Sumas de cuadrados

Antonio Pablo Lozano Vicente

Trabajo de fin del grado de Matemáticas

Universidad de Zaragoza

Prólogo

El propósito de este trabajo es dar una idea general del problema de la suma de cuadrados, y de cómo se ha desarrollado a lo largo del tiempo. Se hace un análisis más profundo de dos de los problemas más clásicos de la teoría de números: el teorema de Fermat y el teorema de Lagrange. El primero trata de la suma de dos cuadrados, mientras que el segundo, trata de la suma de cuatro cuadrados. Sobre estos dos teoremas se ha escrito mucha literatura y se han conseguido demostraciones muy originales, aplicando diferentes herramientas matemáticas. En este trabajo se darán las demostraciones clásicas de estos resultados, haciendo algún comentario sobre otras pruebas cuando sea oportuno y sin profundizar demasiado.

En la primera parte del trabajo, se da una pequeña introducción explicando el origen del estudio de las sumas de cuadrados, y de cómo se fueron desarrollando los dos teoremas a tratar, para acabar con su final demostración.

Se comenta, cómo diversas eminencias en el campo de las matemáticas, han estado involucrados en la resolución y desarrollo de estos problemas.

Los capítulos centrales dan demostraciones rigurosas de estos dos grandes teoremas, incluyendo una segunda demostración en el caso de los dos cuadrados, y un esquema de una demostración alternativa en el caso de los cuatro cuadrados. Remarcar que para estas demostraciones, se suele introducir la ley de reciprocidad cuadrática para probar ciertos resultados clave. Sin embargo, en este trabajo no se hace uso de esta herramienta para probar los teoremas.

El último capítulo, se encarga de dar una visión global de cómo ha evolucionado el problema de la suma de cuadrados. Así, han aparecido varias generalizaciones y se han desarrollado herramientas muy útiles y versátiles para la resolución de diversos problemas, no sólo referentes a la teoría de números. Estas extensiones han dado lugar a campos de estudio completamente nuevos, llegando a crear toda una nueva disciplina matemática como es la teoría de cuerpos de clases.

La realización de este trabajo fue en un principio causa de la obligación. Sin embargo, una vez que estuve enfrascado entre los diferentes textos, resultó agradable e interesante investigar sobre la teoría de números, la cual ha despertado una gran curiosidad en mi y que de otra forma es muy probable que no hubiera llegado a conocer.

Por último agradecer al Dr. Javier Otal, que me ha acompañado en este proyecto dirigiendo este trabajo y ha tenido una gran paciencia conmigo en su labor de corrector. Decir que ha sido una magnífica experiencia, haber realizado esta pequeña escaramuza en el mundo de la teoría de números de la mano de este excelente profesor y amigo.

Antonio P. Lozano

English Summary

The main aim in this final year dissertation is to proof two of the most classic theorems in Number Theory: Fermat's Theorem on the Sum of two Squares and Lagrange's four-square theorem. Both theorems, are about determinate which integers can be expressed as the sum of a given number of squares, that is, which have the form $x_1^2 + \cdots + x_n^2$ where $x_i \in \mathbb{Z}$, for a given n.

In referring to the study of this kind of abstract subjects, people usually ask about why anybody should want to know such facts. While mathematicians rarely raise such questions, they often answer them by pointing out the usefulness of abstract mathematics in physics, engineering, and other scientific disciplines. Also, in the present instance, a case can be made for the usefulness of the study of representations of integers by sums of squares in lattice point problems, crystallography, and certain problems in mechanics.

The theorems

The first theorem to deal with is the Fermat's theorem on sums of two squares wich states that an odd prime p is expressible as

$$p = a^2 + b^2,$$

with a and b integers, if and only if $p \equiv 1 \pmod 4$.

However, Albert Girard had already made a determination of the numbers (not necessarily primes) expressible as a sum of two integral squares. Fermat was the first to claim a proof of it; he announced this theorem in a letter to Marin Mersenne dated December 25, 1640: for this reason this theorem is sometimes called Fermat's Christmas Theorem. Fermat stated that he possessed an irrefutable proof. Elsewhere he stated that his proof was by the method of infinite descent. The method of infinite descent is a method of demonstration which was developed by Fermat. This method relies on the facts that the natural numbers are well ordered and that there are only a finite number of them that are smaller than any given one. Nevertheless, Fermat usually did not write down proofs of his claims, and he did not provide a proof of this statement. The first proof was found by Euler in the middle of the eighteen century after much effort and is based on infinite descent.

In connection with the Lagrange's four-square theorem, it states that any natural number n is expressible as

$$n = a^2 + b^2 + c^2 + d^2$$

with a, b, c, d integers. Bachet remarked in the translation notes of Arithmetica that Diophantus apparently assumed that any number is either a square or the sum of 2, 3 or 4 squares. This result is also known as Bachet's theorem. Fermat, again stated that he possessed a proof that every number is sum of four squares. It was not until 1770 that Lagrange proved the theorem.

The origins

The origin of these results can be found in Diophantus (325-409 A.D.). He propossed a collection of statements connected with characterize of the set of integers, for which the diophantine equation

$n = x^2 + y^2$ has solution. Often those statments are in geometric language, and the meaning of his statements is not always clear. At least some of them appear to be incorrect; one of them, however, is equivalent to the important identities

$$(a_1^2+b_1^2)(a_2^2+b_2^2) = (a_1a_2-b_1b_2)^2+(a_1b_2+b_1a_2)^2 = (a_1a_2+b_1b_2)^2+(a_1b_2-b_1a_2)^2$$

It is easy to prove this identity by applying the multiplicative property of the Gaussian integer's norm (see below). The identity shows that if two numbers are expressible as sum of two squares, its product is too.

Among the mathematicians who studied this problem were Bachet (1581-1638), famous for his own edition of Diophantus, and Girard (1595-1632) who first stated the necessary and sufficient conditions on n for the solvability of $n = x^2 + y^2$ in integers: *n has to be either a square, or a prime* $p \equiv 1$ (mód 4)*, or a product of such numbers, or the product of one of the preceding with a power of 2*.

Shortly afterwards, Fermat (1601-1665) who had a Bachet's copy of Diophantus' Arithmetica, stated, as a condition on n, that $n \equiv 1 \pmod 4$ and that when n is divided by its largest square factor, the quotient should not contain any prime $q \equiv 3 \pmod 4$. In fact, he indicates (in a numerical example), how to compute the number of ways to express such number as sum of two squares. Fermat called the theorem that every prime $4n+1$ is a sum of two squares, the fundamental theorem on right triangles. He worked in solving the proposed problems making some annotations in the margin. In one of those annotations is the famous Fermat's Last Theorem.

The problem of characterizing which numbers can be expressed as a sum of three squares is more difficult than the two squares or the four squares problem, what is more important, requires different methods. However, Gauss proved that n is expressible as a sum of three squares if and only if $n \neq 4^e(8k+7)$; thus $7, 15, 23, 28, \ldots$ are not sums of three squares.

There is no explicit statement of the Lagrange's four-square theorem in Diophantus; however, while he requires conditions for an integer to be a sum of two or of three squares, he states no conditions whatsoever for n to be a sum of four squares. Bachet interpret this as indicating a knowledge of the four squares theorem. Bachet did state the theorem explicitly and mention that he had verified it up to 325, but had no general proof of it. As mentioned earlier, Fermat claimed (again) to have a proof of the theorem. Fermat, in a letter to Carcavi, indicated that his proof (like so many others of his) was based on the method of descent, whose application in this case, he added, required another new idea. After several attempts to prove the four square theorem, especially from Euler and Goldbach, Euler published some relevant results. Euler proved, among others, that there exists integers a, b such that $1 + a^2 + b^2$ is divisible by a given prime p. Also, Euler gave the fundamental formula

$$\begin{cases} (a^2+b^2+c^2+d^2)(p^2+q^2+r^2+s^2) = x^2+y^2+z^2+v^2, \\ x = ap+bq+cr+ds, \qquad y = aq-bp \pm cs \mp dr, \\ z = ar \mp bs - cp \pm dq, \qquad v = as \pm br \mp cq - dp. \end{cases}$$

These results are essential to prove the four square theorem. Finally, Lagrange gave the first proof of the four square theorem and acknowledged his indebtedness to ideas in the preceding results by Euler.

The proofs

The classic proofs for the Fermat's Theorem on the Sum of two Squares and Lagrange's four-square theorem are based on the method of infinite descent. Let's take a closer look at this method.

The method of infinite descent is a particular kind of proof by contradiction which relies on the facts that the natural numbers are well ordered and that there are only a finite number of them that are smaller than any given one. Fermat's own account of it is to be found a letter 'Relation nouvelles entitled decouvertes la science des nombres' which he wrote to Carcavi in 1659. In this letter he tells Carcavi that he has discovered a new method demonstration and applied it successfully to the

solution of a considerable number of problems in the theory of numbers. He calls it the method of infinite unlimited descent (descente infinie indefinie) says that first applied only negative propositions as: *'There is no right triangle whose sides are integers whose area is equal to the square of an integer.'*

He gives the following abstract of his proof of the latter: *'The proof is made by reductio ad absurdum in this manner: If there is a right triangle with integral sides and with an area equal to the square of an integer, then there is a second triangle, smaller than the first, which has the same preperty; and if there is a second triangle, smaller than the first, which has the same property, then there is, by like reasoning, a third smaller than the second; and then a fourth, a fifth, and so on ad infinitum. But there is not an infinite number of integers less than a given integer. From which one concludes thar it is impossible that there should be a right triangle with integral sides and with an area wich is the square of an integer.'*

In the next paragraph Fermat says: 'It was a long time before I was able to apply my method to affirmative questions, because the way and manner of getting at them is much more difficult than that which I employ with negative theorems. So much so that, when I had to prove that every prime number of the form $4k+1$ is made up of two squares, I found myself in much torment. But at last a certain reflection many times repeated gave me the necessay light, and affirmative questions yielded to my method with some new principles by which sheer necessity compelled me to supplement it. This development of my argument in the case of affirmative questions takes the following line. If a prime number of the form $4k+1$ selected at random is not made up of two squares, there will exist another prime number of the same sort but less than given number, and again a third still smaller and so on descending ad infinitum until one comes to the number 5, which i the smallest of all numbers of the kind in question and which the argument would require not to be made up of two squares, although in fact it is so made up. From which one must infer, by reductio ad absurdum, that all numbers of the kind in question are in consequence made up of two squares.'

The foregoing is a good account of Fermat's method although it leaves out all details of the proof. Fermat's proof, like so many others of his, is no extant. As seen above, the proof was made by Euler.

Unlike many other texts of Numbers theory, the proofs in this final year dissertation are independent of the law of quadratic reciprocity.

Before begining the proofs, let's see some useful results and definitions.

Recall that a Gaussian integer is a complex number whose real and imaginary parts are both integers. The Gaussian integers, with ordinary addition and multiplication of complex numbers, form an integral domain, usually written as $\mathbb{Z}[i]$. The norm of a Gaussian integer is defined as the square of its absolute value as a complex number and a natural number:

$$\begin{array}{rcl} N:\mathbb{Z}[i] & \longrightarrow & \mathbb{N}\cup\{0\} \\ z & \longmapsto & N(z)=z\cdot\bar{z} \end{array}$$

The norm is multiplicative i.e. $N(zw)=N(z)N(w)$ and the units of $\mathbb{Z}[i]$ are precisely those elements with norm 1, i.e. the set $\{\pm 1,\pm i\}$.

Definition 1. *For each integer $k\geq 1$, let $S_k=\{n \mid n=x_1^2+...+x_k^2,\ x_1,...,x_k\in\mathbb{Z}\}$, the set of all sums of k squares.*

As a consequence of the mutiplicativity of the norm, it follows the next lemma

Lemma 2. *S_2 is closed under multiplication.*

Demostración. Let $s,t\in S_2$, $s=a_1^2+b_1^2$ and $t=a_2^2+b_2^2$. Let's define $z=a_1+i\cdot b_1$ and $w=a_2+i\cdot b_2$. Now, the lemma follows from the multiplicativity of the norm:

$$N(z)N(w)=(a_1^2+b_1^2)(a_2^2+b_2^2)=(a_1a_2-b_1b_2)^2+(a_1b_2+b_1a_2)^2=N(zw)$$

□

The next result is crucial to proof the Fermat's theorem.

Lemma 3. *For primes $p = 4m+1$ the equation $x^2 \equiv -1 \pmod{p}$ has two solutions $s \in \{1,2,\ldots,p-1\}$, for $p = 2$ there is one such solution, while the primes of the form $p = 4n+3$ there is no solution.*

Demostración. In this proof is where usually other books introduce the law of quadratic reciprocity. However, let's proof the lemma using some simple principles.

For odd p, and $x \in \{1,..,p-1\}$ let P_x be the set $\{x,-x,x^{-1},-x^{-1}\}$ with $x \in \{1,2,...,p-1\}$. with $-x$ the additive inverse and x^{-1} the multiplicative inverse over $\mathbb{Z}_p$. Those sets contains 4 elements with the exception some of the following cases

- $x \equiv -x$ es impossible for odd p.
- $x \equiv x^{-1}$, is equivalent to $x^2 \equiv 1$. This have two solutions, namely $x = 1$ and $x = p-1$, and then $P_1 = P_{p-1} = \{1,-1\} = \{p-1,1-p\} = \{1,p-1\}$
- $x \equiv -x^{-1}$ is equivalent to $x^2 \equiv -1$. This equation may have no solution or two distinct solutions $x = x_0$, $x = p - x_0$; in this case, $P_x = \{x_0, p-x_0\}$.

The set $\{1,..,p-1\}$ has $p-1$ elements, and we partitioned it into quadruples, plus one or two pairs (depending on the size of P_x). For $p-1 = 4m+2$, there is only one pair $\{1,p-1\}$ the rest is quadruples, and thus $x^2 \equiv -1 \pmod{p}$ has no solution. For $p-1 = 4m$ there has to be a second pair, and this contains the two solutions of $x^2 \equiv -1$. □

Theorem 4. *An odd prime p is expressible as $p = x^2+y^2$, with $x,y \in \mathbb{N}$, if and only if $p \equiv 1 \pmod{4}$.*

Demostración. The proof is as follows

1. S_2 is closed under multiplication.
2. Let p to be a prime $4n+1$, p divide to $1+x^2$ for some x, choosing x such that $mp = x^2+1$ with $0 < m < p$
3. Take the less m such that $x^2+y^2 = mp$, with p prime and $1 < m < p$, then find $x_1^2+y_2^2 = m_1 p$, with x_1, y_1 integers y $0 < m_1 < m$ and this contradicts the minimality of m.

