USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
FACULTADE DE MATEMÁTICAS

Traballo Fin de Grao

Un recorrido por el último teorema de Fermat

Lucas López Román

Curso 2024/2025

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

GRADO DE MATEMÁTICAS

Traballo Fin de Grao

Un recorrido por el último teorema de Fermat

Lucas López Román

Junio, 2025

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Trabajo propuesto

Área de Conocimiento: Álgebra
Título: Un recorrido por el último teorema de Fermat
Breve descripción del contenido
Este TFG pretende hacer una aproximación a algunas de las ideas que se desarrollaron durante la exploración y posterior demostración del último teorema de Fermat. Comenzando por los casos $n = 3$ y $n = 4$, se trabajarán algunos aspectos relativos a la aritmética de los cuerpos de números y a los ideales de Dedekind. En la parte final, realizaremos una breve aproximación a algunas de las ideas desarrolladas en el S.XX en torno al concepto de modularidad, que permitirían la demostración del resultado realizada por Andrew Wiles en los años 90.
Recomendaciones
Otras observaciones

Índice

Resumen

Este trabajo realiza una aproximación a algunas de las ideas que se desarrollaron durante la exploración y posterior demostración del último teorema de Fermat. Comenzando por los casos $n = 3$ y $n = 4$, se trabajarán algunos aspectos relativos a la aritmética de los cuerpos de números. La parte central de la memoria se dedica al estudio de la demostración del teorema de Fermat para primos regulares. Finalmente, realizamos un breve acercamiento a algunas de las ideas desarrolladas en el S.XX en torno al concepto de modularidad, que permitirían la demostración del resultado realizada por Andrew Wiles en los años 90.

Abstract

This work presents an approach to some of the ideas that were developed during the exploration and eventual proof of Fermat's Last Theorem. Starting with the cases $n = 3$ and $n = 4$, we will explore certain aspects related to the arithmetic of number fields. The central part of the report focuses on the study of the proof of Fermat's Theorem for regular primes. Finally, we offer a brief overview of some of the ideas developed in the 20th century around the concept of modularity, which ultimately led to Andrew Wiles' proof of the result in the 1990s.

Introducción

Este trabajo se centrará en la demostración del teorema de Fermat, que afirma lo siguiente: la ecuación dada por $x^n + y^n = z^n$ no admite ninguna solución distinta de la trivial para exponentes enteros mayores o iguales que 3, donde x, y y z son números enteros. Se trata de uno de los resultados más conocidos de las matemáticas, que ha obsesionado a toda la comunidad durante más de tres siglos.

Este teorema fue enunciado por primera vez en el siglo XVII, cuando el jurista y matemático aficionado francés, Pierre de Fermat, dejó por escrito el siguiente párrafo en un margen de la *Arithmetica* de Diofanto: *Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi, hanc marginis exiguitas non caperet.*

Comenzaremos por estudiar las soluciones para la expresión anterior cuando $n = 2$, pues en este caso sí que podemos encontrar alguna no trivial. Esto se corresponde con el problema de obtener las ternas pitagóricas, mucho más antiguo que el propio enunciado del teorema de Fermat. Y es que, ya en la civilización babilónica (1800 a.C.), se conocían algunas soluciones particulares para la ecuación $x^2 + y^2 = z^2$, encontradas en la famosa tablilla Plimpton 322. Posteriormente, fue en la Grecia Antigua cuando se formalizó esto con el teorema de Pitágoras, en el seno de la escuela pitagórica. Para este problema, probaremos que cualquier solución se puede escribir como

$$(x, y, z) = (\lambda(p^2 - q^2), 2\lambda pq, \lambda(p^2 + q^2)),$$

donde $p, q \in \mathbb{Z}$ son enteros positivos de distinta paridad con $p > q$ y coprimos entre sí, y λ es un entero positivo.

Una vez hecho esto, probaremos el resultado para los exponentes tres y cuatro, centrándonos así en dos casos concretos. El primero, correspondiente a $n = 4$ y ya estudiado por Fermat, hace uso de la llamada *técnica del descenso*, que es un método clásico que aparece en diferentes problemas de teoría de números. Se basa en suponer la existencia de una solución para la ecuación $x^4+y^4 = z^4$, para luego comprobar que debe haber otra solución menor, pues entonces tendríamos

una sucesión infinita decreciente de enteros positivos, siendo esto imposible. Por su parte, el caso $n = 3$ requiere facorizar la ecuación $x^3 + y^3 = z^3$, empleando la raíz tercera (primitiva) de la unidad $\omega = \frac{-1+\sqrt{-3}}{2}$. Estudiaremos la aritmética del anillo $\mathbb{Z}[\omega]$ y utilizaremos sus propiedades para, supuesta una solución para la ecuación sin factores comunes, encontrar otra menor que contradiga lo anterior.

A continuación, echaremos mano de algunas ideas relativas a la aritmética de los cuerpos de números y de los dominios de Dedekind para demostrar el teorema para los denominados *primos regulares*. En ese sentido, el capítulo 2 sirve para presentar nociones más o menos elementales de teoría algebraica de números, especialmente el cuerpo de clases, que desempeñará un papel crucial en el siguiente capítulo. Más concretamente, el grupo de clases mide cómo de lejos está el anillo de enteros de un cuerpo de números de ser un dominio de ideales principales. En concreto, para la extensión ciclotómica p-ésima, si dicho grupo de clases tiene cardinal coprimo con p, se dice que el primo p es regular. En esos casos, podremos concluir que, si la potencia p-ésima de un ideal, $\mathfrak{a}^p$, es principal, entonces $\mathfrak{a}$ también lo debe ser, lo cual nos permite realizar la demostración del último teorema de Fermat. Para eso, en cualquier caso, distinguiremos dos situaciones: una primera, en la que supondremos que $p \nmid xyz$, y una segunda, en la que relajaremos esa condición. En esta última, la conclusión requiere de un resultado de Kummer de carácter más técnico.

Las ideas empleadas en la demostración del último teorema de Fermat para primos regulares, desarrolladas por la escuela francesa a lo largo del siglo XIX, tienen su motivación en la aritmética clásica. Sin embargo, en lugar de trabajar en el anillo de los enteros $\mathbb{Z}$, se trabaja en el anillo $\mathbb{Z}[\zeta]$, siendo ζ una raíz primitiva p-ésima de la unidad. En ese contexto, no tenemos en general factorización única, pero sí tenemos factorización única de cualquier ideal en producto de ideales primos (teorema de Dedekind). Esto nos permite desarrollar diferentes argumentos aritméticos y realizar una demostración que es válida para un gran número de primos. De hecho, el primer primo irregular es el 37 (si bien es cierto que existen infinitos primos irregulares).

En particular, la idea de demostración empleada para el caso de los primos regulares no se pudo extender al caso general, lo que llevó a una búsqueda de un nuevo enfoque que permitiera abordar la prueba para los casos restantes. Así, en el capítulo 4, presentaremos la teoría de curvas elípticas y formas modulares, dos objetos de naturaleza aparentemente distinta, cuya interacción permitió a Andrew Wiles demostrar el último teorema de Fermat en el siglo XX. Discutiremos sus definiciones y presentaremos sus propiedades básicas, siempre orientadas a las aplicaciones al teorema.

Por último, en el capítulo 5, realizaremos una introducción a la noción de modularidad y enunciaremos la famosa conjetura de Taniyama–Shimura, el resultado que permitió asociar a cada curva elíptica una forma modular. Concluiremos añadiendo la construcción de la curva elíptica de Frey (o de Frey–Hellegouarch), para luego efectuar una pequeña aproximación a la

demostración del último teorema de Fermat a partir de la anterior conjetura, relación que fue establecida por Gerhard Frey y Ken Ribet.

Capítulo 1

El teorema de Fermat para exponentes bajos

En este capítulo, nos centraremos en encontrar las distintas soluciones enteras para la ecuación de Fermat

$$x^n + y^n = z^n, \quad x, y, z \in \mathbb{Z},$$

cuando $n = 2$. Esto irá seguido de la demostración del teorema para dos casos particulares, como son $n = 4$ y $n = 3$. Lo haremos en ese orden atendiendo a la complejidad de sendas demostraciones. En particular, los casos $n = 2$ y $n = 4$ se pueden tratar empleando técnicas de aritmética elemental, mientras que el caso $n = 3$ requiere trabajar con los llamados *enteros de Eisenstein*.

1.1. Las ternas pitagóricas y el caso $n = 2$

Es algo ya conocido que la ecuación dada por

$$x^2 + y^2 = z^2 \tag{1.1}$$

tiene infinitas soluciones en el conjunto de los números enteros. Sin embargo, surge la duda de si es posible encontrar una expresión general para dichas soluciones. Por tanto, en esta sección, nos centraremos en buscarla, tomando como inspiración las referencias [6] y [5].

Definición 1.1. Una *terna* es un punto (x_0, y_0, z_0) del espacio tridimensional $\mathbb{R}^3$. En este trabajo buscaremos únicamente soluciones de números enteros, por lo que, en adelante, las ternas consideradas se referirán a puntos del espacio $\mathbb{Z}^3$.

Observación 1.2. En esta sección veremos que existen distintas ternas que son soluciones de la ecuación $x^2 + y^2 = z^2$. Podemos considerar, por ejemplo, $(0, 0, 0)$, que es la llamada solución trivial, $(3, 4, 5)$ o $(5, 12, 13)$.

Definición 1.3. Una terna se dice *pitagórica* si es solución de la ecuación $x^2 + y^2 = z^2$.

Sea (x, y, z) una terna pitagórica de enteros positivos. Si los tres números tuvieran un factor común d, entonces $(x/d, y/d, z/d)$ también sería una terna pitagórica. Además, si dos de ellos tuviesen un factor común, entonces el tercero también lo tendría. Por tanto, podemos suponer que x, y y z son primos relativos. En ese caso, decimos que la terna pitagórica es *primitiva*.

Observación 1.4. Por lo anterior, x, y y z no pueden ser los tres pares, ni siquiera dos de ellos. Tampoco pueden ser los tres impares, pues en tal caso tendríamos que la suma de dos números impares nos proporcionaría otro impar, lo cual es imposible. Con esto, deducimos que uno es par y los otros dos impares.

Proposición 1.5. *La suma de dos cuadrados de números impares no puede ser un cuadrado perfecto.*

Demostración. Supongamos que fuese cierto. Entonces, tendremos que existen enteros $n, m \in \mathbb{Z}$ con $x = 2m + 1$, $y = 2n + 1$ y $x^2 + y^2 = z^2$. Sustituyendo x e y y elevando al cuadrado, tenemos que:

$$\begin{aligned}(2m+1)^2 + (2n+1)^2 &= 4m^2 + 4m + 1 + 4n^2 + 4n + 1 \\ &= 2(2m^2 + 2m + 2n^2 + 2n + 1) \\ &= z^2.\end{aligned}$$

Así, z debe ser par, y supongamos que $z = 2q$, con $q \in \mathbb{Z}$. Ahora, tenemos:

$$2(2m^2 + 2m + 2n^2 + 2n + 1) = (2q)^2 = 4q^2.$$

Despejando, se obtiene que

$$1 = 2q^2 - 2m^2 - 2m - 2n^2 - 2n,$$

lo cual es imposible, porque $2 \nmid 1$. □

De este modo, hemos visto que z no puede ser par. Supondremos entonces, sin pérdida de generalidad, que x lo es. Continuemos escribiendo la ecuación 1.1 de la siguiente forma:

$$x^2 = z^2 - y^2 = (z + y)(z - y), \tag{1.2}$$

donde $(z + y)$ y $(z - y)$ son pares. Por tanto, existen $u, v, w \in \mathbb{Z}^+$ con $x = 2u$, $z + y = 2v$, $z - y = 2w$. Además, por como están definidos, w y v son coprimos. Así, tenemos:

$$u^2 = vw. \tag{1.3}$$

Proposición 1.6. *Si el producto de dos enteros positivos que son primos relativos es igual a un cuadrado, entonces dichos enteros son cuadrados.*

Demostración. Supongamos entonces que $ab = x^2$, donde a y b son coprimos y $a, b, x \in \mathbb{Z}$. Entonces, a y b no tienen factores comunes, por lo que la factorización de x^2 resultará de la yuxtaposición de las factorizaciones de a y b. Esto es:

$$a = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}, \quad b = y_1^{j_1} y_2^{j_2} \cdots y_m^{j_m}; \quad x^2 = ab = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} y_1^{j_1} y_2^{j_2} \cdots y_m^{j_m}$$

Por otro lado, la factorización de x^2 debe tener todos los exponentes pares (pues estos serán los de la factorización de x multiplicados por dos), por lo que los exponentes en las respectivas factorizaciones de a y b también deberán ser pares. □

De este modo, existirán $p, q \in \mathbb{Z}^+$ de modo que $v = p^2$ y $w = q^2$ (y también son primos relativos). Ahora, como $z+y = 2v$, $z-y = 2w$, se sigue que $z = v+w = p^2+q^2$, $y = v-w = p^2-q^2$. Además, $p > q$ ($y > 0$) y p y q tienen distinta paridad (y, z son impares). Por último,

$$x^2 = z^2 - y^2 = (p^2+q^2)^2 - (p^2-q^2)^2 = 4p^2q^2 = (2pq)^2,$$

de donde $x = 2pq$.

El siguiente resultado, que expresa la forma general de las ternas pitagóricas primitivas (sin factores en común), es inmediato a partir de la discusión anterior.

Teorema 1.7. *Las ternas pitagóricas primitivas son de la forma*

$$(2pq, p^2-q^2, p^2+q^2),$$

donde $p, q \in \mathbb{Z}$ son enteros coprimos, de paridad distinta y tales que $p > q$.

1.2. El método del descenso y el caso $n = 4$

Llegados a este punto y empleando lo demostrado en la sección anterior, ya podemos probar el caso $n = 4$ del teorema de Fermat. De nuevo, seguiremos las referencias [6] y [5].

Teorema 1.8. *No existe ninguna solución distinta de la trivial para la ecuación*

$$x^4 + y^4 = z^4 \tag{1.4}$$

Demostración. Para demostrarlo, emplearemos el método del descenso infinito, una forma de reducción al absurdo.

Supongamos, por el contrario, que sí existen $x, y, z \in \mathbb{Z}^+$ tales que son una solución para la ecuación 1.4. Además, supongamos que son primos relativos (en otro caso, dos de ellos tendrían un factor común y, al verificarse la ecuación 1.4, el tercero también lo tendría, por lo que bastaría con tomar la terna resultante de eliminar los factores comunes). De este modo, (x^2, y^2, z^2) es una terna pitagórica y, por el teorema 1.7, suponiendo que x^2 es par, tenemos que:

$$x^2 = 2pq, \quad y^2 = p^2 - q^2, \quad z^2 = p^2 + q^2,$$

donde los enteros p y q son primos relativos con distinta paridad y $0 < q < p$.

