The integer sides of triangles with a prime side opposite to a $\pi/3$ angle

master's thesis

Utrecht University

author:
Lennert den Besten

supervisor:
Lola Thompson

second reader:
Gunther Cornelissen

Abstract

Consider the non-equilateral triangles with integer sides, one of which is prime and lies opposite a $\pi/3$ angle. Given a positive integer it is easy to predict whether or not it can occur as the length of one of the two remaining sides of any such triangle. We will establish an asymptotic upper bound and conditionally an asymptotic lower bound for the rate of occurrence of such integers. Because it is natural to separate them, we prove results independently for odd and even integers. We expect the correct size to be a double logarithmic factor smaller than the upper bound, and a small logarithmic factor larger than the lower bound. We further expect that the results can be replicated for a wider range of triangles.

Contents

Symbol usage and other miscellaneous conventions

$O\big(g(x)\big)$ The so called big-O relation $f(x) = O\big(g(x)\big)$ between a real valued functions f and a positively real valued function g, signifies that there exists constants $C_0 > 0$ and $x_0 \in \mathbb{R}$ such that for all $x \geq x_0$, the inequality $|f(x)| \leq Cg(x)$ holds. Also often used in the obvious variation where the domain of f (and then possibly also g) is a subset of $\mathbb{R}$.

$o\big(g(x)\big)$ The little-o relation $g(x) = o\big(f(x)\big)$ between a real valued function f and a positively real valued function g, signifies that for every $\varepsilon > 0$, there exists an $x_0 \in \mathbb{R}$ such that for every $x \geq x_0$ the inequality $|f(x)| \leq \varepsilon g(x)$ holds. If $g(x)$ happens to be a function that becomes non-zero from a certain point, then the little-o relation is equivalent to saying that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

$\ll$ An alternative for big-O notation commonly used in analytic number theory. Saying that $f(x) \ll g(x)$ is equivalent to saying that $f(x) = O(g(x))$. Note that it is sometimes also used in reverse $g(x) \gg f(x)$, implying the same big-O relation as before.

$\asymp$ The first type of asymptotic equality we will use; the relation $f(x) \asymp g(x)$ between two positive real valued functions f and g means that $f \ll g \ll f$ (or equivalently that $g \ll f \ll g$).

$\sim$ The second type of asymptotic equality we will use; the relation $f(x) \sim g(x)$ between two real valued function means that the following limit holds

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

$\omega(n)$ The function that counts the number of distinct prime divisors of an integer n. So, if $n = p_1^{k_1} \cdots p_t^{k_t}$ is the prime decomposition of n, we have $\omega(n) = t$.

$\mu(n)$ The Möbius function is defined on the positive integers as follows

$$\mu(n) = \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is square-free} \\ 0 & \text{if } n \text{ is not square-free.} \end{cases}$$

p We use the letter p exclusively to denote a prime number. When used as an index under sums Σ and products Π it is used to indicate that the sum/product runs over the prime

numbers.

$\mathbb{P}$ The set containing all primes.

$P^+(n)$ The largest prime number that divides a positive integer n, with the convention that $P^+(1) = 1$.

$P^-(n)$ The smallest prime number that divides a positive integer n, with the convention that $P^-(1) = 1$.

$\pi(x)$ The function counting the number of primes $\leq x$.

$\lceil x \rceil$ The ceiling function maps a real number x to the smallest integer that is at least as large as x.

$\lfloor x \rfloor$ The floor function maps a real number x to the largest integer that is at least as small as x.

"Mertens' theorem" There seems to be some disagreement in the literature as to what statement should carry the name Mertens' theorem. There are three theorems which are up for the title, all of them proven by the Polish mathematician Franz Mertens. They are concerned with the asymptotic behaviour of one of the three functions

$$\sum_{p \leq N} \frac{1}{p} \qquad \sum_{p \leq N} \frac{\log p}{p} \qquad \prod_{p \leq N} \left(1 - \frac{1}{p}\right).$$

In tradition with Dutch culture, we will stay impartial on this issue and simply refer to all three of the facts as Mertens' theorem. To avoid any confusion, we will always write down the exact statement we are using when applying one of these theorems.

$\triangle$ Where ambiguity would arise from not doing so, we will use this symbol similarly to the qed-square to indicate the end of an example/remark/statement/etc.

1. Introduction

The triangle has been an object of study since ancient history, yet it still frequently is the subject of modern mathematical research. Probably the most well known fact about triangles is the Pythagorean theorem, a discovery popularly attributed to the ancient Greek thinker Pythagoras, even though the content of the theorem was known to members of other civilizations in even earlier history. The theorem states that the side lengths of a triangle a, b, c with a right angle opposite to c, satisfy the equation

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

and that conversely every three positive real numbers that satisfy this equation can be used to form a right triangle.

Of particular interest to number theorists are the right triangles with integer sides, any three positive integers that together form such a triangle are called a Pythagorean triple. As mathematical objects, these triples lie in the intersection of number theory and geometry. It is for example possible to prove that each coprime triple arises from the following identity

$$(x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2$$

by looking at congruences and common divisors, but another possibility is to make a correspondence to the rational points on the unit circle. This is a fairly simple fact, which among other things proves that there are infinitely many Pythagorean triples.

The triples also feature in very deep mathematics; in one of the most acclaimed pieces of mathematics written in recent history [Wile95] Andrew Wiles proved that the Pythagorean triples are the only positive integer solutions to the equation

$$a^n + b^n = c^n \qquad \text{for an } n \geq 2.$$

This statement is popularly known as Fermat's last theorem.

A recent addition to the theory of Pythagorean triples is the 2017 paper [ChPo17] by Sam Chow and Carl Pomerance. In this paper they establish asymptotic upper and lower bounds for the proportion of odd positive integers, as well as the proportion of even positive integers which occur as non-hypotenuse side of a right triangle with prime hypotenuse. The bounds are established using methods from analytic number theory, the most notable of which are Brun's sieve, the distribution of Gaussian primes in narrow circle sectors, and the Hardy-Ramanujan inequality.

In this thesis, which is based on [ChPo17], we will veer away from right triangles and Pythagorean triples, and instead look at integer sides of another class of triangles. The Pythagorean theorem is generalized by the law of cosines which states that the side lengths of a triangle a, b, c with angle γ opposite to c, satisfy the equation

$$a^2 + b^2 - 2ab\cos\gamma = c^2.$$

Besides the angle $\gamma = \pi/2$ which makes the cosine vanish and corresponds to the Pythagorean theorem, the second most standout angle is $\gamma = \pi/3$, which makes the cosine and the factor 2 cancel against each other and results in the following equation

$$a^2 + b^2 - ab = c^2. \tag{1.1}$$

The properties of the integer triples satisfying this equation and the triangles that come along with it are in so far similar to Pythagorean triples and right triangles that we are able to (largely) duplicate the results from [ChPo17] for triangles which have a prime side opposite to a $\pi/3$ angle. That we are no longer concerned with Pythagorean triples though, manifests itself in the fact that we can't make use of the theory on the distribution of Gaussian primes like in [ChPo17] and that we will have to leave a small gap in the proof of the lower bound. Like the lower bound in [ChPo17] though, we expect that our (hypothetical) lower bound can be significantly improved upon anyway. The exact statements we are going to prove are the following.

Theorem 1.1: (a) Let $\mathcal{A}(N)$ be the set containing the even positive integers $n \leq N$ for which there exists a positive integer solution to equation (1.1) with $a = n$ and c a prime. Then the following asymptotic bound holds

$$\#\mathcal{A}(N) \leq \frac{N}{(\log N)^{\eta}} (\log\log N)^{O(1)}$$

where $\eta = 1 - \frac{1+\log\log 2}{\log 2} \approx 0.086$.

(b) If the primes of the form $3x^2+y^2$ are distributed like the Gaussian primes (see conjecture 6.3 for a more precise meaning of this statement), then for every $\varepsilon > 0$, the following asymptotic inequality holds

$$\#\mathcal{A}(N) \gg \frac{N}{(\log N)^{\log 4 - 1 + \varepsilon}}.$$

(c) Let $\mathcal{B}(N)$ be the set containing the odd positive integers $n \leq N$ for which there exists a positive integer solution to equation (1.1) with $b = n$ and c a prime. Then the following asymptotic bound holds

$$\#\mathcal{B}(N) \leq \frac{N}{(\log N)^{\eta}} (\log\log N)^{O(1)}.$$

(d) If the primes of the form $3x^2 + y^2$ are distributed like the Gaussian primes, then for every $\varepsilon > 0$ the following asymptotic inequality holds

$$\#\mathcal{B}(N) \gg \frac{N}{(\log N)^{\log 4 - 1 + \varepsilon}}.$$

△

In our writing, we will assume familiarity with some of the more common results in analytic number theory. In particular the prime number theorem (for arithmetic progressions) and Mertens' theorem(s) will be used frequently. We will spend the upcoming two chapters introducing some of the more advanced tools that are required to establish the bounds which form the main results of this thesis. The first chapter in which we will prove the Hardy-Ramanujan inequality is fairly short and simple. The second is considerably longer; in it we will introduce some sieve theory and prove a weak version of Brun's sieve. Throughout the thesis we will sometimes refer to the appendix to avoid having too many tedious deductions in the main body. There will also be a few instances where we simply call upon existing literature to provide us with the statements we need.

In the final chapter we will say a few words about possible further generalizations and improvements of the results.

2. The Hardy-Ramanujan inequality

In this short preliminary chapter we reproduce a result from the paper [HaRa17] which establishes the Hardy-Ramanujan theorem, which states that the normal order of $\omega(n)$ (see page 3) is $\log\log n$. Of particular use to us will be the following lemma (lemma B on page 9), which is only a small sub-result of the paper.

Lemma 2.1: (The Hardy-Ramanujan inequality) Uniformly for $i \in \mathbb{Z}_{\geq 1}$ and $N \geq 3$,

$$\overline{\omega}_i(N) := \{n \leq N \,|\, \omega(n) = i\} \ll \frac{N}{\log N} \cdot \frac{(\log\log N + c_0)^{i-1}}{(i-1)!}$$

where $c_0 > 0$ is a constant that is independent of i and N.

Proof: We proceed by induction on i. For the base case $i = 1$, we first establish that

$$\begin{aligned}\overline{\omega}_1(N) &= \#\{p \leq N \,|\, p \text{ prime}\} + \#\{p^2 \leq N \,|\, p \text{ prime}\} + \#\{p^3 \leq N \,|\, p \text{ prime}\} + \dots \\ &= \pi(N) + \pi(N^{1/2}) + \pi(N^{1/3}) + \dots\end{aligned}$$

Note that this series is finite, as $\pi(N^{1/k}) = 0$ for $k > \log_2 N$. Using the prime number theorem, we can further bound it as

$$\begin{aligned}\omega_1(N) &\ll \frac{N}{\log N} + \frac{N^{1/2}}{\log(N^{1/2})} + \dots + \frac{N^{1/\lceil \log_2 N\rceil}}{\log(N^{1/\lceil \log_2 N\rceil})} \\ &\leq \frac{1}{\log N}(N + N^{1/2} + \dots + N^{1/\lceil \log_2 N\rceil}) \\ &\leq \frac{1}{\log N}(N + N^{1/2}\log_2 N) \\ &\ll \frac{N}{\log N}.\end{aligned}$$

This confirms that base step of the induction follows for any choice of c_0.

We continue with the induction step and assume that the inequality from the statement holds for a certain $i \geq 1$. How we have to choose c_0 will become clear at the end of the argument. For p prime and k a positive integer, we define $A(p,k)$ to be the set containing all numbers $n \leq N$ that can be written as a product of primes $p^k p_1^{k_1} \dots p_i^{k_i}$ where the numbered primes are strictly increasing in size (i.e., $p_1 < \dots < p_i$) and the largest of which satisfies $p_i \geq p$. By construction, the inequality

$$\#A(p,k) \leq \overline{\omega}_i\left(\frac{N}{p^k}\right)$$

holds. If $q_1^{\ell_1} \dots q_{i+1}^{\ell_{i+1}}$ is the prime factorization of a number counted by $\overline{\omega}_{i+1}(N)$, then it is contained in the sets $A(q_1, \ell_1), \dots, A(q_i, \ell_i)$. By additionally observing that $q_j^{\ell_j+1} \leq N$ holds for $j = 1, \dots, i$, we see that the following inequality holds

$$i\overline{\omega}_{i+1}(N) \leq \sum_{p^{k+1} \leq N} \#A(p,k) \leq \sum_{p^{k+1} \leq N} \overline{\omega}_i\left(\frac{N}{p^k}\right).$$

By dividing by i and applying the induction hypothesis we find the following

$$\begin{aligned}\overline{\omega}_{i+1}(N) &\ll \frac{1}{i} \sum_{p^{k+1} \leq N} \frac{N/p^k}{\log(N/p^k)} \cdot \frac{(\log\log(N/p^k) + c_0)^{i-1}}{(i-1)!} \\ &\leq \frac{N(\log\log N + c_0)^{i-1}}{i!} \sum_{p^{k+1} \leq N} \frac{1}{p^k \log(N/p^k)}.\end{aligned}$$

Note that the constant associated to the $\ll$ can be taken the same as in the base step. So, we can both wrap up the induction, and ensure uniformity by proving that the inequality

$$\sum_{p^{k+1} \leq N} \frac{1}{p^k \log(N/p^k)} \leq \frac{\log\log N + c_0}{\log N} \tag{2.1}$$

holds for an appropriately chosen c_0.

To be able to confirm this inequality we want to use constants that arise in two well known inequalities which are both due to Mertens. The first is a constant $M_1 > 0$ such that

$$\sum_{p \leq N} \frac{1}{p} < \log\log N + M_1.$$

The second is a constant $M_2 > 0$ such that

$$\sum_{p \leq N} \frac{\log p}{p} < M_2 \log N.$$

We now approximate (2.1) in two steps. We start by looking at the first terms of the sum where $k + 1 = 2$. When $0 \leq x \leq 1/2$, the geometric series $1/(1-x) = 1 + x + x^2 + \dots$ is no

larger than $1+2x$. Using this fact we can bound the first terms as follows

$$\begin{aligned}\sum_{p^2\le N}\frac{1}{p\log(N/p)} &= \sum_{p^2\le N}\frac{1}{p\log N}\cdot\frac{1}{1-(\log p/\log N)}\\ &\le \sum_{p^2\le N}\frac{1}{p\log N}\left(1+\frac{2\log p}{\log N}\right)\\ &\le \frac{1}{\log N}\sum_{p^2\le N}\frac{1}{p}+\frac{2}{(\log N)^2}\sum_{p^2\le N}\frac{\log p}{p}\\ &\le \frac{\log\log N+M_1+M_2}{\log N}.\end{aligned}$$

The second step is to look at the remaining terms of (2.1) when $k+1>2$. Note that the condition $p^{k+1}\le N$ holds iff $p^{k+1}N^{1/k}\le N^{(k+1)/k}$ which holds iff $N^{1/(k+1)}\le N/p^k$, we can therefore bound the remaining terms as follows

$$\begin{aligned}\sum_{\substack{p^{k+1}\le N\\ k\neq 1}}\frac{1}{p^k\log(N/p^k)} &\le \sum_{\substack{p^{k+1}\le N\\ k\neq 1}}\frac{1}{p^k\log(N^{1/(k+1)})}\\ &= \frac{1}{\log N}\sum_{\substack{p^{k+1}\le N\\ k\neq 1}}\frac{k+1}{p^k}\\ &\le \frac{1}{\log N}\sum_{k=2}^{\infty}\sum_{m=2}^{\infty}\frac{k+1}{m^k}\\ &\le \frac{3}{\log N}\sum_{m=2}^{\infty}\sum_{k=2}^{\infty}\frac{k-1}{m^k}\\ &\le \frac{3}{\log N}\sum_{m=2}^{\infty}\frac{1}{(m-1)^2}\\ &= \frac{3\zeta(2)}{\log N}\end{aligned}$$

Combining this with the inequality we established in the first step, we conclude that the inequality (2.1) holds if we choose $c_0\ge M_1+M_2+3\zeta(2)$.

□

3. Brun's Sieve

The goal of this preliminary chapter is to establish a weak version of a theorem called Brun's sieve. It will be an important tool that we will utilise multiple times in the main part of this thesis. The Norwegian Viggo Brun is the founder of modern sieve theory. Early in the twentieth century he was able to use his sieve to establish meaningful results for problems which had hitherto been unapproachable, in particular problems related to the twin prime conjecture, a proof of which was his original goal. More than a hundred years and many ideas later, sieve theory has grown into a versatile and indispensable tool for the analytic number theorist.

We start by examining the sieve of Eratosthenes and predominantly carry out ideas that don't quite work as one would hope. Trying to avoid the problems we will run into motivates the alternative approach Brun took. Before we can even properly state the theorem and present its proof, we also need to introduce a lot of notation. To help us motivate and explain our choices we will work through some examples as well. We loosely follow chapter 1 and the first part of chapter 6 of the book [OdC10].

