QUATERNIONS AND THE FOUR SQUARE THEOREM

ALAN LEE

1. Introduction

While we did not focus a lot on noncommutative rings in class, there are many applications to other mathematical subjects in terms of noncommutative rings.

Definition 1.1. A *noncommutative ring* R is a ring that contains at least one pair of elements a, b such that $a \times b \neq b \times a$.

In this paper, we will focus specifically on a noncommutative ring known as quaternions. Using special quaternions called Hurwitz integers, we will prove Lagrange's Four Square Theorem.

2. Background

Trying to decompose every natural number into a minimal amount of perfect squares has always been the subject of much discussion. One of the earliest conjectures is attributed to Fermat and was first proven by Euler in 1747.

Theorem 2.1 (Fermat's Two Square Theorem)**.** *An odd prime number p can be expressed as a sum of two squares if and only if*

$$p \equiv 1 \pmod 4.$$

There is a construction showing that if a and b can be expressed as a sum of two squares, ab can as well. This construction involves complex numbers of the form $a + bi$ and utilizes the behavior of complex number multiplication.

However, this two square method is only limited to numbers that have no prime factors $p \equiv 3 \pmod 4$ of odd power. Lagrange was able to prove a theorem of his own in 1770, which we will focus on in this paper. Analogously, the construction for the product of two numbers that can be expressed as a sum of four sqaures involves quaternions, and will be detailed in a later section.

3. Quaternions and Four Squares

Definition 3.1. The *quaternions* are a number system defined by the fundamental units i, j, and k. All elements are of the form $a + bi + cj + dk$ and multiplication is defined as

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

For the purpose of quaternions, we have a slightly modified definition of conjugates.

Date: July 12, 2020.

Definition 3.2. The *conjugate* of a quaternion $q = a + bi + cj + dk$, denoted as $\bar{q}$, is equal to $a - bi - cj - dk$.

Definition 3.3. We may also represent each quaternion as a matrix, with $1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $j = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$, and $k = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Then each quaternion $a + bi + cj + dk$ can be expressed as $\begin{pmatrix} a+di & b+ci \\ b-ci & a-di \end{pmatrix}$. Setting $\alpha = a + di$ and $\beta = b + ci$ to simplify things, we have $a + bi + cj + dk = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}$. We refer to this matrix as the *representative matrix* of the given quaternion.

Quaternions have a strong presence in number theory, where they may be used to prove theorems as the one below.

Theorem 3.4 (Lagrange's Four Square Theorem)**.** *Every natural number n can be expressed as*

$$n = x_0^2 + x_1^2 + x_2^2 + x_3^2$$

for some $x_0, x_1, x_2, x_3 \in \mathbb{Z}$.

To prove this theorem, we need to define norms and the Hurwitz integers.

Definition 3.5. The *norm* of a quaternion q, denoted as $||q||$, is the square root of the determinant of its representative matrix.

Solving for the square of the norm of an arbitrary quaternion $q = a + bi + cj + dk$, we have

$$||q||^2 = (a + di)(a - di) + (b + ci)(b - ci) = a^2 + b^2 + c^2 + d^2.$$

Norms have a nice property; for any two quaternions q_1 and q_2, we have $||q_1||^2 \cdot ||q_2||^2 = ||q_1 q_2||^2$.

Definition 3.6. *Hurwitz integers* are special quaternions that are elements of the ring $\mathbb{Z}[h, i, j, k]$, where $h = \frac{1+i+j+k}{2}$. We can also think of the Hurwitz integers as the quaternions with only integer coefficients as well as the quaternions with only half-integer coefficients.

Example. $\frac{1+3i+5j+7k}{2}$ is a Hurwitz integer, but $\frac{1+2i+4j+8k}{2}$ is not because not all of the coefficients in the numerator have the same parity.

Note that for every Hurwitz integer h, $||h||^2 \in \mathbb{Z}$.

Proof of Theorem 3.4. The following proof is due to [Ng08]. Our proof will have two parts, first showing that every prime can be expressed as a sum of four squares, then proving that the product of any two numbers that can be expressed as a sum of four squares can also be expressed as a sum of four squares.

The cases $1 = 0^2 + 0^2 + 0^2 + 1^2$ and $2 = 0^2 + 0^2 + 1^2 + 1^2$ are trivial, so let us only focus on proving that all odd primes can be expressed as the sum of four squares.

Lemma 3.7. *All odd primes $p = 2n + 1$ divide $1 + l^2 + m^2$ for some $l, m \in \mathbb{Z}$.*

Proof. It is easy to prove that the numbers $0^2, 1^2, 2^2, \ldots, n^2$ each have different residues modulo p using elementary techniques. Thus, there are exactly $n + 1$ unique quadratic residues modulo p (including 0).

Using similar reasoning, we can deduce that for each $m \in \{0, 1, 2, \ldots, n\}$, $-1 - m^2$ has a unique residue modulo p, not necessarily distinct from any of the residues mentioned above.

