

In the name of the only God Allah, Allah the Lord of the worlds, the most Gracious the most Merciful.

A proof of Legendre's conjecture and Andrica's conjecture

Ahmad Taher Telfah

Jordan-Irbid

December 21,2018

E-mail ahmadt79@hotmail.com

Abstract

this paper carrying a method to introduce the distribution of the densities of the prime numbers and the composite numbers along in natural numbers, the method basically depends on the direct deduction of the composite numbers in a specified intervals that also with using some corrections and modifications to reach maximum and minimum values of the composites and primes densities, this allowed us to detect some special conjectures related to the primes density.

Introduction

Legendre's conjecture states that there is a prime number between m^2 and $(m + 1)^2$ for every positive integer m . The conjecture is one of Landau's problems (1912) on prime numbers, an elementary and asymptotically proofs of Legendre's conjecture are available but not considered as sufficient proofs, the methodology used in this paper by conducting the primes density in a specific intervals for the successive natural numbers $I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, P_n^2 + 3, \dots, P_{n+1}^2 - 1\}$, in this specific intervals the specific primes density is $(d(P_n)) = (\prod_{i=1}^{i=n} (1 - \frac{1}{p_i}))$, as shown bellow by understanding the principle of constructing the primes density we can adding some corrections which allowed us to reach the minimum possible logical primes density which $(d(P_n))_{min} = (\frac{1}{2} \prod_{i=2}^{i=n} (1 - \frac{2}{p_i}))$, the minimum primes density lead us to calculate the largest possible prime gaps as $\frac{1}{dP_{n,min}} + \frac{1}{dP_{n+1,min}} - 1 > g_n$, defined such large gaps allowed us to detect some related conjectures, all of those conjectures in concept are cored by Legendre's conjecture.

keywords: Legendre's conjecture, Landau's problems, prime gaps, Andrica's conjecture,

Legendre's conjecture, proposed by Adrien-Marie Legendre, states that there is a prime number between m^2 and $(m + 1)^2$ for every positive integer m . The conjecture is one of Landau's problems (1912) on prime numbers.

At the 1912 International Congress of Mathematicians, Edmund Landau listed four basic problems about primes. These problems were characterized in his speech as "unattackable at the present state of mathematics" and are now known as Landau's problems. They are as follows:

1. Goldbach's conjecture: Can every even integer greater than 2 be written as the sum of two primes?
2. Twin prime conjecture: Are there infinitely many primes P such that $P + 2$ is prime?
3. Legendre's conjecture: Does there always exist at least one prime between consecutive perfect squares?
4. Are there infinitely many primes P such that $P - 1$ is a perfect square? In other words: Are there infinitely many primes of the form $m^2 + 1$?

Legendre's conjecture is one of a family of results and conjectures related to prime gaps, that is, to the spacing between prime numbers.

A prime number (or a prime) is a natural number greater than 1 that cannot be formed by multiplying two smaller natural numbers. A natural number greater than 1 that is not prime is called a composite number, Every positive integer is composite, prime, or the unit 1.

A prime gap is the difference between two successive prime numbers. The n^{th} prime gap, denoted g_n or g_{p_n} is the difference between the $(n + 1)^{th}$ and the n^{th} prime numbers,

$$g_n = P_{n+1} - P_n$$

N is the group of natural numbers (1,2,3,, m)

$G(C_m)$: the group of composite numbers.

$G(P_m)$: the group of prime numbers

$$N = \{1 + G(C_m) + G(P_m)\}$$

for the sub interval I as, $I \in N$ then $I = \{G'(C_m) + G'(P_m)\}$

as $G'(C_m)$: group of the composite numbers in the interval I .

$G'(P_m)$: group of the prime numbers in the interval I .

