Bachelor Thesis
Prime Generators
10/01/2020

Lucas Hoogendijk (5901413)
Supervisor: Damaris Schindler

Utrecht University

Contents

1 Introduction

In mathematics, prime numbers have always played a major role, especially in number theory. The definition of such a prime number is a number greater than 1 that is divisible solely by one and itself. Numbers which are not prime are called composite numbers. The importance of prime numbers in mathematics has been immortalized in the Fundamental Theorem of Arithmetic:

Theorem 1.1 (Fundamental Theorem of Arithmetic) *Any whole number n greater than 1 is either prime or can be written uniquely as a product of prime numbers:*

$$n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} \tag{1}$$

where $p_1 < p_2 < \cdots < p_k$ are prime numbers and their exponents r_i are whole positive numbers. We refer to this unique product as the prime factorization of n.

The proof of this theorem splits into two parts; the first part is to show that such a factorization exists, the second is to show its uniqueness. To show that such a factorization exists we shall use induction. For the case that $n \in \{2, 3\}$ we know naturally that these are prime numbers. Now assume that for all numbers less than $n+1$ we have proven that such numbers are either prime or have their factorization. If $n+1$ is itself prime then the theorem quickly follows, so now consider that $n+1$ is not prime. Then there must exist two numbers a, b, not necessarily prime, such that $a \cdot b = n+1$. These two numbers are less than $n+1$ so by our induction hypothesis they must have their factorization, whence we can conclude that the factorization of $n+1$ is the product of the factorization of a and b.

Now that we have proven such a factorization exists we are left with the uniqueness. We will prove this through contradiction but we will need the help of the following Lemma:

Lemma 1.2 *Let $a, b, p \in \mathbb{Z}_{>0}$ with p prime and let p be a divisor of ab, then p divides at least one of a and b.*

Proof: Without loss of generality we shall assume that p does not divide a and show that then, p divides b. By Bézout's identity we know that for any integers x, y which share only 1 as their common divisors, we can find integers r, s such that $rx + sy = 1$. As p is only divisible by 1 and itself and p does not divide a we can see that a and p must be such a pair, so that we can find r, s such that $rp + sa = 1$. If we now multiply both sides by b, we obtain $rpb + sab = b$. Now p obviously divides the first term, and the second term sab is divisible by p by assumption that p divides ab. We therefore see that the entire left-hand-side must be divisible by p and therefore so must b, so we may conclude that p divides b. □

Suppose n is smallest number to have two different factorizations, namely

$$n = \prod_{i=1}^{k} p_i^{r_i} \text{ and } n = \prod_{j=1}^{l} q_j^{s_j}. \tag{2}$$

As p_1 divides n it must therefore divide at least one of the prime numbers q_j, this follows from the aforementioned Lemma. By definition of prime numbers q_j can only be divisible by 1 and itself so we can conclude that $p_1 = q_j$. As p_1 has to be the smallest (prime) number greater than 1 to divide n, therefore q_k must also be the smallest (prime) number of the second factorization so we know that $p_1 = q_1$. Then we can divide both sides by p_1 exactly once to end up with a new number n' which has two different factorizations, namely $n' = p_1^{r_1 - 1} \prod_{i=2}^{k} p_i^{r_i}$ and $n' = q_1^{s_1 - 1} \prod_{j=2}^{l} q_j^{s_j}$. This however contradicts the assumption that n was the smallest

number with this property, so we can conclude that the factorization of any whole number greater than 1 must be unique. □

The study of these building blocks of number theory can be traced back to the Greeks, where Euclid had been studying such numbers. In his own work he had already proven the infinitude of prime numbers. His proof is as follows:

Theorem 1.3 (Euclid's proof for the infinitude of prime numbers) *There exist infinitely many prime numbers.*

Proof: Suppose we have a finite set of prime numbers, namely $\{p_1, p_2, \ldots, p_r\}$. Let P be the product of all our prime numbers in this set, plus one. So $P = p_1 \cdot p_2 \cdot p_r + 1$. Then P is either prime or composite. If it is prime then it can't equal any of our initial primes, so we have found a new prime number we can add to our set. If P is composite, then there is at least one prime number p such that p divides P. This p can not be equal to any of our initial primes, otherwise we would come to the conclusion that p divides 1, which we know is impossible. Thus again we can add a prime number to our set of primes and we can keep doing this indefinitely, so there exist infinitely many primes. □

I first encountered this proof in secondary school, aged around 16. Whether it was the simplicity of this proof or how intuitive it felt, this proof has always stuck with me and fueled my interest in number theory. Soon thereafter while bored in class I started to calculate the first few terms, starting with the smallest prime 2, then 3, 7 and 43. As 5 was skipped here, I wanted to know whether it would ever reoccur, however at the time I neither had the knowledge nor the resources to look into this. It would be a few years later that I had the opportunity to delve further into this matter with this thesis. What I did manage to find out at the time was that sequence of course had already been looked at and is known as Sylvester's Sequence, defined as:

$$a_n = 1 + \prod_{i=1}^{n-1} a_i \text{ with } a_1 = 2 \tag{3}$$

It is important to note that this sequence does not generate only prime numbers, as $a_5 = 1807 = 13 \cdot 139$. This sequence however generates numbers coprime, i.e. the only shared divisor is 1, to all previous numbers. The numbers in this sequence and their prime factors have been researched with provable result; it can be shown that infinitely many primes are omitted which we will show with our own proof in Section 4.1.4.

Of the many proofs for the infinitude of primes, there are two more, in my opinion, intriguing ways to show the this result which I will briefly mention here. The first follows from the Riemann zeta function:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}} \tag{4}$$

If we take $s = 1$, we see that we end up with the harmonic series:

$$\sum_{n=1}^{\infty} \frac{1}{n} = \prod_{p \text{ prime}} \frac{1}{1 - \frac{1}{p}} \tag{5}$$

As we know that the harmonic series diverges to infinity, we can therefore conclude that the right-hand-side must too, which is impossible if there are only finitely many primes. Therefore we can use the harmonic series to show that there must exist infinitely many prime numbers. The second method originates from Euler's formula for π which can be derived from the Leibniz series:

$$\frac{\pi}{4} = \frac{3}{4} \cdot \frac{5}{4} \cdot \frac{7}{8} \cdot \frac{11}{12} \cdot \frac{13}{12} \cdot \frac{17}{16} \cdot \frac{19}{20} \cdot \frac{23}{24} \cdots \tag{6}$$

where the numerators are the odd primes and the denominators are the multiples of 4 closest to the respective prime. As π is proven to be irrational, this must be an infinite product as otherwise we would have a rational representation of π, so therefore we can again conclude that there exist infinitely many prime numbers. Another contribution, albeit more recreational, made by Euler was the polynomial

$$n^2 - n + 41 \tag{7}$$

which produces 40 prime numbers for $1 \leq n \leq 40$. Sadly this polynomial and the two proofs above are merely interesting and offer no fruitful way of producing more primes. For polynomials this one turns out to be the most rewarding with 40 primes, with the arithmetic progression

$$L(n) = 43142746595714191 + 5283234035979900n \tag{8}$$

taking the second place for producing 26 primes for $0 \leq n \leq 25$. Even though it has been proven in 2008 by Green and Tao in [29] that for any k there exists a pair a, b such that $L(n) = an + b$ gives primes for $0 \leq n \leq k-1$, few actual linear or polynomial examples are known which possess such progression of primes.

Later I would discover Mullin had modified Euclid's proof into a construction to generate exclusively prime numbers. His construction is as follows; let $p_1 = 2$, then for all $k \geq 1$ we can choose p_{k+1} prime such that p_{k+1} divides $(p_1 p_2 \cdots p_k) + 1$. As this sequence still is dependent on choice, Mullin decided to focus on two of these sequences. Firstly, choose p_{k+1} as small as possible at each step; resulting in the following sequence

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, 5471, \ldots \tag{9}$$

known in the OEIS[4] as A000945. The second possibility is to do the opposite, choose p_{k+1} as large as possible to obtain the sequence

$$2, 3, 7, 43, 139, 50207, 340999, 2365347734339, 4680225641471129, \ldots \tag{10}$$

which can be found in the OEIS as A000946. We call these the first and second Euclid-Mullin sequence respectively. The natural question raised by Mullin was whether these sequences would include every prime. His main question was actually more general than this.

Question 1 (Mullin's Question) *For any given set of primes $\{p_1, \ldots, p_n\}$ and choose p_{k+1} as the smallest, or largest, prime to divide $1 + \prod_{i=1}^{k} p_i$ for $k \geq n$, will every prime number eventually appear in these sequences?*

Most effort into the understanding of such sequences went into variants in a hope to give us a better idea how to even tackle this simply stated but complex question. Sadly, for the sequence mentioned in Mullin's question little is known. Even though empirical evidence by Vardi[10] and a compelling heuristic argument by Shanks [18] would suggest all primes do occur in this first Euclid-Mullin sequence, no real progress has been made in the understanding of the behaviour of this sequence. What more makes this problem so difficult is the apparent randomness in the order of primes and how, after only a few steps, are required to find the factorization of very large numbers; for the 44th term a 180-digit number had to be factored of which the smallest factor is a 68-digit prime. So far only 51 indices of this sequence are known and the smallest prime yet to appear is 41. For the second sequence Cox and van der Poorten [23] were able to show for some small primes that they would never appear and conjectured that infinitely many primes would not appear in the second sequence. This result would later be proven by Andrew R. Booker in [8].

Inspiration for this thesis started with Wooley's article [2], in which he briefly discusses two constructions by Pomerance and Booker which both provably contain every prime in their respective sequences given any starting set, sometimes called the seed of the sequence. Wooley mentions that both of these require a factor

of choice for terms p_{k+1} and proceeds by introducing his own construction which provably generates every prime number and where each p_k is also the k-th smallest prime. We ourselves then delve further into a refined variant merely mentioned in Wooley's variant. From this article we then examined more questions similar to the question raised by Mullin with results from e.g. Booker and Pomerance. We conclude our work with our very own sequence and question:

Question 2 *For any given set of primes $\{p_1, \ldots, p_n\}$ and choose p_{k+1} as the smallest prime to divide $\prod_{i \in I} p_i + \prod_{i \in \{1,\ldots,k\} \setminus I} p_i$ for some $I \subseteq \{1, \ldots, k\}$ and $k \geq n$, will every prime number of the form $4m + 1$ eventually appear in this sequence?*

2 Background

In this section we discuss some essential background knowledge.

It is important to point out slight difference in notation which will be used in this document. Most notably for the cardinality of sets we use:

$$\#S \coloneqq |S|$$

where S is any set. This choice was made to improve readability when used in combination with "$a \mid b$", the symbol to show that a divides b. For the absolute value however we will still use the notation $|n|$.

2.1 Group theory

The following result from group theory which we will make us of is that of Lagrange which is as follows:

Theorem 2.1 (Lagrange's Theorem) *The order of a subgroup of a finite group must divide the order of the group.*

The index of such a subgroup is defined as the order of the group divided by the order of the subgroup. This theorem gives us the following useful corollaries:

Corollary 2.1.1 *The order of every element of a group G divides the order of G.*

Corollary 2.1.2 *If $x \in G$ then $x^{\#G} = e$, where e denotes the identity element.*

Another observation from group theory is on cyclic groups, namely: let G be a cyclic group, then there exists an element $x \in G$ which generates G. This means that for any other element $y \in G$, there exists a positive integer $r \leq \#G$ such that $y = x^r$ and we can thus represent G as $G = \{1, x, x^2, \ldots, x^{\#G-1}\}$.

