University of BRISTOL

Euclidean Prime Generators

Ella Womack

Supervised by Andrew Booker
Level H/6
20 Credit Points

February 18, 2025

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 e. womack.

Date 18/02/2025

1 Introduction

Prime numbers have always fascinated pure mathematicians and the intrigue is well deserved. However, despite the extensive research committed in theory development, still the question of how the primes are generated is unanswered. It is well known that the number of primes extends to an infinite amount, so how can we find each of them in turn?

Euclid's first proof of the infinitude of primes (300BC) led to the development of the Euclid-Mullin sequences. Mullin posed the question over half a century ago, officially conjectured by Shanks [10]: does the first Euclid-Mullin sequence generate every prime? The question remains wide open to this day. However, despite his original question remaining unanswered, much mathematical research has been conducted to explore certain variants of the famous sequence, leading eventually to the discovery of variants which do indeed generate every prime, and moreover variants which can exactly generate the kth smallest prime number.

In this paper, we will examine firstly some short proofs of the infinitude of primes due to Euclid and Stieltjes; following this, we will introduce a discussion of the Euclid-Mullin sequences, along with some results and open conjectures surrounding them. Finally we will link our introductory material to an exposition of the links between the Euclidean Prime Generators and ring theory, along with further variants of the Euclid-Mullin sequences which we will show to generate every prime number.

We note here that some basic results from number theory and ring theory will be stated without proof throughout the paper. Readers are encouraged to convince themselves of these results, but a large amount of prior knowledge of the subject is not necessary.

2 Prior results from number theory

In this section, we provide a number of results without proof for the benefit and understanding of the reader. These results, although stemming predominantly from number theory, come from a range of areas in mathematics, and will be referenced throughout the paper. The following definitions are taken from [1].

Definition 2.1. For a prime q, and $a \in \mathbb{Z}$, the Legendre symbol $\left(\frac{p}{q}\right)$ is defined as follows:

$$\left(\frac{a}{q}\right) = \begin{cases} 0 & \text{if } a \equiv 0 \text{ modulo } q \ , \\ 1 & \text{if } a \equiv x^2 \text{ modulo } q, \\ -1 & \text{otherwise.} \end{cases} \tag{1}$$

Definition 2.2. A **squarefree** integer is an integer which is indivisible by any square number other than 1.

Definition 2.3. Recall that $m \equiv n$ modulo x if $m - n = kx$ for some $k \in \mathbb{Z}$. We define the **residue class** of m modulo n to be

$$[m] = \{x \in \mathbb{Z} : x \equiv m \pmod{n}\}$$

Lemma 2.4. *Each natural number $n \in \mathbb{N}$ such that $n \geq 2$ has at least one prime divisor.*

Theorem 2.5 (Multiplicativity of the Legendre Symbol [1]). *The Legendre symbol satisfies the multiplicative rule:*

$$\left(\frac{m}{p}\right)\left(\frac{n}{p}\right) = \left(\frac{mn}{p}\right).$$

Theorem 2.6 (Fermat's Little Theorem). *For a prime number p, and integer a such that* $\operatorname{hcf}(a, p) = 1$, *then $a^{p-1} \equiv 1 \pmod{p}$. Moreover, for any $a \in \mathbb{Z}$, we have $a^p \equiv a \pmod{p}$.*

3 The infinitude of primes

Euclid first published his famous proof of the infinitude of primes in *Elements* over 2300 years ago, around 300BC. The theorem was originally described "there are more primes than found in any finite list of primes", but has since been adapted to read more concisely as "there are an infinite number of primes".

The original wording of the famous theorem is indeed evident in the structure of Euclid's original proof - upon reading the adaptation of the statement, we may expect to prove by contradiction, but instead we see that Euclid provides an algorithm for finding new primes. This method led to the development in research of the Euclid-Mullin sequences, the main focus of this paper.

Following Euclid, over 200 varying proofs of the infinitude of primes were published [5]. Methods for these subsequent proofs made use of (but were not limited to) the following topics: algebraic number theory, Euler's formula for the Riemann zeta function, and some topological proof of the Euclid's theorem. The list extends, but for this paper we will simply focus on the proofs of Euclid and Stieltjes.

Theorem 3.1. *There are infinitely many prime numbers.*

3.1 Euclid's proof (from *Elements*, C.300BC)

We now begin our first proof, which many regard as one of the more succinct proofs for Theorem 3.1. We follow Pollack and Treviño's proof here from [8] but highlight that they are in fact using Euclid's original work.

Proof. Suppose $\{p_1, \ldots, p_k\}$ is a finite list of distinct primes, with $k \in \mathbb{N}$. Let $P = p_1 p_2 \cdots p_k = \prod_{i=1}^{k} p_i$.

Consider $P + 1 \in \mathbb{N}$. Note that each p_i must divide P, so $P \equiv 0 (\text{mod } p_i)$ for each $i = 1, \ldots, k$.

Since $P + 1 \equiv 1 \pmod{p_i}$, for $i \in \{1, \ldots, k\}$, we have that for all i as above, p_i cannot divide $P + 1$.

However, we have that $P + 1 > 1$ thus by Lemma 2.4, we must have that $P + 1$ has at least one prime divisor. So there exists a prime p which is distinct from our original list $\{p_1, \ldots, p_k\}$. Thus we have discovered a new prime, separate from our finite list.

□

Note: We are somewhat paraphrasing above, since we have that the direct translation into English of the theorem was "the prime numbers are more than any given set of prime numbers". Upon reading Euclid's original proof translated from Greek, we see that his argument is structured around using three lengths A, B and C to prove his result. This is due to the fact that, instead of thinking of the integers as abstract objects as we do today, Euclid would have thought of them as physical lengths, and so his multiplication of the primes had to be defined in terms of areas and volumes.

3.2 Stieltjes's Proof, 1890

Secondly, we take a look at Stieltjes's proof for the infinitude of primes [7]. This proof is focused on decomposition and factors but is still a concise explanation of Theorem 3.1.

Proof. As in Euclid's proof, suppose $\{p_1, \ldots, p_k\}$ is a finite list of distinct primes, with $k \in \mathbb{N}$. Let $P = p_1 p_2 \cdots p_k = \prod_{i=1}^{k} p_i$.

Let $P = AB$ be any decomposition of P with positive factors A, B. So we have $A, B \in \mathbb{N}$.

Suppose there exists a prime p such that $p = p_i$ for some $i \in \{1, \ldots, k\}$. Then we have

$$p \mid AB \Rightarrow (p \mid A \text{ or } p \mid B).$$

Note here that

$$(p \mid A \text{ and } p \mid B) \Rightarrow p^2 \mid AB$$

which is contradictory since all of the primes p_i are distinct. Thus it must be the case that p divides exactly one of A or B. Thus for all $i \in \{1, \ldots, k\}$, we have that since $p \mid A$ or $p \mid B$ but we cannot have $p \mid A$ and $p \mid B$, it follows that $p \nmid A + B$ and so $p_i \nmid A + B$ for all $i \in \{1, \ldots, k\}$.

In other words, none of the primes p_i in our original list can divide the natural number $A + B$.

However, as $A, B \in \mathbb{N}$, we have $A + B \geq 2$, then by Lemma 2.4 it must be that $A + B$ has some prime divisor. So we have discovered a new prime, not included in our original list. □

We see above two of the most important proofs of the infinitude of primes, crucial to our later discussion of the Euclid-Mullin sequences.

