Faculty of Science

Euclid and the infinite number of missing primes

Bachelor Thesis

Daphne Stouthart

Mathematics

Supervisor:

Dr. Lola Thompson
Mathematical Institute

June 2024

Abstract

All the way back to the time of the ancient Greeks, it was the great mathematician Euclid who proved that there are infinitely many primes. In his proof, he presents a method for constructing an infinite sequence of distinct primes. Specifically, the method ensures that given a finite sequence of distinct primes, we can always find a new prime that is different from all the primes in that sequence, which proves that the sequence of all primes cannot be finite. However, Euclid does not specify exactly *which* prime we should choose in each step of the method, but merely *how* we should choose one. Hence, the infinite sequence that eventually arises from following Euclid's method depends on our specific choice of the next prime in each step. This thesis is concerned with one specific such sequence called the second Euclid-Mullin sequence. In particular, it has been proven that infinitely many primes are missing from this sequence. In this thesis, we will analyze and expand upon two such already existing proofs.

Contents

1 Introduction

Around 300 BC, Euclid - one of the most prominent mathematicians of all time - proved that there exist infinitely many primes.

To provide a bit of background on Euclid, it is worth knowing that he pioneered especially in the field of geometry. He is most famous for his treatise called "The Elements" [6]. Although a large part of this treatise concerns geometry, a significant portion of it actually deals with number theory. In particular, it is in The Elements where Euclid's proof of the infinitude of primes appeared for the first time.

Theorem 1.1 (Euclid). *There are infinitely many prime numbers.*

Proof. We construct an infinite sequence $p_1, p_2, p_3, \ldots$ of distinct prime numbers. First, we choose $p_1 = 2$. Second, we choose p_2 to be a prime divisor of $p_1 + 1$, so we obtain $p_2 = 3$. Third, we choose p_3 to be a prime divisor of $p_1 p_2 + 1$, yielding $p_3 = 7$. Suppose we have created a strictly monotonically increasing finite sequence $p_1, p_2, \ldots, p_n$ of prime numbers (where $n \geq 1$). Consider the sum $P \coloneqq 1 + \prod_{j=1}^{n} p_j$, which is either prime or composite. If P is prime, then since it is different from all of the primes in the sequence $p_1, p_2, \ldots, p_n$, we choose $p_{n+1} = P$. If P is composite, we define p_{n+1} to be a prime divisor of P. Then p_{n+1} is different from all of $p_1, p_2, \ldots, p_n$. To see why, suppose that it is not. Then $p_{n+1} = p_j$ for some $j \in \{1, 2, \ldots, n\}$, and thus $p_{n+1} | \prod_{j=1}^{n} p_j$ and also $p_{n+1} | P$. However, using the distributive law, we see that we must then have that $p_{n+1} | 1$. This is a contradiction since no prime is a divisor of 1. So in both the case that P is prime and the case that P is composite, we have found a prime p_{n+1}that is different from all the n primes in our initial sequence, with which the theorem has been proven. □

Note that Euclid's proof is a proof by construction. The method that Euclid presents ensures that given a finite sequence of distinct primes, we can always find a new prime that is different from all the primes in that sequence. This proves that the sequence of all primes cannot be finite.

However, in the case that the sum P is composite, Euclid does not specify exactly *which* prime divisor of P we should choose. That is, the infinite sequence that eventually arises from following Euclid's method depends on our specific choice of the next prime in each step.

The two most intuitive ways of choosing the prime divisor of the sum P are, as proposed by Mullin [14]:

1. always choosing the smallest such one;
2. always choosing the largest such one.

The two sequences arising this way are

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, 5471, \ldots \tag{1}$$

and

$$2, 3, 7, 43, 139, 50207, 340999, 2365347734339, 4680225641471129, 1368845206580129, \ldots, \tag{2}$$

respectively. They appear in the OEIS [11] as the sequences A000945 and A000946 and they are known as Euclid-Mullin sequences.

Several papers have been written on the sequence (1), but in general not much is known about it still. Nonetheless, there is a conjecture by Shanks [13] that says that it contains all the primes.

However, this thesis is concerned with the sequence (2), which is often referred to as the second Euclid-Mullin sequence. Henceforth, we will denote this sequence by $\{P_n\}_{n=1}^{\infty}$. To be specific, the sequence $\{P_n\}_{n=1}^{\infty}$ is formally defined as: $P_1 = 2$, and for $n \geq 1$ the entry P_{n+1} is the largest prime divisor of $1 + \prod_{j=1}^{n} P_j$.

In particular, it has been proven that infinitely many primes are missing from the second Euclid-Mullin sequence. Building on the work of Cox and Van der Poorten [5], it was Booker who, in 2012, was the first

to prove this in [3]. Booker provides an analytic proof which involves quadratic Dirichlet characters, fundamental discriminants, and Burgess's bounds [4].

Then, in 2014, Pollack and Treviño [12] gave an alternative, entirely elementary proof. Their proof uses quadratic residues and nonresidues, and Legendre and Jacobi symbols as well as their properties.

In this thesis, both the proof of Booker and the proof of Pollack and Treviño are analyzed and expanded upon. Therefore, the main theorem of this thesis is the following:

Theorem 1.2 (Booker). *Infinitely many primes are missing from the second Euclid-Mullin sequence.*

In Chapter 3 we will be dealing with the proof of Pollack and Treviño, and in Chapter 4 we will cover the proof of Booker. But before we delve into the proofs of Pollack and Treviño and Booker, we will first establish the necessary prior knowledge in Chapter 2.

2 Preliminaries

This chapter introduces certain concepts that we will need to comprehend before we get started on the proofs presented in Chapters 3 and 4.

We will assume throughout that the reader is familiar with modular arithmetic and some basic concepts and results from linear algebra as well as from group theory.

The results in the first part of this chapter, up to and including Theorem 2.15, are taken from Beukers's book [2] (specifically pages 35, 37, 59 and 60, and Chapter 11), with the exception of the parts that deal with the concept of the (full) (non)square reach.

We start by stating a theorem that can be used to determine the total number of divisors of any integer greater than 1.

Theorem 2.1. *Let $n > 1$ be an integer and let $p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ be the prime factorization of n. Let us denote the denote the total number of divisors of n by $d(n)$. Then*

$$d(n) = \prod_{i=1}^{m} (k_i + 1).$$

The following theorem says something about the solvability of linear congruences.

Theorem 2.2. *Let a, b and M be integers such that $M \geq 2$. Let $d = \gcd(a, M)$. Then the congruence*

$$ax \equiv b \ (mod\ M)$$

has solutions $x \in \mathbb{Z}$ if and only if $d|b$.

Next, we introduce quadratic residues and nonresidues.

Definition 2.3. *Let p be an odd positive prime number and let $a \in \mathbb{Z}$ such that $p \nmid a$. We say that a is a **quadratic residue modulo p** if the congruence $x^2 \equiv a \ (mod\ p)$ has a solution $x \in \mathbb{Z}$. If this congruence does not have a solution $x \in \mathbb{Z}$, the number a is called a **quadratic nonresidue modulo p**.*

Remark 2.4. Note that in order to determine the quadratic residues and nonresidues modulo p, it suffices to determine $x^2 \pmod p$ for $x = 1, 2, \ldots, \frac{p-1}{2}$. This is due to the fact that $(p-x)^2 \equiv (-x)^2 \equiv x^2 \pmod p$.

Example 2.5. As an example, we determine the quadratic residues and nonresidues modulo 23. So, we have to determine $x^2 \pmod{23}$ for $x = 1, 2, \ldots, 11$. We obtain

$$1^2 \equiv 1, 2^2 \equiv 4, 3^2 \equiv 9, 4^2 \equiv 16, 5^2 \equiv 2, 6^2 \equiv 13, 7^2 \equiv 3, 8^2 \equiv 18, 9^2 \equiv 12, 10^2 \equiv 8, 11^2 \equiv 6.$$

Therefore, the quadratic residues modulo 23 are $1, 2, 3, 4, 6, 8, 9, 12, 13, 16$ and 18, whereas the quadratic nonresidues modulo 23 are $5, 7, 10, 11, 14, 15, 17, 19, 20, 21$ and 22.

Throughout Chapter 3 we will denote by $n^*(p)$ the least positive quadratic nonresidue modulo an odd positive prime p. Note that for all odd positive primes p we have $n^*(p) > 1$ (since 1 is a quadratic residue modulo any positive prime) and $n^*(p) < p$.

We now introduce a new notion, namely that of the (full) (non)square reach.

Definition 2.6. *Let $M > 1$. The length of the longest sequence of consecutive quadratic residues (respectively nonresidues) modulo M is called the **square reach of M** (respectively **nonsquare reach of M**) and denoted by $R_{\square}(M)$ (respectively $R_{\boxtimes}(M)$). In both instances, if we allow multiples of M to be included in the longest sequences, we refer to the **full square reach of M** and the **full nonsquare reach of M** and denote these by $\overline{R_{\square}}(M)$ and $\overline{R_{\boxtimes}}(M)$, respectively.*

Remark 2.7. We emphasize that a sequence corresponding to the full square reach of M merely *allows* integers congruent to 0 modulo M, but that does not mean that it necessarily contains such integers. Only in the case that allowing multiples of M results in a sequence of consecutive quadratic residues that has length greater than the square reach of M, the full square reach of M is strictly greater than the square reach of M. Otherwise, the full square reach and square reach of M are equal. The same reasoning holds for the full nonsquare reach of M compared to the nonsquare reach of M. Put more formally, we have $\overline{R_{\square}}(M) \geq R_{\square}(M)$ and $\overline{R_{\boxtimes}}(M) \geq R_{\boxtimes}(M)$ for all $M > 1$.

Example 2.8. In Example 2.5 we determined the quadratic residues and nonresidues modulo 23. If we look at the sequence $1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18$ of quadratic residues, we see that the longest sequence of consecutive terms is $1, 2, 3, 4$, which is of length 4, so $R_{\square}(M) = 4$. If we allow multiples of 23 to be included, the longest sequence of consecutive residues is obtained by adding 0 to the latter sequence, yielding the sequence $0, 1, 2, 3, 4$, so $\overline{R_{\square}}(23) = 5$. Likewise, if we consider the sequence $5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22$ of quadratic nonresidues, we see that $19, 20, 21, 22$ is the longest sequence of consecutive terms. So $R_{\boxtimes}(23) = 4$. Lastly, we have $\overline{R_{\boxtimes}}(23) = 5$, because if we allow multiples of 23 to be included, we obtain the longest sequence of consecutive nonresidues by adding 23 to the latter sequence, yielding the sequence $19, 20, 21, 22, 23$.

