A Proof of Polignac's Conjecture

Marko V. Jankovic

Institute of Electrical Engineering “Nikola Tesla”, Belgrade, Serbia,

Abstract In this paper a proof of the Polignac's Conjecture is going to be presented. The proof represents an extension of the proof of the twin prime conjecture. It will be shown that primes with a gap of size g could be obtained through two stage sieve process, and that will be used to prove the conjecture.

1 Introduction

In number theory, Polignac's conjecture states: For any positive even number g, there are infinitely many prime gaps of size g. In other words: there are infinitely many cases of two consecutive prime numbers with the difference g [1].

In literature it has been proved that exists infinitely many primes with gaps not bigger than 246 [2]. Recently, the conjecture was proved for gaps of the size 2 and 4 [3], 6 [4] and 8 [5]. The problem was addressed in generative space, which means that prime numbers were not analyzed directly, but rather their representatives that can be used to produce them. This paper represents an extension of the previous work [3-5].

Basically, three groups of gaps bigger than 4 exist: the gaps of the size $6k$, the gaps of the size $6k+2$ and gaps of the size $6k+4$, $k \in N$. In the text that follows we mark the prime numbers in the form $6k - 1$ as *mps* primes and prime numbers in the form $6k + 1$ as *mpl* primes, $k \in N$. The gaps of the size $6k$ could be related to the prime pairs in both (*mps, mps*) and (*mpl, mpl*) form. The gaps in the form $6k + 2$ can only be related to the pair of primes in (*mps, mpl*) form, while gaps in the form $6k + 4$ can only be related to the pair of primes in (*mpl, mps*) form. In other words there is not a single prime in *mpl* form that has consecutive prime that is $6k + 2$ apart, and there is not a single prime in *mps* form that has consecutive prime $6k + 4$ apart. It is trivial to show that by simple calculation. Although last two groups are different, they can be treated analogously. Here, three different cases are going to be analyzed in order to explain how conjecture can be proved in general case.

First, the analysis from [5] for the gaps of the size 8 is going to be recapitulated. After that, cases in which gaps are equal to 10 and 12 are going to be analyzed briefly. Then, the proof can easily be generalized for the gaps of any size g. It will be shown that primes with gap of a size g (Gg-primes) could be generated by two stage sieve process. This process will be compared to other two stage recursion sieve process that leaves infinitely many numbers. Fact that sieve process that generate Gg-primes leaves more numbers than the other sieve process will be used to prove that infinitely many Gg-primes exist.

<u>Remark 1:</u> *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, ...).*

<u>Remark 2:</u> *In the text that follows we will use fact that number of even/odd numbers is equal to the one half of the number of natural numbers (also the 1/3, 4/5 and other fractions of the number of natural numbers are used). This is not an usual way of discussing the number of numbers that are infinitely big, but it is quite suitable in this context.* ***It is important to notice that there is no context in which is correct to state that number of natural numbers is equal to the number of odd or even numbers (or that number of natural numbers is equal to the number of numbers divisible by 3, or 5, and so on). What can be said is that it is possible to generate the same number of unique labels for odd or even numbers using the same number of unique labels for natural numbers. However, if we want to produce all even and odd labels at the same time,***

obviously, it is necessary to have two sets of natural numbers at the same time (it is necessary to clone the set of natural numbers), which means that the number of the natural numbers in that moment is two times bigger than the number of even or odd numbers. *This analysis can also be made in quantum probabilistic context, but it is beyond the scope of this paper.*

2 Recapitulation of the proof of Polignac's conjecture for gap equal to eight

As it was already explained in the introduction part, if two consecutive prime numbers have the gap of the size 8 it is clear that smaller of those numbers has to be in *mps* form while the bigger one has to be in *mpl* form. Here, we are going to recapitulate a two stage process, that was presented in [5], that can be used for generation of the smaller primes of the *G*8-prime pairs that are in *mps* form.

In the first stage prime numbers are going to be produced by removal of all composite numbers from the set of natural numbers.

In the second stage, the following numbers are going to be removed:

- Number 2 (since it cannot generate *G*8-prime pair).
- All numbers in *mps* form that represent a smaller numbers in twin prime pairs.
- A quarter of sexy primes (or more precisely one half of *mps* sexy primes). In order to understand why is it so, we are going to consider 5 consecutive numbers $6k$-1, $6k$+1, $6k$+3, $6(k+1)$-1, $6(k+1)$+1. We can see that $6k$-1 and $6(k+1)$+1 can create a *G*8-primes only in the case when they are both primes and when $6k$+1 and $6(k+1)$-1 are not primes ($6k$+3 is obviously composite number divisible by 3). Here, we are going to analyze three situations. When $6k$+1 and $6(k+1)$-1 are both primes, and in the case when only $6k$+1 is prime, removal of twin primes will remove *mps* primes that cannot have *G*8-prime pair. However, when $6k$+1 is not prime and $6(k+1)$-1 is prime, it is also necessary to remove smaller number of sexy prime pair ($6k$-1, $6(k+1)$-1) since it also cannot create a *G*8-prime pair.
- A half of the number of twin primes and quarter of the number of sexy primes from *mpl* primes is going to be removed in such a way that Dirichlet's theorem for arithmetic progressions [8]

still holds after removal of all those numbers that are previously mentioned.

