Carlos Ivorra Castillo

CURVAS ELÍPTICAS

La aritmética es una clase de conocimiento en el que las mejores naturalezas deben ser entrenadas, y que no debe ser abandonado.

PLATÓN

Índice General

Introducción

La teoría de las curvas elípticas es una de las creaciones más interesantes de la matemática del siglo XX, si bien sus antecedentes se remontan hasta la matemática griega. Con la teoría que vamos a desarrollar en este libro podremos tratar problemas como éste (resuelto por Mordell en 1962):

Problema 1 *Demostrar que los únicos números naturales no nulos que pueden expresarse simultáneamente como producto de dos y tres números consecutivos son*

$$6 = 2 \cdot 3 = 1 \cdot 2 \cdot 3 \qquad y \qquad 210 = 14 \cdot 15 = 5 \cdot 6 \cdot 7.$$

Esto equivale a encontrar las soluciones enteras de la ecuación

$$Y(Y+1) = (X-1)X(X+1).$$

El problema puede ser abordado mediante técnicas de la teoría algebraica de números, es decir, utilizando la factorización real o ideal de los anillos de enteros algebraicos de los cuerpos numéricos. Sin embargo, nosotros lo trataremos desde el punto de vista de la geometría algebraica. La ecuación anterior determina una curva proyectiva regular de género 1, y la cuestión es, pues, encontrar los puntos con coordenadas enteras de una curva algebraica dada.

Puntos racionales y enteros En realidad la teoría que vamos a desarrollar se centra principalmente en la búsqueda de puntos con coordenadas racionales,[1] si bien en muchos casos y de forma más o menos indirecta nos permitirá ocuparnos de las soluciones enteras. Sucede que la existencia de puntos enteros o racionales en una curva depende crucialmente de su género. Podemos distinguir tres casos:

- Una curva de género $g = 0$ no tiene puntos racionales o bien tiene infinitos. Sin embargo, puede no tener puntos enteros, tener una cantidad finita de ellos o tener infinitos.
- Una curva de género $g = 1$ no tiene puntos racionales, tiene un número finito de ellos o bien tiene infinitos, pero sólo puede tener una cantidad finita de puntos enteros.

[1] Por soluciones racionales entenderemos soluciones en un cuerpo arbitrario k prefijado, no necesariamente $\mathbb{Q}$.

- Una curva de género $g \geq 2$ sólo puede tener una cantidad finita de puntos racionales.

Veamos un ejemplo que ilustra la situación en género 0:

Problema 2 *Encontrar todas las ternas pitagóricas, es decir, las ternas* (a, b, c) *de números naturales tales que* $a^2 + b^2 = c^2$.

SOLUCIÓN: Ante todo es fácil ver que si dos componentes de una terna pitagórica tienen un divisor primo común p, lo mismo le sucede al tercero, y al eliminarlo de los tres obtenemos otra terna pitagórica. Diremos que una terna es *primitiva* si sus componentes son primas entre sí (en cuyo caso lo son dos a dos). Toda terna pitagórica es múltiplo de una única terna primitiva, luego basta determinar las ternas primitivas.

Más aún, si a y b fueran impares, entonces $c^2 \equiv 2 \pmod 4$, lo cual es imposible, luego una de las dos primeras componentes ha de ser par, y no perdemos generalidad si suponemos que lo es la primera.

Si (a, b, c) es una terna pitagórica, entonces $(a/c, b/c)$ es un punto racional de la cónica $X^2 + Y^2 = 1$. Recíprocamente, todo punto racional de esta cónica determina una terna pitagórica (admitiendo temporalmente ternas con componentes negativas).

Tomamos uno de estos puntos, por ejemplo $(0, 1)$, y consideramos la proyección estereográfica que a cada punto $(t, 0)$ del eje X le asigna el segundo punto donde la recta que pasa por $(0, 1)$ y $(t, 0)$ corta a la cónica. La recta es $X = (1-Y)t$, y el punto de corte con la cónica ha de cumplir $(1-Y)^2t^2+Y^2 = 1$ o, equivalentemente,

$$(1 + t^2)Y^2 - 2t^2Y + t^2 - 1 = 0$$

Una raíz es $Y = 1$, pero buscamos la otra. Si dividimos entre el coeficiente director, el término independiente será el producto de las dos raíces, luego la otra es $(t^2 - 1)/(t^2 + 1)$. De la ecuación de la recta obtenemos el valor de X, con lo que el punto t de la recta se corresponde con el punto

$$\phi(t) = \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right).$$

Es fácil calcular la correspondencia inversa y comprobar que también está definida por funciones racionales en X, Y, de modo que ϕ biyecta los números racionales t con los puntos racionales de la cónica. (El punto $(0, 1)$ se corresponde con el punto infinito de la recta proyectiva.) Si $t = u/v$, con u, $v \in \mathbb{Z}$, primos entre sí, $v \neq 0$, entonces

$$\phi(t) = \left(\frac{2uv}{u^2+v^2}, \frac{u^2-v^2}{u^2+v^2}\right),$$

que se corresponde con la terna pitagórica $(2uv, u^2 - v^2, u^2 + v^2)$. Notemos que si eliminamos la restricción $v \neq 0$ (con lo que estamos admitiendo $t = \infty$)

recuperamos el punto $(0,1)$ que habíamos perdido (o, más exactamente, las ternas triviales asociadas a él).

Vamos a probar que las ternas pitagóricas primitivas (con primera componente par) son exactamente las de la forma

$$(2uv, u^2 - v^2, u^2 + v^2),$$

donde $0 < v < u$ son números naturales primos entre sí de paridades opuestas. Es inmediato comprobar que estas condiciones hacen que la terna sea primitiva. Recíprocamente, si (a,b,c) es una terna primitiva, hemos demostrado que existen enteros (u,v) primos entre sí tales que

$$\frac{a}{c} = \frac{2uv}{u^2+v^2}, \quad \frac{b}{c} = \frac{u^2-v^2}{u^2+v^2}.$$

Obviamente ha de ser $0 < v < u$ (si partimos de una terna de números naturales no nulos). Si u y v fueran ambos impares podríamos simplificar el 2 de la primera fracción y concluiríamos que a es impar, en contra de lo supuesto. ∎

El procedimiento empleado es general: Si una cónica dada por una ecuación con coeficientes en un cuerpo k tiene un punto racional (o sea, con coordenadas en k), entonces tiene infinitos puntos racionales, parametrizables en $\mathrm{P}^1(k)$ mediante la proyección estereográfica.

Más en general aún, veremos que toda curva proyectiva de género 0 definida mediante ecuaciones con coeficientes en k es birracionalmente equivalente a una cónica con coeficientes en k a través de una aplicación birracional definida mediante polinomios con coeficientes en k, de modo que los puntos racionales de la curva dada se corresponden biunívocamente (salvo quizá un número finito de excepciones) con los de la cónica. Si la curva es regular no hay excepciones. Además, toda cónica admite una ecuación homogénea de la forma

$$aX^2 + bY^2 + cZ^2 = 0, \qquad a,b,c \in k.$$

(La ecuación de una cónica en coordenadas homogéneas es una forma cuadrática, y toda forma cuadrática es diagonalizable.)

Dicho esto, no volveremos a ocuparnos de las curvas de género 0. En general, los problemas concernientes a estas curvas se tratan más convenientemente con técnicas de la teoría algebraica de números. El objeto de este libro serán las curvas de género $g = 1$. Una *curva elíptica* sobre un cuerpo k es una curva proyectiva regular de género 1 definida por ecuaciones con coeficientes en k y que tiene al menos un punto racional. En el problema 1 hemos visto un ejemplo de curva elíptica. No es casual que venga dada por una ecuación cúbica. Veremos que toda curva elíptica es isomorfa a una cúbica (plana) regular.

El ejemplo de Selmer Toda cúbica proyectiva regular tiene género $g = 1$. Supuesto que esté definida mediante una ecuación con coeficientes en $\mathbb{Q}$, para

que sea una curva elíptica sobre $\mathbb{Q}$ todavía falta que cumpla una condición adicional, a saber, que tenga un punto racional, y esto dista mucho de ser trivial.

Para curvas de género $g = 0$, el problema puede reducirse a determinar si una forma cuadrática de tipo

$$aX^2 + bY^2 + cZ^2 = 0, \qquad a, b, c \in \mathbb{Q}.$$

tiene una solución no trivial en $\mathbb{Q}$. El teorema de Hasse-Minkowski afirma que esto es equivalente a que tenga solución en $\mathbb{R}$ y en todos los cuerpos p-ádicos $\mathbb{Q}_p$, y a su vez esto puede reducirse a un número finito de comprobaciones en términos de congruencias.

Sin embargo, el teorema de Hasse-Minkowski no es válido para curvas de género 1. Un ejemplo clásico se debe a Selmer, quien demostró que la ecuación

$$3U^3 + 4V^3 + 5W^3 = 0$$

tiene soluciones no triviales en $\mathbb{R}$ y en todos los cuerpos $\mathbb{Q}_p$, pero no tiene soluciones en $\mathbb{Q}$. El estudio local de las curvas elípticas no deja por ello de ser una herramienta valiosa, pero ya no es definitiva, lo cual hace que algunos aspectos de la teoría sean mucho más complejos que los equivalentes en el caso de curvas de género 0. ∎

Veamos otro problema que nos lleva a una familia de curvas elípticas:

Problema 3 Un número natural es *congruente* si es el área de un triángulo rectángulo de lados racionales. *Encontrar un método para decidir si un número dado es o no congruente.*

Observemos en primer lugar que es muy fácil determinar si un número natural n es o no el área de un triángulo rectángulo de lados *enteros.* Dichos lados formarán una terna pitagórica $d(2uv, u^2 - v^2, u^2 + v^2)$, luego

$$n = d^2uv(u^2 - v^2).$$

Teniendo en cuenta que u, v y d han de ser divisores de n, siempre es posible decidir si existen en un número finito de casos. Por ejemplo, es fácil ver que el menor natural que cumple esto es 6, el área del triángulo de lados $(3, 4, 5)$.

El problema ya no es trivial cuando admitimos lados racionales. Observemos que un número es congruente si y sólo si lo es su parte libre de cuadrados, luego basta ocuparse de números libres de cuadrados. Notemos también que no ganaríamos en generalidad si buscáramos números racionales congruentes, pues u/v es el área de un triángulo rectángulo racional si y sólo si lo es uv.

No es cierto que 6 sea el menor número congruente, pues en 1225 Fibonacci descubrió que 5 también lo es. El triángulo correspondiente tiene lados

$$\left(\frac{3}{2}, \frac{20}{3}, \frac{41}{6}\right).$$

Se trata del menor posible, pero esto no es trivial. Tuvieron que pasar cuatro siglos hasta que Fermat demostrara que los números 1, 2 y 3 no son congruentes (luego 4 tampoco). La relación con las curvas elípticas viene dada por el teorema siguiente, que todavía no estamos en condiciones de probar completamente:

Teorema 1 *Si n es un número natural libre de cuadrados, las condiciones siguientes son equivalentes:*

a) *n es congruente.*

b) *Existen tres cuadrados racionales en progresión aritmética de razón n.*

c) *La curva $Y^2 = X^3 - n^2X$ tiene un punto racional (finito) (x, y) distinto de $(-n, 0)$, $(0, 0)$ y $(n, 0)$.*

DEMOSTRACIÓN: De momento probaremos que $a) \Leftrightarrow b) \Rightarrow c)$. La implicación que falta está en la página 51.

$a) \Rightarrow b)$ Si $n = ab/2$, para cierta terna pitagórica (a, b, c), entonces tomemos $x = c^2/4$, de modo que $(a-b)^2/4 = x-n$ y $(a+b)^2/4 = x+n$, luego los números $x - n$, x, $x + n$ forman una progresión aritmética de cuadrados racionales.

$b) \Rightarrow a)$ Si la progresión es $x - n$, x, $x + n$, definimos

$$a = \sqrt{x+n} + \sqrt{x-n}, \quad b = \sqrt{x+n} - \sqrt{x-n}, \quad c = 2\sqrt{x},$$

de modo que a, b, $c \in \mathbb{Q}$, $a^2 + b^2 = c^2$ y $n = ab/2$.

$b) \Rightarrow c)$ Si la progresión es $x - n$, x, $x + n$, entonces su producto es de la forma y^2, para un cierto $y \in \mathbb{Q}$. Así pues, $y^2 = x^3 - n^2x^2$. No puede ser $x = 0$ o $x = \pm n$ porque n es libre de cuadrados. ∎

Aún no estamos en condiciones de aprovechar este teorema para obtener resultados concretos, pero veamos un ejemplo relacionado.

Ejemplo *La ecuación $U^4 + V^4 = W^2$.*

Fermat demostró que esta ecuación no tiene soluciones enteras no triviales (trivial quiere decir con una componente nula). Su argumento es elemental y se basa en aplicar varias veces la fórmula para las ternas pitagóricas. No vamos a verlo aquí. En cambio veremos que el problema se puede reformular en términos de curvas elípticas.

En primer lugar, una solución no trivial cumple $(U/V)^4 + 1 = (W/V^2)^2$. Es inmediato comprobar que la ecuación dada no tiene soluciones enteras no triviales si y sólo si la ecuación $Y^4 + 1 = Z^2$ no tiene soluciones racionales distintas de $(0, \pm 1)$.

Para convertir esta ecuación en una cúbica basta hacer el cambio de variable $Z = Z' + Y^2$, con lo que obtenemos $Z^2 + 2ZY^2 = 1$. También es inmediato que las soluciones racionales de la ecuación anterior se corresponden biunívocamente con las de ésta, y las soluciones triviales siguen siendo $(Y, Z) = (0, \pm 1)$.

Ya tenemos una curva elíptica,[2] pero veremos que toda curva elíptica admite un tipo de ecuación canónica llamada *forma de Weierstrass,* que no es sino una ecuación de la forma

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6,$$

y la ecuación que hemos obtenido no es de esa forma, aunque estamos muy cerca. Para obtener una ecuación de Weierstrass consideramos la clausura proyectiva, cuya ecuación homogénea es

$$XZ^2 + 2ZY^2 = X^3.$$

Los puntos triviales son ahora $(X, Y, Z) = (1, 0, \pm 1)$, a los que hemos de añadir los dos puntos infinitos $(0, 1, 0)$ y $(0, 0, 1)$, pues éstos tampoco proporcionan soluciones a la ecuación de Fermat.

Ahora consideramos la curva afín que resulta de tomar como recta del infinito a $Z = 0$. Es la curva

$$2Y^2 = X^3 - X.$$

Los puntos triviales son ahora $(X, Y) = (\pm 1, 0)$ y $(0, 0)$ (el cuarto ha quedado en el infinito). Para tener una ecuación de Weierstrass basta cambiar el 2 de sitio. Multiplicando por 8 tenemos

$$(4Y)^2 = (2X)^3 - 4(2X),$$

y un cambio de variables obvio transforma la ecuación en:

$$Y^2 = X^3 - 4X.$$

Los puntos triviales son ahora $(\pm 2, 0)$ y $(0, 0)$. Hemos probado que la ecuación $U^4 + V^4 = W^2$ no tiene soluciones enteras no triviales si y sólo si la curva $Y^2 = X^3 - 4X$ no tiene puntos racionales no triviales (donde el sentido de "no trivial" es distinto en cada caso).

Si aceptamos que esto es así —ya hemos dicho que el argumento para la ecuación de Fermat es elemental— tenemos probado que 2 no es un número congruente. ∎

Consideremos ahora la curva elíptica $Y^2 = X^3 - 25X$, relacionada con la congruencia del número 5. Si particularizamos la prueba del teorema 1 al triángulo de Fibonacci obtenemos el punto racional

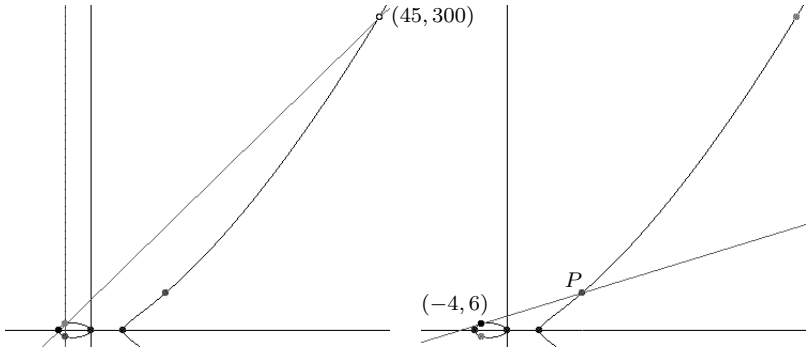
$$\left(\frac{20.172}{1.728}, \frac{62.279}{1.728}\right).$$

Hay soluciones no triviales más simples, como $(-4, \pm 6)$. Fermat descubrió una forma sencilla de obtener nuevos puntos racionales de una cúbica a partir de otros dados. Es fácil ver que si trazamos la secante a la curva por dos puntos

[2] Habría que comprobar que ciertamente lo es, pero esto será evidente en cuanto dispongamos de la teoría elemental.

racionales, la recta cortará a la curva en un tercer punto racional. Aquí hay que considerar que una recta pasa dos veces por cada punto de la curva donde es tangente (o a veces incluso tres veces, según el sentido usual de orden de intersección definido en geometría algebraica).

Por ejemplo, si unimos el punto $(-4, 6)$ con el punto trivial $(-5, 0)$ obtenemos un nuevo punto con coordenadas enteras, a saber, el punto $(45, 300)$. Si trazamos la tangente por $(-4, 6)$, obtenemos el punto P correspondiente al triángulo de Fibonacci:



Este procedimiento geométrico se reduce en la práctica a unas sencillas fórmulas algebraicas explícitas que veremos en su momento. En realidad Diofanto ya lo había utilizado algebraicamente sin darse cuenta de su interpretación geométrica, y a su vez Fermat no podía advertir el profundo significado del proceso. En efecto, veremos que mediante el trazado de secantes y tangentes se puede definir una operación de grupo abeliano sobre toda curva elíptica, de modo que éstas resultan ser los ejemplos más simples de *variedades abelianas* (variedades proyectivas con una estructura de grupo definida por aplicaciones regulares).

Por ejemplo, uno de los resultados fundamentales de la teoría de curvas elípticas (generalizable a variedades abelianas) es el teorema de Mordell-Weil, según el cual, los puntos racionales de una curva elíptica definida sobre un cuerpo numérico forman un grupo abeliano finitamente generado. Esto se traduce en que todos los puntos racionales de la curva pueden obtenerse a partir de un número finito de ellos mediante la construcción sucesiva de secantes y tangentes. Desde un punto de vista algebraico, el teorema de Mordell-Weil implica que el conjunto de puntos racionales de una curva elíptica definida sobre un cuerpo numérico se descompone en suma directa del subgrupo formado por los elementos de torsión (los elementos de orden finito) y un grupo abeliano libre. La solución de los problemas que hemos comentado depende en gran medida de la determinación de esta descomposición. Veremos que es fácil determinar el subgrupo de torsión de una curva elíptica, mientras que el problema de calcular, no ya unos generadores concretos, sino simplemente el rango de la parte libre, es complicado, hasta el punto de que no se conoce ningún método general para resolverlo. La solución de los problemas que hemos comentado depende en gran medida de la posibilidad de calcular este rango en algunos casos concretos.

Supondremos al lector familiarizado con la geometría algebraica y la teoría algebraica de números. No obstante, en el primer capítulo revisaremos los conceptos y los resultados más importantes que vamos a necesitar de geometría algebraica, poniendo especial énfasis (con demostraciones completas) en las consecuencias que podemos extraer del hecho de que una variedad algebraica pueda ser definida mediante ecuaciones en un cuerpo específico (no necesariamente algebraicamente cerrado). Algunos de estos resultados requerirán algunos conocimientos de cohomología de grupos. Estos hechos son esenciales en la prueba de la hipótesis de Riemann para curvas algebraicas, que incluimos en un apéndice porque, siendo un resultado de gran interés en sí mismo, no nos va a ser necesario en todo el libro (el caso particular correspondiente a curvas elípticas admite una demostración mucho más simple que veremos en el capítulo IV).

Los dos capítulos siguientes desarrollan la teoría básica sobre curvas elípticas, y a continuación estudiamos las curvas elípticas definidas sobre cuerpos finitos y cuerpos locales, como requisitos necesarios para abordar las curvas definidas sobre cuerpos numéricos, y en particular el teorema de Mordell-Weil. En el capítulo X nos centramos en las curvas elípticas definidas sobre $\mathbb{C}$, lo que nos conduce de forma natural al estudio de las llamadas funciones modulares, que son unas familias de funciones holomorfas que podrían ser estudiadas de forma independiente, pero que están íntimamente relacionadas con las curvas elípticas. De hecho, los resultados más profundos de la teoría que vamos a exponer (resultados que quedan completamente fuera del alcance de este libro y que conducen, entre otras cosas, a la demostración del Último Teorema de Fermat) consisten en mostrar profundas relaciones entre las curvas elípticas y las funciones modulares.

Finalmente, el último capítulo está dedicado a mostrar una conexión clásica entre las curvas elípticas, las funciones modulares y la teoría de cuerpos de clases.

Capítulo I

Preliminares de geometría algebraica

Suponemos que el lector está familiarizado con la geometría algebraica básica. De todos modos, en este primer capítulo recordaremos los conceptos y resultados más importantes que vamos a necesitar. Al mismo tiempo aprovecharemos para observar que, si bien la geometría algebraica requiere trabajar con cuerpos de constantes algebraicamente cerrados, lo cierto es que si consideramos variedades y funciones definidas por polinomios con coeficientes en un cuerpo menor, esta circunstancia se conserva a través de las construcciones usuales.

A lo largo de todo este capítulo k será un cuerpo perfecto y $\overline{k}$ será una clausura algebraica prefijada. De este modo, la extensión $\overline{k}/k$ es de Galois, y los elementos de k son los fijados por el grupo de k-automorfismos $G(\overline{k}/k)$.

1.1 Variedades afines

El espacio afín n-dimensional sobre un cuerpo k es $A^n(k) = k^n$. Entonces tenemos que $A^n(k) \subset A^n(\overline{k})$. Escribiremos A^n en lugar de $A^n(\overline{k})$. A los puntos de $A^n(k)$ los llamaremos *puntos racionales* de A^n.

Un sistema de referencia afín en $A^n(\overline{k})$ es una $n+1$-tupla $(O, P_1, \ldots, P_n)$ tal que los vectores $\overrightarrow{OP_i}$ sean linealmente independientes sobre $\overline{k}$. Diremos que *está definido sobre* k si los puntos O y P_i son racionales. Los puntos racionales de A^n se caracterizan como los puntos que tienen coordenadas en k^n respecto a cualquier sistema de referencia prefijado definido sobre k.

Siempre que hablemos de un sistema de referencia en A^n se entenderá que está definido sobre k.

Si $S \subset \overline{k}[X_1, \ldots, X_n]$, llamaremos $V(S)$ al conjunto de todos los puntos de A^n cuyas coordenadas en un sistema de referencia prefijado anulan a todos los polinomios de S.

Se dice que un conjunto $C \subset A^n$ es *algebraico* si $C = V(S)$, para cierto conjunto de polinomios $S \subset \overline{k}[X_1, \ldots, X_n]$ (y cierto sistema de referencia). Si $S \subset k[X_1, \ldots, X_n]$ diremos que C *está definido sobre* k.

El carácter algebraico de un conjunto (y el hecho de que esté o no definido sobre k) no depende del sistema de referencia considerado, si bien el conjunto S que lo define variará de un sistema a otro.

Si C es un conjunto algebraico definido sobre k, llamaremos $C(k)$ al conjunto de todos los puntos racionales de C, es decir, $C(k) = C \cap A^n(k)$. Conviene tener presente que $C(k)$ puede ser vacío, por lo que en ningún momento podremos apoyarnos en $C(k)$ para definir conceptos asociados a C. (Basta pensar en $C = V(X^2 + Y^2 + 1)$ con $k = \mathbb{Q}$).

El grupo de Galois $G(\overline{k}/k)$ actúa de forma natural (componente a componente) sobre A^n y, como deja invariantes a los puntos de cualquier sistema de referencia (definido sobre k), es claro que para cada $\sigma \in G(\overline{k}/k)$ y cada $P \in A^n$, las coordenadas de P^σ son las imágenes por σ de las coordenadas de P. En particular, los puntos racionales de A^n son los fijados por todos los automorfismos de $G(\overline{k}/k)$. De aquí obtenemos una caracterización de los conjuntos algebraicos definidos sobre k:

Teorema 1.1 *Un conjunto algebraico $C \subset A^n$ está definido sobre k si y sólo si $\sigma[C] = C$ para todo $\sigma \in G(\overline{k}/k)$.*

Demostración: La condición es claramente necesaria: si $C = V(S)$ con $S \subset k[X_1, \ldots, X_n]$, $P \in C$, $F \in S$ y $\sigma \in G(\overline{k}/k)$, entonces $F(P) = 0$, luego $F(P^\sigma) = 0$, luego $P^\sigma \in C$.

Recíprocamente, sea $C = V(S)$, con $S \subset \overline{k}[X_1, \ldots, X_n]$. La hipótesis hace que si $F \in S$ y $\sigma \in G(\overline{k}/k)$, entonces $F^\sigma(P) = 0$ para todo $P \in C$. En efecto, tenemos que $P^{\sigma^{-1}} \in C$, luego $F(P^{\sigma^{-1}}) = 0$, luego $F^\sigma(P) = 0$. Esto implica que podamos suponer que S es unión de clases de conjugación respecto a la acción de $G(\overline{k}/k)$.

Supongamos que $F_1, \ldots, F_r$ es una de estas clases de conjugación y sean $e_1, \ldots, e_r$ los polinomios simétricos elementales con r indeterminadas. Vamos a probar que $V(S)$ no se altera si sustituimos $F_1, \ldots, F_r$ por $e_i(F_1, \ldots, F_r)$, para $i = 1, \ldots, r$, con lo que el teorema quedará probado, pues estos polinomios tienen sus coeficientes en k. Concretamente, hemos de probar que un punto P es raíz de los polinomios F_i si y sólo si lo es de los $e_i(F_1, \ldots, F_r)$. Una implicación es obvia. Si P es raíz de los $e_i(F_1, \ldots, F_r)$, en particular lo es de

$$e_r(F_1, \ldots, F_r) = F_1 \cdots F_r.$$

Esto significa que $F_i(P) = 0$ para un i. No perdemos generalidad si suponemos $F_1(P) = 0$. Teniendo en cuenta esto y que P es raíz de $e_{m-1}(F_1, \ldots, F_r)$, concluimos que también es raíz de $F_2 \cdots F_r$, luego $F_i(P) = 0$ para $i \geq 2$. Podemos suponer $F_2(P) = 0$. Prosiguiendo de este modo concluimos que P anula a todos los F_i. ∎

Así pues, si C está definido sobre k el grupo $G(\overline{k}/k)$ actúa sobre C. También es claro que

$$C(k) = \{P \in C \mid P^\sigma = P \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Si C es un conjunto algebraico, llamaremos $I(C)$ al ideal de todos los polinomios de $\overline{k}[X_1, \ldots, X_n]$ que se anulan sobre (las coordenadas de) los puntos de C. Si C está definido sobre k definimos además

$$I(C/k) = I(C) \cap k[X_1, \ldots, X_n],$$

que es un ideal de $k[X_1, \ldots, X_n]$.

Una *variedad afín* es un conjunto algebraico $V \subset A^n$ tal que $I(V)$ es un ideal primo. Si V está definida sobre k, es claro que $I(V/k)$ también es un ideal primo, pero del hecho de que $I(V/k)$ sea primo no podemos inferir que V sea una variedad. Basta pensar en $V(X^2 + Y^2)$ con $k = \mathbb{Q}$.

Una aplicación $\phi : V \longrightarrow W$ entre dos variedades afines definidas sobre k es *polinómica* (sobre k) si, fijados sistemas de referencia afines definidos sobre k, existen $F_1, \ldots, F_n \in k[X_1, \ldots, X_m]$ tales que para todo $P \in V$ se cumple que $\phi(P) = (F_1(P), \ldots, F_n(P))$. (Aquí $\phi(P)$ representa en realidad al vector de coordenadas de $\phi(P)$.)

Un *isomorfismo* es una aplicación biyectiva tal que tanto ella como su inversa son polinómicas.

Es claro que la composición de aplicaciones polinómicas es una aplicación polinómica y que las aplicaciones polinómicas transforman puntos racionales en puntos racionales.

Sea V una variedad afín definida sobre k. Llamaremos $k[V]$ al conjunto de las funciones polinómicas (definidas sobre k) $V \longrightarrow \overline{k}$. Aquí estamos considerando a $\overline{k} = A^1(\overline{k})$ como una variedad. Es claro que $k[V]$ es un anillo con las operaciones definidas puntualmente (y aquí consideramos a $\overline{k}$ como cuerpo y no meramente como espacio afín). Además contiene una copia de k (formada por las funciones constantes). Notemos que haciendo $k = \overline{k}$ tenemos definido el anillo $\overline{k}[V]$ como caso particular. Evidentemente $k[V] \subset \overline{k}[V]$.

Fijado un sistema de referencia, cada polinomio $F \in k[X_1, \ldots, X_m]$ define una función polinómica $f \in k[V]$ dada por $f(P) = F(P)$ (entendiendo que el segundo miembro es F actuando sobre las coordenadas de P). La aplicación $F \mapsto f$ es un epimorfismo de anillos y su núcleo es $I(V/k)$. Por consiguiente tenemos la representación

$$k[V] \cong k[X_1, \ldots, X_n]/I(V/k). \tag{1.1}$$

Notemos que lo mismo vale en particular si cambiamos k por $\overline{k}$ e $I(V/k)$ por $I(V)$. El monomorfismo natural

$$k[X_1, \ldots, X_n]/I(V/k) \longrightarrow \overline{k}[X_1, \ldots, X_n]/I(V)$$

se corresponde a través de los isomorfismos (1.1) con la inclusión $k[V] \subset \overline{k}[V]$.

Representaremos por x_i a la clase de X_i en $k[X_1,\dots,X_n]/I(V/k)$. Claramente tenemos que $k[V] = k[x_1,\dots,x_n]$. Como $I(V/k)$ es un ideal primo, la representación (1.1) muestra que $k[V]$ es un dominio íntegro, luego podemos considerar su cuerpo de cocientes, al que llamaremos $k(V)$. Tenemos la inclusión $k(V) \subset \overline{k}(V)$.

Más aún, como ambos cuerpos están generados por las clases $x_1,\dots,x_n$, de hecho $\overline{k}(V) = \overline{k}k(V)$. Esto implica que la extensión $\overline{k}(V)/k(V)$ es de Galois.

Se dice que $\alpha \in k(V)$ es *regular* en $P \in V$ si $\alpha = f/g$ con[1] $f,\ g \in \overline{k}[V]$, $g(P) \neq 0$, y en tal caso definimos $\alpha(P) = f(P)/g(P)$. Así, puede haber representaciones de α para las que el denominador se anule y otras para las que no se anule. No obstante, si α es regular en P, el valor $\alpha(P)$ no depende de la representación con la que se calcula. Si α no es regular en P se dice que es *singular*.

Podemos identificar los elementos de $k(V)$ con las funciones que determinan, en el sentido de que si f/g y f'/g' coinciden sobre el conjunto de puntos regulares para ambas, entonces $f/g = f'/g'$.

En efecto, fijado un sistema de referencia, sea $f = [F]$, $g = [G]$, $f' = [F']$ y $g' = [G']$. Sea $H = FG' - GF'$. Entonces tenemos que $HGG' \in I(V/k)$, pero $G,\ G' \notin I(V/k)$, pues $g,\ g' \neq 0$. Como $I(V/k)$ es primo, ha de ser $H \in I(V/k)$, luego $fg' - gf' = 0$, con lo que $f/g = f'/g'$.

En vista de esto, los elementos de $k(V)$ se llaman *funciones racionales* de V.

Es claro que el grupo de Galois $G(\overline{k}/k)$ actúa de forma natural sobre el anillo de polinomios $\overline{k}[X_1,\dots,X_n]$ (coeficiente a coeficiente), de modo que los polinomios invariantes son precisamente los de $k[X_1,\dots,X_n]$. Observemos que si $F \in k[X_1,\dots,X_n]$, $P \in A^n$ y $\sigma \in G(\overline{k}/k)$, entonces

$$F^\sigma(P) = F(P^{\sigma^{-1}})^\sigma.$$

Si V es una variedad definida sobre k, entonces $\sigma[I(V)] = I(V)$ para todo $\sigma \in G(\overline{k}/k)$, pues si $F \in I(V)$ y $P \in V$, entonces $P^{\sigma^{-1}} \in V$, luego $F(P^{\sigma^{-1}}) = 0$, luego $F^\sigma(P) = 0$, con lo que $F^\sigma \in I(V)$.

Esto nos permite extender cada automorfismo $\sigma \in G(\overline{k}/k)$ a un automorfismo $\sigma : \overline{k}[V] \longrightarrow \overline{k}[V]$ (dado por $[F]^\sigma = [F^\sigma]$) y éste a su vez a un automorfismo $\sigma \in G(\overline{k}(V)/k(V))$. Explícitamente, si $\alpha \in \overline{k}(V)$, $P \in V$ y $\sigma \in G(\overline{k}/k)$, entonces

$$\alpha^\sigma(P) = \alpha(P^{\sigma^{-1}})^\sigma.$$

(Entendiendo que α^σ es regular en P si y sólo si α es regular en $P^{\sigma^{-1}}$.)

Tenemos así un isomorfismo de grupos $G(\overline{k}/k) \cong G(\overline{k}(V)/k(V))$. En efecto, cada $k(V)$-automorfismo de $\overline{k}(V)$ se restringe a un $\overline{k}$-automorfismo de k (porque $\overline{k}/k$ es normal) y está completamente determinado por dicha restricción, dado

[1] Notemos que no exigimos que f y g estén en $k[X_1,\dots,X_n]$. De este modo α es regular en los mismos puntos vista como elemento de $\overline{k}(V)$ o de $k(V)$.

que $\overline{k}(V) = \overline{k}k(V)$. Así pues, cada $k(V)$-isomorfismo de $\overline{k}(V)$ es la (única) imagen de su restricción a $\overline{k}$. En particular vemos que

$$k(V) = \{\alpha \in \overline{k}(V) \mid \alpha^\sigma = \alpha \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

También podemos concluir que k es algebraicamente cerrado en $k(V)$, pues si $\alpha \in k(V)$ es algebraico sobre k, entonces $\alpha \in \overline{k}$ y queda fijo por todo $G(\overline{k}/k)$, luego $\alpha \in k$.

Otra consecuencia es que si $\alpha \in \overline{k}$, sus conjugados sobre k son también conjugados sobre $k(V)$, luego el polinomio mínimo de α sobre k sigue siendo irreducible sobre $k(V)$. De aquí se sigue que si l es una extensión finita de k entonces $|l : k| = |l(V) : k(V)|$, y a su vez esto implica que una k-base de $\overline{k}$ es también una $k(V)$-base de $\overline{k}(V)$. En otros términos, todo elemento de $\overline{k}(V)$ se expresa de forma única como combinación lineal de dicha k-base con coeficientes en $k(V)$. De aquí podemos concluir que existe un isomorfismo natural $\overline{k}(V) \cong k(V) \otimes_k \overline{k}$ y, por consiguiente, $\overline{k}[V] \cong k[V] \otimes_k \overline{k}$.

Tenemos las sucesiones exactas

$$\begin{array}{ccccccccc}
0 & \longrightarrow & I(V/k) \otimes_k \overline{k} & \longrightarrow & k[X_1,\dots,X_n] \otimes_k \overline{k} & \longrightarrow & k[V] \otimes_k \overline{k} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & I(V) & \longrightarrow & \overline{k}[X_1,\dots,X_n] & \longrightarrow & \overline{k}[V] & \longrightarrow & 0
\end{array}$$

donde las dos últimas flechas verticales son los isomorfismos naturales, que forman un cuadrado conmutativo. De aquí concluimos que el isomorfismo central se restringe a un isomorfismo $I(V) \cong I(V/k) \otimes_k \overline{k}$. En otros términos, concluimos que el ideal $I(V)$ es el espacio vectorial (o el ideal) generado por $I(V/k)$.

Tenemos así una caracterización de las variedades definidas sobre k (que enunciamos conjuntamente con el teorema 1.1):

Teorema 1.2 *Si $V \subset A^n$ es una variedad afín, las afirmaciones siguientes son equivalentes:*

a) Existe $S \subset k[X_1,\dots,X_n]$ tal que $V = V(S)$ (es decir, V está definida sobre k).

b) Existe $S \subset k[X_1,\dots,X_n]$ tal que $I(V) = (S)$.

c) $\sigma[V] = V$ para todo $\sigma \in G(\overline{k}/k)$.

Hemos probado que $k(V)$ es el cuerpo fijado por $G(\overline{k}/k)$ en $\overline{k}(V)$. Más delicado es el teorema siguiente:

Teorema 1.3 *Si V es una variedad afín definida sobre k, entonces*

$$k[V] = \{f \in \overline{k}[V] \mid f^\sigma = f \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

DEMOSTRACIÓN: Una inclusión es trivial, pero para la contraria necesitamos usar un resultado de cohomología de grupos.

Tomemos $f = [F] \in \overline{k}[V]$ invariante por $G(\overline{k}/k)$. Para cada $\sigma \in G(\overline{k}/k)$ existe un polinomio $H_\sigma \in I(V)$ tal que $F^\sigma = F + H_\sigma$. Notemos que F tiene un número finito de conjugados, luego en realidad tenemos un número finito de polinomios H_σ.

Teniendo en cuenta que $I(V)$ está generado por $I(V/k)$, es claro que $I(V)$ es, de hecho, el $\overline{k}$-espacio vectorial generado por $I(V/k)$. Los polinomios H_σ se expresarán como combinación lineal de un número finito de polinomios de $I(V/k)$, con coeficientes en una extensión finita de Galois K de k. Sea $M \subset I(V)$ el K-espacio vectorial generado por estos polinomios y sea $B = \{f_1, \dots, f_r\}$ una base de M.

El ideal $I(V)$ tiene un generador formado por polinomios de $k[X_1, \dots, X_n]$. Si multiplicamos los generadores por todos los monomios con coeficiente 1, obtenemos un generador de $I(V)$ como $\overline{k}$-espacio vectorial, del cual podemos extraer una base. En definitiva, tenemos una $\overline{k}$-base de $I(V)$ formada por polinomios de $k[X_1, \dots, X_n]$ (invariantes por $G(\overline{k}/k)$). Podemos suponer además que K contiene a todos los coeficientes de F. Así, los polinomios F^σ con $\sigma \in G(K/k)$ son todos los conjugados de F por $G(\overline{k}/k)$ y los polinomios $H_\sigma = F^\sigma - F$ son los mismos que los que ya teníamos definidos para $\sigma \in G(\overline{k}/k)$. En particular $H_\sigma \in M$.

Ahora bien, para $\sigma, \tau \in G(K/k)$ tenemos que

$$F^{\sigma\tau} = F + H_{\sigma\tau} = (F + H_\sigma)^\tau = F + H_\tau + H_\sigma^\tau,$$

luego $H_{\sigma\tau} = H_\tau + H_\sigma^\tau$.

Más aún, si llamamos $H_\sigma^i \in K$ a la coordenada de H_σ en la base B correspondiente a f_i, tenemos igualmente que $H_{\sigma\tau}^i = H_\tau^i + H_\sigma^{i\tau}$ (aquí usamos que la base es invariante por $G(K/k)$).

Esto significa que la aplicación $\sigma \mapsto H_\sigma^i$ (para un i fijo) es un 1-cociclo respecto de la acción de $G = G(K/k)$ sobre K^+. Ahora usamos un resultado de la cohomología de grupos finitos, a saber, que $H^1(G, K^+) = 0$, lo que significa que todo cociclo es una cofrontera. Explícitamente, existe un $G_i \in K$ tal que $H_\sigma^i = G_i^\sigma - G_i$. El polinomio $G \in I(V)$ que en la base B tiene coordenadas G_i verifica la relación $H_\sigma = G^\sigma - G$, con lo que $F^\sigma - F = G^\sigma - G$, o equivalentemente $(F + G)^\sigma = F + G$, para todo $\sigma \in G(K/k)$.

Concluimos que $F + G \in k[X_1, \dots, X_n]$ y así $[F] = [G] \in k[V]$. ∎

La importancia del teorema anterior se debe a que ahora sabemos que

$$\overline{k}[V] \cap k(V) = k[V].$$

En otros términos: si $\alpha \in \overline{k}[V]$ está definida sobre k como elemento de $\overline{k}(V)$ (es decir, si es el cociente de dos elementos de $k[V]$), entonces está en $k[V]$. En efecto, al estar definida sobre k es invariante por $G(\overline{k}/k)$.

Observemos que para calcular una función $\alpha \in k(V)$ en un punto $P \in V$ podría hacer falta expresarla como $\alpha = f/g$ con $f, g \in \overline{k}[V]$, pero aún así, si P

es un punto racional de V se cumple que $\alpha(P) \in k$, pues $\alpha(P)$ es invariante por el grupo de Galois $G(\overline{k}/k)$.

Para cada punto $P \in V$ definimos el *anillo local* $\mathcal{O}_P(V)$ como el anillo de las funciones racionales de V regulares en P. Claramente $k[V] \subset \mathcal{O}_P(V) \subset k(V)$.

Teorema 1.4 *Sea V una variedad afín definida sobre k. Entonces*

$$k[V] = \bigcap_{P \in V} \mathcal{O}_P(V),$$

es decir, $k[V]$ es el conjunto de las funciones de $k(V)$ que no tienen singularidades.

Demostración: Sea $\alpha \in k(V)$ una función sin singularidades. Consideremos el ideal

$$I = \{G \in \overline{k}[X_1, \ldots, X_n] \mid [G]\alpha \in \overline{k}[V]\}.$$

Se cumple que $V(I) = \varnothing$. En efecto, como claramente $I(V) \subset I$, tenemos que $V(I) \subset V(I(V)) = V$, luego si $P \in V(I)$ entonces P es un punto de V, donde α no es singular, luego $\alpha = [F]/[G]$ con $G(P) \neq 0$, pero entonces $[G]\alpha = [F] \in \overline{k}[V]$ y $G \in I$, contradicción.

Por el teorema de los ceros de Hilbert ha de ser $I = \overline{k}[X_1, \ldots, X_n]$, luego $1 \in I$ y, por lo tanto, $\alpha = [1]\alpha \in \overline{k}[V] \cap k(V) = k[V]$. ∎

Necesitamos un último resultado sobre funciones polinómicas:

Si $\phi : V \longrightarrow W$ es una aplicación polinómica definida sobre k podemos definir el k-homomorfismo de anillos $\overline{\phi} : k[W] \longrightarrow k[V]$ dado por $\overline{\phi}(f) = \phi \circ f$.

Es claro que $\overline{\phi}$ determina a ϕ, pues si ψ es otra aplicación polinómica y se cumple $\overline{\phi} = \overline{\psi}$, entonces, para todo $P \in V$, tenemos que $\overline{\phi}(x_i)(P) = \overline{\psi}(x_i)(P)$, es decir, $x_i(\phi(P)) = x_i(\psi(P))$, luego $\phi(P) = \psi(P)$.

Teorema 1.5 *Si $V \subset A^m$ y $W \subset A^n$ son variedades afines definidas sobre k, la correspondencia $\phi \mapsto \overline{\phi}$ es una biyección entre las aplicaciones polinómicas $\phi : V \longrightarrow W$ (definidas sobre k) y los k-homomorfismos $\overline{\phi} : k[W] \longrightarrow k[V]$.*

Demostración: Hemos probado que la correspondencia es inyectiva. Para probar que es suprayectiva tomamos un k-homomorfismo $\alpha : k[W] \longrightarrow k[V]$. Sea $\alpha(x_i) = [F_i]$, con $F_i \in k[X_1, \ldots, X_m]$. Entonces los F_i determinan una aplicación polinómica $\phi : A^m \longrightarrow A^n$, así como el homomorfismo de anillos $\phi^* : k[X_1, \ldots, X_n] \longrightarrow k[X_1, \ldots, X_m]$ dado por $G \mapsto G(F_1, \ldots, F_n)$.

Si G es un generador de $I(W)$ con coeficientes en k, entonces $\phi^*(G) \in I(V)$, pues la clase de $\phi^*(G)$ módulo $I(V)$ es

$$G([F_1], \ldots, [F_n]) = G(\alpha(x_1), \ldots, \alpha(x_n)) = \alpha([G]) = \alpha(0) = 0.$$

Por lo tanto $\phi[V] \subset W$, pues si $P \in V$ y $G \in I(W)$ es un generador, entonces $G(\phi(P)) = \phi^*(G)(P) = 0$, luego $\phi(P) \in V(I(W)) = W$. Así pues, ϕ se restringe a una función polinómica de V en W definida sobre k, y es fácil ver que cumple $\overline{\phi} = \alpha$. ∎

1.2 Variedades proyectivas

Llamaremos $\mathrm{P}^n(k)$ al espacio proyectivo de dimensión n sobre el cuerpo k. Como en el caso afín podemos considerar $\mathrm{P}^n(k) \subset \mathrm{P}^n(\overline{k})$. Escribiremos simplemente P^n para referirnos a $\mathrm{P}^n(\overline{k})$ y a los puntos de $\mathrm{P}^n(k)$ los llamaremos *puntos racionales* de P^n.

Diremos que un sistema de referencia de P^n está *definido sobre* k si está determinado por puntos racionales. Respecto a un sistema en estas condiciones, los puntos racionales son los que admiten un vector de coordenadas homogéneas en k^{n+1}. En lo sucesivo supondremos siempre que los sistemas de coordenadas que consideramos están definidos sobre k.

El grupo $G(\overline{k}/k)$ actúa de forma natural sobre P^n y fija a los puntos racionales. El recíproco no es trivial:

Teorema 1.6 *En las condiciones anteriores, se cumple que*

$$\mathrm{P}^n(k) = \{P \in \mathrm{P}^n \mid P^\sigma = P \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Demostración: Fijemos un sistema de referencia y consideremos un punto invariante de coordenadas $P = [\alpha_1, \ldots, \alpha_{n+1}]$. Sea K una extensión finita de Galois de k que contenga a todos los α_i. Tenemos entonces que

$$[\alpha_1^\sigma, \ldots, \alpha_{n+1}^\sigma] = [\alpha_1, \ldots, \alpha_{n+1}], \qquad \text{para todo } \sigma \in G(K/k),$$

luego existe un $\lambda_\sigma \in K^*$ tal que $\alpha_i^\sigma = \lambda_\sigma \alpha_i$, para $i = 1, \ldots, n+1$.

Si σ, $\tau \in G(K/k)$, tenemos que

$$\alpha_i^{\sigma\tau} = \lambda_{\sigma\tau}\alpha_i = (\lambda_\sigma \alpha_i)^\tau = \lambda_\sigma^\tau \lambda_\tau \alpha_i.$$

Como algún α_i ha de ser no nulo, concluimos que $\lambda_{\sigma\tau} = \lambda_\tau \lambda_\sigma^\tau$. Esto significa que $\sigma \mapsto \lambda_\sigma$ es un 1-cociclo para la acción de $G = G(K/k)$ sobre K^*, y el teorema de Hilbert-Speiser afirma que $H^1(G, K^*) = 1$, luego existe un $\lambda \in K^*$ tal que $\lambda_\sigma = \lambda^\sigma/\lambda$. Por consiguiente, $\alpha_i^\sigma = \lambda^\sigma \alpha_i/\lambda$ o, lo que es lo mismo, $(\alpha_i/\lambda)^\sigma = \alpha_i/\lambda$ para todo $\sigma \in G(K/k)$. Por consiguiente $\alpha_i/\lambda \in k$ y

$$P = [\alpha_1/\lambda, \ldots, \alpha_{n+1}/\lambda] \in \mathrm{P}^n(k).$$

∎

Si S es un conjunto de formas (polinomios homogéneos) en $\overline{k}[X_1, \ldots, X_{n+1}]$, llamaremos $V(S)$ al conjunto de todos los puntos de P^n cuyas coordenadas homogéneas en un sistema de referencia prefijado anulan a todos los polinomios de S.

Se dice que un conjunto $C \subset \mathrm{P}^n$ es *algebraico* si $C = V(S)$, para cierto conjunto de formas $S \subset \overline{k}[X_1, \ldots, X_{n+1}]$. Si $S \subset k[X_1, \ldots, X_{n+1}]$ diremos que C *está definido sobre* k. El carácter algebraico de un conjunto (y el hecho de que esté o no definido sobre k) no depende del sistema de referencia considerado.

Si C es un conjunto algebraico definido sobre k, llamaremos $C(k)$ al conjunto de todos los puntos racionales de C, es decir, $C(k) = C \cap \mathrm{P}^n(k)$. Notemos que $\sigma[C] = C$ para todo $\sigma \in G(\overline{k}/k)$, así como que

$$C(k) = \{P \in C \mid P^\sigma = P \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

La prueba del teorema 1.1 es válida palabra por palabra en este contexto, de modo que un conjunto algebraico C está definido sobre k si y sólo si el grupo $G(\overline{k}/k)$ actúa sobre C de forma natural.

Llamaremos $I(C)$ al ideal de todos los polinomios de $\overline{k}[X_1, \ldots, X_{n+1}]$ que se anulan[2] sobre (las coordenadas de) los puntos de C. Se prueba que está generado por un conjunto (finito) de formas. Definimos

$$I(C/k) = I(C) \cap k[X_1, \ldots, X_{n+1}],$$

que es un ideal de $k[X_1, \ldots, X_{n+1}]$.

Una *variedad proyectiva* es un conjunto algebraico $V \subset \mathrm{P}^n$ tal que $I(V)$ es un ideal primo. Si V está definida sobre k, es claro que $I(V/k)$ es también un ideal primo.

Fijado un sistema de referencia, definimos el dominio íntegro

$$k_h[V] = k[X_1, \ldots, X_{n+1}]/I(V/k)$$

y el cuerpo $k(V)$ formado por las fracciones de su cuerpo de cocientes determinadas por clases de formas del mismo grado. Similarmente se define $\overline{k}(V)$. Como ambos cuerpos están generados por las fracciones $[X_i/X_{n+1}]$, tenemos que $\overline{k}(V) = \overline{k}k(V)$.

Diremos que un $\alpha \in k(V)$ es *regular* en P si $\alpha = f/g$, con $f, g \in \overline{k}_h[V]$ y $g(P) \neq 0$. (Notemos que $g(P)$ no está bien definido, pero la condición $g(P) \neq 0$ sí lo está.) En tal caso, el valor $\alpha(P) = f(P)/g(P)$ está bien definido. Si α no es regular en P diremos que es *singular* en P.

Así, los elementos de $k(V)$ determinan funciones parciales sobre V. El mismo argumento que en el caso afín justifica que elementos distintos determinan funciones distintas. Más aún, las funciones (parciales) en V obtenidas de este modo no dependen del sistema de referencia con el que se calcula $k(V)$. Por ello consideraremos a $k(V)$ como un cuerpo de funciones parciales en V que admite distintas representaciones como cocientes de clases de polinomios según el sistema de referencia que consideremos. A los elementos de $k(V)$ los llamaremos *funciones racionales* en V (definidas sobre k).

Como en el caso afín, la extensión $\overline{k}(V)/k(V)$ es de Galois y tenemos un isomorfismo natural $G(\overline{k}/k) \cong G(\overline{k}(V)/k(V))$ determinado por

$$\alpha^\sigma(P) = \alpha(P^{\sigma^{-1}})^\sigma.$$

[2]Se entiende que un polinomio se anula sobre un punto del espacio proyectivo si se anula sobre cualquier vector de coordenadas homogéneas, y esto equivale a que se anulen todas las formas que lo componen.

En particular

$$k(V) = \{\alpha \in \overline{k}(V) \mid \alpha^\sigma = \alpha \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Por lo tanto k es algebraicamente cerrado en $k(V)$.

En realidad, estos hechos pueden deducirse de los correspondientes al caso afín. Ello se debe a que si $V \subset \mathrm{P}^n$ es una variedad proyectiva, podemos suponer que no está contenida en el hiperplano $X_{n+1} = 0$, en cuyo caso la intersección $V_* = V \cap A^n$ es una variedad afín. Es fácil ver que una está definida sobre k si y sólo si lo está la otra. De hecho, $I(V_*)$ se obtiene deshomogeneizando los polinomios de $I(V)$ (o sea, haciendo $X_{n+1} = 1$), y esto es un homomorfismo de anillos. Recíprocamente, si V_* está definida sobre k y $F \in I(V)$ es una forma, no puede ocurrir que $X_{n+1} \mid F$, lo cual garantiza que $F = (F_*)^*$ (donde la estrella superior representa la forma que se obtiene al completar con potencias de X_{n+1} el polinomio F_*). Tenemos que $F_* \in I(V_*)$ es combinación lineal de polinomios con coeficientes en k, luego F es combinación lineal de formas con coeficientes en k.

En particular vemos que V está definida sobre k si y sólo si el ideal $I(V)$ está generado por formas con coeficientes en k, luego el teorema 1.2 es válido para variedades proyectivas.

Si $P \in V$, definimos $\mathcal{O}_P(V)$ como el conjunto de las funciones de $k(V)$ regulares en P.

Si V es una variedad afín definida sobre k, es claro que su clausura proyectiva V^* está definida sobre k. Recíprocamente, si V es una variedad proyectiva definida sobre k y tomamos un hiperplano infinito H definido sobre k, la variedad afín $V_* = V \setminus H$ está definida sobre k (y V es su clausura proyectiva).

Si V^* es la clausura proyectiva de V, es conocido que la restricción determina un $\overline{k}$-isomorfismo de $\overline{k}(V^*)$ en $\overline{k}(V)$, que obviamente se restringe a un k-isomorfismo de $k(V^*)$ en $k(V)$. Además, si $P \notin H$, entonces el anillo $\mathcal{O}_P(V^*)$ se corresponde con $\mathcal{O}_P(V)$.

1.3 Variedades cuasiproyectivas

Recordemos que la topología de Zariski en P^n es la que tiene como conjuntos cerrados a los subconjuntos algebraicos. Una *variedad cuasiproyectiva* (o, simplemente, una variedad) en P^n es un abierto V en una variedad proyectiva de P^n respecto a la topología de Zariski. La variedad proyectiva referida es necesariamente la clausura $\overline{V}$ de V en P^n. Las variedades cuasiproyectivas incluyen a las variedades afines y a las proyectivas.

Diremos que V *está definida sobre* k si lo está $\overline{V}$ y para todo automorfismo $\sigma \in G(\overline{k}/k)$ se cumple $\sigma[V] = V$. Esta definición es consistente con la que ya teníamos para variedades afines y proyectivas.

Se demuestra que cada función de $k(\overline{V})$ está determinada por su restricción a V, lo que nos permite definir el cuerpo de las funciones racionales de V como

el formado por las restricciones a V de todas las funciones de $k(\overline{V})$. Tenemos, pues, que $k(V) \cong k(\overline{V})$. De nuevo, esta definición de $k(V)$ es consistente con la que ya teníamos para variedades afines y proyectivas.

Si $\alpha \in \overline{k}(V)$ y $\sigma \in G(\overline{k}/k)$, definimos $\alpha^\sigma \in \overline{k}(V)$ como $\alpha^\sigma(P) = \alpha(P^{\sigma^{-1}})^\sigma$. Si α se extiende a $\overline{\alpha} \in \overline{k}(\overline{V})$, es claro que α^σ es la restricción a V de $\overline{\alpha}^\sigma$. Esto prueba que está bien definida y que la acción de $G(\overline{k}/k)$ sobre $\overline{k}(V)$ es la misma que su acción sobre $\overline{k}(\overline{V})$. En particular

$$k(V) = \{\alpha \in \overline{k}(V) \mid \alpha^\sigma = \alpha \mid \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Para cada punto $P \in V$ definimos $\mathcal{O}_P(V)$ como el subanillo de $k(V)$ formado por las restricciones de las funciones de $\mathcal{O}_P(\overline{V})$. Para cada abierto U en V definimos

$$k[U] = \bigcap_{P \in U} \mathcal{O}_P(V).$$

En particular, $k[V]$ es el anillo de las funciones racionales en V regulares sobre todos los puntos de V. Una vez más, esta definición es consistente con la que ya teníamos para variedades afines. Trivialmente $k[U] = k(V) \cap \overline{k}[U]$.

Puede probarse que toda función $\alpha \in k(V)$ es regular en un abierto $U \subset V$, de modo que $k(V)$ es la unión de los anillos $k[U]$, donde U recorre los abiertos de V.

Una aplicación $\phi : V \longrightarrow W$ entre dos variedades es *regular* si es continua para la topología de Zariski y para todo abierto U de W (tal que $U \cap \phi[V] \neq \varnothing$) y toda función $\alpha \in \overline{k}[U]$, se cumple que $\overline{\phi}(\alpha) = \phi \circ \alpha \in \overline{k}[\phi^{-1}[U]]$. La aplicación ϕ es un *isomorfismo* si es biyectiva y tanto ϕ como ϕ^{-1} son regulares.

Así, toda aplicación regular ϕ induce $\overline{k}$-homomorfismos de anillos

$$\overline{\phi} : \overline{k}[U] \longrightarrow \overline{k}[\phi^{-1}[U]].$$

Si V y W están definidas sobre k, diremos que ϕ está *definida sobre* k si los homomorfismos $\overline{\phi}$ se restringen a k-homomorfismos $\overline{\phi} : k[U] \longrightarrow k[\phi^{-1}[U]]$.

La composición de aplicaciones regulares definidas sobre k es una aplicación regular definida sobre k. Es conocido que las aplicaciones regulares entre variedades afines son precisamente las aplicaciones polinómicas. En virtud del teorema 1.5 tenemos que las aplicaciones regulares definidas sobre k entre variedades afines definidas sobre k son precisamente las aplicaciones polinómicas definidas sobre k.

Teorema 1.7 *Si V es una variedad definida sobre k, entonces $k[V]$ es el conjunto de las aplicaciones regulares $\alpha : V \longrightarrow A^1$ definidas sobre k.*

Demostración: Admitimos que $\overline{k}[V]$ es el conjunto de las aplicaciones regulares $\alpha : V \longrightarrow A^1$. Hemos de ver que α está definida sobre k (en el sentido de que su extensión a $\overline{k}(\overline{V})$ está en $k(\overline{V})$) si y sólo si α está definida sobre k

(en el sentido general para aplicaciones regulares). Una implicación no ofrece dificultad.

Supongamos que α está definida sobre k en el sentido general y vamos a probar que $\alpha^\sigma = \alpha$ para todo $\sigma \in G(\overline{k}/k)$. Si $x \in k[A^1]$ es la identidad, tenemos que $\overline{\alpha}(x) \in k[V]$, luego para todo $P \in V$ se cumple que

$$\overline{\alpha^\sigma}(x)(P) = x(\alpha(P^{\sigma^{-1}})^\sigma) = x^{\sigma^{-1}}(\alpha(P^{\sigma^{-1}}))^\sigma$$

$$= \overline{\alpha}(x)(P^{\sigma^{-1}})^\sigma = \overline{\alpha}(x)^\sigma(P) = \overline{\alpha}(x)(P).$$

Así pues, $\overline{\alpha^\sigma}(x) = \overline{\alpha}(x)$, pero como x es la identidad, esto es lo mismo que $\alpha^\sigma = \alpha$. ∎

Teorema 1.8 *Si $\phi : V \longrightarrow W$ es una aplicación regular definida sobre k y $P \in V$ es un punto racional, entonces $\phi(P)$ también lo es.*

Demostración: Podemos suponer que $x_{n+1}(\phi(P)) \neq 0$. Consideremos el abierto $U = \{Q \in W \mid x_{n+1}(Q) \neq 0\} \subset W$. Entonces $P \in \phi^{-1}[U]$ y $x_i/x_{n+1} \in k[U]$, luego $\overline{\phi}(x_i/x_{n+1}) \in k(\phi^{-1}(U))$. Por consiguiente

$$\overline{\phi}(x_i/x_{n+1})(P) = x_i(\phi(P))/x_{n+1}(\phi(P)) \in k.$$

De aquí se sigue que si tomamos un vector de coordenadas homogéneas para $\phi(P)$ cuya última coordenada valga 1, todas las demás estarán en k. ∎

Una *aplicación racional* $\phi : V \longrightarrow W$ entre dos variedades es una aplicación regular definida en un subconjunto abierto de V que no puede extenderse a una aplicación regular en ningún abierto mayor. Se dice que ϕ es *regular* o *singular* en un punto $P \in V$ según si ϕ está o no definida en P.

Si las variedades V y W están definidas sobre k, diremos que ϕ *está definida sobre* k si lo está la restricción al abierto en el que está definida.

Cualquier aplicación regular de un abierto de V en W se extiende a una única aplicación racional $\phi : V \longrightarrow W$. La composición de aplicaciones racionales (definidas sobre k) es una aplicación racional (definida sobre k). En vista del teorema 1.7, también es evidente que $k(V)$ es el conjunto de las funciones racionales $\alpha : V \longrightarrow A^1$ definidas sobre k.

Sea $\phi : V \longrightarrow W$ una aplicación racional entre variedades definidas sobre k. Para cada $\sigma \in G(\overline{k}/k)$, podemos definir $\phi^\sigma : V \longrightarrow W$ mediante

$$\phi^\sigma(P) = \phi(P^{\sigma^{-1}})^\sigma.$$

Es claro que si ϕ es regular en un abierto $U \subset V$, entonces ϕ^σ es regular en el abierto $\sigma[U] \subset V$, luego ϕ^σ es también una aplicación racional. Además esta definición extiende a la que ya teníamos sobre $\overline{k}(V)$.

Teorema 1.9 *Una aplicación racional $\phi : V \longrightarrow W$ entre variedades definidas sobre k está definida sobre k si y sólo si $\phi^\sigma = \phi$ para todo $\sigma \in G(\overline{k}/k)$.*

DEMOSTRACIÓN: Supongamos que ϕ es invariante por automorfismos. Sea U un abierto en W que corte a la imagen de α y sea $\alpha \in k[U]$. Hemos de probar que $\overline{\phi}(\alpha) \in k[\phi^{-1}[U]]$ o, simplemente, que $\overline{\phi}(\alpha) \in k(V)$, pues ya sabemos que ϕ es racional. En efecto,

$$\overline{\phi}(\alpha)^\sigma = (\phi \circ \alpha)^\sigma = \phi^\sigma \circ \alpha^\sigma = \phi \circ \alpha = \overline{\phi}(\alpha),$$

luego $\overline{\phi}(\alpha) \in k(V)$ y ϕ está definida sobre k. Recíprocamente, si ϕ está definida sobre k y $\sigma \in G(\overline{k}/k)$, sea $V_0 \subset V$ el abierto donde ϕ está definida y sea $P \in V_0 \cap \sigma[V_0]$. Si $\phi(P) \neq \phi^\sigma(P)$, podríamos tomar $\alpha \in k(W)$ que fuera regular en ambos puntos pero que tomara valores distintos en ellos (basta restringir a W un cociente de formas lineales adecuado con coeficientes en k). Entonces $\overline{\phi}(\alpha) \neq \overline{\phi^\sigma}(\alpha)$, mientras que, razonando como antes, $\overline{\phi}(\alpha) = \overline{\phi}(\alpha)^\sigma = \overline{\phi^\sigma}(\alpha)$, contradicción ∎

Si $\phi : V \longrightarrow W$ es una aplicación racional densa (es decir, con imagen densa) entre variedades definidas sobre k, la composición con ϕ induce un $\overline{k}$-monomorfismo de cuerpos $\overline{\phi} : \overline{k}(W) \longrightarrow \overline{k}(V)$. Se cumple que ϕ está definida sobre k si y sólo si $\overline{\phi}$ se restringe a un k-monomorfismo $\overline{\phi} : k(W) \longrightarrow k(V)$. En efecto: si se cumple esto y $U \subset V$ es un abierto donde ϕ es regular, para cada abierto $W' \subset W$ y cada $f \in k[W']$, tenemos que $\overline{\phi}(f) \in \overline{k}[\phi^{-1}[W']] \cap k(V) = k[\phi^{-1}[W']]$, luego $\phi|_U$ está definida sobre k.

Teorema 1.10 *Dadas dos variedades V y W definidas sobre k, la correspondencia $\phi \mapsto \overline{\phi}$ biyecta las aplicaciones racionales densas $\phi : V \longrightarrow W$ definidas sobre k con los k-monomorfismos $\overline{\phi} : k(W) \longrightarrow k(V)$.*

DEMOSTRACIÓN: Admitiendo el teorema cuando k es algebraicamente cerrado, sólo hay que observar que todo k-monomorfismo $\overline{\phi}$ se extiende a un único $\overline{k}$-monomorfismo $\overline{\phi} : \overline{k}(W) \longrightarrow \overline{k}(V)$. ∎

Sea $\phi : V \longrightarrow W$ una función racional definida sobre k, con $V \subset \mathrm{P}^n$ y $W \subset \mathrm{P}^m$ y sea $P_0 \in V$. Podemos suponer sin pérdida de generalidad que $X_{m+1}(P_0) \neq 0$, con lo que, llamando A^m al espacio afín dado por $X_{m+1} \neq 0$, tenemos que $U = \phi^{-1}[A^m]$ es un entorno de P_0 en V, la variedad $W_* = W \cap A^m$ es afín y $\phi|_U : U \longrightarrow W_*$ sigue siendo una aplicación racional sobre k. Al componer ϕ con las funciones coordenadas $x_i \in k[W_*]$ obtenemos funciones racionales $f_i = \phi|_U \circ x_i \in k(V)$. Por lo tanto, la función ϕ viene dada sobre U por una expresión de la forma

$$\phi(P) = [f_1(P), \ldots, f_m(P), 1].$$

Vemos así que toda función racional $\phi : V \longrightarrow W$ definida sobre k admite en un entorno de cada punto una expresión de la forma

$$\phi(P) = [f_1(P), \ldots, f_{m+1}(P)], \qquad f_i \in k(V),$$

definida en el abierto de puntos regulares para todas las funciones f_i y donde alguna de ellas no se anula. Recíprocamente, toda expresión de esta forma

determina una función racional en V definida sobre k que es regular (al menos) donde lo son todas las funciones coordenadas y alguna de ellas es no nula. En principio $\phi : V \longrightarrow \mathrm{P}^m$, pero si la imagen de ϕ está contenida en una variedad W definida sobre k, también podemos considerar $\phi : V \longrightarrow W$ (definida sobre k). (Para comprobar que está definida sobre k observamos que $\phi^\sigma(P) = [f_1^\sigma(P), \ldots, f_{m+1}^\sigma(P)]$.)

Una *aplicación birracional* $\phi : V \longrightarrow W$ es un isomorfismo entre un abierto de V y un abierto de W. Decimos que *está definida sobre* k cuando lo está como aplicación racional.

Es fácil ver que las aplicaciones birracionales $\phi : V \longrightarrow W$ definidas sobre k se corresponden con los k-isomorfismos $\overline{\phi} : k(W) \longrightarrow k(V)$, así como que la inversa de una aplicación birracional definida sobre k está también definida sobre k.

1.4 Variedades complejas

En esta sección demostraremos que si K es un subcuerpo de $\mathbb{C}$, toda variedad proyectiva V/K se extiende a una variedad proyectiva $V/\mathbb{C}$ definida por las mismas ecuaciones. Para ello necesitamos algunos resultados algebraicos.

Definición 1.11 Sea F/K una extensión de cuerpos y sean A y B dos dominios intermedios. Diremos que A y B son *linealmente disjuntos* sobre K si cuando $\{a_i\}_{i\in I} \subset A$ y $\{b_j\}_{j\in J} \subset B$ son conjuntos linealmente independientes sobre K, entonces $\{a_ib_j\}_{(i,j)\in I\times J}$ es linealmente independiente sobre K.

Teorema 1.12 *Sea F/K una extensión de cuerpos y sean A y B dos dominios intermedios linealmente disjuntos sobre K.*

a) Si $\{a_i\}_{i\in I}$ y $\{b_j\}_{j\in J}$ son bases de A y B sobre K respectivamente, entonces $\{a_ib_j\}_{(i,j)\in I\times J}$ es una base de AB sobre K.

b) Los cuerpos de cocientes de A y B son linealmente disjuntos sobre K.

Demostración: a) Por hipótesis $\{a_ib_j\}_{(i,j)\in I\times J}$ es linealmente independiente sobre K, y obviamente es un generador de AB, luego es una base.

b) Sean A' y B' los cuerpos de cocientes. Si no fueran linealmente disjuntos, existirían familias finitas $\{a_i\}_{i\in I} \subset A'$ y $\{b_j\}_{j\in J} \subset B'$ linealmente independientes sobre K tales que $\{a_ib_j\}_{(i,j)\in I\times J}$ es linealmente dependiente sobre K. Esto significa que

$$\sum_{i,j} \lambda_{ij} a_i b_j = 0,$$

para ciertos $\lambda_{ij} \in K$ no todos nulos. Existen $a \in A$, $b \in B$ no nulos tales que $aa_i \in A$, $bb_j \in B$ para todo i, j. Entonces

$$\sum_{i,j} \lambda_{ij} (aa_i)(bb_j) = 0,$$

de donde se sigue que la familia $\{(aa_i)(bb_j)\}_{(i,j)\in I\times J}$ es linealmente dependiente sobre K. Por hipótesis, una de las familias $\{aa_i\}_{i\in I} \subset A$ o $\{bb_j\}_{j\in J} \subset B$ ha de ser linealmente dependiente sobre K, lo cual es absurdo. ∎

Teorema 1.13 *Sea F/K una extensión de cuerpos y sean A y B dos dominios intermedios. Si existen K-bases $\{a_i\}_{i\in I}$ y $\{b_j\}_{j\in J}$ de A y B respectivamente, tales que $\{a_ib_j\}_{(i,j)\in I\times J}$ es linealmente independiente sobre K, entonces A y B son linealmente disjuntos sobre K.*

Demostración: Para probar que A y B son linealmente disjuntos basta tomar familias finitas $\{u_r\}_{r=1}^m \subset A$, $\{v_s\}_{s=1}^n \subset B$ linealmente independientes sobre K y demostrar que el producto es también linealmente independiente. Sean $I' \subset I$, $J' \subset J$ conjuntos finitos tales que

$$\{u_r\}_{r=1}^m \subset \langle a_i \mid i \in I' \rangle\,, \qquad \{v_s\}_{s=1}^n \subset \langle b_j \mid j \in J' \rangle\,.$$

Sean $d = \dim \langle a_i \mid i \in I' \rangle$, $e = \dim \langle b_j \mid j \in J' \rangle$. Podemos extender $\{u_r\}_{r=1}^m$ y $\{v_s\}_{s=1}^n$ a dos bases $\{u_r\}_{r=1}^d$ y $\{v_s\}_{s=1}^e$ de los espacios $\langle a_i \mid i \in I' \rangle$ y $\langle b_j \mid j \in J' \rangle$ respectivamente. Es claro entonces que

$$\langle u_rv_s \mid r = 1,\dots,d,\ s = 1,\dots,e \rangle = \langle a_ib_j \mid i = 1,\dots,d,\ j = 1,\dots,e \rangle\,.$$

Por hipótesis, el segundo espacio tiene dimensión de, luego la familia $\{u_rv_s\}$ ha de ser linealmente independiente. ∎

Teorema 1.14 *Sea K un cuerpo algebraicamente cerrado y sea $K(S)$ una extensión tal que S sea algebraicamente independiente sobre K. Supongamos que $S = S_1 \cup S_2$, con $S_1 = \{x_1,\dots,x_m\}$, $S_2 = \{y_1,\dots,y_n\}$. En una clausura algebraica de $K(S)$, consideremos dos cuerpos L_1 y L_2 tales que las extensiones $L_i/K(S_i)$ sean finitas de Galois. Entonces L_1 y L_2 son linealmente disjuntos sobre K.*

Demostración: Sea $\alpha_i \in L_i$ tal que $L_i = K(S_i)(\alpha_i)$. Consideremos el epimorfismo de anillos $K[X_1,\dots,X_{m+1}] \longrightarrow K[S_1][\alpha_1]$ dado por $X_i \mapsto x_i$, para $i = 1,\dots,m$, y $X_{m+1} \mapsto \alpha_1$. Su núcleo es un ideal primo que determina una variedad afín V_1 tal que $K[V_1] \cong K[S_1][\alpha_1]$ y el isomorfismo se extiende a un isomorfismo $K(V_1) \cong K(S_1)(\alpha_1)$. Las funciones coordenadas $x_1,\dots,x_m$ son algebraicamente independientes, y el polinomio mínimo de x_{m+1} sobre $K(x_1,\dots,x_m)$ se corresponde con el de α_1 sobre $K(S_1)$.

Similarmente, definimos una variedad afín V_2 tal que $K[V_2] \cong K[S_2][\alpha_2]$ y $K(V_2) \cong K(S_2)(\alpha_2)$. El producto $V_1 \times V_2$ es una variedad afín y cumple que $K(V_1 \times V_2) = K(x_1,\dots,x_{m+1},y_1,\dots,y_{n+1})$. Una base de trascendencia de este cuerpo es $\{x_1,\dots,x_m,y_1,\dots,y_n\}$, que nos da un isomorfismo

$$K(x_1,\dots,x_m,y_1,\dots,y_n) \cong K(S).$$

Si extendemos el isomorfismo a dos clausuras algebraicas, tenemos que x_{m+1} ha de corresponderse con un conjugado de α_1, mientras que y_{n+1} ha de corresponderse con un conjugado de α_2. Cambiando los α_i por estos conjugados,

podemos suponer que el isomorfismo $K(V_1 \times V_2) \cong K(S)(\alpha_1, \alpha_2)$ hace corresponder las funciones coordenadas con los generadores del cuerpo de la derecha. En particular se restringe a un isomorfismo $K[V_1 \times V_2] \cong K[S][\alpha_1, \alpha_2]$.

Por otra parte tenemos un isomorfismo $K[V_1] \otimes_K K[V_2] \cong K[V_1 \times V_2]$ dado por $f \otimes g \mapsto fg$, donde en la parte derecha hay que entender que $f(P,Q) = f(P)$ y $g(P,Q) = g(Q)$. Es claro que se trata de un homomorfismo bien definido y trivialmente es suprayectivo, pues la imagen contiene a las funciones coordenadas.

Una K-base de $K[V_1] \times_K K[V_2]$ es de la forma $\{f_i \otimes g_j\}$, donde $\{f_i\}_i$ es una K-base de $K[V_1]$ y $\{g_j\}_j$ es una K-base de $K[V_2]$. Para probar que el epimorfismo es inyectivo basta ver que esta base se transforma en un conjunto linealmente independiente. Ahora bien, si existieran $c_{ij} \in K$ tales que

$$\sum_{i,j} c_{ij} f_i(P) g_j(Q) = 0$$

para todo $(P,Q) \in V_1 \times V_2$, fijando Q tenemos una combinación lineal nula de las funciones f_i, lo que implica que

$$\sum_{i,j} c_{ij} g_j(Q) = 0$$

para todo $Q \in V_2$, de donde se sigue que los c_{ij} son nulos.

Componiendo los isomorfismos tenemos que

$$K[S][\alpha_1, \alpha_2] \cong K[S_1][\alpha_1] \otimes_K K[S_2][\alpha_2].$$

Por lo tanto, el producto de una K-base de $K[S_1][\alpha_1]$ por una K-base de $K[S_2][\alpha_2]$ es linealmente independiente sobre K. Esto significa que los dominios $K[S_i][\alpha_i]$ son linealmente disjuntos, y por el teorema 1.12 también lo son sus cuerpos de cocientes, es decir, L_1 y L_2. ∎

Ahora ya podemos demostrar el resultado básico:

Teorema 1.15 *Sea K un subcuerpo de $\mathbb{C}$ y sea V/K una variedad proyectiva. Entonces el conjunto algebraico $V/\mathbb{C}$ definido por las mismas ecuaciones que V/K es una variedad proyectiva.*

Demostración: No perdemos generalidad si suponemos que K es algebraicamente cerrado. También es claro que basta probar el teorema para una variedad afín. Sea S_2 una base de trascendencia de $K(V)$ formado por funciones coordenadas. Consideremos una extensión $K(V)(S_1)$, donde S_1 es un conjunto de indeterminadas de cardinal igual al grado de trascendencia de la extensión $\mathbb{C}/K$. Es claro entonces que $S = S_1 \cup S_2$ es una base de trascendencia de $K(V)(S_1)$ sobre K. Si en una clausura algebraica de este cuerpo tomamos la clausura algebraica de $K(S_1)$, obtenemos un cuerpo algebraicamente cerrado cuyo grado de trascendencia sobre K es el mismo que el de $\mathbb{C}$. Obviamente es K-isomorfo a $\mathbb{C}$, luego podemos identificarlo con $\mathbb{C}$. Esto nos permite considerar a $\mathbb{C}$ y a $K[V]$ como subdominios de un mismo cuerpo.

Se cumple que $\mathbb{C}$ y $K[V]$ son linealmente disjuntos, pues si tomamos conjuntos finitos linealmente independientes en uno y en otro, podemos formar dos cuerpos L_1 y L_2 que los contengan y que estén en las condiciones del teorema anterior. Por consiguiente, el epimorfismo natural $\mathbb{C}\otimes_K K[V] \longrightarrow \mathbb{C}K[V]$ es un isomorfismo de K-espacios vectoriales.

Esto implica que si transportamos a $\mathbb{C}\otimes_K K[V]$ el producto de $\mathbb{C}K[V]$, esto es, si definimos $(\alpha\otimes f)(\beta\otimes g) = (\alpha\beta)\otimes(fg)$, la definición es consistente y tenemos un dominio íntegro.

Consideremos ahora la sucesión exacta

$$0 \longrightarrow I(V) \longrightarrow K[X_1,\dots,X_n] \longrightarrow K[V] \longrightarrow 0.$$

El producto tensorial es exacto sobre sucesiones de módulos libres, luego tenemos la sucesión exacta

$$0 \longrightarrow \mathbb{C}\otimes_K I(V) \longrightarrow \mathbb{C}\otimes_K K[X_1,\dots,X_n] \longrightarrow \mathbb{C}\otimes_K K[V] \longrightarrow 0.$$

Pero tenemos un isomorfismo natural $\mathbb{C}\otimes_K K[X_1,\dots,X_n] \cong \mathbb{C}[X_1,\dots,X_n]$, a través del cual $\mathbb{C}\otimes_K I(V)$ se corresponde con el ideal P generado por $I(V)$ en $\mathbb{C}[X_1,\dots,X_n]$. Esto nos da el diagrama conmutativo siguiente:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{C}\otimes_K I(V) & \longrightarrow & \mathbb{C}\otimes_K K[X_1,\dots,X_n] & \longrightarrow & \mathbb{C}\otimes_K K[V] & \longrightarrow & 0\\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & P & \longrightarrow & \mathbb{C}[X_1,\dots,X_n] & \longrightarrow & \mathbb{C}[V] & \longrightarrow & 0
\end{array}$$

Aquí $C[V]$ es el anillo de funciones regulares del conjunto algebraico definido por P, es decir, por las mismas ecuaciones que V. Lo que queremos probar es que P es un ideal primo o, lo que es lo mismo, que $\mathbb{C}[V]$ es un dominio íntegro. Pero tenemos que, como espacio vectorial, $\mathbb{C}[V]$ es isomorfo a $\mathbb{C}\otimes_K K[V]$ y si transportamos la estructura de anillo al producto tensorial, obtenemos el mismo producto que hemos considerado antes, y sabemos que con él es un dominio íntegro. ∎

Nota En la prueba del teorema anterior hemos visto que $\mathbb{C}[V] \cong \mathbb{C}\otimes_K K[V]$, luego los dominios $\mathbb{C}$ y $K[V]$ son linealmente disjuntos, luego también lo son los cuerpos $\mathbb{C}$ y $K(V)$, luego $\mathbb{C}(V) \cong \mathbb{C}\otimes_K K(V)$. Así pues, el cuerpo $K(V)$ determina el cuerpo $\mathbb{C}(V)$. ∎

En las condiciones del teorema anterior, cada función de $K[V]$ se extiende a una función de $\mathbb{C}[V]$, lo que nos permite considerar $K[V] \subset \mathbb{C}[V]$, así como $K(V) \subset \mathbb{C}(V)$. Notemos que los cuerpos $K(V)$ y $\mathbb{C}(V)$ están ambos generados por un mismo sistema de funciones coordenadas afines. Si un subconjunto de ellas es algebraicamente dependiente sobre K, también lo es sobre $\mathbb{C}$, luego tenemos que $\dim V/\mathbb{C} \le \dim V/K$. Por otra parte, si $\dim V/K = n$, existe una sucesión de subvariedades cerradas (sobre K):

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V,$$

que se extienden a otras tantas variedades complejas que mantienen las inclusiones no estrictas, luego $\dim V/\mathbb{C} \geq n$. Así pues, al extender una variedad obtenemos otra variedad de la misma dimensión.

También es claro que si V/K es regular, lo mismo le sucede a $V/\mathbb{C}$, pues un punto singular (para cualquiera de las dos variedades) es solución de un sistema de ecuaciones polinómicas con coeficientes en K. Si el sistema no tiene solución en K^n, el teorema de los ceros de Hilbert implica que el polinomio 1 es combinación lineal de los polinomios que definen el sistema, luego tampoco puede haber solución en $\mathbb{C}^n$.

Teorema 1.16 *Si K es un subcuerpo de $\mathbb{C}$ algebraicamente cerrado y V/K es una variedad proyectiva, entonces $V(K)$ es denso en $V(\mathbb{C})$ para la topología de Zariski.*

Demostración: Lo probaremos por inducción sobre la dimensión de V. Para dimensión 1 basta observar que los abiertos (no vacíos) de $V(\mathbb{C})$ son cofinitos, mientras que $V(K)$ es infinito.

Supuesto cierto para variedades de dimensión n, supongamos que V tiene dimensión $n+1$. Sea U un abierto no vacío en $V(\mathbb{C})$. Reduciéndolo podemos suponer que no contiene puntos singulares. Tomemos una función coordenada x que no sea constante en V. Entonces x es una función holomorfa no constante, luego es abierta para la topología compleja. Podemos tomar un abierto conexo no vacío $G \subset U$ (abierto para la topología compleja) tal que $x[G]$ es abierto en $\mathbb{C}$. Como K contiene a la clausura algebraica de $\mathbb{Q}$ y ésta es densa en $\mathbb{C}$, vemos que existe un punto $P \in U$ tal que $x(P) = \alpha \in K$.

Si añadimos a V la ecuación $x = \alpha$ obtenemos una subvariedad W de menor dimensión, que se extiende a una subvariedad de $V/\mathbb{C}$ que contiene al punto P. Por hipótesis de inducción, el abierto $W \cap U$ (que no es vacío porque contiene a P) contiene un punto de $W(K)$, luego U contiene un punto de $V(K)$. ∎

Consideremos ahora una aplicación racional $\phi : V \longrightarrow W$ entre variedades definidas sobre K. En un abierto de V, ϕ viene definida por formas del mismo grado:

$$\phi([X_1, \ldots, X_{m+1}]) = [F_1(X_1, \ldots, X_{m+1}), \ldots, F_{n+1}(X_1, \ldots, X_{m+1})].$$

Si $F \in I(W)$, entonces $F(F_1, \ldots, F_{n+1})$ se anula en $V(K)$, luego está en $I(V)$, luego se anula en todos los puntos de $V/\mathbb{C}$. Esto se traduce en que $F_1, \ldots, F_{n+1}$ definen una aplicación racional entre las variedades extendidas que extiende a ϕ. Teniendo en cuenta que $V(\overline{K})$ (donde $\overline{K}$ es la clausura algebraica de K) es denso en $V(\mathbb{C})$, es fácil ver que la extensión es única.

Teorema 1.17 *Sea V/K una variedad proyectiva definida sobre un subcuerpo $K \subset \mathbb{C}$. Entonces $K(V)$ es el subcuerpo de $\mathbb{C}(V)$ fijado por todos los K-automorfismos de $\mathbb{C}$.*

Demostración: Si representamos por $G(\mathbb{C}/K)$ al grupo de K-automorfismos de $\mathbb{C}$ (aunque la extensión $\mathbb{C}/K$ no sea necesariamente algebraica), es

claro que $G(\mathbb{C}/K)$ actúa sobre el cuerpo $\mathbb{C}(V)$ del modo usual, y ciertamente todos los automorfismos de $G(\mathbb{C}/K)$, extendidos a $\mathbb{C}(V)$, fijan a $K(V)$. Tomemos ahora $\alpha \in \mathbb{C}(V) \setminus K(V)$ y vamos a construir un automorfismo que no lo fije. Supongamos en primer lugar que α es algebraico sobre $K(V)$ y sea α' un conjugado de α en una clausura algebraica $\overline{\mathbb{C}(V)}$ de $\mathbb{C}(V)$. Consideremos el $K(V)$-isomorfismo $\sigma : K(V)(\alpha) \longrightarrow K(V)(\alpha')$ que cumple $\sigma(\alpha) = \alpha'$.

Sea S una base de trascendencia de $\mathbb{C}(V)$ sobre $K(V)(\alpha)$. Se cumple que S es también algebraicamente independiente sobre $K(V)(\alpha')$, pues si tuviéramos una ecuación polinómica $F(s_1, \ldots, s_r) = 0$, con $F \in K(V)(\alpha')[X_1, \ldots, X_r]$ y $s_i \in S$, entonces el producto

$$\prod_\sigma F^\sigma(X_1, \ldots, X_r),$$

donde σ recorre los automorfismos de la clausura normal de $K(V)(\alpha)$ sobre $K(V)$, es un polinomio no nulo con coeficientes en $K(V)$ y anula a $s_1, \ldots, s_r$, luego S sería algebraicamente dependiente sobre $K(V)$ y también sobre $K(V)(\alpha)$.

Es claro entonces que σ se extiende a un $K(V)(S)$-isomorfismo

$$\sigma : K(V)(\alpha)(S) \longrightarrow K(V)(\alpha')(S).$$

Como La extensión $\mathbb{C}(V)/K(V)(\alpha)(S)$ es algebraica, σ se extiende a su vez a un $K(V)$-monomorfismo $\sigma : \mathbb{C}(V) \longrightarrow \overline{\mathbb{C}(V)}$. Ahora bien, $\mathbb{C}(V) = \mathbb{C}K(V)$ y, como $\mathbb{C}$ es algebraicamente cerrado y la extensión $\mathbb{C}/(\mathbb{C} \cap K(V)(S))$ es algebraica, ha de ser $\sigma[\mathbb{C}] = \mathbb{C}$, luego $\sigma : \mathbb{C}(V) \longrightarrow \mathbb{C}(V)$ es la (única) extensión del K-automorfismo $\sigma|_{\mathbb{C}}$, y cumple $\alpha^\sigma = \alpha'$.

Si α es trascendente sobre $K(V)$ repetimos el razonamiento anterior partiendo del $K(V)$-automorfismo $\sigma : K(V)(\alpha) \longrightarrow K(V)(\alpha)$ dado por $\sigma(\alpha) = -\alpha$. ∎

Si V/K y W/K son dos variedades proyectivas y $\phi : V \longrightarrow W$ es una aplicación racional entre las variedades complejas correspondientes, es claro que ϕ es la extensión de una aplicación racional definida sobre K, si y sólo si lo son las composiciones de ϕ con las funciones coordenadas de W, lo que nos permite aplicar el teorema anterior para concluir que esto sucede si y sólo si $\phi^\sigma = \phi$ para todo $\sigma \in G(\mathbb{C}/K)$.

Como consecuencia, si K es algebraicamente cerrado, una aplicación regular $\phi; V \longrightarrow W$ está definida sobre K si y sólo si $\phi[V(K)] \subset W(K)$. En efecto, en tal caso, para todo $\sigma \in G(\mathbb{C}/K)$ tenemos que $\phi^\sigma(P) = \phi(P)$ para todo $P \in V(K)$, luego ϕ^σ y ϕ coinciden en un conjunto denso, luego son iguales.

1.5 Curvas proyectivas

En general, si V es una variedad definida sobre k, tenemos que $\overline{k}(V)$ es una extensión algebraica de $k(V)$, por lo que ambos cuerpos tienen el mismo grado de trascendencia sobre sus respectivos cuerpos de constantes k y $\overline{k}$. A dicho

grado de trascendencia se le llama *dimensión* de la variedad V. Las *curvas* son las variedades (cuasiproyectivas) de dimensión 1.

A partir de aquí usaremos los siguientes convenios de notación:

- Como hasta ahora, k será un cuerpo perfecto y $\overline{k}$ su clausura algebraica.
- C/k indicará que C es una curva proyectiva definida sobre k.
- Como hasta ahora, $\overline{k}(C)$ será el cuerpo de las funciones racionales de C y $k(C)$ será el subcuerpo formado por las funciones definidas sobre k.
- Por el contrario, $\mathcal{O}_P(C)$ representará siempre el anillo de las funciones de $\overline{k}(C)$ (no de $k(C)$) regulares en el punto $P \in C$.
- Si P es un punto de una curva C, llamaremos

$$\mathfrak{m}_P(C) = \{\alpha \in \mathcal{O}_P(C) \mid \alpha(P) = 0\},$$

  que es el único ideal maximal del anillo $\mathcal{O}_P(C)$.
- Como hasta ahora, $C(k)$ representará el conjunto (tal vez vacío) de los puntos racionales de C.

Los cuerpos $k(C)$ y $\overline{k}(C)$ son cuerpos de funciones algebraicas (de una variable) con cuerpo de constantes exacto k y $\overline{k}$ respectivamente. Además $\overline{k}(C)$ es una extensión de constantes de $k(C)$.

Representaremos por $D(C)$ el grupo de los divisores de $\overline{k}(C)$ (a los que nos referiremos también como divisores de C), por $H(C)$ el grupo de clases de $\overline{k}(C)$ y por $H_0(C)$ el grupo de clases de grado 0 de $\overline{k}(C)$.

Recordemos que una curva C es *regular* en un punto $P \in C$ si el ideal $\mathfrak{m}_P(C)$ es principal, en cuyo caso $\mathcal{O}_P(C)$ resulta ser un dominio de ideales principales con $\mathfrak{m}_P(C)$ como único ideal primo, el cual determina una valoración $v_P : \overline{k}(C) \longrightarrow \mathbb{Z} \cup \{+\infty\}$ que tiene a $\mathcal{O}_P(C)$ como anillo de enteros. Los generadores de $\mathfrak{m}_P(C)$ (es decir, las funciones $\alpha \in \mathcal{O}_P(C)$ tales que $v_P(\alpha) = 1$) se llaman *parámetros locales* de C en P.

Si C es una curva proyectiva, para cada valoración v en $\overline{k}(C)$ existe un único punto $P \in C$ tal que $v(f) \geq 0$ para toda $f \in \mathcal{O}_P(C)$ y $v(f) > 0$ si $f \in \mathfrak{m}_P(C)$. Si P es un punto regular, la única valoración de $\overline{k}(V)$ situada sobre P en este sentido es v_P. En particular, si C es una curva proyectiva regular, la aplicación $P \mapsto v_P$ biyecta los puntos de C con las valoraciones de $\overline{k}(C)$ o, lo que es lo mismo, con los divisores primos de $\overline{k}(C)$. La acción natural del grupo $G(\overline{k}/k)$ en C se corresponde a través de esta biyección con la acción natural sobre los divisores primos.

Observemos que si C/k es una curva proyectiva regular y $\mathfrak{p}$ es un divisor primo de $k(C)$, entonces $\mathfrak{p}$ factoriza en $\overline{k}(C)$ como $\mathfrak{p} = \mathfrak{P}_1 \cdots \mathfrak{P}_r$ (pues las extensiones de constantes son no ramificadas). Además los primos $\mathfrak{P}_i$ forman una clase de conjugación respecto al grupo de Galois $G(\overline{k}/k)$. Así pues, los

divisores primos de $k(C)$ se corresponden con las clases de conjugación de puntos de C. A cada divisor primo le corresponden tantos puntos conjugados como indica su grado. En particular, los divisores primos de grado 1 en $k(C)$ se corresponden con los puntos de V invariantes por $G(\overline{k}/k)$, es decir, con los puntos racionales de C.

El ejemplo de Selmer Como aplicación de esta última observación vamos a probar la mitad de lo anunciado en la introducción sobre el ejemplo de Selmer, a saber, que la curva

$$3U^3 + 4V^3 + 5W^3 = 0$$

tiene puntos racionales en $\mathbb{R}$ y en todos los cuerpos p-ádicos $\mathbb{Q}_p$. El caso real es trivial. Para encontrar soluciones p-ádicas usaremos el criterio siguiente:[3]

> *Sea p un número primo, $F \in \mathbb{Z}[X_1, \dots, X_n]$ y $c \in \mathbb{Z}^n$ un punto tal que $v_p(F(c)) > 2v_p(F_i'(c))$ para cierto i (donde F_i representa la derivada respecto a X_i). Entonces existe un punto $\alpha \in \mathbb{Q}_p^n$ tal que $F(\alpha) = 0$ y además $\alpha_j \equiv c_j \pmod p$, para $j = 1, \dots, n$.*

Para $p = 2$ aplicamos el criterio al punto $(1, 0, 1)$ con $i = 1$.
Para $p = 3$ consideramos $(0, 2, -1)$ con $i = 2$.
Para $p = 5$ consideramos $(2, -1, 0)$ con $i = 1$.
Para $p > 5$ llamamos C a la curva definida por la ecuación de Selmer sobre el cuerpo $k = \mathbb{Z}/p\mathbb{Z}$ y consideramos el cuerpo de funciones racionales $k(C)$. Sucede que[4] $k(C)$ tiene al menos un divisor primo de grado 1. Éste se corresponde con un punto racional de C, es decir, con una terna $(x_1, x_2, x_3) \in \mathbb{Z}^3$ tal que $3x_1^3 + 4x_2^3 + 5x_3^3 \equiv 0 \pmod p$ y $x_i \not\equiv 0 \pmod p$ para algún i. Basta aplicar el criterio a este punto y a este índice i. ∎

Una propiedad fundamental afirma que una aplicación racional entre curvas proyectivas es regular sobre los puntos regulares. En particular, si $\phi : C_1 \longrightarrow C_2$ es una aplicación regular entre curvas proyectivas regulares, entonces $\phi[C_1]$ ha de ser un subconjunto algebraico de C_2 (las imágenes de las variedades proyectivas son cerradas), y las únicas posibilidades son que sea finito (en cuyo caso sólo puede constar de un punto, por la irreducibilidad de C_1) o que sea toda C_2. Así pues, toda aplicación regular entre curvas es constante o suprayectiva. En particular "densa" equivale a "suprayectiva" o a "no constante", luego el teorema 1.10 nos da que las aplicaciones no constantes $\phi : C_1 \longrightarrow C_2$ entre curvas proyectivas regulares definidas sobre k se corresponden biunívocamente con los k-monomorfismos $\overline{\phi} : k(C_2) \longrightarrow k(C_1)$ entre sus cuerpos de funciones racionales.

Es fácil ver que todo cuerpo de funciones algebraicas sobre un cuerpo de constantes exacto k es k-isomorfo al cuerpo de funciones racionales de una curva proyectiva C/k. Vamos a demostrar que la curva puede tomarse regular.

[3] Es un caso particular del teorema 7.17 de mi Teoría de Números.
[4] Teorema 9.29 de mi Geometría Algebraica. (Ver también las observaciones tras la definición 9.9.)

Teorema 1.18 *Sea K un cuerpo de funciones algebraicas sobre un cuerpo de constantes exacto k. Entonces K es k-isomorfo al cuerpo de funciones racionales de una curva proyectiva regular C/k.*

DEMOSTRACIÓN: Consideremos la extensión de constantes $\overline{K} = \overline{k}K$, que es un cuerpo de funciones algebraicas sobre el cuerpo de constantes exacto $\overline{k}$. Podemos tomar $x_1 \in K$ no constante tal que la extensión $K/k(x_1)$ sea separable. Si K tiene característica 0 sirve cualquier elemento no constante, mientras que si K tiene característica prima p, basta exigir además que $x_1 \notin K^p$. Notamos esto para observar que $x_2 = 1/x_1$ cumple lo mismo. Por el teorema del elemento primitivo, podemos tomar y_1, $y_2 \in K$ tales que $K = k(x_1, y_1) = k(x_2, y_2)$. De aquí se sigue a su vez que $\overline{K} = \overline{k}(x_1, y_1) = \overline{k}(x_2, y_2)$. Para $i = 1, 2$, sea B_i la clausura entera de $k[x_i]$ en K y sea $\overline{B}_i$ la clausura entera de $\overline{k}[x_i]$ en $\overline{K}$.

La razón de hacerlo todo por duplicado es que si $\mathfrak{P}$ es un divisor primo de $\overline{K}$ entonces $v_{\mathfrak{P}}(x_i) \geq 0$ para $i = 1$ o bien $i = 2$, luego $\overline{k}[x_i] \subset \mathcal{O}_{\mathfrak{P}}(\overline{K})$, donde $\mathcal{O}_{\mathfrak{P}}(\overline{K})$ representa el anillo de enteros de la valoración $v_{\mathfrak{P}}$, que es íntegramente cerrado en $\overline{K}$ (porque tiene factorización única), luego $B_i \subset \overline{B}_i \subset \mathcal{O}_{\mathfrak{P}}(\overline{K})$.

Por simplicidad suprimiremos temporalmente el subíndice i. Vamos a usar que, en general, si D es un dominio íntegro noetheriano íntegramente cerrado y L es una extensión finita separable de su cuerpo de cocientes, entonces la clausura entera de D en L es un D-módulo finitamente generado.[5] Aplicando esto a $D = k[x]$ y $L = K$ obtenemos que B es un $k[x]$-módulo finitamente generado. En particular $B = k[x_1, \ldots, x_n]$.

Por otra parte, $\overline{B}$ es la clausura entera de B en $\overline{K}$. En efecto, todo elemento de $\overline{B}$ es entero sobre $\overline{k}[x]$ y, a su vez, x y los elementos de $\overline{k}$ son enteros sobre $k[x]$, luego sobre B.

La prueba del resultado que hemos citado puede refinarse en estas condiciones para demostrar que $\overline{B} = \overline{k}[x_1, \ldots, x_n]$. En efecto, tomemos $f \in \overline{B}$. Existe una extensión finita l de k tal que $f \in lK$. Tomemos una k-base $w_1, \ldots, w_r$ de l, que será también una K-base de lK. Por lo tanto

$$f = \sum_{j=1}^{r} g_j w_j,$$

para ciertos $g_j \in K$. Basta probar que en realidad $g_j \in B$. Ahora bien, consideremos la base dual $z_1, \ldots, x_r \in l$ de $z_1, \ldots, z_r$, esto es, la base que cumple $\mathrm{Tr}(z_i w_j) = \delta_{ij}$, donde Tr es la traza de la extensión lK/K (que sobre constantes coincide con la de l/k). Basta observar que $g_i = \mathrm{Tr}(z_i f)$ es entero sobre B, luego $g_i \in B$.

Consideremos ahora el epimorfismo $k[X_1, \ldots, X_n] \longrightarrow B$ dado por $X_i \mapsto x_i$, cuyo núcleo determina una variedad afín C/k. Claramente $k[C] \cong B$, luego $k(C) \cong K$. Además $\overline{k}(C) \cong \overline{K}$ (lo que implica que C es una curva) y como $\overline{k}[C] \cong \overline{B}$ es íntegramente cerrado en $\overline{k}(C)$, tenemos que C es regular (como curva afín, si bien su clausura proyectiva puede tener singularidades en el infinito).

[5] Teorema 1.19 de mi geometría algebraica.

Ahora volvemos a introducir el índice i, de modo que en realidad tenemos dos curvas regulares afines C_1/k y C_2/k, con la propiedad de que si $\mathfrak{P}$ es un divisor primo de $\overline{K}$ entonces existe un i tal que $\overline{k}[C_i] \subset \mathcal{O}_{\mathfrak{P}}(\overline{K})$, luego $\mathfrak{P} \cap \overline{k}[C_i]$ es un ideal maximal en $\overline{k}[C_i]$ que se corresponde con un punto (regular) P de C_i, en el sentido de que $\mathfrak{P} \cap \overline{k}[C_i]$ está formado por las funciones de $\overline{k}[C_i]$ que se anulan en P. La regularidad implica que $v_{\mathfrak{P}} = v_P$.

Con más rigor, tenemos dos k-isomorfismos $\phi_i : k(C_i) \longrightarrow K$, que se extienden a $\overline{k}$-isomorfismos $\phi_i : \overline{k}(C_i) \longrightarrow \overline{K}$, de modo que cada divisor primo de K se corresponde a través de uno de estos dos isomorfismos con un punto (regular) de una de las curvas C_i.

Si llamamos $\overline{C}_i$ a la clausura proyectiva de la curva C_i, la composición $\phi_2 \circ \phi_1^{-1}$ induce una aplicación birracional $\phi : \overline{C}_1 \longrightarrow \overline{C}_2$ definida sobre k, de modo que $\overline{\phi} = \phi_2 \circ \phi_1^{-1}$.

Esto se traduce en que si $\mathfrak{P}$ es un divisor primo de $\overline{k}(\overline{C}_1)$, entonces, o bien está situado sobre un punto de C_1, o bien $\overline{\phi}(\mathfrak{P})$ está situado sobre un punto de C_2. Notemos además que ϕ es regular sobre C_1 (porque sus puntos son regulares). Todo esto es cierto por simetría si cambiamos los índices y consideramos ϕ^{-1} en lugar de ϕ.

Consideramos ahora la superficie $V = \overline{C}_1 \times \overline{C}_2$, junto con las aplicaciones $\psi_i : C_i \longrightarrow V$ dadas por $\psi_1(P) = (P, \phi(P))$, $\psi_2(P) = (\phi^{-1}(P), P)$. Definimos $C = \psi_1[C_1] \cup \psi_2[C_2]$. Vamos a probar que C es una curva proyectiva regular definida sobre k tal que $k(C) \cong K$.

En primer lugar observamos que si $U \subset C_1$ es el abierto donde ϕ se restringe a un isomorfismo, entonces $X = \{(P, \phi(P)) \mid P \in U\}$ es una curva contenida en $\psi_1(C_1)$ isomorfa a U. Además $\psi_1[C_1] \subset \overline{X}$. En efecto, si $Q \in C_1$ y W es un abierto en V tal que $\psi_1(Q) \in W$, hemos de probar que $W \cap X \neq \varnothing$. Como $\psi_1^{-1}[W]$ y U son abiertos (no vacíos) en C_1, se cumple que $\psi_1^{-1}[W] \cap U \neq \varnothing$ y, por lo tanto, $W \cap X \neq \varnothing$.

Como $\overline{X} \setminus X$ es finito, lo mismo le sucede a $\overline{X} \setminus \psi_1[C_1]$, luego $\psi_1[C_1]$ es una curva cuasiproyectiva isomorfa a C_1 (el isomorfismo es ψ, cuya inversa es la proyección en la primera componente). En particular $\psi_1[C_1]$ es regular, y el mismo razonamiento se aplica a $\psi_2[C_2]$. Más aún, el conjunto X es el mismo en ambos casos, por lo que $C \subset \overline{X}$ es una curva cuasiproyectiva cubierta por los dos abiertos regulares $\psi_i[C_i]$. Como la regularidad es una propiedad local concluimos que C es regular.

El hecho de que las curvas C_i estén definidas sobre k, al igual que ϕ, implican inmediatamente que C es invariante por $G(\overline{k}/k)$, luego C está definida sobre k. Como las aplicaciones ψ_i son birracionales (y están definidas sobre k), tenemos que $k(C) \cong k(C_i) \cong K$. Falta probar que C es proyectiva.

Tomemos un punto $P \in \overline{C}$ y sea $\mathfrak{P}$ un divisor primo de $\overline{k}(\overline{C})$ situado sobre P. La relación $\psi_1 = \phi \circ \psi_2$ se traduce en la relación $\overline{\psi}_1 = \overline{\psi}_2 \circ \overline{\phi}$ entre los $\overline{k}$-isomorfismos inducidos. La construcción de ϕ implica que $\overline{\psi}_i(\mathfrak{P})$ está situado sobre un punto $Q \in C_i$, para algún índice i, pero entonces $P = \psi_i(Q) \in C$. ∎

Así pues, tenemos una correspondencia biunívoca entre las clases de k-isomorfía de cuerpos de funciones algebraicas sobre un cuerpo de constantes

exacto k y las clases de k-isomorfía de curvas proyectivas regulares definidas sobre k.

Nota Supongamos que C es una curva compleja definida sobre $\mathbb{Q}$. Entonces C es birracionalmente equivalente a una curva compleja regular definida sobre $\mathbb{Q}$. Esto no es una consecuencia directa del teorema anterior, pues la extensión $\mathbb{C}/\mathbb{Q}$ no es algebraica. Ahora bien, si $\mathbb{A}$ es la clausura algebraica de $\mathbb{Q}$, el teorema anterior nos da que $\mathbb{A}(C) \cong \mathbb{A}(C')$, donde $C'/\mathbb{Q}$ es una curva proyectiva regular. Dicha curva se extiende a una curva proyectiva regular $C'/\mathbb{C}$, y es claro entonces que $\mathbb{C}(C) \cong \mathbb{C}(C')$. ∎

Volviendo al estudio de las aplicaciones regulares, si $\phi : C_1 \longrightarrow C_2$ es una aplicación regular no constante entre curvas proyectivas regulares definidas sobre k, el k-monomorfismo $\overline{\phi} : k(C_2) \longrightarrow k(C_1)$ nos permite considerar a $k(C_1)$ como extensión de $k(C_2)$. Puesto que ambos cuerpos tienen grado de trascendencia 1 sobre k, la extensión es algebraica y, como $k(C_1)$ es finitamente generado sobre k —luego también sobre $k(C_2)$—, la extensión es finita.

Definimos el *grado* de ϕ como $\operatorname{grad}\phi = |k(C_1) : k(C_2)| = |\overline{k}(C_1) : \overline{k}(C_2)|$, con el convenio de que el grado de una aplicación constante es 0. En general, diremos que una aplicación (no constante) definida sobre k es separable, puramente inseparable, etc. según lo sea o no la extensión $k(C_1)/k(C_2)$ que determina. Teniendo en cuenta que las aplicaciones racionales son regulares en los puntos regulares es fácil ver que las aplicaciones regulares de grado 1 entre curvas proyectivas regulares son los isomorfismos.

Si $\phi : C_1 \longrightarrow C_2$ es una aplicación regular no constante entre curvas proyectivas regulares y $P \in C_1$, entonces $\phi(P)$ se corresponde con el único divisor primo de $\overline{k}(C_2)$ divisible entre P. Representaremos por $e_\phi(P)$ el índice de ramificación de P sobre $\phi(P)$, de modo que si $P_1, \ldots, P_r$ son las antiimágenes de un mismo punto $Q \in C_2$ se cumple la relación

$$e_\phi(P_1) + \cdots + e_\phi(P_r) = \operatorname{grad}\phi.$$

Si ϕ es separable, el número de puntos ramificados es finito. En tal caso, *fórmula del género* de Hurwitz afirma que si g_1 y g_2 son los géneros de E_1 y E_2, entonces

$$2g_1 - 2 \geq (2g_2 - 2)\operatorname{grad}\phi + \sum_{P \in C_1} (e_\phi(P) - 1),$$

y la igualdad se da si y sólo si la característica de $\overline{k}$ es 0 o bien es prima y no divide a ningún índice $e_\phi(P)$.

La aplicación ϕ se extiende por linealidad a un homomorfismo de grupos $\phi : D(C_1) \longrightarrow D(C_2)$. Es claro que $\phi(\mathfrak{a})$ no es sino la norma de $\mathfrak{a}$ en $D(C_2)$. También tenemos la norma $N_\phi : \overline{k}(C_1) \longrightarrow \overline{k}(C_2)$. La compatibilidad entre ambas normas nos da que $\phi((f)) = (N_\phi(f))$, para toda función $f \in \overline{k}(C_1)$, luego ϕ transforma divisores principales en divisores principales y por lo tanto induce un homomorfismo $\phi : H(C_1) \longrightarrow H(C_2)$ entre los grupos de clases.

Por otra parte, el $\overline{k}$-monomorfismo $\overline{\phi} : \overline{k}(C_2) \longrightarrow \overline{k}(C_1)$ induce un monomorfismo $\overline{\phi} : D(C_2) \longrightarrow D(C_1)$, determinado por

$$\overline{\phi}(Q) = \prod_{P \in \phi^{-1}[Q]} P^{e_\phi(P)}.$$

Si $f \in \overline{k}(C_2)$ tenemos que $\overline{\phi}((f)) = (\overline{\phi}(f))$, luego $\overline{\phi}$ transforma divisores principales en divisores principales, por lo que induce un homomorfismo de grupos $\overline{\phi} : H(C_2) \longrightarrow H(C_1)$.

Las propiedades siguientes son meras transcripciones de propiedades de las extensiones de cuerpos de funciones algebraicas:

a) $\operatorname{grad} \overline{\phi}(\mathfrak{a}) = (\operatorname{grad} \phi)(\operatorname{grad} \mathfrak{a})$, para todo $\mathfrak{a} \in D(C_2)$.

b) $\operatorname{grad} \phi(\mathfrak{a}) = \operatorname{grad} \mathfrak{a}$, para todo $\mathfrak{a} \in D(C_1)$.

En particular vemos que ϕ induce homomorfismos entre los grupos de clases de grado 0:

$$\phi : H_0(C_1) \longrightarrow H_0(C_2) \quad \text{y} \quad \overline{\phi} : H_0(C_2) \longrightarrow H_0(C_1).$$

Representaremos por $\Omega(C_i)$ el espacio de las diferenciales de $\overline{k}(C_i)$. Tenemos que ϕ induce un $\overline{k}$-monomorfismo de espacios vectoriales $\overline{\phi} : \Omega(C_2) \longrightarrow \Omega(C_1)$ dado por

$$\overline{\phi}(f\,dx) = \overline{\phi}(f)\, d\overline{\phi}(x).$$

Dado $x \in \overline{k}(C)$, se cumple que $dx \neq 0$ si y sólo si x es separador, es decir, si y sólo si la extensión $\overline{k}(C)/\overline{k}(x)$ es (finita) separable (se puede probar que siempre existen elementos separadores). En particular, si $\operatorname{car} k = 0$, las únicas funciones con diferencial nula son las constantes. De aquí se deduce un criterio de separabilidad:

Teorema 1.19 *Una aplicación regular no constante $\phi : C_1 \longrightarrow C_2$ entre dos curvas proyectivas regulares es separable si y sólo si $\overline{\phi} : \Omega(C_2) \longrightarrow \Omega(C_1)$ es inyectiva, si y sólo si $\overline{\phi}$ es no nula.*

Demostración: Si $\omega = f\,dx \in \Omega(C_2)$ cumple $\omega \neq 0$, entonces $f \neq 0$ y $dx \neq 0$, luego $\overline{k}(C_2)/\overline{k}(x)$ es separable, al igual que la extensión isomorfa $\overline{\phi}[\overline{k}(C_2)]/\overline{k}(\overline{\phi}(x))$. Por otra parte $\overline{\phi}(f) \neq 0$.

Así pues, $\overline{\phi}(\omega) = \overline{\phi}(f)\, d(\overline{\phi}(x)) \neq 0$ si y sólo si $d(\overline{\phi}(x)) \neq 0$, si y sólo si $\overline{k}(C_1)/\overline{k}(\overline{\phi}(y))$ es separable, si y sólo si $\overline{k}(C_1)/\overline{\phi}[\overline{k}(C_2)]$ es separable, si y sólo si ϕ es separable. ∎

Cada forma diferencial $\omega \in \Omega(C)$ tiene asociado un divisor $(\omega) \in D(C)$ determinado por que

$$v_P(f\,dx) = v_P\left(f\,\frac{dx}{dt}\right),$$

donde t es un parámetro local en P. Es claro que $(\overline{\phi}(\omega)) = \overline{\phi}((\omega))$.

El teorema de Riemann-Roch En el capítulo siguiente necesitaremos el teorema de Riemann-Roch en una única ocasión, pero será un uso crucial (en la prueba de que toda curva elíptica admite una ecuación de Weierstrass), así que vamos a recordar su enunciado.

Si K es un cuerpo de funciones algebraicas sobre el cuerpo de constantes exacto k y $\mathfrak{a}$ es un divisor de K, se define el espacio de *múltiplos* de $\mathfrak{a}$ como el conjunto

$$m(\mathfrak{a}) = \{f \in K \mid v_\mathfrak{p}(f) \geq v_\mathfrak{p}(\mathfrak{a}) \text{ para todo divisor primo } \mathfrak{p} \text{ de } K\}.$$

Se prueba que $m(\mathfrak{a})$ es un k-espacio vectorial de dimensión finita. Teniendo en cuenta que los divisores principales tienen grado 0, vemos que si $f \in m(\mathfrak{a})$ es una función no nula, entonces

$$0 = \operatorname{grad}(f) = \sum_\mathfrak{p} v_\mathfrak{p}(f) \operatorname{grad}\mathfrak{p} \geq \sum_\mathfrak{p} v_\mathfrak{p}(\mathfrak{a}) \operatorname{grad}\mathfrak{p} = \operatorname{grad}\mathfrak{a},$$

luego si $\operatorname{grad}\mathfrak{a} > 0$ entonces $\dim m(\mathfrak{a}) = 0$.

Adoptaremos el convenio usual de definir $\dim \mathfrak{a} = \dim m(\mathfrak{a}^{-1})$. De este modo, los divisores de grado < 0 tienen dimensión nula. Es fácil ver que la dimensión depende únicamente de la clase de similitud de $\mathfrak{a}$, por lo que podemos hablar de la dimensión de una clase de divisores de K.

Se prueba que si K tiene género g, entonces contiene una única clase de divisores W, llamada *clase canónica* de K, caracterizada por que

$$\dim W = g, \qquad \operatorname{grad} W = 2g - 2.$$

El teorema de Riemann-Roch afirma que para toda clase de ideales A de K se cumple la relación

$$\dim A = \operatorname{grad} A - (g - 1) + \dim(W/A).$$

Por ejemplo, si K tiene género $g = 1$ (que es el único caso que nos va a interesar), tenemos que la clase canónica tiene grado 0, luego si $\operatorname{grad} A > 0$ se cumple $\operatorname{grad}(W/A) < 0$ y $\dim(W/A) = 0$, con lo que el teorema de Riemann-Roch se reduce a la igualdad $\dim A = \operatorname{grad} A$.

Aprovechamos la ocasión para demostrar un resultado que hemos citado en la introducción, aunque no nos va a hacer falta en ningún momento:

Teorema 1.20 *Toda curva proyectiva C/k de género 0 es birracionalmente equivalente (sobre k) a una cónica definida sobre k.*

Demostración: Tenemos que $k(C)$ es un cuerpo de funciones algebraicas de género $g = 0$. En este caso la clase canónica cumple $\operatorname{grad} W = -2$, luego si $\operatorname{grad} A \geq 0$ entonces $\operatorname{grad} W/A < 0$ y $\dim W/A = 0$, luego el teorema de Riemann-Roch se reduce a la relación

$$\dim A = \operatorname{grad} A + 1.$$

Si $k(C)$ tiene un divisor de grado 1 podemos tomarlo entero (pues toda clase de divisores de grado ≥ 0 contiene un divisor entero), luego primo, digamos $\mathfrak{p}$. Por el teorema de Riemann-Roch $\dim \mathfrak{p} = 2$, luego existe una función no constante $x \in m(\mathfrak{p}^{-1})$.

Entonces $\mathfrak{p}$ ha de ser el denominador de (x) tanto en $k(C)$ como en $k(x)$ y tiene grado 1 en ambos cuerpos, luego $|k(C) : k(x)| = 1$, es decir, $k(C) = k(x)$. De aquí se sigue que C es birracionalmente equivalente (sobre k) a la recta proyectiva, que a su vez es k-isomorfa, por ejemplo, a la cónica $X^2 + Y^2 = 1$.

Si, por el contrario, $k(C)$ no tiene divisores primos de grado 1, al menos tiene que haber uno $\mathfrak{p}$ de grado 2 (basta tomar un divisor entero de W^{-1}). Entonces tenemos que $\dim \mathfrak{p} = 3$, luego podemos tomar tres funciones linealmente independientes 1, x, $y \in m(\mathfrak{p}^{-1})$.

Ahora $(x) = \mathfrak{q}/\mathfrak{p}$, para cierto divisor primo $\mathfrak{q}$ de grado 2. Como $\mathfrak{p}$ tiene grado 1 en $k(x)$, concluimos que $|k(C) : k(x)| = 2$. Además $y \notin k(x)$, pues en caso contrario sería una función racional de x con a lo sumo un único polo simple en el infinito, luego sería un polinomio en x de grado ≤ 1, pero esto contradice la independencia lineal de 1, x, y. De aquí deducimos que $k(C) = k(x, y)$.

Por otra parte, 1, x, y, x^2, xy, $y^2 \in m(\mathfrak{p}^2)$, y este espacio tiene dimensión 5, luego se cumple una relación

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

con los coeficientes en k no todos nulos. Más aún, alguno de los tres primeros ha de ser no nulo, por la independencia lineal de 1, x, y. La ecuación ha de ser irreducible, o llegaríamos a que $y \in k(x)$ o bien a que $x \in k$.

De aquí se sigue inmediatamente que $k(C)$ es k-isomorfo al cuerpo de las funciones racionales (sobre k) de la cónica definida por la ecuación anterior, luego C es birracionalmente equivalente (sobre k) a dicha cónica. ∎

La aplicación de Frobenius Introducimos ahora una técnica para estudiar aplicaciones regulares inseparables. En lo que sigue supondremos que k es un cuerpo (perfecto) de característica prima p. Así, si $m = p^r$, la correspondencia $a \mapsto a^m$ define un automorfismo de $\overline{k}$ que se restringe a un automorfismo de k. Para cada polinomio $F \in \overline{k}[X_1, \ldots, X_n]$, representaremos por $F^{(m)}$ el polinomio que resulta de elevar a m todos los coeficientes de F. Esto define un automorfismo del anillo de polinomios que se restringe a un automorfismo de $k[X_1, \ldots, X_n]$.

Si $C/k \subset \mathrm{P}^n$ es una curva proyectiva, entonces $I(C)$ es un ideal primo de $\overline{k}[X_1, \ldots X_{n+1}]$ generado por formas de $k[X_1, \ldots X_{n+1}]$, luego $I(C)^{(m)}$ cumple lo mismo y define una curva $C^{(m)} \subset \mathrm{P}^n$ definida sobre k. La curva $C^{(m)}$ está determinada por la relación

$$I(C^{(m)}) = \{F^{(m)} \mid F \in I(C)\}.$$

Ahora definimos la aplicación $\phi : C \longrightarrow C^{(m)}$ mediante

$$\phi([x_1, \ldots, x_{n+1}]) = [x_1^m, \ldots, x_{n+1}^m].$$

Esto es correcto, pues si $[x_1,\dots,x_{n+1}]\in C$ y $F\in I(C)$, entonces

$$F^{(m)}(x_1^m,\dots,x_{n+1}^m)=F(x_1,\dots,x_{n+1})^m=0.$$

Obviamente ϕ es una aplicación regular definida sobre k, conocida como *aplicación de Frobenius de grado m*. A continuación probamos que, efectivamente, tiene grado m.

Teorema 1.21 *Sea C/k una curva proyectiva sobre un cuerpo de característica prima p y sea $\phi: C\longrightarrow C^{(m)}$ la aplicación de Frobenius de grado $m=p^r$.*

a) ϕ es puramente inseparable y tiene grado m.

b) $\overline{\phi}[k(C^{(m)})]=k(C)^m$.

c) Si C es regular, entonces $C^{(m)}$ también lo es.

Demostración: b) Fijado un sistema de referencia, toda $\alpha\in k(C^{(m)})$ es de la forma $\alpha=[F^{(m)}]/[G^{(m)}]$, donde F, $G\in k[X_1,\dots,X_{n+1}]$ son formas del mismo grado. Entonces, para $P=[x_1,\dots,x_{n+1}]$ en un abierto adecuado de C,

$$\overline{\phi}(\alpha)(P)=\frac{F^{(m)}(x_1^m,\dots,x_{n+1}^m)}{G^{(m)}(x_1^m,\dots,x_{n+1}^m)}=\frac{F(x_1,\dots,x_{n+1})^m}{G(x_1,\dots,x_{n+1})^m}=\beta^m(P),$$

donde $\beta=[F]/[G]\in k(C)$. Esto prueba la inclusión $\overline{\phi}[k(C^{(m)})]\subset k(C)^m$. Invirtiendo el razonamiento tenemos la inclusión contraria.

a) Ahora es obvio que ϕ es puramente inseparable. Falta probar que su grado es m. Sea $t\in k(C)$ tal que la extensión $k(C)/k(t)$ sea separable. Esto equivale a que $t\notin k(C)^p$. Entonces, la extensión $k(C)/k(C)^m(t)$ está contenida en la extensión separable $k(C)/k(t)$ y en la extensión puramente inseparable $k(C)/k(C)^m$, luego ha de ser $k(C)=k(C)^m(t)$. Por consiguiente,

$$\operatorname{grad}\phi=|k(C)^m(t):k(C)^m|.$$

Ahora bien, del hecho de que $t\notin k(C)^p$ se sigue que el polinomio mínimo de t sobre $k(C)^m$ es x^m-t^m, luego el grado es m.

c) Supongamos que $C\subset \mathrm{P}^n$ y que $I(C)=(F_1,\dots,F_r)$. La regularidad de C en un punto P equivale a que el espacio tangente $T_P(C)$ tenga dimensión 1, lo cual equivale a que la matriz formada por las derivadas parciales de $F_1,\dots,F_r$ tenga rango $n-1$ en P, lo que a su vez equivale a que cierto menor M de orden $n-1$ no se anule.

La regularidad de $C^{(m)}$ en $\phi(P)$ equivale a que la matriz de las derivadas parciales de $F_1^{(m)},\dots,F_r^{(m)}$ tenga rango $n-1$ en $\phi(P)$, pero es claro que

$$\left.\frac{\partial F_i^{(m)}}{\partial X_j}\right|_{\phi(P)}=\left(\left.\frac{\partial F_i}{\partial X_j}\right|_P\right)^m,$$

(aquí usamos que todo natural s cumple $s \equiv s^m$ (mód p)), luego el menor correspondiente a M es $M^m \neq 0$, luego $C^{(m)}$ es regular en $\phi(P)$ y, como ϕ es suprayectiva por el apartado a), concluimos que $C^{(m)}$ es regular. ∎

Es fácil ver que las aplicaciones de Frobenius son biyectivas, pero no son isomorfismos salvo si $m = 1$ (en cuyo caso convenimos que $C^{(1)} = C$ y que ϕ es la identidad), pues las inversas no son regulares. El interés de la aplicación de Frobenius se debe al teorema siguiente:

Teorema 1.22 *Sea $\psi : C_1 \longrightarrow C_2$ una aplicación regular no constante entre curvas regulares definida sobre un cuerpo k de característica prima p y sea $m = \mathrm{grad}_i\, \psi$ (el grado de inseparabilidad). Entonces ψ es una composición*

$$C_1 \xrightarrow{\phi} C_1^{(m)} \xrightarrow{\chi} C_2,$$

donde ϕ es la aplicación de Frobenius de grado m y χ es una aplicación regular separable definida sobre k.

Demostración: Sea K la clausura separable de $\overline{\psi}[k(C_2)]$ en $k(C_1)$. Entonces $k(C_1)/K$ es puramente inseparable de grado m, luego tenemos las inclusiones

$$\overline{\psi}[k(C_2)] \subset k(C_1)^m \subset K \subset k(C_1),$$

pero por el teorema anterior $k(C_1)/k(C_1)^m$ también tiene grado m, luego

$$K = k(C_1)^m = \overline{\phi}[k(C_1^{(m)})].$$

La inclusión $\overline{\psi}[k(C_2)] \subset \overline{\phi}[k(C_1^{(m)})]$ induce una aplicación racional, luego regular (y definida sobre k), $\chi : C_1^{(m)} \longrightarrow C_2$ tal que $\overline{\chi}[k(C_2)] = \overline{\psi}[k(C_2)]$. Es obvio que χ cumple el teorema. ∎

Comentemos por último que si C es una curva definida sobre un cuerpo finito k de m elementos, entonces el automorfismo $a \mapsto a^m$ es la identidad en k, luego $C^{(m)} = C$ y la aplicación de Frobenius es una biyección $\phi : C \longrightarrow C$.

Capítulo II

La geometría de las curvas elípticas

Empezamos el estudio de las curvas elípticas ocupándonos de sus propiedades relacionadas directamente con la geometría algebraica o, lo que esencialmente es lo mismo, con la posibilidad de trabajar con un cuerpo de constantes algebraicamente cerrado. En capítulos posteriores estudiaremos las peculiaridades de las curvas elípticas definibles sobre cuerpos menores. Mantenemos la notación del capítulo anterior. En particular k será un cuerpo perfecto arbitrario y $\overline{k}$ su clausura algebraica.

2.1 Ecuaciones de Weierstrass

Una curva elíptica sobre un cuerpo k es una curva proyectiva regular de género 1 que tenga al menos un punto racional. Esta última condición es superflua si k es algebraicamente cerrado, pero es crucial en caso contrario. Según veremos, la existencia de un punto racional nos permite dotar a la curva de una estructura de grupo definida geométricamente que resulta ser una representación del grupo de clases de grado 0 de su cuerpo de funciones racionales. Por este motivo, en la definición de curva elíptica no sólo exigimos que haya un punto racional, sino que seleccionamos uno para que represente el papel de elemento neutro. En definitiva, la definición queda como sigue:

Definición 2.1 Una *curva elíptica* sobre un cuerpo k es un par (E, O), donde E es una curva proyectiva regular de género 1 y $O \in E(k)$. Un *isomorfismo* entre dos curvas elípticas (E, O) y (E', O') es un isomorfismo entre E y E' que haga corresponder O con O'.

En la práctica escribiremos E (o E/k) en lugar de (E, O). Si E/k es una curva elíptica, el cuerpo $k(E)$ es un *cuerpo de funciones elípticas,* en el sentido de que es un cuerpo de funciones algebraicas de género 1 y tiene al menos un divisor primo de grado 1, a saber, el correspondiente al punto O.

Vamos a demostrar que toda curva elíptica es isomorfa a una curva elíptica plana. Más aún, a una curva determinada por una ecuación con una forma particular:

Definición 2.2 Una *ecuación de Weierstrass* es una ecuación de la forma

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6. \tag{2.1}$$

Pronto comprenderemos la razón de los subíndices. De momento, como regla mnemotécnica, podemos pensar que cada coeficiente a_i tiene asociado un peso i, que las indeterminadas X e Y tienen pesos 2 y 3 respectivamente y que el peso de un monomio es la suma de los pesos de sus factores. Entonces los índices están elegidos para que todos los monomios tengan peso 6.

Toda ecuación de Weierstrass define una curva proyectiva plana con un único punto infinito, $O = [0, 1, 0]$. En principio dicha curva no tiene por qué ser elíptica, ya que puede tener puntos singulares. Por otra parte, éste es el único inconveniente posible, ya que es conocido que toda cúbica regular tiene género 1. Cuando una curva C definida por una ecuación de Weierstrass sea regular, la consideraremos como curva elíptica tomando como punto O el punto infinito.

El hecho de que toda curva elíptica admita una ecuación de Weierstrass es una consecuencia del teorema de Riemann-Roch:

Teorema 2.3 *Si E/k es una curva elíptica, existen funciones x, $y \in k(E)$ tales que la aplicación $\phi : E \longrightarrow \mathrm{P}^2$ dada por $\phi(P) = [x(P), y(P), 1]$ define un isomorfismo entre E y una curva plana C determinada por una ecuación de Weierstrass con coeficientes en k y $\phi(O) = [0, 1, 0]$. Dicha ecuación es única salvo una transformación afín de la forma*

$$X = u^2X' + r, \qquad Y = u^3Y' + su^2X' + t, \qquad u, r, s, t \in k, \quad u \neq 0.$$

Más aún, cualquier k-isomorfismo entre dos curvas elípticas definidas por ecuaciones de Weierstrass (sobre k) está inducido por una transformación afín de este tipo.

Demostración: Para cuerpos de género 1, el teorema de Riemann-Roch implica que la dimensión de un divisor de grado > 0 es igual a su grado. En particular $\dim O^n = n$, luego podemos encontrar una función $x \in k(E)$ tal que 1, x formen una k-base del espacio $m(O^{-2})$. Igualmente, existe $y \in k(E)$ tal que 1, x, y forman una k-base de $m(O^{-3})$.

Notemos que x tiene un polo doble en O, pues de lo contrario estaría en $m(O^{-1})$ y sería constante. Así, el divisor de x es $(x) = \mathfrak{a}/O^2$ y O^2 es el primo infinito del cuerpo $K = k(x)$, luego $|k(V) : k(x)| = 2$.

Por otra parte, $y \notin k(x)$, ya que en tal caso 1, x, $y \in m_K(O^{-2})$, mientras que $\dim_K O^2 = \operatorname{grad}_K O^2 + 1 = 2$, ya que K tiene género 0. Por lo tanto $k(E) = k(x, y)$. También es claro que y tiene un polo en O de orden exactamente igual a 3.

A continuación observamos que 1, x, x^2, x^3, xy, y, $y^2 \in m(O^{-6})$, y la dimensión de este espacio es 6, luego se cumple

$$A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6y^2 + A_7x^3 = 0,$$

con $A_i \in k$. Los coeficientes A_6 y A_7 no pueden ser nulos, o de lo contrario cada sumando tendría un polo de orden distinto en O, lo cual es imposible. Cambiamos $x = -A_6A_7x'$, $y = A_6A_7^2y'$ y, tras dividir entre $A_6^3A_7^4$, la ecuación queda en la forma del enunciado.

La aplicación ϕ del enunciado es regular salvo quizá en O, pero la expresión

$$\phi(P) = [(x/y)(P), 1, (1/y)(P)]$$

muestra que $\phi(O) = [0, 1, 0]$. Por otra parte, el monomorfismo de cuerpos $\overline{\phi} : k(C) \longrightarrow k(E)$ transforma las funciones coordenadas de C en x, y, luego es un isomorfismo y ϕ también lo es.

Supongamos ahora que x, y son funciones arbitrarias que cumplen el enunciado. A través del isomorfismo ϕ se corresponden con las coordenadas X, e Y de la parte afín de C, luego x e y son regulares en toda la curva E salvo en O, donde tienen polos de grados 2 y 3. Esto es consecuencia de que $|k(E) : k(x)| = 2$ y $|k(E) : k(y)| = 3$, luego el primo infinito de $k(x)$ es O^2 y el de $k(y)$ es O^3.

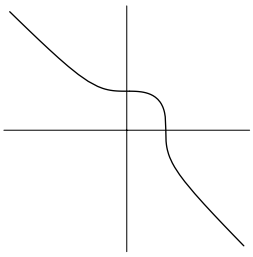
Así pues, si x' e $y' \in k(E)$ dan lugar a una ecuación similar, tenemos que $\{1, x\}$ y $\{1, x'\}$ son bases de $m(O^{-2})$, mientras que $\{1, x, y\}$, $\{1, x', y'\}$ son bases de $m(O^{-3})$. Por consiguiente,

$$x = u_1x' + r, \qquad y = u_2y' + s_2x' + t, \qquad u_1, u_2, r, s_2, t \in k.$$

Además u_1 y u_2 no pueden ser nulos. Si sustituimos estas transformaciones en la ecuación del enunciado, debemos obtener otra ecuación similar, salvo quizá un factor constante que será a la vez el coeficiente de X^3 y el de Y^2. Esto nos da que $u_1^3 = u_2^2$. Haciendo $u = u_2/u_1$ y $s = s_2/u^2$ tenemos $u_1 = u^2$, $u_2 = u^3$, $s_2 = su^2$, y las transformaciones se convierten en las de enunciado.

Por último, supongamos que un isomorfismo $\phi : E \longrightarrow E'$ definido sobre k hace corresponder dos curvas elípticas definidas por ecuaciones de Weierstrass con coeficientes en k. Entonces ϕ induce un k-isomorfismo de cuerpos $\overline{\phi} : k(x', y') \longrightarrow k(x, y)$, donde x, y, x', y' son las funciones coordenadas de E y E' respectivamente. Entonces $\overline{\phi}(x')$, $\overline{\phi}(y')$ son funciones de $k(E)$ que satisfacen una ecuación de Weierstrass sobre k (la misma que x', y'), luego por la parte ya probada ambas ecuaciones están relacionadas por un cambio de variables del tipo indicado. ∎

Ejemplo: La curva de Fermat La *curva de Fermat* es la curva C dada por $U^3+V^3=1$. Recibe este nombre porque su forma homogénea es $U^3+V^3=W^3$, de donde se sigue que el último teorema de Fermat para exponente 3 equivale a que $C(\mathbb{Q})=\{(0,1,1),(1,0,1),(1,-1,0)\}$.



Ciertamente tiene género 1, pues no tiene puntos singulares (las derivadas de la ecuación homogénea no se anulan en ningún punto de C). Tiene tres puntos en el infinito, a saber, $[1,-\omega^i,0]$, donde ω es una raíz cúbica de la unidad. Podemos considerar a C como curva elíptica sobre $\mathbb{Q}$ tomando $O=[1,-1,0]$. Igualmente podríamos definir la curva de Fermat de exponente $n\geq 2$, pero su género sería $g=(n-1)(n-2)/3$, luego sólo obtenemos una curva elíptica cuando $n=3$.

Veamos cómo encontrar una ecuación de Weierstrass para C. Más en general, vamos a considerar la curva C dada por

$$U^3+V^3=aW^3, \qquad a\in k^*.$$

Suponiendo que k no tiene característica 3, vemos que C es también una curva elíptica con $O=[1,-1,0]$. Si k contiene una raíz cúbica de a entonces C es isomorfa sobre k a la curva de Fermat (basta hacer $W'=\sqrt[3]{a}\,W$). En general, C es isomorfa sobre $\overline{k}$ a la curva de Fermat, pero no necesariamente sobre k.

Observemos que la recta tangente a C por O es $U=-V$, luego buscamos una transformación proyectiva $[U,V,W]\mapsto[X,Y,Z]$ que haga corresponder O con $[0,1,0]$ y la recta $U-V=0$ con la recta $Z=0$. La última condición queda garantizada si hacemos $Z=U+V$, y para que se cumpla la primera basta hacer $X=W$. Podemos tomar, por ejemplo,

$$(X,Y,Z)=(W,V-U,V+U).$$

La transformación inversa es $(U,V,W)=(Z-Y,Z+Y,2X)$, y es fácil ver que (tras deshomogeneizar respecto de Z) la ecuación se transforma en

$$3Y^2=4aX^3-1.$$

Para pasar a una ecuación de Weierstrass basta multiplicar por $2^4\cdot 3^2\cdot a^2$ y hacer el cambio $X'=12aX$, $Y'=36aY$, con lo que obtenemos la ecuación

$$Y^2=X^3-432a^2.$$

Si componemos las dos transformaciones que hemos hecho vemos que el isomorfismo entre las dos curvas viene dado por

$$U=\frac{36a-Y}{6X}, \qquad V=\frac{36a+Y}{6X}.$$

En particular, la curva de Fermat es isomorfa (sobre $\mathbb{Q}$) a la curva dada por la ecuación $Y^2=X^3-432$. Sus puntos triviales son $(12,\pm 36)$ y O. ∎

Ejemplo: Curvas de Selmer Se llaman *curvas de Selmer* las curvas dadas por una ecuación homogénea de la forma

$$aU^3 + bV^3 + cW^3 = 0, \qquad abc \neq 0.$$

Evidentemente toda curva de Selmer es regular, luego es una curva elíptica sobre cualquier cuerpo sobre el que tenga un punto racional. Sin embargo, la existencia de puntos racionales en estas curvas no es trivial. En la introducción hemos anunciado que la curva $E/\mathbb{Q}$ dada por

$$3U^3 + 4V^3 + 5W^3 = 0 \tag{2.2}$$

cumple $E(\mathbb{Q}) = \varnothing$, luego no es una curva elíptica sobre $\mathbb{Q}$.

Es fácil decidir si una curva de Selmer contiene puntos racionales con alguna coordenada nula. Por ejemplo, es claro que (2.2) no contiene ninguno. Vamos a probar que una condición necesaria para que una curva de Selmer pueda contener un punto racional (u, v, w) con coordenadas no nulas es que la curva de Selmer

$$R^3 + S^3 + abcT^3 = 0 \tag{2.3}$$

contenga un punto racional con $T \neq 0$. (Trabajamos con $k = \mathbb{Q}$). Sea ρ una raíz cúbica primitiva de la unidad y sea σ el automorfismo no trivial de $\mathbb{Q}(\rho)/\mathbb{Q}$. Definimos

$$\lambda = au^3 + \rho bv^3 + \rho^2 cw^3, \quad \mu = au^3 + \rho^2 bv^3 + \rho cw^3.$$

De este modo $\sigma(\lambda) = \mu$. Teniendo en cuenta que

$$X^3 + Y^3 = (X + Y)(X + \rho Y)(X + \rho^2 Y) = (X + Y)(\rho + \rho^2 Y)(\rho^2 X + \rho Y),$$

obtenemos que $\lambda^3 + \mu^3 = (3au^3)(3bv^3)(3cv^3)$, lo que a su vez implica que los puntos

$$P = (\lambda, \rho\mu, \nu), \qquad P' = (\mu, \rho^2\lambda, \nu),$$

con $\nu = -3uvw$, cumplen la ecuación $R^3 + S^3 + abcT^3 = 0$. Notemos que $P^\sigma = P'$ y que $P \neq P'$. La recta que pasa por P y P' ha de cortar a la cúbica en un tercer punto (R, S, T), que será invariante por σ, luego racional. Hemos de probar que la tercera coordenada de dicho punto es no nula. Si lo fuera, el punto tendría que ser $(1, -1, 0)$. La ecuación de la recta es

$$\begin{vmatrix} 1 & -1 & 0 \\ \lambda & \rho\mu & \nu \\ \mu & \rho^2\lambda & \nu \end{vmatrix} = 0,$$

que se reduce a $\lambda + \rho\mu - \mu - \rho^2\lambda = 0$, o también

$$\lambda - \mu + \rho(\lambda + \mu) = 0.$$

Esto implica $\lambda = \mu = 0$, lo cual es imposible (lleva a $au^3 = bv^3 = cw^3$ y la ecuación original implica $a = 0$ o $u = 0$).

Así pues, si una curva de Selmer tiene un punto racional no trivial (con coordenadas no nulas), la curva (2.3) tiene un punto racional con $T \neq 0$ o, equivalentemente, la cúbica $R^3 + S^3 = -abc$ tiene un punto racional finito. Ahora bien, esta cúbica es del tipo estudiado en el ejemplo anterior, en virtud del cual podemos concluir que la curva elíptica

$$Y^2 = X^3 - 432 \cdot a^2b^2c^2$$

tiene un punto racional distinto de O. En particular, para probar que la curva (2.3) no tiene puntos racionales basta demostrar que la curva

$$Y^2 = X^3 - 432 \cdot 60^2$$

no tiene puntos racionales distintos de O. ∎

Volviendo a la teoría general, bajo ciertas hipótesis podemos simplificar aún más la ecuación que define a una curva elíptica. Por ejemplo, si la característica de k es distinta de 2, el cambio $Y = Y' - \frac{a_1}{2}X' - \frac{a_3}{2}$ transforma una ecuación de Weierstrass general en una ecuación de la forma

$$Y^2 = X^3 + \frac{b_2}{4}X^2 + \frac{b_4}{2}X + \frac{b_6}{4}, \tag{2.4}$$

donde

$$b_2 = a_1^2 + 4a_2, \quad b_4 = 2a_4 + a_1a_3, \quad b_6 = a_3^2 + 4a_6.$$

Si además la característica de k es distinta de 3 el cambio $X = X' + \frac{b_2}{12}$ nos da la ecuación

$$Y^2 = X^3 - \frac{c_4}{48}X - \frac{c_6}{864}, \tag{2.5}$$

donde $c_4 = b_2^2 - 24b_4$, $c_6 = -b_2^3 + 36b_2b_4 - 216b_6$.

Finalmente, con el cambio $Y = Y'/2$ la ecuación adquiere la forma

$$Y^2 = 4X^3 - g_2X - g_3, \tag{2.6}$$

donde $g_2 = 108c_4$ y $g_3 = 216c_6$. Esta ecuación no es de Weierstrass a causa del coeficiente de X^3, pero es la expresión más conveniente cuando se estudian analíticamente las curvas elípticas sobre $\mathbb{C}$. (En realidad son las ecuaciones de esta forma las que consideró Weierstrass.) Debemos tener presente que una de las técnicas que vamos a emplear para estudiar curvas elípticas definidas sobre cuerpos como $\mathbb{Q}$ será reducir los coeficientes módulo diferentes primos p, incluso $p = 2$ o $p = 3$, por lo que no podemos dejar de estudiar las curvas elípticas sobre cuerpos de característica 2 y 3. En la práctica, esto sólo nos supondrá unos pocos cálculos molestos adicionales en unas pocas ocasiones (la mayoría en este capítulo).

Definición 2.4 A las ecuaciones de Weierstrass generales (2.1) las llamaremos ecuaciones de *tipo a,* a las de la forma (2.4) las llamaremos ecuaciones de *tipo b,* a las de la forma (2.5) las llamaremos ecuaciones de *tipo c* y a las de la forma (2.6) las llamaremos *ecuaciones de Weierstrass clásicas* (que no son ecuaciones de Weierstrass según nuestra definición y no las consideraremos como tales salvo que lo indiquemos explícitamente).

En estos términos, hemos probado que toda curva elíptica sobre un cuerpo de característica > 2 admite una ecuación de Weierstrass de tipo b (es decir, con $a_1 = a_3 = 0$), y si la característica es > 3 admite una ecuación de tipo c (con $a_2 = 0$). En la definición siguiente recogemos las cantidades que nos han ido apareciendo hasta ahora junto con algunas más:

Definición 2.5 Para cada ecuación de Weierstrass (2.1) definimos:

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, & c_4 &= b_2^2 - 24b_4,\\ b_4 &= 2a_4 + a_1a_3, & c_6 &= -b_2^3 + 36b_2b_4 - 216b_6,\\ b_6 &= a_3^2 + 4a_6, & \Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6,\\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2, & j &= c_4^3/\Delta. \end{aligned}$$

Alternativamente (en característica $\neq 2, 3$): $\Delta = (c_4^3 - c_6^2)/12^3$.

Observaciones Según hemos visto, las cantidades b_2, b_4 y b_6 son (salvo unos denominadores 4 y 2) los coeficientes de la ecuación de tipo b en que se puede transformar la ecuación dada si la característica del cuerpo es > 2. Sin embargo, hemos de tener presente que están definidas incluso en característica 2. (Los denominadores 4 y 2 no se incluyen en la definición de los b_i para que esto sea cierto.) También conviene observar que si la ecuación de partida está ya en la forma (2.4), al calcular b_2, b_4 y b_6 según las fórmulas anteriores obtenemos las cantidades de partida.

Las observaciones precedentes se aplican igualmente a c_4 y c_6 con los cambios obvios (ecuaciones de tipo c y característica > 3).

La cantidad b_8 se introduce como paso intermedio para definir Δ, el cual se conoce como *discriminante* de la ecuación. Pronto veremos que una ecuación de Weierstrass define una curva regular (y, por lo tanto, elíptica) si y sólo si su discriminante es no nulo. El discriminante tiene una interpretación natural para ecuaciones de tipo b. Recordemos que todo polinomio

$$F(X) = a_0(X - \alpha_1)\cdots(X - \alpha_n)$$

tiene asociado un *discriminante*

$$D = \prod_{1\le i<j\le n} (\alpha_i - \alpha_j)^2,$$

de modo que $F(X)$ tiene raíces simples si y sólo si su discriminante es no nulo. Podemos ver a D como un polinomio simétrico en las indeterminadas α_i, lo que se traduce en que D depende polinómicamente de los polinomios simétricos elementales, los cuales, evaluados en las raíces de F, son —salvo signo— los coeficientes de F, de donde se sigue que D depende polinómicamente de dichos coeficientes. La expresión explícita de D en términos de los coeficientes se obtiene por un cálculo rutinario. En el caso de $F(X) = X^3 + bX^2 + cX + d$ dicha expresión resulta ser

$$D = -4b^3d + b^2c^2 + 18bcd - 4c^3 - 27d^2.$$

Si aplicamos esta fórmula a la ecuación (2.4) y comparamos con la definición de Δ, veremos que $\Delta = 16D$. Así pues, el discriminante de una ecuación de Weierstrass de tipo b es —salvo un factor 16, que se introduce para que la definición sea válida en característica 2— el discriminante de su miembro derecho en el sentido algebraico usual. Si la aplicamos a (2.5) obtenemos la definición alternativa de Δ en términos de c_4 y c_6 (salvo un 16), que sólo es válida en característica > 3. (Naturalmente, dicha definición alternativa exige comprobar que ambas expresiones se corresponden con el mismo polinomio en los coeficientes a_i, lo que no es más que un cálculo rutinario.)

El *invariante* j sólo está definido para ecuaciones con discriminante no nulo. Según hemos comentado, vamos a ver que éstas son precisamente las que definen curvas elípticas, y entonces demostraremos que j depende únicamente de la curva, en el sentido de que todas las ecuaciones de una misma curva tienen el mismo invariante y dos curvas son isomorfas si y sólo si tienen el mismo invariante.

En principio no tenemos definidos Δ y j para una ecuación clásica (2.6), pero podemos definirlos como los correspondientes a la ecuación que resulta de hacer el cambio $Y = 2Y'$. Entonces $a_4 = -g_2/4$, $a_6 = -g_3/4$ y

$$\Delta = g_2^3 - 27g_3^2, \qquad j = \frac{1728\, g_2^3}{g_2^3 - 27g_3^2}.$$

■

El teorema siguiente explica por fin la elección de los subíndices:

Teorema 2.6 *Si aplicamos a una ecuación de Weierstrass un cambio de variables del tipo descrito en el teorema* 2.3, *sus coeficientes y las cantidades que acabamos de definir se transforman según las fórmulas siguientes:*

$$\begin{aligned}
ua_1' &= a_1 + 2s,\\
u^2a_2' &= a_2 - sa_1 + 3r - s^2,\\
u^3a_3' &= a_3 + ra_1 + 2t,\\
u^4a_4' &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st,\\
u^6a_6' &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1,\\
\hline
u^2b_2' &= b_2 + 12r,\\
u^4b_4' &= b_4 + rb_2 + 6r^2,\\
u^6b_6' &= b_6 + 2rb_4 + r^2b_2 + 4r^3,\\
u^8b_8' &= b_8 + 3rb_6 + 3r^2b_4 + r^3b_2 + 3r^4,\\
\hline
u^4c_4' &= c_4,\\
u^6c_6' &= c_6,\\
u^{12}\Delta' &= \Delta,\\
j' &= j.
\end{aligned}$$

Demostración: Se trata de una comprobación rutinaria. ■

Con la ayuda de este teorema podemos encontrar ecuaciones canónicas lo más simples posibles para curvas definidas sobre cuerpos de característica 2 o 3.

En la práctica sólo nos van a interesar las cantidades j, Δ y c_4, así que vamos a calcularlas explícitamente sobre las ecuaciones canónicas:

Teorema 2.7 *Sea E/k una curva (no necesariamente regular) definida mediante una ecuación de Weierstrass. Entonces, bajo las hipótesis indicadas, existe un cambio de variables del tipo descrito en el teorema* 2.3 *que la transforma en otra de la forma indicada en la tabla siguiente:*

$\operatorname{car} k \neq 2,3$	$Y^2 = X^3 + a_4X + a_6$	
$\Delta = -16(4a_4^3 + 27a_6^2)$	$j = 1728\,\dfrac{4a_4^3}{4a_4^3 + 27a_6^2}$	$c_4 = -48a_4$
$\operatorname{car} k = 3,\ c_4 \neq 0$	$Y^2 = X^3 + a_2X^2 + a_6$	
$\Delta = -a_2^3a_6$	$j = -a_2^3/a_6$	$c_4 = a_2^2$
$\operatorname{car} k = 3,\ c_4 = 0$	$Y^2 = X^3 + a_4X + a_6$	
$\Delta = -a_4^3$	$j = 0$	$c_4 = 0$
$\operatorname{car} k = 2,\ c_4 \neq 0$	$Y^2 + XY = X^3 + a_2X^2 + a_6$	
$\Delta = a_6$	$j = 1/a_6$	$c_4 = 1$
$\operatorname{car} k = 2,\ c_4 = 0$	$Y^2 + a_3Y = X^3 + a_4X + a_6$	
$\Delta = a_3^4$	$j = 0,$	$c_4 = 0$

Demostración: El caso en que la característica es distinta de 2 y 3 lo tenemos ya demostrado. Si la característica es 3, hemos visto que podemos llegar a una ecuación de tipo b, que podemos reescribir como

$$Y^2 = X^3 + a_2X^2 + a_4X + a_6,$$

con $c_4 = a_2^2$. Si $c_4 = 0$, entonces $a_2 = 0$ y la ecuación tiene ya la forma indicada. En caso contrario el cambio $X = X' + a_4/a_2$ nos da la ecuación que buscamos.

Partiendo de una ecuación de Weierstrass general, en característica 2 se cumple que $c_4 = a_1^4$, luego si $c_4 = 0$ tenemos $a_1 = 0$ y basta hacer el cambio $X = X' + a_2$. Si $c_4 \neq 0$ entonces $a_1 \neq 0$ y el cambio oportuno es

$$X = a_1^2X' + a_3/a_1, \qquad Y = a_1^3Y' + (a_1^2a_4 + a_3^2)/a_1^3.$$

∎

Conviene observar que cuando j está definido ($\Delta \neq 0$) la distinción $c_4 = 0$ o $c_4 \neq 0$ equivale a $j = 0$ o $j \neq 0$.

Ahora ya podemos caracterizar la regularidad de una ecuación de Weierstrass en términos de su discriminante, tal y como habíamos anunciado:

Teorema 2.8 *Sea C/k una curva definida por una ecuación de Weierstrass. Entonces C es regular si y sólo si $\Delta \neq 0$. En caso contrario C tiene un único punto singular, que es finito y racional.*

DEMOSTRACIÓN: Observemos en primer lugar que el punto infinito O nunca es singular. Para ello homogeneizamos la ecuación de Weierstrass:

$$F(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3,$$

y observamos que

$$\left.\frac{\partial F}{\partial Z}\right|_{(0,1,0)} = 1 \neq 0.$$

El teorema 2.6 muestra que la condición $\Delta \neq 0$ no se altera por cambios de variable (definidos sobre k) y obviamente lo mismo vale para la regularidad de la curva y la racionalidad del punto singular si es que existe. Por lo tanto, no perdemos generalidad si suponemos que la curva viene dada por una de las ecuaciones canónicas descritas en el teorema 2.7.

Supongamos primero que $\operatorname{car} k \neq 2, 3$. Entonces la ecuación es de la forma

$$Y^2 = X^3 + a_4X + a_6,$$

y un punto singular ha de cumplir, además de esta ecuación, las condiciones

$$2Y = 0, \qquad 3X^2 + a_4 = 0.$$

En definitiva, un punto singular ha de ser de la forma $(x_0, 0)$, donde x_0 es una raíz del polinomio $X^3 + a_4X + a_6$ y de su derivada, es decir, una raíz múltiple del polinomio. En otras palabras, la curva es singular si y sólo si este polinomio tiene raíces múltiples. Ahora bien, para una ecuación de tipo c, hemos comprobado que esto equivale a que el discriminante sea nulo. Además, en tal caso, es fácil ver que el punto singular es $(X,Y) = (-3a_6/2a_4, 0)$, luego es único y racional.

En el segundo caso (en particular $\operatorname{car} k = 3$), un punto singular ha de cumplir

$$Y^2 = X^3 + a_2X^2 + a_6, \quad 2a_2X = 0, \quad 2Y = 0.$$

Si $a_2 = 0$ tenemos $\Delta = 0$ y $(X,Y) = (-\sqrt[3]{a_6}, 0)$ es la única solución de estas ecuaciones (es racional porque k es perfecto de característica 3).

Si $a_2 \neq 0$ las dos últimas ecuaciones se cumplen únicamente en $(0,0)$, y la primera se cumple si y sólo si $a_6 = 0$, es decir, si y sólo si $\Delta = 0$.

En el tercer caso (también con $\operatorname{car} k = 3$) las condiciones son

$$Y^2 = X^3 + a_4X + a_6, \quad a_4 = 0, \quad 2Y = 0.$$

Ciertamente hay solución si y sólo si $a_4 = 0$ (si y sólo si $\Delta = 0$), y dicha solución es de nuevo $(X,Y) = (-\sqrt[3]{a_6}, 0)$.

En el cuarto caso (ahora con $\operatorname{car} k = 2$) las condiciones son

$$Y^2 + XY = X^3 + a_2X^2 + a_6, \quad Y = 3X^2, \quad X = 0,$$

que se cumplen únicamente en $(0,0)$ cuando $\Delta = a_6 = 0$. El el último caso tenemos

$$Y^2 + a_3Y = X^3 + a_4X + a_6, \quad 3X^2 + a_4 = 0, \quad a_3 = 0,$$

y esto se cumple únicamente en $(X,Y) = (\sqrt{-a_4}, \sqrt{a_6})$ cuando $\Delta = 0$. ∎

Así pues, si E/k es una curva elíptica y C/k es una curva en las condiciones del teorema 2.3, se cumple que $\Delta_C \neq 0$, por lo que el invariante j_C está definido y, según 2.6, no depende de la curva C con que lo calculamos. Por ello podemos hablar del invariante $j(E) \in k$ de cualquier curva elíptica definida sobre k.

Teorema 2.9 *Dos curvas elípticas son isomorfas (sobre $\overline{k}$) si y sólo si tienen el mismo invariante.*

Demostración: Es claro que dos curvas isomorfas tienen el mismo invariante (pues podemos representarlas por la misma ecuación de Weierstrass). Supongamos ahora que E y E' son dos curvas con el mismo invariante. Supongamos primeramente que $\operatorname{car} k \neq 2, 3$. Entonces las curvas admiten ecuaciones

$$Y^2 = X^3 + a_4X + a_6, \qquad Y'^2 = X'^3 + a'_4X' + a'_6,$$

con

$$\frac{4a_4^3}{4a_4^3 + 27a_6^2} = \frac{4a_4'^3}{4a_4'^3 + 27a_6'^2}.$$

Distinguimos tres casos:

Si $a_4 = 0$, entonces $a'_4 = 0$ y $a_6 \neq 0 \neq a'_6$ (porque $\Delta \neq 0$). Tomamos un $u \in \overline{k}$ tal que $u^6 = a_6/a'_6$ y entonces el cambio $X = u^2X'$, $Y = u^3Y'$ transforma una curva en otra.

Si $a_6 = 0$ entonces $a'_6 = 0$, luego $a_4 \neq 0 \neq a'_4$ y tomamos $u \in \overline{k}$ tal que $u^4 = a_4/a'_4$. De nuevo, el cambio $X = u^2X'$, $Y = u^3Y'$ hace corresponder las curvas.

Si $a_4 \neq 0 \neq a_6$, entonces también $a'_4 \neq 0 \neq a'_6$ (por los casos anteriores). De la igualdad de los invariantes se sigue que $a_4^3a_6'^2 = a_4'^3a_6^2$ o, equivalentemente, $(a_4/a'_4)^3 = (a_6/a'_6)^2$. Tomamos $u \in \overline{k}$ tal que $a^4 = a_4/a'_4$, de manera que $u^{12} = (a_6/a'_6)^2$ y $u^6 = \pm a_6/a'_6$, si el signo es negativo, multiplicamos u por una raíz cuarta primitiva de la unidad, con lo que u^4 sigue siendo el mismo y u^6 cambia de signo. En definitiva, tenemos que $u^4 = a_4/a'_4$ y $u^6 = a_6/a'_6$. El cambio $X = u^2X'$, $Y = u^3Y'$ hace corresponder las ecuaciones.

Supongamos ahora que $\operatorname{car} k \neq 3$ y $j \neq 0$. Podemos considerar ecuaciones de la forma

$$Y^2 = X^3 + a_2X^2 + a_6,$$

con $a_2 \neq 0 \neq a'_2$, $a_6 \neq 0 \neq a'_6$, $a_2^3a'_6 = a_2'^3a_6$. Basta tomar $u^2 = a_2/a'_2$.

El caso siguiente es $\operatorname{car} k = 3$, $j = 0$, con lo que las ecuaciones canónicas son de la forma

$$Y^2 = X^3 + a_4X + a_6,$$

con $a_4 \neq 0 \neq a_4'$. Esta vez hemos de considerar un cambio de coordenadas de la forma $X = u^2X' + r$, $Y = u^3Y'$. Basta tomar u y r de modo que

$$u^4 = a_4'/a_4, \qquad r^3 + a_4r + a_6 - u^6a_6' = 0.$$

Supongamos ahora que $\operatorname{car} k = 2$ y $j \neq 0$. En este caso

$$Y^2 + XY = X^3 + a_2X^2 + a_6$$

con $a_6 = a_6' \neq 0$. Consideramos el cambio $X = X'$, $Y = Y' + sX'$, donde s es una raíz de la ecuación $s^2 + s + a_2 + a_2' = 0$.

Finalmente, si $\operatorname{car} k = 2$, $j = 0$, tenemos

$$Y^2 + a_3Y = X^3 + a_4X + a_6$$

con $a_3 \neq 0 \neq a_3'$. Consideramos un cambio $X = u^2X'+s^2$, $Y = u^3Y'+u^2sX'+t$, de modo que

$$u^3 = a_3/a_3', \quad s^4 + a_3s + a_4 - u^4a_4' = 0,$$
$$t^2 + a_3t + s^6 + a_4s^2 + a_6 - u^6a_6' = 0.$$

■

Este teorema junto con el que demostramos a continuación muestra que hay tantas clases de curvas elípticas como elementos tiene $\overline{k}$.

Teorema 2.10 *Para cada $j_0 \in k$ existe una curva elíptica definida sobre k con invariante j_0.*

Demostración: Si $j_0 \neq 0, 1728$ consideramos la curva

$$Y^2 + XY = X^3 - \frac{36}{j_0 - 1728}\,X - \frac{1}{j_0 - 1728}.$$

Un simple cálculo nos da que $\Delta = j_0^2/(j_0 - 1728)^3$ y $j = j_0$. Para los dos casos que faltan, consideramos las curvas

$$Y^2 + Y = X^3, \quad \Delta = -27, \quad j = 0,$$
$$Y^2 = X^3 + X, \quad \Delta = -64, \quad j = 1728.$$

Notemos que si $\operatorname{car} k = 2, 3$, entonces $1728 = 0$, y una de las dos curvas tiene discriminante no nulo. ■

Definición 2.11 Si E es una curva elíptica, llamaremos $\operatorname{Aut}(E)$ al grupo de los *automorfismos* de E, es decir, los isomorfismos de E en E que fijan a O.

Vamos a usar las ecuaciones de Weierstrass para determinar el grupo de automorfismos de una curva elíptica.

Teorema 2.12 *Si E es una curva elíptica de invariante j sobre un cuerpo de característica p (nula o prima), entonces* $\mathrm{Aut}(E)$ *es un grupo finito y su orden es*

$$|\mathrm{Aut}(E)| = \begin{cases} 2 & \text{si } j \neq 0,\ 1728, \\ 4 & \text{si } j = 1728 \text{ y } p \neq 2,\ 3, \\ 6 & \text{si } j = 0 \quad \text{y } p \neq 2,\ 3, \\ 12 & \text{si } j = 0 = 1728 \text{ y } p = 3, \\ 24 & \text{si } j = 0 = 1728 \text{ y } p = 2. \end{cases}$$

En los tres primeros casos $\mathrm{Aut}(E)$ *es un grupo cíclico.*

Demostración: No perdemos generalidad si suponemos que E viene dada por una ecuación de Weierstrass según el teorema 2.7.

Un automorfismo $\phi : E \longrightarrow E$ está completamente determinado por las funciones $x' = \phi \circ x$, $y' = \phi \circ y$. Dichas funciones han de satisfacer la misma ecuación de Weierstrass que x, y. Por el mero hecho de satisfacer una ecuación de Weierstrass, el teorema 2.3 (más exactamente su demostración) nos da que la relación entre x, y, x', y' ha de ser

$$x = u^2x' + r, \qquad y = u^3y' + su^2x' + t, \qquad u, r, s, t \in k, \quad u \neq 0,$$

pero además u, r, s, t han de cumplir lo necesario para que la nueva ecuación sea la misma. Equivalentemente, lo que tenemos es que ϕ^{-1} es la restricción a E de la transformación afín determinada por las ecuaciones

$$X = u^2X' + r, \qquad Y = u^3Y' + su^2X' + t.$$

Distingamos los cinco casos correspondientes al teorema 2.7. Si $p \neq 2$, 3, la ecuación es de la forma

$$Y^2 = X^3 + a_4X + a_6.$$

Ahora usamos el teorema 2.6. La ecuación de a_1 nos da $s = 0$, la de a_2 nos da $r = 0$ y la de a_3 nos da $t = 0$. Por consiguiente, el isomorfismo se reduce a $X = u^2X'$, $Y = u^3Y'$, y la nueva ecuación pasa a ser

$$Y^2 = X^3 + u^{-4}a_4 + u^{-6}a_6.$$

Si $a_4 \neq 0$ y $a_6 \neq 0$ (lo cual equivale a que $j \neq 0$, 1728), entonces se ha de cumplir que $u^{-4} = 1$ y $u^{-6} = 1$, lo que equivale a que $u^2 = 1$. Por consiguiente tenemos sólo dos automorfismos, dados por $Y = \pm Y'$.

Si $a_6 = 0$, entonces $j = 1728$ y la condición se reduce a que $u^{-4} = 1$, lo que da lugar a cuatro automorfismos distintos. Observemos que la aplicación que a cada raíz cuarta de la unidad u le asigna su automorfismo correspondiente es un isomorfismo de grupos, luego $\mathrm{Aut}(E)$ resulta ser un grupo cíclico.

Si $a_4 = 0$ entonces $j = 0$ y razonamos igualmente, sólo que ahora u es una raíz sexta de la unidad.

Supongamos ahora $p = 3$, $j \neq 0$. Entonces la ecuación es

$$Y^2 = X^3 + a_2X^2 + a_6.$$

Las transformaciones de a_1, a_3 y a_4 nos dan de nuevo que $r = s = t = 0$ y la ecuación se convierte en

$$Y^2 = X^3 + u^{-2}a_2X^2 + u^{-6}a_6,$$

con lo que ha de ser $u^2 = 1$ y tenemos únicamente dos automorfismos.

Si $j = 0$ podemos tomar

$$Y^2 = X^3 + a_4X,$$

pero ahora sólo podemos concluir $s = t = 0$, y la ecuación se transforma en

$$Y^2 = X^3 + u^{-4}a_4X + u^{-6}(a_4r + r^3).$$

Vemos, pues, que $u^4 = 1$, y que $r^3 + a_4r = 0$, lo que da tres valores distintos para r, a saber, 0 y $\pm\sqrt{-a_4}$. Por lo tanto hay 12 automorfismos. Si representamos por (u, r) el automorfismo correspondiente a unos valores dados de u y r, es fácil ver que $g = (\sqrt[4]{-1}, 0)$ tiene orden 4, $a = (1, \sqrt{-a_4})$ tiene orden 3 y que no conmutan, por lo que $a^g = a^{-1}$, de donde g^2 conmuta con a y $h = g^2a$ tiene orden 6. En definitiva,

$$\mathrm{Aut}(E) = \left\langle g, h \mid h^6 = 1,\ g^2 = h^3,\ h^g = h^{-1} \right\rangle.$$

Si $p = 2$ y $j \neq 0$ podemos tomar

$$Y^2 + XY = X^3 + a_6,$$

con lo que ha de ser $u = 1$ (de la ecuación de a_1), $r = 0$ (de la ecuación de a_3), $t = 0$ (de la ecuación de a_4) y $s^2 + s = 0$ (de la ecuación de a_2). En definitiva, hay dos automorfismos, determinados por el valor de $s = 0, 1$.

Por último, si $p = 2$ y $j = 0$ podemos tomar la ecuación

$$Y^2 + a_3Y = X^3.$$

De la ecuación de a_2 sale $r = s^2$ y la ecuación transformada es

$$Y^2 + u^{-3}a_3Y = X^3 + (sa_3 + s^4)X + u^{-6}(t^2 + a_3t + s^6).$$

Se ha de cumplir $u^3 = 1$, $s^4 + a_3s = 0$, $t^2 + a_3t + s^6 = 0$. Tenemos tres posibilidades para u, cuatro para s y, para cada valor de s, la tercera ecuación tiene dos raíces distintas (o sería $a_3 = 0$). En total $\mathrm{Aut}(E)$ consta de 24 automorfismos. Un análisis más detallado muestra que está formado por un subgrupo de orden 8 de tipo cuaternio sobre el que actúa por conjugación un grupo cíclico de orden 3 (que permuta los generadores i, j, k). ∎

El teorema de Riemann-Roch implica que el espacio de las diferenciales de primera clase (diferenciales holomorfas, o sin polos) tiene dimensión 1 sobre $\overline{k}$. Además, los divisores de las formas diferenciales recorren la clase canónica, que en un cuerpo elíptico es la clase principal. Esto significa que las diferenciales de primera clase no tienen ni ceros ni polos. Vamos a calcular explícitamente una diferencial de primera clase para una curva elíptica determinada por una ecuación de Weierstrass:

Definición 2.13 A cada curva C/k definida por una ecuación

$$F(X,Y) = Y^2 + a_1XY + a_3Y - X^3 - a_2X^2 - a_4X - a_6 = 0$$

le asociamos la forma diferencial

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}.$$

En otros términos,

$$\omega = \frac{dx}{F_Y} = -\frac{dy}{F_X}.$$

La igualdad se debe a que las funciones x, $y \in k(C)$ cumplen $F(x,y) = 0$, luego diferenciando queda $F_X\,dx + F_Y\,dy = 0$.

Una comprobación rutinaria muestra que ante un cambio de coordenadas como el del teorema 2.6, la forma que acabamos de definir se transforma según la relación $u^{-1}\omega' = \omega$.

Teorema 2.14 *Si C/k es una curva elíptica definida por una ecuación de Weierstrass, entonces la forma ω es una diferencial de primera clase que no se anula en ningún punto.*

Demostración: Si $P = (x_0, y_0) \in C$ es un punto finito, entonces la tangente a C por P es la recta de ecuación

$$F_X(P)(X - x_0) + F_Y(P)(Y - y_0) = 0.$$

Si $F_X(P) \neq 0$ entonces $Y - y_0 = 0$ no es la recta tangente a C en P, luego $y - y_0$ es un parámetro local de C en P (cualquier recta no tangente induce un parámetro local), y entonces $v_P(\omega) = v_P(F_X^{-1}\,d(y - y_0)) = -v_P(F_X(x,y)) = 0$, pues la función $F_X(x,y)$ es regular y no nula en P.

Similarmente, si $F_Y(P) \neq 0$ entonces $x - x_0$ es un parámetro local de C en P y $v_P(\omega) = -v_P(-F_Y(x,y)) = 0$. Con esto hemos probado que ω no tiene ni ceros ni polos en ningún punto finito de C. Falta, pues, estudiar, el comportamiento de ω en el punto infinito O.

Sea t un parámetro local en O. Sabemos que $x = ft^{-2}$, $y = gt^{-3}$, donde f y g son funciones regulares y no nulas en O. Entonces

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{f't^{-2} - 2t^{-3}f}{2gt^{-3} + a_1ft^{-2} + a_3}\,dt = \frac{f't - 2f}{2g + a_1ft + a_3t^3}\,dt.$$

Si $\operatorname{car} k \neq 0$, es claro que $v_O(\omega) = 0$. En caso contrario llegamos a la misma conclusión razonando igualmente con la expresión de ω en términos de dy. ∎

Terminamos la sección estudiando una clase especial de ecuaciones canónicas de Weierstrass:

Definición 2.15 Una ecuación de Weierstrass está en *forma de Legendre* si puede escribirse como

$$Y^2 = X(X-1)(X-\lambda), \qquad \lambda \in \overline{k}.$$

Llamaremos E_λ a la curva definida por esta ecuación. Un simple cálculo muestra que su discriminante es $\Delta = 16\lambda^2(\lambda-1)^2$, luego es singular cuando $\operatorname{car} k = 2$ o cuando $\lambda = 0, 1$. En los casos restantes se comprueba que

$$j(E_\lambda) = 2^8 \frac{(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}.$$

Teorema 2.16 *Si* $\operatorname{car} k \neq 2$*, toda curva elíptica es isomorfa (sobre* $\overline{k}$*) a una curva* E_λ*, para cierto* $\lambda \in \overline{k}$*,* $\lambda \neq 0, 1$*. La aplicación* $\lambda \mapsto j(E_\lambda)$ *es suprayectiva y toma seis veces cada valor de* $\overline{k}$*, excepto* 0 *(que lo toma sólo dos veces cuando* $\operatorname{car} k \neq 3$ *y una en caso contrario) y* 1728 *(que lo toma sólo tres veces cuando* $\operatorname{car} k \neq 3$ *y una en caso contrario).*

Demostración: Basta tomar como λ una raíz de la ecuación

$$2^8(\lambda^2-\lambda+1)^3 - j\lambda^2(\lambda-1)^2 = 0,$$

donde j es el invariante de la curva dada. Es claro que 0 y 1 no son raíces, luego $\lambda \neq 0, 1$. Como $j(E_\lambda) = j$, tenemos que E_λ es isomorfa a la curva dada.

Si $j(E_\lambda) = j(E_\mu)$ entonces las dos curvas son isomorfas, y es posible transformar sus ecuaciones mediante un cambio de coordenadas afín. Ahora bien, para que un cambio de este tipo transforme una ecuación de Legendre en otra, ha de ser más concretamente de la forma

$$X = u^2X' + r, \qquad Y = u^3Y'.$$

Igualando las ecuaciones tenemos:

$$X(X-1)(X-\mu) = \left(X + \frac{r}{u^2}\right)\left(X + \frac{r-1}{u^2}\right)\left(X + \frac{r-\lambda}{u^2}\right),$$

y hay exactamente seis formas de hacer corresponder los factores. Por ejemplo, si los hacemos corresponder en el orden en que los hemos escrito, queda que $r = 0$, $u^2 = 1$, $\mu = \lambda$. Considerando similarmente las otras posibilidades llegamos a que

$$\mu \in \left\{\lambda, \frac{1}{\lambda}, 1-\lambda, \frac{1}{1-\lambda}, \frac{\lambda}{\lambda-1}, \frac{\lambda-1}{\lambda}\right\}.$$

Así pues, $j(E_\mu)$ toma seis veces distintas el valor $j(E_\lambda)$, salvo que algunos de los valores del conjunto anterior coincidan. Igualándolos dos a dos vemos que

esto sucede cuando $\lambda = -1, 2, 1/2$ (en cuyo caso μ toma tres valores distintos salvo si $\operatorname{car} k = 3$) y cuando $\lambda^2 - \lambda + 1 = 0$ (en cuyo caso μ toma dos valores salvo si $\operatorname{car} k = 3$). Es fácil ver que $j(E_\lambda)$ es 1728 en el primer caso y 0 en el segundo. ∎

Para cuerpos con característica 2, hay una clase de ecuaciones que desempeña un papel similar:

Definición 2.17 Una ecuación de Weierstrass está en *forma de Deuring* si puede escribirse como

$$Y^2 + \alpha XY + Y = X^3, \qquad \alpha \in \overline{k}.$$

Llamaremos D_α a la curva definida por esta ecuación. Un simple cálculo muestra que

$$\Delta = \alpha^3 - 27, \qquad j = \frac{\alpha^3(\alpha^3 - 24)^3}{\alpha^3 - 27}.$$

Así pues, D_α es elíptica cuando $\alpha^3 \neq 27$. Si $\operatorname{car} k = 3$ esto equivale a $\alpha \neq 0$, y en tal caso $j(D_\alpha) = \alpha^9$. Vemos, pues, que en característica 3 toda curva elíptica con invariante $j \neq 0$ admite una ecuación en forma de Deuring. En las demás características no hay excepciones:

Teorema 2.18 *Si $\operatorname{car} k \neq 3$, toda curva elíptica es isomorfa (sobre $\overline{k}$) a una curva D_α, para cierto $\alpha \in \overline{k}$, $\alpha^3 \neq 27$.*

Demostración: Basta tomar como α una raíz de la ecuación

$$\alpha^3(\alpha^3 - 24)^3 - (\alpha^3 - 27)j = 0.$$

No puede ser $\alpha^3 = 27$, pues entonces sería $27^2 = 0$, absurdo. ∎

2.2 La estructura de grupo

Si E es una curva elíptica, podemos considerar la aplicación $E \longrightarrow H_0(E)$ dada por $P \mapsto [P/O]$. Esta aplicación es inyectiva, pues si $[P/O] = [Q/O]$ entonces $[P] = [Q]$, luego existe una función $f \in \overline{k}(E)$ tal que $(f) = P/Q$. Por consiguiente $f \in m(Q^{-1})$, pero $\dim Q = \operatorname{grad} Q = 1$ y $m(Q^{-1})$ contiene a las constantes, lo que nos lleva a un absurdo.

La aplicación también es suprayectiva, pues si $\mathfrak{a}$ es un divisor de grado 0, el teorema de Riemann-Roch nos da que $\dim \mathfrak{a}O = \operatorname{grad} \mathfrak{a}O = 1$, luego existe una función $f \in m(\mathfrak{a}^{-1}O^{-1})$ no nula. Entonces $(f)\mathfrak{a}O$ es un divisor entero de grado 1, luego tiene que ser primo, es decir, un punto $P \in E$. Tenemos así que $[\mathfrak{a}O] = [P]$ o, equivalentemente, que $[\mathfrak{a}] = [P/O]$.

A través de esta biyección podemos trasladar la operación del grupo de clases a la curva E. Notemos antes que si E está definida sobre k, entonces $\overline{k}(E)$ es una extensión de constantes de $k(E)$, por lo que podemos considerar al grupo de

clases de grado 0 de $k(E)$ como un subgrupo de $H_0(E)$, y el argumento anterior muestra que la biyección que hemos definido se restringe a una biyección entre los divisores primos de grado 1 de $k(E)$ (es decir, los puntos de $E(k)$) y las clases de divisores de grado 0 de $k(E)$.

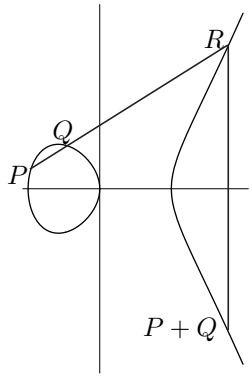
Definición 2.19 Si E es una curva elíptica, definimos en E la ley de composición interna determinada por $[(P+Q)/O] = [P/O][Q/O]$.

El razonamiento precedente muestra que E se convierte con esta operación en un grupo abeliano isomorfo a $H_0(E)$ a través del isomorfismo $P \mapsto [P/O]$. El elemento neutro de E es el punto O. Si E está definida sobre k, entonces $E(k)$ es un subgrupo de E isomorfo al grupo de clases de grado 0 de $k(E)$. Conviene observar que, en términos de la suma en E, el isomorfismo inverso de $E \cong H_0(E)$ viene dado por

$$[P_1^{m_1} \cdots P_r^{m_r}] \mapsto m_1 P_1 + \cdots + m_r P_r.$$

(Porque $[P_1^{m_1} \cdots P_r^{m_r}] = [(P_1/O)^{m_1} \cdots (P_r/O)^{m_r}]$.)

La suma que acabamos de definir tiene una interpretación geométrica simple:



Teorema 2.20 *Sea E una curva elíptica determinada por una ecuación de Weierstrass y sean P, $Q \in E$. Llamemos R al tercer punto donde la recta que pasa por P y Q corta a E. Entonces $P+Q$ es el tercer punto donde la recta que pasa por R y O corta a E.*

Demostración: Recordemos que, para una curva determinada por una ecuación de Weierstrass, las funciones coordenadas cumplen $x \in m(O^{-2})$, $y \in m(O^{-3})$. De hecho, 1, x, y forman una base de $m(O^{-3})$. Una recta proyectiva L es el conjunto de ceros de un polinomio $aX+bY+cZ$. Su ecuación afín es $aX+bY+c=0$.

La función $l = ax + by + c \in \overline{k}(E)$ cumple $l \in m(O^{-3})$, luego tenemos que $(l) = P_1P_2P_3/O^3$, para ciertos puntos $P_i \in E$ no necesariamente distintos entre sí ni distintos de O. Estos tres puntos son precisamente los puntos de corte de la recta L con E. Más precisamente, el número de intersección $I_P(E \cap L)$ es el número de veces que P aparece entre los P_i.

En efecto, recordemos que si F es una forma lineal que no se anula en P y $f = L/G \in \overline{k}(E)$, entonces $I_P(E \cap L) = v_P(f)$. Así, si $P \neq 0$, podemos tomar $F = Z$, con lo que $I_P(E \cap L) = v_P(l)$ es el número de veces que P aparece entre los P_i. Si $P = O$, tomamos $F = Y$, con lo que $f = l(Z/Y) = l/y$. Sea $(y) = Q_1Q_2Q_3/O^3$, donde Q_i son los puntos (finitos) donde la recta $Y = 0$ corta a E. Entonces $I_O(E \cap L) = v_O(P_1P_2P_3/Q_1Q_2Q_3)$ es también el número de veces que O aparece entre los P_i.

Observemos que el razonamiento vale incluso cuando la recta es $Z = 0$, en cuyo caso $l = 1 = O^3/O^3$, y obtenemos que $I_O(E \cap L) = 3$.

En las condiciones del enunciado, si llamamos L a la recta que pasa por P y Q, entonces $(l) = PQR/O^3$ y si L' es la recta que pasa por O y R entonces $(l') = ORS/O^3$, para cierto punto $S \in E$. Hemos de probar que $S = P + Q$. Ahora bien, esto es inmediato: $[P/O][Q/O] = [O/R] = [S/O]$, luego ciertamente $S = P + Q$. ∎

Observemos que el opuesto $-P$ de un punto finito $P \in E$ se calcula como el tercer punto de la recta (vertical) que une O con P. En efecto, si llamamos Q a este punto, el tercer punto de la recta que une P con Q es O, y el tercer punto de la recta que une O con O es O, luego $P + Q = O$.

Vamos a dar fórmulas explícitas para la suma en una curva elíptica E definida por un polinomio

$$F(X,Y) = Y^2 + a_1XY + a_3Y - X^3 - a_2X^2 - a_4X - a_6.$$

Sea $P = (x_0, y_0) \in E$ un punto finito. El punto $-P$ es el tercer punto (x_0, y_0') donde la recta $X = x_0$ corta a E (además de P y O). Tenemos que $F(x_0, Y) = c(Y - y_0)(Y - y_0')$. Comparando los coeficientes de Y^2 sale $c = 1$ y comparando los coeficientes de Y obtenemos que $-P = (x_0, -y_0 - a_1x_0 - a_3)$.

Consideremos ahora dos puntos finitos $P_1 = (x_1, y_1)$ y $P_2 = (x_2, y_2)$ en E. Si $x_1 = x_2$ y $y_1 + y_2 + a_1x_2 + a_3 = 0$, entonces $P_1 + P_2 = O$. Descartamos este caso y consideramos la recta que une P_1 con P_2 (la tangente a E por P_1 si son el mismo punto). Digamos que su ecuación es $Y = \lambda X + \mu$. (Los valores explícitos de λ y μ son fáciles de calcular y están dados en el enunciado del teorema siguiente). Llamemos P_3 al tercer punto en que esta recta corta a E. Se trata de $P_3 = -P_1 - P_2$, luego es finito. Para calcularlo hacemos

$$F(X, \lambda X + \mu) = c(X - x_1)(X - x_2)(X - x_3).$$

Igualando los coeficientes de X^3 queda que $c = -1$, y con X^2 obtenemos

$$x_1 + x_2 + x_3 = \lambda^2 + a_1\lambda - a_2.$$

Sustituyendo en la ecuación de la recta queda que $y_3 = \lambda x_3 + \mu$. Aplicando la fórmula para calcular el opuesto, obtenemos $P_1 + P_2 = -P_3$. En el teorema siguiente damos explícitamente las fórmulas resultantes:

Teorema 2.21 *Sea E una curva elíptica determinada por una ecuación de Weierstrass. Entonces*

a) Si $P_0 = (x_0, y_0) \in E$ es un punto finito, $-P_0 = (x_0, -y_0 - a_1x_0 - a_3)$.

b) Si $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ son puntos finitos de E tales que $x_1 = x_2$ y $y_1 + y_2 + a_1x_2 + a_3 = 0$, entonces $P_1 + P_2 = O$. En caso contrario, sean

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1}, \quad \mu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1}, \quad \text{si } x_1 \neq x_2,$$

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3},$$

$$\mu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}, \qquad \text{si } x_1 = x_2.$$

Entonces, $P_3 = P_1 + P_2$ *viene dado por*

$$\begin{aligned} x_3 &= \lambda^2 + a_1\lambda - a_2 - x_1 - x_2, \\ y_3 &= -(\lambda + a_1)x_3 - \mu - a_3. \end{aligned}$$

c) En particular se cumple la fórmula de duplicación:

$$x(2P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}.$$

(Para probar la fórmula de duplicación podemos partir de una ecuación de tipo b, pues el cambio de variables cumple $X' = X$.)

Estas fórmulas muestran explícitamente que la suma de puntos racionales es de nuevo un punto racional. También nos permiten demostrar el teorema siguiente:

Teorema 2.22 *Si* E/k *es una curva elíptica, entonces las aplicaciones*

$$+ : E \times E \longrightarrow E \quad y \quad - : E \longrightarrow E$$

son regulares y están definidas sobre k*.*

Demostración: No perdemos generalidad si suponemos que E está definida por una ecuación de Weierstrass. El teorema anterior muestra entonces que la restricción de la aplicación inversa a la parte afín de E es polinómica, luego es regular (y definida sobre k), luego determina una aplicación racional en E, pero toda aplicación racional entre curvas proyectivas regulares es regular. En realidad, puesto que es su propia inversa, vemos que es un isomorfismo.

El teorema anterior muestra también que la suma es regular (y definida sobre k) en todos los puntos de $E \times E$ salvo a lo sumo en los de la forma (P, P), $(P, -P)$, (P, O) y (O, P). Para ocuparnos de éstos consideramos las traslaciones $\tau_Q : E \longrightarrow E$ dadas por $\tau_Q(P) = P + Q$. El teorema anterior muestra que son regulares en un abierto (concretamente, para $P \neq O, \pm Q$), y por la regularidad de E son regulares en E. Además, la inversa de una traslación es otra traslación, luego son isomorfismos (pero no de curvas elípticas, pues no conservan el cero).

Dado un par $(P_1, P_2) \in E \times E$, podemos escoger traslaciones τ_1 y τ_2 de modo que $(\tau_1(P_1), \tau_2(P_2))$ esté en el abierto de $E \times E$ donde sabemos que la suma es regular. Ahora basta observar que, en un entorno de (P_1, P_2), la suma se descompone como

$$E \times E \xrightarrow{\tau_1 \times \tau_2} E \times E \xrightarrow{+} E \xrightarrow{\tau_1^{-1}} E \xrightarrow{\tau_2^{-1}} E.$$

Esto prueba que la suma es regular en $E \times E$. ∎

Ejemplo Consideremos la curva $Y^2 = X^3 - 25X$. En la introducción hemos comentado que los puntos $(-4, 6)$, $(-5, 0)$ y $(45, 300)$ están alineados, lo que equivale a que

$$(-4, 6) + (-5, 0) + (45, 300) = 0.$$

Similarmente, $2(-4, -6) = (20.172/1.728, 62.279/1.768)$. Estos resultados pueden comprobarse fácilmente con las fórmulas del teorema 2.21. Por ejemplo, la fórmula de duplicación es

$$x(2P) = \frac{x^4 + 50x^2 + 625}{4x^3 - 100x}.$$

Los puntos triviales $(-5, 0)$, $(0, 0)$ y $(5, 0)$ cumplen $2P = O$. ∎

Ahora podemos demostrar la implicación $c) \Rightarrow b)$ del Teorema 1 de la introducción (la caracterización de los números congruentes):

Demostración: Consideramos un natural n libre de cuadrados, suponemos que la curva elíptica $E/\mathbb{Q}$ dada por $Y^2 = X^3 - n^2X$ tiene un punto racional distinto de O, $(-n, 0)$, $(0, 0)$, $(n, 0)$, y hemos de probar que existen tres cuadrados racionales en progresión aritmética de razón n (lo que a su vez implica que n es congruente).

