Colorings of Pythagorean triples within colorings of the positive integers

Joshua Cooper[*], Michael Filaseta[†], Joshua Harrington, Daniel White

University of South Carolina
Department of Mathematics
1523 Greene Street
Columbia, SC 29208

December 23, 2013

Abstract

The authors investigate k-colorings of the positive integers $\leq n$ for which the triples (a, b, c), where a, b and c are positive integers with $a^2 + b^2 = c^2$ and $c \leq n$, satisfy the condition that a, b and c are colored differently. In particular, they establish for $\varepsilon > 0$ and n sufficiently large, if $k \geq \sqrt{3}^{(1+\varepsilon)\log n/\log\log n}$, then a k-coloring exists such that every triple (a, b, c) as above has a, b and c colored differently. They also consider the problem of finding a similar lower bound on the k for which a k-coloring must exist such that, for all but a small proportion of triples (a, b, c) as above, a, b and c are colored differently.

Mathematics Subject Classification: 11B75 (11B05, 11A07, 05C15)
Key Words: Pythagorean triple, coloring, covering system

1 Introduction

An ordered triple of positive integers (a, b, c) is called a *Pythagorean triple* if a, b and c satisfy the equation $a^2 + b^2 = c^2$. As suggested, with any such triple (a, b, c), we take the third component c to be larger than the other two components a and b. We also order a and b by the largest power of 2 that divides each as follows. An argument modulo 4 can be used to show that the largest power of 2 dividing a cannot equal the largest power of 2 dividing b. We adopt the convention then that in a Pythagorean triple (a, b, c), the largest power of 2 dividing b is greater than the largest power of 2 dividing a . If $\gcd(a, b, c) = 1$, then we say the Pythagorean triple is *primitive*. For fixed positive integers k and n and a fixed k-coloring of the positive integers $\leq n$, we say a Pythagorean triple (a, b, c) is monochromatic if a, b and c are colored the same and we say the triple is 2-colored (or

[*]The first author expresses his gratitude to NSF for support for this research under grant DMS-1001370.
[†]The second author expresses his gratitude to the NSA for funding for this research under grant H98230-12-1-0240.

3-colored) if exactly 2 (or none, respectively) of a, b and c are colored the same. We also note that when using the big-oh notation and Landau's $\ll$ notation in this paper, all implied constants will be absolute.

Fix an integer $n > 1$. Then one may construct $O(\log n)$-colorings of $[n] = \{1, 2, \ldots, n\}$ such that no Pythagorean triple with elements $\leq n$ is monochromatic. A simple argument is given for this coloring at the beginning of Section 4. Such a bound may be crude, as it is not even known if just 2 colors suffice to color $[n]$ so that every Pythagorean triple has this property. Cooper and Poirel [3], and later Kay [4], have in fact constructed such 2-colorings for n as large as 1344 and 1514, respectively.

The main results of this paper are as follows.

Theorem 1. *Fix $\varepsilon > 0$. Let k and n be positive integers with $k \geq \sqrt{3}^{(1+\varepsilon)\log n/\log\log n}$ and $n \geq n_0$, where $n_0 = n_0(\varepsilon)$ is sufficiently large as a function of ε. Then there exist k-colorings of $[n]$ such that every Pythagorean triple with elements $\leq n$ is 3-colored.*

Theorem 2. *Let $\xi(x)$ be an arbitrary positive increasing function that tends to infinity. Let k and n be positive integers with $k \geq c\,\xi(n)\log^2 n/\log\log n$ and $n \geq N$, where c and N are sufficiently large absolute constants. Then there exist k-colorings of $[n]$ such that the proportion of Pythagorean triples with elements $\leq n$ which are not 3-colored is $\ll \max\{2^{-\xi(n)}, 1/\log n\}$.*

Theorem 3. *Let $\xi(x)$ be any positive, increasing function that tends to infinity, and let n be sufficiently large. There exist $\lfloor \xi(n) \rfloor$-colorings of $[n]$ such that the proportion of Pythagorean triples with elements $\leq n$ which are not 3-colored is*

$$\ll \max\left\{\frac{1}{\sqrt{\log \xi(n)}}, \frac{1}{\sqrt{\log\log\log\log n}}\right\}.$$

The emphasis on this last result is that to obtain a coloring with almost all triples being 3-colored only $\lfloor \xi(n) \rfloor$ colors are needed for any function $\xi(x)$ increasing to infinity. For this, we give a simple sieve of Eratosthenes argument. A different sieve approach, using the Brun sieve or the Selberg sieve, would easily allow for an improvement on the proportion above for the number of triples which are not 3-colored.

The above results concern general Pythagorean triples. The following addresses the case of *primitive* Pythagorean triples.

Theorem 4. *There exists a 3-coloring of $\mathbb{N}$ such that, for every sufficiently large positive integer n, the proportion of primitive Pythagorean triples with elements $\leq n$ which are not 3-colored is $\ll 1/\sqrt{\log\log n}$.*

2 Preliminaries

In this section, we give some basic background. For results not specifically referenced, one can consult [1] and [6] for more details.

Throughout this paper, we make use of the following classical result of Euclid.

Theorem 5. *There exists a bijection from the ordered triples (k, s, t) of integers satisfying $k \geq 1$, $\gcd(s,t) = 1$, $s > t > 0$ and exactly one of s and t is even and the Pythagorean triples (a, b, c) given by $a = k(s^2 - t^2)$, $b = 2kst$, and $c = k(s^2 + t^2)$.*

Definition. We call the triple (k, s, t) the *representation* of the Pythagorean triple (a, b, c).

Writet $n = 2^h p_1^{e_1} p_2^{e_2} \dots p_y^{e_y} q_1^{e_{y+1}} \dots q_\ell^{e_{y+\ell}}$, where h is a non-negative integer, $e_1, \dots, e_{y+\ell}$ are positive integers, $p_1, \dots, p_y$ are distinct primes $\equiv 1 \pmod 4$ and $q_1, \dots, q_\ell$ are distinct primes $\equiv 3 \pmod 4$. We will be interested in the number of Pythagorean triples (a, b, c) for which one of a and b equals n and the number of Pythagorean triples (a, b, c) for which $c = n$. These next two results can be found in Beiler [2].

