

# About twin primes and distribution of primes

Gabriele Di Pietro

August 8, 2010

## Abstract

This paper gives us a demonstration of twin primes conjecture using approximation of function  $\Upsilon$  that we introduce in section 6. Sections 1 – 5 give us introduction to terminology and a clarification on  $\Upsilon$  terms. In particular section 5 is really important because of its Lemma. Section 7 reassume foregoing explanations and it give us two theorems and one corollary; the theorem 7.2 give us exact approximation of twin primes counting function.

## 1 Introduction

A number is a *twin* if it can be found at distance two from the previous or following number of a sequence.

A famous algorithm for making table of primes is the sieve of Eratosthenes: sequentially write down the integers from 2 to a number  $n$  that is the last of table; cross out all number greater then 2 which are divisible by 2; the remaining numbers are all twins but only some of them will survive to the next deletions. Then we can find the smallest remaining number greater then 2: it is 3. So we cross out all numbers greater then 3 which are divisible by 3, the remaining ones are twins in the form  $6k + 1$  and  $6k - 1$ . As previous only some of them will survive to the next deletions. We can go as far as  $\lfloor \sqrt{n} \rfloor$  so the numbers remaining are prime.

A famous conjecture affirms that there are an infinite number of twin primes. We can demonstrate this conjecture by the natural distribution of primes using the formula  $\frac{n}{\ln n}$ , discovered by Gauss. This formula approximates the cumulative distribution of the primes. Moreover, paradoxically, the existence of twin primes depends on the following observation: if the mean distance in distribution of primes increases, the amount of twin primes increases too.

## 2 Inaccurate sieves

First we observe that, if a number  $x$  is prime, it will suffice to control if it is divisible by integers under  $\sqrt{x}$ . So, if we want to realize an exact sieve of

Eratosthenes to the number  $x$  we must cross out multiples of primes numerically inferior to  $\sqrt{x}$ .

We name *would-be twin primes* those numbers in the form  $6k+1$ ,  $6k-1$  included between  $x$  and its root. They are

$$\lfloor \frac{x * 2}{6} \rfloor - \lfloor \frac{\sqrt{x} * 2}{6} \rfloor. \quad (1)$$

using (1) we can miscalculate two numbers in excess at most but this is a negligible error, so we don't consider it. To find couple, it will suffice to divide by two.

Primes greater then 3, under  $\sqrt{x}$ , will be named *eliminators* because they will surely cut out some of the would-be twin primes thanks to the sieve; in fact these numbers, being  $6k+1$  or  $6k-1$ , have got a period of 6. This means that they have to "jump" six times to complete the period; so

$$6k + 1 \equiv 1 \pmod{6} \quad (2)$$

$$6k - 1 \equiv -1 \pmod{6} \quad (3)$$

The following figure clarifies the terminology

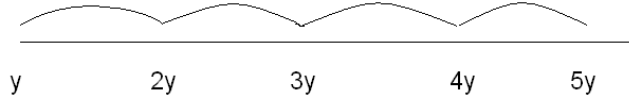


Fig. 0 An eliminator  $y$  and its "jumps"(or multiples).

In the figure,  $y$  is an eliminator and the "jumps" are its multiples. Specifically the eliminator  $6k+1$ , after 4 jumps, cuts out a  $6k-1$  number and, after 6 jumps, it cuts out a  $6k+1$  number; specularly behaves  $6k-1$ .

The very natural conclusion is that every  $y$  eliminator will take  $x/y$  jumps:  $x/6y$  will delete numbers of its own form and  $x/6y$  will cut out numbers of the opposite one. An  $y$  eliminator will delete totally:

$$\lfloor \frac{x}{6y} \rfloor + \lfloor \frac{x}{6y} \rfloor \sim \lfloor \frac{x}{3y} \rfloor \quad (4)$$

would-be twin primes. Among these we can reject the jumps within the eliminator's zone, so we consider only the would-be twin primes over  $\sqrt{x}$ . Finally the formula (4) becomes

$$K = \lfloor \frac{x}{3y} \rfloor - \lfloor \frac{\sqrt{x}}{3y} \rfloor, \quad (5)$$

which doesn't consider possible repetitions, i.e. two or more eliminators could delete the same would-be twin prime, but, by this way, we consider two or more

times the same deleted number; moreover we conjecture that whenever an eliminator makes a jump it deletes a pair of twins, but actually it could delete a number already deprived of its twin (i.e. it is insignificant in the amount of couples).

The final formula is

$$\lfloor \frac{x}{6} \rfloor - \lfloor \frac{\sqrt{x}}{6} \rfloor - \sum_{y \in \pi(\sqrt{x})} K \quad (6)$$

with  $K$  is (5) and  $\pi(\sqrt{x})$  are primes under  $\sqrt{x}$ . Thinking that (6) verifies conjecture (i.e. it is positive) is an utopia, unless we presume that the distribution of eliminators is near  $\sqrt{x}$ . We have only to refine the estimate, by removing repetitions and introducing new remarks. The (6) becomes

$$K \sim \lfloor \frac{1}{3} \frac{x}{y} (1 - \frac{1}{\sqrt{x}}) \rfloor \quad (7)$$

it is not be forgotten that we added the floor functions like simple fractions.

