verificar sobre estas definiciones.

1. Las condiciones se conservan bajo la suma y la multiplicación escalar, es decir, el conjunto de funciones, formas modulares y formas cúspides para algún peso k fijo es un espacio vectorial complejo. Además, si f_1 es una función (o forma) de peso k_1 y si f_2 es una función (o forma) de peso k_2, el producto $f_1 f_2$ es una función (o forma) de peso $k_1 + k_2$. El cociente de una función modular de peso k_1 por una función modular de peso k_2 distinta de cero es una función modular de peso $k_1 - k_2$.

2. Otra observación importante es que si γ_1 y γ_2 son dos transformaciones que generan el grupo $SL_2(\mathbb{Z})$ tales que

$$f(\gamma_1 z) = (c_1 z + d_1)^k f(z) \quad \text{y}$$
$$f(\gamma_2 z) = (c_2 z + d_2)^k f(z)$$

siendo $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ y $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$, entonces se tiene $f(\gamma z) = (cz + d)^k f(z)$ para todo $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$, es decir, se tiene (3.3). En efecto, dadas γ_1 y γ_2 en las condiciones anteriores, tenemos que

$$\begin{aligned} f(\gamma_2 \circ \gamma_1(z)) &= (c_2 \gamma_1(z) + d_2)^k f(\gamma_1(z)) \\ &= \left(\left(c_2 \frac{a_1 z + b_1}{c_1 z + d_1} + d_2 \right) (c_1 z + d_1) \right)^k f(z) \\ &= ((c_2 a_1 + d_2 c_1) z + c_2 b_1 + d_2 d_1)^k f(z). \end{aligned}$$

Ahora bien, la composición $\gamma_2 \circ \gamma_1$ viene dada por el siguiente producto matricial

$$\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} \dots & \dots \\ c_2 a_1 + d_2 c_1 & c_2 b_1 + d_2 d_1 \end{pmatrix},$$

por lo que si se tiene la relación (3.3) para γ_1 y γ_2, también se tiene para el producto $\gamma_1 \gamma_2$. Por ser γ_1 y γ_2 generadoras del grupo modular Γ, ambas transformaciones se pueden escribir en función de S y T y como Γ es generado por S y T, se tiene (3.3). En particular, las relaciones (3.4) y (3.5) implican (3.3).

3. Si k es impar, no hay funciones modulares de peso k para Γ distintas de cero. Podemos ver esta propiedad sustituyendo $\gamma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ en (3.1). Luego, en esta sección supondremos siempre que k es par.

Esta última observación ocurría con las series de Eisenstein $E_k(\tau) = \sum_{\substack{n,m\in\mathbb{Z}\\(n,m)\neq(0,0)}} \frac{1}{(n+m\tau)^k}$. De hecho, si k es par, vamos a probar que E_k son realmente formas modulares de peso k para Γ.

En el teorema 1.3.4 vimos que E_k es holomorfa en $\mathbb{H}$ si $k \geq 3$. Además, en este teorema también vimos que $E_k(\tau)$ satisface las relaciones

$$E_k(\tau+1) = E_k(\tau) \quad \text{y} \quad E_k(\tau) = \tau^{-k} E_k\left(-\frac{1}{\tau}\right)$$

y por tanto, de acuerdo a la segunda observación anterior, E_k satisface la relación (3.3).

Nos queda ver que E_k es holomorfa en el infinito para concluir que es una forma modular. Para ello, nos fijamos en el comportamiento de $E_k(\tau)$ cuando $\Im(\tau) \longrightarrow \infty$. Ya que la serie que define $E_k(\tau)$ converge uniformemente en el semiplano $\Im(\tau) \geq C$ con $C > 0$ (si $k \geq 3$), tenemos que

$$\begin{aligned}
\lim_{\Im(\tau)\longrightarrow\infty} E_k(\tau) &= \lim_{\Im(\tau)\longrightarrow\infty} \sum_{\substack{n,m\in\mathbb{Z}\\(n,m)\neq(0,0)}} \frac{1}{(n+m\tau)^k} \\
&= \sum_{\substack{n,m\in\mathbb{Z}\\(n,m)\neq(0,0)}} \lim_{\Im(\tau)\longrightarrow\infty} \frac{1}{(n+m\tau)^k} \\
&= \sum_{\substack{m=-\infty\\m\neq 0}}^{+\infty} \sum_{\substack{n=-\infty\\n\neq 0}}^{+\infty} \lim_{\Im(\tau)\longrightarrow\infty} \frac{1}{(n+m\tau)^k} \\
&= \sum_{\substack{n=-\infty\\n\neq 0}}^{+\infty} \frac{1}{n^k} = 2\zeta(k),
\end{aligned}$$

pues términos de la forma $(n+m\tau)^{-k}$ con $m \neq 0$ tienden a cero, mientras que los otros tienden a n^{-k}. Luego, E_k es holomorfa en el infinito y además conocemos su valor. Por tanto, si $k \geq 3$, $E_k(\tau) \in M_k(\Gamma)$, esto es, las series de Eisenstein son formas modulares de peso k para Γ.

A lo largo de nuestro trabajo hemos estudiado dos funciones de gran importancia: el discriminante y el J-invariante. Vamos a ver ahora que el discriminante es en realidad una forma cúspide y el J-invariante una función modular.

