TEMat

El problema de las sumas de dos cuadrados

Alberto Cobos Rábano[a]
KU Leuven
albertocobosrabano@gmail.com

Resumen: Con el pretexto de resolver el problema clásico de la representación como suma de dos cuadrados en teoría de números, introduciremos una serie de conceptos fundamentales de este campo, como son las funciones multiplicativas o la descomposición de primos en anillos de enteros. Veremos también la relación entre la teoría de números y otras ramas de las matemáticas, pues la resolución del problema se basa en el estudio de los enteros gaussianos y de la divisibilidad en este anillo. Concluiremos demostrando qué enteros son sumas de dos cuadrados de enteros, y de cuántas maneras distintas.

Abstract: With the pretext of solving the classical number theory problem of representations as a sum of two squares, we shall introduce a series of fundamental concepts of this field, such as multiplicative functions or factorizations of primes in rings of integers. We shall also see the connection between number theory and other fields of mathematics, as the solution of the problem is based on the study of Gaussian integers and their divisibility. We shall finish by showing which integers are a sum of two squares of integers, and in how many different ways.

Palabras clave: teoría de números, sumas de dos cuadrados, función multiplicativa, enteros gaussianos, primos gaussianos.

MSC2010: 11E25.

Recibido: 23 de agosto de 2018.
Aceptado: 13 de septiembre de 2018.

Agradecimientos: Quiero agradecer al profesor Luis Manuel Navas, de la Universidad de Salamanca, todo el tiempo que me ha dedicado y todo lo que he aprendido de él en el desarrollo de mi Trabajo de Fin de Grado, del cual se extrae este artículo.

Referencia: Cobos Rábano, Alberto. «El problema de las sumas de dos cuadrados». En: *TEMat*, 3 (2019), págs. 1-16. ISSN: 2530-9633. URL: https://temat.es/articulo/2019-p1.

[a]El autor estaba afiliado a la Universidad de Salamanca (USAL) durante la realización de este trabajo.

1. Introducción

El objetivo de este artículo es resolver el siguiente problema clásico de la teoría de números.

Problema 1. Determinar qué números naturales son suma de dos cuadrados de enteros. ◂

En otras palabras, queremos determinar qué $n \in \mathbb{N}$ se pueden expresar en la forma $n = a^2 + b^2$ para ciertos $a, b \in \mathbb{Z}$. Con esto ya podemos dar nuestra primera definición.

Definición 2. Diremos que una pareja $(a, b) \in \mathbb{Z}^2$ es una **representación de $n \in \mathbb{N}$ como suma de dos cuadrados** si $a^2 + b^2 = n$. Por abuso de notación, diremos también que $a^2 + b^2$ es una **representación** de n, considerando relevante tanto el orden como el signo de a y b. ◂

Relacionado con el problema de determinar los naturales que son representables, queremos resolver también el siguiente problema.

Problema 3. Dado $n \in \mathbb{N}$ representable como suma de dos cuadrados, dar una fórmula cerrada para el número de representaciones distintas de n como suma de dos cuadrados. ◂

La primera referencia histórica a problemas de sumas de cuadrados es el problema 8 de la *Aritmética* de Diofanto, que pide escribir un cuadrado como suma de dos cuadrado. Dicho problema fue retomado por Fermat, quien enunció también el problema 1. Véase el libro de Weil [9] para más información.

Comencemos con un ejemplo para comprender los problemas 1 y 3.

Ejemplo 4. Para números «pequeños» es fácil hacer las comprobaciones pertinentes a mano. Un ejemplo de número que es suma de dos cuadrados, junto con todas las representaciones posibles, es el siguiente:

$$\begin{aligned} 10 &= 1^2 + 3^2 = (-1)^2 + 3^2 = 1^2 + (-3)^2 = (-1)^2 + (-3)^2 \\ &= 3^2 + 1^2 = 3^2 + (-1)^2 = (-3)^2 + 1^2 = (-3)^2 + (-1)^2. \end{aligned}$$

Este ejemplo nos sirve para aclarar que contamos como representaciones distintas las permutaciones y los cambios de signo; es decir, estamos contando las soluciones $(x, y) \in \mathbb{Z}^2$ de la ecuación $x^2 + y^2 = n$, para cada n fijo. También 4 es suma de dos cuadrados:

$$4 = 2^2 + 0^2 = (-2)^2 + 0^2 = 0^2 + 2^2 = 0^2 + (-2)^2.$$

En este caso existen cuatro representaciones en lugar de ocho, porque cambiar de signo a 0 es no hacer nada. También es fácil comprobar que 3 no es suma de dos cuadrados, luego el problema 1 no es trivial, pues hay números que sí son representables y otros que no. También se puede comprobar que 1996 no es suma de dos cuadrados, o que

$$\begin{aligned} 2018 &= 43^2 + 13^2 = (-43)^2 + 13^2 = 43^2 + (-13)^2 = (-43)^2 + (-13)^2 \\ &= 13^2 + 43^2 = (-13)^2 + 43^2 = 13^2 + (-43)^2 = (-13)^2 + (-43)^2 \end{aligned}$$

son todas las representaciones de 2018 como sumas de dos cuadrados, para lo cual es recomendable leer primero este documento en lugar de hacer todos los cálculos «por fuerza bruta». ◂

Antes de continuar, introducimos algunos conceptos y notaciones.

Definición 5. Dados $n, k \in \mathbb{N}$, se denomina **conjunto de representaciones de n como suma de k cuadrados** y se denota $R_k(n)$ al conjunto

$$R_k(n) = \{(x_1, \dots, x_k) \in \mathbb{Z}^k \,:\, x_1^2 + \dots + x_k^2 = n\}.$$

Estamos interesados en calcular su cardinal, para lo cual consideramos la función que sobre $n \in \mathbb{N}$ toma el valor

$$r_k(n) = \#R_k(n),$$

denominada **función de suma de k cuadrados** o **función de representación** (como suma de k cuadrados). Por tanto, $r_k(n)$ es el número de representaciones de n como suma de k cuadrados de enteros, contando como distintos los cambios de signo y las permutaciones[1]. ◂

[1] También tiene sentido considerar el caso $n = 0$, siendo $R_k(0) = \{(0, \overset{k}{\dots}, 0)\}$ y, por tanto, $r_k(0) = 1$ para todo $k \in \mathbb{N}$.

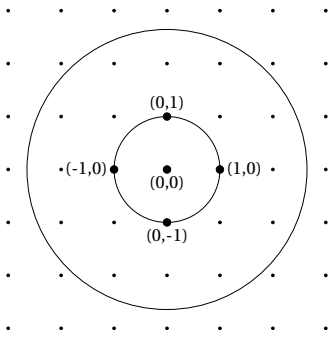

Figura 1: Representación gráfica de $R_2(1) = S^1(0,1) \cap \mathbb{Z}^2$ y $R_2(7) = S^1(0,\sqrt{7}) \cap \mathbb{Z}^2 = \varnothing$.

De la definición surge una duda curiosa: ¿por qué nos interesa contar como distintas las permutaciones y los cambios de signo? Un motivo es que la interpretación geométrica de $R_k(n)$ es que estos son los puntos que pertenecen a la esfera de centro 0 y radio $\sqrt{n}$ en $\mathbb{R}^k$ y tienen coordenadas enteras (véase la figura 1, en la que hemos representado $R_2(1)$ y $R_2(7) = \varnothing$). Por tanto, geométricamente, está claro que las permutaciones y los cambios de signo nos dan puntos distintos y deberían contarse como tales.

Ejemplo 6. Un ejemplo trivial de función de representación es el caso $k = 1$. Para $n \in \mathbb{N}$ se tiene que

$$r_1(n) = \begin{cases} 2 & \text{si } n \text{ es un cuadrado perfecto,} \\ 0 & \text{si } n \text{ no es un cuadrado perfecto,} \end{cases}$$

siendo en el primer caso $\sqrt{n}$ y $-\sqrt{n}$ las dos representaciones distintas. ◂

Sobre las funciones de representación hay muchos resultados conocidos, como por ejemplo el teorema de los cuatro cuadrados de Lagrange, que afirma que $r_4(n) > 0$ para todo $n \in \mathbb{N}$; es decir, que todo entero no negativo es suma de cuatro cuadrados de enteros. De hecho, 4 es el mínimo valor de k con esta propiedad, pues ya hemos visto que 3 no es suma de dos cuadrados y no es complicado comprobar que 7 no es suma de tres cuadrados. También se conoce una fórmula para $r_4(n)$, aunque no abordaremos aquí su demostración. No obstante, no todo está dicho sobre este tipo de problemas. Por ejemplo, el problema de Waring, «para cada $\ell \in \mathbb{N}$, ¿existe un $k \in \mathbb{N}$ tal que cada $n \in \mathbb{N}$ es suma de a lo más k potencias ℓ-ésimas?», no está completamente resuelto, pues se sabe que la respuesta es afirmativa pero se desconocen los valores óptimos de k.

Concluimos esta introducción con una pregunta: ¿qué herramientas son necesarias para resolver los problemas 1 y 3? La base fundamental que vamos a utilizar es el álgebra. Dedicaremos la sección 2 a recordar los conceptos y resultados básicos de la teoría de divisibilidad, que trata de generalizar conceptos como división, máximo común divisor o elemento primo más allá de los números enteros, y se estudia en cursos básicos de cualquier grado en Matemáticas. Aplicaremos esta teoría al estudio de los enteros gaussianos, que son los elementos de la forma $a + bi$ con $a, b \in \mathbb{Z}$ pensados como subanillo de los números complejos. En particular, queremos determinar los primos gaussianos, y relacionarlos con la solución del problema 1. Por último, introduciremos un concepto de teoría de números como son las funciones multiplicativas que nos permitirá, junto con toda la información antes recabada, solucionar el problema 3.

