El diablo de los números

Sección a cargo de

Javier Cilleruelo Mateo

Formas modulares y números casi enteros

por

Fernando Chamizo y Dulcinea Raboso

1. Introducción

Algunos lectores habrán visto en libros de divulgación el curioso cálculo

$$e^{\pi\sqrt{163}} = 262537412640768743{,}99999999999925007\ldots$$

del que se deduce que $e^{\pi\sqrt{163}}$ dista de un entero menos de 10^{-12}. A esta cantidad se le llama a veces *constante de Ramanujan*. En realidad Ramanujan nunca la mencionó (el nombre proviene de una broma del divulgador Martin Gardner [Gar]) pero sí trabajó con otras constantes relacionadas.

¿Hasta qué punto es esto una casualidad? Es difícil con tan pocos caracteres crear un ejemplo no trivial de un número que se parezca tanto a un entero y el lector ya imaginará, sin más que mirar el título, que alguna teoría relacionada con las formas modulares está detrás de este misterio. Jugando un poco con esta constante se obtienen variaciones suyas más feas pero más espectaculares, como por ejemplo

$$\sqrt[3]{e^{\pi\sqrt{163}} - 744} = 640319{,}999999999999999999999939031\ldots$$

que dista de un entero menos de 10^{-24}, y que tomando logaritmos implica que $\log(744 + 640320^3)/\sqrt{163}$ es indistinguible de π hasta el decimal 31º. Este tipo de aproximaciones de π son parte del contenido de [Ram]. Por rizar más el rizo, $e^{\pi\sqrt{163}} + 196884\, e^{-\pi\sqrt{163}}$ dista de un entero menos de $4\cdot 10^{-28}$ y, si restamos 744 y extraemos la raíz cúbica, la distancia al entero más cercano pasa a ser menor que $3\cdot 10^{-40}$.

De la extraña cercanía de la constante de Ramanujan a los enteros surge la inevitable pregunta: ¿por qué?, ¿a qué debemos este comportamiento? Y a medida que nos acercamos a la respuesta, lo que en un principio puede parecer una simple

curiosidad deja de ser tan simple ofreciéndonos a cambio nuevas posibles direcciones con las que apaciguar nuestra sed de conocimientos. Nuestro propósito aquí es dar una explicación para matemáticos no especialistas y utilizarla como pretexto para mostrar unas pinceladas de unos temas fascinantes relacionados de una u otra manera con las formas modulares.

Intentamos que este artículo sea autocontenido en lo esencial (con los conocimientos previos de un estudiante de matemáticas) pero dejando la posibilidad de ir más allá a través de indicaciones y referencias.

2. Retículos

En el desarrollo de la teoría que deseamos presentar vamos a tratar con retículos en $\mathbb{C}$ del tipo

$$\Lambda = \{n\omega_1 + m\omega_2 \ : \ n, m \in \mathbb{Z}\},$$

donde ω_1 y ω_2 son números complejos que apuntan en direcciones distintas, y diremos que son generadores de Λ. Es decir, nuestros retículos pueden ser visualizados como cristales bidimensionales o embaldosados cuyas teselas están hechas con paralelogramos.

Denotaremos el retículo anterior como $[\omega_1, \omega_2]$. Por razones técnicas conviene «orientar» estas presentaciones de los retículos pidiendo que el ángulo de ω_2 a ω_1 sea siempre menor que $180°$, o lo que es lo mismo que $\mathrm{Im}(\omega_1/\omega_2) > 0$.

Evidentemente los generadores no están determinados por el retículo, por ejemplo

$$\mathbb{Z}[i] = \{ni + m \ : \ n, m \in \mathbb{Z}\} = [i, 1] = [i, 1 + 2010i] = [7 + 5i, 3 + 2i] = \cdots$$

¿Cuándo dos pares de generadores dan lugar al mismo retículo? Muy sencillo, debe existir un «cambio de base» entre ellos, pero, para preservar que $n, m \in \mathbb{Z}$, la matriz de tal cambio de base debe operar en $\mathbb{Z}^2$. Las matrices enteras con inversa entera son las que tienen determinante ± 1 y si además conservan la orientación (de ahí la condición técnica anterior) quedan reducidas al grupo de matrices

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ : \ a, b, c, d \in \mathbb{Z}, \ ad - bc = 1 \right\}.$$

De esta forma $[\omega_1, \omega_2] = [\eta_1, \eta_2]$ si y sólo si

$$\begin{cases} a\omega_1 + b\omega_2 = \eta_1 \\ c\omega_1 + d\omega_2 = \eta_2 \end{cases} \tag{1}$$

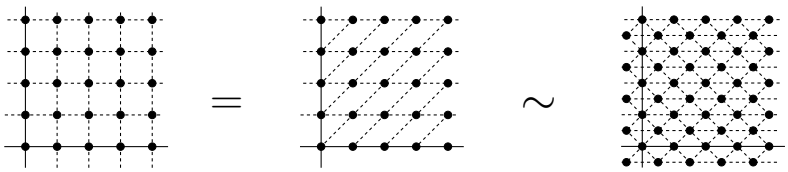
con a, b, c, d como en la definición anterior. Dicho de otro modo, dos vectores de generadores $\vec{v}, \vec{w} \in \mathbb{C}^2$ dan lugar al mismo retículo si y sólo si $\gamma\vec{v} = \vec{w}$ para cierta matriz $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Convengamos que dos retículos son equivalentes si son iguales salvo multiplicar por un número complejo. Geométricamente esto es girar la cabeza y mirar desde más lejos o más cerca. Esta relación será natural más adelante en el ámbito de las formas

cuadráticas. Es decir, $[\omega_1, \omega_2] \sim [\omega_1/\omega_2, 1]$ y módulo esta relación de equivalencia cada retículo es de la forma $[z, 1]$ con z un elemento del semiplano superior

$$\mathbb{H} = \{z \in \mathbb{C} \,:\, \mathrm{Im}(z) > 0\}.$$

Aun así la ambivalencia continúa y por ejemplo $[i, 1] \sim [\frac{1}{2} + \frac{i}{2}, 1]$, como muestra un sencillo dibujo.

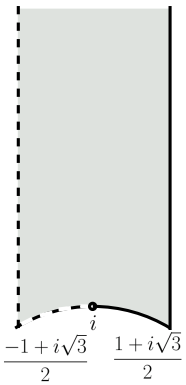


Retículo $[i, 1]$ Retículo $[i, 1 + i]$ Retículo $[\frac{1}{2} + \frac{i}{2}, 1]$

De hecho $[z, 1] \sim [w, 1]$ si y sólo si $\gamma(z) = w$ con $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ definiendo la acción de una matriz sobre un elemento de $\mathbb{H}$ de la forma obvia para que sea coherente con (1):

$$\gamma(z) := \frac{az + b}{cz + d} \qquad \text{para} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{y} \quad z \in \mathbb{H}.$$

Un punto fundamental es que $\gamma\big(\tau(z)\big) = (\gamma\tau)(z)$. Como en el álgebra lineal, componer es multiplicar matrices.