Recordemos que de nuestro estudio de la función $\wp$ de Weirstrass asociada a un retículo Λ, se definen los coeficientes $g_2(\Lambda) = 60E_4(\Lambda)$ y $g_3 = 140E_6(\Lambda)$ que aparecen en la ecuación diferencial que satisface $\wp$ (corolario 1.3.6).[2] Si $\tau \in \mathbb{H}$, ya hemos visto que $g_2(\tau)$ y $g_3(\tau)$ son formas modulares para Γ de peso 4 y 6 respectivamente. Por tanto, el discriminante

$$\triangle(\tau) = g_2^3(\tau) - 27g_3^2(\tau)$$

es una forma modular de peso 12 para Γ, de acuerdo con la primera observación descrita anteriormente. Ahora bien, antes obtuvimos la igualdad

$$\lim_{\Im(\tau)\longrightarrow\infty} E_k(\tau) = 2\zeta(k)$$

y recordemos que

$$\zeta(4) = \frac{\pi^4}{90} \quad \text{y} \quad \zeta(6) = \frac{\pi^6}{945}.$$

Por tanto, pensando en $i\infty$ como el punto en el infinito del semieje positivo y, tenemos que

$$g_2(i\infty) = 60E_4(i\infty) = 120\zeta(4) = \frac{4\pi^4}{3}$$

y

$$g_3(i\infty) = 140E_6(i\infty) = 280\zeta(6) = \frac{8\pi^6}{27}.$$

Luego,

$$\triangle(\infty) = \left(\frac{4\pi^4}{3}\right)^3 - 27\left(\frac{8\pi^6}{27}\right)^2 = 0,$$

lo que muestra que $\triangle(\tau)$ es en efecto una forma cúspide de peso 12 para Γ. Además, en puntos distintos al infinito, $\triangle(\tau)$ no se anula en virtud del teorema 1.3.7.

Luego, la función J-invariante definida como

$$J(\tau) = \frac{g_2^3(\tau)}{\triangle(\tau)},$$

es en principio, una función modular de peso 0 para Γ. Ahora bien, acabamos de ver que

$$\lim_{\Im(\tau)\to\infty} \triangle(\tau) = 0$$

[2] Aquí nos referimos a la función de Weierstrass $\wp$ definida para un retículo general, generado por unos períodos linealmente independientes, digamos w_1 y w_2.

y por ser

$$\lim_{\Im(\tau)\to\infty} g_2^3(\tau) = \left(\frac{4\pi^4}{3}\right)^3 \neq 0,$$

entonces se tiene que

$$\lim_{\Im(\tau)\to\infty} J(\tau) = \infty. \tag{3.6}$$

Luego, J tiene un polo en el infinito, por lo que no puede ser forma modular.

Notemos que el resultado de que J tenga un polo en el infinito también se podía haber deducido si nos fijamos en su expansión de Fourier:

$$J(\tau) = \frac{1}{1728}\left(\frac{1}{q} + 744 + 196884q + ...\right).$$

y de esta expansión vemos además que es un polo simple. Este desarrollo de la función J puede encontrarse en la observación 7.4.1 (capítulo I) de [Sil94].

Una vez más, el calificativo *invariante* queda perfectamente justificado pues al ser de peso 0, para todo automorfismo $\gamma \in SL_2(\mathbb{Z})$ y para todo $\tau \in \mathbb{H}$, se satisface

$$J(\gamma(\tau)) = J(\tau).$$

Ahora probaremos la sobreyectividad de la función J- invariante, lo que nos permitirá concluir la demostración del problema de la inversión de Abel (teorema 2.1.3).

Proposición 3.1.3. *La función $J : \mathbb{H} \longrightarrow \mathbb{C}$ es sobreyectiva.*

Demostración. $J(\tau)$ no es constante en $\mathbb{H}$ ya que hay valores para τ que son distintos bajo la acción de Γ. Además, también sabemos que $J(\tau)$ es holomorfa en $\mathbb{H}$. Por el teorema de la aplicación abierta, la imagen de $J(\tau)$ es un conjunto abierto en $\mathbb{C}$. Dado que el único conjunto que es a la vez abierto y cerrado en $\mathbb{C}$ es él mismo por ser conexo, es suficiente probar que $J(\mathbb{H})$ es cerrado.

Sea $(J(\tau_n))_{n\in\mathbb{N}}$ una sucesión de puntos en $J(\mathbb{H})$ convergente a algún punto $b \in \mathbb{C}$, esto es

$$J(\tau_n) \longrightarrow b \quad \text{si } n \longrightarrow \infty.$$

Veamos que $b \in J(\mathbb{H})$.

Dado $\tau \in \mathbb{H}$, por ser $J(\tau)$ invariante bajo Γ, podemos suponer que todos los τ_n están contenidos en el dominio fundamental F. Distinguimos dos casos.

1. Supongamos que existe una constante $C > 0$ tal que $\Im(\tau_n) \leq C$ para todo $n \in \mathbb{N}$. Entonces el conjunto $\{\tau \in F : \Im(\tau) \leq C\}$ es un compacto. Considerando una subsucesión adecuada se tiene que $(\tau_n)_{n\in\mathbb{N}}$ es convergente, es decir,

$$\tau_n \longrightarrow \tau \in F \subset \mathbb{H}$$

y por la continuidad de la función J, se tiene

$$b = J(\tau) \in J(\mathbb{H}).$$

2. Supongamos que existe una subsucesión de $(\tau_n)_n$ con partes imaginarias convergiendo a ∞. Por (3.6), tenemos que los valores de J de esta subsucesión no estarían acotados, lo que contradice la convergencia de $(J(\tau_n))_n$.

En conclusión, el segundo caso no puede darse y $b \in J(\mathbb{H})$. □

Para finalizar la sección vamos a ver que cada función modular f puede ser expresada como una función racional de J. Para ello, necesitamos el siguiente teorema que además es un resultado de gran interés pues relaciona el peso k de una forma modular con el orden de sus polos.

