UNIVERSITÄT DUISBURG ESSEN

Open-Minded

University of Duisburg-Essen

Faculty of Engineering

Mathematics for Engineers

Institute for Technologies of Metals

A SIEVE FOR FINDING PRIMES IN A RANGE

FIELD: NUMBER THEORY

Authors:
Yasseen Nandally
B.Sc. Computer Engineering

Supervisors:
PD Dr. rer. nat. habil. Robert Martin
Dr. Claudia Gotzes

Winter semester 2020/2021

From: November 2020 To: February 2021

Abstract

What is the next prime number? How to identify if an integer is prime or not? Such questions or the topic prime number in general, have always been a puzzle. The study of prime number trace its route back to around 300 BC by the Greek Mathematician Euclid. Since then, the analysis of prime numbers and its theorems have known an increase in the field of number theory. Based on properties of primes, namely unique factorization, infinitude among others, several methodologies have been derived to find the next prime number or to answer other questions. Those methods include abstract algebra, computational methods etc... Even for defining the properties, there are multiple theorems that have been discovered. The computational method used in this thesis is a sieving method. The idea is to find prime numbers within a given range where only integers lying in this range are of interest. Apropos recognizing the next large prime number, performance and storage have always been the subject of debate when it comes to implementation of an algorithm. The procedure explained in this document is an attempt to improve both performance and storage if it was to be implemented. For a better performance, this is achieved by parallelizing computation as the formulae are independent from each other and concerning storage, variables are used to avoid copying large data every time.

Contents

Chapter 1

Introduction

The aim of the procedure described in this thesis is to identify primes other than trial division method. The initial idea was to discover patterns among primes, but since they are randomly distributed, the motive of pattern recognition which could lead to a formula had gone away. The next idea relies on the fact that prime numbers are non-composite numbers, and thus by eliminating the set of composite numbers from the set of natural numbers, this results in the set of primes. Essentially, this is the beginning point. Composite numbers are emerged from the multiplication of natural numbers, but by listing all the composite values from their factors, this would result in a disordered arrangement. That is where a multiplication grid comes in-to play. The main objective of the grid is to be able to observe composite numbers in an orderly manner. Then comes the captivating part which is analysing the grid thoroughly to be able to deduce patterns encompassed within it, for example, progression within the diagonals. Furthermore, how to read the diagonals itself is not apparent at first sight. In multiplication grid, the most common observable patterns are from the rows, columns and main diagonal which contains perfect squares. However, in this thesis whenever diagonal is addressed, it means when reading in the direction north-east to south-west, unless quoted explicitly. More patterns are explained in the 1st chapter. Then the rest is deriving formulae and reducing computations. A side note to the reader: Many times, between square brackets one will see more or less the following ["as explained above"]. It is an indication telling the reader that at this point, what has explained above is important to be able to understand what follows. In some places, the section to be referred to is explicitly mentioned but when not mentioned, it means the reader needs to have an overall understanding of what has been discussed earlier. Also there are explanations that are repeated, so that the reader doesn't have to refer back, thus making reading easier.

Chapter 2

Literature Review

2.1 Definition and introduction to Prime numbers

A prime number is a positive integer, p having exactly two positive divisors, namely 1 and p. Contrary, an integer n is said to be composite if $n > 1$ and n is not prime. [6]
All prime numbers except 2 are said to be odd prime numbers but not all odd numbers are prime. [24]
In other words, one can say that all prime numbers except 2 are not even numbers or the only even prime number is 2.

2.2 A Historical Review

- History shows prominent study of prime numbers began from ancient Greek Mathematician. Around 300 BC, the famous Greek Mathematician Euclid of Alexandria commonly known as Euclid shows that prime numbers are infinite, and he also gave the unique factorization theorem. [23]

- The Sieve of Eratosthenes named after a Greek Mathematician is another brilliant way of finding prime numbers up to a definite limit. [19]

- Fibonacci's book Liber Abaci (1202) explained the method of trial division. [16]

- In the year 1640, the French Mathematician Pierre de Fermat gave a theorem which is known as Fermat's little theorem stating that: if a is an integer, and p is prime, then $a^p \equiv a \pmod{p}$. [10]

- Pierre de Fermat also formulated the Fermat numbers which is in the form: $F_n = 2^{2^n} + 1$. [26]

- In 1737, the renowned Swiss Mathematician, Leonard Euler proved that the summation of the reciprocals of prime numbers diverges: [17]

$$\sum_{p,\text{prime}} \frac{1}{p} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + ... = \infty$$

- Riemann's Zeta function published in 1859, is of great importance in prime numbers. [5]

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

- Based on the Riemann Zeta function, Euler shows the interrelation between prime numbers and the Zeta function, known as the Euler product formula. [5]

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

- The Riemann function for counting prime numbers to a certain bound based on the zeta zeros is defined as follows: [7]

$$\prod(x) = \sum_{n=1}^{\infty} \frac{1}{n}\pi(x^{\frac{1}{n}})$$

- The Möbius function $\mu(n)$ named after the German Mathematician August Ferdinand Möbius, is the inverse of the Zeta function. [7]
  It is represented as follows:

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

  The rule that describes the $\mu(n)$ is as follows:

  -1 if n contains an odd number of primes.
  1 if n contains an even number of primes.
  0 if n contains a quadratic prime factor.

- A further step in exploring prime numbers was in the 19th century made by the German Mathematician Johann Peter Gustav Lejeune Dirichlet. He stated that there are infinitely many prime generated by some arithmetic progression. The theorem is as follows: [1]

$$kn + h, n = 0, 1, 2, \ldots$$

  where h is the first term of the arithmetic progression and k is the common difference. If h and k have a common factor d, each term of the progression is divisible by d and there can be no more than one prime in the progression if $d > 1$. In other words, there exist infinitely many primes if the condition $(h, k) = 1$ is satisfied, where $(h, k) = 1$ is the gcd.

2.3 Abstract Algebra

Another branch of Mathematics used in finding prime numbers is the abstract algebra. Modular arithmetic and finite fields are one of the approaches among others like group theory, prime elements in a ring and so on. The abstract algebra part is left out as this thesis is oriented towards the computational methods.

2.4 Calculation methods

2.4.1 Trial division

The easiest way to check if an integer, n is prime or not, is by trial division. Starting with 2, n is divided until it is not divisible by 2 anymore. Then the next prime number becomes the next divisor candidate and so on. For example $N = 7399$. Starting with 2, n is not divisible by 2. The same applies for 3 and 5. The next divisor is 7 and the quotient is then 1057. The quotient is further divided using 7 resulting in 151. 151 is again divided by 7, but 7 is not a factor of 151, so the next prime number is chosen which is 11. 11 as well is not a factor of 151. Next prime number is 13, but 13 being greater than $\sqrt{151}$, means it is the stopping point and the prime factorization of 7399 is therefore $7 \cdot 7 \cdot 151$. [19]

2.4.2 Sieving methods

Sieve methodology are used to find factors, patterns in arithmetic progression, identifying prime numbers or for cardinalities approximation of sets based on specific properties such as multiplicative. The sieving process removes unwanted candidates up to a definite limit which later leaves the intended solution set. [15]

Apart from the discussed algorithms below, there are many other methods for generating prime numbers. For example, a linear time algorithm was given by Pritchard also known as the wheel of factorization or wheel sieve. [21]
As for Bengelloun sieve, he showed how a miniscule amount of surplus computation is required to run his algorithm in $O(n)$ time for finding prime numbers up to a defined limit. [4]
For the case of Mairson's sieve, the algorithm takes $O(n)$ time as well to observe prime numbers between 2 and N. [20]

Sieve of Eratosthenes

The sieve of Eratosthenes is used to generate prime numbers up to a bound N. Starting off by the first prime number 2, cross out all multiples of 2 till N. Then move to the next number that has not been crossed out yet and do the same. At the end, a list of prime numbers is obtained. [22]

Segmented sieve

Around the 13th century A.D Bays and Hudson presented their algorithm which is an improved version of the sieve of Eratosthenes by running the process till $\sqrt{N}$. In other words, composite numbers are eliminated iteratively. [2]
Example: Find all prime such that $p \leq 30$.

- A list of integers starting from 2 to 30 is generated:

  2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

- Beginning from 2, cross out every integer by incrementing by 2:
  2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ 9 ~~10~~ 11 ~~12~~ 13 ~~14~~ 15 ~~16~~ 17 ~~18~~ 19 ~~20~~ 21 ~~22~~ 23 ~~24~~ 25 ~~26~~ 27 ~~28~~ 29 ~~30~~

- The same procedure as above is repeated, but this time starting with the next uncrossed number which is 3 and then increments by 3:
  2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ 25 ~~26~~ ~~27~~ ~~28~~ 29 ~~30~~

- Then 5 is the next integer:
  2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ ~~25~~ ~~26~~ ~~27~~ ~~28~~ 29 ~~30~~

- The next chosen number is 7. After every increment of 7, an integer is eliminated but as it can be seen, all the multiples of 7 lesser than 30 are already eliminated and note than $7 \cdot 7$ is 49 which is greater than 30. Thus, the procedure can come to a halt. Finally, the remaining numbers are prime:

  2, 3, 5, 7, 11, 13, 17, 19, 23, 29

Sieve of Sundaram

Sundaram sieve is an array of integers emerging from arithmetic progression, meaning consecutive integers are of definite difference.

- First step is to begin with 4 and increments by 3 infinitely to obtain a sequence of numbers as follows:
  4, 7, 10, 13, 16, 19, 22, 25,...

- Second step is to build another infinite sequence with increment of 5, starting with the number 7 as shown below:
  7, 12, 17, 22, 27, 32, 37, 42,...

- Third sequence starts with the number 10 and the increment is 7:
  10, 17, 24, 31, 38, 45, 52, 59,...

Notice the pattern in the 3 bullet points above. Each sequence is exactly 3 distance from the previous one and the difference between consecutive integers is equal to the previous one plus 2. The sequence can be represented in an array as displayed below:

4	7	10	13	16	19	22	$25\cdots$
7	12	17	22	27	32	37	$42\cdots$
10	17	23	31	38	45	52	$59\cdots$
13	22	31	40	49	58	67	$76\cdots$
16	27	38	49	60	71	82	$93\cdots$
$\cdots$	$\cdots$	$\cdots$					

A remarkable observation from the array is that the difference between 2 consecutive rows gives rise to the arithmetic progression: 3, 5, 7, 9, 11, 13, 15, 17, ...

- Any number x not found in this array, then $2x + 1$ is prime.
- If an integer x is in the array, then $2x + 1$ is not prime.

In other words, $2x + 1$ is prime iff x is not in the array.

[13]

Sieve of Atkin-Bernstein

Invented by A.O.L Atkin and Daniel Bernstein. The sieve is used to find prime numbers up to a certain bound. It is an enhanced version of the sieve of Eratosthenes. It proceeds as follows, in the algorithm:

1. All remainders are modulo-sixty.
2. All numbers are positive integers.
3. Changing the marking to opposite is done by tossing an entry in the sieve.
   a) Create a results list, filled with 2, 3, and 5.
   b) Create a sieve list with an entry for each positive whole number; all entries of this list should initially be marked non-prime.
   c) For every entry l in the sieve list having modulo sixty as remainder r:
      i. If r is 1, 13, 17, 29, 37, 41, 49, or 53, flip it for every possible solution to $4x^2 + y^2 = l$.
      ii. If r is 7, 19, 31, or 43, toss it for each possible solution to $3x^2 + y^2 = l$.
      iii. If r is 11, 23, 47, or 59, flip it for each possible solution to $3x^2 + y^2 = l$ when $x > y$.
      iv. Otherwise, ignore.
   d) Begin with the smallest number of the list.
   e) Take the next number in the sieve list still marked prime.
   f) Insert the number in the results list.

g) Square the number and mark all multiples of that square as non-prime.

h) Repeat steps (d) to (g).

[25]

Incremental sieve

Another more efficient algorithm than the sieve of Eratosthenes is the Incremental sieve. The difference between the two is based on how the crossing of non-prime is done.

Here a table is used to save every prime integer p, that has been discovered, as well as an iterator to keep the next multiple of p to cross out. For example, instead of eliminating all multiples of 17 at one go which is impossible since there is an infinite multiple of 17, the 1st one is saved in the table at $289(17 \cdot 17)$. Upon reaching 289 to verify whether it is prime or not, it is noticed that 289 is a composite integer having 17 as a factor. Thus, it is removed and replaced by $306(289 + 17)$. The above-described algorithm will advance the iterators for 17, 3, and 2 in order for their actual position to be all similar which is 306. When referencing 306, the iterators then separate off. [18]

2.5 Arithmetic on Odd and Even numbers

2.5.1 Addition and subtraction

Operations	Results
Even $\pm$ Even	Even
Even $\pm$ Odd	Odd
Odd $\pm$ Odd	Even

Table 2.1: Addition/subtraction on even and odd numbers

2.5.2 Multiplication

Operations	Results
Even $\times$ Even	Even
Even $\times$ Odd	Even
Odd $\times$ Odd	Odd

Table 2.2: Multiplication on even and odd numbers

[8]

2.6 Sequences

2.6.1 Arithmetic Progression

An arithmetic sequence or progression is a sequence where the difference between 2 successive terms is fixed. [11]
The general formula is:

$$T(n) = a + (n-1)\upsilon$$

a: is the first term
υ: is the common difference
n: is the term needed

2.6.2 Recurrence Relation

A recurrence relation for a progression $a_0, a_1, \ldots,$ is a formula that connects the a_n term to the previous terms of the sequence $a_0, a_1, \ldots a_{n-1}$, provided that the initial values for certain terms in the sequence are given.

