Quadratic Large Sieves for "Algebraically Impossible" Sets

Jaeho Lee

Under the direction of

Samuel Packman
Department of Mathematics
Massachusetts Institute of Technology

Research Science Institute
July 29, 2025

Abstract

Linnik's large sieve provides a bound for the sifted set X constructed by removing a positive proportion of residue classes modulo p. We consider quadratic sieves that classify quadratic residues and take Boolean variations of that sieve. We show that the number of $n \leq N$ such that n and $n+k$ are quadratic residues modulo p for $N^{1/m} < p < 2N^{1/m}$ for some positive integer m is bounded by $(\log N)^{4m+1}$. We then work with the case where at least one of $n+k_i$ for some set of constants k_i is a quadratic residue modulo p for all $N^{1/m} < p < 2N^{1/m}$, showing that the asymptotic bound is identical to the case where n must be a quadratic residue.

Summary

We bound sets related to quadratic residues modulo p, which are elements such that there exists a square number with the same remainder as the element when divided by p. We begin by showing that the number of integers n such that n and $n+k$ are quadratic residues for a set of primes p is as close to a constant as desired. This creates a parallel between the integers and modular arithmetic for a large number of primes, since there are only finitely many integers n such that both n and $n+k$ are squares. We then consider integers n such that at least one of n or $n+k$ is a quadratic residue and show that it is equivalent to the case where only n has to be a quadratic residue for the same set of primes primes.

1 Introduction

Many of the famous theorems and conjectures of number theory can be rewritten in terms of estimating the size of some set; for example, the prime number theorem establishes an asymptotic behavior on the size of the set of primes up to some integer N. Many of these sets share similar characteristics when viewed in terms of residue classes for some prime p, suggesting that they can be obtained by considering some set that only preserves specific elements modulo p that fit certain criteria. This is the essential idea behind sieve theory.

The idea of sieves has existed since Ancient Greece. The Sieve of Eratosthenes [1], which was first described in the 3rd century BCE, leaves only the primes less than N. Other sieves have been fruitfully applied to tackle some famous conjectures. For example, Brun's sieve settled in the affirmative the convergence of the sum of reciprocals of twin primes [2], while Chen's theorem showed that there are infinitely many primes p such that $p + 2$ is either a prime or a semiprime (a multiple of two primes).

The large sieve was first formalized by Linnik [3] in 1941 to study least quadratic non-residues. Since then, the theory of large sieves has been employed as a method in analytic number theory. For example, Rényi [4] used large sieves in 1948 to show that for large enough n, we can write $2n = p + P_k$, where p is some prime and P_k is some product of at most k primes. Further work using the large sieve [5] obtained a value of $k = 2$, which remains one of the strongest results regarding the famous Goldbach's conjecture. Large sieves remain an area of modern research interest.

In this paper, we bound the sifted sets resulting from Boolean variations of the quadratic large sieve. Quadratic large sieves leave only the integers that are a quadratic residue for every prime in some set of primes. We begin with "algebraically impossible" sets, which are sets with conditions satisfied by only finitely many integers when considered over $\mathbb{Z}$. However, these conditions may have infinitely many solutions in modular arithmetic. For example, while no positive integer can satisfy $x^2 + 1$ being a perfect square, it is possible that $x^2 + 1$ is a quadratic residue modulo p for some set of primes p; for instance, 1 and 2 are consecutive quadratic residues for all primes p that are 1 or 7 modulo 8. We also consider less restrictive variants of the quadratic sieve. Specifically, we can consider the sifted set where for some fixed constants $k_1, \ldots, k_a$, the number n is sifted only if every $n + k_i$ is a quadratic non-residue for some prime p.

We begin in Section 2 by introducing the notation used in the paper, then establish the preliminary lemmas and motivating theorems in Section 3. Then, in the next two sections, we bound sifted sets from Boolean variations of the quadratic large sieve. In Section 4, we

bound sets with algebraically impossible conditions and show that they remain improbable to achieve in modular arithmetic. Lastly, in Section 5, we bound the sifted set with the condition that at least one $n + k_i$ for some set of constants k_i is a quadratic residue for each prime.

2 Definitions

In this section, we introduce the notation that we use throughout the paper. We begin by recalling notation in number theory. We write $\mathbb{Z}/n\mathbb{Z}$, the set of residue classes modulo n, as $\mathbb{Z}_n$. Moreover, we write $[N] := \{1, \ldots, N\}$. We now define notation regarding squares and quadratic residues.

Definition 2.1. Let d be an integer. Then, d is *square-free* if $k^2 \mid d$ implies $k = 1$. We call d the *square-free part of* n if $n = c^2 d$ for d square-free.

For ease of notation, we will write $a = \square$ to mean that a is a square number and $a \neq \square$ to mean that a is not a square number. We now introduce the Legendre symbol, which states whether an integer is a quadratic residue modulo p, as well as its extension, the Jacobi symbol.

Definition 2.2. Let $n \in \mathbb{Z}$ and let p be an odd prime. We define the *Legendre symbol* as

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{if } p \mid n, \\ 1 & \text{if } p \nmid n \text{ and } n \equiv x^2 \bmod p \text{ for some } x, \\ -1 & \text{if else.} \end{cases}$$

Definition 2.3. Let $n \in \mathbb{Z}$ and let $q = p_1^{a_1} \ldots p_k^{a_k}$ be odd. We define the *Jacobi symbol* as

$$\left(\frac{n}{q}\right) = \prod_{i=1}^{k} \left(\frac{n}{p}\right)^{a_i}.$$

Notice that the Jacobi symbol is equivalent to the Legendre symbol when q is prime. The Jacobi symbol is a *Dirichlet character*, which we now define.

Definition 2.4. Consider a function $\chi_m : \mathbb{Z} \to \mathbb{C}$. Then, the function χ_m is a *Dirichlet character modulo* m if for all integers a and b, we have

1. $\chi_m(ab) = \chi_m(a)\chi_m(b)$

2. $\chi_m(a) = 0$ if and only if $\gcd(a, m) > 1$

3. $\chi_m(a + m) = \chi_m(a)$.

The most simple Dirichlet character is the *primitive character*.