Sea f una función meromorfa en un abierto de $\mathbb{H}$. Denotaremos por $v_p(f)$ al orden de f en el punto $p \in \mathbb{H}$, es decir, el único entero n tal que $(z-p)^{-n}f(z)$ es holomorfa y no se anula en p. De forma similar, definimos $v_\infty(f)$ como el entero más pequeño tal que $a_n \neq 0$ en el desarrollo de Fourier. Si f es una forma modular de peso k, gracias a que cumple (3.3), el orden tiene sentido en las órbitas de $SL_2(\mathbb{Z})$. Esto quiere decir que puntos de la misma órbita tienen el mismo orden.

Teorema 3.1.4. *Sea $f \neq 0$ una forma modular de peso k, $w = e^{\frac{2\pi i}{3}}$ y F^* el dominio fundamental de $SL_2(\mathbb{Z})$ identificando las línea verticales de los extremos entre ellas y medio arco inferior ($|z| = 1$) con el otro. Entonces*

$$v_\infty(f) + \frac{1}{2}v_i(f) + \frac{1}{3}v_w(f) + \sum_{p\in F^*} v_p(f) = \frac{k}{12}.$$

Demostración. Por el comentario anterior, podemos considerar que todos los representantes de la órbita se encuentran en el dominio fundamental F. Asumamos por simplicidad que no hay ceros ni polos en la frontera de F, excepto los posibles en $i, \rho := \frac{1}{2} + \frac{\sqrt{3}}{2}i$ y $(\rho + 1) := \frac{-1}{2} + \frac{\sqrt{3}}{2}i$. Sea C una constante suficientemente grande de forma que $f(z)$ no tenga ni ceros ni polos en

$\Im(z) \geq C$. Los pequeños arcos alrededor de i, ρ y $\rho + 1$ tienen radio r. El segmento AE tiene parte imaginaria R. Si tuviésemos ceros o polos de f en la frontera de F, el contorno $\mathcal{C}$ que muestra la siguiente figura sería modificado con pequeños arcos alrededor de los polos o ceros.

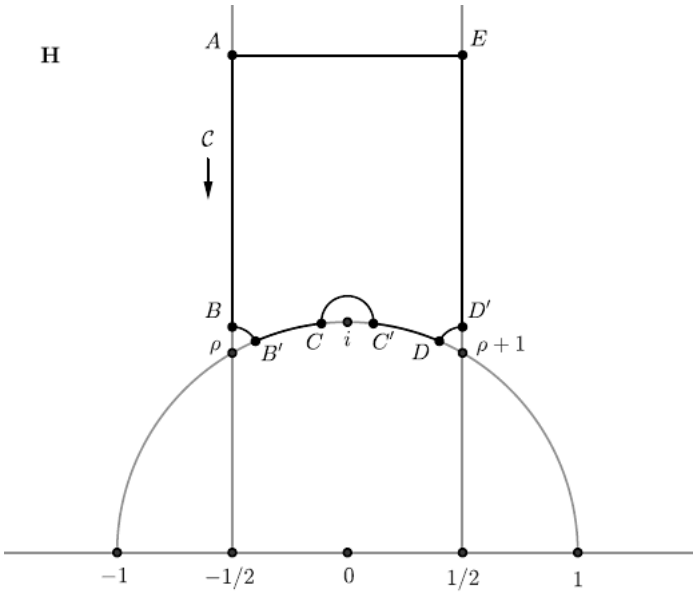

Figura 3.4

Por el principio del argumento,

$$\oint_{\mathcal{C}} \frac{f'(z)}{f(z)} dz = 2\pi i \sum_{p \in F^*} v_p(f). \tag{3.7}$$

Podemos escribir la anterior integral dividiendo el contorno $\mathcal{C}$ en ocho partes, que calcularemos de forma separada.

Primero, las integrales sobre los caminos AB y $D'E$ se cancelan pues

$$\int_{D'}^{E} \frac{f'(z)}{f(z)} dz = \int_{A}^{B} \frac{f'(z+1)}{f(z+1)} dz = -\int_{B}^{A} \frac{f'(z)}{f(z)} dz.$$

Ahora trataremos las integrales sobre los segmentos $B'C$ y $C'D$. Recordemos que la inversión $S(z) = -\dfrac{1}{z}$ satisface

$$f\left(-\frac{1}{z}\right) = z^k f(z).$$

Derivando lo anterior, se tiene que

$$z^{-2}\frac{f'(-1/z)}{f(1/z)} = \frac{k}{z} + \frac{f'(z)}{f(z)}.$$

Recordando que $\frac{d}{dz}(-1/z) = z^{-2}$ y haciendo el cambio de variables $w = -1/z$, obtenemos que

$$\begin{aligned}\int_{C'}^{D} \frac{f'(z)}{f(z)}dz &= \int_{C}^{B'} \frac{f'(-1/w)}{f(-1/w)}w^{-2}dw = \int_{C}^{B'} \left(\frac{k}{w} + \frac{f'(w)}{f(w)}\right)dw \\ &= k\int_{C}^{B'} \frac{1}{z}dz - \int_{C}^{B'} \frac{f'(z)}{f(z)}dz.\end{aligned}$$

Luego,

$$\int_{B'}^{C} \frac{f'(z)}{f(z)}dz + \int_{C'}^{D} \frac{f'(z)}{f(z)}dz \longrightarrow k\frac{\pi i}{6}, \quad \text{si } r \longrightarrow 0,$$

pues el ángulo $\angle C0B'$ tiende a $\frac{\pi}{6}$ si $r \longrightarrow 0$.