Now that we are considering $2n + 2$ residues modulo p: $n + 1$ from each scenario above. However, there are only $2n + 1$ unique residues modulo p since $p = 2n + 1$. Thus, by the Pigeonhole Principle, at least one quadratic residue of the form l^2 must be congruent to a residue of the form $-1 - m^2$ modulo p. Thus, we have

$$\begin{aligned} l^2 &\equiv -1 - m^2 \text{ (mod p)} \\ 1 + l^2 + m^2 &\equiv 0 \text{ (mod p)}. \end{aligned}$$

□

We also have a theorem from [Ng08] that we will not prove in the paper, but can be proved using the Euclidean algorithm.

Theorem 3.8. *If a prime p is irreducible in $\mathbb{Z}[h, i, j, k]$, for any $a, b \in \alpha, \beta \in \mathbb{Z}[h, i, j, k]$ such that $p|\alpha\beta$, we have $p|\alpha$ or $p|\beta$.*

Now consider any odd prime p. We know by the previous lemma that $p|1 + l^2 + m^2$ for some $l, m \in \mathbb{Z}$. Let us factor $1 + l^2 + m^2$ further into $(1 + li + mj)(1 - li - mj)$, which indicates that $1 + l^2 + m^2$ is a product of two Hurwitz integers. Neither of these factors, when divided by p are Hurwitz integers since the $\frac{1}{p}$ term is not a half-integer or an integer.

However, p divides neither $1 + li + mj$ nor $1 - li - mj$ yet divides their product. Using the contrapositive of Theorem 2.8, we can then deduce that p is reducible into a product of two Hurwitz integers.

Therefore, we are able to factor p into $(a+bi+cj+dk)\gamma$ such that both factors are Hurwitz integers and neither of them have norm equal to 1. We also have $p = \bar{p} = (a-bi-cj-dk)\bar{\gamma}$. The product of these two equations yields

$$\begin{aligned} p^2 &= (a + bi + cj + dk)(a - bi - cj - dk)\gamma\bar{\gamma} \\ &= \left(a^2 + b^2 + c^2 + d^2\right)\gamma\bar{\gamma} \\ &= \left(a^2 + b^2 + c^2 + d^2\right)||\gamma||^2. \end{aligned}$$

Then since neither of the new factors are equal to 1, and the only other way to factor p^2 is $p \cdot p$, we must have $p = a^2 + b^2 + c^2 + d^2$ for some integers a, b, c, d or half-integers a, b, c, d. In the case where a, b, c, d are all integers, we are done.

Considering when a, b, c, d are all half-integers, we can do some substitution. In this case, there exists a Hurwitz integer $\omega = \frac{\pm 1 \pm i \pm j \pm k}{2}$ and even a', b', c', d' such that $\omega + a' + b'i + c'j + d'k = a + bi + cj + dk$.

Example. If $a = b = \frac{3}{2}, c = -\frac{1}{2}$ and $d = \frac{13}{2}$, setting $\omega = \frac{1+i+j-k}{2}$ yields

$$\frac{-1-i-j+k}{2} + a' + b'i + c'j + d'k = \frac{3+3i-j+13k}{2}.$$

We can verify that $a' = b' = 2$, $c' = 0$ and $d' = 6$, which are all even.

Since the norm of ω is 1, we have $\omega\bar{\omega} = 1$ as well. We now have

$$\begin{aligned} p &= a^2 + b^2 + c^2 + d^2 \\ &= (a + bi + cj + dk)(a - bi - cj - dk) \\ &= (\omega + a' + b'i + c'j + d'k)\,(\bar{\omega} + a' - b'i - c'j - d'k) \\ &= ((\omega + a' + b'i + c'j + d'k)\,\bar{\omega}) \cdot (\omega\,(\bar{\omega} + a' - b'i - c'j - d'k)) \end{aligned}$$

Multiplying out the first factor, we see that $a'\bar{\omega}$ has integer coefficients since a' is even. This is the same for b', c' and d'. Then we have $\omega\bar{\omega} = 1$. Thus we see that the first factor can be expressed as $w + xi + yj + zk$ for some $w, x, y, z \in \mathbb{Z}$. Since the other factor above is simply the conjugate $w - xi - yj - zk$, we have proven that $p = w^2 + x^2 + y^2 + z^2$ where $w, x, y, z \in \mathbb{Z}$.

Example. Take the prime $p = 31 = \frac{81+25+9+9}{4} = (\frac{9}{2})^2 + (\frac{5}{2})^2 + (\frac{3}{2})^2 + (\frac{3}{2})^2$. We have $a = \frac{9}{2}$, $b = \frac{5}{2}$ and $c = d = \frac{3}{2}$. We can then set up the equation

$$\omega + a' + b'i + c'j + d'k = \frac{9+5i+3j+3k}{2}.$$

In order for $a, b, c, d \equiv 0 \pmod 2$, we must have $\omega = \frac{1+i-j-k}{2}$. Hence we have