2. Un repaso de teoría de divisibilidad

Para facilitar la comprensión de las demás secciones, hemos decidido añadir este repaso sobre teoría de la divisibilidad. Asumimos que el lector está familiarizado con la teoría de divisibilidad en anillos como se estudia en muchos cursos introductorios de álgebra abstracta; en particular, suponemos que conoce los conceptos de elementos primos e irreducibles, unidades y asociados, y máximo común divisor. Recogeremos los principales resultados que nos harán falta posteriormente, aunque no incluiremos su demostración, al no ser este el objetivo de este documento. Puede consultarse el libro de Jacobson [6, capítulo 2].

En lo sucesivo denotaremos por A a un anillo conmutativo con unidad $1 \neq 0$[2] y sin divisores de cero (esto es, un dominio de integridad[3]) y llamaremos indistintamente **unidad** o elemento **invertible** a cualquier $a \in A$ tal que existe $b \in A$ con $ab = 1$.

Definición 7. Un elemento $d \in A$ se dice que **divide** a otro elemento $a \in A$ y se denota $d \mid a$ si $a = md$ para algún $m \in A$. En ese caso, se dice que d es un **divisor** o un **factor** de a, mientras que la igualdad $a = md$ es una **factorización** de a. En términos de ideales, $d \mid a \iff (a) \subseteq (d)$. ◂

Definición 8. Dos elementos $a, b \in A$ se dice que son **asociados** (o que a es un asociado de b, o que b es un asociado de a) si generan el mismo ideal; es decir, si $(a) = (b)$. En términos de divisibilidad, esto equivale a que $a \mid b$ y $b \mid a$. Escribiremos $a \sim b$ si a y b son asociados. En un dominio de integridad, es fácil comprobar que $a \sim b$ si y solo si $b = \epsilon a$ para alguna unidad ϵ. Claramente, ser asociados es una relación de equivalencia en A. ◂

Definición 9. Sea $c \in A$ no nulo y no invertible. Dada una factorización $c = ab$, diremos que a es un **factor (o divisor) no trivial** si a no es invertible y no es asociado de c. En otro caso, decimos que a es un **factor trivial**. En caso de que tanto a como b sean factores no triviales, decimos que la factorización $c = ab$ es una **factorización no trivial**. ◂

Definición 10. Un elemento no invertible $p \in A$, $p \neq 0$, se dice que es **irreducible** si p solamente admite factorizaciones triviales; esto es, si $p = ab$, entonces o bien a o bien b es una unidad. Por el contrario, si $\alpha \in A$ es no nulo y no invertible y admite alguna factorización no trivial, se dice que α es **compuesto**. ◂

Definición 11. Un elemento no invertible $p \in A$, $p \neq 0$, se dice que es **primo** si para cualesquiera $a, b \in A$ tales que $p \mid ab$, o bien $p \mid a$ o bien $p \mid b$. En términos de ideales, esto significa que (p) es un ideal primo (no nulo). ◂

Lema 12. *En un dominio íntegro, los elementos primos son irreducibles.*

Lema 13. *En un dominio de ideales principales*[4] *(DIP) A, un elemento no nulo y no invertible $a \in A$ es irreducible si y solo si el ideal (a) es maximal. En particular, los elementos irreducibles son primos.*

Proposición 14 (unicidad de la factorización en primos en cualquier dominio). *Sea A un dominio de integridad y sea $a \in A$ un elemento no nulo y no invertible. Si a es producto de primos*[5]*, entonces esta factorización es única salvo reordenaciones y asociados. En otras palabras, si*

$$a = p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m,$$

donde los p_i y los q_j son primos (no necesariamente distintos) y $n, m \in \mathbb{N}$, entonces $n = m$ y, tras una reordenación adecuada, $p_i \sim q_i$ para todo i.

Proposición 15 (existencia de factorización en irreducibles en dominios noetherianos). *En un dominio íntegro noetheriano*[6] *A, todo elemento no nulo y no invertible es producto finito de elementos irreducibles. En particular, esto es válido para un DIP.*

Definición 16. Un **dominio de factorización única** (DFU) es un dominio íntegro en el que todo elemento no nulo y no invertible es producto de irreducibles de manera única salvo reordenaciones y asociados. ◂

Lema 17. *En un dominio de factorización única, todo elemento irreducible es primo.*

Teorema 18. *Un dominio de ideales principales es un dominio de factorización única.*

Definición 19. Un **domino euclídeo** es un dominio íntegro A en el que existe una función $\nu : A \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ tal que, si $a, b \in A$ y $b \neq 0$, existen $q, r \in A$ con $a = qb + r$ y o bien $r = 0$ o bien $\nu(r) < \nu(b)$[7]. ◂

[2] Esta condición sirve para excluir el caso del anillo $A = 0$.

[3] Recordamos que un anillo conmutativo con unidad A se dice que es **íntegro** o que es un **dominio de integridad** si para cualesquiera $a, b \in A$ se verifica la condición $ab = 0 \implies a = 0$ o $b = 0$.

[4] Recordamos que un **dominio de ideales principales** es un dominio de integridad A en el que todo ideal es de la forma (a) para algún $a \in A$.

[5] En general, un elemento de un dominio de integridad puede no tener descomposición como producto de primos. Por ello introducimos a continuación el concepto de dominio de factorización única.

[6] Recordamos que un anillo es **noetheriano** si verifica la condición de cadena ascendente; es decir, si para cada cadena de ideales $I_1 \subseteq \ldots \subseteq I_n \subseteq \ldots$ existe un n tal que $I_n = I_{n+m}$ para todo $m \in \mathbb{N}$.

[7] Sin pérdida de generalidad, podemos asumir que $\nu(a) \leq \nu(ab)$ para $a, b \in A$ con $a, b \neq 0$. A veces se considera esta desigualdad como parte de la definición de dominio euclídeo.

Lema 20. *Un domino euclídeo es un dominio de ideales principales.*

Corolario 21. *Un dominio euclídeo es un dominio de factorización única.*

Definición 22. Un máximo común divisor (mcd) de dos elementos a, b en un dominio de integridad A es un divisor común d de a y b que es maximal en la relación de divisibilidad; esto es, un $d \in A$ tal que $d \mid a$ y $d \mid b$, y si δ es cualquier otro elemento con la misma propiedad, entonces $\delta \mid d$[8]. Decimos que a, b son **primos relativos** o **coprimos** si su mcd existe y es una unidad. El máximo común divisor de dos elementos es único salvo asociados. ◂

Teorema 23 (lema de Bezout). *Si A es un DIP, un máximo común divisor de dos elementos $a, b \in A$ es cualquier elemento d tal que $(a) + (b) = (d)$.*

Concluimos con una lista de sencillas observaciones sobre el máximo común divisor y algunos ejemplos de los conceptos que hemos introducido previamente.

Corolario 24 (propiedades de divisibilidad). *Sea A un DIP y sean $a, b, c, a', b' \in A$. Entonces,*

- *Los elementos $a, b \in A$ son coprimos si y solo si existen $s, t \in A$ tales que $sa + tb = 1$.*
- *Si a, b son coprimos y $a \mid bc$, entonces $a \mid c$.*
- *Si a, b son coprimos y $a \mid c$, $b \mid c$, entonces $ab \mid c$.*
- *Los elementos a, b son ambos coprimos con c si y solo si ab es coprimo con c.*
- *Si $ab = a'b'$ y tanto a, b' como a', b son coprimos, entonces $a \sim a'$ y $b \sim b'$.*

Ejemplo 25. El anillo $\mathbb{Z}$ es un dominio euclídeo tomando como ν la función valor absoluto, por lo que también es un DIP (lema 20) y un DFU (corolario 21).

Si k es un cuerpo, es bien conocido que sus únicos ideales son (0) y $k = (1)$, luego k es un DIP y también un DFU (teorema 18); es más, se comprueba fácilmente que el anillo de polinomios en una variable con coeficientes en k, que denotamos por $k[x]$, es un dominio euclídeo tomando $\nu(p(x))$ igual al grado de $p(x)$.

En general, si un anillo A es un DFU, también $A[x]$ es un DFU; en particular, $\mathbb{Z}[x]$ es un DFU. Si $\mathbb{Z}[x]$ fuera un DIP, entonces todo elemento irreducible generaría un ideal maximal (lema 13); en particular, como x es irreducible, (x) sería maximal y, por tanto, $\mathbb{Z}[x]/(x) \simeq \mathbb{Z}$ sería un cuerpo, lo cual es claramente falso. Es decir, $\mathbb{Z}[x]$ es un DFU que no es un DIP, y, por tanto, tampoco es un dominio euclídeo (lema 20). ◂

3. Enteros gaussianos

Tras haber refrescado conceptos como la divisibilidad, los dominios de ideales principales, los dominios de factorización única, etc., es ahora el momento de ponerlos en práctica. Para ello, vamos a introducir los enteros gaussianos. Comenzaremos definiendo este anillo y exponiendo algunas de sus propiedades básicas, para enfrascarnos enseguida en la demostración de que los enteros gaussianos forman un dominio euclídeo, pudiendo, por tanto, aplicar los conocimientos adquiridos en la sección 2. Concluiremos esta sección mostrando que existe una estrecha relación entre divisibilidad en $\mathbb{Z}$ y divisibilidad en los enteros gaussianos. Todos estos detalles constituyen los primeros pasos hacia nuestro objetivo final: resolver los problemas 1 y 3. Para esta sección y la siguiente tomamos como referencia el artículo divulgativo de Conrad [2].

Definición 26. Se denomina **anillo de enteros gaussianos** al subanillo $\mathbb{Z}[\mathrm{i}]$ de $\mathbb{C}$ definido como

$$\mathbb{Z}[\mathrm{i}] = \{a + b\mathrm{i} \,:\, a, b \in \mathbb{Z}\}.$$ ◂

En los números complejos, dado $z = x + y\mathrm{i}$, se define su conjugado como $\overline{z} = x - y\mathrm{i}$. Esta operación se restringe bien a los enteros gaussianos pues, si $\alpha \in \mathbb{Z}[\mathrm{i}]$, entonces $\overline{\alpha} \in \mathbb{Z}[\mathrm{i}]$. Por otro lado, como $\mathbb{C}$ es un cuerpo, $\mathbb{Z}[\mathrm{i}]$ es un anillo íntegro. Además, podemos pensar que los enteros son enteros gaussianos mediante el morfismo inyectivo de anillos $n \mapsto n + 0\mathrm{i}$ para cada $n \in \mathbb{Z}$.