Es claro que para que un punto P de $E/\mathbb{Q}$ tenga orden 2 es necesario y suficiente que la tangente a la curva por P sea vertical, lo que a su vez equivale a que $Y = 0$. En definitiva, los puntos exceptuados son precisamente los puntos de orden 2 de $E/\mathbb{Q}$. Así pues, si la curva tiene otro punto racional P, se cumplirá que $P \neq O \neq 2P$. Pongamos que $P = (x, y)$ y $2P = (x', y')$.

Sea $Y = aX + b$ la tangente a la curva en P. Esta recta pasa por (x, y) (dos veces) y por $(x', -y')$. Esto significa que si sustituimos $Y = aX + b$ en la ecuación de E, el polinomio resultante

$$(X + n)X(X - n) - (aX + b)^2$$

tiene a x como raíz doble y a x' como tercera raíz. Teniendo en cuenta que es mónico, vemos que

$$(X + n)X(X - n) - (aX + b)^2 = (X - x)^2(X - x').$$

Ahora hacemos $X = -n$, con lo que

$$(b - an)^2 = (x + n)^2(x' + n).$$

Sabemos que $x + n \neq 0$ (pues $P \neq (-n, 0)$), luego concluimos que $x' + n$ es un cuadrado. Similarmente llegamos a que x' y $x' - n$ son cuadrados. ∎

A la vista de la demostración anterior, la forma más natural de llegar a que el número 5 es congruente consiste en observar que $(-4, -6)$ es un punto no trivial de $Y^2 = X^3 - 25X$, calcular $2P = (x', y')$ (ver el ejemplo anterior), obtener la progresión aritmética $x' - 5$, x', $x' + 5$ y a partir de aquí calcular la terna (a, b, c) según la demostración del Teorema 1 en la introducción.

Ejemplo *El número* 14 *es congruente.*

La curva elíptica $Y^2 = X^3 - 14^2X$ tiene el punto racional $P = (18, 48)$. (Es fácil encontrarlo con un ordenador.) Aplicando la fórmula de duplicación obtenemos que

$$x(2P) = \frac{4225}{144} = \left(\frac{65}{12}\right)^2,$$

de donde

$$x(2P) - 14 = \left(\frac{47}{12}\right)^2, \quad x(2P) + 14 = \left(\frac{79}{12}\right)^2.$$

Según la prueba del Teorema 1 de la introducción, un triángulo rectángulo racional de área 14 es $(a, b, c) = (21/2, 8/3, 65/6)$, que se obtiene haciendo

$$a = \frac{79}{12} + \frac{47}{12}, \quad b = \frac{79}{12} - \frac{47}{12}, \quad c = 2 \cdot \frac{65}{12}.$$

■

Ejercicio: Demostrar que el número 15 es congruente y encontrar una terna asociada partiendo del punto $P = (-9, 36)$.

Definición 2.23 Si E es una curva elíptica y $P \in E$, definimos la *traslación* por P como la aplicación $\tau_P : E \longrightarrow E$ dada por $\tau_P(Q) = P + Q$.

Claramente, las traslaciones son isomorfismos de curvas, pero, como ya hemos comentado, no son isomorfismos de curvas elípticas porque no cumplen $\tau_P(O) = O$ (salvo en el caso de τ_O, que es la identidad). Es claro que las traslaciones forman con la composición un grupo isomorfo a E.

Teorema 2.24 *Si E es una curva elíptica y ω es una diferencial de primera clase en E, entonces ω es invariante por traslaciones, es decir, se cumple que $\overline{\tau}_P(\omega) = \omega$ para todo $P \in E$.*

Demostración: Como τ_P es un isomorfismo, $\overline{\tau}_P : \Omega(E) \longrightarrow \Omega(E)$ es un $\overline{k}$-isomorfismo, luego $\overline{\tau}_P(\omega) \neq 0$. Por consiguiente existe una función no nula $f_P \in \overline{k}(E)$ tal que $\overline{\tau}_P(\omega) = f_P\omega$. Tomando divisores vemos que

$$(f_P) = (\overline{\tau}_P(\omega))/(\omega) = (\overline{\tau}_P(\omega))/(\omega) = 1/1 = 1,$$

pues las diferenciales de primera clase en un cuerpo de género 1 no tienen ceros ni polos. Esto implica que $f_P \in \overline{k}$ es una constante. Tenemos así definida una función $f : E \longrightarrow \overline{k}^*$ tal que $\overline{\tau}_P(\omega) = f(P)\omega$. De esta relación se sigue que $f(P + Q) = f(P)f(Q)$. En particular $f(O) = 1$. Vamos a ver que f es regular.

Sea $E^* = E \setminus \{O\}$. No perdemos generalidad si suponemos que E es una curva plana definida por una ecuación de Weierstrass. Entonces E^* es una curva afín. Consideremos el producto $E^* \times E^*$, con funciones coordenadas u, v, x, y. Pongamos que $\omega = \alpha\, dx$, con $\alpha \in \overline{k}(E)$.

La función que a cada par $(P,Q) \in E^* \times E^*$ le asigna $\overline{\tau}_P(x)(Q) = x(P+Q)$ es una función racional en $E^* \times E^*$, puesto que la suma es regular y x es racional. Podemos llamarla $\overline{\tau}(x)$. Esto significa que $\overline{\tau}_P(x)(Q) \in A^1$ se calcula (cuando está definido) mediante una función racional $R(U,V,X,Y)$ a partir de las coordenadas (u,v) de P y (x,y) de Q. Para un punto prefijado $P = (u_0, v_0)$, la función $\overline{\tau}_P(x) \in \overline{k}(E^*)$ se calcula mediante la función $R(u_0, v_0, X, Y)$ a partir de las coordenadas (x,y) de Q. Por lo tanto, la función

$$\frac{d\overline{\tau}_P(x)}{dx} \in \overline{k}(E^*)$$

se calcula mediante la derivada parcial de $R(U,V,X,Y)$ respecto de X a partir de las coordenadas (u_0,v_0) de P y las coordenadas (x_0,y_0) de Q. Vemos, pues, que la función $d\overline{\tau}(x)/dx : E^* \times E^* \longrightarrow A^1$ es racional.

El mismo razonamiento prueba que $d\overline{\tau}(\alpha)/dx \in \overline{k}(E^* \times E^*)$, luego llegamos a que

$$\frac{\overline{\tau}(\omega)}{\omega} = \frac{\overline{\tau}(\alpha)\,\frac{d\overline{\tau}(x)}{dx}}{\alpha} \in \overline{k}(E^* \times E^*).$$

Ahora bien, antes hemos probado que, para cada P, esta función es la constante $f(P)$, es decir, que no depende de las coordenadas x e y. En otros términos, fijamos un punto $Q \in E^*$ y consideramos la aplicación $E^* \longrightarrow E^* \times E^*$ dada por $P \mapsto (P,Q)$ (claramente regular), la composición de ésta y la precedente es la función f. Más precisamente, tenemos que f es regular en un abierto de E.

Ahora bien, la relación $f(P+Q) = f(P)f(Q)$ implica que f es regular en toda la curva E, pues, dado $P \in E$, sea $Q \in E$ tal que f es regular en un entorno de $P+Q$, entonces $f(P) = f(P+Q)/f(Q)$, y el segundo miembro es regular en un entorno de P.

Resulta así que $f : E \longrightarrow \mathrm{P}^1$ es una aplicación regular que no toma los valores 0 ni ∞, luego no es suprayectiva y, por consiguiente, es constante. Como $f(O) = 1$, ha de ser $f = 1$. Concluimos que $\overline{\tau}_P(\omega) = \omega$. ∎

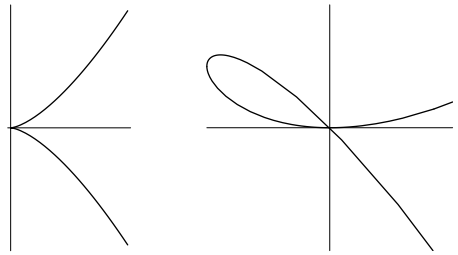
En virtud del teorema anterior, las diferenciales de primera clase en una curva elíptica se llaman también *diferenciales invariantes*. Notemos que una diferencial que no sea de primera clase no puede ser invariante por traslaciones, pues las traslaciones trasladan sus polos.

2.3 Cúbicas singulares

Si tenemos una curva elíptica definida por una ecuación de Weierstrass con coeficientes enteros, una forma de estudiarla es considerar la curva definida por dicha ecuación módulo un primo p. Cuando p divide al discriminante de la ecuación obtenemos una curva singular, por lo que conviene estudiar este tipo de curvas un poco más a fondo. En primer lugar observamos que hay dos tipos de singularidades:

Definición 2.25 Si P es un punto singular de una cúbica plana C, diremos que P es un *nodo* de C si P tiene dos tangentes distintas en P, mientras que P es una *cúspide* si P tiene una única tangente (doble) en P.

Notemos que la cúbica no puede tener tres tangentes, porque entonces sería reducible. Las figuras siguientes muestran cúbicas con una cúspide y un nodo en $(0,0)$.



$Y^2 = X^3$ $\qquad$ $Y^2 + XY - X^3 = 0$

En general, si C/k es una cúbica singular definida por una ecuación de Weierstrass, el teorema 2.8 nos da que su punto singular es finito y racional, digamos (x_0, y_0). Entonces, la traslación

$$X = X' + x_0, \qquad Y = Y' + y_0$$

transforma la ecuación de Weierstrass en otra cuyo punto singular es $(0,0)$. Entonces

$$F(0,0) = a_6 = 0, \qquad \left.\frac{\partial F}{\partial X}\right|_{(0,0)} = a_4 = 0, \qquad \left.\frac{\partial F}{\partial Y}\right|_{(0,0)} = a_3 = 0,$$

luego la ecuación se reduce a

$$Y^2 + a_1XY - a_2X^2 - X^3 = 0.$$

Las tangentes en $(0,0)$ se obtienen factorizando la forma $Y^2 + a_1XY - a_2X^2$. Para ello consideramos la ecuación $T^2 + a_1T - a_2 = 0$. Si llamamos s_1 y s_2 a sus raíces, entonces haciendo $T = Y/X$ obtenemos que

$$Y^2 + a_1XY - a_2X^2 = (Y - s_1X)(Y - s_2X),$$

luego las tangentes son $Y = s_iX$. Ahora hemos de observar que los s_i no están necesariamente en k, sino que en principio pertenecen a una extensión cuadrática de k. Si la singularidad es una cúspide, entonces $s_1 = s_2$, el discriminante del polinomio $T^2 + a_1T - a_2$ es nulo y sí que podemos asegurar que $s_i \in k$ (si $\operatorname{car} k = 2$ el argumento es distinto, pero llegamos a la misma conclusión). El problema puede aparecer cuando la singularidad es un nodo. Esto nos lleva al concepto siguiente:

Definición 2.26 Diremos que una curva C/k definida por una ecuación de Weierstrass tiene un *nodo racional* si tiene un nodo y las pendientes de las tangentes por dicho nodo están en k. En caso contrario diremos que el nodo es *irracional.*

Con este matiz, podemos probar el teorema siguiente:

Teorema 2.27 *Sea C/k una cúbica singular definible por una ecuación de Weierstrass y cuya singularidad sea una cúspide o un nodo racional. Entonces C/k se transforma mediante un cambio de coordenadas sobre k en una de las dos ecuaciones $Y^2 = X^3$ o bien $Y^2 + XY = X^3$. Ambas tienen a $(0,0)$ como único punto singular.*

DEMOSTRACIÓN: La traslación que lleva el punto singular a $(0,0)$ transforma la ecuación en otra con una singularidad en $(0,0)$ del mismo tipo que la de partida. Según hemos visto, dicha ecuación es de la forma

$$Y^2 + a_1XY - a_2X^2 - X^3 = 0,$$

y existe un $s \in k$ tal que $s^2 + a_1 s - a_2 = 0$. Entonces el cambio $Y = Y' + sX'$ reduce la ecuación a

$$Y^2 + AXY - X^3 = 0.$$

Si $A = 0$ tenemos ya una de las curvas del enunciado. Si $A \neq 0$, el cambio $X = A^2X'$, $Y = A^3Y'$ transforma la ecuación en $Y^2 + XY - X^3 = 0$. ∎

Es claro que si la singularidad es un nodo irracional no es posible transformar la ecuación en $Y^2 + XY - X^3$ mediante un cambio de coordenadas en k, pues el cambio inverso transformaría las tangentes de esta curva en $(0,0)$ —que tienen pendiente racional— en las tangentes de la curva dada en su punto singular, que por consiguiente tendrían también pendiente racional.

Teorema 2.28 *Una cúbica C definida por una ecuación de Weierstrass puede clasificarse como sigue:*

a) C es regular si y sólo si $\Delta \neq 0$,

b) C tiene un nodo si y sólo si $\Delta = 0$ y $c_4 \neq 0$,

c) C tiene una cúspide si y sólo si $\Delta = c_4 = 0$.

DEMOSTRACIÓN: El apartado a) ya está demostrado. Supongamos que C es singular. La condición $c_4 \neq 0$ no se altera por cambios de coordenadas (sobre $\overline{k}$), ni tampoco el tipo de singularidad, luego podemos suponer que C viene dada por una de las dos ecuaciones del teorema anterior. Ahora basta observar que la primera cumple $c_4 = 0$ (y tiene una cúspide en $(0,0)$) y la segunda cumple $c_4 = 1$ (y tiene un nodo). ∎

Veamos ahora que la descripción geométrica de la suma en una curva elíptica proporciona también una estructura de grupo sobre cualquier cúbica singular, siempre y cuando eliminemos su punto singular.

Definición 2.29 Si E/k es una curva definida por una ecuación de Weierstrass, llamaremos $E_r(k)$ al conjunto de los puntos regulares de $E(k)$. Sabemos que $E_r(k)$ coincide con $E(k)$ salvo quizá por un punto, que puede ser un nodo (racional o irracional) o una cúspide.

Teorema 2.30 *Sea E/k una cúbica definida por una ecuación de Weierstrass con una cúspide o un nodo racional S. Entonces la suma en $E_r(k)$ definida mediante la construcción descrita en el teorema* 2.20 *convierte a $E_r(k)$ en un grupo abeliano.*

a) Si S es un nodo racional con tangentes $Y = \alpha_1 X + \beta_1$, $Y = \alpha_2 X + \beta_2$, entonces la aplicación $E_r(k) \longrightarrow k^*$ dada por

$$(x, y) \mapsto \frac{y - \alpha_1 x - \beta_1}{y - \alpha_2 x - \beta_2}$$

es un isomorfismo de grupos.

b) Si S es una cúspide con tangente $Y = \alpha X + \beta$, entonces la aplicación $E_r(k) \longrightarrow k^+$ dada por

$$(x, y) \mapsto \frac{x - x(S)}{y - \alpha x - \beta}$$

es un isomorfismo de grupos.

Demostración: Vamos a demostrar que la aplicación descrita en cada apartado es biyectiva, así como que si una recta corta a $E_r(k)$ en dos puntos no necesariamente distintos, el tercer punto también está en $E_r(k)$ y el producto (resp. la suma) de las imágenes de los tres puntos es 1 (resp. 0). De aquí se sigue inmediatamente que la aplicación correspondiente conserva las operaciones, por lo que la suma en $E_r(k)$ cumple los axiomas de grupo.

Es claro que no perdemos generalidad cambiamos de sistema de referencia, por lo que podemos suponer que la ecuación de E/k es una de las dos dadas por el teorema 2.27:

$$Y^2 - X^3 = 0, \qquad Y^2 + XY - X^3 = 0.$$

En ambos casos el punto singular es $S = (0, 0)$. Consideremos primero la curva $Y^2 - X^3 = 0$, que tiene una cúspide con tangente $Y = 0$. La aplicación es

$$(x, y) \mapsto \frac{x}{y}.$$

Si tomamos como coordenadas afines X y Z en lugar de X, Y, la ecuación se transforma en $Z - X^3 = 0$ y la aplicación en $(x, z) \mapsto x$. Ahora el punto singular $S = [0, 0, 1]$ está en el infinito y $O = (0, 0)$. La aplicación biyecta $E_r(k)$ con k^+, pues tiene inversa $t \mapsto (t, t^3)$.

Es claro que las rectas que pasan por S (las rectas verticales) cortan a $E_r(k)$ en un único punto (contando multiplicidades), por lo que una recta que pase por dos puntos de $E_r(k)$ no puede pasar por S (no es vertical). Esto implica que la suma en $E_r(k)$ está bien definida y, dada una recta $Z = aX + b$ que pase por tres puntos P_1, P_2, $P_3 \in E_r$, tenemos que las coordenadas x_i de P_i son las raíces de la ecuación

$$(aX + b) - X^3 = 0.$$

Como no hay término en X^2, concluimos que $x_1 + x_2 + x_3 = 0$.

Pasemos ahora a la ecuación $Y^2 + XY - X^3 = 0$, que tiene un nodo en $(0,0)$ con tangentes $Y = 0$ e $Y = -X$. La aplicación es $(x, y) \mapsto 1 + x/y$.

Si hacemos el cambio $X = X' - Y'$ la ecuación se transforma en

$$XY - (X - Y)^3 = 0$$

y la aplicación en $(x, y) \mapsto x/y$. Conviene observar que ahora la ecuación no es de Weierstrass. Ahora $O = [1, 1, 0]$ y $S = (0, 0)$. Como en el caso anterior, tomamos como variables afines X y Z, con lo que la curva se convierte en

$$XZ - (X - 1)^3 = 0$$

y la aplicación en $(x, z) \mapsto x$. Ahora $O = (1, 0)$ y $S = [0, 0, 1]$. La aplicación es biyectiva (como aplicación $E_r(k) \longrightarrow k^*$), pues tiene inversa dada por

$$t \mapsto (t, (t - 1)^3/t).$$

Como en el caso anterior, las rectas que pasan por S son verticales y cortan a $E_r(k)$ en un solo punto (o en ninguno en el caso de $X = 0$), por lo que la suma está bien definida y si una recta (no vertical) $Z = aX + b$ corta a $E_r(k)$ en tres puntos P_1, P_2, P_3, las coordenadas x_i correspondientes son las raíces de la ecuación

$$X(aX + b) - (X - 1)^3 = 0.$$

Como el coeficiente de X^3 es -1 y el término independiente es 1, concluimos que $x_1 x_2 x_3 = 1$. ∎

Si la singularidad es un nodo irracional, se sigue cumpliendo parte del teorema anterior: el conjunto $E_r(k)$ sigue siendo un grupo, aunque su estructura es un poco más delicada. Llamemos K a la extensión cuadrática de k que contiene a las pendientes de las tangentes a E por su nodo. Sea σ el k-automorfismo no trivial de K. Las pendientes de las tangentes son las raíces de un polinomio de $k[T]$, luego son conjugadas sobre k. Las dos tangentes son rectas que pasan por un mismo punto racional con pendientes conjugadas, luego $\sigma(\alpha_1) = \alpha_2$, $\sigma(\beta_1) = \beta_2$.

La curva E/K tiene un nodo racional, luego podemos aplicarle el teorema anterior. Si llamamos $\phi : E_r(K) \longrightarrow K^*$ al isomorfismo correspondiente, para cada $P = (x, y) \in E_r(K)$ tenemos que

$$\phi(P)^\sigma = \frac{y^\sigma - \alpha_1^\sigma x^\sigma - \beta_1^\sigma}{y^\sigma - \alpha_2^\sigma x^\sigma - \beta_2^\sigma} = \frac{y^\sigma - \alpha_2 x^\sigma - \beta_2}{y^\sigma - \alpha_1 x^\sigma - \beta_1} = \phi(P^\sigma)^{-1}.$$

Tenemos que $P \in E_r(k)$ si y sólo si $P = P^\sigma$, si y sólo si $\phi(P) = \phi(P^\sigma)$, si y sólo si $\phi(P)\phi(P)^\sigma = 1$, si y sólo si $\mathrm{N}_k^K(\phi(P)) = 1$. Así pues, los elementos de $E_r(k)$ se corresponden a través de ϕ con el núcleo de la norma, que es un subgrupo de K^*, luego $E_r(k)$ es un subgrupo de $E_r(K)$. Hemos probado el teorema siguiente:

Teorema 2.31 *Sea E/k una cúbica definida por una ecuación de Weierstrass con un nodo irracional S. Entonces la suma en $E_r(k)$ definida mediante la construcción descrita en el teorema* 2.20 *convierte a $E_r(k)$ en un grupo abeliano isomorfo al núcleo de la norma de la extensión K/k, donde K es la adjunción a k de las pendientes de las tangentes a E por su nodo.*

El caso que más nos interesa es el de curvas definidas sobre cuerpos finitos:

Teorema 2.32 *Si E/k es una cúbica singular definida por una ecuación de Weierstrass sobre un cuerpo finito k de m elementos, entonces*

$$|E_r(k)| = \begin{cases} m-1 & \text{si } E/k \text{ tiene un nodo racional,} \\ m+1 & \text{si } E/k \text{ tiene un nodo irracional,} \\ m & \text{si } E/k \text{ tiene una cúspide.} \end{cases}$$

Demostración: El caso del nodo irracional se sigue de que la norma de una extensión de cuerpos finitos es suprayectiva, luego, con la notación del teorema anterior, $|K^*| = m^2 - 1$ y el núcleo de la norma tiene $(m^2-1)/(m-1) = m+1$ elementos. ∎

Es claro que si la singularidad es un nodo entonces $E_r(k)$ es un grupo cíclico, mientras que si es una cúspide es un grupo elemental (producto de cíclicos de orden primo). Si k tiene orden primo $E_r(k)$ es cíclico en cualquier caso.

2.4 Isogenias

Nos ocupamos ahora de las aplicaciones que conectan adecuadamente dos curvas elípticas.

Definición 2.33 Una *isogenia* $\phi : E_1 \longrightarrow E_2$ entre dos curvas elípticas es una aplicación regular tal que $\phi(O) = O$.

Obviamente la función constante O es una isogenia (la isogenia nula), y es la única isogenia constante. Notemos que si $\psi : E_1 \longrightarrow E_2$ es una aplicación regular entre curvas elípticas y $P = \psi(O)$, entonces $\phi = \psi \circ \tau_{-P}$ es una isogenia, luego toda aplicación regular ψ entre curvas elípticas es composición de una isogenia seguida de una traslación: $\psi = \phi \circ \tau_P$.

Ahora probamos que las isogenias cumplen más de lo que parece indicar la definición:

Teorema 2.34 *Las isogenias son homomorfismos de grupos.*

Demostración: Sea $\phi : E_1 \longrightarrow E_2$ una isogenia entre curvas elípticas. Podemos suponer que es no nula. Consideramos el diagrama siguiente,

$$\begin{array}{ccc} E_1 & \longrightarrow & H_0(E_1) \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ E_2 & \longrightarrow & H_0(E_2) \end{array}$$

donde las flechas horizontales son los isomorfismos $P \mapsto [P/O]$. Obviamente es conmutativo, luego ϕ es un homomorfismo de grupos. ∎

Como consecuencia de este teorema, si $\phi : E_1 \longrightarrow E_2$ es una isogenia no nula, entonces $\mathrm{N}(\phi) = \phi^{-1}(O)$ es un subgrupo finito de E_1, cuyo orden es a lo sumo el grado de ϕ. Los teoremas siguientes precisan esto mucho más:

Teorema 2.35 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia no nula.*

a) Para cada $Q \in E_2$, el cardinal $|\phi^{-1}[Q]|$ es el grado de separabilidad de ϕ.

b) Para cada $P \in E_1$, el índice de ramificación $e_\phi(P)$ es el grado de inseparabilidad de ϕ.

c) La aplicación $\mathrm{N}(\phi) \longrightarrow G(\overline{k}(E_1)/\overline{k}(E_2))$ definida por $T \mapsto \overline{\tau}_T$ es un isomorfismo de grupos.

d) Si ϕ es separable entonces es no ramificada y la extensión $\overline{k}(E_1)/\overline{k}(E_2)$ es finita de Galois, de grado igual al grado de ϕ.

Demostración: a) Toda extensión de cuerpos de funciones algebraicas se descompone en una extensión puramente inseparable (en la que cada primo del cuerpo base es divisible entre un único primo de la extensión) seguida de una extensión separable (en la que los primos no ramificados del cuerpo base son divisibles entre tantos primos de la extensión como indica el grado). Por lo tanto, en cualquier extensión casi todos los primos tienen tantos divisores como indica el grado de separabilidad. En nuestro caso eso significa que casi todos los puntos $Q \in E_2$ cumplen el apartado a), pero ϕ es un epimorfismo de grupos, luego todos los puntos de E_2 tienen el mismo número de antiimágenes.

b) Por el mismo argumento, casi todos los primos de una extensión tienen índice de ramificación igual al grado de inseparabilidad. Ahora bien, cualquier par de puntos de E_1 pueden conectarse por una traslación τ, la cual induce un $\overline{k}$-automorfismo de $\overline{k}(E_1)$ que hace corresponder los respectivos divisores primos. Así, todos los puntos de E_1 han de tener el mismo índice de ramificación, luego éste ha de ser el grado de inseparabilidad de ϕ.

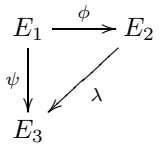
c) Si $T \in \mathrm{N}(\phi)$ y $f \in \overline{k}(E_2)$, entonces

$$\overline{\tau}_T(\overline{\phi}(f)) = \overline{\tau_T \circ \phi}(f) = \overline{\phi}(f),$$

luego $\overline{\tau}_T$ es ciertamente un $\overline{k}(E_2)$-automorfismo de $\overline{k}(E_1)$. Es fácil ver que la aplicación $T \mapsto \overline{\tau}_T$ es un homomorfismo de grupos. La teoría de extensiones de cuerpos nos da que $|G(\overline{k}(E_1)/\overline{k}(E_2))|$ es a lo sumo el grado de separabilidad de la extensión, que es precisamente el orden de $\mathrm{N}(\phi)$. Así pues, si probamos que el homomorfismo es inyectivo, será un isomorfismo. En efecto, si $\overline{\tau}_T = 1$, entonces, para toda $f \in \overline{k}(E_1)$ tenemos que $f(O) = \overline{\tau}_T(f)(O) = f(T)$, lo cual sólo es posible si $T = O$.

d) Si ϕ es separable de grado n entonces es no ramificada por b), el núcleo $\mathrm{N}(\phi)$ tiene n elementos por a) y $G(\overline{k}(E_1)/\overline{k}(E_2))$ tiene también n elementos por c), luego la extensión ha de ser de Galois. ∎

Teorema 2.36 *Si $\phi : E_1 \longrightarrow E_2$ y $\psi : E_1 \longrightarrow E_3$ son isogenias tales que $\mathrm{N}(\phi) \subset \mathrm{N}(\psi)$ y ϕ es separable, entonces existe una isogenia $\lambda : E_2 \longrightarrow E_3$ tal que el diagrama siguiente es conmutativo:*



Demostración: Por el teorema anterior, $\overline{k}(E_1)$ es una extensión finita de Galois de $\overline{k}(E_2)$, y también es una extensión de $\overline{k}(E_3)$. Todo $\overline{k}(E_2)$-automorfismo de $\overline{k}(E_1)$ es de la forma $\overline{\tau}_T$, con $T \in \mathrm{N}(\phi) \subset \mathrm{N}(\psi)$, y si $f \in \overline{k}(E_3)$, entonces $\overline{\tau}_T(\overline{\psi}(f)) = f$. Tenemos, pues, que el grupo de Galois fija a los elementos de $\overline{k}(E_3)$, lo que nos da las inclusiones

$$\overline{\psi}[\overline{k}(E_3)] \subset \overline{\phi}[\overline{k}(E_2)] \subset \overline{k}(E_1).$$

La primera inclusión nos da un $\overline{k}$-monomorfismo $\overline{\lambda} : \overline{k}(E_3) \longrightarrow \overline{k}(E_2)$ tal que $\overline{\lambda} \circ \overline{\phi} = \overline{\psi}$. Por consiguiente, existe una aplicación racional (que, de hecho, será regular) $\lambda : E_2 \longrightarrow E_3$ tal que $\phi \circ \lambda = \psi$. De aquí se sigue que $\lambda(O) = O$, luego λ es una isogenia. ∎

De aquí se obtiene la unicidad del teorema siguiente:

Teorema 2.37 *Sea E una curva elíptica y H un subgrupo finito de E. Entonces existe una única curva elíptica E' (salvo isomorfismo) y una isogenia separable $\phi : E \longrightarrow E'$ tal que $\mathrm{N}(\phi) = H$.*

Demostración: Cada $T \in H$ induce un $\overline{k}$-automorfismo $\overline{\tau}_T$ de $\overline{k}(E)$. Llamemos K al subcuerpo de $\overline{k}(E)$ fijado por todos ellos. Entonces $\overline{k}(E)/K$ es una extensión finita de Galois y su grupo de Galois es isomorfo a H (si $a \in \overline{k}(E)$, entonces $p(X) = \prod_{T \in H} (X - \overline{\tau}_T(a)) \in K[X]$).

Tenemos que $\overline{k} \subset K \subset \overline{k}(E)$, y la extensión superior es finita, de donde se sigue que K es un cuerpo de funciones algebraicas sobre $\overline{k}$, luego existe una curva proyectiva regular E' tal que $\overline{k}(E')$ es $\overline{k}$-isomorfo a K. Componiendo este isomorfismo con la inclusión $K \subset \overline{k}(E)$ obtenemos un $\overline{k}$-monomorfismo $\overline{\phi} : \overline{k}(E') \longrightarrow \overline{k}(E)$, inducido por una aplicación regular $\phi : E \longrightarrow E'$. Por construcción $\overline{\phi}[\overline{k}(E')] = K$.

Vamos a ver que ϕ es no ramificada. Tomemos $P \in E$ y $T \in H$. Para toda $f \in \overline{k}(E')$, se cumple

$$f(\phi(P+T)) = \overline{\tau}_T(\overline{\phi}(f))(P) = \overline{\phi}(f)(P) = f(\phi(P)),$$

donde hemos usado que $\overline{\tau}_T$ fija a $\overline{\phi}(f) \in K$. Esto implica que $\phi(P+T) = \phi(P)$.

Para cada $Q \in E'$ tomemos $P \in E$ tal que $\phi(P) = Q$. Entonces Q tiene a lo sumo tantas antiimágenes como el grado de ϕ, que es $|H|$, pero por otra parte

tiene como antiimágenes a los puntos $P+T$, con $T \in H$, que son distintos dos a dos, luego todos los puntos de E' tienen exactamente $|H|$ antiimágenes. Esto sólo es posible si ϕ es no ramificada y, en virtud del teorema anterior, separable.

Ahora aplicamos la fórmula del género de Hurwitz, que para una extensión no ramificada se reduce a

$$0 = 2g_E - 2 = (2g_{E'} - 2)\operatorname{grad}\phi,$$

luego el género de E' ha de ser $g_{E'} = 1$. Si definimos $O' = \phi(O)$, tenemos que E' es una curva elíptica y ϕ es una isogenia.

Para probar la unicidad observamos que si $\psi : E \longrightarrow E''$ fuera otra isogenia de núcleo H, por el teorema anterior existiría una isogenia $\lambda : E' \longrightarrow E''$ tal que $\phi \circ \lambda = \psi$, pero λ sería un isomorfismo, ya que si $\lambda(P) = O$ podemos expresar $P = \phi(Q)$, con $Q \in E$, pero entonces $\psi(Q) = O$, luego $Q \in H$ y $P = \phi(Q) = O$. ∎

Notemos que la isogenia ϕ no es única, pues, por ejemplo, $-\phi$ cumple también el teorema.

Ejemplo Supongamos que $\operatorname{car} k \neq 2$ y consideremos la curva dada por la ecuación

$$E_1 : Y^2 = X^3 + aX^2 + bX.$$

Su discriminante es $\Delta = 16b^2(a^2 - 4b)$, luego E_1 es una curva elíptica si suponemos que $b \neq 0$ y $b' = a^2 - 4b \neq 0$. En estas condiciones, también es elíptica la curva (de la misma familia) dada por

$$E_2 : Y^2 = X^3 - 2aX^2 + b'X.$$

Definimos $\phi : E_1 \longrightarrow E_2$ como la aplicación dada por

$$\phi(X,Y) = \left(\frac{Y^2}{X^2}, \frac{Y(b-X^2)}{X^2}\right).$$

Veamos que, en efecto, si $(X,Y) \in E_1$ con $X \neq 0$ entonces $\phi(X,Y) \in E_2$. Hemos de ver que

$$\frac{Y^6}{X^6} - 2a\frac{Y^4}{X^4} + (a^2 - 4b)\frac{Y^2}{X^2} = \frac{Y^2(b-X^2)^2}{X^4}.$$

Dividiendo entre Y^2 y multiplicando por X^6 esto equivale a

$$Y^4 - 2aX^2Y^2 + (a^2 - 4b)X^4 = b^2X^2 - 2bX^4 + X^6.$$

Ahora basta sustituir en el miembro izquierdo Y^4 e Y^2 por la expresión que proporciona la ecuación de E_1 y comprobar que tenemos una identidad.

El único punto (finito) de E_1 con $X = 0$ es $(0,0)$. Para calcular su imagen expresamos ϕ en coordenadas homogéneas:

$$\phi(X,Y) = [Y^2, Y(b-X^2), X^2].$$

Vemos que las tres coordenadas se anulan, pero la primera tiene un cero de orden 2 en $(0,0)$, la segunda de orden 1 y la tercera de orden 4. Por lo tanto, basta dividir entre Y:

$$\phi(X,Y) = \left[Y, b - X^2, \frac{YX}{X^2 + aX + b}\right],$$

de donde concluimos que $\phi(0,0) = [0,1,0] = O$. Igualmente podemos comprobar que $\phi(O) = O$, si bien esto se sigue de que ϕ es una isogenia.

Hemos comprobado que el núcleo de ϕ tiene orden 2, pues está formado por O y $(0,0)$.

Observemos ahora que podemos construir una curva E_3 a partir de E_2 igual que hemos construido E_2 a partir de E_1. El resultado es

$$E_3 : Y^2 = X^3 + 4aX^2 + 16bX.$$

Ahora bien, E_3 es isomorfa a E_1 a través de $(X,Y) \mapsto (X/4, Y/8)$. Al componer la isogenia $E_2 \longrightarrow E_3$ análoga a ϕ con este isomorfismo obtenemos una isogenia $\hat{\phi} : E_2 \longrightarrow E_1$ dada por

$$\hat{\phi}(X,Y) = \left(\frac{Y^2}{4X^2}, \frac{Y(b' - X^2)}{8X^2}\right),$$

cuyo núcleo es también de orden 2. ∎

La situación del ejemplo anterior no es casual: Aunque no es evidente en absoluto, veremos que toda isogenia $\phi : E_1 \longrightarrow E_2$ entre dos curvas elípticas tiene asociada una isogenia dual $\hat{\phi} : E_2 \longrightarrow E_1$.

Terminamos la sección con una observación útil para trabajar con isogenias en cuerpos de característica prima:

Teorema 2.38 *Si E es una curva elíptica sobre un cuerpo de característica prima p y $m = p^r$, entonces la aplicación de Frobenius $\phi : E \longrightarrow E^{(m)}$ es una isogenia.*

Demostración: Basta probar que la curva $E^{(m)}$ tiene género 1, pues entonces podemos considerarla como una curva elíptica con neutro igual a $\phi(O)$, lo que convierte a ϕ en una isogenia. Si $K = \overline{k}(E)$, entonces $\overline{k}(E^{(m)}) = K^m$. Todo se reduce a probar que si K es un cuerpo de funciones algebraicas de género 1 sobre un cuerpo de característica p, entonces K^m también tiene género 1. De aquí se sigue que no perdemos generalidad si suponemos que $E \subset \mathrm{P}^2$ está definida por una ecuación de Weierstrass, pero en tal caso $E^{(m)}$ está definido por la ecuación de Weierstrass que resulta de elevar a m los coeficientes de la ecuación de E. El discriminante de la ecuación de $E^{(m)}$ se obtiene también elevando a m el discriminante de la ecuación de E, luego es no nulo. ∎

2.5 Curvas conjugadas

Hemos visto que dos curvas elípticas son isomorfas si y sólo si tienen el mismo invariante. Ahora bien, si estamos estudiando —por ejemplo— el conjunto de puntos racionales de una curva elíptica dada E/k, no podemos sustituirla por otra curva isomorfa cualquiera, pues $E(k)$ puede cambiar completamente, sino que a lo sumo podemos reemplazarla por una curva elíptica k-isomorfa. Esto plantea diversas cuestiones que vamos a tratar aquí, como en cuántas clases distintas de curvas k-isomorfas se descompone una clase de isomorfía de curvas.

Cohomología no abeliana Los problemas que vamos a abordar se tratan más adecuadamente con unos mínimos rudimentos de cohomología no abeliana:

Supongamos que el grupo de Galois $G(\overline{k}/k)$ actúa sobre un grupo M no necesariamente abeliano, es decir, que tenemos un homomorfismo de grupos

$$\rho : G(\overline{k}/k) \longrightarrow \mathrm{Aut}(M).$$

Diremos que la acción es *discreta* si para cada $m \in M$ el estabilizador

$$\mathrm{Est}(m) = \{\sigma \in G(\overline{k}/k) \mid m^\sigma = m\}$$

tiene índice finito en $G(\overline{k}/k)$ o, lo que es lo mismo, es de la forma $G(\overline{k}/l)$, para cierta extensión finita l de k. Esto equivale a que la aplicación

$$\rho : G(\overline{k}/k) \times M \longrightarrow M$$

inducida por la acción sea continua cuando en $G(\overline{k}/k)$ consideramos la topología de Krull y en M la topología discreta. En estas condiciones, definimos el grupo de cohomología

$$H^0(\overline{k}/k, M) = \{m \in M \mid m^\sigma = m \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Un *cociclo* es una aplicación $\xi : G(\overline{k}/k) \longrightarrow M$ que cumple la relación

$$\xi_{\sigma\tau} = \xi_\tau \xi_\sigma^\tau.$$

Diremos que ξ es *continuo* si lo es respecto a la topología de Krull y la topología discreta en M, es decir, si ξ_σ depende únicamente de la clase de σ módulo un subgrupo de índice finito en $G(\overline{k}/k)$ o, dicho de otro modo, si ξ está inducido por una aplicación $\xi : G(l/k) \longrightarrow M$, para cierta extensión finita normal l de k.

Diremos que dos cociclos ξ y ζ son *cohomólogos* si existe un $m \in M$ tal que

$$\xi_\sigma m^\sigma = m\zeta_\sigma \qquad \text{para todo } \sigma \in G(\overline{k}/k).$$

Cuando M es un grupo abeliano, los cociclos continuos forman un grupo con el producto definido puntualmente, y la cohomología de cociclos es la congruencia respecto al subgrupo de las cocadenas (los cociclos cohomólogos al cociclo

constante 1). Sin embargo, cuando M no es abeliano los cociclos no forman un grupo, aunque la cohomología sigue siendo una relación de equivalencia. Por lo tanto podemos definir el conjunto cociente del conjunto de todos los cociclos continuos respecto a la relación de cohomología, y lo representaremos por

$$H^1(\overline{k}/k, M).$$

Grupos de automorfismos Si C/k es una curva proyectiva regular llamaremos $\mathrm{Aut}(C)$ al grupo de automorfismos de C en el sentido usual en geometría algebraica, es decir, al grupo de las biyecciones regulares con inversa regular de C en sí misma.

Cuando C/k sea una curva elíptica escribiremos $\mathrm{Aut}_g(C)$ (donde la g hace referencia a "automorfismos geométricos") para distinguir a este grupo del grupo de automorfismos algebraicos (automorfismos geométricos que además conservan la estructura de grupo). El teorema siguiente muestra la relación entre ambos:

Teorema 2.39 *Si E/k es una curva elíptica, la aplicación*

$$\mathrm{Aut}(E) \times E \longrightarrow \mathrm{Aut}_g(E)$$

dada por $(\phi, P) \mapsto \phi\tau_P$ (donde $\tau_P(Q) = P+Q$) es biyectiva, y es un isomorfismo de grupos si en $\mathrm{Aut}(E) \times E$ consideramos el producto semidirecto inducido por la acción natural de $\mathrm{Aut}(E)$ sobre E, es decir,

$$(\phi, P)(\psi, Q) = (\phi\psi, \psi(P) + Q).$$

Demostración: La suprayectividad se debe a que todo automorfismo geométrico que fije al neutro O es un isomorfismo algebraico. Por consiguiente, dado $\alpha \in \mathrm{Aut}_g(E)$, tomamos $P = \alpha(O)$ y $\phi = \alpha\tau_{-P}$, de modo que $\phi(O) = O$, luego $\phi \in \mathrm{Aut}(E)$ y $(\phi, P) \mapsto \alpha\tau_{-P}\tau_P = \alpha$.

La inyectividad es trivial, pues si $\phi\tau_P = \psi\tau_Q$, evaluando en O obtenemos $P = Q$, de donde también $\phi = \psi$.

Por último vemos que el producto de las imágenes de (ϕ, P) y (ψ, Q) es

$$\phi\tau_P\psi\tau_Q = \phi\psi\psi^{-1}\tau_P\psi\tau_Q = \phi\psi\tau_{\psi(P)}\tau_Q = \phi\psi\tau_{\psi(P)+Q},$$

que es la imagen del par $(\phi\psi, \psi(P) + Q)$. ∎

Curvas conjugadas Introducimos ahora el concepto principal de esta sección:

Definición 2.40 Diremos que dos curvas proyectivas regulares C/k y C'/k son *conjugadas* si son isomorfas (sobre $\overline{k}$). Representaremos por $\mathrm{Conj}(C/k)$ al conjunto cociente del conjunto de curvas conjugadas con C/k respecto de la relación de equivalencia dada por la k-isomorfía.

Por ejemplo, todas las curvas elípticas $E_n/\mathbb{Q}$ dadas por $Y^2 = X^3 - n^2X$ son conjugadas (pues tienen invariante $j = 1728$), pero no son todas $\mathbb{Q}$-isomorfas. Si lo fueran, todos los números naturales serían congruentes (o no lo sería ninguno).

Notemos ahora que $G(\overline{k}/k)$ actúa discretamente sobre $\mathrm{Aut}(C)$ mediante la acción dada por $\phi^\sigma(P) = \phi(P^{\sigma^{-1}})^\sigma$. En efecto, si ϕ está definido sobre l, entonces $G(\overline{k}/l)$ estabiliza a ϕ.

Si C/k es una curva proyectiva regular y C'/k es una curva conjugada con C/k, elegimos un isomorfismo $\phi : C' \longrightarrow C$ y, para cada $\sigma \in G(\overline{k}/k)$, definimos

$$\xi_\sigma = \phi^{-1} \circ \phi^\sigma \in \mathrm{Aut}(C).$$

Vamos a comprobar que se trata de un cociclo continuo. Ciertamente es un cociclo:

$$\xi_{\sigma\tau} = \phi^{-1} \circ \phi^{\sigma\tau} = (\phi^{-1}\phi^\tau)(\phi^{-1}\phi^\sigma)^\tau = \xi_\tau \xi_\sigma^\tau.$$

Además es continuo, pues si ϕ está definido sobre una extensión finita normal l de k, entonces ξ_σ sólo depende de la clase de σ módulo $G(\overline{k}/l)$.

Ahora demostramos que la clase de cohomología $[\xi] \in H^1(\overline{k}/k, \mathrm{Aut}(C))$ sólo depende de la clase de conjugación de C/k. En efecto, si C''/k es una curva k-isomorfa a C/k y $\psi : C'' \longrightarrow C$ es un isomorfismo, hemos de probar que los cociclos $\xi_\sigma = \phi^{-1} \circ \phi^\sigma$ y $\zeta_\sigma = \psi^{-1} \circ \psi^\sigma$ son cohomólogos.

Tomemos un k-isomorfismo $\theta : C' \longrightarrow C''$ y sea $\alpha = \phi^{-1} \circ \theta \circ \psi \in \mathrm{Aut}(C)$. Entonces

$$\xi_\sigma \alpha^\sigma = \phi^{-1} \circ \theta \circ \psi^\sigma = \phi^{-1} \circ \theta \circ \psi \circ \zeta_\sigma = \alpha\zeta_\sigma.$$

Con esto tenemos probada una parte del teorema siguiente:

Teorema 2.41 *Sea C/k una curva proyectiva regular. Para cada curva C'/k conjugada con C/k elegimos un isomorfismo $\phi : C' \longrightarrow C$ y para cada automorfismo $\sigma \in G(\overline{k}/k)$ definimos $\xi_\sigma = \phi^{-1} \circ \phi^\sigma$. Entonces la correspondencia $C'/k \mapsto \xi$ induce una biyección*

$$\mathrm{Conj}(C/k) \longrightarrow H^1(\overline{k}/k, \mathrm{Aut}(C)).$$

Demostración: Tenemos probada la existencia de la aplicación inducida. Falta ver que es biyectiva. Supongamos que C'/k y C''/k son curvas conjugadas con C/k que determinan la misma clase de cohomología. Elegimos isomorfismos $\phi : C' \longrightarrow C$ y $\psi : C'' \longrightarrow C$ y formamos los cociclos $\xi_\sigma = \phi^{-1} \circ \phi^\sigma$, $\zeta_\sigma = \psi^{-1} \circ \psi^\sigma$. Por hipótesis existe $\alpha \in \mathrm{Aut}(C)$ tal que $\xi_\sigma \alpha^\sigma = \alpha\zeta_\sigma$, para todo $\sigma \in G(\overline{k}/k)$. Consideramos el isomorfismo $\theta = \phi \circ \alpha \circ \psi^{-1} : C' \longrightarrow C''$. Vamos a probar que está definido sobre k. En efecto, si $\sigma \in G(\overline{k}/k)$, tenemos que

$$\theta^\sigma = \phi^\sigma \circ \alpha^\sigma \circ (\psi^{-1})^\sigma = \phi \circ \xi_\sigma \alpha^\sigma \circ (\psi^{-1})^\sigma = \phi \circ \alpha\zeta_\sigma \circ (\psi^{-1})^\sigma = \phi \circ \alpha \circ \psi^{-1} = \theta.$$

Esto prueba que la correspondencia es inyectiva. La parte más delicada es la suprayectividad. Consideremos un cociclo $\xi : G(\overline{k}/k) \longrightarrow \mathrm{Aut}(C)$. Cada

$\sigma \in G(\overline{k}/k)$ determina un automorfismo $\xi_\sigma : C \longrightarrow C$, el cual determina a su vez un $\overline{k}$-automorfismo de cuerpos $\bar{\xi}_\sigma : \overline{k}(C) \longrightarrow \overline{k}(C)$.

Para cada $f \in \overline{k}(C)$ definimos $f\sigma = \bar{\xi}_\sigma(f^\sigma)$. Se comprueba inmediatamente que se trata de una acción discreta de $G(\overline{k}/k)$ sobre $\overline{k}(C)$ (distinta de la usual, denotada por f^σ). Definimos

$$L = \{f \in \overline{k}(C) \mid f\sigma = f \text{ para todo } \sigma \in G(\overline{k}/k)\}.$$

Es claro que L es un subcuerpo de $\overline{k}(C)$. Además $L \cap \overline{k} = k$, pues si $f \in \overline{k}$ se cumple que $f\sigma = f^\sigma$. Veamos ahora que $\overline{k}L = \overline{k}(C)$. Tomamos $f \in \overline{k}(C)$ y sea l/k una extensión finita normal tal que $f\sigma = f$ para todo $\sigma \in G(\overline{k}/l)$. Sea $\alpha_1, \ldots, \alpha_n$ una k-base de l y sea $G(l/k) = \{\sigma_1, \ldots, \sigma_n\}$. Los elementos

$$g_i = \sum_{j=1}^{n} (\alpha_i f)\sigma_j = \sum_{j=1}^{n} \alpha_i^{\sigma_j} (f\sigma_j)$$

están claramente en L.

Por otra parte, la matriz $(\alpha_i^{\sigma_j})$ es regular, pues el cuadrado de su determinante es el discriminante de la base. Esto nos permite despejar cada f^{σ_j} (en particular f) como combinación lineal de los g_i, lo que prueba que $f \in \overline{k}L$.

En particular tenemos que L tiene grado de trascendencia 1 sobre k, luego es un cuerpo de funciones algebraicas sobre el cuerpo de constantes (exacto) k. Esto implica que L es k-isomorfo a $k(C')$, donde C'/k es una curva proyectiva regular.

El isomorfismo se extiende a un $\overline{k}$-isomorfismo $\overline{\phi} : \overline{k}(C) \longrightarrow \overline{k}(C')$. A través de $\overline{\phi}$, la acción que hemos definido en $\overline{k}(C)$ se corresponde con una acción de $G(\overline{k}/k)$ en $\overline{k}(C')$ que fija a $k(C')$ y coincide con la natural sobre $\overline{k}$. Ahora bien, estas dos propiedades determinan completamente a dicha acción, luego ésta ha de coincidir con la acción usual de $G(\overline{k}/k)$ sobre $\overline{k}(C')$. En otros términos, para toda función $f \in \overline{k}(C)$ y todo $\sigma \in G(\overline{k}/k)$ tenemos que $\overline{\phi}(f)^\sigma = \overline{\phi}(f\sigma) = \overline{\phi}(\xi_\sigma \circ f^\sigma)$.

Sea $\phi : C' \longrightarrow C$ el isomorfismo de curvas que induce el $\overline{k}$-isomorfismo $\overline{\phi}$, es decir, tal que $\overline{\phi}(f) = \phi \circ f$. Entonces

$$\phi^\sigma \circ f^\sigma = \phi \circ \xi_\sigma \circ f^\sigma,$$

para toda $f \in \overline{k}(C)$, de donde podemos concluir que $\phi^\sigma = \phi \circ \xi_\sigma$ o, lo que es lo mismo, $\xi_\sigma = \phi^{-1} \circ \phi^\sigma$. Esto significa que la clase de la curva C'/k en $\mathrm{Conj}(C'/k)$ se corresponde con la clase de ξ en $H^1(\overline{k}/k, \mathrm{Aut}(C))$. ∎

Espacios Homogéneos Vamos a ver que si E/k es una curva elíptica, las curvas conjugadas con E/k determinadas por cociclos con valores en E tienen una estructura adicional. En primer lugar describiremos dicha estructura y luego la relacionaremos con la cohomología.

Definición 2.42 Sea E/k una curva elíptica un *espacio homogéneo* (principal) para E/k es una curva proyectiva regular C/k junto con una aplicación regular $+ : C \times E \longrightarrow C$ definida sobre k que cumpla las propiedades siguientes:

a) $p + O = p$, para todo $p \in C$,

b) $(p + P) + Q = p + (P + Q)$, para todo $p \in C$, P, $Q \in E$,

c) Para cada $p \in C$, la aplicación $\theta_p : E \longrightarrow C$ dada por $\theta_p(P) = p + P$ es un isomorfismo.

En particular, la propiedad c) implica que todo espacio homogéneo sobre E/k es una curva conjugada con E/k. Otra consecuencia es que para cada par de puntos p, $q \in C$ existe un único punto $P \in E$ tal que $p + P = q$. Representaremos a este punto por $P = q - p$.

Notemos que la resta $- : C \times C \longrightarrow E$ es una aplicación regular definida sobre k. En efecto, fijamos un punto $p_0 \in C$ y observamos que[1]

$$\theta_{p_0}^{-1}(q) - \theta_{p_0}^{-1}(p) = p_0 + \theta_{p_0}^{-1}(q) - (p_0 + \theta_{p_0}^{-1}(p)) = q - p.$$

Como $\theta_{p_0}^{-1}$ (y la resta en E) es regular, concluimos que la resta en C también lo es. Por otra parte, si $\sigma \in G(\overline{k}/k)$, vemos que

$$\theta_p^\sigma(P) = \theta_p(P^{\sigma^{-1}})^\sigma = p^\sigma + P = \theta_{p^\sigma}(P),$$

luego $\theta_p^\sigma = \theta_{p^\sigma}$ y, en consecuencia, $(\theta_p^{-1})^\sigma = \theta_{p^\sigma}^{-1}$. Por consiguiente:

$$(p - q)^\sigma = \theta_q^{-1}(p)^\sigma = (\theta_q^\sigma)^{-1}(p^\sigma) = \theta_{q^\sigma}^{-1}(p^\sigma) = p^\sigma - q^\sigma,$$

lo que significa que la resta está definida sobre k.

Diremos que dos espacios homogéneos C/k y C'/k para una curva elíptica E/k son *equivalentes* si existe un isomorfismo $\theta : C \longrightarrow C'$ definido sobre k tal que

$$\theta(p + P) = \theta(p) + P,$$

para todo $p \in C$, $P \in E$.

Definimos el *grupo de Weil-Châtelet* de una curva elíptica E/k como el conjunto de clases de equivalencia de espacios homogéneos para E/k. (Enseguida veremos que tiene una estructura natural de grupo.) Lo representaremos por $\mathrm{WC}(E/k)$.

Es claro que E/k es un espacio homogéneo para sí misma considerando la acción definida por traslaciones. Los espacios homogéneos equivalentes a E/k se llaman *triviales.* Veremos que la clase trivial es el elemento neutro del grupo de Weil-Châtelet. En primer lugar damos una caracterización sencilla:

Teorema 2.43 *Si C/k es un espacio homogéneo para una curva elíptica E/k, entonces C/k es trivial si y sólo si $C(k) \neq \varnothing$.*

[1] Aquí usamos que, en general, $(q+Q)-(p+P) = (q-p)+(Q-P)$, pues esto es equivalente a $q + Q = p + P + (q - p) + Q - P$, que es una consecuencia inmediata de la definición de espacio homogéneo.

DEMOSTRACIÓN: Si C/k es trivial y $\theta : E \longrightarrow C$ es una equivalencia de espacios homogéneos, entonces $\theta(O) \in C(k)$.

Recíprocamente, si $p_0 \in C(k)$, el isomorfismo $\theta_{p_0} : E \longrightarrow C$ está definido sobre k y es una equivalencia de espacios homogéneos, pues

$$\theta_{p_0}(P+Q) = p_0 + P + Q = \theta_{p_0}(P) + Q.$$

■

Ejemplo *Si E/k es una curva elíptica sobre un cuerpo finito k, entonces* $\mathrm{WC}(E/k) = 0$.

Esto es consecuencia de que todo cuerpo de funciones algebraicas sobre un cuerpo finito tiene al menos un divisor primo de grado 1. En particular, si C/k es un espacio homogéneo para E/k, tenemos que $k(C)$ tiene un divisor primo de grado 1, que se corresponde con un punto de $C(k)$ y, por consiguiente, C/k determina la clase trivial en $\mathrm{WC}(E/k)$. ■

Teorema 2.44 *Si E/k es una curva elíptica, existe una biyección*

$$\mathrm{WC}(E/k) \longrightarrow H^1(G(\overline{k}/k), E)$$

definida como sigue: para cada clase $[C/k] \in \mathrm{WC}(E/k)$ elegimos $p_0 \in C$ y le asignamos la clase del cociclo $\{p_0^\sigma - p_0\}_\sigma$.

DEMOSTRACIÓN: Veamos que la aplicación está bien definida. En primer lugar, $\{p_0^\sigma - p_0\}_\sigma$ es un cociclo:

$$p_0^{\sigma\tau} - p_0 = (p_0^{\sigma\tau} - p_0^\tau) + (p_0^\tau - p_0) = (p_0^\sigma - p_0)^\tau + (p_0^\tau - p_0).$$

Además es continuo, pues si $p_0 \in k'$, depende sólo de la clase de σ módulo $G(\overline{k}/k')$.

En segundo lugar, la imagen no depende de la elección del representante C/k ni del punto p_0, pues si C'/k es un espacio equivalente y $p_0' \in C'(k)$, sea $\theta : C \longrightarrow C'$ la equivalencia. Entonces

$$p_0^\sigma - p_0 = \theta(p_0^\sigma) - \theta(p_0) = (p_0'^\sigma - p_0') + (\theta(p_0)^\sigma - p_0'^\sigma) - (\theta(p_0) - p_0'),$$

luego los cociclos $\{p_0^\sigma - p_0\}_\sigma$ y $\{p_0'^\sigma - p_0'\}_\sigma$ se diferencian en la cocadena inducida por $m = \theta(p_0) - p_0'$ e inducen la misma clase de cohomología.

Ahora veamos que la aplicación es inyectiva. Si dos espacios homogéneos C/k y C'/k determinan cociclos cohomólogos $\{p_0^\sigma - p_0\}_\sigma$ y $\{p_0'^\sigma - p_0\}_\sigma$, entonces existe un $m \in E$ tal que

$$p_0^\sigma - p_0 = (p_0'^\sigma - p_0') + (m^\sigma - m),$$

para todo $\sigma \in G(\overline{k}/k)$. Definimos $\theta : C \longrightarrow C'$ mediante $\theta(p) = p_0' + (p - p_0) + m$. Es claro que θ es un isomorfismo. Además, si $\sigma \in G(\overline{k}/k)$ vemos que

$$\theta^\sigma(p) = \theta(p^{\sigma^{-1}})^\sigma = (p_0' + (p^{\sigma^{-1}} - p_0) + m)^\sigma = p_0'^\sigma + (p - p_0^\sigma) + m^\sigma$$

$$= p'_0 + (p'^{\sigma}_0 - p'_0) + (p - p^{\sigma}_0) + (m^{\sigma} - m) + m$$
$$= p'_0 + (p^{\sigma}_0 - p_0) + (p - p^{\sigma}_0) + m = p'_0 + (p - p_0) + m = \theta(p),$$

luego θ está definido sobre k. Se comprueba inmediatamente que θ es compatible con la acción de E, luego C/k y C'/k son equivalentes.

Por último, veamos que la aplicación es suprayectiva. El teorema 2.39 nos permite ver a E como subgrupo de $\mathrm{Aut}_g(E)$, luego también podemos ver un cociclo $\{\xi_\sigma\}_\sigma$ sobre E como cociclo sobre $\mathrm{Aut}_g(E)$. Enseguida veremos que hemos de pasar al cociclo $\{-\xi_\sigma\}_\sigma$. El teorema 2.41 nos da una curva C/k junto con un isomorfismo $\phi : C \longrightarrow E$ tal que para todo $\sigma \in G(\overline{k}/k)$ se cumple $\phi^{-1} \circ \phi^{\sigma} = -\xi_\sigma$. En particular,

$$(\phi^{-1})^{\sigma} \circ \phi = -\xi^{\sigma}_{\sigma^{-1}} = \xi_\sigma,$$

donde hemos usado la ecuación de los cociclos teniendo en cuenta que $\xi_1 = O$.

Definimos $+ : C \times E \longrightarrow C$ mediante $p + P = \phi^{-1}(\phi(p) + P)$. Veamos que la suma está definida sobre k. Para ello tomamos $\sigma \in G(\overline{k}/k)$ y calculamos

$$(p + P)^{\sigma} = (\phi^{-1})^{\sigma}(\phi^{\sigma}(p^{\sigma}) + P^{\sigma}) = \phi^{-1}(\phi(p^{\sigma}) - \xi_\sigma + P^{\sigma} + \xi_\sigma) = p^{\sigma} + P^{\sigma}.$$

Ahora es inmediato comprobar que C/k es un espacio homogéneo con esta suma (su estructura es la trasladada a través de ϕ de la estructura trivial de espacio homogéneo en E/k). Para calcular su cociclo, tomamos $p_0 = \phi^{-1}(O)$, con lo que

$$p^{\sigma}_0 - p_0 = (\phi^{-1})^{\sigma}(O) - \phi^{-1}(O) = \phi^{-1}(O + \xi_\sigma) - \phi^{-1}(O) = \xi_\sigma,$$

pues, ciertamente, $\phi^{-1}(O) + \xi_\sigma = \phi^{-1}(O + \xi_\sigma)$ por la definición de la suma. ∎

Como E es un grupo abeliano, $H^1(G(\overline{k}/k), E)$ es un grupo, luego la biyección descrita en el teorema anterior induce una estructura de grupo en el conjunto $\mathrm{WC}(E/k)$. Es inmediato que el elemento neutro es la clase trivial.

Curvas elípticas conjugadas Nos ocupamos ahora de las curvas conjugadas con una curva elíptica dada E/k correspondientes a cociclos con valores en el grupo $\mathrm{Aut}(E)$ de los automorfismos de E como curva elíptica.

Notemos que el teorema 2.39 nos da homomorfismos de grupos

$$i : \mathrm{Aut}(E) \longrightarrow \mathrm{Aut}_g(E), \quad j : \mathrm{Aut}_g(E) \longrightarrow \mathrm{Aut}(E)$$

tales que $i \circ j = 1$, donde i es la inclusión. Estos homomorfismos inducen a su vez aplicaciones

$$i : H^1(G(\overline{k}/k), \mathrm{Aut}(E)) \longrightarrow H^1(G(\overline{k}/k), \mathrm{Aut}_g(E)),$$
$$j : H^1(G(\overline{k}/k), \mathrm{Aut}_g(E)) \longrightarrow H^1(G(\overline{k}/k), \mathrm{Aut}(E))$$

tales que $i \circ j = 1$, por lo que i es inyectiva y podemos ver a $H^1(G(\overline{k}/k), \mathrm{Aut}(E))$ como subconjunto de $H^1(G(\overline{k}/k), \mathrm{Aut}_g(E))$.

Diremos que dos curvas elípticas E/k y E'/k son *conjugadas* sobre k *como curvas elípticas* (no sólo como curvas) si son isomorfas sobre $\overline{k}$ (a través de un isomorfismo de curvas elípticas, es decir, un isomorfismo que hace corresponder los elementos neutros respectivos). Representaremos por $\mathrm{Conj}_O(E/k)$ el conjunto de clases de k-isomorfía de curvas elípticas E'/k conjugadas con E en este sentido.

Teorema 2.45 *Si E/k es una curva elíptica, la biyección del teorema* 2.41 *se restringe a una biyección*

$$\mathrm{Conj}_O(E/k) \longrightarrow H^1(G(\overline{k}/k), \mathrm{Aut}(E)).$$

Demostración: Sea E'/k una curva elíptica conjugada con E/k. Sea $\phi : E' \longrightarrow E$ un isomorfismo. Entonces

$$(\phi^{-1} \circ \phi^\sigma)(O) = \phi^\sigma(\phi^{-1}(O)) = \phi^\sigma(O') = \phi(O')^\sigma = O^\sigma = O,$$

luego $\phi^{-1} \circ \phi^\sigma \in \mathrm{Aut}(E)$.

Recíprocamente, si una curva E'/k determina un cociclo $\{\phi^{-1} \circ \phi^\sigma\}_\sigma$ con imagen en $\mathrm{Aut}(E)$ (donde $\phi : E' \longrightarrow E$ es un isomorfismo de curvas), entonces, llamando $O' = \phi^{-1}(O)$, vemos que

$$\phi^\sigma(\phi^{-1}(O)) = O,$$

luego $O'^\sigma = O'$, lo que significa que $O' \in E(k)$ y podemos considerar a E'/k como curva elíptica sobre k con neutro O, y ϕ se convierte en un isomorfismo de curvas elípticas. ∎

El teorema 2.12 muestra que si la característica de k es distinta de 2 y 3 entonces $\mathrm{Aut}(E)$ es un grupo cíclico, luego $\mathrm{Conj}_O(E/k)$ tiene también una estructura natural de grupo. Es fácil calcularlo explícitamente:

Teorema 2.46 *Sea E/k una curva elíptica sobre un cuerpo k de característica distinta de* 2 *y* 3*. Sea*

$$n = \begin{cases} 2 & \text{si } j(E) \neq 0,\ 1728, \\ 4 & \text{si } j(E) = 1728, \\ 6 & \text{si } j(E) = 0. \end{cases}$$

Entonces $\mathrm{Conj}_O(E/k) \cong k^*/k^{*n}$. Explícitamente, si E/k admite una ecuación de Weierstrass $Y^2 = X^3 + a_4X + a_6$, entonces cada clase d (mód k^{*n}) se corresponde con la curva elíptica dada por

$$\begin{array}{ll} Y^2 = X^3 + d^2a_4X + d^3a_6 & \text{si } j(E) \neq 0,\ 1728, \\ Y^2 = X^3 + da_4X & \text{si } j(E) = 1728, \\ Y^2 = X^3 + da_6 & \text{si } j(E) = 0. \end{array}$$

DEMOSTRACIÓN: En la demostración del teorema 2.12 se ve que $\mathrm{Aut}(E)$ es un grupo cíclico de orden n. Por lo tanto, si llamamos C_n al grupo de las raíces n-simas de la unidad de $\overline{k}$, tenemos que $\mathrm{Aut}(E) \cong C_n$ (aquí usamos la hipótesis sobre la característica). Consideramos la sucesión exacta

$$0 \longrightarrow C_n \longrightarrow \overline{k}^* \xrightarrow{n} \overline{k}^* \longrightarrow 0,$$

de la que extraemos la sucesión de cohomología:

$$\longrightarrow k^* \xrightarrow{n} k^* \xrightarrow{\delta} H^1(G(\overline{k}/k), C_n) \longrightarrow H^1(G(\overline{k}/k), \overline{k}^*) \longrightarrow$$

El último grupo es trivial por el teorema de Hilbert-Speiser, luego δ induce un isomorfismo $H^1(G(\overline{k}/k), C_n) \cong k^*/k^{*n}$.

Para calcular explícitamente el isomorfismo, supongamos primeramente que $j \neq 0$, 1728. Si $d \in k^*$ y $\sigma \in G(\overline{k}/k)$, definimos $\chi_d(\sigma) = (\sqrt{d})^\sigma/\sqrt{d} \in \{\pm 1\}$. Es claro que $\chi_d(\sigma)$ no depende de la elección de la raíz cuadrada. De hecho, es fácil ver que $\delta([d]) = [\{\chi_d(\sigma)\}_\sigma]$.

Si llamamos E'/k a la curva indicada en el enunciado, observamos que un isomorfismo $\phi : E' \longrightarrow E$ viene dado por $\phi(X, Y) = (d^{-1}X, d^{-3/2}Y)$, de modo que $\phi^\sigma(X, Y) = (d^{-1}X, \chi_d(\sigma)d^{-3/2}Y)$ y

$$(\phi^{-1} \circ \phi^\sigma)(X, Y) = (X, \chi_d(\sigma)Y) = \chi_d(\sigma)(X, Y).$$

Por consiguiente, la curva del enunciado se corresponde con $[d]$.

Veamos ahora el caso $j(E) = 1728$. Ahora $\mathrm{Aut}(E)$ consta de cuatro automorfismos de la forma $(X, Y) \mapsto (u^2X, u^3Y)$, donde u es una raíz cuarta de la unidad.

Partimos de una clase $[d] \in k^*/k^{*4}$, elegimos $\sqrt[4]{d} \in \overline{k}$ y formamos el cociclo $\xi_\sigma = (\sqrt[4]{d})^\sigma/\sqrt[4]{d}$. Entonces $\delta([d]) = [\{\xi_\sigma\}_\sigma]$. En este caso el isomorfismo es $\phi(X, Y) = (d^{-1/2}X, d^{-3/4}Y)$, con lo que $\phi^\sigma(X, Y) = (\xi_\sigma^2 d^{-1/2}X, \xi_\sigma^3 d^{-3/4}Y)$ y

$$(\phi^{-1} \circ \phi^\sigma)(X, Y) = (\xi_\sigma^2 X, \xi_\sigma^3 Y) = \xi_\sigma(X, Y).$$

El caso $j(E) = 0$ es análogo. ∎

Recíprocamente, es fácil decidir si dos curvas dadas son o no k-isomorfas:

Teorema 2.47 *Sean E/k y E'/k dos curvas elípticas conjugadas sobre un cuerpo k de característica $\neq 2$, 3 definidas mediante ecuaciones canónicas:*

$$E : Y^2 = X^3 + a_4X + a_6, \qquad E' : Y^2 = X^3 + a_4'X + a_6'.$$

Entonces E y E' son k-isomorfas si y sólo si se cumple la condición siguiente:

a) Si $j \neq 0$, 1728, entonces $\sqrt{a_6/a_6'} \in k$.

b) Si $j = 1728$, entonces $\sqrt[4]{a_6/a_6'} \in k$.

c) Si $j = 0$, entonces $\sqrt[6]{a_6/a_6'} \in k$.

DEMOSTRACIÓN: Que las curvas sean conjugadas significa simplemente que tienen el mismo invariante j. Observemos que $j = 0$ equivale a que $a_4 = a_4' = 0$, mientras que $j = 1728$ equivale a que $a_6 = a_6' = 0$.

Los únicos cambios de variables que transforman una ecuación de este tipo en otra son los de la forma $X = u^2X'$, $Y = u^3Y'$. Por lo tanto, las dos curvas serán k-isomorfas si y sólo si existe un $u \in k^*$ tal que

$$a_4' = u^{-4}a_4, \qquad a_6' = u^{-6}a_6. \tag{2.7}$$

En el caso $j \neq 0$, 1728, si existe el isomorfismo entonces $\sqrt{a_6/a_6'} = u^3 \in k$. Recíprocamente, supongamos que $\sqrt{a_6/a_6'} \in k$. La igualdad de los invariantes se traduce en la relación

$$\left(\frac{a_6}{a_6'}\right)^2 = \left(\frac{a_4}{a_4'}\right)^3.$$

En general, si tenemos una relación $v^2 = w^3$, podemos llamar $r = v/w$ y entonces $r^3 = v$. En nuestro caso concluimos que a_6/a_6' tiene raíz cúbica en k. Por otra parte, si $u = v^2 = w^3$, entonces $u = r^6$, con $r = v/w$. Nosotros tenemos que a_6/a_6' tiene raíz cuadrada (por hipótesis) y raíz cúbica en k, luego tiene una raíz sexta: $a_6/a_6' = u^6$. A su vez, $(a_4/a_4')^3 = (u^4)^3$, luego $a_4/a_4' = \omega u^4$, donde $\omega \in k$ cumple $\omega^3 = 1$. Cambiamos u por ωu y entonces se cumple $a_4/a_4' = u^4$ sin dejar de cumplirse $a_6/a_6' = u^6$.

En los casos $j = 0$ ($a_4 = 0$), o $j = 1728$ ($a_6 = 0$) las condiciones (2.7) son exactamente las del enunciado. ∎

Curvas de género 1 Si C/k es una curva proyectiva regular de género 1, no es necesariamente una curva elíptica sobre k, pues para ello hace falta además que tenga un punto racional. En cualquier caso, será una curva elíptica sobre $\overline{k}$ tomando como elemento neutro cualquier punto $O \in C$. Podemos encontrar un isomorfismo $\phi : C \longrightarrow E$, donde E es una curva dada por una ecuación de Weierstrass. Ahora bien, si $\sigma \in G(\overline{k}/k)$ entonces tenemos un isomorfismo $\phi^\sigma : C \longrightarrow E^\sigma$, luego las curvas E y E^σ son isomorfas, luego tienen el mismo invariante $j(E) = j(E)^\sigma$. Esto prueba que $j(E) \in k$. Por el teorema 2.10 existe una curva elíptica E'/k con invariante $j(E)$, luego hemos llegado a que C/k es conjugada con una curva elíptica E'/k.

Sea $\phi : C \longrightarrow E'$ un isomorfismo y consideremos el cociclo $\xi_\sigma = \phi^{-1} \circ \phi^\sigma$. Por el teorema 2.39, podemos descomponer $\xi_\sigma = \zeta_\sigma \tau_{P_\sigma}$, con $\zeta_\sigma \in \mathrm{Aut}(E')$ y $P_\sigma \in E'$. Como la proyección sobre $\mathrm{Aut}(E')$ es un homomorfismo de grupos, es fácil ver que $\{\zeta_\sigma\}_\sigma$ es también un cociclo, el cual determina una curva elíptica E/k conjugada con E. Sea $\psi : E \longrightarrow E'$ un isomorfismo tal que $\psi^{-1} \circ \psi^\sigma = \zeta_\sigma$. Vamos a calcular el cociclo correspondiente al isomorfismo $\phi \circ \psi^{-1} : C \longrightarrow E$:

$$(\phi \circ \psi^{-1})^{-1} \circ (\phi \circ \psi^{-1})^\sigma = \psi \circ \phi^{-1} \circ \phi^\sigma \circ (\psi^{-1})^\sigma = \psi \circ \zeta_\sigma \tau_{P_\sigma} \circ (\psi^{-1})^\sigma$$

$$= \psi^\sigma \circ \tau_{P_\sigma} \circ (\psi^{-1})^\sigma = \tau_{(\psi^{-1})^\sigma(P_\sigma)}.$$

Así pues, C/k se corresponde con una clase de $H^1(G(\overline{k}/k), E)$, y por el teorema 2.44 admite una estructura de espacio homogéneo para E/k. Con esto casi hemos demostrado el teorema siguiente:

Teorema 2.48 *Si C/k es una curva proyectiva regular de género 1, existe una curva elíptica E/k (única salvo k-isomorfismo) tal que C/k admite una estructura de espacio homogéneo para E/k.*

DEMOSTRACIÓN: Sólo falta probar la unicidad. Supongamos que C/k es un espacio homogéneo para las dos curvas E/k y E'/k. Fijemos un punto $p_0 \in C$ y consideremos los isomorfismos $\theta_{p_0} : E \longrightarrow C$, $\theta'_{p_0} : D' \longrightarrow C$ dados por $\theta_{p_0}(P) = p_0 + P$, $\theta'_{p_0}(P') = p_0 + P'$.

Entonces $\phi : \theta_{p_0} \circ \theta'^{-1}_{p_0} : E \longrightarrow E'$ es un isomorfismo de curvas elípticas, pues $\phi(O) = \theta'^{-1}_{p_0}(p_0) = O'$, y está definido sobre k, pues

$$\phi^\sigma(P) = \phi(P^{\sigma^{-1}})^\sigma = \theta'^{-1}_{p_0^\sigma}(\theta_{p_0^\sigma}(P^{\sigma^{-1}}))^\sigma = \theta'^{-1}_{p_0^\sigma}(p_0^\sigma + P^{\sigma^{-1}})^\sigma = P.$$

Así pues, E/k y E'/k son k-isomorfas. ∎

Capítulo III

El álgebra de las curvas elípticas

En este capítulo estudiamos la estructura de grupo de una curva elíptica definida sobre un cuerpo algebraicamente cerrado, junto con las estructuras relacionadas con ella. Por ejemplo, observamos que las isogenias de una curva elíptica en sí misma forman un anillo con la suma definida puntualmente y la composición como producto. En la primera sección definimos este anillo de endomorfismos, pero sólo al llegar a la última sección estaremos en condiciones de determinar su estructura.

3.1 Las multiplicaciones enteras

Definición 3.1 Si E y E' son dos curvas elípticas, llamamos $\mathrm{Hom}(E, E')$ al conjunto de todas las isogenias de E en E', que es un grupo abeliano con la suma definida puntualmente (es obvio que la suma de isogenias es una isogenia). Definimos $\mathrm{End}(E)$ como el conjunto de todas las isogenias de E en E, que es un anillo (no necesariamente conmutativo) tomando como producto la composición de aplicaciones.

De momento lo único que conocemos del anillo de endomorfismos de una curva elíptica es su grupo de unidades, que es el grupo $\mathrm{Aut}(E)$ descrito en el teorema 2.12.

Si E es una curva elíptica y $m \in \mathbb{Z}$, representaremos también por m a la isogenia $m : E \longrightarrow E$ definida por la multiplicación por m en el sentido usual de la teoría de grupos $m \mapsto mP$. Ciertamente es una isogenia (definida sobre k si E lo está). La aplicación $\mathbb{Z} \longrightarrow \mathrm{End}(E)$ que a cada $m \in \mathbb{Z}$ le asigna la multiplicación por m es claramente un homomorfismo de anillos. No es evidente en absoluto, pero vamos a probar que es inyectivo. Para ello necesitaremos el teorema siguiente que, en contra de lo que podría parecer a primera vista, no es una consecuencia inmediata de las definiciones, sino una conexión no trivial entre la suma en una curva elíptica (una operación específica de las curvas

elípticas) y la suma de formas diferenciales (una operación general de las curvas proyectivas).

Teorema 3.2 *Sean $\phi, \psi : E \longrightarrow E'$ dos isogenias entre curvas elípticas y sea ω una diferencial invariante en E'. Entonces*

$$\overline{\phi+\psi}(\omega) = \overline{\phi}(\omega) + \overline{\psi}(\omega).$$

DEMOSTRACIÓN: Adoptamos el convenio de que si ϕ es la isogenia nula entonces $\overline{\phi} : \Omega(E') \longrightarrow \Omega(E)$ es la aplicación nula. Entonces la igualdad es trivialmente cierta si $\phi = 0$ o $\psi = 0$. Supongamos ahora que ambas son no nulas pero que $\phi + \psi = 0$. Entonces $\psi = -\phi$, es decir, ψ es la composición de $\phi : E \longrightarrow E'$ con $-1 : E' \longrightarrow E'$ (la aplicación $Q \mapsto -Q$). Por lo tanto, $\overline{\psi} = \overline{-1} \circ \overline{\phi}$ y en este caso basta probar que $\overline{-1}(\omega) = -\omega$.

Sea $\overline{k}(E') = \overline{k}(x', y')$, donde las funciones x', y' satisfacen una ecuación de Weierstrass. Entonces el teorema 2.21 nos da que

$$\overline{-1}(x') = x', \qquad \overline{-1}(y') = -y' - a_1x' - a_3,$$

y por el teorema 2.14 tenemos que, salvo una constante que podemos despreciar,

$$\omega = \frac{dx'}{2y' + a_1x' + a_3},$$

luego

$$\overline{-1}(\omega) = \frac{dx'}{2(-y' - a_1x' - a_3) + a_1x' + a_3} = -\frac{dx'}{2y' + a_1x' + a_3} = -\omega.$$

Así pues, podemos suponer que las tres isogenias ϕ, ψ y $\phi + \psi$ son no nulas. Como en el caso anterior, sea $\overline{k}(E') = \overline{k}(x', y')$, donde las funciones x', y' satisfacen una ecuación de Weierstrass $F(x', y') = 0$. Según el teorema 2.14, tenemos que

$$\omega = \frac{dx'}{F_Y(x', y')} = -\frac{dy'}{F_X(x', y')}.$$

Sean x_1, y_1, x_2, y_2 dos pares de funciones en $\overline{k}(E)$ que cumplan $F(x_i, y_i) = 0$. Luego tomaremos

$$x_1 = \overline{\phi}(x'), \quad y_1 = \overline{\phi}(y'), \quad x_2 = \overline{\psi}(x'), \quad y_2 = \overline{\psi}(y'), \tag{3.1}$$

pero el argumento requiere razonar primero con funciones arbitrarias. Derivando la ecuación obtenemos que

$$\omega(x_i, y_i) = \frac{dx_i}{F_Y(x_i, y_i)} = -\frac{dy_i}{F_X(x_i, y_i)}.$$

El teorema 2.21 nos da dos funciones racionales

$$R(X_1, Y_1, X_2, Y_2) \qquad \text{y} \qquad S(X_1, Y_1, X_2, Y_2)$$

tales que, para cada par de puntos (P,Q) con $x'(P) \neq x'(Q)$ se cumple

$$x'(P+Q) = R(x'(P), y'(P), x'(Q), y'(Q)),$$
$$y'(P+Q) = S(x'(P), y'(P), x'(Q), y'(Q)).$$

Definimos $x_3 = R(x_1, y_1, x_2, y_2)$, $y_3 = S(x_1, y_1, x_2, y_2) \in \overline{k}(E)$. En el caso particular (3.1) se cumple

$$x_3 = \overline{\phi + \psi}(x'), \qquad y_3 = \overline{\phi + \psi}(y'). \tag{3.2}$$

En general:

$$\begin{aligned} dx_3 &= \frac{\partial R}{\partial x_1}\,dx_1 + \frac{\partial R}{\partial y_1}\,dy_1 + \frac{\partial R}{\partial x_2}\,dx_2 + \frac{\partial R}{\partial y_2}\,dy_2 \\ &= \left(\frac{\partial R}{\partial x_1} + \frac{\partial R}{\partial y_1} - \frac{F_X(x_1,y_1)}{F_Y(x_1,y_1)}\right)dx_1 + \left(\frac{\partial R}{\partial x_2} + \frac{\partial R}{\partial y_2} - \frac{F_X(x_2,y_2)}{F_Y(x_2,y_2)}\right)dx_2. \end{aligned}$$

De aquí llegamos a que

$$\omega(x_3, y_3) = M_1(x_1, y_1, x_2, y_2)\,\omega(x_1, y_1) + M_2(x_1, y_1, x_2, y_2)\,\omega(x_2, y_2), \tag{3.3}$$

para ciertas funciones racionales M_1 y M_2.

Tomemos $x_1 = \overline{\phi}(x')$, $y_1 = \overline{\phi}(y')$, $x_2 = x'(Q)$, $y_2 = y'(Q)$, donde $Q \in E'$ es un punto sobre el que están definidas x' e y'. Notemos que las funciones constantes x_2, y_2 cumplen ciertamente $F(x_2, y_2) = 0$. Con esta elección, para cada $P \in E$ tal que $x'(\phi(P)) \neq x_2$, tenemos que

$$\begin{aligned} x_3(P) &= R(\overline{\phi}(x'), \overline{\phi}(y'), \overline{\phi}(x_2), \overline{\phi}(y_2))(P) = \overline{\phi}(R(x', y', x_2, y_2))(P) \\ &= R(x'(\phi(P), y'(\phi(P), x'(Q), y'(Q)) = x'(\phi(P) + Q) \\ &= x'(\tau_Q(\phi(P))) = \overline{\phi}(\overline{\tau}_Q(x'))(P). \end{aligned}$$

Por lo tanto $x_3 = \overline{\phi}(\overline{\tau}_Q(x'))$, e igualmente $y_3 = \overline{\phi}(\overline{\tau}_Q(y'))$. Teniendo en cuenta el teorema 2.24, de aquí llegamos a que

$$\omega(x_3, y_3) = \overline{\phi}(\overline{\tau}_Q(\omega(x', y'))) = \overline{\phi}(\overline{\tau}_Q(\omega)) = \overline{\phi}(\omega) = \omega(\overline{\phi}(x'), \overline{\phi}(y')) = \omega(x_1, y_1).$$

Por otra parte, como x_2 es constante se cumple $dx_2 = 0$, luego $\omega(x_2, y_2) = 0$ y (3.3) se reduce a

$$\omega(x_1, y_1) = M_1(\overline{\phi}(x'), \overline{\phi}(y'), x'(Q), y'(Q))\,\omega(x_1, y_1),$$

luego $M_1(\overline{\phi}(x'), \overline{\phi}(y'), x'(Q), y'(Q)) = 1$ para todo punto $Q \in E'$ donde estén definidas x' e y'. Obviamente entonces $M_1(\overline{\phi}(x'), \overline{\phi}(y'), \overline{\psi}(x'), \overline{\psi}(y')) = 1$. Análogamente se llega a la misma conclusión para M_2, con lo que (3.1), (3.2) y (3.3) nos dan la relación

$$\omega(\overline{\phi + \psi}(x'), \overline{\phi + \psi}(y')) = \omega(\overline{\phi}(x'), \overline{\phi}(y')) + \omega(\overline{\psi}(x'), \overline{\psi}(y'))$$

o, lo que es lo mismo: $\overline{\phi + \psi}(\omega) = \overline{\phi}(\omega) + \overline{\psi}(\omega)$. ∎

Al aplicar este teorema a las multiplicaciones por enteros obtenemos lo siguiente:

Teorema 3.3 *Si E es una curva elíptica, ω es una diferencial invariante en E y $m \in \mathbb{Z}$, entonces $\overline{m}\omega = m\omega$.*

DEMOSTRACIÓN: El teorema es cierto para $m = 0$ por definición y claramente también para $m = 1$. El teorema anterior nos da las relaciones

$$\overline{m+1}\omega = \overline{m}\omega + \omega, \qquad \overline{-m}\omega = -\overline{m}\omega.$$

La primera nos da la conclusión para $m \geq 0$ por inducción, y la segunda la extiende a los números negativos. ∎

Como consecuencia:

Teorema 3.4 *Sea E una curva elíptica y sea $m \in \mathbb{Z}$, $m \neq 0$. Supongamos que $\operatorname{car}\overline{k} = 0$ o que $\operatorname{car}\overline{k}$ es un primo que no divide a m. Entonces la multiplicación por m es una isogenia separable (no nula).*

DEMOSTRACIÓN: El teorema anterior y las hipótesis sobre la característica implican que $\overline{m}\omega = m\omega \neq 0$, luego la multiplicación por m no es constante y, por el teorema 1.19 es separable. ∎

Teorema 3.5 *Si E es una curva elíptica, entonces la aplicación $\mathbb{Z} \longrightarrow \operatorname{End}(E)$ que a cada entero m le asigna la multiplicación por m es un monomorfismo de anillos.*

DEMOSTRACIÓN: Basta probar que si $p \in \mathbb{Z}$, $p \neq 0$, entonces la multiplicación por p es una isogenia no nula. Observemos que una composición de isogenias no nulas (suprayectivas) es suprayectiva y, por consiguiente no nula. Así pues, podemos suponer que p es primo.

Supongamos primero que $\operatorname{car}\overline{k} \neq 2$. Si $p = 2$ basta aplicar el teorema anterior. Si p es impar, basta probar que E contiene un punto P de orden 2, es decir, tal que $P \neq O$ y $2P = O$. En tal caso, es obvio que $pP \neq 0$.

En efecto, por el teorema 2.7 podemos suponer que E está definida por una ecuación de Weierstrass de la forma $Y^2 = F(X)$. Si $a \in \overline{k}$ es una raíz de $F(X)$, tenemos que $P = (a, 0) \in E$ y la tangente a E en P es $X = a$, de donde se sigue que $P + P = O$.

Supongamos ahora que $\operatorname{car}\overline{k} = 2$. Por el teorema anterior, basta considerar el caso $p = 2$. Si la multiplicación por 2 fuera nula, todos los puntos de E tendrían orden 2, así que basta probar que no es así. De nuevo podemos suponer que E está definida por una ecuación de Weierstrass $F(X, Y) = 0$ de uno de los dos tipos que indica el teorema 2.7 para característica 2. Notemos que para que un punto $P = (a, b) \in E$ tenga orden 2, la tangente a E en P ha de ser vertical, lo cual equivale a que $F_Y(P) = 0$. Considerando las ecuaciones del teorema 2.7, esta condición equivale a $X = 0$ en el primer caso y a $a_3 = 0$ en el segundo caso. La segunda posibilidad se descarta inmediatamente, pues implica $\Delta = 0$. La primera, junto a $F(P) = 0$ nos da $Y^2 = a_6$, luego E tiene únicamente dos puntos de orden 2. ∎

De aquí se deducen a su vez propiedades más generales:

Teorema 3.6 *Si E_1 y E_2 son curvas elípticas, entonces el grupo de isogenias* $\mathrm{Hom}(E_1, E_2)$ *es un $\mathbb{Z}$-módulo libre de torsión. Si E es una curva elíptica, entonces* $\mathrm{End}(E)$ *es un anillo (no necesariamente conmutativo) de característica* 0 *y sin divisores de* 0.

DEMOSTRACIÓN: Según hemos observado en la prueba del teorema anterior, la composición de isogenias no nulas es no nula. Por lo tanto, si $m \in \mathbb{Z}$ y $\phi \in \mathrm{Hom}(E_1, E_2)$ cumplen $m\phi = 0$, esto puede verse como la composición de ϕ con la multiplicación por m, luego $\phi = 0$ o bien $m = 0$. El teorema anterior garantiza que $m = 0$ como isogenia equivale a $m = 0$ como entero, luego no hay elementos de torsión. El mismo razonamiento prueba la segunda parte. ∎

Nota En muchos casos, el anillo de endomorfismos de una curva elíptica es simplemente $\mathrm{End}(E) = \mathbb{Z}$. Cuando no sucede así, se dice que la curva E tiene *multiplicación compleja.* Existe toda una teoría sobre curvas elípticas con multiplicación compleja en la que no vamos a entrar aquí. ∎

Ejemplo Sea $E/\mathbb{Q}$ la curva elíptica dada por $Y^2 = X^3 - X$. Si i es la unidad imaginaria, llamemos también i a la isogenia $(x, y) \mapsto (-x, iy)$. Puesto que (como isogenia) cumple $i^2 = -1$, no es la multiplicación por ningún número entero. Así pues, E tiene multiplicación compleja. Claramente podemos definir de forma natural un homomorfismo de anillos $\mathbb{Z}[i] \longrightarrow \mathrm{End}(E)$. Puede probarse que es un isomorfismo.(Ver el ejemplo de la página 302.) ∎

Seguidamente vamos a calcular el grado de la multiplicación por m. Para ello necesitamos un teorema cuya demostración requiere más geometría algebraica de la que podemos explicar con la suficiente brevedad.[1] La prueba requiere cierta familiaridad con la teoría de divisores sobre superficies algebraicas (concretamente, sobre la superficie $E \times E$, donde E es una curva elíptica).

Teorema 3.7 *Si E es una curva elíptica, existe una aplicación*

$$(\ ,\ ) : \mathrm{End}\, E \times \mathrm{End}\, E \longrightarrow \mathbb{Q}$$

que verifica las propiedades

$$(\phi, \psi) = (\psi, \phi), \quad (\phi + \chi, \psi) = (\phi, \psi) + (\chi, \psi)$$

y además $(\phi, \phi) = \mathrm{grad}\, \phi$, para toda isogenia ϕ.

De aquí se sigue inmediatamente la relación

$$\mathrm{grad}(\phi + \psi) + \mathrm{grad}(\phi - \psi) = 2(\mathrm{grad}\, \phi + \mathrm{grad}\, \psi).$$

[1] En el último capítulo de mi Geometría Algebraica se da una prueba detallada. Allí figura la hipótesis de que el cuerpo de constantes tenga característica distinta de 2 o 3, pero la prueba vale literalmente sin esta hipótesis, pues sólo se usa al apelar a resultados previos que aquí hemos demostrado en general.

Si la aplicamos a $\phi = m \geq 1$ y $\psi = 1$, obtenemos

$$\operatorname{grad}(m+1) + \operatorname{grad}(m-1) = 2(\operatorname{grad} m + 1),$$

y ahora una simple inducción demuestra el teorema siguiente:

Teorema 3.8 *Si E es una curva elíptica y $m \in \mathbb{Z}$, entonces el grado de la multiplicación por m en E es* $\operatorname{grad} m = m^2$.

DEMOSTRACIÓN: Para $m = 0$ y $m = 1$ es obvio. Para $m > 1$ se prueba por inducción a partir de la fórmula previa al teorema. Para $m < 0$ basta usar que -1 es un automorfismo, luego tiene grado 1. ∎

Definición 3.9 Si E es una curva elíptica y $m \in \mathbb{N}$, definimos

$$E[m] = \{P \in E \mid mP = 0\},$$

de modo que

$$E_t = \bigcup_{n=1}^{\infty} E[m]$$

es el subgrupo de torsión de E (el subgrupo formado por los elementos de orden finito).

Observemos que $E[m]$ es el núcleo de la multiplicación por m. Los teoremas 2.35 y 3.4 implican que si m no es divisible entre la característica del cuerpo de constantes entonces el grupo $E[m]$ tiene m^2 elementos. Más precisamente, bajo dicha hipótesis se cumple que

$$E[m] \cong (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z}).$$

En efecto, si descomponemos $E[m]$ en producto de grupos cíclicos de orden potencia de primo, cada primo $p \mid m$ no puede aparecer más que en el orden de dos factores, pues de lo contrario $|E[p]| \geq p^3$. Por otra parte, si $m = p^r m'$, con $(m, m') = 1$, entonces $E[m]$ no puede contener un subgrupo de orden p^{2r}, luego ha de haber exactamente dos factores de orden p^r.

Si la característica del cuerpo de constantes es un primo p y $m = p^r m'$ con $(m, m') = 1$, es evidente que

$$E[m] = E[p^r] \oplus E[m'].$$

Queda pendiente, pues, estudiar el orden de los subgrupos $E[p^r]$ cuando p es la característica de k. Para ello necesitamos la noción de isogenia dual, de la que nos ocupamos en la sección siguiente.

3.2 La isogenia dual

Consideremos una isogenia no nula $\phi : E_1 \longrightarrow E_2$ entre dos curvas elípticas. La prueba de que las isogenias son homomorfismos de grupos se basa en que ϕ se corresponde con el homomorfismo $\phi : H_0(E_1) \longrightarrow H_0(E_2)$ que induce entre los grupos de clases de grado 0 (que esencialmente es la norma). Ahora bien, ϕ también induce un homomorfismo $\overline{\phi} : H_0(E_2) \longrightarrow H_0(E_1)$ (la inclusión). Componiéndolo con los isomorfismos naturales entre las curvas y sus grupos de clases obtenemos un homomorfismo $\hat{\phi} : E_2 \longrightarrow E_1$. No es trivial, pero vamos a probar que es una isogenia no nula.

Teorema 3.10 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia no nula de grado m.*

a) Existe una única isogenia $\hat{\phi} : E_2 \longrightarrow E_1$ tal que $\phi \circ \hat{\phi} = m$.

b) La isogenia $\hat{\phi}$ es la composición de los homomorfismos

$$E_2 \longrightarrow H_0(E_2) \xrightarrow{\overline{\phi}} H_0(E_1) \longrightarrow E_1,$$

donde la primera y la última flecha representan los isomorfismos naturales.

Demostración: Veamos primero la unicidad. Si $\hat{\phi}$ y $\hat{\phi}'$ cumplen a), entonces $\phi \circ (\hat{\phi} - \hat{\phi}') = m - m = 0$. Como ϕ no es nula, ha de serlo $\hat{\phi} - \hat{\phi}'$.

Supongamos ahora que tenemos dos isogenias $E_1 \xrightarrow{\phi} E_2 \xrightarrow{\psi} E_3$ de grados m y n y que sabemos que $\hat{\phi}$ y $\hat{\psi}$ existen. Entonces $\hat{\psi} \circ \hat{\phi}$ cumple a) para $\phi \circ \psi$. En efecto:

$$\phi \circ \psi \circ \hat{\psi} \circ \hat{\phi} = \phi \circ m \circ \hat{\phi} = m \circ \phi \circ \hat{\phi} = m \circ n = mn.$$

Así pues, en virtud del teorema 1.22, basta probar la existencia de $\hat{\phi}$ bajo el supuesto de que ϕ es separable o una aplicación de Frobenius. Supongamos primero que ϕ es separable. Entonces su núcleo $\mathrm{N}(\phi)$ es un grupo finito de orden m (por el teorema 2.35), luego todos sus elementos tienen orden divisor de m, es decir, $\mathrm{N}(\phi) \subset \mathrm{N}(m)$. Ahora basta aplicar el teorema 2.36, que nos da una isogenia $\hat{\phi} : E_2 \longrightarrow E_1$ tal que $\phi \circ \hat{\phi} = m$.

Ahora supongamos que ϕ es una aplicación de Frobenius. Es claro que la aplicación de Frobenius para $m = p^r$ es la composición r veces consigo misma de la aplicación de Frobenius para p, luego podemos suponer que $\phi : E \longrightarrow E^{(p)}$, donde p es la característica de $\overline{k}$.

Los teoremas 1.19 y 3.3 implican que la multiplicación por p no es separable, ya que si ω es una diferencial invariante de E entonces $\overline{p}\omega = p\omega = 0$. Por lo tanto, la descomposición de p según el teorema 1.22 es de la forma $p = \phi^e \circ \psi$, donde ψ es una isogenia separable y $e \geq 1$. Podemos tomar $\hat{\phi} = \phi^{e-1} \circ \psi$.

Con esto queda demostrado a). Para probar b) llamamos $\hat{\phi}'$ a la composición indicada y vamos a ver que $\phi \circ \hat{\phi}' = m$. En efecto, si $P \in E_1$ tenemos que

$$P \mapsto \phi(P) \mapsto [\phi(P)/O] \mapsto \sum_{P' \in \phi^{-1}[\phi(P)]} e_\phi(P')P' - \sum_{T \in \phi^{-1}[O]} e_\phi(T)T$$

$$= \operatorname{grad}_i \phi \left(\sum_{T \in \mathrm{N}(\phi)} (P+T) - \sum_{T \in \mathrm{N}(\phi)} T \right) = (\operatorname{grad} \phi)P.$$

Aquí hemos usado el teorema 2.35, apartados a) y b).

Por lo tanto, para todo $Q \in E_2$, podemos elegir $P \in E_1$ tal que $\phi(P) = Q$ y entonces $\hat{\phi}(Q) = \hat{\phi}(\phi(P)) = mP = \hat{\phi}'(\phi(P)) = \hat{\phi}'(Q)$. En suma: $\hat{\phi}' = \hat{\phi}$. ∎

Definición 3.11 Si $\phi : E_1 \longrightarrow E_2$ es una isogenia de grado m entre dos curvas elípticas, definimos la *isogenia dual* $\hat{\phi} : E_2 \longrightarrow E_1$ como la única isogenia que cumple $\phi \circ \hat{\phi} = m$ si $m \neq 0$ y $\hat{\phi} = 0$ si $m = 0$ (es decir, si ϕ es la isogenia nula).

Si la isogenia ϕ está definida sobre un cuerpo k, entonces $\hat{\phi}$ también lo está (porque obviamente la multiplicación por m lo está, y la unicidad de la definición implica que ϕ es invariante por $G(\overline{k}/k)$).

Diremos que dos curvas elípticas E_1 y E_2 (definidas sobre k) son *isógenas* (sobre k) si existe una isogenia no nula $\phi : E_1 \longrightarrow E_2$ (definida sobre k). La existencia de la isogenia dual prueba que esta relación es de equivalencia, lo cual no es evidente a partir de la mera definición.

Ejemplo Un cálculo rutinario muestra que las isogenias del ejemplo de la página 61 son duales. Basta comprobar que $\phi \circ \hat{\phi} = 2$. ∎

El teorema siguiente recoge las propiedades básicas de la isogenia dual:

Teorema 3.12 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia de grado m entre dos curvas elípticas.*

a) Se cumple que $\phi \circ \hat{\phi} = m$ en E_1 y $\hat{\phi} \circ \phi = m$ en E_2.

b) Si $\phi : E_2 \longrightarrow E_3$ es otra isogenia, entonces $\widehat{\phi \circ \psi} = \hat{\psi} \circ \hat{\phi}$.

c) En una curva elíptica E, la multiplicación por $m \in \mathbb{Z}$ cumple $\hat{m} = m$.

d) $\operatorname{grad} \phi = \operatorname{grad} \hat{\phi}$.

e) $\hat{\hat{\phi}} = \phi$.

Demostración: Si ϕ es nula todas las propiedades se cumplen trivialmente, así que supondremos que no lo es.

a) La primera relación es la definición de isogenia dual. La segunda resulta de cancelar ϕ en la igualdad siguiente:

$$\phi \circ (\hat{\phi} \circ \phi) = m \circ \phi = \phi \circ m.$$

b) Consecuencia inmediata de la definición.

c) Puesto que $m \circ m = m^2 = \operatorname{grad} m$, la unicidad de la definición implica que $\hat{m} = m$.

d) Basta observar que

$$m^2 = \operatorname{grad} m = \operatorname{grad}(\phi \circ \hat{\phi}) = (\operatorname{grad} \phi)(\operatorname{grad} \hat{\phi}) = m \operatorname{grad} \hat{\phi}.$$

e) Consecuencia inmediata de a), d) y la unicidad de la definición. ∎

3.3 Curvas supersingulares

Ahora podemos estudiar la estructura de los grupos de torsión $E[p^e]$, donde E es una curva elíptica sobre un cuerpo de característica prima p. Sucede que hay dos posibilidades:

Teorema 3.13 *Sea E una curva elíptica sobre un cuerpo de característica p. Entonces $E[p^e] = 0$ para todo $e \geq 0$ o bien $E[p^e] \cong \mathbb{Z}/p^e\mathbb{Z}$ para todo $e \geq 0$.*

Demostración: Sea $\phi : E \longrightarrow E$ la aplicación de Frobenius de orden p con lo que ϕ^e es la aplicación de Frobenius de orden p^e. Puesto que $E[p^e]$ es el núcleo de la multiplicación por p^e, el teorema 2.35 nos da que

$$|E[p^e]| = \operatorname{grad}_s p^e = \operatorname{grad}_s(\phi^e \circ \widehat{\phi^e}) = \operatorname{grad}_s(\phi^e \circ \hat{\phi}^e) = (\operatorname{grad}_s \hat{\phi})^e.$$

Ahora basta tener en cuenta que $\hat{\phi}$ tiene grado p, luego su grado de separabilidad sólo puede ser 1 o p.

Si $E[p^e]$ tiene orden p^e, entonces tiene que ser un grupo cíclico, pues si no tuviera elementos de orden p^e tendríamos que $E[p^e] = E[p^{e-1}]$, lo cual es imposible por la parte ya probada. ∎

Definición 3.14 Sea E una curva elíptica definida sobre un cuerpo de característica prima p. Diremos que E es *supersingular* si $E[p] = 0$ y que E es *ordinaria* en caso contrario, es decir, si $E[p]$ tiene orden p.

Observemos que el término "supersingular" es capcioso, pues las curvas supersingulares no son singulares[2] (son curvas elípticas). Según el teorema anterior, si E es una curva supersingular se cumple que todos los grupos de torsión $E[p^e]$ son triviales, mientras que si E es ordinaria entonces $E[p^e] \cong \mathbb{Z}/p^e\mathbb{Z}$.

El teorema siguiente está implícito en la demostración de 3.13:

[2]El nombre está relacionado con los anillos de endomorfismos. Lo "más normal" es que el anillo de endomorfismos de una curva elíptica sea isomorfo a $\mathbb{Z}$, un caso menos frecuente es que sea isomorfo a un orden de un cuerpo cuadrático imaginario, y a las curvas con dicha propiedad se las llama "singulares". Veremos que las curvas supersingulares sobre cuerpos finitos tienen anillos de endomorfismos aún mayores.

Teorema 3.15 *Sea E una curva elíptica sobre un cuerpo de característica prima p y sea $\phi_e : E \longrightarrow E^{(p^e)}$ la aplicación de Frobenius de grado p^e. Las afirmaciones siguientes son equivalentes:*

a) E es supersingular.

b) $\hat{\phi}_e$ es (puramente) inseparable para un (para todo) $e \geq 1$.

c) La multiplicación $p : E \longrightarrow E$ es puramente inseparable.

DEMOSTRACIÓN: En la prueba de 3.13 hemos visto que el carácter supersingular u ordinario de E depende de si $\hat{\phi}_1$ tiene grado de separabilidad 1 o p, lo que equivale a que $\hat{\phi}_e$ tenga grado de separabilidad 1 o p^e (es decir, a que sea puramente inseparable o separable). Esto prueba la equivalencia entre a) y b). También hemos visto que $|E[p]| = \operatorname{grad}_s p$, lo que nos da la equivalencia entre a) y c). ∎

Es obvio que el carácter supersingular u ordinario de una curva elíptica es invariante por isomorfismos, luego depende únicamente de su invariante j. Podemos hablar así de invariantes supersingulares y ordinarios. El teorema siguiente demuestra que realmente existen curvas de ambos tipos:

Teorema 3.16 *En cuerpos de característica $p = 2$, el único invariante supersingular es $j = 0$.*

DEMOSTRACIÓN: Según la tabla del teorema 2.7, una curva con invariante $j = 0$ es, por ejemplo,

$$E : Y^2 + Y = X^3.$$

Un punto no nulo en $E[2]$ ha de tener tangente vertical, es decir, la derivada parcial de la ecuación respecto de Y ha de ser 0, pero dicha derivada es constante igual a 1, luego $E[2] = 0$ y E es supersingular.

Por el contrario, una curva con invariante no nulo admite una ecuación de la forma

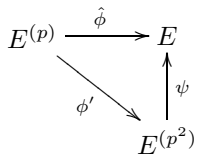
$$E : Y^2 + XY = X^3 + a_6,$$

la cual tiene un único punto de orden 2, a saber, $(0, \sqrt{a_6})$, luego $E[2] \neq 0$ y E es ordinaria. ∎

Más en general, el teorema siguiente muestra que el número de invariantes supersingulares (en una característica fija) es siempre finito:

Teorema 3.17 *Si E es una curva elíptica supersingular sobre un cuerpo de característica prima p, entonces $j(E)$ está en el cuerpo de p^2 elementos.*

DEMOSTRACIÓN: Sea $\hat{\phi} : E^{(p)} \longrightarrow E$ la isogenia dual de la aplicación de Frobenius de orden p. Tenemos que $\hat{\phi}$ es puramente inseparable de grado p, luego por el teorema 1.22 se descompone en la forma



donde ϕ' es la aplicación de Frobenius de $E^{(p)}$ y ψ tiene grado 1, luego es un isomorfismo. Por consiguiente, $j(E) = j(E^{(p^2)})$.

Ahora bien, si tomamos una ecuación de Weierstrass para E, tenemos que una ecuación de Weierstrass para $E^{(p^2)}$ se obtiene elevando a p^2 todos los coeficientes de la primera, luego $j(E^{(p^2)}) = j(E)^{p^2}$. En definitiva, hemos probado que $j(E) = j(E)^{p^2}$, y esto implica que $j(E)$ está en el cuerpo de p^2 elementos. ∎

Por último observamos que la supersingularidad es también invariante por isogenia:

Teorema 3.18 *Toda curva isógena a una curva supersingular es supersingular.*

DEMOSTRACIÓN: Sea $\psi : E_1 \longrightarrow E_2$ una isogenia no nula entre dos curvas definidas sobre un cuerpo de característica prima p. Entonces $p\circ\psi = \psi\circ p$. Si E_1 es supersingular, entonces p es puramente inseparable en E_1, luego, comparando los grados en la igualdad anterior, también lo es sobre E_2, luego también ésta es supersingular. ∎

3.4 Los módulos de Tate

En las secciones anteriores hemos estudiado los grupos de torsión $E[m]$ de una curva elíptica E. Estos grupos contienen mucha información sobre E. Por ejemplo, observemos que si dos isogenias coinciden sobre el grupo de torsión E_t, entonces son iguales, pues el núcleo de la diferencia es infinito. En realidad vemos que es suficiente con que coincidan sobre el grupo

$$E[q^\infty] = \bigcup_{e=0}^{\infty} E[q^e],$$

donde q es un primo distinto de la característica del cuerpo de constantes. En esta sección veremos cómo "encajar" los grupos $E[q^e]$ en una estructura más conveniente que su unión. Para ello necesitamos la noción de límite proyectivo de un sistema de módulos:

Definición 3.19 Si A es un anillo conmutativo y unitario, un *sistema proyectivo* de A-módulos es una sucesión $\{M_n\}_{n=1}^{\infty}$ de A-módulos junto con una sucesión de homomorfismos $\phi_n : M_n \longrightarrow M_{n-1}$. Definimos el *límite proyectivo* del sistema como el submódulo $M = \varprojlim_n M_n$ del producto $\prod_n M_n$ formado por las sucesiones

$$x = (x_1, x_2, x_3, \dots)$$

tales que $\phi_n(x_n) = x_{n-1}$, para todo $n > 1$. Llamaremos $\pi_n : M \longrightarrow M_n$ a las restricciones de las proyecciones. Obviamente $\pi_{n+1} \circ \phi_{n+1} = \pi_n$.

Ejercicio: Demostrar que si $(M', \{\pi'_n\})$ es otro módulo junto con homomorfismos $\pi'_n : M' \longrightarrow M_n$ tales que $\pi'_{n+1} \circ \phi_{n+1} = \pi'_n$, entonces existe un único homomorfismo $\phi : M' \longrightarrow M$ tal que $\phi \circ \pi_n = \pi'_n$, así como que esta propiedad determina el límite proyectivo salvo isomorfismo.

Es claro que si las aplicaciones ϕ_n son suprayectivas lo mismo sucede con las proyecciones π_n. Sin más que cambiar "módulo" por "anillo" podemos definir el límite proyectivo de un sistema proyectivo de anillos y homomorfismos de anillos.

Ejemplo Sea p un número primo y consideremos el sistema proyectivo formado por los anillos $\mathbb{Z}/p^n\mathbb{Z}$ con los epimorfismos naturales $\mathbb{Z}/p^{n+1}\mathbb{Z} \longrightarrow \mathbb{Z}/p^n\mathbb{Z}$ dados por $[x] \mapsto [x]$. Llamemos $\mathbb{Z}_p$ al límite proyectivo.

Se trata de un anillo conmutativo y unitario, cuya unidad es $1 = (1, 1, 1, \dots)$. Además, para todo $m \in \mathbb{Z}$ no nulo tenemos que $m1 = ([m], [m], \dots) \neq 0$, pues si $m1 = 0$ entonces $p^n \mid m$ para todo $n \geq 1$. Así pues, $\mathbb{Z}_p$ tiene característica 0.

Es claro que una sucesión $([x_1], [x_2], [x_3], \dots)$ es una unidad de $\mathbb{Z}_p$ si y sólo si $p \nmid x_n$ para todo n, pero esto equivale a que $p \nmid x_1$. Así pues, el grupo de las unidades de $\mathbb{Z}_p$ es

$$U_p = \{x \in \mathbb{Z}_p \mid \pi_1(x) \neq 0\}.$$

Para cada $x \in \mathbb{Z}_p$ no nulo, definimos $v_p(x)$ como el mínimo natural m tal que $\pi_{m+1}(x) \neq 0$. Si $v_p(x) = m$ y $\pi_n(x) = [x_n]$, entonces, para todo $n > m$,

$$x_n \equiv 0 \ (\text{mód } p^m) \qquad \text{y} \qquad x_n \not\equiv 0 \ (\text{mód } p^{m+1}),$$

luego $x_n = y_n p^m$, con $(y_n, p) = 1$. Obviamente, si $n \leq m$ podemos elegir $x_n = x_{m+1}$ y se cumple lo mismo. Entonces $\epsilon = (y_1, y_2, \dots) \in U_p$ y $x = \epsilon p^m$.

En resumen, hemos probado que todo elemento $x \in \mathbb{Z}_p$ no nulo se expresa como $x = \epsilon p^m$, donde $\epsilon \in U_p$ y $m \geq 0$. La descomposición es única, pues necesariamente $m = v_p(x)$.

De aquí se sigue inmediatamente que $\mathbb{Z}_p$ es un dominio íntegro, pues un producto de elementos no nulos es de la forma $(\epsilon p^m)(\epsilon' p^n)$. Si fuera nulo, entonces sería $p^{m+n} = 0$, lo cual es imposible porque $\mathbb{Z}_p$ tiene característica 0.

Llamemos $\mathbb{Q}_p$ al cuerpo de cocientes de $\mathbb{Z}_p$. Es fácil ver que v_p se extiende a una valoración en $\mathbb{Q}_p$ cuyo anillo de enteros es $\mathbb{Z}_p$ y que restringida a $\mathbb{Q}$ es la valoración p-ádica. Vamos a probar que $\mathbb{Q}_p$ es completo con dicha valoración y que $\mathbb{Q}$ es denso en $\mathbb{Q}_p$. De este modo, $\mathbb{Q}_p$ resultará ser el cuerpo de los números p-ádicos.

En efecto, si $\{x_n\}$ es una sucesión de Cauchy en $\mathbb{Z}_p$, para cada $r \geq 0$ existe un $m \geq 0$ tal que si $n \geq m$ entonces $v_p(x_n - x_m) \geq r$, lo que implica que $\pi_r(x_n) = \pi_r(x_m)$. En particular, la sucesión $\pi_r(x_n)$ es finalmente constante igual a un cierto $y_r \in \mathbb{Z}/p^r\mathbb{Z}$. En realidad, la condición de Cauchy implica que para cada $r \geq 0$ existe un $m \geq 0$ tal que si $n \geq m$ entonces $\pi_s(x_n) = y_s$ para todo $s \leq r$. De aquí se sigue que $y = (y_1, y_2, \dots) \in \mathbb{Z}_p$, así como que $y = \lim_n x_n$.

Hemos probado que toda sucesión de Cauchy en $\mathbb{Z}_p$ es convergente (lo cual prueba que $\mathbb{Z}_p$ es cerrado en $\mathbb{Q}_p$), pero necesitamos demostrar que lo mismo vale para toda sucesión de Cauchy en $\mathbb{Q}_p$. Ahora bien, toda sucesión de Cauchy

$\{x_n\}$ está acotada, luego existe un $r \in \mathbb{Z}$ tal que $v_p(x_n) \geq r$ para todo $n \geq 0$. Entonces es claro que $\{x_n p^{-r}\}$ es una sucesión de Cauchy en $\mathbb{Z}_p$, que converge a un $y \in \mathbb{Z}_p$, y la sucesión original converge a yp^r.

Por último, si $x \in \mathbb{Z}_p$ y $\pi_n(x) = [x_n]$ entonces $v_p(x - x_n) \geq n$, luego $\mathbb{Z}$ es denso en $\mathbb{Z}_p$. Podemos identificar al cuerpo de los números p-ádicos con la clausura de $\mathbb{Q}$ en $\mathbb{Q}_p$, y entonces el anillo de los enteros p-ádicos es la clausura de $\mathbb{Z}$ en $\mathbb{Q}_p$, o sea, $\mathbb{Z}_p$, según acabamos de probar. Por consiguiente, $\mathbb{Q}_p$ es el cuerpo de cocientes del anillo de los enteros p-ádicos, y éste es el cuerpo de los números p-ádicos. ∎

Observemos que la topología de $\mathbb{Z}_p$ puede describirse fácilmente en términos de su estructura de límite proyectivo: Una base de $\mathbb{Z}_p$ la forman los conjuntos

$$B(n,a) = \{x \in \mathbb{Z}_p \mid \pi_n(x) = a\}, \qquad n \geq 1, \quad a \in \mathbb{Z}/p^n\mathbb{Z}.$$

En efecto, basta observar que si $x_0 \in \mathbb{Z}_p$ cumple $\pi_n(x_0) = a$, entonces

$$B(n,a) = \{x \in \mathbb{Z}_p \mid v_p(x - x_0) > n\}.$$

Definición 3.20 Sea E una curva elíptica y sea $l \in \mathbb{Z}$ un número primo. El *módulo de Tate* l-ádico de E es el límite proyectivo de $\mathbb{Z}$-módulos

$$T_l(E) = \varprojlim_n E[l^n],$$

respecto de los homomorfismos naturales $P \mapsto lP$.

En principio $T_l(E)$ es un $\mathbb{Z}$-módulo, pero podemos dotarlo de una estructura mejor: tenemos que $E[l^n]$ es un $\mathbb{Z}/l^n\mathbb{Z}$-módulo, luego si

$$\alpha = (a_1, a_2, \ldots) \in \mathbb{Z}_l, \qquad x = (P_1, P_2, \ldots) \in T_l(E),$$

podemos definir

$$\alpha x = (a_1 P_1, a_2 P_2, \ldots) \in T_l(E),$$

y es claro que así $T_l(E)$ se convierte en un $\mathbb{Z}_l$-módulo. Su estructura es fácil de determinar:

Supongamos primero que l es distinto de la característica del cuerpo de constantes. Las aplicaciones $E[l^{e+1}] \longrightarrow E[l^e]$ tienen núcleo de orden l^2, luego son suprayectivas. Es claro que si P y P' forman una base de $E[l^{e+1}]$ como $\mathbb{Z}/l^{e+1}\mathbb{Z}$-módulo, entonces los puntos lP y lP' forman una base de $E[l^e]$ como $\mathbb{Z}/l^e\mathbb{Z}$-módulo. Más aún, si fijamos una base en $E[l^e]$, aplicando a P y P' un cambio de base adecuado podemos exigir que lP y lP' sea la base prefijada.

Fijamos una base P_1, P'_1 de $E[l]$ y a partir de ella ir formamos bases P_e, P'_e de $E[l^e]$, de modo que $P_e = lP_{e+1}$, $P'_e = lP'_{e+1}$. Con ellas podemos formar dos elementos P, $P' \in T_l(E)$ tales que $T_l(E) = \langle P \rangle \oplus \langle P' \rangle$. Así pues, $T_l(E) \cong \mathbb{Z}_l \times \mathbb{Z}_l$.

Si el cuerpo de constantes tiene característica l podemos razonar igualmente, con lo que llegamos al teorema siguiente:

Teorema 3.21 *Si E es una cueva elíptica sobre un cuerpo k y l es un número primo, entonces $T_l(E) \cong \mathbb{Z}_l \times \mathbb{Z}_l$ si $l \neq \operatorname{car} k$, mientras que si $l = \operatorname{car} k$ entonces $T_l(E) \cong \mathbb{Z}_l$ o bien $T_l(E) = 0$ (según si E es ordinaria o supersingular).*

Definimos en $T_l(E)$ la topología que tiene por base a los conjuntos de la forma

$$B(n, P) = \{x \in T_l(E) \mid \pi_n(x) = P\}, \qquad n \geq 1, \quad P \in E[l^n].$$

Si $l \neq \operatorname{car} k$ y fijamos una $\mathbb{Z}_l$-base P, P' de $T_l(E)$, entonces el isomorfismo inducido $\mathbb{Z}_l \times \mathbb{Z}_l \longrightarrow T_l(E)$ resulta ser un homeomorfismo cuando en $\mathbb{Z}_l \times \mathbb{Z}_l$ consideramos la topología producto. En efecto, una base de $\mathbb{Z}_l \times \mathbb{Z}_l$ viene dada por los conjuntos

$$\{(\alpha, \beta) \in \mathbb{Z}_l \times \mathbb{Z}_l \mid \pi_n(\alpha) = a,\ \pi_n(\beta) = b\}, \qquad n \geq 1,\ a,\ b \in \mathbb{Z}/l^n\mathbb{Z},$$

que claramente se corresponden con los abiertos básicos

$$\{x \in T_l(E) \mid \pi_n(x) = aP_n + bP'_n\}.$$

Igualmente se razona que si $l = \operatorname{car} k$ y $T_l(E) \neq 0$ entonces $T_l(E)$ es topológicamente isomorfo a $\mathbb{Z}_l$.

Si $\phi : E_1 \longrightarrow E_2$ es una isogenia entre dos curvas elípticas, es claro que ϕ se restringe a homomorfismos de $\mathbb{Z}/l^n\mathbb{Z}$-módulos $\phi_n : E_1[l^n] \longrightarrow E_2[l^n]$, los cuales inducen un homomorfismo de $\mathbb{Z}_l$-módulos $\phi_l : T_l(E_1) \longrightarrow T_l(E_2)$. Tenemos así un monomorfismo de grupos

$$\operatorname{Hom}(E_1, E_2) \longrightarrow \operatorname{Hom}(T_l(E_1), T_l(E_2)).$$

Es fácil ver que es inyectivo (supuesto que $l \neq \operatorname{car} k$), pero necesitamos un hecho más fuerte. Observemos que $\operatorname{Hom}(E_1, E_2)$ es simplemente un grupo abeliano, mientras que $\operatorname{Hom}(T_l(E_1), T_l(E_2))$ tiene una estructura natural de $\mathbb{Z}_l$-módulo con las operaciones definidas puntualmente.

Teorema 3.22 *Si E_1 y E_2 son curvas elípticas sobre un cuerpo de característica distinta de l, entonces el homomorfismo natural de $\mathbb{Z}_l$-módulos*

$$\mathbb{Z}_l \otimes_{\mathbb{Z}} \operatorname{Hom}(E_1, E_2) \longrightarrow \operatorname{Hom}(T_l(E_1), T_l(E_2))$$

es inyectivo.

Demostración: En primer lugar demostraremos que si M es un subgrupo finitamente generado de $\operatorname{Hom}(E_1, E_2)$, entonces el grupo

$$M^* = \{\phi \in \operatorname{Hom}(E_1, E_2) \mid m\phi \in M \text{ para algún } m \geq 1\}$$

también es finitamente generado.

En efecto, observemos que por 3.6 el $\mathbb{Z}$-módulo M es libre de torsión, luego es un $\mathbb{Z}$-módulo libre de rango finito. Por consiguiente $\mathbb{R} \otimes_{\mathbb{Z}} M$ es un espacio

vectorial real de dimensión finita. La aplicación $(\ ,\ )$ definida en 3.7 se extiende a una forma bilineal en $\mathbb{R}\otimes_{\mathbb{Z}} M$, que a su vez nos da una extensión continua de la aplicación grado. Por lo tanto, podemos considerar el conjunto

$$U = \{\phi \in \mathbb{R}\otimes_{\mathbb{Z}} M \mid \operatorname{grad}\phi < 1\},$$

que es un entorno abierto de 0.

Por otra parte, si $\phi \in M^*$ y $m\phi = \psi \in M$, la aplicación $\phi \mapsto \frac{1}{m}\otimes\psi$ nos permite identificar a M^* con un subgrupo de $\mathbb{R}\otimes_{\mathbb{Z}} M$.

Obviamente $M^* \cap U = \{0\}$, luego M^* es un subgrupo discreto de $\mathbb{R}\otimes_{\mathbb{Z}} M$. Esto implica que es finitamente generado (es un retículo).

Usando una vez más que $\operatorname{Hom}(E_1, E_2)$ es libre de torsión concluimos que, de hecho, M^* es un $\mathbb{Z}$-módulo libre.

Pasemos ya a la prueba del teorema. Si $\phi \in \mathbb{Z}_l \otimes_{\mathbb{Z}} \operatorname{Hom}(E_1, E_2)$ tiene imagen nula en $\operatorname{Hom}(T_l(E_1), T_l(E_2))$, podemos tomar un subgrupo M de $\operatorname{Hom}(E_1, E_2)$ finitamente generado tal que $\phi \in \mathbb{Z}_l \otimes_{\mathbb{Z}} M$. Construimos el subgrupo M^* y fijamos una base $\phi_1, \ldots, \phi_t \in \operatorname{Hom}(E_1, E_2)$. Pongamos que

$$\phi = \alpha_1\phi_1 + \cdots + \alpha_t\phi_t, \qquad \alpha_i \in \mathbb{Z}_l.$$

Entonces, la imagen de ϕ en $\operatorname{Hom}(T_l(E_1), T_l(E_2))$ es

$$\phi_l = \alpha_1(\phi_1)_l + \cdots + \alpha_t(\phi_t)_l = 0.$$

Elijamos $a_i \in \mathbb{Z}$ tales que $\pi_n(\alpha_i) = [a_i]$. Para cada $P \in E_1[l^n]$, tomamos $x \in T_l(E_1)$ tal que $\pi_n(x) = P$, y entonces

$$\pi_n(\phi_l(x)) = a_1\phi_1(P) + \cdots + a_t\phi_t(P) = 0,$$

luego la isogenia

$$\psi = a_1\phi_1 + \cdots + a_t\phi_t \in \operatorname{Hom}(E_1, E_2)$$

se anula sobre $E_1[l^n]$. Según el teorema 2.36 existe $\lambda \in \operatorname{Hom}(E_1, E_2)$ tal que $\psi = l^n\lambda$, pero entonces $\lambda \in M^*$, luego

$$\lambda = b_1\phi_1 + \cdots + b_t\phi_t, \qquad b_i \in \mathbb{Z}.$$

Por la unicidad de las coordenadas, ha de ser $a_i = l^n b_i$, luego $\pi_n(\alpha_i) = 0$ para todo n, luego $\alpha_i = 0$ y $\phi = 0$. ∎

Ahora podemos precisar la estructura de los grupos de isogenias:

Teorema 3.23 *Si E_1 y E_2 son dos curvas elípticas, entonces* $\operatorname{Hom}(E_1, E_2)$ *es un $\mathbb{Z}$-módulo libre de rango menor o igual que* 4.

DEMOSTRACIÓN: Fijemos un primo l distinto de la característica del cuerpo de constantes. Sabemos que $T_l(E_1)$ y $T_l(E_2)$ son $\mathbb{Z}_l$-módulos libres de rango 2, luego $\mathrm{Hom}(T_l(E_1), T_l(E_2))$ es isomorfo al grupo de matrices 2×2 con coeficientes en $\mathbb{Z}_l$, luego es un $\mathbb{Z}_l$-módulo libre de rango 4.

El teorema anterior nos permite ver a $\mathbb{Z}_l \otimes_{\mathbb{Z}} \mathrm{Hom}(E_1, E_2)$ como $\mathbb{Z}_l$-submódulo de $\mathrm{Hom}(T_l(E_1), T_l(E_2))$, luego es libre y de rango menor o igual que 4. Por consiguiente, todo submódulo finitamente M generado de $\mathrm{Hom}(E_1, E_2)$ es libre y determina un $\mathbb{Z}_l$-submódulo $\mathbb{Z}_l \otimes_{\mathbb{Z}} M$ del mismo rango en $\mathrm{Hom}(T_l(E_1), T_l(E_2))$, por lo que $\operatorname{rang} M \leq 4$.

Si $\mathrm{Hom}(E_1, E_2)$ no fuera finitamente generado, podríamos encontrar una sucesión de submódulos finitamente generados

$$M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots$$

cuyo rango está acotado por 4, luego a partir de uno dado —que podemos suponer que es M_0— todos tendrán el mismo rango. Esto implica que los cocientes M_n/M_0 son finitos, luego existe un $m_n \geq 1$ tal que $m_n M_n \subset M_0$. Así pues, $M_n \subset M_0^*$, donde M_0^* es el módulo definido en la demostración del teorema anterior, donde hemos probado que es libre de rango finito. Ahora bien, esto es imposible, pues entonces la unión de los módulos M_n también sería un módulo finitamente generado, algún M_n contendría un generador de la unión y entonces $M_n = M_{n+1}$, contradicción.

Así pues, $\mathrm{Hom}(E_1, E_2)$ es finitamente generado, luego es libre y de rango menor o igual que 4. ∎

Ahora vamos a definir una forma bilineal en $T_l(E)$ que nos aportará información sobre las isogenias duales. Primeramente definiremos formas bilineales sobre los grupos $E[l^n]$ y luego las combinaremos entre sí. Vamos a usar con frecuencia el siguiente hecho elemental:

Teorema 3.24 *Un divisor $\mathfrak{a} = P_1^{m_1} \cdots P_r^{m_r}$ en una curva elíptica E es principal si y sólo si $\sum m_i = 0$ y $\sum m_i P_i = O$.*

DEMOSTRACIÓN: Basta observar que $\operatorname{grad} \mathfrak{a} = \sum m_i$ y, supuesto que $\mathfrak{a}$ tenga grado 0, entonces $\sum m_i P_i$ es el punto de E correspondiente a $[\mathfrak{a}]$ a través del isomorfismo $H_0(E) \cong E$. Ciertamente, $\mathfrak{a}$ es principal si y sólo si su clase es el elemento neutro de $H_0(E)$, si y sólo si su imagen en E es O. ∎

Dada una curva elíptica E/k, fijamos un natural $m > 1$ no divisible entre $\operatorname{car} k$. Consideremos un punto $T \in E[m]$. Por el teorema anterior existe una función $f \in \overline{k}(E)$ tal que $(f) = T^m/O^m$. La función f está determinada por m y T salvo una constante.

Podemos tomar $T' \in E$ tal que $mT' = T$, y entonces existe $g \in \overline{k}(E)$ tal que

$$(g) = \overline{m}(T)/\overline{m}(O) = \prod_{R \in E[m]} \frac{T' + R}{R}.$$

(Aquí usamos que $|E[m]| = m^2$ y $m^2 T' = O$.)

Ahora observamos que $(m \circ f) = \overline{m}((f)) = (g^m)$, luego multiplicando f por una constante podemos suponer que $m \circ f = g^m$. De nuevo, la función g está determinada por m y T salvo una constante.

Sea ahora $S \in E[m]$ un punto arbitrario y $X \in E$. Tenemos que

$$g(X+S)^m = f(mX + mS) = f(mX) = g(X)^m.$$

Esto implica que las funciones $g(X+S)$ y $g(X)$ (para un S fijo) tienen los mismos ceros y polos, luego el cociente $g(X+S)/g(X)$ es constante. Más aún, es una raíz m-sima de la unidad en $\overline{k}$.

Definición 3.25 Sea E una curva elíptica sobre un cuerpo k, sea $m > 1$ un natural no divisible entre $\operatorname{car} k$ y sea U_m el grupo de las raíces m-simas de la unidad en $\overline{k}$. Definimos el *producto de Weil* de orden m en E como la aplicación

$$e_m : E[m] \times E[m] \longrightarrow U_m$$

dada por

$$e_m(S,T) = \frac{g(X+S)}{g(X)},$$

donde $g \in \overline{k}(E)$ es cualquier función que cumpla $(g) = \overline{m}(T)/\overline{m}(O)$ y X es cualquier punto de E donde g no tenga un cero ni un polo.

Observemos que, aunque g esté definida salvo una constante, ésta se cancela en el cociente $e_m(S,T)$, luego $e_m(S,T)$ está unívocamente determinado por m, S y T. El teorema siguiente recoge las propiedades básicas de este producto.

Teorema 3.26 *Sea E una curva elíptica sobre un cuerpo k. Entonces:*

a) $e_m(S_1+S_2,T) = e_m(S_1,T)e_m(S_2,T)$,
$e_m(S,T_1+T_2) = e_m(S,T_1)e_m(S,T_2)$.

b) $e_m(S,T) = e_m(T,S)^{-1}$.

c) Si $e_m(S,T) = 1$ *para todo* $S \in E[m]$, *entonces* $T = O$.

d) Para cada $\sigma \in G(\overline{k}/k)$, *se cumple que* $e_m(S,T)^\sigma = e_m(S^\sigma, T^\sigma)$.

e) Si $S \in E[mm']$, $T \in E_m$, *entonces* $e_{mm'}(S,T) = e_m(m'S,T)$.

Demostración: a) Claramente

$$e_m(S_1+S_2,T) = \frac{g(X+S_1+S_2)}{g(X+S_2)}\frac{g(X+S_2)}{g(X)} = e_m(S_1,T)e_m(S_2,T).$$

Para demostrar la linealidad en la segunda componente tomamos funciones f_1, f_2, f_3, g_1, g_2, g_3 correspondientes a los puntos T_1, T_2 y $T_3 = T_1 + T_2$.

Sea $h \in \overline{k}(E)$ tal que $(h) = T_3O/T_1T_2$. Entonces $(f_3/f_1f_2) = (h^m)$, luego existe $c \in \overline{k}^*$ tal que $f_3 = cf_1f_2h^m$. Componiendo con la multiplicación por m obtenemos que $g_3^m = cg_1^m g_2^m (m \circ h)^m$, luego $g_3 = c' g_1 g_2 (m \circ h)$. Así,

$$\begin{aligned} e_m(S, T_1 + T_2) &= \frac{g_3(X+S)}{g_3(X)} = \frac{g_1(X+S)g_2(X+S)h(mX+mS)}{g_1(X)g_2(X)h(mX)} \\ &= e_m(S,T_1)e_m(S,T_2). \end{aligned}$$

b) Por el apartado anterior tenemos que

$$e_m(S+T, S+T) = e_m(S,S)e_m(T,T)e_m(S,T)e_m(T,S).$$

Basta probar que $e_m(T,T) = 1$ para todo $T \in E[m]$. Recordemos que τ_P representa a la traslación $X \mapsto X + P$. Observemos que

$$\Big(\prod_{i=0}^{m-1} \tau_{iT} \circ f\Big) = \prod_{i=0}^{m-1} \overline{\tau}_{iT}((f)) = \prod_{i=0}^{m-1} \frac{((1+i)T)^m}{(iT)^m} = 1,$$

luego $\prod_{i=0}^{m-1} \tau_{iT} \circ f$ es constante y, si $mT' = T$ la función $\prod_{i=0}^{m-1} \tau_{iT'} \circ g$ también lo es, ya que su potencia m-sima es la multiplicación por m seguida de la función anterior. Evaluándola en X y $X + T'$ hemos de obtener el mismo resultado:

$$\prod_{i=0}^{m-1} g(X + iT') = \prod_{i=0}^{m-1} g(X + (i+1)T').$$

Cancelando términos, la igualdad se reduce a

$$g(X) = g(X + mT') = g(X + mT).$$

Así pues, $e_m(T,T) = 1$.

c) Si $e_m(S,T) = 1$ para todo $S \in E[m]$, entonces $g(X+S) = g(X)$, para todo $S \in E[m]$ y todo $X \in E$ donde g esté definida. Equivalentemente, $\overline{\tau}_S(g) = g$, para todo $S \in E[m]$. Por 2.35 c) tenemos que $g \in \overline{m}[\overline{k}(E)]$, es decir, que $g = m \circ h$, para cierta $h \in \overline{k}(E)$.

Así, $(m \circ h)^m = g^m = m \circ f$, luego $f = h^m$, luego $(h)^m = (f) = T^m/O^m$, luego $(h) = T/O$, lo cual sólo es posible si $T = O$.

d) Si f y g definen e_m para T, es claro que f^σ y g^σ definen e_m para T^σ, luego

$$e_m(S^\sigma, T^\sigma) = \frac{g^\sigma(X^\sigma + S^\sigma)}{g^\sigma(X^\sigma)} = \left(\frac{g(X+S)}{g(X)}\right)^\sigma = e_m(S,T)^\sigma.$$

e) Si f y g definen e_m para T, entonces

$$(f^{m'}) = T^{mm'}/O^{mm'}, \qquad (m' \circ g)^{mm'} = (mm' \circ f)^{m'},$$

luego $f^{m'}$ y $m' \circ g$ definen $e_{mm'}$ para T. Por consiguiente

$$e_{mm'}(S,T) = \frac{(m' \circ g)(X+S)}{(m' \circ g)(X)} = \frac{g(Y + m'S)}{g(Y)} = e_m(m'S, T).$$

■

Una consecuencia sencilla de estas propiedades es el teorema siguiente:

Teorema 3.27 *Si E/k es una curva elíptica y $m > 1$ no es divisible entre* $\operatorname{car} k$*, una condición necesaria para que $E[m] \subset E(k)$ es que k contenga una raíz m-sima primitiva de la unidad.*

DEMOSTRACIÓN: La imagen de e_m en U_m es un subgrupo, digamos de orden $d \mid m$. Entonces

$$1 = e_m(S,T)^d = e_m(dS,T),$$

para todo $T \in E[m]$, luego $dS = O$ para todo $S \in E[m]$, lo cual sólo es posible si $d = m$. En definitiva, la imagen de e_m es todo U_m.

Si $E[m] \subset E(k)$, entonces cada $e_m(S,T)$ es invariante por $G(\overline{k}/k)$, luego $e_m(S,T) \in k$, es decir, $U_m \subset k$. ∎

Veamos ahora la conexión de el producto de Weil con las isogenias duales:

Teorema 3.28 *Si $\phi : E_1 \longrightarrow E_2$ es una isogenia entre dos curvas elípticas, m es un natural no divisible entre la característica del cuerpo de constantes y $S \in E_1[m]$, $T \in E_2[m]$, entonces*

$$e_m(S, \hat{\phi}(T)) = e_m(\phi(S), T).$$

DEMOSTRACIÓN: Podemos suponer que ϕ no es nula. Sean f y g las funciones que definen e_m para T en E_2. Según el teorema 3.10, la isogenia dual está determinada por la relación

$$\hat{\phi}(T)/O = \overline{\phi}(T/O),$$

luego podemos tomar $h \in \overline{k}(E_1)$ tal que

$$(h) = \frac{O\overline{\phi}(T)}{\overline{\phi}(O)\hat{\phi}(T)}.$$

Claramente,

$$\left(\frac{\phi \circ f}{h^m}\right) = \frac{\overline{\phi}((f))}{(h)^m} = \frac{\overline{\phi}(T)^m}{\overline{\phi}(O)^m (h)^m} = \frac{\hat{\phi}(T)^m}{O^m}$$

y

$$\left(\frac{\phi \circ g}{m \circ h}\right)^m = \frac{m \circ \phi \circ f}{(m \circ h)^m} = m \circ \left(\frac{\phi \circ f}{h^m}\right).$$

Esto significa que las funciones $\phi \circ f/h^m$ y $(\phi \circ g)/(m \circ h)$ definen e_m para $\hat{\phi}(T)$. Así pues,

$$\begin{aligned} e_m(S, \hat{\phi}(T)) &= \frac{(\phi \circ g)/(m \circ h)(X+S)}{(\phi \circ g)/(m \circ h)(X)} \\ &= \frac{g(\phi(X)+\phi(S))}{g(\phi(X))} \frac{h(mX)}{h(mX+mS)} \\ &= \frac{g(\phi(X)+\phi(S))}{g(\phi(X))} \frac{h(mX)}{h(mX)} = e_m(\phi(S), T). \end{aligned}$$

∎

De aquí deducimos una propiedad no trivial de las isogenias duales:

Teorema 3.29 *Si ϕ, $\psi : E_1 \longrightarrow E_2$ son dos isogenias entre curvas elípticas, entonces $\widehat{\phi+\psi} = \hat{\phi} + \hat{\psi}$.*

DEMOSTRACIÓN: Sea $\chi = \widehat{\phi+\psi} - \hat{\phi} - \hat{\psi}$. Hemos de probar que $\chi = 0$. Para ello basta probar que χ se anula sobre todos los grupos $E_2[m]$, donde m no es divisible entre la característica del cuerpo de constantes, ya que entonces su núcleo será infinito. Tomamos $S \in E_1[m]$, $T \in E_2[m]$, de modo que

$$\begin{aligned} e_m(S, \chi(T)) &= e_m(S, \widehat{\phi+\psi}(T)) e_m(S, \hat{\phi}(T))^{-1} e_m(S, \hat{\psi}(T))^{-1} \\ &= e_m((\phi+\psi)(S), T) e_m(\phi(S), T)^{-1} e_m(\psi(S), T)^{-1} \\ &= e_m(\phi(S), T) e_m(\psi(S), T) e_m(\phi(S), T)^{-1} e_m(\psi(S), T)^{-1} = 1, \end{aligned}$$

luego $\chi(T) = 0$, para todo $T \in E_2[m]$. ∎

Como aplicación probamos lo siguiente:

Teorema 3.30 *Sea E una curva elíptica y $\phi \in \mathrm{End}(E)$. Entonces*

$$\phi + \hat{\phi} = 1 + \mathrm{grad}\,\phi - \mathrm{grad}(1-\phi) \in \mathbb{Z}.$$

DEMOSTRACIÓN: El teorema se cumple trivialmente si $\phi = 0$. En caso contrario

$$\mathrm{grad}(1-\phi) = (1-\phi)\widehat{(1-\phi)} = (1-\phi)(1-\hat{\phi}) = 1 - \phi - \hat{\phi} + \mathrm{grad}\,\phi.$$

∎

Finalmente, veamos que podemos unir todos los productos e_{l^n} en un único producto sobre $T_l(E)$.

Sea E una curva elíptica sobre un cuerpo k y sea l un primo distinto de $\mathrm{car}\,k$. Sea U_{l^n} el grupo de las raíces l^n-ésimas de la unidad en $\overline{k}$. Observemos que, al igual que hemos obtenido el anillo de enteros l-ádicos $\mathbb{Z}_l$ como límite proyectivo de los anillos $\mathbb{Z}/l^n\mathbb{Z}$, podemos obtenerlo igualmente a partir de los grupos U_{l^n} con las aplicaciones $U_{l^{n+1}} \longrightarrow U_{l^n}$ dadas por $\zeta \mapsto \zeta^l$.

Concretamente, tomamos una raíz l-ésima primitiva $\zeta_1 \in U_l$ y vamos tomando inductivamente raíces $\zeta_n \in U_{l^n}$ de modo que $\zeta_{n+1}^l = \zeta_n$. De este modo tenemos diagramas conmutativos

$$\begin{array}{ccc} \mathbb{Z}/l^{n+1}\mathbb{Z} & \longrightarrow & U_{l^{n+1}} \\ \downarrow & & \downarrow \\ \mathbb{Z}/l^n\mathbb{Z} & \longrightarrow & U_{l^n} \end{array}$$

donde las flechas horizontales son los isomorfismos de grupos dados por $[a] \mapsto \zeta_n^a$. Estos isomorfismos inducen un isomorfismo de grupos $\varprojlim_n \mathbb{Z}/l^n\mathbb{Z} \longrightarrow \varprojlim_n U_{l^n}$.

Así pues, podemos identificar el grupo aditivo de $\mathbb{Z}_l$ con el grupo (multiplicativo) $\varprojlim_n U_{l^n}$.

Ahora fijemos x, $y \in T_l(E)$, de modo que $\pi_n(x)$, $\pi_n(y) \in E[l^n]$, luego podemos calcular $e_{l^n}(\pi_n(x), \pi_n(y)) \in U_{l^n}$. Estos elementos determinan un elemento del límite proyectivo, pues

$$e_{l^{n+1}}(\pi_{n+1}(x), \pi_{n+1}(y))^l = e_{l^{n+1}}(\pi_{n+1}(x), l\pi_{n+1}(y))$$
$$= e_{l^n}(l\pi_{n+1}(x), l\pi_{n+1}(y)) = e_{l^n}(\pi_n(x), \pi_n(y)).$$

Llamamos $e(x, y) \in \mathbb{Z}_l$ a la imagen de este elemento por el isomorfismo entre el límite proyectivo y $\mathbb{Z}_l$. El teorema siguiente es inmediato:

Teorema 3.31 *Si E es una curva elíptica y l es un primo distinto de la característica del cuerpo de constantes, entonces existe una aplicación*

$$e : T_l(E) \times T_l(E) \longrightarrow \mathbb{Z}_l$$

que verifica:

a) $e(S_1 + S_2, T) = e(S_1, T) + e(S_2, T)$,
$e(S, T_1 + T_2) = e(S, T_1) + e(S, T_2)$.

b) $e(S, T) = -e(T, S)$.

c) Si $e(S, T) = 0$ *para todo* $S \in T_l(E)$, *entonces* $T = 0$.

d) Para cada $\sigma \in G(\overline{k}/k)$, *se cumple que* $e(S, T)^\sigma = e(S^\sigma, T^\sigma)$.

e) Si $\phi : E_1 \longrightarrow E_2$ *es una isogenia entre curvas elípticas,* $S \in T_l(E_1)$ *y* $T \in T_l(E_2)$, *entonces* $e(\phi(S), T) = e(S, \hat{\phi}(S))$.

Como aplicación demostramos lo siguiente:

Teorema 3.32 *Sea E una curva elíptica, $\phi \in \mathrm{End}(E)$ y l un primo distinto de la característica del cuerpo de constantes. Entonces*

$$\mathrm{N}(\phi_l) = \det \phi = \phi\hat{\phi}, \qquad \mathrm{Tr}(\phi_l) = 1 + \mathrm{grad}\, \phi - \mathrm{grad}(1 - \phi) = \phi + \hat{\phi}.$$

En particular la norma y la traza de ϕ_l son enteros independientes de l.

Demostración: Tomemos una base S, $T \in T_l(E)$ y sea

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

la matriz de ϕ_l en dicha base. Entonces

$$(\mathrm{grad}\, \phi)e(S, T) = e((\mathrm{grad}\, \phi)S, T) = e(\hat{\phi}_l(\phi_l(S)), T) = e(\phi_l(S), \phi_l(T))$$
$$= e(aS + bT, cS + dT) = (ad - bc)e(S, T) = (\det \phi_l)e(S, T).$$

Por consiguiente $\mathrm{grad}\, \phi = \det \phi_l$. Un simple cálculo muestra que toda matriz 2×2 cumple $\mathrm{Tr}(A) = 1 + \det A - \det(1 - A)$. ∎

3.5 El anillo de endomorfismos

Finalmente estamos en condiciones de determinar la estructura de los anillos de endomorfismos de las curvas elípticas. Necesitamos algunas definiciones:

Definición 3.33 Un *álgebra de cuaternios* es una $\mathbb{Q}$-álgebra A (es decir, un anillo que contiene a $\mathbb{Q}$ como subcuerpo) que tiene una $\mathbb{Q}$-base $\{1, \alpha, \beta, \gamma\}$ tal que

$$\alpha^2,\ \beta^2 \in \mathbb{Q}, \quad \alpha^2 < 0, \quad \beta^2 < 0, \quad \gamma = \alpha\beta = -\beta\alpha.$$

Un *orden* en un álgebra de cuaternios A es un subanillo (unitario) que como $\mathbb{Z}$-módulo es libre de rango 4. Esto hace que una $\mathbb{Z}$-base de $\mathcal{O}$ sea también una $\mathbb{Q}$-base de A.

Notemos que si K es un cuerpo cuadrático también podemos verlo como un álgebra de grado 2 sobre $\mathbb{Q}$ y la definición anterior de orden (cambiando rango 4 por rango 2) es la usual en teoría de números.

Teorema 3.34 *Sea $\mathcal{O}$ un anillo unitario de característica 0 sin divisores de 0 que cumpla las propiedades siguientes:*

a) $\mathcal{O}$ es un $\mathbb{Z}$-módulo libre de rango menor o igual que 4.

b) $\mathcal{O}$ tiene una antiinvolución $\alpha \mapsto \hat{\alpha}$, esto es, una aplicación que cumple las propiedades siguientes:

$$\widehat{\alpha+\beta} = \hat{\alpha} + \hat{\beta}, \quad \widehat{\alpha\beta} = \hat{\beta}\hat{\alpha}, \quad \hat{\hat{\alpha}} = \alpha.$$

Además, si $\alpha \in \mathbb{Z}$ entonces $\hat{\alpha} = \alpha$.

c) Si $\alpha \in \mathcal{O}$ entonces $\alpha\hat{\alpha} \in \mathbb{Z}$, $\alpha\hat{\alpha} \geq 0$ y $\alpha\hat{\alpha} = 0$ si y sólo si $\alpha = 0$.

Entonces $\mathcal{O}$ es isomorfo a $\mathbb{Z}$, a un orden de un cuerpo cuadrático imaginario o a un orden en un álgebra de cuaternios.

Demostración: Sea $A = \mathbb{Q} \otimes \mathcal{O}$ (de modo que A es una $\mathbb{Q}$-álgebra tal que toda $\mathbb{Z}$-base de $\mathcal{O}$ es una $\mathbb{Q}$-base de A). Ciertamente $\mathcal{O}$ es un orden en A. Basta probar que A es $\mathbb{Q}$, un cuerpo cuadrático imaginario o un álgebra de cuaternios.

La antiinvolución de $\mathcal{O}$ se extiende de forma única a una aplicación $\mathbb{Q}$-lineal en A y es claro que conserva todas las propiedades del enunciado (salvo que $\alpha\hat{\alpha}$ es, en general, un número racional y no un número entero).

Definimos una *norma* y una *traza* $\mathrm{N}, \mathrm{Tr} : A \longrightarrow \mathbb{Q}$ mediante

$$\mathrm{N}(\alpha) = \alpha\hat{\alpha}, \qquad \mathrm{Tr}(\alpha) = \alpha + \hat{\alpha}.$$

El hecho de que la traza está en $\mathbb{Q}$ se sigue de que

$$\mathrm{N}(\alpha - 1) = \alpha\hat{\alpha} - \alpha - \hat{\alpha} + 1 = \mathrm{N}(\alpha) - \mathrm{Tr}(\alpha) + 1.$$

Obviamente la traza es $\mathbb{Q}$-lineal y si $\alpha \in \mathbb{Q}$ entonces $\mathrm{Tr}(\alpha) = 2\alpha$. Más aún, si un $\alpha \in A$ cumple $\mathrm{Tr}(\alpha) = 0$ entonces

$$0 = (\alpha - \alpha)(\alpha - \hat{\alpha}) = \alpha^2 - \mathrm{Tr}(\alpha)\alpha + \mathrm{N}(\alpha) = \alpha^2 + \mathrm{N}(\alpha),$$

luego o bien $\alpha = 0$ o bien $\alpha^2 = -\mathrm{N}(\alpha)$ es un número racional negativo.

Si $A = \mathbb{Q}$ no hay nada que probar. En caso contrario existe $\alpha \in A \setminus \mathbb{Q}$. Reemplazándolo por $\alpha - \frac{1}{2}\mathrm{Tr}(\alpha)$ podemos exigir que $\mathrm{Tr}(\alpha) = 0$, con lo que $\alpha^2 < 0$ y $\mathbb{Q}(\alpha)$ es un cuerpo cuadrático imaginario. Si $A = \mathbb{Q}(\alpha)$ ya no hay nada que probar. En caso contrario existe $\beta \in A \setminus \mathbb{Q}(\alpha)$. Podemos reemplazarlo por

$$\beta - \frac{1}{2}\mathrm{Tr}(\beta) - \frac{\mathrm{Tr}(\alpha\beta)}{2\alpha^2}\alpha,$$

lo que garantiza que $\mathrm{Tr}(\beta) = \mathrm{Tr}(\alpha\beta) = 0$. En particular $\beta^2 < 0$. El hecho de que las trazas se anulen equivale a

$$\alpha = -\hat{\alpha}, \quad \beta = -\hat{\beta}, \quad \alpha\beta = -\hat{\beta}\hat{\alpha},$$

y al sustituir las dos primeras igualdades en la tercera obtenemos que $\alpha\beta = -\beta\alpha$.

Si probamos que 1, α, β y $\gamma = \alpha\beta$ son linealmente independientes, tendrá que ser $A = \langle 1, \alpha, \beta, \gamma\rangle$, luego A será un álgebra de cuaternios. Supongamos que

$$a + b\alpha + c\beta + d\gamma = 0, \qquad a, b, c, d \in \mathbb{Q}.$$

Tomando trazas obtenemos que $2a = 0$, luego $a = 0$. Multiplicamos por α por la izquierda y por β por la derecha:

$$b\alpha^2\beta + c\beta^2\alpha + d\alpha^2\beta^2 = 0,$$

pero esto implica que $\beta \in \mathbb{Q}(\alpha)$ salvo si $b = 0$, pero entonces $\alpha \in \mathbb{Q}$ salvo si $c = 0$, lo que obliga también a que $d = 0$. ∎

Como consecuencia inmediata:

Teorema 3.35 *El anillo de endomorfismos de una curva elíptica es isomorfo a $\mathbb{Z}$, a un orden en un cuerpo cuadrático imaginario o a un orden en un álgebra de cuaternios.*

En efecto, si E es una curva elíptica hemos probado que el anillo $\mathrm{End}(E)$ satisface todas las hipótesis del teorema 3.34. Observemos que la norma y la traza de una isogenia en el sentido definido en la prueba de 3.34 coinciden con las del teorema 3.32.

Es conocido que el grupo de unidades de un orden cuadrático tiene orden 2, 4 o 6, por lo que en los casos en los que el teorema 2.12 afirma que el grupo de automorfismos tiene orden 12 o 24 (es decir, las curvas de invariante $j = 0$ en característica 2 y 3), podemos concluir que el anillo de endomorfismos tiene rango 4.

Terminamos la sección con una última propiedad:

Teorema 3.36 *Dos curvas isógenas tienen anillos de endomorfismos del mismo rango.*

DEMOSTRACIÓN: Sea $\lambda : E_1 \longrightarrow E_2$ una isogenia de grado $n > 0$ entre dos curvas elípticas. Entonces podemos definir un monomorfismo de anillos

$$\phi : \mathbb{Q} \otimes \mathrm{End}(E_1) \longrightarrow \mathbb{Q} \otimes \mathrm{End}(E_2)$$

mediante $\phi(\alpha) = \frac{1}{n}(\hat{\lambda}\alpha\lambda)$. Es fácil ver que efectivamente es un homomorfismo (teniendo en cuenta que $\lambda\hat{\lambda} = n$). Si $\phi(\alpha) = 0$ entonces $\hat{\lambda}\alpha\lambda = 0$, luego $\hat{\lambda}\alpha$ toma valores en el núcleo de λ, que es finito, lo cual sólo es posible si α es la isogenia nula.

Esto prueba que el rango de $\mathrm{End}(E_1)$ es menor o igual que el de $\mathrm{End}(E_2)$, pero como la relación de isogenia es simétrica se ha de dar la igualdad. ∎

Capítulo IV

Curvas elípticas sobre cuerpos finitos

Nos ocupamos ahora de las curvas elípticas definidas sobre cuerpos finitos. Cuando tenemos una curva elíptica definida mediante una ecuación con coeficientes enteros (o, más en general, enteros algebraicos) podemos tomar congruencias módulo un primo para pasar a una ecuación definida en un cuerpo finito. Si el primo no divide al discriminante de la curva, el resultado será de nuevo una curva elíptica y podremos aplicarle los resultados que vamos a obtener aquí.

4.1 Puntos racionales

En general, una curva elíptica E/k definida sobre un cuerpo k no algebraicamente cerrado no tiene por qué tener puntos racionales (distintos de O). Sin embargo, vamos a probar que si k es un cuerpo finito con $|k| > 4$ entonces E/k tiene al menos un punto racional no trivial. Más aún, vamos a dar una estimación del número de puntos racionales de E/k.

Lo primero que observamos es que, si k tiene m elementos, entonces los coeficientes de una ecuación de E/k quedan invariantes al elevarlos a m, por lo que la imagen de la aplicación de Frobenius ϕ_m es la propia curva E y por lo tanto ϕ_m resulta ser un endomorfismo. Explícitamente, $\phi_m(X,Y) = (X^m, Y^m)$, luego un punto $P \in E$ está en $E(k)$ si y sólo si $\phi(P) = P$. En otros términos:

$$E(k) = \mathrm{N}(1 - \phi_m).$$

Vamos a probar que $1 - \phi_m$ es separable, con lo que el orden del núcleo será el grado. Más en general:

Teorema 4.1 *Sea E/k una curva elíptica definida sobre un cuerpo k de m elementos. Sea $\phi_m : E \longrightarrow E$ la aplicación de Frobenius de orden m y sean r, $s \in \mathbb{Z}$. Entonces la aplicación $r + s\phi_m$ es separable si y sólo si $p \nmid r$.*

DEMOSTRACIÓN: Vamos a aplicar el teorema 1.19. Para ello tomamos una diferencial invariante ω en E y calculamos (usando 3.2 y 3.3)

$$\overline{r + s\phi}(\omega) = r\omega + s\overline{\phi}(\omega) = r\omega.$$

(Notemos que $\overline{\phi}(\omega) = 0$ por 1.19, ya que ϕ es puramente inseparable.) Así pues, concluimos que $r + s\phi$ es separable si y sólo si $r\omega \neq 0$, si y sólo si $p \nmid r$. ∎

De este modo nos encontramos con que el número de puntos racionales de una curva E/k es

$$|E(k)| = |\,\mathrm{N}(1 - \phi_m)| = \mathrm{grad}(1 - \phi_m).$$

A menudo es conveniente expresar esto en los términos siguientes:

Definición 4.2 Sea E/k una curva elíptica definida sobre un cuerpo k de m elementos. La *traza de Frobenius* de E/k es

$$T(E/k) = m + 1 - \mathrm{grad}(1 - \phi_m) \in \mathbb{Z},$$

donde ϕ_m es la aplicación de Frobenius de orden m en E.

Según el teorema 3.32, tenemos que $T(E/k)$ es la traza del endomorfismo inducido por ϕ_m en cualquier módulo de Tate $T_l(E)$ para cualquier primo l distinto de la característica de k.

Si llamamos $N_k = |E(k)|$, hemos probado la relación

$$N_k = m + 1 - T(E/k).$$

El teorema siguiente nos permitirá estimar $T(E/k)$, y con ello el número de puntos racionales de E/k.

Teorema 4.3 *Bajo las hipótesis del teorema* 3.7, *si* ϕ, $\psi \in \mathrm{End}\, E$, *se cumple*

$$|(\phi, \psi)| \leq \sqrt{(\phi, \phi)(\psi, \psi)}.$$

DEMOSTRACIÓN: Si m, $n \in \mathbb{Z}$, se cumple que

$$0 \leq \mathrm{grad}(m\phi + n\psi) = m^2(\phi, \phi) + 2mn(\phi, \psi) + n^2(\psi, \psi).$$

En particular, tomando $m = (\phi, \psi)$, $n = -(\phi, \phi)$ queda

$$0 \leq -(\phi, \psi)^2(\phi, \phi) + (\phi, \phi)^2(\psi, \psi).$$

Podemos suponer $\phi \neq 0$ y simplificar (ϕ, ϕ), con lo que $(\phi, \psi)^2 \leq (\phi, \phi)(\psi, \psi)$. ∎

Volviendo al caso en de la aplicación de Frobenius ϕ_m de E/k, tenemos que

$$\mathrm{grad}(1 - \phi_m) = \mathrm{grad}\, 1 + \mathrm{grad}\, \phi_m - 2(1, \phi_m) = 1 + m - 2(1, \phi_m),$$

luego

$$T(E/k) = m + 1 - \operatorname{grad}(1 - \phi_m) = 2(1, \phi_m)$$

y por el teorema anterior

$$|T(E/k)| \leq 2\sqrt{\operatorname{grad} 1 \operatorname{grad} \phi} = 2\sqrt{m}.$$

En resumen:

Teorema 4.4 (Hasse) *Sea E una curva elíptica definida sobre el cuerpo k de m elementos. Entonces*[1]

$$\big||E(k)| - m - 1\big| \leq 2\sqrt{m}.$$

En particular, el número de puntos de E en el cuerpo de m elementos es asintóticamente igual a m.

Ahora es inmediato que si $|k| > 4$ entonces toda curva elíptica definida sobre k tiene al menos un punto racional no trivial.

Otra consecuencia notable de la relación $|E(k)| = \operatorname{grad}(1 - \phi_m)$ es que el número de puntos racionales se conserva por isogenias. En efecto:

Teorema 4.5 *Dos curvas elípticas definidas sobre un mismo cuerpo finito k e isógenas sobre k tienen el mismo número de puntos racionales.*

Demostración: Sea $\lambda : E_1 \longrightarrow E_2$ una isogenia no nula entre dos curvas elípticas definidas sobre k y sean ϕ_1 y ϕ_2 las correspondientes aplicaciones de Frobenius de grado m. El hecho de que λ esté definida sobre k hace que conmute el diagrama siguiente:

$$\begin{array}{ccc} E_1 & \xrightarrow{\lambda} & E_2 \\ {\scriptstyle\phi_1}\downarrow & & \downarrow{\scriptstyle\phi_2} \\ E_1 & \xrightarrow[\lambda]{} & E_2 \end{array}$$

Por consiguiente, también conmuta el diagrama

$$\begin{array}{ccc} E_1 & \xrightarrow{\lambda} & E_2 \\ {\scriptstyle 1-\phi_1}\downarrow & & \downarrow{\scriptstyle 1-\phi_2} \\ E_1 & \xrightarrow[\lambda]{} & E_2 \end{array}$$

Por consiguiente, $\operatorname{grad}(1 - \phi_1) \operatorname{grad} \lambda = \operatorname{grad} \lambda \operatorname{grad}(1 - \phi_2)$ y simplificando el grado de λ tenemos el teorema. ∎

[1] Esto es un caso particular de la Hipótesis de Riemann. Ver el Apéndice A.

Si conocemos cuántos puntos racionales tiene una curva elíptica E sobre el cuerpo k de m elementos, podemos calcular cuántos tiene sobre el cuerpo K de m^d elementos. Se trata de encontrar la relación entre las trazas de Frobenius

$$t = T(E/k) = \phi_m + \hat{\phi}_m, \qquad t' = T(E/K) = \phi_m^d + \hat{\phi}_m^d.$$

Ahora bien, ϕ_m y $\hat{\phi}_m$ son las raíces del polinomio $X^2 - tX + m$, luego

$$t' = \left(\frac{t + \sqrt{t^2 - 4m}}{2}\right)^d + \left(\frac{t - \sqrt{t^2 - 4m}}{2}\right)^d.$$

Para terminar vamos a dar una fórmula útil para contar los puntos racionales de una curva elíptica definida sobre un cuerpo k de característica $p \neq 2$. El grupo multiplicativo k^* es cíclico, luego tiene un único subgrupo de índice 2 que define un único carácter no trivial $\chi : k^* \longrightarrow \{\pm 1\}$. Si $x \in K^*$, entonces $\chi(x) = 1$ si y sólo si x es un cuadrado en k^*. Extendemos χ a k tomando $\chi(0) = 0$.

Consideremos ahora una curva elíptica E/k dada por una ecuación de Weierstrass de la forma $Y^2 = F(X)$, donde $F(X) \in k[X]$ tiene raíces simples en $\overline{k}$. Entonces, cada $x \in k$ dará lugar a un punto de E si $F(x) = 0$, a dos puntos si $F(x)$ es un cuadrado no nulo y a ninguno en otro caso. Añadiendo además el punto infinito vemos que

$$|E(k)| = 1 + \sum_{x \in k} (\chi(F(x)) + 1) = m + 1 + \sum_{x \in k} \chi(F(x)),$$

donde m es el número de elementos de k. Equivalentemente,

$$T(E/k) = -\sum_{x \in k} \chi(F(x)). \tag{4.1}$$

4.2 Curvas supersingulares

En esta sección calcularemos el número de curvas elípticas supersingulares (salvo isomorfismo) definidas sobre un cuerpo finito. En primer lugar daremos una caracterización sencilla de las curvas supersingulares:

Teorema 4.6 *Una curva elíptica E/k definida sobre un cuerpo finito k de característica p es supersingular si y sólo si $p \mid T(E/k)$.*

Demostración: Sea m el número de elementos de k y sea ϕ_m la aplicación de Frobenius de grado m. Entonces, como endomorfismos, $\hat{\phi}_m = T(E/k) - \phi_m$.

Según el teorema 4.1 tenemos que $m \mid T(E/k)$ si y sólo si $\hat{\phi}_m$ es inseparable, lo que equivale a que E sea supersingular. ∎

Así pues, podemos determinar si una curva es supersingular sin más que contar sus puntos racionales. Para curvas definidas sobre un cuerpo primo el criterio anterior se puede refinar:

Teorema 4.7 *Sea $p > 3$ un número primo y E/k una curva definida sobre el cuerpo de p elementos. Entonces E es supersingular si y sólo si $T(E/k) = 0$, si y sólo si $|E(k)| = p + 1$.*

DEMOSTRACIÓN: Una implicación es el teorema anterior. Supongamos que E es supersingular. Entonces la multiplicación por p es puramente inseparable, luego $p = \phi^2 h$, donde ϕ es la aplicación de Frobenius de grado p y h tiene grado 1, luego es un isomorfismo. Sea $t = T(E/k)$. Por el teorema anterior $t = pr$, para cierto $r \in \mathbb{Z}$. Supongamos que $r \neq 0$.

Por la unicidad de la isogenia dual ha de ser $\hat{\phi} = \phi h$. Tenemos que

$$t = pr = \phi^2 hr = \phi + \hat{\phi} = \phi + \phi h = \phi(1 + h).$$

Como ϕ es biyectiva podemos simplificar $\phi hr = 1 + h$. Sea ω la diferencial invariante de E. Por el teorema 1.19 tenemos que $\overline{(1+h)}(\omega) = 0$, y por 3.2 esto equivale a $\overline{h}(\omega) = -\omega$. En particular $h \neq 1$.

Podemos suponer que E viene dada por una ecuación canónica

$$E : Y^2 = X^3 + a_4 X + a_6,$$

y entonces el teorema 2.12 nos da que h es de la forma

$$h(X, Y) = (u^2 X, u^3 Y),$$

donde u es una raíz de la unidad, y queremos probar que es $u = -1$. Según la definición 2.13,

$$\omega = \frac{dx}{2y},$$

luego

$$\overline{h}(\omega) = \frac{d(u^2 x)}{2u^3 y} = u^{-1}\omega = -\omega,$$

luego $u^{-1} = -1$ y, por lo tanto, $u = -1$, lo que prueba que h es el automorfismo de orden 2, o sea, $h = -1$. Concluimos que $t = \phi + \hat{\phi} = \phi + \phi(-1) = 0$, en contra de lo supuesto. ∎

Ejercicio: Mostrar que el teorema anterior es falso en característica 2 o 3.

Pasemos ya a la cuestión de determinar si una curva es supersingular.

Teorema 4.8 *Sea k un cuerpo finito de característica $p > 2$.*

a) Si E/k es una curva elíptica dada por una ecuación de Weierstrass de la forma $Y^2 = F(X)$, donde $F(X) \in k[X]$ es un polinomio con raíces simples en $\overline{k}$. Entonces E es supersingular si y sólo si el coeficiente de X^{p-1} en $F(X)^{(p-1)/2}$ es nulo.

b) Sea $r = (p-1)/2$. El polinomio

$$H_p(T) = \sum_{i=0}^{r} \binom{r}{i}^2 T^i$$

tiene raíces distintas en $\overline{k}$ y si $\lambda \in \overline{k}$, $\lambda \neq 0, 1$, entonces la curva elíptica

$$E_\lambda : Y^2 = X(X-1)(X-\lambda)$$

es supersingular si y sólo si $H_p(\lambda) = 0$.

c) El número de invariantes supersingulares en característica p es

$$1 + [p/12] + \epsilon_p,$$

donde $[\ ]$ denota la parte entera y

$$\epsilon_p = \begin{cases} -1 & \text{si } p \equiv \ \ 1 \text{ (mód 12)}, \\ \ \ 1 & \text{si } p \equiv -1 \text{ (mód 12)}, \\ \ \ 0 & \text{en otro caso.} \end{cases}$$

Demostración: a) Sea $\chi : k^* \longrightarrow \{\pm 1\}$ el único carácter no trivial de orden 2 de k^*, extendido con $\chi(0) = 0$. Según hemos visto al final de la sección anterior,

$$|E(k)| = 1 + m + \sum_{x \in k} \chi(F(x)),$$

donde m es el número de elementos de k. Ahora bien, como k^* es un grupo cíclico de orden $m-1$, tenemos que $\chi(x) = x^{(m-1)/2}$ (entendiendo que χ toma sus valores en k). Por lo tanto, viendo a $|E(k)|$ como elemento de k (es decir, identificándolo con su resto módulo p) tenemos que

$$|E(k)| = 1 + m + \sum_{x \in k} F(x)^{(m-1)/2}.$$

Observemos ahora que si x_0 es un generador de k^* tenemos

$$\sum_{x \in k} x^i = \sum_{j=0}^{m-2} x_0^{ij} = \frac{x_0^{(m-1)i} - 1}{x_0^i - 1} = 0$$

salvo si $m-1 \mid i > 0$, en cuyo caso $x_0^i = 1$ y la suma da $m - 1 = -1$. (Notemos que si $i = 0$ entonces no podemos eliminar el sumando correspondiente a $x = 0$ y la suma da $m = 0$.)

Como $F(X)$ tiene grado 3, al multiplicar $F(X)^{(m-1)/2}$ obtenemos un polinomio de grado $3(m-1)/2$. Al aplicar lo anterior a cada monomio cuando x recorre k, el único monomio de grado $i > 0$ múltiplo de $m-1$ es el de grado $m-1$, luego si llamamos T_m al coeficiente de grado $m-1$ en $F(X)^{(m-1)/2}$, entonces (en k) $|E(k)| = 1 - T_m$.

Esto significa que T_m es $T(E/k)$ (como elemento de k). Así pues, E es supersingular si y sólo si $T(E/k) \equiv 0$ (mód p) si y sólo si $T(E/k) = 0$ en k, si y sólo si $T_m = 0$.

Falta probar que $T_m = 0$ si y sólo si $T_p = 0$ (pues el enunciado del teorema hace referencia a T_p). Para ello observamos que

$$F(X)^{(p^{n+1}-1)/2} = F(X)^{(p^n-1)/2}(F(X)^{(p-1)/2})^{p^n}.$$

Al igualar coeficientes obtenemos la relación $T_{p^{n+1}} = T_{p^n}T_p^{p^n}$. En efecto, el segundo factor es un polinomio cuyos monomios tienen grado múltiplo de p^n. El coeficiente de grado $(p-1)p^n$ es $T_p^{p^n}$, que multiplicado por el coeficiente de grado $p^n - 1$ del primer factor (que es T_{p^n}) proporciona un término de grado $p^{n+1} - 1$ en el producto, pero ya no hay más, pues el siguiente término del segundo factor tiene grado p^{n+1}, que excede a $p^{n+1} - 1$, y el término anterior tiene grado $(p-2)p^n$, pero en el primer factor no hay ya términos de grado $2p^n - 1 > 3(p^n - 1)/2$.

A partir de la relación obtenida, un simple argumento inductivo nos da la conclusión.

b) Vamos a aplicar el apartado anterior a $F(X) = X(X-1)(X-\lambda)$. Buscamos el coeficiente de X^{p-1} en $(X(X-1)(X-\lambda))^r$, que es el coeficiente de X^r en $(X-1)^r(X-\lambda)^r$. Dicho coeficiente es

$$\sum_{i=0}^{r} \tbinom{r}{i}(-\lambda)^i \tbinom{r}{r-i}(-1)^{r-i} = (-1)^r H_p(\lambda).$$

Falta probar que $H_p(T)$ tiene raíces simples. Para ello comprobaremos la identidad siguiente:

$$4T(1-T)\frac{d^2H_p}{dT^2} + 4(1-2T)\frac{dH_p}{dT} - H_p(T) = 0. \tag{4.2}$$

En efecto, tenemos que

$$\frac{dH_p}{dT} = \sum_{i=1}^{r} \tbinom{r}{i}^2 iT^{i-1}, \qquad \frac{d^2H_p}{dT^2} = \sum_{i=2}^{r} \tbinom{r}{i}^2 i(i-1)T^{i-2},$$

$$\begin{aligned}
4(1-2T)\frac{dH_p}{dT} &= \sum_{i=1}^{r} \tbinom{r}{i}^2 4iT^{i-1} - \sum_{i=1}^{r} \tbinom{r}{i}^2 8iT^i \\
&= \sum_{i=0}^{r-1} \tbinom{r}{i+1}^2 4(i+1)T^i - \sum_{i=0}^{r} \tbinom{r}{i}^2 8iT^i, \\
4T(1-T)\frac{d^2H_p}{dT^2} &= \sum_{i=2}^{r} \tbinom{r}{i}^2 4i(i-1)T^{i-1} - \sum_{i=2}^{r} \tbinom{r}{i}^2 4i(i-1)T^i \\
&= \sum_{i=0}^{r-1} \tbinom{r}{i+1}^2 4i(i+1)T^i - \sum_{i=0}^{r} \tbinom{r}{i}^2 4i(i-1)T^i.
\end{aligned}$$

Ahora hacemos $\binom{r}{i+1} = \frac{r-i}{i+1}\binom{r}{i}$, y queda

$$4(1-2T)\frac{dH_p}{dT} = \sum_{i=0}^{r} \tbinom{r}{i}^2 \frac{(r-i)^2}{i+1} 4T^i - \sum_{i=0}^{r} \tbinom{r}{i}^2 8iT^i,$$

$$4T(1-T)\frac{d^2H_p}{dT^2} = \sum_{i=0}^{r}\binom{r}{i}^2\frac{(r-i)^2}{i+1}4iT^i - \sum_{i=0}^{r}\binom{r}{i}^2 4i(i-1)T^i.$$

El miembro izquierdo de (4.2) es de la forma $\sum_{i=0}^{r}\binom{r}{i}^2 C_iT^i$, donde

$$\begin{aligned} C_i &= \frac{(r-i)^2}{i+1}4i + \frac{(r-i)^2}{i+1}4 - 4i(i-1) - 8i - 1 \\ &= 4(r-i)^2 - 4i^2 - 4i - 1 = 4r^2 - 8ri - 4i - 1 \\ &= p^2 - 2p - 4pi = 0, \end{aligned}$$

pues k tiene característica p. Esto prueba (4.2). Derivando sucesivamente esta expresión obtenemos identidades de la forma

$$4T(1-T)\frac{d^nH_p}{dT^n} = \text{polinomio en las derivadas de orden } < n,$$

para $n \geq 2$. Si $H_p(T)$ tuviera una raíz $t \neq 0, 1$ de orden $n \geq 2$, entonces la derivada n-sima de H debería ser no nula en t, mientras que las derivadas de orden inferior deberían ser todas nulas, pero esto contradice la relación precedente. Por lo tanto, sólo falta justificar que 0 y 1 no son raíces de $H_p(T)$. Ahora bien,

$$H_p(0) = 1, \qquad H_p(1) = \sum_{i=0}^{r}\binom{r}{i}^2 = \binom{p-1}{r} \not\equiv 0 \ (\text{mód } p).$$

(La suma de cuadrados puede evaluarse así: para escoger r objetos de entre $2r$, podemos partir los $2r$ en dos grupos de r, escoger i del primer grupo — de $\binom{r}{i}$ formas distintas— y $r-i$ del segundo —también de $\binom{r}{i}$ formas distintas—.)

c) Por el teorema 2.16, sabemos que toda curva elíptica admite una ecuación de Weierstrass E_λ en forma de Legendre, cuyo invariante es

$$j(E_\lambda) = 2^8\,\frac{(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}.$$

Ahora bien, cada invariante j se corresponde con 6 valores de λ, excepto $j = 0$, que se corresponde con 2 y $j = 1728$, que se corresponde con 3, esto último salvo si $p = 3$, caso que tratamos aparte. Vemos que

$$H_3(T) = 1 + T,$$

luego el único invariante supersingular es el correspondiente a $\lambda = -1$, que a su vez se corresponde con el invariante $j = 0$.

Si $p \geq 5$ definimos $\delta_p(j) = 1$ si la curva elíptica de invariante j es supersingular en característica p y $\delta_p(j) = 0$ si es ordinaria. Entonces, para $p \geq 5$, el número de invariantes supersingulares es

$$\begin{aligned} &\frac{1}{6}\left(\frac{p-1}{2} - 2\delta_p(0) - 3\delta_p(1728)\right) + \delta_p(0) + \delta_p(1728) \\ &\qquad = \frac{p-1}{12} + \frac{2}{3}\delta_p(0) + \frac{1}{2}\delta_p(1728). \end{aligned}$$

Falta por determinar los valores de $\delta_p(0)$ y $\delta_p(1728)$, es decir, cuándo los invariantes $j = 0$ y $j = 1728$ son supersingulares. Como el resultado tiene interés en sí mismo lo enunciamos como teorema independiente a continuación. Observemos que el criterio depende únicamente del resto de p módulo 12. Llevando los cuatro casos posibles a la fórmula anterior se obtiene el criterio del enunciado. ∎

Teorema 4.9 *El invariante $j = 0$ es supersingular en característica $p > 3$ si y sólo si $p \equiv -1$ (mód 3), mientras que el invariante $j = 1728$ es supersingular si y sólo si $p \equiv -1$ (mód 4).*

DEMOSTRACIÓN: Consideramos la curva $Y = X^3 + 1$, cuyo invariante es $j = 0$ en cualquier característica. Aplicamos el apartado a) del teorema anterior, para lo cual necesitamos el coeficiente de $X^{(p-1)/2}$ en $(X^3+1)^{(p-1)/2}$. Éste será nulo salvo si $p \equiv 1$ (mód 3), en cuyo caso vale $\binom{(p-1)/2}{(p-1)/3}$, que no es nulo módulo p (pues se obtiene como productos y cocientes de elementos no nulos).

Para $j = 1728$ consideramos la curva $Y^2 = X^3 + X$. Buscamos el coeficiente de $X^{(p-1)/2}$ en $(X^2 + 1)^{(p-1)/2}$. Como antes, será nulo salvo si $p \equiv 1$ (mód 4), en cuyo caso vale $\binom{(p-1)/2}{(p-1)/4}$, que no es cero módulo p. ∎

Para terminar demostramos que el carácter ordinario o supersingular de una curva elíptica está muy relacionado con la estructura de su anillo de endomorfismos:

Teorema 4.10 *Sea E/k una curva elíptica definida sobre un cuerpo finito k. Si E es supersingular entonces* $\mathrm{End}(E)$ *tiene rango* 4 *como $\mathbb{Z}$-módulo y si E es ordinaria entonces tiene rango* 2*.*

DEMOSTRACIÓN: Supongamos primeramente que E es ordinaria, digamos que $|k| = p^e$ y sea ϕ_{p^e} la aplicación de Frobenius de grado p^e. Veamos que $\phi_{p^e} \notin \mathbb{Z}$. En caso contrario, comparando grados habría de ser $\phi_{p^e} = \pm p^{e/2}$ (con e par), pero entonces $p^{e/2}$ sería inseparable y E sería supersingular, en contra de lo supuesto.

Con esto tenemos que el rango de $\mathrm{End}(E)$ es 2 o 4. Si probamos que $\mathrm{End}(E)$ es conmutativo, el rango tendrá que ser 2. Observemos que el homomorfismo natural $\mathrm{End}(E) \longrightarrow \mathrm{End}(T_p(E))$ es inyectivo, pues si una isogenia ϕ tiene imagen nula entonces se anula sobre todos los subgrupos $E[p^r]$, y como E es ordinaria cada uno de ellos tiene p^r elementos, luego ϕ tiene núcleo infinito y, por consiguiente, es nula.

Por otra parte, como E es ordinaria tenemos que $T_p(E) \cong \mathbb{Z}_p$, luego hemos probado que $\mathrm{End}(E)$ es isomorfo a un subanillo de $\mathrm{End}(\mathbb{Z}_p) \cong \mathbb{Z}_p$. (Notemos que $\mathrm{End}(\mathbb{Z}_p)$ es el anillo de los homomorfismos de $\mathbb{Z}_p$-módulos de $\mathbb{Z}_p$ en sí mismo, y que $\mathbb{Z}_p$ es un $\mathbb{Z}_p$-módulo libre de base 1.)

Consideremos ahora el caso en que E es supersingular y sea $A = \mathbb{Q}\otimes\mathrm{End}(E)$. Supongamos, por reducción al absurdo, que el álgebra A tiene rango menor que 4, con lo que es isomorfa a $\mathbb{Q}$ o a un cuerpo cuadrático imaginario. Por el

teorema 3.18 toda curva isógena a E es supersingular, por 3.17 sólo hay (salvo isomorfismo) una cantidad finita de tales curvas y en la prueba de 3.36 vemos que el anillo de endomorfismos de cualquiera de ellas es isomorfo a un orden de A.

Si $A = \mathbb{Q}$ tomamos un primo cualquiera $l \neq p$, mientras que si A es un cuerpo cuadrático, exigimos además que l se conserve primo en todos los órdenes de A correspondientes a curvas isógenas a E. Precisemos esto:

Si el orden maximal de A es $\mathbb{Z}[\alpha]$, entonces los demás órdenes son de la forma $\mathbb{Z}[f\alpha]$, donde al entero f se le llama conductor del orden correspondiente. Hay infinitos primos racionales l que se conservan primos en el orden maximal. De entre ellos, tomamos uno que no divida a ninguno de los conductores de los órdenes indicados.

Puesto que $E[l^i] \cong \mathbb{Z}/l^i\mathbb{Z} \times \mathbb{Z}/l^i\mathbb{Z}$, podemos tomar una sucesión de subgrupos

$$H_1 \leq H_2 \leq \cdots \leq E, \qquad H_i \cong \mathbb{Z}/l^i\mathbb{Z}.$$

Por el teorema 2.37 existen curvas elípticas E_i junto con isogenias separables $\lambda_i : E \longrightarrow E_i$ de núcleo H_i. Según hemos comentado, hay un número finito de posibles curvas E_i, luego dos de ellas han de ser isomorfas, $E_i \cong E_{i+j}$. Por otra parte, el teorema 2.36 nos da una isogenia $E_{i+j} \longrightarrow E_i$ tal que $\lambda_{i+j} \circ \lambda = \lambda_i$, luego el núcleo de λ es isomorfo a $H_{i+j}/H_i \cong \mathbb{Z}/l^j\mathbb{Z}$. Al componer λ con el isomorfismo obtenemos una isogenia separable $\lambda : E_i \longrightarrow E_i$ cuyo núcleo es cíclico de orden l^j. Por consiguiente, λ tiene grado l^j y $\lambda\hat{\lambda} = l^j$.

Si $\mathrm{End}(E_i) \cong \mathbb{Z}$ la factorización única implica que j es par y $\lambda = \epsilon l^{j/2}$, donde $\epsilon = \pm 1$. Si $\mathrm{End}(E_i)$ es un orden cuadrático podemos llegar a la misma conclusión (ahora con $\epsilon \in \mathrm{Aut}(E_i)$) gracias a la elección de l. En efecto, tenemos que l es primo con el conductor del orden, y todo ideal de un orden cuadrático de norma prima con el conductor se descompone de forma única como producto de ideales primos. Como (l) es primo, la factorización de λ ha de ser $\lambda = \epsilon l^r$ y, aplicando la involución, de hecho $\lambda = \epsilon l^{j/2}$.

Esto nos lleva a que el núcleo de $l^{j/2}$ es cíclico, pero por otra parte sabemos que es producto de dos grupos cíclicos, lo que nos da una contradicción. ∎

4.3 El número de curvas sobre un cuerpo

Según el teorema 2.46, si E/k es una curva elíptica de invariante $j \neq 0, 1728$, hay tantas clases de k-isomorfía de curvas conjugadas con E/k como elementos en k^*/k^{*2}. En general esto nos da infinitas clases, pero si k es un cuerpo finito k^{*2} tiene índice 2, luego a cada invariante j le corresponden sólo 2 clases de conjugación de curvas con invariante j. En esta sección calcularemos el número de clases de k-isomorfía de curvas conjugadas a una dada a partir de una fórmula sobre grupos finitos, que nos permitirá considerar incluso las características 2 y 3, exceptuadas en 2.46.

Recordemos que un grupo G *actúa* sobre un conjunto X si cada elemento de G tiene asociada una permutación de X de forma consistente con la operación

en G, es decir, si tenemos un homomorfismo $\rho : G \longrightarrow \Sigma_X$, donde Σ_X denota el grupo de las permutaciones de X. En lugar de $\rho(g)(x)$ escribimos simplemente xg. Si G actúa sobre X, a cada $g \in G$ le podemos asociar el conjunto fijado

$$F(g) = \{x \in X \mid xg = x\},$$

y a cada $x \in X$ le podemos asociar su estabilizador:

$$G_x = \{g \in G \mid xg = x\}.$$

Por otra parte, en X podemos establecer la relación de equivalencia en virtud de la cual dos elementos x, y están relacionados si existe un $g \in G$ tal que $xg = y$. Las clases de equivalencia se llaman *órbitas* y el conjunto cociente se representa por X/G. El resultado que necesitamos es el siguiente:

Teorema 4.11 (Fórmula de Burnside) *Si un grupo finito G actúa sobre un conjunto finito X, entonces*

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |F(g)|.$$

DEMOSTRACIÓN: Veamos primeramente que la fórmula se cumple cuando $|X/G| = 1$. Se dice entonces que la acción es *transitiva.* Consideramos el conjunto $A = \{(x, g) \in X \times G \mid xg = x\}$.

Agrupando los pares de A según su primera o su segunda componente obtenemos que

$$\sum_{g \in G} |F(g)| = \sum_{x \in X} |G_x|.$$

Ahora bien, si $x \in X$, la aplicación $\phi : G \longrightarrow X$ dada por $\phi(g) = xg$ es suprayectiva (porque sólo hay una órbita) y $\phi(g) = \phi(h)$ si y sólo si $xg = xh$, si y sólo si $xhg^{-1} = x$, si y sólo si $hg^{-1} \in G_x$. Por lo tanto, cada $x \in X$ tiene exactamente $|G_x|$ antiimágenes por ϕ y concluimos que $|G| = |X||G_x|$. Esto nos da que

$$\sum_{g \in G} |F(g)| = \sum_{x \in X} |G|/|X| = |G|,$$

como había que probar.