Despejando ahora en la segunda de las igualdades: $y^2 + q^2 = p^2$, de donde y, q y p conforman otra terna pitagórica primitiva. Como y es impar y por las deducciones de la sección anterior, tendremos que q es par y p es impar. Aplicamos entonces el teorema 1.7 nuevamente:

$$q = 2ab, \quad y = a^2 - b^2, \quad p = a^2 + b^2,$$

donde a y b son enteros positivos con paridades opuestas, primos relativos y tales que $0 < b < a$.

Aquí, podemos escribir:

$$x^2 = 2pq = 4ab(a^2 + b^2),$$

de donde se deduce que $ab(a^2 + b^2)$ es un cuadrado. Además, ab y $(a^2 + b^2)$ son primos relativos, pues si un primo k divide a ab, entonces también divide a a o a b, pero no a ambos (a y b son primos relativos). De esto, se sigue que ab y $(a^2 + b^2)$ también deben ser cuadrados, de donde a y b también lo son.

Supongamos, por tanto, que $a = m^2$ y $b = n^2$, con $m, n \in \mathbb{Z}$. Entonces:

$$m^4 + n^4 = a^2 + b^2 = p < p^2 + q^2 = z^2 < z^4 = x^4 + y^4.$$

Realizando este procedimiento de forma recursiva, obtenemos una sucesión infinita de enteros positivos decreciente, lo cual es imposible. Hemos llegado, por tanto, a una contradicción, por lo que no puede existir ninguna solución distinta de la trivial para la ecuación 1.4.

□

1.3. Los enteros de Eisenstein y el caso $n = 3$

En esta sección, nos centraremos en demostrar que la ecuación

$$x^3 + y^3 = z^3 \tag{1.5}$$

no tiene soluciones en $\mathbb{Z}^3$ distintas de la trivial, $(0, 0, 0)$. Para ello, nos guiaremos empleando la referencia [6].

Sea $\omega = \frac{-1+\sqrt{-3}}{2}$ una raíz cúbica primitiva de la unidad.

Definición 1.9. Los enteros de Eisenstein, $\mathbb{Z}[\omega]$, se definen como

$$\mathbb{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbb{Z}\}.$$

Observación 1.10. Esta definición viene motivada por la factorización del polinomio:

$$x^3 + y^3 = (x + y)(x^2 - xy + y^2),$$

pues desarrollando la segunda parte de la expresión anterior cuando $y \neq 0$ se tiene:

$$x^2 - xy + y^2 = y^2 \left(\frac{x^2 - xy + y^2}{y^2} \right) = y^2 \left(\left(\frac{x}{y} \right)^2 - \frac{x}{y} + 1 \right).$$

Tomando ahora $\gamma = x/y$ y resolviendo la ecuación de segundo grado correspondiente, podemos reescribir la igualdad anterior:

$$x^2 - xy + y^2 = y^2(\gamma^2 - \gamma + 1) = y^2 \left(\gamma - \frac{1+\sqrt{-3}}{2} \right) \left(\gamma - \frac{1-\sqrt{-3}}{2} \right),$$

y deshaciendo el cambio de variable realizado:

$$\begin{aligned} x^3 + y^3 &= (x + y)y^2 \left(\frac{x}{y} - \frac{1+\sqrt{-3}}{2} \right) \left(\frac{x}{y} - \frac{1-\sqrt{-3}}{2} \right) \\ &= (x + y) \left(x - y\frac{1+\sqrt{-3}}{2} \right) \left(x - y\frac{1-\sqrt{-3}}{2} \right). \end{aligned}$$

Observación 1.11. Nótese que $\omega = \frac{-1+\sqrt{-3}}{2}$ es una raíz cúbica primitiva de la unidad ($\omega^3 = 1$). Tal y como se desarrolla en [6, Cap. 1], se tiene que $\mathbb{Z}[\omega]$ es un dominio euclídeo con la norma dada por

$$N(a + b\omega) = (a + b\omega)(a + b\overline{\omega}) = a^2 - ab + b^2.$$

Además, por ser dominio euclídeo, también es DIP (dominio de ideales principales) y DFU (dominio de factorización única).

Para encontrar las unidades de $\mathbb{Z}[\omega]$, imponemos la condición necesaria de que $N(a+b\omega) = 1$, esto es, $a^2 - ab + b^2 = 1$; alternativamente,

$$3(a - b)^2 + (a + b)^2 = 4.$$

De aquí, un análisis caso a caso muestra que los elementos que verifican tal condición son 1, -1, ω, $-\omega$, ω^2 y $-\omega^2$, donde hemos usado que $\omega^2 = -1 - \omega$. Por último, basta ver que cada uno de ellos tiene inverso (que es otro de ellos):

- 1 es su propio inverso, al igual que -1.

- $\omega\omega^2 = \omega(-1-\omega) = -\omega - \omega^2 = 1$, luego ω y ω^2 son inversos.
- $-\omega(-\omega^2) = -\omega(1+\omega) = -\omega - \omega^2 = 1$, de donde $-\omega$ y $-\omega^2$ son inversos.

Observación 1.12. Nótese que $\omega\overline{\omega} = 1$ y $\omega + \overline{\omega} = -1$.

De este modo, bastará con probar que la ecuación 1.5 no tiene soluciones en $\mathbb{Z}[\omega]$. Además, de existir una solución, podemos suponer que dos cualesquiera de los números que la componen no tienen factores comunes (en caso de tenerlos, también estarían en el tercero y podríamos obtener otra solución eliminándolos).

Proposición 1.13. *Sea $\rho = 1 - \omega$. Entonces, tendremos que*

$$\mathbb{Z}[\omega]/\rho\mathbb{Z}[\omega] \simeq \mathbb{Z}/3\mathbb{Z},$$

por lo que $a + b\omega \equiv -1, 0, 1 \pmod{\rho}$.

Demostración. Como bien hemos dicho, sea $\rho = 1 - \omega$. Entonces, tenemos que

$$N(\rho) = N(1-\omega) = (1-\omega)(1-\overline{\omega}) = 3,$$

por lo que 3 es múltiplo de ρ. Además, recordemos que para el cociente $\mathbb{Z}[\omega]/\rho\mathbb{Z}[\omega]$:

$$[\alpha] = [\beta] \Leftrightarrow \alpha - \beta \in \rho\mathbb{Z}[\omega].$$

Dado entonces $a + b\omega \in \mathbb{Z}[\omega]$, sean a' y b' los enteros entre 0 y 2 tales que $a \equiv a' \pmod 3$ y $b \equiv b' \pmod 3$. Tendremos, por tanto, nueve opciones distintas de clases de equivalencia módulo 3 en $\mathbb{Z}[\omega]$: $0, 1, 2, \omega, 1+\omega, 2+\omega, 2\omega, 2\omega+1$ y $2\omega+2$.

Ahora, como $\rho = 1-\omega$, entonces $1-\omega \equiv 0 \pmod{\rho}$, y tendremos las siguientes congruencias:

- $1 \equiv \omega \pmod{\rho}$,
- $2 \equiv \omega + 1 \pmod{\rho}$,
- $0 \equiv \omega - 1 \equiv \omega + 2 \pmod{\rho}$ (recordemos que, como 3 es múltiplo de ρ, las congruencias módulo 3 también son ciertas),
- $2 \equiv 2\omega \pmod{\rho}$,
- $0 \equiv 2\omega + 1 \pmod{\rho}$,
- $1 \equiv 2\omega + 2 \pmod{\rho}$.

De este modo, hemos visto que todas las posibles clases de equivalencia en $\mathbb{Z}[\omega]$ planteadas anteriormente se pueden reducir a las clases $[0], [1]$ y $[2]$ cuando el módulo es ρ, por lo que es claro el isomorfismo entre $\mathbb{Z}[\omega]/\rho\mathbb{Z}[\omega]$ y $\mathbb{Z}/3\mathbb{Z}$.

□

Lema 1.14. *Sea $\alpha = a + b\omega$ de modo que $\alpha \equiv \pm 1 \pmod{\rho}$. Entonces, tenemos que $\alpha^3 \equiv \pm 1 \pmod{\rho^4}$.*

Demostración. Centrémonos en el caso en que $\alpha \equiv 1 \pmod{\rho}$. Queremos ver que $\alpha^3 \equiv 1 \pmod{\rho^4}$ o, lo que es lo mismo, que $\alpha^3 - 1 \equiv 0 \pmod{\rho^4}$. Bastará entonces con ver que $\alpha^3 - 1$ es múltiplo de ρ^4. Tenemos:

$$\alpha^3 - 1 = (\alpha - 1)(\alpha^2 + \alpha + 1) = (\alpha - 1)(\alpha - \omega)(\alpha - \omega^2).$$

Ahora, veamos para cada uno de los factores anteriores:

- $\alpha \equiv 1 \pmod{\rho}$ por hipótesis, de donde $\alpha - 1 \equiv 0 \pmod{\rho}$;
- $1 \equiv \omega \pmod{\rho}$ por la proposición anterior, luego $\alpha \equiv \omega \pmod{\rho}$ y $\alpha - \omega \equiv 0 \pmod{\rho}$;
- $\omega^2 = \overline{\omega} = -1 - \omega \equiv 2 + 2\omega \equiv 1 \pmod{\rho}$, por lo que $\alpha - \omega^2 \equiv 0 \pmod{\rho}$.

De este modo, tenemos que los tres factores son múltiplos de ρ. Veamos ahora que los tres tienen restos distintos con módulo ρ^2. Para ello, veremos que no puede darse que dos de los restos con dicho módulo coincidan. Entonces:

- $\alpha - 1 \equiv \alpha - \omega \pmod{\rho^2}$, luego $1 - \omega = \rho \equiv 0 \pmod{\rho^2}$, lo cual es imposible;
- $\alpha - 1 \equiv \alpha - \omega^2 \pmod{\rho^2}$, luego $1 - \omega^2 \equiv 0 \pmod{\rho^2}$. Como $1 - \omega^2 = (1 + \omega)(1 - \omega)$ y $1 + \omega \equiv 2 \pmod{\rho}$ es invertible, nos queda que $\rho \equiv 0 \pmod{\rho^2}$, lo cual es imposible;
- $\alpha - \omega \equiv \alpha - \omega^2 \pmod{\rho^2}$, luego $\omega - \omega^2 = \omega(1 - \omega) = \rho\omega \equiv 0 \pmod{\rho^2}$, pero $1 \equiv \omega \pmod{\rho}$, por lo que $\rho \equiv 0 \pmod{\rho^2}$, lo cual es imposible.

Observamos ahora que $\mathbb{Z}[\omega]/\rho^2\mathbb{Z}[\omega]$ es un anillo con 9 elementos en el que exactamente 3 de ellos son múltiplos de ρ. En efecto,

$$\rho^2 = (1 - \omega)^2 = \omega^2 - 2\omega + 1 = -1 - \omega - 2\omega + 1 = -3\omega,$$

por lo que, multiplicando por $-\omega^2$, tenemos que $3 \equiv 0 \pmod{\rho^2}$.

Esto quiere decir que un conjunto de representantes de $\mathbb{Z}[\omega]/\rho^2\mathbb{Z}[\omega]$ viene dado por los elementos $a + b\omega$, con $a, b \in \{0, 1, 2\}$, donde únicamente 0, $1 + 2\omega$ y $2 + \omega$ son múltiplos de ρ. Por

lo tanto, los tres factores $\alpha - 1$, $\alpha - \omega$ y $\alpha - \omega^2$ son múltiplos de ρ, y cada uno de ellos es congruente con uno de los tres factores diferentes módulo ρ^2 que son múltiplos de ρ; en particular, exactamente uno de ellos será congruente con 0, lo que quiere decir que uno (y solo uno) será múltiplo también de ρ^2.

Alternativamente, lo que hemos visto es que tenemos un isomorfismo de anillos

$$\mathbb{Z}[\omega]/\rho^2\mathbb{Z}[\omega] \simeq \mathbb{Z}[\omega]/3\mathbb{Z}[\omega] \simeq \mathbb{F}_3[X]/(X^2).$$

□

Proposición 1.15. *Exactamente uno entre x, y y z es múltiplo de ρ.*

Demostración. Supongamos que ninguno de los tres números es múltiplo de ρ. Tendremos tres casos distintos, y, aplicando el lema anterior, podemos concluir lo siguiente:

- $x, y \equiv 1 \pmod{\rho}$: Así, $x^3, y^3 \equiv 1 \pmod{\rho^4}$ y $z^3 = x^3 + y^3 \equiv 1 + 1 \equiv -1 \pmod{\rho}$, con $z^3 \equiv -1 \pmod{\rho^4}$. Por tanto, despejando, tendríamos que $3 \equiv 0 \pmod{\rho^4}$, pero $3 = \rho\overline{\rho}$, por lo que es imposible.
- $x, y \equiv -1 \pmod{\rho}$: En este caso, $x^3, y^3 \equiv -1 \pmod{\rho^4}$ y $z^3 = x^3 + y^3 \equiv -1 - 1 = -2 \equiv 1 \pmod{\rho}$, con $z^3 \equiv 1 \pmod{\rho^4}$. De nuevo, despejando, tendríamos que $3 \equiv 0 \pmod{\rho^4}$, lo cual es imposible.
- $x \equiv 1 \pmod{\rho}$, $y \equiv -1 \pmod{\rho}$: Ahora, $x^3 \equiv 1 \pmod{\rho^4}$, $y^3 \equiv -1 \pmod{\rho^4}$ y $z^3 = x^3 + y^3 \equiv 1 - 1 = 0 \pmod{\rho}$, contradiciendo que z no es múltiplo de ρ.

Por tanto, como en todos los casos llegamos a una contradicción, tenemos que alguno tiene que ser múltiplo de ρ. Por otro lado, si $x, y \equiv 0 \pmod{\rho}$, entonces $z^3 = x^3 + y^3 \equiv 0 \pmod{\rho}$, y z también sería múltiplo, contradiciendo que los números de la solución no tienen factores comunes. Así, es claro que uno, y solo uno, es múltiplo de ρ. □

Proposición 1.16. *Supongamos que $\rho | z$. Entonces $\rho^2 | z$.*

Demostración. Como solo uno entre x, y y z es múltiplo de ρ y $\rho | z$ por hipótesis, entonces $y \equiv -1 \pmod{\rho}$ y $x \equiv 1 \pmod{\rho}$ (o al revés: $y \equiv 1 \pmod{\rho}$ y $x \equiv -1 \pmod{\rho}$). Así, $y^3 \equiv -1 \pmod{\rho^4}$ y $x^3 \equiv 1 \pmod{\rho^4}$, de donde $z^3 = x^3 + y^3 \equiv 0 \pmod{\rho^4}$. Si z fuese múltiplo de ρ, pero no de ρ^2, entonces z^3 sería múltiplo de ρ^3, pero no de ρ^4, contradiciendo lo anterior. Por lo tanto, z debe ser múltiplo de ρ^2.