A common example of a recurrence is the Fibonacci sequence. The recurrence relation for Fibonacci sequence is as follows: [3]

$$f_n = f_{n-1} + f_{n-2}, \quad n \geq 3, \quad f_1 = f_2 = 1$$

2.7 Literature gap

So far, most research done for finding prime numbers, the initial point is to begin with 2 and then the designated algorithm comes into play. In other words, the algorithms used can find prime numbers up to a defined bound that is $p \leq x$.

The procedure in this thesis allows to find prime numbers within a range, for example finding all primes lying in the range $y \leq p \leq x$ provided y and x are both even integers. If the list of prime numbers required is $300 \leq p \leq 340$, only integers in this range will be dealt with, however in most algorithms seen, one has to compute for all numbers till 340 even though integers lesser than 300 will be discarded on the first go. We will see how to close the gap in the next coming few chapters.

Chapter 3

Prime numbers

3.1 Prime number and cake distribution

The idea of cake distribution is used to give an overview of how composite numbers occur in nature and in the process of forming composite numbers, prime numbers do not appear. Also, it is used for explaining why we stopped at the half-value. The notion of half-value may be unclear now of what is meant by that, but after (section 3.1), it would be clear.

Intuitively

Suppose some cakes have to be distributed among people and the number of people starts from 2 and keep increasing, i.e 2,3,4,... For 2 people, to satisfy a fair distribution, at least 2 cakes are needed, then the number increases by 2, that is 4,6,8 and so on. The same goes for 3. Starting from 3, the number of cakes increases by 3. Similarly, the other numbers get incremented accordingly. For every number of people, the minimum cakes required is the number of people itself, i.e for 7 people, minimum 7 cakes are required, and this is referred as the base. The base by itself only is not counted, instead it is summed up together with the next increment and further. For example, if there are 9 people, minimum 9 cakes are needed, but this only is not counted instead the base is counted together with the next and further increments which forms the numbers 18,27,36 and so on. Like so, if one does this till infinity for every integer, notice that prime numbers are never formed as their existence is not based upon other numbers, rather they exist by themselves. The base alone is ignored as it represents division of a number by itself. A number divided by itself only and the number 1 is a prime number and this a reason why the starting number of people is 2 and not 1. But the goal here is to find composite numbers.

The next two explanations, namely explanation 1 and 2, explain the idea of half value by keeping the above clarification in mind. The explanations 1 and 2 might seem repetitive but they show the relation and consistency with the above statement.

Explanation 1:

Consider a group of people receiving cakes.

Example: Suppose we distribute 8 cakes.

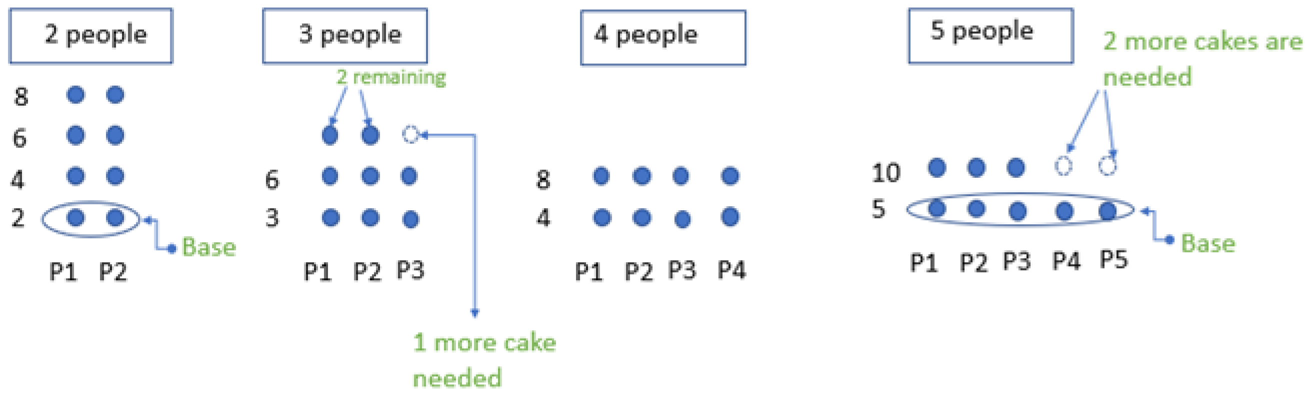

Figure 3.1: Distribution of 8 cakes

Figure 3.1 shows the distribution of cakes. It only shows up to 5 people, the rest [i.e numbers like 6,7 and 8] follow the same logic. The explanation below gives an idea what the diagram represents.

Consider distributing 8 cakes. In order to distribute the cakes, the number of people chosen begins from 2 [a good reason for it, is discussed in the paragraph just below and why not start with 1]. Starting from 2, the first row gives 2, second row gives a total of 4 and so on. The first row is the base, that is in order to distribute 2 cakes, minimum 2 people are needed. The base is ignored because of the definition of prime numbers.

For example: Having 2 cakes distributed among 2 people, in other words it means a number divide by itself, and prime numbers are numbers that divide by themselves and 1. But here the aim is to find composite numbers. Also 1 is ignored for the same reason.

The same procedure is followed using 3,4… people. Not in every case there exists even distribution, that is there are cakes that remain as leftovers in certain cases. To make the distribution fair, more cakes are needed to fill the gap. Notice that as soon as the half-way is surpassed [4 here], more cakes are needed to fill the second row for an even distribution. It is remarkable that the number will keep on increasing, for example in the case of 6 people, the number of cakes required is 12 so as not to be left with the base only and the cake demand increases as the number of people increases. Thus, the half-way is a guaranteed position for the number to occur [8 here]. Note that this applies to even numbers only. As we try to make the distribution fair and by avoiding the base only, it can be observed that numbers such as 2,3,5,7... do not appear since the base only is not considered and those numbers are prime, that is if we do not consider only the base [the first row as explained above why], but instead when we sum up all the numbers [to build up composite numbers] including the base, numbers like 2,3,5,7,... do not appear as a result of the sum.

Why the second row only?

Because it is only by the removal of the second row that we are left with the state where a number is divided by itself and this is not included.

Definition: Prime numbers are numbers that are not divisible by any other numbers except by the number itself and 1.

- Suppose there are 10 cakes and 10 people, the 10 cakes are divisible.
- It is also the case if there was only 1 person, he gets all the cakes.

The 2 statements above reflect the definition of prime numbers. But the aim here is to find all numbers that are not prime, and the missing ones are then prime. That is why division by 1 and the number itself is ignored.

Explanation 2:

The cake distribution represents division. Distributing 8 cakes between 2 people results in 4 cakes each and in terms of the diagram this represents 4 rows for each person separately. Likewise, if there are 3 people, each one obtains 2 cakes and we are left with 2 cakes, represented by 2 rows with equal distribution and 1 row with 2 leftovers. In other words, 8 divided by 3 gives 2 and a remainder of 2. One can choose to read the diagram column-wise as well, but the idea of distribution remains the same.

We notice 2 things here in order to satisfy a fair distribution:

- Either there are leftovers.
- Or we need more cakes.

Furthermore, we observe that as the number of people reaches more than half the number of cakes available, we need more cakes to satisfy the second-row distribution, but more cakes are not available. Thus, we stop at half-way as further division will not be possible. We have to consider leftovers after the half-way, because if we remove them to make the cake distribution fair, we will be left with the first row only and as seen earlier the distribution represents division and having the base as the only row represents a number divided by itself. Also, numbers after the half-value are not taken into account, because if the second row is not completely filled, it will result in numbers that can be prime when summing up. The reason the first row is ignored is based on the definition of prime numbers, where division by a number itself is not counted.

A further remark is when more cakes are added to satisfy a fair distribution, some numbers like 2,3,5,7... do not appear except on the first row [but first row is not considered as said above] and those numbers are prime.

We conclude 2 things here:

1. The half-value $x/2$ is a guaranteed position for x to occur. Summing over all the rows (Figure 3.1) gives x, where x is even.
2. For a fair distribution, prime numbers do not occur.

3.2 Prime numbers being non-composite numbers

Since prime numbers are not composite numbers, building a multiplication table from all possible combinations will give all composite numbers except prime numbers. Multiplication by 1 is not considered. In this way, prime numbers can be identified.

3.2.1 Remarks about the table

- Firstly, based on the explanation above regarding the half-value, the latter is where calculations start. This is the starting point.
- The table is read based on the diagonals. The patterns that appear and which are useful can be found on those diagonals.
- The values within each diagonal are incremented, for example: 20, 27, 32, 35.
- The values that cause the values within a diagonal to increment, they themselves however are decremented [for example 7, 5, 3, 1] by following an Arithmetic progression with a common difference of -2.

- The Arithmetic progression within a diagonal is either odd or even throughout that diagonal.
- Each diagonal follows either an even or odd Arithmetic progression and they alternate.
- The table is symmetrical, with values in the main diagonal being perfect squares. The lower part can be neglected.

Finding the term and the difference corresponding to the mid-value of 40 (Figure 3.2).

$T = 4 + (n - 1)$
$10 = 3 + n$
$n = 7$
$\therefore$ difference, d is 7 too since Term = difference. The difference here is the difference between the 1st and the 2nd value of the diagonal where the half-value occurs. Notice also that the difference is equal to the term, n. The term n and the difference d can be seen in dark blue in Figure 3.2.

The second value of the diagonal is then $10 \times 2 = 20 + difference \Rightarrow 20 + 7 = 27$
$n =$term
$d =$difference

The formula $T = 4 + (n - 1)$ can be simplified to $T = 3 + n$. Also, since the difference d is same as the term n, n can be simply replaced by d, thus giving the formula $T = 3 + d$. Since T is the half-value, it can be replaced by $\frac{x}{2}$ resulting in the following equation:

$$\frac{x}{2} = 3 + d \tag{3.1}$$

Note: d is not a common difference. Instead, it is the difference between the first two values of the half-value diagonal.

3.3 Possibilities for searching Prime numbers

There are 2 ways when searching for prime numbers from the table:

- Finding prime numbers between a range: $y < P < x$.

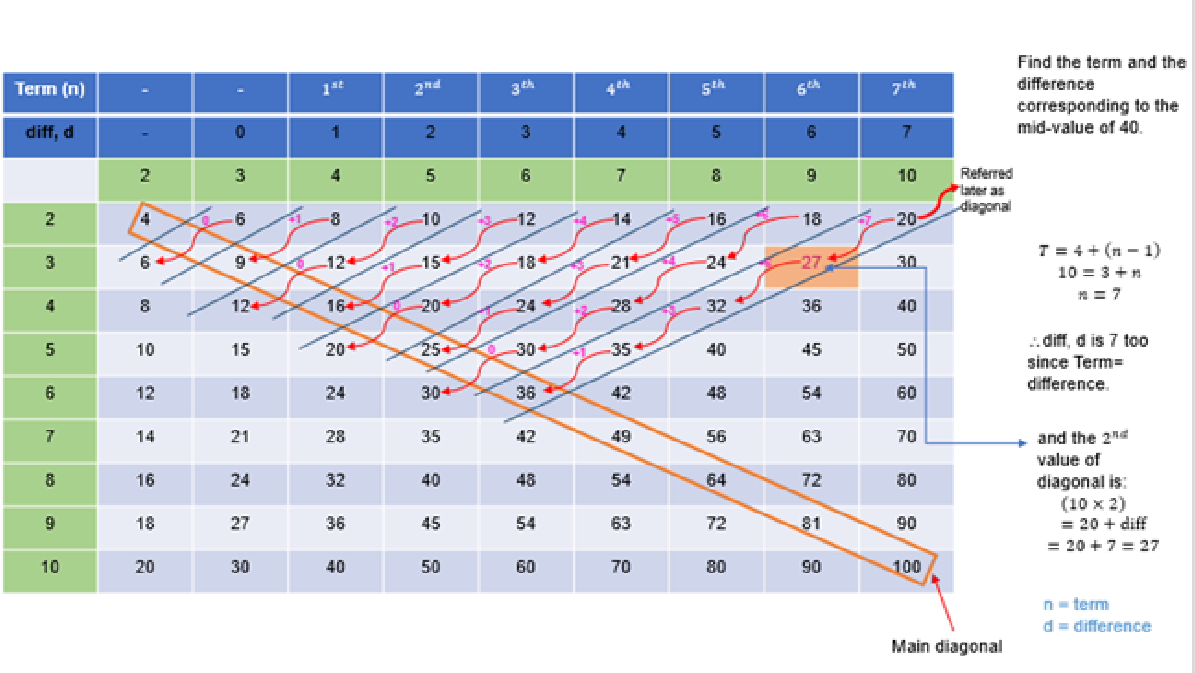

Figure 3.2: Multiplication grid for generating composite numbers

- Prime numbers less than a specific number: $P < x$.

In both cases, x and y have to be even numbers because:

1. The half-value of an even number is the maximum guaranteed position where that number will occur.

2. Odd numbers do not divide exactly by 2. So, the concept of half-value cannot be applied here.

3.3.1 Example for searching prime numbers less than a specific number $P < x$

For $P < 20$ (figure 3.2). The starting point is the half-value that is $\frac{20}{2} = 10$ (see the top green row of Figure 3.2).

Take note that numbers can be repeated in the table. Taking the half-value as the starting point is a guarantee that no numbers will be missed. Because if we were to start on the diagonal (14 18 20) where values within that diagonal are lesser than 20, clearly number like (16) from previous diagonal would be missed.

- Moving to the next diagonal.
  $\Rightarrow 10 - 1 = 9$

  For 10, the difference is less by 1 than what was calculated before [i.e 7], so the difference is 6.

  $\Rightarrow 9 \cdot 2 = 18, 18 + 6 = 24$ and $24 > 20$.

  The diagonal is neglected as the values are out of the interested range except for the first value, those are always considered.