Por otro lado, si $r \longrightarrow 0$, tenemos que

$$\begin{aligned}\int_{B}^{B'} \frac{f'(z)}{f(z)}dz &\longrightarrow -\frac{\pi i}{3}v_w(f) \\ \int_{C}^{C'} \frac{f'(z)}{f(z)}dz &\longrightarrow -\pi i v_i(f) \\ \int_{D}^{D'} \frac{f'(z)}{f(z)}dz &\longrightarrow -\frac{\pi i}{3}v_{\rho+1}(f) = -\frac{\pi i}{3}v_\rho(f).\end{aligned}$$

Finalmente, para evaluar la integral del segmento EA, hacemos el cambio de variable $q(z) = e^{2\pi i z}$, por lo que $dq = 2\pi i q dz$. El camino $q(EA)$ recorre una circunferencia de radio $e^{-2\pi R}$ centrado en el origen en el sentido de las agujas del reloj. Por tanto,

$$\begin{aligned}\int_{E}^{A} \frac{f'(z)}{f(z)}dz &= \int_{q(EA)} \frac{f'(q)q'(z)}{f(q)}\frac{dq}{2\pi i q} = \int_{q(EA)} \frac{f'(q)}{f(q)}dq \\ &= -2\pi i v_0(f(q)) = -2\pi i v_\infty(f(z)).\end{aligned}$$

La última igualdad se sigue del desarrollo de Fourier $f = \sum_{n=0}^{+\infty} a_n e^{2\pi i n z} = \sum_{n=0}^{+\infty} a_n q^n$, considerando f en función de z o q respectivamente.

Sumando los ochos caminos se tiene que

$$\oint_{\mathcal{C}} \frac{f'(z)}{f(z)}dz = k\frac{\pi i}{6} - \pi i v_i(f) - \frac{2\pi i}{3}v_w(f) - 2\pi i v_\infty(f),$$

y por (3.7) obtenemos que

$$2\pi i \sum_{p \in F^*} v_p(f) = k\frac{\pi i}{6} - \pi i v_i(f) - \frac{2\pi i}{3} v_w(f) - 2\pi i v_\infty(f).$$

Dividiendo por $2\pi i$ obtenemos el resultado. □

El teorema anterior también se conoce como fórmula $k/12$. Este teorema es de gran relevancia porque nos permite estudiar además la dimensión del espacio vectorial $M_k(\Gamma)$ sobre $\mathbb{C}$. Una fórmula general para la dimensión está disponible en [DS05] (teorema 3.5.1) o en [Shi71] (sección 2.6).

Exponemos ahora una serie de consecuencias de la fórmula $k/12$ que nos serán de ayuda para demostrar que toda función modular J se puede expresar como una función racional de J.

Proposición 3.1.5. *$M_k(\Gamma) = 0$ si k es negativo. Además, cualquier forma modular de peso 0 es constante.*

Demostración. Si $k < 0$ no hay forma de que la suma de los términos no negativos en la izquierda de la fórmula sea igual a $\frac{k}{12}$.

La segunda parte es una aplicación de la fórmula a la función $f(z) - f(i)$. □

Corolario 3.1.6. *Una forma modular de peso $k > 0$ tiene al menos un cero en $\mathbb{H}_\infty$.*

Demostración. Si f no tuviera ningún cero, entonces $\frac{1}{f}$ sería también una forma modular, pero $\frac{1}{f}$ tiene peso negativo.

Si $f \neq 0$ es una forma modular de peso k y si $a \in \mathbb{H}_\infty$ es un cero de f, por la fórmula obtenemos que

$$\frac{k}{12} \geq \frac{v_a(f)}{e(a)} \geq \frac{1}{3},$$

siendo $e(a) = \frac{1}{2}, \frac{1}{3}$ o 1, según a sea i, w o algún $p \in F$ distinto a i y w respectivamente. Notar que hemos extendido la definición de e por $e(i\infty) = 1$. De hecho, de esta última desigualdad deducimos que no existen formas modulares no constantemente nulas de peso $k = 2$. □

Por otro lado, podemos considerar la fórmula $k/12$ como un análogo del teorema 1.1.6 que nos decía que toda función elíptica no constante tiene la misma cantidad de polos que de ceros. En realidad, podemos reformular la fórmula $k/12$ en el caso especial en el que $k = 0$ como sigue.

Teorema 3.1.7. *Si f es una función modular y no idénticamente nula, el número de ceros es el mismo que el número de polos en el dominio fundamental F (ponderados con los factores $\frac{1}{e(a)}$).*

Teorema 3.1.8. *Cada función racional de J es una función modular para Γ. Recíprocamente, cada función modular f para Γ puede ser expresada como una función racional de J.*

Demostración. La primera parte es clara. Para probar el recíproco, sean $z_1, ..., z_n$ los ceros y $p_1, ..., p_n$ los polos de f en el dominio fundamental F. Sea

$$g(\tau) \coloneqq \prod_{k=1}^{n} \frac{J(\tau) - J(z_k)}{J(\tau) - J(p_k)}$$

donde introducimos 1 como factor siempre que z_k o p_k sea ∞. Luego, g es una función racional de J y por tanto modular. Por el teorema anterior, g tiene los mismos ceros y polos que f en el dominio fundamental F. Por tanto, $\frac{f}{g}$ no tiene ceros ni polos, por lo que debe ser constante. En conclusión, f es una función racional de J.