– All prime numbers in *mpl* form that are left.

– All prime numbers in *mps* form that have a bigger odd $G8$-neighbor (odd number that is by 8 bigger than the prime of interest) that is a composite number.

At the end, only the prime numbers in the *mps* form, that represent the smaller numbers of the $G8$-prime pairs, are going to stay. Their number is a half of the number of $G8$-primes. It is going to be shown that that number is infinite.

STAGE 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

$$a_c = 2(2ij + i + j) + 1 = 2((2j+1)i + j) + 1. \tag{1}$$

where $i, j \in N$. It is simple to conclude that all odd composite numbers could be represented by product $(2i + 1)(2j + 1)$, where $i, j \in N$. If it is checked how that formula looks like for the odd numbers, after simple calculation, equation (1) 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

$$2m + 1 = (2i+1)(2j+1) \quad .$$

From that, it follows that m must be in the form

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

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 (1) will stay. This process is equivalent to the sieve of Sundaram [6].

Let us denote the numbers used for the generation of odd prime numbers with *m2* (here we ignore number 2). Those are the numbers that are left after the implementation of Sundaram sieve. The number of those numbers that are smaller than some natural number n, is equivalent to the number of prime numbers smaller than n. If we denote with $\pi(n)$ number of primes smaller than n, than the following equation holds [7]

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

From [7] we also know that following holds

$$\pi(n) > \frac{n}{\ln(n)}, \quad n \geqslant 17. \qquad (3)$$

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 their bigger odd $G8$-neighbor must be in the form $2n + 9$, $n \in N$. Now, we should implement a second stage in which we are going to **remove**:

A. Number 2 (since 2 cannot make a $G8$-prime pair);

B. All twin primes in *mps* form, a quarter of sexy primes (smaller sexy numbers in *mps* form). Also, a number of primes in *mpl* form that is equal to the number that represents the sum of the half of the numbers of twin primes and quarter of the numbers of sexy primes, in such a way that Dirichlet's theorem for arithmetic progressions [8] still holds after removal of all those numbers that are mentioned– so number of numbers that is going to be left is the number of primes minus the number of twin primes minus one half of the number of sexy primes (number 2 is ignored, and that has no impact on the analysis that follows). It is not difficult to prove that number of numbers left, is infinite. It is not difficult to understand it having in mind that relative density of twin and sexy primes to the density of all primes is zero.

C. The rest of *mpl* primes – 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, one half of the numbers that are left after step *B;*

D. All odd primes in the form $2m + 1$ such that $2m + 9$, $m \in N$ represents a composite number (all primes whose bigger odd *G*8-neighbor is composite number). If we make the same analysis, like in the Stage 1, it is simple to understand that m must be in the form

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

All numbers (in observational space) that are going to stay must be numbers in *mps* form and they represent a smaller primes of the *G*8-prime pairs that is in *mps* form. 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 additional number from the numbers left.

Let us mark the number of *G*8-primes with π_{G8}. Also, let us define the number of numbers that is left after two consecutive implementations of Sundaram sieve as *pd8*. The prime numbers that are obtained after removal of number 2, twin primes in *mps* form and a quarter of sexy primes (a half of sexy primes in *mps* form) and the corresponding number of *mpl* primes (equal to the sum of a half of the number of twin primes and a quarter of the number of the sexy primes) are going to be called *G*8-fs primes. The numbers obtained after recursive implementation of two Sundaram sieves (where the second Sundaram sieve is implemented on *G*8-fs primes) are going to be called *G*8-double primes. The second stage sieve that is identical to the first stage sieve can be obtained if the *G*8-fs primes are lined up next to each other and then the numbers are removed from the exactly same positions like in the first stage – the second stage sieve is applied on the indexes of *G*8-fs primes. In that case it is easy to understand that the following equation would holds ($n \in N$)

$$pd8(n) \approx \frac{\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)}{\ln\left(\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)\right)} \quad , \tag{5}$$

where $pd8(n)$ represents the number of $G8$-double primes smaller than some natural number n *(this number can be easily found out (by simple counting), once it is known which mpl numbers are removed in the process of obtaining the G8-fs primes)*. Here, it will be shown that number of $G8$-double primes, marked as *pd8,* is smaller than the number $\pi_{G8}/2$. In order to understand why it is so, we are going to analyze (2) and (4) in more detail.