The Legendre symbol, which we will define next, is used for detecting quadratic residues and nonresidues modulo odd primes p.

Definition 2.9. *Let p be an odd prime number and let $a \in \mathbb{Z}$. The **Legendre symbol** is defined as follows:*

$$\left(\frac{a}{p}\right) \coloneqq \begin{cases} 1, & \text{if } a \text{ is a quadratic residue modulo } p \\ -1, & \text{if } a \text{ is a quadratic nonresidue modulo } p \\ 0, & \text{if } p \mid a \end{cases}$$

Rather than mentioning all properties of the Legendre symbol, we will point out one that will be of particular importance in this thesis.

Property 2.10. *The Legendre symbol is a completely multiplicative function of its first argument. Precisely, this means that for all $a, b \in \mathbb{Z}$ and all odd primes p we have*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

In particular, Property 2.10 implies the following theorem.

Theorem 2.11. *Let p be an odd positive prime. Then we have the following:*

1. *The product of two quadratic residues modulo p is again a quadratic residue.*
2. *The product of two quadratic nonresidues modulo p is a quadratic residue.*
3. *The product of a quadratic residue and a quadratic nonresidue is a quadratic nonresidue modulo p.*

The Jacobi symbol is a generalization of the Legendre symbol in the sense that its second argument is defined for all odd natural numbers. It is defined in terms of the Legendre symbol, as follows:

Definition 2.12. *Let $n \in \mathbb{N}_{>0}$ be odd and let $a \in \mathbb{Z}$. Let $n = p_1^{k_1} \cdots p_m^{k_m}$ be the prime factorization of n. The **Jacobi symbol** $\left(\frac{a}{n}\right)$ is defined as*

$$\left(\frac{a}{n}\right) \coloneqq \left(\frac{a}{p_1}\right)^{k_1}\left(\frac{a}{p_2}\right)^{k_2}\cdots\left(\frac{a}{p_m}\right)^{k_m}.$$

Example 2.13. Consider $n = 33$ and $a = 4$. The prime factorization of 33 is $3 \cdot 11$. The number 4 is a quadratic residue modulo 3, because $1^2 \equiv 4 \pmod 3$, so $\left(\frac{4}{3}\right) = 1$. And, since $2^2 \equiv 4 \pmod{11}$, the number 4 is also a quadratic residue modulo 11, so $\left(\frac{4}{11}\right) = 1$. Therefore, we have $\left(\frac{4}{33}\right) = \left(\frac{4}{3}\right)\left(\frac{4}{11}\right) = 1 \cdot 1 = 1$.

The following property for the Jacobi symbol follows directly from Definition 2.12 and Property 2.10.

Property 2.14. *The Jacobi symbol is a completely multiplicative function of its first argument as well as of its second argument. More formally, this means that for all $a, b \in \mathbb{Z}$ and all positive odd integers m and n we have*

$$\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right) \quad \text{and} \quad \left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right),$$

respectively.

In particular, Property 2.14 implies that

$$\left(\frac{a^2}{n}\right) = \left(\frac{a}{n}\right)^2 \text{ as well as } \left(\frac{a}{n^2}\right) = \left(\frac{a}{n}\right)^2,$$

ensuring that both of these expressions are equal to either 0 or 1.

We state the following important theorem.

Theorem 2.15 (Law of quadratic reciprocity for the Jacobi symbol). *Let m and n be positive odd integers such that* $\gcd(m,n) = 1$. *Then,*

1. $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$
2. $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$
3. $\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}$.

The next part of this chapter deals with Dirichlet characters. The content of this part, up to and including Theorem 2.28, is taken from Apostol's book [1], in particular page 24 and Chapters 6 and 8.

Before we state the definition of a Dirichlet character, we first need the notions of characters of a group and arithmetic functions.

Definition 2.16. *Let G be any group. A* ***character*** *of G is a complex-valued function f defined on G that satisfies the following two conditions:*

1. *For all $a, b \in G$ we have $f(ab) = f(a)f(b)$.*
2. *For some $c \in G$ we have $f(c) \neq 0$.*

Definition 2.17. *An* ***arithmetic function*** *is a real- or complex-valued function defined on $\mathbb{N}_{>0}$.*

Definition 2.18. *Let G be the group of reduced residue classes modulo k. We associate to each character f of G an arithmetic function $\chi = \chi_f$ such that*

$$\chi(m) = \begin{cases} 0, & \text{if } \gcd(m,k) > 1 \\ f(\hat{m}), & \text{if } \gcd(m,k) = 1, \end{cases}$$

where $\hat{m} = \{a : a \equiv m \pmod{k}\}$ is the residue class for m. We call the function χ a ***Dirichlet character modulo k****.*

In particular, a ***quadratic Dirichlet character*** *takes on only real values.*

Remark 2.19. We note that the only non-zero values that a Dirichlet character takes on are *roots of unity*, which are complex numbers that result in the value 1 when they are taken to the power of some positive integer. More precisely, an n-th root of unity is a complex number z such that $z^n = 1$, where n is a positive integer. Therefore, the only non-zero values that a quadratic Dirichlet character takes on are 1 and -1, since these are the only two real roots of unity.

Theorem 2.20. *Let χ be a Dirichlet character modulo k. Then the following two statements hold:*

1. *The character χ is completely multiplicative. That is, we have $\chi(mn) = \chi(m)\chi(n)$ for all m and n.*
2. *The character χ is periodic with period k, meaning that $\chi(m+k) = \chi(m)$ for all m.*

Example 2.21. The most basic example of a quadratic Dirichlet character is the *principal character* (sometimes called the trivial character), denoted by χ_0, which is defined as

$$\chi_0(m) = \begin{cases} 0, & \text{if } \gcd(m,k) > 1 \\ 1, & \text{if } \gcd(m,k) = 1, \end{cases}$$

where k is the modulus of χ_0.

Example 2.22. Let p be an odd prime and let $m \in \mathbb{Z}$. Consider the Legendre symbol $\left(\frac{m}{p}\right)$ (see Definition 2.9), which is 0 if $\gcd(m,p) > 1$ and either 1 or -1 if $\gcd(m,p) > 1$, and moreover, which is periodic with period p. Furthermore, recall from Property 2.10 that the Legendre symbol is completely multiplicative. We see that it follows that $\chi(m) \coloneqq \left(\frac{m}{p}\right)$ is a non-principal quadratic Dirichlet character modulo p.

Next, we will define the notions of induced moduli, conductors, and primitive Dirichlet characters.

Definition 2.23. *Let χ be a Dirichlet character modulo k. Let $d > 0$ be such that $d|k$. If we have $\chi(a) = 1$ for all a such that $\gcd(a,k) = 1$ and $a \equiv 1 \pmod{d}$, the number d is said to be an **induced modulus** for the character χ.*

In particular, for every Dirichlet character χ modulo k, the modulus k is an induced modulus for χ.

Definition 2.24. *Let χ be a Dirichlet character modulo k. The **conductor** of χ is the smallest induced modulus for χ.*

Definition 2.25. *A Dirichlet character χ modulo k is called **primitive** if there exists no integer $d < k$ that is an induced modulus for χ.*

In other words, a Dirichlet character χ modulo k is primitive precisely when the conductor of χ is equal to k.

Next, we state the following result regarding the conductor of a Dirichlet character.

Theorem 2.26. *The conductor of a Dirichlet character χ is a divisor of every induced modulus for χ.*

Remark 2.27. Note that Theorem 2.26 implies in particular that the conductor of a Dirichlet character χ modulo k is a divisor of k.

An important result for Dirichlet characters is the following:

Theorem 2.28. *Let χ be a Dirichlet character modulo k. Then χ can be expressed as*

$$\chi(n) = \psi(n)\chi_0(n) \text{ for all } n,$$

where χ_0 is the principal character modulo k and ψ is a primitive character modulo the conductor of χ.

Let us introduce some notation. When we write $f(x) \ll_\epsilon g(x)$ for two functions f and g, we will understand this to mean that there exists a constant $C_\epsilon > 0$ such that $|f(x)| \leq C_\epsilon |g(x)|$. We emphasize that the constant C_ϵ is dependent on ϵ. The following theorem uses this notation.

Theorem 2.29 ([10, Theorem 1]). *Let $\epsilon > 0$ be arbitrarily small. Let χ be a non-principal Dirichlet character modulo q, and let n_χ be the least positive integer for which χ takes on a value that is not 0 or 1. Then for all integers $q \geq 2$, we have*

$$n_\chi \ll_\epsilon \begin{cases} q^{\frac{1}{4\sqrt{e}}+\epsilon} & \text{if } q \text{ is cubefree,} \\ q^{\frac{1}{3\sqrt{e}}+\epsilon} & \text{otherwise.} \end{cases}$$

Lastly, we define fundamental discriminants and we state an important result that connects these to primitive quadratic Dirichlet characters.

Definition 2.30 ([9]). *An integer D is called a **fundamental discriminant** if one of the following three statements holds:*

1. *$D \equiv 1 \pmod 4$ and D is squarefree.*
2. *$D = 4m$, where $m \equiv 2 \pmod 4$ and m is squarefree.*
3. *$D = 4m$, where $m \equiv 3 \pmod 4$ and m is squarefree.*

The next theorem involves the concept of *Kronecker symbols*. We will not give a formal definition of the Kronecker symbol, but we will merely state that, similar to the Jacobi symbol being a generalization of the Legendre symbol, the Kronecker symbol is a generalization of the Jacobi symbol. That is, the Kronecker symbol $\left(\frac{a}{n}\right)$ is defined for all $a, n \in \mathbb{Z}$.

Theorem 2.31 ([7, Theorem 1.1]). *Let D be a fundamental discriminant. Then the Kronecker symbol $\left(\frac{D}{\cdot}\right)$ is a primitive quadratic Dirichlet character modulo $|D|$ (which is also the conductor of χ). Conversely, let χ be a primitive quadratic Dirichlet character. Then there exists a unique fundamental discriminant D such that $\chi = \left(\frac{D}{\cdot}\right)$.*

3 An elementary proof

This chapter is completely based on the article [12] by Pollack and Treviño; all results are from [12] and we follow the treatment and proof strategies of [12].