Definition 2.5. The character $\chi_{m,0} : \mathbb{Z} \to \{0, 1\}$ such that

$$\chi_{m,0}(a) = \begin{cases} 0 & \text{if } \gcd(a, m) > 1, \\ 1 & \text{if } \gcd(a, m) = 1 \end{cases}$$

is called the *primitive character*.

Notice that the Jacobi symbol is a Dirichlet character [6].

2.1 Sieves

We now define the sieve. Broadly, a sieve removes some set of residue classes for some set of primes from the original set, leaving us with the sifted set.

Definition 2.6. Let $N \in \mathbb{N}$, let P be a set of primes, and let E_p be a set of residue classes modulo p. Then, we define the *sifted set* X as

$$X = \{n \in [N] \mid n \not\equiv k \pmod{p} \text{ for any } k \in E_p \text{ for any } p \in P\}.$$

In this paper, we study sieves that remove a positive proportion of the residue classes modulo p for each prime.

Remark. We often write our sifted sets in terms of the elements $x \in X$ rather than in terms of the sets E_p. One may note that it is possible to state every sifted set in this paper in the form of Definition 2.6.

An example of a sieve is the sieve that removes all quadratic nonresidues for every prime p, which results in the sifted set being the set of squares. Given some value of N, the size of this sifted set is approximately $N^{1/2}$. We now formalize this estimate following the notation of [7].

Definition 2.7 ([7])**.** Consider functions $f, g : \mathbb{N} \to \mathbb{R}^+$. Then, we write $f \asymp g$ if and only if there exist positive constants c and C such that $cg(n) \leq f(n) \leq Cg(n)$ for all $n \in \mathbb{N}$. Similarly, we write $f \ll g$ if there exists a constant C such that $f(n) \leq Cg(n)$ for all $n \in \mathbb{N}$

and $f \gg g$ if there exists a constant c such that $f(n) \geq Cg(n)$ for all $n \in \mathbb{N}$. If $f \ll g$, we write $f = O(g)$.

We also work with primes bounded by a constant factor of M.

Definition 2.8. Let $M \in \mathbb{Z}$. We define $P_M := \{p \text{ prime} \mid M \leq p \leq 2M\}$.

Remark. The bound of M and $2M$ for P_M are not significant and can be replaced by cM and CM, respectively.

With the established notation, we now present bounds for sifted sets of quadratic residues. Other definitions and notation necessary for algebraically impossible quadratic sieves, which have properties satisfied by only finitely many integers over $\mathbb{Z}$, are later introduced in Section 4.

3 Preliminary Lemmas

In this section, we introduce the preliminary lemmas and other results that are used later in the paper.

3.1 Quadratic Reciprocity

We begin by presenting the general version of quadratic reciprocity.

Lemma 3.1 ([6])**.** *Let P and Q be odd numbers such that* $\gcd(P, Q) = 1$*. Then,*

$$\left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = (-1)^{\frac{P-1}{2}\frac{Q-1}{2}}.$$

Moreover, we have

$$\left(\frac{2}{P}\right) = (-1)^{\frac{P^2-1}{8}}$$

The version of quadratic reciprocity presented above, however, means that

$$\left(\frac{P}{Q}\right) = -\left(\frac{Q}{P}\right)$$

when $P, Q \equiv 3 \pmod 4$. We wish to find some character such that it is always equal to the value of $\left(\frac{P}{Q}\right)$ for all P and Q, which would, for example, allow us to switch from a complicated modulus to a simpler one. To do this, we define a new character.

Definition 3.2. We define the function $\chi_{4n} : \mathbb{Z} \to \{-1, 0, 1\}$ as

$$\chi_{4n}(q) = \begin{cases} \left(\frac{n}{q}\right) & \text{if } q \equiv 1 \pmod 2, \\ 0 & \text{if } q \equiv 0 \pmod 2. \end{cases}$$

Lemma 3.3. *The function χ_{4n} is a non-primitive Dirichlet character modulo $4n$.*

Proof. We begin by noticing that $\chi_{4n}(q) = 0$ if and only if $\gcd(4n, q) > 1$. Indeed, if q is even then $\chi_{4n}(q) = 0$ by definition, and if $\gcd(n, q) > 1$ then $\chi_{4n}(q) = \left(\frac{n}{q}\right) = 0$. Moreover, it is multiplicative, since for $\gcd(4n, p) = \gcd(4n, q) = 1$, we have

$$\chi_{4n}(p)\chi_{4n}(q) = \left(\frac{n}{p}\right)\left(\frac{n}{q}\right) = \left(\frac{n}{pq}\right) = \chi_{4n}(pq).$$

Now, we show that χ_{4n} is $4n$-periodic. First, let $n = 2^{2k}m$ with m odd. Then, it follows from Lemma 3.1 that

$$\chi_{4n}(q) = \left(\frac{2^{2k}m}{q}\right) = \left(\frac{m}{q}\right) = \left(\frac{q}{m}\right)(-1)^{\frac{q-1}{2}\frac{m-1}{2}}.$$

Since $\left(\frac{q}{m}\right)$ is m-periodic and $(-1)^{\frac{q-1}{2}}$ is 4-periodic, the function $\chi_{4n}(q)$ is $4m$-periodic. Since $m \mid n$, the function $\chi_{4n}(q)$ is also $4n$-periodic.

Now, let $n = 2^{2k+1}m$ with m odd. Then, it follows from Lemma 3.1 that

$$\chi_{4n}(q) = \left(\frac{2^{2k+1}m}{q}\right) = \left(\frac{2}{q}\right)\left(\frac{m}{q}\right) = \left(\frac{q}{m}\right)(-1)^{\frac{q-1}{2}\frac{n-1}{2}+\frac{q^2-1}{8}}.$$

Since $\left(\frac{q}{m}\right)$ is m-periodic and $(-1)^{\frac{q-1}{2}\frac{n-1}{2}+\frac{q^2-1}{8}}$ is 8-periodic, the function $\chi_{4n}(q)$ is $8m$-periodic. Since $8m \mid 4n$, the function $\chi_{4n}(q)$ is also $4n$-periodic. ∎

Moreover, χ_{4n} follows the same orthogonality property as the Jacobi symbol.