□

Generalizing for any integer:

Theorem 5. *A positive integer n is a sum of two squares if and only if every prime $q \equiv 3 \pmod{4}$ divides n to an even power.*

Demostración. It is a rather easy consequence of the unique factorization property in the ring $\mathbb{Z}$, Lemma 2 and Theorem 4 at once. □

Alternatively, Fermat's theorem can be proven by using Gaussian integers.

First at all, recall some properties of $\mathbb{Z}[i]$. We can define the divisibility and factorization in $\mathbb{Z}[i]$ in the same way that in the integers. So, like in $\mathbb{Z}$, we obtain that if p is irreducible in $\mathbb{Z}[i]$, it is prime too. Also, $\mathbb{Z}[i]$ is a unique factorization domain (UFD).

Theorem 6. *An odd prime p is expressible as $p = x^2+y^2$, with $x,y \in \mathbb{N}$, if and only if $p \equiv 1 \pmod{4}$.*

Demostración. The alternative proof is as follows

1. S_2 is closed under multiplication.

2. Let p to be a prime $4n+1$, p divide to $1+x^2$ for some x, choosing x such that $mp = x^2+1$ with $0 < m < p$

3. Let x such that $p|1+x^2 = (1+xi)(1-xi)$ and then prove that $p \nmid (1+xi)$ and $p \nmid (1-xi)$ then p is not prime in $\mathbb{Z}[i]$.

4. If p is not prime in $\mathbb{Z}[i]$ then $p \in S_2$.

□

Just as the two-squares identity can be explained in terms of complex numbers, a similar four-squares identity can be derived from a generalisation of complex numbers known as the quaternions $\mathbb{H}$. The elements of $\mathbb{H}$ are the points $q = (a,b,c,d) \in \mathbb{R}^4$, and addition and subtraction are performed by the usual method for vectors. To define multiplication, it is useful to write each quaternion in the form $q = a1 + bi + cj + dk$, where $a,b,c,d \in \mathbb{R}$ and $1,i,j,k$ denote the standard basis vectors of $\mathbb{R}^4$. Defining the products of the basis vectors by

$$i^2 = j^2 = k^2 = -1$$

$$ij = -ij = k$$

$$jk = -kj = i$$

$$ki = -ik = j$$

Notice that multiplication is not commutative, since $ij = -ji$ for example. By assuming distributivity (that is, $q(q' + q") = qq' + qq"$ and $(q' + q")q = q'q + q"q$ for all $q, q', q"$), we find that the product of any pair of quaternions $q_1 = (a_1^2 + b_1^2 + c_1^2 + d_1^2)$ and $q_2 = (a_2^2 + b_2^2 + c_2^2 + d_2^2)$

$$\begin{aligned} q_1q_2 = (a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2) &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)^2 \\ &+ (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)^2 \\ &+ (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)^2 \\ &+ (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)^2 \end{aligned}$$

The conjugate of a quaternion $q = a + bi + cj + dk$ is the quaternion $\overline{q} = a - bi - cj - dk$. The norm of a Quaternion q is defined as the square root of the product $q\overline{q}$. It will be denoted by $||q||$. Since conjugation is an automorphism, it follows:

$$||q_1q_2|| = ||q_1|| \cdot ||q_2||$$

So the norm is multiplicative. As a consequence of this, it is easy to proof the next lemma.

Lemma 7. *S_4 is closed under multiplication.*

Demostración. Let $x_1 = a_1^2 + b_1^2 + c_1^2 + d_1^2 = ||1 \cdot a_1 + i \cdot b_1 + j \cdot c_1 + k \cdot d_1||^2 = ||q_1||^2$ and $x_2 = a_2^2 + b_2^2 + c_2^2 + d_2^2 = ||1 \cdot a_2 + i \cdot b_2 + j \cdot c_2 + k \cdot d_2||^2 = ||q_2||^2$. Now, the lemma follows from the multiplicativity of the norm:

$$x_1x_2 = ||q_1||^2||q_2||^2 = ||q_1q_2||^2$$

Then

$$\begin{aligned} (a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2) &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)^2 \\ &+ (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)^2 \\ &+ (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)^2 \\ &+ (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)^2 \end{aligned}$$

□

In the same way as in the two squares theorem the proof of Lagrange's theorem is supported by the following result.

Lemma 8. *If p is an odd prime, then $a^2+b^2+1=kp$ for some integers a,b,k with $0<k<p$.*

Demostración. Let $p=2n+1$. Consider the sets $A:=\{a^2 \mid a=0,1,\ldots,n\}$ and $B:=\{-b^2-1 \mid b=0,1,\ldots,n\}$. We have the following facts:

- No two elements in A are congruent mod p, for if $a^2 \equiv c^2 \pmod{p}$, then either $p \mid (a-c)$ or $p \mid (a+c)$ by unique factorization of primes. Since $a-c, a+c \leq 2n < p$, and $0 \leq a, c$, we must have $a=c$.
- Similarly, no two elements in B are congruent mod p.
- Furthermore, $A \cap B = \emptyset$ since elements of A are all non-negative, while elements of B are all negative.
- Therefore, $C := A \cup B$ has $2n+2$, or $p+1$ elements.

Therefore, two elements in C must be congruent mod p. In addition, by the first two facts, the two elements must come from different sets. As a result, we have $a^2+b^2+1=kp$ for some k. Clearly k is positive. Also, $p^2=(2n+1)^2>2n^2+1 \geq a^2+b^2+1=kp$, so $p>k$. □

Theorem 9. *Every non-negative integer is a sum of four squares.*

Demostración. The proof is as follows

1. S_4 is closed under multiplication.
2. Let p to be a prime, p divide to $1+x^2+y^2$ for some x,y.
3. Take the less m such that $x^2+y^2=mp$, with p prime and $1<m<p$, then find $x_1^2+y_2^2=m_1 p$, with x_1, y_1 integers y $0<m_1<m$ and this contradicts the minimality of m.

□

As in the case of the Fermat's theorem, we can give a more algebraic proof of four squares theorem using the quaternions. For this purpose, we define the Hurwitz quaternion (also known as Hurwitz integers) as $\mathbb{H}_1 := \{a+bi+cj+dk \mid a,b,c,d \in \mathbb{Z} \text{ or } a,b,c,d \in \mathbb{Z}+1/2\}$ which is a subring of the ring of all quaternions $\mathbb{H}$. Every irreducible number over $\mathbb{H}_1$ is prime over $\mathbb{H}_1$.

Theorem 10. *Every non-negative integer is a sum of four squares.*

Demostración. The proof is as follows

1. S_4 is closed under multiplication.
2. Let p to be a prime, p divide to $1+x^2+y^2$ for some x,y.
3. Let x,y such that $p|1+x^2+y^2=(1+xi+yj)(1-xi-yj)$ and then prove that $p \nmid (1+xi+yj)$ and $p \nmid (1-xi-yj)$ then p is not prime in $\mathbb{H}_1$.
4. If p is not prime in $\mathbb{H}_1$ then $p \in S_4$.

□

Other sums of powers

There exist many extensions and generalizations of these results. Once the theorems are proven, one of the questions may emerge is how to express a number as sums of a determinated number of squares or how many representations as sums of squares are possibles. In 1834, Jacobi found an exact formula for the number of ways a given positive integer can be represented as the sum of four squares.

One of the generalization is the Fermat polygonal number theorem. This theorem states that every positive integer is a sum of at most n n-gonal numbers. That is, every positive integer can be written as the sum of three or fewer triangular numbers, and as the sum of four or fewer square numbers, and as the sum of five or fewer pentagonal numbers, and so on. Gauss proved the triangular case in 1796, commemorating the occasion by writing in his diary the line *'ΕΥΡΗΚΑ! num* $= \triangle + \triangle + \triangle$*'*. The full polygonal number theorem was finally proven by Cauchy in 1813. A shorter proof given by Nathanson (1987) is based on the following lemma due to Cauchy:

> For odd positive integers a and b such that $b^2 < 4a$ and $3a < b^2 + 2^b + 4$ we can find nonnegative integers s,t,u, and v such that $a = s^2 + t^2 + u^2 + v^2$ and $b = s + t + u + v$.

We also can generalize these theorems taking linear combinations of sums of squares, i.e. to caracterize which numbers can be expressed as $x^2 + ay^2$ where a is a given number and $x, y \in \mathbb{Z}$. This question is partially answered by Gauss Reciprocity Law, and has ultimately lead to a whole new mathematical discipline, namely the Class Field Theory. Another generalization is about sums of higher powers, that is, to determinate which numbers are expressibles as sums of k-th powers. This problem is known as Waring's problem.

Waring stated *'Every integer is a cube or the sum of two, three, ... nine cubes; every integer is also the square of a square, or the sum of up to nineteen such; and so forth.'*.

It is presumed that by this, in modern notation, Waring meant that for every $k \geq 3$ there are numbers s such that every natural number is the sum of at most s k-th powers of natural numbers and that the smallest such number $g(k)$ satisfies $g(3) = 9$, $g(4) = 19$.

Hilbert proved the existence of a such number $g(k) < \infty$ for any k, but the proof did not show how to find those numbers. Some of those values are known. Obviously the case $k = 2$ is Lagrange theorem. Other known values are $g(3) = 9$, $g(4) = 19$, $g(5) = 37$, $g(6) = 73$. However, in the case $k = 3$, only 23 and 239 requires as many as nine cubes. That fact, the Waring problem is a bit more generalizated. $G(k)$ is defined to be the least positive integer s such that every sufficiently large integer (i.e. every integer greater than some constant) can be represented as a sum of at most s kth powers of positive integers. There are not a way to determine those numbers but there exists some bounds.

A variation of this problem known as the 'easier' Waring problem in which sums and differences of powers are taken ($n = \pm x_1^k \pm \cdots \pm x_s^k$ where n is given and x_i are integers).Another problem related to the sums of squares is the Taxicab problem. The name is derived from a conversation involving mathematicians Hardy and Ramanujan. As told by Hardy:

I remember once going to see him when he was lying ill at Putney. I had ridden in taxi-cab No. 1729, and remarked that the number seemed to be rather a dull one, and that I hoped it was not an unfavourable omen. "No", he replied, ït is a very interesting number; it is the smallest number expressible as the sum of two (positive) cubes in two different ways.

Typically denoted $Ta(n)$, the n-th Taxicab number is defined as the smallest number that can be expressed as a sum of two positive algebraic cubes in n distinct ways. The restriction of the summands to positive numbers is necessary, because allowing negative numbers allows for more (and smaller) instances of numbers that can be expressed as sums of cubes in n distinct ways. The concept of a taxicab number has been introduced to allow for alternative, less restrictive definitions of this nature. In a sense, the specification of two summands and powers of three is also restrictive; a generalized taxicab number allows for these values to be other than two and three, respectively. The generalized taxicab number $Taxicab(k, j, n)$ is the smallest number which can be expressed as the sum of j k-th positive powers in n different ways.

Índice general

Capítulo 1

Introducción

Esta pequeña introducción se divide en dos secciones. En la primera parte se describe cómo surgió el problema y fueron apareciendo diversos resultados que acabarían en los enunciados del teorema de Fermat y del teorema de Lagrange.

En la segunda parte aparece una explicación detallada del método del descenso infinito, una forma de demostración introducida por Pierre de Fermat en el s.XVII y muy utilizada por éste para la resolucion de diversos problemas en teoría de números. Este método se basa en el principio de buena ordenación y tiene cierta similitud con el método de inducción.

1.1. Sumas de cuadrados: El origen

El origen de estos resultados se encuentra en la resolución de diversos problemas planteados por Diofanto (s.III) quien es considerado como *"el padre del álgebra"*.

A continuación se muestran algunos de estos problemas que llevaron finalmente a los enunciados y demostraciones de los teoremas conocidos como teorema de Fermat y teorema de Lagrange.

Diofanto propuso encontrar cuatro números x_i tales que cada una de las ocho expresiones $E = (\sum x_j)^2 \pm x_i$ fuera un cuadrado. Ahora bien, en un triángulo rectángulo de lados $p,q,h \in \mathbb{N}$, que se denotará en adelante como $\triangle(p,q,h)$, se tiene que $h^2 \pm 2pq$ es un cuadrado, es decir $k^2 = h^2 \pm 2pq$ para algún k entero.

Si para $i = 1,\ldots,4$, se tiene $h^2 = p_i^2 + q_i^2$, definiendo $x = \frac{h}{\sum 2p_iq_i}$ y $x_i = 2p_iq_ix^2$ se tiene que $\sum x_i = hx$. Es sencillo comprobar que las ocho expresiones $(\sum x_j)^2 \pm x_i$ son cuadrados de números racionales $(\sum x_j)^2 \pm x_i = h^2x^2 \pm 2x^2a_ib_i = x^2(a_i \pm b_i)^2$.

Por tanto, para dar respuesta al problema planteado, basta encontrar cuatro triángulos rectángulos con igual hipotenusa, es decir, un cuadrado que pueda ser expresado como suma de dos cuadrados de cuatro maneras distintas. Por ejemplo, se consideran los triángulos $\triangle(3,4,5)$ y $\triangle(5,12,13)$. Si se multiplica la hipotenusa de cada uno de ellos por cada lado del otro triángulo, se obtienen dos triángulos con la misma hipotenusa $\triangle(39,52,65)$ y $\triangle(25,60,65)$.

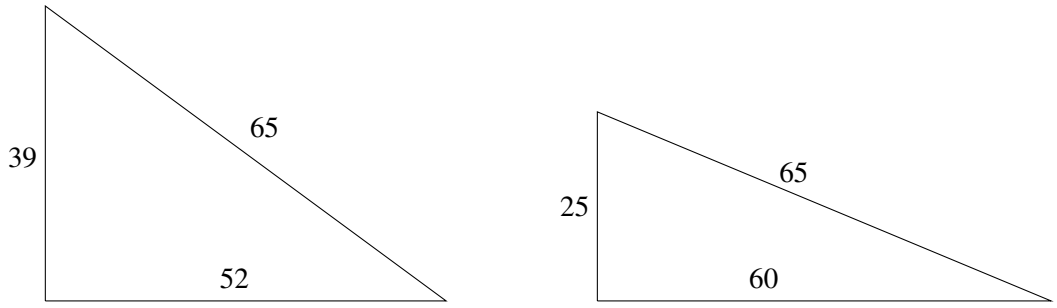

Figura 1.1: Dos triángulos rectángulos con igual hipotenusa.