This chapter deals with Pollack Treviño's elementary proof of Theorem 1.2 as provided in their article [12]. We start by proving a bound on $n^*(p)$.

Lemma 3.1 ([12, Lemma 1]). *If p is a positive odd prime, then $n^*(p) < \frac{1}{2} + \sqrt{p}$.*

Proof. Let p be a positive odd prime, and let us write $n := n^*(p)$. Since p is prime, the number $\frac{p}{n}$ is not an integer. Therefore, we know that $\left\lceil \frac{p}{n} \right\rceil > \frac{p}{n}$. Multiplying both sides of this inequality by n yields

$$n\left\lceil \frac{p}{n} \right\rceil > p. \tag{1}$$

By definition of the ceiling function, we have that

$$\left\lceil \frac{p}{n} \right\rceil < \frac{p}{n} + 1, \tag{2}$$

or equivalently (since $n > 0$),

$$n\left\lceil \frac{p}{n} \right\rceil < p + n. \tag{3}$$

Subtracting p from both sides of inequalities (1) and (3) and combining the resulting inequalities, we obtain $0 < n\left\lceil \frac{p}{n} \right\rceil - p < n$. All integers in the open interval $(0, n)$ are strictly smaller than p and hence not divisible by p, meaning each of them is either a quadratic residue or a quadratic nonresidue modulo p. But, since n is the *least* positive quadratic nonresidue modulo p, all integers in the open interval $(0, n)$ must be quadratic residues modulo p. Therefore, particularly the number $n\left\lceil \frac{p}{n} \right\rceil - p$ is a quadratic residue modulo p. Equivalently, this means that $n\left\lceil \frac{p}{n} \right\rceil$ is a quadratic residue modulo p (since $n\left\lceil \frac{p}{n} \right\rceil - p \equiv n\left\lceil \frac{p}{n} \right\rceil \pmod{p}$). Since $n\left\lceil \frac{p}{n} \right\rceil$ is a quadratic residue and n is a quadratic nonresidue, it follows from Theorem 2.11 that $\left\lceil \frac{p}{n} \right\rceil$ must be a quadratic nonresidue modulo p. Hence by definition of n (and since $\left\lceil \frac{p}{n} \right\rceil$ is positive), we have

$$\left\lceil \frac{p}{n} \right\rceil \geq n \tag{4}$$

Combining inequalities (2) and (4), we obtain

$$\begin{aligned} 1 + \frac{p}{n} > \left\lceil \frac{p}{n} \right\rceil \geq n &\implies 1 + \frac{p}{n} > n \\ &\implies n + p > n^2 \qquad \text{(since } n > 0\text{)} \\ &\implies p > n^2 - n \\ &\implies p \geq n^2 - n + 1, \end{aligned}$$

where the last step follows from the fact that both p and $n^2 - n$ are integers. Therefore we obtain

$$\left(n - \frac{1}{2}\right)^2 = n^2 - n + \frac{1}{4} < n^2 - n + 1 \leq p.$$

So we have found in particular that $\left(n - \frac{1}{2}\right)^2 < p$, from which it follows that $n - \frac{1}{2} < \sqrt{p}$, or equivalently $n < \frac{1}{2} + \sqrt{p}$, as desired. □

Next, we show that the square reach of a positive odd prime is bounded.

Lemma 3.2 ([12, Lemma 2]). *If p is a positive odd prime and $2 \leq n < p$ is a quadratic nonresidue modulo p, then $R_{\square}(p) \leq \max\left\{\frac{p}{n}, n - 1\right\}$.*

Proof. Let p be a positive odd prime, and let $2 \leq n < p$ be a quadratic nonresidue modulo p. Define $R := R_{\square}(p)$. We want to prove that either $R \leq \frac{p}{n}$, or $R \leq n-1$. To this end, we proceed by assuming that $R > \frac{p}{n}$, in which case we need to show that $R \leq n-1$. Let us fix a number $a \in \mathbb{Z}$ such that $a+1, a+2, \ldots, a+R$ is a sequence of consecutive quadratic residues modulo p. If we multiply each term in this sequence by n, this yields the sequence $na+n$, $na+2n$, ..., $na+Rn$. Each term in the latter sequence is a product of a quadratic nonresidue (n) and a quadratic residue ($a+i$, where $i \in \{1, 2, \ldots, R\}$) and hence by Theorem 2.11 is a quadratic nonresidue modulo p.
Consider the intervals

$$(na+jn, na+(j+1)n) \quad \text{with} \quad 1 \leq j < \left\lceil \frac{p}{n} \right\rceil, \quad \text{and} \quad \left(na + \left\lceil \frac{p}{n} \right\rceil n, na+n+p \right). \tag{$*$}$$

We will show that each quadratic residue can be viewed modulo p in such a way that it is contained in exactly one of these intervals.
Since p is prime, the number $\frac{p}{n}$ is not an integer. But R is an integer. So $R > \frac{p}{n}$ is equivalent to $R \geq \lceil \frac{p}{n} \rceil$. This means that $\{na+n, na+2n, \ldots, na+\lceil \frac{p}{n} \rceil n\} \subset \{na+n, na+2n, \ldots, na+Rn\}$, where the latter set is a set of quadratic nonresidues. Therefore, $na+jn$ for $1 \leq j \leq \lceil \frac{p}{n} \rceil$ are quadratic nonresidues. Note that since $na+n+p \equiv na+n \pmod{p}$, this means that all boundaries of the intervals in $(*)$ are quadratic nonresidues modulo p.
Consider the union of the intervals in $(*)$. That is, consider the set $U := (na+n, na+2n) \cup (na+2n, na+3n) \cup \cdots \cup \big(na + (\lceil \frac{p}{n} \rceil - 1)n, na + \lceil \frac{p}{n} \rceil n\big) \cup \big(na + \lceil \frac{p}{n} \rceil n, na+n+p\big)$. The set U ranges over exactly one period of length p; as stated above, the left boundary of the first interval is congruent modulo p to the right boundary of the last interval. Every integer that is not contained in the set U is a quadratic nonresidue (these are the boundaries). Put differently, considered modulo p, (at least) all of the integers in the sequence $0, 1, \ldots, p-1$ that are not quadratic nonresidues are contained in the set U, and no two distinct integers in the set U are congruent modulo p. Therefore, by each quadratic residue must be contained in exactly one of the intervals that make up the set U.
The number of integers in an interval of the form $(na+jn, na+(j+1)n)$ is $na+(j+1)n-(na+jn)-1 = n-1$, and the number of integers in the interval $\big(na+\lceil \frac{p}{n} \rceil n, na+n+p\big)$ is $na+n+p-(na+\lceil \frac{p}{n} \rceil n)-1 = n+p-\lceil \frac{p}{n} \rceil n-1$. Note that since p is prime, we have $\lceil \frac{p}{n} \rceil > \frac{p}{n}$, or equivalently, $\lceil \frac{p}{n} \rceil n > p$. It follows that $n+p-\lceil \frac{p}{n} \rceil n - 1 < n-1$.
The construction of the intervals in $(*)$ ensures that each sequence of consecutive quadratic residues is contained in exactly one of these intervals. Hence we see that the square reach of p is at most $n-1$, which completes the proof. □

This allows us to show that also the full square reach of a positive odd prime is bounded.

Lemma 3.3 ([12, Proposition 3]). *If p is a positive odd prime, then $\overline{R_{\square}}(p) < 2\sqrt{p}$.*

Proof. Let p be a positive odd prime, and let us define $n := n^*(p)$. We give a proof by cases.

Case 1 Assume that $\overline{R_{\square}}(p) > R_{\square}(p)$. Then, if we allow integers congruent to 0 modulo p to be included when determining the longest sequence of consecutive quadratic residues, this sequence must contain a multiple of p. Hence we can view this sequence modulo p in such a way that it contains 0, which is exactly what we will do in the following two subcases:

Case 1a Assume that -1 is not a quadratic residue modulo p. Then we can regard our sequence as starting at 0. Moreover, it follows from the way that n is defined that the least positive integer that is not included in the sequence is n. Thus, we can consider our sequence modulo p as the sequence $0, 1, 2, \ldots, n-1$, which has length n.

Case 1b Assume that -1 is a quadratic residue modulo p. Then, again, the integer n is the least positive integer that is not included in our sequence. Moreover, each $-k = -1 \cdot k$, where $k = 1, 2, \ldots, n-1$, is the product of two quadratic residues and hence by Theorem 2.11 is a quadratic residue. Note that $-n = -1 \cdot n$ is the product of a quadratic residue and a nonresidue, so again by Theorem 2.11 it is a quadratic nonresidue and therefore it is not included in the sequence. Hence, we obtain the sequence $-(n-1), -(n-2), \ldots, -2, -1, 0, 1, 2 \ldots, n-2, n-1$ of length $2n-1$.

From the fact that $n > 1$ it follows that $2n > 1 + n$, or equivalently $2n - 1 > n$. So in both cases 1a and 1b we see that $\overline{R_\square}(p) \leq 2n - 1$. Moreover, by Lemma 3.1 we have $n < \frac{1}{2} + \sqrt{p}$, which implies that $2n < 1 + 2\sqrt{p}$, and hence $2n - 1 < 2\sqrt{p}$. This shows that in case 1 we indeed have $\overline{R_\square}(p) < 2\sqrt{p}$.