It is not difficult to understand that m in (2) and (4) is represented by the threads that are defined by odd prime numbers. For details see Appendix A. Now we are going to compare stages 2 for generation of $G8$-double primes and *mps* $G8$-primes, step by step, for a few initial steps (analysis can be easily extended to any number of steps). Starting point for the second stage is the moment when the primes $G8$-fs are generated.

Table 1 Comparison of the stage 2 for generation of $G8$-double primes and *mps* $G8$-primes – threads defined by a few smallest primes

Step	Stage 2 – $G8$-double primes	Step	Stage 2 – *mps* $G8$-primes
1	Remove even numbers (except 2) amount of numbers left is 1/2	1	Remove numbers defined by thread defined by 3 (obtained for $i = 1$) amount of numbers left is 1/2
2	Remove numbers defined by thread defined by 3 (obtained for $i = 1$) amount of numbers left is 2/3 of the numbers that are left after previous step	2	Remove numbers defined by thread defined by 5 (obtained for $i = 2$) amount of numbers left is 3/4 of the numbers that are left after previous step
3	Remove numbers defined by thread defined by 5 (obtained for $i = 2$) amount of numbers left is 4/5 of the numbers that are left after previous step	3	Remove numbers defined by thread defined by 7 (obtained for $i = 3$) amount of numbers left is 5/6 of the numbers that are left after previous step
4	Remove numbers defined by thread defined by 7 (obtained for $i = 3$) amount of numbers left is 6/7 of the numbers that are left after previous step	4	Remove numbers defined by thread defined by 11 (obtained for $i = 5$) amount of numbers left is 9/10 of the numbers that are left after previous step

It should be kept in mind that in the case of $G8$-double primes sieve is implemented on indexes of

*G*8-fs prime numbers (indexes are consecutive numbers 1, 2, 3, …).

Values of the fractions presented in the Table 1 are asymptotically correct (in the finite case those values are only approximately correct - for details see [3]). **Here is important to understand that it is assumed, without a proof, that Dirichlet's theorem on arithmetic progressions [8] holds on the subset of primes marked as *G*8-fs primes (in other words, after removal of twin primes in *mps* form and after removal of the quarter of the sexy primes (half of the sexy primes in *mps* form) from prime numbers, that it is possible to select the same number of *mpl* primes that are going to be removed in such a way that Dirichlet's theorem on arithmetic progressions [8] still holds). For the arithmetic progressions that are defined by finite numbers this is a simple task, but it requires additional attention when we are speaking about progressions defined by infinite numbers.**

From the table, it can be noticed that threads defined by the same number in first and second column will not remove the same percentage of numbers. The reason is obvious – consider for instance the thread defined by 3: in the first column it will remove 1/3 of the numbers left, but in the second column 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 is asymptotically the same number of numbers that give residual -1 and numbers that give residual 1, when the number is divided by 3 (see [3, Appendix C]). Same way of reasoning can be applied for all other threads defined by the same prime in different columns.

From Table 1 can be seen that in every step, except step 1, threads in the second column will leave bigger percentage of numbers than the corresponding threads in the first column. This could be easily understood from the analysis that follows:

- suppose that we have two natural numbers j, k such that $j - 1 \geq k$ ($j, k \in N$), then the following set of equations is trivially true

$$j+k-1 \geqslant 2k$$

$$-j-k+1 \leqslant -2k$$

$$jk-j-k+1 \leqslant jk-2k$$

$$(j-1)(k-1) \leqslant (j-2)k$$

$$\frac{k-1}{k} \leqslant \frac{j-2}{j-1}$$

The equality sign holds only in the case $j = k + 1$. In the set of prime numbers there is only one case when $j = k + 1$ and that is in the case of primes of 2 and 3. In all other cases $p(i) - p(i - 1) > 1$, ($i > 1$, $i \in N$, $p(i)$ is i-th prime number). So, in all cases $i > 2$

$$\frac{p(i-1)-1}{p(i-1)} < \frac{p(i)-2}{p(i)-1} \quad .$$

From Table 1 (or last equation) we can see 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). From previous analysis we can safely conclude that the following equation holds

$$\pi_{G8} > \frac{\pi_{G8}}{2} > pd8 = \lim_{n \to \infty} pd8(n) \quad .$$

Having in mind (3) and (5), we can say that for some n big enough the following inequality holds

$$pd8(n) > \frac{\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)}{\ln\left(\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)\right)} . \tag{6}$$