- Moving to next diagonal
  $\Rightarrow 9 - 1 = 8$, for 8 the difference is less than by 1 than the previous diagonal.

  Therefore, the difference is 5; $8 \cdot 2 = 16; 16 + 5 = 21 > 20$ meaning the values get eliminated except for 16 which is the first value.

- Moving further,
  $\Rightarrow 8 - 1 = 7$; difference is 4; $7 \cdot 2 = 14$ and $14 + 4 = 18 < 20$ and therefore this diagonal is considered until a value larger than 20 is reached, but here it is not the case, thus the whole diagonal is considered.

From the starting diagonal on-wards, the Arithmetic progression is used; either even or odd in order to obtain all the values that are less than the number we are looking for (in this case, it is 20).

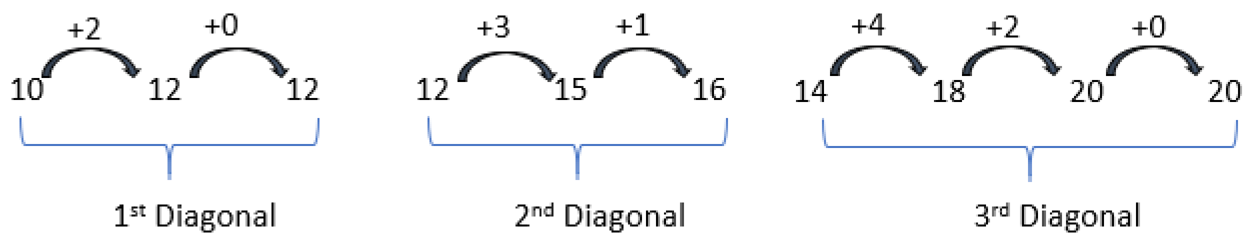

Figure 3.3: Illustration of the diagonal progressions

The missing numbers are 2,3,5,7,9,11,13,17,19 which are prime. The above procedure is a brief overview. Some computations for the remaining diagonals are not shown.

3.3.2 Rules to follow when searching prime numbers found between a range:

When searching for prime numbers lying in a range $y < P < x$, there is a checklist of procedures to be followed.

- Starting from $\frac{x}{2}$ [discussed earlier why], find out which diagonal has to be considered [same procedures as above in example 1(section 3.3.1)], but this time when crossing diagonals, not the whole diagonal is crossed out, instead only some values are eliminated while others are considered. Because there can be numbers lying on those diagonals that fall in the range between y and x. For example, consider a diagonal where x lies somewhere in between on the diagonal. Numbers after x will be greater than x and therefore ignored, whereas number before will be lesser, thus falling in range. Therefore, one has to compute the values lying on a diagonal until a value is reached such that it is larger than x. At this point, computation on that particular diagonal is stopped as numbers after will be larger than x and thus out of range.

- Keep moving to the left from $\frac{x}{2}$ until a diagonal is reached such that its last value is less than y [later in the thesis, we will see how to find the stopping point]. This ensures that there is no number in that diagonal as well as diagonals after that falls in the range between y and x. Also, that diagonal whose last value is less than y is not considered at all as the previous values are lesser than the last one. This indicates that we do not have to move to the next diagonal and search further.

- Starting from $\frac{x}{2}$, diagonals where the second number is greater than x are crossed out, since all the numbers coming after in that diagonal will be greater than x and therefore out of range. The second number is mentioned, because the first number of every diagonal up-to the half value is always in the range.

- One must not stop if the value y is found anywhere on a diagonal unless it is the last value. The stopping point must be such that the last value of the diagonal must be lesser than y. If one stops as soon as y is found on any diagonal at any position other than the last value, this is wrong as the values coming later after y will be greater than y and therefore found in the interested range.

- All numbers lying on the diagonals which were considered are listed down and the missing numbers are then prime.

Note: Upon eliminating the diagonals, the 1st value of every diagonal is considered as all those number from the half value to the stopping diagonal are found in the range .

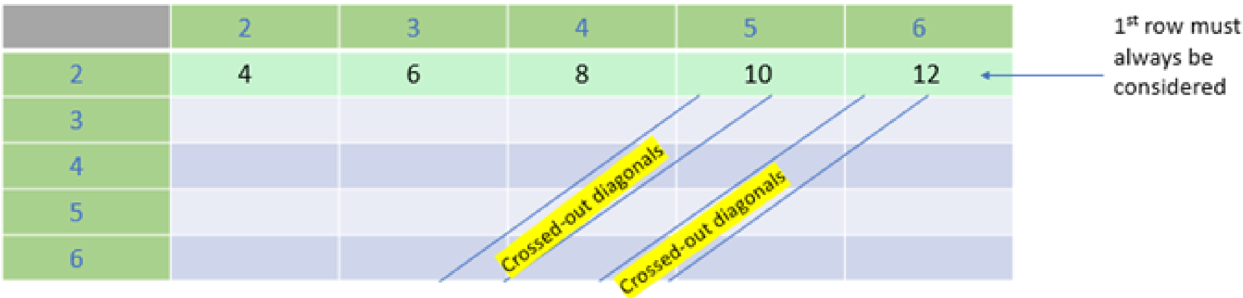

Figure 3.4: Crossed-out diagonals illustration

3.4 Proof regarding the end value of a diagonal and values in the next diagonal

Claim: There are no greater values than the end value of a diagonal that can be found in further diagonals[moving from right to left].

				N	N+1
	2	3	4	5	6
2	4	6	8	10	12
3	6	9	12	15	18
4	8	12	16	20	24
N=5	10	15	20	25	30
6	12	18	24	30	36

Figure 3.5: End value and next diagonal values relation

There is no number greater than 30 in the next diagonal

$30 = N(N+1) \Rightarrow N^2 + N = 30$

The end value for the next diagonal which is the largest value of the diagonal is $N(N) = N^2$; and $N^2 + N > N^2$ [Q.E.D].

This also proves the point that once the stopping diagonal has been identified with the last value less than y, we can be sure that there are no values in further diagonals which are in the range $y < P < x$. The last value of the next diagonal will be even lesser than that of the stopping the diagonal as well as the previous values on that diagonal. This applies to all the diagonals that follow after the stopping diagonal. So by no chance a number in the interested range will be missed.

Chapter 4

A more effective and quicker method

It is quite remarkable that as the number gets bigger, more computations are required, and it is time-consuming as well. To reduce computation cost, the following approach is used:

From the fact that all prime numbers are odd, diagonals that are incremented evenly are eliminated as the values on those diagonals are all even.

Observation: Notice that between 2 odd or 2 even diagonals there is a difference of 4. It is the difference between the first values of the 2 even or odd diagonal that is being referred here.

There are 2 possible cases, the difference d, calculated using the formula $\frac{x}{2} = d + 3$ is either an even number or odd.

4.1 Odd difference

Steps:

Suppose it is asked to find prime numbers between $x - 20 < P < x$. Note that the value $x - 20$ must be even as discussed earlier.

- Begin at the very first diagonal that increments oddly.
- Keep going down the diagonal, then move to the next odd diagonal and proceed the same.
- When to stop?
  Upon reaching a diagonal whose end value is smaller than $x - 20$, then stop. Here is the reason why: Consider the example $y < P < x$. If a diagonal is reached where the end value is less than y, then it is of no use to move to the next diagonal [when moving from right to left] as the values in the next diagonal will be even smaller and therefore out of range, that is smaller than y [see section 3.4].

  Do not confuse with previous method discussed as it was for the case $P < x$, while here it is $y < P < x$.

- A set of values is obtained which contains both odd and even numbers. Odd numbers will be in the form $x - b$ where b is odd i.e. b can be 1,3,5... and even numbers are of the form $x - a$ where a is even i.e. 0,2,4...

  Note: x is always even, thus:

  (i) $Even - even = even$

  (ii) $Even - odd = odd$

- From the set of values obtained in the previous step, the even ones are eliminated and out of the remaining numbers, the missing odd numbers are prime numbers. They should also be in range. For the example considered $x - 20 < P < x$, only those in between x and $x - 20$ should be considered, i.e. $x - 1, x - 3, x - 5, ..., x - 19$.

As a general explanation, the list [let's call it list 1] obtained after the above procedure will be of the form $x - odd_1, x - odd_2, x - odd_3, ...$ where odd_i can be 1,3,5,... depending on the range.

If another list [list 2] is constructed containing all the odd numbers found in the range, for example, a list of all odd numbers between x and $x - 20$ would look as given above [last bullet point].

Notice that all the numbers from list 1 are found in list 2. Those numbers that are present in list 2 but not in list 1, they are prime numbers and of course the list is based on the range we are looking for.

The 3 paragraph above is a general example on how prime numbers are obtained based on the lists.

Example 1:
Find prime $20 < P < 40$

- Highlighted values in pink are rejected as they are either greater or lesser than values within the range

- The values that are highlighted in yellow fall in the range:
  $20 < P < 40 \Rightarrow x - 20 < P < x$

Using the formula $T = d + 3$ with $T = \frac{x}{2}$, the difference d can be calculated.
$x = 40$, thus $\frac{40}{2} = d + 3 \Rightarrow d = 17$

- The set of values obtained are as follows:
  $E = \{x - 4, x - 8, x - 12, x - 16, x - 20, x - 8, x - 16\}$
  $O = \{x - 1, x - 7, x - 13, x - 5, x - 19, x - 15\}$
  $E \cup O =$
  $\{x-4, x-8, x-12, x-16, x-20, x-8, x-16, x-1, x-7, x-13, x-5, x-19, x-15\}$

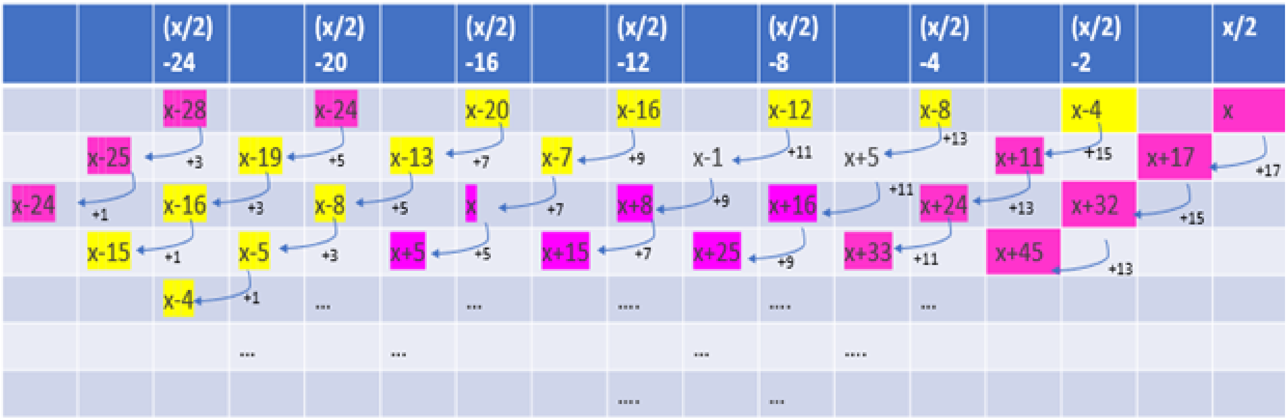

Figure 4.1: Pictorial representation for $20 < P < 40$

Only the set O is considered:

$O = \{x-1, x-7, x-13, x-5, x-19, x-15\}$

- The missing odd values that is within the range and not from the above set O are:
  1. $x - 3 \Rightarrow 40 - 3 = 37$
  2. $x - 9 \Rightarrow 40 - 9 = 31$
  3. $x - 11 \Rightarrow 40 - 11 = 29$
  4. $x - 17 \Rightarrow 40 - 17 = 23$

All the above numbers are prime numbers.

Example 2:
Below is another example with some skipped steps.
Find $60 < P < 80$.

- The list of values is as follows:

$R = E \cup O = \{x-5, x-11, x-17, x-8, x-16, x-15, x-3\}$ The missing odd values are:

1. $x - 1 \Rightarrow 80 - 1 = 79$
2. $x - 7 \Rightarrow 80 - 7 = 73$
3. $x - 9 \Rightarrow 80 - 9 = 71$
4. $x - 13 \Rightarrow 80 - 13 = 67$
5. $x - 19 \Rightarrow 80 - 19 = 61$

All the above values are prime.

							...		...		...		...		...		...		...		x/2
					x-50		x-52		x-48		x-44		x-40		x-36		x-32		x-28		x
				x-47		x-41		x-35		x-29		x-23		x-17		x-11		x-5		x+37	
			x-40		x-32		x-24		x-16		x-8		x		x+8		x+16				
		...		x-25		x-15		x-5		x+5		x+15		...							
	....		x-20		x-8		x-4		...		...		....		....		....				
		x-17		x-3				...		...				...		....					
	x-16												....		....						