Methodology:

Defined the interval $I(P_n)$, as $I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, P_n^2 + 3, \dots, P_{n+1}^2 - 1\}$

P_n represent the maximum prime number in the consecutive prime numbers 2,3,5,7,11,, P_n

$$P_1 = 2, P_2 = 3, P_3 = 5, P_4 = 7, \dots, P_n = P_n$$

the expectation value of the number of the composites in the interval of length $L \in I(P_n)$ is $\langle \#C_n \rangle$
 $\langle \#C_n \rangle = L(d(C_n)) = L(1 - \prod_{i=1}^{i=n} (1 - \frac{1}{P_i}))$ (6)

$$(d(P_n)) + (d(C_n)) = 1 \quad (7)$$

The weakness points of the distribution function :

1. There is some deviation when we deduct the composite numbers due to the boundary of the interval and the length of the interval.
2. The distribution function depends on the distribution of the primes which actually we don't know it.

To avoid and fix the first weakness point we have to deal and fix the deduction factor $\frac{1}{P_i}$ in every part of the composite density distribution part to be matching with the boundary of the interval and its length, as an example take the interval $\{10,11,12,13,14\}$ with length 5 the expected numbers factored by 2 is $5/2$ but actually it is 3, and the interval $\{11,12,13,14,15\}$ with length 5 and the expected numbers factored by 2 is $5/2$ but actually it is 2, and so on, for other deduction factors we may face such deviation, there are many methods to fix this deviation and to reach the exact correction we have to fix every deduction factor independently, but in this work we care about maximizing the composite numbers to reach the least primes density, actually the reality of the deduction factor $\frac{1}{P_i}$ is as an expression of an order to deduct one number from every P_i length of the consecutive numbers in the interval I of length L , from this expression we can direct fix the deduction factor $\frac{1}{P_i}$ by make it $\frac{2}{P_i}$ which imply an order to deduct two numbers from every P_i length of consecutive numbers in the interval I of length L , and that is especially for long intervals more than sufficient to maximize the composite numbers and reduce the primes, by applying this but with excluding the factor $\frac{1}{2}$ as it is destroy the distribution and make it with zero primes, we will get that, the corrected density:

$$\text{maximum composites density} \quad (d(C_n))_{\max} = (\frac{1}{2} + \frac{2 \cdot 1}{3 \cdot 2} + \frac{2 \cdot 1 \cdot 1}{5 \cdot 2 \cdot 3} + \frac{2 \cdot 1 \cdot 1 \cdot 3}{7 \cdot 2 \cdot 3 \cdot 5} + \dots + \frac{2}{P_n} \frac{1}{2} \prod_{i=2}^{i=n-1} (1 - \frac{2}{P_i}))$$

And the corrected density of primes:

$$\text{The minimum primes density} \quad (d(P_n))_{\min} = (\frac{1}{2} \prod_{i=2}^{i=n} (1 - \frac{2}{P_i})) \quad (8)$$

$$\text{The maximum composite density} \quad (d(C_n))_{\max} = (1 - \frac{1}{2} \prod_{i=2}^{i=n} (1 - \frac{2}{P_i})) \quad (9)$$

But for more accuracy we will construct another correction factor to clarify the meaning and the effect of the deduction factor, by studying the deduction factor $\frac{1}{P_i}$ Actually for a length L as P_i , the factor $\frac{1}{P_i}$ able for correction for all $\frac{1}{P_i}$ as P_i is prime and not factored by any P_i except itself, now to correct any factor $\frac{1}{P_i}$ multiplied by L to reach the maximum expected deduction value we have to insert the

correction value $\frac{P_i-1}{P_i}$ to the deduction factor $\frac{1}{P_i}$ to make it $\frac{1+(\frac{P_i-1}{P_i})}{P_i}$ so when we multiply the corrected

deduction factor by L as that $\frac{1+(\frac{P_i-1}{P_i})}{P_i} L$ we will reach the maximum expected composite numbers

factored by the prime P_i , before applying this we will simplify the corrected factor $\frac{1+(\frac{P_i-1}{P_i})}{P_i}$ to