(The n for which equations holds can be specified once the numbers in *mpl* form that are going to be eliminated are known – for instance it can be said that $n \geq 317$ can be chosen, since 317 is the 17th prime left, when number 2, all twin primes which means that *mps* twin primes are balanced by *mpl* twin primes, and a quarter of the sexy primes are eliminated from the prime numbers set, and if we assume that all corresponding *mpl* numbers (that "balance" a quarter of sexy primes that are numbers in *mps* form) that are going to be removed are all bigger than 317). It is not difficult to prove that the number of primes that are left when 2 and twin primes in *mps* form and quarter of sexy primes, and corresponding number of *mpl* primes are eliminated, is infinite. One way to understand it, is the fact that the relative density of twin primes and all primes is zero as well as the relative density of sexy primes and all primes is zero.

Having that in mind, it it easy to show that following holds

$$\lim_{n\to\infty} \frac{\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)}{\ln\left(\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)\right)} = \infty .$$

Then, the following equation holds

$$pd8 = \lim_{n\to\infty} pd8(n) = \infty .$$

Now, we can safely conclude that the number of *G*8-primes is infinite. That concludes the proof.

Here we will state the following conjecture: for *n* big enough, number of *g*8-primes is given by the following equation

$$\pi_{G8}(n) \sim 4C_2 \cdot \left(\frac{\left(\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)\right)}{\ln\left(\pi(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n)\right)} \right).$$

Constant C_2 is known as twin prime constant [9]. The equation can be expressed by using only *n*, but it would be cumbersome. Why it is reasonable to make such conjecture can be understood from the procedure that is similar to one used in [3, Appendix B]. If we mark the number of primes smaller than some natural number *n* with $\pi(n) \approx f(n)$, where function $f(n)$ gives good estimation of the number of primes smaller than *n,* than $\pi_{G8}(n)$, for *n* big enough, is given by the following equation

$$\pi_{G8}(n) \sim 4C_2 \cdot \left(f\left(f(n) - \pi_{G2}(n) - 0.5\pi_{G6}(n) \right) \right).$$

Here we can see that constant C_2 has a misleading name. It is connected with repeated implementation of a sieve that produces prime numbers which is also, but not exclusively, connected to the twin primes.

3 Analysis of the Polignac's conjecture for gaps equal to ten

As it was already explained in the introduction part, if two consecutive prime numbers have the gap of the size 10 it is clear that smaller those numbers has to be in *mpl* form while the bigger one has to be in *mps* form. Proof that number of the *G*10-primes is infinite can be done by trivial extension of the proof that infinitely many *G*8-primes exist. Here we are only going to analyze a two stage process that can be used for generation of the smaller primes of *G*10-prime pairs (*mpl* G10-primes), and how it differs from two stage process used for generation of smaller primes of *G*8-prime pairs. The first stage in both processes are equivalent and they produce prime numbers by removing all composite numbers from the set of natural numbers. Here we are going to analyze the second stage of the process that produces *mpl* *G*10-primes.

In order to do that we will consider 6 consecutive numbers

$$6k+1,\ 6k+3,\ 6(k+1)-1,\ 6(k+1)+1,\ 6(k+1)+3 \text{ and } 6(k+2)-1.$$

We can see that $6(k+1)$ and $6(k+2)-1$ can create a *G*10-primes only in the case when neither $6(k+1)-1$ nor $6(k+1)+1$ are primes, and at the same time both 6k+1 and 6(k+2) - 1 are primes (numbers $6k+3$ and $6(k+1)+3$ are obviously composite). Here we are going to analyze three cases:

1. Both $6(k+1)-1$ and $6(k+1)+1$ are primes. We can see that removal of all smaller primes of cousin prime pairs, removes primes in *mpl* form that potentially cannot have a *G*10-prime pair.

2. $6(k+1)-1$ is a prime number and $6(k+1)+1$ is a composite number. Again, we can see that removal of all smaller primes of cousin prime pairs removes primes in *mpl* form that potentially cannot have a *G10*-prime pair.

3. $6(k+1)-1$ is composite number and $6(k+1)+1$ is prime number. Now, we can see that beside

the removal of all smaller primes in cousin prime pairs, it is necessary to remove smaller primes in *mpl* sexy prime pairs (which is one half of *mpl* sexy primes and a quarter of all sexy primes), in order to remove all primes in *mpl* form that potentially cannot have a *G*10-prime pair.

4. Remove appropriate number of primes in *mps* form that match the number of *mpl* numbers that are removed in steps 1-3 of the second stage in such a way that Dirichlet's theorem on arithmetic progressions still holds.