Pero, en la práctica, ¿qué cuentas tendríamos que hacer para saber si dos elementos de $\mathbb{H}$ generan el mismo retículo? Aplicando varias veces la traslación $T(z) = z + 1$, $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, o su inversa, conseguiremos llevar z a la banda $-1/2 < \mathrm{Re}(z) \leq 1/2$. Una vez allí, aplicando la inversión $S(z) = -1/z$, $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, si fuera necesario, pasará a la zona $|z| > 1$ o a $|z| = 1$ con $\mathrm{Re}(z) \geq 0$. Aplicando nuevas traslaciones e inversiones no es difícil probar que el proceso termina (de ahí deriva una prueba de que $\mathrm{SL}_2(\mathbb{Z})$ está generado por T y S), y en un número finito de pasos obtenemos un γ tal que $\gamma(z)$ está en el *dominio fundamental restringido* $\mathcal{F}^*$, que es el *dominio fundamental*

$$\mathcal{F} := \{z \in \mathbb{H} \,:\, |\,\mathrm{Re}(z)| \leq 1/2,\ |z| \geq 1\},$$

omitiendo la parte de la frontera con $\mathrm{Re}(z) < 0$.

3. Formas cuadráticas

Gauss dedicó gran parte de su obra maestra *Disquisitiones Arithmeticae* [Gau] a crear una teoría aritmética de formas cuadráticas (véase en [Ba 1] el contexto y un resumen del contenido de esta obra).

Las *formas cuadráticas binarias* (*primitivas*) de discriminante $-D$ son los elementos del conjunto

$$\mathcal{Q}_D = \{ax^2 + bxy + cy^2 \ : \ a \in \mathbb{Z}^+, \ \mathrm{mcd}(a,b,c) = 1, \ D = 4ac - b^2\}.$$

El discriminante de una forma no puede ser cualquier número ya que $D \equiv -b^2$ (mód 4), así las posibilidades se reducen a $-D \equiv 0, 1$ (mód 4). Nosotros consideraremos únicamente las formas definidas positivas, es decir, con $D > 0$.

Al factorizar $Q = ax^2 + bxy + cy^2 \in \mathcal{Q}_D$ en $\mathbb{C}[x,y]$ se obtiene

$$Q = a(x + z_Q y)(x + \bar{z}_Q y) \qquad \text{con} \qquad z_Q = \frac{b + i\sqrt{D}}{2a} \in \mathbb{H}.$$

A la forma cuadrática Q le asignamos el número complejo z_Q y el retículo $[z_Q, 1]$.

C. F. Gauss

Supongamos que Q representa ciertos enteros, esto es, que todos ellos admiten una expresión del tipo $Q(x_0, y_0)$ para ciertos $x_0, y_0 \in \mathbb{Z}$; entonces está claro que $Q \circ \gamma$ representa los mismos enteros para cualquier cambio de variable lineal tal que él y su inversa estén definidos sobre $\mathbb{Z}$, es decir, para cualquier $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ (de nuevo se excluye el determinante -1 por razones técnicas).

Con esta motivación, diremos que dos formas $Q_1, Q_2 \in \mathcal{Q}_D$ son *equivalentes*, y escribiremos $Q_1 \sim Q_2$ si son iguales salvo un cambio de variable en $\mathrm{SL}_2(\mathbb{Z})$. Es fácil ver que esto ocurre si y sólo si sus retículos son equivalentes en el sentido de la sección anterior. Por otro lado, sabíamos que cualquier retículo es equivalente a exactamente uno de la forma $[z, 1]$ con $z \in \mathcal{F}^*$ y, traduciendo esto al lenguaje de las formas cuadráticas, cuando $z = (b + i\sqrt{D})/(2a)$ se tiene que

$$\mathcal{H}_D = \{ax^2+bxy+cy^2 \in \mathcal{Q}_D \ : \ D = 4ac-b^2 \ \text{ con } -a < b \le a < c \ \text{ ó } \ 0 \le b \le a = c\}$$

es un conjunto completo de representantes de $\mathcal{Q}_D/\!\sim$. A su cardinal se le llama *número de clases* y se denota con $h(-D)$. No es difícil deducir que $h(-D) < \infty$.

Nota histórica: Gauss impuso que b fuera par en su definición de formas cuadráticas, pero esta condición dificulta la relación con otros temas y no se contempla en la actualidad. Por otro lado, Gauss no siempre considera que los coeficientes sean coprimos ([Gau, Art. 226]) lo que da lugar a dos posibles definiciones del número de clases. La literatura moderna prefiere la indicada aquí, aunque la relación con la factorización en anillos de enteros conduce en muchas ocasiones a restringirse a los llamados *discriminantes fundamentales*, aquéllos para los que la condición de coprimalidad es superflua y por tanto ambas definiciones coinciden (esto ocurre por ejemplo cuando D es libre de cuadrados).

3.1. Número de clases uno

Para $D = 3$ o $D = 4$, jugando un poco con las condiciones impuestas a a, b, c en $\mathcal{H}_D$ tenemos que el número de clases es uno. Con un pequeño programa o a mano (¡Gauss calculó varios miles de números de clases!, [Gau, Art. 302]), llegaríamos a la siguiente lista con $h(-D) = 1$:

$$D = 3,\ 4,\ 7,\ 8,\ 11,\ 12,\ 16,\ 19,\ 27,\ 28,\ 43,\ 67,\ 163$$

(a menudo se omiten 12, 16, 27 y 28 por no ser discriminantes fundamentales).

La pregunta de si la lista es completa parece de naturaleza combinatoria elemental, pero no lo es en absoluto, y tiene una peculiar historia [Go], desde Gauss hasta nuestros días, en la que han participado temas tan dispares como la hipótesis de Riemann generalizada, curvas elípticas o aproximación diofántica. En 1952 K. Heegner, un profesor de instituto, probó que no había más discriminantes con esta propiedad, pero su prueba no recibió la atención merecida y el teorema fue redescubierto por H. Stark en 1967.

Veamos aquí un ejemplo rápido de número de clases uno en el contexto de las representaciones por una forma cuadrática.

Consideremos $Q = x^2 + xy + 2y^2 \in \mathcal{Q}_7$. Si -7 es residuo cuadrático módulo un primo p, existen a y b tales que $b^2 = -7 + 4ap$; entonces $Q' = ax^2 + bxy + py^2 \in \mathcal{Q}_7$ y se tiene $Q \sim Q'$ porque $h(-7) = 1$. En particular Q y Q' representan los mismos enteros. Notando que $Q'(0, 1) = p$ hemos probado en unas líneas un resultado en absoluto trivial:

$$-7 \text{ residuo cuadrático módulo } p \ \Rightarrow\ \exists x, y \in \mathbb{Z} \text{ tales que } p = x^2+xy+2y^2.$$

El argumento se puede invertir para probar el recíproco, salvo incluir el caso $p = 7$. Por otro lado la ley de reciprocidad cuadrática permite caracterizar cuándo -7 es residuo cuadrático en términos de congruencias, obteniéndose:

> Los primos p de la forma $x^2 + xy + 2y^2$ son exactamente $p = 7$ y los que verifican $p \equiv 1, 2, 4 \pmod 7$.

