An Elementary Proof of Green-Tao Theorem

Marko V. Jankovic

Institute of Electrical Engineering "Nikola Tesla", Belgrade, Serbia,
Swiss Rockets, Basel, Switzerland

Abstract In this paper an elementary proof of Green-Tao theorem is going to be presented. The proof represents an extension of the proof of the Polignac's conjecture (or twin prime, or Sophie Germain primes conjecture). It will be shown that arithmetic progressions that consist of prime numbers and that are of the length k (k is natural number), could be obtained through k-stage recursive type sieve process, and that their number is infinite.

1 Introduction

The Green-Tao theorem states that the sequence of prime numbers contains arbitrarily long arithmetic progressions [1]. This means that for every natural number k exists arithmetic progression of primes with k terms. In [1] it has been shown that exists infinitely many such sequences.

Here, an elementary proof of that theorem is going to be presented. The proof is based on extension of recently proposed proofs of the Sophie Germain prime conjecture, twin prime conjecture or Polignac's conjecture [2, 3, 4]. The major difference is that in the case of Green- Tao theorem recursion has the depth that is equal to the length of the arithmetic progression, while the depth of the recursion in the case of Polignac's or Sophie Germain conjecture is 2.

Basically, three groups of differences between prime numbers exists (here we ignore number 2): the prime numbers that are $6f$ far apart, $6f$-2 far apart and $6f$-4 far apart, $f \in N$. In the text that follows we mark the prime numbers in the form $6f - 1$ as *mps* primes and prime numbers in the form $6f + 1$ as *mpl* primes, $f \in N$. The gaps of the size $6f$ could be related to the prime pairs in both (*mps, mps*) and (*mpl, mpl*) form. The gaps in the form $6f$ - 2 can only be related to the pair of primes in (*mpl,* *mps*) form, while gaps in the form $6f$ - 4 can only be related to the pair of primes in (*mps, mpl*) form. In other words there is not a single prime in *mpl* form that has consecutive prime that is $6f$ - 4 apart, and there is not a single prime in *mps* form that has consecutive prime $6f$ - 2 apart. That means that all arithmetic progression longer than 2 have primes that are $6f$ far apart (for 2 terms it is proved in [3] and [4]).

It will be shown that exist lower bound for the number of arithmetic progressions with k ($k \in N$) terms that are smaller than some natural number n, $n \in N$, and that will be used to show that for every k exists infinite number of such progressions. To be precise, when in this paper is said number of progressions that are smaller than n, it is considered that the first term of the progression is smaller than n.

Remark 1: *In this paper any infinite series in the form $c_1 \cdot l \pm c_2$ is going to be called a thread defined by number c_1 (in literature these forms are known as linear factors – however, it seems that the term thread is probably better choice in this context). Here c_1 and c_2 are numbers that belong to the set of natural numbers (c_2 can also be zero and usually is smaller than c_1) and l represents an infinite series of consecutive natural numbers in the form (1, 2, 3, …).*

2 An elementary proof of the Green-Tao theorem

It is well known that all prime numbers can be expressed in one of the following forms

$$ps_k = 6k - 1$$

$$pl_k = 6k + 1,\ k \in N.$$

As it was already explained, we will call numbers ps_k - numbers in *mps* form and numbers pl_k - numbers in *mpl* form.

As it was already explained in the introduction, if we want to have an arithmetic progression that has more than two terms, the difference between two consecutive terms must be in the form $6f$, $f \in N$. It is very simple to understand, that if we want to have more than 4 terms, the difference must be divisible by 5. The only exception to this rule is the quintuplet (5, 11, 17, 23, 29) since the 5 is the only prime number that is divisible by 5. In all other cases if the difference between consecutive terms is not divisible by 5, one of the 5 consecutive terms has to be divisible by 5, which means that it is a composite number. This rule can be easily extended to all other primes – if we want to have progression that has at least 11 terms, difference has to be divisible by all primes that are smaller or equal to 11, difference has to be multiple of primorial(11) = 11#; or if we want progression that has at least 97 terms, difference has to be multiple of 97#. In general case if we want to have a progression of at least k terms, we are going to use a difference that is multiple of p#, where p is the smallest prime number that is not smaller than k. What is also clear is the fact that if first number of the progression is in *mps* form all other members are going to be in *mps* form. Also, if the first member of a progression is in *mpl* form, all other members of a progression are in *mpl* form. In the text that follows we are going to analyze only progressions in *mps* form – progressions in *mpl* form can be done analogously.