Por otra parte $65 = 1^2 + 8^2 = 4^2 + 7^2$ y teniendo en cuenta que $\triangle(a^2 - b^2, 2ab, a^2 + b^2)$ es un triángulo rectángulo (se comprueba directamente), se obtienen cuatro triángulos rectángulos con igual hipotenusa y lados enteros: $\triangle(39,52,65)$, $\triangle(25,60,65)$, $\triangle(33,56,65)$ y $\triangle(16,63,65)$.

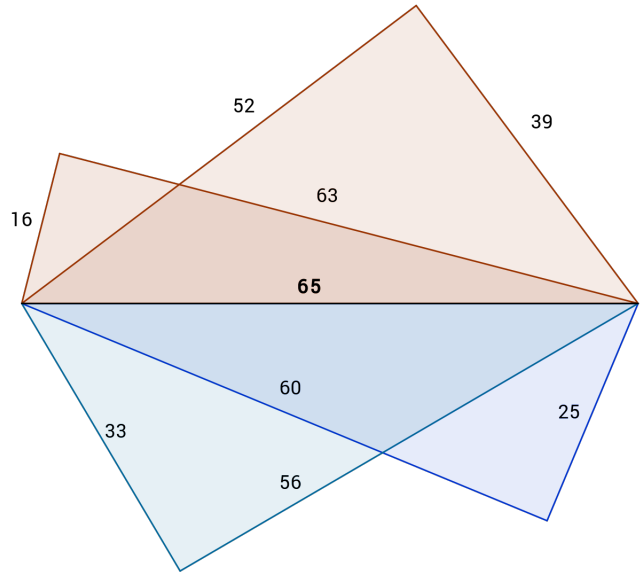

Figura 1.2: Cuaro triángulos rectángulos con igual hipotenusa.

Y con estos triángulos con igual hipotenusa se obtiene la solución al problema planteado

$$x_i = \frac{17136600}{163021824}, \frac{12675000}{163021824}, \frac{15615600}{163021824}, \frac{8517600}{163021824}.$$

Otro de los problemas propuestos por Diofanto fue el de descomponer la unidad en dos partes de modo que al añadir a cada una el mismo número dado se forme un cuadrado, es decir, dado a, encontrar x, y tales que $x+y=1$ y $x+a$ y $y+a$ sean cuadrados. Para ello el número dado no debe ser impar y el doble más uno, $2a+1$, no debe ser divisible por ningún número primo de la forma $4n-1$ Desafortunadamente la condición necesaria no se lee con claridad en los manuscritos de Diofanto.

Respecto esto último Fermat postuló lo siguiente: *'La verdadera condición (aquella que es general y excluye los números que son inadmisibles) es que el número dado a no puede ser impar y que* $2a+1$ *tras dividirlo por el mayor cuadrado como factor, no puede ser divisible por un número primo de la forma* $4n-1$*'.*

Realmente fué Girard en 1632 quien por primera vez dió una caracterización de los numeros que podian expresarse como suma de dos cuadrados: *«Todo cuadrado, todo número primo de la forma* $4n+1$*, un producto formado por éstos y el doble de uno de los anteriores»* Por esta razón, en algunos textos el teorema de los dos cuadrados es conocido como Teorema de Girard. A raíz de estos problemas y diversas soluciones que fueron surgiendo, sobre todo con casos particulares, Fermat comunicó en una carta a Mersenne el 25 de Diciembre de 1640 (Por lo que en ocasiones el teorema de la suma de dos cuadrados también es conocido como el Teorema de la Navidad) las siguientes observaciones:

- Todo número primo de la forma $4n+1$ es la hipotenusa de sólo un triángulo rectángulo, su cuadrado de dos, su cubo de tres y así sucesivamente.
- Un número primo de la forma $4n+1$ y su cuadrado son suma de dos cuadrados de manera única, su cubo y potencia cuarta es suma de dos cuadrados de dos formas distintas, su potencia quinta y sexta de tres y así sucesivamente.

- Es sencillo hallar cuantos triángulos rectángulos se pueden formar con una hipotenusa w dada. Se pueden formar $p^a q^b r^c s$ donde p,q,r son primos de la forma $4n+1$ y s es un cuadrado sin ninguno de los factores anteriores y definir $w = 2c(2ab+a+b)+2ab+a+b+c$
- Para encontrar un número que sea la hipotenusa de un número determinado w de triángulos rectángulos, basta tomar los factores primos de $2w+1$ y restar uno a cada uno de éstos y tomar la mitad de lo que queda como exponente para cualquier primo de la forma $4n+1$. Por ejemplo, si $w=7$, $2w+1 = 15 = (2+1)(2\cdot 2+1)$ por tanto pq^2 resuelve el problema con p,q primos de la forma $4n+1$.

Fermat llamó al teorema de que todo número primo de la forma $4n+1$ es una suma de dos cuadrados el teorema fundamental de los triángulos rectángulos. Y afirmó que tenía una prueba irrefutable de éste basada en el método del descenso infinito. Sin embargo, no fue hasta más de un siglo despúes cuando Euler publica una demostración tras varios intentos y resultados intermedios por parte de varios matemáticos como Goldbach o Jaquemet. Posteriormente se publicaron diversas demostraciones de este resultado relacionando varias ramas de las matemáticas.

En los problemas propuestos por Diofanto también se pueden encontrar cuestiones sobre representaciones superiores, decir qué números pueden ser expresados como suma de tres, cuatro, cinco cuadrados. Gauss demostró que un número es suma de tres cuadrados si y sólo si $n \neq 4^e(8k+7)$; por tanto $7,15,23,28,\ldots$ no son suma de tres cuadrados. Es fácil probar que ningún entero $n = 4^e(8k+7)$ puede expresarse como suma de tres cuadrados. Probar el recíproco es más complicado, principalmente porque el conjunto de números expresables como suma de tres cuadrados no es cerrado bajo la multiplicación.

Al igual que ocurre con la suma de dos cuadrados, se puede encontrar el origen del enunciado del teorema de los cuatro cuadrados en Diofanto quien propuso encontrar cuatro números x_i tales que la suma de sus cuadrados sumados (restados) con la suma de los x_i, diera como resultado un número dado n. Tomó como ejemplo $n=12$ ($n=4$). Para ello observar que $x^2 \pm x + 1/4 = (x \pm 1/2)^2$ es un cuadrado. Por tanto la suma de los cuatro cuadrados más la suma de sus lados más 1 es la suma de otros cuatro cuadrados, luego tendrá que ser igual a $13(5)$. Es decir $\sum x_i^2 \pm \sum x_i + 1$ es una suma de cuatro cuadrados, en este caso $13(5)$. Así pues si se divide $13(5)$ en cuatro cuadrados. Después se resta (suma) $1/2$ a cada uno de sus lados y se obtienen los lados de los cuadrados pedidos.

$$13 = 4+9 = \tfrac{64}{25}+\tfrac{36}{25}+\tfrac{144}{25}+\tfrac{81}{25}, \qquad \left(5 = \tfrac{9}{25}+\tfrac{16}{25}+\tfrac{64}{25}+\tfrac{36}{25}\right)$$

Por tanto, la solución al problema es

$$\tfrac{11}{10}+\tfrac{7}{10}+\tfrac{19}{10}+\tfrac{13}{10}, \qquad \left(\tfrac{11}{10}+\tfrac{13}{10}+\tfrac{21}{10}+\tfrac{17}{10}\right)$$

Bachet observó que Diofanto, tanto aquí como en otros enunciados, parecía asumir que cualquier número puede ser expresado como suma de dos, tres o cuatro cuadrados (resultado conocido como Teorema de Bachet) y afirmó haber demostrado la proposicion para números menores que 325 y hallado la descomposición en suma de cuatro cuadrados para números menores que 120. Fermat por su parte, aseguraba tener una demostracion de que todo número podía ser expresado como suma de cuatro cuadrados y que seguramente este teorema ya era conocido por Diofanto.

Descartes enunció el teorema (*'cuya demostración juzgó de tal dificultad que no se atrevió a intentarla'*): cualquier número que es suma de tres cuadrados y mayor que 41, puede también ser expresado como suma de cuatro cuadrados. Euler admitió que no pudo probar el teorema de Bachet de que cualquier número puede expresarse como suma de cuatro cuadrados, ni dar una regla general para expresar n^2+7 como suma de cuatro cuadrados. Sin embargo demostró el teorema para ciertos conjuntos de números y dió la fórmula fundamental

$$\begin{cases} (a^2+b^2+c^2+d^2)(p^2+q^2+r^2+s^2) = x^2+y^2+z^2+v^2, \\ x = ap+bq+cr+ds, \qquad y = aq-bp \pm cs \mp dr, \\ z = ar \mp bs - cp \pm dq, \qquad v = as \pm br \mp cq - dp. \end{cases}$$

A mediados del siglo XVIII, Euler declaró haber demostrado que si p es un número primo cualquiera, entonces existen cuatro números enteros a,b,c,d no divisibles por p, tales que $a^2+b^2+c^2+d^2$ es divisible por p. Después de diversos intentos para demostrar el teorema de Bachet, sobre todo por parte de Euler y Goldbach, Euler publicó algunos resultados sobre éste. Euler probó, entre otros, que existen enteros a,b tales que $1+a^2+b^2$ es divisible por un número primo p dado.

Fué finalmente Lagrange quien en 1770 dió la primera demostración del teorema de Bachet reconociendo que no habría sido posible sin las ideas y aportaciones de Euler. Lagrange añadió una generalización para la identidad dada por Euler:

$$\begin{gathered}(p^2-Bq^2-Cr^2+BCs^2)(p_1^2-Bq_1^2-Cr_1^2+BCs_1^2)=\\(pp_1+Bqq_1\pm C(rr_1\mp Bsq_1))^2-B(pq_1+qp_1\pm C(rs_1+sr_1))^2-\\C(pr_1-Bqs_1\pm rp_1\mp Bsq_1)^2+BC(qr_1-ps_1\pm sp_1\mp rq_1)^2\end{gathered}$$

Posteriormente, en 1773, Euler dió una demostración mucho más simple que la de Lagrange. A parte de las generalizaciones que surgieron de los resultados probados, estos teoremas han sido mirados desde muy diferentes ángulos dando con demostraciones ingeniosas y creando nexos entre diferentes aspectos de las matemáticas.

1.2. Método del descenso infinito

La primera vez que se menciona este método es en una carta de Fermat enviada a Pierre de Carcavi en 1659 cuyo titulo rezaba *"Relation des nouvelles découvertes en la science des nombres."*. En esta carta comunicaba a Carcavi que había descubierto un nuevo método de demostración y que lo había aplicado con exito en la resolución de un número considerable de problemas de la teoría de números. A esta forma de demostración la llamó método de descenso infinito o ilimitado y que en un principio sólo lo había aplicado a proposiciones negativas como:

«No existe ningún triángulo rectángulo cuyos lados sean números enteros y su área sea el cuadrado de un número entero»

Fermat argumentó de forma abstracta la demostración de la siguiente manera: *'La demostración se realiza por reducción al absurdo como sigue: si existe un triángulo de lados enteros y área el cuadrado de un entero, entonces existe un segundo triángulo menor que el primero con esta misma propiedad; y si existe un segundo triángulo menor que el primero con la misma propiedad, entonces existe un tercero menor que el segundo, y entonces hay un cuarto, un quinto y así ad infinitum. Pero no hay un número ilimitado de enteros menor que uno dado. Por tanto es imposible que pueda haber un triángulo rectángulo cuyos lados sean números enteros y su área sea el cuadrado de un número entero'*

En el siguiente parrafo de la carta, Fermat comentaba que había encontrado grandes dificultades para aplicar este método a enunciados no negativos, lo cual le produjo grandes quebraderos de cabeza a la hora de intentar demostrar el teorema de los dos cuadrados. Para este último afirmaba tener finalmente una demostración basada en el metodo del descenso infinito aunque según decía, necesitó emplear algunos nuevos principios. Expuso en esta carta un esquema de los pasos que había seguido en la demostración pero ésta estaba falta de detalles.

Notar que este método se basa principalmente la aplicación del principio de buena ordenación (todo conjunto no vacío de números naturales posee un número mínimo). Otro ejemplo algo más detallado de este tipo de razonamiento se ve en el siguiente ejemplo.

Se consideran ternas pitagóricas, es decir aquellas ternas de números naturales (a,b,c) tales que $c^2=b^2+a^2$, o dicho de otra forma, son lados de un triángulo rectángulo.

Teorema 1.2.1. *No existen ternas pitagóricas que se correspondan con un triángulo isosceles, o lo que es lo mismo que no existen ternas pitagóricas de la forma* (a,a,c).

Demostración. Si existiera una tal tupla (a,a,c) se tendría $c^2 = 2a^2$, esto querría decir que c^2 es par luego c también es par. Sea pues $c = 2c_1$ con $c_1 \in \mathbb{N}$. Se tiene la igualdad $4c_1^2 = 2a^2$, es decir $a^2 = 2c_1^2$, por tanto a es par; sea ahora $a = 2a_1$, es decir $c_1^2 = 2a_1^2$ de donde se tiene otra terna (a_1,a_1,c_1) que es un triángulo estrictamente menor que el primero. Aplicando este proceso sobre el nuevo triángulo se obtendría otro (a_2,a_2,c_2) menor que el anterior. Repitiendo este proceso, se tiene una secuencia infinita de tuplas cuyos primeros valores forman una sucesión infinita decreciente

$$a > a_1 > a_2 > a_3 > \dots$$

de valores enteros lo cual es imposible ya que en una sucesión tal, tarde o temprano tendrían que aparecer números negativos (La correspondiente sucesión de triángulos pitagóricos se muestra en la figura 1.2; es claro que no pueden ser valores negativos pues son medidas). Luego no existen tuplas pitagóricas de la forma (a,a,c). □

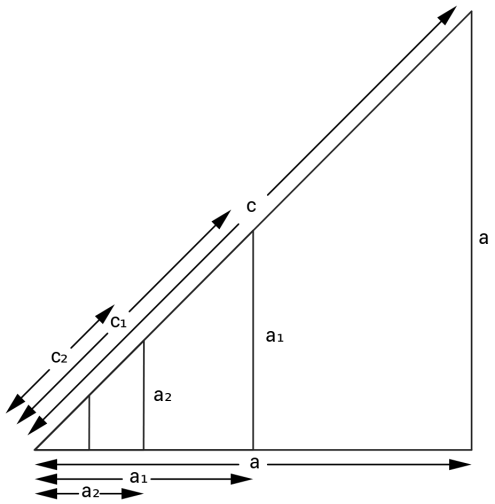

Figura 1.3: Una secuencia de triángulos pitagóricos.