Figure 4.2: Example for $60 < P < 80$

4.1.1 Finding a general formula for odd difference

Diagonal 1:
$x, x + d, x + d + (d - 2), x + d + (d - 2) + (d - 4), x + (d - 2) + (d - 4) + (d - 6), \ldots$

$\Rightarrow x, x + d, x + 2d - 2, x + 3d - 6, x + 4d - 12, x + 5d - 20, x + 6d - 30, \ldots$

$\Rightarrow f_{o,1}(n) = x + nd - n(n - 1); n \geq 0$

Diagonal 2:
$x - 4, x - 4 + (d - 2), x - 4 + (d - 2) + (d - 4), x - 4 + (d - 2) + (d - 4) + (d - 6), \ldots$

$\Rightarrow x - 4, x + d - 6, x + 2d - 10, x + 3d - 16, \ldots$

$\Rightarrow f_{o,2}(n) = x - 4 + nd - n(n + 1); n \geq 0$

Diagonal 3: $x - 8, x - 8 + (d - 4), x - 8 + (d - 4) + (d - 6), \ldots$

$\Rightarrow x - 8, x + d - 12, x + 2d - 18, \ldots$

$\Rightarrow f_{o,3}(n) = x - 8 + nd - n(n + 3); \geq 0$

Combining all together gives:

$$\boxed{F_o(R, n) = x - 4(R - 1) + nd - n(n + 2R - 3); R \geq 1, n \geq 0} \tag{4.1}$$

R: is the diagonal
n: is the nth term on the diagonal

d: is the difference obtained using the equation $\frac{x}{2} = d + 3$

4.1.2 Proof of correctness for the odd diagonals

- Diagonal 1:$x, x+d, x+d+(d-2), x+d+(d-2)+(d-4), x+(d-2)+(d-4)+(d-6), \ldots$

  $\Rightarrow x, x+d, x+2d-2, x+3d-6, x+4d-12, x+5d-20, \ldots$

  $$f_{o,1}(n) = x + nd - n(n-1); n \geq 0, n \leq h$$

  where h gives the number of elements in the diagonal.

  $$f_{o,1}(n+1) = f_{o,1}(n) + d - 2(n)$$

  $$f_{o,1}(0) = x$$

  $$f_{o,1}(1) = f_{o,1}(0) + d - 2(n) \Rightarrow f_{o,1}(0) + d + 0 \Rightarrow x + d$$

  $$f_{o,1}(2) = f_{o,1}(1) + d - 2 \Rightarrow x + d + d - 2 \Rightarrow x + 2d - 2$$

  $$f_{o,1}(3) = f_{o,1}(2) + d - 4 \Rightarrow x + 2d - 2 + d - 4 \Rightarrow x + 3d - 6$$

  Now to prove that $f_{o,1}(n) = x + nd - n(n-1)$ for all $n \geq 0$.

  Base case is true: $n = 0, f_{o,1}(0) = x + 0(d) - 0(0-1) = x$.

  Induction assumption: Suppose that $f_{o,1}(n) = x + nd - n(n-1)$, then

  Induction step:

  $$f_{o,1}(n+1) = f_{o,1}(n) + d - 2(n)$$

$$= x + nd - n(n-1) + d - 2(n)$$

$$= x + nd - n(n-1) + d - 2n$$

$$= x + nd - n^2 + n + d - 2n$$

$$= x + d(n+1) - n^2 - n$$

$$= x + d(n+1) - n(n+1)$$

And now consider $f_{o,1}(n) = x + nd - n(n-1)$, substituting n by $n+1$ gives:

$f_{o,1}(n+1) = x + (n+1)d - (n+1)[(n+1) - 1]$

$\Rightarrow x + d(n+1) - n(n+1)$

[QED]

- Diagonal 2:
  $x-4, x-4+(d-2), x-4+(d-2)+(d-4), x-4+(d-2)+(d-4)+(d-6), \ldots$

  $\Rightarrow x-4, x+d-6, x+2d-10, x+3d-16, \ldots$

$$f_{o,2}(n) = x - 4 + nd - n(n+1); n \geq 0, n \leq h$$

$$f_{o,2}(0) = x - 4$$

$$f_{o,2}(n+1) = f_{o,2}(n) + d - 2(n+1)$$

$$f_{o,2}(0) = x - 4$$

$$f_{o,2}(1) = f_{o,2}(0) + d - 2(n+1) \Rightarrow f_{o,2}(0) + d - 2(n+1) \Rightarrow x - 4 + d - 2 \Rightarrow x + d - 6$$

$$f_{o,2}(2) = f_{o,2}(1) + d - 2(1+1) \Rightarrow x + d - 6 + d - 4 \Rightarrow x + 2d - 10$$

$$f_{o,2}(3) = f_{o,2}(2) + d - 2(2+1) \Rightarrow x + 2d - 10 + d - 6 \Rightarrow x + 3d - 16$$

Now proving that $f_{o,2}(n) = x - 4 + nd - n(n+1)$ for all $n \geq 0$.

Base case is true: $n = 0, f_{o,2}(0) = x - 4 + 0(d) - 0(0+1) = x - 4$.

Induction assumption: Suppose that $f_{o,2}(n) = x - 4 + nd - n(n+1)$, then

Induction step:

$$f_{o,2}(n+1) = f_{o,2}(n) + d - 2(n+1)$$

$$= x - 4 + nd - n(n+1) + d - 2(n+1)$$

$$= x - 4 + nd - n^2 - n + d - 2n - 2$$

$$= x + d(n+1) - n^2 - 3n - 6$$

$$= x - 6 + d(n+1) - n(n+3)$$

And now consider $f_{o,2}(n) = x - 4 + nd - n(n+1)$, substituting n by $n+1$ gives:

$$f_{o,2}(n+1) = x - 4 + (n+1)d - (n+1)[(n+1)+1]$$

$$= x - 4 + d(n+1) - [(n+1)(n+2)]$$

$$= x - 4 + d(n+1) - (n^2 + 3n + 2)$$

$$= x - 4 + d(n+1) - n^2 - 3n - 2$$

$$= x - 6 + d(n+1) - n(n+3)$$

$[QED]$

Analogously, the same process is done for the other diagonals.

4.2 Even difference

4.2.1 Finding a general formula for even difference

When the difference is odd, one starts counting the starting point from the half-value diagonal. But in the case of even difference, to find the starting diagonal, the starting point for counting the diagonal is not the half-value diagonal. Instead the starting point for counting the diagonal begins one diagonal just after the half-value. It is because to avoid counting even diagonals as starting on the even diagonal will result in even numbers only since adding the even differences to even numbers [numbers on the diagonal] will result in even and prime numbers are odd, thus resulting in useless calculation. What is meant by "counting the starting point" is discussed in section (5.2). There is it is observed that not every time the starting diagonal is the half-value diagonal. One has to find the value of R for that.

Diagonal 1: $f_{e,1}(n) = x - 2 + nd - n(n); n \geq 0$

Diagonal 2: $f_{e,2}(n) = x - 6 + nd - n(n+2); n \geq 0$

Diagonal 3: $f_{e,3}(n) = x - 10 + nd - n(n+4); n \geq 0$

Diagonal 4: $f_{e,4}(n) = x - 14 + nd - n(n+6); n \geq 0$

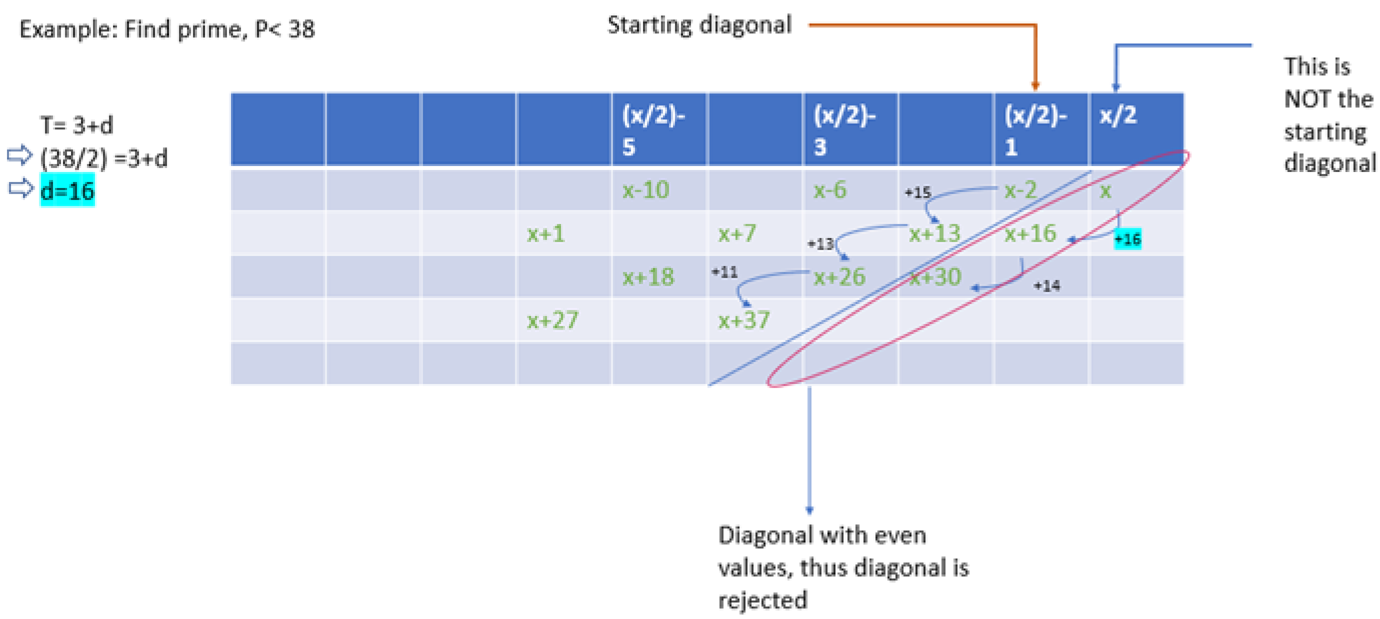

Figure 4.3: A graphical representation for the starting diagonal for even difference

Combining all together gives:

$$\boxed{F_e(R,n) = x - 2(2R-1) + nd - n(n+2R-2); R \geq 1, n \geq 0} \tag{4.2}$$

Remarks:
Even though the difference is even, the searched values are still of the form $[x - odd]$. The formula gives any number on any diagonal, so by substituting the proper value, the concerning numbers are retrieved [interested numbers are odd numbers]. Thus, whether the difference is odd or even, it does not matter as long as the proper values are substituted. In chapter 5 we will see how values of n can affect the result obtained. That's why it is explicitly mentioned to substitute proper values.

4.2.2 Proof of correctness for the even diagonals

- Diagonal 1: $x-2, x+d-3, x+2d-6, x+3d-11, x+4d-18, \ldots$

  $$f_{e,1}(n) = x - 2 + nd - n(n); n \geq 0, n \leq h$$

  where h gives the number of elements in the diagonal.

  $$f_{e,1}(n+1) = f_{e,1}(n) + d - (2n+1)$$

  $$f_{e,1}(0) = x - 2$$

$$f_{e,1}(1) = f_{e,1}(0) + d - (2n+1) \Rightarrow f_{e,1}(0) + d - (0+1) \Rightarrow x + d - 3$$

$$f_{e,1}(2) = f_{e,1}(1) + d - (2+1) \Rightarrow x + d - 3 + d - 3 \Rightarrow x + 2d - 6$$

$$f_{e,1}(3) = f_{e,1}(2) + d - (4+1) \Rightarrow x + 2d - 6 + d - 5 \Rightarrow x + 3d - 11$$

Now to prove that $f_{e,1}(n) = x - 2 + nd - n(n)$ for all $n \geq 0$.

Base case is true: $n = 0, f_{e,1}(0) = x - 2 + 0(d) - 0(0) = x - 2$.

Induction assumption: Suppose that $f_{e,1}(n) = x - 2 + nd - n(n)$, then
Induction steps:

$$f_{e,1}(n+1) = f_{e,1}(n) + d - (2n+1)$$

$$= x - 2 + nd - n^2 + d - 2n - 1$$

$$= x + d(n+1) - n^2 - 2n - 3$$

$$= x - 3 + d(n+1) - n(n+2)$$

And now consider $f_{e,1}(n) = x - 2 + nd - n(n)$, substituting n by $n+1$ gives:

$$f_{e,1}(n+1) = x - 2 + (n+1)d - (n+1)^2$$

$$= x - 2 + d(n+1) - (n^2 + 2n + 1)$$

$$= x - 2 + d(n+1) - n^2 - 2n - 1$$

$$= x - 3 + d(n+1) - n(n+2)$$

[QED]

- Diagonal 2: $x-6, x+d-5, x+2d-6, x+3d-9, x+4d-14, \ldots$

$$f_{e,2}(n) = x - 6 + nd - n(n-2); n \geq 0, n \leq h$$

$$f_{e,2}(n+1) = f_{e,2}(n) + d - (2n-1)$$

$$f_{e,2}(0) = x - 6$$

$$f_{e,2}(1) = f_{e,2}(0) + d - (2n+1) \Rightarrow f_{e,2}(0) + d - (0-1) \Rightarrow x + d - 5$$

$$f_{e,2}(2) = f_{e,2}(1) + d - (2-1) \Rightarrow x + d - 5 + d - 1 \Rightarrow x + 2d - 6$$

$$f_{e,2}(3) = f_{e,2}(2) + d - (4-1) \Rightarrow x + 2d - 6 + d - 3 \Rightarrow x + 3d - 9$$

Now to prove that $f_{e,2}(n) = x - 6 + nd - n(n-2)$ for all $n \geq 0$

Base case is true: $n = 0, f_{e,2}(0) = x - 6 + 0(d) - 0(0-2) = x - 6$

Induction assumption: Suppose that $f_{e,2}(n) = x - 6 + nd - n(n-2)$, then
Induction steps:

$$f_{e,2}(n+1) = f_{e,2}(n) + d - (2n-1)$$

$$= x - 6 + nd - n^2 + 2n + d - 2n + 1$$

$$= x - 5 + d(n+1) - n^2$$

Now consider $f_{e,2}(n) = x - 6 + nd - n(n-2)$. Substituting n by $n+1$ gives:

$$f_{e,2}(n+1) = x - 6 + (n+1)d - (n+1)[(n+1) - 2]$$

$$= x - 6 + d(n+1) - (n^2 - n + n - 1)$$

$$= x - 6 + d(n+1) - n^2 + 1$$

$$= x - 5 + d(n+1) - n^2$$

[QED]

Similarly, the other diagonals can be proven.