Lemma 3.4. *The character $\chi_{4n}\chi_{4m}$ is primitive if and only if nm is a square number.*

Proof. Notice that

$$\chi_{4n}\chi_{4m}(q) = \left(\frac{nm}{q}\right) = \left(\frac{4nm}{q}\right),$$

so $\chi_{4n}\chi_{4m}$ is primitive if and only if $\chi_{q,2}(4nm) = 1$ for all q with $\gcd(q, nm) = 1$. However, this means that nm must be a square number, so $\chi_{4n}\chi_{4m}$ is primitive if and only if $nm = k^2$ for some $k \in \mathbb{Z}$. ∎

This is analogous to the fact that $\left(\frac{n}{q}\right) = 1$ for all n with $\gcd(n, q) = 1$ if q is a square number.

3.2 Quadratic Sieves

We now look at quadratic sieves, which has the sifted set X as

$$X = \left\{ n \in [N] \mid \left(\frac{n}{p}\right) = 1 \text{ for all } p \in P_M \right\}$$

for some M. Recall that P_M is the set of primes such that $M \leq p \leq 2M$. We begin by presenting two large sieve inequalities due to Linnik [3].

Theorem 3.5 ([3])**.** *Let $N \in \mathbb{Z}$ and consider all primes $p < N^{1/2}$. Moreover, let $\varepsilon \in \mathbb{R}$ be a fixed constant and let $|E_p| > \varepsilon p$ for all p. Then the set S sifted by E_p has size $|S| \ll N^{1/2}$.*

Theorem 3.6 (Linnik multiplicative large sieve [3])**.** *Let $P, Q \in \mathbb{Z}$. Moreover, let $f : \mathbb{N} \to \mathbb{C}$ be a function with $f(n) = 0$ for all $n > N$ and let φ denote the Euler totient function. Then,*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sum_{\substack{\chi \neq \chi_0 \\ mod\ q}} \left| \sum_{n=1}^{N} f(n)\chi_q(n) \right|^2 \leq (N + Q^2)\|f\|_{L^2}^2.$$

The multiplicative large sieve inequality allows us to find the distribution of quadratic residues modulo p for all primes $p \in P_M$. We wish to show that the set of quadratic residues is similar in size to the set of squares, which is the statement of Theorem 3.7. The following theorem follows from Lemma 6.1 in [8].

Theorem 3.7. *Let $N \in \mathbb{Z}$ and $M = N^{1/m}$ for some $m \in \mathbb{Z}$ with $m \geq 1$. Consider the set $[N]$ and all primes $p \in P_M$. Then, the set*

$$X = \left\{ n \in [N] \mid \left(\frac{n}{p}\right) = 1 \text{ for all } p \in P_M \right\}$$

has size $|X| \ll N^{1/2}(\log N)^{2m}$.

Proof. We begin by considering the set $Y \subseteq X$ defined as

$$Y = \{n \in X \mid n \text{ square-free}\}.$$

Since there are at most $N^{1/2}$ squares less than N, we have $|X| \ll N^{1/2}|Y|$, so it is sufficient to show that $|Y| \ll (\log N)^{2m}$. Now, consider the set

$$\mathscr{L} = \left\{ q \in \mathbb{Z} \mid q = \prod_{i=1}^{2m} p_i,\ p_i \neq p_j \text{ for } i \neq j,\ p_i \in P_M \right\},$$

where $q < 2^{2m}N^2$ for any $q \in \mathscr{L}$ from $p_i \leq 2N^{1/m}$. It follows from the prime number theorem that $|\mathscr{L}| \asymp \left(\frac{M}{\log M}\right)^{2m}$. Since $n \in Y$ implies that $\left(\frac{n}{q}\right) = \prod_{i=1}^{2m} \left(\frac{n}{p_i}\right) = 1$ for all $q \in \mathscr{L}$, we add over all $q \in \mathscr{L}$ to get

$$\left(\frac{M}{\log M}\right)^{2m} \mathbb{1}_Y(n) \ll \left| \sum_{q \in \mathscr{L}} \left(\frac{n}{q}\right) \right|.$$

After squaring both sides and summing over all square-free n, it follows that

$$\left(\frac{M}{\log M}\right)^{4m} |Y| \ll \sum_{\substack{n=1 \\ n \text{ square-free}}}^{N} \left| \sum_{q \in \mathscr{L}} \left(\frac{n}{q}\right) \right|^2.$$

By the statement of quadratic reciprocity in Lemma 3.3, we have $\left(\frac{n}{q}\right) = \chi_{4n}(q)$. Then, we can rewrite the sum over $q \in \mathscr{L}$ as a sum from 1 to $2^{2m}N^2$ and use an indicator function on $\mathscr{L}$, from which we have

$$\left(\frac{M}{\log M}\right)^{4m} |Y| \ll \sum_{\substack{n=1 \\ n \text{ square-free}}}^{N} \left| \sum_{\ell=1}^{2^{2m}N^2} \mathbb{1}_{\mathscr{L}}(\ell) \chi_{4n}(\ell) \right|^2.$$

We now apply the Linnik multiplicative large sieve inequality and clear constants to get

$$\left(\frac{M}{\log M}\right)^{4m} |Y| \ll N^2 \sum_{\ell} |1_{\mathscr{L}}(\ell)|^2 = N^2 \left(\frac{M}{\log M}\right)^{2m}.$$

Since $M = N^{1/m}$ and $\log N \asymp \log M$, this gives that $|Y| \ll (\log N)^{2m}$, so

$$|X| \ll N^{1/2} (\log N)^{2m}. \qquad \blacksquare$$

Remark. Another approach to bounding the quadratic sieve is to use the Linnik additive

large sieve. If m is even, the additive large sieve can result in a bound of

$$|X| \ll N^{1/2}(\log N)^{m/2},$$

resulting in savings of $(\log N)^{3m/2}$. However, if $m = 2k + 1$, the additive large sieve gives a bound of

$$|X| \ll N^{\frac{k+1}{2k+1}}(\log N)^k.$$

This bound is especially bad for small m. For example, when $M = N^{1/3}$, the additive bound is $N^{2/3}(\log N)$, significantly larger compared to the multiplicative bound of $N^{1/2}(\log N)^6$.