□

De aquí en adelante, trabajaremos suponiendo que $\rho|z$. En caso de que no se tuviera esa condición, bastaría reescribir la ecuación 1.5 como sigue:

$$z^3 - y^3 = z^3 + (-y)^3 = x^3,$$

en caso de que $\rho|x$. El caso en que $\rho|y$ sería idéntico. Después, los resultados posteriores serían análogos.

Como bien hemos visto en la observación 1.10 y desarrollando la ecuación 1.5, tenemos que:

$$x^3 + y^3 = (x + y)(x^2 - xy + y^2) = (x + y)(x + y\omega)(x + y\omega^2),$$

y como $\omega^3 = 1$, podemos llegar a que:

$$(x + y)(x + y\omega)(x + y\omega^2) = (x + y)(x\omega + y\omega^2)(x\omega^2 + y\omega).$$

Sea, ahora, $k \geq 2$ el menor entero para el que se cumple que $\rho^k|z$, pero $\rho^{k+1} \nmid z$.

Proposición 1.17. *Los factores $x+y$, $x\omega+y\omega^2$ y $x\omega^2+y\omega$ son congruentes módulo ρ. Además, la suma de los tres es 0.*

Demostración. Veamos las congruencias dos a dos.

- Consideremos los dos primeros factores. Ver que $x + y \equiv x\omega + y\omega^2 \pmod{\rho}$ equivale a ver que $x + y - x\omega - y\omega^2 \equiv 0 \pmod{\rho}$. Reescribiendo:

  $$x+y-x\omega-y\omega^2 = x(1-\omega)+y(1-\omega^2) = (1-\omega)[x+y(1+\omega)] = \rho[x+y(1+\omega)] \equiv 0 \pmod{\rho}.$$

- Sean ahora el primer y el tercer factor. Ver que $x + y \equiv x\omega^2 + y\omega \pmod{\rho}$ equivale a ver que $x + y - x\omega^2 - y\omega \equiv 0 \pmod{\rho}$. Ahora:

  $$x+y-x\omega^2-y\omega = x(1-\omega^2)+y(1-\omega) = (1-\omega)[x(1+\omega)+y] = \rho[x(1+\omega)+y] \equiv 0 \pmod{\rho}.$$

- Por último, tomemos los dos últimos factores. Ver que $x\omega + y\omega^2 \equiv x\omega^2 + y\omega \pmod{\rho}$ equivale a ver que $x\omega + y\omega^2 - x\omega^2 - y\omega \equiv 0 \pmod{\rho}$. Veamos:

  $$\begin{aligned} x\omega + y\omega^2 - x\omega^2 - y\omega &= x(\omega - \omega^2) + y(\omega^2 - \omega) = x\omega(1 - \omega) - y\omega(1 - \omega) \\ &= (1 - \omega)(x\omega - y\omega) = \rho(x\omega - y\omega) \equiv 0 \pmod{\rho}. \end{aligned}$$

Para demostrar que su suma vale cero, tenemos que:

$$x + y + x\omega + y\omega^2 + x\omega^2 + y\omega = x(1 + \omega + \omega^2) + y(1 + \omega + \omega^2) = x \cdot 0 + y \cdot 0 = 0.$$

□

Como $\rho|z$, entonces $\rho^3|z^3 = (x+y)(x\omega+y\omega^2)(x\omega^2+y\omega)$. De este modo, como los tres factores son congruentes con módulo ρ, cada uno de ellos es divisible por ρ.

Lema 1.18. *En las condiciones anteriores, se cumplen las siguientes afirmaciones:*

- *El factor $x + y$ es coprimo con $x + y(1 + \omega)$.*
- *El factor $x + y$ es coprimo con $x(1 + \omega) + y$.*
- *El factor $x\omega^2 + y\omega$ es coprimo con $x\omega - y\omega$.*

Demostración. Veamos cada caso por separado, aunque siguiendo una misma idea:

- Si no fueran coprimos, entonces existiría un primo p tal que $p|x + y$ y $p|x + y(1 + \omega)$. Así, tendríamos que $p|x + y(1 + \omega) - (x + y) = y\omega$, y como ω es una unidad en $\mathbb{Z}[\omega]$, entonces tendríamos que $p|y$. Sin embargo, $p|(x + y) - y = x$, contradiciendo que x e y no tienen factores comunes.
- Si no fueran coprimos, entonces existiría un primo p tal que $p|x + y$ y $p|x(1 + \omega) + y$. Así, tendríamos que $p|x(1 + \omega) + y - (x + y) = x\omega$, y como ω es una unidad en $\mathbb{Z}[\omega]$, entonces tendríamos que $p|x$. Sin embargo, $p|(x + y) - x = y$, contradiciendo que x e y no tienen factores comunes.
- Si no fueran coprimos, entonces existiría un primo p tal que $p|x\omega^2 + y\omega$ y $p|x\omega - y\omega$. La primera, como ω es una unidad, se puede escribir: $p|x\omega + y$. Así, tendríamos que $p|x\omega + y - (x\omega - y\omega) = y(1 + \omega)$. Como $1 + \omega = -\omega^2$ es una unidad del anillo, tenemos que $p|y$, por lo que también divide a x, contradiciendo que x e y no tienen factores comunes.

□

Proposición 1.19. *Dados los tres factores de la proposición 1.17, ρ es su máximo común divisor.*

Demostración. Veámoslo para los factores dos a dos:

- Sea $d = \gcd(x + y, x\omega + y\omega^2)$. Entonces, tenemos que $d|x + y$ y $d|x\omega + y\omega^2$, luego:

$$d|(x + y) - (x\omega + y\omega^2) = \rho(x + y(1 + \omega)).$$

  Como $d|x + y$ y $x + y$ es coprimo con $x + y(1 + \omega)$, entonces $d \nmid (x + y(1 + \omega))$, por lo que $d|\rho$. Por otro lado, como ρ es divisor de cada uno de los factores, $\rho|d$, y así $d = \rho$.

- Sea $d = \gcd(x + y, x\omega^2 + y\omega)$. Entonces, tenemos que $d|x + y$ y $d|x\omega^2 + y\omega$, luego:

$$d|(x + y) - (x\omega^2 + y\omega) = \rho(x(1 + \omega) + y).$$

  Como $d|x + y$ y $x + y$ es coprimo con $x(1 + \omega) + y$, entonces $d \nmid (x(1 + \omega) + y)$, por lo que $d|\rho$. Por otro lado, como ρ es divisor de cada uno de los factores, $\rho|d$, y así $d = \rho$.

- Sea $d = \gcd(x\omega^2 + y\omega, x\omega + y\omega^2)$. Entonces, tenemos que $d|x\omega^2 + y\omega$ y $d|x\omega + y\omega^2$, luego:

$$d|(x\omega + y\omega^2) - (x\omega^2 + y\omega) = \rho(x\omega - y\omega).$$

  Como $d|x\omega^2 + y\omega$ y $x\omega^2 + y\omega$ es coprimo con $x\omega - y\omega$, entonces $d \nmid (x\omega - y\omega)$, por lo que $d|\rho$. Por otro lado, como ρ es divisor de cada uno de los factores, $\rho|d$, y así $d = \rho$.

□

Con todo esto, ahora podremos reescribir los factores anteriores como:

$$x + y = A\rho,\ x\omega + y\omega^2 = B\rho,\ x\omega^2 + y\omega = C\rho.$$

Además, por los resultados anteriores, tendremos que $A + B + C = 0$, donde A, B y C son coprimos dos a dos. También se cumple que $z^3 = \rho^3 ABC$, por lo que cada uno será, salvo unidad, un cubo. Podemos escribir, por tanto:

$$A = \alpha\zeta^3,\ B = \beta\eta^3,\ C = \gamma\xi^3,$$

donde α, β y γ son unidades de $\mathbb{Z}[\omega]$ y ζ, η y ξ son coprimos.

Como $\rho^k|z$ y k es el mayor exponente que lo verifica, entonces es claro que $\rho^{k-1}|(z/\rho)$ y $\rho^k \nmid (z/\rho)$. Con esto, $ABC = (z/\rho)^3$ es divisible por $\rho^{3(k-1)}$.

Como bien hemos visto en la proposición 1.16, $\rho^4|z^3$, de donde se sigue que $\rho|(z/\rho)^3 = ABC$. Así, es claro que uno entre $\zeta,\ \eta$ y ξ es divisible por ρ, y supondremos, sin pérdida de generalidad, que es ξ. También, por ser unidades, tendremos que $\alpha\beta\gamma = \pm 1$, dado que 1 y -1 son las únicas unidades que son cubos en $\mathbb{Z}[\omega]$.

Ahora, como $\xi^3|ABC = (z/\rho)^3$, tendremos que $\rho^{k-1}|\xi$ y $\rho^k \nmid \xi$. Por otro lado, al ser coprimos con ξ, podemos garantizar que $\eta, \zeta \equiv \pm 1 \pmod{\rho}$, de donde $\eta^3, \zeta^3 \equiv \pm 1 \pmod{\rho^4}$ por el lema 1.14. En concreto, también se cumple que $\eta^3, \zeta^3 \equiv \pm 1 \pmod{\rho^3}$.

Proposición 1.20. *En estas condiciones,* $\alpha \equiv \pm\beta \pmod{\rho^3}$.

Demostración. Como $A+B+C = 0$, tenemos que $A+B+C \equiv 0 \pmod{\rho^3}$. Por otra parte, como $\rho|\xi$, podemos escribir que $C = \gamma\xi^3 \equiv 0 \pmod{\rho^3}$, con lo que $A+B \equiv 0 \pmod{\rho^3}$. Desarrollando ahora la expresión anterior: $\alpha\zeta^3 \equiv -\beta\eta^3 \pmod{\rho^3}$, y como $\eta^3, \zeta^3 \equiv \pm 1 \pmod{\rho^3}$, entonces tenemos que $\alpha \equiv \pm\beta \pmod{\rho^3}$. □

Como α y β son unidades en $\mathbb{Z}[\omega]$, llegamos a que la única forma de tener $\alpha \equiv \pm\beta \pmod{\rho^3}$ es que $\alpha = \pm\beta$.

Observación 1.21. Recordemos que las unidades en $\mathbb{Z}[\omega]$ son $1, -1, \omega, -\omega, \omega^2, -\omega^2$, cuyas diferencias no son múltiplos de ρ^3:

- $1 - (-1) = 2$ no es múltiplo de ρ,
- $1 - \omega = \rho$ no es múltiplo de ρ^2,
- $1 - (-\omega) = 1 + \omega$ no es múltiplo de ρ,
- $1 - \omega^2 = (1 + \omega)(1 - \omega)$ no es múltiplo de ρ^2...

Con esto, la igualdad $\alpha\beta\gamma = \pm 1$ se puede reescribir como $\pm\alpha^2\gamma = \pm 1$, y multiplicando por α a ambos lados: $\alpha\alpha^2\gamma = \gamma = \pm\alpha$.

Hemos llegado, con todo, a que se cumple:

$$\alpha\zeta^3 \pm \alpha\eta^3 \pm \alpha\xi^3 = 0,$$

y como α es una unidad, es claro que

$$\zeta^3 \pm \eta^3 \pm \xi^3 = 0,$$

que sería otra solución para la ecuación 1.5 para la que $\rho^{k-1} | \xi$, pero $\rho^k \nmid \xi$. Esto contradice la elección de la solución inicial, pues habíamos pedido de hipótesis que no tuviera factores comunes y fuese lo más simplificada posible. Por tanto, no existirá tal solución y solo existe la trivial.

Capítulo 2

Teoría algebraica de números

Llegados a este punto, hemos encontrado una forma general para las soluciones de la ecuación con $n = 2$. También hemos comprobado que el teorema se cumple para los exponentes 3 y 4 (y, por lo tanto, para todos sus múltiplos). Sin embargo, ante la imposibilidad de hacer una prueba para cada primo, debemos intentar dar una demostración que abarque más casos.

Con esto, nos centraremos ahora en los denominados *primos regulares*. Para ello, y antes de dar la definición de dicho concepto, necesitamos introducir previamente algunas nociones sobre cuerpos de números y anillos de enteros. En este caso, utilizaremos principalmente las referencias [7], [9] y [10].

2.1. Cuerpos de números y anillos de enteros

Definición 2.1. Un cuerpo de números es una extensión de cuerpos de $\mathbb{Q}$ de orden finito.

Definición 2.2. Dado $K|\mathbb{Q}$ un cuerpo de números, se define su anillo de enteros, $\mathcal{O}_K$, como

$\mathcal{O}_K = \{\alpha \in K \mid \text{ existen } a_{n-1}, \ldots, a_0 \in \mathbb{Z} \text{ con } \alpha^n + a_{n-1}\alpha^{n-1} + ... + a_0 = 0, \text{ con } n \text{ el grado de } \alpha\}$.

Ejemplo 2.3. Vamos a ilustrar la definición anterior con los siguientes ejemplos:

- Si $K = \mathbb{Q}(i)$, entonces $\mathcal{O}_K = \mathbb{Z}[i]$.
- Si $K = \mathbb{Q}(\sqrt{-3})$, entonces $\mathcal{O}_K = \mathbb{Z}[\omega] = \mathbb{Z}[\frac{-1+\sqrt{-3}}{2}]$ es el anillo de los enteros de Eisenstein, con el que hemos trabajado en el capítulo anterior.

Además, sobre el anillo de enteros de un cuerpo de números se tienen las siguientes propiedades, que nos resultarán de gran utilidad. La primera de ellas, de hecho, justifica el porqué del nombre escogido.

Proposición 2.4. *Se tiene que $\mathcal{O}_K$ es un anillo.*

Demostración. Véase en la referencia [7, Cap.2,3]. □

Observación 2.5. Nótese que $\mathcal{O}_K$ no tiene por qué ser un dominio de factorización única (DFU) y, en particular, tampoco un dominio de ideales principales (DIP). Basta considerar, por ejemplo, el anillo $\mathbb{Z}[\sqrt{-5}]$, en el que se tiene la siguiente igualdad:

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}),$$

donde los factores 2, 3 y $1 \pm \sqrt{5}$ son irreducibles todos ellos.

