_____

The Twin Primes Seen from a Different Perspective

A. Dinculescu[1*]

[1]*Formerly with: "Esperion" at Western Michigan University, Kalamazoo, MI 49008, USA.*

Research Article

_____

Abstract

The paper presents a framework for the construction of an elementary proof of the infinitude of twin primes. It starts from the fact that all positive integers can be divided into numbers that can lead directly to a pair of twin primes (called twin ranks) and numbers (called non-ranks) that do not have this property. While the twin ranks cannot be directly calculated, the non-ranks can be easily calculated with a simple equation based on ordinary primes. They present a series of properties that once rigorously proven make the finiteness of twin prime an impossibility. Foremost among these properties is the fact that they can be arranged in an infinite number of sets called groups and super-groups. These sets have a built-in symmetry, a precise interval length and a well-defined number of terms. Another important property is that the depletion of twin primes via non-ranks goes in steps from one "basic" interval to another. As one goes higher up in the number series, these intervals grow larger and larger while the prime numbers required for their depletion become more and more sparse.

1 Introduction

Two prime numbers P_j and P_{j+1} are called twin primes if $P_{j+1} - P_j = 2$. Because they have a tendency to thin out compared with the usual primes as one goes higher up in the number series, for many years the set of twin primes was considered to be most likely finite. Although nowadays there is a strong consensus that there are infinitely many twin primes, a formal proof of this conjecture (called the Twin Prime Conjecture [1]) was not found yet. The main difficulty is the fact that probabilistic events for consecutive primes are not truly independent [2].

In this paper we look at the problem from a different perspective and show that once the interdependence of the twin primes with the other primes is understood, it is hard to see how their number can be finite. This is because, as one goes higher up in the number series, the prime numbers which play a key role in their depletion become more and more sparse [3], while the intervals that have to be covered by a single prime grow larger and larger.

_____

*Corresponding author: adinculescu@cox.net;

2 Basic Properties

We begin by showing that with regard to the twin primes all positive integers can be divided into two and only two categories: twin ranks and non-ranks. These two concepts were introduced and discussed in a previous paper [4], but in order to facilitate the exposition we recall here some of the characteristics that are essential in understanding their role in the formation of the twin primes.

One can associate to each pair of twin primes P and $P+2$ a "**twin index**" $K = P+1$ representing the number between them. Since all twin primes with the exception of 3 and 5 are of the form $P = 6n \pm 1$ (where n is a positive integer) all twin indices except 4 are of the form $K = 6n$. For reasons that will become apparent shortly, one can define a "**twin rank**" as a number of the form $k_* = K / 6$ and work with twin ranks instead of twin indices. All twin ranks lead to a twin index (and hence to a pair of twin primes) by a single algebraic operation: multiplication by 6. Since $K^2 - 1 = P(P+2)$ is divisible only by P and $P + 2$, it follows that that if a number $(6n)^2 - 1 = (6n+1)(6n-1)$ is not divisible by any prime $P \le \sqrt{6n+1}$, then $n = k_*$ is a twin rank and $6k_* \pm 1$ are twin primes. Conversely, all numbers k that satisfy one of the two variants of the following equation

$$k = nP \pm [P / 6] \tag{1}$$

where $[x]$ means the nearest integer to x, cannot lead to a pair of twin primes by a single algebraic operation. They were called "**non-ranks**". As shown, there are no other numbers in the set of positive integers besides twin ranks and non-ranks [4]. With the above formula one can find all non-ranks smaller than $M_{j+1} = \left(P_{j+1}^2 - 1 \right) / 6$ by using all primes $5 \le P \le P_j$. By subtracting these non-ranks from the set of positive integers $n < \left(P_{j+1}^2 - 1 \right) / 6$ one obtains all twin ranks $k_* < \left(P_{j+1}^2 - 1 \right) / 6$ and, hence, all twin primes with indices $K < P_{j+1}^2 - 1$. (Note: Because the reminder from the division of a number a by a number b can be equal to 0.5, the expression "the nearest integer to a / b" is usually avoided in number theory. Nonetheless, we are using it here because the reminder from the division of a prime number $P \ge 5$ by 6 is either 1 / 6 or 5 / 6).

It is important to realize that while many of the non-ranks k_j determined with (1) using a prime P_j can be found with the help of primes smaller than P_j, none of them can be found using primes larger than P_j. This property is very important because it ensures that once a number was shown to be a twin rank by using in (1) all primes up to a certain prime P_j, that number is not going to be "**covered**" (i.e. shown to be a non-rank) by primes larger than P_j. In other words, while the covering process started by a prime at the points $P \pm [P / 6]$ goes forward to infinity, it does not go backward.