Lemma 1. *Given the prime factorization of a positive integer n as above, the number of Pythagorean triples (a, b, c) for which either $a = n$ or $b = n$ is given by*

$$P_L(n) = \frac{1}{2}\left(|2h-1| \prod_{j=1}^{y+\ell} (2e_j - 1) - 1\right).$$

Lemma 2. *Given the prime factorization of a positive integer n as above, the number of Pythagorean triples (a, b, c) for which $c = n$ is given by*

$$P_H(n) = \frac{1}{2}\left(\prod_{j=1}^{y} (2e_j - 1) - 1\right).$$

The following consequence is immediate.

Theorem 6. *The number of Pythagorean triples (a, b, c) for which one of a, b or c equals n is*

$$P(n) = P_L(n) + P_H(n) = \frac{1}{2}\left(|2h-1| \prod_{j=1}^{y+\ell} (2e_j - 1) - 1\right) + \frac{1}{2}\left(\prod_{j=1}^{y} (2e_j - 1) - 1\right).$$

We will want to make use of some elementary notions and results from graph theory. All graphs referenced in this paper are taken to be undirected and simple. The *chromatic number* of a graph G is the least number of colors required to color the vertex set $V(G)$ such that no adjacent vertices share the same color. We let $\chi(G)$ denote this number. Let $E(G)$ denote the set of edges of G, the elements of which may be represented uniquely as a two element set $\{v_i, v_j\}$, for vertices $v_i, v_j \in V(G)$. For $v \in V(G)$, the *degree of* v is the number of edges $\{v_i, v_j\} \in E(G)$ for which one of v_i and v_j is v. We denote this degree by $\deg(v)$. Define $\Delta(G)$ to be $\max_{v \in V(G)}\{\deg(v)\}$.

Lemma 3. *For any graph G as above, $\chi(G) \leq \Delta(G) + 1$.*

Proof. We color the vertices $v_1, \dots, v_n$ of G using the following algorithm.

1. Let $j = 0$.
2. Let $j \leftarrow j + 1$.
3. If $j > n$, then terminate.
4. Define $\alpha(v_j)$ to be the minimal element of $[1 + \Delta(G)] \setminus \bigcup_{(v_i, v_j) \in E(G), i<j} \{\alpha(v_i)\}$.
5. Go to step 2.

Note that $\alpha(v_j)$ is well-defined for each $j \leq n$ since no vertex has more than $\Delta(G)$ neighbors. By the definition of α, no adjacent vertices are the same color. This completes the proof. $\square$

The following result of Wigert [8] gives the maximal growth rate of the divisor function.

Theorem 7. *For m a positive integer, define $d(m)$ as the number of divisors of m. Then*

$$\limsup_{m\to\infty} \frac{\log d(m)}{\log m/\log\log m} = \log 2.$$

We will make use of covering systems when proving Theorem 2. A *covering system* (or simply a *covering*) $\mathcal{C}$ is a set of residue classes $\{r_i \pmod{m_i}\}_{i\in I}$, where I is a finite indexing set, such that every integer is contained in at least one residue class in $\mathcal{C}$. We say that a covering system $\mathcal{C}$ is *exact* if each integer resides in exactly one residue class in $\mathcal{C}$.

Let $\Omega(x)$ be the number of Pythagorean triples (a, b, c) with $c \leq x$. The following is due to Sierpiński [7].

Theorem 8. *There is a constant B such that*

$$\Omega(x) = \frac{4}{\pi} x \log x + Bx + O(x^{2/3}).$$

For primitive Pythagorean triples, we have the following corresponding result of D. H. Lehmer [5].

Theorem 9. *The number of primitive Pythagorean triples (a, b, c) with $c \leq x$ is $\sim x/(2\pi)$.*

Let s be a positive integer. For the proof of Theorem 3, we introduce a completely multiplicative function ρ_s which we define for a prime p by

$$\rho_s(p) = \begin{cases} 2 & \text{if } p \nmid s \\ 1 & \text{if } p \mid s. \end{cases}$$

In particular, we make use of the following estimate.

Lemma 4. *Let s and d be positive integers, with d square-free. The number of positive integers $t \leq s$ such that d divides $s^2 - t^2$ is*

$$\frac{\rho_s(d)s}{d} + O(\rho_s(d)).$$

Proof. The definition of $\rho_s(p)$, for a prime p, is the number of incongruent t modulo p satisfying $t \equiv \pm s \pmod{p}$. Since d is square-free, the Chinese Remainder Theorem implies that there are exactly $\rho_s(d)$ incongruent t modulo d satisfying $t \equiv \pm s \pmod{d}$. Thus, in every block of d consecutive integers, there are exactly $\rho_s(d)$ integers t for which $t \equiv \pm s \pmod{d}$. It follows that the number of positive integers $t \leq s$ such that d divides $s^2 - t^2$ is

$$\rho_s(d)\left\lfloor \frac{s}{d} \right\rfloor + O(\rho_s(d)) = \frac{\rho_s(d)s}{d} + O(\rho_s(d)),$$

completing the proof. $\square$

An indispensable tool used throughout our arguments is the Prime Number Theorem, stated below.

Theorem 10. *Let $\pi(x)$ denote the number of primes $\leq x$. Then $\pi(x) \sim x/\log x$. Equivalently,*

$$\sum_{p \leq x} \log p \sim x.$$

Here and throughout, sums over p as above are understood to be sums over primes p. That the two assertions in the above theorem are equivalent is not completely trivial, but the proof of their equivalence is elementary and typically done in proofs of the Prime Number Theorem. Another asymptotic we will use is a form of Dirichlet's theorem on primes in arithmetic progressions.

Theorem 11. *Let a and m be positive integers satisfying gcd$(a, m) = 1$. Then*

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod{m}}} \frac{1}{p} \sim \frac{\log \log x}{\varphi(m)}.$$

3 The Proof of Theorem 1

Let n be a positive integer. We define a graph G_n on n vertices v_j with $1 \leq j \leq n$ by setting the edge set to be the $\{v_i, v_j\}$ for which there exists a Pythagorean triple (a, b, c) with $c \leq n$ and with i and j distinct elements of the set $\{a, b, c\}$. Observe that $\chi(G_n)$ is the least number of colors required to color $[n]$ so that each Pythagorean triple $\leq n$ is 3-colored. From Theorem 6 and Lemma 3, we obtain

$$\begin{aligned}
\chi(G_n) &\leq 1 + \Delta(G_n) \\
&= 1 + \max_{m \in [n]} \{P(m)\} \\
&= 1 + \frac{1}{2} \max_{m \in [n]} \left\{ |2h - 1| \prod_{j=1}^{y+\ell} (2e_j - 1) + \prod_{j=1}^{y} (2e_j - 1) \right\} \qquad (1)
\end{aligned}$$

where h, y, ℓ and the e_j's are dependent on m as in Lemmas 1 and 2. The above is less than

$$\max_{m \leq n} \left\{ (2h + 2) \prod_{j=1}^{y+\ell} (2e_j + 2) \right\} = \max_{m \leq n} \left\{ 2^{\omega(m)+1} d(m) \right\} \leq \max_{m \leq n} \{2d(m)^2\}.$$

A direct application of Theorem 7 implies $\chi(G_n) \leq 4^{(1+o(1)) \log n/\log \log n}$.