### 3 The $w(n)$ function

Now we can introduce  $w(n)$  function that represent the number of distinct primes factor of a number  $n$ . Hence we can avoid repetitions, in fact all numbers hit by an eliminator are deleted necessarily by another eliminator too (because of factorization); so we can easily halve number of jumps. The question is that a would-be twin primes could be deleted by more than two numbers, for example if all numbers would be hit by three eliminators we should divide total amount of jumps by three. All we need now is a function that shows how many different primes is made a number. Now we consider mean increase of  $w(n)$  (it is be considered that the growth of  $w(n)$  is very irregular). This is the number eliminators jumps should be divided by.

So, using Hardy and Wright's estimate (1979) we can right

$$\frac{\sum_{i=1}^n w(i)}{n} = \ln \ln(n) + B_1 + O(\frac{1}{\ln(n)}) \quad (8)$$

with  $B_1 = 0.261497128$ . Lower bound for the mean function is

$$\ln \ln(n) + B_1 - \frac{1}{\ln(n)} \leq \frac{\sum_{i=1}^n w(i)}{n} \quad (9)$$

which allows the following calculation

$$\frac{\sum_{i=\sqrt{n}}^n w(i)}{n - \sqrt{n}}. \quad (10)$$

Nevertheless, (10) or

$$\frac{\sum_{i=\sqrt{n}}^n w(i) - \frac{n}{\ln n} + \frac{\sqrt{n}}{\ln \sqrt{n}}}{n - \sqrt{n} - \frac{n}{\ln n} + \frac{\sqrt{n}}{\ln \sqrt{n}}} \quad (11)$$

(the correct version of (10)) formulas are not enough to give an answer to our question: in fact we are not considering every number factorization, but we have to consider only factorization of  $6k + 1$  and  $6k - 1$  (i.e. numbers hit by eliminators). Using (9) we have

$$w(z) = z \ln \ln(z) + B_1 z + O\left(\frac{z}{\ln z}\right) - (z-1) \ln \ln(z-1) - B_1(z-1) - O\left(\frac{z-1}{\ln z-1}\right) \quad (12)$$

with  $z = 6k - 1$  or  $z = 6k + 1$ . The

$$L = \frac{1}{\#z} \sum_{z \in [1, n]} w(z) \quad (13)$$

give us the  $w(n)$  mean we were finding, of course different from (11). It has to be considered that we are making linear transformations: so (13) has to be a linear transformation of (8), by this way, we can go on till the least square regression of a some type of solution.

## 4 Distribution of primes

Using prime numbers theorem:

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x / \ln(x)} = 1$$

we obtain this assert: if we have  $x$  numbers (considering that  $\frac{x}{\ln(x)}$  are primes) we can choose as mean distance among close primes precisely  $\ln(x)$ . That means  $\lim_{x \rightarrow \infty} \ln x = \infty$ . This simple consideration is also supported by Rosser & Schoenfeld's estimate (1962). They assert that if  $x \geq 17$  then

$$\frac{x}{\ln x} < \pi(x) < 1.25506 \frac{x}{\ln x} \quad (14)$$

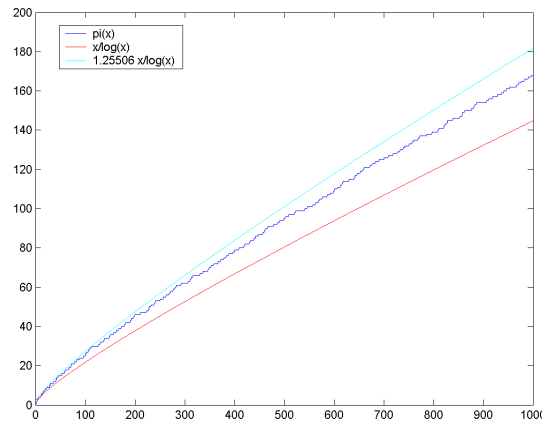


Fig. 1 Estimate function of  $\pi(x)$  with functions of (14).

The conclusion is that the exact mean distance can be calculated with  $\epsilon(x) \frac{x}{\ln(x)}$  with

$$\epsilon(x) \in [1, 1.25506] \tag{15}$$

$$\lim_{x \rightarrow \infty} \epsilon(x) = 1 \tag{16}$$

So, considering  $1.25506 \frac{x}{\ln(x)}$ , primes distance will be lower then  $\epsilon(x) \frac{x}{\ln(x)}$ , eliminators are lower then the exact ones, therefore they will take more "jumps". This is obviously the worst possible hypothesis; instead if we consider  $\epsilon(x) = 1$  we are acting for the best.

$$1.25506 \frac{\mu}{\ln \mu} = 1.25506 \frac{x}{\ln x} + 1 \tag{17}$$

setting  $\mu > 0$ .(17) answers this question: considering  $x$  prime (and remembering that primes counting function increases, according to estimate obtained from choosing  $\epsilon(x) = 1.25506$ ) which one will be next prime number  $\mu$  where primes counting function will be bigger of an unit ? Now, taking in account that primes counting function is bigger of an unit, this means that there is a new prime. Therefore,  $x - \mu$  is the distance between two next primes.