El siguiente resultado, que será una de las claves para el próximo capítulo, afirma que, si bien $\mathcal{O}_K$ no tiene por qué ser un dominio de factorización única, sí se cumple que todo ideal no nulo se puede escribir de forma única como producto de ideales primos no nulos. Es decir, los resultados sobre factorización tienen que trabajarse a nivel de ideales, y no elemento a elemento.

Proposición 2.6 (Dedekind)**.** *En $\mathcal{O}_K$, todo ideal no nulo se puede escribir de forma única como producto de ideales primos.*

Demostración. Demostración disponible en [9, Cap. 3]. □

2.2. El grupo de clases

Vamos a introducir en esta sección la noción de grupo de clases, que desempeñará un papel crucial en la demostración del teorema de Fermat para primos regulares.

Definición 2.7. Sea K un cuerpo de números. Se dice que I es un ideal fraccionario de $\mathcal{O}_K$ si es un $\mathcal{O}_K$-submódulo de K para el que existe $\alpha \in \mathcal{O}_K$ de modo que $\alpha I \subset \mathcal{O}_K$.

Definición 2.8. En las condiciones anteriores, un ideal principal J de $\mathcal{O}_K$ es un $\mathcal{O}_K$-submódulo de K que está generado por un elemento, es decir, $J = (\alpha)$.

Ahora, tomando en $\mathcal{O}_K$ los ideales fraccionarios no nulos, podemos definir una operación de grupo con la multiplicación. Así, los ideales principales forman un subgrupo y podemos considerar la siguiente relación de equivalencia.

Definición 2.9. Definimos una relación de equivalencia de la siguiente forma: dados dos ideales I, J, diremos que $I \sim J$ si existen $\alpha, \beta \in \mathcal{O}_K$ verificando que $(\alpha I) = (\beta J)$.

Esta definición da sentido a la siguiente:

Definición 2.10. El grupo de clases de un cuerpo de números es el cociente:

$$\mathrm{Cl}(K) = \{\text{ideales fraccionarios no nulos}\}/\{\text{ideales principales no nulos}\}$$

Teorema 2.11 (Dirichlet, Minkowski)**.** *Sea K un cuerpo de números. Se tiene que* $\mathrm{Cl}(K)$ *es un grupo finito.*

Demostración. Está probado en [7, Cap. 2,3]. □

Ahora, nos interesa definir el concepto de primo regular, en torno al cual se construirá la demostración del capítulo siguiente.

Definición 2.12. Dado p un número primo, se dice que es regular si se verifica que

$$p \nmid |\mathrm{Cl}(\mathbb{Q}(\zeta_p))|.$$

Ejemplo 2.13. Veamos algunos ejemplos para entender las definiciones previas:

- Si $n = 2$, tenemos que $\zeta_2 = e^{2\pi i/2} = e^{\pi i} = -1$, luego $\mathbb{Q}(\zeta_2) = \mathbb{Q}$. Por tanto, $\mathcal{O}_K = \mathbb{Z}$, todos los ideales son fraccionarios y principales (estamos en un DIP), de donde $|\mathrm{Cl}(\mathbb{Q}(\zeta_p))| = 1$. Es evidente que $2 \nmid 1$, por lo que 2 es un primo regular.
- En particular, como en el caso anterior, si $\mathbb{Q}(\zeta_p)$ es un DIP, entonces tendremos que todos los ideales fraccionarios serán principales. Así, $|\mathrm{Cl}(\mathbb{Q}(\zeta_p))| = 1$, en cuyo caso, es claro que $p \nmid 1$ y p es un primo regular.

Observación 2.14. Todos los primos menores que 100 son regulares, salvo los números 37 (pues $|\mathrm{Cl}(\mathbb{Q}(\zeta_{37}))| = 37$), 59 (ya que $|\mathrm{Cl}(\mathbb{Q}(\zeta_{59}))| = 359233$) y 67 (porque $|\mathrm{Cl}(\mathbb{Q}(\zeta_{67}))| = 6712739$). Además, hay infinitos primos no regulares y se conjetura que existen también infinitos primos regulares.

2.3. Cuerpos ciclotómicos

Por último, para poder trabajar con los primos regulares, también introduciremos la noción de cuerpo ciclotómico y algunas de sus propiedades.

Definición 2.15. Dado n un número entero, denotamos por $\zeta_n = e^{\frac{2\pi i}{n}}$ a la raíz n-ésima de la unidad. Entonces, la extensión ciclotómica n-ésima se define como la extensión finita de $\mathbb{Q}$, $K_n = \mathbb{Q}(\zeta_n)$.

Proposición 2.16. *El grupo de Galois de las extensiones ciclotómicas es tal que*

$$\mathrm{Gal}\,(\mathbb{Q}(\zeta_n)|\mathbb{Q}) \simeq (\mathbb{Z}/n\mathbb{Z})^{\times},$$

donde este último es el grupo de unidades enteras módulo n.

Demostración. Disponible en [7, Cap. 2]. □

Ahora, centrándonos en el caso que nos concierne, supongamos que $n = p$ es un número primo. Esto nos permite obtener el siguiente resultado:

Proposición 2.17. *Sean p un primo, ζ_p una raíz p-ésima de la unidad y $K = \mathbb{Q}(\zeta_p)$. Entonces, se tiene que $\mathcal{O}_K = \mathbb{Z}[\zeta_p]$.*

Demostración. Véase la referencia [9, Cap. 6]. □

Capítulo 3

El teorema de Fermat para primos regulares

Una vez discutido el teorema de Fermat para $n = 3$ y $n = 4$, pasamos ahora a tratar un caso más general. Para esta parte, hemos seguido principalmente [1].

3.1. Lemas previos

En este capítulo vamos a realizar la demostración del último teorema de Fermat para primos regulares. Consideraremos, por tanto, la ecuación

$$x^p + y^p = z^p, \tag{3.1}$$

donde p es un primo regular. Dividiremos la prueba en dos casos:

- Caso I: ninguno de los enteros x, y, z es múltiplo de p.
- Caso II: exactamente uno de los enteros x, y, z es múltiplo de p.

Observamos que eso es suficiente: si los tres fuesen divisibles por p, entonces, dividiendo por la mayor potencia de p que divide a los tres, podemos llegar a que al menos uno de ellos no es múltiplo de p; si exactamente dos son múltiplos de p, a partir de la ecuación dada tenemos que el tercero también lo sería.

De ahora en adelante, denotaremos $\zeta := \zeta_p$ cuando no quepa lugar a confusión, por simplicidad. Veamos un par de lemas que utilizaremos más adelante en este capítulo.

Lema 3.1. *En el anillo $\mathbb{Z}[\zeta]$ se verifican:*

1. $(1-\zeta) = (1-\zeta^2) = \cdots = (1-\zeta^{p-1})$ *y* $1+\zeta$ *es una unidad.*
2. $(1-\zeta)$ *es el único ideal primo que divide a* p*. Además,* $p = u(1-\zeta)^{p-1}$*, con* u *una unidad.*

Demostración. Recordemos que estamos trabajando en el anillo $\mathbb{Z}[\zeta]$.

1. Veamos que cada uno de los ideales $(1-\zeta^j)$ coincide con el ideal $(1-\zeta)$:
   Por un lado, $1-\zeta^j = (1-\zeta)(1+\zeta+\cdots+\zeta^{j-1})$.
   Por otro lado, dado $j \in \{1,\dots,p-1\}$, sabemos que existe k de modo que $jk \equiv 1 \pmod{p}$.
   Entonces, en $\mathbb{Z}[\zeta]$, tenemos que $\zeta^1 = \zeta^{jk}$. Así, basta ver que:
   $$\frac{1-\zeta^{jk}}{1-\zeta^j} = 1+\zeta^j+\cdots+\zeta^{(k-1)j}.$$
   De este modo, como $(1-\zeta)$ es múltiplo de $(1-\zeta^j)$ y viceversa, los ideales que generan son el mismo.
   Para comprobar que $1+\zeta$ es una unidad, basta tener en cuenta que $1-\zeta^2 = (1-\zeta)(1+\zeta)$.
   De aquí, como $(1-\zeta) = (1-\zeta^2)$, es claro que $1+\zeta$ debe ser una unidad.

2. Consideramos el polinomio:
   $$X^p - 1 = (X-1)(X^{p-1}+\cdots+X+1) = (X-1)\prod_{i=1}^{p-1}(X-\zeta^i).$$
   Sustituyendo $X=1$ en el segundo factor, tenemos que
   $$p = \prod_{i=1}^{p-1}(1-\zeta^i),$$
   por lo que se cumple que $(p) = (1-\zeta)(1-\zeta^2)\dots(1-\zeta^{p-1})$ como producto de ideales.
   Ahora, por el apartado anterior, podemos reescribir $(p) = (1-\zeta)^{p-1}$.
   Sabiendo que $\zeta^{p-1} = -\zeta^{p-2} - \cdots - \zeta - 1$, podemos tomar el siguiente anillo:
   $$\mathbb{Z}[\zeta]/(p) = \left\{a_0 + a_1\zeta + \cdots + a_{p-2}\zeta^{p-2}, \text{con } a_0,\dots,a_{p-2} \in \{0,1,\dots,p-1\}\right\},$$
   cuyo cardinal es p^{p-1}.
   Por otro lado, en el anillo $\mathbb{Z}[\zeta]/(1-\zeta)$ se tiene que $1 \equiv \zeta$ y $p \equiv 0$, luego se verifica que:
   $$a_0 + a_1\zeta + \cdots + a_k\zeta^k \equiv a_0 + a_1 + \cdots + a_k \pmod{1-\zeta}.$$
   Como $(p) = (1-\zeta)^{p-1}$, entonces el anillo $\mathbb{Z}[\zeta]/(1-\zeta)$ tiene p elementos, de donde deducimos que es un cuerpo. Esto, a su vez, implica que $(1-\zeta)$ es un ideal maximal y, en consecuencia es primo, como queríamos demostrar.

□

Observación 3.2. La demostración anterior se hace considerando que las raíces $1, \zeta, \ldots, \zeta^{p-2}$ son linealmente independientes tanto sobre $\mathbb{Q}$ como sobre $\mathbb{Z}/p\mathbb{Z}$. Sin embargo, esto no es trivial, por lo que lo probaremos en el siguiente resultado.

Proposición 3.3. *Las raíces $1, \zeta, \ldots, \zeta^{p-2}$ son linealmente independientes en el anillo $\mathbb{Z}/p\mathbb{Z}$.*

Demostración. Para probar este resultado, hay que tener en cuenta el siguiente isomorfismo:

$$\mathbb{Z}[\zeta] \cong \mathbb{Z}[X]/\Phi_p(X),$$

donde como de costumbre $\Phi_p(X)$ denota el polinomio ciclotómico. Además, también es importante notar que $\Phi_p(X) = \frac{X^p-1}{X-1} \equiv \frac{(X-1)^p}{X-1} = (X-1)^{p-1} \pmod{p}$. De este modo, por la definición de cociente, tenemos que:

$$\mathbb{Z}[\zeta]/p = \mathbb{Z}[X]/(p, \Phi_p(X)) = (\mathbb{Z}[X]/p)/(\Phi_p(X)) = \mathbb{F}_p[X]/(X-1)^{p-1}.$$

De esto último se deduce que los elementos $1, \zeta, \ldots, \zeta^{p-2}$ son linealmente independientes sobre el cuerpo $\mathbb{F}_p$. □

Lema 3.4. *Sea $v \in \mathbb{Z}[\zeta]^\times$. Entonces, $v/\overline{v}$ es una raíz de la unidad.*

Demostración. Véase la referencia [14, Cap. 1]. □

3.2. Demostración del caso I

Inicialmente, veamos un ejemplo de cómo se puede abordar el problema.

Ejemplo 3.5. Aunque ya está demostrado, volvamos a considerar el caso de $n = p = 3$, con su correspondiente ecuación $x^3 + y^3 = z^3$. En concreto, de haber una solución, se tiene que $x^3 + y^3 \equiv z^3 \pmod{9}$. Además, supongamos que $3 \nmid xyz$, por lo que $3 \nmid x$. Entonces, x podrá tomar el conjunto de valores $\{1, 2, 4, 5, 7, 8\}$. Trabajando módulo nueve, tenemos que

$$1^3 \equiv 1,\ 2^3 \equiv 8,\ 4^3 \equiv 1,\ 5^3 \equiv 8,\ 7^3 \equiv 1,\ 8^3 \equiv 8 \pmod{9};$$

luego x^3 es congruente con 8 o 1 en estas condiciones. De modo análogo, se tiene lo mismo para y^3. Así, podemos encontrarnos en los siguientes casos:

- $x^3, y^3 \equiv 1 \pmod{9}$, de donde $z^3 \equiv 2 \pmod{9}$.
- $x^3, y^3 \equiv 8 \pmod{9}$, de donde $z^3 \equiv 7 \pmod{9}$.

- $x^3 \equiv 8,\ y^3 \equiv 1 \pmod 9$ o viceversa, de donde $z^3 \equiv 0 \pmod 9$, contradiciendo que z no es múltiplo de 3.

Hemos llegado a que z^3 es congruente con 2 o 7 módulo 9. Sin embargo, no hay ningún cubo de un número entero que verifique esto, por lo que no puede existir la solución a la ecuación inicial.

Observación 3.6. Tomando ahora el número primo 5, el procedimiento seguido en el ejemplo anterior también nos lleva, aunque con cálculos algo más tediosos, a que tampoco hay soluciones para la ecuación $x^5 + y^5 = z^5$. No obstante, el método ya no funciona para el exponente 7.

Al comprobar que trabajar con los números enteros no es suficiente, lo que haremos será emplear las extensiones ciclotómicas de la forma $K = \mathbb{Q}(\zeta)$, cuyos anillos de enteros son (por un resultado visto en el capítulo anterior) $\mathcal{O}_K = \mathbb{Z}[\zeta]$. Por tanto, en lo que sigue y salvo que se diga lo contrario, trabajaremos en $\mathbb{Z}[\zeta]$.

Entonces, dado p un primo regular, supongamos que se verifica la ecuación

$$x^p + y^p = z^p, \tag{3.2}$$

con x, y, z enteros y coprimos entre sí. Además, como hemos dicho, asumamos que $p \nmid xyz$.