[8] Como es de esperar, si tomamos $A = \mathbb{Z}$ el concepto de máximo común divisor que acabamos de definir coincide con el habitual, y es único salvo cambio de signo.

Ejemplo 27. Dados $\alpha, \beta \in \mathbb{Z}[\mathrm{i}]$, podemos determinar si $\alpha \mid \beta$ en $\mathbb{Z}[\mathrm{i}]$ dividiendo en $\mathbb{C}$. Esto nos dará un cociente de elementos de $\mathbb{Z}[\mathrm{i}]$ (que pertenece, por tanto, al cuerpo de fracciones, $\mathbb{Q}[\mathrm{i}]$) y se trata de comprobar si este cociente vuelve a ser un elemento de $\mathbb{Z}[\mathrm{i}]$. Por ejemplo,

$$\frac{14+3\mathrm{i}}{4+5\mathrm{i}} = \frac{(14+3\mathrm{i})(4-5\mathrm{i})}{16+25} = \frac{71-58\mathrm{i}}{41} \notin \mathbb{Z}[\mathrm{i}],$$

lo cual demuestra que $(4+5\mathrm{i}) \nmid (14+3\mathrm{i})$ en $\mathbb{Z}[\mathrm{i}]$. ◂

También podemos restringir a los enteros gaussianos la norma que a $z = x + y\mathrm{i} \in \mathbb{C}$ le asigna el valor

$$N(z) = z\overline{z} = (x+y\mathrm{i})\,(x-y\mathrm{i}) = x^2 + y^2.$$

Está claro, además, que si $\alpha \in \mathbb{Z}[\mathrm{i}]$, entonces $N(\alpha) \in \mathbb{Z}_{\geq 0}$. La norma N nos facilita información sobre el anillo de enteros gaussianos. Además, utilizando que la conjugación compleja es multiplicativa (es decir, $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$) se comprueba fácilmente que también es multiplicativa la norma N.

Una propiedad fundamental sobre $\mathbb{Z}[\mathrm{i}]$ es que se trata de un dominio euclídeo y, por tanto, hay factorización única como producto de irreducibles (que son lo mismo que los primos) a partir de los resultados mencionados en la sección previa.

Teorema 28. *$\mathbb{Z}[\mathrm{i}]$ es un dominio euclídeo con respecto a la norma N. En otras palabras, dados $\alpha, \beta \in \mathbb{Z}[\mathrm{i}]$ con $\beta \neq 0$, existen $\gamma, \rho \in \mathbb{Z}[\mathrm{i}]$ tales que $\alpha = \beta\gamma + \rho$ y $N(\rho) < N(\beta)$. De hecho, podemos elegir ρ tal que $N(\rho) \leq \frac{1}{2}N(\beta)$.*

Demostración. Sean $\alpha, \beta \in \mathbb{Z}[\mathrm{i}]$ con $\beta \neq 0$. Dividimos en $\mathbb{C}$ para obtener que

$$\frac{\alpha}{\beta} = \frac{\alpha\overline{\beta}}{\beta\overline{\beta}} = \frac{\alpha\overline{\beta}}{N(\beta)} = \frac{m+n\mathrm{i}}{N(\beta)},$$

donde $m + n\mathrm{i} = \alpha\overline{\beta} \in \mathbb{Z}[\mathrm{i}]$. Usando que $\mathbb{Z}$ es un dominio euclídeo, podemos encontrar $q_1, q_2, r_1, r_2 \in \mathbb{Z}$ tales que

$$m = N(\beta)q_1 + r_1, \qquad n = N(\beta)q_2 + r_2$$

y $0 \leq |r_1|, |r_2| \leq \frac{1}{2}N(\beta)$. Esto es consecuencia de permitir restos negativos; es decir, en lugar de tomar restos $r_1, r_2 \in \{0, 1, \ldots, N(\beta)-1\}$, permitimos elegir los restos de entre los enteros del intervalo $[-N(\beta)/2, \ldots, N(\beta)/2)$. Entonces,

$$\frac{\alpha}{\beta} = \frac{N(\beta)q_1 + r_1 + (N(\beta)q_2 + r_2)\mathrm{i}}{N(\beta)} = q_1 + q_2\mathrm{i} + \frac{r_1 + r_2\mathrm{i}}{N(\beta)}.$$

Tomamos $\gamma = q_1 + q_2\mathrm{i}$, reordenamos y multiplicamos por β, para obtener que

$$\alpha - \beta\gamma = \frac{r_1 + r_2\mathrm{i}}{\overline{\beta}}.$$

Basta con comprobar que $N(\alpha - \beta\gamma) \leq \frac{1}{2}N(\beta)$, y tomar $\rho = \alpha - \beta\gamma$. Ahora bien, tomando normas a ambos lados, como $N(\overline{\beta}) = N(\beta)$, queda que

$$N(\alpha - \beta\gamma) = \frac{r_1^2 + r_2^2}{N(\beta)} \leq \frac{\frac{1}{4}N(\beta)^2 + \frac{1}{4}N(\beta)^2}{N(\beta)} = \frac{1}{2}N(\beta). \qquad \blacksquare$$

A continuación, presentamos una serie de resultados que relacionan la divisibilidad en $\mathbb{Z}$ y en $\mathbb{Z}[\mathrm{i}]$ por medio de la norma N. Algunas de estas propiedades serán claves para resolver el problema de los dos cuadrados.

Lema 29. *Sean $c \in \mathbb{Z}$ y $\alpha = a + b\mathrm{i} \in \mathbb{Z}[\mathrm{i}]$. Entonces, $c \mid \alpha$ en $\mathbb{Z}[\mathrm{i}]$ si y solo si $c \mid a$ y $c \mid b$ en $\mathbb{Z}$. En particular, tomando $b = 0$ se deduce que $c \mid a$ en $\mathbb{Z}[\mathrm{i}]$ si y solo si $c \mid a$ en $\mathbb{Z}$.*

Demostración. Si $\beta = m + n\mathrm{i} \in \mathbb{Z}[\mathrm{i}]$, $c\beta = \alpha \iff cm + cn\mathrm{i} = a + b\mathrm{i} \iff cm = a$ y $cn = b$. $\blacksquare$

Proposición 30 (propiedades de divisibilidad sobre la norma). *Sean $\alpha, \beta \in \mathbb{Z}[\mathrm{i}]$.*

1. *Si $\beta \mid \alpha$, entonces $N(\beta) \mid N(\alpha)$.*
2. *α es una unidad si y solo si $N(\alpha) = 1$.*
3. *Las unidades gaussianas son $\mathbb{Z}[\mathrm{i}]^* = \{\pm 1, \pm \mathrm{i}\}$.*
4. *Si $\alpha \neq 0$ y $\beta \mid \alpha$, entonces $N(\beta) = N(\alpha)$ si y solo si $\beta \sim \alpha$.*
5. *$N(\mathrm{mcd}_{\mathbb{Z}[\mathrm{i}]}(\alpha, \beta)) \mid \mathrm{mcd}_{\mathbb{Z}}(N(\alpha), N(\beta))$.*
6. *Si $\alpha, \beta \neq 0$, entonces cualquier divisor común de α, β de norma máxima es un mcd.*
7. *Si α, β tienen normas coprimas en $\mathbb{Z}$, entonces α, β son coprimos en $\mathbb{Z}[\mathrm{i}]$.*
8. *Si $N(\alpha)$ es primo en $\mathbb{Z}$, entonces α es primo en $\mathbb{Z}[\mathrm{i}]$.*

Demostración.

1. Es consecuencia inmediata de ser N multiplicativa.
2. Si $\alpha\beta = 1$, entonces $N(\alpha)N(\beta) = N(\alpha\beta) = N(1) = 1$ y, como $N(\alpha), N(\beta) \in \mathbb{Z}_{\geq 0}$, se debe dar que $N(\alpha) = N(\beta) = 1$. Recíprocamente, $1 = N(\alpha) = \alpha\overline{\alpha}$ indica que $\overline{\alpha}$ es el inverso de α.
3. Basta resolver $a^2 + b^2 = 1$ para $a, b \in \mathbb{Z}$.
4. Dado $\alpha \neq 0$, si $\beta \mid \alpha$ tenemos que $\alpha = \gamma\beta$. Entonces, $N(\alpha) = N(\gamma)N(\beta)$, luego $N(\alpha) = N(\beta)$ si y solo si $N(\gamma) = 1$, que por la propiedad 2 equivale a que $\gamma \in \mathbb{Z}[\mathrm{i}]^*$ y, por tanto, $\beta \sim \alpha$.
5. Por la propiedad 2, la norma es independiente de la elección de asociados. Si $\delta \mid \alpha, \beta$, por la propiedad 1 tenemos que $N(\delta) \mid N(\alpha), N(\beta)$ y, por tanto, $N(\delta) \mid \mathrm{mcd}_{\mathbb{Z}}(N(\alpha), N(\beta))$. En particular, esto es válido para $\delta = \mathrm{mcd}_{\mathbb{Z}[\mathrm{i}]}(\alpha, \beta)$.
6. Sea δ un mcd de α, β y sea d un divisor común de α, β de norma máxima. Por definición, $d \mid \delta$, luego por la propiedad 1 tenemos que $N(d) \mid N(\delta)$ y, por tanto, $N(d) \leq N(\delta)$, así que por maximalidad de $N(d)$ se tiene que $N(d) = N(\delta)$. De la propiedad 4 deducimos que $d \sim \delta$, luego d también es un mcd de α, β.
7. Por la propiedad 5, si $\mathrm{mcd}_{\mathbb{Z}}(N(\alpha), N(\beta)) = 1$, entonces $N(\mathrm{mcd}_{\mathbb{Z}[\mathrm{i}]}(\alpha, \beta)) = 1$, luego de la propiedad 2 se deduce que $\mathrm{mcd}_{\mathbb{Z}[\mathrm{i}]}(\alpha, \beta)$ es una unidad, es decir, α y β son coprimos en $\mathbb{Z}[\mathrm{i}]$.
8. Si $N(\alpha) = p$ para algún primo p y $\alpha = \beta\gamma$, entonces $p = N(\beta)N(\gamma)$. Utilizando la factorización única en $\mathbb{N}$ deducimos que $N(\beta) = 1$ o $N(\gamma) = 1$; por tanto, o bien β o bien γ es una unidad. ∎

Observación 31. La propiedad 1 no es una equivalencia. En el ejemplo 27 vimos que $(4 + 5\mathrm{i}) \nmid (14 + 3\mathrm{i})$, pero $N(4 + 5\mathrm{i}) = 41 \mid 205 = N(14 + 3\mathrm{i})$.