In order to prove that exist an infinite number of such progressions, we are going to create k stages recursive type process.

If we start with all natural numbers, the procedure looks as follows:

STAGE 1: Remove all composite numbers. So, only prime numbers are left. If we denote with $\pi(n)$ number of prime numbers smaller than n, the following equation holds [5]

$$\pi(n) \approx \frac{n}{\ln(n)}.$$

From [5] we also know that following holds

$$\pi(n) > \frac{n}{\ln(n)}, \quad n \geqslant 17. \tag{1}$$

Prime numbers can be obtained in the following way:

First, we remove all even numbers (except 2) from the set of natural numbers. Then, it is necessary to remove the composite odd numbers from the rest of the numbers. In order to do that, the formula for the composite odd numbers is going to be analyzed. It is well known that odd numbers bigger than 1, here denoted by a, can be represented by the following formula

$$a = 2n + 1,$$

where $n \in N$. It is not difficult to prove that all composite odd numbers a_c can be represented by the following formula

$$a_c = 2(2ij + i + j) + 1 = 2((2j+1)i + j) + 1. \qquad (2)$$

where $i, j \in N$. It is simple to conclude that all composite numbers could be represented by product $(i + 1)(j + 1)$, where $i, j \in N$. If it is checked how that formula looks like for the odd numbers, after simple calculation, equation (2) is obtained. This calculation is presented here. The form $2m + 1$, $m \in N$ will represent odd numbers that are composite. Then the following equation holds

$$2m + 1 = (i_1 + 1)(j_1 + 1) \quad ,$$

where $i_1, j_1 \in N$. Now, it is easy to see that the following equation holds

$$m = \frac{i_1 j_1 + i_1 + j_1}{2}.$$

In order to have $m \in N$, it is easy to check that i_1 and j_1 have to be in the forms

$$i_1 = 2i \text{ and } j_1 = 2j,$$

where $i, j \in N$. From that, it follows that m must be in the form

$$m = 2ij + i + j = (2i + 1)j + i. \qquad (3)$$

When all numbers represented by m are removed from the set of odd natural numbers bigger than 1, only the numbers that represent odd prime numbers are going to stay. In other words, only odd numbers that cannot be represented by (2) will stay. This process is equivalent to the sieve of Sundaram [6].

STAGE 2

What was left after the first stage are prime numbers. With the exception of number 2, all other prime numbers are odd numbers. All odd primes can be expressed in the form $2n + 1$, $n \in N$. It is simple to understand that if we want to have an arithmetic progression in which the difference between consecutive terms is $d = p(f)\#$, the difference between first term of the progression and all other terms of the progression is the following series $D = (d, 2d, 3d, \ldots, (k\text{-}1)d)$, and d is selected according to the previous analysis ($p(f)$ is f^{th} prime number that is not smaller than k). So, we are going to search for the prime numbers that have feature to have colleague primes at distances D. In order to do that in this step we are going to remove all primes that do not have a bigger colleague prime at distance d. If we mark prime with $2n$+1, the colleague must be in the form $2n + 1 + d$, $n \in N$. Now, we should implement a second stage in which we are going to **remove**:

A. number 2 (since 2 can make only two term progressions), but this has no impact on the analysis,

B. the primes in *mpl* form – it is trivial to see that it can be done by one thread that is defined by 3 – so in this step it is going to be removed, approximately (having in mind that the number of *mps* primes is a bit bigger than the number of *mpl* primes), one half of the numbers that are left after step *A*,

C. now all odd numbers in the form $2m + 1$ that have bigger colleague d apart, that is in the form $2m + 1 + d$, $m \in N$ and that is composite. If we make analysis similar to one in STAGE 1, it is simple to understand that m must be in the form

$$m = 2ij + i + j - 1 = (2i + 1)\,j + i - d/2. \tag{4}$$

All numbers (in observational space) that are going to stay must be numbers in *mps* form and they represent primes in *mps* form that have prime bigger colleagues that are d apart. What has to be noticed is that thread in (4) that is defined by prime number 3 (for $i = 1$) is not going to remove any number from the numbers left, since it will remove same numbers as the thread defined by 3 used in STAGE 1, since 3 divides $d/2$. Same holds for all threads defined by the primes that divide $d/2$.