Inicialmente, de forma análoga a lo realizado en el capítulo 1 para el caso $p = 3$, desarrollamos la anterior ecuación:

$$\begin{aligned} z^p = x^p + y^p = y^p \left(\left(\frac{x}{y} \right)^p + 1 \right) &= y^p \left(\left(1 + \frac{x}{y} \right) \left(\zeta + \frac{x}{y} \right) \dots \left(\zeta^{p-1} + \frac{x}{y} \right) \right) \\ &= (y + x)\left(\zeta y + x\right)\left(\zeta^2 y + x\right) \dots \left(\zeta^{p-1} y + x\right) = \prod_{j=0}^{p-1} (\zeta^j y + x). \end{aligned} \tag{3.3}$$

A continuación, veremos que estos factores son todos primos relativos entre ellos.

Proposición 3.7. *Los factores $(x + \zeta^j y)$, con $j \in \{0, 1, \dots, p-1\}$, son coprimos entre sí.*

Demostración. Supongamos que no fuese cierto, es decir, que existen $j, k \in \{0, 1, \dots, p-1\}$ con $j \neq k$ y q de modo que:

$$q | x + \zeta^j y,\ q | x + \zeta^k y.$$

Entonces, también tendremos que $q | (x + \zeta^j y) - (x + \zeta^k y) = y(\zeta^j - \zeta^k)$. Como $q \nmid y$ (si $q|y$, como $q|(x + \zeta^j y)$, se cumpliría que $q|x$, contradiciendo que x, y no tienen factores comunes), tiene que darse que $q | \zeta^j - \zeta^k$. Tomando, sin pérdida de generalidad, $k = j + \beta$, podemos reescribir lo anterior de la siguiente manera:

$$q | \zeta^j - \zeta^k = \zeta^j - \zeta^{j+\beta} = \zeta^j (1 - \zeta^\beta).$$

Como ζ^j es una unidad, deducimos que $q|1-\zeta^\beta$. De aquí, por el lema 3.1, se sigue que $q|1-\zeta$ y, por tanto, $q = 1-\zeta$. Por otro lado, como $q|x+\zeta^j y$ y $x+\zeta^j y|x^p+y^p = z^p$, tenemos que $q = 1-\zeta|z^p$. Ahora, al ser p primo entero y verificarse que $q = 1-\zeta|p$ y $q|z^p$, esto implica que $p|z^p$, lo cual es una contradicción con que $p \nmid xyz$. □

De este modo, teniendo en cuenta la ecuación 3.3 y que los factores del producto son todos coprimos por la proposición anterior, podremos escribir

$$(x+\zeta y) = \mathfrak{a}^p,$$

para algún ideal $\mathfrak{a}$. Además, de esto se deduce que $\mathfrak{a}^p$ es un ideal principal en $\mathbb{Z}[\zeta]$. Así, en el grupo de clases, $\mathfrak{a}^p$ coincidirá con el ideal neutro (1). Por otro lado, el orden de $\mathfrak{a}$ en $\mathrm{Cl}(\mathbb{Q}(\zeta_p))$ tendrá que ser 1 o p, pero como p es primo regular, este último no es posible.

Hemos llegado entonces a que $\mathfrak{a}$ será un ideal de orden 1 en el grupo de clases, por lo que también será el elemento neutro y, en consecuencia, un ideal principal. Entonces podremos escribir $\mathfrak{a} = (t)$, con $t \in \mathbb{Z}[\zeta]$, de donde se sigue que $(x+\zeta y) = \mathfrak{a}^p = (t)^p = (t^p)$. Más concretamente, podemos tomar u unidad en $\mathbb{Z}[\zeta]$ tal que

$$x+\zeta y = ut^p.$$

Ahora, desarrollemos la expresión de t en $\mathbb{Z}[\zeta]$:

$$t = b_0 + b_1\zeta + \cdots + b_{p-2}\zeta^{p-2}, \text{con } b_0, \ldots, b_{p-2} \in \mathbb{Z}.$$

A continuación, elevando a p y trabajando módulo p:

$$t^p \equiv (b_0)^p + (b_1\zeta)^p + \cdots + (b_{p-2}\zeta^{p-2})^p \equiv b_0 + b_1 + \cdots + b_{p-2} \pmod{p}.$$

Observación 3.8. Nótese que las congruencias anteriores son consecuencia del binomio de Newton, el pequeño teorema de Fermat ($a^p \equiv a \pmod{p}$) y de que ζ^j es una unidad y raíz p-ésima de 1 para cualquier $j \in \{1, \ldots, p-2\}$.

De este modo, trabajando módulo p, como $b_0, \ldots, b_{p-2} \in \mathbb{Z}$, tenemos que $t^p \equiv \overline{t}^p$. Además, por el lema 3.4, vemos que $u/\overline{u} = \pm\zeta^j$ para alguna $j \in \{0, 1, \ldots, p-1\}$, pues es una raíz de la unidad en $\mathbb{Z}[\zeta]$.

Fijemos un valor de j y un signo, esto es, $u/\overline{u} = \zeta^j$. Entonces, se cumple

$$x+\zeta y = ut^p = \zeta^j \overline{u}t^p \equiv \zeta^j \overline{u}\overline{t}^p = \zeta^j \overline{ut^p} = \zeta^j \overline{x+\zeta y} = \zeta^j(x+\overline{\zeta}y) \pmod{p}.$$

Teniendo en cuenta que $\overline{\zeta} = \zeta^{-1}$, podemos reescribir la expresión anterior como

$$x+\zeta y-\zeta^j x-\zeta^{j-1}y \equiv 0 \pmod{p}.$$

Análogamente, en el caso de fijar $u/\overline{u} = -\zeta^j$, llegamos a que

$$x + \zeta y + \zeta^j x + \zeta^{j-1} y \equiv 0 \pmod{p}.$$

Será suficiente, por tanto, probar que ninguna de estas congruencias tiene solución para los distintos valores de $j \in \{0, 1, \ldots, p-1\}$ cuando x, y son enteros coprimos entre sí y con p para alcanzar una contradicción.

Ya hemos visto en la proposición 3.3 que las raíces $1, \zeta, \ldots, \zeta^{p-2}$ son linealmente independientes. Por tanto, para los valores de j distintos de $0, 1, 2, p-1$, es imposible encontrar una combinación de las raíces congruente con 0, llegando así a la contradicción deseada.

Ahora, solo falta ver qué ocurre en los cuatro casos concretos determinados por dichos valores de j. Por comodidad, además, supondremos que el primo regular p considerado es tal que $p \geq 5$, pues ya hemos probado que para 3 no hay soluciones, tanto en el ejemplo 3.5 como en el capítulo 1.

Proposición 3.9. *En $\mathbb{Z}[\zeta]$, para $j = p-1$, la congruencia $x + \zeta y - \zeta^j x - \zeta^{j-1} y \equiv 0 \pmod{p}$ no tiene soluciones en los enteros.*

Demostración. De tenerla, teniendo presente que $1+\zeta+\cdots+\zeta^{p-2}+\zeta^{p-1} = 0$, podemos desarrollar el lado izquierdo de la siguiente forma:

$$\begin{aligned} x + \zeta y - \zeta^j x - \zeta^{j-1} y &= x(1 - \zeta^{p-1}) + y(\zeta - \zeta^{p-2}) \\ &= x(1 - (-1 - \zeta - \cdots - \zeta^{p-2})) + y(\zeta - \zeta^{p-2}) \\ &= 2x + (x+y)\zeta + x(\zeta^2 + \cdots + \zeta^{p-3}) + (x-y)\zeta^{p-2} \equiv 0 \pmod{p}. \end{aligned}$$

Sin embargo, de ser x, y no nulos, esto contradiría la independencia lineal de las raíces $1, \zeta, \ldots, \zeta^{p-2}$, por lo que no hay soluciones no triviales. □

Observación 3.10. En caso de considerar la otra congruencia, $x + \zeta y + \zeta^j x + \zeta^{j-1} y \equiv 0 \pmod{p}$, con el procedimiento de antes llegamos a que

$$x + \zeta y + \zeta^j x + \zeta^{j-1} y = (y-x)\zeta - x(\zeta^2 + \cdots + \zeta^{p-3}) + (y-x)\zeta^{p-2} \equiv 0 \pmod{p},$$

obteniendo a la misma contradicción (basta fijarse en el coeficiente de, por ejemplo, ζ^2).

Proposición 3.11. *En $\mathbb{Z}[\zeta]$, para $j = 0$, la congruencia $x + \zeta y - \zeta^j x - \zeta^{j-1} y \equiv 0 \pmod{p}$ no tiene soluciones en los enteros.*

Demostración. De tenerla, podemos desarrollar el lado izquierdo de la siguiente forma:

$$x + \zeta y - \zeta^j x - \zeta^{j-1} y = x(1-1) + y(\zeta - \zeta^{-1}) = y(\zeta - \zeta^{-1}) \equiv 0 \pmod{p},$$

y multiplicando por ζ en ambos lados, se sigue que

$$y(\zeta^2 - 1) \equiv 0 \pmod{p}.$$

A continuación, como y no es múltiplo de p, tendríamos que $(\zeta^2-1) \equiv 0 \pmod{p}$, contradiciendo la independencia lineal de 1 y ζ^2. □

Observación 3.12. En caso de considerar la otra congruencia, $x+\zeta y+\zeta^j x+\zeta^{j-1}y \equiv 0 \pmod{p}$, llegaríamos a que

$$x + \zeta y + \zeta^j x + \zeta^{j-1}y = 2x + (\zeta + \zeta^{-1})y \equiv 0 \pmod{p},$$

y multiplicando por ζ veríamos que $2x\zeta + y\zeta^2 + y \equiv 0 \pmod{p}$, obteniendo de nuevo una contradicción.

Proposición 3.13. *En $\mathbb{Z}[\zeta]$, para $j = 2$, la congruencia $x + \zeta y - \zeta^j x - \zeta^{j-1}y \equiv 0 \pmod{p}$ no tiene soluciones en los enteros.*

Demostración. De tenerla, podemos desarrollar el lado izquierdo de la siguiente forma:

$$x + \zeta y - \zeta^j x - \zeta^{j-1}y = x(1 - \zeta^2) + y(\zeta - \zeta) = x(1 - \zeta^2) \equiv 0 \pmod{p}.$$

A continuación, como x no es múltiplo de p, tendríamos que $(1-\zeta^2) \equiv 0 \pmod{p}$, contradiciendo la independencia lineal de 1 y ζ^2. □

Observación 3.14. En caso de considerar la otra congruencia, $x+\zeta y+\zeta^j x+\zeta^{j-1}y \equiv 0 \pmod{p}$, llegaríamos a que

$$x + \zeta y + \zeta^j x + \zeta^{j-1}y = x(1 + \zeta^2) + 2y\zeta \equiv 0 \pmod{p},$$

obteniendo otra contradicción con la independencia lineal de $1, \zeta, \zeta^2$.

Proposición 3.15. *En $\mathbb{Z}[\zeta]$, para $j = 1$, la congruencia $x + \zeta y - \zeta^j x - \zeta^{j-1}y \equiv 0 \pmod{p}$ no tiene soluciones en los enteros.*

Demostración. De tenerla, podemos desarrollar el lado izquierdo de la siguiente forma:

$$x + \zeta y - \zeta^j x - \zeta^{j-1}y = x(1 - \zeta) + y(\zeta - 1) = (1 - \zeta)(x - y) \equiv 0 \pmod{p}.$$

Ahora, por la primera parte del lema 3.1, podemos escribir $p = u(1-\zeta)^{p-1}$, por lo que se cumple que $(1-\zeta)(x-y)$ es múltiplo de $(1-\zeta)^{p-1}$, de donde $(x-y)$ es múltiplo de $(1-\zeta)^{p-2}$, o lo que es lo mismo, $x \equiv y \pmod{(1-\zeta)^{p-2}}$.

Por otro lado, como $p \geq 3$, se tiene que $p - 2 \geq 1$. Además, al ser x, y enteros, esto fuerza a que $x \equiv y \pmod{p}$. También podemos hacer toda la demostración de forma análoga tras

intercambiar y por $-z$, obteniendo de igual modo que $x \equiv -z \pmod{p}$. Entonces, con todo esto, debería cumplirse que

$$0 = x^p + y^p - z^p \equiv 3x^p \pmod{p}.$$

Sin embargo, como $3 \nmid p$ y x es coprimo por hipótesis con p, llegamos a una contradicción. □

Observación 3.16. En caso de considerar la otra congruencia, $x+\zeta y+\zeta^j x+\zeta^{j-1}y \equiv 0 \pmod{p}$, llegaríamos a que

$$x + \zeta y + \zeta^j x + \zeta^{j-1} y = x(1+\zeta) + y(1+\zeta) = (x+y)(1+\zeta) \equiv 0 \pmod{p}.$$

Por el lema 3.1, $1+\zeta$ es una unidad, luego tendrá inverso y

$$x + y \equiv 0 \pmod{p}.$$

De esto se seguiría que $z^p = x^p + y^p \equiv (x+y)^p \equiv 0 \pmod{p}$, por lo que p dividiría a z, contradiciendo nuestra hipótesis inicial.

Con todo esto, ya hemos probado que, para cualquier valor de $j \in \{0, 1, \ldots, p-1\}$, no existen soluciones para las congruencias planteadas anteriormente en $\mathbb{Z}[\zeta]$, por lo que tampoco existirán en $\mathbb{Z}$. Así, por reducción al absurdo, podemos concluir que no existen ternas de números enteros no nulas verificando la ecuación 3.2, siempre que ninguno de esos enteros sea divisible por p, siendo p un primo regular.

3.3. Demostración del caso II

Tras probar que, para un primo regular p, no existen soluciones para la ecuación 3.2 si $p \nmid xyz$, ahora nos falta ver qué ocurre en el caso contrario. Supondremos, por tanto, que existe una solución no nula de enteros verificando la ecuación y tal que $p|xyz$. Por supuesto, sin pérdida de generalidad, podremos suponer nuevamente que p solo divide a uno de dichos enteros (en caso contrario, dividiría a los tres y podríamos dividir entre p, obteniendo una nueva terna). En concreto, tomaremos como hipótesis que $p|z$ (de no ser así, basta despejar y hacer un cambio de variable para tenerlo). Escribiremos, por tanto, $z = p^r z_0$, donde z_0 es coprimo con p y $r \geq 1$. Así, la ecuación 3.2 se puede reescribir

$$x^p + y^p + \omega(1-\zeta)^{(p-1)rp} z_0^p = 0, \tag{3.4}$$

con ω una unidad en $\mathbb{Z}[\zeta]$ y $p \nmid xyz_0$. En particular, como $(1-\zeta)$ es el único primo dividiendo a p por el lema 3.1, se cumple que $(1-\zeta) \nmid xyz_0$.