De este resultado, se desprende fácilmente el hecho de que $\sqrt{2}$ no es un número racional, ya que si lo fuera se tendría que existen a,c números enteros tales que $\sqrt{2} = c/a$ y por tanto $c^2 = 2a^2$ lo cual es imposible. El descubrimiento de números irracionales fue un verdadro trauma para los Pitagoricos, quienes basaban toda su ciencia y filosofía en las propiedades de los números enteros y racionales.

Capítulo 2

El Teorema de los dos cuadrados

En este capítulo se demuestra el resultado más conocido como el Teorema de Fermat. Aunque en realidad fué Albert Girard en 1632 quien formula por primera vez que *"todo número primo de la forma* $4n+1$ *es suma de dos cuadrados"* (como ya se ha comentado en la introducción). Para ello se verán inicialmente algunas definiciones y resultados previos necesarios para la demostración.

Como se puede observar, se darán dos demostraciones distintas del mismo teorema. Suele ser de gran utilidad tener más de una demostración de un mismo resultado, no porque le añada validez (una sóla demostración correcta es suficiente), sino que puede servir para comprender mejor el resultado e incluso dar opción para desarrollarlo en diversas direcciones.

Como se comentará en su momento, a diferencia de muchos textos de Teoría de Números, las demostraciones en este trabajo son independientes de la ley de reciprocidad cuadrtica de Legendre-Gauss.

2.1. Aritmética modular

Se recordarán a continuación, algunos resultados básicos de la aritmética modular, la cual será de gran ayuda a la hora de trabajar con los problemas a abordar.

Dados enteros a,b,m con $m>0$, se dice que a es congruente con b módulo m si se cumple que $m|a-b$ y se escribe como

$$a \equiv b \pmod{m}.$$

Una expresión como la anterior se denomina congruencia. Se tiene que $a \equiv b \pmod{m}$ cuando la diferencia $a-b$ pertenece al conjunto de los múltiplos de m. Todavía cabe otra definición, basada en que el resto de la división de a por m es único: a es congruente con b módulo m si, y solo si, dan el mismo resto al dividirlos por m. Dicho de otra forma, $a \equiv b \pmod{m}$ quiere decir que a es de la forma $mk+b$ con $k \in \mathbb{Z}$. La relación de congruencia es una relación de equivalencia en el conjunto $\mathbb{Z}$ de los números enteros cuyas clases de equivalencia son conjuntos formados por todos los números que al dividirlos por m se obtiene el mismo resto. La clase de equivalencia de un entero a se denota como $[a]_m$ (o simplemente $[a]$ si sobreentendemos el módulo) y hay exactamente m clases de equivalencia $\{[0],[1],[2],...,[m-1]\}$; este conjunto suele denotarse como $\mathbb{Z}_m$. En ocasiones es más conveniente utilizar como representantes de las clases de equivalencia aquellos números con menor valor absoluto,

$$\begin{cases} \{0,\pm1,\pm2,...,\pm(m-1)/2\} & \text{si } m \text{ es impar} \\ \{0,\pm1,\pm2,...,\pm(m-2)/2\} & \text{si } m \text{ par} \end{cases} \tag{2.1}$$

de modo que para cualquier resto r, se tendrá $-m/2 \leq r \leq m/2$.

Además las sumas y productos de enteros congruentes también son congruentes, es decir si $a \equiv b \pmod{m}$ y $x \in \mathbb{Z}$ se tiene

$$a+x \equiv b+x \pmod{m}; \quad ax \equiv bx \pmod{m}; \quad -a \equiv -b \pmod{m},$$

lo que muestra que la suma y la multiplicación son operaciones bien definidas sobre el conjunto de las clases de equivalencia $[a]+[b]=[a+b]$ y $[a]\cdot[b]=[a\cdot b]$. Así el conjunto $\mathbb{Z}_m$ es un anillo conmutativo con unidad.

Si a y b son dos enteros cualesquiera y m es un entero positivo, la congruencia $ax \equiv b \pmod{m}$ tiene una solución en x si y sólo si b es divisible por el máximo común divisor $mcd(a,m)$ de a y m. Cuando éste es el caso, y x_0 es una solución, el conjunto de todas las soluciones viene dado por $\{x_0 + k\frac{m}{d} \mid k \in \mathbb{Z}\}$. En particular, existen exactamente $d = mcd(a,n)$ soluciones en el conjunto de residuos $\{0,1,2,...,n-1\}$.

Como consecuencia, si $m = p$ es primo, $\mathbb{Z}_p$ es un cuerpo, pues la congruencia $ax \equiv 1 \pmod{p}$ tiene una única solución si $[0] \neq [a] \in \mathbb{Z}_p$.

2.2. Enteros de Gauss

Es de sobra conocido el conjunto de los números enteros $\mathbb{Z}$ y el de los números complejos $\mathbb{C}$ ambos muy interesantes y muy útiles en gran cantidad de problemas relacionados con muchas ramas de las matemáticas.

Un entero de Gauss es exactamente un número complejo cuyas partes real e imaginarias son números enteros. Los enteros de Gauss fueron introducidos por Gauss en 1832 motivado por el estudio de las sumas de cuadrados. El conjunto de los enteros de Gauss se representa como $\mathbb{Z}[i]$ y por tanto es el siguiente:

$$\mathbb{Z}[i] = \{x+yi \mid x,y \in \mathbb{Z}\} \quad (\text{donde} \quad i^2 = -1)$$

Con las operaciones heredadas de $\mathbb{C}$ se tiene que la suma y la multiplicación de enteros de Gauss es un entero de Gauss:

$$z,w \in \mathbb{Z}[i] \text{ con } z = a_1 + ib_1 \text{ y } w = a_2 + ib_2,$$

$$z+w = (a_1+a_2)+i(b_1+b_2) \in \mathbb{Z}[i] \quad \text{y} \quad z\cdot w = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1) \in \mathbb{Z}[i]$$

En consecuencia $\mathbb{Z}[i]$ *es un dominio de integridad*, cuyo cuerpo de fracciones es la extensión simple $\mathbb{Q}(i)$ formada por los números complejos cuya parte real e imaginaria son números racionales.

Se puede ver una estrecha relación entre estos números y las sumas de cuadrados ya que si se toma $z = a+b\cdot i \in \mathbb{Z}[i]$, se tiene que $|z|^2 = z\cdot\bar{z} = a^2+b^2$. Se define la norma en el dominio $\mathbb{Z}[i]$ como:

$$\begin{array}{rcl} N : \mathbb{Z}[i] & \longrightarrow & \mathbb{N}\cup\{0\} \\ z & \longmapsto & N(z) = z\cdot\bar{z} \end{array}$$

Y se demuestra entonces que la norma así definida satisface las propiedades

- $N(z) = 0$ si y sólo si $z = 0$
- $N(zw) = N(z)N(w)$ para todos z,w en $\mathbb{Z}[i]$

La primera propiedad es trivial y para la segunda basta aplicar la definición

$$N(zw) = zw\overline{zw} = z\bar{z}w\overline{w} = N(z)N(w)$$

2.3. Primera demostración del teorema de Fermat

Se dan primero algunas definiciones que ayudarán a hacer más comprensible la demostración.

Definición 2.3.1. *Por cada entero $k \geq 1$, sea $S_k = \{n \mid n = x_1^2 + ... + x_k^2,\ x_1, ..., x_k \in \mathbb{Z}\}$, el conjunto de todas las sumas de k cuadrados.*

Ejemplo 2.3.1. *$S_1 = \{0, 1, 4, 9, ...\}$ es el conjunto de todos los cuadrados, se comprueba directamente que S_2, el conjunto de las sumas de dos cuadrados, contiene a 0, 1, 2, 4, 5 y 8 pero no a 3, 6 o 7.*

Lema 2.3.2. *El conjunto S_2, es multiplicativamente cerrado, es decir, si $s, t \in S_2$, entonces $st \in S_2$*

Demostración. Sea $s = a_1^2 + b_1^2$ y $t = a_2^2 + b_2^2$ elementos de S_2 con $a_1, b_1 . a_2, b_2 \in \mathbb{Z}$, luego aplicando que los enteros de Gauss $z = a_1 + i \cdot b_1$ y $w = a_2 + i \cdot b_2$ tienen como norma s y t respectivamente, aplicando ahora que $N(zw) = N(z)N(w)$ se tiene:

$$N(z)N(w) = (a_1^2 + b_1^2)(a_2^2 + b_2^2) = (a_1 a_2 - b_1 b_2)^2 + (a_1 b_2 + b_1 a_2)^2 = N(zw) \tag{2.2}$$

y por tanto $st \in S_2$ ya que $a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2 \in \mathbb{Z}$. □

Observación 2.3.1. *Aplicando inducción, se demuestra que el producto de cualquier numero finito de elementos de S_2 también está en S_2.*

Observación 2.3.2. *Reemplazando z por $\overline{z}$ en 2.2 obtenemos la igualdad*

$$(a_1^2 + b_1^2)(a_2^2 + b_2^2) = (a_1 a_2 + b_1 b_2)^2 + (a_1 b_2 - b_1 a_2)^2 \tag{2.3}$$

una identidad que se utiliza más adelante.

El Lema 2.3.2 sugiere que para determinar los elementos de S_2, se pueden estudiar primero los números primos ya que todo entero $n \geq 2$ es producto de números primos y si todos sus factores estan en S_2 entonces n también lo está. Notar que no todos los números primos son suma de dos cuadrados, por ejemplo el 3.

El resultado que aparece a continuación se prueba sin emplear la ley de reciprocidad cuadrática y será clave para demostrar el teorema de Fermat.

Lema 2.3.3. *Para números primos p de la forma $p = 4n + 1$ la ecuación $s^2 \equiv -1 \pmod{p}$ tiene dos soluciones $s \in \{1, 2, \ldots, p-1\}$, si $p = 2$ tiene solución única, mientras que para primos de la forma $4n + 3$ la ecuación no tiene solución.*

Demostración. Si $p = 2$ basta tomar $s = 1$. Sea p impar. Dado $x \in \{1, .., p-1\}$, se cosidera el conjunto $P_x := \{x, -x, x^{-1}, -x^{-1}\}$ con $x \in \{1, 2, ..., p-1\}$ siendo $-x$ la inversa aditiva y x^{-1} la inversa multiplicativa en $\mathbb{Z}_p$. Estos conjuntos tienen cuatro elementos salvo que alguno de ellos coincida, es decir, estos casos serán:

- $x \equiv -x$ es imposible por ser p impar.
- $x \equiv x^{-1}$, es decir $x^2 \equiv 1$, que tiene dos soluciones $x = 1$ y $x = p - 1$, con lo que se tiene que la única tupla cumpliendo esta relación es $P_1 = P_{p-1} = \{1, -1\} = \{p-1, 1-p\} = \{1, p-1\}$
- $x \equiv -x^{-1}$ por tanto, $x^2 \equiv -1$. Esta ecuación puede tener dos soluciones distintas o no tener ninguna. De tener dos soluciones, serían $x = x_0$ y $x = p - x_0$ por lo el conjunto tendría dos elementos $\{x_0, p - x_0\}$.

El conjunto $\{1,2,\ldots,p-1\}$ tiene $p-1$ elementos y se ha realizado una partición en cuadruplas (cuando P_x tiene 4 elementos) y una o dos tuplas (los casos en los que P_x tiene 2 elementos). Ahora, si $p-1=4m+2$ entonces sólo puede haber una tupla y lo demás son cuadruplas, por tanto $s^2 \equiv -1 \pmod{p}$ no tiene solución. Si $p-1=4m$ debe exisitir una segunda tupla y esta contiene las dos soluciones de $s^2 \equiv -1 \pmod{p}$ □

Teorema 2.3.4. *Todo número primo de la forma $4n+1$, con $n \in \mathbb{N}$, es suma de dos cuadrados.*

Demostración. Supongamos que $p \equiv 1 \pmod{4}$, por el teorema 2.3.3 se tiene que para algún u entero $u^2+1 = rp$ con $r \in \mathbb{Z}$. Se puede elegir un u tal que $0 \leq u \leq p-1$, con $0 \leq r \leq p$ y se tiene que $rp = u^2+1^2 \in S_2$. Sea m el menor entero tal que $mp \in S_2$ y $0 < m < p$. Si $m=1$ entonces $p \in S_2$ y el resultado est demostrado.

Sea $m > 1$. Como $mp \in S_2$, se tiene que $mp = a_1^2+b_1^2$ para algunos enteros a_1 y b_1. Sean $a_2, b_2 \in \mathbb{Z}$ los representantes de las clases de a_1 y b_1 con menor valor absoluto como en 2.1, de modo que $a_2 \equiv a_1$ y $b_2 \equiv b_1 \pmod{m}$ y $\mid a_1 \mid, \mid b_1 \mid < m/2$, se tiene $a_2^2+b_2^2 \equiv a_1^2+b_1^2 \equiv 0 \pmod{m}$; Así $a_2^2+b_2^2 = sm$ para algún $s \in \mathbb{Z}$. Como $\mid a_1 \mid, \mid b_1 \mid < m/2$, se tiene $a_1^2+b_1^2 \leq 2(m/2)^2 = m^2/2$ luego $s \leq m/2$ y por tanto $s < m$.

También se tiene $s > 0$, pues si $s=0$, entonces $a_2^2+b_2^2=0$, y esto implica que $a_1 \equiv b_1 \equiv 0 \pmod{m}$, esto es m divide a a_1 y b_1. Luego m^2 divide a $a_1^2+b_1^2 = mp$ y por tanto m divide a p. Esto no puede ser ya que p es primo y $1 < m < p$, luego $0 < s < m$.