Con un poco más de esfuerzo y notando que $4(x^2 + xy + 2y^2) = (2x + y)^2 + 7y^2$, es posible caracterizar los enteros de la forma $x^2 + 7y^2$ y expresar el número de representaciones mediante una fórmula sencilla.

Un análisis similar se puede llevar a cabo siempre que $h(-D) = 1$.

3.2. Más allá del número de clases uno

Cuando el número de clases es mayor que uno, en general no es posible dar una fórmula sencilla para el número de representaciones de un entero por una forma cuadrática específica, sino que debemos considerar todas las de $\mathcal{H}_D$. A veces ocurre una casualidad y las formas en diferentes clases representan números en diferentes clases de congruencias, permitiendo separar el número de representaciones. Todo esto tiene su explicación dentro de la división que hizo Gauss de las formas cuadráticas en *géneros* ([Gau, Art. 230, 231]). Ya Gauss y Euler (con otra motivación) conocían

que había al menos 65 casos buenos, con fórmulas cerradas para el número de representaciones mediante $x^2 + Dy^2$, los llamados *números idóneos*. Hoy en día se sabe que hay a lo más otro ejemplo y se conjetura que tal nuevo ejemplo no existe, pero nadie ha conseguido probarlo.

Una visión más amplia del concepto de número de clases la da la teoría de ideales. En el caso de un cuerpo cuadrático imaginario los ideales son retículos generados por un par de enteros algebraicos. Las formas cuadráticas correspondientes a estos retículos tienen el mismo discriminante y el número de clases nos aporta información sobre la factorización única en el anillo de enteros del cuerpo cuadrático. Así, cuando el número de clases es uno, se tiene factorización única y a medida que crece estamos más alejados de ella (véase la última sección).

4. Por fin, formas modulares

Esencialmente una función modular es una función homogénea (de grado cero) en los retículos considerados en la sección 2. Es decir, verificando $f([\omega_1, \omega_2]) = f([\alpha\omega_1, \alpha\omega_2])$. Según hemos visto, tal función puede considerarse que depende sólo de $z = \omega_1/\omega_2 \in \mathbb{H}$ y debe cumplir $f(\gamma(z)) = f(z)$ para todo $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Equivalentemente f es invariante por T y S, en particular es 1-periódica en z y por tanto tiene un desarrollo de Fourier, que exigiremos sea absolutamente convergente y de la forma

$$f(z) = \sum_{n=-N}^{\infty} a_n e^{2\pi i n z} \tag{2}$$

para algún $N \in \mathbb{Z}$ (suponemos $a_{-N} \neq 0$). En cierto sentido esto significa que permitimos que esta función holomorfa tenga un «polo» en $i\infty$. Definimos por tanto una *función modular* como una función holomorfa en $\mathbb{H}$ de la forma (2) que satisface además $f(-1/z) = f(z)$. Nótese que es invariante por T y por S, y consecuentemente por la acción de $\mathrm{SL}_2(\mathbb{Z})$.

Una bella y sencilla aplicación del principio del argumento ([Ap, § 2.4]) permite probar que si f no es constante entonces N es el número de soluciones $z \in \mathcal{F}^*$ de $f(z) = c$ para cualquier $c \in \mathbb{C}$, con cierto convenio para las multiplicidades de i y de $(1 + i\sqrt{3})/2$, en particular $N \in \mathbb{Z}^+$ y se deduce un *teorema de Liouville modular*:

Las únicas funciones modulares acotadas son las constantes.

Para simplificar hemos divergido de la terminología más aceptada [Se] llamando aquí función modular a un sustituto de las formas modulares de peso cero. Las *formas modulares de peso k* responden a la definición anterior con $N = 0$ pero ahora exigiendo $f(-1/z) = z^k f(z)$, o lo que es lo mismo, pidiendo que la función sea homogénea de grado k en los retículos. Añadiendo la función nula, las formas modulares tienen estructura de espacio vectorial de dimensión finita sobre $\mathbb{C}$ para cada k, y es posible dar una fórmula sencilla para su dimensión ([Se, VII, § 3]).

¿Puede una definición aparentemente tan rebuscada tener aplicaciones fuera de la propia teoría de formas modulares? La respuesta es un rotundo sí, y en [Za] se discuten algunas de ellas de nivel asequible. Además siempre está el último teorema

de Fermat para acallar cualquier duda y, desde una perspectiva más general, basta echar un vistazo a las publicaciones o a las bases de datos para darse cuenta de su tremenda participación e impacto en la teoría de números contemporánea.

4.1. Sólo hay una función modular

Es fácil construir formas modulares de peso par $k > 2$ forzando las simetrías. Unas de las más famosas son las series de Eisenstein:

$$E_k(z) = c_k \sum_{(m,n)\in\mathbb{Z}^2\setminus\{\vec{0}\}} \frac{1}{(m+nz)^k},$$

donde se elige la constante $c_k = \left(2\sum_{m=1}^{\infty}\frac{1}{m^k}\right)^{-1} = \frac{1}{2\zeta(k)}$ para que $E_k(i\infty) = 1$. Claramente $E_k(z)$ es 1-periódica y $E_k(-1/z) = z^k E_k(z)$. A partir de estas formas modulares podemos crear otras, por ejemplo $E_4^3(z) - E_6^2(z)$ que es de peso 12 y además se puede probar que no se anula en $\mathbb{H}$ ([Ap, Th. 1.14]). Entonces cualquier forma modular de peso 12 dividida entre ella dará lugar a una función modular. Así podemos considerar

$$j(z) = \frac{1728E_4^3(z)}{E_4^3(z) - E_6^2(z)}. \tag{3}$$

En principio esta elección es arbitraria, pero en seguida veremos que otras construcciones no generan funciones esencialmente diferentes.

Un ejercicio no sencillo pero asequible prueba que, salvo una constante, los coeficientes de Fourier de E_k son una función divisor y por tanto enteros, y así $j(z)$ es el cociente de dos series de coeficientes enteros. La del denominador, después de alguna ingeniosa simplificación (por eso se incluye el 1728), comienza por 1. En esta situación. dividiendo los «polinomios trigonométricos infinitos» se llega a que los coeficientes de Fourier de $j(z)$ son, sorprendentemente, enteros. Calculando los primeros:

$$j(z) = e^{-2\pi iz} + 744 + 196884\, e^{2\pi iz} + 21493760\, e^{4\pi iz} + \cdots \tag{4}$$

Según la consecuencia del principio del argumento antes citada, $N = 1$ implica que j define una aplicación biyectiva del dominio fundamental restringido en $\mathbb{C}$. En el marco de las formas cuadráticas, este hecho se traduce en que $j(z_Q) = j(z_{Q'})$ si y sólo si $Q \sim Q'$.

En cierto modo j es la función modular más sencilla porque, fuera de las constantes, tiene el menor valor de N posible. De hecho es la única función modular, o dicho con más rigor:

Todas las funciones modulares vienen dadas por polinomios en $j(z)$.

La prueba, basada en el principio de inducción sobre el valor N, muestra cómo dicho valor coincide con el grado del polinomio. Imaginemos por ejemplo que tenemos una función modular con $N = 7$ y $a_7 = 2010$; entonces $f(z) - 2010\big(j(z)\big)^7$ tendrá $N = 6$, y repitiendo el procedimiento llegaríamos a una constante en el caso $N = 0$.