$$\frac{P_i+P_i-1}{P_i P_i} = \frac{2P_i-1}{P_i^2}$$

By applying the corrected factor and converting the deduction factor $\frac{1}{P_i}$ to $\frac{2P_i-1}{P_i^2}$ we will get

$$1 - \left\{ \frac{2P_1-1}{P_1^2} + \frac{2P_2-1}{P_2^2} \left(\frac{P_1-1}{P_1^2} \right) + \frac{2P_3-1}{P_3^2} \left(\frac{P_1-1}{P_1^2} \frac{P_2-1}{P_2^2} \right) + \dots + \frac{2P_n-1}{P_n^2} \prod_{i=1}^{i=n-1} \left(\frac{P_i-1}{P_i^2} \right) \right\} = \prod_{i=1}^{i=n} \left(\frac{P_i-1}{P_i^2} \right)$$

So

$$(d(P_n))_{\min} = \prod_{i=1}^{i=n} \left(\frac{P_i-1}{P_i^2} \right) \quad (10)$$

$$(d(C_n))_{\max} = 1 - \left\{ \prod_{i=1}^{i=n} \left(\frac{P_i-1}{P_i^2} \right) \right\} \quad (11)$$

In this case we fix the first weakness point by reaching the minimum possible prime density and maximum possible composite density, but we still stokes with the primes distribution problem to be able to determine any value for those densities.

Now

$$\begin{aligned} (d(P_n))_{\min} &= \prod_{i=1}^{i=n} \left(\frac{P_i-1}{P_i^2} \right) \\ &= \prod_{i=1}^{i=n} \frac{P_i - (1 + (\frac{P_i-1}{P_i}))}{P_i} \\ &= \frac{2 - (1 + (\frac{2-1}{2}))}{2} \frac{3 - (1 + (\frac{3-1}{3}))}{3} \frac{5 - (1 + (\frac{5-1}{5}))}{5} \dots \frac{P_n - (1 + (\frac{P_n-1}{P_n}))}{P_n} \end{aligned}$$

Always $\left(\frac{P_i-1}{P_i} \right) < 1$ as $i \in (1, 2, 3, 4, \dots, n)$

Then $\left(1 + \left(\frac{P_i-1}{P_i} \right) \right) < 2$ as $i \in (1, 2, 3, 4, \dots, n)$

For long intervals or intervals of length with the order of $2n$ we can exclude the correction of the first

term for factor $\frac{2 - (1 + (\frac{2-1}{2}))}{2}$ then we will get

$$\left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i} \right) \right) \cong \frac{1}{2} \prod_{i=2}^{i=n} \left(\frac{P_i-1}{P_i^2} \right)$$

$$\left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right) \leq \frac{1}{2} \prod_{i=2}^{i=n} \left(\frac{(P_i-1)^2}{P_i^2}\right)$$

So we will take in the consideration the following densities

$$\text{The minimum primes density} \quad (d(P_n))_{\min} = \left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right) \quad (12)$$

$$\text{The maximum composites density} \quad (d(C_n))_{\max} = \left(1 - \frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right) \quad (13)$$

Which is approximately the same result as the first direct correction.

$$\langle \#P_n \rangle_{\min=L} = L \left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right) \quad (14)$$

$$\langle \#C_n \rangle_{\max=L} = L \left(1 - \left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right)\right) \quad (15)$$

As $\langle \#P_n \rangle_{\min}$ is the minimum possible expectation value of the number of the primes in the interval of length $L \in I(P_n)$.

And $\langle \#C_n \rangle_{\max}$ is the maximum possible expectation value of the number of the composites in the interval of length $L \in I(P_n)$.

