Prime Number Algorithm

Nicholas M. Roma

June 23, 2021

Abstract

A sieve will reveal the prime numbers by enumerating the natural numbers and filtering multiples. It is possible to generate the primes by successively generating *non-multiples*, i.e., *gaps*, of the primes already generated. The algorithm that generates primes based on this method uses approximately *fewer* iterations than the optimal Sieve of Atkin at $\mathcal{O}(N/\log\log N)$. This paper will

1. define *gaps* and establish the method of computing *compound gaps*,
2. derive an expression for primes as compound gaps,
3. define the algorithm for generating primes, and
4. demonstrate the approximate number of iterations.

1 Computation of Non-multiples, or *Gaps*

Throughout the development of the expression for the sequence of primes, the term *gaps* has been used to mean *non-multiples.* The term *gaps* is traditionally used to refer to the differences between successive primes, but below it will mean *non-multiples.*

Gaps The set of gaps of a number n, denoted γ_n, is the set of all non-multiples $g \not\equiv 0 \mod n$. For example, the gaps of 2 are

$$\gamma_2 = \{1 \mod 2\}.$$[1]

The gaps of 3 are

$$\gamma_3 = \{1, 2 \mod 3\}.$$

In general the gaps of $n \in \mathbb{N}$ for $n > 1$ are

$$\gamma_n = \{1, 2, \ldots n-1 \mod n\}.$$

It's interesting to note: the gaps of 0 are

$$\gamma_0 = \mathbb{N}^0 - \{0\},$$

[1] Where the braces nicely explicate the *class* of such equivalents.

and the set of gaps of 1 is

$$\gamma_1 = \varnothing.$$

Similarly, for any n,

$$0 \notin \gamma_n,$$

$$1 \in \gamma_n,$$

and, of course,

$$n \notin \gamma_n,$$

in as much as all multiples of $n \notin \gamma_n$ and therefore $n \cdot 1 \notin \gamma_n$.

Bases of Gaps Obviously, gaps are *modulo* n, or equivalently, gaps are relative to a *base* n. This correlates to a description of the gaps of n as a set of expressions. For example, with $x \in \mathbb{N}^0$, the gaps of 2 are

$$\gamma_2 = \{2x + 1\},$$

and in general the gaps of base $n \in \mathbb{N}$ for $n > 1$ are

$$\gamma_n = \{nx + 1, nx + 2, \ldots nx + m_{n-1}\}.$$

The gaps of n constitute a set of equivalence classes *modulo* n, that is, all the classes not equivalent to zero.

Characteristic of an Equivalence Class The *characteristic* is the expression (in base b) that yields the elements of an equivalence class $m_i \mod b$, where each element in the class *conforms* to the characteristic. For example, the equivalence class determined by modulus b and some m_i has the *characteristic* of $bx + m_i$. A *proper* characteristic is given as $bx + (m_i \mod b)$. In these terms, *multiples* of b have the characteristic $bx + 0$ and *gaps* of b have characteristics like $bx + m_i$ where $m_i \neq 0$. Let m_i be called the *module*, [2] and let the *equivalents* mean the output of a characteristic expression, i.e., the elements of that class. A characteristic is determined by the *base* and the *module*, but with a known base, a module m_i can determine and be determined by a characteristic (and the terms could be interchangeable).

Rank of Equivalent The value of x in the characteristic is the *rank* of that equivalent. If $e_x = bx + m_i$ then e_x has *rank* of x in the class $m_i \mod b$. The *unranked* equivalent is the equivalent at rank $x = 0$.

[2] Where a *module* corresponds to a column when the $\mathbb{N}^0$ are written in base b many columns.

Characteristics of Gaps Gaps, then, are sets of characteristics, i.e., those whose equivalents are not multiples of the base. Let the *characteristics of gaps* be denoted as $\chi\gamma_b$, for example,

$$\chi\gamma_5 = \{\text{“}5x+1\text{”}, \text{“}5x+2\text{”}, \text{“}5x+3\text{”}, \text{“}5x+4\text{”}\}.$$