This bound, however, can be improved with a revised argument. Let n be sufficiently large. For an $m \leq n$, we consider the factorization

$$m = \alpha_1^{\gamma_1} \ldots \alpha_u^{\gamma_u} \beta_1^{\gamma_{u+1}} \ldots \beta_v^{\gamma_{u+v}},$$

where the γ_i are all positive integers and the α_i and β_i are $u + v$ distinct primes satisfying

$$\alpha_i \leq \frac{\log n}{(\log \log n)^2} \quad \text{and} \quad \beta_i > \frac{\log n}{(\log \log n)^2}.$$

Note that the expressions α_i, β_i, γ_i, u and v appearing here all depend on m. From Theorem 10, the bound on α_i implies that

$$u \le \pi\left(\frac{\log n}{(\log\log n)^2}\right) \le \frac{2\log n}{(\log\log n)^3}. \tag{2}$$

The bound on β_i gives

$$\left(\frac{\log n}{(\log\log n)^2}\right)^{\gamma_{u+1}+\cdots+\gamma_{u+v}} \le \prod_{i=1}^{v}\beta_i^{\gamma_{u+i}} \le m \le n$$

$$\implies \quad \gamma_{u+1}+\cdots+\gamma_{u+v} \le \frac{(1+o(1))\log n}{\log\log n}.$$

From our previous argument, we know

$$\chi(G_n) \le 1 + \max_{m\in[n]}\left\{\prod_{j=1}^{u}(2\gamma_j-1)\prod_{j=u+1}^{u+v}(2\gamma_j-1)\right\}. \tag{3}$$

A direct computation gives $2\gamma-1 \le 3^{\gamma/2}$ for $\gamma\in\{1,2\}$. If $\gamma\ge 2$ is replaced by $\gamma+1$, then the left-hand side of this inequality increases by 2 and the right-hand side increases by $3^{\gamma/2}\big(\sqrt{3}-1\big) \ge 3\big(\sqrt{3}-1\big) > 2$. Hence, $2\gamma-1 \le \sqrt{3}^{\gamma}$ holds for every positive integer γ. We deduce then that

$$\prod_{j=u+1}^{u+v}(2\gamma_j-1) \le \sqrt{3}^{\gamma_{u+1}+\cdots+\gamma_{u+v}} \le \sqrt{3}^{(1+o(1))\log n/\log\log n}. \tag{4}$$

Now, we bound the first product in (3). For every $j\le u$, we have $2^{\gamma_j}\le n$ so that

$$2\gamma_j - 1 \le 2\log_2 n \le 3\log n \le \exp(2\log\log n).$$

From (2), we obtain

$$\begin{aligned}\prod_{j=1}^{u}(2\gamma_j-1) &\le \prod_{j=1}^{u}\exp(2\log\log n) = \exp(2u\log\log n)\\ &\le \exp\left(\frac{4\log n}{(\log\log n)^2}\right) \le 3^{4\log n/(\log\log n)^2}\end{aligned} \tag{5}$$

Since (4) and (5) hold for every $m\le n$, we obtain from (3) the desired bound on $\chi(G_n)$.

4 The Beginning of the Proof of Theorems 2 and 3

Let n be a positive integer and $\xi(x)$ be some positive, increasing function tending to infinity. Define $\nu_2(m)$ to be the exponent in the largest power of 2 dividing m. For clarification, $\nu_2(n)=h$ in the context of Lemma 1.

We will assign to each positive integer $m\le n$ an ordered pair $(u,v)=\big(u(m),v(m)\big)$. This will represent our coloring of the positive integers $\le n$. For example, if we assign the ordered pair

$(\nu_2(m), 1)$ to each integer $m \leq n$, then we see that we have a coloring that uses at most $O(\log n)$ colors. Observe that in Theorem 5, since s and t have opposite parity, we have $\nu_2(b) > \nu_2(a) = \nu_2(c)$, so this coloring of the positive integers $\leq n$ has the property that no Pythagorean triple (a, b, c) with $c \leq n$ is monochromatic.

To prove Theorems 2 and 3, we begin in a similar way by letting u be $\nu_2(m)$ if $\nu_2(m) \leq \xi(n)$ and $\xi(n) + 1$ otherwise. As we will see, Theorem 5 then implies that most Pythagorean triples (a, b, c) with $c \leq n$ have a and c colored differently from b. We turn now to clarifying what we mean by "most" here.

Observe that $u(a) = u(c) \neq u(b)$ for every Pythagorean triple (a, b, c) with $c \leq n$ and $\nu_2(c) \leq \xi(n)$. We consider now the case that $c \leq n$ and $\nu_2(c) > \xi(n)$. For the moment fix such a triple (a, b, c). Let $\alpha = \nu_2(c)$. Since $2^\alpha \leq c \leq n$, we deduce

$$\xi(n) < \alpha \leq \log_2 n.$$

Let (k, s, t) be the representation of the Pythagorean triple (a, b, c) given in Theorem 5. Since s and t have opposite parity, we deduce $2^\alpha | k$. It follows that $(k/2^\alpha, s, t)$ is the representation of the Pythagorean triple $(a/2^\alpha, b/2^\alpha, c/2^\alpha)$ where the last component $c/2^\alpha$ is at most $n/2^\alpha$. In other words, the number of Pythagorean triples (a, b, c) with $c \leq n$ and $\nu_2(c) = \alpha$ is bounded above by the number of Pythagorean triples (a, b, c) with $c \leq n/2^\alpha$. We sum over the possible values of α and apply Theorem 8 to get an upper bound on the number of Pythagorean triples (a, b, c) with $c \leq n$ and $\nu_2(c) > \xi(n)$. This upper bound is

$$\sum_{\alpha=\lceil \xi(n) \rceil}^{\lfloor \log_2 n \rfloor} \Omega\left(\frac{n}{2^\alpha}\right) \ll n \sum_{\alpha=\lceil \xi(n) \rceil}^{\lfloor \log_2 n \rfloor} \frac{\log(n/2^\alpha)}{2^\alpha} \ll \frac{n \log n}{2^{\xi(n)}}.$$

Observe that Theorem 8 now implies that the proportion of Pythagorean triples (a, b, c) with $c \leq n$ for which $u(b) = u(a)$ (or, equivalently, $u(b) = u(c)$) is $O(1/2^{\xi(n)})$. Since $\xi(n)$ is an increasing function tending to infinity, we see that this proportion tends to 0 as n tends to infinity.