To handle some problems related to that, such as Legendre's conjecture and since Legendre's interval $\in I(P_n)$ and its length $2m \geq 2P_n$ so we will apply $\langle \#P_n \rangle_{\min=L} = L \left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right)$

Any Legendre interval $I_{Leg} = (m^2, m^2 + 1, m^2 + 2, \dots, (m+1)^2)$ then $I_{Leg} \in I(P_n)$ as $P_{n+1}^2 > m^2 \geq P_n^2$

Defined L_{Leg} is the length of the Legendre interval. $L_{Leg} = (m+1)^2 - m^2 - 1 = 2m$.

$$\langle \#P_n \rangle_{\min} = 2m \left\{ \frac{1}{2} \frac{1}{3} \frac{3}{5} \frac{5}{7} \frac{9}{11} \dots \frac{P_i-2}{P_i} \dots \frac{P_n-2}{P_n} \right\}$$

$$= 2m \frac{1}{2} \frac{1}{P_n} \left\{ \frac{3}{3} \frac{5}{5} \frac{9}{7} \frac{11}{11} \dots \frac{P_{i+1}-2}{P_i} \dots \frac{P_n-2}{P_{n-1}} \right\}$$

$$\mu = \left\{ \frac{3}{3} \frac{5}{5} \frac{9}{7} \frac{11}{11} \dots \frac{P_{i+1}-2}{P_i} \dots \frac{P_n-2}{P_{n-1}} \right\}, \mu \geq 1 \text{ since the prime gap } g_n \geq 2 \text{ for all } i \geq 2$$

$$m \geq P_n \text{ then } m/P_n \geq 1$$

$$\langle \#P_n \rangle_{\min} = \mu m/P_n \geq 1$$

Which mean the sufficient length to obtain one prime number in the interval $I(P_n)$ is $2m$

Which proof that Legendre's conjecture there is a prime number between m^2 and $(m + 1)^2$ for every positive integer m , but this still weak proof as we don't know the exact location of the prime number in the length L and we need to proof the successively of the intervals lengths, a stronger proof will be along with the proof of Andrica's conjecture.

to handle **Andrica's conjecture**

When $L(d(P_n)_{\min} \geq 1$ it means we found the sufficient length to find the first prime number as long as L is starting with the lower bound P_n^2 of the interval $I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, \dots, P_{n+1}^2 - 1\}$ so we can say the interval of such L to be contained at least one prime number with unknown prime location if we take a consecutive such length we will find another prime number with unknown location, such scenario lead us to conclude that may some large gap defined with length of $2(L - 1)$ is bounded by two primes P_i, P_{i+1} then the maximum possible prime gap g_i in the depth of the interval $I(P_n)$ will be within L as $L=2/(d(P_n)_{\min}$, but from another point of view and if the interval of length L is near the upper bound of the interval $I(P_n)$ which is $(P_{n+1}^2 - 1)$ that may also introduce an interference with the next interval $I(P_{n+1})$ which carry possibility to create another large gap due to the interference in this area that kind of interference in this area is possible between the primes in the consecutive intervals $I(P_n)$ and $I(P_{n+1})$, if such gap with such interference is exist so it is possible to be the largest ever gap if and only if its length is less than the gap between P_n^2 and P_{n+1}^2 which is already detected by the weak proof of Legendre's conjecture, even by considering both P_n and P_{n+1} are twin primes, this is mean there are two kinds of large gap the first in the depth of the interval $I(P_n)$ which we will call it the normal large gap and the second which conformed by the interference of two intervals $I(P_n)$ and $I(P_{n+1})$ which we will call it the interference large gap,

$$\frac{2}{dP_{n,\min}} - 1 = g_{large} \quad \text{normal large gap} \quad (16)$$

$$\frac{1}{dP_{n,\min}} + \frac{1}{dP_{n+1,\min}} - 1 = g_l \quad \text{interference large gap} \quad (17)$$

When **Andrica's conjecture** states that $\sqrt{P_{n+1}} - \sqrt{P_n} < 1$, or $g_n < 2\sqrt{P_n} + 1$ so that what we will detect to insure this conjecture, by detecting when $g_{large} < 2P_n + 1$ and so for more certainty we have to detect when $g_l < 2P_n + 1$ or for interference large gap.