Notice that the number of square-free elements that are a quadratic residue for all primes $p \in P_M$ is some power of $\log N$, and is thus smaller than N^ε for any ε. This suggests that the structure of squares is largely maintained even when considered modulo p, since the property that x is square-free and square-like is only satisfied by a small proportion of integers. In the next section, we demonstrate that other algebraically impossible properties are also highly improbable when considered modulo p.

4 Algebraically Impossible Quadratic Sieves

We now turn to the case of "algebraically impossible" quadratic sieves. Recall that "algebraically impossible" conditions are only satisfied by finitely many integers. Specifically, one such condition is that $x^2 + k$ is a square number for some fixed $k \in \mathbb{Z}$. While there are only finitely many solutions over the integers, there may be more integers x that fit this criteria if we consider the case where x and $x + k$ are both quadratic residues modulo p. In fact, by Theorem 3.7, we know that there are $O(N^{1/2}(\log N)^{2m})$ values of x such that x is a quadratic residue for all primes $p \in P_M$, where $M = N^{1/m}$. Comparatively, there are only $O(N^{1/2})$ squares in the integers. To find such pairs x and $x + k$, we begin by reviewing quadratic Diophantine equations.

4.1 Ideals in Rings of Quadratic Integers

In this section, we work with quadratic integers over $\mathbb{Z}[\sqrt{D}]$ for some square-free D. It is well known that for most values of D, the ring $\mathbb{Z}[\sqrt{D}]$ no longer preserves unique factorization for primes [9]. However, if we instead consider ideal factorization, we regain unique factorization. This section culminates in relating the number of ideal divisors to the

number of divisors over the integers.

Definition 4.1. A complex number z is called *a quadratic integer* if it is the solution of $z^2 + bz + c$, where $b, c \in \mathbb{Z}$. Every quadratic integer lies in a uniquely defined quadratic field $\mathbb{Q}[\sqrt{D}]$ where D satisfies $b^2 - 4c = Dk^2$ for some integer k. The set of quadratic integers that lie in $\mathbb{Q}[\sqrt{D}]$ is called the *ring of integers of* $\mathbb{Q}[\sqrt{D}]$ and denoted $\mathcal{O}_D$.

While $\mathcal{O}_D$ is closely related to $\mathbb{Z}[\sqrt{D}]$, they are not necessarily equal.

Lemma 4.2 ([9])**.** *Let $r \in \mathcal{O}_D$ and let*

$$\omega = \begin{cases} \sqrt{D} & \text{if } D \equiv 2, 3 \pmod 4, \\ \frac{1+\sqrt{D}}{2} & \text{if } D \equiv 1 \pmod 4. \end{cases}$$

Then $r = x + y\omega$ for some $x, y \in \mathbb{Z}$.

We also define the conjugate of r by sending $\sqrt{D}$ to $-\sqrt{D}$. We denote this by $\overline{r}$. One may observe that this is identical to complex conjugation when $D < 0$. This then allows us to define the conjugate ideal $\overline{I}$ as the image of an ideal I under conjugation.

Notice that every ideal is generated by at most two elements. Indeed, for any $r_1, r_2, r_3 \in \mathcal{O}_D$, there exist $c_1, c_2, c_3 \in \mathcal{O}_D$ such that $c_1 r_1 + c_2 r_2 + c_3 r_3 = 0$ by $\mathcal{O}_D$ being generated by two linearly independent elements. Thus, if $I = (r, s)$ then $\overline{I} = (\overline{r}, \overline{s})$. Conjugate ideals allow us to split principal ideals generated by primes.

Lemma 4.3 ([9])**.** *Let p be a prime and consider the ideal (p) of $\mathcal{O}_D$. Then, either (p) is a prime ideal or there exists some prime ideal π such that*

$$\pi\overline{\pi} = (p).$$

In order to factor ideals, we write $I \mid J$ if and only if $I \supset J$. The next lemma shows that ideal factorization is indeed unique.

Lemma 4.4 (Unique Prime Ideal Factorization [9])**.** *Let I be an ideal in $\mathcal{O}_D$. Then, there exist prime ideals $P_1, \ldots, P_k$ such that*

$$I = P_1 P_2 \cdots P_k.$$

This factorization is unique up to ordering.

It is possible that the prime ideal and its conjugate are not distinct; that is, $\pi = \overline{\pi}$. While it is possible to explicitly state the prime ideal π, it is unnecessary for the purposes of this paper. The following lemma bounds the number of ideal divisors.

Lemma 4.5. *Let $n \in \mathbb{Z}$. Then, the set of all ideals π such that $\pi \mid (n)$ has size at most $d(n)^2$.*

Proof. Let $n = p_1^{a_1} \cdots p_m^{a_m}$ for distinct primes p_i with $a_i > 0$ for all i. Then, we have $(n) = (p_1)^{a_1} \cdots (p_m)^{a_m}$. For each $1 \leq i \leq m$, either (p_i) is prime or $(p_i) = \pi_i \overline{\pi_i}$ for some prime ideal π. If (p_i) is prime, there are $a_i + 1$ ideals that divide $(p_i)^{a_i}$. If $(p_j) = \pi_j \overline{\pi_j}$, we have

$$(p_j)^{a_j} = \pi_j^{a_j} \overline{\pi_j}^{a_j}$$

and there are $(a_j + 1)^2$ ideals that divide $(p_j)^{a_j}$. Therefore, since $(a_i + 1)^2 > (a_i + 1)$, we upper bound the number of divisors of (n) as

$$\prod_{i=1}^{m} (a_i + 1)^2 = \left(\prod_{i=1}^{m} (a_i + 1) \right)^2 = d(n)^2. \qquad \blacksquare$$

This bound is not generally sharp, especially when there are many prime divisors of n that form prime ideals in $\mathcal{O}_D$. However, this bound is sufficient for our purposes, as $d(n)^2$ is a constant when n is fixed.

4.2 Generalized Pell Equations

We now turn our attention to generalized Pell equations. The structure of positive solutions, that is, $x, y > 0$, to the Pell equation

$$x^2 - Dy^2 = 1 \tag{1}$$

for D not a square number have been understood since the 12th century [10]. Specifically, the Indian mathematician Bhāskara II devised a method to find all solutions to Eq. (1). We restate the result as a corollary of Dirichlet's unit theorem, as stated in Chapter 13 of [11].