2.2 Number theory

The motivation for using congruence is that it allows for generalization on theorems when dealing with divisibility. We call two numbers $a, b \in \mathbb{Z}$ congruent modulo M if their difference is a multiple of M. Note

that this is an equivalence relation, with common notation for the aforementioned being $a \equiv b \mod M$. These congruence classes are denoted with $(\mathbb{Z}/M\mathbb{Z})$ with representatives of the classes being $\{0, 1, \ldots, M-1\}$. Under addition we see that $(\mathbb{Z}/M\mathbb{Z})$ is a group, however for multiplication alone we instead need a subset of these classes. For those familiar with rings, $(\mathbb{Z}/M\mathbb{Z})$ is of course an example of a ring.

2.2.1 The set of invertible congruence classes

We denote the set of invertible congruence classes modulo M as $(\mathbb{Z}/M\mathbb{Z})^\times$. A number a is invertible modulo M if there exists a number b such that $a \cdot b \equiv 1 \mod M$. This happens to be the case if and only if such a is coprime to M. If we take M to be 6, then we see we get $(\mathbb{Z}/6\mathbb{Z})^\times = \{1, 5\}$. If we take M to be prime, say 7, then for its invertible congruence classes we are left with all except zero: $\{1, \ldots, M-1\}$.

2.2.2 Euler's totient function

Euler's totient function, $\phi(M)$, is defined to denote the number of invertible congruence classes modulo M, i.e. the amount of numbers less than M coprime to M, so $\phi(M) = \#(\mathbb{Z}/M\mathbb{Z})^\times$. Again we are mostly interested in M prime, for which we know:

$$\phi(M) = \#\{1, 2, \ldots, M-1\} = M - 1. \tag{11}$$

2.2.3 Legendre's symbol

Not only are we interested in the congruence classes but also in the quadratic character of a number modulo p, with p an odd prime. This is where Legendre's symbol comes into play, which is defined as follows: for p an odd prime and $a \in \mathbb{Z}$, $\left(\frac{a}{p}\right) = 1$ if there exists an $r \in \mathbb{Z}$ such that $r^2 \equiv a \mod p$, $\left(\frac{a}{p}\right) = -1$ if such r does not exist and $\left(\frac{a}{p}\right) = 0$ if $p \mid a$. It is important to point out that whilst $\frac{a}{b}$ is used for fractions, in this article $\left(\frac{a}{b}\right)$ instead is used solely to denote Legendre's Symbol.

An important property of this symbol is that it is multiplicative, i.e. for all $a, b \in \mathbb{Z}$ and p an odd prime, $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$. It is also useful to remind oneself of the following property: for $a, b \in \mathbb{Z}$ and p an odd prime, $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$ if $a \equiv b \mod p$.
The following theorem was proven by Euler and later used by Legendre as a definition for his symbol for the quadratic character:

Theorem 2.2 (Euler) *Let p be an odd prime and $a \in \mathbb{Z}$ not a multiple of p. Then we have that*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p. \tag{12}$$

This theorem especially is interesting with the following corollary for the special case that $a = -1$:

Corollary 2.2.1 *Let p be an odd prime. Then $\left(\frac{-1}{p}\right) = 1$ if and only if $p \equiv 1 \mod 4$.*

We only use the special case $a = -1$ however this theorem 2.2 can be useful for other choices of a. We would be remiss if we were to leave out one of the more famous results considering the quadratic character, namely the law of quadratic reciprocity:

Theorem 2.3 (The law of quadratic reciprocity) *If p, q are odd distinct primes then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \tag{13}$$

This law is attributed to Gauss, as he was the first to prove it, which he did in 6 different ways. From this law we can see that $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ if and only if at least one of p and q is equivalent to 1 modulo 4, as then the exponent is an even number.

2.2.4 The primorial $p\#$

The name primorial originates from Dubner. Where factorial is defined as

$$n! = \prod_{i=1}^{n} i \tag{14}$$

for $n \in \mathbb{Z}_{\geq 1}$, or the product of all integers less than or equal to n, the factorial is defined as

$$p\# = \prod_{\substack{q \text{ prime} \\ q \leq p}} q \tag{15}$$

or the product of all prime numbers less than or equal to p.

2.2.5 The p-adic valuation

The p-adic valuation $\nu_p(m)$, with p prime and $m \in \mathbb{Z}_{\geq 0}$, is defined to be the exponent of p in the prime factorization of n; it is the exponent such that $p^{\nu_p(m)}$ divides m, but $p^{\nu_p(m)+1}$ does not.
It is important to note that for $a, b \in \mathbb{Z}_{\geq 0}$:

$$\nu_p(a \cdot b) = \nu_p(a) + \nu_p(b) \tag{16}$$

and that

$$\nu_p(\frac{a}{b}) = \nu_p(a) - \nu_p(b). \tag{17}$$

2.2.6 Legendre's Formula

Also known as De Polignac's Formula, the p-adic valuation of factorials, so $\nu_p(n!)$ with $n \in \mathbb{Z}_{\geq 0}$, can be expressed as the following series:

$$\nu_p(n!) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor. \tag{18}$$

We will not prove this formula but instead show its workings with the following neat question:
What is the number of closing zeroes of 2020!, i.e. what is the number of zeroes at the end of this number?
To tackle this problem recall that the number of closing zeroes depends on what number of times 10 is a divisor of our number 2020!. As $10 = 2 \cdot 5$ and 2 will definitely occur as a factor in 2020! more often than 5 will, this reduces the question to how many times 5 divides 2020!, or $\nu_5(2020!)$. The number of multiples

of 5 less than 2020 has to be $\lfloor \frac{2020}{5} \rfloor$ so we can add one factor of 5 for each of those. However 25 not only is a multiple of 5, it contributes two factors of 5, so we still need to count an extra factor for the multiples of $25 = 5^2$, which in turn will equal $\lfloor \frac{2020}{5^2} \rfloor$. We repeat this for $125 = 5^3$ and $625 = 5^4$ and see that the number of factors 5 in 2020! must equal $\lfloor \frac{2020}{5} \rfloor + \lfloor \frac{2020}{5^2} \rfloor + \lfloor \frac{2020}{5^3} \rfloor + \lfloor \frac{2020}{5^4} \rfloor = 404 + 80 + 16 + 3 = 503$.

2.2.7 p-smooth

We call a number n p-smooth if all prime factors in its factorization are less than or equal to p. We will also consider squarefree p-smooth numbers, where each of the prime factors in the prime factorization can only occur once. An interesting relation between the primorial and squarefree p-smooth numbers is that the divisors of the primorial $p\#$ and squarefree p-smooth numbers are the same.

2.2.8 Dirichlet's theorem

There are many proofs to show there exist infinitely many primes of some specific form, e.g. for primes $4k - 1$ or $6k - 1$, however in 1837 Dirichlet managed to prove this for all forms $a + nd$,

Theorem 2.4 (Dirichlet's theorem on arithmetic progressions) *For a, d coprime, there exist infinitely many prime numbers in the arithmetic progression $a + nd$.*

2.3 Prime Number Theorem

It would be wrong to discuss prime numbers and never mention the Prime Number Theorem (PNT).

Theorem 2.5 (Prime Number Theorem) *Let $\pi(x)$ denote the prime counting function, i.e. the number of primes less than or equal to x. Then $\pi(x) \sim \frac{x}{\ln x}$. In asymptotic notation this is equivalent to saying $\pi(x) = O(\frac{x}{\ln x}) = \frac{x}{\ln x} + o(\frac{x}{\ln x})$.*

Because of the prime number theorem we can interestingly obtain an asymptotic formula for the aforementioned primorial $p_n\#$. We do this by taking the logarithm of the primorial so we can instead consider the following sum

$$\ln\left(\prod_{i=1}^{n} p_i\right) = \sum_{i=1}^{n} \ln p_i \tag{19}$$

We first calculate $\sum_{p \leq x} (\ln x - \ln p)$ as we will need it in the following result:

$$\sum_{p \leq x \text{ prime}} (\ln x - \ln p) = \sum_{p \leq x \text{ prime}} \int_p^x \frac{1}{t} \mathrm{d}t \tag{20}$$

This step we obtain from the Fundamental Theorem of Calculus

Theorem 2.6 (Fundamental Theorem of Calculus)

$$F(b) - F(a) = \int_a^b f(x) \mathrm{d}x \tag{21}$$

where $f(x) = F'(x)$.

If we consider the summation in 20, $\sum_{p\leq x \text{ prime}} \int_p^x \frac{1}{t}\mathrm{d}t$, we see that a number k between 1 and x is included in the summation exactly the same number of times we take the integral such that $p \leq k \leq x$, which occurs $\pi(k)$ times, once for every prime less than k. This means we can rewrite the the summation to:

$$\sum_{p\leq x \text{ prime}} \int_p^x \frac{1}{t}\mathrm{d}t = \int_1^x \frac{\pi(t)}{t}\mathrm{d}t \tag{22}$$

From the prime number theorem we know there exists a $c > 0$ such that $\pi(t) \leq \frac{ct}{\ln t}$, as $\pi(x) = O(\frac{t}{\ln t})$, so we can bound the integral with $\frac{\pi(t)}{t} = \frac{c}{\ln t}$:

$$\sum_{p\leq x \text{ prime}} (\ln x - \ln p) = \int_1^x \frac{\pi(t)}{t}\mathrm{d}t = O\left(\int_1^x \frac{1}{\ln t}\mathrm{d}t\right) = o(x) \tag{23}$$

Now we can return to calculating our original sum from 19

$$\sum_{i=1}^{n} \ln p_i = \sum_{i=1}^{n} (\ln p_n - \ln p_n + \ln p_i) = \pi(p_n)\ln p_n - \sum_{i=1}^{n} (\ln p_n - \ln p_i) \tag{24}$$

The PNT gives us the estimation that $\pi(p_n) = \frac{p_n}{\ln p_n} + o(\frac{p_n}{\ln p_n})$ and for the sum we calculated the estimation above.

$$\sum_{i=1}^{n} \ln p_i = \left(\frac{p_n}{\ln p_n} + o(\frac{p_n}{\ln p_n})\right)\ln p_n - o(p_n) = p_n + o(p_n) = p_n(1 + o(1)) \tag{25}$$

The PNT too allows us to estimate the above with $p_n = n\ln n + o(n\ln n)$

$$\ln(p_n\#) = \sum_{i=1}^{n} \ln p_i = n\ln n(1 + o(1)) \tag{26}$$

Now we can take the exponents of the left and right-hand-side to obtain an asymptotic formula for the primorial:

$$p_n\# = e^{n\ln n(1+o(1))} \tag{27}$$

2.4 Bertrand-Chebyshev Theorem

The following theorem will show essential in our proof in Section 3.1. This actually follows from the PNT, too: $\frac{p_{n+1}}{p_n} \to \frac{(n+1)\ln(n+1)}{n\ln n} \to 1$ as $n \to \infty$, however the This result was first proven by Chebyshev in 1850, however our proof will follow Erdös' proof, as described in [5] by Victor H. Moll.