Since the methods that are applied in the first and the second stage are similar, it can be intuitively concluded that the number of numbers left after the second "Sundaram" sieve, should be comparable to $gt2(n)$ defined by the following equation ($n \in N$)

$$gt2(n)=\frac{\pi(n)}{\ln(\pi(n))}>\frac{n}{(\ln(n))^2} \quad . \tag{5}$$

The $gt2(n)$ would be obtained in the case when second stage sieve would produce the same amount of numbers removed with each thread, like the original Sundaram sieve. However, the assumption is not correct and formula requires some compensation terms since the second “Sundaram” sieve is applied on an incomplete set, that is depleted by previously implemented Sundaram sieve. Actually, $gt2(n)$ represents a lower bound for the number of primes that have a prime colleague that is d apart and that are smaller than some number n and that are left after stage 2. In order to understand why it is so, we are going to analyze stages 1 and 2 in more detail.

It is not difficult to be seen that m in (3) and (4) is represented by the threads that are defined by odd prime numbers (see Appendix A). Now, we are going to compare sieves in stages 1 and 2. Starting point in the second stage is point B (the number of numbers left is number of primes; 2 is ignored).

Table 1 Comparison of the stages 1 and 2 for a few threads defined by smallest primes

Step	Stage 1	Step	Stage 2
1	Remove even numbers (except 2) amount of numbers left 1/2	1	Remove the rest of *mpl* primes amount of numbers left 1/2
2	Remove numbers defined by thread defined by 3 (obtained for $i = 1$) amount of numbers left 2/3	2	Remove numbers defined by thread defined by $p(f+1)$; amount of numbers left $(p(f+1)-2)/(p(f+1)-1)$
3	Remove numbers defined by thread defined by 5 (obtained for $i = 2$) amount of numbers left 4/5	3	Remove numbers defined by thread defined by $p(f+2)$; amount of numbers left $(p(f+2)-2)/(p(f+2)-1)$
4	Remove numbers defined by thread defined by 7 (obtained for $i = 3$) amount of numbers left 6/7	4	Remove numbers defined by thread defined by $p(f+3)$; amount of numbers left $(p(f+3)-2)/(p(f+3)-1)$
5	Remove numbers defined by thread defined by 11 (obtained for $i = 5$) amount of numbers left 10/11	5	Remove numbers defined by thread defined by $p(f+4)$; amount of numbers left $(p(f+4)-2)/(p(f+4)-1)$
6	Remove numbers defined by thread defined by 13 (obtained for $i = 6$) amount of numbers left 12/13	6	Remove numbers defined by thread defined by $p(f+5)$; amount of numbers left $(p(f+5)-2)/(p(f+5)-1)$

From Table 1 it can be seen that, in every step, except step 1, threads in the second stage will leave bigger percentage of numbers than the corresponding threads in the first stage. Of course it holds for all other threads of interest (not only those presented in Table 1). This is going to be analyzed in Appendix B. Based on analysis of (4), it is known that threads defined by primes that are smaller or equal to prime $p(f)$ will not remove any number in this and any subsequent stages. Only threads defined by primes that are bigger or equal to prime $p(f+1)$ will remove some additional numbers. It can be noticed that threads defined by the same number in the first and the second stage will not remove the same percentage of numbers. The reason is obvious – consider for instance the thread defined by 3: in the first stage it will remove 1/3 of the numbers left, but in the second stage it will remove ½ of the numbers left, since the thread defined by 3 in stage 1 has already removed one third of the numbers (odd numbers divisible by 3 in observation space). So, only odd numbers (in observational space) that give residual 1 and -1 when they are divided by 3 are left, and there are approximately same number of numbers that give residual -1 and numbers that give residual 1, when the number is divided by 3. Same way of reasoning can be applied for all other threads defined by same prime in different stages.

So, from previous paragraph we know that bigger number of numbers is left in every step of stage 2 then in the stage 1 (except 1st step). From that, we can conclude that after every step bigger than 1, part of the numbers that is left in stage 2 is bigger than number of numbers left in the stage 1 (that is also noticeable if we consider amount of numbers left after removal of all numbers generated by threads that are defined by all prime numbers smaller than some natural number). Let us mark the number of primes that have prime colleague d apart ($D1$-primes) smaller than some natural number n with $\pi_{D1}(n)$. From previous analysis we can safely conclude that the following equation holds for some n big enough (having in mind (1))

$$\pi_{D1}(n) > gt2(n) \quad .$$

Having in mind (1), by some elementary calculation it can be realized that n that is big enough is n

≥ 73.

Since it it easy to show that following holds

$$\lim_{n \to \infty} gt2(n) = \lim_{n \to \infty} \frac{\pi(n)}{\ln(\pi(n))} = \infty \quad ,$$

we can safely conclude that the number of $D1$-primes is infinite.