Ahora $(a_1^2+b_1^2)(a_2^2+b_2^2) = mpsm = m^2sp$ y de la identidad 2.3

$$(a_1^2+b_1^2)(a_2^2+b_2^2) = (a_1a_2+b_1b_2)^2+(a_1b_2-b_1a_2)^2$$

se tiene

$$(a_1a_2+b_1b_2)^2+(a_1b_2-b_1a_2)^2 = m^2sp.$$

Como $a_1a_2+b_1b_2 = a_2^2+b_2^2 = 0 \pmod{m}$ y $a_1b_2-b_1a_2 = 0 \pmod{m}$ se puede dividir esta ecuación por m^2 quedando

$$\left(\frac{a_1a_2+b_1b_2^2}{m}\right)^2+\left(\frac{a_1b_2-b_1a_2}{m}\right)^2 = sp$$

donde ambos sumandos son enteros. Luego $sp \in S_2$ con $0 < s < m$ lo cual es una contradicción por la minimalidad de m. Por tanto $m=1$ □

Teorema 2.3.5. *Un número entero positivo n es suma de dos cuadrados si y sólo el exponente en la factorización de n de los primos $q \equiv 3 \pmod{4}$ es par.*

Demostración. ($\Leftarrow$) Supongamos que

$$n = 2^e p_1^{e_1}\cdots p_k^{e_k} q_1^{2f_1}\cdots q_l^{2f_l} = 2^e p_1^{e_1}\cdots p_k^{e_k}(q_1^2)^{f_1}\cdots(q_1^2)^{f_l}$$

para algunos primos $p_i \equiv 1 \pmod{4}$ y $q_j \equiv 3 \pmod{4}$, donde los exponentes son enteros $e \geq 0, e_i > 0$ y $f_j > 0$. Ahora bien $2 = 1^2+1^2 \in S_2$, por el teorema anterior se tiene que $p_i \in S_2$ y también que $q_j^2 = q_j^2+0^2 \in S_2$. Por tanto n es producto de elementos de S_2 y por el lema 2.3.2 y por tanto $n \in S_2$

($\Rightarrow$) Sea $n \in S_2$, y tomemos como hipotesis que $n = x^2+y^2$. Sea q cualquier número primo tal que $q \equiv 3 \pmod{4}$, sea q^f la mayor potencia de q que divide a n ($q^f \mid n$) y supongamos que f es impar. Sea d el máximo común divisor de x e y, es decir, $x = ad$ y $y = bd$ con $mcd(x,y) = 1$. Se tiene que $n = (a^2+b^2)d^2$ y por tanto $nd^{-2} = a^2+b^2$. Si $q^e \mid d$, entonces $q^{f-2e} \mid nd^{-2}$. Ahora bien $f-2e$ es impar luego distinto de cero, así que $q \mid nd^{-2} = a^2+b^2$ y por tanto $a^2 \equiv -b^2 \pmod{q}$. Pero b no puede ser divisible por q (ya que si así fuera, q dividiría a a y b contradiciendo m.c.d.$(x,y) = 1$), así que b es una unidad $\pmod{q}$. Si $c = b^{-1}$ módulo q, $bc \equiv 1 \pmod{q}$. Entonces multiplicando por c^2 se tiene $(ac)^2 \equiv -1 \pmod{q}$. Pero esto es imposible para un q primo tal que $q \equiv 3 \pmod{4}$ por el lema 2.3.3 luego f tiene que ser par. □

Ejemplo 2.3.2. *El entero* $60(=2^2\cdot 3\cdot 5)$ *no es suma de dos cuadrados ya que el exponente de 3 en la factorización es impar. Sin embargo,* $180(=2^2\cdot 3^2\cdot 5)$ *si que es suma de dos cuadrados. Para encontrarlos, primero se expresa 5 como suma de dos cuadrados:* $5=1^2+2^2$. *Al multiplicar por* $2^2\cdot 3^2$ *y se obtiene* $180=2^2\cdot 3\cdot 5=(2\cdot 3\cdot 2)^2+(2\cdot 3\cdot 1)^2=12^2+6^2$.

Ejemplo 2.3.3. *El entero* $221(=13\cdot 17)$ *es suma de dos cuadrados ya que* $13\equiv 17\equiv 1 \pmod 4$. *Para encontrar estos cuadrados, se utiliza la demostración del Lema 2.3.2 sustituyendo* $13=3^2+2^2$ *y* $17=4^2+1^2$*:*

$$221=st=(a_1a_2-b_1b_2)^2+(a_1b_2+b_1a_2)^2=(3\cdot 4-2\cdot 1)^2+(3\cdot 1+2\cdot 4)^2=10^2+11^2.$$

Notar que la ecuación 2.3 puede dar diferentes expresiones, en este caso, $221=14^2+5^2$. De manera similar, se puede expresar $6409(=221\cdot 29)$ repitiendo el proceso: $221=10^2+11^2$ y $29=5^2+2^2$, por tanto, $6409=(10\cdot 5-11\cdot 2)^2+(10\cdot 2+11\cdot 5)^2=28^2+75^2$.

Corolario 2.3.1. *Un número primo* p *es suma de dos cuadrados si y sólo si* $p=2$ *ó* $p\equiv 1 \pmod 4$.

2.4. Aritmética de los Enteros de Gauss

En esta sección se hace un estudio mas profundo de los enteros de Gauss. Ya se había visto que $\mathbb{Z}[i]$ es un dominio de integridad al que se le ha dotado de una norma. El siguiene paso razonable es ver si se puede dividir y factorizar de la misma forma que con los números enteros.

Se tiene que las unidades en $\mathbb{Z}[i]$ (elementos invertibles, es decir, los $\alpha\in\mathbb{Z}[i]$ tales que $\exists\beta\in\mathbb{Z}[i]$ y $\alpha\beta=\beta\alpha=1$) tienen norma 1, de hecho, un entero de Gauss tiene norma 1 si y solo si es una unidad. Se demuestra de forma sencilla que las unidades de $\mathbb{Z}[i]$ son $\{\pm 1,\pm i\}$. Dados $z,w\in\mathbb{Z}[i]$, se dice que z divide a w, o $z\mid w$ si $z=sw$ para algún $s\in\mathbb{Z}[i]$ y dado un $z\in\mathbb{Z}[i]$ se recuerda que z que es irreducible en $\mathbb{Z}[i]$ si z no es una unidad y sus únicos divisores son

$$\{\pm 1,\pm i,\pm z\pm iz\}$$

Dicho de otra forma, si sus únicos divisores son $z,1$ y sus asociados ($w=uz$ con u unidad).

No todos los números primos en $\mathbb{Z}$ son irreducibles en $\mathbb{Z}[i]$; por ejemplo 2 es primo en $\mathbb{Z}$ y sin embargo no lo es en $\mathbb{Z}[i]$ ya que $2=(1+i)(1-i)$.

Se tienen los siguientes resultados:

Lema 2.4.1. *Si* $z\mid w$ *en* $\mathbb{Z}[i]$, *entonces* $N(z)\mid N(w)$ *en* $\mathbb{Z}$

Demostración. Si $z\mid w$, entonces $w=sz$, tomando normas $N(w)=N(sz)=N(s)N(z)$ y se tiene que $N(z)\mid N(w)$. □

Este lema puede ser útil a la hora de determinar si un número es irreducible en $\mathbb{Z}[i]$, por ejemplo, $z=4+i$ es irreducible ya que $N(z)=4^2+1^2=17$ que es irreducible en $\mathbb{Z}$ y por tanto si $w\mid z$ se tendrá que $z=sw$, aplicando normas $N(z)=N(w)N(s)=17$ y como la norma es un entero positivo, las únicas opciones son $N(w)=1$, $N(s)=17$ o $N(s)=1$, $N(w)=17$. Por simetrá se puede tomar cualquiera de las dos opciones y se tiene que w es una unidad, luego $s=w^{-1}z$, es decir, los únicos divisores de z son las unidades y asociados de z. Por tanto $z=4+i$ es irreducible en $\mathbb{Z}[i]$.

Se demuestra a continuación que todo entero de Gauss se factoriza como producto de enteros de Gauss irreducibles.

Teorema 2.4.2. *Sea* $0\neq z\in\mathbb{Z}[i]$ *y* $N(z)\neq 1$, *entonces* $\exists p_1\dots p_k\in\mathbb{Z}[i]$ *irreducibles tales que* $z=p_1\cdots p_k$.

Demostración. La demostración es similar a la de los números enteros. Sea $z \in \mathbb{Z}[i]$, si z es irreducible, ya está si no $z = sw$ con $N(s), N(w) < N(z)$. Si se repite este proceso con cada factor, o w es irreducible, o $w = w_1 w_2$ con $N(w_1), N(w_2) < N(w)$ y así sucesivamente, como las normas son cada vez más pequeas y como poco tienen que valer 1 este proceso termina en algún punto y se tiene una factorización en factores primos de $\mathbb{Z}[i]$. □

Recordar que para probar la factorización única en $\mathbb{Z}$ se utilizaba que todo irreducible en $\mathbb{Z}$ verifica la condición de primo (el recíproco se verifica siempre), esto es que si p es irreducible y $p \mid ab \Rightarrow p \mid a$ o $p \mid b$ y este resultado se probaba con el algoritmo euclídeo. Para $\mathbb{Z}[i]$ se procede del mismo modo.

Lema 2.4.3. *Sean $z, w \in \mathbb{Z}[i]$, $w \neq 0$, entonces existe un cociente $s \in \mathbb{Z}[i]$ y un resto $t \in \mathbb{Z}[i]$ tales que*

$$z = sw + t \text{ con } N(t) < N(w)$$

Demostración. Sean $z, w \in \mathbb{Z}[i]$ y $w \neq 0$. Se considera su cociente en el cuerpo de fracciones de $\mathbb{Z}[i]$, $z/w = a' + b'i \in \mathbb{Q}(i)$ $(a', b' \in \mathbb{Q})$.

Sean $a, b \in \mathbb{Z}$ las mejores aproximaciones de a', b', es decir, $|a - a'| \leq 1/2$, $|b - b'| \leq 1/2$. Se define $s = a + bi \in \mathbb{Z}[i]$, teniendo así $z = sw + ((a' - a) + (b' - b)i)w$ y

$$N(((a' - a) + (b' - b)i)w) = ((a' - a)^2 + (b' - b)^2)N(w) \leq N(w)/2 < N(w)$$

($w \neq 0$ implica $N(w) \neq 0$)

□

Un dominio de integridad que posee una norma tal que se cumple lo anterior se conoce como *Dominio Eucldeo*.

Definición 2.4.4. *Sean $z, w, s \in \mathbb{Z}[i]$ tales que $s \mid z$ y $s \mid w$ entonces se dice que s es un divisor común de z y w. Se dirá que s es un máximo común divisor (mcd) de z y w si es undivisor común con máxima norma posible.*

Lema 2.4.5. *Si $z = sw + r$ donde $N(r) < N(w)$. Si $r = 0$, w es un mcd de z y w. Si no, un mcd de z y r es también un mcd de z y w (y viceversa).*

Demostración. Lo primero es claro ya que ningún divisor de w puede tener norma mayor que $N(w)$.

Sea $r \neq 0$. Cualquier divisor común de z y w tendrá que ser un divisor común de z y r, y viceversa. Esto implica cualquier *mcd* de z y w será *mcd* de z y r, y viceversa. □

Este lema muestra que el siguiente algoritmo para determinar el *mcd* de z y w en $\mathbb{Z}[i]$ funciona siempre. Al igual que en el caso de $\mathbb{Z}$, el algoritmo termina en un número finito de pasos.

Algoritmo Eucldeo en $\mathbb{Z}[i]$

1. Supongamos que $N(z) \geq N(w)$. Sea $(z_1, w_1) = (z, w)$
2. Se define $z_j = s_j w_j + r_j$ con $N(r_j) < N(w_j)$. Si $r_j = 0$, entonces w_j es un *mcd* para z y w. Si no continuar.
3. Sea $z_{j+1} = w_j$, $w_{j+1} = r_j$. Se tiene que $N(z_{j+1}) > N(w_{j+1})$ por el paso 2. Incrementar j en 1 y repetir paso 2.

Se puede recorrer el algoritmo anterior en orden inverso y obtener que el *mcd* de z y w tiene que ser de la forma $uz + vw$ para algún $u, v \in \mathbb{Z}[i]$ (Identidad de Bezout).

Lema 2.4.6. *Todo p irreducible en $\mathbb{Z}[i]$ es primo en $\mathbb{Z}[i]$, es decir que para todo p irreducible en $\mathbb{Z}[i]$, $p \mid zw$ entonces $p \mid z$ o $p \mid w$.*

Demostración. Se supone que $p \mid zw$ pero $p \nmid z$. Como p no divide a z y p es irreducible en $\mathbb{Z}[i]$, cualquier *mcd* de p y de z tendrá que ser una unidad. Por lo comentado anteriormente, esta unidad u es de la forma $u = sp + tz$ con $s,t \in \mathbb{Z}[i]$. Luego

$$uw = spw + tzw$$

y se tiene que p divide a ambos terminos en la derecha, $p \mid uw$ por tanto $uw = pk$ con $k \in \mathbb{Z}[i]$ como u es una unidad, $w = u^{-1}pk$, es decir $p \mid w$. □

Ahora es sencillo ver que si p y p' son irreducibles en $\mathbb{Z}[i]$, se tiene que si $p \mid p'$ implica que $p = up'$ donde u es una unidad en $\mathbb{Z}[i]$, es decir son asociados. Este resultado se utiliza en la demostración de la factorización única.

Teorema 2.4.7. *Sea $0 \neq z \in \mathbb{Z}[i]$ y sean $z = p_1 \dots p_m$ y $z = p'_1 \dots p'_n$ dos factorizaciones de z en irreducibles Gaussianos p_j y p'_j. Entonces $m = n$, y salvo reordenaciones de p'_j se tiene que $p_j = u_j p'_j$ para cada j, donde u_j es una unidad en $\mathbb{Z}[i]$*

Demostración. La demostración es similar a la de $\mathbb{N}$. Se supone que el teorema es falso. Cancelando cualquier factor común, se puede asumir

$$p_1 \dots p_m = p'_1 \dots p'_n$$

donde los p_k no son asociados a los p'_j, es decir $p_k \neq up'_j$ para algún k, j y u una unidad. Es claro que $p_1 \mid z = p'_1 \dots p'_n$. Dado que p_1 es primo, por el lema 2.4.6 $p_1 \mid p'_1$ o $p_1 \mid (p'_2 \dots p'_n)$. Ahora, $p_1 \mid p'_1$ es imposible ya que $p_1 \neq up'_1$ para cualquier unidad u. Por tanto, $p_1 \mid (p'_2 \dots p'_n)$. Al repetir este razonamiento eventualmente se obtendrá que $p_1 \mid p'_n$ lo cual es una contradicción. □

En consecuencia el dominio euclideo $\mathbb{Z}[i]$ es un *Dominio de Factorización única* (DFU)

2.5. Segunda demostración del teorema de Fermat

En esta demostración se pone de manifiesto la estrecha relación entre los números expresables como suma de cuadrados y los enteros de Gauss.

Teorema 2.5.1. *Un número primo p es suma de dos cuadrados si y sólo si $p = 2$ ó $p \equiv 1 \pmod 4$.*

Demostración. $\Rightarrow$) Es inmediato, basta observar que si $n \in \mathbb{Z}$, $n^2 \equiv 0, 1 \pmod 4$ según sea n par o no y que un número primo no puede ser suma de dos pares o dos impares.