3 The Case for the Infinity of Twin Primes

Let us examine the set of natural numbers and see how the twin ranks arise as a consequence of the fact that there are not enough primes to cover all positive integers. The first characteristic we notice is that there are two non-ranks associated to a prime P in each interval of length $\Delta n = P$ starting with $P \pm [P/6]$. They are symmetrically distributed at equal distances $[P/6]$ from all numbers that are a multiple of P. One can associate, therefore, to each non-rank a "**parent prime**" defined as the smallest prime required by (1) to find it. **Example**: Here are the first 10 non-ranks of parent prime $P_{17} = 59$: 777, 1013, 1052, 1288, 1347, 1603, 1642, 1760, 1937, 1957. All of them are of the form $k_{17} = 59n \pm [59/6]$.

Starting with 4, the smallest non-rank, all non-ranks k_j of the same parent prime P_j form an infinite number of consecutive groups of length $L_j = \prod_{i=3}^{j} P_i$, each of them containing G_j terms. The totality of non-ranks of parent primes $5 \leq P_i \leq P_j$ form an infinite number of consecutive super-groups of equal interval lengths L_j, each of them containing S_j terms. (Note: Although L_j, G_j and S_j are whole numbers, for simplicity in the following we will also use them to designate the corresponding intervals, groups or super-groups. The same is true for R_j, a set of numbers to be defined shortly). Because L_j is divisible by all primes $5 \leq P_i \leq P_j$, each super-group S_j has an integer number of groups G_i with $i \leq j$, and the gaps between its terms will come out in the same order regardless from which side of the interval one starts to count them. This makes the distribution of gaps symmetric with respect to a central gap between the numbers $L_{c1} = (L_j - 3)/2$ and $L_{c2} = (L_j + 3)/2$. It also has as a consequence the fact that the sum of two non-ranks situated at equal distances from the central gap equals L_j. Obviously, the size of the gaps and their distribution will be the same in all the other super-groups obtained by adding to each term in the first super-group the number L_j multiplied by an integer. As shown in [4] one has

$$G_j = 2\prod_{i=3}^{j-1}(P_i - 2) \tag{2}$$

$$S_j = L_j \left(1 - \prod_{i=3}^{j} \frac{P_i - 2}{P_i}\right) \tag{3}$$

As mentioned, the number of terms in a super-group is the sum of the number of terms in the constituent groups. However, this does not mean that S_j is given by $\sum_{i=3}^{j} G_i$ with G_j as in (2). In order to arrive at (3), one has to take into account the fact that in each super-group of order j there are nested $G_{ij} = (L_j / L_i) G_i$ groups of order $3 \leq i \leq j$. (Note that all these numbers are exact integers).

Although per the fundamental theorem of arithmetic all natural numbers are multiples of a prime *P,* this does not guaranty that one can associate a prime to every number according to (1). The number 10 for example is a multiple of 2 and 5, but there is no number *n* that allows it to be written as either $nP + [P/6]$ or $nP - [P/6]$. Consequently, 10 is a twin rank. It gives the twin index 60 when multiplied by 6.

As we go further up in the natural number series, we notice that the covering process is not monotonous. It goes in steps from one "**basic**" interval $\Delta M_j = (P_{j+1}^2 - P_j^2)/6$ to the next basic interval $\Delta M_{j+1} = (P_{j+2}^2 - P_{j+1}^2)/6$. Any time one covers a basic interval ΔM_j and goes to the next interval ΔM_{j+1}, one needs a larger prime to cover the numbers in that interval left uncovered by previous primes. And here is an important aspect of the problem: Not only a prime P cannot cover more than about a fraction $2/P$ of the remnants in a basic interval, but very often there is no number n which together with P can satisfy (1) for that interval. In this case all the remnants are twin ranks.