From previous analysis we can conclude that in this part of the second stage it is necessary to remove all number of prime numbers that is equal to the number of cousin primes and one half of all sexy primes. After this has been done, the procedure can follow the line of reasoning that is used in proof that infinitely many *G*8-primes exist. The only small difference would be to prove that exist infinitely many primes that are not cousin primes or one half of sexy primes (and that follows from the fact that the ratio of density of cousin and sexy primes to all primes is zero). The other difference comes from the fact that 10/2=5 and that means that the thread that is defined by 5, in the second stage, will not remove any additional numbers. That will be analyzed in Appendix C.

This analysis can easily be extended to gaps g, where $g=6k+2$, $k \in N$.

4. Analysis of the Polignac's conjecture for gaps equal to twelve

The *G*12-primes are successive prime numbers with the gap 12. It was already mentioned that in the case when gap $g=6k$, where k is natural number, we can have *G*12-prime pairs in (*mps*, *mps*) as well as (*mpl*, *mpl*) forms.

It is clear that we can generate *G12*-primes using the two stage sieve process that is similar to the process used for generation of *G*6-primes [4]. The only difference will be in the part of the second stage that is going to be analyzed here. The first stages are identical. The difference in the second stage is related to the removal of the primes that potentially cannot produce a *G*12-prime pairs.

In order to see what kind of primes should be removed from the set of prime numbers. Here, we are going to analyze *mps* $G12$-primes (*mpl* $G12$-primes can be analyzed analogously). We are going to analyze the following 7 consecutive odd numbers

$$6k-1,\ 6k+1,\ 6k+3,\ 6(k+1)-1,\ 6(k+1)+1,\ 6(k+1)+3 \text{ and } 6(k+2)-1.$$

So, if want to have *G12*-prime pair ($6k-1$, $6(k+2)-1$) it is clear that numbers $6k+1$, $6(k+1)-1$ and $6(k+1)+1$ cannot be prime numbers. Here we are going to analyze all possible cases.

1. All three numbers $6k+1$, $6(k+1)-1$ and $6(k+1)+1$, are primes. In that cases it is clear that removal of all smaller primes in the twin prime pairs remove *mps* primes that potentially cannot have *G12*-prime pair.

2. $6k+1$, $6(k+1)-1$ are prime numbers and $6(k+1)+1$ is composite number. In this case beside removal of twin primes it is necessary to remove a smaller number in all *mps* sexy prime pairs (that represent half of *mps* sexy primes, or a quarter of all sexy primes).

3. $6k+1$, $6(k+1)+1$ are prime numbers and $6(k+1)-1$ is composite number. Removal of the smaller primes in twin prime pairs is enough,

4. $6(k+1)+1$, $6(k+1)-1$ are prime numbers and $6k+1$ is composite number. In this case beside removal of twin primes it is necessary to remove a smaller number in all *mps* sexy prime pairs (that represent half of *mps* sexy primes, or a quarter of all sexy primes).

5. $6k+1$ is prime number and $6(k+1)+1$, $6(k+1)-1$ are composite numbers. What is required is removal of twin primes.

6. $6(k+1)-1$ is prime number and $6k+1$, $6(k+1)+1$ are composite numbers. In this case removal of smaller number in the *mps* sexy prime pairs is required.

7. $6(k+1)+1$ is prime number and $6k+1$, $6(k+1)-1$ are composite numbers. In this case it is necessary to remove smaller prime from all $G8$-prime pairs.

Now it is necessary to remove appropriate number of *mpl* primes that is equal to the number of *mps* primes removed in all cases (1-7) that were analyzed. That should be done in such a way that

Dirichlet's theorem on arithmetic progressions holds (this is assumed as possible without proof). Now, the rest of the procedure that is applied in the G6-primes case has to be done in G12-primes case.

Case of *mpl* $G12$-primes can be analyzed analogously. Of course in that case it would be necessary to remove all cousin and $G10$-primes, as well as *mpl* sexy primes, instead of twin primes, G8-primes and *mps* sexy primes.

This analysis can easily be extended to gaps g, where $g=6n$, $n \in N$.

4. Proof that the number of *Gg*-primes is infinite

Here we are going to analyze the general case of the gap of the size $g > 6$. From previous analysis, it is very simple to understand that *Gg*-primes could be obtained by two stage process.

STAGE 1

Using the same methodology as previously, generate all prime numbers. In order to do that, from the set of all natural numbers bigger than 1, remove all even numbers (except 2) and all odd numbers generated by equation (2).

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 their bigger *Gg*-odd neighbor must be in the form $2n + 1 + g$, $n, g \in N, g > 6$. Now, we should implement a second step in which we are going to **remove**:

A. Number 2 (since 2 cannot make *Gg*-prime pair), but this has no impact on the analysis,

B. **If mod (*g*, 6) = 2:** all smaller primes of twin prime pairs, all smaller primes in *s*-prime pairs, where s is even natural number in the form $6x + 2$ and $s < g$ and a quarter of all *k*-primes (half of the number of *k*-primes in *mps* form), where k represents number in the form $6y$ and $k < g$ ($x, y \in N$).

