Euclid-Mullin Sequences in Arithmetic Progressions

Sheridan Heywood
Supervisor: Andrew R. Booker

25th March 2025
Level 7, 20cp

Abstract

The study of the infinitude of primes is one that has fascinated mathematicians for millennia dating back to Euclid's original proof in c. 300 BCE. His proof can be used to construct infinite sequences of primes, these are the so called Euclid-Mullin sequences. The first two sections will cover what is currently known about these sequences. The third section will discuss a generalisation to the construction of sequences of primes 1 (mod m), and finally in section four, we will discuss when exactly similar sequences exist for $a \pmod{m}$, and cover some examples.

Acknowledgements. The author would like to thank Andrew Booker for his supervision and advice on this project, as well as their peer Arron Willis for assistance with grammatical and structural errors.

Acknowledgement of Sources

Acknowledgement of Sources

For all ideas taken from other sources (books, articles, internet), the source of the ideas is mentioned in the main text and fully referenced at the end of the report.

All material which is quoted essentially word-for-word from other sources is given in quotation marks and referenced.

Pictures and diagrams copied from the internet or other sources are labelled with a reference to the web page or book, article etc.

Signed S. Heywood

Date 25/03/2025

Contents

1 Introduction

1.1 Euclid's Theorem

Recall that a prime number is a number divisible only by 1 and itself. The study of such numbers has been ongoing for thousands of years. Of particular note is the study of the infinitude of the primes. The first recorded proof that there are infinitely many prime numbers is attributed to Euclid circa 300 BCE (Book IX, Proposition 20 of Euclid's Elements). His argument was as follows.

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

Proof. Suppose that there is a finite list of primes, $p_1, p_2, ..., p_n$. Let P denote the product of these primes and consider $P + 1 = \prod_{i=1}^{n} p_i + 1$. If $P + 1$ is prime the case is trivial; if not, then $P + 1$ must have some prime divisor, say q. This divisor cannot be any of $p_1, p_2, ..., p_n$ as if it were then it would divide both P and $P + 1$, and therefore the difference between them, but no prime numbers divide 1 and thus q must be a new prime not in the original list. □

In 1963, Mullin used the idea of this proof to suggest two sequences [1]. The first of these is generated by taking the next term of the sequence to be the least prime divisor of the product of the previous terms plus one, and the second by instead taking the largest prime divisor. These sequences have been named the Euclid-Mullin sequences.

Definition 1.2 (The First Euclid-Mullin Sequence)**.** Starting with $p_1 = 2$, take p_{n+1} to be the least prime divisor of $\prod_{i=1}^{n} p_i + 1$. The first few terms of this sequence are:

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, ... \tag{EM1}$$

Definition 1.3 (The Second Euclid-Mullin Sequence)**.** Starting with $q_1 = 2$, take q_{n+1} to be the largest prime divisor of $\prod_{i=1}^{n} q_i + 1$. The first few terms of this sequence are:

$$2, 3, 7, 43, 139, 50207, 340999, 2365347734339, ... \tag{EM2}$$

See Appendix A.1 for a full list of currently known terms.

Remark 1.4. These sequences can in fact be generated by starting from the empty list using the convention that the empty product is equal to one.

1.2 The First Euclid-Mullin Sequence

Little is known about the first Euclid-Mullin sequence. In his 1963 work, Mullin asked two questions: (i) is $\{p_i : i \in \mathbb{N}\}$ recursive? and (ii) does $\{p_i : i \in \mathbb{N}\}$ generate all primes? Also arguing that if (i) is false, then (ii) would be too. Mullin was actually a logician, and his first question refers to concepts from Computability Theory which are outside the scope of this project. In essence he was asking if there was a way to tell whether or not a specific prime occurs in the sequence without having to calculate the entire sequence. As of present, this is still unknown.

The answer to question (ii) is also still unknown. Of the 51 known terms, the smallest prime not known to occur is 41. In 1991, Shanks conjectured that the sequence does in fact generate every prime number, and in [2] provided the following heuristic argument.

Conjecture 1.5 (Shanks, [2])**.** The first Euclid-Mullin sequence, p_n, is a rearrangement of the sequence of all primes so each prime q equals p_n for one, and only one, index n.

Heuristic Argument. Denote as q the smallest prime that has not occured in the sequence up to term p_n. Define $P_n = \prod_{i=1}^{n} p_n$, and set r_1, r_2 such that

$$P_{n-1} \equiv r_1 \pmod{q}, \text{ and } p_n \equiv r_2 \pmod{q},$$

then by definition, $r_1, r_2 \not\equiv 0 \pmod{q}$. Then we have that

$$q = p_{n+1} \iff r_1 r_2 + 1 \equiv 0 \pmod{q}. \tag{1}$$

since if $q = p_{n+1}$ then $q \mid p_1 \cdots p_n + 1$. Then we have $p_1 \cdots p_n + 1 = P_{n-1} p_n + 1 \equiv r_1 r_2 + 1 \pmod{q}$.

Note that $r_1 r_2 \pmod{q}$ can be congruent to any number (residue) between 1 and $q - 1$ inclusive. If (1) is not true, then we replace n with $n + 1$ then $n + 2$ and so on. If we repeat this for $k(q - 1)$ steps, on average each residue will be represented by $r_1 r_2 \pmod{q}$ k times, assuming an equidistribution of the residues. Given that this process can be repeated infinitely, it is highly improbable that

$$r_1 r_2 \equiv q - 1 \text{ (mod } q)$$

will never occur. Given that it does, then (1) will be true, and so q will appear in the sequence. Further, q will not be able to occur again in the sequence since henceforth $r_1 \equiv 0$ (mod q). □

Despite further progress in this field, this conjecture remains open.

2 The Second Euclid-Mullin Sequence

When Mullin introduced the second sequence, he asked the questions: "(iii) Does the process [defined in Definition 1.3] generate $\{q_i : i \in \mathbb{N}\}$ in increasing order", "(iv) Is $\{q_i : i \in \mathbb{N}\}$ still recursive even though (iii) may be answered negatively?", and "(v) If (iii) is answered negatively, does $\{q_i : i \in \mathbb{N}\}$ generate all primes?". Similarly to before, question (iv) is outside the scope of this project, and the answer is unknown.

The first of these questions was answered in 1984 by Naur by manual calculation of the terms, proving that $q_{10} < q_9$ and so the sequence was in fact not increasing [3]. The second of these questions had been tackled prior to this. In 1967 Cox and Van Der Poorten presented a proof that certain primes were missed by the sequence, in particular that the only primes less than 53 that occur are precisely $2, 3, 7$, and 43 [4]. They then further conjectured that the sequence in fact misses infinitely many primes.

This conjecture proved more difficult, finally being solved in 2012 by Booker, who showed that the sequence does in fact miss infinitely many primes [5].

Theorem 2.1 (Booker, [5])**.** *The second Euclid-Mullin sequence* (EM2) *omits infinitely many primes.*

This section will explore Cox and Van Der Poorten's condition for the non-occurence of a prime in (EM2) and a version of Booker's proof presented by Pollack and Treviño in 2014 [6].

But first, we cover some necessary preliminaries.

2.1 Preliminaries

Definition 2.2 (Quadratic Residue)**.** For p an odd prime number, an integer a, such that $p \nmid a$, is called a *quadratic residue modulo* p if there exists an $x \in \mathbb{Z}$ such that $x^2 \equiv a \pmod{p}$, and a *quadratic non-residue modulo* p otherwise.

Example 2.3. Consider $p = 19$. Note that we only need check x^2 for $x = 1, 2, \ldots, 9$ since for $y = 10, 11, \ldots, 18$, we have that $y = 19 - x \equiv -x \pmod{19}$ for one of the x, and then $y^2 \equiv (-x)^2 = x^2 \pmod{19}$.

So we check x^2 for $x = 1, 2, \ldots, 9$:

x	x^2	$x^2 \pmod{19}$
± 1	1	1
± 2	4	4
± 3	9	9
± 4	16	16
± 5	25	6
± 6	36	17
± 7	49	11
± 8	64	7
± 9	81	5

Thus {1, 4, 5, 6, 7, 9, 11, 16, 17} are all quadratic residues modulo p, whilst {2, 3, 8, 10, 12, 13, 14, 15, 18} are quadratic non-residues.

Remark 2.4. The argument in the above example infact applies to all odd primes p, in that to find the quadratic residues modulo p, you only need check $x = 1, 2, \dots, \frac{p-1}{2}$, since each $\frac{p+1}{2}, \dots, p-1$ can be considered as $p - x$ for one of the x, in which case $p - x \equiv -x \pmod{p}$, and then $(-x)^2 = x^2$.

Definition 2.5. Denote by $\ell(\square, p)$ the *longest run* $a+1, a+2, \dots, a+\ell$ *of consecutive quadratic residues modulo* p. Similarly, denote by $\ell(\boxtimes, p)$ the *longest run of consecutive quadratic non-residues modulo* p. If we wish to include 0 and p in the run, we denote each of these as $\ell'(\square, p)$ and $\ell'(\boxtimes, p)$.

Example 2.6. Going back to our previous example with $p = 19$, since the quadratic residues are $\{1, 4, 5, 6, 7, 9, 11, 16, 17\}$, and the quadratic non-residues are $\{2, 3, 8,$
$10, 12, 13, 14, 15, 18\}$, we have that $\ell(\square, 19) = \#\{4, 5, 6, 7\} = 4$, and $\ell(\boxtimes, 19) = \#\{12, 13, 14, 15\} = 4$ (where the # represents the cardinality, or size, of the set). In this case $\ell'(\square, p)$ and $\ell'(\boxtimes, p)$ are the same.

Instead consider $p = 23$ which has quadratic residues $\{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$ and quadratic non-residues $\{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}$. Clearly $\ell(\square, 23) = \#\{1, 2,$
$3, 4\} = 4$, and $\ell(\boxtimes, 23) = \#\{19, 20, 21, 22\}$, however since these runs appear at the start and end of the sequences respectively, we have that $\ell'(\square, 23) = \#\{0, 1, 2, 3, 4\} = 5$, and $\ell'(\boxtimes, 23) = \#\{19, 20, 21, 22, 23\} = 5$.

A nice shorthand to represent whether or not an integer is a quadratic residue modulo p was introduced by Legendre.

Definition 2.7 (Legendre Symbol)**.** For p an odd prime, and a an integer, we define the *Legendre Symbol* to be:

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

Notation. The Legendre symbol may also be written as $(a \mid p)$.

Example 2.8. From our previous examples, we have that $\left(\frac{6}{19}\right) = 1$, whilst $\left(\frac{7}{23}\right) = -1$.

We now state some useful properties of the Legendre symbol that will come up often in our later proofs, but first we state Euler's Criterion, a method for evaluating the symbol. We will state it without proof, but if a proof is desired see for example [7, p. 65-66].

Theorem 2.9 (Euler's Criterion)**.**

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}$$

Theorem 2.10. *For p an odd prime and a, b integers, we have:*

(i) for all a and b,

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

(ii) if $a \equiv b \pmod{p}$, then

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

(iii) if $p \nmid a$, then

$$\left(\frac{a^2}{p}\right) = 1 \text{ and } \left(\frac{a^2 b}{p}\right) = \left(\frac{b}{p}\right),$$

(iv) we also have

$$\left(\frac{1}{p}\right) = 1 \text{ and } \left(\frac{-1}{p}\right) = (-1)^{(p-1)/2},$$

(v) and

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

Proof. Parts (i-iv) follow directly from Euler's Criterion, whilst (v) is more complicated. For a proof see [7, p. 67-70]. □

This theorem provides us with the following corollary, which follows directly from part (i).

Corollary 2.11. *The product of two quadratic residues or two quadratic non-residues is always a quadratic residue, whilst the product of a quadratic residue and a quadratic non-residue is always a quadratic non-residue.*

Next we state a very important result from number theory known as Quadratic Reciprocity. This provides a helpful tool for calculating Legendre symbols by 'inversion' of the symbol. For a proof see [7, p. 67-72].

Theorem 2.12 (Quadratic Reciprocity)**.** *Suppose that p and q are odd distinct primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Remark 2.13. If we know that $\left(\frac{q}{p}\right) \neq 0$, we must have that $\left(\frac{q}{p}\right)^2 = 1$, and so we can multiply both sides of (2.12) by $\left(\frac{p}{q}\right)$ to obtain a more useful form of the theorem:

$$\left(\frac{p}{q}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{q}{p}\right).$$

We now present a generalisation of the Legendre Symbol.

Definition 2.14 (Jacobi Symbol)**.** For Q an odd positive integer with prime decomposition $Q = q_1 \cdots q_s$ not necessarily distinct, and a an integer, we define the *Jacobi Symbol*, $\left(\dfrac{a}{Q}\right)$ to be:

(i) $\left(\dfrac{a}{1}\right) = 1$,

(ii) $\left(\dfrac{a}{Q}\right) = 0$ if $\gcd(a, Q) > 1$,

(iii) $\left(\dfrac{a}{Q}\right) = \left(\dfrac{a}{q_1}\right)\left(\dfrac{a}{q_2}\right)\cdots\left(\dfrac{a}{q_s}\right)$ if $\gcd(a, Q) = 1$.

Remark 2.15. The properties of the Legendre symbol outlined in Theorem 2.10, as well as Quadratic Reciprocity, all generalise to the Jacobi symbol as expected.

2.2 The Conjecture

We will now explore Cox and Van Der Poorten's proof of a sufficient condition for the non-occurrence of a prime in (EM2). This section will follow their reasoning in [4] with extra exposition added.

Denote by $\{q_i\} = \{2, 3, 7, 43, 139, 50207, 340999, ...\}$ the second Euclid-Mullin sequence (EM2), and by $\{p_i\} = \{2, 3, 5, 7, 11, 13, ...\}$ the sequence of all primes in monotonically increasing order.

Before we get into their argument however, let us consider two rather simple cases in order to understand the basics of their reasoning.

Theorem 2.16. *The prime* 5 *does not appear in the Second Euclid-Mullin Sequence* (EM2)*.*

Proof. Suppose for the sake of contradiction, that 5 does appear in (EM2), then for some k we must have that $1 + q_1 \cdots q_k = p_1^{e_1} \cdots p_r^{e_r}$ with $p_r = 5$ being the largest prime appearing on the right-hand side of the above equation. However we already know that $2, 3$ appear in (EM2) and so must appear on the left-hand side of the equation, and so cannot appear on the right. This tells us that $p_1^{e_1} \cdots p_r^{e_r} = 5^{e_r}$ so we now have the equation

$$2 \cdot 3 \cdot q_3 \cdots q_k + 1 = 5^e$$

for some $e \in \mathbb{N}$. Now, we have that the right-hand side is $5^e \equiv 1^e \equiv 1 \pmod 4$. However, the left-hand side is 1 more than 2 times an odd number. Since an odd number must be $\equiv \pm 1 \pmod 4$, we must have that the left-hand side is $\equiv 3 \pmod 4$. We thus have a contradiction, telling us that 5 does not appear in (EM2) as claimed. □

This shows a nice contradiction that can arise for the early terms in (EM2) and is the motivation for congruence (8) below. However, once you get to numbers even just a little bit higher, this congruence alone is not enough.