Para afrontar nuestro problema, buscaremos demostrar el siguiente resultado:

Teorema 3.17. *La ecuación general*

$$\alpha^p + \beta^p + \varepsilon(1-\zeta)^{pn}\gamma^p = 0 \tag{3.5}$$

no tiene soluciones no nulas en $\mathbb{Z}[\zeta]$*, con* $\varepsilon \in \mathbb{Z}[\zeta]^\times$*,* $n \geq 1$ *y* $(1-\zeta) \nmid \alpha\beta\gamma$*. En consecuencia, tampoco existen soluciones enteras en el caso particular de la ecuación 3.4.*

Teniendo en cuenta el desarrollo de $\alpha^p + \beta^p$ según se hizo en la ecuación 3.3, llegamos a la ecuación de ideales dada por

$$\prod_{j=0}^{p-1}(\alpha + \zeta^j\beta) = (1-\zeta)^{pn}(\gamma)^p. \tag{3.6}$$

Como γ es no nulo, es evidente que los factores $\alpha+\beta, \alpha+\zeta\beta, \ldots, \alpha+\zeta^{p-1}\beta$ también lo son. Una primera observación es que todos los factores del producto anterior son múltiplos de $1-\zeta$.

Proposición 3.18. *Los factores* $\alpha+\beta, \alpha+\zeta\beta, \ldots, \alpha+\zeta^{p-1}\beta$ *son todos múltiplos de* $(1-\zeta)$*.*

Demostración. Inicialmente, comprobemos que todos los factores son congruentes módulo $(1-\zeta)$:

$$\alpha+\beta \equiv \alpha+\zeta^j\beta \Leftrightarrow \beta(1-\zeta^j) = \beta(1-\zeta)(1+\zeta+\cdots+\zeta^{j-1}) \equiv 0 \pmod{(1-\zeta)},$$

$$\alpha+\zeta^k\beta \equiv \alpha+\zeta^j\beta \Leftrightarrow \beta(\zeta^k-\zeta^j) = \beta\zeta^j(\zeta^{k-j}-1) = \beta\zeta^j(\zeta-1)(\zeta^{k-j-1}+\cdots+1) \equiv 0 \pmod{(1-\zeta)}.$$

Además, por la ecuación 3.6, sabemos que alguno de los factores del lado izquierdo es divisible por $(1-\zeta)$. Así, visto lo anterior, todos lo son. □

Sin embargo, a partir del resultado anterior podemos observar algo más general:

$$\begin{aligned}\gcd(\alpha+\beta, \alpha+\zeta^j\beta) &= \gcd(\zeta^j\beta-\beta, \alpha+\beta) = \gcd(\beta(\zeta^j-1), \alpha+\beta)\\ &= \gcd(\beta(1-\zeta)(\zeta^{j-1}+\cdots+1), \alpha+\beta) = \gcd(\beta(1-\zeta), \alpha+\beta)\\ &= (1-\zeta)\gcd(\beta, \alpha+\beta) = (1-\zeta)\gcd(\beta, \alpha).\end{aligned}$$

Esta observación se puede abstraer de la siguiente forma.

Proposición 3.19. *Se tiene que* $\gcd(\alpha+\zeta^k\beta, \alpha+\zeta^j\beta) = (1-\zeta)\gcd(\alpha, \beta)$*.*

Demostración. Emulando el cálculo anterior, se tiene lo siguiente:

$$\begin{aligned}\gcd(\alpha+\zeta^k\beta, \alpha+\zeta^j\beta) &= \gcd(\zeta^j\beta-\zeta^k\beta, \alpha+\zeta^k\beta) = \gcd(\beta(\zeta^j-\zeta^k), \alpha+\zeta^k\beta)\\ &= \gcd(\beta\zeta^k(\zeta^{j-k}-1), \alpha+\zeta^k\beta) = \gcd(\beta\zeta^k(1-\zeta), \alpha+\zeta^k\beta)\\ &= \gcd(\beta(1-\zeta), \alpha+\zeta^k\beta) = (1-\zeta)\gcd(\beta, \alpha+\zeta^k\beta)\\ &= (1-\zeta)\gcd(\alpha, \beta).\end{aligned}$$

□

Finalmente, recordemos que, por el desarrollo de la demostración del lema 3.1, $\mathbb{Z}[\zeta]/(1-\zeta) \simeq \mathbb{Z}/p\mathbb{Z}$. De igual modo, también se cumple que $\mathbb{Z}[\zeta]/(1-\zeta)^2 \simeq (\mathbb{Z}/p\mathbb{Z})[X]/(1-X)^2$.

Proposición 3.20. *Entre los factores $\alpha+\beta, \alpha+\zeta\beta, \ldots, \alpha+\zeta^{p-1}\beta$ uno, y solo uno de ellos, es también múltiplo de $(1-\zeta)^2$.*

Demostración. Hemos visto anteriormente que todo los factores son múltiplos de $(1-\zeta)$. Además, como hay exactamente p restos posibles módulo $(1-\zeta)^2$, en caso de no ser uno de ellos múltiplo de $(1-\zeta)^2$, habrá algún resto que se debería repetir, es decir, al menos dos de ellos deberían ser congruentes con dicho módulo. Para ello, observamos que si

$$\alpha+\zeta^j\beta \equiv \alpha+\zeta^k\beta \pmod{(1-\zeta)^2},$$

entonces

$$\beta(\zeta^j-\zeta^k) = \beta\zeta^j(1-\zeta^{k-j}) \equiv \beta(1-\zeta^{k-j}) \equiv 0 \pmod{(1-\zeta)^2}.$$

De este modo, como $(1-\zeta^{k-j})$ coincide con $(1-\zeta)$ salvo una unidad, esto forzaría que β fuese divisible por $(1-\zeta)$, lo que contradice nuestra hipótesis. Así, es evidente que todos los restos deben ser distintos, y exactamente uno de los factores será múltiplo de $(1-\zeta)^2$. Los divisores comunes calculados anteriormente también ponen de manifiesto que no puede haber dos factores en estas condiciones (pues ninguno contiene al factor $(1-\zeta)$ más de una vez). □

Visto esto y volviendo a la ecuación 3.6, queda claro que $n \geq 2$, y existe un factor tal que $\alpha+\zeta^{j_0}\beta \equiv 0 \pmod{(1-\zeta)^2}$. Sin pérdida de generalidad, podemos suponer que dicho factor es $\alpha+\beta$. Además, si escribimos $\mathfrak{a} = \gcd(\alpha,\beta)$, tendremos que:

$$(\alpha+\beta) = (1-\zeta)^{pn-(p-1)}\mathfrak{a}\mathfrak{c}_0^p,$$

$$(\alpha+\zeta^j\beta) = (1-\zeta)\mathfrak{a}\mathfrak{c}_j^p.$$

Observación 3.21. Estas expresiones se tienen fijándonos en que $\gcd(\alpha,\beta)$ tiene que aparecer p veces y solo divide a $(\alpha+\beta)$ una vez. También, los términos $\mathfrak{c}_j$ denotan la parte resultante de sacar los factores $(1-\zeta)$ y $\mathfrak{a}$, y deben ser potencias de exponente p.

Por otro lado, tenemos que los ideales de la forma $\frac{(\alpha+\zeta^j\beta)}{(\alpha+\beta)}$ son principales, pues son el cociente de dos ideales principales (en el grupo de clases, están en el neutro, por lo que su cociente también está en el neutro). En concreto,

$$\frac{(\alpha+\zeta^j\beta)}{(\alpha+\beta)} = \left(\frac{\alpha+\zeta^j\beta}{\alpha+\beta}\right) = (1-\zeta)^{p-pn}\mathfrak{c}_j^p\mathfrak{c}_0^{-p},$$

de donde se sigue que $\mathfrak{c}_j^p\mathfrak{c}_0^{-p}$ es un ideal principal. Como el orden de $(\mathfrak{c}_j\mathfrak{c}_0^{-1})$ en $\mathrm{Cl}(\mathbb{Q}(\zeta_p))$ tendrá que ser 1 o p, y p es primo regular, llegamos entonces a que $(\mathfrak{c}_j\mathfrak{c}_0^{-1})$ será un ideal de orden 1,

por lo que también será el neutro y, en consecuencia, un ideal principal. Así, podemos escribir $\mathfrak{c}_j\mathfrak{c}_0^{-1} = (t_j)$, o $\mathfrak{c}_j = t_j\mathfrak{c}_0$, donde $t_j \in \mathbb{Q}(\zeta)$ y es coprimo con $1-\zeta$.

Podemos escribir entonces la ecuación de ideales

$$(\alpha + \zeta^j\beta)(\alpha+\beta)^{-1} = (t_j)^p(1-\zeta)^{-p(n-1)},$$

que se traduce en la ecuación de elementos, donde $\varepsilon \in \mathbb{Z}[\zeta]^\times$, dada por

$$\frac{\alpha+\zeta^j\beta}{\alpha+\beta} = \frac{\varepsilon_j t_j^p}{(1-\zeta)^{p(n-1)}}. \tag{3.7}$$

Por otro lado, consideremos ahora la identidad en $\mathbb{Z}[\zeta]$ dada por

$$\zeta(\alpha+\overline{\zeta}\beta) + (\alpha+\zeta\beta) - (1+\zeta)(\alpha+\beta) = 0.$$

Como ζ es una raíz p-ésima de la unidad, se cumple que $\overline{\zeta} = \zeta^{p-1}$, así que esta igualdad se cumple para cualesquiera valores de α y β, por lo que también es cierta para los que satisfacían el teorema.

Dividiendo ahora la identidad por $(\alpha+\beta)$ (que es no nulo por la existencia de una solución no trivial que hemos supuesto), llegamos a que

$$\frac{\zeta(\alpha+\zeta^{p-1}\beta)}{\alpha+\beta} + \frac{\alpha+\zeta\beta}{\alpha+\beta} - (1+\zeta) = 0,$$

y sustituyendo según la expresión obtenida en la igualdad 3.7, se sigue que

$$\frac{\zeta\varepsilon_{p-1}t_{p-1}^p}{(1-\zeta)^{p(n-1)}} + \frac{\varepsilon_1 t_1^p}{(1-\zeta)^{p(n-1)}} - (1+\zeta) = 0.$$

Tomando denominador común y limpiando los denominadores, tenemos que

$$\zeta\varepsilon_{p-1}t_{p-1}^p + \varepsilon_1 t_1^p - (1+\zeta)(1-\zeta)^{p(n-1)} = 0. \tag{3.8}$$

Ahora, para cada j, podemos escribir $t_j = x_j/y_j$, con $x_j, y_j \in \mathbb{Z}[\zeta]$. Al ser t_j coprimo con el factor $(1-\zeta)$, entonces podemos suponer que x_j e y_j también lo son (de no serlo, serían divisibles por la misma potencia de $(1-\zeta)$ y podríamos eliminar dichas potencias). Introduciendo esto en la ecuación 3.8, obtenemos la siguiente igualdad:

$$\zeta\varepsilon_{p-1}\frac{x_{p-1}^p}{y_{p-1}^p} + \varepsilon_1\frac{x_1^p}{y_1^p} - (1+\zeta)(1-\zeta)^{p(n-1)} = 0.$$

De nuevo, sin más que tomar denominadores comunes, obtenemos

$$\zeta\varepsilon_{p-1}(x_{p-1}y_1)^p + \varepsilon_1(x_1y_{p-1})^p - (1+\zeta)(1-\zeta)^{p(n-1)}(y_1y_{p-1})^p = 0. \tag{3.9}$$

Por comodidad, denotaremos $c_{p-1} = x_{p-1}y_1$, $c_1 = x_1 y_{p-1}$ y $c_0 = y_1 y_{p-1}$. Dividiendo la expresión anterior por la unidad que acompaña a c_{p-1}^p, llegamos a

$$c_{p-1}^p + \frac{\varepsilon_1}{\zeta\varepsilon_{p-1}} c_1^p - \frac{1+\zeta}{\zeta\varepsilon_{p-1}}(1-\zeta)^{p(n-1)} c_0^p = 0. \tag{3.10}$$

Esta última ecuación es notablemente similar a la descrita en el enunciado del teorema 3.17, cambiando n por $n-1$. Además, cabe destacar que el término $\frac{1+\zeta}{\zeta\varepsilon_{p-1}}$ es una unidad en $\mathbb{Z}[\zeta]$. También se tiene que c_{p-1}, c_1 y c_0 son coprimos con $(1-\zeta)$.

Para poder eliminar el coeficiente que acompaña a c_1^p, necesitamos comprobar que este último es una potencia p-ésima, de modo que podríamos meterla dentro de la potencia. De ser así, habríamos llegado a la misma ecuación planteada en el enunciado del terorema, salvo el cambio de n a $n-1$ en el exponente de $(1-\zeta)$.

Introducimos ahora un lema que nos será de gran utilidad para poder concluir este razonamiento.

Lema 3.22 (Kummer)**.** *Dados p un primo regular y u una unidad en $\mathbb{Z}[\zeta]$, si hay un $m \in \mathbb{Z}$ tal que $u \equiv m \pmod{p}$ en $\mathbb{Z}[\zeta]$, entonces u es una potencia p-ésima de una unidad en $\mathbb{Z}[\zeta]$.*

Demostración. Disponible en [14, Cap. 1]. □

Aplicando el lema, veamos que el coeficiente $\frac{\varepsilon_1}{\zeta\varepsilon_{p-1}}$ es una potencia p-ésima de una unidad. Para ello, consideraremos la igualdad 3.10 módulo p:

$$c_{p-1}^p + \frac{\varepsilon_1}{\zeta\varepsilon_{p-1}} c_1^p \equiv 0 \pmod{p} \text{ en } \mathbb{Z}[\zeta].$$

Podemos escribir c_1 y c_{p-1} en términos de la base de $\mathbb{Z}[\zeta]$ tal y como sigue:

$$c_1 = a_0 + a_1\zeta + \cdots + a_{p-2}\zeta^{p-2},$$

$$c_{p-1} = b_0 + b_1\zeta + \cdots + b_{p-2}\zeta^{p-2},$$

y desarrollando sus potencias p-ésimas módulo p tenemos

$$c_1^p \equiv a_0^p + a_1^p + \cdots + a_{p-2}^p \equiv a_0 + a_1 + \cdots + a_{p-2} \pmod{p},$$

$$c_{p-1}^p \equiv b_0^p + b_1^p + \cdots + b_{p-2}^p \equiv b_0 + b_1 + \cdots + b_{p-2} \pmod{p}.$$

De este modo, como c_1^p y c_{p-1}^p son congruentes con enteros módulo p y c_1 es coprimo con $(1-\zeta)$, podemos invertirlo y llegar a que

$$\frac{\varepsilon_1}{\zeta\varepsilon_{p-1}} \equiv -\frac{c_{p-1}^p}{c_1^p} \equiv \text{ un número entero} \pmod{p}.$$

Llegados a este punto, sin más que recurrir al lema de Kummer, queda probado que $\frac{\varepsilon_1}{\zeta\varepsilon_{p-1}}$ es una potencia p-ésima.

