Sums of Squares Expo

Swastika Dey, Anderson Hao, Jessie Wang, Jeffrey Xu, Yuanrui Zhao
Mentor: Clyde Kertzer

August 3, 2025

1 Distances Between Lattice Points

We will first motivate our exploration into the sums of squares and r-gonal numbers with an interesting geometric problem–finding the number of distinct nonnegative distances between lattice points in a grid.

Definition 1.1. An $n \times n$ grid is the set of all points (x, y) such that $x, y \in \mathbb{Z}$ and $0 \leq x, y \leq n$.

Lemma 1.2

The number of distinct nonnegative distances between two lattice points on an $n \times n$ grid is bound above by $\binom{n+2}{2}$.

Proof. Consider two lattice points (a_1, b_1) and (a_2, b_2) on an $n \times n$ grid. The distance between (a_1, b_1) and (a_2, b_2) depends only on the horizontal and vertical distance between the two points; namely, $|a_1 - a_2|$ and $|b_1 - b_2|$. Let the horizontal distance be H and the vertical distance be V. We know that

$$0 \leq a_1, a_2, b_1, b_2 \leq n.$$

As a result,

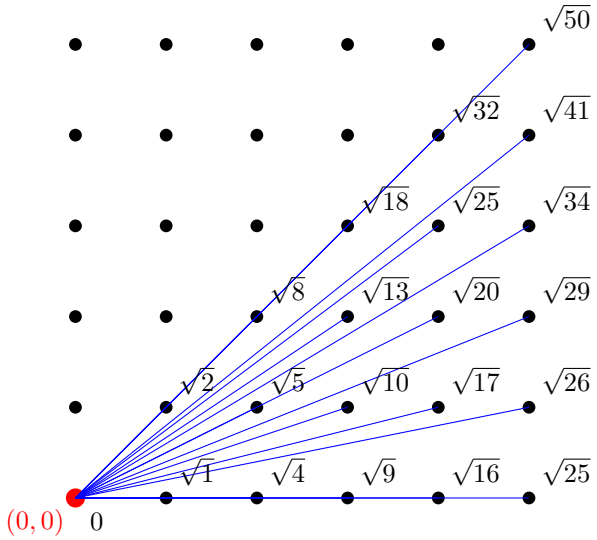
$$0 \leq |a_1 - a_2| \text{ and } |b_1 - b_2| \leq n.$$

If $|a_1 - a_2| = |b_1 - b_2|$, we have a total of $n + 1$ possible pairs (H, V). If $|a_1 - a_2| \neq |b_1 - b_2|$, notice that swapping the values of H and V does not affect the distance. Hence, we can divide the number of pairs (H, V) by 2 to get $n(n + 1)/2$ distances. Adding our counts, we have a maximum of

$$n + 1 + \frac{n(n+1)}{2} = \frac{(n+2)(n+1)}{2} = \binom{n+2}{2}$$

distinct nonnegative distances between two lattice points on an $n \times n$ grid. This is an upper bound and not an exact count because some distances can be counted by multiple pairs (H, V). □

Total distinct distances: 20



Remark 1.3. In the grid above, 25 is counted by $(0, 5)$ and $(3, 4)$, so there are only 20 distinct distances on a 5×5 grid. This is the smallest grid where our upper bound is not the exact count. To answer this question, we must determine the number of ways to express a number as the sum of two squares...

2 Sums of Two Squares

We will now fully characterize the natural numbers that can be expressed as the sum of two squares, as well as the number of distinct expressions possible. We begin by constructing a multiplicatively closed subset of the natural numbers whose elements are all expressible as the sum of two squares.

Proposition 2.1 (Diophantus-Brahmagupta-Fibonacci-Kertzer Identity)
For all $a, b, c, d \in \mathbb{Z}$, we have that

$$(a^2 + b^2)(c^2 + d^2) = (ad + bc)^2 + (ac - bd)^2.$$

Proof. Let $a, b, c, d \in \mathbb{Z}$. Expanding yields

$$\begin{aligned}(a^2 + b^2)(c^2 + d)^2 &= a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 \\ &= a^2c^2 + 2abcd + b^2d^2 + a^2d^2 - 2abcd + b^2c^2 \\ &= (ac + bd)^2 + (ad - bc)^2.\end{aligned}$$

□

Two useful results follow:

Corollary 2.1.1

The set of naturals that are expressible as the sum of two squares is closed under multiplication.

Corollary 2.1.2

For all $x, y \in \mathbb{Z}[i]$, we have that $N(xy) = N(x)N(y)$.

We begin with the first result, which reduces the problem of constructing the subset to determining which primes are expressible as the sum of two squares. Note that all even primes are expressible as the sum of two squares:

Remark 2.2. $2 = 1^2 + 1^2$.

The only primes that remain are odd primes, which can be split into two groups by their remainders mod 4. The primes that are 3 mod 4 cannot be expressed as the sum of two squares due to the following two lemmas.

Lemma 2.3

Let $k \in \mathbb{Z}$. Then k^2 is 0 or 1 mod 4.