La propiedad 4 solo es cierta bajo la hipótesis de que $\beta \mid \alpha$. Es fácil encontrar enteros gaussianos que tengan igual norma pero no sean asociados. Por ejemplo, $\alpha = 2 + \mathrm{i}$, $\beta = 2 - \mathrm{i}$ tienen norma 5; $\alpha = 3 + 4\mathrm{i}$, $\beta = 5$ tienen norma 25, y $1 + 8\mathrm{i}$, $4 + 7\mathrm{i}$ tienen norma 65.

Del mismo modo, la propiedad 8 permite comprobar la primalidad en $\mathbb{Z}[\mathrm{i}]$ a través de la norma, pero no es una equivalencia. Por ejemplo, es fácil comprobar que 3 es primo en $\mathbb{Z}[\mathrm{i}]$ pero tiene norma 9. Esto muestra, además, que no es cierto en general que $N(\mathrm{mcd}_{\mathbb{Z}[\mathrm{i}]}(\alpha, \beta)) = \mathrm{mcd}_{\mathbb{Z}}(N(\alpha), N(\beta))$. Por ejemplo, $2 - \mathrm{i}$ y $2 + \mathrm{i}$ son coprimos (pues son primos como consecuencia de la propiedad 8 y, como hemos dicho anteriormente, no son asociados), pero ambos tienen norma igual a 5. También $\alpha = 3 + 4\mathrm{i}$ y $\beta = 3 + \mathrm{i}$ son coprimos (pues $\alpha = (2 + \mathrm{i})^2$ y $\beta = (2 - \mathrm{i})(1 + \mathrm{i})$), mientras que $N(\alpha) = 25$, $N(\beta) = 10$, de modo que $\mathrm{mcd}_{\mathbb{Z}}(N(\alpha), N(\beta)) = 5$. ◂

Corolario 32. *Se verifican las siguientes afirmaciones:*

1. *Un entero gaussiano α es no nulo y no invertible si y solo si $N(\alpha) > 1$.*
2. *Supongamos que $N(\alpha) > 1$. Un factor β de α es trivial si y solo si $N(\beta) = 1$ (unidad) o $N(\beta) = N(\alpha)$ (asociado). Por tanto, existen ocho factores triviales de α, dados por las cuatro unidades $\pm 1, \pm \mathrm{i}$ y los cuatro asociados $\pm\alpha, \pm\mathrm{i}\alpha$. Además, un divisor β de α es no trivial si y solo si $1 < N(\beta) < N(\alpha)$.*

Demostración. Basta aplicar que N valora en $\mathbb{Z}_{\geq 0}$, la equivalencia entre tener norma nula y ser 0, y la propiedad 2 de la proposición 30 para demostrar la primera afirmación, mientras que la segunda afirmación se deduce de las propiedades 1, 2, 3 y 4 de la proposición 30. ∎

4. Primos gaussianos y sumas de dos cuadrados

Nos proponemos ahora resolver el problema 1. Para ello, estudiaremos primero qué primos enteros son representables como sumas de dos cuadrados, lo cual está íntimamente relacionada con la divisibilidad en $\mathbb{Z}[\mathrm{i}]$ y da sentido al estudio que hemos hecho de este anillo. Esta relación motivará que dediquemos parte de nuestro tiempo a explorar los primos enteros que son primos gaussianos, la factorización en $\mathbb{Z}[\mathrm{i}]$ y a determinar todos los primos gaussianos. Concluiremos determinando la factorización de cualquier entero como producto de primos gaussianos, y utilizando este resultado para resolver el problema 1.

Comenzamos con una sencilla observación que nos ayuda a entender por qué los enteros gaussianos son esenciales para nuestro estudio.

Proposición 33. *Un entero n es representable como suma de dos cuadrados si y solo si existe un entero gaussiano α tal que $n = N(\alpha)$.*

Demostración. Esto es inmediato a partir de la expresión $N(a + b\mathrm{i}) = a^2 + b^2$. ∎

La proposición previa nos permite hacer una de las observaciones fundamentales, que era conocida por Fermat y Euler: la propiedad de ser representable como suma de dos cuadrados es una propiedad multiplicativa. Es decir, si $n_1, n_2, \dots, n_s \in \mathbb{N}$ son representables con $n_i = N(\alpha_i)$ para cada i, entonces $n_1 n_2 \cdots n_s = N(\alpha_1 \cdots \alpha_s)$ también es representable[9]. ¡Ojo!, no estamos diciendo que se trate de una equivalencia; es decir, puede suceder que $n_1 n_2 \cdots n_s$ sea representable como suma de dos cuadrados mientras que alguno de los n_i no sea representable. Siguiendo este razonamiento, estamos interesados en estudiar la representabilidad de los divisores y, por tanto, la representabilidad de los primos enteros, y en relacionar esta con la divisibilidad en $\mathbb{Z}[\mathrm{i}]$.

Teorema 34. *Un primo entero p es suma de dos cuadrados si y solo si p es compuesto en $\mathbb{Z}[\mathrm{i}]$. Es decir, p permanece primo en $\mathbb{Z}[\mathrm{i}]$ si y solo si p no es suma de dos cuadrados.*

Demostración. Si un primo entero p es de la forma $p = a^2 + b^2$, entonces $p = (a + b\mathrm{i})(a - b\mathrm{i})$ es una descomposición no trivial en $\mathbb{Z}[\mathrm{i}]$ por el corolario 32, pues $1 < N(a \pm b\mathrm{i}) = p < p^2 = N(p)$.

Recíprocamente, sea p un primo en $\mathbb{Z}$ que es compuesto en $\mathbb{Z}[\mathrm{i}]$, y sea $p = \alpha\beta$ una descomposición no trivial. Tomando normas, $p^2 = N(\alpha)N(\beta)$. Por el corolario 32, necesariamente se debe cumplir que $N(\alpha) = N(\beta) = p$. Por tanto, si $\alpha = a + b\mathrm{i}$, entonces $p = a^2 + b^2$. ∎

Corolario 35. *Sea $p \in \mathbb{N}$ un primo impar que es compuesto en $\mathbb{Z}[\mathrm{i}]$. Salvo asociados, p tiene dos factores primos distintos en $\mathbb{Z}[\mathrm{i}]$. Además, dichos factores son conjugados y tienen norma p.*

Demostración. De la demostración del teorema 34 sabemos que $N(a \pm b\mathrm{i}) = p$, luego por la propiedad 8 de la proposición 30, $a \pm b\mathrm{i}$ son primos gaussianos. Si $a + b\mathrm{i} = \epsilon(a - b\mathrm{i})$ para alguna unidad $\epsilon \in \mathbb{Z}[\mathrm{i}]$, entonces

$$\begin{cases} \epsilon = 1 & \implies a + b\mathrm{i} = a - b\mathrm{i} \implies b = 0, \quad p = a^2, \\ \epsilon = -1 & \implies a + b\mathrm{i} = -a + b\mathrm{i} \implies a = 0, \quad p = b^2, \\ \epsilon = \mathrm{i} & \implies a + b\mathrm{i} = b + a\mathrm{i} \implies b = a, \quad p = 2a^2, \\ \epsilon = -\mathrm{i} & \implies a + b\mathrm{i} = -b - a\mathrm{i} \implies b = -a, \quad p = 2a^2. \end{cases}$$

En cualquier caso, esto es imposible para un primo impar p, por lo que $a \pm b\mathrm{i}$ no son asociados. ∎

Comenzaremos estudiando los primos enteros que son sumas de dos cuadrados. Para ello, el teorema 34 nos indica que es conveniente estudiar la factorización en $\mathbb{Z}[\mathrm{i}]$ de los primos enteros. El siguiente lema nos indica que de este modo se recuperan todos los primos gaussianos.

Lema 36. *Todo primo gaussiano π divide a algún primo entero p en $\mathbb{Z}[\mathrm{i}]$.*

Demostración. Basta observar que $N(\pi) = \pi\overline{\pi}$ en $\mathbb{Z}[\mathrm{i}]$. Como $N(\pi) > 1$ y $N(\pi) \in \mathbb{N}$, existe la descomposición en primos enteros $\pi\overline{\pi} = N(\pi) = p_1 \cdots p_r$, y como π es primo, debe dividir a alguno de los p_i. ∎

[9] Esta sencilla idea, basada en la multiplicatividad de N, resulta clave para demostrar el teorema de los cuatro cuadrados de Lagrange que hemos enunciado más adelante como el teorema 58, pues en la práctica se demuestra que todo primo entero es representable.

Observación 37. El primo $p = 2$ factoriza como $2 = (1+\mathrm{i})(1-\mathrm{i})$, donde $1 \pm \mathrm{i}$ son primos gaussianos de norma 2, pero son asociados pues $1 - \mathrm{i} = -\mathrm{i}(1+\mathrm{i})$, luego $2 = -\mathrm{i}(1+\mathrm{i})^2$, apareciendo dos veces el primo gaussiano $1+\mathrm{i}$. El hecho de que en una factorización aparezca un primo con multiplicidad mayor que 1 se conoce como **ramificación**. En general, se utiliza la siguiente terminología para clasificar la factorización de un primo entero en $\mathbb{Z}[\mathrm{i}]$. ◂

Definición 38. Sea $p \in \mathbb{N}$ un primo entero.