4 A discussion of the Euclid-Mullin sequences

In this section, we will use the method demonstrated in Euclid's proof of Theorem 3.1 to generate the Euclid-Mullin sequences. Further, we will note the differences between the two sequences along with two results due to Cox and Van der Poorten and Booker.

4.1 Generating the Euclid-Mullin sequences

Recall our finite set of distinct primes $\{p_1, \ldots, p_k\}$, with the previously defined product of said primes $P = p_1 p_2 \cdots p_k = \prod_{i=1}^{k} p_i$. Then by Euclid's proof we can find some new prime p such that p divides $P+1$.
If we choose our new prime p to be the smallest prime which divides $P+1$, we generate the first Euclid-Mullin sequence (A000945 from OEIS) [8]:

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, \ldots$$

It is conjectured by Shanks [10] that the first Euclid-Mullin sequence eventually generates every prime number, however little has been proven rigorously to confirm this conjecture, nor to provide counterexample.
If we take p to be the largest prime dividing $P+1$ we generate the second Euclid-Mullin sequence, of which only fourteen values have been found; see below for the first nine (A000946 from OEIS):

$$2, 3, 7, 43, 139, 50207, 340999, 2365347734339, 4680225641471129, \ldots$$

To demonstrate the difference in results of the separate Euclid-Mullin sequences, see the tables below for the first six primes generated in each:

First Euclid-Mullin Sequence		
Set of primes	$N = P + 1$	New prime p_{k+1}
$\varnothing$	2	2
{2}	3	3
{2,3}	7	7
{2,3,7}	43	43
{2,3,7,43}	$1807 = 13 \cdot 139$	13
{2,3,7,43,13}	$23479 = 53 \cdot 443$	53

And now the second sequence:

Second Euclid-Mullin Sequence		
Set of primes	$N = P + 1$	New prime p_{k+1}
$\varnothing$	2	2
{2}	3	3
{2,3}	7	7
{2,3,7}	43	43
{2,3,7,43}	$1807 = 13 \cdot 139$	139
{2,3,7,43,139}	$251035 = 5{\cdot}50207$	50207

We see from these tables the difference in results using the distinct methods of each of the Euclid-Mullin sequences. Moreover, observe that it immediately seems that certain primes are omitted from the second Euclid-Mullin sequence; in fact, this observation extends to an infinite amount of prime numbers which are not generated by the second Euclid-Mullin sequence. We discuss these results in the following sections.

Remark: The second Euclid-Mullin sequence is in fact *not* monotonic increasing [6]; if we continue our method above to find the next terms of the sequence we generate:

$$\begin{aligned}
a_7 &= 340999 \\
a_8 &= 2365347734339 \\
a_9 &= 4680225641471129 \\
a_{10} &= 1368845206580129 \\
a_{11} &= 889340324577880670089824574922371 \\
a_{12} &= 20766142440959799312827873190033784610984957267051218394040721
\end{aligned}$$

thus we see that in fact $a_{10} < a_9$, demonstrating that the sequence is not monotonic. This observation was published by Naur, based on computations of the values in the second Euclid-Mullin sequence along with their factorisations. These factorisations were carried out by the likes of Morrison and Brillhart, Pollard and Brillhart, and Lehmer and Selfridge. [6].

Following this observation, how can we be sure that the second Euclid-Mullin sequence omits any, if not *infinitely* many primes?

4.2 The Cox Van Der Poorten investigation

Cox and Van Der Poorten concentrated their investigation on the second Euclid-Mullin sequence, confirming the following result:

Theorem 4.1. *The primes 5, 11, 13, 17, 19, 23, 29, 31, 37, 41 and 47 do not occur in the second Euclid-Mullin sequence.*

We talk briefly through their method here. Their exposition in [4] obtains sufficent conditions in determining whether a given prime does not occur in the second Euclid-Mullin sequence.

Beginning by defining $\{q_i\} = \{2, 3, 5, 7, \ldots\}$ as the sequence of all primes in monotonic increasing order and $\{p_i\} = \{2, 3, 7, 43, 139, \ldots\}$ as the second Euclid-Mullin sequence, we firstly note that for $r > 1$ then q_r occurs in $\{p_i\}$ if and only if for some positive integer k,

$$p_1 \cdots p_k + 1 = q_1^{k_1} \cdots q_r^{k_r}$$

for $k_1, \ldots, k_{r-1} \geq 0$ and $k_r > 0$. This is due to the construction of $\{p_i\}$ as defined in Section 4.1.

Following this we have that:

$$p_1 \cdots p_k + 1 = q_1^{k_1} \cdots q_r^{k_r} \equiv 1 \pmod{p_i}.$$

Thus by distinctness of the p_i,

$$q_1^{k_1} \cdots q_r^{k_r} \not\equiv 1 \pmod{p_i^2}.$$

These two observations lead to the useful results:

- When q_i is one of the primes in $\{p_i\}$ then $k_i = 0$. Thus for $r > 1$ we indeed have $q_1^{k_1} = 2^{k_1} = 2^0 = 1$ and thus $q_1^{k_1} \cdots q_r^{k_r} \equiv 1 \pmod{2}$.
- Since $q_1^{k_1} \cdots q_r^{k_r} \not\equiv 1 \pmod{p_i^2}$ then $q_2^{k_2} \cdots q_r^{k_r} \not\equiv 1 \pmod{4}$.

The above results can be used immediately to show that 5 does not occur in $\{p_i\}$:

Example (Nonoccurrence of 5 in $\{p_i\}$)**.** For $\{q_i\} = \{2, 3, 5, 7, \ldots\}$ we have $5 = q_3$. Suppose 5 occurs in $\{p_i\}$. So for some $k > 0$,

$$p_1 \cdots p_k + 1 \equiv q_1^{k_1} \cdots q_r^{k_r} \equiv 2^{k_1} \cdot 3^{k_2} \cdot 5^{k_3}$$

with $k_1, k_2 \geq 0, k_3 > 0$, since $q_3 = 5$ must be the largest prime factor of $p_1 \cdots p_k + 1$. We then note that both 2 and 3 occur in $\{p_i\}$. Consider the prime product $p_1 \cdots p_k$. Note that in $p_1 \cdots p_k$, we have $p_1 = 2$, and p_j is odd for all $2 \leq j \leq k$. Thus it must be that

$$p_1 \cdots p_k \equiv 2 \pmod 4$$

since clearly $2^2 = 4 \nmid p_1 \cdots p_k$ but $p_1 \cdots p_k$ is even. Then immediately we have $p_1 \cdots p_k + 1 \equiv 2 + 1 \equiv 3 \pmod 4$. Now note by the previous two observations and the fact that both 2 and 3 occur in $\{p_i\}$, it must be that $k_1 = k_2 = 0$ and so we must have that

$$q_1^{k_1} \cdots q_r^{k_r} = 2^0 \cdot 3^0 \cdot 5^{k_3}.$$

Then since $5 \equiv 1 \pmod 4$ we have that $5^{k_3} \equiv 1^{k_3} \equiv 1 \pmod 4$ for any $k_3 > 0$. Thus $q_1^{k_1} \cdots q_r^{k_r} \equiv 1 \pmod 4$ which is of course contradictory following our earlier argument that $p_1 \cdots p_k + 1 \equiv q_1^{k_1} \cdots q_r^{k_r} \equiv 3 \pmod 4$. So 5 is nonoccurring in the second Euclid-Mullin sequence.