Proof. Let $k \in \mathbb{Z}$. If k is even, then $k = 2k_1$ for some $k_1 \in \mathbb{Z}$. Squaring gives

$$k^2 \equiv (2k_1)^2 \equiv 4k_1^2 \equiv 0 \pmod 4.$$

Now, if k is odd, then $k = 2k_1 + 1$ for some $k_1 \in \mathbb{Z}$. Squaring gives

$$k^2 \equiv (2k_1 + 1)^2 \equiv 4k_1^2 + 4k_1 + 1 \equiv 1 \pmod 4.$$ □

Lemma 2.4

Let $n \in \mathbb{N}$. If $n \equiv 3 \pmod 4$, then n cannot be expressed as a sum of two squares.

Proof. Let $k \in \mathbb{Z}$. By Lemma 2.3, k^2 is 0 or 1 mod 4. Thus, the sum of two squares can only be $0 + 0 = 0$, $0 + 1 = 1 + 0 = 1$, or $1 + 1 = 2$ mod 4. □

On the other hand, every prime that is 1 mod 4 can be expressed as the sum of two squares.

Theorem 2.5

If p is a prime and $p \equiv 1 \pmod 4$, then there exist $a, b \in \mathbb{Z}$ such that $p = a^2 + b^2$.

We begin by proving supporting lemmas.

Lemma 2.6
Given a prime p and a unit $u \in U_p$, there exists a unique $u^{-1} \in U_p$ such that $uu^{-1} = 1$.

Proof. By the definition of a unit, there exists some $u^{-1} \in U_p$ such that $uu^{-1} = 1$. Now, assume that $uv = 1$. Then, $v = u^{-1}uv = u^{-1}$. □

Lemma 2.7
Given $u \in U_p$, we have that $u = u^{-1}$ if and only if $u = \pm 1$.

Proof. If $u = \pm 1$, then $u \cdot u = u^2 = (\pm 1)^2 = 1$, so $u^{-1} = u$ by Lemma 2.6.

Conversely, if $u = u^{-1}$, then $1 \equiv uu^{-1} \equiv u^2 \pmod{p}$. Subtracting 1 from both sides and factoring gives

$$(u-1)(u+1) \equiv u^2 - 1 \equiv 0 \pmod{p}.$$

Since p is prime, one of $u-1$ and $u+1$ must be divisible by p, forcing u to be 1 or -1. □

Lemma 2.8 (Wilson's Theorem)
For all primes p, we have that $(p-1)! \equiv -1 \pmod{p}$.

Proof. For $p = 2$, the statement is true, as $1! \equiv 1 \equiv -1 \pmod{2}$. Since 2 is the only even prime, we may now assume p is odd. Now, because p is prime, each of $1, 2, \ldots, p-1$ is relatively prime to p; consequently, all of them are units in $\mathbb{Z}_p$. Now, take the units $2, 3, \ldots, p-2$. None of the units are 1 or -1, so Lemmas 2.6 and 2.7 imply that each of them has a unique multiplicative inverse that is neither 1 nor -1. As a result, due to $p-3$ being even, we may pair up multiplicative inverses to get that

$$(p-1)! \equiv (p-1)(p-2)\cdots(2)(1) \equiv (p-1)\left(1^{\frac{p-3}{2}}\right) \equiv -1 \pmod{p}.$$ □

Lemma 2.9
Let p be an odd prime. There exists a $j \in U_p$ such that $j^2 \equiv -1 \pmod{p}$ if and only if $p \equiv 1 \pmod{4}$.

Proof. If $p \equiv 1 \pmod{4}$, we have that $\frac{p-1}{2}$ is even. Then, it follows from Lemma 2.8 that:

$$\begin{aligned}(p-1)! &\equiv (p-1)(p-2)\cdots\left(\frac{p+1}{2}\right)\left(\frac{p-1}{2}\right)\cdots(2)(1)\\ &\equiv (-1)(-2)\ldots\left(-\frac{p-1}{2}\right)\left(\frac{p-1}{2}\right)\cdots(2)(1)\\ &\equiv \left(-1^2\right)\left(-2^2\right)\cdots\left(-\left(\frac{p-1}{2}\right)^2\right)\end{aligned}$$

$$\equiv (-1)^{\frac{p-1}{2}} \left(1^2\right)\left(2^2\right)\cdots\left(\left(\frac{p-1}{2}\right)^2\right)$$
$$\equiv (1)\left(\left(\frac{p-1}{2}\right)!\right)^2$$
$$\equiv \left(\left(\frac{p-1}{2}\right)!\right)^2$$

Taking $j = \left(\frac{p-1}{2}\right)!$ thus satisfies the lemma statement.

Now, if there exists a $j \in U_p$ such that $j^2 \equiv -1 \pmod 4$, it follows that the order of j is 4, as $j^4 \equiv 1 \pmod p$ and no smaller power of j produces 1. Then, because $j^{p-1} \equiv 1 \pmod p$ by Fermat's Little Theorem, we have that $4|(p-1)$, implying that p is 1 mod 4. □

We are now ready to show that all primes that are 1 mod 4 can be represented as the sum of two squares.