□

Es frecuente definir la función j como

$$j(z) \coloneqq 1728J(z) = 1728\frac{g_2^3(z)}{g_2^3(z) - 27g_3^2(z)},$$

ya que en este caso los coeficientes de la serie de Fourier de la función $j(z)$ resultan ser enteros.[3] El teorema anterior nos dice que toda función modular f es una función racional de J, y por tanto de j. Algunas veces, esta función racional es un polinomio con coeficientes enteros, dándonos una identidad de la forma

$$f(z) = a_1 j(z) + a_2 j^2(z) + ... + a_k j^k(z).$$

Sin embargo, no todas las funciones son invariantes bajo las transformaciones del grupo modular Γ y por tanto no puede expresarse en términos de

[3] Una prueba de este resultado se puede encontrar en la proposición 7.4 (Capítulo 1) de [Sil94].

$j(\tau)$. Ahora bien, dependiendo de la función en cuestión, puede que encontremos que f sí sea invariante bajo transformaciones de cierto subgrupo G de Γ. Por ejemplo, el conjunto de las funciones invariantes bajo el subgrupo $\Gamma_0(N)$ del grupo modular Γ definido como

$$\Gamma_0(N) \coloneqq \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0 \pmod{\text{N}} \right\}$$

jugará un papel importante en la última sección.

3.2 Funciones L

En lo que resta de trabajo vamos a relacionar nuestros tres objetos matemáticos principales: funciones elípticas, curvas elípticas y formas modulares. Esta parte del trabajo será a modo de complemento, por lo que omitiremos algunas pruebas, permitiéndonos cierta relajación. Decir también que esta omisión de demostraciones, en cierta medida, se debe a que sólo soy una entusiasta del tema, quedando el posible desarrollo por completar en próximos años o simplemente, por satisfacción personal.

En esta sección definimos las *funciones L* asociadas a curvas elípticas definidas sobre $\mathbb{Q}$ y a formas modulares, en concreto a formas cúspides. En el teorema de Modularidad veremos que las funciones L son en realidad un puente que nos relaciona ambos objetos. Para escribir esta sección, nos hemos basado en [Loz11].

Digamos que tenemos una secuencia de números $a_1, a_2, a_3, ...$ en $\mathbb{C}$ que nos gustaría estudiar. Una práctica común en teoría de números es considerar la función generadora

$$L(s) = a_1 + \frac{a_2}{2^s} + \frac{a_3}{3^s} + ... = \sum_{n \geq 1} \frac{a_n}{n^s}.$$

La variable s puede tomar cualquier valor complejo siempre y cuando la serie sea convergente.

A la función $L(s)$ definida anteriormente la llamaremos **función *L*** y tal función permite recoger información algebraica que puede ser muy interesante.

Las funciones L surgieron por el trabajo de Dirichlet en su intento y logro por demostrar la infinitud de números primos en progresiones aritméticas.

Por ejemplo, la secuencia constante 1,1,1,... da origen a quizás, la función L más famosa, esta es la función zeta de Riemann,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

que aparece naturalmente al estudiar la distribución de los números primos p a través de la fórmula de Euler

$$\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}}.$$[4]

Para llegar a nuestro objetivo, necesitamos conocer los tipos de singularidades de una curva elíptica.

Sea $\mathbb{K}$ un cuerpo. En general, si $x_0, y_0 \in \mathbb{K}$ son las coordenadas de un punto de una curva $\mathcal{C}$ definida por una ecuación $F(x, y) = 0$, diremos que $\mathcal{C}$ es ***suave*** en (x_0, y_0) si las dos derivadas parciales dF/dx y dF/dy no son nulas en el punto (x_0, y_0). Notemos que esta definición clásica es independiente del cuerpo $\mathbb{K}$, pues la derivada parcial de un polinomio $F(x, y)$ se define por la fórmula usual, que tiene sentido sobre cualquier cuerpo. Si $\mathbb{K}$ es el cuerpo $\mathbb{R}$ de los números reales, esto coincide con la típica condición de que la curva $\mathcal{C}$ tenga una línea tangente en (x_0, y_0). En el caso en el que $F(x, y)$ venga dada por $F(x, y) = y^2 - f(x)$, las derivadas parciales en (x_0, y_0) son $2y_0$ y $-f'(x_0)$. Si suponemos que $\mathbb{K}$ no es un cuerpo de característica 2, las derivadas se anulan a la vez si y sólo si $y_0 = 0$ y x_0 es una raíz múltiple de $f(x)$. Por tanto, la curva $\mathcal{C}$ no es suave en un punto si y sólo si $f(x)$ tiene una raíz múltiple. Ahora bien, dada una curva elíptica $E(\mathbb{K})$, hemos visto que puede ser parametrizada por la ecuación

$$E(\mathbb{K}) : y^2 = 4x^3 - g_2 x - g_3$$

con $g_2(\Lambda)$ y $g_3(\Lambda)$ invariantes de cierto retículo Λ y que $E(\mathbb{K})$ no tiene puntos singulares si y sólo si el discriminante de $f(x) \coloneqq 4x^3 - g_2 x - g_3$ no se anula, lo que equivale a que $f(x)$ no tenga raíces múltiples. Por tanto, en relación con lo anterior, si $P \in E(\mathbb{K})$ es un punto singular de una curva elíptica, sus derivadas parciales evaluadas en P son nulas.

Consideremos ahora una curva elíptica $E(\mathbb{K})$ con ecuación normal de Weierstrass $F(x, y) = 0$, siendo $F(x, y)$ dada por

$$F(x, y) = y^2 + a_1 xy + a_3 - x^3 - a_2 x^2 - a_4 x - a_6.$$

[4]Esta relación la vimos en la asignatura Teoría Analítica de Números.