En el caso general, si $X/G = \{X_1, \ldots, X_n\}$, entonces G actúa sobre cada órbita X_i y por la parte ya probada

$$|G| = \sum_{g \in G} |F_i(g)|,$$

donde $F_i(g) = \{x \in X_i \mid xg = x\}$. Al sumar sobre i obtenemos

$$n|G| = \sum_{g \in G} |F(g)|,$$

lo que prueba el teorema. ∎

Como aplicación de esta fórmula vamos a demostrar lo siguiente:

Teorema 4.12 *Sea $\nu = \nu(m, j)$ el número de curvas elípticas salvo k-isomorfismo definidas sobre el cuerpo k de m elementos con invariante j. Entonces $\nu = 2$ excepto en los casos indicados en la tabla siguiente:*

m	j	ν
$1, 5$ (mód 12)	1728	4
1 (mód 6)	0	6
3^{2r+1}	0	4
3^{2r}	0	6
2^{2r+1}	0	3
2^{2r}	0	7

DEMOSTRACIÓN: Sea $m = p^n$. Supongamos en primer lugar que $p > 3$ y que $j \neq 0, 1728$. Entonces toda curva elíptica definida sobre k admite una ecuación canónica según el teorema 2.7, es decir, de la forma $Y^2 = X^3 + a_4X + a_6$ con

$$j = 1728\,\frac{4a_4^3}{4a_4^3 + 27a_6^2}.$$

Observemos que $a_4 \neq 0 \neq a_6$, o de lo contrario sería $j = 0, 1728$. El número total de ecuaciones correspondientes a un j dado es el número de pares $(a_4, a_6) \in k^* \times k^*$ que cumplen la ecuación

$$(1728 - j)(a_4/3)^3 = (a_6/2)^2.$$

Hay $(m-1)/2$ valores de a_4 para los que el miembro izquierdo tiene raíz cuadrada en k. (Serán los valores para a_4 con o sin raíz cuadrada, según si 3 y $1728 - j$ tienen o no raíz cuadrada en k.) Para cada valor de a_4 hay dos valores para a_6. En total hay $m - 1$ ecuaciones con invariante j.

Dos ecuaciones corresponden a curvas k-isomorfas si y sólo si existe una transformación del tipo dado por el teorema 2.3 que transforma una en la otra. Del teorema 2.6 (comparar con la prueba del teorema 2.12) se sigue que las transformaciones que convierten una ecuación del tipo considerado en otra del mismo tipo son las de la forma

$$X = u^2X', \qquad Y = u^3Y', \qquad u \in k^*.$$

Estas transformaciones constituyen un grupo G de orden $m-1$ y una transformación dada fija a una ecuación si y sólo si $u^2 = 1$, en cuyo caso las fija a todas. Por lo tanto, hay dos transformaciones que fijan a las $m - 1$ ecuaciones y $m - 3$ que no fijan a ninguna. La fórmula de Burnside nos da que

$$\nu = \frac{1}{m-1}(m - 1 + m - 1) = 2.$$

Los demás casos se tratan de forma similar. Como ilustración detallaremos el más complicado, el correspondiente a $m = 2^n$, n par, $j = 0$. Según el teorema 2.7, las curvas en estas condiciones admiten ecuaciones de la forma

$$Y^2 + a_3Y = X^3 + a_4X + a_6, \qquad a_3 \neq 0.$$

Hay un total de $(m-1)m^2 = (2^n-1)2^{2n}$ ecuaciones. Del teorema 2.6 se sigue que las transformaciones que conservan la forma canónica de la ecuación son las del tipo

$$X = u^2X' + s^2, \qquad Y = u^3Y' + su^2X' + t, \qquad u \neq 0.$$

Éstas forman un grupo G de orden $(m-1)m^2 = (2^n-1)2^{2n}$. Además, las que fijan a una ecuación dada son las que cumplen

$$u^3 = 1, \quad (u-1)a_4 + sa_3 + s^4 = 0, \quad s^2a_4 + s^6 + ta_3 + t^2 = 0.$$

Hemos de calcular $|F(g)|$ para cada $g = (u, s, t)$. Distinguimos casos:

a) $u = 1$, $s = 0$.

- $t = 0$. Entonces g es la identidad y $|F(g)| = (2^n-1)2^{2n}$.
- $t \neq 0$. Entonces g fija a las ecuaciones que cumplen $ta_3+t^2 = 0$, lo que determina completamente a a_3, mientras que a_4 y a_6 son arbitrarios. Por consiguiente, $|F(g)| = m^2 = 2^{2n}$. Hay $m-1$ transformaciones g en este caso, luego la contribución total a la fórmula de Burnside es $(2^n-1)2^{2n}$.

b) $u = 1$, $s \neq 0$. Las ecuaciones son $sa_3 + s^4 = 0$, $s^2a_4 + s^6 + ta_3 + t^2 = 0$. La primera determina a a_3 y la segunda a a_4, mientras que a_6 es arbitrario. Así pues, $|F(g)| = m = 2^n$. Hay $(m-1)m = (2^n-1)2^n$ transformaciones de este tipo, luego la contribución total a la fórmula es $(2^n-1)2^{2n}$.

c) $u \neq 1$. Aquí usamos que n es par, pues de lo contrario este caso no podría darse. En efecto, el menor cuerpo de característica 2 que contiene a las raíces cúbicas de la unidad es el de 4 elementos, luego éstas están contenidas en los cuerpos de 2^n elementos con n par.

- $s = 0$, $t = 0$. Las condiciones se reducen a $a_4 = 0$, luego concluimos que $|F(g)| = (m-1)m = (2^n-1)2^n$. Hay 2 transformaciones en este caso (correspondientes a los dos valores de u), luego en total la contribución es de $(2^n-1)2^{n+1}$.
- $s = 0$, $t \neq 0$. A la condición $a_4 = 0$ hay que añadir $a_3 = -t$, lo que reduce las ecuaciones a $|F(g)| = m = 2^n$. El número de transformaciones es $2(2^n-1)$, luego la contribución es de $(2^n-1)2^{n+1}$.
- $s \neq 0$. Las condiciones son

  $$(u-1)a_4 + sa_3 + s^4 = 0, \quad s^2a_4 + s^6 + ta_3 + t^2 = 0.$$

  Despejamos a_4 en la primera (notemos que $1/(u-1) = u$) y sustituimos en la segunda:

  $$a_3(us^3 + t) + (u+1)s^6 + t^2 = 0,$$

o también

$$a_3(us^3+t)+(us^3+t)^2=0.$$

— Si $t=us^3$ se cumplen las condiciones para cualquier a_3 y cualquier a_6, mientras que a_4 está determinado. Esto nos da

$$|F(g)|=(m-1)m=(2^n-1)2^n,$$

y el número de transformaciones es $2(m-1)$. La contribución es de $(2^n-1)^2 2^{n+1}$.

— Si $t\neq us^2$ entonces a_3 está fijo, al igual que a_4, luego $|F(g)|=2^n$ y el número de transformaciones es $2(m-1)^2$. La contribución es de $(2^n-1)^2 2^{n+1}$.

La fórmula de Burnside nos da, finalmente,

$$\begin{aligned}\nu &= \frac{1}{(2^n-1)2^{2n}}(3(2^n-1)2^{2n}+2(2^n-1)2^{n+1}+2(2^n-1)^2 2^{n+1})\\ &= \frac{1}{2^{2n}}(3\cdot 2^{2n}+2^{n+2}+4\cdot 2^{2n}-2^{n+2})=7.\end{aligned}$$

■

A partir de aquí es fácil calcular el número total de curvas elípticas sobre un cuerpo dado:

Teorema 4.13 *El número N de curvas elípticas salvo k-isomorfismo sobre el cuerpo k de m elementos viene dado por la tabla siguiente:*

m	N
2^{2r+1}	$2^{2r+2}+1$
2^{2r}	$2^{2r+1}+5$
3^{2r+1}	$2(3^{2r+1}+1)$
3^{2r}	$2(3^{2r\ \ 1}+3)$
1 (mód 12)	$2m+6$
5 (mód 12)	$2m+2$
7 (mód 12)	$2m+4$
11 (mód 12)	$2m$

Demostración: Es una comprobación sencilla. Veamos por ejemplo el primer caso: hay 2^{2r+1} invariantes posibles, cada uno de los cuales da lugar a 2 curvas, excepto $j=0$, que aporta 3. En total:

$$N=2(2^{2r+1}-1)+3=2^{2r+2}+1.$$

■

Dada una curva

$$E: Y^2=X^3+a_4X+a_6$$

definida sobre un cuerpo finito de característica > 3, el teorema 4.12 afirma que —salvo casos excepcionales— sólo hay dos clases de conjugación de curvas isomorfas a E. El teorema 2.46 nos da que un representante de la otra clase es la curva dada por

$$E' = Y^2 = X^3 + d^2a_4X + d^3a_6,$$

donde d no es un cuadrado en k^*.

Vamos a determinar la relación entre las trazas de E y E'. Para ello consideramos el carácter cuadrático $\chi : k^* \longrightarrow \{\pm 1\}$, aplicamos la fórmula (4.1) y usamos que cuando x recorre k lo mismo hace dx:

$$\begin{aligned} T(E'/k) &= -\sum_{x\in k} \chi(x^3 + d^2a_4x + d^3a_6) = -\sum_{x\in k} \chi(d^3x^3 + d^3a_4x + d^3a_6) \\ &= \chi(d)^3T(E/k) = -T(E/k). \end{aligned}$$

(Notemos que $\chi(d) = -1$ porque hemos tomado d sin raíz cuadrada en k.)

Vemos así que E y E' tienen traza distinta supuesto que ésta sea no nula (en particular si son curvas ordinarias). Con esto tenemos casi demostrado el teorema siguiente:

Teorema 4.14 *Dos curvas elípticas ordinarias sobre un cuerpo finito k son k-isomorfas si y sólo si tienen el mismo invariante y el mismo número de puntos racionales.*

Demostración: Las condiciones son obviamente necesarias. Si se cumplen y $j \neq 0, 1728$, entonces sólo hay dos clases de curvas k-isomorfas con invariante j. Si las dos curvas no fueran k-isomorfas, una de ellas sería k-isomorfa a una curva E y otra a la correspondiente curva E' en las condiciones de la discusión precedente, pero ya hemos visto que en tal caso tendrían trazas opuestas (no nulas, por ser ordinarias), luego no tendrían el mismo número de puntos racionales.

Falta considerar los casos $j = 0$ y $j = 1728$. Estos invariantes son supersingulares en las características 2 y 3, luego podemos suponer que $\operatorname{car} k > 3$. Si $j = 1728$ ambas curvas admiten ecuaciones de la forma

$$E : Y^2 = X^3 + a_4X, \qquad E' : Y^2 = X^3 + a_4'X.$$

Por 4.9 sabemos además que $p \equiv 1$ (mód 4), luego también $m \equiv 1$ (mód 4). Como de costumbre, llamamos m al cardinal de k. Trabajando en k, se cumple

$$T(E/k) = -\sum_{x\in k} \chi(x^3 + a_4x) = -\sum_{x\in k^*} (x^3 + a_4x)^{(m-1)/2}$$

$$= -\sum_{i=0}^{(m-1)/2} \sum_{x\in k^*} \binom{(m-1)/2}{i} x^{2i+(m-1)/2} a_4^{(m-1)/2-i}.$$

Ahora usamos la relación

$$\sum_{x\in k^*} x^i = \begin{cases} -1 & \text{si } m-1 \mid i, \\ 0 & \text{en caso contrario,} \end{cases}$$

que ya nos ha aparecido en la prueba del teorema 4.8. El único exponente múltiplo de $m-1$ en nuestra suma corresponde a $i=(m-1)/4$, luego

$$T(E/k) = \binom{(m-1)/2}{(m-1)/4} a_4^{(m-1)/4}.$$

Como las curvas son ordinarias, las trazas son no nulas módulo p, luego el número combinatorio es no nulo en k. La misma relación vale para $T(E'/k)$, y simplificando los números combinatorios concluimos que $a_4^{(m-1)/4} = a_4'^{(m-1)/4}$ o, lo que es lo mismo, $(a_4/a_4')^{(m-1)/4} = 1$. Esto implica que $\sqrt[4]{a_4/a_4'} \in k$. (En general, si $k^* = \langle g\rangle$ y $(g^i)^{(m-1)/4} = 1$, entonces $m-1 \mid i(m-1)/4$, luego $4 \mid i$, luego g^i tiene raíz cuarta $g^{i/4}$.) Por el teorema 2.47, las dos curvas son k-isomorfas.

Si $j=0$ hemos de considerar ecuaciones de la forma $Y^2 = X^3 + a_6$. El razonamiento es análogo. Ahora tenemos que $m \equiv 1 \pmod 3$, y la expresión para la traza resulta ser

$$T(E/k) = \binom{(m-1)/2}{(m-1)/3} a_6^{(m-1)/6},$$

de donde concluimos que $(a_6/a_6')^{(m-1)/6} = 1$ y de aquí que $\sqrt[6]{a_6/a_6'} \in k$. ∎

Ejercicio: Razonar que el teorema anterior es falso para curvas supersingulares.

Capítulo V

Grupos formales

En este capítulo introducimos una nueva técnica para estudiar las curvas elípticas. Esencialmente consiste en estudiar el desarrollo en serie de Taylor de la suma en un entorno del punto (O, O). Esto nos llevará a la noción de grupo formal, que es una serie de potencias en dos variables que satisface las condiciones necesarias obvias para poder ser una serie de Taylor de una operación de grupo. Antes de introducir la noción general de grupo formal dedicaremos la primera sección a ver con detalle cómo aparecen estos grupos en el estudio de las curvas elípticas.

5.1 Desarrollos de Taylor en O

Consideremos una curva elíptica E definida por una ecuación de Weierstrass

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6.$$

Puesto que queremos estudiar desarrollos de Taylor alrededor del punto O, conviene hacer un cambio de sistema de referencia para que sus coordenadas pasen a ser $(0, 0)$. Dado que $O = [0, 1, 0]$, basta con homogeneizar respecto de Z y deshomogeneizar respecto de Y. La ecuación pasa a ser

$$Z + a_1XZ + a_3Z^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

Para eliminar signos conviene hacer el cambio $Z' = -Z$, $X' = -X$, con lo que la ecuación se convierte en

$$Z = X^3 + a_1XZ + a_2X^2Z + a_3Z^2 + a_4XZ^2 + a_6Z^3.$$

Si las coordenadas afines originales eran $x = X/Z$, $y = Y/Z$, las nuevas son $t = -X/Y$, $z = -Z/Y$, luego

$$t = -x/y, \qquad z = -1/y,$$

y están ligadas por la ecuación

$$z = t^3 + (a_1t + a_2t^2)z + (a_3 + a_4t)z^2 + a_6z^3.$$

Sabemos que x tiene un polo doble en O, mientras que y tiene un polo triple. Por consiguiente, t y z tienen ceros en O de órdenes 1 y 3, respectivamente. En particular t es un parámetro local de E en O y podemos considerar el desarrollo de z en potencias de t, que será de la forma

$$z = t^3(A_0 + A_1t + A_2t^2 + \cdots), \qquad A_i \in \overline{k}.$$

Vamos a demostrar que en realidad $A_n \in \mathbb{Z}[a_1, \ldots, a_6]$. Esto es una consecuencia de un resultado general:

Teorema 5.1 *Sea A un dominio íntegro y $A[[T]]$ su anillo de series formales de potencias. Sean $F \in A[[T]][Z]$ y $n \geq 1$ tales que*

$$F(0) \equiv 0 \pmod{T^n}, \qquad F'(0) \equiv 1 \pmod{T}.$$

Entonces la sucesión

$$z_0 = 0, \qquad z_{m+1} = z_m - F(z_m)$$

converge a una serie $z \in A[[T]]$, que es la única que cumple

$$F(z) = 0, \qquad z \equiv 0 \pmod{T^n}.$$

Demostración: Llamemos $\mathfrak{m} = (T)$ al ideal maximal de $A[[T]]$. Estamos suponiendo que el término independiente de F está en $\mathfrak{m}^n$, luego si $z_m \in \mathfrak{m}^n$, también $F(z_m) \in \mathfrak{m}^n$, luego $z_{m+1} \in \mathfrak{m}^n$. Así pues, concluimos que $z_m \in \mathfrak{m}^n$ para todo m.

Veamos ahora que $z_m \equiv z_{m+1} \pmod{\mathfrak{m}^{m+n}}$. Para $m = 0$ esto significa que $F(0) \equiv 0 \pmod{\mathfrak{m}^n}$, lo cual es cierto por hipótesis. Supongamos que la congruencia se cumple para naturales menores que m.

Consideremos dos indeterminadas Z_1 e Z_2. Claramente

$$F(Z_1) - F(Z_2) = (Z_1 - Z_2)(F'(0) + Z_1G(Z_1, Z_2) + Z_2H(Z_1, Z_2)),$$

con G, $H \in A[[T]][Z_1, Z_2]$. Por lo tanto,

$$z_{m+1} - z_m = z_m - F(z_m) - (z_{m-1} - F(z_{m-1})) = z_m - z_{m-1} - (F(z_m) - F(z_{m-1}))$$
$$= (z_m - z_{m-1})(1 - F'(0) - z_mG(z_m, z_{m-1}) - z_{m-1}H(z_m, z_{m-1})).$$

Por hipótesis de inducción $z_m - z_{m-1} \in \mathfrak{m}^{m+n-1}$, por la hipótesis del teorema $1 - F'(0) \in \mathfrak{m}$ y, según hemos visto antes, z_{m-1}, $z_m \in \mathfrak{m}^n$, luego podemos concluir que $z_{m+1} - z_m \in \mathfrak{m}^{m+n}$.

Con esto hemos probado que la sucesión $z_{m+1} - z_m$ tiende a 0, luego la sucesión z_m es de Cauchy en $A[[T]]$, luego converge a una serie z. Tomando

límites en la relación recurrente que define a z_m obtenemos que $F(z) = 0$. Como cada $z_m \in \mathfrak{m}^n$, la continuidad de la valoración implica que $z \in \mathfrak{m}^n$.

Por último, si $z' \in A[[T]]$ es otra serie que cumple $F(z') = 0$ y $z' \in \mathfrak{m}^n$, entonces

$$0 = F(z) - F(z') = (z - z')(F'(0) + zG(z, z') + z'H(z, z')).$$

Si fuera $z \neq z'$ entonces $F'(0) + zG(z, z') + z'H(z, z') = 0$, de donde podríamos concluir que $F'(0) \in \mathfrak{m}$, contradicción. ∎

Volviendo a la discusión previa al teorema, tomamos $A = \mathbb{Z}[a_1, \ldots, a_6]$ y

$$F(Z) = Z - T^3 - (a_1T + a_2T^2)Z - (a_3 + a_4T)Z^2 - a_6Z^3.$$

Ciertamente

$$F(0) = T^3 \equiv 0 \text{ (mód } T^3), \qquad F'(0) = 1 - (a_1T + a_2T^2) \equiv 1 \text{ (mód } T),$$

luego la serie de potencias $S(T) \in \mathbb{Z}[a_1, \ldots, a_6][[T]]$ que nos da el teorema cumple la ecuación

$$S(T) = T^3 + (a_1T + a_2T^2)S(T) + (a_3 + a_4T)S(T)^2 + a_6S(T)^3,$$

así como que $T^3 \mid S(T)$. Ahora bien, si aplicamos el teorema con $A = \overline{k}$ obtenemos la misma serie $S(T)$, pero ahora la serie de Taylor de z cumple las mismas propiedades, luego por la unicidad concluimos que $S(T)$ es dicha serie de Taylor.

Más aún, podemos aplicar el teorema con $A = A[a_1, \ldots, a_6]$, considerando ahora los a_i's como indeterminadas. Así obtenemos una única serie

$$S(T) \in \mathbb{Z}[a_1, \ldots, a_6][[T]]$$

con la propiedad de que cuando sustituimos las indeterminadas a_i por valores concretos en un cuerpo $\overline{k}$, la serie $S(T)$ se particulariza a la serie de Taylor de z respecto de t correspondiente a la curva elíptica determinada por la ecuación de Weierstrass de coeficientes a_i. Pongamos que

$$S(T) = T^3(A_0 + A_1T + A_2T^2 + \cdots),$$

con $A_i \in \mathbb{Z}[a_1, \ldots, a_6]$. Es fácil calcular explícitamente los polinomios A_m. Por ejemplo:

$$A_0 = 1, \quad A_1 = a_1, \quad A_2 = a_1^2 + a_2, \quad A_3 = a_1^3 + 2a_1a_2 + a_3,$$
$$A_4 = a_1^4 + 3a_1^2a_2 + 3a_1a_3 + a_2^2 + a_4, \ \ldots$$

Vamos a demostrar que, tal y como se observa, la suma de los índices de cada monomio de A_m es igual a m. Para ello definimos el *grado* de un monomio de $\mathbb{Z}[a_1, \ldots, a_6]$ como la suma de sus índices y llamamos G_r al $\mathbb{Z}$-módulo generado por los monomios de grado r. De este modo,

$$\mathbb{Z}[a_1, \ldots, a_6] = \bigoplus_{r=0}^{\infty} G_r$$

y $G_rG_s \subset G_{r+s}$. Vamos a demostrar que $A_m \in G_m$.

Para ello observemos cómo se construye concretamente la serie $S(T)$. Si llamamos

$$P(T,Z) = T^3 + (a_1T + a_2T^2)Z + (a_3 + a_4T)Z^2 + a_6Z^3,$$

la sucesión z_n construida en 5.1 es

$$z_0 = 0, \qquad z_{m+1} = P(T, z_m).$$

Más detalladamente:

$$z_1 = P(T,0), \quad z_2 = P(T,P(T,0)), \quad z_3 = P(T,P(T,P(T,0))),\dots$$

Si definimos

$$P_1(T,Z) = P(T,Z), \qquad P_{m+1}(T,Z) = P(Z,P_m(T,Z)),$$

entonces $z_m = P_m(T,0)$.

Extendamos la aplicación grado a $\mathbb{Z}[a_1,\dots,a_6,T,Z]$ estableciendo que T tiene grado -1 y Z tiene grado -3. Ahora

$$\mathbb{Z}[a_1,\dots,a_6,T,Z] = \bigoplus_{r=-\infty}^{+\infty} G_r.$$

Es claro que $P(T,Z) \in G_{-3}$. Más aún, una simple inducción prueba que $P_m(T,Z) \in G_{-3}$ para todo $m \geq 1$. Por consiguiente,

$$P_m(T,0) = T^3(1 + B_1^mT + B_2^mT^2 + \cdots + B_N^mT^N),$$

con $B_r^m \in G_r$. Como $S(T) = \lim_m P_m(T,0)$, tenemos que $A_r = B_r^m$ para m suficientemente grande, luego A_r tiene grado r.

El teorema siguiente resume lo que hemos obtenido hasta aquí:

Teorema 5.2 *Existe una serie formal de potencias*

$$S(T) = T^3(1 + A_1T + A_2T^2 + \cdots\ ) \in \mathbb{Z}[a_1,\dots,a_6][[T]]$$

tal que si E es una curva elíptica determinada por una ecuación de Weierstrass con coeficientes $a_1,\dots,a_6$, entonces la serie de Taylor de $z = -1/y$ en O alrededor del parámetro $t = -x/y$ es $S(T)$. Además, los índices de cada monomio de $A_m \in \mathbb{Z}[a_1,\dots,a_6]$ suman m.

Ahora podemos volver a las coordenadas de Weierstrass: observemos que la serie $1 + A_1T + A_2T^2 + \cdots$ es una unidad de $\mathbb{Z}[a_1,\dots,a_6][[T]]$, luego las series de Laurent de $x = t/z$ e $y = -1/z$ alrededor de O tienen sus coeficientes en el anillo $\mathbb{Z}[a_1,\dots,a_6]$. Un cálculo rutinario muestra que

$$\begin{aligned} x &= \frac{1}{t^2} - \frac{a_1}{t} - a_2 - a_3t - (a_4 + a_1a_3)t^2 + \cdots \\ y &= -\frac{1}{t^3} + \frac{a_1}{t^2} + \frac{a_2}{t} + a_3t + (a_4 + a_1a_3)t^2 \ + \cdots \end{aligned}$$

Consideramos ahora el producto $E \times E$. Sean p_i las proyecciones, $t_i = p_i \circ t$, $z_i = p_i \circ z$, para $i = 1, 2$. Las funciones t_1, t_2 forman un sistema local de parámetros en (O, O). Sea $S : E \times E \longrightarrow E$ la suma en E y sea $s = S \circ t$. Entonces $s \in k(E \times E)$. Vamos a estudiar su desarrollo en serie de Taylor alrededor del punto (O, O).

Consideremos el plano afín determinado por $Y \neq 0$, en el cual podemos tomar a t y z como coordenadas afines. Fijados dos puntos distintos P_1, $P_2 \in E$, de coordenadas (t_1, z_1), (t_2, z_2), la pendiente de la recta que los une es

$$\lambda = \frac{z_2 - z_1}{t_2 - t_1}.$$

Podemos ver a λ como una función racional en la superficie $E \times E$. Su desarrollo en serie alrededor de (O, O) es[1]

$$\lambda(T_1, T_2) = \sum_{n=3}^{+\infty} A_{n-3} \frac{T_2^n - T_1^n}{T_2 - T_1} \in \mathbb{Z}[a_1, \ldots, a_6][[T_1, T_2]].$$

Observemos que λ no tiene términos de grado 0 o 1. La recta que pasa por P_1 y P_2 viene dada por la ecuación $Z = \lambda T + \mu$, donde $\mu = z_1 - \lambda t_1$. Si vemos a μ como función racional en $E \times E$, tenemos también que su desarrollo en serie es $\mu(T_1, T_2) = z_1(T_1) - \lambda(T_1, T_2)T_1 \in \mathbb{Z}[a_1, \ldots, a_6][[T_1, T_2]]$.

Sustituyendo la ecuación $Z = \lambda T + \mu$ en la ecuación de Weierstrass (más exactamente, en su transformada en términos de T y Z) obtenemos una ecuación cúbica en T, dos de cuyas raíces son t_1 y t_2. El cociente de los coeficientes de grado 2 y 3 será la suma de las raíces cambiada de signo, luego la tercera raíz es

$$t_3 = -t_1 - t_2 - \frac{a_1\lambda + a_3\lambda^2 + a_2\mu + 2a_4\lambda\mu + 3a_6\lambda^2\mu}{1 + a_2\lambda + a_4\lambda^2 + a_6\lambda^3}.$$

Así, t_3 es la coordenada t del tercer punto donde la recta corta a E. Dicho punto es $-P_1 - P_2$. Si consideramos a t_3 como función racional en $E \times E$, tenemos que t_3 es la composición de la función $(P_1, P_2) \mapsto -P_1 - P_2$ con la función $t \in k(E)$. La expresión anterior muestra que su desarrollo en serie alrededor de (O, O) cumple $t_3(T_1, T_2) \in \mathbb{Z}[a_1, \ldots, a_6][[T_1, T_2]]$, pues el desarrollo en serie del denominador es una unidad del anillo de series de potencias (porque λ no tiene término independiente). Más aún, como λ no tiene términos de grado 1, vemos que

$$t_3(T_1, T_2) = -T_1 - T_2 + \cdots$$

En términos de las coordenadas de Weierstrass, el opuesto de un punto (x, y) viene dado por $(x, -y - a_1x - a_3)$. Teniendo en cuenta que $t = -x/y$, la

[1] En principio no sabemos si λ es regular en (O, O), pero el monomorfismo de anillos $\mathcal{O}_{(O,O)}(E \times E) \longrightarrow \overline{k}[[T_1, T_2]]$ se extiende a un monomorfismo de cuerpos de $\overline{k}(E \times E)$ en el cuerpo de cocientes de $\overline{k}[[T_1, T_2]]$, y en este caso vemos que la imagen de λ es entera.

coordenada t de dicho opuesto es $i = x/(y+a_1x+a_3)$. Viendo a i como función racional en E, su desarrollo en serie alrededor de O es

$$i(T) = \frac{x(T)}{y(T)+a_1x(T)+a_3} = \frac{T^{-2}-a_1T^{-1}-\cdots}{-T^{-3}+a_1T^{-2}+\cdots},$$

luego

$$i(T) = -T+\cdots \in \mathbb{Z}[a_1,\dots,a_6][[T]].$$

Claramente $s = t_3 \circ i$, luego[2]

$$s(T_1,T_2) = i(t_3(T_1,T_2)) = T_1+T_2+\cdots \in \mathbb{Z}[a_1,\dots,a_6][[T_1,T_2]].$$

Con esto tenemos probada una parte del teorema siguiente:

Teorema 5.3 *Sea E una curva elíptica definida por una ecuación de Weierstrass con coeficientes $a_1,\dots,a_6$. Sean x, y las coordenadas de Weierstrass y $t = -x/y$. Sean $s(P_1,P_2) = t(P_1+P_2)$, $i(P) = t(-P)$, consideradas como funciones en $\overline{k}(E\times E)$ y $\overline{k}(E)$ respectivamente. Entonces sus series de Taylor tienen sus coeficientes en $\mathbb{Z}[a_1,\dots,a_6]$ y son de la forma*

$$s(T_1,T_2) = T_1+T_2+\cdots, \qquad i(T) = -T+\cdots$$

Además cumplen las propiedades siguientes:

a) $s(T_1,s(T_2,T_3)) = s(s(T_1,T_2),T_3)$,

b) $s(T_1,0) = T_1$, $s(0,T_2) = T_2$,

c) $s(T_1,i(T_1)) = 0$,

d) $s(T_1,T_2) = s(T_2,T_1)$.

Demostración: Falta demostrar las cuatro propiedades. Para la primera consideramos la función $u \in \mathcal{O}_{(O,O,O)}(E\times E\times E)$ definida mediante $u(P_1,P_2,P_3) = t(P_1+P_2+P_3)$.

Si $\phi_1 : E\times E\times E \longrightarrow E\times E$ viene dada por $\phi_1(P_1,P_2,P_3) = (P_1+P_2,P_3)$, tenemos que $u = \phi_1\circ s$, luego la serie de Taylor de u en (O,O,O) es

$$\tau(u) = \tau(s)(\tau(\overline{\phi}_1(t_1)),\tau(\overline{\phi}_1(t_2))) = s(s(T_1,T_2),T_3),$$

pero igualmente podemos probar que

$$\tau(u) = s(T_1,s(T_2,T_3)),$$

luego ambas series coinciden.

Para las otras propiedades se razona análogamente. Así, para probar b) consideramos $\phi(P) = (P,O)$, para c) tomamos $\phi(P) = (P,-P)$ y para d) tomamos $\phi(P_1,P_2) = (P_2,P_1)$. ∎

[2]Aquí usamos que si $\phi : V \longrightarrow W$ es regular y $\alpha \in \mathcal{O}_{\phi(P)}(W)$, la serie de Taylor de $\phi\circ\alpha$ en P es la composición de la serie de Taylor de α en $\phi(P)$ con las series de Taylor de $\phi\circ t_i$, donde $t_1,\dots,t_n$ es el sistema de parámetros locales considerado en $\phi(P)$. Concretamente, tomamos $\phi(P_1,P_2) = -P_1-P_2$ y $\alpha = i$, con lo que $\overline{\phi}(t) = t_3$.

5.2 Grupos formales

Los resultados de la sección anterior nos llevan a esta definición:

Definición 5.4 Un *grupo formal* (abeliano) sobre un dominio íntegro A es una serie de potencias $F \in A[[X,Y]]$ que satisface las propiedades siguientes:

a) $F(X,Y) = X + Y + \cdots$

b) $F(X,F(Y,Z)) = F(F(X,Y),Z)$,

c) $F(X,0) = X$, $F(0,Y) = Y$,

d) Existe una serie $i(X) \in A[[X]]$ tal que $F(X,i(X)) = 0$,

e) $F(X,Y) = F(Y,X)$.

Observemos que la serie $i(X)$ cuya existencia postula la definición es única, pues si $j(X)$ cumple lo mismo, entonces

$$j(X) = F(j(X),0) = F(j(X),F(X,i(X))) = F(F(j(X),X),i(X))$$
$$= F(F(X,j(X)),i(X)) = F(0,i(X)) = i(X).$$

En estos términos, lo que hemos demostrado en la sección anterior es que la serie de Taylor en (O,O) de la suma de una curva elíptica dada por una ecuación de Weierstrass con coeficientes $a_1,\dots,a_6$ es un grupo formal sobre el anillo $\mathbb{Z}[a_1,\dots,a_6]$.

Dos ejemplos más elementales de grupos formales son los siguientes:

El grupo formal aditivo Se trata del grupo dado por

$$F(X,Y) = X + Y.$$

Obviamente cumple todas las propiedades que exige la definición. La serie inversa es $i(X) = -X$. ∎

El grupo formal multiplicativo Se trata del grupo dado por

$$F(X,Y) = X + Y + XY = (1+X)(1+Y) - 1.$$

La serie inversa es $i(X) = \dfrac{1}{1+X} - 1 = \displaystyle\sum_{n=1}^{\infty} (-1)^n X^n$. ∎

Vemos que un grupo formal es una serie de potencias que "definiría" una operación de grupo en caso de converger. Más adelante veremos que en algunos casos podemos garantizar la convergencia y obtenemos ciertamente estructuras de grupo. Para no llegar a expresiones forzadas conviene usar letras distintas cuando pensamos en un grupo formal como tal y cuando lo usamos como la serie de potencias que es. Así hablaremos de un grupo formal G con suma $F \in A[X,Y]$, si bien, conjuntistamente $G = F$.

Definición 5.5 Un *homomorfismo* $h : G_1 \longrightarrow G_2$ entre dos grupos formales sobre un dominio íntegro A es una serie de potencias $h = H \in A[[T]]$ sin término independiente tal que

$$H(F_1(X,Y)) = F_2(H(X), H(Y)).$$

Observemos que si G es un grupo formal, entonces la serie $H(T) = T$ define un homomorfismo formal $1 : G \longrightarrow G$ (la identidad en G). Así mismo, dados dos homomorfismos formales $h_1 : G_1 \longrightarrow G_2$ y $h_2 : G_2 \longrightarrow G_3$ podemos definir $h_1 \circ h_2 : G_1 \longrightarrow G_3$ mediante la composición de series $H_2(H_1(T))$.

Diremos que $h : G_1 \longrightarrow G_2$ es un *isomorfismo* si existe otro homomorfismo $h' : G_2 \longrightarrow G_1$ tal que $h \circ h' = h' \circ h = 1$.

Ejemplo Consideremos una isogenia $\phi : E_1 \longrightarrow E_2$ entre dos curvas elípticas sobre un cuerpo k con coordenadas de Weierstrass x_1, y_1 y x_2, y_2 respectivamente. Sean $t_i = -x_i/y_i \in \overline{k}(E_i)$ y consideremos los grupos formales G_i determinados por las series de Taylor $s_i(X,Y)$ según el teorema 5.3. Sea $H(T)$ la serie de Taylor en O de la función $\phi \circ t_2$. Vamos a comprobar que define un homomorfismo formal $h : G_1 \longrightarrow G_2$.

Ante todo, como $t_2(\phi(O)) = O$, es claro que H no tiene término independiente. Ahora basta observar la coincidencia de las funciones

$$E_1 \times E_1 \xrightarrow{+} E_1 \xrightarrow{\phi} E_1 \xrightarrow{t_2} k, \qquad E_1 \times E_1 \xrightarrow{\phi\times\phi} E_1 \times E_1 \xrightarrow{+} E_2 \xrightarrow{t_2} k$$

La serie de Taylor en (O,O) de la primera es $s_1(H(X), H(Y))$, mientras que la de la segunda es $H(s_2(X,Y))$.

Si tenemos isogenias $\phi_1 : E_1 \longrightarrow E_2$ y $\phi_2 : E_2 \longrightarrow E_3$, entonces el homomorfismo formal asociado a $\phi_1 \circ \phi_2$ es la composición $h_1 \circ h_2$ de los homomorfismos formales asociados a cada ϕ_i. En efecto, se trata del homomorfismo formal determinado por la serie de Taylor de $\phi_1 \circ \phi_2 \circ t_3$, que es la composición de la serie de Taylor de $\phi_2 \circ t_3$ (o sea, $H_2(T)$), con la serie de Taylor de $\phi_1 \circ t_2$ (que es $H_1(T)$).

Es obvio que la isogenia identidad induce el homomorfismo formal identidad, luego si una isogenia es un isomorfismo, su homomorfismo formal inducido también lo es. ∎

Si G es un grupo formal y $m \in \mathbb{Z}$, podemos definir inductivamente homomorfismos $\overline{m} : G \longrightarrow G$ mediante

$$\overline{0}(T) = 0, \quad \overline{m+1}(T) = F(\overline{m}(T), T), \quad \overline{m-1}(T) = F(\overline{m}(T), i(T)).$$

Una inducción rutinaria muestra que realmente son homomorfismos. Para estudiarlos conviene demostrar un resultado general:

Teorema 5.6 *Sea A un dominio íntegro y $a \in A$ una unidad. Entonces, para toda serie de potencias $F(T) = aT + \cdots$ existe una única serie $G(T) \in A[T]$ tal que $F(G(T)) = T$. Dicha serie cumple también que $G(F(T)) = T$.*

DEMOSTRACIÓN: Vamos a construir una sucesión $G_n(T) \in A[T]$ tal que

$$F(G_n(T)) \equiv T \text{ (mód } T^{n+1}), \qquad G_{n+1}(T) = G_n(T) \text{ (mód } T^{n+1}).$$

La segunda condición implica que la sucesión $G_n(T)$ es de Cauchy en $A[[T]]$, luego converge a una serie $G(T)$ que, en virtud de la primera condición, cumple $F(G(T)) = T$ (la aplicación $G \mapsto F(G)$ es claramente continua).

En primer lugar tomamos $G_1(T) = a^{-1}T$. Supuesta definida $G_n(T)$, vamos a probar que podemos tomar $G_{n+1}(T) = G_n(T) + cT^{n+1}$ para cierto $c \in A$. Esto garantiza la segunda propiedad cualquiera que sea c. Respecto a la primera, calculamos

$$\begin{aligned} F(G_{n+1}(T)) &= F(G_n(T) + cT^{n+1}) \equiv F(G_n(T)) + acT^{n+1} \\ &\equiv T + bT^{n+1} + acT^{n+1} \text{ (mód } T^{n+2}), \end{aligned}$$

para cierto $b \in A$, por hipótesis de inducción. Basta tomar $c = -b/a$.

Así tenemos la existencia de una serie G tal que $F(G(T)) = T$. Ahora bien, la construcción muestra que $G(T) = a^{-1}T + \cdots$ luego podemos aplicar la parte ya probada a G y obtener una serie H tal que $G(H(T)) = T$. Entonces $F(G(H(T))) = F(T)$, luego $H(T) = F(T)$ y, por consiguiente, $G(F(T)) = T$.

Finalmente, si G' es cualquier otra serie que cumpla $F(G'(T)) = T$, entonces

$$G(T) = G(F(G'(T))) = G'(T).$$

■

Veamos ahora un par de propiedades sobre la multiplicación por un entero en un grupo formal que serán importantes después.

Teorema 5.7 *Sea G un grupo formal sobre un dominio íntegro A y $m \in \mathbb{Z}$.*

a) La multiplicación por m en G cumple $\overline{m}(T) = mT + \cdots$

b) Si m es una unidad en A, entonces $\overline{m} : G \longrightarrow G$ es un isomorfismo.

DEMOSTRACIÓN: El apartado a) se demuestra fácilmente por inducción sobre m. Para $m < 0$ observamos que

$$0 = F(T, i(T)) = T + i(T) + \cdots,$$

luego $i(T) = -T + \cdots$

Para probar b) usamos el teorema anterior, que nos da una serie $G(T)$ tal que $\overline{m}(G(T)) = G(\overline{m}(T)) = T$. Sólo hemos de probar que G es un homomorfismo. En efecto,

$$\overline{m}(F(G(X), G(Y))) = F(\overline{m}(G(X)), \overline{m}(G(Y))) = F(X, Y),$$

luego

$$F(G(X), G(Y)) = G(\overline{m}(F(G(X), G(Y)))) = G(F(X, Y)).$$

■

Diferenciales invariantes Ahora vamos a demostrar que todo grupo formal tiene una forma diferencial invariante por traslaciones. El espacio de formas diferenciales sobre $A[[X]]$ puede identificarse con el propio $A[[X]]$, pues una forma diferencial es simplemente una expresión de tipo $\omega = H(X)\,dX$, con $H(X) \in A[[X]]$.

Por otra parte, la traslación en un grupo formal G es simplemente su suma vista como función de una variable, es decir, $\tau_Y(X) = F(X,Y)$.

La transformada de una forma ω por la traslación de G es la forma diferencial

$$\overline{\tau}_Y(\omega) = H(F(X,Y))\frac{\partial F}{\partial X}\,dT \in A[[X,Y]]$$

Diremos que una forma diferencial ω sobre $A[[X]]$ es *invariante* respecto de un grupo formal G sobre A si $\overline{\tau}_Y(\omega) = \omega$. Diremos que $\omega = H(X)\,dX$ está *normalizada* si $H(0) = 1$.

Ejemplo Es fácil ver que una diferencial invariante para el grupo formal aditivo $F(X,Y) = X + Y$ es $\omega = dX$, mientras que una diferencial invariante para el grupo formal multiplicativo $F(X,Y) = (1+X)(1+Y) - 1$ es

$$\omega = \frac{dX}{1+X} = (1 - X + X^2 - X^3 + \cdots)\,dX.$$

∎

Teorema 5.8 *Si G es un grupo formal sobre un dominio íntegro A, entonces existe una única diferencial invariante normalizada, dada por la fórmula*

$$\omega = \left(\left.\frac{\partial F}{\partial X}\right|_{(0,X)}\right)^{-1} dX.$$

Cualquier otra diferencial invariante es de la forma $a\omega$, con $a \in A$.

Demostración: Supongamos que $\omega = H(X)\,dX$ es una diferencial invariante para G. Entonces

$$H(F(X,Y))\frac{\partial F}{\partial X} = H(X).$$

En particular, haciendo $X = 0$ tenemos que

$$H(Y)\left.\frac{\partial F}{\partial X}\right|_{(0,Y)} = H(0).$$

Cambiando la Y por la X queda

$$H(X)\left.\frac{\partial F}{\partial X}\right|_{(0,X)} = H(0).$$

La derivada de F tiene término independiente 1, luego $H(X)$ está completamente determinado por $H(0)$ y ha de ser

$$H(X) = \left(\left.\frac{\partial F}{\partial X}\right|_{(0,X)} \right)^{-1} H(0).$$

Basta probar que la diferencial del enunciado es invariante. Esto equivale a demostrar que

$$\left(\left.\frac{\partial F}{\partial X}\right|_{(0,F(X,Y))} \right)^{-1} \frac{\partial F}{\partial X} = \left(\left.\frac{\partial F}{\partial X}\right|_{(0,X)} \right)^{-1}.$$

Para ello derivamos respecto de Z (aplicando la regla de la cadena) la igualdad $F(Z,F(X,Y)) = F(F(Z,X),Y)$, con lo que obtenemos la relación

$$\left.\frac{\partial F}{\partial X}\right|_{(Z,F(X,Y))} = \left.\frac{\partial F}{\partial X}\right|_{(F(Z,X),Y)} \left.\frac{\partial F}{\partial X}\right|_{(Z,X)}.$$

Ahora hacemos $Z = 0$, con lo que la expresión se reduce a

$$\left.\frac{\partial F}{\partial X}\right|_{(0,F(X,Y))} = \frac{\partial F}{\partial X} \left.\frac{\partial F}{\partial X}\right|_{(0,X)},$$

que es la relación que necesitábamos. ∎

Teorema 5.9 *Sea E una curva elíptica sobre un cuerpo k con coordenadas de Weierstrass x, y, sea $t = -x/y$, sea $\omega = h\,dt$ una diferencial invariante de E y sea $H(T)$ la serie de Taylor de h respecto del parámetro t. Entonces $\omega' = H(T)\,dT$ es una diferencial invariante del grupo formal asociado a E.*

Demostración: Sabemos que ω es una diferencial de primera clase de E, es decir, no tiene ceros ni polos, por lo que podemos suponer que $h(O) = 1$. Para cada punto $P \in E$ tenemos que

$$h = (h \circ \tau_P)\frac{d\tau_P}{dt},$$

luego

$$1 = h(O) = h(P)\left.\frac{d\tau_P}{dt}\right|_O.$$

La derivada de τ_P es, por definición, la única función en $\overline{k}(E)$ cuya serie de Taylor en O es la derivada de la serie de Taylor de τ_P. Así pues, la serie de Taylor de h es

$$H(T) = \left(\left.\frac{\partial s}{\partial X}\right|_{(0,X)} \right)^{-1},$$

donde $s(X,Y)$ es la serie de Taylor de la suma en E. Según el teorema anterior, $H(T)\,dT$ es la diferencial invariante normalizada del grupo formal de E. ∎

Si $h : G_1 \longrightarrow G_2$ es un homomorfismo entre grupos formales sobre un dominio íntegro A, a cada forma diferencial $\omega = G(X)\,dX$ sobre $A[[X]]$ le podemos asignar otra dada por

$$\overline{h}(\omega) = G(H(x))\frac{dH}{dX}\,dX.$$

Esta transformación relaciona de forma natural las diferenciales invariantes:

Teorema 5.10 *Si $h : G_1 \longrightarrow G_2$ es un homomorfismo entre grupos formales, entonces las diferenciales invariantes normalizadas ω_{G_1} y ω_{G_2} verifican la relación*

$$\overline{h}(\omega_{G_2}) = H'(0)\,\omega_{G_1}.$$

DEMOSTRACIÓN: Vamos a ver que $\overline{h}(\omega_{G_2})$ es una diferencial invariante para G_1. En efecto, si $\omega_{G_2} = G(X)\,dX$, entonces

$$\begin{aligned}
\overline{\tau}_Y(\overline{h}(\omega_{G_2})) &= \overline{\tau}_Y\left(G(H(X))\frac{dH}{dX}dX\right)\\
&= G(H(F_1(X,Y)))\left.\frac{dH}{dX}\right|_{F_1(X,Y)}\frac{\partial F_1}{\partial X}\,dX\\
&= G(F_2(H(X),H(Y)))\frac{\partial H(F_1(X,Y))}{\partial X}\,dX\\
&= G(F_2(H(X),H(Y)))\frac{\partial F_2(H(X),H(Y))}{\partial X}\,dX\\
&= G(F_2(H(X),H(Y)))\left.\frac{\partial F_2}{\partial X}\right|_{(H(X),H(Y))}\frac{dH}{dX}\,dX\\
&= G(H(X))\frac{dH}{dX}\,dX = \overline{h}(\omega_{G_2}).
\end{aligned}$$

El término independiente de

$$G(H(X))\frac{dH}{dX}$$

es $H'(0)$, luego el teorema anterior implica la igualdad del enunciado. ∎

Veamos otra aplicación de la diferencial invariante que nos será útil un poco más adelante:

Teorema 5.11 *Si G es un grupo formal sobre un dominio íntegro A y $p \in \mathbb{Z}$ es un primo, entonces existen series de potencias $f(T)$, $g(T) \in A[[T]]$ tales que $f(0) = g(0) = 0$ y $\overline{p}(T) = pf(T) + g(T^p)$.*

DEMOSTRACIÓN: Sea $\omega(T) = H(T)\,dT$ la diferencial invariante normalizada de G. Según el teorema 5.7 tenemos que $\overline{p}(T) = pT + \cdots$, luego el teorema anterior implica que

$$pH(T)\,dT = p\omega(T) = H(\overline{p}(T))\overline{p}'(T)\,dT.$$

La serie $H(\overline{p}(T))$ tiene término independiente igual a 1, luego tiene inversa en $A[[T]]$, luego $\overline{p}'(T) \in pA[[T]]$. Esto significa que cada término a_nT^n en $\overline{p}(T)$ cumple $p \mid a_n$ o bien $p \mid n$. ∎

Logaritmos y exponenciales En este apartado trabajaremos con grupos formales sobre un dominio íntegro A de característica 0 y llamaremos $K = A\otimes\mathbb{Q}$. Podemos identificar a K con el subanillo del cuerpo de cocientes de A formado por las fracciones con denominador entero.

Definición 5.12 Sea G un grupo formal sobre un dominio íntegro A de característica 0, sea $K = A \otimes \mathbb{Q}$ y sea $\omega = (1 + c_1X + c_2X^2 + \cdots)dX$ la diferencial invariante normalizada. El *logaritmo formal* de G es la serie de potencias

$$\log_G(X) = \int \omega = X + \frac{c_1}{2}X^2 + \frac{c_2}{3}X^3 + \cdots \quad \in K[[X]].$$

La *exponencial formal* de G es la única serie de potencias $\exp_G(X) \in K[[X]]$ que cumple (según 5.6)

$$\log_G(\exp_G(X)) = \exp_G(\log_G(x)) = X.$$

Por ejemplo, si G es el grupo formal multiplicativo, sabemos que su diferencial invariante normalizada es

$$\omega = (1 - X + X^2 - X^3 + \cdots)\, dX,$$

luego

$$\log_G(X) = X - \frac{X^2}{2} + \frac{X^3}{3} - \frac{X^4}{4} + \cdots$$

Observemos que si $A = \mathbb{C}$ ésta es la serie de potencias de la función holomorfa $\log(1 + X)$, cuya inversa es la función $e^X - 1$, cuya serie de potencias es a su vez

$$\exp_G(X) = \sum_{n=1}^{\infty} \frac{X^n}{n!}.$$

Vamos a probar que esta serie es realmente $\exp_G(X)$ (para cualquier anillo de coeficientes A). Lo que sabemos es que, como series en $\mathbb{C}[[X]]$, se cumple

$$\exp_G(\log_G(X)) = X,$$

luego esta identidad se cumple también en $\mathbb{Q}[[X]]$, luego en $K[[X]]$.

El interés de los logaritmos formales se debe en gran parte al teorema siguiente:

Teorema 5.13 *Sea G un grupo formal sobre un dominio íntegro A de característica 0 y sea G_a el grupo formal aditivo sobre A. Si consideramos a ambos como grupos formales sobre $K = A \otimes \mathbb{Q}$, entonces $\log_G : G \longrightarrow G_a$ es un isomorfismo.*

Demostración: Sea $\omega = P(X)\, dX$ la diferencial invariante normalizada de G. Entonces

$$P(F(X,Y))\frac{\partial F}{\partial X} = P(X).$$

Equivalentemente,

$$\frac{\partial}{\partial X}\log_G F(X,Y) = \frac{\partial}{\partial X}\log_G(X).$$

Esto implica que

$$\log_G F(X,Y) = \log_G(X) + H(Y),$$

para cierta $H(Y) \in K[[Y]]$. Haciendo $X = 0$ vemos que $H(Y) = \log_G Y$, luego

$$\log_G F(X,Y) = \log_G(X) + \log_G(Y).$$

Esto demuestra que $\log_G$ es un homomorfismo. Cambiando X e Y por $\exp_G(X)$ y $\exp_G(Y)$ obtenemos

$$\log_G F(\exp_G(X), \exp_G(Y)) = X + Y,$$

y tomando exponenciales en ambos miembros

$$F(\exp_G(X), \exp_G(Y)) = \exp_G(X+Y),$$

luego $\exp_G$ también es un homomorfismo y es el inverso de $\log_G$, luego el logaritmo es un isomorfismo. ∎

Nota Es fácil ver que en la construcción de las funciones exponencial y logarítmica de un grupo formal G no hemos usado en ningún momento la conmutatividad de G (es decir, la propiedad $F(X,Y) = F(Y,X)$), pero tenemos la relación

$$F(X,Y) = \exp_G(\log_G(X) + \log_G(Y)),$$

de donde se sigue que todo grupo formal sobre un dominio íntegro de característica 0 es necesariamente abeliano. Esto no es cierto en característica prima. ∎

La restricción a característica nula es obviamente necesaria para trabajar con logaritmos debido a que en la definición aparecen necesariamente denominadores enteros. Dichos denominadores los conocemos explícitamente en el caso de la función logaritmo, pero no sucede lo mismo con su inversa, la exponencial. El teorema siguiente nos permitirá obtener una estimación.

Teorema 5.14 *Sea A un dominio íntegro de característica 0 y sea*

$$F(X) = \sum_{n=1}^{\infty} \frac{a_n}{n!} X^n$$

una serie de potencias con $a_n \in A$ y de modo que a_1 sea una unidad. Entonces la serie inversa $G(X)$ determinada por $F(G(X)) = X$ es de la forma

$$G(X) = \sum_{n=1}^{\infty} \frac{b_n}{n!} X^n,$$

con $b_n \in A$.

DEMOSTRACIÓN: Derivando en $F(G(X)) = X$ obtenemos que

$$\left.\frac{dF}{dX}\right|_{G(X)} \frac{dG}{dX} = 1,$$

y haciendo $X = 0$ vemos que $a_1 b_1 = 1$. Como a_1 es una unidad de A, concluimos que $b_1 \in A$. Derivando de nuevo queda

$$\left.\frac{d^2F}{dX^2}\right|_{G(X)} \left(\frac{dG}{dX}\right)^2 + \left.\frac{dF}{dX}\right|_{G(X)} \frac{d^2G}{dX^2} = 0.$$

Una simple inducción muestra que para todo $n \geq 2$ podemos expresar

$$\left.\frac{dF}{dX}\right|_{G(X)} \frac{d^nG}{dX^n}$$

como polinomio con coeficientes enteros de las series

$$\left.\frac{d^iF}{dX^i}\right|_{G(X)},\ i = 1, \ldots, n, \quad \frac{d^jG}{dX^j},\ j = 1, \ldots, n-1.$$

Al evaluar en $X = 0$ obtenemos $a_1 b_n$ como polinomio con coeficientes enteros de los a_i y los b_j, luego otra inducción nos da que $b_n \in A$. ∎

En particular:

Teorema 5.15 *Sea G un grupo formal sobre un dominio íntegro A de característica 0. Entonces*

$$\log_G(X) = \sum_{n=1}^{\infty} \frac{a_n}{n} X^n, \qquad \exp_G(X) = \sum_{n=1}^{\infty} \frac{b_n}{n!} X^n,$$

con a_n, $b_n \in A$, $a_1 = b_1 = 1$.

5.3 Grupos formales sobre cuerpos métricos

En esta sección K será un cuerpo métrico discreto completo, $\mathcal{O}$ será su anillo de enteros, $\mathfrak{P}$ el ideal maximal de $\mathcal{O}$ y $k = \mathcal{O}/\mathfrak{P}$ y G será un grupo formal sobre $\mathcal{O}$. La peculiaridad de este contexto es que cualquier serie de potencias $H(X_1, \ldots, X_n) \in \mathcal{O}[[X_1, \ldots, X_n]]$ converge obviamente en $\mathfrak{P} \times \cdots \times \mathfrak{P}$.

Por consiguiente podemos definir el grupo $G(\mathfrak{P})$ formado por el conjunto $\mathfrak{P}$ con la suma determinada por $\alpha +_G \beta = F(\alpha, \beta)$ (y el inverso $-_G\alpha = i(\alpha)$). Más aún, es claro que $G(\mathfrak{P}^n) = \mathfrak{P}^n$ es un subgrupo de $G(\mathfrak{P})$.

También es claro que todo homomorfismo $h : G_1 \longrightarrow G_2$ entre grupos formales sobre $\mathcal{O}$ induce un homomorfismo de grupos $h : G_1(\mathfrak{P}^n) \longrightarrow G_2(\mathfrak{P}^n)$, que será un isomorfismo si y sólo si h lo es. Si $m \in \mathbb{Z}$, la multiplicación $\overline{m} : G \longrightarrow G$ induce la multiplicación por m en $G(\mathfrak{P}^n)$.

Observemos que si G_a es el grupo formal aditivo, entonces $G_a(\mathfrak{P}^n)$ es simplemente $\mathfrak{P}^n$ con la estructura de grupo que posee como ideal de $\mathfrak{O}$. Por otra parte, si G_m es el grupo formal multiplicativo, entonces el grupo (aditivo) $G_m(\mathfrak{P}^n)$ es isomorfo al grupo (multiplicativo) de unidades $U_n = 1 + \mathfrak{P}^n$, a través del isomorfismo natural dado por $a \mapsto 1 + a$.

No podemos usar el teorema 5.13 para probar que todos los grupos $G(\mathfrak{P})$ son isomorfos a $\mathfrak{P}$ porque el isomorfismo formal entre G y el grupo aditivo G_a está definido sobre un anillo mayor que $\mathfrak{O}$, sobre el cual no convergen necesariamente las series exponencial y logarítmica. Luego veremos que, bajo ciertas hipótesis, ambas series convergen sobre subgrupos suficientemente pequeños, con lo que lo máximo que tendremos será un isomorfismo $G(\mathfrak{P}^r) \cong \mathfrak{P}^r$, para r suficientemente grande. De momento observemos que si G es un grupo formal arbitrario, el grupo cociente

$$G(\mathfrak{P}^n)/G(\mathfrak{P}^{n+1})$$

es independiente de G. En efecto, si $x, y \in \mathfrak{P}^n$, entonces

$$x +_G y = F(x, y) = x + y + \cdots \equiv x + y \ (\text{mód } \mathfrak{P}^{n+1}),$$

luego la identidad es un isomorfismo $G(\mathfrak{P}^n)/G(\mathfrak{P}^{n+1}) \cong \mathfrak{P}^n/\mathfrak{P}^{n+1}$.

Ahora vamos a estudiar los elementos de torsión de los grupos $G(\mathfrak{P}^n)$. El teorema siguiente representará un papel importante en el estudio de la aritmética de las curvas elípticas.

Teorema 5.16 *Sea $p \geq 0$ la característica del cuerpo de restos k. Entonces todo elemento de torsión de $G(\mathfrak{P})$ tiene orden potencia de p. (En particular, si $p = 0$ no hay elementos de torsión).*

Demostración: Basta probar que no existen elementos de torsión en $G(\mathfrak{P})$ de orden primo con p (lo cual no es ninguna restricción si $p = 0$). En efecto, si existe un elemento de torsión $x \in G(\mathfrak{P})$ y su orden m no es potencia de p, entonces, o bien $p = 0$ y entonces m es primo con p, o bien multiplicando x por la mayor potencia de p que divide a m obtenemos un nuevo elemento de torsión de orden primo con p.

Supongamos, pues, que $x \in G(\mathfrak{P})$ cumple $mx = 0$, con $(m, p) = 1$, y hemos de probar que $x = 0$. Como p es la característica de k, el hecho de que m sea primo con p equivale a que $m \notin \mathfrak{P}$, luego es una unidad de $\mathfrak{O}$ y el teorema 5.7 nos da que la multiplicación (formal) por m es un isomorfismo formal de G en sí mismo. Por consiguiente, la multiplicación por m es también un isomorfismo $m : G(\mathfrak{P}) \longrightarrow G(\mathfrak{P})$. Esto nos permite concluir que $x = 0$. ∎

Estudiemos ahora con más detalle el caso en que k tiene característica prima.

Teorema 5.17 *Supongamos que K tiene característica 0 pero que la característica del cuerpo de restos k es un primo p. Sea $x \in G(\mathfrak{P})$ un elemento de torsión de orden p^n. Entonces*

$$v(x) \leq \frac{v(p)}{p^n - p^{n-1}}.$$

DEMOSTRACIÓN: Por el teorema 5.11 sabemos que $\overline{p}(T) = pf(T) + g(T^p)$, con $f(0) = g(0) = 0$, y por 5.7 ha de ser $f(T) = T + \cdots$

Demostramos el teorema por inducción sobre n. Si $x \neq 0$ pero $\overline{p}(x) = 0$, entonces $0 = pf(x) + g(x^p)$. Esto implica que

$$v(px) = v(pf(x)) = v(g(x^p)) \geq v(x^p) = pv(x),$$

luego $v(p) \geq (p-1)v(x)$, como teníamos que probar.

Supongamos que el teorema es cierto para n y que x tiene orden $n+1$. Entonces

$$v(\overline{p}(x)) = v(pf(x) + g(x^p)) \geq \text{mín}\{v(px), v(x^p)\}.$$

Como $\overline{p}(x)$ tiene orden n, por hipótesis de inducción

$$v(\overline{p}(x)) \leq \frac{v(p)}{p^n - p^{n-1}},$$

luego

$$\text{mín}\{v(px), v(x^p)\} \leq \frac{v(p)}{p^n - p^{n-1}},$$

pero no puede ser que

$$v(px) = v(p) + v(x) \leq \frac{v(p)}{p^n - p^{n-1}},$$

luego ha de ser

$$v(x^p) = pv(x) \leq \frac{v(p)}{p^n - p^{n-1}},$$

como había que probar. ∎

Así, por ejemplo, un grupo formal sobre $\mathbb{Z}_p$ no puede tener elementos de torsión si $p \geq 3$, y si $p = 2$ tendrá a lo sumo elementos de orden 2.

Vamos a estudiar ahora la convergencia de las series exponencial y logarítmica. Nos apoyaremos en dos resultados técnicos.

Teorema 5.18 *Sea $p \in \mathbb{Z}$ un primo tal que $v(p) \geq 1$. Entonces, para cada natural $n \geq 1$ se cumple*

$$v(n!) \leq \frac{(n-1)v(p)}{p-1}.$$

DEMOSTRACIÓN: En efecto:

$$v(n!) = \sum_{i=1}^{\infty} E(n/p^i)v(p),$$

donde $E(x)$ denota la parte entera (observemos que $E(n/p^i)$ es el número de múltiplos de p^i menores o iguales que n). Por consiguiente

$$v(n!) \leq \sum_{i=1}^{E(\log_p n)} nv(p)/p^i = \frac{nv(p)}{p-1}(1 - p^{-E(\log_p n)}) \leq \frac{(n-1)v(p)}{p-1},$$

donde usamos que $np^{-E(\log_p n)} \geq 1$, pues equivale a que $E(\log_p n) \leq \log_p n$. ∎

Teorema 5.19 *Supongamos que K tiene característica 0 y sea $p \in \mathbb{Z}$ un primo tal que $v(p) > 0$.*

a) Toda serie de la forma

$$f(T) = \sum_{n=1}^{\infty} \frac{a_n}{n} T^n, \qquad a_n \in \mathcal{O}$$

converge en la bola abierta de centro 0 y radio 1 (esto es, sobre los puntos $x \in \mathcal{O}$ que cumplen $v(x) > 0$).

b) Toda serie de la forma

$$g(T) = \sum_{n=1}^{\infty} \frac{b_n}{n!} T^n, \qquad b_n \in \mathcal{O}$$

converge en la bola abierta de centro 0 formada por los puntos $x \in \mathcal{O}$ tales que $v(x) > v(p)/(p-1)$. Además $v(g(x)) = v(x)$.

Demostración: Para la serie $f(T)$ basta observar que

$$v(a_n x^n/n) = v(a_n) + nv(x) - v(n) \geq nv(x) - v(p)\log_p n,$$

luego

$$\lim_n v(a_n x^n/n) = +\infty,$$

lo que implica la convergencia de $f(x)$.

Igualmente,

$$v(b_n x^n/n!) = v(b_n) + nv(x) - v(n!) \geq nv(x) - \frac{(n-1)v(p)}{p-1}$$

$$= v(x) + (n-1)\left(v(x) - \frac{v(p)}{p-1}\right).$$

Si x cumple la hipótesis de b), entonces

$$\lim_n v(b_n x^n/n!) = +\infty$$

y $g(x)$ converge. Más aún, tenemos que si $n \geq 2$ entonces

$$v(b_n x^n/n!) > v(x),$$

luego el valor de todas las sumas parciales de $g(x)$ es $v(x)$, de donde podemos concluir que $v(g(x)) = v(x)$. ∎

Con esto podemos probar el siguiente resultado sobre logaritmos:

Teorema 5.20 *Si K tiene característica 0 y $p \geq 2$ es un primo tal que $v(p) > 0$, entonces el logaritmo formal induce un homomorfismo $\log_G : G(\mathfrak{P}) \longrightarrow K$. Si r es un natural $r > v(p)/(p-1)$, entonces el logaritmo se restringe a un isomorfismo $\log_G : G(\mathfrak{P}^r) \longrightarrow G_a(\mathfrak{P}^r)$ (donde G_a es el grupo formal aditivo).*

DEMOSTRACIÓN: El teorema anterior se aplica claramente a la serie de potencias $\log_G T$, luego ésta converge en $G(\mathfrak{P})$. El hecho de que la serie sea un homomorfismo de grupos formales se traduce inmediatamente en que la aplicación inducida sobre $G(\mathfrak{P})$ es un homomorfismo de grupos. Veamos ahora que se restringe a un homomorfismo $\log_G : G(\mathfrak{P}^r) \longrightarrow G_a(\mathfrak{P}^r)$.

Si $x \in G(\mathfrak{P}^r)$ y $n \geq 1$, o bien $1 \leq n \leq p-1$, en cuyo caso

$$v(x^n/n) = nv(x) \geq nr \geq r,$$

o bien $2 \leq p \leq n$, en cuyo caso

$$\begin{aligned} v\left(\frac{x^n}{n}\right) - r &\geq nv(x) - v(n) - r \geq (n-1)r - v(p)\log_p n \\ &> \frac{n-1}{p-1}v(p) - v(p)\,\frac{\log n}{\log p} \\ &= \frac{v(p)(n-1)}{\log p}\left(\frac{\log p}{p-1} - \frac{\log n}{n-1}\right) \geq 0, \end{aligned}$$

donde usamos que la función $\log t/(t-1)$ es monótona decreciente para $t \geq 2$.

Esto prueba que cada término de la serie $\log_G x$ está en $\mathfrak{P}^r$, luego lo mismo le sucede a la suma.

El teorema anterior implica también que serie exponencial $\exp_G(T)$ define un homomorfismo

$$\exp_G : G_a(\mathfrak{P}^r) \longrightarrow G(\mathfrak{P}^r).$$

El hecho de que $\exp_G(T)$ y $\log_G(T)$ sean formalmente inversos se traduce en que las aplicaciones inducidas también lo son. ∎

5.4 Grupos formales en característica prima

En esta sección A será un dominio íntegro de característica prima p.

Definición 5.21 Sea $h : G_1 \longrightarrow G_2$ un homomorfismo formal entre dos grupos formales sobre A. La *altura* de h (representada por $\operatorname{alt} h$) es el mayor número natural a tal que $H(T) = H^*(T^{p^a})$, para cierta serie $H^*(T) \in A[[T]]$. Convenimos que si $h = 0$ entonces $\operatorname{alt} h = +\infty$. La *altura* de un grupo formal G es la altura de la multiplicación por p en G.

Por ejemplo, el teorema 5.7 implica que si $(m,p) = 1$ entonces $\operatorname{alt} \overline{m} = 0$, mientras que el teorema 5.11 implica que $\operatorname{alt} \overline{p} \geq 1$. Por lo tanto, la altura de un grupo formal (sobre un dominio de característica prima) es siempre mayor o igual que 1.

El teorema siguiente es útil para calcular alturas:

Teorema 5.22 *Sea $h : G_1 \longrightarrow G_2$ un homomorfismo formal entre grupos formales sobre A.*

a) Si $H'(0) = 0$ *entonces* $H(T) = H^*(T^p)$*, para cierto* $H^*(T) \in A[[T]]$.

b) Si $H(T) = H^*(T^{p^a})$*, con* $a = \operatorname{alt} h$*, entonces* $(H^*)'(0) \neq 0$.

Demostración: a) Sean ω_1 y ω_2 las diferenciales invariantes normalizadas de G_1 y G_2. Entonces, por el teorema 5.10,

$$0 = H'(0)\,\omega_1 = \overline{h}(\omega_2) = (1 + \cdots)H'(T)\,dT,$$

luego $H'(T) = 0$, de donde se sigue claramente que $H(T) = H^*(T^p)$.

b) Llamemos $q = p^a$. Sea $F_1(X,Y) = \sum a_{ij}X^iY^j$ la suma en G_1 y definamos $F_1^{(q)}(X,Y) = \sum a_{ij}^q X^iY^j$. Vamos a ver que esta serie determina un grupo formal $G_1^{(q)}$. En efecto, consideremos unas nuevas indeterminadas S, T y llamemos $X = S^q$, $Y = Y^q$, de modo que podemos identificar $A[[X,Y]] = A[[S^q, T^q]]$. En estos términos, $F_1^{(q)}(X,Y) = F_1(S,T)^q$.

Igualmente podemos llamar $Z = U^q$. Así, elevando a q la identidad

$$F_1(F_1(S,T),U)) = F_1(S, F_1(T,U))$$

obtenemos la identidad correspondiente para $F_1^{(q)}$. Igualmente se comprueban las propiedades restantes.

Veamos ahora que H^* determina un homomorfismo $h^* : G_1^{(q)} \longrightarrow G_2$. En efecto,

$$H^*(F_1^{(q)}(X,Y)) = H^*(F_1(S,T)^q) = H(F_1(S,T)) = F_2(H(S),H(T))$$

$$= F_2(H^*(S^q), H^*(T^q)) = F_2(H^*(X), H^*(Y)).$$

Por lo tanto, si $(H^*)'(0) = 0$, el apartado anterior nos da que podemos expresar $H^*(T) = H^{**}(T^p)$, pero entonces $H(T) = H^{**}(T^{p^{a+1}})$ y a no sería la altura de h, contradicción. ∎

Con esto podemos probar que las alturas se comportan bien con la composición de homomorfismos:

Teorema 5.23 *Sean* $G_1 \xrightarrow{h_1} G_2 \xrightarrow{h_2} G_3$ *dos homomorfismos formales entre grupos formales sobre* A*. Entonces* $\operatorname{alt}(h_1 \circ h_2) = \operatorname{alt} h_1 + \operatorname{alt} h_2$.

Demostración: Sean $H_1(T) = H_1^*(T^{p^{a_1}})$, $H_2(T) = H_2^*(T^{p^{a_2}})$, donde a_1 y a_2 son las alturas correspondientes. Entonces

$$H_2(H_1(T)) = H_2^*(\tilde{H}_1^*(T^{p^{a_1+a_2}})),$$

donde $\tilde{H}_1^*$ resulta de elevar a p^{a_2} los coeficientes de H_1^*. Por el teorema anterior los coeficientes de grado 1 de H_1^* y H_2^* son no nulos, luego lo mismo sucede con el de $\tilde{H}_1^*$ y el de la composición $H_2^*(\tilde{H}_1^*(T))$ (notemos que las series no tienen término independiente por la definición de homomorfismo formal). El apartado a) del teorema anterior implica entonces que $\operatorname{alt}(h_1 \circ h_2) = a_1 + a_2$. ∎

Ahora probamos que la noción de altura se corresponde con el grado de inseparabilidad de una isogenia:

Teorema 5.24 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia entre dos curvas elípticas definidas por ecuaciones de Weierstrass sobre un cuerpo de característica prima p y sea $h : G_1 \longrightarrow G_2$ el homomorfismo inducido entre los grupos formales de las curvas. Entonces el grado de inseparabilidad de ϕ es $p^{\operatorname{alt} h}$.*

DEMOSTRACIÓN: Del teorema 1.22 se sigue que toda isogenia se descompone como composición de una aplicación de Frobenius y una isogenia separable. Como el grado de inseparabilidad es multiplicativo y la altura es aditiva, basta demostrar el teorema cuando ϕ está en uno de estos dos casos.

Si ϕ es la aplicación de Frobenius de grado p^r, entonces su grado de inseparabilidad es p^r. Por otra parte, el homomorfismo h es el determinado por la serie de Taylor de $\phi \circ t_2 = t_1^{p^r}$, que es T^{p^r}, luego $\operatorname{alt} h = r$.

Si ϕ es separable, sea $\omega = h\,dt$ una diferencial invariante de E_2 y sea $\omega(T)$ la diferencial correspondiente del grupo formal de E_2. El teorema 5.10 nos da que una diferencial invariante del grupo formal de E_1 es

$$\bar{h}(\omega(T)) = H'(0)\,\omega.$$

Por otra parte, de la definición de $\bar{h}(\omega(T))$ se sigue fácilmente que es la diferencial formal asociada a $\bar{\phi}(\omega)$, que es una forma no nula por el teorema 1.19. Así pues, $H'(0) \neq 0$, luego $\operatorname{alt} h = 0$. ∎

En particular, la altura del grupo formal de una curva elíptica sobre un cuerpo de característica p (es decir, la altura de la multiplicación por p) sólo puede ser 1 o 2, ya que el grado de la multiplicación por p es p^2, luego su grado de inseparabilidad ha de ser p o p^2. Según el teorema 3.15 esta altura será 1 si la curva es ordinaria y 2 si es supersingular.

Capítulo VI

Curvas elípticas sobre cuerpos locales

Si tenemos una curva elíptica E definida mediante una ecuación de Weierstrass con coeficientes enteros, una forma de estudiarla es considerar las curvas (más simples) definidas por la misma ecuación sobre el cuerpo $\mathbb{Z}/p\mathbb{Z}$, para cada primo p. La curva obtenida de este modo se llama reducción módulo p de la curva E. Lo primero que hemos de tener presente es que no tiene por qué ser una curva elíptica, pero ello dependerá únicamente de si p divide o no al discriminante de la ecuación. En particular, los primos con "mala reducción" serán siempre un número finito.

Este proceso de reducción presenta otros inconvenientes que hemos de considerar. Por ejemplo, una misma curva elíptica puede admitir distintas ecuaciones de Weierstrass con coeficientes enteros, cada una con un discriminante diferente, por lo que la buena o mala reducción en un primo puede depender de la ecuación que elijamos.

Para tratar estas cuestiones conviene considerar primero las curvas definidas sobre los cuerpos p-ádicos $\mathbb{Q}_p$ como paso intermedio de la reducción, es decir, una ecuación con coeficientes enteros puede verse también como una ecuación con coeficientes en el anillo de enteros p-ádicos $\mathbb{Z}_p$, la cual a su vez puede reducirse a una ecuación en el cuerpo $\mathbb{Z}_p/(p) \cong \mathbb{Z}/p\mathbb{Z}$.

Más en general, ahora nos proponemos estudiar las curvas elípticas definidas sobre un cuerpo métrico discreto completo. En todo el capítulo, salvo que se indique lo contrario, K denotará un cuerpo métrico completo[1] respecto de una valoración v, llamaremos $\mathcal{O}$ a su anillo de enteros, U a su grupo de unidades, $\mathfrak{P}$ a su ideal primo, $k = \mathcal{O}/\mathfrak{P}$ a su cuerpo de restos y π denotará un primo de $\mathcal{O}$, de modo que $\mathfrak{P} = (\pi)$. Supondremos además que K y k son perfectos. En todos los casos de interés el cuerpo K tendrá característica 0 y k será finito, luego estas hipótesis no son restrictivas.

[1] En realidad la completitud no es necesaria en la primera sección.

6.1 Ecuaciones minimales

Sea E/K una curva elíptica y sean x, $y \in K(E)$ dos generadores que satisfagan una ecuación de Weierstrass

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \qquad a_i \in K.$$

Si $u \in K$, el teorema 2.6 nos da que los generadores $(u^{-2}x, u^{-3}y)$ satisfacen una ecuación similar con coeficientes $u^i a_i$. Así, si tomamos $v(u)$ suficientemente grande podemos obtener una ecuación de Weierstrass para E con coeficientes en $\mathcal{O}$. En consecuencia, el discriminante Δ también será entero.

Definición 6.1 Una ecuación de Weierstrass es *entera* si sus coeficientes están en $\mathcal{O}$. Un cambio de variables de la forma

$$X = u^2X' + r, \qquad Y = u^3Y' + su^2X' + t, \qquad u \neq 0$$

es *semientero* si u, r, s, $t \in \mathcal{O}$. Diremos que es *entero* si además $u \in U$ (lo que equivale a que el cambio inverso sea también semientero).

Acabamos de ver que toda curva elíptica admite una ecuación de Weierstrass entera. Es obvio que un cambio de variables semientero transforma una ecuación de Weierstrass entera en otra ecuación de Weierstrass entera. Respecto al recíproco, tenemos el teorema siguiente:

Teorema 6.2 *Consideremos dos curvas elípticas K-isomorfas E/K y E'/K definidas por ecuaciones de Weierstrass enteras con discriminantes Δ y Δ' tales que $v(\Delta) \geq v(\Delta')$. Entonces existe un cambio de variables, necesariamente semientero, que transforma la ecuación de E en la de E'. El cambio será entero si y sólo si $v(\Delta) = v(\Delta')$.*

Demostración: Por el teorema 2.3, existen u, r, s, $t \in K$, con $u \neq 0$ que determinan un cambio de variables que transforma E en E'. El teorema 2.6 nos da que sus discriminantes satisfacen la relación $u^{12}\Delta' = \Delta$, luego tenemos que $v(\Delta) = v(\Delta') + 12v(u)$, de donde $v(u) \geq 0$ y $v(u) = 0$ si y sólo si $v(\Delta) = v(\Delta')$. Falta probar que r, s, $t \in \mathcal{O}$.

De la relación $u^4b'_6 = b_6 + 2rb_4 + r^2b_2 + 4r^3$ deducimos que

$$v(r) + v(2b_4 + rb_2 + 4r^2) \geq 0.$$

Si fuera $v(r) < 0$ tendríamos que tener $v(2b_4 + rb_2 + 4r^2) \geq 0$, luego

$$v(rb_2 + 4r^2) = v(r) + v(b_2 + 4r) \geq 0,$$

luego $v(b_2 + 4r) \geq 0$ y, en cualquier caso, tendría que ser $v(4r) \geq 0$. Razonando igualmente con la transformación de b_8 concluimos que $v(3r) \geq 0$. Ahora bien, como ha de ser $v(2) = 0$ o $v(3) = 0$, podemos afirmar que $v(r) \geq 0$. La transformación de a_2 nos da que $v(s) \geq 0$ y la de a_6 nos permite concluir que $v(t) \geq 0$. ∎

Dada una ecuación de Weierstrass entera, tiene sentido considerar su reducción módulo $\mathfrak{p}$, que determinará una curva elíptica sobre k si y sólo si $v(\Delta) = 0$. Si no es así, tenemos dos posibilidades: que la reducción tenga una cúspide o que tenga un nodo, el cual a su vez puede ser racional o irracional. Demostraremos que si elegimos la ecuación con $v(\Delta)$ mínimo no sólo tenemos más posibilidades de que el resultado sea una curva elíptica, sino que, cuando no lo es, el tipo de la curva resultante no depende de la elección de la ecuación. Esto lo veremos más adelante. Ahora discutiremos la noción de ecuación minimal:

Definición 6.3 Si E/K es una curva elíptica, una *ecuación de Weierstrass minimal* para E es una ecuación de Weierstrass entera para E tal que $v(\Delta)$ es el menor posible. Representaremos por $\delta_{\mathfrak{p}}$ a dicho valor mínimo.

Del teorema anterior se desprende que toda ecuación de Weierstrass entera de una curva elíptica se puede transformar en una ecuación minimal mediante un cambio de variables semientero, así como que dos ecuaciones minimales de una misma curva elíptica están relacionadas por un cambio de variables entero.

En general, para determinar si una ecuación entera es o no minimal hemos de considerar su terna de *valores covariantes* $(v(c_4), v(c_6), v(\Delta))$. (Las cantidades c_4, c_6 y, a veces, Δ se llaman *covariantes* de la ecuación.) Notemos que $v(c_4)$ y $v(c_6)$ son números naturales o bien $+\infty$.

El teorema siguiente muestra que es muy fácil decidir si una ecuación dada es o no minimal, salvo si el cuerpo de restos cumple $\operatorname{car} k \leq 3$. En tal caso necesitaremos un análisis más detallado, pero el teorema siguiente será de todos modos el punto de partida.

Teorema 6.4 *Sea E/K una curva elíptica definida por una ecuación entera con valores covariantes $(a, b, c) = (v(c_4), v(c_6), v(\Delta))$.*

a) La ecuación es minimal si y sólo si E no admite otra ecuación entera con valores covariantes $(a-4, b-6, c-12)$. En particular esto sucede si $a < 4$ o bien $b < 6$ o bien $c < 12$.

b) Si $\operatorname{car} K > 3$, entonces la ecuación es minimal si y sólo si no existe una ecuación de Weierstrass entera con covariantes $c_4' = c_4\pi^{-4}$, $c_6' = c_6\pi^{-6}$.

c) Si $\operatorname{car} k > 3$, entonces la ecuación es minimal si y sólo si se cumple una de las tres desigualdades $v(c_4) < 4$, $v(c_6) < 6$, $v(\Delta) < 12$.

Demostración: a) Si existe tal ecuación, entonces la ecuación dada no es minimal, pues la nueva cumple $v(\Delta') = v(\Delta) - 12 < v(\Delta)$. Recíprocamente, si la ecuación dada no es minimal, se puede transformar en una minimal mediante un cambio de variables semientero, pero entonces los covariantes de la ecuación minimal cumplen

$$v(c_4') = v(c_4) - 4v(u), \quad v(c_6') = v(c_6) - 6v(u), \quad v(\Delta') = v(\Delta) - 6v(u).$$

Como la ecuación de partida no es minimal, la tercera igualdad implica que $v(u) > 0$, luego $v(c_4) \geq 4$, $v(c_6) \geq 6$, $v(\Delta) \geq 12$. Por consiguiente, el cambio $Y = \pi^2 Y'$, $X = \pi^3 X'$ transforma la ecuación de partida en otra con valores covariantes $(a-4, b-6, c-12)$.

b) Supongamos que existe una ecuación con los covariantes indicados. Entonces su discriminante es $\Delta' = \Delta\pi^{-12}$ y su invariante es $j' = j$, luego dicha ecuación define una curva elíptica E' isomorfa a E. Por otra parte, E y E' son k-isomorfas a curvas definidas por las ecuaciones de tipo b dadas por (2.5). Estas ecuaciones no necesariamente enteras, pero que satisfacen las hipótesis del teorema 2.47, luego las curvas E y E' son K-isomorfas. Vemos, pues, que E es K-isomorfa a una curva elíptica definida por una ecuación de Weierstrass con discriminante $v(\Delta') < v(\Delta)$, luego la ecuación de E no es minimal.

Recíprocamente, si la ecuación de E no es minimal, por a) sabemos que ha de ser $v(c_4) \geq 4$ y $v(c_6) \geq 6$, y entonces el cambio $X = \pi^2 X'$, $Y = \pi^3 Y'$ nos da una ecuación con los covariantes requeridos.

c) El apartado a) muestra que la condición es suficiente. Supongamos que $v(c_4) \geq 4$, $v(c_6) \geq 6$, $v(\Delta) \geq 12$. Entonces $c_4' = c_4\pi^{-4}$ y $c_6' = c_6\pi^{-6}$ son enteros.

La hipótesis $\operatorname{car} k \neq 2, 3$ equivale a que $\operatorname{car} K \neq 2, 3$ (lo que nos permite usar el apartado anterior) y además $v(2) = v(3) = 0$. Ahora bien, si $\operatorname{car} K > 3$ cualquier par de valores c_4', c_6' son los covariantes de una ecuación de Weierstrass —la dada por (2.5)—, que en nuestro caso es entera y con discriminante no nulo, luego por b) concluimos que la ecuación dada no es minimal. ∎

Observemos que la relación $\Delta = (c_4^3 - c_6^2)/12^3$ hace que si $\operatorname{car} k > 3$ las condiciones $v(\Delta) \geq 12$, $v(c_4) \geq 4$ impliquen $v(c_6) \geq 6$, luego la minimalidad equivale de hecho a que $v(\Delta) < 12$ o $v(c_4) < 4$.

Ejemplo La ecuación $Y^2 + XY + Y = X^3 + X^2 + 22X - 9$ sobre el cuerpo $\mathbb{Q}_p$ de los números p-ádicos tiene discriminante $\Delta = -2^{15} \cdot 5^2$ y $c_4 = -5 \cdot 211$. Por lo tanto, el teorema anterior garantiza que es minimal para todo primo p. ∎

Ejemplo Si $p > 3$ es un primo, la ecuación $Y^2 = X^3 - 3(1+p^N)X - 2(1+p^N)$ sobre el cuerpo $\mathbb{Q}_p$ cumple $\Delta = 2^6 \cdot 3^3 \cdot (1 + p^N) \cdot p^N$, luego $v_p(c_4) = 0$ y $v_p(\Delta) = N$. La primera condición implica que es minimal, y de este modo vemos que existen curvas elípticas con δ_p arbitrariamente grande. ∎

Nos falta encontrar un criterio para decidir si una ecuación dada es minimal cuando $\operatorname{car} k \leq 3$. Abordaremos el problema a través del apartado b) del teorema anterior (limitándonos al caso en que $\operatorname{car} K = 0$). La cuestión es, pues, determinar bajo qué condiciones existe una curva elíptica con unos covariantes dados c_4, $c_6 \in \mathcal{O}$.

Si $\operatorname{car} k > 3$ la condición necesaria y suficiente es que $c_4^3 - c_6^2 \neq 0$. La condición es necesaria porque el discriminante de la ecuación correspondiente ha de ser $\Delta = (c_4^3 - c_6^2)/1728$, y es suficiente porque basta tomar la curva dada por la ecuación (2.5). Ahora bien, el caso que necesitamos resolver es justo el opuesto: $\operatorname{car} k \leq 3$.

Teorema 6.5 (Kraus) *Supongamos que* $\operatorname{car} K = 0$ *y que* $\operatorname{car} k = p \leq 3$. *Sean* c_4, c_6, $\Delta \in \mathcal{O}$ *tales que* $c_4^3 - c_6^2 = 1728\Delta \neq 0$. *Consideramos los polinomios*

$$\Psi_2(X) = X^3 - 3c_4X - 2c_6, \qquad \Psi_3(X) = X^4 - 6c_4X^2 - 8c_6X - 3c_4^2.$$

Entonces, c_4 *y* c_6 *son los covariantes de una ecuación de Weierstrass entera si y sólo si se cumple una de las dos condiciones siguientes:*

a) $p = 3$ *y existe* $\theta \in \mathcal{O}$ *tal que* $\Psi_2(\theta) \equiv 0 \pmod{27}$.

b) $p = 2$ *y existen* θ, $\tau \in \mathcal{O}$ *tales que*

$$\Psi_3(\theta^2) \equiv 0 \pmod{256}, \qquad \Psi_2(\theta^2) \equiv -16\tau^2 \pmod{64}.$$

DEMOSTRACIÓN: a) Supongamos que existe una ecuación con covariantes c_4 y c_6. Entonces basta tomar $\theta = b_2$, pues entonces

$$\Psi_2(\theta) = 27 \cdot 16b_6 \equiv 0 \pmod{27}.$$

Supongamos ahora que existe $\theta \in \mathcal{O}$ tal que $\Psi_2(\theta) \equiv 0 \pmod{27}$. La ecuación (2.5) tiene covariantes c_4 y c_6, pero no es necesariamente entera. Si le aplicamos el cambio $X = X' + \theta/12$ se convierte en

$$Y^2 = X^3 + \frac{\theta}{4}X^2 + \frac{\theta^2 - c_4}{48}X + \frac{\Psi_2(\theta)}{1728}.$$

Los covariantes de esta ecuación son los mismos. Vamos a probar que los coeficientes son enteros, para lo cual sólo hemos de ver que $\theta^2 \equiv c_4 \pmod{3}$.

Tenemos que $\theta^3 \equiv 3c_4\theta + 2c_6 \pmod{27}$ (por la hipótesis sobre $\Psi_2(\theta)$) y por la relación entre c_4, c_6 y Δ se cumple también que $c_4^3 \equiv c_6^2 \pmod{27}$. De aquí se sigue que

$$(c_4\theta + c_6)^3 \equiv 3c_6(c_4\theta + c_6)^2 \pmod{27}.$$

(Basta desarrollar el cubo y sustituir las dos congruencias precedentes.) Por consiguiente $c_4\theta + c_6 \equiv 0 \pmod{3}$. Por último, una comprobación rutinaria nos da la identidad

$$(X^2 - c_4)^3 = \Psi_2(X)(X^3 + 2c_6) + 3(c_4X + c_6)^2 + 1728\Delta,$$

que al ser evaluada en $X = \theta$ nos da la congruencia buscada: $\theta^2 \equiv c_4 \pmod{3}$.

b) Si existe la ecuación, aplicando si es necesario el cambio de variables entero $X = X' - a_2/3$ podemos suponer que $a_2 = 0$. Tomamos $\theta = a_1$, $\tau = a_3$. Entonces

$$\Psi_3(\theta^2) = 2^6 \cdot 3^3(b_6\theta^2 - b_4^2) = 2^8 \cdot 3^3 \cdot b_8 \equiv 0 \pmod{256}.$$

(Para comprobar la primera igualdad sustituimos c_4 y c_6 por sus valores en términos de b_2, b_4, b_6 y hacemos $b_2 = \theta^2$, para la segunda sustituimos b_6 y b_4 por sus valores en términos de los a_i.) Por otra parte,

$$\Psi_2(\theta^2) = 2^4 \cdot 3^3 \cdot b_6 \equiv -16\tau^2 \pmod{64}.$$

Supongamos ahora que ciertos enteros θ y τ cumplen las congruencias. Elegimos $a_1 = \theta$, $a_2 = 0$ y $a_3 = \tau$, con lo que basta probar que existen enteros a_4, a_6 tales que

$$c_4 = \theta^4 - 24(\tau\theta + 2a_4), \quad c_6 = -\theta^6 + 36\theta^2(\tau\theta + 2a_4) - 216(\tau^2 + 4a_6). \quad (6.1)$$

La primera ecuación nos da que

$$a_4 = \frac{\theta^4 - c_4 - 24\tau\theta}{48}.$$

Hemos de probar que a_4 es entero, para lo cual basta ver que

$$\theta^4 - c_4 - 8\tau\theta \equiv 0 \text{ (mód 16)}.$$

Usamos la identidad $\Psi_3(X) = 4X\Psi_2(X) - 3(X^2 - c_4)^2$, de la que obtenemos

$$4\theta^2\Psi_2(\theta^2) - 3(\theta^4 - c_4)^2 \equiv 0 \text{ (mód 256)}. \quad (6.2)$$

Por otro lado, las hipótesis implican que

$$4\Psi_2(\theta^2) \equiv -64\tau^2 \text{ (mód 256)}, \quad (6.3)$$

luego

$$(\theta^4 - c_4)^2 - (8\tau\theta)^2 \equiv 4(\theta^4 - c_4)^2 \text{ (mód 256)}.$$

Ahora bien, (6.2) y (6.3) implican que $64 \mid (\theta^4 - c_4)^2$, luego el miembro derecho de la última congruencia es 0. En definitiva:

$$(\theta^4 - c_4 + 8\tau\theta)(\theta^4 - c_4 - 8\tau\theta) \equiv 0 \text{ (mód 256)}.$$

Uno de los dos factores ha de ser 0 módulo 16, pero como $8 \equiv -8$ (mód 16), en realidad han de serlo los dos, lo que prueba la integridad de a_4.

La segunda ecuación de (6.1) determina el valor de a_6. Falta probar que también éste es entero. Multiplicándola por 2 queda

$$\begin{aligned} 1728a_6 &= -2c_6 - 2\theta^6 + 72\tau\theta^3 + 3\theta^2(\theta^4 - c_4 - 24\tau\theta) - 432\tau^2 \\ &\equiv \Psi_2(\theta^2) + 16\tau^2 \equiv 0 \text{ (mód 64)}. \end{aligned}$$

De aquí se sigue que $v(a_6) \geq 0$. ∎

Las condiciones del teorema de Kraus se simplifican notablemente cuando $v(p) = 1$ (por ejemplo, si $K = \mathbb{Q}_2$ o $K = \mathbb{Q}_3$). En efecto:

Teorema 6.6 *Supongamos que* $\operatorname{car} K = 0$, $\operatorname{car} k = p \leq 3$ *y que* $v(p) = 1$. *Sean* c_4, c_6, $\Delta \in \mathcal{O}$ *tales que* $c_4^3 - c_6^2 = 1728\Delta \neq 0$. *Entonces,* c_4 *y* c_6 *son los covariantes de una ecuación de Weierstrass entera si y sólo si se cumple una de las dos condiciones siguientes:*

a) $p = 3$ *y* $v(c_6) \neq 2$.

b) $p = 2$ y se cumple una de las dos condiciones siguientes:

- $v(c_4) = 0$ *y existe* $x \in \mathcal{O}$ *tal que* $c_6 \equiv -x^2$ (mód 4),
- $v(c_4) \geq 4$ *y existe* $x \in \mathcal{O}$ *tal que* $c_6 \equiv 8x^2$ (mód 32).

DEMOSTRACIÓN: Consideremos $p = 3$. La relación $c_4^3 - c_6^2 = 1728\Delta \neq 0$ implica $v(c_4) = v(c_6) = 0$ o bien $v(c_6) \geq 2$ (pues si 3 divide a uno de los dos covariantes, entonces los divide a ambos y $3^3 \mid c_6^2$, luego $3^2 \mid c_6$).

Si $v(c_4) = 0$ entonces $\theta = -c_6/c_4$ cumple

$$\Psi_2(\theta^2) = 1728\Delta c_6/c_4^3 \equiv 0 \text{ (mód 27)},$$

luego por el teorema anterior existe una ecuación entera con los invariantes dados. Si $v(c_6) \geq 3$ basta tomar $\theta = 0$. Supongamos ahora que $v(c_6) = 2$ (con lo que $v(c_4) \geq 1$) y veamos que no hay enteros θ y z tales que $\theta^3 - 3c_4\theta - 2c_6 = 27z$. En efecto, tendría que ser $v(\theta^3) \geq 2$, luego $v(\theta) \geq 1$, pero entonces $v(c_6) \geq 3$, contradicción.

Consideremos ahora $p = 2$. Como antes, ha de ser $v(c_4) = 0$ o bien $v(c_4) \geq 2$.

Si $v(c_4) = 0$ y existe una ecuación entera con los invariantes dados, entonces

$$-c_6 \equiv b_2^3 \equiv a_1^6 \text{ (mód 4)}.$$

Recíprocamente, si existe $x \in \mathcal{O}$ tal que $x^2 \equiv -c_6$ (mód 4), en particular c_6 es una unidad de $\mathcal{O}$, luego podemos expresar la congruencia en la forma

$$x^2 = -c_6(1 + 4a), \qquad a \in \mathcal{O}.$$

Similarmente, la relación entre c_4, c_6 y Δ puede expresarse en la forma

$$c_4^3 = c_6^2(1 + 64b), \qquad b \in V.$$

Tomamos $\theta = c_4/x$, con lo que

$$c_4 = \frac{\theta^4(1 + 4a)^2}{1 + 64b}, \qquad c_6 = -\frac{\theta^6(1 + 4a)^3}{(1 + 64b)^2}.$$

Al sustituir estas dos expresiones en las definiciones de $\Psi_2(X)$ y $\Psi_3(X)$ obtenemos que

$$\Psi_3(\theta^2) \equiv 0 \text{ (mód 256)}, \qquad \Psi_2(\theta^2) \equiv -16a^2\theta^6 \text{ (mód 64)},$$

luego basta tomar $\tau = a\theta^3$ y el teorema de Kraus garantiza la existencia de la ecuación.

Si $v(c_4) \geq 4$ y existe la ecuación, entonces la definición de c_4 implica que $v(b_2) \geq 2$ y $v(b_4) \geq 1$, luego

$$c_6 \equiv -216b_6 \equiv 8a_3^2 \text{ (mód 32)}.$$

Recíprocamente, si $c_6 \equiv 8x^2$ (mód 32), basta tomar $\theta = 0$ y $\tau = x$.

Falta probar que si $v(c_4) \geq 2$ y existe ecuación, entonces $v(c_4) \geq 4$. En efecto, tiene que ser $v(c_3) \geq 3$ y de la relación

$$\theta^8 - 6c_4\theta^4 - 8c_6\theta^2 - 3c_4^2 = 256z$$

se sigue que $v(\theta) \geq 1$ y de aquí que $v(c_4^2) \geq 7$, luego $v(c_4) \geq 4$. ∎

Notemos que $\mathbb{Z}_2/(4) \cong \mathbb{Z}/4\mathbb{Z}$ y $\mathbb{Z}_2/(32) \cong \mathbb{Z}/32\mathbb{Z}$, luego las condiciones del teorema anterior se reducen a resolver congruencias de enteros (racionales). Más aún, si $v_2(c_4) = 0$ entonces c_6 es impar, y la congruencia de este caso equivale a $c_6 \equiv -1$ (mód 4). Similarmente, si $v_2(c_4) \geq 4$ la congruencia se reduce a $c_6 \equiv 0,\ 8$ (mód 32).

Ejemplo La curva $E/\mathbb{Q}_3$ dada por la ecuación

$$Y^2 = X^3 - 324X - 243$$

cumple $c_4 = 2^4 \cdot 3^5$, $c_6 = 2^5 \cdot 3^8$, $\Delta = 2^4 \cdot 3^{12} \cdot 11 \cdot 23$. Vemos que $v_3(c_4) = 5 \geq 4$, $v_3(c_6) = 8 \geq 6$, $v_3(\Delta) = 12 \geq 12$, luego no cumple la condición 6.4 c), pero a pesar de ello es minimal por el apartado b) de dicho teorema. En efecto, hemos de ver que no existe ninguna ecuación de Weierstrass entera con covariantes $c_4' = 2^4 \cdot 3$, $c_6' = 2^5 \cdot 3^2$, lo cual es cierto por el teorema anterior, ya que $v_3(c_6') = 2$. ∎

Ejercicio: Usar el teorema anterior para construir una ecuación de Weierstrass entera sobre $\mathbb{Q}_2$ tal que $v_2(c_4) = 6$, $v_2(c_6) = 10$, $v_2(\Delta) = 12$. Probar que es minimal.

6.2 Reducción de curvas elípticas

Definición 6.7 Sea E/K una curva elíptica en $\mathrm{P}^2(K)$ definida por una ecuación de Weierstrass entera. Llamaremos *reducción de E/K módulo $\mathfrak{P}$* a la curva (tal vez singular) en $\mathrm{P}^2(k)$ definida por la ecuación resultante de tomar restos módulo $\mathfrak{P}$ en los coeficientes de la ecuación de E/K. La representaremos por $\tilde{E}/k$.

Aquí es crucial observar que la reducción $\tilde{E}/k$ depende de la ecuación de Weierstrass considerada, de modo que dos curvas elípticas isomorfas sobre K pueden tener reducciones no isomorfas sobre k. Sin embargo, el teorema 6.2 garantiza que si tomamos dos ecuaciones de Weierstrass minimales para una misma curva elíptica E/K, una se transforma en la otra mediante un cambio de coordenadas entero, el cual induce un cambio de coordenadas en $\mathrm{P}^2(k)$ que hace corresponder las dos reducciones.

En definitiva, para una curva elíptica arbitraria E/K, definimos la *reducción* $\tilde{E}/k$ de E/K módulo $\mathfrak{P}$ como la reducción de cualquier curva elíptica K-isomorfa a E/K definida por una ecuación de Weierstrass minimal, de modo que la reducción no depende (salvo k-isomorfismo) de la elección de la ecuación minimal. No obstante, conviene tener presente que cualquier ecuación de Weierstrass entera, minimal o no, admite una reducción.

En la práctica, cuando hablemos de la reducción de una curva elíptica definida por una ecuación de Weierstrass entera entenderemos que se trata de la reducción asociada a dicha ecuación, mientras que si no hablamos de ninguna ecuación en particular entenderemos que se trata de la reducción respecto a una ecuación minimal cualquiera de la curva dada.

Consideremos una curva E/K definida por una ecuación de Weierstrass entera. Multiplicando las coordenadas homogéneas de un punto de $\mathrm{P}^2(K)$ por una potencia adecuada de π podemos hacer que todas sean enteras y al menos una de ellas unitaria. Esto se traduce en que los restos módulo $\mathfrak{P}$ de estas coordenadas forman una terna no nula en k^3 y, en definitiva, tenemos una aplicación natural $\mathrm{P}^2(K) \longrightarrow \mathrm{P}^2(k)$, la cual se restringe a su vez a una aplicación $E(K) \longrightarrow \tilde{E}(k)$. A esta aplicación la llamaremos también *reducción* módulo $\mathfrak{P}$ y la representaremos por $P \mapsto \tilde{P}$.

Llamaremos $\tilde{E}_r(k)$ al conjunto de puntos regulares de $\tilde{E}/k$ con coordenadas en k. Sabemos que puede ser todo $\tilde{E}(k)$ o bien $\tilde{E}(k)$ menos un punto (que en ningún caso podrá ser el neutro). El teorema 2.30 nos da que $\tilde{E}_r(k)$ tiene estructura de grupo aunque la curva $\tilde{E}/k$ sea singular. Definimos

$$E_0(K) = \{P \in E(K) \mid \tilde{P} \in \tilde{E}_r(k)\}, \qquad E_1(K) = \{P \in E(K) \mid \tilde{P} = O\}.$$

El conjunto $E_0(K)$ es el de los puntos con *reducción regular,* mientras que $E_1(K)$ es el *núcleo* de la reducción.

Teorema 6.8 *Dada una curva E/K definida por una ecuación de Weierstrass entera, la reducción módulo $\mathfrak{P}$ determina una sucesión exacta de grupos abelianos:*

$$0 \longrightarrow E_1(K) \longrightarrow E_0(K) \longrightarrow \tilde{E}_r(k) \longrightarrow 0.$$

Demostración: Hemos de probar que $E_0(K)$ es un subgrupo de $E(K)$ y que la reducción es un epimorfismo de grupos. Evidentemente $\tilde{O} = O$, luego $O \in E_0(K)$. La fórmula

$$-(x, y, z) = (x, -y - a_1x - a_3z, z)$$

prueba que $\widetilde{-P} = -\tilde{P}$, de donde concluimos que si $P \in E_r(K)$ entonces también $\tilde{P} \in E_r(K)$.

Consideremos ahora tres puntos P, Q, $R \in E(K)$ tales que $P + Q + R = O$. Esto significa que son la intersección con E de una recta r. Pongamos que la ecuación de E/K es $F(X, Y, Z) = 0$.

Representamos cada uno de los tres puntos por una terna de coordenadas homogéneas enteras y al menos una de ellas unitaria. Consideremos en primer lugar el caso en que dos de los puntos son distintos, por ejemplo $P \neq Q$. Sea $Q - P = \pi^r v$, donde $r \geq 0$ y v tiene coordenadas enteras, al menos una unitaria. Entonces la recta r está formada por los puntos $L = P + tv$, con $t \in \overline{K}$ y se cumple que

$$F(P + Tv) = aT(T - \pi^r)(T - t_1),$$

donde $t_1 \in K$ es el parámetro que cumple $R = P + t_1 v$. Despejándolo a partir de la coordenada unitaria de v vemos que $t_1 \in \mathcal{O}$.

La reducción módulo $\mathfrak{P}$ del polinomio $F(P + Tv)$ tiene por raíces a los parámetros de los puntos donde la recta $\tilde{r}$ corta a $\tilde{E}$. En particular no puede ser idénticamente nula ($a \notin \mathfrak{P}$) y vemos entonces que dicha intersección la forman los puntos $\tilde{P}$, $\tilde{Q}$, $\tilde{R}$.

Llegamos a la misma conclusión si los tres puntos de partida son iguales. En tal caso tenemos que $3P = O$. Llamamos r a la tangente a E por P, cuyos puntos serán de la forma $L = P + tv$, para cierto vector v que podemos tomar de coordenadas enteras y al menos una unitaria. Ahora $F(P + Tv) = aT^3$ y como antes razonamos que $\tilde{P}$ es el único punto de intersección de $\tilde{r}$ con $\tilde{E}$.

En cualquier caso, si suponemos que P, $Q \in E_r(K)$, podemos concluir que $\tilde{P} + \tilde{Q} + \tilde{R} = O$, luego $\tilde{R} \in \tilde{E}(k)$, $R \in E_r(K)$ y

$$\widetilde{P + Q} = \widetilde{-R} = -\tilde{R} = \tilde{P} + \tilde{Q}.$$

Con esto queda probado que $E_r(K)$ es un subgrupo y que la reducción es un homomorfismo. Veamos que es suprayectivo.

Sea $f(X, Y) = Y^2 + a_1XY + a_3Y - X^3 - a_2X^2 - a_4X - a_6 = 0$ una ecuación de Weierstrass minimal para E/K. Tomemos un punto en $\tilde{E}_r(k)$, que podemos suponer distinto de $\tilde{O}$. Digamos que sus coordenadas de Weierstrass son (α, β). Entonces $\tilde{f}(\alpha, \beta) = 0$ y la regularidad se traduce —por ejemplo— en que

$$\frac{\partial \tilde{f}}{\partial X}(\alpha, \beta) \neq 0.$$

(La alternativa es que la derivada respecto de Y sea no nula.) Fijemos $y_0 \in \mathcal{O}$ tal que $\tilde{y}_0 = \beta$ y consideremos el polinomio $f(X, y_0) \in \mathcal{O}[X]$. Su imagen en $k[X]$ tiene una raíz simple en α, luego el lema de Hensel implica que existe $x_0 \in \mathcal{O}$ tal que $f(x_0, y_0) = 0$ y $\tilde{x}_0 = \alpha$. El punto $P = (x_0, y_0)$ está en E/K y $\tilde{P} \in \tilde{E}_r(k)$ es el punto de partida, luego la reducción es suprayectiva. ∎

Las curvas elípticas se pueden clasificar por su tipo de reducción módulo $\mathfrak{P}$:

Definición 6.9 Consideremos una curva elíptica E/K dada por una ecuación de Weierstrass entera.

a) E tiene *buena reducción* sobre K si $\tilde{E}/k$ es regular. En caso contrario se dice que tiene *mala reducción* sobre K.

b) E tiene *reducción multiplicativa* sobre K si $\tilde{E}/k$ tiene un nodo. En tal caso la reducción puede ser *racional* o *irracional* según lo sea dicho nodo.

c) E tiene *reducción aditiva* si $\tilde{E}/k$ tiene una cúspide.

Los términos "reducción multiplicativa" y "reducción aditiva" hacen referencia al teorema 2.30. Cuando hablemos del tipo de reducción de una curva elíptica en general, se entenderá que nos referimos al tipo de reducción de cualquiera de sus ecuaciones minimales.

Es fácil saber el tipo de reducción de una curva elíptica. El teorema siguiente es una consecuencia inmediata de 2.28 y 2.30.

Teorema 6.10 *Sea E/K una curva elíptica determinada por una ecuación de Weierstrass entera. Entonces:*

a) E tiene buena reducción si y sólo si $v(\Delta) = 0$.

b) E tiene reducción multiplicativa si y sólo si $v(\Delta) > 0$ y $v(c_4) = 0$.

c) E tiene reducción aditiva si y sólo si $v(\Delta) > 0$ y $v(c_4) > 0$.

Ejemplo Si $p \geq 5$ es un número primo, las curvas siguientes tienen las propiedades indicadas:

Ecuación minimal	Δ	c_4	Reducción en $\mathbb{Q}_p$
$Y^2 = X^3 + pX^2 + 1$	$-2^4 \cdot 3^3$	$2^8 \cdot p^2$	buena
$Y^2 = X^3 + X^2 + p$	$-2^4 \cdot 3^3 \cdot p^2$	2^8	multiplicativa
$Y^2 = X^3 + p$	$-2^4 \cdot 3^3 \cdot p^2$	0	aditiva

■

Observemos que si K'/K es una extensión no ramificada entonces la valoración de K' extiende a la de K y la reducción $E_0(K') \longrightarrow \tilde{E}_r(k')$ extiende a la reducción $E_0(K) \longrightarrow \tilde{E}_r(k)$. Estas reducciones determinan una única aplicación $E_0(K_{nr}) \longrightarrow \tilde{E}_r$, donde K_{nr} es la máxima extensión no ramificada de K y no indicamos el cuerpo de restos en $\tilde{E}_r$ porque éste es $\overline{k}$, la clausura algebraica de k, luego $\tilde{E}_r(\overline{k})$ es en realidad toda la curva $\tilde{E}$ (salvo el punto singular si lo hay). Tenemos que K_{nr} es un cuerpo métrico discreto no necesariamente completo, pero las propiedades de las reducciones sobre los grupos $E_0(K')$ implican las propiedades análogas sobre $E_0(K_{nr})$. Por ejemplo, la reducción es suprayectiva. Así, al extender K hasta K_{nr} obtenemos todos los puntos de $\tilde{E}_r$. Esto es necesario para tener la unicidad en el teorema siguiente:

Teorema 6.11 *Sean E/K y E'/K dos curvas elípticas definidas mediante ecuaciones de Weierstrass enteras con buena reducción módulo $\mathfrak{P}$ y $\phi : E \longrightarrow E'$ una isogenia definida sobre K. Entonces existe una única isogenia $\tilde{\phi} : \tilde{E} \longrightarrow \tilde{E}'$ definida sobre k tal que para todo punto $P \in E(K_{nr})$ se cumple $\tilde{\phi}(\tilde{P}) = \widetilde{\phi(P)}$.*

Demostración: La unicidad es evidente. Sólo hemos de probar la existencia. Sea $\pi \in \mathcal{O}$ tal que $v_{\mathfrak{P}}(\pi) = 1$. Digamos que ϕ viene definida por

$$\phi(X, Y, Z) = [F_1(X, Y, Z), F_2(X, Y, Z), F_3(X, Y, Z)],$$

para ciertas formas $F_i \in \mathcal{O}[X, Y, Z]$ del mismo grado. Llamemos $\tilde{F}_i \in k[X, Y, Z]$ a la reducción de F_i. Si las tres reducciones son idénticamente nulas sobre $\tilde{E}$, entonces $\tilde{F}_i = \tilde{H}_i \tilde{F}$, donde $F(X, Y, Z) = 0$ es la ecuación de Weierstrass que define a E/K. Por lo tanto $F_i = H_i F + \pi G_i$, para ciertas formas $G_i \in \mathcal{O}[X, Y, Z]$.

En tal caso $\phi(X, Y, Z) = [G_1(X, Y, Z), G_2(X, Y, Z), G_3(X, Y, Z)]$ es una definición alternativa de ϕ. Ahora bien, si llamamos N al menor natural tal que π^N

divide a todos los $F_i(P)$ cuando P recorre $E(K_{nr})$ (y consideramos coordenadas homogéneas enteras para P), vemos que el N correspondiente a las formas G_i es una unidad menor que el correspondiente a las formas F_i, luego tras un número finito de pasos llegaremos a unas formas que definen a ϕ con $N = 0$. Conservando la notación original F_i, lo que hemos probado es que podemos suponer que una de las formas $\tilde{F}_i$ no es idénticamente nula en $\tilde{E}$. Esto hace que

$$\tilde{\phi}(X,Y,Z) = [\tilde{F}_1(X,Y,Z), \tilde{F}_2(X,Y,Z), \tilde{F}_3(X,Y,Z)]$$

define una aplicación racional $\tilde{\phi} : \tilde{E} \longrightarrow \tilde{E}'$. Como las curvas son elípticas, será —de hecho— una aplicación regular. En principio no podemos asegurar que sea una isogenia, pues para ello hace falta que $\tilde{\phi}(O) = O$. El diagrama

$$\begin{array}{ccc} E(K_{nr}) & \xrightarrow{\phi} & E'(K_{nr}) \\ \downarrow & & \downarrow \\ \tilde{E} & \xrightarrow[\tilde{\phi}]{} & \tilde{E}' \end{array}$$

será conmutativo sobre todos los puntos $P \in E(K_{nr})$ tales que $\tilde{F}_i(\tilde{P}) \neq 0$ para algún i. Esto lo cumplen todos los puntos $\tilde{P}$ salvo a lo sumo un número finito de ellos. Fijamos uno $\tilde{P}$ y observamos que cuando $\tilde{Q}$ recorre los infinitos puntos que no anulan a alguna de las formas $\tilde{F}_i$, la suma $\tilde{P} + \tilde{Q}$ toma infinitos valores distintos, luego para alguno de ellos se cumplirá que $\tilde{P} + \tilde{Q}$ tampoco anula a alguna de las formas. Esto implica que el diagrama conmuta en P, en Q y en $P + Q$. Por consiguiente:

$$\tilde{\phi}(\tilde{P} + \tilde{Q}) = \widetilde{\phi(P + Q)} = \widetilde{\phi(P)} + \widetilde{\phi(Q)} = \tilde{\phi}(\tilde{P}) + \tilde{\phi}(\tilde{Q}).$$

Ahora bien, si $\tilde{\phi}(O) = U$, tenemos que $\phi^* = \tilde{\phi} - U$ es una aplicación regular que cumple $\phi^*(O) = O$, luego es una isogenia, luego $\phi^*(\tilde{P}+\tilde{Q}) = \phi^*(\tilde{P})+\phi^*(\tilde{Q})$. Explícitamente:

$$\tilde{\phi}(\tilde{P} + \tilde{Q}) - U = \tilde{\phi}(\tilde{P}) - U + \tilde{\phi}(\tilde{Q}) - U,$$

pero esto implica que $U = 0$, luego $\tilde{\phi}$ es una isogenia. Sólo falta demostrar que el diagrama conmuta sobre todos los puntos de $E(K_{nr})$. Fijado $P \in E(K_{nr})$, como antes podemos encontrar un punto Q tal que ni $\tilde{Q}$ ni $\tilde{P} + \tilde{Q}$ anulen a las tres formas $\tilde{F}_i$. Esto significa que el diagrama conmuta sobre $P + Q$ y sobre Q. Por consiguiente:

$$\tilde{\phi}(\tilde{P}) + \tilde{\phi}(\tilde{Q}) = \tilde{\phi}(\tilde{P} + \tilde{Q}) = \widetilde{\phi(P + Q)} = \widetilde{\phi(P)} + \widetilde{\phi(Q)} = \widetilde{\phi(P)} + \tilde{\phi}(\tilde{Q}),$$

luego $\tilde{\phi}(\tilde{P}) = \widetilde{\phi(P)}$, y el diagrama conmuta en P. ∎

De este teorema no se sigue que si las curvas E y E' son isógenas entonces sus reducciones también lo son, pues para ello faltaría justificar que la reducción

de una isogenia no nula es no nula. Esto es cierto, pero lo demostraremos en la sección siguiente.

Volvamos ahora al ejemplo de la página 147. Consideramos las curvas en sentido abstracto, es decir, sin prefijar ninguna ecuación de Weierstrass. Observemos que la tercera curva $Y^2 = X^3 + p$ tiene reducción aditiva en $\mathbb{Q}_p$, pero tiene buena reducción sobre $K = \mathbb{Q}_p(\sqrt[6]{p})$. En efecto, sobre K la ecuación ya no es minimal, sino que el cambio

$$X = \sqrt[3]{p}\,X', \qquad Y = \sqrt{p}\,Y'$$

la transforma en $Y^2 = X^3 + 1$, con discriminante $\Delta = -2^4 \cdot 3^3$, y así $v(\Delta) = 0$. Esto no es casual, sino que ilustra el teorema siguiente, para el que conviene introducir una clasificación alternativa de los tipos de reducción:

Definición 6.12 Diremos que una curva elíptica E/K tiene *reducción estable* módulo $\mathfrak{P}$ si tiene buena reducción módulo $\mathfrak{P}$. Diremos que tiene reducción *semiestable* si tiene buena reducción o bien reducción multiplicativa. Diremos que la reducción es *inestable* si es aditiva.

El teorema siguiente explica por qué conviene agrupar bajo un mismo concepto la buena reducción y la reducción multiplicativa:

Teorema 6.13 *Sea E/K una curva elíptica.*

a) Si E tiene reducción semiestable sobre K, entonces tiene el mismo tipo de reducción (buena o multiplicativa) sobre cualquier extensión finita de K.

b) Si E tiene reducción inestable sobre K, entonces existe una extensión finita de K sobre la que E tiene reducción semiestable.

Demostración: a) Sea K' una extensión finita de K. Fijemos una ecuación de Weierstrass minimal para E sobre K y consideremos un cambio de variables semientero

$$X = u^2X' + r, \qquad Y = u^3Y' + su^2X' + t, \qquad u,\ r,\ s,\ t \in \mathcal{O}'$$

que la transforme en una ecuación minimal sobre K'. Entonces

$$0 \le v'(\Delta') = v'(u^{-12}\Delta), \qquad 0 \le v'(c_4') = v'(u^{-4}c_4).$$

Por consiguiente,

$$0 \le v'(u) \le \min\{v'(\Delta)/12, v'(c_4)/4\}.$$

Si E tiene reducción semiestable, entonces tenemos que $v(\Delta) = 0$ o bien $v(c_4) = 0$, luego también $v'(\Delta) = 0$ o bien $v'(c_4) = 0$ y en ambos casos $v'(u) = 0$. Así pues, $v'(\Delta') = v'(\Delta)$ y $v'(c_4') = v'(c_4)$. Esto nos permite concluir que la reducción de E/K' es la misma que la de E/K.

b) Supongamos primeramente que $\operatorname{car} k \neq 2$. Entonces podemos aplicar el teorema 2.16, que nos permite sustituir a K por una extensión finita sobre la que E admita una ecuación de Weierstrass en forma normal de Legendre:

$$Y^2 = X(X-1)(X-\lambda), \qquad \lambda \neq 0, 1.$$

Es fácil ver que $\Delta = 16\lambda^2(\lambda-1)^2$, $c_4 = 16(\lambda^2-\lambda+1)$. Distinguimos tres casos:

Caso 1: $\lambda \in \mathcal{O}$, $\lambda \not\equiv 0, 1 \pmod{\mathfrak{P}}$. Entonces $v(\Delta) = 0$, por lo que la ecuación es minimal y E tiene buena reducción.

Caso 2: $\lambda \in \mathcal{O}$, $\lambda \equiv 0, 1 \pmod{\mathfrak{P}}$. Entonces $v(\Delta) > 0$ y $v(c_4) = 0$, luego la ecuación es minimal y su reducción es multiplicativa.

Caso 3: $\lambda \notin \mathcal{O}$. Sea $r = -v(\lambda)$, de modo que $v(\pi^r\lambda) = 0$. Adjuntando a K si es necesario el elemento $\pi^{1/2}$ podemos hacer el cambio $X = \pi^{-r}X'$, $Y = \pi^{-3r/2}Y'$, lo que nos da la ecuación

$$Y^2 = X(X-\pi^r)(X-\pi^r\lambda).$$

Esta ecuación tiene coeficientes enteros y cumple $v(\Delta) > 0$, $v(c_4) = 0$, luego es minimal y su reducción es multiplicativa.

Si $\operatorname{car} k = 2$ razonamos igualmente con la forma normal de Deuring. En una extensión de K la curva E admite una ecuación de la forma

$$Y^2 + \alpha XY + Y = X^3, \qquad \alpha^3 \neq 27.$$

Se cumple que $\Delta = \alpha^3 - 27$, $c_4 = \alpha(\alpha^3 - 24)$. De nuevo distinguimos tres casos:

Caso 1: $\alpha \in \mathcal{O}$, $\alpha \not\equiv 27 \pmod{\mathfrak{P}}$. Entonces $v(\Delta) = 0$, luego la ecuación es minimal y tiene buena reducción.

Caso 2: $\alpha \in \mathcal{O}$, $\alpha \equiv 27 \pmod{\mathfrak{P}}$. Entonces $v(\Delta) > 0$, $c_4 \equiv 81 \not\equiv 0 \pmod{\mathfrak{P}}$, luego $v(c_4) = 0$ y así la ecuación es minimal y la reducción multiplicativa.

Caso 3: $\alpha \notin \mathcal{O}$. Tomamos $r = -v(\alpha)$ y hacemos el cambio $X = \pi^{-2r}X'$, $Y = \pi^{-3r}Y'$, con lo que la nueva ecuación es

$$Y^2 + \pi^r\alpha XY + \pi^{3r}Y = X^3,$$

que cumple $\Delta = \pi^{9r}((\pi^r\alpha)^3 - 27\pi^{3r}) \equiv 0 \pmod{\mathfrak{P}}$ y

$$c_4 = \pi^r\alpha((\pi^r\alpha)^3 - 24\pi^{3r}) \equiv (\pi^r\alpha)^4 \not\equiv 0 \pmod{\mathfrak{P}}.$$

Así pues, $v(\Delta) > 0$, $v(c_4) = 0$, la ecuación es minimal y su reducción es multiplicativa. ∎

Todavía podemos decir más sobre el tipo de reducción de una curva en una extensión:

Teorema 6.14 *Una curva elíptica E/K tiene reducción estable en una extensión finita de K si y sólo si $j(E) \in \mathcal{O}$.*

DEMOSTRACIÓN: Supongamos que E tiene reducción estable en una extensión K' de K. Entonces $v'(\Delta') = 0$ y $v'(c_4') \geq 0$, luego $v'(j(E)) \geq 0$ y esto implica que $v(j(E)) \geq 0$.

Recíprocamente, supongamos que $j(E) \in \mathcal{O}$. Consideremos primeramente el caso en que $\operatorname{car} k \neq 2$. Entonces podemos extender K de modo que E admita una ecuación en forma normal de Legendre E_λ. Entonces

$$j(E) = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}.$$

Así pues, λ cumple

$$(\lambda^2 - \lambda + 1)^3 - 2^{-8}\lambda^2(\lambda - 1)^2 = 0.$$

De aquí se sigue que $\lambda \in \mathcal{O}$, $\lambda \not\equiv 0, 1 \pmod{\mathfrak{P}}$ y como en el teorema anterior concluimos que E tiene reducción estable.

Si $\operatorname{car} k = 2$ consideramos una ecuación en forma normal de Deuring, de forma que

$$j(E) = \frac{\alpha^3(\alpha^3 - 24)^3}{\alpha^3 - 27}.$$

Razonando análogamente vemos que $\alpha \in \mathcal{O}$ y $\alpha^3 \not\equiv 27 \pmod{\mathfrak{P}}$, luego la ecuación es minimal y su reducción es estable. ∎

Por último observamos que para que cambie el tipo de reducción de una curva con reducción inestable es necesario que la extensión de K sea ramificada:

Teorema 6.15 *Sea E/K una curva elíptica dada por una ecuación de Weierstrass minimal y sea K'/K una extensión no ramificada. Entonces la ecuación de E sigue siendo minimal sobre K' y, en consecuencia, el tipo de reducción es el mismo sobre ambos cuerpos.*

DEMOSTRACIÓN: Supongamos que $\operatorname{car} k \neq 2, 3$. Entonces el teorema 6.4 nos da que la ecuación de E/K cumple $v(\Delta) < 12$ o $v(c_4) < 4$ (ver la observación tras el teorema). Como la extensión K'/K es no ramificada, la valoración de K' extiende a la de K, luego la ecuación también es minimal sobre K'.

Supongamos ahora que $\operatorname{car} k = 3$. Supongamos que existe un cambio de variables semientero (respecto de $\mathcal{O}'$) que transforma la ecuación minimal (sobre K) en otra ecuación entera con menor $\delta_{\mathfrak{P}}$. Diremos que tal cambio es *reductor* y vamos a probar que entonces existe un cambio de variables reductor definido sobre $\mathcal{O}$, lo que contradice la minimalidad sobre K de la ecuación de partida.

Como $\operatorname{car} k \neq 2$, podemos aplicar cambios de variables enteros para transformar tanto la ecuación minimal de partida como la ecuación reducida en ecuaciones de Weierstrass de tipo b. La composición del cambio reductor con estos dos cambios enteros sigue siendo un cambio reductor, luego podemos suponer que éste pasa de una ecuación de tipo b en otra de tipo b.

Si el cambio reductor está determinado por u, r, s, $t \in \mathcal{O}$, el teorema 2.6 muestra que el hecho de que sea reductor equivale a que $v(u) > 0$. Más aún, dicho teorema implica que podemos hacer $u = \pi$, pues con ello estamos multiplicando cada coeficiente a'_i por un entero, luego los a'_i siguen estando en $\mathcal{O}'$ y el cambio sigue cumpliendo $v(u) > 0$, luego sigue siendo reductor.

La ecuación de a_1 implica que $s = 0$ y la de a_3 que $t = 0$. El cambio se reduce, pues, a

$$X = \pi^2 X' + r, \qquad Y = \pi^3 Y', \qquad r \in \mathcal{O}'.$$

Sea $S \subset \mathcal{O}$ un conjunto de representantes de las clases de k, de modo que todo elemento de $\mathcal{O}$ se expresa de forma única como $\sum_{i\geq 0} s_i \pi^i$. Podemos extender S a un conjunto similar S' para $\mathcal{O}'$.

Vamos a probar que si $\rho \in \mathcal{O}'$ podemos sustituir r por $r + \rho\pi^2$ de modo que el cambio de variables sigue siendo reductor. Esto implica que podemos suponer que $r = r_0 + r_1\pi$, con r_0, $r_1 \in S'$.

En efecto, nos basamos en las ecuaciones

$$\begin{aligned}
\pi^2 b'_2 &= b_2 + 12r,\\
\pi^4 b'_4 &= b_4 + rb_2 + 6r^2,\\
\pi^6 b'_6 &= b_6 + 2rb_4 + r^2 b_2 + 4r^3.
\end{aligned}$$

Llamamos $g_2(r)$, $g_4(r)$, $g_6(r)$ a los miembros derechos. Hemos de probar que $\pi^i \mid g_i(r + \rho\pi^2)$, para $i = 2, 4, 6$, lo que justifica que los coeficientes de la ecuación transformada siguen siendo enteros. Ahora bien:

$$\begin{aligned}
g_2(r + \rho\pi^2) &= g_2(r) + 12\rho\pi^2 = (b'_2 + 12\rho)\pi^2,\\
g_4(r + \rho\pi^2) &= g_4(r) + \rho\pi^2(b_2 + 12r) + 6\rho^2\pi^4\\
&= (b'_4 + \rho b'_2 + 6\rho^2)\pi^4,\\
g_6(r + \rho\pi^2) &= (b'_6 + 2b'_4\rho + b'_2\rho^2 + 4\rho^3)\pi^6.
\end{aligned}$$

Ahora demostramos que r_0, $r_1 \in S$, con lo que $r \in \mathcal{O}$ y tenemos que el cambio de variables está definido sobre $\mathcal{O}$, que es la contradicción que buscábamos.

La ecuación de b_2 muestra que $v(b_2) \geq 1$, luego $b_2 = \beta_2\pi$, con $\beta_2 \in \mathcal{O}$, y ahora la ecuación de b_4 nos permite concluir que $b_4 = \beta_4\pi$ con $\beta_4 \in \mathcal{O}$. De aquí se sigue que

$$g_6(r) \equiv 4r_0^3 + b_6 \equiv 0 \ (\text{mód } \pi)$$

(la última congruencia porque $\pi^6 \mid g_6(r)$). Esto implica que $\tilde{r}_0$ es puramente inseparable sobre k, luego $\tilde{r}_0 \in k$ y por la unicidad de los desarrollos en serie

$r_0 \in S$. (Existe un $\alpha \in O$ tal que $\tilde{\alpha} = \tilde{r}_0$, luego r_0 es el primer coeficiente del desarrollo en serie de α.)

Factoricemos $3 = \kappa\pi$, con $\kappa \in \mathcal{O}$ (si $\operatorname{car} K = 3$ es $\kappa = 0$). Un simple cálculo nos da que

$$g_4(r) = b_4 + r_0 b_2 + 6r_0^2 + \pi^2(r_1\beta_2 + 4\kappa r_0 r_1 + 6r_1^2),$$

$$g_6(r) = (b_6 + 2r_0 b_4 + r_0^2 b_2 + 4r_0^3) + 2\pi r_1(b_4 + b_2 r_0 + 6r_0^2) + \pi^3(r_1^2\beta_2 + 4\kappa r_0 r_1^2 + 4r_1^3).$$

Sustituyendo la primera ecuación en la segunda obtenemos

$$\pi^6 b_6' = (b_6 + 2r_0 b_4 + r_0^2 b_2 + 4r_0^3) + 2r_1 b_4'\pi^5 - \pi^3(8r_1^3 + r_1^2(4\kappa r_0 + \beta_2)).$$

De aquí deducimos que el primer paréntesis del segundo miembro ha de ser de la forma $\beta\pi^3$, con $\beta \in \mathcal{O}$. Dividiendo entre π^3 pasamos a que

$$8r_1^3 + r_1^2(4\kappa r_0 + \beta_2) \equiv -\beta \pmod{\pi},$$

pero por otra parte $\pi^2 b_2' = (\beta_2 + 4\kappa r_0)\pi + 4\kappa r_1\pi^2$, de donde

$$4\kappa r_0 + \beta_2 \equiv 0 \pmod{\pi}.$$

En definitiva, llegamos a la congruencia $r_1^3 \equiv \beta \pmod{\pi}$, de la que se sigue que $\tilde{r}_1$ es puramente inseparable sobre k, luego está en k y $r_1 \in S$.

El caso $\operatorname{car} k = 2$ es análogo, sólo que requiere más cálculos. El planteamiento es el mismo, pero ahora no podemos suponer que las ecuaciones sean de tipo b y hemos de trabajar con las fórmulas para los a_i en lugar de los b_i. De todos modos, mediante un cambio entero con $u = 1$ y $s = t = 0$ podemos transformar la ecuación para exigir que $a_2 = 0$. Como antes, podemos suponer que $u = \pi$, si bien s y t no tienen por qué ser nulos.

Ahora demostramos que, para todo ρ, σ, $\tau \in \mathcal{O}'$, si cambiamos

$$r \mapsto r + \rho\pi^3, \quad s \mapsto s + \sigma\pi, \quad t \mapsto t + \tau\pi^3,$$

el nuevo cambio de variables sigue siendo reductor, lo que nos permite suponer que

$$r = r_0 + r_1\pi + r_2\pi^2, \quad s = s_0, \quad t = t_0 + t_1\pi + t_2\pi^2,$$

donde todas las variables con subíndices representan elementos de S'.

Ahora las ecuaciones son:

$$\begin{aligned}
\pi a_1' &= a_1 + 2s,\\
\pi^2 a_2' &= -sa_1 + 3r - s^2,\\
\pi^3 a_3' &= a_3 + ra_1 + 2t,\\
\pi^4 a_4' &= a_4 - sa_3 - (t + rs)a_1 + 3r^2 - 2st,\\
\pi^6 a_6' &= a_6 + ra_4 + r^3 - ta_3 - t^2 - rta_1.
\end{aligned}$$

Llamamos $f_i(r, s, t)$ a los miembros derechos. Las tres primeras ecuaciones implican que $a_1 = \alpha_1\pi$, $a_3 = \alpha_3\pi$, con $\alpha_i \in \mathcal{O}$, así como que $r \equiv s^2 \pmod{\pi}$. Expresamos $2 = \kappa\pi$, con $\kappa \in \mathcal{O}$.

Realizamos las sustituciones una a una. Primero cambiamos $s \mapsto s + \sigma\pi$. Hemos de probar que los coeficientes a'_i siguen siendo enteros o, lo que es lo mismo, que $\pi^i \mid f_i(r, s + \sigma\pi, t)$. Por ejemplo,

$$f_4(r, s + \sigma\pi, t) = f_4(r, s, t) - \sigma\pi^2(\alpha_3 + r\alpha_1 + \kappa t) = \pi^4 a'_4 - \sigma\pi^4 a'_3.$$

Ahora cambiamos $r \mapsto r + \rho\pi^3$. Detallamos por ejemplo el caso $i = 6$:

$$\begin{aligned} f_6(r + \rho\pi^3) &= \pi^6 a'_6 + \rho\pi^3 a_4 + 3r^2\rho\pi^3 + 3r\rho^2\pi^6 + \rho^3\pi^9 - ta_1\rho\pi^3 \\ &\equiv \rho\pi^3(a_4 + 3r^2 - ta_1) \\ &\equiv \rho\pi^3(\pi^4 a'_4 + s(a_3 + ra_1 + 2t)) \\ &\equiv \rho\pi^3 s\pi^3 a'_3 \equiv 0 \ (\text{mód } \pi^6). \end{aligned}$$

El cambio $t \mapsto t + \tau\pi^3$ se trata similarmente.

A continuación demostramos que r_0, r_1, s_0, t_0, $t_1 \in S$. La ecuación de a_4 nos da que $3r^2 + a_4 \equiv 0$ (mód π), luego $r_0^2 + a_4 \equiv 0$ (mód π). Esto implica que $\tilde{r}_0$ es puramente inseparable sobre k, luego $r_0 \in S$.

La ecuación de a_6 implica que $a_6 + ra_4 + r^3 - t^2 \equiv 0$ (mód π), con lo que $a_6 + r_0 a_4 + r_0^3 - t_0^2 \equiv 0$ (mód π). Esto prueba que $\tilde{t}_0$ es puramente inseparable sobre k, luego $t_0 \in S$. De la ecuación de a_2 obtenemos que $s_0 \in S$.

Ahora reducimos la ecuación de a_2 módulo π^2:

$$-s\alpha_1\pi + 3r_1\pi + (3r_0 - s_0) \equiv 0 \ (\text{mód } \pi^2).$$

La congruencia $r \equiv s^2$ (mód π) implica $3r_0 - s_0^2 \equiv 0$ (mód π). Pongamos que $3r_0 - s_0^2 = \lambda\pi$, con $\lambda \in \mathcal{O}$. Entonces

$$-s\alpha_1 + r_1 + \lambda \equiv 0 \ (\text{mód } \pi).$$

Esto implica que $\tilde{r}_1 \in k$, luego $r_1 \in S$. Por último reducimos la ecuación de a_6 módulo π^3:

$$\pi^2(r_2(a_4 + 3\rho^2) - t_1(\alpha_3 + \kappa t_0 + \rho\alpha_1) - t_1^2) \equiv \gamma \ (\text{mód } \pi^3),$$

donde $\rho = r_0 + r_1\pi$ y $\gamma \in \mathcal{O}$. Esto implica que $\gamma = \pi^2\gamma_0$, con $\gamma_0 \in \mathcal{O}$, con lo que

$$r_2(a_4 + 3\rho^2) - t_1(\alpha_3 + \kappa t_0 + \rho\alpha_1) - t_1^2 \equiv \gamma_0 \ (\text{mód } \pi).$$

La ecuación de a_4 implica que el primer paréntesis es nulo módulo π, y la de a_3 implica lo mismo para el segundo. Así pues, $t_1^2 + \gamma_0 \equiv 0$ (mód π) y concluimos como siempre que $t_1 \in S$.

Por el contrario, no podemos asegurar que r_2 y t_2 estén en S. Veamos lo máximo que podemos decir. En la ecuación de a_6 sustituimos $r = \rho + r_2\pi^2$, $t = \tau + t_2\pi^2$ (con ρ, $\tau \in \mathcal{O}$), tomamos congruencias módulo π^6 y agrupamos en un único término δ todos los sumandos que están en $\mathcal{O}$:

$$\delta \equiv r_2\pi^2(a_4 - ta_1 + s\rho r_2\pi^2 + 3\rho^2) - t_2\pi^3(\rho\alpha_1 + \alpha_3 + \kappa\tau) - t_2^2\pi^4 \ (\text{mód } \pi^6).$$

Sustituimos a_4 por la fórmula de a_4 y $\alpha_3 + \kappa\tau$ por la de a_3:

$$\delta \equiv r_2\pi^2(sa_3 + 2st + rsa_1 - 3r^2 + 3\rho r_2\pi^2 + 3\rho^2) - t_2\pi^3(\rho\alpha_1 + \pi^2 a_3' - r\alpha_1) - t_2^2\pi^4.$$

Volvemos a sustituir $r = \rho + r_2\pi^2$:

$$\delta \equiv r_2\pi^2(sa_3 + 2st + rsa_1 - 3\rho r_2\pi^2) - t_2\pi^3(\pi^2 a_3' - r_2\alpha_1\pi^2) - t_2^2\pi^4 \text{ (mód } \pi^6).$$

Equivalentemente:

$$\delta \equiv r_2\pi^2 s(a_3 + 2t + ra_1) - 3\rho r_2^2\pi^4 - t_2\pi^5(a_4' - r_2\alpha_1) - t_2^2\pi^4 \text{ (mód } \pi^6).$$

Aplicamos de nuevo la ecuación de a_3:

$$\delta \equiv r_2 sa_3'\pi^5 - 3\rho r_2^2\pi^4 - t_2\pi^5(a_4' - r_2\alpha_1) - t_2^2\pi^4 \text{ (mód } \pi^6).$$

De aquí se sigue que $\delta = \epsilon\pi^4$, con $\epsilon \in \mathcal{O}$. Dividiendo entre π^4 queda

$$\epsilon \equiv -3\rho r_2^2 - t_2^2 \equiv r_0 r_2^2 + t_2^2 \text{ (mód } \pi).$$

Ahora usamos que $r \equiv s^2$ (mód π), de donde $r_0 \equiv s_0^2$ (mód π):

$$\epsilon \equiv (s_0 r_2 + t_2)^2 \text{ (mód } \pi).$$

El argumento usual de inseparabilidad nos da que $\tilde{s}_0\tilde{r}_2 + \tilde{t}_2 \in k$. En otros términos, existe un $u \in \mathcal{O}$ tal que $s_0 r_2 + t_2 \equiv u$ (mód π). Más explícitamente, digamos que $t_2 + s_0 r_2 = u + v\pi$, con $v \in \mathcal{O}'$.

El teorema quedará probado si justificamos que podemos cambiar

$$r \mapsto r - r_2\pi^2 = r_0 + r_1\pi, \qquad t \mapsto t + s_0 r_2\pi^2 = t_0 + t_1\pi + u\pi^2 + v\pi^3$$

de modo que el cambio de variables sigue siendo reductor, pues en tal caso sabemos que el término $v\pi^3$ también puede eliminarse, con lo que llegamos finalmente a un cambio definido sobre $\mathcal{O}$.

En la práctica hemos de probar que $\pi^i \mid f_i(r - r_2\pi^2, s_0, t + s_0 r_2\pi^2)$, para todo índice i. Veámoslo por ejemplo para $i = 4$.

$$\begin{aligned} f_4(r - r_2\pi^2, s_0, t + s_0 r_2\pi^2) &= f_4(r, s_0, t) - 6rr_2\pi^2 + 3r_2^2\pi^4 - 2s_0^2 r_2\pi^2 \\ &\equiv -\kappa r_2\pi^3(3r + s_0^2) \equiv -\kappa r_2\pi^3 2s_0^2 \\ &\equiv -\kappa^2 r_2 s_0^2\pi^4 \equiv 0 \text{ (mód } \pi^4). \end{aligned}$$

(Aquí hemos usado que $3r \equiv r_0$ (mód π).) ∎

6.3 Puntos enteros y puntos de torsión

En esta sección extraeremos numerosas consecuencias del teorema siguiente, que es el motivo que nos ha llevado a estudiar los grupos formales en el capítulo anterior:

Teorema 6.16 *Sea E/K una curva elíptica dada por una ecuación de Weierstrass entera, sea G el grupo formal de E (que por 5.3 es un grupo formal sobre el anillo $\mathcal{O}$) y sea $z(T)$ la serie de Taylor de $z = -1/y$ en O respecto del parámetro $t = -x/y$ (que por 5.2 está en $\mathcal{O}[[T]]$). Entonces la aplicación $G(\mathfrak{P}) \longrightarrow E_1(K)$ dada por $t \mapsto (t/z(t), -1/z(t))$ (entendiendo que la imagen de 0 es O) es un isomorfismo de grupos.*

Demostración: Observemos en primer lugar que

$$E_1(K) = \{P \in E(K) \mid v(t(P)) \geq 1,\ v(z(P)) \geq 1\}.$$

En efecto, si $\tilde{P} = [0,1,0]$, entonces $P = [a,b,c]$, donde $v(a) \geq 1$, $v(b) = 0$, $v(c) \geq 1$, y entonces $t(P) = -a/b$, $z(P) = -c/b$ son enteros no unitarios. El recíproco se prueba igualmente.

Llamemos $\phi : G(\mathfrak{P}) \longrightarrow \mathrm{P}^2(K)$ a la aplicación dada por

$$\phi(t) = [t, -1, z(t)].$$

La definición es correcta, pues la serie $z(T)$ tiene coeficientes enteros, y esto implica que converge en los puntos de $G(\mathfrak{P})$.

Sabemos que $z(T) = T^3(1+\cdots)$, de donde se sigue que $z(T)$ sólo se anula en $t = 0$. Por lo tanto, en términos de las coordenadas de Weierstrass tenemos que $\phi(0) = O$ y, para los demás puntos, ϕ es la aplicación descrita en el enunciado.

La serie $z(T)$ cumple la ecuación

$$z = T^3 + (a_1T + a_2T^2)z + (a_3 + a_4T)z^2 + a_6z^3,$$

donde los coeficientes a_i son los de la ecuación de Weierstrass minimal que estamos considerando. Se aquí se sigue que si $t \in G(\mathfrak{P})$, entonces el punto $\phi(t)$ cumple la ecuación de Weierstrass, luego $\phi : G(\mathfrak{P}) \longrightarrow E(K)$. Más aún, como $z(t) \in \mathfrak{P}^3$, es obvio que $\phi(t)$ se reduce al punto O, luego $\phi : G(\mathfrak{P}) \longrightarrow E_1(K)$.

Vamos a ver que ϕ tiene por inversa a la aplicación $t : E_1(K) \longrightarrow G(\mathfrak{P})$. Es inmediato que $t(\phi(t_0)) = t_0$, para todo $t_0 \in G(\mathfrak{P})$. Tomemos ahora un punto $P \in E_1(K)$ y hemos de probar que $\phi(t(P)) = P$. Podemos suponer que $P \neq O$, y entonces $P = (x(P), y(P)) = (t(P)/z(P), -1/z(P))$, luego basta demostrar que $z(t(P)) = z(P)$. (Notemos que la z del miembro izquierdo es la serie $z(T)$, mientras que la del miembro derecho es la función $z \in K(E)$).

Si llamamos $t_0 = t(P) \in \mathfrak{P}$, los dos números $z(t_0)$ y $z(P)$ están en $\mathfrak{P}$ y cumplen la ecuación

$$z = t_0^3 + (a_1t_0 + a_2t_0^2)z + (a_3 + a_4t_0)z^2 + a_6z^3.$$

Por consiguiente, la diferencia $z(t_0) - z(P)$ es

$$(a_1t_0 + a_2t_0^2)(z(t_0) - z(P)) + (a_3 + a_4t_0)(z(t_0)^2 - z(P)^2) + a_6(z(t_0)^3 - z(P)^3).$$

Si fuera $z(t_0) \neq z(P)$ podríamos dividir:

$$1 = a_1t_0 + a_2t_0^2 + (a_3 + a_4t_0)(z(t_0) + z(P)) + a_6(z(t_0)^2 + z(t_0)z(P) + z(P)^2),$$

y de aquí se concluye que $1 \in \mathfrak{P}$, lo cual es absurdo.

Con esto tenemos probado que ϕ y t son mutuamente inversas. Vamos a probar, por último, que t es un homomorfismo de grupos. Sea $s : E \times E \longrightarrow K$ la composición de la suma en E con la función t. Se trata de una función racional en $E \times E$, luego $s = R(t_1, z_1, t_2, z_2)$, para cierta $R \in K[T_1, Z_1, T_2, Z_2]$. Entonces su serie de Taylor en (O, O) respecto de los parámetros t_1, t_2 es claramente $s(T_1, T_2) = R(T_1, z(T_1), T_2, z(T_2))$. Por lo tanto, si P, $P' \in E_1(K)$, se cumple

$$\begin{aligned} t(P + P') &= R(t(P), z(P), t(P'), z(P')) = R(t(P), z(t(P)), t(P'), z(t(P'))) \\ &= s(t(P), t(P')) = t(P) + t(P'). \end{aligned}$$

Esto demuestra que t es un isomorfismo de grupos y, por consiguiente, ϕ también lo es. ∎

El isomorfismo dado por el teorema anterior nos permite definir una cadena de subgrupos de $E_1(K)$, correspondientes a los subgrupos $G(\mathfrak{P}^n)$. Vamos a ver que éstos tienen una interpretación natural.

Definición 6.17 Sea E/K una curva elíptica definida por una ecuación de Weierstrass entera $F(X, Y, Z) = 0$, donde

$$F(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

Definimos como sigue una aplicación $v : E(K) \longrightarrow \mathbb{N} \cup \{+\infty\}$ (que no hay que confundir con la valoración en K, pese a que usamos la misma notación). Para cada punto $P \in E(K)$, fijamos unas coordenadas homogéneas (x, y, z) que sean enteras y al menos una de ellas unitaria. (Así, $v(x)$, $v(y)$, $v(z) \geq 0$ no dependen de la elección de la terna.)

- Si $v(z) = 0$ diremos que el punto P es *entero* y definimos $v(P) = 0$.
- Si $v(z) > 0$, despejando x^3 en la ecuación vemos que $v(x) > 0$, luego por la elección de la terna ha de ser $v(y) = 0$ y al despejar x^3 el término y^2z tiene valor $\mathfrak{P}$-ádico estrictamente menor que cualquier otro de su miembro, luego $3v(x) = v(z)$. En este caso definimos $v(P) = v(x)$.

Para comprender el significado de v observamos en primer lugar que

$$v(P) = +\infty \quad \text{si y sólo si} \quad P = O.$$

En efecto, $O = [0, 1, 0]$ está en el caso $v(z) = +\infty > 0$ y entonces tenemos que $v(O) = v(x) = v(0) = +\infty$. Recíprocamente, si $v(P) = +\infty$ es porque $v(z) > 0$ y $v(x) = +\infty$. Más aún, $v(z) = 3v(x) = +\infty$, luego $P = [0, y, 0] = O$.

Para un punto finito $P = (x, y) = [x, y, 1]$, tenemos que

$$v(P) = 0 \quad \text{si y sólo si} \quad x,\ y \in \mathcal{O}.$$

En otras palabras, los puntos enteros son los que tienen coordenadas afines enteras. Finalmente, si $v(P) > 0$, tenemos que $P = (x, y) = [x', y', z']$, donde $x = x'/z'$, $y = y'/z'$ y

$$v(x) = v(x') - v(z') = -2v(P),$$

$$v(y) = v(y') - v(z') = -3v(P).$$

Podemos decir, pues, que $v(P)$ mide lo que dista P de tener coordenadas afines enteras. Notemos ahora que de la propia definición de $E_1(K)$ se sigue que

$$E_1(K) = \{P \in E(K) \mid v(P) \geq 1\}.$$

Para cada natural $m \geq 1$ definimos

$$E_m(K) = \{P \in E(K) \mid v(P) \geq m\}.$$

No es evidente en absoluto que los conjuntos $E_m(K)$ sean subgrupos de $E(K)$, pero lo cierto es que lo son, como se deduce del teorema siguiente:

Teorema 6.18 *Sea E/K una curva elíptica dada por una ecuación de Weierstrass entera, sea G su grupo formal y sea $\phi : G(\mathfrak{P}) \longrightarrow E_1(K)$ el isomorfismo dado por el teorema* 6.16. *Entonces, para todo $t \in G(\mathfrak{P})$ se cumple que $v(t) = v(\phi(t))$.*

Demostración: Basta observar que

$$v(\phi(t)) = -\frac{1}{3}v(-1/z(t)) = \frac{1}{3}v(z(t)) = v(t).$$

■

Esto prueba que, en efecto los conjuntos $E_m(K)$ son las imágenes por el isomorfismo de los subgrupos $G(\mathfrak{P}^m) = \mathfrak{P}^m = \{t \in K \mid v(t) \geq m\}$, luego tenemos una cadena de subgrupos

$$E(K) \geq E_0(K) \geq E_1(K) \geq \cdots \geq E_m(K) \geq \cdots \geq \bigcap_{m=1}^{\infty} E_m(K) = 0.$$

Otra consecuencia inmediata es que si P, $Q \in E(K)$, entonces

$$v(P + Q) \geq \text{mín}\{v(P), v(Q)\},$$

y se da la igualdad si $v(P) \neq v(Q)$. (Observemos que esto es trivial si alguno de los puntos es entero o infinito, y en los casos restantes es consecuencia del teorema anterior.)

Otra observación sobre los subgrupos $E_m(K)$ es que el cociente de dos consecutivos es isomorfo a k^+ o, en otros términos, que tenemos sucesiones exactas (para $m \geq 1$):

$$0 \longrightarrow E_{m+1}(K) \longrightarrow E_m(K) \longrightarrow k^+.$$

En efecto (ver las observaciones de la página 5.3):

$$E_m(K)/E_{m+1}(K) \cong G(\mathfrak{P}^m)/G(\mathfrak{P}^{m+1}) \cong \mathfrak{P}^m/\mathfrak{P}^{m+1} \cong k^+.$$

El último isomorfismo es el inducido por el epimorfismo $\mathfrak{P}^m \longrightarrow k^+$ dado por $\alpha \mapsto [\alpha\pi^{-m}]$.

Para $m = 0$ también tenemos una sucesión exacta, pero su último grupo no es k^+, sino $\tilde{E}_r(k)$, tal y como hemos demostrado en 6.8.

Finalmente, el teorema 5.20 nos da que si $\operatorname{car} K = 0$ entonces, para todo natural m suficientemente grande se cumple que

$$E_m(K) \cong G(\mathfrak{P}^m) \cong \mathfrak{P}^m \cong \mathfrak{O}^+.$$

El último isomorfismo es el dado por $\alpha \mapsto \alpha\pi^{-m}$. De hecho, si $p = \operatorname{car} k$ y $v(p) = 1$, tenemos el isomorfismo para $m = 2$, e incluso para $m = 1$ si $p > 2$.

La consecuencia más importante que vamos a extraer del teorema 6.16 es la siguiente:

Teorema 6.19 *Sea E/K una curva elíptica dada por una ecuación de Weierstrass entera y $m \geq 2$ un número natural primo con* $\operatorname{car} k$. *Entonces $E_1(K)$ no tiene elementos de orden m y, si $\tilde{E}/k$ es regular, entonces la reducción $E(K)[m] \longrightarrow \tilde{E}(k)$ es inyectiva.*

Demostración: Sea G el grupo formal de E. Por el teorema 5.16 sabemos que $G(\mathfrak{P})$ no tiene elementos de orden m. Por 6.16 lo mismo vale para $E_1(K)$. Si $\tilde{E}/k$ es regular entonces $E_0(K) = E(K)$ y $\tilde{E}_r(k) = \tilde{E}(k)$. La restricción a $E(K)[m]$ de la reducción módulo $\mathfrak{P}$ tiene núcleo trivial, luego es inyectiva. ∎

Este teorema permite en muchos casos determinar los elementos de torsión de una curva elíptica definida sobre $\mathbb{Q}$ o, más en general, sobre un cuerpo numérico.

Ejemplo Consideremos la curva elíptica $E/\mathbb{Q}$ dada por

$$Y^2 + Y = X^3 - X + 1.$$

Su discriminante es $\Delta = -13 \cdot 47$. Esta ecuación define también una curva elíptica sobre $\mathbb{Q}_p$ con el mismo determinante, de modo que $E(\mathbb{Q})$ es un subgrupo de $E(\mathbb{Q}_p)$. Además $v_2(\Delta) = v_3(\Delta) = 0$, luego la ecuación es minimal para $\mathbb{Q}_2$ y $\mathbb{Q}_3$ y la reducción es regular. Una simple comprobación muestra que $\tilde{E}(\mathbb{Z}/2\mathbb{Z}) = \{\tilde{O}\}$ y $\tilde{E}(\mathbb{Z}/3\mathbb{Z}) = \{\tilde{O}\}$ (la ecuación no tiene soluciones en $\mathbb{Z}/2\mathbb{Z}$ ni en $\mathbb{Z}/3\mathbb{Z}$), luego el teorema anterior implica que $E(\mathbb{Q})[m] = 0$ para todo m primo con 2 o con 3. En particular, $E(\mathbb{Q})$ no tiene elementos de torsión de orden primo, luego no tiene elementos de torsión en absoluto. En otras palabras, $E(\mathbb{Q}) = \{O\}$. ∎

Ejemplo Consideremos ahora la curva $E/\mathbb{Q}$ dada por $Y^2 = X^3 + 3$, cuyo discriminante es $\Delta = -2^4 \cdot 3^5$. Esta misma ecuación determina curvas elípticas sobre $\mathbb{Q}_p$ para todo primo p y como $v_p(\Delta) < 12$ es minimal. Además la reducción módulo p es regular si $p \geq 5$.

Se comprueba fácilmente que $|\tilde{E}(\mathbb{Z}/5\mathbb{Z})| = 6$, $|\tilde{E}(\mathbb{Z}/7\mathbb{Z})| = 13$. Si p es un primo distinto de 5 y de 7, el teorema anterior implica que $E(\mathbb{Q})[p]$ puede verse como subgrupo de $\tilde{E}(\mathbb{Z}/5\mathbb{Z})$ y $\tilde{E}(\mathbb{Z}/7\mathbb{Z})$, lo que obliga a que $E(\mathbb{Q})[p] = 0$. Por otra parte, $E(\mathbb{Q})[5]$ es subgrupo de $\tilde{E}(\mathbb{Z}/7\mathbb{Z})$, pero este último grupo no puede tener elementos de orden 5, luego $E(\mathbb{Q})$ tampoco tiene elementos de orden 5. Igualmente se concluye que no hay elementos de orden 7. Concluimos que $E(\mathbb{Q})$ no tiene elementos de torsión.

Puesto que $(1, 2) \in E(\mathbb{Q})$, este punto ha de tener orden infinito, luego podemos afirmar que la ecuación $Y^2 = X^3 + 3$ tiene infinitas soluciones en $\mathbb{Q}^2$, lo cual no es evidente en absoluto. ∎

Ejemplo Consideremos ahora la curva $E/\mathbb{Q}$ dada por $Y^2 = X^3 + X$, cuyo discriminante es $\Delta = -2^6$. Es fácil ver que $(0, 0) \in E(\mathbb{Q})[2]$. Vamos a ver que es el único elemento de torsión de $E(\mathbb{Q})$ (aparte de O).

Es fácil comprobar que los grupos $\tilde{E}(\mathbb{Z}/3\mathbb{Z})$ y $\tilde{E}(\mathbb{Z}/5\mathbb{Z})$ tienen orden 4. Esto implica que $E(\mathbb{Q})$ todo elemento de torsión de $E(\mathbb{Q})$ no trivial ha de tener orden 2 o 4. Más concretamente, vemos que

$$\tilde{E}(\mathbb{Z}/3\mathbb{Z}) = \{O, (0,0), (2,1), (2,2)\}, \qquad \tilde{E}(\mathbb{Z}/5\mathbb{Z}) = \{O, (0,0), (2,0), (3,0)\}.$$

En ambas curvas, los elementos de orden 2 son los que cumplen $Y = 0$, luego

$$\tilde{E}(\mathbb{Z}/3\mathbb{Z}) \cong \mathbb{Z}/4\mathbb{Z}, \qquad \tilde{E}(\mathbb{Z}/5\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

El segundo grupo nos muestra que $E(\mathbb{Q})$ no tiene elementos de orden 4 y el primero que a lo sumo hay un elemento de orden 2. Así pues, $(0, 0)$ es el único elemento de torsión (no nulo) de $E(\mathbb{Q})$. ∎

Ejercicio: Probar que si K es un cuerpo numérico y E/K es una curva elíptica, entonces $E(K)$ tiene un número finito de elementos de torsión.

Los puntos de orden finito están muy cerca de tener coordenadas enteras. El teorema siguiente describe la situación general:

Teorema 6.20 *Supongamos que K tiene característica 0 y que k tiene característica prima p. Sea E/K una curva elíptica determinada por una ecuación de Weierstrass entera y sea $P \in E(K)$ un punto de orden $m \geq 2$.*

a) Si m no es potencia de p, entonces $v(P) = 0$, es decir, $x(P)$, $y(P) \in \mathcal{O}$.

b) Si $m = p^n$, entonces $v(P) \leq v(p)/(p^n - p^{n-1})$.

DEMOSTRACIÓN: a) Sea G el grupo formal de E. El teorema 5.16 implica que $G(\mathfrak{P})$ no tiene elementos de torsión de orden m, luego $P \notin E_1(K)$, es decir, $v(P) = 0$.

b) O bien $v(P) = 0$ o bien $P \in E_1(K)$, en cuyo caso el teorema 5.17 nos da la desigualdad del enunciado. ∎

En las condiciones del teorema anterior, si $v(p) = 1$ el apartado b) implica también $v(P) = 0$ salvo si $p = 2$ y $n = 1$, luego concluimos que todos los puntos de torsión de $E(K)$ tienen coordenadas enteras salvo a lo sumo los de orden 2 (cuando $\operatorname{car} k = 2$), que pueden cumplir $v(P) = 1$, con lo que serán de la forma $(a/4, b/8)$, con a, $b \in \mathcal{O}$.

Ejemplo El punto $(-1/4, 1/8)$ tiene orden 2 en la curva $E/\mathbb{Q}_2$ dada por la ecuación $Y^2 + XY = X^3 + 4X + 1$. ∎

Por último probamos que curvas isógenas tienen reducciones isógenas. Conviene introducir la notación siguiente:

Definición 6.21 Si E/K y E'/K son dos curvas elípticas definidas sobre K, representaremos por $\operatorname{Hom}_K(E, E')$ al grupo de las isogenias de E en E' definidas sobre K. Igualmente, $\operatorname{End}_K(E)$ será el anillo de las isogenias de E en sí misma definidas sobre K.

Teorema 6.22 *Sean E/K y E'/K dos curvas elípticas con buena reducción módulo $\mathfrak{P}$. Entonces la reducción de isogenias dada por el teorema* 6.11 *es un monomorfismo de grupos $\operatorname{Hom}_K(E, E') \longrightarrow \operatorname{Hom}_k(E, E')$ (y también de anillos cuando $E = E'$).*

DEMOSTRACIÓN: Es inmediato que la reducción es un homomorfismo. Sólo hemos de probar que es inyectiva. Ahora bien, si $\phi \in \operatorname{Hom}_K(E, E')$ cumple $\tilde{\phi} = 0$, fijamos un primo $p \neq \operatorname{car} k$, y observamos que si $P \in E[p^n]$, entonces $P \in E(L)[p^n]$, para cierta extensión finita L de K, luego $\phi(P) \in E'(L)[p^n]$ y

$$\widetilde{\phi(P)} = \tilde{\phi}(\tilde{P}) = O.$$

Por el teorema 6.19 podemos concluir que $\phi(P) = O$, luego $E[p^n]$ está en el núcleo de ϕ para todo n, luego dicho núcleo es infinito y $\phi = 0$. ∎

6.4 La topología métrica

En esta sección mostraremos que la métrica de K induce una topología sobre las variedades algebraicas definidas sobre K exactamente igual que en el caso complejo. Empezamos definiendo la topología métrica sobre el espacio proyectivo $\mathrm{P}^n(K)$. Para ello definimos

$$A_i = \{P \in \mathrm{P}^n(K) \mid x_i(P) \neq 0\}$$

y llamamos $p_i : A_i \longrightarrow K^n$ a la aplicación que a cada $P \in A_i$ le asigna la n-tupla resultante de eliminar x_i en la única $n+1$-tupla de coordenadas homogéneas de P que cumple $x_i = 1$. Se comprueba inmediatamente que las composiciones $p_i^{-1} \circ p_j$ son homeomorfismos de K^n en sí mismo (dotado de la topología producto), lo cual permite definir una única topología en $\mathrm{P}^n(K)$ respecto a la cual los conjuntos A_i son abiertos y las aplicaciones p_i son homeomorfismos.

Definición 6.23 Llamaremos *topología métrica* en $\mathrm{P}^n(K)$ a la única topología respecto a la cual los conjuntos A_i son abiertos y las funciones $p_i : A_i \longrightarrow K^n$ son homeomorfismos.

Notemos que todo el razonamiento precedente es válido si K es un cuerpo métrico arbitrario, no necesariamente discreto. Es obvio que si L es una extensión algebraica de K considerada como cuerpo métrico con la única extensión posible del valor absoluto de K, entonces la topología inducida en $\mathrm{P}^n(L)$ extiende a la de $\mathrm{P}^n(K)$. En particular esto se aplica cuando L es la clausura algebraica $\overline{K}$ de K.

Si $V \subset \mathrm{P}^n(\overline{K})$ es una variedad proyectiva, llamaremos *topología métrica* en V a la restricción de la topología métrica de $\mathrm{P}^n(\overline{K})$. En particular, si V está definida sobre K, la restricción a $V(K)$ de la topología métrica de V coincide con la restricción a $V(K)$ de la topología métrica de $\mathrm{P}^n(K)$.

Si $F \in \overline{K}[X_1, \ldots, X_{n+1}]$ es una forma, entonces el conjunto

$$V(F) = \{P \in \mathrm{P}^n(\overline{K}) \mid F(P) = 0\}$$

es cerrado para la topología métrica, pues $V(F) \cap A_{n+1}$ se corresponde a través de la aplicación p_{n+1} con el cerrado $\{X \in \overline{K}^n \mid F(X, 1) = 0\}$, y lo mismo vale para los demás abiertos A_i.

Como consecuencia, todo subconjunto algebraico de $\mathrm{P}^n(\overline{K})$ es cerrado para la topología métrica o, equivalentemente, todo abierto para la topología de Zariski es abierto para la topología métrica (en $\mathrm{P}^n(\overline{K})$ y, por consiguiente, en cualquier variedad proyectiva).

Se comprueba inmediatamente que la aplicación $p : \overline{K}_{\mathfrak{p}}^{n+1} \setminus \{0\} \longrightarrow \mathrm{P}^n(\overline{K})$ que a cada $n+1$-tupla le asigna el punto con tales coordenadas homogéneas es continua. De aquí se sigue a su vez que toda aplicación regular $\phi : V \longrightarrow W$ entre variedades proyectivas en $\overline{K}$ es continua respecto de la topología métrica.

En efecto, es claro que no perdemos generalidad si consideramos

$$\phi : V \longrightarrow \mathrm{P}^m(\overline{K}).$$

En un entorno de P para la topología de Zariski, la aplicación ϕ viene dada por $\phi(Q) = [F_1(Q), \ldots, F_{m+1}(Q)]$, donde los $F_i \in \overline{K}[X_1, \ldots, X_{m+1}]$ son formas del mismo grado que no se anulan simultáneamente en P. Supongamos sin pérdida de generalidad que $x_1(P) \neq 0$ y $F_{m+1}(P) \neq 0$. Entonces, el conjunto

$$U = \{Q \in V(\overline{K}) \mid x_1(Q) \neq 0,\ F_{m+1}(Q) \neq 0\}$$

es un entorno de P para la topología de Zariski, luego también para la topología métrica. Sobre este entorno U, la aplicación ϕ puede obtenerse como composición de tres funciones continuas:

a) La aplicación $U \longrightarrow \overline{K}^n$ dada por $[X_1, \ldots, X_n] \mapsto (X_2/X_1, \ldots, X_n/X_1)$ (que es la restricción del homeomorfismo $A_1 \longrightarrow \overline{K}^n$).

b) La aplicación polinómica $X \mapsto (F_1(1, X), \ldots, F_{m+1}(1, X))$,

c) La aplicación $p : \overline{K}^{m+1} \setminus \{0\} \longrightarrow \mathrm{P}^m(\overline{K})$.

Ahora es fácil ver que la topología métrica en un producto $V \times W$ es el producto de las topologías métricas. En efecto, basta demostrarlo para el caso de $\mathrm{P}^n(K) \times \mathrm{P}^m(K)$, y a su vez basta probar que la topología métrica en un abierto $A_i \times A_j$ es la topología producto. Ahora bien, la aplicación natural $A_i \times A_j \longrightarrow K^{n+m}$ es un homeomorfismo para la topología métrica (porque es regular con inversa regular) y también para la topología producto.

Por último, observemos que si K es localmente compacto y V/K es una variedad proyectiva definida sobre K, entonces el conjunto de puntos racionales $V(K)$ es compacto respecto a la topología métrica. Como $V(K)$ es cerrado en un espacio proyectivo $\mathrm{P}^n(K)$, basta probar que éste es compacto, pero ello es evidente, ya que es espacio $\mathrm{P}^n(K)$ es la imagen por la aplicación continua $p : K^{n+1} \setminus \{0\} \longrightarrow \mathrm{P}^n(K)$ del compacto

$$C = \{x \in K^{n+1} \mid \sum_i |x_i| = 1\}.$$

Como primera aplicación demostramos el teorema siguiente:

Teorema 6.24 *Si K es localmente compacto y E/K es una curva elíptica, entonces el núcleo $E_1(K)$ de la reducción módulo $\mathfrak{P}$ tiene índice finito en $E(K)$.*

Demostración: En efecto, observemos que $E(K)$ es un grupo topológico compacto respecto de la topología métrica (es decir, la suma y la aplicación $P \mapsto -P$ son continuas, porque son regulares). En la prueba de 6.16 hemos visto que

$$E_1(K) = \{P \in E(K) \mid v(t(P)) \geq 1,\ v(z(P)) \geq 1\}.$$

Si llamamos U al abierto en $E(K)$ donde están definidas t y z, tenemos que $E_1(K) \subset U$ y las funciones $t, z : U \longrightarrow K$ son continuas, al igual que lo es la valoración $v : K \longrightarrow \mathbb{Z} \cup \{+\infty\}$. De aquí concluimos que $E_1(K)$ es abierto en $E(K)$.

Aunque no nos va a hacer falta, notemos que todo subgrupo abierto de un grupo topológico es también cerrado, ya que su complementario es una unión de trasladados, que también son abiertos. Así pues, $E_1(K)$ es abierto y cerrado en $E(K)$.

La conclusión del teorema es ahora inmediata, pues el compacto $E(K)$ se descompone en unión disjunta de las clases (abiertas) módulo $E_1(K)$, luego ha de haber un número finito de clases. ∎

La compacidad local de K equivale a que el cuerpo de restos k sea finito. Notemos que en tal caso el índice $|E_0(K) : E_1(K)| = |E_r(k)|$ es claramente finito, luego el teorema anterior equivale a la finitud de $|E(K) : E_0(K)|$. Cuando k no es finito el índice $|E(K) : E_1(K)|$ ya no tiene por qué ser finito, pero $|E(K) : E_0(K)|$ lo es igualmente. De momento no estamos en condiciones de probarlo.

6.5 El criterio de Néron-Ogg-Shafarevich

Como es habitual, llamaremos $\overline{K}$ a una clausura algebraica de K y K_{nr} será la máxima extensión no ramificada de K. La valoración de K se extiende de forma única a cada extensión finita no ramificada, luego también a K_{nr}. Así pues, K_{nr} es también un cuerpo métrico discreto. Su cuerpo de restos $\overline{k}$ es una clausura algebraica de k.

Definimos el *subgrupo de inercia* de $G(\overline{K}/K)$ como $I = G(\overline{K}/K_{\mathrm{nr}})$, de modo que $G(K_{\mathrm{nr}}/K) \cong G(\overline{K}/K)/I$. Cada elemento de $G(\overline{K}/K)$ induce un automorfismo de $\overline{k}/k$, de modo que I es el núcleo del epimorfismo $G(\overline{K}/K) \longrightarrow G(\overline{k}/k)$. En otras palabras, I está formado por los automorfismos que actúan trivialmente sobre $\overline{k}$.

Si E/K es una curva elíptica, sabemos que $G(\overline{K}/K)$ actúa también sobre los grupos de torsión $E[m]$ y sobre los módulos de Tate $T_l(E)$. El resultado principal de esta sección es una caracterización de la buena reducción de E en términos de la restricción a I de estas acciones.

Aceptaremos sin demostración un hecho que hemos comentado en la sección anterior: Si E/K es una curva elíptica entonces el índice $|E(K) : E_0(K)|$ es finito (aunque K no sea localmente compacto).

Teorema 6.25 (Criterio de Néron–Ogg–Shafarevich) *Consideremos una curva elíptica E/K y sea $p = \operatorname{car} k$. Las condiciones siguientes son equivalentes:*

a) E tiene buena reducción sobre K.

b) El grupo de inercia I actúa trivialmente sobre $E[m]$, para todo $m \geq 2$ primo con p.

c) El grupo de inercia I actúa trivialmente sobre el módulo de Tate $T_l(E)$ para algún primo $l \neq p$ (o para todo primo $l \neq p$).

d) El grupo de inercia I actúa trivialmente sobre $E[m]$, para infinitos valores de m primos con p.

Demostración: Veamos que a) $\Rightarrow$ b). Sea K' una extensión finita de K tal que $E[m] \subset E(K')$. Si E tiene buena reducción sobre K y Δ es el discriminante de una ecuación de Weierstrass minimal de E, entonces $v(\Delta) = 0$, pero la valoración de K' es un múltiplo de la de K, luego también $v'(\Delta) = 0$, lo que prueba que la ecuación sigue siendo minimal sobre K' y E tiene también

buena reducción sobre K'. Podemos aplicar el teorema 6.19 para concluir que la reducción $E[m] \longrightarrow E(k')$ es inyectiva.

Tomemos ahora $\sigma \in I$ y $P \in E[m]$. Hemos de probar que $P^\sigma = P$. Ahora bien, esto equivale a que $\tilde{P}^\sigma = \tilde{P}$, lo cual es cierto porque σ actúa trivialmente sobre $E(k')$ por definición de I.

Es evidente que b) $\Rightarrow$ c) $\Rightarrow$ d). Veamos, pues, que d) $\Rightarrow$ a). Sea K_{nr} la mayor extensión no ramificada de K. Tomemos un número natural m que cumpla las propiedades siguientes:

a) m es primo con p,

b) $m > |E(K_{\mathrm{nr}}) : E_0(K_{\mathrm{nr}})|$,

c) I actúa trivialmente sobre $E[m]$.

Ahora consideramos las sucesiones exactas

$$0 \longrightarrow E_0(K_{\mathrm{nr}}) \longrightarrow E(K_{\mathrm{nr}}) \longrightarrow E(K_{\mathrm{nr}})/E_0(K_{\mathrm{nr}}) \longrightarrow 0,$$

$$0 \longrightarrow E_1(K_{\mathrm{nr}}) \longrightarrow E_0(K_{\mathrm{nr}}) \longrightarrow E_r(\overline{k}) \longrightarrow 0.$$

Por c) tenemos que $E[m] \subset E(K_{\mathrm{nr}})$. Sabemos que $E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. Para cada primo $q \mid m$, sea q^e la mayor potencia de q que divide a m. Podemos tomar un subgrupo $\langle P\rangle \times \langle Q\rangle$ de $E[m]$ de orden $q^e \times q^e$. Si tuviéramos que $q^{e-1}P \notin E_0(K_{\mathrm{nr}})$ o $q^{e-1}Q \notin E_0(K_{\mathrm{nr}})$ (pongamos el primer caso), entonces la primera sucesión exacta inyectaría $\langle P\rangle$ en el cociente $E(K_{\mathrm{nr}})/E_0(K_{\mathrm{nr}})$, con lo que el índice $|E(K_{\mathrm{nr}}) : E_0(K_{\mathrm{nr}})|$ sería múltiplo de q^e. Por b) esto no puede ocurrir para todo primo q, luego existe un primo $q \mid m$ tal que $q^{e-1}P \in E_0(K_{\mathrm{nr}})$ y $q^{e-1}Q \in E_0(K_{\mathrm{nr}})$. Así pues, el grupo $E_0(K_{\mathrm{nr}})$ contiene un subgrupo isomorfo a $\mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$.

Ahora consideramos la segunda sucesión exacta. Por 6.19 sabemos que $E_1(K_{\mathrm{nr}})$ no contiene elementos de orden q, luego $E_r(\overline{k})$ contiene un subgrupo isomorfo a $\mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$.

Supongamos que E tiene mala reducción sobre K_{nr}. Si la reducción es multiplicativa aplicamos 2.30 a $E(\overline{k})$ (y aquí usamos que $\overline{k}$ es algebraicamente cerrado) de modo que $E_r(\overline{k}) \cong \overline{k}^*$, pero entonces los elementos de orden q en $E_r(\overline{k})$ formarían un subgrupo isomorfo a $\mathbb{Z}/q\mathbb{Z}$ (al grupo de las raíces q-ésimas de la unidad de $\overline{k}$). Si la reducción es aditiva, sería $E_r(\overline{k}) \cong \overline{k}$ y no habría elementos de torsión.

Así pues, E ha de tener buena reducción sobre K_{nr}. El teorema 6.15 nos da que E también tiene buena reducción sobre K. ∎

Veamos una aplicación:

Teorema 6.26 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia no nula definida sobre K entre dos curvas elípticas definidas sobre K. Entonces E_1 tiene buena reducción sobre K si y sólo si la tiene E_2.*

DEMOSTRACIÓN: Sea $m \geq 2$ un número natural primo con la característica de k y con el grado de ϕ. Entonces la restricción $\phi : E_1[m] \longrightarrow E_2[m]$ es inyectiva, pues el núcleo de ϕ tiene orden primo con m, luego de hecho es biyectiva, pues ambos grupos tienen orden m^2. Además la restricción es un isomorfismo de $G(\overline{K}/K)$-módulos, luego el grupo de inercia actúa trivialmente en $E_1[m]$ si y sólo si actúa trivialmente en $E_2[m]$. Basta aplicar el teorema anterior. ∎

Combinando este teorema con 6.22 vemos que si una curva elíptica E/K tiene buena reducción módulo $\mathfrak{P}$, entonces todas las curvas elípticas isógenas a E/K sobre K tienen también buena reducción módulo $\mathfrak{P}$ y las restricciones son isógenas. Si el cuerpo de restos k es finito, el teorema 4.5 nos da además que el cardinal $|\tilde{E}(k)|$ es el mismo para todas las curvas isógenas a E sobre K.

Capítulo VII

Curvas elípticas sobre cuerpos numéricos

Nos ocupamos ahora de las curvas elípticas definidas sobre cuerpos numéricos, en particular sobre $\mathbb{Q}$. En todo el capítulo, K será un cuerpo numérico, $\mathcal{O}$ será su anillo de enteros algebraicos y, para cada divisor primo no arquimediano $\mathfrak{P}$ de K, representaremos por $K_\mathfrak{P}$ la compleción de K respecto de la valoración $v_\mathfrak{P}$, por $\mathcal{O}_\mathfrak{P}$ el anillo de enteros de $K_\mathfrak{P}$, por $k_\mathfrak{P}$ el cuerpo de restos, etc.

Si E/K es una curva elíptica definida sobre K mediante una ecuación de Weierstrass, para cada divisor primo no arquimediano $\mathfrak{P}$ de K podemos considerar la curva elíptica $E/K_\mathfrak{P}$ definida mediante la misma ecuación. Es obvio que un cambio de variables sobre K que haga corresponder dos ecuaciones de Weierstrass para una misma curva E puede verse también como un cambio de variables sobre $K_\mathfrak{P}$. Por lo tanto, a cada curva elíptica E/K le podemos asociar una extensión $E/K_\mathfrak{P}$ a través de una ecuación de Weierstrass, sin que importe la elección de ésta (dos ecuaciones para la misma curva dan lugar a extensiones isomorfas). Más aún, toda isogenia no nula entre dos curvas elípticas se extiende (mediante las mismas ecuaciones) a una isogenia no nula entre las extensiones, luego curvas isógenas sobre K se extienden a curvas isógenas sobre $K_\mathfrak{P}$.

A su vez, a partir de la extensión a $K_\mathfrak{P}$ podemos construir la reducción $\tilde{E}(k_\mathfrak{P})$, que es una curva (tal vez singular) definida sobre el cuerpo finito $k_\mathfrak{P}$. Podemos distinguir entre primos $\mathfrak{P}$ sobre los que E/K tiene buena reducción, reducción multiplicativa o reducción aditiva.

7.1 El discriminante mínimo

Diremos que una ecuación de Weierstrass para una curva elíptica E/K es *entera* si tiene sus coeficientes en $\mathcal{O}$, de modo que también es una ecuación entera para todas las extensiones $E/K_\mathfrak{P}$. Es claro que toda curva elíptica E/K admite una ecuación de Weierstrass de tipo c que, mediante un cambio $X = u^2X'$,

$Y = u^3Y'$ para un $u \in K^*$ adecuado, se transforma en una ecuación entera. Toda ecuación entera cumple obviamente que $\Delta \in \mathcal{O}$.

Fijada una ecuación de Weierstrass entera para una curva elíptica E/K, una condición necesaria que E/K tenga mala reducción módulo un primo $\mathfrak{P}$ es que $\mathfrak{P} \mid \Delta$, luego el conjunto de primos con mala reducción es siempre finito. Sin embargo, la condición no es suficiente, pues la ecuación no tiene por qué ser minimal sobre $\mathfrak{P}$. En esta sección vamos a estudiar si es posible encontrar una ecuación de Weierstrass para E/K que sea minimal para todos los divisores primos no arquimedianos de K, de modo que los primos con mala reducción sean exactamente los que dividen al discriminante de la ecuación.

En principio, para cada divisor primo no arquimediano $\mathfrak{P}$ de K, podemos encontrar una ecuación de Weierstrass minimal sobre $K_\mathfrak{P}$. Si su discriminante es Δ, entonces el natural $\delta_\mathfrak{P} = v_\mathfrak{P}(\Delta)$ es un invariante de E/K, de modo que E/K tiene buena reducción módulo $\mathfrak{P}$ si y sólo si $\delta_\mathfrak{P} = 0$. Esto nos lleva a la definición siguiente:

Definición 7.1 Llamamos *discriminante mínimo* de una curva E/K al ideal de K dado por

$$\mathcal{D}_{E/K} = \prod_\mathfrak{P} \mathfrak{P}^{\delta_\mathfrak{P}},$$

donde $\mathfrak{P}$ recorre los primos no arquimedianos de K.

En estos términos, E/K tiene mala reducción módulo un primo no arquimediano $\mathfrak{P}$ de K si y sólo si $\mathfrak{P} \mid \mathcal{D}_{E/K}$. Naturalmente, esta definición no resuelve nuestro problema. La cuestión es si E/K admite una ecuación minimal global en el sentido siguiente:

Definición 7.2 Una *ecuación de Weierstrass minimal* para una curva elíptica E/K es una ecuación de Weierstrass entera para E cuyo discriminante Δ cumpla $(\Delta) = \mathcal{D}_{E/K}$, es decir, una ecuación que sea minimal para todas las curvas $E/K_\mathfrak{P}$, para todo primo no arquimediano $\mathfrak{P}$ de K.

En principio, si E/K es una curva elíptica definida por una ecuación de Weierstrass

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6, \quad a_i \in \mathcal{O},$$

para cada primo no arquimediano $\mathfrak{P}$ de K existe un cambio de variables

$$X' = u_\mathfrak{P}^2X + r_\mathfrak{P}, \quad Y = u_\mathfrak{P}^3Y + s_\mathfrak{P}u_\mathfrak{P}^2X + t_\mathfrak{P}, \qquad u_\mathfrak{P},\ s_\mathfrak{P},\ t_\mathfrak{P} \in \mathcal{O}_\mathfrak{P},$$

que transforma la ecuación en una ecuación minimal para $E/K_\mathfrak{P}$. La relación entre los discriminantes será $\Delta = u_\mathfrak{P}^{12}\Delta_\mathfrak{P}$. Para todos los primos $\mathfrak{P}$ salvo a lo sumo un número finito de ellos tendremos $v_\mathfrak{P}(\Delta) = v_\mathfrak{P}(\Delta_\mathfrak{P}) = 0$, luego también $v_\mathfrak{P}(u_\mathfrak{P}) = 0$ y podemos definir el ideal

$$\mathfrak{a}_\Delta = \prod_\mathfrak{P} \mathfrak{P}^{v_\mathfrak{P}(u_\mathfrak{P})},$$

de modo que $\mathcal{D}_{E/K} = (\Delta)/\mathfrak{a}_\Delta^{12}$. Ahora observamos que la clase $[\mathfrak{a}_\Delta]$ no depende de Δ. En efecto, si consideramos otra ecuación de Weierstrass entera con discriminante Δ', entonces $\Delta = u^{12}\Delta'$, para cierto $u \in K^*$, luego

$$(\Delta')\mathfrak{a}_{\Delta'}^{12} = \mathcal{D}_{E/K} = (\Delta)\mathfrak{a}_\Delta^{12} = (\Delta')(u\mathfrak{a}_\Delta)^{12},$$

con lo que $\mathfrak{a}_{\Delta'} = (u)\mathfrak{a}_\Delta$.

Definición 7.3 Llamaremos *clase de Weierstrass* de una curva elíptica E/K sobre un cuerpo numérico K a la clase de ideales de K determinada por cualquier ideal $\mathfrak{a}_\Delta$ correspondiente a cualquier ecuación de Weierstrass para E. La representaremos por $W_{E/K}$.

Teorema 7.4 *Una curva elíptica E/K sobre un cuerpo numérico K tiene una ecuación de Weierstrass minimal si y sólo si la clase de Weierstrass $W_{E/K}$ es la clase principal.*

Demostración: Si E/K tiene una ecuación minimal con discriminante Δ, entonces $(\Delta) = \mathcal{D}_{E/K} = (\Delta)/\mathfrak{a}_\Delta^{12}$, luego $\mathfrak{a}_\Delta = 1$ y $W = 1$.

Recíprocamente, supongamos que $W = 1$. Tomemos una ecuación de Weierstrass entera para E y discriminante Δ. Para cada divisor primo no arquimediano $\mathfrak{P}$ de K consideremos un cambio de variables

$$X = u_\mathfrak{P}^2 X' + r_\mathfrak{P}, \quad Y = u_\mathfrak{P}^3 Y' + s_\mathfrak{P} u_\mathfrak{P}^2 X_\mathfrak{P} + t_\mathfrak{P}$$

que la transforme en una ecuación minimal para $\mathfrak{P}$, digamos con coeficientes $a_{i\mathfrak{P}}$ y con discriminante $\Delta_\mathfrak{P}$. La ecuación de partida será ya minimal para todos los primos $\mathfrak{P}$ salvo un número finito de ellos. Si llamamos S al conjunto de estos primos, podemos suponer que $u_\mathfrak{P} = 1$ y $r_\mathfrak{P} = s_\mathfrak{P} = t_\mathfrak{P} = 0$ para todo $\mathfrak{P} \notin S$. En cualquier caso, $u_\mathfrak{P}$, $r_\mathfrak{P}$, $s_\mathfrak{P}$ y $t_\mathfrak{P}$ son $\mathfrak{P}$-enteros.

Por hipótesis, existe $u \in K^*$ tal que $\prod_\mathfrak{P} \mathfrak{P}^{v_\mathfrak{P}(u_\mathfrak{P})} = (u)$. Esto significa que $v_\mathfrak{P}(u_\mathfrak{P}) = v_\mathfrak{P}(u)$, para todo $\mathfrak{P}$.

Por el teorema chino del resto podemos tomar r, s, $t \in \mathcal{O}$ tales que para todo $\mathfrak{P} \in S$ se cumpla que

$$v_\mathfrak{P}(r - r_\mathfrak{P}),\ v_\mathfrak{P}(s - s_\mathfrak{P}), v_\mathfrak{P}(t - t_\mathfrak{P}) > \max_i\{v_\mathfrak{P}(u_\mathfrak{P}^i a_{i\mathfrak{P}})\}.$$

Consideremos ahora la ecuación de Weierstrass para E determinada por el cambio de variables

$$X = u^2 X' + r, \quad Y = u^3 Y' + s u^2 X' + t.$$

Vamos a comprobar que tiene coeficientes enteros (llamémoslos a_i'). Basta probar que $v_\mathfrak{P}(a_i') \geq 0$ para todo primo $\mathfrak{P}$. Si $\mathfrak{P} \notin S$, entonces $u \in U_\mathfrak{P}$, luego la conclusión es clara. Supongamos que $\mathfrak{P} \in S$ y consideremos, por ejemplo, a_2', que cumple

$$u^2 a_2' = a_2 - s a_1 + 3r - s^2.$$

También tenemos que $u_\mathfrak{P}^2 a_{2\mathfrak{P}} = a_2 - s_\mathfrak{P} a_1 + 3r_\mathfrak{P} - s_\mathfrak{P}^2$, luego

$$v_\mathfrak{P}(u^2 a_2') = v_\mathfrak{P}(u_\mathfrak{P}^2 a_{2\mathfrak{P}} - (s - s_\mathfrak{P})a_1 - 3(r - r_\mathfrak{P}) - (s^2 - s_\mathfrak{P}^2))$$

$$= v_\mathfrak{P}(u_\mathfrak{P}^2 a_{2\mathfrak{P}} - (s - s_\mathfrak{P})(a_1 + s + s_\mathfrak{P}) - 3(r - r_\mathfrak{P})) = v_\mathfrak{P}(u_\mathfrak{P}^2 a_{2\mathfrak{P}}),$$

por la elección de r y s. Concluimos que $v_\mathfrak{P}(a_2') = v_\mathfrak{P}(a_{2\mathfrak{P}}) \geq 0$.