Proof of Theorem 2.5. From Lemma 2.9, there exists $j \in U_p$ such that $j^2 \equiv -1 \pmod p$. Two cases arise based on the size of j.

Case 1: $1 \leq j \leq \sqrt{p}$

If $j \leq \sqrt{p}$, then $j^2 + 1 \leq p + 1$. As j is a positive integer and $j^2 + 1 \equiv 0 \pmod p$, it follows that $j^2 + 1$ must equal p. Thus, $a = j$ and $b = 1$ satisfy $a^2 + b^2 = p$.

Case 2: $j > \sqrt{p}$

Let $q = \lfloor\sqrt{p}\rfloor$. The smallest prime that is 1 mod 4 is 5, so we may assume $p \geq 5$ and $q \geq 2$. From the definition of floor, we have that

$$k \leq \sqrt{p} < k + 1$$

which when squared yields

$$k^2 \leq p < k^2 + 2k + 1.$$

Since p is prime and an integer, $k(k+2) = k^2 + 2k \neq p$. Thus, we may improve the upper bound on p to

$$p \leq k^2 + 2k - 1.$$

Now, for each $1 \leq i \leq k$, let $x_i \equiv ij \pmod p$ and $0 \leq x_i < p$. We have for each i that

$$i^2(j^2 + 1) \equiv i^2 j^2 + i^2 \equiv x_i^2 + i^2 \equiv 0 \pmod p.$$

Then, if there exists a $x_i \leq k$, the size of $x_i^2 + i^2$ can be bounded:

$$0 < x_i^2 + i^2 \leq 2k^2 < 2k^2 + 4 = 2p.$$

As $x_i^2 + i^2 \equiv 0 \pmod p$, the inequality forces $x_i^2 + i^2$ to be p. Taking $a = x_i$ and $b = i$ concludes; therefore, we may assume that all x_i are at least $k + 1$.

Next, if there exists a $x_i \geq p-k$, we have that $x_i \equiv -i' \pmod{p}$ for some $1 \leq i' \leq k$. It follows that

$$\begin{aligned} x_i &\equiv -i' && \pmod{p} \\ ij &\equiv (j^2)i' && \pmod{p} \\ ij^4 &\equiv i'j^5 && \pmod{p} \\ i &\equiv i'j && \pmod{p} \\ i &\equiv x_{i'} && \pmod{p}. \end{aligned}$$

However, because $1 \leq i \leq k$ implies that $1 \leq x_{i'} \leq k$, this contradicts our earlier assumption. Consequently, we may assume that $k+1 \leq x_i \leq p-k-1$ for each x_i.

Finally, using the Pigeonhole Principle with the k x_i's and the range they all lie in, we have that there exist distinct x_i and x_i' such that

$$\begin{aligned} |x_i - x_{i'}| &\leq \frac{p-k-1-(k+1)}{k-1} \\ &\leq \frac{p-2k-2}{k-1} \\ &\leq \frac{(k^2+2k-1)-2k-2}{k-1} \\ &\leq \frac{k^2-3}{k-1} \\ &< \frac{k^2-1}{k-1} \\ &< k+1 \\ &\leq k \end{aligned}$$

Without loss of generality, let $i > i'$ and $1 \leq |x_i - x_i'| = k' \leq k$. Then, $ji - ji' \equiv \pm k' \pmod{p}$, so $j(i-i') \equiv \pm k' \pmod{p}$. Thus, $x_{i-i'} = k'$ or $p-k'$. Since $p-k' \geq p-k > p-k-1$, we have that $x_{i-i'}$ must be k'. Taking $a = k'$ and $b = i-i'$ thus concludes. □

Theorem 2.10

A natural k can be expressed as the sum of two squares if and only if

$$k = 2^\ell \prod_{q \equiv 3 \bmod 4} q^2 \prod_{p \equiv 1 \bmod 4} p,$$

where $\ell \in \mathbb{Z}_{\geq 0}$. Here, p and q are prime.

Proof. Let p be a 1 mod 4 prime and q be a 3 mod 4 prime.

($\Rightarrow$) All numbers of the form of p, q^2, and 2 work: p by Theorem 2.5, $0^2 + q^2 = q^2$, and $1^2 + 1^2 = 2$. Then all numbers that are the product of 2's, p's, and q^2's are achievable by Corollary 2.1.1, as desired.

($\Leftarrow$) Take any odd prime r that divides k. We will show that if $r^2 \nmid k$, then $r \equiv 1 \bmod 4$. This

means that all k that can be expressed as the sum of two squares are of the claimed form.

Write $k = a^2 + b^2$. If $r \mid \gcd(a, b)$, then $r \mid a$, $a \mid b$, so $r^2 \mid a^2$, $r^2 \mid b^2$, and $r^2 \mid a^2 + b^2 = k$. But $r^2 \mid k$, so we discard this case. Note that if r divides one of a and b then r must divide the other, as $r \mid k$. Since r does not divide both, we may now assume that $r \nmid a$, $r \nmid b$.