The idea now is to define $v(m)$ in such a way that typically we have $v(a) \neq v(c)$ for Pythagorean triples (a, b, c) with $c \leq n$. In the next two sections, we give different values of $v(m)$ that provide us with the proofs needed for Theorems 2 and 3, respectively.

5 The Proof of Theorem 2

Fix $\varepsilon > 0$ sufficiently small; the choice $\varepsilon = 1/3$ will suffice. Set $\mathcal{S} = \left\{ \prod_{x<p\leq 2x} p : x > 2 \right\}$. The next lemma implies that every sufficiently large integer is between s and $s^{1+\varepsilon}$ for some $s \in \mathcal{S}$.

Lemma 5. *Fix $\varepsilon > 0$ as above. There exists an infinite sequence $\{s_i\}_{i=1}^\infty$ with each $s_i \in \mathcal{S}$ such that*

$$s_{i-1} < s_i < s_{i-1}^{1+\varepsilon} \quad \textit{for } i \in \{2, 3, \dots\}.$$

Proof. Let $\delta > 0$ be sufficiently small that

$$\frac{(1+\delta)^2(1+\varepsilon/2)}{(1-\delta)^2(1+\varepsilon)} < 1.$$

Let x_1 be sufficiently large, and choose

$$x_i = x_{i-1}\frac{1+\delta}{1-\delta}(1+\varepsilon/2) \quad \text{for } i \in \{2,3,\dots\}.$$

Then

$$\begin{aligned}(1+\delta)\,x_{i-1} &< (1+\delta)(1+\varepsilon/2)\,x_{i-1} = (1-\delta)\,x_i \\ &< (1+\delta)\,x_i = \frac{(1+\delta)^2}{1-\delta}(1+\varepsilon/2)\,x_{i-1} < (1-\delta)(1+\varepsilon)\,x_{i-1}.\end{aligned}$$

Hence,

$$e^{(1+\delta)x_{i-1}} < e^{(1-\delta)x_i} < e^{(1+\delta)x_i} < e^{(1-\delta)(1+\varepsilon)x_{i-1}}. \tag{6}$$

From Theorem 10, if $n = \prod_{x<p\le 2x} p \in \mathcal{S}$, then

$$\log n = \sum_{p|n} \log p = \sum_{x<p\le 2x} \log p \sim x. \tag{7}$$

Thus, if x is sufficiently large, then

$$e^{(1-\delta)x} < n < e^{(1+\delta)x}.$$

Therefore, defining

$$s_i = \prod_{x_i<p\le 2x_i} p \quad \text{for } i \in \{1,2,\dots\},$$

and recalling that x_1 is sufficiently large, we see from (6) that

$$s_{i-1} < s_i < s_{i-1}^{1+\varepsilon} \quad \text{for } i \in \{2,3,\dots\},$$

finishing the proof. □

Let n be sufficiently large, and let n_0 be the largest element of $\mathcal{S}$ that is $\le n$. Then Lemma 5 implies that $n \le n_0^{1+\varepsilon}$. Since $n_0 \in \mathcal{S}$, we can write $n_0 = \prod_{i=1}^{r} p_i$ where the prime factors are indexed so that $p_i < p_j$ for $i < j$. Since n is large, n_0 is large and each p_j is large. For each integer ℓ with $1 \le \ell \le r$, we define the following set of residue classes.

$$K_{n_0,\ell} := \left\{ \lambda \prod_{i<\ell} p_i \pmod{\prod_{i\le\ell} p_i} \;:\; \lambda \in \{1,2,\dots,p_\ell - 1\} \right\}$$

We have the following lemma.

Lemma 6. *The collection of residue classes*

$$K_{n_0} := \left(\bigcup_{1\le\ell\le r} K_{n_0,\ell} \right) \cup \{0 \pmod{n_0}\}$$

is an exact covering.

Proof. Each integer is contained in $\lambda \pmod{p_1}$ for exactly one $\lambda \in \{1, 2, \ldots, p_1 - 1\}$ except for the integers which are divisible by p_1. Each such integer divisible by p_1 is contained in $p_1\lambda \pmod{p_1p_2}$ for exactly one $\lambda \in \{1, 2, \ldots, p_2 - 1\}$ except for those divisible by p_1p_2. Continuing in this manner, we find that each integer is contained in exactly one residue class of exactly one $K_{n_0,\ell}$ except for the integers divisible by n, which are contained in $0 \pmod{n_0}$. $\square$

Now, for such a covering K_{n_0}, we define $\kappa : \mathbb{N} \to K_{n_0}$ by sending each natural number to the residue class in K_{n_0} that the natural number belongs to. Note κ is well-defined by Lemma 6. We will show that taking $v(m) = \kappa(m)$ in the previous section has the desired properties needed to complete our proof of Theorem 2.

Consider a Pythagorean triple (a, b, c) with a representation (k, s, t) that satisfies the following three conditions for some positive integer $\ell \le r$:

(i) $\prod_{i=1}^{\ell-1} p_i \mid k$,

(ii) $p_\ell \nmid k$,

(iii) $p_\ell \nmid t$.

If $\kappa(a) = \kappa(c)$, then $k(s^2 - t^2) \equiv k(s^2 + t^2) \pmod{p_\ell}$. From (ii), one obtains $p_\ell \mid t$, contradicting (iii). Hence, if a representation (k, s, t) of a Pythagorean triple (a, b, c) satisfies (i), (ii) and (iii), then $\kappa(a) \ne \kappa(c)$.