Si $P = (x_0, y_0)$ es un punto singular de E, por lo anterior, se tiene que

$$\left(\frac{\partial F}{\partial x}\right)_P = 0 = \left(\frac{\partial F}{\partial y}\right)_P.$$

Podemos escribir el desarrollo de Taylor de $F(x, y)$ en el punto (x_0, y_0) como sigue

$$\begin{aligned} F(x,y) - F(x_0,y_0) &= \lambda_1(x-x_0)^2 + \lambda_2(x-x_0)(y-y_0) + \lambda_3(y-y_0)^2 - (x-x_0)^3 \\ &= ((y-y_0) - \alpha(x-x_0))((y-y_0) - \beta(x-x_0)) - (x-x_0)^3 \end{aligned}$$

para ciertos $\lambda_i, \alpha, \beta \in \mathbb{K}$.

Diremos que el punto singular $P \in E$ es un **nodo** si $\alpha \neq \beta$. Si $\alpha = \beta$, diremos que P es una **cúspide**.
Geométricamente, si P es un nodo, tenemos dos rectas tangentes a E en P y si P es una cúspide, sólo tendremos una. Por ejemplo, en la figura 2.2, tenemos una cúspide y un nodo respectivamente.

Antes de definir el concepto de función L asociada a una curva elíptica definida sobre $\mathbb{Q}$, definamos el concepto de reducción módulo p de E siendo p un primo.

Definición 3.5. *Sea E una curva elíptica definida sobre $\mathbb{Q}$ dada en la forma simplificada de Weierstrass $y^2 = x^3 + Ax + B$. Sea $p \geq 2$ un primo y denotemos por $\widetilde{E}$ la reducción de la curva elíptica E módulo p, es decir, $\widetilde{E}$ definida sobre el cuerpo $\mathbb{F}_p = \mathbb{Z}/\mathbb{Z}_p$.*

*Diremos que E tiene **buena reducción módulo p** si $\widetilde{E}$ es una curva suave sobre $\mathbb{F}_p$. Si $\widetilde{E}$ tiene un punto singular $P \in E(\mathbb{F}_p)$ diremos que E tiene mala reducción módulo p y distinguiremos dos casos:*

- *Si $\widetilde{E}$ tiene una cúspide en P, diremos que E tiene una reducción aditiva (o inestable)*
- *Si $\widetilde{E}$ tiene un nodo en P, diremos que E tiene una reducción multiplicativa (o semiestable). Además, si las pendientes de las dos rectas tangentes están en $\mathbb{F}_p$, se dice que la reducción es split multiplicativa.*

Estamos ya preparados para definir la función L de una curva elíptica racional.

Sea E una curva elíptica sobre $\mathbb{Q}$ dada por la ecuación de Weierstrass

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

con los coeficientes $a_i \in \mathbb{Z}$. Para un primo $p \in \mathbb{Z}$ de buena reducción, definimos N_p como el número de puntos de E con coordenadas en $\mathbb{F}_p$. En otras palabras, N_p es el número de puntos de

$$\{\mathcal{O}\} \cup \Big\{(x,y) \in \mathbb{F}_p^2 : y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 \equiv 0 \pmod{\text{p}}\Big\},$$

siendo $\mathcal{O}$ el punto del infinito.

Sea $a_p = p+1-N_p$. Definimos la **parte local en el punto P de la serie L** como

$$L_p(T) = \begin{cases} 1 - a_pT + pT^2 & \text{si } E \text{ tiene buena reducción en } P \\ 1 - T & \text{si } E \text{ tiene split multiplicativa en } P \\ 1 + T & \text{si } E \text{ no tiene split multiplicativa en } P \\ 1 & \text{si } E \text{ tiene reducción aditiva en } P. \end{cases}$$

Definición 3.6. *Sea $E(\mathbb{Q})$ una curva elíptica racional.* ***La función L de E*** *o la función de Hasse-Weil de E se define como*

$$L(E,s) = \prod_{p \geq 2} \frac{1}{L_p(p^{-s})},$$

donde el producto se efectúa sobre los primos $p \geq 2$ y $L_p(T)$ es el factor definido anteriormente.

El matemático Helmut Hasse probó en 1930 que el producto que define $L(E,s)$ converge y proporciona una función analítica en el semiplano $\Re(s) > \dfrac{3}{2}$. No obstante, varios matemáticos conjeturaron que $L(E,s)$ tiene una extensión analítica a todo el plano complejo. Hoy en día esta extensión es un teorema debido a la conjetura de Taniyama-Shimura, actual teorema de Modularidad, que enunciaremos en la próxima sección.

Ahora bien, al principio de la sección dijimos que una función L se define a partir de una secuencia de números $a_1, a_2, a_3, ...$ El siguiente teorema nos muestra que existe una sucesión $\{a_n\}_n$ asociada a $L(E,s)$ y que en efecto se trata de una función L de acuerdo con la definición dada.

Teorema 3.2.1. *Sea $E(\mathbb{Q})$ una curva elíptica y sea $L(E,s)$ su función L. Definamos los coeficientes a_n para $n \geq 1$ como sigue. Si $p \geq 2$ es un primo, definimos*

$$a_p = \begin{cases} p+1-N_p & \textit{si } E \textit{ tiene buena reducción en } P \\ 1 & \textit{si } E \textit{ tiene split multiplicativa en } P \\ -1 & \textit{si } E \textit{ no tiene split multiplicativa en } P \\ 0 & \textit{si } E \textit{ tiene reducción aditiva en } P. \end{cases}$$

Si $n = p^r$ para algún $r \geq 1$, definimos a_{p^r} recursivamente usando la relación

$$a_p a_{p^r} = a_{p^{r+1}} + p a_{p^{r-1}} \quad \textit{si } E \textit{ tiene buena reducción en } P$$

y

$$a_{p^r} = (a_p)^r \quad \textit{si } E \textit{ tiene mala reducción en } P.$$

Finalmente, si $(m,n) = 1$, definimos $a_{mn} = a_m a_n$.