STAGE 3:

In this stage, from primes left after the STAGE 2, we are going to remove all primes that do not have a prime colleague that is $2d$ apart, or all primes in the form $2n + 1$, that do not have a bigger prime colleague in the form $2n + 1 + 2d$, $n \in N$. So, in this stage we are going to **remove all primes in the form 2*m* + 1 such that** $2m + 1 + 2d$, $m \in N$ are composite. If we make the same analysis like in previous stages, it is simple to understand that m must be in the form

$$m = 2ij + i + j - 1 = (2i + 1)j + i - d. \tag{6}$$

All numbers (in observational space) that are going to stay must be numbers in *mps* form and they represent primes that have prime colleagues at distances d and $2d$. What has to be noticed is that threads in (6) that are defined by primes that divide d, will not remove any number in this stage.

Since the methods that are applied in the first and the third stage are similar, it can be intuitively concluded that the number of numbers left after the second "Sundaram" sieve, should be comparable to $gt3(n)$ defined by the following equation ($n \in N$)

$$gt3(n) = \frac{\pi_{D1}(n)}{\ln(\pi_{D1}(n))} > \frac{\pi(n)}{\ln(\pi(n))\left(\ln(\pi(n)) - \ln(\ln(\pi(n)))\right)} > \frac{n}{(\ln(n))^3} \quad . \tag{7}$$

The $gt3(n)$ would be obtained in the case when third stage sieve would produce the same amount of numbers removed with each thread, like the original Sundaram sieve. However, the assumption is not correct and formula requires some compensation terms since the second "Sundaram" sieve is applied on an incomplete set, that is depleted by previously implemented Sundaram sieve. Actually, $gt3(n)$ represents a lower bound for the number of $D2$-primes that have a prime colleagues that are

d and $2d$ apart and that are smaller than some number n and that are left after stage 3.

Again, we are going briefly to compare sieves in stages 1 and 3.

Table 2 Comparison of the stages 1 and 3 for a few threads defined by smallest primes

Step	Stage 1	Step	Stage 3
1	Remove even numbers (except 2) amount of numbers left 1/2	1	Remove numbers defined by thread defined by $p(f+1)$; amount of numbers left $(p(f+1)-3)/(p(f+1)-2)$
2	Remove numbers defined by thread defined by 3 (obtained for $i = 1$) amount of numbers left 2/3	2	Remove numbers defined by thread defined by $p(f+2)$; amount of numbers left $(p(f+2)-3)/(p(f+2)-2)$
3	Remove numbers defined by thread defined by 5 (obtained for $i = 2$) amount of numbers left 4/5	3	Remove numbers defined by thread defined by $p(f+3)$; amount of numbers left $(p(f+3)-3)/(p(f+3)-2)$
4	Remove numbers defined by thread defined by 7 (obtained for $i = 3$) amount of numbers left 6/7	4	Remove numbers defined by thread defined by $p(f+4)$; amount of numbers left $(p(f+4)-3)/(p(f+4)-2)$
5	Remove numbers defined by thread defined by 11 (obtained for $i = 5$) amount of numbers left 10/11	5	Remove numbers defined by thread defined by $p(f+5)$; amount of numbers left $(p(f+5)-3)/(p(f+5)-2)$
6	Remove numbers defined by thread defined by 13 (obtained for $i = 6$) amount of numbers left 12/13	6	Remove numbers defined by thread defined by $p(f+6)$; amount of numbers left $(p(f+6)-3)/(p(f+6)-2)$

Like in the STAGE 2, we will state that in Table 2, in every step, except may be step 1, threads in the second stage will leave bigger percentage of numbers than the corresponding threads in the first stage (this is explained in Appendix B). Of course it holds for all other threads of interest (not only those presented in Table 2).

Let us mark the number of $D2$-primes smaller than some natural number n with $\pi_{D2}(n)$. From previous analysis we can safely conclude that the following equation holds for some n big enough (having in mind (1))

$$\pi_{D2}(n) > gt3(n) \quad .$$

Since it it easy to show that following holds

$$\lim_{n \to \infty} gt3(n) > \lim_{n \to \infty} \frac{n}{(\ln(n))^3} = \infty \quad ,$$

we can safely conclude that the number of $D2$-primes is infinite.

…

STAGE k.