It was in ancient Greece that Eratosthenes created the first mathematical sieve. Around 200 BC he conceived of a quick and simple way to find, given a number n, what the prime numbers up to n are. The easiest way to explain his method is through an example, let's take $n = 30$. The first step is to write down all the integers up to 30. The next step is to go through this list and cross out all the multiples of 2 excluding 2 itself. We then iteratively repeat this general recipe, we take the next number on the list that's not crossed out and then cross out all it's multiples excluding the number itself. Once we reach an integer that is greater than $\sqrt{n}$ we can stop, any composite number $\leq n$ must have a divisor in the numbers up to and including $\sqrt{n}$. For $n = 30$ we end up with the following table

1	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~

Ignoring 1, the numbers that remain are exactly the primes. Note that we crossed out the multiples of 2 with a /, the multiples of 3 with a \, and the multiples of 5 with a $-$. This helps us explain how we can make this quantitative, i.e. how we can use this to compute $\pi(n)$ (see page 3). We started out with 29 numbers, we then crossed out $\lfloor \frac{30}{2} \rfloor - 1 = 14$ numbers with a /, crossed out $\lfloor \frac{30}{3} \rfloor - 1 = 9$ numbers with a \ and $\lfloor \frac{30}{5} \rfloor - 1 = 5$ numbers with a $-$. If we now compute $29 - 14 - 9 - 5 = 1$ we end up underestimating $\pi(n)$, this is caused by the fact that we crossed out some numbers more than once. To compensate we count the number of double crossing outs and add them back. There are $\lfloor \frac{30}{2 \cdot 3} \rfloor = 5$ double $\times$ crossings, there are $\lfloor \frac{30}{2 \cdot 5} \rfloor = 3$

double $\not$ crossings, and finally $\lfloor \frac{30}{3\cdot 5} \rfloor = 2$ double $\diagdown$ crossings. Computing $1+5+3+2=11$ now gives us an overestimation of $\pi(n)$ by one. This is because 30 can be crossed out twice in two different ways, but thrice in only one way. What we are doing here is called the principle of inclusion-exclusion. Generalizing this idea for an arbitrary n gives us the following formula

$$\pi(n) = \pi(\sqrt{n}) + (n-1) - \sum_{p \leq \sqrt{n}} \left\lfloor \frac{n}{p} \right\rfloor + \sum_{p_1 < p_2 \leq \sqrt{n}} \left\lfloor \frac{n}{p_1 p_2} \right\rfloor - \sum_{p_1 < p_2 < p_3 \leq \sqrt{n}} \left\lfloor \frac{n}{p_1 p_2 p_3} \right\rfloor + \dots$$

Note that it becomes clear here why we have to use the floor function, that 30 is divisible by all primes $\leq \sqrt{30}$ is just a nice coincidence. Moving the $\pi(\sqrt{n}) - 1$ to the left hand side and using the well known Möbius function μ (see page 3), we can write this succinctly as

$$\pi(n) - \pi(\sqrt{n}) + 1 = \sum_{\substack{d \\ p \mid d \Rightarrow p \leq \sqrt{n}}} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor$$

Note that moving $\pi(\sqrt{n}) - 1$ translates to counting 1 and including the primes themselves in the crossing out of their multiples, so for our example this just means we would have crossed out 2,3 and 5 as well.

An interesting question to explore is whether or not we can use this formula to give an accurate approximation of $\pi(n)$. A natural first step is to use $\lfloor \frac{n}{d} \rfloor = \frac{n}{d} - \{\frac{n}{d}\}$ and absorb the sum of the fractional parts in an error term R of yet unknown size. The remaining main-term is a well-known (finite) Dirichlet series which can be rewritten as a (finite) Euler product

$$\sum_{\substack{d \\ p \mid d \Rightarrow p \leq \sqrt{n}}} \mu(d) \left\lfloor \frac{n}{d} \right\rfloor = n \sum_{\substack{d \\ p \mid d \Rightarrow p \leq \sqrt{n}}} \frac{\mu(d)}{d} + R = n \prod_{p \leq \sqrt{n}} \left(1 - \frac{1}{p}\right) + R$$

Mertens' theorem tells us that the product asymptotically equals $e^{-\gamma}/\log(\sqrt{n})$, where $\gamma \approx 0.577$ is Euler's constant. Likewise, we can use the prime number theorem to approximate $\pi(\sqrt{n})$. We end up with the following approximation

$$\pi(n) = 2e^{-\gamma} \frac{n}{\log n} + R + O\left(\frac{\sqrt{n}}{\log n}\right)$$

We know from the prime number theorem though that $\pi(n)$ is asymptotically equal to $n/\log n$. This basically tells us that the constant we found for our main term is wrong, and correspondingly that R which we dubbed as error term is actually the same order of magnitude as the main term, which is a problem. No matter how well we estimate R, it will never be negligible compared to the main term. Taking another look at this error term

$$R = -\sum_{\substack{d \\ p \mid d \Rightarrow p \leq \sqrt{n}}} \mu(d) \left\{ \frac{n}{d} \right\} \tag{3.1}$$

we actually see that by crudely estimating its terms with 1, and hence estimating the sum by its number of terms, we end up with a gross overestimation of its size. Each prime $p \leq \sqrt{n}$ can either divide or not divide d, so the sum has $2^{\pi(\sqrt{n})}$ terms, which is almost exponential in our main term. Controlling error terms like these represents a fundamental problem of sieves. It also represents the fact that sieves are typically not used for their accuracy, the strength of the sieve tends to lie in its the ability to handle more general questions like 'how many primes are there of the form $P(x)$?' where P is some polynomial.

Brun was the first who thought of generalizing the above approach and modifying it in a way that produces error terms that we can control. As mentioned earlier, we will introduce a lot of notation before we discuss his methods. This notation is standard for stating so called sieve problems. We will look back at how the sieve of Eratosthenes translates into a sieve problem, as well as look at a few other examples that will help explain how more general problems can be described.

The first thing we do is fix two positive real numbers $z < x$ which will act respectively as the upper bound for the primes we are going to sieve with and the upper bound for the numbers we are going to sieve for. We define a sequence

$$\mathcal{A} := (a_m)_{m \leq x} \qquad a_m \in \mathbb{R}_{\geq 0}$$

whose entries will act as weights, and which is usually just defined as the characteristic function of a set of natural numbers. Next, we let $\mathcal{P}$ be the set containing the primes we want to sieve with, and define the product

$$P(z) := \prod_{\substack{p<z \\ p \in \mathcal{P}}} p.$$

This enables us to define our object of interest, the *sifting function*

$$S(\mathcal{A}, z) := \sum_{\substack{m \leq x \\ \gcd(m, P(z))=1}} a_m, \tag{3.2}$$

which sifts out the a_m which we are not interested in. The idea is that appropriately choosing $\mathcal{A}, \mathcal{P}$ and z (which we will typically take as function of variable x) will make the sifting function $S(\mathcal{A}, z)$ an interesting quantity. We can for example define the sieve of Eratosthenes as follows, take $\mathcal{A}$ to be 1 everywhere, $\mathcal{P}$ to be the set of all primes and take $z = \sqrt{x}$. The condition $\gcd(m, P(z)) = 1$ then acts as a detector of sufficiently large primes and the sifting function equals $\pi(x) - \pi(\sqrt{x})$. We will have a look at more illuminating examples in a moment, but first we have more to say about the sifting function.

The following well known property of the Möbius function

$$\sum_{d|m} \mu(d) = \begin{cases} 1 & \text{if } m = 1 \\ 0 & \text{else} \end{cases}$$

allows us to reformulate the condition on the gcd as follows

$$S(\mathcal{A}, z) = \sum_{\substack{m \leq x \\ \gcd(m, P(z))=1}} a_m = \sum_{m \leq x} \sum_{d \mid \gcd(m, P(z))} \mu(d) a_m = \sum_{d \mid P(z)} \mu(d) \sum_{\substack{m \leq x \\ m \equiv 0 (d)}} a_m.$$

What we are doing here is more or less the same as what we did above in the sieve of Eratosthenes example, but now in a more general fashion. We want to further examine the inner sums, which we denote as

$$A_d(x) := \sum_{\substack{m \leq x \\ m \equiv 0 (d)}} a_m.$$

In the example we would have $A_d(x) = \frac{x}{d} + O(1)$ with a small error term of size not greater than 2. So, we are able to write $A_d(x)$ as the sum of a nice function and some small error term. This is a common occurrence which we introduce notation for. In general we assume that we can approximate these sums as follows

$$A_d(x) = g(d)X + r_d(\mathcal{A})$$

where $X \asymp A_1(x) = \sum_{m \leq x} a_m$ is a convenient approximation of the total weight of the sequence $\mathcal{A}$, where the *density function* $g : \mathbb{Z}_{\geq 1} \to [0, 1)$ is assumed to be multiplicative and $r_d(\mathcal{A})$ is an error term. In the example, the obvious choice would be

$$X = x, \qquad g(d) = \frac{1}{d}, \qquad r_d(\mathcal{A}) = 2.$$

What we more or less assume, is that we can closely describe the distribution of the total weight over the congruence classes with the density function g. The values of this density function can be interpreted as the probability of a member of the sequence $\mathcal{A}$ being divisible by d. The assumption for it to be multiplicative reflects that this is an independent event for two coprime numbers.

Note that all this is defined very loosely. Especially this last piece of notation is meant to be suggestive rather than restrictive, we are merely agreeing on a framework in which we can nicely phrase and study sieve problems.

A pair of examples will demonstrate how this notation comes together, and shed some more light on the strengths and shortcomings of sieves.

Example 3.1: A famous unsolved problem in number theory asks how often the polynomial $m^2 + 1$ equals a prime. It is not even known whether this happens infinitely often. With the above notation we can rephrase this as a sieve problem.

Let $\mathcal{A}$ be the characteristic function of the set $\{m^2 + 1 \leq x\}$. As set of primes we can take $\mathcal{P} = \{p \text{ prime} \mid p \not\equiv 3 \pmod 4\}$ because $m^2 + 1$ is never divisible by a prime that is $3 \pmod 4$.

We are looking for primes, so we take $z = \sqrt{x}$ like before. The sifting function $S(\mathcal{A}, z)$ is then equal to the number of primes of the form $m^2 + 1$ in the interval $(z, x]$. The total weight of the sequence is

$$A_1(x) = \lfloor \sqrt{x - 1} \rfloor$$

which we can approximate nicely by taking $X = \sqrt{x}$. For p prime, the size of the sets

$$A_p(x) = \#\{m^2 + 1 \leq x \,|\, m^2 \equiv -1 \pmod{p}\}$$

depends on whether or not -1 is a quadratic residue mod p. This is only happens when $p \not\equiv 3 \pmod 4$. In that case there are two solutions contained in each residue class for odd p and one per residue class when $p = 2$. The Chinese remainder theorem generalizes this to squarefree d. This motivates defining $g(d)$ by putting

$$\begin{aligned} g(2) &= \frac{1}{2} \\ g(p) &= \frac{2}{p} \quad \text{for } p \equiv 1 \pmod 4, \end{aligned}$$

which uniquely defines a multiplicative function supported on the squarefree numbers without prime divisors equal to 3 mod 4. It is good to note that we could have equivalently taken $\mathcal{P}$ to be the set of all primes and additionally put $g(p) = 0$ for the primes 3 mod 4. We have as approximation

$$A_d(x) = g(d)X + r_d(\mathcal{A})$$

with an error term $r_d(\mathcal{A})$ which is bounded absolutely by $2^{\omega(d)}$ where $\omega(d)$ counts the number of distinct prime factors of d.

Producing a lower bound for $S(\mathcal{A}, z)$ that grows to infinity would prove that there are infinitely many primes of the form $m^2 + 1$. As mentioned before, current methods have not proven to be powerful enough to generate such a bound. With Brun's sieve though, it is possible to produce the non-trivial upper bound $O(\sqrt{x}/\log x)$.

By slightly changing the approach or widening the search to include more numbers than just primes, one is often able to derive sharper results. See for example [OdC10] example 1.6, where an asymptotic formula for the number of squarefree integers of the form $m^2 + 1$ follows from a moderate amount of effort.

Example 3.2: Another famous unsolved problem is the twin prime conjecture, which proposes that there are infinitely many primes p such that $p + 2$ is prime as well. We can also rephrase this as a sieve problem.

If two primes form a twin pair, their product contains exactly two prime divisors. Therefore, one way to set up our sieve is such that it excludes the numbers that have more than two prime divisors from the set $\{m(m + 2) \leq x\}$. We take $\mathcal{A}$ as the characteristic function of this set.

There is no class of primes we can exclude beforehand, so we take $\mathcal{P} = \mathbb{P} = \{\text{all primes}\}$. If one of the factors m or $m+2$ is composite it must have a prime divisor of size at most $x^{1/4}$. So if we take $z = x^{1/4}$, the sifting function $S(\mathcal{A}, z)$ is equal to the number of twin prime pairs in the interval $(z, x]$. Because $m(m+2) = (m+1)^2 - 1$, the total weight is equal to the number of squares which aren't greater than $x+1$, i.e.

$$A_1(x) = \lfloor \sqrt{x+1} \rfloor.$$

This approximates nicely as $X = \sqrt{x}$. The size of $A_d(x)$ directly depends on the number of solutions to the equation $m(m+2) \equiv 0 \pmod{d}$. Because d is squarefree, the Chinese remainder theorem reduces this to the case where d is prime. For odd primes there are two distinct solutions and for 2 there is only one. We let $g(d)$ be the unique multiplicative function supported on squarefree numbers defined by

$$\begin{aligned} g(2) &= \frac{1}{2} \\ g(p) &= \frac{2}{p} \quad \text{for } p \text{ odd.} \end{aligned}$$

Then the approximation

$$A_d(x) = g(d)X + r_d(\mathcal{A})$$

has an error term that is of size at most $|r_d(\mathcal{A})| \leq 2^{\omega(d)}$.

As in the previous example, there is no known method that produces a lower bound that is strong enough to confirm that $S(\mathcal{A}, z)$ grows infinitely large. Like before though, Brun's sieve produces a non-trivial upper bound, now of size $O(x/(\log x)^2)$. At the end of the chapter we will establish the slightly weaker bound $O\big(x(\log\log x/\log x)^2\big)$ using the likewise weaker version of Brun's sieve which we are working towards.

△

The first problem we ran into when trying to approximate $\pi(n)$ with the sieve of Eratosthenes, was that the error term (3.1) is of the same order of size as the main term. This prevents the main term from producing an accurate asymptotic formula. However, (3.1) is in theory still small enough to produce correct upper and lower bounds. While the prime number theorem is of course a lot more powerful for counting primes, and makes the sieve of Eratosthenes useless in this regard, no such theorems exist in current mathematics for problems like in the above examples.

The other problem we noticed is the fact that the approach we took to estimate (3.1) greatly overestimates its size. A big cause of this problem is that the number of terms that are being summed is enormous compared to the size of the main term. In the above two examples the situation isn't particularly better. Remember that in general the complete error term equals

$$S(\mathcal{A}, z) - X \sum_{p \mid P(z)} \mu(d) g(d) = \sum_{d \mid P(z)} \mu(d) r_d(\mathcal{A}).$$

In the examples this sum contains roughly the same number of terms as we had with the sieve of Eratosthenes (in example 1, $\mathcal{P}$ only contains about half the primes) while the terms themselves are typically larger, the bound $2^{\omega(d)}$ is even a lot worse than before. So the obvious approach of trying reduce the error term by improving the approximations seems hopeless. In reality, we don't really run into bounds better than $O(1)$ for the terms.

Brun succeeded to control the error term by approaching the main term with a truncated version in which only a limited number of steps of inclusion/exclusion have been carried out. What we will concretely do is compare the sifting function with its approximation by comparing them both with the following truncated version of the approximation

$$X \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)g(d)$$

where $r > 1$ is an integer.

We start with comparing the sifting function with its truncated version

$$\sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)A_d(x) \tag{3.3}$$

Like we saw in the numerical example of the sieve of Eratosthenes, this gives either an overestimation or underestimation of $S(\mathcal{A}, z)$ depending on the parity of r. That this is true in this generality isn't directly clear. A straightforward way to prove this is through the identity

$$\sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d) = (-1)^{r+1}\binom{\omega(n)-1}{r-1}.$$

It is however also an easy consequence of the following formula, which we need anyway to help us quantify how far off (3.3) is from the complete sum.

Lemma 3.3: (Buchstab formula) For an integer d, define $\mathcal{A}_d$ to be the sequence which is equal to $\mathcal{A}$ on multiples of d and 0 elsewhere. The following equality holds

$$S(\mathcal{A}, z) = A_1(x) - \sum_{p|P(z)} S(\mathcal{A}_p, p)$$

Proof: Comparing the right hand side to the original definition of $S(\mathcal{A}, z)$, equation (3.2), we see that we have to confirm that the sum which is subtracted from $A_1(x)$ exactly deletes the a_m for which m has a common divisor with $P(z)$. Fix such an m and let p be its smallest

prime divisor. For primes $p' < p$, we have that a_m is not featured in the sequence $\mathcal{A}_{p'}$ and hence is not a summand of $S(\mathcal{A}_{p'}, p')$.