Con los coeficientes restantes se razona de forma similar. El discriminante de la nueva ecuación es $\Delta' = u^{-12}\Delta$, luego

$$v_\mathfrak{P}(\Delta') = v_\mathfrak{P}(u^{-12}\Delta) = v_\mathfrak{P}((u_\mathfrak{P}/u)^{12}\Delta_\mathfrak{P}) = v_\mathfrak{P}(\Delta_\mathfrak{P}).$$

Así pues, hemos encontrado una ecuación minimal para E/K. ∎

Como consecuencia inmediata tenemos el teorema siguiente:

Teorema 7.5 *Si K es un cuerpo numérico con número de clases $h = 1$, entonces toda curva elíptica E/K admite una ecuación de Weierstrass minimal.*

Puede probarse que el recíproco es cierto, es decir, que si $h > 1$ entonces existen curvas elípticas que no admiten ecuaciones minimales.

Ejemplo Consideremos el cuerpo $K = \mathbb{Q}(\sqrt{-10})$, que tiene número de clases $h = 2$ y consideremos la curva E/K dada por

$$Y^2 = X^3 + 125.$$

Su discriminante es $\Delta = -2^4 3^3 5^6$. Es fácil ver que en (el anillo de enteros de) K se cumple que $2 = \mathfrak{p}^2$, $5 = \mathfrak{q}^2$, mientras que 3 se conserva primo. Esto hace que la ecuación de E sea minimal para todo primo de K salvo quizá para el primo $\mathfrak{q}$. El cambio de coordenadas

$$X = \sqrt{-10}^2 X', \qquad Y = \sqrt{-10}^3 Y'$$

transforma la ecuación en $Y^2 = X^3 - 1/8$, que es minimal y tiene buena reducción en $\mathfrak{q}$. Por consiguiente, $\mathcal{D}_{E/K} = (2^4 3^3)$ y $W_{E/K}^{12} = [5^6]$, con lo que $W_{E/K} = [\mathfrak{q}]$.

Es claro que $\mathbb{Z}[\sqrt{-10}]$ no tiene elementos de norma 5, luego el ideal $\mathfrak{q}$ no es principal, $W_{E/K} \neq 1$ y así E/K no tiene una ecuación minimal global. ∎

Los teoremas siguientes nos permitirán calcular ecuaciones minimales para curvas elípticas definidas sobre $\mathbb{Q}$. En realidad el método puede generalizarse para cuerpos numéricos arbitrarios (es decir, a un método para decidir si existe ecuación minimal y que permite calcularla en caso afirmativo), pero el caso de $\mathbb{Q}$ es mucho más sencillo. Empezamos por probar una versión global del teorema de Kraus:

Teorema 7.6 *Sean c_4, c_6, $\Delta \in \mathbb{Z}$ tales que $c_4^3 - c_6^2 = 1728\Delta \neq 0$. Entonces existe una curva elíptica $E/\mathbb{Q}$ con una ecuación de Weierstrass entera con covariantes c_4 y c_6 (y discriminante Δ) si y sólo si $v_3(c_6) \neq 2$ y se cumple una de las dos condiciones siguientes:*

a) $c_6 \equiv -1 \pmod 4$,

b) $c_4 \equiv 0 \pmod{16}$ *y* $c_6 \equiv 0,\ 8 \pmod{32}$.

DEMOSTRACIÓN: Si existe tal ecuación ha de cumplir el teorema 6.6 para $K = \mathbb{Q}_2$ y $K = \mathbb{Q}_3$, luego cumple las condiciones del enunciado (ver la observación tras 6.6).

Recíprocamente, si se cumplen estas condiciones, el teorema 6.6 nos da que existen curvas elípticas $E_2/\mathbb{Q}_2$ y $E_3/\mathbb{Q}_3$ que admiten ecuaciones de Weierstrass enteras con covariantes c_4 y c_6. Mediante un cambio de variables adecuado con $u = 1$, ambas se transforman en la misma ecuación de tipo c:

$$Y^2 = X^3 - \frac{c_4}{48}X - \frac{c_6}{864}. \tag{7.1}$$

Equivalentemente, para $p = 2,\ 3$, existen cambios de variable determinados por $r_p,\ s_p,\ t_p \in \mathbb{Q}_p$ y $u = 1$ que transforman esta ecuación en dos ecuaciones enteras sobre $\mathbb{Q}_p$. Vamos a ver que dado un número natural arbitrariamente grande N existen $r,\ s,\ t \in \mathbb{Q}$ tales que

$$v_p(r - r_p) \geq N, \quad v_p(s - s_p) \geq N, \quad v_p(s - s_p) \geq N \quad \text{para } p = 2, 3$$

y $v_p(r),\ v_p(s),\ v_p(t) \geq 0$ para todo primo $p > 3$.

En efecto, por la densidad de $\mathbb{Q}$ en $\mathbb{Q}_p$ podemos aproximar r_p por $r'_p \in \mathbb{Q}$ y por el teorema de aproximación podemos aproximar r'_2 y r'_3 por un mismo número $r' \in \mathbb{Q}$. Sólo falta probar que podemos aproximar r' respecto a v_2 y v_3 mediante un $r \in \mathbb{Q}$ que sea entero respecto de los demás primos. Pongamos que $r' = a/b$. Por el teorema chino del resto existe un $m \in \mathbb{Z}$ tal que

$$v_p(a - m) \geq N + v_p(b), \quad \text{para } p = 2, 3,$$

$$v_p(m) \geq v_p(b) \quad \text{para todo primo } p \mid b, \quad p > 3.$$

Entonces $r = m/b$ cumple lo pedido. Igualmente se construyen s y t. Ahora basta observar que si tomamos N suficientemente grande, el cambio de variables determinado por $r,\ s,\ t$ y $u = 1$ transforma la ecuación (7.1) en una ecuación entera (que obviamente tendrá también covariantes c_4 y c_6). En efecto, las fórmulas de 2.6 (teniendo en cuenta que los polinomios son funciones continuas) muestran que los coeficientes de esta ecuación estarán cerca de los de las ecuaciones obtenidas con $r_p,\ s_p,\ t_p$, y todo número racional suficientemente próximo a un entero p-ádico es un entero p-ádico. En conclusión, el cambio de variable lleva a una ecuación con coeficientes enteros diádicos y triádicos. Obviamente también son enteros para los demás primos, luego están en $\mathbb{Z}$. ∎

Con esto podemos calcular fácilmente los covariantes (y el discriminante) de las ecuaciones minimales de una curva dada:

Teorema 7.7 *Sea $E/\mathbb{Q}$ una curva elíptica dada por una ecuación de Weierstrass entera con covariantes c_4 y c_6 y discriminante Δ. Sea u el mayor número natural tal que $c'_4 = u^{-4}c_4$ y $c'_6 = u^{-6}c_6$ son enteros y satisfacen las condiciones del teorema anterior. Entonces toda ecuación de Weierstrass entera con covariantes c'_4 y c'_6 es una ecuación minimal para $E/\mathbb{Q}$.*

DEMOSTRACIÓN: El cambio de variables $X = u^2X'$, $Y = u^3Y'$ transforma la ecuación de $E/\mathbb{Q}$ en una ecuación (no necesariamente entera) de covariantes c_4', c_6', la cual puede transformarse a su vez en la ecuación (7.1) (con c_i' en lugar de c_i).

Por otra parte, cualquier ecuación entera de covariantes c_i' —y existe al menos una por el teorema anterior— puede transformarse en esa misma ecuación, luego también en la ecuación de partida.

En definitiva, hemos probado que la curva $E/\mathbb{Q}$ admite una ecuación de Weierstrass entera de covariantes c_4', c_6' (y discriminante Δ'). La maximalidad de u hace que dicha ecuación sea necesariamente minimal. ∎

Así pues, para encontrar una ecuación minimal a una curva dada sólo nos falta encontrar una ecuación entera con unos covariantes dados.

Definición 7.8 Una ecuación de Weierstrass con coeficientes en $\mathbb{Z}$ está *reducida* si a_1, $a_3 \in \{0,1\}$ y $a_2 \in \{-1,0,1\}$.

Teorema 7.9 *Toda ecuación de Weierstrass con coeficientes en $\mathbb{Z}$ se transforma mediante un cambio de variables entero en una única ecuación reducida. Dados c_4, c_6, $\Delta \in \mathbb{Z}$ en las condiciones del teorema anterior, la ecuación reducida con tales covariantes tiene los coeficientes siguientes:*

$$\begin{array}{rcll} a_1 & \equiv & c_4 \text{ (mód 2)}, & [a_1 \in \{0,1\}], \\ a_2 & \equiv & -c_6 - a_1 \text{ (mód 3)}, & [a_2 \in \{-1,0,1\}], \\ a_3 & \equiv & (b_2^3 - 3c_4b_2 - 2c_6)/16 \text{ (mód 2)}, & [a_3 \in \{0,1\}, b_2 = a_1 + 4a_2], \\ a_4 & = & (b_2 - 24a_1a_3 - c_4)/48, & \\ a_6 & = & (-b_2^3 - c_6 + 36b_2(a_1a_3 + 2a_4) - 216a_3)/864. & \end{array}$$

DEMOSTRACIÓN: Consideremos una ecuación de Weierstrass entera con coeficientes $a_1, \ldots, a_6$ y vamos a construir una nueva ecuación reducida con coeficientes $a_1', \ldots, a_6'$, al mismo tiempo que construimos un cambio de variables con $u = 1$ y r, s, $t \in \mathbb{Z}$ que transforme una en la otra. Según el teorema 2.6, ha de ser $a_1' = a_1 + 2s$, luego para que la ecuación que obtengamos sea reducida hemos de tomar necesariamente

$$a_1' \equiv a_1 \text{ (mód 2)}, \quad a_1' \in \{0,1\}, \quad s = (a_1' - a_1)/2.$$

Similarmente, ha de ser

$$\begin{array}{rcllll} a_2' & \equiv & a_2 - a_1 - s^2 \text{ (mód 3)}, & a_2' \in \{-1,0,1\}, & r = (a_2' - a_2 + a_1 + s^2)/3, \\ a_3' & \equiv & a_3 + ra_1 \text{ (mód 2)}, & a_3' \in \{0,1\}, & t = (a_3' - a_3 - ra_1)/2, \\ a_4' & = & a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 3r^2 - 2st, & & \\ a_6' & = & a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1. & & \end{array}$$

Esto prueba la existencia de la ecuación. Ahora vamos a probar que toda ecuación de Weierstrass reducida puede recuperarse a partir de los covariantes c_4 y c_6 como indica el enunciado, lo que en particular nos dará la unicidad. Para adecuarnos a la notación del enunciado, a partir de aquí los coeficientes de la ecuación reducida serán a_i en lugar de a_i'.

Como $a_1 = a_1^2$, tenemos que

$$a_1 \equiv b_2 \equiv b_2^2 \equiv c_4 \text{ (mód 2)},$$

luego a_1 ha de ser el que indica el enunciado. Igualmente,

$$a_2 = b_2 - a_1^2 - 3a_2 \equiv b_2 - a_1^2 = b_2 - a_1 \equiv b_2^3 - a_1 \equiv -c_6 - a_1 \text{ (mód 3)},$$

$$(b_2^3 - 3c_4b_2 - 2c_6)/16 = 27b_6 \equiv b_6 \equiv a_3^2 \equiv a_3 \text{ (mód 2)},$$

$$\frac{b_2^2 - 24a_1a_3 - c_4}{48} = \frac{-24a_1a_3 + 24b_4}{48} = \frac{-a_1a_3 + b_4}{2} = a_4,$$

e igualmente se justifica la fórmula para a_6. ∎

Como consecuencia inmediata:

Teorema 7.10 *Toda curva elíptica $E/\mathbb{Q}$ admite una única ecuación de Weierstrass minimal reducida.*

Demostración: Cuando transformamos una ecuación minimal en una ecuación reducida según el teorema anterior, la ecuación que obtenemos sigue siendo minimal, pues el cambio de variables tiene $u = 1$. Una ecuación minimal se transforma en otra mediante un cambio de variables entero con $u = 1$ (por el teorema 6.2 aplicado a todos los primos). Por consiguiente ambas tienen los mismos covariantes, luego si ambas son reducidas han de ser la misma. ∎

En definitiva, dada una curva elíptica $E/\mathbb{Q}$ dada por una ecuación de Weierstrass, podemos aplicar un cambio de variables $X = u^2X'$, $Y = u^3Y'$ para pasar a una ecuación entera con covariantes c_4 y c_6, luego calculamos los covariantes minimales c_4' y c_6' según el teorema 7.7 y construimos la ecuación minimal según el teorema anterior.

Resulta natural preguntarse si puede existir una curva elíptica E/K con $\mathcal{D}_{E/K} = 1$, es decir, una curva con buena reducción módulo todos los primos no arquimedianos de K. El teorema siguiente muestra que la respuesta es negativa cuando $K = \mathbb{Q}$:

Teorema 7.11 *Sea $E/\mathbb{Q}$ una curva elíptica dada por una ecuación de Weierstrass entera de discriminante $\Delta \in \mathbb{Z}$. Entonces Δ no puede ser de la forma d^3, donde d es un entero cuyos divisores primos sean todos congruentes con 1 módulo 8. En particular no puede ser $\Delta = \pm 1$.*

Demostración: En caso contrario $\Delta = d^3$, con $d \equiv \pm 1$ (mód 8). Veamos que a_1 ha de ser impar. En caso contrario $v_2(b_2) \geq 2$, $v_2(b_4) \geq 1$, $v_2(c_4) \geq 4$, y la relación

$$c_6^2 = c_4^3 + 1728d^3 \tag{7.2}$$

implica que $c_6 = 8c$, pero entonces $c^2 \equiv 27d^3 \equiv \pm 3$ (mód 8), y esta congruencia no tiene solución.

Así pues, a_1 es impar, luego b_2 también y $c_4 \equiv b_2^2 \equiv 1$ (mód 8). Sustituyendo en (7.2) $x = c_4 + 12d$, $y = c_6$ obtenemos

$$y^2 = x(x^2 - 36dx^2 + 432d^2),$$

donde $x \equiv 5$ (mód 8). En particular, $x \neq 0$. Ahora observamos que

$$Q = x^2 - 36dx^2 + 432d^2 = (x - 18d)^2 + 108d^2 > 0,$$

luego también $x = y^2/Q > 0$. Descompongamos

$$x = 3^u \prod_p p^{v_p} \prod_q q^{v_q},$$

donde p recorre los primos que dividen al $\mathrm{mcd}(x, d)$ y q los divisores de x distintos de 3 y que no dividen a d. Claramente $v_q(Q) = 0$, luego $v_q = v_q(y^2)$ es par, luego $q^{v_q} \equiv 1$ (mód 8). Por hipótesis también $p \equiv 1$ (mód 8), luego

$$x \equiv 3^u \equiv 1,\ 3 \text{ (mód 8)},$$

cuando teníamos que $x \equiv 5$ (mód 8). ∎

Ejemplo Las curvas siguientes muestran que la restricción sobre los divisores primos de Δ es necesaria en el teorema anterior:

$$\begin{array}{ll} Y^2 = X^3 - X, & \Delta = \ \ 2^6, \\ Y^2 + Y = X^3, & \Delta = -3^3, \\ Y^2 + XY = X^3 - X^2 - 2X - 1, & \Delta = -7^3. \end{array}$$

Por el contrario, si $K = \mathbb{Q}(\sqrt{29}\,)$ y $\epsilon = (5+\sqrt{29})/2$ es la unidad fundamental de K, la curva $Y^2 + XY + \epsilon^2 Y = X^3$ cumple $\Delta = -\epsilon^{10}$, luego tiene buena reducción módulo todos los primos de K. ∎

Terminaremos la sección justificando la existencia de ecuaciones minimales en un sentido más débil que no requiere ninguna hipótesis sobre el número de clases de K. Para ello recordamos previamente algunos resultados sobre cuerpos numéricos:

Teorema 7.12 *Sea K un cuerpo numérico y S un conjunto finito de divisores primos de K que contenga a todos los primos arquimedianos y sea*

$$K_S = \{\alpha \in K \mid v_{\mathfrak{P}}(\alpha) \geq 0 \text{ para todo } \mathfrak{P} \notin S\}.$$

Entonces:

a) K_S es la localización del anillo de enteros $\mathcal{O}_K$ respecto del conjunto multiplicativo formado por los enteros divisibles a lo sumo entre primos de S, es decir:

$$K_S = \{a/b \mid a,\, b \in \mathcal{O}_K,\ v_{\mathfrak{P}}(b) = 0 \text{ para todo } \mathfrak{P} \notin S\}.$$

b) K_S es un dominio de Dedekind (en particular es íntegramente cerrado) y S puede extenderse a un conjunto finito mayor de modo que K_S sea un dominio de ideales principales.

DEMOSTRACIÓN: a) Una inclusión es obvia y para probar la contraria usamos que todo ideal de $\mathcal{O}_K$ elevado al número de clases h es principal. Así, si $\alpha \in K_S$, el ideal fraccional (α) admite una descomposición

$$(\alpha) = \frac{\mathfrak{a}}{\mathfrak{q}_1^h \cdots \mathfrak{q}_s^h},$$

donde $\mathfrak{q}_i \in S$ y $\mathfrak{a}$ es un ideal de K. Ahora bien, los ideales $\mathfrak{q}_i^h$ son principales y $\mathfrak{a}$ también tiene que serlo. En definitiva, $\alpha = a/b$, con a, $b \in \mathcal{O}_K$ y b sólo es divisible entre primos de S.

b) En general, toda localización de un dominio de Dedekind es un dominio de Dedekind. Además, los ideales de K_S son de la forma $\mathfrak{a}_S = \mathfrak{a}K_S$, donde $\mathfrak{a}$ es un ideal de $\mathcal{O}$.

Fijamos un conjunto de representantes $\mathfrak{a}_1, \ldots, \mathfrak{a}_h$ de las clases de ideales de K, en cada uno de ellos tomamos un elemento $\alpha_i \in \mathfrak{a}_i$ no nulo y añadimos a S los primos que lo dividen. Esto hace que α_i sea una unidad en K_S y que, por consiguiente, $\mathfrak{a}_{iS} = 1$.

Para cada ideal $\mathfrak{a}_S$ de K_S existe un i tal que $(\alpha)\mathfrak{a} = (\beta)\mathfrak{a}_i$, para ciertos α, $\beta \in \mathcal{O}_K$, de donde $(\alpha)\mathfrak{a}_S = (\beta)$, lo que implica que $\mathfrak{a}_S$ es principal. ∎

Teorema 7.13 *Sea K un cuerpo numérico y S un conjunto finito de valores absolutos de K que contenga a todos los primos arquimedianos y a todos los divisores de 2 y 3. Supongamos además que la localización K_S es un dominio de ideales principales. Entonces toda curva elíptica E/K admite una ecuación de Weierstrass $Y^2 = X^3 + AX + B$ con A, $B \in K_S$, tal que, en K_S, $(\mathcal{D}_{E/K}) = (\Delta)$, donde el discriminante es $\Delta = -16(4A^3 + 27B^2)$.*

DEMOSTRACIÓN: Tomemos cualquier ecuación de Weierstrass para E/K de la forma indicada en el enunciado con coeficientes enteros y discriminante Δ. Para cada primo $\mathfrak{p} \notin S$ podemos conseguir una ecuación minimal mediante un cambio de la forma

$$X = u_\mathfrak{p}^2 X', \qquad Y = u_\mathfrak{p}^3 Y', \qquad u_\mathfrak{p} \in K^*.$$

En efecto, con este cambio la ecuación se transforma en

$$Y^2 = X^3 + u_\mathfrak{p}^{-4} Ax + u_\mathfrak{p}^{-6} B,$$

y por 6.4 la ecuación será minimal si $v_\mathfrak{p}(u_\mathfrak{p}^{-4}A) \geq 0$, $v_\mathfrak{p}(u_\mathfrak{p}^{-6}B) \geq 0$ y o bien $v_\mathfrak{p}(u_\mathfrak{p}^{-4}A) < 4$ o bien $v_\mathfrak{p}(u_\mathfrak{p}^{-6}B) < 6$. Así, basta tomar $u_\mathfrak{p}$ de modo que $v_\mathfrak{p}(u_\mathfrak{p})$ sea el máximo que respete las dos primeras condiciones y necesariamente se cumplirá una de las dos últimas.

Para los posibles primos no arquimedianos en S el cambio necesario puede ser más complicado, pero en cualquier caso determina unos números $u_\mathfrak{p}$ con los que podemos formar el ideal

$$\mathfrak{a}_\Delta = \prod_{\mathfrak{p}\nmid\infty} \mathfrak{p}^{v_\mathfrak{p}(u_\mathfrak{p})}$$

que verifica

$$\mathcal{D}_{E/K} = (\Delta)/\mathfrak{a}_\Delta^{12}.$$

Si consideramos estos ideales como ideales de K_S, entonces los primos de S se vuelven triviales, con lo que tenemos

$$\mathcal{D}_{E/K} = (\Delta)/\prod_{\mathfrak{p}\notin S} \mathfrak{p}^{12v_\mathfrak{p}(u_\mathfrak{p})}$$

Como K_S es un dominio de ideales principales, existe un $u \in K_S$ tal que

$$\prod_{\mathfrak{p}\notin S} \mathfrak{p}^{v_\mathfrak{p}(u_\mathfrak{p})} = (u).$$

Esto significa que $v_\mathfrak{p}(u) = v_\mathfrak{p}(u_\mathfrak{p})$ para todo primo $\mathfrak{p} \notin S$, luego la ecuación

$$Y^2 = X^3 + u^{-4}Ax + u^{-6}B$$

tiene sus coeficientes en K_S y es minimal para todos los primos $\mathfrak{p} \notin S$. Su discriminante es $\Delta' = u^{-12}\Delta$, luego en K_S tenemos

$$\mathcal{D}_{E/K} = (\Delta)/(u)^{12} = (\Delta').$$

■

7.2 El subgrupo de torsión

Estudiamos ahora los puntos de torsión de una curva elíptica E/K, es decir, del subgrupo de torsión $E_{\mathrm{tor}}(K)$ del grupo $E(K)$. El teorema siguiente es una consecuencia directa de 6.20:

Teorema 7.14 *Sea E/K una curva elíptica determinada por una ecuación de Weierstrass entera y sea $P \in E(K)$ un punto de orden $m \geq 2$.*

a) Si m no es potencia de primo, entonces $x(P)$, $y(P) \in \mathcal{O}$.

b) Si $m = p^n$, para cada divisor primo no arquimediano $\mathfrak{P}$ de K, sea $r_\mathfrak{P}$ la parte entera de $v_\mathfrak{P}(p)/(p^n - p^{n-1})$. Entonces

$$v_\mathfrak{P}(x(P)) \geq -2r_\mathfrak{P}, \qquad v_\mathfrak{P}(y(P)) \geq -3r_\mathfrak{P}.$$

Para $K = \mathbb{Q}$ el teorema anterior implica que los puntos de torsión tienen coordenadas enteras salvo quizá por un denominador 4 en la x y un denominador 8 en la y (y en tal caso el punto tiene orden 2). Más precisamente:

Teorema 7.15 *Sea $E/\mathbb{Q}$ una curva elíptica dada por una ecuación de Weierstrass entera. Entonces todo punto de torsión de $E(\mathbb{Q})$ (distinto de O) tiene coordenadas enteras, salvo quizá un único punto de orden 2 con coordenadas $(a/4, b/8)$, con a, $b \in \mathbb{Z}$ impares. Para que pueda existir tal punto es necesario que a_1 sea impar.*

DEMOSTRACIÓN: Si existe un punto $P = (x, y)$ de orden 2, entonces la tangente a la curva en dicho punto ha de ser vertical, luego la derivada respecto de Y de la ecuación de Weierstrass se ha de anular. Esto significa que

$$y = -\frac{a_1 x + a_3}{2}.$$

El teorema anterior afirma que $v_2(P) \leq 1$, donde v_2 es la aplicación definida en 6.17 para la curva $E/\mathbb{Q}_2$ definida por la misma ecuación de Weierstrass. Para que P sea fraccionario es necesario que $v_2(P) = 1$, con lo que $v_2(x) = -2$ y $v_2(y) = -3$. Entonces

$$-2 = v_2(2y) = v_2(a_1 x + a_3) \geq \min\{v_2(a_1 x), v_2(a_3)\}.$$

Como $v_2(a_3) \geq 0$, el mínimo se alcanza en $v_2(a_1 x) = v_2(a_1) + v_2(x) = -2$, luego $v_2(a_1) = 0$ y así a_1 es impar.

Si $E(K)$ contiene otro punto de orden 2, de hecho ha de contener a los tres que hay en $E(\overline{K})$, digamos $P_i = (x_i, y_i)$, para $i = 1, 2, 3$. Vamos a probar que en tal caso sólo uno de ellos es fraccional. Si lo fueran dos de ellos, también lo sería el tercero, pues los puntos con coordenadas no enteras diádicas forman un subgrupo de $E(\mathbb{Q}_2)$ (el núcleo de la reducción módulo 2). El cambio

$$Y = Y' - \frac{a_1}{2} X' - \frac{a_3}{2}$$

transforma la ecuación en una de tipo b. Los puntos de orden 2 de esta ecuación se caracterizan por tener la segunda coordenada nula, y han de corresponderse con los puntos de orden 2 de la ecuación original, luego serán $(x_i, 0)$. Esto implica que x_1, x_2, x_3 son las raíces del miembro derecho de la ecuación de Weierstrass de tipo b. En particular

$$x_1 x_2 x_3 = \frac{b_6}{4} = a_6 + \frac{a_3^2}{4}.$$

Por una parte, $v_2(x_1 x_2 x_3) = -6$, mientras que por otra $v_2(a_6 + a_3^2/4) \geq -2$, y esta contradicción prueba el teorema. ∎

Ejemplo La curva $Y^2 = X^3 + 8$ contiene los puntos de coordenadas enteras $(1, \pm 3)$ y $(2, \pm 4)$ que, no obstante, tienen orden infinito, pues, por ejemplo

$$2(1, 3) = (-14/8, -13/8), \qquad 2(2, 4) = (-14/8, 13/8),$$

y esta curva no puede tener puntos de torsión con coordenadas fraccionarias. ∎

El teorema siguiente generaliza un resultado obtenido independientemente por Lutz y Nagell y permite en muchos casos calcular el subgrupo de torsión de una curva elíptica $E/\mathbb{Q}$:

Teorema 7.16 *Sea $E/\mathbb{Q}$ una curva elíptica dada por una ecuación de Weierstrass entera, sea $P = (x, y)$ un punto de torsión no nulo y llamemos*

$$y' = y + (a_1x + a_3)/2.$$

Entonces $y' = 0$ si y sólo si P tiene orden 2. En caso contrario $2y' \in \mathbb{Z}$ y $(2y')^2 \mid 4\Delta$. Si además a_1 es par entonces $(2y')^2 \mid \Delta$ y si a_3 también es par entonces $y' \in \mathbb{Z}$ y $16y'^2 \mid \Delta$.

Demostración: Observemos que el cambio de variables que transforma la ecuación de $E/\mathbb{Q}$ en una ecuación de tipo b hace corresponder el punto P con el punto $P' = (x, y')$. El punto P tiene orden 2 si y sólo si lo tiene P', lo que ciertamente equivale a que $y' = 0$.

Si P no tiene orden 2, el teorema anterior nos da que $x,\ y \in \mathbb{Z}$, luego también $2y' = 2y + a_1x + a_3 \in \mathbb{Z}$. El punto $2P$ también es de torsión, luego sus coordenadas son enteras salvo quizá si tiene orden 2, pero en cualquier caso $4x(2P) \in \mathbb{Z}$.

Ahora necesitamos algunos cálculos. Recordemos la fórmula de duplicación del teorema (2.21):

$$x(2P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}.$$

Observemos que el denominador es $(2y')^2$. Por otra parte, una comprobación rutinaria nos da la siguiente identidad de polinomios:

$$u(X)(4X^3 + b_2X^2 + 2b_4X + b_6) - v(X)(X^4 - b_4X^2 - 2b_6X - b_8) = \Delta, \quad (7.3)$$

donde

$$\begin{aligned} u(X) &= 12X^3 - b_2X^2 - 10b_4X + b_2b_4 - 27b_6, \\ v(X) &= 48X^2 + 8b_2X + 32b_4 - b_2^2. \end{aligned}$$

Combinando ambas relaciones obtenemos que

$$(2y')^2(u(x) - x(2P)v(x)) = \Delta.$$

Multiplicando por 4 ambos miembros, el segundo factor es entero, con lo que $(2y')^2 \mid 4\Delta$. Si a_1 es par entonces $x(2P)$ ha de ser entero y no hace falta multiplicar por 4.

Finalmente, si a_3 también es par vemos que $4 \mid b_2$, $2 \mid b_4$, $4 \mid b_6$, luego $4 \mid u(x)$ y $4 \mid v(x)$, de donde concluimos que $16y'^2 \mid \Delta$. ∎

Ejemplo Consideremos la curva $E/\mathbb{Q}$ dada por la ecuación

$$Y^2 + XY + Y = X^3, \qquad \Delta = -2 \cdot 1.$$

Si (x, y) es un punto de torsión de orden > 2 entonces $(2y')^2 \mid 2 \cdot 13$, con lo que $2y' = \pm 1$. La ecuación en forma b es

$$Y'^2 = X^3 + \frac{1}{4}X^2 + \frac{1}{2}X + \frac{1}{4},$$

y al hacer $Y' = \pm 1/2$ obtenemos

$$X^3 + \frac{1}{4}X^2 + \frac{1}{2}X = 0,$$

cuya única raíz entera es $x = 0$. La relación $2y' = 2y + x + 1$ nos da los puntos $(0, 0)$ y $(0, -1)$.

Por otra parte, si (x, y) fuera un punto de orden 2, la coordenada x cumpliría la ecuación

$$X^3 + \frac{1}{4}X^2 + \frac{1}{2}X + \frac{1}{4} = 0,$$

y $4x$ sería una raíz entera de la ecuación

$$X^3 + X^2 + 8X + 16 = 0.$$

Se comprueba que no existen raíces enteras, luego $E(\mathbb{Q}) = \{O, (0, 0), (0, -1)\}$. ∎

Ejemplo Consideremos la curva $E/\mathbb{Q}$ dada por la ecuación

$$Y^2 = X^3 - 43X + 166, \qquad \Delta = -2^{19} \cdot 13.$$

El polinomio de la derecha no tiene raíces enteras, por lo que no hay puntos de orden 2. Si (x, y) es un punto de torsión, se ha de cumplir que $y^2 \mid 2^{15} \cdot 13$, luego $y \mid 128$.

Esto nos da un máximo de seis posibles puntos de torsión (además de O):

$$(3, \pm 8), \quad (-5, \pm 16), \quad (11, \pm 32).$$

Usando la fórmula de duplicación del teorema 2.21 comprobamos que si $P = (3, 8)$ entonces

$$x(P) = 3, \quad x(2P) = -5, \quad x(4P) = 11, \quad x(8P) = 3,$$

luego $8P = \pm P$, lo que implica que P tiene orden 7, 3 o 9. No puede tener orden 9 porque a lo sumo hay 7 puntos de torsión y si tuviera orden 3 sería $x(2P) = x(-P) = 3$, luego P tiene orden 7 y concluimos que $E_{\mathrm{tor}}(\mathbb{Q})$ es un grupo cíclico de orden 7. ∎

Del teorema anterior se sigue que si $E/\mathbb{Q}$ es una curva elíptica, entonces $E_{\mathrm{tor}}(\mathbb{Q})$ es un grupo finito. En realidad esto es también una consecuencia sencilla del teorema 6.19. A continuación damos una versión más cómoda en la práctica:

Teorema 7.17 *Sea E/Q una curva elíptica dada por una ecuación de Weierstrass entera y sea S el conjunto de los primos con buena reducción. Entonces el orden de $E_{\rm tor}(\mathbb{Q})$ divide a*

$$\operatorname*{mcd}_{p\in S}(\epsilon_p|\tilde{E}(\mathbb{Z}/p\mathbb{Z})|),$$

donde $\epsilon_p = 1$ salvo si $p = 2$ y $E(\mathbb{Q})$ tiene un punto fraccional de orden 2, en cuyo caso $\epsilon_2 = 2$.

Demostración: Consideramos la curva $E/\mathbb{Q}_p$ definida por la misma ecuación. El núcleo de la reducción módulo p es $E_1(\mathbb{Q}_p)$, que está formado por los puntos de coordenadas no enteras p-ádicas. Como los puntos de torsión tienen coordenadas enteras (racionales), vemos que la reducción módulo p es inyectiva en $E_{\rm tor}(\mathbb{Q})$, salvo si $p = 2$ y hay un punto fraccionario de orden 2, en cuyo caso el núcleo de la reducción tiene orden 2 y $|E_{\rm tor}(\mathbb{Q})|$ divide a $2|\tilde{E}(\mathbb{Z}/2\mathbb{Z})|$. ∎

La tabla siguiente contiene 15 ejemplos de curvas elípticas sobre $\mathbb{Q}$ con la estructura de su grupo de torsión. Un notable teorema debido a Mazur asegura que el grupo de torsión de una curva elíptica sobre $\mathbb{Q}$ ha de ser necesariamente uno de los 15 grupos que aparecen en la tabla.

Ecuación	Δ	$E_{\rm tor}(\mathbb{Q})$	Generadores
$Y^2 = X^3 - 2$	$-2^6 \cdot 3^3$	1	O
$Y^2 = X^3 + 8$	$-2^{10} \cdot 3^3$	C_2	$(-2, 0)$
$Y^2 = X^3 + 4$	$-2^8 \cdot 3^3$	C_3	$(0, 2)$
$Y^2 = X^3 + 4X$	-2^{12}	C_4	$(2, 4)$
$Y^2 - Y = X^3 - X^2$	-11	C_5	$(0, 0)$
$Y^2 = X^3 + 1$	$-2^4 \cdot 3^3$	C_6	$(2, 3)$
$Y^2 = X^3 - 43X + 166$	$-2^{19} \cdot 13$	C_7	$(3, 8)$
$Y^2 + 7XY = X^3 + 16X$	$2^8 \cdot 3^4 \cdot 17$	C_8	$(-2, 4)$
$Y^2 + XY + Y = X^3 - X^2 - 14X + 29$	$-2^9 \cdot 3^5$	C_9	$(3, 1)$
$Y^2 + XY = X^3 - 45X + 81$	$2^{10} \cdot 3^5 \cdot 11$	C_{10}	$(0, 9)$
$Y^2 + 43XY - 210Y = X^3 - 210X^2$	$2^{12} \cdot 3^6 \cdot 5^3 \cdot 7^4 \cdot 13$	C_{12}	$(0, 0)$
$Y^2 = X^3 - 4X$	-2^{12}	$C_2 \oplus C_2$	$(0, 0), (0, 2)$
$Y^2 = X^3 + 2X^2 - 3X$	$2^8 \cdot 3^2$	$C_2 \oplus C_4$	$(-3, 0), (-1, 2)$
$Y^2 + 5XY - 6Y = X^3 - 3X^2$	$2^2 \cdot 3^6 \cdot 5^2$	$C_2 \oplus C_6$	$(2, -2), (0, 0)$
$Y^2 + 17XY - 120Y = X^3 - 60X^2$	$2^8 \cdot 3^8 \cdot 5^4 \cdot 7^2$	$C_2 \oplus C_8$	$(-40, 400), (0, 0)$

Ejemplo Veamos cómo se calcula el grupo de torsión de la última curva de la tabla. Es claro que $(0, 0)$ está en la curva, y aplicando la fórmula de duplicación se observa que su orden es exactamente 8. Así tenemos ya un subgrupo C_8 de $E_{\rm tor}(\mathbb{Q})$. Por otro lado, resolviendo el sistema de ecuaciones $F(X, Y) = 0$ y $F_Y(X, Y) = 0$ encontramos todos los puntos de orden 2, que son $(24, -144)$, $(-40, 400)$ y $(15/4, 225/8)$. El primero es $4 \cdot (0, 0)$, mientras que con el segundo obtenemos un nuevo subgrupo C_2, con lo que en total tenemos un subgrupo $C_2 \oplus C_8$. Ahora hemos de probar que no hay más elementos de torsión.

Como E tiene buena reducción módulo 11, sabemos que $E(\mathbb{Q})$ se sumerge en $\tilde{E}(\mathbb{Z}/11\mathbb{Z})$. Ahora bien, el número de puntos de esta reducción es a lo sumo 23 (cada valor de x da a lo sumo dos valores posibles para y, luego hay a lo sumo 22 puntos finitos). Por otra parte, dicho número ha de ser múltiplo de 16, luego $|\tilde{E}(\mathbb{Z}/11\mathbb{Z})| = 16$ y también $|E_{\rm tor}(\mathbb{Q})| = 16$. ∎

Finalmente vamos a determinar los grupos de torsión sobre $\mathbb{Q}$ de dos familias de curvas elípticas especialmente simples, las de la forma $Y^2 = X^3 + AX$ y las de la forma $Y^2 = X^3 + B$.

Observemos que para estudiar la torsión de $Y^2 = X^3 + AX$ podemos suponer que A está libre de potencias cuartas, pues si $A = m^4A'$, el cambio $X = m^2X$, $Y = m^3Y'$ transforma la curva en $Y^2 = X^3 + AX$ y la estructura del grupo de torsión no varía.

Teorema 7.18 *Sea $E/\mathbb{Q}$ la curva dada por $Y^2 = X^3 + AX$, con $A \in \mathbb{Z}$ (de modo que $\Delta = -2^6A^3$ y $j = 1728$). Supongamos que A está libre de potencias cuartas. Entonces*

$$E_{\mathrm{tor}}(\mathbb{Q}) \cong \begin{cases} C_2 \oplus C_2 & \text{si } -A \text{ es un cuadrado,} \\ C_4 & \text{si } A = 4, \\ C_2 & \text{en otro caso.} \end{cases}$$

Demostración: Sea $p \equiv -1 \pmod 4$ un primo que no divida a Δ. (Existe por el teorema de Dirichlet sobre primos en progresiones aritméticas.) Por el teorema 7.17 tenemos que $|E_{\mathrm{tor}}(\mathbb{Q})|$ divide a $|\tilde{E}(\mathbb{Z}/p\mathbb{Z})| = p + 1$, donde hemos usado el teorema 4.9, según el cual la reducción módulo p es supersingular.

Si en particular tomamos $p \equiv 3 \pmod 8$ vemos que $8 \nmid |E_{\mathrm{tor}}(\mathbb{Q})|$, pues en caso contrario tendríamos que $8 \mid p + 1$.

Si elegimos $p \equiv 7 \pmod{12}$ vemos que $3 \nmid |E_{\mathrm{tor}}(\mathbb{Q})|$.

Si $q > 3$ es primo, podemos tomar $p \equiv 3 \pmod{4q}$, de donde se sigue que $q \nmid |E_{\mathrm{tor}}(\mathbb{Q})|$, ya que en caso contrario $q \mid p + 1 \equiv 4 \pmod q$ y $q \mid 4$.

Con esto podemos concluir que $|E_{\mathrm{tor}}(\mathbb{Q})| \mid 4$.

En cualquier caso, $E(\mathbb{Q})$ tiene un punto de orden 2, a saber, $(0, 0)$. Tendrá tres puntos de orden 2 si y sólo si $X^3 + AX$ se escinde en $\mathbb{Q}[X]$, es decir, si y sólo si $-A$ es un cuadrado.

Sólo falta ver que $E(\mathbb{Q})$ tiene un punto de orden 4 si y sólo si $A = 4$. Ciertamente, para $A = 4$ está el punto $(2, 4)$. En general, la fórmula de duplicación nos da que si existe un punto (x, y) que cumple $2(x, y) = (0, 0)$, entonces

$$x^4 - 2Ax^2 + A^2 = (x^2 - A)^2 = 0.$$

Así pues, ha de ser $x^2 = A$ y, como A no tiene potencias cuartas, x ha de ser libre de cuadrados. Ahora bien, x cumple la ecuación

$$y^2 = x^3 + Ax = 2x^3.$$

De aquí se sigue que x no puede ser divisible entre primos impares, y la única posibilidad resulta ser $x = 2$, luego $A = 4$. ∎

Para estudiar la torsión de $Y^2 = X^3 + B$ podemos suponer que B está libre de potencias sextas.

Teorema 7.19 *Sea $E/\mathbb{Q}$ la curva dada por $Y^2 = X^3 + B$, con $B \in \mathbb{Z}$ (de modo que $\Delta = -2^4 \cdot 3^3 \cdot B^2$, $j = 0$). Supongamos que B está libre de potencias sextas.*

Entonces

$$E_{\mathrm{tor}}(E) \cong \begin{cases} C_6 & \text{si } B = 1, \\ C_3 & \text{si } B = -2^4 \cdot 3^3 \text{ o } B \text{ es un cuadrado } B \neq 1, \\ C_2 & \text{si } B \text{ es un cubo, } B \neq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

DEMOSTRACIÓN: Si $p \equiv -1$ (mód 3) es un primo que no divide a Δ, entonces los teoremas 7.17 y 4.9 nos dan que $|E_{\mathrm{tor}}(\mathbb{Q})|$ divide a $|\, p+1$.

Si tomamos $p \equiv 5$ (mód 12) vemos que 4 no puede dividir a $|E_{\mathrm{tor}}(\mathbb{Q})|$. Si $p \equiv 2$ (mód 9) vemos que 9 tampoco divide a $|E_{\mathrm{tor}}(\mathbb{Q})|$. Por último, para todo primo $q > 3$, podemos tomar $p \equiv 2$ (mód $3q$) y concluir que q no divide a $|E_{\mathrm{tor}}(\mathbb{Q})|$.

En definitiva, tenemos que $|E_{\mathrm{tor}}(\mathbb{Q})| \mid 6$. Claramente $E/\mathbb{Q}$ tiene un punto de orden 2 si y sólo si $X^3 + B$ tiene una raíz en $\mathbb{Q}$, si y sólo si B es un cubo. La cuestión es cuándo existe un punto de orden 3. Un punto P tiene orden 3 si y sólo si $2P = -P$, si y sólo si $x(2P) = x(P)$, pues esta última condición implica $2P = \pm P$, pero $2P = P$ sólo lo cumple $P = O$. La fórmula de duplicación nos da

$$\frac{x^4 - 8Bx}{4(x^3 + B)} = x.$$

Equivalentemente, $x^4 = -4Bx$. Una solución es $x = 0$, $y = B^2$. Así pues, si B es un cuadrado hay un elemento de orden 3. La otra posibilidad es $x^3 = -4B$, con lo que $y^2 = -3B$. En particular $B < 0$. Como B es libre de potencias sextas, sólo puede ser divisible entre los primos 2 y 3. Más concretamente, ha de ser $B = -2^4 \cdot 3^3$. En resumen, $E/\mathbb{Q}$ tiene un punto de orden 3 si y sólo si B es un cuadrado o $B = -2^4 \cdot 3^3$. A partir de aquí el teorema es inmediato. ∎

Observemos que la curva $Y^2 = X^3 - 2^4 \cdot 3^3 = X^3 - 432$ que aparece como caso excepcional en el teorema anterior es precisamente la curva que consideramos en el capítulo II en relación con el último teorema de Fermat para exponente 3. Es fácil ver que sus puntos de torsión son precisamente los puntos $(12, \pm 36)$ y O que allí llamábamos soluciones triviales.

7.3 El teorema débil de Mordell-Weil

Esta sección está dedicada a demostrar el teorema siguiente:

Teorema 7.20 (Teorema débil de Mordell-Weil) *Si K es un cuerpo numérico, E/K es una curva elíptica y $m \geq 2$ es un número natural, entonces el grupo $E(K)/mE(K)$ es finito.*

Se trata de un paso intermedio para demostrar uno de los resultados básicos sobre la aritmética de las curvas elípticas sobre cuerpos numéricos: que el grupo de puntos racionales es finitamente generado.

Cohomología continua Necesitaremos algunos resultados sobre cohomología de grupos. Llamaremos $\mathbb{A}$ a la clausura algebraica de $\mathbb{Q}$. El grupo de Galois $G(\mathbb{A}/K)$ es un grupo topológico con la topología de Krull, respecto a la cual una base de entornos del neutro la forman los subgrupos (normales) de índice finito.

Sea M un grupo abeliano sobre el que $G(\mathbb{A}/K)$ actúa de forma continua cuando en M consideramos la topología discreta, es decir, de modo que la acción

$$M \times G(\mathbb{A}/K) \longrightarrow M$$

sea una aplicación continua. Esto equivale a que las aplicaciones $\sigma \mapsto m^\sigma$ sean continuas en $G(\mathbb{A}/K)$ para cada $m \in M$ prefijado. En particular la antiimagen de $\{m\}$ ha de ser abierta, lo cual significa que existe una extensión finita normal L de K tal que $G(\mathbb{A}/L)$ fija a m. Así pues, si llamamos

$$M^L = \{m \in M \mid m^\sigma = m \text{ para todo } \sigma \in G(\mathbb{A}/L)\},$$

tenemos que M es la unión de los submódulos M^L. Recíprocamente, esta propiedad implica la continuidad de la acción, pues cada aplicación $\sigma \mapsto m^\sigma$ es constante en un entorno $G(\mathbb{A}/L)\sigma$ de cada automorfismo σ.

Vamos a considerar los grupos de cohomología de orden 0 y 1 de $G(\mathbb{A}/K)$ sobre M. Recordemos que el grupo de cohomología de orden 0 es simplemente $H^0(\mathbb{A}/K, M) = M^K$. El grupo $H^1(\mathbb{A}/K, M)$ es el cociente del grupo de *cociclos,* es decir, aplicaciones $a : G(\mathbb{A}/K) \longrightarrow M$ tales que

$$a_{\sigma\tau} = a_\sigma^\tau + a_\tau,$$

sobre el grupo de *cofronteras,* aplicaciones de la forma ∂m para $m \in M$ dadas por $(\partial m)_\sigma = m^\sigma - m$.

La continuidad de la acción de $G(\mathbb{A}/K)$ sobre M implica claramente que todas las cofronteras son continuas. Sin embargo, dentro del grupo de cociclos podemos considerar el subgrupo formado por los cociclos continuos, cuyo cociente sobre las cofronteras determina lo que llamaremos el grupo de *cohomología continua* $H^1_c(\mathbb{A}/K, M) \leq H^1(\mathbb{A}/K, M)$.

La ecuación de los cociclos implica que $a_1 = 0$, y si es continuo existe una extensión finita normal L de K tal que a toma el valor 0 sobre $G(\mathbb{A}/L)$, luego a es constante sobre cada clase $G(\mathbb{A}/L)\sigma$. En particular, a toma un número finito de valores, luego eligiendo L suficientemente grande podemos exigir que todos ellos estén en M^L, con lo que $a_\sigma = \tilde{a}_{\sigma|_L}$, para un cierto cociclo de $G(L/K)$ sobre M^L. Recíprocamente, todo cociclo de esta forma es continuo, pues es constante sobre cada clase $G(\mathbb{A}/L)\sigma$.

Para cada extensión finita normal L de K podemos considerar la *inflación*

$$\mathrm{Inf}_L : H^1(L/K, M^L) \longrightarrow H^1(\mathbb{A}/K, M)$$

dada por $\mathrm{Inf}_L([\tilde{a}]) = [a]$, donde $a_\sigma = \tilde{a}_{\sigma|_L}$.

En estos términos, hemos probado que $H^1_c(\mathbb{A}/K, M)$ es la unión de las imágenes de todas las inflaciones Inf_L. Es conocido que las inflaciones

$$\mathrm{Inf}_{L'/L} : H^1(L/K, M^L) \longrightarrow H^1(L'/K, M^{L'})$$

son inyectivas, luego las inflaciones Inf_L también lo son. Si las identificamos con inclusiones, podemos considerar que $H^1_c(\mathbb{A}/K, M)$ es la unión de los subgrupos $H^1(L/K, M)$ (más precisamente, es su límite inductivo).

Si L es una extensión finita normal de K, también podemos considerar la restricción $\mathrm{Res}_L : H^1(\mathbb{A}/K, M) \longrightarrow H^1(\mathbb{A}/L, M)$ definida restringiendo los cociclos de $G(\mathbb{A}/K)$ a $G(\mathbb{A}/L)$. Claramente se restringe a un homomorfismo

$$\mathrm{Res}_L : H^1_c(\mathbb{A}/K, M) \longrightarrow H^1_c(\mathbb{A}/L, M)$$

que conmuta con las restricciones

$$\mathrm{Res}_{L'/L} : H^1(L'/K, M^{L'}) \longrightarrow H^1(L'/L, M^L).$$

La exactitud de las sucesiones

$$1 \longrightarrow H^1(L/K, M^L) \xrightarrow{\mathrm{Inf}} H^1(L'/K, M) \xrightarrow{\mathrm{Res}} H^1(L'/L, M).$$

implica inmediatamente la de la sucesión

$$1 \longrightarrow H^1(L/K, M^L) \xrightarrow{\mathrm{Inf}} H^1_c(\mathbb{A}/K, M) \xrightarrow{\mathrm{Res}} H^1_c(\mathbb{A}/L, M).$$

Por último, sabemos que toda sucesión exacta de $G(\mathbb{A}/K)$-módulos

$$0 \longrightarrow S \xrightarrow{f} M \xrightarrow{g} N \longrightarrow 0$$

induce una sucesión exacta

$$0 \longrightarrow H^0(\mathbb{A}/K, S) \longrightarrow H^0(\mathbb{A}/K, M) \longrightarrow H^0(\mathbb{A}/K, N)$$

$$\xrightarrow{\delta} H^1(\mathbb{A}/K, S) \longrightarrow H^1(\mathbb{A}/K, M) \longrightarrow H^1(\mathbb{A}/K, N).$$

Vamos a ver que todos los homomorfismos se restringen a una sucesión exacta

$$0 \longrightarrow H^0(\mathbb{A}/K, S) \longrightarrow H^0(\mathbb{A}/K, M) \longrightarrow H^0(\mathbb{A}/K, N)$$

$$\xrightarrow{\delta} H^1_c(\mathbb{A}/K, S) \longrightarrow H^1_c(\mathbb{A}/K, M) \longrightarrow H^1_c(\mathbb{A}/K, N).$$

De hecho sólo hay que probar que la imagen de δ está en $H^1_c(\mathbb{A}/K, S)$, pues f y g son trivialmente continuas, luego transforman cociclos continuos en cociclos continuos (y la exactitud de la sucesión restringida es trivial).

Si $n \in N^K = H^0(\mathbb{A}/K, N)$, tomamos $m \in M$ tal que $g(m) = n$ y entonces $\delta(n) = [a_\sigma]$ está determinado por que $f(a_\sigma) = \partial m_\sigma = m^\sigma - m$. Es claro entonces que si $m \in M^L$ se cumple también que $\delta(n) \in H^1(L/K, M^L)$, pues si $\sigma|_L = \tau|_L$ entonces $f(a_\sigma) = f(a_\tau)$ y $a_\sigma = a_\tau$. Así pues, $\delta(n) \in H^1_c(\mathbb{A}/K, M)$.

A partir de aquí trabajaremos únicamente con los grupos de cohomología continua, por lo que suprimiremos el subíndice c.

El producto de Kummer Sea E/K una curva elíptica, $m \geq 2$ y consideremos la sucesión exacta

$$0 \longrightarrow E[m] \longrightarrow E(\mathbb{A}) \xrightarrow{m} E(\mathbb{A}) \longrightarrow 0$$

Sabemos que $G(\mathbb{A}/K)$ actúa sobre $E(\mathbb{A})$, y la acción es continua, pues para todo $P \in E(\mathbb{A})$ podemos tomar una extensión finita L de K tal que las coordenadas de P estén en L, y entonces $G(\mathbb{A}/L)$ fija a P. Por consiguiente podemos formar la sucesión exacta de cohomología

$$0 \longrightarrow E(K)[m] \longrightarrow E(K) \xrightarrow{m} E(K)$$

$$\xrightarrow{\delta} H^1(\mathbb{A}/K, E[m]) \longrightarrow H^1(\mathbb{A}/K, E(\mathbb{A})) \xrightarrow{m} H^1(\mathbb{A}/K, E(\mathbb{A})).$$

De aquí podemos extraer la llamada *sucesión exacta de Kummer* para E/K:

$$0 \longrightarrow E(K)/mE(K) \xrightarrow{\delta} H^1(\mathbb{A}/K, E[m]) \longrightarrow H^1(\mathbb{A}/K, E(\mathbb{A}))[m] \longrightarrow 0,$$

donde el último módulo es el conjunto de los elementos de $H^1(\mathbb{A}/K, E(\mathbb{A}))$ de orden divisor de m.

Nuestro objetivo es demostrar que el grupo $E(K)/mE(K)$ es finito. Tomemos una extensión finita normal L de K tal que $E[m] \subset E(L)$ y llamemos F al núcleo de la aplicación natural

$$E(K)/mE(K) \longrightarrow E(L)/mE(L).$$

Si demostramos que F es finito, bastará probar que $E(L)/mE(L)$ es también finito o, equivalentemente, podremos suponer que $E[m] \subset E(K)$. Consideramos el diagrama conmutativo con filas exactas

$$\begin{array}{ccccccc}
0 & \longrightarrow & F & \longrightarrow & E(K)/mE(K) & \longrightarrow & E(L)/mE(L) \\
 & & \downarrow & & \downarrow{\scriptstyle \delta} & & \downarrow{\scriptstyle \delta} \\
0 & \longrightarrow & H^1(L/K, E[m]) & \xrightarrow{\inf} & H^1(\mathbb{A}/K, E[m]) & \xrightarrow{\mathrm{res}} & H^1(\mathbb{A}/L, E[m])
\end{array}$$

Notemos que el segundo cuadrado es claramente conmutativo, lo que implica que δ hace corresponder el núcleo de la fila superior con el de la fila inferior, es decir, se restringe a un homomorfismo de F en $H^1(L/K, E[m])$. Como los grupos $G(L/K)$ y $E[m]$ son finitos, concluimos que $H^1(L/K, E[m])$ también lo es, al igual que F.

En lo que sigue supondremos que $E[m] \subset E(K)$. Esto significa que $G(\mathbb{A}/K)$ actúa trivialmente sobre $E[m]$, luego las cofronteras son nulas y los cociclos son homomorfismos de grupos. En definitiva,

$$H^1(\mathbb{A}/K, E[m]) = \mathrm{Hom}(G(\mathbb{A}/K), E[m]),$$

y tenemos una inclusión

$$\delta : E(K)/mE(K) \longrightarrow \mathrm{Hom}(G(\mathbb{A}/K), E[m]).$$

Definimos el *producto de Kummer*

$$\kappa : E(K) \times G(\mathbb{A}/K) \longrightarrow E[m]$$

mediante $\kappa(P,\sigma) = \delta([P])(\sigma)$. Claramente κ es una aplicación bilineal cuyo núcleo izquierdo es $mE(K)$. Vamos a calcular su núcleo derecho. Para ello observamos que $\kappa(P,\sigma) = Q^\sigma - Q$, donde $Q \in E(\mathbb{A})$ es cualquier punto que cumpla $mQ = P$. (Esto se sigue de la definición de los homomorfismos de conexión.)

Ahora es fácil ver que el núcleo derecho de κ es $G(\mathbb{A}/L)$, donde L es la adjunción a K de las coordenadas de todos los puntos $Q \in E(\mathbb{A})$ tales que $mQ \in E(K)$. En efecto, si $\sigma \in G(\mathbb{A}/K)$ cumple que $\kappa(P,\sigma) = 0$ para todo $P \in E(K)$, entonces, para todo $Q \in E(\mathbb{A})$ tal que $mQ \in E(K)$ tenemos que

$$0 = \kappa(mQ,\sigma) = Q^\sigma - Q,$$

luego $\sigma \in G(\mathbb{A}/L)$. La otra inclusión es análoga.

Observemos que L es una extensión normal de K (no sabemos si finita), pues $\sigma[L] = L$ para todo $\sigma \in G(\mathbb{A}/K)$. Esto nos permite concluir que el producto de Kummer induce una forma bilineal regular

$$\kappa : E(K)/mE(K) \times G(L/K) \longrightarrow E[m].$$

Si probamos que la extensión L/K es finita, podremos concluir que el grupo $E(K)/mE(K)$ también lo es.

La extensión L/K Lo primero que podemos decir de la extensión L/K es que es abeliana de exponente m (es decir, todos los elementos de $G(L/K)$ tienen orden divisor de m). En efecto, basta tener en cuenta que el producto de Kummer induce un monomorfismo de grupos

$$G(L/K) \longrightarrow \mathrm{Hom}(E(K), E[m]).$$

Llamemos S al conjunto (finito) de los divisores primos $\mathfrak{p}$ de K que cumplan alguna de las condiciones siguientes:

a) E tiene mala reducción en $\mathfrak{p}$,

b) $v_\mathfrak{p}(m) \neq 0$,

c) $\mathfrak{p}$ es arquimediano.

Vamos a demostrar que si $\mathfrak{p} \notin S$, entonces la extensión L/K es no ramificada en $\mathfrak{p}$, es decir, que $\mathfrak{p}$ no se ramifica en ninguna extensión finita intermedia K'.

Como el producto de extensiones no ramificadas es no ramificada, no perdemos generalidad si suponemos que K' es la adjunción a K de las coordenadas de un punto $Q \in E(\mathbb{A})$ tal que $P = mQ \in E(K)$.

Sea $\mathfrak{P}$ un divisor de $\mathfrak{p}$ en K'. Hemos de probar que la extensión $K'_{\mathfrak{P}}/K_{\mathfrak{p}}$ es no ramificada. El índice de ramificación de la extensión es el orden del grupo de inercia de $\mathfrak{P}$ (el grupo de los $K_{\mathfrak{p}}$-automorfismos de $K'_{\mathfrak{P}}$ que inducen la identidad sobre el cuerpo de restos $k'_{\mathfrak{P}}$), luego basta ver que este grupo es trivial.

Como $E/K_{\mathfrak{p}}$ tiene buena reducción en $\mathfrak{p}$, tenemos que $E/K'_{\mathfrak{P}}$ tiene buena reducción en $\mathfrak{P}$ (podemos tomar la misma ecuación de Weierstrass), luego podemos considerar la reducción $E(K') \longrightarrow E(k'_{\mathfrak{P}})$. Por otra parte, como $v_{\mathfrak{p}}(m) = 0$ y E tiene buena reducción módulo $\mathfrak{p}$, el teorema 6.19 nos da que la reducción $E(K)[m] \longrightarrow E(k_{\mathfrak{p}})$ es inyectiva.

Tomemos σ en el grupo de inercia. Esto significa que $\tilde{Q}^\sigma = \tilde{Q}$, luego la reducción de $Q^\sigma - Q$ en $k'_{\mathfrak{P}}$ es nula.

Pero $m(Q^\sigma - Q) = (mQ)^\sigma - mQ = P^\sigma - P = 0$, pues $P \in E(K)$. Concluimos que $Q^\sigma - Q \in E[m] \subset E(K)$, luego $Q^\sigma - Q \in E(K)[m]$ y tiene reducción nula en $E(k'_{\mathfrak{P}})$, luego también en $E(k_{\mathfrak{p}})$. Por la inyectividad de la reducción tenemos que $Q^\sigma - Q = 0$.

Acabamos de probar que el grupo de inercia deja fijo a Q, luego deja fijas a sus coordenadas y, por consiguiente, a todo $K'_{\mathfrak{P}}$. En suma, dicho grupo de inercia es trivial.

Extensiones de Kummer El teorema débil de Mordell-Weil quedará probado si demostramos que toda extensión abeliana de exponente m de K no ramificada fuera de un conjunto finito S de primos de K que contenga al menos a los primos no arquimedianos es finita. (Pues la extensión L/K cumple estas propiedades.)

En primer lugar observamos que si el resultado es cierto para una extensión finita K' de K, entonces también es cierto para K, pues si L es una extensión abeliana de K de exponente m no ramificada fuera de un conjunto finito S de primos, entonces LK'/K' es una extensión abeliana de exponente m de K' no ramificada fuera del conjunto S' formado por los divisores de los primos de S en K'. Así pues, LK'/K' es finita, LK'/K también y L/K también. Por consiguiente podemos extender el cuerpo K y suponer que contiene las raíces m-simas de la unidad.

En segundo lugar, tampoco perdemos generalidad si extendemos el conjunto de primos S. Podemos hacerlo de tal modo que $v_{\mathfrak{p}}(m) = 0$ para todo primo $\mathfrak{p} \notin S$. Por último, el teorema 7.12 nos permite suponer además que la localización K_S es un dominio de ideales principales.

El grupo de unidades de K_S es

$$U_S = \{\alpha \in K \mid v_{\mathfrak{p}}(\alpha) = 0 \text{ para todo } \mathfrak{p} \notin S\}.$$

El teorema de las unidades de Hasse afirma que U_S es un $\mathbb{Z}$-módulo finitamente generado.

Si un cuerpo K tiene característica 0 y contiene a la raíces m-simas de la unidad, las extensiones abelianas de K de exponente m se llaman *extensiones de Kummer* de K y se obtienen adjuntando a K raíces m-simas de elementos de K^*. Además, una extensión $K(\sqrt[m]{a})$ es no ramificada en un primo $\mathfrak{p}$ que no divida a m si y sólo si $a = b^m u$, donde b, $u \in K^*$, $v_\mathfrak{p}(u) = 0$.

Equivalentemente, $K(\sqrt[m]{a})/K$ es no ramificada en un primo $\mathfrak{p} \notin S$ si y sólo si $m \mid v_\mathfrak{p}(a)$. En efecto, una implicación es obvia y si $v_\mathfrak{p}(a) = mr$, tomamos $\pi \in K$ tal que $v_\mathfrak{p}(\pi) = 1$, $b = \pi^r$, $u = a/b^m$. Así $a = b^m u$ tiene la forma requerida para que la extensión sea no ramificada.

Teniendo en cuenta que si dos elementos de K^* se diferencian en un factor de K^{*m} sus raíces m-simas determinan la misma extensión, el teorema quedará demostrado si probamos la finitud del grupo

$$K(S,m) = \big\{[\alpha] \in K^*/K^{*m} \;\big|\; m \mid v_\mathfrak{p}(\alpha) \text{ para todo } \mathfrak{p} \notin S\big\}.$$

Veamos que la aplicación natural $U_S \longrightarrow K(S,m)$ es suprayectiva. En efecto, si $[\alpha] \in K(S,m)$, entonces el ideal fraccional (α) en K_S es una potencia m-sima. Como K_S es un dominio de ideales principales, existe $\beta \in K^*$ tal que $(\alpha) = (\beta)^m$, luego existe $u \in U_S$ tal que $\alpha = u\beta^m$ y $[\alpha] = [u]$.

Claramente, tenemos un epimorfismo $U_S/U_S^m \longrightarrow K(S,m)$, y como U_S es finitamente generado, U_S/U_S^m es un grupo finito y $K(S,m)$ también. ∎

7.4 Alturas

Acabamos de probar que si K es un cuerpo numérico y E/K es una curva elíptica, entonces los grupos $E(K)/mE(K)$ son finitamente generados, lo cual es una condición necesaria pero no suficiente para que el grupo $E(K)$ sea finitamente generado. El teorema siguiente contiene lo que nos falta para llegar a esta conclusión:

Teorema 7.21 *Sea G un grupo abeliano en el que hay definida una función $a : G \longrightarrow \mathbb{R}$ que cumpla las tres propiedades siguientes:*

a) Para cada $Q \in G$ existe una constante $C_1 = C_1(Q)$ tal que para todo $P \in G$ se cumple $a(P+Q) \le 2a(P) + C_1$.

b) Existen un natural $m \ge 2$ y una constante C_2 tales que para todo $P \in G$ se cumple $a(mP) \ge m^2 a(P) - C_2$.

c) Para toda constante C_3, el conjunto $\{P \in G \mid a(P) \le C_3\}$ es finito.

Si, para el natural m considerado en b), el grupo G/mG es finito, entonces el grupo G es finitamente generado.

DEMOSTRACIÓN: Tomemos representantes $Q_1, \ldots, Q_r$ de las clases del grupo G/mG. Sea $P \in G$ un elemento arbitrario. Entonces $P = mP_1 + Q_{i_1}$, para cierto $P_1 \in G$ y cierto índice i_1. Similarmente, $P_1 = mP_2 + Q_{i_2}$ y, en general, obtenemos una sucesión de puntos determinada recurrentemente por la relación $P_{n-1} = mP_n + Q_{i_n}$. Por las propiedades de la función a tenemos que

$$a(P_j) \le \frac{1}{m^2}(a(mP_j) + C_2) = \frac{1}{m^2}(a(P_{j-1} - Q_{i_j}) + C_2)$$

$$\le \frac{1}{m^2}(2a(P_{j-1}) + C_1' + C_2),$$

donde C_1' es la mayor de las constantes C_1 dadas por la hipótesis a) para los elementos $Q = -Q_i$. Notemos que las constantes C_1' y C_2 no depende de P.

Ahora aplicamos esta desigualdad a un P_n, lo que nos acota $a(P_n)$ en términos de $a(P_{n-1})$. A su vez, acotamos $a(P_{n-1})$ en términos de $a(P_{n-2})$ hasta llegar a P. El resultado es:

$$a(P_n) \le \left(\frac{2}{m^2}\right)^n a(P) + \left(\frac{1}{m^2} + \frac{2}{m^4} + \cdots + \frac{2^{n-1}}{m^{2n}}\right)(C_1' + C_2)$$

$$< \left(\frac{2}{m^2}\right)^n a(P) + \frac{C_1' + C_2}{m^2 - 2} \le 2^{-n} a(P) + (C_1' + C_2)/2.$$

Para n suficientemente grande se cumple que $a(P_n) \le 1 + (C_1' + C_2)/2$. Además,

$$P = m^n P_n + \sum_{j=1}^{n} m^{j-1} Q_{i_j},$$

luego hemos probado que un generador de G es

$$\{Q_1, \ldots, Q_r\} \cup \{Q \in G \mid a(Q) \le 1 + (C_1' + C_2)/2\},$$

que es un conjunto finito por la hipótesis c). ∎

A continuación vamos a definir una función sobre el espacio proyectivo $\mathrm{P}^N(K)$ cuya restricción a las curvas elípticas nos permitirá aplicar el teorema anterior.

Recordemos que el *valor absoluto canónico* asociado a un primo $p \in \mathbb{Z}$ es el valor absoluto de $\mathbb{Q}$ dado por $|n|_p = 1/p^{v_p(n)}$, mientras que el valor absoluto canónico asociado a ∞ es el valor absoluto usual en $\mathbb{Q}$, que representaremos por $|\ |_\infty$. Similarmente, el valor absoluto canónico asociado a un divisor primo $\mathfrak{p}$ de un cuerpo numérico K es el único valor absoluto asociado a $\mathfrak{p}$ que extiende a un valor absoluto canónico de $\mathbb{Q}$. Lo representaremos por $|\ |_\mathfrak{p}$.

Si $\mathfrak{p}$ es un divisor primo (arquimediano o no) en un cuerpo numérico K, y p es el divisor primo racional al cual divide, representaremos por

$$n_\mathfrak{p} = |K_\mathfrak{p} : \mathbb{Q}_p|$$

al *grado local* de K en $\mathfrak{p}$. Si L/K es una extensión de cuerpos numéricos y $\mathfrak{p}$ es un divisor primo en K, entonces

$$\sum_{\mathfrak{P}|\mathfrak{p}} n_{\mathfrak{P}} = \sum_{\mathfrak{P}|\mathfrak{p}} n(\mathfrak{P}/\mathfrak{p}) n_{\mathfrak{p}} = |L:K| n_{\mathfrak{p}}, \tag{7.4}$$

donde $n(\mathfrak{P}/\mathfrak{p}) = |L_{\mathfrak{P}} : K_{\mathfrak{p}}|$. Así mismo tenemos la *fórmula del producto:*

Para todo $x \in K^*$ se cumple la relación

$$\prod_{\mathfrak{p}} |x|_{\mathfrak{p}}^{n_{\mathfrak{p}}} = 1,$$

donde $\mathfrak{p}$ recorre todos los divisores primos de K, incluidos los arquimedianos.

Definición 7.22 Sea K un cuerpo numérico y $P \in \mathrm{P}^N(K)$ un punto con coordenadas homogéneas $P = [x_0, \ldots, x_N]$, $x_i \in K$. Definimos la *altura* de P en K como

$$A_K(P) = \prod_{\mathfrak{p}} \max\{|x_0|_{\mathfrak{p}}^{n_{\mathfrak{p}}}, \ldots, |x_N|_{\mathfrak{p}}^{n_{\mathfrak{p}}}\}.$$

La fórmula del producto implica claramente que esta definición no depende de la elección de las coordenadas homogéneas del punto P. Otro hecho elemental es que $A_K(P) \geq 1$. En efecto, siempre podemos elegir un vector de coordenadas homogéneas con una coordenada igual a 1, con lo que todos los factores que aparecen en la definición de $A_K(P)$ son como mínimo 1.

Observemos que todo punto de $\mathrm{P}^N(\mathbb{Q})$ se expresa en forma única como $P = [x_0, \ldots, x_N]$, donde los x_i son enteros primos entre sí. Entonces, para todo primo arquimediano p, el factor correspondiente en la definición de altura vale 1, luego

$$A_{\mathbb{Q}}(P) = \max\{|x_0|_\infty, \ldots, |x_N|_\infty\}.$$

La definición de altura para cuerpos numéricos arbitrarios es una forma de generalizar esta definición teniendo en cuenta que los anillos de enteros algebraicos no tienen necesariamente factorización única, por lo que en el caso general no podemos hablar de coordenadas enteras primas entre sí.

De la fórmula (7.4) se sigue fácilmente que si L/K es una extensión de cuerpos numéricos y $P \in \mathrm{P}^N(K)$, entonces

$$A_L(P) = A_K(P)^{|L:K|}.$$

Esto hace que podamos definir la *altura absoluta* de un punto P como

$$A(P) = A_K(P)^{1/|K:\mathbb{Q}|}$$

(tomando la raíz positiva), de modo que $A(P)$ es independiente de K. Seguidamente demostraremos varios resultados técnicos sobre las alturas que acabamos de definir.

Definición 7.23 Una aplicación regular de *grado d* entre espacios proyectivos es una aplicación $\phi : \mathrm{P}^N(\mathbb{A}) \longrightarrow \mathrm{P}^M(\mathbb{A})$ tal que existen polinomios homogéneos $F_0, \ldots, F_M$ de grado d que no se anulan simultáneamente en ningún punto de $\mathbb{A}^{N+1}$ distinto de O y para todo $P \in \mathrm{P}^N(\mathbb{A})$ se cumple

$$\phi(P) = [F_0(P), \ldots, F_M(P)].$$

Observemos que toda aplicación regular entre espacios proyectivos es de esta forma para ciertos polinomios unívocamente determinados salvo una constante, pues si pudiéramos expresar también $\phi(P) = [G_0(P), \ldots, G_M(P)]$, tendríamos que $\alpha = F_i/G_i$ sería independiente de i. Si expresamos $\alpha = U/V$, con U y V primos entre sí, tenemos que $VF_i = UG_i$, luego $V \mid G_i$ y $U \mid F_i$ para todo i, luego U y V son constantes (o los F_i y G_i tendrían ceros comunes).

Así pues, lo que hemos definido es el grado de una aplicación regular entre espacios proyectivos.

Para el caso de una aplicación $\phi : \mathrm{P}^1(\mathbb{A}) \longrightarrow \mathrm{P}^1(\mathbb{A})$ este grado coincide con el que ya teníamos definido para aplicaciones entre curvas. En efecto, el cuerpo de funciones racionales de $\mathrm{P}^1(\mathbb{A})$ es $\mathbb{A}(X)$, y su imagen por $\overline{\phi}$ es $\mathbb{A}(\phi \circ X)$. Si identificamos a $\alpha = \phi \circ X$ con una función racional en $A^1(\mathbb{A})$, tenemos que $\alpha = F_{1*}(X)/F_{0*}(X)$, donde los polinomios $F_{i*}(X) = F_i(X, 1)$ son primos entre sí y el de mayor grado tiene grado d. Hemos de probar que $|\mathbb{A}(X) : \mathbb{A}(\alpha)| = d$. Cambiando α por $1/\alpha$ si es preciso podemos suponer que $\operatorname{grad} F_{1*} = d$. Entonces $F_{1*}(T) - \alpha F_{0*}(T) \in \mathbb{A}(\alpha)[T]$ es un polinomio de grado d que anula a X. Además es irreducible, pues si se descompusiera en producto de dos factores, uno de ellos tendría grado 0 en α y dividiría a F_{1*} y F_{2*}, luego sería constante.

Teorema 7.24 *Sea $\phi : \mathrm{P}^N(\mathbb{A}) \longrightarrow \mathrm{P}^M(\mathbb{A})$ una aplicación regular de grado d. Entonces existen constantes C_1, $C_2 > 0$ tales que para todo $P \in \mathrm{P}^N(\mathbb{A})$ se cumple*

$$C_1 A(P)^d \leq A(\phi(P)) \leq C_2 A(P)^d.$$

Demostración: Sea $\phi = [F_0, \ldots, F_M]$ y $P = [x_0, \ldots, x_N]$. Tomemos un cuerpo numérico K que contenga a todos los x_i y a todos los coeficientes de los polinomios F_i. Para cada divisor primo $\mathfrak{p}$ de K llamamos $|\phi|_\mathfrak{p}$ al máximo de los valores absolutos respecto a $\mathfrak{p}$ de los coeficientes de los F_i. Definimos también

$$|P|_\mathfrak{p} = \max_{0 \leq i \leq N} |x_i|_\mathfrak{p}, \qquad |\phi(P)|_\mathfrak{p} = \max_{0 \leq j \leq M} |F_j(P)|_\mathfrak{p}.$$

Entonces

$$A_K(P) = \prod_\mathfrak{p} |P|_\mathfrak{p}^{n_\mathfrak{p}}, \qquad A_K(\phi(P)) = \prod_\mathfrak{p} |\phi(P)|_\mathfrak{p}^{n_\mathfrak{p}}.$$

Definimos $A_K(\phi) = \prod_\mathfrak{p} |\phi|_\mathfrak{p}^{n_\mathfrak{p}}$. Si hacemos $\epsilon(\mathfrak{p}) = 1$ si $\mathfrak{p}$ es arquimediano y $\epsilon(\mathfrak{p}) = 0$ en caso contrario, la desigualdad triangular para $\mathfrak{p}$ se puede expresar como

$$|t_1 + \cdots + t_n|_\mathfrak{p} \leq n^{\epsilon(\mathfrak{p})} \max\{|t_1|_\mathfrak{p}, \ldots, |t_n|_\mathfrak{p}\}.$$

Las letras C_1, $C_2, \ldots$ representarán constantes independientes de P. La desigualdad triangular nos da que

$$|F_i(P)|_\mathfrak{p} \leq C_1^{\epsilon(\mathfrak{p})}|\phi|_\mathfrak{p}|P|_\mathfrak{p}^d,$$

donde la constante C_1 es el número de monomios de grado d en $N+1$ variables con coeficiente 1. Tomando el máximo sobre i vemos que

$$|\phi(P)|_\mathfrak{p} \leq C_1^{\epsilon(\mathfrak{p})}|\phi|_\mathfrak{p}|P|_\mathfrak{p}^d.$$

Ahora elevamos a $n_\mathfrak{p}$, multiplicamos para todo $\mathfrak{p}$ y elevamos a $|K:\mathbb{Q}|^{-1}$, con lo que queda

$$A(\phi(P)) \leq C_1 A(\phi)A(P)^d,$$

pues

$$\sum_\mathfrak{p} \epsilon(\mathfrak{p})n_\mathfrak{p} = \sum_{\mathfrak{p}|\infty} n_\mathfrak{p} = |K:\mathbb{Q}|.$$

Con esto tenemos probada la mitad del teorema. Estamos suponiendo que las formas F_i no tienen ceros comunes en $\mathbb{A}^{N+1}$ salvo el punto $(0,\ldots,0)$. En otros términos, que

$$V(F_0,\ldots,F_M) = \varnothing.$$

Por el teorema de los ceros de Hilbert, esto implica que el ideal $(F_0,\ldots,F_M)$ contiene todas las formas de grado e, para un cierto e suficientemente grande. En particular, existen polinomios $G_{ij} \in \mathbb{A}[X_0,\ldots,X_N]$ tales que

$$X_i^e = \sum_{j=0}^M G_{ij}F_j.$$

Eliminando términos, podemos suponer que los G_{ij} son formas de grado $e-d$. Extendiendo K si es necesario, podemos suponer también que todos los polinomios G_{ij} tienen sus coeficientes en K. Llamamos $|G|_\mathfrak{p}$ al máximo valor absoluto respecto de $\mathfrak{p}$ de un coeficiente de algún G_{ij} y

$$A_K(G) = \prod_\mathfrak{p} |G|_\mathfrak{p}^{n_\mathfrak{p}}.$$

Ahora, las coordenadas de P cumplen

$$|x_i|_\mathfrak{p}^e = \left|\sum_{j=0}^M G_{ij}(P)F_j(P)\right|_\mathfrak{p} \leq C_2^{\epsilon(\mathfrak{p})} \max_{0\leq j\leq M} |G_{ij}(P)F_j(P)|_\mathfrak{p}.$$

Tomando el máximo sobre i llegamos a que

$$|P|_\mathfrak{p}^e \leq C_2^{\epsilon(\mathfrak{p})} \max_{i,j} |G_{ij}(P)|_\mathfrak{p}|\phi(P)|_\mathfrak{p}.$$

Como cada G_{ij} tiene grado $e-d$, el número de monomios que lo componen está acotado independientemente de P, luego

$$|G_{ij}(P)|_\mathfrak{p} \leq C_3^{\epsilon(\mathfrak{p})}|G|_\mathfrak{p}|P|_\mathfrak{p}^{e-d}.$$

Sustituimos arriba y multiplicamos por $|P|_\mathfrak{p}^{d-e}$:

$$|P|_\mathfrak{p}^d \le C_4^{\epsilon(\mathfrak{p})}|G|_\mathfrak{p}|\phi(P)|_\mathfrak{p}.$$

Por último, elevamos a $n_\mathfrak{p}$, multiplicamos para todo $\mathfrak{p}$ y elevamos a $|K:\mathbb{Q}|^{-1}$, con lo que obtenemos la desigualdad buscada. ∎

Si $\alpha \in \mathbb{A}$, llamaremos $A(\alpha) = A([\alpha, 1])$. El teorema siguiente relaciona la altura de los coeficientes de un polinomio con la altura de sus raíces:

Teorema 7.25 *Sea*

$$F(T) = a_0T^d + a_1T^{d-1} + \cdots + a_d = a_0(T-\alpha_1)\cdots(T-\alpha_d) \in \mathbb{A}[T]$$

un polinomio de grado d. Entonces

$$2^{-d}\prod_{j=1}^d A(\alpha_j) \le A([a_0,\ldots,a_d]) \le 2^{d-1}\prod_{j=1}^d A(\alpha_j).$$

Demostración: Las desigualdades no se alteran si dividimos F entre a_0, luego podemos suponer que $a_0 = 1$. Sea $K = \mathbb{Q}(\alpha_1,\ldots,\alpha_d)$. Para cada divisor primo $\mathfrak{p}$ de K definimos $\epsilon(\mathfrak{p}) = 2$ si $\mathfrak{p}$ es arquimediano y $\epsilon(\mathfrak{p}) = 1$ en caso contrario. De este modo la desigualdad triangular es

$$|x+y|_\mathfrak{p} \le \epsilon(\mathfrak{p})\,\text{máx}\{|x|_\mathfrak{p}, |y|_\mathfrak{p}\}.$$

Vamos a demostrar que

$$\epsilon(\mathfrak{p})^{-d}\prod_{j=1}^d \text{máx}\{|\alpha_j|_\mathfrak{p}, 1\} \le \max_{0\le i\le d}|a_i|_\mathfrak{p} \le \epsilon(\mathfrak{p})^{d-1}\prod_{j=1}^d \text{máx}\{|\alpha_j|_\mathfrak{p}, 1\}.$$

Admitiendo esto, basta elevar a $n_\mathfrak{p}$, multiplicar $\mathfrak{p}$ y elevar a $|K:\mathbb{Q}|^{-1}$ y obtenemos las desigualdades del enunciado. Razonamos por inducción sobre d.

Si $d = 1$ tenemos que $F(T) = T - \alpha_1$ y todo es trivial. Suponemos el resultado cierto para polinomios de grado $d-1$ con raíces en K. No perdemos generalidad si suponemos que

$$|\alpha_d|_\mathfrak{p} = \min_{0\le j\le d}|\alpha_j|_\mathfrak{p}.$$

Definimos $G(T) = (T-\alpha_1)\cdots(T-\alpha_{d-1}) = T^{d-1} + b_1T^{d-2} + \cdots + b_{d-1}$, de modo que $F(T) = (T-\alpha_d)G(T)$.

Comparando los coeficientes vemos que $a_i = b_i - \alpha_d b_{i-1}$. Esta relación vale para $0 \le i \le d$ si convenimos en que $b_{-1} = b_d = 0$. Ahora tenemos que

$$\max_{0\le i\le d}|a_i|_\mathfrak{p} = \max_{0\le i\le d}|b_i - \alpha_d b_{i-1}|_\mathfrak{p} \le \epsilon(\mathfrak{p})\max_{0\le i\le d}\{|b_i|_\mathfrak{p}, |\alpha_d b_{i-1}|_\mathfrak{p}\}$$

$$\le \epsilon(\mathfrak{p})\max_{0\le i\le d}|b_i|_\mathfrak{p}\,\text{máx}\{|\alpha_d|_\mathfrak{p}, 1\} \le \epsilon(\mathfrak{p})^{d-1}\prod_{j=1}^d \text{máx}\{|\alpha_j|_\mathfrak{p}, 1\},$$

donde en la última desigualdad hemos aplicado la hipótesis de inducción.

Para la otra desigualdad distinguimos dos casos: Si $|\alpha_d|_\mathfrak{p} \le \epsilon(\mathfrak{p})$, entonces

$$\prod_{j=1}^{d} \text{máx}\{|\alpha_j|_\mathfrak{p}, 1\} \le \text{máx}\{|\alpha_d|_\mathfrak{p}, 1\}^d \le \epsilon(\mathfrak{p})^d$$

y el resultado es claro, pues $|a_0|_\mathfrak{p} = 1$. Supongamos ahora que $|\alpha_d|_\mathfrak{p} > \epsilon(\mathfrak{p})$ y sea j el índice para el que $|b_{j-1}|_\mathfrak{p}$ es máximo. Entonces

$$\max_{0\le i\le d} |\alpha_i|_\mathfrak{p} = \max_{0\le i\le d} |b_i - \alpha_d b_{i-1}|_\mathfrak{p} \ge |b_j - \alpha_d b_{j-1}|_\mathfrak{p}.$$

Si $\mathfrak{p}$ no es arquimediano el último término es

$$|\alpha_d|_\mathfrak{p}|b_{j-1}|_\mathfrak{p} = \epsilon(\mathfrak{p})^{-1} \max_{0\le i\le d-1} |b_i|_\mathfrak{p}|\alpha_d|_\mathfrak{p},$$

y si es arquimediano entonces

$$|b_j - \alpha_d b_{j-1}|_\mathfrak{p} \ge |\alpha_d|_\mathfrak{p}|b_{j-1}|_\mathfrak{p} - |b_j|_\mathfrak{p} \ge (|\alpha_d|_\mathfrak{p} - 1)|b_{j-1}|_\mathfrak{p}$$

$$= (|\alpha_d|_\mathfrak{p} - 1) \max_{0\le i\le d-1} |b_i|_\mathfrak{p} > \epsilon(\mathfrak{p})^{-1}|\alpha_d|_\mathfrak{p} \max_{0\le i\le d-1} |b_i|_\mathfrak{p},$$

pues $|\alpha_d|_\mathfrak{p} > 2$.

En cualquier caso tenemos que

$$\max_{0\le i\le d} |\alpha_i|_\mathfrak{p} > \epsilon(\mathfrak{p})^{-1} \max_{0\le i\le d-1} |b_i|_\mathfrak{p}\, \text{máx}\{|\alpha_d|_\mathfrak{p}, 1\},$$

y basta aplicar la hipótesis de inducción. ∎

Finalmente podemos probar:

Teorema 7.26 *Si K es un cuerpo numérico y $C > 0$ una constante, entonces el conjunto*

$$\{P \in \mathrm{P}^N(K) \mid A_K(P) \le C\}$$

es finito.

Demostración: Teniendo en cuenta la relación $A_L(P) = A_K(P)^{|L:K|}$, es claro que podemos sustituir K por una extensión, luego podemos suponer que K es una extensión normal de $\mathbb{Q}$, digamos de grado n.

Sea $P = [x_0, \dots, x_N]$. Podemos suponer que algún $x_i = 1$. Entonces

$$A_K(P) = \prod_\mathfrak{p} \max_{0\le i\le N} |x_i|_\mathfrak{p}^{n_\mathfrak{p}} \ge \max_{0\le i\le N} \prod_\mathfrak{p} \text{máx}\{|x_i|_\mathfrak{p}, 1\}^{n_\mathfrak{p}} = \max_{0\le i\le N} A_K(x_i).$$

Así, todo punto en el conjunto del enunciado tiene un vector de coordenadas homogéneas cuyas componentes están en el conjunto

$$\{x \in K \mid A_K(x) \le C\}.$$

Si probamos que este conjunto es finito, el teorema estará demostrado. (Equivalentemente, hemos reducido el teorema al caso $N = 1$.)

Sea $x \in K$ tal que $A_K(x) \leq C$ y sea

$$F_x(T) = (T - x_1) \cdots (T - x_d) = T^d + a_1 T^{d-1} + \cdots + a_d$$

su polinomio mínimo sobre $\mathbb{Q}$ (de modo que $d \leq n = |K : \mathbb{Q}|$). Entonces, por el teorema anterior,

$$A([1, a_1, \ldots, a_d]) \leq 2^{d-1} \prod_{j=1}^{d} A(x_j).$$

Los x_j son conjugados de x, y es fácil ver que los automorfismos conservan la altura (porque permutan los divisores primos), luego

$$A([1, a_1, \ldots, a_d]) \leq 2^{d-1} A(x)^d \leq (2C)^{d/n}.$$

Si probamos que el conjunto

$$\{P \in \mathrm{P}^d(\mathbb{Q}) \mid A(P) \leq C\}$$

es finito, entonces para cada $d \leq n$ habrá una cantidad finita de polinomios F_x posibles, luego habrá una cantidad finita de números x posibles. Con esto hemos reducido el teorema al caso $K = \mathbb{Q}$ y por la reducción al caso $N = 1$ en realidad basta probar la finitud del conjunto

$$\{x \in \mathbb{Q} \mid A(x) \leq C\}.$$

Ahora bien, si x está en este conjunto entonces $|x|_p \leq C$ para todo primo p de $\mathbb{Q}$, luego $v_p(x) \geq -\log C / \log p$ para todo primo no arquimediano. Tomando p suficientemente grande para que $-\log C / \log p > -1$ concluimos que $v_p(x) \geq 0$ para todo primo p suficientemente grande (independientemente de x). Así pues, existe un $m \in \mathbb{Z}$ no nulo tal que $mx \in \mathbb{Z}$ para todo x del conjunto. En definitiva

$$\{x \in \mathbb{Q} \mid A(x) \leq C\} \subset (1/m)\mathbb{Z}$$

y está acotado respecto del valor absoluto usual, luego es un conjunto finito. ∎

7.5 El teorema de Mordell-Weil

Ahora ya tenemos los elementos necesarios para demostrar el teorema de Mordell-Weil, según el cual el grupo $E(K)$ es finitamente generado. Lo que hemos de hacer es definir una altura en una curva elíptica que cumpla las propiedades del teorema 7.21. Para ello usaremos la altura en $\mathrm{P}^N(\mathbb{A})$ que hemos estudiado en la sección anterior combinada con una función racional de la curva. Conviene además tomar logaritmos:

Definición 7.27 La *altura (absoluta) logarítmica* en $\mathrm{P}^N(\mathbb{A})$ es la aplicación $a : \mathrm{P}^N(\mathbb{A}) \longrightarrow \mathbb{R}$ dada por $a(P) = \log A(P)$.

Como sabíamos que $A(P) \geq 1$, ahora tenemos que $a(P) \geq 0$.

Si K es un cuerpo numérico, E/K es una curva elíptica y $f \in \mathbb{A}(E)$, la *altura* en E relativa a f es la función $a_f : E(\mathbb{A}) \longrightarrow \mathbb{R}$ dada por $a_f(P) = a(f(P))$.

El teorema 7.26 puede reformularse como sigue:

Teorema 7.28 *Si K es un cuerpo numérico, E/K es una curva elíptica y $f \in K(E)$ es una función no constante, entonces para todo $C > 0$ el conjunto*

$$\{P \in E(K) \mid a_f(P) \leq C\}$$

es finito.

Demostración: Si P está en el conjunto del enunciado, entonces $f(P)$ está en el conjunto

$$\{Q \in \mathrm{P}^1(K) \mid A(Q) \leq e^C\}.$$

Este conjunto es finito por 7.26 y cada uno de sus puntos tiene a lo sumo un número finito de antiimágenes por f. Así pues, el conjunto del enunciado es finito. ∎

El teorema siguiente nos da la relación fundamental entre las alturas en las curvas elípticas y su estructura de grupo:

Teorema 7.29 *Sea K un cuerpo numérico, sea E/K una curva elíptica y sea $f \in K(E)$ una función racional par no constante (es decir, que verifique $f(-P) = f(P)$ para todo $P \in E(\mathbb{A})$). Entonces, para todo par de puntos P, $Q \in E(\mathbb{A})$ se cumple*

$$a_f(P+Q) + a_f(P-Q) = 2a_f(P) + 2a_f(Q) + O(1).$$

Demostración: Consideremos una ecuación de Weierstrass para E de la forma

$$Y^2 = X^3 + AX + B.$$

Probaremos primero el teorema para $f = x$. Ciertamente es una función par. Si P y Q son puntos finitos de E tales que $P \neq \pm Q$, las fórmulas del teorema 2.21 nos dan que

$$x(P+Q) = \left(\frac{y(Q)-y(P)}{x(Q)-x(P)}\right)^2 - 2x(P) - 2x(Q),$$

$$x(P-Q) = \left(\frac{-y(Q)-y(P)}{x(Q)-x(P)}\right)^2 - 2x(P) - 2x(Q),$$

y operando se llega a que

$$x(P+Q) + x(P-Q) = \frac{2(x(P)+x(Q))(A+x(P)x(Q)) + 4B}{(x(P)+x(Q))^2 - 4x(P)x(Q)},$$

$$x(P)x(Q) = \frac{(x(P)x(Q) - A)^2 - 4B(x(P) + x(Q)}{(x(P) + x(Q))^2 - 4x(P)x(Q)}.$$

Con estas fórmulas se prueba que el diagrama siguiente es conmutativo:

$$\begin{array}{ccc} E \times E & \xrightarrow{\phi} & E \times E \\ \downarrow & & \downarrow \\ \mathrm{P}^1 \times \mathrm{P}^1 & & \mathrm{P}^1 \times \mathrm{P}^1 \\ \downarrow & & \downarrow \\ \mathrm{P}^2 & \xrightarrow[\psi]{} & \mathrm{P}^2 \end{array}$$

donde $\phi(P, Q) = (P + Q, P - Q)$,

$$\psi([t, u, v]) = [u^2 - 4tv, 2u(At + v) + 4Bt^2, (v - At)^2 - 4Btu],$$

y las flechas verticales son $(P, Q) \mapsto (x(P), x(Q))$ seguida de

$$([a_1, b_1], [a_2, b_2]) \mapsto [b_1b_2, a_1b_2 + a_2b_1, a_1a_2].$$

Es fácil ver que esta última aplicación es regular. La regularidad de ψ la probaremos más abajo y la de las aplicaciones restantes es inmediata. Admitiendo la regularidad de ψ, para probar que el diagrama es conmutativo basta verlo sobre un abierto de $E \times E$, por ejemplo, en el abierto donde son válidas las expresiones para $x(P+Q)+x(P-Q)$ y $x(P+Q)x(P-Q)$ que hemos obtenido, y la comprobación es inmediata.

Ahora vamos a comprobar que ψ es regular, para aplicar 7.24. Supongamos que $\psi([t, u, v]) = [0, 0, 0]$. Si $t = 0$, entonces $u^2 = 0$, $2uv = 0$, $v^2 = 0$, luego $t = u = v = 0$. Supongamos ahora que $t \neq 0$. No perdemos generalidad si suponemos $t = 1$. Entonces tenemos:

$$u^2 - 4v = 0, \qquad 2u(A + v) + 4B = 0, \qquad (v - A)^2 - 4Bu = 0.$$

Como motivación del paso siguiente podemos pensar que si, de acuerdo con el diagrama conmutativo, $(t, u, v) = (1, x(P) + x(Q), x(P)x(Q))$, entonces la primera ecuación es $(x(P) - x(Q))^2 = 0$, luego $x(P) = x(Q) = u/2$. No vamos a usar esto para nada, pero es la razón de fondo por la que conviene hacer el cambio $x = u/2$, con lo que las ecuaciones se convierten en

$$x^2 - v = 0, \qquad 4x(A + v) + 4B = 0, \qquad (v - A)^2 - 8Bx = 0.$$

Sustituimos la primera en las otras dos:

$$4x^3 + 4Ax + 4B = 0, \qquad x^4 - 2Ax^2 - 8Bx + A^2 = 0. \tag{7.5}$$

Los polinomios

$$F(X) = X^3 + AX + B, \qquad G(X) = X^4 - 2AX^2 - 8BX + A^2$$

son bien conocidos en la teoría clásica de curvas elípticas. La fórmula de duplicación se reduce en este caso particular —para una ecuación de Weierstrass tipo c— a $x(2P) = G(x(P))/4F(x(P))$. La identidad (7.3) es ahora

$$4(3X^3 - 5AX - 27B)4F(X) - 16(3X^2 + 4A)G(X) = \Delta,$$

pero entonces (7.5) implica que $\Delta = 0$, lo cual es absurdo.

Volvamos al diagrama y llamemos σ a la aplicación vertical. Entonces

$$a(\sigma(P+Q, P-Q)) = a(\psi(\sigma(P,Q))) = 2a(\sigma(P,Q)) + O(1),$$

donde hemos usado 7.24 (teniendo en cuenta que $a = \log A$).

Ahora basta probar que

$$a(\sigma(P,Q)) = a_x(P) + a_x(Q) + O(1),$$

pues al aplicar esto a los dos miembros de la igualdad precedente obtenemos la relación del enunciado para $f = x$.

Es fácil ver que si $P = O$ o $Q = O$ entonces $a(\sigma(P,Q)) = a_x(P) + a_x(Q)$, luego basta probar la relación cuando $P \neq O \neq Q$. Pongamos que $P = [\alpha_1, 1]$, $Q = [\alpha_2, 1]$, con lo que

$$a(\sigma(P,Q)) = a([1, \alpha_1 + \alpha_2, \alpha_1\alpha_2]), \quad a_x(P) + a_x(Q) = a(\alpha_1) + a(\alpha_2).$$

Basta aplicar 7.25 al polinomio $(T + \alpha_1)(T + \alpha_2)$, lo que nos da

$$a(\alpha_1) + a(\alpha_2) - \log 4 \leq a([1, \alpha_1 + \alpha_2, \alpha_1\alpha_2]) \leq a(\alpha_1) + a(\alpha_2) + \log 2.$$

Esto termina la prueba para $f = x$. El caso general es inmediato una vez hayamos demostrado que

$$a_f = \frac{1}{2}(\operatorname{grad} f)a_x + O(1).$$

En efecto, el cuerpo de las funciones pares de $K(E)$ es $K(x)$, pues toda función de $K(E)$ se expresa en la forma $u + vy$, con u, $v \in K(x)$, y será par si y sólo si

$$u(P) + v(P)y(P) = u(-P) + v(-P)y(-P) = u(P) - v(P)y(P)$$

para todo punto $P \in E(\mathbb{A})$, lo que equivale a que $v = 0$.

Así pues, existe $R(X) \in K(X)$ tal que $f = R(x)$. Podemos ver a $R(X)$ como una función racional $\rho : \mathrm{P}^1(\mathbb{A}) \longrightarrow \mathrm{P}^1(\mathbb{A})$ que, de hecho, será regular (toda aplicación racional entre curvas proyectivas es regular). Así pues, el teorema 7.24 nos da que

$$a_f = (\operatorname{grad} \rho)a_x + O(1).$$

Como $f = x \circ \rho$, tenemos que $\operatorname{grad} f = \operatorname{grad} x \operatorname{grad} \rho = 2 \operatorname{grad} \rho$, luego

$$2a_f = (\operatorname{grad} f)a_x + O(1).$$

∎

Ahora tenemos ya todo lo necesario para demostrar el teorema que perseguíamos:

Teorema 7.30 (Teorema de Mordell-Weil) *Si K es un cuerpo numérico y E/K es una curva elíptica, entonces el grupo $E(K)$ es finitamente generado.*

DEMOSTRACIÓN: Fijemos una función par no constante $f \in K(E)$. (Por ejemplo la coordenada x respecto de una ecuación de Weierstrass.) Basta probar que la altura a_f cumple las propiedades del teorema 7.21. Enunciadas de forma ligeramente distinta, éstas son:

a) Fijado $Q \in E(K)$, para todo $P \in E(K)$ se cumple

$$a_f(P + Q) \leq 2a_f(P) + O(1),$$

donde $O(1)$ depende de Q.

b) Fijado $m \in \mathbb{Z}$, para todo $P \in E(K)$ se cumple

$$a_f(mP) = m^2 a_f(P) + O(1).$$

c) Para toda constante C, el conjunto $\{P \in G \mid a_f(P) \leq C\}$ es finito.

La propiedad a) se sigue inmediatamente del teorema anterior, teniendo en cuenta que $a_f(P - Q) \geq 0$.

Teniendo en cuenta que f es par, basta probar b) para $m \geq 0$. Los casos $m = 0$, 1 son triviales. Razonamos por inducción sobre m. Si el resultado es cierto para $m - 1$ y m, aplicamos el teorema anterior a los puntos mP y P, con lo que

$$a_f((m+1)P) = -a_f((m-1)P) + 2a_f(mP) + 2a_f(P) + O(1)$$

$$= (-(m-1)^2 + 2m^2 + 2)a_f(P) + O(1) = (m+1)^2 a_f(P) + O(1).$$

La propiedad c) es el teorema 7.28.

Así pues, el teorema 7.21 junto con 7.20 implica que $E(K)$ es finitamente generado. ∎

Aunque no nos ha hecho falta para demostrar el teorema de Mordell-Weil, es interesante notar que el teorema 7.29 se puede refinar eliminando el término $O(1)$. Para probarlo nos basaremos en el teorema siguiente:

Teorema 7.31 *Sea K un cuerpo numérico y E/K una curva elíptica. Para cada función par no constante $f \in K(E)$ y cada punto $P \in E(\mathbb{A})$, el límite*

$$\frac{1}{\operatorname{grad} f} \lim_N 4^{-N} a_f(2^N P)$$

existe y es independiente de f.

Demostración: Según hemos visto en la prueba del teorema de Mordell-Weil, (la propiedad b para $m = 2$), existe una constante C tal que para todo $Q \in E(\mathbb{A})$ se cumple

$$|a_f(2Q) - 4a_f(Q)| \leq C.$$

Vamos a probar que la sucesión del enunciado es de Cauchy. Consideremos naturales $N \geq M \geq 0$. Entonces

$$|4^{-N} a_f(2^N P) - 4^{-M} a_f(2^M P)| = \left| \sum_{n=M}^{N-1} 4^{-n-1} a_f(2^{n+1} P) - 4^{-n} a_f(2^n P) \right|$$

$$\leq \sum_{n=M}^{N-1} 4^{-n-1} |a_f(2^{n+1} P) - 4a_f(2^n P)| \leq \sum_{n=M}^{N-1} 4^{-n-1} C \leq C/4^{M+1}.$$

Así pues, la sucesión es de Cauchy, luego converge. Supongamos ahora que g es otra función par. En la prueba de 7.29 hemos visto que

$$a_f = \frac{1}{2}(\operatorname{grad} f) a_x + O(1),$$

donde x es la primera coordenada respecto de una ecuación de Weierstrass de E. Usando el mismo resultado para g obtenemos la relación

$$(\operatorname{grad} g) a_f = (\operatorname{grad} f) a_g + O(1).$$

Por consiguiente,

$$(\operatorname{grad} g) 4^{-N} a_f(2^N P) - (\operatorname{grad} f) 4^{-N} a_g(2^N P) = 4^{-N} O(1) \to 0,$$

luego el límite del enunciado no depende de f. ∎

Definición 7.32 Si K es un cuerpo numérico y E/K es una curva elíptica, definimos la *altura canónica* de E/K como la función $\hat{a}_E : E(\mathbb{A}) \longrightarrow \mathbb{R}$ dada por

$$\hat{a}_E(P) = \frac{1}{\operatorname{grad} f} \lim_N 4^{-N} a_f(2^N P),$$

donde $f \in K(E)$ es cualquier función par no constante.

Teorema 7.33 *Sea K un cuerpo numérico, E/K una curva elíptica y $\hat{a} = \hat{a}_E$ la altura canónica en E. Entonces:*

a) Para todo par de puntos P, $Q \in E(\mathbb{A})$,

$$\hat{a}(P+Q) + \hat{a}(P-Q) = 2\hat{a}(P) + 2\hat{a}(Q).$$

b) Para todo $P \in E(\mathbb{A})$ y todo $m \in \mathbb{Z}$,

$$\hat{a}(mP) = m^2\hat{a}(P).$$

c) La aplicación $\langle \ , \ \rangle : E(\mathbb{A}) \times E(\mathbb{A}) \longrightarrow \mathbb{R}$ dada por

$$\langle P, Q\rangle = \frac{1}{2}(\hat{a}(P+Q) - \hat{a}(P) - \hat{a}(Q))$$

es bilineal y $\hat{a}(P) = \langle P, P\rangle$.

d) Si $P \in E(\mathbb{A})$ entonces $\hat{a}(P) \geq 0$ y $\hat{a}(P) = 0$ si y sólo si P es un punto de torsión.

e) Si $f \in K(E)$ es una función par no constante, entonces

$$(\operatorname{grad} f)\hat{a} = a_f + O(1).$$

Además, $\hat{a}$ es la única función que cumple a la vez e) para cierta función f y b) para cierto natural $m \geq 2$.

Demostración: En primer lugar demostramos e). En la prueba del teorema anterior hemos visto que existe una constante C (que depende de f) tal que

$$|4^{-N}a_f(2^N P) - 4^{-M}a_f(2^M P)| \leq C/4^{M+1}.$$

Basta tomar $M = 0$ y hacer tender N a infinito:

$$|(\operatorname{grad} f)\hat{a}(P) - a_f(P)| \leq C/4.$$

Para probar a) partimos de 7.29:

$$a_f(P+Q) + a_f(P-Q) = 2a_f(P) + 2a_f(Q) + O(1),$$

sustituimos P, Q por $2^N P$, $2^N Q$, multiplicamos por $4^{-N}/\operatorname{grad} f$ y hacemos tender N a infinito.

b) se demuestra análogamente partiendo de la relación

$$a_f(mP) = m^2 a_f(P) + O(1),$$

demostrada en la prueba del teorema de Mordell-Weil.

La propiedad c) es consecuencia directa de a). En efecto, observemos que

$$\langle P, Q\rangle = \frac{1}{2}(\hat{a}(P+Q) - \hat{a}(P) - \hat{a}(Q)) = \frac{1}{2}(\hat{a}(P) + \hat{a}(Q) - \hat{a}(P-Q)).$$

Obviamente $\langle P,Q\rangle = \langle Q,P\rangle$. Basta probar que la expresión

$$S(P,R,Q) = (P+R,Q) - (P,Q) - (R,Q)$$

es idénticamente nula.

Como $\hat{a}(0) = 0$ y $\hat{a}(-Q) = \hat{a}(Q)$ vemos que $\langle P,-Q\rangle = -\langle P,Q\rangle$ y, por simetría, $\langle -P,Q\rangle = -\langle P,Q\rangle$. De aquí a su vez obtenemos que

$$S(P,R,-Q) = -S(P,R,Q), \quad S(-P,-R,Q) = -S(P,R,Q).$$

Ahora bien, desarrollando la definición de S vemos que

$$2S(P,R,Q) = \hat{a}(P+R+Q) - n(P+R) - \hat{a}(P+Q) - \hat{a}(R+Q) + \hat{a}(P) + \hat{a}(R) + \hat{a}(Q)$$

es una expresión simétrica en sus tres variables, luego

$$-S(P,R,Q) = S(-P,-R,Q) = -S(P,-R,Q) = S(P,R,Q),$$

de donde podemos concluir que $S(P,R,Q) = 0$.

La única parte no trivial de d) es que si $\hat{a}(P) = 0$ entonces P es un punto de torsión. En efecto, para todo $m \in \mathbb{Z}$ tenemos que $\hat{a}(mP) = m^2\hat{a}(P) = 0$. Ahora bien, de e) se deduce que el conjunto

$$\{P \in E(\mathbb{A}) \mid \hat{a}(P) < 1\}$$

es finito, pero contiene a todos los mP, luego P tiene orden finito.

Finalmente, supongamos que $a' : E(\mathbb{A}) \longrightarrow \mathbb{R}$ cumple

$$a'(mP) = m^2 a'(P), \qquad (\operatorname{grad} f)a' = a_f + O(1),$$

para cierto $m \geq 2$ y cierta $f \in K(E)$ par no constante. Entonces

$$a'(m^N P) = m^{2N} a'(P), \qquad a - \hat{a} = O(1).$$

Por lo tanto,

$$a'(P) = m^{-2N} a'(m^N P) = m^{-2N}\hat{a}(m^N P) + m^{-2N}O(1) = \hat{a}(P) + m^{-2N}O(1).$$

Haciendo tender N a infinito obtenemos $\hat{a} = a'$. ∎

Los teoremas de estructura para $\mathbb{Z}$-módulos finitamente generados nos dicen que $E(K) = E_{\mathrm{tor}}(K) \oplus L$, donde L es un $\mathbb{Z}$-módulo libre. El rango r de este módulo es un invariante de E/K a cuyo estudio dedicaremos todo el capítulo siguiente. La altura canónica es una forma cuadrática en $E(K)$, o también en el $\mathbb{Z}$-módulo libre

$$E(K)/E_{\mathrm{tor}}(K) \cong L \cong \mathbb{Z}^r.$$

En este cociente además es regular, es decir, $\hat{a}([P]) = 0$ si y sólo si $[P] = 0$.

Es claro que $\hat{a}$ se extiende a una forma cuadrática en

$$\mathbb{R} \otimes E(K)/E_{\mathrm{tor}}(K) \cong \mathbb{R}^r.$$

(Tomar el producto tensorial no es más que sumergir $\mathbb{Z}^r$ en $\mathbb{R}^r$.) Sucede que esta extensión es definida positiva, pero esto no es una consecuencia inmediata del apartado d) del teorema anterior.

Teorema 7.34 *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{R}$ y sea $R \subset V$ un retículo completo. Sea $q : V \longrightarrow \mathbb{R}$ una forma cuadrática con las propiedades siguientes:*

a) Si $P \in R$ entonces $q(P) = 0$ si y sólo si $P = 0$.

b) Para toda constante C, el conjunto $\{P \in R \mid q(P) \leq C\}$ es finito.

Entonces q es definida positiva.

DEMOSTRACIÓN: Podemos encontrar una base de V respecto a la cual, para todo punto $P \in V$ de coordenadas $(x_1, \ldots, x_r)$ se cumple

$$q(P) = \sum_{i=1}^{s} x_i^2 - \sum_{i=1}^{t} x_{s+i}^2,$$

donde $s + t \leq r$. Podemos identificar V con $\mathbb{R}^r$ a través de esta base. Vamos a aplicar el teorema de Minkowski, según el cual todo subconjunto de $\mathbb{R}^n$ absolutamente convexo de volumen suficientemente grande contiene un elemento no nulo de R. Los conjuntos

$$B(\epsilon, \delta) = \{P \in \mathbb{R}^r \mid \sum_{i=1}^{s} x_i^2 \leq \epsilon, \quad \sum_{i=1}^{t} x_{s+i}^2 \leq \delta\}$$

son absolutamente convexos para todo ϵ, $\delta > 0$. Sea

$$\lambda = \text{ínf}\{q(P) \mid P \in R, \ P \neq 0\}.$$

Por las hipótesis sobre q tenemos que $\lambda > 0$ (notemos que la propiedad b implica que $q(P) \geq 0$ en R). Si q no es definida positiva, entonces $s < r$ y el conjunto $B(\lambda/2, \delta)$ contiene un punto no nulo $P \in R$ para δ suficientemente grande (pues su volumen es infinito si $s + t < r$ y tiende a infinito cuando δ tiende a infinito si $s + t = r$). Ahora bien, entonces

$$q(P) = \sum_{i=1}^{s} x_i^2 - \sum_{i=1}^{t} x_{s+i}^2 \leq \lambda/2,$$

en contradicción con la elección de λ. ∎

Claramente el teorema es aplicable a $\hat{a}_E$ considerando a $E(K)/E_{\text{tor}}(K)$ como retículo en $\mathbb{R} \otimes E(K)$.

Capítulo VIII

El rango de una curva elíptica

Según el teorema de Mordell-Weil, si E/K es una curva elíptica sobre un cuerpo numérico, el grupo $E(K)$ es finitamente generado, por lo que

$$E(K) \cong E_{\mathrm{tor}}(K) \oplus \mathbb{Z}^r,$$

para cierto natural r, invariante por K-isomorfismos, al que se denomina *rango* de E/K. Más aún, teniendo en cuenta que el núcleo de una isogenia es un grupo finito, es claro que el rango de una curva elíptica es invariante por isogenias.

Ejemplo En el capítulo II vimos que el último teorema de Fermat para exponente 3 es cierto si y sólo si la curva de Fermat $Y^2 = X^3 - 432$ no tiene más puntos racionales que los de torsión (ver la observación tras el teorema 7.19). En otras palabras, la ecuación $X^3 + Y^3 = Z^3$ no tiene soluciones enteras no triviales si y sólo si el rango de la curva de Fermat es 0. ∎

Ejemplo Si n es un número natural libre de cuadrados, el teorema 7.18 afirma que la curva $Y^2 = X^3 - n^2X$ tiene 4 puntos de torsión, que obviamente son O, $(\pm n, 0)$ y $(0, 0)$. En virtud del Teorema 1 de la introducción, tenemos que n es congruente si y sólo si la curva tiene rango > 0. (Es fácil ver que la hipótesis de que n sea libre de cuadrados se puede eliminar.) ∎

Mientras el cálculo del grupo de torsión de una curva elíptica es relativamente simple y siempre puede obtenerse en un número finito de pasos (al menos para curva sobre $\mathbb{Q}$), no se conoce ningún método para calcular el rango de una curva elíptica arbitraria. Se conocen técnicas que funcionan en muchos casos concretos, pero no un procedimiento general.

Hemos demostrado el teorema de Mordell-Weil a través del teorema 7.21. Más concretamente podemos tomar $m = 2$, pues al aplicarlo hemos visto que

todo natural $m \geq 2$ cumple las hipótesis necesarias. Si analizamos detenidamente la prueba, veremos que toda ella es constructiva excepto la determinación de los generadores del grupo $E(K)/2E(K)$. Así, si conocemos un sistema generador de este grupo, podemos encontrar explícitamente una constante C de modo que al añadir —de acuerdo con la prueba de 7.21— todos los puntos de altura menor o igual que C obtenemos un generador de $E(K)$, que a su vez puede ser refinado hasta una base.

De todos modos, el cálculo explícito de una base es muy laborioso, así que en este capítulo nos limitaremos a proporcionar técnicas para determinar el rango y sólo eventualmente obtendremos bases explícitas.

Podemos descomponer el grupo de torsión en producto de grupos cíclicos de orden potencia de primo (dichos órdenes son los *divisores elementales* de $E(K)$):

$$E_{\mathrm{tor}}(K) \cong C_{p_1^{m_1}} \times \cdots \times C_{p_s^{m_s}}.$$

Es claro entonces que

$$E(K)/2E(K) \cong (\mathbb{Z}/2\mathbb{Z})^r \times C_{p_1^{m_1}}/2C_{p_1^{m_1}} \times \cdots \times C_{p_s^{m_s}}/2C_{p_s^{m_s}},$$

donde

$$C_{p_j^{m_j}}/2C_{p_j^{m_j}} \cong \begin{cases} C_2 & \text{si } p_j = 2, \\ 0 & \text{si } p_j \neq 2. \end{cases}$$

Así pues, $|E(K)/2E(K)| = 2^{r+t}$, donde t es el número de divisores elementales pares (el número de índices j tales que $p_j = 2$).

En definitiva, para calcular r nos basta conocer la estructura de $E_{\mathrm{tor}}(K)$ y el cardinal de $E(K)/2E(K)$.

8.1 Curvas con tres puntos de orden 2

Sea E/K una curva elíptica sobre un cuerpo numérico. En esta sección veremos un método para calcular el índice $|E(K) : 2E(K)|$ bajo el supuesto de que $E(K)$ contenga a todos los puntos de orden 2. (Ésta será una condición necesaria, aunque no suficiente, para que el método sea aplicable.)

La teoría necesaria puede desarrollarse más en general para cualquier número natural $m \geq 2$ tal que $E[m] \subset E(K)$. Entonces (ver página 185) tenemos definido el producto de Kummer $\kappa : E(K) \times G(\mathbb{A}/K) \longrightarrow E[m]$, dado por $\kappa(P, \sigma) = Q^\sigma - Q$, donde $Q \in E(\mathbb{A})$ es cualquier punto con $mQ = P$. Su núcleo izquierdo es $mE(K)$, luego el producto de Kummer induce un homomorfismo

$$\delta_E : E(K)/mE(K) \longrightarrow \mathrm{Hom}(G(\mathbb{A}/K), E[m]). \tag{8.1}$$

Sea U_m el grupo de las raíces m-simas de la unidad. El teorema 3.27 nos da que $U_m \subset K^*$. Así, todo homomorfismo $c : G(\mathbb{A}/K) \longrightarrow U_m$ tiene un núcleo de la forma $G(\mathbb{A}/L)$, para cierta extensión abeliana L/K, y c puede verse como un cociclo de $G(L/K)$. Puesto que $H^1(L/K) = 1$, existe un $\beta \in L^*$ tal que

$c(\sigma) = \beta^\sigma/\beta$, para todo $\sigma \in G(\mathbb{A}/K)$. El hecho de que $c(\sigma)$ ha de tener orden m para todo σ se traduce en que $\beta^m \in K^*$. Por consiguiente tenemos un isomorfismo

$$\delta_K : K^*/K^{*m} \longrightarrow \mathrm{Hom}(G(\mathbb{A}/K), U_m) \tag{8.2}$$

dado por $\delta_K(b)(\sigma) = \beta^\sigma/\beta$, donde $\beta \in \mathbb{A}$ es cualquier número que cumpla $\beta^m = b$.

Por último, recordemos que contamos con el producto de Weil

$$e_m : E[m] \times E[m] \longrightarrow U_m$$

definido en 3.25. El método que vamos a emplear para calcular el rango de E/K se basa en el teorema siguiente:

Teorema 8.1 *En las condiciones anteriores, existe una forma bilineal*

$$b : E(K)/mE(K) \times E[m] \longrightarrow K^*/K^{*m}$$

tal que $e_m(\delta_E(P), T) = \delta_K(b(P,T))$ (Entendiendo ambos miembros como homomorfismos $G(\mathbb{A}/K) \longrightarrow U_m$). Su núcleo izquierdo es trivial y su imagen está contenida en el grupo

$$K(S,m) = \big\{[\alpha] \in K^*/K^{*m} \;\big|\; m \mid v_\mathfrak{p}(\alpha) \text{ para todo } \mathfrak{p} \notin S\big\},$$

donde S es el conjunto formado por los primos arquimedianos de K, los primos donde E tiene mala reducción y los que dividen a m. Además, $b(P,T)$ puede calcularse como sigue:

Tomamos funciones f_T, $g_T \in K(E)$ tales que

$$(f_T) = T^m/O^m, \qquad m \circ f_T = g_T^m.$$

Entonces, para $P \neq T, O$ se cumple que

$$b(P,T) = f_T(P)K^{*m}.$$

(Notemos que $b(T,T)$ puede calcularse por linealidad.)

Demostración: Ciertamente $e_m(\delta_E(P), T)$ puede verse como un elemento de $\mathrm{Hom}(G(\mathbb{A}/K), U_m)$, el cual será la imagen por δ_K de un único elemento de K^*/K^{*m}, al que llamamos $b(P,T)$. Es inmediato que b es bilineal.

Hemos de ver que si $b(P,T) = 1$ para todo $T \in E[m]$ entonces $P \in mE(K)$. En efecto, tenemos que para todo $T \in E[m]$ y todo $\sigma \in G(\mathbb{A}/K)$ se cumple que $e_m(\kappa(P,\sigma), T) = 1$. Las propiedades del producto de Weil nos dan que $\kappa(P,\sigma) = O$, luego $P \in mE(K)$ (pues el núcleo izquierdo del producto de Kummer es $mE(K)$).

Sabemos que el núcleo derecho del producto de Kummer es $G(\mathbb{A}/L)$, donde L es la adjunción a K de las coordenadas de los puntos $Q \in E(\mathbb{A})$ tales que $mQ \in E(K)$. También sabemos que la extensión L/K es finita y no ramificada fuera de S.

Sea $\beta \in \mathbb{A}$ tal que $b(P,T) = [\beta^m]$. Entonces, para todo $\sigma \in G(\mathbb{A}/L)$ tenemos que

$$\beta^\sigma/\beta = \delta_K(b(P,T))(\sigma) = e_m(\kappa(P,\sigma),T) = e_m(O,T) = 1,$$

luego $\beta \in L$. Por lo tanto, la extensión $K(\beta)/K$ es no ramificada fuera de S. Si $\mathfrak{p} \notin S$ y $\mathfrak{P}$ es cualquier divisor de $\mathfrak{p}$ en $K(\beta)$, entonces

$$v_\mathfrak{p}(\beta^m) = v_\mathfrak{P}(\beta^m) = m v_\mathfrak{P}(\beta),$$

luego $m \mid v_\mathfrak{p}(\beta^m)$ y, por consiguiente, $b(P,T) \in K(S,m)$.

Para calcular $b(P,T)$ tomamos un punto $Q \in E(\mathbb{A})$ tal que $mQ = P$ y $\beta \in \mathbb{A}$ tal que $b(P,T) = [\beta^m]$. Así, para cada $\sigma \in G(\mathbb{A}/K)$, tenemos que

$$\beta^\sigma/\beta = \delta_K(b(P,T))(\sigma) = e_m(\delta_E(P)(\sigma),T) = e_m(Q^\sigma - Q, T).$$

Ahora aplicamos la definición de e_m, en virtud de la cual

$$\beta^\sigma/\beta = g_T(X + Q^\sigma - Q)/g_T(X),$$

donde X es cualquier punto donde g_T no tenga un cero ni un polo. Si $P \neq T, O$ podemos hacer $X = Q$, con lo que obtenemos

$$\beta^\sigma/\beta = g_T(Q)^\sigma/g_T(Q).$$

Esto implica que $g_T(Q)/\beta \in K^*$, luego $g_T(Q)^m \equiv \beta^m \pmod{K^{*m}}$. Por la elección de f_T esto equivale a que $f_T(P) \equiv \beta^m \pmod{K^{*m}}$, luego concluimos que $b(P,T) = f_T(P)K^{*m}$. ∎

Las aplicaciones de este teorema se basan en que en la prueba del teorema débil de Mordell-Weil hemos visto que el grupo $K(S,m)$ es finito, y el teorema anterior nos proporciona un monomorfismo de $E(K)/mE(K)$ en el grupo finito $\mathrm{Hom}(E[m], K(S,m))$.

A partir de aquí nos restringiremos al caso $m = 2$. Consideremos una ecuación de Weierstrass $Y^2 = X^3 + a_2X^2 + a_4X + a_6$ para la curva E/K. Los puntos de orden 2 en E son los puntos que cumplen $Y = 0$, es decir, las raíces del miembro derecho de la ecuación. La hipótesis $E[2] \subset E(K)$ equivale, pues, a que la ecuación factoriza en la forma

$$Y^2 = (X - e_1)(X - e_2)(X - e_3), \qquad e_1, e_2, e_3 \in K,$$

de modo que los puntos de orden 2 de E son $T_i = (e_i, 0)$. Si llamamos $T = (e, 0)$ a uno cualquiera de los tres, la función f_T del teorema anterior es $f_T = x - e$. En efecto, tiene un único cero en T y un único polo en O, y su divisor es de la forma $(f_T) = T^2/O^2$, ya que $v_O(f_T) = v_O(x) = 2$. Además, la fórmula de duplicación se reduce a

$$2 \circ x = \frac{x^4 - 2a_4x^2 - 8a_6x + a_4^2 - 4a_2a_6}{(2y)^2}.$$

Por lo tanto, $2 \circ f_T = 2 \circ x - e$ resulta ser

$$\frac{x^4 - 4ex^3 + (-2a_4 - 4ea_2)x^2 + (-8a_6 - 4ea_4)x + a_4 - 4a_2a_6 - 4ea_6}{(2y)^2}.$$

Comparando el numerador con

$$(x^2 + Px + Q)^2 = x^2 + 2Px^3 + (P^2 + 2Q)x^2 + 2PQx + Q^2$$

e igualando los coeficientes de x^3 y x vemos que con $P = -2e$, $Q = 2a_6/e + a_4$ resulta una identidad. Para comprobar que el coeficiente de x^2 y el término independiente coinciden usamos las relaciones

$$a_2 = -e_1 - e_2 - e_3, \quad a_4 = e_1e_2 + e_1e_3 + e_2e_3, \quad a_6 = -e_1e_2e_3$$

y suponemos, sin pérdida de generalidad, que $e = e_1$. En definitiva, concluimos que $f_T = g_T^2$ para cierta función $g_T \in K(E)$, como había que comprobar.

La forma bilineal b del teorema anterior nos da un monomorfismo de grupos

$$E(K)/2E(K) \longrightarrow \mathrm{Hom}(E(K), K(S,2)).$$

Podemos ver a $E(K)$ y a $K(S,2)$ como espacios vectoriales sobre $\mathbb{Z}/2\mathbb{Z}$, de modo que los homomorfismos de grupos entre ambos coinciden con las aplicaciones lineales. Una base del primer espacio es (T_1, T_2), luego el grupo de homomorfismos es isomorfo al grupo $K(S,2) \times K(S,2)$, donde identificamos cada homomorfismo f con el par $(f(T_1), f(T_2))$.

Nuestro problema es identificar los pares $(b_1, b_2) \in K(S,2) \times K(S,2)$ que se corresponden con clases $[P] \in E(K)/2E(K)$, es decir, tales que existe un $P \in E(K)$ de modo que $b(P, T_1) = b_1$, $b(P, T_2) = b_2$. El número de tales pares es el orden de $E(K)/2E(K)$.

Según el teorema anterior, un par (b_1, b_2) se corresponderá con un punto $P \neq O$, T_1, T_2 si existen números $(x, y, z_1, z_2) \in K \times K \times K^* \times K^*$ tales que

$$y^2 = (x - e_1)(x - e_2)(x - e_3), \qquad b_1z_1^2 = x - e_1, \qquad b_2z_2^2 = x - e_2.$$

Podemos sustituir las dos últimas ecuaciones en la primera y definir una nueva variable z_3 mediante $y = b_1b_2z_1z_2z_3$, con lo que las ecuaciones anteriores equivalen a

$$b_1b_2z_3^2 = x - e_3, \qquad b_1z_1^2 = x - e_1, \qquad b_2z_2^2 = x - e_2.$$

Por último, eliminamos x, con lo que el sistema equivale a

$$b_1z_1^2 - b_2z_2^2 = e_2 - e_1, \qquad b_1z_1^2 - b_1b_2z_3^2 = e_3 - e_1.$$

Cada par (b_1, b_2) para el que existe una solución (z_1, z_2, z_3) de este sistema de ecuaciones se corresponde con un punto $[P] \in E(K)/2E(K)$ dado por

$$x = b_1z_1^2 + e_1, \qquad y = b_1b_2z_1z_2z_3.$$

Aquí nos falta añadir el par $(b_1, b_2) = (1, 1)$, correspondiente a $P = O$ y los pares $(b(T_1, T_1), b(T_1, T_2))$, $(b(T_2, T_1), b(T_2, T_2))$. Podemos calcularlos por linealidad:

$$b(T_1, T_1) = b(T_1, T_1 + T_2)b(T_1, T_2)^{-1} = b(T_1, T_3)b(T_1, T_2)^{-1}$$

$$= \frac{e_1 - e_3}{e_1 - e_2} = (e_1 - e_2)(e_1 - e_3),$$

e igualmente

$$b(T_2, T_2) = (e_2 - e_1)(e_2 - e_3).$$

El teorema siguiente resume lo que hemos obtenido:

Teorema 8.2 *Sea E/K una curva elíptica sobre un cuerpo numérico K dada por una ecuación de Weierstrass*

$$y^2 = (x - e_1)(x - e_2)(x - e_3), \qquad \text{con } e_1, e_2, e_3 \in K.$$

Sea S el conjunto de los primos arquimedianos de K más los primos que dividen a 2 más los primos donde E tiene mala reducción. Sea

$$K(S, 2) = \big\{[\alpha] \in K^*/K^{*2} \;\big|\; 2 \mid v_{\mathfrak{p}}(\alpha) \text{ para todo } \mathfrak{p} \notin S\big\}.$$

Entonces existe un monomorfismo de grupos

$$E(K)/2E(K) \longrightarrow K(S, 2) \times K(S, 2)$$

dado por

$$P = (x, y) \mapsto \begin{cases} (x - e_1, x - e_2) & \text{si } x \neq e_1,\ e_2, \\ ((e_1 - e_2)(e_1 - e_3), e_1 - e_2) & \text{si } x = e_1, \\ (e_2 - e_1, (e_2 - e_1)(e_2 - e_3)) & \text{si } x = e_2, \\ (1, 1) & \text{si } x = \infty \text{ (o sea, si } P = O). \end{cases}$$

Si un par $(b_1, b_2) \in K(S, 2) \times K(S, 2)$ no es la imagen de uno de los tres puntos O, $(e_1, 0)$, $(e_2, 0)$, entonces es la imagen de un punto (x, y) si y sólo si las ecuaciones

$$b_1 z_1^2 - b_2 z_2^2 = e_2 - e_1, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = e_3 - e_1$$

tienen una solución $(z_1, z_2, z_3) \in K^* \times K^* \times K$. Si tal solución existe, entonces $(x, y) = (b_1 z_1^2 + e_1, b_1 b_2 z_1 z_2 z_3)$.

Ejemplo Vamos a calcular el rango de la curva elíptica $E/\mathbb{Q}$ dada por

$$Y^2 = X^3 - 12X^2 + 20X = X(X - 2)(X - 10).$$

Su discriminante es $\Delta = 2^{14}5^2$, luego tiene buena reducción en todos los primos excepto en 2 y 5. Por consiguiente

$$S = \{2, 5, \infty\}.$$

De aquí se sigue inmediatamente que

$$\mathbb{Q}(S,2) = \{[\pm 1], [\pm 2], [\pm 5], [\pm 10]\}.$$

Ahora hemos de determinar cuáles de los 64 pares de $\mathbb{Q}(S,2) \times \mathbb{Q}(S,2)$ están asociados a elementos $[P] \in E(\mathbb{Q})/2E(\mathbb{Q})$. Pongamos, por ejemplo,

$$e_1 = 0, \qquad e_2 = 2, \qquad e_3 = 10.$$

Los puntos de torsión O, $(0,0)$, $(2,0)$, $(10,0)$ se corresponden con los pares

$$(1,1), (5,-2), (2,-1), (10,2).$$

Cualquier otro par (b_1, b_2) se corresponderá con un punto P si y sólo si las ecuaciones

$$b_1 z_1^2 - b_2 z_2^2 = 2, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = 10$$

tienen soluciones en $\mathbb{Q}$ (con $z_1,\ z_2 \neq 0$). La tabla siguiente contiene un punto P cuando existe o un cuerpo donde las ecuaciones del teorema anterior no tienen solución (las columnas corresponden a valores de b_1 y las filas de b_2):

<table>
<tr><td></td><td>1</td><td>2</td><td>5</td><td>10</td><td>$-1,-2,-5,-10$</td></tr>
<tr><td>1</td><td>O</td><td>$(18,-48)$</td><td colspan="2">$\mathbb{Q}_5$</td><td rowspan="3">$\mathbb{R}$</td></tr>
<tr><td>2</td><td>$\mathbb{Q}_5$</td><td>$\mathbb{Q}_5$</td><td>$(20,60)$</td><td>$(10,0)$</td></tr>
<tr><td>$5,10$</td><td colspan="2">$\mathbb{Q}_5$</td><td colspan="2">$\mathbb{Q}_5$</td></tr>
<tr><td>-1</td><td>$(1,-3)$</td><td>$(2,0)$</td><td colspan="2">$\mathbb{Q}_5$</td><td rowspan="3">$\mathbb{R}$</td></tr>
<tr><td>-2</td><td colspan="2">$\mathbb{Q}_5$</td><td>$(0,0)$</td><td>$(10/9,-80/27)$</td></tr>
<tr><td>$-5,-10$</td><td colspan="2">$\mathbb{Q}_5$</td><td colspan="2">$\mathbb{Q}_5$</td></tr>
</table>

En efecto:

a) En primer lugar situamos los cuatro puntos de torsión.

b) Si $b_1 < 0$ y $b_2 > 0$ la ecuación $b_1 z_1^2 - b_2 z_2^2 = 2$ no tiene soluciones en $\mathbb{R}$.

c) Si $b_1 < 0$ y $b_2 < 0$ la ecuación $b_1 z_1^2 - b_1 b_2 z_3^2 = 10$ no tiene soluciones en $\mathbb{R}$.

d) Para $(b_1, b_2) = (1,-1)$, las ecuaciones son

$$z_1^2 + z_2^2 = 2, \qquad z_1^2 + z_3^2 = 10.$$

Es claro que una solución es $(1,1,3)$, que se corresponde con el punto $(1,-3) \in E(\mathbb{Q})$.

e) Teniendo en cuenta que estamos calculando la imagen de un homomorfismo, concluimos que dicha imagen ha de contener los productos

$$(1,-1)(5,-2) = (5,2),\ (1,-1)(2,-1) = (2,1),\ (1,-1)(10,2) = (10,-2),$$

que se corresponden con los puntos

$$(20,60), \quad (18,-48), \quad (10/9,-80/27) \in E(\mathbb{Q}).$$

f) Si $5 \nmid b_1$ y $5 \mid b_2$ observamos que

$$v_5(b_1 z_1^2) = 2v_5(z_1), \qquad v_5(-b_2 z_2^2) = 2v_5(z_2) + 1.$$

En particular las dos cantidades son distintas, luego

$$0 = v_5(2) = v_5(b_1 z_1^2 - b_2 z_2^2) = \min\{2v_5(z_1), 2v_5(z_2) + 1\},$$

de donde se sigue que $v_5(z_1)$, $v_5(z_2) \geq 0$, es decir, z_1, z_2 son enteros en $\mathbb{Q}_5$. Considerando la segunda ecuación vemos que z_3 también ha de ser entero en $\mathbb{Q}_5$. Además $v_5(z_1) \geq 1$ y de la primera se sigue que $5 \mid 2$. Así pues, no hay solución.

g) Al multiplicar por $(5, 2)$ los ocho pares que acabamos de descartar obtenemos otros ocho pares que no pertenecen a la imagen de $E(\mathbb{Q})/2E(\mathbb{Q})$. Concretamente son

$$(5,2)(1,5) = (5,10), \quad (5,2)(1,10) = (5,5), \quad (5,2)(1,-5) = (5,-10),$$

$$(5,2)(1,-10) = (5,-5), \quad (5,2)(2,5) = (10,10), \quad (5,2)(2,10) = (10,5),$$

$$(5,2)(2,-5) = (10,-10), \quad (5,2)(2,-10) = (10,-5).$$

h) Para $(b_1, b_2) = (1, 2)$, las ecuaciones son

$$z_1^2 - 2z_2^2 = 2, \qquad z_1^2 - 2z_3^2 = 10.$$

Pongamos que $z_1 = u/w$, $z_3 = v/w$, con $(u, v, w) = 1$. Entonces tenemos que $5 \mid u^2 - 2v^2$, pero 2 es un resto no cuadrático módulo 5, luego $5 \mid u$, $5 \mid v$, y por lo tanto $5 \nmid w$. Así pues, z_1 y z_3 son enteros en $\mathbb{Q}_5$ y múltiplos de 5, pero entonces llegamos a que $25 \mid 10$, luego no hay solución.

i) Al multiplicar $(1, 2)$ por los siete pares no triviales que están en la imagen de $E(\mathbb{Q})/2E(\mathbb{Q})$ tenemos siete puntos más que no están en dicha imagen, con lo que la tabla queda completa.

En total tenemos que $|E(\mathbb{Q})/2E(\mathbb{Q})| = 2^3$. Por otra parte, se comprueba que $|E(\mathbb{Z}/3\mathbb{Z})| = 4$, luego el teorema 7.17 nos da que

$$E_{\mathrm{tor}}(\mathbb{Q}) = E[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

Esto nos permite concluir que $E/\mathbb{Q}$ tiene rango 1, es decir,

$$E(\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}.$$

■

Ejemplo Vamos a calcular el rango de la curva elíptica $E/\mathbb{Q}$ dada por

$$Y^2 = X(X-1)(X+3).$$

Su discriminante es $\Delta = 2^8 \cdot 3^2$, luego $S = \{2, 3, \infty\}$. Se comprueba que $|E(\mathbb{Z}/5\mathbb{Z})| = 8$, por lo que $E/\mathbb{Q}$ tiene a lo sumo 8 puntos de torsión. En efecto, es fácil ver que

$$\begin{aligned} E_{\mathrm{tor}}(\mathbb{Q}) &= \{O, (-1,2), (1,0), (-1,-2), (0,0), (-3,0), (3,6), (3,-6)\} \\ &= \langle (-1,2)\rangle \oplus \langle (0,0)\rangle \cong C_4 \oplus C_2. \end{aligned}$$

Tenemos que $\mathbb{Q}(S,2) = \{[\pm 1], [\pm 2], [\pm 3], [\pm 6]\}$. Tomaremos por ejemplo $e_1 = 0$, $e_2 = 1$, $e_3 = -3$. Las imágenes en $\mathbb{Q}(S,2) \times \mathbb{Q}(S,2)$ de los puntos de torsión son los pares

$$(1,1), \qquad (-1,-2), \qquad (-3,-1), \qquad (3,2).$$

Vamos a ver que no hay más imágenes, con lo que $|E(\mathbb{Q})/2E(\mathbb{Q})| = 4$ y $E/\mathbb{Q}$ resulta tener rango 0.

Para que un par (b_1, b_2) distinto de los ya considerados tenga una antiimagen se han de cumplir las ecuaciones

$$b_1 z_1^2 - b_2 z_2^2 = 1, \qquad b_1 z_1^2 - b_1 b_2 x_3^2 = -3, \qquad z_1,\ z_2 \neq 0.$$

La primera ecuación no tiene soluciones reales cuando $b_1 < 0$, $b_2 > 0$, mientras que la segunda no tiene soluciones reales cuando $b_1 > 0$, $b_2 < 0$. Así pues, b_1 y b_2 han de tener el mismo signo.

Si $3 \nmid b_1$, $3 \mid b_2$, como en el ejemplo anterior deducimos que z_1, z_2, z_3 son enteros en $\mathbb{Q}_3$. Por la primera ecuación $3 \nmid z_1$, mientras que la segunda implica lo contrario. Esto nos descarta los ocho pares

$$(\pm 1, \pm 3), \quad (\pm 1, \pm 6), \quad (\pm 2, \pm 3), \quad (\pm 2, \pm 6).$$

Multiplicando estos pares por $(3,2)$, que sí tiene antiimagen, obtenemos ocho pares más sin antiimagen:

$$(\pm 3, \pm 6), \quad (\pm 3, \pm 3), \quad (\pm 6, \pm 6), \quad (\pm 6, \pm 3).$$

Para $(3,1)$ tenemos la ecuación $3z_1^2 - z_2^2 = 1$, que implica que z_1 y z_2 son enteros en $\mathbb{Q}_3$ y $z_2^2 \equiv -1$ (mód 3), lo cual es imposible. Al multiplicar este par por los cuatro pares con antiimagen obtenemos cuatro pares sin antiimagen:

$$(3,1), \quad (-3,-2), \quad (-1,-1), \quad (1,2).$$

Si $2 \mid b_1$, $b_2 = \pm 1$, la primera ecuación nos da que z_1 y z_2 son enteros en $\mathbb{Q}_2$. Por otra parte, $v_2(b_1 z_1^2) = 2v_2(z_1) + 1$, $v_2(-b_1 b_2 z_3^2) = 2v_2(z_3) + 1$. Si ambos valores coinciden entonces z_3 también es entero, y si son distintos la segunda ecuación nos da que el menor de ellos es 0, luego en cualquier caso z_3 es entero.

La segunda ecuación nos da entonces la contradicción $2 \mid 3$. Esto descarta los cuatro puntos

$$(\pm 2, \pm 1), \quad (\pm 6, \pm 1).$$

Al multiplicarlos por $(3, 2)$, que sí que tiene antiimagen, obtenemos los cuatro puntos

$$(\pm 2, \pm 2), \quad (\pm 6, \pm 2),$$

con lo que sólo nos quedan los cuatro puntos iniciales.

En particular, las únicas soluciones racionales de la ecuación dada son los ocho puntos de torsión. ∎

Ejemplo *El número 2 no es congruente.*

Aunque ya conocemos una prueba de este hecho basada en que la ecuación de Fermat $X^4 + Y^4 = Z^2$ no tiene soluciones no triviales, vamos a dar ahora una prueba basada en curvas elípticas, con lo que tendremos una demostración alternativa del resultado de Fermat sobre esta ecuación.

Hemos de probar que la curva $Y^2 = X^3 - 4X$ tiene rango 0. En este caso $S = \{2, \infty\}$ y $\mathbb{Q}(S, 2) = \{[\pm 1], [\pm 2]\}$. Sabemos que los puntos de torsión de la curva son O, $(-2, 0)$, $(0, 0)$ y $(2, 0)$, cuyas imágenes en $\mathbb{Q}(S, 2) \times \mathbb{Q}(S, 2)$ (tomando, por ejemplo, $e_1 = -2$, $e_2 = 0$, $e_3 = 2$) son

$$(1, 1), \quad (2, -2), \quad (2, -1), \quad (1, 2).$$

La tabla correspondiente es:

<table>
<tr><th></th><th>1</th><th>2</th><th>−1 −2</th></tr>
<tr><td>1
2</td><td>O
$(2,0)$</td><td rowspan="2">$\mathbb{Q}_2$</td><td rowspan="2">$\mathbb{R}$</td></tr>
<tr><td>−1
−2</td><td rowspan="2">$\mathbb{Q}_2$</td><td>$(0,0)$
$(-2,0)$</td><td rowspan="2">$\mathbb{R}$</td></tr>
</table>

En efecto, un par (b_1, b_2) es la imagen de un punto de orden infinito si y sólo si se cumplen las ecuaciones

$$b_1 z_1^2 - b_2 z_2^2 = 2, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = 4, \qquad (z_1, z_2, z_3) \in \mathbb{Q}^* \times \mathbb{Q}^* \times \mathbb{Q}.$$

Los casos descartados con $\mathbb{R}$ son obvios. Basta probar que, por ejemplo, las ecuaciones correspondientes a $(1, -2)$ no tienen solución en $\mathbb{Q}_2$. Éstas son:

$$z_1^2 + 2z_2^2 = 2, \qquad z_1^2 + 2z_3^2 = 4.$$

Si hubiera solución:

$$2 = v_2(4) = \min\{2v_2(z_1), 2v_2(z_3) + 1\},$$

de donde $v_2(z_1) = 1$. Con la primera ecuación obtenemos similarmente que $v_2(z_2) = 0$. Haciendo $z_1 = 2\epsilon$, tenemos que $2\epsilon^2 + z_2^2 = 1$. Por consiguiente

$$1 \equiv 2 + 1 \pmod 4,$$

lo cual es absurdo. ∎

Ejemplo *El número* 10 *no es congruente.*

Hemos de considerar la curva $Y^2 = X^3 - 100X$. Tomamos $e_1 = -10$, $e_2 = 0$, $e_3 = 10$, con lo que las ecuaciones son

$$b_1 z_1^2 - b_2 z_2^2 = 10, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = 20.$$

Vamos a completar la tabla siguiente:

	1	5	2	10	$-1\ -5\ -2\ -10$
1	O	$\mathbb{Q}_2$	$\mathbb{Q}_2$	$\underline{\mathbb{Q}_2}$	
5	$\mathbb{Q}_5$	$\mathbb{Q}_2$	$\mathbb{Q}_2$	$\mathbb{Q}_5$	
2	$\mathbb{Q}_2$	$\mathbb{Q}_5$	$\mathbb{Q}_5$	$\mathbb{Q}_2$	
10	$\mathbb{Q}_2$	$(10,0)$	$\underline{\mathbb{Q}_2}$	$\mathbb{Q}_2$	$\mathbb{R}$
-1	$\underline{\mathbb{Q}_2}$	$\mathbb{Q}_2$	$\mathbb{Q}_2$	$(0,0)$	
-5	$\mathbb{Q}_5$	$\mathbb{Q}_2$	$\mathbb{Q}_2$	$\mathbb{Q}_5$	
-2	$\mathbb{Q}_2$	$\mathbb{Q}_5$	$\mathbb{Q}_5$	$\mathbb{Q}_2$	
-10	$\mathbb{Q}_2$	$\underline{\mathbb{Q}_2}$	$(-10,0)$	$\mathbb{Q}_2$	

a) Primero situamos los cuatro puntos de torsión.

b) Si $b_1 < 0$ las ecuaciones no tienen soluciones en $\mathbb{R}$.

c) Consideremos el caso $b_1 = 2$ y $v_2(b_2) = 0$.

La ecuación $2z_1^2 - b_2 z_2^2 = 10$ implica que $v_2(z_1) = 0$, $v_2(z_2) \geq 1$. Haciendo $z_2 = 2u$ obtenemos $z_1^2 - 2b_2u^2 = 5$, de donde $2b_2u^2 \equiv 0$ (mód 4), lo que implica que $2 \mid u$, luego $z_1^2 \equiv 5$ (mód 8), y esto es imposible.

Con esto descartamos todas las entradas de la columna del 2 marcadas con $\mathbb{Q}_2$ (sin subrayar). Al multiplicar estos puntos por los de torsión descartamos todas las entradas de la tabla marcadas con $\mathbb{Q}_2$.

d) Consideremos $b_1 = 5$, $b_2 = \pm 2$. La primera ecuación nos da que $v_5(z_1) = 0$ y $v_5(z_2) \geq 1$. Haciendo $z_2 = 5u$ queda $z_1^2 - 5b_2u^2 = 2$, de donde llegamos a la contradicción $z_1^2 \equiv 2$ (mód 5). Esto nos descarta todas las entradas de la tabla marcadas con $\mathbb{Q}_5$.

e) Finalmente consideramos el par $(10, 1)$. La primera ecuación nos da que $v_2(z_1) = 0$ y entonces la segunda, que es $z_1^2 - z_3^2 = 2$ nos permite concluir que $z_3 \in \mathbb{Z}_2$, así como que $z_3^2 \equiv -1$ (mód 4), lo cual es absurdo. Esto descarta las cuatro entradas de la tabla marcadas con $\underline{\mathbb{Q}_2}$.

Vemos, pues, que la curva tiene rango 0. ∎

Similarmente podríamos demostrar que 1 y 3 no son congruentes, pero vamos a deducirlo de resultados más generales. Observemos que en una curva elíptica

$$Y^2 = (X - e_1)(X - e_2)(X - e_3), \qquad e_1, e_2, e_3 \in \mathbb{Z}$$

dos de las raíces e_i han de ser congruentes módulo 2, luego la curva tiene mala reducción módulo 2 y, por consiguiente, $2 \mid \Delta$. Si llamamos s al número de primos que dividen a Δ, concluimos que el conjunto S definido en 8.2 cumple $|S| = s+1$, luego $|\mathbb{Q}(S,2)| = 2^{s+1}$ y

$$|E(\mathbb{Q})/2E(\mathbb{Q})| \le 2^{2s+2}.$$

Por otra parte, $E(\mathbb{Q})$ contiene al menos un subgrupo $C_2 \oplus C_2$, luego el número de divisores elementales pares es $t \ge 2$. Si llamamos r al rango de E, tenemos la relación $2^{r+t} = |E(\mathbb{Q})/2E(\mathbb{Q})| \le 2^{2s+2}$, luego $r \le 2s+2-t \le 2s$. Tenemos así una cota sencilla para el rango de la curva, que no obstante puede ser mejorada:

Teorema 8.3 *Sea $E/\mathbb{Q}$ una curva elíptica de la forma*

$$Y^2 = (X-e_1)(X-e_2)(X-e_3), \qquad e_1, e_2, e_3 \in \mathbb{Z}.$$

Sean m y a el número de primos sobre los que E tiene reducción multiplicativa y aditiva respectivamente. Entonces el rango r de $E/\mathbb{Q}$ satisface la relación

$$r \le m + 2a - 1.$$

Demostración: Observemos en primer lugar que si $p \mid \Delta$ y las tres raíces son congruentes módulo p, entonces la reducción es de la forma $Y^2 = (X-e)^3$, y claramente es aditiva. Por el contrario, si sólo dos de las raíces son congruentes módulo p, entonces la reducción es de la forma $Y^2 = (x-e)^2(X-e')$, que tras un cambio de variables se convierte en $Y^2 = X^3 - eX^2$. Esta curva tiene un nodo en $(0,0)$, luego la reducción es multiplicativa.

La aplicación del teorema 8.2 inyecta $E(\mathbb{Q})/2E(\mathbb{Q})$ en un grupo que es suma directa de grupos tantos grupos $C_2 \oplus C_2$ como elementos tiene el conjunto S. Por ejemplo, la coordenada asociada al primo $\infty \in S$ viene dada por

$$\phi_\infty(x,y) = \begin{cases} (\operatorname{sig}(x-e_1), \operatorname{sig}(x-e_2)) & \text{si } x \ne e_1, e_2, \\ (\operatorname{sig}((e_1-e_2)(e_1-e_3)), \operatorname{sig}(e_1-e_2)) & \text{si } x = e_1, \\ (\operatorname{sig}(e_2-e_1), \operatorname{sig}((e_2-e_1)(e_2-e_3)) & \text{si } x = e_2, \\ (1,1) & \text{si } x = \infty. \end{cases}$$

Podemos suponer que $e_1 < e_2 < e_3$. Así, si $x \notin \{e_1, e_2, e_3\}$ se cumple $x-e_1 > x-e_2 > x-e_3$ y además

$$(x-e_1)(x-e_2)(x-e_3) = y^2 > 0.$$

Por lo tanto ha de ser $x - e_1 > 0$. Concluimos que

$$\phi_\infty(x,y) = \begin{cases} (1, \operatorname{sig}(x-e_2)) & \text{si } x \ne e_1, e_2, \\ (1,-1) & \text{si } x = e_1, \\ (1,-1) & \text{si } x = e_2, \\ (1,1) & \text{si } x = \infty. \end{cases}$$

Vemos, pues, que la imagen de $E(\mathbb{Q})/2E(\mathbb{Q})$ está contenida en el subgrupo de $\mathbb{Q}(2,S)\times\mathbb{Q}(2,S)$ cuya primera componente es positiva, lo que reduce a la mitad el máximo orden posible.

Consideremos ahora un primo p donde E tenga reducción multiplicativa. Supongamos en primer lugar que $p \mid e_1 - e_2$. Sea (x,y) un punto finito de $E(\mathbb{Q})$ y supongamos además que $x \notin \{e_1, e_2, e_3\}$. Llamemos $u_i = v_p(x-e_i)$. Entonces $u_1+u_2+u_3 = v_p(y^2) = 2v_p(y)$. Observemos además que si un $u_i < 0$, entonces, como e_i es entero, $-u_i$ ha de ser el exponente de p en el denominador de x, luego $u_1 = u_2 = u_3$ y los tres son pares.

La coordenada de la imagen de (x,y) asociada a p (vista como elemento de $\mathbb{Z}/2\mathbb{Z}\oplus\mathbb{Z}/2\mathbb{Z}$) es

$$\phi_p(x,y) = (u_1 \text{ (mód 2)}, u_2 \text{ (mód 2)}).$$

Acabamos de probar que si un $u_i < 0$ entonces $\phi_p(x,y) = (0,0)$. Lo mismo sucede si los tres u_i son nulos. Si $u_1 > 0$, entonces p divide al numerador de $x - e_1$, luego también al de $x - e_2$, pero no al de $x - e_3$ (porque la reducción es multiplicativa). Así pues, $u_3 = 0$, $u_1 + u_2$ es par y $\phi_p(x,y) \in \{(0,0),(1,1)\}$. Lo mismo sucede si $u_2 > 0$. Por el contrario, si $u_3 > 0$ ha de ser $u_1 = u_2 = 0$ y $\phi_p(x,y) = (0,0)$.

En cualquier caso, tenemos que $\phi_p(x,y) \in \{(0,0),(1,1)\}$. Esto también es cierto si $x \in \{e_1,e_2,e_3\}$. En efecto, $\phi_p(e_3,0) = (0,0)$ y

$$\phi_p(e_1,0) = \phi_p(e_2,0) = (v_p(e_1-e_2) \text{ (mód 2)}, v_p(e_1-e_2) \text{ (mód 2)}).$$

Así pues, el sumando $C_2 \oplus C_2$ asociado a p puede sustituirse por un sumando C_2. Si $p \mid e_1 - e_3$ la imagen de ϕ_p resulta estar en $\{(0,0),(1,0)\}$, mientras que si $p \mid e_2 - e_3$ la imagen resulta estar en $\{(0,0),(0,1)\}$, pero la conclusión es la misma.

En definitiva, hemos sumergido $E(\mathbb{Q})/2E(\mathbb{Q})$ en un grupo de 2^{m+2a+1} elementos, y el razonamiento previo al teorema aplicado a esta cota nos da la desigualdad del enunciado. ∎

Ejemplo *El número* 1 *no es congruente.*

Hemos de probar que la curva $Y^2 = X^3 - X$ tiene rango 0. Ahora bien, sólo tiene mala reducción en $p = 2$, donde la reducción es multiplicativa. El teorema anterior nos da que el rango es $r \le 1 - 1 = 0$. ∎

Las cotas que proporciona el teorema anterior no siempre se alcanzan. Por ejemplo, si lo aplicamos a la curva $Y^2 = X^3 - 4X$, que tiene reducción aditiva en 2, la cota que obtenemos es $r \le 1 + 2\cdot 1 = 1$, mientras que el rango es $r = 0$. Veamos, no obstante, que en ciertos casos podemos refinarlas:

Teorema 8.4 *Sea p un primo impar y $E/\mathbb{Q}$ la curva elíptica dada por la ecuación $Y^2 = X^3 - p^2X$. Entonces, su rango r cumple la desigualdad*

$$r \le \begin{cases} 2 & \text{si } p \equiv 1 \text{ (mód 8)},\\ 0 & \text{si } p \equiv 3 \text{ (mód 8)},\\ 1 & \text{si } p \equiv 5,7 \text{ (mód 8)}. \end{cases}$$

DEMOSTRACIÓN: Tenemos que $S = \{[\pm 1], [\pm 2], [\pm p], [\pm 2p]\}$. Según el teorema anterior, tomamos $e_1 = -p$, $e_2 = 0$. Las ecuaciones para los puntos de orden infinito son:

$$b_1 z_1^2 - b_2 z_2^2 = p, \qquad b_1 z_1^2 - b_1 b_2 z_3^2 = 2p.$$

Esta tabla recoge las conclusiones de los razonamientos que le siguen:

	1	p	2	$2p$	$-1, -p, -2, -2p$
1	O	$1, 5$	$1, 7$	1	$\mathbb{R}$
p	1	$1, 7$	$1, 5$	$(p, 0)$	
$2, 2p$	Teorema anterior				
-1	$1, 5$	$(0, 0)$	1	$1, 7$	$\mathbb{R}$
$-p$	$1, 7$	1	$(-p, 0)$	$1, 5$	
$-2, -2p$	Teorema anterior				

En primer lugar hemos situado en la tabla los puntos de torsión, cuyas imágenes son

$$(1, 1), \quad (p, -1), \quad (2, -p), \quad (2p, p).$$

De la prueba del teorema anterior se desprende que b_1 ha de ser positivo. Marcamos los pares descartados con $\mathbb{R}$, pues son también los pares para los que las ecuaciones no tienen soluciones en $\mathbb{R}$. Así mismo, teniendo en cuenta que 2 tiene reducción multiplicativa, en la prueba del teorema anterior se ve que $v_2(b_2) = 0$. Esto nos descarta las ocho entradas de la tabla marcadas con "Teorema anterior".

Consideremos la primera ecuación para el par $(b_1, b_2) = (1, -1)$ y hagamos $z_1 = u/w$, $z_2 = v/w$, con u, v, $w \in \mathbb{Z}$:

$$u^2 + v^2 = w^2 p.$$

Si $p \mid u$, entonces $p \mid v$ y $p \mid w$, podemos simplificar p^2 en ambos miembros y repetir el proceso hasta que $p \nmid u$, e igualmente $p \nmid v$. Entonces resulta que $u^2 + v^2 \equiv 0 \pmod{p}$, luego -1 es un resto cuadrático módulo p, lo que equivale a que $p \equiv 1, 5 \pmod 8$. En este caso la segunda ecuación coincide con la primera.

Así pues, si $p \not\equiv 1, 5 \pmod 8$, el par $(1, -1)$, y con él los pares

$$(1, -1), \quad (2p, -p), \quad (p, 1), \quad (2, p),$$

no están en la imagen. En la tabla están marcados con $1, 5$.

Razonando igualmente con $(2, 1)$ llegamos a que $2u^2 - v^2 \equiv 0 \pmod{p}$, con lo que 2 es un resto cuadrático módulo p, lo que equivale a que $p \equiv \pm 1 \pmod 8$.

La segunda ecuación se convierte en $u^2 - v^2 = w^2 p$, que no aporta ninguna información.

Concluimos que si $p \not\equiv 1, 7 \pmod 8$ entonces el par $(2, 1)$, y con él los pares

$$(2, 1), \quad (p, p), \quad (2p, -1), \quad (1, -p)$$

no están en la imagen. En la tabla están marcados con $1, 7$.

Similarmente, la primera ecuación para el punto $(2,-1)$ nos lleva a que -2 es un resto cuadrático módulo p, lo cual equivale a que $p \equiv 1,3$ (mód 8). La segunda ecuación nos lleva de nuevo a que $p \equiv 1,5$ (mód 8), como en el primer caso. Por consiguiente, para que los cuatro puntos restantes de la tabla estén en la imagen es necesario que $p \equiv 1$ (mód 8).

De aquí se sigue inmediatamente el teorema. ∎

Como consecuencia:

Teorema 8.5 *Si p es un primo $p \equiv 3$ (mód 8) entonces p no es congruente.*

Con esto ya tenemos casi justificado por completo que los primeros números congruentes son

$$5, \quad 6, \quad 7, \quad 13, \quad 14, \quad 15.$$

En realidad nos falta probar que 7 y 13 son congruentes. Para $n = 7$ no es difícil encontrar el punto $P = (25, 120)$ en la curva $Y^2 = X^3 - 49X$. El 13 es ligeramente más complicado. Vamos a ver cómo encontrar un punto racional en la curva correspondiente.

Ejemplo Consideremos un primo $p \equiv 5$ (mód 8) y vamos a buscar puntos racionales en la curva $Y^2 = X^3 - pX$ distintos de los cuatro puntos de torsión. En la prueba de 8.4 hemos visto que $p \equiv 5$ (mód 8) era una condición necesaria para que el par $(b_1, b_2) = (1,-1)$ se correspondiera con un punto de la curva. Las ecuaciones para este punto son

$$z_1^2 + z_2^2 = p, \qquad z_1^2 + z_3^2 = 2p,$$

que equivalen a

$$p = z_1^2 + z_2^2, \qquad z_3^2 = z_1^2 + 2z_2^2.$$

La segunda ecuación corresponde a una cónica proyectiva. Obviamente no contiene puntos racionales con $z_1 = 0$, luego no perdemos soluciones si deshomogeneizamos respecto de z_1 y pasamos a la ecuación afín $z_3^2 - 2z_2^2 = 1$.

Un punto racional es $(z_2, z_3) = (0, 1)$, que no nos sirve, pues no cumple la primera ecuación, pero podemos usarlo para parametrizar las demás soluciones racionales según explicamos en la introducción.

Las rectas que pasan por $(0, 1)$ son de la forma $z_3 = tz_2 + 1$. El segundo punto de intersección con la cónica es

$$(z_2, z_3) = \left(\frac{2t}{2-t^2}, \frac{2+t^2}{2-t^2}\right).$$

Haciendo $t = r/s$, con $(r, s) = 1$, vemos que si la curva tiene un punto racional, han de existir enteros r y s tales que

$$\frac{1}{z_1}(z_1, z_2, z_3) = \left(1, \frac{2t}{2-t^2}, \frac{2+t^2}{2-t^2}\right) = \left(1, \frac{2rs}{2s^2-r^2}, \frac{2s^2+r^2}{2s^2-r^2}\right),$$

luego

$$z_1 = \lambda(2s^2 - r^2), \quad z_2 = \lambda 2rs, \quad z_3 = \lambda(2s^2 + r^2),$$

para cierto $\lambda \in \mathbb{Q}$.

Veamos que $\lambda = 1/n$, con $n \in \mathbb{Z}$. Si $\lambda = m/n$, con $(m, n) = 1$, entonces m divide a los enteros nz_1 y nz_2 (son enteros por las expresiones dadas por los miembros derechos), pero

$$p = \frac{(nz_1)^2 + (nz_2)^2}{n^2},$$

luego $m^2 \mid p$, lo que sólo es posible si $m = \pm 1$.

En conclusión, si encontramos enteros r, s, n tales que $(r, s) = 1$ y

$$p = z_1^2 + z_2^2 = \frac{(2s^2 - r^2)^2 + 4r^2s^2}{n^2} = \frac{4s^4 + r^4}{n^2},$$

entonces la curva contendrá un punto racional determinado por

$$x = b_1 z_1^2 + e_1 = \frac{(2s^2 - r^2)^2}{n^2} - p.$$

Ahora es fácil programar un ordenador para que recorra los pares (r, s) y compruebe si $(4s^4 + r^4)/p$ es un entero cuadrado perfecto (que será n^2). La tabla siguiente contiene los resultados para los primos $p \equiv 5$ (mód 8) menores que 150:

p	r	s	n	$x(P)$
5	1	1	1	-4
13	1	3	5	$-\frac{36}{25}$
29	7	5	13	$-\frac{4.900}{169}$
37	1	21	145	$-\frac{1.764}{21.025}$
53	119	143	5.945	$-\frac{1.158.313.156}{35.343.025}$
61	41	39	445	$-\frac{10.227.204}{198.025}$
101	97	325	21.041	$-\frac{42.630.008.164}{442.723.681}$
109	6	7	10	$-\frac{1.764}{25}$
149	14	17	50	$-\frac{56.644}{625}$

Observamos que casi todas las soluciones se obtienen con valores relativamente pequeños de r y s. ∎

Ejercicio: Encontrar primos $p \equiv 7$ (mód 8) que sean números congruentes.

8.2 Los grupos de Selmer y Tate-Shafarevich

En esta sección veremos una técnica para calcular el rango de una curva elíptica que generaliza a la que acabamos de describir, al tiempo que nos permitirá comprender mejor el problema.

Sea K un cuerpo numérico, $\mathfrak{p}$ un divisor primo de K, sea $K_\mathfrak{p}$ la compleción correspondiente y $\overline{K}_\mathfrak{p}$ una clausura algebraica. Fijemos un K-monomorfismo $i : \mathbb{A} \longrightarrow \overline{K}_\mathfrak{p}$, que a su vez induce inclusiones $i : \mathrm{P}^n(\mathbb{A}) \longrightarrow \mathrm{P}^n(\overline{K}_\mathfrak{p})$.

Consideremos una curva elíptica E/K, que podemos suponer definida por una ecuación de Weierstrass, la cual define también una curva elíptica $E/K_\mathfrak{p}$. La inclusión i en $\mathrm{P}^2(\mathbb{A})$ se restringe a un monomorfismo $i : E/K \longrightarrow E/K_\mathfrak{p}$.

Observemos que este monomorfismo depende de i, pero la curva $E/K_\mathfrak{p}$ está completamente determinada por E/K.

Si C/K es un espacio homogéneo para E/K, las ecuaciones de C definen también un conjunto algebraico $C/K_\mathfrak{p}$. Las funciones racionales que definen un isomorfismo (sobre $\mathbb{A}$) $C \longrightarrow E$ determinan también una aplicación birracional entre las curvas sobre $\overline{K}_\mathfrak{p}$, luego $C/K_\mathfrak{p}$ es una curva proyectiva de género 1. Por el teorema de los ceros de Hilbert, si C no tiene singularidades sobre $\mathbb{A}$, tampoco las puede tener sobre $\overline{K}_\mathfrak{p}$, luego $C/K_\mathfrak{p}$ es regular e isomorfa a $E/K_\mathfrak{p}$. Más aún, las funciones racionales que definen la acción $C \times E \longrightarrow C$ sobre $\mathbb{A}$, también definen una acción sobre $\overline{K}_\mathfrak{p}$ (los axiomas de espacio homogéneo equivalen a que las funciones racionales que definen la acción cumplan ciertas relaciones, y esto no depende del cuerpo de coeficientes que consideremos). En definitiva, vemos que cada espacio homogéneo C/K determina un espacio homogéneo $C/K_\mathfrak{p}$ de modo que el diagrama siguiente conmuta:

$$\begin{array}{ccc} C(\mathbb{A}) \times E(\mathbb{A}) & \longrightarrow & C(\mathbb{A}) \\ {\scriptstyle i\times i}\downarrow & & \downarrow{\scriptstyle i} \\ C(\overline{K}_\mathfrak{p}) \times E(\overline{K}_\mathfrak{p}) & \longrightarrow & C(\overline{K}_\mathfrak{p}) \end{array}$$

De nuevo, la curva $C/K_\mathfrak{p}$ y la aplicación $C\times E \longrightarrow C$ (sobre $\overline{K}_\mathfrak{p}$) no dependen de la elección de i.

Es claro que una equivalencia entre espacios homogéneos sobre $\overline{k}$ induce una equivalencia entre los correspondientes espacios sobre $\overline{k}_\mathfrak{p}$, luego tenemos una aplicación

$$\mathrm{WC}(E/K) \longrightarrow \mathrm{WC}(E/K_\mathfrak{p})$$

independiente de i. Recordemos ahora el teorema 2.44, según el cual tenemos isomorfismos

$$\mathrm{WC}(E/K) \cong H^1(G(\mathbb{A}/K), E(\mathbb{A})), \quad \mathrm{WC}(E/K_\mathfrak{p}) \cong H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p})).$$

Observemos que i induce un monomorfismo $\bar{\imath} : G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}) \longrightarrow G(\mathbb{A}/K)$ (si identificamos i con una inclusión, se trata de la restricción a $\mathbb{A}$). Esto nos

permite definir un homomorfismo de grupos

$$H^1(G(\mathbb{A}/K), E(\mathbb{A})) \longrightarrow H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p}))$$

dado por $[\{\xi_\sigma\}_\sigma] \mapsto [\{i(\xi_{\bar{\imath}(\sigma)}\}_\sigma]$. (Se comprueba inmediatamente que está bien definido.) Además, el diagrama siguiente es conmutativo:

$$\begin{array}{ccc} \mathrm{WC}(E/K) & \longrightarrow & \mathrm{WC}(E/K_\mathfrak{p}) \\ \downarrow & & \downarrow \\ H^1(G(\mathbb{A}/K), E(\mathbb{A})) & \longrightarrow & H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p})) \end{array}$$

En efecto: si C/K es un espacio homogéneo para E/K y $p \in C(\mathbb{A})$, entonces su imagen en $H^1(G(\mathbb{A}/K), E(\mathbb{A}))$ es la clase del cociclo $\{p^\sigma - p\}_\sigma$, que a su vez se corresponde con la clase del cociclo $\{i(p^{\bar{\imath}(\sigma)}) - i(p)\}_\sigma = \{i(p)^\sigma - i(p)\}_\sigma$, que, por otra parte, es el cociclo asociado al espacio homogéneo $C/K_\mathfrak{p}$.

Puesto que la flecha horizontal superior y las dos verticales no dependen de la elección de i, concluimos que la horizontal inferior tampoco lo hace, lo cual no es evidente a priori. A partir de aquí consideraremos $\mathbb{A} \subset \overline{K}_\mathfrak{p}$, con lo que i será la inclusión e $\bar{\imath}$ la restricción.

Consideremos ahora dos curvas elípticas E/K, E'/K y una isogenia no nula $\phi : E \longrightarrow E'$ definida sobre K. Si representamos por $E[\phi]$ el núcleo de ϕ, tenemos la sucesión exacta

$$0 \longrightarrow E[\phi] \longrightarrow E \xrightarrow{\phi} E' \longrightarrow 0,$$

de la que derivamos la sucesión exacta larga

$$0 \longrightarrow E(K)[\phi] \longrightarrow E(K) \xrightarrow{\phi} E'(K)$$

$$\xrightarrow{\delta} H^1(G(\mathbb{A}/K), E[\phi]) \longrightarrow H^1(G(\mathbb{A}/K), E) \longrightarrow H^1(G(\mathbb{A}/K), E') \longrightarrow \cdots$$

De aquí obtenemos a su vez la sucesión exacta corta

$$0 \longrightarrow E'(K)/\phi[E(K)] \xrightarrow{\delta} H^1(G(\mathbb{A}/K), E[\phi]) \longrightarrow H^1(G(\mathbb{A}/K), E)[\phi] \longrightarrow 0,$$

donde el último grupo es el núcleo de la aplicación inducida por ϕ sobre el grupo de cohomología. Representaremos por $\mathrm{WC}(E/K)[\phi]$ el subgrupo correspondiente del grupo de Weil-Châtelet de E/K. Tenemos así una sucesión exacta

$$0 \longrightarrow E'(K)/\phi[E(K)] \xrightarrow{\delta} H^1(G(\mathbb{A}/K), E[\phi]) \longrightarrow \mathrm{WC}(E/K)[\phi] \longrightarrow 0.$$

Esta sucesión no es útil para estudiar el grupo $E'(K)/\phi[E(K)]$ porque el grupo $H^1(G(\mathbb{A}/K), E[\phi])$ resulta ser infinito. Sin embargo, podemos sumergir a $E'(K)/\phi[E(K)]$ en un grupo menor considerando el homomorfismo

$$\mathrm{WC}(E/K) \longrightarrow \prod_\mathfrak{p} \mathrm{WC}(E/K_\mathfrak{p}).$$

Definición 8.6 En las condiciones anteriores, llamaremos *grupo de Selmer* de la isogenia ϕ al núcleo $S^\phi(E/K)$ del homomorfismo

$$H^1(G(\mathbb{A}/K), E[\phi]) \longrightarrow \prod_{\mathfrak{p}} \mathrm{WC}(E/K_\mathfrak{p}).$$

Llamaremos *grupo de Tate-Shafarevich* de la curva E/K al núcleo $Ш(E/K)$ del homomorfismo

$$\mathrm{WC}(E/K) \longrightarrow \prod_{\mathfrak{p}} \mathrm{WC}(E/K_\mathfrak{p}).$$

(La letra cirílica Ш es la inicial de Shafarevich.) De las propias definiciones se sigue inmediatamente que tenemos una sucesión exacta

$$0 \longrightarrow E'(K)/\phi[E(K)] \xrightarrow{\delta} S^\phi(E/K) \longrightarrow Ш(E/K)[\phi] \longrightarrow 0, \tag{8.3}$$

donde $Ш(E/K)[\phi] = Ш(E/K) \cap \mathrm{WC}(E/K)[\phi]$.

La ventaja es que —como veremos enseguida— el grupo de Selmer es finito. Antes observemos que el grupo de Tate-Shafarevich tiene una interpretación muy simple: sus elementos representan a las clases de K-isomorfía de espacios homogéneos C/K que tienen puntos racionales en todas las compleciones $C/K_\mathfrak{p}$. Así pues, se cumple $Ш(E/K) = 0$ si y sólo si el principio de Hasse es válido para los espacios homogéneos sobre E/K, es decir, si la existencia de un punto racional en un espacio homogéneo para E/K equivale a que los haya en todas sus compleciones.

Definición 8.7 Sea K un cuerpo numérico y M un grupo abeliano finito sobre el que actúa (continuamente) el grupo $G(\mathbb{A}/K)$. Sea $\mathfrak{p}$ un divisor primo de K y llamemos $I_\mathfrak{p} = G(\overline{K}_\mathfrak{p}/K_{\mathfrak{p}\mathrm{nr}}) \leq G(\mathbb{A}/K)$ al grupo de inercia de $\mathfrak{p}$. (Aquí identificamos los elementos de $G(\overline{K}_\mathfrak{p}/K_\mathfrak{p})$ con sus restricciones a $\mathbb{A}$.) Diremos que una clase $[\{\xi_\sigma\}_\sigma] \in H^1(G(\mathbb{A}/K), M)$ es *no ramificada* en $\mathfrak{p}$ si su restricción a $H^1(I_\mathfrak{p}, M)$ es trivial. Si S es un conjunto de primos de K, definimos

$$H^1(G(\mathbb{A}/K), M; S)\{\alpha \in H^1(G(\mathbb{A}/K), M) \mid \alpha \text{ es no ramificada en todo } \mathfrak{p} \notin S\}.$$

El primer paso para demostrar la finitud del grupo de Selmer es el siguiente:

Teorema 8.8 *Sea $\phi : E/K \longrightarrow E'/K$ una isogenia de grado m definida sobre un cuerpo numérico K y sea S el conjunto (finito) de los divisores primos $\mathfrak{p}$ de K que cumplen alguna de las propiedades siguientes:*

a) $\mathfrak{p}$ es arquimediano,

b) $\mathfrak{p} \mid m$,

c) E/K tiene mala reducción en $\mathfrak{p}$.

Entonces, para todo primo $\mathfrak{p} \notin S$, una clase de $H^1(G(\mathbb{A}/K), E[\phi])$ es trivial en $H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p}))$ si y sólo si es no ramificada en $\mathfrak{p}$. En particular

$$S^\phi(E/K) \leq H^1(G(\mathbb{A}/K), E[\phi]; S).$$

DEMOSTRACIÓN: Tomemos un primo $\mathfrak{p} \notin S$ y una clase de cohomología $\alpha = [\{\xi_\sigma\}_\sigma] \in H^1(G(\mathbb{A}/K), E[\phi])$ no ramificada en $\mathfrak{p}$. Esto significa que existe un punto $P \in E[\phi]$ tal que $\xi_\sigma = P^\sigma - P$, para todo $\sigma \in I_\mathfrak{p}$. Fijemos una extensión normal L/K tal que $E[\phi] \subset E[m] \subset E(L)$ y de modo que ξ_σ dependa únicamente de la clase de σ en $G(L/K)$. Sea $\mathfrak{P}$ un primo de L que divida a $\mathfrak{p}$.

Los automorfismos $\sigma \in I_\mathfrak{p}$ se caracterizan por que inducen la identidad en el cuerpo de restos l de $L_\mathfrak{P}$. Por lo tanto, cumplen que

$$\tilde{\xi}_\sigma = \tilde{P}^\sigma - \tilde{P} = \tilde{O}.$$

Por otra parte, como $\mathfrak{P} \nmid m$, el teorema 6.19 nos da que la reducción módulo $\mathfrak{P}$ es inyectiva sobre $E(L_\mathfrak{P})[m]$, luego $\xi_\sigma = O$ para todo $\sigma \in I_\mathfrak{p}$. De aquí se sigue a su vez que si $\tau \in G(\mathbb{A}/K)$ es arbitrario, entonces

$$\xi_{\sigma\tau} = \xi_\sigma^\tau + \xi_\tau = \xi_\tau,$$

es decir, que ξ_σ sólo depende de la clase de σ módulo $I_\mathfrak{p}$. Por consiguiente, si dos automorfismos $\sigma, \tau \in G(\overline{K}_\mathfrak{p}/k_\mathfrak{p})$ inducen el mismo automorfismo en $G(\overline{k}/k)$, resulta que $\xi_\sigma = \xi_\tau$. Esto nos permite definir $[\{\tilde{\xi}_\sigma\}_\sigma] \in H^1(G(\overline{k}/k), \tilde{E})$ mediante $\tilde{\xi}_{\tilde{\sigma}} = \widetilde{\xi_\sigma}$.

Ahora bien, $\mathrm{WC}(\tilde{E}/k) = 0$, porque el cuerpo de restos k es finito. Por el teorema 2.44, también $H^1(G(\overline{k}/k), \tilde{E}) = 0$, lo que significa que existe un punto $p \in \tilde{E}(k)$ tal que $\tilde{\xi}_\sigma = p^\sigma - p$, para todo $\sigma \in G(\overline{K}_\mathfrak{p}/K_\mathfrak{p})$.

Como $\xi_\sigma \in E[\phi] \subset E[m]$, vemos que $(mp)^\sigma = mp$, para todo $\sigma \in G(\overline{k}/k)$, luego $q = mp \in E(k)$. Tomemos un punto $Q \in E(K)$ tal que $\tilde{Q} = q$ y fijemos un $R \in E(\overline{K})$ tal que $mR = Q$.

De nuevo por 6.19, cuando T recorre el grupo $E[m]$, los puntos $\tilde{R}+\tilde{T}$ recorren m puntos distintos de $\tilde{E}$ tales que $m(\tilde{R}+\tilde{T}) = \tilde{Q} = q$. Sólo hay m puntos que cumplen esto, y uno de ellos es p, luego eligiendo T adecuadamente se cumple que $P = R + T \in E(\overline{K})$ verifica $\tilde{P} = p$ y además $mP = Q \in E(K)$.

Así tenemos que $\tilde{\xi}_\sigma = \widetilde{P^\sigma - P}$, para todo $\sigma \in G(\overline{K}_\mathfrak{p}/K_\mathfrak{p})$. Ahora bien, $m(P^\sigma - P) = (mP)^\sigma - mP = Q^\sigma - Q = O$, luego $\xi_\sigma - (P^\sigma - P) \in E[m]$. Aplicando 6.19 una vez más, llegamos a que $\xi_\sigma = P^\sigma - P$ para todo $\sigma \in G(\overline{K}_\mathfrak{p}/K_\mathfrak{p})$, luego α es trivial en $H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p}))$.

Recíprocamente, supongamos ahora que $\alpha = [\{\xi_\sigma\}_\sigma]$ es trivial en el grupo $H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p}))$. Esto significa que existe un punto $P \in E(\overline{K}_\mathfrak{p})$ tal que $\xi_\sigma = P^\sigma - P$ para todo $\sigma \in G(\overline{K}_\mathfrak{p}/K_\mathfrak{p})$ y, en particular, para todo $\sigma \in I_\mathfrak{p}$.

Consideremos una extensión finita L de $K_\mathfrak{p}$ tal que $P \in E(L)$ y sea $\mathfrak{P}$ el primo de L. Razonando como antes vemos que si $\sigma \in I_\mathfrak{p}$, entonces

$$\tilde{\xi}_\sigma = \tilde{P}^\sigma - \tilde{P} = \tilde{O}.$$

Por otra parte, $\xi_\sigma \in E[\phi] \subset E[m]$ y el teorema 6.19 nos da que $\xi_\sigma = O$. Ciertamente entonces α es no ramificada en $\mathfrak{p}$. ∎

Más precisamente, acabamos de probar que $S^\phi(E/K)$ está formado por las clases $\alpha \in H^1(G(\mathbb{A}/K), E[\phi]; S)$ que son triviales en $H^1(G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}), E(\overline{K}_\mathfrak{p}))$ (o,

equivalentemente, en $\mathrm{WC}(E/K_\mathfrak{p})$) para todo primo $\mathfrak{p} \in S$. En otras palabras, para comprobar si $\alpha \in S^\phi(E/K)$ sólo hay que estudiar si un número finito de espacios homogéneos son o no triviales.

Según el teorema anterior, para demostrar que el grupo de Selmer es finito basta probar que el grupo $H^1(G(\mathbb{A}/K), E[\phi]; S)$ lo es. Para ello nos basaremos en el mismo resultado de finitud que usamos para probar el teorema débil de Mordell-Weil. Es natural que así sea, pues éste se recupera haciendo $E = E'$ y tomando como ϕ la multiplicación por m.

Teorema 8.9 *Si $\phi : E/k \longrightarrow E'/k$ es una isogenia definida sobre K, entonces el grupo de Selmer $S^\phi(E/K)$ es finito.*

Demostración: Por el teorema anterior, basta probar que si S es un conjunto finito de primos de K, el grupo $H^1(G(\mathbb{A}/K), E[\phi]; S)$ es finito. Tomemos una extensión finita L de K tal que $E[\phi] \subset E(L)$ y sea S' el conjunto (finito) de los primos de L que dividen a los primos de S. La inflación y la restricción determinan una sucesión exacta

$$0 \longrightarrow H^1(G(L/K), E[\phi]) \longrightarrow H^1(G(\mathbb{A}/K), E[\phi]) \longrightarrow H^1(G(\mathbb{A}/L), E[\phi]),$$

y es fácil ver que la restricción se restringe a un homomorfismo

$$H^1(G(\mathbb{A}/K), E[\phi]; S) \longrightarrow H^1(G(\mathbb{A}/L), E[\phi]; S').$$

La sucesión exacta muestra que el núcleo de la restricción es finito, luego basta probar la finitud del grupo $H^1(G(\mathbb{A}/L), E[\phi]; S')$. Equivalentemente, podemos suponer que $E[\phi] \subset E(K)$, con lo que $G(\mathbb{A}/K)$ actúa trivialmente sobre $E[\phi]$. En tal caso

$$H^1(G(\mathbb{A}/K), E[\phi]) = \mathrm{Hom}(G(\mathbb{A}/K), E[\phi]).$$

Llamemos ahora m al exponente de $E[\phi]$ (es decir, al mínimo común múltiplo de los órdenes de los elementos de $E[\phi]$). Notemos que si $\alpha \in H^1(G(\mathbb{A}/K), E[\phi])$, su núcleo ha de ser un grupo $G(\mathbb{A}/L)$, para cierta extensión L de K tal que $G(L/K)$ es isomorfo a un subgrupo de $E[\phi]$. En particular L/K es una extensión finita abeliana de exponente m (es decir, todos los elementos de $G(L/K)$ tienen orden divisor de m).

Además $\alpha \in H^1(G(\mathbb{A}/K), E[\phi]; S)$ si y sólo si $I_\mathfrak{p} \leq G(\mathbb{A}/L)$ para todo primo $\mathfrak{p} \notin S$. Si $\mathfrak{P}$ es un divisor de $\mathfrak{p}$ en L, esto equivale a que

$$I_\mathfrak{p} \leq G(\mathbb{A}/L) \cap G(\overline{K}_\mathfrak{p}/K_\mathfrak{p}) = G(\overline{K}_\mathfrak{p}/L_\mathfrak{P})$$

o, lo que es lo mismo, a que $L_\mathfrak{P} \subset K_{\mathfrak{p}\mathrm{nr}}$. En definitiva, $\alpha \in H^1(G(\mathbb{A}/K), E[\phi]; S)$ si y sólo si la extensión L/K es no ramificada fuera de S.

Sea L la mayor extensión abeliana de K de exponente m no ramificada fuera de S. En la prueba del teorema débil de Mordell-Weil hemos visto que

la extensión L/K es finita, y acabamos de demostrar que $H^1(G(\mathbb{A}/K), E[\phi]; S)$ está contenido en la imagen de la inflación

$$\mathrm{Hom}(G(L/K), E[\phi]) \longrightarrow H^1(G(\mathbb{A}/K), E[\phi]; S).$$

(De hecho, es fácil ver que $H^1(G(\mathbb{A}/K), E[\phi]; S)$ es exactamente la imagen de la inflación, y que ésta es un isomorfismo.)

Concluimos que $H^1(G(\mathbb{A}/K), E[\phi]; S)$ es finito, como queríamos probar. ∎

Ejemplo Vamos a reinterpretar en los términos que estamos considerando en esta sección el caso tratado en la sección 8.1, es decir, el de una curva elíptica E/K tal que $E[m] \subset E(K)$ para cierto natural m. No daremos los detalles porque son laboriosos y no nos aportarían nada nuevo. En la sección siguiente desarrollaremos con todo detalle un caso más general.

Tomamos como $\phi : E \longrightarrow E$ la multiplicación por m, con lo que

$$S^m(E/K) \leq H^1(G(\mathbb{A}/K), E[m]; S),$$

donde S es el conjunto formado por los primos arquimedianos de K, los primos donde E tiene mala reducción y los que dividen a m. Por otra parte,

$$H^1(G(\mathbb{A}/K), E[m]) = \mathrm{Hom}(G(\mathbb{A}/K), E[m])$$

y es fácil ver que el monomorfismo (8.1) inducido por el producto de Kummer es precisamente el homomorfismo δ que hemos considerado en la discusión del grupo de Selmer. Fijada una base de $E[m]$, el producto de Weil e_m y el isomorfismo (8.2) determinan isomorfismos

$$\mathrm{Hom}(G(\mathbb{A}/k), E[m]) \cong \mathrm{Hom}(G(\mathbb{A}/K), U_m \times U_m) \cong (K^*/K^{*m}) \times (K^*/K^{*m}),$$

donde $U_m \subset K$ es el grupo de las raíces m-simas de la unidad. El isomorfismo

$$E(K)/mE(K) \longrightarrow K(S,m) \times K(S,m)$$

inducido por la forma bilineal b considerada en 8.1 no es más que la composición con δ de estos isomorfismos. Se comprueba así mismo que el grupo $H^1(G(\mathbb{A}/K), E[m]; S)$ se corresponde precisamente con $K(S,m) \times K(S,m)$.

Consideremos ahora el caso $m = 2$. Según acabamos de comentar, el grupo $K(S,2) \times K(S,2)$ no es más que una representación sencilla del grupo $H^1(G(\mathbb{A}/K), E[2]; S)$. A través de los isomorfismos naturales que hemos considerado, cada par $(b_1, b_2) \in K(S,2) \times K(S,2)$ se corresponde con un elemento de $H^1(G(\mathbb{A}/K), E(K))$, el cual determina un espacio homogéneo C/K para E/K. No es trivial, pero puede demostrarse que el espacio C/K es trivial (esto es, tiene un punto racional) en una compleción $K_\mathfrak{p}$ si y sólo si las ecuaciones indicadas en el teorema 8.2 tienen solución en $K_\mathfrak{p}$.