Definition 4.6. Consider some $r \in \mathcal{O}_D$ and write $r = x + \sqrt{D}y$ for appropriate choices of x and y. We define the *norm* on $\mathcal{O}_D$ by

$$N(r) = x^2 - Dy^2.$$

Moreover, r is a *positive element* if $x, y > 0$.

Lemma 4.7 ([11])**.** *There exists an element u_0 such that for all $u \in \mathcal{O}_D$ with $N(u) = 1$, there exists some k such that $u = \pm u_0^k$.*

Similarly, the solutions to the generalized Pell equation

$$x^2 - Dy^2 = n$$

for some integer n are also well-understood. We present the structure of such solutions in terms of elements of $\mathcal{O}_D$.

Lemma 4.8 ([12])**.** *Let $u_0 \in \mathcal{O}_D$ be the same as u_0 from Lemma 4.7 and consider all $r \in \mathcal{O}_D$ such that $N(r) = n$. Then, there exists a set $\{\nu_1, \ldots, \nu_m\}$ such that*

$$r = \pm \nu_i u_0^k$$

for some i and k, but

$$\nu_i \neq \pm \nu_j u_0^k$$

for any $i \neq j$ and $k \in \mathbb{Z}$.

Therefore, the solutions to the generalized Pell equation can be partitioned into equivalence classes given by $[\nu_i] := \{r \in \mathcal{O}_D \mid r = \pm \nu_i u_0^k\}$. In each equivalence class, it is clear that the size of the positive solutions grows exponentially. However, to bound the total number of solutions, we must also bound the number of equivalence classes. The following result uses Lemma 4.5 and applies it for the equation $x^2 - Dy^2 = n$.

Lemma 4.9. *Consider the equation*

$$x^2 - Dy^2 = n \tag{2}$$

and consider the set $S = \{\nu_1, \ldots, \nu_m\}$ from Lemma 4.8. Then,

$$|S| \leq d(n)^2.$$

Proof. We show that there exists an injective mapping from the equivalence classes $[\nu_i]$ to ideals π that divide (k). Indeed, consider mapping each $[\nu_i]$ to the ideal (ν_i) and notice that $[\nu_i] \subset (\nu_i)$.

Then, if $(\nu_i) = (\nu_j)$, there exists some $r \in \mathcal{O}_D$ such that $\nu_i = \nu_j r$. However, we have $N(\nu_i) = N(\nu_j)$, which implies that $N(r) = 1$. By Lemma 4.7, we have $r = \pm u_0^k$ for some $k \in \mathbb{Z}$, so $\nu_i = \pm \nu_j u_0^k$. By Lemma 4.8, this implies that $i = j$. Thus, the mapping is injective and $|S| \leq d(k)^2$. ■

We are ready to prove the bound on the number of bounded solutions.

Theorem 4.10. *Let x, y be positive integers such that $x^2 - Dy^2 = k$ and $x + \sqrt{D}y < n$. Then, there are at most $O(\log n)$ solutions.*

Proof. We show that for each equivalence class $[\nu_i]$, there are at most $\log n$ solutions. Indeed, since $x + \sqrt{D}y < n$, it follows that the element $\nu_i u_0^k$ must satisfy

$$\nu_i u_0^k < (\sqrt{D} + 1)n.$$

Since ν_i and u_0 are both greater than 1, we have

$$k < \frac{\log n}{\log u_0} \ll \log n$$

since $\log u_0 > \log 2$. Therefore, there are at most $O(\log n)$ choices for k for each equivalence class. By Lemma 4.9, we know that the number of equivalence classes is at most $d(k)^2$. Thus, the total number of solutions is at most $d(k)^2 O(\log n) = O(\log n)$, as desired. ■

4.3 Algebraic Near-Impossibility

The generalized Pell equation is one type of quadratic form, which is any polynomial of the type $ax^2 + bxy + cy^2$ with $a, b, c \in \mathbb{Z}$. We consider the form $ax^2 - cy^2$ for $a, c > 0$.

Lemma 4.11. *Let (u, v) be a positive solution to $u^2 A - v^2 B = C$ with $A, B, C \in \mathbb{Z}$ and $A, B, C > 0$ and let $m = u\sqrt{A} + v\sqrt{B}$. Then, $N(m^2) = C^2$.*

Proof. We have

$$m^2 = (u^2 A + v^2 B) + 2uv\sqrt{AB}.$$

Therefore, it follows that

$$\begin{aligned} N(m^2) &= (u^2 A + v^2 B)^2 - 4u^2 v^2 AB \\ &= (u^2 A - v^2 B)^2 \\ &= C^2. \end{aligned}$$

■

Since the squaring map is injective for positive m, we know that every solution to $u^2A - v^2B = C$ corresponds uniquely to a solution to $x^2 - ABy^2 = C^2$.

Lemma 4.12. *Consider the set of all solutions (u, v) to $Au^2 - Bv^2 = C$ such that $u, v < n$. Then, the size of this set is at most $O(\log n)$.*

Proof. Let $m = u\sqrt{A} + v\sqrt{B}$ be a solution. Then, since $u, v < n$, we have $m < n(\sqrt{A} + \sqrt{B})$, so $m^2 < n^2(\sqrt{A} + \sqrt{B})^2$. Then, by Theorem 4.10, we know that the size of the solution set is at most

$$O(\log(n^2(A + B + 2\sqrt{AB})) = O(\log n). \qquad \blacksquare$$

We now connect the solutions to $Au^2 - Bv^2 = C$ to solutions of $(ya^2 + k)(yb^2 + k) = \ell^2$ for some $l \in \mathbb{Z}$, which we state in terms of their square-free parts. Specifically, we know that if $(ya^2 + k)(yb^2 + k) = \ell^2$ then $ya^2 + k$ and $yb^2 + k$ must share the same square-free part.