Because $p|m$ we have that a_m is a member of the sequence $\mathcal{A}_p$. The product $P(p)$ by definition factors into the primes smaller than p, so it has no common divisor with m. Hence a_m is a summand of $S(\mathcal{A}_p, p)$.

For primes $p' > p$, we have that p is a common divisor of m and $P(p')$, which implies that a_m is not a summand of $S(\mathcal{A}_{p'}, p')$.

For an m that doesn't have a common divisor with $P(z)$, it follows tautologically that m also doesn't have a common divisor with any of the primes dividing $P(z)$. Therefore a_m is never a term of the $S(\mathcal{A}_p, p)$. □

The reason why this is a useful formula is that $S(\mathcal{A}, z)$ becomes easier to understand when z gets smaller, as there will be fewer terms to exclude. Basically, this formula lets us subdivide our problem into easier parts. By repeatedly applying it, we obtain the following.

Corollary 3.4: We use $P^-(d)$ to denote the smallest prime divisor of an integer d. The following equality holds

$$S(\mathcal{A}, z) = \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)A_d(x) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} S\big(\mathcal{A}_d, P^-(d)\big). \tag{3.4}$$

Proof: We use induction on r. The base case $r = 1$ is exactly the Buchstab formula. Assuming the equality holds for an r, we apply the Buchstab formula to find that

$$\begin{aligned}
S(\mathcal{A}, z) &= \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)A_d(x) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} S\big(\mathcal{A}_d, P^-(d)\big) \\
&= \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)A_d(x) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} \Big(A_d(x) - \sum_{p|P(P^-(d))} S(\mathcal{A}_{dp}, p)\Big) \\
&= \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)A_d(x) + \sum_{\substack{d|P(z) \\ \omega(d)=r}} \mu(d)A_d(x) + (-1)^{r+1} \sum_{\substack{d|P(z) \\ \omega(d)=r}} \sum_{p|P(P^-(d))} S(\mathcal{A}_{dp}, p) \\
&= \sum_{\substack{d|P(z) \\ \omega(d)<r+1}} \mu(d)A_d(x) + (-1)^{r+1} \sum_{\substack{d|P(z) \\ \omega(d)=r+1}} S\big(\mathcal{A}_d, P^-(d)\big)
\end{aligned}$$

To see that equality of the rightmost sums in the last step holds it is best to fix an a_m and compare how often it is a summand in both sums. The expression $S(\mathcal{A}_{dp}, p)$ sums a_m exactly once iff $P^-(m) = p$ and $d|m$. So the coefficient of a_m in the double sum is equal to the number

of divisors $d|P(z)$ for which $d|m$ and $P^-(m) < P^-(d)$ and $\omega(d) = r$. Similarly, the expression $S\big(\mathcal{A}_d, P^-(d)\big)$ sums a_m exactly once iff $P^-(d) = P^-(m)$ and $d|m$. So the coefficient of a_m in final sum in the equation is equal to the number of divisors $d|P(z)$ for which $d|m$ and $P^-(d) = P^-(m)$ and $\omega(d) = r + 1$. □

Because the quantities $S\big(\mathcal{A}_d, P^-(d)\big)$ are all positive, (3.3) gives an upper bound for $S(\mathcal{A}, z)$ when r is odd and a lower bound when r is even.

We want to expand the sifting function one step further than (3.4) and replace the A_d with their approximations $\mu(d)g(d) + r_d(\mathcal{A})$. Doing this gives us

$$S(\mathcal{A}, z) = X \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)g(d) + E_1 + E_2$$

where E_1 and E_2 are defined as

$$E_1 := (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} S\big(\mathcal{A}_d, P^-(d)\big) \quad \text{and} \quad E_2 := \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)r_d(\mathcal{A})$$

and are error terms which we will deal with later. Like we wanted, we now have a comparison of the sifting function with a truncated version of its approximation.

By only slightly adapting the proof of the Buchstab formula we can similarly expand the complete approximation of the sifting function $XV(z)$, where

$$V(z) := \sum_{d|P(z)} \mu(d)g(d)$$

Repeating the arguments we now obtain as Buchstab formula

$$V(z) = 1 - \sum_{\substack{p|P(z) \\ d|P(p)}} \mu(d)g(d)g(p) = 1 - \sum_{p|P(z)} g(p)V(p)$$

and repeated application of this now gives

$$V(z) = \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)g(d) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d)V\big(P^-(d)\big).$$

As promised, we use this to compare the sifting function with its approximation, subtracting the two yields

$$S(\mathcal{A}, z) - XV(z) = E_1 + E_2 + E_3$$

where the third error term equals

$$E_3 = (-1)^{r+1} X \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d) V\big(P^-(d)\big).$$

The error terms we are left with seem a lot more reasonable than what our naïve approach gave us before. The expressions for E_1 and E_3 look somewhat similar; we will be able to relate both of them to something roughly like $XV(z)$. The term E_2 will be collected in a larger application-dependent error term together with some similar terms that we have to shave off E_1 along the way. What we are able to prove is the following.

Theorem 3.5: (Brun's sieve weak version) Let c be the solution to the equation $(c/e)^c = e$. Let $z \geq 2$ and $r \geq c|\log V(z)|$, then

$$|S(\mathcal{A}, z) - XV(z)| \leq e^{-r-1} V(z)^{1-c} X + R(\mathcal{A}, z^r)$$

where

$$R(\mathcal{A}, D) := \sum_{\substack{d|P(z) \\ d<D}} |r_d(\mathcal{A})|$$

for $D > 1$.

Proof: With all of the above, what we have left to do is appropriately bound the three error terms. We find the following bound for the first error term

$$\begin{aligned} |E_1| &\leq \sum_{\substack{d|P(z) \\ \omega(d)=r}} S\big(\mathcal{A}_d, P^-(d)\big) \\ &\leq \sum_{\substack{d|P(z) \\ \omega(d)=r}} A_d(x) \\ &= \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d)X + r_d(\mathcal{A}) \\ &= XG_r + \sum_{\substack{d|P(z) \\ \omega(d)=r}} r_d(\mathcal{A}). \end{aligned}$$

where we write

$$G_r = \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d).$$

The series $V(z)$ has the following product expansion

$$V(z) = \sum_{d|P(z)} \mu(d)g(d) = \prod_{p<z} \big(1 - g(p)\big).$$

Because $1 - g(p) \leq 1$ for all p, we have the trivial bound $V(z) \leq 1$. Therefore we have the following bound for the third error term

$$|E_3| \leq X \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d)V\big(P^{-1}(d)\big) \leq XG_r$$

which is, up to the $r_d(\mathcal{A})$ terms, the same as the bound for E_1. We collect these $r_d(\mathcal{A})$ terms together with the second error term E_2. Summed together they are not greater than $R(\mathcal{A}, z^r)$ which will act as an application-dependent error term. Also using that the signs of E_1 and E_3 are opposite to one another, we may conclude that

$$|E_1 + E_2 + E_3| \leq G_r X + R(\mathcal{A}, z^r).$$

So, if we now appropriately bound G_r for $r \geq c|\log V(z)|$, the theorem statement follows.

Using that $g(p) < 1$ we find the following

$$G_1 = \sum_{p|P(z)} g(p) \leq \sum_{p|P(z)} -\log(1 - g(p)) = -\log V(z).$$

For larger r we can use the multiplicative property of g to relate G_r to G_1. Given a $d|P(z)$ with $\omega(d) = r$, the multinomial theorem tells us that the coefficient of $g(d)$ in the polynomial

$$G_1^r = \left(\sum_{p|P(z)} g(p)\right)^r$$

equals $\binom{r}{1,1,\ldots,1} = r!$. Therefore the inequality $G_r \leq G_1^r/r!$ holds. Applying the inequality $1/r! \leq e^{-1}(e/r)^r$ (which, without going into detail, follows from Stirlings approximation of the factorial) allows us to write

$$G_r \leq \frac{1}{e}\left(\frac{eG_1}{r}\right)^r \leq \frac{1}{e}\left(\frac{e|\log V(z)|}{r}\right)^r. \tag{3.5}$$

Heuristically, G_r should represent the portion of the total weight $A_1(x)$ that gets added/subtracted in the r-th step of inclusion/exclusion. So we expect this inequality to become useful from around the point where r is least as big as $e|\log V(z)|$, because then the fraction is ≤ 1. The slightly stronger assumption that $r \geq c\,|\log V(z)|$ (note that $c \approx 3.591 > e$) additionally allows

us to nicely deal with the term r^{-r}. For any number $b \geq c$ we have that $b^b \geq e^{2b-c+1}$ (for $b = c$ we have equality and the LHS grows faster). This true, in particular, when we put $b = r/|\log V(z)|$. With this (3.5) becomes

$$G_r \leq e^{r-1} b^{-b|\log V(z)|} \leq e^{r-1} \left(e^{2\frac{r}{|\log V(z)|} - c + 1} \right)^{-|\log V(z)|} \leq e^{-r-1} V(z)^{1-c}.$$

The theorem statement now follows.

□

By making a small assumption on the size of the error terms $r_d(\mathcal{A})$, we are able to prove an alternative version of the above theorem which is less general but easier to use. This is the statement we will use later on in the main part of this thesis. For a stronger version of Brun's sieve we refer to [OdC10] corollary 6.10.

Corollary 3.6: Suppose the remainder terms $r_d(\mathcal{A})$ satisfy $r_d(\mathcal{A}) \leq g(d)d$ for all $d|P(z)$, then

$$|S(\mathcal{A}, z) - XV(z)| \leq (1 + e)V(z)X(\log X)^{-1} + X^{3/4}$$

for $X \geq 16$ and z in the range $4 \leq z \leq X^{1/\log(V(z)^{-c}\log X)}$, where c is the same constant as in theorem 3.5 .

Proof: In the proof of theorem 3.5 we bounded the error term

$$R_r := \sum_{\substack{d|P(z) \\ \omega(d) \leq r}} |r_d(\mathcal{A})|$$

as $R(\mathcal{A}, z^r)$. The assumption on the size of $r_d(\mathcal{A})$ allows us to give a bound for R_r that is more explicit. One thing we will use is the bound $G_k \leq G_1/k!$ for $k \in \mathbb{Z}_{\geq 1}$, which we established in

the proof of theorem 3.5. We further put $A = \max\{1, zG_1/r\}$, and estimate

$$\begin{aligned} R_r &\leq \sum_{\substack{d|P(z)\\ \omega(d)\leq r}} g(d)d \\ &\leq \sum_{\substack{d|P(z)\\ \omega(d)\leq r}} g(d)z^{\omega(d)} \\ &= \sum_{k=0}^{r} G_k z^k \\ &\leq \sum_{k=0}^{r} \frac{(zG_1)^k}{k!} \\ &\leq A^r e^{zG_1/A} \\ &\leq \left(\frac{ezG_1}{r}\right)^r + e^r \end{aligned}$$

where the second to last inequality follows from expanding the exponential to at least its r-th term and using that $A \geq 1$. From the proof of theorem 3.5 we take the bounds for E_1, E_2 and E_3 as well as the bound $G_r \leq e^{-1}(eG_1/r)^r$, combining these with our new bound for R_r we obtain

$$|E_1 + E_2 + E_3| \leq G_r X + R_r \leq \left(\frac{eG_1}{r}\right)^r (e^{-1}X + z^r) + e^r. \tag{3.6}$$

We choose $r = \lfloor \log X / \log z \rfloor$, which together with the bounds for X and z specified in the statement ensures that we can further bound this as desired. We highlight how each of these choices comes into effect. Firstly, the choice of r directly implies that

$$e^{-1}X + z^r \leq e^{-1}X + z^{\log X/\log z} = (e^{-1} + 1)X \tag{3.7}$$

and together with the lower bound $z \geq 4$ it ensures that

$$e^r \leq e^{\log X/\log 4} \leq X^{1/\log 4} \leq X^{3/4}. \tag{3.8}$$

Because of the upper bound $z \leq X^{1/\log(V(z)^{-c}\log X)}$ and $X \geq 16$, we have

$$\begin{aligned} r &= \left\lfloor \frac{\log X}{\log z} \right\rfloor \\ &\geq \left\lfloor \frac{\log X}{\log(X^{1/\log(V(z)^{-c}\log X)})} \right\rfloor \\ &= \left\lfloor \log(V(z)^{-c}\log X) \right\rfloor \\ &\geq \lfloor -c\log V(z) + \log\log 16 \rfloor \\ &\geq c|\log V(z)|. \end{aligned}$$

which allows us to deal with the r^{-r} term as before in the proof of theorem 3.5. Following the same steps we now obtain

$$\left(\frac{eG_1}{r}\right)^r \leq e^{-r}V(z)^{1-c},$$

and using $r \geq \log(V(z)^{-c}\log X) - 1$, we can further bound this as

$$\left(\frac{eG_1}{r}\right)^r \leq \frac{eV(z)}{\log X} \tag{3.9}$$

Substituting the equations (3.7),(3.8) and (3.9) back into (3.6) proves the corollary.

□

Let's finish with an example of how these results can be applied. We return to the twin prime problem from example 3.2. Using the corollary we can fairly quickly establish a meaningful upper bound for the number of twin primes.

Example 3.7: To have our results nicely formulated we slightly rephrase the sieve problem from example 3.2. We take $\mathcal{A}$ to be the characteristic function of the set $\{m(m+2)\,|\,m \leq x\}$ and we keep $\mathcal{P}$ as the set of all primes. The sifting function then counts at least all the twin prime pairs whose smallest prime lies in the interval $(z, x]$. We emphasize the fact that the sifting function can count more than just the number of twin primes pairs when z is not taken to be sufficiently large, in this case when $z < \sqrt{x+2}$. The total weight now becomes

$$A_1\big(x(x+2)\big) = x$$

which we approximate accordingly as $X = x$. For $d|P(z)$ we approximate

$$A_d\big(x(x+2)\big) = \sum_{\substack{m\leq x \\ d|m(m+2)}} 1 = Xg(d) + r_d(\mathcal{A})$$

with the same the density function g and remainder terms $r_d(\mathcal{A})$ as before. Note that the size of $|r_d(\mathcal{A})|$ is bounded by $g(d)d$ which equals the number of solutions to the equation $m(m+2) \equiv 0 \pmod{d}$ that are contained in one residue system. Therefore we may apply corollary 3.6.

To be able to quantify the statement of the corollary we would like some sort of asymptotic expression for $V(z)$. We can find this by looking at the product expansion of $V(z)$. To help us we need two facts, the first of which is Mertens' theorem, which says that

$$\prod_{p<z}\left(1-\frac{1}{p}\right) = \frac{e^\gamma}{\log z}\left(1+O\left(\frac{1}{\log z}\right)\right).$$

The second fact is the absolute convergence of the product

$$\prod_{p\neq 2}\left(1-\frac{2}{p}\right)\left(1-\frac{1}{p}\right)^{-2}$$

which we will establish ourselves. Recognizing the geometric series we can approximate the square as

$$\left(1-\frac{1}{p}\right)^{-2}=(1+p^{-1}+p^{-2}+\dots)^2=1+2p^{-1}+O(p^{-2}).$$

Multiplying this with $(1-2p^{-1})$ gives something with order of size $1+O(p^{-2})$, the infinite product over this is absolutely convergent. A sufficient condition for absolute convergence of an infinite product $\prod_n 1+b_n$ is the absolute convergence of the infinite series $\sum_n b_n$. Now, using the product expansion of $V(z)$ we find its order of growth

$$\begin{aligned}V(z)&=\prod_{p<z}(1-g(p))\\&=\frac{1}{2}\prod_{3\leq p<z}\left(1-\frac{2}{p}\right)\\&=\frac{1}{2}\prod_{3\leq p<z}\left(1-\frac{1}{p}\right)^2\prod_{3\leq p<z}\left(1-\frac{2}{p}\right)\left(1-\frac{1}{p}\right)^{-2}\\&\asymp\frac{1}{(\log z)^2}.\end{aligned}$$

With this piece of information we can determine how we want to choose the size of z. We want the main term to be accurate, the error terms to be small, and we of course need z to stay within the bounds the corollary prescribes.

For better accuracy of our main term $XV(z)=xV(z)$ we basically just want z to be as close to $x^{1/2}$ as possible; when $z=\sqrt{x+2}$ the sifting function counts all of the twin prime pairs in $(z,x]$ and nothing more. As $z\leq x$, it follows from the above asymptotic expression that $V(z)\gg(\log x)^{-2}$. Therefore, the second summand of the error term, $X^{3/4}$ will be negligible compared to the main term no matter how we choose z. The upper bound $z\leq x^{1/\log(V(z)^{-c}\log x)}$ from the corollary ensures that the first summand of the error term, $(1+e)V(z)x(\log x)^{-1}$ is also negligible compared to the main term. As this upper bound is not as good as $x^{1/2}$ (it actually moves away from $x^{1/2}$ as x increases), we just take z as large as this bound permits us. Then $\log z$ is of size

$$\log z=\log(x^{1/\log(V(z)^{-c}\log x)})=\frac{\log x}{\log\log x-c\log\big(V(z)\big)}\asymp\frac{\log x}{\log\log x}$$

where we use that $|\log V(z)| \asymp |\log\log z| \ll |\log\log x|$. The main term is therefore bounded as

$$XV(z) \ll x\left(\frac{\log\log x}{\log x}\right)^2$$

We conclude that the number of twin primes is $O\big(x(\log\log x/\log x)^2\big)$.