Entonces, la función $L(E,s)$ puede ser escrita como la serie

$$L(E,s) = \sum_{n \geq 0} \frac{a_n}{n^s}.$$

Una demostración del teorema anterior es análoga a la demostración de la igualdad $\zeta(s) = \prod_p \frac{1}{1-p^{-s}}$, donde se usa el teorema fundamental de la aritmética y el desarrollo $\frac{1}{1-x} = \sum_{n \geq 0} x^n$. En este caso, nuestra función es la función de Hasse-Weil $L(E,s)$, por lo que la demostración, no por ello menos complicada, se reduce a calcular el producto de Euler asociado a $L(E,s)$. Otra demostración es posible gracias al teorema de Modularidad. En esta demostración necesitamos los operadores de Hecke y las llamadas autoformas, que en resumen son formas modulares que son autovectores para los operadores de Hecke. Esta parte se puede ver con más detalle en la sección *§2* del capítulo 2 de [Sil94].

Ya tenemos definidas las funciones L asociadas a las curvas elípticas racionales. Queda definir las funciones L de las formas modulares. No obstante, como dijimos al principio de la sección, nos centraremos en las formas cúspides para el subgrupo $\Gamma_0(N)$ de Γ, donde recordemos que $\Gamma_0(N)$ es

$$\Gamma_0(N) \coloneqq \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0 \pmod{\text{N}} \right\}.$$

Definición 3.7. *La **función L de una forma cúspide** $f(z) = \sum_{n \geq 1} a_n q^n \in S_k(\Gamma_0(N))$ se define como*

$$L(f,s) = \sum_{n \geq 1} \frac{a_n}{n^s}.$$

Antes de dar paso a la conjetura de Taniyama-Shimura, necesitamos definir el conductor de una curva elíptica racional.

Para cada primo $p \in \mathbb{Z}$, definimos la cantidad f_p como sigue

$$f_p = \begin{cases} 0 & \text{si } E \text{ tiene buena reducción en } P \\ 1 & \text{si } E \text{ tiene reducción multiplicativa en } P \\ 2+\delta_p & \text{si } E \text{ tiene reducción aditiva en } P \text{ y } p = 2, 3. \end{cases}$$

donde δ_p es una "pequeña" (≤ 6) corrección que vale cero para los primos distintos de 2 y 3. Este invariante técnico se estudia con detalle en *§10* del capítulo IV del libro [Sil94].

Definición 3.8. *Dada una curva elíptica $E(\mathbb{Q})$, definimos el conductor N_E de $E(\mathbb{Q})$ como*

$$N_E = \prod_p p^{f_p}$$

donde el producto es sobre los primos p.

En cierto sentido, el conductor es un número que mide la mala reducción de la curva E. A través del teorema de Modularidad, vamos a ver que se trata de una constante importante que conecta las curvas elípticas con las formas cúspides. De hecho, los primos que dividen al conductor N_E son exactamente los primos que dividen al discriminate de E y los posibles primos a mala reducción son aquellos que dividen al discriminante. Las pruebas de estas observaciones se pueden encontrar en el libro de Álvaro Robledo que comentábamos al principio de la sección.

3.3 La conjetura de Taniyama-Shimura

Existen numerosos ejemplos que muestran las funciones L de ciertas curvas elípticas racionales y las funciones L de ciertas formas cúspides y ambas funciones parecen ser idénticas.[5] Esta sorprendente casualidad estimuló a muchos matemáticos del siglo XX a hallar un resultado que relacionara las curvas elípticas y las formas cúspides de forma general. En un principio, se propuso la siguiente definición.

Definición 3.9. *Diremos que una curva elíptica racional $E(\mathbb{Q})$ es modular si existe una forma cúspide $f(z) \in S_2(\Gamma_0(N))$ tal que*

$$L(E, s) = L(f, s).$$

[5] Un ejemplo se puede encontrar en el libro de Álvaro Lozano.

Por tanto, si E es modular, podemos escribir

$$L(E,s) = L(f,s) = \sum_{n\geq 1} \frac{a_n}{n^s},$$

siendo a_n los coeficientes del desarrollo de Fourier de la cierta forma cúspide f. Ahora bien, los coeficientes a_n deben de coincidir con los valores definidos en el teorema 3.2.1. Luego, dada una curva elíptica, tenemos una candidata clara para una forma cúspide f. No obstante, no es nada sencillo demostrar que esta función candidata es en efecto una forma modular. Es en este punto donde intervienen Taniyama y Shimura, afirmando en la siguiente conjetura, planteada en 1955, que todas las curvas elípticas son modulares.

Conjetura 3.3.1. ***(Teorema de Modularidad).*** *Sea $E(\mathbb{Q})$ una curva elíptica racional de conductor N_E con función L dada por $L(E,s) = \sum_{n\geq 1} \frac{a_E(n)}{n^s}$.*

Entonces los coeficientes a_n son los coeficientes de Fourier de una forma cúspide f_E de peso 2 y nivel N_E, es decir, $f_E(z) \in S_2(\Gamma_0(N_E))$,

$$f_E(z) = a_E(1)q + a_E(2)q^2 + a_E(3)q^3 + ... \in S_2(\Gamma_0(N_E)),$$

siendo $q = e^{2\pi i z}$.

La conjetura se demostró en varias etapas.

En 1993, Andrew Wiles hizo pública una demostración para el caso en el que N_E tiene todos sus factores primos con exponente 1 (caso semi-estable), pero la demostración tenía un error. Richard Taylor asistió a Wiles para lograr superar las dificultades técnicas y en 1995 ambos publicaron una demostración correcta para el caso semi-estable. Alrededor del año 2000, una serie de publicaciones por otros matemáticos lograron demostrar todos los casos restantes.