Case 2 Assume that $\overline{R_\square}(p) = R_\square(p)$. We again distinguish between two subcases:

Case 2a Assume that the interval $\left(\frac{1}{2}\sqrt{p}, 2\sqrt{p}\right]$ contains a quadratic nonresidue modulo p, say m. We will show that in this case the bound on $\overline{R_\square}(p)$ results from Lemma 3.2.
As m is an integer, requiring m to satisfy $\frac{1}{2}\sqrt{p} < m \leq 2\sqrt{p}$ means the same as requiring m to satisfy $\left\lceil \frac{1}{2}\sqrt{p} \right\rceil \leq m \leq \left\lfloor 2\sqrt{p} \right\rfloor$. In order to apply Lemma 3.2, we need to verify that $2 \leq m < p$. That is, we have to check that $\left[\left\lceil \frac{1}{2}\sqrt{p} \right\rceil, \left\lfloor 2\sqrt{p} \right\rfloor\right] \subseteq [2, p)$. If the prime p is greater than or equal to 5 we obtain the following:

$$\begin{aligned} p \geq 5 &\implies \sqrt{p} \geq \sqrt{5} \\ &\implies \frac{1}{2}\sqrt{p} \geq \frac{1}{2}\sqrt{5} \\ &\implies \left\lceil \frac{1}{2}\sqrt{p} \right\rceil \geq \left\lceil \frac{1}{2}\sqrt{5} \right\rceil = 2 \end{aligned}$$

and

$$\begin{aligned} p \geq 5 > 4 &\implies \frac{1}{4}p > 1 \\ &\implies \frac{1}{4}p^2 > p \\ &\implies \frac{1}{2}p > \sqrt{p} \\ &\implies p > 2\sqrt{p} > \left\lfloor 2\sqrt{p} \right\rfloor. \end{aligned}$$

If $p = 3$, we note that $\left[\left\lceil \frac{1}{2}\sqrt{p} \right\rceil, \left\lfloor 2\sqrt{p} \right\rfloor\right] = [1, 3] \not\subseteq [2, 3) = [2, p)$. However, the only quadratic nonresidue modulo 3 in the interval $[1, 3]$ is 2, and $2 \in [2, 3)$ so also if $p = 3$ we can apply Lemma 3.2.
So, since p is a positive odd prime and $2 \leq m < p$ is a quadratic nonresidue modulo p, it follows from Lemma 3.2 that $\overline{R_\square}(p) = R_\square(p) \leq \max\left\{\frac{p}{m}, m-1\right\}$. If $\max\left\{\frac{p}{m}, m-1\right\} = \frac{p}{m}$, then we obtain

$$\begin{aligned} m > \frac{1}{2}\sqrt{p} &\implies \frac{1}{m} < \frac{2}{\sqrt{p}} \\ &\implies \frac{p}{m} < 2\sqrt{p}. \end{aligned}$$

On the other hand, if $\max\left\{\frac{p}{m}, m-1\right\} = m - 1$, then $m \leq 2\sqrt{p}$ implies that $m - 1 < 2\sqrt{p}$.
Hence, we see that in case 2a we have $\overline{R_\square}(p) = \max\left\{\frac{p}{m}, m-1\right\} < 2\sqrt{p}$.

Case 2b Assume that the interval $\left(\frac{1}{2}\sqrt{p}, 2\sqrt{p}\right]$ does not contain a quadratic nonresidue modulo p. Since p is a positive prime, we have

$$\begin{aligned} p > 1 &\implies \sqrt{p} > 1 \\ &\implies 2\sqrt{p} > 1 + \sqrt{p} > \frac{1}{2} + \sqrt{p} > n, \end{aligned}$$

where the last inequality of the compound inequality follows from Lemma 3.1. Consequently, since $n < 2\sqrt{p}$ and since n is a quadratic nonresidue and is therefore not contained in the interval $\left(\frac{1}{2}\sqrt{p}, 2\sqrt{p}\right]$, we must have that

$$n \leq \frac{1}{2}\sqrt{p}. \tag{5}$$

Each of the integers $1, 2, \ldots, p-1$ is either a quadratic residue or a quadratic nonresidue modulo p (because none of these integers is a mulitple of p). Therefore, by Theorem 2.11 each of the squares $1^2, 2^2, \ldots, (p-1)^2$ is a quadratic residue.

Consider the integers k^2n, where $1 \leq k < p$. As each of these is the product of a quadratic residue (k^2) and a quadratic nonresidue (n), again it follows from Theorem 2.11 that each is a quadratic nonresidue modulo p.

Let us choose k to be the greatest integer such that

$$k^2n \leq \frac{1}{2}\sqrt{p}, \tag{6}$$

or equivalently,

$$-k^2n \geq -\frac{1}{2}\sqrt{p}. \tag{7}$$

(Note that this statement is validated by inequality (5), which guarantees that inequality (6) holds in any case at the very least for $k = 1$.) Then we have $(k+1)^2n > \frac{1}{2}\sqrt{p}$, where $(k+1)^2n$ is a quadratic nonresidue. And because there are no quadratic nonresidues in the interval $\left(\frac{1}{2}\sqrt{p}, 2\sqrt{p}\right]$, it follows that

$$(k+1)^2n > 2\sqrt{p}. \tag{8}$$

Adding inequality (7) to inequality (8), we find

$$\begin{aligned}(k+1)^2n - k^2n > 2\sqrt{p} - \frac{1}{2}\sqrt{p} &\implies k^2n + 2kn + n - k^2n > \frac{3}{2}\sqrt{p} \\ &\implies (2k+1)n > \frac{3}{2}\sqrt{p}.\end{aligned}$$

Moreover, if we multiply both sides of inequality (6) by 3, we obtain $\frac{3}{2}\sqrt{p} \geq 3k^2n$. Hence, we have found that $(2k+1)n > 3k^2n$. Since n is positive, the latter inequality is equivalent to

$$2k + 1 > 3k^2. \tag{9}$$

On the other hand, we have $1 \leq k$, which implies that

$$1 + 2k \leq 3k. \tag{10}$$

And, $1 \leq k$ also implies that $k \leq k^2$, or equivalently,

$$3k \leq 3k^2. \tag{11}$$

Combining inequalities (10) and (11), we obtain $1+2k \leq 3k^2$. However, this contradicts inequality (9).

The above shows that in case 2, there must be a quadratic nonresidue modulo p in the interval $\left(\frac{1}{2}\sqrt{p}, 2\sqrt{p}\right]$, in which case we have proved that $R_{\square}(p) < 2\sqrt{p}$.

Since case 1 and case 2 exhaust all possibilities (see Remark 2.7), we have shown that $R_{\square}(p) < 2\sqrt{p}$, concluding the proof. □

Remark 3.4. The careful reader will notice that if $k = p-1$ is the largest k such that inequality (6) holds, then $(k+1)^2n = p^2n$ is not a quadratic nonresidue since $p|p^2n$, such that inequality (8) need not necessarily be true. However, although this might not be intuitively clear directly, we will show that it is impossible for k to be equal to $p-1$. To that end, let us assume for the sake of arriving at a contradiction, that $k = p-1$ is the largest k such that inequality (6) holds. We first check that

$$(p-1)^2 > \sqrt{p} \tag{12}$$

holds for positive odd primes p. Solving this inequality by hand would require our use of the quartic formula, but this is a considerably complicated and very extensive formula. Therefore, we will

rather enter the inequality into a calculator. Doing so, we find that its solution is given by the set $\{p \mid p \in [0, 0.275508) \cup (2.220744, \infty)\}$. Hence we see that inequality (12) holds in particular for all positive odd primes p. Moreover, we know that $n > 1 > \frac{1}{2}$. Combining inequality (12) and the fact that $n > \frac{1}{2}$, we find that, since $(p-1)^2$, $\sqrt{p}$ and n are all positive,

$$n(p-1)^2 > \frac{1}{2}\sqrt{p}$$

for positive odd primes p, contradicting inequality (6) for $k = p - 1$.

Next, we will prove that the full nonsquare reach of a positive odd prime is bounded. The idea of the proof is very similar to that of the proof of Lemma 3.2.

Lemma 3.5 ([12, Proposition 4]). *If p is a positive odd prime, then $\overline{R_{\boxtimes}}(p) < 2\sqrt{p}$.*

Proof. Let p be a positive odd prime. Consider the intervals

$$\left(j^2, (j+1)^2\right), \text{ where } 1 \leq j < \left\lfloor \sqrt{p} \right\rfloor, \text{ and } \left(\left\lfloor \sqrt{p} \right\rfloor^2, p+1\right). \tag{$**$}$$

We will prove that each quadratic nonresidue and each multiple of p can be regarded modulo p as being included in exactly one of these intervals.

First of all, note that every multiple of p can be viewed modulo p as being contained in the interval $\left(\left\lfloor \sqrt{p} \right\rfloor^2, p+1\right)$ since $p \in \left(\left\lfloor \sqrt{p} \right\rfloor^2, p+1\right)$. The integers $1, 2, \ldots, \left\lfloor \sqrt{p} \right\rfloor$ are not divisible by p; hence by Theorem 2.11 their squares are quadratic residues modulo p. Thus, noting that $1^2 \equiv p+1 \pmod{p}$, we see that all boundaries of the intervals in $(**)$ are quadratic residues modulo p.

Define $U \coloneqq \left(1^2, 2^2\right) \cup \left(2^2, 3^2\right) \cup \cdots \left(\left\lfloor \sqrt{p} \right\rfloor^2, p+1\right)$, the union of the intervals in $(**)$. The set U encompasses one whole period of length p; as already mentioned, the left boundary of the most left interval and the right boundary of the most right interval are congruent modulo p. Each integer that is excluded from the set U is a quadratic residue; these are represented by the boundaries. In other words, the set U can be regarded modulo p to contain (at least) all of the integers in the sequence $2, 3, \ldots, p$ that are not quadratic residues. Also, there are no two different integers in the set U that are congruent modulo p. Consequently, each quadratic nonresidue and each multiple of p is included in precisely one of the intervals in $(**)$.

The number of integers in an interval of the form $\left(j^2, (j+1)^2\right)$ is $(j+1)^2 - j^2 - 1 = 2j$. We have $j < \left\lfloor \sqrt{p} \right\rfloor < \sqrt{p}$, and so $2j < 2\sqrt{p}$. The number of integers in the interval $\left(\left\lfloor \sqrt{p} \right\rfloor^2, p+1\right)$ is

$$p + 1 - \left\lfloor \sqrt{p} \right\rfloor^2 - 1 = p - \left\lfloor \sqrt{p} \right\rfloor^2 < p - (\sqrt{p} - 1)^2 = 2\sqrt{p} - 1 < 2\sqrt{p},$$

where the first inequality follows from the fact that $\left\lfloor \sqrt{p} \right\rfloor > \sqrt{p} - 1$ by definition of the floor function.