$\Leftarrow$) Sea $p \neq 2$ un número primo tal que $p \equiv 1 \pmod 4$, por 2.3.3, existe un entero tal que $p \mid n^2 + 1 = (n+i)(n-i)$ y ocurre que $p \nmid (n+i)$ y $p \nmid (n-i)$ ya que en caso contrario se tendría que p divide a n. Luego p no es primo en $\mathbb{Z}[i]$. Como $\mathbb{Z}[i]$ es un DFU, existe $z \in \mathbb{Z}[i]$ primo tal que $z \mid p$, es decir $p = zw$ con z y w enteros de Gauss no unidades. Tomando normas $p^2 = N(p) = N(z)N(w)$ y como $N(z), N(w) \neq 1$ tiene que ser $N(z) = N(w) = p = a^2 + b^2$.

□

Corolario 2.5.1. *Si p un primo entero, entonces son equivalentes:*

1. *p es expresable como suma de dos cuadrados.*
2. *$\exists z \in \mathbb{Z}[i]$ cuya norma es p.*
3. *p no es irreducible en $\mathbb{Z}[i]$.*

Demostración. 1) $\Rightarrow$ 2) Sea $p = a^2 + b^2$. Basta tomar $z = a + bi$.

2) $\Rightarrow$ 3) Sea z tal que $N(z) = p = z\bar{z} = a^2 + b^2$. Como $z, \bar{z}$ no son unidades, pues tendrían norma 1, se tiene que p se descompone en producto de dos elementos de $\mathbb{Z}[i]$, luego no es irreducible.

3) $\Rightarrow$ 1) Si p no es irreducible en $\mathbb{Z}[i]$, existen $z = a + bi, w = c + di \in \mathbb{Z}[i]$ no unidades tales que $p = zw$. Tomando normas, se obtiene $p^2 = N(p) = N(z)N(w)$, como z y w no son unidades $N(z)$, $N(w) \neq 1$ luego sólo puede ser $N(z) = N(w) = p = a^2 + b^2$.

□

Corolario 2.5.2. *Un número primo $p \in \mathbb{N}$ es irreducible en $\mathbb{Z}[i]$ si y sólo si $p \equiv 3 \pmod 4$.*

Lema 2.5.2. *Un entero de Gauss es irreducible si y sólo si es de una de las dos formas siguientes:*

1. *Es asociado de un número primo $p > 0$ entero con $p \equiv 3 \pmod 4$, es decir $p, -p, ip, -ip$.*
2. *Tiene norma prima.*

Demostración. Si p es como en 1 es claro que es primo por el resultado anterior.

Sea $z \in \mathbb{Z}[i]$ tal que su norma sea un primo de $\mathbb{Z}$, $p = N(z)$.

Sea $z = ab$ cualquier factorización de z, tomando normas $p = N(a)N(b)$. Es una ecuación de enteros positivos, y p es primo en $\mathbb{N}$ y por lo tanto o $N(a)$ o $N(b)$ es 1, pero entonces uno de los dos sería una unidad y se tiene que z es primo.

Queda probar que si $z = a + bi \in \mathbb{Z}[i]$ es irreducible, entonces es de una de las formas enunciadas. Si $a = 0$ o $b = 0$, entonces z es asociado con p.

Si $a = 0$ y $b = 0$, entonces $N(z) = a^2 + b^2 = (a + bi)(a - bi)$ es una descomposición en factores irreducibles. Por ser $\mathbb{Z}[i]$ un DFU, $N(z)$ debe ser un número primo, porque en otro caso una descomposición en $\mathbb{Z}$ sería distinta de la anterior. □

Teorema 2.5.3. *Un número natural n es suma de dos cuadrados si y sólo si cualquier primo tal que $p \mid n$ y $p \equiv 3 \pmod 4$ aparece como potencia par en la factorización de n.*

Demostración. Sea $n = m^2 q$, de modo que los factores primos de q son 2 o congruentes con 1 módulo 4. Por el teorema 2.5.1 y dado que S_2 es cerrado con la multiplicación, se tiene que n es suma de dos cuadrados. Recíprocamente, supongamos que $n = a^2 + b^2$. La descomposición en $\mathbb{Z}[i]$ en factores irreducibles de $a + bi$ será, por el lema 2.5.2, $a + bi = up_1 \cdots p_r(c_1 + d_1) \cdots (c_s + id_s)$, donde u es una unidad, $p_1, \ldots, p_r \in \mathbb{Z}$ son números primos, tales que $p_i \equiv 3 \pmod 4$, para $i = 1, \ldots, r$, y $c_j + id_j \in \mathbb{Z}[i]$ son elementos de norma $c^2 + d^2$, para $j = 1, \ldots, s$, que por el lema 2.5.2, no puede ser congruente con 3 módulo 4. Conjugando la expresión anterior, se obtiene una descomposición de $a - bi$, y multiplicándolas, queda:

$$n = (a + bi)(a - bi) = p_1^2 \cdots p_1^2 (c_1^2 + d_1^2) \cdots (c_s^2 + d_s^2),$$

que verifica el enunciado. □

Capítulo 3

Teorema de los cuatro cuadrados

Como ya se ha comentado, es más sencillo estudiar las sumas de cuatro cuadrados que las de tres. Antes de nada, se presenta a los cuaternios de Hamilton, que al igual que ocurre con los enteros de Gauss para las sumas de dos cuadrados, es una herramienta útil e interesante para las demostraciones de esta sección. Los cuaternios son una extensión de los números complejos. En general, a las extensiones de los números complejos se las conoce como números hipercomplejos.

3.1. Cuaternios de Hamilton

El descubrimiento de los cuaternios en 1843 fué clasificado como un importante hito en la historia del algebra abstracta, puesto que permitió calcular algebraicamente las rotaciones de los cuerpos sólidos.

En 1833 Sir William Rowan Hamilton logró dar estructura algebráica a las parejas de números reales, haciendo corresponder la estructura de $\mathbb{R}^2$ con la de $\mathbb{C}$. Esto fue un triunfo del álgebra pues se empezaron a vislumbrar otro tipo de estructuras que no eran números en el sentido usual del término, pero estos nuevos objetos obedecían a ciertas reglas de operación, de manera similar a las reglas de operación acostumbradas que se conocían para los números.

Dado el éxito que Hamilton tuvo al darle una estructura algebraica a $\mathbb{R}^2$, los cuaternios surgieron de los intentos de Hamilton por generalizar las operaciones (aritmética) de los números complejos de una manera que fuese aplicable en $\mathbb{R}^3$. Hamilton llevaba años trabajando con tres términos - uno por cada dimensión del espacio - pero podía sólo hacerlos rotar en un plano. No fue hasta el año de 1843, a la edad de 38 años, que Hamilton, en un chispazo de inspiración, inventó en un instante un sistema de 3 partes 'imaginarias' que se convertiría en el álgebra de los cuaternios. Según una historia relatada por el propio Hamilton, la solución al problema que le ocupaba le sobrevino un día que estaba paseando con su esposa, bajo la forma de la ecuación:

$$i^2 = j^2 = k^2 = ijk = -1$$

Inmediatamente, grabó esta expresión en el lateral del puente de Brougham, que estaba muy cerca del lugar.

Cuando el introdujo el cuarto término, encontró las rotaciones tridimensionales que venía buscando, pero tuvo problemas al conceptualizar el significado de este término extra. Como casi todos los victorianos, supuso que este término debería significar algo, así que en el prefacio de sus Conferencias sobre Cuaterniones de 1853 adicionó en una nota de pie de página: *'Parecía (y aun me parece) normal conectar esta unidad espacial extra con la concepción del tiempo.'*

Con respecto a estos números, Melanie Bayley hace una interpretación de Alicia en el País de las Maravillas en la que observa un paralelismo entre los cuaternios de Hamilton y la fiesta de té del Sombrerero Loco.

Lewis Carroll fue el seudónimo del reverendo Charles Lutwidge Dodgson, un matemático de la Universidad de Oxford. Se asume razonablemente, que muchos de los pasajes del libro tuvieron una inspiración matemática, aunque Dodgson no hizo ningún comentario al respecto. Por lo que se sabe Dodgson tenía una opinión muy tradicionalista de las matemáticas, basadas en el enfoque axiomático de los Elementos de Euclides. Bayley lo describe como 'aferrado conservador en matemáticas', quien se sentía frustrado al ver que las matemáticas decaían según él en sus estándares de rigor.

La situación en el libro es la siguiente: Alicia, comparte la mesa con tres extrañas figuras: el Sombrerero, la Liebre de Marzo y el Lirón. La figura Tiempo, que ha discutido con el Sombrerero, está ahora ausente y en un arranque de locura no permite al Sombrerero mover los relojes pasadas las seis.

Según interpreta Bayley, los participantes en la fiesta de té, representan tres términos de un cuaternio, en el cual el cuarto término, el tiempo, no aparece. Sin el Tiempo, los personajes quedan atascados en la mesa de té, moviéndose constantemente alrededor, buscando platos y vasos limpios. Su movimiento alrededor de la mesa podría ser la rememoración de los intentos iniciales de Hamilton de calcular el movimiento, el cual se limitaba a rotaciones en el plano antes de adicionar el tiempo. Aun cuando Alicia se une a la fiesta, ella no puede parar al Sombrerero, ni a la Liebre, ni al Lirón que giran erráticamente en torno a la mesa, porque ella no es una unidad espacial extra como lo es el Tiempo.

Por otra parte, las respuestas de Alicia a las adivinanzas del Sombrerero son no conmutativas, al igual que ocurre con los cuaternios. La escena es la siguiente:

-Entonces debes decir lo que piensas -siguió la Liebre de Marzo.

-Ya lo hago -se apresuró a replicar Alicia-. O al menos... al menos pienso lo que digo... Viene a ser lo mismo, ¿no?

-¿Lo mismo? ¡De ninguna manera! -dijo el Sombrerero-. ¡En tal caso, sería lo mismo decir «veo lo que como» que «como lo que veo»!

-¡Y sería lo mismo decir –añadió la Liebre de Marzo- «me gusta lo que tengo» que «tengo lo que me gusta»!

-¡Y será lo mismo decir –añadió el Lirón, que parecía hablar en medio de sus sueños- «respiro cuando duermo» que «duermo cuando respiro»!

Cuando la escena termina, la Liebre y el Sombrerero tratan de poner al Lirón en la jarra de té. Ésta podría ser la ruta hacia la libertad. Si pudieran deshacerse del Lirón, podrían existir independientemente, como un número complejo con dos términos. Aun absurdo, según Dodgson, pero ya libres de una rotación permanente alrededor de la mesa.

Históricamente un cuaternio está descrito por una cuaterna de números reales $q = (a,b,c,d)$, y el conjunto $\mathbb{H}$ de todos ellos es exactamente el espacio vectorial real $\mathbb{R}^4$. Llamando $\{1,i,j,k\}$ a la base canónica de $\mathbb{H}$, se puede escribir $\mathbb{H} = \{a+bi+cj+dk \mid a,b,c,d \in \mathbb{R}\}$. Este espacio vectorial admite una multiplicación distributiva y asociativa (que se denota por yuxtaposición, i.e. qq') inducida por las reglas $i^2 = j^2 = k^2 = ijk = -1$ que no es conmutativa pero hace que $\mathbb{H}$ sea un álgebra de división, es decir todo elemento de $\mathbb{H}$ no nulo tiene inverso. Al igual que en la construcción de $\mathbb{C}$ a partir de $\mathbb{R}$, siendo $\mathbb{C}$ isomorfo a un $\mathbb{R}$-espacio vectorial de dimensión 2, se puede construir $\mathbb{H}$ a partir de $\mathbb{C}$, siendo $\mathbb{H}$ isomorfo a un $\mathbb{C}$-espacio vectorial de dimensión 2 ($q = a+bi+cj+dk = (a+bi)+(c+di)j$). Esta interpretación se vuelve a obtener con la siguiente definición.

Definición 3.1.1. *El álgebra $\mathbb{H}$ de cuaternios es la subálgebra del álgebra de las matrices complejas 2×2 formada por las matrices de la forma*

$$q = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$$

donde α y β son números complejos. Un elemento de $\mathbb{H}$ se denomina cuaternio, cuaternión o número cuaternio.

Dados $\alpha = a+bi$ y $\beta = c+di$, donde $a,b,c,d \in \mathbb{R}$, se toman como $1,i,j,k$ a las siguientes matrices

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, j = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Resulta obvio que cualquier elemento $q = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix}$ de $\mathbb{H}$ puede expresarse como combinación lineal de las anteriores $q = a \cdot 1 + b \cdot i + c \cdot j + d \cdot k$

Definición 3.1.2. *La norma de un cuaternio $||q||$ se define como la raíz cuadrada de su determinante, por tanto $||q||^2$ es*

$$det \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} = \alpha\overline{\alpha} + \beta\overline{\beta} = |\alpha| + |\beta|$$

$$q = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix} = a^2 + b^2 + c^2 + d^2$$

Los cuaternios $1,i,j,k$ tienen norma 1 y satisfacen

$$i^2 = j^2 = k^2 = -1$$

$$ij = -ij = k$$

$$jk = -kj = i$$

$$ki = -ik = j$$

Los cuaternios son un grupo abeliano con la suma. Los cuaternios no nulos son grupo con la multiplicación, pero el producto de cuaterniones generalmente no es conmutativo ($q_1q_2 \neq q_2q_1$) aunque si que son asociativos y distributivos sobre la suma.

Debido a la propiedad multiplicativa de los determinantes $det(q_1q_2) = det(q_1)det(q_2)$ se tiene que la norma cumple

$$||q_1q_2|| = ||q_1|| \cdot ||q_2||$$

Se define el conjugado de un cuaternio $q = a \cdot 1 + b \cdot i + c \cdot j + d \cdot k$ como $\overline{q} = a \cdot 1 - b \cdot i - c \cdot j - d \cdot k$. Se tienen las siguientes propiedades:

$$q\overline{q} = ||q||^2$$

$$\overline{q_1 + q_2} = \overline{q_1} + \overline{q_2}$$

$$\overline{q_1q_2} = \overline{q_2q_1}$$

(Debido a que el producto no es conmutativo).

Es interesante remarcar el hecho de que los únicos n tales que $\mathbb{R}^n$ tiene multiplicación distributiva sobre la suma y norma multiplicativa son $n = 1,2,4,8$. Para $n = 1$ son los números reales $\mathbb{R}$, $n = 2$ corresponde con los complejos $\mathbb{C}$, con $n = 4$ son los cuaternios $\mathbb{H}$ y para $n = 8$ son los octoniones $\mathbb{O}$.

3.2. Teorema de los cuatro cuadrados

Como se ve a continuación existe una cierta analogía entre la demostración de este teorema y el teorema de Fermat. Se han ordenado los razonamientos de forma que este parecido sea más evidente.