Since the base will be known, the characteristics can equivalently be given by just the *modules*, for example,

$$\chi\gamma_5 = \{1, 2, 3, 4\}.$$

Compound Gaps Let *compound gaps* mean the gaps of all of several bases, denoted $\gamma(b_1, b_2, \ldots b_n)$. As *expressions*, it is possible to compose [3] the characteristics in order to describe compound gaps. For a characteristic of the gaps of b, like $bx + m_i$, it is necessary to consider which characteristic k_i of a coprime base c will result in a multiple of c:

$$b(cx + k_i) + m_i.$$

Since k_i is being multiplied by b then added to m_i, it follows that $b \cdot k_i$ should be equivalent to $c - m_i \mod c$, and therefore k_i should be equivalent to $(b^{-1} \mod c) \cdot (c - m_i) \mod c$. Since b and c are coprime, there is necessarily exactly one $b^{-1} \mod c \in \{0 \ldots c-1\}$, given by

$$b^{-1} \mod c \equiv b^{\phi(c)-1} \mod c,^{4}$$

so

$$k_i \equiv [(b^{\phi(c)-1} \mod c) \cdot (c - m_i \mod c)] \mod c.^{5}$$

This k_i results in

$$\begin{aligned} & b[cx + b^{-1} \cdot (c - m_i)] + m_i \\ = {} & bcx + 1 \cdot (c - m_i) + m_i \mod c \\ = {} & bcx + c \\ = {} & c(bx + 1) \end{aligned}$$

which is certainly divisible by c. Moreover, if

$$k_i \not\equiv -m_i/b \mod c \implies$$

$$b \cdot k_i \not\equiv -m_i \mod c \implies$$

[3] Where the terms *compose gaps* and *composition of gaps* are fine, the term *composite gap* should only mean a gap which is a composite number, and the result of composing gaps should be meant by *compound gaps*

[4] By Euler

[5] Note when c is *prime* then

$$k_i \equiv [(b^{c-2} \mod c) \cdot (c - m_i \mod c)] \mod c.$$

$$b \cdot k_i + m_i \not\equiv 0 \mod c \implies$$
$$bcx + (b \cdot k_i + m_i) \not\equiv 0 \mod c.$$

Therefore, *all of the other* $k_i \in \{0 \dots c-1\}$ *result in gaps of* b *and of* c. [6] For the gaps $\gamma(b_1, b_2, \dots b_n)$, it is interesting to note:

$$0 \notin \gamma(b_1, b_2, \dots b_n),$$
$$1 \in \gamma(b_1, b_2, \dots b_n),$$

and

$$\{b_1, b_2, \dots b_n\} \cap \gamma(b_1, b_2, \dots b_n) = \varnothing,$$

by extension of the notes above in the case of *simple* gaps. Further, for all multiples of all the b_i,

$$b_i \cdot x \notin \gamma(b_1, b_2, \dots b_n).$$

Characteristics of Compound Gaps The characteristics of compound gaps have the form

$$\gamma(b_1, b_2, \dots b_n) = \{(b_1 \cdot b_2 \dots \cdot b_n)x + m_i\} \mid 1 \le m_i \le (b_1 \cdot b_2 \dots \cdot b_n) - 1.$$

In particular, the compound gaps of successive prime bases will have the form

$$\gamma(2, 3, \dots p_n) = \{(p_n\#)x + m_i\} \mid 1 \le m_i \le p_n\# - 1.$$

Number of Characteristics of Compound Gaps The k_i which bears a multiple of c has been referred to (in the utmost *formal* situations) as the *magic mod*. By extension, all of the other k_i in c's modules would be *muggle mods*. There is always *one* magic mod and $c-1$ muggles (from $0 \le k_i \le c-1$). When all of the characteristics are composed, then, for each of the characteristics of gaps of b, there will be $c-1$ many resultant characteristics, one for each muggle mod of c. [7] In other words, the number of characteristics for the compound gaps will be like the muggles of c for each m_i of b's. So,

$$|\chi\gamma(b, c)| = (b-1) \cdot (c-1).$$

And in general for $\gamma(b_1, b_2, \dots b_n)$

$$|\chi\gamma(b_1, b_2, \dots b_n)| = (b_1 - 1) \cdot (b_2 - 1) \cdot (b_3 - 1) \dots (b_n - 1).$$

In particular, the number of characteristics for compound gaps of successive primes to p_n, denoted $g_\pi(p_n)$, is

$$\begin{aligned} g_\pi(p_n) &= (2-1) \cdot (3-1) \cdot (5-1) \dots \cdot (p_n - 1) \\ &= \prod_{i=1}^{n} p_i - 1. \end{aligned}$$

[6] In fact, in this case, $bcx + (b \cdot k_i + m_i) \equiv m_i \mod b$ and $bcx + (b \cdot k_i + m_i) \equiv (b \cdot k_i + m_i) \mod c$

