REPRESENTING INTEGERS AS SUMS OF SQUARES

MICHAEL WONG

Abstract. We study in detail the special case of Waring's problem when the power $k = 2$. Ultimately, we prove that four is the least number of squares needed to represent any integer. To this end, we prove that some numbers cannot be represented as sums of two squares, some cannot be represented as sums of three, and all can be represented as sums of four. We also show that numbers of a certain form can be represented as sums of two squares. Though we mostly use classic methods of number theory, we venture into group theory to prove a few preliminary theorems.

Contents

1. Introduction

The Diophantine equation $N = \sum_{i=1}^{s} x_i^k$ is of special interest in mathematics. Elegant in its simplicity, this equation can be applied to advanced topics in mathematics and physics. Beginning math students, or just someone with a curious mind, will find it an accessible introduction to number theory.

The equation itself has a rich history apart from the general theory of Diophantine equations. Such celebrated mathematicians as Euler, Lagrange, Legendre, and Gauss have contributed to our knowledge of it. Among the problems they tackled was what integers can be represented by the sum. Providing a concise statment of this problem, Edward Waring conjectured that every integer is the sum of a fixed number s of k^{th} powers. The function $g(k)$, depending only on k, is defined as the least number of k^{th} powers needed for all integers. Hilbert first proved Waring's conjecture, and mathematicians are still calculating the values of $g(k)$ for larger k. The first few values have been known for some time: $g(2) = 4$, $g(3) = 9$, and $g(4) = 19$.

In the following exposition, we prove that $g(2) = 4$. Obviously, we must show that some numbers cannot be represented as sums of two squares; some numbers cannot be represented as sums of three squares, and all numbers *can* be represented

Date: August 21, 2009.

as sums of four squares. Along the way, we also prove that numbers satisyfing certain conditions can be represented as sums of two squares. Throughout, we assume the reader has a basic understanding of congruences, quadratic residues, fields, and groups.

2. Quadratic Character of -1

For the two- and four-square theorems, we must know whether -1 is a quadratic residue modulo the prime p. Our approach to this point involves a few basic theorems in field theory and group theory.

First, we establish the relation of congruence to finite fields. Let $\mathbb{Z}_p$ be the set of all congruence classes modulo p. We denote the congruence class corresponding to an integer n as $[n]$. Clearly, the set consists only of p elements. Now, we define addition and multiplication on these elements in the usual manner:

$$[m] + [n] = [m + n]$$

$$[m] \cdot [n] = [m \cdot n]$$

Then $\{\mathbb{Z}_p, +, \cdot\}$ is a field, as is easily verified. For convenience, we represent $[0]$ by 0 and every other congruence class by its least postive residue, allowing us to write $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$.

Let $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$. From the fact that $\{\mathbb{Z}_p, +, \cdot\}$ is a field, we can immediately deduce that $\{\mathbb{Z}_p^*, \cdot\}$ is an abelian group. We call this group the *multiplicative group* of $\mathbb{Z}_p$. Let us examine its structure in a general context. For any x in a group G, we define the *order* of x as the least positive integer n such that $x^n = 1$ and denote it as $ord(x) = n$. It is easily shown that the elements $x^0, x^1, \ldots, x^{n-1}$ are distinct and, combined with the operation $\{\cdot\}$, make an abelian subgroup of G, denoted as G'. We say that G' is the *cyclic* subgroup *generated* by x. Thus, we arrive at an alternative definition of the order of x: the order of G'.

We recall here the definition of 'quadratic residue' in the special case that the modulus is prime:

Definition 2.1. The element a of $\mathbb{Z}_p^*$ is a *quadratic residue* (mod p) if the polynomial $\lambda^2 - a$ has a zero over the field $\mathbb{Z}_p$. Otherwise, a is a *quadratic non-residue*.