5. ¿Por qué $e^{\pi\sqrt{163}}$ está tan cerca de un entero?

El resultado central que probaremos, y que explica la constante de Ramanujan y otras constantes relacionadas, es el siguiente:

> Teorema. *Sean $Q_1, Q_2, \dots\ Q_{h(-D)}$ representantes de las clases de $\mathcal{Q}_D$, entonces el llamado* polinomio de clases
>
> $$P_D(x) = \prod_{k=1}^{h(-D)} \big(x - j(z_{Q_k})\big)$$
>
> *es un polinomio (mónico) con coeficientes enteros.*

El corolario inmediato es que si $h(-D) = 1$ entonces $j(z_Q)$ es un entero para cualquier $Q \in \mathcal{Q}_D$. De todos los discriminantes con número de clases uno, el mayor corresponde a $D = 163$, y en el caso $Q = x^2+xy+41y^2 \in \mathcal{Q}_{163}$, $z_Q = (1+i\sqrt{163})/2$. Al sustituir en (4) se obtiene

$$j(z_Q) = -e^{\pi\sqrt{163}} + 744 - 196884\, e^{-\pi\sqrt{163}} + 21493760\, e^{-2\pi\sqrt{163}} + \dots \in \mathbb{Z}.$$

A pesar de que los coeficientes de Fourier son grandes, $e^{-n\pi\sqrt{163}}$ tiende tan rápido a cero que es natural que los dos primeros términos ya den una buena aproximación. Usando que $j(z_Q) = -640320^3$ (véase la tabla más abajo), se tiene

$$e^{\pi\sqrt{163}} \approx 744 + 640320^3.$$

Esto explica todos los ejemplos de la introducción. Por cierto, la razón de que $j(z_Q)$ sea un cubo perfecto tiene que ver con que $\sqrt[3]{j(z)}$ «casi» es una función modular.

La elección del mayor discriminante hace que la parte imaginaria del z_Q correspondiente sea lo más grande posible, acelerando así la convergencia de la serie de Fourier. Con $D = 67$ sólo obtenemos $e^{\pi\sqrt{67}} = 147197952743{,}999998662\dots$ que mejoramos con un término más a $e^{\pi\sqrt{67}} + 196884e^{-\pi\sqrt{67}} = 147197952744{,}000000000000000991\dots$ (véanse otros cálculos relacionados en la sección 6).

Valores de $j(z_Q)$ para $Q \in \mathcal{Q}_D$ con $h(-D) = 1$ y $-D$ discriminante fundamental									
D	3	4	7	8	11	19	43	67	163
$j(z_Q)$	0	12^3	-15^3	20^3	-32^3	-96^3	-960^3	-5280^3	-640320^3

5.1. La ecuación modular

En este apartado vamos a construir polinomios mónicos de coeficientes enteros cuyas raíces son de la forma $j(z_Q)$. En cierto modo la demostración del teorema se basará en extraer P_D de ellos empleando el máximo común divisor.

En anteriores secciones hemos trabajado con matrices de $\mathrm{SL}_2(\mathbb{Z})$, pero ahora iremos un poco más lejos considerando elementos de M_n, matrices enteras 2×2 con

determinante n positivo. Dentro de M_n caben varias copias de $\mathrm{SL}_2(\mathbb{Z})$ en el sentido de que

$$M_n = \bigcup_{\delta\in\Delta_n} \mathrm{SL}_2(\mathbb{Z})\delta \tag{5}$$

es una partición, donde

$$\Delta_n = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \,:\, ad = n,\ a > 0,\ 0 \le b < d \right\} \subset M_n.$$

Sabemos que $j(z)$ es modular, es decir, $j(\gamma(z)) = j(z)$ para toda matriz $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, pero si en lugar de γ tomamos $\lambda \in M_n$, en general j no tiene por qué satisfacer esta relación y es aquí donde el conjunto Δ_n se transforma en un arma poderosa para poder recuperar la preciada modularidad.

Para hacernos una idea pensemos en el caso $n = 2$ y consideremos $\delta_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, $\delta_2 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ y $\delta_3 = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ en M_2. Tal y como el lector puede adivinar por la notación empleada, estas matrices no están escogidas al azar, sino que son los elementos de Δ_2. Un cálculo sencillo bastaría para comprobar que $j(z/2)$, $j((z+1)/2)$, $j(2z)$ no son modulares, sin embargo, y esto requiere un poco más de ingenio y un redoble de tambor, la función $j(z/2) + j((z+1)/2) + j(2z)$ vuelve a ser modular. Nótese que la traslación $z \mapsto z+1$ permuta los dos primeros sumandos y deja invariante el último, y de forma parecida actúa $z \mapsto -1/z$.

En general, dado $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ y $\delta_i \in \Delta_n$, $\delta_i\gamma$ es un elemento de M_n, y por (5) admitirá una expresión del tipo $\gamma_i'\delta_i'$ para ciertos $\gamma_i' \in \mathrm{SL}_2(\mathbb{Z})$ y $\delta_i' \in \Delta_n$, de modo que si δ_i y δ_j son elementos distintos, necesariamente δ_i' y δ_j' también lo serán. Por tanto, para cualquier $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ existen $\gamma_1, \dots \gamma_s \in \mathrm{SL}_2(\mathbb{Z})$ tales que $\{\delta_1\gamma, \dots, \delta_s\gamma\} = \{\gamma_1\delta_1, \dots, \gamma_s\delta_s\}$ donde $\Delta_n = \{\delta_1, \delta_2, \dots, \delta_s\}$. De aquí, la función $P\big(j(\delta_1(z)), \dots, j(\delta_s(z))\big)$ es modular para cualquier polinomio simétrico P, ya que cambiar z por $\gamma(z)$ se traduce en una permutación de las funciones $j(\delta_i(z))$.

Aplicando este hecho a los coeficientes de

$$\prod_{i=1}^{s}(X - j(\delta_i(z))) = \prod_{\substack{ad=n \\ a,d>0}} \prod_{b=0}^{d-1} \Big(X - j\Big(\frac{az+b}{d}\Big)\Big),$$

visto como polinomio en X, todos ellos serán modulares y por tanto polinomios en $j(z)$. Así pues el producto coincide con algún polinomio $\Phi_n(X, Y) \in \mathbb{C}[X, Y]$ evaluado en $Y = j(z)$, al que llamaremos *n-ésimo polinomio modular*. Nótese que la *ecuación modular*, $\Phi_n(X, Y) = 0$, establece una relación entre $X = j(nz)$ e $Y = j(z)$.