The adequate number of primes in *mpl* form that is equal to the number of *mps* primes that is removed in this step. The number of numbers that is left after this is number of primes minus the number of twin primes minus all *s*-primes minus one half of *k*-primes. It is not difficult to prove that number of primes left after this step is infinite.

If mod(*g*, 6) = 4: all smaller primes in cousin prime pairs and all smaller primes in *l*-prime pairs, where l is even natural number in the form $6x + 4$ and $l < g$, and *a quarter* primes of all *k*-prime pairs (a half of the *k*-primes in *mpl* form), where k represents number in the form $6y$ and $k < g$ ($x, y \in N$). The adequate number of primes in *mps* form that is equal to the number of *mpl* primes that is removed in this step. The number of numbers that is left after this is number of primes minus the number of cousin primes, minus all *l*-primes, minus one half of all *k*-primes. It is not difficult to prove that number of primes left after this step is infinite.

If mod(*g*, 6) = 0: Assume that we are working with numbers in *mps* form, we remove smaller primes from *k*-primes pairs, k is even natural number in the form $6x$, and $k < g$; smaller prime of all twin prime pairs and *s*-prime pairs, where s is even natural number in the form $6x + 2$ and $s < g$. The adequate number of primes in *mpl* form that is equal to the number of *mps* primes that is removed in this step.

If we work with primes in *mpl* form, we remove smaller primes from *k*-primes pairs, k is even natural number in the form $6x$, and $k < g$, smaller prime of all cousin prime pairs and all *l*-prime pairs, where l is even natural number in the form $6x + 4$ and $l < g$, The adequate number of primes in *mps* form that is equal to the number of *mpl* primes that is removed in this step.

It is not difficult to prove that number of primes left after this step is infinite.

C. **If mod (*g*, 6) = 2**, all *mpl* primes – 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, one half of the numbers that are left after first stage;

If mod (*g*, 6) = 4, all *mps* primes with same consequences as previously;

If mod (*g*, 6) = 0, if we are working with *mps* primes, remove rest of *mpl* primes. If we work with *mpl* primes remove rest of *mps* primes.

D. Assume now that we are working with numbers in *mps* form (everything should be done analogously for numbers in *mpl* form). Remove all odd primes in the form $2m + 1$ such that $2m + g$, $m \in N$, represents a composite number (those that have composite bigger *g*-odd neighbor). If we make the same analysis again, it is simple to understand that *m* must be in the form

$$m = 2ij + i + j - 1 = (2i + 1)\, j + i - g/2. \tag{7}$$

All numbers (in observational space) that are going to stay must be numbers in *mps* form and they represent a smaller primes of the *Gg*-prime pairs.

It should be noticed that if g/2 is divisible by the prime (2i+1) that defines thread in (7), there will be no additional removals of the numbers defined by that thread in STAGE2. That will lead to an additional correction of the formula for the estimation of number of Gg-primes smaller than some natural number n (see Appendix C).

Following identical procedure like in the case of sexy or *G8*-primes it is possible to it possible to prove that the number of *Gg*-primes is infinite.

Here, we are going to conjecture several formulas for the number of *Gg*-primes. Why this is reasonable can be seen in Appendix B and C. If we mark the number of primes smaller than some natural number *n* with $\pi(n) \approx f(n)$, where function $f(n)$ gives good estimation of the number of primes smaller than *n,* than $\pi_{Gg}(n)$, for *n* big enough, is given by the following equation

a. Case mod(g, 6) = 2

$$\pi_{Gg}(n) \sim C_G(g) \cdot 4C_2 \cdot \left(f\left(f(n) - \pi_{G2}(n) - \sum_{2<S<g} \left(\pi_{GS}(n)\right) - 0.5 \sum_{0<K<g} \left(\pi_{GK}(n)\right)\right)\right),$$

where *S* are even numbers in the form 6*x*+2 and *K* is an even number in the form 6*y*, $x, y \in N,$ and $C_G(g)$ is correction constant for gap *g* (for details, see Appendix C). If we us $f(n) = MoLi(n)$ [10]

and we define $P_g(n)$ as

$$P_g(n) = \pi(n) - \pi_{G2}(n) - \sum_{2<S<g} \left(\pi_{GS}(n)\right) - 0.5 \sum_{0<K<g} \left(\pi_{GK}(n)\right) \quad ,$$

then we have

$$\pi_{Gg}(n) \sim C_G(n)\, 4C_2 \cdot \int_2^{P_g(n)} \frac{dx}{\ln\left(x+\sqrt{P_g(n)}\right)} \quad .$$

b. Case mod(g, 6) = 4

$$\pi_{Gg}(n) \sim C_G(g) \cdot 4C_2 \cdot \left(f\left(f(n) - \pi_{G4}(n) - \sum_{4<L<g} \left(\pi_{GL}(n)\right) - 0.5 \sum_{0<K<g} \left(\pi_{GK}(n)\right)\right)\right),$$

where L are even numbers in the form $6x+4$ and K is an even number in the form $6y$, $x,y \in N$, and $C_G(g)$ is correction constant for gap g (for details, see Appendix C).