We will bound the number of Pythagorean triples (a, b, c) with $c \le n$ having a representation (k, s, t) not satisfying (i), (ii) and (iii) for all $\ell \in \{1, 2, \ldots, r\}$. First, we estimate the number of such triples that do not satisfy (i) and (ii) for all $\ell \in \{1, 2, \ldots, r\}$. By Lemma 6 or a direct argument, we deduce $n_0 \mid k$. Since (k, s, t) is a representation for (a, b, c), we have

$$a = k(s^2 - t^2), \quad b = 2kst \quad \text{and} \quad c = k(s^2 + t^2).$$

Hence, $k \le k(s^2 + t^2) = c \le n$, so that there are at most n/n_0 values of k divisible by n_0. Furthermore, $(1, s, t)$ is a representation for the primitive Pythagorean triple $\left(s^2 - t^2, 2st, s^2 + t^2\right)$, and

$$s^2 + t^2 = \frac{c}{k} \le \frac{n}{n_0}.$$

By Theorem 9, we deduce that there are at most n/n_0 such primitive Pythagorean triples and, hence, at most n/n_0 possibilities for $(1, s, t)$. We deduce then that there are at most

$$\left(\frac{n}{n_0}\right)^2 \le \left(\frac{n_0^{1+\varepsilon}}{n_0}\right)^2 = n_0^{2\varepsilon} \le n^{2/3}$$

Pythagorean triples (a, b, c) with $c \le n$ having a representation (k, s, t) with k divisible by n_0.

Suppose now that a Pythagorean triple (a, b, c) with $c \le n$ has a representation (k, s, t) with $n_0 \nmid k$. Lemma 6 implies that (i) and (ii) hold for a unique $\ell \in \{1, 2, \ldots, r\}$. We bound from above the proportion of Pythagorean triples (a, b, c) with $c \le n$ having representation (k, s, t) for which some $\ell \in \{1, 2, \ldots, r\}$ satisfies (i) and (ii) but not (iii). We will obtain a proportion that is $\ll 1/p_1$.

Lemma 7. *Fix a positive integer k, and let $A_1, A_2, \ldots, A_k$ be non-negative real numbers satisfying $\sum_{i=1}^{k} A_i \leq 1$. For arbitrary non-negative real numbers $r_1, r_2, \ldots r_k$, we have*

$$\sum_{i=1}^{k} r_i A_i \leq \max_{1 \leq i \leq k}\{r_i\}.$$

Proof. The proof of this lemma is immediate upon replacing each r_i in the last sum above with $\max_{1 \leq i \leq k}\{r_i\}$. □

Fix n as above and $\ell \in \{1, 2, \ldots, r\}$. Let $\Omega = \Omega(n)$. Let S be the set of Pythagorean triples (a, b, c) with $c \leq n$ and with the corresponding representation (k, s, t) satisfying $n_0 \nmid k$ and at least one of (i), (ii) and (iii) does not hold for each $\ell \in \{1, 2, \ldots, r\}$. Let $\Omega' = \Omega'(n)$ be the size of the set S, Ω_ℓ be the number of $(a, b, c) \in S$ which have a representation (k, s, t) satisfying (i) and (ii), and Ω'_ℓ be the number of $(a, b, c) \in S$ which have a representation (k, s, t) satisfying (i) and (ii) but not (iii). By Lemma 7, we have

$$\frac{\Omega'}{\Omega} = \sum_{\ell=1}^{r} \frac{\Omega'_\ell}{\Omega} = \sum_{\ell=1}^{r} \frac{\Omega'_\ell}{\Omega_\ell} \cdot \frac{\Omega_\ell}{\Omega} \leq \max_{1 \leq \ell \leq r}\left\{\frac{\Omega'_\ell}{\Omega_\ell}\right\}. \tag{8}$$

Thus, if we can find a uniform upper bound on Ω'_ℓ/Ω_ℓ, then this bound also serves as an upper bound for $\Omega'(n)/\Omega(n)$.

We first obtain an upper bound on Ω'_ℓ. We set $M = n/P$ where $P = P_\ell = \prod_{i=1}^{\ell-1} p_i$. The implied constants in our $\ll$ notation for our arguments in this section depend only on M being sufficiently large, so we observe that M is large since p_r is large, $p_r \mid (n_0/P)$ and $n_0/P \leq n/P = M$. Since each Pythagorean triple (a, b, c) counted by Ω'_ℓ has a representation (k, s, t) satisfying

$$c = k(s^2 + t^2) \leq n \quad \text{and} \quad P \mid k,$$

we deduce that $k = k'P$ for some positive integer k' with $k'(s^2 + t^2) \leq M$. In particular, (k, s, t) also satisfies $s^2 + t^2 \leq M$ and, hence, s and t are each $\leq \sqrt{M}$. Also, since (iii) does not hold for the representation of an (a, b, c) counted by Ω'_ℓ, we have $t = p_\ell \tau$ for some positive integer τ. Hence,

$$\begin{aligned}
\Omega'_\ell &\leq \sum_{1 \leq s \leq \sqrt{M}} \sum_{\substack{1 \leq t \leq \sqrt{M} \\ p_\ell \mid t}} \sum_{1 \leq k' \leq M/(s^2+t^2)} 1 \\
&\leq \sum_{1 \leq s \leq \sqrt{M}} \sum_{1 \leq \tau \leq \sqrt{M}} \frac{M}{s^2 + p_\ell^2 \tau^2} \\
&\leq \sum_{1 \leq \tau \leq \sqrt{M}} \sum_{1 \leq s \leq p_\ell \tau} \frac{M}{s^2 + p_\ell^2 \tau^2} + \sum_{1 \leq \tau \leq \sqrt{M}} \sum_{p_\ell \tau < s \leq \sqrt{M}} \frac{M}{s^2 + p_\ell^2 \tau^2} \\
&\leq \sum_{1 \leq \tau \leq \sqrt{M}} \sum_{1 \leq s \leq p_\ell \tau} \frac{M}{p_\ell^2 \tau^2} + \sum_{1 \leq \tau \leq \sqrt{M}} \sum_{p_\ell \tau < s < \infty} \frac{M}{s^2} \\
&\leq \sum_{1 \leq \tau \leq \sqrt{M}} \frac{M}{p_\ell \tau} + \sum_{1 \leq \tau \leq \sqrt{M}} \frac{M}{p_\ell \tau}
\end{aligned}$$