4. Triangles with integer sides

We define $\mathcal{T}$ to be the set consisting of the sets $\{p, a, b\}$ containing three positive integers which satisfy

$$a^2 + b^2 - ab = p^2$$

and for which p is prime. The elements in $\mathcal{T}$ represent the non-equilateral triangles with integer sides which have a $\pi/3$ angle opposite to a prime side. We will call the two sides a, b which do not lie opposite to the $\pi/3$ angle the *legs* of such a triangle. Our interest lies in the proportion of positive integers which occur as a leg of a triangle in $\mathcal{T}$. We will in particular distinguish between the even and the odd numbers that occur as a leg. The reason why we want to ignore the side p is that these are exactly the primes equal to 1 modulo 3; their rate of occurrence is known through the prime number theorem for arithmetic progressions.

In this chapter we will establish explicit descriptions of both the set of odd numbers that occur as legs in $\mathcal{T}$ and the set of even numbers that occur as a leg in $\mathcal{T}$. We will rely on the following lemma, which gives an alternative construction of the triangles in $\mathcal{T}$.

Lemma 4.1: For a triangle p, a, b in $\mathcal{T}$ with b odd, there exists a coprime pair of positive integers x, y which satisfy $3x^2 + y^2 \in 4\mathbb{P}$, such that the sides of the triangle can be expressed in x, y as

$$p = \frac{3x^2 + y^2}{4},$$

$$b = xy,$$

$$a = \frac{3x^2 - y^2 + 2xy}{4} \quad \text{or} \quad a = \frac{y^2 - 3x^2 + 2xy}{4}$$

And conversely, if we have a coprime pair of positive integers x, y that satisfy $3x^2 + y^2 \in 4\mathbb{P}$, then defining p, a, b like this gives a triangle in $\mathcal{T}$ as long as $a > 0$, which is the case for at least one of the two expressions for a.

Proof: We start by proving the latter statement which is the more straightforward of the two. To confirm that $a^2 + b^2 - ab = p^2$ holds we just need to carry out a few computations. First we see that

$$p^2 = \frac{9x^4 + 6x^2y^2 + y^4}{16}$$

so that for the first of the possible expressions for a we have

$$a^2 = \frac{(3x^2 - y^2 + 2xy)^2}{4^2} = \frac{9x^4 + 12x^3y - 2x^2y^2 - 4xy^3 + y^4}{16} = p^2 + \frac{3x^2 - y^2 + 2xy}{4}b - b^2$$

and for the second we similarly have

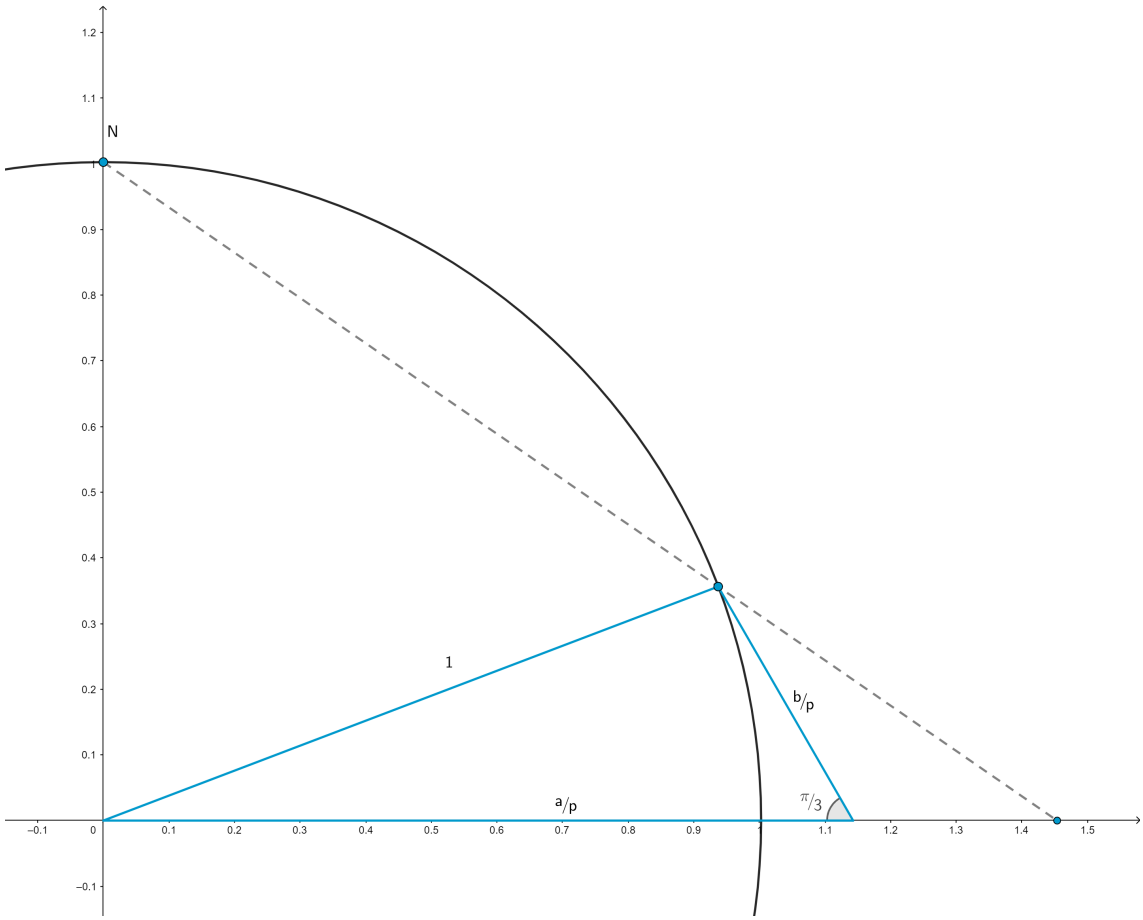
$$a^2 = \frac{(y^2 - 3x^2 + 2xy)^2}{4^2} = \frac{9x^4 - 12x^3y - 2x^2y^2 + 4xy^3 + y^4}{16} = p^2 + \frac{y^2 - 3x^2 + 2xy}{4}b - b^2$$

so that indeed $a^2 + b^2 - ab = p^2$ holds in both cases. To see that at least one of the expressions for a is positive, note that they add up to xy which is positive.

We move on to proving the first of the two statements of the lemma. Let p, a, b be a triangle in $\mathcal{T}$. Like with right triangles, if we divide all the side lengths by p and put the side with length a/p on the horizontal axis, we find a unique point on the positive part of the unit circle S^1. The law of sines gives us the coordinates of this point, they are

$$\left(\frac{\sqrt{3}b}{2p}, \sqrt{1 - \frac{3b^2}{4p^2}}\right) = \left(\frac{\sqrt{3}b}{2p}, \frac{2a - b}{2p}\right).$$

The stereographic projection from the north pole $N = (0, 1)$ on the circle maps a point (x, y) on $S^1 \setminus N$ to $\left(\frac{x}{1-y}, 0\right)$ on the horizontal axis (see illustration below).



The point corresponding to our triangle gets mapped to

$$\frac{\frac{\sqrt{3}b}{2p}}{1-\frac{2a-b}{2p}} = \frac{\sqrt{3}b}{2p-2a+b}.$$

Note that the resulting fraction is reduced only when $3 \nmid 2p+2a-b$. Alternatively, it can be represented as

$$\begin{aligned}\frac{\sqrt{3}b}{2p-2a+b} &= \frac{\sqrt{3}b(2p+2a-b)}{4p^2-(2a-b)^2} \\ &= \frac{2p+2a-b}{\sqrt{3}b}.\end{aligned}$$

Where we use that the equality $a^2+b^2-ab=p^2$ holds to obtain the last step. In either case we end up with something of the form $\sqrt{3}x/y$ or $y/\sqrt{3}x$ where $3 \nmid y$.

We look at how a point on the real line of the same general shape maps back to the circle under the inverse stereographic projection, which maps an arbitrary point r on the real line to the coordinate $\left(\frac{2r}{r^2+1}, \frac{r^2-1}{r^2+1}\right)$ on S^1. For x, y coprime and $3 \nmid y$, it maps the relevant fractions as follows

$$\frac{\sqrt{3}x}{y} \mapsto \left(\frac{2\sqrt{3}xy}{3x^2+y^2}, \frac{3x^2-y^2}{3x^2+y^2}\right) \qquad \text{and} \qquad \frac{y}{\sqrt{3}x} \mapsto \left(\frac{2\sqrt{3}xy}{3x^2+y^2}, \frac{y^2-3x^2}{3x^2+y^2}\right).$$

For certain x, y one these coordinates should correspond to the earlier coordinates on S^1 that we found using the law of sines. Comparing these will allow us to express p, a, b in terms of x and y. We need to take care though, because not all fractions are necessarily reduced in their above forms.

As b is odd and $b \neq p$, the fraction $\sqrt{3}b/2p$ is reduced. This implies that $3x^2+y^2$ is a multiple of $2p$. This prevents either of the coprime x and y from being even. Any odd prime that divides xy divides either x or y and because $3 \nmid y$, this prime then can't divide $3x^2+y^2$. Therefore, the only common divisor of $2xy$ and $3x^2+y^2$ is 2. This allows us to compare the first coordinates, which because of the identical denominators also allows us to compare the second coordinates. The expressions from the lemma statement follow directly from this comparison. □

Remark 4.2: It is insightful to note a few facts regarding the triangles in $\mathcal{T}$ and lemma 4.1. Associated to any triangle p, a, b in $\mathcal{T}$, is the triangle $p, \max(a,b), |a-b|$ which also lies in $\mathcal{T}$. This implies that the triangles come in pairs and that for each such pair there are three integers occurring as leg lengths, two of which are odd and one of which is even.

Related to this is the fact that (under the assumption that b is an odd leg) there are two possible expressions for a given in the lemma. The value of the 'remaining' expression is not

some random number, its absolute value is equal to the differing leg of the associate triangle. This follows from the fact that adding up both expressions for a gives $xy = b$.

Lastly, for a triangle pair, there are two different odd legs which can be taken as b in the lemma. Therefore, there are also two different coprime pairs x, y which generate this triangle pair. They in turn give four expressions for a, at least one of which the even leg has to occur as. There is some irregularity as to which of these four expressions is going to equal the even leg. Similarly, there is irregularity with regards to the parity of the leg that occurs in both of the triangles of the pair.

The following numerical examples illustrate these observations.

	$x = 1, y = 5$	$x = 3, y = 1$	$x = 5, y = 1$	$x = 3, y = 7$
$(3x^2 + y^2)/4$	7	7	19	19
$(3x^2 - y^2 + 2xy)/4$	-3	8	21	5
$(y^2 - 3x^2 + 2xy)/4$	8	-5	-16	16
xy	5	3	5	21

△

We are interested in either the odd or even numbers that occur as legs of triangles in $\mathcal{T}$. Lemma 4.1 immediately gives us a characterization of the set of odd legs. We define $\mathcal{B}(N)$ to be the set that contains the odd numbers $\leq N$ that occur as a leg. From the lemma it follows that

$$\mathcal{B}(N) = \{xy \leq N \mid 3x^2 + y^2 \in 4\mathbb{P} \text{ and } x, y \text{ are coprime}\}.$$

At first glance, lemma 4.1 doesn't seem to give us as nice a description of the even legs. We define $\mathcal{A}(N)$ to be the set that contains all the even numbers $\leq N$ that occur as a leg. From the lemma and the comments in the remark, it only follows that $\mathcal{A}(N)$ equals the union of the following two sets

$$\mathcal{A}_1(N) := \left\{ \frac{3x^2 - y^2 + 2xy}{4} \leq N \;\middle|\; xy \in \mathcal{B}(\infty) \text{ and } xy \equiv -1, 3 \pmod 8 \right\}$$
$$\mathcal{A}_2(N) := \left\{ \frac{y^2 - 3x^2 + 2xy}{4} \leq N \;\middle|\; xy \in \mathcal{B}(\infty) \text{ and } xy \equiv 1, -3 \pmod 8 \right\}.$$

By a change of variables we will be able to give a more convenient description of $\mathcal{A}(N)$ though. Consider the following set

$$\mathcal{A}'(N) := \{uv \leq N \,|\, 3u^2 + v^2 \in 4\mathbb{P} \text{ and } \gcd(u, v) = 2\},$$

which at first sight seems to be more related to the odd legs $\mathcal{B}(N)$ than to the even legs $\mathcal{A}(N)$. Applying the substitution $u = (x+y)/2$ and $v = (3x-y)/2$ to an even leg in $\mathcal{A}_1(N)$, we find that

$$\begin{aligned} uv &= \frac{3x^2+2xy-y^2}{4} = \text{ the even leg,} \\ 3u^2+v^2 &= \frac{3x^2+6xy+3y^2+9x^2-6xy+y^2}{4} = 3x^2+y^2 \in 4\mathbb{P}. \end{aligned}$$

Note that the above equalities force both u and v to be even but also prevents them from having a larger common divisor than 2. It follows that $\mathcal{A}_1(N) \subseteq \mathcal{A}'(N)$.

Similarly, if we take an even leg inside $\mathcal{A}_2(N)$ and apply the substitution $u = (y-x)/2$ and $v = (3x+y)/2$, we find

$$\begin{aligned} uv &= \frac{y^2+2xy-3x^2}{4} = \text{the even leg,} \\ 3u^2+v^2 &= \frac{3y^2-6xy+3x^2+9x^2+6xy+y^2}{4} = 3x^2+y^2 \in 4\mathbb{P}. \end{aligned}$$

from which we likewise conclude that $\mathcal{A}_2(N) \subseteq \mathcal{A}'(N)$.

We can do one of two changes of variables back to prove that $\mathcal{A}'(N) \subseteq \mathcal{A}_1(N) \cup \mathcal{A}_2(N)$. Given an element $uv \in \mathcal{A}'(N)$, first consider the substitution $x = (u+v)/2$ and $y = (3u-v)/2$. We have that

$$\begin{aligned} \frac{3x^2-y^2+2xy}{4} &= \frac{3u^2+6uv+3v^2-9u^2+6uv-v^2+6u^2+4uv-2v^2}{16} = uv, \\ 3x^2+y^2 &= \frac{3u^2+6uv+3v^2+9u^2-6uv+v^2}{4} = 3u^2+v^2 \in 4\mathbb{P}. \end{aligned}$$

This corresponds to an element in $\mathcal{A}_1(N)$ if x and y are positive integers such that $\gcd(x,y) = 1$. Note that $\{u,v\} = \{0,2\} \pmod 4$, otherwise $(3u^2+v^2)/4$ is an even number. Therefore x and y are both (odd) integers. Suppose that $d|x$ and $d|y$, then $d|(x+y) = 2u$ and $d|(3x-y) = 2v$, which implies that $d|\gcd(u,v) = 2$. As d is odd, indeed $d = 1$.

In the case that $3u < v$, the above substitution doesn't give a positive integer value for y. We can alternatively make the substitution $x = (v-u)/2$ and $y = (3u+v)/2$. We similarly have that

$$\begin{aligned} \frac{y^2-3x^2+2xy}{4} &= \frac{9u^2+6uv+v^2-3u^2+6uv-3v^2+2v^2+4uv-6u^2}{16} = uv, \\ 3x^2+y^2 &= \frac{3u^2-6uv+3v^2+9u^2+6uv+v^2}{4} = 3u^2+v^2 \in 4\mathbb{P}. \end{aligned}$$

which corresponds to an element in $\mathcal{A}_2(N)$ because $\gcd(x,y) = 1$ like before. We conclude that $\mathcal{A}'(N) = \mathcal{A}(N)$.

There is an obvious bijection between $\mathcal{A}'(N)$ and $\mathcal{C}(4N)$, where the latter is defined as follows

$$\mathcal{C}(N) := \{uv \leq N \,|\, 3u^2 + v^2 \in \mathbb{P}\}.$$

In a sense this is the most natural of the three sets $\mathcal{A}(N), \mathcal{B}(N), \mathcal{C}(N)$, even though it doesn't necessarily contain any legs of triangles in $\mathcal{T}$. Not coincidentally, this is also the most workable of the three, and luckily for us their likeness is so strong that we will be able to permit ourselves to mainly concern ourselves with that one. We will say a few more words as to why this is the case in the following chapters.

What we would ultimately like is an asymptotic expression for the size of $\mathcal{C}(N)$. While this will remain out of reach for now, we will be able to fully establish an asymptotic upper bound and conditionally establish a lower bound in the following two chapters. We expect the upper bound to be tight up to a factor of around $\log\log N$, while the lower bound can probably be improved by small power of $\log N$ to match the upper bound.

Before we start off our effort, we want to go through some heuristics to form an idea of what we could realistically expect the size of $\mathcal{C}(N)$ to be. The first useful observation to make is the following.