- Si p permanece primo en $\mathbb{Z}[\mathrm{i}]$, se dice que p es un primo **inerte**.
- Si p es compuesto en $\mathbb{Z}[\mathrm{i}]$, decimos que p es un primo que **descompone**.
- Si $p = 2$, se dice que 2 es el primo **ramificado**. ◂

Utilizando esta terminología, hemos visto hasta el momento que p es inerte si y solo si p no es suma de dos cuadrados; que los compuestos impares p descomponen con dos factores primos no asociados, y que $p = 2$ es ramificado y es asociado del cuadrado del primo $1+\mathrm{i}$. Por otro lado, el teorema 34 ha resultado ser una herramienta útil en nuestro estudio, pero no resuelve el problema por sí solo. Necesitamos la caracterización de primos que son sumas de dos cuadrados, debida a Fermat.

Teorema 39 (clasificación de los primos que son sumas de dos cuadrados). *Sea p un primo entero. Las siguientes afirmaciones son equivalentes:*

1. $p = 2$ *o* $p \equiv 1 \bmod 4$.
2. *La ecuación* $x^2 \equiv -1 \bmod p$ *tiene solución; es decir,* -1 *es un cuadrado módulo* p.
3. $p = a^2 + b^2$ *para ciertos* $a, b \in \mathbb{Z}$.

En particular, los primos enteros p congruentes con 3 mód 4 *son precisamente los primos enteros que no son suma de dos cuadrados.*

Demostración. **3** $\implies$ **1**: Basta utilizar que los cuadrados módulo 4 son 0 y 1, por lo que cualquier entero $n \equiv 3 \bmod 4$ no puede ser suma de dos cuadrados en $\mathbb{Z}/(4)$, y menos aún en $\mathbb{Z}$.

1 $\implies$ **2**: Para $p = 2$, $x^2 \equiv -1 \bmod p$ tiene solución $x = 1$. Para un primo impar p, podemos considerar la siguiente factorización en $\mathbb{F}_p[x]$:

$$x^{p-1} - 1 = (x^{\frac{(p-1)}{2}} - 1)(x^{\frac{(p-1)}{2}} + 1). \tag{1}$$

Aplicando el teorema de Fermat[10], todo elemento a no nulo de $\mathbb{F}_p$ verifica que $a^{p-1} - 1 \equiv 0 \bmod p$ o, lo que es lo mismo, existen al menos $p-1$ raíces distintas en $\mathbb{F}_p$ del polinomio $x^{p-1} - 1$. En el lado derecho de la ecuación (1), al ser $\mathbb{F}_p[x]$ un DFU (ejemplo 25), el primer polinomio del lado derecho no puede tener más raíces que su grado, que es $(p-1)/2$, luego el segundo polinomio del lado derecho debe tener alguna raíz sobre $\mathbb{F}_p$; en otras palabras, existe $a \in \mathbb{Z}$ tal que $a^{(p-1)/2} \equiv -1 \bmod p$. Si, además, $p \equiv 1 \bmod 4$, entonces $(p-1)/4 \in \mathbb{N}$ y $x = a^{(p-1)/4}$ satisface que $x^2 \equiv a^{(p-1)/2} \equiv -1 \bmod p$.

2 $\implies$ **3**: Por el teorema 34, basta demostrar que p es compuesto en $\mathbb{Z}[\mathrm{i}]$. Si $x \in \mathbb{Z}$ es una solución de $x^2 \equiv -1 \bmod p$, entonces $p \mid (x^2+1)$ en $\mathbb{Z}$, luego $p \mid (x^2+1) = (x+\mathrm{i})(x-\mathrm{i})$ en $\mathbb{Z}[\mathrm{i}]$. Si p fuera primo en $\mathbb{Z}[\mathrm{i}]$, tendríamos que $p \mid (x+\mathrm{i})$ o $p \mid (x-\mathrm{i})$, y del lema 29 deduciríamos que $p \mid 1$, lo cual es imposible. Por tanto, p es compuesto en $\mathbb{Z}[\mathrm{i}]$. ∎

Resumimos los resultados previos en los siguientes teoremas.

Teorema 40 (factorización de primos enteros como producto de primos gaussianos). *Sea $p \in \mathbb{N}$ un primo. La descomposición de p en $\mathbb{Z}[\mathrm{i}]$ queda determinada por la congruencia de p módulo* 4 *como sigue:*

1. *Si* $p \equiv 3 \bmod 4$, *entonces p permanece primo en* $\mathbb{Z}[\mathrm{i}]$ *(inerte).*
2. *Si* $p \equiv 1 \bmod 4$, *entonces* $p = \pi\overline{\pi}$, *donde* $\pi, \overline{\pi}$ *son primos conjugados y no asociados (descompone).*
3. $2 = (1+\mathrm{i})(1-\mathrm{i}) = -\mathrm{i}\,(1+\mathrm{i})^2$ *(ramificado).*

Demostración. Basta combinar la observación 37, los teoremas 39 y 34 y el corolario 35. ∎

[10] Si $a \not\equiv 0 \bmod p$, entonces $a^{p-1} \equiv 1 \bmod p$. Aunque se pueden dar demostraciones sencillas y directas, también es un corolario inmediato del teorema de Lagrange en teoría de grupos, pues el grupo de unidades $\mathbb{F}_p^*$ tiene orden $p-1$.

Teorema 41. *Sea $\alpha \in \mathbb{Z}[\mathrm{i}]$ un primo gaussiano. Entonces, salvo asociados, α debe ser de uno de los siguientes tipos:*

1. *$p \in \mathbb{N}$ es un primo entero con $p \equiv 3 \bmod 4$,*
2. *π o $\overline{\pi}$ con $N(\pi) = p \in \mathbb{N}$ primo y $p \equiv 1 \bmod 4$, o*
3. *$1 + \mathrm{i}$.*

Además, todos los tipos de enteros gaussianos arriba descritos, al igual que sus asociados, son primos gaussianos, por lo que hemos descrito todos los primos gaussianos.

Demostración. Por el lema 36, cada primo gaussiano divide a un primo entero. Como $\mathbb{Z}[\mathrm{i}]$ tiene factorización única, los primos de $\mathbb{Z}[\mathrm{i}]$ deben ser los primos gaussianos que aparecen en el teorema 40. Por último, está claro que los enteros gaussianos que aparecen en el enunciado son primos, pues en los dos últimos tipos $N(\alpha)$ es primo y la propiedad 8 de la proposición 30 lo concluye, y en el primer tipo es consecuencia del teorema 40. ∎

En particular, del teorema 41 se deduce que la norma de cualquier primo gaussiano es p o p^2, siendo p un primo entero. Por último, del teorema 40 se deduce la factorización de cualquier entero dentro de $\mathbb{Z}[\mathrm{i}]$.

Corolario 42. *Sea $n \geq 2$ un entero, y sean $p_1, \dots, p_r$ los factores primos de n congruentes con $1 \bmod 4$ y $q_1, \dots, q_s$ los factores primos de n congruentes con $3 \bmod 4$, de modo que la descomposición de n como producto de primos en $\mathbb{Z}$ es la siguiente:*

$$n = 2^c p_1^{n_1} \cdots p_r^{n_r} q_1^{m_1} \cdots q_s^{m_s}$$

para ciertos $n_\ell, m_j \geq 1$ y $c \in \mathbb{Z}_{\geq 0}$. Para cada $1 \leq \ell \leq r$, sea $p_\ell = \pi_\ell \overline{\pi}_\ell$ la descomposición de p_ℓ como producto de dos primos conjugados no asociados en $\mathbb{Z}[\mathrm{i}]$. Entonces, n factoriza en $\mathbb{Z}[\mathrm{i}]$ del siguiente modo:

$$n = (1+\mathrm{i})^{2c} \pi_1^{n_1} \overline{\pi}_1^{n_1} \cdots \pi_r^{n_r} \overline{\pi}_r^{n_r} q_1^{m_1} \cdots q_s^{m_s}.$$

Además, dicha factorización es única salvo el orden de los factores y asociados.

Demostración. Es consecuencia de ser $\mathbb{Z}[\mathrm{i}]$ un DFU (por el teorema 28 y el corolario 21) y del teorema 41. ∎

Terminamos esta sección con el resultado principal, la solución completa al problema 1, que viene dada por medio de los factores primos del entero en cuestión. En el teorema está excluido el caso $n = 1$, pero está claro que $1 = 1^2 + 0^2 = (-1)^2 + 0^2 = 0^2 + 1^2 = 0^2 + (-1)^2$ son todas sus representaciones, y, en particular, es representable.

Teorema 43. *Un entero $n > 1$ es suma de dos cuadrados si y solo si cada factor primo p de n verificando $p \equiv 3 \bmod 4$ tiene multiplicidad par.*

Demostración. Ya hemos comentado que ser suma de dos cuadrados es una propiedad multiplicativa tras la proposición 33. El primo $p = 2$ y cualquier primo $p \equiv 1 \bmod 4$ son sumas de dos cuadrados por el teorema 39; por tanto, también lo son sus potencias. Por otro lado, un primo $p \equiv 3 \bmod 4$ no es suma de dos cuadrados, pero p^2 sí lo es y, por tanto, cualquier potencia par de p es también suma de dos cuadrados. En definitiva, cualquier n en el que los primos $p \equiv 3 \bmod 4$ aparecen con multiplicidad par es una suma de dos cuadrados.