Theorem 2.7 (Bertrand-Chebyshev Theorem) *For all $n \in \mathbb{Z}_{\geq 1}$, there exists a prime number p such that $n < p \leq 2n$.*

We shall prove this result by examining the following product:

$$\prod_{n<p\leq 2n} p,$$

where this product, and further products in this proof, runs over p prime. According to the stated theory this is never an empty product, i.e. it is always greater than 1. We will show that this is the case through contradiction; assume for a certain n that there is no prime p such that $n < p \leq 2n$ and show that this

leads to an impossible situation. More precisely, we will show that an upper bound will be less than a lower bound.
We will examine this product with the help of the prime factorization of $\binom{2n}{n}$. In this proof with use $\nu_p(m)$, the p-adic valuation of a number m, as defined in 2.2.5:

$$\begin{aligned}\binom{2n}{n} &= \prod_{p\leq 2n} p^{\nu_p(\binom{2n}{n})} \\ &= \prod_{p\leq\sqrt{2n}} p^{\nu_p(\binom{2n}{n})} \prod_{\sqrt{2n}<p\leq\frac{2}{3}n} p^{\nu_p(\binom{2n}{n})} \prod_{\frac{2}{3}n<p\leq n} p^{\nu_p(\binom{2n}{n})} \prod_{n<p\leq 2n} p^{\nu_p(\binom{2n}{n})}.\end{aligned}$$

Where it now may seem we arbitrarily split up this product, we hope to show why it makes sense to split the range up into these pieces. For each of these products, our aim is to find a closed expression by which the products are bounded so that we may find a closed form upper bound for $\binom{2n}{n}$.
For $p \leq \sqrt{2n}$, we will use the following inequality:

$$\prod_{p\leq\sqrt{2n}} p^{\nu_p(\binom{2n}{n})} \leq (2n)^{\sqrt{2n}}. \tag{28}$$

This inequality follows from the following lemma:

Lemma 2.8 *If p is prime and divides $\binom{2n}{n}$, then $p^{\nu_p(\binom{2n}{n})} \leq 2n$.*

Proof: Let $l \in \mathbb{Z}_{\geq 0}$ such that $p^l \leq 2n < p^{l+1}$. Then

$$\begin{aligned}\nu_p\left(\binom{2n}{n}\right) &= \nu_p((2n)!) - \nu_p(n!) &(29)\\ &= \sum_{i\geq 1}\left\lfloor\frac{2n}{p^i}\right\rfloor - 2\sum_{i\geq 1}\left\lfloor\frac{n}{p^i}\right\rfloor &(30)\\ &= \sum_{i=1}^{l}\left(\left\lfloor\frac{2n}{p^i}\right\rfloor - 2\left\lfloor\frac{n}{p^i}\right\rfloor\right) &(31)\end{aligned}$$

By choice of l we know that $\left\lfloor\frac{2n}{p^m}\right\rfloor = 0$ for $m > l$, so we know it makes no sense to sum any further than $i = l$. Moreover for the floor function, there are two cases to distinguish for $x \in \mathbb{Q}$ (note that the following also holds true for $x \in \mathbb{R}$ but that $\mathbb{Q}$ is sufficient): let $\lfloor x\rfloor = n$ for some $n \in \mathbb{Z}$, then either $x \in \left[n, n+\frac{1}{2}\right)$ or $x \in \left[n+\frac{1}{2}, n+1\right)$. These two cases give us different outcomes for $\lfloor 2x\rfloor - 2\lfloor x\rfloor$:
For $x \in \left[n, n+\frac{1}{2}\right)$:

$$\lfloor 2x\rfloor - 2\lfloor x\rfloor = 2n - 2n = 0 \tag{32}$$

and for $x \in \left[n+\frac{1}{2}, n+1\right)$:

$$\lfloor 2x\rfloor - 2\lfloor x\rfloor = 2n + 1 - 2n = 1 \tag{33}$$

which means that we can bound every $\left(\left\lfloor\frac{2n}{p^i}\right\rfloor - 2\left\lfloor\frac{n}{p^i}\right\rfloor\right)$ by 1:

$$\begin{aligned}\nu_p\left(\binom{2n}{n}\right) &\leq \sum_{i=1}^{l} 1 \\ &\leq l. &(34)\end{aligned}$$

So indeed we have

$$\nu_p\left(\binom{2n}{n}\right) \leq l \tag{35}$$

$$p^{\nu_p(\binom{2n}{n})} \leq p^l \leq 2n \tag{36}$$

$$\prod_{p\leq\sqrt{2n}} p^{\nu_p(\binom{2n}{n})} \leq \prod_{p\leq\sqrt{2n}} 2n = (2n)^{\sqrt{2n}} \qquad \square$$

which concludes the upper bound for the first product.

Note that by Lemma 2.8 we now see that for all $p > \sqrt{2n}$ we have $0 \leq \nu_p(\binom{2n}{n}) \leq 1$, as for $\nu_p(\binom{2n}{n}) \geq 2$ we have that $p_p^{\nu}(\binom{2n}{n}) > 2n$. Whence we can ignore the p-adic valuation by simply bounding the p-adic valuation by 1:

$$\prod_{\sqrt{2n}<p\leq\frac{2}{3}n} p^{\nu_p(\binom{2n}{n})} \leq \prod_{\sqrt{2n}<p\leq\frac{2}{3}n} p \tag{37}$$

For the second of the products we will make us of the following bound:

$$\prod_{\sqrt{2n}<p\leq\frac{2}{3}n} p \leq 2^{\frac{4}{3}n} \tag{38}$$

which will follow from the following Lemma:

Lemma 2.9 *For all* $m \in \mathbb{Z}_{\geq 1}$

$$\prod_{p\leq m} p \leq 2^{2m} \tag{39}$$

We will prove this through induction. We can check by hand that this is true for $m = 2$ and that it also holds true for $m = 1$ as we have an empty product on the left-hand-side; which equals 1. As is usual for induction we now assume the statement holds true for values less than m.
In the case that m is even, we know that m itself can't be prime:

$$\prod_{p\leq m} p = \prod_{p\leq m-1} p \tag{40}$$

and because we know that the Lemma holds true for values less than m:

$$\prod_{p\leq m} p \leq 2^{2(m-1)} \tag{41}$$

$$\leq 2^{2m} \tag{42}$$

which concludes the case when m is even.
We are left with the case that m is odd, i.e. $m = 2k + 1$ for some $k \in \mathbb{Z}_{\geq 1}$. Note that for $k + 2 \leq p \leq 2k + 1$ we have that $\nu_p(\binom{2k+1}{k}) = 1$. We will use this fact as follows; as for all $k + 2 \leq p \leq 2k + 1$ we have that $p \mid \binom{2k+1}{k}$, we know that the product of all these primes p must divide $\binom{2k+1}{k}$ and therefore we have that:

$$\prod_{k+2\leq p\leq 2k+1} p \leq \binom{2k+1}{k} \tag{43}$$

Moreover, for the binomial we have that

$$2\binom{2k+1}{k} = \binom{2k+1}{k} + \binom{2k+1}{k+1} \tag{44}$$

$$\leq \sum_{i=0}^{2k+1} \binom{2k+1}{i} = 2^{2k+1} \tag{45}$$

so that we have $\binom{2k+1}{k} \leq 2^{2k}$. Now we have all tools necessary for the case that $m = 2k+1$:

$$\prod_{p \leq m} p = \prod_{p \leq k+1} p \prod_{k+2 \leq p \leq 2k+1} p \tag{46}$$

$$\leq 2^{2(k+1)} \prod_{k+2 \leq p \leq 2k+1} p \tag{47}$$

We use our induction hypothesis here as $k+1 < m$. We will now use the last two bounds from equations 43 and 45:

$$\prod_{p \leq m} p \leq 2^{2(k+1)} \binom{2k+1}{k} \tag{48}$$

$$\leq 2^{2(k+1)} \cdot 2^{2k} \tag{49}$$

$$\leq 2^{2(2k+1)} = 2^{2m} \tag{50}$$

$$\prod_{p \leq m} p \leq 2^{2m} \qquad \square$$

This means that our inequality indeed holds true:

$$\prod_{\sqrt{2n} < p \leq \frac{2}{3}n} p \leq \prod_{p \leq \frac{2}{3}n} p \leq 2^{\frac{4}{3}n}. \tag{51}$$

The third product follows quite quickly, as for p prime and $\frac{2}{3}n < p \leq n$ we know that $\nu_p(2n!) = 2$ and that $\nu_p(n!) = 1$, so $\nu_p(\binom{2n}{n}) = 0$. This means that $\prod_{\frac{2}{3}n < p \leq n} p^{\nu_p(\binom{2n}{n})} = \prod_{\frac{2}{3}n < p \leq n} p^0 = 1$.
The fourth and final product happens to be the core of our prove, which we assume to be the empty product; an assumption which will result in a falsehood.
We can now replace all product with previously discussed inequalities:

$$\binom{2n}{n} = \prod_{p \leq \sqrt{2n}} p^{\nu_p(\binom{2n}{n})} \prod_{\sqrt{2n} < p \leq \frac{2}{3}n} p^{\nu_p(\binom{2n}{n})} \prod_{\frac{2}{3}n < p \leq n} p^{\nu_p(\binom{2n}{n})} \prod_{n < p \leq 2n} p^{\nu_p(\binom{2n}{n})} \tag{52}$$

$$\leq (2n)^{\sqrt{2n}} \cdot 2^{\frac{4}{3}n} \tag{53}$$

Now that we have nothing but closed-form expressions we can check when this inequality fails. As we pointed out in the beginning we wish to prove an upper bound for this binomial to be less than a lower bound. We will use the following lower bound to manipulate the inequality:

$$2^{2n} = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} \leq \sum_{i=0}^{2n} \binom{2n}{n} = (2n+1)\binom{2n}{n}$$

Which will give us our last inequality to work with:

$$\frac{2^{2n}}{2n+1} \leq \binom{2n}{n} \leq (2n)^{\sqrt{2n}} \cdot 2^{\frac{4}{3}n} \tag{54}$$

$$\frac{2^{2n}}{2n+1} \leq (2n)^{\sqrt{2n}} \cdot 2^{\frac{4}{3}n} \tag{55}$$

$$2^{\frac{2}{3}n} \leq (2n)^{\sqrt{2n}}(2n+1) \tag{56}$$

$$2^{\frac{2}{3}n} < (2n)^{\sqrt{2n}}(2n)^2 \tag{57}$$

$$\tag{58}$$

In the step above we used the fact that $(2n)^2 > 2n+1$ for $n \geq 1$. In the following line we will use a similar result but to show that for $n \geq 18$ we have that $\frac{1}{3}\sqrt{2n} \geq 2$, which will be applied to the exponent of the last $(2n)$:

$$2^{\frac{2}{3}n} < (2n)^{\sqrt{2n}}(2n)^{\frac{1}{3}\sqrt{2n}} \tag{59}$$

$$2^{2n} < (2n)^{4\sqrt{2n}} \tag{60}$$

$$2^{\sqrt{2n}} < \sqrt{2n}^8 \tag{61}$$

This final inequality definitely fails for $\sqrt{2n} \geq 2^6$ as that results in $2^{(2^6)} = 2^{64} \leq (2^6)^8 = 2^{48}$. As the left-hand-side grows quicker than the right-hand-side, this means we have proven the theorem for $n \geq 2^{11} = 2048$. Note that this also justifies the choice to use the inequality $\frac{1}{3}\sqrt{2n} \geq 2$ for $n \geq 18$.
For n less than 2048 we can simply see that the set of primes $\{2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503\}$ is enough to prove the case for $n < 2047$, as each prime in this set is less than or equal to double the previous prime. This means that for all $n \in \mathbb{Z}_{\geq 1}$ there exists a prime such that $n < p \leq 2n$. □

3 Wooley's article

In his article [2], Wooley discusses two alternate iterative methods to generate not only an infinite amount of prime numbers, but all prime numbers. This article is aptly named "A Superpowered Euclidean Prime Generator."