To determine the quadratic character of -1, we need to prove that the the multiplicative group of $\mathbb{Z}_p$ is cyclic. In the trivial case $p = 2$, -1 is a quadratic residue. Now, assuming p is odd and $\{\mathbb{Z}_p^*, \cdot\}$ is cyclic, let x be a generator of the group. So $x^{p-1} = 1$, implying that $x^{\frac{1}{2}(p-1)} = p - 1 \equiv -1 \pmod{p}$. Clearly, then, -1 is a quadratic residue if and only if $P = \frac{1}{2}(p-1)$ is even. P is even for all primes congruent to 1 and odd for those congruent to 3 (mod 4). Therefore, we have the following proposition:

Proposition 2.2. *Suppose p is an odd prime. -1 is a quadratic residue (mod p) if and only if p is of the form $4k + 1$.*

We proceed to the theorems necessary to prove this statement.

Lemma 2.3. *Suppose x, y are elements of the abelian group $\{G, \cdot\}$. Furthermore, suppose $ord(x) = a$, $ord(y) = b$, and $(a, b) = 1$. Then $ord(xy) = ab$.*

Proof. Let $ord(xy) = j$. j must divide ab, for $(xy)^{ab} = x^{ab}y^{ab} = 1$. By contradiction, assume $j < ab$. $j \mid ab$ and $(a, b) = 1$ implies that $j \mid a$ or $j \mid b$. Without loss

of generality, assume the latter. $(xy)^j = x^j y^j = 1$, implying that $y^j = (x^j)^{-1}$. Raising both sides to the ath power, we have $y^{aj} = (x^{aj})^{-1} = 1$. Because $j \mid b$, the order of y^j is b/j. So a must be a multiple of b/j; that is, for some integer m, we have $a = m(b/j)$, contradicting our assumption that $(a, b) = 1$. A similar argument follows if we assume $j \mid a$. Therefore, $j = ab$. □

This lemma leads to its own generalization. In the next lemma, we remove the condition that $(a, b) = 1$.

Lemma 2.4. *Suppose x, y are elements of the abelian group $\{G, \cdot\}$. Furthermore, suppose $ord(x) = a$ and $ord(y) = b$. Then there exists $z \in G$ such that $ord(z) = lcm(a, b)$.*

Proof. For the moment, assume that there exist m and n such that $m \mid a$ and $n \mid b$, $(m, n) = 1$, and $mn = lcm(a, b)$. We will prove afterwards that m and n exist. So $\text{ord}(x^{a/m}) = m$ and $\text{ord}(y^{b/n}) = n$. By the previous lemma, $ord(x^{a/m}y^{b/n}) = mn$.

Now, we shall construct m and n. Consider a and b in standard form:

$$a = \prod_{i \in \mathbb{N}} p_i^{\alpha_i}$$

$$b = \prod_{i \in \mathbb{N}} p_i^{\beta_i}$$

For each i, if $\alpha_i < \beta_i$, divide a by $p_i^{\alpha_i}$; if $\alpha_i > \beta_i$, divide b by $p_i^{\beta_i}$; if $\alpha_i = \beta_i$, perform either operation. In the end, we are left with two numbers satisfying the conditions for m and n. □

It may be helpful to illustrate this construction of m and n by a numerical example.

Example 2.5. Suppose $a = 120$ and $b = 36$. In standard form,

$$a = 120 = 5 \cdot 3 \cdot 2 \cdot 2 \cdot 2$$

$$b = 36 = 3 \cdot 3 \cdot 2 \cdot 2$$

Completing the procedure in the previous lemma, we have

$$m = 40 = 5 \cdot 2 \cdot 2 \cdot 2$$

$$n = 9 = 3 \cdot 3$$

with $mn = 360 = lcm(120, 36)$. The reader may easily verify that m and n satisfy the other conditions in Lemma 2.4.

From the previous lemma, we can immediately draw an important conclusion that will help us in the next and final theorem. Suppose x and y are elements of the abelian group $\{G, \cdot\}$, and suppose $ord(x) = a$ and $ord(y) = b$. Then $a \mid b$, $b \mid a$, or there exists $z \in G$ such that $ord(z) > a, b$.