$x - \mu$  represents mean increase, in fact Gauss' estimate does not into account the actual distribution of primes otherwise we had noticed that at a certain point twin primes would be disappeared because  $x - \mu$  is monotonous rising function. The same could be done considering equation with  $\epsilon(x) = 1$ , obtaining the following limits:

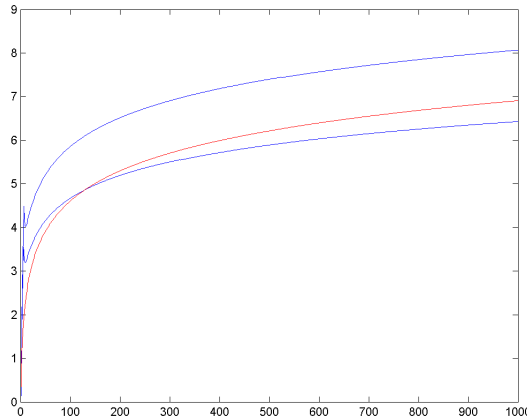


Fig. 2 With:  
 Blue  $\rightarrow$  functions  $\mu - x$  with  $\epsilon(x) = 1.25506$ ,  $\epsilon(x) = 1$  and  
 Red  $\rightarrow \ln(x)$ .

Solving (17) we have a function which follows the state of  $\ln(x)$  in fact we can demonstrate that this function is a logarithm shifted by a constant and rotated on a point. According to our pursuit is always simpler making a least square regression with  $x \geq 13$  on  $1e5$  samples where we can find  $\mu - x = 0.7851 \ln(x) + 1.0047$ . Solving the same function considering  $\epsilon(x) = 1$  we can approximate trough  $\ln(x) + 1.1985$  and we can guarantee, following the expectations, that  $\ln(x)$  remains between the blue functions in fig.2 with  $x > 150$ .

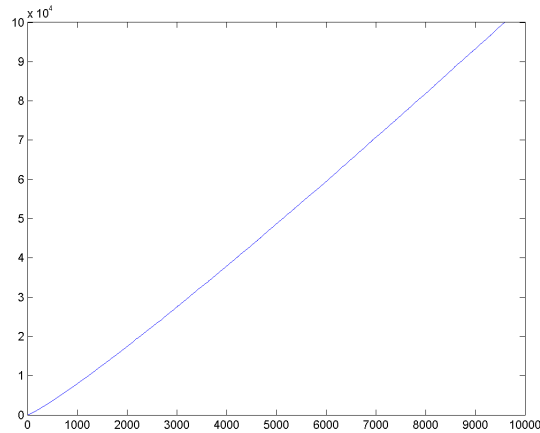


Fig. 3 The function is made by a lot of couples  $(i, p)$  where  $i$  is the  $i$ -th ordinal number end  $p$  is the related prime; for example  $(1, 2), (2, 3), (3, 5), (4, 7) \dots$

Calculating the distance where  $x \geq 3$  is prime, the function  $\mu = x + 0.7851 \ln(x) + 1.0047$  gives a good approximation of fig.3 function. The following graph represents the difference between  $\mu = x + 0.7851 \ln(x) + 1.0047$  and the function shown in fig.3 .

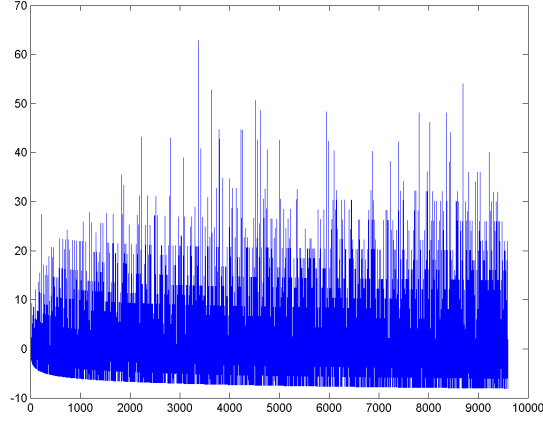


Fig. 4 Difference between  $\mu = x + 0.7851 \ln(x) + 1.0047$  and the function shown in fig.3 choosing  $x$  as prime.

From these results we can get the approximate formula that generates  $y$  eliminators of (5) which are used in (6). The recursive formula becomes

$$\begin{cases} x_0 = 2 \\ x_{n+1} = x_n + 0.7851 \ln(x_n) + 1.0047. \end{cases} \quad (18)$$

The elements generated by algorithm approximate exact eliminators from the below.

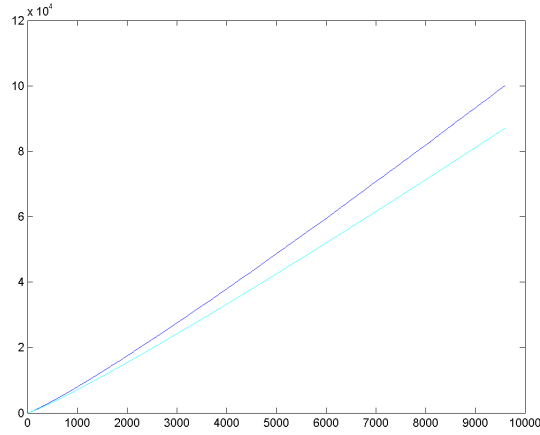


Fig. 4 Function of figure 3 (blue) and function defined by algorithm (18) (cyan).