The way that the intervals in $(**)$ are defined guarantees that each sequence of consecutive quadratic nonresidues modulo p allowing multiples of p is included in precisely one of these intervals. Therefore, we conclude that the full nonsquare reach of p is striclty smaller than $2\sqrt{p}$, as desired. □

Finally, now that we have verified that the full square and full nonsquare reaches of any odd positive prime p are strictly bounded by $2\sqrt{p}$, we are ready to prove the main result, Theorem 1.2.

The upcoming theorem implies Theorem 1.2: In particular, we will show that given any finite sequence of distinct primes that are missing from the second Euclid-Mullin sequence, we can find a prime that is different from all of these primes and that is also missing from the second Euclid-Mullin sequence. This proves that the sequence of all primes that are missing from the second Euclid-Mullin sequence cannot be finite.

Theorem 3.6 ([12, Proposition 5]). *If $Q_1, Q_2, \ldots, Q_r$ (where $r \geq 1$) are the smallest r primes missing from the sequence $\{P_n\}_{n=1}^{\infty}$, then there is another missing prime smaller than*

$$12^2 \left(\prod_{i=1}^{r} Q_i\right)^2.$$

Proof. Let $Q_1, Q_2, \ldots, Q_r$, where $r \geq 1$, be the smallest r primes missing from the sequence $\{P_n\}_{n=1}^{\infty}$. Define $U \coloneqq 12^2\left(\prod_{i=1}^{r} Q_i\right)^2$. Assume, to the contrary, that all primes smaller than U - other than $Q_1, Q_2, \ldots, Q_r$ - are contained in the sequence $\{P_n\}_{n=1}^{\infty}$. Consider the interval $(2, U]$. Let q be the prime in the interval $(2, U]$ that is latest to turn up in the sequence $\{P_n\}_{n=1}^{\infty}$. In particular, say that q is the n-th entry of the sequence $\{P_n\}_{n=1}^{\infty}$, so $q = P_n$. Then, by construction of the sequence $\{P_n\}_{n=1}^{\infty}$, the prime q is the largest prime divisor of $p \coloneqq 1 + \prod_{j=1}^{n-1} P_j$.
We will show that the only possible prime factors of p are, aside from q, the primes $Q_1, Q_2, \ldots, Q_r$. Since q is the prime in $(2, U]$ that is latest to appear in the sequence $\{P_n\}_{n=1}^{\infty}$, all other primes in the interval $(2, U]$, except for the primes $Q_1, Q_2, \ldots, Q_r$, turn up in the sequence $\{P_n\}_{n=1}^{\infty}$ before q does. Hence, the primes in $(2, U]$ are precisely $P_1, P_2, \ldots, P_n$ and $Q_1, Q_2, \ldots, Q_r$. Moreover, since q is the largest prime factor of p, all possible other prime factors of p must be in the interval $(2, q)$, which is a subset of $(2, U]$. So certainly, any prime factor of p that is not q must be one of $P_1, P_2, \ldots, P_{n-1}$ or one of $Q_1, Q_2, \ldots, Q_r$. But a prime factor of p can never be one of $P_1, P_2 \ldots, P_{n-1}$. (To see why, suppose that $P_i | p$ for some $i \in \{1, 2, \ldots, n-1\}$. Then since $P_i | \prod_{j=1}^{n-1} P_j$, it follows from the distributive law that $P_i | 1$, which is a contradiction since no prime is a divisor of 1.) Hence, the only remaining candidates for possible prime factors of p, aside from q, are the primes $Q_1, Q_2 \ldots, Q_r$. Therefore, we can write

$$p = Q_1^{e_1} Q_2^{e_2} \cdots Q_r^{e_r} q^e \tag{13}$$

for certain integers $e_1, \ldots, e_r \geq 0$ and $e \geq 1$.
Next, we state and prove the following claim:

Claim. *There exists a natural number $m \leq U$ satisfying the following three conditions:*

1. $m \equiv 1 \ (mod\ 4)$
2. $m \equiv -1 \ (mod\ Q_1 \cdots Q_r)$
3. $\left(\frac{m}{q}\right) = \left(\frac{-1}{q}\right)$

Proof of claim. Define $X \coloneqq 2Q_1 \cdots Q_r - 1$ and $Y \coloneqq 4Q_1 \cdots Q_r$. Let m be a natural number that is congruent to X modulo Y. Then,

$$m = kY + X = 4kQ_1 \cdots Q_r + 2Q_1 \cdots Q_r - 1 = (4k+2)Q_1 \cdots Q_r - 1 \equiv -1 \pmod{Q_1 \cdots Q_r}$$

for some integer k, so m satisfies condition 2 of the claim.
Note that we must have $2Q_1 \cdots Q_r \equiv 2 \pmod 4$; if $4 | 2Q_1 \cdots Q_r$, then $2 | Q_1 \cdots Q_r$, which is not possible because $Q_1 \cdots Q_r$ is a product of odd primes. Therefore,

$$m = 4kQ_1 \cdots Q_r + 2Q_1 \cdots Q_r - 1 \equiv 2Q_1 \cdots Q_r - 1 \ \equiv 1 \pmod 4,$$

so the number m satisfies condition 1.
Now that we have verified that an integer of the form $m = kY + X$ satisfies conditions 1 and 2, condition 3 boils down to finding a (small) non-negative integer k such that

$$\left(\frac{kY + X}{q}\right) = \left(\frac{-1}{q}\right). \tag{14}$$

The prime divisors of Y are precisely 2 and $Q_1, \ldots, Q_r$. The number q appears in the sequence $\{P_n\}_{n=1}^{\infty}$, so it is different from all of the Q_i (where $i \in \{1, \ldots, r\}$). Also, $q \neq 2$. Therefore, we have $\gcd(Y, q) = 1$.
From Theorem 2.2 it follows that the congruence $YY' \equiv 1 \pmod q$ has solutions $Y' \in \mathbb{Z}$. Let us fix such a number Y'. Then $YY'k + XY' \equiv k + XY' \pmod q$. Thus, the equation in (14) is equivalent to

$$\left(\frac{k + XY'}{q}\right) = \left(\frac{-Y'}{q}\right). \tag{15}$$

Hence, we seek a non-negative integer k such that equation (15) holds. If $q | -Y'$, then there exists a non-zero integer a such that $q \cdot a = -Y'$. But then $q \cdot -a \cdot Y = YY'$, meaning that $q | YY'$, which contradicts the fact that $YY' \equiv 1 \pmod q$. Therefore, we have $q \nmid -Y'$ (or equivalently, $\left(\frac{-Y'}{q}\right) \neq 0$). There are two cases:

Case 1. The number $-Y'$ is a quadratic residue modulo q.

Case 2. The number $-Y'$ is a quadratic nonresidue modulo q.

Assume that we are in case 1. Then, we seek a non-negative integer k such that $\left(\frac{-Y'}{q}\right) = 1$. Suppose that XY' is congruent to $l \pmod q$. Then $k + XY' \equiv k + l \pmod q$. So for $k = 0, 1, 2, \ldots$, we obtain the following:

$$\begin{aligned} XY' &\equiv l \pmod q \\ 1 + XY' &\equiv l + 1 \pmod q \\ 2 + XY' &\equiv l + 2 \pmod q \\ 3 + XY' &\equiv l + 3 \pmod q \\ &\vdots \end{aligned}$$

where the terms $k + XY'$ for $k = 0, 1, 2, \ldots$, are consecutive terms.
In Lemma 3.5 we showed that full nonsquare reach of any odd positive prime p' is strictly smaller than $2\sqrt{p'}$. So in particular, since q is an odd positive prime, we have $\overline{R_{\boxtimes}}(q) < 2\sqrt{q}$. Equivalently, since the largest integer that is strictly smaller than $2\sqrt{q}$ is $\lfloor 2\sqrt{q} \rfloor$, we have $\overline{R_{\boxtimes}}(q) \leq \lfloor 2\sqrt{q} \rfloor$. Therefore, in the 'worst' case, we have

$$\left(\frac{XY'}{q}\right) \neq 1, \left(\frac{1 + XY'}{q}\right) \neq 1, \left(\frac{2 + XY'}{q}\right) \neq 1, \ldots, \left(\frac{\lfloor 2\sqrt{q} \rfloor - 1 + XY'}{q}\right) \neq 1. \tag{16}$$

But then, since this is a sequence of consecutive quadratic nonresidues modulo q of length $\lfloor 2\sqrt{q} \rfloor$, we must necessarily have for $k = \lfloor 2\sqrt{q} \rfloor$, that $\left(\frac{k + XY'}{q}\right) = 1$. Note that, since $\lfloor 2\sqrt{q} \rfloor$ is striclty smaller than one period of q, the sequence in (16) can contain at most one zero. If it contains a zero, we see that in this 'worst' case, we have $k = \overline{R_{\boxtimes}}(q)$. If it does not contain a zero, we have $k = R_{\boxtimes}(q)$. Since $\overline{R_{\boxtimes}}(q) \geq R_{\boxtimes}(q))$, we see that $k \leq \overline{R_{\boxtimes}}(q)$.
We can make a similar analysis of case 2, applying Lemma 3.3 instead of Lemma 3.5 in the above analysis, in which case we obtain $k \leq \overline{R_{\square}}(q)$. Therefore, we see that there indeed exists a non-negative integer $k \leq \max\{\overline{R_{\boxtimes}}(q), \overline{R_{\square}}(q)\} < 2\sqrt{q}$ such that equation (14) holds, so the number m satisfies condition 3 of the claim.
Using that $k < 2\sqrt{q}$, we obtain

$$0 < m = kY + X < 2Y\sqrt{q} + Y < 2Y\sqrt{q} + Y\sqrt{q} = 3Y\sqrt{q} \leq 3Y\sqrt{U} = U,$$

where in the last step we use that $3Y = 12Q_1 \cdots Q_r = \sqrt{U}$. Hence, $m < U$. This proves the claim. ■