Theorem 2.17. *The prime 11 does not appear in the Second Euclid-Mullin Sequence* (EM2).

Proof. Suppose for the sake of contradiction that 11 does appear in (EM2). Similarly to the proof for 5, we thus must have that for some k, $1 + q_1 \cdots q_k = p_1^{e_1} \cdots p_r^{e_r}$ with $p_r = 11$ being the largest prime on the right-hand side. Further, we know that $2, 3$, and 7 all appear in (EM2) and so must appear on the left-hand side, thus not the right-hand side. So

$$2 \cdot 3 \cdot 7 \cdot q_4 \cdots q_k + 1 = 5^{e_3} \cdot 11^{e_5}. \tag{2}$$

By the same argument as before, the left-hand side is $\equiv 3 \equiv -1 \pmod 4$. We also have that the right-hand side is $\equiv 1^{e_3} \cdot (-1)^{e_5} \equiv (-1)^{e_5} \pmod 4$. However this is not enough to get a contradiction, since if e_5 is odd both sides of the equation above will have the same congruence ($-1 \pmod 4$) which gives us the congruence $e_5 \equiv 1 \pmod 2$. To get to a contradiction, we need to use quadratic residues. First, note that

$$\left(\frac{5^{e_3} \cdot 11^{e_5}}{q_i}\right) = \left(\frac{1}{q_i}\right) = 1$$

for all $i = 2, \ldots, k$ by Theorem 2.10(ii) and (2) since $5^{e_3} \cdot 11^{e_5} \equiv 1 \pmod{q_i}$. Then also note that

$$\left(\frac{5^{e_3} \cdot 11^{e_5}}{q_i}\right) = \left(\frac{5}{q_i}\right)^{e_3} \left(\frac{11}{q_i}\right)^{e_5}. \tag{3}$$

So we clearly need that if $\left(\frac{5}{q_i}\right)$ or $\left(\frac{11}{q_i}\right)$ are quadratic non-residues, then their exponents e_3, e_5 respectively are even and if they are both quadratic non-residues, then their exponents must both add to an even number. Now we calculate the quadratic residues for a few values of i.

For $i = 2$, we have that $q_i = 3$, so $\left(\frac{5}{3}\right) = \left(\frac{2}{3}\right) = -1$ and $\left(\frac{11}{3}\right) = \left(\frac{2}{3}\right) = -1$. This is because the quadratic residues mod 3 are $0^2 = 0, 1^2 = 1, 2^2 = 4 \equiv 1 \pmod 3$, and so 2 is clearly not a quadratic residue. So we have that the right-hand side of (3) is $(-1)^{e_3} \cdot (-1)^{e_5}$ which is equal to 1 only when $e_3 + e_5 \equiv 0 \pmod 2$.

Doing the same for $i = 3$ ($q_3 = 7$) we get $\left(\frac{5}{7}\right) = -1, \left(\frac{11}{7}\right) = \left(\frac{4}{7}\right) = 1$, since the quadratic residues of 7 are $0^2 = 0, 1^2 = 1, 2^2 = 4, 3^2 = 9 \equiv 2 \pmod 7$, this time giving us the equation $(-1)^{e_3} \cdot 1^{e_5} = (-1)^{e_3}$ so we need $e_3 \equiv 0 \pmod 2$.

Combining all this we have the 3 congruences:

$$\begin{aligned} e_5 &\equiv 1 \pmod 2, \\ e_3 + e_5 &\equiv 0 \pmod 2, \\ e_3 \quad &\equiv 0 \pmod 2. \end{aligned}$$

Adding any two of these together gives us a contradiction, e.g. adding the bottom two gives the congruence $2e_3 + e_5 \equiv e_5 \equiv 0 \pmod 2$ which is a direct contradiction to the top congruence. Thus 11 does not appear in (EM2). □

Cox and Van Der Poorten conjectured that these additional congruences utilising quadratic residues (combined with the previous congruence utilising the properties of the primes modulo 4) are sufficient to reveal every number that does not appear in (EM2), given that enough terms of (EM2) are known. These additional arguments are the motivation for the congruences (9) in Cox and Van Der Poorten's argument as presented below.

Cox and Van Der Poorten's argument in [4]. We aim to find sufficient conditions to determine whether or not a prime p_r appears in (EM2) or not.

Let $r > 1$. If p_r occurs in (EM2) then for some $k \in \mathbb{N}$, we must have that p_r is the greatest divisor of $1 + q_1 \cdots q_k$, i.e.

$$1 + q_1 \cdots q_k = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r} \tag{4}$$

for some integers $e_1, \ldots, e_{r-1} \geq 0$ and $e_r \geq 1$. Thus we must have that

$$p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r} \equiv 1 \pmod{q_i}, \tag{5}$$

for all $i = 1, 2, \ldots, k$. Then since the $q_1, \ldots, q_k$ are all distinct (since Euclid's proof always generates a new prime), we have

$$p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r} \not\equiv 1 \pmod{q_i^2}, \tag{6}$$

for all $i = 1, 2, \ldots, k$. These congruences are helpful but imply weaker yet even more useful conditions. In particular, (5) tells us that when a prime, say p_i, occurs in (EM2) then $e_i = 0$, since clearly a number divisible by p_i cannot also be congruent to $1 \pmod{p_i}$. In particular we must have that $e_1 = 0$, since we know that 2 occurs in (EM2) and $p_1 = 2$. Then (6), in the specific case when $i = 1$ ($q_1 = 2$) tells us that

$$p_2^{e_2} \cdots p_r^{e_r} \not\equiv 1 \pmod{4}. \tag{7}$$

Combining this with the fact that $p_2^{e_2} \cdots p_r^{e_r}$ must be odd as 2 does not divide it, we get that $p_2^{e_2} \cdots p_r^{e_r} \equiv -1 \pmod{4}$. Therefore, $p_2^{e_2} \cdots p_r^{e_r}$ must contain an odd number of prime divisors that are congruent to $-1 \pmod{4}$ when counted according to multiplicity (i.e. if p_i is odd with exponent e_i then it is counted e_i times). This is because if there were an even number of prime divisors $\equiv -1 \pmod{4}$, say $2m$ for some m, then $p_2^{e_2} \cdots p_r^{e_r} \equiv (-1)^{2m} \equiv 1 \pmod{4}$, a contradiction of (7).

We must also have that $p_2^{e_2} \cdots p_r^{e_r}$ is a quadratic residue modulo q_i for all $i = 1, 2, \ldots, k$ by (5), since 1 is always a quadratic residue modulo any natural number. This tells us that in particular $\left(\frac{p_2^{e_2} \cdots p_r^{e_r}}{q_i}\right) = \left(\frac{p_2}{q_i}\right)^{e_2} \cdots \left(\frac{p_r}{q_i}\right)^{e_r} = 1$, and so $p_2^{e_2} \cdots p_r^{e_r}$ must contain an even number of prime divisors (again counted according to multiplicity) that are quadratic non-residues of q_i for $i = 2, \ldots, k$.

We can combine these two facts to get a set of congruences that must be solvable if p_r occurs in (EM2): For some non-negative integers $e_2, \ldots, e_r$, with

$e_i = 0$ when p_i is one of $q_2, \dots, q_k$, we have from the first fact (odd number of prime divisors $\equiv -1 \pmod 4$) that

$$\begin{aligned} &\text{taking} \quad a_{1j} = \begin{cases} 1, & \text{if } p_j \equiv -1 \pmod 4, \\ 0, & \text{if } p_j \equiv 1 \pmod 4, \end{cases} \\ &\text{then} \quad a_{12}e_2 + a_{13}e_3 + \cdots + a_{1r}e_r \equiv 1 \pmod 2, \end{aligned} \tag{8}$$

and from the second fact (an even number of prime divisors that are quadratic non-residues) we get

$$\begin{aligned} &\text{taking} \quad 2a_{ij} = 1 - \left(\frac{p_j}{q_i}\right) \quad (i = 2, \dots, k;\ j = 2, \dots, r), \\ &\text{then} \quad a_{i2}e_2 + a_{i3}e_3 + \cdots + a_{ir}e_r \equiv 0 \pmod 2 \quad (i = 2, \dots, k). \end{aligned} \tag{9}$$

□

These congruences may seem a bit abstract, but upon further examination are rather simple. In (8), the a_{1j} acts as an indicator function for whether the corresponding prime p_j is $\equiv -1 \pmod 4$ or not, and then the summation simply counts each prime with respect to it's multiplicity. (9) is similar with the a_{ij} being an indicator for whether p_j is a quadratic non-residue modulo q_i. These congruences therefore establish the following theorem.

Theorem 2.18 ([4, Theorem 1])**.** *For the prime p_r, if for some k the congruences (8, 9) are inconsistent then*

(i) the prime p_r does not occur in (EM2)*, and*

(ii) neither do the primes p_j $(j < r)$ unless p_j is already one of $q_1, q_2, \dots, q_k$.

Where part (ii) follows immediately, since if the congruences (8, 9) are inconsistent, then they are also inconsistent with $e_{j+1} = e_{j+2} = \cdots = e_r = 0$.

Cox and Van Der Poorten used this initially to show that the primes 5, 11, 13, 17, 19, 23, 29, 31, 37, 41, and 47 do not occur in (EM2), however there was a mistake in their working and what they claimed was not actually provable at the time. In their working, they calculated $\left(\frac{47}{139}\right) = -1$ when it should be 1. From this error, they claimed that an inconsistency in the congruences (8, 9) could be obtained from looking at the first congruence $(i = 1)$, and $i = 4, 5, 6$ in the second. However from the author's working[1], it seems to actually require $i = 1, 4, 5, 6, 7, 11, 12, 13$. A correction of their example will be presented below.

Example 2.19. Take $p_r = 53$, and $k = 13$ (so we use the first 13 terms of (EM2), see section A.1). Of the primes less than 53, we know that $2, 3, 7, 43$ appear in the first 13 terms of (EM2), so we do not need to consider these. Thus if 53 appears in (EM2), we have that

$$1 + 2 \cdot 3 \cdots q_{13} = 5^{e_3} \cdot 11^{e_5} \cdots 41^{e_{13}} \cdot 47^{e_{15}} \cdot 53^{e_{16}}.$$

[1] The author's working consists of python code that is available to view in a github repository found at [8].

For some $e_3, e_5, \ldots, e_{13}, e_{15}, e_{16}$ non-negative integers. The congruence (8) gives us

$j =$	3	5	6	7	8	9	10	11	12	13	15	16
$p_j =$	5	11	13	17	19	23	29	31	37	41	47	53
$p_j \pmod 4 \equiv$	1	-1	1	1	-1	-1	1	-1	1	1	-1	1
$a_{1j} =$	0	1	0	0	1	1	0	1	0	0	1	0

and the congruences (9) give (we will only consider the rows we need, $i = 4, 5, 6, 7, 11, 12, 13$)

$j =$	3	5	6	7	8	9	10	11	12	13	15	16
$p_j =$	5	11	13	17	19	23	29	31	37	41	47	53
$\left(\frac{p_j}{q_4}\right) =$	-1	1	1	1	-1	1	-1	1	-1	1	1	1
$\left(\frac{p_j}{q_5}\right) =$	1	1	1	-1	-1	-1	1	1	1	1	1	-1
$\left(\frac{p_j}{q_6}\right) =$	-1	-1	1	-1	-1	1	-1	-1	-1	1	1	1
$\left(\frac{p_j}{q_7}\right) =$	1	1	1	1	-1	-1	-1	1	1	1	-1	-1
$\left(\frac{p_j}{q_{11}}\right) =$	1	-1	-1	-1	1	-1	-1	1	-1	-1	-1	1
$\left(\frac{p_j}{q_{12}}\right) =$	1	-1	1	-1	1	1	1	-1	1	1	-1	1
$\left(\frac{p_j}{q_{13}}\right) =$	1	1	-1	1	-1	1	1	-1	-1	-1	1	1

Which gives us the following series of congruences for $e_3, \ldots, e_{16}$

$$\begin{aligned}
i &= 1 & e_5 + e_8 + e_9 + e_{11} + e_{15} &\equiv 1 \pmod 2,\\
i &= 4 & e_3 + e_8 + e_{10} + e_{12} &\equiv 0 \pmod 2,\\
i &= 5 & e_7 + e_8 + e_9 + e_{16} &\equiv 0 \pmod 2,\\
i &= 6 & e_3 + e_5 + e_7 + e_8 + e_{10} + e_{11} + e_{12} &\equiv 0 \pmod 2,\\
i &= 7 & e_8 + e_9 + e_{10} + e_{15} + e_{16} &\equiv 0 \pmod 2,\\
i &= 11 & e_5 + e_6 + e_7 + e_9 + e_{10} + e_{12} + e_{13} + e_{15} &\equiv 0 \pmod 2,\\
i &= 12 & e_5 + e_7 + e_{11} + e_{15} &\equiv 0 \pmod 2,\\
i &= 13 & e_6 + e_8 + e_{11} + e_{12} + e_{13} &\equiv 0 \pmod 2.
\end{aligned}$$

Now if we add every congruence except the first together we get

$$\begin{aligned}
2e_3 + 3e_5 + 2e_6 + 4e_7 + 5e_8 + 3e_9 + 4e_{10} + 3e_{11} + 4e_{12} + 2e_{13} + 3e_{15} + 2e_{16}\\
\equiv e_5 + e_8 + e_9 + e_{11} + e_{15} \equiv 0 \pmod 2.
\end{aligned}$$

But then the first congruence tells us that the same sum is $\equiv 1 \pmod 2$. These cannot both be true and thus we have an inconsistency. Therefore, the only primes smaller than 53 that appear in (EM2) are exactly $2, 3, 7, 43$.

The largest prime that can be shown to be omitted with the 14 currently known terms of (EM2) is 73. More interestingly however is that the calculation of the 14th term did not increase this bound, since an inconsistency arises for 73 with just the first 13 terms.