4.3 Why d is not substituted by $\frac{x}{2} - 3$?

From the equation $\frac{x}{2} = d + 3$, if d is made the subject of formula the equation becomes $d = \frac{x}{2} - 3$ and can thus be substituted in both equations (4.1) and (4.2). But that's not the case, instead it is replaced by the value calculated for it. Why?

Because when substituting d in either equation (4.1) or (4.2) , it turns out that the answer obtained is no more in the form $x - odd_i$, instead it is in the form $kx - odd_i$ where k is an integer other than 1.

Consider [the example $60 < x < 76$ from section 6.1], after substituting n by 1 and d by $\frac{x}{2} - 3$, the result obtained is $\frac{3x}{2} - 39$ instead of $x - 1$. Though both gives the same answer 75 when replacing x by 76, $\frac{3x}{2} - 39$ cannot be easily identified from the list of odd numbers obtained [in step 5, chapter 4.1] unless x is substituted.

The form $x - odd_i$ is desirable since a list of odd numbers is obtained and from that list, the missing prime numbers can be easily identified. But with the constant other than 1 being added in front of x, the missing odd numbers cannot be found by just cross checking with the list.

If by substituting d, the constant did not appear in the result, this would result in a better algorithm since computation speed would increase and also the amount of space required for saving d would be minimized.

Chapter 5

Reducing computation

To avoid unnecessary computations on a diagonal, for example, numbers that do not fall in range add to computation cost, it is important to find out the starting, stopping point and avoid even numbers.

To begin with, based on whether the difference is odd or even, the appropriate formula is chosen:

Odd: $F_o(R, n) = x - 4(R - 1) + nd - n(n + 2R - 3); R \geq, n \geq 0$

Even: $F_e(R, n) = x - 2(2R - 1) + nd - n(n + 2R - 2); R \geq 1, n \geq 0$

5.1 Finding stopping point on a diagonal

When going down a diagonal [$direction$:$\swarrow$] , the number increases. So whenever performing trial and error operation [see section 5.1.1], as soon as the constant c becomes positive $x + c$, further numbers do not need to be calculated as they will be greater.

When a negative constant $x - c$ is repeated twice consecutively [see in (section 6: Full examples), where $x - 13$ is repeated twice], this indicates that the lower half of the table has been reached that is the part below the main diagonal. Thus, further calculations on that diagonal will just add to unnecessary computation cost. No doubt positive constant also does occur but when it comes to positive constant, calculation is halted upon the **first** positive constant as it will be out of range, as well as the numbers that follow after.

5.1.1 Case 1: Odd difference

Example 1: $20 < P < 40$.

Rewriting $\Rightarrow x - 20 < P < x; x = 40$.

- The aim is to find which values of n when substituted by trial and error results in values less than x and greater than $x - 20$.

For $x = 40, d = 17$. Using the formula $T = d + 3$ with $T = \frac{x}{2}$, the difference d can be calculated as follows: $x = 40$, thus $\frac{40}{2} = d + 3 \Rightarrow d = 17$.

1. Consider diagonal 2: $f_2(n) = x - 4 + nd - n(n-1); n \geq 0$, this gives:
   $x - 4 + 17n - n(n+1) < x$

   $$\Rightarrow -n^2 + 16n - 4 < 0 \tag{5.1}$$

   Trial and error when replacing different values of n in (5.1):

   $n = 0 \Rightarrow -4 < 0$ [accept]$\Rightarrow x - 4$

   $n = 1 \Rightarrow 11 > 0$ [reject]$\Rightarrow x + 11$ [rejected because it is greater than x and thus falls out of range]

2. Consider diagonal 3:$f_3(n) = x - 8 + nd - n(n+3); n \geq 0$, substituting d and equating gives:

   $$\Rightarrow x - 4 + 17n - n(n+3) < x \Rightarrow -n^2 + 14n - 8 < 0 \tag{5.2}$$

   Trial and error when replacing different values of n in (5.2):

   $n = 0 \Rightarrow -8 < 0$ [accept]$\Rightarrow x - 8$

   $n = 1 \Rightarrow 5 > 0$ [reject]$\Rightarrow x + 5$ [rejected because it is greater than x and thus falls out of range]

3. Now consider diagonal 8: $f_8(n) = x - 28 + nd - n(n+13); n \geq 0$, substituting d and equating gives:

   $$x - 28 + 17n - n(n+13) < x \Rightarrow -n^2 + 4n - 28 < 0 \tag{5.3}$$

   Trial and error when replacing different values of n in (5.3):

   $n = 0 \Rightarrow -28 < 0$ [reject]$\Rightarrow x - 28$ [rejected because it is less than y and thus falls out of range]

   $n = 1 \Rightarrow -25 < 0$ [reject]$\Rightarrow x - 25$ [rejected because it is less than y and thus falls out of range]

- Some of the values above are rejected even though they are less than 0 [condition imposed by equations 5.1, 5.2, 5.3] because they are out of range. The range is $x - 20 < P < x$, $x - 28$ and $x - 25$ are lesser than $x - 20$, thus rejected. An advantage is that it is not necessary to solve the equation for both less than x and then greater than $x - 20$. It is sufficient only to compare the constant obtained after performing the calculations as shown above with the range, and then decide whether to accept or reject.

5.1.2 Choices for n in case of odd difference

For an odd difference, there are 2 choices for n, the latter can be odd or even. Below is a further analysis.

e represents even.
o represents odd.

1. n is Even.

$$
\begin{array}{lllllll}
F_o(R,n) & = & x & -\overbrace{4(R-1)}^{e} & +\overbrace{nd}^{e\cdot o} & -\overbrace{n(n+2R-3)}^{e(e\mp o)}, & R \geq 1, \quad n \geq 0 \\
\Rightarrow & = & x & -e & +e\cdot o & -\overbrace{e(e\mp o)}^{e\cdot o} & \\
\Rightarrow & = & x & -e & +\overbrace{e\cdot o}^{e} & \mp\overbrace{e\cdot o}^{e} & \\
\Rightarrow & = & x & -e & +\overbrace{e\mp e}^{e} & & \\
\Rightarrow & = & x & -\overbrace{e\mp e}^{e} & & & \\
\Rightarrow & = & x & \mp e & & &
\end{array}
$$

2. n is Odd.

$$
\begin{array}{lllllll}
F_o(R,n) & = & x & -\overbrace{4(R-1)}^{e} & +\overbrace{nd}^{o\cdot o} & -\overbrace{n(n+2R-3)}^{o(o\mp o)}, & R \geq 1, \quad n \geq 0 \\
\Rightarrow & = & x & -e & +o\cdot o & -\overbrace{o(o\mp o)}^{o\cdot e} & \\
\Rightarrow & = & x & -e & +\overbrace{o\cdot o}^{o} & \mp\overbrace{o\cdot e}^{e} & \\
\Rightarrow & = & x & -e & +\overbrace{o\mp e}^{o} & & \\
\Rightarrow & = & x & -\overbrace{e\mp o}^{o} & & & \\
\Rightarrow & = & x & \mp o & & &
\end{array}
$$

Difference	n	result
odd	even	x-even [reject]
odd	odd	x-odd [accept]

Table 5.1: Possible values of n for odd difference

- $x - even \Rightarrow even$, because note that x also is even.
- $x - odd \Rightarrow odd$, this one is considered as it is odd, and the aim is to look for odd numbers only.

5.1.3 Case 2: Even difference

The same procedures are followed when the difference is even. Below is a short example.

Example 1: $28 < P < 38$, difference $d = 16$.

Note: $d = 16$ is the difference for the 1st diagonal [starting point] between the first two values of that diagonal only. The difference is always between the 1st two values of the starting diagonal. For the rest calculation, it follows arithmetic progression.

Considering diagonal 2: $f_2(n) = x - 6 + nd - n(n-1); n \geq 0$, substituting d and equating it to less than x gives:

$$x - 6 + 16n - n^2 - 2n < x \Rightarrow -n^2 + 14n - 6 < 0 \tag{5.4}$$

Trial and error when replacing different values of n in (5.4):

$n = 0 \Rightarrow -6 < 0$ [accept]$\Rightarrow x - 6$

$n = 1 \Rightarrow 7 < 0$ [reject]$\Rightarrow x + 7$ [rejected because it is greater than x and thus falls out of range.

5.1.4 Possibilities of n for even difference

As for the case of odd difference there were 2 possibilities of choosing n, the same idea goes here.

1. n is Even

$$
\begin{aligned}
F_e(R,n) &= x \quad -\overbrace{2(2R-1)}^{e} \quad + \overbrace{nd}^{e\cdot e} \quad -\overbrace{n(n+2R-2)}^{e(e\mp e)}, \quad R\geq 1, \quad n\geq 0\\
\Rightarrow \quad &= x \quad -e \quad +e\cdot e \quad -\overbrace{e(e\mp e)}^{e\cdot e}\\
\Rightarrow \quad &= x \quad -e \quad +\overbrace{e\cdot e}^{e} \quad \mp\overbrace{e\cdot e}^{e}\\
\Rightarrow \quad &= x \quad -e \quad +\overbrace{e\mp e}^{e}\\
\Rightarrow \quad &= x \quad -\overbrace{e\mp e}^{e}\\
\Rightarrow \quad &= x \quad \mp e
\end{aligned}
$$

2. n is Odd

$$
\begin{aligned}
F_e(R,n) &= x \quad -\overbrace{2(2R-1)}^{e} \quad + \overbrace{nd}^{o\cdot e} \quad -\overbrace{n(n+2R-2)}^{o(o\mp e)}, \quad R\geq 1, \quad n\geq 0\\
\Rightarrow \quad &= x \quad -e \quad +o\cdot e \quad -\overbrace{o(o\mp e)}^{o\cdot o}\\
\Rightarrow \quad &= x \quad -e \quad +\overbrace{o\cdot e}^{e} \quad \mp\overbrace{o\cdot o}^{o}\\
\Rightarrow \quad &= x \quad -e \quad +\overbrace{e\mp o}^{o}\\
\Rightarrow \quad &= x \quad -\overbrace{e\mp o}^{o}\\
\Rightarrow \quad &= x \quad \mp o
\end{aligned}
$$

Difference	n	result
even	even	x-even [reject]
even	odd	x-odd [accept]

Table 5.2: Possible values of n for even difference

The same idea applies here as above for rejecting $x - even$ and accepting $x - odd$.

5.2 Finding starting point, i.e on which diagonal to start

Notice that not every time the starting point is the 1st or 2nd diagonal. Generally, the values on the 1st, 2nd or the diagonal after [moving from $\frac{x}{2}$ to the left] are larger than the range. For example, consider $y < P < x$, in many cases the values are larger than x when it comes to starting on the 2nd or 3rd diagonal. Thus, starting on diagonals just after the half-value can result in useless and extra computation. Therefore, finding the starting diagonal can be considered as pre-processing.

5.2.1 Case 1: Odd difference

The idea is to find R such that when replacing $n = 1$ in $F_o(R, n) = x - 4(R-1) + nd - n(n+2R-3)$, $F_o(R, n)$ is less than x. That is, $F_o(R, n) < x$.

- Why $n = 1$?

  Because the goal is to get the 2nd value of diagonal which is lesser than x. Take note that the 1st value of every diagonal starts with $n = 0$, but because the 1st value [i.e. when $n = 0$] is even, that is why it is ignored.

Solution:

$$x - 4(R-1) + nd - n(n+2R-3) < x; n = 1 \tag{5.5}$$

From (3.1) $\frac{x}{2} = 3 + d$, making d the subject formula, the equation becomes:

$$\Rightarrow d = \frac{x}{2} - 3 \tag{5.6}$$

substituting (5.6) in (5.5) gives:

$\Rightarrow x - 4R + 4 + \frac{x}{2} - 3 - 1 - 2R + 3 < x$

$\Rightarrow -6R + \frac{x}{2} + 3 < 0$

The starting diagonal for odd difference can be found using:

$$\boxed{-6R + \frac{x}{2} + 3 < 0} \tag{5.7}$$

Example 1: For $x = 80, d = 37$.

Solving for R gives: $R > 7.166\ldots, \Rightarrow R = 8$ [see table in figure 5.1 to cross check].

Notice that the answer obtained is rounded up to the next natural number as the diagonals of the tables are natural numbers.

The starting point for counting up-to $R = 8$ which is the starting diagonal for the case of odd is the half value[$\frac{80}{2} = 40$ in this example] and it is clearly seen that when $R = 8$, the value considered [75 here encircled in red] on that diagonal is lesser than x. The first row is neglected as all the values are even on that row, that is why 52 is not considered.