Lema 3.2.1. *El conjunto S_4, es multiplicativamente cerrado, es decir, si $s,t \in S_4$, entonces $st \in S_4$*

Demostración. Sea $x_1 = a_1^2+b_1^2+c_1^2+d_1^2 = ||1\cdot a_1 + i\cdot b_1 + j\cdot c_1 + k\cdot d_1||^2 = ||q_1||^2$ y $x_2 = a_2^2+b_2^2+c_2^2+d_2^2 = ||1\cdot a_2 + i\cdot b_2 + j\cdot c_2 + k\cdot d_2||^2 = ||q_2||^2$. Por la propiedad multiplicativa de la norma, se tiene que

$$x_1x_2 = ||q_1||^2||q_2||^2 = ||q_1q_2||^2$$

Por tanto se tiene:

$$\begin{aligned}(a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2) &= (a_1a_2-b_1b_2-c_1c_2-d_1d_2)^2 \\ &+ (a_1b_2+b_1a_2+c_1d_2-d_1c_2)^2 \\ &+ (a_1c_2-b_1d_2+c_1a_2+d_1b_2)^2 \\ &+ (a_1d_2+b_1c_2-c_1b_2+d_1a_2)^2 \end{aligned} \tag{3.1}$$

□

Observación 3.2.1. *Aplicando inducción, se demuestra que el producto de cualquier numero finito de elementos de S_4 también está en S_4.*

Observación 3.2.2. *Si en 3.1 en vez de $q = a\cdot 1 + b\cdot i + c\cdot j + d\cdot k$ se toma su conjugado se obtiene la siguiente identidad*

$$\begin{aligned}(a_1^2+b_1^2+c_1^2+d_1^2)(a_2^2+b_2^2+c_2^2+d_2^2) &= (a_1a_2+b_1b_2+c_1c_2+d_1d_2)^2 \\ &+ (a_1b_2-b_1a_2-c_1d_2+d_1c_2)^2 \\ &+ (a_1c_2+b_1d_2-c_1a_2-d_1b_2)^2 \\ &+ (a_1d_2-b_1c_2+c_1b_2-d_1a_2)^2 \end{aligned} \tag{3.2}$$

Esta última identidad se utiliza más adelante.

Lema 3.2.2. *Sea p un número primo, entonces, existen x,y enteros tales que p divide a x^2+y^2+1 (la ecuación $x^2+y^2 \equiv -1 \pmod{p}$ tiene solución). Además se pueden tomar x,y con $x^2+y^2+1 = mp$ tales que $0 \le m < p$.*

Demostración. Si $p = 2$ es trivial. Sea p un primo impar. Sea P el conjunto de todos los cuadrados módulo p. Notar que si se toman clases de equivalencia como en 2.1, es decir, con menor valor absoluto, los cuadrados módulo p serán $\{0,1,\ldots,(p-1)/2\}$ y por tanto hay $(p+1)/2$ elementos en P. Sea P' el conjunto de los números de la forma $-1-x$ con $x \in P$. P' también tiene $(p+1)/2$ elementos. Como hay p elementos distintos módulo p, se tiene que P y P' tienen algún elemento en común, lo que implica que $x^2+y^2 \equiv -1 \pmod{p}$ tiene solución. Además, para esta pareja concreta se tiene:

$$0 < 1+x^2+y^2 \le 1+((p-1)/2)^2+((p-1)/2)^2 \le p^2/4+p^2/4 < p^2$$

Y por tanto $0 < mp < p^2$. □

De otra forma

Demostración. Sea $A = \{1+x^2 \mid x \in 0,1,\ldots,(p-1)/2\}$. No hay dos elementos de A congruentes módulo p. En efecto si $1+x^2 \equiv 1+x_1^2 \pmod{p}$, se tiene que p divide a $(x-x_1)(x+x_1)$. Se tiene por tanto que $p \mid (x-x_1)$ o $p \mid (x+x_1)$ es decir, como $x, x_1 \in \{0,1,\ldots,(p-1)/2\}$, se tiene que $x+x_1 < p$ y solo puede ser $x = x_1$. Luego, los restos al dividir cada uno de los $(p+1)/2$ elementos de P por p son distintos.

De forma similar tomando $B = \{-y^2 \mid y \in 0,1,\ldots,(p-1)/2\}$ también dan restos diferentes al dividirlos por p. Estos conjuntos son disjuntos pues los elementos de B son negativos y los de A son mayores que 1.

Se tiene que la unión de ambos conjuntos tiene $p+1$ elementos. Como sólo hay p posibles restos distintos, alguno de los restos es común en ambos conjuntos, es decir $x^2+1 \equiv -y^2 \pmod{p}$. Además, para esta pareja concreta se tiene:

$$0 < 1+x^2+y^2 \leq 1+((p-1)/2)^2+((p-1)/2)^2 \leq p^2/4+p^2/4 < p^2$$

Y por tanto $0 < mp < p^2$. □

Lema 3.2.3. *Si p es un primo impar y m un entero positivo par tal que mp es suma de cuatro cuadrados, entonces $(m/2)p$ es también suma de cuatro cuadrados.*

Demostración. Sean m y p como en el enunciado y sea $mp = a^2+b^2+c^2+d^2$ donde a,b,c,d son enteros. Como m es par, se da uno de los siguientes casos:

1. a,b,c,d son todos pares.
2. a,b,c,d son todos impares.
3. Dos de ellos son pares y otros dos impares.

En cualquiera de los casos se tiene que $a+b$, $a-b$, $c+d$ y $c-d$ son pares. Por tanto:

$$\frac{mp}{2} = \left(\frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2 + \left(\frac{c+d}{2}\right)^2 + \left(\frac{c-d}{2}\right)^2$$

Y se tiene que $(m/2)p \in S_4$ □

Teorema 3.2.4. *Todo entero no negativo es suma de cuatro cuadrados.*

Demostración. Es claro que $0,1,2 \in S_4$, luego por el Lema 3.2.1, es suficiente probar que todo primo impar p está en S_4. Sea p primo entero impar, por el lema 3.2.2, para algún $0 < m < p$ se tiene

$$mp = x_1^2+x_2^2+x_3^2+x_4^2 \tag{3.3}$$

Se observa que si un número es suma de tres cuadrados también es suma de cuatro cuadrados ($S_3 \subseteq S_4$). Sea m_0 el menor m que cumple (3,3). Por el lema 3.2.1, se puede suponer m impar. Si $m_0 = 1$ se tiene el el resultado. Supongamos $1 < m_0 < p$. No todos los x_j son divisibles por p, ya que si lo fueran se tendría $m_0 p = p^2(y_1^2+y_2^2+y_3^2+y_4^2)$, pero $m_0 < p$. Como $m_0 \geq 3$, se puede suponer por el algoritmo de la división

$$x_j = m_0 b_j + y_j, \quad j = 1,2,3,4.$$

Donde $|y_i| < m_0/2$ y $y_1^2+y_2^2+y_3^2+y_4^2 > 0$. Además $x_j \equiv y_j \pmod{m_0}$. Por tanto

$$0 < y_1^2+y_2^2+y_3^2+y_4^2 < 4(m_0/2)^2 = m_0^2$$

luego

$$y_1^2+y_2^2+y_3^2+y_4^2 \equiv x_1^2+x_2^2+x_3^2+x_4^2 \equiv 0 \pmod{m_0}$$

Luego se tiene

$$y_1^2+y_2^2+y_3^2+y_4^2 = m_0 m_1 \tag{3.4}$$

con $0 < m_1 < m_0$. Multiplicando 3,3 y 3,4 y aplicando la identidad 3.2, se tiene

$$\begin{array}{rcl}(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2) & = & (x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2 \\ & + & (x_1y_2-x_2y_1-x_3y_4+x_4y_3)^2 \\ & + & (x_1y_3+x_2y_4-x_3y_1-x_4y_2)^2 \\ & + & (x_1y_4-x_2y_3+x_3y_2-x_4y_1)^2 \\ & = & z_1^2+z_2^2+z_3^2+z_4^2 = m_0^2 m_1 p\end{array}$$

Como $x_j \equiv y_j \pmod{m_0}$, se tiene

$$z_1 = x_1y_1+x_2y_2+x3_1y_3+x_4y_4 \equiv x_1^2+x_2^2+x_3^2+x_4^2 \equiv 0 \pmod{m_0}$$

$$z_2 = x_1y_2-x_2y_1-x_3y_4+y_4x_3 \equiv x_1y_2-y_2x_1-x_3y_4+y_4x_3 \equiv 0 \pmod{m_0}$$

y así con cada uno de los z_i. Luego se puede dividir en ambos lados por m_0^2 quedando

$$m_1 p = (\frac{z_1}{m_0})^2+(\frac{z_2}{m_0})^2+(\frac{z_3}{m_0})^2+(\frac{z_4}{m_0})^2$$

Es decir, $m_1 p \in S_4$ con $0 < m_1 < m_0$, lo que contradice la minimalidad de m_0 y se tiene $m_0 = 1$ lo que finaliza la demostración. □

La demostración de este teorema se basa en el método del descenso infinito al igual que la primera demostración del teorema de Fermat en la sección anterior. En esencia, ambas demostraciones son similares, basta recordar a grandes rasgos los pasos empleados en la del teorema de Fermat

1. S_2 es multiplicativamente cerrado.
2. Si p es un primo de la forma $4n+1$ p divide a $1+x^2$ para algún x, pudiendo elegir este x de forma que $mp = x^2+1$ con $0 < m < p$
3. Se toma el menor m para el que $x^2+y^2 = mp$, con p primo y $1 < m < p$, entonces encontrar $x_1^2+y_2^2 = m_1 p$, con x_1, y_1 enteros y $0 < m_1 < m$ contradiciendo la minimalidad de m.

Al igual que en el teorema de los dos cuadrados existe una demosración alternativa utilizando los enteros de Gauss, se puede dar una damostración de este teorema empleando los cuaternios de Hamilton. Esta demostración no ha sido incluida dada la limitada extensión para este trabajo pero se da una idea esquemática.

Para tal demostración, se definen los cuaternios de Hurwitz (también llamados enteros de Hurwitz) como $\mathbb{H}_1 := \{a+bi+cj+dk \mid a,b,c,d \in \mathbb{Z} \text{ o } a,b,c,d \in \mathbb{Z}+1/2\}$ que es un subanillo de los cuaternios de Hamilton. Se tiene que todo irreducible en $\mathbb{H}_1$ es primo en $\mathbb{H}_1$ y la demostración sigue el siguiente esquema:

1. S_4 es multiplicativamente cerrado.
2. Si p es un entero primo, p divide a $1+x^2+y^2$ para algún x, y.
3. Dado p, se toman x, y tales que $p|1+x^2+y^2 = (1+xi+yj)(1-xi-yj)$ y se prueba que $p \nmid (1+xi+yj)$ y $p \nmid (1-xi-yj)$ luego p no es primo en $\mathbb{H}_1$.
4. Si p no es primo en $\mathbb{H}_1$ se tiene que $p \in S_4$.

Una demostración completa se puede ver en *Quaternions and the four square theorem* de Jia Hong Ray Ng [Ji08].

Capítulo 4

Representaciones superiores

Los teoremas sobre las sumas de dos y cuatro cuadrados, originaron diversas generalizaciones sobre estos. Algunas de estas generalizaciones surgieron antes de las demostraciones del teorema de Fermat y el de Lagarnge.

Una de las preguntas que sugieren estos teoremas es cómo representar un número dado como suma de cuadrados y de cuantas formas puede hacerse. En 1834 Jacobi encontró el número exacto de formas en las que se puede expresar un número entero positivo n como suma de cuatro cuadrados. Este número es 8 veces la suma de los divisores de n si n es impar y 24 veces la suma de los divisores impares de n si n es par.

Una de las generalizaciones es el teorema de los números poligonales de Fermat. Recordar que un número poligonal es aquel que puede ser representado como puntos dispuestos en forma de polígono regular, empezando por el 1. El teorema de los números poligonales, dice que todo número entero positivo puede ser expresando como suma de, como mucho, 3 números triangulares, 4 números cuadrados, 5 números pentagonales, etc. Gauss demostró el resultado para números triangulares en 1796 y anotó en su diario *'EΥPHKA! num* $= \triangle + \triangle + \triangle$*'* conmemorando tal ocasión. El teorema de los números poligonales fué finalmente probado por Cauchy en 1813. Nathanson en 1987 da una demostración más corta basada en un lema probado por Cauchy que dice:

> Para números naturales impares a y b tales que $b^2 < 4a$ y $3a < b^2 + 2b + 4$ se pueden encontrar enteros no negativos s,t,u y v tales que $a = s^2 + t^2 + u^2 + v^2$ y $b = s + t + u + v$.

Otra generalización es tomar combinaciones de cuadrados, es decir, buscar números expresables como $x^2 + ay^2$ donde a es un número dado y $x, y \in \mathbb{Z}$. La ley de reciprocidad cuadrática de Gauss, responde parcialmente a este problema y conduce en último término, a desarrollar toda una nueva disciplina conocida como Teoría de cuerpos de clases.

Del mismo modo, con la suma de cuatro cuadrados, se puede plantear para que a,b,c,d números naturales, la ecuación $n = ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2$, tiene solución para todo n siendo x_1, x_2, x_3, x_4 enteros.

El teorema de los cuatro cuadrados es el caso $a = b = c = d = 1$. La solución general de este problema fue dada por Ramanujan:

Asumiendo que $a \leq b \leq c \leq d$ entonces hay exactamente 54 elecciones de a,b,c,d tal que la ecuación anterior puede resolverse en enteros x_1, x_2, x_3, x_4 para todo n (en realidad Ramanujan dio una elección más, $a = 1, b = 2, c = d = 5$, pero en este caso la ecuación no tiene solución para $n = 15$).

En los teoremas demostrados en este trabajo, se deja la posibilidad, de que en las representaciones de números como suma de cuadrados, alguno de los sumandos sea nulo. Encontrar descomposiciones en sumas de cuadrados tales que ningún sumando se anule, es un problema interesante pero escapa a los contenidos de este trabajo. Puede encontrarse un análisis de este problema en *Representations of Integers as Sums of Squares* de Emil Grosswald [Gros85].

Otra forma de generalizar estos teoremas consiste en buscar representaciones de números como potencias superiores, es decir, buscar números que se puedan expresar como sumas de potencias k-ésimas. Este problema es conocido como el problema de Waring.

4.1. El problema de Waring

'Omnis integer numerus vel est cubus, vel e duobus,
tribus, 4, 5, 6, 7, 8, vel novem cubis compositus,
est etiam quadrato-quadratus vel e duobus, tribus &c.
usque ad novemdecim compositus & sic deinceps.'