In [4], Cox and Van Der Poorten conjectured that there are infinitely many primes missed by the Second Euclid-Mullin Sequence, and further that every missed prime would produce an inconsistency in congruences (8, 9). They also showed that both conjectures cannot be false.

In 2012 the first of their conjectures was finally answered by Booker [5] who provided an analytic proof that any list of primes omitted from (EM2) admits another omitted prime below a certain bound. In 2014 Pollack and Treviño provided a more elementary proof of the same theorem, albeit with a weaker bound [6]. An exposition of their proof will follow in the next two subsections.

2.3 Bounding Runs of Quadratic Residues

This section will follow section 2 from Pollack and Treviño's paper "The Primes That Euclid Forgot" [6] with additional exposition.

In this section, the goal will be to prove an upper bound for $\ell(\square, p), \ell(\boxtimes, p)$, $\ell'(\square, p)$ and $\ell'(\boxtimes, p)$, namely that each one is smaller than $2\sqrt{p}$.

Lemma 2.20 ([6, Lemma 1])**.** *Denote by $n_2(p)$ the smallest quadratic non-residue modulo p. Then $n_2(p) < \frac{1}{2} + \sqrt{p}$.*

Proof. Set $n = n_2(p)$, then note that since p is prime, we have $\frac{p}{n} < \lceil \frac{p}{n} \rceil < \frac{p}{n} + 1$. If we then multiply through by n and subtract p, we get the inequality

$$0 < n\left\lceil \frac{p}{n} \right\rceil - p < n.$$

By definition of n, we must have that $x \in (0, n) \cap \mathbb{Z}$ are all quadratic residues modulo p, and in particular $n\lceil \frac{p}{n} \rceil$ is a quadratic residue. Then by Corollary 2.11, we must have that $\lceil \frac{p}{n} \rceil$ is a quadratic non-residue. Minimality of n then tells us that $\lceil \frac{p}{n} \rceil \geq n$ so $1 + \frac{p}{n} > \lceil \frac{p}{n} \rceil \geq n$. We thus have

$$1 + \frac{p}{n} > n \implies n + p > n^2 \implies p > n^2 - n \implies p \geq n^2 - n + 1,$$

where the final inequality holds since $p, n \in \mathbb{Z}$. Therefore, we get

$$\begin{aligned} \left(n - \frac{1}{2}\right)^2 < n^2 - n + 1 \leq p &\implies n - \frac{1}{2} < \sqrt{p} \\ &\implies n < \frac{1}{2} + \sqrt{p} \end{aligned}$$

as required. □

Lemma 2.21 ([6, Lemma 2])**.** *Let $1 \leq n < p$ be a quadratic non-residue modulo p. Then*

$$\ell(\square, p) \leq \max\left\{\frac{p}{n}, n - 1\right\}.$$

Proof. Let $\ell = \ell(\square, p)$, and choose $a \in \mathbb{Z}$ such that $a + 1, a + 2, \ldots, a + \ell$ are all quadratic residues modulo p. Multiply each of these by n to get the sequence

$$na + n, na + 2n, \ldots, na + \ell n. \tag{10}$$

Note that since n is a quadratic non-residue, and each $a+i$ is a quadratic residue, we must have that each term in (10) is a quadratic non-residue by Corollary 2.11.

Now suppose that $\ell > \frac{p}{n}$ and consider the intervals

$$(na + jn, na + (j+1)n), \ \forall j = 1, 2, \dots, \left\lceil \frac{p}{n} \right\rceil - 1, \tag{11}$$

$$(na + \left\lceil \frac{p}{n} \right\rceil n, na + n + p). \tag{12}$$

Each interval in (11) has width n, and since $\ell > \frac{p}{n}$, we have that $\ell \geq \lceil \frac{p}{n} \rceil$, so that the end points of each interval occur in the sequence (10) and are thus quadratic non-residues. Further, since $\lceil \frac{p}{n} \rceil n > p$, we have that $(na+n+p)-(na+\lceil \frac{p}{n} \rceil n) = n + p - \lceil \frac{p}{n} \rceil n < n$, and so the interval in (12) has width less than n.

Now consider the union of all of the intervals in (11) and (12), i.e.

$$\begin{aligned}(na + n, na + 2n) &\cup (na + 2n, na + 3n) \cup \cdots \\ &\cdots \cup (na + (\left\lceil \frac{p}{n} \right\rceil - 1)n, na + \left\lceil \frac{p}{n} \right\rceil n) \cup (na + \left\lceil \frac{p}{n} \right\rceil n, na + n + p).\end{aligned} \tag{13}$$

Since each successive interval has starting point equal to the end point of the previous interval, it is clear that (13) is exactly the set $[na+n, na+n+p]$ with the points in (10) taken out, as well as $na+n+p$. Thus (13) has width p.

Therefore, up to modulo p, we can consider each quadratic residue modulo p as lying inside one of these intervals, which each contains at most $n-1$ integers. Thus the longest run of quadratic residues can be at most $n-1$ integers long.

Combining this with our assumption, we either have that $\ell < \frac{p}{n}$, or $l < n-1$, as required. □

Lemma 2.22 ([6, Proposition 3])**.** *If p is an odd prime, then $\ell'(\square, p) < 2\sqrt{p}$.*

Proof. Begin by noting that $\ell'(\square, p) \geq \ell(\square, p)$, and in particular, $\ell'(\square, p) = \ell(\square, p)$ only if the longest run of quadratic residues does not contain a multiple of p, as such we begin by ruling out long runs of quadratic residues containing a multiple of p. Any such run can be considered modulo p as containing 0, and we consider two cases:

(i) -1 is not a quadratic residue modulo p,

(ii) -1 is a quadratic residue modulo p.

In the first case, any run containing a multiple of p can be viewed as a subset of $[0, n_2(p))$, and thus has length at most $n_2(p)$. In the second case, any run containing a multiple of p can be viewed as a subset of $(-n_2(p), n_2(p))$, and thus has length at most $2n_2(p) - 1$. Note that from Lemma 2.20, $2n_2(p) - 1 < 2(\frac{1}{2} + \sqrt{p}) - 1 = 2\sqrt{p}$, so we have that

$$\ell'(\square, p) \leq \max\{2n_2(p) - 1, \ell(\square, p)\} \leq \max\{2\sqrt{p}, \ell(\square, p)\}.$$

It therefore suffices to show that $\ell(\square, p) < 2\sqrt{p}$. To that end, consider the interval $(\frac{1}{2}\sqrt{p}, 2\sqrt{p}]$. Suppose that there is a quadratic non-residue in the interval, call it m, then by Lemma 2.21, $\ell(\square, p) \leq \max\{\frac{p}{m}, m-1\}$. Clearly $m - 1 < 2\sqrt{p}$ since $m \leq 2\sqrt{p}$, and $\frac{p}{m} < \frac{p}{\frac{1}{2}\sqrt{p}} = 2\sqrt{p}$, so $\ell(\square, p) < 2\sqrt{p}$ as required.

Now suppose that there is no quadratic non-residue in the interval. From Lemma 2.20 we know that $n_2(p) < \frac{1}{2} + \sqrt{p}$ and $\frac{1}{2} + \sqrt{p} < 2\sqrt{p}$. Combining this with the fact that $n_2(p)$ can't be in the interval $(\frac{1}{2}\sqrt{p}, 2\sqrt{p}]$, we must have that $n_2(p) \leq \frac{1}{2}\sqrt{p}$. Call $n := n_2(p)$, then each of k^2n with $1 \leq k < p$ is a quadratic non-residue modulo p by Corollary 2.11 (since k^2 is trivially a quadratic residue but n is not). Pick k as large as possible such that

$$k^2 n \leq \frac{1}{2}\sqrt{p} \tag{14}$$

then since no integer in $(\frac{1}{2}\sqrt{p}, 2\sqrt{p}]$ is a quadratic non-residue, we must have that

$$(k+1)^2 n > 2\sqrt{p}. \tag{15}$$

Since $2\sqrt{p} = 4 \cdot \frac{1}{2}\sqrt{p}$, combine (14) and (15) to get

$$4k^2 n \leq 2\sqrt{p} < (k+1)^2 n \implies 3k^2 n \leq 2\sqrt{p} < (2k+1)n$$

and thus $3k^2 < 2k + 1$. But this inequality does not hold for any $k \geq 1$ and we get a contradiction. Thus there must be a quadratic residue in the interval $(\frac{1}{2}\sqrt{p}, 2\sqrt{p}]$, and so as we have already proven, we must have that $\ell(\square, p) < 2\sqrt{p}$ and the result follows. □

Lemma 2.23 ([6, Proposition 4]). *For each odd prime p, we have $\ell'(\boxtimes, p) < 2\sqrt{p}$.*

Proof. Consider the intervals

$$(j^2, (j+1)^2), \ \forall j = 1, 2, \ldots, \lfloor \sqrt{p} \rfloor - 1, \tag{16}$$

$$(\lfloor \sqrt{p} \rfloor^2, p+1). \tag{17}$$

First note that the start and end points of each of the intervals in (16) are trivially quadratic residues (since they are squares), and that the union of all of the intervals covers $[1, \lfloor \sqrt{p} \rfloor^2]$ with every square removed. Then since $(j+1)^2 - j^2 = 2j+1$, there lie exactly $2j$ integers in each interval, and $2j < 2\lfloor \sqrt{p} \rfloor < 2\sqrt{p}$.

Now also note that the interval (17) contains $p - \lfloor \sqrt{p} \rfloor^2 < p - (\sqrt{p} - 1)^2 < 2\sqrt{p}$ integers, and that the interval when combined with the union of all of the intervals in (16), covers all of $[1, p)$ with the squares taken out. This tells us that every quadratic non-residue or multiple of p must be contained in one of the intervals in (16) or (17), and in both cases we have shown that there are less than $2\sqrt{p}$ integers in the interval, meaning that we must have that $\ell'(\boxtimes, p) < 2\sqrt{p}$ as required. □

Remark 2.24. Note that the above proof transitively applies the same limit to $\ell(\boxtimes, p)$, since $\ell(\boxtimes, p) \leq \ell'(\boxtimes, p)$.

2.4 The Proof

We are now ready to move on to the proof of Theorem 2.1. In a similar fashion to Euclid's original proof of the infinitude of primes, the proof shows that if you know the r smallest primes missing from the sequence, then there must be another within a certain bound. However the proof is not constructive in the same way as Euclid's, meaning that it does not provide a method for finding another missing prime not in the list. Booker's original proof in [5] provides such a bound for another missing prime, as does Pollack's and Treviño's, however due to the more elementary nature of their proof, the bound is not as strong. Their proof follows below.

Notation. Throughout the following proof, denote the second Euclid-Mullin sequence (EM2) as $q_1, q_2, q_3, \ldots$.

Theorem 2.25 ([6, Proposition 5]). *Let $p_1, p_2, \ldots, p_r$ be the smallest r primes omitted from the second Euclid-Mullin sequence, where $r \geq 0$. Then there is another omitted prime smaller than*

$$12^2 \left(\prod_{i=1}^{r} p_i \right)^2 . \tag{18}$$

Proof. Set $X = 12^2(\prod_{i=1}^{r} p_i)^2$. For the sake of contradiction, let us suppose that every prime $p \leq X$ except $p_1, \ldots, p_r$ appears in the second Euclid-Mullin sequence (EM2). Take p to be the prime in $[2, X)$ that is last to appear in the sequence $\{q_i\}$, say as the nth term, q_n. Then by construction of (EM2), p is the largest prime dividing $1 + q_1 \cdots q_{n-1}$. Since p is the last prime in $[2, X)$ to appear in (EM2), every other prime in $[2, X)$ must have either already appeared or will not appear, i.e. is one of the p_i. In particular, we must have that every prime smaller than p that is not one of the p_i, is one of $q_1, \ldots, q_{n-1}$ and thus the only other possible factors of $1 + q_1, \ldots, q_{n-1}$ are $p_1, \ldots, p_r$. So we have that

$$1 + q_1 \cdots q_{n-1} = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r} p^e \tag{19}$$

for some integers $e_1, \ldots, e_r \geq 0$ and $e \geq 1$.

We next endeavour to find a natural number $d \leq X$ that satisfies each of the following conditions

$$d \equiv 1 \pmod 4, \quad d \equiv -1 \pmod{p_1 \cdots p_r}, \tag{20}$$

and

$$\left(\frac{d}{p} \right) = \left(\frac{-1}{p} \right). \tag{21}$$

Begin by defining $a := 2p_1 \cdots p_r - 1$ and $m := 4p_1 \cdots p_r$, and then consider a d of the form $d = km + a$. Since we know that 2 appears in (EM2) we must have that the p_i are all odd, and so $p_1 \cdots p_r \equiv \pm 1 \pmod 4$. So

$$d = km + a = p_1 \cdots p_r \cdot (4k + 2) - 1 \equiv 2 \cdot p_1 \cdots p_r - 1 \equiv 1 \pmod 4,$$

where the final equivalence holds since for any $b \equiv \pm 1 \pmod 4$, we have $2b \equiv \pm 2 \equiv 2 \pmod 4$. We also have that

$$d = km + a = -1 + (4k+2) \cdot p_1 \cdots p_r \cdot \equiv -1 \pmod{p_1 \cdots p_r}.$$

Thus this d satisfies (20) for any integer k. It thus remains to find such a k that satisfies (21). We look for a small non-negative integer k, such that $\left(\frac{d}{p}\right) = \left(\frac{km+a}{p}\right) = \left(\frac{-1}{p}\right)$, or equivalently if we fix some m' satisfying $mm' \equiv 1 \pmod p$ (which exists since $\gcd(m,p) = 1$), then multiplying both sides of the previous equation by $\left(\frac{m'}{p}\right)$, we get $\left(\frac{km+a}{p}\right) \cdot \left(\frac{m'}{p}\right) = \left(\frac{kmm'+am'}{p}\right) = \left(\frac{k+am'}{p}\right)$, and $\left(\frac{-1}{p}\right) \cdot \left(\frac{m'}{p}\right) = \left(\frac{-m'}{p}\right)$. Thus we look for a non-negative integer k with

$$\left(\frac{k+am'}{p}\right) = \left(\frac{-m'}{p}\right). \tag{22}$$

Note that since $\gcd(m,p) = 1$, by the construction of m' we must also have that $\gcd(m',p) = 1$, and so $\left(\frac{-m'}{p}\right) \neq 0$ and we must have that $-m'$ is either a quadratic residue or a quadratic non-residue modulo p. Then further observe that when we run through $k \in \mathbb{N}$, $k + am'$ is a run of consecutive integers, and so from our results in section 2.3 the entire run of $k + am'$ can only contain runs of quadratic residues or non-residues of length $< 2\sqrt{p}$, and so we can find some $k \leq \max\{\ell'(\square, p), \ell'(\boxtimes, p)\} < 2\sqrt{p}$ such that (22) is satisfied, and in turn so is (21). Then the corresponding d satisfies

$$0 < d = km + 1 < 2m\sqrt{p} + m < 3m\sqrt{p} < 3m\sqrt{X},$$

and since $3m = 12p_1 \cdots p_r = \sqrt{X}$, we have that $d < X$ and thus this d satisfies everything we need it to.