We continue with the proof of the theorem. Let $m \leq U$ be a natural number that satisfies the three conditions of the claim. Let us write $Q \coloneqq Q_1 \cdots Q_r$.
We determine the possible prime factors of m. Since in particular the prime 2 is contained in the sequence $\{P_n\}_{n=1}^{\infty}$, it follows that $Q_i > 2$ for all $i \in \{1, \ldots, r\}$. Combining this fact with the fact that $m \leq U$ and condition 2 of the claim, we see that $m \in (2, U]$. Recall that each prime in the interval $(2, U]$ is one of $P_1, P_2, \ldots, P_n$ or one of $Q_1, Q_2, \ldots, Q_r$. Since $q \nmid -1$, it follows from condition 3 in the claim that $P_n = q \nmid m$. Condition 2 of the claim says that there exists some non-zero integer b such that $m = bQ - 1$. Suppose that $Q_i | m$ for some $i \in \{1, 2, \ldots, r\}$. Then we have $Q_i | bQ - 1$ and $Q_i | bQ$. Again, using the distributive law we see that $Q_i | 1$, which is a contradiction. Therefore, the only possible prime factors of m are $P_1, P_2, \ldots, P_{n-1}$. Any integer d can be written in the form st^2, where s and t are integers and s is squarefree. So in particular, since m is odd (and positive), we can write $m = xy^2$ for certain positive odd integers x and y, where x is squarefree. Since $m \equiv 1 \pmod 4$, it follows from the law of quadratic reciprocity for the Jacobi symbol (Theorem 2.15) that $\left(\frac{m}{p}\right)\left(\frac{p}{m}\right) = 1$. This means that either $\left(\frac{m}{p}\right)$ and $\left(\frac{p}{m}\right)$ are both equal to 1, or they are both equal to -1. Either way, we have $\left(\frac{m}{p}\right) = \left(\frac{p}{m}\right)$. Hence,

$$\left(\frac{m}{p}\right) = \left(\frac{p}{xy^2}\right) = \left(\frac{p}{x}\right)\left(\frac{p}{y^2}\right) = \left(\frac{p}{x}\right)\left(\frac{p}{y}\right)^2, \tag{17}$$

where the last two steps follow from the complete multiplicativity of the Jacobi symbol (Property 2.14).
From the fact that x is squarefree it follows that $x = P_1^{l_1} P_2^{l_2} \cdots P_{n-1}^{l_{n-1}}$, where each l_i (with $i \in \{1, 2, \ldots, n-1\}$) is either equal to 0 or 1. Hence, $x | P_1 P_2 \cdots P_{n-1}$, from which it follows that $p \equiv 1 \pmod{x}$. Consequently, we have $\left(\frac{p}{x}\right) = \left(\frac{1}{x}\right)$. The right hand side of this equation is equal to 1, because the number 1 is a quadratic residue modulo any natural number greater than 1.
Recall that the only possible prime factors of m are $P_1, P_2, \ldots, P_{n-1}$. So certainly, each prime divisor of y is among $P_1, P_2, \ldots, P_{n-1}$. Moreover, the corresponding exponent of each of the prime divisors of y is greater than or equal to 1 in the prime factorization of y. It follows that $y \nmid p$. Hence, $\left(\frac{p}{y}\right)$ equals 1 or -1, which means that $\left(\frac{p}{y}\right)^2 = 1$.
Therefore, since $\left(\frac{p}{x}\right) = \left(\frac{p}{y}\right)^2 = 1$, it follows from equation (17) that $\left(\frac{m}{p}\right) = 1 \cdot 1 = 1$.
On the other hand, using equation (13) and repeatedly applying the complete multiplicativity of Jacobi symbols (Property 2.14), we obtain

$$\begin{aligned}
\left(\frac{m}{p}\right) &= \left(\frac{m}{Q_1^{e_1} Q_2^{e_2} \cdots Q_r^{e_r} q^e}\right) \\
&= \left(\frac{m}{Q_1^{e_1}}\right)\left(\frac{m}{Q_2^{e_2}}\right) \cdots \left(\frac{m}{Q_r^{e^r}}\right)\left(\frac{m}{q^e}\right) \\
&= \left(\frac{m}{Q_1}\right)^{e_1}\left(\frac{m}{Q_2}\right)^{e_2} \cdots \left(\frac{m}{Q_r}\right)^{e_r}\left(\frac{m}{q}\right)^{e} \\
&= \left[\prod_{i=1}^{r}\left(\frac{m}{Q_i}\right)^{e_i}\right] \cdot \left(\frac{m}{q}\right)^{e}. \qquad (18)
\end{aligned}$$

Using condition 2 of the claim and the fact that $Q_i | Q$ for all $i \in \{1, 2, \ldots, r\}$, we see that $m \equiv -1 \pmod{Q_i}$ for all i. Hence, we have $\left(\frac{m}{Q_i}\right) = \left(\frac{-1}{Q_i}\right)$ for all i, from which it follows together with condition 3 in the claim that the expression in (18) is equivalent to $\left[\prod_{i=1}^{r}\left(\frac{-1}{Q_i}\right)^{e_i}\right]\left(\frac{-1}{q}\right)^{e}$. We can rewrite this expression as follows:

$$\begin{aligned}
\left[\prod_{i=1}^{r}\left(\frac{-1}{Q_i}\right)^{e_i}\right] \cdot \left(\frac{-1}{q}\right)^{e} &= \left[\prod_{i=1}^{r}\left(\frac{-1}{Q_i^{e_i}}\right)\right] \cdot \left(\frac{-1}{q^e}\right) \\
&= \left(\frac{-1}{Q_1^{e_1} \cdots Q_r^{e^r} q^e}\right) \\
&= \left(\frac{-1}{p}\right),
\end{aligned}$$

where we again used the complete multiplicativity of the Jacobi symbol (Property 2.14).
Furthermore, since $P_1 = 2$, we have $p = 1 + P_1 \cdots P_{n-1} = 1 + 2\prod_{j=2}^{n-1} P_j$. We must have that either $2\prod_{j=2}^{n-1} P_j \equiv 2 \pmod 4$, or $4 | 2\prod_{j=2}^{n-1} P_j$. However, $4 | 2\prod_{j=2}^{n-1} P_j$ implies that $2 | \prod_{j=2}^{n-1} P_j$, which is a contradiction since $\prod_{j=2}^{n-1} P_j$ is a product of odd primes and therefore must be odd. Thus, $2\prod_{j=2}^{n-1} P_j \equiv 2 \pmod 4$, so $p = 1 + 2\prod_{j=2}^{n-1} P_j \equiv 3 \pmod 4$. But then by the law of quadratic reciprocity for the Jacobi symbol (Theorem 2.15), we have $\left(\frac{m}{p}\right) = \left(\frac{-1}{p}\right) = -1$. However, this contradicts our previously found result that $\left(\frac{m}{p}\right) = 1$. Therefore, there must be a prime $Q_{r+1} < U$ that is different from all of the primes $Q_1, Q_2, \ldots, Q_r$ and that is not contained in the sequence $\{P_n\}_{n=1}^{\infty}$, which concludes the proof. □

4 An analytic proof

This entire chapter is based on Booker's paper [3]; all results presented in this chapter are acquired from [3] and we follow the approach, including the proof strategies, of [3].

This chapter covers the analytic proof of Theorem 1.2 that Booker provides in [3].

In preparation for proving the main result, we state and prove two lemmas.

Lemma 4.1 ([3, Lemma 1]). *Let $\epsilon > 0$ be arbitrarily small. If χ is a non-principal quadratic Dirichlet character modulo q, then there is a positive prime number $n \ll_\epsilon q^{\frac{1}{4\sqrt{e}}+\epsilon}$ such that $\chi(n) = -1$.*

Proof. Let χ be a non-principal quadratic Dirichlet character of modulus q. Because the character χ is quadratic, it only takes on the values 0, 1, and -1. Moreover, since the character χ is non-principal, there must be certain integers (coprime to q) for which χ takes on the value -1. Define n to be the least positive integer for which this is the case. We will show that n must be prime. To that end, suppose that n is composite. Let $n = q_1^{m_1} q_2^{m_2} \cdots q_k^{m_k}$ be the prime factorization of n. In particular, if we set $a = q_2^{m_2} \cdots q_k^{m_k}$, we can write $n = q_1^{m_1} a$. Hence, since χ is completely multiplicative (Theorem 2.20), we find that $-1 = \chi(n) = \chi(q_1^{m_1} a) = \chi(q_1^{m_1})\chi(a)$. This implies that either $\chi(q_1^{m_1}) = 1$ and $\chi(a) = -1$, or $\chi(q_1^{m_1}) = -1$ and $\chi(a) = 1$. In either case we have found a positive integer (namely a or $q_1^{m_1}$, respectively) that is smaller than n and for which χ takes the value -1, contradicting the minimality of n.
To complete the proof of the lemma, it remains to verify the upper bound on n. There are two cases:

Case 1 Assume that q is cubefree.

Theorem 2.29 applies to all non-principal Dirichlet characters, so it applies in particular to quadratic characters such as χ. Adopting the notation of Theorem 2.29, we have $n = n_\chi$, by definition of n. Then, since q is cubefree, this theorem says that $n \ll_\epsilon q^{\frac{1}{4\sqrt{e}}+\epsilon}$.

Case 2 Assume that q is not cubefree.

Let q_1 be the conductor of χ. Recall from Theorem 2.26 (and Remark 2.27) that the conductor of a Dirichlet character is a divisor of the corresponding modulus, meaning in this case that $q_1 | q$. So, we can write $q = q_0 q_1$ for some positive integer q_0.

According to Theorem 2.28, we can write χ as a product $\chi(n) = \psi(n)\chi_0(n)$ for all $n \in \mathbb{Z}$, where ψ is a primitive character modulo q_1 and χ_0 is the principal character modulo q. Let $\widetilde{\chi}_0$ be the principal character modulo q_0, and define a character $\widetilde{\chi} := \psi\widetilde{\chi}_0$. We will show that $\widetilde{\chi} = \chi$.
Since $q_0 | q$, we have $\gcd(m, q) > 1$ for any integer m such that $\gcd(m, q_0) > 1$. For all integers m, this means that if $\widetilde{\chi}_0(m) = 0$, then $\chi_0(m) = 0$.
Furthermore, from the fact that $q_0 | q$, it follows that $\gcd(m, q_0) = 1$ for any integer m satisfying $\gcd(m, q) = 1$. So, for all integers m we have that if $\chi_0(m) = 1$, then $\widetilde{\chi}_0(m) = 1$.
There is one remaining possibility that we need to check. Suppose that m is an integer such that $\gcd(m, q) > 1$ while $\gcd(m, q_0) = 1$. Then, since $q = q_0 q_1$, this implies that we must necessarily have that $\gcd(m, q_1) > 1$. From $\gcd(m, q) > 1$ it follows that $\chi_0(m) = 0$ and from $\gcd(m, q_0) = 1$ it follows that $\widetilde{\chi}_0(m) = 1$. Moreover, the fact that $\gcd(m, q_1) > 1$ implies that $\psi_0(m) = 0$. Therefore, we obtain $\chi(m) = \psi(m)\chi_0(m) = 0 \cdot 0 = 0$ and $\widetilde{\chi}(m) = \psi(m)\widetilde{\chi}_0(m) = 0 \cdot 1 = 0$.
We have considered each possibility; hence, we see that $\widetilde{\chi} = \chi$.