Hemos probado, por tanto, que si se cumple la ecuación 3.17, también debe cumplirse si cambiamos n por $n-1$, por lo que tenemos una especie de descenso infinito en el grado de la potencia $(1-\zeta)$ de tal ecuación. Sin embargo, como al inicio de la sección habíamos visto que necesariamente $n \geq 2$, esto supone una contradicción, pues la ecuación no se cumple para $n = 1$.

Así, hemos demostrado el teorema planteado al principio de la sección, poniendo en manifiesto que la ecuación 3.2 no tiene soluciones enteras no triviales.

Capítulo 4

Curvas elípticas y formas modulares

El objetivo de este capítulo es doble. Por una parte, vamos a contextualizar históricamente los resultados presentados en capítulos anteriores, enfatizando las limitaciones de los mismos. A partir de ahí, explicaremos cómo los avances en torno a la demostración del teorema motivaron el estudio de la interacción entre dos objetos que ya habían sido introducidos previamente por los matemáticos, pero que en apariencia eran muy distintos: las curvas elípticas y las formas modulares. Las referencias principales de este capítulo son [2], [8], [12] y [13].

4.1. Desarrollo histórico

Los primeros acercamientos al teorema de Fermat se produjeron durante el desarrollo de las matemáticas griegas del siglo V a.C. con las ternas pitagóricas. Aunque parece que otras culturas, como la babilónica, ya conocían algunos ejemplos, se atribuye su descubrimiento y estudio a la escuela pitagórica, al estar relacionadas con el conocido teorema de Pitágoras. Posteriormente, otros matemáticos, como Euclides o Diofanto, lograron deducir fórmulas que permitieron obtener las distintas ternas existentes, como la que se ha desarrollado en el primer capítulo. A pesar de todo esto, no fue hasta el siglo XVII que Fermat escribió el enunciado de su teorema en el margen del libro *Arithmetica* de Diofanto, acompañado de la demostración para el caso de $n = 4$ que encontramos el capítulo 1 de este trabajo.

Más adelante, muchos matemáticos emprendieron el desafío de probar el resultado para el resto de exponentes mayores o iguales que tres. La mayoría de ellos fracasó. Algunos consiguieron probar casos particulares, como es el caso de Euler y Gauss, que desarrollaron una elaborada demostración para el caso $n = 3$. El primero, aunque su prueba está incompleta, consiguió transformar la ecuación haciendo uso de los enteros de Eisenstein. El segundo, por su parte, se dedicó a desarrollar la aritmética modular.

Sin embargo, fue fundamentalmente el trabajo de los matemáticos franceses y alemanes que desarrollaron la teoría algebraica de números en el siglo XIX el que condujo a la prueba del teorema para los primos regulares. Entre ellos, debemos destacar las contribuciones de Lamé, Dirichlet y, sobre todo, Kummer. El primero de ellos, extendió la factorización del caso n=3 empleando los enteros ciclotómicos, pero falló al pensar que en todos estos anillos se tendría un dominio de factorización única. Dirichlet, por su parte, trató de corregir los vacíos y errores de Lamé, a parte de desarrollar la teoría de las extensiones ciclotómicas. Finalmente, fue Kummer el que definió los primos regulares y consiguió formalizar la prueba del teorema de Fermat para estos casos. Posteriormente, se adentró en el intento de desarrollar la teoría de números empleando ideales para tratar los casos de los primos irregulares, aunque sin éxito.

Tras publicarse esta demostración del teorema de Fermat para los primos regulares, hubo un largo período de tiempo en el que distintos matemáticos intentaron seguir las líneas de esa prueba para descifrar los casos restantes, siempre sin éxito alguno. El problema reside en que, en los primos irregulares, no encontramos un objeto que nos permita trabajar con ciertas propiedades de factorización, como era el anillo de enteros para el caso ya probado. Así, durante más de un siglo, no hubo avances y se dio por imposible encontrar una solución.

Hacia mediados del siglo XX, en un contexto totalmente independiente de la demostración del teorema de Fermat, dos matemáticos japoneses trabajaban con las curvas elípticas. Estos eran Taniyama y Shimura, que acabaron enunciando un resultado que revolucionaría todas las matemáticas. El primero, trabajando con una curva concreta y su función L, se percató de que esta última coincidía con la correspondiente función asociada a una forma modular. Esto le llevó a preguntarse si ocurriría lo mismo para el resto de curvas elípticas. Aunque, inicialmente, su teoría fue considerada vaga o especulativa, acabó formalizando con Shimura la conjetura a la que dan nombre. Esto supuso un enorme impacto, al asociar dos objetos supuestamente tan distintos: las curvas elípticas, de naturaleza geométrica, y las formas modulares, de un área más analítica.

En las siguientes dos secciones introduciremos las nociones de curva elíptica y forma modular, y en el siguiente capítulo explicaremos cómo la conexión entre ambas llegó a poder demostrar, finalmente, el último teorema de Fermat.

4.2. Curvas elípticas

Antes de introducir propiamente las curvas elípticas, vamos a incidir en el hecho de que estas son curvas planas proyectivas, es decir, son curvas que están en un espacio proyectivo de dimensión dos. Sin embargo, de cara a su estudio, se acostumbra a trabajar en el espacio afín. Esto es porque en la definición impondremos que tengan un punto sobre el cuerpo, y se puede demostrar que, de hecho, tienen un punto de inflexión. En ese caso, se puede escoger un sistema

de coordenadas en el que la curva tiene un único punto en la recta del infinito, que es precisamente el único punto de inflexión.

Observación 4.1. Para esta sección, nos será de utilidad tener una idea de lo que es una curva no singular. Informalmente, esta propiedad hace referencia al hecho de que la curva no se autointerseca (no tiene nodos) y que no puede tener "picos" o cúspides. Para las definiciones precisas, véase [4, Cap. 3].

Definición 4.2. Sea K un cuerpo. Una *curva elíptica* sobre K es una curva plana proyectiva no singular E/K, de grado 3 y con un punto base $O \in E(K)$, siendo $E(K)$ los puntos de la curva elíptica E con coeficientes en K.

Como anticipamos, se puede probar que E tiene un punto de inflexión. Tomándolo en la recta del infinito, es posible llegar a una ecuación afín de la forma

$$y^2 + y = x^3 + ax^2 + bx + c.$$

Es decir, en coordenadas proyectivas la curva se escribe como

$$Y^2Z + YZ^2 = X^3 + aX^2Z + bXZ^2 + cZ^3.$$

En la recta del infinito, que corresponde a $Z = 0$, hay un único punto, que en coordenadas proyectivas se escribe como $[0 : 1 : 0]$. Si la característica del cuerpo K es distinta de 2 o de 3, podemos llegar a una ecuación reducida de la forma

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

en la que la condición de que sea no singular se escribe como $4A^3 + 27B^2 \neq 0$. Para más detalles, véanse [13] y [2].

Proposición 4.3. *Dada una curva elíptica E, consideramos dos puntos en ella, P y Q. Entonces, la recta que pasa por tales puntos interseca a la curva E en un tercer punto, $P * Q$.*

Demostración. Véase en [13, Cap. 1]. □

Cabe destacar que, en este tipo de curvas, podemos definir una operación que las dota de una estructura de grupo. Para ello, consideramos nuevamente el "punto en el infinito", $\mathcal{O}$.

Si tomamos P y Q puntos de una curva elíptica E, consideramos el punto que se encuentra en la misma recta como indica la proposición anterior, $P * Q$. Si $P = Q$, la recta $P * Q$ se define como la tangente a la curva elíptica por el punto P. En cualquier caso, definimos la suma de los dos puntos, $P + Q$ como el punto simétrico a $P * Q$ respecto del eje horizontal.

Observación 4.4. La operación consistente en tomar el punto simétrico con respecto al eje horizontal se corresponde con considerar la intersección de la curva con respecto a la recta que también pasa por el punto del infinito $\mathcal{O}$. Por lo tanto, tenemos que

$$P + Q = \mathcal{O} * (P * Q).$$

Para dotar a la curva elíptica de una estructura de grupo, necesitamos también fijar un elemento neutro y definir el inverso de cualquier elemento. Para el neutro se toma el punto del infinito, ya que

$$P + \mathcal{O} = \mathcal{O} * (P * \mathcal{O}) = P,$$

y análogamente $\mathcal{O} + P = P$. Además, dado que gráficamente estamos entendiendo que una recta que pasa por un punto P y por $\mathcal{O}$ es una recta vertical por P, tenemos que el elemento inverso para cualquier punto se obtiene mediante la simetría por el eje horizontal.

Cabe destacar también que se cumplen las propiedades de conmutatividad y asociatividad, lo que resulta en el siguiente resultado.

Proposición 4.5. *El conjunto de puntos de una curva elíptica, $E(K)$, con la operación que hemos definido, es un grupo abeliano.*

De cara al estudio posterior de las curvas elípticas nos interesa estudiar sus propiedades sobre un cuerpo finito, que por simplicidad vamos a suponer que es $\mathbb{F}_p$, para algún primo p. El número de soluciones de una ecuación (afín) de la forma

$$y^2 = x^3 + ax + b$$

está acotado por $2p$: para cada uno de los p valores de x, y puede tomar a lo sumo dos valores, dado que sobre un cuerpo cualquiera un elemento tiene a lo sumo dos raíces cuadradas. A esos $2p$ puntos hay que sumar el punto del infinito.

Sin embargo, el siguiente teorema asegura que la cota de $2p + 1$ se puede mejorar y que el número de puntos es, aproximadamente, $p + 1$.

Teorema 4.6 (Hasse)**.** *Si E es una curva elíptica definida sobre el cuerpo finito $\mathbb{F}_p$, entonces el número de puntos de E con coordenadas en $\mathbb{F}_p$ viene dado por $p+1-a_p$, donde a_p es un término de error satisfaciendo que $|a_p| < 2\sqrt{p}$.*

Demostración. Para la demostración, se puede consultar la referencia [13, Cap. 4]. □

4.3. Formas modulares

En esta última sección del capítulo vamos a introducir el concepto de forma modular, que resulta esencial para entender la demostración del último teorema de Fermat. Para ello, veamos unas nociones necesarias, usando principalmente [3] y [8].

Definición 4.7. El semiplano de Poincaré (o semiplano positivo), se define como el subconjunto de $\mathbb{C}$ dado por $\mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\} = \{x + iy : x, y \in \mathbb{R}, y > 0\}$.

Definición 4.8. El grupo lineal general, $\mathrm{GL}_2(\mathbb{R})$, consiste en el grupo de las matrices de dimensión 2×2, con coeficientes reales, que son invertibles. Se denota por $\mathrm{GL}_2^+(\mathbb{R})$ al subgrupo de matrices con determinante positivo:

$$\mathrm{GL}_2^+(\mathbb{R}) = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}) : ad - bc > 0 \right\}.$$

Escribiremos $\mathrm{SL}_2(\mathbb{R})$ para referirnos al subgrupo de $\mathrm{GL}_2(\mathbb{R})$ formado por las matrices cuyo determinante es 1.

Observación 4.9. Podemos construir una acción $\psi\colon \mathrm{GL}_2^+(\mathbb{R}) \times \mathbb{H} \to \mathbb{H}$ como sigue. Si $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, con $\gamma \in GL_2^+(\mathbb{R})$, entonces

$$\psi(\gamma, z) := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az + b}{cz + d}.$$

Es un sencillo ejercicio ver que está bien definida. Como $\gamma \in \mathrm{GL}_2^+(\mathbb{R})$, sabemos que $ad - bc > 0$. De este modo, vemos que $\Im(\gamma \cdot z) > 0$:

$$\begin{aligned} \Im\left(\frac{az + b}{cz + d}\right) &= \Im\left(\frac{(az + b)(c\overline{z} + d)}{|cz + d|^2}\right) \\ &= \frac{\Im(ac|z|^2 + adz + bc\overline{z} + bd)}{|cz + d|^2} \\ &= \frac{(ad - bc)\Im(z)}{|cz + d|^2}. \end{aligned}$$

En base a lo anterior, tiene sentido realizar la siguiente definición.

Definición 4.10. Se define el factor de automorfía como la aplicación $j\colon \mathrm{GL}_2^+(\mathbb{R}) \times \mathbb{H} \to \mathbb{C}$, dada por

$$j(\gamma, z) = j\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, z\right) = cz + d \in \mathbb{C},$$

donde $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\gamma \in \mathrm{GL}_2^+(\mathbb{R})$.

Con estos conceptos previos, ya estamos en condiciones de introducir las siguientes nociones relativas a las formas modulares.

Definición 4.11. Una función holomorfa $f : \mathbb{H} \to \mathbb{C}$ se dice débilmente modular de peso $k \in \mathbb{Z}$ en $\mathrm{SL}_2(\mathbb{Z})$ si, para todo $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, con $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ se cumple lo siguiente:

$$f(\psi(\gamma, z)) = f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z), \ \forall z \in \mathbb{H}.$$

Ejemplo 4.12. En la definición anterior, si tomamos $\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, resulta que

$$f(\psi(\gamma, z)) = f(z+1) = f(z).$$

Con esto, podemos ver que esta es una función periódica. También veremos que admite una expansión de Fourier de la forma

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

Definición 4.13. Una función es holomorfa en el infinito si admite el desarrollo anterior con los coeficientes a_n iguales a cero para $n < 0$, esto es,

$$f(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z}.$$

Decimos que la función se anula en el infinito si, además de lo anterior, $a_0 = 0$.

Definición 4.14. Una forma modular de peso $k \in \mathbb{Z}$ se define como una aplicación $f : \mathbb{H} \to \mathbb{C}$ que verifica las siguientes condiciones:

1. Es holomorfa.
2. Es débilmente modular.
3. Es holomorfa en el infinito.

Si además $a_0 = 0$ decimos que se trata de una forma modular cuspidal.

A continuación, hablaremos brevemente de los subgrupos de congruencia, y de cómo se puede extender el concepto de función modular a ellos.

Definición 4.15. El subgrupo de congruencia principal de nivel N es el conjunto

$$\Gamma(N) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

El subgrupo de congruencia de nivel $\Gamma_0(N)$ se define como

$$\Gamma_0(N) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{N} \right\},$$

donde $*$ denota un entero arbitrario.

Aunque el ejemplo más importante que consideraremos nosotros de cara a las aplicaciones a la teoría de curvas elípticas es el de $\Gamma_0(N)$, esta noción se puede generalizar a un contexto más arbitrario.