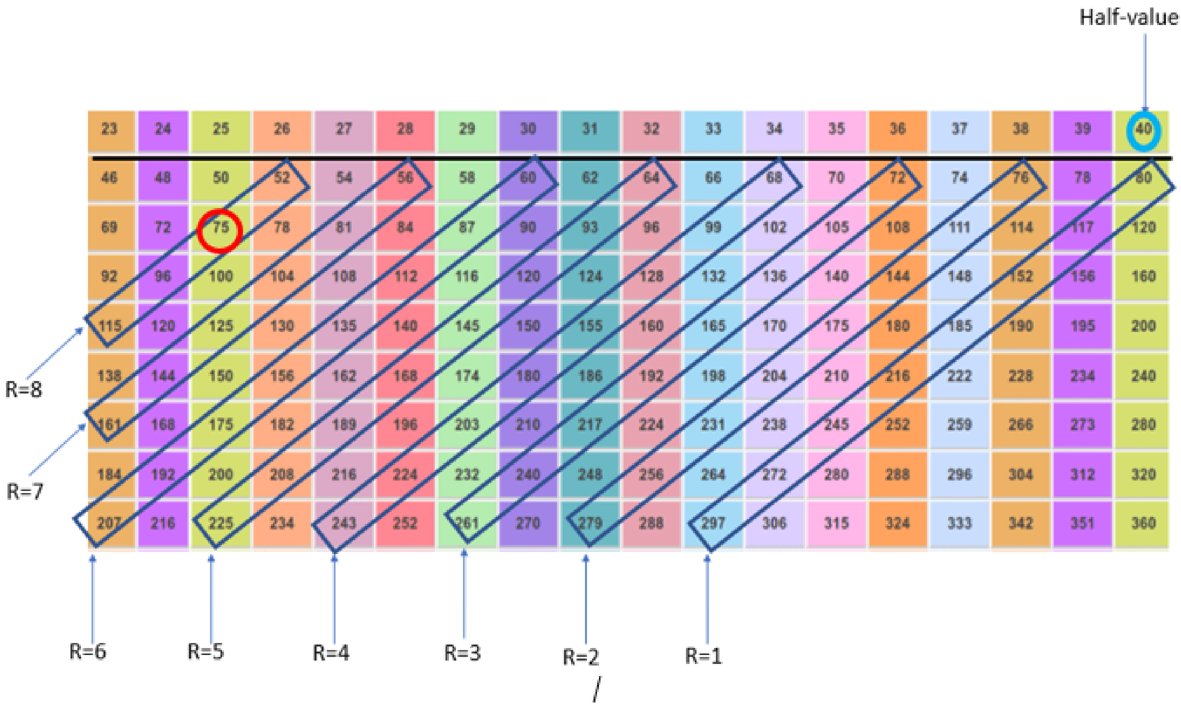

Figure 5.1: Illustration for the starting diagonal when x is 80

Example 2: For $x = 40, d = 17, R > 3.8\ldots, \Rightarrow R = 4$

Example 3: For $x = 60, d = 27, R > 5.5\ldots, \Rightarrow R = 6$

The same idea applies for the above examples 2, 3 and for others with odd differences. For the above examples 2 and 3, cross check with their tables respectively.

5.2.2 Case 2: Even difference

Finding R such that $F_e(R, n) < x$:

$$x - 2(2R - 1) + nd - n(n + 2R - 2) < x \tag{5.8}$$

From (3.1) $\frac{x}{2} = 3 + d$, making d the subject formula, the equations becomes:

$$\Rightarrow d = \frac{x}{2} - 3 \tag{5.9}$$

substituting (5.9) in (5.8) gives:

$\Rightarrow x - 4R + 2 + \frac{x}{2} - 3 - 1 - 2R + 2 < x$

$\Rightarrow -6R + \frac{x}{2} < 0$

For even difference, starting diagonal can be calculated using:

$$\boxed{-6R + \frac{x}{2} < 0} \tag{5.10}$$

Examples:

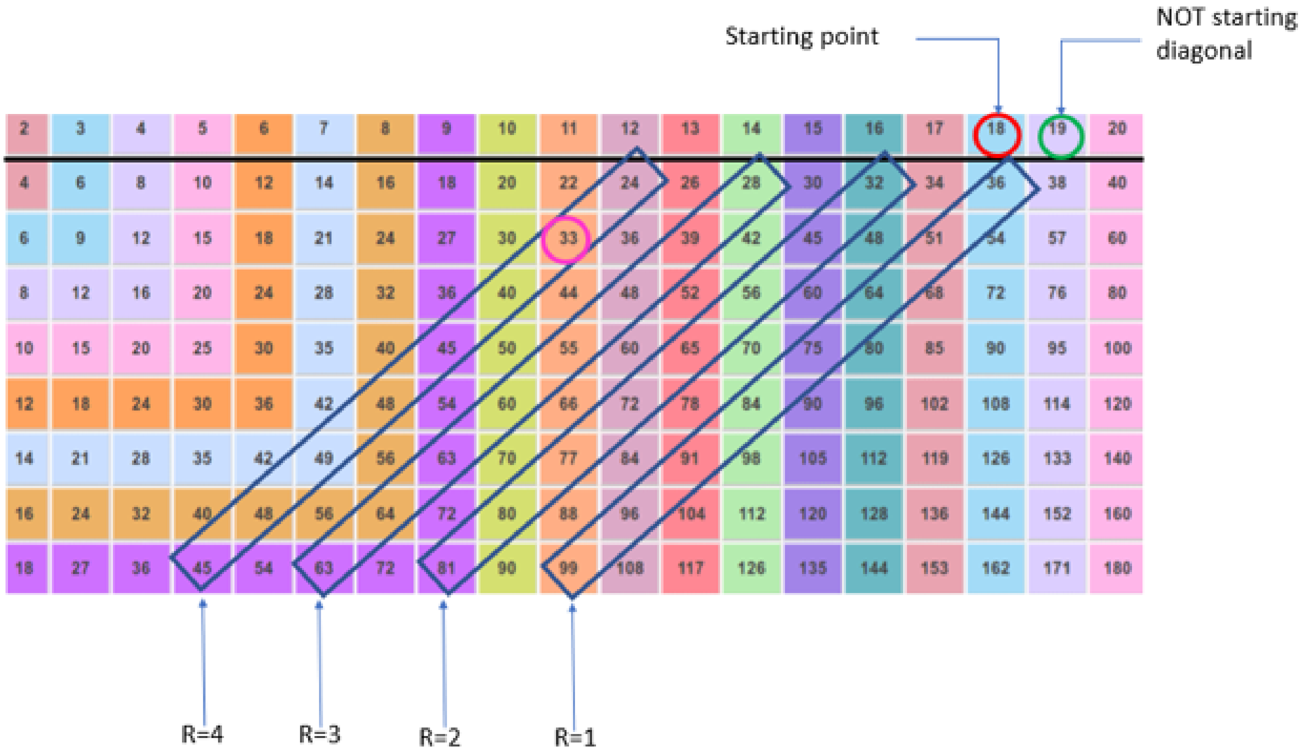

Figure 5.2: Illustration for the starting diagonal when x is 38

1. For $x = 38, R > 3.16\ldots, \Rightarrow R = 4$

   From figure 5.2, it can be observed that the starting point is not the half-value, 19 i.e. $\frac{38}{2}$, instead it is 18 to avoid even diagonal. Then the rest is like the odd case.

2. For $x = 78 \Rightarrow R > 6.5\ldots, R = 7$

3. For $x = 46 \Rightarrow R > 3.8\ldots, R = 4$

4. For $x = 66 \Rightarrow R > 5.5\ldots, R = 6$

5. For $x = 70 \Rightarrow R > 5.8\ldots, R = 6$

Note: The starting point encircled with the label "starting point" in figure 5.2 is actually the beginning point for counting R and not where calculations begin.

5.3 Finding on which diagonal to stop in terms of R:

Previously, [section 5.1: Finding stopping point on a diagonal], the stopping point was about where to stop on a particular diagonal, but this time the aim is to find which diagonal is the stopping diagonal. It is useful when implemented, where each diagonal can be computed independently, thus finding on which diagonal to stop serves as a condition for the algorithm to stop instead of going to the next diagonal which is of no purpose.

Consider the same problem $y < P < x$.

The idea is to find which diagonal's last value is lesser than y. That diagonal is rejected as the previous values in it will be even lesser. See section 3.4.

Goal: Find R such that $F(R, n) < y$. n must be calculated and then substituted in $F(R, n)$ such that it is the last value. The following formula [Arithmetic Progression formula] is used for finding n:

$$\boxed{T(n) = a + (n-1)\upsilon} \tag{5.11}$$

a is the difference between the two starting numbers of a diagonal. That is a is the 1st term of the sequence $\ldots$,9,7,5,$\ldots$. It can be any diagonal based on the value of R. Note that a depends on which diagonal is being considered and not the difference that is calculated using the formula $d = \frac{x}{2} - 3$. Here υ is the common difference having a value of -2.

5.3.1 Case 1: Odd difference

$F_o(R, n) = x - 4(R-1) + nd - n(n + 2R - 3)$.

As seen earlier, there is an arithmetic progression in the diagonal and in order to find the value of n, the values of the arithmetic progression are taken into account and not the values lying on the diagonal. The odd arithmetic progression follows the following pattern: $\ldots$,9,7,5,3,1 with the last term being 1 and the progression has a common difference of -2.

For odd difference, a_o is given by: $\boxed{\frac{x-4(R-1)}{2}-3}$

a_o is obtained as follows:

- Starting from the half-value diagonal, the 1st value of the diagonal is x.
- To move to further diagonal, one has to subtract 4 iteratively based on the diagonal. This gives $x - 4(R-1), R \geq 1$.
- To obtain the difference between the 1st two values of a diagonal depending on x, the result from the second step is divided by 2 and then subtract 3 from it giving: $\frac{x-4(R-1)}{2} - 3$. Refer to diagram (3.2) and equation (3.1) for better understanding why a division by 2 and a subtraction of 3 is performed.

Substituting $T(n)$, a and v in: $T(n) = a + (n-1)v$:

$$T(n) = a + (n-1)v$$

$$\Rightarrow 1 = \frac{x-4(R-1)}{2} - 3 + (n-1)(-2)$$

$$\Rightarrow 2 = x - 4R + 4 - 6 - 4n - 4$$

$$\boxed{n = \frac{x-4R}{4}} \tag{5.12}$$

Substituting (5.12) in $F(R) < y$:

$$\Rightarrow x - 4\,(R-1) + \left(\frac{x-4R}{4}\right)\left(\frac{x}{2} - 3\right) - \left(\frac{x-4R}{4}\right)\left(\frac{x-4R}{4} + 2R - 3\right) < y$$

$$\Rightarrow x - 4R + 4 + \frac{x^2}{8} - \frac{xR}{2} - \frac{3x}{4} + 3R - \left(\frac{x}{4} - R\right)\left(\frac{x}{4} + R - 3\right) < y$$

$$\Rightarrow \frac{x^2}{8} - R + \frac{x}{4} - \frac{xR}{4} + 4 - \left[\frac{x^2}{16} + \frac{xR}{4} - \frac{3x}{4} - \frac{xR}{4} - R^2 + 3R\right] < y$$

$$\Rightarrow \frac{x^2}{16} + R^2 - \frac{xR}{2} + x - 4R + 4 < y$$

$$\boxed{\frac{x^2}{16} + R^2 - \frac{xR}{2} + x - 4R + 4 < y} \tag{5.13}$$

Example 1: $20 < P < 40$.

Using the above equation (5.13):

$x = 40, y = 20$

$$\frac{40^2}{16} + R^2 - 20R + 40 - 4R + 4 < 20$$

$$\Rightarrow R^2 - 24R + 124 < 0$$

$$\Rightarrow 7.52786\ldots < R < 16.47213\ldots$$

The above calculated range indicates that for any values falling in that range, the last value is less than y [here value of y is 20].

The range is a general answer, but the correct value has to be picked in such a way that it corresponds to the table and for the above example, R should be 7 [cross check from table]. That is, take the lowest number from the range and round it down.

The range obtained above [$7.52786\ldots < R < 16.47213\ldots$] is a general mathematical solution for the quadratic inequality [$R^2 - 24R + 124 < 0$]. Here what the range tells us is that pick any R lying in the range, the values on that diagonal are all lesser than y. Note that not all the values of R really makes sense as in many cases as for here, the 16th diagonal does not even exist. But what matters here is that we are looking for the very 1st diagonal whose last number is smaller than y, thus that particular diagonal indicates that we do not need to compute further diagonals as the values will be obviously lesser than y. Rounding down exclude the computation of that diagonal whereas rounding up will include that diagonal but at the end the values on that diagonal are lesser than y, thus of no use in the case of rounding up.

Example 2: $60 < P < 80$.

$$\frac{80^2}{16} + R^2 - 40R + 80 - 4R + 4 < 60$$

$$\Rightarrow R^2 - 44R + 424 < 0$$

$$R = 14$$

Proof of correctness for a

To prove that a is correct for both cases (odd and even), equation (4.1) and (4.2) are used, respectively. Since a is the difference between the first two values of a diagonal, thus subtracting the 2nd value from the 1st by using the general equation for odd and even, should give a. See below.

The formula below is used in both cases:

$$\boxed{d = \frac{x}{2} - 3} \tag{5.14}$$

d is substituted in both general equations for even and odd cases.

Odd difference:

$$\boxed{a_o = \frac{x - 4(R-1)}{2} - 3} \tag{5.15}$$

Using equation (4.1), d is substituted by (5.14).

$$F_o(R,n) = x - 4(R-1) + nd - n(n + 2R - 3); R \geq 1, n \geq 0$$

Using the three above equations, the following condition is checked if true or not:

Check if:

$$F_o(R,1) - F_o(R,0) \stackrel{?}{=} a$$

$$\begin{aligned}
\tfrac{3x}{2} - 6R + 3 - (x - 4R + 4) &= \tfrac{x-4(R-1)}{2} - 3 \quad | \cdot 2 \\
&\Leftrightarrow \\
3x - 12R + 6 - 2(x - 4R + 4) &= x - 4R + 4 - 6 \\
&\Leftrightarrow \\
3x - 12R + 6 - 2x + 8R - 8 &= x - 4R + 4 - 6 \quad | - x \\
&\Leftrightarrow \\
-4R - 2 &= -4R - 2 \\
&\Leftrightarrow \\
-2 &= -2 \\
&\Leftrightarrow \\
0 &= 0
\end{aligned}$$

[QED]

5.3.2 Case 2: Even difference

$F(R) = x - 2(2R - 1) + nd - n(n + 2R - 2)$

For even difference, a is given by: $\boxed{a_e = \frac{x - 2(2R-1)}{2} - 3}$

The same logic as in odd difference [section 5.3.1] is followed to obtained a_e.