Returning to the main proof, we aim to attain a contradiction through the conditions that d satisfies.

Since $d \leq X$, and d is coprime to each p_i by (20) and also to p by (21) (since $\left(\frac{-1}{p}\right) \neq 0$ as $\gcd(-1,p) = 1$), so every prime dividing d must be among the primes $q_1, \ldots, q_{n-1}$. Write $d = d_0 d_1^2$, where d_0 is squarefree (not divisible by any square numbers), then $d_0 \mid q_1 \cdots q_{n-1}$, and so

$$\begin{aligned}
\left(\frac{d}{1 + q_1 \cdots q_{n-1}}\right) &= \left(\frac{1 + q_1 \cdots q_{n-1}}{d}\right) (-1)^{(d-1)((1+q_1 \cdots q_{n-1})-1)/4} \\
&= \left(\frac{1 + q_1 \cdots q_{n-1}}{d_0 d_1^2}\right) \\
&= \left(\frac{1 + q_1 \cdots q_{n-1}}{d_0}\right) \left(\frac{1 + q_1 \cdots q_{n-1}}{d_1^2}\right) \\
&= \left(\frac{1}{d_0}\right) \cdot \left(\left(\frac{1 + q_1 \cdots q_{n-1}}{d_1}\right)\right)^2 \\
&= 1 \cdot 1 = 1.
\end{aligned}$$

The first equality uses Quadratic Reciprocity, and the second holds since $d-1 \equiv 0 \pmod 4$ by (20) leaving $q_1 \cdots q_{n-1} = 2 \cdot q_2 \cdots q_{n-1}$ which is even, so

$$(-1)^{(d-1)((1+q_1\cdots q_{n-1})-1)/4} = (-1)^{\frac{d-1}{4}\cdot 2\cdot q_2\cdots q_{n-1}} = (-1)^{2(\frac{d-1}{4}q_2\cdots q_{n-1})} = 1.$$

Finally, the fourth equality holds since $1 + q_1 \cdots q_{n-1} \equiv 1 \pmod{d_0}$.

On the other hand, (20) tells us that $\left(\frac{d}{p_i}\right) = \left(\frac{-1}{p_i}\right)$ (since $d \equiv -1 \pmod{p_i}$ $\forall i = 1, \ldots, r$), and (21) gives us that $\left(\frac{d}{p}\right) = \left(\frac{-1}{p}\right)$. So combining this with (19), we have

$$\begin{aligned}\left(\frac{d}{1+q_1\cdots q_{n-1}}\right) &= \left(\frac{d}{p_1^{e_1}p_2^{e_2}\cdots p_r^{e_r}p^e}\right)\\ &= \left(\frac{d}{p_1}\right)^{e_1}\left(\frac{d}{p_2}\right)^{e_2}\cdots\left(\frac{d}{p_r}\right)^{e_r}\left(\frac{d}{p}\right)^{e}\\ &= \left(\frac{-1}{p_1}\right)^{e_1}\left(\frac{-1}{p_2}\right)^{e_2}\cdots\left(\frac{-1}{p_r}\right)^{e_r}\left(\frac{-1}{p}\right)^{e}\\ &= \left(\frac{-1}{p_1^{e_1}p_2^{e_2}\cdots p_r^{e_r}p^e}\right)\\ &= \left(\frac{d}{1+q_1\cdots q_{n-1}}\right)\\ &= (-1)^{((1+q_1\cdots q_{n-1})-1)/2}\\ &= -1.\end{aligned}$$

Where the final equality holds since $\frac{(1+q_1\cdots q_{n-1})-1}{2} = \frac{q_1\cdots q_{n-1}}{2} = \frac{2\cdot q_2\cdots q_{n-1}}{2} = q_2\cdots q_{n-1}$, which is odd.

Thus we have proven that given our supposition, we have that $\left(\frac{d}{1+q_1\cdots q_{n-1}}\right) = 1$ and $\left(\frac{d}{1+q_1\cdots q_{n-1}}\right) = -1$, a contradiction. Therefore we must have that $p_1, \ldots, p_r$ are not the only primes $\leq X$ that are omitted from (EM2). □

Remark 2.26. As mentioned previously, Booker's original proof provides a stronger bound in the form:

$$\limsup_{n\to\infty} \frac{\log p_{n+1}}{\log(p_1\cdots p_n)} \leq \frac{1}{4\sqrt{e}-1} = 0.1787....$$

This tells us that the exponent 2 in (18) can be replaced by any real number larger than $\frac{1}{4\sqrt{e}-1} = 0.1787...$ provided that the constant 12 be replaced by some larger constant.

3 Euclid-Mullin Sequences in $mk + 1$

There is a well known proof of the infinitude of primes in the arithmetic progression $mk + 1$ that utilises the cyclotomic polynomials. In this section we will cover this proof, as well as consider how it can be utilised to generalise the Euclid-Mullin sequences.

3.1 Cyclotomics

This section will follow Lemma 1.17 to Corollary 1.20 in Pollack's book "Not Always Buried Deep" [9]. The goal of this section will be to prove the following theorem:

Theorem 3.1 ([9, Corollary 1.20])**.** *For each natural number m, there are infinitely many primes $p \equiv 1 (\text{mod } m)$.*

To this end, we first introduce a few definitions and prove a couple necessary lemmata that will lead to our proof.

Definition 3.2 (Root of Unity)**.** A complex number $z \in \mathbb{C}$ is referred to as a *root of unity* if there exists a natural number, $m \in \mathbb{N}$, such that $z^m = 1$. In this case, z may be referred to as an *mth root of unity.* Each root takes the form $\zeta_m^k = e^{2\pi ik/m}$ for $1 \leq k \leq m$.

Further, an mth root of unity is called *primitive* if it is not an nth root of unity for some $n \in \mathbb{N}$ such that $n < m$, or equivalently if $\gcd(k, m) = 1$.

Remark 3.3. The mth roots of unity form a cyclic group generated by ζ_m: $\{1, \zeta_m, \zeta_m^2, \zeta_m^3, \ldots, \zeta_m^{m-1}\}$.

Example 3.4. The second roots of unity are exactly 1 and -1, which are the only solutions to $x^2 = 1$ in $\mathbb{C}$, whilst the fourth roots of unity are $1, -1, i, -i$, so the primitive fourth roots of unity are $\pm i$.

Definition 3.5 (Cyclotomic Polynomial)**.** The *mth cyclotomic polynomial* is defined as

$$\Phi_m(x) := \prod_{\substack{1 \leq k \leq m \\ \gcd(k,m)=1}} \left(x - e^{2\pi ik/m} \right). \tag{23}$$

i.e. $\Phi_m(x)$ is a monic polynomial in $\mathbb{C}$ which has the primitive mth roots of unity as its roots, each with multiplicity 1.

Remark 3.6. Note that there are exactly $\varphi(m)$ primitive mth roots of unity where $\varphi(m) = \#\{k \in \mathbb{N} \mid k < m \text{ and } \gcd(k, m) = 1\}$ is *Euler's Totient Function*, which counts the number of integers less than m which are coprime to m. Therefore it is clear that the degree of $\Phi_m(x)$ is exactly $\varphi(m)$. Further, $\varphi(m)$ is even for $m > 2$ and so cyclotomic polynomials have even degree excluding Φ_1 and Φ_2.

Another interesting consequence of the construction of these polynomials is that for $m > 1$, the sequence of coefficients is a palindrome, meaning that it reads the same forwards and backwards. This is because cyclotomic polynomials are actually a special case of polynomials known as self-reciprocals which satisfy the following condition: for f a polynomial with degree d, we have $x^d f(x^{-1}) = f(x)$. This will be shown for cyclotomic polynomials in Corollary 3.15.

Example 3.7. The first few cyclotomic polynomials are:
$\Phi_1(x) = (x - e^{2\pi i}) = x - 1$,
$\Phi_2(x) = (x - e^{2\pi i/2}) = x + 1$,
$\Phi_3(x) = (x - e^{2\pi i/3})(x - e^{4\pi i/3}) = (x + \frac{1}{2} - \frac{\sqrt{3}i}{2})(x + \frac{1}{2} + \frac{\sqrt{3}i}{2}) = x^2 + x + 1$,
$\Phi_4(x) = (x - e^{2\pi i/4})(x - e^{6\pi i/4}) = (x - i)(x + i) = x^2 + 1$.

This is not the only way to construct the cyclotomic polynomials, in fact there is a recursive formula that allows you to construct any of them through polynomial division over $\mathbb{Z}[x]$. This recursive method relies primarily on the following relation

$$\begin{aligned} x^m - 1 &= \prod_{1 \le k \le m} \left(x - e^{2\pi i k/m} \right) \\ &= \prod_{d|m} \prod_{\substack{1 \le k \le m \\ \gcd(k,m)=d}} \left(x - e^{2\pi i k/m} \right) \\ &= \prod_{d|m} \Phi_{m/d}(x) = \prod_{d|m} \Phi_d(x). \end{aligned}$$

The second line follows since every integer k less than m will have $\gcd(k, m)$ equal to some divisor of m, and the third line follows from the fact that the roots of unity with $\gcd(k, m) = d$ are exactly the primitive $\frac{m}{d}$th roots of unity. Then finally summing over d or m/d will both cover every divisor of m.

This can then be rearranged to give the formula

$$\Phi_m(x) = \frac{x^m - 1}{\prod_{\substack{d|m \\ d<m}} \Phi_d(x)}. \tag{24}$$

So starting with $m = 1$, there are no divisors so we just recover $\Phi_m(x) = x - 1$. Then for $m = 2$, $\Phi_2(x) = \frac{x^2-1}{x-1} = x + 1$, $\Phi_3(x) = \frac{x^3-1}{x-1} = x^2 + x + 1$, $\Phi_4(x) = \frac{x^4-1}{(x-1)(x+1)} = x^2 + 1$ and so on. This is often an easier method for calculating $\Phi_m(x)$ for large m if you know $\Phi_d(x)$ for all divisors d of m as it doesn't involve calculating relations of primitive roots of unity.

Remark 3.8. One interesting consequence of this is that it provides a clear form for $\Phi_p(x)$ where p is prime. Since the only divisors of p are 1 and p itself, we have $\Phi_p(x) = \frac{x^p-1}{x-1} = 1 + x + x^2 + \cdots + x^{p-1}$

Example 3.9. Consider $m = 15$. 15 has proper divisors $1, 3, 5$ and from the above remark we know that $\Phi_3(x) = x^2 + x + 1, \Phi_5(x) = x^4 + x^3 + x^2 + x + 1$, so

$$\begin{aligned}\Phi_{15}(x) &= \frac{x^{15}-1}{\Phi_1(x)\Phi_3(x)\Phi_5(x)} \\ &= \frac{x^{15}-1}{(x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)} \\ &= x^8 - x^7 + x^5 - x^4 + x^3 - x + 1.\end{aligned}$$

A classic infinitude of primes proof involving cyclotomic polynomials is the proof that there are infinitely many primes such that $p \equiv 1 \pmod 4$.

Theorem 3.10. *There are infinitely many primes p satisfying $p \equiv 1$ (mod 4).*

Proof. Take a finite list of primes that satisfy the congruence relation, say $p_1, p_2, ..., p_k$. Then consider the number $a = 2p_1p_2 \cdots p_k$. Plugging this into $\Phi_4(x) = x^2 + 1$ we get $\Phi_4(a) = (2p_1p_2 \cdots p_k)^2 + 1 = 4(p_1p_2 \cdots p_k)^2 + 1 =: N$. It is clear that $N \equiv 1(\text{mod } 4)$ and $N > 1$. We thus have that there exists some prime p that divides N. This p cannot be any of the $p_1, p_2, ..., p_k$ since $N \equiv 1(\text{mod} p_i)$ $\forall$ $1 \leq i \leq k$. So N has some prime divisor not in the list. We then have

$$\begin{aligned}p|N \implies p|a^2+1 &\implies a^2 \equiv -1 \pmod p \\ &\implies a^4 \equiv \quad 1 \pmod p\end{aligned}$$

By Fermat's Little Theorem we have that $4|p-1$ so $p \equiv 1 \pmod 4$ as required. Thus no finite list of primes of the form $4k + 1$ is complete and we have the desired result. □

Definition 3.11 (Prime Divisor of a polynomial)**.** Let f be a nonzero polynomial with integer coefficients ($f(x) \in \mathbb{Z}[x]$). We say that a prime p is a *prime divisor* of f if p divides $f(n)$ for some $n \in \mathbb{Z}$.

Lemma 3.12 ([9, Lemma 1.17])**.** *Let $f(x) \in \mathbb{Z}[x]$ be a non-constant polynomial. Then f has infinitely many prime divisors.*

Proof. If $f(0) = 0$, then every prime divides f and we are done. So assume otherwise. Let a_0 denote the constant term of f, such that $f(0) = a_0$ and then consider $g(x) = f(x)/a_0$ which is such that $g(0) = 1$. We will show that g has infinitely many prime divisors. Let $p_1, ..., p_k$ be a finite list of prime divisors of g, and consider $g(mp_1 \cdots p_k) =: A$ where m is taken large enough that $A \in \mathbb{Z}$ and $|A| > 1$. Thus A has some prime divisor not in $p_1, ..., p_k$ since $A \equiv 1(\text{mod } p_i)$ $\forall i$ such that $1 \leq i \leq k$. Thus any finite list of prime divisors of g is incomplete and so g has infinitely many prime divisors. Since $f(a_0x) = a_0g(x)$ it follows that f also has infinitely many prime divisors. □

Lemma 3.13 ([9, Lemma 1.18])**.** *For $m \in \mathbb{N}$, $\Phi_m(x)$ has integer coefficients.*

Despite this Lemma being from [9], the proof there uses Möbius Inversion which would require some significant introduction. Instead we follow a different proof that uses induction, as presented in [10].