Actually algorithm approximates better function of fig. 3 if we choose  $x_{n+1} = x_n + \ln(x_n)$ ; as a consequence the distance between two primes is  $\ln(x)$ , recalling what we deduced at the beginning using prime numbers theorem. Assuming that the distance between two primes is a logarithm without a coefficient lower then one is basic to go further with our analysis .

## 5 Completely deleted twins

When the distance between two primes, that according the foregoing estimate is

$$\mu - x = \ln(x), \quad (19)$$

increases as far as it becomes bigger then 8, we know that among two primes is a completely deleted couple of would-be twin primes (remembering that period is 6) i.e. a couple of twin doesn't survive eliminator's jumps. This is a good thing because, thanks to this sacrifice, the others couples of primes will survive proving our claim. As a matter of fact if a couple is completely erased, this means that we needed to use two jumps to eliminate it. So we take into account the jumps that will hit numbers already deprived of their twin, from the whole number of jumps. These jumps, in fact, are irrelevant to count the would-be couples of twin that will be deleted (indeed, we can consider a couple of twin already destroyed, from the moment it misses one of the two members)

Mathematically we can right

$$T = \sum_{p \in [\sqrt{x}, x]} \lfloor \frac{dist(p) - 3}{6} \rfloor \quad (20)$$

where  $dist(p)$  gives distance function between  $p$  (prime) and the following prime number. Now

$$\tilde{T} = \sum_{p=\sqrt{n}}^n (\lfloor \frac{\ln(p) - 3}{6} \rfloor)^+ \quad (21)$$

is the number of jumps needed to completely erase the twin couples between  $\sqrt{n}$  and  $n$ . Every  $x$  in (21) is, in fact, a prime hence if we develop the formula above we obtain

$$\sum_{p=\sqrt{n}}^n \lfloor \frac{\ln(p)}{6} \rfloor - \lfloor \frac{3}{6} * (\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}) \rfloor \sim \frac{1}{6} (\sum_{p=\sqrt{n}}^n \ln(p) - 5.5 (\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}})) \quad (22)$$

clearly we added here floor functions so we have to take away at least a number of elements equal to decimals you may obtain from formula above. For this considering numerator divided by 6 the following decimals can be obtained:  $\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}$ ; adding these numbers we obtain 2.5, thus we have to add to the numerator  $-2.5 * (\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}})$ : this clarifies 5.5. We can approximate logarithmic part with

$$J = \sum_{p=\sqrt{n}}^n \ln(p) \sim (\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}) \ln(n) = \tilde{J} \quad (23)$$



clearly explained according to following graph

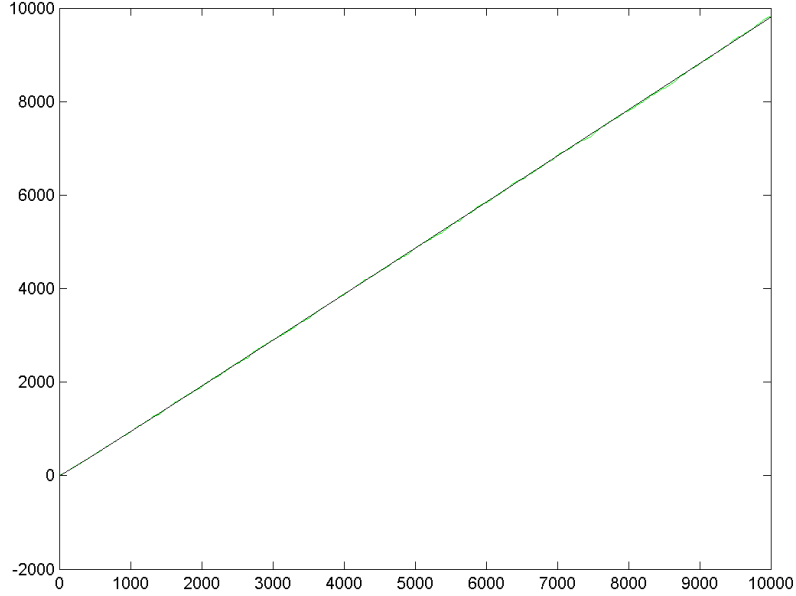


Fig 5. With:  
 Green  $\rightarrow$  J summation function of logarithm.  
 Black  $\rightarrow$  approximation  $\tilde{J}$

To prove this conjecture we need a lemma using properties of logarithms

$$\sum_{p=\sqrt{n}}^n \ln(p) = \ln(n\#) - \ln(\sqrt{n}\#) = \ln\left(\frac{n\#}{\sqrt{n}\#}\right) \quad (24)$$

hence (23) becomes

$$\frac{n\#}{\sqrt{n}\#} \sim n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln\sqrt{n}}} = \frac{n^{\frac{n}{\ln n}}}{n^{\frac{\sqrt{n}}{\ln\sqrt{n}}}} \quad (25)$$

where  $n\#$  is primorial of  $n$ .