If we us $f(n) = MoLi(n)$ [10] and we define $P_g(n)$ as

$$P_g(n) = \pi(n) - \pi_{G4}(n) - \sum_{4<L<g} \left(\pi_{GL}(n)\right) - 0.5 \sum_{0<K<g} \left(\pi_{GK}(n)\right) \quad ,$$

then we have

$$\pi_{Gg}(n) \sim C_G(n)\, 4C_2 \cdot \int_2^{P_g(n)} \frac{dx}{\ln\left(x+\sqrt{P_g(n)}\right)} \quad .$$

c. Case mod(g, 6) = 0

$$\begin{aligned}\pi_{Gg}(n) \sim\ & C_G(g) \cdot 4C_2 \cdot \left(f\left(f(n) - \pi_{G2}(n) - \sum_{1\leqslant K<g} \left(0.5\,\pi_{GK}(n)\right) - \sum_{2<S<g} \left(\pi_{GS}(n)\right)\right)\right) \\ & + C_G(g) \cdot 4C_2 \cdot \left(f\left(f(n) - \pi_{G4}(n) - \sum_{1\leqslant K<g} \left(0.5\,\pi_{GK}(n)\right) - \sum_{4<L<g} \left(\pi_{GL}(n)\right)\right)\right),\end{aligned}$$

and S, L, K and $C_G(g)$ are defined as previously.

If we us $f(n) = MoLi(n)$ [10] and we define $P_{gS}(n)$ as

$$P_{gS}(n) = \pi(n) - \pi_{G2}(n) - \sum_{1\leqslant K<g} \left(0.5\,\pi_{GK}(n)\right) - \sum_{2<S<g} \left(\pi_{GS}(n)\right) \quad ,$$

and $P_{gL}(n)$ as

$$P_{gL}(n) = \pi(n) - \pi_{G4}(n) - \sum_{1 \leqslant K < g} \left(0.5\,\pi_{GK}(n)\right) - \sum_{4 < L < g} \left(\pi_{GL}(n)\right) \quad ,$$

then we have

$$\pi_{Gg}(n) \sim C_G(g) \cdot 4C_2 \cdot \int_2^{P_{gS}(n)} \frac{dx}{\ln\left(x + \sqrt{P_{gS}(n)}\right)} + C_G(g) \cdot 4C_2 \cdot \int_2^{P_{gL}(n)} \frac{dx}{\ln\left(x + \sqrt{P_{gL}(n)}\right)} \quad .$$

References

[1] A. de Polignac. (1849). Recherches nouvelles sur les nombres premiers. Comptes Rendus des Séances de l'Académie des Sciences.

[2] D.H.J. Polymath. (2014) Variants of the Selberg sieve and bounded intervals containing many primes. Research in the Mathematical Sciences. 1(12), arxiv:1407.4897.

[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 15.

[4] M. V. Jankovic. (2021). Proof of Polignac's Conjecture for Gap Equal to Six (Proof that exists infinitely many sexy primes), hal-03323045, version 2.

[5] M. V. Jankovic. (2022). Proof of Polignac's Conjecture for Gap Equal to Eight , hal-03656705, version 1.

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

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

[8] P. G. L. Dirichlet. (1837) Bewies des Satzes, dass jede unbregenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftluchen Factor sind, unendlich

viele Primzahlen enthält. Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften zu Berlin, 48, pp. 45-71.

[9] G.H. Hardy, J.E. Littlewood. (1923) Some Problems of 'Partitio Numerorum.' III. On the Expression of a Number as a Sum of Primes, Acta Math. 44, pp.1-70.

[10] M.V. Jankovic. (2021) A New Approximation of Prime Counting Function Based on Modified Logarithmic Integral. viXra: 2110.0133v2.

APPENDIX A.

Here it is going to be proved that m in (2) is represented by threads defined by odd prime numbers. Now, the form of (2) for some values of i will be checked.