Substituting $T(n)$, a and υ in: $T(n) = a + (n - 1)\upsilon$:

$$T(n) = a + (n - 1)\upsilon$$

$$\Rightarrow 1 = \frac{x - 2(2R - 1)}{2} - 3 + (n + 1)\upsilon$$

$$\Rightarrow 1 = \frac{x - 2(2R - 1)}{2} - 3 - 2n + 2$$

$$\boxed{n = \frac{x}{4} - R - \frac{1}{2}} \tag{5.16}$$

Substituting (5.14) and (5.16) in (4.2) $[F_e(R)]$ such that
$F_e(R) < y$:

$$\Rightarrow x - 4R + 2 + \frac{x^2}{8} - \frac{xR}{2} - \frac{x}{4} - \frac{3x}{4} + 3R + \frac{3}{2} - \left[\frac{x^2}{16} - R^2 - \frac{3x}{4} + 2R + \frac{5}{4}\right] < y$$

$$\Rightarrow \frac{x^2}{8} - R - \frac{xR}{2} - \frac{7}{2} - \left[\frac{x^2}{16} - R^2 - \frac{3x}{4} + 2R + \frac{5}{4}\right] < y$$

$$\Rightarrow \frac{x^2}{8} - R - \frac{7}{2} - \frac{x^2}{16} + R^2 - \frac{3x}{4} - 2R - \frac{5}{4} < y$$

$$\Rightarrow \frac{x^2}{16} + R^2 - \frac{xR}{2} - 3R + \frac{3x}{4} + \frac{9}{4} < y$$

General formula for even-difference:

$$\boxed{\frac{x^2}{16} + R^2 - \frac{xR}{2} - 3R + \frac{3x}{4} + \frac{9}{4} < y} \tag{5.17}$$

Example 1: $28 < P < 38$

$$\Rightarrow 90.25 + R^2 - 19R - 3R + 30.75 < 28$$

$$\Rightarrow R^2 - 22R + 93 < 0$$

$$\Rightarrow 5.70849\ldots < R < 16.29150\ldots$$

$$R = 5$$

Example 2: $26 < P < 66$

$$\Rightarrow 272.25 + R^2 - 33R - 3R + 49.5 + 2.25 < 28$$

$$\Rightarrow R^2 - 36R + 298 < 0$$

$$\Rightarrow 12.90098\ldots < R < 23.09901\ldots$$

$$R = 12$$

Proof of correctness for a

The formula(5.14) below is used again:

$$d = \frac{x}{2} - 3$$

Even difference:

Using

$$a = \frac{x - 2(2R-1)}{2} - 3$$

and (4.2) $[F_e(R)]$, d is substituted by (5.14) in

$$F_e(R) = x - 2(2R-1) + nd - n(n + 2R - 2); R \geq 1, n \geq 0$$

to see if the the following condition is true or not, check if:

$$F_e(R,1) - F_e(R,0) \stackrel{?}{=} a$$

$$\begin{array}{ccll}
\begin{array}{c}\left[x - 2(2R-1) + (\frac{x}{2} - 3) - (1 + 2R - 2)\right] \\ -\left[x - 2(2R-1)\right]\end{array} & = & \frac{x-2(2R-1)}{2} - 3 & \\
& \Leftrightarrow & & \\
x - 4R + 2 + \frac{x}{2} - 3 + 1 - 2R - [x - 4R + 2] & = & \frac{x-2(2R-1)}{2} - 3 & \\
& \Leftrightarrow & & \\
\frac{3x}{2} - 6R - (x - 4R + 2) & = & \frac{x-2(2R-1)}{2} - 3 & | \cdot 2 \\
& \Leftrightarrow & & \\
3x - 12R - 2(x - 4R + 2) & = & x - 2(2R-1) - 6 & \\
& \Leftrightarrow & & \\
3x - 12R - 2x + 8R - 4 & = & x - 4R + 2 - 6 & | - x \\
& \Leftrightarrow & & \\
4R - 2 & = & -4R - 2 & \\
& \Leftrightarrow & & \\
-2 & = & -2 & \\
& \Leftrightarrow & & \\
0 & = & 0 &
\end{array}$$

[QED]

Chapter 6

Full examples

6.1 Odd difference

Find all prime numbers $60 < P < 76$.
Rewriting: $y < P < x$, where $y = 60$ and $x = 76$.
Range: $x - 16 < P < x$, where $y = x - 16 \Rightarrow 60$.

1. Finding the difference:

$$T = d + 3, \;\; T = \frac{x}{2}$$

$$\Rightarrow \frac{76}{2} = d + 3$$

$$\Rightarrow d = 35$$

2. Finding on which diagonal to start using the formula below:

$$-6R + x/2 + 3 < 0$$

$$\Rightarrow -6R + 38 + 3 < 0$$

$$\Rightarrow R > 6.833\ldots$$

$$\Rightarrow R = 7$$

3. Finding on which diagonal to stop using the following formula:

$$\frac{x^2}{16} + R^2 - \frac{xR}{2} + x - 4R + 4 < y$$

$$\Rightarrow \frac{76^2}{16} + R^2 - 38R + 76 - 4R + 4 < 60$$

$$\Rightarrow R^2 - 42R + 381 < 0$$

$$\Rightarrow 13.5\ldots < R < 28.7\ldots$$

$$\Rightarrow R = 13$$

4. Using the general formula below for finding values on each diagonal:

$$F_o(R, n) = x - 4(R-1) + nd - n(n + 2R - 3); R \geq 1, n \geq 0$$

Beginning from the starting diagonal, $R = 7$ as calculated in step (2) and with $d = 35$ calculated in step (1):

(i)

$$\begin{aligned} F_o(R = 7) &= x - 4(7-1) + 35n - n(n + 2(7) - 3) \\ &= x - 24 + 35n - n(n + 11) \end{aligned}$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 24 + 35 - 12 \Rightarrow\ x - 1[accept]$$

$$n = 3 \Rightarrow x + 39\ [reject]$$

Note: on each diagonal the stopping point is as soon as the constant c in the equation $x + c$ becomes positive as the numbers further down the diagonal will be definitely larger.

Why positive? Because then it becomes larger than x and thus lies out of range. The same follows for every diagonal.

(ii)

$$F_o(R = 8) = x - 28 + 35n - n(n + 13)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 7\ [accept]$$

$$n = 3 \Rightarrow x + 29\ [reject]$$

(iii)

$$F_o(R = 9) = x - 32 + 35n - n(n + 15)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 13\ [accept]$$

$$n = 3 \Rightarrow x + 19\ [reject]$$

(iv)

$$F_o(R = 10) = x - 36 + 35n - n(n + 17)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 19 \ [reject]$$

$$n = 3 \Rightarrow x + 9 \ [reject]$$

(v)

$$F_o(R = 11) = x - 40 + 35n - n(n + 19)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 25 \ [reject]$$

$$n = 3 \Rightarrow x - 11 \ [accept]$$

$$n = 5 \Rightarrow x + 1 \ [reject]$$

(vi)

$$F_o(R = 12) = x - 44 + 35n - n(n + 21)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 31 \ [reject]$$

$$n = 3 \Rightarrow x - 11 \ [accept]$$

$$n = 5 \Rightarrow x + 1 \ [reject]$$

(vii)

$$F_o(R = 13) = x - 48 + 35n - n(n + 23)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 37 \ [reject]$$

$$n = 3 \Rightarrow x - 21 \ [reject]$$

$$n = 5 \Rightarrow x - 13 \ [accept]$$

$$n = 7 \Rightarrow x - 13 \ [accept] \quad \leftarrow STOP!$$

Note: $x - 13$ is repeated twice, indicating the stopping point. See explanation in section (5.1).

5. List of odd values that are not prime found in the table are:
$x - 1, x - 7, x - 11, x - 13$

6. Missing odd values from list obtained in step (5) lying within the range that we are looking for, are prime:
   - $x - 3 = 73$
   - $x - 5 = 71$
   - $x - 9 = 67$
   - $x - 15 = 61$

6.2 Even difference

Find all prime numbers $50 < P < 70$.

Rewriting: $y < P < x$,where $y = 50$ and $x = 70$.
Range: $x - 20 < P < x$, where $y = x - 20 \Rightarrow 50$.

1. Finding the difference:

$$T = d + 3, \;\; T = \frac{x}{2}$$

$$\Rightarrow \frac{70}{2} = d + 3$$

$$\Rightarrow d = 32$$

2. Finding on which diagonal to start using the formula below:

$$-6R + \frac{x}{2} < 0$$

$$\Rightarrow -6R + 35 < 0$$

$$\Rightarrow R > 5.833\ldots$$

$$\Rightarrow R = 6$$

3. Finding on which diagonal to stop using the following formula:

$$\frac{x^2}{16} + R^2 - \frac{xR}{2} - 3R + \frac{3x}{4} + \frac{9}{4} < y$$

$$\Rightarrow 306.25 + R^2 - 35R - 3R + 52.5 + 2.25 < 50$$

$$\Rightarrow R^2 - 38R + 311 < 0$$

$$\Rightarrow 11.928833\ldots < R < 26.07101\ldots$$

$$\Rightarrow R = 11$$

4. Using the general formula below to finding values on each diagonal:

$$F_e(R,n) = x - 2(2R-1) + nd - n(n+2R-2); R \geq 1, n \geq 0$$

Beginning from the starting diagonal, $R = 6$ as calculated in step (2) and with $d = 32$ calculated in step (1):

(i)

$$F_e(R=6) = x - 22 + 32n - n(n+10)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 1 \ [accept]$$

$$n = 3 \Rightarrow x + 35 \ [reject]$$

(ii)

$$F_e(R=7) = x - 26 + 32n - n(n+12)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 7 \ [accept]$$

$$n = 3 \Rightarrow x + 25 \ [reject]$$

(iii)

$$F_e(R=8) = x - 30 + 32n - n(n+14)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 13 \ [accept]$$

$$n = 3 \Rightarrow x + 15 \ [reject]$$

(iv)

$$F_e(R=9) = x - 34 + 32n - n(n+16)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 19 \ [accept]$$

$$n = 3 \Rightarrow x + 5 \ [reject]$$

(v)

$$F_e(R=10) = x - 38 + 32n - n(n+18)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 25 \ [reject]$$

$$n = 3 \Rightarrow x - 5 \ [accept]$$

$$n = 5 \Rightarrow x + 7 \ [reject]$$

(vi)

$$F_e(R = 12) = x - 42 + 32n - n(n + 20)$$

Trial and error by substituting n with odd values:

$$n = 1 \Rightarrow x - 31 \; [reject]$$

$$n = 3 \Rightarrow x - 15 \; [accept]$$

$$n = 5 \Rightarrow x - 7 \; [accept]$$

$$n = 7 \Rightarrow x - 7 \; [accept] \leftarrow STOP!$$

5. List of odd values that are not prime and found within the interested range are:
   $x - 1, x - 5, x - 7, x - 13, x - 15, x - 19$

6. Missing odd values from list obtained in step (5) lying within the range that we are looking for, are prime.
   - $x - 3 = 67$
   - $x - 9 = 61$
   - $x - 11 = 59$
   - $x - 17 = 53$

Chapter 7

Comparison to other sieves

1. **Sieve of Eratosthenes**
   - Instead of listing all the numbers, the algorithm above lists only numbers that lie in the range we are looking for, thus having the form $y < P < x$, whereas the sieve of Eratosthenes is in the form $P < x$ which involves a lot of numbers that need to be discarded.
   - Operations begin from the mid-value of x then proceeds till y as seen earlier while the sieve of Eratosthenes starts from the 1st prime number which is 2.
   - The multiplication grid (figure 3.2) generates composite numbers and considering the pattern of the diagonals, they are incremented in an oddly or evenly way and it alternates. Furthermore, each diagonal follows a progression.
   - There is no need to write every natural number starting from 2 as in the sieve of Eratosthenes, instead, it can be substituted by variables and each diagonal can, in turn, be related by the same variable (usually x) which is helpful as the number gets larger.
2. **Sieve of Sundaram**

   As seen in the literature review, the sieve of Sundaram as well makes use of arithmetic progression. The use of the arithmetic progression is the only similarity with the procedure described in this thesis but how it is used and the rest of this thesis, are completely different to that of the sieve of Sundaram.

Regarding the other sieving methods, there is not much to compare as they are completely different to that of the thesis.

Chapter 8

Time Complexity and Space Analysis

It is always desirable to know the complexity of an algorithm. Thus, one has to consider the computational steps involved. In this thesis, only 3 basic arithmetic operations are involved namely, addition, subtraction and multiplication. The complexity of multiplication is more complex than that of addition for the the same input given. Addition and subtraction with two n digit numbers, N and N, has a complexity of $O(n)$, and the output is one $n+1$ digit number whereas the complexity for multiplication of two n digit having an output of one $2n$ digit number can vary depending on the algorithm used. Harvey-Hoeven algorithm has a complexity of $O(n \log n)$, n being the digit number. [12][14]
There are other algorithms as well such as Führer's algortihm [9] and others.

8.1 Complexity of the Algorithm

First of all, finding how many arithmetic operations does the general equation $F_{e/o}$ have:

$$F_o(R,n) = x \overbrace{-\underbrace{4(\underbrace{R-1}_{6})}_{7}+\overbrace{\underbrace{nd}_{5}\overbrace{-n\overbrace{(n+\overbrace{\overbrace{2R}^{1}-3}^{2})}^{3}}^{4}}^{8}}^{9} \quad R \geq 1, \quad n \geq 0$$

In total, the odd equation $F_o(R,n)$ has 9 arithmetic operations. Note that the subtraction $x-k$, where k is a constant is not counted because $x-k$ is the required form in order to proceed with elimination from the set of odd integers and if the subtraction operation was to be performed, an integer would be obtained which is not the aim here. The different numbers of arithmetic operations are as follows: 3 subtractions, 4 multiplications and 2 additions.