Theorem 2.6. *The multiplicative group of $\mathbb{Z}_p$ is cyclic*

Proof. Let k be the highest order of all elements in $\mathbb{Z}_p^*$. Consider the polynomial $\lambda^k - 1$ over the field $\mathbb{Z}_p$. By what we just stated, the order of any element in $\mathbb{Z}_p^*$ divides k. Therefore, every such element is a zero of the polynomial, implying $k \geq p - 1$. But the order of an element of a group cannot be greater than the order of the group. So $k = p - 1$, and any element whose order is k is a generator of $\{\mathbb{Z}_p^*, \cdot\}$. □

By what was discussed before Lemma 2.3, this theorem implies Proposition 2.2.

3. Representing Integers as Sums of Two Squares

To familiarize the reader with the problem, we begin with a simple geometric interpretation of the Diophantine equation

$$N = x^2 + y^2 \tag{3.1}$$

Consider a circle with radius N, centered at $(0, 0)$ in $\mathbb{R}^2$. Then N can be represented as a sum of two integer squares if and only if there exist integer coordinates (x, y) on the circle. Obviously, if the coordinates (x, y) in the first quadrant satisfy Equation 3.1, then $(-x, y)$, $(-x, -y)$, and $(x, -y)$ are also solutions. Because they differ only in sign, these solutions are *not essentially distinct.*

Clearly, not all circles of radius N have points of integer coordinates. Our task for this section is to develop a criterion for two-square representability. The first step is to demonstrate a necessary condition for a number to be representable.

Theorem 3.2. *Suppose $p = 4k+3$ is prime. If a number N divisible by p is a sum of two squares, then the power of p in the standard form of N is even.*

Proof. Let $N = x^2 + y^2$. Observe that $p \mid N$ implies $x^2 \equiv -y^2 \pmod{p}$. We know from Proposition 2.2 that -1 is a quadratic non-residue (mod p). Thus, the only possible solutions to the congruence are $x \equiv y \equiv 0 \pmod{p}$. Then $p \mid x$ and $p \mid y$, implying $p^2 \mid N$. Now, let $N = p^2 d$. By the same argument, $p \mid d$ implies $p^2 \mid d$. Therefore, any divisor of N divisible by p is divisible by p^2. So the power of p in the standard form of N is even. □

An immediate corollary of this theorem is that *any* number of the form $4k + 3$ is not representable. The proof is left as an exercise for the reader.

We now want to show that the conclusion of Theorem 3.2 is also a sufficient condition. Reformulating the statement accordingly, we have the following criterion:

Theorem 3.3 (Two-Squares Theorem)**.** *A number N is a sum of two squares if and only if the power a_i of a factor $p_i = 4k + 3$ in the standard form of N is even.*

The proof of the remaining half of this theorem is greatly simplified by an identity,

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2 \tag{3.4}$$

By this equation, if each factor in the standard form of a number is representable, then the number itself is representable. So let us examine the standard form of a number N satisfying the condition of the theorem. Trivially, $2 = 1^2 + 1^2$, so any power of 2 is representable. By assumption, the power a_i of a prime $p_i = 4k + 3$ is even, so $p_i^{a_i}$ is representable. All other primes are of the form $4k + 1$. Therefore, to prove Theorem 3.3, we must show that a prime $p = 4k + 1$ is a sum of two squares.

Our method to prove the last statement is Fermat's 'method of descent.' First, we prove that a multiple of p is representable. Second, we prove that the least representable multiple is p itself.

Theorem 3.5. *A prime $p = 4k + 1$ is a sum of two squares.*

Proof. By Proposition 2.2, -1 is a quadratic residue (mod p). Therefore, the congruence $z^2 + 1 \equiv 0 \pmod{p}$ is soluble. So there exists $m > 0$ such that $mp = z^2 + 1^2$. Now, consider the more general congruence, $x^2 + y^2 \equiv 0 \pmod{p}$. Let x_1, y_1 be residues of x, y (mod p) such that $|x_1|, |y_1| < \frac{1}{2}p$ Then

$$0 < m = \frac{1}{p}(x_1^2 + y_1^2) < \frac{1}{p}(\frac{1}{2}p^2) < p \tag{3.6}$$

Let $m = m_0$ be the least value for which mp is representable as a sum of two squares.