De este modo, el grupo de Selmer $S^2(E/K)$ resulta ser isomorfo al subgrupo de $K(S,2) \times K(S,2)$ formado por los pares cuyas ecuaciones asociadas tienen un

punto racional en todas las compleciones de K (más los asociados a los puntos de orden 2). En realidad basta considerar las compleciones respecto de los primos de S.

En todos los ejemplos que hemos considerado en la sección 8.1, el grupo de Selmer $S^2(E/\mathbb{Q})$ ha coincidido con la imagen de $E(\mathbb{Q})/2E(\mathbb{Q})$ en el grupo $\mathbb{Q}(S,2)\times\mathbb{Q}(S,2)$. Dicho de otro modo: para todos los pares que no tenían antiimagen en $E(\mathbb{Q})/2E(\mathbb{Q})$ hemos encontrado una compleción $\mathbb{Q}_p$ (tal vez con $p=\infty$) en la que el espacio homogéneo correspondiente no tenía puntos racionales. Esto significa que en todos ellos se cumplía $Ш(E/\mathbb{Q})[2]=0$. ∎

El teorema siguiente nos permite calcular la antiimagen de un elemento de $S^\phi(E/K)$ en $E'(K)/\phi[E(K)]$ cuando existe, es decir, cuando el espacio homogéneo correspondiente es trivial.

Teorema 8.10 *Consideremos una isogenia $\phi: E/K\longrightarrow E'/K$ definida sobre K, sea $[\{\xi_\sigma\}_\sigma]\in H^1(G(\mathbb{A}/K),E[\phi])$ y sea C/K un espacio homogéneo para E/K que represente a la clase de $\mathrm{WC}(E/K)$ correspondiente a $[\{\xi_\sigma\}_\sigma]$. Sea $\theta: C\longrightarrow E$ un isomorfismo tal que $\theta^{-1}\circ\theta^\sigma=\tau_{\xi_\sigma}$. Entonces la aplicación $\theta\circ\phi: C\longrightarrow E'$ está definida sobre K y, si $P\in C(K)$, entonces la imagen de $(\theta\circ\phi)(P)\in E'(K)$ por el homomorfismo $\delta: E'(K)\longrightarrow S^\phi(E/K)$ es $[\{\xi_\sigma\}_\sigma]$.*

Demostración: Fijemos $\sigma\in G(\overline{K}/K)$ y $P\in C$. Entonces

$$(\phi(\theta(P))^\sigma=\phi(\theta^\sigma(P^\sigma))=\phi(\theta(P^\sigma)+\xi_\sigma)=\phi(\theta(P^\sigma)),$$

pues $\xi_\sigma\in E[\phi]$. Esto prueba que $(\theta\circ\phi)^\sigma=\theta\circ\phi$. Si $P\in C(K)$, entonces

$$\delta(\phi(\theta(P)))_\sigma=\theta(P)^\sigma-\theta(P)=\theta^\sigma(P)-\theta(P)=\theta(P)+\xi_\sigma-\theta(P)=\xi_\sigma.$$

∎

Los resultados que hemos obtenido nos permiten calcular el orden del grupo $E'(K)/\phi[E(K)]$ de forma similar a como calculábamos el de $E(K)/mE(K)$ en la sección 8.1. Aparentemente no ganamos nada generalizando, pues necesitamos calcular este orden cuando ϕ es la multiplicación por m, para algún m. Ahora bien, si m es el grado de ϕ y consideramos la isogenia dual $\hat\phi: E'\longrightarrow E$, también podemos calcular el orden de $E(K)/\hat\phi[E'(K)]$, y la sucesión exacta

$$0\longrightarrow\frac{E'(K)[\hat\phi]}{\phi[E(K)[m]]}\longrightarrow\frac{E'(K)}{\phi[E(K)]}\xrightarrow{\hat\phi}\frac{E(K)}{mE(K)}\longrightarrow\frac{E(K)}{\hat\phi[E'(K)]}\longrightarrow 0 \tag{8.4}$$

nos permite calcular el orden de $E(K)/mE(K)$.

En la sección siguiente detallamos los cálculos en un caso particular. Para terminar citaremos la siguiente conjetura:

Conjetura *Si E/K es una curva elíptica sobre un cuerpo numérico K, entonces el grupo $Ш(E/K)$ es finito.*

No vamos a justificarlo, pero si la conjetura fuera cierta sería posible dar un algoritmo para calcular (al menos en teoría) el rango de cualquier curva elíptica. En realidad bastaría con que $Ш(E/K)$ no tuviera elementos infinitamente divisibles entre 2.

8.3 Curvas con un punto de orden 2

Consideremos una curva elíptica E/K sobre un cuerpo numérico K con un punto racional T de orden 2. Tomando una ecuación de Weierstrass de tipo b, las coordenadas de T serán de la forma $(x, 0)$ y, aplicando una traslación, podemos suponer que $T = (0, 0)$, con lo que la ecuación será

$$E:\ Y^2 = X^3 + aX^2 + bX, \qquad a, b \in K.$$

Su discriminante es $\Delta = 16b^2b'$, donde $b' = a^2 - 4b$. En la página 61 hemos visto que si llamamos

$$E':\ Y^2 = X^3 - 2aX^2 + b'X,$$

entonces la isogenia $\phi : E \longrightarrow E'$ dada por

$$\phi(X, Y) = \left(\frac{Y^2}{X^2}, \frac{Y(b - X^2)}{X^2}\right)$$

tiene núcleo $E[\phi] = \{O, T\}$ y la isogenia dual es

$$\hat{\phi}(X, Y) = \left(\frac{Y^2}{4X^2}, \frac{Y(b' - X^2)}{8X^2}\right).$$

Sea S el conjunto de los primos de K que dividen a 2Δ más los primos arquimedianos. Puesto que $E[\phi]$ y $\{\pm 1\}$ son isomorfos como $G(\mathbb{A}/K)$-módulos (ambos son triviales), tenemos que

$$H^1(G(\mathbb{A}/K), E[\phi]) \cong H^1(G(\mathbb{A}/K), \{\pm 1\}) \cong K^*/K^{*2}.$$

Veamos que este isomorfismo se restringe a

$$H^1(G(\mathbb{A}/K), E[\phi]; S) \cong K(S, 2).$$

En efecto: tomemos $\alpha \in H^1(G(\mathbb{A}/K), E[\phi])$, digamos $\alpha \neq 0$. Entonces, el núcleo de α es un grupo $G(\mathbb{A}/L)$ de índice 2 en $G(\mathbb{A}/K)$, es decir, tal que $|L : K| = 2$. Digamos que $L = K(\sqrt{d}\,)$, con $d \in K^*$. Entonces

$$\alpha(\sigma) = \begin{cases} O & \text{si } \sigma(\sqrt{d}\,) = \ \sqrt{d}, \\ T & \text{si } \sigma(\sqrt{d}\,) = -\sqrt{d}. \end{cases} \tag{8.5}$$

Es claro entonces que la imagen de α en K^*/K^{*2} es $[d]$. En la prueba del teorema 8.9 hemos visto que $\alpha \in H^1(G(\mathbb{A}/K), E[\phi]; S)$ si y sólo si la extensión L/K es no ramificada fuera de S. Ahora basta probar que L/K es no ramificada en un primo no arquimediano $\mathfrak{p}$ tal que $\mathfrak{p} \nmid 2$ si y sólo si $2 \mid v_\mathfrak{p}(d)$.

En efecto: sea $\mathfrak{P}$ un primo de L que divida a $\mathfrak{p}$. La extensión L/K es no ramificada en $\mathfrak{P}$ si y sólo si lo es $L_\mathfrak{P}/K_\mathfrak{p}$. Si $2 \nmid v_\mathfrak{p}(d)$, entonces

$$2v_\mathfrak{P}(\sqrt{d}\,) = v_\mathfrak{P}(d) = e(\mathfrak{P}/\mathfrak{p})v_\mathfrak{p}(d),$$

luego $2 \mid e(\mathfrak{P}/\mathfrak{p})$ y, por consiguiente, L/K es ramificada en $\mathfrak{P}$. Recíprocamente, si $2 \mid v_\mathfrak{p}(d)$, multiplicando d por el cuadrado de un primo de $K_\mathfrak{p}$ podemos suponer que d es una unidad de $K_\mathfrak{p}$, en cuyo caso, la base $\{1, \sqrt{d}\}$ tiene discriminante $D = 4d$, luego $v_\mathfrak{p}(D) = 0$ y el discriminante de la extensión divide a D, luego no es divisible entre $\mathfrak{p}$, lo que implica que $\mathfrak{p}$ no se ramifica. ∎

Así pues, cada $d \in K(S,2)$ se corresponde con el elemento del grupo de cohomología $H^1(G(\mathbb{A}/K), E[\phi]; S)$ determinado por el cociclo (8.5). Para saber si este elemento está en el grupo de Selmer $S^\phi(E/K)$ hemos de calcular su espacio homogéneo correspondiente en $\mathrm{WC}(E/K)$.

Para ello consideramos la curva

$$C' : dW^2 = d^2 - 2adZ^2 + (a^2 - 4b)Z^4.$$

Es fácil ver que no tiene puntos singulares en el plano afín, pues ello es equivalente a que el polinomio de miembro derecho tenga sus raíces simples, y ciertamente las tiene. Sin embargo C' tiene un punto singular en el infinito, a saber $(U, W, Z) = (0, 1, 0)$. Por el teorema 1.18 existe una curva proyectiva regular C/K y una aplicación birracional $r : C \longrightarrow C'$ definida sobre K. Vamos a probar que C/K es el espacio homogéneo que buscamos.

Definimos una aplicación racional $\theta : E \longrightarrow C'$ mediante

$$\theta(x,y) = (z,w) = \left(\sqrt{d}\,x/y, \sqrt{d}\,(x - b/x)(x/y)^2\right).$$

Como $x/y = xy/y^2 = y/(x^2 + ax + b)$, otra expresión para θ es

$$\theta(x,y) = \left(\frac{\sqrt{d}\,y}{x^2 + ax + b}, \frac{\sqrt{d}\,(x^2 - b)}{x^2 + ax + b}\right).$$

Veamos que la imagen está ciertamente en C':

$$d(w/z^2)^2 = (x - b/x)^2 = (x + b/x)^2 - 4b = ((y/x)^2 - a)^2 - 4b = (d/z^2 - a)^2 - 4b.$$

Multiplicando por z^4 queda

$$dw^2 = (d - az^2)^2 - 4b = d^2 - 2adz^2 + (a^2 - 4b)z^4.$$

Se cumple que θ es birracional, pues

$$\sqrt{d}\,w/z^2 = (x - b/x) = 2x - (x + b/x) = 2x - ((y/x)^2 - a) = 2x - (d/z^2 - a).$$

De aquí obtenemos x en términos de z y w, y entonces $y = \sqrt{d}\,x/z$. Explícitamente:

$$\theta^{-1}(z,w) = \left(\frac{\sqrt{d}\,w - az^2 + d}{2z^2}, \frac{dw - a\sqrt{d}\,z^2 + d\sqrt{d}}{2z^3}\right).$$

Vemos, pues, que $\theta \circ r^{-1} : E \longrightarrow C$ es también birracional, y como ambas curvas son regulares, de hecho es un isomorfismo. Con esto tenemos que C/K es una curva conjugada con E/K. Su cociclo asociado es

$$\xi_\sigma = (\theta \circ r^{-1}) \circ (r \circ \theta^{-1})^\sigma = \theta \circ (\theta^{-1})^\sigma.$$

Obviamente, si $\sigma(\sqrt{d}\,) = \sqrt{d}$, entonces $\xi_\sigma = 1 = \tau_O$. Por el contrario, si $\sigma(\sqrt{d}\,) = -\sqrt{d}$, entonces

$$\begin{aligned}
\xi_\sigma(x,y) &= (\theta^{-1})^\sigma(z,w) = \left(\frac{-\sqrt{d}\,w - az^2 + d}{2z^2}, \frac{dw + a\sqrt{d}\,z^2 - d\sqrt{d}}{2z^3}\right) \\
&= \left(-\frac{\sqrt{d}\,w}{z^2} + x, \frac{2a\sqrt{d}\,z^2 - 2d\sqrt{d}}{2z^3} + y\right) \\
&= \left(\frac{b}{x}, \frac{y(ax^2 - y^2 + x^3)}{x^3}\right) = \left(\frac{b}{x}, -\frac{by}{x^2}\right) = \tau_T(x,y).
\end{aligned}$$

Así pues, C/K se corresponde con un elemento de $H^1(G(\mathbb{A}/K), E)$, luego admite estructura de espacio homogéneo, y su cociclo asociado es el dado por (8.5), esto es, el correspondiente a d.

Siguiendo el teorema 8.10, ahora calculamos $\psi = r \circ \theta^{-1} \circ \phi : C \longrightarrow E'$. Si $r(P) = (z,w)$, entonces,

$$\psi(P) = (d/z^2, -dw/z^3).$$

El teorema siguiente recoge lo que hemos obtenido y algunos hechos adicionales:

Teorema 8.11 *Sean E/K y E'/K las curvas elípticas dadas por las ecuaciones*

$$E:\ Y^2 = X^3 + aX^2 + bX, \qquad E' : Y^2 = X^3 - 2aX + (a^2 - 4b)X$$

y sea $\phi : E \longrightarrow E'$ la isogenia dada por

$$\phi(X,Y) = \left(\frac{Y^2}{X^2}, \frac{Y(b - X^2)}{X^2}\right).$$

Sea S el conjunto de los primos arquimedianos de K más los divisores del discriminante $\Delta = 16b^2(a^2 - 4b)$ más los divisores de 2. Entonces existe una sucesión exacta

$$0 \longrightarrow E'(K)/\phi[E(K)] \xrightarrow{\delta} K(S,2) \longrightarrow \mathrm{WC}(E/K)[\phi]$$

dada por

$$\delta([P]) = \begin{cases} [1] & \text{si } P = O, \\ [a^2 - 4b] & \text{si } P = (0,0), \\ [x(P)] & \text{en otro caso,} \end{cases} \qquad [d] \mapsto [C_d],$$

donde C_d es la regularización de la curva

$$C'_d:\ dW^2 = d^2 - 2adZ^2 + (a^2 - 4b)Z^4.$$

Se cumple que

$$S^\phi(E/K) \cong \{[d] \in K(S,2) \mid C_d(K_\mathfrak{p}) \neq \varnothing \text{ para todo } \mathfrak{p} \in S\}.$$

La aplicación $\psi : C'_d \longrightarrow E'$ *dada por*

$$\psi(z,w) = (d/z^2, -dw/z^3)$$

cumple que si $P \in C'_d(K)$ *es regular, entonces* $\delta(\psi(P)) = [d]$.

El rango r *de las curvas cumple*

$$2^{r+2} = |E'(K) : \phi[E(K)]| \cdot |E(K) : \hat{\phi}[E'(K)]|.$$

DEMOSTRACIÓN: Sólo falta probar que δ se calcula como indica el enunciado y la fórmula para el rango. Si $P = (x,y) \neq O, (0,0)$, para calcular $\delta([P])$ tomamos $Q \in E$ tal que $\phi(Q) = P$, consideramos el cociclo $\xi_\sigma = Q^\sigma - Q$ y calculamos su imagen en K^*. Hemos de probar que es x.

Sea $Q = (u,v)$. Entonces $x = v^2/u^2$, $y = v(b-u^2)/u^2$. Pongamos que $\sqrt{x}^\sigma = \epsilon\sqrt{x}$, con $\epsilon = \pm 1$. Entonces $(v/u)^\sigma = \epsilon v/u$. Como $P \in E'(K)$, tenemos que $y^\sigma = y$, luego

$$(v(b-u^2)/u^2)^\sigma = \epsilon(v/u)((b-u^2)/u)^\sigma = v(b-u^2)/u^2.$$

Concluimos que

$$((b-u^2)/u)^\sigma = \epsilon(b-u^2)/u.$$

Como (u,v) cumple la ecuación de E, vemos que $x = v^2/u^2 = (u^2+b)/u+a$, luego

$$((u^2+b)/u)^\sigma = (u^2+b)/u.$$

Restando las dos últimas ecuaciones vemos que

$$(2u)^\sigma = \frac{(1-\epsilon)b + (1+\epsilon)u^2}{u}.$$

Así pues,

$$u^\sigma = \begin{cases} u & \text{si } \sigma(\sqrt{x}\,) = \sqrt{x}, \\ b/u & \text{si } \sigma(\sqrt{x}\,) = -\sqrt{x}. \end{cases}$$

De aquí se sigue que

$$Q^\sigma - Q = \begin{cases} O & \text{si } \sigma(\sqrt{x}\,) = \sqrt{x}, \\ (0,0) & \text{si } \sigma(\sqrt{x}\,) = -\sqrt{x}. \end{cases}$$

Por consiguiente, a través de las identificaciones naturales, $\delta([P]) = [x]$.

Si $P = (0,0)$ razonamos igualmente, sólo que ahora las antiimágenes de P son los puntos $(u,0)$, donde u varía en las raíces no nulas del polinomio $X^3 + aX^2 + bX$, es decir, en las raíces de $X^2 + aX + b$, que son

$$u = \frac{-a \pm \sqrt{a^2 - 4b}}{2}.$$

En inmediato que $Q^\sigma - Q$ es 0 si y sólo si $\sigma(\sqrt{a^2-4b}\,) = \sqrt{a^2-4b}$, de donde $\delta([P]) = [a^2-4b]$.

Vamos a calcular el rango de las curvas. Sea t el número de divisores elementales pares de $E(K)$. Es claro que $t = 1$ si $E(K)$ tiene un único punto de orden 2 y $t = 2$ si hay tres puntos de orden 2 (a su vez, esto depende de si $a^2 - 4b \in K^{*2}$). De la sucesión exacta (8.4) deducimos que

$$2^{r+t} = |E(K) : 2E(K)| = \frac{|E(K) : \hat{\phi}[E'(K)]| \cdot |E'(K) : \phi[E(K)]|}{|E'(K)[\hat{\phi}] : \phi[E(K)[2]]|}.$$

Hemos de probar que

$$|E'(K)[\hat{\phi}] : \phi[E(K)[2]]| = \begin{cases} 1 & \text{si } t = 2, \\ 2 & \text{si } t = 1. \end{cases}$$

Puesto que $|E'(K)[\hat{\phi}]| = 2$, esto equivale a que $|\phi[E(K)[2]]| = t$. Ahora bien, es claro que $|E(K)[2]| = 2t$, luego la imagen tiene orden t. ∎

Aunque nos nos va a hacer falta, no es difícil mostrar explícitamente la curva C_d. Basta tener presente el teorema siguiente:

Teorema 8.12 *Consideremos una curva C'/K dada por una ecuación*

$$W^2 = aZ^4 + bZ^2 + c,$$

donde el polinomio de la derecha tiene cuatro raíces distintas en $\mathbb{A}$. Su regularización es la curva C/K en $\mathrm{P}^3(\mathbb{A})$ dada por las ecuaciones

$$W^2 = aV^2 + bV + c, \quad V = Z^2.$$

Demostración: Es claro que las aplicaciones

$$(V, W, Z) \mapsto (W, Z) \quad \text{y} \quad (W, Z) \mapsto (Z^2, W, Z)$$

son regulares y mutuamente inversas en la parte afín de ambas curvas, luego C y C' son birracionalmente equivalentes y C es regular salvo a lo sumo en los puntos infinitos. Si homogeneizamos con una variable U y hacemos $U = 0$ vemos que dichos puntos son $(U, V, W, Z) = (0, 1, \pm\sqrt{a}, 0)$. Si deshomogeneizamos respecto de V las ecuaciones se convierten en

$$W^2 = a + bU + cU^2, \quad U = Z^2.$$

La parte afín de esta curva es isomorfa a la de $W^2 = a + bZ^2 + cZ^4$, que es regular en $(\pm\sqrt{a}, 0)$, luego C es regular. ∎

En particular, la curva C_d del teorema 8.11 viene dada por las ecuaciones

$$dW^2 = d^2 - 2adV + (a^2 - 4b)V^2, \quad V = Z^2.$$

Los puntos infinitos de C_d (que se corresponden con el punto $(0, 1, 0)$ de C'_d) son $(0, 1, \pm\sqrt{(a^2-4b)/d}, 0)$.

De todos modos, notemos que para comprobar que $C_d(K_\mathfrak{p}) = \varnothing$ basta probar que C'_d no tiene puntos racionales regulares (o sea, finitos), ya que si C_d tiene puntos en $K_\mathfrak{p}$, entonces es una curva elíptica y tiene infinitos de ellos, luego C'_d tendrá puntos racionales regulares.

Ejemplo Vamos a calcular el rango de la curva elíptica $E/\mathbb{Q}$ dada por

$$E:\ Y^2 = X^3 - 6X^2 + 17X.$$

Su discriminante es $\Delta = -2^9 \cdot 17^2$, luego $S = \{2, 17, \infty\}$ y por consiguiente $\mathbb{Q}(S,2) = \{\pm 1, \pm 2, \pm 17, \pm 34\}$. La curva isógena es

$$E' : Y^2 = X^3 + 12X^2 - 32X,$$

y, para cada $d \in \mathbb{Q}(S,2)$, el espacio homogéneo a estudiar es

$$C'_d:\ dW^2 = d^2 + 12dZ^2 - 32Z^4.$$

En primer lugar observamos que $\delta(0,0) = -32 \equiv -2 \pmod{\mathbb{Q}^{*2}}$.

Para determinar si $d = 2$ está en la imagen de δ estudiamos la curva

$$C'_2:\ 2W^2 = 4 + 24Z^2 - 32Z^4.$$

Como buscamos puntos racionales, podemos sustituir $Z = Z'/2$, con lo que la ecuación se reduce a

$$W^2 = 2 + 3Z^2 - Z^4.$$

No es difícil encontrar el punto $(Z, W) = (1, 2)$, luego $2 \in \operatorname{Im}\delta$, al igual que $-2 \cdot 2 = -1$.

Consideremos ahora

$$C'_{17}:\ 17W^2 = 17^2 + 12 \cdot 17Z^2 - 32Z^4.$$

Supongamos que esta ecuación tiene solución en $\mathbb{Q}_{17}$. Entonces

$$2 = v_{17}(17^2) = \min\{2v_{17}(w) + 1, v_{17}(12 \cdot 17z^2 - 32z^4)\}.$$

Concluimos que $2v_{17}(w) + 1 > 2$, luego $v_{17}(w) \geq 0$. Por consiguiente,

$$\min\{2v_{17}(z) + 1, 4v_{17}(z)\} \geq 0,$$

lo que implica que también $v_{17}(z) \geq 0$. Más aún, la ecuación muestra que $v_{17}(z) \geq 1$, de donde $v_{17}(w) \geq 1$ y, por último, que $v_{17}(17^2) \geq 3$, contradicción.

Así pues, $17 \notin S^{\phi}(E/\mathbb{Q})$. Teniendo en cuenta que $S^{\phi}/E/\mathbb{Q})$ es un grupo, podemos concluir que $S^{\phi}(2, \mathbb{Q}) = \{\pm 1, \pm 2\} = \operatorname{Im}\delta$. En particular vemos que $Ш(E/\mathbb{Q})[\phi] = 0$.

Ahora intercambiamos los papeles de E y E'. Ahora la ecuación es

$$C'_d : dW^2 = d^2 - 24dZ^2 + 272Z^4.$$

En primer lugar tenemos que $\delta(0,0) = 272 = 17$. En este caso observamos que si $d < 0$ la ecuación no tiene soluciones en $\mathbb{R}$, lo que nos descarta ya cuatro posibilidades.

Si $d = 2$, tenemos la ecuación

$$C_2' : \ 2W^2 = 4 - 48Z^2 + 272Z^4.$$

Con el cambio $Z = Z'/2$ se reduce a

$$2W^2 = 4 - 12Z^2 + 17Z^4.$$

Se comprueba inmediatamente que si esta ecuación tiene solución en $\mathbb{Q}_2$, dicha solución ha de ser entera, pero entonces $v_2(z) \geq 1$, $v_2(w) \geq 1$, $v_2(4) \geq 3$, contradicción.

Resulta que $2 \notin S^{\hat{\phi}}(E'/\mathbb{Q})$, de donde podemos concluir que

$$S^{\hat{\phi}}(E'/\mathbb{Q}) = \{1, 17\}.$$

Según 8.11, el rango r de las curvas satisface la relación $2^{r+2} = 4 \cdot 2$, luego $r = 1$. ∎

En el caso de curvas definidas sobre $\mathbb{Z}$ podemos dar varias condiciones necesarias para que un elemento de $\mathbb{Q}(S, 2)$ pueda estar en la imagen de δ. Por ejemplo:

Teorema 8.13 *Sean $E/\mathbb{Z}$ y $E'/\mathbb{Z}$ las curvas elípticas dadas por las ecuaciones*

$$E : \ Y^2 = X^3 + aX^2 + bX, \qquad E' : Y^2 = X^3 - 2aX^2 + b'X, \qquad b' = a^2 - 4b,$$

y sea $[P] \in E'(\mathbb{Q})/\phi[E(\mathbb{Q})]$, donde ϕ es la isogenia considerada en 8.11. *Si $\delta([P]) = [d]$, donde $d \in \mathbb{Z}$ es libre de cuadrados, entonces $d \mid b'$.*

Demostración: Si $P = O$ entonces $d = 1$ y si $P = (0, 0)$ entonces d es la parte libre de cuadrados de b', luego también $d \mid b'$.

Para cualquier otro punto $P = (x, y)$ tenemos que $\delta([P]) = [x]$. Partimos de que P cumple la ecuación

$$y^2 = x^3 - 2ax^2 + b'x.$$

Los razonamientos tras la definición 6.17 muestran que si p es un primo entonces $v_p(x) < 0$ si y sólo si $v_p(y) < 0$, y en tal caso $v_p(x) = -2v_p(P)$, $v_p(y) = -3v_p(P)$. Equivalentemente, podemos expresar

$$x = \frac{m}{e^2}, \quad y = \frac{n}{e^3}, \qquad (m, e) = (n, e) = 1.$$

(Se entiende que si $n = 0$ entonces $e = 1$.) La ecuación equivale a

$$n^2 = m(m^2 - 2ame^2 + b'e^4).$$

Sea $u = (m, b')$, de modo que podemos descomponer $m = um_1$, $b' = ub_1$, con $(m_1, b_1) = 1$. Elegimos u con el mismo signo que m, de modo que $m_1 > 0$.

De la ecuación se sigue que $u^2 \mid n^2$, digamos $n = un_1$. Así, la ecuación pasa a ser

$$n_1^2 = m_1(um_1^2 - 2am_1e^2 + b_1e^4).$$

Como $(m_1, b_1) = (m_1, e) = 1$, los dos factores de la derecha son primos entre sí, luego ambos son cuadrados (notemos que no son negativos). Pongamos que $m_1 = k^2$. Entonces $\delta([P]) = [k^2u/e^2] = [u]$, y $u \mid b'$. El número d del enunciado es la parte libre de cuadrados de u, luego también $d \mid b'$. ∎

Observemos que si queremos aplicar este teorema a la curva E' en lugar de E', entonces E' pasa a ser la curva

$$E'' : \ Y^2 = X^3 + 4aX^2 + 16bX,$$

luego el teorema nos da que $d \mid 16b$. Sin embargo podemos concluir que en realidad $d \mid b$. En efecto, si $P = O$ es $d = 1$, si $P = (0, 0)$ entonces d es la parte libre de cuadrados de $16b$, que ciertamente divide a b y, en el caso general, tenemos que $P = (x, y)$ cumple la ecuación de E'', pero entonces $(x/4, y/8)$ cumple la ecuación de E, luego por el teorema $x/4$ es congruente módulo $\mathbb{Q}^{*2}$ con un entero divisor de b, luego lo mismo le sucede a x.

Otra condición necesaria sencilla para que un entero d pueda estar en el grupo de Selmer es la siguiente:

Teorema 8.14 *Sean $E/\mathbb{Z}$ y $E'/\mathbb{Z}$ las curvas elípticas dadas por las ecuaciones*

$$E : \ Y^2 = X^3 + aX^2 + bX, \qquad E' : Y^2 = X^3 - 2aX^2 + b'X, \qquad b' = a^2 - 4b.$$

Si $[d] \in S^\phi(E/\mathbb{Q})$ y $b' > 0$ entonces también $d > 0$.

Demostración: En efecto, la ecuación

$$dW^2 = d^2 - 2adZ^2 + b'Z^4$$

ha de tener una solución en $\mathbb{R}$, lo cual es imposible si $d < 0$. ∎

Combinando los dos últimos teoremas obtenemos una cota para el rango:

Teorema 8.15 *Sea $E/\mathbb{Z}$ una curva elíptica dada por una ecuación*

$$E : \ Y^2 = X^3 + aX^2 + bX.$$

Sea $b' = a^2 - 4b$ y sean ν_1 y ν_2 el número de divisores primos de b y b' respectivamente. Entonces el rango r de $E/\mathbb{Q}$ cumple la relación

$$r \le \nu_1 + \nu_2 - 1.$$

Demostración: Por el teorema 8.13 y la observación que le sigue, la imagen de $E'(\mathbb{Q})/\phi[E(\mathbb{Q})]$ en $\mathbb{Q}(S, 2)$ tiene a lo sumo tantos elementos como divisores libres de cuadrados tiene b', que son 2^{ν_2+1}. Similarmente, al razonar con E' obtenemos a lo sumo un grupo de orden 2^{ν_1+1}. Ahora bien, como $b' + 4b = a^2$, no puede ocurrir que b' y b sean ambos negativos, luego el teorema anterior nos reduce a la mitad uno de los dos órdenes, y en total

$$2^{r+2} \le 2^{\nu_1+\nu_2+1},$$

lo que nos da la relación indicada. ∎

Ejemplo El teorema anterior puede usarse para concluir directamente que una curva tiene rango 0 cuando $b = \pm 1$ y b' es potencia de primo. Si $b = 1$ tenemos que $b' = a^2 - 4 = (a+2)(a-2)$ no es potencia de primo salvo en los casos $(a, b) = (0, 1)$, $(\pm 1, 1)$, $(\pm 3, 1)$, $(\pm 6, 1)$.

Con $b = -1$ se ha de cumplir $a^2 + 4 = p^n$, lo cual sucede, por ejemplo, para

$$a = 0,\ \pm 1,\ \pm 2,\ \pm 3,\ \pm 5,\ \pm 7,\ \pm 11,\ \pm 13,\ \pm 17,\ \pm 27, \ldots$$

Esto no agota las posibilidades, pues también están las curvas con $b' = \pm 1$, como es el caso de $(a, b) = (\pm 3, 2)$. ∎

Veamos otra condición necesaria para que una clase $[d] \in \mathbb{Q}(S, 2)$ pueda estar en el grupo de Selmer. Recordemos que el *símbolo de Hilbert* en $\mathbb{Q}_p^*$ (admitiendo $p = \infty$) puede definirse como

$$(\alpha, \beta)_p = \begin{cases} 1 & \text{si } \alpha x^2 + \beta y^2 = z^2 \text{ tiene una solución no trivial en } \mathbb{Q}_p, \\ -1 & \text{en caso contrario.} \end{cases}$$

Por ejemplo, es inmediato que $(\alpha, -\alpha)_p = 1$ para todo $\alpha \in \mathbb{Q}_p^*$. No es trivial, pero se prueba que el símbolo de Hilbert $\mathbb{Q}_p^* \times \mathbb{Q}_p^* \longrightarrow \{\pm 1\}$ es multiplicativo y conmutativo. Obviamente entonces sólo depende del resto módulo $\mathbb{Q}_p^{*2}$ de sus argumentos.

Teorema 8.16 *Sean $E/\mathbb{Z}$ y $E'/\mathbb{Z}$ las curvas elípticas dadas por las ecuaciones*

$$E:\ Y^2 = X^3 + aX^2 + bX, \qquad E' : Y^2 = X^3 - 2aX^2 + b'X, \qquad b' = a^2 - 4b.$$

Si $[d] \in S^\phi(E/\mathbb{Q})$, entonces $(d, b)_p = 1$ para todo primo p.

Demostración: La ecuación

$$dW^2 = d^2 - 2adZ^2 + b'Z^4$$

ha de tener una solución en $\mathbb{Q}_p$. Hacemos $x = 1$ y llamamos $y = Z^2$, con lo que

$$W^2 = dx^2 - 2axy + (b'/d)y^2 = d(x - (a/d)y)^2 + (-4b/d)y^2.$$

Por consiguiente

$$1 = (d, -4b/d)_p = (d, -d)_p (d, b)_p = (d, b)_p.$$

∎

Para usar este teorema tendremos en cuenta las siguientes reglas de cálculo: Fijamos $p < \infty$ y sean ϵ, η unidades p-ádicas. Entonces

a) $(\alpha, -\alpha)_p = 1$, $\quad (\alpha, \alpha)_p = (\alpha, -1)_p$.

b) Si $p \neq 2$ entonces $(p, \epsilon)_p = (\epsilon/p)$ (símbolo de Legendre), $(\epsilon, \eta)_p = 1$.

c) $(2, \epsilon)_2 = 1$ si y sólo si $\epsilon \equiv \pm 1$ (mód 8),
$(\epsilon, \eta)_2 = 1$ si y sólo si $\epsilon \equiv 1$ (mód 4) o bien $\eta \equiv 1$ (mód 4).

Notemos por último que $(\alpha,\beta)_\infty = 1$ si y sólo si $\alpha > 0$ o $\beta > 0$.

Teorema 8.17 *Sea D un entero no nulo y $E/\mathbb{Q}$ la curva elíptica dada por la ecuación $Y^2 = X^3 + D^3$. Supongamos que todo primo $p \mid D$ cumple $p = 3$ o bien $p \equiv 5 \pmod{12}$. Entonces el rango de $E/\mathbb{Q}$ es $r = 0$.*

DEMOSTRACIÓN: No perdemos generalidad si suponemos que D es libre de cuadrados, pues si $D = d^2D'$, el cambio de variables $X = d^2X'$, $Y = d^3Y'$ transforma la ecuación en $Y^2 = X^3 + D'^3$.

Como el punto de orden 2 es $(-D, 0)$, aplicamos la traslación $X = X' - D$, con lo que la ecuación se convierte en

$$E:\ Y^2 = X^3 - 3DX^2 + 3D^2,$$

que se corresponde con

$$E':\ Y^2 = X^3 + 6DX^2 - 3D^2.$$

En principio vemos que la imagen de E' contiene a $\{1, -3\}$, mientras que la de E contiene a $\{1, 3\}$. Vamos a probar que no hay más posibilidades. Cualquier elemento de $\mathbb{Q}(S, 2)$ que esté en la imagen de una de las dos curvas ha de ser de la forma $[d]$, donde d es un entero libre de cuadrados divisor de $3D^2$.

Si d es divisible entre un primo $p \equiv 5 \pmod{12}$, entonces

$$(d, \pm 3D^2)_p = (d, \pm 3)_p = (p, 3)_p = (3/p) = (p/3) = (2/3) = -1,$$

luego $[d]$ no puede estar en las imágenes de E o E'.

La clase $[-1]$ tampoco puede estar en ninguna de las dos imágenes, pues

$$(-1, \pm 3D^2)_3 = (-1, 3)_3 = (-1/3) = -1.$$

Así pues, el rango r cumple $2^{r+2} = 2 \cdot 2$, luego es $r = 0$. ∎

Ahora vamos a acotar el rango que una familia de curvas:

Teorema 8.18 *Para cada primo impar p, consideremos la curva elíptica*

$$E_p:\ Y^2 = X^3 + pX.$$

Entonces

$$\operatorname{rang} E_p(\mathbb{Q}) + \dim_2 Ш(E_p/\mathbb{Q})[2] = \begin{cases} 0 & \text{si } p \equiv 7, 11 \pmod{16}, \\ 1 & \text{si } p \equiv 3, 5, 13, 15 \pmod{16}, \\ 2 & \text{si } p \equiv 1, 9 \pmod{16}. \end{cases}$$

DEMOSTRACIÓN: Se entiende que el segundo sumando es la dimensión de $Ш(E_p/\mathbb{Q})[2]$ como espacio vectorial sobre $\mathbb{Z}/2\mathbb{Z}$. Por simplicidad llamaremos E a E_p y E' a la curva

$$Y^2 = X^3 - 4pX.$$

Tenemos que $\mathbb{Q}(S,2) = \{\pm 1, \pm 2, \pm p, \pm 2p\}$. Además $1, p \in S^{\hat{\phi}}(E'/\mathbb{Q})$ y, como $p > 0$, el teorema 8.14 implica que las clases con representante negativo no están en el grupo de Selmer. Además, el espacio homogéneo

$$C_2' : \ 2W^2 = 4 + pZ^4$$

cumple $C_2'(\mathbb{Q}_2) = \varnothing$, pues todo punto racional ha de ser entero, pero entonces $v_2(Z) \geq 1$, luego $v_2(W) \geq 1$, luego $v_2(4) \geq 3$, contradicción.

Con esto podemos concluir que $S^{\hat{\phi}}(E'/\mathbb{Q}) = \{1, p\}$.

Por otra parte tenemos que $1, -p \in S^{\phi}(E/\mathbb{Q})$. Vamos a estudiar qué sucede con los elementos restantes de $\mathbb{Q}(S,2)$.

***d* = −1** Se cumplirá que $-1 \in S^{\phi}(E/\mathbb{Q})$ si y sólo si la ecuación

$$W^2 + 1 = 4pZ^4$$

tiene solución en $\mathbb{Q}_2$ y en $\mathbb{Q}_p$. Es claro que una solución en $\mathbb{Q}_p^2$ ha de estar de hecho en $\mathbb{Z}_p^2$, y entonces $W^2 \equiv -1 \pmod{p}$. Esto último sucede si y sólo si $p \equiv 1 \pmod 4$. Recíprocamente, si la congruencia tiene solución, el criterio enunciado en la página 21 (que vamos a usar repetidamente sin más referencia) nos da que la ecuación tiene solución en $\mathbb{Q}_p^2$.

Concluimos que $C_{-1}(\mathbb{Q}_2) \neq \varnothing$ si y sólo si $p \equiv 1 \pmod 4$. Bajo esta hipótesis, estudiamos la existencia de soluciones en $\mathbb{Q}_2$. Si $p \equiv 1 \pmod 8$ hacemos el cambio $(Z, W) = (Z'/4, W'/8)$, con lo que la ecuación se convierte en

$$W^2 + 64 = pZ^4.$$

Tenemos que $(Z, W) = (1, 1)$ es una solución módulo 8, que da lugar a una solución en $\mathbb{Q}_2^2$.

Si $p \equiv 5 \pmod 8$ hacemos $(Z, W) = (Z/2, W/2)$, con lo que la ecuación pasa a ser

$$W^2 + 4 \equiv pZ^4,$$

y la solución $(Z, W) = (1, 1)$ módulo 8 da lugar a una solución en $\mathbb{Q}_2^2$. En definitiva, si $p \equiv 1 \pmod 4$, no sólo $C_{-1}(\mathbb{Q}_p) \neq \varnothing$, sino que también $C_{-1}(\mathbb{Q}_2) \neq \varnothing$. Así pues:

$$-1 \in S^{\phi}(E/\mathbb{Q}) \text{ si y sólo si } p \equiv 1 \pmod 4.$$

***d* = −2** El espacio homogéneo C_{-2} tiene por ecuación

$$W^2 + 2 = 2pZ^4.$$

Todo punto racional en $\mathbb{Q}_p$ es entero, y da lugar a la congruencia

$$W^2 \equiv -2 \pmod{p}.$$

Recíprocamente, una solución de esta congruencia proporciona un punto racional. Así pues,

$$C_{-2}(\mathbb{Q}_p) \neq \varnothing \text{ si y sólo si } p \equiv 1,\ 3 \text{ (mód 8)}.$$

Por otra parte, una solución en $\mathbb{Q}_2$ ha de ser entera y además ha de cumplir $v_2(W) \geq 1$. Haciendo $(Z, W) = (Z', 2W')$, basta ver si la ecuación

$$2W^2 + 1 = pZ^4$$

tiene soluciones en $\mathbb{Z}_2$. Si $p \equiv 11$ (mód 16), la ecuación no tiene soluciones módulo 16, luego tampoco las tiene en $\mathbb{Z}_2$. En los otros casos, para pasar a una solución en $\mathbb{Z}_2$ necesitamos soluciones módulo 32 con Z impar. La tabla muestra que siempre existen (suponiendo $p \equiv 1, 3$ (mód 8)):

p (mód 32)	1	3	9	17	19	25
(Z, W)	$(1, 0)$	$(3, 11)$	$(1, 2)$	$(3, 0)$	$(1, 3)$	$(3, 2)$

La conclusión es que $-2 \in S^\phi(E/\mathbb{Q})$ si y sólo si $p \equiv 1, 3, 9$ (mód 16).

$d = 2$ Ahora la ecuación es

$$W^2 = 2 - 2pZ^4.$$

La discusión es muy similar a la del caso anterior. La existencia de solución en $\mathbb{Q}_p$ equivale a que la congruencia $W^2 \equiv 2$ (mód 0) tenga solución, lo cual equivale a que $p \equiv 1, 7$ (mód 8).

Si $p \equiv 1$ (mód 8) sabemos que -1, $-2 \in S^\phi(E/\mathbb{Q})$, luego también se cumple que $2 \in S^\phi(E/\mathbb{Q})$. Falta estudiar el caso $p \equiv 7$ (mód 8). Podemos hacer el cambio $(Z, W) = (Z', 2W')$ y pasar a $2W^2 = 1 - pZ^4$.

Si $p \equiv 7$ (mód 16) no hay soluciones módulo 16 y si $p \equiv 15$ (mód 16) entonces tenemos las soluciones siguientes módulo 32:

$$(Z, W) = \begin{cases} (1, 3) & \text{si } p \equiv 15 \text{ (mód 32)}, \\ (1, 1) & \text{si } p \equiv 31 \text{ (mód 32)}. \end{cases}$$

La conclusión es que

$$2 \in S^\phi(E/\mathbb{Q}) \text{ si y sólo si } p \equiv 1, 9, 15 \text{ (mód 16)}.$$

Con esto tenemos información suficiente para concluir que

$$|S^\phi(E/\mathbb{Q})| = \begin{cases} 2 & \text{si } p \equiv 7, 11 \text{ (mód 16)}, \\ 4 & \text{si } p \equiv 3, 5, 13, 15 \text{ (mód 16)}, \\ 8 & \text{si } p \equiv 1, 9 \text{ (mód 16)}. \end{cases}$$

Las sucesiones exactas (8.4) y (8.3) nos dan que

$$\dim E(\mathbb{Q})/2E(\mathbb{Q}) + \dim_2 E'(\mathbb{Q})[\hat{\phi}]/\phi[E(K)[2]]$$

$$= \dim_2 E(\mathbb{Q})/\hat{\phi}[E'(\mathbb{Q})] + \dim_2 E'(\mathbb{Q})/\phi[E(\mathbb{Q})]$$

$$= \dim_2 S^{\hat{\phi}}(E'/\mathbb{Q}) - \dim_2 Ш(E'/\mathbb{Q})[\hat{\phi}] + \dim_2 S^\phi(E/\mathbb{Q}) - \dim_2 Ш(E/\mathbb{Q})[\phi].$$

El teorema 7.18 muestra que tanto $E(\mathbb{Q})$ como $E'(\mathbb{Q})$ tienen exactamente dos puntos de torsión, luego

$$\dim_2 E'(\mathbb{Q})[\hat{\phi}]/\phi[E(K)[2]] = 1, \quad \dim_2 E(\mathbb{Q})/2E(\mathbb{Q}) = r+1,$$

donde r es el rango de $E(\mathbb{Q})$. Por otra parte, hemos visto que todos los elementos de $S^{\hat{\phi}}(E'/\mathbb{Q})$ están en la imagen de δ, luego $Ш(E'/\mathbb{Q})[\hat{\phi}] = 0$. Ahora nos fijamos en la sucesión exacta

$$0 \longrightarrow Ш(E/\mathbb{Q})[\phi] \longrightarrow Ш(E/\mathbb{Q})[2] \xrightarrow{\phi} Ш(E'/\mathbb{Q})[\hat{\phi}] = 0,$$

de la que obtenemos que

$$\dim_2 Ш(E/\mathbb{Q})[\phi] = \dim_2 Ш(E/\mathbb{Q})[2].$$

En total llegamos a que

$$r + 2 = 1 + \dim_2 S^{\phi}(E/\mathbb{Q}) - \dim_2 Ш(E/\mathbb{Q})[2].$$

De aquí se sigue inmediatamente la fórmula del enunciado. ∎

Vamos a usar el teorema anterior para dar ejemplos de curvas cuyo grupo de Tate-Shafarevich no es trivial. Vamos a usar una consecuencia de la ley de reciprocidad bicuadrática: Todo primo $p \equiv 1 \pmod 8$ puede expresarse en la forma $p = A^2 + B^2$. Se cumple que 2 es un resto bicuadrático módulo p si y sólo si $8 \mid AB$.

Teorema 8.19 *Sea $p \equiv 1 \pmod 8$ un primo tal que 2 no sea un resto bicuadrático módulo p. Entonces la curva $E/\mathbb{Q}$ dada por $Y^2 = X^3 + pX$ tiene rango 0 y $Ш(E/\mathbb{Q})[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.*

Demostración: En la prueba del teorema anterior hemos visto que

$$S^{\phi}(E/\mathbb{Q}) = \{\pm 1, \pm 2, \pm p, \pm 2p\}.$$

Basta probar que la imagen de δ está formada únicamente por los dos puntos obvios: 1 y $-p$. A su vez, para ello basta probar que los espacios homogéneos C_{-1}, C_2 y C_{-2} no tienen puntos racionales en $\mathbb{Q}$.

Consideramos primeramente los espacios $C_{\pm 2}$, dados por

$$\pm W^2 = 2 - 2pZ^4.$$

Haciendo $Z = r/t$, con r, $t \in \mathbb{Z}$ primos entre sí, la ecuación se convierte en

$$\pm(t^2W)^2 = 2t^4 - 2pr^4,$$

de donde se sigue que $t^2W \in \mathbb{Z}$ y además es par. Digamos, pues, que $W = 2s/t^2$. Ahora la ecuación es

$$\pm 2s^2 = t^4 - pr^4.$$

Observemos que p no puede dividir a s ni a t, pues si dividiera a uno dividiría a los dos, y también a r. Por lo tanto, tenemos la congruencia

$$\pm 2s^2 \equiv t^4 \text{ (mód } p),$$

donde los dos miembros son no nulos.

Si q es un primo impar que divida a s, entonces $(p/q) = 1$. (Notemos que $q \nmid r$.) Por lo tanto $(q/p) = 1$ y, como también $(2/p) = 1$, concluimos que $(s/p) = 1$. Así pues, $s \equiv u^2$ (mód p), y llegamos a que

$$\pm 2u^4 \equiv t^4 \text{ (mód } p).$$

Ahora bien, -1 es un resto bicuadrático módulo p. Basta considerar $v^{(p-1)/8}$, donde v es una raíz primitiva de la unidad módulo p. Concluimos que 2 es un resto bicuadrático módulo p, en contra de lo supuesto.

Consideremos ahora el espacio C_{-1}, cuya ecuación es $-W^2 = 1 - 4pZ^4$. Expresemos $Z = r/2t$, donde r, $t \in \mathbb{Z}$ y $(r, t) = 1$ (pero admitimos la posibilidad de que r sea par). Entonces $-(2t^2W)^2 = 4t^4 - pr^4$, de donde $s = 2t^2W \in \mathbb{Z}$ y la ecuación se convierte en

$$s^2 + 4t^4 = pr^4.$$

Ahora es fácil ver que r ha de ser impar, pues si $r = 2r'$, entonces $s = 2s'$, la ecuación se convierte en $s'^2 + t^4 = 16pr'^4$, con lo que s' es impar (ya que t ha de ser impar por ser primo con r), y esto es imposible módulo 8.

Por otra parte, el hecho de que $(r, t) = 1$ implica que $p \nmid s$ y $p \nmid t$, de donde a su vez $(s, t) = 1$. Por último, la ecuación módulo 8 nos da que $t = 2u$, con lo que

$$s^2 + 64u^4 = pr^4, \qquad (s, u) = 1.$$

Vamos a trabajar en el anillo $\mathbb{Z}[i]$. Tenemos que

$$pr^4 = s^2 + 64u^4 = (s + 8u^2 i)(s - 8u^2 i).$$

Los factores de la derecha no son divisibles entre enteros, luego r no puede ser divisible entre primos $q \equiv -1$ (mód 4) (pues también son primos en $\mathbb{Z}[i]$). Como además es impar, su factorización ha de ser de la forma

$$r = \prod_j \pi_j \bar{\pi}_j,$$

para ciertos primos de Gauss imaginarios π_j.

Por otra parte, $p = A^2 + B^2 = (A + Bi)(A - Bi)$, donde A, $B \in \mathbb{Z}$ y podemos suponer que A es impar. Si probamos que $8 \mid B$ tendremos que 2 es un resto bicuadrático módulo p, en contra de lo supuesto, y el teorema quedará probado. Para ello observamos que

$$pr^4 = (A + Bi)(A - Bi)\prod_j \pi_j^4 \bar{\pi}_j^4 = (s + 8u^2 i)(s - 8u^2 i).$$

Como los factores de la derecha no son divisibles entre enteros, ajustando la notación podemos suponer que

$$(A+Bi)\prod_j \pi_j^4 = s + 8u^2 i.$$

Podemos suponer que $\pi_j = a_j + 2b_j i$, con a_j impar. (Ello implica multiplicar π_j por una unidad adecuada, pero como en la expresión que tenemos aparece elevado a la cuarta, la unidad desaparece.) Entonces $\pi_j^4 \equiv a_j^4 \equiv 1$ (mód 8), luego $A + Bi \equiv s$ (mód 8), lo que implica que $8 \mid B$. ∎

Ejemplo El teorema anterior se aplica, por ejemplo, a los primos

$$17 = 1^2 + 4^2, \quad 41 = 5^2 + 4^2, \quad 97 = 9^2 + 4^2, \quad 193 = 7^2 + 12^2.$$

Así pues, $Y^2 = X^3 + 17X$ es un ejemplo de curva elíptica sobre $\mathbb{Q}$ con grupo de Tate-Shafarevich no trivial. ∎

Ejercicio: Para cada primo impar p, sea $E_p/\mathbb{Q}$ la curva $Y^2 = X^3 - pX$. Entonces

$$\mathrm{rang} E_p(\mathbb{Q}) + \dim_2 Ш(E_p/\mathbb{Q})[2] = \begin{cases} 0 & \text{si } p \equiv 3, 11, 13 \text{ (mód 16)}, \\ 1 & \text{si } p \equiv 5, 7, 9, 15 \text{ (mód 16)}, \\ 2 & \text{si } p \equiv 1 \text{ (mód 16)}. \end{cases}$$

8.4 Curvas sin puntos de orden 2

En esta sección daremos una cota para el rango de una curva elíptica $E/\mathbb{Q}$ sin puntos racionales de orden 2. El punto de partida será el teorema siguiente:

Teorema 8.20 *Sea $E/\mathbb{Q}$ una curva elíptica dada por una ecuación de Weierstrass de tipo b: $Y^2 = f(X)$. Supongamos que $E(\mathbb{Q})$ no contiene puntos de orden 2, es decir, que el polinomio $f(X)$ no tiene raíces en $\mathbb{Q}$. Sea λ una raíz de $f(X)$ y $L = \mathbb{Q}(\lambda)$. Entonces, la aplicación $\delta : E(\mathbb{Q}) \longrightarrow L^*/L^{*2}$ dada por*

$$\delta(P) = \begin{cases} [x(P) - \lambda] & \text{si } P \neq O, \\ [1] & \text{si } P = O, \end{cases}$$

es un homomorfismo de grupos, que induce un monomorfismo

$$\delta : E(\mathbb{Q})/2E(\mathbb{Q}) \longrightarrow L^*/L^{*2}.$$

Demostración: Como $E(L)$ tiene un punto de orden 2 (a saber, el punto $T = (\lambda, 0)$), podemos aplicarle el teorema 8.11, que nos da un homomorfismo $E(L) \longrightarrow L^*/L^{*2}$ cuya restricción a $E(\mathbb{Q})$ es el homomorfismo del enunciado. (Notemos que en 8.11 es $\delta(P) = [x(P)]$ porque hemos trasladado la curva para tomar $\lambda = 0$.)

Hemos de probar que el núcleo de δ es $2E(Q)$. Ciertamente, si $P \in 2E(\mathbb{Q})$ se cumple $\delta(P) = 1$. Recíprocamente, si se cumple $\delta(P) = 1$, tenemos que

$x(P) - \lambda$ es un cuadrado en L. Si llamamos L' a la adjunción a $\mathbb{Q}$ de las tres raíces de $f(X)$, aplicando automorfismos de $L'/\mathbb{Q}$ obtenemos que lo mismo vale si cambiamos λ por cualquiera de las otras dos raíces, y a su vez esto implica que P está en el núcleo del homomorfismo descrito en el teorema 8.2 para la curva E/L', es decir, que $P \in 2E(L')$. Digamos que $P = 2Q$, con $Q \in E(L')$. Falta probar que podemos tomar Q en $E(\mathbb{Q})$.

Llamemos $G = G(L'/\mathbb{Q})$. Para cada $\sigma \in G$ se cumple $2Q = P = P^\sigma = 2Q^\sigma$, luego $Q^\sigma = Q + T_\sigma$, con $T_\sigma \in E[2]$. Se comprueba inmediatamente que $\{T_\sigma\}_\sigma$ es un cociclo que determina una clase en $H^1(G, E[2])$. Si probamos que este grupo es trivial, entonces existe un $T \in E'[2]$ tal que $T_\sigma = T - T^\sigma$, para todo $\sigma \in G$, luego $(Q+T)^\sigma = Q + T$, luego $Q' = Q + T \in E(\mathbb{Q})$ y $P = 2Q'$.

Si el grupo $G = \langle\sigma\rangle$ es cíclico de orden 3, entonces el grupo de cohomología es trivial, pues puede calcularse como

$$H^1(G, E[2]) = \{P \in E[2] \mid P + P^\sigma + P^{\sigma^2} = O\} \,/\, \{P \in E[2] \mid P^\sigma - P = O\}.$$

Si G es el grupo completo de las permutaciones de las tres raíces de $f(X)$, también puede verse como el grupo de las permutaciones de los tres puntos de orden 2, con lo que el subgrupo cíclico H de orden 3 cumple también que $H^1(H, E[2]) = 0$, y en general tenemos la sucesión exacta

$$0 \longrightarrow H^1(G/H, 0) \xrightarrow{\text{Inf}} H^1(G, E[2]) \xrightarrow{\text{Res}} H^1(H, E[2]) = 0,$$

donde el 0 del grupo de la izquierda es el subgrupo de los elementos de $E[2]$ fijados por H, y claramente concluimos que $H^1(G, E[2]) = 0$. ∎

Nuestro problema es, una vez más, acotar la imagen de δ. Consideremos ahora una ecuación de Weierstrass general con coeficientes enteros:

$$E : Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6, \qquad a_i \in \mathbb{Z}.$$

De ésta se puede pasar a una ecuación en forma b:

$$Y^2 = X^3 + \frac{b_2}{4}X^2 + \frac{b_4}{2}X + \frac{b_6}{4}, \qquad b_i \in \mathbb{Z},$$

y, a su vez, un cambio de la forma $X = 2^{-2i}X'$, $Y = 2^{-3i}Y'$ la transforma en una ecuación

$$E' : Y^2 = X^3 + 2^{2i-2}b_2X^2 + 2^{4i-1}b_4X + 2^{6i-2}b_6,$$

donde $i \in \mathbb{Z}$ es el mínimo para el cual los coeficientes siguen siendo enteros. Diremos que esta ecuación es 2-minimal de tipo b (lo que no significa que sea 2-minimal en el sentido de que la curva $E'/\mathbb{Q}_2$ tenga discriminante mínimo, pues puede haber ecuaciones de tipo a con discriminante menor). Llamemos $f(X)$ al miembro derecho de esta última ecuación. Según las observaciones tras la definición 2.5, tenemos que el discriminante D_f de este polinomio es $D_f = \Delta/16$, donde Δ es el discriminante de la ecuación 2-minimal.

Supongamos ahora que $E(\mathbb{Q})[2] = 0$, que es equivalente a $E'(\mathbb{Q})[2] = 0$ y a su vez se traduce en que $f(X)$ no tiene raíces en $\mathbb{Q}$, luego es irreducible en $\mathbb{Q}[X]$. Llamemos λ, λ' y λ'' a las raíces complejas de $f(X)$. Sea $\mathcal{O}$ el anillo de enteros algebraicos del cuerpo cúbico $L = \mathbb{Q}(\lambda)$. Tenemos que $\mathbb{Z}[\lambda] \subset \mathcal{O}$, pero no tiene por qué darse la igualdad. Consideramos el índice

$$I = |\mathcal{O} : \mathbb{Z}[\lambda]| = \sqrt{D_f/D_L},$$

donde D_L es el discriminante del cuerpo L. Por último necesitamos una definición adicional:

Definición 8.21 Una curva elíptica $E/\mathbb{Z}$ es *cuasisupersingular* en 2 si cumple las cuatro condiciones siguientes:

a) $E(\mathbb{Q})[2] = 0$, con lo que, con la notación precedente, $L = \mathbb{Q}(\lambda)$ es un cuerpo cúbico.

b) 2 se ramifica en L, es decir, existe un (único) primo no arquimediano $\mathfrak{p}$ en L tal que $v_\mathfrak{p}(2) \geq 2$.

c) Si $\mathfrak{p}$ es el primo considerado en b), entonces $v_\mathfrak{p}(\lambda) \geq 2$.

d) Existe un natural impar N tal que para todo $(x, y) \in NE'(\mathbb{Q})$, donde $E'/\mathbb{Z}$ es la ecuación 2-minimal de tipo b de $E/\mathbb{Q}$, se cumple $v_\mathfrak{p}(x) \leq 0$.

La unicidad en b) se sigue de que cada primo racional se descompone a lo sumo en tres primos de L. Notemos que estas propiedades no dependen de la elección de λ. Vamos a probar que las curvas con reducción supersingular en 2 cumplen estas propiedades, si bien encontraremos ejemplos de curvas cuasisupersingulares en 2 con reducción ordinaria.

Teorema 8.22 *Si $E/\mathbb{Z}$ es una curva elíptica dada por una ecuación de Weierstrass con reducción supersingular en 2, entonces es cuasisupersingular en 2. Más aún, 2 se ramifica totalmente en L y $\pi = \lambda/2$ cumple $v_\mathfrak{p}(\pi) = 1$.*

Demostración: Que la curva $\tilde{E}(\mathbb{Z}/2\mathbb{Z})$ sea supersingular equivale a que su discriminante sea 1 y su invariante 0. Esto implica que a_2 es par, con lo que $b_2 = 4b_2'$ y $b_4 = 2b_4'$, con b_2', $b_4' \in \mathbb{Z}$. Por otra parte, b_6 ha de ser impar (o el discriminante sería nulo módulo 2). Esto implica que en $f(X)$ ha de ser $i = 1$, luego

$$f(X) = X^3 + 4b_2'X^2 + 16b_4'X + 16b_6.$$

La igualdad $f(\lambda) = 0$ equivale a

$$\pi^3 + 2b_2'\pi^2 + 4b_4'\pi + 2b_6 = 0.$$

Vemos, pues, que π satisface una ecuación de Eisenstein. Esto prueba que π es primo en $\mathbb{Q}_2(\pi)$ y que la extensión $\mathbb{Q}_2(\pi)/\mathbb{Q}_2$ es totalmente ramificada. Además $f(X)$ resulta ser irreducible, luego se cumplen las tres primeras propiedades. Para la última podemos tomar $N = 15$. En efecto, es fácil ver que

de las 32 ecuaciones de Weierstrass con coeficientes en $\mathbb{Z}/2\mathbb{Z}$ sólo 8 corresponden a curvas supersingulares, y el número de puntos varía entre 1, 3 y 5. Así pues, $15E(\mathbb{Q}) \subset E_1(\mathbb{Q})$, donde $E_1(\mathbb{Q})$ es el núcleo de la reducción módulo 2. Según las observaciones tras la definición 6.17, tenemos que $v_2(P) \geq 1$ para todo $P \in 15E(\mathbb{Q})$, luego $v_2(x) = -2v_2(P) \leq -2$. El isomorfismo entre E y E' cumple $X' = 4X$, luego si $(x, y) \in 15E'(\mathbb{Q})$, entonces $v_2(x) \leq 0$. ∎

El resultado que perseguimos es el siguiente:

Teorema 8.23 *Sea $E/\mathbb{Z}$ una curva elíptica tal que $E(\mathbb{Q})[2] = 0$ y sea $f(X)$ el miembro derecho de su ecuación 2-minimal de tipo b. Sea λ una raíz compleja de $f(X)$ y $L = \mathbb{Q}(\lambda)$. Sea D_L el discriminante del cuerpo L, sea $\mathcal{O}$ su anillo de enteros algebraicos, sea I el índice de $\mathbb{Z}[\lambda]$ en $\mathcal{O}$ y sea H el grupo de clases. Entonces, el rango r de $E(\mathbb{Q})$ satisface la desigualdad*

$$r \leq n_L + 2n_p + n_q + n_h,$$

donde

a) $n_L = \begin{cases} 1 & \text{si } D_L < 0, \\ 2 & \text{si } D_L > 0. \end{cases}$

b) n_p *es el número de primos* $p \mid I$ *con tres divisores distintos en* L.

c) n_q *es el número de primos* $q \mid I$ *con exactamente dos divisores distintos en* L.

d) $n_h = \dim_2 H[2]$.

La cota anterior puede reducirse en una unidad si además E es cuasisupersingular en 2 y, llamando $\mathfrak{p}$ al primo de L que cumple $v_{\mathfrak{p}}(2) \geq 2$, se cumple que $\mathcal{O}$ contiene una unidad $\epsilon \equiv 1 \pmod{\mathfrak{p}^2}$.

Demostración: Sea $E'/\mathbb{Z}$ la curva dada por la ecuación 2-minimal de tipo b de E. Podemos aplicarle el teorema 8.20 y considerar el monomorfismo

$$\delta : E'(\mathbb{Q})/2E'(\mathbb{Q}) \longrightarrow L^*/L^{*2}.$$

El mismo razonamiento que en la prueba de 8.13 nos da que todo punto no nulo de $E'/\mathbb{Q}$ es de la forma $(m/e^2, n/e^3)$, con $m, n, e \in \mathbb{Z}$, $e \geq 1$, $(m, e) = (n, e) = 1$. Por lo tanto, la imagen de δ es el $\mathbb{Z}/2\mathbb{Z}$-espacio vectorial

$$V = \{[m - \lambda e^2] \mid (m/e^2, n/e^3) \in E'(\mathbb{Q})\}.$$

(En principio faltaría añadir $\delta(O) = [1]$, pero esta clase aparece ciertamente como imagen de otros puntos no nulos.) Según hemos visto al principio del capítulo, el rango r de $E/\mathbb{Q}$ (o de $E'/\mathbb{Q}$) cumple $|V| = 2^r$. De la relación

$$n^2 = (m - \lambda e^2)(m - \lambda' e^2)(m - \lambda'' e^2)$$

deducimos que todo ideal $\mathfrak{p}$ de L que tenga exponente impar en $m - \lambda e^2$ debe dividir a $\gamma = (m - \lambda' e^2)(m - \lambda'' e^2)$, luego también a $(m - \lambda e^2, \gamma)$.

Pongamos que $f(X) = X^3 + aX^2 + bX + c$. Entonces, el desarrollo de Taylor de $f(X)$ en λ es

$$f(X) = (X-\lambda)f'(\lambda) + (3\lambda + 2a)(X-\lambda)^2 + (x-\lambda)^3.$$

Evaluamos en $X = m/e^2$ y multiplicamos por e^6, con lo que obtenemos

$$(m-\lambda e^2)\gamma = n^2 = (m-\lambda e^2)e^4 f'(\lambda) + (3\lambda+2a)e^2(m-\lambda e^2)^2 + (m-\lambda e^2)^3,$$

de donde

$$e^4 f'(\lambda) = \gamma - (m-\lambda e^2)\big((3\lambda+2a)e^2 + (m-\lambda e^2)\big).$$

Observemos que si $\mathfrak{p} \mid (m-\lambda e^2, \gamma)$, como $(m,e) = 1$, tenemos que $\mathfrak{p} \nmid e$, y la igualdad anterior nos permite concluir que

$$v_{\mathfrak{p}}(f'(\lambda)) \geq \min\{v_{\mathfrak{p}}(m-\lambda e^2), v_{\mathfrak{p}}(\gamma)\}.$$

Esto prueba que $(m-\lambda e^2, \gamma) \mid f'(\lambda) = (\lambda-\lambda')(\lambda-\lambda'')$, y en particular $(m-\lambda e^2, \gamma) \mid \mathrm{N}(f'(\lambda)) = -D_f$. (La norma es la de la extensión $L/\mathbb{Q}$.)

En definitiva hemos probado que

$$m - \lambda e^2 = \mathfrak{a}\mathfrak{b}^2,$$

donde $\mathfrak{a}$ es un ideal libre de cuadrados $\mathfrak{a} \mid D_f$. Además

$$\mathrm{N}(m-\lambda e^2) = n^2 = \mathrm{N}(\mathfrak{a})\,\mathrm{N}(\mathfrak{b})^2,$$

luego $\mathrm{N}(\mathfrak{a})$ es un cuadrado.

Un primo racional p puede factorizar en L de cinco formas distintas (donde $\mathfrak{p}_i$ representa a un primo de norma p^i):

$$p = \mathfrak{p}_3, \quad p = \mathfrak{p}_1^3, \quad p = \mathfrak{p}_1^2\mathfrak{q}_1, \quad p = \mathfrak{p}_1\mathfrak{q}_2, \quad p = \mathfrak{p}_1\mathfrak{q}_1\mathfrak{r}_1.$$

En los dos primeros casos ha de ser $(p,\mathfrak{a}) = 1$, o de lo contrario $v_p(\mathrm{N}(\mathfrak{a}))$ sería impar y la norma no sería un cuadrado.

En el tercer caso, o bien $(p,\mathfrak{a}) = 1$ o bien $\mathfrak{p}_1\mathfrak{q}_1 \mid \mathfrak{a}$ y $v_p(\mathrm{N}(\mathfrak{a})) = 2$.

En el cuarto caso o bien $(p,\mathfrak{a}) = 1$ o bien $\mathfrak{p}_1 \nmid \mathfrak{a}$, $\mathfrak{q}_2 \mid \mathfrak{a}$ y $v_p(\mathrm{N}(\mathfrak{a})) = 2$.

En el quinto caso, o bien $(p,\mathfrak{a}) = 1$ o bien $\mathfrak{a}$ es divisible exactamente entre dos de los tres factores primos, y entonces $v_p(\mathrm{N}(\mathfrak{a})) = 2$.

Vamos a ver que si $p \mid \mathrm{N}(\mathfrak{a})$, entonces $p \mid I$. Razonamos según la factorización de p:

En el tercer caso tenemos que $\mathfrak{p}_1\mathfrak{q}_1 \mid m - \lambda e^2$, luego

$$p \mid \mathfrak{p}_1^2\mathfrak{q}_1^2 \qquad \text{y} \qquad \mathfrak{p}_1^2\mathfrak{q}_1^2 \mid (m-\lambda e^2)^2,$$

luego

$$\delta = \frac{m^2 - 2\lambda m e^2 + \lambda^2 e^4}{p} \in \mathcal{O}, \qquad p\delta \in \mathbb{Z}[\lambda].$$

Además $\delta \notin \mathbb{Z}[\lambda]$, pues ello implicaría que $p \mid e$, pero $(e, n) = 1$. Esto implica que $p \mid I$, pues basta considerar el módulo $M = \langle 1, \lambda, \lambda^2, \delta \rangle$ y examinar la sucesión

$$pM \subset \mathbb{Z}[\lambda] \subset M \subset \mathcal{O}.$$

Tenemos que $|M : pM| = p^3$, $|\mathcal{O} : \mathbb{Z}[\lambda]| = I$, $|M : \mathbb{Z}[\lambda]| \neq 1$.

En los casos cuarto y quinto p no se ramifica en L, luego $p \nmid D_L$, aunque $p \mid D_f$, luego $p \mid I$.

Esto hace que $\mathfrak{a}$ sólo pueda variar en un conjunto de $2^{2n_p+n_q}$ ideales, donde n_p es el número de divisores de I con factorización de tipo 5 y n_q el número de divisores con factorizaciones de tipo 3 o 4. (Notemos que son los números definidos en el enunciado.)

Para cada ideal $\mathfrak{a}$ que pueda obtenerse a partir de un punto (no nulo) de $E'(\mathbb{Q})$, escogemos un $\mu_\mathfrak{a} = m - \lambda e^2$, de modo que $\mu_\mathfrak{a} = \mathfrak{a}\mathfrak{b}_\mathfrak{a}^2$. De este modo, para un punto (no nulo) arbitrario de $E'(\mathbb{Q})$ tenemos que

$$m - \lambda e^2 = \mu_\mathfrak{a}(\mathfrak{b}\mathfrak{b}_\mathfrak{a}^{-1})^2.$$

Vemos que $\mathfrak{b}\mathfrak{b}_\mathfrak{a}^{-1} \in H[2]$. Sean $\mathfrak{c}_1, \ldots, \mathfrak{c}_{n_h}$ ideales cuyas clases formen una base de $H[2]$ y sea $\mathfrak{c}_i^2 = (\xi_i)$. Entonces

$$(\mathfrak{b}\mathfrak{b}_\mathfrak{a}^{-1})^2 = \alpha^2 \xi_1^{j_1} \cdots \xi_{n_h}^{j_{n_h}},$$

para cierto $\alpha \in L^*$ y ciertos $j_i \in \{0, 1\}$. Así,

$$m - \lambda e^2 = \epsilon \mu_\mathfrak{a} \alpha^2 \xi_1^{j_1} \cdots \xi_{n_h}^{j_{n_h}}, \tag{8.6}$$

para cierta unidad ϵ de L.

Ahora observamos que podemos tomar λ real y, si $D_f > 0$, es decir, si f tiene tres raíces reales, podemos tomar como λ la menor de ellas. Esto implica que $m - \lambda e^2 > 0$, pues si λ es la única raíz real entonces $m - \lambda' e^2$ y $m - \lambda'' e^2$ son conjugados, luego su producto es positivo. En particular $\mu_\mathfrak{a} > 0$. Cambiando ξ_i por $-\xi_i$ si es necesario, podemos suponer que todos son positivos, (son reales porque están en L) luego concluimos que $\epsilon > 0$.

Como L es real, el subgrupo de torsión de su grupo de unidades tiene orden 2, luego el subgrupo de las unidades positivas es libre y su rango es n_L, luego hay 2^{n_L} unidades módulo L^*.

Con esto concluimos que $|V| \leq 2^n$, donde $n = n_L + 2n_p + n_q + n_h$.

Por último veamos que con las hipótesis adicionales podemos probar que la unidad ϵ del enunciado no puede aparecer en (8.6) con $\mathfrak{a} = 1$ y $j_i = 0$. Esto implica que $|V| < 2^n$, luego $|V| \leq 2^{n-1}$.

Sea N el número natural dado por 8.21. La restricción de δ a $NE'(\mathbb{Q})$ tiene núcleo $2NE'(\mathbb{Q})$, y $NE'(\mathbb{Q})/2NE'(\mathbb{Q}) \cong E'(\mathbb{Q})/2E'(\mathbb{Q})$. Esto hace que la imagen (finita) de $NE'(\mathbb{Q})$ sea la misma que la de $E'(\mathbb{Q})$, es decir,

$$V = \{[m - \lambda e^2] \mid (m/e^2, n/e^3) \in NE'(\mathbb{Q})\}.$$

Todos los razonamientos precedentes siguen siendo válidos si partimos de esta representación de V, pero ahora tenemos además que m y n son impares (por la propiedad de N).

Como $\mathfrak{p}^2 \mid 2$ y $\mathfrak{p}^2 \mid \lambda$, vemos que $m - \lambda e^2 \equiv 1 \pmod{\mathfrak{p}^2}$. En particular tenemos que $\mu_1 \equiv 1 \pmod{\mathfrak{p}^2}$. Basta ver que la congruencia

$$\epsilon\alpha^2 \equiv 1 \pmod{\mathfrak{p}^2}$$

es imposible para todo $\alpha \in L^*$. Ello se debe a que todo $\alpha \in L^*$ es congruente módulo $\mathfrak{p}^2$ con uno de los restos 0, 1, π, $1+\pi$, luego $\alpha^2 \equiv 0, 1 \pmod{\mathfrak{p}^2}$ y por lo tanto $\epsilon\alpha^2 \equiv 0, 1+\pi \pmod{\mathfrak{p}^2}$. ∎

Ejemplo *La curva elíptica dada por la ecuación*

$$Y^2 + Y = X^3 - X^2$$

tiene rango 0.

En efecto, el discriminante es $\Delta = -11$ y $j = -2^{12}/11$, luego la curva tiene reducción supersingular en 2. La ecuación 2-minimal de tipo b es

$$Y^2 = X^3 - 4X^2 + 16, \qquad D_f = -2^8 \cdot 11.$$

Si λ es una raíz del miembro derecho y $\pi = \lambda/2$, entonces π satisface la ecuación de Eisenstein

$$\pi^3 - 2\pi^2 + 2 = 0.$$

El discriminante del orden $\mathbb{Z}[\pi]$ es $-2^2 \cdot 11$, y una comprobación rutinaria muestra que el 2^2 no puede eliminarse, de modo que $\mathcal{O} = \mathbb{Z}[\pi]$ y $D_L = -2^2 \cdot 11$. La constante de Minkowski para el cuerpo L es $M_{1,1} = 0.28295$, por lo que todo ideal de L es equivalente a uno de norma menor que $M_{1,1}\sqrt{|D_L|} = 1.87$, luego el número de clases es $h = 1$. Además, ahora es claro que $I = 2^3$ y por el teorema 8.22 sabemos que 2 se ramifica completamente en L, luego $n_p = n_q = n_h = 0$. Además $n_L = 1$ y con esto obtenemos la cota $r \le 1$ para el rango.

La cota puede ser reducida una unidad tomando $\epsilon = 1 - \pi$, pues, llamando $g(X) = X^3 - 2X^2 + 2$, es claro que $\mathrm{N}(1-\pi) = g(1) = 1$, luego ϵ es una unidad y, como $\mathfrak{p} \mid 2$, se cumple $\epsilon \equiv 1 + \pi \pmod{\mathfrak{p}^2}$. Esto nos permite concluir que el rango es 0. ∎

Ejercicio: Demostrar que las únicas soluciones racionales (y a fortiori enteras) de la ecuación $Y^2 + Y = X^3 - X^2$ son $(0,0)$, $(1,0)$, $(1,-1)$, $(0,-1)$.

Ejemplo *La curva elíptica dada por la ecuación*

$$Y^2 + Y = X^3 - 7$$

tiene rango 0.

Tenemos $\Delta = -3^9$, $j = 0$, luego la curva tiene reducción supersingular en 2. La ecuación 2-minimal es $Y^2 = X^3 - 2^4 \cdot 3^3$, con $D_f = -2^8 \cdot 3^9$. Obviamente $L = \mathbb{Q}(\sqrt[3]{2})$, y entonces $\mathcal{O} = \mathbb{Z}[\sqrt[3]{2}]$, $D_L = -2^2 \cdot 3^3$, con lo que $I = 6^3$. Es fácil comprobar que $n_L = 1$, $n_p = n_q = n_h = 0$, lo que nos da la cota $r \leq 1$ para el rango, que puede ser rebajada a $r = 0$ mediante la unidad $\epsilon = 1 + \sqrt[3]{2} + \sqrt[3]{4}$. ∎

Acabamos de probar que la curva $Y^2 = X^3 - 432$ tiene rango 0, y en el primer ejemplo de este capítulo hemos visto que esto equivale a que la ecuación de Fermat $X^3 + Y^3 = Z^3$ no tiene soluciones no triviales.

Ejemplo *La curva elíptica dada por la ecuación*

$$Y^2 + Y = X^3 - X$$

tiene rango 1.

Ahora tenemos $\Delta = 37$, $j = 2^{12} \cdot 3^3/37$, luego de nuevo la reducción en 2 es supersingular. La ecuación minimal es $Y^2 = X^3 - 16X + 16$, con $D_f = 2^8 \cdot 37$. Si λ es una ecuación del miembro derecho y $\pi = \lambda/2$, la ecuación de π es

$$\pi^3 - 4\pi + 2 = 0.$$

El orden $\mathbb{Z}[\pi]$ tiene discriminante $2^2 \cdot 37$, y una comprobación rutinaria muestra que el 2^2 no se puede eliminar, con lo que $\mathcal{O} = \mathbb{Z}[\pi]$, $D_L = 2^2 \cdot 37$, $I = 8$.

Todo ideal de L es equivalente a otro de norma menor que $M_{3,0}\sqrt{D_L} = 0.22223 \cdot \sqrt{2^2 \cdot 37} = 2.7$, pero es obvio que $\mathrm{N}(\pi) = -2$, luego $2 = (\pi)$ y concluimos que el número de clases es $h = 1$. Por lo tanto $n_L = 2$, $n_p = n_q = n_h = 0$. Esto nos proporciona una cota $r \leq 2$ para el rango, que puede ser rebajada a $r \leq 1$ con la unidad $\epsilon = -3 + \pi^2$. (Una comprobación rutinaria muestra que $\mathrm{N}(\epsilon) = 1$.)

Por otra parte, es claro que $(0, 0) \in E(\mathbb{Q})$ y la fórmula de duplicación nos da que

$$8(0,0) = \left(\frac{21}{25}, \frac{-69}{125}\right).$$

Este punto no puede ser de torsión, luego $r = 1$. ∎

El ejemplo de Selmer En la introducción hemos anunciado que la curva

$$3U^3 + 4V^3 + 5W^5 = 0$$

tiene puntos racionales en todas las compleciones $\mathbb{Q}_p$ de $\mathbb{Q}$ pero no tiene puntos racionales en $\mathbb{Q}$. La primera parte la hemos probado en la página 21, y en la página 35 hemos visto que para probar la segunda basta ver que la curva elíptica

$$Y^2 = X^3 - 432 \cdot 60^2$$

no tiene puntos racionales distintos de O. El teorema 7.19 prueba que la curva no tiene puntos de torsión distintos de O, luego basta probar que su rango es 0. La ecuación es ya de tipo b, pero no es minimal. La ecuación minimal es

$$Y^2 = X^3 - 2^2 \cdot 3^5 \cdot 5^2,$$

luego $D_f = -2^4 \cdot 3^{13} \cdot 5^4$.

Si llamamos $\lambda = \sqrt[3]{2^2 \cdot 3^5 \cdot 5^2}$ y $\mu = \sqrt[3]{30}$, entonces $\lambda = 3\mu^2$ y $\mu = \lambda^2/270$, luego $L = \mathbb{Q}(\lambda) = \mathbb{Q}(\mu)$ es un cuerpo cúbico puro, cuyo anillo de enteros es[1] $\mathcal{O} = \mathbb{Z}[\mu]$, $D_L = -2^2 \cdot 3^5 \cdot 5^2$, $I = 2 \cdot 3^4 \cdot 5$.

Todo ideal de L es equivalente a uno de norma menor que $M_{11}\sqrt{|D_L|} = 44{,}1$. La tabla siguiente contiene todos los ideales (p, η) primos de L de norma menor que 44 junto con $\mathrm{N}(\eta)$. Para calcular las normas observamos que

$$\mathrm{N}(\mu + a) = a^3 + 30.$$

(Más en general, se cumple que

$$\mathrm{N}(a + b\mu + c\mu^2) = a^3 + 30b^3 - 90abc + 900c^3.)$$

(p, η)	$\mathrm{N}(\eta)$	(p, η)	$\mathrm{N}(\eta)$	(p, η)	$\mathrm{N}(\eta)$
$(2, \mu)$	$2 \cdot 3 \cdot 5$	$(3, \mu)$	$2 \cdot 3 \cdot 5$	$(5, \mu)$	$2 \cdot 3 \cdot 5$
$(11, \mu - 2)$	$2 \cdot 11$	$(17, \mu - 4)$	$-2 \cdot 17$	$(19, \mu + 2)$	$2 \cdot 19$
$(19, \mu + 3)$	$3 \cdot 19$	$(19, \mu - 5)$	$-5 \cdot 19$	$(23, \mu + 9)$	$3 \cdot 11 \cdot 23$
$(29, \mu - 1)$	29	$(31, \mu + 1)$	31	$(31, \mu + 5)$	$5 \cdot 31$
$(31, \mu - 6)$	$-2 \cdot 3 \cdot 31$	$(41, \mu + 6)$	$2 \cdot 3 \cdot 41$		

Por otra parte, $\mathrm{N}(\mu - 3) = 3$, luego $3 = (\mu - 3)$ es principal. Pongamos que $2 = \mathfrak{p}^3$ y $5 = \mathfrak{q}^5$. Con la ayuda de un ordenador no es difícil ver que $\mathrm{N}(\mu^2 + 3\mu + 10) = 10$, luego ha de ser $\mu^2 + 3\mu + 10 = \mathfrak{pq}$ y, en el grupo de clases, $[\mathfrak{q}] = [\mathfrak{p}]^{-1}$.

Si $\mathfrak{r} = (11, \mu - 2)$, vemos que $\mu - 2 = \mathfrak{pr}$, luego $[\mathfrak{r}] = [\mathfrak{p}]^{-1} = [\mathfrak{q}]$. De mismo modo se descartan los restantes primos de la tabla, con lo que el grupo de clases es $H = \{[1], [\mathfrak{p}], [\mathfrak{q}]\}$. Como $[\mathfrak{p}]^3 = [1]$, el número de clases es $h = 1$ o $h = 3$. Puede probarse que $h = 3$, pero a nosotros nos basta con saber que es impar, pues entonces $n_h = 0$. También es claro que $n_p = n_q = 0$, $n_L = 1$. Con esto obtenemos para el rango la cota $r \leq 1$.

Vamos a ver que la curva es cuasisupersingular en 2 (si bien no tiene reducción supersingular). La única propiedad de la definición que no es obvia es la última. Vamos a ver que se cumple con $N = 1$, es decir, hemos de probar que si un punto $(x, y) \in \mathbb{Q}^2$ cumple $y^2 = x^3 - 2^2 \cdot 3^5 \cdot 5^2$, entonces $v_2(x) \leq 0$.

Supongamos que, por el contrario, $x = 2^a x'$, con $v_2(x') = 0$, $a > 0$. Entonces $y = 2y'$, donde $v_2(y') = 0$, y la ecuación se transforma en

$$y'^2 = 2^{3a-2}x'^3 - 3^5 \cdot 5^2,$$

de donde

$$2^{3a-2}x' - 3 \equiv 1 \text{ (mód 8)}.$$

[1] Ver el teorema 2.27 de mi Teoría de Números.

Esta congruencia es obviamente imposible si $a > 1$, y para $a = 1$ se reduce a

$$2x' \equiv 4 \text{ (mód 8)},$$

lo que nos da la contradicción $v_2(x') \geq 1$.

Para reducir la cota basta considerar la unidad $\epsilon = 1 + 9\mu - 3\mu^2$. ∎

Capítulo IX

Puntos enteros

Hasta ahora hemos estudiado esencialmente el grupo de puntos racionales de una curva elíptica. Ahora nos ocuparemos de los puntos enteros. El resultado más importante de este capítulo será un teorema de Siegel en virtud del cual toda curva elíptica contiene a lo sumo un número finito de puntos con coordenadas enteras. Dicho teorema se basa en un profundo resultado sobre aproximación diofántica conocido como teorema de Roth. Antes de entrar en estas cuestiones, dedicaremos la primera sección a probar algunos hechos más sencillos.

9.1 Resultados elementales

Los resultados de esta sección son consecuencias del teorema siguiente:

Teorema 9.1 *Sea K un cuerpo numérico y $\mathcal{O}$ su anillo de enteros. Sea E/K una curva elíptica dada por una ecuación de Weierstrass con coeficientes en $\mathcal{O}$ y $P \in E(K)$ un punto tal que nP tiene coordenadas en $\mathcal{O}$ para cierto $n \in \mathbb{Z}$. Entonces P tiene también sus coordenadas en $\mathcal{O}$.*

Demostración: Si $P = (x, y)$ no tiene coordenadas en $\mathcal{O}$, entonces existe un primo $\mathfrak{p}$ no arquimediano tal que $v_\mathfrak{p}(x) < 0$ o bien $v_\mathfrak{p}(y) < 0$. Esto implica que $v_\mathfrak{p}(P) = m > 0$ (ver 6.17) o, equivalentemente, que $P \in E_m(K)$. Como este conjunto es un subgrupo, también $nP \in E_m(K)$, lo que implica que nP no tiene coordenadas enteras, contradicción. ∎

De aquí podemos deducir un criterio que a veces nos permite encontrar una base del grupo de puntos racionales de una curva elíptica sobre $\mathbb{Z}$ de rango 1. Consideremos en principio una curva elíptica $E/\mathbb{R}$ definida mediante una ecuación de Weierstrass de tipo b. Si su discriminante cumple $\Delta > 0$, entonces el miembro derecho de la ecuación es un polinomio con tres raíces reales, digamos $\lambda < \lambda' < \lambda''$. Esto hace que $E(\mathbb{R})$ tenga dos componentes conexas:

$$E^0(\mathbb{R}) = \{P \in E(\mathbb{R}) \mid x(P) \geq \lambda''\} \cup \{O\},$$

$$E^1(\mathbb{R}) = \{P \in E(\mathbb{R}) \mid \lambda \leq x(P) \leq \lambda'\}.$$

Se cumple que $E^0(\mathbb{R})$ es un subgrupo de índice 2 en $E(\mathbb{R})$. (Esto es un caso particular de un hecho general sobre grupos topológicos: las componentes conexas son las clases de congruencia del subgrupo formado por la componente conexa del elemento neutro.)

Si la ecuación de Weierstrass no es de tipo b todo sigue siendo cierto, pues el isomorfismo que la transforma en una ecuación de tipo b conserva el discriminante y la coordenada x.

Si la curva está definida sobre $\mathbb{Q}$, podemos considerar igualmente el subgrupo $E^0(\mathbb{Q}) = E^0(\mathbb{R}) \cap E(\mathbb{Q})$, así como $E^1(\mathbb{Q}) = E^1(\mathbb{R}) \cap E(\mathbb{Q})$. O bien $E^0(\mathbb{Q})$ tiene índice 2 en $E(\mathbb{Q})$, o bien $E^1(\mathbb{Q}) = \varnothing$.

Teorema 9.2 *Sea $E/\mathbb{Z}$ una curva elíptica dada por una ecuación de Weierstrass que cumpla las condiciones siguientes:*

a) El rango de $E(\mathbb{Q})$ es 1.

b) $E(\mathbb{Q})$ contiene un punto Q de orden infinito tal que $Q+T$ es entero para todo punto $T \in E_{\mathrm{tor}}(\mathbb{Q})$.

c) $\Delta > 0$.

d) $E^1(\mathbb{Q}) \neq \varnothing$.

Entonces $E(\mathbb{Q}) = E_{\mathrm{tor}}(\mathbb{Q}) \oplus \langle P \rangle$, para cierto punto $P \in E^1(\mathbb{Q})$ con coordenadas enteras.

Demostración: En principio sabemos que existe un punto $R \in E(\mathbb{Q})$ tal que $E(\mathbb{Q}) = E_{\mathrm{tor}}(\mathbb{Q}) \oplus \langle R \rangle$. Entonces existe un $n \in \mathbb{Z}$ y un $T \in E_{\mathrm{tor}}(\mathbb{Q})$ tal que $nR = T + Q$. Por hipótesis nR tiene coordenadas enteras y, por el teorema anterior, R también.

Es claro que $E(\mathbb{Q}) = E_{\mathrm{tor}}(\mathbb{Q}) \oplus \langle R+T \rangle$, para todo $T \in E_{\mathrm{tor}}(\mathbb{Q})$, luego en realidad hemos demostrado que $R+T$ tiene coordenadas enteras cualquiera que sea T. Basta probar que podemos elegir T de modo que $R+T \in E^1(\mathbb{Q})$. En caso contrario, en particular tendríamos que $R \in E^0(\mathbb{Q})$ y como $E^0(\mathbb{Q})$ es un subgrupo, también $E_{\mathrm{tor}}(\mathbb{Q}) \subset E^0(\mathbb{Q})$, de donde llegamos a que $E(\mathbb{Q}) \subset E^0(\mathbb{Q})$, en contra de lo supuesto. ∎

Observemos que $E^1(\mathbb{Q})$ puede contener a lo sumo un número finito de puntos con coordenadas enteras, y que éstos son fáciles de encontrar.

Ejemplo La curva $E'/\mathbb{Q}$ dada por

$$Y^2 = X^3 + 12X^2 - 32X, \qquad \Delta = 2^{18} \cdot 17,$$

tiene rango 1, según hemos visto en el ejemplo de la página 233, donde nos ha aparecido como "pareja" de la curva $E/\mathbb{Q}$ de ecuación

$$Y^2 = X^3 - 6X^2 + 17X, \qquad \Delta = -2^9 \cdot 17^2.$$

Es fácil ver que $E'_{\rm tor}(\mathbb{Q}) = \{(0,0)\}$, así como que los únicos puntos enteros de $E'^1(\mathbb{Q})$ son $(-4, \pm 16)$. Además, $(-4,16)+(0,0) = (8,32)$, luego por el teorema anterior

$$E'(\mathbb{Q}) = \langle (0,0)\rangle \oplus \langle (-4,16)\rangle\,.$$

No podemos aplicar el teorema anterior a la curva E, pues tiene discriminante negativo. Sin embargo, vamos a ver que

$$E(\mathbb{Q}) = \langle (0,0)\rangle \oplus \langle (4,-6)\rangle\,.$$

Para ello consideramos la isogenia $\phi : E \longrightarrow E'$, cuya isogenia dual viene dada por

$$\hat\phi(X,Y) = \left(\frac{Y^2}{4X^2}, \frac{Y(-32-X^2)}{8X^2}\right),$$

Aplicándola al punto $(-4,16)$ obtenemos $\hat\phi(-4,16) = (4,-6)$. Este punto tiene la propiedad de que $\phi(4,-6) = 2(-4,16)$.

Pongamos que $E(\mathbb{Q}) = \langle (0,0)\rangle \oplus \langle P\rangle$. Entonces $(4,-6) = u(0,0) + vP$, con $u,\ v \in \mathbb{Z}$. Aplicando ϕ obtenemos que $2(-4,16) = v\phi(P)$, de donde se sigue fácilmente que $v = \pm 1,\ \pm 2$. Basta descartar las posibilidades $v = \pm 2$. Si se diera una de las dos, tendríamos que

$$(4,-6) \in 2E(\mathbb{Q}), \qquad \text{o bien} \qquad (4,-6)+(0,0) = \left(\frac{17}{4}, -\frac{51}{8}\right) \in 2E(\mathbb{Q}).$$

Veamos que el segundo caso es imposible, y el primero se descarta análogamente. La fórmula de duplicación para E se reduce a

$$x(2(X,Y)) = \left(\frac{X^2-17}{2Y}\right)^2.$$

Si $17/4$ fuera de esta forma, para un punto $(X,Y) \in E(\mathbb{Q})$, tendríamos que

$$\frac{(X^2-17)^2}{17} = Y^2 = X^3 - 6X^2 + 17X.$$

Al operar obtenemos que X ha de ser raíz de un polinomio mónico con coeficientes enteros, luego X ha de ser entero, pero al mismo tiempo se comprueba que dicho polinomio no tiene raíces enteras. ∎

Ejemplo La curva $E/\mathbb{Q}$ dada por la ecuación

$$Y^2 + Y = X^3 - X, \qquad \Delta = 37,$$

cumple $E(\mathbb{Q}) = \langle (0,0)\rangle$.

En efecto, en el capítulo anterior hemos visto que tiene rango 1. El teorema 7.16 implica que un punto de torsión (x,y) ha de cumplir que $2y+1 \in \mathbb{Z}$

y $(2y+1)^2 \mid 37$, lo que obliga a que $2y+1 = \pm 1$, con lo que $y = 0, -1$. Las únicas posibilidades son los puntos

$$(-1,0), \quad (0,0), \quad (1,0), \quad (-1,1), \quad (0,-1), \quad (1,-1),$$

pero ninguno de ellos resulta ser de torsión. Así pues, $E_{\mathrm{tor}}(\mathbb{Q}) = 0$. Los puntos enteros en $E^1(\mathbb{Q})$ son

$$(-1,0), \quad (0,0), \quad (-1,-1), \quad (0,-1),$$

luego sabemos que uno de ellos genera $E(\mathbb{Q})$. Los dos últimos son los opuestos de los dos primeros, luego podemos prescindir de ellos. Se comprueba que $3(0,0) = (-1,-1)$, luego $(-1,-1)$ no puede ser un generador, luego su opuesto $(-1,0)$ tampoco. Concluimos que $(0,0)$ es el generador que buscamos. ∎

Cuando se cumplen las hipótesis del teorema anterior y no hay puntos de orden 2, es fácil encontrar todos los puntos enteros de la curva:

Teorema 9.3 *Sea $E/\mathbb{Z}$ una curva elíptica definida por una ecuación de Weierstrass que cumpla las hipótesis del teorema 9.2 y además $E(\mathbb{Q})[2] = 0$. Entonces todo punto entero de $E(\mathbb{Q})$ es de la forma $2^i P$, donde $i \in \mathbb{N}$ y $P \in E^1(\mathbb{Q})$ es entero.*

Demostración: Por el teorema 9.2 tenemos que $E(\mathbb{Q}) = E_{\mathrm{tor}}(\mathbb{Q}) \otimes \langle Q \rangle$, donde $Q \in E^1(\mathbb{Q})$ es entero. Sea $R \in E(\mathbb{Q})$ un punto entero. Podemos expresarlo como $R = T + 2^i mQ$, donde $m \in \mathbb{Z}$ es impar. Como el grupo de torsión es impar, podemos expresar $T = 2T_1$, para cierto $T_1 \in E_{\mathrm{tor}}(\mathbb{Q})$. Si $i \geq 1$, entonces $R = 2(T_1 + 2^{i-1}mQ)$. Repitiendo el proceso, llegamos a que $R = 2^i(T_i + mQ)$, donde $T_i \in E_{\mathrm{tor}}(\mathbb{Q})$. Por 9.1 sabemos que $P = T_i + mQ$ es entero. Además, como $[Q]$ es no trivial en $E(\mathbb{Q})/E^0(\mathbb{Q})$ y este grupo tiene orden 2, lo mismo le sucede a mQ, es decir, $mQ \in E^1(\mathbb{Q})$. Por otra parte, $T_i = 2T_{i+1} \in E^0(\mathbb{Q})$, luego $P \in E^1(\mathbb{Q})$. ∎

Ejemplo Vamos a calcular los puntos enteros de la curva

$$Y^2 + Y = X^3 - X.$$

En el ejemplo anterior hemos visto que cumple las hipótesis del teorema 9.2, y además $E_{\mathrm{tor}}(\mathbb{Q}) = 0$, luego podemos aplicar el teorema anterior. Hay cuatro puntos enteros en $E^1(\mathbb{Q})$, a saber:

$$(-1,0), \quad (0,0), \quad (-1,-1), \quad (0,-1).$$

Calculamos:

$$2(0,0) = (1,0), \quad 2(1,0) = (2,-3), \quad 2(2,-3) = \left(\frac{21}{25}, -\frac{69}{125}\right),$$

$$2(-1,0) = (6,-15), \quad 2(6,-15) = \left(\frac{1357}{841}, -\frac{53277}{24389}\right).$$

El teorema 9.1 garantiza que ninguno de los múltiplos siguientes será entero. Si partimos de los puntos $(-1,-1)$ y $(0,-1)$ obtendremos los inversos de los puntos que hemos obtenido. Concluimos que $E(\mathbb{Q})$ contiene exactamente 10 puntos enteros:

$$(0,0), \quad (1,0), \quad (2,-3), \quad (-1,0), \quad (6,-15),$$

$$(0,-1), \quad (1,-1), \quad (2,2), \quad (-1,-1), \quad (6,14).$$

Con esto hemos resuelto el Problema 1 planteado en la introducción: encontrar todos los números naturales que pueden expresarse simultáneamente como producto de dos y tres números consecutivos. De los 10 puntos que hemos encontrado, los únicos que dan lugar a soluciones positivas no triviales son $(2,2)$, que corresponde a

$$2 \cdot 3 = 6 = 1 \cdot 2 \cdot 3,$$

y $(6,14)$, que corresponde a

$$14 \cdot 15 = 210 = 5 \cdot 6 \cdot 7.$$

■

Puede probarse que, tras un número finito de multiplicaciones por 2, siempre se llega a un punto no entero, pero esto es un caso particular de un resultado mucho más general que vamos a demostrar. Como ya hemos anunciado, el conjuntos de puntos enteros es necesariamente finito.

9.2 Aproximación diofántica

La prueba del resultado principal que vamos a ver sobre puntos enteros en curvas elípticas se basa en un profundo teorema de aproximación diofántica. En esta sección presentaremos el contexto en el que este teorema surge de forma natural.

La idea básica es que, puesto que $\mathbb{Q}$ es denso en $\mathbb{R}$, cualquier número irracional puede ser aproximado con la precisión deseada por un número racional adecuado. Sin embargo, las aproximaciones pueden ser "malas" en el sentido de tener un denominador muy grande. Por ejemplo, siempre podemos aproximar un irracional salvo una centésima mediante una fracción con denominador 100. Esto es un ejemplo de una mala aproximación. En cambio,

$$\frac{355}{113} = 3,1415929\ldots$$

es una aproximación muy buena de π, pues con un denominador del orden de 100 lo aproxima con 7 cifras decimales exactas. Se trata de un caso particular del teorema siguiente:

Teorema *Si α es un número irracional y p/q es un convergente de la fracción continua de α, entonces*

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

luego existen infinitas aproximaciones de α (con precisión arbitraria) en estas condiciones.

Si α es un número algebraico, Liouville demostró que el exponente 2 del teorema anterior no puede incrementarse indefinidamente:

Teorema 9.4 (Liouville) *Sea $\alpha \in \mathbb{A}$ un número algebraico de grado $d \geq 2$ sobre $\mathbb{Q}$ (es decir, $|\mathbb{Q}(\alpha) : \mathbb{Q}| = d$). Entonces existe una constante $C > 0$ dependiente α tal que para todo número racional p/q se cumple*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{C}{q^d}.$$

Demostración: Sea $f(T) = a_0T^d + a_1T^{d-1} + \cdots + a_d \in \mathbb{Z}[T]$ el polinomio mínimo de α (multiplicado por un $a_0 \in \mathbb{Z}$ para que sus coeficientes sean enteros). Sea

$$C_1 = \sup\{f'(T) \mid \alpha - 1 \leq t \leq \alpha + 1\}.$$

Si un número racional cumple

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

podemos aplicar el teorema del valor medio:

$$|f(p/q)| = |f(\alpha) - f(p/q)| \leq C_1|\alpha - p/q|,$$

pero $q^d f(p/q) \in \mathbb{Z}$ y $f(p/q) \neq 0$, pues f no puede tener raíces racionales, luego $|q^d f(p/q)| \geq 1$. Así pues,

$$\frac{1}{q^d} \leq C_1 \left|\alpha - \frac{p}{q}\right|,$$

siempre y cuando p/q aproxime a α con error menor que 1, pero tomando $C = \text{mín}\{1, 1/C_1\}$ tenemos una constante que vale para todo p/q. ∎

Conviene observar que podemos eliminar la constante C del enunciado a cambio de admitir un número finito de contraejemplos:

Teorema 9.5 *Si α es un número algebraico de grado d, entonces hay a lo sumo una cantidad finita de números racionales p/q que cumplen*

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

Demostración: Para cada valor de q hay a lo sumo una cantidad finita de valores de p que cumplan la aproximación, luego si existieran infinitas aproximaciones, las habría con denominador arbitrariamente grande, luego podríamos tomar una tal que $1/q^d < C/q^d$, donde C es la constante del teorema anterior, lo cual es imposible. ∎

Ejemplo Liouville usó su teorema para mostrar ejemplos concretos sencillos de números trascendentes. Por ejemplo, el número

$$\alpha = 2^{-1!} + 2^{-2!} + 2^{-3!} + \cdots$$

es trascendente, pues si $j \geq j_0$, podemos aproximarlo con p_j/q_j, donde

$$p_j = 2^{j!}(2^{-1!} + 2^{-2!} + \cdots + 2^{-j!}), \qquad q_j = 2^{j!},$$

y la diferencia es

$$\left|\alpha - \frac{p_j}{q_j}\right| = 2^{-(j+1)!} + 2^{-(j+2)!} + \cdots < \sum_{k=0}^{\infty} \frac{1}{2^{(j+1)!+k}} = \frac{1}{2^{(j+1)!-1}} < \frac{1}{q_j^j} \leq \frac{1}{q_j^{j_0}}.$$

Así, p_j/q_j aproxima a α con exponente j_0. Si α fuera algebraico, su grado debería ser $d \geq j_0$ para todo j_0, lo cual es absurdo. ∎

En definitiva, tenemos que cualquier número irracional α puede aproximarse arbitrariamente bien mediante números racionales p/q con un error menor que $1/q^2$, pero que si α es algebraico de grado d no es posible conseguir aproximaciones arbitrariamente buenas con error menor que $1/q^d$. Esto plantea la cuestión de si el exponente 2 es el mejor posible o si, por el contrario, podemos conseguir aproximaciones arbitrariamente buenas con exponente $2 < e < d$. Esto ha sido objeto de numerosos trabajos. Entre otras contribuciones citamos que Thue demostró en 1909 que el teorema de Liouville se cumple también con un exponente $d/2 + 1 + \epsilon$, para todo $\epsilon > 0$. Siegel demostró en 1921 que vale $2\sqrt{d} + \epsilon$, mientras que Gelfond y Dyson demostraron independientemente en 1947 que el teorema se sigue cumpliendo con $\sqrt{2d} + \epsilon$. Finalmente, Roth demostró en 1955 que es suficiente un exponente $2 + \epsilon$, con lo que, a fin de cuentas, el exponente 2 no se puede mejorar (para números algebraicos). Esto es lo que queremos demostrar. En realidad probaremos un resultado más general válido para cuerpos numéricos arbitrarios.

9.3 El teorema de Roth

La versión del teorema de Roth que vamos a demostrar es la siguiente:

Teorema 9.6 (Teorema de Roth) *Sea K un cuerpo numérico, sea $\mathfrak{p}$ un divisor primo de K (arquimediano o no) extendido arbitrariamente a $\mathbb{A}$, sea $\alpha \in \mathbb{A}$ y sea $\epsilon > 0$. Entonces, para cada constante $C > 0$ existe a lo sumo una cantidad finita de números $r \in K$ tales que*

$$|\alpha - r|_{\mathfrak{p}} < \frac{C}{A_K(r)^{2+\epsilon}}.$$

Observemos que este teorema recoge el caso clásico que hemos discutido en la sección anterior, pues si $K = \mathbb{Q}$ y $\mathfrak{p} = \infty$ la altura viene dada por $A_{\mathbb{Q}}(p/q) = \max\{|p|, q\}$. Si para una constante C existieran infinitas aproximaciones con

$|\alpha - p/q| < C/q^{2+\epsilon}$, las habría con q suficientemente grande como para que $C/q^{2+\epsilon} < 1$, pero entonces $|\alpha q - p| < q$ y $|p| < 2|\alpha|q$, luego tendríamos que $A_{\mathbb{Q}}(p/q) < 2|\alpha|q$ y no se cumpliría el teorema para la constante $C(2|\alpha|)^{2+\epsilon}$.

Por otra parte, eligiendo la constante C suficientemente pequeña podemos asegurar que para todo número racional p/q se cumpla $|\alpha - p/q| \geq C/q^{2+\epsilon}$, con lo que tenemos exactamente el teorema de Liouville con $2+\epsilon$ en lugar de d.

Como primer paso en la demostración del teorema de Roth probaremos que es suficiente demostrar una versión ligeramente más débil:

Teorema 9.7 *Sea K un cuerpo numérico, sea $\mathfrak{p}$ un divisor primo de K (arquimediano o no) extendido arbitrariamente a $\mathbb{A}$, sea $\alpha \in \mathbb{A}$ un entero algebraico y sea $\epsilon > 0$. Entonces el conjunto de las alturas de los números $r \in K$ tales que*

$$|\alpha - r|_{\mathfrak{p}} < \frac{1}{A_K(r)^{2+\epsilon}}$$

está acotado.

Demostración (de que este teorema equivale al teorema de Roth): Es evidente que el teorema de Roth implica 9.7. Recíprocamente, supongamos que se cumple 9.7 y veamos que se cumple el teorema de Roth. En primer lugar supongamos que α es un entero algebraico. Si existieran infinitas aproximaciones $r \in K$ para una constante dada C, entonces

$$\frac{C}{A_K(r)^{2+\epsilon}} = \frac{1}{A_K(r)^{2+\epsilon/2}} \cdot \frac{C}{A_K(r)^{\epsilon/2}}$$

y tendríamos infinitas aproximaciones para las que $C/A_K(r)^{\epsilon/2} < 1$, las cuales también incumplirían el teorema con $C = 1$ y $\epsilon/2$ (al ser infinitas, sus alturas no estarían acotadas).

Así pues, tenemos probado el teorema de Roth para enteros algebraicos. Si α no es necesariamente entero, existe un entero racional $m \neq 0$ tal que $m\alpha$ es un entero algebraico. Aplicamos el teorema 7.24 a la aplicación $\mathrm{P}^1 \longrightarrow \mathrm{P}^1$ dada por $[x, y] \mapsto [mx, y]$, con lo que obtenemos constantes tales que

$$C_1 A_K(r) \leq A_K(mr) \leq C_2 A_K(r),$$

para todo $r \in K$. (El teorema está enunciado para la altura absoluta, pero es obvio que también vale para la altura de K.) Así, si

$$|\alpha - r|_{\mathfrak{p}} < \frac{C}{A_K(r)^{2+\epsilon}},$$

entonces

$$|m\alpha - mr|_{\mathfrak{p}} < \frac{C|m|_{\mathfrak{p}}}{A_K(r)^{2+\epsilon}} < \frac{C|m|_{\mathfrak{p}} C_1^{-2-\epsilon}}{A_K(mr)^{2+\epsilon}},$$

luego cada aproximación a α da lugar a una aproximación a $m\alpha$ con otra constante. La conclusión es ahora obvia. ∎

Vamos a enunciar dos teoremas que demostraremos en la sección siguiente en los cuales nos basaremos para probar el teorema de Roth. Para ello hemos de introducir cierta notación: Sea K un cuerpo numérico y consideremos un polinomio $G(X_1,\dots,X_m)\in K[X_1,\dots,X_m]$. Sean $r_1,\dots,r_m$ números naturales de modo que $\operatorname{grad}_{X_i} G\le r_i$. Si T es otra indeterminada, podemos considerar $G\in K[T][X_1,\dots,X_m]$ y desarrollar G como polinomio de Taylor alrededor del punto $(T,\dots,T)$:

$$G(X_1,\dots,X_m)=\sum_i G^{(i)}(T,\dots,T)(T-X_1)^{i_1}\cdots(T-X_n)^{i_m},$$

donde $i=(i_1,\dots,i_m)$ recorre las m-tuplas $0\le i_j\le r_j$ y

$$G^{(i)}=\frac{1}{i_1!\cdots i_m!}\left.\frac{\partial^{i_1+\cdots i_m}G}{\partial X_1^{i_1}\cdots\partial X_m^{i_m}}\right|_{(T,\dots,T)}.$$

Observemos que los coeficientes de $G^{(i)}$ son combinaciones enteras de los coeficientes de G, pues si $k\ge i$

$$\frac{1}{i!}\frac{\partial^i X_j^k}{\partial X_j^k}=\binom{k}{i}X_j^{k-i}.$$

Una cota de los coeficientes enteros que aparecen es

$$\binom{r_1}{i_1}\cdots\binom{r_m}{i_m}\le 2^{r_1+\cdots+r_m}.$$

Si $\mathfrak{p}$ es un divisor primo en K, representamos por $|G|_{\mathfrak{p}}$ el máximo de los valores absolutos de los coeficientes de G. Definimos también $|G|_\infty=\max_{\mathfrak{p}|\infty}|G|_{\mathfrak{p}}$.

Enunciamos ahora los dos teoremas que vamos a usar:

Teorema 9.8 *Sea K un cuerpo numérico, sea $\mathcal{O}$ su anillo de enteros y sea $F(Y)\in K[Y]$ un polinomio mónico de grado n. Para cada número $0<\eta<1/2$, cada natural $m>(2n/\eta)^2$ y cada sucesión de naturales $r_1,\dots,r_m$ no nulos existe un polinomio $G(X_1,\dots,X_m)\in\mathcal{O}[X_1,\dots,X_m]$ no nulo que verifica:*

a) $\operatorname{grad}_{X_j} G\le r_j$.

b) Si

$$\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\le m\left(\frac{1}{2}-\eta\right),$$

entonces toda raíz α de F cumple $G^{(i)}(\alpha,\dots,\alpha)=0$.

c) Existe una constante C (que depende de F, pero no de η, m ni de los r_j) tal que $|G|_\infty\le C^{r_1+\cdots+r_m}$.

Teorema 9.9 *Sea K un cuerpo numérico de grado N, sea $\mathcal{O}$ su anillo de enteros, $m > 0$ un número natural, $0 < \delta < 1/16^{m+N}$, sean $r_1, \dots, r_m$ naturales no nulos tales que*

$$\delta r_1 > r_2,\ \delta r_2 > r_3, \dots, \delta r_m > 10. \tag{9.1}$$

Sea $G(X_1, \dots, X_m) \in \mathcal{O}[X_1, \dots, X_m]$ no nulo tal que $\operatorname{grad}_{X_j} G \le r_j$. Consideremos $\beta_1, \dots, \beta_m \in K$ números de altura $A_1, \dots, A_m$ respectivamente y supongamos:

$$N \log 4 + 4m \le \delta \log A_1, \tag{9.2}$$

$$r_1 \log A_1 \le r_j \log A_j, \qquad j = 1, \dots, m, \tag{9.3}$$

$$|G|_\infty \le A_1^{\delta r_1}. \tag{9.4}$$

Entonces existe una m-tupla $s = (s_1, \dots, s_m)$ de números $0 \le s_j \le r_j$ tal que

$$s_1/r_1 + \cdots + s_m/r_m \le 20^m \delta^{(1/2)^m} \tag{9.5}$$

y $G^{(s)}(\beta_1, \dots, \beta_m) \neq 0$.

Admitiendo estos resultados, vamos a probar el teorema 9.7:

Partimos de un cuerpo numérico K de grado N y un entero algebraico $\alpha \in \mathbb{A}$. Sea $F(Y) \in \mathcal{O}[Y]$ su polinomio mínimo, digamos de grado n. Supongamos que existen elementos $\beta \in K$ que aproximan a α según el enunciado (respecto de un primo $\mathfrak{p}$) con altura arbitrariamente grande. Fijemos un número real $0 < \eta < 1/10$ y vamos a elegir m, δ, $\beta_1, \dots, \beta_m$, $r_1, \dots, r_m$ en este orden con los criterios siguientes:

Tomamos $m > (2n/\eta)^2$, de modo que podemos aplicar el teorema 9.8 a F con η y m.

Luego tomamos $0 < \delta < \eta$ tal que $\delta < 1/16^{m+N}$, $20^m \delta^{(1/2)^m} < \eta$.

A continuación elegimos $\beta_1 \in K$ que aproxime a α y con altura A_1 suficientemente grande como para que cumpla (9.2) y

$$\delta(1 - \delta) \log A_1 > \log C$$

(la constante de 9.8 c).

Seguidamente elegimos $\beta_2, \dots, \beta_m \in K$ que aproximen a α y cuyas alturas $A_2, \dots, A_m$ cumplan

$$\delta \log A_2 > 2 \log A_1, \dots, \delta \log A_m > 2 \log A_{m-1}.$$

Luego elegimos un natural r_1 tal que

$$r_1 \delta \log A_1 > 10 \log A_j, \qquad j = 2, \dots, m.$$

Y por último elegimos naturales $r_2, \dots, r_m$ de modo que

$$r_1 \frac{\log A_1}{\log A_j} \le r_j < 1 + r_1 \frac{\log A_1}{\log A_j}.$$

De este modo

$$r_1 \log A_1 \le r_j \log A_j \le \log A_j + r_1 \log A_1 < (1+\delta/10) r_1 \log A_1. \tag{9.6}$$

La primera desigualdad es (9.3). Ahora observamos que los r_j cumplen (9.1). En efecto,

$$\frac{r_{j+1}}{r_j} < \left(1+\frac{\delta}{10}\right)\frac{\log A_j}{\log A_{j+1}} < \left(1+\frac{\delta}{10}\right)\frac{\delta}{2} < \delta$$

y

$$\delta r_m \ge \delta r_1 \frac{\log A_1}{\log A_m} > 10.$$

Podemos aplicar 9.8, que nos da un polinomio $G(X_1,\dots,X_n)$ que cumple

$$|G|_\infty \le C^{r_1+\cdots+r_m} \le C^{r_1(1+\delta+\cdots+\delta^{m-1})} \le C^{r_1/(1-\delta)} < A_1^{r_1\delta},$$

por la elección de A_1, luego se cumplen todas las hipótesis de 9.9. Así pues, existe una m-tupla $s = (s_1,\dots,s_m)$ de manera que $s_1/r_1 + \cdots + s_m/r_m \le \eta$ y $G^{(s)}(\beta_1,\dots,\beta_m) \ne 0$. Si llamamos $\overline{G} = G^{(s)}$, tenemos que si una m-tupla i cumple

$$i_1/r_1 + \cdots + i_m/r_m \le m(1/2-\eta) - \eta,$$

entonces $s+i$ está en las hipótesis del teorema 9.8, luego $\overline{G}^{(i)}(\alpha,\dots,\alpha) = 0$.

Consideramos el desarrollo de Taylor

$$\overline{G}(\beta_1,\dots,\beta_m) = \sum_i \overline{G}^{(i)}(\alpha,\dots,\alpha)(\alpha-\beta_1)^{i_1}\cdots(\alpha-\beta_m)^{i_m},$$

en el que son nulos todos los términos salvo los correspondientes a m-tuplas i tales que

$$i_1/r_1 + \cdots + i_m/r_m \ge m(1/2-2\eta). \tag{9.7}$$

(Notemos que $m(1/2-2\eta) > m(1/2-\eta)-\eta$.) De todos modos, el número total de sumandos es a lo sumo

$$2^{r_1+\cdots+r_m} \le 2^{mr_1} \le A_1^{\delta r_1 m}.$$

(En la elección de A_1 podemos exigir que $\delta \log A_1 \ge \log 2$.)

Vamos a acotar $|\overline{G}^{(i)}(\alpha,\dots,\alpha)|_\mathfrak{p}$. Para ello observamos que F está formado por a lo sumo $(r_1+1)\cdots(r_m+1) \le 2^{r_1+\cdots+r_m} \le 2^{mr_1}$ monomios. Al pasar a $\overline{G} = F^{(s)}$ y luego a $\overline{G}^{(i)}$ lo que hacemos es reducir el grado de cada monomio, multiplicarlo por un entero $\le 2^{r_1+\cdots+r_m}$ y luego evaluar en α, por lo que una cota del valor absoluto de cada monomio reducido de esta forma es

$$|F|_\mathfrak{p}|2\alpha|_\mathfrak{p}^{r_1+\cdots+r_m} \le |F|_\mathfrak{p}|2\alpha|_\mathfrak{p}^{mr_1}.$$

En total,

$$|\overline{G}^{(i)}(\alpha,\dots,\alpha)|_\mathfrak{p} \le 2^{mr_1}|F|_\mathfrak{p}|2\alpha|_\mathfrak{p}^{mr_1} \le C_1^{mr_1} \le A_1^{\delta r_1 m},$$

exigiendo $\log C_1 \le \delta \log A_1$ en la elección de A_1. Así pues,

$$|\overline{G}(\beta_1,\dots,\beta_m)|_\mathfrak{p} \le A_1^{2\delta m r_1} \max_i |\alpha-\beta_1|_\mathfrak{p}^{i_1}\cdot|\alpha-\beta_m|_\mathfrak{p}^{i_m},$$

donde i recorre las m-tuplas que cumplen (9.7).

Como los β_j aproximan a α, tenemos que

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}} \le A_1^{2\delta m r_1}\frac{1}{(A_1^{i_1}\cdots A_m^{i_m})^{2+\epsilon}},$$

para cierta m-tupla i. Si elevamos al grado local $N_{\mathfrak{p}}$, el último factor se hace menor, luego no hace falta elevarlo:

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \le A_1^{2\delta m r_1 N_{\mathfrak{p}}}\frac{1}{(A_1^{i_1}\cdots A_m^{i_m})^{2+\epsilon}},$$

Ahora cambiamos cada i_j por $r_j i_j/r_j$ y usamos que $A_1^{r_1} \le A_j^{r_j}$ por (9.6). Luego sustituimos $i_1/r_1 + \cdots + i_m/r_m$ por $m(1/2-2\eta)$, con lo que obtenemos

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \le A_1^{2\delta m r_1 N_{\mathfrak{p}}}\frac{1}{A_1^{m r_1(1/2-2\eta)(2+\epsilon)}}. \tag{9.8}$$

Sea ahora $\mathfrak{p}$ un primo arquimediano de K distinto del que estamos usando para aproximar α. De nuevo tenemos en cuenta que $\overline{G}$ está formado por a lo sumo 2^{mr_1} monomios cuyos coeficientes están acotados por $2^{mr_1}|F|_\infty$. Por consiguiente,

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}} \le 4^{mr_1}|F|_\infty |\beta_1|_{\mathfrak{p}}^{r_1}\cdots|\beta_m|_{\mathfrak{p}}^{r_m}.$$

Podemos definir A_1 de modo que $4^{mr_1}|F|_\infty \le A_1^{2\delta m r_1}$, con lo que

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \le A_1^{2\delta m r_1 N_{\mathfrak{p}}}\,\text{máx}\{|\beta_1|_{\mathfrak{p}}^{N_{\mathfrak{p}}},1\}^{r_1}\cdots\text{máx}\{|\beta_m|_{\mathfrak{p}}^{N_{\mathfrak{p}}},1\}^{r_m}. \tag{9.9}$$

Si $\mathfrak{p}$ es un primo no arquimediano, entonces no hemos de incluir la cota 2^{mr_1} del número de sumandos gracias la desigualdad triangular no arquimediana y, como F tiene los coeficientes enteros, en lugar de $2^{mr_1}|F|_\infty$ tenemos simplemente un 1. Así pues,

$$|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \le \text{máx}\{|\beta_1|_{\mathfrak{p}}^{N_{\mathfrak{p}}},1\}^{r_1}\cdots\text{máx}\{|\beta_m|_{\mathfrak{p}}^{N_{\mathfrak{p}}},1\}^{r_m}. \tag{9.10}$$

Si multiplicamos para todo primo $\mathfrak{p}$ de K la expresión correspondiente (9.8), (9.9) y (9.10), en el miembro derecho aparecerá la altura de cada β_j salvo por el factor $\text{máx}\{|\beta_j|_{\mathfrak{p}}^{N_{\mathfrak{p}}},1\}$ correspondiente al primo del enunciado, pero podemos añadirlo porque con ello hacemos mayor el producto. El resultado es:

$$\prod_{\mathfrak{p}}|\overline{G}(\beta_1,\dots,\beta_m)|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \le \frac{A_1^{2\delta m r_1 N'}A_1^{r_1}\cdots A_m^{r_m}}{A_1^{m r_1(1/2-2\eta)(2+\epsilon)}},$$

donde N' es la suma de los grados locales de los primos arquimedianos y el del enunciado (si es que éste no es arquimediano). Por el teorema 9.9 sabemos que $\overline{G}(\beta_1,\dots,\beta_m) \ne 0$, luego por la fórmula del producto el miembro izquierdo resulta ser igual a 1. En el miembro derecho usamos la relación (9.6), de modo que $A_j^{r_j} < A_1^{r_1(1+\eta)}$ y, por consiguiente,

$$1 \le \frac{A_1^{2\delta m r_1 N'}A_1^{m r_1(1+\eta)}}{A_1^{m r_1(1/2-2\eta)(2+\epsilon)}}.$$

Tomando logaritmos:

$$(1/2 - 2\eta)(2 + \epsilon) \leq 2\delta N' + 1 + \eta.$$

Ahora bien, fijado η, esto es válido para δ arbitrariamente pequeño, por lo que

$$(1/2 - 2\eta)(2 + \epsilon) \leq 1 + \eta.$$

A su vez, η es arbitrariamente pequeño, luego llegamos a que $\epsilon \leq 0$, lo cual es absurdo. ∎

9.4 Resultados auxiliares

En esta sección demostraremos los teoremas 9.8 y 9.9 que hemos usado para probar el teorema de Roth. Para el primero de ellos necesitamos dos resultados previos:

Teorema 9.10 *Si K es un cuerpo numérico de grado N y $\mathcal{O}$ es su anillo de enteros, existen constantes C_1, C_2, $B_0 > 0$ tales que para todo $B \geq B_0$ el conjunto de los números $\alpha \in \mathcal{O}$ tales que $|\alpha|_\mathfrak{p} \leq B$ para todo primo arquimediano $\mathfrak{p}$ de K es finito y su cardinal $\lambda(B)$ cumple*

$$C_1 B^N \leq \lambda(B) \leq C_2 B^N.$$

Demostración: Esto se debe a que la representación geométrica transforma $\mathcal{O}$ en un retículo completo en $\mathbb{R}^s \times \mathbb{C}^t \cong \mathbb{R}^N$, donde s y t son el número de primos arquimedianos reales y complejos de K, respectivamente. Si llamamos

$$T = \{x \in \mathbb{R}^s \times \mathbb{C}^t \mid |x_i| \leq 1, \text{ para } i = 1, \ldots, s + t\},$$

entonces T es un subconjunto acotado de $\mathbb{R}^N$ y el conjunto del enunciado está formado por los puntos cuya representación geométrica está en el homotético BT. Obviamente el conjunto es finito y la estimación de su cardinal se basa en un resultado general sobre retículos completos: Si v es el volumen de T y V es el volumen del paralelepípedo fundamental del retículo, entonces[1]

$$\lambda(B) = \frac{v}{V} B^N + O(B^{N-1}).$$

Explícitamente, existe una constante $C > 0$ tal que

$$\frac{v}{V} - \frac{C}{B} < \frac{\lambda(B)}{B^N} < \frac{v}{V} + \frac{C}{B}.$$

Si tomamos B_0 suficientemente grande como para que $C/B_0 < v/2V$, entonces, para todo $B \geq B_0$ se cumple

$$0 < \frac{v}{2V} < \frac{\lambda(B)}{B^N} < \frac{3v}{2V},$$

lo que prueba el teorema. ∎

[1] Teorema 11.6 de mi Teoría de números. La hipótesis sobre que la frontera sea parametrizable Lipschitz se cumple trivialmente en este caso.

Teorema 9.11 *Sean $r_1, \dots, r_m$ naturales no nulos y $\eta > 0$. El número e de m-tuplas $(i_1, \dots, i_m)$ que cumplen $0 \le i_j \le r_j$,*

$$\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} \le m\left(\frac{1}{2} - \eta\right)$$

es a lo sumo $(r_1+1)\cdots(r_m+1)/\eta m^{1/2}$.

Demostración: Por inducción sobre m. Para $m = 1$ es trivial, pues si $\eta \ge 1/2$ no hay ninguno y si $\eta < 1/2$ entonces $(r_1+1)/\eta > r_1 + 1$.

Supongamos ahora que $m > 1$ y tomemos λ determinado por

$$m\left(\frac{1}{2} - \eta\right) = \frac{1}{2}(m - \lambda),$$

es decir, $\lambda = 2m\eta$. En términos de λ la cota superior es

$$\frac{2m^{1/2}}{\lambda}(r_1+1)\cdots(r_m+1).$$

La afirmación es trivial si $\lambda \le 2m^{1/2}$. Supongamos, pues, que $\lambda > 2m^{1/2}$. El número de m-tuplas i será la suma del número de $m-1$-tuplas que cumplen

$$\frac{i_1}{r_1} + \cdots + \frac{i_{m-1}}{r_{m-1}} \le \frac{1}{2}(m-\lambda) - \frac{i_m}{r_m} = \frac{m-1}{2} - \frac{\lambda - 1 + 2i_m/r_m}{2},$$

para cada $i_m = 0, \dots, r_m$. Por hipótesis de inducción, el número de m-tuplas será a lo sumo

$$\sum_{i_m=0}^{r_m} \frac{2(m-1)^{1/2}}{\lambda - 1 + 2i_m/r_m}(r_1+1)\cdots(r_{m-1}+1).$$

Ahora observamos que

$$\sum_{i_m=0}^{r_m} \frac{2}{\lambda - 1 + 2i_m/r_m} = \sum_{i_m=0}^{r_m}\left(\frac{1}{\lambda - 1 + 2i_m/r_m} + \frac{1}{\lambda + 1 - 2i_m/r_m}\right)$$

(Los dos sumandos recorren la misma sucesión de números en sentidos opuestos.)

$$= \sum_{i_m=0}^{r_m} \frac{2\lambda}{\lambda^2 - (1 - 2i_m/r_m)^2} \le \sum_{i_m=0}^{r_m} \frac{2\lambda}{\lambda^2 - 1} = (r_m+1)\frac{2\lambda}{\lambda^2-1}.$$

Así pues, el número de m-tuplas está acotado por

$$\frac{2\lambda(m-1)^{1/2}}{\lambda^2-1}(r_1+1)\cdots(r_m+1).$$

Falta comprobar que

$$\frac{2\lambda(m-1)^{1/2}}{\lambda^2-1} \le \frac{2m^{1/2}}{\lambda}$$

o, equivalentemente, que

$$\left(\frac{m-1}{m}\right)^{1/2} \leq \frac{\lambda^2-1}{\lambda^2}.$$

Notemos que $(1-1/m)^{1/2} \leq 1-1/2m$ (basta elevar al cuadrado), luego es suficiente probar que $1-1/2m \leq 1-1/\lambda^2$, pero esto se sigue de la condición $\lambda \geq 2m^{1/2}$ que estamos suponiendo. ∎

Demostración de 9.8: Sean B_0, C_1 y C_2 las constantes asociadas a K por el teorema 9.10. Fijemos un número $B \geq B_0$. El número de polinomios $G \in \mathcal{O}[X_1,\ldots,X_m]$ que cumplen $\operatorname{grad}_{X_j} G \leq r_j$ y $|G|_\infty \leq B$ es como mínimo $(C_1B)^{N(r_1+1)\cdots(r_m+1)}$.

Para cada m-tupla i que cumpla la hipótesis b) del teorema, sea $RG^{(i)}$ el resto de la división de $G^{(i)}(Y,\ldots,Y)$ entre $F(Y)$. Observemos que G consta de un máximo de $(r_1+1)\cdots(r_m+1) \leq 2^{r_1+\cdots+r_m}$ monomios de grado menor o igual que $r_1+\cdots+r_m$. Al calcular $G^{(i)}$ obtenemos una suma de no más de $2^{r_1+\cdots+r_m}$ monomios de grado menor o igual que $r_1+\cdots+r_m$ con coeficientes acotados (respecto de cualquier valor absoluto arquimediano) por $2^{r_1+\cdots+r_m}|G|_\infty$. Al sustituir las variables por Y, se agruparán los monomios del mismo grado, luego quedarán a lo sumo $2^{r_1+\cdots+r_m}$ monomios de grado menor o igual que $r_1+\cdots+r_m$ con coeficientes acotados por $4^{r_1+\cdots+r_m}|G|_\infty$.

Llamemos $C_0 = \max\{|F|_\infty, 1\}$. Para dividir $G^{(i)}(Y,\ldots,Y)$ entre $F(Y)$, el primer paso es calcular $G^{(i)}(Y,\ldots,Y) - aF(Y)$, donde a es el coeficiente director del dividendo. El polinomio que resulta tiene sus coeficientes acotados por

$$4^{r_1+\cdots+r_m}|G|_\infty + 4^{r_1+\cdots+r_m}|G|_\infty C_0 \leq 2\cdot 4^{r_1+\cdots+r_m} C_0|G|_\infty.$$

En el siguiente paso de la división euclídea, al polinomio resultante le restamos otro polinomio de la forma $aF(Y)$, con lo que el nuevo polinomio tiene sus coeficientes acotados por

$$2\cdot 4^{r_1+\cdots+r_m} C_0|G|_\infty + 2\cdot 4^{r_1+\cdots+r_m} C_0|G|_\infty C_0 \leq 2^2\cdot 4^{r_1+\cdots+r_m} C_0^2|G|_\infty.$$

La división euclídea se ha de realizar a lo sumo en tantos pasos como el grado del dividendo, o sea, en no más de $r_1+\cdots+r_m$ pasos, luego llegamos a que

$$|RG^{(i)}|_\infty \leq 2^{r_1+\cdots+r_m}\cdot 4^{r_1+\cdots+r_m} C_0^{r_1+\cdots+r_m}|G|_\infty = C_3^{r_1+\cdots+r_m}|G|_\infty,$$

donde la constante C_3 sólo depende de F.

En resumen, a cada polinomio G le hemos asociado un conjunto de a lo sumo e restos $RG^{(i)}$ (donde e está acotado según el teorema anterior) de grado menor que n (el grado de F) y tales que $|RG^{(i)}|_\infty \leq C_3^{r_1+\cdots+r_m}B$. Por el teorema 9.10, tales restos varían en un conjunto de a lo sumo $C_2(C_3^{r_1+\cdots+r_m}B)^{Nn}$ polinomios, luego los conjuntos de restos asociados a cada G varían en un conjunto de a lo sumo

$$(C_4^{r_1+\cdots+r_m}B)^{Nne} \leq (C_4^{r_1+\cdots+r_m}B)^{Nn(r_1+1)\cdots(r_m+1)/\eta m^{1/2}}$$

conjuntos, donde la constante C_4 sólo depende de F.

El teorema quedará probado si justificamos que tomando B suficientemente grande podemos conseguir dos polinomios distintos G_1 y G_2 con los mismos restos, pues entonces el polinomio $G = G_1 - G_2$ seguirá cumpliendo las propiedades a) y c) de 9.8, pero además cumplirá b), pues los $G^{(i)}(Y,\dots,Y)$ correspondientes a índices en las hipótesis de b) serán múltiplos de F, luego se anularán en sus raíces.

Dichos polinomios existirán si

$$(C_4^{r_1+\cdots+r_m}B)^{Nn(r_1+1)\cdots(r_m+1)/\eta m^{1/2}} < (C_1B)^{N(r_1+1)\cdots(r_m+1)}.$$

Equivalentemente, $(C_4^{r_1+\cdots+r_m}B)^{Nn/\eta m^{1/2}} < (C_1B)^N$. Tenemos como hipótesis que $m > (2n/\eta)^2$, luego nos basta conseguir

$$(C_4^{r_1+\cdots+r_m}B)^{N/2} < (C_1B)^N.$$

Esto equivale a

$$C_5^{r_1+\cdots+r_m} < C_1^N B^{N/2}.$$

Podemos suponer que $C_1 < 1$, con lo que $C_1^{-N} > 1$ y por consiguiente $C_5^{r_1+\cdots+r_m} < (C_1^{-N}C_5)^{r_1+\cdots+r_m}$, luego basta conseguir

$$C_6^{r_1+\cdots+r_m} < B^{N/2},$$

o también $C_7^{r_1+\cdots+r_m} < B$. Ahora, el apartado c) del enunciado se cumple con la constante $C = C_7 + 1$, pues si tomamos

$$C_7^{r_1+\cdots+r_m} < B < C^{r_1+\cdots+r_m}$$

encontramos un polinomio G que cumple a) y b) y $|G|_\infty \le B < C^{r_1+\cdots+r_m}$, tal y como exige c). ∎

La demostración de 9.9 requiere varios resultados adicionales. Los teoremas siguientes valen para cuerpos arbitrarios K de característica 0, si bien los necesitamos únicamente para cuerpos numéricos.

Un *operador diferencial* en $K[X_1,\dots,X_m]$ es un operador de la forma:

$$D = \frac{1}{i_1!\cdots i_m!}\frac{\partial^r}{\partial X_1^{i_1}\cdots\partial X_m^{i_m}},$$

donde $r = i_1+\cdots+i_m$ es el *grado* del operador. Claramente D es una aplicación lineal en $K[X_1,\dots,X_m]$ que cumple además la relación

$$D(FG) = G\,DF + F\,DG.$$

Admitimos la identidad como único operador diferencial de grado 0.

Teorema 9.12 *Sean $F_1,\dots,F_n \in K[X_1,\dots X_m]$ polinomios linealmente independientes sobre K. Entonces existen operadores diferenciales $D_1,\dots,D_n$ tales que cada D_j tiene grado menor que j y $\det(D_jF_i) \neq 0$.*

DEMOSTRACIÓN: Lo probaremos por inducción sobre n. Si $n = 1$ ha de ser $F_1 \neq 0$ y basta tomar como D_1 la identidad. Supongamos el teorema para $n-1$. Dados n polinomios linealmente independientes, por hipótesis de inducción existen operadores diferenciales $D_1, \ldots, D_{n-1}$ tales que el grado de cada D_j es menor que j y

$$\begin{vmatrix} F_1 & F_2 & \cdots & F_{n-1} \\ D_2F_1 & D_2F_2 & \cdots & D_2F_{n-1} \\ \vdots & \vdots & & \vdots \\ D_{n-1}F_1 & D_{n-1}F_2 & \cdots & D_{n-1}F_{n-1} \end{vmatrix} \neq 0.$$

Supongamos que para todo operador diferencial D de grado menor que n se cumple que

$$\begin{vmatrix} F_1 & F_2 & \cdots & F_{n-1} & F_n \\ D_2F_1 & D_2F_2 & \cdots & D_2F_{n-1} & D_2F_n \\ \vdots & \vdots & & \vdots & \vdots \\ D_{n-1}F_1 & D_{n-1}F_2 & \cdots & D_{n-1}F_{n-1} & D_{n-1}F_n \\ DF_1 & DF_2 & \cdots & DF_{n-1} & DF_n \end{vmatrix} = 0.$$

Pensemos en las columnas de este determinante como vectores del espacio $K(X_1, \ldots, X_m)^n$. Las primeras $n-1$ columnas son linealmente independientes, mientras que la última es combinación lineal de las anteriores. Ahora bien, los coeficientes de la combinación lineal no dependen de D, pues son los únicos que expresan la columna $(F_n, D_2F_n, \ldots, D_{n-1}F_n)$ como combinación lineal de las columnas anteriores en $K(X_1, \ldots, X_m)^{n-1}$. Concluimos que existen funciones racionales $G_1, \ldots, G_n$ tales que $G_n = 1$ y para todo operador diferencial D de grado menor que n se cumple

$$G_1DF_1 + \cdots + G_nDF_n = 0. \tag{9.11}$$

Derivando respecto de X_j obtenemos:

$$G_1\frac{\partial}{\partial X_j}DF_1 + \cdots + G_n\frac{\partial}{\partial X_j}DF_n + \frac{\partial G_1}{\partial X_j}DF_1 + \cdots + \frac{\partial G_n}{\partial X_j}DF_n = 0.$$

Apliquemos esto al caso particular en que $D = D_i$, $i = 1, \ldots, n-1$. Entonces

$$\frac{1}{i_j + 1}\frac{\partial}{\partial X_j}D$$

es un operador diferencial de grado menor que $i+1$, luego cumple (9.11), y lo sigue cumpliendo si eliminamos el coeficiente. Por lo tanto

$$\frac{\partial G_1}{\partial X_j}D_iF_1 + \cdots + \frac{\partial G_n}{\partial X_j}D_iF_n = 0, \qquad i = 1, \ldots, n-1.$$

Notemos que como $G_n = 1$ en estas $n-1$ ecuaciones no aparece realmente D_iF_n. Si las consideramos como un sistema de $n-1$ ecuaciones lineales con

incógnitas las derivadas de los G_k, la matriz de coeficientes (D_iF_k) tiene determinante no nulo, luego podemos concluir que todas las derivadas parciales son nulas, luego las funciones racionales G_k son todas constantes. Aplicando (9.11) al operador identidad concluimos que los polinomios F_i son linealmente dependientes sobre K, en contra de lo supuesto. ∎

Teorema 9.13 *Sea $G \in K[X_1,\dots,X_m]$ un polinomio no nulo, $m \geq 2$, tal que $\operatorname{grad}_{X_j} G \leq r_j$ para cada $j = 1,\dots,m$. Entonces existe un natural $1 \leq l \leq r_m+1$ y operadores diferenciales $D_0,\dots,D_{l-1}$ en las variables $X_1,\dots,X_{m-1}$ tales que el grado de cada D_i es menor o igual que i y si*

$$F(X_1,\dots,X_m) = \det\left(D_i \frac{1}{j!}\frac{\partial^j G}{\partial X_m^j}\right)_{i,j=0}^{l-1}$$

entonces $F \neq 0$ y $F(X_1,\dots,X_m) = U(X_1,\dots,X_{m-1})V(X_m)$, donde U y V son polinomios con $\operatorname{grad}_{X_j} U \leq lr_j$, para $j = 1,\dots,m-1$ y $\operatorname{grad} V \leq lr_m$.

Demostración: Consideremos todas las representaciones posibles de G de la forma

$$G = \phi_0(X_m)\psi_0(X_1,\dots,X_{m-1}) + \cdots + \phi_{l-1}(X_m)\psi_{l-1}(X_1,\dots,X_{m-1}),$$

donde los polinomios ϕ_i y ψ_i tienen grado en X_j menor o igual que r_j.

Siempre podemos encontrar una representación así haciendo $l-1 = r_m$ y $\phi_k(X_m) = X_m^k$, pero de entre todas las posibles tomamos una con l mínimo. Obviamente $0 \leq l \leq r_m + 1$.

Los polinomios $\phi_i(X_m)$ son linealmente independientes sobre K, pues de lo contrario podríamos conseguir una representación con l menor. Por el mismo motivo, los polinomios ψ_i son linealmente independientes.

Aplicamos el teorema anterior a ambos conjuntos de polinomios. Para los ψ_i encontramos operadores $D_0,\dots,D_{l-1}$ en las variables $X_1,\dots,X_{m-1}$ tales que el grado de cada D_i es menor o igual que i y

$$U(X_1,\dots,X_{m-1}) = \det(D_i\psi_k) \neq 0.$$

En el caso de los ϕ_i, como los operadores diferenciales han de ser distintos dos a dos (o si no el determinante formado con ellos sería nulo trivialmente), éstos no pueden ser sino las derivadas parciales de orden $0,1,\dots,l-1$ (con los coeficientes correspondientes). Reordenando los polinomios si es preciso concluimos que

$$V(X_m) = \det\left(\frac{1}{j!}\frac{\partial^j \phi_k}{\partial X_m^j}\right) \neq 0.$$

Como ambos determinantes son $l \times l$, es claro que U y V tienen los grados que indica el enunciado. También es obvio que $F = UV$. ∎

Teorema 9.14 *En las hipótesis del teorema anterior, sea* $\mathfrak{p}$ *un divisor primo de* K*. Si* $\mathfrak{p}$ *no es arquimediano, entonces* $|F|_\mathfrak{p} \le |G|_\mathfrak{p}^l$*, mientras que si* $\mathfrak{p}$ *es arquimediano, entonces*

$$|F|_\mathfrak{p} \le |(r_1+1)\cdots(r_m+1)|^l\, l!\, 2^{l(r_1+\cdots+r_m)}|G|_\mathfrak{p}^l.$$

DEMOSTRACIÓN: Si $\mathfrak{p}$ no es arquimediano el teorema es trivial, pues sólo hay que estimar cada sumando del determinante que define a F, que es un producto de l polinomios con valor absoluto igual al de G (ya que los operadores diferenciales no aumentan el valor absoluto de los coeficientes).

Si $\mathfrak{p}$ es arquimediano consideramos a G como suma de $(r_1+1)\cdots(r_m+1)$ monomios, con lo que podemos desarrollar el determinante que define a F como suma de $|(r_1+1)\cdots(r_m+1)|^l$ determinantes, cada uno de los cuales tiene coeficientes de la forma

$$aD_i\frac{1}{j!}\frac{\partial^j}{\partial X_m^j}X_1^{i_1}\cdots X_m^{i_m},$$

donde a es un coeficiente de G. Las derivadas forman un operador diferencial que produce un coeficiente entero menor o igual que $2^{r_1+\cdots+r_m}$, luego cada término del determinante tiene valor absoluto menor o igual que $2^{l(r_1+\cdots+r_m)}|G|_\mathfrak{p}^l$. Por último, el determinante tiene $l!$ sumandos, lo que nos lleva a la cota del enunciado. ∎

Consideremos un polinomio no nulo $F \in K[X_1,\ldots,X_m]$, sean $r_1,\ldots,r_m$ naturales no nulos y $\beta_1,\ldots,\beta_m \in K$. Tenemos que F admite un desarrollo de Taylor

$$F(X_1,\ldots,X_m) = \sum_i F^{(i)}(\beta_1,\ldots,\beta_m)(X_1-\beta_1)^{i_1}\cdots(X_m-\beta_m)^{i_m}.$$

Por abreviar, llamaremos $\bar{r} = (r_1,\ldots,r_m)$ y $\bar{\beta} = (\beta_1,\ldots,\beta_m)$. Definimos el *índice* $\theta = \mathrm{ind}_{\bar{r},\bar{\beta}}F$ como el menor número $\theta = i_1/r_1 + \cdots + i_m/r_m$ tal que $F^{(i)}(\beta_1,\ldots,\beta_m) \neq 0$.

Obviamente el índice es ≥ 0 y es igual a 0 si y sólo si $F(\beta_1,\ldots,\beta_m) \neq 0$. Además cumple las propiedades siguientes:

Ind 1 *Si* F *y* G *son dos polinomios no nulos, entonces*

$$\mathrm{ind}(F+G) \ge \min\{\mathrm{ind}F, \mathrm{ind}\,G\}, \qquad \mathrm{ind}FG = \mathrm{ind}\,F\,\mathrm{ind}\,G.$$

La fórmula para la suma es obvia. Veamos la del producto. De entre todas las m-tuplas $i = (i_1,\ldots,i_m)$ tales que $\mathrm{ind}\,F = i_1/r_1 + \cdots + i_m/r_m$, consideremos las que tienen el menor valor posible de i_1; de entre ellas, las que tienen el menor valor posible de i_2; y así sucesivamente, hasta quedarnos con una m-tupla mínima i. Similarmente, tomamos una m-tupla mínima j de manera que $\mathrm{ind}\,G = j_1/r_1 + \cdots + j_m/r_m$. Entonces

$$(FG)^{(i)+(j)}(\beta_1,\ldots,\beta_m) = \sum_{u+v=i+j} F^{(u)}G^{(v)}(\beta_1,\ldots,\beta_m),$$

pero todos los sumandos son nulos excepto el correspondiente a $u = i$, $v = j$. En efecto, si $u + v = i + j$, entonces

$$\frac{u_1}{r_1} + \cdots + \frac{u_m}{r_m} + \frac{v_1}{r_1} + \cdots + \frac{v_m}{r_m} = \operatorname{ind} F + \operatorname{ind} G.$$

Si $u_1/r_1 + \cdots + u_m/r_m < \operatorname{ind} F$, entonces $F^{(u)} = 0$, y si, por el contrario, $u_1/r_1 + \cdots + u_m/r_m > \operatorname{ind} F$, entonces ha de ser $v_1/r_1 + \cdots + v_m/r_m < \operatorname{ind} G$, luego $G^{(v)} = 0$. Por consiguiente, para que un sumando sea no nulo ha de cumplir

$$\frac{u_1}{r_1} + \cdots + \frac{u_m}{r_m} = \operatorname{ind} F, \qquad \frac{v_1}{r_1} + \cdots + \frac{v_m}{r_m} = \operatorname{ind} G.$$

Si fuera $u \neq i$, consideremos el menor índice s tal que $u_s \neq i_s$. Por la minimalidad de i ha de ser $u_s > i_s$. Entonces $v_k = j_k$ para $k < s$ y $v_s < j_s$, lo que contradice la minimalidad de j. Concluimos que $u = i$ e, igualmente, $v = j$.

De aquí se sigue que $FG^{(i)+(j)} \neq 0$. Por otra parte, si

$$\frac{u_1}{r_1} + \cdots + \frac{u_m}{r_m} + \frac{v_1}{r_1} + \cdots + \frac{v_m}{r_m} < \operatorname{ind} F + \operatorname{ind} G,$$

entonces

$$\frac{u_1}{r_1} + \cdots + \frac{u_m}{r_m} < \operatorname{ind} F, \qquad \text{o bien} \qquad \frac{v_1}{r_1} + \cdots + \frac{v_m}{r_m} < \operatorname{ind} G,$$

luego $F^{(u)}G^{(v)} = 0$, lo que nos permite concluir que el índice de FG se alcanza con la m-tupla $(i) + (j)$. ∎

Ind 2 *Si*

$$\frac{\partial^{j_1+\cdots+j_m} F}{\partial X_1^{j_1} \cdots \partial X_m^{j_m}} \neq 0,$$

entonces

$$\operatorname{ind} \frac{\partial^{j_1+\cdots+j_m} F}{\partial X_1^{j_1} \cdots \partial X_m^{j_m}} \geq \operatorname{ind} F - \frac{j_1}{r_1} - \cdots - \frac{j_m}{r_m}.$$

Esto se comprueba sin dificultad, al igual que la propiedad siguiente:

Ind 3 *Si l es un natural no nulo, entonces*

$$\operatorname{ind}_{l\bar{r},\bar{\beta}} F = \frac{1}{l} \operatorname{ind}_{\bar{r},\bar{\beta}} F.$$

He aquí el último resultado que necesitamos para demostrar el teorema 9.9:

Teorema 9.15 *En las condiciones del teorema* 9.13, *tomemos* $0 \leq \delta \leq 1$ *y supongamos que* $\delta r_1 \geq r_2, \ldots, \delta r_{m-1} \geq r_m, \delta r_m \geq 10$. *Sean* $\beta_1, \ldots, \beta_m \in K$ *y sea* $\theta = \operatorname{ind}_{\bar{r},\bar{\beta}} G$. *Entonces*

$$\min\{\theta, \theta^2\} \leq \frac{4}{l} \operatorname{ind}_{\bar{r},\bar{\beta}} F + 4\delta.$$

DEMOSTRACIÓN: Consideremos un operador diferencial

$$D_i = \frac{1}{i_1! \cdots i_{m-1}!} \frac{\partial^{i_1+\cdots+i_{m-1}}}{\partial X_1^{i_1} \cdots \partial X_{m-1}^{i_{m-1}}}$$

de grado $\leq l-1$. Si el polinomio

$$D_i \frac{1}{j!} \frac{\partial^j G}{\partial X_m^j}$$

no es idénticamente nulo, su índice es mayor o igual que

$$\theta - \frac{i_1}{r_1} - \cdots - \frac{i_{m-1}}{r_{m-1}} - \frac{j}{r_m} \geq \theta - \frac{r_m}{r_{m-1}} - \frac{j}{r_m},$$

donde hemos usado que $r_i \geq r_{m-1}$ y que $i_1 + \cdots + i_{m-1} \leq l-1 \leq r_m$. Como $r_m/r_{m-1} \leq \delta$ y el índice no puede ser negativo, debe ser mayor o igual que

$$\text{máx}\{0, \theta - j/r_m\} - \delta.$$

El determinante que define a F es una suma de $l!$ términos, cada uno de los cuales es un producto de l polinomios como el que acabamos de considerar. El índice de cada uno de estos términos (supuesto que no sea idénticamente nulo) será la suma de los índices de los factores, luego mayor o igual que

$$\sigma = \sum_{j=0}^{l-1} \text{máx}\{0, \theta - j/r_m\} - l\delta.$$

Por la propiedad del índice de una suma, concluimos que $\text{ind}\, F \geq \sigma$.

Podemos suponer que $\theta r_m > 10$, pues en caso contrario $\theta \leq 10/r_m \leq \delta$ y la desigualdad que queremos probar se cumple trivialmente. Distingamos dos casos:

Si $\theta r_m < l$ entonces (representando por $[\ ]$ la parte entera)

$$\sum_{j=0}^{l-1} \text{máx}\{0, \theta - j/r_m\} = \frac{1}{r_m} \sum_{j=0}^{l-1} \text{máx}\{0, \theta r_m - j\} \geq \frac{1}{r_m} \sum_{j=0}^{[\theta r_m]} ([\theta r_m] - j)$$

$$= \frac{[\theta r_m]([\theta r_m]+1)}{2r_m} \geq \frac{[\theta r_m]^2}{2r_m}.$$

Es fácil ver que si $a > 10$ entonces $(a-1)^2/2 > a^2/3$. Aplicamos esto a $\theta r_m > 10$, con lo que

$$\frac{[\theta r_m]^2}{2r_m} \geq \frac{(\theta r_m - 1)^2}{2r_m} > \frac{\theta^2 r_m}{3}.$$

En resumen, $\text{ind} F \geq \theta^2 r_m/3 - l\delta$, luego

$$\theta^2 \leq \frac{3l}{r_m} \frac{1}{l} \text{ind}\, F + \frac{3l}{r_m}\delta \leq \frac{4}{l} \text{ind}\, F + 4\delta,$$

pues $r_m \geq 10$ y $l \leq r_m + 1$.

Supongamos ahora que $\theta r_m \geq l$. Entonces

$$\sum_{j=0}^{l-1} \max\{0, \theta - j/r_m\} = \sum_{j=0}^{l-1} (\theta - j/r_m) = l\theta - \frac{l(l-1)}{2r_m} \geq \frac{l\theta}{2}.$$

Por lo tanto $\operatorname{ind} F \geq l\theta/2$ y

$$\theta \leq \frac{2}{l} \operatorname{ind} F \leq \frac{4}{l} \operatorname{ind} F + 4\delta.$$

■

El teorema siguiente es una versión ligeramente más general de 9.9 donde no suponemos que el polinomio G tenga coeficientes enteros (necesitamos suprimir esta hipótesis para que funcione la inducción en que se basa la prueba). Observemos que si G tiene coeficientes enteros y $|G|_\infty \leq A_1^{\delta r_1}$, entonces

$$A_K(G) = \prod_{\mathfrak{p}} |G|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \leq \prod_{\mathfrak{p}|\infty} |G|_{\mathfrak{p}}^{N_{\mathfrak{p}}} \leq \prod_{\mathfrak{p}} |G|_\infty^{N_{\mathfrak{p}}} = |G|_\infty^N \leq A_1^{N\delta r_1}.$$

Esto justifica el cambio de la hipótesis (9.4) por (9.15).

Teorema 9.16 *Sea K un cuerpo numérico de grado N, sea $m > 0$ un número natural, sea $0 < \delta < 1/16^{m+N}$, sean $r_1, \ldots, r_m$ naturales no nulos tales que*

$$\delta r_1 > r_2, \ \delta r_2 > r_3, \ldots, \delta r_m > 10. \tag{9.12}$$

Sea $G(X_1, \ldots, X_m) \in K[X_1, \ldots, X_m]$ no nulo tal que $\operatorname{grad}_{X_j} G \leq r_j$. Consideremos $\beta_1, \ldots, \beta_m \in K$ números de altura $A_1, \ldots, A_m$ respectivamente y supongamos:

$$N \log 4 + 4m \leq \delta \log A_1, \tag{9.13}$$

$$r_1 \log A_1 \leq r_j \log A_j, \qquad j = 1, \ldots, m, \tag{9.14}$$

$$A_K(G) \leq A_1^{N\delta r_1}. \tag{9.15}$$

Entonces $\operatorname{ind}_{\bar{r},\bar{\beta}} G \leq 20^m \delta^{(1/2)^m}$.

Demostración: Razonamos por inducción sobre m. Supongamos en primer lugar $m = 1$. Observemos que para un polinomio de una variable, el índice es $\operatorname{ind} G = e/r_1$, donde $e \geq 0$ es la multiplicidad de β_1 en $G(X_1)$. Por el teorema 7.25, teniendo en cuenta que las alturas son ≥ 1, vemos que

$$A_1^e = A_K(\beta_1)^e \leq 2^{r_1} A_K(G) \leq 4^{r_1 N} A_K(G) \leq 4^{r_1 N} A_1^{N\delta r_1},$$

donde hemos usado (9.15). Usando (9.13) obtenemos

$$e \log A_1 \leq r_1 N \log 4 + N\delta r_1 \log A_1 \leq \delta r_1 \log A_1 + N\delta r_1 \log A_1,$$

luego $e \leq (N+1)\delta r_1$ y, por consiguiente,

$$\operatorname{ind} G = e/r_1 \leq (N+1)\delta \leq \delta^{1/2} \leq 20\delta^{1/2},$$

donde hemos usado que $\delta < 1/16^{N+1}$, luego

$$(N+1)\delta^{1/2} < (N+1)/4^{N+1} < 1.$$

Supongamos ahora que el teorema es cierto para polinomios de grado menor que $m > 1$. Aplicamos a G el teorema 9.13, lo que nos da un polinomio $F = UV$. Si $\mathfrak{p}$ es un primo arquimediano de K, el teorema 9.14 nos da que

$$|F|_{\mathfrak{p}} \le 2^{mr_1 l} \cdot 2^{r_1 l} \cdot 2^{mr_1 l} |G|_{\mathfrak{p}}^l \le 8^{mr_1 l} |G|_{\mathfrak{p}}^l,$$

donde hemos acotado $r_j + 1 \le 2^{r_j} \le 2^{r_1}$ y $l! \le l^l \le 2^{r_1 l}$. Para los primos no arquimedianos es simplemente $|F|_{\mathfrak{p}} \le |G|_{\mathfrak{p}}^l$. Elevamos al grado local $N_{\mathfrak{p}}$ y multiplicamos para todos los primos, con lo que resulta

$$A_K(F) \le 8^{Nmr_1 l} A_K(G)^l \le A_1^{2\delta N r_1 l},$$

donde hemos usado (9.13), que nos da $8^m \le e^{4m} \le A_1^{\delta}$, así como (9.15).

Como $F = UV$ y las variables están separadas, cada coeficiente de F es el producto de un coeficiente de U por un coeficiente de V. De aquí se sigue inmediatamente que $A_K(F) = A_K(U)A_K(V)$ y, como las alturas son ≥ 1, tenemos que $A_K(U)$ y $A_K(V)$ están ambas acotadas por $A_1^{2\delta N r_1 l}$.

Vamos a aplicar la hipótesis de inducción a los polinomios U y V con 2δ en lugar de δ. Veamos que se cumplen todas las hipótesis:

Ciertamente $0 < 2\delta < 1/16^{m-1+N}$ (para U) y $0 < 2\delta < 1/16^N + 1$ (para V).

Los grados parciales de U y V están acotados por lr_j, y estos números cumplen obviamente (9.12) con 2δ en vez de δ. Las hipótesis (9.13) y (9.14) se cumplen claramente y ya hemos comprobado (9.15).

Así pues, la hipótesis de inducción nos da que

$$\operatorname{ind}_{l\bar{r},\bar{\beta}} U \le 20^{m-1}(2\delta)^{(1/2)^{m-1}}, \qquad \operatorname{ind}_{lr_m,\beta_m} V \le 20(2\delta)^{1/2}.$$

Al multiplicar por 20 o sacar raíces cuadradas sucesivas a 2δ obtenemos números mayores, luego $\operatorname{ind} V \le \operatorname{ind} U$. Aunque no es exactamente la propiedad del índice de un producto que hemos demostrado anteriormente (porque las variables de los factores no son las mismas), es fácil ver que igualmente se cumple

$$\operatorname{ind}_{l\bar{r},\bar{\beta}} F = \operatorname{ind}_{l\bar{r},\bar{\beta}} U + \operatorname{ind}_{l\bar{r},\bar{\beta}} V$$
$$\le 2 \cdot 20^{m-1}(2\delta)^{(1/2)^{m-1}} \le 2\sqrt{2} \cdot 20^{m-1}\delta^{(1/2)^{m-1}}.$$

Por consiguiente,

$$\operatorname{ind}_{\bar{r},\bar{\beta}} F \le l \cdot 2\sqrt{2} \cdot 20^{m-1}\delta^{(1/2)^{m-1}}.$$

Ahora aplicamos el teorema 9.15, según el cual $\theta = \operatorname{ind} G$ cumple

$$\min\{\theta, \theta^2\} \le 8\sqrt{2} \cdot 20^{m-1}\delta^{(1/2)^{m-1}} + 4\delta$$
$$\le (8\sqrt{2}+4)20^{m-1}\delta^{(1/2)^{m-1}} \le 20^m \delta^{(1/2)^{m-1}},$$

donde hemos usado que $4\delta \le 4\delta^{(1/2)^{m-1}} \le 4 \cdot 20^{m-1}\delta^{(1/2)^{m-1}}$.

Por último, si $\theta \le \theta^2$ hacemos

$$\theta \le 20^{m-1}\delta^{(1/2)^m} \le 20^m \delta^{(1/2)^m},$$

mientras que si $\theta^2 \le \theta$ concluimos con

$$\theta \le 20^{m/2}\,\delta^{(1/2)^m} \le 20^m \delta^{(1/2)^m}.$$

■

Con esto queda demostrado el teorema de Roth.

9.5 El teorema de Siegel

Como primera aplicación del teorema de Roth vamos a demostrar un teorema de Siegel con el que probaremos que una curva elíptica tiene a lo sumo un número finito de puntos con coordenadas enteras.

Observemos que si K es un cuerpo numérico, $\mathfrak{p}$ es uno de sus divisores primos, α es un número algebraico y $\epsilon > 0$, el teorema de Roth nos permite encontrar una constante C tal que para todo $r \in K$ se cumpla

$$|\alpha - r|_{\mathfrak{p}} \ge \frac{C}{A_K(r)^{2+\epsilon}}.$$

Por consiguiente

$$\frac{\log|\alpha - r|_{\mathfrak{p}}}{\log A_K(r)} \ge \frac{\log C}{\log A_K(r)} - (2+\epsilon).$$

Supongamos que podemos aproximarnos a α desde puntos $r \in K$ con la topología $\mathfrak{p}$-ádica (es decir, que $\alpha \in K_{\mathfrak{p}}$). Entonces $A_K(r)$ tiende a infinito, por lo que

$$\liminf_{r\to\alpha} \frac{\log|\alpha - r|_{\mathfrak{p}}}{\log A_K(r)} \ge -(2+\epsilon)$$

y, como esto vale para todo $\epsilon > 0$, concluimos que

$$\liminf_{r\to\alpha} \frac{\log|\alpha - r|_{\mathfrak{p}}}{\log A_K(r)} \ge -2.$$

El primer paso para demostrar el teorema de Siegel será obtener un resultado similar pero referente a puntos de una curva elíptica en lugar de a números. Si K es un cuerpo numérico, $\mathfrak{p}$ es un divisor primo (arquimediano o no), $K_{\mathfrak{p}}$ es la compleción y $\mathbb{K}_{\mathfrak{p}}$ es una clausura algebraica de $K_{\mathfrak{p}}$, cada curva elíptica E/K puede verse también como curva $E/K_{\mathfrak{p}}$, lo que nos permite considerar la topología métrica en $E(\mathbb{K}_{\mathfrak{p}})$ (entendiendo como tal a la topología compleja si $\mathfrak{p}$ es arquimediano).

Definición 9.17 Sea K un cuerpo numérico, $\mathfrak{p}$ uno de sus divisores primos, E/K una curva elíptica, $Q \in E(K_\mathfrak{p})$ y $t \in K_\mathfrak{p}(E)$ una función racional con un cero de orden $e \geq 1$ en Q. Para cada $P \in E(K_\mathfrak{p})$ definimos

$$d_\mathfrak{p}(P,Q) = \min\{|t(P)|_\mathfrak{p}^{1/e}, 1\},$$

entendiendo que si t_Q tiene un polo en P, entonces $d_\mathfrak{p}(P,Q) = 1$.

Observemos que $d_\mathfrak{p}$ depende de la elección de t. No es realmente una distancia en $E(K_\mathfrak{p})$, pero es claro que si $P \to Q$ en la topología métrica entonces $d_\mathfrak{p}(P,Q) \to 0$, y esto será suficiente. El teorema siguiente implica que la elección de t_Q no será relevante en los resultados que vamos a obtener:

Teorema 9.18 *Sea K un cuerpo numérico, sea $\mathfrak{p}$ un divisor primo en K, sea E/K una curva elíptica, sea $Q \in E(K_\mathfrak{p})$ y sean t, $t' \in E(K_\mathfrak{p})$ dos funciones no idénticamente nulas con un cero en Q. Llamemos $d_\mathfrak{p}(\ \ ,Q)$ y $d'_\mathfrak{p}(\ \ ,Q)$ a las funciones dadas por la definición precedente. Si Q es un punto de acumulación de $E(K_\mathfrak{p})$, entonces*

$$\lim_{P\to Q} \frac{\log d'_\mathfrak{p}(P,Q)}{\log d_\mathfrak{p}(P,Q)} = 1.$$

Demostración: Sean e y e' respectivamente el orden de t y t' en Q. Observemos que la continuidad de t y t' (junto con el hecho de que tienen un número finito de ceros) implica que si $P \neq Q$ varía en un entorno de Q suficientemente pequeño entonces

$$d_\mathfrak{p}(P,Q) = |t(P)|_\mathfrak{p}^{1/e} > 0, \qquad d_\mathfrak{p}(P,Q) = |t'(P)|_\mathfrak{p}^{1/e} > 0,$$

luego podemos tomar logaritmos. La función $f = t'^e/t^{e'} \in K_\mathfrak{p}(E)$ no tiene ni un cero ni un polo en Q. Además

$$\log|f(P)|_\mathfrak{p} = e\log|t'(P)|_\mathfrak{p} - e'\log|t(P)|_\mathfrak{p},$$

luego

$$\frac{\log d'_\mathfrak{p}(P,Q)}{\log d_\mathfrak{p}(P,Q)} = \frac{e\log|t'(P)|_\mathfrak{p}}{e'\log|t(P)|_\mathfrak{p}} = 1 + \frac{\log|f(P)|_\mathfrak{p}}{e'\log|t(P)|_\mathfrak{p}}.$$

La función $|f|_\mathfrak{p}$ es continua y no se anula en un entorno compacto de P, luego toma valores en un intervalo $[\epsilon, M]$ y, por consiguiente, el logaritmo está acotado. Así, en el último término de la igualdad anterior, el numerador está acotado y el denominador tiende a infinito cuando $P \to Q$, lo que prueba el teorema. ∎

Necesitamos una última propiedad básica de las distancias:

Teorema 9.19 *Sea K un cuerpo numérico, sea $\mathfrak{p}$ un divisor primo de K, sea $\phi : E_1 \longrightarrow E_2$ una aplicación regular no constante definida sobre K entre dos curvas elípticas sobre K. Sea $Q \in E_1(K_\mathfrak{p})$ un punto de acumulación. Entonces*

$$\lim_{P\to Q} \frac{\log d_\mathfrak{p}(\phi(P),\phi(Q))}{\log d_\mathfrak{p}(P,Q)} = e_\phi(Q),$$

donde $e_\phi(Q)$ es el índice de ramificación de ϕ en Q.

DEMOSTRACIÓN: Sean t y t' parámetros locales de E_1 y E_2 en Q y $\phi(Q)$ respectivamente. Vamos a usar estas funciones para calcular $d_\mathfrak{p}$. El teorema anterior implica que el resultado será cierto para cualquier otra elección.

Entonces $\phi \circ t' = \epsilon t^{e_\phi(Q)}$, donde $\epsilon \in K_\mathfrak{p}(E_2)$ no tiene ni un cero ni un polo en Q. Así pues, para puntos P suficientemente cercanos a Q,

$$\frac{\log d_\mathfrak{p}(\phi(P),\phi(Q))}{\log d_\mathfrak{p}(P,Q)} = \frac{\log |t'(\phi(P))|_\mathfrak{p}}{\log |t(P)|_\mathfrak{p}} = \frac{\log |\epsilon(P)|_\mathfrak{p} + e_Q(P)\log |t(P)|_\mathfrak{p}}{\log |t(P)|_\mathfrak{p}}$$

La función $|\epsilon|_\mathfrak{p}$ es continua y no nula en un entorno compacto de Q, luego $\log |\epsilon(P)|_\mathfrak{p}$ está acotado en dicho entorno, mientras que $\log |t(P)|_\mathfrak{p}$ tiende a infinito cuando $P \to Q$. La conclusión es ahora inmediata. ∎

Pasemos ahora a reformular el teorema de Roth en términos de distancias sobre curvas:

Teorema 9.20 *Sea K un cuerpo numérico, $\mathfrak{p}$ un divisor primo en K, sea E/K una curva proyectiva definida sobre K, sea $Q \in E(\mathbb{A})$ un punto de acumulación de $E(K)$ respecto de la topología $\mathfrak{p}$-ádica y $f \in K(E)$ una función racional no constante. Entonces*

$$\liminf_{P\to Q} \frac{\log d_\mathfrak{p}(P,Q)}{\log A_K(f(P))} \geq -2,$$

donde P varía en $E(K)$.

DEMOSTRACIÓN: Notemos que para que Q pueda ser un punto de acumulación de $E(K)$ es necesario que $Q \in E(K_\mathfrak{p})$ (pues $E(K_\mathfrak{p}) \cap E(\mathbb{A})$ es cerrado en $E(\mathbb{A})$, viendo ambos conjuntos como subespacios de $E(\mathbb{K}_\mathfrak{p})$). Por lo tanto la distancia está definida. También es claro que no importa con qué función $t \in K_\mathfrak{p}(E)$ la calculamos.

Aplicando el teorema 7.24 a la aplicación $1/x$ (extendida de forma natural a una aplicación regular $\mathrm{P}^1(K) \longrightarrow \mathrm{P}^1(K)$) vemos que el límite inferior del enunciado no se altera si cambiamos f por $1/f$, luego podemos suponer que f no tiene un polo en Q. Esto nos permite tomar $t = f - f(Q)$ para calcular $d_\mathfrak{p}$. Llamemos e al orden de t en Q. Entonces, para todo $\epsilon > 0$,

$$\liminf_{P\to Q} \frac{\log d_\mathfrak{p}(P,Q)}{\log A_K(f(P))} = \liminf_{P\to Q} \frac{\log |f(P)-f(Q)|_\mathfrak{p}}{e\log A_K(f(P))}$$

$$= \frac{1}{e}\liminf_{P\to Q}\left(\frac{\log\big(A_K(f(P))^{2+\epsilon}|f(P)-f(Q)|_\mathfrak{p}\big)}{\log A_K(f(P))} - (2+\epsilon)\right).$$

El teorema de Roth aplicado a $r = f(P) \in K$ y $\alpha = f(Q) \in \mathbb{A}$ vemos que para todo $P \in E(K)$ salvo a lo sumo un número finito de excepciones se cumple

$$A_K(f(P))^{2+\epsilon}|f(P)-f(Q)|_\mathfrak{p} \geq 1.$$

(Aquí usamos además que f no toma nunca el mismo valor sobre infinitos puntos.) Por otra parte, como el conjunto de puntos cuya altura no supera cualquier

cota prefijada es finito, podemos encontrar entornos de Q en los que $A_K(f(P))$ sea arbitrariamente grande, luego

$$\liminf_{P\to Q} \frac{\log d_{\mathfrak{p}}(P,Q)}{\log A_K(f(P))} \geq -\frac{2+\epsilon}{e}.$$

Como $\epsilon > 0$ es arbitrario y $e \geq 1$, la conclusión es obvia. ∎

Con esto ya podemos probar el teorema de Siegel. Se trata de un refinamiento sustancial del teorema anterior que se obtiene combinándolo con el teorema (débil) de Mordell-Weil. Recordemos que sobre una curva elíptica hemos definido

$$a_f(P) = \log A(P) = \frac{\log A_K(P)}{|K:\mathbb{Q}|}.$$

Teorema 9.21 (Siegel) *Sea K un cuerpo numérico, sea $\mathfrak{p}$ un divisor primo sobre K, sea E/K una curva elíptica, sea $f \in K(E)$ una función par no constante y $Q \in E(\mathbb{A})$ un punto de acumulación de $E(K)$. Entonces*

$$\lim_{P\to Q} \frac{\log d_{\mathfrak{p}}(P,Q)}{a_f(P)} = 0,$$

donde P varía en $E(K)$.

Demostración: Tomemos una sucesión de puntos distintos $P_i \in E(K)$ que converja a Q y tal que exista

$$\lim_i \frac{\log d_{\mathfrak{p}}(P_i,Q)}{a_f(P_i)} = L.$$

Basta probar que $L = 0$. Ciertamente $L \leq 0$, luego sólo hemos de probar que $L \geq 0$.

Sea $m \geq 2$ un número natural. El teorema débil de Mordell-Weil nos da que el grupo $E(K)/mE(K)$ es finitamente generado, luego tomando una subsucesión podemos suponer que todos los puntos P_i son de la forma $P_i = mP_i' + R$, para ciertos puntos $P_i' \in E(K)$ y un $R \in E(K)$ fijo. Entonces, usando las propiedades a) y b) de a_f demostradas en la prueba del teorema de Mordell-Weil obtenemos que

$$m^2 a_f(P_i') = a_f(mP_i') + O(1) = a_f(P_i - R) + O(1) \leq 2a_f(P_i) + O(1),$$

donde la función $O(1)$ es independiente de i.

Por otra parte, como las operaciones de E son continuas para la topología $\mathfrak{p}$-ádica (porque son aplicaciones regulares), tenemos que $mP_i' \to Q - R$. Como $E(K_{\mathfrak{p}})$ es compacto, la sucesión P_i' ha de tener una subsucesión convergente a un punto $Q' \in E(K_{\mathfrak{p}})$, pero dicho punto cumplirá $mQ' = Q - R$, lo que implica que $Q' \in E(\mathbb{A})$. En efecto, la multiplicación por m es suprayectiva en $E(\mathbb{A})$ (como toda isogenia no constante), luego existe un punto $Q' \in E(\mathbb{A})$ que cumple $mQ' = Q - R$ y, concretamente, hay m^2 puntos en estas condiciones.

Lo mismo es válido en $E(\mathbb{K}_\mathfrak{p})$, luego los m^2 puntos $Q' \in E(\mathbb{K}_\mathfrak{p})$ que cumplen $mQ' = Q - R$ son precisamente los de $E(\mathbb{A})$. Así pues, tomando una subsucesión podemos suponer que $P_i' \to Q' \in E(\mathbb{A})$.

El teorema 2.35 implica que la multiplicación por m es no ramificada, al igual que la traslación por R (porque es un isomorfismo). Por consiguiente, el teorema 9.19 nos da que

$$\lim_i \frac{\log d_\mathfrak{p}(P_i, Q)}{\log d_\mathfrak{p}(P_i', Q')} = 1.$$

Combinando esto con la desigualdad que hemos obtenido para las alturas vemos que

$$L = \lim_i \frac{\log d_\mathfrak{p}(P_i, Q)}{a_f(P_i)} \geq \liminf_i \frac{\log d_\mathfrak{p}(P_i', Q')}{\frac{m^2}{2} a_f(P_i') + O(1)}.$$

(Notemos que $\log d_\mathfrak{p} \leq 0$, lo cual invierte la desigualdad.) El teorema anterior nos da que

$$\liminf_i \frac{\log d_\mathfrak{p}(P_i', Q')}{|K : \mathbb{Q}| a_f(P_i')} \geq -2,$$

de donde concluimos que

$$L \geq -\frac{4|K : \mathbb{Q}|}{m^2}.$$

Como m es arbitrario, ha de ser $L \geq 0$. ∎

Como ya hemos comentado, de este teorema se desprende que las curvas elípticas sobre cuerpos numéricos no pueden tener infinitos puntos con coordenadas enteras. Empezamos demostrando un caso particular:

Teorema 9.22 *Sea K un cuerpo numérico, sea S un conjunto finito de divisores primos de K que contenga al menos a todos los primos arquimedianos y sea*

$$K_S = \{\alpha \in K \mid |\alpha|_\mathfrak{p} \leq 1 \text{ para todo } \mathfrak{p} \in S\}.$$

Sea E/K una curva elíptica y $x \in K(E)$ la primera coordenada respecto de una ecuación de Weierstrass. Entonces el conjunto

$$\{P \in E(K) \mid x(P) \in K_S\}$$

es finito.

Demostración: Aplicamos el teorema anterior con $f = x$. Supongamos que hubiera infinitos puntos $P_i \in E(K)$ tales que $x(P_i) \in K_S$. Entonces

$$a_x(P_i) = \frac{1}{|K : \mathbb{Q}|} \sum_{\mathfrak{p} \in S} \log \max\{1, |x(P_i)|_\mathfrak{p}^{n_\mathfrak{p}}\}. \tag{9.16}$$

Observemos que el conjunto

$$\{\alpha \in K \mid |\alpha|_\mathfrak{p} \leq 1 \text{ para todo } \mathfrak{p}\}$$

es finito (sus elementos son enteros algebraicos con su representación geométrica acotada). Tomando una subsucesión, podemos suponer que el mayor de los sumandos de (9.16) se alcanza siempre en un mismo primo $\mathfrak{p} \in S$ (y es mayor que 1 por la observación precedente), con lo que

$$a_x(P_i) \le |S| \log |x(P_i)|_{\mathfrak{p}} \qquad \text{para todo } i.$$

En particular $|x(P_i)|_{\mathfrak{p}} \to \infty$ y, como x tiene un único polo en O, esto implica que $P_i \to O$. (Dado un entorno U de O, la función $|x|_{\mathfrak{p}}$ está acotada en su complementario, luego la sucesión P_i está finalmente en U. Podemos calcular distancias a O mediante la función $t = 1/x$ con $e = 2$, Así,

$$d_{\mathfrak{p}}(P_i, O) = \min\{1, |x(P_i)|_{\mathfrak{p}}^{-1/2}\}.$$

Si i es suficientemente grande,

$$-\frac{\log d_{\mathfrak{p}}(P_i, O)}{a_x(P_i)} = \frac{\log |x(P_i)|_{\mathfrak{p}}}{2a_x(P_i)} \ge \frac{1}{2|S|}.$$

Esto contradice al teorema anterior. ∎

Más en general:

Teorema 9.23 *Sea K un cuerpo numérico, sea S un conjunto finito de divisores primos de K que contenga al menos a todos los primos arquimedianos y sea*

$$K_S = \{\alpha \in K \mid |\alpha|_{\mathfrak{p}} \le 1 \text{ para todo } \mathfrak{p} \in S\}.$$

Sea E/K una curva proyectiva regular de género 1 y $f \in K(E)$ una función racional no constante. Entonces el conjunto

$$\{P \in E \mid f(P) \in K_S\}$$

es finito.

Demostración: Ciertamente f tiene un polo $O \in E(\mathbb{A})$. No perdemos generalidad si extendemos K para garantizar que $O \in K$. Entonces podemos considerar a E/K como una curva elíptica con neutro O. Sean $x, y \in K(E)$ dos funciones que satisfagan una ecuación de Weierstrass

$$y^2 = x^3 + Ax + B.$$

Como $|K(x, y) : K(x)| = 2$, podemos expresar

$$f = \frac{F(x) + G(x)y}{H(x)},$$

con polinomios $F(X)$, $G(X)$, $H(X) \in K[X]$. Como $v_O(x) = -2$, $v_O(y) = -3$, $v_O(f) < 0$, podemos concluir que

$$2 \operatorname{grad} H < \max\{2 \operatorname{grad} F, 2 \operatorname{grad} G + 3\}. \tag{9.17}$$

Por otra parte,

$$(fH(x) - F(x))^2 = (G(x)y)^2 = G(x)^2(x^3 + Ax + B).$$

Tenemos así que x es raíz de un polinomio con coeficientes en $K[f]$. La mayor potencia de x estará en los términos

$$f^2H(x)^2, \quad F(x)^2, \quad fH(x)F(x), \quad G(x)^2x^3.$$

El grado del tercero será menor que el de uno de los dos primeros salvo si $\operatorname{grad} H = \operatorname{grad} F$, pero en tal caso (9.17) implica que los tres tienen grado menor que el cuarto. Así pues, podemos descartar $fH(x)F(x)$ supuesto que justifiquemos que los términos de mayor grado de los otros tres no se cancelan.

Igualmente, (9.17) implica que el primero tiene grado menor que el segundo o el cuarto, luego podemos descartarlo si justificamos que los términos de mayor grado de $F(x)^2$ y $G(x)^2x^3$ no se cancelan, pero no pueden cancelarse, pues uno tiene grado par y otro impar.

En resumen, la mayor potencia de x aparece únicamente en $F(x)^2$ o en $G(x)^2x^3$. En particular el término de mayor grado tiene su coeficiente en K (es decir, no contiene a f). Dicho de otro modo, hemos probado que x es la raíz de un polinomio mónico con coeficientes en $K[f]$. Pongamos que es

$$a_0X^N + a_1(f)X^{N-1} + \cdots + a_{N-1}(f)X + a_N(f) = 0,$$

donde podemos exigir que $a_0 \in K_S$ y $a_i(f) \in K_S[f]$ (sin más que multiplicar por un elemento adecuado de K). Basta probar el teorema para un conjunto de primos mayor que S, luego podemos añadir a S los primos necesarios para que a_0 sea una unidad de K_S. Así podemos dividir entre a_0 y suponer que $a_0 = 1$.

Ahora, si $P \in E(K)$ cumple $f(P) \in K_S$, tenemos que $x(P)$ es la raíz de un polinomio mónico con coeficientes en K_S, pero K_S es íntegramente cerrado (teorema 7.12), luego concluimos que $x(P) \in K_S$. Es decir,

$$\{P \in E(K) \mid f(P) \in K_S\} \subset \{P \in E(K) \mid x(P) \in K_S\},$$

y el último conjunto es finito por el teorema anterior. ∎

Observemos que si S es el conjunto de los primos arquimedianos de K entonces K_S es su anillo de enteros. Más en particular, todo sistema de ecuaciones con coeficientes racionales (con cualquier número de variables) que defina una curva proyectiva de género 1 tiene a lo sumo un número finito de soluciones enteras.

Las técnicas que hemos empleado no son efectivas, en el sentido de que no nos permiten acotar las coordenadas de las posibles soluciones enteras de una ecuación dada, con lo que no podemos calcular explícitamente todas las soluciones (ni determinar si existe alguna). En 1966, Baker desarrolló una teoría de formas lineales en logaritmos que permite obtener cotas explícitas, aunque éstas son muy elevadas. Por ejemplo, para una ecuación

$$Y^2 = AX^3 + BX^2 + CX + D$$

con A, B, C, $D \in \mathbb{Z}$ y $M = \max\{|A|, |B|, |C|, |D|\}$, las técnicas de Baker permiten probar que toda solución entera ha de cumplir

$$\max\{|x|, |y|\} < \exp((10^6 M)^{10^6}).$$

Ejemplo Vamos a ver que hay curvas elípticas de la forma $Y^2 = X^3 + m$ con una cantidad arbitrariamente grande de soluciones enteras.

En la página 177 hemos visto que la curva $Y^2 = X^3 + 8$ contiene puntos racionales de orden infinito. En particular contiene infinitas soluciones racionales. Tomemos n de ellas $(p_i/r, q_i/r)$, para $i = 1, \ldots, n$. Entonces

$$\frac{q_i^2}{r^2} = \frac{p_i^3}{r^3} + 8,$$

luego

$$(r^2 q_i)^2 = (r p_i)^3 + 8r^6.$$

Esto significa que la ecuación $Y^2 = X^3 + 8r^6$ tiene al menos las n soluciones enteras distintas dadas por $(rp_i, r^2 q_i)$, para $i = 1, \ldots, n$. ∎

Como aplicación del teorema de Siegel podemos probar un resultado con numerosas aplicaciones a la teoría de las curvas elípticas.

Teorema 9.24 (Shafarevich) *Sea K un cuerpo numérico y S un conjunto finito de divisores primos de K que contenga a todos los primos arquimedianos. Entonces, salvo isomorfismo, hay un número finito de curvas elípticas E/K con buena reducción en los primos que no están en S.*

DEMOSTRACIÓN: Es claro que podemos sustituir S por un conjunto mayor. En particular podemos suponer que S contiene a todos los primos que dividen a 2 y 3 así como que el anillo K_S es un dominio de ideales principales (teorema 7.12).

Supongamos que existen infinitas curvas elípticas E_i/K no isomorfas dos a dos y todas con buena reducción fuera de S. A cada una de ellas podemos aplicarle el teorema 7.13, lo que nos da una ecuación

$$Y^2 = X^3 + A_i X + B_i, \qquad A_i,\ B_i \in K_S$$

con discriminante $\Delta_i = -16(4A_i^3 + 27B_i^2)$ no divisible entre ningún primo de S (es decir, Δ_i es una unidad de K_S).

Por el teorema de las unidades de Hasse, el grupo U_S de las unidades de K_S es finitamente generado, luego el cociente U_S/U_S^{12} es finito. Extrayendo una subsucesión de E_i, podemos suponer que todos los discriminantes Δ_i tienen el mismo resto módulo U_S^{12}, es decir, que existe un $C \in U_S$ de modo que

$$\Delta_i = C D_i^{12}, \qquad D_i \in U_S.$$

La relación entre Δ_i, A_i y B_i implica que $(-12A_i/D_i^4, 27B_i/D_i^6)$ es un punto de la curva elíptica $Y^2 = X^3 - 27C$ con coordenadas en K_S. Por el teorema de

Siegel sólo hay un número finito de tales puntos, luego ha de haber al menos dos índices i, j tales que

$$A_i/D_i^4 = A_j/D_j^4, \qquad B_i/D_i^6 = B_j/D_j^6,$$

pero entonces el cambio de variables

$$X = (D_i/D_j)^2 X', \qquad Y = (D_i/D_j)^3 Y'$$

determina un isomorfismo entre E_i y E_j, contradicción. ∎

Uniendo esto al teorema 6.26 obtenemos la consecuencia siguiente:

Teorema 9.25 *Si K es un cuerpo numérico y E/K es una curva elíptica, existen, salvo isomorfismo, un número finito de curvas isógenas a E sobre K.*

Capítulo X

Curvas elípticas complejas

Nos ocupamos ahora de las curvas elípticas definidas sobre el cuerpo $\mathbb{C}$ de los números complejos, en cuyo estudio podemos emplear técnicas analíticas, más concretamente, la teoría de funciones de variable compleja. La idea básica es que las curvas elípticas complejas son superficies de Riemann de género 1, esto es, homeomorfas a un toro. Topológicamente, pueden ser obtenidas identificando los lados opuestos de un cuadrado. Vamos a ver que si en lugar de un cuadrado tomamos un paralelogramo adecuado en $\mathbb{C}$, entonces esta identificación puede hacerse mediante funciones holomorfas. Para precisar esta idea dedicamos la primera sección a estudiar los toros complejos que se obtienen identificando los lados opuestos de un paralelogramo.

10.1 Retículos y toros complejos

Definición 10.1 Un *retículo complejo* es un grupo $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$, donde ω_1 y ω_2 son dos números complejos linealmente independientes sobre $\mathbb{R}$.

Es claro que la independencia lineal de ω_1 y ω_2 equivale a que sean no nulos y $\tau = \omega_2/\omega_1$ tenga parte imaginaria no nula. Eligiendo el orden de ω_1 y ω_2 podemos exigir que $\operatorname{Im}\tau > 0$. Diremos entonces que el par (ω_1, ω_2) es una *base orientada* del retículo R. En lo sucesivo sobrentenderemos siempre que las bases con las que trabajamos están orientadas.