Then take $a^2 = k - b^2 \mod r$ to get $a^2 \equiv -b^2 \bmod r$. Multiplying by $(b^{-1})^2$ on both sides gives $(ab^{-1})^2 \equiv -1 \bmod r$. Since $-1 \not\equiv 1 \pmod r$, $\operatorname{ord}_r(ab^{-1}) \neq 1, 2$. Squaring gives $(ab^{-1})^4 \equiv 1 \bmod r$, so $\operatorname{ord}_r(ab^{-1}) \mid 4$. Thus, $\operatorname{ord}_r(ab^{-1}) = 4$. Therefore, $4 \nmid r - 1$, so $r \equiv 1 \bmod 4$ as needed. □

Lemma 2.11

In all three of the following cases, α is prime in $\mathbb{Z}[i]$.

(a) If $N(\alpha) = p$ where p is a 1 mod 4 prime, then there exist exactly two primes $\alpha \in \mathbb{Z}[i]$ (up to units), and these are α and $\overline{\alpha}$.

(b) If $p = 2$, only $\alpha = 1 + i$ works.

(c) If $N(\alpha) = q^2$ where q is a 3 mod 4 prime, then the only $\alpha \in \mathbb{Z}[i]$ is $\alpha = q$.

Proof. Let $\alpha = a + bi$.

(a) By Theorem 2.5, there exists some $\beta \in \mathbb{Z}[i]$ with $N(\beta) = p$. Then β must be prime, otherwise write $p = N(\beta) = N(a)N(b)$ with $a, b \in \mathbb{Z}[i]$ non-unit, which contradicts the primality of p, since $N(a), N(b) > 1$. Likewise, α is prime.

We know $N(\beta) \mid N(\alpha)$. We may also write $\beta\overline{\beta} \mid \alpha\overline{\alpha}$, which implies $\beta \mid \alpha\overline{\alpha}$. Then by FTA, since β is prime, either $\beta \mid \alpha$ or $\beta \mid \overline{\alpha}$. $N(\alpha) = N(\beta) = N(\overline{\beta})$, so $\alpha = \beta$ or $\alpha = \overline{\beta}$, as needed.

(b) We require $N(\alpha) = a^2 + b^2 = 2$, which is only possible when $\alpha = 1 + i$ (and associates).

(c) If α is not prime, write $q^2 = N(\alpha) = N(a)N(b)$ for some non-unit $a, b \in \mathbb{Z}[i]$. Then $N(a) = N(b) = q$ since q is prime. Let $a = x + yi$, so that $x^2 + y^2 = q$. This is impossible since q is 3 mod 4. Take $x^2 + y^2 = q \bmod 4$ to find $x^2 + y^2 \equiv \bmod 4$. Thus α is prime.

Now to show there is only one α. Let $\alpha = a + bi$. Then we know that $a^2 + b^2 = q^2$ from taking norms. Using a similar method to Theorem 2.10, and since $q \equiv 3 \bmod 4$, we must have that $q \mid \gcd(a, b)$. The only possible solution is then $(a, b) = (q, 0)$ and its permutations and negations. Thus, $\alpha = q$ up to multiplication by units. □

Lemma 2.12

The three types of norms in Lemma 2.11 are the only ones that correspond to primes in $\mathbb{Z}[i]$. In other words, if $N(\alpha) \neq 2, p, q^2$, then α cannot be prime.

Proof. Let $N(\alpha) = a_1 a_2 \cdots a_n$, where $a_1, \ldots, a_n \in \mathbb{Z}$ are of the form $2, p$, or q^2. Then by Lemma 2.11, we may find primes $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}[i]$ with $N(\alpha_i) = a_i$ for all $1 \leq i \leq n$. Thus, $\alpha = \alpha_1 \cdots \alpha_n$, so α is not prime, since none of the α_i are units. □

Theorem 2.13

Let

$$k = 2^\ell \prod_{q \equiv 3 \bmod 4} q^2 \prod_{p \equiv 1 \bmod 4} p$$

where $\ell \in \mathbb{Z}_{\geq 0}$. If the number of ordered pairs $(a, b) \in \mathbb{Z}^2$ that satisfy $a^2 + b^2 = k$ is defined as $r(k)$, then

$$r(k)/4 = \tau\left(\prod_{p \equiv 1 \bmod 4} p\right).$$

Proof. Let p be a 1 mod 4 prime and q be a 3 mod 4 prime, and let

$$m = \prod_{1 \bmod 4} p.$$

Note that $r(k)$ counts the number of $\alpha \in \mathbb{Z}[i]$ such that $N(\alpha) = k$. There are exactly $r(k)/4$ solutions α to $N(\alpha) = k$ up to multiplication by units, since there are 4 units in $\mathbb{Z}[i]$.

Now prime factorize α in $\mathbb{Z}[i]$. Since norm in $\mathbb{Z}[i]$ is multiplicative by Corollary 2.1.2, we have that

$$N(\alpha) = N\left(\prod_i \alpha_i\right) = \prod_i N(\alpha_i),$$

where each α_i is prime. This is unique by UPF in $\mathbb{Z}[i]$. Each $N(\alpha_i)$ is p, q^2, or 2, by Lemma 2.12.