$$m_0 p = x_1^2 + y_1^2 \tag{3.7}$$

By contradiction, assume $m_0 > 1$. Let x_2, y_2 be residues of x_1, y_1 (mod m_0) such that $|x_2|, |y_2| \leq \frac{1}{2}m_0$. Observe that

$$x_2^2 + y_2^2 \equiv x_1^2 + y_1^2 \equiv 0 \pmod{m_0}$$

This implies that there exists r such that

$$m_0 r = x_2^2 + y_2^2 \tag{3.8}$$

We wish to show that $r \neq 0$. If $r = 0$, then

$$x_2 = y_2 = 0 \Rightarrow m_0 \,|\, x_1,\ m_0 \,|\, y_1 \Rightarrow m_0^2 \,|\, m_0 p \Rightarrow m_0 \,|\, p$$

But by Equation 3.6, $m_0 < p$, and by assumption, $1 < m_0$. So $r = 0$ contradicts that p is prime. Observe that

$$r = \frac{1}{m_0}(x_2^2 + y_2^2) \leq \frac{1}{m_0}(\frac{1}{2}m_0^2) < m_0$$

Multiplying Equation 3.7 and Equation 3.8 and applying Equation 3.4, we have

$$m_0^2 rp = (x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1x_2 + y_1y_2)^2(x_1y_2 - y_1x_2)^2 \tag{3.9}$$

Note that each factor on the right is divisble by m_0:

$$x_1x_2 + y_1y_2 \equiv x_1^2 + y_1^2 \equiv 0 \pmod{m_0}$$

$$x_1y_2 - y_1x_2 \equiv x_1y_1 - y_1x_1 \equiv 0 \pmod{m_0}$$

So let $m_0X = (x_1x_2 + y_1y_2)$ and $m_0Y = (x_1y_2 - y_1x_2)$. Then dividing Equation 3.9 by m_0^2, we have

$$rp = X^2 + Y^2$$

Consequently, there exists $r < m_0$, $r \neq 0$, such that rp is representable as a sum of two squares. This contradicts the definition of m_0. Therefore, $m_0 = 1$. □

From the the two-squares theorem, we can deduce that $g(2) > 2$.

4. Representing Integers as Sums of Four Squares

Perhaps the first question that comes to the reader's mind is why we address four squares before we address three. As will become evident, the proofs for the two- and four-square theorems are very similar, while the three-square theorem requires completely different methods.

To hint at the similarity, we again touch on the geometry of the problem. Consider the Diophantine equation

$$N = w^2 + x^2 + y^2 + z^2 \tag{4.1}$$

and imagine a 4-sphere of radius N, centered at $(0,0,0,0)$ in $\mathbb{R}^4$. Then N can be represented as a sum of four integer squares if and only if there exist integer coordinates (w,x,y,z) on the 4-sphere. With regard to sign, there are sixteen solutions that are not essentially distinct.

As mentioned in the introduction, all numbers can be represented as a sum of four squares. We formalize this statement thus:

Theorem 4.2 (Four-Squares Theorem)**.** *Any integer is a sum of four squares.*

Our work is simplified by an analog to Equation 3.4:

$$(a^2+b^2+c^2+d^2)(A^2+B^2+C^2+D^2) = \begin{array}{ll} (aA+bB+cC+dD)^2 & + \\ (aB-bA-cD+dC)^2 & + \\ (aC+bD-cA-dB)^2 & + \\ (aD-bC+cB-dA)^2 & \end{array} \tag{4.3}$$

So if each factor in the standard form of a number is representable, then the number itself is representable. Once again, let us examine the standard form of a number N; here, however, N is completely arbitrary. Trivially, $2 = 1^2+1^2+0^2+0^2$, so any power of 2 is representable. It immediately follows from Theorem 3.5 that any power of a prime $p = 4k+1$ is representable. Thus, in order to prove the above theorem, we must show that any prime of the form $4k+3$ is a sum of four squares.

We again employ the 'method of descent,' making the structure of the proof nearly identical to that of the two-squares proof. Just as we did before, we show that a multiple of p is representable from a soluble congruence: $x^2+y^2+1 \equiv 0 \pmod{p}$. Because the solubility of this congruence is not immediately obvious, however, we prove it in a separate lemma.