[7] It is important to note all the k_i of c's modules get *a shot* to be the magic mod for a given m_i of b's. In other words, they take turns being magic or muggle depending on the m_i

2 Primes as Compound Gaps

Nearest Prime Functions It will be helpful to define the *nearest prime predecessor* of a number n as

$$n' = \text{the largest prime } p \mid p \leq n,$$

and the *least greater prime* of n as

$$n^* = \text{the smallest prime } p \mid p > n.$$

Note

$$n'^* = n^*,$$

and

$$(n')' = n'.$$

To pull a prime number down to the previous prime, it would be neccessary find the predecessor of the number minus one:

$$p_{i-1} = (p_i - 1)',$$

where

$$(p_i)' = p_i.$$

Prime Gaps on P_n For the gaps of several prime bases, $\gamma(p_1, p_2, \ldots p_n)$, none of the bases nor their multiples are elements in the gaps. In particular, for *successive* primes, $\{2, 3, 5, \ldots p_n\}$, none of them nor their multiples are elements in the gaps. Therefore, all $g \in \gamma(2, 3, 5 \ldots p_n)$ have a prime factorization like

$$g = p_{n+1}^{x_1} \cdot p_{n+2}^{x_2} \cdot p_{n+3}^{x_3} \cdot p_{n+4}^{x_4} \cdots$$

When all the $x_i = 0$, the result is $g = 1$, which is known to be an element of the gaps. The smallest non-trivial element is when $x_1 = 1$ and the remaining exponents are all 0, which is p_{n+1}. Similarly, the smallest composite gap is when $x_1 = 2$ and the remaining exponents are all 0, that is, p_{n+1}^2. What this means is

1. *the smallest non-trivial element of the gaps of successive primes is always the next prime,*

2. *the smallest composite number in the gaps is* p_{n+1}^2, and therefore

3. *all* $g \in \gamma(2, 3, \ldots p_n) \mid p_{n+1} \leq g < p_{n+1}^2$ *are prime!*

Noting that $p_{n+1} = p_n^*$, let the *prime gaps on* p_n, denoted $\bar{\gamma}_{p_n}^*$, be

$$\bar{\gamma}_{p_n}^* = g \in \gamma(2, 3, \ldots p_n) \mid p_n^* \leq g < (p_n^*)^2.$$

The careful reader will notice that in the case of $\bar{\gamma}_{p_n}^*$, the subscript p_n denotes *all* the successive prime bases, i.e., $2 \ldots p_n$.

Proper Prime Gaps on P_n The prime gaps on p_{n-1} are

$$\bar{\gamma}^*_{p_{n-1}} = g \in \gamma(2, 3, \ldots p_{n-1}) \mid p_n \leq g < p_n^2$$

Since $p_n < 2 \cdot p_{n-1} \implies p_n < p_{n-1}^2$, [8] there is always an overlap of the prime gaps on p_{n-1} with those on p_n, namely

$$\bar{\gamma}^*_{p_{n-1}} \cap \bar{\gamma}^*_{p_n} = \{p_n^*, \ldots (p_n^2)'\}.$$

Let the *novel* prime gaps on p_n, or the *proper* prime gaps on p_n, denoted $\gamma^*_{p_n}$, be the gaps

$$\gamma^*_{p_n} = g \in \gamma(2, 3, \ldots p_n) \mid p_n^2 < g < (p_n^*)^2.$$

Regular Prime Gaps At a certain point p_n, the base $p_n\#$ is large enough that all of the prime gaps are always *unranked*. If the characteristics of the prime gaps are

$$\chi\gamma^*_{p_n} = \{(p_n\#)x + m_i\} \mid p_n^2 < m_i < (p_n^*)^2,$$

since this base $p_n\# >> (p_n^*)^2$ then all primes from $(p_n^2)'$ to $((p_n^*)^2)'$ correspond *singly* [9] to a characteristic, because $x > 0 \implies (p_n\#)x + m_i > (p_n^*)^2$, thus x can only be 0, and therefore all the prime gaps are *unranked*. In fact, at this same p_n, the base is also large enough that *all* primes thru $(p_n\# - 1)'$ are *all* unranked. The prime gaps which are always unranked are the *regular* prime gaps, i.e., $p_n \geq 7$.