Theorem 3.1 (Theorem) *Let $\pi_1 = 2$ and for $k \geq 1$, define π_{k+1} to be the least prime divisor of $n^{n^n} - 1$, where $n = \prod_{i=1}^{k} \pi_i$. Then for all k, π_k will be the k-th smallest prime.*

In order to prove this theorem, we will first prove the following:

Lemma 3.2 (Euler's congruence theorem) *Let $M \in \mathrm{Z}_{\geq 2}$. Then for all $a \in \mathrm{Z}$ with $\gcd(a, M) = 1$:*

$$a^{\phi(M)} \equiv 1 \mod M \tag{62}$$

If the reader wishes to refresh their knowledge on $\phi(M)$ and congruence classes, this can be found under Section 2. This Lemma follows directly from 2.1.2, however the following proof is surprisingly elegant and worth mentioning.

Proof: Consider $(\mathbb{Z}/M\mathbb{Z})^\times$ with the representatives of the congruence classes $b_1, b_2, \ldots, b_{\phi(M)}$. Because $\gcd(a, M) = 1$, we know that a is invertible. This means that for all $1 \leq i \leq \phi(M)$, $a \cdot b_i$ again is invertible modulo M. So if we multiply all representatives of $(\mathbb{Z}/M\mathbb{Z})^\times$ with a, we will still be left with $(\mathbb{Z}/M\mathbb{Z})^\times$ but now with representatives $ab_1, ab_2, \ldots . ab_{\phi(M)}$. Because these represent the same congruence classes modulo M, the product of these multiplications will also be equivalent modulo M:

$$\prod_{i=1}^{\phi(M)} b_i \equiv \prod_{i=1}^{\phi(M)} ab_i \mod M \tag{63}$$

$$\equiv a^{\phi(M)} \prod_{i=1}^{\phi(M)} b_i \mod M \tag{64}$$

Because for all i these $b_i \in (\mathbb{Z}/M\mathbb{Z})^\times$, we know that the products are unequal to zero and therefore we can conclude from this that $1 \equiv a^{\phi(M)} \mod M$. Again, for our sake we are more interested in M prime, so:

Corollary 3.2.1 *For p prime, we have:*

$$1 \equiv a^{p-1} \mod p$$

For any a not a multiple of p.

One might be familiar with this result as it is also known as Fermat's little theorem.

In order to prove Theorem 3.1, we have to show that the smallest prime π that doesn't divide n does divide $n^{n^n} - 1$. We can rewrite this to the equivalence $n^{n^n} \equiv 1 \mod \pi$. Our goal here is to prove that $\pi - 1$ divides n^n, as we can then rewrite n^n to be $\lambda(\pi - 1)$ for some $\lambda \in \mathbb{Z}_{\geq 0}$ and apply Corollary 3.2.1.

Proof of Theorem 3.1: Let π be the smallest prime that doesn't divide n, where n was defined as $n = \prod_{i=1}^{k} \pi_i$ with π_i the i-th smallest prime. As π is the smallest prime to not divide n, all prime divisors of $\pi - 1$ must also divide n. Similarly to $n = \pi_1 \cdot \pi_2 \cdots \pi_k$, $\pi - 1$ must have its representation as product of primes as $\pi - 1 = \pi_1^{r_1} \cdot \pi_2^{r_2} \cdots \pi_k^{r_k}$, where $r_k \in \mathbb{Z}_{\geq 0}$ represents the multiplicity of that prime factor. For n^n, the prime representation would be $n^n = \pi_1^n \cdot \pi_2^n \cdots \pi_k^n$. As we stated above, our goal is to show that $\pi - 1$ divides n^n. This can be done by considering the multiplicity of the prime factors of $\pi - 1$; if for all i, or all prime factors of $\pi - 1$, we have $r_i \leq n$ then $\pi - 1$ must divide n^n, with

$$\frac{n^n}{\pi - 1} = \pi_1^{n-r_1} \cdot \pi_2^{n-r_2} \cdots \pi_k^{n-r_k} \tag{65}$$

where this fraction of course is a whole number if $n \geq r_i$ for all i. Now obviously, for any $k, n \geq 2$ we have $k^n \geq n + 1$. As all prime numbers are greater than or equal to 2, we have $\pi_i^n \geq n + 1 \geq \pi - 1$. This shows us that for all prime divisors π_i of $\pi - 1$, $r_i \leq n$ so we can indeed conclude that $\pi - 1$ divides n^n. Consequently we may write that $n^n = \lambda(\pi - 1)$ for some $\lambda \in \mathbb{Z}_{\geq 0}$. Now, what we wanted to check was if the equivalence $n^{n^n} \equiv 1 \mod \pi$ holds:

$$n^{n^n} \equiv n^{\lambda(\pi-1)} \equiv (n^\lambda)^{\pi-1} \mod \pi \tag{66}$$

Now as π is a prime number not dividing n by construction, 3.2.1 indeed shows that the above is equivalent to $1 \mod \pi$, so the smallest prime number that does not divide n does divide $n^{n^n} - 1$. Now as n already was the product of the k smallest primes, this new prime number π must be the $k + 1$-th smallest prime. □

As one may have noted, in this iterative process we need knowledge of the prime factorization of n. However, for computational or theoretical reasons we could also wonder if we could identify a prime with any given lower bound $N \in \mathbb{Z}_{\geq 0}$. Of course, a very inefficient way to use our previously proven theorem is to choose n to be $N!$, then the smallest prime number to divide $N!^{N!^{N!}} - 1$ will be the smallest prime number to exceed N. In the article[2] a refinement to the above was given as a comment by Andrew Booker and Andrew Granville, which stated:

Theorem 3.3 (Refinement) *Let N be any whole number greater than 1. Then the smallest prime to exceed N is the least prime divisor of $N!^{N!} - 1$*

As this theorem was merely stated and not proven in this article, we have given our own proof.

3.1 Our proof of the zero-knowledge refinement

Our aim is to prove that the smallest prime π to exceed N divides $N!^{N!} - 1$. Analogous to the proof of Theorem 1.2, we will rewrite this to the equivalence:

$$N!^{N!} \equiv 1 \mod \pi \tag{67}$$

Again, we will use Corollary 3.2.1 and prove that $\pi - 1$ divides $N!$ as we can then rewrite our equivalence to $N!^{\lambda(\pi-1)} \equiv 1 \mod \pi$ (for $N > 3$). For $N = 2, 3$ we will simply skip this Lemma and check the theorem by hand:

$$2!^{2!} - 1 = 2^2 - 1 = 3 \tag{68}$$

which checks out as 3 is indeed the smallest prime to exceed 2.

$$3!^{3!} - 1 = 6^6 - 1 = 35 = 5 \cdot 7 \tag{69}$$

which also checks out as 5 is indeed the smallest prime to exceed 3. In order to prove this for $N \geq 4$ we will use the following Lemma:

Lemma 3.4 (Bertrand-Chebyshev Theorem) *For all $n \in \mathbb{Z}_{\geq 1}$, there exists a prime number p with $n < p \leq 2n$.*

We will prove 3.3 for different cases considering the factorization of $\pi - 1$. Firstly, the case when $\pi - 1$ is divisible by more than one prime. Let p be any prime number that divides $\pi - 1$. From Lemma 3.4, we know that $\pi - 1 < 2N$. From this we can deduce that $\frac{\pi-1}{p} < \frac{2}{p}N$. This means that both $\frac{\pi-1}{p}$ and p are less than N and thus are both factors of $N!$, meaning $\pi - 1$ divides $N!$ if $\pi - 1$ has more than one prime number in its factorization. The only cases left to prove now are when $\pi - 1$ has only one prime number in its factorization, or: there exists a $q \in \mathbb{Z}_{\geq 0}$ prime such that for some $k \in \mathbb{Z}_{\geq 2}$, $\pi - 1 = q^k$. We can rule out $k = 1$ as we know that π must be odd and greater than 3. If we are to observe the parity of this equations, i.e. the equivalence $\mod 2$, we see that the left-hand-side $\pi - 1$ has to be even. The right-hand-side must therefore also be even, which implies that q must be 2. Now we distinguish another two cases, $k > 2$ and $k = 2$.

The first case, $k > 2$. From our assumptions we know that $q^k = \pi - 1$. Similarly as before, $q^{k-1} = \frac{\pi-1}{q} < \frac{2}{q}N < N$. Moreover, as $k > 2$ and $q > 1$ we know $q \neq q^{k-1}$ and again both of these are less than N, which means that they both are factors of $N!$, so this case too follows our theorem. Now the case when $k = 2$. Then, $\pi - 1 = 2^2 = 4$ and $\pi = 5$. As we stated that π must be the smallest prime to exceed N, and as $N \geq 4$ we see the only possible value for this case to occur is when $N = 4$. This gives us $N! = 24$, which is divisible by $\pi - 1 = 4$ so we can also check this one by hand:

$$\begin{aligned} N!^{N!} &\equiv 4!^{4!} && \mod \pi && (70)\\ &\equiv 24^{6 \cdot 4} && \mod \pi && (71)\\ &\equiv \left(24^6\right)^{\pi-1} && \mod \pi && (72)\\ &\equiv 1 && \mod \pi && (73) \end{aligned}$$

This means that for any number N we indeed find that the smallest prime to exceed N, $\pi - 1$ does divide $N!$, which means that we can use 3.2.1 to prove the theorem. □

4 Prime generating sequences

As the infinitude of prime numbers is so important, many different proofs of this Fundamental Theorem are known. Many of these follow in the steps of Euclid, i.e. given a finite set, or seed, of prime numbers, show we can construct a number coprime to all prime numbers in our seed. As it turns out such variations on Euclid's proof have been the inspiration for the upcoming generating sequences. The previous generator by Wooley also uses this fact as any prime to divide n is obviously coprime to $n^{n^n} - 1$ and has the advantage of generating all primes in order. In general such constructions generate infinitely many (co)prime numbers,

however whether such Euclidean Sequences[1] contain *all* prime numbers is a natural question that arises and was first posed by Mullin in [16].

4.1 Sequences analogous to Euclid's proof

There are three sequences which most resemble Euclid's proof which we will focus on in this section. These sequences are the first and second Euclid-Mullin sequences and Sylvester's sequence.

4.1.1 The first Euclid-Mullin sequence

Recall that the first Euclid-Mullin sequence is defined as follows: let $p_1 = 2$ and let p_{k+1} be the smallest divisor of $P_k = 1 + \prod_{i=1}^{k} p_i$. In 1974 the first 9 terms were known through computation by Guy and Nowakowski [15] and it would be Shanks in 1991 who conjectured that this sequence would contain every prime [18]. He did this with the following heuristic argument; consider the product of all primes in our sequence modulo p for some p prime not yet in our sequence. As more terms get added to our sequence, our product should attain random values modulo p. Once our product is equivalent to -1 modulo p for some n we know that p then divides P_n and therefore must occur in the sequence. If however p were to be a prime such that it never occurs in our sequence, the product will never be equivalent to -1 which would not follow the randomness of the product modulo p. As it indeed seems that the values modulo p are random, it is compelling to believe this heuristic argument. More terms have been calculated since then with a lot of credit to Wagstaff [17], not only for the first Euclid-Mullin sequence but also for the second and the two sequences most akin to the Euclid-Mullin sequences where the leading $+1$ is substituted by -1.