**Lemma 5.1** *Assuming mean distance among two primes  $\ln(n)$ , the quantity  $J - \tilde{J}$  contains endless zeros with*

$$J = \sum_{p=\sqrt{n}}^n \ln(p) \sim \left(\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln\sqrt{n}}\right) \ln(n) = \tilde{J} \quad (26)$$

hence  $|J - \tilde{J}|$  is not bounded .

**Proof:**

Supposing that mean distance among two primes is  $\ln(n)$ ,  $n + \log(n)$  represents following prime of  $n$ . Assuming  $n$  is a zero thus we have using (25)

$$\frac{n\#}{\sqrt{n\#}} = n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}} \quad (27)$$

Before reaching  $n + \ln(n)$  it holds following inequality

$$n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}} < (n + k)^{\frac{n+k}{\ln(n+k)} - \frac{\sqrt{n+k}}{\ln \sqrt{n+k}}} \quad (28)$$

where  $k < \ln n$  with  $n > 1$ . If  $k = \ln n$  we obtain a new prime; hence  $\frac{n\#}{\sqrt{n\#}}$  becomes

$$\frac{(n + \ln n)\#}{\sqrt{n + \ln n\#}} = \frac{n\#(n + \ln n)}{\sqrt{n\#}} = n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}} (n + \ln n) \quad (29)$$

by this way, we obtain that denominator will not increase, considering  $\sqrt{n} + \ln \sqrt{n} > \sqrt{(n + \ln n)}$ . It means that while numerator will be multiplied by a new prime number, the denominator will be always the same one.

Thanks to (29) for each  $n > 0$  we can assert:

$$n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}} (n + \ln n) > (n + \ln n)^{\frac{n+\ln n}{\ln(n+\ln n)} - \frac{\sqrt{n+\ln n}}{\ln \sqrt{n+\ln n}}} \quad (30)$$

deducing that when primorial's fraction exceeds exponential's fraction, we find a prime. Furthermore we have to see which  $k$  value we have to choose to obtain

$$n^{\frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}} (n + \ln n) = k^{\frac{k}{\ln(k)} - \frac{\sqrt{k}}{\ln \sqrt{k}}} \quad (31)$$

If fixing  $k > n + \ln n$ , we arrive to verify the foregoing equation, we could running the risk of finding a prime (in mean) before reaching zero. In this case  $\tilde{J}$  is a perfect approximation because  $\tilde{J} \leq J$ . However the value

$$E = \liminf_{n \rightarrow \infty} \frac{p_{n+1} - p_n}{\ln p_n} \quad (32)$$

has been proved to be

- $E < 1$  Erdős
- $E \leq 1/2$ ,  $E \leq 0.46650 \dots$  Bombieri, Davenport
- $E \leq 0.44254 \dots$  Huxley
- $E \leq 0.2486 \dots$  Maier
- $E = 0$  Goldston, Pintz e Yildirim

this indicates that there are endless primes and the distance among these is lesser than logarithm; on the other hand, to obtain mean value, you may find endless primes which have a distance among themselves bigger than logarithm. Continuing, we don't find a prime before  $\ln n$  then there are endless values that verify (31) guaranteeing function endless oscillations of  $J - \tilde{J}$ , remembering that  $J - \tilde{J}$  contains endless zeros .

The following graph shows "difference function" oscillation, on logarithmic scale, with  $n < 404$

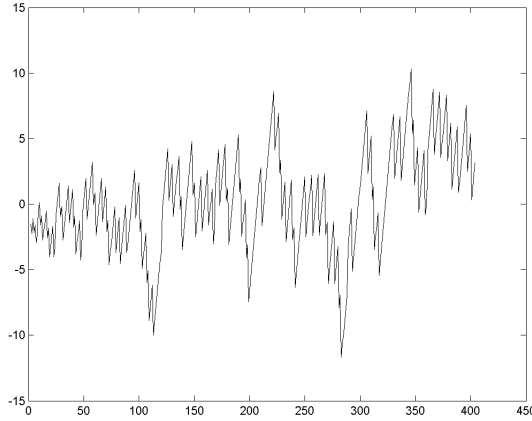


Fig. 6 This is the "difference function" between  $\ln \frac{\#n}{\#\sqrt{n}}$  and  $\ln n \frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}$

On the contrary, about bounds we deduce that:

$$n \frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}} (n + \ln n) - (n + \ln n) \frac{n + \ln n}{\ln(n + \ln n)} - \frac{\sqrt{n + \ln n}}{\ln \sqrt{n + \ln n}} \quad (33)$$

and

$$(n + \ln n - 1) \frac{n + \ln n - 1}{\ln(n + \ln n - 1)} - \frac{\sqrt{n + \ln n - 1}}{\ln \sqrt{n + \ln n - 1}} - n \frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}} \quad (34)$$

represent  $|J - \tilde{J}|$  max-values.