Proof (from Proposition 45 of [10]). For $n = 1$, $\Phi_1(x) = x - e^{2\pi i} = x - 1 \in \mathbb{Z}[x]$. Assume that $\Phi_n(x)$ has integer coefficients for $n < m$. We then claim that

$$x^m - 1 = \prod_{d|m} \Phi_d(x) = \underbrace{\left(\prod_{\substack{d|m \\ d<m}} \Phi_d(x) \right)}_{(\ddagger)} \cdot \Phi_m(x) \tag{25}$$

To see the first equality, note that the mth roots of unity are exactly the disjoint union of the primitive dth roots of unity for all d such that $d|m$. The first product obtained in the second equality, ($\ddagger$), is monic (leading coefficient is 1) by definition of $\Phi_d(x)$, and also has integer coefficients by the induction hypothesis. So $\Phi_m(x)$ is obtained by dividing a polynomial with integer coefficients by a monic polynomial with integer coefficients which implies that $\Phi_m(x)$ has integer coefficients. To see this, consider that the product of a monic polynomial in $\mathbb{Z}[x]$ with another polynomial with any coefficients in $\mathbb{Q} \setminus \mathbb{Z}$ would not have integer coefficients. The induction is thus complete. □

Before we move on, first note that Lemma 3.13 has the following corollary:

Corollary 3.14. *For all positive integers, $m \in \mathbb{N}$, $\Phi_m(x)$ has constant term ± 1. In particular, for $m \geq 2$, $\Phi_m(x)$ has constant term 1.*

Proof. By Lemma 3.13, we know that $\Phi_m(x)$ has integer coefficients, then using equation (25) established in the proof of the same Lemma, it is clear that

$$\Phi_m(x) \mid x^m - 1.$$

It follows that the constant term of $\Phi_m(x)$ must divide -1, and the only integers that do so are ± 1.

For the particular case, note that $\Phi_2(x) = x + 1$, so take $m > 2$. If m is odd, then -1 is not an mth root of unity $((-1)^m = -1$, and if m is even, then -1 is an mth root of unity, but is not primitive since $\gcd(m/2, m) > 1$. Therefore $\Phi_m(x)$ does not have ± 1 as roots (since 1 would also not be primitive). Thus the only roots are the complex primitive roots of unity, which come in reciprocal pairs. To see this note that $(e^{2\pi ik/m})^{-1} = e^{-2\pi ik/m} = e^{2\pi i(m-k)/m}$ since e is 2π periodic. Then note that $\gcd(k, m) = 1 \implies \gcd(m - k, m) = 1$ since gcd is preserved when one variable is taken modulo the other, and when one variable is negated. This means that if $e^{2\pi ik/m}$ is a primitive root of unity, so is its reciprocal. Finally, since the constant term is the product of all of the roots, which can be split up into reciprocal pairs each pair multiplying together to give 1, it is clear that $\Phi_m(0) = 1$ for $m \geq 2$. □

This Corollary allows us to prove our previous claim that the cyclotomic polynomials are self-reciprocal for $m \geq 2$.

Corollary 3.15. *For $m \geq 2$, $\Phi_m(x)$ is self-reciprocal, i.e. it satisfies:*

$$\Phi_m(x) = x^{\varphi(m)} \Phi_m(x^{-1}).$$

Proof. Take $m \geq 2$. By construction $\Phi_m(x)$ is a monic polynomial and has constant term 1 by Corollary 3.14. We showed in the proof of that same Corollary that all mth primitive roots of unity come in reciprocal pairs, such that the reciprocal of any root of $\Phi_m(x)$ is also a root. Thus consider $x^{\varphi(m)}\Phi_m(x^{-1})$, noting that $\varphi(m)$ is the degree of Φ_m. So we have

$$\begin{aligned} x^{\varphi(m)}\Phi_m\left(\frac{1}{x}\right) &= x^{\varphi(m)} \prod_{\substack{1\leq k\leq m \\ \gcd(k,m)=1}} \left(\frac{1}{x} - e^{2\pi ik/m}\right) \\ &= \prod_{\substack{1\leq k\leq m \\ \gcd(k,m)=1}} \left(1 - xe^{2\pi ik/m}\right). \end{aligned}$$

Clearly the roots of $x^{\varphi(m)}\Phi_m(x^{-1})$ are the values where each bracket in the final product above is equal to 0, i.e. $1 - xe^{2\pi ik/m} = 0$ which are exactly $(e^{2\pi ik/m})^{-1}$. Finally note that since $\Phi_m(x)$ has constant term 1, $x^{\varphi(m)}\Phi_m(x^{-1})$ is monic.

Combining all this, we have that $\Phi_m(x)$ and $x^{\varphi(m)}\Phi_m(x^{-1})$ are both monic, and have the same roots, and thus must be the same polynomial. □

We now prove another interesting case that allows easier computation of some cyclotomic polynomials. This will be required in a later proof.

Theorem 3.16. *For p an odd prime*

$$\Phi_{2p}(x) = \Phi_p(-x) = 1 - x + x^2 - \cdots - x^{p-2} + x^{p-1}.$$

Proof. Note that the proper divisors of $2p$ are exactly $1, 2,$ and p, and that $\Phi_1(x)\Phi_p(x) = x^p - 1$ by Remark 3.8. So we have

$$\begin{aligned} \Phi_{2p}(x) &= \frac{x^{2p} - 1}{\Phi_1(x)\Phi_2(x)\Phi_p(x)} \\ &= \frac{(x^p + 1)(x^p - 1)}{\Phi_2(x) \cdot (x^p - 1)} \\ &= \frac{x^p + 1}{x + 1} \\ &= 1 - x + x^2 - \cdots + x^{p-1}, \end{aligned}$$

where the final equality follows from the well known identity

$$a^n + b^n = (a + b)(a^{n-1} - a^{n-2}b + \cdots - b^{n-2}a + b^{n-1}).$$ □

Remark 3.17. This actually holds for general $m > 1$ odd, i.e.:

$$\Phi_{2m}(x) = \Phi_m(-x) \quad \text{for } m > 1 \text{ odd}$$

but this is more arduous to prove and not needed for the later proof. It follows from the fact that if $\zeta, \zeta^2, \ldots, \zeta^k$ are the primitive mth roots of unity for m odd, then the primitive $2m$th roots of unity are exactly $-\zeta, -(\zeta^2), \ldots, -(\zeta^k)$.

We are now ready to categorise the prime divisors of Φ_m.

Lemma 3.18 ([9, Lemma 1.19])**.** *If p is a prime divisor of $\Phi_m(x)$, then either $p|m$ or $p \equiv 1(\text{mod } m)$.*

The following proof is adapted from [9, 10].

Proof. If p is a prime divisor of $\Phi_m(x)$, then $p|\Phi_m(a)$ for some $a \in \mathbb{Z}$ such that $\Phi_m(a) \neq 0$. Take k to be the order of a modulo p, i.e. $a^k \equiv 1 \pmod{p}$. By (25) we have that $p| \prod_{d|m} \Phi_d(a) = a^m - 1$. We therefore have that $k|m$.

Suppose now that p does not divide m. We will show that in this case, we have $k = m$, and therefore Fermat's Little Theorem tells us that $m|p-1$ so $p \equiv 1 \pmod{m}$. For the sake of contradiction suppose that $k < m$. Then again by (25), there exists some $d|k$ such that $p|\Phi_d(a)$. Since $p|\Phi_d(a)$ and $p|\Phi_m(a)$, a is a root of order ≥ 2 of $\Phi_d\Phi_m \pmod{p}$ and thus also of $x^m - 1 \pmod{p}$. Writing $x^m - 1 = (x-a)^2 g(x)$ for some function g, it is clear that a is therefore also a root of the derivative of $x^m - 1$. So $ma^{m-1} \equiv 0 \pmod{p}$ but we know that $p \nmid m$ and $p \nmid a$ (since the constant term of a cyclotomic polynomial is always ± 1) so we have a contradiction. Therefore, either $p|m$ or $p \equiv 1 \pmod{m}$ as required. □

We now have all we need to prove Theorem 3.1, which will be rather trivial using the Lemmas we have provided.

Proof of Theorem 3.1. By Lemma 3.13, we know that $\Phi_m(x)$ has integer coefficients for all $m \in \mathbb{N}$. We can therefore apply Lemma 3.12 to see that $\Phi_m(x)$ has infinitely many prime divisors. Finally, applying Lemma 3.18 tells us that all prime divisors of $\Phi_m(x)$, p, either divide m or satisfy $p \equiv 1 \pmod{m}$. To bring it all together, observe that m can only have finitely many prime divisors, and thus there must be infinitely many primes satisfying the congruence relation. □

3.2 The Method

The proof of Lemma 3.12 is reminiscent of Euclid's original proof of the infinitude of primes. We now outline a method capable of generating a sequence of primes all congruent to 1 modulo m, to that end we will provide another proof of Theorem 3.1 to emphasise its constructive nature, similarly to Euclid's original proof.

Theorem 3.1. *For each natural number m, there are infinitely many primes $p \equiv 1(\text{mod } m)$.*

Proof. Suppose that there is a finite list of primes $p_1, p_2, \ldots, p_n$ all congruent to 1 modulo m. Call a the product of these primes and consider $N := \Phi_m(m \cdot a)$. If $|N| < 1$ then instead consider $N := \Phi_m(lma)$ where l is taken large enough to ensure that $|N| > 1$. Now we have that $N \equiv \pm 1 \pmod{p_i}$ for all $i = 1, \ldots, n$ and $|N| > 1$ so N must have some prime divisor p that is not equal to any of the

p_i. Now by Lemma 3.18 we must have that $p \mid m$ or $p \equiv 1 \pmod{m}$, but since $N \equiv \pm 1 \pmod{m}$ we must have that $p \nmid m$ and so $p \equiv 1 \pmod{m}$. Thus we have found another prime congruent to 1 modulo m not in our original list. □

And so, similar to Mullin we can use this constructive proof to outline a method of generating a sequence of primes all congruent to 1 modulo m.

Method 1.

1. Consider some positive integer m and a list of primes $p_1, \ldots, p_k$ that all satisfy $p_i \equiv 1 \pmod{m}$,
2. Consider $N = \Phi_m(m \cdot a)$,
3. If $|N| \leq 1$, then set $a = l \cdot p_1 \cdots p_k$ for some l large enough that $|N| > 1$, and repeat from step 2,
4. Since $|N| > 1$, N must have some prime divisor, say p, not equal to any of the $p_1, \ldots, p_k$ and also not a divisor of m, so $p \equiv 1 \pmod{m}$.
5. Set $p_{k+1} = p$.

The method outlined above can be performed on any list of primes that satisfy the congruence, including the empty list. In the case that the list is empty, we would have $a = 1$, in which case the congruences still hold and the method can be continued as normal.

In most cases that we consider, step 3 will not be necessary, rather it is just included for the odd outlier case. One example of this is when $m = 1$. This case is already strange since every prime number satisfies $p \equiv 1 \equiv 0 \pmod{1}$, but nonetheless we can still apply the method. Starting with the empty list, we get $\Phi_1(1) = 1 - 1 = 0$. Thus we need to use step 3. Note that $l = 3$ is the smallest value we can take, which generates the prime 2. To generate the second prime in the list, the use of step 3 is once again required, this time with $l \geq 2$. With $l = 2$, $\Phi_1(2 \cdot 2) = 4 - 1 = 3$. Beyond this point, step 3 is not required.

Notice that we run into the same problem that Mullin discussed, that as to which prime divisor we take of N in step 4. This similarly allows for differing lists of primes that satisfy the congruence to emerge based on which prime divisor is chosen to be the next term in the sequence.

We propose a name for these sequences.

Definition 3.19. Call the sequence of primes all congruent to 1 modulo m generated by taking the smallest (respectively largest) prime divisor in Method 1, *the First (respectively second) Euclid-Mullin Sequence in* $mk + 1$.

Example 3.20. Consider applying Method 1 to generate the Euclid-Mullin sequence in $2k + 1$ ($m = 2$). Starting with the empty list, we get $a = 1$, so we plug $2 \cdot 1$ into $\Phi_2(x) = x + 1$, to get $\Phi_2(2) = 2 + 1 = 3$. Clearly $3 \equiv 1 \pmod{2}$ so we take $p_1 = 3$. Continue this to get $\Phi_2(2 \cdot 3) = 7$, so take $p_2 = 7$. Then $\Phi_2(2 \cdot 3 \cdot 7) = 43$, so $p_3 = 43$. Continue with $\Phi_2(2 \cdot 3 \cdot 7 \cdot 43) = 1807 = 13 \cdot 139$

so we get a choice for p_4. The astute reader will recognise this sequence as the original Euclid-Mullin sequence with a slight difference, that being that the sequences start from 3, excluding 2. This is accounted for however since at each step we are considering $2 \cdot p_1 \cdots p_k$. Appending 2 to the start of the sequence, and taking $p_4 = 13$ will give us the first Euclid-Mullin sequence, (EM1), whislt $p_4 = 139$ will give us the second, (EM2).

Example 3.21. Method 1 is a slight alteration of the one outlined in the proof of Theorem 3.10, using $a = 4 \cdot p_1 \cdots p_k$ rather than $a = 2 \cdot p_1 \cdots p_k$. In the proof outlined for the theorem, using the former a rather than the latter would also produce a valid proof.

If we wanted to generate the first Euclid-Mullin sequence in $4k+1$ we can simply apply Method 1.

Primes	$\Phi_4(4 \cdot p_1 \cdots p_k)$	Factors
$\emptyset$	$4^2 + 1 = 17$	17
$\{17\}$	$68^2 + 1 = 4625$	$5, 37$
$\{17,5\}$	$(4 \cdot 17 \cdot 5)^2 + 1 = 115601$	115601

On the other hand, if we follow the method outlined in the proof of theorem 3.10, we would get:

Primes	$\Phi_4(2 \cdot p_1 \cdots p_k)$	Factors
$\emptyset$	$2^2 + 1 = 5$	5
$\{5\}$	$10^2 + 1 = 101$	101
$\{5, 101\}$	$1010^2 + 1 = 1020101$	1020101

Example 3.22. Now consider an odd m, we will take $m = 3$ for this example. Applying the method with $\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$ and taking the greatest factor at each step, we get the Second Euclid-Mullin Sequence in $5k+1$.

Starting with the empty list, we calculate $\Phi_5(5) = 781$, with factors $11, 71$, so we take $p_1 = 71$. Then $\Phi_5(5 \cdot 71) = 15927165881 = 11 \cdot 1447924171$, so $p_2 = 1447924171$. The next step of the iteration, $\Phi_5(5 \cdot 71 \cdot 1447924171)$ gives us a number with 47 digits, with factors $11, 31, 41, 311, 12011, a, b$ (where a, b have 9 and 28 digits respectively), and we take the largest of these to be the next term. The next iteration gives us a number with 158 digits, with largest divisor having 136 digits. One more iteration would require us to factorise a number with 699 digits.