Following the noted observations, Cox and Van der Poorten then claimed that by multiplicativity and counting that:

- The natural number $q_2^{k_2} \cdots q_r^{k_r}$ contains an odd number of prime divisors congruent to $-1 \pmod 4$.
- The same natural number $q_2^{k_2} \cdots q_r^{k_r}$ contains an even number of prime divisors which are quadratic non-residues of p_i since indeed $q_2^{k_2} \cdots q_r^{k_r}$ is a quadratic residue modulo p_i.

Hence the below congruences must be solvable for non-negative integers $k_2, \ldots, k_r$ with $k_i = 0$ when q_i occurs in the second Euclid-Mullin sequence:

$$a_{1j} = 1 \text{ if } q_j \equiv -1 \pmod 4;$$

$$a_{1j} = 0 \text{ if } q_j \equiv 1 \pmod 4.$$

$$a_{12}k_2 + \ldots + a_{1r}k_r \equiv 1 \pmod 2;$$

$$a_{i2}k_2 + \ldots + a_{ir}k_r \equiv 0 \pmod 2.$$

$$2a_{ij} = 1 - \left(\frac{q_j}{p_i}\right), i = 2, 3, \ldots, k; j = 2, 3, \ldots, k.$$

Notes: We can equally take a_{ij} to be such that $\left(\frac{q_j}{p_i}\right) = (-1)^{a_{ij}}$. Upon recalling that $p_1 = 2$, then it follows that $q_j \equiv (-1)^{a_{1j}} \pmod 4$.

We have $a_{i2}k_2 + \ldots + a_{ir}k_r \equiv 0 \pmod 2$ for the following reasoning:

$$\begin{aligned} 1 = \left(\frac{p_1 \cdots p_k + 1}{p_i}\right) &= \left(\frac{q_1^{k_1} \cdots q_r^{k_r}}{p_i}\right) \\ &= \left(\frac{q_2}{p_i}\right)^{k_2} \cdots \left(\frac{q_r}{p_i}\right)^{k_r} \\ &= (-1)^{a_{i2}k_2} \cdots (-1)^{a_{ir}k_r} \\ &= (-1)^{a_{i2}k_2 + \ldots + a_{ir}k_r} \end{aligned}$$

So indeed we must have that $a_{i2}k_2 + \ldots + a_{ir}k_r \equiv 0 \pmod 2$.

Moreover we have $a_{12}k_2 + \ldots + a_{1r}k_r \equiv 1 \pmod 2$ for the following reasoning: Recall $q_j \equiv (-1)^{a_{1j}} \pmod 4$. So the following congruence is satisfied:

$$\begin{aligned} 1 + p_1 \cdots p_k &= q_2^{k_2} \cdots q_r^{k_r} \\ &\equiv ((-1)^{a_{12}})^{k_2} \cdots ((-1)^{a_{1r}})^{k_r} \pmod 4 \\ &\equiv (-1)^{a_{12}k_2 \cdots a_{1r}k_r} \pmod 4. \end{aligned}$$

Then simply note we have $1 + p_1 \cdots p_k \equiv -1 \pmod 4$ from the previous argument. Thus it must be that $(-1)^{a_{12}k_2 \cdots a_{1r}k_r} \equiv -1 \pmod 4$. Thus $a_{12}k_2 \cdots a_{1r}k_r \equiv 1 \pmod 2$ as required.

With these conditions, the following theorem was proved:

Theorem 4.2. *If for some k the congruences above are unsatisfied then:*

1. *The prime q_r does not occur in the second Euclid-Mullin sequence;*
2. *Nor do the primes q_j for $j < r$ unless q_j is already one of the values in the sequence.*

Note that in the second clause of the above result, the word ‘already’ is used merely to rule out if a prime is known to be occurring in the sequence. This result was used to provide proof for Theorem 4.1. Full proofs for Theorem 4.1 and Theorem 4.2 can be found at [4].

Cox and Van der Poorten also emphasised the likelihood that an infinite amount of primes fail to occur in the second Euclid-Mullin sequences. In fact their observations led to a proof by Booker [3] of this very conjecture.

4.3 Booker’s theorem

Theorem 4.3 (Booker)**.** *The second Euclid-Mullin sequence omits infinitely many primes.*

Booker’s proof of Theorem 4.2 was monumental for the long discussion of the Euclid-Mullin sequences. We omit the proof from this paper due to the complexity of some number theoretic results, however full expositions can be found at [3],[8] .

4.4 The remaining question

So we have discussed how the second Euclid-Mullin sequence in fact omits infinitely many primes and thus cannot possibly generate every prime number. However, mathematicians still contemplate the possible success of the first Euclid-Mullin sequence to this day; could it be that this sequence generates every prime number? How could one go about investigating this result further?

We follow this section with our final discussions in this paper, where we contemplate generalisations of the Euclid-mullin sequences and their strong links to abstract ring theory. We will then conclude by discussing another variant which we show succinctly to generate every prime.

5 Links to ring theory

We now consider a generalisation of Euclid’s proof of Theorem 3.1. Following this, we discuss variants of the Euclid-Mullin sequences and, using ring theoretic results, we will prove that one particular generalisation indeed generates every prime.

5.1 Preliminaries from ring and group theory

Let us state a selection of results from ring theory which we will refer to throughout this section:

Definition 5.1. We define $\mathbb{F}_p$ to be the field consisting of exactly p elements for prime number $p \in \mathbb{N}$, with $\mathbb{F}_p^\times$ defined as the units in the field $\mathbb{F}_p$. Note that as $\mathbb{F}_p$ is a field, its units are exactly defined as $\mathbb{F}_p^\times = \mathbb{F}_p \setminus \{0\}$, a standard result in ring theory.

Definition 5.2. Let G be a group, and let $x \in G$ be any element of said group G. The left coset of a subgroup $H \leq G$ is defined as follows:

$$xH = \{xh : h \in H\}.$$

Theorem 5.3 (Theorem 7.37 of [1])**.** *For prime number p, the multiplicative group of integers modulo p, $((\mathbb{Z}/p\mathbb{Z})^\times, \times)$, is cyclic. (Note that $((\mathbb{Z}/p\mathbb{Z})^\times$ is a field).*

5.2 A generalisation of Euclid's proof

For the remainder of section 5, we will follow the exposition from [2]. Consider some finite set of primes $\{p_1, \ldots, p_k\}$, and take as usual $P = p_1 \cdots p_k$, the product of all primes in the list above. Our generalisation states that for any subset $I \subseteq \{1, \ldots, k\}$, if we define

$$N_I = \prod_{i \in I} p_i + \prod_{i \in \{1,\ldots,k\} \setminus I} p_i$$

then the number N_I is coprime to P. In other words, the highest common factor of P and N_I is 1, and moreover, N_I has at least one prime factor by Lemma 2.4 as $N_I \in \mathbb{N}$.