Now break the product in terms of the prime factors of $N(\alpha)$ to find that

$$N(\alpha) = \prod_{\substack{i, \\ N(\alpha_i) = p}} N(\alpha_i) \prod_{\substack{i, \\ N(\alpha_i) = q^2}} N(\alpha_i) \prod_{\substack{i, \\ N(\alpha_i) = 2}} N(\alpha_i).$$

By Lemma 2.11, for the products where $N(\alpha_i) = q^2$ or 2, there is only one α_i with $N(\alpha_i) = q^2$ or 2, respectively. Thus, we need only count the number of ways to choose α_i when $N(\alpha_i) = p$. Thus, $r(k) = r(m)$.

As $N(\alpha)$ has only 1 mod 4 factors, we write it

$$N(\alpha) = \prod_{\substack{i, \\ N(\alpha_{i,j}) = p_i}} (N(\alpha_{i,1}) N(\alpha_{i,2}) \cdots N(\alpha_{i,k_i})),$$

where the p_i are distinct 1 mod 4 primes, and $\alpha_{i,j}$ are primes in $\mathbb{Z}[i]$. This comes from representing

$$\alpha = \prod_{\substack{i, \\ N(\alpha_{i,j}) = p_i}} (\alpha_{i,1} \alpha_{i,2} \cdots \alpha_{i,k_i}).$$

By Lemma 2.11, there are exactly 2 choices for each $\alpha_{i,j}$, where $1 \leq j \leq k_i$. These two choices

are β_i and $\overline{\beta_i}$.

We have exactly $k_i + 1$ ways to assign the $\alpha_{i,j}$ in each term: we may have x of the $\alpha_{i,j}$ be β_i and the other $k_i - x$ of the $\alpha_{i,j}$ be $\overline{\beta_i}$ for all $0 \leq x \leq k_i$.

Multiplying over i, the number of ways to choose all $\alpha_{i,j}$ where i and j are both free is

$$\prod_i (k_i + 1).$$

We will show these are distinct. Each of β_i and $\overline{\beta_i}$ are prime in $\mathbb{Z}[i]$ while UPF tells us distinct prime factorizations cannot result in the same α. Thus each of the

$$\prod_i (k_i + 1)$$

ways specified above are distinct, as needed. □

3 Sums of Three Squares

Lemma 3.1

If $n \equiv 7 \pmod 8$, then n cannot be written as the sum of three integer squares. That is, there do not exist integers x, y, z such that

$$n = x^2 + y^2 + z^2.$$

Proof. Every integer can be written in the form $8k + r$ where $r \in \{0, 1, 2, 3, 4, -3, -2, -1\}$. Also, note that $x^2 \equiv (-x)^2 \pmod 8$, so we only need to consider $x \in \{0, 1, 2, 3, 4\}$. Their squares mod 8 are

$$\begin{aligned}
(8k)^2 &= 64k^2 \equiv 0 \pmod 8 \\
(8k+1)^2 &= 64k^2 + 16k + 1 \equiv 1 \pmod 8 \\
(8k+2)^2 &= 64k^2 + 32k + 4 \equiv 4 \pmod 8 \\
(8k+3)^2 &= 64k^2 + 48k + 9 \equiv 1 \pmod 8 \\
(8k+4)^2 &= 64k^2 + 64k + 16 \equiv 0 \pmod 8
\end{aligned}$$

So for any integer a, we have that

$$a^2 \equiv 0, 1, 4 \pmod 8.$$

Now we consider all possible sums $x^2 + y^2 + z^2 \pmod 8$, where each square is in $\{0, 1, 4\}$. The representative combinations are

$$0+0+0=0, \quad 0+0+1=1, \quad 0+0+4=4, \quad 0+1+1=2, \quad 0+1+4=5,$$

$$0+4+4=0, \quad 1+1+1=3, \quad 1+1+4=6, \quad 1+4+4=1, \quad 4+4+4=4.$$

So, the possible values of $x^2 + y^2 + z^2$ mod 8 are $0, 1, 2, 3, 4, 5, 6$. Since 7 is not in this set, n cannot be written as $n = x^2 + y^2 + z^2$. □

The following lemma will be useful to prove our characterization of the sum of three squares.

Lemma 3.2

For integers x, y and z, if $x^2 + y^2 + z^2$ is divisible by 4, then x, y, z are all even.

Proof. By Lemma 2.3, the only possible square residues are 0 and 1 and that an odd number squared is 1 mod 4. The possible sums are 1, 2, or 3 mod 4. Hence, for the sum to be 0 mod 4, all three must be even. □

Theorem 3.3

Let $n = 4^a(8k+7)$ be a positive integer, where $a, k \in \mathbb{Z}_{\geq 0}$. Then n cannot be represented as the sum of three integer squares; that is, there do not exist integers x, y, z such that

$$n = x^2 + y^2 + z^2.$$

Proof. (Induction). We will induct on a. For the base case, when $a = 0$, we have $n = 8k + 7$, which cannot be written as the sum of 3 squares by Lemma 3.1. Now, assume that for some $a \geq 1$, no integer of the form $4^a(8k + 7)$ can be written as the sum of three squares.