This example helps to illuminate two things. Firstly, that the second Euclid-Mullin sequence takes significantly more computing power to calculate than the first, as the main way to find the largest prime divisor of a number is to fully factorise it. In comparison, the first Euclid-Mullin sequence only requires the smallest prime divisor of a number, one that is often quite small. For example for that 699-digit number, it is easy to verify that the smallest prime divisor is 1171 by simple brute-force checking whilst it is much harder to fully factorise such a large number.

The second thing it shows us is that when we apply the method to an odd m, it actually only generates primes $p \equiv 1 \pmod{2m}$. This is simple to see when you consider that if $p \equiv 1 \pmod{m}$ then $p \equiv 1 \pmod{2m}$ or $p \equiv m+1 \pmod{2m}$ and it is easy to see that in the second case p must be even, and thus not prime (except for 2). In our example it is clear that every factor found is $1 \pmod{10}$. This tells us that we have in fact defined two methods of generating primes $p \equiv 1 \pmod{10}$, one using $\Phi_5(x)$ and the other $\Phi_{10}(x)$. This may motivate us to ask whether or not these sequence generate the same primes, or in fact different ones. This question will be considered further in the next section.

Example 3.21 shows that the method outlined in Method 1 is not the only way to generate the Euclid-Mullin sequence in $mk+1$. This may motivate us to question for which multiples of a as defined in step 1 the method holds. In the method, a is defined as $a = p_1 \cdots p_k$ and then the value of $N := \Phi_m(m \cdot a)$ is considered. We multiply a by m to ensure that $N \equiv \pm 1 \pmod{m}$, so that $p \nmid m$ and by Lemma 3.18 we must have that $p \equiv 1 \pmod{m}$ but in many, if not all, cases this m is not the only choice we can use. Note that since $p_1, \ldots, p_k \equiv 1 \pmod{m}$ we must have that $a = p_1 \cdots p_k \equiv 1 \pmod{m}$ and so it follows that $\Phi_m(n \cdot a) \equiv \Phi_m(n) \pmod{m}$. This shows us that this problem reduces to understanding what values $\Phi_m(n)$ can take on the residue classes of m (i.e. what values we get when we put $0, 1, 2, \ldots, m-1$ into $\Phi_m(n)$). This may sound simple, and whilst in some cases it is, in others it is not. We consider a few simple cases below.

Theorem 3.23. *Take p to be any prime,*

(i) $\Phi_p(n) \equiv 1 \pmod{p}$ *for all* $n \not\equiv 1 \pmod{p}$. *When* $n \equiv 1 \pmod{p}$ *we have* $\Phi_p(n) \equiv 0 \pmod{p}$.

(ii) *If p is odd then* $\Phi_{2p}(n) \equiv 1 \pmod{2p}$ *for all* $n \not\equiv -1 \pmod{p}$. *When* $n \equiv -1 \pmod{p}$ *we have* $\Phi_{2p}(n) \equiv p \pmod{2p}$.

Proof. (i) Recall from Remark 3.8 that we have $\Phi_p(x) = \frac{x^p-1}{x-1} = 1 + x + x^2 + \cdots + x^{p-1}$. In the case when $n \equiv 1 \pmod{p}$, we have that $\Phi_p(n) \equiv \Phi_p(1) \equiv 1 + 1 + 1 + \cdots + 1 \pmod{p}$. It is clear that there are exactly p 1's in this sum, and so $\Phi_p(n) \equiv 0 \pmod{p}$ as claimed. Now if $n \not\equiv 1 \pmod{p}$, then by Fermat's Little Theorem we have that $n^p \equiv n \pmod{p}$ so $\Phi_p(n) = \frac{n^p-1}{n-1} \equiv \frac{n-1}{n-1} \equiv 1 \pmod{p}$.

(ii) From Theorem 3.16, we know that

$$\begin{aligned}\Phi_{2p}(x) = \Phi_p(-x) &= 1 - x + x^2 - \cdots + x^{p-1} \\ &= \frac{(-x)^p - 1}{x-1} = \frac{-x^p-1}{x-1}\end{aligned}$$

since p is odd. Now take $n \not\equiv -1 \pmod{p}$. Fermat's Little Theorem tells us that $n^p \equiv n \pmod{p}$ so

$$\Phi_{2p}(n) = \frac{-n^p-1}{-n-1} \equiv \frac{-n-1}{-n-1} = 1 \pmod{p}.$$

Note that powers preserve congruence modulo 2 (since $0^n = 0, 1^n = 1$) so $-x + x^2 - \cdots + x^{p-1}$, a sum of an even number of terms that are either all even or all odd, is even. We must have that $\Phi_{2p}(n) \equiv 1 \pmod 2$, which implies that $\Phi_{2p}(n) \equiv 1 \pmod{2p}$ since it cannot be $\equiv p + 1 \pmod{2p}$.

Now take $n \equiv -1 \pmod p$, then

$$\Phi_{2p}(n) = \Phi_p(-n) \equiv 0 \pmod p$$

by part (i). Then by the same argument as before, $\Phi_{2p}(n) \equiv 1 \pmod 2$ so we must have that $\Phi_{2p}(n) \equiv p \pmod{2p}$ since it is divisible by p but not 2.

□

Some other observations that seem to hold are:[2]

Conjecture 3.24. (i) $\Phi_m(n)$ is always congruent to either $0, 1,$ or $q \pmod m$ for every $m, n \in \mathbb{N}$ where q is the largest prime divisor of m.

(ii) $\Phi_m(n) \equiv 0 \pmod m$ only in the case presented in the previous theorem (and trivially when $m = 1$).

There are also some cases where $\Phi_m(n) \equiv 1 \pmod m$ for any $n \in \mathbb{Z}$. Take for example $m = 24$ with $\Phi_{24}(x) = x^8 - x^4 + 1$.

Theorem 3.25. $\Phi_{24}(n) \equiv 1 \pmod{24}$ *for any integer* n.

Proof. By Lemma 3.18 we must have that any prime divisor p of $\Phi_{24}(x)$ must satisfy $p \mid 24$ or $p \equiv 1 \pmod{24}$ so if we can show that $2, 3 \nmid \Phi_{24}(n)$ for all $n \in \mathbb{Z}$ then we must have that all prime divisors satisfy the latter condition. First note that for $n = -1, 0, 1$, $\Phi_{24}(n) = 1$ so it does not have any prime divisors. If $|n| > 1$ then $\Phi_{24}(n) > 3$ so must have some prime divisor. Observe that $n^8 - n^4$ is always even and so $\Phi_{24}(n)$ must always be odd, and thus $2 \nmid \Phi_{24}(n)$ for all $n \in \mathbb{Z}$. Now consider $n \equiv -1, 0, 1 \pmod 3$. Similarly to before, in this case $\Phi_{24}(n) \equiv 1 \pmod 3$ so $3 \nmid \Phi_{24}(n)$ for $n \equiv -1, 0, 1 \pmod 3$. Since all $n \in \mathbb{Z}$ fits into one of these congruences we must have that $3 \nmid \Phi_{24}(n)$ for all $n \in \mathbb{Z}$. Thus we must have that any prime divisor of $\Phi_{24}(x)$ satisfies $p \equiv 1 \pmod{24}$ and so $\Phi_{24}(n) \equiv 1 \pmod{24}$ for all $n \in \mathbb{Z}$ as required. □

Φ_{24} is not the only cyclotomic polynomial to satisfy this property. The first few are $m = 12, 15, 24, 28, 30, 33, ...$ and in fact 569 of the first 1000 cyclotomic polynomials satisfy this property. It is likely that a similar proof to above is possible in each case.

Notation. We may wish to address this ambiguity when talking about a sequence. To that end, we may refer to *The (first/second) Euclid-Mullin Sequence in* $mk+1$ *generated by* $f(x)$, where $f(x)$ is the function that we put the product of our list of primes, x, into.

[2] These observations come from looking at the values that the first 1000 cyclotomic polynomials can take, see [8].

Example 3.26. The sequences considered in Example 3.21 would be the first Euclid-Mullin sequences in $4k+1$ generated by $\Phi_4(4x)$ and $\Phi_4(2x)$ respectively. In a similar manner, the sequences generated by Method 1 would be the Euclid-Mullin sequences in $mk+1$ generated by $\Phi_m(mx)$.

3.3 Questions

In a similar vein to Mullin in [1], we can ask some questions about the Euclid-Mullin sequences in $mk+1$.

3.3.1 For the First Sequence

Throughout this section, refer to $(p_i)_{m,a}$ as the First Euclid-Mullin sequence in $mk+1$ generated by $\Phi_m(ax)$. If a is not specified then $a = m$.

Question 1. Does $(p_i)_m$ contain all primes congruent to 1 modulo m? If not, is there an $a \in \mathbb{N}$ such that $(p_i)_{m,a}$ contains all primes congruent to 1 modulo m?

We know that $(p_i)_2$ is the original First Euclid-Mullin sequence (except 2), which is conjectured to contain all primes. It seems likely that a similar conjecture can be given for any m. In fact, in Shanks' conjecture (Conjecture 1.5) the only difference would be in (1), where we would instead have that (with the notation from the conjecture)

$$q = p_{n+1} \iff \Phi_m(mr_1r_2) \equiv 0 \pmod{q}$$

and if $\Phi_m(mr_1r_2)$ can take any residue between 1 and $q-1$ inclusive then again the conjecture would hold. If $\Phi_m(mr_1r_2)$ cannot take any residue, there may be an a such that $\Phi_m(ar_1r_2)$ can, and then again the conjecture could be applied there.

Question 2. Is $(p_i)_{nm}$ a subsequence of $(p_i)_m$?

Clearly if question 1 does hold, then every term in $(p_i)_{nm}$ will be contained in $(p_i)_m$ since if a number is $1 \pmod{nm}$ then it will also be $1 \pmod{m}$. For example, we have seen that the method for $3k+1$ generates a sequence in $6k+1$ due to the nature of the congruences, so every term in $(p_i)_3$ would be contained in $(p_i)_6$ if both contain all the primes satisfying their congruences. But more interestingly is to ask if one is a subsequence of the other, or if the terms in $(p_i)_{nm}$ appear in $(p_i)_m$ in the same order (possibly with some other terms in between).

If question 1 does not hold, then we cannot be sure that every term in $(p_i)_{nm}$ will occur in $(p_i)_m$ and so this question will likely be more difficult to answer, unless for example a condition can be found to show that there exists a prime in $(p_i)_{nm}$ that does not exist in $(p_i)_m$. One way to do this would be to discover whether or not each sequence is recursive like Mullin originally asked, i.e. if there is a way to tell whether or not a prime will occur in each sequence.

Question 3. If question 1 is answered negatively, does $\bigcup_{m\in\mathbb{N}}(p_i)_m$ contain all primes?

If question 1 proves true, then clearly the union will contain all primes since $(p_i)_2$ would contain all odd primes, and $(p_i)_1$ contains 2 as seen in the discussion after Method 1.

If not, then this question will prove more difficult. Every odd prime will satisfy $p \equiv 1 \pmod{m}$ for multiple $m \in \mathbb{N}$ (e.g. $m = p-1, \frac{p-1}{2}$) and if the prime occurs in $(p_i)_m$ for one of these m then the question will hold, but this is likely rather difficult to show.

An extension of this would be to consider the union of all sequences generated by $\Phi_m(ax)$ for any a such that $\Phi_m(a) \equiv 1 \pmod{m}$ for all m, i.e.

$$\bigcup_{\substack{m\in\mathbb{N}\\ a\in A_m}} (p_i)_{m,a} \quad \text{for} \quad A_m = \{a \in \{0,1,\dots,m-1\} : \Phi_m(a) \equiv 1 \pmod{m}\}.$$

It seems more likely that this would contain all primes compared to the original union, since significantly more sequences are contained in the second union.

3.3.2 For the Second Sequence

Throughout this section, refer to $(q_i)_{m,a}$ as the Second Euclid-Mullin sequence in $mk+1$ generated by $\Phi_m(ax)$. If a is not specified then $a = m$.

Question 4. Is $(q_i)_m$ ever an increasing sequence?

We have seen a few examples such that $(q_i)_m$ is not monotonically increasing, take for example $(q_i)_2$, the original second Euclid-Mullin sequence. Another example is $(q_i)_4$ as seen in Example 3.21. But is the sequence ever increasing? This seems quite unlikely but is difficult to rule out as it would likely require manual computation of each sequence until a contradiction is found.

Question 5. Does $(q_i)_m$ contain all primes congruent to 1 modulo m? If not, is there an a such that $(q_i)_{m,a}$ contains all primes congruent to 1 modulo m?

Yet again, we know of at least one case where this does not hold, namely with the original second Euclid-Mullin sequence $(p_i)_2$ and so it seems likely that this will not hold in any case. In fact this is likely provable using a similar method to Cox and Van Der Poorten's original proof and conjecture in [4]. It is also likely that each sequence misses infinitely many primes, with a proof likely existing in a similar vein to Booker's in [5].

Question 6. If question 5 is answered negatively, is $(q_i)_{nm}$ a subsequence of $(q_i)_m$?

This question is likely more difficult to rule out for every case. We can rule it out for example in one case, where we know from Example 3.21 that $(q_i)_4$ contains 17 yet from section 2 we know that $(q_i)_2$ does not. This was found through manual calculation, and it is likely that all cases will require this. However if the sequences can be shown to be recursive, then this will likely be easier.

Question 7. If question 5 is answered negatively, does $\bigcup_{m\in\mathbb{N}}(q_i)_m$ contain all primes?

This question is likely fairly straight forward to answer if there is a small prime that can be proven to not occur in any sequences. For example the number 5 does not appear in $(q_i)_2$ and it's only other chance to appear will be in $(q_i)_4$, so if it can be shown to not occur then this question is false.

Similarly to above though we can ask whether the union

$$\bigcup_{\substack{m\in\mathbb{N}\\ a\in A_m}}(q_i)_{m,a} \quad \text{for} \quad A_m=\{a\in\{0,1,\dots,m-1\}:\Phi_m(a)\equiv 1 \pmod{m}\}.$$

hits all primes. If we go back to 5 then it is contained in $(q_i)_{4,2}$ as seen in Example 3.21 so we would have to find a larger prime to check. We know 7 occurs in $(q_i)_2$ so the next smallest to check would be 11 in $(q_i)_{5,a}$ or $(q_i)_{10,b}$ for all $a\in A_5$ and $b\in A_{10}$ (since we know it doesn't occur in $(q_i)_2$.