Note that the above generalisation is actually a construction in Steiltjes' proof; indeed, we decompose P into two positive factors, so $P = AB$ where $A = \prod_{i \in I} p_i$ and $B = \prod_{i \in \{1,\ldots,k\} \setminus I} p_i$. Upon supposing there exists a prime divisor p of AB such that $p \in \{p_1, \ldots, p_k\}$ then we must have via following Steiltjes' proof that p divides exactly one of A or B and so for all $p \in \{p_1, \ldots, p_k\}$, we have $p \nmid A + B = N_I$ as previously defined. So there must exist a new prime which is a divisor of N_I, not included in the original list of primes.

We include two examples to illustrate the above idea.

Example. Consider a set of primes $A = \{2, 3, 5, 7, 11\}$ so the product

$$P_A = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310.$$

Take $I = \{1, 2, 3\} \subseteq \{1, 2, 3, 4, 5\}$. So then

$$N_I = \prod_{i \in \{1,2,3\}} p_i + \prod_{i \in \{4,5\}} p_i = 2 \cdot 3 \cdot 5 + 7 \cdot 11 = 107.$$

Then in fact, 107 is prime itself, and so indeed it is coprime to P_A.

Example. Consider a set of primes $B = \{2, 7, 13, 17\}$ so the product

$$P_B = 2 \cdot 7 \cdot 13 \cdot 17 = 3094.$$

Take $J = \{1, 2\} \subseteq \{1, 2, 3, 4\}$. So then

$$N_J = \prod_{j \in \{1,2\}} p_j + \prod_{j \in \{3,4\}} p_j = 2 \cdot 7 + 13 \cdot 17 = 235.$$

Then indeed $N_J = 235$ is coprime to P_B, even if N_J is not prime itself.

We observe here that Euclid's construction of his proof of the infinitude of primes (Section 3.1) is exactly the case above when we set $I = \varnothing$.

5.2.1 Revisiting Euclid's proof

Proof. (From [2]). Taking $\{p_1, \ldots, p_k\}$ as our finite list of primes, and taking $P = p_1 \cdots p_k$, the product of all primes in the list. If we take $I = \varnothing$ in the method above we obtain

$$\begin{aligned} N_I &= \prod_{i \in \varnothing} p_i + \prod_{i \in \{1,\ldots,k\}} p_i && (2) \\ &= 1 + P && (3) \end{aligned}$$

under the convention that the empty product is equal to 1; thus $N_I = P + 1$ has at least one prime factor and is indeed coprime to P if we follow the remainder of Euclid's proof.

□

We call the initially defined sequence in this section the generalised Euclid sequence with seed $\{p_1, \ldots, p_k\}$. We proceed by showing that this construction is provably general enough to generate every prime.

5.3 Generating every prime

Theorem 5.4. *For any finite set X of primes, there exists a generalised Euclid sequence with seed X containing every prime.*

In this section, we will state, prove and discuss a number of results to provide stepping stones which will lead to a construction of our main proof of Theorem 5.4.

For the following proofs, let us define $S_q \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ to be the set of residue classes generated by squarefree, positive integers (see Definitions 2.2, 2.3). In other words,

$$S_q = \left\{ d + q\mathbb{Z} : q \in \mathbb{N}, d \mid \prod_{p<q} p \right\}.$$

Remark: What does S_q look like? It contains all of the integers of the form $d + kq$ for $k \in \mathbb{Z}$. Since the product $\prod_{p<q} p$ is a product of every such prime smaller than q it is easy to see that as q becomes large, so does this product. Thus the number of divisors d of this product will also increase rapidly:

Define $\pi(n)$ to be the function as follows:

$$\pi(x) = \sum_{p \leq x, p \text{ prime}} 1$$

and define the function $\tau(n)$ to be the number of divisors $d \in \mathbb{N}$ of n. Note that the τ function is indeed multiplicative. In other words,

$$\tau(mn) = \tau(m)\tau(n).$$

Then we have that $\tau(\prod_{p<q} p) = 2^{\pi(q-1)}$ noting that each prime has exactly two divisors. So this function is *almost* exponential, and is rapidly increasing. This suggests that S_q should be as large as possible and in fact we will later show for $q > 7$ that $S_q = (\mathbb{Z}/q\mathbb{Z})^\times$.

So the main aspect to note before we embark on constructing the main proof is that S_q is large. It is important to make this observation because, again considering our generalisation stated previously, if for example q is the smallest prime not yet generated in $\{p_1, \ldots, p_k\}$ then there is a good chance that q will divide $d + \frac{n}{d}$ for $\frac{d}{n} = p_1 \ldots p_k = P$. We will see later why this is relevant.

We now state two additional number-theoretic results to aid our understanding of our main proof.

Lemma 5.5. *For any prime $q \in \mathbb{N}$, we have that $|S_q| > \frac{1}{2}(q-1)$.*

This lemma ties in nicely with our remark above; as our prime q increases, it is clear that $|S_q|$ will increase alongside it. The proof can be found at [9].

Lemma 5.6. *Let q be an odd prime. Let $a \in (\mathbb{Z}/q\mathbb{Z})^\times$. Then the following hold:*

1. *Suppose $q \neq 5$, or $q = 5$ and $a \neq 3 + 5\mathbb{Z}$ then there exists some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that*

$$\left(\frac{x + a/x}{q}\right) \neq 1.$$

2. *Suppose $q \notin \{7, 13\}$. Then there exists some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that*

$$\left(\frac{x^6 + a}{q}\right) \neq 1.$$

A full proof of Lemma 5.6 can be found at [2].

Finally, we approach the main result required to complete our proof of Theorem 5.4. Once we have proved the following lemma, the main theorem simply follows from induction of this result.

Lemma 5.7. *For a finite set of primes $X = \{p_1, \dots, p_k\}$, if q is the smallest prime not contained in X then there exists a generalised Euclid sequence with seed X which contains q.*

Proof. Let $X = \{p_1, \dots, p_k\}$, the finite set of primes defined above, and set $P = p_1 \cdots p_k \in \mathbb{N}$.

Firstly we consider the case where q is an even prime not contained in the set X, so $q = 2$. Then we will have P is odd and so $P + 1$ is even so we simply choose our new prime $q = 2$ since then q will indeed divide the even natural number $P + 1$. Thus without loss of generality we assume for the remainder of the proof that our missing prime q is odd.

For $d \in \mathbb{N}$, let

$$S = \{d + q\mathbb{Z} : d \mid P\}$$

and as above let

$$S_q = \left\{ d + q\mathbb{Z} : d \mid \prod_{p<q} p \right\}$$

.

So clearly, as $d \mid \prod_{p<q} p$ then this implies $d \mid \prod_{p \in X} p$ thus $d \mid P$ and so $S_q \subseteq S$.

We consider two cases below:

Cases for S :

Case 1: $S = (\mathbb{Z}/q\mathbb{Z})^\times$.

Recall from our earlier section covering preliminaries from number theory the Legendre symbol (covered in Definition 2.1). Then the theorem of quadratic reciprocity implies that indeed

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right) \cdot (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

and for $x \in \mathbb{Z}/q\mathbb{Z}$, the quadratic residues of $\mathbb{Z}/q\mathbb{Z}$ are precisely the values of $x^2(\text{mod } q) \in \mathbb{Z}/q\mathbb{Z}$ for each x such that $q \nmid x$.