Lemma 4.4. *Suppose $p = 4k+3$ is prime. The congruence $x^2+y^2+1 \equiv 0 \pmod{p}$ is soluble.*

Proof. Consider the congruence in this form: $x^2+1 \equiv -y^2 \pmod{p}$. $-1 \equiv p-1$ is a quadratic non-residue modulo $p = 4k+3$. Now that we have proven that a positive quadratic non-residue exists, let z be the least postive non-residue (mod p). Trivially, $z > 1$. Then there exists a residue w such that $w+1 = z$. Observe that there exists y such that $z \equiv -y^2 \pmod{p}$. Therefore, the congruence $x^2+1 \equiv -y^2 \pmod{p}$ is soluble. □

We are now ready for the main proof.

Theorem 4.5. *A prime $p = 4k+3$ is a sum of four squares.*

Proof. By the previous lemma, the congruence $x^2+y^2+1 \equiv 0 \pmod{p}$ is soluble. So there exists $m > 0$ such that $mp = x^2+y^2+1$. Now, consider the more general congruence, $x_1^2+x_2^2+x_3^2+x_4^2 \equiv 0 \pmod{p}$. Let y_i, $i \in \{1,2,3,4\}$, be a residue of $x_i \pmod{p}$ such that $|y_i| < \frac{1}{2}p$. Then

$$0 < m = \frac{1}{p}\sum_{i=1}^{4} y_i^2 < \frac{1}{p}(p^2) = p \tag{4.6}$$

Let $m = m_0$ be the least value for which mp is representable as a sum of four squares.

$$m_0 p = y_1^2+y_2^2+y_3^2+y_4^2 \tag{4.7}$$

By contradiction, assume $m_0 > 1$. Let z_i be a residue of $y_i \pmod{m_0}$ such that $|z_i| \leq \frac{1}{2}m_0$. Observe that

$$\sum_{i=1}^{4} z_i^2 \equiv \sum_{i=1}^{4} y_i^2 \equiv 0 \pmod{m_0}$$

This implies that there exists r such that

$$m_0 r = z_1^2 + z_2^2 + z_3^2 + z_4^2 \tag{4.8}$$

We wish to show that $r \neq 0$. If $r = 0$, then

$$\forall i, z_i = 0 \;\Rightarrow\; m_0 \,|\, y_i \;\Rightarrow\; m_0^2 \,|\, m_0 p \;\Rightarrow\; m_0 \,|\, p$$

But by Equation 4.6, $m_0 < p$, and by assumption, $1 < m_0$. So $r = 0$ contradicts that p is prime. Observe that

$$r = \frac{1}{m_0}\sum_{i=1}^{4} z_i^2 \leq \frac{1}{m_0}(m_0^2) = m_0$$

But we want to show that r is *strictly* less than m_0. If $r = m_0$, then $|z_i| = \frac{1}{2}m_0$ and $\sum_{i=1}^{4} z_i^2 \equiv 0 \pmod{m_0^2}$. Now, $|z_i| = \frac{1}{2}m_0$ implies that $z \equiv \frac{1}{2}m_0 \equiv -\frac{1}{2}m_0 \pmod{m_0}$. Recall $z_i \equiv y_i \pmod{m_0}$. Therefore,

$$m_0 \,|\, y_i - z_i,\; m_0 \,|\, y_i + z_i \;\Rightarrow\; m_0^2 \,|\, y_i^2 - z_i^2 \;\Rightarrow\; y_i^2 \equiv z_i^2 \pmod{m_0^2}$$

Then $\sum_{i=1}^{4} y_i^2 \equiv \sum_{i=1}^{4} z_i^2 \equiv 0 \pmod{m_0^2}$. So $m_0^2 \,|\, m_0 p$, which implies $m_0 \,|\, p$. As before, this contradicts that p is prime. Now, multiplying Equation 4.7 and Equation 4.8 and applying Equation 4.3,

$$m_0^2 rp = (y_1^2+y_2^2+y_3^2+y_4^2)(z_1^2+z_2^2+z_3^2+z_4^2) = \begin{array}{ll} (y_1z_1 + y_2z_2 + y_3z_3 + y_4z_4)^2 & + \\ (y_1z_2 - y_2z_1 - y_3z_4 + y_4z_3)^2 & + \\ (y_1z_3 + y_2z_4 - y_3z_1 - y_4z_2)^2 & + \\ (y_1z_4 - y_2z_3 + y_3z_2 - y_4z_1)^2 & \end{array} \tag{4.9}$$