Para ver que los coeficientes de Φ_n están en $\mathbb{Z}$ se utiliza el desarrollo de Fourier de $j(z)$, que tiene la forma $j(z) = \sum_{k=-1}^{\infty} a_k\, e^{2\pi i k z}$ con $a_k \in \mathbb{Z}$. Así

$$\Phi_n(X, j(z)) = \prod_{\substack{ad=n \\ a,d>0}} \prod_{b \pmod d} \Big(X - \sum_{k=-1}^{\infty} a_k \zeta_d^{bk}\, e^{2\pi i k a z/d}\Big),$$

donde $\zeta_d = e^{2\pi i/d}$ es una raíz d-ésima de la unidad. De aquí que los coeficientes de $\Phi_n(X, j(z))$, visto como polinomio en X, tengan un desarrollo de la forma $\sum_{k=-N}^{\infty} b_k e^{2\pi ikz/n}$, con $b_k \in \mathbb{Z}[\zeta_n]$, para cierto N. Los automorfismos de Galois $\zeta_n \mapsto \zeta_n^r$, $\mathrm{mcd}(r, n) = 1$, permutan las series de $j((az + b)/d)$ y por tanto fijan los b_k, que estarán entonces en $\mathbb{Z}[\zeta_n] \cap \mathbb{Q} = \mathbb{Z}$. Las potencias fraccionarias de $e^{2\pi iz}$ desaparecen puesto que, al ser modular, cada coeficiente es invariante bajo $z \mapsto z+1$.

En resumen, los coeficientes de $\Phi_n(X, j(z))$ son polinomios en $j(z)$ con coeficientes enteros, por lo que $\Phi_n(X, Y) \in \mathbb{Z}[X, Y]$, y teniendo en cuenta que, por (5), todo $\lambda \in M_n$ se puede escribir como $\gamma\delta$ para ciertos $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ y $\delta \in \Delta_n$, las soluciones de $\Phi_n(X, j(z)) = 0$ son $X = j(\lambda(z))$ con $\lambda \in M_n$.

Consideramos ahora la restricción del polinomio $\Phi_n(X, Y)$ a la diagonal $X = Y$, $\Psi_n(X) = \Phi_n(X, X)$, descartando el caso en el que n es un cuadrado, puesto que si $n = m^2$ entonces $\delta = \left(\begin{smallmatrix} m & 0 \\ 0 & m \end{smallmatrix}\right)$ sería uno de los elementos de Δ_n y cumpliría $\delta(z) = z$, por lo que $\Phi_n(X, X)$ sería idénticamente nulo. Las soluciones de $\Psi_n(X) = 0$ son aquellos $j(z)$, $z \in \mathbb{H}$, para los que existe un $\lambda \in M_n$ tal que $\lambda(z) = z$, y esta ecuación implica que z es de la forma z_Q. Recíprocamente cada z_Q satisface por definición una ecuación de segundo grado sobre $\mathbb{Z}$, $az^2 + bz + c = 0$ con $a > 0$, por lo que siempre podemos encontrar una matriz $\lambda \in M_{2\times 2}(\mathbb{Z})$ que fije z_Q; tomando por ejemplo $\lambda = \left(\begin{smallmatrix} b+k & c \\ -a & k \end{smallmatrix}\right)$ con k entero, la matriz tendrá determinante positivo y $\lambda(z_Q)$ estará de nuevo en el semiplano superior. Si de entre todas las posibles matrices elegimos una cuyo determinante no sea un cuadrado, entonces bastaría ver que $\Psi_n(X)$ es mónico para poder asegurar que $j(z_Q)$ es un entero algebraico. Se tiene

$$\Psi_n(j(z)) = \prod \left(e^{-2\pi iz} + \cdots - e^{-2\pi iaz/d} e^{-2\pi ib/d} + \cdots\right) = c_N\, e^{-2\pi iNz} + \cdots$$

donde los puntos suspensivos representan términos de orden inferior. Está claro que c_N debe ser una raíz de la unidad, y como además c_N es entero, pues $\Psi_n \in \mathbb{Z}[X]$, las opciones se reducen a $c_N = \pm 1$, por lo que $\Psi_n(X)$ es mónico salvo un signo.

En el caso que nos ocupa, para $z_Q = (1+i\sqrt{163})/2$, la matriz $\lambda = \left(\begin{smallmatrix} -1 & 41 \\ -1 & 0 \end{smallmatrix}\right) \in M_{41}$ satisface $\lambda(z_Q) = z_Q$, y por tanto $\Psi_{41}(j(z_Q)) = 0$. Es decir, $j(z_Q)$ es raíz de Ψ_{41} y por lo anterior entero algebraico.

Conviene señalar que, incluso en los casos más sencillos, Ψ_n es computacionalmente difícil de calcular porque sus coeficientes son gigantescos. Por ejemplo, para $n = 2$ se tiene

$$\Psi_2(X) = -X^4 + 2978X^3 + 40595175X^2 + 17496000000X - 157464000000000.$$

5.2. La prueba del teorema

Antes de nada, veamos los dos primeros casos, $D = 4$ y $D = 3$.

Para $Q = x^2+y^2 \in \mathcal{Q}_4$ se tiene $z_Q = i$, mientras que para $Q = x^2+xy+y^2 \in \mathcal{Q}_3$, $z_Q = \omega$ con $\omega = (-1 + i\sqrt{3})/2$. En la definición de $E_6(i)$ los sumandos $(n + mi)^{-6}$ y $(-m + ni)^{-6}$ son iguales salvo el signo, porque $-m + in = i(n + mi)$, por tanto $E_6(i) = 0$. De la misma forma, como $\omega(n + m\omega) = \omega(n - m) - m$, los sumandos $(n+m\omega)^{-4}$ y $\omega(-m+(n-m)\omega)^{-4}$ son también iguales y por tanto $(1-\omega)E_4(\omega) = 0$. Cuando sustituimos en (3) obtenemos $j(i) = 1728$ y $j(\omega) = 0$, y se deduce el teorema para $D = 4$ y $D = 3$ que son casos de número de clases uno.

El paso clave para deducir el teorema en el caso general es el siguiente lema con sabor aritmético:

$$\text{Sea } Q \in \mathcal{Q}_D,\ \ \Psi_n(j(z_Q)) = 0 \Leftrightarrow x^2 + Dy^2 = 4n \text{ tiene solución } (x, y) \in \mathbb{Z}^2.$$

Una consecuencia es que, o bien $\Psi_n(j(z_Q)) = 0$ se cumple para todas las formas de $\mathcal{Q}_D$, o bien no se cumple para ninguna. Concluimos que $P_D \mid \Psi_n$ si y sólo si $x^2 + Dy^2 = 4n$ tiene solución, mientras que si no hay solución, P_D y Ψ_n no tienen factores comunes.

El lema parece muy profundo pero no lo es en absoluto. Suponemos que n no es un cuadrado (si no, Ψ_n es nulo). Del apartado anterior, $j(z_Q)$ es raíz de Ψ_n si y sólo si existe γ con $\det\gamma = n$ y $\gamma(z_Q) = z_Q$. Si escribimos los elementos de γ como incógnitas y resolvemos la sopa de letras dada por estas ecuaciones (dos ecuaciones con 4 incógnitas), veremos que las soluciones dependen de dos parámetros enteros x e y ligados por $x^2 + Dy^2 = 4n$. ¡Ésa es toda la prueba!