4.1.2 The second Euclid-Mullin sequence

The second sequence is similar to the first, except that for the choice of p_{k+1} instead the largest divisor is chosen. In 1968 the question whether all prime numbers appear was answered negatively; Cox and van der Poorten had been able to show in [23] that $2, 3, 7$ and 43 are the only primes less than 53 not omitted from this sequence. After Naur showed in [21] that the sequence is not monotone increasing (he calculated that $p_{10} < p_9$) the question whether infinitely many primes would be omitted remained open and would be answered in the affirmative by Booker in 2013 [8]. This proof has its core in finding an upper bound for an omitted prime q_n in terms of all primes less than q_n also omitted. An alternate proof to Booker's can be found in [12], where Pollack and Treviño presented a proof less dependent on analytic number theory and more based around the distributions of (non)quadratic residues.

4.1.3 Alternative results

The two other sequences of which Wagstaff [17] had calculated multiple terms are very similar to the Euclid-Mullin sequences; let $r_1 = 3$ and let r_{k+1} be some prime factor of $\prod_{i=1}^{k} r_i - 1$. For the variant where the smallest prime divisor is chosen each step it is also conjectured it will contain all primes. For the variant where the choice of prime is the largest prime divisor, thanks to Selfridge[2] it is known that some primes are omitted, however the question whether infinitely many primes are omitted remains unanswered.

[1] The term Euclidean Sequences is used here following Ribenboim [19] to denote sequences closely resembling Euclid's proof.
[2] This was never published but mentioned by Guy and Nowakowski in [15]

4.1.4 Sylvester's sequence

The final sequence is exactly Euclid's construction, i.e. not all numbers in the sequence are prime. Instead of asking whether all primes occur in the sequence, one could ask whether all prime numbers divide one of the numbers in this sequence. Even though we essentially don't exclude any primes at any step, Odoni had already showed in 1985 in [27] that infinitely many primes do not divide any of the a_n.

Theorem 4.1 *Let $a_1 = 2$ and for $n \geq 1$*

$$a_{n+1} = 1 + \prod_{i=1}^{n} a_i \tag{74}$$

then any prime p that divides some a_n must be equivalent to 1 modulo 6.

Proof: The first step in our proof is to see that our Sylvester's Sequence can also be defined as the recurrence relation

$$a_{n+1} = a_n^2 - a_n + 1 \tag{75}$$

and let q be an odd prime such that $q \mid a_{n+1}$, then we have the following equivalence:

$$a_n^2 - a_n + 1 \equiv 0 \mod q \tag{76}$$
$$4a_n^2 - 4a_n + 4 \equiv 0 \mod q \tag{77}$$
$$(2a_n - 1)^2 \equiv -3 \mod q \tag{78}$$

This means that for q to divide a_{n+1}, $\left(\frac{-3}{q}\right) = 1$, so by multiplicity of Legendre's symbol, $\left(\frac{-1}{q}\right)\left(\frac{3}{q}\right) = 1$. This can only happen if both are 1 or both -1. For the first, we know that $\left(\frac{-1}{q}\right) = 1$ if and only if $q \equiv 1 \mod 4$. Now we use quadratic reciprocity to see that $1 = \left(\frac{3}{q}\right) = \left(\frac{q}{3}\right)$, which only occurs when $q \equiv 1 \mod 3$. So when $q \equiv 1 \mod 12$ we know that $\left(\frac{-3}{q}\right) = 1$.

For the second case, when both Legendre symbols are -1, we now know that $\left(\frac{-1}{q}\right) = -1$ and therefore that $-1 = \left(\frac{3}{q}\right) = -\left(\frac{q}{3}\right)$, again by quadratic reciprocity but now both q and 3 are $3 \mod 4$ so we see a change of sign. Just as before we must have $\left(\frac{q}{3}\right) = 1$ so $q \equiv 1 \mod 3$, so in this case for $q \equiv 7 \mod 12$ we have that $\left(\frac{-3}{q}\right) = 1$. Combining the two gives us that all primes that divide any a_n must be of the form $6k + 1$. This concludes the proof of the theorem. □

From this we can also see that any prime of the form $6k + 5$ divides none of the a_n and by Dirichlet's theorem on arithmetic progression we know that there must be infinitely many primes of this form, which is sufficient to show that there are infinitely many primes which divide none of the a_n.

Guy and Nowakowski had already shown this in 1974 [15], however they went one step further. As Sylverster's Sequence is recursive, the values of $a_n \mod p$ for any prime p must be periodical. They used this fact to identify that the only primes less than 1000 to divide any a_n are

$$7, 13, 43, 73, 139, 181, 547, 607. \tag{79}$$

Vardi [10] has contributed to this specific sequence by finding all primes less than 2×10^8 to divide some a_n for $n \leq 200$.

4.2 Alternate sequences of coprime numbers

As mentioned before there are proofs which vary from Euclid's proof but with the same principle of generating coprime numbers. In this section we discuss some specific kinds of (prime) numbers and whether they can or have been used to construct such a sequence of coprime numbers.

4.2.1 Factorial

Looking back at Wooley's construction to find the smallest prime larger than a given N, one can also instead consider $N! + 1$ for a more general construction. Obviously, $N! + 1$ is coprime to all numbers less than or equal to N. Here we discard the property to generate the smallest next prime however we don't have to deal with the explosiveness of $N!^{N!}$. One could ask too if all primes divide some $a_N = N! + 1$ and the answer in this case is yes. This follows directly from Wilson's theorem:

Theorem 4.2 (Wilson's theorem) *For every prime p the equivalence $(p-1)! \equiv -1 \mod p$ holds true.*

Proof: Consider the congruence classes in $(\mathbb{Z}/p\mathbb{Z})^\times$. Every one of these classes has its unique inverse, so we can pair up factors in $(p-1)!$. However, only 1 and $p-1$ don't pair up with any other as they are their own inverses. This means that $(p-1)! \equiv 1 \cdot (p-1) \equiv -1 \mod p$. □

Therefore for every prime p, we see that they must divide $a_{p-1} = (p-1)! + 1$. This sequence however does not have the property that all elements are coprime, i.e. $N! + 1$ need not be coprime to $n! + 1$ for some $N, n \in \mathbb{Z}_{>0}$. For a sequence of coprime numbers based on factorials, we thought of our own construction. Let $a_1 = 2$ and for $k \geq 1$ let $a_{k+1} = d! + 1$, where d is the largest prime factor of a_k. Then all a_k are coprime. This sequence too has an absurd growth rate, as $a_5 = 71! + 1$, a number with over a hundred digits. Sadly this sequence is not very interesting as it leaves out quite a lot of primes, whichever way you look at it.

Another question considering such factorials is whether there are infinitely many numbers N for each of the forms $N! \pm 1$ such that they are prime. The answer to this, too, is still unknown and too conjectured to be true by Chris Caldwell and Yves Gallot.[22]

4.2.2 Primorial

Not only did Caldwell investigate such factorial primes, he also investigated primorial primes, so primes of the form $p\# \pm 1$. This again closely resembles Wooley's method but now for $n^{n^n} - 1$. Because his method found the k-th smallest prime in the k-th step, n became the product of the first k primes, i.e. $n = p_k\#$. For this too Caldwell and Gallot conjectured that there are infinitely many primes of each of the forms $p\# \pm 1$. Interestingly, it is also unknown but suspected that there are infinitely many composite numbers of the form $p\# \pm 1$.

4.2.3 Pomerance's variant

While Pomerance's variant differs slightly from Euclid's method as all the choices in each step need not be coprime to all previous primes in the starting seed, it is worth mentioning as it provably generates all prime numbers, in order from a certain point. Let $p_1 = 2$. We choose p_{k+1} as follows. Let $n = \prod_{i=1}^{k} p_i$. Then we choose p_{k+1} as the smallest prime to divide $d+1$ for some $d \mid n$. As the choice for d is very liberal Pomerance

proved but never published that for $k \geq 5$, p_k is the k-th smallest prime. One can see that this d indeed need not be coprime to all p_j for $0 \leq j \leq k$.

In [11] Crandall and Pomerance swiftly pointed out another sequence which is a slight restriction on Pomerance's variant, raised by M. Newman from the Australian National University. Let $p_1 = 2$ and $p_2 = 3$, and choose p_{k+1} as the least prime not yet in our sequence to divide $p_i p_j + 1$ for some $1 \leq i < j \leq k$. This sequence has yet to be tackled, from some minor computations we see that e.g. 19 and 5 do not appear in order, however whether this sequence generates infinitely many primes is still unknown. It is even unknown but highly suspected whether this sequence is even infinite. The first few terms of this sequence are:

$$2, 3, 7, 5, 11, 13, 17, 23, 29, 19, 31, 37, 41, 43, 47, 53, 59, \ldots \tag{80}$$

4.2.4 Stieltje's proof

Stieltje's proof for the infinitude of primes is as follows.
Proof: Assume that $p_1, \ldots, p_k$ is a finite set of all the primes, and let $N = p_1 \cdot p_2 \cdots p_k$ be their product. Choose any factorization for $N = d \cdot d'$, then every prime p_i divides either d or d'. However, then $m+n$ is not divisible by any of the p_i, which contradicts the assumption that the number of primes is finite, as $m+n > 1$. Again, this proof for the infinitude of primes can be used to generate a sequence of prime numbers which is exactly what Chua's sequence entails.

Booker has shown in [6] that a sequence constructed with the method of Stieltje's proof does generate all primes, the proof of which we will discuss in full.

4.3 A variant of the Euclid-Mullin sequence containing every prime

Using Stieltje's proof as a baseline, Chua's sequence is as follows. With $P = \{p_1, \ldots, p_k\}$ a finite set of primes, we now choose p_{k+1} as the smallest prime to divide

$$N_I = \prod_{i \in I} p_i + \prod_{i \in \{1,\ldots,k\} \setminus I} p_i \tag{81}$$

where $I \subseteq \{1, \ldots, k\}$, i.e. we consider all $I \subseteq \{1, \ldots, k\}$ with their subsequent N_I and choose p_{k+1} as the smallest prime to divide some N_I. Chua's sequence starts with the seed $P = \{2\}$ to give us the following sequence:

$$2, 3, 5, 11, 37, 13, 7, 29, 17, 19 \tag{82}$$

It can be easily seen that this construction allows for a more steady growth and leaves out fewer primes than the first Euclid-Mullin sequence, as the original Euclid-Mullin sequence has a 14-digit-number as its 9th element. Chua considered the sequence where the choice of p_{k+1} is fixed as the smallest prime divisor, however allowing for the choice of any prime divisor of any N_I to be p_{k+1} we obtain what we will refer to from now on as a generalized Euclid sequence. The freedom of choice in the Euclid-Mullin sequence only applied to the choice of which prime divisors we select as p_{k+1}, whereas the generalized Euclid sequences grant us more freedom for choosing our next prime p_{k+1} as the number of choices for I and thus N_I is proportional to $\mathcal{P}(I)$, where $\mathcal{P}$ is the powerset. This freedom proves to be sufficient that this construction will eventually contain every prime.

Theorem 4.3 *For any finite set P of prime numbers, there exists a generalized Euclid sequence with seed P containing every prime.*

We will examine the following set $S_q = \left\{ d + q\mathbb{Z} : d \in \mathbb{Z}_{>0}, d \mid \prod_{p<q} p \right\}$. This set is a subset of $(\mathbb{Z}/q\mathbb{Z})^\times$, more precisely the residue classes of squarefree, $(q-1)$-smooth, positive integers. We will use the fact that this set is large, large enough that it is impossible to avoid q showing up in this sequence. Preferably we would like S_q to be equal to the multiplicative residue classes, however it is sufficient to have that $\#S_q > \frac{1}{2}(q-1)$.