Lemma 4.13. *Let $d \in \mathbb{Z}$ and let $[d]$ be the set of all $a < N$ such that $ya^2 + k$ have the same square-free part d. Then, the size of $[d]$ is $O(\log N)$.*

Proof. Let $ya^2 + k$ has square-free part d. Then, we can rewrite it as $ya^2 + k = dc^2$ for some c. Since $n < N$, we have $a, c < N$, so by Lemma 4.12, there are at most $O(\log n)$ choices for a, as desired. $\blacksquare$

We are now ready to bound the "algebraically impossible" sifted set.

Theorem 4.14. *Let $N, k \in \mathbb{Z}$ and let $M = N^{1/m}$. Consider the set $[N]$ and all primes $p \in P_M$ for some $m \in \mathbb{Z}$ with $m \geq 1$. Then, the set*

$$X = \left\{ n \in [N] \mid \left(\frac{n}{p}\right) = \left(\frac{n+k}{q}\right) = 1 \text{ for all } p \in P_M \right\}$$

has size $|X| \ll (\log N)^{4m+1}$.

Proof. We begin by classifying n based on its square-free part y. Let

$$Y = \left\{ y \text{ square-free} \mid \left(\frac{y}{p}\right) = 1 \text{ for all } p \in P_M \right\}.$$

If $\left(\frac{n}{p}\right) = 1$ for all $p \in P_M$ then there exists some $y \in Y$ such that $n = ya^2$ for some $a \in \mathbb{Z}$. Let

$$X_y = \left\{ ya^2 \mid a \leq \sqrt{\frac{N}{y}} \right\} \cap X$$

and notice that $X = \bigcup_{y \in Y} X_y$. We claim that $|X_y| \ll (\log N)^{2m+1}$. Consider the set

$$\mathscr{L} = \left\{ q \in \mathbb{Z} \mid q = \prod_{i=1}^{2m} p_i,\ p_i \neq p_j \text{ for } i \neq j,\ p_i \in P_M \right\},$$

the same set as in the proof of Theorem 3.7. For $n + k = ya^2 + k$ to also be a quadratic residue, it must satisfy $\left(\frac{ya^2+k}{p}\right) = 1$ for all $p \in P_M$. Then, since χ is multiplicative, we know that

$$\left(\frac{M}{\log M}\right)^{2m} \mathbb{1}_{X_y}(n) \ll \left| \sum_{q \in \mathscr{L}} \left(\frac{ya^2 + k}{q}\right) \right|.$$

After taking the square of both sides and summing over $a < \sqrt{\frac{N}{y}}$, it follows that

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll \sum_{a < \sqrt{\frac{N}{y}}} \left| \sum_{q \in \mathscr{L}} \left(\frac{ya^2 + k}{q}\right) \right|^2.$$

Lemma 3.3 implies that $\left(\frac{ya^2+k}{q}\right) = \chi_{4(ya^2+k)}(q)$. Further notice that instead of summing over $q \in \mathscr{L}$, we may equivalently sum over 1 through $2^{2m}N^2$ and use an indicator function to pick out $q \in \mathscr{L}$. As a result,

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll \sum_{a} \left| \sum_{\ell=1}^{2^{2m}N^2} \mathbb{1}_{\mathscr{L}}(\ell) \chi_{4(ya^2+k)}(\ell) \right|^2.$$

Now, we split a into equivalence classes $[d]$ as defined in Lemma 4.12. Let D denote the set of all possible options for the square-free part d. Then,

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll \sum_{d \in D} \sum_{a \in [d]} \left| \sum_{\ell=1}^{2^{2m}N^2} \mathbb{1}_{\mathscr{L}}(\ell) \chi_{4(ya^2+k)}(\ell) \right|^2.$$

We can then switch the order of summation. Let $\sum_n^*$ denote the sum over all n such that each n is in a unique class $[d]$. By Lemma 4.12, there are at most $O(\log n)$ elements in $[d]$. Thus, we have

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll \log N \sum_{a}^{*} \left| \sum_{\ell=1}^{2^{2m}N^2} \mathbb{1}_{\mathscr{L}}(\ell) \chi_{4(ya^2+k)}(\ell) \right|^2.$$

We now apply the Linnik multiplicative large sieve inequality and clear constants to get

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll (\log N)N^2 \sum_{\ell=1}^{2^{2m}N^2} |\mathbb{1}_{\mathscr{L}}(\ell)|^2.$$

Since $|\mathscr{L}| \asymp \left(\frac{M}{\log M}\right)^{2m}$ and $\log M \asymp \log N$, we have

$$\left(\frac{M}{\log M}\right)^{4m} |X_y| \ll (\log N)N^2 \frac{N^2}{(\log N)^{2m}},$$

which implies that

$$|X_y| \ll (\log N)^{2m+1}.$$

Therefore, by Theorem 3.7, we have $|Y| \ll (\log N)^{2m}$, so

$$|X| = \bigcup_{y \in Y} |X_y| \ll |Y|(\log N)^{2m+1} \ll (\log N)^{4m+1},$$

showing our desired bound. ■

Remark. When m is even, the bound can be sharped to

$$|X| \ll (\log N)^{\frac{5}{2}m+1}$$

from $|Y| \ll (\log N)^{m/2}$ for even m. It is likely that the bound can be further sharpened to $|X| \ll (\log N)^{m+1}$ through the use of the additive large sieve rather than the multiplicative large sieve, since the sharper bound for Y was also shown through the additive large sieve.

5 Boolean Quadratic Sieve

In the previous section, we looked at the intersection of two quadratic sieves; that is, the set with all elements n such that n and $n+k$ were both quadratic residues modulo p for all $p \in P_M$ for $M = N^{1/m}$. In this section, we consider the union of quadratic sieves. Specifically, we bound the size of the set

$$|X| = \left\{ n \in [N] \mid \text{for all } p \in P_M, \text{ there exists some } i \text{ such that } \left(\frac{n+k_i}{p}\right) = 1 \right\}$$

for some fixed set $\{k_1, \ldots, k_a\} \subset \mathbb{Z}$. We begin by defining a function that replicates the behavior of the indicator function on X while being mean 0. Specifically, consider the function

$$f_q(n) = 1 - \prod_{i=1}^{a} \left(1 - \left(\frac{n+k_i}{q}\right)\right)$$

Notice that $f_q(n) \leq 1$. Moreover, $f_q(n) = 1$ if and only if there exists some i such that $\left(\frac{n+k_i}{q}\right) = 1$, and if $\left(\frac{n}{q}\right) = -1$ for all q then $f_q(n) = 1 - 2^a$. Therefore, if we let

$$\mathscr{L}_a = \left\{ q \in \mathbb{Z} \mid q = \prod_{i=1}^{2ma} p_i, p_i \neq p_j \text{ for } i \neq j, p_i \in P_M \right\}.$$

it follows that

$$\left(\frac{M}{\log M}\right)^{2ma} \mathbb{1}_X(n) \ll \left| \sum_{q \in \mathscr{L}_a} f_q(n) \right|. \tag{3}$$

We now present a series of lemmas to bound the right-hand side. The following lemma is a special case of Theorem 11.23 in [7].