Se llama *paralelogramo fundamental* asociado a una base ω_1, ω_2 de un retículo complejo R al conjunto

$$P = \{\alpha\omega_1 + \beta\omega_2 \mid 0 \leq \alpha,\ \beta < 1\}.$$

Es claro entonces que el plano complejo se descompone en unión disjunta de los trasladados de P, es decir,

$$\mathbb{C} = \bigcup_{\omega \in R} \omega + P.$$

Un *toro complejo* es un cociente $T = \mathbb{C}/R$, donde R es un retículo complejo. Es claro que la proyección natural $p : \mathbb{C} \longrightarrow T$ es inyectiva sobre cada trasladado $z + P$ de un paralelogramo fundamental de R. En particular p es localmente inyectiva. Se comprueba sin dificultad que T admite una única estructura de superficie de Riemann respecto a la cual p es localmente conforme (simplemente, tomamos como cartas a las inversas locales de p). En particular tenemos definida una topología en T, que resulta ser compacta. Más precisamente, una aplicación $\mathbb{R}$-lineal $\mathbb{C} \longrightarrow \mathbb{R}^2$ que transforme R en $\mathbb{Z}^2$ induce un difeomorfismo

$$T \longrightarrow \mathbb{R}^2/\mathbb{Z}^2 \cong S^1 \times S^1$$

que es, de hecho, un isomorfismo de grupos.

Vemos así que los toros complejos son grupos de Lie algebraicamente difeomorfos todos ellos al grupo $S^1 \times S^1$. No obstante, el difeomorfismo que hemos construido no es en general una transformación conforme y es que, según veremos enseguida, existen infinitos toros complejos no conformemente equivalentes dos a dos.

Para ello observemos que la proyección $p : \mathbb{C} \longrightarrow T$ es lo que en topología se conoce como un cubrimiento, es decir, una aplicación continua y suprayectiva tal que cada punto $[z] \in T$ tiene un entorno abierto V con la propiedad de que p se restringe a un homeomorfismo $p|_U : U \longrightarrow V$ sobre cada componente conexa U de $p^{-1}[V]$. En efecto, basta tomar como V cualquier entorno abierto de $[z]$ donde p tenga inversa. Con esto podemos probar:

Teorema 10.2 *Sean R y S dos retículos complejos y $\phi : \mathbb{C}/R \longrightarrow \mathbb{C}/S$ una aplicación holomorfa. Entonces existen constantes α, $\beta \in \mathbb{C}$ tales que el diagrama siguiente es conmutativo:*

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\tilde{\phi}} & \mathbb{C} \\ {\scriptstyle p_R}\downarrow & & \downarrow{\scriptstyle p_S} \\ \mathbb{C}/R & \xrightarrow[\phi]{} & \mathbb{C}/S \end{array}$$

donde $\tilde{\phi}$ viene dada por $\tilde{\phi}(z) = \alpha z + \beta$.

Demostración: Sea $f : \mathbb{C} \longrightarrow \mathbb{C}/S$ dada por $f = p_R \circ \phi$. Un teorema básico de la topología algebraica (el criterio de elevación) afirma que, como p_S es un cubrimiento de un espacio topológico y su dominio $\mathbb{C}$ es localmente compacto y simplemente conexo, la aplicación continua (arbitraria) f se eleva a $\mathbb{C}$, es decir, existe una aplicación continua $\tilde{\phi} : \mathbb{C} \longrightarrow \mathbb{C}$ tal que $\tilde{\phi} \circ p_S = f$. En concreto, esto significa que $\phi([z]) = [\tilde{\phi}(z)]$.

El hecho de que p_R y p_S sean localmente conformes implica que $\tilde{\phi}$ es holomorfa. Para cada $\omega \in R$, la aplicación $\tilde{\phi}(z+\omega) - \tilde{\phi}(z)$ toma valores en S. Como $\mathbb{C}$ es conexo y S es discreto, ha de ser constante, y su derivada será nula. Así pues, $\tilde{\phi}'(z + \omega) = \tilde{\phi}'(z)$, para todo $z \in \mathbb{C}$ y todo $\omega \in R$.

Esto nos permite definir $\tilde{\phi}' : \mathbb{C}/R \longrightarrow \mathbb{C}$ mediante $\tilde{\phi}'([z]) = \tilde{\phi}'(z)$ y de nuevo por la conformidad local de p_R tenemos que se trata de una aplicación

holomorfa. Concluimos que $\tilde{\phi}'$ es constante, ya que de no serlo su imagen sería abierta en $\mathbb{C}$, pero también compacta, lo cual es absurdo.

Sea, pues, $\alpha \in \mathbb{C}$ tal que $\tilde{\phi}'(z) = \alpha$ para todo $z \in \mathbb{C}$. Obviamente entonces, existe $\beta \in \mathbb{C}$ tal que $\tilde{\phi}(z) = \alpha z + \beta$ para todo $z \in \mathbb{C}$. ∎

De aquí deducimos consecuencias muy importantes.

Definición 10.3 Un *homomorfismo analítico* $\phi : \mathbb{C}/R \longrightarrow \mathbb{C}/S$ entre dos toros complejos es un homomorfismo de grupos que además es una aplicación holomorfa entre las dos superficies de Riemann. Un *isomorfismo analítico* es un homomorfismo analítico biyectivo (en particular una transformación conforme).

Si en el teorema anterior suponemos que $\phi(0) = 0$, entonces se ha de cumplir que $\beta \in S$, y en tal caso $\tilde{\phi}(z) = \alpha z$ induce también la aplicación ϕ,luego podemos tomar $\beta = 0$, con lo que resulta que ϕ es un homomorfismo analítico. Esto demuestra el teorema siguiente:

Teorema 10.4 *Si $\phi : \mathbb{C}/R \longrightarrow \mathbb{C}/S$ es una aplicación holomorfa entre dos toros complejos que cumple $\phi(0) = 0$, entonces ϕ es un homomorfismo analítico, inducido por la multiplicación por un cierto $\alpha \in \mathbb{C}$.*

En particular, una transformación conforme $\phi : \mathbb{C}/R \longrightarrow \mathbb{C}/S$ entre dos toros complejos que cumpla $\phi(0) = 0$ es un isomorfismo analítico. Ahora bien, si $\phi(0) = [\beta]$, podemos considerar la aplicación $\psi : \mathbb{C}/S \longrightarrow \mathbb{C}/S$ dada por $\psi([z]) = [z-\beta]$. Es claro que está bien definida y es una transformación conforme del toro en sí mismo. Por consiguiente, $\phi \circ \psi$ es también una transformación conforme y conserva el 0, luego es un isomorfismo analítico. Con esto hemos probado:

Teorema 10.5 *Dos toros complejos son conformemente equivalentes si y sólo si son analíticamente isomorfos.*

Si $\phi : \mathbb{C}/R \longrightarrow \mathbb{C}/S$ es un isomorfismo analítico entre dos toros complejos, el número α que lo representa en $\mathbb{C}$ ha de cumplir $\alpha R = S$ y, recíprocamente, todo $\alpha \in \mathbb{C}$ que cumpla $\alpha R = S$ induce un isomorfismo analítico entre los toros. Esto nos lleva a la definición siguiente:

Definición 10.6 Dos retículos complejos R y S son *linealmente equivalentes* si existe un número complejo α tal que $\alpha R = S$.

Geométricamente, esto significa que uno puede obtenerse de otro mediante una rotación y una homotecia. Los teoremas anteriores muestran que dos toros $\mathbb{C}/R$ y $\mathbb{C}/S$ son conformemente equivalentes (o analíticamente isomorfos) si y sólo si los retículos R y S son linealmente equivalentes.

Explícitamente, si $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$ y $S = \langle \omega_1', \omega_2' \rangle_{\mathbb{Z}}$, tenemos que R y S son equivalentes si y sólo si existe un $\alpha \in \mathbb{C}^*$ tal que $\langle \alpha\omega_1', \alpha\omega_2' \rangle_{\mathbb{Z}} = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$, lo que a su vez equivale a que existan enteros a, b, c, d tales que

$$\alpha\omega_1' = d\omega_1 + c\omega_2, \quad \alpha\omega_2' = b\omega_1 + a\omega_2, \qquad ad - bc = \pm 1.$$

Esto equivale a que

$$\frac{\omega_2'}{\omega_1'} = \frac{a\omega_2 + b\omega_1}{c\omega_2 + d\omega_1}, \qquad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1.$$

Si llamamos $\tau = \omega_2/\omega_1$ y $\tau' = \omega_2'/\omega_1'$, concluimos que R y S son linealmente equivalentes si y sólo si existen números a, b, c, d tales que

$$\tau' = \frac{a\tau + b}{c\tau + d}, \qquad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1.$$

Observemos que

$$\operatorname{Im}\tau' = \frac{\operatorname{Im}((a\tau + b)(c\bar{\tau} + d)))}{|c\tau + d|^2} = \frac{bc\operatorname{Im}\bar{\tau} + ad\operatorname{Im}\tau}{|c\tau + d|^2} = \frac{ad - bc}{|c\tau + d|^2}\operatorname{Im}\tau. \quad (10.1)$$

Si suponemos que las bases están orientadas entonces $\operatorname{Im}\tau$ e $\operatorname{Im}\tau'$ son positivos, luego $ad - bc > 0$. Con esto hemos probado el teorema siguiente:

Teorema 10.7 *Sean $R = \langle\omega_1, \omega_2\rangle_{\mathbb{Z}}$ y $S = \langle\omega_2', \omega_2'\rangle_{\mathbb{Z}}$ dos retículos complejos, donde las bases están orientadas. Entonces R y S son linealmente equivalentes si y sólo si existen números enteros a, b, c, d tales que*

$$\tau' = \frac{a\tau + b}{c\tau + d}, \qquad ad - bc = 1.$$

Ahora es claro que existen infinitos retículos no equivalentes dos a dos, luego, tal y como afirmábamos, hay infinitos toros complejos no conformemente equivalentes dos a dos.

En particular tenemos que todo retículo complejo $R = \langle\omega_1, \omega_2\rangle_{\mathbb{Z}}$ es linealmente equivalente a otro de la forma $\langle 1, \tau\rangle_{\mathbb{Z}}$, con $\operatorname{Im}\tau > 0$ (sólo tenemos que tomar $\tau = \omega_2/\omega_1$). Por ello en muchas ocasiones podremos limitarnos a trabajar con retículos de esta forma.

10.2 Las funciones de Weierstrass

Dos números complejos ω_1 y ω_2 linealmente independientes sobre $\mathbb{R}$ determinan un paralelogramo en $\mathbb{C}$. El retículo R generado por ω_1 y ω_2 es un artificio algebraico para tratar más cómodamente el toro que resulta de identificar los lados opuestos de dicho paralelogramo. En efecto, dicho toro puede identificarse con el cociente $\mathbb{C}/R$, donde por comodidad trabajamos con todo el plano complejo en lugar de con el paralelogramo determinado por ω_1 y ω_2.

En esta sección veremos que todo toro complejo es conformemente equivalente (e isomorfo como grupo) a una curva elíptica compleja. El isomorfismo induce de forma natural un isomorfismo entre el cuerpo de funciones racionales de la curva elíptica y el cuerpo de funciones meromorfas del toro, las cuales a su vez pueden identificarse de forma natural con las funciones meromorfas en $\mathbb{C}$ con periodos ω_1 y ω_2. Esto nos lleva a las funciones elípticas en el sentido clásico:

Definición 10.8 Una *función elíptica* respecto a un retículo complejo R es una función meromorfa $f : \mathbb{C} \longrightarrow \mathbb{C}^\infty$ tal que $f(z+\omega) = f(z)$ para todo $z \in \mathbb{C}$ y todo $\omega \in R$.

Un número complejo ω que cumpla $f(z+\omega) = f(z)$ para todo $z \in \mathbb{C}$ se llama *periodo* de f. Así pues, los elementos de R son periodos de f, pero no exigimos que sean los únicos periodos de f. Es obvio que para que f sea elíptica respecto de R basta con que tenga por periodos a una base de R. En otras palabras, una función elíptica (respecto a algún retículo) no es más que una función meromorfa doblemente periódica (una función con dos periodos linealmente independientes).

Es evidente que si $\bar{f} : \mathbb{C}/R \longrightarrow \mathbb{C}^\infty$ es una función meromorfa en un toro complejo y $p : \mathbb{C} \longrightarrow \mathbb{C}/R$ es la proyección natural, entonces $f = p \circ \bar{f}$ es una función elíptica sobre R y, recíprocamente, toda función elíptica sobre R induce una función meromorfa $\bar{f}$ tal que $f = p \circ \bar{f}$. En lo sucesivo no distinguiremos entre f y $\bar{f}$.

El hecho de que las funciones elípticas puedan verse como funciones meromorfas sobre una superficie de Riemann compacta tiene muchas consecuencias. Por lo pronto tenemos que las funciones elípticas sobre un retículo forman un cuerpo con las operaciones definidas puntualmente. Más aún, el cuerpo $\mathcal{M}(T)$ de las funciones meromorfas sobre una superficie de Riemann compacta T es un cuerpo de funciones algebraicas de una variable.[1] Además, los puntos de T se corresponden biunívocamente con los divisores primos de $\mathcal{M}(T)$: un punto P se corresponde con el primo $\mathfrak{P}$ tal que $v_{\mathfrak{P}}(f)$ es el orden de la función f en el punto P.

Si $f \in \mathcal{M}(T)$ no es constante, podemos considerarla como elemento del cuerpo $k = \mathbb{C}(f)$, donde su divisor asociado es de la forma $(f) = \mathfrak{p}/\mathfrak{q}$, para ciertos divisores primos $\mathfrak{p}$ y $\mathfrak{q}$ de k. Éstos factorizarán en $\mathcal{M}(T)$ como $\mathfrak{p} = \mathfrak{P}_1 \cdots \mathfrak{P}_n$ y $\mathfrak{q} = \mathfrak{Q}_1 \cdots \mathfrak{Q}_n$, donde los primos $\mathfrak{P}_i$ y $\mathfrak{Q}_i$ (no necesariamente distintos) se corresponden respectivamente con los ceros y polos de f en T (y el número de veces que aparece repetido cada primo es el orden del cero o polo correspondiente). El número n de factores es el mismo para $\mathfrak{p}$ y $\mathfrak{q}$ porque los divisores principales tienen grado 0. Esto significa que (contando multiplicidades) f tiene el mismo numero de ceros que de polos en T. A este número se le llama *orden* de la función f. Observemos que, como $\mathfrak{p}$ tiene grado 1 en k y grado n en $\mathcal{M}(T)$, podemos concluir que

$$|\mathcal{M}(T) : \mathbb{C}(f)| = \text{orden de } f. \tag{10.2}$$

En el caso particular en que $T = \mathbb{C}/R$ es un toro complejo y f es, por lo tanto, una función elíptica, el orden de f es el número de ceros o de polos que tiene f en un paralelogramo fundamental de R (o en cualquiera de sus trasladados), contados siempre con su multiplicidad.

[1] Demostrar esta afirmación en toda su generalidad implica probar un resultado nada trivial, a saber, que en toda superficie de Riemann existen funciones meromorfas que separan puntos. No obstante, para el caso específico de un toro complejo T, construiremos enseguida tales funciones de forma explícita (cf. la nota al pie de la página 292).

De (10.2) se sigue que una función elíptica no constante no puede tener orden 1, ya que entonces sería $\mathcal{M}(T) = \mathbb{C}(f)$, cuando el cuerpo de la derecha tiene género 0 y el de la izquierda tiene género 1 (ya que su superficie de Riemann es homeomorfa a un toro).

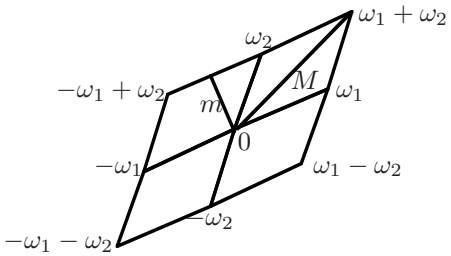
Vamos a probar que todo retículo complejo tiene asociadas funciones elípticas no constantes. El punto de partida es el teorema siguiente:

Teorema 10.9 *Si R es un retículo en $\mathbb{C}$ y $\alpha \in \mathbb{R}$, entonces la serie*

$$\sum_{\omega \in R\setminus\{0\}} \frac{1}{\omega^\alpha}$$

converge (absolutamente) si y sólo si $\alpha > 2$.

Demostración: Notemos que, puesto que no hemos especificado un orden en los sumandos, sólo tiene sentido hablar de convergencia absoluta. Sea ω_1, ω_2 una base de R y llamemos m y M a las distancias mínima y máxima, respectivamente, de 0 a la frontera del paralelogramo indicado en la figura:



Si $\omega \neq 0$ es cualquiera de los 8 elementos de R representados en la figura, tenemos que $m \leq |\omega| \leq M$. Si dibujamos los 12 paralelogramos que rodean a los 4 que aparecen en la figura, encontraremos 16 nuevos elementos $\omega \in R$ tales que $2m \leq |\omega| \leq 2M$. A continuación encontraremos 24 nuevos elementos tales que $3m \leq |\omega| \leq 3M$, etc.

En general, encontramos $8k$ elementos de R que satisfacen las desigualdades

$$\frac{1}{(kM)^\alpha} \leq \frac{1}{|\omega|^\alpha} \leq \frac{1}{(km)^\alpha}.$$

Si llamamos $S(n) = \sum |\omega|^{-\alpha}$, donde ω recorre los $8(1+2+\cdots+n)$ elementos de R más cercanos a 0, tenemos que

$$\frac{8}{M^\alpha} + \frac{2\cdot 8}{(2M)^\alpha} + \cdots + \frac{n\cdot 8}{(nM)^\alpha} \leq S(n) \leq \frac{8}{m^\alpha} + \frac{2\cdot 8}{(2m)^\alpha} + \cdots + \frac{n\cdot 8}{(nm)^\alpha},$$

lo cual equivale a

$$\frac{8}{M^\alpha}\sum_{k=1}^{n} \frac{1}{k^{\alpha-1}} \leq S(n) \leq \frac{8}{m^\alpha}\sum_{k=1}^{n} \frac{1}{k^{\alpha-1}}.$$

Si $\alpha > 2$, la serie de la derecha es convergente, luego $S(n)$ también, mientras que si $\alpha \leq 2$ la serie de la izquierda es divergente, luego $S(n)$ también. La convergencia de $S(n)$ equivale a la convergencia absoluta de la serie del enunciado.

∎

La definición siguiente nos da la función más elemental que puede asociarse a un retículo complejo de forma natural. No es una función elíptica, pero nos permitirá construir otras que sí lo son.

Definición 10.10 La *función sigma de Weierstrass* asociada a un retículo complejo R es la función $\sigma : \mathbb{C} \longrightarrow \mathbb{C}$ dada por

$$\sigma(z) = z \prod_{\omega \in R\setminus\{0\}} \left(1 - \tfrac{z}{\omega}\right) e^{\frac{z}{\omega} + \frac{z^2}{2\omega^2}}.$$

Según la teoría básica sobre productos infinitos, la convergencia de la serie del teorema anterior para $\alpha = 3$ implica la convergencia absoluta y casi uniforme del producto en todo el plano complejo. Tenemos así una función entera con ceros simples en todos los puntos de R (y sólo en ellos). Se trata de una función impar, pues $-\omega$ recorre $R \setminus \{0\}$ cuando ω lo hace, y claramente

$$\lim_{z\to 0} \frac{\sigma(z)}{z} = 1.$$

Todos estos hechos determinan una cierta analogía entre las funciones $\sigma(z)$ y $\operatorname{sen} z$. Observemos que $\sigma(z)$ no puede ser una función elíptica porque tendría un único cero y ningún polo en cada paralelogramo fundamental de R.

Introducimos ahora la *función dseta de Weierstrass,* definida por

$$\zeta(z) = \frac{\sigma'(z)}{\sigma(z)}.$$

Se trata de una función impar meromorfa en $\mathbb{C}$ con polos simples en los puntos de R (lo que ya implica que no puede ser elíptica). La convergencia absoluta del producto que define la función sigma equivale a la convergencia absoluta y casi uniforme de la serie

$$\log \frac{\sigma(z)}{z} = \sum_{\omega \in R\setminus\{0\}} \left(\log\left(1 - \frac{z}{\omega}\right) + \frac{z}{\omega} + \frac{z^2}{2\omega^2}\right),$$

la cual determina un logaritmo holomorfo de $\sigma(z)/z$ en un entorno de 0. El logaritmo que aparece en la serie es el que toma partes imaginarias en $]-\pi, \pi[$. Derivando queda

$$\frac{\sigma'(z)}{\sigma(z)} - \frac{1}{z} = \sum_{\omega \in R\setminus\{0\}} \left(\frac{-1/\omega}{1 - z/\omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right),$$

luego

$$\zeta(z) = \frac{1}{z} + \sum_{\omega \in R\setminus\{0\}} \left(\frac{1}{z - \omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right).$$

En principio tenemos probada esta igualdad en un entorno de 0, pero por el principio de prolongación analítica se cumple en todo $\mathbb{C}$, dado que la serie converge uniformemente en todo compacto que no contenga puntos de R. En efecto, el término general es $z^2/(z-\omega)\omega^2$, y basta compararlo con $1/\omega^3$.

Vemos que ζ tiene residuo 1 en todos sus polos. Volviendo a derivar llegamos a que

$$\zeta'(z) = -\frac{1}{z^2} - \sum_{\omega\in R\setminus\{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right).$$

Definimos la *función $\wp$ de Weierstrass* como

$$\wp(z) = -\zeta'(z) = \frac{1}{z^2} + \sum_{\omega\in R\setminus\{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right).$$

Es claro que $\wp$ es meromorfa en $\mathbb{C}$, con polos dobles en los puntos de R. La serie que la define converge uniformemente en todo compacto que no contenga puntos de R, al igual que sucede con la serie que resulta de derivar término a término:

$$\wp'(z) = -2\sum_{\omega\in R} \frac{1}{(z-\omega)^3}.$$

Es evidente que $\wp'(z)$ es una función elíptica sobre R (de orden 3). Vamos a ver que $\wp(z)$ también lo es (pero de orden 2). Observemos para ello que $\wp(z)$ es par. En efecto, usamos que

$$(-z-\omega)^2 = (z+\omega)^2 = (z-(-\omega))^2,$$

sumamos sobre ω y tenemos en cuenta que $-\omega$ recorre R cuando ω lo hace.

Como $\wp'$ es elíptica sobre R, si $\omega \in R$ la función $\wp(z+\omega)-\wp(z)$ es constante. Ahora bien, para $z = -\omega/2$ ha de ser

$$\wp(z+\omega) - \wp(z) = \wp(\omega/2) - \wp(-\omega/2) = 0,$$

luego $\wp(z+\omega) - \wp(z)$ es la función nula.

Así pues, para cara retículo complejo R hemos encontrado dos funciones elípticas no constantes[2] asociadas a R, las funciones $\wp(z)$ y $\wp'(z)$. Ambas se conocen como *funciones de Weierstrass* de R.

Hemos demostrado que la función $\wp$ es par, es decir, cumple $\wp(-z) = \wp(z)$, mientras que de la propia definición como serie infinita se sigue que $\wp'$ es impar, es decir, que $\wp'(-z) = -\wp'(z)$.

Como los órdenes de $\wp$ y $\wp'$ son 2 y 3 respectivamente, la relación (10.2) nos da que $|\mathcal{M}(T) : \mathbb{C}(\wp)| = 2$ y $|\mathcal{M}(T) : \mathbb{C}(\wp')| = 3$, de donde concluimos inmediatamente que $\mathcal{M}(T) = \mathbb{C}(\wp, \wp')$, en otras palabras:

[2] En particular vemos que $\wp(z)$ separa el 0 del toro $T = \mathbb{C}/R$ de cualquier otro punto $p \in T$ (en el sentido de que $\wp(0) \neq \infty \neq \wp(p)$). Componiendo con traslaciones vemos que T tiene funciones meromorfas que separan cualquier par de puntos, tal y como habríamos anticipado.

Teorema 10.11 *Toda función elíptica sobre un retículo complejo R se expresa como una función racional de las funciones de Weierstrass $\wp$ y $\wp'$ de R.*

Las funciones $\wp$ y $\wp'$ deben satisfacer una ecuación polinómica (pues ambas pertenecen a un mismo cuerpo de funciones algebraicas). Vamos a encontrar explícitamente esta ecuación.

Definición 10.12 Sea R un retículo en $\mathbb{C}$. Para cada natural $n \geq 3$, la *serie de Eisenstein* de R de orden n es

$$G_n = \sum_{\omega \in R\setminus\{0\}} \frac{1}{\omega^n}.$$

El teorema 10.9 prueba la convergencia de las series de Eisenstein. En la prueba del teorema siguiente se ve que $G_{2k+1} = 0$ para todo n:

Teorema 10.13 *La serie de Laurent en 0 de la función de Weierstrass $\wp$ de un retículo R es*

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} (2n+1) G_{2n+2}\, z^{2n}.$$

Demostración: Usamos el desarrollo de Taylor

$$\frac{1}{(1-z)^2} = \sum_{n=0}^{\infty} (n+1) z^n, \qquad |z| < 1.$$

Sea $m > 0$ el menor módulo de un elemento no nulo de R. Si $0 < |z| < m$ y $\omega \in R$ es no nulo, entonces $|z/\omega| < 1$ y

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

luego

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

Sumando sobre ω y teniendo en cuenta que todas las series convergen absolutamente,

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} (n+1) \sum_{\omega \in R\setminus\{0\}} \frac{1}{\omega^{n+2}}\, z^n = \frac{1}{z^2} + \sum_{n=1}^{\infty} (n+1) G_{n+2}\, z^n.$$

Como $\wp$ es una función par, las series de Eisenstein G_{2k+1} han de ser nulas, con lo que queda la expresión del enunciado. ∎

Ahora podemos probar:

Teorema 10.14 *La función $\wp$ de un retículo R satisface la ecuación diferencial*

$$\wp'^2 = 4\wp^3 - 60 G_4 \wp - 140 G_6.$$

DEMOSTRACIÓN: Derivando la serie de Laurent de $\wp$ obtenemos que (para todo z cercano a 0)

$$\wp'(z) = -\frac{2}{z^3} + 6G_4 z + 20G_6 z^3 + \cdots$$

Por lo tanto

$$\wp'(z)^2 = \frac{4}{z^6} - \frac{24G_4}{z^2} - 80G_6 + \cdots$$

donde los puntos suspensivos representan una función holomorfa que se anula en 0. Por otra parte,

$$\wp(z)^3 = \frac{1}{z^6} + \frac{9G_4}{z^2} + 15G_6 + \cdots$$

Por consiguiente

$$\wp'(z)^2 - 4\wp^3 + 60G_4\wp(z) = -140G_6 + \cdots$$

Esta función es elíptica y no tiene polos, luego ha de ser constante, lo que nos da la ecuación diferencial que buscamos. ∎

Si llamamos $g_2 = 60G_4$ y $g_3 = 140G_6$ tenemos que la función $\wp$ satisface la ecuación diferencial

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

Equivalentemente, las funciones $\wp$ y $\wp'$ satisfacen la ecuación algebraica

$$Y^2 = 4X^3 - g_2X - g_3.$$

Esto es lo que en 2.4 hemos llamado una ecuación de Weierstrass clásica, que no es una ecuación de Weierstrass en el sentido general, pero que en realidad es el tipo de ecuación que consideró Weierstrass al estudiar las funciones elípticas debido a que, como acabamos de ver, es el tipo de ecuación que relaciona a las funciones $\wp$ y $\wp'$. Recordemos que hemos definido el discriminante y el invariante de una ecuación de este tipo como los correspondientes a la ecuación que resulta de hacer $Y = 2Y'$, a saber,[3]

$$\Delta = g_2^3 - 27g_3^2, \qquad j = \frac{1728\, g_2^3}{g_2^3 - 27g_3^2}.$$

Observemos que el cuerpo $\mathbb{C}(\wp, \wp')$ de las funciones elípticas sobre el retículo R es isomorfo al cuerpo de las funciones racionales de la curva proyectiva E determinada por la ecuación. Como $\mathbb{C}(\wp, \wp')$ es también isomorfo al cuerpo de las funciones meromorfas del toro complejo $\mathbb{C}/R$, tiene género 1, luego lo mismo le ha de suceder a E, luego E ha de ser una cúbica regular, es decir, una curva elíptica. Equivalentemente, se ha de cumplir que $\Delta \neq 0$, lo que a su vez equivale a que el polinomio $4X^3 - g_2X - g_3$ tenga tres raíces distintas. Vamos a calcularlas explícitamente:

[3]Es frecuente definir $J = g_2^3/(g_2^7 - 27g_3^2)$, lo cual es más natural en algunos aspectos, pero no en otros. Por ejemplo, la serie de Fourier del teorema 10.32 no son enteros si usamos con J en lugar de j.

Definición 10.15 Si $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$ es un retículo complejo y $\wp$ es su correspondiente función de Weierstrass, definimos

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

El teorema siguiente muestra, entre otras cosas, que estos números no dependen (salvo el orden) de la elección de la base de R:

Teorema 10.16 *Si R es un retículo en $\mathbb{C}$, entonces*

$$4X^3 - g_2X - g_3 = 4(X-e_1)(X-e_2)(X-e_3).$$

Además, las raíces e_1, e_2, e_3 son distintas dos a dos.

DEMOSTRACIÓN: Sea α_i uno de los tres números $\omega_1/2$, $\omega_2/2$, $(\omega_1+\omega_2)/2$, de modo que $\wp(\alpha_i) = e_i$. Claramente $\alpha_i \notin R$, pero $2\alpha_i \in R$. La función $\wp(z) - e_i$ tiene un cero doble en α_i, pues, usando que $\wp'$ es impar, vemos que

$$-\wp'(\alpha_i) = \wp'(-\alpha_i) = \wp'(-\alpha_i + 2\alpha_i) = \wp'(\alpha_i),$$

luego ha de ser $\wp'(\alpha_i) = 0$ y así α_i es un cero de $\wp - e_i$ de orden al menos dos. Contando los polos vemos que la función tiene orden 2, luego α_i ha de ser su único cero y su orden ha de ser exactamente 2.

Por consiguiente, la función $g(z) = 4(\wp-e_1)(\wp-e_2)(\wp-e_3)$ tiene exactamente tres ceros dobles. Lo mismo podemos decir de la función $\wp'^2$, que tiene orden 6 (sólo tiene polos de orden 6 en los puntos de R) y acabamos de ver que tiene ceros de orden al menos 2 en los mismos puntos que g. Así pues, dichos ceros han de ser exactamente de orden 2 y no puede tener más. El cociente

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

ha de ser constante y, calculando su límite en 0 (dividiendo entre $\wp^3$), vemos que la constante ha de ser 1, y así tenemos la igualdad

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

Es claro que la función $\wp$ toma infinitos valores distintos, luego los dos miembros de la igualdad del enunciado son polinomios que coinciden en infinitos puntos, luego son iguales.

Para terminar observemos que, aunque ya sabemos que los e_i han de ser distintos, ahora podemos dar una sencilla prueba directa: si —por ejemplo— fuera $e_1 = e_2$, entonces $\wp(z) - e_1$ tendría un cero doble en $\omega_1/2$ y otro en $\omega_2/2$, pero ya hemos visto que su orden es 2, luego esto es imposible. ∎

Según hemos visto hasta aquí, a cada retículo complejo R le hemos asociado una función de Weierstrass $\wp$ y unos números complejos g_2 y g_3 que a su vez determinan una ecuación de Weierstrass que a su vez determina una curva elíptica plana. A continuación demostramos que las funciones $\wp$ y $\wp'$ parametrizan la curva de forma natural:

Teorema 10.17 *Sea $T = \mathbb{C}/R$ un toro complejo, sea $\wp$ su función de Weierstrass y sea E la curva elíptica dada por $Y^2 = 4X^3 - g_2X - g_3$. Entonces la aplicación $\phi : T \longrightarrow E$ dada por $\phi(P) = (\wp(P), \wp'(P))$ para $P \neq 0$ y $\phi(0) = O$ es una transformación conforme y un isomorfismo de grupos.*

DEMOSTRACIÓN: Obviamente $\phi|_{T\setminus\{0\}} : T \setminus \{0\} \longrightarrow \mathbb{C}^2$ es holomorfa, luego también lo es como aplicación en E. Así pues, ϕ es holomorfa salvo a lo sumo en 0. Para probar que también aquí lo es, hemos de componerla con una carta de E alrededor de O. Sirve la restricción de cualquier función $f \in \mathbb{C}(E)$ tal que $v_O(f) = 1$. Por ejemplo, podemos tomar $f = x/y$. Claramente, $(\phi \circ f)(P) = \wp(P)/\wp'(P)$, que es una función holomorfa en un entorno de 0.

Observemos ahora que ϕ es biyectiva. En efecto, si $\wp(P_1) = \wp(P_2) = \alpha$ y $\wp'(P_1) = \wp'(P_2)$ con $P_1 \neq P_2$, entonces $P_i \neq 0$, pues $\wp$ sólo tiene un polo en 0. Así pues, α es finito y podemos considerar la función $\wp - \alpha$. Vemos que tiene ceros en P_1 y P_2. Como $\wp$ es par, también $-P_2$ es un cero de $\wp - \alpha$.

Distinguimos dos casos: si $P_2 \neq -P_2$, entonces, dado que la función $\wp - \alpha$ tiene orden 2, ha de ser $P_1 = -P_2$, pero entonces $\wp'(P_1) = -\wp'(P_2)$, lo que obliga a que $\wp'(P_1) = \wp'(P_2) = 0$, pero esto es imposible, ya que entonces $\wp - \alpha$ tendría (contando multiplicidades) al menos cuatro ceros.

La otra posibilidad es $P_2 = -P_2$, y entonces $\wp'(P_2) = \wp'(-P_2) = -\wp'(P_2)$, luego ha de ser $\wp'(P_2) = 0$ y así $\wp - \alpha$ tiene un cero en P_1 y dos en P_2, en contradicción nuevamente con que su orden es 2.

La suprayectividad de ϕ es ahora trivial, pues la imagen ha de ser abierta porque ϕ es holomorfa y ha de ser cerrada porque T es compacto. Sólo falta demostrar que ϕ es un homomorfismo de grupos.

Consideramos la aplicación $f : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}/R$ dada por

$$f(z_1, z_2) = \phi^{-1}(\phi([z_1]) + \phi([z_2])).$$

Obviamente es holomorfa. En particular es continua, por lo que el mismo criterio de elevación que hemos usado en la prueba de 10.2 nos da ahora una aplicación continua $\tilde{f} : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ tal que $f(z_1, z_2) = [\tilde{f}(z_1, z_2)]$. Como la proyección de $\mathbb{C}$ en el toro es localmente conforme, tenemos que $\tilde{f}$ es holomorfa.

Pongamos que $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$. Es claro que para m_1, $m_2 \in \mathbb{Z}$,

$$f(z_1, z_2 + m_1\omega_1 + m_2\omega_2) = f(z_1),$$

luego existen n_1, $n_2 \in \mathbb{Z}$ dependientes de z_1, z_2, m_1 y m_2, tales que

$$\tilde{f}(z_1, z_2 + m_1\omega_1 + m_2\omega_2) = \tilde{f}(z_1, z_2) + n_1\omega_1 + n_2\omega_2.$$

Fijados m_1 y m_2, la aplicación $(z_1, z_2) \mapsto (n_1, n_2)$ es claramente continua, luego constante, es decir, n_1 y n_2 no dependen de z_1 y z_2. Por lo tanto, podemos derivar:

$$\frac{\partial \tilde{f}}{\partial z_1}(z_1, z_2 + m_1\omega_1 + m_2\omega_2) = \frac{\partial \tilde{f}}{\partial z_1}(z_1, z_2),$$

$$\frac{\partial \tilde{f}}{\partial z_2}(z_1, z_2 + m_1\omega_1 + m_2\omega_2) = \frac{\partial \tilde{f}}{\partial z_2}(z_1, z_2).$$

Esto significa que las derivadas (para un z_1 fijo) son funciones elípticas, pero no tienen polos, luego son constantes o, dicho de otro modo, no dependen de z_2. Igualmente se demuestra que no dependen de z_1. Esto se traduce en que

$$\tilde{f}(z_1, z_2) = az_1 + bz_2 + c,$$

para ciertos a, b, $c \in \mathbb{C}$. Por lo tanto:

$$\phi([az_1] + [bz_2] + [c]) = \phi([z_1]) + \phi([z_2]).$$

Evaluando en $(0,0)$ queda que $\phi([c]) = O$, luego $[c] = 0$. Evaluando en $(z_1, 0)$ tenemos que $\phi([az_1]) = \phi([z_1])$, luego $[az_1] = [z_1]$ para todo z_1, e igualmente $[bz_2] = [z_2]$, luego la ecuación se reduce a

$$\phi([z_1] + [z_2]) = \phi([z_1]) + \phi([z_2]).$$

■

Extendiendo la definición 10.3 a aplicaciones entre toros complejos y curvas elípticas, podemos decir que la aplicación ϕ dada por el teorema anterior es un *isomorfismo analítico*. Es obvio que dos toros complejos son analíticamente isomorfos si y sólo si lo son sus curvas elípticas asociadas, lo cual equivale a que sean isomorfas como curvas elípticas, lo cual equivale a su vez a que tengan el mismo invariante. De este modo, hemos asociado un invariante $j \in \mathbb{C}$ a cada retículo complejo (o a cada toro complejo) de tal forma que dos retículos son linealmente equivalentes (y dos toros son analíticamente isomorfos o conformemente equivalentes) si y sólo si tienen el mismo invariante.

Ejemplo Vamos a determinar el invariante del retículo $R = \langle 1, i\rangle_{\mathbb{Z}}$.

En primer lugar observamos que cuando (m,n) recorre $\mathbb{Z}^2 \setminus (0,0)$, entonces tanto $m + ni$ como $i(m + ni) = -n + mi$ recorren $R \setminus \{0\}$. Por lo tanto,

$$G_6 = \sum_{(m,n)\neq(0,0)} \frac{1}{(i(m+ni))^6} = -\sum_{(m,n)\neq(0,0)} \frac{1}{(m+ni)^6} = -G_6,$$

luego $G_6 = 0$ y $g_3 = 140G_6 = 0$. Respecto a G_4 tenemos que

$$\begin{aligned} G_4 &= \sum_{m\in\mathbb{Z}\setminus\{0\}} \frac{1}{m^4} + \sum_{n>0} \left(\frac{1}{(m+ni)^4} + \frac{1}{(m-ni)^4} \right) \\ &= \sum_{m\in\mathbb{Z}\setminus\{0\}} \frac{1}{m^4} + \sum_{n>0} \frac{(m-ni)^4 + (m+ni)^4}{(m^2+n^2)^4}. \end{aligned}$$

Desarrollando el numerador concluimos fácilmente que $G_4 > 0$, luego también $g_2 = 60G_4 > 0$.

Así pues, el toro $\mathbb{C}/R$ es isomorfo a una curva elíptica dada por una ecuación de la forma

$$Y^2 = 4X^3 - g_2X.$$

En particular su invariante es $j = 1728$. Así, aunque no tengamos el valor de g_2, tenemos determinada la curva salvo isomorfismo. (Dicho valor puede aproximarse numéricamente, y resulta ser $g_2 = 0{,}738565313004819736\ldots$) ■

Ejemplo Sea ρ una raíz cúbica primitiva de la unidad. Vamos a calcular el invariante del retículo $R = \langle 1, \rho\rangle_{\mathbb{Z}}$.

Teniendo en cuenta que $\rho^2 + \rho + 1 = 0$, calculamos

$$G_4 = \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{(m+n\rho)^4} = \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{(m\rho^3+n\rho)^4}$$

$$= \frac{1}{\rho^4} \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{(m\rho^2+n)^4} = \frac{1}{\rho} \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{((n-m)-m\rho)^4}$$

$$= \frac{1}{\rho} \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{(m+n\rho)^4} = \frac{1}{\rho} G_4,$$

de donde se sigue que $G_4 = 0$, e igualmente $g_2 = 60G_4 = 0$. Por consiguiente el invariante es $j = 0$. ∎

Terminamos la sección citando un resultado que demostraremos más tarde, en el capítulo XII:

Teorema 10.18 (Teorema de Uniformización) *Toda curva elíptica es analíticamente isomorfa a un toro complejo.*

Equivalentemente, se trata de probar que para todo $j \in \mathbb{C}$ existe un retículo complejo cuya curva elíptica asociada tiene invariante j (y así lo demostraremos en 12.3). Esto significa que toda curva elíptica $E/\mathbb{C}$ es isomorfa a una curva $E'/\mathbb{C}$ asociada a un toro complejo T.

Componiendo los isomorfismos $T \xrightarrow{\phi} E' \longrightarrow E$ obtenemos un isomorfismo analítico entre T y E. Las funciones coordenadas de este isomorfismo son funciones racionales en $\wp$ y $\wp'$, luego son funciones elípticas sobre R. En particular, toda curva elíptica $E/\mathbb{C}$ se puede parametrizar (uniformizar) con dos funciones elípticas sobre un retículo apropiado.

10.3 Isogenias complejas

En esta sección estudiaremos los homomorfismos analíticos entre dos toros complejos. Usaremos el teorema de uniformización que acabamos de enunciar para traducir algunas de las consecuencias que obtendremos a resultados sobre isogenias entre curvas elípticas complejas.

Antes de entrar propiamente en esta cuestión, observemos que una curva elíptica sobre $\mathbb{A}$ tiene menos puntos y, por consiguiente, menos traslaciones, que su extensión a $\mathbb{C}$, pero vamos a ver que la extensión no introduce más isogenias:

Teorema 10.19 *Si $E_1/\mathbb{A}$ y $E_2/\mathbb{A}$ son dos curvas elípticas, entonces toda isogenia entre las extensiones $E_1/\mathbb{C}$ y $E_2/\mathbb{C}$ es la extensión de una isogenia entre $E_1/\mathbb{A}$ y $E_2/\mathbb{A}$.*

DEMOSTRACIÓN: Sea $\phi : E_1 \longrightarrow E_2$ una isogenia entre las extensiones. Según las observaciones tras el teorema 1.17, basta ver que $\phi[E_1(\mathbb{A})] \subset E_2(\mathbb{A})$. Supongamos, por el contrario, que existe un punto $P \in E_1(\mathbb{A})$ de manera que $\phi(P) \notin E_2(\mathbb{A})$. Para cada $\mathbb{A}$-automorfismo σ de $\mathbb{C}$ tenemos que $\phi^\sigma(P) = \phi(P)^\sigma$. Si $\phi(P)$ tiene una coordenada trascendente, entonces $\phi(P)^\sigma$ recorre una cantidad no numerable de puntos de $E_2(\mathbb{C})$ al variar σ, luego ϕ^σ recorre una cantidad no numerable de isogenias en $\mathrm{Hom}(E_1, E_2)$, pero esto es imposible, pues el grupo de isogenias es numerable por el teorema 3.23. ∎

Así pues, en las condiciones del teorema anterior, el grupo $\mathrm{Hom}(E_1, E_2)$ es el mismo tanto si consideramos las curvas sobre $\mathbb{A}$ o sobre $\mathbb{C}$. (Más precisamente, la extensión y la restricción son isomorfismos mutuamente inversos entre ambos grupos.) En particular lo mismo sucede con el anillo de endomorfismos $\mathrm{End}(E)$ de una curva elíptica $E/\mathbb{A}$.

Si $E_1/\mathbb{C}$ y $E_2/\mathbb{C}$ son dos curvas complejas, podemos representarlas mediante dos toros complejos $E_i \cong T_i = \mathbb{C}/R_i$, y entonces es claro que tenemos un isomorfismo $\mathrm{Hom}(E_1, E_2) \cong \mathrm{Hom}(T_1, T_2)$, donde el segundo grupo es el de los homomorfismos analíticos entre los toros. El isomorfismo se obtiene componiendo cada isogenia con sendos isomorfismos analíticos fijos entre las curvas y los toros.

El interés de esta representación de las isogenias se debe al teorema 10.4, según el cual cada homomorfismo analítico entre los toros complejos es de la forma $[z] \mapsto [\alpha z]$, donde $\alpha \in \mathbb{C}$ ha de cumplir que $\alpha R_1 \subset R_2$. Notemos que α está completamente determinado por el homomorfismo, pues si $[\alpha z] = [\beta z]$ para todo $z \in \mathbb{C}$, entonces $(\alpha - \beta)z \in R_2$ para todo $z \in \mathbb{C}$, lo cual sólo es posible si $\alpha - \beta = 0$.

Esto nos permite identificar cada homomorfismo analítico entre T_1 y T_2 con el número complejo α que lo determina. De este modo podemos ver a $\mathrm{Hom}(T_1, T_2)$ como un subgrupo de $\mathbb{C}$. (Notemos que, en efecto, la suma de números complejos se corresponde con la suma de homomorfismos definida puntualmente.)

Si los dos toros son el mismo $T = \mathbb{C}/R$, entonces la composición de endomorfismos de T se corresponde con el producto de números complejos, por lo que $\mathrm{End}(T)$ resulta ser un subanillo de $\mathbb{C}$.

En particular vemos que $\mathrm{End}(T)$ es un anillo conmutativo. Esto es un hecho no trivial sobre curvas elípticas: Si $E/\mathbb{C}$ es una curva elíptica, tenemos definido un monomorfismo de anillos $\mathrm{End}(E) \longrightarrow \mathbb{C}$ que prueba que $\mathrm{End}(E)$ es conmutativo.

Esto restringe al teorema 3.35, pues ahora sabemos que $\mathrm{End}(E)$ no puede ser un orden en un álgebra de cuaternios, sino que ha de ser isomorfo a $\mathbb{Z}$ o bien a un orden cuadrático imaginario. El teorema siguiente lo muestra de forma explícita:

Teorema 10.20 *Sea $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$ un retículo complejo, sea $T = \mathbb{C}/R$ y llamemos $\tau = \omega_2/\omega_1$. Se cumple* $\mathrm{End}(T) \neq \mathbb{Z}$ *si y sólo si $\mathbb{Q}(\tau)$ es un cuerpo cuadrático imaginario, y en tal caso* $\mathrm{End}(T)$ *es un orden de $\mathbb{Q}(\tau)$.*

DEMOSTRACIÓN: Supongamos que existe un $\alpha \in \mathrm{End}(T)$ que no es entero. Entonces $\alpha R \subset R$, y existen a, b, c, $d \in \mathbb{Z}$ tales que

$$\alpha\omega_1 = a\omega_1 + b\omega_2, \qquad \alpha\omega_2 = c\omega_1 + d\omega_2.$$

Dividiendo entre ω_1 obtenemos que $\alpha = a + b\tau$, luego $\mathbb{Q}(\alpha) = \mathbb{Q}(\tau)$. Más aún:

$$\begin{vmatrix} \alpha - a & -b \\ -c & \alpha - d \end{vmatrix} = 0,$$

y desarrollando el determinante concluimos que α es raíz de un polinomio mónico de grado 2 con coeficientes enteros. En particular $\mathbb{Q}(\tau)$ es un cuerpo cuadrático (imaginario, porque si τ fuera real entonces ω_1 y ω_2 serían linealmente dependientes sobre $\mathbb{R}$) y α está en el anillo de enteros $\mathcal{O}$ de $\mathbb{Q}(\tau)$. Concluimos que $\mathrm{End}(T) \subset \mathcal{O}$, de donde se sigue inmediatamente que $\mathrm{End}(T)$ es un orden de $\mathbb{Q}(\tau)$.

Supongamos ahora que $\mathbb{Q}(\tau)$ es un cuerpo cuadrático imaginario. Pongamos que τ cumple la ecuación

$$a\tau^2 + b\tau + c = 0, \qquad a,\ b,\ c \in \mathbb{Z},\ a \neq 0.$$

Definimos $\alpha = a\tau \notin \mathbb{R}$. Entonces

$$\alpha\omega_1 = a\omega_2, \qquad \alpha\omega_2 = a\tau^2\omega_1 = -b\tau\omega_1 - c\omega_1 = -c\omega_1 - b\omega_2,$$

luego $\alpha R \subset R$ y, por lo tanto, $\alpha \in \mathrm{End}(T)$. ∎

Por otra parte, si $\mathcal{O}$ es un orden cuadrático imaginario, entonces $\mathcal{O} = \langle 1, \tau \rangle_{\mathbb{Z}}$, para cierto entero cuadrático τ, luego $\mathcal{O}$ es un retículo complejo. Obviamente $\alpha\mathcal{O} \subset \mathcal{O}$ si y sólo si $\alpha \in \mathcal{O}$, luego se cumple que $\mathrm{End}(\mathbb{C}/\mathcal{O}) = \mathcal{O}$. Vemos así que todo orden cuadrático imaginario puede representarse como el anillo de endomorfismos de un toro complejo (o de una curva elíptica).

Si $\alpha \in \mathrm{End}(T)$ no es entero, es costumbre decir que α es una *multiplicación compleja* de T. Más en general, si $E/\mathbb{C}$ es una curva elíptica analíticamente isomorfa a un toro T, las multiplicaciones complejas de T se llaman también multiplicaciones complejas de E. Esto explica por qué —como ya habíamos comentado en la página 79— cuando una curva elíptica cumple $\mathrm{End}(E) \neq \mathbb{Z}$, se dice que tiene *multiplicación compleja.*

Conviene observar que las multiplicaciones complejas de una curva elíptica $E/\mathbb{C}$ no dependen de la elección toro complejo T isomorfo a E (ni de la elección del isomorfismo). En efecto, si $E \cong T_1 \cong T_2$, entonces tenemos un isomorfismo $\sigma : \mathrm{End}(T_1) \longrightarrow \mathrm{End}(T_2)$, que se extiende a un isomorfismo entre los correspondientes cuerpos de cocientes, pero dos cuerpos cuadráticos isomorfos son iguales, digamos a un mismo cuerpo K, y entonces σ es la restricción a

$\mathrm{End}(T_1)$ de un automorfismo de K, que sólo puede ser la identidad o la conjugación compleja, y es claro que un orden cuadrático es cerrado para ambas, luego $\mathrm{End}(T_1) = \mathrm{End}(T_2)$.

Así pues, si $E/\mathbb{C}$ es una curva elíptica con multiplicación compleja, existe un único orden cuadrático imaginario $\mathcal{O}$ tal que $\mathrm{End}(E) \cong \mathcal{O}$ y, más aún, existen únicamente dos isomorfismos $\phi : \mathrm{End}(E) \longrightarrow \mathcal{O}$, pues si consideramos dos de ellos, ϕ y ψ, vemos que $\phi^{-1} \circ \psi$ es un automorfismo de $\mathcal{O}$ que se extiende a un automorfismo del cuerpo de cocientes K, luego $\phi^{-1} \circ \psi$ ha de ser la identidad o la conjugación compleja, luego $\psi = \phi$ o bien $\psi = \overline{\phi}$.

El teorema siguiente muestra que podemos escoger canónicamente uno de los dos isomorfismos:

Teorema 10.21 *Si $E/\mathbb{C}$ es una curva elíptica con multiplicación compleja cuyo anillo de endomorfismos es isomorfo a un orden cuadrático $\mathcal{O}$, entonces existe un único isomorfismo $\phi : \mathcal{O} \longrightarrow \mathrm{End}(E)$ tal que*

$$\overline{\phi(\alpha)}(\omega) = \alpha\omega,$$

para todo $\alpha \in \mathcal{O}$ y toda diferencial de primera clase ω de $E/\mathbb{C}$.

Demostración: Fijemos un isomorfismo analítico $f : \mathbb{C}/R \longrightarrow E$ entre un toro complejo y la curva dada. Vamos a demostrar que el isomorfismo ϕ inducido por f tiene la propiedad requerida.

La proyección $\pi : \mathbb{C}/R \longrightarrow \mathbb{C}$ es localmente inyectiva, y sus inversas locales $z = (\pi|_U)^{-1}$ son cartas en el toro. Si tomamos dos cartas de esta forma z_1 y z_2 con parte de su dominio en común, en la intersección $z_1 - z_2$ toma valores en R, luego es localmente constante y $dz_1 = dz_2$. Esto nos permite definir dz como la única forma diferencial en $\mathbb{C}/R$ que restringida a un entorno de cada punto coincide con la diferencial de cualquier inversa local de la proyección. (Notemos que, globalmente, dz no es la diferencial de ninguna función meromorfa.)

Es claro que dz es una diferencial de primera clase en el toro. Recordemos que el espacio de las diferenciales de primera clase en una superficie de Riemann compacta tiene dimensión igual a su género, que en nuestro caso es 1. Por consiguiente, si ω es una diferencial de primera clase en E, ha de ser $\overline{f}(\omega) = c\,dz$, para cierto $c \in \mathbb{C}^*$. Ahora basta calcular:

$$\overline{\phi(\alpha)}(\omega) = \overline{f^{-1}}(\overline{\alpha}(\overline{f}(\omega)) = \overline{f^{-1}}(\overline{\alpha}(c\,dz)) = \overline{f^{-1}}(c\,d(\alpha z)) = \alpha\overline{f^{-1}}(c\,dz) = \alpha\omega.$$

(Aquí hemos usado que localmente se cumple $\alpha \circ z = \alpha z$, luego también $d(\alpha \circ z) = \alpha\,dz$, y esto vale globalmente.)

El otro isomorfismo $\psi : \mathcal{O} \longrightarrow \mathrm{End}(E)$ es la composición de la conjugación compleja con ϕ, luego cumple $\overline{\psi(\alpha)}(\omega) = \phi(\overline{\alpha})(\omega) = \overline{\alpha}\,\omega$, luego ϕ es el único que satisface la relación del enunciado. ∎

En lo sucesivo, si una curva elíptica $E/\mathbb{C}$ tiene multiplicación compleja, identificaremos sus endomorfismos con números complejos a través del isomorfismo

determinado por el teorema anterior.(Notemos que basta con que la condición se cumpla sobre una diferencial de primera clase no nula arbitraria.) El teorema 3.3 muestra que de este modo seguimos identificando cada entero n con la multiplicación por n en E.

Obviamente la existencia de multiplicaciones complejas se conserva por isomorfismos, luego podemos hablar de invariantes j con multiplicación compleja. Los ejemplos siguientes muestran que $j = 0$ y $j = 1728$ la tienen:

Ejemplo Consideremos una curva elíptica $E/\mathbb{C}$ con invariante $j = 1728$. En los ejemplos de la sección anterior hemos visto que $E/\mathbb{C}$ es analíticamente isomorfa al toro complejo $\mathbb{C}/R$ determinado por $R = \langle 1, i\rangle_{\mathbb{Z}} = \mathbb{Z}[i]$. Por consiguiente $\mathrm{End}(E) \cong \mathbb{Z}[i]$.

Un ejemplo concreto es la curva $Y^2 = X^3 - X$ considerada en el ejemplo de la página 79, donde llamamos i a la isogenia dada por $i(x, y) = (-x, iy)$. Vamos a comprobar que esta identificación está de acuerdo con el convenio que hemos adoptado. Para ello tomamos la diferencial de primera clase $\omega = dx/y$ y calculamos:

$$\overline{i}(\omega) = \frac{d(-x)}{iy} = i\omega.$$

■

Ejemplo Si ρ es una raíz cúbica primitiva de la unidad, el retículo $R = \langle 1, \rho\rangle_{\mathbb{Z}}$ es el anillo de enteros algebraicos de $\mathbb{Q}(\sqrt{-3}\,)$ y es el único orden cuadrático con 6 unidades, luego $\mathbb{C}/R$ tiene 6 automorfismos, y lo mismo le sucede a su curva elíptica asociada. Por el teorema 2.12 podemos concluir que el invariante de R es $j = 0$. En la sección anterior hemos obtenido esto mismo explícitamente.

Un ejemplo concreto lo proporciona la curva $Y^2 = X^3 + 1$, y es fácil ver que $\rho(x, y) = (\rho x, y)$. ■

Las isogenias entre curvas con las mismas multiplicaciones complejas resultan ser también "lineales" para éstas:

Teorema 10.22 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia entre dos curvas elípticas con el mismo anillo $\mathcal{O}$ de multiplicaciones complejas. Entonces, para todo $\alpha \in \mathcal{O}$ y todo $P \in E_1$ se cumple que $\phi(\alpha P) = \alpha\phi(P)$.*

No es cierto que dos curvas isógenas tengan necesariamente las mismas multiplicaciones complejas. No obstante, el teorema 3.36 muestra que dos curvas elípticas isógenas tienen ambas multiplicación compleja o ninguna la tiene. Más aún, en la prueba se ve que, en cualquier caso, tenemos el isomorfismo $\mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E_1) \cong \mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E_2)$. (Observemos que $\mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E)$ es simplemente el cuerpo de cocientes de $\mathrm{End}(E)$.) Ahora podemos probar el recíproco para curvas con multiplicación compleja:

Teorema 10.23 *Sean $E_1/\mathbb{C}$ y $E_2/\mathbb{C}$ dos curvas elípticas y supongamos que $E_1/\mathbb{C}$ tiene multiplicación compleja. Entonces las curvas son isógenas si y sólo si $\mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E_1) \cong \mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E_2)$.*

DEMOSTRACIÓN: Una implicación es el teorema 3.36. Si se da el isomorfismo, entonces ambas curvas tienen multiplicación compleja. Sean $\mathcal{O}_1$ y $\mathcal{O}_2$ los anillos respectivos de multiplicaciones complejas. Ambos son órdenes de un mismo cuerpo cuadrático K. Sea $\mathcal{O}$ su orden maximal y sea $E/\mathbb{C}$ una curva elíptica isomorfa a $\mathbb{C}/\mathcal{O}$. Entonces $E_i/\mathbb{C}$ es analíticamente isomorfa a un toro complejo $\mathbb{C}/R_i$ y es fácil encontrar un entero no nulo m tal que $mR_i \subset \mathcal{O}_i \subset \mathcal{O}$, luego existe un homomorfismo de $\mathbb{C}/R_i$ en $\mathbb{C}/\mathcal{O}$, luego las dos curvas E_i son ambas isógenas a E y, por consiguiente, son isógenas entre sí. ∎

Ahora también es claro que dos curvas isógenas no tienen necesariamente el mismo anillo de endomorfismos. Basta pensar en dos curvas isomorfas a dos toros asociados a dos órdenes de un mismo cuerpo cuadrático.

Si K es un cuerpo cuadrático complejo, un *módulo completo* en K es un $\mathbb{Z}$-módulo libre $R \subset K$ de rango 2. Observemos que esto es lo mismo que un retículo $R \subset K$, pues si $\mathcal{O}_K$ es el orden maximal de K, siempre es posible encontrar un entero no nulo m tal que $mR \subset \mathcal{O}_K$, el hecho de que $\mathcal{O}_K$ sea un retículo implica que mR también lo es, y por consiguiente R también.

Cada módulo completo R en K tiene asociado un *anillo de coeficientes*

$$\mathcal{O}_R = \{\alpha \in K \mid \alpha R \subset R\},$$

que resulta ser un orden en K. Es evidente entonces que el toro complejo $\mathbb{C}/R$ tiene a $\mathcal{O}_R$ como anillo de endomorfismos. El interés de esto se debe a que el conjunto de todos los módulos completos con un mismo anillo de coeficientes $\mathcal{O}$ se convierte en un grupo abeliano con el producto dado por

$$\langle \omega_1, \omega_2 \rangle \langle \eta_1, \eta_2 \rangle = \langle \omega_1\eta_1, \omega_2\eta_2, \omega_1\eta_2, \omega_2\eta_1 \rangle .$$

El neutro es $\mathcal{O}$. Además, los módulos de la forma $\alpha\mathcal{O}$, con $\alpha \in K^*$ forman un subgrupo, de modo que dos módulos R_1 y R_2 son congruentes módulo este subgrupo si y sólo si son similares, es decir, si y sólo si existe $\alpha \in K^*$ (equivalentemente, en $\mathbb{C}^*$) tal que $R_1 = \alpha R_2$.

Puesto que módulos similares determinan toros complejos isomorfos, concluimos que las clases de isomorfía de curvas elípticas con un anillo $\mathcal{O}$ de multiplicaciones complejas se corresponden con las clases de similitud de módulos con anillo de coeficientes $\mathcal{O}$. Ahora bien, este número es siempre finito y es lo que se conoce como *número de clases* de $\mathcal{O}$. En particular vemos que hay un número finito de clases de isomorfía de curvas elípticas con un mismo anillo de endomorfismos (no trivial).

De esta observación deducimos una propiedad de las curvas con multiplicación compleja:

Teorema 10.24 *Si una curva elíptica $E/\mathbb{C}$ tiene multiplicación compleja, entonces su invariante j es un número algebraico (luego E es isomorfa a una curva definida sobre un cuerpo numérico).*

DEMOSTRACIÓN: Consideremos una ecuación de Weierstrass para $E/\mathbb{C}$ y un automorfismo σ de $\mathbb{C}$. Al aplicarlo a los coeficientes de la ecuación obtenemos

una curva E^σ con invariante j^σ. Es claro que σ induce un isomorfismo entre los anillos de endomorfismos de ambas curvas, luego, según acabamos de ver, al variar σ tenemos que j^σ sólo puede recorrer un número finito de valores complejos. Esto implica que j es algebraico. ∎

Nota De la demostración anterior se deduce que si el anillo de multiplicaciones complejas de E es $\mathcal{O}$ y su número de clases es $h_\mathcal{O}$, entonces $|\mathbb{Q}(j) : \mathbb{Q}| \leq h_\mathcal{O}$. ∎

Observemos ahora que si α es una multiplicación compleja de una curva elíptica $E/\mathbb{C}$ y σ es un automorfismo de $\mathbb{C}$, entonces α^σ tiene dos interpretaciones distintas: la obvia y también el endomorfismo $f \in \mathrm{End}(E^\sigma)$ dado por $f(P) = (\alpha P^{\sigma^{-1}})^\sigma$. Veamos que en realidad son la misma, en el sentido de que este endomorfismo se corresponde con la multiplicación por α^σ.

Para ello usaremos que σ induce de forma natural un isomorfismo entre los espacios de diferenciales de primera clase de ambas curvas. Comprobar esto supone recorrer todos los pasos que conlleva la definición algebraica de las formas diferenciales: en primer lugar, σ induce un automorfismo del cuerpo de series formales de potencias $\mathbb{C}((X))$ que conserva las derivadas formales, luego induce un automorfismo del espacio de formas diferenciales sobre $\mathbb{C}((X))$ tal que $(\alpha\, d\beta)^\sigma = \alpha^\sigma\, d\beta^\sigma$. A continuación observamos que si $P \in E$, $f \in \mathcal{O}_P(E)$ y π es un parámetro local de E en P, entonces la serie de Taylor de f^σ en P^σ respecto a π^σ es la imagen por σ de la serie de Taylor de f en P respecto de π (porque la serie se define algebraicamente a partir de $\mathcal{O}_P(E)$, y σ induce un isomorfismo entre este anillo y $\mathcal{O}_{P^\sigma}(E^\sigma)$). De aquí se sigue que si f, $g \in \mathbb{C}(E)$, entonces $(f\, dg)^\sigma_P = (f^\sigma\, dg^\sigma)_{P^\sigma}$, lo que a su vez nos permite definir $(f\, dg)^\sigma = f^\sigma\, dg^\sigma$, y así tenemos un isomorfismo entre los espacios de formas diferenciales de ambas curvas, y es fácil ver que se restringe a un isomorfismo entre los espacios de formas diferenciales de primera clase.

Notemos también que si $\phi : E \longrightarrow E'$ es una isogenia, entonces

$$\overline{\phi^\sigma}(f^\sigma\, dg^\sigma) = (\phi^\sigma \circ f^\sigma)d(\phi^\sigma \circ g^\sigma) = (\phi \circ f)^\sigma d(\phi \circ g)^\sigma = (\overline{\phi}(f\, dg))^\sigma.$$

Tomemos una diferencial de primera clase ω en E. Entonces, por la propiedad que acabamos de probar,

$$\overline{\alpha^\sigma}(\omega^\sigma) = (\overline{\alpha}(\omega))^\sigma = (\alpha\omega)^\sigma = \alpha^\sigma\omega^\sigma.$$

Esto demuestra que la isogenia α^σ se corresponde con la multiplicación compleja α^σ. ∎

Como aplicación de este hecho podemos refinar el teorema 10.19:

Teorema 10.25 *Sea E/K una curva elíptica sobre un cuerpo $K \subset \mathbb{C}$ y con multiplicaciones complejas en un cuerpo cuadrático L. Entonces todo endomorfismo de E está definido sobre el cuerpo KL.*

DEMOSTRACIÓN: Los endomorfismos de E son multiplicaciones por números complejos α, enteros o imaginarios. Si σ es un automorfismo que fija a KL, entonces $\alpha^\sigma = \alpha$ no sólo en el sentido trivial, sino también viendo a α como endomorfismo, luego α está definido sobre KL (teorema 1.17). ∎

La dualidad entre isogenias tiene una interpretación muy simple en términos de toros complejos. Dado un homomorfismo analítico $\alpha : \mathbb{C}/R_1 \longrightarrow \mathbb{C}/R_2$ entre dos toros complejos, existe un único homomorfismo $\hat{\alpha} : \mathbb{C}/R_2 \longrightarrow \mathbb{C}/R_1$ tal que $\alpha\hat{\alpha} = \operatorname{grad}\alpha$ (donde $\operatorname{grad}\alpha$ es el orden del núcleo de α). En efecto, la única posibilidad es $\hat{\alpha} = (\operatorname{grad}\alpha)\alpha^{-1}$.

Es fácil comprobar que $\hat{\alpha}$ así definido es realmente un homomorfismo: para cada $z \in R_2$, tenemos que $\alpha^{-1}z$ está en el núcleo de α, luego $(\operatorname{grad}\alpha)\alpha^{-1}z \in R_1$. Esto significa que $\hat{\alpha}R_2 \subset R_1$.

Si α es una multiplicación compleja de una curva elíptica, entonces el teorema 3.30 nos da que $\alpha + \hat{\alpha} \in \mathbb{Z}$, al igual que $\alpha\hat{\alpha}$, luego α y $\hat{\alpha}$ son raíces de un mismo polinomio cuadrático con coeficientes enteros, luego $\hat{\alpha}$ es el conjugado complejo de α. Esto es trivialmente cierto si α es entero. En particular vemos que $\operatorname{grad}\alpha = |\alpha|^2$.

10.4 Funciones modulares asociadas

Hasta aquí hemos asociado varios números complejos G_{2k}, g_2, g_3, Δ y j, a cada retículo complejo R. Con propiedad deberíamos escribir $G_{2k}(R)$, $g_2(R)$, $g_3(R)$, etc. Si nos restringimos a retículos de la forma $R = \langle 1, \tau\rangle_{\mathbb{Z}}$, estas funciones pueden verse como funciones de τ, y así nos encontramos con los ejemplos básicos de lo que en el capítulo XII llamaremos funciones modulares.

La restricción a retículos $\langle 1, \tau\rangle_{\mathbb{Z}}$ no supone una pérdida de generalidad porque —según sabemos— todo retículo complejo es linealmente equivalente a uno de esta forma. Con más detalle, podemos empezar por expresar R en función de una base. Así, si ω_1 y ω_2 son números complejos tales que $\tau = \omega_2/\omega_1$ cumple $\operatorname{Im}\tau > 0$, definimos $G_{2k}(\omega_1, \omega_2) = G_{2k}(\langle\omega_1, \omega_2\rangle_{\mathbb{Z}})$, $g_2(\omega_1, \omega_2) = g_2(\langle\omega_1, \omega_2\rangle_{\mathbb{Z}})$, etc. Se cumplen las siguientes relaciones de homogeneidad:

$$G_{2k}(\alpha\omega_1, \alpha\omega_2) = \alpha^{-2k}G_{2k}(\omega_1, \omega_2) \qquad \text{para todo } \alpha \in \mathbb{C}^*.$$

Esto se sigue inmediatamente de la definición de las series de Eisenstein, que en términos de una base es

$$G_{2k}(\omega_1, \omega_2) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(m\omega_1 + n\omega_2)^{2k}}.$$

De aquí se deducen a su vez las relaciones

$$g_2(\alpha\omega_1, \alpha\omega_2) = \alpha^{-4}g_2(\omega_1, \omega_2), \quad g_3(\alpha\omega_1, \alpha\omega_2) = \alpha^{-6}g_3(\omega_1, \omega_2),$$

$$\Delta(\alpha\omega_1, \alpha\omega_2) = \alpha^{-12}\Delta(\omega_1, \omega_2), \quad j(\alpha\omega_1, \alpha\omega_2) = j(\omega_1, \omega_2).$$

En particular tenemos que $G_{2k}(\omega_1,\omega_2) = \omega_1^{-2k} G_{2k}(1,\tau)$, de donde a su vez

$$g_2(\omega_1,\omega_2) = \omega_1^{-4} g_2(1,\tau), \quad g_3(\omega_1,\omega_2) = \omega_1^{-6} g_3(1,\tau),$$

$$\Delta(\omega_1,\omega_2) = \omega_1^{-12}\Delta(1,\tau), \quad j(\omega_1,\omega_2) = j(1,\tau).$$

Es por esto que no perdemos generalidad si nos limitamos a estudiar las funciones $G_{2k}(\tau) = G_{2k}(1,\tau)$, $g_2(\tau) = g_2(1,\tau)$, etc., definidas sobre el semiplano

$$H = \{\tau \in \mathbb{C} \mid \operatorname{Im}\tau > 0\}.$$

Como ya hemos dicho, todas ellas son funciones modulares, si bien no introduciremos la noción general de función modular hasta el capítulo XII. La función $j : H \longrightarrow \mathbb{C}$ se conoce como *función modular de Klein,* y tenemos pendiente demostrar que toma todos los valores complejos, pues ello se traduce en que toda curva elíptica compleja es analíticamente isomorfa a un toro complejo. Más aún, si admitimos este hecho, podemos demostrar que existen retículos complejos cuyos invariantes g_2 y g_3 toman cualquier par de valores prefijados (siempre y cuando no den lugar a un discriminante nulo):

Teorema 10.26 *Si c_2 y c_3 son números complejos tales que $c_2^3 - 27c_3^2 \neq 0$, entonces existen números complejos ω_1, ω_2 linealmente independientes sobre $\mathbb{R}$ tales que $g_2(\omega_1,\omega_2) = c_2$ y $g_3(\omega_1,\omega_2) = c_3$.*

Demostración: Vamos a admitir que existe un número complejo $\tau \in H$ tal que $j(\tau) = j(1,\tau) = 1728c_2^3/(c_2^3 - 27c_3^2)$ (teorema 12.3).

Si $c_2 = 0$ entonces $j(1,\tau) = 0$, luego $g_2(1,\tau) = 0$ y $g_3(1,\tau) \neq 0$. Sea $\alpha \in \mathbb{C}$ tal que $\alpha^{-6} g_3(1,\tau) = c_3$. Tomamos $\omega_1 = \alpha$, $\omega_2 = \alpha\tau$, de modo que $g_2(\omega_1,\omega_2) = \alpha^{-4} g_2(1,\tau) = 0 = c_2$ y $g_3(\omega_1,\omega_2) = \alpha^{-6} g_3(1,\tau) = c_3$.

Si $c_2 \neq 0$ entonces $j(1,\tau) \neq 0$, luego $g_2(1,\tau) \neq 0$. Tomamos $\alpha \in \mathbb{C}$ tal que $\alpha^{-4} g_2(1,\tau) = c_2$ y $\omega_1 = \alpha$, $\omega_2 = \alpha\tau$. Entonces $g_2(\omega_1,\omega_2) = \alpha^{-4} g_2(1,\tau) = c_2$. Por otra parte,

$$\frac{1728c_2^3}{c_2^3 - 27c_3^2} = j(1,\tau) = j(\omega_1,\omega_2) = \frac{1728c_2^3}{c_2^3 - 27g_3^2(\omega_1,\omega_2)}.$$

Por lo tanto $g_3^2(\omega_1,\omega_2) = c_3^2$ y $g_3(\omega_1,\omega_2) = \pm c_3$. Cambiando ω_1 y ω_2 por $-\omega_1$, $-\omega_2$ no alteramos g_2 pero cambiamos el signo a g_3, con lo que podemos garantizar la igualdad $g_3(\omega_1,\omega_2) = c_3$. ∎

Lo primero que podemos probar de las funciones $G_{2k}(\tau)$, etc. es que son holomorfas. Ello es consecuencia inmediata del siguiente teorema general:

Teorema 10.27 *Para cada $\alpha > 2$, la función $f_\alpha : H \longrightarrow \mathbb{C}$ dada por*

$$f_\alpha(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(m+n\tau)^\alpha}$$

es holomorfa en H.

DEMOSTRACIÓN: Basta ver que la serie converge absoluta y uniformemente en cada conjunto de la forma

$$S = \{x + yi \mid |x| \le A, \quad y \ge \delta\}, \qquad A > 0, \quad \delta > 0.$$

Por el criterio de mayoración de Weierstrass y teniendo en cuenta 10.9, basta ver que existe una constante $M > 0$ tal que

$$\frac{1}{|m + n\tau|^\alpha} \le \frac{M}{|m + ni|^\alpha},$$

para todo $\tau \in S$ y todo $(m, n) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$. A su vez, basta ver que existe una constante K tal que $|m + n\tau|^2 > K|m + ni|^2$ para todo $\tau \in S$. (Si $T^2 > K$, entonces $T^\alpha > K^{\alpha/2}$.) Si $\tau = x + iy$, esto equivale a

$$(m + nx)^2 + (ny)^2 > K(m^2 + n^2).$$

Si $n = 0$ esto se cumple siempre que $0 < K < 1$. Si $n \neq 0$ llamamos $q = m/n$, con lo que la desigualdad equivale a

$$\frac{(q + x)^2 + y^2}{1 + q^2} > K.$$

Veamos que esto se cumple para todo q con

$$K = \frac{\delta^2}{1 + (A + \delta)^2}.$$

Si $|q| \le A + \delta$ es trivial, pues $(q + x)^2 \ge 0$ e $y^2 \ge \delta^2$. Si $|q| > A + \delta$, entonces $|x/q| < |x|/(A + \delta) \le A/(A + \delta) < 1$, luego

$$\left|1 + \frac{x}{q}\right| \ge 1 - \left|\frac{x}{q}\right| > 1 - \frac{A}{A + \delta} = \frac{\delta}{A + \delta}.$$

Por consiguiente, $|q + x| \ge \dfrac{q\delta}{A + \delta}$ y

$$\frac{(q + x)^2 + y^2}{1 + q^2} > \frac{\delta^2}{(A + \delta)^2} \frac{q^2}{1 + q^2} \ge \frac{\delta^2}{(A + \delta)^2} \frac{(A + \delta)^2}{1 + (A + \delta)^2} = K,$$

donde hemos usado que la función $t/(1 + t)$ es creciente. ∎

Como consecuencia inmediata:

Teorema 10.28 *Las funciones $G_{2k}(\tau)$ (para $k \ge 2$), $g_2(\tau)$, $g_3(\tau)$, $\Delta(\tau)$ y $j(\tau)$ son holomorfas en el semiplano* $\operatorname{Im}\tau > 0$.

Al tratar con las funciones del teorema anterior, podemos identificar cada punto $\tau \in H$ con el retículo complejo $\langle 1, \tau\rangle_{\mathbb{Z}}$, pero hemos de tener presente que,

según el teorema 10.7, dos retículos $\langle 1,\tau\rangle_{\mathbb{Z}}$ y $\langle 1,\tau'\rangle_{\mathbb{Z}}$ son linealmente equivalentes si y sólo si existen números enteros a, b, c, d tales que

$$\tau' = \frac{a\tau+b}{c\tau+d}, \qquad ad-bc=1.$$

Por ejemplo, teniendo en cuenta que dos retículos linealmente equivalentes tienen el mismo invariante j, de aquí deducimos la siguiente propiedad fundamental de la función modular de Klein: Si a, b, c, $d\in\mathbb{Z}$ cumplen $ad-bc=1$, entonces

$$j\left(\frac{a\tau+b}{c\tau+d}\right) = j(\tau).$$

Para las demás funciones tenemos una relación similar, aunque algo más elaborada. Teniendo en cuenta que $\langle 1,\tau\rangle_{\mathbb{Z}} = \langle c\tau+d, a\tau+b\rangle_{\mathbb{Z}}$, vemos que

$$G_{2k}\left(\frac{a\tau+b}{c\tau+d}\right) = G_{2k}\left(1,\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{2k}G_{2k}(c\tau+d,a\tau+b)$$

$$= (c\tau+d)^{2k}G_{2k}(1,\tau) = (c\tau+d)^{2k}G_{2k}(\tau). \tag{10.3}$$

Como consecuencia:

$$g_2\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^4 g_2(\tau), \quad g_3\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^6 g_3(\tau),$$

$$\Delta\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{12}\Delta(\tau).$$

En particular, tomando $a=b=d=1$, $c=0$, vemos que $j(\tau+1)=j(\tau)$, $G_{2k}(\tau+1)=G_{2k}(\tau)$, etc., es decir, que todas las funciones que estamos considerando tienen periodo 1. Esto nos permite desarrollarlas en serie de Fourier, en el sentido del teorema siguiente:

Teorema 10.29 *Sea $f : H_r \longrightarrow \mathbb{C}$ una función holomorfa de periodo 1 en el semiplano*

$$H_r = \{\tau\in\mathbb{C} \mid \operatorname{Im}\tau > r\}.$$

Entonces f admite un único desarrollo en serie de la forma

$$f(\tau) = \sum_{n=-\infty}^{+\infty} a_n e^{2n\pi i\tau}, \qquad a_n\in\mathbb{C},$$

que converge uniformemente en cada subconjunto compacto de H_r.

Demostración: Llamemos $D = D(0,e^{-2\pi r})$. Basta observar que la función $q : H_r \longrightarrow D$ dada por $q(\tau) = e^{2\pi i\tau}$ tiene periodo 1 y transforma cada recta horizontal de H_r en una circunferencia. Además q es localmente conforme, luego podemos definir una función holomorfa $f^* : D\setminus\{0\} \longrightarrow \mathbb{C}$ de manera que $f(\tau) = f^*(e^{2\pi i\tau})$, para todo $\tau\in H_r$. La función $f^*(z)$ admite un desarrollo en serie de Laurent en D, que da lugar al desarrollo del enunciado al sustituir

$z = q(\tau)$. La unicidad es consecuencia inmediata de la unicidad del desarrollo de Laurent. ∎

El desarrollo del teorema anterior se conoce como *desarrollo en serie de Fourier* de la función f. Vamos a calcular los desarrollos correspondientes a las funciones que estamos estudiando. Partimos de la factorización del seno:

$$\operatorname{sen} \pi z = \pi z \prod_{m=1}^{\infty} \left(1 - \frac{z}{m}\right)\left(1 + \frac{z}{m}\right).$$

La teoría de los productos infinitos nos da que, para $z \notin \mathbb{Z}$, un logaritmo de la función $(\operatorname{sen} \pi z)/(\pi z)$ viene dado por

$$\log \frac{\operatorname{sen} \pi z}{\pi z} = \sum_{m=1}^{\infty} \log\left(1 - \frac{z^2}{m^2}\right),$$

donde la serie es absolutamente convergente y el logaritmo es cualquier rama uniforme del logaritmo (por ejemplo, si la tomamos en $\mathbb{C}$ menos el semieje real negativo, el logaritmo está definido para $\operatorname{Im} z > 0$). Derivando llegamos a que

$$\pi \frac{\cos \pi z}{\operatorname{sen} \pi z} = \frac{1}{z} + \sum_{m=1}^{\infty} \left(\frac{1}{z-m} + \frac{1}{z+m}\right).$$

De la relación entre las funciones trigonométricas y la exponencial se sigue que

$$\cos \pi z = \frac{1}{2} e^{-i\pi z}(e^{2\pi i z} + 1), \qquad \operatorname{sen} \pi z = \frac{1}{2i} e^{-i\pi z}(e^{2\pi i z} - 1).$$

Por consiguiente, si $\operatorname{Im} z > 0$,

$$\pi \frac{\cos \pi z}{\operatorname{sen} \pi z} = \pi i \, \frac{e^{2\pi i z} + 1}{e^{2\pi i z} - 1} = \pi i + \frac{2\pi i}{e^{2\pi i z} - 1} = \pi i - 2\pi i \sum_{m=0}^{\infty} e^{2m\pi i z}.$$

Así pues,

$$\frac{1}{z} + \sum_{m=1}^{\infty} \left(\frac{1}{z-m} + \frac{1}{z+m}\right) = \pi i - 2\pi i \sum_{m=0}^{\infty} e^{2m\pi i z}.$$

Derivando $k-1$ veces:

$$(-1)^{k-1}(k-1)! \left(\frac{1}{z^k} + \sum_{m=1}^{\infty} \left(\frac{1}{(z-m)^k} + \frac{1}{(z+m)^k}\right)\right) \tag{10.4}$$

$$= -(2\pi i)^k \sum_{m=1}^{\infty} m^{k-1} e^{2m\pi i z}.$$

Si cambiamos k por $2k$ y z por $n\tau$, con $n \in \mathbb{Z}$, $n \neq 0$, la serie del miembro izquierdo se descompone en suma de dos series absolutamente convergentes (pues son subseries de $G_{2k}(\tau)$), que a su vez se pueden agrupar en la forma:

$$(2k-1)! \sum_{m=-\infty}^{+\infty} \frac{1}{(m+n\tau)^{2k}} = (2\pi i)^{2k} \sum_{m=1}^{\infty} m^{2k-1} e^{2mn\pi i \tau}.$$

(En principio tenemos la igualdad para τ suficientemente pequeño, pero ambos miembros son funciones holomorfas en H, luego la igualdad vale para todo $\tau \in H$.)

Así pues, separando en $G_{2k}(\tau)$ los sumandos correspondientes a $n = 0$ de los correspondientes a $n \neq 0$ obtenemos que

$$G_{2k}(\tau) = 2\zeta(2k) + 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(2\pi i)^{2k}m^{2k-1}}{(2k-1)!}e^{2mn\pi i\tau},$$

donde $\zeta(\tau)$ es la función dseta de Riemann. Si llamamos

$$\sigma_k(n) = \sum_{d|n} d^k,$$

tenemos probado el teorema siguiente:

Teorema 10.30 *Si $k > 1$ es un número natural, la serie de Eisenstein*

$$G_{2k}(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}^2\\(m,n)\neq(0,0)}} \frac{1}{(m+n\tau)^{2k}}$$

admite el desarrollo en serie de Fourier

$$G_{2k}(\tau) = 2\zeta(2k) + 2\frac{(2\pi i)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)e^{2n\pi i\tau}.$$

En particular, teniendo en cuenta que

$$\zeta(4) = \frac{\pi^4}{90}, \qquad \zeta(6) = \frac{\pi^6}{945},$$

vemos que

$$g_2(\tau) = 60G_4(\tau) = \frac{(2\pi)^4}{2^2\cdot 3}\left(1 + 240\sum_{n=1}^{\infty}\sigma_3(n)e^{2n\pi i\tau}\right),$$

$$g_3(\tau) = 140G_6(\tau) = \frac{(2\pi)^6}{2^3\cdot 3^3}\left(1 - 504\sum_{n=1}^{\infty}\sigma_5(n)e^{2n\pi i\tau}\right).$$

De aquí deducimos a su vez:

Teorema 10.31 *La función $\Delta(\tau)$ tiene un desarrollo en serie de Fourier en el semiplano H de la forma*

$$\Delta(\tau) = (2\pi)^{12}\sum_{n=1}^{\infty}\tau(n)e^{2n\pi i\tau},$$

donde la función $\tau(n)$ (que no hay que confundir con la variable τ) toma valores enteros y $\tau(1) = 1$.

DEMOSTRACIÓN: Llamemos

$$A = \sum_{n=1}^{\infty} \sigma_3(n) e^{2n\pi i\tau}, \qquad B = \sum_{n=1}^{\infty} \sigma_5(n) e^{2n\pi i\tau}.$$

Entonces

$$\Delta(\tau) = g_2^3(\tau) - 27g_3^2(\tau) = \frac{64\pi^{12}}{27}((1+240A)^3 - (1-504B)^2).$$

Notemos que

$$(1+240A)^3-(1-504B)^2 = 1+720A+3\cdot 240^2A^2+240^3A^3-1+1.008B-504^2B^2$$

$$= 12^2(5A+7B) + 12^3(100A^2 - 147B^2 + 8.000A^3).$$

Además, para todo entero d se cumple que

$$5d^3 + 7d^5 = d^3(5+7d^2) \equiv \begin{cases} d^3(d^2-1) \equiv 0 \text{ (mód 3)}, \\ d^3(1-d^2) \equiv 0 \text{ (mód 4)}, \end{cases}$$

y al sumar sobre todos los divisores de $n \in \mathbb{Z}$ obtenemos que $12 \mid 5\sigma_3(n)+7\sigma_5(n)$, por lo que 12 divide a todos los coeficientes de la serie $5A+7B$ y 12^3 divide a todos los coeficientes de $(1+240A)^3 - (1-504B)^2$. Así pues,

$$\Delta(\tau) = \frac{64\pi^{12}}{27}\, 12^3 \sum_{n=1}^{\infty} \tau(n) e^{2n\pi i\tau} = (2\pi)^{12} \sum_{n=1}^{\infty} \tau(n) e^{2n\pi i\tau},$$

donde los coeficientes $\tau(n)$ son enteros.

El coeficiente de $e^{2\pi i\tau}$ tanto en A como en B es $\sigma_3(1) = \sigma_5(1) = 1$, con lo que el coeficiente de $e^{2\pi i\tau}$ en $(1+240A)^3 - (1-504B)^2$ es $12^2(5+7) = 12^3$. Concluimos que $\tau(1) = 1$. ∎

La función $\tau(n)$ se conoce como *función tau de Ramanujan.* Sus primeros valores son

$$\tau(1) = 1, \quad \tau(2) = -24, \quad \tau(3) = 252, \quad \tau(4) = -1472, \quad \tau(5) = 4.830,$$

$$\tau(6) = 6.048, \quad \tau(7) = -16.744, \quad \tau(8) = 84.480, \quad \tau(9) = -113.643,$$

$$\tau(10) = -115.920.$$

La función τ tiene muchas propiedades aritméticas que no son fáciles de probar, muchas de las cuales fueron conjeturadas por Ramanujan. Por ejemplo, es multiplicativa y satisface varias congruencias del estilo de

$$\tau(n) \equiv \sigma_{11}(n) \text{ (mód 691)}.$$

Nos ocuparemos de estos hechos en el capítulo XII, donde daremos también una fórmula sencilla para calcular la función. (El carácter multiplicativo de la función τ lo probaremos en el apéndice B.) Ahora pasamos a la función de Klein:

Teorema 10.32 *La función de Klein j tiene un desarrollo en serie de Fourier en el semiplano H de la forma*

$$j(\tau) = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} c(n)e^{2n\pi i\tau},$$

donde los coeficientes $c(n)$ son enteros.

DEMOSTRACIÓN: Convenimos que E, E', etc. representan series de Fourier

$$E = \sum_{n=0}^{\infty} a_n e^{2\pi ni\tau}$$

con coeficientes enteros. Entonces

$$g_2^3(\tau) = \frac{64}{27}\pi^{12}(1+E)^3 = \frac{64}{27}\pi^{12}(1+E'),$$

$$\Delta(\tau) = \frac{64}{27}\pi^{12}\,12^3 e^{2\pi i\tau}(1+E'').$$

Por consiguiente:

$$j(\tau) = \frac{12^3 g_2^3(\tau)}{\Delta(\tau)} = \frac{1+E'}{e^{2\pi i\tau}(1+E'')}.$$

Ahora bien, si $E'' = 1 + a_1 e^{2\pi i\tau} + a_2 e^{4\pi i\tau} + \cdots$, entonces la serie de potencias $1 + a_1 z + a_2 z^2 + \cdots$ converge en un entorno de 0 a una función holomorfa f tal que $f(0) = 1$, luego $1/f$ es holomorfa en un entorno de 0 y su desarrollo en serie de potencias $1/f = 1 + b_1 z + b_2 z^2 + \cdots$ satisface la identidad

$$(1 + a_1 z + a_2 z^2 + \cdots)(1 + b_1 z + b_2 z^2 + \cdots) = 1.$$

Esto nos da las igualdades $a_1 + b_1 = 0$, $a_2 + a_1 b_1 + b_2 = 0$, ... de las que se sigue que los coeficientes b_n son enteros, con lo que

$$\frac{1}{1+E''} = 1 + E''',$$

para cierta serie de Fourier E''' con coeficientes enteros. En definitiva,

$$j(\tau) = \frac{(1+E')(1+E''')}{e^{2\pi i\tau}} = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} c(n)e^{2n\pi i\tau}.$$

■

Los primeros valores de la función $c(n)$ son:

$$c(0) = 744, \quad c(1) = 196.884, \quad c(2) = 21.493.760, \quad c(3) = 864.299.970,$$

$$c(4) = 20.245.856.256, \quad c(5) = 333.202.640.600, \quad c(6) = 4.252.023.300.096.$$

Vemos que los coeficientes crecen muy rápidamente. Esto no es casual. Por ejemplo, $c(2) = 2^{11} \cdot 5 \cdot 2.099$, y el factor 2^{11} es obligado, ya que la función $c(n)$ satisface muchas congruencias, tales como

$$\begin{aligned} c(2^m n) &\equiv 0 \ (\text{mód } 2^{3m+8}), \\ c(3^m n) &\equiv 0 \ (\text{mód } 3^{2m+3}), \\ c(5^m n) &\equiv 0 \ (\text{mód } 5^{m+1}), \\ c(7^m n) &\equiv 0 \ (\text{mód } 7^m). \end{aligned} \tag{10.5}$$

Nos ocuparemos de estos hechos en el capítulo XII.

Finalmente, vamos a probar que las funciones $e_1(\tau)$, $e_2(\tau)$ y $e_3(\tau)$ (ver 10.15) son holomorfas en H. Llamemos $\wp(z;\tau)$ a la función $\wp$ del retículo $\langle 1, \tau\rangle_{\mathbb{Z}}$. Entonces,

$$\wp(z;\tau) = \frac{1}{z^2} + \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \left(\frac{1}{(z - m\tau - n)^2} - \frac{1}{(m\tau + n)^2} \right).$$

Desarrollamos esta expresión usando (10.4) para $k = 2$, es decir:

$$\sum_{n=-\infty}^{+\infty} \frac{1}{(z+n)^2} = (2\pi i)^2 \sum_{n=1}^{\infty} n e^{2n\pi i z} = (2\pi i)^2 \frac{e^{2\pi i z}}{(1 - e^{2\pi i z})^2}. \tag{10.6}$$

(Esto es válido si $\operatorname{Im} z > 0$.)

En la expresión de $\wp(z;\tau)$ separamos la suma para $m = 0$ de la suma para $m \neq 0$ y queda

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

Aplicamos (10.6) a z, $z + m\tau$ y $-z + m\tau$. Para que las partes imaginarias sean positivas se ha de cumplir $0 < \operatorname{Im} z < \operatorname{Im} \tau$:

$$\wp(z;\tau) = (2\pi i)^2 \frac{e^{2\pi i z}}{(1 - e^{2\pi i z})^2} - \frac{\pi^2}{3}$$

$$+(2\pi i)^2 \sum_{m=1}^{+\infty} \sum_{n=1}^{\infty} \left(n e^{2n\pi i(z+m\tau)} + n e^{2n\pi i(-z+m\tau)} - 2n e^{2nm\pi i\tau} \right).$$

Equivalentemente:

$$\frac{1}{(2\pi i)^2} \wp(z;\tau) = \frac{1}{12} + \frac{e^{2\pi i z}}{(1 - e^{2\pi i z})^2} + \sum_{m=1}^{+\infty} \sum_{n=1}^{\infty} n e^{2\pi i n m \tau} (e^{2n\pi i z} + e^{-2n\pi i z} - 2).$$

Teorema 10.33 *Las funciones*

$$e_1(\tau) = \wp(1/2;\tau), \qquad e_2(\tau) = \wp(\tau/2;\tau), \qquad e_3(\tau) = \wp((1+\tau)/2;\tau)$$

son holomorfas en H.

DEMOSTRACIÓN: Si hacemos $z = \tau/2$ en la fórmula que acabamos de obtener, podemos agrupar los sumatorios como una serie de Laurent $f(w)$, donde $w = e^{i\pi\tau}$, y sabemos que $f(w)$ es convergente si $\tau \in H$, es decir, si $w \in D(0,1) \setminus \{0\}$. Puesto que una serie de Laurent convergente determina siempre una función holomorfa, concluimos que $e_2(\tau) = \wp(\tau/2;\tau)$ es holomorfa en H. Lo mismo vale para $e_3(\tau)$. En cambio, no podemos aplicar el mismo razonamiento a $e_1(\tau)$, pues ahora sería $z = 1/2$, cuando necesitamos $\operatorname{Im} z > 0$. No obstante, $e_1(\tau) = -e_2(\tau) - e_3(\tau)$, luego también es holomorfa. ∎

10.5 El grupo modular

En esta última sección estudiaremos un caso particular que servirá como motivación para una teoría mucho más general que trataremos en los capítulos siguientes. Hemos visto que la función modular de Klein satisface $j(\tau) = j(\tau')$ si y sólo si

$$\tau' = \frac{a\tau + b}{c\tau + d},$$

para ciertos enteros a, b, c, d tales que $ad - bc = 1$. Se define el *grupo modular* como el grupo Γ formado por todas las matrices 2×2 con coeficientes enteros y determinante 1. A cada matriz

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

le podemos asociar la transformación de Möbius $\sigma_A : \mathbb{C}^\infty \longrightarrow \mathbb{C}^\infty$ dada por

$$\sigma_A(\tau) = \frac{a\tau + b}{c\tau + d}.$$

Si identificamos $\mathbb{C}^\infty = \mathrm{P}^1(\mathbb{C})$, vemos que

$$\sigma_A([u,v]) = \sigma_A([u/v,1]) = [a(u/v)+b, c(u/v)+d] = [au+bv, cu+dv] = [(u,v)A],$$

de donde se sigue inmediatamente que la aplicación $A \mapsto \sigma_A$ es un homomorfismo de Γ en el grupo de las transformaciones conformes de $\mathbb{C}^\infty$ en sí mismo. Se comprueba inmediatamente que su núcleo está formado por las matrices $\pm I$ (donde I es la identidad), de modo que podemos identificar los elementos de Γ con las transformaciones de Möbius que inducen, teniendo presente que dos matrices A y B determinan la misma transformación si y sólo si $A = \pm B$.

Para transformaciones $\sigma \in \Gamma$, la fórmula (10.1) se reduce a

$$\operatorname{Im}\sigma(\tau) = \frac{\operatorname{Im}\tau}{|c\tau + d|^2}, \qquad \text{donde } \sigma = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \tag{10.7}$$

de donde concluimos que todos los elementos de Γ se restringen a transformaciones conformes del semiplano H en sí mismo.

Ahora podemos decir que la función j coincide sobre dos puntos τ, $\tau' \in H$ si y sólo si existe un $\sigma \in \Gamma$ tal que $\tau' = \sigma(\tau)$. En tal caso diremos que τ y τ' son *equivalentes* respecto de Γ.

Conviene observar aquí una analogía con las funciones elípticas: Si f es elíptica respecto de un retículo R, la aplicación que a cada $\omega \in R$ le asocia la traslación $\sigma_\omega(\tau) = \tau+\omega$ es un monomorfismo de R en el grupo de las traslaciones del plano. Si llamamos G a su imagen, tenemos que f toma el mismo valor sobre cada par de puntos τ, τ' equivalentes respecto de G, es decir, tales que existe un $\sigma \in G$ tal que $\tau' = \sigma(\tau)$. Un paralelogramo fundamental P de R es un subconjunto del plano con la propiedad de que todo punto es equivalente por G a un único punto de P. Si lo tomamos cerrado, la única excepción es que cada punto de la frontera de P es equivalente a su opuesto. Vamos a probar que el conjunto

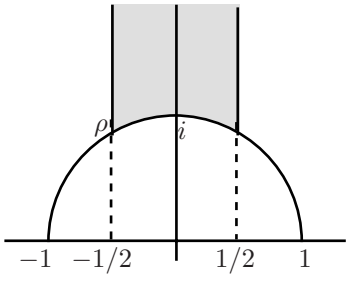
$$D = \{\tau \in H \mid |\operatorname{Re}\tau| \le 1/2, \quad |\tau| \ge 1\}$$

cumple la misma propiedad para el grupo Γ (esto se expresa diciendo que D es un *dominio fundamental* de Γ):

Teorema 10.34 *Todo punto $\tau \in H$ es equivalente respecto al grupo modular Γ a un punto de D, y dos puntos τ, $\tau' \in D$ no son equivalentes entre sí excepto si están en la frontera y cumplen $\operatorname{Re}\tau = \pm\operatorname{Re}\tau'$, $\operatorname{Im}\tau = \operatorname{Im}\tau'$. Además, el grupo modular está generado por las matrices*

$$t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad y \quad s = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(cuyas transformaciones asociadas son $t(\tau) = \tau + 1$ y $s(\tau) = -1/\tau$).



Demostración: El dominio D es la parte sombreada en la figura. El número ρ es la raíz cúbica de la unidad

$$\rho = \frac{-1+\sqrt{-3}}{2}.$$

Llamemos Γ' al subgrupo de Γ generado por s y t. Notemos que $-1 = s^2 \in \Gamma'$. Fijado $\tau \in \mathbb{C}$, los números de la forma $c\tau + d$ con c, $d \in \mathbb{Z}$ forman un retículo complejo, que es un subconjunto cerrado y discreto de $\mathbb{C}$, luego el conjunto

$$\{|c\tau + d| \mid (c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}\}$$

tiene un mínimo elemento. La fórmula (10.7) muestra entonces que podemos tomar $z \in H$ equivalente a τ respecto a Γ' con parte imaginaria maximal. Aplicando varias veces la traslación t o bien t^{-1} no alteramos la parte imaginaria (y pasamos a puntos equivalentes respecto a Γ'), luego podemos suponer que z

cumple $|\operatorname{Re} z| \leq 1/2$. Si fuera $|z| < 1$ entonces $s(z)$ tendría parte imaginaria mayor, luego ha de ser $|z| \geq 1$.

Con esto hemos probado que todo punto de H es equivalente a un punto de D respecto a Γ' . Ahora veamos que dos z, $z' \in D$ no son equivalentes respecto de Γ salvo que estén en la frontera y sean simétricos respecto al eje imaginario. Ciertamente, un par de puntos simétricos en estas condiciones son equivalentes (a través de t si $\operatorname{Re} z = \pm 1/2$ o a través de s si $|z| = 1$).

Pongamos que $\operatorname{Im} z \leq \operatorname{Im} z'$ y que $z' = g(z)$, con $g \in \Gamma$ dada por

$$g(\tau) = \frac{a\tau + b}{c\tau + d}.$$

Cambiando los signos podemos suponer que $c \geq 0$. La fórmula (10.7) nos da que $|cz + d| \leq 1$, pero $\operatorname{Im} z \geq \sqrt{3}/2$, luego ha de ser $c \leq 2/\sqrt{3}$, luego $c = 0$, 1.

Si $c = 0$ entonces $ad = 1$, podemos suponer $a = d = 1$ y así $g(\tau) = \tau + b$, con lo que $g = t^b$. Esto ya implica que $\operatorname{Im} z = \operatorname{Im} z'$, y además $|\operatorname{Re} z - \operatorname{Re} z'| = |b|$, lo cual sólo es posible si $b = \pm 1$ y z, z' tienen partes reales de $\pm 1/2$.

Si $c = 1$, la relación $|z + d| \leq 1$ con $|z| \geq 1$ sólo puede darse si $d = -1$, 0, 1.

Si $d = 0$ ha de ser $b = -1$, con lo que $|z| = 1$ y $g(\tau) = a - 1/\tau$. Vemos que $-1/z$ tiene también módulo 1. Para que $z' = a - 1/z$ esté en D, las únicas posibilidades son que $a = 0$ (y entonces z y z' son simétricos respecto al eje imaginario) o bien que $a = \pm 1$ (y entonces z y z' han de ser ρ y $s(\rho)$). En cualquiera de los dos casos se cumple lo que queremos probar.

Si $d = \pm 1$, la relación $|z + d| \leq 1$ sólo puede darse si $z = \rho$ o bien $z = s(\rho)$. Si es $z = \rho$ entonces $d = 1$ y

$$g(\tau) = \frac{a\tau + a - 1}{\tau + 1} = a - \frac{1}{\tau + 1}.$$

Por lo tanto $z' = a + \rho$. Necesariamente $a = 1$ y se cumple lo que queremos probar. El caso $z = s(\rho)$ es análogo.

Sólo queda probar $\Gamma = \Gamma'$. En efecto, si $g \in \Gamma$, para cada z en el interior de D, acabamos de probar que $g(z)$ es equivalente a un $z' \in D$ respecto a Γ', luego existe $g' \in \Gamma'$ tal que $g'(g(z)) = z'$. Como z no está en la frontera de D, sabemos que esto sólo puede ocurrir si $g'(g(z)) = z$. Por el principio de prolongación analítica ha de ser $g \circ g' = 1$, luego $g = g'^{-1} \in \Gamma'$. (Notemos que, viendo a g y g' como matrices, lo que hemos probado es que $g = \pm g'^{-1}$ pero, como $-1 \in \Gamma'$, la conclusión es correcta.) ∎

Si identificamos los lados opuestos de un paralelogramo fundamental de un retículo R obtenemos una superficie de Riemann T de género 1 tal que las funciones elípticas sobre R pueden verse como funciones holomorfas sobre T. En el capítulo siguiente veremos que podemos hacer lo mismo con la función modular: No es evidente, pero si identificamos los puntos de D simétricos respecto al eje imaginario obtenemos una superficie de Riemann S, de modo que j puede verse como una función holomorfa sobre S. Intuitivamente está claro que S es topológicamente un cilindro con un extremo cerrado o, también, una esfera menos un punto. Sucede además que si compactificamos S con un punto infinito

obtenemos una esfera S^* y resulta que j se extiende a una función holomorfa $j : S^* \longrightarrow \mathbb{C}^\infty$. Más concretamente, veremos que los coeficientes de Laurent de j respecto a una carta adecuada alrededor del punto infinito son precisamente sus coeficientes de Fourier, y el hecho de que el coeficiente de $e^{-2\pi i\tau}$ sea no nulo se traducirá en que j tiene un polo simple en el infinito. Sabemos que j es inyectiva en S, luego, admitiendo todo esto, es claro que j ha de ser una transformación conforme entre dos esferas, lo que nos permite concluir que j toma todos los valores complejos y, según ya hemos explicado, esto demuestra el teorema de uniformización.

Capítulo XI

Superficies modulares

En el capítulo anterior hemos visto que las propiedades básicas de las funciones elípticas complejas son consecuencia de que pueden verse como funciones meromorfas sobre una superficie de Riemann compacta (un toro complejo). Así mismo, nos han aparecido una serie de funciones holomorfas asociadas $\Delta(\tau)$, $j(\tau)$, etc., definidas sobre el semiplano

$$H = \{\tau \in \mathbb{C} \mid \operatorname{Im}\tau > 0\}.$$

Son ejemplos de lo que en el capítulo siguiente llamaremos funciones modulares, que están muy relacionadas con ciertas superficies compactas llamadas superficies modulares, pero que no son tan fáciles de construir como los toros complejos, sino que para ello vamos a necesitar todo este capítulo.

En la última sección del capítulo anterior esbozamos la construcción de la más simple de las superficies modulares: la que se obtiene al identificar los puntos de H equivalentes respecto al grupo modular. Las superficies modulares en general se obtienen del mismo modo a partir de grupos de transformaciones adecuados. Dedicamos la primera sección a introducir los grupos con los que vamos a trabajar.

11.1 Transformaciones de Möbius

La teoría de funciones de variable compleja prueba que las únicas transformaciones conformes de la esfera de Riemann en sí misma son las llamadas *transformaciones de Möbius,* es decir, las transformaciones de la forma

$$z \mapsto \frac{az+b}{cz+d}, \qquad a,b,c,z \in \mathbb{C}, \quad ad-bc \neq 0.$$

Si pensamos en la esfera de Riemann como en la recta proyectiva compleja, las transformaciones de Möbius pueden verse también como las transformaciones proyectivas de la recta en sí misma. El interés de este enfoque es que nos da un aparato algebraico para manejarlas:

En general, si A es un anillo, el *grupo lineal general* $\mathrm{LG}(n, A)$ es el grupo de todas las matrices regulares $n \times n$ con coeficientes en A. La regularidad de una matriz equivale a que su determinante sea una unidad de A. Si A es un cuerpo, esto es tanto como decir que el determinante ha de ser no nulo.

A cada matriz

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LG}(2, \mathbb{C})$$

le hacemos corresponder la transformación de Möbius $\phi_A : \mathbb{C}^\infty \longrightarrow \mathbb{C}^\infty$ dada por

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

Si consideramos coordenadas homogéneas en $\mathbb{C}^\infty$ (de modo que cada $z \in \mathbb{Z}$ se identifica con $(z, 1)$ y el punto infinito es $\infty = (1, 0)$) la aplicación ϕ_A es

$$\phi_A([z_1, z_2]) = [(z_1, z_2)A].$$

Visto así, es claro que ϕ constituye un epimorfismo de $\mathrm{LG}(2, \mathbb{C})$ en el grupo de las transformaciones de Möbius. Su núcleo lo forman las matrices diagonales.

Por otra parte, la teoría de funciones de variable compleja también prueba que toda transformación conforme de un semiplano en sí mismo es la restricción de una transformación de Möbius.[1]

Vamos a probar que las únicas transformaciones conformes del semiplano H en sí mismo son las asociadas al grupo $\mathrm{LG}_+(2, \mathbb{R})$, es decir, las transformaciones de Möbius con coeficientes reales y determinante positivo.

Ciertamente, las transformaciones de este tipo conservan H. Ello es debido a la relación (10.1), que podemos reescribir como sigue:

$$\operatorname{Im} \phi_A(\tau) = \frac{\det A}{|c\tau + d|^2} \operatorname{Im} \tau, \quad \text{para } A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LG}(2, \mathbb{R}).$$

Por otra parte, si f es una transformación de Möbius que deja invariante a H, claramente f ha de transformar la recta real en sí misma (incluyendo en ella el punto ∞). Supongamos que $f(\infty) = r \in \mathbb{R}$. Entonces tomamos

$$B = \begin{pmatrix} 0 & 1 \\ -1 & -r \end{pmatrix} \in \mathrm{LG}_+(2, \mathbb{R}),$$

de modo que $\phi_B(r) = \infty$ y $f \circ \phi_B(\infty) = \infty$. Si probamos que $f \circ \phi_B$ tiene la forma indicada, lo mismo valdrá para f. Equivalentemente, podemos suponer que $f(\infty) = \infty$. Esto se traduce en que $f = \phi_A$ para una matriz A con $c = 0$ o, equivalentemente, en que $f(\tau) = r\tau + s$, donde necesariamente r, $s \in \mathbb{R}$ (pues f transforma $\mathbb{R}$ en $\mathbb{R}$).

[1] Este resultado se enuncia más habitualmente para discos, pero cualquier disco puede transformarse en cualquier semiplano mediante una transformación de Möbius, luego es válido igualmente para semiplanos.

Finalmente notamos que multiplicando una matriz de $\mathrm{LG}_+(2,\mathbb{R})$ por una matriz diagonal adecuada obtenemos otra que induce la misma transformación de Möbius y tiene determinante 1.

Si A es un anillo, definimos el *grupo lineal especial* $\mathrm{LE}(n,A)$ como el subgrupo de $\mathrm{LG}(n,A)$ formado por las matrices de determinante 1.

Acabamos de justificar que toda transformación conforme de H en sí mismo puede obtenerse a partir de una matriz de $\mathrm{LE}(2,\mathbb{R})$. Por ello, a partir de este momento sólo trataremos con matrices de este grupo. Notemos que dos matrices determinan la misma transformación de Möbius si y sólo si se diferencian a lo sumo en un factor ± 1. Equivalentemente, el grupo de las transformaciones conformes de H es isomorfo al cociente $\mathrm{LE}(2,\mathbb{R})/\{\pm I\}$.

Nota Observemos que la representación matricial en $\mathrm{LE}(2,\mathbb{R})$ de una transformación de H no siempre es la obvia. Por ejemplo, la representación de $f(\tau) = 2\tau$ no es

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \text{sino} \qquad \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{pmatrix}.$$

En lo sucesivo no distinguiremos entre las matrices de $\mathrm{LE}(2,\mathbb{R})$ y sus transformaciones de Möbius asociadas, si bien hemos de tener siempre presente que los pares de matrices $\pm A$ se identifican como transformaciones de Möbius.

Recapitulando, si

$$f = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LE}(2,\mathbb{R}),$$

tenemos que f tiene un único polo en $\tau = -d/c$, que es un número real, o bien el punto ∞. Para cualquier otro $\tau \in \mathbb{C}$ se cumple la relación (10.7). Esto implica que f se restringe a una transformación conforme en el semiplano H.

Nota La teoría que vamos a desarrollar tiene una interpretación geométrica basada en el hecho de que el semiplano H es un modelo de la geometría hiperbólica (no euclídea) cuando se consideran como rectas hiperbólicas las rectas verticales y las semicircunferencias ortogonales al eje real. En H pueden definirse las nociones de distancia y medida de ángulos (en particularidad la perpendicularidad), de modo que se cumplen los axiomas básicos de la geometría euclídea excepto el postulado de las paralelas. Las transformaciones asociadas a las matrices de $\mathrm{LE}(2,\mathbb{R})$ son precisamente los movimientos hiperbólicos, es decir, las biyecciones de H en sí mismo que conservan las distancias y la orientación (y por consiguiente la medida de ángulos).

Conviene señalar también que, mediante una transformación de Möbius que haga corresponder el semiplano H con el disco unitario Δ, es posible desarrollar toda la teoría en Δ en lugar de en H. Este enfoque es más intuitivo porque desaparece la asimetría entre el punto ∞ y los puntos del eje real (que, aunque no forman parte del plano hiperbólico, sí intervienen en las definiciones de recta, movimiento, etc. a modo de "puntos infinitos"). Sin embargo, dicha asimetría resulta ser técnicamente muy conveniente, pues, como veremos, el punto ∞

resulta ser mucho más cómodo de manejar y a menudo será posible reducir las cuestiones concernientes a puntos finitos al caso especial de ∞. ∎

Las superficies modulares que queremos definir se obtendrán identificando los puntos de H a través de ciertos subgrupos de $\mathrm{LE}(2,\mathbb{R})$, análogamente a como los toros complejos se obtienen identificando los puntos de $\mathbb{C}$ a través de ciertos subgrupos de traslaciones. Como ya hemos comentado, el caso más simple lo proporciona el grupo modular, que con la notación que acabamos de introducir no es sino $\Gamma = \mathrm{LE}(2,\mathbb{Z})$. Las transformaciones de Möbius asociadas a los elementos de Γ se conocen como *transformaciones unimodulares.*

Diremos que dos puntos de H son *equivalentes* bajo un subgrupo Γ de $\mathrm{LE}(2,\mathbb{R})$ si uno puede transformarse en el otro mediante un elemento de Γ. Es obvio que se trata de una relación de equivalencia en H. El análogo al paralelogramo fundamental de un retículo es el concepto que introducimos seguidamente:

Un subconjunto cerrado $D \subset H$ es un *dominio fundamental* para Γ si todo punto de H es equivalente a un punto de D y dos puntos de D no son equivalentes salvo quizá si ambos están en la frontera.

Ciertamente, los paralelogramos fundamentales de un retículo complejo R cumplen esta definición respecto al grupo de traslaciones asociadas a R. El teorema 10.34 nos da que

$$D = \{\tau \in H \mid |\operatorname{Re}\tau| \leq 1/2, \quad |\tau| \geq 1\}$$

es un dominio fundamental para el grupo modular. Desde el punto de vista de la geometría hiperbólica se trata de un triángulo (de lados rectilíneos) no acotado.

Nuestro propósito es dotar de estructura de superficie de Riemann al cociente de H sobre la equivalencia respecto de un subgrupo Γ de $\mathrm{LE}(2,\mathbb{R})$. El subgrupo no puede ser arbitrario, pues el cociente podría reducirse a un punto, o no ser un espacio de Hausdorff, etc. Esto nos lleva a las consideraciones topológicas de la sección siguiente.

11.2 Grupos topológicos

Definición 11.1 Un *grupo topológico* es un grupo G dotado de estructura de espacio topológico (de Hausdorff) de modo que el producto $G \times G \longrightarrow G$ y la aplicación $G \longrightarrow G$ dada por $\alpha \mapsto \alpha^{-1}$ son continuas.

Ejemplos El grupo $\mathrm{LG}(2,\mathbb{C})$ puede identificarse con un subconjunto de $\mathbb{C}^4$ de forma natural:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LG}(2,\mathbb{C}) \leftrightarrow (a,b,c,d) \in \mathbb{C}^4.$$

La imagen la forman las cuádruplas $(a,b,c,d) \in \mathbb{C}^4$ tales que $ad - bc = 1$, luego es un abierto. A través de esta biyección podemos dotar a $\mathrm{LG}(2,\mathbb{C})$ de

una topología con la que se convierte en un grupo topológico, pues el producto se corresponde con una aplicación polinómica y la aplicación $\alpha \mapsto \alpha^{-1}$ con una función racional cuyo denominador no se anula. Claramente $\mathrm{LG}(2,\mathbb{C})$ es localmente compacto y tiene una base numerable. ∎

Es obvio que todo subgrupo de un grupo topológico hereda una estructura natural de grupo topológico. En particular, $\mathrm{LE}(2,\mathbb{R})$ es un subgrupo cerrado de $\mathrm{LG}(2,\mathbb{C})$ (su imagen en $\mathbb{C}^4$ es un subconjunto cerrado de $\mathbb{R}^4$). ∎

Un subgrupo de $\mathrm{LE}(2,\mathbb{R})$ que nos va a interesar especialmente es el *grupo ortogonal especial* $\mathrm{OE}(2,\mathbb{R})$, formado por las matrices de $\mathrm{LE}(2,\mathbb{R})$ cuya inversa coincide con su traspuesta. Explícitamente, es fácil ver que $\mathrm{OE}(2,\mathbb{R})$ está formado por las matrices

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

tales que $a^2 + b^2 = 1$. Obviamente $\mathbb{R}$ es un grupo topológico con la suma, y la aplicación $\mathbb{R} \longrightarrow \mathrm{OE}(2,\mathbb{R})$ dada por

$$t \mapsto \begin{pmatrix} \cos 2\pi t & \operatorname{sen} 2\pi t \\ -\operatorname{sen} 2\pi t & \cos 2\pi t \end{pmatrix}$$

es un homomorfismo de grupos continuo y suprayectivo. Su núcleo es $\mathbb{Z}$. ∎

A través del homeomorfismo de $\mathrm{LG}(2,\mathbb{C})$ en $\mathbb{C}^4$ que define la topología del primer grupo, el grupo modular $\mathrm{LE}(2,\mathbb{Z})$ se corresponde con un subconjunto de $\mathbb{Z}^4$, luego $\mathrm{LE}(2,\mathbb{Z})$ es un subgrupo cerrado y discreto de $\mathrm{LE}(2,\mathbb{R})$. ∎

El grupo T de las traslaciones del plano complejo puede identificarse con el subgrupo de $\mathrm{LE}(2,\mathbb{C})$ compuesto por las matrices de la forma

$$t_b = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}, \qquad b \in \mathbb{C}.$$

La aplicación $b \mapsto t_b$ define obviamente un isomorfismo topológico $\mathbb{C} \cong T$ (es decir, un isomorfismo de grupos que además es un homeomorfismo). ∎

Definición 11.2 Una *acción (continua)* de un grupo topológico G sobre un espacio topológico S es una aplicación continua $S \times G \longrightarrow S$ tal que $s1 = s$ para todo $s \in S$ y $(s\alpha)\beta = s(\alpha\beta)$ para todo $s \in S$ y todos los α, $\beta \in G$.

En estas condiciones, es claro que la multiplicación por un $\alpha \in G$ fijo es un homeomorfismo de S en sí mismo.

Si un grupo G actúa sobre un conjunto S, diremos que dos elementos $s, t \in S$ son *equivalentes* respecto a G si existe $\alpha \in G$ tal que $s\alpha = t$. Claramente se trata de una relación de equivalencia en S. Las clases de equivalencia se llaman *órbitas* y el conjunto de todas las órbitas se representa por S/G. Esto generaliza a las definiciones análogas que hemos dado en la sección anterior. La acción es *transitiva* si todos los puntos de S son equivalentes entre sí, es decir, si S/G consta de una única órbita.

Ejemplos Es fácil ver que la acción $\mathbb{C}^\infty \times \mathrm{LG}(2,\mathbb{C}) \longrightarrow \mathbb{C}^\infty$ descrita en la sección anterior es continua, si bien no vamos a necesitar este hecho general. Es más fácil comprobar directamente la continuidad de la acción que realmente nos va a interesar, que es la acción $H \times \mathrm{LE}(2,\mathbb{R}) \longrightarrow H$. Identificando el grupo con un subconjunto de $\mathbb{R}^4$, está determinada por funciones racionales cuyos denominadores no se anulan. La acción es transitiva, pues para todo $a+bi \in H$, la matriz

$$\alpha = \begin{pmatrix} \sqrt{b} & 0 \\ a/\sqrt{b} & 1/\sqrt{b} \end{pmatrix}$$

cumple que $i\alpha = a+bi$, luego la órbita de i es todo el semiplano. ∎

También es continua la acción $\mathbb{C} \times T \longrightarrow \mathbb{C}$ del grupo de las traslaciones del plano complejo (identificando T con $\mathbb{C}$, se trata de la suma de números complejos). La acción es obviamente transitiva. Si R es un retículo complejo, a través del isomorfismo $\mathbb{C} \cong T$ se identifica con un subgrupo cerrado y discreto de T, de modo que el cociente $\mathbb{C}/R$ que acabamos de definir es el toro complejo definido por R en el sentido usual. ∎

Si G es un grupo topológico y K es un subgrupo de G, entonces K actúa sobre G por multiplicación (por la derecha). El conjunto cociente es G/K (en el sentido usual de la teoría de grupos). ∎

Definición 11.3 Si G es un grupo topológico que actúa sobre un espacio S y $p : S \longrightarrow S/G$ es la proyección natural en el cociente, definimos en S/G la topología respecto a la cual un conjunto $A \subset S/G$ es abierto si y sólo si $p^{-1}[A]$ es abierto en S. Nos referiremos a ella como la *topología cociente.*

Es claro que la topología cociente es ciertamente una topología en S/G respecto a la cual p es continua. Más aún, p es abierta, pues si $X \subset S$ es abierto, entonces $p[X]$ es abierto en S/G, ya que $p^{-1}[p[X]] = \bigcup_{g\in G} Xg$ es unión de abiertos.

En el caso en que G es un retículo complejo R, la topología en $\mathbb{C}/R$ que acabamos de definir es la topología usual en el toro complejo. Un retículo complejo es un subgrupo discreto de $\mathbb{C}$ (que podemos identificar con un subgrupo discreto del grupo de las traslaciones del plano). Nuestro objetivo es obtener una construcción similar a la de los toros complejos a partir de un cociente del semiplano H respecto de un subgrupo discreto Γ de $\mathrm{LE}(2,\mathbb{R})$. Si no exigimos que Γ sea discreto podemos encontrarnos muchas patologías. Por ejemplo, si Γ actúa transitivamente sobre H el cociente se reduce a un punto, o si una órbita tiene un punto de acumulación en otra órbita (en particular si es densa en H) entonces el cociente no tiene la propiedad de Hausdorff, etc.

En esta sección demostraremos que si Γ es discreto entonces H/Γ es un espacio de Hausdorff, y más adelante veremos que podemos dotarlo de estructura analítica. Como primera observación sobre los subgrupos discretos demostramos que son siempre cerrados:

Teorema 11.4 *Sea G un grupo topológico y Γ un subgrupo localmente compacto. Entonces Γ es cerrado en G.*

DEMOSTRACIÓN: Sea C un entorno compacto de 1 en Γ. Sea U un entorno abierto de 1 en G tal que $U \cap \Gamma \subset C$. Tomemos $x \in \overline{\Gamma}$. Consideremos la aplicación $f(\alpha, \beta) = \alpha^{-1}\beta$. Como $f(x, x) = 1 \in U$, tenemos que $f^{-1}[U]$ contiene un entorno $V \times V$ de (x, x), es decir, V es un entorno de x tal que $V^{-1}V \subset U$. Entonces $(V \cap \Gamma)^{-1}(V \cap \Gamma) \subset C$. Como x está en la clausura de Γ, podemos tomar $y \in V \cap \Gamma$. Claramente $V \cap \Gamma \subset yC$. Notemos que si W es un entorno de x, entonces $W \cap V \cap \Gamma \neq \varnothing$, luego $x \in \overline{V \cap \Gamma} \subset yC \subset \Gamma$ (pues yC es compacto). Concluimos que Γ es cerrado. ∎

Antes de estudiar los cocientes que realmente nos interesan vamos a obtener un par de resultados sobre cocientes de un grupo topológico respecto de un subgrupo cerrado:

Teorema 11.5 *Sea G un grupo topológico y K un subgrupo cerrado de G. Entonces la topología cociente en G/K es de Hausdorff. Si K es un subgrupo normal, entonces G/K es un grupo topológico con dicha topología.*

DEMOSTRACIÓN: La aplicación $f : G \times G \longrightarrow G$ dada por $f(\alpha, \beta) = \alpha^{-1}\beta$ es continua. Si $p(\alpha), p(\beta) \in G/K$ son clases distintas, entonces $(\alpha, \beta) \notin f^{-1}[K]$, luego existe un abierto $U \times V$ en $G \times G$ tal que

$$(\alpha, \beta) \in U \times V, \qquad (U \times V) \cap f^{-1}[K] = \varnothing.$$

Los conjuntos $p[U]$ y $p[V]$ son entornos de $p(\alpha)$ y $p(\beta)$ respectivamente, y es fácil ver que son disjuntos.

Supongamos ahora que K es normal y sea U un abierto en G/K. Consideramos el diagrama siguiente:

$$\begin{array}{ccc} G \times G & \xrightarrow{f} & G \\ {\scriptstyle p\times p}\downarrow & & \downarrow{\scriptstyle p} \\ G/K \times G/K & \xrightarrow[g]{} & G/K \end{array}$$

Las flechas horizontales son el producto en G y G/K. Es fácil ver que $g^{-1}[U] = (p \times p)[f^{-1}[p^{-1}[[U]]]$, con lo que $g^{-1}[U]$ es abierto y g es continua. Similarmente se comprueba la continuidad de $K\alpha \mapsto K\alpha^{-1}$. ∎

Si K es un subgrupo cerrado de G, no necesariamente normal, el conjunto cociente G/K considerado en el teorema anterior está formado por las clases gK generadas por multiplicación por la derecha por elementos de K. Es claro que podemos definir igualmente una topología de Hausdorff sobre el conjunto de las clases Kg correspondientes a la multiplicación por la izquierda. Vamos a usar este hecho en el contexto siguiente:

Si un grupo topológico G actúa transitivamente sobre un espacio S, fijado un $s \in S$, podemos considerar el *estabilizador* $K = \{\alpha \in G \mid s\alpha = s\}$, que es un subgrupo cerrado de G. Si llamamos G/K al espacio de clases Kg, podemos definir la biyección $G/K \longrightarrow S$ dada por $Kg \mapsto sg$.

Ciertamente es continua, pues si U es abierto en S, su antiimagen en G/K es $p[f_s^{-1}[U]]$, donde $f_s : G \longrightarrow S$ es la aplicación continua determinada por $f_s(g) = sg$. Sin embargo, en general no podemos garantizar que la biyección sea un homeomorfismo. El teorema siguiente nos da una condición suficiente.

Teorema 11.6 *Sea G un grupo topológico con una base numerable que actúa transitivamente sobre un espacio topológico S localmente compacto. Sea $s \in S$ y $K = \{\alpha \in G \mid s\alpha = s\}$. Sea $G/K = \{Kg \mid g \in G\}$. Entonces la aplicación $G/K \longrightarrow S$ dada por $Kg \mapsto sg$ es un homeomorfismo.*

Demostración: Ya sabemos que es continua. Basta probar que es abierta. Un abierto en G/K es de la forma $p[U]$, donde U es un abierto en G. Hemos de ver que su imagen en S, es decir, $sU = \{sg \mid g \in U\}$ es abierta en S. Fijamos $g \in U$ y hemos de ver que sg es un punto interior de sU.

Tenemos que Ug^{-1} es un entorno de 1. Su antiimagen por el producto en G contiene un entorno $V \times V$ de $(1, 1)$. Así, V es un entorno de 1 con la propiedad de que $VV \subset Ug^{-1}$. Puesto que V^{-1} también es un entorno de 1, cambiando V por $V \cap V^{-1}$, podemos suponer que $V = V^{-1}$.

Si sV contiene un punto interior sv, entonces $sg = svv^{-1}g$ será un punto interior de $sVVg \subset U$, como queremos probar. Así pues, basta ver que sV no tiene interior vacío.

Obviamente, $G = \bigcup_{\alpha \in G} V\alpha$. Como G tiene una base numerable, podemos encontrar $\{\alpha_n\}_{n=0}^{\infty} \subset G$ de modo que $G = \bigcup_{n=0}^{\infty} V\alpha_n$. Entonces $S = \bigcup_{n=0}^{\infty} sV\alpha_n$.

Por el teorema de Baire, algún $sV\alpha_n$ ha de tener interior no vacío, luego sV también tiene interior no vacío (ya que la multiplicación por α_n es un homeomorfismo). ∎

En particular esto se aplica al teorema de isomorfía: si $f : G \longrightarrow H$ es un epimorfismo continuo entre dos grupos topológicos, podemos considerar la acción de G en H dada por $hg = hf(g)$. Claramente es transitiva, pues la órbita de 1 es la imagen de G. El estabilizador de 1 es el núcleo $\mathrm{N}(f)$ y la aplicación considerada en el teorema anterior es el isomorfismo $G/\mathrm{N}(f) \cong H$, luego, suponiendo que G tenga una base numerable y que H sea localmente compacto, el isomorfismo que proporciona el teorema de isomorfía es topológico (es decir, es también un homeomorfismo).

Ejemplos La aplicación $\mathbb{R} \longrightarrow S^1$ dada por $t \mapsto e^{2\pi it}$ es un epimorfismo continuo de $\mathbb{R}$ en la esfera unidad compleja cuyo núcleo es $\mathbb{Z}$. Por consiguiente, el grupo $\mathbb{R}/\mathbb{Z}$ es topológicamente isomorfo a S^1. ∎

Antes hemos construido también un epimorfismo continuo $\mathbb{R} \longrightarrow \mathrm{OE}(2, \mathbb{R})$ con núcleo $\mathbb{Z}$. Concluimos que el grupo ortogonal especial es topológicamente

isomorfo a $\mathbb{R}/\mathbb{Z}$, y también a la circunferencia unidad S^1. En particular $\mathrm{OE}(2,\mathbb{R})$ es un grupo topológico compacto. ∎

Consideremos ahora la acción de $G = \mathrm{LE}(2,\mathbb{R})$ sobre el semiplano H, que ya hemos visto que es transitiva. Vamos a calcular el estabilizador de i. Para ello observamos que una matriz

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

cumple $i\alpha = i$ si y sólo si $a = d$ y $b = -c$.

Concluimos que el estabilizador de i es $K = \mathrm{OE}(2,\mathbb{R})$. El teorema anterior prueba que H es homeomorfo al cociente $\mathrm{LE}(2,\mathbb{R})/\mathrm{OE}(2,\mathbb{R})$. Más aún, si Γ es un subgrupo de G, podemos considerar su acción natural en H inducida por la de G y su acción en G/K dada por $(Kg)\gamma = K(g\gamma)$. Claramente tenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccc} G/K \times \Gamma & \longrightarrow & G/K \\ \downarrow & & \downarrow \\ H \times \Gamma & \longrightarrow & H \end{array}$$

Vemos así que es equivalente estudiar la acción de Γ sobre H que estudiar su acción sobre G/K. ∎

El mismo planteamiento es aplicable trivialmente a la acción del grupo G de las traslaciones del plano complejo sobre $S = \mathbb{C}$. Esta vez los estabilizadores son triviales, luego, tomando $K = 1$, vemos que la acción de un subgrupo $\Gamma \subset G$ sobre $\mathbb{C}$ es equivalente a la acción de Γ sobre G/K (o sobre G) por multiplicación por la derecha. ∎

A partir de aquí supondremos que G es un grupo topológico localmente compacto, K un subgrupo compacto de G y Γ un subgrupo discreto de G. Llamaremos $S = G/K$ y $p : G \longrightarrow S$ será la proyección natural. Por el teorema 11.5 tenemos que S es un espacio de Hausdorff. Como p es abierta, de hecho S es localmente compacto. Según acabamos de ver, tomando $G = \mathrm{LE}(2,\mathbb{R})$ y $K = \mathrm{OE}(2,\mathbb{R})$, todo cuanto digamos se aplicará en particular a la acción de Γ sobre el semiplano H (y también a la acción de un retículo sobre el plano complejo, aunque lo que sucede en este caso ya lo sabemos).

Teorema 11.7 *Si $A \subset S$ es un conjunto compacto, entonces $p^{-1}[A]$ es compacto.*

Demostración: Tomemos un cubrimiento de G formado por abiertos de clausura compacta. Sus imágenes por p son abiertos en S y cubren el conjunto A. Así pues,

$$A \subset \bigcup_{i=1}^{n} p[V_i],$$

donde cada V_i es un abierto en G de clausura compacta. Por lo tanto,

$$p^{-1}[A] \subset \bigcup_{i=1}^{n} KV_i \subset \bigcup_{i=1}^{n} K\overline{V}_i.$$

El conjunto $p^{-1}[A]$ es cerrado y cada $K\overline{V}_i$ es compacto, luego $P^{-1}[A]$ es compacto. ∎

Teorema 11.8 *Si A y B son subconjuntos compactos de S, entonces el conjunto $\{g \in \Gamma \mid Ag \cap B \neq \varnothing\}$ es finito.*

Demostración: Sean $C = p^{-1}[A]$ y $D = p^{-1}[B]$. Por el teorema anterior, C y D son subconjuntos compactos de G. Si $Ag \cap B \neq \varnothing$, entonces $Cg \cap D \neq \varnothing$, luego $g \in \Gamma \cap (C^{-1}D)$. Ahora bien, $\Gamma \cap (C^{-1}D)$ es a la vez compacto y discreto, luego es finito. ∎

En particular, tomando $A = B = \{z\}$, vemos que, para todo $s \in S$, el conjunto $\{g \in \Gamma \mid sg = s\}$ es finito.

Teorema 11.9 *Para cada $s \in S$, existe un entorno U de s tal que*

$$\{g \in \Gamma \mid Ug \cap U \neq \varnothing\} = \{g \in \Gamma \mid sg = s\}.$$

Demostración: Sea V un entorno compacto de z. Por el teorema anterior

$$\{g \in \Gamma \mid Vg \cap V \neq \varnothing\} = \{g_1, \ldots, g_r\}.$$

Podemos suponer que $sg_i = s$ exactamente para $1 \leq i \leq r'$. Para cada $i > r'$ podemos tomar entornos disjuntos V_i de s y W_i de sg_i. Basta tomar

$$U = V \cap \bigcap_{i>r'} (V_i \cap W_i g_i^{-1}).$$

∎

Teorema 11.10 *Si s_1, $s_2 \in S$ no son equivalentes respecto de Γ, entonces existen entornos U_i de s_i tales que $U_1 g \cap U_2 = \varnothing$ para todo $g \in \Gamma$.*

Demostración: Sean X_1 y X_2 entornos compactos de s_1 y s_2. Por 11.8 sabemos que

$$\{g \in \Gamma \mid X_1 g \cap X_2 \neq \varnothing\} = \{g_1, \ldots, g_r\}.$$

Como s_1 y s_2 no son equivalentes, ha de ser $s_1 g_i \neq s_2$, luego podemos encontrar entornos disjuntos U_{i1} y U_{i2} de $s_1 g_i$ y s_2 respectivamente. Basta tomar

$$U_1 = X_1 \cap U_{11} g_1^{-1} \cap \cdots \cap U_{r1} g_r^{-1}, \quad U_2 = X_2 \cap U_{12} \cap \cdots \cap U_{r2}.$$

∎

Ahora es claro que S/Γ es un espacio de Hausdorff. En particular hemos demostrado lo que necesitábamos:

Teorema 11.11 *Si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, entonces el conjunto cociente H/Γ es un espacio de Hausdorff.*

Ejercicio: Demostrar que si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, entonces las órbitas de los puntos de H respecto de Γ son discretas.

11.3 Puntos elípticos y parabólicos

Si R es un retículo complejo, el cociente $\mathbb{C}/R$ es compacto y la proyección natural $p : \mathbb{C} \longrightarrow \mathbb{C}/R$ se restringe a un homeomorfismo en un entorno de cada punto. Esta última propiedad permite definir fácilmente la estructura analítica del toro, pero ninguno de estos dos hechos tiene por qué ser cierto para un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$. Basta pensar en el grupo modular $\Gamma = \mathrm{LE}(2,\mathbb{Z})$. El teorema 10.34 muestra que todo entorno de i contiene infinitos pares de puntos simétricos con la misma imagen en H/Γ. Los puntos como éste son excepcionales, pero tenemos que caracterizarlos y tratarlos aparte en muchas ocasiones. Por otro lado, ya hemos comentado que el cociente H/Γ no es compacto, sino que necesitamos compactificarlo con un punto para obtener una esfera.

En general, si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, en esta sección estudiaremos los puntos donde la proyección $p : H \longrightarrow H/\Gamma$ no es localmente inyectiva (a los que llamaremos puntos elípticos), al tiempo que veremos cómo introducir puntos adicionales que compactifiquen el cociente (a los que llamaremos puntos parabólicos). Ambas clases de puntos surgen de forma natural a partir de una clasificación sencilla de los elementos de $\mathrm{LE}(2,\mathbb{R})$:

Si $g \in \mathrm{LE}(2,\mathbb{R})$, sus puntos fijos en $\mathbb{C}^\infty$ (es decir, los de su transformación de Möbius asociada) son sus vectores propios (los vectores propios de g como matriz) interpretados como coordenadas homogéneas de la recta proyectiva. Si g tiene un espacio de vectores propios de dimensión 2, entonces $g = \pm 1$ y su transformación de Möbius asociada es la identidad. Descartando este caso, las posibilidades son que g tenga uno o dos espacios de vectores propios de dimensión 1 (según si su polinomio característico tiene una raíz doble o simple). Por consiguiente, g tiene uno o dos puntos fijos en $\mathbb{C}^\infty$. El hecho de que los coeficientes de g sean reales implica que si $z \in \mathbb{C}$ es un punto fijo, también lo es su conjugado $\bar{z}$. Con esto llegamos a la clasificación siguiente:

Definición 11.12 Sea $g \in \mathrm{LE}(2,\mathbb{R})$ una transformación $g \neq \pm 1$.

a) Diremos que g es *parabólica* si tiene un único punto fijo en $\mathbb{R} \cup \{\infty\}$ (que será, de hecho, su único punto fijo en $\mathbb{C}^\infty$).

b) Diremos que g es *elíptica* si no tiene puntos fijos en $\mathbb{R} \cup \{\infty\}$ (en cuyo caso tendrá dos puntos fijos, un $z \in H$ y su conjugado $\bar{z} \notin H$).

c) Diremos que g es *hiperbólica* si tiene dos puntos fijos en $\mathbb{R} \cup \{\infty\}$ (que serán, de hecho, todos sus puntos fijos en $\mathbb{C}^\infty$).

La discusión precedente muestra que cada transformación $g \in \mathrm{LE}(2,\mathbb{R})$ distinta de ± 1 es de uno de estos tres tipos. El teorema siguiente permite determinar fácilmente el carácter de una matriz:

Teorema 11.13 *Si $g \in \mathrm{LE}(1,\mathbb{R})$, $g \neq \pm 1$, entonces*

a) g es parabólica si y sólo si $|\operatorname{Tr}(g)| = 2$,

b) g es elíptica si y sólo si $|\operatorname{Tr}(g)| < 2$,

c) g es hiperbólica si y sólo si $|\operatorname{Tr}(g)| > 2$.

DEMOSTRACIÓN: El polinomio característico de g es $x^2 - \operatorname{Tr}(g)x + 1$, cuyo discriminante es $\operatorname{Tr}(g)^2 - 4$. Es claro que g es parabólica si y sólo si tiene un único valor propio, lo cual equivale a que $\operatorname{Tr}(g) = \pm 2$. Si $|\operatorname{Tr}(g)| > 2$ entonces g tiene dos valores propios reales, luego tiene dos vectores propios reales y, por consiguiente, es hiperbólica. Igualmente se tiene el recíproco. Por exclusión tenemos también la caracterización de las matrices elípticas. ∎

Nota Esta clasificación tiene una interpretación en términos de la geometría hiperbólica. Recordemos que las transformaciones de $\mathrm{LE}(2,\mathbb{R})$ son los movimientos hiperbólicos. De entre ellos, las transformaciones hiperbólicas (movimientos sin puntos fijos) son el equivalente no euclídeo de las traslaciones, las transformaciones elípticas (movimientos con un punto fijo) son el equivalente no euclídeo de los giros, mientras que las transformaciones parabólicas son un tipo de movimientos sin equivalente euclídeo, llamados también giros infinitos (giros con centro en el infinito).

Consideremos por ejemplo los generadores del grupo modular: la transformación $t(z) = z + 1$ es una traslación euclídea, pero desde el punto de vista de la geometría hiperbólica es un giro alrededor de ∞ (tiene traza 2 y fija a ∞). En cambio, la transformación $s(z) = -1/z$ es un giro de 180° alrededor de i. ∎

Definición 11.14 Si Γ es un subgrupo de $\mathrm{LE}(2,\mathbb{R})$, diremos que $z \in \mathbb{R} \cup \{\infty\}$ es un punto *parabólico* (respecto de Γ) si existe un elemento parabólico $g \in \Gamma$ tal que $g(z) = z$. Diremos que $z \in H$ es *elíptico* (respecto de Γ) si existe un elemento elíptico $g \in \Gamma$ tal que $g(z) = z$.

Notemos que si $z \in \mathbb{R} \cup \{\infty\}$ es parabólico para Γ y $g \in \Gamma$, entonces $g(z)$ también lo es, pues si $f \in \Gamma$ es parabólico y $f(z) = z$, tenemos que $f^g \in \Gamma$ también es parabólico (por ejemplo, porque tiene la misma traza) y claramente $f^g(g(z)) = g(z)$. Igualmente, la imagen de un punto elíptico $z \in H$ por un elemento de Γ es de nuevo un punto elíptico. Definimos

$$H^* = H \cup \{z \in \mathbb{R} \cup \{\infty\} \mid z \text{ es parabólico (respecto de } \Gamma)\}.$$

Según acabamos de observar, si $z \in H^*$ y $g \in \Gamma$, entonces $g(z) \in H^*$, de modo que Γ actúa sobre H^*. Además la órbita de un punto parabólico está formada exclusivamente por puntos parabólicos y la órbita de un punto elíptico está formada exclusivamente por puntos elípticos. En la práctica hablaremos de puntos parabólicos y elípticos de H^*/Γ para referirnos a las órbitas formadas por tales puntos de H^*. Así, el cociente H^*/Γ contiene los puntos de H/Γ más los puntos parabólicos, que son los puntos que nos proponíamos añadir.

Seguidamente calculamos como ejemplo los puntos elípticos y parabólicos del grupo modular.

Teorema 11.15 *Los puntos parabólicos del grupo modular Γ son ∞ y los números racionales. Todos ellos constituyen una única clase de equivalencia. Los puntos elípticos son los equivalentes a i y a $\rho = (1+\sqrt{-3})/2$.*

DEMOSTRACIÓN: Claramente ∞ queda fijo por la transformación parabólica $g(z) = z+1$. Si $r = p/q$ es un número racional (con p y q primos entre sí), tomamos enteros u y v tales que $up + vq = 1$. Sea

$$g = \begin{pmatrix} p & q \\ u & v \end{pmatrix} \in \Gamma.$$

Claramente $g(\infty) = r$, luego r es un punto parabólico equivalente a ∞. Supongamos ahora que $z \in \mathbb{R}$ es parabólico y veamos que es un número racional. Sea

$$g = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \Gamma$$

una transformación parabólica tal que $g(z) = z$.

Esto significa que $cz^2 + (d-a)z - b = 0$ (con $c \neq 0$, o si no $g = \pm 1$), y además z es la única solución de esta ecuación. Por consiguiente el polinomio es reducible en $\mathbb{Q}[x]$ y su raíz z ha de ser racional.

Ocupémonos ahora de los puntos elípticos. Una transformación elíptica g ha de cumplir $|\operatorname{Tr}(g)| < 2$, luego $\operatorname{Tr}(g) = -1, 0, 1$ y el polinomio característico de g ha de ser x^2+1 o bien $x^2 \pm x + 1$. Teniendo en cuenta que g y $-g$ fijan a los mismos puntos (pues determinan la misma aplicación), podemos suponer que el polinomio característico de g es x^2+1 o bien x^2+x+1. Ambos polinomios son irreducibles en $\mathbb{Q}[x]$, por lo que g, vista como matriz racional, tiene un único factor invariante y su forma canónica correspondiente es una de las matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{o bien} \quad \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

Esto significa que g es conjugada a una de estas dos matrices, pero necesitamos garantizar que la matriz de paso está en Γ. Para ello consideramos el anillo $\mathbb{Z}[g]$ (subanillo del anillo de matrices 2×2 con coeficientes enteros, donde $\mathbb{Z}$ se identifica con el anillo de matrices diagonales). Teniendo en cuenta que g es raíz de su polinomio característico, es fácil ver que todo elemento de $\mathbb{Z}[g]$ es de la forma $m + ng$, con $m, n \in \mathbb{Z}$. La expresión es única, pues si $m + ng = 0$ entonces g sería diagonal, luego sería $g = \pm 1$. Es fácil ver entonces que $\mathbb{Z}[g]$ es isomorfo a $\mathbb{Z}[i]$ o bien a $\mathbb{Z}[\rho]$. En particular $\mathbb{Z}[g]$ es un dominio de ideales principales, y esto permite aplicar toda la teoría de endomorfismos de espacios vectoriales al caso de g como endomorfismo de $\mathbb{Z}^2$.

Concretamente, $\mathbb{Z}^2$ es un $\mathbb{Z}[g]$-módulo finitamente generado y libre de torsión, pues si $(m+ng)v = 0$, para un $v \in \mathbb{Z}^2$ no nulo, entonces $g(v) = -(m/n)v$, luego $-m/n$ sería un valor propio de g y, en particular una raíz (racional) de su polinomio característico, lo cual es imposible. Por consiguiente, $\mathbb{Z}^2$ es un $\mathbb{Z}[g]$-módulo libre. No puede tener rango 2, ya que entonces g tendría un espacio de vectores propios de dimensión 2 o bien dos vectores propios reales, en contradicción con

su carácter elíptico. Si v es una base de $\mathbb{Z}^2$ como $\mathbb{Z}[g]$-módulo, entonces v, $g(v)$ es una base de $\mathbb{Z}^2$ como $\mathbb{Z}$-módulo y en dicha base, la matriz de g es una de las dos formas canónicas anteriores.

Ahora sabemos que la matriz de cambio de base es entera y tiene determinante ± 1. Si tiene determinante -1 consideramos la base $g(v)$, v, en la cual la matriz de g es

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{o bien} \quad \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Concluimos que g es conjugada en Γ a una de las cuatro matrices anteriores, y es claro que todo punto fijado por g es equivalente respecto de Γ a un punto fijado por una de estas matrices. Así pues, los puntos elípticos de Γ son los equivalentes a los puntos fijados por ellas. Éstos han de cumplir $z^2 + 1 = 0$ o bien $z^2 - z + 1 = 0$. Las soluciones en H son i y $\rho + 1$, y el último punto es equivalente a ρ. ∎

Para comprender la diferencia geométrica entre los puntos elípticos, los puntos parabólicos y los demás puntos de H hemos de considerar los estabilizadores:

Definición 11.16 Si Γ es un subgrupo de $\mathrm{LE}(2,\mathbb{R})$ y $z \in H^*$, definimos

$$\Gamma_z = \{g \in \Gamma \mid zg = z\}.$$

Representaremos por $\bar{\Gamma}_z$ al correspondiente grupo de transformaciones de Möbius o, equivalentemente, a la imagen de Γ_z por el epimorfismo

$$\mathrm{LE}(2,\mathbb{R}) \longrightarrow \mathrm{LE}(2,\mathbb{R})/\{\pm 1\}.$$

Es inmediato comprobar que si $g \in \Gamma$ y $z_1 g = z_2$, entonces $\Gamma_{z_2} = \Gamma_{z_1}^g$, luego puntos equivalentes tienen estabilizadores isomorfos.

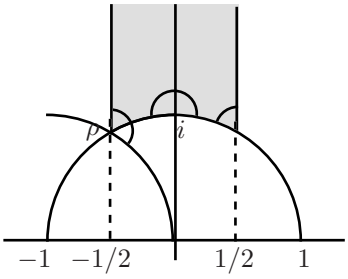
Claramente, si $z \in H$ y $g \in \Gamma_z$ cumple $g \neq \pm 1$, entonces g ha de ser una transformación elíptica, luego z ha de ser un punto elíptico. En otros términos, si $z \in H$ no es elíptico entonces $\Gamma_z \subset \{\pm 1\}$, luego $\bar{\Gamma}_z = 1$.

Enseguida nos ocuparemos de determinar la estructura de $\bar{\Gamma}_z$ cuando z es un punto elíptico o parabólico, pero de momento veamos el caso concreto del grupo modular. Unas simples comprobaciones nos dan que

$$\begin{aligned} \bar{\Gamma}_i &= \langle s \rangle, \quad \text{donde} \quad s(\tau) = -\tfrac{1}{\tau}, \\ \bar{\Gamma}_\rho &= \langle g \rangle, \quad \text{donde} \quad g(\tau) = -\tfrac{1}{\tau+1}, \\ \bar{\Gamma}_\infty &= \langle t \rangle, \quad \text{donde} \quad t(\tau) = \tau + 1. \end{aligned}$$

El grupo $\bar{\Gamma}_i$ tiene dos elementos. Notemos que $ds_i = -d\tau$, lo que significa que en un entorno de i la transformación s se comporta como un giro de 180°, luego los puntos de un entorno de i se agrupan en parejas de puntos equivalentes respecto a Γ. Por lo tanto, si queremos pensar en un entorno de i en el cociente H^*/Γ, no hemos de mirar todo un entorno de i en H^*, pues entonces vemos cada

punto repetido dos veces, sino que hemos de mirar sólo "la mitad", por ejemplo, la mitad que queda dentro del dominio fundamental y que está señalada en la figura.



Similarmente, el grupo $\bar{\Gamma}_\rho$ tiene tres elementos y, como $dg_\rho = \rho^2\, d\tau$, en un entorno de ρ la transformación g se comporta como un giro de 240°, de modo que cada punto del entorno correspondiente en H^*/Γ está representado por tres puntos alrededor de ρ. Si queremos pensar en un entorno de ρ en H^*/Γ hemos de considerar únicamente la tercera parte de un entorno de ρ en H^*, tal y como indica la figura o, si no queremos salirnos del dominio fundamental, podemos mirar la sexta parte que queda dentro del dominio más la otra sexta parte marcada cerca de $\rho + 1$.

El grupo $\bar{\Gamma}_\infty$ es infinito. Todavía no estamos en condiciones de hablar con precisión de lo que ocurre en este caso, pero básicamente es que al mirar un entorno de infinito en H^* estamos viendo cada punto de H^*/Γ repetido infinitas veces.

Finalmente, si z es cualquier punto de H no elíptico, en un entorno suficientemente pequeño no hay puntos equivalentes. Esto es trivial si z está en el interior del dominio fundamental, pero si está en la frontera (y no es elíptico) también se cumple. Por ejemplo, si z está en la recta $\operatorname{Re} z = -1/2$, los puntos a la izquierda de z se corresponden con los puntos a la izquierda de su simétrico respecto al eje imaginario, pero no con los puntos a su derecha, luego un entorno de z en H^* suficientemente pequeño se corresponde unívocamente con un entorno de z en H^*/Γ.

El resto de esta sección lo dedicamos a precisar y generalizar las ideas que acabamos de discutir, pues en la sección siguiente nos basaremos en ellas para definir la estructura analítica de H^*/Γ. Empezamos analizando los estabilizadores de los puntos elípticos:

Teorema 11.17 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ y $z \in H$ un punto elíptico respecto de Γ. Entonces Γ_z (y, por consiguiente, $\bar{\Gamma}_z$) es un grupo cíclico finito.*

Demostración: Recordemos que el subgrupo de $\mathrm{LE}(2,\mathbb{R})$ formado por los elementos que fijan a i es el grupo ortogonal especial $\mathrm{OE}(2,\mathbb{R})$, que es compacto. El grupo $K = \{g \in \mathrm{LE}(2,\mathbb{R}) \mid zg = z\}$ es un conjugado de $\mathrm{OE}(2,\mathbb{R})$, luego es topológicamente isomorfo a la circunferencia unidad S^1. Entonces $\Gamma_z = K \cap \Gamma$ es a la vez compacto y discreto, luego es finito. De hecho, es isomorfo a un subgrupo finito de $\mathbb{R}/\mathbb{Z}$ y es fácil ver que todo subgrupo finito de $\mathbb{R}/\mathbb{Z}$ es cíclico (su antiimagen en $\mathbb{R}$ ha de ser discreta y todo subgrupo discreto de $\mathbb{R}$ es cíclico infinito). ∎

Teorema 11.18 *Si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, entonces los elementos de orden finito de Γ son ± 1 y los elementos elípticos.*

DEMOSTRACIÓN: El teorema anterior prueba que los elementos elípticos tienen orden finito. Recíprocamente, supongamos que $g \in \Gamma$ tiene orden finito. Toda matriz es conjugada en $\mathrm{LG}(2,\mathbb{C})$ a una matriz diagonal

$$\begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix}.$$

Si g tiene orden finito, lo mismo le sucede a esta matriz, lo cual se traduce en que ζ es una raíz de la unidad. Si $\zeta = \pm 1$, entonces $g = \pm 1$. En caso contrario, como la traza se conserva, es claro que $|\operatorname{Tr} g| < 2$, luego g es elíptico. ∎

Definición 11.19 Si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ y $z \in H$ es un punto elíptico, llamaremos *orden* de z al número de elementos del grupo $\bar{\Gamma}_z$.

Por ejemplo, el grupo modular determina dos clases de puntos elípticos: los equivalentes a i, de orden 2, y los equivalentes a ρ, de orden 3.

Nos ocupamos ahora de los puntos parabólicos. Si $s \in \mathbb{R} \cup \{\infty\}$ es un punto parabólico de un subgrupo discreto Γ de $\mathrm{LE}(2,\mathbb{R})$, el grupo Γ_s contiene ciertamente a todas las transformaciones parabólicas de Γ que fijan a s, pero en principio podría contener también transformaciones hiperbólicas que tuvieran a s como uno de sus dos puntos fijos. Vamos a ver que no es así, sino que Γ_s consta únicamente de transformaciones parabólicas. (En el caso en que z es un punto elíptico, es evidente que Γ_z consta únicamente de transformaciones elípticas, pues se trata de un grupo finito.)

Para cada $s \in \mathbb{R} \cup \{\infty\}$, llamamos

$$P(s) = \{g \in \mathrm{LE}(2,\mathbb{R}) \mid sg = s,\ g \text{ es parabólico}\} \cup \{\pm 1\}.$$

Queremos probar que si s es un punto parabólico de un subgrupo discreto Γ de $\mathrm{LE}(2,\mathbb{R})$, entonces $\Gamma_s = P(s) \cap \Gamma$. Al mismo tiempo veremos que el grupo de transformaciones de Möbius $\bar{\Gamma}_s$ es un grupo cíclico infinito. Razonaremos con ∞ y después trasladaremos los resultados a un punto parabólico arbitrario por conjugación.

Supongamos, pues, que ∞ es parabólico respecto a Γ. Es fácil ver que las matrices de $\mathrm{LE}(2,\mathbb{R})$ que fijan a ∞ son las de la forma

$$\begin{pmatrix} a & 0 \\ t & a^{-1} \end{pmatrix}$$

y, de entre ellas, las parabólicas (de traza ± 2) son las de la forma

$$\pm \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}.$$

Ahora es claro que el grupo $P(\infty)$ es topológicamente isomorfo a $\mathbb{R} \times \{\pm 1\}$. El grupo $\bar{P}(\infty)$ está formado por las traslaciones $z \mapsto z+t$, y es topológicamente isomorfo a $\mathbb{R}$.

El grupo $\bar{P}(\infty)\cap\bar{\Gamma}$ es un subgrupo discreto de $\bar{P}(\infty)$. Es claro que un subgrupo discreto de $\mathbb{R}$ es cíclico, pues está generado por el menor de sus elementos positivos. Por consiguiente, el grupo $\bar{P}(\infty)\cap\bar{\Gamma}$ está generado por la traslación

$$\sigma = \begin{pmatrix} \pm 1 & 0 \\ t & \pm 1 \end{pmatrix} \in P(\infty)\cap\Gamma$$

con $|t|$ mínimo. Supongamos ahora que Γ_∞ contiene un elemento hiperbólico

$$\tau = \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix}, \qquad |a| \neq 1.$$

Cambiando τ por τ^{-1} si es necesario, podemos suponer que $|a| < 1$. Así

$$\tau^{-1}\sigma\tau = \begin{pmatrix} \pm 1 & 0 \\ a^2 t & \pm 1 \end{pmatrix} \in P(\infty)\cap\Gamma,$$

pero $|a^2 t| < |t|$, contradicción.

Con esto hemos probado que $\Gamma_\infty = P(\infty)\cap\Gamma$, así como que el grupo $\bar{\Gamma}_\infty$ está generado por una traslación horizontal $z \mapsto z + t$. En general:

Teorema 11.20 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ y sea s un punto parabólico de Γ. Entonces $\Gamma_s = P(s)\cap\Gamma$ y $\bar{\Gamma}_s \cong \mathbb{Z}$.*

Demostración: Lo tenemos probado para $s = \infty$ y el caso general se reduce a éste por conjugación. El razonamiento para ello es típico y no lo volveremos a repetir: La matriz

$$\alpha_s = \begin{pmatrix} 0 & -1 \\ 1 & s \end{pmatrix}$$

cumple $s\alpha_s = \infty$, de donde se sigue que ∞ es un punto parabólico de Γ^{α_s}. Sabemos, pues, que $\Gamma_\infty^{\alpha_s} = P(\infty)\cap\Gamma^{\alpha_s}$ y que $\bar{\Gamma}_\infty^{\alpha_s} \cong \mathbb{Z}$. Conjugando por α_s^{-1} obtenemos los hechos análogos para s. ∎

Terminamos con un hecho técnico que necesitaremos más adelante:

Teorema 11.21 *Si $\alpha \in \mathrm{LE}(2,\mathbb{R})$ es elíptico o parabólico, entonces α no es conjugado con α^{-1} en $\mathrm{LE}(2,\mathbb{R})$.*

Demostración: Supongamos que α es elíptico. Entonces α fija a un $z \in H$. Como $\mathrm{LE}(2,\mathbb{R})$ actúa transitivamente sobre H, existe $\tau \in \mathrm{LE}(2,\mathbb{R})$ tal que $z\tau = i$. Así, α^τ es un elemento elíptico que fija a i. Sabemos que esto equivale a que $\alpha^\tau \in \mathrm{OE}(2,\mathbb{R})$.

Es claro que si α es conjugado con su inverso, lo mismo le sucede a α^τ, luego podemos suponer que $\alpha \in \mathrm{OE}(2,\mathbb{R})$. Pongamos que $\alpha^\sigma = \alpha^{-1}$. Más explícitamente:

$$\alpha\sigma = \begin{pmatrix} p & q \\ -q & p \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} p & -q \\ q & p \end{pmatrix} = \sigma\alpha^{-1}.$$

Teniendo en cuenta que $q \neq 0$ (o si no $\alpha = \pm 1$ no sería elíptico), de aquí se sigue que $b = c$ y $a = -d$, pero entonces $1 = \det\sigma = -a^2 - b^2$, lo cual es absurdo.

Supongamos ahora que α es parabólico. En el teorema anterior hemos visto que $\mathrm{LE}(2,\mathbb{R})$ actúa transitivamente sobre $\mathbb{R} \cup \{\infty\}$, por lo que podemos exigir que el punto fijo de α sea 0, y entonces α tiene la forma

$$\alpha = \pm \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

El mismo razonamiento del caso anterior nos lleva ahora a que

$$\pm \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \pm \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix},$$

de donde $b = 0$ y $a = -d$, con lo que $1 = \det\sigma = -a^2$, contradicción. ∎

Ejercicio: Demostrar que todo elemento hiperbólico de $\mathrm{LE}(2,\mathbb{R})$ es conjugado con su inverso.

11.4 La estructura analítica

Finalmente estamos en condiciones de definir una estructura analítica sobre los cocientes H^*/Γ. Empezamos definiendo una topología sobre H^* más adecuada que la usual para tratar con los puntos parabólicos. Más en general, si llamamos

$$\overline{H} = \{z \in \mathbb{C} \mid \mathrm{Im}\, z \geq 0\} \cup \{\infty\},$$

definimos en $\overline{H}$ la topología siguiente:

a) Una base de entornos de un punto $z \in H$ está formada por los discos abiertos usuales $D(z,r)$ contenidos en H.

b) Una base de entornos de un punto $z \in \mathbb{R}$ está formada por los conjuntos $\{z\} \cup T(z,r)$, donde $T(z,r) \subset H$ es el disco abierto de radio r tangente al eje real en z.

c) Una base de entornos de ∞ está formada por los conjuntos

$$\{\infty\} \cup \{z \in H \mid \mathrm{Im}\, z > r\}.$$

Es fácil ver que estas condiciones determinan ciertamente una topología (de Hausdorff) en $\overline{H}$, la cual induce en H la topología usual, mientras que en $\mathbb{R} \cup \{\infty\}$ induce la topología discreta.

Observemos que si $g \in \mathrm{LE}(2,\mathbb{R})$, entonces g induce un homeomorfismo sobre $\overline{H}$, pues g transforma circunferencias de la esfera de Riemann en circunferencias (entendiendo que las circunferencias que pasan por ∞ son las rectas). De esta

forma, si z_1, $z_2 \in \mathbb{R}$ y $z_1 g = z_2$, tenemos que g transforma la circunferencia de un entorno básico de z_1 en una circunferencia que pasa por z_1 y está contenida en H (salvo en z_2), luego dicha circunferencia ha de ser una circunferencia tangente al eje real por z_2. Los puntos del entorno de z_1 han de transformarse en el interior o en el exterior de la circunferencia imagen, pero, como han de quedar dentro de H, han de corresponderse con los puntos interiores. En definitiva, g transforma cada entorno básico de z_1 en un entorno básico de z_2 (y considerando g^{-1} tenemos también el recíproco). Similarmente, si $zg = \infty$, la circunferencia de un entorno de z ha de corresponderse con una circunferencia que pasa por ∞ —es decir, con una recta— completamente contenida en H (salvo en ∞), luego ha de ser una recta horizontal. Los puntos del entorno de z han de corresponderse con uno de los semiplanos determinados por la recta imagen y, como han de estar contenidos en H, dicho semiplano ha de ser el superior, luego el entorno de z se transforma en un entorno básico de ∞. Igualmente se prueba el recíproco.

Esto no significa que el grupo $\mathrm{LE}(2,\mathbb{R})$ actúe (continuamente) sobre $\overline{H}$ con esta topología, pues no es cierto que la acción $\overline{H} \times \mathrm{LE}(2,\mathbb{R}) \longrightarrow \overline{H}$ sea continua. No obstante, si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, la restricción $\overline{H} \times \Gamma \longrightarrow \overline{H}$ sí que es continua. De hecho, es fácil ver que la continuidad de una acción de un grupo discreto Γ sobre un espacio arbitrario S equivale a que cada aplicación $s \mapsto sg$ sea un homeomorfismo.

En particular, tenemos que Γ actúa sobre H^* cuando en H^* consideramos la topología inducida desde $\overline{H}$. Conviene tener presente que H^* no es localmente compacto salvo en el caso en que $H^* = H$.

Supongamos ahora que $\infty \in H^*$ y vamos a estudiar sus entornos. Para cada $g \in \mathrm{LE}(2,\mathbb{R})$, llamaremos c_g al coeficiente superior izquierdo de la matriz g. Así, por ejemplo, sabemos que el estabilizador de ∞ es

$$\Gamma_\infty = \{g \in \Gamma \mid c_g = 0\}.$$

Una simple comprobación muestra que $|c_g|$ depende únicamente de la clase doble $\Gamma_\infty g \Gamma_\infty$, es decir, que si multiplicamos g por elementos de Γ_∞ no alteramos $|c_g|$.

Teorema 11.22 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ respecto al cual ∞ sea un punto parabólico y sea $M > 0$. Entonces hay un número finito de clases $\Gamma_\infty g \Gamma_\infty$ con $g \in \Gamma$ tales que $|c_g| \le M$.*

Demostración: Según 11.20, el grupo $\bar{\Gamma}_\infty$ es cíclico y está generado por una traslación

$$\tau = \pm \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix}.$$

Sea $g = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \Gamma$ tal que $0 < c_g \le |M|$. Vamos a encontrar un $g'' \in \Gamma_\infty g \Gamma_\infty$ tal que ig'' pertenezca a un compacto K que depende sólo de M y h.

En primer lugar, tomamos un entero n tal que $1 \le d + nhc \le 1 + |hc|$. Una simple comprobación muestra que $g' = \tau^n g = \begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix}$ cumple $|c'| = |c|$, $|d'| = d + nhc$. Por (10.7) tenemos que

$$1 \ge \operatorname{Im} g'(i) = \frac{1}{c'^2 + d'^2} \ge \frac{1}{M^2 + (1 + |h|M)^2}.$$

Por otra parte, $\tau(z) = z + h$, luego podemos tomar un entero m adecuado para que $g'' = \tau^n g \tau^m$ cumpla $0 \le \operatorname{Re} ig'' \le |h|$, sin cambiar con ello la parte imaginaria de ig'. En resumen, ig'' está en el compacto

$$K = \{z \in \mathbb{C} \mid 0 \le \operatorname{Re} z \le |h|,\ (M^2 + (1 + |h|M)^2)^{-1} \le \operatorname{Im} z \le 1\}.$$

Con esto hemos probado que toda clase $\Gamma_\infty g \Gamma_\infty$ con $0 < c_g \le |M|$ tiene un representante g'' tal que $ig'' \in K$. El teorema 11.8 aplicado a los compactos $\{i\}$ y K nos da que sólo un número finito de $g'' \in \Gamma$ cumplen $ig'' \in K$, luego hay un número finito de clases en estas condiciones. Finalmente observamos que sólo hay una clase con $c_g = 0$, a saber, Γ_∞, luego el teorema está probado. ∎

Teorema 11.23 *Si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ respecto al cual ∞ es un punto parabólico, existe un $r > 0$ tal que $|c_g| \ge r$ para todo $g \in \Gamma \setminus \Gamma_\infty$. Además, para todo $z \in H$ y todo $g \in \Gamma \setminus \Gamma_\infty$ se cumple*

$$\operatorname{Im} z \operatorname{Im} g(z) \le \frac{1}{r^2}.$$

Demostración: La existencia de r es inmediata a partir del teorema anterior. Además, según (10.7),

$$\operatorname{Im} g(z) = \frac{1}{|cz + d|^2} \operatorname{Im} z \le \frac{\operatorname{Im} z}{(c \operatorname{Im} z)^2} \le \frac{1}{r^2 \operatorname{Im} z}.$$

∎

Ahora podemos traducir estos hechos a resultados sobre puntos parabólicos arbitrarios:

Teorema 11.24 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ y sea s un punto parabólico respecto a Γ. Entonces existe un entorno U de s en H^* tal que*

$$\Gamma_s = \{g \in \Gamma \mid Ug \cap U \ne \varnothing\}.$$

Demostración: No perdemos generalidad si suponemos que $s = \infty$, en cuyo caso basta tomar $U = \{\infty\} \cup \{z \in H^* \mid \operatorname{Im} z > 1/r\}$, donde r viene dado por el teorema anterior. Así, si $g \in \Gamma \setminus \Gamma_\infty$ y $z \in U$ tenemos que $\operatorname{Im}(zg) < 1/r$, luego $zg \notin U$. ∎

Teorema 11.25 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$. Para cada punto parabólico s de Γ y cada compacto K de H existe un entorno U de s tal que $U \cap Kg = \varnothing$ para todo $g \in \Gamma$.*

DEMOSTRACIÓN: Como es habitual, podemos suponer $s = \infty$. Tomemos dos números reales A y B tales que todo $z \in K$ cumpla $0 < A < \operatorname{Im} z < B$. Sea r según 11.23 y tomemos

$$U = \{\infty\} \cup \{z \in H \mid \operatorname{Im} z > \max\{B, 1/Ar^2\}\}.$$

Así, si $z \in K$ y $g \in \Gamma \setminus \Gamma_\infty$, tenemos que $\operatorname{Im} zg < 1/Ar^2$, luego $zg \notin U$. Si, por el contrario $g \in \Gamma_\infty$, entonces $\operatorname{Im} zg = \operatorname{Im} z < B$, luego también $zg \notin U$. ∎

Con esto ya podemos probar algo básico, pero no trivial:

Teorema 11.26 *Si* Γ *es un subgrupo discreto de* $\mathrm{LE}(2,\mathbb{R})$*, entonces el cociente* H^*/Γ *es un espacio de Hausdorff.*

DEMOSTRACIÓN: Sea $p : H^* \longrightarrow H^*/\Gamma$ la proyección natural. Observemos que $H/\Gamma \subset H^*/\Gamma$, además H/Γ es abierto en H^*/Γ, pues $p^{-1}[H/\Gamma] = H$ es abierto en H^*. Además, la topología cociente de H/Γ es la inducida por H^*/Γ, pues un subconjunto de H/Γ es abierto para cualquiera de las dos si y sólo si su antiimagen por p es abierta en H.

Al espacio H/Γ podemos aplicarle el teorema 11.11 y concluir que es un espacio de Hausdorff. El teorema anterior (aplicado a un entorno compacto de un punto de H) muestra que todo punto parabólico puede separarse de todo punto de H/Γ. Sólo falta probar que dos puntos parabólicos distintos tienen entornos disjuntos.

Tomamos, pues, dos puntos parabólicos s y t no relacionados. No perdemos generalidad si suponemos que $t = \infty$. Sea $z \mapsto z + h$ la traslación que genera Γ_∞. Tomemos $u > 0$ y consideremos los conjuntos siguientes:

$$\begin{aligned} L &= \{z \in \mathbb{C} \mid \operatorname{Im} z = u\},\\ K &= \{z \in L \mid 0 \le \operatorname{Re} z \le |h|\},\\ V &= \{z \in H^* \mid \operatorname{Im} z > u\} \cup \{\infty\}. \end{aligned}$$

Como K es compacto, el teorema anterior nos da un entorno U de s (podemos tomarlo básico) tal que $K \cap U\Gamma = \varnothing$. Veamos que $V \cap U\Gamma = \varnothing$. En caso contrario, existe un $g \in \Gamma$ tal que $V \cap Ug \neq \varnothing$. Por hipótesis $sg \neq \infty$, luego la frontera de Ug ha de ser una circunferencia tangente a $\mathbb{R}$, luego claramente $L \cap Ug \neq \varnothing$. Por consiguiente, Ug corta a una traslación de K por un elemento de Γ_∞, es decir, existe $\sigma \in \Gamma_\infty$ tal que $Ug \cap K\sigma \neq \varnothing$. Así, $Ug\sigma^{-1} \cap K \neq \varnothing$, contradicción. ∎

A continuación definimos una estructura analítica en H^*/Γ. Si $z \in H^*$, los teoremas 11.9 y 11.24 nos dan que existe un entorno abierto U de z en H^* tal que

$$\Gamma_z = \{g \in \Gamma \mid Ug \cap U \neq \varnothing\}.$$

En particular, dos elementos de U se corresponden por un elemento de Γ si y sólo si se corresponden por un elemento de Γ_z.

Observemos que el abierto U no puede contener puntos parabólicos o elípticos distintos de z. En efecto, si $z' \in U$ fuera un punto elíptico o parabólico $z' \neq z$, entonces todo $g \in \Gamma_{z'}$ cumpliría $z' \in Ug \cap U \neq \varnothing$, luego sería $g \in \Gamma_z$, luego g tendría dos puntos fijos, z y z', lo cual es imposible. Esto prueba que el conjunto de los puntos elípticos y parabólicos es un subespacio discreto de H^*, luego sus órbitas forman a su vez un subespacio discreto de H^*/Γ.

Si z no es un punto elíptico ni parabólico, entonces $\bar{\Gamma}_z = 1$ y U no puede contener puntos equivalentes, luego la proyección natural $p : H^* \longrightarrow H^*/\Gamma$ se restringe a una aplicación inyectiva $p|_U : U \longrightarrow H^*/\Gamma$. Obviamente, dicha restricción es un homeomorfismo en la imagen. Tomamos a su inversa como carta de H^*/Γ, es decir, las cartas alrededor de los puntos de H^*/Γ que no son elípticos ni parabólicos serán por definición las inversas de las restricciones de p a los abiertos donde p es inyectiva. Al componer dos cartas de este tipo obtenemos la identidad, que ciertamente es holomorfa, luego estas cartas son compatibles entre sí.

Supongamos ahora que z es un punto elíptico de orden n. Esto significa que el grupo $\bar{\Gamma}_z$ es cíclico de orden n. Sea $\lambda : H \longrightarrow D(0,1)$ una transformación conforme que cumpla $\lambda(z) = 0$. Entonces $\lambda^{-1}\bar{\Gamma}_z\lambda$ es un grupo de n transformaciones conformes del disco unitario en sí mismo que fijan a 0. Es conocido que las únicas transformaciones en estas condiciones son de la forma $w \mapsto \zeta w$, donde $|w| = 1$, y para que formen un grupo cíclico de orden n la única posibilidad es que ζ tome los valores $e^{2k\pi i/n}$.

Así, dos puntos w_1, $w_2 \in U$ se corresponden por un elemento de Γ si y sólo si se corresponden por un elemento de Γ_z, si y sólo si $\lambda(w_1) = e^{2k\pi i/n}\lambda(w_2)$, si y sólo si $\lambda(w_1)^n = \lambda(w_n)^n$.

La aplicación $p[U] \longrightarrow D(0,1)$ dada por $p(w) \mapsto \lambda(w)^n$ está bien definida y es inyectiva. De hecho, su imagen es un abierto y la aplicación es un homeomorfismo en la imagen. Para probarlo observamos que si A es abierto en $p[U]$, entonces $p|_U^{-1}[A]$ es un abierto en U, su imagen por λ es un abierto en el disco unitario y su imagen por $w \mapsto w^n$ también es abierta. El recíproco se prueba igualmente. Tomamos esta aplicación como carta alrededor de $p(z)$.

Si componemos dos cartas correspondientes a dos elecciones de λ obtenemos una transformación conforme. Ello se debe a que $\lambda_1^{-1} \circ \lambda_2$ es una rotación del disco unidad, de donde se sigue que la composición de las cartas es (la restricción) de una rotación de ángulo n veces mayor.

Por otra parte, si componemos una de estas cartas con la carta correspondiente a un punto no elíptico $p|_{U'}^{-1} : p[U'] \longrightarrow U'$, obtenemos la aplicación $w \mapsto \lambda(w)^n$ que es holomorfa y biyectiva (en su dominio), luego una transformación conforme.

No es necesario considerar la composición de dos cartas correspondientes a dos puntos elípticos distintos, pues hemos visto que $p[U]$ no tiene más puntos elípticos. Por consiguiente, un punto en la intersección de los dominios de cartas de este tipo sería no elíptico, y la composición de las cartas elípticas podría desdoblarse en la composición de cada una de ellas con una carta no elíptica.

Consideremos finalmente el caso en que z es un punto parabólico. Tomemos $\alpha \in \mathrm{LE}(2,\mathbb{R})$ tal que $\alpha(z) = \infty$. Entonces $\alpha^{-1}\bar{\Gamma}_z\alpha$ está generado por una traslación $w \mapsto w + h$, con $h > 0$. Por consiguiente, dos puntos w_1, $w_2 \in U$ se corresponden por un elemento de Γ si y sólo si se corresponden por un elemento de Γ_z, si y sólo si $\alpha(w_1) = \alpha(w_2) + kh$, si y sólo si $e^{2\pi i\alpha(w_1)/h} = e^{2\pi i\alpha(w_2)/h}$. Tomamos como carta de dominio $p[U]$ la aplicación $p(w) \mapsto e^{2\pi i\alpha(w)/h}$ (con el convenio de que $p(z) \mapsto 0$). Notemos que la imagen de $p[U]$ es un disco abierto de centro 0. Todas las comprobaciones son análogas a las del caso anterior.

Con esto tenemos probada la mayor parte del teorema siguiente:

Teorema 11.27 *Si Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$, existe una única estructura analítica sobre H^*/Γ (considerado como espacio topológico con la topología cociente) tal que la aplicación natural $p : H \longrightarrow H^*/\Gamma$ es holomorfa.*

Demostración: Para probar la existencia sólo nos falta comprobar que, respecto de la estructura que acabamos de definir, la aplicación p es holomorfa. Sabemos que lo es en cada punto no elíptico, pues se restringe a la inversa de una carta. Si $z \in H$ es elíptico, la composición de p con la carta que hemos construido alrededor de $p(z)$ es $z \mapsto \lambda(z)^n$, que también es holomorfa.

Supongamos que tenemos dos estructuras analíticas en H^*/Γ respecto a las cuales $p|_H$ es holomorfa. Hemos visto que p se restringe a un homeomorfismo (y, por consiguiente, a una transformación conforme para ambas estructuras) en un entorno de cada punto de H no elíptico, luego las cartas de ambas estructuras alrededor de puntos no elípticos ni parabólicos son compatibles entre sí. En otras palabras, la identidad en H^*/Γ es holomorfa (como aplicación entre ambas estructuras) salvo a lo sumo en los puntos elípticos y los parabólicos. Ahora bien, estos puntos son aislados, luego una lectura de la identidad alrededor de uno de ellos es una aplicación holomorfa en un disco salvo una posible singularidad en su centro, que resulta ser evitable porque la aplicación está acotada. ∎

Definición 11.28 Una *superficie modular* es una superficie de Riemann (compacta) de la forma H^*/Γ, donde Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$.

Notemos que, en general, las variedades analíticas H^*/Γ no tienen por qué ser compactas, luego no todo grupo Γ determina una superficie modular. Hemos visto que los puntos elípticos y los parabólicos en una variedad H^*/Γ forman un conjunto discreto, luego concluimos que una superficie modular tiene un número finito de puntos elípticos y parabólicos. En otras palabras, una condición necesaria para que un grupo Γ determine una superficie modular es que determine una cantidad finita de órbitas de puntos elípticos y parabólicos.

Si Γ es el grupo modular, teniendo en cuenta cómo es su dominio fundamental, es claro que el cociente H^*/Γ es topológicamente una esfera, luego efectivamente Γ determina una superficie modular. Sabemos que tiene un único punto parabólico y dos puntos elípticos.

A partir del grupo modular podemos encontrar muchos otros grupos que determinan superficies modulares. Para ello conviene obtener algunos resultados generales.

Definición 11.29 Diremos que dos subgrupos G_1 y G_2 de un grupo G son *conmensurables* si $G_1 \cap G_2$ tiene índice finito en ambos.

Se trata de una relación de equivalencia, pues si G_1 es conmensurable con G_2 y G_2 es conmensurable con G_3, entonces tenemos una inyección natural

$$(G_1 \cap G_2) \,/\, (G_1 \cap G_2 \cap G_3) \longrightarrow G_2 \,/\, (G_2 \cap G_3)$$

que prueba que $G_1 \cap G_2 \cap G_3$ tiene índice finito en $G_1 \cap G_2$, luego en G_1, de donde $G_1 \cap G_3$ también tiene índice finito en G_1. Igualmente se prueba que $G_1 \cap G_3$ tiene índice finito en G_3.

Si G_1 y G_2 son subgrupos conmensurables de un mismo grupo topológico, entonces G_1 es discreto si y sólo si lo es G_2. En efecto, por la transitividad, no perdemos generalidad si suponemos que G_2 es un subgrupo de índice finito en G_1. Así, si G_1 es discreto también lo es G_2 trivialmente. Si G_2 es discreto, entonces G_1 es una unión finita de clases módulo G_2, es decir, una unión finita de espacios cerrados discretos disjuntos, luego es discreto.

Teorema 11.30 *Si Γ y Γ' son subgrupos conmensurables discretos de $\mathrm{LE}(2,\mathbb{R})$ entonces ambos determinan el mismo conjunto de puntos parabólicos.*

Demostración: No perdemos generalidad si suponemos que Γ' es un subgrupo de índice finito en Γ. Obviamente, todo punto parabólico de Γ' lo es de Γ. Recíprocamente, si s es un punto parabólico de Γ, entonces existe un elemento parabólico $g \in \Gamma$ tal que $sg = s$. Como el índice $|\Gamma : \Gamma'|$ es finito, $g^n \in \Gamma'$ para cierto natural $n > 0$, con lo que obtenemos un elemento parabólico de Γ' que también fija a s. Así pues, s es un punto parabólico de Γ'. ∎

Notemos que en las hipótesis del teorema puede ocurrir que dos puntos parabólicos estén relacionados para un subgrupo y no para el otro.

Teorema 11.31 *Si Γ y Γ' son subgrupos conmensurables discretos de $\mathrm{LE}(2,\mathbb{R})$ entonces H^*/Γ es compacto si y sólo si lo es H^*/Γ'.*

Demostración: Podemos suponer que Γ' es un subgrupo de índice finito en Γ. En estas condiciones, la aplicación natural $H^*/\Gamma' \longrightarrow H^*/\Gamma$ es continua y suprayectiva. Por lo tanto, si H^*/Γ' es compacto, también lo es H^*/Γ. Supongamos ahora que H^*/Γ es compacto. Vamos a ver que todo $x \in H^*/\Gamma$ tiene un entorno que es imagen de un subconjunto compacto de H^* por la aplicación $p : H^* \longrightarrow H^*/\Gamma$. Si x no es parabólico, entonces $x = p(z)$, con $z \in H$, y basta tomar un entorno compacto de z en H. Si x es parabólico, entonces $x = p(s)$, donde $s \in \mathbb{R} \cup \{\infty\}$, y el problema es que s no tiene entornos compactos en H^*. No perdemos generalidad si suponemos que $s = \infty$.

Sea $U = \{\infty\} \cup \{z \in \mathbb{C} \mid \operatorname{Im} z \geq 1\}$ un entorno de ∞. El grupo $\bar{\Gamma}_\infty$ está generado por una traslación $g(z) = z + h$, con $h > 0$. Por lo tanto, todo punto de U es equivalente respecto a Γ_∞ con un punto del compacto $K = \{\infty\} \cup \{z \in U \mid 0 \leq \operatorname{Re} z \leq h\}$. Así pues, el entorno $p[U] = p[K]$ cumple lo pedido.

Como H^*/Γ es compacto $H^*/\Gamma = \bigcup_{j=1}^{n} p[K_j]$, donde $K_j \subset H^*$ es compacto. Si $\Gamma/\Gamma' = \{g_1\Gamma', \ldots, g_m\Gamma'\}$, entonces

$$H^*/\Gamma' = \bigcup_{j=1}^{n}\bigcup_{l=1}^{m} p'[K_j g_l],$$

luego H^*/Γ' es compacto. ∎

En particular, todo subgrupo de $\mathrm{LE}(2,\mathbb{R})$ conmensurable con el grupo modular determina una superficie modular. Consideremos con más detalle el caso en que Γ es un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ que define una superficie modular y Γ' es un subgrupo de Γ de índice finito. Tal y como hemos comentado en la prueba del teorema anterior, la aplicación natural $\phi : H^*/\Gamma' \longrightarrow H^*/\Gamma$ es continua y suprayectiva. Más aún, es holomorfa. Para probarlo tomamos un punto arbitrario $z \in H^*$. Claramente, $\bar{\Gamma}'_z \leq \bar{\Gamma}_z$.

Si $\bar{\Gamma}_z = 1$, entonces z no es elíptico ni parabólico para ninguna de las dos superficies modulares. Podemos tomar como cartas alrededor de $p'(z)$ y $p(z)$ las aplicaciones $p'(w) \mapsto w$ y $p(w) \mapsto w$, con lo que la lectura de ϕ es la identidad.

Supongamos ahora que z es un punto elíptico para Γ. Sean

$$n = |\bar{\Gamma}_z|, \quad m = |\bar{\Gamma}'_z|, \quad e = |\bar{\Gamma}_z : \bar{\Gamma}'_z|.$$

Notemos que z será un punto elíptico para Γ' si $m > 1$ y será un punto normal (ni elíptico ni parabólico) si $m = 1$.

Para construir cartas alrededor de $p'(z)$ y $p(z)$ tomamos una transformación conforme $\lambda : H \longrightarrow D(0,1)$. Las cartas vienen dadas por $p'(w) \mapsto \lambda(w)^n$ y $p'(w) \mapsto \lambda(w)^m$, de modo que la lectura de ϕ en estas cartas es $w \mapsto w^e$, que es una función holomorfa.

Si z es un punto parabólico para Γ, entonces también lo es para Γ', con lo que $\bar{\Gamma}_z$ y $\bar{\Gamma}'_z$ son dos grupos cíclicos infinitos. Sea $e = |\bar{\Gamma}_z : \bar{\Gamma}'_z|$. Tomamos $\alpha \in \mathrm{LE}(2,\mathbb{R})$ tal que $\alpha(z) = \infty$. Claramente $|\bar{\Gamma}^\alpha_\infty : \bar{\Gamma}'^\alpha_\infty| = e$ y si $\bar{\Gamma}^\alpha_\infty$ está generado por la traslación $g(w) = w + h$, con $h > 0$, entonces $\bar{\Gamma}'^\alpha_\infty$ está generado por la traslación $g'(w) = w + eh$.

Ahora las cartas son $p'(w) \mapsto e^{2\pi i \alpha(w)/eh}$ y $p(w) \mapsto e^{2\pi i \alpha(w)/h}$, luego la lectura de ϕ en estas cartas es $w \mapsto w^e$, que es una función holomorfa.

Casi tenemos probado el teorema siguiente:

Teorema 11.32 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ tal que H^*/Γ sea compacto y sea Γ' un subgrupo de índice finito. Sea $\phi : H^*/\Gamma' \longrightarrow H^*/\Gamma$ la aplicación natural. Entonces ϕ es una aplicación holomorfa de grado $n(\phi) = |\bar{\Gamma} : \bar{\Gamma}'|$. Para cada $z \in H^*$, el índice de ramificación de ϕ en $p'(z)$ es*

$$e(\phi, p'(z)) = |\bar{\Gamma}_z : \bar{\Gamma}'_z|.$$

DEMOSTRACIÓN: Sólo falta comprobar que el grado de ϕ es el indicado. Dicho grado es el número de antiimágenes de un punto de H^*/Γ no ramificado, por ejemplo, cualquier punto $p(z)$ que no sea ni elíptico ni parabólico. Sea

$$\bar\Gamma/\bar\Gamma' = \{g_1\bar\Gamma', \dots, g_n\bar\Gamma'\}.$$

Entonces, $\phi^{-1}[p(z)] = \{p'(zg_1), \dots, p'(zg_n)\}$ y los n puntos son distintos dos a dos. En efecto, si $p'(zg_i) = p'(zg_j)$, entonces existe un $h \in \bar\Gamma'$ tal que $zg_ih = zg_j$, es decir, $zg_ihg_j^{-1} = z$, luego $g_ihg_j^{-1} \in \bar\Gamma_z = 1$, luego $g_i^{-1}g_j = h \in \bar\Gamma'$, luego $g_i\bar\Gamma' = g_j\bar\Gamma'$, luego $i = j$.