Proposition 4.3: Exactly the primes $p \equiv 1 \pmod 3$ can be expressed (uniquely) as $p = 3u^2 + v^2$ with $u, v \in \mathbb{Z}_{\geq 1}$.

Proof: For a prime $p \equiv 1 \pmod 3$, quadratic reciprocity dictates that -3 is a quadratic remainder modulo p. This implies that there is an integer x such that p divides $x^2 + 3 = (x + \sqrt{-3})(x - \sqrt{-3})$. As p divides neither of these factors inside $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$, this prevents p from being prime in this unique factorization domain (see appendix A for a proof of this fact). So, this ring must contain two non-unit elements α and α' in such that $p = \alpha\alpha'$. Mapping this equality with the multiplicative norm map $N : \mathbb{Z}[\frac{1+\sqrt{-3}}{2}] \to \mathbb{Z}$ tells us that $N(\alpha) = N(\alpha') = p$. By possibly multiplying with an appropriate unit (one of $\frac{1\pm\sqrt{-3}}{2}$) we may assume that $\alpha \in \mathbb{Z}[\sqrt{-3}]$, so that we can express it as $\alpha = u\sqrt{-3} + v$. It then follows that

$$p = N(\alpha) = \alpha\overline{\alpha} = 3u^2 + v^2$$

The uniqueness of u, v follows from the unique factorization inside $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ and the fact that the unit with which an element $\alpha \in \mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ has to be multiplied with to end up in $\mathbb{Z}[\sqrt{-3}]$ is unique up to sign.

□

By the prime number theorem for arithmetic progressions, the number of primes equal to 1 mod 3 up to N grows asymptotically as $N/2\log N$. This at least ensures that $\mathcal{C}(N)$ grows infinitely large as N increases.

As mentioned earlier, the Hardy-Ramanujan theorem states that the expected number of distinct prime divisors of a positive integer n is $\log\log n$. Therefore, n should have around $2^{\log\log n} = (\log n)^{\log 2}$ divisors $d|n$ such that d and n/d are coprime. Each such divisor represents a pair which we can fill into the equation $3u^2 + v^2$ and try to hit a prime with. The odds of hitting such a prime are around $(\log n)^{-1}$. Because $\log 2 - 1 < 0$, a reasonable expectation for $\#\mathcal{C}(N)/N$ would be some negative power of $\log N$.

5. An upper bound

In this chapter we will establish an asymptotic upper bound for the count of odd/even numbers that occur as legs of triangles in $\mathcal{T}$. More precisely, we will spend this chapter proving the following statement.

Theorem 5.1: The following asymptotic inequality holds

$$\#\mathcal{C}(N) \leq \frac{N}{(\log N)^{\eta}}(\log\log N)^{O(1)}$$

where $\eta = 1 - \frac{1+\log\log 2}{\log 2} \approx 0.086$.

△

This constant η is not uncommon in mathematics, its first occurrence seems to be in a paper by Paul Erdös concerning the number of distinct integers in an $n \times n$ multiplication table [Erdö60]. It is commonly known as the Erdös-Ford-Tenenbaum constant, referring to two other mathematicians who came across this constant in their works [Ford08] [Tene84].

As we mentioned before, we can permit ourselves to mainly work with $\mathcal{C}(N)$. This is because firstly, the relation $\#\mathcal{A}(N) = \#\mathcal{C}(4N)$ directly implies that theorem 5.1 also holds for $\mathcal{A}(N)$. Furthermore, most of the things we are going to do in this chapter to prove theorem 5.1 immediately work when we take the odd legs $\mathcal{B}(N)$ instead of $\mathcal{C}(N)$, and where needed, only slight adaptations are needed to count quadruples of primes instead of just primes. We will quickly discuss these adaptations at the end of the chapter, when more context is in place.

The general idea we will be executing is to remove the numbers from $\mathcal{C}(N)$ which have a large number of prime divisors or are divisible by only small primes. The set that remains is then approximable with Brun's sieve. The bound obtained from the sieve as well as the size of the removed sets will depend on how we quantify our notions of 'small' and even more so 'large'. In order to optimize the result from our methods, some fine tuning will be needed at the end.

Throughout this chapter we will use N to denote an arbitrarily large positive integer. We start by defining and establishing bounds for the (sub)sets that are to be removed. Let $1 < \alpha < 2$ be a real parameter which will be chosen optimally at the end of the chapter. Why it has to satisfy these bounds will become clear somewhere along the way. Further put $L = \lfloor \alpha \log\log N \rfloor$ and define

$$\mathcal{E}_1(N) := \{n \leq N \,|\, \omega(n) > L\}$$

which is the first set of numbers we are going to exclude from $\mathcal{C}(N)$. Note that the expected number of distinct prime factors of a positive integer n is $\log\log n$. The Erdös-Kac theorem even states that $\omega(n)$ has a normal distribution.

Lemma 5.2: The following bound holds

$$\#\mathcal{E}_1(N) \ll \frac{N}{(\log N)^{1-\alpha+\alpha\log\alpha}}$$

Proof: In lemma 2.1 we uniformly determined an asymptotic upper bound for the count of numbers not greater than N which have exactly n distinct prime divisors. Because of the uniformity, we can turn this into a bound for $\mathcal{E}_1(N)$ by summing over i. Writing $k = \log\log N$ to shorten notation, we find

$$\#\mathcal{E}_1(N) \ll \sum_{i>L} \frac{N}{\log N} \cdot \frac{(k+c_0)^{i-1}}{(i-1)!} = \frac{N}{\log N} \sum_{i\geq L} \frac{(k+c_0)^i}{i!}$$

In the sum on the right hand side we recognize the tail end of the Taylor expansion of the exponential e^{k+c_0}. In appendix B we will elaborate on the fact that L being larger than $k+c_0$ (which is true for large N and the reason why we require $\alpha > 1$) ensures that asymptotically the entire sum can be approximated by its largest term (which is its first term). We will actually use this and similar approximations a few more times. For now it tells us that

$$\begin{aligned}
\sum_{i\geq L} \frac{(k+c_0)^i}{i!} &\ll \frac{(k+c_0)^L}{L!} \\
&\leq \left(\frac{(k+c_0)e}{L}\right)^L \\
&= \left(\frac{e}{\alpha} + O\left(\frac{1}{k}\right)\right)^L \\
&\ll \left(\frac{e}{\alpha}\right)^L \\
&\leq (\log N)^{\alpha-\alpha\log\alpha}
\end{aligned} \tag{5.1}$$

where we used that $1/L! \leq e^L/L^L$ (which follows from looking at the L-th term of e^L) and where the second to last step follows in the same fashion as

$$\left(1+\frac{1}{k}\right)^{\alpha k} = e^{k\log(1+\frac{1}{k})} = e^{1+O(k^{-1})} \leq \text{some constant}$$

follows for large k. Multiplying equation (5.1) with $N/\log N$ gives us the promised inequality. □

We move on to the second set of numbers we want to exclude from $\mathcal{C}(N)$. We define

$$\mathcal{E}_2(N) := \{n \leq N \,|\, P^+(n) \leq N^{1/\log\log N}\}$$

For this set we have a fixed bound that we won't need to worry about when optimizing later on.

Lemma 5.3: The following bound holds

$$\#\mathcal{E}_2(N) \leq \frac{N}{(\log N)^2}$$

Proof: For positive integers $x \geq 1$ and $y \geq 2$, define

$$\Psi(x,y) := \#\{n \leq x \,|\, P^+(n) \leq y\}.$$

In equation 1.6 in the paper [Brui51] (not to be confused with the very similar equation 1.6 in the second version of this paper from 1966), the following inequality is established

$$\Psi(x,y) \leq x(\log y)^2 \exp\big(-z\log z - z\log\log z + O(z)\big) \qquad \text{for } 3 < z < \frac{4\sqrt{y}}{\log y} \tag{5.2}$$

where $z = (\log x)/\log y$.

With $x = N$ and $y = N^{1/\log\log N}$, the condition on z becomes

$$3 < \log\log N < \frac{4(\log\log N)N^{1/2\log\log N}}{\log N},$$

which is satisfied for large enough N, to easily see that this is true for the upper bound one may rewrite it to $1 < 4\exp\big(\frac{\log N}{2\log\log N} - \log\log N\big)$. Now applying the inequality, we find that

$$\begin{aligned}\mathcal{E}_2(N) &= \Psi(N, N^{1/\log\log N}) \\ &\leq \frac{N(\log N)^2}{(\log\log N)^2}(\log N)^{-\log\log\log N - \log\log\log\log N + O(1)} \\ &\leq \frac{N}{(\log N)^2}.\end{aligned}$$

Again, the last step just follows when N is taken to be large enough.

□

Note that we could have taken any fixed power of $\log N$ in the denominator. We just choose 2 because this makes the bound strong enough that it won't get in our way during the optimization at the end of the chapter.

This might raise the legitimate question as to why we bother with excluding such a small set. The reason for this is that it will clean up our upcoming application of Brun's sieve. Sieves are set up in a way that the sifting function counts numbers we are interested in inside the interval $(z, x]$, but ignores the numbers inside the interval $(0, z]$. In practice this isn't a big

problem because z is always such a small number that it is negligible compared to the sifting function, but it isn't something we can just ignore. By excluding elements from $\mathcal{E}_2(N)$ we are preemptively dealing with this, which somewhat simplifies the application of the sieve.

Having bounded the sizes of the sets we want to exclude from $\mathcal{C}(N)$, we will now bound the size of the set that remains. We define

$$\mathcal{C}^*(N) := \mathcal{C}(N) \setminus \big(\mathcal{E}_1(N) \cup \mathcal{E}_2(N)\big).$$

Excluding the elements in $\mathcal{E}_2(N)$ from $\mathcal{C}(N)$ ensures that an integer n in the remaining $\mathcal{C}^*(N)$ can be decomposed in at least one of two ways. The first possibility is that $n = u_0\ell v$ where ℓ is a prime larger than $N^{1/\log\log N}$ and $3u_0^2l^2 + v^2$ is prime, the other possibility is that $n = uv_0\ell$ where ℓ is a prime larger than $N^{1/\log\log N}$ and $3u^2 + v_0^2\ell^2$ is prime. The removal of elements from $\mathcal{E}_1(N)$ additionally ensures that the remainder n/ℓ has at most L distinct prime divisors. This confirms the following inequality

$$\#\mathcal{C}^*(N) \leq \sum_{\substack{u_0 v \leq N^{1-1/\log\log N} \\ \omega(u_0 v)\leq L}} S_1(u_0, v) + \sum_{\substack{uv_0 \leq N^{1-1/\log\log N} \\ \omega(uv_0)\leq L}} S_2(u, v_0) \tag{5.3}$$

where

$$S_1(u_0, v) = \sum_{\substack{N^{1/\log\log N} < \ell \leq N/u_0 v \\ \ell,\, 3u_0^2l^2+v^2 \in \mathbb{P}}} 1 \quad \text{and} \quad S_2(u, v_0) = \sum_{\substack{N^{1/\log\log N} < \ell \leq N/uv_0 \\ \ell,\, 3u^2+v_0^2l^2 \in \mathbb{P}}} 1.$$

This is useful to us because S_1 and S_2 can be bounded with Brun's sieve.

Lemma 5.4: For $j = 1, 2$, we have

$$S_j(u, v) \leq \frac{N(\log\log N)^{O(1)}}{(\log N)^2 uv}$$

Proof: The proofs for $j = 1, 2$ are almost identical. To avoid confusion and having to interrupt the proof too much, we will proceed under the assumption that $j = 1$. We will say a few words at the end about where the differences lie.

Note that we may assume that $2|uv$ and that u and v are coprime, because otherwise $S_1(u, v) = 0$. We are going to define a sieve problem for which the sifting function is at least as large as $S_1(u, v)$, and then we will use corollary 3.6 to give us an asymptotic upper bound.

We take $x = N/uv$, take $\mathcal{P} = \mathbb{P} = \{\text{all primes}\}$ and define $\mathcal{A}$ as the characteristic function of the set

$$\{m(3u^2m^2 + v^2) \,|\, 1 \leq m \leq x\}.$$

We further choose $z = N^{(\log\log N)^{-3}}$, which is simultaneously both small and large enough to make the corollary work, to give us a useful bound, and such that

$$S_1(u, v) \leq \#\big\{m \in (z, x] \,\big|\, \gcd\big(m(3u^2m^2 + v^2), P(z)\big) = 1\big\} = S(\mathcal{A}, z).$$

As in example 3.7, we have chosen $\mathcal{A}$ such that the total weight can be approximated nicely by putting $X = x$. For $d|P(z)$, the size of $A_d\big(x(3u^2x^2+v^2)\big)$ is directly dependent on the number of solutions m to the equation $m(3u^2m^2+v^2) \equiv 0 \pmod{d}$. By the Chinese remainder theorem this reduces to the case where $d = p$ is prime. If $p|3uv$, the equation has only the one solution $m \equiv 0$ per residue system (note that this includes the case $p = 2$). When this isn't the case there might be more solutions, depending on the number of solutions of the equation $3u^2m^2 + v^2 \equiv 0 \pmod{p}$. Using quadratic reciprocity we find that

$$\left(\frac{-v^2 3^{-1} u^{-2}}{p}\right) = \left(\frac{-3}{p}\right) = \left(\frac{p}{3}\right),$$

which implies that when $p \equiv 1 \pmod 3$ there are two additional solutions to the equation, and in the other case when $p \equiv 2 \pmod 3$ there are none. This motivates defining $g(d)$ by putting

$$g(p) = \begin{cases} \frac{1}{p} & \text{if } p|3uv \text{ or } p \equiv 2 \pmod 3 \\ \frac{3}{p} & \text{if } p \nmid 3uv \text{ and } p \equiv 1 \pmod 3. \end{cases}$$

With this the size of

$$|r_d(\mathcal{A})| = |g(d)X - A_d\big(x(3u^2x^2 + v^2)\big)|$$

is bounded by the number of solutions m to the equation $3u^2m^2 + v^2 \equiv 0 \pmod{d}$ contained in one residue system, which is equal to $g(d)d$. This confirms that the main condition to apply corollary 3.6 is satisfied. To be able to quantify the statement and verify that the z we chose isn't too large, we want an asymptotic expression for

$$V(z) = \prod_{p<z} (1 - g(p)).$$

Firstly, Mertens' theorem generalized for arithmetic progressions (see theorem 1 in [Will74]) tells us that

$$\prod_{\substack{p<z \\ p\equiv 1 (\mathrm{mod} 3)}} \left(1 - \frac{1}{p}\right) \asymp \prod_{\substack{p<z \\ p\equiv 2 (\mathrm{mod} 3)}} \left(1 - \frac{1}{p}\right) \asymp \frac{1}{(\log z)^{1/2}}.$$

Secondly, the product

$$\prod_{p>3} \left(1 - \frac{3}{p}\right) \left(1 - \frac{1}{p}\right)^{-3}$$

is absolutely convergent; after expanding the geometric series the cube can be approximated as $1 + 3p^{-1} + O(p^{-2})$, hence the terms of the infinite product are of size $1 + O(p^{-2})$. Using

these two facts we find

$$\begin{aligned}
V(z) &= \prod_{p\mid 3uv}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p<z\\ p\nmid 3uv\\ p\equiv 2(\mathrm{mod}3)}}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p<z\\ p\nmid 3uv\\ p\equiv 1(\mathrm{mod}3)}}\Big(1-\frac{3}{p}\Big)\\
&\asymp \prod_{\substack{p<z\\ p\equiv 2(\mathrm{mod}3)}}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p<z\\ p\equiv 1(\mathrm{mod}3)}}\Big(1-\frac{3}{p}\Big)\\
&\asymp \prod_{\substack{p<z\\ p\equiv 2(\mathrm{mod}3)}}\Big(1-\frac{1}{p}\Big)\prod_{\substack{p<z\\ p\equiv 1(\mathrm{mod}3)}}\Big(1-\frac{3}{p}\Big)\Big(1-\frac{1}{p}\Big)^{-3}\prod_{\substack{p<z\\ p\equiv 1(\mathrm{mod}3)}}\Big(1-\frac{1}{p}\Big)^{3}\\
&\asymp \frac{1}{(\log z)^2}.
\end{aligned}$$

Before we can apply the corollary we need to check that the upper bound $z \le X^{1/\log(V(z)^{-c}\log X)}$ holds, which we can only properly do now that we have an asymptotic expression for $V(z)$. Let C_1 be a constant such that for (large) N the inequality

$$-c\log V(z) \le C_1 \log\log z$$

holds. Then, also using that $X \ge N^{1/\log\log N}$ we compare logarithms

$$\begin{aligned}
\log\big(X^{1/\log(V(z)^{-c}\log X)}\big) &= \frac{\log X}{\log\log X - c\log V(z)}\\
&\ge \frac{\log N}{\log\log N\big(\log\log N - c\log V(z)\big)}\\
&\ge \frac{\log N}{(1+C_1)(\log\log N)^2}\\
&\ge \frac{\log N}{(\log\log N)^3}\\
&= \log z
\end{aligned}$$

which confirms that z is indeed small enough for the corollary to apply.