4 Euclid-Mullin Sequences in $mk + a$

In the previous section, we have defined a method of generating a Euclid-Mullin like sequence of primes all congruent to 1 modulo m. A natural extension of this is to question whether or not similar sequences exist for primes congruent to a modulo m.

There is a famous theorem from Dirichlet regarding primes in general arithmetic sequences which is stated below.

Theorem 4.1 (Dirichlet's Theorem)**.** *If a and m are coprime positive integers, then there exist infinitely many primes p such that $p \equiv a \pmod{m}$. In other words, there are infinitely many primes in the arithmetic sequence*

$$a, a + m, a + 2m, a + 3m, \ldots$$

Proof. See, for example, [11]. □

Note that if a and m are not coprime, then the arithmetic sequence $a, a + m, a+2m, \ldots$ cannot contain infinitely many primes as each term $a+km$, $k > 0$ will be divisible by $\gcd(a, m)$ and thus not prime.

More relevant to us however is considering when exactly such a sequence permits a Euclidean proof, and therefore a method to construct a Euclid-Mullin sequence in $mk+a$. But first it will be helpful to consider and eventually define what exactly is meant by a Euclidean proof.

4.1 Euclidean Proof

This section will lightly follow the exposition of [12].

Consider the Euclidean style proofs that we have looked at so far. Euclid's original proof used the polynomial $x + 1$, and the proof for $1 \pmod{m}$ used the cyclotomic polynomials. This suggests that the utilisation of polynomials may play an important role. However, these cases are both for the congruence $p \equiv 1 \pmod{m}$, is there any difference for $p \equiv a \pmod{m}$? Consider the following example.

Theorem 4.2. *There are infinitely many primes p satisfying $p \equiv 3 \pmod{4}$.*

Proof. Consider a finite list of primes $p_1, \ldots, p_k$ all congruent to 3 modulo 4. Now consider the polynomial $g(x) = 4x-1$. Then $g(p_1 \cdots p_k) = 4(p_1 \cdots p_k)-1 \equiv 3 \pmod{4}$ must have prime divisors each congruent to 1 or 3 modulo 4 and none equal to any of the $p_1, \ldots, p_k$. If all of the prime divisors of $N := g(p_1 \cdots p_k)$ are congruent to 1 modulo 4, then their product would also be congruent to 1 modulo 4. Thus N must have some prime divisor congruent to 3 modulo 4 not equal to any of the $p_1, \ldots, p_k$. □

This is another clear example of a Euclidean style proof, once again utilising a polynomial. Note further that in each proof the polynomial is constructed in such a way that its values at integer points are divisible by primes satisfying the required congruence. In the case of the proof for Theorem 3.10 which used $f(x) = 4x^2 + 1$, all (but finitely many) prime divisors take the form $p \equiv 1 \pmod 4$, whilst in the case of $g(x) = 4x - 1$ at least one prime divisor at each integer point is $p \equiv 3 \pmod 4$. In either case however, clearly the set of prime divisors of the polynomial contains an infinite number of primes satisfying the required congruence, a condition also satisfied by each cyclotomic polynomial, as outlined in Theorem 3.1. Thus a logical condition for a Euclidean Proof of the infinitude of primes in the progression $km + a$ is that a polynomial exists with infinitely many prime divisors $p \equiv a \pmod m$.

We have already seen in Lemma 3.12 that a non-constant polynomial with integer coefficients has infinitely many primes divisors. We can combine this with the following theorem to get a more complete understanding of what a Euclidean Proof could entail.

Definition 4.3. For $f \in \mathbb{Z}[x]$, call $P(f)$ *the set of prime divisors of* f*.*

Theorem 4.4 ([12, Theorem 3])**.** *If* $f, g \in \mathbb{Z}[x]$ *are non-constant, then* $P(f) \cap P(g)$ *is infinite.*

This theorem has some significant implications. If we take one of the polynomials to be the mth cyclotomic polynomial, which has finitely many prime divisors $p \mid m$ and infinitely many $p \equiv 1 \pmod m$, this implies that every polynomial must have infinitely many prime divisors $p \equiv 1 \pmod m$. This leads Murty and Thain to give the following definition of a Euclidean Proof in [12]:

Definition 4.5 (Euclidean Proof)**.** A *Euclidean Proof* for the infinitude of primes in the arithmetic progression $a \pmod m$ is the demonstration of the existence of a polynomial $f \in \mathbb{Z}[x]$ such that all prime divisors of f (apart from finitely many) are either $p \equiv 1 \pmod m$ or $p \equiv a \pmod m$.

Call such a polynomial a *Euclidean polynomial*, and we can assume that it is irreducible.

So now we can return to our question, that of when exactly a Euclidean Proof of the infinitude of primes $p \equiv a \pmod m$ exists. In 1912, Schur [13] proved that if $a^2 \equiv 1 \pmod m$ then a Euclidean polynomial exists for $a \pmod m$, which seems like quite a significant step toward answering our question, but infact his statement actually applies in both directions, as proved by Murty in 1988 [14]. Schur and Murty's proofs were later collected into a more accessible paper by Murty and Thain [12].

Theorem 4.6 ([12, Theorem 1])**.** *A Euclidean Proof exists for the arithmetic progression* $a \pmod m$ *if and only if* $a^2 \equiv 1 \pmod m$*.*

Proof. See [12]. □

This exactly provides the answer to our question, showing the case in which a Euclidean Proof exists whilst also telling us that in every other case, one definitely does not exist.

4.2 The Euclidean Bound of Dirichlet's Theorem

Consider the case $m = 24$. Note that the integers less than 24 that are coprime to 24 are $1, 5, 7, 11, 13, 17, 19, 23$, and observe that $1^2 \equiv 1 \pmod{24}$, $5^2 = 25 \equiv 1 \pmod{24}, 7^2 = 49 \equiv 1 \pmod 4, \dots$. This in fact holds for every integer coprime to 24 so by Theorem 4.6 there must be a Euclidean proof for $a \pmod{24}$ for every a coprime to 24. 24 is in fact special in this case, as all integers k that satisfy this condition (that every coprime number has square congruent to 1 modulo k) are exactly 24 and all of it's divisors, i.e. $k = 1, 2, 3, 4, 6, 8, 12$, and 24 [15]. So 24 is the largest integer for which a Euclidean proof exists for all cases of Dirichlet's Theorem 4.1 and is thus the natural bound for which Dirichlet's Theorem has a Euclidean proof (although note that there are cases for $m < 24$ and a coprime to m which do not have Euclidean Proofs).

This case is considered by Bateman and Low in [16] where they give a detailed proof from scratch. We will cover their proof in this section.

Lemma 4.7 ([16, Lemma 1])**.** *If p is a prime greater than 3, then we have*

$$\begin{aligned}
(-1 \mid p) = 1 \textit{ if and only if } p &\equiv 1 \pmod 4,\\
(2 \mid p) = 1 \textit{ if and only if } p &\equiv 1, 7 \pmod 8,\\
(-2 \mid p) = 1 \textit{ if and only if } p &\equiv 1, 3 \pmod 8,\\
(3 \mid p) = 1 \textit{ if and only if } p &\equiv 1, 11 \pmod{12},\\
(-3 \mid p) = 1 \textit{ if and only if } p &\equiv 1 \pmod 6,\\
(6 \mid p) = 1 \textit{ if and only if } p &\equiv 1, 5, 19, 23 \pmod{24},\\
(-6 \mid p) = 1 \textit{ if and only if } p &\equiv 1, 5, 7, 11 \pmod{24}.
\end{aligned}$$

Proof. These assertions follow directly from quadratic reciprocity (Theorem 2.12) and Theorem 2.10:

- $(-1 \mid p) = (-1)^{(p-1)/2} = 1 \iff p \equiv 1 \pmod 4$.
- $(2 \mid p) = (-1)^{(p^2-1)/8} = 1 \iff (p^2 - 1)/8 \equiv 0 \pmod 2 \iff p^2 \equiv 1 \pmod{16}$, and $1^2 = 1, 3^2 = 9, 5^2 = 25 \equiv 9, 7^2 = 49 \equiv 1, 9^2 = 81 \equiv 1, 11^2 = 121 \equiv 9, 13^2 = 169 \equiv 9, 15^2 = 225 \equiv 1 \pmod{16}$ (we only check odd residues since p is odd) so $p \equiv 1, 7, 9, 15 \pmod{16} \iff p \equiv 1, 7 \pmod 8$.
- $(-2 \mid p) = (-1 \mid p) \cdot (2 \mid p)$. Both of these are 1 iff $p \equiv 1 \pmod 4$ and $p \equiv 1, 7 \pmod 8$, thus $p \equiv 1 \pmod 8$. Both are -1 iff $p \equiv 3 \pmod 4$ and $p \equiv 3, 5 \pmod 8$ so $p \equiv 3 \pmod 8$. Thus $p \equiv 1, 3 \pmod 8$.
- $(3 \mid p) = (-1)^{(p-1)/2} \cdot (p \mid 3)$. Both of these are 1 iff $p \equiv 1 \pmod 4$ and $p \equiv 1 \pmod 3$ so $p \equiv 1 \pmod{12}$. Both are -1 iff $p \equiv 1 \pmod 4$ and $p \equiv 2 \pmod 3$ so $p \equiv 11 \pmod{12}$. Thus $p \equiv 1, 11 \pmod{12}$.
- $(-3 \mid p) = (-1 \mid p) \cdot (3 \mid p) = (p \mid 3) = 1 \iff p \equiv 1 \pmod 3 \iff p \equiv 1 \pmod 6$ since p odd.

- $(6 \mid p) = (2 \mid p) \cdot (3 \mid p)$. Both of these are 1 iff $p \equiv 1, 7$ and $p \equiv 1, 11 \pmod{12}$ so $p \equiv 1, 23 \pmod{24}$. Both are -1 iff $p \equiv 3, 5 \pmod 8$ and $p \equiv 5, 7 \pmod{12}$ so $p \equiv 5, 19 \pmod{24}$. Thus $p \equiv 1, 5, 19, 23 \pmod{24}$.
- $(-6 \mid p) = (-1 \mid p) \cdot (6 \mid p)$. Both of these are 1 iff $p \equiv 1 \pmod 4$ and $p \equiv 1, 5, 19, 23 \pmod{24}$ so $p \equiv 1, 5 \pmod{24}$. Both are -1 iff $p \equiv 3 \pmod 4$ and $p \equiv 7, 11, 13, 17 \pmod{24}$ so $p \equiv 7, 11 \pmod{24}$. Thus $p \equiv 1, 5, 7, 11 \pmod{24}$.

□

We now define the following 7 polynomials and note some identities that they satisfy.

$$\begin{aligned}
f_5(x) &= x^4 + 9 = (x^2)^2 + 3^2 = (x^2+3)^2 - 6x^2 = (x^2-3)^2 + 6x^2\\
f_7(x) &= x^4 + 2x^2 + 4 = (x^2+2)^2 - 2x^2 = (x^2+1)^2 + 3 = (x^2-2)^2 + 6x^2\\
f_{11}(x) &= x^4 + 4x^2 + 1 = (x^2+1)^2 + 2x^2 = (x^2+2)^2 - 3 = (x^2-1)^2 + 6x^2\\
f_{13}(x) &= x^4 - x^2 + 1 = (x^2-1)^2 + x^2 = (x^2 - \tfrac{1}{2})^2 + 3(\tfrac{1}{2})^2 = (x^2+1)^2 - 3x^2\\
f_{17}(x) &= x^4 + 1 = (x^2)^2 + 1 = (x^2+1)^2 - 2x^2 = (x^2-1)^2 + 2x^2\\
f_{19}(x) &= x^4 - 2x^2 + 4 = (x^2-2)^2 + 2x^2 = (x^2-1)^2 + 3 = (x^2+2)^2 - 6x^2\\
f_{23}(x) &= x^4 - 4x^2 + 1 = (x^2-1)^2 - 2x^2 = (x^2-2)^2 - 3 = (x^2+1)^2 - 6x^2
\end{aligned}$$

The motivation for use of these polynomials comes from the 24th cyclotomic polynomial, $\Phi_{24}(x) = x^8 - x^4 + 1$. Considering the identities:

$$\begin{aligned}
x^8 - x^4 + 1 &= (x^2 + x^2 + 1)^2 - 2(x^3 + x)^2 = (x^4 - x^2 - 1)^2 + 2(x^3 - x)^2\\
&= (x^4 + 1)^2 - 3(x^2)^2 = (x^4 - \tfrac{1}{2})^2 + 3(\tfrac{1}{2})^2\\
&= (x^4 + 3x^2 + 1)^2 - 6(x^3 + x)^2 = (x^4 - 3x^2 + 1)^2 + 6(x^3 - x)^2
\end{aligned}$$

we see that the field generated by the primitive 24th roots of unity ($\mathbb{Q}[\zeta_{24}]$) contains the quadratic fields generated by the square roots of $-1, \pm 2, \pm 3, \pm 6$. One can check that the zeros of the above polynomials also generate these same quadratic fields.

Lemma 4.8 ([16, Lemma 3])**.** *Suppose a is one of $5, 7, 11, 13, 17, 19, 23$. Then if p is a prime divisor of f_a greater than 3, we have $p \equiv 1, a \pmod{24}$.*

Proof. In each case, this follows from Lemma 4.7 and the identities given for each polynomial above. Bateman and Low present a proof for $a = 11$ in [16]. We will present a proof for $a = 19$ with proofs for the other a following similarly.