In this stage, from primes left after the STAGE k-1, we are going to remove all primes that do not have a colleague prime that is $(k\text{-}1)d$ apart, or all primes $2n + 1$, that do not have bigger prime colleague in the form $2n + 1 + (k\text{-}1)d$, $n \in N$. So, in this stage we are going to **remove all primes in the form 2*m* + 1 such that** $2m + 1 + (k\text{-}1)d$, $m \in N$ *is* composite. If we make the same analysis like in the previous stages, it is simple to understand that m must be in the form

$$m = 2ij + i + j - 1 = (2i + 1)\,j + i - (k - 1)d/2. \tag{8}$$

All numbers (in observational space) that are going to stay must be numbers in *mps* form and they represent primes in *mps* form that have prime bigger colleagues that are D apart. What has to be noticed in (8) is that threads defined by primes that divide $(k\text{-}1)d/2$ will not remove numbers in this stage.

Since the methods that are applied in the first and the k^{th} stage are similar, it can be intuitively concluded that the number of numbers left after the k^{th} stage "Sundaram" sieve, should be comparable to $gtk(n)$ defined by the following equation ($n \in N$, $\pi_{D\text{-}1}(n)$ denotes the number of primes smaller than n, left after stage k-1)

$$gtk(n) = \frac{\pi_{D-1}(n)}{\ln(\pi_{D-1}(n))} > \frac{n}{(\ln(n))^k} \quad . \tag{9}$$

The $gtk(n)$ would be obtained in the case when k stage sieve would produce the same amount of numbers removed with each thread, like the original Sundaram sieve. However, the assumption is not correct and formula requires some compensation terms since the second "Sundaram" sieve is applied on an incomplete set, that is depleted by previously implemented Sundaram sieve. Actually,

$gtk(n)$ represents a lower bound for the number of D-primes that have prime colleagues that are D apart and that are smaller than some number n and that are left after stage k.

Let us mark the number of D-primes smaller than some natural number n with $\pi_D(n)$. From previous analysis we can safely conclude that the following equation holds for some n big enough (having in mind (1))

$$\pi_D(n) > gtk(n) \quad .$$

Since it it easy to show that following holds

$$\lim_{n \to \infty} gtk(n) > \lim_{n \to \infty} \frac{n}{(\ln(n))^k} = \infty \quad ,$$

we can safely conclude that the number of D-primes is infinite. That completes the proof. What has to be said is that actually one more step should be performed and number of arithmetic progression with k+1 terms and same difference should be calculated and subtracted from the number of arithmetic progressions with k elements. However, this can be ignored for large n, since the number of progressions of length k+1 is very small comparing to the number of progressions with k terms (and same difference) for n big enough.

Number of arithmetic progressions smaller than some natural number n is going to be analyzed in the next version of this paper. Here we will just say that similar analysis like in the case of the Polignac's conjecture can be done with some additional analysis that is results of the depth of the recursion that has to be applied in the case of Green-Tao theorem.

References

[1] B. Green, T. Tao. (2008) The primes contain arbitrarily long arithmetic progressions, Annals of Mathematics, 167 (2), pp. 481-547.

[2] M. V. Jankovic. (2021). Proof of the Sophie Germain Primes Conjecture, hal-02169242, version

8.

[3] M. V. Jankovic. (2021). A Proof of the Twin Prime Conjecture (together with the Proof of Polignac's Conjecture for Cousin Primes), hal-02549967, version 9

[4] M. V. Jankovic. (2021). A Proof of the Polignac's Conjecture for gaps bigger than four, hal-03170059, version 3.

[5] J.B. Rosser, L. Schoenfeld. (1962) Approximate formulas for some functions of prime numbers. Illinois Journal of Mathematics, 6(1), pp. 64-94.

[6] V. Ramaswami Aiyar. (1934) Sundaram's Sieve for prime numbers, The Mathematics Student, 2(2) , 73.

APPENDIX A.

Here it is going to be proved that m in (3) is represented by threads defined by odd prime numbers. Now, the form of (3) for some values of i will be checked.