Sea ahora $n > 1$ suma de dos cuadrados que tiene un factor primo $p \equiv 3 \bmod 4$ (pues en otro caso no hay nada que demostrar). Si $n = a^2 + b^2$, entonces $p \mid n = (a + b\mathrm{i})(a - b\mathrm{i})$ en $\mathbb{Z}[\mathrm{i}]$, y como p es inerte, o bien $p \mid (a + b\mathrm{i})$ o bien $p \mid (a - b\mathrm{i})$. En cualquier caso, la conclusión es que $p \mid a$ y $p \mid b$ en $\mathbb{Z}$ por el lema 29. De $a = pA$ y $b = pB$ se deduce que $n = a^2 + b^2 = p^2(A^2 + B^2)$, luego $p^2 \mid n$ y $n/p^2 = A^2 + B^2$ es suma de dos cuadrados. Si m es la multiplicidad de p como divisor de n, escribiendo $m = 2k + r$ para $r = 0$ o 1 y aplicando este resultado recursivamente k veces, concluimos que $n' = n/p^{2k}$ es suma de dos cuadrados. El caso $r = 1$ implicaría que $p \mid n'$ pero $p^2 \nmid n'$, pero esto no es posible por el razonamiento previo (sustituyendo n por n'). Por tanto, se tiene que $r = 0$ o, lo que es lo mismo, la multiplicidad de p es par. ∎

5. Número de representaciones como suma de dos cuadrados

Para concluir, queremos utilizar los resultados anteriores para resolver el problema 3. La fórmula buscada dependerá, como uno podría imaginar, de los divisores de n que son congruentes con 1 y con 3 mód 4; más concretamente, de su cardinal. La fórmula que queremos demostrar es la siguiente:

$$r_2(n) = 4(d_1(n) - d_3(n)), \tag{2}$$

donde $d_j(n)$ denota el número de divisores de n congruentes con j mód 4 para $j \in \{1,3\}$. Para la demostración necesitamos realizar varias comprobaciones, por lo que hemos decidido dividir la prueba en una serie de pasos.

- Partiendo de la factorización de n, observar que podemos eliminar las potencias pares de primos $p \equiv 3$ mód 4, al igual que el factor 2 (con su potencia).
- Interpretar el factor 4 de la ecuación (2), así como las funciones $r_2(n)$ y $r_2(n)/4$, en términos de ideales del anillo $\mathbb{Z}[i]$.
- Demostrar que, tras dividir la ecuación (2) por 4, ambos términos son multiplicativos en n; es decir, que si demostramos la fórmula para n, m coprimos, entonces es válida para nm.
- Demostrar que la fórmula es cierta para las potencias de primos enteros $n = p^e$.

La mayoría de estas ideas se encuentran en el libro de Hardy y Wright [5, sección 20] con un enfoque ligeramente distinto.

Proposición 44. *Dado $n \in \mathbb{N}$, $r_2(n) = r_2(2n)$.*

Demostración. Sea $n = a^2+b^2$ una representación de n. Entonces, se puede comprobar que $(a-b)^2+(a+b)^2$ es una representación de $2n$. Recíprocamente, sea $2n = c^2+d^2$. Entonces, $c \equiv d$ mód 2 y $n = \left(\frac{c+d}{2}\right)^2+\left(\frac{c-d}{2}\right)^2$. Por tanto, $(a,b) \mapsto (a-b, a+b)$ es una biyección entre $R_2(n)$ y $R_2(2n)$. ∎

Observación 45. El hecho de que si $n = a^2+b^2$, entonces $2n = (a-b)^2+(a+b)^2$, es también consecuencia de la igualdad $(a^2+b^2)(c^2+d^2) = (ac-bd)^2+(ad+bc)^2$ para cualesquiera $a, b, c, d \in \mathbb{R}$, lo cual se deduce de que N sea multiplicativa y, por tanto, $N(a+bi)N(c+di) = N((a+bi)(c+di))$. Dicha fórmula nos indica que el producto de enteros representables es representable, y nos da un modo de obtener una representación de nm conocidas representaciones de n y de m, que es lo que aplicamos para $m = 2 = 1^2+1^2$. Solamente existen fórmulas similares que expresan el producto de dos sumas de k cuadrados como una suma de k cuadrados para $k \in \{1,2,4,8\}$, como afirma el problema de Hurwitz; para más información, véase el libro de Jacobson [6, sección 7.6]. ◂

La proposición 44 nos indica que, a la hora de calcular $r_2(n)$, podemos primero simplificar n dividiendo por 2 tantas veces como sea posible. El análogo para los primos $p \equiv 3$ mód 4 que presentamos a continuación nos permite simplificar las potencias de p por pares, lo cual tiene sentido porque recordamos que, si p es divisor de n con multiplicidad impar, entonces n no es representable.

Proposición 46. *Sean $n, p \in \mathbb{N}$ con p primo y $p \equiv 3$ mód 4. Entonces, $r_2(n) = r_2(p^2n)$.*

Demostración. Si $n = a^2+b^2$, entonces $p^2n = (pa)^2+(pb)^2$. Recíprocamente, si $p^2n = c^2+d^2$, entonces, repitiendo el argumento del teorema 43 (es decir, que como p es inerte, $p \mid c^2+d^2 = (c+di)(c-di)$ en $\mathbb{Z}[i]$ y, por tanto, $p \mid c, d$ en $\mathbb{Z}$), vemos que $n = (c/p)^2+(d/p)^2$ es una representación de n. Por tanto, $(a,b) \mapsto (pa, pb)$ es una biyección entre $R_2(n)$ y $R_2(p^2n)$. ∎

Hasta el momento no hemos tenido que lidiar con el tema de las permutaciones y cambios de signo entre representantes. Con el fin de evitar este detalle, consideramos una variación sobre el conjunto de representaciones.

Definición 47. Una representación $n = a^2+b^2$ se dice **positiva** si $a > 0$ y $b \geq 0$. Se define el **conjunto de representaciones positivas de** n como sigue:

$$R_2^+(n) = \{(x,y) \in \mathbb{Z}^2 : x^2+y^2 = n, x > 0, y \geq 0\}.$$

Definimos también la función $r_2^+(n) = \#R_2^+(n)$ que cuenta las representaciones positivas de n. ◂

Vía la biyección $\mathbb{Z}^2 \to \mathbb{Z}[\mathrm{i}] : (a,b) \mapsto \gamma = a + b\mathrm{i}$, podemos pensar que $R_2(n)$ y $R_2^+(n)$ son subconjuntos de $\mathbb{Z}[\mathrm{i}]$, y diremos que $a + b\mathrm{i}$ es un entero gaussiano **positivo** si $(a,b) \in R_2^+(n)$. Bajo esta correspondencia, $r_2(n)$ es claramente el número de enteros gaussianos γ tales que $N(\gamma) = n$. El siguiente resultado nos dice que al considerar solamente generadores positivos estamos eliminando la multiplicación por 4 que aparece al considerar asociados (pues hay precisamente cuatro unidades), de modo que $r_2(n)$ cuenta los cuatro asociados de un $\alpha \neq 0$ como representaciones distintas, mientras que $r_2^+(n)$ cuenta solamente una de ellas.

Lema 48. *Sea $n \in \mathbb{N}$. Si pensamos $(x,y) \in R_2(n)$ como el entero gaussiano $\alpha = x + y\mathrm{i} \neq 0$, entonces α tiene un único asociado β que pertenece a $R_2^+(n)$. Por tanto, $r_2(n) = 4r_2^+(n)$.*

Demostración. Que los cuatro asociados distintos de α,

$$\{\alpha = x + y\mathrm{i}, \quad \mathrm{i}\alpha = -y + x\mathrm{i}, \quad -\alpha = -x - y\mathrm{i}, \quad -\mathrm{i}\alpha = y - x\mathrm{i}\},$$

pertenecen a $R_2(n)$, pero solo uno de ellos pertenece a $R_2^+(n)$, es una mera comprobación. La fórmula es una consecuencia inmediata. ∎

Observación 49. Recordemos que $\alpha \sim \beta \iff (\alpha) = (\beta)$. Por tanto, estamos seleccionando el único generador positivo de cada ideal (no nulo). Es más, podemos definir una norma N' sobre el conjunto de ideales del siguiente modo: para $I = (\alpha)$, se define $N'(I) = N(\alpha)$. N' está bien definida porque $(\alpha) = (\beta) \iff \alpha = \epsilon\beta$ para alguna unidad ϵ, y $N(\alpha) = N(\epsilon)N(\beta) = N(\beta)$ por ser N multiplicativa y por la propiedad 2 de la proposición 30. Además, N' es multiplicativa por serlo N. Por abuso de notación, denotaremos a ambas normas como N. Teniendo esto en cuenta, $r_2^+(n)$ contabiliza el número de ideales de $\mathbb{Z}[\mathrm{i}]$ de norma n (vía la correspondencia que asocia a cada $(a,b) \in R_2^+(n)$ el ideal $(a + b\mathrm{i})$), a diferencia de $r_2(n)$, que contabiliza el número de elementos de norma n. ◂

Intentemos ahora entender el factor 4 de la ecuación (2), así como el hecho de que en el ejemplo 4 el número 4 tuviera cuatro representaciones mientras que el número 10 tenía ocho representaciones, cuando en ambos casos hay una sola representación salvo permutaciones y asociados. Para ello consideramos el conjunto de elementos no nulos de $\mathbb{Z}[\mathrm{i}]$, y la acción sobre él (por multiplicación) del grupo de unidades $\mathbb{Z}[\mathrm{i}]^*$, que identificamos con el grupo cíclico C_4 generado por la multiplicación por i. La acción es libre[11] y, por tanto, cada órbita[12] tiene longitud 4; de hecho, la órbita de $\alpha \neq 0$ es $\{\mathrm{i}\alpha, -\alpha, -\mathrm{i}\alpha, \alpha\}$. Podemos considerar también la acción (sobre el conjunto de elementos no nulos de $\mathbb{Z}[\mathrm{i}]$) del grupo diédrico D_4, tomando como generadores la multiplicación por i y la conjugación. En ese caso, la órbita de $\alpha = x + y\mathrm{i} \neq 0$ es

$$\begin{aligned} \alpha &= x + y\mathrm{i} \leftrightarrow (x,y), & \overline{\alpha} &= x - y\mathrm{i} \leftrightarrow (x,-y), \\ \mathrm{i}\alpha &= -y + x\mathrm{i} \leftrightarrow (-y,x), & \mathrm{i}\overline{\alpha} &= y + x\mathrm{i} \leftrightarrow (y,x), \\ -\alpha &= -x - y\mathrm{i} \leftrightarrow (-x,-y), & -\overline{\alpha} &= -x + y\mathrm{i} \leftrightarrow (-x,y), \\ -\mathrm{i}\alpha &= y - x\mathrm{i} \leftrightarrow (y,-x), & -\mathrm{i}\overline{\alpha} &= -y - x\mathrm{i} \leftrightarrow (-y,-x). \end{aligned}$$

En términos de ideales, C_4 permuta los generadores de (α), mientras que D_4 permuta los generadores de (α) y los de su ideal conjugado $(\overline{\alpha})$. Como siempre hay cuatro asociados distintos, la órbita puede tener longitud 4 u 8, y tiene longitud 4 si y solo si $(\alpha) = (\overline{\alpha})$. Esto último ocurre si y solo si

- o bien $x = 0$ o $y = 0$, caso en que (α) está generado por un número natural,
- o bien $x = y$, caso en que (α) está generado por $m(1 + \mathrm{i})$ para algún $m \in \mathbb{N}$.