According to Theorem 2.31, the number q_1 is a fundamental discriminant. Recall (from Definition 2.30) that this means that one of the following holds:

- $q_1 \equiv 1 \pmod 4$ and q_1 is squarefree.
- $q_1 = 4m$, where $m \equiv 2 \pmod 4$ and m is squarefree.
- $q_1 = 4m$, where $m \equiv 3 \pmod 4$ and m is squarefree.

Let us define the number $q_0' := \prod_{\substack{p \text{ prime} \\ p | q_0 \\ p \nmid q_1}} p$, which is squarefree and coprime to q_1. Let χ_0' be the principal character modulo q_0', and define $\chi' := \psi\chi_0'$. We will show that $\chi = \chi'$. Recall that

$$\chi_0(m) = \begin{cases} 1, & \text{if } \gcd(m, q_0) = 1, \\ 0, & \text{if } \gcd(m, q_0) > 1, \end{cases}$$

and similarly,

$$\chi_0'(m) = \begin{cases} 1, & \text{if } \gcd(m, q_0') = 1, \\ 0, & \text{if } \gcd(m, q_0') > 1. \end{cases}$$

Furthermore, we have

$$\psi(m) \begin{cases} \neq 0, & \text{if } \gcd(m, q_1) = 1, \\ = 0, & \text{if } \gcd(m, q_1) > 1. \end{cases}$$

First, note that if $\gcd(q_0, q_1) = 1$, then $q_0' = q_0$. In that case, the characters χ_0' and χ_0 are equivalent, from which it follows that $\chi' = \chi$.
So let us assume that $\gcd(q_0, q_1) > 1$. Let m be a random integer. If $\gcd(m, q_1) > 1$, then we have $\psi(m) = 0$, which implies that $\chi(m) = 0 = \chi'(m)$. On the other hand, if $\gcd(m, q_1) = 1$ then we must have that $\gcd(m, q_0) = \gcd(m, q_0')$. This means that $\chi_0(m) = \chi_0'(m)$, and thus $\chi(m) = \chi'(m)$.
So, we indeed have $\chi = \chi'$.

What the above argument shows us is that, if the modulus q_0 of χ_0 was not already squarefree as well as coprime to q_1, then we may replace the character $\chi = \psi\chi_0$ with the character $\chi' = \psi\chi_0'$, where the modulus q_0' of χ_0' is squarefree and comprime to q_1. Therefore, we may restrict ourselves to the case in which q_0 is squarefree and coprime to q_1.

Since by assumption q is not cubefree, it must be that $q_1 = 4m$, where $m \equiv 2 \pmod 4$ and m is squarefree. Consequently, q_1 contains a factor of 2^3 in its prime factorization, and, moreover, we see that this is the only cube factor in its prime factorization. Combining this with the facts that $q = q_0 q_1$ and that q_0 is squarefree and coprime to q_1, we see that also the only cube factor of q is 2^3. And, as it turns out, even with this being the case we can still apply Theorem 2.29. (The explanation hereof goes beyond the scope of this thesis. We refer the interested reader to [8, (12.56)], which, as Booker says, accurately expresses how Burgess's bounds [4, Theorem 2] that (according to Booker) Theorem 2.29 rests on still applies, the only difference being the implicit constant). Hence, applying Theorem 2.29 we find also in this case that $n \ll_\epsilon q^{\frac{1}{4\sqrt{e}}+\epsilon}$.

So the bound on n holds in both cases, which concludes the proof. □

Lemma 4.2 ([3, Lemma 2]). *Let $q_1, \dots, q_r$ be positive integers that are pairwise coprime. For every $i \in \{1, \dots, r\}$, let χ_i be a non-principal quadratic Dirichlet character of modulus q_i. Furthermore, let $\epsilon_i \in \{\pm 1\}$. Then there exists a squarefree positive integer n with no more than r prime factors, that are each $\ll_\epsilon (q_1 \cdots q_r)^{\frac{1}{4\sqrt{e}}+\epsilon}$, and such that $\chi_i(n) = \epsilon_i$ for all $i \in \{1, \dots, r\}$.*

Proof. For each $i \in \{1, \dots, r\}$ we denote by ψ_i the principal character modulo q_i. Let $q = q_1 \cdots q_r$. Furthermore, let χ_S be a character modulo q such that

$$\chi_S(n) = \prod_{i=1}^{r} \begin{cases} \chi_i(n) & \text{if } i \in S, \\ \psi_i(n) & \text{if } i \notin S, \end{cases}$$

which is defined for every non-empty subset S of $\{1, \cdots, r\}$. Then χ_S is, as a product of quadratic characters, a quadratic character. Furthermore, we see that the fact that the integers $q_1, q_2, \dots, q_r$ are pairwise coprime ensures that the character χ_S is non-principal.

Hence, since χ_S is a non-principal quadratic Dirichlet character, we can apply Lemma 4.1, which says that there is a positive prime number $n_S \ll_\epsilon q^{\frac{1}{4\sqrt{e}}+\epsilon}$ such that $\chi_S(n_S) = -1$.

Consider the field $\mathbb{F}_2^r$. This is the field of dimension r in which addition and multiplication are modulo 2. That is, the entries of each vector take on only the values 0 and 1. Fix a non-empty subset S of the set $\{1, \ldots, r\}$. Let $v_S = (a_1, \cdots, a_r) \in \mathbb{F}_2^r$ be the indicator vector of S. Precisely, this means that $a_i = 1$ if $i \in S$ and $a_i = 0$ if $i \notin S$.
Each $\chi_i(n_S)$ either equals 1 or -1: If for some $i \in S$ we would have $\chi_i(n_S) = 0$, this would imply that $\chi_S(n_S) = 0$, which is a contradiction. On the other hand, if $i \notin S$, then $\psi_i(n_S) \neq 0$ because $\psi_i(n_S)$ is a factor of the product $\chi_S(n_S)$. Moreover, $\psi_i(n_S) \neq 0$ implies that $\gcd(n_S, q_i) = 1$, since ψ_i is a principal character. But $\gcd(n_S, q_i) = 1$ implies that $\chi_i(n_S) \neq 0$. Hence, we can define the vector $w_S = (b_1, \ldots, b_r) \in \mathbb{F}_2^r$ to be the vector for which $\chi_i(n_S) = (-1)^{b_i}$ for $i \in \{1, \ldots, r\}$. Note that the vector w_S is unique.

We will show that the scalar product of the vectors v_S and w_S equals 1. The scalar product of w_S and v_S is $w_S \cdot v_S = a_1b_1 + a_2b_2 + \cdots + a_rb_r$. Since a_i is defined to be zero whenever $i \notin S$, we can reduce this sum to include only the terms a_ib_i for which $i \in S$. In other words, we have $w_S \cdot v_S = \sum_{i \in S} a_ib_i$. Given that $a_i = 1$ for all $i \in S$ this sum can be reduced further to $\sum_{i \in S} b_i$. Moreover, since $\chi_S(n_S) = -1$, the principal characters $\psi_i(n)$ for which $i \notin S$ must all be equal to 1, so that we can write $\chi_S(n_S)$ as the product $\prod_{i \in S} \chi_i(n_S)$. Because this product is equal to -1, an odd number of the $\chi_i(n_S)$ must be equal to -1. This means that, considered modulo 2, an odd number of the b_i must be equal to 1 (while the rest of the b_i are equal to 0). Therefore, the scalar product $w_S \cdot v_S$ is equal to the sum of an odd number of 1's and hence is an odd number. Any odd number considered modulo 2 is equal to 1. Thus, $w_S \cdot v_S = 1$.

Next, we show that the vector field $\mathbb{F}_2^r$ is spanned by the set $\{w_S : \emptyset \neq S \subset \{1, \ldots, r\}\}$. Suppose to the contrary that the set $\{w_S : \emptyset \neq S \subset \{1, \ldots, r\}\}$ does not span $\mathbb{F}_2^r$. Then there is a non-zero vector v in $\mathbb{F}_2^r$ that cannot be written as a linear combination of the vectors w_S. This implies that v is not a scalar multiple of any of the w_S. (Note that, in the context of $\mathbb{F}_2^r$, this means that there is no vector w_S such that $v = \lambda w_S$, where $\lambda \in \{0, 1\}$. Since the vector v is non-zero it follows that there is no vector w_S such that $v = w_S$.) Put differently, the vector v is linearly independent of each of the vectors w_S (considered separately), so $v \cdot w_S = 0$ for each non-empty subset S of $\{1, \ldots, r\}$. On the other hand, all non-zero vectors in $\mathbb{F}_2^r$ are covered by the vectors v_S; hence, the vector v must be equal to one of the vectors v_S and as we have seen above, $v_S \cdot w_S = 1$ (for all non-empty subsets S of $\{1, \ldots, r\}$). So in particular, we have $v \cdot w_S = 1$. This is a contradiction.

Every spanning set of a vector space can be reduced to a basis for that vector space. Hence, there exists a subset of $\{w_s : \emptyset \neq S \subset \{1, \ldots, r\}\}$ that forms a basis for $\mathbb{F}_2^r$. In other words, we can find a set T of non-empty subsets of $\{1, \ldots, r\}$ for which $\{w_s : S \in T\}$ forms a basis for $\mathbb{F}_2^r$. Let T be such a set.