$$\leq \frac{2M}{p_\ell}\left(1+\frac{1}{2}\log M\right),$$

where the inequalities in the last two lines are justified by comparing sums to integrals. Thus,

$$\Omega'_\ell \ll \frac{M}{p_\ell}\log M. \tag{9}$$

Now we find a lower bound on Ω_ℓ. We use M, P and k' as defined in our argument for (9) above. We proceed in a similar manner, but instead consider only even values of t and odd values of s, writing $t = 2v$ and $s = 2u-1$. We also impose the conditions $1 \leq v \leq \sqrt{M/32}$ and $v < u \leq 2v$. With these restrictions, we have

$$u^2+v^2 \leq 5v^2 \leq \frac{5M}{32}. \tag{10}$$

Also, $t < s \leq 4v$ so that

$$s^2+t^2 < 32v^2 \leq M.$$

Observe that if $\gcd(2v, 2u-1) = 1$, then Theorem 5 implies that $(k'P, 2u-1, 2v)$ is a representation for some Pythagorean triple (a, b, c) counted by Ω_ℓ provided that $p_\ell \nmid k'$ and

$$1 \leq k' \leq \frac{n}{P\big((2u-1)^2+(2v)^2\big)} = \frac{M}{(2u-1)^2+(2v)^2}.$$

Since $M/\big(4(u^2+v^2)\big)$ is smaller than the right-hand side above, $(k'P, 2u-1, 2v)$ remains a representation for some Pythagorean triple (a, b, c) counted by Ω_ℓ if $p_\ell \nmid k'$ and

$$1 \leq k' \leq \frac{M}{4(u^2+v^2)}. \tag{11}$$

Recalling p_ℓ is an odd prime, the number of positive integers k' satisfying (11) divisible by p_ℓ is at most

$$\frac{M}{p_\ell\big(4(u^2+v^2)\big)} \leq \frac{M}{12(u^2+v^2)}.$$

We deduce from (10) that the number of k' satisfying (11) not divisible by p_ℓ is at least

$$\begin{aligned}\frac{M}{4(u^2+v^2)} - \frac{M}{12(u^2+v^2)} - 1 &= \frac{M}{6(u^2+v^2)} - 1 \\ &\geq \frac{M}{6(u^2+v^2)} - \frac{5M}{32(u^2+v^2)} \gg \frac{M}{u^2+v^2}.\end{aligned}$$

Therefore,

$$\begin{aligned}\Omega_\ell &\gg M \sum_{1\leq v\leq\sqrt{M/32}} \sum_{\substack{v<u\leq 2v \\ \gcd(2v,2u-1)=1}} \frac{1}{u^2+v^2} \\ &\gg M \sum_{1\leq v\leq\sqrt{M/32}} \sum_{\substack{v<u\leq 2v \\ \gcd(2v,2u-1)=1}} \frac{1}{5v^2}.\end{aligned}$$

This gives us

$$\Omega_\ell \gg M \sum_{1\le v\le\sqrt{M/32}} \frac{1}{v^2} \sum_{\substack{v<u\le 2v\\ \gcd(2v,2u-1)=1}} 1. \tag{12}$$

With some justification, we show next that we can drop the condition $\gcd(2v, 2u-1) = 1$ in this last expression. Observe that

$$\sum_{1\le v\le\sqrt{M/32}} \frac{1}{v^2} \sum_{v<u\le 2v} 1 = \sum_{1\le v\le\sqrt{M/32}} \frac{1}{v}. \tag{13}$$

On the other hand, if $\gcd(2v, 2u-1) > 1$, then v and $2u-1$ have a common odd prime factor $\le v$. Thus,

$$\begin{aligned}
\sum_{1\le v\le\sqrt{M/32}} \frac{1}{v^2} \sum_{\substack{v<u\le 2v\\ \gcd(2v,2u-1)>1}} 1 &\le \sum_{\substack{3\le q\le v\\ q \text{ prime}}} \sum_{\substack{1\le v\le\sqrt{M/32}\\ q|v}} \frac{1}{v^2} \sum_{\substack{v<u\le 2v\\ q|(2u-1)}} 1 \\
&\le \sum_{\substack{3\le q\le v\\ q \text{ prime}}} \sum_{\substack{1\le v\le\sqrt{M/32}\\ q|v}} \frac{1}{v^2}\left(\frac{v}{q}+1\right) \\
&\le \sum_{\substack{3\le q\le v\\ q \text{ prime}}} \sum_{\substack{1\le v\le\sqrt{M/32}\\ q|v}} \frac{2}{qv} \\
&\le \sum_{\substack{3\le q\le v\\ q \text{ prime}}} \sum_{1\le qv'\le\sqrt{M/32}} \frac{2}{q^2 v'} \\
&\le \sum_{\substack{3\le q\le v\\ q \text{ prime}}} \sum_{1\le v'\le\sqrt{M/32}} \frac{2}{q^2 v'} \\
&\le \Bigg(\sum_{\substack{q\ge 3\\ q \text{ prime}}} \frac{2}{q^2}\Bigg) \sum_{1\le v\le\sqrt{M/32}} \frac{1}{v}.
\end{aligned}$$

Overestimating the sum over q by allowing q to be composite gives that

$$\sum_{\substack{q\ge 3\\ q \text{ prime}}} \frac{2}{q^2} \le 2\left(\frac{\pi^2}{6} - 1 - \frac{1}{4}\right) < 0.8.$$

Combining the upper bound just obtained with the lower bound in (13) and using (12), we now deduce

$$\Omega_\ell \gg M \sum_{1\le v\le\sqrt{M/32}} \frac{1}{v} \gg M\log M,$$

by comparing the sum above to an appropriate integral.

Combining this lower bound for Ω_ℓ with the upper bound for Ω'_ℓ given in (9), we deduce that $\max_\ell(\Omega'_\ell/\Omega_\ell) \ll 1/p_1$ for every $\ell \in \{1, 2, \ldots, r\}$. Thus, from (8), we deduce

$$\Omega' \ll \Omega/p_1,$$

where the implied constant is absolute.