Por último, si incluimos al elemento $0 = 0 + 0\mathrm{i} \leftrightarrow (0,0)$, se trataría del único punto fijo de las acciones antes mencionadas.

Continuamos simplificando el problema 3: veremos que $r_2^+(n)$ es una función multiplicativa en n; es decir, si $n, m \in \mathbb{N}$ y $\mathrm{mcd}(m,n) = 1$, entonces $r_2^+(mn) = r_2^+(m)r_2^+(n)$. La multiplicatividad es una propiedad importante de funciones sobre $\mathbb{N}$ que reduce considerablemente la dificultad de determinar su valor, y aparece frecuentemente al tomar enfoques aritméticos, como es nuestro caso. Por este motivo es conveniente que tratemos algunos resultados básicos sobre estas funciones.

[11] Recordamos que una acción de un grupo G en un conjunto X se dice **libre** si de la igualdad $g \cdot x = h \cdot x$, siendo $g, h \in G, x \in X$, se deduce que $g = h$. En nuestro caso esto está claro pues $\epsilon\alpha = \epsilon'\alpha \implies (\epsilon - \epsilon')\alpha = 0$ y, por integridad, se tiene que $\epsilon = \epsilon'$.

[12] La **órbita** de un elemento x (siendo G una acción en X) es el conjunto $\{g \cdot x\}_{g \in G}$ de los trasladados de x por G.

5.1. Apunte sobre funciones multiplicativas

Definición 50. Sea M un monoide[13] con elemento unidad 1. Una **función aritmética M-valorada** (o simplemente una **función aritmética** cuando no haya ambigüedad sobre M) es una función $f : \mathbb{N} \to M$. ◂

Definición 51. Una función aritmética f se dice que es **multiplicativa** si $f(1) = 1$ y si para cualesquiera $n, m \in \mathbb{N}$ tales que $\mathrm{mcd}(n, m) = 1$ se verifica que $f(nm) = f(n)f(m)$. Si, además, f verifica que $f(nm) = f(n)(m)$ para cualesquiera $n, m \in \mathbb{N}$ (independientemente de su mcd), entonces se dice que f es **completamente multiplicativa**. ◂

Un ejemplo trivial de función completamente multiplicativa es la función constante 1. Para $M = \mathbb{C}$ también lo es la exponenciación compleja $f(n) = n^s$. Un ejemplo menos obvio y que está relacionado con las representaciones por dos cuadrados es el siguiente.

Ejemplo 52. Es fácil comprobar que la función[14]

$$\chi(n) = \begin{cases} (-1)^{\frac{n-1}{2}} & \text{si } n \equiv 1 \bmod 2, \\ 0 & \text{si } n \equiv 0 \bmod 2, \end{cases} = \begin{cases} 1 & \text{si } n \equiv 1 \bmod 4, \\ -1 & \text{si } n \equiv 3 \bmod 4, \\ 0 & \text{si } n \equiv 0 \bmod 2 \end{cases} \tag{3}$$

es una función completamente multiplicativa con valores en el submonoide $\{-1, 0, 1\}$ de $(\mathbb{N}, \cdot, 1)$. ◂

También es fácil comprobar la siguiente proposición, por lo que dejamos la demostración para el lector.

Proposición 53.

- *Una función $f : \mathbb{N} \to M$ es multiplicativa si y solo si $f(1) = 1$ y para cualesquiera $p, n \in \mathbb{N}$ y $e \in \mathbb{N}$, siendo p primo y $\mathrm{mcd}(p, n) = 1$, se verifica que $f(p^e n) = f(p^e)f(n)$.*
- *Dos funciones multiplicativas $f, g : \mathbb{N} \to M$ son iguales si y solo si $f(p^e) = g(p^e)$ para cada primo $p \in \mathbb{N}$ y cada $e \in \mathbb{N}$.*

Pasamos ahora a introducir la convolución de Dirichlet, que es una operación de funciones aritméticas y que preserva la multiplicatividad. Este hecho nos será muy útil para determinar la expresión de $r_2^+(n)$. De ahora en adelante consideramos solamente funciones aritméticas $\mathbb{C}$-valoradas.

Definición 54. Sean f, g funciones aritméticas. Definimos la **convolución de Dirichlet** de f y g como

$$(f * g)(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right) = \sum_{ab=n} f(a)g(b),$$

donde la suma se toma sobre todos los divisores positivos de n. ◂

Proposición 55. *Si f, g son funciones multiplicativas, entonces $f * g$ también es multiplicativa.*

Demostración. Está claro que $(f * g)(1) = 1$. Basta demostrar que $(f * g)(p^e n) = (f * g)(p^e)(f * g)(n)$ si p es primo, $e \in \mathbb{N}$ y $\mathrm{mcd}(p, n) = 1$. Observamos que todo divisor d de $p^e n$ es de la forma $d = p^c d'$ para $0 \le c \le e$ y $d' \mid n$ únicos. Por tanto, se tiene que

$$\begin{aligned}(f * g)(p^e n) &= \sum_{d|p^e n} f(d)g\left(\frac{p^e n}{d}\right) = \sum_{\substack{0\le c\le e \\ d'|n}} f(p^c d')g\left(\frac{p^e n}{p^c d'}\right) = \sum_{\substack{0\le c\le e \\ d'|n}} f(p^c)f(d')g(p^{e-c})g\left(\frac{n}{d'}\right) \\ &= \left(\sum_{0\le c\le e} f(p^c)g\left(\frac{p^e}{p^c}\right)\right)\left(\sum_{d'|n} f(d')g\left(\frac{n}{d'}\right)\right) = (f * g)(p^e)\,(f * g)(n).\end{aligned}$$

■

[13] Un **monoide** es un conjunto M con una operación interna asociativa y con elemento neutro. El lector puede pensar que se trata de un grupo pues tomaremos $M = \mathbb{C}$, pero no es necesaria la existencia de elemento opuesto para la definición.

[14] En teoría de números, χ se conoce como **símbolo de reciprocidad cuadrática de** -1 o como el **carácter de Dirichlet no trivial módulo** 4.

5.2. Expresión de la función de representación

Veamos ahora la principal razón por la que hemos introducido las funciones multiplicativas.

Teorema 56. *La función $r_2^+(n)\colon \mathbb{N} \to \mathbb{Z}_{\geq 0}$ es multiplicativa.*

Demostración. Está claro que $r_2^+(1) = 1$. Queremos demostrar que $r_2^+(p^e n) = r_2^+(p^e) r_2^+(n)$ para cualquier primo p y cualesquiera $n, e \in \mathbb{N}$ con $\mathrm{mcd}(p, n) = 1$. Pensando los elementos de $R_2^+(m)$ como ideales de norma m, la multiplicatividad de la norma demuestra que la aplicación de multiplicación de ideales $R_2^+(p^e) \times R_2^+(n) \to R_2^+(p^e n)\colon ((\beta), (\gamma)) \mapsto (\beta\gamma)$ está bien definida. Basta demostrar que se trata de una biyección.

La inyectividad se sigue de que β, γ deben ser coprimos en $\mathbb{Z}[\mathrm{i}]$ porque sus normas son coprimas en $\mathbb{Z}$ (propiedad 7 de la proposición 30) y del corolario 24. En efecto, si $\beta\gamma = \beta'\gamma'$ con $N(\beta) = N(\beta') = p^e$ y $N(\gamma) = N(\gamma') = n$, entonces $\mathrm{mcd}(N(\beta), N(\gamma')) = 1$ implica que $\mathrm{mcd}(\beta, \gamma') = 1$, luego $\beta \mid \beta'$. Del mismo modo se deduce que $\beta' \mid \beta$, luego β, β' son asociados, y lo análogo es cierto para γ, γ'. En términos de ideales, esto significa que $(\beta) = (\beta')$ y $(\gamma) = (\gamma')$.

La epiyectividad equivale a demostrar que si $\alpha \in \mathbb{Z}[\mathrm{i}]$ tiene norma $p^e n$ entonces se puede factorizar $\alpha = \beta\gamma$ con $N(\beta) = p^e$ y $N(\gamma) = n$. Consideramos la factorización de α en $\mathbb{Z}[\mathrm{i}]$, $\alpha = \pi_1 \cdots \pi_r$, siendo π_i primos gaussianos contados con multiplicidad y únicos salvo asociados. Entonces, tenemos que $p^e n = N(\alpha) = N(\pi_1) \cdots N(\pi_r)$. Algunas de las normas $N(\pi_j)$ deben ser divisibles por p. Reordenando, podemos suponer que son $\pi_1, \ldots, \pi_s$, con $1 \leq s \leq r$, y entonces $p \nmid N(\pi_{s+1}), \ldots, N(\pi_r)$. Sabemos que $N(\pi_j)$ debe ser de la forma q o q^2 para algún primo entero q, luego $N(\pi_j) = p$ o p^2 para $1 \leq j \leq s$, mientras que $\mathrm{mcd}(p, N(\pi_{s+1} \cdots \pi_r)) = 1$. Necesariamente $N(\pi_1 \cdots \pi_s) = p^e$ y $N(\pi_{s+1} \cdots \pi_r) = n$. Por tanto, debe darse que $\beta = \pi_1 \cdots \pi_s$ y $\gamma = \pi_{s+1} \cdots \pi_r$ salvo asociados. ∎

Finalmente, podemos resolver el problema 3 utilizando los resultados previos.