En lugar de escribir los detalles del cálculo, vamos a dar unos indicios geométricos de por qué sale. Si $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, sabemos que $\gamma(z_Q) = z_Q$ equivale a $Q \circ \gamma = Q$, pero si $\det\gamma = n$, entonces γ multiplica por n el área de las celdas del retículo correspondiente a Q y por tanto la equivalencia es con $Q \circ \gamma = nQ$. Fijémonos en $Q = x^2 + y^2 \in \mathcal{Q}_4$; como $Q = \|\cdot\|^2$ en $\mathbb{R}^2$, los cambios orientables que fijan Q son los giros, que tienen por matriz $\left(\begin{smallmatrix}\cos\alpha & -\operatorname{sen}\alpha\\ \operatorname{sen}\alpha & \cos\alpha\end{smallmatrix}\right)$, y eso no deja muchas posibilidades enteras, pues $x = \cos\alpha$, $y = \operatorname{sen}\alpha$ implica $x^2 + y^2 = 1$, pero el caso $\det\gamma = n > 1$ permite tomar múltiplos y las matrices serán de la forma $\gamma = \left(\begin{smallmatrix}x & -y\\ y & x\end{smallmatrix}\right)$ con $x^2+y^2 = n$. Desde el punto de vista de los números reales, pasar de $\mathcal{Q}_4$ a $\mathcal{Q}_D$ es pasar de unas elipses a otras multiplicando una dirección por $\sqrt{D/4}$ y los nuevos «giros» en $\mathcal{Q}_D$ estarán asociados a la ecuación $x^2 + (y\sqrt{D/4})^2 = n$. Renombrando las variables se tiene la ecuación indicada.

Ahora pasamos a probar el teorema. Se procede por inducción. Ya hemos visto que se cumple para 3 y 4, y queremos probarlo para cierto $D > 4$ suponiendo que se verifica en los casos anteriores. Como cada $j(z_Q)$ es raíz de algún Ψ_n, siempre existe un polinomio $P \in \mathbb{Z}[x]$ mónico, con todas sus raíces de esta forma, tal que $P_D \mid P$. Escogiendo P de grado mínimo lo que hay que probar es que toda raíz $j(z_Q)$ de P cumple $Q \in \mathcal{Q}_D$, de esta forma se deduce $P_D = P$. Para ello procederemos por reducción al absurdo suponiendo $Q \in \mathcal{Q}_{D'}$ con $D' \neq D$ para alguna raíz $j(z_Q)$.

Si $D' < D$ entonces, usando la hipótesis de inducción, $P_{D'} \in \mathbb{Z}[x]$ implica $\widetilde{P} = P/\operatorname{mcd}(P, P_{D'}) \in \mathbb{Z}[x]$, y $P_D \mid \widetilde{P}$ contradice que el grado sea mínimo. Si $D' > D$ se puede probar (en seguida indicaremos cómo) que existe n tal que $x^2 + Dy^2 = 4n$ tiene solución, mientras que $x^2 + D'y^2 = 4n$ no la tiene. Por el lema $\Psi_n(j(z_Q)) \neq 0$, y $\widetilde{P} = \operatorname{mcd}(P, \Psi_n)$ vuelve a contradecir la minimalidad del grado.

El único cabo suelto es que para $D < D'$ hay que encontrar un entero n tal que $x^2 + Dy^2 = 4n$ tenga solución y $x^2 + D'y^2 = 4n$ no la tenga (para $D = 3$, $D' = 12$ tal entero no existe, pero estamos bajo la hipótesis $D > 4$). Digamos por ejemplo que D es impar, entonces $D \equiv 3 \pmod 4$. Tomando $n = (D + 1)/4$, la primera ecuación tiene solución $x = 1,\ y = 1$, y la segunda no tiene solución a no ser que $D' = D + 1$ o que $D + 1$ sea un cuadrado. Si ocurre alguna de estas situaciones

tomamos $n = (D+9)/4$ o $n = (D+25)/4$. El único problema aparecería si $D+1$, $D+9$ y $D+25$ fueran los tres cuadrados, pero tal cosa es imposible (los cuadrados pares se espacian más de 20 a partir de 4^2). En el caso D par se procede de forma similar asignando a n el valor $D/4$, $(D+4)/4$ ó $(D+16)/4$.

Con más trabajo se puede demostrar que, de hecho, los factores irreducibles de los polinomios Ψ_n son polinomios de clases y hay una fórmula sencilla debida a L. Kronecker ligando ambas familias de polinomios ([Cox, § 11], [Za]).

6. Más números casi enteros

Ya hemos mencionado que $-D = -163$ es el discriminante mayor en valor absoluto con número de clases 1, por tanto produce en este sentido el mejor ejemplo con $e^{\pi\sqrt{D}}$ casi entero. El mayor discriminante con número de clases 2 es $-D = -427$, y el teorema implica que $j\big((1+i\sqrt{427})/2\big)$ tiene que ser raíz de un polinomio mónico de segundo grado con coeficientes enteros y por tanto se expresará mediante números enteros y una raíz cuadrada. De hecho se tiene[1]

$$j\left(\frac{1+i\sqrt{427}}{2}\right) = -5280^3\big(236674+30303\sqrt{61}\big)^3,$$

y podemos jugar con este valor como hicimos en el caso $D = 163$. Ahora 427 es mucho mayor en valor absoluto y la aproximación por los dos términos $-e^{\pi\sqrt{427}}+744$ de la serie de Fourier será mejor. De esta forma se construyen números casi enteros tales como

$$e^{\pi\sqrt{427}} - 99942102751737734859571200\sqrt{61} \\ = 780572775626189195990630 4743,99999999999999999999987388\ldots,$$

que se mejora al dividir entre 5280^3,

$$\frac{e^{\pi\sqrt{427}}-744}{5280^3} - 6789639488444631\sqrt{61} \\ = 530287796141477 01,99999999999999999999999999999999914322\ldots,$$

o todavía más al extraer raíces cúbicas:

$$\frac{\sqrt[3]{e^{\pi\sqrt{427}}-744}}{5280} - 30303\sqrt{61} \\ = 236673,99999999999999999999999999999999999999999987253\ldots$$

En la teoría de formas modulares abundan las identidades asombrosas debido a la rigidez impuesta por la estructura analítica compleja (o algebraica, si se prefiere),

[1] El valor está tomado de `http://mathworld.wolfram.com/j-Function.html`.

que impide que construyamos muchas formas modulares sin que haya relaciones entre ellas. Parte de esas identidades tienen que ver con los llamados q-productos (aquí q es $e^{2\pi iz}$). Por ejemplo, se puede probar, aunque no es fácil, que la función

$$F(z) = \frac{\left(1 + 256e^{2\pi iz}P(z)\right)^3}{e^{2\pi iz}P(z)} \qquad \text{con} \qquad P(z) = \prod_{n=1}^{\infty}\left(1 + e^{2\pi inz}\right)^{24}$$

es invariante por la acción de $\mathrm{SL}_2(\mathbb{Z})$. Además, con un poco de paciencia se deduce que $F(z) = e^{-2\pi iz} + 744 + \cdots$, por tanto el teorema de Liouville modular aplicado a $j(z) - F(z)$ implica $j(z) = F(z)$. Tomando los primeros términos del producto P se consiguen aproximaciones de $j(z)$ sin tener que enfrentarse a su misterioso desarrollo de Fourier. Por ejemplo, para $z = (1 + i\sqrt{163})/2$ sabemos que $j(z) = -640320^3$, y con solo el primer factor del producto y extrayendo la raíz cúbica se obtiene

$$\frac{1 - 256\, e^{-\pi\sqrt{163}}(1 - e^{-\pi\sqrt{163}})^{24}}{e^{-\pi\sqrt{163}/3}(1 - e^{-\pi\sqrt{163}})^{8}} = 640320{,}000000000000000000000000000743198\ldots$$