Por otra parte, si $p'(w) \in \phi^{-1}[p(z)]$, entonces $p(w) = p(z)$, luego existe un $g \in \bar\Gamma$ tal que $w = zg$. Podemos expresar $g = g_ih$, con $h \in \bar\Gamma'$, y entonces $p'(w) = p'(zg_ih) = p'(zg_i)$. ∎

En el caso en que Γ es el grupo modular, podemos usar la fórmula de Hurwitz para determinar el género de la superficie definida por el subgrupo Γ':

Teorema 11.33 *Sea Γ' un subgrupo de índice finito del grupo $\Gamma = \mathrm{LE}(2,\mathbb{Z})$. Sea $\mu = |\bar\Gamma : \bar\Gamma'|$ y sean ν_2, ν_3 y ν_∞ el número de puntos elípticos de orden 2, puntos elípticos de orden 3 y de puntos parabólicos de H^*/Γ' respectivamente. Entonces el género g de H^*/Γ' es*

$$g = 1 + \frac{\mu}{12} - \frac{\nu_2}{4} - \frac{\nu_3}{3} - \frac{\nu_\infty}{2}.$$

DEMOSTRACIÓN: Sea $\phi : H^*/\Gamma' \longrightarrow H^*/\Gamma$ la aplicación natural. La fórmula de Hurwitz afirma que

$$2 - 2g = 2n(\phi) + \sum_a (1 - e(\phi, a)),$$

donde a recorre los puntos de H^*/Γ'. Por el teorema anterior sabemos que $n(\phi) = \mu$ y que si $a = p'(z)$, entonces $e(\phi, a) = |\bar\Gamma_z : \bar\Gamma'_z|$. Vemos, pues, que $e(\phi, a) = 1$ siempre que $\bar\Gamma_z = 1$, es decir, cuando a no es ni un punto elíptico ni parabólico de H^*/Γ. En otras palabras, salvo si $\phi(a)$ es $p(i)$, $p(\rho)$ o $p(\infty)$.

Si $\phi(a) = p(i)$, entonces $|\bar\Gamma_z| = 2$, con lo que $|\bar\Gamma'_z|$ puede ser 1 o 2. Por hipótesis $p(i)$ tiene ν_2 antiimágenes a para las cuales $|\bar\Gamma'_z| = 2$, y entonces $e(\phi, a) = 1$, y digamos que tiene ν_2' antiimágenes con $|\bar\Gamma'_z| = 1$, con lo que $e(\phi, a) = 2$.

La suma de los índices de las antiimágenes ha de ser el grado de ϕ, es decir, $\nu_2 + 2\nu_2' = \mu$, luego

$$\sum_{\phi(a)=p(i)} (1 - e(\phi, a)) = \frac{\mu - \nu_2}{2}(-1) = \frac{\nu_2 - \mu}{2}.$$

El mismo razonamiento nos da que

$$\sum_{\phi(a)=p(\rho)} (1 - e(\phi, a)) = \frac{\mu - \nu_3}{3}(-2) = \frac{2(\nu_3 - \mu)}{3}.$$

Finalmente observamos que todas las antiimágenes de $p(\infty)$ son puntos parabólicos de H^*/Γ, luego hay ν_∞ de ellas. Por consiguiente,

$$\sum_{\phi(a)=p(\infty)} (1 - e(\phi, a)) = \nu_\infty - \sum_{\phi(a)=p(\infty)} e(\phi, a) = \nu_\infty - \mu.$$

Así pues,

$$2 - 2g = 2\mu + \frac{\nu_2 - \mu}{2} + \frac{2(\nu_3 - \mu)}{3} + \nu_\infty - \mu.$$

Al despejar obtenemos la fórmula del enunciado. ∎

Terminamos con un teorema que necesitaremos en la sección siguiente:

Teorema 11.34 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ tal que H^*/Γ sea compacto y sea Γ' un subgrupo de índice finito. Sea $\phi : H^*/\Gamma' \longrightarrow H^*/\Gamma$ la aplicación natural. Sea $z \in H^*$ y supongamos que $\phi^{-1}(p(z)) = \{p'(z_1), \ldots, p'(z_k)\}$. Sea $\sigma_i \in \Gamma'$ tal que $z_i\sigma_i = z$. Entonces*

$$\Gamma = \bigcup_{i=1}^{k} \Gamma_z \sigma_i \Gamma' \qquad \textit{(unión disjunta).}$$

Es decir, hay tantas antiimágenes de $p(z)$ como clases dobles $\Gamma_z\sigma_i\Gamma'$.

Demostración: Tomemos $\sigma \in \Gamma$. Entonces $p'(z\sigma) \in \phi^{-1}(p(z))$, luego $p'(z\sigma) = p'(z_i)$, para algún i, luego existe $\tau \in \Gamma$ tal que $z\sigma = z_i\tau = z\sigma_i\tau$. Consecuentemente, $\tau' = \sigma_i\tau\sigma^{-1} \in \Gamma_z$, luego $\sigma = \tau'^{-1}\sigma_i\tau \in \Gamma_z\sigma_i\Gamma'$.

Por otra parte, las clases son distintas dos a dos, ya que si $\sigma_j = \tau'\sigma_i\tau$, con $\tau' \in \Gamma_z$ y $\tau \in \Gamma'$, entonces $z_j = z\sigma_j = z\tau'\sigma_i\tau = z\sigma_i\tau = z_i\tau$, luego sería $p'(z_j) = p'(z_i)$, contradicción. ∎

11.5 Ejemplos de superficies modulares

Definición 11.35 Para cada número natural $N \geq 1$, llamaremos

$$\Gamma(N) = \left\{ \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LE}(2,\mathbb{Z}) \;\middle|\; a \equiv d \equiv 1 \ (\text{mód } N),\ b \equiv c \equiv 0 \ (\text{mód } N) \right\}.$$

A $\Gamma(N)$ se le llama *subgrupo de congruencias principal de nivel N*. En general, un *grupo de congruencias* de nivel N es un grupo Γ que satisfaga las inclusiones $\Gamma(N) \leq \Gamma \leq \mathrm{LE}(2,\mathbb{Z})$.

Observemos que $\Gamma(1) = \mathrm{LE}(2,\mathbb{Z})$ es simplemente el grupo modular. El teorema siguiente prueba que $\Gamma(N)$ tiene índice finito en $\Gamma(1)$, con lo que todos los subgrupos de congruencias son conmensurables con el grupo modular y, por consiguiente, definen superficies modulares.

Llamaremos $X(N) = H^*/\Gamma(N)$.

Teorema 11.36 *Sea $f : \mathrm{LE}(2,\mathbb{Z}) \longrightarrow \mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$ la aplicación definida tomando restos coeficiente a coeficiente. Entonces la sucesión*

$$1 \longrightarrow \Gamma(N) \longrightarrow \mathrm{LE}(2,\mathbb{Z}) \xrightarrow{f} \mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z}) \longrightarrow 1$$

es exacta.

DEMOSTRACIÓN: Sólo hay que probar que f es suprayectiva. Para ello tomamos una matriz

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

con coeficientes enteros tal que $ad - bc \equiv 1$ (mód N). Hemos de probar que es congruente (coeficiente a coeficiente) con una matriz de $\mathrm{LE}(2,\mathbb{Z})$.

Por la teoría de divisores elementales, existen matrices B, $B' \in \mathrm{LG}(2,\mathbb{Z})$ tales que BAB' es diagonal. Que B y B' estén en $\mathrm{LG}(2,\mathbb{Z})$ significa que son inversibles, es decir, que tienen determinante ± 1. Multiplicándolas si es necesario por una matriz diagonal de determinante -1, podemos exigir que tengan determinante 1, es decir, que estén en $\mathrm{LE}(2,\mathbb{Z})$.

Si encontramos una matriz $A' \in \mathrm{LE}(2,\mathbb{Z})$ congruente con BAB' módulo N, entonces $B^{-1}A'B'^{-1}$ es congruente con A módulo N y el teorema queda probado. Por consiguiente, podemos suponer que A es diagonal, digamos

$$A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix},$$

donde $ad = 1 + rN$. Basta encontrar enteros x e y tales que la matriz

$$\begin{pmatrix} a + xN & yN \\ N & d \end{pmatrix}$$

tenga determinante 1. Dicho determinante es

$$ad + xdN - yN^2 = 1 + (r + xd - yN)N,$$

luego basta con que $r + xd - yN = 0$. Esto tiene solución porque $(d, N) = 1$. ∎

Así pues, el índice de $\Gamma(N)$ en el grupo modular es precisamente el orden de $\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$. Vamos a calcularlo.

Si $N = \prod p_i^{e_i}$ es la descomposición de N en factores primos, entonces es sabido que

$$\mathbb{Z}/N\mathbb{Z} \cong \bigoplus \mathbb{Z}/p_i^{e_i}\mathbb{Z}.$$

El término derecho es el producto cartesiano de los sumandos con la suma y el producto definidas componente a componente. Es claro que el grupo lineal general 2×2 de este anillo es isomorfo al producto de los grupos lineales generales de los factores:

$$\mathrm{LG}(2,\mathbb{Z}/N\mathbb{Z}) \cong \mathrm{LG}(2,\bigoplus \mathbb{Z}/p_i^{e_i}\mathbb{Z}) \cong \prod \mathrm{LG}(2,\mathbb{Z}/p_i^{e_i}\mathbb{Z}).$$

El 1 de $\mathbb{Z}/N\mathbb{Z}$ se corresponde con $(1,\dots,1)$ en la suma directa, luego las matrices de determinante 1 se corresponden con las n-tuplas de matrices de determinante 1, es decir,

$$\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z}) \cong \prod \mathrm{LE}(2,\mathbb{Z}/p_i^{e_i}\mathbb{Z}).$$

Ahora consideramos la sucesión exacta

$$1 \longrightarrow X \longrightarrow \mathrm{LG}(2,\mathbb{Z}/p^e\mathbb{Z}) \longrightarrow \mathrm{LG}(2,\mathbb{Z}/p\mathbb{Z}) \longrightarrow 1,$$

donde el epimorfismo es el que a cada matriz módulo p^e le hace corresponder la matriz con los mismos coeficientes módulo p.

El núcleo X está formado por todas las matrices

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \ (\text{mód } p^e)$$

tales que $a \equiv d \equiv 1 \ (\text{mód } p)$, $b \equiv c \equiv 0 \ (\text{mód } p)$. (El determinante de una matriz que cumpla esto es $\equiv 1 \ (\text{mód } p)$, luego es una unidad módulo p^e.) Hay p^{e-1} restos módulo p^e congruentes con 0 módulo p y otros tantos congruentes con 1 módulo p, luego $|X| = p^{4(e-1)}$.

Por otro lado, $|\mathrm{LG}(2,\mathbb{Z}/p\mathbb{Z})| = (p^2-1)(p^2-p)$. (Hay tantas matrices regulares como bases ordenadas en un espacio vectorial de dimensión 2 sobre $\mathbb{Z}/p\mathbb{Z}$. El espacio tiene p^2 vectores y todos menos 0 pueden formar parte de una base. Una vez escogido uno, cualquiera menos sus p múltiplos puede completar la base.) Así pues,

$$|\mathrm{LG}(2,\mathbb{Z}/p^e\mathbb{Z})| = p^{4(e-1)}(p^2-1)(p^2-p) = p^{4e}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^2}\right).$$

Ahora consideramos la sucesión exacta

$$1 \longrightarrow \mathrm{LE}(2,\mathbb{Z}/p^e\mathbb{Z}) \longrightarrow \mathrm{LG}(2,\mathbb{Z}/p^e\mathbb{Z}) \xrightarrow{\det} U_{p^e} \longrightarrow 1,$$

de la que deducimos que

$$|\mathrm{LE}(2,\mathbb{Z}/p^e\mathbb{Z})| = p^{3e}\left(1-\frac{1}{p^2}\right).$$

En definitiva:

$$|\Gamma(1):\Gamma(N)| = |\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})| = \prod_i p_i^{3e_i}\left(1-\frac{1}{p_i^2}\right) = N^3\prod_{p|N}\left(1-\frac{1}{p^2}\right).$$

En realidad nos interesa el índice $|\bar\Gamma(1):\bar\Gamma(N)|$. Para calcularlo observamos que $-1 \in \Gamma(2)$ pero $-1 \notin \Gamma(N)$ para $N > 2$. Por lo tanto:

$$\mu_N = |\bar\Gamma(1):\bar\Gamma(N)| = \begin{cases} (N^3/2)\prod_{p|N}\left(1-\frac{1}{p^2}\right) & \text{si } N > 2, \\ 6 & \text{si } N = 2. \end{cases} \tag{11.1}$$

Vamos a calcular el género de la superficie $X(N)$ aplicando el teorema 11.33. Acabamos de calcular el valor de μ. El teorema siguiente nos da que $\nu_2 = \nu_3 = 0$:

Teorema 11.37 *Si $N > 1$ la superficie $X(N)$ no tiene puntos elípticos.*

DEMOSTRACIÓN: Hemos de probar que $\Gamma(N)$ no contiene elementos elípticos. Observemos en primer lugar que $\Gamma(N)$ es claramente un subgrupo normal de $\Gamma(1)$. En la prueba del teorema 11.15 hemos visto que todo elemento elíptico de $\Gamma(1)$ es conjugado en $\Gamma(1)$ con una de las matrices

$$\pm\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Ninguna de ellas está en $\Gamma(N)$ y, por ser normal, tampoco lo están sus conjugadas, luego, en efecto, $\Gamma(N)$ no contiene ningún elemento elíptico. ∎

Respecto a ν_∞ tenemos lo siguiente:

Teorema 11.38 *La superficie $X(N)$ tiene exactamente $\nu_\infty = \mu_N/N$ puntos parabólicos.*

DEMOSTRACIÓN: Claramente

$$\bar\Gamma(1)_\infty = \left\langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle, \quad \bar\Gamma(N)_\infty = \bar\Gamma(N) \cap \bar\Gamma(1)_\infty = \left\langle \begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix} \right\rangle,$$

luego $|\bar\Gamma(1)_\infty : \bar\Gamma(N)_\infty| = N$. Según el teorema 11.32 este índice es el índice de ramificación de ∞ para la aplicación $\phi : X(N) \longrightarrow X(1)$. Ahora bien, si s es cualquier punto parabólico, sabemos que existe $g \in \Gamma(1)$ tal que $\infty g = s$, luego $\bar\Gamma(N)^g_\infty$ es el estabilizador de s en el grupo $\bar\Gamma(N)^g = \bar\Gamma(N)$. Así pues,

$$|\bar\Gamma(1)_s : \bar\Gamma(N)_s| = |\bar\Gamma(1)^g_\infty : \bar\Gamma(N)^g_\infty| = |\bar\Gamma(1)_\infty : \bar\Gamma(N)_\infty| = N.$$

Así pues, $\infty \in X(1)$ tiene exactamente ν_∞ antiimágenes con índice de ramificación N. Consecuentemente, $\mu_N = N\nu_\infty$. ∎

El teorema 11.33 nos da ahora que el género de la superficie $X(N)$ (para $N > 1$) es

$$g_N = 1 + \mu_N \frac{N-6}{12N},$$

donde μ_N es el dado por (11.1). La tabla siguiente contiene los primeros valores de g_N:

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
g_N	0	0	0	0	0	1	3	5	10	13	26	25	50	49	73

Veamos ahora otra familia de superficies modulares de gran interés en teoría de números y en geometría algebraica.

Definición 11.39 Para cada número natural $N \geq 1$ sea

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LE}(2,\mathbb{Z}) \;\middle|\; c \equiv 0 \ (\text{mód } N) \right\}.$$

Claramente $\Gamma(N) \leq \Gamma_0(N) \leq \mathrm{LE}(2,\mathbb{Z})$, luego $\Gamma_0(N)$ define una superficie modular, que representaremos por $X_0(N) = H^*/\Gamma_0(N)$. Observemos que $\Gamma_0(1) = \mathrm{LE}(2,\mathbb{Z})$.

Vamos a calcular los elementos que intervienen en la fórmula de 11.33. En primer lugar observamos que la función f del teorema 11.36 define un isomorfismo entre $\Gamma_0(N)/\Gamma(N)$ y el grupo de las matrices de $\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$ de la forma

$$\begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix}.$$

Este grupo tiene orden $N\phi(N)$, donde ϕ es la función de Euler. Teniendo en cuenta que $-1 \in \Gamma_0(N)$, concluimos que

$$|\bar{\Gamma}(1) : \bar{\Gamma}_0(N)| = |\Gamma(1) : \Gamma_0(N)| = \frac{|\Gamma(1) : \Gamma(N)|}{|\Gamma_0(N) : \Gamma(N)|} = N \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

Así tenemos probada la primera afirmación del teorema siguiente:

Teorema 11.40 *Con la notación de* 11.33 *para* $\Gamma' = \Gamma_0(N)$*, se cumple:*

a) $\mu = N \prod_{p|N} \left(1 + \frac{1}{p}\right)$,

b) $\nu_2 = \begin{cases} 0 & \text{si } 4 \mid N, \\ \prod_{p|N} \left(1 + \left(\frac{-1}{p}\right)\right) & \text{si } 4 \nmid N, \end{cases}$

c) $\nu_3 = \begin{cases} 0 & \text{si } 9 \mid N, \\ \prod_{p|N} \left(1 + \left(\frac{-3}{p}\right)\right) & \text{si } 9 \nmid N, \end{cases}$

d) $\nu_\infty = \sum_{d|N} \phi(\mathrm{mcd}(d, N/d))$,

donde ϕ *es la función de Euler y*

$$\left(\frac{-1}{p}\right) = \begin{cases} 0 & \text{si } p = 2, \\ 1 & \text{si } p \equiv 1 \ (\mathrm{mód}\ 4), \\ -1 & \text{si } p \equiv 3 \ (\mathrm{mód}\ 4), \end{cases} \qquad \left(\frac{-3}{p}\right) = \begin{cases} 0 & \text{si } p = 3, \\ 1 & \text{si } p \equiv 1 \ (\mathrm{mód}\ 3), \\ -1 & \text{si } p \equiv 2 \ (\mathrm{mód}\ 3), \end{cases}$$

son los símbolos de Legendre.

Demostración: Calculamos en primer lugar ν_∞. Es fácil ver que el miembro derecho de d) es multiplicativo, es decir, si lo llamamos $f(N)$ y $N = mn$ con $(m,n) = 1$, entonces $f(N) = f(m)f(n)$. Vamos a probar que $\nu_\infty(N)$ también es una función multiplicativa, con lo que bastará demostrar la igualdad en el caso en que N es potencia de primo.

Supongamos, pues que $N = mn$, con $(m,n) = 1$. Según 11.34, sabemos que ν_∞ es el número de clases dobles $\Gamma_0\sigma\Gamma_0(N)$ (notemos que aquí Γ_0 es el estabilizador del punto parabólico 0). Es claro que éste coincide con el número de clases

dobles en $\Gamma/\Gamma(N)$ respecto a los subgrupos $\Gamma_0\Gamma(N)/\Gamma(N)$ y $\Gamma_0(N)/\Gamma(N)$. El isomorfismo $\Gamma/\Gamma(N) \cong \mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$ dado por el teorema 11.36 nos permite concluir que dicho número coincide con el número de clases dobles en $\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$ respecto a las imágenes de los subgrupos indicados.

Observemos que Γ_0 está formado por las matrices de la forma

$$\begin{pmatrix} \pm 1 & t \\ 0 & \pm 1 \end{pmatrix}, \qquad t \in \mathbb{Z},$$

luego la proyección de $\Gamma_0\Gamma(N)/\Gamma(N)$ módulo N está compuesta por las matrices de esta misma forma pero con $t \in \mathbb{Z}/N\mathbb{Z}$.

Por otra parte, la imagen de $\Gamma_0(N)/\Gamma(N)$ está formada por las matrices de la forma

$$\begin{pmatrix} a & 0 \\ b & d \end{pmatrix}, \qquad ad = 1, \quad a, b, d \in \mathbb{Z}/N\mathbb{Z}.$$

Es claro que, a través de la factorización

$$\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z}) \cong \mathrm{LE}(2,\mathbb{Z}/m\mathbb{Z}) \times \mathrm{LE}(2,\mathbb{Z}/n\mathbb{Z}),$$

ambos subgrupos factorizan como el producto de los subgrupos correspondientes para m y n, luego el número de clases dobles que determinan en $\mathrm{LE}(2,\mathbb{Z}/N\mathbb{Z})$ es el producto de los números correspondientes para m y n, que es lo que queríamos probar.

Así pues, a partir de aquí suponemos que $N = p^e$ es potencia de primo. Consideramos todos los pares (c, d) de números naturales que verifican

$$(c, d) = 1, \quad d \mid N, \quad 0 < c \leq N/d.$$

Para cada uno de ellos, elegimos enteros a y b tales que $ad - bc = 1$ y formamos la matriz

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \Gamma.$$

Vamos a ver que las matrices así formadas constituyen un sistema completo de representantes de las clases $\sigma\Gamma_0(N)$.

Para $d = 1$ hay $N = p^e$ valores posibles para c, mientras que si $d = p^i$, con $0 < i \leq e$, las posibilidades son $\phi(p^{e-i})$. En total, el número de matrices que tenemos es

$$p^e + \sum_{i=0}^{e-1} \phi(p^i) = p^e + p^{e-1} = p^e\left(1 + \frac{1}{p}\right)$$

que es justamente el valor de $|\Gamma : \Gamma_0(N)|$ que teníamos calculado. Por consiguiente sólo hemos de probar que estas matrices no son congruentes módulo $\Gamma_0(N)$. Supongamos que dos de ellas lo fueran:

$$\begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}\begin{pmatrix} u & x \\ v & y \end{pmatrix} = \begin{pmatrix} * & cy + ax \\ * & dy + bx \end{pmatrix},$$

donde $N \mid x$ y $uy \equiv 1 \pmod N$ (en particular $(y, N) = 1$). Sea $N = dk$. Tenemos que $N = dk = (dy + bx)r = dy + bx'dk$, luego $k = y + bx'k$, luego $k \mid y$ y también $k \mid N$. Así pues, $k = 1$ y por lo tanto $d = N$. Por simetría, también $d' = N$, con lo que $d = d'$ y, necesariamente, $c = c' = 1$.

Digamos que dos pares (c, d) y (c', d') están relacionados si

$$\begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix} = \begin{pmatrix} \pm 1 & t \\ 0 & \pm 1 \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} * & \pm c + td \\ * & \pm d \end{pmatrix}.$$

Notemos que la matriz de paso es una matriz arbitraria de Γ_0. El número de clases de equivalencia es precisamente ν_∞. Para calcularlo observamos en primer lugar que la igualdad $d' = \pm d$ obliga a que el signo sea positivo. Vemos que clases de equivalencia distintas corresponden a valores distintos de d, luego ν_∞ será la suma del número de clases de equivalencia correspondientes a cada $d \mid N$. Según el enunciado del teorema, hemos de probar que hay $\phi(\mathrm{mcd}(d, N/d))$.

En efecto, para un $d = p^i$ fijo, hay tantos pares como números naturales $0 < c \leq N/d$ primos con d, y dos de ellos están relacionados si y sólo si los valores correspondientes de c son congruentes módulo d.

Si $i \leq e - i$, entonces $(d, N/d) = d$, con lo que cada c es congruente módulo d con un único natural $0 < c' \leq d$, y ciertamente hay $\phi((d, N/d))$ clases de equivalencia.

Si, por el contrario, $e - i \leq i$, entonces $(d, N/d) = N/d$ y hay $\phi((d, N/d))$ posibilidades para c. Ningún par está relacionado con ningún otro, luego también hay $\phi((d, N/d))$ clases.

Vamos a calcular ahora ν_3. La situación es un poco más delicada porque si $z \in H$ es un punto elíptico de orden 3 respecto a Γ, no es necesariamente cierto que también lo sea respecto a $\Gamma_0(N)$. Lo será si y sólo si existe un $\sigma \in \Gamma_z \cap \Gamma_0(N)$ tal que $\sigma \neq \pm 1$. El grupo Γ_z es cíclico de orden 6, pero si un tal σ tiene orden 6, entonces $-\sigma$ cumple lo mismo y tiene orden 3. Así pues, z es elíptico respecto de $\Gamma_0(N)$ si y sólo si el grupo $\Gamma_0(N)$ contiene un elemento elíptico de orden 3 que fija a z.

Los elementos elípticos de orden 3 son los que tienen polinomio mínimo $x^2 + x + 1$, y en la prueba de 11.15 hemos visto que cualquiera de ellos es conjugado en Γ con una de las matrices

$$\tau = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \qquad \tau^2 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Llamemos S_1 al conjunto de todas las matrices conjugadas con τ y S_2 al conjunto de las conjugadas con τ^2. Por 11.21 sabemos que $S_1 \cap S_2 = \varnothing$. Cada punto z elíptico de orden 3 respecto de Γ es fijado exactamente por dos elementos de orden 3, uno de los cuales está en S_1 y el otro (su cuadrado) está en S_2. Concluimos que z es un punto elíptico respecto de $\Gamma_0(N)$ si y sólo si existe un (único) $\sigma \in S_1 \cap \Gamma_0(N)$ que fija a z.

Si z_1 y z_2 son fijados por σ_1 y $\sigma_2 \in S_1$ y son equivalentes respecto de $\Gamma_0(N)$, entonces $z_1 = z_2 g$, para cierto $g \in \Gamma_0(N)$, con lo que también $z_1 \sigma_2^g = z_1$,

con $\sigma_2^g \in S_1$. Por la unicidad, $\sigma_1 = \sigma_2^g$. Recíprocamente, los puntos fijos de dos elementos de S_1 conjugados en $\Gamma_0(N)$ son equivalentes respecto de $\Gamma_0(N)$. Concluimos que ν_3 es el número de clases de conjugación en que $\Gamma_0(N)$ divide a S_1.

Conviene observar también que si

$$\alpha = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

entonces $\det \alpha = -1$ y $\tau^\alpha = \tau^2$. De aquí se sigue inmediatamente que si conjugamos una matriz de S^2 por una matriz entera de determinante -1 obtenemos una matriz de S_1 (pues si la conjugada estuviera en S_2, podríamos pasar de τ^2 a τ conjugando con una matriz de Γ, en contradicción con 11.21).

Llamemos $\rho = e^{2\pi i/3}$. En la prueba del teorema 11.15 hemos visto que si $g \in S_1 \cup S_2$, entonces el anillo $\mathbb{Z}[g]$ es isomorfo a $\mathbb{Z}[\rho]$, así como que $\mathbb{Z}^2$ adquiere una estructura natural de $\mathbb{Z}[g]$-módulo libre de rango 1. Además, en cierta base de $\mathbb{Z}^2$ como $\mathbb{Z}$-módulo, la matriz de la multiplicación por g es τ, que es la misma que la matriz de la multiplicación por ρ en el $\mathbb{Z}[\rho]$-módulo $\mathbb{Z}[\rho]$ respecto a la base $\{1, \rho\}$. Haciendo corresponder ambas bases obtenemos un $\mathbb{Z}$-isomorfismo $f : \mathbb{Z}[\rho] \longrightarrow \mathbb{Z}^2$ con la propiedad de que

$$f(x\rho) = f(x)g, \qquad \text{para todo } x \in \mathbb{Z}[\rho]. \tag{11.2}$$

Recíprocamente, si $f : \mathbb{Z}[\rho] \longrightarrow \mathbb{Z}^2$ es cualquier $\mathbb{Z}$-isomorfismo, existe un único $g \in S_1 \cup S_2$ para el que se cumple (11.2). En efecto, definimos $g : \mathbb{Z}^2 \longrightarrow \mathbb{Z}^2$ mediante $g(x) = g(f^{-1}(x)\rho)$, con lo que tenemos un isomorfismo de $\mathbb{Z}^2$ de orden 3 cuyo determinante es el mismo que el de la multiplicación por ρ en $\mathbb{Z}[\rho]$, o sea, 1. Identificando a g con su matriz en la base canónica tenemos que $g \in \Gamma$ y cumple obviamente (11.2). La unicidad es obvia.

Llamamos T al conjunto de todos los $\mathbb{Z}$-isomorfismos de $f : \mathbb{Z}[\rho] \longrightarrow \mathbb{Z}^2$. Entonces $T = T_1 \cup T_2$, donde T_1 contiene a los isomorfismos que se corresponden con elementos $g \in S_1$ y T_2 contiene a los isomorfismos que se corresponden con elementos de S_2. Sea $L = \mathbb{Z} \times N\mathbb{Z}$. Se comprueba inmediatamente que

$$\Gamma_0(N) = \{g \in \Gamma \mid Lg = L\}.$$

Para cada $g \in S_1$, sea $f \in T_1$ el isomorfismo que cumple (11.2). Definimos $\mathfrak{a}_g = f^{-1}[L]$. Obviamente $\mathfrak{a}_g$ es un subgrupo aditivo de $\mathbb{Z}[\rho]$ y $\mathfrak{a}_g$ es un ideal de $\mathbb{Z}[\rho]$ si y sólo si $g \in \Gamma_0(N)$.

En tal caso, $\mathbb{Z}[\rho]/\mathfrak{a}_g \cong \mathbb{Z}^2/L \cong \mathbb{Z}/N\mathbb{Z}$ (isomorfismo de grupos), de donde se sigue que $\mathrm{N}(\mathfrak{a}_g) = N$.

Recíprocamente, supongamos que $\mathfrak{a}$ es un ideal de $\mathbb{Z}[\rho]$ de norma N tal que el grupo $\mathbb{Z}[\rho]/\mathfrak{a}$ es cíclico. Vamos a ver que existe un $g \in S_1$ tal que $\mathfrak{a} = \mathfrak{a}_g$.

Por las propiedades de los $\mathbb{Z}$-módulos, existe una base v_1, v_2 de $\mathbb{Z}[\rho]$ y unos números naturales d_1, d_2 tales que $d_1 v_1$, $d_2 v_2$ es una base de $\mathfrak{a}$. El cociente es isomorfo a $\mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z}$, luego ha de ser $N = d_1 d_2$, $(d_1, d_2) = 1$. Tomemos r, $s \in \mathbb{Z}$ tales que $rd_1 - sd_2 = 1$. Entonces $w_1 = d_1 v_1 + d_2 v_2$, $w_2 = sv_1 + rv_2$

forman otra base de $\mathbb{Z}[\rho]$ y $w_1 = (d_1v_1) + (d_2v_2)$, $w_2' = sd_2(d_1v_1) + rd_1(d_2v_2)$ forman otra base de $\mathfrak{a}$, de modo que $w_2' = Nw_2$. El isomorfismo $f : \mathbb{Z}[\rho] \longrightarrow \mathbb{Z}^2$ que hace corresponder w_1, w_2 con la base canónica cumple $f[\mathfrak{a}] = L$.

En el caso en que $f \in T_2$ consideramos $\alpha : \mathbb{Z}^2 \longrightarrow \mathbb{Z}^2$ de matriz

$$\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

que cumple $\alpha[L] = L$, luego $f' = f \circ \alpha$ sigue cumpliendo $f[\mathfrak{a}] = L$ y si f está asociado a $g \in S_2$, entonces f' está asociado a $g^\alpha \in S_1$. Así pues, en cualquier caso $\mathfrak{a} = \mathfrak{a}_g$ para un cierto $g \in S_1$ (además $g \in \Gamma_0(N)$ porque $\mathfrak{a}$ es un ideal).

Ahora probamos que dos elementos de $S_1 \cap \Gamma_0(N)$ se corresponden con el mismo ideal si y sólo si son conjugados en $\Gamma_0(N)$.

Supongamos que g, $g' \in S_1 \cap \Gamma_0(N)$ se corresponden con el mismo ideal $\mathfrak{a}$. Sean f y f' los isomorfismos asociados. Entonces $\alpha = f^{-1} \circ f' \in \Gamma$, luego $f' = f \circ \alpha$. Aplicando (11.2) vemos que $g' = g^\alpha$ y si $u \in L$, entonces $u = f(v)$ con $v \in \mathfrak{a}$, luego $u\alpha = f(v)\alpha = f'(v) \in L$, así pues, $L\alpha \subset L$, e igualmente se prueba la otra inclusión, luego $L\alpha = L$ y, por consiguiente, $\alpha \in \Gamma_0(N)$.

Recíprocamente, supongamos que $g' = g^\alpha$ para cierto $\alpha \in \Gamma_0(N)$. Si f cumple (11.2) para g, entonces $f \circ \alpha$ lo cumple para g', luego

$$\mathfrak{a}_{g'} = (f \circ \alpha)^{-1}[L] = f^{-1}[\alpha^{-1}[L]] = f^{-1}[L] = \mathfrak{a}_g.$$

Con esto hemos probado que ν_3 es el número de ideales $\mathfrak{a}$ de $\mathbb{Z}[\rho]$ de norma N tales que el cociente $\mathbb{Z}[\rho]/\mathfrak{a}$ es cíclico. Vamos a ver que la última condición se da exactamente cuando $\mathfrak{a}$ no es divisible entre naturales mayores que 1.

Si $\mathfrak{a}$ es divisible entre un número natural, entonces es divisible entre un primo p, es decir, $\mathfrak{a} \subset (p)$. Entonces el cociente $\mathbb{Z}[\rho]/(p)$ es una imagen de $\mathbb{Z}[\rho]/\mathfrak{a}$, luego debería ser un grupo cíclico, pero como anillo tiene característica p, contradicción.

Supongamos ahora que $\mathfrak{a}$ no es divisible entre naturales mayores que 1. Recordemos que un primo racional p factoriza en $\mathbb{Z}[\rho]$ de una de las tres formas siguientes:

a) Si $(-3/p) = 0$ (o sea, $p = 3$) entonces $p = \mathfrak{p}^2$.

b) Si $(-3/p) = 1$ entonces $p = \mathfrak{p}\bar{\mathfrak{p}}$, $\mathfrak{p} \neq \bar{\mathfrak{p}}$.

c) Si $(-3/p) = -1$ entonces p se conserva primo.

La descomposición de $\mathfrak{a}$ en ideales primos ha de ser de la forma $\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$, donde cada $\mathfrak{p}_i$ tiene norma prima p_i con los p_i distintos dos a dos (es decir, $\mathfrak{a}$ no puede ser divisible entre dos primos conjugados distintos). Además, si $p_i = 3$, entonces $e_i = 1$. Por el teorema chino del resto,

$$\mathbb{Z}[\rho]/\mathfrak{a} \cong (\mathbb{Z}[\rho]/\mathfrak{p}_1^{e_1}) \times \cdots \times (\mathbb{Z}[\rho]/\mathfrak{p}_r^{e_r})$$

y cada factor es un grupo cíclico de orden $p_i^{e_i}$ y como anillo tiene característica $p_i^{e_i}$, ya que $p_i^{e_i}$ es el menor natural divisible entre $p_i^{e_i}$. Esto equivale a que el cociente es cíclico y por lo tanto el producto también lo es.

El análisis que acabamos de realizar nos dice también cuántos ideales hay en las condiciones indicadas. Si N es divisible entre 9, todo ideal de norma N ha de ser divisible entre 3, luego no existen ideales de esta forma. Si N es divisible entre un primo que se conserva primo en $\mathbb{Z}[\rho]$, entonces tampoco existe ninguno. En caso contrario, para cada primo p que divida a N con exponente e y $p = \mathfrak{p}\bar{\mathfrak{p}}$, tenemos dos posibilidades: que $\mathfrak{a}$ contenga a $\mathfrak{p}^e$ o que contenga a $\bar{\mathfrak{p}}^e$. Ahora es claro que la fórmula del enunciado para ν_3 nos da el número de ideales.

El cálculo de ν_2 es análogo al de ν_3, considerando ahora las matrices

$$\tau = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \tau^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

y el anillo $\mathbb{Z}[i]$ en lugar de $\mathbb{Z}[\rho]$. ∎

El teorema 11.33 nos permite ahora calcular el género de cualquier superficie modular $X_0(N)$. La tabla siguiente contiene los primeros casos:

N	1	2	3	4	5	6	7	8	9	10	11	12
g	0	0	0	0	0	0	0	0	0	0	1	0
N	13	14	15	16	17	18	19	20	21	22	23	24
g	0	1	1	0	1	0	1	1	1	2	2	1
N	25	26	27	28	29	30	31	32	33	34	35	36
g	0	2	1	2	2	3	2	1	1	3	3	1

En general si $p > 3$ es primo, es fácil ver que el género de $X_0(p)$ es

$$g = \begin{cases} n-1 & \text{si } p = 12n+1, \\ n & \text{si } p = 12n+5 \text{ o } p = 12n+7, \\ n+1 & \text{si } p = 12n+11. \end{cases}$$

Terminamos la sección comentando un resultado clásico que no vamos a demostrar: toda superficie de Riemann de género $g \geq 2$ es conformemente equivalente a una superficie modular determinada por un grupo hiperbólico Γ, es decir, un grupo sin elementos parabólicos o elípticos. Geométricamente, Γ es un grupo de traslaciones del plano hiperbólico que tiene por dominio fundamental un polígono convexo (cuyos lados son rectas hiperbólicas). Los trasladados del polígono por Γ "teselan" el plano hiperbólico igual que los trasladados de un paralelogramo de periodos de un toro complejo teselan $\mathbb{C}$.

11.6 La medida de una superficie modular

Para terminar con las propiedades generales de las superficies modulares veremos que en cada una de ellas es posible definir una medida de forma natural.

El punto de partida es la geometría hiperbólica. Puede probarse que el elemento de longitud hiperbólico es $ds = y^{-1}|dz|$, donde $z = x + iy$. Es decir, que si definimos la longitud de un segmento hiperbólico como la integral de ds sobre dicho segmento, se cumplen los mismos resultados básicos que en el caso de la geometría euclídea. En particular, las transformaciones de $\mathrm{LE}(2,\mathbb{R})$ conservan las longitudes. Notemos que ds es el elemento de longitud más simple que hace que un segmento de longitud euclídea fija sea más largo cuanto más cerca esté del eje real.

Similarmente, si los lados de un rectángulo infinitesimal tienen longitudes euclídeas dx y dy, entonces, sus longitudes hiperbólicas son $y^{-1}dx$ e $y^{-1}dy$, luego su área hiperbólica es $y^{-2}dxdy$. La forma diferencial $\omega = y^{-2}\,dx\wedge dy$ es el elemento de área hiperbólico y, en particular, es invariante por las isometrías de $\mathrm{LE}(2,\mathbb{R})$. Vamos a probar directamente este hecho, junto con algunos resultados técnicos adicionales:

Teorema 11.41 *Sea η la forma diferencial en H dada por $\eta = y^{-1}\,dz$ y sea*

$$\alpha = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LE}(2,\mathbb{R}).$$

Entonces:

a) $\alpha^\sharp(\eta) - \eta = -2i\,d\log(cz+d)$,

b) $d\eta = y^{-2}\,dx\wedge dy = (i/2y^2)\,dz\wedge d\bar{z}$,

c) $\alpha^\sharp(d\eta) = d\eta$.

Demostración: Tenemos que

$$\alpha^\sharp(\eta) = (\operatorname{Im}\alpha)^{-1}\,d\alpha = \frac{|cz+d|^2}{y}\,\frac{dz}{(cz+d)^2} = \frac{c\bar{z}+d}{cz+d}\,\eta,$$

luego

$$\alpha^\sharp(\eta) - \eta = \left(\frac{c\bar{z}+d}{cz+d} - 1\right)\eta = -\frac{2yi}{cz+d}\,\eta = -2i\,d\log(cz+d).$$

El apartado b) es un cálculo directo y c) se sigue de a):

$$\alpha^\sharp(d\eta) = d(\alpha^\sharp(\eta)) = d\eta.$$

∎

Definimos la *medida hiperbólica* en H como

$$m(A) = \int_A y^{-2}\,dxdy,$$

donde A es cualquier subconjunto de H medible Lebesgue. Del teorema anterior se sigue que esta medida es invariante por $\mathrm{LE}(2,\mathbb{R})$, es decir, si $\alpha \in \mathrm{LE}(2,\mathbb{R})$, entonces

$$m(\alpha[A]) = \int_{\alpha[A]} d\eta = \int_A \alpha^\sharp(d\eta) = \int_A \eta = m(A).$$

Esta invarianza nos permite elevar la medida m a cualquier superficie modular. En efecto, sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ que determine una superficie modular H^*/Γ.

Para cada punto $z \in H^*$ llamamos $\Gamma_z = \{g \in \Gamma \mid zg = z\}$. Sabemos que existe un entorno abierto U_z de z tal que

$$\Gamma_z = \{g \in \Gamma \mid U_z g \cap U_z \neq \varnothing\}.$$

Este entorno no puede contener puntos elípticos o parabólicos distintos del propio z. Podemos exigir que, para todo $g \in \Gamma$, se cumpla $U_{zg} = U_z g$. Sea $p : H^* \longrightarrow H^*/\Gamma$ la proyección natural. Como es abierta, tenemos que $p[U_z]$ es un entorno de $[z]$ en H^*/Γ.

Si z no es elíptico ni parabólico, entonces $p|_{U_z}$ es inyectiva y, para cada $A \subset p[U_z]$, podemos definir $\mu(A) = m(p|_{U_z}^{-1}[A])$. Esta definición no depende de la elección del representante z de la clase $[z]$, pues si $z' = gz$, entonces $p|_{U_z'}^{-1}[A] = g[p|_{U_z}^{-1}[A]]$, y ambos conjuntos tienen la misma medida.

Si z es un punto elíptico de orden e, entonces cada punto de $p[U_z]$ tiene e exactamente antiimágenes en U_z, por lo que definimos la medida de un conjunto $A \subset p[U_z]$ como $\mu(A) = m(p|_{U_z}^{-1}[A])/e$. Como en el caso anterior, esta definición no depende de la elección de z.

Si z es un punto parabólico, tomamos $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $zg = \infty$. Así, ∞ es un punto parabólico de Γ^g. El grupo Γ^g_∞ está generado por una traslación $z \mapsto z + h$ y podemos tomar U_z de modo que $U_z g = \{w \in H \mid \operatorname{Im} w > c\} \cup \{\infty\}$. Para cada $A \subset p[U_z]$, definimos

$$\mu(A) = m(Ag \cap \{w \in H \mid \operatorname{Im} w > 0,\ 0 \leq \operatorname{Re} w < h\}).$$

La invarianza de m implica una vez más que esta definición no depende de la elección de z o g. Notemos que todas las medidas $\mu(A)$ que hemos definido son finitas. Esto es evidente en todos los casos salvo en el último, donde sólo hay que comprobar que la integral

$$\int_{[0,h]\times[c,+\infty]} \frac{dxdy}{y^2}$$

es finita.

De este modo, tenemos cubierta la superficie modular H^*/Γ por una familia de abiertos $p[U_z]$, en cada uno de los cuales tenemos definida una medida finita μ_z. Se comprueba que estas medidas son compatibles dos a dos, en el sentido de que si $A \subset p[U_z] \cap p[U_{z'}]$, entonces $\mu_z(A) = \mu_{z'}(A)$. Por ejemplo, si z es un punto elíptico de orden e y z' no es elíptico ni parabólico, tomamos un punto $z_0 \in p|_{U_{z'}}^{-1}[A]$. Entonces $p(z_0)$ tiene e antiimágenes en $p|_{U_z}^{-1}[U_z]$, que serán de la forma $z_0 g_i$, con $g_i \in \Gamma$, $i = 1, \ldots, e$. Claramente

$$p|_{U_z}^{-1}[A] = \bigcup_{i=1}^{e} p|_{U_{z'}}^{-1}[A] g_i,$$

con lo que $m(p|_{U_z}^{-1}[A]) = e\, m(p|_{U_{z'}}^{-1}[A])$, luego $\mu_z(A) = \mu_{z'}(A)$.

Como la superficie H^*/Γ es compacta, podemos cubrirla con un número finito de abiertos $p[U_z]$, y es claro que las medidas μ_z determinan una única medida μ en H^*/Γ que las extiende a todas (podemos tomar una partición de la unidad subordinada al cubrimiento y definir de forma obvia la integral respecto de μ de una función continua). La medida μ extiende —de hecho— a todas las medidas μ_z (incluso a las que no hemos tomado en el cubrimiento finito, pues siempre podríamos haber añadido un abierto más). Obviamente μ es una medida finita, pues H^*/Γ está cubierto por un número finito de abiertos de medida finita. Ahora es claro el teorema siguiente:

Teorema 11.42 *Si H^*/Γ es una superficie modular, existe una única medida regular μ en H^*/Γ caracterizada por que la proyección natural $p : H \longrightarrow H^*/\Gamma$ restringida a cualquier abierto U donde sea inyectiva transforma μ en la medida hiperbólica m.*

Si $\Gamma = \mathrm{LE}(2,\mathbb{Z})$ y D es su dominio fundamental, es fácil ver que

$$\mu(H^*/\Gamma) = \int_D \frac{dxdy}{y^2} = \frac{\pi}{3}.$$

A partir de aquí, el teorema siguiente nos permite calcular la medida de las superficies modulares que hemos estudiado en las secciones precedentes:

Teorema 11.43 *Sea Γ un subgrupo discreto de $\mathrm{LE}(2,\mathbb{R})$ que defina una superficie modular H^*/Γ y sea $\Gamma' \leq \Gamma$ un subgrupo de índice finito $|\Gamma : \Gamma'| = m$. Entonces $\mu(H^*/\Gamma') = m\mu(H^*/\Gamma)$.*

Demostración: Consideramos la aplicación natural $\phi : H^*/\Gamma' \longrightarrow H^*/\Gamma$, que tiene grado m. Podemos triangular H^*/Γ de modo que las antiimágenes por ϕ de los triángulos formen una triangulación de H^*/Γ', y cada triángulo de H^*/Γ dé lugar a m triángulos en H^*/Γ' (los detalles son los mismos que los de la prueba de la fórmula de Hurwitz para funciones holomorfas entre superficies de Riemann). Si exigimos que los puntos elípticos y parabólicos de cualquiera de las dos superficies sean vértices de la triangulación y que los triángulos sean suficientemente pequeños, entonces cada una de las m antiimágenes de un triángulo tiene la misma medida que el triángulo de partida (porque sus proyecciones en H son triángulos equivalentes respecto a Γ), luego la medida de H^*/Γ' es m veces la de H^*/Γ. ∎

Capítulo XII

Funciones modulares

Ya estamos en condiciones de definir las funciones modulares. La idea básica es que las funciones modulares son las funciones meromorfas sobre una superficie modular, pero hemos de bajarlas a funciones sobre el semiplano H, y la cuestión es entonces qué condiciones garantizan que una función meromorfa en H se eleva a una función meromorfa en una superficie dada. Dichas condiciones serán dos: la invarianza por el grupo de transformaciones que define a la superficie y otra en términos de series de Fourier que asegure la meromorfía en los puntos parabólicos. No obstante, con todo esto tendremos un concepto de función modular que incluirá a funciones como $j(\tau)$ pero no a otras como $\Delta(\tau)$. Por ello después daremos una definición más general que incluirá a las demás funciones cuyo carácter modular habíamos anunciado en el capítulo X.

12.1 Funciones modulares de grado cero

Definición 12.1 Una *función automorfa* asociada a un subgrupo discreto Γ de $\mathrm{LE}(2,\mathbb{R})$ que defina una superficie modular H^*/Γ es una función $f : H \longrightarrow \mathbb{C}^\infty$ tal que existe una función meromorfa $\bar{f} : H^*/\Gamma \longrightarrow \mathbb{C}^\infty$ de manera que $f = p \circ \bar{f}$, donde $p : H \longrightarrow H^*/\Gamma$ es la proyección natural.

Las funciones automorfas asociadas a subgrupos de índice finito en el grupo modular $\mathrm{LE}(2,\mathbb{Z})$ se llaman *funciones modulares.* Cuando no se especifica el grupo se entiende que es todo el grupo modular.

Evidentemente, una función modular $f : H \longrightarrow \mathbb{C}^\infty$ respecto de un grupo Γ es una función meromorfa en H invariante por Γ, es decir, que verifica

$$f(zg) = \bar{f}(p(zg)) = \bar{f}(p(z)) = f(z)$$

para todo $z \in H$ y todo $g \in \Gamma$.

Recíprocamente, toda función f meromorfa en H e invariante por Γ induce una función meromorfa $\bar{f} : H/\Gamma \longrightarrow \mathbb{C}^\infty$ tal que $f = p \circ \bar{f}$.

En efecto, observemos en primer lugar que $\bar{f}$ es continua en H/Γ, pues si U es abierto en $\mathbb{C}^\infty$ entonces $\bar{f}^{-1}[U] = p[f^{-1}[U]]$ es abierto en H/Γ. También es claro que $\bar{f}$ es meromorfa en todo punto no elíptico, porque al componerla con la inversa de p en un entorno del punto (que es una carta) obtenemos f, que es meromorfa. Por último, la lectura de $\bar{f}$ en un entorno de un punto elíptico (respecto a cartas adecuadas de H/Γ y $\mathbb{C}^\infty$) es una función de un disco de centro 0 en un entorno acotado de 0 holomorfa en todo el disco salvo a lo sumo en 0, pero continua en 0, luego también es holomorfa en 0 y, por consiguiente $\bar{f}$ es meromorfa en el punto elíptico.

La función f será modular si y sólo si $\bar{f}$ se extiende a una función meromorfa en los puntos parabólicos, es decir, si y sólo si éstos son singularidades evitables o polos de $\bar{f}$. (Notemos que una función $\bar{f} : H^*/\Gamma \longrightarrow \mathbb{C}^\infty$ que satisfaga la definición de función modular ha de coincidir con la $\bar{f}$ que estamos considerando en el abierto H/Γ).

Vamos a ver cómo expresar esto en términos de f. En primer lugar consideramos el caso del punto $p(\infty)$. El estabilizador $\bar{\Gamma}_\infty$ está generado por una traslación $z \mapsto z+h$, para un cierto natural $h > 0$. Por consiguiente la invarianza de f por Γ implica en particular que f tiene periodo h, es decir, $f(z+h) = f(z)$, para todo $z \in H$. La función $q(z) = e^{2\pi i z/h}$ tiene también periodo h y transforma H en el disco unidad $D(0,1) \setminus \{0\}$, luego podemos definir $f^*(q) = f(z)$, donde z es cualquier antiimagen de q por la función q_z. Dos antiimágenes cualesquiera se diferencian en un múltiplo de h, luego $f(z)$ es el mismo para ambas. Tenemos así una función $f^* : D(0,1) \setminus \{0\} \longrightarrow \mathbb{C}^\infty$ caracterizada por que $f^*(q(z)) = f(z)$ para todo $z \in H$. La derivada de $q(z)$ no se anula en ningún punto, luego q se restringe a una transformación conforme en un entorno de cada punto de H, de donde se sigue inmediatamente que f^* es una función meromorfa en el disco unidad con una singularidad en 0.

Para que $\bar{f}$ sea meromorfa en $p(\infty)$ necesitamos en primer lugar que $p(\infty)$ sea una singularidad aislada de $\bar{f}$, lo cual equivale a que $\bar{f}$ no tenga polos en un entorno reducido de ∞, lo cual equivale a que f no tenga polos en un semiplano $\operatorname{Im} z > R$, lo cual equivale a que f^* no tenga polos en un entorno de 0, lo cual equivale a que f^* tenga una singularidad aislada en 0.

Admitiendo que es así, vamos a ver que la singularidad de f^* en 0 es del mismo orden que la de $\bar{f}$ en $p(\infty)$, de modo que $\bar{f}$ será meromorfa en $p(\infty)$ si y sólo si la función f^* es meromorfa en 0.

En efecto, ello se debe a que la función $\bar{q}$ dada por $p(z) \mapsto q(z)$ (con el convenio de que $p(\infty) \mapsto 0$) es una carta de H^*/Γ alrededor de $p(\infty)$, y si $q \neq 0$ está en el rango de $\bar{q}$, digamos $q = \bar{q}(p(z)) = q(z)$, entonces

$$(\bar{q}^{-1} \circ \bar{f})(q) = \bar{f}(p(z)) = f(z) = f^*(q).$$

Así pues, f^* es la lectura de $\bar{f}$ alrededor de $p(\infty)$, luego el orden de f^* en 0 es, por definición, el mismo que el de $\bar{f}$ en $p(\infty)$.

En cualquier caso, la función f^* admite un desarrollo en serie de Laurent de la forma

$$f^*(q) = \sum_{n=-\infty}^{+\infty} c_n q^n, \qquad 0 < |q| < r,$$

para ciertos coeficientes $c_n \in \mathbb{C}$, lo cual equivale a que f admite un (único) desarrollo en serie de Fourier de la forma

$$f(z) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi n i z/h}, \qquad \operatorname{Im} z > R,$$

con los mismos coeficientes c_n. Concluimos que $\bar{f}$ es meromorfa en $p(\infty)$ si y sólo si f es holomorfa en un semiplano $\operatorname{Im} z > r$ y existe un $k \in \mathbb{Z}$ tal que los coeficientes de Fourier c_n de f son nulos para $n < k$.

Si Γ tiene otros puntos parabólicos s no equivalentes a ∞, podemos aplicar el criterio anterior tomando $\alpha \in \mathrm{LE}(2,\mathbb{Z})$ tal que $s\alpha = \infty$. Es claro que α induce una transformación conforme $\bar{\alpha} : H^*/\Gamma \longrightarrow H^*/\Gamma^\alpha$ mediante $p(z) \mapsto p_\alpha(z\alpha)$. En particular $p(s) \mapsto p_\alpha(\infty)$. Si f es una función meromorfa en H invariante por Γ, entonces $f_\alpha(z) = f(z\alpha^{-1})$ es una función meromorfa en H invariante por Γ^α y $\bar{f} = \bar{\alpha} \circ \bar{f}_\alpha$. Por consiguiente, $\bar{f}$ es meromorfa en $p(s)$ si y sólo si $\bar{f}_\alpha$ es meromorfa en $\bar{\alpha}(p(s)) = p_\alpha(\infty)$, si y sólo si la serie de Fourier de f_α tiene coeficientes nulos para índices suficientemente pequeños. En definitiva:

Teorema 12.2 *Una función $f : H \longrightarrow \mathbb{C}^\infty$ es modular respecto a un subgrupo Γ de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ si y sólo si:*

a) f es meromorfa en H,

b) para todo $g \in \Gamma$ y todo $z \in H$ se cumple $f(zg) = f(z)$,

c) para todo punto parabólico $s \in \mathbb{Q} \cup \{\infty\}$ y toda función $\alpha \in \mathrm{LE}(2,\mathbb{Z})$ tal que $s\alpha = \infty$, la función $f_\alpha(z) = f(z\alpha^{-1})$ es holomorfa en un semiplano $\operatorname{Im} z > R$ y su desarrollo en serie de Fourier es de la forma

$$f_\alpha(z) = \sum_{n=k}^{\infty} c_n e^{2\pi n i z/h}, \qquad \operatorname{Im} z > R, \quad k \in \mathbb{Z}.$$

(En realidad basta con que esto se cumpla para un punto parabólico de cada órbita y para un α fijo.)

Ahora ya es evidente que la función modular de Klein $j(\tau)$ es ciertamente una función modular: es holomorfa, invariante por el grupo modular y su desarrollo en serie de Fourier calculado en 10.32 muestra que tiene un polo simple en ∞ con residuo 1. El teorema siguiente recoge las propiedades básicas de j. Como ya explicamos en el capítulo X, de él se sigue inmediatamente el teorema de uniformización, cuya prueba teníamos pendiente.

Teorema 12.3 *Si j es la función modular de Klein, entonces el cuerpo de todas las funciones modulares respecto al grupo modular $\Gamma = \mathrm{LE}(2,\mathbb{Z})$ es $\mathbb{C}(j)$. La función j toma todos los valores complejos. Si $\rho = e^{2\pi i/3}$, se cumple que $j(\rho) = 0$ y $j(i) = 1718$. Además ρ es un cero triple de j, mientras que i es un cero doble de $j - 1728$.*

DEMOSTRACIÓN: Sabemos que la función j induce en la superficie modular $X(1)$ una función meromorfa con un único polo simple en ∞. Esto implica que $j : X(1) \longrightarrow \mathbb{C}^\infty$ es una aplicación holomorfa de grado 1. (En particular es suprayectiva, luego j toma todos los valores complejos.)

Si llamamos K al cuerpo de todas las funciones meromorfas sobre $X(1)$, entonces el divisor de j es de la forma $(j) = \mathfrak{p}\mathfrak{q}^{-1}$, donde $\mathfrak{q}$ es el divisor correspondiente a ∞. Por otra parte, el divisor de j en el cuerpo $k = \mathbb{C}(j)$ tiene la misma forma, luego concluimos que $\mathfrak{q}$ es el primo infinito de k y así,

$$1 = \operatorname{grad}_K \mathfrak{q} = |K : k| \operatorname{grad}_k \mathfrak{q} = |K : k|.$$

En los ejemplos al final de la sección 10.2 hemos calculado los valores de $j(i)$ y $j(\rho)$. Para calcular el orden de j en ρ hemos de tener cuidado: la función inducida por j en la superficie modular $X(1)$ tiene orden 1 en el punto $p(\rho)$, pero esto no significa que j tenga orden 1 en ρ. Por el contrario, sabemos que la proyección $p : H \longrightarrow X(1)$ toma tres veces un mismo valor en cualquier entorno suficientemente pequeño de ρ, luego j toma tres veces un mismo valor en cualquier entorno suficientemente pequeño de ρ. Esto implica claramente que el orden de j en ρ es 3. Similarmente se razona con $j - 1728$. ∎

A partir del desarrollo en serie de Fourier de la función modular es fácil deducir ahora otras propiedades destacables. Por ejemplo, el hecho de que sus coeficientes sean reales implica que $j(-\bar{\tau}) = \overline{j(\tau)}$, es decir, que j transforma puntos simétricos respecto al eje imaginario en puntos simétricos respecto al eje real. Más aún, es fácil ver que la mitad izquierda del dominio fundamental descrito en 10.34 se corresponde conformemente con el semiplano H, mientras que la mitad derecha se corresponde con el semiplano complementario.

Es fácil determinar las funciones modulares de los grupos $\Gamma_0(N)$:

Teorema 12.4 *El cuerpo de las funciones modulares respecto al grupo $\Gamma_0(N)$ es $\mathbb{C}(j, j_N)$, donde $j_N(\tau) = j(N\tau)$.*

DEMOSTRACIÓN: Consideremos

$$\alpha = \begin{pmatrix} \sqrt{N} & 0 \\ 0 & 1/\sqrt{N} \end{pmatrix} \in \mathrm{LE}(2, \mathbb{R}),$$

que induce la transformación $\tau\alpha = N\tau$. Sabemos que $\alpha : H^* \longrightarrow H^*$ es un homeomorfismo. Además, si dos puntos τ, $\tau' \in H^*$ son equivalentes respecto a $\Gamma_0(N)$, entonces $\tau\alpha$ y $\tau'\alpha$ son equivalentes respecto a $\Gamma(1)$. En efecto, tenemos que

$$\tau' = \frac{a\tau + b}{cN\tau + d}, \qquad a, b, c, d \in \mathbb{Z}, \quad ad - bcN = 1$$

luego

$$\tau'\alpha = \frac{aN\tau + Nb}{cN\tau + d} = (N\tau)g = \tau\alpha g,$$

donde

$$g = \begin{pmatrix} a & Nb \\ c & d \end{pmatrix} \in \Gamma(1).$$

Por consiguiente α induce una aplicación continua $\bar{\alpha} : X_0(N) \longrightarrow X_0(1)$, que claramente es holomorfa salvo a lo sumo en los puntos elípticos e hiperbólicos, pero, como éstos son puntos aislados, concluimos que $\bar{\alpha}$ es holomorfa en toda la superficie modular $X_0(N)$. Ahora es claro que $\bar{j}_N = \bar{\alpha} \circ \bar{j}$, luego j_N es ciertamente una función modular respecto de $\Gamma_0(N)$.

Según 11.32, sabemos que la aplicación $\phi : X_0(N) \longrightarrow X_0(1)$ tiene grado $n = |\bar{\Gamma}_0(1) : \bar{\Gamma}_0(N)|$, luego induce un monomorfismo $\mathcal{M}(X_0(1)) \longrightarrow \mathcal{M}(X_0(N))$ a través del cual las funciones modulares respecto a $\Gamma_0(N)$ se identifican con una extensión de grado n del cuerpo de las funciones modulares respecto a $\Gamma_0(1)$. Más detalladamente, tenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccc} \mathcal{M}(X_0(1)) & \xrightarrow{\bar{\phi}} & \mathcal{M}(X_0(N)) \\ \uparrow & & \uparrow \\ \mathbb{C}(j) & \longrightarrow & \mathbb{C}(j, j_N) \end{array}$$

La flecha horizontal inferior es la inclusión, la flecha vertical izquierda es un isomorfismo y la derecha un monomorfismo. Si probamos que $\mathbb{C}(j, j_N)$ tiene grado n sobre $\mathbb{C}(j)$, la flecha vertical derecha será un isomorfismo y el teorema estará probado.

Sea $\bar{\Gamma}_0(1)/\bar{\Gamma}_0(N) = \{g_1\bar{\Gamma}_0(N), \ldots, g_n\bar{\Gamma}_0(N)\}$ y consideremos las funciones $f_i(\tau) = j_N(\tau g_i)$ (holomorfas en H). Podemos suponer que $g_1 = 1$, y por consiguiente $f_1 = j_N$.

Sea $S(x_1, \ldots, x_n)$ un polinomio simétrico con coeficientes en $\mathbb{C}$. La función $S(f_1, \ldots, f_n)$ es holomorfa en H e invariante por el grupo modular $\Gamma_0(1)$, pues si $g \in \Gamma_0(1)$ se cumple que

$$S(f_1, \ldots, f_n)(\tau g) = S(j_N(\tau g g_1), \ldots, j_N(\tau g g_n))$$

$$= S(j_N(\tau g_1), \ldots, j_N(\tau g_n)) = S(f_1, \ldots, f_n)(\tau).$$

Por otra parte, como j_N es meromorfa en el punto parabólico $s = \infty g_i$, tenemos que f_i tiene un desarrollo de Fourier en ∞ con coeficientes nulos para índices pequeños, y lo mismo es válido para $S(f_1, \ldots, f_n)$ (al transformar H en el disco unidad sin el origen, estamos usando que una combinación polinómica de funciones meromorfas es meromorfa). Concluimos que $S(f_1, \ldots, f_n)$ es una función modular, luego $S(f_1, \ldots, f_n) \in \mathbb{C}(j)$. En particular, el polinomio

$$F(Y) = \prod_{i=1}^{n} (Y - f_i)$$

tiene sus coeficientes en $\mathbb{C}(j)$, pues todos ellos son polinomios simétricos en $f_1, \ldots, f_n$. Más aún, son funciones holomorfas en H, ya que todas las funciones f_i lo son, luego en realidad F tiene sus coeficientes en $\mathbb{C}[j]$. Así pues, podemos expresarlo en la forma $F(j, Y)$, donde $F[X, Y] \in \mathbb{C}[X, Y]$.

Si probamos que las funciones f_i son distintas dos a dos y que además son conjugadas sobre $\mathbb{C}(j)$, tendremos que $F(j, Y)$ será un polinomio irreducible de

grado n, luego será el polinomio mínimo de j_N sobre $\mathbb{C}(j)$ y el teorema quedará probado.

Sea $G(j,Y)$ el polinomio mínimo de $f_1 = j_N$ sobre $\mathbb{C}(j)$, de modo que para todo $\tau \in H$ se cumple $G(j(\tau), j_N(\tau)) = 0$. Teniendo en cuenta que j es invariante por g_i, vemos que

$$G(j(\tau), f_i(\tau)) = G(j(\tau g_i), j_N(\tau g_i)) = G(j(\tau'), j_N(\tau')) = 0,$$

luego $G(j, f_i) = 0$. Esto prueba que f_1 es conjugada de cada f_i.

Supongamos ahora que $f_i = f_j$, es decir, que $j(N(\tau g_i)) = j(N(\tau g_j))$, para todo $\tau \in H$. Esto significa que los puntos $N(\tau g_i)$ y $N(\tau g_j)$ son equivalentes respecto del grupo $\Gamma_0(1)$. En principio, el elemento del grupo que los relaciona depende de τ, pero tiene que haber un $g \in \Gamma_0(1)$ tal que $(N(\tau g_i))g = N(\tau g_j)$ para un conjunto no numerable de puntos τ. Por prolongación analítica la igualdad vale para todo τ.

Si llamamos

$$\alpha = \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix},$$

tenemos que $\tau g_i \alpha g = \tau g_j \alpha$, luego $g_i \alpha g = \pm g_j \alpha$ o, también, $g_i^{-1} g_j = \pm g^{\alpha^{-1}}$. Una simple comprobación muestra que el miembro derecho tiene la entrada c divisible entre N y, como el miembro izquierdo está en $\mathrm{LE}(2,\mathbb{Z})$, concluimos que $g_i^{-1} g_j \in \Gamma_0(N)$, lo cual sólo es posible si $i = j$. ∎

Es claro que las funciones f_i no dependen de la elección de los representantes g_i de las clases módulo $\bar{\Gamma}_0(N)$, por lo que el polinomio $F(X,Y)$ tampoco depende de dicha elección y está canónicamente determinado por la superficie modular $X_0(N)$. Dedicamos la sección siguiente a estudiarlo con más detalle.

12.2 La ecuación modular

Sabemos que el cuerpo de las funciones meromorfas sobre una superficie de Riemann es un cuerpo de funciones algebraicas, pero en general no es fácil determinar una ecuación que lo determine. Sin embargo, acabamos de ver que para el caso de la superficie modular $X_0(N)$ sí sabemos encontrar explícitamente unos generadores de su cuerpo de funciones meromorfas, a saber j y j_N, junto con una ecuación polinómica irreducible $F(j, j_N) = 0$ que los relaciona.

Definición 12.5 Sea $N > 0$ un número natural y sea

$$\bar{\Gamma}_0(1)/\bar{\Gamma}_0(N) = \{g_1\bar{\Gamma}_0(N), \ldots, g_n\bar{\Gamma}_0(N)\}.$$

El *polinomio modular* de orden N es el polinomio $F_N(X,Y) \in \mathbb{C}[X,Y]$ caracterizado por que $F_N(j,Y) = \prod_{i=1}^{n} (Y - f_i)$, donde $f_i(\tau) = j_N(\tau g_i)$.

Hemos visto que $F_N(j,Y)$ es el polinomio mínimo de j_N sobre $\mathbb{C}(j)$, luego la *ecuación modular* $F_N(X,Y)=0$ determina salvo isomorfismo el cuerpo de funciones modulares respecto a $\Gamma_0(N)$.

Es posible dar algoritmos explícitos para calcular los polinomios modulares, pero los resultados no son muy manejables. Por ejemplo, puede probarse que

$$F_2(X,Y)=X^3+Y^3-X^2Y^2+1.488XY(X+Y)-162.000(X^2+Y^2)$$

$$+40.773.375XY+8.748.000.000(X+Y)-157.464.000.000.000.$$

Sin embargo, es posible obtener mucha información valiosa sobre estos polinomios. En esta sección probaremos entre otras cosas que F_N tiene coeficientes enteros y que es un polinomio simétrico en X e Y. Llamemos

$$\alpha=\begin{pmatrix} N & 0\\ 0 & 1\end{pmatrix}.$$

De este modo, las funciones f_i son $f_i(\tau)=j_N(\tau g_i)=j(\tau g_i\alpha)=j(\tau g_i\alpha g_i')$, para cualquier $g_i'\in \mathrm{LE}(2,\mathbb{Z})$. Vamos a ver que podemos elegir g_i' de modo que $\alpha_i=g_i\alpha g_i'$ sea de la forma

$$\alpha_i=\begin{pmatrix} a & 0\\ b & d\end{pmatrix},$$

con $(a,b,d)=1$, $ad=N$, $a,\ d\geq 0$, $0\leq b<d$. Más en general, tenemos el teorema siguiente:

Teorema 12.6 *Dada una matriz*

$$\alpha=\begin{pmatrix} u & w\\ v & x\end{pmatrix}$$

tal que $u,v,w,x\in\mathbb{Z}$, $(u,v,w,x)=1$, *existe* $g\in \mathrm{LE}(2,\mathbb{Z})$ *tal que*

$$\alpha g=\begin{pmatrix} a & 0\\ b & d\end{pmatrix},$$

con $(a,b,d)=1$, $a,\ d\geq 0$, $0\leq b<d$.

Demostración: Observemos que si multiplicamos

$$\begin{pmatrix} u & w\\ v & x\end{pmatrix}\begin{pmatrix} p & r\\ q & s\end{pmatrix}=\begin{pmatrix} up+wq & ur+ws\\ vp+xq & vr+xs\end{pmatrix},$$

donde $(u,v,w,x)=1$ y $ps-qr=1$, entonces el producto tiene también entradas primas entre sí. En efecto, un primo que las dividiera a todas dividiría a

$$p(ur+ws)-r(up+wq)=w(ps-qr)=w,$$

e igualmente se prueba que dividiría a las cuatro entradas del primer factor.

Así pues, la condición $(a,b,d)=1$ se va a cumplir necesariamente. Multiplicando si es preciso por -1 conseguimos $a,\ d\geq 0$. Si observamos además que

$$\begin{pmatrix} a & 0\\ b & d\end{pmatrix}\begin{pmatrix} 1 & 0\\ k & 1\end{pmatrix}=\begin{pmatrix} a & 0\\ b+dk & d\end{pmatrix},$$

concluimos que nos bastará conseguir que $c=0$.

Si $w=0$ no queda nada que probar y si $u=0$ tomamos

$$g=\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}.$$

Supongamos que $u\neq 0\neq w$ y sea $e=(u,w)$, de modo que $(w/e,u/e)=1$, luego existen enteros r y s tales que $ru/e+sw/e=1$. Tomamos

$$g=\begin{pmatrix} r & -w/e\\ s & u/e\end{pmatrix}.$$

Ciertamente, $g\in \mathrm{LE}(2,\mathbb{Z})$ y αg cumple $c=-uw/e+uw/e=0$. ∎

El número de matrices α_i en las condiciones del teorema anterior con determinante $N=ad$ para un $d\mid N$ dado es $(d/e)\phi(e)$, donde $e=(d,N/d)$. En efecto, $(a,b,d)=1$ equivale a $(b,e)=1$, y cada uno de los $\phi(e)$ enteros $0\leq b_0<e$ primos con e da lugar a d/e valores posibles $b=b_0+ke$, con $0\leq k<d/e$. Así pues, el número total de matrices posibles α_i es

$$n(N)=\sum_{d\mid N}\frac{d}{e}\phi(e).$$

Es inmediato comprobar que esta función es multiplicativa, y para $N=p^r$ se reduce a

$$n(p^r)=1+p^r+\sum_{i=1}^{r-1}\frac{p^i}{e}e\left(1-\frac{1}{p}\right)=1+p^r+\sum_{i=1}^{r-1}(p^i-p^{i-1})$$

$$=p^r+p^{r-1}=p^r\left(1+\frac{1}{p}\right).$$

En total,

$$n(N)=N\prod_{p\mid N}\left(1+\frac{1}{p}\right).$$

Ahora bien, según 11.40, este número es precisamente $n=|\bar\Gamma_0(1):\bar\Gamma_0(N)|$, luego concluimos que los α_i recorren todas las matrices de la forma indicada, con lo que tenemos una expresión explícita para el polinomio modular:

$$F_N(j,Y)=\prod\left(Y-j\left(\tfrac{a\tau+b}{d}\right)\right),\tag{12.1}$$

donde a, b, d recorren todos los números naturales tales que $ad=N$, $0\leq b<d$, $(a,b,d)=1$.

Sustituyendo en el desarrollo de Fourier

$$j(\tau) = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} c(n)e^{2n\pi i\tau},$$

vemos que, llamando $\zeta_d = e^{2\pi i/d}$,

$$f_i(\tau) = j\left(\frac{a\tau+b}{d}\right) = \zeta_d^{-b}e^{-2\pi ai\tau/d} + \sum_{n=0}^{\infty} c(n)\zeta_d^b e^{2na\pi i\tau/d}. \tag{12.2}$$

Haciendo $z = e^{2\pi i\tau}$, a cada una de estas funciones le podemos asignar la serie de potencias fraccionarias

$$\bar{f}_i(z) = \zeta_d^{-b}z^{a/d} + \sum_{n=0}^{\infty} c(n)\zeta_d^b z^{na/d} \in \mathbb{Q}(\zeta_N)\{z\}.$$

Los coeficientes de $F_N(j, Y)$ son de la forma $q(\tau) = S(f_1, \ldots, f_n)$, donde el polinomio $S(X_1, \ldots, X_n)$ es simétrico. Es claro que

$$S(\bar{f}_1, \ldots, \bar{f}_n) = \sum_{n=k}^{\infty} c_n z^{n/N}, \qquad k \in \mathbb{Z}, \quad c_n \in \mathbb{Z}[\zeta_N].$$

Por otro lado, cada $r \in U_N$ define un $\mathbb{Q}$-automorfismo de $\mathbb{Q}(\zeta_N)$, el cual se extiende a un $\mathbb{Q}\{z\}$ automorfismo de $\mathbb{Q}(\zeta_N)\{z\}$ (y así obtenemos todos los $\mathbb{Q}\{z\}$ automorfismos). Dichos automorfismos permutan las series $\bar{f}_i(z)$, luego fijan a la serie $S(\bar{f}_1, \ldots, \bar{f}_n)$. Por consiguiente, $S(\bar{f}_1, \ldots, \bar{f}_n) \in \mathbb{Q}\{z\}$. Concluimos que los coeficientes c_n son en realidad números enteros y

$$q(\tau) = S(f_1, \ldots, f_n) = \sum_{n=k}^{\infty} c_n e^{2n\pi i\tau/N}.$$

Ahora bien, sabemos que $q(\tau)$ es una función modular, luego tiene periodo 1. Esto hace que

$$q(\tau+1) = \sum_{n=k}^{\infty} c_n\zeta_N^n z^n = \sum_{n=k}^{\infty} c_n z^n = q(\tau),$$

y la unicidad de las series de Laurent hace que $c_n = 0$ siempre que $\zeta_N^n \neq 1$, es decir, siempre que $N \nmid n$. Por consiguiente,

$$q(\tau) = \sum_{n=k}^{\infty} c_n e^{2n\pi i\tau}, \qquad k \in \mathbb{Z}, \quad c_n \in \mathbb{Z}.$$

De aquí se sigue ahora que $q(\tau)$ es un polinomio en j con coeficientes enteros. En efecto, si $k < 0$ consideramos $q(\tau) - c_n j^k$, que es una función modular holomorfa en H cuya serie de Fourier tiene coeficientes enteros pero su singularidad en ∞ tiene orden menor. Repitiendo el razonamiento llegamos a un polinomio $P(X)$ con coeficientes enteros tal que $q(\tau) - P(j)$ es una función modular holomorfa en H y nula en ∞, lo que obliga a que sea idénticamente nula.

En definitiva, los coeficientes del polinomio $F_N(j, Y)$ son polinomios en j con coeficientes enteros, luego concluimos que $F_N(X, Y) \in \mathbb{Z}[X, Y]$.

Veamos ahora que $F_N(X,Y)$ es simétrico. Partimos de que j_N es raíz de $F_N(j,Y)$, luego $F_N(j(\tau),j(N\tau))=0$ para todo $\tau\in H$. Aplicando esto a τ/N vemos que $F_N(j(\tau/N),j(\tau))=0$ para todo $\tau\in H$, luego $j(\tau/N)$ es raíz del polinomio $F_N(X,j)$. Ahora bien, también es raíz de $F_N(j,Y)$, la correspondiente a

$$\alpha_i=\begin{pmatrix}1&0\\0&N\end{pmatrix}.$$

Como $F_N(j,Y)$ es irreducible, ha de ser $F_N(j,X)\mid F_N(X,j)$. La división es en $\mathbb{Q}(j)[X]$, pero por el lema de Gauss también en $\mathbb{Z}[j][X]$. Digamos que $F_N(X,j)=F_N(j,X)G(j,X)$, para cierto polinomio $G(j,X)\in\mathbb{Z}[j][X]$. Como j es trascendente sobre $\mathbb{Q}$, de hecho $F_N(X,Y)=F_N(Y,X)G(Y,X)$. Ahora bien:

$$F_N(X,Y)=F_N(X,Y)G(X,Y)G(Y,X),$$

de donde $G(X,Y)G(Y,X)=1$, con lo que $G(X,Y)=\pm1$. Si el signo fuera $G(X,Y)=-1$ tendríamos que $F_N(j,j)=-F_N(j,j)$, luego $F_N(j,j)=0$, es decir, $F_N(j,Y)$ tendría una raíz en $\mathbb{Q}(j)$, en contradicción con el hecho de que es irreducible. Así pues, $G=1$ y concluimos que $F_N(X,Y)=F_N(Y,X)$.

Resumimos en un teorema lo que hemos obtenido:

Teorema 12.7 *El polinomio modular $F_N(X,Y)$ tiene coeficientes enteros y es simétrico en X e Y.*

Otra propiedad de interés es la siguiente:

Teorema 12.8 *Si N no es un cuadrado perfecto, entonces el polinomio modular $F_N(X,X)$ es mónico.*

Demostración: Si N no es un cuadrado perfecto, en todos los factores de (12.1) se ha de cumplir $a\neq d$. Teniendo en cuenta (12.2), vemos que

$$j(\tau)-j\left(\frac{a\tau+b}{d}\right)=e^{-2\pi i\tau}-\zeta_d^{-b}e^{-2\pi ai\tau/d}+\sum_{n=0}^{\infty}c(n)e^{2n\pi i\tau}-\sum_{n=0}^{\infty}c(n)\zeta_d^{b}e^{2na\pi i\tau/d}.$$

La clave está en que los dos términos no singulares no se cancelan, y el de menor grado (sea cual sea de los dos) tiene por coeficiente una raíz de la unidad. Multiplicando sobre todas las ternas (a,b,d) obtenemos el desarrollo en serie de Fourier de $F_N(j,j)$, que, por otra parte, ha de tener coeficientes enteros. El menor coeficiente de Fourier no nulo es el producto de los coeficientes análogos de cada factor, que son todos raíces de la unidad, luego es a la vez entero y raíz de la unidad. Por consiguiente es ±1. Así pues:

$$F_n(j,j)=\pm e^{-2\pi im\tau}+c_{m-1}e^{-2\pi i(m-1)\tau}+\cdots,$$

de donde se sigue que

$$F_n(X,X)=\pm X^m+c_{m-1}X^{m-1}+\cdots$$

∎

Esto nos permite mejorar el teorema 10.24:

Teorema 12.9 *Si $\tau \in H$ es un irracional cuadrático, entonces $j(\tau)$ es un entero algebraico.*

DEMOSTRACIÓN: Sea $K = \mathbb{Q}(\tau)$ y sea $\mathcal{O} = \langle 1, z\rangle_{\mathbb{Z}}$ su orden maximal. Tomemos $\lambda \in \mathcal{O}$ tal que $N = \mathrm{N}(\lambda)$ sea libre de cuadrados. (Si $K = \mathbb{Q}(i)$ basta tomar $\lambda = 1 + i$ y si $K = \mathbb{Q}(\sqrt{d}\,)$ tomamos $\lambda = \sqrt{d}$.) Pongamos que

$$\lambda z = az + b, \qquad \lambda = cz + d, \qquad a, b, c, d \in \mathbb{Z},$$

y definamos

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

Entonces $N = \mathrm{N}(\lambda) = |\mathcal{O} : (\lambda)| = ad - bc$. El hecho de que N sea libre de cuadrados implica que $(a, b, c, d) = 1$. Por 12.6 existe $g \in \mathrm{LE}(2, \mathbb{Z})$ tal que las entradas de αg cumplen las condiciones de (12.1), por lo que

$$F_N(j(\tau), j(\tau\alpha g)) = F_N(j(\tau), j(\tau\alpha)) = 0$$

para todo $\tau \in H$. En particular, para $\tau = z$ tenemos que

$$z\alpha = \frac{az + b}{cz + d} = \frac{\lambda z}{\lambda} = z,$$

luego $F_N(j(z), j(z)) = 0$. Así pues, $j(z)$ es raíz del polinomio $F_N(X, X)$ y, como N es libre de cuadrados, el teorema anterior nos da que el polinomio es mónico, luego $j(z)$ es un entero algebraico. El τ del enunciado es de la forma $\tau = rz + s$, con r, $s \in \mathbb{Q}$, $r > 0$. Equivalentemente, $\tau = z\beta$, con

$$\beta = \begin{pmatrix} r & 0 \\ s & 1 \end{pmatrix}.$$

Multiplicando por el producto de los denominadores de r y s podemos suponer que β tiene entradas enteras (aunque la última ya no sea 1 necesariamente). Dividiendo entre el máximo común divisor de las tres podemos suponer que son primas entre sí. Esto no altera la transformación de Möbius, luego sigue siendo cierto que $\tau = z\beta$ y podemos aplicar a β el teorema 12.6, según el cual existe $g \in \mathrm{LE}(2, \mathbb{Z})$ tal que βg está en las condiciones de (12.1) para $N = \det\beta$. Así pues:

$$F_N(j(z), j(z\beta g)) = F_N(j(z), j(\tau)) = 0.$$

Esto significa que $j(\tau)$ es raíz del polinomio $F_N(j(z), X)$, que es mónico y sus coeficientes son enteros algebraicos, luego $j(\tau)$ también es un entero algebraico. ∎

Veamos ahora otras consecuencias de la existencia de la ecuación modular. Podemos considerar la curva proyectiva plana $V/\mathbb{Q}$ definida por $F_N(X, Y)$. Tenemos un $\mathbb{C}$-isomorfismo $\mathbb{C}(j, j_N) \cong \mathbb{C}(V)$ por el que j y j_N se corresponden, respectivamente, con las funciones coordenadas x e y. A través de este isomorfismo, el cuerpo $\mathbb{Q}(j, j_N)$ se corresponde con $\mathbb{Q}(x, y) = \mathbb{Q}(V)$. La curva V es,

en general, singular, pero por el teorema 1.18 sabemos que existe una curva proyectiva regular $M_N/\mathbb{Q}$ tal que $\mathbb{Q}(M_N)$ es $\mathbb{Q}$-isomorfo a $\mathbb{Q}(V)$, y el isomorfismo se extiende a un $\mathbb{C}$-isomorfismo entre $\mathbb{C}(M_N)$ y $\mathbb{C}(V)$. Componiendo estos isomorfismos, obtenemos un $\mathbb{C}$-isomorfismo $\mathbb{C}(j, j_N) \cong \mathbb{C}(M_N)$ que se restringe a un $\mathbb{Q}$-isomorfismo $\mathbb{Q}(j, j_N) \cong \mathbb{Q}(M_N)$. A su vez, este isomorfismo determina una transformación conforme $\phi : M_N \longrightarrow X_0(N)$ de modo que el isomorfismo no es más que la composición con ϕ. En definitiva:

Teorema 12.10 *Existe una curva proyectiva regular $M_N/\mathbb{Q}$ y una transformación conforme $\phi : M_N \longrightarrow X_0(N)$ de modo que el cuerpo $\mathbb{Q}(j, j_N)$ se corresponde a través de ϕ con el cuerpo $\mathbb{Q}(M_N)$.*

El cuerpo $\mathbb{Q}(j, j_N)$ admite una caracterización sencilla:

Teorema 12.11 *Si K es un subcuerpo de $\mathbb{C}$, entonces $K(j, j_N)$ está formado por las funciones de $\mathbb{C}(j, j_N)$ con coeficientes de Fourier en K.*

Demostración: Tenemos que $\mathbb{C}(j, j_N)$ es isomorfo a un cuerpo de funciones racionales $\mathbb{C}(V)$, de modo que $\mathbb{Q}(j, j_N)$ se corresponde con $\mathbb{Q}(V)$, con lo que $K(j, j_N)$ se corresponde con $K(V)$. El grupo de K-automorfismos $G(\mathbb{C}/K)$ actúa sobre $\mathbb{C}(V)$, y la acción puede trasladarse a $\mathbb{C}(j, j_N)$ a través del isomorfismo, de modo que el cuerpo fijado por $G(\mathbb{C}/K)$ en $\mathbb{C}(j, j_N)$ es $K(j, j_N)$. Basta probar que si $f \in \mathbb{C}(j, j_N)$ tiene como desarrollo de Fourier

$$f(\tau) = \sum_{n=r}^{+\infty} c_n e^{2\pi n i \tau/h}$$

y $\sigma \in G(\mathbb{C}/K)$, entonces

$$f^\sigma(\tau) = \sum_{n=r}^{+\infty} c_n^\sigma e^{2\pi n i \tau/h}. \tag{12.3}$$

Admitiendo esto, la conclusión es inmediata, pues el teorema 1.17 nos da que $K(j, j_N)$ es el cuerpo fijado por $G(\mathbb{C}/K)$, que está formado por las funciones con coeficientes de Fourier fijados por $G(\mathbb{C}/K)$, y el mismo argumento del teorema 1.17 prueba que el cuerpo fijado por $G(\mathbb{C}/K)$ en $\mathbb{C}$ es K. (De hecho, es 1.17 en el caso en que V es un punto.)

La afirmación es inmediata para funciones de $\mathbb{C}[j]$, pues si

$$f = c_n j^n + \cdots + c_1 j + c_0, \qquad c_i \in \mathbb{C},$$

entonces $f^\sigma = c_n^\sigma j^n + \cdots + c_1^\sigma j + c_0^\sigma$ y los coeficientes de Fourier de f (resp. f^σ) son las combinaciones lineales de los coeficientes de Fourier de $j^n, \ldots, j, 1$ (que son enteros) con coeficientes c_i (resp. c_i^σ).

Cada automorfismo de $G(\mathbb{C}/K)$ induce dos automorfismos en $\mathbb{C}(j)$, el definido a través de $\mathbb{C}(V)$ y el dado por (12.3). En efecto, para comprobar que (12.3) es ciertamente un automorfismo basta observar que podemos sumergir $\mathbb{C}(j)$ en

el cuerpo de series formales de potencias $\mathbb{C}((z))$ asignando a cada función modular su serie de Laurent en ∞ (cuyos coeficientes son los coeficientes de Fourier), y (12.3) es la restricción a la imagen de $\mathbb{C}(j)$ del automorfismo inducido por σ en $\mathbb{C}((z))$ de forma natural. (Sabemos que σ transforma elementos de la imagen de $\mathbb{C}[j]$ en elementos de la imagen de $\mathbb{C}[j]$, luego lo mismo vale para la imagen de $\mathbb{C}(j)$.)

El hecho de que los dos automorfismos coincidan sobre $\mathbb{C}[j]$ implica claramente que ambos coinciden sobre $\mathbb{C}(j)$.

Finalmente, todo elemento de $\mathbb{C}(j, j_N)$ es de la forma

$$f = g_n j_N^n + \cdots + g_1 j_N + g_0, \qquad g_i \in \mathbb{C}(j),$$

luego basta probar que los coeficientes de Fourier de $(g_i j_N^i)^\sigma = g_i^\sigma j_N^i$ son las imágenes por σ de los de $g_i j_N^i$, sabiendo que esto es cierto para g_i y que los coeficientes de j_N^i son enteros. Basta tener en cuenta que los coeficientes del producto se calculan como los de un producto de series de Laurent. ∎

Para terminar demostramos una relación entre la ecuación modular y las isogenias entre curvas elípticas:

Teorema 12.12 *Sean $E_1/\mathbb{C}$ y $E_2/\mathbb{C}$ dos curvas elípticas con invariantes j_1 y j_2 respectivamente. Entonces existe una isogenia $\phi : E_1 \longrightarrow E_2$ con núcleo cíclico de orden N si y sólo si $F_N(j_1, j_2) = 0$.*

Demostración: Podemos sustituir las curvas por dos toros complejos $\mathbb{C}/R_1$ y $\mathbb{C}/R_2$. Más aún, podemos suponer que $R_1 = \langle 1, \tau\rangle_{\mathbb{Z}}$, para cierto $\tau \in H$. Entonces $j_1 = j(\tau)$. Evaluando en τ y en j_2 la definición de F_N vemos que

$$0 = F_N(j(\tau), j_2) = \prod_{i=1}^{n} (j_2 - j_N(\tau g_i)),$$

luego existe un $g = g_i$ en el grupo modular tal que $j_2 = j_N(\tau g) = j(Ng(\tau))$. Por lo tanto, podemos tomar $R_2 = \langle 1, Ng(\tau)\rangle_{\mathbb{Z}}$. Sea $\tau' = g(\tau)$, de modo que $j_1 = j(\tau) = j(\tau')$. Esto significa que podemos cambiar R_1 por $R_1 = \langle N, N\tau'\rangle_{\mathbb{Z}}$ y así resulta que $R_1 = \langle N, N\tau'\rangle_{\mathbb{Z}} \subset \langle 1, N\tau'\rangle_{\mathbb{Z}} = R_2$ y la aplicación natural $\mathbb{C}/R_1 \longrightarrow \mathbb{C}/R_2$ tiene núcleo cíclico de orden N.

Recíprocamente, si existe un homomorfismo $\alpha : \mathbb{C}/R_1 \longrightarrow \mathbb{C}/R_2$ con núcleo cíclico de orden N, (donde $\alpha \in \mathbb{C}^*$), entonces $\alpha R_1 \subset R_2$, y cambiando R_1 por αR_1 podemos suponer que $R_1 \subset R_2$, de modo que el cociente R_2/R_1 es cíclico de orden N.

Fijando bases de R_1 y R_2, la matriz con las coordenadas de la base de R_1 en la base de R_2 tiene coeficientes enteros y determinante no nulo. Multiplicándola por matrices de $\mathrm{LE}(2, \mathbb{Z})$ podemos pasar a una matriz diagonal

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \qquad a \mid d.$$

Esto significa que podemos expresar $R_2 = \langle \omega_1, \omega_2\rangle_{\mathbb{Z}}$ y $R_1 = \langle a\omega_1, d\omega_2\rangle_{\mathbb{Z}}$ (y claramente podemos exigir que a, $d > 0$). El cociente es isomorfo a $C_a \times C_d$,

luego ha de ser $a = 1$, $d = N$. Si llamamos $\tau = \omega_2/\omega_1$, podemos sustituir los retículos por $R_1 = \langle 1, N\tau\rangle_{\mathbb{Z}}$ y $R_2 = \langle 1, \tau\rangle_{\mathbb{Z}}$, luego $j_1 = j(N\tau) = j_N(\tau)$, $j_2 = j(\tau)$, de donde se sigue que $F_N(j_1, j_2) = 0$. ∎

12.3 Funciones modulares de grados superiores

De acuerdo con la definición que hemos dado hasta ahora, las funciones de Eisenstein $G_{2k}(\tau)$ o la función discriminante $\Delta(\tau)$ no son funciones modulares, ya que no son invariantes por el grupo modular. En su lugar verifican relaciones del tipo (10.3). Tal y como habíamos anunciado, vamos a generalizar la noción de función modular para incluir a estas y otras muchas funciones de interés. Empezamos admitiendo (10.3) como regla de transformación:

Definición 12.13 Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y $k \in \mathbb{Z}$. Una *función cuasimodular* de grado $2k$ respecto de Γ es una aplicación meromorfa $f : H \longrightarrow \mathbb{C}$ tal que para todo

$$\alpha = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \Gamma$$

y todo $\tau \in H$ se cumple

$$f(\tau\alpha) = (c\tau + d)^{2k} f(\tau).$$

(Notemos que la única función modular de grado impar sería la forma nula, pues sólo ella cumpliría la relación anterior con $\alpha = -I$.)

La relación (10.3) prueba que la serie de Eisenstein G_{2k} es una función cuasimodular de grado $2k$ (en particular g_2 y g_3 son funciones cuasimodulares de grados 4 y 6, respectivamente), e inmediatamente después de (10.3) hemos visto que Δ es cuasimodular de grado 12.

Las funciones modulares en sentido amplio serán las funciones cuasimodulares meromorfas en los puntos parabólicos, pero todavía no estamos en condiciones de definir esta noción de meromorfía. Primeramente conviene expresar las reglas de transformación que estamos considerando como la invarianza de otras funciones asociadas. Recordemos que $\mathrm{LG}_+(2,\mathbb{R})$ es el subgrupo de $\mathrm{LG}(2,\mathbb{R})$ formado por las matrices de determinante positivo.

Dada

$$\alpha = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{LG}_+(2,\mathbb{R}),$$

definimos

$$j_\alpha(\tau) = \frac{d\alpha}{d\tau} = (\det\alpha)(c\tau + d)^{-2}.$$

Aplicando la regla de la cadena vemos que si α, $\beta \in \mathrm{LG}_+(2,\mathbb{R})$, entonces

$$j_{\alpha\beta}(\tau) = \frac{d\alpha\beta}{d\tau} = \frac{d\beta}{d\tau}(\tau\alpha)\cdot\frac{d\alpha}{d\tau} = j_\beta(\tau\alpha)j_\alpha(\tau). \tag{12.4}$$

Notemos también que si

$$\alpha = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},$$

entonces

$$j_\alpha(\tau) = a^2 a^{-2} = 1.$$

Vemos así que j_α no se altera si multiplicamos α por un número real o, equivalentemente, que j_α depende únicamente de la transformación de Möbius inducida por α.

Si $f \in \mathcal{M}(H)$, definimos

$$(\alpha|_{2k} f)(\tau) = j_\alpha(\tau)^k f(\tau\alpha) = (\det\alpha)^k (c\tau + d)^{-2k} f(\tau\alpha).$$

Así, la condición de cuasimodularidad de f respecto a un subgrupo Γ de $\mathrm{LE}(2,\mathbb{Z})$ es que $\alpha|_{2k} f = f$ para todo $\alpha \in \Gamma$ (observemos que en este caso $\det\alpha = 1$). Como consecuencia de (12.4) tenemos que

$$(\alpha|_{2k}(\beta|_{2k} f))(\tau) = (\beta|_{2k} f)(\tau\alpha) j_\alpha(\tau)^k = f(\tau\alpha\beta) j_\beta(\tau\alpha)^k j_\alpha(\tau)^k$$

$$= f(\tau\alpha\beta) j_{\alpha\beta}(\tau)^k = ((\alpha\beta)|_{2k} f)(\tau),$$

de modo que

$$\alpha|_{2k}(\beta|_{2k} f) = (\alpha\beta)|_{2k} f.$$

En otros términos, el grupo $\mathrm{LE}(2,\mathbb{R})$ actúa sobre $\mathcal{M}(H)$ (con una acción distinta para cada k). En particular vemos que si la condición de invarianza de la definición de función cuasimodular se cumple para unos generadores del grupo Γ, entonces se cumple para todos sus elementos.

Por último conviene observar lo siguiente:

Teorema 12.14 *Sean Γ y Γ' dos subgrupos de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y sea $\alpha \in \mathrm{LG}_+(2,\mathbb{R})$ tal que $\Gamma'^\alpha \subset \Gamma$. Si f es una función cuasimodular de grado $2k$ respecto de Γ, entonces $\alpha|_{2k} f$ es cuasimodular de grado $2k$ respecto de Γ'.*

Demostración: Tomemos $\beta \in \Gamma'$, de modo que $\beta = \alpha\gamma\alpha^{-1}$, para cierto $\gamma \in \Gamma$. Así, para todo $\tau \in H$,

$$\beta|_{2k}(\alpha|_{2k} f) = (\beta\alpha)|_{2k} f = (\alpha\gamma)|_{2k} f = \alpha|_{2k}(\gamma|_{2k} f) = \alpha|_{2k} f.$$

■

Ahora ya podemos ocuparnos del problema de definir la meromorfía en los puntos parabólicos de una función cuasimodular. La cuestión es que una función cuasimodular respecto de un grupo Γ no es invariante por Γ, por lo que no induce una función sobre la superficie modular asociada. Por ello, la noción de meromorfía en los puntos parabólicos no tiene sentido literalmente. No obstante, vamos a ver que la condición c) del teorema 12.2 sí tiene una generalización sencilla. Para ello vemos que si Γ es un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$,

entonces f es una función cuasimodular de grado $2k$ respecto de Γ y $s \in \mathbb{Q}\cup\{\infty\}$, podemos tomar $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $sg = \infty$, y el teorema anterior garantiza que $g^{-1}|_{2k}f$ es una función cuasimodular respecto de $\Gamma^{g^{-1}}$. Este grupo ha de contener una traslación

$$\alpha = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix},$$

donde podemos suponer que $h > 0$ es el menor posible. (Este h no depende de la elección de g, pues claramente es $h = |\bar{\Gamma} : \bar{\Gamma}_s|$). Puesto que $j_\alpha = 1$, tenemos que

$$(g^{-1}|_{2k}f)(\tau + h) = (g^{-1}|_{2k}f)(\tau\alpha) = (\alpha|_{2k}(g^{-1}|_{2k}f))(\tau) = (g^{-1}|_{2k}f)(\tau),$$

de modo que $g^{-1}|_{2k}f$ tiene periodo h. Si no tiene polos en un semiplano de la forma $\operatorname{Im}\tau > R$ podemos considerar el desarrollo en serie de Fourier

$$(g^{-1}|_{2k}f)(\tau) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi n i\tau/h}, \qquad \operatorname{Im}\tau > R.$$

Definición 12.15 Diremos que una función cuasimodular f de grado $2k$ respecto a un subgrupo Γ de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ es *meromorfa* en un punto parabólico $s \in \mathbb{Q}\cup\{\infty\}$ si cuando $g \in \mathrm{LE}(2,\mathbb{Z})$ cumple $sg = \infty$ entonces la función $g^{-1}|_{2k}f$ admite un desarrollo en serie de Fourier cuyos coeficientes son nulos para índices menores que un cierto $r \in \mathbb{Z}$.

Notemos que la condición no depende de la elección de g, ya que si tenemos $sg_1 = sg_2 = \infty$, entonces $\beta = g_2^{-1}g_1$ cumple $\infty\beta = \infty$, luego

$$\beta = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}.$$

Sea $(g^{-1}|_{2k}f)(\tau) = f^*(e^{2\pi i\tau/h})$, para cierta función f^* meromorfa en 0. Entonces

$$\begin{aligned}(g_2^{-1}|_{2k}f)(\tau) &= (\beta g_1^{-1}|_{2k}f)(\tau) = (\beta|_{2k}(g_1^{-1}|_{2k}f))(\tau)\\ &= f^*(e^{2\pi i(\tau+t)/h}) = g^*(e^{2\pi i\tau/h}),\end{aligned}$$

donde $g^*(z) = f^*(e^{2\pi it/h}z)$ es una función meromorfa en 0. Además el orden de f^* en 0 es el mismo que el de g^*, por lo que podemos hablar del orden de una función cuasimodular meromorfa en un punto parabólico.

Definición 12.16 Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y sea $k \in \mathbb{Z}$. Una *función modular* de grado $2k$ respecto de Γ es una aplicación $f : H \longrightarrow \mathbb{C}^\infty$ que cumpla las propiedades siguientes:

a) f es meromorfa en H,

b) para todo $\alpha \in \Gamma$ se cumple $\alpha|_{2k}f = f$,

c) f es meromorfa en los puntos parabólicos de Γ.

Es claro que cada una de las condiciones de esta definición generaliza a la correspondiente del teorema 12.2, por lo que —tal y como pretendíamos— las funciones modulares de grado 0 son exactamente lo que hasta ahora llamábamos funciones modulares.

Se comprueba trivialmente que una función cuasimodular f es meromorfa en un punto parabólico si y sólo si lo es en todos los puntos equivalentes, y además f tiene el mismo orden en todos ellos. En particular, para comprobar que una función es modular respecto a $\mathrm{LE}(2,\mathbb{Z})$ basta ver que es meromorfa en ∞.

Así, ahora es claro que las funciones G_{2k}, g_2, g_3, Δ y j son modulares de grados $2k$, 4, 6, 12 y 0 respectivamente.

Definición 12.17 Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$. Para cada $k \in \mathbb{Z}$, llamaremos $A_{2k}(\Gamma)$ al conjunto de todas las funciones modulares de grado $2k$ respecto de Γ.

Es claro que los conjuntos A_{2k} son $\mathbb{C}$-subespacios vectoriales del cuerpo $\mathcal{M}(H)$. Vamos a ver que forman suma directa. En efecto, supongamos que

$$f_m + f_{m+1} + \cdots + f_n = 0, \qquad f_k \in A_{2k}(\Gamma),$$

y vamos a probar que todas las f_k son nulas. En caso contrario, supongamos que i es el menor índice tal que $f_i \neq 0$. Tenemos entonces que

$$f_i = -f_{i+1} - \cdots - f_n.$$

Consideremos la transformación $s(\tau) = -1/\tau$. Como Γ tienen índice finito en el grupo modular, existe un natural $l > 0$ tal que $s^l \in \Gamma$. Aplicando la igualdad anterior a $s^{mr}(\tau)$ vemos que

$$\tau^{2lri} f_i(\tau) = -\tau^{2lr(i+1)} f_{i+1}(\tau) - \cdots - \tau^{2lrn} f_n(\tau),$$

de donde

$$f_i(\tau) = -\tau^{2lr} f_{i+1}(\tau) - \cdots - \tau^{2lr(n-i)} f_n(\tau).$$

Haciendo $r \to \infty$ queda que $f_i(\tau) = 0$ para todo $\tau \in H$ con $|\tau| < 1$. Por consiguiente $f_i = 0$, contradicción.

Esto nos permite definir la suma directa

$$A(\Gamma) = \bigoplus_{k=-\infty}^{+\infty} A_{2k}(\Gamma),$$

que es una $\mathbb{C}$-subálgebra graduada del cuerpo $\mathcal{M}(H)$, pues claramente

$$A_{2k}(\Gamma)A_{2l}(\Gamma) \subset A_{2k+2l}(\Gamma).$$

Conviene definir algunas clases especiales de funciones modulares:

Definición 12.18 Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y sea $k \in \mathbb{Z}$. Una *forma modular* de grado $2k$ respecto de Γ es una función modular de grado $2k$ holomorfa en H y en los puntos parabólicos de Γ (lo cual significa que su orden en éstos es ≥ 0). Una *forma parabólica* respecto de Γ es una forma modular que se anula en los puntos parabólicos (es decir, que tiene orden > 0 en ellos).

Por ejemplo, el desarrollo en serie de Fourier de $G_{2k}(\tau)$ muestra que no es parabólica, al igual que g_2 y g_3. Por el contrario, Δ sí que es una forma parabólica de grado 12. La función modular j no es una forma modular, pues tiene un polo en ∞. (De hecho, es claro que no hay formas modulares no constantes de grado 0.)

Llamaremos $M_{2k}(\Gamma)$ al espacio vectorial de las formas modulares de grado $2k$ respecto de Γ y $M^0_{2k}(\Gamma)$ al subespacio de las formas parabólicas. Así mismo podemos considerar las subálgebras graduadas de $A(\Gamma)$ dadas por

$$M(\Gamma) = \bigoplus_{k=-\infty}^{+\infty} M_{2k}(\Gamma), \quad \text{y} \quad M^0(\Gamma) = \bigoplus_{k=-\infty}^{+\infty} M^0_{2k}(\Gamma).$$

El teorema siguiente explica en parte la razón por la que tienen interés las formas parabólicas. Recordemos que las diferenciales de primera clase en una superficie de Riemann son las diferenciales holomorfas:

Teorema 12.19 *Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y consideremos la proyección natural $p : H \longrightarrow H^*/\Gamma$. Para cada diferencial de primera clase ω en H^*/Γ sea $p^\sharp(\omega) = f\,dz$. Entonces la correspondencia $\omega \mapsto f$ es un isomorfismo entre el espacio de las diferenciales de primera clase de H^*/Γ y el espacio de las formas parabólicas de grado 2 respecto de Γ.*

Demostración: Ciertamente, si ω es una diferencial de primera clase en H^*/Γ, entonces $p^\sharp(\omega)$ es una forma diferencial holomorfa en H, luego es de la forma $f\,dz$ para cierta función f holomorfa en H. Si $\alpha \in \Gamma$, entonces $p = \alpha \circ p$, luego $p^\sharp = p^\sharp \circ \alpha^\sharp$. Por consiguiente,

$$f\,dz = \alpha^\sharp(f\,dz) = (\alpha \circ f)\,d\alpha = (\alpha \circ f) j_\alpha\,dz = \alpha|_2 f\,dz.$$

Así pues, $\alpha|_2 f = f$. Nos falta comprobar que f es holomorfa (y nula) en los puntos parabólicos de Γ. Como es usual, no perdemos generalidad si nos restringimos a ∞.

Por la construcción de la estructura analítica de la superficie modular, existe una carta q en un entorno de ∞ tal que $p \circ q = r$, donde $r(z) = e^{2\pi i z/h}$, para cierto natural $h > 0$. Sea $\omega = q^\sharp(f^*\,dw)$, donde w es la identidad en el disco unidad y $f^*(w)$ es una función holomorfa en dicho disco. Así,

$$f\,dz = p^\sharp(\omega) = p^\sharp(q^\sharp(f^*\,dw)) = r^\sharp(f^*\,dw) = (r \circ f^*)dr = (r \circ f^*)r\,dz.$$

Concluimos que, para todo z en un entorno de ∞,

$$f(z) = f^*(e^{2\pi i z/h}) e^{2\pi i z/h}.$$

Desarrollando f^* en serie de Taylor obtenemos el desarrollo de f en serie de Fourier, y vemos que tiene orden ≥ 1, luego f es nula en ∞ y, por consiguiente, es una forma parabólica.

Recíprocamente, consideremos ahora una forma parabólica f de grado 2. Cada punto $z \in H$ que no sea elíptico tiene un entorno U donde $p|_U$ es una transformación conforme, luego la forma diferencial $f|_U\,dz$ induce una forma diferencial ω_U en $p[U]$ tal que $f|_U\,dz = p|_U^\sharp(\omega_U)$. Las formas ω_U coinciden en los puntos comunes de sus dominios. En efecto, si $p[U] = p[U']$, entonces existe $\alpha \in \Gamma$ tal que $U' = \alpha[U]$. De la relación $p|_U = \alpha \circ p|_{U'}$ se sigue que $p|_U^\sharp = p|_{U'}^\sharp \circ \alpha^\sharp$, luego

$$p|_U^\sharp(\omega_{U'}) = \alpha^\sharp(f|_{U'}\,dz) = (f|_U \circ \alpha)\,d\alpha = (\alpha|_2 f)|_U\,dz = f|_U\,dz.$$

Por consiguiente $\omega_{U'} = \omega_U$. Con esto tenemos definida una forma diferencial holomorfa ω sobre los puntos de la superficie modular H^*/Γ que no son elípticos ni parabólicos y tal que $p^\sharp(\omega) = f\,dz$. Veamos que ω también es holomorfa en dichos puntos.

Si $p(z)$ es un punto elíptico, entonces existe una carta $q : W \longrightarrow D(0,1)$ alrededor de $p(z)$ tal que $p \circ q = \lambda \circ w^n$, donde $\lambda : H \longrightarrow D(0,1)$ es una transformación conforme. Sea $\omega|_W = q^\sharp(f^*\,dw)$, donde f^* es una función holomorfa en $D(0,1)$ salvo una singularidad en 0, que hemos de ver que es evitable. En efecto, tenemos que

$$f\,dz = p^\sharp(\omega) = \lambda^\sharp((w^n)^\sharp(f^*\,dw)) = \lambda^\sharp(f^*(w^n)\,dw^n) = \lambda^\sharp(nf^*(w^n)w^{n-1}\,dw).$$

Como λ es una transformación conforme y f es holomorfa en H, concluimos que $nf^*(w^n)w^{n-1}$ es holomorfa en 0, pero el orden de esta función en 0 es

$$n - 1 + no(f^*, 0) \geq 0,$$

luego $o(f^*, 0) \geq -1 + 1/n > -1$, luego $o(f^*, 0) \geq 0$, como había que probar.

El razonamiento para los puntos parabólicos es una combinación del caso anterior y del caso correspondiente de la implicación contraria. ∎

Es conocido que el espacio de las diferenciales de primera clase de un cuerpo de funciones algebraicas tiene dimensión igual al género, luego tenemos que la dimensión del espacio de las formas parabólicas de grado 2 respecto de un subgrupo Γ de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ es igual al género de la superficie modular H^*/Γ. El teorema de Riemann-Roch permite calcular la dimensión de los espacios de formas modulares y parabólicas de cualquier grado, pero no vamos a entrar en ello. Probaremos únicamente la finitud de la dimensión:

Teorema 12.20 *Si Γ es un subgrupo de índice finito en $LE(2,\mathbb{Z})$, entonces los espacios $M_{2k}(\Gamma)$ de formas modulares de grado $2k$ respecto de Γ tienen dimensión finita sobre $\mathbb{C}$.*

Demostración: Consideremos en primer lugar una función modular no nula $F_0 \in A_{2k}(\Gamma)$. No podemos considerar a F_0 como una función sobre la

superficie modular $X = H^*/\Gamma$, pues F_0 toma valores distintos sobre puntos equivalentes, pero sí que podemos definir el orden de F_0 en un punto de X. En efecto, si τ_1, $\tau_2 \in H$ son puntos equivalentes respecto de Γ, digamos $\tau_2 = \tau_1\gamma$, con $\gamma \in \Gamma$, entonces, $F_0(\tau\gamma) = (c\tau+d)^{2k}F_0(\tau)$, lo que se traduce en que el orden de F_0 en τ_1 es el mismo que en τ_2. Esto justifica que hablemos del orden de F_0 en un punto no parabólico de X, y ya hemos comentado anteriormente que el orden de F_0 en un punto parabólico también está bien definido. Puesto que los ceros y polos de F_0 en H han de ser un conjunto discreto, es claro que el conjunto de puntos de X donde F_0 tiene orden no nulo ha de ser finito, luego podemos asociar a F_0 un divisor (F_0) de la superficie de Riemann X estableciendo que $v_P((F_0))$ es el orden de F_0 en el punto $P \in X$.

Fijemos ahora una forma modular no nula $F_0 \in M_{2k}(\Gamma)$ (si no existe, entonces el espacio tiene dimensión 0 y no hay nada que probar). Si $F \in M_{2k}(\Gamma)$ es cualquier otra forma modular no nula, entonces el cociente $f = F/F_0$ es una función modular de grado 0, es decir, una función meromorfa en X. Además, como F es holomorfa, es claro que f tiene a lo sumo polos en los puntos donde F_0 tiene ceros, y el orden de éstos no puede ser superior al orden de los ceros correspondientes de F_0. En otras palabras, f ha de ser un múltiplo del divisor (F_0). Recíprocamente, cualquier función f en el espacio $m((F_0))$ de los múltiplos de F_0 define una forma modular $F = fF_0 \in M_{2k}(\Gamma)$. Es inmediato comprobar que la correspondencia $f \leftrightarrow F$ (extendida de forma que $0 \leftrightarrow 0$) es un isomorfismo entre $m((F_0))$ y $M_{2k}(\Gamma)$. Es sabido que los espacios de múltiplos de un divisor tienen dimensión finita, luego lo mismo le sucede a $M_{2k}(\Gamma)$. ∎

12.4 Funciones modulares de $\mathbf{LE(2,\mathbb{Z})}$

En esta sección estudiamos más detalladamente las funciones y formas modulares respecto al grupo modular completo $\Gamma = \mathrm{LE}(2,\mathbb{Z})$. Vamos a determinar la estructura del álgebra de las formas modulares. Como consecuencia obtendremos varios resultados notables sobre los coeficientes de Fourier de Δ y j.

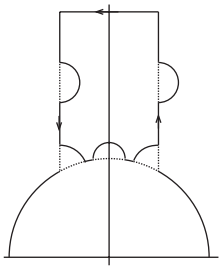
Si f es una función modular y $z \in H$, representamos por $o(f,z)$ al orden de f en z en el sentido usual de la teoría de funciones de variable compleja, mientras que si $z \in \mathbb{Q} \cup \{\infty\}$ es un punto parabólico, el orden $o(f,z)$ es el determinado por la serie de Fourier de f. El resultado básico es el siguiente:

Teorema 12.21 *Sea f una función modular no nula de grado $2k$. Entonces*

$$o(f,\infty) + \frac{1}{3}o(f,\rho) + \frac{1}{2}o(f,i) + \sum_{z \neq i,\,\rho} o(f,z) = \frac{k}{6},$$

donde la suma recorre los puntos $z \in H$ módulo Γ no equivalentes a ρ ni a i.

Demostración: Vamos a considerar la integral de la función meromorfa f'/f sobre el ciclo γ indicado en la figura:



donde los arcos se introducen para bordear posibles polos de f en la frontera del dominio fundamental y el segmento horizontal se toma suficientemente alto como para rodear todos los polos de f. (Si hubiera un polo sobre un punto no elíptico la circunferencia unidad estaría repetido simétricamente respecto al eje imaginario, y lo bordearíamos por arriba en un caso y por abajo en otro, de modo que γ sólo encerraría uno de los dos). Por simplicidad supondremos que no hay polos salvo a lo sumo en los puntos i y ρ, que son los casos más delicados. La prueba se modifica fácilmente para tratar otros polos. El teorema de los residuos nos da que

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(\zeta)}{f(\zeta)}\,d\zeta = \sum_{z\neq i,\,\rho} o(f,z).$$

Calculamos en primer lugar la integral sobre el segmento vertical, para lo cual hacemos el cambio de variable $z = e^{2\pi i\zeta}$, con lo que la integral se convierte en

$$-\frac{1}{2\pi i}\int_{|z|=r} \frac{f^{*\prime}(z)}{f^*(z)}\,dz = -o(f,\infty).$$

Las integrales sobre los segmentos verticales se cancelan porque f tiene periodo 1 y los segmentos se recorren en sentidos opuestos (esto sigue siendo cierto si hemos tenido que sortear polos).

Consideremos ahora la integral alrededor de ρ y su simétrico. Si f tiene orden m en ρ, entonces

$$\frac{f'(\zeta)}{f(\zeta)} = \frac{m}{\zeta-\rho} + h(\zeta),$$

donde h es holomorfa en ρ. Llamemos c_r al arco alrededor de ρ correspondiente a un radio r. Como h está acotada en un entorno de ρ, es claro que

$$\lim_{r\to 0}\frac{1}{2\pi i}\int_{c_r} h(\zeta) = 0.$$

Por otra parte,

$$\lim_{r\to 0}\frac{m}{2\pi i}\int_{c_r}\frac{d\zeta}{\zeta-\rho} = \frac{m}{2\pi i}\lim_{r\to 0}(\log c_r(1) - \log c_r(0)) = -\frac{m}{2\pi i}\frac{\pi i}{3} = -\frac{m}{6}.$$

El mismo resultado se obtiene sobre el punto simétrico de ρ respecto al eje imaginario, luego, en total, el límite cuando $r \to 0$ de la integral sobre los dos arcos es $-o(f,\rho)/3$.

El mismo argumento prueba que si hacemos tender a 0 el radio del arco alrededor de i la integral tiende a $-o(f,i)/2$.

Nos falta estudiar la integral sobre los dos arcos simétricos situados sobre la circunferencia unidad. Llamémoslos AB y $A'B'$. La aplicación $s(\tau) = -1/\tau$ transforma uno en el otro invirtiendo la orientación (y esto sigue siendo cierto si hemos tenido que sortear más polos con arcos intermedios). Por consiguiente, el cambio $\zeta = s(\xi)$ nos da que

$$\frac{1}{2\pi i}\int_A^B \frac{f'(\zeta)}{f(\zeta)}\,d\zeta = -\frac{1}{2\pi i}\int_{A'}^{B'} \frac{f'(s(\xi))}{\xi^2 f(s(\xi))}\,d\xi.$$

El hecho de que f sea modular se traduce en que $f(s(\xi)) = \xi^{2k} f(\xi)$, luego

$$\frac{df(s(\xi))}{d\xi} = \frac{f'(s(\xi))}{\xi^2} = \xi^{2k} f'(\xi) + 2k\xi^{2k-1} f(\xi).$$

Por lo tanto,

$$\frac{f'(s(\xi))}{\xi^2 f(s(\xi))} = \frac{f'(\xi)}{f(\xi)} + \frac{2k}{\xi}.$$

Concluimos que

$$\frac{1}{2\pi i}\int_A^B \frac{f'(\zeta)}{f(\zeta)}\,d\zeta = -\frac{1}{2\pi i}\int_{A'}^{B'} \frac{f'(\xi)}{f(\xi)}\,d\xi - \frac{2k}{2\pi i}\int_{A'}^{B'} \frac{d\xi}{\xi},$$

con lo que la suma de las integrales sobre los dos arcos considerados resulta ser

$$\frac{2k}{2\pi i}\int_{B'}^{A'} \frac{d\xi}{\xi}.$$

(Notemos que hemos invertido los extremos, de modo que ahora el arco se recorre en sentido positivo).

Si hubiéramos rodeado otros polos situados sobre el arco, es claro que la integral sobre el arco modificado coincide con la integral sobre el arco de circunferencia sin modificar (porque ahora el integrando $1/\xi$ no tiene polos sobre la circunferencia), y la integral es simplemente $k/(2\pi)$ por la longitud del arco $A'B'$. Cuando hacemos tender a 0 los radios de los arcos alrededor de los puntos elípticos, dicha longitud tiende a $2\pi/12$, luego la integral tiende a $k/6$.

En definitiva, si llamamos γ_r al ciclo formado con arcos de radio r alrededor de los polos, hemos probado que

$$\lim_{r\to 0} \frac{1}{2\pi i}\int_{\gamma_r} \frac{f'(\zeta)}{f(\zeta)}\,d\zeta = \frac{k}{6} - \frac{1}{3}o(f,\rho) - \frac{1}{2}o(f,i).$$

La conclusión es ahora inmediata. ∎

Como primera consecuencia vemos que no existen formas modulares de grado negativo (pues el miembro izquierdo en la fórmula del teorema anterior es no negativo si la función es una forma modular). En realidad podemos decir mucho más:

Teorema 12.22 *Las funciones g_2 y g_3 son algebraicamente independientes y el álgebra de las formas modulares es $M = \mathbb{C}[g_2, g_3]$.*

DEMOSTRACIÓN: Vamos a estudiar primeramente la estructura de los espacios M_{2k}. Observemos en general que si f es una forma modular no nula de grado $2k$ los órdenes que aparecen en el teorema anterior son todos mayores o iguales que 0.

Sabemos que $M_0 = \mathbb{C}$, pues las formas modulares de grado 0 se corresponden con funciones holomorfas (sin polos) sobre una esfera, luego han de ser constantes. También podemos obtenerlo del teorema anterior, que muestra que si $f \in M_0$ no es idénticamente nula no tiene ceros, de modo que $f - f(i)$ ha de ser idénticamente nula (pues se anula en i), luego f es constante.

Si $f \in M_2$, el término derecho de la fórmula del teorema anterior es $1/6$, luego la igualdad es imposible, por lo que ha de ser $f = 0$ y así $M_2 = 0$.

Veamos ahora que $M_4 = \langle g_2 \rangle$. Si $f \in M_2$ no es nula, el miembro derecho de la fórmula del teorema anterior es $1/3$, luego f ha de tener un único cero simple en ρ. En particular, g_2 sólo se anula en ρ, luego existe una constante c tal que $f - cg_2$ se anula en ∞, lo que sólo puede suceder si $f - cg_2 = 0$, con lo que $f \in \langle g_2 \rangle$.

El mismo argumento prueba que $M_6 = \langle g_3 \rangle$.

Veamos ahora que $M_8 = \langle g_2^2 \rangle$. Si $f \in M_8$ es no nula, el miembro derecho de la fórmula es $2/3$, que sólo puede conseguirse si f tiene un único cero doble en ρ. Por lo tanto, tomando una constante para la cual $f - cg_2^2$ tenga un cero adicional, llegamos a que $f \in \langle g_2^2 \rangle$.

Análogamente vemos que $M_{10} = \langle g_2 g_3 \rangle$: Si $f \in M_{10}$ no es nula, el miembro derecho de la fórmula es $5/6$, que sólo puede conseguirse si $o(f, \rho) = o(f, i) = 1$.

Supongamos ahora que $k \geq 6$ y sea $f \in M_{2k}$ una forma no nula. Supongamos como hipótesis de inducción que toda forma de grado menor que $2k$ se expresa como polinomio en g_2 y g_3. Sea $k = 2r$ o $k = 2r + 1$. Podemos tomar una constante c tal que $f - cg_2^r$ o bien $f - cg_2^{r-1} g_3$ se anule en ∞. Equivalentemente, podemos suponer que f se anula en ∞. En tal caso, como Δ tiene un cero simple en ∞, el cociente f/Δ es holomorfo en ∞, con lo que claramente $f/\Delta \in M_{2k-12}$. Por hipótesis de inducción f/Δ es un polinomio en g_2 y g_3, luego f también.

Falta probar que g_2 y g_3 son algebraicamente independientes. Supongamos que verifican una relación polinómica $P(g_2, g_3) = 0$, donde podemos suponer que el polinomio $P(X, Y)$ tiene grado mínimo. Puesto que los espacios M_{2k} tienen suma directa, todos los monomios que aparecen en $P(g_2, g_3)$ han de corresponder a formas del mismo grado m. No puede aparecer una potencia de g_2, pues entonces la relación sería $g_2^r + g_3 Q(g_2, g_3) = 0$, lo cual es absurdo, ya que $g_2(i) \neq 0$, $g_3(i) = 0$. Por consiguiente g_3 aparece en todos los monomios, lo que nos permite simplificarlo y llegar a una relación de grado menor, contradicción. ∎

Como consecuencia:

Teorema 12.23 *Para todo número natural k, se cumple que*

$$\dim M_{2k} = \begin{cases} E[k/6] & \text{si } k \equiv 1 \text{ (mód 6)}, \\ E[k/6]+1 & \text{si } k \not\equiv 1 \text{ (mód 6)}. \end{cases}$$

DEMOSTRACIÓN: En vista del teorema anterior, una base de M_{2k} la forman las funciones $g_2^r g_3^s$ tales que $4r+6s = 2k$, luego la dimensión es igual al número de naturales s tales que $2 \mid k-3s \geq 0$. Si k es par, s ha de ser par, luego la dimensión es igual al número de múltiplos de 6 menores o iguales que k, que es $E[k/6]+1$. Si k es impar s ha de ser impar, y la dimensión es el número de números de la forma $6t+3 \leq k$, que también es $E[k/6]+1$ excepto en el caso en que $k = 6k+1$, pues entonces falta $6k+3$ y hay uno menos. ∎

Observemos que para $k \geq 2$ el espacio M_{2k}^0 de las formas parabólicas de grado $2k$ tiene dimensión igual a una unidad menos que M_{2k}, pues la serie de Eisenstein $G_{2k}(\tau)$ es una forma modular de grado $2k$ que no se anula en ∞. La aplicación $M_{2k} \longrightarrow \mathbb{C}$ dada por $f \mapsto f(\infty)$ es, pues, un epimorfismo con núcleo M_{2k}^0.

Terminamos la sección con varias aplicaciones de los resultados que acabamos de obtener. Ante todo, conviene normalizar las funciones de Eisenstein:

Definición 12.24 Las *funciones de Eisenstein normalizadas* son las funciones

$$E_{2k}(\tau) = \frac{G_{2k}(\tau)}{2\zeta(2k)}.$$

Teniendo en cuenta 10.30 y la fórmula

$$\zeta(2k) = \frac{(-1)^{k+1} 2^{2k-1} \pi^{2k} B_{2k}}{(2k)!}$$

(donde B_{2k} son los números de Bernoulli), tenemos que

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) e^{2n\pi i\tau}.$$

Los primeros números de Bernoulli son

$$B_2 = 1/6, \quad B_4 = -1/30, \quad B_6 = 1/42,$$
$$B_8 = -1/30, \quad B_{10} = 5/66, \quad B_{12} = 691/2730.$$

Así, por ejemplo,

$$\begin{aligned} E_4(\tau) &= 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) e^{2n\pi i\tau}, \\ E_6(\tau) &= 1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) e^{2n\pi i\tau}, \\ E_8(\tau) &= 1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) e^{2n\pi i\tau}, \\ E_{12}(\tau) &= 1 + \frac{65520}{691} \sum_{n=1}^{\infty} \sigma_{11}(n) e^{2n\pi i\tau}. \end{aligned}$$

Igualmente, la función discriminante normalizada es

$$\frac{\Delta(\tau)}{(2\pi)^{12}} = \sum_{n=1}^{\infty} \tau(n) e^{2\pi i n}.$$

Al conocer la dimensión de los espacios M_{2k} podemos establecer muchas relaciones entre estas funciones y, por consiguiente, entre sus coeficientes de Fourier. Por ejemplo, tenemos que $\dim M_8 = 1$, luego la función $f \mapsto f(\infty)$ es un isomorfismo y, en particular, $E_8 = E_4^2$, pues ambas formas toman el valor 1 en ∞. Al igualar los coeficientes de Fourier obtenemos la relación

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m).$$

Un caso más notable se obtiene al considerar el espacio M_{12}, que tiene dimensión 2. Tenemos que E_{12}, E_6^2, $\Delta/(2\pi)^{12} \in M_{12}$, luego tienen que ser linealmente dependientes. Más concretamente, viendo la serie de Fourier

$$E_6^2(\tau) = 1 - 1008 \sum_{n=1}^{\infty} \sigma_5(n) e^{2n\pi i\tau} + 254016 \sum_{n=1}^{\infty} \sum_{m=2}^{n-1} \sigma_5(m)\sigma_5(n-m) e^{2n\pi i\tau},$$

es claro que E_{12} y E_6^2 son linealmente independientes, luego ha de ser

$$\Delta/(2\pi)^{12} = \lambda E_{12} + \mu E_6^2,$$

para ciertos escalares λ, $\mu \in \mathbb{C}$. Igualando los coeficientes para $n = 1$ y $n = 2$ obtenemos las ecuaciones

$$\lambda + \mu = 0, \qquad \frac{65520}{691}\lambda - 1008\mu = 1,$$

de donde obtenemos $\lambda = -\mu = \dfrac{691}{762048}$. Igualando los coeficientes n-simos resulta

$$\tau(n) = \frac{65}{756}\sigma_{11}(n) - \frac{691}{756}\sigma_5(n) - \frac{691}{3} \sum_{m=1}^{n-1} \sigma_5(m)\sigma_5(n-m). \tag{12.5}$$

Tenemos así una fórmula explícita para calcular la función de Ramanujan $\tau(n)$. En particular, $756 \cdot 3 \cdot \tau(n) \equiv 3 \cdot 65\sigma_{11}(n)$ (mód 691) o, lo que es lo mismo:

$$\tau(n) \equiv \sigma_{11}(n) \text{ (mód 691)}, \tag{12.6}$$

que es una de las congruencias de Ramanujan sobre la función $\tau(n)$.

Ahora vamos a encontrar una fórmula para calcular los coeficientes de Fourier de la función modular

$$j(\tau) = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} c(n) e^{2\pi i n\tau}.$$

Para ello consideramos las funciones E_6^2, $\Delta/(2\pi)^{12}$, $j\Delta/(2\pi)^{12} \in M_{12}$. La dimensión del espacio es 2 y las series de Fourier

$$E_6^2(\tau) = 1 - 1008\sum_{n=1}^{\infty}\sigma_5(n)e^{2\pi in\tau} + 504^2\sum_{n=2}^{\infty}\sum_{k=1}^{n-1}\sigma_5(k)\sigma_5(n-k)e^{2\pi in\tau},$$

$$j(\tau)\Delta(\tau)/(2\pi)^{12} = 1 + \sum_{n=1}^{\infty}\tau(n+1)e^{2\pi in\tau} + \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}c(k)\tau(n-k)e^{2\pi in\tau},$$

muestran que $\Delta/(2\pi)^{12}$ y $j\Delta/(2\pi)^{12}$ son linealmente independientes.

Por consiguiente ha de ser $E_6^2 = \lambda j\Delta/(2\pi)^{12} + \mu\Delta$. Igualando coeficientes para $n = 0$ obtenemos que $\lambda = 1$, mientras que para $n = 1$ resulta:

$$-1008 = \lambda(\tau(2) + c(0)\tau(1)) + \mu\tau(1).$$

Necesitamos los valores $\tau(2) = -24$, que se calcula con (12.5), y $c(0) = 744$, que puede calcularse refinando la prueba de 10.32. Con esto obtenemos $\mu = -12^3$. Así pues, hemos probado que

$$E_6^2 = (j - 12^3)\Delta/(2\pi)^{12}.$$

Al igualar los coeficientes n-simos obtenemos la relación

$$-1008\sigma_5(n) + 504^2\sum_{k=0}^{n-1}\sigma_5(k)\sigma_5(n-k) = \tau(n+1) + \sum_{k=0}^{n-1}c(k)\tau(n-k) - 12^3\tau(n).$$

La expresión se simplifica si en (12.5) despejamos

$$-1008\sigma_5(n) + 504^2\sum_{k=0}^{n-1}\sigma_5(k)\sigma_5(n-k) = \frac{65520}{691}\sigma_{11}(\tau) - \frac{762048}{691}\tau(n).$$

Al igualar las dos fórmulas (y usando de nuevo que $c(0) = 744$) obtenemos

$$\tau(n+1) + \sum_{k=1}^{n-1}c(k)\tau(n-k) - 984\tau(n) = \frac{65520}{691}(\sigma_{11}(\tau) - \tau(n)) - 1008\tau(n),$$

o también

$$\sum_{k=1}^{n-1}c(k)\tau(n-k) + \tau(n+1) + 24\tau(n) = \frac{65520}{691}(\sigma_{11}(\tau) - \tau(n)).$$

Esta fórmula permite calcular recurrentemente los coeficientes $c(n)$. Notemos que, como el miembro izquierdo es entero y el primo 691 no divide a 65520, esta fórmula implica también la congruencia (12.6).

12.5 La función eta de Dedekind

Vamos a estudiar ahora una función introducida por Dedekind en 1877 y que, aunque no es modular, está estrechamente relacionada con las funciones modulares, especialmente con la función Δ. Con su ayuda demostraremos, entre otras cosas, una fórmula debida a Jacobi, según la cual

$$\Delta(\tau) = (2\pi)^{12} e^{2\pi i\tau} \prod_{n=1}^{\infty} (1 - e^{2n\pi i\tau})^{24}.$$

Precisamente, la función eta es la que se obtiene al eliminar el exponente 24 en el miembro derecho (y la constante):

Definición 12.25 La *función eta de Dedekind* es la función

$$\eta(\tau) = e^{\pi i\tau/12} \prod_{n=1}^{\infty} (1 - e^{2n\pi i\tau}),$$

definida en el semiplano H.

Observemos que el producto infinito $\prod_{n=1}^{\infty} (1 - z^n)$ es absolutamente convergente para $|z| < 1$, luego ciertamente $\eta(\tau)$ es una función holomorfa en H que no se anula en ningún punto. En términos de la función η, la fórmula de Jacobi es

$$\Delta(\tau) = (2\pi)^{12} \eta^{24}(\tau).$$

Esencialmente, el problema es demostrar que η^{24} es una forma modular de grado 12. En primer lugar notamos que

$$\eta(\tau + 1) = e^{\pi i(\tau+1)/12} \prod_{n=1}^{\infty} (1 - e^{2n\pi i(\tau+1)}) = e^{\pi i/12} \eta(\tau),$$

luego $\eta^{24}(\tau + 1) = \eta^{24}(\tau)$.

El punto más delicado es determinar el comportamiento de η bajo la transformación $\tau \mapsto -1/\tau$. La prueba clásica utiliza funciones elípticas, pero aquí daremos una prueba corta debida a Siegel.

Teorema 12.26 *Para todo $\tau \in H$ se cumple*

$$\eta(-1/\tau) = (-i\tau)^{1/2} \eta(\tau),$$

donde la raíz cuadrada es la rama uniforme que es positiva sobre el semieje real positivo.

Demostración: Basta probar la relación cuando $\tau = yi$, con $y > 0$, pues el caso general se sigue entonces por prolongación analítica. Hemos de ver, pues, que $\eta(i/y) = y^{1/2}\eta(iy)$. Es inmediato que tanto $\eta(i/y)$ como $\eta(iy)$ son números reales positivos. Por lo tanto, esta relación equivale a

$$\log \eta(i/y) - \log \eta(iy) = \frac{1}{2} \log y.$$

Ahora bien,

$$\log\eta(iy) = -\frac{\pi y}{12} + \sum_{n=1}^{\infty}\log(1-e^{-2ny}) = -\frac{\pi y}{12} - \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{e^{-2\pi mny}}{m}$$

$$= -\frac{\pi y}{12} - \sum_{m=1}^{\infty}\frac{1}{m}\frac{e^{-2\pi my}}{1-e^{-2\pi my}} = -\frac{\pi y}{12} + \sum_{m=1}^{\infty}\frac{1}{m(1-e^{2\pi my})}.$$

Por consiguiente, hemos de probar que

$$\sum_{m=1}^{\infty}\frac{1}{m(1-e^{2\pi my})} - \sum_{m=1}^{\infty}\frac{1}{m(1-e^{2\pi m/y})} - \frac{\pi}{12}\left(y-\frac{1}{y}\right) = -\frac{1}{2}\log y. \tag{12.7}$$

Definimos

$$F_n(z) = -\frac{1}{8z}\cot(\pi iNz)\,\cot\frac{\pi Nz}{y}, \qquad N = n+1/2.$$

Sea C el paralelogramo de vértices $\pm y$, $\pm i$ recorrido en sentido positivo. Dentro de C, la función F_n tiene polos simples en los puntos $z = im/N$ y $z = my/N$, para $m = \pm1, \pm2, \ldots, \pm n$, así como un polo triple en $z = 0$.

Teniendo en cuenta que $\cot z = 1/z - z/3 + \cdots$ es fácil ver que el residuo de $F_n(z)$ en 0 es $i(y - y^{-1})/24$.

El residuo en $z = im/N$ es

$$\frac{1}{8\pi m}\cot\frac{\pi im}{y}.$$

Como esta expresión en par en m, vemos que

$$\sum_{\substack{m=-n\\ m\neq0}}^{n}\operatorname*{Res}_{z=im/N}F_n(z) = 2\sum_{m=1}^{n}\frac{1}{8\pi m}\cot\frac{\pi im}{y}.$$

Pero

$$\cot i\theta = \frac{\cos i\theta}{\operatorname{sen} i\theta} = i\frac{e^{-\theta}+e^{\theta}}{e^{-\theta}-e^{\theta}} = -i\frac{e^{2\theta}+1}{e^{2\theta}-1} = \frac{1}{i}\left(1-\frac{2}{1-e^{2\theta}}\right).$$

Tomando $\theta = \pi m/y$ queda

$$\sum_{\substack{m=-n\\ m\neq0}}^{n}\operatorname*{Res}_{z=im/N}F_n(z) = \frac{1}{4\pi i}\sum_{m=1}^{n}\frac{1}{m} - \frac{1}{2\pi i}\sum_{m=1}^{n}\frac{1}{m(1-e^{2\pi m/y})}.$$

Similarmente,

$$\sum_{\substack{m=-n\\ m\neq0}}^{n}\operatorname*{Res}_{z=my/N}F_n(z) = \frac{i}{4\pi}\sum_{m=1}^{n}\frac{1}{m} - \frac{i}{2\pi}\sum_{m=1}^{n}\frac{1}{m(1-e^{2\pi my})}.$$

Resulta, pues, que la suma de los residuos de F_n dentro de C multiplicados por $2\pi i$ es precisamente el miembro izquierdo de (12.7). El teorema quedará probado si vemos que

$$\lim_{n\to\infty}\int_C F_n(z)\,dz = -\frac{1}{2}\log y.$$

De la definición de cotangente se sigue fácilmente que

$$\cot \pi i N z = i\,\frac{(e^{-\pi z})^{2n+1}+1}{(e^{-\pi z})^{2n+1}-1}.$$

De aquí se sigue que

$$\lim_n \cot \pi i N Z = \begin{cases} i & \text{si } \operatorname{Re} z > 0,\\ -i & \text{si } \operatorname{Re} z < 0.\end{cases}$$

Similarmente,

$$\cot\frac{\pi N z}{y} = i\,\frac{(e^{i\pi z/y})^{2n+1}+1}{(e^{i\pi z/y})^{2n+1}-1},$$

de donde

$$\lim_n \frac{\pi N z}{y} = \begin{cases} i & \text{si } \operatorname{Im} z < 0,\\ -i & \text{si } \operatorname{Im} z > 0.\end{cases}$$

Esto implica que $\lim_n zF_n(z) = \pm 1/8$ cuando z está sobre el paralelogramo C salvo quizá en sus vértices. El signo es positivo en los lados que unen y con i y $-y$ con $-i$, mientras que es negativo en los otros dos.

Las expresiones anteriores muestran también que las funciones $zF_n(z)$ están uniformemente acotadas en C. Por ejemplo, si $\operatorname{Re} z > 0$ tenemos que $|e^{-\pi z}| < 1$ y, como z varía en un compacto (uno de los lados de C) existe, de hecho, un $\epsilon > 0$ tal que $|e^{-\pi z}| < 1-\epsilon$, luego también $|(e^{-\pi z})^{2n+1}| < 1-\epsilon$, luego $|(e^{-\pi z})^{2n+1} - 1| \ge \epsilon$ y así obtenemos una cota de $\cot \pi i N z$ independiente de n.

Puesto que la función $1/z$ está acotada en C, concluimos que las funciones $F_n(z)$ están uniformemente acotadas en C y podemos aplicar el teorema de la convergencia acotada de Lebesgue:

$$\lim_n \int_C F_n(z)\,dz = \int_C \lim_n F_n(z)\,dz =$$

$$\frac{1}{8}\left(-\int_{-i}^{y}\frac{dz}{z}+\int_y^i\frac{dz}{z}-\int_i^{-i}\frac{dz}{z}+\int_{-y}^{-i}\frac{dz}{z}\right) = \frac{1}{4}\left(-\int_{-i}^{y}\frac{dz}{z}+\int_y^i\frac{dz}{z}\right)$$

$$= \frac{1}{4}\left(-\left(\log y+\frac{\pi i}{2}\right)+\left(\frac{\pi i}{2}-\log y\right)\right) = -\frac{1}{2}\log y.$$

■

Como consecuencia tenemos que $\eta^{24}(-1/\tau) = \tau^{12}\eta^{24}(\tau)$. Antes hemos visto que η^{24} es invariante por $\tau \mapsto \tau+1$. Como la función Δ cumple también estos dos hechos, vemos que la función $f(\tau) = \Delta(\tau)/\eta^{24}(\tau)$ es invariante por los dos

generadores del grupo modular, luego es, de hecho, invariante por todo el grupo. Claramente es holomorfa en H y no se anula. Veamos su comportamiento en el punto ∞. En primer lugar:

$$\eta^{24}(\tau) = e^{2\pi i\tau}\prod_{n=1}^{\infty}(1-e^{2n\pi i\tau})^{24} = g^*(e^{2\pi i\tau}),$$

donde $g^*(z) = z\prod_{n=1}^{\infty}(1-x^n)^{24}$ es holomorfa con un cero simple en $z=0$. Además $g^*(z)/z$ toma el valor 1 en 0

Teniendo en cuenta la serie de Fourier de Δ, es claro que f es holomorfa en ∞ y $f(\infty) = (2\pi)^{12}$. En definitiva, f es una función modular de grado 0 sin ceros. Esto sólo puede ser si f es constante. Concretamente, ha de ser $f = (2\pi)^{12}$. Con esto queda probada la fórmula de Jacobi:

Teorema 12.27 *Para todo $\tau \in H$ se cumple*

$$\Delta(\tau) = (2\pi)^{12}e^{2\pi i\tau}\prod_{n=1}^{\infty}(1-e^{2n\pi i\tau})^{24}.$$

En las secciones siguientes veremos otras aplicaciones de la función η.

12.6 Funciones modulares respecto a $\Gamma_0(N)$

En esta sección mostraremos ejemplos de funciones modulares respecto a los grupos $\Gamma_0(N)$. Como aplicación demostraremos algunos casos particulares de las congruencias (10.5) sobre los coeficientes de la función de Klein.

Ya conocemos un ejemplo de función modular respecto de $\Gamma_0(N)$, a saber, la función $j_N(\tau) = j(N\tau)$. De hecho en 12.4 hemos visto que el cuerpo de las funciones modulares respecto de $\Gamma_0(N)$ es $\mathbb{C}(j, j_N)$. Más aún, conocemos el polinomio mínimo de j_N sobre $\mathbb{C}(j)$, dado por (12.1). En el caso particular en que n es un primo p, dicho polinomio es

$$F_p(j,Y) = (Y - j_p(\tau))\prod_{b=0}^{p-1}\left(Y - j\left(\tfrac{\tau+b}{p}\right)\right).$$

Así pues, la traza de j_p es

$$\operatorname{Tr} j_p(\tau) = j_p(\tau) + \sum_{b=0}^{p-1} j\left(\tfrac{\tau+b}{p}\right) \in \mathbb{C}(j).$$

Definición 12.28 Si p es un número primo, definimos la función

$$f_p(\tau) = \frac{1}{p}\sum_{b=1}^{p-1} j\left(\tfrac{\tau+b}{p}\right) = \frac{1}{p}(\operatorname{Tr} j_p(\tau) - j_p(\tau)).$$

Obviamente se trata de una función modular respecto de $\Gamma_0(p)$. Su interés radica en su serie de Fourier. Recordemos que

$$j(\tau) = e^{-2\pi i\tau} + \sum_{n=0}^{\infty} c(n)e^{2n\pi i\tau},$$

luego

$$f_p(\tau) = \frac{1}{p}\sum_{b=1}^{p-1}\sum_{n=-1}^{\infty} c(n)e^{2n\pi i(\tau+b)/p} = \frac{1}{p}\sum_{n=-1}^{\infty} c(n)e^{2n\pi i\tau/p}\sum_{b=1}^{p-1} e^{2n\pi ib/p}.$$

Ahora usamos que

$$\sum_{b=1}^{p-1}(e^{2\pi ib/p})^n = \begin{cases} 0 & \text{si } p \nmid n, \\ p & \text{si } p \mid n. \end{cases}$$

Así,

$$f_p(\tau) = \sum_{n=0}^{\infty} c(pn)e^{2n\pi i\tau}.$$

Vamos a usar esta función para obtener las congruencias (10.5) para $m = 1$. El caso general se obtiene aplicando repetidamente los argumentos que vamos a emplear. De momento hacemos una disgresión que nos llevará a un método general de construcción de funciones modulares.

Sea f una función cuasimodular de grado $2k$ respecto a un subgrupo Γ de índice finito en $\mathrm{LE}(2,\mathbb{Z})$. Para comprobar si es meromorfa en un punto parabólico $s \in \mathbb{Q} \cup \{\infty\}$ tomamos una transformación $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $sg = \infty$ y estudiamos la serie de Fourier de $g^{-1}|_{2k}f$. Vamos a probar que en realidad podemos tomar cualquier $\alpha \in \mathrm{LE}(2,\mathbb{R})$ (que cumpla $s\alpha = \infty$), si bien entonces el orden de la serie de Fourier que obtengamos no tendrá por qué coincidir con el orden de f en s.

En efecto, se comprueba inmediatamente que $\alpha^{-1}|_{2k}f$ es invariante por Γ^α, y en particular tiene periodo $h = |\bar\Gamma^\alpha : \bar\Gamma^\alpha_\infty| = |\bar\Gamma : \bar\Gamma_s|$. Suponemos que $\alpha^{-1}|_{2k}f$ no tiene polos en un semiplano y que admite un desarrollo en serie de Fourier $(\alpha^{-1}|_{2k}f)(\tau) = f^*(e^{2\pi i\tau/h})$, donde $f^*(z)$ es una función meromorfa en un entorno de 0.

Tomemos, por otra parte, $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $sg = \infty$. El hecho de que $\alpha^{-1}|_{2k}f$ no tenga polos en un semiplano se traduce en que f no tiene polos en un entorno (reducido) de s, y esto a su vez en que $g^{-1}|_{2k}f$ no tiene polos en un semiplano. Consideramos su serie de Fourier $(g^{-1}|_{2k}f)(\tau) = g^*(e^{2\pi i\tau/h})$. Hemos de probar que $g^*(z)$ es meromorfa en 0.

Ahora, la matriz $\beta = g^{-1}\alpha$ fija a ∞, luego es de la forma

$$\beta = \begin{pmatrix} a & 0 \\ b & 1/a \end{pmatrix}, \qquad a, b \in \mathbb{R},$$

luego $\tau\beta = a^2\tau + ab$. Consecuentemente:

$$g^*(e^{2\pi i\tau/h}) = (g^{-1}|_{2k}f)(\tau) = (\beta\alpha^{-1}|_{2k}f)(\tau) = (\beta|_{2k}(\alpha^{-1}|_{2k}f))(\tau)$$

$$= \frac{1}{a}(\alpha^{-1}|_{2k}f)(a^2\tau + ab) = \frac{1}{a}\, f^*(e^{2\pi iab/h}e^{2\pi i\tau a^2/h}) = \phi(e^{2\pi i\tau a^2/h}),$$

donde $\phi(z) = (1/a)f^*(e^{2\pi iab/h}z)$ es una función meromorfa en 0.

Vemos así que la función $g^*(e^{2\pi i\tau/h}) = \phi(e^{2\pi i\tau a^2/h})$ tiene periodos h y h/a^2. Esto sólo es posible si ambos son múltiplos enteros de un mismo periodo t, es decir, $h = ut$, $h/a^2 = vt$, donde u, $v \in \mathbb{N}$. Esto nos da un nuevo desarrollo en serie de Fourier $g^*(e^{2\pi i\tau/h}) = \phi(e^{2\pi i\tau a^2/h}) = \psi(e^{2\pi i\tau/t})$, donde $\psi(z)$ es una función holomorfa con una singularidad en 0.

Ahora bien, $\phi(e^{2\pi i\tau a^2/h}) = \psi((e^{2\pi i\tau a^2/h})^v)$, de donde se sigue inmediatamente que $\phi(z) = \psi(z^v)$ y, por lo tanto, ψ es meromorfa en 0. Igualmente, $g^*(z) = \psi(z^u)$, luego también g^* es meromorfa en 0, como había que probar.

Observemos que el orden de f en s es u/v por el orden de la serie de Fourier dada por α. Así pues, aunque el orden de la serie no es necesariamente el de f, al menos nos permite determinar si éste es mayor, menor o igual a 0.

Con ayuda de este hecho podemos extender el teorema 12.14:

Teorema 12.29 *Sean Γ y Γ' dos subgrupos de índice finito en $\mathrm{LE}(2,\mathbb{Z})$ y sea $\alpha \in \mathrm{LE}(2,\mathbb{R})$ tal que $\Gamma'^{\alpha} \subset \Gamma$. Si f es una función modular de grado $2k$ respecto de Γ, entonces $\alpha|_{2k}f$ es modular de grado $2k$ respecto de Γ'. Además $\alpha|_{2k}f$ es una forma modular o parabólica si y sólo si lo es f.*

Demostración: El teorema 12.14 nos da que $\alpha|_{2k}f$ es cuasimodular. Sólo falta probar que es meromorfa en los puntos parabólicos de Γ'. Ahora, bien, α biyecta los puntos parabólicos de Γ con los de Γ'^{α}, que son los mismos que los de Γ, a saber, los puntos de $\mathbb{Q} \cup \{\infty\}$.

Si $s \in \mathbb{Q} \cup \{\infty\}$, llamamos $s' = s\alpha \in \mathbb{Q} \cup \{\infty\}$ y tomamos un $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $s'g = \infty$. Como f es meromorfa en s', tenemos que $g^{-1}|_{2k}f$ tiene un desarrollo en serie de Fourier con coeficientes nulos por debajo de un cierto índice. Ahora bien,

$$g^{-1}|_{2k}f = (g^{-1}\alpha^{-1}\alpha)|_{2k}f = \beta^{-1}|_{2k}(\alpha|_{2k}f),$$

donde $\beta = \alpha g \in \mathrm{LE}(2,\mathbb{R})$ cumple $s\beta = s\alpha g = s'g = \infty$. Por el argumento previo al teorema, esto prueba que $\alpha|_{2k}f$ es meromorfa en s. Más aún, podemos decir que $\alpha|_{2k}f$ tiene un cero o un polo en s si y sólo si f lo tiene en s', de donde se sigue la última parte del enunciado. ∎

Veamos un caso particular:

Teorema 12.30 *Si f es una función modular de grado $2k$ respecto al grupo $\Gamma_0(1) = \mathrm{LE}(2,\mathbb{Z})$ y $N > 0$ es un número natural, entonces $f_N(\tau) = f(N\tau)$ es una función modular de grado $2k$ respecto a $\Gamma_0(N)$. Además f_N es una forma modular o una forma parabólica si y sólo si lo es f.*

Demostración: Basta aplicar el teorema anterior con

$$\alpha = \begin{pmatrix} 0 & \sqrt{N} \\ -1/\sqrt{N} & 0 \end{pmatrix}.$$

En efecto, si

$$\beta = \begin{pmatrix} a & Nc \\ b & d \end{pmatrix} \in \Gamma_0(N),$$

entonces

$$\alpha^{-1}\beta\alpha = \begin{pmatrix} 0 & -\sqrt{N} \\ 1/\sqrt{N} & 0 \end{pmatrix} \begin{pmatrix} a & Nc \\ b & d \end{pmatrix} \begin{pmatrix} 0 & \sqrt{N} \\ -1/\sqrt{N} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} d & -Nb \\ -c & a \end{pmatrix} \in \Gamma_0(1).$$

Además,

$$(\alpha|_{2k} f)(\tau) = f(-1/N\tau)(\sqrt{N}\tau)^{-2k} = f(N\tau)(N\tau)^{2k}(N\tau)^{-2k}N^k = N^k f(N\tau).$$

Obviamente entonces, $f(N\tau)$ también es una función modular respecto de $\Gamma_0(N)$. ∎

Consideremos en particular el caso de la función $\Delta(N\tau)$. Se trata de una forma parabólica respecto a $\Gamma_0(N)$. Teniendo en cuenta que

$$\Delta(\tau) = (2\pi)^{12} \sum_{n=1}^{\infty} \tau(n) e^{2\pi n i \tau},$$

vemos que

$$\Delta(N\tau) = (2\pi)^{12} \sum_{n=1}^{\infty} \tau(n) e^{2\pi N n i \tau},$$

luego $\Delta(N\tau)$ tiene un cero de orden N en ∞. La función

$$\phi(\tau) = \frac{\Delta(N\tau)}{\Delta(\tau)}$$

es claramente una función modular (de grado 0) respecto a $\Gamma_0(N)$ y tiene un cero de orden $N-1$ en ∞. Además no tiene ceros ni polos finitos. Si suponemos que N es primo, entonces el teorema 11.40 nos da que la superficie $X_0(N)$ tiene exactamente $\nu_\infty = 2$ puntos parabólicos, luego $\phi(\tau)$ tiene un único cero de orden $N-1$ y un único polo (en el otro punto parabólico) también de orden $N-1$.

Supongamos además que $X_0(N)$ tiene género 0, lo cual sucede exactamente para los primos $N = 2, 3, 5, 7, 11, 13$. Una función meromorfa en una esfera con un único cero y un único polo, ambos de orden $N-1$ tiene obviamente una raíz de orden $N-1$, por lo que podemos definir (salvo una raíz de la unidad) la función $\sqrt[N-1]{\phi(\tau)}$, que genera todo el cuerpo de las funciones modulares (de grado 0) respecto a $\Gamma_0(N)$. Exceptuando el caso $N = 11$, tenemos una expresión explícita para este generador:

Teorema 12.31 *Consideremos uno de los primos* $p = 2, 3, 5, 7, 13$. *Entonces el cuerpo de las funciones modulares de grado* 0 *respecto del grupo* $\Gamma_0(p)$ *está generado por la función*

$$\Phi(\tau) = \frac{\eta^r(p\tau)}{\eta^r(\tau)},$$

donde $r = 24/(p-1)$.

DEMOSTRACIÓN: Según hemos observado, las funciones modulares respecto de $\Gamma_0(p)$ están generadas por cualquier raíz $p-1$-ésima de $\phi(\tau) = \Delta(p\tau)/\Delta(\tau)$. En términos de la función η de Dedekind es

$$\phi(\tau) = \frac{\eta^{24}(p\tau)}{\eta^{24}(\tau)} = \Phi^{p-1}(\tau).$$

■

Observemos que hemos tenido que excluir $p = 11$ porque $11 - 1 = 10 \nmid 24$.

Sabemos que la función $\Phi(\tau)$ tiene un único 0 simple en ∞ y un polo simple en el otro punto parabólico de $X_0(p)$. Es fácil ver que ninguna función de $\Gamma_0(p)$ transforma 0 en ∞, luego $\Phi(\tau)$ tiene su polo en 0.

Obtendremos las congruencias que buscamos expresando la función f_p como polinomio en Φ. Usando dos veces la definición de f_p vemos que

$$f_p(-1/p\tau) = \frac{1}{p}\operatorname{Tr}(j_p)(-1/p\tau) - \frac{1}{p}j(-1/\tau) =$$

$$= \frac{1}{p}\operatorname{Tr}(j_p)(p\tau) - \frac{1}{p}j(\tau) = f_p(p\tau) + \frac{1}{p}j(p^2\tau) - \frac{1}{p}j(\tau),$$

luego

$$pf_p(-1/p\tau) = j(\tau) + pf_p(p\tau) + j(p^2\tau) = f^*(e^{2\pi i\tau}),$$

con

$$f^*(z) = z^{-p^2} + z^{-1} + F(z), \tag{12.8}$$

donde $F(z)$ es holomorfa en 0 y su serie de Taylor tiene coeficientes enteros.

Por otra parte, según la definición de $\eta(\tau)$ tenemos que

$$\Phi(\tau) = \frac{e^{2\pi irp\tau/24}\prod_{n=1}^{\infty}(1-e^{2n\pi ip\tau})^r}{e^{2\pi i\tau/24}\prod_{n=1}^{\infty}(1-e^{2n\pi i\tau})^r} = e^{2\pi i\tau}\prod_{n=1}^{\infty}\left(\frac{1-e^{2\pi i\tau pn}}{1-e^{2\pi i\tau n}}\right)^r = g^*(e^{2\pi i\tau}),$$

donde

$$g^*(z) = z\prod_{n=1}^{\infty}\left(\frac{1-z^{pn}}{1-z^n}\right)^r = z\prod_{n=1}^{\infty}(1+z^n+\cdots+z^{n(p-1)})^r.$$

Es fácil ver que

$$g^*(z) = z + d_2z^2 + d_3z^3 + \cdots$$

donde los coeficientes d_i son enteros. En otros términos, la serie de Fourier de $\Phi(\tau)$ es

$$\Phi(\tau) = e^{2\pi i\tau} + \sum_{n=2}^{\infty} d_n e^{2\pi i\tau n}, \qquad d_n \in \mathbb{Z}.$$

El teorema 12.26 nos da la relación

$$\Phi(-1/p\tau) = \frac{\eta^r(-1/\tau)}{\eta^r(-1/p\tau)} = \frac{(-i\tau)^{r/2}\eta^r(\tau)}{(-ip\tau)^{r/2}\eta^r(p\tau)} = \frac{1}{p^{r/2}\Phi(\tau)}.$$

Definimos

$$\psi(\tau) = p^{r/2}\Phi(-1/p\tau) = \frac{1}{\Phi(\tau)} = h^*(e^{2\pi i\tau}),$$

donde

$$h^*(z) = \frac{1}{z} + H(z), \tag{12.9}$$

y la función $H(z)$ es holomorfa en 0, cumple $H(0) = 1$ y su serie de Taylor tiene coeficientes enteros. Combinando (12.8) y (12.9) vemos que

$$pf_p(-1/p\tau) - \psi^{p^2}(\tau) = K_0(e^{2\pi i\tau}),$$

donde $K_0(z)$ tiene un polo en 0 de orden $\leq p^2 - 1$ y su serie de Laurent tiene coeficientes enteros. Por consiguiente, existe $b_1 \in \mathbb{Z}$ tal que

$$pf_p(-1/p\tau) - \psi^{p^2}(\tau) - b_1\psi^{p^2-1}(\tau) = K_1(e^{2\pi i\tau}),$$

donde $K_1(z)$ tiene un polo en 0 de orden $\leq p^2 - 2$ y su serie de Laurent tiene coeficientes enteros. Al cabo de p^2 pasos llegamos a una función

$$\chi(-1/p\tau) = pf_p(-1/p\tau) - \psi^{p^2}(\tau) - b_1\psi^{p^2-1} - \cdots - b_{p^2-1}\psi(\tau) = K_{p^2-1}(e^{2\pi i\tau}),$$

donde $K_{p^2-1}(z)$ es holomorfa en 0 y los coeficientes b_n son enteros. Cambiando τ por $-1/p\tau$ vemos que

$$\chi(\tau) = pf_p(\tau) - (p^{r/2}\Phi(\tau))^{p^2} - b_1(p^{r/2}\Phi(\tau))^{p^2-1} - \cdots - b_{p^2-1}p^{r/2}\Phi(\tau).$$

Esta función $\chi(\tau)$ es modular respecto a $\Gamma_0(p)$, es holomorfa en H y también en ∞, pues tanto Φ como f_p lo son. Por último, sabemos que $\chi(-1/p\tau)$ es holomorfa en ∞, lo que se traduce en que $\chi(\tau)$ es también holomorfa en 0. (Aquí usamos que para comprobar la holomorfía en 0 podemos llevarlo a ∞ mediante la función $-1/p\tau$, aunque no esté en $\mathrm{LE}(2,\mathbb{Z})$.)

En conclusión χ es una función modular sin polos, luego es constante. Teniendo en cuenta que $\phi(\infty) = 0$, la constante es concretamente $\chi(\infty) = pc(0)$. Con esto tenemos:

$$f_p(\tau) = p^{r/2-1}(a_1\Phi(\tau) + \cdots + a_{p^2}\Phi^{p^2}(\tau)) + c(0).$$

Igualando los coeficientes de Fourier de ambos miembros vemos que

$$c(pn) \equiv 0 \ (\text{mód } p^{r/2-1}), \qquad n > 0.$$

Teniendo en cuenta que $r = 24/(p-1)$, esto equivale a

$$\begin{aligned} c(2n) &\equiv 0 \ (\text{mód } 2^{11}),\\ c(3n) &\equiv 0 \ (\text{mód } 3^5),\\ c(5n) &\equiv 0 \ (\text{mód } 5^2),\\ c(7n) &\equiv 0 \ (\text{mód } 7). \end{aligned}$$

Notemos que para $p = 13$ se obtiene una congruencia trivial. De hecho, se comprueba que $13 \nmid c(13)$, por lo que no existen congruencias similares para $p = 13$. Por el contrario, aunque nos hemos visto obligados a excluir el caso, lo cierto es que también se da la congruencia $c(11^m n) \equiv 0 \pmod{11^m}$.

12.7 Funciones modulares respecto a $\Gamma(2)$

Vamos a mostrar algunos ejemplos de funciones modulares respecto del grupo $\Gamma(2)$. Entre ellas se encuentran las funciones $e_i(\tau)$, para $i = 1, 2, 3$. Recordemos que en 10.33 hemos visto que son holomorfas en H. Ahora probamos que son invariantes por $\Gamma(2)$. Llamemos $\wp(z;\tau)$ a la función de Weierstrass del retículo $\langle 1, \tau\rangle_{\mathbb{Z}}$. De la propia definición se sigue que, para una transformación unimodular arbitraria,

$$\wp\left(\frac{z}{c\tau+d}; \frac{a\tau+b}{c\tau+d}\right) = \frac{(c\tau+d)^2}{z^2}$$
$$+ \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \left(\frac{(c\tau+d)^2}{(z-m(a\tau+b)-n(c\tau+d))^2} - \frac{(c\tau+d)^2}{(m(a\tau+b)+n(c\tau+d))^2}\right).$$

Como la transformación es unimodular, $\langle 1, \tau\rangle_{\mathbb{Z}} = \langle a\tau+b, c\tau+d\rangle_{\mathbb{Z}}$, de donde llegamos claramente a que

$$\wp\left(\frac{z}{c\tau+d}; \frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2\wp(z;\tau).$$

Eligiendo adecuadamente los valores de z resulta:

$$\begin{aligned}
e_1\left(\frac{a\tau+b}{c\tau+d}\right) &= \wp\left(\frac{1}{2}; \frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2\wp\left(\frac{c\tau+d}{2};\tau\right), \\
e_2\left(\frac{a\tau+b}{c\tau+d}\right) &= \wp\left(\frac{(a\tau+b)/2}{c\tau+d}; \frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2\wp\left(\frac{a\tau+b}{2};\tau\right), \\
e_3\left(\frac{a\tau+b}{c\tau+d}\right) &= (c\tau+d)^2\wp\left(\frac{(a+c)\tau+(b+d)}{2};\tau\right).
\end{aligned}$$

La periodicidad de $\wp$ permite reducir módulo 2 los coeficientes a, b, c, d, con lo que los miembros derechos son de la forma $(c\tau+d)^2 e_i(\tau)$, pero el índice i depende de la paridad de los coeficientes.

En particular, si la transformación está en $\Gamma(2)$ tenemos que a y d son impares, mientras que b y c son pares, y entonces las relaciones se reducen a

$$e_i\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 e_i(\tau), \qquad i = 1, 2, 3.$$

Esto prueba que las funciones e_i son cuasimodulares respecto de $\Gamma(2)$. Conviene observar la acción concreta de los generadores del grupo modular sobre las funciones e_i:

$$e_1(\tau+1) = e_1(\tau), \quad e_2(\tau+1) = e_3(\tau), \qquad e_3(\tau+1) = e_2(\tau),$$

$$e_1(-1/\tau) = \tau^2 e_2(\tau), \quad e_2(-1/\tau) = \tau^2 e_1(\tau), \quad e_3(-1/\tau) = \tau^2 e_3(\tau).$$

Teorema 12.32 *Las funciones $e_1(\tau)$, $e_2(\tau)$, $e_3(\tau)$ son formas modulares de grado 2 respecto del grupo $\Gamma(2)$.*

DEMOSTRACIÓN: Sólo nos falta comprobar que son holomorfas en los puntos parabólicos. Calcular directamente las series de Fourier sería farragoso, así que lo comprobaremos indirectamente. Observemos que la menor traslación en $\Gamma(2)$ es $\tau \mapsto \tau + 2$, luego las series de Fourier son de la forma $f^*(e^{\pi i \tau})$.

Sea $s \in \mathbb{Q} \cup \{\infty\}$ un punto parabólico y tomemos $h \in \mathrm{LE}(2, \mathbb{Z})$ de modo que $sh = \infty$. Tenemos que $e_i(\tau)$ satisface la ecuación

$$4e_i^3(\tau) - g_2(\tau)e_i(\tau) - g_3(\tau) = 0,$$

luego también

$$4e_i^3(\tau h^{-1}) - g_2(\tau h^{-1})e_i(\tau h^{-1}) - g_3(\tau h^{-1}) = 0.$$

Multiplicando por $j_{h^{-1}}(\tau)^6$ llegamos a que

$$4(h^{-1}|_2 e_i)^3(\tau) - (h^{-1}|_4 g_2)(\tau)(h^{-1}|_2 e_i)(\tau) - (h^{-1}|_6 g_3)(\tau) = 0.$$

Pongamos que

$$(h^{-1}|_2 e_i)(\tau) = f^*(e^{\pi i \tau}), \quad (h^{-1}|_4 g_2)(\tau) = g_2^*(e^{\pi i \tau}), \quad (h^{-1}|_6 g_3)(\tau) = g_3^*(e^{\pi i \tau}),$$

donde $g_2^*(z)$ y $g_3^*(z)$ son funciones holomorfas en 0 y $f^*(z)$ es holomorfa con una singularidad en 0. Tenemos que, para $z \neq 0$,

$$4f^{*3}(z) = g_2^*(z)f^*(z) + g_3^*(z),$$

luego

$$\frac{g_2^*(z)}{4f^{*2}(z)} + \frac{g_3^*(z)}{4f^{*3}(z)} = 1.$$

Basta probar que $f^*(z)$ está acotada en un entorno de 0, y sabemos que g_2^* y g_3^* lo están. Si no fuera así, podríamos tomar una sucesión $\{z_n\}$ convergente a 0 de modo que $f^*(z_n)$ tendiera a ∞, con lo que la igualdad anterior nos daría $0 = 1$. ∎

En la prueba del teorema anterior ha sido crucial que la ecuación satisfecha por las funciones e_i fuera homogénea (de grado 6) respecto a los grados de las funciones involucradas. Si nos restringimos a funciones de grado 0 no necesitamos dicha homogeneidad, y podemos enunciar un resultado general:

Teorema 12.33 *Sea Γ un subgrupo de índice finito en $\mathrm{LE}(2, \mathbb{Z})$ y sea f una función cuasimodular de grado 0 respecto de Γ algebraica sobre el cuerpo de las funciones modulares de grado 0 respecto de Γ. Entonces f es modular respecto de Γ.*

DEMOSTRACIÓN: Sólo hay que probar que es meromorfa en los puntos parabólicos. Supongamos que

$$f^n(\tau) + a_{n-1}(\tau)f^{n-1}(\tau) + \cdots + a_0(\tau) = 0,$$

donde las funciones $a_i(\tau)$ son modulares respecto de Γ. Sea $s \in \mathbb{Q} \cup \{\infty\}$ un punto parabólico y $g \in \mathrm{LE}(2,\mathbb{Z})$ tal que $sg = \infty$. Entonces

$$f^n(\tau g^{-1}) + a_{n-1}(\tau g^{-1})f^{n-1}(\tau g^{-1}) + \cdots + a_0(\tau g^{-1}) = 0,$$

o también

$$1 = -\frac{a_{n-1}(\tau g^{-1})}{f(\tau g^{-1})} - \cdots - \frac{a_0(\tau g^{-1})}{f^n(\tau g^{-1})}.$$

Sabemos que $a_i(\tau g^{-1}) = a_i^*(e^{2\pi i\tau/h})$, para cierta función $a_i^*(z)$ meromorfa en 0. Las funciones $a_i(\tau g^{-1})$ no tienen polos en un cierto semiplano $\operatorname{Im}\tau > r$, luego $f(\tau g^{-1})$ tampoco puede tenerlos (en un polo la igualdad sería $1 = 0$). Por lo tanto $f(\tau g^{-1}) = f^*(e^{2\pi i\tau/h})$, para una cierta función $f^*(z)$ holomorfa en un entorno reducido de 0.

Tenemos que

$$1 = -\frac{a_{n-1}^*(z)}{f^*(z)} - \cdots - \frac{a_0^*(z)}{f^{*n}(z)}.$$

Sea k un natural mayor que el orden de cualquier polo en 0 de cualquiera de las funciones $a_i^*(z)$. Entonces

$$1 = -\frac{z^k a_{n-1}^*(z)}{z^k f^*(z)} - \cdots - \frac{z^{kn} a_0^*(z)}{(z^k f^*)^n(z)}.$$

Los numeradores están acotados en un entorno de 0, luego la función $z^k f^*(z)$ también ha de estarlo, o de lo contrario llegaríamos de nuevo a que $1 = 0$. Esto prueba que $f^*(z)$ es meromorfa en 0, luego $f(\tau)$ es meromorfa en s. ∎

Volviendo a las funciones $e_i(\tau)$, tenemos que cada $g \in \Gamma(1) = \mathrm{LE}(2,\mathbb{Z})$ las permuta mediante $e_i \mapsto g^{-1}|_2 e_i$. En otros términos, tenemos un homomorfismo de grupos $\rho : \bar\Gamma(1) \longrightarrow \Sigma_{\{e_1,e_2,e_3\}}$, cuya imagen contiene al menos las trasposiciones (e_2, e_3) y (e_1, e_2) (las imágenes de $\tau \mapsto \tau+1$ y $\tau \mapsto -1/\tau^2$). Estas trasposiciones generan todo el grupo de permutaciones, luego ρ es un epimorfismo. Su núcleo contiene a $\bar\Gamma(2)$ y la fórmula (11.1) nos da que $|\bar\Gamma(1) : \bar\Gamma(2)| = 6$, luego el núcleo es exactamente $\bar\Gamma(2)$. En definitiva tenemos que $\bar\Gamma(1)/\bar\Gamma(2) \cong \Sigma_{\{e_1,e_2,e_3\}}$.

Esta representación nos permite calcular la acción de $\Gamma(1)$ sobre el cuerpo $K = A_0(\Gamma(2))$ de las funciones modulares de grado 0 respecto de $\Gamma(2)$. Nos referimos a la acción $\rho : \bar\Gamma(1) \longrightarrow G(K/k)$ dada por $\rho(g)(f) = g^{-1}|_0 f$, donde $k = A_0(\Gamma(1))$.

Ante todo, recordemos que $A_0(\Gamma(2))$ y $A_0(\Gamma(1))$ son los cuerpos de funciones meromorfas de las superficies modulares $X(2)$ y $X(1)$, así como que —según el teorema 11.32— el grado $|\bar\Gamma(1) : \bar\Gamma(2)|$ coincide con el grado de la aplicación natural $\phi : X(2) \longrightarrow X(1)$, que a su vez coincide con el grado de la extensión de

los cuerpos de funciones meromorfas. En resumen: $|K:k|=6$. Si demostramos que ρ es un epimorfismo, tendremos entonces que induce una representación $\bar{\Gamma}(1)/\bar{\Gamma}(2)\cong G(K/k)$. Para ello consideramos la función

$$\lambda(\tau)=\frac{e_1(\tau)-e_3(\tau)}{e_2(\tau)-e_3(\tau)}.$$

Obviamente $\lambda\in K$ y, dado que $\Gamma(1)$ permuta las formas e_i de todas las formas posibles, es claro que transforma a λ en las seis funciones

$$\lambda=\frac{e_1-e_3}{e_2-e_3},\quad 1-\lambda=\frac{e_1-e_2}{e_3-e_2},\quad \frac{1}{\lambda}=\frac{e_2-e_3}{e_1-e_3},\quad \frac{1}{1-\lambda}=\frac{e_3-e_2}{e_1-e_2},$$

$$\frac{\lambda-1}{\lambda}=\frac{e_2-e_1}{e_3-e_1},\quad \frac{\lambda}{\lambda-1}=\frac{e_3-e_1}{e_2-e_1}.$$

Así pues, estas seis funciones son k-conjugadas de λ y son distintas dos a dos (o de lo contrario λ sería una función algebraica sobre $\mathbb{C}$, luego sería constante, y no es el caso). Concluimos que $K=k(\lambda)=\mathbb{C}(j,\lambda)$. Además, esto prueba que $\Gamma(1)$ induce al menos seis automorfismos distintos, por lo que la representación ρ es un epimorfismo, como queríamos probar.

La función λ tiene una interpretación geométrica muy sencilla: para cada $\tau\in H$, el valor $\lambda(\tau)$ cumple que la ecuación de Legendre

$$Y^2=X(X-1)(X-\lambda(\tau))$$

determina una curva elíptica de invariante $j(\tau)$. En efecto, en principio sabemos que una curva con invariante $j(\tau)$ es la dada por la ecuación de Weierstrass

$$Y^2=4(X-e_1)(X-e_2)(X-e_3).$$

Queremos otra con raíces 0, 1 y $\lambda(\tau)$. El cambio de variables $X'=X-e_3$ nos da una curva isomorfa de ecuación

$$Y^2=4X(X-(e_1-e_3))(X-(e_2-e_3)).$$

Ahora hacemos $X=(e_2-e_3)X'$ y obtenemos

$$Y^2=4(e_2-e_3)^3X(X-1)(X-\lambda(\tau)).$$

Finalmente, el cambio $Y=\sqrt{4(e_2-e_3)^3}Y'$ (para una elección arbitraria de la raíz cuadrada) nos da la ecuación de Legendre que buscábamos.

Ahora, la observación tras la definición 2.15 nos da que

$$j=2^8\,\frac{(\lambda^2-\lambda+1)^2}{\lambda^2(\lambda-1)^2}.$$

Operando se obtiene el polinomio mínimo de λ sobre $\mathbb{C}(j)$, pero es más interesante esta expresión, pues muestra que $\mathbb{C}(j,\lambda)=\mathbb{C}(\lambda)$, es decir, que el

cuerpo de las funciones modulares de grado 0 respecto de $\Gamma(2)$ es simplemente $\mathbb{C}(\lambda)$.

Otros ejemplos de funciones modulares respecto de $\Gamma(2)$ son

$$\sqrt{\Delta} = (2\pi)^6\eta^{12} = \pm 4(e_1-e_2)(e_1-e_3)(e_2-e_3),$$

que es una forma modular de grado 6, y

$$\sqrt{j-1728} = \frac{\sqrt{27\cdot 1728}g_3}{\sqrt{\Delta}},$$

que es una función modular de grado 0.

El ejemplo siguiente nos da otra prueba de la modularidad de $\sqrt{\Delta}$ que no depende de los razonamientos anteriores, a la vez que nos proporciona un ejemplo de función modular respecto de $\Gamma(3)$.

Ejemplo Teniendo en cuenta que

$$j = \frac{1728g_2^3}{\Delta} = \frac{1728g_2^3}{(2\pi)^{12}\eta^{24}}, \quad j-1728 = \frac{27\cdot 1728g_3^2}{\Delta} = \frac{27\cdot 1728g_3^2}{(2\pi)^{12}\eta^{24}},$$

podemos definir las funciones

$$j^{1/3} = \frac{12g_2}{(2\pi)^4\eta^8}, \qquad \sqrt{j-1728} = \frac{\sqrt{27\cdot 1728}g_3}{(2\pi)^6\eta^{12}}.$$

Ciertamente son holomorfas en H. Se comprueba inmediatamente que

$$j^{1/3}(\tau+1) = e^{-2\pi i/3}j^{1/3}(\tau), \qquad j^{1/3}(-1/\tau) = j^{1/3}(\tau),$$

$$\sqrt{j(\tau+1)-1728} = -\sqrt{j(\tau)-1728}, \qquad \sqrt{j(1/\tau)-1728} = -\sqrt{j(\tau)-1728}.$$

Vamos a probar que $j^{1/3}$ es modular respecto a $\Gamma(3)$ y $\sqrt{j-1728}$ es modular respecto a $\Gamma(2)$. En efecto, las funciones $f(\tau) = j^{1/3}(\tau)$ y $g(\tau) = \sqrt{j(\tau)-1728}$ son linealmente independientes sobre $\mathbb{C}$ (o en caso contrario $\tau \mapsto \tau+1$ actuaría igual sobre ambas), luego el espacio vectorial $V = \langle f, g\rangle$ tiene dimensión 2. Si llamamos $\Gamma = \mathrm{LE}(2,\mathbb{Z})$ tenemos un homomorfismo $\phi : \bar{\Gamma} \longrightarrow \operatorname{Aut} V$ dado por $\phi(\alpha)(f)(\tau) = f(\tau\alpha)$.

La imagen es el subgrupo de $\operatorname{Aut} V$ generado por los automorfismos que en la base (f,g) tienen matrices

$$\begin{pmatrix} e^{-2\pi i/3} & 0 \\ 0 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

La segunda es el cubo de la primera, luego la imagen es un grupo cíclico de orden 6. Si llamamos N al núcleo de ϕ, tenemos que $\bar{\Gamma}/N$ es un grupo abeliano de 6 elementos. En particular, el subgrupo derivado de $\bar{\Gamma}$ cumple $\bar{\Gamma}' \leq N$.

Por otra parte, $\bar{\Gamma}$ está generado por $t(\tau) = \tau + 1$ y $s(\tau) = -1/\tau$, luego también por ts y s, que tienen orden 3 y 2 respectivamente, luego $\bar{\Gamma}/\bar{\Gamma}'$ tiene a lo sumo 6 elementos, y concluimos que $N = \bar{\Gamma}'$.

Ahora es claro que los elementos de $\bar{\Gamma}$ que fijan a f forman el único subgrupo $\bar{\Gamma}' \leq H \leq \bar{\Gamma}$ de índice 3 en $\bar{\Gamma}$, mientras que los elementos de $\bar{\Gamma}$ que fijan a g forman el único subgrupo $\bar{\Gamma}' \leq K \leq \bar{\Gamma}$ de índice 2 en $\bar{\Gamma}$.

La fórmula (11.1) nos da que $|\bar{\Gamma} : \bar{\Gamma}(2)| = 6$, mientras que $|\bar{\Gamma} : \bar{\Gamma}(3)| = 12$. Existen subgrupos $\bar{\Gamma}(2) \leq K^* \leq \bar{\Gamma}$ y $\bar{\Gamma}(3) \leq H^* \leq \bar{\Gamma}$, normales en $\bar{\Gamma}$ y de índices 2 y 3 respectivamente (todo grupo de orden 12 tiene un subgrupo normal de orden 4). El homomorfismo natural $\bar{\Gamma} \longrightarrow (\bar{\Gamma}/H^*) \times (\bar{\Gamma}/K^*)$ es claramente un epimorfismo, luego su núcleo es $\bar{\Gamma}' = H^* \cap K^*$, luego concluimos que $H^* = H$, $K^* = K$ y, por consiguiente, los elementos de $\bar{\Gamma}(3)$ fijan a f y los de $\bar{\Gamma}(2)$ fijan a g.

Con esto hemos probado que f y g son cuasimodulares respecto a los grupos correspondientes. La modularidad se sigue ahora del teorema 12.33. ∎

Capítulo XIII

Multiplicación compleja

Ahora vamos a estudiar más profundamente las curvas elípticas con multiplicación compleja, es decir, las curvas con endomorfismos distintos de las multiplicaciones por enteros. Los hechos básicos sobre las multiplicaciones complejas de una curva elíptica las discutimos en la sección 10.3. Recordemos que toda curva elíptica $E/\mathbb{C}$ es analíticamente isomorfa a un toro complejo $\mathbb{C}/R$, con $R = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$. Según el teorema 10.20, la curva $E/\mathbb{C}$ tiene multiplicaciones complejas si y sólo si $\tau = \omega_2/\omega_1$ pertenece a un cuerpo cuadrático imaginario K. Puesto que podemos sustituir el retículo R por $\langle 1, \tau \rangle_{\mathbb{Z}}$, vemos que toda curva con multiplicación compleja puede representarse como un toro complejo determinado por un retículo R contenido en un cuerpo cuadrático imaginario K. Recíprocamente, si R es un retículo en estas condiciones, toda curva elíptica isomorfa al toro $\mathbb{C}/R$ tiene multiplicaciones complejas, y su anillo de endomorfismos es isomorfo al orden cuadrático

$$\mathcal{O}_R = \{\alpha \in K \mid \alpha R \subset R\}.$$

Recordemos así mismo que los retículos contenidos en un cuerpo cuadrático imaginario K son lo que en teoría de números se llaman módulos completos de K y $\mathcal{O}_R$ es lo que se llama el anillo de coeficientes del módulo R. Así pues, el objeto de este tema será estudiar las curvas elípticas representables por toros complejos determinados por los módulos de los cuerpos cuadráticos imaginarios.

13.1 Multiplicaciones ideales

Dos curvas elípticas asociadas a dos módulos completos de un cuerpo cuadrático imaginario son isomorfas si y sólo si los módulos son linealmente equivalentes o, en el lenguaje de la teoría de números, similares. Por lo tanto, las clases de isomorfía de curvas elípticas con un mismo anillo de multiplicaciones complejas $\mathcal{O}$ se corresponden biunívocamente con las clases de similitud de módulos con anillo de coeficientes $\mathcal{O}$.

Los módulos con anillo de coeficientes $\mathcal{O}$ tienen una estructura natural de grupo dada por

$$\langle \omega_1, \omega_2 \rangle \langle \eta_1, \eta_2 \rangle = \langle \omega_1\eta_1, \omega_2\eta_2, \omega_1\eta_2, \omega_2\eta_1 \rangle .$$

El conjunto de clases de similitud es precisamente el grupo cociente respecto del subgrupo formado por los módulos similares a $\mathcal{O}$. Es un grupo finito y su orden se llama número de clases de $\mathcal{O}$. Toda clase de similitud admite como representante a un ideal de $\mathcal{O}$.

A partir de aquí vamos a suponer que $\mathcal{O}$ es el orden maximal de K. Todo lo que sigue se generaliza sin dificultad a órdenes arbitrarios, pero sólo vamos a necesitar este caso y las pruebas resultan técnicamente más simples.

Consideremos una curva elíptica $E/\mathbb{C}$ cuyo anillo de endomorfismos sea isomorfo al orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Entonces E puede representarse como un toro complejo $\mathbb{C}/R$, donde R es un módulo completo cuyo anillo de coeficientes es $\mathcal{O}_K$. Si T es cualquier otro módulo completo con el mismo anillo de coeficientes, podemos definir $T * E$ como una curva elíptica analíticamente isomorfa al toro complejo $\mathbb{C}/T^{-1}R$. (Enseguida comprenderemos por qué tomamos el inverso de T.)

Observemos que $T * E$ sólo está definida salvo isomorfismo, y que, siempre salvo isomorfismo, depende únicamente de la clase de similitud de T y de la clase de isomorfía de E. Por ello no perdemos generalidad si en lugar de un módulo completo arbitrario T consideramos un ideal $\mathfrak{a}$ de $\mathcal{O}_K$. En tal caso $\mathfrak{a}R \subset R$, luego $R \subset \mathfrak{a}^{-1}R$ y tenemos un homomorfismo natural entre los toros complejos $\mathbb{C}/R$ y $\mathbb{C}/\mathfrak{a}^{-1}R$, que induce una isogenia $\mathfrak{a} : E \longrightarrow \mathfrak{a}*E$. (Esto explica el inverso en la definición de $*$.)

Hay que tener presente que, fijada una curva $\mathfrak{a} * E$, la isogenia $\mathfrak{a}$ sólo está determinada salvo automorfismos de las dos curvas (pues depende de la elección de los isomorfismos entre éstas y los toros complejos). La curva $\mathfrak{a} * E$ será isomorfa a E si y sólo si los módulos R y $\mathfrak{a}^{-1}R$ son similares, lo que equivale a que el ideal $\mathfrak{a}$ sea principal, digamos $\mathfrak{a} = (\alpha)$. En tal caso podemos tomar $(\alpha) * E = E$ y como isogenia inducida por (α) a la multiplicación por α. Así pues, la noción de "multiplicación por ideales" que acabamos de definir extiende a la de multiplicación compleja, con la peculiaridad de que la multiplicación por un ideal no principal no es un endomorfismo, sino una isogenia con imagen en otra curva.

Definimos

$$E[\mathfrak{a}] = \{P \in E \mid \alpha P = O \text{ para todo } \alpha \in \mathfrak{a}\}.$$

Aquí αP ha de entenderse como la acción sobre P del endomorfismo de E asociado a α a través del isomorfismo canónico dado por el teorema 10.21.

Al contrario que las definiciones anteriores, $E[\mathfrak{a}]$ no depende de ninguna elección arbitraria. Vamos a probar que es precisamente el núcleo de la multiplicación por $\mathfrak{a}$.

Fijemos isomorfismos $\phi : E \longrightarrow \mathbb{C}/R$ y $\psi : \mathfrak{a} * E \longrightarrow \mathbb{C}/\mathfrak{a}^{-1}R$. Llamemos $i : \mathbb{C}/R \longrightarrow \mathbb{C}/\mathfrak{a}^{-1}R$ al homomorfismo natural dado por $i([z]) = [z]$. Entonces la multiplicación por $\mathfrak{a}$ es la isogenia $\phi \circ i \circ \psi^{-1}$. Tomemos un punto $P \in E$ y sea $\phi(P) = [z]$. Tenemos que $\mathfrak{a}P = O$ si y sólo si $\psi^{-1}(i([z])) = 0$, si y sólo si $[z] = 0$ en $\mathbb{C}/\mathfrak{a}^{-1}R$, si y sólo si $z \in \mathfrak{a}^{-1}R$, si y sólo si $\mathfrak{a}z \subset R$, si y sólo si $\alpha z \in R$ para todo $\alpha \in \mathfrak{a}$, si y sólo si $\alpha[z] = 0$ para todo $\alpha \in \mathfrak{a}$, si y sólo si $\alpha P = O$ para todo $\alpha \in \mathfrak{a}$, si y sólo si $P \in E[\mathfrak{a}]$.