Lemma 5.1 ([7])**.** *Consider a polynomial $g \in \mathbb{Z}_p[X]$ with r distinct roots and not a perfect square. Then,*

$$\left| \sum_{x=1}^{p-1} \left(\frac{g(x)}{p}\right) \right| \leq (r-1)p^{1/2} \ll p^{1/2}$$

The above lemma allows us to show a similar result to Lemma 4.12 for the value $\prod_{i=1}^{a}(n+k_i)$.

Lemma 5.2. *Let $d \in \mathbb{N}$ and let $[d]$ be the set of all $n < N$ such that $\prod_{i=1}^{a}(n+k_i)$ have square-free part d. Then, the size of $[d]$ is $O(N^{1/2})$.*

Proof. Let $g(n) = \prod_{i=1}^{a}(n+k_i)$. Then, if $n \in [d]$ then

$$\left(\frac{dg(n)}{p}\right) = 1$$

for all primes p. Consider the set

$$M = \left\{ m \in [N] \mid \left(\frac{dg(m)}{p}\right) = 1 \text{ for all } p < N^{1/2} \right\}.$$

The set M is the set sifted by $E_p = \left\{m \in \mathbb{Z}_p \mid \left(\frac{dg(m)}{p}\right) = -1\right\}$ for all $p < N^{1/2}$. By Lemma 5.1, we have

$$|E_p| \gg \frac{p}{2}.$$

Therefore, by Theorem 3.5, we have $|M| \ll N^{1/2}$. Since $n \in [d]$ implies $n \in M$, the size of $[d]$ is $O(N^{1/2})$. ■

Remark. If $a = 2$, a bound of $O(\log N)$ is possible. It is possible that with further work, we can prove that solutions (x, y) to the equation

$$d\prod_{i=1}^{a}(x + k_i) = y^2$$

grow exponentially, which would give us a bound of $O(\log N)$. However, the bound of $O(N^{1/2})$ is sufficient for the sake of this paper. Note that this is the best possible bound when $a = 1$.

With this lemma in hand, we now bound the sum over quadratic characters of the polynomial $\prod(n + k_i)$. Recall that $\mathscr{L}_a = \left\{q \in \mathbb{Z} \mid q = \prod_{i=1}^{2ma} p_i,\ p_i \neq p_j \text{ for } i \neq j,\ p_i \in P_M\right\}$.

Lemma 5.3. *Let $N \in \mathbb{Z}$, let $a, b \in \mathbb{N}$ with $a < b$, let $\{k_1, \ldots, k_a\} \subset \mathbb{Z}$, and let $M = N^{1/m}$ for some $m \in \mathbb{Z}$ with $m \geq 1$. Then, we have*

$$S_a = \sum_{n<N}\left|\sum_{q\in\mathscr{L}_b}\prod_{i=1}^{a}\left(\frac{n+k_i}{q}\right)\right|^2 \ll N^{2a+2b+1/2}(\log N)^{-2mb}.$$

Proof. To ease notation, we write $K_n := \prod_{i=1}^{a}(n + k_i)$. We begin by rewriting the sum using quadratic reciprocity from Lemma 3.3 and the indicator function. Specifically, we have

$$S_a = \sum_{n}\left|\sum_{\ell=1}^{2^{2mb}N^{2b}} \mathbb{1}_{\mathscr{L}_b}(\ell)\chi_{4K_n}(\ell)\right|^2.$$

Now, we split the sum over n. Let D be the set of all distinct classes $[d]$ defined in Lemma 5.2. Then, we have

$$S_a = \sum_{d\in D}\sum_{a\in[d]}\left|\sum_{\ell=1}^{2^{2mb}N^{2b}} \mathbb{1}_{\mathscr{L}_b}(\ell)\chi_{4K_n}(\ell)\right|^2.$$

We switch the order of summation. Let $\sum_n^*$ denote the sum over all n such that each n is in a unique class $[d]$. Since Lemma 5.2 states that there are at most $O(N^{1/2})$ elements in $[d]$,

we have

$$S_a \ll N^{1/2} \sum_n^* \left| \sum_{\ell=1}^{2^{2mb}N^{2b}} \mathbb{1}_{\mathscr{L}_b}(\ell)\chi_{4K_n}(\ell) \right|^2,$$

By the Linnik multiplicative large sieve in Theorem 3.6, it follows that

$$S_a \ll N^{1/2}N^{2a} \sum_{\ell=1}^{2^{2mb}N^{2b}} |\mathbb{1}_{\mathscr{L}_b}(\ell)|^2.$$

Since $|\mathscr{L}_b| \asymp \left(\frac{M}{\log M}\right)^{2mb}$ and $\log M \asymp \log N$, we have

$$S_a \ll N^{2a+2b+1/2}(\log N)^{-2mb}. \qquad \blacksquare$$

This allows us to bound the Boolean quadratic sieve.

Theorem 5.4. *Let $N, k_i \in \mathbb{Z}$ for $1 \leq i \leq a$ and let $M = N^{1/m}$ for some $m \in \mathbb{Z}$ with $m \geq 1$. Then, the set*

$$X_a = \left\{ n \in [N] \mid \text{for all } p \in P_M, \text{ there exists some } i \text{ such that } \left(\frac{n+k_i}{p}\right) = 1 \right\}$$

has size $|X| \ll N^{1/2}(\log N)^{2ma}$.