Definición 4.16. De forma más general, un subgrupo $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ se dice subgrupo de congruencia si existe un $N \geq 1$ tal que $\Gamma(N) \subseteq \Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$. Además, se llama nivel del subgrupo de congruencia al menor N verificando que $\Gamma(N) \subseteq \Gamma$.

Observación 4.17. Es importante destacar el hecho de que toda forma modular admite una expansión en serie de Fourier, ya que siempre existirá algún N para el cual la matriz $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}$ pertenezca al subgrupo de congruencia.

A modo de notación, para una matriz $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, definimos el operador barra de peso k, $f|_k\alpha$, como

$$(f|_k\alpha)(z) = (cz + d)^{-k} f(\alpha \cdot z),$$

siendo $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Definición 4.18. Una función $f : \mathbb{H} \to \mathbb{C}$ es una forma modular de peso $k \in \mathbb{Z}$ sobre un subgrupo de congruencia Γ si:

1. f es holomorfa en $\mathbb{H}$.
2. $f|_k\gamma = f$ para todo $\gamma \in \Gamma$.
3. $f|_k\alpha$ es holomorfa en el infinito para cada $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

Si $f|_k\alpha$ se anula en el infinito para cada $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, decimos que la forma modular es cuspidal. Para más comodidad, el conjunto de formas modulares cuspidales de peso k para el subgrupo $\Gamma_0(N)$ se denota por $S_k(N)$.

Observación 4.19. La tercera condición de la definición, en el caso en el que $\alpha \in \Gamma$, simplemente indica que f es holomorfa en infinito.

Uno de los aspectos importantes de los espacios de formas modulares es que están equipados de una colección de aplicaciones lineales, llamadas operadores de Hecke, cuya teoría espectral proporciona resultados importantes sobre la estructura de estos espacios. En particular, estas aplicaciones lineales son hermíticas con respecto al llamado producto de Petersson y conmutan entre sí, por lo que existe una base en la que todos los operadores de Hecke diagonalizan a la vez. Los autovectores de estas aplicaciones lineales se llaman formas propias o autoformas y desempeñan un papel importante en la conexión de las formas modulares con las curvas elípticas.

Capítulo 5

El último teorema de Fermat: idea de la demostración

En este último capítulo de la memoria, explicaremos la idea de algunos de los ingredientes esenciales para la demostración del último teorema de Fermat, dando también una perspectiva histórica y explicando cómo se llegó a la prueba.

5.1. La noción de modularidad

Una noción clave de cara a comparar las curvas elípticas y las formas modulares es la de función L o serie L. Se puede entender como una generalización de la función zeta de Riemann, y se puede asociar tanto a una curva elíptica como a una forma modular. Sin embargo, veremos que la función L asociada a una forma modular tiene, a priori, mejores propiedades analíticas que la de una curva elíptica. Para ello, volveremos a apoyarnos en las referencias [2], [8] y [13].

Sea $E/\mathbb{Q}$ una curva elíptica con una ecuación afín de la forma

$$y^2 = x^3 + ax + b, \quad 4a^3 + 27b^2 \neq 0,$$

y supongamos también que $a, b \in \mathbb{Z}$.

Definición 5.1. Se define la función L de la curva elíptica $E/\mathbb{Q}$ como

$$L(E,s) := \prod_{p \nmid N} (1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p|N} (1 - a_p p^{-s})^{-1},$$

donde $s \in \mathbb{C}$ es la variable compleja, N es el conductor de la curva y los a_p son los términos de error introducidos en el teorema de Hasse.

Conviene observar que la definición anterior consta de un producto (infinito) con dos partes: la primera tiene un número infinito de factores, por lo que es la responsable de controlar la convergencia o no del producto. La siguiente se corresponde con un número finito de primos, correspondiente a los factores de mala reducción. De hecho, se pueden encontrar expresiones explícitas para los a_p dependiendo del tipo de reducción que tienen sobre $\mathbb{F}_p$ (es decir, dependiendo de si es un nodo o una cúspide).

Por otro lado, el teorema de Hasse permite demostrar que $L(E,s)$ converge para $\Re(s) > 3/2$. Para una discusión más detallada, véase [2, §1.4].

Sea ahora $f \in S_2(N)$ una forma modular cuspidal de nivel N (y peso $k = 2$). Consideramos la expansión de Fourier de f, que es de la forma

$$f(q) = \sum_{n=1}^{\infty} a_n q^n, \quad q = e^{2\pi i z}.$$

Definición 5.2. Se define la función L de dicha forma modular f como

$$L(f,s) := \sum_{n=1}^{\infty} a_n n^{-s} = \prod_{p \nmid N} (1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p|N} (1 - a_p p^{-s})^{-1},$$

donde $s \in \mathbb{C}$ es la variable compleja y, en este caso, los a_p son los coeficientes de su expansión de Fourier.

Para esta serie L asociada a una forma modular, se cumplen algunas propiedades que resultan de gran utilidad:

- Convergencia. Se cumple que f converge absolutamente (si $f \in S_2(N)$) para cualquier s verificando que $\Re(s) > k/2 + 1 = 2$.
- Es posible extender la función L de forma analítica a todo $\mathbb{C}$.
- La función L, así como su extensión compleja, satisfacen la siguiente ecuación funcional:

  $$\Lambda(f,s) = \pm N^{s-1} \Lambda(f, 2-s),$$

  donde el término anterior es $\Lambda(f,s) := (2\pi)^{-s}\Gamma(s)L(f,s)$, siendo $\Gamma(s)$ la función Gamma.

Se pueden consultar estas propiedades más extendidamente con sus respectivas pruebas en [8]. En todos los casos, las demostraciones son de tipo analítico y se basan en representaciones integrales.

Tras establecer estos conceptos, nos interesa ver si es posible establecer alguna relación entre las series L de una curva elíptica y una forma modular. Dicha relación podría también dotar a las series de las curvas elípticas de estas propiedades que acabamos de detallar.

Teorema 5.3 (Taniyama–Shimura)**.** *Dada $E/\mathbb{Q}$ una curva elíptica, existe una forma modular cuspidal $f \in S_2(N)$ de modo que $L(E,s) = L(f,s)$, donde N es el conductor de la curva.*

Observación 5.4. Esto se traduce en que los coeficientes a_p asociados a la serie L de la forma modular f (obtenidos a partir de la expansión de Fourier) coinciden con los coeficientes asociados a la función L de la curva elíptica (que son los términos de error).

Observación 5.5. El conductor de una curva elíptica $E/\mathbb{Q}$ es un entero positivo que proporciona una medida sobre las malas reducciones de la curva con módulo los distintos primos. Entenderemos que una reducción de la curva E módulo p es mala si presenta nodos o cúspides.

5.2. La curva elíptica de Frey

El anterior resultado, en el momento de su formulación, no fue probado, por lo que no dejaba de ser una conjetura. Esta había surgido hacia mediados del siglo XX, cuando el matemático japonés Taniyama, se fijó detenidamente en los primeros términos de la serie L de una forma modular concreta. Fue entonces cuando se dio cuenta de que estos coincidían exactamente con los de la función L de una curva elíptica conocida. Tras calcular los siguientes términos y probar con otras formas modulares, se percató de que esto seguía ocurriendo. Esto, junto con su trabajo con el también japonés Shimura (que lo respaldó y prosiguió con sus investigaciones tras su suicidio), le llevo a formalizar la conjetura.

Esto supuso una revolución en las matemáticas, y se empezaron a publicar multitud de artículos en los que se resolvían problemas suponiendo que el resultado era cierto. Así, se fue construyendo una abundancia de resultados que dependían de la certeza de la conjetura, volviendo totalmente necesario probarla.

Más adelante, en el año 1984, Gerhard Frey afirmó que probar Taniyama–Shimura era suficiente para demostrar el teorema de Fermat. Para ello, introdujo la famosa curva elíptica de Frey, como detallaremos a continuación:

Definición 5.6. Dado $p > 2$, sea una solución para la ecuación $x^p + y^p = z^p$, esto es, una terna (a,b,c) verificando que $a^p + b^p = c^p$. Entonces, se define la curva elíptica de Frey como aquella que tiene ecuación afín dada por

$$y^2 = x(x - a^p)(x + b^p).$$

Observación 5.7. La curva de Frey cumple las siguientes propiedades:

- Se trata de una curva elíptica *imaginaria*, pues nace a partir de una solución a la ecuación del teorema de Fermat, que puede no existir (y, de hecho, hoy en día sabemos que no existe).

- Esta expresión se obtiene a partir de la ecuación original de Fermat y la hipotética solución, tras una serie de tediosas manipulaciones.
- La curva es elíptica (es decir, no singular).

Además, Frey afirmó que esta curva tiene unas "extrañas propiedades" (malas reducciones para muchos primos...) que no permiten asociarla con ninguna forma modular, contradiciendo así la conjetura de Taniyama–Shimura. Sin embargo, su prueba de que tal curva no era modular no estaba completa y, tras muchos meses de investigación, fue Ken Ribet quien lo acabó demostrando correctamente con su idea de *bajar el nivel*.

El argumento de Frey también se puede enunciar al revés:

- Si se cumple Taniyama–Shimura, entonces toda curva elíptica debe ser modular.
- Como la curva de Frey no se asocia a ninguna forma modular, entonces no debe existir.
- En consecuencia, no puede existir una solución no trivial para la ecuación de Fermat, por lo que el teorema es cierto.

5.3. Andrew Wiles y Richard Taylor

Hacia finales del siglo XX, Andrew Wiles era un matemático que ejercía como profesor en la Universidad de Princeton (Nueva Jersey, Estados Unidos) y uno de los mayores expertos del mundo en el campo de las curvas elípticas. Había mostrado interés en el teorema de Fermat y acabó asumiendo el reto de demostrarlo. Así, la prueba realizada por Ribet lo llevó a centrarse en la conjetura de Taniyama–Shimura.

Decidió dejar de lado los congresos y todas las actividades que no estuvieran directamente relacionadas con su objetivo, para poder concentrarse en el problema, aunque tuvo que mantener sus tutorías y seminarios en la universidad. Entonces, Wiles se dedicó a estudiar y familiarizarse con los resultados relativos a las curvas elípticas y a las formas modulares, así como con las matemáticas construidas a partir de ellos. Terminó recluyéndose y trabajando en secreto en el ático de su casa.

Parte de la dificultad que presentaba abordar una demostración de la conjetura radicaba en la existencia de infinitas curvas elípticas y formas modulares, que no son fáciles de contar, y mucho menos de ordenar. Por ello, todos los intentos previos de realizar una demostración por inducción habían fracasado. Se intentaba probar que los infinitos términos de las series L coincidían para una curva elíptica y una forma modular concretas, pero era difícil entender con qué ecuaciones seguir el argumento.

Wiles, por su parte, y tras muchos meses de reflexión, intentó llevar a cabo el proceso de inducción de una forma diferente: procuró demostrar que, para toda curva elíptica, cada término de la serie L coincidía con el correspondiente de una forma modular. Esto le permitía establecer un orden natural para la demostración: primero tomaría el primer término y, después, lo probaría para uno arbitrario, suponiendo que los anteriores cumplían el resultado por hipótesis.

El caso $n = 1$, es decir, la prueba para el primer término de la serie, lo realizó inspirándose en resultados clásicos de la teoría de Galois. Posteriormente, para abordar el caso general, volvió a asistir regularmente a congresos y buscó resultados recientes que le permitieran atacar el problema. Fue entonces cuando conoció el llamado método de Kolygavin–Flach y comenzó a estudiarlo y adaptarlo a su trabajo. Descubrió un modo de clasificar las curvas elípticas en distintas familias y extendió dicho método de diversas formas, con el fin de poder demostrar el resultado dentro de cada familia. Así, tras siete años de trabajo, anunció en una conferencia que había demostrado la conjetura de Taniyama–Shimura y, en consecuencia, el último teorema de Fermat.

Tras dicho anuncio, los expertos en el campo comenzaron a revisar la demostración de Wiles minuciosamente, en busca de algún error que la dejase incompleta o que supusiera una contradicción. En efecto, tal error apareció, y fue entonces cuando el propio Andrew Wiles, trabajando junto con Richard Taylor (quien anteriormente había sido su alumno), lo corrigió, concluyendo así la demostración correcta del resultado. Dicha prueba definitiva apareció en 1995 en dos artículos publicados en la revista *Annals of Mathematics* [15], [11]: el primero (y más extenso), firmado en solitario, y el segundo, en colaboración con Richard Taylor, en el que explican cómo solventaron el problema que había surgido durante el primer proceso de revisión.

Bibliografía

[1] Conrad, K. (2025). *Fermat's last theorem for regular primes*, online notes.

[2] Darmon, H. (2004). *Rational Points on Modular Elliptic Curves*, American Mathematical Society.

[3] Diamond, F. y Shurman, J. (2005). *A First Course in Modular Forms*, Graduate Texts in Mathematics, Springer-Verlag, New York.

[4] Fulton, W. (1971). *Curvas algebraicas*, Editorial Reverté.

[5] Hardy, G. H. y Wright, E. M. (2009). *An Introduction to the Theory of Numbers*, 6th ed., Oxford University Press.

[6] Ireland, K. y Rosen, M. (1990). *A Classical Introduction to Modern Number Theory*, 2nd ed., Graduate Texts in Mathematics, Springer-Verlag, New York.

[7] Marcus, D. A. (2018). *Number Fields*, 2nd ed., Universitext, Springer, Cham.

[8] Masdeu, M. (2015). *Modular Forms (MA4H9)*, notes on a course at the University of Warwick.

[9] Milne, J. S. (2017). *Algebraic Number Theory (v3.07)*, online notes.

[10] Neukirch, J. (1999). *Algebraic Number Theory*, Grundlehren der Mathematischen Wissenschaften, Springer, Berlin–Heidelberg.

[11] Taylor, R. y Wiles, A. (1995). *Ring-theoretic properties of certain Hecke algebras*, Ann. Math., **141**, 553–572.

[12] Silverman, J. H. (2009). *The Arithmetic of Elliptic Curves*, 2nd ed., Graduate Texts in Mathematics, Springer-Verlag, New York.

[13] Silverman, J. H. y Tate, J. T. (2015). *Rational Points on Elliptic Curves*, 2nd ed., Universitext, Springer.

[14] Washington, L. C. (1997). *Introduction to Cyclotomic Fields*, 2nd ed., Graduate Texts in Mathematics, Springer.

[15] Wiles, A. (1995). *Modular elliptic curves and Fermat's last theorem*, Ann. Math., **141**, 443–551.