Both our choice of X and the asymptotic we have found for $V(z)$ is the same as in example 3.7. So, for the same reasons as we gave there, the error term $XV(z)(\log X)^{-1} + X^{3/4}$ is negligible compared to the main term $XV(z)$. Therefore the sifting function is asymptotically

equal to this main term. We conclude that

$$\begin{aligned} S_1(u,v) &\leq S(\mathcal{A},z) \\ &\ll XV(z) \\ &\ll \frac{X}{(\log z)^2} \\ &= \frac{N(\log\log N)^{O(1)}}{uv(\log N)^2} \end{aligned}$$

which finishes the proof for the case $j = 1$.

At the beginning we promised we would point out what's different for the case $j = 2$. We have to change the sieve such that it detects primes of the form $3u^2 + v^2m^2$ instead of ones of the form $3u^2m^2 + v^2$, under the same restrictions for m. Really the only point where we have to take care doing this is when defining $g(d)$, but because $\left(\frac{-3u^2v^{-2}}{p}\right) = \left(\frac{p}{3}\right)$ the definition of g remains unchanged and from there on the rest of the proof can essentially be repeated word for word.

□

By substituting these bounds into (5.3) we obtain

$$\#\mathcal{C}^*(N) \leq \frac{N(\log\log N)^{O(1)}}{(\log N)^2} I \tag{5.4}$$

where

$$I = \sum_{\substack{uv \leq N^{1-1/\log\log N} \\ \omega(uv) \leq L}} \frac{1}{uv},$$

which we bound as follows.

Lemma 5.5: The following asymptotic bound holds

$$I \ll (\log N)^{\alpha(1+\log 2-\log\alpha)}.$$

Proof: We start with the following inequality

$$I \leq \sum_{i+j\leq L} \sum_{\substack{u\leq N \\ \omega(u)=i}} \frac{1}{u} \sum_{\substack{v\leq N \\ \omega(v)=j}} \frac{1}{v}.$$

Summing up to N instead of $N^{1/\log\log N}$ in the second sum and N instead of $N^{1/\log\log N}/u$ in the third is just for ease of notation, it will not worsen the bound we obtain.

Given a $u \leq N$ with $\omega(u) = i$, the multinomial theorem implies that u^{-1} is a summand in the polynomial

$$\left(\sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^i$$

with coefficient $\binom{i}{1,1,\ldots,1} = i!$. Making a similar observation for $v \leq N$ with $\omega(v) = j$ allows us to further bound I as

$$\begin{aligned} I &\leq \sum_{i+j \leq L} \frac{1}{i!} \left(\sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^i \frac{1}{j!} \left(\sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^j \\ &= \sum_{i+j \leq L} \frac{1}{(i+j)!} \binom{i+j}{i} \left(\sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^{i+j}. \end{aligned}$$

Making the substitution $n = i + j$ and using the binomial identity $\sum_{i=0}^{n} \binom{n}{i} = 2^n$, we can further bound this as

$$I \leq \sum_{n \leq L} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} \left(\sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^n = \sum_{n \leq L} \frac{1}{n!} \left(2 \sum_{p^\nu \leq N} \frac{1}{p^\nu} \right)^n.$$

Mertens' theorem tells us that there exists a constant c_1 such that for all $N \geq 2$ the difference between the sum of prime reciprocals $\sum_{p \leq N} p^{-1}$ and $\log \log N$ does not exceed c_1. The sum of the higher power prime reciprocals is convergent, and it is not larger than $\sum_{n=2}^{\infty} (\zeta(n) - 1) = 1$ (this equality follows by expanding $\zeta(n)$ and switching the order of summation). Using this we can further bound I as

$$I \leq \sum_{n \leq L} \frac{(2 \log \log N + 2c_1 + 2)^n}{n!}, \tag{5.5}$$

which is a sum that should look familiar to us. Earlier on in the chapter we gave an upper bound for the tail end of an exponential. The sum above however contains the first terms, but because $L < 2 \log \log N + 2c_1 + 2$ (using that $\alpha < 2$), this sum can likewise be asymptotically approximated by its largest term (which is now its last term). We will prove this in detail in

appendix B. Applying this and reusing the inequality $1/L! \leq e^L/L^L$, we finally bound I as

$$\begin{aligned} I &\ll \frac{(2e\log\log N + 2ec_1 + 2e)^L}{L!} \\ &\leq \left(\frac{2e\log\log N + 2ec_1 + 2e}{L}\right)^L \\ &= \left(\frac{2e}{\alpha} + O\left(\frac{1}{L}\right)\right)^L \\ &\ll \left(\frac{2e}{\alpha}\right)^L \\ &\leq (\log N)^{\alpha(1+\log 2-\log\alpha)}. \end{aligned}$$

□

By substituting this bound for I into (5.4) we finalize our approximation of $\#\mathcal{C}^*(N)$ as follows

$$\#\mathcal{C}^*(N) \leq \frac{N(\log\log N)^{O(1)}}{(\log N)^{2-\alpha(1+\log 2-\log\alpha)}}.$$

Now, after a lot of bounds and approximations we can wrap up the proof of theorem 5.1 and with it the chapter. We have established asymptotic bounds for sizes of the three sets $\mathcal{E}_1(N), \mathcal{E}_2(N), \mathcal{C}^*(N)$ which unioned together contain $\mathcal{C}(N)$. So, the size of $\mathcal{C}(N)$ is bounded by the size of the largest of the three. The bounds for $\#\mathcal{E}_1(N)$ and $\#\mathcal{C}^*(N)$ are dependent of our choice of $1 < \alpha < 2$. The bound for $\#\mathcal{E}_2(N)$ is the best bound of the three, regardless of how α is chosen. Comparing the bounds for the other two sets we see that to find the best bound we need to pick α such that it maximizes

$$\min\{2 - \alpha(1 + \log 2 - \log\alpha), 1 - \alpha + \alpha\log\alpha\}.$$

These functions respectively have derivatives $\log\alpha - \log 2$ and $\log\alpha$. With this it is easy to see that the first function is decreasing in the interval $(1, 2)$ while the second is increasing in the interval $(1, 2)$. Therefore their minimum is maximal at their intersection $\alpha = 1/\log 2$. We conclude that

$$\#\mathcal{C}(N) = \frac{N(\log\log N)^{O(1)}}{(\log N)^{\eta}},$$

where $\eta = 1 - (1 + \log\log 2)/\log 2$.

Remark 5.6: In the introduction of this chapter we promised to end the chapter with a few words on how to adapt the proof for the set of odd integers $\mathcal{B}(N)$ instead of $\mathcal{C}(N)$. Remember that we were able to characterize $\mathcal{B}(N)$ as follows

$$\mathcal{B}(N) = \{uv \leq N \,|\, 3u^2 + v^2 \in 4\mathbb{P} \text{ and } u, v \text{ are coprime}\}.$$

The only place in the proof that needs attention is where we apply Brun's sieve, starting at page 37. First of all the definitions of S_1 and S_2 need to be redefined as

$$S_1(u_0, v) = \sum_{\substack{N^{1/\log\log N} < \ell \leq N/u_0 v \\ \ell,\, (3u_0^2 l^2 + v^2)/4 \in \mathbb{P}}} 1 \quad \text{and} \quad S_2(u, v_0) = \sum_{\substack{N^{1/\log\log N} < \ell \leq N/u v_0 \\ \ell,\, (3u^2 + v_0^2 l^2)/4 \in \mathbb{P}}} 1.$$

We need to accordingly change a few things in the proof of lemma 5.5. The first change is a bit subtle; we may now assume that $2 \nmid uv$ and that u and v are coprime. Looking mod 8 ensures us that $(3u^2 + v^2)/4$ is never even. Therefore, we can take $\mathcal{P} = \mathbb{P}_{\text{odd}}$. The next change to make is to define $\mathcal{A}$ as the characteristic function of the set

$$\left\{ m(3u^2m^2 + v^2)/4 \,\middle|\, 1 \leq m \leq x \right\} \cap \mathbb{Z}_{\geq 1}.$$

All the other setup parameters stay the same. Because we excluded 2 from $\mathcal{P}$, all $d|P(z)$ are odd and 4 is invertible modulo d. So the solutions m to the equation $m(3u^2m^2 + v^2)/4 \equiv 0 \pmod{d}$ occur in the exact same manner as in the original proof.

6. A (conditional) lower bound

As mentioned before, the general idea we follow comes from the paper [ChPo17] in which the authors establish asymptotic upper and lower bounds for the number of the odd/even integers that occur as a non-hypotenuse side of right triangles with a prime hypotenuse. In the previous chapter we successfully adapted the methods from one of their chapters to replicate the upper bound established there for legs of triangles in $\mathcal{T}$.

Sadly though, not all of the tools that are used to produce the lower bound in [ChPo17] are general enough to be applicable in the same manner for legs of triangles in $\mathcal{T}$. The obstruction is the usage of a theorem which allows us to count Gaussian primes (which are pairs (u, v) such that $u^2 + v^2$ is prime) in an area of the positive quadrant. What we would like is to be able to instead count the pairs (u, v) in the same area for which $3u^2 + v^2$ is prime. Numerical evidence suggests that a result similar to the theorem on Gaussian primes should be true, but sadly no such generalization seems to exist in the literature. Since the rest of the methods used to establish the lower bound in [ChPo17] do carry over nicely, we choose to still present the proof of an analogous lower bound which works up to this result. When a bit more context is in place we will provide the exact details of what it is that we are assuming, as well as the numerical data which supports it. The exact statement we are going to prove is the following.

Theorem 6.1: If the primes of the form $3u^2 + v^2$ are distributed like the Gaussian primes, then for every $\varepsilon > 0$, the following asymptotic inequality holds

$$\#\mathcal{C}(N) \gg \frac{N}{(\log N)^{\log 4 - 1 + \varepsilon}}.$$

△

As in the previous chapter, the relation $\#\mathcal{A}(N) = \#\mathcal{C}(4N)$ implies that this bound also holds for $\mathcal{A}(N)$, and all the methods we will use also work for $\mathcal{B}(N)$ with only minor adaptations.

It unfortunately isn't the case that we can just give lower bounds for the same subsets of $\mathcal{C}(N)$ that we bounded from above in the previous chapter. The approach we take is slightly different, but still similar enough that it conveniently allows us to reuse a lot of the tools from the previous chapter.

As in the previous chapter, N is an integer of arbitrary size and $1 < \beta < 2$ a parameter.

We start by defining the following sets

$$\begin{aligned}
\mathcal{L}_0 &= \{(u,v) \in \mathbb{Z}^2_{\geq 1} \mid 1 \leq uv \leq N, 3u^2 + v^2 \in \mathbb{P}\},\\
\mathcal{L}_1 &= \{(u,v) \in \mathcal{L}_0 \mid P^+(uv) \leq N^{1/\log\log N}\},\\
\mathcal{L}_2 &= \{(u,v) \in \mathcal{L}_0 \setminus \mathcal{L}_1 \mid \omega(u) > \beta \log\log N\},\\
\mathcal{L}_3 &= \{(u,v) \in \mathcal{L}_0 \setminus \mathcal{L}_1 \mid \omega(v) > \beta \log\log N\},\\
\mathcal{L} &:= \mathcal{L}_0 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3).
\end{aligned}$$

All of which should look quite familiar, as these sets are very similar to the ones we bounded in the previous chapter. We shall be looking for a lower bound for $\#\mathcal{L}$, which means that we need a lower bound for $\#\mathcal{L}_0$ and we need upper bounds for the cardinalities $\#\mathcal{L}_i$ for $i = 1, 2, 3$.

The reason we are interested in $\#\mathcal{L}$ is because we can relate it to the size of $\mathcal{C}(N)$ via the following set

$$\mathcal{S}(N) := \big\{\big((u,v),(s,t)\big) \in \mathcal{L}^2 \,|\, uv = st\big\}.$$

Through a clever application of the Cauchy-Schwarz inequality we can prove the following.

Lemma 6.2: The following inequality holds

$$\#\mathcal{C}(N) \geq (\#\mathcal{L})^2/\#\mathcal{S}(N). \tag{6.1}$$

Proof: We will prove the stronger inequality with the smaller set

$$\mathcal{C}'(N) := \{n \in \mathcal{C}(N) \mid \exists (u,v) \in \mathcal{L} \text{ such that } uv = n\}$$

in the place of $\mathcal{C}(N)$.

Say that $\mathcal{C}'(N)$ contains τ elements $n_1, \dots n_\tau$. We can then partition $\mathcal{L}$ into τ parts according to what the product uv of an element (u,v) is. Denote with $k_1, \dots k_\tau$ the sizes of each subset of the partition, i.e., there are k_i pairs $(u,v) \in \mathcal{L}$ such that $uv = n_i$. With this we have $\#\mathcal{L} = k_1 + \cdots + k_\tau$. We can partition $\mathcal{S}(N)$ in the same manner, but now the i-th subset of the partition has size k_i^2. So that $\#\mathcal{S}(N) = k_1^2 + \cdots + k_\tau^2$. Cauchy-Schwarz relates the two cardinalities as follows

$$(\#\mathcal{L})^2 = \left(\sum_{i=1}^{\tau} k_i\right)^2 \leq \tau \sum_{i=1}^{\tau} k_i^2 = \#\mathcal{C}'(N)\#\mathcal{S}(N).$$

The inequality (6.1) follows immediately.

□

Because $\#\mathcal{S}(N)$ features in the denominator in (6.1), we will also have to bound it from above. We will do this at the end of this chapter by yet again using Brun's sieve.

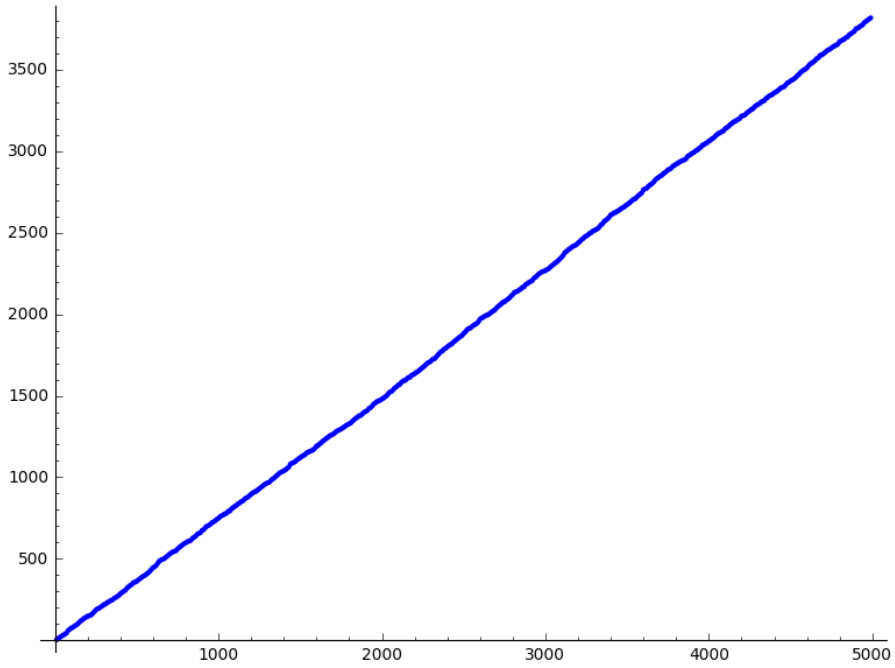
The irony now is that $\mathcal{L}_0$ is the only set we want to give a direct lower bound for, but that this is the point where we are going to have to leave a gap in the proof. The lower bound we expect to hold is the following.

Conjecture 6.3: The following asymptotic bound holds

$$\#\mathcal{L}_0 \gg N.$$

△

The set $\mathcal{L}_0$ contains the positive integer pairs (u, v) under the curve $v = N/u$ for which $3u^2 + v^2$ is prime. For convenience we will call such pairs $(3, 1)$-Gaussian primes. In [ChPo17] the analogous statement (equation (3.2)) is that the same $\gg$ bound holds when counting the number of Gaussian primes (a, b) under the curve $b = N/a$. The proof relies on a theorem which, under a few mild conditions, gives a lower bound for the number of Gaussian primes lying in a given circle sector. As mentioned before, we could not find such a statement for $(3, 1)$-Gaussian primes in the literature. Other than the analogous statement being true, we also have some numerical evidence which backs up the conjecture. In the figure below we have graphed the number of $(3, 1)$-Gaussian primes under the curve $v = N/u$ up to $N = 5000$. The linearity of the graph needs no mention. The few lines of (Sage) code that produced this graph can be found in appendix C.



Leaving this difficulty behind, we are now going to establish upper bounds like we promised. To show that $\mathcal{L}$ is of the same size as $\mathcal{L}_0$, we are first going to prove that $\mathcal{L}_i$ for $i = 1, 2, 3$ are all negligibly small compared $\mathcal{L}_0$. Starting with the following.