Incluso con discriminantes pequeños, por ejemplo 7, se consiguen resultados casi enteros visibles tomando sólo el primer factor:

$$\frac{1 - 256\, e^{-\pi\sqrt{7}}(1 - e^{-\pi\sqrt{7}})^{24}}{e^{-\pi\sqrt{7}/3}(1 - e^{-\pi\sqrt{7}})^{8}} = 15{,}00000868\ldots$$

Cualquiera puede construir sus propios ejemplos con las recetas aquí explicadas y retar a su paquete matemático favorito a que calcule un número suficiente de cifras decimales que desvele su naturaleza casi entera.

7. Epílogo

El experto posiblemente encontrará en las páginas anteriores omisiones que juzgará inexcusables. Le remitimos a la introducción donde se indican el objetivo y las limitaciones del artículo. No obstante, utilizaremos esta última sección para excitar la curiosidad del lector no especialista mencionando brevemente algunos de los conceptos omitidos deliberadamente en nuestra exposición.

Curvas elípticas y multiplicación compleja

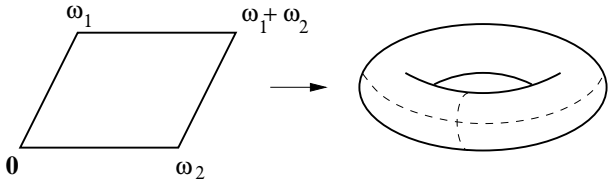
Las formas modulares están vinculadas a las curvas elípticas, curvas cúbicas no singulares que guardan una estrecha relación con las funciones elípticas y las integrales elípticas que se estudiaron durante el siglo XIX. En teoría de números son especialmente importantes las definidas sobre $\mathbb{Q}$, las cuales, con cambios de variable adecuados, siempre se pueden reducir a la forma canónica

$$y^2 = x^3 + Ax + B,$$

con A y B racionales (o incluso enteros).

Si tomamos un polinomio F irreducible, $F(x,y)=0$ define una curva algebraica, a partir de la cual se obtiene una superficie compacta sobre $\mathbb{C}$ llamada *superficie de Riemann de F*. La superficie de Riemann asociada a una curva elíptica es un toro complejo.

La forma más simple de construir un toro T es a partir de un paralelogramo P cuyos lados opuestos identificamos. Si P es el paralelogramo en $\mathbb{C}$ de vértices $0, \omega_1, \omega_2, \omega_1+\omega_2$ entonces la «forma» de P está determinada por $z=\omega_1/\omega_2$, que podemos suponer que está en el semiplano superior (intercambiando ω_1 y ω_2 si fuese necesario). De este modo el conjunto de números complejos de la forma $n\omega_1+m\omega_2$ es un retículo Λ generado por ω_1 y ω_2, y el toro T obtenido a partir de identificar los lados opuestos del paralelogramo P es el cociente $\mathbb{C}/\Lambda$ (aquí el cociente se hace de la manera obvia: $z_1 \sim z_2 \Leftrightarrow z_1 - z_2 \in \Lambda$). Por lo tanto el toro, considerado como una superficie de Riemann, está determinado por el número complejo z, y dos toros T y T' serán isomorfos si y sólo si $\Lambda \sim \Lambda'$.



El toro complejo $\mathbb{C}/\Lambda$ hereda la estructura de grupo de $(\mathbb{C},+)$. En la curva algebraica esta ley de grupo se expresa con funciones racionales en las coordenadas y los coeficientes, por tanto en una curva elíptica definida sobre $\mathbb{Q}$ al operar puntos de coordenadas racionales se obtienen puntos del mismo tipo.

Dado un retículo, es posible encontrar una curva, y a la inversa, cada curva produce un único retículo salvo isomorfismos. Dado $\Lambda=[\omega_1,\omega_2]$, si definimos

$$v_1 = 15\sum_{\omega\in\Lambda}\frac{1}{\omega^4}, \qquad v_2 = 35\sum_{\omega\in\Lambda}\frac{1}{\omega^6},$$

a Λ le corresponde la curva elíptica $y^2 = x^3 - v_1x - v_2$, donde los coeficientes dependen del retículo y la función j se puede relacionar con ellos,

$$E: y^2 = x^3 - v_1x - v_2 \quad \leftrightarrow \quad j(\omega_1/\omega_2) = \frac{6912v_1^3}{4v_1^3 - 27v_2^2}.$$

Inversamente, dada la curva $y^2 = x^3 + Ax + B$, calculando $\frac{6912A^3}{4A^3+27B^2}$ obtenemos un número complejo que corresponde a $j(z)$ para algún $z \in \mathcal{F}^*$, de donde podemos encontrar el retículo que, salvo equivalencias, corresponde a la curva.

Hay otra relación más profunda entre curvas elípticas sobre $\mathbb{Q}$ y formas modulares que ha sido la base para la demostración del *Último Teorema de Fermat*. En cierto modo las formas modulares explican cómo hay que «pegar» todas las informaciones de las curvas reducidas módulo p donde p recorre los primos. Algunos ejemplos explícitos pueden verse en [Ko], [Rab].

Algunos retículos tienen una simetría especial llamada *multiplicación compleja* que consiste en que se aplican en sí mismos al multiplicar por un número complejo (no real). Esto es lo que ocurre por ejemplo con $\mathbb{Z}+i\mathbb{Z}$ multiplicando por cualquier z del propio retículo, y no ocurre en $\mathbb{Z}+i\sqrt[3]{2}\,\mathbb{Z}$ para ningún $z\in\mathbb{C}\setminus\mathbb{R}$. Los retículos con multiplicación compleja están ligados a las extensiones cuadráticas imaginarias de $\mathbb{Q}$, y todos los del tipo $[z_Q,1]$ tienen esta propiedad. Un retículo con multiplicación compleja está casi determinado, salvo equivalencias, por su anillo de endomorfismos, en el sentido de que el número de posibilidades está acotado por el número de clases. Por ejemplo, $\alpha=(1+i\sqrt{163})/2$ es z_Q para cierta Q en $\mathcal{Q}_{163}$ que tiene número de clases uno, por tanto $[\alpha,1]$ es el único retículo salvo equivalencias con anillo de endomorfismos $\mathbb{Z}[\alpha]$. Esto implica que sólo hay una curva elíptica E, salvo isomorfismos, con $\mathrm{End}(E)\cong\mathbb{Z}[\alpha]$. Por otro lado, si $j(\alpha)\in\mathbb{C}$ no fuera racional, aplicando un automorfismo de $\mathbb{C}$ a los coeficientes de E obtendríamos una nueva curva con el mismo anillo de endomorfismos. Este esquema permite dar una prueba «rápida» de que $j(\alpha)\in\mathbb{Q}$ (véase [Sh, § 4.6, p. 108]) y por tanto de que $e^{\pi\sqrt{163}}$ es casi racional, dando por supuesto conocimientos sobre curvas elípticas, y lo suficiente de teoría de ideales en extensiones cuadráticas, para deducir que un retículo está determinado, salvo equivalencias, siempre que su anillo de endomorfismos sea un dominio de factorización única.