In fact the first max-value is  $J - \tilde{J}$ , so we have to multiply  $\frac{n\#}{\sqrt{n}\#}$  numerator by  $n + \ln n$  (a new prime); the second max-value represents  $\tilde{J} - J$  distance where we find no prime and we assume  $\frac{\#n}{\#\sqrt{n}} = n \frac{n}{\ln(n)} - \frac{\sqrt{n}}{\ln \sqrt{n}}$ , i.e. this function is a zero of  $J - \tilde{J}$ .

Studying the functions (33) and (34) we have noticed that they are increasing functions greater than zero, with  $n > 1$ , and they are not bounded from above. These considerations prove our thesis .  $\diamond$

Assuming that mean distance between two prime numbers is  $\ln(x)$  is basic to demonstrate our theory.

In fact, as we show at the end of theorem 7.1, our theory is proved if we set

logarithmic constant bigger than one. Furthermore presence of endless zeros does not allow us to correct the mistake given by approximation of (26) in fact studying (31) for each  $n$  value, we observe that  $k$  increases on an average slower than logarithm. By this way  $\tilde{J}$  approximates in excess twin prime amount because mistakes of  $J - \tilde{J}$  negative part will be greater than those of the positive one. To avoid this fact you may choose a different approximation. A good solution is

$$\tilde{J} = \left( \frac{n}{\ln(n) + \alpha} - \frac{\sqrt{n}}{\ln \sqrt{n}} \right) \ln(n) \quad (35)$$

with  $\alpha > 0.4$ .

## 6 Alternative approach to use of $w(n)$ function and $L$ term

Considering would-be twin primes between  $x$  and its root, we find that these numbers are  $\lfloor \frac{x*2}{6} \rfloor - \lfloor \frac{\sqrt{x}*2}{6} \rfloor$  according to (1). If we cut out from this formula primes up to  $\lfloor \sqrt{x}, x \rfloor$  included we get the primes erased by the sieve without repetitions. Hence we obtain effortlessly the (13) and third member of (6). Therefore

$$\frac{1}{L} \sum_{y \in \pi(\sqrt{x})} K = \lfloor \frac{x}{3} \rfloor - \lfloor \frac{\sqrt{x}}{3} \rfloor - \pi(x) + \pi\sqrt{x} \quad (36)$$

without computing  $w(n)$  for each would-be prime numbers. Therefore using all these approximations we can right twin primes computing function :

$$\Upsilon(x) = \lfloor \frac{x}{6} \rfloor - \lfloor \frac{\sqrt{x}}{6} \rfloor - \frac{1}{L} \sum_{y \in \pi(\sqrt{x})} K + T \quad (37)$$

with

$$\frac{1}{L} \sum_{y \in \pi(\sqrt{x})} K = \lfloor \frac{x}{3} \rfloor - \lfloor \frac{\sqrt{x}}{3} \rfloor - \pi(x) + \pi\sqrt{x} \quad (38)$$

$$T = \sum_{p \in [\sqrt{x}, x]} \lfloor \frac{dist(p) - 3}{6} \rfloor \quad (39)$$

It represents the amount of twin primes included between  $\sqrt{x}$  and  $x$ ; if this amount is bigger than zero, our conjecture (affirming that *there are endless twin primes couples*) is proved.

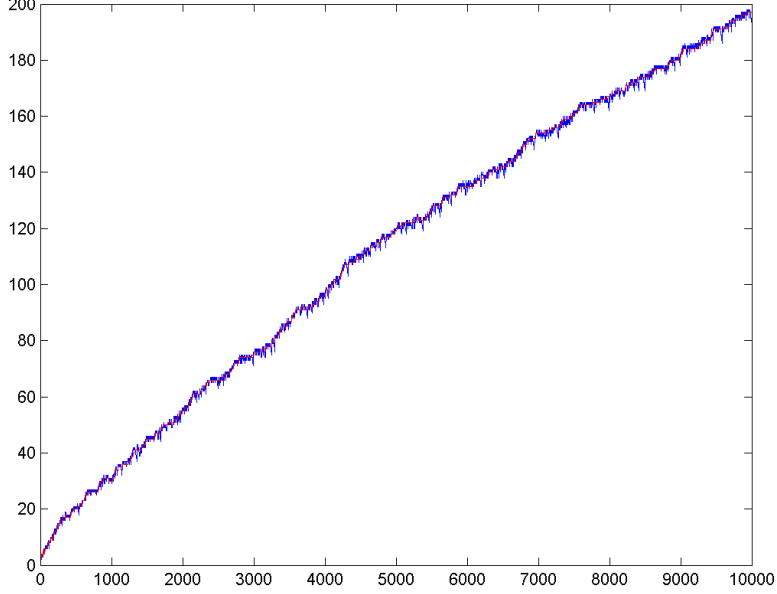


Fig. 7 With:  
 Red  $\rightarrow$  exact twin primes function into  $[\sqrt{x}, x]$  and  
 Blue  $\rightarrow \Upsilon(x)$ .