Case i = 1: $m = 3j + 1$,

Case i = 2: $m = 5j + 2$,

Case i = 3: $m = 7j + 3$,

Case i = 4: $m = 9j + 4 = 3(3j + 1) + 1$,

Case i = 5: $m = 11j + 5$,

Case i = 6: $m = 13j + 6$,

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

Case i = 8: $m = 17j + 8$,

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), (5) and (7).

Note: *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 asymptotic density of numbers left, after implementation of the first and second Sundaram sieve is calculated.

After first k steps of the first Sundaram sieve, after removal of all composite even numbers, density of numbers left is given by the following equation

$$c_k = \frac{1}{2}\prod_{j=2}^{k+1}\left(1 - \frac{1}{p(j)}\right),$$

where $p(j)$ is j-th prime number.

In the case of second "Sundaram" sieve the density of numbers left after the first k-steps is given by the following equation

$$c2_k = \prod_{j=2}^{k+1}\left(1 - \frac{1}{p(j)-1}\right) = \prod_{j=2}^{k+1}\left(\frac{p(j)-2}{p(j)-1}\right).$$

So, if implementation of first sieve will result in the number of prime numbers smaller than n which we denote as $\pi(n)$, than implementation of the second sieve on some set of size $\pi(n)$ should result in the number of numbers $gp(n)$ that are defined by the following equation (for some big enough n)

$$gp(n) = r_{S2S1}(n) \cdot \frac{\pi(n)}{\ln(\pi(n))} \quad ,$$

where $r_{S2S1}(n)$ is defined by the following equation (k is the number of primes smaller or equal to $n1 = sqrt(n)$, where $sqrt$ marks square root function)

$$r_{S2S1}(n) = \frac{c2_k}{c_k} = \frac{\prod_{p>2,\, p\leq n1}\left(\frac{p-2}{p-1}\right)}{\prod_{p\leq n1}\left(\frac{p-1}{p}\right)} = 2\prod_{p>2,\, p\leq n1}\left(\frac{p-2}{p-1}\right)\left(\frac{p}{p-1}\right) \approx 2\mathrm{C}_2 .$$

For n that is not big, $gp(n)$ should be defined as

$$gp(n) = f_{COR}(n) \cdot 2\mathrm{C}_2 \cdot \frac{\pi(n)}{\ln(\pi(n))},$$

where $f_{COR}(n)$ represents correction factor that asymptotically tends toward 1 when n tends to infinity.

APPENDIX C.

Here we are going to repeat equation (7)

$$m = 2ij + i + j - 1 = (2i + 1)\,j + i - g/2.$$

From equation is quite clear that thread defined by the prime $(2i + 1)$ in STAGE 2 will not produce new removals of numbers in addition to removals produced by the thread defined by the same prime in the STAGE 1, in the case when $(2i+1)$ divides $g/2$. That means that r_{S2S1} defined in Appendix B has to be changed. From Appendix B we know that r_{S2S1} is defined by the following equation

$$r_{S2S1}(n) = \frac{c2_k}{c_k} = \frac{\prod_{p>2,\, p\leq n}\left(\frac{p-2}{p-1}\right)}{\prod_{p\leq n}\left(\frac{p-1}{p}\right)} = 2\prod_{p>2,\, p\leq n}\left(\frac{p-2}{p-1}\right)\left(\frac{p}{p-1}\right) \approx 2\mathrm{C}_2,$$

while $c2_k$ is defined by the following equation

$$c2_k = \prod_{j=2}^{k+1}\left(1 - \frac{1}{p(j)-1}\right) = \prod_{j=2}^{k+1}\left(\frac{p(j)-2}{p(j)-1}\right).$$

In the case when prime $p(j)$ divides $g/2$ corresponding term in $c2_k$ is actually 1 instead of

$$\frac{p(j)-2}{p(j)-1}.$$

That means that $c2_k$ has to be multiplied by the following term

$$\frac{p(j)-1}{p(j)-2},$$

in order to obtain the correct equation.

If we mark with Z set of all primes $3 < p \leq n1$ ($n1 = sqrt(n)$) that divide $g/2$, a constant $C_G(g)$ is defined by the following equation

$$C_G(g) = \prod_{p\in Z}\left(\frac{p-1}{p-2}\right).$$

Now, we can write the proper values for $c2_k$ and r_{S2S1} as

$$c2_k = C_G(g) \prod_{j=2}^{k+1} \left(\frac{p(j) - 2}{p(j) - 1} \right),$$

and

$$r_{S2S1}(n) = C_G(g) \cdot \frac{c2_k}{c_k} = C_G(g) \cdot 2 \prod_{p>2, p\leq n} \left(\frac{p-2}{p-1} \right) \left(\frac{p}{p-1} \right) \approx C_G(g) \cdot 2\mathrm{C}_2.$$