Primes as Compound Gaps of Primes These terms allow for the expression of the sequence of primes. Let n be such that $p_r^2 \leq n < p_{r+1}^2$.[10] The sequence of primes *on* n, denoted α_n, is given by

$$\begin{aligned} \alpha_n &= \{2, 3, 5, \ldots (p_r^2)'\} \cup \{(p_r^2)', \ldots n', \ldots (p_{r+1}^2)'\} \\ &= \alpha_{p_r^2} \cup \gamma^*_{p_r} \\ &= \{2, 3, 5, \ldots (p_{r-1}^2)'\} \cup \{(p_{r-1}^2)', \ldots (p_r^2)'\} \cup \gamma^*_{p_r} \\ &= \alpha_{p_{r-1}^2} \cup \gamma^*_{p_{r-1}} \cup \gamma^*_{p_r}. \end{aligned}$$

By continuation,

$$\alpha_n = \gamma^*_{p_0} \cup \gamma^*_{p_1} \ldots \gamma^*_{p_{r-1}} \cup \gamma^*_{p_r}.$$

[8] by Chebyshev

[9] i.e., one-to-one

[10] So p is n's *prime root correspondent.*

3 Prime Number Algorithm

Tactic of the Algorithm From the expression of primes on n above, the tactic of computing primes would be to successively compose the characteristics of primes to the *prime root correspondent* of n, at each step yielding the proper prime gaps on p_i. There is a variation of the algorithm which keeps the base smaller (but still astronomical) and probably scales better. This variation utilizes the fact that a non-trivial n' will be a regular prime gap when the base $p_c\#$ becomes greater than n', because at that point, the characteristics span $p_c^* \le n' \le p_c\# - 1$, and therefore all primes thru n' will be computed [11], after *removing the composite modules.*

Inverse Primorial Correspondent Let the *inverse primorial correspondent*, denoted $n\#_c$, be the smallest prime p_c such that $p_c\# \ge n$. Note

$$(n\#_c)\# \ge n,$$

and

$$(p_c\#)\#_c = p_c.$$

The characteristics of the gaps of primes thru p_c contain $p_c\#x + n'$, for non-trivial n.

Characteristics with a Composite Module After computing all characteristics to n', it will be important to the algorithm to know which characteristics will be composite. Excluding the characteristics with composite modules will leave only the primes, and, assuming the characteristics are *regular*, all primes in the range p_c^* to $(p_c\# - 1)'$ will be represented by the remaining characteristics. The modules from p_c^* to $(p_c^*)^2$ are already known to be prime (as the prime gaps on p_c). The composite modules in the range $(p_c^*)^2$ to $p\# - 1$ are given by all combinations of prime factors from p^* to $(\frac{p\#-1}{p^*})'$. It's possible to express all combinations of factors as pairs, by first considering $m_i = f_1 \cdot w_1$, then considering $m_j = m_i \cdot w_2$. To this end I want to consider the factors f_i as a factor and as a tuple of *subfactors*. When I mean f_i as a factor I'll say $|\vec{f_i}|$, which is the product of all the components of $\vec{f_i}$, and when I mean all the subfactors of f_i I'll say $\langle c_1, c_2, \dots c_n\rangle$. In particular when I mean the maximum subfactor I'll say $\lceil \vec{f_i} \rceil$ [12]. In this way the algorithm can carry the prime factorization with the current factor being considered.

[11] Incidentally, n' will also occur in other compound gaps, not *just* when $p_c\#$ becomes greater than n'.

[12] If you can forgive my taking such notational liberties (but I've already gone *this* far; you should've raised your objection before now)

Pseudocode

1. **Set** $\phi \leftarrow \{\langle p^* \rangle, \ldots \langle \sqrt{p\# - 1}' \rangle\}$, where ϕ is a set of tuples of prime factors.

2. **For Each** $\vec{f_i} \in \phi$

   (a) **If** $\frac{p\#-1}{|\vec{f_i}|} \geq p^*$ **Then**

      i. **For Each** $\omega_j \in \{p^*, \ldots (\frac{p\#-1}{|\vec{f_i}|})'\}$ where $\omega_j \geq \lceil \vec{f_i} \rceil$

         A. **Set** $m_i \leftarrow |\vec{f_i}| \cdot \omega_j$

         B. **Add** m_i to *compositeModules*

         C. **If** $\frac{p\#-1}{m_i} \geq p^*$ **Then Add** $\vec{m_i} = \langle c_1, c_2, \ldots c_n, \omega_j \rangle$ to ϕ, where $\langle c_1, c_2, \ldots c_n \rangle = \vec{f_i}$.