To complete the proof, we recall that

$$n_0 \in \mathcal{S} = \left\{ \prod_{x<p\le 2x} p : x > 2 \right\} \quad \text{and} \quad n_0 \le n \le n_0^{1+\varepsilon}.$$

Let x be such that $n_0 = \prod_{x<p\le 2x} p$. As in (7), we have $\log n_0 \sim x$. Therefore, the number of residue classes in K_{n_0} is

$$\le \sum_{p|n_0} p < (\pi(2x) - \pi(x))2x \ll \frac{x^2}{\log x} \sim \frac{(\log n_0)^2}{\log\log n_0} \ll \frac{(\log n)^2}{\log\log n}.$$

Recalling that our coloring consisted of pairs (u, v) where u is a non-negative integer $\le \xi(n) + 1$ and v is associated with residue classes in K_{n_0}, we deduce that the total number of colorings used is

$$\ll \frac{\xi(n)(\log n)^2}{\log\log n},$$

as claimed.

The proportion of the Pythagorean triples (a, b, c) with $c \le n$ that are not 3-colored is

$$\begin{aligned}
&\ll \frac{1}{2^{\xi(n)}} + \frac{n^{2/3}}{n\log n} + \frac{\Omega'(n)}{\Omega(n)} \ll \frac{1}{2^{\xi(n)}} + \frac{1}{n^{1/3}\log n} + \frac{1}{x} \\
&\ll \frac{1}{2^{\xi(n)}} + \frac{1}{n^{1/3}\log n} + \frac{1}{\log n_0} \ll \max\left\{ \frac{1}{2^{\xi(n)}}, \frac{1}{\log n} \right\}.
\end{aligned}$$

This finishes the proof of Theorem 2.

Letting $\xi(n) = \log\log n$, we have the following corollary to Theorem 2.

Corollary 1. *Let n be sufficiently large. There exist $O(\log^2 n)$-colorings of $\{1, 2, \ldots, n\}$ for which the proportion of Pythagorean triples (a, b, c) with $c \le n$ which are not 3-colored is $\ll (\log n)^{-\log 2}$.*

6 The Proof of Theorem 3

We take n sufficiently large and assign to each positive integer $m \le n$ an ordered pair (u, v). We define $u = u(m)$ as in Section 4 except with $\xi(m)$ there replaced by $\sqrt{\xi(m)}$. Thus, there are $\ll \sqrt{\xi(m)}$ different values of $u(m)$ and the proportion of Pythagorean triples (a, b, c) with $c \le n$ for which $u(b) = u(a)$ (or $u(b) = u(c)$) is $\ll 1/2^{\sqrt{\xi(n)}}$. For $v = v(m)$, we use that for a prime p congruent to 3 modulo 4, the number -1 is not a quadratic residue modulo p. The following is an easy consequence (and the proof is left to the reader).

Lemma 8. *Any number of the form s^2+t^2 with $\gcd(s,t)=1$ cannot be divisible by a prime congruent to 3 modulo 4.*

We set

$$z=\min\{\log\log\log n,\sqrt{\xi(n)}\}. \tag{14}$$

This choice of z will allow us to give a self-contained argument using the sieve of Eratosthenes. We take $v(m)$ to be the number of distinct primes $\le z$ which are congruent to 3 modulo 4 and divide m. For a Pythagorean triple (a,b,c) with $v(a)=v(c)$ and with representation (k,s,t), we deduce from Lemma 8 that each prime $\le z$ congruent to 3 modulo 4 dividing s^2-t^2 also divides k.

Let $A(s,n)$ be the number of positive integer pairs (t,k), with $t\le s$ and $k(s^2+t^2)\le n$, satisfying if p is a prime $\le z$ with $p\equiv 3 \pmod 4$ and $p\mid (s^2-t^2)$, then $p\mid k$. Note that we do not require here that $\gcd(s,t)=1$. Recall Theorem 5. As an overcount of the number of Pythagorean triples (a,b,c) where $v(a)=v(c)$, we use $\sum_{s\le\sqrt{n}}A(s,n)$. In addition, to overestimate $A(s,n)$, we instead consider (t,k) above with $ks^2\le n$ rather than $k(s^2+t^2)\le n$.

Let $\mathcal{P}$ be the product of all primes $\le z$ that are congruent to 3 modulo 4. Let $A_d(s,n)$ denote the number of positive integer pairs (t,k), with $t\le s$ and $ks^2\le n$, satisfying $d\mid (s^2-t^2)$ and $\gcd(k,d)=1$. The principle of inclusion-exclusion implies that

$$A(s,n)=\sum_{d\mid\mathcal{P}}\mu(d)A_d(s,n),$$

where μ is the Möbius function. Lemma 4 gives us

$$\begin{aligned}
A_d(s,n)&=\sum_{\substack{k\le n/s^2\\ \gcd(k,d)=1}}|\{t\le s: d\mid (s^2-t^2)\}|\\
&=\left(\frac{\rho_s(d)s}{d}+O(\rho_s(d))\right)\sum_{\substack{k\le n/s^2\\ \gcd(k,d)=1}}1\\
&=\left(\frac{\rho_s(d)s}{d}+O(\rho_s(d))\right)\left(\frac{\varphi(d)n}{ds^2}+O(\varphi(d))\right)\\
&=\frac{\rho_s(d)\varphi(d)n}{d^2s}+O\left(\rho_s(d)\varphi(d)\left(\frac{n}{ds^2}+\frac{s}{d}+1\right)\right).
\end{aligned}$$

Thus,

$$A(s,n)\le\sum_{d\mid\mathcal{P}}\mu(d)\left(\frac{\rho_s(d)\varphi(d)n}{d^2s}+O\left(\rho_s(d)\varphi(d)\left(\frac{n}{ds^2}+\frac{s}{d}+1\right)\right)\right).$$

We use that there are at most $\pi(z)\le z$ primes dividing $\mathcal{P}$. For the error term above, observe that $d\mid\mathcal{P}$ implies that $\rho_s(d)\le 2^z$ and that $d\le z^z$. Therefore,

$$\rho_s(d)\varphi(d)\left(\frac{n}{ds^2}+\frac{s}{d}+1\right)\le 2^zd\left(\frac{n}{ds^2}+\frac{s}{d}+1\right)\le\frac{n\,2^z}{s^2}+2^zs+2^zz^z.$$

Also,

$$\sum_{d|\mathcal{P}} |\mu(d)| = \sum_{d|\mathcal{P}} 1 \le 2^z.$$

Hence,

$$A(s,n) \le \left(\sum_{d|\mathcal{P}} \mu(d) \frac{\rho_s(d)\varphi(d)}{d^2} \right) \frac{n}{s} + O\left(\frac{n\, 2^{2z}}{s^2} + 2^{2z}s + 2^{2z}z^z \right). \tag{15}$$