Lemma 6.4: The following relation holds

$$\mathcal{L}_1 = o(N).$$

Proof: Remember that for a pair $(u, v) \in \mathcal{L}_0$, we have that $uv \leq N$. Therefore $u \leq \sqrt{N}$ or $v \leq \sqrt{N}$. A quick application of equation (5.2) from the proof of lemma 5.3 in the previous chapter, we find

$$\#\mathcal{L}_1 \leq \sum_{a \leq \sqrt{N}} \sum_{\substack{b \leq a^{-1}N \\ P^+(b) \leq N^{1/\log\log N}}} 1 = \sum_{a \leq \sqrt{N}} \Psi(N/a, N^{1/\log\log N}) \ll \sum_{a \leq \sqrt{N}} \frac{N}{a(\log N)^2} \ll \frac{N}{\log N} = o(N).$$

□

A similar result holds for the remaining two $\mathcal{L}_i$.

Lemma 6.5: The following relation holds

$$\mathcal{L}_i = o(N) \text{ for } i = 2, 3.$$

Proof: The two sets $\mathcal{L}_2$ and $\mathcal{L}_3$ are very similar, but because of the asymmetry of the equation $3u^2 + v^2$ they aren't equal. They are however similar enough that we will continue the proof under the assumption that $i = 2$. Only minimal changes are required to adapt the proof to work for $\mathcal{L}_3$.

It might be somewhat surprising that we will be able to very closely follow the steps we took in the previous chapter to establish the upper bound for $\mathcal{C}^*(N)$. After all, $\mathcal{L}_2$ contains numbers with a minimum number of distinct prime divisors while $\mathcal{C}^*(N)$ excludes numbers with too many distinct prime divisors. The key difference is that the restriction on the number of distinct prime divisors in $\mathcal{L}_2$ is there only for u instead of for the product uv. After the computations ahead, we will say a few more words about this, which will hopefully make it clear why this makes such a difference.

That an element $(u, v) \in \mathcal{L}_2$ cannot lie in $\mathcal{L}_1$ implies that u or v must contain a prime factor $\ell > N^{1/\log\log N}$. We can therefore bound $\#\mathcal{L}_2$ as

$$\#\mathcal{L}_2 \leq \sum_{\substack{uv \leq N^{1-1/\log\log N} \\ \omega(u) > \beta \log\log N}} S_1(u, v) + S_2(u, v)$$

where S_1 and S_2 are the same as on page 37. In lemma 5.4, we used Brun's sieve to bound S_1 and S_2. Applying this we further bound $\#\mathcal{L}_2$ as

$$\begin{aligned}\#\mathcal{L}_2 &\leq \frac{N(\log\log N)^{O(1)}}{(\log N)^2} \sum_{\substack{u\leq N^{1-1/\log\log N}\\ \omega(u)>\beta\log\log N}} \frac{1}{u} \sum_{v\leq u^{-1}N^{1-1/\log\log N}} \frac{1}{v} \\ &\leq \frac{N(\log\log N)^{O(1)}}{\log N} \sum_{\substack{u\leq N^{1-1/\log\log N}\\ \omega(u)>\beta\log\log N}} \frac{1}{u}. \end{aligned} \tag{6.2}$$

This sum of reciprocals can be bounded in more or less the same manner as we bounded I in lemma 5.5. Writing $T = \lfloor \beta \log\log N \rfloor$, we quickly go through the steps and now find

$$\begin{aligned}\sum_{\substack{u\leq N^{1-1/\log\log N}\\ \omega(u)>\beta\log\log N}} \frac{1}{u} &\leq \sum_{i>T} \frac{1}{i!} \left(\sum_{p^\nu \leq N} \frac{1}{p^\nu}\right)^i \\ &\leq \sum_{i\geq T} \frac{(\log\log N + c_1 + 1)^i}{i!} \\ &\ll \frac{(\log\log N + c_1 + 1)^T}{T!} \\ &\leq \left(\frac{e\log\log N + ec_1 + e}{T}\right)^T \\ &\ll \left(\frac{e}{\beta}\right)^T \\ &\ll (\log N)^{\beta(1-\log\beta)}. \end{aligned}$$

Note that the fact that only $\omega(u)$ is restricted manifests itself here. The power of the exponential which we are taking the tail end of its series expansion of, is a factor 2 smaller compared to when we bounded I (looking in particular at equation (5.5)). Along with this, the position of the large terms of the sum is shifted towards its beginning. Therefore we can still approximate this sum by its biggest term, even though we are looking at its tail end instead of at its first terms. Additionally note that $\beta > 1$ is a requirement to be able to make this approximation.

Substituting the bound back into (6.2), we conclude that

$$\#\mathcal{L}_2 \ll \frac{N\log\log N^{O(1)}}{(\log N)^{1-\beta(1-\beta\log\beta)}} = o(N)$$

as $1 - \beta(1 - \log\beta) > 0$.

□

We combine lemma 6.5 with lemma 6.4 and conjecture 6.3 to conclude that

$$\mathcal{L} \gg N. \tag{6.3}$$

The only set left that is still unbounded is $\mathcal{S}(N)$. We plan to establish an upper bound for it with yet another application of Brun's sieve. Because we now have to deal with two integer pairs at the same time, it will take some effort to arrive at something that is nicely siftable.

Note that $\mathcal{S}(N)$ contains $\mathcal{L}$ as a diagonal. Using that $u \leq \sqrt{N}$ or $v \leq \sqrt{N}$ as in lemma 6.4, and applying lemma 5.4, we bound $\#\mathcal{L}$ form above as follows

$$\begin{aligned}
\#\mathcal{L} &\leq \sum_{a \leq \sqrt{N}} \sum_{b \leq a^{-1} N^{1-1/\log\log N}} \big(S_1(a,b) + S_2(a,b)\big) \\
&\leq \frac{N(\log\log N)^{O(1)}}{(\log N)^2} \sum_{a,b \leq N} \frac{1}{ab} \\
&\leq N(\log\log N)^{O(1)}.
\end{aligned} \tag{6.4}$$

Note that according to (the data of) our conjecture this bound is a bit sloppy, but it will be good enough to be able to show that the diagonal, which satisfies the same bound, is negligibly small compared to the rest of $\mathcal{S}(N)$.

Ignoring the diagonal and factoring out a prime of size greater than $N^{1/\log\log N}$, we define

$$\mathcal{G} := \big\{(u,v,s,t) \in \mathbb{Z}^4_{\geq 1} \,\big|\, uv = st \leq N^{1-1/\log\log N}; u \neq s; \omega(u), \omega(v), \omega(s), \omega(t) \leq T\big\}$$

where $T = \lfloor \beta \log\log N \rfloor$, as before. With this, we have the following inequality

$$\#\mathcal{S}(N) \leq \#\mathcal{L} + \sum_{(u,v,s,t)\in\mathcal{G}} S_1'(u,v,s,t) + S_2'(u,v,s,t) + S_3'(u,v,s,t) + S_4'(u,v,s,t) \tag{6.5}$$

where

$$\begin{aligned}
S_1'(u,v,s,t) &:= \sum_{\substack{N^{1/\log\log N} < \ell \leq N/uv \\ \ell, 3u^2\ell^2+v^2, 3s^2\ell^2+t^2 \in \mathbb{P}}} 1, & S_2'(u,v,s,t) &:= \sum_{\substack{N^{1/\log\log N} < \ell \leq N/uv \\ \ell, 3u^2\ell^2+v^2, 3s^2+t^2\ell^2 \in \mathbb{P}}} 1, \\
S_3'(u,v,s,t) &:= \sum_{\substack{N^{1/\log\log N} < \ell \leq N/uv \\ \ell, 3u^2+v^2\ell^2, 3s^2\ell^2+t^2 \in \mathbb{P}}} 1, & S_4'(u,v,s,t) &:= \sum_{\substack{N^{1/\log\log N} < \ell \leq N/uv \\ \ell, 3u^2+v^2\ell^2, 3s^2+t^2\ell^2 \in \mathbb{P}}} 1.
\end{aligned}$$

Just like the S_j in the previous chapter, we can bound these S_j' using Brun's sieve. Having excluded elements from the diagonal ensures that the three primality conditions in the sums

defining S_j' are distinct. This results in a slightly improved bound compared to the ones we established for the S_j.

Lemma 6.6 : For $j = 1, 2, 3, 4$ we have the bound

$$S_j'(u, v, s, t) \ll \frac{N(\log\log N)^{O(1)}}{uv(\log N)^3}$$

Proof: As in the proof of lemma 5.4, the proofs for the different j are almost identical, so we will only treat the case $j = 1$. Furthermore, a lot of the parts of the setup of the sieve are the same as in lemma 5.4. We will omit detailed explanations accordingly.

We take $x = N/uv$, take $\mathcal{P} = \mathbb{P} = \{\text{all primes}\}$ and take $\mathcal{A}$ to be the characteristic function of the set

$$\{m(3u^2m^2 + v^2)(3s^2m^2 + t^2) \,|\, 1 \leq m \leq x\}.$$

We further put $z = N^{(\log\log N)^{-3}}$ and $X = x$.

The point where this proof diverges from that of lemma 5.4 is when determining the size of $A_d\big(x(3u^2x^2 + v^2)(3s^2x^2 + t^2)\big)$. We want to know, for a prime p, how many solutions to the equation $m(3u^2m^2 + v^2)(3s^2m^2 + t^2) \equiv 0 \pmod p$ exist in one residue system. The extra primality condition means that there are at most two extra solutions compared to lemma 5.4. As before the solution $m \equiv 0;$ always exists, whether there are zero, two or four more solutions depends on the existence and distinctness of solutions to the equations $3u^2m^2 + v^2 \equiv 0$ and $3s^2m^2 + t^2 \equiv 0$. The existence condition for both equations is the same as in lemma 5.4, two solutions for each equation exist exactly when $p \equiv 1 \pmod 3$ as long as p doesn't divide $3uv$. These pairs of solutions are distinct iff $-v^2/3u^2 \not\equiv -t^2/3s^2 \pmod p$, which happens iff p doesn't divide $ut \pm vs$. This motivates defining the density function as follows

$$g(p) = \begin{cases} \frac{1}{p} & \text{if } p \,|\, 3uv \text{ or } p \equiv 2 \pmod 3 \\ \frac{3}{p} & \text{if } p \nmid 3uv \text{ and } p \,|\, (ut \pm vs) \text{ and } p \equiv 1 \pmod 3 \\ \frac{5}{p} & \text{if } p \nmid 3uv \text{ and } p \nmid \pm vs \text{ and } p \equiv 1 \pmod 3. \end{cases}$$

It is easy to glance over the fact that this would be a problematic state of affairs if we didn't exclude the diagonal. To get the logarithm cubed in the bound from the statement, we basically need $g(p)$ to equal $3/p$ on average. An important part in coming to this average is that the second case $g(p) = 3/p$ happens only finitely often, which is true because $ut \pm vs$ is non-zero. For suppose that it does equal zero, then $ut = sv$ because both ut and sv are positive. A multiplication with u then shows that $u^2t = suv = s^2t$, which implies that $u = s$, but this is only possible for $\big((u, v), (s, t)\big)$ on the diagonal of $\mathcal{S}(N)$. Applying Mertens' theorem for arithmetic progressions in the same manner as in the proof of lemma 5.4 now yields

$$V(z) \asymp \frac{1}{(\log z)^3}.$$

The final steps from lemma 5.4 can now be repeated to conclude that

$$S_1'\big((u,v),(s,t)\big) \ll XV(z) \ll \frac{N(\log\log N)^{O(1)}}{uv(\log N)^3}.$$

□

By substituting the bounds from the lemma and equation (6.4) into (6.5) we obtain

$$\#\mathcal{S}(N) \leq N(\log\log N)^{O(1)} + \frac{N(\log\log N)^{O(1)}}{(\log N)^3}\mathcal{I} \tag{6.6}$$

where

$$\mathcal{I} = \sum_{(u,v,s,t)\in\mathcal{G}} \frac{1}{uv}.$$

After restructuring the sum slightly, we can bound $\mathcal{I}$ similarly to how we bounded I in lemma 5.5. Because the sum now has more terms, the bound becomes slightly worse.

Lemma 6.7: The following asymptotic bound holds

$$\mathcal{I} \ll (\log N)^{2\beta(1+\log 2-\log\beta)}.$$

Proof: Factorizing the elements of $\mathcal{G}$ will allow us to conveniently rewrite $\mathcal{I}$. Fix two pairs $\big((u,v)(s,t)\big) \in \mathcal{G}$ and let $g = \gcd(u,s)$. Then there are coprime integers x, y such that

$$u = gx \quad \text{and} \quad s = gy.$$

Because $uv = st$, there is an additional integer z such that

$$v = yz \quad \text{and} \quad t = xz,$$

where $z = \gcd(v,t)$ and the four integers g, x, y, z are pairwise coprime. This proves that the set

$$\mathcal{G}' := \big\{(g,x,y,z) \in \mathbb{Z}_{\geq 1}^4 \,\big|\, gxyz \leq N^{1-1/\log\log N}; x \neq y; \omega(gx), \omega(yz), \omega(gy), \omega(xz) \leq T\big\}$$

surjects onto $\mathcal{G}$. This allows us to bound $\mathcal{I}$ as follows

$$\mathcal{I} \leq \sum_{(g,x,y,z)\in\mathcal{G}'} \frac{1}{gxyz} \leq \sum_{i_1+\cdots+i_4\leq 2T} \prod_{j=1}^{4} \Bigg(\sum_{\substack{n\leq N \\ \omega(n)=i_j}} \frac{1}{n} \Bigg).$$

Using the multinomial theorem we can further bound this as

$$\begin{aligned}\mathcal{I} &\leq \sum_{i_1+\cdots+i_4\leq 2T} \prod_{j=1}^{4} \frac{1}{i_j!}\left(\sum_{p^\nu\leq N}\frac{1}{p^\nu}\right)^{i_j} \\ &= \sum_{i_1+\cdots+i_4\leq 2T} \frac{1}{(i_1+i_2+i_3+i_4)!}\binom{i_1+i_2+i_3+i_4}{i_1,i_2,i_3,i_4}\left(\sum_{p^\nu\leq N}\frac{1}{p^\nu}\right)^{i_1+i_2+i_3+i_4}.\end{aligned}$$

After making the substitution $n = i_1 + \cdots + i_4$ and applying the multinomial identity $4^n = \sum_{i_1+\cdots+i_4=n}\binom{n}{i_1,\ldots,i_4}$, we follow the final few steps from the proof of lemma 5.5 (omitting details and using the same constant c_1), to conclude that

$$\begin{aligned}\mathcal{I} &\leq \sum_{n\leq 2T}\frac{1}{n!}\left(4\sum_{p^\nu\leq N}\frac{1}{p^\nu}\right)^n \\ &\leq \sum_{n\leq 2T}\frac{(4\log\log N+4c_1+4)^n}{n!} \\ &\ll \frac{(4\log\log N+4c_1+4)^{2T}}{(2T)!} \\ &\leq \left(\frac{4e\log\log N+4ec_1+4e}{2T}\right)^{2T} \\ &\leq \left(\frac{2e}{\beta}+O\left(\frac{1}{T}\right)\right)^{2T} \\ &\ll (\log N)^{2\beta(1+\log 2-\log\beta)}\end{aligned}$$

Note that it is here that we use $\beta < 2$ to be able to approximate the sum by its largest term. □

Note that $2\beta(1+\log 2-\log\beta) > 3$, so that substituting this into equation (6.6) finalizes our bound for $\mathcal{S}(N)$ as

$$\#\mathcal{S}(N) \leq N(\log\log N)^{O(1)}(\log N)^{2\beta(1+\log 2-\log\beta)-3}.$$

By combining this with lemma 6.2 and inequality (6.3), and making the substitution $\beta = 1+\varepsilon$, we deduce theorem 6.1.

7. Closing comments

There are two possible improvements of the results in this thesis that we want to highlight. The first has already been mentioned in the previous chapters: it is the improvement of the upper bound by removing the factor $(\log\log N)^{O(1)}$ and the improvement of the lower bound by a factor of $\log N$ to bring it up to the anticipated upper bound. The sole reason we expect this to be possible is that the authors of [ChPo17] formulated an analogous conjecture in their final chapter. For a suggestion of a possible route leading to these improved bounds, we refer to their comments at the end of their paper.

The second improvement we have in mind is to determine whether there are (and if so which) other angles γ that correspond to triangles whose side values behave in the same manner as right triangles and the triangles in $\mathcal{T}$.

Some (limited) numerical experimentation suggests that for a general fraction $\cos\gamma = s/t$, things do not necessarily work as nicely as for $\cos\gamma = 0$ and $\gamma = 1/2$. The law of cosines becomes

$$ta^2 + tb^2 - 2sab = tp^2.$$

For all the fractions that we checked (all the reduced fractions $s/t < 1$ with $t \leq 6$) there seem to be plenty of solutions. Additionally, the set of primes for which there is a solution seems predictable as well.

To illustrate this we will consider the example value $s/t = 1/3$. This will bring some fundamental differences to light, but also show enough similarities to suggest that a generalization to 'nicely behaved', if not all, fractions s/t is possible.