Computation of the Sequence of Primes Generating the primes to N' then becomes the same as computing the gaps of the primes from 2 to $N\#_{c-1}$, then composing these characteristics with $N\#_c$ but only until reaching the characteristic for N', then iterating the characteristics from $(p_{c+1}^2)'$ to N' *except* those with a composite module.

Pseudocode

1. **Set** $primes \leftarrow \{2, 3\}$; **Set** $p_i \leftarrow 3$, **Set** $base \leftarrow 2\#$, **Set** $p_{i+1} \leftarrow 3$

2. **Set** $partition \leftarrow \{2\#x + 1\}$

3. **Do Until** $base \geq N$, i.e., for $base \in \{2, \ldots N\#_{c-1}\}$.

   (a) **Add** to *primes* all equivalents e_{prime} in the *partition* where $p_i^2 < e_{prime} < p_{i+1}^2$ and $e_{prime} \leq N$

   (b) **Set** $p_i \leftarrow p_{i+1}$, **Set** $p_{i+1} \leftarrow primes.Next$

   (c) **For** $r_i = 0 \ldots p_i$

      i. **For Each** characteristic m_i **in** *partition*

         A. **Compute** k_i from $base(p_i x + k_i) + m_i$. ($base$ and p_i must be coprime because p_i is a prime and $base$ is the product of the primes less than p_i.)

         B. **If** k_i is negative **Then Set** $k_i \leftarrow k_i + p_i$. (This k_i should be $0 \leq k_i < p_i$.)

         C. **If** $k_i = r_i$ **Then Continue**

         D. **If** $base \cdot p_i x + base \cdot r_i + m_i \leq N$ *for any rank* ≥ 0 **Then Add** it to the *partition*.

         E. **Else Break**. (All m_i after this will also be greater than N.)

   (d) **Set** $base \leftarrow base \cdot p_i$

4. **Compute** the $\chi_c\gamma(2, 3, \ldots N\#_c)$ where the modules are less than or equal to N'

5. **Add** $\chi_p = \chi\gamma(2, 3, \ldots N\#_c) - \chi_c\gamma(2, 3, \ldots N\#_c)$ to *partition*.

Implementation Notes An implementation may need to create a partition *per round*, otherwise the enumeration will alter the set it is enumerating. The *base* would need to be a non-primitive type capable of holding an arbitrarily large number, in fact it will become $N\#_c\#$. This can cause a long time delay as a component operation of computing k_i is finding $p_i\# \mod p_j$. Since *base* is constant per step of the do-until loop, storing a copy of *base* per characteristic in the partition is not necessary. This is why there is a separate variable for *base* above, despite the fact characteristics are described as complete expressions. Since each r_i of p_i is iterated over each m_i of *base*, the primes will be generated *in order*, and as *proper prime gaps* on *base*, they are not duplicated. It is necessary to compute the range of ranks for *irregular* prime gaps, i.e., for primes less than $(11^2)'$ which will have ranks from about $0 \le x \le 4$. Once the *base* is greater than or equal to 7, it can be assumed they are *unranked*.

4 Big O PNA

Number of Computed Characteristics The $\mathcal{O}(PNA)$ is concerned with the number of compositions of characteristics - the same as the number of k_i's computed. The algorithm effectively iterates over $base = 2\#\ldots N\#_c$, but it will break as soon as the resultant characteristic is larger than N. This effectively means all characteristics of $\gamma(2,3,\ldots N\#_{c-1})$ are computed, and all characteristics for $N\#_c$ up to N', so the number of compositions is

$$\sum_{p_i=2}^{N\#_{c-1}} g_\pi(p_i) + g_\pi(N\#_c)\frac{N}{N\#_c\# - 1}.$$

To remove the composite modules, the algorithm effectively removes all non-primes from the characteristics:

$$g_\pi(N\#_c)\frac{N}{N\#_c\# - 1} - [\pi(N) - \pi(N\#_{c+1})] - 1.$$

The total number of steps of the algorithm is approximately

$$\sum_{p_i=2}^{N\#_{c-1}} g_\pi(p_i) + g_\pi(N\#_c)\frac{2N}{N\#_c\# - 1} - \pi(N) + \pi(N\#_{c+1}) - 1.$$

5 Conclusions

TBD: We want some kind of proof of correctness of the given formula and an analytic comparison of Omega PNA with Omega Atkin, i.e. $\mathcal{O}(N/\log\log N)$.