Meditationes Algebraicæ 1770, pp. 204-5(Edward Waring)

Así enunció Waring, la conjetura que afirma que todo entero es expresable como suma de 9 cubos, 19 potencias cuartas y en general un número finito de potencias k-ésimas. Es decir, todo entero positivo puede expresarse como suma de a lo sumo s potencias k-ésimas positivas, siendo s dependiente de k (se entiende que k es un número entero no negativo).

$$n = x_1^k + \ldots + x_s^k$$

Se denota por $g(k)$ al mínimo s que verifica la condición anterior. El teorema de Lagrange prueba esta conjetura para el caso $k = 2$. No fue hasta 1909 cuando David Hilbert, resolvió el problema de la existencia de $g(k) < \infty$, para todo k. La demostración se basa en la siguiente identidad:

Lema 4.1.1. *Dados $k \geq 1$, $n \geq 1$, sea $N = \binom{2k+n-1}{2k}$. Existen N números racionales $\lambda_1, \ldots, \lambda_N$ y enteros $a_{1i}, \ldots, a_{ni}$, $i = 1, \ldots, N$ tales que*

$$(x_1^2 + \cdots + x_n^2)^k = \sum_{i=1}^{N} \lambda_i (a_{1i}x_1 + \cdots + a_{ni}x_n)^{2k}$$

para todo x_i entero.

Aunque la conjetura quedaba demostrada, Hilbert no determinó la forma de calcular el valor numérico de $g(k)$ para cualquier k. Se conocen algunos valores de estos números. Como ya se ha comentado el caso $g(2) = 4$, es el teorema de Lagrange. Otros valores conocidos son

- $g(3) = 9$, demostrado por Wieferich y Kempner entre 1909 y 1912.
- $g(4) = 19$, demostrado por Balasubramanian, Dress y Deshouillers en 1986.
- $g(5) = 37$, demostrado por Chen Jingrun en 1964.
- $g(6) = 73$, demostrado por Pillai en 1940.

Como se puede ver, lo que conjeturó Waring en su día, resultó ser cierto. Es curioso el hecho de que $g(6)$ se encontrara antes que $g(5)$, y éste a su vez antes que $g(4)$. Sin embargo, aunque hacen falta 9 cubos para expresar cualquier entero, únicamente son necesarios 9 cubos para dos números, el 23 y el 239, para el resto es suficiente con 8 sumandos. Por este curioso hecho, se generaliza algo más el problema de Waring. Se denota como $G(k)$ al menor número de k-ésimas potencias, necesario para representar cualquier entero, salvo un número finito de ellos. Es inmediato que $G(k) \leq g(k)$.

Aunque no se conoce ninguna fórmula para calcular estos números, si se pueden establecer cotas

Teorema 4.1.2. *Dado un entero positivo k se tiene*

$$g(k) \geq \lfloor (3/2)^k \rfloor + 2^k - 2$$

donde $\lfloor x \rfloor$ es el mayor entero menor o igual que x.

Demostración. Sea $3^k = q2^k + r$, donde $1 \leq r < 2^k$ y $q = \lfloor (3/2)^k \rfloor$, como el entero $n = 2^k q - 1 < 3^k$, solo puede ser representado por potencias k-éimas de 1 y 2, y el menor número de éstas es $\lfloor (3/2)^k \rfloor - 1$ potencias k-éimas de 2 y $2^k - 1$ potencias k-éimas de 1 esto es, $n = (q-1)2^k + (2^k - 1)1^k$, así que se requieren $2^k + q - 2$ potencias k-éimas. □

Este resultado fue probado por Johann Albrecht Euler, hijo de Leonhard Euler.

Si se toman los primeros valores conocidos de $g(k)$ se observa que coincide con su cota superior. Por este motivo se cree que la cota da el valor exacto. De hecho se sabe que es así exceptuando un número finito de casos.

Teorema 4.1.3. *Si $k > 6$ y se cumple*

$$3^k - 2^k + 2 < (2^k - 1)\lfloor (3/2)^k \rfloor \tag{4.1}$$

se tiene $g(k) = \lfloor (3/2)^k \rfloor + 2^k - 2$.

Además, si no se cumple 4.1, definiendo $N(k) = \lfloor (3/2)^k \rfloor \cdot \lfloor (4/3)^k \rfloor + \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor$ se tiene

$$\begin{cases} g(k) = \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + 2^k - 3 & \text{si } N(k) > 2^k \\ g(k) = \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + 2^k - 2 & \text{si } N(k) = 2^k \end{cases}$$

En 1957 Mahler demostró, que si existen valores de k para los cuales 4.1 es falsa, entonces sólo pueden ser un número finito. Posteriormente Stemmler (1964) verificó, en ordenador, que 4.1 se cumple para $k \leq 200000$. En 1990, Kubina y Wunderlich ya lo habían extendido a $k \leq 471600000$.

Igualmente, también existe una acotación para los $G(k)$. Se tiene que si $k \geq 2$, entonces se verifica la desigualdad $G(k) \geq k+1$. Resultado probado por Maillet en 1908. Esto significa que existen números naturales arbitrariamente grandes que no son suma de k-éimas potencias. Así que para todo $k \geq 2$ se verifican las desigualdades $k+1 \leq G(k) \leq g(k)$.

Para el caso de sumas de cuadrados se tiene por el teorema de Lagrange que $g(2) = 4$ y por tanto $G(2) \leq 4$. Ahora bien, como se ha comentado, los números de la forma $n = 4^e(8k+7)$ no pueden expresarse como suma de tres cuadrados. Por tanto se tiene $G(k) = 4$

Para los cubos (caso $k = 3$), no se conoce el valor exacto de $G(3)$. En 1949, Yu V. Linnik, demostró que $G(3) \leq 7$. Aplicando el resultado de Maillet, se tiene que $4 \leq G(3) \leq 7$, aunque los cálculos numéricos, parecen indicar que $G(3) = 4$.

Igualmente se puede extender el problema de Waring permitiendo sumar y restar potencias k-ésimas. Es decir un número n puede ser expresado como suma o diferencia de s potencias k-ésimas si

$$n = \pm x_1^k \pm \cdots \pm x_s^k$$

De forma similar a lo anterior, se denota por $w(k)$ al mínimo s que verifica lo anterior para cualquier n. Igualmente, se define como $W(k)$ al menor número de k-ésimas potencias, necesario para representar cualquier entero, salvo un número finito de ellos.

El problema de determinar $w(k)$ y $W(k)$ se suele conocer como el problema 'fácil' de Waring, aunque en realidad es más dificil.

Un desarrollo más amplio del problema de Waring se puede ver en *Waring's Problem: A Survey*, de Vaughan y Wooley [VW02]

4.2. Números Taxicab

El problema de los números taxicab, profundiza en el estudio de la suma de cubos. Da comienzo con una anécdota, bastante conocida, entre los matemáticos Hardy y Ramanujan.

Srinivasa Ramanujan, nació en 1887 en Erode, India, en el seno de una familia de brahmanes pobre y ortodoxa. Ramanujan fue autodidacta de las matemáticas. Adquirió prácticamente la totalidad de su conocimiento sobre éstas, a través de los libros La Trigonometría plana de S. Looney, y Synopsis of Elementary Results in Pure Mathematics de S. Carr que contenían un listado de unos 6000 teoremas sin demostración.

El 16 de enero de 1913, el matemático británico Godfrey Harold Hardy, recibe una carta de Ramanujan, un joven indio de 25 años. En esta carta, Ramanujan, explicaba su situación, admitiendo

que carecía de una formación universitaria en matemáticas, habiendo seguido una trayectoria propia. En esta carta se adjuntaban 120 fórmulas, alguna de las cuales desbordaban al propio Hardy, el cual comentó: 'Forzoso es que sean verdaderas, porque si no lo fueran, nadie habría tenido la suficiente imaginación para inventarlas'.

Hardy le invitó a trasladarse a Cambridge, a lo que Ramanujan en un principio se mostró reticente. Por fin, tras la intervención de su madre y de la diosa Namagiri, de la que Ramanujan afirmaba que le dictaba sus resultados en sueños, y de una beca de 250 libras, Ramanujan llega al Trinity College en la primavera de 1913. Durante su estancia en Cambridge, en plena guerra mundial, Ramanujan enfermó de tuberculosis, por lo que acabó ingresado en diversas ocasiones. En 1919 tras el fin de la contienda, y gravemente enfermo, decide regresar a la India. Morirá a los pocos meses. A pesar de ello, su trabajo con Hardy ha dejado una increíble producción de resultados matemáticos sorprendentes, en forma de 'Cuadernos'. Algunos de ellos todavía están siendo estudiados.

Una de las veces que Ramanujan fue ingresado, Hardy fue a visitarlo, y le llamó la atención el número del taxi, 1729, pues lo encontraba muy poco interesante. Debió de estar pensando en ello porque entró en la habitación del hospital en donde estaba Ramanujan tumbado en la cama y, con un 'hola' seco, expresó su desilusión acerca de este número. Era, según él, un número aburrido, agregando que esperaba que no fuese un mal presagio. 'No, Hardy, dijo Ramanujan, es un número muy interesante. Es el número más pequeño expresable como la suma de dos cubos positivos de dos formas diferentes.' Hardy, a continuación, le preguntó si conocía la respuesta para las cuartas potencias. Ramanujan contestó, tras pensarlo un momento, que no podía ver la respuesta, pero que pensaba que debía ser un número extremadamente grande. De hecho, la respuesta, obtenida mediante cálculos con ordenador, es $635318657 = 134^4 + 133^4 = 158^4 + 59^4$.

De esta anécdota surgen los llamados números taxicab. Se define el enésimo número taxicab ($Ta(n)$) como el menor entero que se puede expresar como suma de dos cubos de n formas distintas. Para $n = 1$ es sencillo comprobar que $Ta(1) = 2 = 1^3 + 1^3$. Otros valores son

- $Ta(2) = 1729 = 1^3 + 12^3 = 9^3 + 10^3$ conocido como el número Hardy-Ramanujan, fue publicado por primera vez por Bernard Frènicle de Bessy en 1657.

- $Ta(3) = 87539319 = 167^3 + 436^3 = 228^3 + 423^3 = 255^3 + 414^3$ obtenido por John Leech en 1967 mediante cálculos con supercomputadoras.

Se conocen 6 números taxicab, siendo $Ta(6) = 24153319581254312065344$ hallado por Uwe Hollerbach en 2008. Para el resto, Christian Boyer en 2006 encontró cotas superiores para $Ta(7), \ldots, Ta(12)$.

Incluyendo alguna restricción más, se encuentran los llamados números 'cubefree taxicab' $T(n)$ siendo éste, el menor número que se puede descomponer como n sumas distintas de la forma $T(n) = x^3 + y^3$ siendo x e y coprimos. Se tiene que $Ta(1) = T(1)$ y $Ta(2) = T(2)$. Otros valores conocidos

- $T(3) = 15170835645$ hallado por Paul Vojta en 1981.

- $T(4) = 1801049058342701083$ hallado por Stuart Gascoigne e independientemente por Duncan Moore en 2003.

Se puede generalizar los números taxicab permitiendo mayor número de sumandos y potencias mayores. Así, se denota como $Taxicab(k, j, n)$ al menor entero positivo que puede expresarse como la suma de j potencias positivas de k de n formas diferentes. Los números taxicab son el caso $k = 3$ y $j = 2$. Por ahora no se conoce nigún $Taxicab(5, 2, n)$ para $n \geq 2$.

Así mismo, también pueden generalizarse los números taxicab, permitiendo sumas y restas de potencias cúbicas. A éstos se los conoce como números cabtaxi, siendo $Cabtaxi(n)$ el menor entero positivo que puede descomponerse como suma o resta de dos potencias cúbicas de n formas. Para estos números se conocen hasta $n = 10$. $Cabtaxi(5)$, $Cabtaxi(6)$ y $Cabtaxi(7)$ fueron hallados por Randall L. Rathbun; $Cabtaxi(8)$ hallado por Daniel J. Bernstein; $Cabtaxi(9)$ hallado por Duncan

Moore. Para $Cabtaxi(10)$ Christian Boyer encontró una cota superior en 2006 y ésta fue verifcada como Cabtaxi(10) por Uwe Hollerbach en mayo de 2008.

Como se observa existen infinidad de maneras de abordar y generalizar este problema. Las generalizaciones proporcionan una visión más amplia de resultados anteriores e incluso en ocasiones permiten demostrar resultados con mayores restricciones.

Bibliografía

[AZ09] M. Aigner y G.M. Ziegler. *Proofs from THE BOOK*, Springer-Verlag, Berlin, New York, 2009.

[Bus18] W. H. Bussey. *Fermat's Method of Infinite Descent*, The American Mathematical Monthly, Vol. 25, No. 8, 1918.

[Cal11] C. Calderón-García. *Sobre el problema clásico de Waring*, Revista de la Academia de Ciencias Exactas, Físicas, Químicas y Naturales de Zaragoza, 2011.

[Dic19] L.E. Dickson. *History of the theory of numbers*, Washington, Carnegie Institution of Washington, 1919.

[Gros85] E. Grosswald. *Representations of Integers as Sums of Squares*, Springer-Verlag New York Inc., 1985.

[Gauss] Gaussianos. *http://gaussianos.com/*

[Ji08] Jia Hong Ray Ng. *Quaternions and the four square theorem*, Pappers from the University of Chicago, http://www.math.uchicago.edu/~may/VIGRE/VIGRE2008/REUPapers/Ng.pdf, 2008.

[JJ98] G.A. Jones y J.M. Jones. *Elementary Number Theory*,Springer undergraduate mathematics series, 1998.

[Kos09] K. Kostadinov. *Introduction to Number Theory notes*,Boston University, Department of Mathematics and Statistics, 2009.

[Mar09] K. Martin. *Introduction to Number Theory*,University of Oklahoma, Department of Mathematics, 2009.

[Rod] J.J. Rodríguez-Padilla. *El Algebra y la geometría de los cuaternios y algunas de sus aplicaciones*, http://www.bibliotecadigital.uson.mx/, Tesis digitales, Tesis 21070.

[Rib72] P. Ribenboim. *Algebraic numbers*, John Wiley & Sons, Inc. 1972.

[Ros95] H.E.Rose. *A Course in Number Theory*,Oxford Science Publications, 1995.

[Tat05] Leung Tat-Wing. *The Method of Infinite Descent*, Mathematical Excalibur, Vol. 10, No. 4, 2005.

[VW02] R.C. Vaughan y T.D. Wooley. *Waring's Problem: A Survey*, University of Bristol, 2002.

[Wiki] Wikipedia. *http://www.wikipedia.org/*