In addition using prime number theorem:

$$\pi(x) + \pi\sqrt{x} \sim \left( \frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}} \right). \quad (40)$$

## 7 Approximation of twin primes' counting function

Hence considering the previous conclusions we reach the following approximation

$$\tilde{\Upsilon}(x) = \lfloor \frac{x}{6} \rfloor - \lfloor \frac{\sqrt{x}}{6} \rfloor - \frac{1}{\tilde{L}} \sum_{\tilde{y} \in \pi(\sqrt{x})} \tilde{K} + \tilde{T} \quad (41)$$

with

$$\frac{1}{\tilde{L}} \sum_{\tilde{y} \in \pi(\sqrt{x})} \tilde{K} = \lfloor \frac{x}{3} \rfloor - \lfloor \frac{\sqrt{x}}{3} \rfloor - \left( \frac{x}{\ln(x)} + \frac{\sqrt{x}}{\ln\sqrt{x}} \right) \quad (42)$$

$$\tilde{T} = \frac{(\ln(x) - 5.5)}{6} \left( \frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}} \right) \quad (43)$$

**Theorem 7.1** *The function  $\tilde{\Upsilon}(x)$ , defined by (40)-(42), is bigger then zero for each  $x > 141.83$ . This function approximates  $\Upsilon(x)$ , twin primes counting function, defined by (36)-(38).*

**Proof:**

$$\begin{aligned}
\frac{x}{6} - \frac{\sqrt{x}}{6} - \frac{x}{3} + \frac{\sqrt{x}}{3} + \left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) + \frac{(\ln(x) - 5.5)}{6} \left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) &\geq 0 \\
\frac{\sqrt{x}}{6} - \frac{x}{6} + \left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) \left(1 + \frac{(\ln(x) - 5.5)}{6}\right) &\geq 0 \\
\sqrt{x} - x + \left(\frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln\sqrt{x}}\right) (6 + \ln(x) - 5.5) &\geq 0 \\
\frac{\ln(x) \ln\sqrt{x}(\sqrt{x} - x) + (x \ln\sqrt{x} - \sqrt{x} \ln(x))(0.5 + \ln(x))}{\ln(x) \ln\sqrt{x}} &\geq 0 \\
\frac{\sqrt{x} \ln(x) \ln\sqrt{x} + 0.5(x \ln\sqrt{x} - \sqrt{x} \ln(x)) - \ln^2(x)\sqrt{x}}{\ln(x) \ln\sqrt{x}} &\geq 0 \\
\frac{\sqrt{x} \ln\sqrt{x}(\ln(x) + 0.5\sqrt{x}) - \sqrt{x} \ln(x)(0.5 + \ln(x))}{\ln(x) \ln\sqrt{x}} &\geq 0
\end{aligned}$$

using logarithms properties

$$\begin{aligned}
\frac{\sqrt{x} \left(\frac{\ln(x) + 0.5\sqrt{x}}{2} - 0.5 - \ln(x)\right)}{\ln\sqrt{x}} &\geq 0 \\
\frac{\sqrt{x}(0.5\sqrt{x} - 1 - \ln(x))}{\ln x} &\geq 0
\end{aligned}$$

So the denominator is bigger then zero for each  $x > 1$ , and the numerator is a logarithmic equation having as solutions 0.52 and 141.83. This equation is positive for external values of the following range:  $[0.52, 141.83]$ . Hence  $\tilde{\Upsilon}(x)$  is positive for each  $x \in [0.52, 1] \cup [141.83, \infty)$  proving our theorem.

◇

For each  $x < 20$  you may count the twin primes by hand; the theorem assures that our conjecture is proved, because for each new  $x$ , the twins will increase more and more, assuming that the function is always positive. This assumption guarantees that twin primes are endless.

The very basic computing hypothesis to demonstrate the theorem is the assumption that the distance among primes is, in mean, bigger or equal then  $\ln(x)$ : the same statement is valid for the estimate of  $J$  function in (27). If we assumed that the distance among primes were  $k \ln(x)$  with  $0 < k < 1$  from Gauss' approximation our theorem had not been proved. For instance, with  $k = 0.7851$  we would have:

$$\begin{aligned}
\frac{\sqrt{x}(\sqrt{x}0.5 - 1 - 0.5702 \ln(x) - 0.2149\sqrt{x} \ln(x))}{\ln(x)} &\geq 0 \\
\frac{\sqrt{x}(\sqrt{x}(0.5 - 0.2149 \ln(x)) - 1 - 0.5702 \ln(x))}{\ln(x)} &\geq 0
\end{aligned}$$

So, positiveness of numerator depends from the following expression

$$\sqrt{x}(0.5 - 0.2149 \ln(x)) \quad (44)$$

This quantity is positive at the very beginning, but it will become soon negative because of  $\ln(x)$ , by this way, the numerator will become negative too, considering that we will find a sum of all negative terms; and this fact will reject our conjecture.

**Teorema 7.2** *Using hypothesis of theorem 7.1 and considering*

$$\tilde{T} = \frac{(\ln(x) - 5.5)}{6} \left( \frac{x}{\ln(x) + \alpha} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) \quad (45)$$

for each  $0.4 \leq \alpha \leq 0.5$  the approximating function  $\tilde{Y}(x)$  is positive for each  $x > 90$ .