The multiplicativity of ρ_s and φ implies that

$$\begin{aligned} \sum_{d|\mathcal{P}} \mu(d) \frac{\rho_s(d)\varphi(d)}{d^2} &= \prod_{\substack{p\le z \\ p\equiv 3 \pmod 4}} \left(1 - \frac{\rho_s(p)\varphi(p)}{p^2}\right) \\ &\le \prod_{\substack{p\le z \\ p\equiv 3 \pmod 4}} \left(1 - \frac{p-1}{p^2}\right) \\ &= \prod_{\substack{p\le z \\ p\equiv 3 \pmod 4}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p\le z \\ p\equiv 3 \pmod 4}} \left(1 + \frac{1}{p(p-1)}\right). \end{aligned}$$

The last product above approaches a constant as z tends to infinity, and Theorem 11 can be used to see that

$$\prod_{\substack{p\le z \\ p\equiv 3 \pmod 4}} \left(1 - \frac{1}{p}\right) \ll \frac{1}{\sqrt{\log z}}.$$

From (15), we obtain

$$A(s,n) \ll \frac{n}{s\sqrt{\log z}} + \frac{n\, 2^{2z}}{s^2} + 2^{2z}s + 2^{2z}z^z.$$

Using the above estimate for $A(s,n)$, our bound on the number of Pythagorean triples (a,b,c) satisfying $v(a) = v(c)$ is

$$\sum_{s\le\sqrt{n}} A(s,n) \ll \frac{n\log n}{\sqrt{\log z}} + n\, 2^{2z} + 2^{2z}z^z\sqrt{n}. \tag{16}$$

Recalling (14), we get

$$\sum_{s\le\sqrt{n}} A(s,n) \ll \frac{n\log n}{\sqrt{\log\log\log\log n}} + \frac{n\log n}{\sqrt{\log \xi(n)}}. \tag{17}$$

The needed estimates are complete. We have colored each integer $m \le n$ by a pair (u,v) of non-negative integers with $u \le \sqrt{\xi(n)}$ and v bounded by the number of primes up to $\sqrt{\xi(n)}$. For n large, there are less than $\lfloor \xi(n) \rfloor$ such pairs. The number of Pythagorean triples (a,b,c) with $c \le n$ is $\Omega(n) \sim (4/\pi) n\log n$ by Theorem 8. Since $u(b) = u(a)$ or $u(b) = u(c)$ for $\ll n\log n/2^{\sqrt{\xi(n)}}$ of these triples and since (17) provides a bound on the number of triples with $v(a) = v(c)$, Theorem 3 follows.

7 The Proof of Theorem 4

We describe the coloring of $\mathbb{N}$ that we use to establish Theorem 4. First, color all positive even integers red. Next, color each odd positive integer divisible by at least one prime congruent to 3 modulo 4 white. Color the remaining elements of $\mathbb{N}$ blue. By Lemma 8 and Theorem 5, we see that if a primitive Pythagorean triple (a, b, c) with $c \leq n$ is not 3-colored, it must be that $a = s^2 - t^2$ is odd and has no prime divisors congruent to 3 modulo 4. Call these the *bad* triples. We need only count then how many bad triples exist.

Let $z = \log n$, and define $\mathcal{P}$ as the product of primes $\leq z$ that are $\equiv 3 \pmod 4$. From Theorem 5, we see that an upper bound for the number of bad triples is

$$\sum_{s \leq \sqrt{n}} \left|\left\{\text{the number of } t \leq s \text{ such that } p \mid (s-t) \implies p \equiv 1 \pmod 4 \text{ or } p > z\right\}\right|.$$

By an inclusion-exclusion argument similar to (but easier than) the one used in the previous section, we see that

$$\begin{aligned}&\left|\left\{\text{the number of } t \leq s \text{ such that } p \mid (s-t) \implies p \equiv 1 \pmod 4 \text{ or } p > z\right\}\right| \\ &\qquad = \sum_{p \mid (2\mathcal{P})} \mu(d) \left(\frac{s}{d} + O(1)\right) = \frac{s}{2} \prod_{\substack{p \equiv 3 \pmod 4 \\ p \leq z}} \left(1 - \frac{1}{p}\right) + O\big(2^{\pi(z)}\big).\end{aligned}$$

Hence, an upper bound for the number of bad triples is

$$\sum_{s \leq \sqrt{n}} \left(\frac{s}{2} \prod_{\substack{p \equiv 3 \pmod 4 \\ p \leq z}} \left(1 - \frac{1}{p}\right) + O\big(2^{\pi(z)}\big)\right) \ll \frac{n}{\sqrt{\log \log n}}.$$

Recalling Theorem 9, we deduce Theorem 4.

References

[1] T. Apostol, Introduction to Analytic Number Theory, Undergraduate Texts in Mathematics, New York, Springer-Verlag, 1976.

[2] A. Beiler, The Eternal Triangle, Recreations in the Theory of Numbers: The Queen of Mathematics Entertains, New York, Dover, 1966.

[3] J. Cooper and C. Poirel, *Note on the pythagorean triple system*, 2008, URL `http://www.math.sc.edu/~cooper/pth.pdf`

[4] W. Kay, An Overview of the Constructive Local Lemma, Master's Thesis, University of South Carolina, 2012.

[5] D.N. Lehmer, *Asymptotic evaluation of certain totient sums*, Amer. J. Math. 22 (1900), 293–335.

[6] I. Niven, H. S. Zuckerman and H. L. Montgomery, An Introduction to the Theory of Numbers, ed. 5, New York, John Wiley & Sons, Inc., 1991.

[7] W. Sierpiński, *Sur la sommation de la série* $\sum_{a<n\leq b}\tau(n)F(n)$, *où* $\tau(n)$ *signifie le nombre de décompositions du nombre* n *en une somme de deuz carrés de nombres entiers*, Prace Mat. Fiz. 18 (1908) 1–59 (in Polish); Oeuvres Choises, Vol. 1 (Varsovie, 1974) 109–154 (in French).

[8] S. Wigert, *Sur l'ordre de grandeur du nombre des diviseurs d'un entier*, Arkiv Mat. Astr. Fys., 3 (1907), paper 18, 1–9.