$$\frac{1+i-j-k}{2} + a' + b'i + c'j + d'k = \frac{9+5i+3j+3k}{2}$$

$$a' + b'i + c'j + d'k = 4 + 2i + 2j + 2k.$$

Using the equation found earlier, we can now express p as

$$\begin{aligned} 31 &= ((\omega + a' + b'i + c'j + d'k)\,\bar{\omega}) \cdot (\omega\,(\bar{\omega} + a' - b'i - c'j - d'k)) \\ &= \left(\frac{1+i-j-k}{2} + 4 + 2i + 2j + 2k\right) \cdot \frac{1-i+j+k}{2} \\ &\cdot \frac{1+i-j-k}{2} \cdot \left(\frac{1-i+j+k}{2} + 4 - 2i - 2j - 2k\right) \\ &= (1 + (2+i+j+k)\,(1-i+j+k)) \cdot (1 + (2-i-j-k)\,(1+i-j-k)) \\ &= (2 - i + j + 5k) \cdot (2 + i - j - 5k) \\ &= 2^2 + 1^2 + 1^2 + 5^2. \end{aligned}$$

Now that we have shown every prime can be expressed as a sum of four squares, we prove the second part. Take two integers x_1 and x_2 such that $x_1 = a_1^2 + b_1^2 + c_1^2 + d_1^2 = ||a_1 + b_1 i + c_1 j + d_1 k||^2$ and $x_2 = a_2^2 + b_2^2 + c_2^2 + d_2^2 = ||a_2 + b_2 i + c_2 j + d_2 k||^2$.

Since norms are multiplicative, we have

$$\begin{aligned} x_1 x_2 &= ||(a_1 + b_1 i + c_1 j + d_1 k)(a_2 + b_2 i + c_2 j + d_2 k)||^2 \\ &= (a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2)^2 + (a_1 b_2 + a_2 b_1 + c_1 d_2 - c_2 d_1)^2 \\ &+ (a_1 c_2 + a_2 c_1 - b_1 d_2 + b_2 d_1)^2 + (a_1 d_2 + a_2 d_1 + b_1 c_2 - b_2 c_1)^2. \end{aligned}$$

Since all of the a_i, b_i, c_i, and d_i were integers, the product $x_1 x_2$ is also a sum of four squares. Hence, all positive integers can be expressed as the sum of four integer squares.

■

4. Example and Conclusion

We have seen how a specific noncommutative ring called quaternions allowed us to derive a proof of the Lagrange Four Square Theorem. Now let us try out an example to see how the four squares are constructed for an arbitrary number.

Example. Take $31 = 3^2+3^2+3^2+2^2$ and $57 = 7^2+2^2+2^2+0^2$. Their product, $31*57 = 1767$, can be expressed as

$$\begin{aligned} 1767 &= 31 \cdot 57 \\ &= ||(a_1 + b_1 i + c_1 j + d_1 k)(a_2 + b_2 i + c_2 j + d_2 k)||^2 \\ &= (a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2)^2 + (a_1 b_2 + a_2 b_1 + c_1 d_2 - c_2 d_1)^2 \\ &+ (a_1 c_2 + a_2 c_1 - b_1 d_2 + b_2 d_1)^2 + (a_1 d_2 + a_2 d_1 + b_1 c_2 - b_2 c_1)^2 \\ &= ((3)(7) - (3)(2) - (3)(2) - (2)(0))^2 + ((3)(2) + (7)(3) + (3)(0) - (2)(2))^2 \\ &+ ((3)(2) + (7)(3) - (3)(0) + (2)(2))^2 + ((3)(0) + (7)(2) + (3)(2) - (2)(3))^2 \\ &= (21 - 6 - 6)^2 + (6 + 21 - 4)^2 + (6 + 21 + 4)^2 + (14 + 6 - 6)^2 \\ &= 9^2 + 23^2 + 31^2 + 14^2. \end{aligned}$$

Fermat also went on to write about a generalized version of the four square theorem. However, it was Cauchy who was able to prove it fully in 1813. See [Nat87] for a quick proof of the theorem.

Theorem 4.1 (Fermat Polygonal Number Theorem)**.** *Every positive integer n is the sum of at most n n-gonal numbers.*

Example. Take 58 as we try to break it up into smaller parts.

(1) Triangular: $58 = 55 + 3$
(2) Square: $58 = 49 + 9$
(3) Pentagonal: $58 = 51 + 5 + 1 + 1$
(4) Hexagonal: $58 = 45 + 6 + 6 + 1$

Returning to the broader focus of noncommutative rings in general, while they are often overlooked due to a lack of "good" properties, they are useful in their own right. There are many classes of noncommutative rings such as division, semisimple, semiprimitive, and simple rings, as well as specific types of rings such as the quaternions, matrix rings, and group rings. Quite a few theorems take an interest in these aforementioned classes, including the Artin-Wedderburn Theorem and Wedderburn's Little Theorem. There is still a lot to explore regarding noncommutative rings since these types of rings have less restrictions (not having commutativity as a property). Many results regarding these rings are fairly recent, having been conjectured or proven in the past few decades.

References

[Nat87] Melvyn B Nathanson. A short proof of cauchy's polygonal number theorem. *Proceedings of the American Mathematical Society*, pages 22–24, 1987.

[Ng08] Jia Hong Ray Ng. Quaterions and the four square theorem. 2008.