Let us first consider the case of $\left(\frac{-P}{q}\right) = 1$ thus $-P \equiv x^2(\text{mod } q)$ for some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$. Thus, taking d as this value, we have:

$$d^2 \equiv -P \iff d^2 + P \equiv 0 \iff d + \frac{P}{d} \equiv 0(\text{mod } q)$$

so ineed there exists some divisor of P, d, such that $d + \frac{P}{d} \equiv 0(\text{mod } q)$. Then we can choose q as our next term. This is since q is clearly a divisor of $d + \frac{P}{d}$ by the above implications, and since this expression involves the product P, the requirements of the sequence are fulfilled and so we can choose $p_{k+1} = q$ to extend our sequence.

Now we consider the case of $\left(\frac{-P}{q}\right) = -1$. Thus $-P$ is not a quadratic residue modulo q. In other words, there does not exist any $x \in \mathbb{Z}/q\mathbb{Z}$ such that $-P \equiv x^2(\mod q)$. Now we notice this implies there is no divisor d of P such that $d + \frac{P}{d} \equiv 0(\text{mod } q)$, since

$$d + \frac{P}{d} \equiv 0(\text{mod } q) \iff d^2 + P \equiv 0(\text{mod } q) \iff -P \equiv d^2(\text{mod } q)$$

which is clearly contradictory by the case requirements.

We highlight here that these choices must be in keeping with the conditions of Lemma 5.6. In other words we must have that $q \neq 5$ or $P \not\equiv 3(\text{mod } 5)$. Note that although q does not divide $d + \frac{P}{d}$ there does exist some other prime, say p', which indeed is a divisor of $d + \frac{P}{d}$; this is due to Lemma 2.4 and the fact that $d + \frac{P}{d}$ is a natural number. Now recall that the Legendre symbol satisfies the multiplicative rule (stated in section 2). Thus this prime p' can indeed be chosen specifically to satisfy the condition that $\left(\frac{p'}{q}\right) = -1$. This is due to the fact that by Lemma 5.6, there exists some d such that

$$\left(\frac{d + \frac{P}{d}}{q}\right) = -1$$

since clearly we have $\left(\frac{d+\frac{P}{d}}{q}\right) \neq 0$ and since $\left(\frac{-P}{q}\right) = -1$, there must be some prime p' such that $\left(\frac{p'}{q}\right) = -1$, by multiplicativity.

The availability of quadratic non-residues modulo q means we can successfully choose this p' as our next term in the sequence to satisfy $\left(\frac{p'}{q}\right) = -1$. Consider the new product $P' = p' \cdot P$. Then by multiplicativity of the Legendre symbol, we have

$$\left(\frac{-P'}{q}\right) = \left(\frac{p' \cdot -P}{q}\right) = \left(\frac{p}{q}\right) \cdot \left(\frac{-P}{q}\right) = (-1) \cdot (-1) = 1.$$

Then we are back to the case where P' is the product of primes in the sequence $\{p_1, \ldots, p_k, p'\}$ and $\left(\frac{P'}{q}\right) = 1$ and so we follow through as above to choose q as the next prime in the sequence. So indeed we have constructed a generalised Euclid sequence with seed X containing q.

If the conditions of Lemma 5.6 are not satisfied, in other words if $q = 5$ and $P \equiv 3 (\text{mod } 5)$, we choose our divisor d of P to be $d = 1$. Then $d + \frac{P}{d} = 1 + P$. Now, upon assumption that $P \equiv 3 (\text{mod } 5)$, it follows that $P + 1 \equiv 3 + 1 \equiv 4 (\text{mod } 5)$. So it is clear that there must exist some prime divisor p' of $P + 1$ satisfying $p' \not\equiv 1 (\text{mod } 5)$ and so by considering $P' = p' \cdot P$, then

$$P' = p' \cdot P \not\equiv 1 \cdot 3 \equiv 3 (\text{mod } 5)$$

and so the previous limitation is broken and we can proceed as initially in this case.

Case 2: $S \neq (\mathbb{Z}/q\mathbb{Z})^{\times}$.

This case is more involved than our initial case and does require a few more steps. We will slowly work through each point explaining thoroughly along the way.

Motivationally, if $S \neq (\mathbb{Z}/q\mathbb{Z})^{\times}$ we still have that $S \subsetneq (\mathbb{Z}/q\mathbb{Z})^{\times}$. So perhaps our best bet would be to attempt to *enlarge* S so it becomes more and more similar to $(\mathbb{Z}/q\mathbb{Z})^{\times}$. If we can do this successfully, eventually reaching that $S = (\mathbb{Z}/q\mathbb{Z})^{\times}$, we revert back to our initial case where it was proven that indeed we may choose q, the missing prime in question, as the next prime in our sequence. As q is an arbitrary odd prime, this indicates that this generalisation will generate every prime.

We begin by considering the set

$$T = \left\{ p \in \mathbb{N} : p \text{ prime and } p \mid \left(d + \frac{P}{d} \right) \right\}.$$

We remark that T indeed contains potential additions of primes to our sequence; we hope, in fact, that q is contained somewhere in T. And we make one further note in describing pS: we think of pS as the residue class obtained by multiplying each element of S by $p(\text{mod } q)$. We can think of this loosely as a coset of S. (Note this is a loose way to think about it since we do not necessarily have that S is a group, but this may aid with understanding).

Following this setup, let us consider the following iterative process which we use to 'expand' S:

- Choose a new prime $p \in T$ not contained in our original seed X and add it to X. In other words, we replace X with $X \cup \{p\}$.
- Recall our original product $P = p_1 \cdots p_k$ of the primes contained in the seed X. Replace this product P with $P \cdot p$ where p is as described above.
- Replace S with $S \cup pS$.

Then we essentially have expanded S by replacing S each time, since we add a new set of residue classes modulo q. By doing this, we increase our chances of q being a factor of $d + \frac{P}{d}$ since each time we iterate we are adding new prime factors to the product P. So eventually, we will either reach a point where $q \in T$ (in which case we simply choose q as the next term), or S will *stabilise* thus cannot expand any further. In other words, we eventually reach a point where $pS \subseteq S$ for each $p \in T$.

Consider an element $s \in S$, and the subgroup $H \leq (\mathbb{Z}/q\mathbb{Z})^{\times}$ where H is generated by the set $\{p + q\mathbb{Z} : p \in T\}$. It is natural then to consider the left coset sH contained within $(\mathbb{Z}/q\mathbb{Z})^{\times}$. Since S has stabilised, we may then conclude that S is exactly the union of cosets $\bigcup_{s \in S} sH$. This is due to the fact that the subgroup H is generated by elements of the form $p + q\mathbb{Z}$, and in our iterative process used earlier to 'expand' S, we essentially were multiplying elements of the original S by residue classes modulo q. Particularly, we are multiplying S by an element in H each time. Note that $s \in S$ implies that $sH \subseteq S$ by the above observation that $pS \subseteq S$ for every $p \in T$. $S = \bigcup_{s\in S} sH$ as above.

Now we observe that $|H|$ must divide $|S|$.

We now consider the case where S has not stabilised. We consider the possibility that S is *restricted* to stay within a subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$; if this is the case then our proof has failed, since then there is no chance of S growing to become $(\mathbb{Z}/q\mathbb{Z})^\times$. To begin, let $K \leq (\mathbb{Z}/q\mathbb{Z})^\times$ such that $[(\mathbb{Z}/q\mathbb{Z})^\times : K] \geq 4$.