Como decíamos, la conjetura de Taniyama-Shimura es hoy en día conocida como el teorema de Modularidad, teorema que ha servido como puente para demostrar el Último Teorema de Fermat.

En el margen de una página del libro *Arithmetica de Diofanto*, Fermat escribió en 1637 que ningún cubo es suma de dos cubos, ninguna potencia cuarta es suma de dos potencias cuartas,... y que él había encontrado una demostración *maravillosa* pero que lamentablemente no cabía en el margen de

una página. Esta demostración nunca fue encontrada entre los documentos de Fermat, y el enunciado pasó a la historia como el Último Teorema de Fermat, a la espera de una demostración. En otras palabras, este teorema puede enunciarse de la siguiente manera

> Si n es un número entero mayor o igual que 3, no existen números enteros positivos x, y, z tales que $x^n + y^n = z^n$.

Como comentamos anteriormente, en 1995 Andrew Wiles, junto con la ayuda de Richard Taylor, demuestran en un artículo de 98 páginas el teorema de Modularidad para el caso semi-estable, caso que implica el teorema de Fermat.

Puede ser que nunca sepamos que tenía en mente Fermat, pero al menos la demostración actual no cabe en el margen de una página.

3.4 La conjetura de Swinnerton-Dyer

Finalizamos el trabajo comentando brevemente uno de los siete problemas del milenio: la conjetura de Swinnerton-Dyer. Para ello, recordaremos antes qué significa un punto de un grupo de orden finito y enunciaremos también el teorema de Mordell-Weil.

Dado $P \in E(\mathbb{Q})$, diremos que P tiene orden finito si sumado consigo mismo finitas veces vuelve al punto P. En caso contrario, diremos que P tiene orden infinito.

Por otro lado, el teorema de Mordell-Weil afirma que los puntos racionales de las curvas elípticas sobre $\mathbb{Q}$ no singurales forman un grupo abeliano finitamente generado. La prueba de este resultado se puede encontrar en el capítulo 6 de [Hus87].

Luego, para una curva elíptica $E(\mathbb{Q})$, podemos aplicar el teorema de estructura para grupos finitamente generados (resultado visto en la asignatura Estructuras Algebraicas), obteniendo una descomposición de la forma

$$E(\mathbb{Q}) = \mathbb{Z}^g \times Tors(E(\mathbb{Q}))$$

donde g es un entero llamado el **rango** de $E(\mathbb{Q})$ y $Tors(E(\mathbb{Q}))$ es un grupo abeliano finito formado por todos los elementos de orden finito de $E(\mathbb{Q})$. En otras palabras, el rango de $E(\mathbb{Q})$ es el número de puntos racionales $P \in E(\mathbb{Q})$

con orden infinito.

Hay ejemplos de curvas conocidas con rango de hasta 12. Se desconoce si el rango está acotado o no, aunque encontrar una cota en general es considerada como improbable. Con nuestra comprensión actual de las curvas elípticas, el rango g es un valor misterioso y difícil de calcular para un caso particular. No obstante, Birch y Swinnentor-Dyer encontraron gran evidencia para formular en 1965 la siguiente conjetura acerca del valor de g.

Conjetura 3.4.1. ***Conjetura de Swinnerton-Dyer.*** *El rango g de una curva elíptica E definida sobre $\mathbb{Q}$ es igual al orden del cero de $L_E(s)$ en $s = 1$, siendo $L_E(s)$ la función L de la curva elíptica E.*

Actualmente, esta conjetura no ha podido ser probada ni refutada. En caso de que sea cierta, al igual que ocurrió con la conjetura de Taniyama-Shimura, ¿será necesario desarrollar nuevos conceptos matemáticos?

Bibliografía

[Apo89] Tom M. Apostol. *Modular Functions and Dirichlet Series in Number Theory*. Springer-Verlag, New York, 1989.

[BF09] Rolf Busam y Eberhard Freitag. *Complex analysis*. Springer, 2009.

[DS05] Fred Diamond y Jerry Michael Shurman. *A first course in modular forms*. Vol. 228. Springer, 2005.

[Ful08] William Fulton. "Algebraic curves". En: *An Introduction to Algebraic Geom* 54 (2008).

[Hus87] Dale Husemöller. *Elliptic curves, volume 111*. Vol. 99. Graduate Texts in Mathematics, 1987.

[Kob12] Neal I Koblitz. *Introduction to elliptic curves and modular forms*. Vol. 97. Springer Science & Business Media, 2012.

[Loz11] Álvaro Lozano-Robledo. *Elliptic curves, modular forms, and their L-functions*. Vol. 58. American Mathematical Soc., 2011.

[Mar70] Alekséi Markushévich. *Teoría de las funciones analíticas. Tomo II*. Moscú: MIR, 1970.

[Shi71] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*. Vol. 1. Princeton university press, 1971.

[Sil09] Joseph H Silverman. *The arithmetic of elliptic curves*. Vol. 106. Springer, 2009.

[Sil94] Joseph H Silverman. *Advanced topics in the arithmetic of elliptic curves*. Vol. 151. Springer Science & Business Media, 1994.

[SS03] Elias M. Stein y Rami Shakarchi. *Complex Analysis*. Princeton University Press, 2003.

[WW28] E. T. Whittaker y G. N. Watson. *A Course of Modern Analysis: an Introduction to the General Theory of Infinite Processes and of Analytic Functions; with an Account of the Principal Transcendental Functions.* Cambridge University Press, 1928.