Consider $n = 4^{a+1}(8k + 7)$. Suppose for the sake of contradiction, that

$$4^{a+1}(8k + 7) = x^2 + y^2 + z^2$$

for integers x, y, and z. By Lemma 3.2, x, y, and z are even. Write $x = 2x_0, y = 2y_0, z = 2z_0$ for integers x', y', z'. Substituting gives

$$4^{a+1}(8k + 7) = 4(x^2 + y_0^2 + z_0^2),$$

and dividing by 4 gives

$$4^a(8k + 7) = x^2 + y_0^2 + z_0^2,$$

which contradicts the inductive hypothesis. Therefore, no integer of the form $4^a(8k + 7)$ can be written as the sum of three integer squares. □

4 Sums of Four Squares

We show that all natural numbers are representable as the sum of four squares.

Theorem 4.1

For all $n \in \mathbb{N}$, there exists $a, b, c, d \in \mathbb{Z}$ such that $n = a^2 + b^2 + c^2 + d^2$.

We begin by proving supporting lemmas.

Lemma 4.2
The set of natural numbers that can be written as a sum of four squares is closed under multiplication.

Proof. Let $n = a^2 + b^2 + c^2 + d^2$ and $m = w^2 + x^2 + y^2 + z^2$. Then
$nm = (aw+bx+cy+dz)^2 + (ax-bw+cz-dy)^2 + (ay-bz-cw+dx)^2 + (az+by-cx-dw)^2$
so nm is also a sum of four squares. □

Lemma 4.3
Let p be an odd prime. If $p \equiv 3 \pmod 4$, then there exist $a, b, c, d \in U_p$ such that
$$a^2 + b^2 + c^2 + d^2 \equiv 0 \pmod p.$$

Proof. Let p be an odd prime that is 3 mod 4. Note that
$$\left(\frac{p-1}{2}\right)^2 \equiv \frac{p^2-2p+1}{4} \equiv \frac{1}{4} \equiv \frac{p+1}{4} \pmod p$$
and that $\frac{p+1}{4}$ is an integer, due to p being 3 mod 4. Then, we have that
$$\left(\frac{p-1}{2}\right)^2 + \left(\frac{p-1}{2}\right)^2 + \frac{p-1}{2} \equiv 0 \pmod p.$$
If there exist quadratic residues mod p that sum to $\frac{p-1}{2}$, or $\frac{p-1}{2}$ is itself a quadratic residue, we are done. Thus, assume that $\frac{p-1}{2}$ is a quadratic non-residue mod p. Now, consider the pairwise sums
$$1 + \frac{p-3}{2} \equiv 2 + \frac{p-5}{2} \equiv \cdots \equiv \frac{p-3}{4} + \frac{p+1}{4} \equiv \frac{p-1}{2} \pmod p$$
and
$$\frac{p+1}{2} + p - 1 \equiv \frac{p+3}{2} + p - 2 \equiv \cdots \equiv \frac{3p-1}{4} + \frac{3p-1}{4} \equiv \frac{p-1}{2} \pmod p$$
In total, there are $\frac{p-1}{2}$ pairs that sum to $\frac{p-1}{2}$, with every unit except for $\frac{p-1}{2}$ included in one and only one sum. Because $\frac{p-1}{2}$ is a quadratic non-residue by assumption, the $\frac{p-1}{2}$ quadratic residues mod p must all be present in the pairwise sums. If two of them are in the same pairwise sum, we are done. However, if no two of them are in the same pairwise sum, then there must be one in each pairwise sum, including the $\frac{3p-1}{4} + \frac{3p-1}{4}$ sum, so we are done. □

Lemma 4.4
For all $a, b \in \mathbb{N}$, there exist $q, r \in \mathbb{Z}$ such that $a = bq + r$ and $-\frac{b}{2} < r \leq \frac{b}{2}$.

Proof. By the division algorithm, there exist $q', r' \in \mathbb{Z}$ such that $a = bq' + r'$ and $0 \leq r' < b$. If $0 \leq r' \leq \frac{b}{2}$, we are done. If $\frac{b}{2} < r' < b$, then taking $r = r' - b$ and $q = q' + 1$ concludes. □

Lemma 4.5

For all $k \in \mathbb{N}$, we have that k is a sum of four squares if and only if $2k$ is a sum of four squares.