We will prove that the primes n_S for which $S \in T$ are distinct. Suppose that $n_{S_1} = n_{S_2}$ for certain $S_1, S_2 \in T$ such that $S_1 \neq S_2$. For $i \in \{1, \ldots, r\}$, let $w_{S_1} = (x_1, \ldots, x_r)$ be the vector for which $\chi_i(n_{S_1}) = (-1)^{x_i}$, and $w_{S_2} = (y_1, \ldots, y_r)$ the vector for which $\chi_i(n_{S_2}) = (-1)^{y_i}$. Then, since $n_{S_1} = n_{S_2}$, we have that $\chi_i(n_{S_1}) = \chi_i(n_{S_2})$ for all $i \in \{1, \ldots, r\}$, which implies that $(-1)^{x_i} = (-1)^{y_i}$ for all i. However, this means that $x_i \equiv y_i \pmod 2$, so considered in $\mathbb{F}_2^r$, we have $w_{S_1} = (x_1, \ldots, x_r) = (y_1, \ldots, y_r) = w_{S_2}$. This contradicts the fact that $\{w_s : S \in T\}$ is a basis for $\mathbb{F}_2^r$.

Therefore, we have found a positive integer $m \coloneqq \prod_{S \in T} n_S$, which is squarefree as it is the product of distinct primes. Furthermore, since $\dim(\mathbb{F}_2^r) = r$, there are r vectors w_S in the basis $\{w_s : S \in T\}$, each corresponding to a specific set $S \in T$. Hence, the number m has exactly r prime factors which, moreover, each are $\ll_\epsilon (q_1 \cdots q_r)^{\frac{1}{4\sqrt{e}}+\epsilon}$. So we can write $m = n_{S_1}n_{S_2} \cdots n_{S_r}$, where the n_{S_i} (with $i \in \{1, \ldots, r\}$) are distinct primes. Then, by Theorem 2.1, the total number of divisors of m is

$$d(m) = \prod_{i=1}^{r} (1 + 1) = \prod_{i=1}^{r} 2 = 2^r.$$

Moreover, each divisor n of m corresponds to a unique sequence $(\chi_1(n), \chi_2(n), \ldots, \chi_r(n))$, and vice versa. Each of these sequences is a sequence whose entries are either 1 or -1; hence, there are 2^r of these sequences.

Therefore, we have a bijective relation between the divisors of m and $\{\pm 1\}^r$. In conclusion, the number m is a squarefree positive integer with no more than r prime factors, that are each $\ll_\epsilon (q_1 \cdots q_r)^{\frac{1}{4\sqrt{e}}+\epsilon}$, and such that $\chi_i(m) = \epsilon_i$ for all $i \in \{1, \ldots, r\}$, exactly as claimed in the lemma. □

We are now ready to prove the main result that infinitely many primes are missing from the second Euclid-Mullin sequence. Similar to the proof of Theorem 3.6, we show that given a finite sequence of distinct primes that are omitted from the second Euclid-Mullin sequence, there exists a prime that is different from the primes in our finite sequence and which, moreover, is omitted from the second Euclid-Mullin sequence. This shows that it is impossible for the sequence of all omitted primes to be finite.

Theorem 4.3 ([3, Theorem 1]). *The sequence $\{P_n\}_{n=1}^\infty$ omits infinitely many primes. If $\{Q_n\}_{n=1}^\infty$ denotes the sequence of omitted primes in increasing order, then for n sufficiently large, we have*

$$Q_{n+1} \leq (Q_1 \cdots Q_n)^{\frac{1}{4\sqrt{e}-1}}.$$

Proof. We give a proof by contradiction. Let $Q_1, Q_2, \ldots, Q_r$, where $Q_1 < Q_2 < \cdots < Q_r$ (and $r \geq 1$), be the smallest r primes that are omitted from the sequence $\{P_n\}_{n=1}^\infty$. Suppose that all primes up to a certain number $x \geq 3$, with the exception of $Q_1, \ldots, Q_r$, are contained in the sequence $\{P_n\}_{n=1}^\infty$. Let $p \leq x$ be the prime that is last to turn up in the sequence $\{P_n\}_{n=1}^\infty$. In particular, say that $p = P_{n+1}$. Recall that by construction of the sequence $\{P_n\}_{n=1}^\infty$, we must then have that p is the largest prime that divides $q \coloneqq 1 + \prod_{j=1}^n P_j$. So possible other prime factors of q are strictly smaller than p. Any prime that is strictly smaller than p must be one of $P_1, \ldots, P_n$, or one of $Q_1, \ldots, Q_r$. It follows from the distributive law that $P_i \nmid q$ for all $i \in \{1, \ldots, n\}$. This means that, if $p \neq q$, then the other prime divisor(s) of q must be one of $Q_1, \ldots, Q_r$. Therefore, we can write

$$q = Q_1^{k_1} Q_2^{k_2} \cdots Q_r^{k_r} p^k$$

for certain integers $k \geq 1$ and $k_1, \ldots, k_r \geq 0$.

Recall from Example 2.22 that the Legendre symbol is a non-principal quadratic Dirichlet character. Consider the Legendre symbols $\left(\frac{\cdot}{Q_1}\right), \left(\frac{\cdot}{Q_2}\right), \ldots, \left(\frac{\cdot}{Q_r}\right)$ and $\left(\frac{\cdot}{p}\right)$. Then $\left(\frac{\cdot}{Q_i}\right)$ (where $i = 1, 2, \ldots, r$) and $\left(\frac{\cdot}{p}\right)$ are non-principal quadratic Dirichlet characters modulo Q_i, and modulo p, respectively. Furthermore, the integers $Q_1, \ldots, Q_r$ and p are pairwise coprime positive integers, as they are all distinct positive primes. Therefore, we can apply Lemma 4.2 to the integers $Q_1, \ldots, Q_r$ and p and the Legendre symbols $\left(\frac{\cdot}{Q_i}\right)$ (with $i = 1, 2, \ldots, r$) and $\left(\frac{\cdot}{p}\right)$. This lemma says that there exists a squarefree positive integer d such that each prime divisor of d is $\ll_\epsilon (pQ_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}}+\epsilon}$, and such that each of $\left(\frac{d}{Q_i}\right)$ for $i = 1, 2, \ldots, r$ and $\left(\frac{d}{p}\right)$ is either 1 or -1. So in particular, there is such a $d \equiv 1 \pmod 4$ that satisfies

$$\left(\frac{d}{Q_i}\right) = \left(\frac{-4}{Q_i}\right) \text{ for all } i \in \{1, \ldots, r\}, \text{ and } \left(\frac{d}{p}\right) = \left(\frac{-4}{p}\right).$$

Note that indeed $\left(\frac{d}{Q_i}\right) \neq 0$ for all i and also $\left(\frac{d}{p}\right) \neq 0$, because for all i we have $Q_i \nmid -4$, and $p \nmid -4$ since the Q_i and p are odd primes. This means that the only possible prime divisors of d are $P_1, P_2, \ldots, P_n$. In other words, we can write d as the product of the distinct primes in some non-empty subset of $\{P_1, P_2, \ldots, P_n\}$. However, if x is sufficiently large, then the bound $(pQ_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}}+\epsilon}$ (which is dependent on ϵ) is smaller than x, because $p \leq x$ and $\frac{1}{4\sqrt{e}} = 0.15163... < 1$. Actually, Booker claims that we can find such an $x \ll_\epsilon (Q_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}-1}+\epsilon}$ based on our choice of ϵ. However, this means that d must have a small prime factor which is not one of $P_1, \ldots, P_n$. This contradicts the fact that each prime factor of d is one of $P_1, \ldots, P_n$. Therefore, there exists a prime $Q_{r+1} \ll_\epsilon (Q_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}-1}+\epsilon}$ that is different from all of the primes $Q_1, \ldots, Q_r$, and that is not contained in the sequence $\{P_n\}_{n=1}^\infty$. This proves the theorem. □

Remark 4.4. Recall (from Chapter 2, page 6) that the notation $Q_{r+1} \ll_\epsilon (Q_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}-1}+\epsilon}$ that appears in the proof of Theorem 4.3 means that that there exists a constant $C_\epsilon > 0$ such that $Q_{r+1} \leq (Q_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}-1}} \cdot C_\epsilon$. This is equivalent to saying that there exists an $\epsilon > 0$ such that $Q_{r+1} \leq (Q_1 \cdots Q_r)^{\frac{1}{4\sqrt{e}-1}+\epsilon}$.

References

[1] T.M. Apostol, *Introduction to analytic number theory.* Gehring, F.W., Halmos, P.R. (eds), Springer-Verlag, New York, Heidelberg, Berlin (1976).

[2] F. Beukers, *Getaltheorie - Een Inleiding.* Epsilon Uitgaven, Amsterdam (2018).

[3] A.R. Booker, *On Mullin's Second Sequence of Primes*, Integers **12** (2012), no. 6, 1167 – 1177.

[4] D.A. Burgess, *On Character Sums and L-Series. II*, Proceedings of the London Mathematical Society **s3-13** (1963), no. 1, 524 – 536.

[5] C.D. Cox and A.J. Van der Poorten, *On a sequence of prime numbers*, Journal of the Australian Mathematical Society **8** (1968), no. 3, 571 – 574.

[6] Euclid, *Euclid's Elements: All Thirteen Books in One Volume.* Densmore, D. (ed), Green Lion Press (2002). Translated by T. L. Heath.

[7] L. Goldmakher, *Legendre, Jacobi, and Kronecker symbols*, `https://web.williams.edu/Mathematics/lg5/Kronecker.pdf`, p.2.

[8] H. Iwaniec and E. Kowalski, *Analytic number theory.* Friedlander, S.J., Manin, Y., Sarnak, P. (eds), Colloquium publications (American Mathematical Society), vol **53**. American Mathematical Society, Providence RI (2004).

[9] J. Klaise, *Orders in quadratic imaginary fields of small class number*, preprint (2012), p.2.

[10] Y. Lau and J. Wu, *On the least quadratic non-residue*, International Journal of Number Theory **04** (2008), no.3, 423 – 435.

[11] The On-Line Encyclopedia of Integer Sequences, Published electronically at https://oeis.org (2024).

[12] P. Pollack and E. Treviño, *The Primes that Euclid Forgot*, The American Mathematical Monthly **121** (2014), 433 – 437.

[13] D. Shanks, *Euclid's primes*, Bulletin of the Institute of Combinatorics and its Applications **1** (1991), 33 – 36.

[14] A.A. Mullin, *Recursive function theory. (A modern look at a Euclidean idea.)*, Research problems, Bulletin of the Institute of Combinatorics and its Applications **69** (1963), p.737.