Factorización, ideales y formas cuadráticas

En $\mathbb{Z}[i]$ hay una teoría de factorización en primos similar a la de $\mathbb{Z}$, pero si intentamos reproducir esta teoría, por ejemplo en $\mathbb{Z}[\sqrt{-5}]$, decretando que un número es primo si y sólo si es divisible por sí mismo y la unidad (salvo signos), en seguida encontramos dificultades como

$$3\cdot 7=(1+2\sqrt{-5})\cdot(1-2\sqrt{-5}),$$

donde el primo $1+2\sqrt{-5}$ no divide ni a 3 ni a 7 pero sí a su producto. La explicación moderna es que cada número $z\in\mathbb{Z}[\sqrt{-5}]$ tiene asociado el retículo $[z]=[z\sqrt{-5},z]$, y definiendo la multiplicación de retículos como el (menor) retículo que contiene a los productos de los generadores, se puede concluir

$$[3]=[1+2\sqrt{-5},3]\cdot[3,1-2\sqrt{-5}]\quad\text{y}\quad[1+2\sqrt{-5}]=[1+2\sqrt{-5},3]\cdot[1+2\sqrt{-5},7].$$

Entonces 3 y $1+2\sqrt{-5}$ tienen un factor en común $[1+2\sqrt{-5},3]$ que no vemos porque no corresponde a un número, es un *ideal* (un retículo). Admitiendo la factorización en ideales todo vuelve a funcionar como en $\mathbb{Z}$.

Gauss consideró, en lugar de retículos, las formas cuadráticas, y definió una manera extraña de multiplicarlas, que llamó *composición* ([Gau, Art. 235]). Las clases de retículos corresponden a clases de formas cuadráticas que con la composición adquieren una estructura de grupo abeliano de orden el número de clases. Cuando este grupo, llamado *grupo de clases*, es trivial, la correspondencia entre retículos y números de verdad es biyectiva.

Cuerpos de clases y funciones modulares

Aunque los ideales constituyen la manera más simple de resolver los problemas de factorización, la *teoría de cuerpos de clases* permite salvar nuestra intuición y «ver» como números los factores comunes inexistentes en $\mathbb{Z}[\sqrt{-5}]$ del apartado anterior a cambio de introducir nuevos elementos en el anillo (de cierta forma canónica y minimal en la que no entraremos). En nuestro ejemplo, la teoría dice que los problemas con la factorización desaparecerán si se extiende $\mathbb{Z}[\sqrt{-5}]$ al anillo de enteros de $K = \mathbb{Q}(\sqrt{-5}, j(\sqrt{-5}))$, que es igual a $\mathbb{Z}[i, r]$ con r la razón áurea. Se dice que K es el *cuerpo de clases* (de Hilbert). En $\mathbb{Z}[i, r]$ ya se puede definir el (máximo) común divisor de 3 y $1 + 2\sqrt{-5}$, y es $1 - r - ir$. En los anillos de enteros de cuerpos cuadráticos imaginarios siempre hay que añadir un valor especial de la función j al cuerpo [Coh], pero no se sabe bien cuál es el análogo para el caso cuadrático real,[2] y esta pregunta es parte de la lectura que se hace del 12º problema de Hilbert [Ba 2].

La teoría de cuerpos de clases excede este burdo ejemplo de factorización y, por resumir su propósito en una línea, intenta representar propiedades aritméticas en el grupo de Galois de una extensión de cuerpos. Por dar un ejemplo concreto, en el caso anterior $\mathrm{Gal}(K/\mathbb{Q}(\sqrt{-5}))$ es isomorfo al grupo de clases. Una de las motivaciones de la teoría es generalizar la ley de reciprocidad cuadrática.

Agradecimientos. Agradecemos a Adrián Ubis la colaboración prestada en la elaboración de este artículo, así como a David Torres por la lectura y sugerencias realizadas. También queremos agradecer a Javier Cilleruelo, encargado de la sección, por su atenta lectura y su ayuda en la preparación del manuscrito.

Referencias

[Ap] T. M. Apostol, *Modular functions and Dirichlet series in number theory*, Springer-Verlag, New York, 1990.

[Ba 1] P. Bayer, Les Disquisicions aritmètiques de Gauss, *Conferències FME, Volum III. Curs Gauss,* 87–123, Universitat Politècnica de Catalunya, 2006.

[Ba 2] P. Bayer, El sueño de juventud de Kronecker, *Notes del Seminari de Teoria de Nombres*, UB-UAB-UPC, Barcelona, 2005.

[Coh] H. Cohn, *Introduction to the construction of class fields*, Cambridge University Press, Cambridge, 1985.

[Cox] D.A. Cox, *Primes of the form $x^2 + ny^2$. Fermat, class field theory and complex multiplication*, John Wiley & Sons, New York, 1989.

[Gar] M. Gardner, Mathematical Games, *Scientific American* **232** (4) (1975), 127.

[Gau] C. F. Gauss, *Disquisitiones arithmeticae*, Springer-Verlag, New York, 1986.

[Go] D. Goldfeld, Gauss's class number problem for imaginary quadratic fields, *Bull. Amer. Math. Soc. (N.S.)* **13** (1985), no. 1, 23–37.

[2]En las Terceras Jornadas de Teoría de Números, V. Rotger presentó algunos de sus avances en el problema en una bella charla que posiblemente aparecerá en [Te].

[Ko] N. Koblitz, *Introduction to elliptic curves and modular forms*, Second edition, Graduate Texts in Mathematics 97, Springer-Verlag, New York, 1993.

[Rab] D. Raboso, Formas modulares y la curva $y^2 = x^3 - 35x - 98$. *Trabajo de fin de máster*, Universidad Autónoma de Madrid, 2009.

[Ram] S. Ramanujan, Modular equations and approximations to π, *Quart. J. Math.* **45** (1914), 350–372.

[Se] J.-P. Serre, *A course in arithmetic*, Graduate Texts in Mathematics 7, Springer-Verlag, New York-Heidelberg, 1973.

[Sh] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, Princeton, N.J., 1971.

[Te] *Proceedings of «Terceras Jornadas de Teoría de Números»*, Salamanca, 29 junio 2009 – 3 julio 2009 (por aparecer).

[Za] D. Zagier, Elliptic modular forms and their applications, *The 1-2-3 of modular forms,* 1–103, Universitext, Springer, Berlin, 2008.

Fernando Chamizo, Dpto. de Matemáticas, Universidad Autónoma de Madrid
Correo electrónico: fernando.chamizo@uam.es
Página web: http://www.uam.es/fernado.chamizo

Dulcinea Raboso, Dpto. de Matemáticas, Universidad Autónoma de Madrid
Correo electrónico: dulcinea.raboso@uam.es