Proof. If k is a sum of four squares, then $2k$ is a sum of four squares by Lemma 4.2, as $2 = 1^2 + 1^2 + 0^2 + 0^2$. Now, if $2k = a^2 + b^2 + c^2 + d^2$, there must be an even number of odd squares in the sum, so the sum can be split into two pairs of squares with the same parity. Without loss of generality, let the pairs be $a^2 + b^2$ and $c^2 + d^2$. Then, we have that

$$\left(\frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2 + \left(\frac{c+d}{2}\right)^2 + \left(\frac{c-d}{2}\right)^2 = \frac{2a^2 + 2b^2 + 2c^2 + 2d^2}{4} = k.$$

Since a^2 and b^2 are the same parity, it follows that $a+b$ and $a-b$ are both even; the same is true for c and d. Thus, we have that k is also expressible as the sum of four squares. □

Theorem 4.6

Let p be an odd prime. If $p \equiv 3 \pmod 4$, then there exist $a, b, c, d \in \mathbb{Z}$ such that

$$p = a^2 + b^2 + c^2 + d^2.$$

Proof. By Lemma 4.3, there exist multiples of p that are expressible as the sum of four squares. Let

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

be the smallest positive multiple of p expressible as the sum of four squares. If $n = 1$, we are done; thus, assume for the sake of contradiction that $n > 1$. Furthermore, because Lemma 4.3 constructs a multiple of p that is the sum of squares of units, and $u^2 \equiv (p-u)^2$ for all units $u \in U_p$, we have that

$$np \leq 4\left(\frac{p-1}{2}\right)^2 < 4\left(\frac{p}{2}\right)^2 = p^2,$$

so $n < p$. Now, by Lemma 4.5, we have that n must be odd, or else $np/2$ would be a smaller multiple of p. Then, applying Lemma 4.4 to a, b, c, and d, let

$$\begin{aligned} a &= a_1 n + a_2 \\ b &= b_1 n + b_2 \\ c &= c_1 n + c_2 \\ d &= d_1 n + d_2. \end{aligned}$$

Substituting into the sum of four squares, we find

$$\begin{aligned} np &= (a_1 n + a_2)^2 + (b_1 n + b_2)^2 + (c_1 n + c_2)^2 + (d_1 n + d_2)^2 \\ &= n^2(a_1^2 + b_1^2 + c_1^2 + d_1^2) + 2n(a_1 a_2 + b_1 b_2 + c_1 c_2 + d_1 d_2) + (a_2^2 + b_2^2 + c_2^2 + d_2^2). \end{aligned}$$

Then, isolating $a_2^2 + b_2^2 + c_2^2 + d_2^2$ and dividing both sides by n, we have that

$$\frac{a_2^2 + b_2^2 + c_2^2 + d_2^2}{n} = p - (a_1^2 + b_1^2 + c_1^2 + d_1^2)n - 2(a_1 a_2 + b_1 b_2 + c_1 c_2 + d_1 d_2).$$

The sum of squares $a_2^2 + b_2^2 + c_2^2 + d_2^2$ is thus divisible by n. Next, it follows from Lemma 4.4 that

$$-\frac{n}{2} < a_2, b_2, c_2, d_2 \leq \frac{n}{2}.$$

However, since n is odd, the bounds can be improved to

$$-\frac{n}{2} < -\frac{n-1}{2} \leq a_2, b_2, c_2, d_2 \leq \frac{n-1}{2} < \frac{n}{2}.$$

Squaring the inequality yields

$$0 \leq a_2^2, b_2^2, c_2^2, d_2^2 \leq \left(\frac{n-1}{2}\right)^2 < \frac{n^2}{4}.$$

Dividing by n gives

$$\frac{a_2^2 + b_2^2 + c_2^2 + d_2^2}{n} < \frac{4\left(\frac{n}{2}\right)^2}{n} = n.$$

Let $k = \left(a_2^2 + b_2^2 + c_2^2 + d_2^2\right)/n$. If $k = 0$, then each of a_2, b_2, c_2, and d_2 must be 0. Consequently, we have that n divides a, b, c, and d, so n^2 divides $a^2 + b^2 + c^2 + d^2 = np$, a contradiction to the primality of p. Therefore, assume $k > 0$. Then, multiplying k by p, we have that

$$\begin{aligned}
kp &= \frac{a_2^2 + b_2^2 + c_2^2 + d_2^2}{n} \cdot \frac{a^2 + b^2 + c^2 + d^2}{n} \\
&= \frac{(a_2a + b_2b + c_2c + d_2d)^2}{n^2} + \frac{(a_2b - b_2a + c_2d - d_2c)^2}{n^2} \\
&\quad + \frac{(a_2c - b_2d - c_2a + d_2b)^2}{n^2} + \frac{(a_2d + b_2c - c_2b - d_2a)^2}{n^2} \\
&= \left(\frac{a_2(na_1 + a_2) + b_2(nb_1 + b_2) + c_2(nc_1 + c_2) + d_2(nd_1 + d_2)}{n}\right)^2 \\
&\quad + \left(\frac{a_2(nb_1 + b_2) - b_2(na_1 + a_2) + c_2(nd_1 + d_2) - d_2(nc_1 + c_2)}{n}\right)^2 \\
&\quad + \left(\frac{a_2(nc_1 + c_2) - b_2(nd_1 + d_2) - c_2(na_1 + a_2) + d_2(nb_1 + b_2)}{n}\right)^2 \\
&\quad + \left(\frac{a_2(nd_1 + d_2) + b_2(nc_1 + c_2) - c_2(nb_1 + b_2) - d_2(na_1 + a_2)}{n}\right)^2
\end{aligned}$$