Lemma 4.4 *For any prime q, $\#S_q > \frac{1}{2}(q-1)$.*

In article [28] Kenneth Roger has shown that this lower bound can actually be set to $\frac{53}{88}(q-1)$, however for our proof $\frac{1}{2}(q-1)$ will suffice.

Lemma 4.5 *For q an odd prime and $a \in (\mathbb{Z}/q\mathbb{Z})^\times$*

(i) If $q = 5$ or $q \neq 5$ and $a = 3 + 5\mathbb{Z}$ then there exists an $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $\left(\frac{x+a/x}{q}\right) \neq 1$

(ii) If $q \notin \{7, 13\}$, then there exist $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $\left(\frac{x^6+a}{q}\right) \neq 1$

Proof of (i): If it were the case that such x doesn't exist, then that would mean that for all $x \in (\mathbb{Z}/q\mathbb{Z})^\times$, $\left(\frac{x+a/x}{q}\right) = 1$. Then the following identical sum would be:

$$\sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x + a/x}{q}\right) = \#(\mathbb{Z}/q\mathbb{Z})^\times = q - 1 \tag{83}$$

So our goal to prove (i) will be to show that the equality above will not hold but instead that the sum is less than $q-1$. Note that $\left(\frac{x^2}{q}\right)$=1, so that we instead can consider the following sum:

$$\sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x + a/x}{q}\right) = \sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x(x^2 + a)}{q}\right) \tag{84}$$

The Legendre symbol in this context can also be interpreted to be 1 when there exists a solution to the curve $y^2 = x(x^2 + a)$, with $(x, y) \in (\mathbb{Z}/q\mathbb{Z})^2$. This curve happens to be an elliptic curve with one point at infinity. This means we can use the Hasse-bound for the number of affine points on an elliptic curve if this curve has no repeated root $\mod q$. Recall that for a repeated root its derivative will have the same root. Let $f(x) = x(x^2 + a)$, then $f'(x) = 3x^2 + a$. See that the roots of $f(x)$ are $x = 0$ and x such that $x^2 = -a$. Note that these values are not roots for $f'(x)$ for $\mod q$ when $q \geq 3$. So indeed we have no repeated roots and we can use the Hasse-bound[3]:

$$|N - q| \leq 2\sqrt{q} \tag{85}$$

where N is the number of solutions to our curve. We will use the following resulting bound $N \leq q + 2\sqrt{q}$. We can translate our sum to number of solution as follows; if for x we find that $\left(\frac{x(x^2+a)}{q}\right) = 1$, then there exists a solution y such that $y^2 = x(x^2 + a)$. As q is an odd prime, y and $-y$ are two unique solutions to this equation. If instead we find $\left(\frac{x(x^2+a)}{q}\right) = 0$ then that means that $q \mid x(x^2 + a)$ and thus $q \mid y$. In $(\mathbb{Z}/q\mathbb{Z})$ however this gives us exactly one solution, namely $y = q\mathbb{Z}$. Lastly, if $\left(\frac{x(x^2+a)}{q}\right) = -1$ then there exist no solution. The number of solutions for any $x \in (\mathbb{Z}/q\mathbb{Z})$ therefore must equal $\left(\frac{x(x^2+a)}{q}\right) + 1$ and so the total

number of solutions N can be written as:

$$N = \sum_{x\in(\mathbb{Z}/q\mathbb{Z})} \left[\left(\frac{x(x^2+a)}{q}\right)+1\right] \leq 2\sqrt{q}+q \tag{86}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})} \left[\left(\frac{x(x^2+a)}{q}\right)+1\right] \leq 2\sqrt{q}+q \tag{87}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x(x^2+a)}{q}\right) + \sum_{x\in(\mathbb{Z}/q\mathbb{Z})} 1 \leq 2\sqrt{q}+q \tag{88}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x(x^2+a)}{q}\right) \leq 2\sqrt{q} \tag{89}$$

After equation 87 we switched from $x \in (\mathbb{Z}/q\mathbb{Z})$ to $x \in (\mathbb{Z}/q\mathbb{Z})^\times$. This leaves out $x = q\mathbb{Z}$ but in this case $\left(\frac{x(x^2+a)}{q}\right) = \left(\frac{x}{q}\right)\left(\frac{x^2+a}{q}\right) = 0$, so it doesn't influence the sum. Our goal was to show that the left-hand-side is less than $q-1$, and for its upper bound $2\sqrt{q}$ this is true for $q \geq 7$. For $q \in \{3,5\}$ it can be checked directly that $\left(\frac{x(x^2+a)}{q}\right) \neq 1$ for some $x \in (\mathbb{Z}/q\mathbb{Z})^\times$. This concludes the proof of (i).
For (ii) the proof is the same but for the fact that we no longer have an elliptic curve but a curve of genus 2, with two points at infinity, so instead we can use the Weil bound[3]:

$$N + 1 - q \leq 2g\sqrt{q} \tag{90}$$

where g is the genus of the curve. To apply this bound we again need that our function $f(x) = x^6 + a$ has no repeated roots $\mod q$. We will check this with its derivative $f'(x) = 6x^5$. Note that this function has roots $\mod q$ only for $x = 0$, or when $6 \equiv 0 \mod q$, as then $f'(x)$ becomes the zero function. For $x = 0$ we see that this is not a zero of $f(x)$ as $a \in (\mathbb{Z}/q\mathbb{Z})^\times \not\ni 0$ so we are not dealing with a repeated root. The case when $6 \equiv 0 \mod q$ only happens for $q \in \{2,3\}$ so for $q \geq 5$ our function f has no repeated roots. This means we can indeed apply the Weil bound:

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})} \left[\left(\frac{x^6+a}{q}\right)+1\right] \leq 4\sqrt{q}+q-1 \tag{91}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x^6+a}{q}\right) + \left(\frac{a}{q}\right) + \sum_{x\in(\mathbb{Z}/q\mathbb{Z})} 1 \leq 4\sqrt{q}+q-1 \tag{92}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x^6+a}{q}\right) \leq 4\sqrt{q}-1-\left(\frac{a}{q}\right) \tag{93}$$

$$\sum_{x\in(\mathbb{Z}/q\mathbb{Z})^\times} \left(\frac{x^6+a}{q}\right) \leq 4\sqrt{q} \tag{94}$$

Just as before we left $x = q\mathbb{Z}$ out of the sum, which now affects our equation with the addition of $\left(\frac{a}{q}\right)$. The bound on the right-hand-side now is less than $q-1$ for $q \geq 19$ and just as before for $q \in \{3,5,11,17\}$ it can be checked directly. This concludes the proof of (ii) and with that 4.5. □

From here on we let $n = \prod_{p\in P} p$, where $P = \{p_1, \ldots, p_k\}$. It can be seen empirically that construction 81 doesn't necessarily generate primes in order, i.e. p_{k+1} is not the smallest prime not in P. As S_q is $(q-1)$-smooth it does not cover all our options for N_I, so we shall consider the following set S instead:

$$S = \{d + q\mathbb{Z} : d \in \mathbb{Z}_{>0}, d \mid n\} \tag{95}$$

where $S_q \subseteq S \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$. For future reference it will be easier to consider N_I as follows:

$$N_d \coloneqq N_I = d + \frac{n}{d} \tag{96}$$

where $d = \prod_{i \in I} p_i$. Our aim is to show that for any finite set of primes P not containing q, we will encounter an N_d such that $q \mid N_d$, or equivalently that $d + n/d \equiv 0 \mod q$. If $q = 2$, then both n and d will be odd, so that $d + n/d$ must be even and so we can set p_{k+1} to be q. Thus we may assume q to be an odd prime. Firstly we will prove that a $d \in S$ can be chosen such that $d + n/d \equiv 0 \mod q$ in the case that $S = (\mathbb{Z}/q\mathbb{Z})^\times$. Important to see is that we can rewrite the equivalence to:

$$d + n/d \equiv 0 \mod q \tag{97}$$

$$d \equiv -\frac{n}{d} \mod q \tag{98}$$

$$d^2 \equiv -n \mod q \tag{99}$$

which means we can instead consider the quadratic character of $-n$, or $\left(\frac{-n}{q}\right)$. If $\left(\frac{-n}{q}\right)$ happens to be 1, that means that there exists an $r \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $r^2 \equiv -n \mod q$. As $S = (\mathbb{Z}/q\mathbb{Z})^\times$, there exists a $d \equiv r \mod q$ and therefore $q \mid d + n/d$, our desired result. On the other hand, the case could be that $\left(\frac{-n}{q}\right) = -1$. By Lemma 4.5(i) we know that not all of $\left(\frac{d+n/d}{q}\right)$ can equal 1 and thus that there exists a d such that $\left(\frac{d+n/d}{q}\right) = -1$ given that $q \neq 5$ or that $n \neq 3 + 5\mathbb{Z}$. However if $\left(\frac{d+n/d}{q}\right) = -1$ then there must be a prime p in the prime factorization of $d + n/d$ such that $\left(\frac{p}{q}\right) = -1$. As we set out to prove that q would eventually pop up in our sequence, we are free to choose p_{k+1} as the aforementioned p. This means we find ourselves a new $n' = pn$, for which

$$\left(\frac{-n'}{q}\right) = \left(\frac{p}{q}\right)\left(\frac{-n}{q}\right) = (-1)(-1) = 1 \tag{100}$$

This result lands us back in the case that $\left(\frac{-n'}{q}\right) = 1$ which we have proven above. [3] What's left to prove for $S = (\mathbb{Z}/q\mathbb{Z})^\times$ is the case that $q = 5$ and that $n = 3 + 5\mathbb{Z}$, which can be resolved through the same method, namely to choose d to be 1. Due to the fact that $n \equiv 3 \mod 5$ and therefore $n + 1 \equiv -1 \mod 5$ there exists a $p \mid (n + 1)$ for which $p \not\equiv -1 \mod 5$ and by again choosing $n' = pn$ we no longer have that $n' \equiv 3 \mod 5$ which feeds back to what we have proven above. Note that this p is necessarily coprime to n as $d = 1$ so $N_d = n + 1$ and as $n = \prod_{\pi \in P} \pi$ we find that we can indeed choose $p_{k+1} = p$.

Secondly rests us the case that $S \neq (\mathbb{Z}/q\mathbb{Z})^\times$. Just as above we must choose intermediary primes to add to our set P. We will define the set of primes from which we can choose our new prime $p = p_{k+1}$ as follows:

$$T = \{p : p \text{ prime and } p \mid (d + n/d) \text{ for some } d \mid n\}. \tag{101}$$

Having chosen this p, we thus replace P, n and S by $P \cup \{p\}$, pn and $S \cup \{pS\}$ respectively. As $S \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ is finite there are two cases upon which this procedure can come to a standstill; $q \in T$ for which we can choose $q = p_{k+1}$ and are done, or that S can't be increased any more, meaning that for every $p \in T$ our $pS \subseteq S$ and thus $S = S \cup pS$. In this second case we can see that S must contain sG, where $s \in S$ and G is the subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$ generated by:

$$G \coloneqq \{p + q\mathbb{Z} : p \in T\}. \tag{102}$$

This means that S must be a union of cosets sG and that $\#G \mid \#S$. In order to conclude something about $\#G$ it is useful to first look at subgroups of $(\mathbb{Z}/q\mathbb{Z})^\times$ in general, in particular their index.

[3] Observant reader might notice that in [6] the author instead wrote $\left(\frac{-n'}{p}\right)$; this is a typo.