Suppose that p is a prime divisor of f_{19} greater than 3. Then $f_{19}(n) \equiv 0 \pmod p$ for some $n \in \mathbb{Z}$. So from the identities above we have that

$$n^4 - 2n^2 + 4 \equiv (n^2-2)^2 + 2n^2 \equiv (n^2-1)^2 + 3 \equiv (n^2+2)^2 - 6n^2 \equiv 0 \pmod p.$$

Each of these equivalences gives us a different identity. The first ($(n^2-2)^2 + 2n^2$) tells us that $(n^2-2)^2 \equiv -2n^2 \pmod p$ so that $-2n^2$ must be a quadratic residue modulo p, and since n^2 is trivially a quadratic residue, in particular we must have that -2 is a quadratic residue, i.e. that $\left(\frac{-2}{p}\right) = 1$. Similarly, the second $((n^2-1)^2+3)$ tells us that $\left(\frac{-3}{p}\right) = 1$, and the third $((n^2+2)^2-6n^2)$ tells us that $\left(\frac{6}{p}\right) = 1$. Combining all this we have that

$$\left(\frac{-2}{p}\right) = \left(\frac{-3}{p}\right) = \left(\frac{6}{p}\right) = 1,$$

so by Lemma 4.7 we must have that $p \equiv 1, 3 \pmod 8$, $p \equiv 1 \pmod 6$, and $p \equiv 1, 5, 19, 23 \pmod{24}$. The only p that satisfies all 3 congruences is when $p \equiv 1, 19 \pmod{24}$ as required. □

The rest of this proof follows in a manner similar to the proof for $3 \pmod 4$. Consider the following polynomials.

$$\begin{aligned}
g_5(x) &= \tfrac{1}{2} f_5(12x+1) = 24(432x^4 + 144x^3 + 18x^2 + x) + 5,\\
g_7(x) &= f_7(6x+1) = 24(54x^4 + 36x^3 + 12x^2 + 2x) + 7,\\
g_{11}(x) &= \tfrac{1}{3} f_{11}(6x+2) = 24(18x^4 + 24x^3 + 14x^2 + 4x) + 11,\\
g_{13}(x) &= f_{13}(12x+2) = 24(864x^4 + 576x^3 + 138x^2 + 14x) + 13,\\
g_{17}(x) &= f_{17}(6x+2) = 24(54x^4 + 72x^3 + 36x^2 + 7x) + 17,\\
g_{19}(x) &= \tfrac{1}{12} f_{19}(12x+4) = 24(72x^4 + 96x^3 + 47x^2 + 10x) + 19,\\
g_{23}(x) &= \tfrac{1}{2} f_{23}(12x+3) = 24(432x^4 + 432x^3 + 150x^2 + 21x) + 23.
\end{aligned}$$

The coefficients of each polynomial do not particularly matter. We just need that the constant term of g_a is a and that the rest of the coefficients are divisible by 24.

Lemma 4.9 ([16, Lemma 4])**.** *Suppose a is one of $5, 7, 11, 13, 17, 19, 23$ and $n \in \mathbb{Z}$. Then $g_a(n)$ has at least one prime divisor $p \equiv a \pmod{24}$.*

Proof. By definition of each g_a, it is clear that $g_a(n) \equiv a \pmod{24}$. It follows from this that $2, 3 \nmid g_a(n)$. Then by Lemma 4.8 and the definition of g_a, it follows that any prime divisor of $g_a(n)$ must be $p \equiv 1, a \pmod{24}$ but since $g_a(n) \equiv a \pmod{24}$, they cannot all be $\equiv 1 \pmod{24}$ and so there must be at least one prime divisor $\equiv a \pmod{24}$. □

Theorem 4.10 ([16, Theorem])**.** *Let a be an integer coprime to 24, then there are infinitely many primes p such that $p \equiv a \pmod{24}$.*

Proof. The case when $a = 1$ is covered by Theorem 3.1, so suppose a is one of $5, 7, 11, 13, 17, 19, 23$. Take $S = \{p_1, p_2, \ldots, p_k\}$ to be some finite set of primes all greater than 24, and call P the product of all primes in S. Now for each a take some $c_a \in \mathbb{N}$ such that

$$g_a(c_a) \not\equiv 0 \pmod a.$$

For example,

$$c_a = \begin{cases} 1, & \text{for } a = 7, 11, 13, 19, \\ 2, & \text{for } a = 5, 7, 23. \end{cases}$$

Now consider $N = g_a(c_a P^{a-1})$. By definition of g_a, c_a, and P we must have that $N > 1$ and so has some prime divisors. By Lemma 4.9 we know that at least one of these must be $\equiv a \pmod{24}$. Now since $a \notin S$ we have $\gcd(a, P) = 1$. Combining this with the fact that $g(0) = a$, we must have that N is not divisible by any of the primes in S. Since each a is prime, Fermat's Little Theorem tells us that $P^{a-1} \equiv 1 \pmod{a}$, and so in particular we have that

$$N = g_a(c_a P^{a-1}) \equiv g_a(c_a) \not\equiv 0 \pmod{a}$$

and so N is not divisible by a. Thus N has a prime divisor congruent to $a \pmod{24}$ which is not equal to a or any of the primes in S. Hence S cannot include all primes greater than 24 which are congruent to $a \pmod{24}$.

Since S can be taken to be any finite set of primes greater than 24, we have the desired result. □

4.3 Sequences

So we now know exactly when a Euclidean proof exists for the arithmetic progression $mk + a$, and that 24 is the largest number such that one exists for all a coprime to m. Now we can construct the (first/second) Euclid-Mullin sequences in $mk + a$. But how can we know what Euclidean polynomials to use for each $mk + a$? For that we can look to the proof that Murty and Thain provided in [12], as well as a further example they gave for the progression $15k + 4$.

In their proof, they introduce the polynomial

$$f(x)^2 = \prod_{\gcd(k,m)=1} (x - (u - \zeta^k)(u - \zeta^{ka})) \tag{26}$$

where u is taken to be a non-zero multiple of m, and m is assumed to be greater than 1 (this is not a problem since the only congruence modulo 1 is 0 which is trivial). The reason for the choice of the u is to ensure that $f(0) = \Phi_m(u) \equiv 1 \pmod{m}$. As we have seen in our discussion around Theorem 3.23, this is not the only choice for such a u, but rather just one of them. The proof then goes on to prove that this polynomial has all prime divisors (except finitely many) congruent to 1 or a modulo m. After concluding their proof, they provide an example that proves a Euclidean polynomial for the progression $15k + 4$. The polynomial used however does not appear to take the form of (26), rather taking the following form:

$$f_{a,m}(x)^2 = \prod_{\gcd(k,m)=1} (x - (\zeta^k + \zeta^{ka})). \tag{27}$$

In particular, they had

$$f_{4,15}(x) = (x - (\zeta + \zeta^4))(x - (\zeta^2 + \zeta^8))(x - (\zeta^7 + \zeta^{13}))(x - (\zeta^{11} + \zeta^{14}))$$

which simplifies to

$$f_{4,15}(x) = x^4 - x^3 + 2x^2 + x + 1.$$[3]

This polynomial can thus be used to find the Euclid-Mullin sequences in $15k+4$ generated by $f_{4,15}(15x+1)$, where this value is taken to ensure that there is always at least one prime divisor 4 (mod 15). The first sequence starts $1999, 97789, 1842764404729, 19, 619, \dots$.

Note that the polynomials used in the previous section for the proof of the infinitude of primes $24k+a$ utilises many polynomials of the form in (27). In particular we have that $f_a(x) = f_{a,24}(x)$ for $a = 5, 7, 11, 17, 19, 23$. This does not work for every congruence however, as $f_{13,24}(x) = x^4 \neq f_{13}$. However, f_{13} does take the form in (26) with $u = 0$.

So to construct the (first/second) Euclid-Mullin sequence in $mk+a$, a Euclidean polynomial is required that has infinitely many prime divisors congruent to 1 or a modulo m, say $f_{a,m}(x)$, and then a specific value of the polynomial is taken (e.g. $f_{a,m}(bx+c)$), to be used at each step to ensure that there is at least one prime divisor of the latter kind for any input. The method then follows in the usual Euclidean way. To find such a polynomial, first try the form in (27) as this form is usually simpler to calculate, and then if that does not work then try different u in the form in (26).

Example 4.11. Consider $f_{6,7}(x) = x^3 + x^2 - 2x - 1$ from (27). The Euclid-Mullin sequences in $7k+6$ can be generated by $f_{6,7}(7x)$. The first sequence for example has the first 4 terms $13, 761669, 937, 310228066732684181095239674203$.

Example 4.12. Recall the proof for the infinitude of primes in $4k+3$ (Theorem (4.2)) which used the polynomial $g(x) = 4x - 1$. The forms (26), (27) give $x - 1$ and x respectively, which can both be used to give this g. The sequence generated by $g(x)$ gives the first 5 terms $3, 11, 131, 17291, 298995971$. One thing of interest is that these are the first 5 terms of both the first and second Euclid-Mullin sequence in $4k+3$ generated by $g(x)$ since whilst at each step there are multiple prime divisors, only one of each is congruent to 3 modulo 4. The 6th term is where the sequences diverge.

Another thing to note is that the polynomial $h(x) = 4x + 3$ would also work in this case since the same congruences required for the proof of Theorem 4.2 hold, and so we could also generate the Euclid-Mullin sequences in $4k+3$ with $h(x)$. These sequences have first 8 terms $7, 31, 67, 19, 179, 197788559, 39120313477930807, 4381368653643102103430885023$, which are the same for both sequences, and the 9th term is where they diverge.

In fact these sequences could be generated by $f(x) = 4x + b$ for any $b \equiv 3 \pmod 4$ which emphasises the fact that these sequences are not unique, and in fact can be generated in any number of ways.

[3] Note that [12] appears to have a printing error as this polynomial is claimed to simplify to $x^4 - x^3 + 2x^2 + 1$. Otherwise, their argument is sound.

Similarly to the example above, polynomials of the form (26) will work for any applicable u and so similarly any sequence in $mk+a$ is not unique, and will change depending on the Euclidean polynomial used to generate it.

4.4 Questions

Many of the same questions as discussed in section 3.3 apply here with similar discussions so we will only consider a couple of new questions.

Question 8. Does the union of the Euclid-Mullin sequences in $mk+a$ ranging over the possible generating polynomials contain all primes $a \pmod m$?

This seems likely for both cases, since there are a potentially infinite number of generating polynomials of the form (26) for each arithmetic progression. However these will get large very quickly and so this is probably a hard question to answer.

Question 9. Is there any combination of arithmetic progression $mk+a$ and generating polynomial f such that the first and second Euclid-Mullin sequences in $mk+a$ generated by f are identical?

Example 4.12 showed that in some cases the first and second sequences would match for the first few terms which motivates us to ask this question. This would require that at each step of the method, the value of $f(x)$ has only one prime divisor that matches the congruence $a \pmod m$ and so the ambiguity inherent in Euclid's Method is eliminated. This seems unlikely however as primes are unpredictable.

A Appendix

A.1 The Euclid-Mullin Sequences

As of the time of writing, the first 51 terms of the First Euclid-Mullin Sequence (EM1) are known. These are as follows (A000945 on OEIS [17])

i	p_i	i	p_i	i	p_i
1	2	18	37	35	89
2	3	19	1741	36	19
3	7	20	1313797957	37	577
4	43	21	887	38	223
5	13	22	71	39	139703
6	53	23	7127	40	457
7	5	24	109	41	9649
8	6221671	25	23	42	61
9	38709183810571	26	87	43	4357
10	139	27	159227	44	(see below)
11	2801	28	(see below)	45	107
12	11	29	103	46	127
13	17	30	1079990819	47	3313
14	5471	31	9539	48	(see below)
15	52662739	32	3143065813	49	59
16	23003	33	29	50	31
17	30693651606209	34	3847	51	211

i	p_i
28	643679794963466223081509857
44	87991098722552272708281251793312351581099392851768893748012603709343
48	227432689108589532754984915075774848386671439568260420754414940780761245893

In comparison, only the first 14 terms of the Second Euclid-Mullin Sequence (EM2) are known. These are as follows (A000946 on OEIS [17])

i	q_i	i	q_i
1	2	6	50207
2	3	7	340999
3	7	8	2365347734339
4	43	9	4680225641471129
5	139	10	1368845206580129

i	q_i
11	889340324577880670089824574922371
12	20766142440959799312827873190033784610984957267051218394040721
13	348654613352373829454902145370501700873487314509264314920485482161426646699 8637603378972254923344607825545244648001799
14	264025908176651231151241967831104868143619302344557880597101834841512474609 60172672371287819122033451

References

[1] A. A. Mullin, (1963), *"Recursive Function Theory. (A modern look at a Euclidean idea.)"*, Bulletin AMS 69, p. 737. https://doi.org/10.1090/S0002-9904-1963-11017-4

[2] D. Shanks, (1991), *"Euclid's Primes"*, Bulletin of the Institute of Combinatorics and Its Applications 1, p. 33-36.

[3] T. Naur, (1984), *"Mullin's Sequence of Primes is not Monotonic"*, Proceedings of the American Mathematical Society 90, p. 43-44. https://doi.org/10.2307/2044665

[4] C. D. Cox, A. J. Van Der Poorten, (1968), *"On a sequence of prime numbers."* Journal of the Australian Mathematical Society 8, p. 571-574. https://doi.org/10.1017/S1446788700006236

[5] A. R. Booker, (2012), *"On Mullin's second sequence of primes"*. https://arxiv.org/abs/1107.3318.

[6] P. Pollack, E. Treviño, (2014), *"The primes that Euclid forgot"*, American Mathematical Monthly 121 (5), p. 433–437. https://doi.org/10.4169/amer.math.monthly.121.05.433

[7] G. Everest, T. Ward, (2005), *"An Introduction to Number Theory"*, Graduate Texts in Mathematics.

[8] Author's GitHub Repository: *"Euclidean Prime Generators"*. https://github.com/SHeywood8/euclidean-prime-generators.

[9] P. Pollack, (2009), *"Not Always Buried Deep"*, p. 23-25. https://pollack.uga.edu/NABDofficial.pdf

[10] J. Lee., A. Kumar, (2012), *"Theory of Numbers"* Lecture Notes, Section 13. Obtained from MIT OpenCourseWare: https://ocw.mit.edu/courses/18-781-theory-of-numbers-spring-2012/.

[11] Z. Wang, (2017), *"Elementary Proof of Dirichlet Theorem"*. https://math.uchicago.edu/~may/REU2017/REUPapers/WangZijian.pdf.

[12] M. R. Murty, N. Thain, (2006), *"Primes in Certain Arithmetic Progressions."* Functiones et Approximatio Commentarii Mathematici 35, p. 249-359. https://doi.org/10.7169/facm/1229442627.

[13] I. Schur, (1912), *"Über die Existenz unendlich vieler Primzahlen in einigen speziellen arithmetischen Progressionen"*, Sitzungber. Berliner Math. Ges. 11, p. 40–50.

[14] M. R. Murty, (1988), *"Primes in Certain Arithmetic Progressions."* J Madras University, p. 161-169.

[15] S. K. Chebolu, (2012), *“What Is Special about the Divisors of 24?”* Mathematics Magazine, 85(5), p.366-372. https://doi.org/10.4169/math.mag.85.5.366.

[16] P. T. Bateman, M. E. Low, (1965), *“Prime Numbers in Arithmetic Progressions with Difference 24.”* The American Mathematical Monthly, 72(2), p. 139-143. https://doi.org/10.2307/2310975.

[17] *The On-Line Encyclopedia of Integer Sequences (OEIS)*. https://oeis.org/.