We examine each of the four squares individually.

$$\begin{aligned}
&\left(\frac{a_2(na_1 + a_2) + b_2(nb_1 + b_2) + c_2(nc_1 + c_2) + d_2(nd_1 + d_2)}{n}\right)^2 \\
&= \left(\frac{n(a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2) + (a_2^2 + b_2^2 + c_2^2 + d_2^2)}{n}\right)^2 \\
&= \left(a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 + \frac{a_2^2 + b_2^2 + c_2^2 + d_2^2}{n}\right)^2
\end{aligned}$$

$$= (a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 + k)^2$$

The first square is thus the square of an integer.

$$\left(\frac{a_2(nb_1 + b_2) - b_2(na_1 + a_2) + c_2(nd_1 + d_2) - d_2(nc_1 + c_2)}{n}\right)^2$$
$$= \left(\frac{n(a_2b_1 - b_2a_1 + c_2d_1 - d_2c_1) + a_2b_2 - b_2a_2 + c_2d_2 - d_2c_2}{n}\right)^2$$
$$= \left(\frac{n(a_2b_1 - b_2a_1 + c_2d_1 - d_2c_1)}{n}\right)^2$$
$$= (a_2b_1 - b_2a_1 + c_2d_1 - d_2c_1)^2$$

The second square is thus the square of an integer.

$$\left(\frac{a_2(nc_1 + c_2) - b_2(nd_1 + d_2) - c_2(na_1 + a_2) + d_2(nb_1 + b_2)}{n}\right)^2$$
$$= \left(\frac{n(a_2c_1 - b_2d_1 - c_2a_1 + d_2b_1) + a_2c_2 - b_2d_2 - c_2a_2 + d_2b_2}{n}\right)^2$$
$$= \left(\frac{n(a_2c_1 - b_2d_1 - c_2a_1 + d_2b_1)}{n}\right)^2$$
$$= (a_2c_1 - b_2d_1 - c_2a_1 + d_2b_1)^2$$

The third square is thus the square of an integer.

$$\left(\frac{a_2(nd_1 + d_2) + b_2(nc_1 + c_2) - c_2(nb_1 + b_2) - d_2(na_1 + a_2)}{n}\right)^2$$
$$= \left(\frac{n(a_2d_1 + b_2c_1 - c_2b_1 - d_2a_1) + a_2d_2 + b_2c_2 - c_2b_2 - d_2a_2}{n}\right)^2$$
$$= \left(\frac{n(a_2d_1 + b_2c_1 - c_2b_1 - d_2a_1)}{n}\right)^2$$
$$= (a_2d_1 + b_2c_1 - c_2b_1 - d_2a_1)^2$$

The fourth square is thus the square of an integer.

As a result, we have that kp is a smaller multiple of p that can be written as the sum of four squares, which is a contradiction. Therefore, $n > 1$ is impossible, so $n = 1$. □

We are now ready to show that all naturals can be expressed as the sum of four squares.

Proof of Theorem 4.1. Note that 1 is a sum of four squares, as

$$1 = 1^2 + 0^2 + 0^2 + 0^2.$$

Now, from Remark 2.2, we have that 2 is a sum of two squares and thus also a sum of four squares. Then, by Theorem 2.5, we have that all primes that are 1 mod 4 are sums of two squares and thus also sums of four squares. Finally, by Theorem 4.6, we have that all primes that are 3 mod 4 are sums of four squares. Consequently, all primes are expressible as the sum of four squares; since any natural greater than one is a product of primes, Lemma 4.3

concludes. □

5 r-gonal Numbers

Definition 5.1. An r-gonal number is a number that counts dots arranged in the shape of a regular r-gon. The n-th r-gon is formed by extending two adjacent sides of the previous r-gon by one dot and adding other necessary dots to form a regular r-gon.

Lemma 5.2

The nth r-gonal number is given by the formula

$$f(n,r) := \frac{n(2+(r-2)(n-1))}{2}.$$

Proof. Let $n, r \in \mathbb{N}$. To find the difference between $f(n,r)$ and $f(n-1,r)$, we count two parts: the new dots as vertices and the new dots as edges. First, for the vertices, we add $r-1$ dots because we have r dots as vertices in a r-gon and 1 dot is already from the $(n-1)$th r-gon. Next, there are $r-2$ newly added edges with each one having $n-2$ dots on the edge. So the total number of new dots added is

$$(r-1)+(n-2)(r-2) = 1+(n-1)(r-2).$$

Therefore,

$$f(n,r) = f(n-1,r)+(1+(n-1)(r-2)).$$

So we know that

$$\begin{aligned} f(n,r) &= f(r,0)+\sum_{k=0}^{n-1}(1+k(r-2)) \\ &= 0+\sum_{k=0}^{n-1}(1+k(r-2)) \\ &= \frac{n(1+(1+(n-1)(r-2)))}{2} \\ &= \frac{n(2+(n-1)(r-2))}{2}. \end{aligned}$$

□