When trying to replicate chapter 4 with $s/t = 1/3$, things only seem to half work. The core of the problem lies in the fact that the statement analogous to lemma 4.1 only works in one direction. Given a triangle, the trick with the stereographic projection does dictate that there are coprime positive integers x, y such that the prime side is of the form $p = (2x^2 + y^2)/3$, one of the legs is of the form xy and the remaining leg equals one of the following expressions

$$\frac{2x^2 - y^2 + xy}{3} \qquad \text{or} \qquad \frac{y^2 - 2x^2 + xy}{3}$$

Given such a coprime pair x, y for which $(2x^2 + y^2)/3$ is a prime, it is not always the case that choosing one of above expressions always gives back a triangle with integer sides though. We have written down a few numerical examples below which support these claims, as well as bring other important differences to light.

	$x=2,$ $y=1$	$x=4,$ $y=1$	$x=2,$ $y=5$	$x=5,$ $y=1$	$x=1$ $y=7$	$x=7$ $y=5$	$x=1$ $y=11$
$(2x^2+y^2)/3$	3	11	11	17	17	41	41
$(2x^2-y^2+xy)/3$	3	35/3	-7/3	18	-40/3	36	-36
$(y^2-2x^2+xy)/3$	5/3	-9	9	-44/3	18	-38/3	130/3
xy	2	4	10	5	7	35	11

For example, the column $x=4, y=1$ does not even feature three positive integers, so it cannot correspond to a triangle with integer sides.

The contents of the bottom three rows do not seem as nicely ordered as in chapter 4. First of all the distribution of even/odd integers is no longer predictable. We have to look modulo 3 instead to bring back structure. Ignoring the fractions and negative numbers in the middle two rows leaves us exactly with the integers divisible by 3 which feature as a leg. This suggests that we can define a set $\mathcal{A}(N)$ which counts exactly the legs divisible by 3 in a very similar manner to how we defined the set $\mathcal{A}(N)$ in chapter 4 which exactly contained the even legs (remember that in chapter 4 we had to similarly filter out odd numbers from the middle two rows). The appropriate definition seems to be

$$\mathcal{A}(N) = \{uv \leq N \,|\, 2u^2+v^2 \in \mathbb{P}\}.$$

If we analogously to chapter 4 define $\mathcal{B}(N)$ to contain all the integers from the bottom row, we run into a bit of a problem though. Not all the integers in this row (e.g. 4 and 11) seem to feature in a triangle with only integer sides. This means that if we take the nice definition

$$\mathcal{B}(N) := \{xy \leq N \,|\, 2x^2+y^2 \in 3\mathbb{P}\},$$

that $\mathcal{B}(N)$ gets contaminated with 'bad' numbers that we are not interested in. This might not be a huge problem though. For the primes of the form $(2x^2+y^2)/3$ (which are exactly the primes $1,3 \bmod 8$) there seems to be at least one such pair x, y (while there are typically two pairs) that does correspond to a triangle with integer sides. This suggests that $\mathcal{B}(N)$ is at most half-filled with 'bad' numbers. Since the bounds we established are asymptotically accurate up to an integer anyway, this factor $1/2$ wouldn't meaningfully affect the results.

The thing that changes the most when adapting the proof of the upper and lower bounds to work for these new definitions of $\mathcal{A}(N)$ and $\mathcal{B}(N)$ is the application of Brun's sieve, which should be able to handle modifications like these quite well.

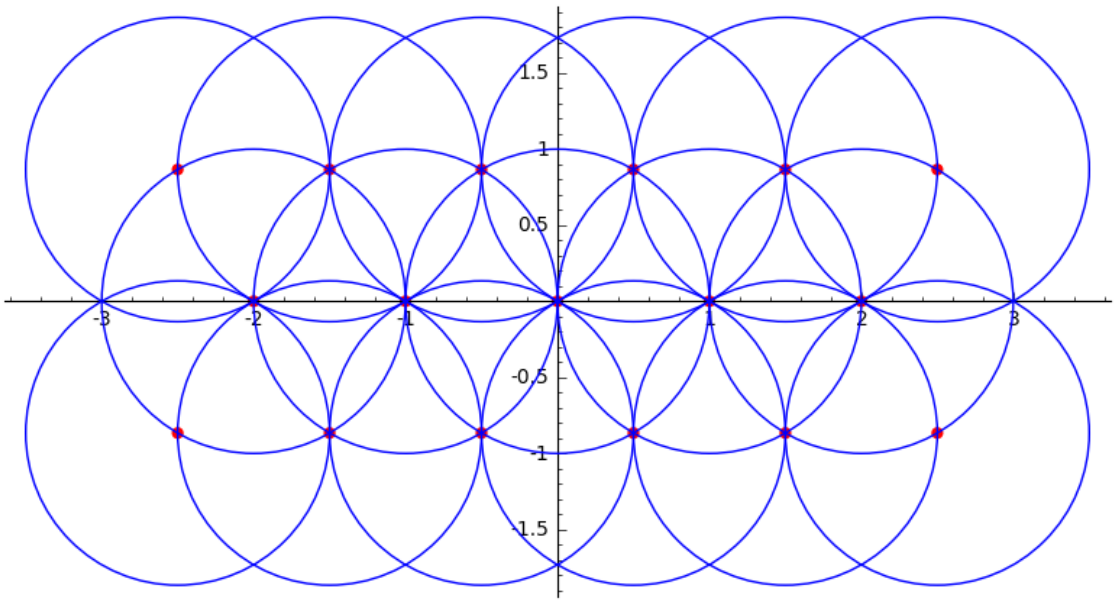
All of this bodes well for generalization to a wider range of angles. How wide this range is exactly needs further investigation though. While the problems we run into with $s/t = 1/3$ seem manageable, other things might start breaking down when (one of) s and t gets larger.

8. Appendix

A

In this section we provide a proof of the fact that $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ is a unique factorization domain. We will do this by directly proving that it admits a Euclidean algorithm It is then automatically a principal ideal domain and hence has unique factorization.

Given two elements α, β in $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ with $\beta \neq 0$, we need to confirm that there exist two more elements q, r in $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ such that $\alpha = q\beta + r$ and such that $|r| < |\beta|$. If we divide by β this translates into $\alpha/\beta = q + r/\beta$ with $|r/\beta| < 1$. From this we deduce that the condition is fulfilled if every fraction α/β in $\mathbb{Q}(\sqrt{-3})$ can be approximated by an element q in $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ such that $|q - \alpha/\beta| < 1$. This is indeed possible because the collection of open discs of radius 1 centered at the elements of $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ covers the entire complex plane, as can be deduced from the picture below.



B

In this section we will justify the approximations we made of the partial exponential sums throughout chapters 5 and 6. It suffices to prove the following statement.

Lemma B1: For $\eta > 1 > \mu$, the asymptotic inequalities

$$\sum_{m \leq L} \frac{(\eta L)^m}{m!} \ll \frac{(\eta L)^L}{L!} \quad \text{and} \quad \sum_{j \geq L} \frac{(\mu L)^j}{j!} \ll \frac{(\mu L)^L}{L!}$$

hold as L grows to ∞.

△

The truth of this statement relies on the way the size of the exponential is distributed over the terms of its Taylor expansion. For an integer n, by far the largest term of the expansion of e^n is the term $n^n/n!$, Stirlings formula approximately says that the entire series is only a factor $\sqrt{2\pi n}$ larger. The terms are almost normally distributed around the largest term $n^n/n!$, though not quite like a Gaussian e^{-x^2}, but more like $e^{-x \log x}$. This will still mean that the largest term in the sums in lemma B1 is so dominant that it asymptotically approaches the entire sum. The symmetrical character of the distribution is also reflected in the following related asymptotic equality

$$\sum_{n \leq L} \frac{L^n}{n!} \sim e^L/2.$$

One way to properly prove this is through a technique known as Laplace's method, which can be used to approximate the integral representation of the remainder of this truncated Taylor series. We will prove lemma B1 basically following the same idea, but we need a slight variation of the statement that is commonly referred to as Laplace's method. The statement we will use is the following, we will also provide the proof.

Lemma B2: (Laplace's method) Let $U \subseteq \mathbb{R}$ be an open that contains the interval $[a, b)$ and let $f : U \to \mathbb{R}$ a continuously differentiable function decreasing at a and such that for every $\delta > 0$, the supremum of f on $(a + \delta, b)$ is strictly less than $f(a)$ and such that $\int_a^b e^{f(x)} dx$ converges. Then

$$\int_a^b e^{Mf(x)} dx \sim \frac{e^{Mf(a)}}{M|f'(a)|}$$

as M goes to ∞.

Proof: To prove that the $\sim$ relation holds we will show that the limit of the quotient of the RHS and LHS is bounded from below and above by 1.

We first prove the lower bound. Let $\varepsilon > 0$, by Taylor's theorem there is a $\delta > 0$ such that for $x \in (a, a+\delta)$ the inequality

$$f(x) \geq f(a) + (f'(a) - \varepsilon)(x - a)$$

holds. From this we derive the following inequality

$$\begin{aligned}\int_a^b e^{Mf(x)} dx &\geq \int_a^{a+\delta} e^{M(f(a)+(f'(a)-\varepsilon)(x-a))} dx \\ &= e^{Mf(a)} \frac{e^{-\delta M(\varepsilon - f'(a))} - 1}{M(f'(a) - \varepsilon)}\end{aligned}$$

Dividing both sides of this inequality by $e^{Mf(a)}/M|f'(a)|$ and taking the limit $M \to \infty$, we find

$$\lim_{M\to\infty} \frac{\int_a^b e^{Mf(x)} dx}{e^{Mf(a)}/M|f'(a)|} \geq \frac{|f'(a)|}{f'(a) - \varepsilon} \lim_{M\to\infty} (e^{-\delta M(\varepsilon - f'(a))} - 1) = \frac{|f'(a)|}{\varepsilon - f'(a)}.$$

Because this is true for any $\varepsilon > 0$ and $f'(a)$ is negative by assumption, we may conclude that

$$\lim_{M\to\infty} \frac{\int_a^b e^{Mf(x)} dx}{e^{Mf(a)}/M|f'(a)|} \geq 1.$$

To prove that the upper bound also holds we will follow a similar strategy, but we need to work around the fact that we will not be able to ignore the remainder part $\int_{a+\delta}^b$ of the larger integral like before.

Let $\varepsilon > 0$, by Taylor's theorem there is a $\delta > 0$ such that for $x \in (a, a+\delta)$ the inequality

$$f(x) \leq f(a) + (f'(a) + \varepsilon)(x - a)$$

holds. Because of the condition on the supremum, there must exist some $\gamma > 0$ such that $f(x) \leq f(a) - \gamma$ for all $x \in (a+\delta, b)$. From this we now derive the inequality

$$\begin{aligned}\int_a^b e^{Mf(x)} dx &\leq \int_a^{a+\delta} e^{M(f(a)+(f'(a)+\varepsilon)(x-a))} dx + \int_{a+\delta}^b e^{Mf(x)} dx \\ &\leq e^{Mf(a)} \frac{e^{-\delta M(-f'(a)-\varepsilon)} - 1}{M(f'(a) + \varepsilon)} + e^{(M-1)(f(a)-\gamma)} \int_{a+\delta}^b e^{f(x)} dx.\end{aligned}$$

Like before we want to divide both sides by $e^{Mf(a)}/M|f'(a)|$ and let $M \to \infty$. Note that the right summand on the right hand side vanishes in this process as $e^{-M\gamma} M \cdot \text{constant} \to 0$ as

$M \to \infty$. So, ignoring this summand and proceeding as described, we get close to the situation we had before, we now find

$$\lim_{M\to\infty} \frac{\int_a^b e^{Mf(x)}dx}{e^{Mf(a)}/M|f'(a)|} \leq \frac{|f'(a)|}{-\varepsilon - f'(a)}.$$

Because this is true for every $\varepsilon > 0$, we may conclude that

$$\lim_{M\to\infty} \frac{\int_a^b e^{Mf(x)}dx}{e^{Mf(a)}/M|f'(a)|} \leq 1.$$

□

Note that a mirrored version where $f(b)$ is the maximum instead of $f(a)$ and where $f'(b) > 0$ follows by symmetry. We will use both versions to prove lemma B1.

Proof of lemma B1: The remainder of the n-th order Taylor expansion of a function f is

$$\int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt.$$

We write $x = \eta L$ and use this to rewrite first of the two sums from the lemma as follows

$$\begin{aligned}
\sum_{m\leq L} \frac{x^m}{m!} &= e^x - \int_0^x \frac{(x-t)^L}{L!} e^t dt \\
&= e^x - \frac{e^x}{L!} \int_0^x y^L e^{-y} dy \\
&= e^x - \frac{e^x}{L!} \left(L! - \int_x^\infty y^L e^{-y} dy \right) \\
&= \frac{e^x}{L!} \int_x^\infty y^L e^{-y} dy.
\end{aligned}$$

We can use Laplace's method to find a useful asymptotic for the integral. After carrying out the substitution $y = Lz$ and applying Laplace, we find the following

$$\begin{aligned}
\int_x^\infty y^L e^{-y} dy &= L^{L+1} \int_\eta^\infty z^L e^{-Lz} dz \\
&= L^{L+1} \int_\eta^\infty e^{L(\log z - z)} dz \\
&\sim L^{L+1} \frac{e^{L(\log \eta - \eta)}}{L|\eta^{-1} - 1|} \\
&= (\eta L)^L e^{-x} \frac{\eta}{\eta - 1}
\end{aligned}$$

substituting this back into the above, we find

$$\sum_{m\leq L}\frac{(\eta L)^m}{m!}\sim\frac{\eta}{\eta-1}\cdot\frac{(\eta L)^L}{L!}\ll\frac{(\eta L)^L}{L!}$$

which is exactly the approximation we wanted to establish for the first sum.

The second sum we want to approximate already is the remainder part of the Taylor series. In order to get nicer formulas we ignore the first term of the sum for now, we will add it back at the end. Writing $x=\mu L$ and making similar substitutions as before, we now see that

$$\begin{aligned}\sum_{j\geq L+1}\frac{(\mu L)^j}{j!}&=\int_0^x\frac{(x-t)^L}{L!}e^t dt\\&=\frac{e^x}{L!}\int_0^x y^L e^{-y}dy.\end{aligned}\tag{8.1}$$

We substitute $y=Lz$ and apply the mirrored version of the Laplace method to find the following asymptotic for the integral

$$\begin{aligned}\int_0^x y^L e^{-y}dy&=L^{L+1}\int_0^\mu z^L e^{-Lz}dz\\&=L^{L+1}\int_0^\mu e^{L(\log z-z)}dz\\&\sim L^{L+1}\frac{e^{L(\log\mu-\mu)}}{L|\mu^{-1}-1|}\\&=(\mu L)^L e^{-x}\frac{\mu}{1-\mu}.\end{aligned}$$

By substituting this back into (8.1), we obtain

$$\sum_{j\geq L+1}\frac{(\mu L)^j}{j!}\sim\frac{(\mu L)^L}{L!}\cdot\frac{\mu}{1-\mu}\ll\frac{(\mu L)^L}{L!}.$$

Adding back to both sides the term $(\mu L)^L/L!$ which we chose to ignore at the start, we conclude that the second inequality from the lemma also holds.

□

C

```
L = []
for N in range(1,5000):
    Counter = 0
    if (N/10).is_integer():
        for u in range(1,N):
            for v in range(1,floor(N/u)):
                if (3*u^2 + v^2) in Primes():
                    Counter = Counter+1
        L.append([N,Counter])
list_plot(L, color = 'blue')
```

This piece of Sage code was used to generate the graph on page 46. Note that for this range of N, this computation might take a few minutes on a present-day consumer grade PC.

9. References

Books

[OdC10] Opera de Cribro - Friedlander, J.B., Iwaniec, H - American Mathematical Society Colloquium Publications Volume 57 - (2010)

Papers

[Brui51] de Bruijn, N.G. - On the number of positive integers $\leq x$ and free of prime factors $> y$ - Proceedings of the Koninklijke Nederlandse Akadademie van Wetenschappen Series A **54** (1), 50–60 - (1951)

[ChPo17] Chow, S., Pomerance C. - Triangles with prime hypothenuse - Research in Number Theory **3**, 21 - (2017)

[Erdö60] Erdős, P. - An asymptotic inequality in the theory of numbers - Vestnik Leningrad University **15**, 41–49 - (1960)

[Ford08] Ford, K. - The distribution of integers with a divisor in a given interval - Annals of Mathematics **168** (2), 367–433 - (2008)

[HaRa17] Hardy, G.H., Ramanujan, S. - The normal number of prime factors of a number n - Quarterly Journal of Mathematics, XLVIII, 76–92 - (1917)

[Tene84] Tenenbaum, G. - Sur la probabilité qu'un entier possède un diviseur dans un intervalle donné - Compositio Mathematica **51** (2), 243–263 - (1984)

[Wile95] Wiles, A. - Modular elliptic curves and Fermat's Last Theorem - Annals of Mathematics **142**, 443-551 - (1995)

[Will74] Williams, K.S. - Mertens' theorem for arithmetic progressions - Journal of Number Theory **6**, 353-359 - (1974)