En particular, si $\alpha \in \mathcal{O}_K$, entonces $E[\alpha]$ es el núcleo de la multiplicación por α y si $m \in \mathbb{Z}$ entonces $E[m]$ en este sentido es el grupo que ya teníamos definido.

Ahora vamos a calcular el orden de $E[\mathfrak{a}]$. Para ello observamos primeramente que no perdemos generalidad si suponemos que R es un ideal $\mathfrak{b}$ de $\mathcal{O}_K$. Claramente $|E[\mathfrak{a}]| = |\mathfrak{a}^{-1}\mathfrak{b} : \mathfrak{b}|$.

Para calcular este índice observamos que, en general, si $\mathfrak{r} \subset \mathfrak{s}$ son ideales fraccionales de K y $\mathfrak{t}$ es un ideal de K, se cumple que $|\mathfrak{s} : \mathfrak{r}| = |\mathfrak{ts} : \mathfrak{tr}|$. En efecto: podemos tomar un $\alpha \in K^*$ tal que $\alpha\mathfrak{r}$ y $\alpha\mathfrak{s}$ sean ideales de $\mathcal{O}_K$. Es fácil ver que la multiplicación por α induce isomorfismos de grupos $\mathfrak{s}/\mathfrak{r} \cong (\alpha\mathfrak{s})/(\alpha\mathfrak{r})$ y $\mathfrak{ts}/\mathfrak{tr} \cong (\mathfrak{t}\alpha\mathfrak{s})/(\mathfrak{t}\alpha\mathfrak{r})$, luego podemos suponer que $\mathfrak{r}$ y $\mathfrak{s}$ son ideales de $\mathcal{O}_K$, en cuyo caso $|\mathfrak{ts} : \mathfrak{tr}| = \mathrm{N}(\mathfrak{tr})/\mathrm{N}(\mathfrak{ts}) = \mathrm{N}(\mathfrak{r})/\mathrm{N}(\mathfrak{s}) = |\mathfrak{s} : \mathfrak{r}|$.

Ahora es inmediato que $|E[\mathfrak{a}]| = |\mathfrak{a}^{-1}\mathfrak{b} : \mathfrak{b}| = |\mathfrak{b} : \mathfrak{ab}| = |\mathcal{O}_K : \mathfrak{a}| = \mathrm{N}(\mathfrak{a})$. Más aún, ahora es fácil determinar la estructura de $E[\mathfrak{a}]$: Fijemos un punto $P \in E[\mathfrak{a}]$ no nulo y consideremos la aplicación $\mathcal{O}_K \longrightarrow E[\mathfrak{a}]$ dada por $\alpha \mapsto \alpha P$. Obviamente es un homomorfismo de grupos y su núcleo es $\mathfrak{a}$, luego ha de ser $E[\mathfrak{a}] \cong \mathcal{O}_K/\mathfrak{a}$ (la suprayectividad se debe a que ambos grupos tienen el mismo cardinal).

Resumimos en un teorema lo que hemos obtenido:

Teorema 13.1 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Para cada ideal $\mathfrak{a}$ de $\mathcal{O}_K$, la multiplicación $\mathfrak{a} : E \longrightarrow \mathfrak{a} * E$ tiene núcleo $E[\mathfrak{a}] \cong \mathcal{O}_K/\mathfrak{a}$, luego su grado es igual a $\mathrm{N}(\mathfrak{a})$.*

Ahora necesitamos un resultado auxiliar de álgebra conmutativa:

Teorema 13.2 *Sea D un dominio de Dedekind, sea $\mathfrak{a}$ un ideal de D y sea M un D-módulo contenido en una extensión del cuerpo de cocientes de D. Entonces la aplicación $\phi : \mathfrak{a}^{-1}M \longrightarrow \mathrm{Hom}_D(\mathfrak{a}, M)$ dada por $\phi(x)(\alpha) = \alpha x$ es un isomorfismo.*

Demostración: Es claro que ϕ está bien definida y es un homomorfismo de D-módulos. Si $\phi(x) = 0$, entonces $\alpha x = 0$ para todo $\alpha \in \mathfrak{a}$, lo que sólo puede suceder si $x = 0$, luego ϕ es inyectiva.

Si $f \in \mathrm{Hom}_D(\mathfrak{a}, M)$, para todos los α, $\beta \in \mathfrak{a}$ tenemos que

$$\alpha f(\beta) = f(\alpha\beta) = \beta f(\alpha),$$

luego $x = \alpha^{-1}f(\alpha) = \beta^{-1}f(\beta) \in \mathfrak{a}^{-1}M$ es independiente del elemento $\alpha \in \mathfrak{a}$ con que se calcula y cumple

$$f(\alpha) = \alpha x = \phi(x)(\alpha).$$

Así pues, $f = \phi(x)$, luego ϕ es suprayectiva. ∎

Recordemos que si una curva elíptica $E/\mathbb{C}$ tiene multiplicación compleja, entonces su invariante $j(E)$ es un número algebraico (teorema 10.24), luego en su clase de isomorfía podemos encontrar un representante definido sobre $\mathbb{A}$. En particular, cada producto $\mathfrak{a} * E$ puede tomarse definido sobre $\mathbb{A}$.

Teorema 13.3 *Sea $E/\mathbb{A}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Para cada $\sigma \in G(\mathbb{A}/\mathbb{Q})$ y cada ideal $\mathfrak{a}$ de $\mathcal{O}_K$ se cumple:*

$$(\mathfrak{a} * E)^\sigma = \mathfrak{a}^\sigma * E^\sigma.$$

Demostración: Consideremos una sucesión exacta

$$\mathcal{O}_K^m \xrightarrow{A} \mathcal{O}_K^n \longrightarrow \mathfrak{a} \longrightarrow 0,$$

donde A es una aplicación lineal determinada por una matriz $m \times n$ con coeficientes en $\mathcal{O}_K$. (Tomamos un generador de $\mathfrak{a}$ con n elementos y llevamos la base canónica de $\mathcal{O}_K^n$ a dichos generadores; como $\mathcal{O}_K$ es noetheriano, el núcleo del epimorfismo así construido es finitamente generado, luego podemos repetir el proceso.) Por otra parte, si $E \cong \mathbb{C}/R$, tenemos una sucesión exacta

$$0 \longrightarrow R \longrightarrow \mathbb{C} \longrightarrow E \longrightarrow 0.$$

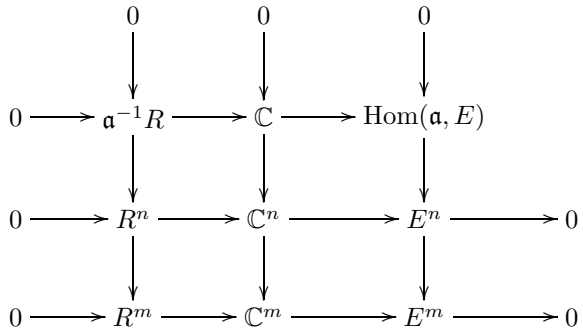
De ambas sucesiones obtenemos el diagrama siguiente, con todas las filas y columnas exactas:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \mathrm{Hom}(\mathfrak{a}, R) & \longrightarrow & \mathrm{Hom}(\mathfrak{a}, \mathbb{C}) & \longrightarrow & \mathrm{Hom}(\mathfrak{a}, E) \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^n, R) & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^n, \mathbb{C}) & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^n, E) \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^m, R) & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^m, \mathbb{C}) & \longrightarrow & \mathrm{Hom}(\mathcal{O}_K^m, E)
\end{array}$$

(Aquí "Hom" siempre hace referencia a homomorfismos de $\mathcal{O}_K$-módulos.) Claramente, para todo $\mathcal{O}_K$-módulo M se cumple que $\mathrm{Hom}(\mathcal{O}_K^n, M) \cong M^n$, mientras que el teorema anterior nos da los isomorfismos

$$\mathrm{Hom}(\mathfrak{a}, R) \cong \mathfrak{a}^{-1}R, \qquad \mathrm{Hom}(\mathfrak{a}, \mathbb{C}) \cong \mathfrak{a}^{-1}\mathbb{C} = \mathbb{C}.$$

Esto nos permite reescribir el diagrama anterior en la forma



Notemos que las dos últimas filas son obviamente exactas en E^n y E^m respectivamente. Las aplicaciones de la última fila de flechas verticales están determinadas por la matriz traspuesta A^t. En los dos primeros casos viendo a sus entradas como números complejos, en el tercero viéndolas como endomorfismos de E.

Llamemos H al núcleo del homomorfismo $E^n \longrightarrow E^m$. Del diagrama anterior podemos extraer la sucesión exacta

$$0 \longrightarrow \mathfrak{a}^{-1}R \longrightarrow \mathbb{C} \longrightarrow H \longrightarrow R^m/A^tR^n.$$

El último homomorfismo se obtiene tomando una antiimagen en $\mathbb{C}^n$, pasando a $\mathbb{C}^m$ y tomando una antiimagen en R^m. Una comprobación rutinaria muestra que esto siempre es posible, que la aplicación resultante no depende de las elecciones intermedias, que es un homomorfismo, así como la exactitud de la sucesión.

Tenemos que E^n es una variedad proyectiva regular. En particular es una variedad analítica compleja. La aplicación $\mathbb{C} \longrightarrow H \subset E^n$ es holomorfa, también lo es el homomorfismo inducido $\mathbb{C}/\mathfrak{a}^{-1}R \longrightarrow H \subset E^n$, así como la composición de éste con un isomorfismo $\mathfrak{a} * E \cong \mathbb{C}/\mathfrak{a}^{-1}R$. En definitiva, tenemos una aplicación holomorfa (que además es un monomorfismo de grupos) $\mathfrak{a} * E \longrightarrow H \subset E^n$. Llamemos $V \subset H$ a la imagen.

Tanto $\mathfrak{a} * E$ como E^n son variedades proyectivas regulares, toda aplicación holomorfa entre variedades proyectivas regulares es una aplicación regular, y la imagen de una variedad proyectiva regular por una aplicación regular es un conjunto algebraico, luego V es una variedad proyectiva (es irreducible porque la aplicación es biyectiva). Además V es un subgrupo de E^n, y las operaciones algebraicas vienen dadas por funciones regulares. En particular las traslaciones son isomorfismos (de variedades, no de grupos) y de aquí se sigue que V es regular (pues ha de tener al menos un punto regular y cualquier otro es la imagen de éste por una traslación).

En definitiva, V es una variedad proyectiva conformemente equivalente (luego isomorfa en todos los sentidos) a $\mathfrak{a} * E$. Veamos ahora que V es abierta en H respecto a la topología compleja. En efecto, tomemos un entorno U de 0 en $\mathbb{C}^m$ que no contenga ningún otro punto de R^m. Su antiimagen en $\mathbb{C}^n$ es un

entorno abierto de 0 y su imagen V' en E^n (que es abierta, porque las funciones holomorfas lo son) cumple que $V = V' \cap H$. En efecto si $P \in V' \cap H$, podemos tomarle una antiimagen en $\mathbb{C}^n$ que esté en la antiimagen de U, tomamos su imagen en U y, entonces, el único elemento de R^m al que podemos llegar es el 0, luego P está en el núcleo del homomorfismo $H \longrightarrow R^m/A^tR^n$, que es V.

Recíprocamente, si $P \in V$, tomamos una antiimagen $z \in \mathbb{C}^n$ y su imagen $z' \in \mathbb{C}^m$, y se cumple que, de hecho, $z' \in R^m$. Más aún, el hecho de que P esté en V significa que z' es la imagen de un $w \in R^n$. Si restamos a z la imagen de w en $\mathbb{C}^n$ obtenemos otro z con el que calcular la imagen de P, pero ahora su imagen en $\mathbb{C}^n$ es 0, luego está en U, luego z está en la antiimagen de U, luego $P \in V'$.

Con esto tenemos que V es abierto y cerrado en H, y es conexo, pues es una curva proyectiva, luego V es la componente conexa del elemento neutro (viendo a H como grupo topológico). El cociente H/V es isomorfo a un subgrupo del grupo numerable R^m/A^tR^n, luego H es unión disjunta de una cantidad numerable de componentes conexas isomorfas a V. En realidad sólo puede haber una cantidad finita de ellas, pues si fueran infinitas tendrían un punto de acumulación en el compacto E^n (fuera de H, pues cada componente es abierta y cerrada en H), pero entonces la aplicación $E^n \longrightarrow E^m$ tendría que anularse fuera de su núcleo, lo cual es absurdo.

Así pues, H es unión de un número finito de curvas disjuntas isomorfas a V. Dichas curvas han de ser las componentes irreducibles del conjunto algebraico H, luego V puede caracterizarse como la componente irreducible de H a la que pertenece el elemento neutro.

Tratemos de entender lo que hemos obtenido: en principio, la construcción de la curva $\mathfrak{a} * E$ es analítica, pues requiere tomar un toro complejo $\mathbb{C}/R$ analíticamente isomorfo a E, pasar al toro $\mathbb{C}/\mathfrak{a}^{-1}R$ y tomar una curva elíptica analíticamente isomorfa a este segundo toro. Lo que acabamos de probar es que existe una forma alternativa de construir $\mathfrak{a}*E$, más técnica, pero completamente algebraica. Concretamente, para construir $\mathfrak{a}*E$ obtenemos una matriz A a partir del ideal $\mathfrak{a}$. La obtención de A es algebraica y así, si tomamos $\sigma \in G(\mathbb{A}/\mathbb{Q})$, para el ideal $\mathfrak{a}^\sigma$ sirve la matriz A^σ.

A partir de la matriz A construimos la aplicación $\phi : E^n \longrightarrow E^m$ que resulta de interpretar las entradas de A como multiplicaciones complejas de E. La aplicación entre las curvas $(E^\sigma)^m \longrightarrow (E^\sigma)^n$ determinada por A^σ es precisamente ϕ^σ, en el sentido usual. Ahora tomamos su núcleo H (y observamos que el núcleo de la aplicación conjugada es H^σ), y dentro de H consideramos la componente irreducible V que contiene al elemento neutro (que en H^σ será V^σ). Lo que hemos demostrado es que podemos tomar $\mathfrak{a}*E = V$, luego también $\mathfrak{a}^\sigma * E^\sigma = V^\sigma$, y esto prueba el teorema. ∎

Consideremos ahora una curva elíptica compleja $E/\mathbb{A}$ cuyo anillo de multiplicaciones sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Si $\sigma \in G(\mathbb{A}/K)$, entonces E^σ es una curva elíptica con el mismo anillo de multiplicaciones complejas que E. La clase de isomorfía de E se corresponde con una clase de ideales C del grupo de clases $\mathcal{H}_K$ de K, y la clase de E^σ se corresponde

con otra clase C'. Si llamamos $F(\sigma) = CC'^{-1}$, entonces tenemos definida una aplicación

$$F : G(\mathbb{A}/K) \longrightarrow \mathcal{H}_K$$

con la propiedad de que $E^\sigma = F(\sigma) * E$ para todo $\sigma \in G(\mathbb{A}/K)$.

En principio podríamos haber definido F sobre $G(\mathbb{A}/\mathbb{Q})$ (no hemos usado en ningún momento que los automorfismos considerados fijen a K), pero vamos a ver que, definida sobre $G(\mathbb{A}/K)$, la aplicación F no depende de E. En efecto:

Teorema 13.4 *Si K es un cuerpo cuadrático imaginario, existe un único homomorfismo de grupos $F : G(\mathbb{A}/K) \longrightarrow \mathcal{H}_K$ que cumple*

$$E^\sigma = F(\sigma) * E$$

para todo $\sigma \in G(\mathbb{A}/K)$ y toda curva elíptica $E/\mathbb{A}$ con anillo de multiplicaciones complejas $\mathcal{O}_K$.

Demostración: Hemos probado que, fijada una curva E, existe una aplicación F que cumple lo pedido para todo σ. Veamos que es independiente de E. Para ello tomamos dos curvas elípticas $E_1/\mathbb{A}$ y $E_2/\mathbb{A}$ con anillo de multiplicaciones complejas $\mathcal{O}_K$ y llamamos F_1 y F_2 a las aplicaciones asociadas a cada una de ellas. Existe un ideal $\mathfrak{a}$ de $\mathcal{O}_K$ tal que $E_2 = \mathfrak{a} * E_1$. Para cada $\sigma \in G(\mathbb{A}/K)$, tenemos que

$$([\mathfrak{a}] * E_1)^\sigma = E_2^\sigma = F_2(\sigma) * E_2 = F_2(\sigma) * (\mathfrak{a} * E_1) = (F_2(\sigma)[\mathfrak{a}]F_1(\sigma)^{-1}) * E_1^\sigma.$$

Por el teorema anterior el término de la izquierda es $[\mathfrak{a}]^\sigma * E_1^\sigma = [\mathfrak{a}] * E_1^\sigma$, luego

$$[\mathfrak{a}] = F_2(\sigma)[\mathfrak{a}]F_1(\sigma)^{-1},$$

y concluimos que $F_1(\sigma) = F_2(\sigma)$. Ahora es fácil ver que se trata de un homomorfismo de grupos (notemos que en el razonamiento siguiente estamos usando que la F para E es la misma que la F para $F(\sigma) * E$):

$$F(\sigma\tau) * E = E^{\sigma\tau} = (F(\sigma) * E)^\tau = F(\tau) * (F(\sigma) * E) = (F(\sigma)F(\tau)) * E,$$

de donde podemos concluir que $F(\sigma\tau) = F(\sigma)F(\tau)$.

∎

Terminamos la sección con algunos resultados sobre reducción de isogenias. En general, puede demostrarse que si $\phi : E \longrightarrow E'$ es una isogenia entre dos curvas elípticas definidas sobre un cuerpo numérico K y $\mathfrak{p}$ es un primo no arquimediano de K donde ambas tienen buena reducción, entonces la isogenia reducida $\tilde{\phi} : \tilde{E} \longrightarrow \tilde{E}'$ tiene el mismo grado que ϕ. No obstante, la prueba es complicada y aquí nos limitaremos a demostrar algunos casos particulares.

Teorema 13.5 *Sea E/K una curva elíptica definida sobre un cuerpo numérico K y sea $\mathfrak{p}$ un primo no arquimediano de K donde E tenga buena reducción. Entonces, para cada isogenia $\phi \in \mathrm{End}(E)$ se cumple* $\operatorname{grad} \tilde{\phi} = \operatorname{grad} \phi$.

DEMOSTRACIÓN: Veamos en primer lugar que si E es una curva elíptica arbitraria y $\phi \in \mathrm{End}(E) \setminus \mathbb{Z}$, entonces la isogenia dual $\hat{\phi}$ es la única $\psi \in \mathrm{End}(E)$ que conmuta con ϕ y es raíz del mismo polinomio mónico de grado 2 con coeficientes enteros.

Ciertamente $\hat{\phi}$ cumple estas propiedades y si $\psi \neq \phi$ las cumple, entonces ϕ y ψ generan un cuerpo K en el anillo de división $\mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{End}(E)$. Dicho cuerpo ha de contener a $\phi^{-1} = (\operatorname{grad}\phi)^{-1}\hat{\phi}$, luego también a $\hat{\phi}$, pero entonces ψ y $\hat{\phi}$ son raíces de un mismo polinomio irreducible sobre $\mathbb{Q}$ con raíz ϕ, luego ha de ser $\psi = \hat{\phi}$.

En las hipótesis del teorema, es claro que $\tilde{\hat{\phi}}$ cumple las condiciones anteriores respecto de $\hat{\phi}$, luego ha de ser $\tilde{\hat{\phi}} = \hat{\tilde{\phi}}$. En consecuencia, si llamamos $m = \operatorname{grad}\phi$, tenemos que

$$m = \tilde{m} = \widetilde{\phi\hat{\phi}} = \tilde{\phi}\hat{\tilde{\phi}},$$

luego $\operatorname{grad}\tilde{\phi} = m$. (Notemos que $\tilde{m} = m$ porque la reducción es un homomorfismo de anillos). ∎

Teorema 13.6 *Sea $\phi : E_1 \longrightarrow E_2$ una isogenia definida sobre un cuerpo numérico K entre dos curvas elípticas definidas sobre K y sea $\mathfrak{p}$ un primo de K en el que ambas curvas tengan buena reducción. Sea p la característica del cuerpo de restos. Si $p \nmid \operatorname{grad}\phi$, entonces la reducción $\tilde{\phi} : \tilde{E}_1 \longrightarrow \tilde{E}_2$ cumple* $\operatorname{grad}\tilde{\phi} = \operatorname{grad}\phi$.

DEMOSTRACIÓN: Si H es el núcleo de ϕ, el teorema 6.19 nos da que la reducción es inyectiva sobre ϕ, luego el núcleo de $\tilde{\phi}$ contiene un subgrupo isomorfo a H, lo que demuestra que $\operatorname{grad}\phi \mid \operatorname{grad}\tilde{\phi}$.

Tomando $m = \operatorname{grad}\phi = \operatorname{grad}\hat{\phi}$, tenemos que $\phi\circ\hat{\phi} = m$, con lo que $\tilde{\phi}\circ\tilde{\hat{\phi}} = m$.

Por consiguiente $(\operatorname{grad}\tilde{\phi})(\operatorname{grad}\tilde{\hat{\phi}}) = m^2$. Puesto que $m \mid \operatorname{grad}\tilde{\phi}$ y $m \mid \operatorname{grad}\tilde{\hat{\phi}}$, ha de ser $m = \operatorname{grad}\tilde{\phi}$. ∎

13.2 El cuerpo de clases de Hilbert

En esta sección demostraremos que si $E/\mathbb{C}$ es una curva elíptica cuyo anillo de multiplicaciones complejas es el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K, entonces el cuerpo $K(j(E))$ es el cuerpo de clases de Hilbert de K, es decir, la máxima extensión abeliana no ramificada de K. Empezamos demostrando que la extensión es abeliana:

Teorema 13.7 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Entonces la extensión $K(j(E))/K$ es abeliana.*

DEMOSTRACIÓN: No perdemos generalidad si suponemos que E está definida sobre $\mathbb{A}$. Consideramos el homomorfismo $F : G(\mathbb{A}/K) \longrightarrow \mathcal{H}_K$ dado por

el teorema 13.4. Su núcleo es un subgrupo normal de $G(\mathbb{A}/K)$ de índice finito, luego es de la forma $G(\mathbb{A}/H)$, para cierta extensión normal H de K, que será abeliana, pues $G(H/K)$ es isomorfo a un subgrupo del grupo de clases $\mathcal{H}_K$.

Tenemos que un automorfismo $\sigma \in G(\mathbb{A}/K)$ cumple $\sigma \in G(\mathbb{A}/H)$ si y sólo si $F(\sigma) = 1$, si y sólo si $E^\sigma \cong E$, si y sólo si $j(E)^\sigma = j(E^\sigma) = j(E)$, si y sólo si $\sigma \in G(\mathbb{A}/K(j(E)))$.

Así pues, $G(\mathbb{A}/H) = G(\mathbb{A}/K(j(E)))$ y, por consiguiente, $H = K(j(E))$. ∎

Según hemos visto en la demostración del teorema anterior, a partir de aquí podemos considerar el monomorfismo

$$F : G(H/K) \longrightarrow \mathcal{H}_K,$$

donde $H = K(j(E))$ (y E es cualquier curva elíptica cuyo anillo de multiplicaciones complejas sea $\mathcal{O}_K$, pues todos los invariantes $j(E)$ generan el mismo cuerpo H).

El conjunto de clases de isomorfía de curvas elípticas con anillo de multiplicaciones complejas $\mathcal{O}_K$ es finito. Tomemos un conjunto de representantes $E_1, \ldots, E_h$. Sea L una extensión finita de K tal que todas las curvas E_i estén definidas sobre L. Más aún, podemos tomarlo de forma que toda isogenia entre dos curvas E_i esté definida sobre L. Ello se debe a que los anillos $\mathrm{Hom}(E_i, E_j)$ son $\mathbb{Z}$-módulos finitamente generados, luego basta asegurar que la condición se cumpla para un número finito de isogenias.

Llamamos S al conjunto finito formado por los primos siguientes:

a) Los primos que se ramifican en L.

b) Los primos tales que alguna curva E_i tiene mala reducción en uno de sus divisores en L.

c) Los divisores de uno de los números $\mathrm{N}_{\mathbb{Q}}^{L}(j(E_i) - j(E_k))$, para $i \neq k$.

Observemos que las normas que aparecen en la propiedad c) son números enteros, porque sabemos que los invariantes $j(E_i)$ son enteros algebraicos. Por simplicidad podemos añadir a S los primos 2 y 3.

Fijemos ahora un primo $p \notin S$ y supongamos además que se escinde en K, digamos $p = \mathfrak{p}\mathfrak{p}'$.

La clase de ideales $[\mathfrak{p}]^{-1}$ contiene infinitos primos. Tomemos uno $\mathfrak{q}$ tal que $\mathrm{N}(\mathfrak{q}) \neq p$. Entonces $\mathfrak{p}\mathfrak{q} = (\alpha)$, para cierto $\alpha \in \mathcal{O}_K$. Consideremos las aplicaciones dadas por el esquema siguiente:

$$\begin{array}{ccccccc}
\mathbb{C}/R & \longrightarrow & \mathbb{C}/\mathfrak{p}^{-1}R & \longrightarrow & \mathbb{C}/\alpha R & \xrightarrow{\alpha} & \mathbb{C}/R \\
\downarrow & & \downarrow & & & & \downarrow \\
E & \xrightarrow[\mathfrak{p}]{} & \mathfrak{p} * E & & \xrightarrow[\mathfrak{q}]{} & & E
\end{array}$$

Aquí R es un módulo completo de K tal que $\mathbb{C}/R$ es isomorfo a E, la primera flecha horizontal es la aplicación natural $[z] \mapsto [z]$, que se corresponde

a través de los isomorfismos verticales con la multiplicación por $\mathfrak{p}$. La segunda flecha horizontal es también la aplicación natural $[z] \mapsto [z]$ entre $\mathbb{C}/\mathfrak{p}^{-1}R$ y $\mathbb{C}/\mathfrak{q}^{-1}\mathfrak{p}^{-1}R = \mathbb{C}/\alpha R$. Este toro es isomorfo a $\mathbb{C}/R$ a través de la multiplicación por α, luego podemos tomar como $\mathfrak{q} * \mathfrak{p} * E$ la propia curva E.

Al componer toda la fila horizontal obtenemos la aplicación $[z] \mapsto [\alpha z]$, luego la fila vertical es la multiplicación compleja $\mathfrak{p} \circ \mathfrak{q} = \alpha$.

Sea $\mathfrak{P}$ un divisor de $\mathfrak{p}$ en L. Si tomamos ecuaciones de Weierstrass para E y $\mathfrak{p} * E$ minimales en $\mathfrak{P}$, entonces ambas curvas tienen buena reducción en $\mathfrak{P}$ (notemos que $\mathfrak{p} * E$ es una E_i). Los dos últimos teoremas de la sección anterior garantizan que

$$\operatorname{grad}\tilde{\alpha} = \operatorname{grad}\alpha = \mathrm{N}_{\mathbb{Q}}^{K}(\alpha), \quad \operatorname{grad}\tilde{\mathfrak{q}} = \operatorname{grad}\mathfrak{q} = \mathrm{N}_{\mathbb{Q}}^{K}(\mathfrak{q}),$$

y, teniendo en cuenta que el grado de la composición es el producto de los grados, también se ha de cumplir que

$$\operatorname{grad}\tilde{\mathfrak{p}} = \operatorname{grad}\mathfrak{p} = p.$$

Una diferencial de primera clase de E es

$$\omega = \frac{dx}{2y + a_1x + a_3},$$

donde a_1 y a_3 son los coeficientes de la ecuación de Weierstrass. Sabemos que

$$\overline{\alpha}(\omega) = \frac{d(\alpha \circ x)}{2(\alpha \circ y) + a_1(\alpha \circ x) + a_3} = \alpha\omega.$$

Aquí, $\alpha \circ x$ y $\alpha \circ y$ son funciones racionales en x e y, y

$$d(\alpha \circ x) = \frac{d(\alpha \circ x)}{dx}\, dx$$

puede calcularse por las reglas usuales de derivación, teniendo en cuenta que, por la ecuación de Weierstrass,

$$\frac{dy}{dx} = \frac{3x^2 + 2a_2x + a_4 - a_1y - a_3}{2y + a_1x}.$$

Sin más que aplicar estas operaciones se llega a una expresión de la forma

$$F(x, y)\, dx = \alpha\, \frac{dx}{y},$$

donde F es una función racional con coeficientes en $\mathcal{O}_L$. La igualdad equivale a que $yF(x, y) = \alpha$, lo cual a su vez se reduce a que cierto polinomio esté en el ideal que define a E. Todos estos pasos pueden seguirse igualmente para la curva $\tilde{E}$, y el resultado es que $\tilde{\bar{\alpha}}(\tilde{\omega}) = \tilde{\alpha}\tilde{\omega} = 0$, pues $\alpha \in \mathfrak{p} \subset \mathfrak{P}$.

Según el teorema 1.19, esto significa que la isogenia $\tilde{\alpha}$ no es separable, pero $\tilde{\mathfrak{q}}$ sí que lo es, ya que su grado es primo con p. Por consiguiente, ha de ser $\tilde{\mathfrak{p}}$ la

reducción inseparable y, como su grado es p, el teorema 1.22 implica que $\tilde{\mathfrak{p}}$ es precisamente la aplicación de Frobenius de grado p.

Concluimos que la curva $\widetilde{\mathfrak{p} * E}$ es isomorfa a la curva de Frobenius de $\tilde{E}$, luego su invariante es $j(\widetilde{\mathfrak{p} * E}) = j(\tilde{E})^p$. Equivalentemente:

$$j(\mathfrak{p} * E) \equiv j(E)^p \pmod{\mathfrak{P}}.$$

Consideramos ahora el cuerpo $H = K(j(E)) \subset L$ y el símbolo de Artin

$$\sigma = \left(\frac{H/K}{\mathfrak{p}}\right) \in G(H/K).$$

Si $\mathfrak{P}'$ es el primo de H que divide a $\mathfrak{P}$, se cumple que

$$j(E)^\sigma \equiv j(E)^{\mathrm{N}(\mathfrak{P}')} = j(E)^p \pmod{\mathfrak{P}'}.$$

Por consiguiente

$$j(\mathfrak{p} * E) \equiv j(E)^\sigma = j(E^\sigma) = j(F(\sigma) * E) \pmod{\mathfrak{P}}.$$

Las curvas $\mathfrak{p} * E$ y $F(\sigma) * E$ son isomorfas a dos de las curvas E_i, luego la condición c) en la definición del conjunto S implica que la congruencia ha de ser una igualdad:

$$j(\mathfrak{p} * E) = j(F(\sigma) * E).$$

Así pues, $\mathfrak{p} * E \cong F(\sigma) * E$ y esto implica que $F(\sigma) = [\mathfrak{p}]$. El teorema siguiente recoge lo que hemos obtenido:

Teorema 13.8 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K y sea $H = K(j(E))$. Entonces, para todos los primos $\mathfrak{p}$ de grado 1 de K salvo un número finito de ellos, se cumple que*

$$F\left(\frac{H/K}{\mathfrak{p}}\right) = [\mathfrak{p}].$$

Vamos a afinar considerablemente este resultado. Sea Δ el discriminante de la extensión H/K, de modo que el homomorfismo de Artin está definido sobre el grupo $I(\Delta)$ de los ideales fraccionales de K primos con Δ. Vamos a demostrar que para todo $\mathfrak{a} \in I(\Delta)$ se cumple

$$F\left(\frac{H/K}{\mathfrak{a}}\right) = [\mathfrak{a}].$$

Sea $\mathfrak{f}$ el conductor de la extensión H/K. Se trata de un ideal[1] de $\mathcal{O}_K$ tal que el subgrupo

$$P_\mathfrak{f} = \{(\alpha)/(\beta) \mid \alpha, \beta \in \mathcal{O}_K,\ (\alpha\beta, \Delta) = 1, \alpha \equiv \beta \pmod{\mathfrak{f}}\}$$

[1]En principio podría contener factores primos arquimedianos, pero no es el caso porque K no tiene divisores primos reales.

está contenido en el núcleo del homomorfismo de Artin. La clase de $\mathfrak{a}$ en $I(\Delta)/P_{\mathfrak{f}}$ contiene infinitos primos de grado 1. Podemos tomar uno $\mathfrak{p}$ para el que se cumpla el teorema anterior. Entonces $\gamma = \alpha/\beta$ tal que $(\gamma) \in P_{\mathfrak{f}}$ y $\mathfrak{a} = (\gamma)\mathfrak{p}$. Se cumple que

$$F\left(\frac{H/K}{\mathfrak{a}}\right) = F\left(\frac{H/K}{(\gamma)\mathfrak{p}}\right) = F\left(\frac{H/K}{\mathfrak{p}}\right) = [\mathfrak{p}] = [\mathfrak{a}].$$

Llamemos $N \leq I(\Delta)$ al núcleo del homomorfismo de Artin, es decir, al grupo de clases de H sobre K. Si $\mathfrak{p}$ es un ideal primo de K tal que $\mathfrak{p} \nmid \Delta$, tenemos que

$$\mathfrak{p} \in N \Leftrightarrow \left(\frac{H/K}{\mathfrak{p}}\right) = 1 \Leftrightarrow F\left(\frac{H/K}{\mathfrak{p}}\right) = [\mathfrak{p}] = 1 \Leftrightarrow \mathfrak{p} \text{ es principal.}$$

Así pues, salvo a lo sumo para un número finito de primos (los divisores de Δ) tenemos que $\mathfrak{p} \in N$ si y sólo si $\mathfrak{p}$ está en el grupo de los ideales fraccionales principales de K, que es precisamente el grupo de clases del cuerpo de clases de Hilbert de K. Si dos grupos de ideales contienen casi los mismos primos, entonces determinan el mismo cuerpo de clases, luego podemos concluir que H es el cuerpo de clases de Hilbert de K. El teorema siguiente recoge este hecho y otras consecuencias sencillas:

Teorema 13.9 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Llamemos h_K al número de clases de K. Entonces*

a) $K(j(E))$ es el cuerpo de clases de Hilbert de K.

b) $|\mathbb{Q}(j(E)) : \mathbb{Q}| = |K(j(E)) : K| = h_K$.

c) Para todo ideal $\mathfrak{a}$ de $\mathcal{O}_K$ se cumple

$$\left(\frac{K(j(E))/K}{\mathfrak{a}}\right)(j(E)) = j(\mathfrak{a} * E).$$

d) Si $E_1, \ldots, E_h$ son un sistema completo de representantes de las clases de isomorfía de curvas elípticas con anillo de multiplicaciones complejas $\mathcal{O}_K$, entonces $j(E_1), \ldots, j(E_h)$ son los conjugados de $j(E)$ sobre K (y también sobre $\mathbb{Q}$).

Demostración: Ya tenemos probado a). La segunda igualdad en b) es consecuencia inmediata de a). Respecto a la primera, es obvio que

$$h_K = |K(j(E)) : K| \leq |\mathbb{Q}(j(E)) : \mathbb{Q}|,$$

y la desigualdad contraria la conocíamos ya (ver la nota tras el teorema 10.24).

Como consecuencia de a), ahora sabemos que el discriminante de la extensión $K(j(E))/K$ es $\Delta = 1$, luego el isomorfismo de Artin está definido sobre todo el grupo de clases $\mathcal{H}_K$, y hemos probado que es el inverso de F. La propiedad c) es equivalente a la propiedad que define a F. Por último toda E_i es de la forma $\mathfrak{a} * E$, para cierto ideal $\mathfrak{a}$, luego d) es consecuencia de c). ∎

Ejemplo Vamos a calcular el cuerpo de clases de Hilbert de $K = \mathbb{Q}(\sqrt{-5}\,)$. El número de clases es $h = 2$ y un ideal no principal es

$$\mathfrak{p} = (2, 1 + \sqrt{-5}) = \langle 2, 1 + \sqrt{-5}\rangle_{\mathbb{Z}}\,.$$

El teorema anterior nos da que el cuerpo de clases H está generado sobre K por el par de números conjugados

$$j(\sqrt{-5}\,) \approx 1264538.9094751405\ldots$$

$$j\left(\frac{1+\sqrt{-5}}{2}\right) \approx -538.9094751405\ldots$$

Las aproximaciones se calculan analíticamente. Calculando la suma y el producto obtenemos el polinomio mínimo de $j(\sqrt{-5}\,)$, que resulta ser

$$X^2 - 1264000X - 681472000$$

Calculando sus raíces obtenemos el valor exacto

$$j(\sqrt{-5}\,) = 632000 + 282880\sqrt{5}.$$

Ahora es inmediato que $H = K(\sqrt{5}\,) = \mathbb{Q}(\sqrt{-5}, \sqrt{5}\,)$. ∎

El procedimiento del ejemplo anterior se puede usar para calcular fácilmente el cuerpo de clases de Hilbert de cualquier cuerpo cuadrático imaginario a partir de un conjunto de representantes del grupo de clases.

13.3 La máxima extensión abeliana

En esta sección veremos cómo obtener la máxima extensión abeliana de un cuerpo cuadrático imaginario K a partir de una curva elíptica $E/\mathbb{C}$ con anillo de multiplicaciones complejas $\mathcal{O}_K$. Dicha extensión está relacionada con los puntos de torsión de E, tal y como muestra el teorema siguiente:

Teorema 13.10 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático K. Consideremos una ecuación de Weierstrass para E con coeficientes en $H = K(j(E))$. Entonces, el cuerpo $L = K(j(E), E_{\mathrm{tor}})$ que resulta de adjuntar a K el invariante $j(E)$ y las coordenadas (afines) de los puntos de torsión (no nulos) de E es una extensión abeliana de H.*

Demostración: Para cada $m > 1$, sea $L_m = K(j(E), E[m]) = H(E[m])$ la extensión de H generada por las coordenadas de los puntos (no nulos) de $E[m]$. Consideremos el monomorfismo de grupos

$$\rho : G(L_m/H) \longrightarrow \mathrm{Aut}(E[m])$$

dada por $\rho(\sigma)(T) = T^\sigma$. En principio, $\rho(\sigma)$ es un automorfismo de $E[m]$ visto como grupo abeliano, pero además se cumple que

$$\rho(\sigma)(\alpha T) = (\alpha T)^\sigma = \alpha^\sigma T^\sigma = \alpha T^\sigma = \alpha\rho(\sigma)(T),$$

luego $\rho(\sigma)$ es un automorfismo de $\mathcal{O}_K/(m)$-módulos.

Fijemos un punto $P \in E[m]$. Según el teorema 13.1 (ver las observaciones precedentes), la aplicación $[\alpha] \mapsto \alpha P$ determina un isomorfismo $\mathcal{O}_K/(m) \cong E[m]$ (claramente un isomorfismo de $\mathcal{O}_K/(m)$-módulos), luego $E[m]$ es un $\mathcal{O}_K/(m)$-módulo libre de rango 1 y su grupo de automorfismos es isomorfo al grupo abeliano $E[m]$. Así pues, ρ inyecta $G(L_m/H)$ en un subgrupo de $E[m]$, lo que prueba que la extensión L_m/H es abeliana, para todo m y, por consiguiente, L también lo es. ∎

En las condiciones del teorema anterior, no es cierto en general que el cuerpo L sea una extensión abeliana de K. Para conseguir extensiones abelianas de K tendremos que "manipular" levemente las coordenadas de los puntos de torsión. Dicha manipulación es muy simple, pero antes de describirla vamos a demostrar algunos resultados previos que necesitaremos para justificar que, ciertamente, nos lleva a la máxima extensión abeliana de K.

Teorema 13.11 *Sea E/L una curva elíptica definida sobre un cuerpo numérico L y cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de un cuerpo cuadrático imaginario K. Sea $\mathfrak{P}$ un ideal primo de L en el que E tenga buena reducción. Si una isogenia $\gamma \in \mathrm{End}(\tilde{E})$ conmuta con todas las isogenias de la imagen de la reducción $\mathrm{End}(E) \longrightarrow \mathrm{End}(\tilde{E})$, entonces γ está en dicha imagen.*

Demostración: Según el teorema 3.35, el anillo $\mathrm{End}(\tilde{E})$ es un orden cuadrático o bien un orden en un álgebra de cuaternios. Si es un orden cuadrático no hay nada que probar porque, como $\mathcal{O}_K$ es maximal, la reducción de isogenias ha de ser un isomorfismo. Supongamos, pues, que es un orden en un álgebra de cuaternios A. Si identificamos a $\mathcal{O}_K$ con la imagen de la reducción de isogenias, entonces tenemos que $K \subset A$. Pongamos que $K = \mathbb{Q}(\sqrt{d}\,)$ y llamemos $\alpha = \sqrt{d}$. De la demostración del teorema 3.34 se sigue que podemos encontrar $\beta \in A$ de modo que

$$A = \langle 1, \alpha, \beta, \alpha\beta\rangle_{\mathbb{Q}}\,, \quad \alpha^2,\ \beta^2,\ (\alpha\beta)^2 \in \mathbb{Q}, \quad \alpha\beta = -\beta\alpha.$$

La isogenia del enunciado será de la forma $\gamma = a + b\alpha + c\beta + d\alpha\beta$ y ha de cumplir $\alpha\gamma = \gamma\alpha$. Operando se concluye fácilmente que esto sólo es posible si $c = d = 0$, con lo que $\gamma \in K$. Además, todo endomorfismo es raíz de un polinomio mónico con coeficientes enteros, luego $\gamma \in \mathcal{O}_K$. ∎

Como aplicación demostramos el teorema siguiente:

Teorema 13.12 *Sea K un cuerpo cuadrático imaginario, sea H su cuerpo de clases de Hilbert y sea E/H una curva elíptica cuyo anillo de multiplicaciones*

complejas sea el orden maximal $\mathcal{O}_K$. Para todo primo $\mathfrak{p}$ de grado 1 en K salvo a lo sumo un número finito de ellos, si llamamos

$$\sigma = \left(\frac{H/K}{\mathfrak{p}}\right)$$

existe una isogenia $\lambda : E \longrightarrow E^\sigma$ de grado $p = \mathrm{N}(\mathfrak{p})$ definida sobre una extensión L de H cuya reducción módulo cualquier divisor $\mathfrak{P}$ de $\mathfrak{p}$ en L es la aplicación de Frobenius de grado p.

DEMOSTRACIÓN: Sea S el conjunto finito de primos de K que hemos considerado en la prueba del teorema 13.8. Allí hemos visto que si $\mathfrak{p} \notin S$ la isogenia $\mathfrak{p} : E \longrightarrow E*\mathfrak{p}$ tiene grado p y su reducción módulo $\mathfrak{P}$ es puramente inseparable de grado p.

Por otra parte, sabemos que $E * \mathfrak{p} \cong E^\sigma$ y, de hecho, podemos suponer la igualdad (pues $E * \mathfrak{p}$ sólo está definido salvo isomorfismo). La reducción se descompone como

$$\tilde{E} \xrightarrow{\phi} \tilde{E}^{(p)} \xrightarrow{\epsilon} \widetilde{E^\sigma},$$

donde ϕ es la aplicación de Frobenius y ϵ es un isomorfismo. Ahora bien, σ módulo $\mathfrak{P}$ es el automorfismo de Frobenius, luego $\widetilde{E^\sigma} = \tilde{E}^{(p)}$ y ϵ es un automorfismo de $\widetilde{E^\sigma}$. Basta probar que $\epsilon = \tilde{\epsilon}_0$, para cierto automorfismo ϵ_0 de E^σ, pues entonces la isogenia buscada será $\lambda = \mathfrak{p} \circ \epsilon_0^{-1}$. Notemos que si $\epsilon = \tilde{\epsilon}_0$, automáticamente ϵ_0 es un automorfismo, pues la reducción conserva el grado (teorema 13.5). Por el teorema anterior basta ver que ϵ conmuta con todas las reducciones.

Observemos, en general, que si $f : V \longrightarrow W$ es una aplicación racional entre dos curvas definidas sobre un cuerpo de característica prima p y ϕ_V, ϕ_W son las respectivas aplicaciones de Frobenius de grado p, entonces $f^\sigma : V^{(p)} \longrightarrow W^{(p)}$, donde σ es el automorfismo de Frobenius y se cumple la relación $f \circ \phi_W = \phi_V \circ f^\sigma$. La comprobación es inmediata.

Un endomorfismo arbitrario de E viene determinado por un $\alpha \in \mathcal{O}_K$. Si aplicamos la observación anterior a $\tilde{\alpha}$, vemos que $\tilde{\alpha} \circ \phi = \phi \circ \widetilde{\alpha^\sigma} = \phi \circ \tilde{\alpha}$. Notemos que la α de los dos primeros términos es un endomorfismo de E, mientras que la última es un endomorfismo de E^σ. Ahora calculamos:

$$\phi \circ \epsilon \circ \tilde{\alpha} = \tilde{\mathfrak{p}} \circ \tilde{\alpha} = \widetilde{\mathfrak{p} \circ \alpha} = \widetilde{\alpha \circ \mathfrak{p}} = \tilde{\alpha} \circ \phi \circ \epsilon = \phi \circ \tilde{\alpha} \circ \epsilon.$$

Aquí hemos usado el teorema 10.22. Como ϕ es biyectiva esto implica que $\epsilon \circ \tilde{\alpha} = \tilde{\alpha} \circ \epsilon$, como teníamos que probar. ∎

Si el ideal $\mathfrak{p}$ es principal podemos decir más:

Teorema 13.13 *Sea K un cuerpo cuadrático imaginario, sea H su cuerpo de clases de Hilbert y E/H una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$. Para cada ideal principal primo $\mathfrak{p}$ de grado 1 en K, salvo a lo sumo un número finito de ellos, existe un único $\pi \in \mathcal{O}_K$ tal que $\mathfrak{p} = (\pi)$ y la multiplicación por π se reduce a la aplicación de Frobenius de grado $p = \mathrm{N}(\mathfrak{p})$.*

DEMOSTRACIÓN: Como ahora $\mathfrak{p}$ es principal, en el teorema anterior tenemos $\sigma = 1$, por lo que λ es la multiplicación por un $\pi \in \mathcal{O}_K$ está definido sobre H (teorema 10.25). Es claro entonces que podemos tomar $L = H$.

La norma de π es el grado de la multiplicación por π, que es p, luego π genera el ideal $\mathfrak{p}$ o bien su conjugado. Ahora bien, en la prueba del teorema 13.8 hemos visto (para una multiplicación compleja arbitraria α) que si $\tilde{\omega}$ es una diferencial invariante en $\tilde{E}$, entonces $\bar{\tilde{\pi}}(\tilde{\omega}) = \tilde{\pi}\tilde{\omega}$, y esta forma diferencial ha de ser nula porque $\tilde{\pi}$ es la aplicación de Frobenius, luego es inseparable. Así pues, $\tilde{\pi} = 0$ y, por consiguiente, $\pi \in \mathfrak{P} \cap K = \mathfrak{p}$, luego $\mathfrak{p} = (\pi)$.

La unicidad se debe a que la reducción de isogenias es inyectiva. ∎

Ahora ya podemos volver al problema de describir la máxima extensión abeliana de un cuerpo cuadrático imaginario K. Dada una curva elíptica $E/\mathbb{C}$, una *función de Weber* para E es cualquier función regular $h \in \mathbb{C}(E)$ con la propiedad de que para cada par de puntos $P, Q \in E$ se cumpla que $h(P) = h(Q)$ si y sólo si existe un automorfismo α de E tal que $\alpha(P) = Q$.

Es fácil ver que toda curva $E/\mathbb{C}$ tiene una función de Weber. Si

$$Y^2 = 4X^3 - g_2X - g_3$$

es una ecuación de Weierstrass de $E/\mathbb{C}$, entonces una función de Weber viene dada por

$$h = \begin{cases} x & \text{si } g_2g_3 \neq 0, \\ x^2 & \text{si } \quad g_3 = 0, \\ x^3 & \text{si } \quad g_2 = 0. \end{cases}$$

En efecto, si $g_2g_3 \neq 0$ (o, equivalentemente, si $j \neq 0$, 1728) entonces los únicos automorfismos de E son ± 1, donde $-(X, Y) = (X, -Y)$ luego $h = x$ es una función de Weber.

Si $g_3 = 0$ (o, equivalentemente, $j = 1728$) el anillo de multiplicaciones complejas es $\mathbb{Z}[i]$, los automorfismos son ± 1, $\pm i$, donde $i(X, Y) = (-X, iY)$, y es fácil ver que podemos tomar como función de Weber $h = x^2$.

Si $g_2 = 0$ (o, equivalentemente, $j = 0$) el anillo de multiplicaciones complejas es $\mathbb{Z}[\rho]$, donde $\rho^3 = 1$, el grupo de automorfismos está generado por $-\rho$ y $-\rho(X, Y) = (\rho X, -Y)$, con lo que es fácil comprobar que una función de Weber es $h = x^3$.

Aunque éstas son las funciones de Weber más simples, vamos a complicarlas ligeramente a cambio de hacerlas independientes de elecciones arbitrarias:

Definición 13.14 Sea $E/\mathbb{C}$ una curva elíptica compleja dada por una ecuación de Weierstrass $Y^2 = X^3 - g_2X - g_3$. Definimos la *función de Weber* de E como la función $h \in \mathbb{C}(E)$ dada por

$$h = \begin{cases} \dfrac{g_2g_3}{\Delta}x & \text{si } j \neq 0, 1728, \\ \dfrac{g_2^2}{\Delta}x^2 & \text{si } j = 1728, \\ \dfrac{g_3}{\Delta}x^3 & \text{si } j = 0. \end{cases}$$

Esta función h no depende de la ecuación de Weierstrass considerada, pues las coordenadas relativas a dos ecuaciones de Weierstrass están relacionadas en la forma $x = u^2x'$, $y = u^3y'$, para cierto $u \in \mathbb{C}^*$, y los coeficientes se transforman según las relaciones

$$g_2 = u^4 g_2', \quad g_3 = u^6 g_3', \quad \Delta = u^{12}\Delta',$$

de donde se sigue inmediatamente que ambas ecuaciones dan lugar a la misma función de Weber.

Notemos que si E está definida sobre un cuerpo L, entonces $h \in L(E)$.

Recordemos que si $\mathfrak{m}$ es un ideal de un cuerpo cuadrático imaginario K, entonces el *cuerpo de clases radial* de K módulo $\mathfrak{m}$ es el cuerpo de clases que tiene por grupo de clases al grupo

$$P_{\mathfrak{m}} = \{(\alpha)/(\beta) \in I(\mathfrak{m}) \mid \alpha,\, \beta \in \mathcal{O}_K, (\alpha\beta, \mathfrak{m}) = 1,\ \alpha \equiv \beta \ (\text{mód } \mathfrak{m})\}.$$

Teorema 13.15 *Sea K un cuerpo cuadrático imaginario, sea H su cuerpo de clases de Hilbert, sea E/H una curva elíptica cuyo anillo de multiplicaciones complejas sea $\mathcal{O}_K$ y sea $h \in H(E)$ la función de Weber de E. Para cada ideal $\mathfrak{m}$ de $\mathcal{O}_K$, el cuerpo de clases radial de K módulo $\mathfrak{m}$ es $K(j(E), h[E[\mathfrak{m}]])$.*

Demostración: Llamemos $L = K(j(E), h[E[\mathfrak{m}]])$. Aquí hay que entender que $h(O) = \infty$ no se considera a la hora de calcular L. Sea $H = K(j(E))$ el cuerpo de clases de Hilbert de K. Es evidente que L está contenido en el cuerpo descrito en el teorema 13.10, luego la extensión L/H es abeliana y la extensión L/K es normal. Llamemos A al cuerpo de clases radial de K módulo $\mathfrak{m}$, de modo que $K \subset H \subset A$ y la extensión A/K también es abeliana.

Fijemos una ecuación de Weierstrass para E con coeficientes en H. Sea L' una extensión normal de K que contenga a L y a las coordenadas de los puntos de $E[\mathfrak{m}]$.

Vamos a demostrar que, con a lo sumo un número finito de excepciones, un primo $\mathfrak{p}$ de grado 1 en K se escinde completamente en L si y sólo si se escinde completamente en A.

Supongamos primeramente que $\mathfrak{p}$ se escinde completamente en A, lo que significa que $\mathfrak{p} \in P_{\mathfrak{m}}$, luego $\mathfrak{p} = (\alpha)$, para un cierto $\alpha \in \mathcal{O}_K$, $\alpha \equiv 1$ (mód $\mathfrak{m}$). Consideremos un divisor $\mathfrak{P}'$ de $\mathfrak{p}$ en L' y el símbolo de Frobenius

$$\sigma = \left(\frac{L'/K}{\mathfrak{P}'}\right).$$

(Esto supone que $\mathfrak{p}$ no se ramifica en L', lo que descarta sólo un número finito de primos $\mathfrak{p}$.) Se cumple que

$$\sigma|_L = \left(\frac{L/K}{\mathfrak{P}}\right),$$

donde $\mathfrak{P}$ es el divisor de $\mathfrak{P}'$ en L. El orden de $\sigma|_L$ es el grado de inercia $f(\mathfrak{P}/\mathfrak{p})$, luego si probamos que $\sigma = 1$ tendremos que $\mathfrak{p}$ se escinde completamente en L.

La restricción $\sigma|_H$ es el símbolo de Artin de $\mathfrak{p}$ para la extensión L/K. Como $\mathfrak{p}$ es principal, $\sigma|_H = 1$, luego σ fija a $j(E)$, y sólo hemos de probar que σ fija también a $h[E[\mathfrak{m}]]$. Fijemos, pues, $T \in E[\mathfrak{m}]$.

Eliminando un número finito de posibilidades para $\mathfrak{p}$, podemos suponer que cumple el teorema 13.13, lo que nos da un generador $\mathfrak{p} = (\pi)$ tal que la multiplicación por π se reduce módulo $\mathfrak{P}$ al automorfismo de Frobenius. En particular $\tilde{\pi}$ coincide con $\tilde{\sigma}$ sobre $\tilde{E}(L'_{\mathfrak{P}'})$. Por otra parte, como $(\pi) = (\alpha)$, existe una unidad $\epsilon \in \mathcal{O}_K$ tal que $\pi = \epsilon\alpha$.

Ahora calculamos: $\widetilde{T^\sigma} = \tilde{T}^{\tilde{\sigma}} = \tilde{\pi}(\tilde{T}) = \widetilde{\pi T}$. Si excluimos los primos $\mathfrak{p}$ cuya norma divide a $|E[\mathfrak{m}]|$, tenemos que la reducción módulo $\mathfrak{P}'$ es inyectiva sobre $E[\mathfrak{m}]$, luego podemos concluir que $T^\sigma = \pi T$. Observemos que $h^\sigma = h$, porque $h \in H(E)$ y $\sigma|_H = 1$. Así pues:

$$h(T)^\sigma = h(T^\sigma) = h(\pi T) = h(\epsilon\alpha T) = h(\alpha T) = h(T).$$

En el último paso hemos usado que $\alpha \equiv 1 \pmod{\mathfrak{m}}$ y $T \in E[\mathfrak{m}]$. Esto prueba que $\sigma = 1$.

Recíprocamente, supongamos que $\mathfrak{p}$ se escinde completamente en L. Entonces se escinde completamente en H, luego $\mathfrak{p}$ es principal. Por el teorema 13.13 tenemos que $\mathfrak{p} = (\pi)$, de modo que la multiplicación por π se reduce (módulo cualquier divisor $\mathfrak{P}'$ de $\mathfrak{p}$ en L') a la aplicación de Frobenius. Ahora sabemos que el símbolo de Frobenius $\sigma|_L$ es la identidad y, como antes, tenemos que $\tilde{\pi}$ coincide con $\tilde{\sigma}$ en $\tilde{E}(L'_{\mathfrak{P}'})$.

Llamemos $\tilde{h}$ a la aplicación definida según 13.14 para la curva reducida $\tilde{E}$. Es evidente que $\widetilde{h(P)} = \tilde{h}(\tilde{P})$, para todo punto $P \in E$. Si $T \in E[\mathfrak{m}]$, tenemos que

$$\tilde{h}(\tilde{\pi}\tilde{T}) = \tilde{h}(\widetilde{\pi T}) = \tilde{h}(\widetilde{T^\sigma}) = \widetilde{h(T^\sigma)} = \widetilde{h(T)^\sigma} = \widetilde{h(T)} = \tilde{h}(\tilde{T}).$$

De la definición de h y $\tilde{h}$ se sigue que existe una unidad $\epsilon \in \mathcal{O}_K$ tal que $\tilde{\pi}\tilde{T} = \tilde{\epsilon}\tilde{T}$. En efecto, consideremos, por ejemplo, el caso $j = 0$. Entonces tenemos que $\tilde{x}^3(\tilde{\pi}\tilde{T}) = \tilde{x}^3(\tilde{T})$. De la ecuación de Weierstrass se sigue que a lo sumo hay seis puntos en $(\tilde{x}, \tilde{y}) \in \tilde{E}$ tales que $\tilde{x}^3 = \tilde{x}^3(\tilde{T})$, y estos puntos han de ser los puntos $\tilde{\epsilon}\tilde{T}$, cuando ϵ recorre las seis unidades de $\mathcal{O}_K$ (que son distintos dos a dos, porque la reducción es inyectiva en $E[\mathfrak{m}]$). Por consiguiente $\tilde{\pi}\tilde{T}$ es uno de estos seis puntos.

Concluimos que $\widetilde{(\pi - \epsilon)T} = \tilde{O}$, luego $(\pi - \epsilon)T = O$, luego $\pi T = \epsilon T$, luego $\pi\epsilon^{-1}T = T$, luego $(\pi\epsilon^{-1} - 1)T = O$.

Por el teorema 13.1, tenemos que la aplicación $\alpha \mapsto \alpha T$ induce un isomorfismo $\mathcal{O}_K/(\mathfrak{m}) \cong E[\mathfrak{m}]$, luego $\pi\epsilon^{-1} \equiv 1 \pmod{\mathfrak{m}}$ y además $\mathfrak{p} = (\pi) = (\pi\epsilon^{-1})$, luego $\mathfrak{p} \in P_{\mathfrak{m}}$. Por consiguiente $\mathfrak{p}$ se escinde completamente en A.

Los cuerpos L y A son extensiones normales de K, y la teoría de cuerpos de clases garantiza[2] que si los primos de grado 1 de K que se escinden completamente en ambas son casi los mismos, necesariamente $L = A$, como queríamos probar. ∎

Como consecuencia inmediata:

Teorema 13.16 *Sea K un cuerpo cuadrático imaginario, sea H su cuerpo de clases de Hilbert, sea E/H una curva elíptica cuyo anillo de multiplicaciones complejas sea $\mathcal{O}_K$ y sea $h \in H(E)$ la función de Weber de E. Entonces, la máxima extensión abeliana de K es el cuerpo $K(j(E), h[E_{\mathrm{tor}}])$.*

En particular, si $j \neq 0$, 1728, la máxima extensión abeliana se obtiene adjuntando el invariante $j(E)$ y las coordenadas x de todos los puntos de torsión de E.

Ejemplo Consideremos la curva $E/\mathbb{Q}$ dada por $Y^2 = X^3 + X$, cuyo anillo de multiplicaciones complejas es $\mathbb{Z}[i]$. Su invariante es $j = 1728$ y su función de Weber es $h = -x^2/64$ (aunque, para nuestros fines, $h = x^2$ sirve igual).

Los puntos finitos de orden 2 son $(0,0)$, $(i,0)$, $(-i,0)$, luego el cuerpo de clases radiales de $\mathbb{Q}(i)$ módulo 2 es

$$K_2 = \mathbb{Q}(i)(1728, h(0), h(i), h(-i)) = \mathbb{Q}(i).$$

Si $T = (x,y) \in E$, la fórmula de duplicación es

$$2T = \left(\frac{x^4 - 2x^2 + 1}{4y^2}, \frac{x^6 + 5x^4 - 5x^2 - 1}{8y^3}\right).$$

Si $T \in E[3]$ entonces $2T = T$, y al igualar las coordenadas x y sustituir la ecuación de la curva obtenemos que

$$T \in E[3] \Rightarrow 3x^4 + 6x^2 - 1 = 0.$$

Esta ecuación tiene a lo sumo cuatro raíces, y para cada valor de x hay a lo sumo dos valores de y correspondientes a puntos de E, luego a lo sumo hay ocho puntos de E que cumplen la ecuación. Como tiene que haber exactamente ocho (pues $E[3]$ tiene nueve puntos contando O), concluimos que

$$T \in E[3] \Leftrightarrow 3x^4 + 6x^2 - 1 = 0.$$

Las raíces de la ecuación son

$$\alpha = \pm\sqrt{\frac{-3 \pm 2\sqrt{3}}{3}}.$$

[2]Ver el teorema 9.21 de mi *Teoría de cuerpos de clases*. Nótese que en él "casi" significa salvo un conjunto con densidad de Dirichlet nula.

Por lo tanto, el cuerpo de clases radiales de $\mathbb{Q}(i)$ módulo 3 es $K_3 = \mathbb{Q}(i, \sqrt{3}\,)$.

Un punto $T = (x, y) \in E$ tiene (exactamente) orden 4 si y sólo si $y(2T) = 0$, es decir, si y sólo si

$$x^6 + 5x^4 - 5x^2 - 1 = 0.$$

(Los restantes puntos de $E[4]$ son los de $E[2]$, y ya hemos visto que no dan lugar a extensiones de $\mathbb{Q}(i)$.)

La ecuación factoriza como $(x+1)(x-1)(x^4+6x^2+1) = 0$, y las raíces del último factor son $\alpha = \pm\sqrt{-3 \pm 2\sqrt{2}}$. Por lo tanto, el cuerpo de clases radiales módulo 4 es $K_4 = \mathbb{Q}(i, \sqrt{2}\,)$. ∎

13.4 El teorema fundamental

En esta sección demostraremos el teorema principal de la multiplicación compleja, un teorema del que se siguen fácilmente los resultados que hemos obtenido hasta ahora y que, debidamente generalizado, nos permitirá obtener resultados análogos para curvas cuyo anillo de multiplicaciones complejas sea un orden no maximal. Necesitamos algunos resultados que reunimos en un teorema previo.

Sea K un cuerpo numérico y $\mathcal{O}_K$ su anillo de enteros algebraicos. Para cada ideal primo $\mathfrak{p}$ de K representaremos por $K_\mathfrak{p}$ la compleción de K respecto de $\mathfrak{p}$ y $\mathcal{O}_\mathfrak{p}$ será su anillo de enteros. Si $\mathfrak{a}$ es un ideal fraccional de K, representaremos por $\mathfrak{a}_\mathfrak{p}$ la clausura de $\mathfrak{a}$ en $K_\mathfrak{p}$, que no es sino el ideal fraccional que resulta de eliminar de $\mathfrak{a}$ todos los divisores distintos de $\mathfrak{p}$.

Si M es un $\mathcal{O}_K$-módulo, definimos su *parte $\mathfrak{p}$-primaria* como el $\mathcal{O}_K$-módulo

$$M[\mathfrak{p}^\infty] = \{m \in M \mid \mathfrak{p}^e m = (0) \text{ para un } e \geq 0\}.$$

Teorema 13.17 *Con la notación precedente:*

a) Sea M un $\mathcal{O}_K$-módulo de torsión. Entonces la suma natural

$$S : \bigoplus_\mathfrak{p} M[\mathfrak{p}^\infty] \longrightarrow M$$

es un isomorfismo de $\mathcal{O}_K$-módulos.

b) Si $\mathfrak{a}$ es un ideal fraccional de K, para cada primo $\mathfrak{p}$ de K, la inclusión $K \longrightarrow K_\mathfrak{p}$ induce un isomorfismo $(K/\mathfrak{a})[\mathfrak{p}^\infty] \cong K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$.

c) Si $\mathfrak{a}$ es un ideal fraccional de K, entonces $K/\mathfrak{a} \cong \bigoplus_\mathfrak{p} K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$.

Demostración: a) Supongamos que $S(\mu) = 0$. Para cada $\mathfrak{p}$, sea $e(\mathfrak{p}) \geq 0$ el menor natural tal que $\mathfrak{p}^{e(\mathfrak{p})}\mu_\mathfrak{p} = (0)$. Notemos que casi todo $\mu_\mathfrak{p} = 0$, luego casi todo $e(\mathfrak{p}) = 0$. Esto nos permite definir (fijado un ideal primo $\mathfrak{q}$) el ideal

$$\mathfrak{d} = \prod_{\mathfrak{p}\neq\mathfrak{q}} \mathfrak{p}^{e(\mathfrak{p})}.$$

Claramente $\mathfrak{d}\mu_\mathfrak{p} = (0)$ para todo $\mathfrak{q} \neq \mathfrak{p}$. Por otro lado,

$$(0) = \mathfrak{d}S(\mu) = \mathfrak{d}\sum_\mathfrak{p}\mu_\mathfrak{p} = \sum_\mathfrak{p}\mathfrak{d}\mu_\mathfrak{p} = \mathfrak{d}\mu_\mathfrak{q}.$$

Pero $\mathfrak{d}$ es primo con $\mathfrak{q}$, luego $\mathfrak{d} + \mathfrak{q}^{e(\mathfrak{q})} = 1$. Así pues,

$$(\mu_\mathfrak{q}) = (\mathfrak{d} + \mathfrak{q}^{e(q)})\mu_\mathfrak{q} = \mathfrak{d}\mu_\mathfrak{q} + \mathfrak{q}^{e(\mathfrak{q})}\mu_\mathfrak{q} = (0).$$

Esto prueba que $\mu_\mathfrak{q} = 0$, para todo $\mathfrak{q}$, luego $\mu = 0$ y S es inyectiva.

Para probar la suprayectividad, tomamos $m \in M$. Por hipótesis existe un $\alpha \in \mathcal{O}_K$ tal que $\alpha m = 0$. Pongamos que $(\alpha) = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$. Los ideales $\alpha\mathfrak{p}_i^{-e_i}$ son primos entre sí, luego

$$\alpha\mathfrak{p}_1^{-e_1} + \cdots + \alpha\mathfrak{p}_r^{-e_r} = (1).$$

Podemos tomar $\epsilon_i \in \alpha\mathfrak{p}_i^{-e_i}$ de modo que $\epsilon_1 + \cdots + \epsilon_r = 1$. Así, si $i \neq j$ tenemos que $\epsilon_i \equiv 0 \pmod{\mathfrak{p}_j^{e_j}}$, y la igualdad implica entonces que $\epsilon_i \equiv 1 \pmod{\mathfrak{p}_i^{e_i}}$. Como $\mathfrak{p}_i^{e_i}\epsilon_i \subset (\alpha)$, tenemos que $\mathfrak{p}_i^{e_i}\epsilon_i m = (0)$, luego $\epsilon_i m \in M[\mathfrak{p}_i^\infty]$.

Los elementos $\epsilon_i m$ (completados con ceros) definen un elemento μ del dominio de S que claramente cumple $S(\mu) = m$.

b) Tomemos $[\alpha] \in (K/\mathfrak{a})[\mathfrak{p}^\infty]$ y supongamos que está en el núcleo del homomorfismo inducido por la inclusión, es decir, que $\alpha \in \mathfrak{a}_\mathfrak{p}$. Por otra parte, tenemos que existe un $e \geq 0$ tal que $\mathfrak{p}^e\alpha \subset \mathfrak{a}$. Estos dos hechos implican respectivamente que $v_\mathfrak{p}(\alpha) \geq v_\mathfrak{p}(\mathfrak{a})$ y $v_\mathfrak{q}(\alpha) \geq v_\mathfrak{q}(\mathfrak{a})$ para todo $\mathfrak{q} \neq \mathfrak{p}$. Así pues, $\alpha \in \mathfrak{a}$ y $[\alpha] = 0$. Esto demuestra la inyectividad.

Tomemos ahora $[\beta] \in K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$. El teorema de aproximación nos da un $\alpha \in K$ tal que $v_\mathfrak{p}(\alpha - \beta) \geq v_\mathfrak{p}(\mathfrak{a})$ y $v_\mathfrak{q}(\alpha) \geq v_\mathfrak{q}(\mathfrak{a})$ para todo[3] $\mathfrak{q} \neq \mathfrak{p}$. La primera desigualdad implica que $\alpha - \beta \in \mathfrak{a}_\mathfrak{p}$, luego $[\beta] = [\alpha]$. Por otra parte, si tomamos $e \geq v_\mathfrak{p}(\mathfrak{a}) - v_\mathfrak{p}(\alpha)$, entonces $v_\mathfrak{q}(\mathfrak{p}^e\alpha) \geq v_\mathfrak{q}(\mathfrak{a})$ para todos los primos $\mathfrak{q}$, luego $\mathfrak{p}^e\alpha \subset \mathfrak{a}$ y así $[\alpha] \in (K/\mathfrak{a})[\mathfrak{p}^\infty]$ es una antiimagen de $[\beta]$.

c) es una consecuencia inmediata de a) y b). ∎

Representaremos por J_K al grupo de los elementos ideales de K. Recordemos que un elemento ideal $s \in J_K$ está determinado por sus componentes $s_\mathfrak{p} \in K_\mathfrak{p}^*$, donde $\mathfrak{p}$ recorre los divisores primos (arquimedianos o no) de K y $|s_\mathfrak{p}|_\mathfrak{p} = 1$ para todos los divisores $\mathfrak{p}$ salvo a lo sumo un número finito de ellos.

Para cada elemento ideal $s \in J_K$ tenemos definido el ideal fraccional

$$(s) = \prod_\mathfrak{p}\mathfrak{p}^{v_\mathfrak{p}(s_\mathfrak{p})},$$

[3]En principio el teorema de aproximación nos permite imponer condiciones sobre un número finito de primos (por ejemplo, $\mathfrak{p}$ y los divisores de $\mathfrak{a}$), pero si $\alpha \in K$ cumple lo pedido para estos primos, el teorema chino del resto nos permite tomar un $u \in \mathcal{O}_K$ tal que $|u-1|_\mathfrak{q} < \epsilon$ para estos primos y $|u|_\mathfrak{q} < \epsilon$ para los restantes primos que cumplan $|\alpha|_\mathfrak{q} > 1$ (donde ϵ es un número real suficientemente pequeño). Entonces αu cumple las mismas condiciones de aproximación que α y además $|\alpha u|_\mathfrak{q} \leq 1$ para todos los otros primos. Por ejemplo:

$$|\alpha u - \beta|_\mathfrak{p} \leq \text{máx}\{|\alpha - \beta|_\mathfrak{p}|u|_\mathfrak{p}, |\beta|_\mathfrak{p}|u - 1|_\mathfrak{p}\}.$$

donde $\mathfrak{p}$ recorre ahora los primos no arquimedianos de K. Si $\mathfrak{a}$ es un ideal fraccional de K, definimos $s\mathfrak{a} = (s)\mathfrak{a}$. Es claro que $(s\mathfrak{a})_\mathfrak{p} = s_\mathfrak{p}\mathfrak{a}_\mathfrak{p}$. El teorema anterior nos da los isomorfismos

$$K/\mathfrak{a} \cong \bigoplus_\mathfrak{p} K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p}, \qquad K/s\mathfrak{a} \cong \bigoplus_\mathfrak{p} K_\mathfrak{p}/s_\mathfrak{p}\mathfrak{a}_\mathfrak{p}.$$

Definimos la multiplicación por s como el isomorfismo de $\mathcal{O}_K$-módulos determinado por el siguiente diagrama conmutativo:

$$\begin{array}{ccc} K/\mathfrak{a} & \xrightarrow{\ s\ } & K/s\mathfrak{a} \\ \downarrow & & \downarrow \\ \displaystyle\bigoplus_\mathfrak{p} K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p} & \longrightarrow & \displaystyle\bigoplus_\mathfrak{p} K_\mathfrak{p}/s_\mathfrak{p}\mathfrak{a}_\mathfrak{p} \end{array}$$

donde la flecha inferior es el homomorfismo $(x_\mathfrak{p}) \mapsto (s_\mathfrak{p}x_\mathfrak{p})$.

Ahora ya podemos demostrar el teorema fundamental de la multiplicación compleja:

Teorema 13.18 *Sea K un cuerpo cuadrático imaginario, sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea el orden maximal $\mathcal{O}_K$ de K, sea $f : \mathbb{C}/\mathfrak{a} \longrightarrow E$ un isomorfismo analítico, donde $\mathfrak{a}$ es un ideal fraccional de K, sea $s \in J_K$ y $\sigma \in \mathrm{Aut}(\mathbb{C})$ un automorfismo cuya restricción a la máxima extensión abeliana de K sea el símbolo de Artin (K/s). Entonces existe un único isomorfismo analítico $f' : \mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow E^\sigma$ que hace conmutativo al diagrama siguiente:*

$$\begin{array}{ccc} K/\mathfrak{a} & \xrightarrow{\ s^{-1}\ } & K/s^{-1}\mathfrak{a} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ E & \xrightarrow{\ \sigma\ } & E^\sigma \end{array}$$

Demostración: Observemos que $K/\mathfrak{a}$ es el subgrupo de torsión del toro complejo $\mathbb{C}/\mathfrak{a}$, luego su imagen por f es E_{tor}. Todas las flechas del diagrama son inyectivas, luego f' está completamente determinada sobre $K/s^{-1}\mathfrak{a}$ y, como este subgrupo es denso en $\mathbb{C}/s^{-1}\mathfrak{a}$, la unicidad es inmediata.

Supongamos ahora que $E_1/\mathbb{C}$ es una curva elíptica isomorfa a $E/\mathbb{C}$ y que $f_1 : \mathbb{C}/\mathfrak{a}_1 \longrightarrow E_1$ es un isomorfismo analítico. Vamos a probar que si (E_1, f_1) cumple el teorema, entonces (E, f) también lo cumple. Suponemos, pues, que existe un isomorfismo analítico $f_1' : \mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow E_1^\sigma$ y un diagrama conmutativo

$$\begin{array}{ccc} K/\mathfrak{a}_1 & \xrightarrow{\ s^{-1}\ } & K/s^{-1}\mathfrak{a}_1 \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ E_1 & \xrightarrow{\ \sigma\ } & E_1^\sigma \end{array}$$

Fijemos también un isomorfismo $i : E_1 \longrightarrow E$. Además ha de existir un $\gamma \in K^*$ tal que $\mathfrak{a}_1 = \gamma\mathfrak{a}$. Es fácil ver que todos los cuadrados del diagrama siguiente son conmutativos:

$$\begin{array}{ccc}
K/\mathfrak{a} & \xrightarrow{s^{-1}} & K/s^{-1}\mathfrak{a} \\
{\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\gamma} \\
K/\mathfrak{a}_1 & \xrightarrow{s^{-1}} & K/s^{-1}\mathfrak{a}_1 \\
{\scriptstyle f_1}\downarrow & & \downarrow{\scriptstyle f_1'} \\
E_1 & \xrightarrow{\sigma} & E_1^\sigma \\
{\scriptstyle i}\downarrow & & \downarrow{\scriptstyle i^\sigma} \\
E & \xrightarrow{\sigma} & E^\sigma
\end{array}$$

Por lo tanto, podemos tomar como f' la composición de las flechas verticales de la derecha.

Así pues, podemos sustituir E por una curva isomorfa y $\mathfrak{a}$ por un ideal fraccional similar. Esto nos permite suponer que E está definida sobre $\mathbb{Q}(j(E))$ y que $\mathfrak{a}$ es un ideal (entero) de $\mathcal{O}_K$.

Fijemos un natural $m \geq 3$ y sea L/K una extensión finita de Galois que cumpla $j(E) \in L$ y $E[m] \subset E(L)$. El teorema 13.15 implica que L contiene el cuerpo de clases radiales K_m módulo m.

Vamos a construir un isomorfismo topológico f'_m que cumpla el teorema sobre los puntos de torsión de orden divisor de m. Empezamos eligiendo un primo $\mathfrak{P}$ de L que cumpla las condiciones siguientes:

a) $\sigma|_L = \left(\frac{L/K}{\mathfrak{P}}\right)$,

b) $\mathfrak{p} = \mathfrak{P} \cap K$ es un primo de grado 1, es decir, $p = \mathrm{N}_{\mathbb{Q}}^K(\mathfrak{p})$ es un primo racional.

c) $\mathfrak{p}$ es no ramificado en L.

d) $\mathfrak{p}$ no es ninguno de los primos del conjunto S considerado en la prueba del teorema 13.8.

e) $\mathfrak{P} \nmid m$.

La existencia de $\mathfrak{P}$ es consecuencia inmediata del teorema de Tchebotarev,[4] según el cual existen infinitos primos que cumplen las tres primeras condiciones, mientras que las dos últimas excluyen sólo a un número finito de ellos. Como L contiene el cuerpo radial K_m, la hipótesis sobre σ nos da que

$$\left(\frac{K}{s}\right)\Big|_{K_m} = \sigma|_{K_m} = \left(\frac{L/K}{\mathfrak{P}}\right)\Big|_{K_m} = \left(\frac{K_m/K}{\mathfrak{P}}\right) = \left(\frac{K_m/K}{\mathfrak{p}}\right).$$

[4]Teorema 9.19 de mi *Teoría de cuerpos de clases.*

Si $\pi \in J_K$ es un elemento ideal tal que $v_{\mathfrak{p}}(\pi) = 1$ y $\pi_{\mathfrak{q}} = 1$ para todo primo $\mathfrak{q} \neq \mathfrak{p}$, entonces

$$\left(\frac{K}{\pi}\right)\Big|_{K_m} = \left(\frac{K_m/K}{(\pi)}\right) = \left(\frac{K_m/K}{\mathfrak{p}}\right).$$

Por consiguiente

$$\left(\frac{K_m/K}{s\pi^{-1}}\right) = 1.$$

Ahora bien, el núcleo del homomorfismo de Artin para el cuerpo radial K_m es el grupo K^*W_m, de modo que $s\pi^{-1} = \alpha u$, para cierto $\alpha \in K^*$ y cierto $u \in W_m$. Esto último significa que

$$u_{\mathfrak{q}} \in U_{\mathfrak{q}}, \qquad \text{para todo ideal primo } \mathfrak{q}, \text{ y}$$

$$u_{\mathfrak{q}} \equiv 1 \ (\text{mód } \mathfrak{q}^{v_{\mathfrak{q}}(m)}), \qquad \text{para todo } \mathfrak{q} \mid m.$$

Por otra parte, L contiene el cuerpo de clases de Hilbert $H = K(j(E))$ y

$$\sigma|_H = \left(\frac{L/K}{\mathfrak{P}}\right)\Big|_H = \left(\frac{H/K}{\mathfrak{P}}\right) = \left(\frac{H/K}{\mathfrak{p}}\right).$$

Podemos aplicar el teorema 13.12, que nos da una isogenia $\lambda : E \longrightarrow E^\sigma$ cuya reducción módulo $\mathfrak{P}$ es la aplicación de Frobenius de grado p. (Notemos que al elegir el cuerpo L podemos exigir que sea lo suficientemente grande como para que la isogenia λ que tomamos ahora esté definida sobre L, pues, a priori, las posibles curvas E^σ son un número finito y los grupos de isogenias son finitamente generados.)

Observemos además que en la demostración de 13.12 se ve que podemos tomar $\mathfrak{p} * E = E^\sigma$, y que la isogenia λ es entonces la composición de la multiplicación por $\mathfrak{p}$ con un automorfismo ϵ de E^σ. Ahora bien, la multiplicación por $\mathfrak{p}$ está definida salvo automorfismos, porque depende de la elección de los isomorfismos $f : \mathbb{C}/\mathfrak{a} \longrightarrow E$ y $f'' : \mathbb{C}/\mathfrak{p}^{-1}\mathfrak{a} \longrightarrow E^\sigma$. Si cambiamos f'' por su composición con ϵ, entonces λ es simplemente la multiplicación por $\mathfrak{p}$, y tenemos el diagrama conmutativo siguiente:

$$\begin{array}{ccccc} \mathbb{C} & \longrightarrow & \mathbb{C}/\mathfrak{a} & \xrightarrow{f} & E \\ {\scriptstyle 1}\downarrow & & \downarrow & & \downarrow{\scriptstyle \lambda} \\ \mathbb{C} & \longrightarrow & \mathbb{C}/\mathfrak{p}^{-1}\mathfrak{a} & \xrightarrow[f'']{} & E^\sigma \end{array}$$

Vamos a probar que λ y σ actúan igual sobre $E[m]$. Para ello tomamos un punto $P \in E[m]$ y consideramos la reducción módulo $\mathfrak{P}$. Entonces

$$\widetilde{\lambda(T)} = \tilde{\lambda}(\tilde{T}) = \widetilde{T^\sigma},$$

pues tanto $\tilde{\lambda}$ como $\sigma|_L = \left(\frac{L/K}{\mathfrak{P}}\right)$ actúan sobre los puntos de $E(L)$ elevando a p las coordenadas.

Como $\mathfrak{P} \nmid m$, el teorema 6.19 nos da que la reducción es inyectiva sobre $E[m]$, luego $\lambda(T) = T^\sigma$.

Recordemos ahora la descomposición $s = \alpha\pi u$ del elemento ideal del enunciado. Claramente $(s) = (\alpha)(\pi) = (\alpha)\mathfrak{p}$, luego $s^{-1}\mathfrak{a} = \alpha^{-1}\mathfrak{p}^{-1}\mathfrak{a}$. Así pues, la multiplicación por α^{-1} nos da un isomorfismo $\mathbb{C}/\mathfrak{p}^{-1}\mathfrak{a} \longrightarrow \mathbb{C}/s^{-1}\mathfrak{a}$, con lo que eligiendo adecuadamente un isomorfismo $f'_m : \mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow E^\sigma$ tenemos el siguiente diagrama conmutativo:

$$\begin{array}{ccccc}
\mathbb{C} & \longrightarrow & \mathbb{C}/\mathfrak{a} & \xrightarrow{f} & E \\
{\scriptstyle 1}\downarrow & & \downarrow & & \downarrow{\scriptstyle \lambda} \\
\mathbb{C} & \longrightarrow & \mathbb{C}/\mathfrak{p}^{-1}\mathfrak{a} & \xrightarrow{f''} & E^\sigma \\
{\scriptstyle \alpha^{-1}}\downarrow & & \downarrow & & \downarrow{\scriptstyle 1} \\
\mathbb{C} & \longrightarrow & \mathbb{C}/s^{-1}\mathfrak{a} & \xrightarrow{f'_m} & E^\sigma
\end{array}$$

Vamos a demostrar que f'_m cumple lo que queríamos, es decir, la conmutatividad del diagrama

$$\begin{array}{ccc}
m^{-1}\mathfrak{a}/\mathfrak{a} & \xrightarrow{f} & E \\
{\scriptstyle s^{-1}}\downarrow & & \downarrow{\scriptstyle \sigma} \\
m^{-1}s^{-1}\mathfrak{a}/s^{-1}\mathfrak{a} & \xrightarrow[f'_m]{} & E^\sigma
\end{array}$$

Concretamente, si $t \in m^{-1}\mathfrak{a}/\mathfrak{a}$, hemos de probar que $f(t)^\sigma = f'_m(s^{-1}t)$. Lo que sabemos es que $f(t)^\sigma = \lambda(f(t)) = f'_m(\alpha^{-1}t)$. Como f'_m es un isomorfismo, basta ver que $\alpha^{-1}t = s^{-1}t$.

Para cada ideal primo $\mathfrak{q}$ de K, es claro que la componente $\mathfrak{q}$-primaria de $m^{-1}\mathfrak{a}/\mathfrak{a}$ es un submódulo de la componente $\mathfrak{q}$-primaria de $K/\mathfrak{a}$, lo que significa que la descomposición de t en suma de componentes primarias como elemento de ambos $\mathcal{O}_K$-módulos es la misma. Teniendo en cuenta que tanto la multiplicación por α^{-1} como la multiplicación por s^{-1} son homomorfismos de $\mathcal{O}_K$-módulos, basta probar la igualdad $\alpha^{-1}t = s^{-1}t$ cuando t pertenece a una componente $\mathfrak{q}$-primaria. En tal caso ambos miembros están en la componente $\mathfrak{q}$-primaria de $K/s^{-1}\mathfrak{a}$ y, por 13.17 b) basta probar la igualdad de sus imágenes en $K_\mathfrak{q}/s_\mathfrak{q}^{-1}\mathfrak{a}_\mathfrak{q}$. Si $t = [\tau]$, lo que hemos de ver es que $[\alpha^{-1}\tau] = [s_\mathfrak{q}^{-1}\tau]$ o, equivalentemente, que $\alpha^{-1}\tau - s_\mathfrak{q}^{-1}\tau \in s_\mathfrak{q}^{-1}\mathfrak{a}_\mathfrak{q}$, para todo $\tau \in m^{-1}\mathfrak{a}_\mathfrak{q}$.

Teniendo en cuenta la descomposición $s_\mathfrak{q} = \alpha\pi_\mathfrak{q}u_\mathfrak{q}$, esto es equivalente a que $\tau - \pi_\mathfrak{q}u_\mathfrak{q}^{-1}\tau \in \pi_\mathfrak{q}^{-1}u_\mathfrak{q}^{-1}\mathfrak{a}_\mathfrak{q}$, o también a que $\pi_\mathfrak{q}u_\mathfrak{q}\tau - \tau \in \mathfrak{a}_\mathfrak{q}$, para todo $\tau \in m^{-1}\mathfrak{a}_\mathfrak{q}$. A su vez, esto equivale a la inclusión $(\pi_\mathfrak{q}u_\mathfrak{q} - 1)m^{-1}\mathfrak{a}_\mathfrak{q} \subset \mathfrak{a}_\mathfrak{q}$, o también a ésta otra: $(\pi_\mathfrak{q}u_\mathfrak{q} - 1)\mathfrak{a}_\mathfrak{q} \subset m\mathfrak{a}_\mathfrak{q}$.

Observemos que $u_\mathfrak{q} \equiv 1$ (mód m) (esto es trivial si $\mathfrak{q} \nmid m$), luego basta probar la inclusión $(\pi_\mathfrak{q} - 1)\mathfrak{a}_\mathfrak{q} \subset m\mathfrak{a}_\mathfrak{q}$. Si $\mathfrak{q} \neq \mathfrak{p}$ entonces $\pi_\mathfrak{q} = 1$ y la inclusión es trivial.

Así pues, sólo hemos de demostrar que $(\pi_\mathfrak{p} - 1)\mathfrak{a}_\mathfrak{p} \subset \mathfrak{a}_\mathfrak{p}$, donde hemos eliminado m porque $\mathfrak{p} \nmid m$ y, por consiguiente, $m\mathfrak{a}_\mathfrak{p} = \mathfrak{a}_\mathfrak{p}$. La inclusión es ahora trivial (pues $\mathfrak{a}_\mathfrak{p}$ es un ideal).

El diagrama conmutativo de f'_m es el que requiere el enunciado excepto porque está restringido a $m^{-1}\mathfrak{a}/\mathfrak{a}$. El teorema quedará probado si demostramos que todos los isomorfismos f'_m son el mismo.

Sea $n \geq 1$ un número natural y sea $f'_{mn} : \mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow E^\sigma$ el isomorfismo correspondiente a mn. La composición $f'^{-1}_m \circ f'_{mn}$ es un automorfismo de E^σ, luego es la multiplicación por una cierta unidad $\epsilon \in \mathcal{O}_K$. Así pues, $f'_{mn} = f'_m \circ \epsilon$. Para cada $t \in m^{-1}\mathfrak{a}/\mathfrak{a}$ tenemos

$$\epsilon(f'_m(s^{-1}t)) = f'_{mn}(s^{-1}t) = f(t)^\sigma = f'_m(s^{-1}t).$$

Cuando t recorre $m^{-1}\mathfrak{a}/\mathfrak{a}$, el punto $f'_m(s^{-1}t)$ recorre $E^\sigma[m]$, luego tenemos que $\epsilon T = T$ para todo $T \in E^\sigma[m]$. Hemos tomado $m \geq 3$, luego la isogenia $\epsilon - 1$ tiene al menos 9 puntos en su núcleo. Esto obliga a que $\epsilon = 1$, pues en caso contrario sería $\mathrm{N}(\epsilon - 1) \geq 9$, cuando es fácil ver que si ϵ es una unidad cuadrática $\mathrm{N}(\epsilon - 1) \leq 4$. ∎

A partir del teorema fundamental pueden deducirse fácilmente los resultados de las secciones precedentes, pero en lugar de hacer esto, que no nos aportaría ninguna información nueva, vamos a ver que el teorema fundamental se generaliza fácilmente a curvas elípticas cuyo anillo de multiplicaciones complejas es un orden cuadrático arbitrario (no necesariamente maximal) y a partir de aquí generalizaremos correspondientemente los resultados que ya conocemos sobre multiplicación compleja. Para ello necesitamos unos preliminares que exponemos en la sección siguiente:

13.5 Módulos completos

El único problema que presenta la generalización del teorema fundamental a órdenes arbitrarios es que hay que definir el producto de un elemento ideal por un módulo completo arbitrario, mientras que ahora sólo lo tenemos definido para ideales fraccionales. Nos ocupamos de ello en esta sección.

Sea K un cuerpo numérico de grado n y sea p un primo racional. Para cada divisor primo $\mathfrak{p}$ de K representaremos por $K_\mathfrak{p}$ la compleción de K correspondiente a $\mathfrak{p}$. En particular, $\mathbb{Q}_p$ será el cuerpo de los números p-ádicos. Definimos la $\mathbb{Q}_p$-álgebra (conmutativa y unitaria, pero con divisores de cero)

$$K_p = \bigoplus_{\mathfrak{p}|p} K_\mathfrak{p}.$$

Es claro que la dimensión de K_p sobre $\mathbb{Q}_p$ es n. Podemos definir una norma en K_p mediante $\|x\|_p = \max_{\mathfrak{p}|p} |x_\mathfrak{p}|_\mathfrak{p}$.

En un espacio vectorial de dimensión finita sobre un cuerpo métrico completo, todas las normas son equivalentes, y determinan un espacio normado completo. Podemos identificar cada $\alpha \in \mathbb{Q}_p$ con el vector $(\alpha)_\mathfrak{p} \in K_p$, de modo que el producto escalar (externo) de elementos de K_p por elementos de $\mathbb{Q}_p$ coincide con el producto interno. Notemos además que si $\alpha \in \mathbb{Q}_p$ entonces $\|\alpha\|_p = |\alpha|_p$, luego la topología inducida en $\mathbb{Q}_p$ desde K_p es la usual. En particular $\mathbb{Q}_p$ es la clausura de $\mathbb{Q}$.

También podemos identificar cada $\alpha \in K$ con el vector $(\alpha)_\mathfrak{p} \in K_p$, de modo que $K \subset K_p$. Esta identificación es consistente con la identificación $\mathbb{Q}_p \subset K_p$ para elementos de $\mathbb{Q}$. El teorema de aproximación nos da que K es denso en K_p, pues dados $\epsilon > 0$ y $x \in K_p$, existe un $\alpha \in K$ tal que $|x_\mathfrak{p} - \alpha|_\mathfrak{p} < \epsilon$ para todo $\mathfrak{p}$, luego $\|x - \alpha\|_p < \epsilon$.

Más aún, si llamamos $\mathcal{O}_\mathfrak{p}$ al anillo de enteros de $K_\mathfrak{p}$ y

$$\mathcal{O}_p = \bigoplus_{\mathfrak{p}|p} \mathcal{O}_\mathfrak{p} = \{x \in K_p \mid \|x\|_p \leq 1\},$$

se cumple que el orden maximal $\mathcal{O}_K$ de K es denso en $\mathcal{O}_p$. En efecto, todo $x \in \mathcal{O}_p$ se aproxima por un $\alpha \in K$ tal que $\|\alpha\|_p \leq 1$, luego basta probar que todo α en estas condiciones puede aproximarse por un $\beta \in \mathcal{O}_K$. Por el teorema chino del resto podemos tomar $u \in \mathcal{O}_K$ tal que $|u - 1|_\mathfrak{p} < \epsilon$ para todo $\mathfrak{p} \mid p$ y $|u|_\mathfrak{q} < |\alpha|_\mathfrak{q}^{-1}$ para todo ideal primo $\mathfrak{q}$ de K tal que $|\alpha|_\mathfrak{q} > 1$. Así, $\beta = \alpha u$ cumple que $|\alpha - \beta|_\mathfrak{p} = |\alpha|_\mathfrak{p}|u - 1|_\mathfrak{p} < \epsilon$, para todo $\mathfrak{p} \mid p$, (luego $\|\alpha - \beta\|_p < \epsilon$) y $|\beta|_\mathfrak{q} \leq 1$ para todo ideal primo $\mathfrak{q}$ de K, luego[5] $\beta \in \mathcal{O}_K$.

Veamos ahora que toda $\mathbb{Q}$-base de K es también una $\mathbb{Q}_p$-base de K_p. Basta probarlo para una base entera $\alpha_1, \ldots, \alpha_n$ (pues cualquier otra base se diferencia de ésta en un cambio de variables con coeficientes racionales y determinante no nulo, y esto prueba que también es una $\mathbb{Q}_p$-base). Como las dimensiones son las mismas, basta ver que la base dada es un sistema generador de K_p. Para ello tomamos un $x \in K_p$ arbitrario. Existe un $m \in \mathbb{Z}$ no nulo tal que $mx \in \mathcal{O}_p$ Sea $\{\alpha_i\}_i$ una sucesión en $\mathcal{O}_K$ convergente a mx. Pongamos que

$$\alpha_i = m_{i1}\alpha_1 + \cdots + m_{in}\alpha_n, \qquad m_{ij} \in \mathbb{Z}.$$

Como $\mathbb{Z}_p$ es compacto, pasando a una subsucesión podemos suponer que las sucesiones $\{m_{ij}\}_i$ convergen a enteros p-ádicos $\beta_i \in \mathbb{Z}_p$, de donde se sigue que $mx = \beta_1\alpha_1 + \cdots + \beta_n\alpha_n$ y, por consiguiente, $x \in \langle \alpha_1, \ldots, \alpha_n \rangle_{\mathbb{Q}_p}$.

El argumento que acabamos de emplear demuestra, más en general, que si $M = \langle \alpha_1, \ldots, \alpha_n \rangle_\mathbb{Z}$, para ciertos $\alpha_1, \ldots, \alpha_n \in K$, entonces la clausura de M en K_p es $M_p = \langle \alpha_1, \ldots, \alpha_n \rangle_{\mathbb{Z}_p}$. En particular, si M es un módulo completo de K, es decir, si los α_i son linealmente independientes (sobre $\mathbb{Z}$ o, equivalentemente, sobre $\mathbb{Q}$), entonces M_p es un $\mathbb{Z}_p$-módulo libre de rango n con la misma base.

[5] Si $\beta = a/b$, con a, $b \in \mathcal{O}_K$, entonces $v_\mathfrak{q}(b) \leq v_\mathfrak{q}(a)$ para todo $\mathfrak{p}$, luego $(b) \mid (a)$, luego $(a) \subset (b)$ y $\beta \in \mathcal{O}_K$.

Si M es un módulo completo de K, definimos $M_{(p)} = M_p \cap K$. Vamos a ver que

$$M_{(p)} = \{m/a \mid m \in M,\ a \in \mathbb{Z},\ p \nmid a\}.$$

Si m/a es de la forma indicada, tenemos que $1/a \in \mathbb{Z}_p$, luego $m/a \in M_p \cap K$.

Para probar el recíproco fijamos una $\mathbb{Z}$-base $\alpha_1, \dots, \alpha_n$ de M, que también es una $\mathbb{Q}$-base de K y una $\mathbb{Q}_p$-base de K_p. Así, todo $x \in K_p$ se expresa de forma única como

$$x = \beta_1\alpha_1 + \cdots + \beta_n\alpha_n, \qquad \beta_i \in \mathbb{Q}_p.$$

Consideramos la norma de K_p dada por $\|x\| = \max_i |\beta_i|_p$. Como todas las normas de K_p son equivalentes, M_p también es la clausura de M respecto de esta norma. Si $x \in M$, entonces las coordenadas β_i son enteros racionales, luego $\|x\| \leq 1$ y, como la norma es continua, todo $x \in M_p$ cumple $\|x\| \leq 1$. En particular, si $x \in M_{(p)}$ tenemos que sus coordenadas son números racionales $\beta_i \in \mathbb{Q}$ tales que $|\beta_i|_p \leq 1$, luego son de la forma $\beta_i = b_i/a$, con b_i, $a \in \mathbb{Z}$, $p \nmid a$. De aquí se sigue inmediatamente que $x = m/a$, con $m \in M$.

Notemos que si llamamos $\mathbb{Z}_{(p)} = \mathbb{Z}_p \cap \mathbb{Q}$, entonces toda $\mathbb{Z}$-base de M es también una $\mathbb{Z}_{(p)}$-base de $M_{(p)}$.

Hasta aquí hemos mantenido fijo el primo p. Ahora vamos a estudiar la relación entre las compleciones de un mismo módulo M respecto a primos distintos. En primer lugar observamos que un módulo completo está determinado por sus compleciones:

$$M = \bigcap_p M_{(p)}.$$

En efecto, fijada una base de M, consideramos como antes la norma en K_p determinada por el máximo valor absoluto p-ádico de sus coordenadas. Si x está en la intersección, hemos visto que tiene norma ≤ 1 para todos los primos p, luego sus coordenadas $\beta_i \in \mathbb{Q}$ cumplen $|\beta_i|_p \leq 1$ para todo primo p, luego necesariamente $\beta_i \in \mathbb{Z}$ y, por lo tanto, $x \in M$.

Las distintas compleciones de un mismo módulo completo están sometidas a una restricción: para todo primo p salvo a lo sumo un número finito de ellos, se ha de cumplir que

$$M_{(p)} = (\mathcal{O}_K)_{(p)}, \qquad M_p = \mathcal{O}_p.$$

En efecto, si fijamos una base de $\mathcal{O}_K$ y una base de M, la matriz A de cambio de base tiene entradas racionales y determinante no nulo. Para todos los primos p salvo a lo sumo un número finito de ellos, los coeficientes de A serán enteros p-ádicos y el determinante será una unidad p-ádica. De aquí se siguen inmediatamente ambas igualdades.

El resultado principal que necesitamos es el recíproco:

Teorema 13.19 *Si $\{N_p\}_p$ es una sucesión de $\mathbb{Z}_{(p)}$-módulos libres de rango n tales que $N_p \subset K$ y $N_p = (\mathcal{O}_K)_{(p)}$ para todo primo p salvo a lo sumo un número finito de ellos, entonces existe un único módulo completo M tal que $M_{(p)} = N_p$ para todo primo p.*

DEMOSTRACIÓN: La unicidad es obvia. De hecho, ha de ser $M = \bigcap_p N_p$. Sólo hemos de comprobar que M cumple lo pedido.

Sea S el conjunto finito de primos p para los que $N_p \neq (\mathcal{O}_K)_{(p)}$. Fijemos una base B de $\mathcal{O}_K$. Las coordenadas en B de cualquier elemento de M son números racionales con denominador divisible a lo sumo entre primos de S. Vamos a ver que el exponente de cada primo $p \in S$ en los denominadores de dichas coordenadas está acotado superiormente. Esto equivale a que el valor absoluto p-ádico de las coordenadas esté acotado superiormente. Tomamos una $\mathbb{Z}_{(p)}$-base B' de N_p y consideramos la matriz A de cambio de base con respecto a B. Las coordenadas en B de un elemento de M son combinación lineal (con coeficientes en A) de sus coordenadas en B', que están en $\mathbb{Z}_{(p)}$, luego el valor absoluto p-ádico de las coordenadas respecto a B está acotado por el máximo valor absoluto p-ádico de los coeficientes de A.

Como consecuencia, existe un $m \in \mathbb{Z}$ no nulo (suficientemente divisible entre los primos de S) tal que $mM \subset \mathcal{O}_K$. Esto implica que mM es un $\mathbb{Z}$-módulo finitamente generado.

Por otra parte, para cada $p \in S$ existe un $r \geq 0$ tal que $p^r\mathcal{O}_K \subset N_p$ (basta expresar los elementos de B como combinaciones lineales racionales de una $\mathbb{Z}_{(p)}$-base de N_p y considerar la potencia de p que hace enteros p-ádicos a todos los coeficientes). Como S es finito, existe un $m' \in \mathbb{Z}$ no nulo tal que $m'\mathcal{O}_K \subset M$. Esto implica que el $\mathbb{Z}$-módulo $\mathcal{O}_K/mM$ es de torsión, luego el rango de mM ha de ser n.

Tenemos probado que M es un módulo completo de K. Es evidente que $M_{(p)} \subset N_p$ para todo primo p, y se da la igualdad salvo a lo sumo para un conjunto finito de primos. Si existe $x \in N_p \setminus M_{(p)}$, razonando igual que antes podemos encontrar un $m' \in \mathbb{Z}$ tal que $p \nmid m'$ y $m'x \in M$. En particular $m'x \in N_p$, pero esto es imposible, pues el hecho de que $x \notin M_{(p)}$ equivale a que alguna de las coordenadas de x en una $\mathbb{Z}$-base de M no está en $\mathbb{Z}_{(p)}$, y a $m'x$ le sucede lo mismo. Así pues $M_{(p)} = N_p$ y M cumple lo que pide el enunciado. ∎

En realidad nos interesa la versión para compleciones:

Teorema 13.20 *Si $\{N_p\}_p$ es una sucesión de $\mathbb{Z}_p$-módulos libres de rango n tales que $N_p \subset K_p$ y $N_p = \mathcal{O}_p$ para todo primo p salvo a lo sumo un número finito de ellos, entonces existe un único módulo completo M tal que $M_p = N_p$ para todo primo p.*

DEMOSTRACIÓN: Veamos en primer lugar que $N_{(p)} = N_p \cap K$ es un $\mathbb{Z}_{(p)}$-módulo libre de rango n. Es claro que existe un $r \geq 0$ tal que $p^rN_p \subset \mathcal{O}_p$, luego $p^rN_{(p)} \subset \mathcal{O}_p \cap K = (\mathcal{O}_K)_{(p)}$. Esto implica que $p^rN_{(p)}$ es un $\mathbb{Z}_{(p)}$-módulo libre de rango $\leq n$. Igualmente, existe un $s \geq 0$ tal que $p^s\mathcal{O}_p \subset N_p$, luego $p^s(\mathcal{O}_K)_{(p)} \subset N_{(p)}$, de donde se sigue que el módulo $(\mathcal{O}_K)_{(p)}/p^rN_{(p)}$ es de torsión, luego $p^rN_{(p)}$ tiene rango n y $N_{(p)}$ también.

Podemos aplicar el teorema anterior, que nos da un módulo completo M tal que $M_{(p)} = N_{(p)}$ para todo primo p. Ahora observamos que N_p es abierto y

cerrado en K_p, pues si consideramos la norma asociada a una $\mathbb{Z}_p$-base de N_p (la definida como el máximo de los valores absolutos de las coordenadas), tenemos que N_p es el conjunto de los elementos de K_p de norma ≤ 1. Por consiguiente, puesto que K es denso en K_p, podemos concluir que $N_{(p)}$ es denso en N_p (y lo mismo vale para $M_{(p)}$ y M_p), luego $N_p = \overline{N_{(p)}} = \overline{M_{(p)}} = M_p$. La unicidad es evidente. ∎

Más adelante necesitaremos el teorema siguiente:

Teorema 13.21 *Sea K un cuerpo cuadrático y M un módulo completo de K con anillo de coeficientes $\mathcal{O}$. Entonces M es localmente principal, es decir, para cada primo p existe un $\alpha \in K^*$ tal que $M_{(p)} = \alpha\mathcal{O}_{(p)}$, luego también $M_p = \alpha\mathcal{O}_p$.*

Demostración: Llamemos $f \in \mathbb{Z}$ al conductor de $\mathcal{O}$. Esto significa que si $\mathcal{O}_K = \langle 1, \alpha\rangle_{\mathbb{Z}}$, entonces $\mathcal{O} = \langle 1, f\alpha\rangle_{\mathbb{Z}}$, luego la matriz de cambio de base tiene determinante f y si $p \nmid f$ entonces $\mathcal{O}_{(p)} = (\mathcal{O}_K)_{(p)}$ es un dominio de ideales principales.

Es conocido que M es similar a un ideal de $\mathcal{O}$ de norma prima con cualquier entero prefijado, por ejemplo f. Esto significa que existe un $\mu \in K^*$ tal que μM es un ideal en dichas condiciones.

Si $p \nmid f$, entonces $\mu M_{(p)} = \beta\mathcal{O}_{(p)}$ para cierto $\beta \in \mathcal{O}_{(p)}$, (puesto que $\mathcal{O}_{(p)}$ es un dominio de ideales principales), luego el teorema se cumple tomando $\alpha = \mu^{-1}\beta$.

Si $p \mid f$ entonces $\mu M_{(p)} = \mathcal{O}_{(p)}$. En efecto, todo elemento de $\mathcal{O}_{(p)}$ es de la forma β/b, donde $\beta \in \mathcal{O}$, $b \in \mathbb{Z}$, $p \nmid b$. Como el cociente $\mathcal{O}/\mu M$ es finito y tiene cardinal primo con p, existe un número natural c tal que $p \nmid c$ y $c\alpha \in \mu M$. Entonces $\beta/b = c\beta/cb \in \mu M_{(p)}$. El teorema se cumple con $\alpha = \mu^{-1}$. ∎

Sean

$$K_p^* = \bigoplus_{\mathfrak{p}|p} K_{\mathfrak{p}}^*, \qquad U_p = \bigoplus_{\mathfrak{p}|p} U_{\mathfrak{p}}$$

los grupos de unidades de K_p y $\mathcal{O}_p$, respectivamente. El grupo de elementos ideales J_K puede verse como el subgrupo de

$$\prod_p K_p^* \times \prod_{\mathfrak{p}|\infty} K_{\mathfrak{p}}^*$$

formado por los elementos cuyas componentes están todas en U_p salvo a lo sumo un número finito de ellas.

Dado un módulo completo M de K y un elemento ideal $s \in J_K$, para cada primo $p \in \mathbb{Z}$ tenemos el $\mathbb{Z}_p$-módulo M_p (de rango n) y $s_p \in K_p^*$. Es claro que $s_p M_p$ es también un $\mathbb{Z}_p$-módulo de rango n, así como que casi todos ellos son iguales a $\mathcal{O}_p$. Por el teorema anterior existe un único módulo completo, al que representaremos por sM, tal que $(sM)_p = s_p M_p$ para todo primo p.

Aplicamos el teorema 13.17 a) con $K = \mathbb{Q}$, $\mathcal{O}_K = \mathbb{Z}$. Tenemos que K/M es un $\mathbb{Z}$-módulo de torsión, luego podemos descomponerlo como

$$K/M \cong \bigoplus_p (K/M)[p^\infty].$$

Ahora bien, $(K/M)[p^\infty] \cong K/M_{(p)}$, pues si $[\alpha] \in (K/M)[p^\infty]$ está en el núcleo del homomorfismo inducido por la identidad, entonces $\alpha \in M_{(p)}$, es decir, $\alpha = m/a$, con $m \in M$, $a \in \mathbb{Z}$, $p \nmid a$, pero por otra parte existe un $r \geq 0$ tal que $p^r\alpha \in M$. Esto implica que las coordenadas de m en una base de M han de ser divisibles entre a, luego $\alpha \in M$ y $[\alpha] = 0$.

Por otra parte, dado $[\alpha] \in K/M_{(p)}$, podemos encontrar un $a \in \mathbb{Z}$ no nulo tal que $m = a\alpha \in M$. Pongamos que $a = p^r b$, donde $p \nmid b$. Por el teorema de Bezout existen u, $v \in \mathbb{Z}$ tales que $p^r u + bv = 1$, luego

$$\alpha = \frac{m}{p^r b} = \frac{p^r um + bvm}{p^r b} = \frac{u}{b}\, m + \frac{1}{p^r}(vm).$$

El primer sumando está en $M_{(p)}$, luego llamando $\beta = vm/p^r$ tenemos que $[\alpha] = [\beta] \in K/M_{(p)}$, y la clase $[\beta] \in K/M$ cumple $p^r[\beta] = 0$, luego podemos concluir que $[\beta] \in (K/M)[p^\infty]$ es una antiimagen de $[\alpha]$.

Ahora tenemos el isomorfismo

$$K/M \cong \bigoplus_p K/M_{(p)}.$$

Por último observamos que la inclusión induce un isomorfismo

$$K/M_{(p)} \cong K_p/M_p.$$

Obviamente es inyectivo y, como M_p es abierto en K_p y K es denso, dado $\alpha \in K_p$ podemos encontrar un $\beta \in K$ tal que $\beta \in \alpha + M_p$, de modo que $[\alpha] = [\beta]$.

En definitiva tenemos un isomorfismo natural

$$K/M \cong \bigoplus_p K_p/M_p.$$

Dado un elemento ideal $s \in J_K$, tenemos igualmente

$$K/sM \cong \bigoplus_p K_p/s_pM_p,$$

y la multiplicación por s_p determina homomorfismos $K_p/M_p \longrightarrow K_p/s_pM_p$. Así pues, podemos definir la multiplicación $s : K/M \longrightarrow K/sM$ como el único homomorfismo que hace conmutativo el diagrama

$$\begin{array}{ccc} K/M & \xrightarrow{\ s\ } & K/sM \\ \downarrow & & \downarrow \\ \bigoplus_p K_p/M_p & \xrightarrow[s_p]{} & \bigoplus_p K_p/s_pM_p \end{array}$$

Sólo nos falta demostrar que el producto sM que acabamos de definir y el homomorfismo $s : K/M \longrightarrow K/sM$ coinciden con los que ya teníamos definidos

cuando $\mathfrak{a}$ es un ideal fraccional de K. Para ello empezamos observando que si $\mathfrak{a}$ es un ideal fraccional, entonces

$$\mathfrak{a}_p = \bigoplus_{\mathfrak{p}|p} \mathfrak{a}_\mathfrak{p}.$$

En efecto, supongamos primeramente que $\mathfrak{a}$ es un ideal entero. Recordemos que $\mathfrak{a}_\mathfrak{p} = \mathfrak{p}^{v_\mathfrak{p}(\mathfrak{a})}$ es la clausura de $\mathfrak{a}$ en $K_\mathfrak{p}$. Esto nos da inmediatamente una inclusión: el producto de las clausuras es un cerrado en M_p que contiene a $\mathfrak{a}$, luego también a su clausura $\mathfrak{a}_p$. Consideremos ahora un punto $(\alpha_\mathfrak{p})_\mathfrak{p}$ en el producto de las clausuras. Podemos aproximar cada $\alpha_\mathfrak{p}$ respecto al valor absoluto $\mathfrak{p}$-ádico por un elemento $\alpha'_\mathfrak{p} \in \mathfrak{a}$. El teorema chino del resto nos da un $\alpha \in \mathcal{O}_K$ tal que

$$\alpha \equiv \alpha'_\mathfrak{p} \ (\text{mód } \mathfrak{p}^r), \qquad \alpha \equiv 0 \ (\text{mód } \mathfrak{q}^r),$$

para todo $\mathfrak{p} \mid p$ y todo $\mathfrak{q} \mid \mathfrak{a}$, $\mathfrak{q} \nmid p$. Si r es suficientemente grande tenemos que $\alpha \in \mathfrak{a}$ y es una aproximación del punto de partida.

Un ideal fraccional arbitrario es de la forma $\mathfrak{a}/m$, donde $\mathfrak{a}$ es un ideal entero y $m \in \mathbb{Z}$. Teniendo en cuenta que la multiplicación por $1/m$ es un homeomorfismo de K_p en sí mismo, el resultado se sigue inmediatamente del caso ya probado.

Si $s \in J_K$, entonces $s_p = (s_\mathfrak{p})_\mathfrak{p}$ y

$$s_p\mathfrak{a}_p = \bigoplus_{\mathfrak{p}|p} s_\mathfrak{p}\mathfrak{a}_\mathfrak{p} = \bigoplus_{\mathfrak{p}|p} \mathfrak{p}^{v_\mathfrak{p}(s_\mathfrak{p})}\mathfrak{a}_\mathfrak{p}.$$

En la sección anterior hemos definido $s\mathfrak{a}$ como $(s)\mathfrak{a}$ y este ideal cumple que

$$((s)\mathfrak{a})_p = \bigoplus_{\mathfrak{p}|p} ((s)\mathfrak{a})_\mathfrak{p} = \bigoplus_{\mathfrak{p}|p} \mathfrak{p}^{v_\mathfrak{p}(s_\mathfrak{p})}\mathfrak{a}_\mathfrak{p} = s_p\mathfrak{a}_p.$$

Así pues, según la definición de esta sección, también se cumple $s\mathfrak{a} = (s)\mathfrak{a}$. Claramente

$$K_p/\mathfrak{a}_p = \bigoplus_{\mathfrak{p}|p} K_\mathfrak{p}/\mathfrak{a}_\mathfrak{p},$$

de donde se sigue fácilmente que las dos definiciones de la multiplicación por s determinan en realidad el mismo homomorfismo.

En la sección siguiente necesitaremos esta caracterización de los módulos cuadráticos con un anillo de coeficientes dado:

Teorema 13.22 *Sea $\mathcal{O}$ un orden de un cuerpo cuadrático K y sea M un módulo completo en K. Entonces $\mathcal{O}$ es el anillo de coeficientes de M si y sólo si existe un elemento ideal $s \in J_K$ tal que $M = s\mathcal{O}$.*

Demostración: Ciertamente, los módulos de la forma $M = s\mathcal{O}$ tienen anillo de coeficientes $\mathcal{O}$, pues un $\alpha \in K^*$ es un coeficiente de M si y sólo si cumple $\alpha s\mathcal{O} = s\alpha\mathcal{O} \subset s\mathcal{O}$, si y sólo si $\alpha\mathcal{O} \subset \mathcal{O}$, si y sólo si $\alpha \in \mathcal{O}$.

Recíprocamente, si M tiene anillo de coeficientes $\mathcal{O}$, por el teorema 13.21 sabemos que para cada primo p existe un $s_p \in K^*$ tal que $M_p = s_p\mathcal{O}_p$. Además podemos exigir que $s_p = 1$ salvo a lo sumo para un número finito de primos p. Los s_p determinan un $s \in J_K$ tal que $M = s\mathcal{O}$. ∎

13.6 Órdenes arbitrarios

Ya estamos en condiciones de generalizar el teorema fundamental a curvas elípticas cuyo anillo de multiplicaciones complejas sea un orden cuadrático imaginario arbitrario:

Teorema 13.23 *Sea K un cuerpo cuadrático imaginario, sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea un orden $\mathcal{O}$ de K, sea $f : \mathbb{C}/R \longrightarrow E$ un isomorfismo analítico, donde $R \subset K$ es un retículo, sea $s \in J_K$ y $\sigma \in \mathrm{Aut}(\mathbb{C})$ un automorfismo cuya restricción a la máxima extensión abeliana de K sea el símbolo de Artin (K/s). Entonces existe un único isomorfismo analítico $f' : \mathbb{C}/s^{-1}R \longrightarrow E^\sigma$ que hace conmutativo al diagrama siguiente:*

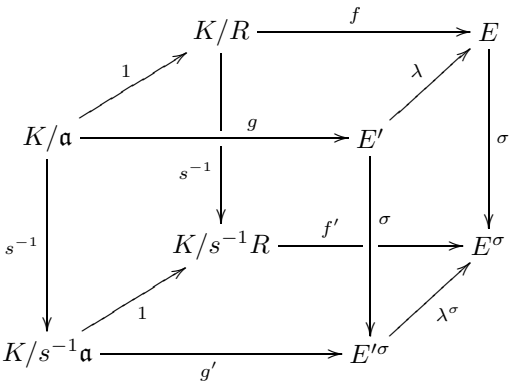
$$\begin{array}{ccc} K/R & \xrightarrow{s^{-1}} & K/s^{-1}R \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ E & \xrightarrow{\sigma} & E^\sigma \end{array}$$

Demostración: Sea $\mathcal{O}_K$ el orden maximal de K. Expresando una base de $\mathcal{O}_K$ en términos de una base de R se concluye que existe un $m \in \mathbb{Z}$ no nulo tal que $m\mathcal{O}_K \subset R$. Sea $\mathfrak{a}$ un ideal de $\mathcal{O}_K$ (por ejemplo $m\mathcal{O}_K$) tal que $\mathfrak{a} \subset R$ y sea $E'/\mathbb{C}$ una curva elíptica isomorfa a $\mathbb{C}/\mathfrak{a}$. Fijemos un isomorfismo $g : \mathbb{C}/\mathfrak{a} \longrightarrow E'$.

Observemos que se da también la inclusión $s^{-1}\mathfrak{a} \subset s^{-1}R$. En efecto, de la inclusión $\mathfrak{a} \subset R$ se sigue que $\mathfrak{a}_p \subset R_p$ para todo primo p, luego $s_p^{-1}\mathfrak{a}_p \subset s_p^{-1}R_p$ y, tomando la intersección primero con K y luego sobre p, obtenemos, en efecto, $s^{-1}\mathfrak{a} \subset s^{-1}R$.

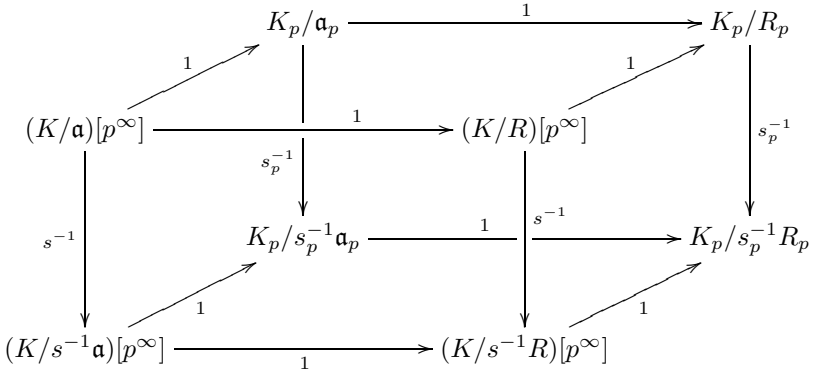
La identidad induce un homomorfismo $\mathbb{C}/\mathfrak{a} \longrightarrow \mathbb{C}/R$, que se corresponde a través de los isomorfismos fijados con una isogenia $\lambda : E' \longrightarrow E$. Por otra parte, el teorema 13.18 nos da un isomorfismo $g' : \mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow E'^\sigma$.

El diagrama siguiente resume la situación:



Aquí f' es el isomorfismo que queremos definir. La cara anterior es conmutativa por la construcción de g', la cara superior lo es por definición de λ, la

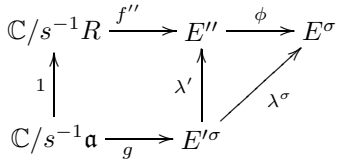
cara derecha lo es obviamente y vamos a probar que la izquierda también es conmutativa. Como la multiplicación por s^{-1} está definida independientemente sobre cada parte p-primaria, basta probar la conmutatividad sobre $(K/\mathfrak{a})[p^\infty]$. Consideramos el diagrama siguiente:



Queremos probar la conmutatividad de la cara anterior, y ésta se sigue inmediatamente de la conmutatividad de las restantes.

El núcleo de λ es $g[R/\mathfrak{a}]$, luego el núcleo de λ^σ es $\sigma[g[R/\mathfrak{a}]] = g'[s^{-1}[R/\mathfrak{a}]]$. Por otra parte, $R/\mathfrak{a}$ es el núcleo de la flecha superior de la cara izquierda del primer cubo. Teniendo en cuenta que la multiplicación por s^{-1} es un isomorfismo (su inversa es la multiplicación por s), es claro que $s^{-1}[R/\mathfrak{a}]$ ha de ser el núcleo de la flecha inferior de dicha cara, es decir, $s^{-1}R/s^{-1}\mathfrak{a}$. En definitiva, el núcleo de λ^σ es $g'[s^{-1}R/s^{-1}\mathfrak{a}]$.

Sea E'' una curva elíptica isomorfa a $\mathbb{C}/s^{-1}R$ y sea $f'' : \mathbb{C}/s^{-1}R \longrightarrow E''$ un isomorfismo. El homomorfismo $\mathbb{C}/s^{-1}\mathfrak{a} \longrightarrow \mathbb{C}/s^{-1}R$ se corresponde a través de g' y f'' con una isogenia $\lambda' : E'^\sigma \longrightarrow E''$. Su núcleo es el mismo que el de λ^σ, luego el teorema 2.37 nos da que E'' es isomorfa a E. Más aún, existe un isomorfismo ϕ que hace conmutativo el diagrama siguiente:



Llamamos $f' = f'' \circ \phi : \mathbb{C}/s^{-1}R \longrightarrow E^\sigma$, con lo que la cara inferior del primer cubo es conmutativa. La conmutatividad de las caras restantes implica la de la cara posterior, que es la que había que probar. ∎

Si $E/\mathbb{C}$ es una curva elíptica cuyo anillo de multiplicaciones complejas es un orden $\mathcal{O}$ de un cuerpo cuadrático K, podemos tomar un retículo R de K tal que $E \cong \mathbb{C}/R$. Para cada elemento ideal $s \in J_K$ podemos definir $s * E$ como una curva elíptica isomorfa a $\mathbb{C}/s^{-1}R$. Obviamente $s * E$ sólo está definida salvo isomorfismo y no depende de la elección de R. En particular, $j(s * E)$ está unívocamente determinado. El teorema siguiente generaliza a 13.9 c):

Teorema 13.24 *Si E es una curva elíptica cuyo anillo de multiplicaciones complejas es un orden de un cuerpo cuadrático imaginario K, entonces la extensión $K(j(E))/K$ es abeliana y para todo elemento ideal $s \in J_K$ se cumple*

$$j(s * E) = \left(\frac{K}{s}\right)(j(E)).$$

DEMOSTRACIÓN: Aplicamos el teorema anterior con $s = 1$, lo que nos da que si σ es cualquier automorfismo de $\mathbb{C}$ que fije a la máxima extensión abeliana de K, entonces $E^\sigma \cong E$, luego $j(E)^\sigma = j(E)$, luego $j(E)$ está en dicha extensión.

Ahora consideramos un s arbitrario y concluimos que $E^\sigma \cong s * E$, luego

$$j(s * E) = j(E)^\sigma = \left(\frac{K}{s}\right)(j(E)).$$

∎

Teorema 13.25 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea un orden $\mathcal{O}$ de un cuerpo cuadrático imaginario K y sea R un módulo completo en K tal que $\mathbb{C}/R \cong E$. Sea $W = \{s \in J_E \mid sR = R\}$. Entonces el grupo de clases de $K(j(E))/K$ es K^*W.*

DEMOSTRACIÓN: Observemos que $s \in W$ equivale a que $s_p R_p = R_p$ para todo primo p, lo que a su vez equivale a que $s_p \in \mathcal{O}_p^*$ para todo primo p (donde $\mathcal{O}_p^*$ es el grupo de unidades de $\mathcal{O}_p$). En efecto, por el teorema 13.21 existe un $\alpha \in K^*$ tal que $R_p = \alpha\mathcal{O}_p$, luego $s_p\alpha\mathcal{O}_p = \alpha\mathcal{O}_p$, luego $s_p\mathcal{O}_p = \mathcal{O}_p$, y esto implica claramente que $s_p \in \mathcal{O}_p^*$. Igualmente se prueba el recíproco. En definitiva:

$$W = K_\infty^* \times \prod_p \mathcal{O}_p^*.$$

Observemos que para casi todo primo p se cumple que $\mathcal{O}_p = (\mathcal{O}_K)_p$, luego $\mathcal{O}_p^* = U_p$ (el grupo de unidades de $\mathcal{O}_K$). Para los primos restantes, $\mathcal{O}_p$ es cerrado en $(\mathcal{O}_K)_p$ y tiene índice finito, luego es abierto y cerrado en $(\mathcal{O}_K)_p$ y, por consiguiente, en K_p^*. Esto implica que $\mathcal{O}_p^*$ es abierto y cerrado en K_p^*, luego W es abierto y cerrado en J_K, luego K^*W es un subgrupo abierto y cerrado de J_K.

Sea L el cuerpo de clases de K^*W. Hemos de probar que $L = K(j(E))$. Para ello tomamos un K-automorfismo σ de la mayor extensión abeliana de K. Si σ fija a L, entonces existe un $s \in W$ tal que $\sigma = (K/s)$. Por definición de W tenemos que $s^{-1}R = R$, luego $s * E = E$, y por el teorema anterior $j(E)^\sigma = j(E)$. Esto demuestra que $K(j(E)) \subset L$.

Recíprocamente, si $j(E)^\sigma = j(E)$, tomemos $s \in J_E$ tal que $\sigma = (K/s)$. El teorema anterior implica que $s * E \cong E$, luego existe un $\alpha \in K^*$ tal que $\alpha s^{-1}R = R$. Llamamos $s' = \alpha s^{-1} \in J_K$, de modo que $s' \in W$ y $s \in K^*W$, luego σ fija a L. Con esto tenemos la igualdad. ∎

Observemos ahora que si $\mathcal{O}$ es un orden cuadrático imaginario y $\mathcal{H}_\mathcal{O}$ es el grupo de las clases de similitud módulos completos con anillo de coeficientes $\mathcal{O}$,

el teorema 13.22 nos da un epimorfismo $\phi : J_K \longrightarrow \mathcal{H}_\mathcal{O}$ dado por $s \mapsto [s\mathcal{O}]$. Un elemento ideal s está en el núcleo si y sólo si $s\mathcal{O} = \alpha\mathcal{O}$, para cierto $\alpha \in K^*$, luego el núcleo es K^*W, donde W es el grupo definido en el teorema anterior. Así pues, $J_K/K^*W \cong \mathcal{H}_\mathcal{O}$. En particular, en las condiciones del teorema anterior, $|K(j(E)) : K| = h_\mathcal{O}$ (donde $h_\mathcal{O}$ es el número de clases de $\mathcal{O}$).

Más aún, sea f el conductor de $\mathcal{O}$ y sea H el grupo de clases de $\mathcal{O}$ módulo f. Esto significa que $I(f)/H$ es isomorfo al grupo de Galois del cuerpo de clases de $\mathcal{O}$. Por otra parte, tenemos un isomorfismo $\mathcal{H}_\mathcal{O} \cong I(f)/H$ que se calcula como sigue: a cada clase de $\mathcal{H}_\mathcal{O}$ le escogemos como representante un ideal de norma prima con f y le asignamos la clase en $I(f)/H$ determinada por el ideal de $\mathcal{O}_K$ generado por dicho representante. Componiendo obtenemos un isomorfismo $J_K/K^*W \cong I(f)/H$.

Vamos a determinar explícitamente este isomorfismo. Tomemos una clase $[\mathfrak{a}] \in \mathcal{H}_\mathcal{O}$, donde podemos suponer que $\mathfrak{a}$ es un ideal de $\mathcal{O}$ de norma prima con f.

Consideremos un primo p. Si $p \mid f$, entonces $p \nmid \mathrm{N}(\mathfrak{a}) = |\mathcal{O} : \mathfrak{a}| = |\mathcal{O}_K : \mathfrak{a}_K|$. Estos índices son, respectivamente, el valor absoluto del determinante de la matriz de coordenadas de una base de $\mathfrak{a}$ respecto a una base de $\mathcal{O}$ y el análogo con $\mathfrak{a}_K$ y $\mathcal{O}_K$. Por consiguiente, $\mathcal{O}_p = \mathfrak{a}_p$ y $(\mathcal{O}_K)_p = (\mathfrak{a}_K)_p$, de donde a su vez deducimos que podemos expresar $\mathfrak{a} = s\mathcal{O}$ para un $s \in J_K$ tal que $s_p = 1$ siempre que $p \mid f$, y entonces $(\mathfrak{a}_K)_p = s_p(\mathcal{O}_K)_p$ para tales primos.

Si $p \nmid f$, entonces $p \nmid |\mathcal{O}_K : \mathcal{O}| = |\mathfrak{a}_K : \mathfrak{a}|$, luego $(\mathcal{O}_K)_p = \mathcal{O}_p$, $(\mathfrak{a}_K)_p = \mathfrak{a}_p$ y, como $\mathfrak{a}_p = s_p\mathcal{O}_p$, también $(\mathfrak{a}_K)_p = s_p(\mathcal{O}_K)_p$. En definitiva, tenemos a la vez $\mathfrak{a} = s\mathcal{O}$ y $\mathfrak{a}_K = s\mathcal{O}_K$.

De aquí deducimos que toda clase de J_K/K^*W es de la forma $[s]$ para un cierto $s \in J_K$ tal que $s\mathcal{O}_K = (s) \in I(f)$ y de modo que la imagen de $[s]$ en $I(f)/H$ es precisamente $[(s)]$. En definitiva, el isomorfismo $J_K/K^*W \cong I(f)/H$ es el inducido por el epimorfismo natural $J_K/K^* \longrightarrow I(f)/H$, lo que nos permite concluir que H es el grupo de clases módulo f correspondiente a K^*W, así como que $K(j(E))$ es el cuerpo de clases de $\mathcal{O}$.

Ahora ya podemos generalizar fácilmente el teorema 13.9. Únicamente necesitamos una definición adicional: si $E/\mathbb{C}$ es una curva elíptica cuyo anillo de multiplicaciones complejas es un orden $\mathcal{O}$ de un cuerpo cuadrático imaginario K y M es un módulo completo de K cuyo anillo de coeficientes es $\mathcal{O}$, podemos definir la curva $M * E$ como una curva elíptica isomorfa a $\mathbb{C}/M^{-1}R$, donde R es un módulo completo de K tal que $\mathbb{C}/R \cong E$ y M^{-1} es el módulo inverso de M en el grupo de módulos completos de K con anillo de coeficientes $\mathcal{O}$. Como es habitual, la curva $M * E$ está definida salvo isomorfismo.

Teorema 13.26 *Sea $E/\mathbb{C}$ una curva elíptica cuyo anillo de multiplicaciones complejas sea un orden $\mathcal{O}$ de un cuerpo cuadrático imaginario K. Sea h el número de clases de $\mathcal{O}$ y f su conductor. Entonces*

a) $K(j(E))$ es el cuerpo de clases de $\mathcal{O}$.

b) $|\mathbb{Q}(j(E)) : \mathbb{Q}| = |K(j(E)) : K| = h$.

c) Para todo ideal $\mathfrak{a}$ de $\mathcal{O}$ de norma prima con f se cumple

$$\left(\frac{K(j(E))/K}{\mathfrak{a}_K}\right)(j(E)) = j(\mathfrak{a} * E).$$

d) Si $E_1, \dots, E_h$ son un sistema completo de representantes de las clases de isomorfía de curvas elípticas con anillo de multiplicaciones complejas $\mathcal{O}$, entonces $j(E_1), \dots, j(E_h)$ son los conjugados de $j(E)$ sobre K (y también sobre $\mathbb{Q}$).

Demostración: Ya hemos probado el apartado a) y, como consecuencia inmediata, la segunda igualdad del apartado b). En las condiciones de c), tenemos que $\mathfrak{a}$ tiene a $\mathcal{O}$ como anillo de coeficientes, y hemos probado que existe un elemento ideal $s \in J_K$ tal que $s\mathcal{O} = \mathfrak{a}$ y $(s) = s\mathcal{O}_K = \mathfrak{a}_K$. Así

$$\left(\frac{K(j(E))/K}{\mathfrak{a}_K}\right)(j(E)) = \left(\frac{K}{s}\right)(j(E)) = j(s * E).$$

Si $E \cong \mathbb{C}/R$, para un cierto módulo completo R de K, entonces es claro que $s^{-1}R = s^{-1}\mathcal{O}R = \mathfrak{a}^{-1}R$, luego $s * E = \mathfrak{a} * E$ y tenemos probado c).

Por c) sabemos que todo K-conjugado de $j(E)$ es de la forma $j(E_i)$, para un cierto i. Como en total ha de haber h conjugados, tenemos d) para K.

Veamos ahora la primera igualdad de b). Dado $\sigma \in \mathrm{Aut}(\mathbb{C})$, tenemos que E^σ tiene el mismo anillo de endomorfismos que E, luego $j(E)^\sigma = j(E^\sigma) = j(E_i)$ para algún i. Esto prueba que $|\mathbb{Q}(j(E)) : \mathbb{Q}| \le h = |K(j(E)) : K|$. La otra desigualdad es obvia. Ahora es inmediata la afirmación d) para $\mathbb{Q}$. ∎

Todavía podemos hacer algunas precisiones adicionales. En general, si M es un módulo completo de un cuerpo cuadrático imaginario K y representamos por $\overline{M}$ el módulo conjugado, las series que definen a $g_2(M)$ y $g_3(M)$ muestran que $g_2(\overline{M}) = \overline{g_2(M)}$ y $g_3(\overline{M}) = \overline{g_3(M)}$, de donde

$$j(\overline{M}) = \overline{j(M)}.$$

Si $\mathcal{O}$ es un orden de K tenemos que $\overline{\mathcal{O}} = \mathcal{O}$, luego $j(\mathcal{O}) \in \mathbb{R}$. Esto implica que $\mathbb{Q}(j(\mathcal{O})) = K(j(\mathcal{O})) \cap \mathbb{R}$, pues tenemos una inclusión y K tiene grado 2 sobre ambos cuerpos.

Un resultado no trivial de la teoría algebraica de números afirma que sólo existen nueve cuerpos cuadráticos imaginarios con número de clases $h = 1$, a saber, los cuerpos $\mathbb{Q}(\sqrt{d}\,)$ con

$$d = -1, \quad -2, \quad -3, \quad -7, \quad -11, \quad -19, \quad -43, \quad -67, \quad -163.$$

A partir de aquí es fácil concluir que sólo existen trece órdenes cuadráticos imaginarios con número de clases $h = 1$, que son los órdenes maximales de estos nueve cuerpos más los órdenes de conductor $f = 2$ en $\mathbb{Q}(\sqrt{-1}\,)$, $\mathbb{Q}(\sqrt{-3}\,)$, $\mathbb{Q}(\sqrt{-7}\,)$ y el orden de conductor $f = 3$ en $\mathbb{Q}(\sqrt{-3}\,)$.

Por el teorema anterior, una condición necesaria y suficiente para que exista una curva elíptica $E/\mathbb{Q}$ cuyo anillo de multiplicaciones complejas sea un orden cuadrático imaginario dado $\mathcal{O}$ es que el número de clases de $\mathcal{O}$ sea 1. En tal caso vemos además que $E/\mathbb{Q}$ es única salvo isomorfismo. Así pues, concluimos que existen exactamente trece clases de isomorfía de curvas elípticas $E/\mathbb{Q}$ con multiplicaciones complejas. La tabla siguiente contiene una ecuación minimal para cada una de ellas:

d	f	Ecuación minimal	Δ	j
-1	1	$Y^2 = X^3 + X$	2^6	$2^6 \cdot 3^3$
	2	$Y^2 = X^3 - 11X + 14$	2^9	$2^3 \cdot 3^3 \cdot 11^3$
-2	1	$Y^2 = X^3 + 4X^2 + 2X$	2^9	$2^6 \cdot 5^3$
-3	1	$Y^2 + Y = X^3$	3^3	0
	2	$Y^2 = X^3 - 15X + 22$	$2^8 \cdot 3^3$	$2^4 \cdot 3^3 \cdot 5^3$
	3	$Y^2 + Y = X^3 - 30X + 63$	3^5	$-2^{15} \cdot 3 \cdot 5^3$
-7	1	$Y^2 + XY = X^3 - X^2 - 2X - 1$	7^3	$-3^3 \cdot 5^3$
	2	$Y^2 = X^3 - 595X + 5586$	$2^{12} \cdot 7^3$	$3^3 \cdot 5^5 \cdot 17^3$
-11	1	$Y^2 + Y = X^3 - X^2 - 7X + 10$	11^3	-2^{15}
-19	1	$Y^2 + Y = X^3 - 38X + 90$	19^3	$-2^{15} \cdot 3^3$
-43	1	$Y^2 + Y = X^3 - 860X + 9707$	43^3	$-2^{18} \cdot 3^3 \cdot 5^3$
-67	1	$Y^2 + Y = X^3 - 7370X + 243528$	67^3	$-2^{15} \cdot 3^3 \cdot 5^3 \cdot 11^3$
-163	1	$Y^2 + Y = X^3 - 2174420X + 1234136692$	163^3	$-2^{18} \cdot 3^3 \cdot 5^3 \cdot 23^3 \cdot 29^3$

Apéndice A

La hipótesis de Riemann

Si K es un cuerpo de funciones algebraicas de género g sobre un cuerpo de constantes exacto k de q elementos, se define la *norma absoluta* de un divisor $\mathfrak{a}$ de K como $\mathrm{N}(\mathfrak{a}) = q^{\operatorname{grad}\mathfrak{a}}$. La función dseta de K es la función

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{\mathrm{N}(\mathfrak{a})^s},$$

donde $\mathfrak{a}$ recorre los divisores enteros de K. Los resultados siguientes son consecuencias relativamente sencillas del teorema de Riemann-Roch:[1]

Convergencia *La función $\zeta_K(s)$ converge en el semiplano $\operatorname{Re} s > 1$ a una función holomorfa.*

Racionalidad *Dicha función se extiende a una función meromorfa en todo el plano complejo dada por*

$$\zeta_K(s) = \frac{L(q^{-s})}{(1-q^{-s})(1-q^{1-s})},$$

donde $L(x) \in \mathbb{Z}[x]$ es un polinomio de grado $2g$ tal que $L(0) = 1$ y $L(1)$ es el número de clases (de divisores de grado 0) de K.

Ecuación funcional *La función $\zeta_K(s)$ satisface la ecuación funcional*

$$q^{(g-1)s}\zeta_K(s) = q^{(g-1)(1-s)}\zeta_K(1-s), \qquad \text{para todo } s \in \mathbb{C}.$$

Producto de Euler *En el semiplano $\operatorname{Re} s > 1$, la función $\zeta_K(s)$ admite el desarrollo en producto infinito*

$$\zeta_K(s) = \prod_{\mathfrak{p}} \frac{1}{1 - \frac{1}{\mathrm{N}(\mathfrak{p})^s}}.$$

[1] Éstos y los demás resultados citados aquí sin prueba están demostrados en el capítulo IX de mi Geometría Algebraica.

De esta última propiedad se sigue que $\zeta_K(s)$ no se anula en el semiplano $\operatorname{Re} s > 1$, y por la ecuación funcional tampoco lo hace en el semiplano $\operatorname{Re} s < 0$. En otras palabras, sus ceros han de estar en la banda $0 \le \operatorname{Re} s \le 1$.

En este apéndice demostraremos el análogo en este contexto a la famosa hipótesis de Riemann, es decir:

Hipótesis de Riemann *Los ceros de la función $\zeta_K(s)$ están todos situados sobre la recta* $\operatorname{Re} s = 1/2$.

Si llamamos $\alpha_1, \dots, \alpha_{2g}$ a los inversos de los ceros del polinomio $L(x)$, entonces s es un cero de ζ_K si y sólo si $q^s = \alpha_i$, para cierto i (y todo α_i puede ponerse de esta forma, porque $L(0) = 1$). La hipótesis de Riemann equivale, pues, a que $|\alpha_i| = q^{1/2}$ para todo i.

Esto es especialmente interesante porque puede probarse que si llamamos N_n al número de primos de grado 1 de la extensión de constantes de K de grado n, entonces

$$N_n = q^n + 1 - \sum_{i=1}^{2g} \alpha_i^n,$$

por lo que la hipótesis de Riemann implica la estimación

$$|N_n - q - 1| \le 2gq^{n/2}.$$

De hecho, la hipótesis de Riemann es equivalente a la estimación

$$N_n = q^n + O(q^{n/2}),$$

que es precisamente lo que vamos a demostrar aquí.

Necesitaremos un resultado adicional, y es que un cuerpo de funciones algebraicas K cumple la hipótesis de Riemann si y sólo si la cumple una extensión finita de constantes de K o, lo que es lo mismo:

> *Para probar que un cuerpo cumple la hipótesis de Riemann podemos sustituirlo por una extensión finita de constantes.*

La demostración original es de André Weil, si bien nosotros vamos a seguir un argumento mucho más elemental debido a Enrico Bombieri.[2]

El punto de partida será representar el cuerpo K como el cuerpo de funciones racionales de una curva proyectiva regular C/k. Llamaremos $\phi : C \longrightarrow C$ a la aplicación de Frobenius de grado q, de modo que N_n es el cardinal de $C(k_n)$ (donde k_n es la extensión de grado n de k) o, equivalentemente, el número de puntos fijos de ϕ^n.

Notemos que ϕ puede verse también como la aplicación inducida por el automorfismo de Frobenius de la extensión $\overline{k}/k$ (dado por $\phi(\alpha) = \alpha^q$). No

[2]Aquí seguimos la exposición que figura en el apéndice de Rosen, M. *Number theory in function fields.* Springer, New York, 2002.

obstante, no hemos de confundir el $k(C)$-automorfismo $\phi : \overline{k}(C) \longrightarrow \overline{k}(C)$ que extiende a ϕ como elemento del grupo $G(\overline{k}/k)$, con el $\overline{k}$-monomorfismo $\bar{\phi} : \overline{k}(C) \longrightarrow \overline{k}(C)$ inducido por ϕ como aplicación regular entre curvas.

Fijemos un punto racional $P \in C(k)$ o, equivalentemente, un divisor primo $\mathfrak{p} \in k(C)$ de grado 1 (puede probarse que siempre existe al menos uno, pero no necesitaremos este hecho, pues si no existe todo lo que vamos a concluir se cumplirá trivialmente.)

Para cada $r \geq 0$ definimos $M_r = m(P^r)$, es decir, el conjunto de las funciones racionales en $\overline{k}(C)$ que tienen a lo sumo un polo en P de orden a lo sumo r. Vamos a demostrar algunas propiedades:

Teorema A.1 *Con la notación precedente, se cumple:*

a) $\dim M_{r+1} \leq \dim M_r + 1$.

b) $\dim M_r \leq r + 1$.

c) $\dim M_r \geq m - g + 1$, *y si* $r > 2g - 2$ *se da la igualdad.*

d) Si $f \in M_r$, *entonces* $\phi \circ f = (f^{\phi^{-1}})^q$.

e) $\phi \circ M_r \subset M_{rq}$.

f) $\dim M_r^{p^e} = \dim M_r$, *donde* $p = \operatorname{car} k$ *y* $e \geq 0$.

g) $\dim \phi \circ M_r = \dim M_r$.

Demostración: a) Si f y g tienen un polo de orden exactamente $m+1$ en el punto P (y ningún otro polo), entonces f/g tiene orden 0 en P, luego existe $\gamma \in \overline{k}$, $\gamma \neq 0$ tal que $v_P(f/g - \gamma) \geq 1$. Así, $f - \gamma g = g(f/g - \gamma) \in M_r$.

Esto implica que si $M_r \neq M_{r+1}$ y a una base de M_r le añadimos una función de $M_{r+1} \setminus M_r$, obtenemos una base de M_{r+1}.

b) se deduce de a) por inducción. Notemos que M_0 está formado por las funciones constantes, luego tiene dimensión 1.

c) Es una consecuencia inmediata del teorema de Riemann-Roch.

d) Llamemos $\lambda = \phi^{-1} \in G(\overline{k}/k)$. Tomemos una función $f \in M_r$ y un punto $Q \in C(\overline{k})$, $Q \neq P$. Tomemos dos formas del mismo grado F y G que definan a f en un entorno de $\phi(Q)$. Entonces

$$(\phi \circ f)(Q) = \frac{F(\phi(Q))}{G(\phi(Q))} = \frac{F^\lambda(Q)^q}{G^\lambda(Q)^q} = f^\lambda(Q)^q.$$

e) Notemos que si $f \in M_r$ entonces $f^{\phi^{-1}} \in M_r$, pues ϕ fija a P. Por el apartado anterior $\phi \circ f \in M_{rq}$.

f) Aquí $M_r^{p^e}$ representa el espacio formado por las funciones f^{p^e} con $f \in M_r$. Es claro que elevando a p^e los elementos de una base de M_r obtenemos una base de $M_r^{p^e}$.

g) La aplicación $M_r \longrightarrow \phi \circ M_r$ dada por $f \mapsto \phi \circ f$ es obviamente lineal y suprayectiva. Basta ver que es inyectiva. Ahora bien, si $\phi \circ f = \phi \circ g$, entonces $(f^{\phi^{-1}})^q = (g^{\phi^{-1}})^q$, luego $f^{\phi^{-1}} = g^{\phi^{-1}}$, luego $f = g$. ∎

Si A es un subespacio de M_r y B un subespacio de R_s, llamaremos AB al subespacio de M_{r+s} generado por los productos fg con $f \in A$ y $g \in B$.

Teorema A.2 *Si $lp^e < q$, entonces el epimorfismo natural*

$$M_l^{p^e} \otimes_{\overline{k}} (\phi \circ M_r) \longrightarrow M_l^{p^e} (\phi \circ M_r)$$

es un isomorfismo.

Demostración: Por la propiedad a) del teorema anterior podemos encontrar una base $f_1, \dots, f_t$ de M_r tal que $v_P(f_i) < v_P(f_{i+1})$, para $i = 1, \dots, t-1$. Entonces $\phi \circ f_i$ es una base de $\phi \circ M_r$, y cada elemento del producto tensorial se expresa de forma única como

$$\sum_{i=1}^{t} g_i^{p^e} \otimes (\phi \circ f_i), \qquad g_i \in M_l.$$

Si este elemento está en el núcleo del epimorfismo, entonces

$$\sum_{i=1}^{t} g_i^{p^e} (\phi \circ f_i) = 0,$$

y basta probar que en tal caso todos los g_i son nulos. Si alguno no lo es, sea j el menor índice posible. Entonces

$$g_j^{p^e} (\phi \circ f_j) = -\sum_{i=j+1}^{t} g_i^{p^e} (\phi \circ f_i).$$

Por consiguiente, teniendo en cuenta que $v_P(\phi \circ f) = q v_P(f^{\phi^{-1}}) = q v_P(f)$,

$$p^e v_P(g_j) + q v_P(f_j) \geq \min_{i>j} \{p^e v_P(g_i) + q v_P(f_i)\} \geq -p^e l + q v_P(f_{j+1}),$$

luego

$$p^e v_P(g_j) \geq -lp^e + q(v_P(f_{j+1}) - v_P(f_j)) \geq q - lp^e > 0.$$

Esto significa que g_j tiene un cero en P, pero está en M_l, luego no tiene polos fuera de P, luego no tiene ningún polo, luego $g_j = 0$, contradicción. ∎

Como consecuencia inmediata:

Teorema A.3 *Si $lp^e < q$ entonces*

$$\dim M_l^{p^e} (\phi \circ M_r) = (\dim M_l)(\dim M_r).$$

Ahora podemos demostrar más o menos "la mitad" de la hipótesis de Riemann:

Teorema A.4 *Supongamos que* $(g+1)^4 < q$ *y que* q *es una potencia par de la característica* p*. Entonces*

$$N_1 \leq q + 1 + (2g+1)\sqrt{q}.$$

DEMOSTRACIÓN: (Notemos que el teorema se cumpliría trivialmente si fuera $N_1 = 0$, aunque ya hemos comentado que este caso no puede darse.) Tomemos l y e tales que $lp^e < q$. Mantenemos la notación empleada en la prueba de A.2. Definimos

$$\delta : M_l^{p^e}(\phi \circ M_r) \longrightarrow M_l^{p^e} M_r$$

mediante

$$\delta\left(\sum_{i=1}^{t} g_i^{p^e}(\phi \circ f_i)\right) = \sum_{i=1}^{t} g_i^{p^e} f_i.$$

El teorema A.2 garantiza que δ es una aplicación $\overline{k}$-lineal bien definida. Supongamos que l, $r \geq g$. Entonces, la dimensión del dominio es

$$(\dim M_l)(\dim M_r) \geq (l-g+1)(r-g+1)$$

y la de $\operatorname{Im}\delta \subset M_{lp^e+r}$ es a lo sumo $lp^e + r - g + 1$. Por consiguiente, el núcleo de δ tiene dimensión mayor o igual que

$$(l-g+1)(m-g+1) - (lp^e + r - g + 1).$$

Supongamos que esta cantidad es > 0, en cuyo caso δ tiene núcleo no trivial. Sea

$$f = \sum_{i=1}^{t} g_i^{p^e}(\phi \circ f_i)$$

un elemento no nulo del núcleo. Si $Q \in C(k)$, $Q \neq P$, entonces

$$f(Q) = \sum_{i=1}^{t} g_i(Q)^{p^e} f_i(\phi(Q)) = \sum_{i=1}^{t} g_i(Q)^{p^e} f_i(Q) = 0.$$

Así pues, f se anula en los puntos de $C(k)$ salvo quizá en P. Ahora bien, todo elemento de $\phi \circ M_r$ es una potencia q-ésima (por A.1 d) y como $p^e < q$ concluimos que f es una potencia p^e-ésima. Esto significa que f tiene (contando multiplicidades) al menos $p^e(N_1 - 1)$ ceros, y como $\phi \circ M_r \subset M_{rq}$, el número de polos es a lo sumo $lp^e + rq$. Así pues, $p^e(N_1 - 1) \leq lp^e + rq$. De aquí llegamos a que

$$N_1 \leq 1 + l + rqp^{-e}. \tag{A.1}$$

Recordemos que esta desigualdad es válida bajo las hipótesis siguientes:

a) $lp^e < q$,

b) l, $r \geq g$,

c) $(l-g+1)(r-g+1) > lp^e + r - g + 1$.

Vamos a elegir l, r, e de modo que se cumplan estas hipótesis y $(A.1)$ se convierta en la desigualdad del enunciado.

Estamos suponiendo que $q = p^{2b}$. Tomamos $e = b$ y $r = p^b + 2g$. Nos falta elegir l para que se cumpla c). Notemos que con las elecciones precedentes c) se convierte en

$$(l - g)(p^b + g + 1) > lp^b,$$

o equivalentemente,

$$l > \frac{g}{g+1}p^b + g.$$

Tomamos como l el menor natural mayor que el miembro derecho, con lo que se cumplen b) y c). Ahora usaremos la hipótesis $(g+1)^4 < q$ para probar que también se cumple a).

En efecto, tenemos que $(q+1)^2 < p^b$, luego $gp^b + (g+1)^2 < (g+1)p^b$, luego

$$\frac{g}{g+1}p^b + g + 1 < p^b.$$

Por la elección de l tenemos que $l < p^b$, luego $lp^b < p^{2b} = q$. Finalmente sustituimos las definiciones de e, r y l en la desigualdad (A.1), recordando además que $l < p^b$:

$$N_1 < 1 + p^b + (p^b + 2g)p^b = q + 1 + (2g+1)\sqrt{q}.$$

■

Notemos que si K cumple las hipótesis del teorema anterior, también las cumple la extensión de constantes de grado n de K, por lo que en realidad tenemos que

$$N_n \leq q^n + O(q^{n/2}).$$

Consideremos ahora una extensión finita de Galois L de K cuyo cuerpo de constantes exacto siga siendo k. Podemos considerar $K = k(C)$, $L = k(C')$, donde C y C' son curvas proyectivas regulares definidas sobre k. La inclusión $K \subset L$ puede verse como una aplicación regular $C' \longrightarrow C$ definida sobre k. Sean $\overline{K} = \overline{k}K$, $\overline{L} = \overline{k}L$ y $G = G(L/K) \cong G(\overline{L}/\overline{K})$. Es claro que el automorfismo de Frobenius $\phi \in G(\overline{L}/L)$ se restringe al de $G(\overline{K}/K)$. Además, considerando $\phi \in G(\overline{L}/K)$ y $G \leq G(\overline{L}/K)$, tenemos que ϕ conmuta con G. Ello se debe a que $\overline{L} = L\overline{K}$, de modo que si $\sigma \in G$, $a \in L$, $b \in \overline{K}$, entonces

$$\sigma(\phi(ab)) = \sigma(a)\phi(b) = \phi(\sigma(ab)).$$

Llamemos T al conjunto de los primos de grado 1 de K (considerados como primos de $\overline{K}$). Así $N_1 = |T|$. Llamemos $\tilde{T}$ a los primos de $\overline{L}$ que dividen a los de T.

Si $\mathfrak{p} \in T$, entonces G actúa transitivamente sobre los primos de $\tilde{T}$ que dividen a $\mathfrak{p}$. Por otra parte ϕ fija a este conjunto, pues $\mathfrak{p}^\phi = \mathfrak{p}$. Así pues, para cada $\mathfrak{P} \in \tilde{T}$ ha de existir un $\sigma \in G$ tal que $\mathfrak{P}^\phi = \mathfrak{P}^\sigma$. Habrá tantas elecciones

posibles para σ como automorfismos fijen a $\mathfrak{P}$, es decir, tantos como el índice de ramificación $e(\mathfrak{P}/\mathfrak{p})$. En particular, si llamamos $\tilde{T}'$ al conjunto de los primos de $\tilde{T}$ no ramificados, tenemos una aplicación $\eta : \tilde{T}' \longrightarrow G$ dada por $\mathfrak{P} \mapsto \sigma$.

Llamaremos $\tilde{T}(\sigma)$ al conjunto de los primos $\mathfrak{P} \in \tilde{T}'$ tales que $\eta(\mathfrak{P}) = \sigma$ y $N_1(\sigma, \overline{L}/\overline{K})$ al cardinal de $\tilde{T}(\sigma)$.

Cada primo $\mathfrak{p} \in T$ no ramificado en L es divisible entre $|G|$ primos de $\tilde{T}'$, luego $|\tilde{T}| = |G|N_1 + O(1)$, donde el error $O(1)$ depende del número de primos ramificados en $\overline{L}/\overline{K}$, pero no de q (es decir, si sustituimos k por una extensión finita la cota $O(1)$ sigue siendo la misma). Por otra parte,

$$\tilde{T}' = \bigcup_{\sigma \in G} \tilde{T}(\sigma),$$

y la unión es disjunta, luego

$$\sum_{\sigma \in G} N_1(\sigma, \overline{L}/\overline{K}) = |G|N_1 + O(1). \tag{A.2}$$

Representaremos por $\tilde{g}$ el género de $\overline{L}$. Ahora necesitamos una variante del teorema A.4.

Teorema A.5 *Con la notación precedente, supongamos que q es una potencia par de p, que $(\tilde{g}+1)^4 < q$ y sea $\sigma \in G$. Entonces*

$$N_1(\sigma, \overline{L}/\overline{K}) \leq q + 1 + (2\tilde{g}+1)\sqrt{q}.$$

Demostración: Podemos suponer que existe un punto $P \in C'(k)$ y definimos $M_r = m(P)$. Sea

$$\delta_\sigma : M_l^{p^e}(\phi \circ M_r) \longrightarrow M_r^{p^e}(\sigma \circ M_r)$$

la aplicación dada por

$$\delta_\sigma\left(\sum_{i=1}^{t} g_i^{p^e}(\phi \circ f_i)\right) = \sum_{i=1}^{t} g_i^{p^e}(\sigma \circ f_i).$$

Aquí usamos la notación de la prueba del teorema A.2, el cual justifica que δ_σ está bien definida (suponiendo $lp^e < q$).

Observemos que si $f \in M_r$, entonces $\sigma \circ f \in m((P^{\sigma^{-1}})^r)$, luego la imagen de δ_σ está contenida en $m(P^{lp^e}(P^{\sigma^{-1}})^r)$ y la dimensión de este espacio es a lo sumo $lp^e + r - g + 1$.

Bajo las mismas hipótesis que en el teorema A.4 podemos obtener un elemento no nulo f del núcleo de δ_σ, sólo que ahora se anula únicamente sobre los puntos de $\tilde{T}(\sigma)$ (es decir, los puntos $Q \in C'(k)$ tales que $\phi(Q) = \sigma(Q)$) distintos de P. Esto nos lleva a la misma conclusión pero cambiando N_1 por $N_1(\sigma, \overline{L}/\overline{K})$ y g por $\tilde{g}$. ∎

Ahora veremos el argumento que nos permite invertir la desigualdad:

Teorema A.6 *Bajo las hipótesis del teorema anterior, para cada $\sigma \in G$, se cumple*

$$q + 1 + |G|(N_1 - q - 1) + O(\sqrt{q}) \leq N_1(\sigma, \overline{L}/\overline{K}).$$

DEMOSTRACIÓN: Por el teorema anterior

$$0 \leq q + 1 + (2\tilde{g} + 1)\sqrt{q} - N_1(\sigma, \overline{L}/\overline{K}).$$

Sumamos sobre σ y usamos (A.2):

$$\begin{aligned} 0 &\leq \sum_{\sigma \in G} (q + 1 + (2\tilde{g} + 1)\sqrt{q} - N_1(\sigma, \overline{L}/\overline{K})) \\ &\leq |G|(q + 1 + (2\tilde{g} + 1)\sqrt{q}) - |G|N_1 + O(1). \end{aligned}$$

Como cada sumando es ≥ 0, ha de ser

$$q + 1 + (2\tilde{g} + 1)\sqrt{q} - N_1(\sigma, \overline{L}/\overline{K}) \leq |G|(q + 1 + (2\tilde{g} + 1)\sqrt{q}) - |G|N_1 + O(1),$$

de donde

$$q + 1 + |G|(N_1 - q - 1) - (|G| - 1)(2\tilde{g} + 1)\sqrt{q} + O(1) \leq N_1(\sigma, \overline{L}/\overline{K}).$$

De aquí obtenemos la desigualdad del enunciado. ∎

Ahora ya podemos probar la hipótesis de Riemann bajo ciertas condiciones:

Teorema A.7 *Sea K un cuerpo de funciones algebraicas de género g sobre un cuerpo de constantes exacto k de q elementos. Supongamos que q es una potencia par de la característica p y que $(g + 1)^4 < q$. Supongamos así mismo que existe $x \in K$ tal que $K/k(x)$ es separable y la clausura normal L de $k(x)$ sobre K tiene a k como cuerpo de constantes exacto. Entonces K cumple la hipótesis de Riemann.*

DEMOSTRACIÓN: Por el teorema A.4 (ver la observación posterior) tenemos que $N_1 \leq q + O(\sqrt{q})$. Sea $G = G(\overline{L}/\overline{k}(x))$ y $H = G(\overline{L}/\overline{K})$. Notemos que $k(x)$ tiene exactamente $q + 1$ divisores primos de grado 1. Para cada $\sigma \in G$, el teorema anterior aplicado a la extensión $\overline{L}/\overline{k}(x)$ nos da

$$q + O(\sqrt{q}) \leq N_1(\sigma, \overline{L}/\overline{k}(x)).$$

Enseguida veremos que si $\tau \in H$ entonces $N_1(\tau, \overline{L}/\overline{k}(x)) = N_1(\tau, \overline{L}/\overline{K})$. Aceptándolo de momento, sumamos en $\tau \in H$ y, usando (A.2), obtenemos

$$|H|q + O(\sqrt{q}) \leq \sum_{\tau \in H} N_1(\tau, \overline{L}/\overline{K}) = |H|N_1 + O(1),$$

de donde $q + O(\sqrt{q}) \leq N_1$. Ahora observamos que si K cumple las hipótesis del teorema, lo mismo vale para cualquier extensión finita de constantes de K y la cota del error $O(\sqrt{q})$ no depende de q, luego en realidad hemos probado que

$$q^n + O(q^{n/2}) \leq N_n,$$

y uniendo las dos desigualdades tenemos la relación $N_n = q^n + O(q^{n/2})$, que, según hemos comentado, equivale a la hipótesis de Riemann.

Falta probar que, en efecto, si $\tau \in H$ entonces $N_1(\tau, \overline{L}/\overline{k}(x)) = N_1(\tau, \overline{L}/\overline{K})$.

Sea $\mathfrak{P}$ un divisor primo de $\overline{L}$ que divida a un primo de grado 1 $\mathfrak{p}$ de $k(x)$ y tal que $\mathfrak{P}^\phi = \mathfrak{P}^\tau$. Basta probar que $\mathfrak{P}$ divide a un primo de grado 1 de K. Sea $\mathfrak{p}'$ el primo de $\overline{K}$ divisible entre $\mathfrak{P}$. Claramente $\mathfrak{p}'^\phi = \mathfrak{p}'^\tau = \mathfrak{p}'$, pero esto significa que el punto de la curva C asociado a $\mathfrak{p}'$ está en $C(k)$ (porque lo fija la aplicación de Frobenius). Así pues, $\mathfrak{p}'$ es un divisor primo de grado 1 de K.

■

Para terminar la demostración sólo hemos de ver que todo cuerpo K tiene una extensión finita de constantes que satisface las hipótesis del teorema anterior. Ahora bien, para cada $n \geq 1$, llamemos k_n a la extensión de grado n de k. Tomemos n suficientemente grande para que $(g+1)^4 < q^n$. Podemos elegir n par y así q^n es una potencia par de p. Existe $x \in k_n K$ tal que $k_n K/k_n(x)$ es separable. Sea L la clausura normal de $k_n(x)$ sobre K y sea k_m el cuerpo de constantes exacto de L. Entonces $n \mid m$, luego q^m sigue siendo una potencia par de p y es claro que L sigue siendo la clausura normal de $k_m(x)$ sobre $k_m K$. Así pues, $k_m K$ cumple el teorema anterior, y esto termina la prueba.

Apéndice B

Operadores de Hecke

En este apéndice demostraremos el carácter multiplicativo de la función τ de Ramanujan. Para ello introduciremos los llamados operadores de Hecke del grupo modular.

Una *correspondencia* entre dos conjuntos S y S' es, en el sentido usual de la teoría de conjuntos, un subconjunto $Z \subset S \times S'$. Podemos pensar que Z determina un criterio para asignar a cada elemento de S algunos elementos de S' (tal vez ninguno). En particular, toda aplicación $f : S \longrightarrow S'$ es una correspondencia, con la peculiaridad de que asigna un único elemento de S' a cada elemento de S.

Una *correspondencia finita* entre S y S' es una correspondencia Z tal que para todo $s \in S$ el conjunto $Z[s] = \{s' \in S' \mid (s, s') \in Z\}$ es finito, es decir, una correspondencia que a cada elemento de S le asigna un número finito (tal vez ninguno) de elementos de S'. Toda aplicación es una correspondencia finita.

Las correspondencias se pueden componer igual que las aplicaciones: dadas $Z \subset S \times S'$ y $Z' \subset S' \times S''$, entonces

$$Z \circ Z' = \{(s, s'') \in S'' \mid \text{existe } s' \in S' \text{ tal que } (s, s') \in Z,\ (s', s'') \in Z'\}.$$

La composición de correspondencias finitas es de nuevo una correspondencia finita. Ahora vamos a definir la noción de correspondencia finita algebraica, que generaliza la noción de correspondencia finita (conjuntista) a un marco algebraico más adecuado para nuestros fines.

Definición B.1 Una *correspondencia finita algebraica* entre dos conjuntos S y S' es un homomorfismo de módulos $Z : \langle S\rangle_{\mathbb{Z}} \longrightarrow \langle S'\rangle_{\mathbb{Z}}$ entre los $\mathbb{Z}$-módulos libres generados por S y S'.

Ciertamente, esta definición generaliza a la noción conjuntista de correspondencia finita, pues si Z es una correspondencia en el sentido conjuntista, podemos identificarla con el homomorfismo dado por

$$Z(s) = \sum_{s' \in Z[s]} s'.$$

No obstante, la noción algebraica es más general, pues para una correspondencia arbitraria Z y cada $s \in S$ tenemos que

$$Z(s) = \sum_{s' \in S'} n_{ss'} s', \qquad n_{ss'} \in \mathbb{Z},$$

con lo que podemos hablar de la multiplicidad $n_{ss'}$ de cada imagen s' de un elemento $s \in S$. Es decir, una correspondencia algebraica no sólo asigna a cada elemento $s \in S$ varios elementos de S', sino que a cada imagen de s le asigna una multiplicidad.

El conjunto $C(S, S')$ de todas las correspondencias de S en S' tiene una estructura obvia de $\mathbb{Z}$-módulo con la suma definida puntualmente. En el conjunto $C(S) = C(S, S)$ tenemos además la composición de homomorfismos, que extiende a la composición conjuntista de correspondencias, con lo que $C(S)$ resulta ser un anillo, el *anillo de las correspondencias finitas en S.*

A continuación observamos que las formas modulares respecto a $\mathrm{LE}(2, \mathbb{Z})$ pueden verse como funciones definidas sobre retículos en lugar de como funciones en H. En efecto, consideremos una forma modular f de grado $2k$. Podemos verla como una función sobre el conjunto S de todos los retículos complejos mediante

$$f(\langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}) = \omega_1^{-2k} f(\omega_2/\omega_1). \tag{B.1}$$

Esta definición no depende de la elección de la base, pues si ω_1', ω_2' es otra base del mismo retículo, entonces $\omega_1' = d\omega_1 + c\omega_2$, $\omega_2' = b\omega_1 + a\omega_2$, donde la matriz de coeficientes α está en $\mathrm{LE}(2, \mathbb{Z})$. Por consiguiente, llamando $\tau = \omega_2/\omega_1$,

$$f(\langle \omega_1', \omega_2' \rangle_{\mathbb{Z}}) = (c\omega_2 + d\omega_1)^{-2k} f\left(\frac{a\omega_2 + b\omega_1}{c\omega_2 + d\omega_1}\right)$$

$$= \omega_1^{-2k}(c\tau + d)^{-2k} f\left(\frac{a\tau + b}{c\tau + d}\right) = \omega_1^{-2k} f(\omega_2/\omega_1) = f(\langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}).$$

Es claro que, como función en S, la función f cumple

$$f(\lambda L) = \lambda^{-2k} f(L), \qquad \lambda \in \mathbb{C} \setminus \{0\}. \tag{B.2}$$

Recíprocamente, toda función f en S que cumpla esta relación de homogeneidad puede obtenerse mediante (B.1) a partir de una función en el semiplano H dada por

$$f(\tau) = f(\langle 1, \tau \rangle_{\mathbb{Z}}).$$

Esta función no es necesariamente holomorfa, pero cumple la relación de invarianza de las funciones modulares:

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = f\left(\left\langle 1, \frac{a\tau + b}{c\tau + d} \right\rangle_{\mathbb{Z}}\right)$$

$$= (c\tau + d)^{2k} f(\langle c\tau + d, a\tau + b \rangle_{\mathbb{Z}}) = (c\tau + d)^{2k} f(\tau).$$

Así pues, si f es una forma modular de grado $2k$, podemos verla como una función $f : S \longrightarrow \mathbb{C}$, o también como un homomorfismo de $\mathbb{Z}$-módulos $f : \langle S \rangle_{\mathbb{Z}} \longrightarrow \mathbb{C}$.

Definición B.2 Sea S el conjunto de todos los retículos complejos. Para cada natural $n \geq 1$ llamamos $T(n) \in C(S)$ a la correspondencia finita dada por

$$T(n)(L) = \sum_{|L:L'|=n} L'.$$

En otros términos, $T(n)$ asigna a cada retículo L el conjunto de todos sus subretículos de índice n. Este conjunto es finito, pues necesariamente se ha de cumplir $nL \leq L' \leq L$ y $L/nL \cong (\mathbb{Z}/n\mathbb{Z})^2$, luego hay tantos subretículos L' como subgrupos tiene el grupo $(\mathbb{Z}/n\mathbb{Z})^2$.

Para cada $\lambda \in \mathbb{C}^*$ consideramos también la correspondencia $R_\lambda \in C(S)$ dada por $R_\lambda(L) = \lambda L$.

Las correspondencias $T(n)$ y R_λ son elementos del anillo $C(S)$, donde el producto es la composición de homomorfismos. El teorema siguiente recoge algunas propiedades básicas:

Teorema B.3 *Se cumple:*

a) $R_\lambda R_\mu = R_{\lambda\mu}$,

b) $R_\lambda T(n) = T(n) R_\lambda$,

c) $T(m)T(n) = T(mn)$, *si* $(m, n) = 1$,

d) $T(p)T(p^n) = T(p^{n+1}) + pR_pT(p^{n-1})$, *para todo primo* p *y* $n \geq 1$.

Demostración: a) y b) son obvias. Respecto a c), observemos que

$$T(n)(T(m)(L)) = \sum_{|L:L'|=m} \sum_{|L':L''|=n} L''.$$

Basta probar que siempre que $|L : L''| = mn$ existe un único retículo intermedio $L'' \leq L' \leq L$ tal que $|L : L'| = m$ y $|L' : L''| = n$. Equivalentemente, hay que probar que el grupo L/L'' tiene un único subgrupo de orden n. Esto es un hecho general: todo grupo abeliano de orden mn con $(m, n) = 1$ tiene un único subgrupo de orden n (el formado por los elementos de orden divisor de n).

Para probar d) observamos que

$$T(p^n)(T(p)(L)) = \sum_{|L:L'|=p^{n+1}} a_{L'} L',$$

donde $a_{L'}$ es el número de retículos $L' \leq L'' \leq L$ tales que $|L : L''| = p$, $|L'', L'| = p^n$. Por otra parte,

$$T(p^{n+1})(L) = \sum_{|L:L'|=p^{n+1}} L',$$

$$pT(p^{n-1})(R_p(L)) = pT(p^{n-1})(pL) = p \sum_{|L:L'|=p^{n+1}} b_{L'} L',$$

donde

$$b_{L'} = \begin{cases} 1 & \text{si } L' \subset pL, \\ 0 & \text{en otro caso.} \end{cases}$$

(Notemos que $|L : pL| = p^2$.)

Hemos de demostrar que $a_{L'} = pb_{L'} + 1$. Distingamos dos casos:

a) Si $L' \not\subset pL$, hemos de ver que $a_{L'} = 1$, es decir, que hay un único retículo $L' \le L'' \le L$ de índice p en L. Un tal L'' ha de cumplir $pL \le L'' \le L$. Entonces, $(pL + L')/pL$ y L''/pL son dos grupos de orden p y uno está contenido en el otro, luego coinciden. Esto prueba la unicidad. La existencia es obvia.

b) Si $L' \subset pL$ hemos de ver que hay exactamente $p+1$ retículos $L' \le L'' \le L$ de índice p en L. Estos retículos cumplen, de hecho, $L' \le pL \le L'' \le L$ y se corresponden con los subgrupos de orden p en $L/pL \cong (\mathbb{Z}/p\mathbb{Z})^2$, que son $p+1$, como es fácil comprobar. ∎

De este teorema se desprende que las correspondencias $T(n)$ dependen polinómicamente de las correspondencias $T(p)$ y R_p, donde p recorre los números primos. Como éstas conmutan entre sí, concluimos que todos los operadores $T(n)$ y R_λ conmutan entre sí.

Ahora probamos un resultado técnico que necesitaremos para relacionar las correspondencias $T(n)$ con las funciones modulares:

Teorema B.4 *Sea $L = \langle \omega_1, \omega_2 \rangle_{\mathbb{Z}}$ un retículo complejo y $n \ge 1$ un número natural. Para cada matriz*

$$\alpha = \begin{pmatrix} a & 0 \\ b & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Z}), \qquad ad = n, \quad a \ge 1, \quad 0 \le b < d,$$

sea $L_\alpha = \langle \omega_1', \omega_2' \rangle_{\mathbb{Z}}$, con $\omega_1' = d\omega_1, \quad \omega_2' = b\omega_1 + a\omega_2$. Entonces, los retículos L_α son distintos dos a dos y son todos los subretículos de L de índice n.

Demostración: Es claro que cada L_α tiene índice n en L. Si $L' \le L$ es un retículo de índice n, definimos $Y_1 = L/(L' + \langle \omega_1 \rangle_{\mathbb{Z}})$, $Y_2 = \langle \omega_1 \rangle_{\mathbb{Z}} / (L' \cap \langle \omega_1 \rangle_{\mathbb{Z}})$, que son dos grupos cíclicos generados por las clases de ω_2 y ω_1 respectivamente. Sean a y d sus órdenes. La sucesión exacta

$$0 \longrightarrow Y_2 \longrightarrow L/L' \longrightarrow Y_1 \longrightarrow 0$$

prueba que $ad = n$.

Si $\omega_1' = d\omega_1$, tenemos que $\omega_1' \in L'$. Por otra parte, existe un $\omega_2' \in L'$ tal que $\omega_2' \equiv a\omega_2 \pmod{\langle \omega_1 \rangle_{\mathbb{Z}}}$. Digamos que $\omega_2' = b\omega_1 + a\omega_2$. Podemos exigir que $0 \le b < d$ y entonces ω_2' y b quedan completamente determinados. Se cumple que $L' = \langle \omega_1', \omega_2' \rangle_{\mathbb{Z}}$, pues ambos retículos tienen índice n en L.

Así pues, tenemos una aplicación que a cada retículo L' le asigna una matriz α de modo que $L' = L_\alpha$. Es claro que las aplicaciones $L' \mapsto \alpha$ y $\alpha \mapsto L'$ son mutuamente inversas, lo que prueba el teorema. ∎

Según hemos visto, toda función modular f de grado $2k$ puede verse como un homomorfismo de $\mathbb{Z}$-módulos $f : \langle S\rangle_{\mathbb{Z}} \longrightarrow \mathbb{C}$, donde S es el conjunto de todos los retículos complejos. Por otra parte tenemos definidas las correspondencias $T(n) : \langle S\rangle_{\mathbb{Z}} \longrightarrow \langle S\rangle_{\mathbb{Z}}$, que nos permiten formar la composición $T(n)f = T(n)\circ f$, cuya restricción a S cumple la relación (B.2):

$$T(n)f(\lambda L) = f(T(n)(\lambda L)) = f(\lambda T(n)(L))$$
$$= \lambda^{-2k} f(T(n)(L)) = \lambda^{-2k} T(n)f(L).$$

Esto nos permite a su vez ver a $T(n)f$ como una función definida sobre H que verifica la relación de invarianza de las formas modulares de grado $2k$. Vamos a ver que es, de hecho, una forma modular. Explícitamente, teniendo en cuenta el teorema B.4, vemos que

$$(T(n)f)(\tau) = f(T(n)(\langle 1,\tau\rangle_{\mathbb{Z}})) = f\Big(\sum_{a,b,d} \langle d, a\tau+b\rangle_{\mathbb{Z}}\Big)$$

$$= \sum_{a,b,d} d^{-2k} f\Big(\frac{a\tau+b}{d}\Big) = \frac{1}{n^k}\sum_{\alpha}(\alpha|_{2k}f)(\tau),$$

donde α recorre las matrices del teorema B.4.

Esta expresión muestra que $T(n)f$ es una función holomorfa en H. Falta probar que es holomorfa en ∞. Sea

$$f(\tau) = \sum_{m=0}^{\infty} c(m) e^{2\pi i m\tau}$$

la serie de Fourier de f. Entonces

$$(T(n)f)(\tau) = \sum_{a,b,d} d^{-2k} \sum_{m=0}^{\infty} c(m) e^{2\pi i m(a\tau+b)/d}.$$

Ahora bien,

$$\sum_{0\le b<d} e^{2\pi i m b/d} = \begin{cases} d & \text{si } d \mid m,\\ 0 & \text{si } d \nmid m.\end{cases}$$

Por consiguiente,

$$T(n)f(\tau) = \sum_{a,d} d^{-2k+1} \sum_{m'=0}^{\infty} c(m'd) e^{2\pi i m' a\tau}.$$

Llamando $m = m'a$, tenemos que

$$T(n)f(\tau) = \sum_{m=0}^{\infty}\Big(\sum_{a|(n,m)} (n/a)^{-2k+1} c(mn/a^2)\Big) e^{2\pi i m\tau}.$$

Esta expresión muestra que $T(n)f$ es holomorfa en ∞, así como que si $c_0 = 0$, entonces el primer coeficiente de la serie de $T(n)f$ también es nulo, luego $T(n)$ transforma formas parabólicas en formas parabólicas.

Para evitar denominadores conviene modificar ligeramente la definición de los operadores $T(n)$. Llamemos $T(n)^*$ a lo que hasta ahora hemos llamado $T(n)$.

Definición B.5 Sea M_{2k} el espacio de las formas modulares de grado $2k$. Para cada natural $n \geq 1$, definimos el *operador de Hecke* $T(n) : M_{2k} \longrightarrow M_{2k}$ como el dado por

$$T(n)f = n^{2k-1}T(n)^*f = n^{k-1}\sum_\alpha \alpha|_{2k} f,$$

donde α recorre las matrices del teorema B.4.

De este modo,

$$T(n)\Big(\sum_{m=0}^\infty c(m)e^{2\pi i\tau m}\Big) = \sum_{m=0}^\infty b(m)e^{2\pi i\tau m},$$

donde

$$b(m) = \sum_{a|(n,m)} a^{2k-1}c(mn/a^2).$$

(Se entiende que en el sumatorio $a \geq 1$.)

En particular, $b(0) = \sigma_{2k-1}(n)c(0)$ y $b(1) = c(n)$. De aquí se sigue —como ya habíamos comentado— que $T(n)$ se restringe a un operador $T(n) : M^0_{2k} \longrightarrow M^0_{2k}$ en el espacio de las formas parabólicas de grado $2k$.

Los operadores de Hecke heredan las propiedades de las correspondencias que los inducen:

Teorema B.6 *Los operadores de Hecke en el espacio M_{2k} conmutan entre sí y satisfacen las relaciones siguientes:*

a) $T(m)T(n) = T(mn)$, *si* $(m,n) = 1$,

b) $T(p^n)T(p) = T(p^{n+1}) + p^{2k-1}T(p^{n-1})$, *si p es primo y $n \geq 1$.*

Demostración: La conmutatividad se sigue inmediatamente de la definición de los operadores de Hecke y de la conmutatividad de las correspondencias que los inducen. Comprobamos b). La prueba de a) es similar y más sencilla. Si $f \in M_{2k}$, entonces

$$T(p^n)T(p)f = T(p)(T(p^n)f) = p^{(n+1)(2k-1)}T(p)^*(T(p^n)^*f) =$$

$$p^{(n+1)(2k-1)}(T(p)^* \circ T(p^n)^* \circ f) = p^{(n+1)(2k-1)}(T(p^{n+1})^* + pR_pT(p^{n-1})^*) \circ f$$

$$= T(p^{n+1})f + p^{4k-1}R_p \circ T(p^{n-1}) \circ f.$$

Ahora observamos que, para un retículo L, se cumple

$$(R_p \circ T(p^{n-1}) \circ f)(L) = f(p(T(p^{n-1})(L))$$

$$= p^{-2k}f(T(p^{n-1})(L)) = p^{-2k}(T(p^{n-1}) \circ f)(L).$$

Así pues,

$$T(p^n)T(p)f = T(p^{n+1})f + p^{2k-1}T(p^{n-1})f.$$

■

Ahora podemos entender la relación entre los operadores de Hecke y los coeficientes de Fourier de las formas modulares. Pensemos, por ejemplo, en el espacio M_{12}^0. Sabemos que tiene dimensión 1 y que está generado por la función discriminante $\Delta(\tau)$. Esto hace que Δ sea necesariamente un vector propio de todos los operadores de Hecke, es decir, que existen números complejos $\lambda(n)$ tales que $T(n)\Delta = \lambda(n)\Delta$. Vamos a ver que esto implica el carácter multiplicativo de la función de Ramanujan.

Teorema B.7 *Sea f una forma modular de grado $2k > 0$ cuyo desarrollo en serie de Fourier sea*

$$f(\tau) = \sum_{n=0}^{\infty} c(n)e^{2\pi i n\tau}.$$

Supongamos que f es simultáneamente un vector propio de todos los operadores de Hecke $T(n)$, es decir, tal que $T(n)f = \lambda(n)f$ para ciertos $\lambda(n) \in \mathbb{C}$. Entonces

a) $c(1) \neq 0$.

b) Para cada $n \geq 1$, se cumple $c(n) = \lambda(n)c(1)$.

c) Si $c(1) = 1$, entonces, para m, $n \geq 1$,

$$c(m)c(n) = \sum_{d|(m,n)} d^{2k-1} c\left(\frac{mn}{d^2}\right).$$

En particular, si $(m, n) = 1$ se cumple $c(mn) = c(m)c(n)$.

Demostración: El m-simo coeficiente de Fourier de $T(n)f$ es $\lambda(n)c(m)$, y según las observaciones posteriores a la definición B.5 es

$$\lambda(n)c(m) = \sum_{d|(n,m)} d^{2k-1} c\left(\frac{mn}{d^2}\right). \tag{B.3}$$

En particular, si $m = 1$, queda $\lambda(n)c(1) = c(n)$. Si fuera $c(1) = 0$ entonces sería $c(n) = 0$ para todo $n \geq 1$, luego f sería una constante, lo cual es absurdo (tendría que ser $k = 0$). Así pues, $c(1) \neq 0$. Esto prueba a) y b). Bajo la hipótesis de c) tenemos $\lambda(n) = c(n)$ y (B.3) se convierte en la fórmula del enunciado. ∎

Ejemplo Según la observación previa al teorema, esto se aplica a la función Δ y, por consiguiente, a la función τ de Ramanujan. Tenemos, pues, demostrada la relación:

$$\tau(m)\tau(n) = \sum_{d|(m,n)} d^{2k-1} \tau\left(\frac{mn}{d^2}\right).$$

En particular de aquí se desprende su carácter multiplicativo:

$$\tau(mn) = \tau(m)\tau(n), \qquad \text{si } (m, n) = 1,$$

así como una relación recurrente para calcular $\tau(p^n)$ (supuesto conocido $\tau(p)$):

$$\tau(p^n)\tau(p) = \tau(p^{n+1}) + p^{11}\tau(p^{n-1}).$$

∎

La teoría de los operadores de Hecke puede generalizarse a otros grupos de congruencias, y es una herramienta muy útil para el estudio de las formas modulares, especialmente para el estudio de las series de Dirichlet formadas con los coeficientes de una forma parabólica. Por ejemplo, para el caso concreto de la función Δ, puede probarse que la serie

$$L(\Delta, s) = \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}.$$

converge en el semiplano complejo $\operatorname{Re} s > 6.5$, así como que admite un desarrollo en producto de Euler de la forma

$$\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \prod_p \frac{1}{1 - \tau(p)p^{-s} + p^{11-2s}},$$

donde p recorre los números primos. Además, $L(\Delta, s)$ se prolonga analíticamente a todo el plano complejo y satisface la ecuación funcional

$$(2\pi)^{-s}\ \frac{1}{s}\,\Pi(s)L(\Delta, s) = (2\pi)^{s-12}\ \frac{1}{12-s}\,\Pi(12-s)L(\Delta, 12-s).$$

Bibliografía

[1] Apostol, T.M. *Modular Functions and Dirichlet Series in Number Theory.* Springer, New York, 1976.

[2] Cassels, J.W.S. *Diophantine equations with special reference to elliptic curves.* J. London Math. Soc. **41** (1966), 193–291.

[3] Connell, I. *Elliptic Curve Handbook.* Manuscrito, 1996.

[4] Dolgachev, I. *Lectures on Modular Forms.* Apuntes 1997.

[5] Knapp, A.W. *Elliptic Curves.* Princeton University Press, Princeton, New Jersey, 1992.

[6] Lang, S. *Diophantine Geometry.* John Wiley & Sons, New York, 1962.

[7] Lang, S. *Elliptic Functions.* Springer, New York, 1987.

[8] Milne, J.S. *Modular Functions and Modular Forms.* Apuntes, 1990.

[9] Milne, J.S. *Elliptic Curves.* Apuntes, 1996.

[10] Serre, J.P., *Cours d'arithmetique.* Presses universitaires de France, Paris, 1970.

[11] Shafarevich, I. R. *Basic Algebraic Geometry 1.* Springer, New York, 1994.

[12] Shimura, G. *Introduction to the Arithmetic Theory of Automorphic Functions.* Iwanami Shoten and Princeton University Press, 1971.

[13] Siegel, C. L. *Topics in Complex Function Theory.* (3 volúmenes) John Whiley & sons, New York, 1969, 1971, 1973.

[14] Silverman, J.H. *The Arithmetic of Elliptic Curves.* Springer, New York, 1986

[15] Silverman, J.H. *Advanced Topics in the Arithmetic of Elliptic Curves.* Springer, New York, 1994

Índice de Materias