Remark: We note here that over $(\mathbb{Z}/q\mathbb{Z})^\times$, the quadratic equation as we know it still holds. This is since, if we are searching for $d \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $d + \frac{P}{d} = h$, then

$$\begin{aligned} d + \frac{P}{d} = h &\iff d^2 - hd + P \equiv 0 \text{ simply by multiplying by } d; \\ &\iff d = \frac{h \pm \sqrt{h^2 - 4P}}{2} \text{ by the quadratic formula.} \end{aligned}$$

Thus, $d + \frac{P}{d} = h$ has solutions in $(\mathbb{Z}/q\mathbb{Z})^\times$ if and only if $\left(\frac{h^2-4P}{q}\right) \neq -1$.

In particular, $h^2 - 4P \equiv x^2 (\text{mod } q)$ for some x implies that the equation $d + \frac{P}{d} = h$ is quadratic in $(\mathbb{Z}/q\mathbb{Z})^\times$ thus has at most two solutions. Thus,

$$\left| \left\{ d \in (\mathbb{Z}/q\mathbb{Z})^\times : d + \frac{P}{d} \in K \right\} \right| \leq 2 \cdot |K| \leq \frac{1}{2}(q-1)$$

since there are at least 4 distinct left cosets of K in $(\mathbb{Z}/q\mathbb{Z})^\times$ and so it follows that $|K| \leq \frac{1}{4}(q-1)$. So, supposing that S is confined to the subgroup K, we end up contradicting Lemma 5.5 where it is stated that $|S_q| > \frac{1}{2}(q-1)$. So we dismiss the case that S is confined to this K.

Now consider any prime divisor r of $q-1$. We consider the subgroup

$$K_r = \{x^r : x \in (\mathbb{Z}/q\mathbb{Z})^\times\}$$

of $(\mathbb{Z}/q\mathbb{Z})^\times$. Note K_r has index r in $(\mathbb{Z}/q\mathbb{Z})^\times$ since $(\mathbb{Z}/q\mathbb{Z})^\times$ is cyclic under multiplication so via Lagrange's theorem,

$$[(\mathbb{Z}/q\mathbb{Z})^\times : K_r] = \frac{|(\mathbb{Z}/q\mathbb{Z})^\times|}{|K_r|} = \frac{(q-1)}{(q-1)/r} = r.$$

Our aim in the following steps is to show that S will eventually cover all of $(\mathbb{Z}/q\mathbb{Z})^\times$; if we succeed we may then follow our previous arguments to obtain our result. We reach this result by analysing the structure of H, defined above. To begin, let $r_1^{e_1} \cdots r_m^{e_m}$ be the prime factorisation of $q-1$.

Consideration 1: $r_i \geq 5$:

Let $K = K_{r_i}$ as defined above. Note we can equivalently define

$$K_{r_i} = \{k \in (\mathbb{Z}/q\mathbb{Z})^\times : r_i^{e_i} \text{ does not divide the order of } k\}.$$

We can redefine K_{r_i} in this way because, taking a generator g of $(\mathbb{Z}/q\mathbb{Z})^\times$, we have that for each $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ is such that $x = g^m$ for some $m \in \mathbb{N}$. Thus it follows that

$$K_{r_i} = \{x^{r_i} : x \in (\mathbb{Z}/q\mathbb{Z})^\times\} = \{(g^m)^{r_i} : m \in \mathbb{N}\} = \{(g^{r_i})^m : m \in \mathbb{N}\} = \langle g^{r_i} \rangle .$$

So $K_{r_i} = \langle g^{r_i} \rangle$ for generator g of $(\mathbb{Z}/q\mathbb{Z})^\times$. Then we have $k \in (\mathbb{Z}/q\mathbb{Z})^\times$ is a power of g^{r_i} if and only if $r_i^{e_i}$ does not divide the order of k.

Then we have that $|H|$ is indeed divisible by $r_i^{e_i}$ since H includes at least one element of the form $p + q\mathbb{Z}$ which is not included in K_{r_i}. We indeed obtain that $r_i^{e_i}$ divides $|H|$ by Lagrange's theorem.
Since this works for all prime factors at least 5, the only possible prime factors of the index $[(\mathbb{Z}/q\mathbb{Z})^\times : H]$ are 2 and 3.

Consideration 2: $r_i \in \{2, 3\}$:

The above argument does not work with K_2 or K_3 due to the small index size. So we consider $K_{{r_i}^2}$. If r_i^2 divides $q-1$ then the index of $K_{{r_i}^2}$ in $(\mathbb{Z}/q\mathbb{Z})^\times$ is exactly r_i^2. This is due to the fact that $K_{{r_i}^2}$ contains all elements whose orders divide $\frac{q-1}{r_i^2}$, thus $|K_{{r_i}^2}| = \frac{q-1}{r_i^2}$. The result then follows by Lagrange's theorem. Then by definition of H and noting the fact that H contains at least one element outside of $K_{{r_i}^{e_i-1}}$ we obtain that ${r_i}^{e_i-1}$ divides

$|H|$. So the index of H is not divisible by r_i^2, and then noting that this index also divides $\frac{q-1}{|H|}$ it follows that the index of H divides 6. The necessity of this point will become clear in the following case analysis of our odd prime q.

Cases for q :

Case 1: $q \not\equiv 1 (\text{mod } 3)$:

Note here that the index of H in $(\mathbb{Z}/q\mathbb{Z})^\times$ is equal to $\frac{q-1}{|H|}$ as above. Since 3 does not divide $q-1$ (via our case conditions) it follows that H cannot have index 3. But the index of H must divide $q-1$ by Lagrange's theorem so cannot be divisible by 3. Since the index of H must divide 6 by the above argument, it follows that H must have index less than or equal to 2.

Now note the following:

$$[(\mathbb{Z}/q\mathbb{Z})^\times : H] = \frac{q-1}{|H|} \leq 2 \Rightarrow \frac{1}{2}(q-1) \mid |H|$$

and the fact that S is a union of cosets of H implies that $|H| \mid |S|$, so simply noting that $|S|>\frac{1}{2}(q-1)$by Lemma 5.5 we have that the order of S is exactly equal to $q-1$ thus $S = (\mathbb{Z}/q\mathbb{Z})^\times$ as required.

Case 2: $q \equiv 1 (\text{mod } 3)$:

We consider $K = K_6$ in order to obtain that there in fact exists some $p \in T$ such that $p^{\frac{q-1}{6}} \not\equiv 1 (\text{mod } q)$. Using the choice of p, recall the following facts:

$$K_2 = \{k \in (\mathbb{Z}/q\mathbb{Z})^\times : k^{\frac{q-1}{2}} = 1\},$$

$$K_3 = \{k \in (\mathbb{Z}/q\mathbb{Z})^\times : k^{\frac{q-1}{3}} = 1\},$$

and

$$p^{\frac{q-1}{6}} = (p^{\frac{q-1}{2}} / p^{\frac{q-1}{3}}) \not\equiv 1 (\text{mod } q).$$

Indeed, this implies immediately that $p + q\mathbb{Z}$ is excluded from at least one of K_2 and K_3 since it cannot lie in the intersection $K_2 \cap K_3$ the subgroup of elements whose orders divide $\frac{q-1}{6}$.