Proof. Consider the set

$$\mathscr{L}_a = \left\{ q \in \mathbb{Z} \mid q = \prod_{i=1}^{2ma} p_i, p_i \neq p_j \text{for } i \neq j, p_i \in M \right\},$$

the same set as in Lemma 5.3. Equation (3) states that

$$\left(\frac{M}{\log M}\right)^{2ma} \mathbb{1}_X(n) \ll \left| \sum_{q \in \mathscr{L}_a} f_q(n) \right|.$$

We take the L^2 norm of both sides and split the right-hand side using the triangle inequality, giving us

$$\left(\left(\frac{M}{\log M}\right)^{4ma} |X| \right)^{1/2} \ll \sum_{j=0}^{a} \left(\binom{a}{j} \left(\sum_{n<N} \left| \sum_{q \in \mathscr{L}_a} \prod_{i=1}^{j} \left(\frac{n+k_i}{q}\right) \right|^2 \right)^{1/2} \right) \tag{4}$$

since for any sets of integers $\{k_1, \dots, k_j\}$ and $\{\ell_1, \dots, \ell_j\}$, we have

$$\sum_{n<N} \left| \sum_{q \in \mathscr{L}_a} \prod_{i=1}^{b} \left(\frac{n+k_i}{q} \right) \right|^2 \asymp \sum_{n<N} \left| \sum_{q \in \mathscr{L}_a} \prod_{i=1}^{b} \left(\frac{n+\ell_i}{q} \right) \right|^2.$$

We now use the bound from Lemma 5.3, which states that

$$\sum_{n<N} \left| \sum_{q \in \mathscr{L}_a} \prod_{i=1}^{j} \left(\frac{n+k_i}{q} \right) \right|^2 \ll N^{2a+2j+1/2} (\log N)^{-2ma} \ll N^{4a+1/2} (\log N)^{-2ma}.$$

Substituting the bound into Equation (4), we have

$$\left(\left(\frac{M}{\log M} \right)^{4ma} |X| \right)^{1/2} \ll \sum_{j=0}^{a} \left(\binom{a}{j} (N^{4a+1/2} (\log N)^{-2ma})^{1/2} \right).$$

Since $\sum_{b=0}^{a} \binom{a}{b} = 2^a$, we can simplify the right-hand side to

$$\left(\left(\frac{M}{\log M} \right)^{4ma} |X| \right)^{1/2} \ll 2^a (N^{4a+1/2} (\log N)^{-2ma})^{1/2}$$

After squaring both sides and noting that $\log M \asymp \log N$, we get

$$\frac{N^{4a}}{(\log N)^{4ma}} |X| \ll N^{4a+1/2} (\log N)^{-2ma},$$

which implies our desired bound of

$$|X| \ll N^{1/2} (\log N)^{2ma}.$$ ∎

Remark. A better bound is likely possible for even m due to the alternative proof using additive characters rather than the multiplicative character. In particular, for $a = 2$, we use the bound

$$\sum_{n<N} \left| \sum_{q \in \mathscr{L}_2} \left(\frac{n+k_1}{q} \right) \left(\frac{n+k_2}{q} \right) \right|^2 \ll (\log N)^{4m+1}$$

mentioned in the Remark following Lemma 5.2. Moreover, by the bound in the Remark

following Theorem 3.7, we have

$$\sum_{n<N} \left| \sum_{q \in \mathscr{L}_2} \left(\frac{n+k_i}{q} \right) \right|^2 \ll N^{1/2} (\log N)^m.$$

Therefore, using these two bounds, we get that

$$|X| \ll N^{1/2} (\log N)^m.$$

We conjecture that the same bound is possible for all values of a.

6 Acknowledgements

I would like to thank my mentor, Samuel Packman, for his guidance, patience, and tremendous help throughout this project, as well as Professor Larry Guth in helping formulate the idea behind this project. I would like to thank Dr. Tanya Khovanova for her support and insightful comments regarding how to best present my work, as well as Prof. Roman Bezrukavnikov and Dr. Jonathan Bloom for helping with my project and for general supervision. In addition, I would like to give gratitude to my tutor, AnaMaria Perez, for her helpful suggestions and supporting me through pieces of advice throughout the program. I would also like to thank Dimitar Chakarov, Allen Lin, Mihika Dusad, and Stanislav Ivanov Atanasov for their valuable feedback on this paper, as well as Ashley Malkin, Evan Lim, and Joonsoo Lee for their feedback on my presentation. I would like to thank the Massachusetts Institute of Technology and the Center for Excellence in Education for hosting me and allowing me to work on this project. I thank Sophia Jin for her feedback on my paper and helping me with my presentation. I express my sincerest gratitude to Ramp for sponsoring my summer at RSI. Lastly, I would like to thank my parents, Dillon Sexton, and Kayman McIver for their continuous support and allowing me to come to the Research Science Institute.

References

[1] H. Helfgott. An improved sieve of eratosthenes. *Math. Comp.*, 89(321):333–350, 2020.

[2] R. P. Brent. Irregularities in the distribution of primes and twin primes. *Math. Comp.*, 29(129):43–56, 1975.

[3] Y. V. Linnik. The large sieve. *CR Acad. Sci. URSS (NS)*, 30:292–294, 1941.

[4] A. Rényi. On the representation of an even number as the sum of a prime and an almost prime. *Izv. Math.*, 12:57–78, 1948.

[5] J. R. Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica.*, 16(2):157–176, 1973.

[6] I. Niven, H. S. Zuckerman, and H. L. Montgomery. *An introduction to the theory of numbers*. John Wiley & Sons, 1991.

[7] H. Iwaniec and E. Kowalski. *Analytic number theory*. American Mathematical Soc., 2021.

[8] D. R. Heath-Brown. A mean value estimate for real character sums. *Acta Arith.*, 72(3):235–275, 1995.

[9] P. Ribenboim. *Classical theory of algebraic numbers*. Springer Science & Business Media, 2013.

[10] H. W. Lenstra Jr. Solving the pell equation. *Notices Amer. Math. Soc.*, 49(2):182–192, 2002.

[11] S. Alaca and K. S. Williams. *Introductory Algebraic Number Theory*. Cambridge University Press, 2003.

[12] M. J. Jacobson and H. C. Williams. *Solving the Pell equation*. Springer, 2009.