Now

$$\begin{aligned} g_{large} &= \frac{2}{dP_n} - 1 \\ &= 2 \frac{1}{dP_n} - 1 \\ &= 2 \frac{1}{\left(\frac{1}{2} \prod_{i=2}^{i=n} \left(1 - \frac{2}{P_i}\right)\right)} - 1 \\ &= \left\{2 \left(\frac{2}{1} \frac{3}{1} \frac{5}{3} \frac{7}{5} \dots \dots \frac{P_i}{P_{i-2}} \dots \dots \frac{P_n}{P_{n-2}}\right)\right\} - 1 \end{aligned}$$

$$= 4P_n \left(\frac{3}{3} \frac{5}{5} \frac{7}{9} \dots \dots \frac{P_i}{P_{i+1}-2} \dots \dots \frac{P_{n-1}}{P_{n-2}} \right) - 1$$

$$\text{Consider } \varepsilon = \left(\frac{3}{3} \frac{5}{5} \frac{7}{9} \dots \dots \frac{P_i}{P_{i+1}-2} \dots \dots \frac{P_{n-1}}{P_{n-2}} \right) \quad (18)$$

ε is decreasing along with n and $\varepsilon < 1 \forall n \geq 5$ or $P_n \geq 11$

$$g_{large} = 4P_n \varepsilon - 1 \quad (19)$$

$$\varepsilon < \frac{1}{2} \quad \forall n \geq 12 \text{ or } P_n \geq 37$$

So

$$g_{large} < 2P_n + 1 \quad \forall n \geq 12 \text{ or } P_n \geq 37$$

g_{large} is the greatest possible prime gap in the depth of the interval

$$I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, \dots, P_{n+1}^2 - 1\}$$

Which mean Andrica's conjecture working for all primes $> 37^2$ in the depth of the interval $I(P_n)$ $\forall n \geq 12$. and so it is a stronger proof for Legendre's conjecture in the same area.

For the large gap of interference,

$$g_l = \frac{1}{dP_{n,min}} + \frac{1}{dP_{n+1,min}} - 1$$

$$g_l = \left(\frac{2}{1} \frac{3}{1} \frac{5}{3} \frac{7}{5} \dots \dots \frac{P_i}{P_{i-2}} \dots \dots \frac{P_n}{P_{n-2}} \right) + \left(\frac{2}{1} \frac{3}{1} \frac{5}{3} \frac{7}{5} \dots \dots \frac{P_i}{P_{i-2}} \dots \dots \frac{P_n}{P_{n-2}} \frac{P_{n+1}}{P_{n+1}-2} \right) - 1$$

$$g_l = \left\{ \left(\frac{2}{1} \frac{3}{1} \frac{5}{3} \frac{7}{5} \dots \dots \frac{P_i}{P_{i-2}} \dots \dots \frac{P_n}{P_{n-2}} \right) \left(1 + \frac{P_{n+1}}{P_{n+1}-2} \right) \right\} - 1$$

$$g_l = \left\{ 2P_n \left(\frac{3}{3} \frac{5}{5} \frac{7}{9} \dots \dots \frac{P_i}{P_{i+1}-2} \dots \dots \frac{P_{n-1}}{P_{n-2}} \right) \left(1 + \frac{P_{n+1}}{P_{n+1}-2} \right) \right\} - 1$$

$$\text{Consider } \varepsilon = \left(\frac{3}{3} \frac{5}{5} \frac{7}{9} \dots \dots \frac{P_i}{P_{i+1}-2} \dots \dots \frac{P_{n-1}}{P_{n-2}} \right)$$