For the even case, there is one multiplication extra, thus totalling to 10 arithmetic operations. The extra operation is denoted by 10 as shown below:

$$F_o(R,n) \quad = \quad x \quad -2(\overbrace{2R}^{10} -1) \quad +nd \quad -n(n+2R-2), \quad R \geq 1, \quad n \geq 0$$

The complexity of the algorithm described in the thesis depends on the complexity of the algorithm used for the arithmetic operations. As discussed earlier in this same chapter paragraph one, addition and subtraction have a complexity of $O(n)$. For multiplication, one has a variety of choice. Considering the Harvey-Hoeven algorithm which has a complexity of $O(n \log n)$, the complexity of $F_{e/o}(R,n)$ approximates to as follows :

Note: Below when the assumption is made based on the size of two integers, the larger one is considered and thus, the other one is assumed to be of the same size. Also, one can consider the largest variable among all and then set all the other variables size to be the same as the largest one. For example, suppose R is the largest variable, for complexity analysis, the formula
$F_o(R,n) = x - 4(R-1) + nd - n(n+2R-3)$ is assumed to be
$F_o(R,n) = x - 4(R) + R \cdot R - R(R+2R)$. Below is a more detailed analysis.

1. $2R$ can be split as $R + R$, which is the addition of two $n-$digit numbers. Thus, having a complexity of $O(n)$.

2. $2R - 3$ can be ignored as the second operand is a constant.

3. $n + (2R - 3)$: assuming that n and $2R - 3$ has the same number of digits, $n + (2R - 3)$ can be approximated to $O(n)$.

4. $n(n + 2R - 3)$: Based on the assumption in (3), $n + 2R - 3$ is considered to be a $\tau-$digit number and therefore $n(n + 2R - 3)$ is estimated to be $\tau \cdot \tau$. Thus, $n(n + 2R - 3)$ has a complexity of $O(n \log n)$.

5. $n \cdot d$: Assuming both variables to be of the same number of digits, the complexity is $O(n \log n)$.

6. The -1 in $R - 1$ can be neglected as it is a constant.

7. $4(R - 1)$ can be split into two additions approximating to $2R + 2R$. Each $2R$ can in turn be split further into $R + R$, having a complexity of $O(n)$. Then both are added which is one more addition, resulting in $3 \cdot O(n)$.

8. $nd - n(n + 2R - 3)$: Based on the previous assumption in (3) and (5), where both $n \cdot d$ and $n(n + 2R - 3)$ are taken to be $n-$digit numbers,

$nd - n(n + 2R - 3)$ is therefore considered as a subtraction between two $n-$digit numbers having the complexity $O(n)$.

9. $4(R-1) - nd - n(n+2R-3)$: Assuming $4(R-1)$ and $nd - n(n+2R-3)$ are both two $n-$digit numbers, the subtraction operation has a complexity of $O(n)$.

Combining all together gives:

$O(n) + O(n) + O(n \log n) + O(n \log n) + 3 \cdot O(n) + O(n) + O(n)$

$\approx 7 \cdot O(n) + 2 \cdot O(n \log n)$

$\approx k \cdot O(n) + m \cdot O(n \log n)$, where k and m are constants.

$\approx O(n) + O(n \log n)$

The above complexity is for one value of a diagonal, but this is also the complexity of the algorithm since each single value can be computed independently and parallel.

8.2 Avoiding large multiplication by trading off Parallelism

From the general equations for odd and even case $F_o(R, n)$ and $F_e(R, n)$ respectively which are:

$F_o(R, n) = x - 4(R-1) + nd - n(n + 2R - 3)$ and
$F_e(R, n) = x - 2(2R-1) + nd - n(n + 2R - 2)$,

as n grows larger, the more complex will the multiplication nd be. Here n is referring to the n^{th} element of a diagonal. Multiplication of large numbers can be costly.

Avoiding multiplication is possible but this snatch away the power of parallelization within a diagonal. Below is an example:

Consider the example $60 < P < 76 \Rightarrow x - 16 < P < x$

The difference, d is $d = 35$.
The starting diagonal is $R = 7$.

Substituting $R = 7$ and different values of n in the equation
$F_o(R, n) = x - 4(R-1) + nd - n(n + 2R - 3)$.

$n = 0 : x - 4(6) \Rightarrow x - 24$
$n = 1 : x - 24 + d - 1(1 + 11) \Rightarrow x - 1$
Notice a difference of $+23$ $[-1 - -24 = 23]$ between the first two values.
$n = 2$ can be written directly by adding a difference which is lesser by 2 than the previously calculated difference (i.e $23 - 2 = 21$) to the previous result $(x - 1)$ resulting in $x - 1 + 21 \Rightarrow x + 20$.

Notice here that even numbers as well are being computed, which then have to be discarded straight away.

This can be done for all the diagonals. Firstly, get the first value of the diagonal which is obtained by substituting $n = 0$ and then the second value by replacing n by 1 in the general equation. Find the difference between them. To obtain the next value, subtract 2 from the last difference obtained and add it to the precedence result. Keep doing so as long as the result obtained after the operation falls in the interested range.

The first two multiplication namely for $n = 0$ and $n = 1$ are not complex one.

One can therefore choose either to have parallelization within the diagonal or having less complex operation.

Mathematical equation:

$$\boxed{F_{e/o}(R, n) = F_{e/o}(R, n - 1) + d_{(n-1,n-2)} - 2; n \geq 2}$$

where $d_{(n-1,n-2)}$ is the difference between the previous two values of that diagonal.

$F_e(R, 0) = x - 2(2R - 1)$; $F_e(R, 1) = x - 2(2R - 1) + d - (2R - 1)$
$F_o(R, 0) = x - 4(R - 1)$; $F_o(R, 1) = x - 4(R - 1) + d - (2R - 2)$

8.3 Memory analysis

The analysis below is based on the assumption that, each memory unit can store an integer entirely, neglecting the architecture of the machine. When the algorithm runs, at the start, the set of odd numbers falling in the interested range should be in memory, that is the set $O = \{x - 1, x - 3, x - 5, \ldots\}$. The number of elements in the set of odd numbers is equal to the range divided by 2 as even and odd numbers alternate.

8.3.1 Small Range

Find $999,990 < p < 1,000,000$

1. Finding the difference:

   $\frac{x}{2} = d + 3$

   $\Rightarrow d = \frac{x}{2} - 3$

   $\Rightarrow \frac{1{,}000{,}000}{2} - 3 \Rightarrow d = 499,997$[Odd difference]
2. Finding on which diagonal to start:

   $-6R + \frac{x}{2} + 3 < 0$

   $\Rightarrow -6R + 500003 < 0$

   $\Rightarrow R > 83,333.83$

   $\Rightarrow R = 83,334$

3. Stopping diagonal:

   Using $\frac{x^2}{16} + R^2 - \frac{xR}{2} + x - 4R + 4 < y$

   $\Rightarrow 6.25 \cdot 10^{10} + R^2 - 500,000R + 1,000,000 - 4R + 4 < 999,900$

   $\Rightarrow R^2 - 500,004R + 6.250000011 \cdot 10^{10} < 0$

   $\Rightarrow 249,002.05 < R < 251,001.948$

   $\Rightarrow R = 249,002$ [Taking the smallest value of the range]

Cross checking to have a better perception:

Finding n being the last value of the stopping diagonal using:

$n = \frac{x-4R}{4}$

Solving gives: $n = 998$

Replacing the appropriate value in the odd general equation:

$F_o(R,n) = x - 4(R-1) + nd - n(n+2R-3)$

$n = 998,\ d = 499,997,\ R = 249,002$

$F_o(R,n) = 1,000,000$, this is the last value of the stopping diagonal. One might think, why this is the stopping diagonal when the last value of the stopping diagonal is not lesser than $999,900$. Just to have a feeling, if the next diagonal is taken with the last value which are $R = 249,003$, $n = 997$ [use the appropriate formula to calculate n] respectively, and replaced in the general equation for odd, the following result is obtained, $F_o(R,n) = 998,001$, meaning that all previous values on that diagonal is even lesser and definitely out of range since the lowest boundary we are interested in is $999,900$.

Memory required = 50 units.

8.3.2 Large Range

Find $100 < p < 1,000,000$

1. Finding the difference:

   $\frac{x}{2} = d + 3$

   $\Rightarrow d = \frac{x}{2} - 3$

   $\Rightarrow \frac{1,000,000}{2} - 3 \Rightarrow d = 499,997$[Odd difference]

2. Finding on which diagonal to start:

   $-6R + \frac{x}{2} + 3 < 0$

   $\Rightarrow -6R + 500003 < 0$

   $\Rightarrow R > 83,333.83$

   $\Rightarrow R = 83,334$

3. Stopping diagonal:

   Using $\frac{x^2}{16} + R^2 - \frac{xR}{2} + x - 4R + 4 < y$

$\Rightarrow 6.25 \cdot 10^{10} + R^2 - 500,000R + 1,000,000 - 4R + 4 < 100$

$\Rightarrow R^2 - 500,004R + 6.25009999 \cdot 10^{10} < 0$

$\Rightarrow 249,992 < R < 250,012$

$\Rightarrow R = 249,992$ [Taking the smallest value of the range]

Cross checking:

Finding n being the last value of the stopping diagonal using:

$n = \frac{x-4R}{4}$

Solving gives: $n = 8$.

Replacing the necessary values in the equation for odd:

$F_o(R, n) = x - 4(R - 1) + nd - n(n + 2R - 3)$

$n = 8,\ d = 499,997,\ R = 249,992$

$F_o(R, n) = 100$, this is the last value of the stopping diagonal which is exactly the lower boundary of the range we are interested in; this also means all values before are lesser.

Memory needed $= 499,950$ units.

Chapter 9

Conclusion

As mentioned in the introduction, the aim was to realize a new method for finding primes. Step by steps throughout the document, an algorithm has demonstrated so, which forms part of the sieve family. Although the method explained is comparable to sieving methodologies, however the whole procedure is a new finding. Starting from the distribution of prime numbers, going over the elementary properties and the use a multiplication grid. Also, the method shows the inter-connectivity between recurrence relation, arithmetic progression and the multiplication grid where all are used in order to achieve the goal of finding prime numbers by avoiding the trial division. Lastly, an overview of the complexity and storage was given. Further research in this field is possible, such as implementation and optimization of the algorithm.

List of Figures

List of Tables

Bibliography

[1] Apostol, T. M. Dirichlet's theorem on primes in arithmetical progressions. In *Introduction to Analytic Number Theory*. Springer, 1976, pp. 146–156.

[2] Bays, C., and Hudson, R. H. The segmented sieve of eratosthenes and primes in arithmetic progressions to 10 12. *BIT Numerical Mathematics 17*, 2 (1977), 121–127.

[3] Beck, M., and Geoghegan, R. *The art of proof: basic training for deeper mathematics*. Springer Science & Business Media, 2010.

[4] Bengelloun, S. An incremental primal sieve. *Acta informatica 23*, 2 (1986), 119–125.

[5] Borwein, P. B., Borwein, P., Choi, S., Rooney, B., and Weirathmueller, A. The riemann hypothesis: a resource for the afficionado and virtuoso alike. vol. 27. Springer Science & Business Media, 2008, pp. 10–11.

[6] Crandall, R., and Pomerance, C. B. Prime numbers: a computational perspective. vol. 182. 2006, pp. 1,118.

[7] Dittrich, W. Reassessing riemann's paper: On the number of primes less than a given magnitude. Springer, 2018, pp. 8–9.

[8] Dora, R. *Figuring Out Mathematics*. Pearson Education India.

[9] Fürer, M. Faster integer multiplication. *SIAM Journal on Computing 39*, 3 (2009), 979–1005.

[10] Gillespie, F. S. A generalization of fermat's little theorem. *Fibonacci Quart 27*, 2 (1989), 109–115.

[11] Grigorieva, E. Introduction to sequences and series. In *Methods of Solving Sequence and Series Problems*. Springer, 2016, pp. 1–63.

[12] Harvey, D., and Van Der Hoeven, J. Integer multiplication in time o (n log n).

[13] Havil, J. Sundaram's sieve.

[14] Klarreich, E. Multiplication hits the speed limit, 2019.

[15] Mollin, R. An overview of sieve methods. *Int. J. Contemp. Math. Sci. 5*, 1-4 (2010), 67–80.

[16] Mollin, R. A. A brief history of factoring and primality testing bc (before computers). *Mathematics magazine 75*, 1 (2002), 18–29.

[17] Narkiewicz, W. The development of prime number theory: from euclid to hardy and littlewood. Springer Science & Business Media, 2000, p. 11.

[18] O'neill, M. e. The genuine sieve of eratosthenes. *Journal of Functional Programming 19*, 1 (2009), 95.

[19] Pomerance, C. The search for prime numbers. *Scientific American 247*, 6 (1982), 136–147.

[20] Pritchard, P. A sublinear additive sieve for finding prime number. *Communications of the ACM 24*, 1 (1981), 18–23.

[21] Pritchard, P. Fast compact prime number sieves (among others). *Journal of algorithms 4*, 4 (1983), 332–344.

[22] Sorenson, J. An introduction to prime number sieves. Tech. rep., University of Wisconsin-Madison Department of Computer Sciences, 1990.

[23] Stillwell, J. Greek number theory. In *Mathematics and Its History*. Springer, 2010, p. 40.

[24] Stillwell, J. Numbers and geometry. Springer Science & Business Media, 2012, p. 9.

[25] Wirian, D. J. Parallel prime sieve: Finding prime numbers. *Institute of Information & Mathematical Sciences Massey University at Albany, Auckland, New Zealand* (2009).

[26] Zhou, P. On the existence of infinitely many mersenne primes and finitely many fermat primes.

DOI: 10.17185/duepublico/74412
URN: urn:nbn:de:hbz:464-20210517-125113-3