<u>Case i = 1:</u> $m = 3j + 1$,

<u>Case i = 2:</u> $m = 5j + 2$,

<u>Case i = 3:</u> $m = 7j + 3$,

<u>Case i = 4:</u> $m = 9j + 4 = 3(3j + 1) + 1$,

<u>Case i = 5:</u> $m = 11j + 5$,

<u>Case i = 6:</u> $m = 13j + 6$,

<u>Case i = 7:</u> $m = 15y + 7 = 5(3j + 1) + 2$,

<u>Case i = 8:</u> $m = 17j + 8$,

It can be seen that m is represented by the threads that are defined by odd prime numbers. From examples (cases $i = 4$, $i = 7$), it can be seen that if $(2i + 1)$ represent a composite number, m that is represented by thread defined by that number also has a representation by the the thread defined by one of the prime factors of that composite number. That can be proved easily in the general case, by direct calculation, using representations similar to (2). Here, that is going to be analyzed. Assume that $2i + 1$ is a composite number, the following holds

$$2i + 1 = (2l + 1)(2s + 1)$$

where ($l, s \in N$). That leads to

$$i = 2ls + l + s.$$

The simple calculation leads to

$$m = (2l + 1)\,(2s + 1)\,j + 2ls + l + s = (2l + 1)(2s+1)j + s(2l + 1) + l$$

or

$$m = (2l+1)((2s+1)j + s) + l$$

which means

$$m = (2l + 1)f + l,$$

and that represents the already exiting form of the representation of m for the factor $(2l + 1)$, where

$$f = (2s + 1)j + s.$$

In the same way this can be proved for (4), (6) and (8).

<u>Note</u>: *It is not difficult to understand that after implementation of stage 1, the number of numbers in residual classes of some specific prime number are equal. In other words, after implementation of stage 1, for example, all numbers divisible by 3 (except 3, but it does not affect the analysis) are removed. However, the number of numbers in the forms 3k + 1 and 3k + 2 (alternatively, 3k – 1) are equal. The reason is that the thread defined by any other prime number (bigger than 2) will remove the same number of numbers from the numbers in the form 3k + 1 and from the numbers in the form 3k + 2. It is simple to understand that, for instance, thread defined by number 5, is going*

to remove 1/5 of the numbers in form $3k + 1$ and 1/5 of the numbers in form $3k + 2$. This can be proved by elementary calculation. That will hold for all other primes and for all other residual classes.

APPENDIX B.

Here we are going to analyze two sequences that consist of prime numbers, that have same length k, $k \in N$. The first sequence S2 consists of numbers

$p(1), p(2), \ldots, p(k)$,

while the second sequence Ss consists of the prime numbers

$p(s+1), p(s+2), \ldots, p(s+k)$, where $s \in N$.

Now we are going to form two new sequences :

$pd(1) = p(2) - p(1), pd(2) = p(3) - p(1), \ldots, pd(k\text{-}1) = p(k) - p(1)$ and

$psd(1) = p(s+2) - p(s+1), psd(2) = p(s+3) - p(s+1), \ldots, psd(k\text{-}1) = p(s+k) - p(s+1)$.

Here we are going to prove that for all $s \in N$, the following set of inequalities holds

$$pd(1) < psd(1)$$

$$pd(2) < psd(2)$$

$$\ldots$$

$$pd(k\text{-}1) < psd(k\text{-}1).$$

So, the difference from between every term in the sequence and the first term in the sequence is always smaller for the sequence of primes S2 that starts with two than any other sequence Ss that starts with some other prime number. The reasons for that are quite simple

- only the sequence S2 contains one even prime, and difference between first and second term in the sequence is one. In all other sequences the minimal difference between first and second term is 2.

- only the sequence S2 contain the second, third and forth term that are 1, 3 and 5 apart from the first term. In all other sequences Ss the second, third and forth term are at least 2, 6, and 8 apart from the first term.
- and so on … For every additional term you can prove by direct comparison that r^{th} term in the S2 sequence is at smaller distance from the first term in the sequence S2 than in any other sequence Ss. Also, it can be proved by choosing the starting term in the sequence Ss, and than start the Eratosthen sieve procedure and compare the sequences that are obtained.

Here we will also say that if we create sequence Ss-α, $\alpha < s$, where s marks the first prime in sequence Ss, the sequence *psd* of this new sequence will not be changed. It is trivial to prove.

Having all this in mind we can easily prove that for all $k > 1$ and $d \geq 0$, the following inequality holds

$$\frac{p(1+d)-1}{p(1+d)} \leqslant \frac{p(k+d)-\alpha}{p(k+d)-\alpha+1},$$

since

$$p(1+d)p(k+d)-p(1+d)\alpha+p(1+d)-p(k+d)+\alpha-1 \leqslant p(1+d)p(k+d)-\alpha p(1+d),$$

or

$$p(1+d) \leqslant p(k+d)-\alpha+1,$$

and equality sign is possible only for the case $d = 0$ and $p(k+d)-\alpha+1 = 2$, or when the first term of the sequence Ss is shifted to coincide with 2.

Having in mind what was previously said, it is clear that statements after Tables 1 and 2 were true.