Let H be a subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$ with an index of at least 4. For any $h \in H$ we know that the number of d such that $d + n/d = h$ must be less than or equal to 2 as d and $d' = n/d$ both have the same result. therefore we can relate the cardinality of the set of these d and H as follows:

$$\#\{d \in (\mathbb{Z}/q\mathbb{Z})^\times : d + n/d \in H\} \leq 2\#H \leq \frac{1}{2}(q-1), \tag{103}$$

where the final inequality stems from the choice that the index of H is at least 4. We know from Lemma 4.4 that the number of $d + n/d$ is too large such that all $d + n/d \in H$, so there exists $d \mid n$ such that $(d + n/d) + q\mathbb{Z} \notin H$. This also implies that there exists $p \in T$ such that $p + q\mathbb{Z} \notin H$. As $\#(\mathbb{Z}/q\mathbb{Z})^\times = (q-1)$ and that the index of a subgroup must always divide the size of the supergroup, we can define specific subgroups through the prime representation of the size of $(\mathbb{Z}/q\mathbb{Z})^\times$:

$$H_r \coloneqq \{h \in (\mathbb{Z}/q\mathbb{Z})^\times : h^{\frac{q-1}{r}} = 1\} = \{x^r : x \in (\mathbb{Z}/q\mathbb{Z})^\times\}, \tag{104}$$

where $r \mid (q-1)$. Note that r is then index of h with respect to $(\mathbb{Z}/q\mathbb{Z})^\times$. This can be more intuitive with the fact that $(\mathbb{Z}/q\mathbb{Z})^\times$ is cyclic for q prime, so we can represent $(\mathbb{Z}/q\mathbb{Z})^\times$ with $\{a^m : a(\mathbb{Z}/q\mathbb{Z})^\times\}$ for some a. Then $H_r = \{x^r : x \in (\mathbb{Z}/q\mathbb{Z})^\times\}$ becomes $\{a^{rm} : a \in (\mathbb{Z}/q\mathbb{Z})^\times\}$, which now only attains unique values for $0 \leq m < \frac{q-1}{r}$ so indeed we see the index of H_r is r. Let $q - 1 = \prod_{i=1}^{m} r_i^{e_i}$ be the prime factorization of $q-1$ with $e_i = \nu_{r_i}(q-1)$. For $r_i \geq 5$ we find we can use the same reasoning as in 103 to show for

$$H = H_{r_i} = \{h \in (\mathbb{Z}/q\mathbb{Z})^\times : r_i^{e_i} \nmid \operatorname{ord}(h)\} \tag{105}$$

that there exist some prime $p \in T$ such that p has order divisible by $r_i^{e_i}$. This means that G, which was the subgroup generated by $\{p + q\mathbb{Z} : p \in T\}$, also has order divisible by $r_i^{e_1}$. For $r_i \in \{2, 3\}$ however the index of H_{r_i} would be less than 4, however we instead use the same argument as above for $H_{r_i^2}$ to see that the order of G is divisible by $r_i^{e_i - 1}$ for $r_i \in \{2, 3\}$. The index of G can be calculated as follows:

$$\frac{\#(\mathbb{Z}/q\mathbb{Z})^\times}{\#G}, \tag{106}$$

and as we've seen for all $r_i \geq 5$ these prime numbers divide both $\#(\mathbb{Z}/q\mathbb{Z})^\times$ and G, except for $r_i \in \{2, 3\}$. From this we can then conclude that the order of the subgroup G must divide 6. In the case that $q \neq 1 + 3\mathbb{Z}$, 3 does not divide the order of $(\mathbb{Z}/q\mathbb{Z})^\times$, which is $q - 1$. Therefore the index of G must divide 2 which implies that $\frac{1}{2}(q-1) \mid \#G \mid \#S$. As $\frac{1}{2} < \#S_q$ by Lemma 4.4, therefore $\frac{1}{2} < \#S$ which would imply that $\#S = (q-1)$. This however contradicts our assumption that $\#S \neq (\mathbb{Z}/q\mathbb{Z})^\times$, so we know this case, namely that $q \not\equiv 1 \mod 3$, cannot occur. In the case that $q \equiv 1 \mod 3$, we can use the same reasoning for H as above, however we now take $r = 6$ to see that there exists a prime $p \in T$ such that $p^{\frac{q-1}{6}} \not\equiv 1 \mod q$. As $\frac{q-1}{6} = \frac{q-1}{2} - \frac{q-1}{3}$, we can conclude from $p^{\frac{q-1}{6}} = p^{\frac{q-1}{2}} / p^{\frac{q-1}{3}}$ that at least one of H_2 and H_3 does not contain this p. If $p \notin H_3$ we find ourselves in the same situation as before when we assumed that $q \not\equiv 1 \mod 3$, namely that G has order divisible by 2 which again would lead to the conclusion that $\#S = q - 1$. This means we can assume that it is H_2 which does not include $p + q\mathbb{Z}$. Hence the index of G must divide 3. If the index of G were to be 1, then that would mean $G = (\mathbb{Z}/q\mathbb{Z})^\times$ which in turn would imply that $S = (\mathbb{Z}/q\mathbb{Z})^\times$, again contradicting our assumption that $S \neq (\mathbb{Z}/q\mathbb{Z})^\times$, so we can safely assume that the index of G is 3, i.e. $G = H_3$ and $\#G = \frac{1}{3}(q-1)$. Combining Lemma 4.4 with the fact that $S = \bigcup_{s \in S} sG$ i.e. that S is a union of cosets sG, we find that $S = G \cup sG$ for some $s \notin G$. Note here that $\#S = 2\#G = 2\#H_3$. We now observe the function $f : d \mapsto d + n/d$. We know that $p \in H_3$, however we also know that S has stabilized and that by our construction of T and consequently G, $f : S \to H_3$, where this mapping must be 2-1 onto if we wish to avoid $p + q\mathbb{Z} \notin H_3$, i.e. for $h \in H_3$ we can find d such that $d + n/d = h + q\mathbb{Z}$. We can rewrite this to give

us the following result:

$$d + n/d \equiv h \mod q \tag{107}$$
$$d^2 + n \equiv dh \mod q \tag{108}$$
$$4d^2 - 4dh + 4n \equiv 0 \mod q \tag{109}$$
$$(2d - h)^2 + 4n - h^2 \equiv 0 \mod q \tag{110}$$
$$(2d - h)^2 \equiv h^2 - 4n \mod q \tag{111}$$

As we want this to have a solution for every $h \in H_3$, we must therefore have that $\left(\frac{h^2-4n}{q}\right)=1$. However, as we might recall from 104, we know that $H_3 = \{x^3 : x \in (\mathbb{Z}/q\mathbb{Z})^\times\}$, meaning that for every $x \in (\mathbb{Z}/q\mathbb{Z})^\times$, $\left(\frac{x^6-4n}{q}\right)=1$. This contradicts our Lemma 4.5 for $q \notin \{7, 13\}$ so for $q \notin \{7, 13\}$ we have proven that we cannot avoid there being a $p \in T$ such that $p + q\mathbb{Z} \notin H_3$ and thus we find ourselves again at the conclusion that $S = (\mathbb{Z}/q\mathbb{Z})^\times$. Lastly for $q \in \{7, 13\}$ we can directly check that $\#S_7 = \#\{1, 2, 3, 5, 6\} > \frac{2}{3}(7 - 1)$ and that $\#S_{13} = \#(\mathbb{Z}/q\mathbb{Z})^\times = q - 1 > \frac{2}{3}(13 - 1)$, where the fact that $\#S_q > \frac{2}{3}(q - 1)$ in turn implies that S does not map 2-1 onto H_3 and thus that there exists $p \in T$ such that $p + q\mathbb{Z} \notin H_3$. For every case we have thus shown that $S = (\mathbb{Z}/q\mathbb{Z})^\times$, which was sufficient to show that q will show up as a choice for p_{k+1} in T. □

This means we have proven that for any starting seed and any prime q not yet in the sequence, there exists a sequence in which this q appears. We can do this for every smallest prime q not yet in our sequence so we may state that there exists a sequence containing every prime number. The question whether Chua's specific sequence, choosing the smallest prime each step, will contain every prime however still remains unanswered.

4.4 Recent development

As we mentioned at the start of the proof in the section above we would like S_q to be equal to the multiplicative residue classes as that would shorten the proof considerably; the case $S \neq (\mathbb{Z}/q\mathbb{Z})^\times$ would be ruled out. This result was not known at the time this proof had been established, however it has since been proven by Booker and Pomerance in [13] and stated here verbatim.

Theorem 4.6 (Theorem 1 in [13]) *Let p be a prime different from 5 and 7, and $a \in \mathbb{Z}$. Then there is a squarefree, p-smooth, positive integer n such that $n \equiv a \mod p$.*

With its corollary

Corollary 4.6.1 *Let p be a prime different from 5 and 7, and $a \in (\mathbb{Z}/p\mathbb{Z})^\times$. Then there is a squarefree, $(p-1)$-smooth, positive integer n such that $n \equiv a \mod p$.*

This corollary states exactly that $S_q = (\mathbb{Z}/q\mathbb{Z})^\times$ for $q \notin \{5, 7\}$. With this development, there now exists a published proof that Pomerance's variant generates all primes in order as it follows quite quickly from this. *Proof of Pomerance's variant:* First, observe that the first few terms are, respectively, 2,3,7 and 5. Now let p be the smallest prime not yet in our sequence, and let n be the product of all primes less than p. Notice that the set of squarefree, $(p-1)$-smooth, positive integers is a subset of the set of the divisors of n so our Corollary states that we can choose a $d \mid n$ such that $d \equiv -1 \mod p$ and thus that we can indeed choose $p_{k+1} = p \mid d + 1$. □

5 A new attempt

In my attempt to make yet another variant of the Euclid-Mullin sequence I was inspired by the previous construction, the generalized Euclid sequence, which provably contains every prime. One question that arose was whether it would be possible to focus on a property which some primes shared. The following construction is the one we focus on:
With $P = \{p_1, \ldots, p_k\}$ a finite set of primes, we choose p_{k+1} to be the smallest prime that divides

$$N_I = \prod_{i \in I} p_i^2 + \prod_{i \in \{1,\ldots,k\} \setminus I} p_i^2 \tag{112}$$

for some $I \subseteq \{1, \ldots, k\}$, which is almost the same as the previous construction. This particular construction was chosen to focus on primes of the form $4k+1$ for some $k \in \mathbb{Z}_{\geq 1}$. After considering various forms to focus on, primes equivalent to 1 modulo 4 became the choice mainly because of the following property noted in 2.2.1: $\left(\frac{-1}{p}\right) = 1$ if and only if p is equivalent to 1 modulo 4. We use this property to ensure that all divisors of N_I are such primes equivalent to 1 modulo 4. To show this we again look at the following expression equivalent to 112:

$$N_d \coloneqq N_I = d^2 + \frac{n^2}{d^2} \tag{113}$$

where $d = \prod_{i \in I} p_i$. Let q be a prime such that $q \mid N_d$, then we have:

$$d^2 + \frac{n^2}{d^2} \equiv 0 \mod q \tag{114}$$

After rearranging the terms we end up with the following equivalence:

$$d^4 \equiv -n^2 \mod q \tag{115}$$

What is significant here is that the left-hand-side is definitely a quadratic residue mod q and that the right-hand-side can only be a quadratic residue if and only if $\left(\frac{-1}{q}\right) = 1$. As q divides $d^2 + \frac{n^2}{d^2}$, -1 must be a quadratic residue modulo q so q is a prime such that $q \equiv 1 \mod 4$. Choosing $p_{k+1} = q$ where q is the smallest prime to divide N_d, we end up with the following sequence:

$$2, 5, 29, 17, 41, 13, 37, 53, 61, 97, 101, 73, 89, 109, 149, 137, 113, 181, 173, 157, 229, 197, 241, 257, 233 \ldots \tag{116}$$

Contrary to the Euclid-Mullin sequence, these primes appear almost in order; with 13 being 4 indices away from being in numerical order. We have a heuristic argument to explain why our sequence picks up the smallest prime close to numerical order which is similar to Shank's [18] heuristic argument for the first Euclid-Mullin sequence. Shank's reasoning was that n, the product of all found primes so far, should attain random values modulo a prime p not yet in our sequence thus there is no reason to believe $n \equiv -1 \mod p$ will never occur. We can make the same argument here if we were to fix d at some point. For every prime we find from thereon now n^2/d^2 will presumably attain random values modulo p, where p is the smallest prime yet to be found. Again there is no reason to believe n^2/d^2 will never be equivalent to $-d^2$ for some p. Now as our sequence only obtains values which are square residues, our n^2/d^2 can only attain half the equivalences modulo p, which is exactly why we expect our sequence picks up primes at a faster rate than the first Euclid-Mullin sequence. Moreso, this is for only one fixed $d \mid n$ and we have many d to choose from.