ε is decreasing along with n , and $\varepsilon < 1 \forall n \geq 5$ or $P_n \geq 11$

$$\beta = \left(1 + \frac{P_{n+1}}{P_{n+1}-2} \right) \quad \beta > 2 \quad \forall n \quad \beta \text{ is greater when } P_n \text{ and } P_{n+1} \text{ are twin primes so,}$$

$$\beta_{max} = \left(1 + \frac{P_n + 2}{P_n + 2 - 2} \right)$$

$$\beta_{max} = \left(2 + \frac{2}{P_n} \right)$$

$$g_l = \{2P_n \left(2 + \frac{2}{P_n}\right) \varepsilon\} - 1$$

$$g_l = \{(4P_n + 4)\varepsilon\} - 1 \tag{20}$$

$$\varepsilon < \frac{1}{2} \quad \forall n \geq 12 \text{ or } P_n \geq 37$$

So
$$g_l < 2P_n + 1 \quad \forall n \geq 12 \text{ or } P_n \geq 37$$

g_l is the greatest possible prime gap around the interval $I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, \dots, P_{n+1}^2 - 1\}$

Which mean Andrica's conjecture working for all primes $> 37^2$. and so on Legendre conjecture working strongly in this area.

And so for many conjectures related to the field of the large prime gaps it can be detected by the minimum prime density and the large prime gaps.

Discussion and Conclusion.

As shown we can specify the primes density of the long intervals in the depth of the interval

$I(P_n) = \{P_n^2, P_n^2 + 1, P_n^2 + 2, \dots, P_{n+1}^2 - 1\}$ by $(d(P_n)) = \left(\prod_{i=1}^n \left(1 - \frac{1}{P_i}\right)\right)$, which lead us to better understanding to the primes distribution, regarding the weak proof of Legendre's conjecture it is giving the same result of the primes density as the strong proof but the specific of the strong proof is giving more accuracy to the existing of a prime number between m^2 and $(m + 1)^2$, perhaps the sufficient lengths to obtain the primes as the expression of the successive boxes each box has one (particle) prime number with unknown exact location in the length of the box so it does not mean that if the large gap

$g_l = P_{l+1} - P_l = r$ then $P_{l+1} - P_{l-1} = 2r$, no it is mean $P_{l+1} - P_{l-1} = \frac{3}{2}r$, and the exact space you need to find (P_{l-1}, P_l, P_{l+1}) is $\left(\frac{4}{2}r - 1\right)$, and you can find (P_{l-1}, P_l, P_{l+1}) alone in a space within $\left(\frac{5}{2}r - 2\right)$.

So with this understanding, and the same methodology of proofing Andrica's conjecture then

from (14) $\langle \#P_n \rangle_{\min} = L \left(\frac{1}{2} \prod_{i=2}^n \left(1 - \frac{2}{P_i}\right)\right)$ and from (18) $\varepsilon = \left(\frac{3}{3} \frac{5}{5} \frac{7}{9} \dots \dots \frac{P_i}{P_{i+1}-2} \dots \dots \frac{P_{n-1}}{P_n-2}\right)$ then

$\varepsilon = \left(\frac{L}{2P_n \langle \#P_n \rangle_{\min}}\right)$ and for Legendre's interval of length $L_{Leg} = 2m$, and $m \geq P_n$ then we can say

$\varepsilon = \left(\frac{1}{\langle \#P_n \rangle_{\min}}\right)$ so in Legendre's interval if we want to find tow primes then ε should be $\varepsilon = \frac{1}{3}$ and if

we want to find 3 primes then $\varepsilon = \frac{1}{4}$ and so if we want to find K number of primes then ε should be

within $\varepsilon = \frac{1}{K+1}$.

.....

References

https://en.wikipedia.org/wiki/Legendre%27s_conjecture

https://en.wikipedia.org/wiki/Landau%27s_problems

https://en.wikipedia.org/wiki/Prime_number

https://en.wikipedia.org/wiki/Andrica%27s_conjecture

https://en.wikipedia.org/wiki/Prime_gap