Note that each factor on the right is divisble by m_0:

$$\left.\begin{array}{rcl} y_1z_1 + y_2z_2 + y_3z_3 + y_4z_4 & \equiv & y_1^2 + y_2^2 + y_3^2 + y_4^2 \\ y_1z_2 - y_2z_1 - y_3z_4 + y_4z_3 & \equiv & y_1y_2 - y_2y_1 - y_3y_4 + y_4y_3 \end{array}\right\} \equiv 0 \pmod{m_0}$$

and similarly for the other two factors. Then dividing Equation 4.9 by m_0^2, we have

$$rp = X_1^2 + X_2^2 + X_3^2 + X_4^2,\; X_i \in \mathbb{Z}$$

So there exists $r < m_0$, $r \neq 0$, such that rp is representable as a sum of four squares. This contradicts the definition of m_0. Therefore, $m_0 = 1$. □

From the four-square theorem, we deduce that $g(2) \leq 4$.

5. Representing Integers as Sums of Three Squares

To prove finally that $g(2) = 4$, we must show that certain numbers cannot be represented as sums of three squares. To start, consider the Diophantine equation

$$N = x^2 + y^2 + z^2 \tag{5.1}$$

We suspect that the parities of x, y, and z give rise to distinct sets of representable numbers. Three variables, each with two possibilites for partiy, yield eight possibile sets of values for N. Thus, we are naturally lead to consider the sum in Equation 5.1

to the modulus 8. Clearly, if n is even, then n^2 is congruent to 0 or 4 (mod 8). It remains to show the possible congruences where n is odd.

Lemma 5.2. *If n is odd, then $n^2 \equiv 1 \pmod 8$.*

Proof. n is odd implies that $n+1$ and $n-1$ are even. Consider the equation

$$n^2 - 1 = 4(\frac{n+1}{2})(\frac{n-1}{2})$$

Now, $(\frac{n+1}{2})$ and $(\frac{n-1}{2})$ are consecutive integers. So either $(\frac{n+1}{2})$ or $(\frac{n-1}{2})$ is even. Therefore, $n^2 - 1 \equiv 0 \pmod 8$, implying $n^2 \equiv 1 \pmod 8$. □

By these considerations, we can directly calculate the possible congruences for a sum of three squares:

$$x^2 + y^2 + z^2 \equiv \begin{cases} 0 \text{ or } 4 & \text{if x, y, z even} \\ 1 \text{ or } 5 & \text{if x, y even} \\ 2 \text{ or } 6 & \text{if x even} \\ 3 & \text{if x, y, z odd} \end{cases} \pmod 8 \tag{5.3}$$

Because 7 is not a residue (mod 8), no number of the form $8k+7$ can be represented as a sum of three squares. This conclusion is sufficient to prove that $g(2) = 4$. But to show which numbers *are* representable, we must prove a stronger theorem.

Theorem 5.4. *No number $N = 4^n(8k+7)$, where $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, is the sum of three squares.*

Proof. By contradiction, assume

$$N = 4^n(8k+7) = x^2 + y^2 + z^2 \tag{5.5}$$

where x, y, and z are integers. Let $n = n_0$ be the least power for which N is representable. Note that $n_0 > 0$, for if $n_0 = 0$, then $N = (8k+7) \equiv 7 \pmod 8$. Now, by Equation 5.3, N is even *and* $4 \mid N$ implies that x, y, and z are even. So let $X = \frac{1}{2}x$, $Y = \frac{1}{2}y$, and $Z = \frac{1}{2}z$. Then dividing Equation 5.5 by 4, we have

$$N = 4^{n_0-1}(8k+7) = X^2 + Y^2 + Z^2$$

implying $n = n_0 - 1$ is a power for which N is representable. This contradicts the definition of n_0. Therefore, N is not representable. □