Suppose initially that $p+q\mathbb{Z} \notin K_3$. Then H has index at most 2. Then by an earlier argument, it follows that $S=(\mathbb{Z}/q\mathbb{Z})^\times$ as required. Now suppose $p+q\mathbb{Z} \notin K_2$. So H has index dividing 3. Suppose $|S|=q-1$. Then it follows that $S=(\mathbb{Z}/q\mathbb{Z})^\times$ so our result is immediate. So assume from now that $|S|<q-1$. Recall from Lemma 5.5 that $|S|>|H|$. Note that since $|H|=\frac{1}{3}(q-1)$ that $|S|=2|H|=\frac{2}{3}(q-1)$ since if S stabilises but doesn't cover the entirety of $(\mathbb{Z}/q\mathbb{Z})^\times$ then it must consist of two cosets of H. Then we make use of our earlier argument with $K=K_3$, noting that $d+\frac{P}{d}=k$ has at most two solutions for fixed k, and that $|S|=2|K_3|$. Then since $d \mapsto d+\frac{P}{d}$ must map S 2 to 1 onto K_3, note that the quadratic formula implies that again for fixed k, we have that $\left(\frac{k^2-4P}{q}\right)=1$ for all $k \in K_3$. In other words, $k^2-4P \equiv x^2 (\text{mod } q)$ for some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$. Thus for all $x \in (\mathbb{Z}/q\mathbb{Z})^\times$, Thus we have $\left(\frac{x^6-4P}{q}\right)=1$. Note this is due to two points: firstly, since $|S|=2|K_3|$ then for $k \in K_3$, each equation $d+\frac{P}{d}=k$ has two solutions, so $\left(\frac{k^2-4P}{q}\right)=1$ each time. Secondly, since $x^3 \in K_3$ for every $x \in (\mathbb{Z}/q\mathbb{Z})^\times$, we indeed have that $\left(\frac{x^6-4P}{q}\right)=1$ for such x.

Consider $q \notin \{7,13\}$. By Lemma 5.6 we then have that there exists some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $\left(\frac{x^6-4P}{q}\right) \neq 1$, which of course is contradictory by our previous note. So now assume $q \in \{7,13\}$. Recall our definition of S_q:

$$S_q=\left\{d+q\mathbb{Z} : q \in \mathbb{N}, d \mid \prod_{p<q} p\right\}.$$

We will compute $|S_q|$ directly. If $q=7$, we have that for the definition of S_q as above, that the primes $p<q$ are exactly 2, 3 and 5. Thus noting $2 \cdot 3 \cdot 5 = 30$, we have $S_q=\{d+q\mathbb{Z} : d \mid 30\}$. Then the square-free divisors of 30 are $\{1,2,3,5,6,10,15,30\}$ and under modulo 7 these divisors are reduced to the list $\{1,2,3,5,6\}$ of disinct residue classes. Thus $|S_7|=5>4=\frac{2}{3}|(\mathbb{Z}/7\mathbb{Z})^\times|$. So indeed S_7 is exactly $(\mathbb{Z}/7\mathbb{Z})^\times$. By a similar method, it can be verified that for $q=13$ that $S_{13}=(\mathbb{Z}/13\mathbb{Z})^\times$. Thus the proof is concluded.

□

Theorem 5.4 then follows from induction of this result. We again highlight that the proof of Lemma 5.7 and subsequently Theorem 5.4 are due to [2].

6 Further variants of the Euclid-Mullin Sequences

Here we discuss a proof of one further variant of the Euclid-Mullin sequence, taking ideas from Pomerance, Booker and Wooley to define a sequence which in fact generates every prime and moreover defines the kth smallest prime within a sequence.

To begin, we state a Theorem due to Wooley [11].

Theorem 6.1. *Consider the sequence where $p_1 = 2$ and p_{k+1} is defined to be the least divisor exceeding 1 of $P^{P^P} - 1$ where $P = p_1 \cdots p_k$ as in our previous sections. Then p_{k+1} is a prime number and moreover for each k, p_k is the kth smallest prime.*

The proof for this theorem follows immediately from the following lemma, which we will prove.

Lemma 6.2. *For $n \in \mathbb{N}$, the least prime divisor of $n^{n^n} - 1$ is the smallest prime which does not divide n.*

Proof. For $n = 1$ the result is immediate, since $1^{1^1} - 1 = 0$ thus vacuously has no prime divisor. Suppose for the remainder of the proof that $n \geq 2$.

Let q be the smallest prime not dividing n. Let the primes dividing n be in the list $\{p_1, \ldots, p_k\}$. Then the inequality $q \leq p_1 \cdots p_k + 1 \leq n + 1$ holds. Then note that all prime divisors of $q - 1$ lie in the list $\{p_1, \ldots, p_k\}$.

We claim that for any natural n that $2^n \geq n + 1$. The claim follows by a simple induction proof on n. For $n = 0$, we have $2^0 = 1 \geq 0 + 1$. So suppose true for some n that $2^n \geq n + 1$. Then considering $n + 1$, we have that

$$2^{n+1} = 2 \cdot 2^n \geq 2 \cdot (n + 1) = 2n + 2 \geq n + 2 = (n + 1) + 1$$

by the Inductive Hypothesis and the fact that the multiplication function on natural numbers is strictly increasing. Thus indeed we have that $2^{n+1} \geq (n + 1) + 1$, proving the claim.

Following this claim, we note that for each $1 \leq i \leq n$, that

$$p_i{}^n \geq 2^n \geq n+1 \geq q.$$

So we have $(q-1) \mid (p_1 \dots p_k)^n$ and so $(q-1) \mid n^n$. Now define $\alpha \in \mathbb{Z}$ by $\alpha = \frac{n^n}{q-1}$. Note $q \nmid n$. Then Fermat's Little Theorem implies that $n^{n^n} = (n^\alpha)^{q-1} \equiv 1(\text{mod } q)$. In other words, $q \mid (n^{n^n}-1)$. So q, the least prime not dividing n in fact is a divisor of $n^{n^n}-1$.

□

We highlight that the proof above is again due to Wooley, [11]. Now note by Lemma 6.2, we have that in the setting of Theorem 6.1, that indeed the least prime divisor of $P^{P^P}-1$ is the smallest prime which does not divide P. But we defined $P = p_1 \cdots p_k$. Thus we can always choose our next term in our sequence to be the smallest prime missing from our list $\{p_1, \dots, p_k\}$. Thus indeed p_k is the kth smallest prime as desired. The theorem is proved.

Although intriguing, we can see that the applications of Theorem 6.1 are unpractical to say the least. Indeed, Wooley himself described his own result to have computational value "less than nanoscopic"[11]. We illustrate this below.

Example (Application of Theorem 6.1)**.** First let $n = 2$ and consider the natural number

$$n^{n^n} - 1 = 2^{2^2} = 16 - 1 = 15.$$

Then indeed the least divisor exceeding 1 of 15 is 3, the smallest prime which does not divide 2 thus verifying the result of the theorem.

Now let $n = 3$. Then, approximately

$$n^{n^n} - 1 = 3^{3^3} - 1 \approx 7.62559748 \times 10^{12}.$$

Clearly, finding the least prime divisor of this immediately immense natural number will prove to be very difficult. So it is clear that the theorem has limited application in a practical sense.

We make one final note on the sequence variants by stating one final theorem. This is another direct consequence of Lemma 6.2, and requires no knowledge of previously found primes.