If we were to follow the same steps taken in Booker's proof for the generalized Euclid, to try and show that a general case, i.e. p_{k+1} does not have to be the smallest prime but could be any prime dividing $d^2 + n^2/d^2$, our proof will already diverge at the case where $S_q = (\mathbb{Z}/q\mathbb{Z})^\times$, where q is the smallest prime of the form $4k+1$ not yet in our sequence. In fact, because our n by our construction can only be divisible by finitely

many primes of the form $4k+3$, we can't guarantee S_q will ever equal $(\mathbb{Z}/q\mathbb{Z})^\times$ as we never add primes of the form $4k+3$. This also means we can't use the newer result by Booker and Pomerance, as we are not free to choose from all squarefree $(p-1)$-smooth numbers. Another essential part in their proof is to disprove the case that S stabilizes as anything less than $(\mathbb{Z}/q\mathbb{Z})^\times$. Essential in disproving this is that for any $h \in H$ the number of d such that $d + n/d = h$ must be less than or equal to 2 to allow the proof to focus on subgroups of $(\mathbb{Z}/q\mathbb{Z})^\times$ with index less than 4, however in our case we would have $d^2 + \frac{n^2}{d^2} = h$, where the number of d for any h must be less than or equal to 4. This would result in our version having to focus on index less than 8, which we have been unable to prove. The main difficulty this presents is that for H_7 we do not have a way to disprove that S ever stabilizes as anything less than $(\mathbb{Z}/q\mathbb{Z})^\times$, which is required to show we can choose a sequence in which every q not yet in our sequence will be included. If we however were able to disprove S stabilizes as anything less than $(\mathbb{Z}/q\mathbb{Z})^\times$ our proof would be as follows:

This will be the setup for a proof of the generalized variant of our sequence, so where we have freedom to choose any p dividing $d^2 + \frac{n^2}{d^2}$ for some $d \mid n$ and n the product of all primes in our sequence. The aim is to prove our generalized sequence will attain every prime p of the form $4k+1$ not yet in our sequence. As stated above we assume $S = (\mathbb{Z}/p\mathbb{Z})^\times$ at some point, which is the only step we were unable to prove.

Setup for the proof: Let p be the smallest prime yet to appear in our seed and $n = p_1 p_2 \cdots p_k$. By our construction we know that $\left(\frac{-n^2}{p}\right) = 1$, so there exists some k such that $k + \frac{n^2}{k} \equiv 0 \mod p$. However our construction has d^2 instead of k, so if we wish to choose a d such that this is the case we need that $\left(\frac{k}{p}\right) = 1$. If $\left(\frac{k}{p}\right) = 1$ then we can choose this d, from the assumption that $S = (\mathbb{Z}/p\mathbb{Z})^\times$, such that $d^2 \equiv k \mod p$. If $\left(\frac{k}{p}\right) = -1$ then we can't choose such d, as k (and therefore also $-k$) are not square residues modulo p.

Lemma 5.1 *For q an odd prime and $a \in (\mathbb{Z}/q\mathbb{Z})^\times$ there exists an $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $\left(\frac{x^2 + a/x^2}{q}\right) \neq 1$*

The proof for this is the same as Lemma 4.5*(ii)* so we will not repeat it, except we now have the curve $y^2 = x^4 + a$ which only has repeated roots for $q = 3$. This however won't be a problem as we only focus on primes equivalent to 1 modulo 4 anyway. From this lemma we know that there must exist a d such that $\left(\frac{d^2 + n^2/d^2}{p}\right) = -1$. By multiplicity of Legendre's symbol there must therefore be a prime divisor q of $d^2 + \frac{n^2}{d^2}$ for which $\left(\frac{q}{p}\right) = -1$. This is the prime we choose to add to our sequence of primes, so we now continue not only with $n' = nq$ but also with $k' = kq$ and we keep the equivalence $k' + \frac{n'^2}{k'} \equiv 0 \mod p$. Now we find that $\left(\frac{k'}{p}\right) = 1$ so we may now choose a d, by our assumption that $S = (\mathbb{Z}/p\mathbb{Z})^\times$, such that $d^2 \equiv k' \mod p$ and therefore there now exists a d such that p divides $d^2 + \frac{n^2}{d^2}$. Notice here that it could even be sufficient to show that $S^2 \coloneqq \{s^2 : s \in S\}$ attains all square residues modulo p; because we square our choice of d it doesn't matter whether $-d$ was even an option to begin with.

Even though we were unable to fully prove that our general sequence provably will contain every prime of the form $4k+1$, we do believe this to be the case.

6 Conclusion

In this thesis, we have investigated various methods to generate infinitely many primes through iterative methods. After understanding these sequences and why some show better or more provable results than others we came up with our own sequence. Even though we were unable to prove the general case for our

sequence we still conjecture the following

Conjecture 6.1 *Let $P = \{p_1, \ldots, p_k\}$ be a finite set of prime numbers and choose p_{k+1} as the smallest prime to, for some $I \subseteq \{1, \ldots, k\}$, divide*

$$N_I = \prod_{i \in I} p_i + \prod_{\{1,\ldots,k\} \setminus I} p_i \tag{117}$$

Then this sequence omits no prime equivalent to 1 modulo 4.

A Appendix

The following code, written in Python, is the very naive and manual code used to generate our own sequence 112. Small refinements we made to the code are e.g. checking only half the combinations, as the set I and I^C give the same N_I in our sequence.

```
x = [int(x) for x in input("Enter seed: ").split()]

from itertools import combinations

def diff(first, second):
    second = set(second)
    return [item for item in first if item not in second]

modnum = 13

l = len(x)
for i in range(int(l/2)+1):
        prod = 1
        k = combinations(x,i)
        for j in k:
                d1 = 1
                d2 = 1
                for l in j:
                        d1 *= l
                        d1 %= modnum
                d1 = d1*d1
                d1 %= modnum
                opp = diff(x,j)
                for m in opp:
                        d2 *= m
                        d2 %= modnum
                d2 = d2*d2
                d2 %= modnum
                res = (d1+d2)%modnum
                if res == 0:
                        print(res,j,opp)
```

Add info (publisher) to books

References

[1] Beukers, Frits. **Getaltheorie - Een inleiding**. ISBN 978-90-5041-147-9, fifth publication, 2015

[2] Wooley, Trevor D. **A Superpowered Euclidean Prime Generator.** `https://arxiv.org/pdf/1607.05267.pdf`

[3] Schmidt, Wolfgang M. **Equations over Finite Fields An Elementary Approach**. ISBN 3-540-07855-X

[4] Online Encyclopeida of Integer Sequences (OEIS) `https://oeis.org`

[5] Moll, Victor H. **Numbers and Functions: From a Classical-experimental Mathematician's Point of View.** ISBN 978-0821887950, 65th pucblication, 2012

[6] Booker, Andrew R. **A variant of the Euclid-Mullin sequence containing every prime.** Journal of Integer Sequences, Volume 19, Article 16.6.4, 2016

[7] Booker, Andrew R. Irvine, Sean A., **The Euclid-Mullin graph.** Journal of Number Theory, Volume 165, pages 30-57, 2016

[8] Booker, Andrew R. **On Mullin's Second Sequence of Primes.** Integers, Volume 12A, Article A4, 2012

[9] Hardy, Godfrey H. Wright, Edward M., **An Introduction to the Theory of Numbers (Sixth edition).** ISBN 978-7-115-21427-0

[10] Vardi, Ilan. **Computational Recreations in Mathematica.** ISBN 0-201-52989-0

[11] Crandall, Richard. Pomerance, Carl. **Prime Numbers A Computational Perspective Second Edition.** ISBN 0-387-25282-7

[12] Pollack, Paul. Treviño, Enrique. **The Primes that Euclid Forgot.**

[13] Booker, Andrew R. Pomerance, Carl. **Squarefree smooth numbers and Euclidean prime generators.** Proc. Amer. Math. Soc. Volume 145, 2017

[14] Armstrong, Mark A. **Groups and Symmetry.** ISBN 0-387-96675-7

[15] Guy, Richard K. Nowakowski, Richard J. **Discovering Primes with Euclid.** Delta, Volume 5, 49-63, 1975

[16] Mullin, Albert A. **Recursive Function Theory.** Bull. Amer. Math. Soc. Volume 69, 737, 1963

[17] Wagstaff, Samuel S. **Computing Euclid's Primes.** Bull. Inst. Combin. Appl. Volume 8, 1993

[18] Shanks, Dan. **Euclid's primes.** Bull. Inst. Combin. Appl. Volume 1, 1991

[19] Ribenboim, Paulo. **The New Book of Prime Number Records.** ISBN 978-0-387-94457-9, 3rd publication, 1996

[20] Ribenboim, Paulo. **The Little Book of Bigger Primes Second Edition.** ISBN 978-0-387-20169-6, 2004

[21] Naur, Thorkil. **Mullin's Sequence of Primes is not Monotonic.** Proc. Amer. Math. Soc. Volume 90, Number 1, 1984

[22] Caldwell, Chris K. Gallot, Yves. **On the Primality of $n! \pm 1$ and $2 \times 3 \times 5 \times \cdots \times p \pm 1$.** Math. Comp. Volume 71:237, 441-448, 2002

[23] Cox, C. D. van der Poorten, A. J. **On a sequence of prime numbers.** J. Austral. Math. Soc. Volume 8, 571-574, 1968

[24] Caldwell, Chris K. Dubner, Harvey. **Primorial, factorial, and multifactorial primes.** Math. Spectrum 26, 1-7, 1993/4

[25] Caldwell, Chris K. **On the primality of $n! \pm 1$ and $2 \cdot 3 \cdot 5 \cdots p \pm 1$.** Math. Comp. Volume 64:2, 889-890, 1995

[26] Dubner, Harvey. **Factorial and primorial primes.** J. Recr. Math. Volume 19, 197-203, 1987

[27] Odoni, R. W. K. **On the prime divisors of the sequence $w_{n+1} = 1 + w_1 w_2 \cdots w_n$.** J. Lonon Math. Soc. Volume s2-32, Issue 1, 1-11, 1985

[28] Rogers, Kenenth. **The Schnirelmann density of the squarefree integers.** Proc. Amer. Math. Soc. Issue 15, 515-516, 1964

[29] Green, Ben. Tao, Terence. **The primes contain arbitrarily long arithmetic progressions.** Annals of Mathematics. Issue 167(2), 481-547, 2008