Teorema 57. *Para cada $n \in \mathbb{N}$, $r_2^+(n) = d_1(n) - d_3(n)$ y, por tanto, $r_2(n) = 4(d_1(n) - d_3(n))$.*

Demostración. Basta probar la fórmula para $r_2^+(n)$ y aplicar el lema 48. Para ello, consideramos de nuevo la función χ del ejemplo 52. Por definición se tiene que

$$(\chi * 1)(n) = \sum_{d \mid n} \chi(d) 1(n/d) = \sum_{d \mid n} \chi(d) = \sum_{\substack{d \mid n \\ d \equiv 1 \bmod 4}} 1 \; - \sum_{\substack{d \mid n \\ d \equiv 3 \bmod 4}} 1 \; = d_1(n) - d_3(n).$$

Por tanto, tenemos que demostrar que $r_2^+(n) = (\chi * 1)(n)$. Ambas funciones son multiplicativas: para $r_2^+(n)$ es consecuencia del teorema 56, mientras que $\chi * 1$ es multiplicativa por serlo las funciones χ y 1, como consecuencia de la proposición 55. De la proposición 53 se deduce que basta con demostrar que ambas funciones coinciden sobre las potencias de primos, para lo cual utilizaremos la siguiente propiedad: si p es primo y $e \in \mathbb{N}$, entonces los divisores de p^e son precisamente p^c variando c en el conjunto $\{0, 1, \ldots, e\}$. Separamos a continuación la demostración según las congruencias módulo 4.

- Para $p = 2$, tenemos que $r_2^+(2^e) = r_2^+(1) = 1$ por la proposición 44. Por otro lado, todos los divisores de 2^e distintos de $d = 1$ son pares, luego $d_1(2^e) - d_3(2^e) = 1$ y se da la igualdad.
- Para un primo $p \equiv 3 \bmod 4$, aplicando la proposición 46 y el teorema 43 tenemos que

$$r_2^+(p^e) = \begin{cases} r_2^+(1) = 1 & \text{si } e \equiv 0 \bmod 2, \\ 0 & \text{si } e \equiv 1 \bmod 2. \end{cases}$$

  Observamos que $p^c \equiv 1 \bmod 4$ si y solo si c es par y que $p^c \equiv 3 \bmod 4$ si y solo si c es impar, por lo que $d_1(p^e) - d_3(p^e)$ es la diferencia entre la cantidad de números pares e impares en el conjunto $\{0, 1, \ldots, e\}$, que claramente coincide con el valor arriba indicado en la expresión de $r_2^+(p^e)$.
- Si $p \equiv 1 \bmod 4$, entonces todos los divisores de p^e son congruentes con 1 mód 4. Así, $d_1(p^e) - d_3(p^e)$ es precisamente el número de divisores de p^e, que es $e+1$. Por otro lado, $r_2^+(p^e)$ es el número de ideales (α) de norma p^e. Como p es un primo que descompone, esto es, $p = \pi\overline{\pi}$ con $N(\pi) = N(\overline{\pi}) = p$, tenemos que $\alpha\overline{\alpha} = \pi^e \overline{\pi}^e$ y, por la factorización única, los ideales mencionados son de la forma $(\alpha) = (\pi)^r (\overline{\pi})^s$ con $r + s = e$ y $0 \leq r, s$, luego hay precisamente $e + 1$ de ellos. ∎

6. Conclusiones

En este artículo hemos visto cómo la teoría de números puede nutrirse del álgebra y hemos dado unas pinceladas de algunos conceptos de la teoría algebraica de números, como son los primos que descomponen, que ramifican o que son inertes, y también de algunos conceptos de teoría analítica de números, como las funciones multiplicativas y la convolución de Dirichlet. Todos estos detalles nos han servido para resolver un problema clásico que muestra muy bien cómo es la teoría de números, en la que para la demostración de un enunciado aparentemente inocente uno debe valerse de muy diversas herramientas.

No queremos terminar sin comentar algunos problemas relacionados con la suma de dos cuadrados que no hemos podido tratar. Por ejemplo, hemos estudiado el número representaciones que existen de un entero n, pero no hemos hablado de cómo obtenerlas. Puede encontrarse un ejemplo de este cálculo en el libro de Niven, Zuckerman y Montgomery [7, capítulo 3, ejemplo 3]. En general, el problema se reduce a calcular las representaciones para los primos que dividan a n y utilizar la observación 45. Claramente es necesario el uso de ordenadores para valores grandes de n, siendo interesante estudiar los algoritmos de resolución de este problema.

Similar al estudio que hemos hecho para dos cuadrados, uno puede preguntarse por el valor de la función $r_k(n)$ para distintos valores de k. Está claro que en nuestro estudio ha resultado fundamental utilizar los enteros gaussianos $\mathbb{Z}[\mathrm{i}]$. El caso $k = 4$ se puede estudiar de manera similar, sustituyendo $\mathbb{C}$ por los cuaterniones[15] y $\mathbb{Z}[\mathrm{i}]$ por los cuaterniones con coeficientes enteros[16]. Uno puede de este modo demostrar el siguiente teorema.

Teorema 58 (teorema de los cuatro cuadrados de Lagrange). *Todo entero no negativo es suma de cuatro cuadrados de enteros.*

El teorema es equivalente a afirmar que $R_4(n) \neq \varnothing$ para todo $n \in \mathbb{N}$, y también a que $r_4(n) \geq 1$ para todo $n \in \mathbb{N}$. Se conoce, además, la fórmula explícita de $r_4(n)$, dada por Jacobi.

Teorema 59. *Para $n \in \mathbb{N}$, $r_4(n) = 8 \sum_{4 \nmid d \mid n} d$.*

La suma se toma sobre los divisores de n que no son divisibles por 4. Existen varias maneras de demostrar esta fórmula: una es observando que una suma de cuatro cuadrados no es más que dos sumas de dos cuadrados, luego la fórmula de $r_4(n)$ puede deducirse si conocemos la expresión de $r_2(n)$ sin excesivo trabajo (véase el libro de Davidoff, Sarnak y Valette [4, sección 2.4], donde se trata el caso n impar, pero el caso n par es consecuencia de los resultados que allí aparecen); otra manera consiste en estudiar funciones modulares, que son interesantes por sí mismas y están muy presentes en la actualidad en la teoría de números; también se puede intentar seguir los pasos que hemos dado en este artículo y utilizar la factorización sobre los cuaterniones con coeficientes enteros, aunque esta opción es bastante complicada porque se trata de un anillo no conmutativo (para estudiar dicha factorización puede consultarse el libro de Conway y Smith [3, capítulo 5]), y también existen demostraciones elementales, como, por ejemplo, el artículo de Spearman y Williams [8] o el libro de Williams [10, capítulo 9], donde se hace una demostración basada en unas identidades de Liouville nada evidentes, aunque elementales.

Referencias

[1] Cobos Rábano, Alberto. *On certain algebraic, arithmetic and topological properties of quaternions.* Trabajo de Fin de Grado. Universidad de Salamanca, 2018.

[2] Conrad, Keith. *The gaussian integers.* Expository papers. 2016. URL: http://www.math.uconn.edu/~kconrad/blurbs/.

[15] Por cuaterniones nos referimos a los **cuaterniones de Hamilton**, que son, salvo isomorfismo algebraico, el único $\mathbb{R}$-álgebra de dimensión finita no conmutativa. Más explícitamente, se trata de elementos de la forma $a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ tales que $a, b, c, d \in \mathbb{R}$ y donde $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$.

[16] Para ser más precisos, es conveniente considerar un subanillo de los cuaterniones con coeficientes enteros, denominado **anillo de cuaterniones de Hurwitz**, por ser este un dominio euclídeo. Se muestra así que la teoría de divisibilidad es relevante en el problema de determinar $r_k(n)$ en general, y no solo en el caso $k = 2$.

[3] Conway, John H. y Smith, Derek A. *On quaternions and octonions: their geometry, arithmetic, and symmetry*. A K Peters, Ltd., 2003, págs. xii+159. isbn: 978-1-56881-134-5.

[4] Davidoff, Giuliana; Sarnak, Peter, y Valette, Alain. *Elementary number theory, group theory, and Ramanujan graphs*. Vol. 55. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 2003, págs. x+144. https://doi.org/10.1017/CBO9780511615825.

[5] Hardy, Godfrey H. y Wright, Edward M. *An introduction to the theory of numbers*. 3rd ed. Oxford, at the Clarendon Press, 1954, págs. xvi+419.

[6] Jacobson, Nathan. *Basic algebra. I*. Second edition. W. H. Freeman y Company, New York, 1986, págs. xviii+499. isbn: 978-0-7167-1480-4.

[7] Niven, Ivan; Zuckerman, Herbert S., y Montgomery, Hugh L. *An introduction to the theory of numbers*. Fifth. John Wiley & Sons, Inc., New York, 1991, págs. xiv+529. isbn: 978-0-471-62546-9.

[8] Spearman, Blair K. y Williams, Kenneth S. «The simplest arithmetic proof of Jacobi's four squares theorem». En: *Far East Journal of Mathematical Sciences (FJMS)* 2.3 (2000), págs. 433-439. issn: 0972-0871.

[9] Weil, André. *Number theory. An approach through history from Hammurapi to Legendre*. Reprint of the 1984 edition. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007, págs. xxii+377. isbn: 978-0-8176-4565-6.

[10] Williams, Kenneth S. *Number theory in the spirit of Liouville*. Vol. 76. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 2011, págs. xviii+287. isbn: 978-0-521-17562-3.