Theorem 6.3. *Let $N \in \mathbb{N}$. Then the smallest prime exceeding N is the least divisor (exceeding* 1*) of* $N!^{N!^{N!}} - 1$.

Once again, this theorem's practical application is incredibly limited: computing $N!^{N!^{N!}} - 1$ for any $N \geq 2$ will produce an intimidatingly large number to factor.

7 Shanks' conjecture

Previously we touched on Shanks' published conjecture regarding the first Euclid-Mullin sequence [10]; we state it below in his own words.

Note: Let us denote the first Euclid-Mullin sequence as $\{p_i\}$ for the entirety of this section. We introduce this notation for simplicity in explanation of Shanks' arguments.

Conjecture 7.1 (Shanks' Conjecture, 1984)**.** *The first Euclid-Mullin sequence, $\{p_i\}$, is a rearrangement of all of the primes. In other words, each prime number q is equal to p_n for one and only one index n.*

Shanks' published his heuristic argument [10] for this conjecture in 1984. In this section we will outline his observations.

Let us begin by defining q to be the smallest prime which has not occurred up to p_N for some index N. For example, for $N = 6$ we have $q = 5$; for $N = 14$, we have that $q = 19$. Alongside this we note that $P_N = p_1 \cdots p_N$. We then write

$$P_{N-1} \equiv r_1 \pmod{q}; p_N \equiv r_2 \pmod{q}.$$

It follows that $r_1 \not\equiv 0 \pmod{q}$ and $r_2 \not\equiv 0 \pmod{q}$. Then, simply by recollection of how we generate terms in $\{p_i\}$, we note the following:

$$q = p_{N+1} \iff r_1 r_2 + 1 \equiv 0 \pmod{q}. \tag{4}$$

Proof. Firstly suppose that $q = p_{N+1}$. Then immediately we have

$$r_1 r_2 + 1 \equiv P_{N-1} \cdot p_N + 1 \equiv P_N + 1 \equiv 0 \pmod{q}$$

since we must have $q \mid P_N + 1$ by definition of the inclusion of q as term p_{N+1} in $\{p_i\}$. Conversely, supposing that $r_1 r_2 + 1 \equiv 0 \pmod{q}$, then the reverse of the above congruence is clearly satisfied, thus

$$P_N + 1 \equiv r_1 r_2 + 1 \equiv 0 \pmod{q}.$$

Thus $q = p_{N+1}$. □

We have that if q occurs as p_{N+1} in $\{p_i\}$ then indeed $q \mid (P_N + 1)$ as demonstrated above. If q never occurs, we have that:

1. Each P_N is always coprime to q;

2. $P_N + 1$ is divisible by q at most finitely many times, since q is defined to be the smallest prime not occurring up to p_N thus if $q \mid (P_N + 1)$ but q does not occur in $\{p_i\}$ then there must be some smaller prime divisor of $P_N + 1$ which instead occurs.

Now note there are $q - 1$ choices for $P_N \pmod{q}$; thus it is highly probable that $P_N \equiv -1 \pmod{q}$ should occur at least once. This is equivalently stated as $P_N + 1 \equiv 0 \pmod{q}$. Then unless a prime smaller than q is not included in $\{p_i\}$ we take q as p_{N+1}. This is immediate from definition of q: the smallest prime not included in $\{p_i\}$ up to p_N.

Let us consider the highly improbable outcome that $P_N \equiv -1 \pmod{q}$ never occurs. Then this means there are infinitely many choices of P_N such that P_N 'avoids' hitting the residue class of $-1 \pmod{q}$. We observe that this outcome would be more than a little strange; why would the seemingly random behaviours of primes continually follow a pattern in avoiding this residue class?

Shanks finishes his argument by highlighting that once the occurrence of $q = p_{N+1}$ takes place, the prime q can never occur again in the sequence since following this, $r_1 \equiv 0 \pmod{q}$ is always satisfied. Finally, he observes some slight number theoretic and probabilistic setbacks, but claims that these are minor and should not disturb his argument overall. We dismiss further mention of them here.

It seems appropriate to finish in the same manner in which Shanks concludes his article: with a remark taken from a personal correspondence between himself and Dyson, who indeed suggests, one further time, “our conjecture may in fact be undecidable”[10].

References

[1] Menny Aka, Manfred Einsiedler, and Thomas Ward. *A journey through the realm of numbers—from quadratic equations to quadratic reciprocity.* Springer Undergraduate Mathematics Series. Springer, Cham, [2020] ©2020, pp. xix+344. ISBN: 978-3-030-55233-6; 978-3-030-55232-9. DOI: 10.1007/978-3-030-55233-6. URL: https://doi.org/10.1007/978-3-030-55233-6.

[2] Andrew R. Booker. “A variant of the Euclid-Mullin sequence containing every prime”. In: *J. Integer Seq.* 19.6 (2016), Article 16.6.4, 6. ISSN: 1530-7638.

[3] Andrew R. Booker. “On Mullin’s second sequence of primes”. In: *Integers* 12.6 (2012), pp. 1167–1177. ISSN: 1867-0652,1867-0660. DOI: 10.1515/integers-2012-0034. URL: https://doi.org/10.1515/integers-2012-0034.

[4] C. D. Cox and A. J. Van der Poorten. “On a sequence of prime numbers”. In: *J. Austral. Math. Soc.* 8 (1968), pp. 571–574.

[5] Romeo Meštrović. *Euclid’s theorem on the infinitude of primes: a historical survey of its proofs (300 B.C.–2022) and another new proof.* 2023. arXiv: 1202.3670 [math.HO]. URL: https://arxiv.org/abs/1202.3670.

[6] Thorkil Naur. “Mullin’s sequence of primes is not monotonic”. In: *Proc. Amer. Math. Soc.* 90.1 (1984), pp. 43–44. ISSN: 0002-9939,1088-6826. DOI: 10.2307/2044665. URL: https://doi.org/10.2307/2044665.

[7] Paul Pollack. *Not always buried deep.* A second course in elementary number theory. American Mathematical Society, Providence, RI, 2009, pp. xvi+303. ISBN: 978-0-8218-4880-7. DOI: 10.1090/mbk/068. URL: https://doi.org/10.1090/mbk/068.

[8] Paul Pollack and Enrique Treviño. "The primes that Euclid forgot". In: *Amer. Math. Monthly* 121.5 (2014), pp. 433–437. ISSN: 0002-9890,1930-0972. DOI: 10.4169/amer.math.monthly.121.05.433. URL: https://doi.org/10.4169/amer.math.monthly.121.05.433.

[9] Kenneth Rogers. "The Schnirelmann density of the squarefree integers". In: *Proc. Amer. Math. Soc.* 15 (1964), pp. 515–516. ISSN: 0002-9939,1088-6826. DOI: 10.2307/2034736. URL: https://doi.org/10.2307/2034736.

[10] Daniel Shanks. "Euclid's primes". In: *Bull. Inst. Combin. Appl.* 1 (1991), pp. 33–36. ISSN: 1183-1278,2689-0674.

[11] Trevor D. Wooley. "A superpowered Euclidean prime generator". In: *Amer. Math. Monthly* 124.4 (2017), pp. 351–352. ISSN: 0002-9890,1930-0972. DOI: 10.4169/amer.math.monthly.124.4.351. URL: https://doi.org/10.4169/amer.math.monthly.124.4.351.