We assert that all numbers *not* of the form $4^n(8k+7)$ are representable. Combining this statement with the previous theorem, we can formulate a criterion for three-square representability:

Theorem 5.6 (Three-Squares Theorem)**.** *A number N is a sum of three squares if and only if N is not of the form $4^n(8k+7)$.*

Unfortunately, there is no elementary proof for the remaining half of this theorem because no identity like Equation 3.4 or Equation 4.3 exists for three squares. For a counterexample, take any number congruent to 3 and any number congruent to 5 (mod 8). Trivially, $3 = 1^2 + 1^2 + 1^2$, and $5 = 2^2 + 1^2 + 0^2$. However, $3 \cdot 5 = 15 \equiv 7 \pmod 8$, which we just declared to be not representable. Consequently, the proof requires methods we have not discussed thus far. We simply leave it as a topic for further investigation, providing only a brief outline to serve as a starting point.

The proof we have in mind is based on the theory of *quadratic forms*, viewed through the lens of linear algebra. The following is a list of informal definitions of the principal terms.

- A *quadratic form* is a homogeneous polynomial of degree 2 in k variables. In the general case, we denote a quadratic form as $Q(x_1, x_2, \ldots, x_k) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$. One may regard the input of Q as a k-tuple $\mathbf{x}$ in $\mathbb{R}^k$. For the proof, we assume that the coefficients a_{ij} are integers.
- Let $Q_1(\mathbf{x})$ and $Q_2(\mathbf{y})$ be quadratic forms. They are *equivalent* if for some $k \times k$ matrix C with determinant 1, $x_i = \sum_{j=1}^{k} y_j$ for all i. The reader may verify that this relation indeed constitutes an equivalence relation.
- A *positive definite* quadratic form is one whose values are all positive; that is, for all $\mathbf{x}$, $Q(\mathbf{x}) > 0$.
- The *discriminant* of a quadratic form Q is the determinant of the coefficient matrix $A = (a_{ij})$, denoted as $d(Q)$.

From these definitions, one must demonstrate certain elementary properties of binary (two-variable) and ternary (three-variable) quadratic forms, ultimately proving the following lemma:

Lemma 5.7. *Every positive definite ternary quadratic form Q with $d(Q) = 1$ is equivalent to a sum of three squares.*

.

To begin proving Theorem 5.6, let us re-examine the possible congruences of a sum of three squares modulo 8. If a number $N \equiv 7 \pmod 8$, then N is not representable, so we ignore this case. Suppose now that $N = 4^n d$, where 4 does not divide d. Clearly, if d is representable, then N is representable. Therefore, it is sufficient to prove that any N *not* congruent to 0, 4, or 7 (mod 8) is representable. Operating under the previous lemma, one must then show that such N can be represented by a positive definite ternary quadratic form with discriminant 1. The proof would then be complete.

Acknowledgments. I would like to thank my graduate student mentors, Asaf Hadari and Rita Jiménez. Their boundless knowledge, patience, and enthusiasm enabled me to extract the most out of this project. I would also like to thank Professor Peter May for organizing the 2009 REU, providing the perfect opportunity to pursue extracurricular mathematics.

References

[1] T. Apostol. Calculus. Vol. 2. 2nd ed., New York, John Wiley & Sons. 1969.
[2] H. Davenport. The Higher Arithmetic. Cambridge, Cambridge University Press. 1982.
[3] W. Deskins. Abstract Algebra. New York, Dover Publications, Inc. 1995.
[4] P. Erdős and J. Surányi. Topics in the Theory of Numbers. New York, Springer. 2003.
[5] E. Grosswald. Representations of Integers as Sums of Squares. New York, Springer-Verlag. 1985.
[6] G. Hardy and E. Wright. An Introduction to the Theory of Numbers. 5th ed., Oxford, Oxford University Press. 1979.
[7] D. Loeffler. (27 April, 2003). Is the multiplicative group mod p necessarily cyclic? Message posted to https://nrich.maths.org/discus/messages/2069/5977.html?1051613517.