Following the step of foregoing demonstration we have

$$\begin{aligned} \frac{x}{6} - \frac{\sqrt{x}}{6} - \frac{x}{3} + \frac{\sqrt{x}}{3} + \left( \frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) + \frac{(\ln(x) - 5.5)}{6} \left( \frac{x}{\ln(x) + \alpha} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) &\geq 0 \\ \sqrt{x} - x + 6 \left( \frac{x}{\ln(x)} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) + (\ln(x) - 5.5) \left( \frac{x}{\ln(x) + \alpha} - \frac{\sqrt{x}}{\ln \sqrt{x}} \right) &\geq 0 \\ \sqrt{x} - x + 6 \frac{x}{\ln(x)} - \frac{12\sqrt{x}}{\ln x} + \frac{x \ln x}{\ln(x) + \alpha} - 2\sqrt{x} - \frac{5.5x}{\ln(x) + \alpha} + \frac{11\sqrt{x}}{\ln x} &\geq 0 \\ \sqrt{x} \left( 1 - \sqrt{x} + \frac{6\sqrt{x}}{\ln x} - \frac{12}{\ln x} + \frac{\sqrt{x} \ln x}{\ln(x) + \alpha} - 2 - \frac{5.5\sqrt{x}}{\ln(x) + \alpha} + \frac{11}{\ln x} \right) &\geq 0 \\ \sqrt{x} \left( -\sqrt{x} + \frac{6\sqrt{x}}{\ln x} - \frac{1}{\ln x} + \frac{\sqrt{x} \ln x}{\ln(x) + \alpha} - 1 - \frac{5.5\sqrt{x}}{\ln(x) + \alpha} \right) &\geq 0 \\ \frac{\sqrt{x}(-\sqrt{x} \ln x(y) + 6\sqrt{x}(y) - (y) + \ln^2 x \sqrt{x} - \ln x(y) - 5.5\sqrt{x} \ln x)}{\ln x(y)} &\geq 0 \end{aligned}$$

with  $y = \ln x + \alpha$  Regarding this calculus we can compute the numerator as follows

$$\begin{aligned} -(\ln x \alpha \sqrt{x} + \ln x \alpha + \ln^2 x + 5.5\sqrt{x} \ln x + \ln x + \alpha) + (6\sqrt{x} \ln x + 6\sqrt{x} \alpha) &\geq 0 \\ -(\ln^2 x + \alpha \sqrt{x} \ln x + \ln x(\alpha + 1) + \alpha) + (0.5\sqrt{x} \ln x + 6\sqrt{x} \alpha) &\geq 0 \end{aligned}$$

The very basic terms are  $-\alpha \sqrt{x} \ln x$  and  $0.5\sqrt{x} \ln x$ , these ones prevail over all others terms; considering that  $6\sqrt{x} \alpha$  prevails over  $\ln^2 x + \ln x(\alpha + 1) + \alpha$ , but  $6\sqrt{x} \alpha$  is passed by  $(0.5 - \alpha)\sqrt{x} \ln x$ . The positiveness of the prevailing term is given by  $\alpha$  that consequently has to be lesser or equal to 0.5. Fixing when  $0.4 \leq \alpha \leq 0.5$   $\tilde{Y}(x)$  is positive for each  $x > 90$ .  $\diamond$

**Corollary 7.3** *Given hypothesis that mean distance between two primes is  $\ln(x)$  there are endless compensation terms like  $2 \ln(x) - 2$ .*

**Proof:** Using theorem 7.1 we can affirm that there are endless twin primes. So considering  $x$  numbers mean distance between two primes is  $\ln(x)$ . Thus

$$\ln(x) = \frac{2+y}{2} \rightarrow y = 2\ln(x) - 2 \quad (46)$$

where  $y$  compensates the mean and it can be a linear combination of the distance among primes.  $\diamond$

## References

- [1] Rosser, J. B. and Schoenfeld, L. *Approximate Formulas for Some Functions of Prime Numbers*. Illinois J. Math. 6, 64-97, 1962
- [2] Apostol, T.M., *Introduction to analytic number theory*, Springer (1975)
- [3] Erdős, P., *The difference of consecutive primes*, Duke Math. J.6:438-441 (1940)
- [4] Naldi, G., Pareschi, L., *Matlab, concetti e progetti*, Apogeo (2008)
- [5] Goldston, D. A., Motohahi, Y. J., Pintz, and Yıldırım, C. Y.. *Small gaps between primes exist*. arXiv website, <http://xxx.sissa.it/pdf/math.NT/0505300> (2005)
- [6] Greaves, G. *Sieves in Number Theory*. Springer, Berlin, (2001)
- [7] Hardy, G.H., Wright, E.M. *An introduction to the theory of numbers* Oxford Science Publications, Oxford, fifth edition (1979)
- [8] Pintz, J., *Very large gaps between consecutive primes* J. Number Theory, 63:286301, (1997)
- [9] Derbyshire, J. *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics* New York: Penguin, (2004)
- [10] Gauss, C. F. Werke, Band 10, Teil I. p. 10, (1863)

**Author:** Di Pietro Gabriele

**Address:** Giulianova (TE), Via Annunziata 42

**Telephone:** 3471160919

**E-mail:** petreius@libero.it