As an example, let us start from the number $M_{17} = (P_{17}^2 - 1)/6$ (where $P_{17} = 59$ is the 17th prime) and let us apply eq. (1) with $5 \leq P \leq 59$ to the basic interval situated between $M_{17} = (59^2 - 1)/6 = 580$ and $M_{18} = (61^2 - 1)/6 = 620$. We obtain the non-ranks: 580, 581, 582, 583, 584, 585, 586, 587, 589, 591, 592, 594, 595, 596, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 613, 614, 615, 616, 617, 618, 619. After subtracting these non-ranks from the set of positive integers in the interval, we obtain the remnants: 588, 590, 593, 597, 612. As expected (because we used all required parent primes) all these remnants are twin ranks. They correspond to the following pairs of twin primes: (3527, 3529); (3539, 3541); (3557, 3559); (3581, 3583) and (3671, 3673). Now, if we apply the procedure to the next basic interval situated between $M_{18} = (61^2 - 1)/6 = 620$ and $M_{19} = (67^2 - 1)/6 = 748$ using the same primes as before, we obtain all non-ranks in the interval (starting with 620) except one: 742. In order to cover this number we have to use the prime next to P_{17}, i.e. $P_{18} = 61$. Indeed, $742 = 12 \times 61 + 10$ in accordance with eq. (1). In the next interval situated between $M_{19} = (67^2 - 1)/6 = 748$ and $M_{20} = (71^2 - 1)/6 = 840$ the new prime $P_{19} = 67$ manages to cover two more remnants $793 = 12 \times 67 - 11$ and $815 = 12 \times 67 + 11$, but in the succeeding three intervals the primes $P_{20} = 71$, $P_{21} = 73$, and $P_{22} = 79$ are not able to cover any remnants

because there are no numbers n that together with these primes can satisfy eq. (1) for the corresponding intervals. Consequently, all remnants in these intervals are twin ranks. What one sees is the following: On one hand, as one goes further up in the number series, one has at one's disposal more and more primes that can be used to cover the remnants in the incoming intervals. On the other hand, the primes needed for the covering process become more and more sparse and the basic intervals which have to be covered by a single prime grow larger and larger. The main questions are:

- *If P_z is the zth prime, is it possible to have all integers larger than $M_{z+1} = \left(P_{z+1}^2 - 1 \right)/6$ in a super-group $L_z = \prod_{i=3}^{z} P_i$ covered by primes smaller than P_{z+1} ?* (Recall that the covering process started by a prime at the points $P \pm \left[P/6 \right]$ goes forward to infinity).

- *If the answer to the first question is "no", are the remaining primes $P_{z+1} \le P \le \sqrt{6L_z + 1}$ able to cover all the subsequent remnants in the super-group?*

In order to answer the first question we notice that the interval L_z contains a number S_z of non-ranks of parent primes $5 \le P_i \le P_z$ given by the same equation (3) regardless of the value of z. Because $L_z > S_z$ there will always be in L_z a number R_z of remnants not included in S_z. One has

$$R_z = L_z - S_z = \prod_{i=3}^{z} \left(P_i - 2 \right) \tag{4}$$

The uniform distribution of terms in L_z allows the symmetry in S_z with respect to the central gap to be preserved after the subtraction of its terms from L_z. Consequently, the gaps between consecutive terms in R_z are symmetrically distributed with respect to the central gap. (This is valid for any remnant R_j). All remnants smaller than $M_{z+1} = \left(P_{z+1}^2 - 1 \right)/6$ are twin ranks because the non-ranks smaller than M_{z+1} were covered by the primes $5 \le P_i \le P_z$ and are not in R_z.

Let us assume there are no twin ranks after M_{z+1} because they have been covered by primes smaller than P_{z+1}. This means that all numbers in R_z after M_{z+1} are non-ranks. These non-ranks must be of parent primes larger than P_z because if they were smaller they would have been included in S_z and no numbers would have been left for the interval between M_{z+1} and L_z. If this would be the case, the number of remnants would be much smaller than the value given in (4) and there would be no symmetry. Since this negates the basic properties of non-ranks, we conclude that:

Given an arbitrary prime P_{z+1} and a super-group of length $L_z = \prod_{i=3}^{z} P_i$, the primes smaller than

P_{z+1} cannot cover all numbers in the super-group larger than $M_{z+1} = \left(P_{z+1}^2 - 1 \right) / 6$.

In order to answer the second question we recall that the gaps between consecutive terms in any remnant R_j are symmetrically distributed on each side of the central gap. This means the fraction of remnants in two intervals of equal length situated at equal distances from the central gap are equal, and the fraction of remnants at the beginning does not differ too much from the fraction at the end, with both of them not significantly different from the average value

$$x_j = \frac{R_j}{L_j} = \prod_{i=3}^{j} \frac{P_i - 2}{P_i} \tag{5}$$

Example: Let $j = 7$. One has $P_j = 17$, $L_j = 85085$, and $L_{c1} = \left(L_j - 3 \right) / 2 = 42541$. If one divides the interval between 541 and 42541 in 42 intervals of equal length and measures the fraction of remnants in each of them, one obtains the following values: 0.265, 0.261, 0.263, 0.257, 0.262, 0,261 ,0.262, 0.259, 0.257, 0.264, 0.259, 0.264, 0.257, 0.261, 0.265, 0.257, 0.261, 0.266, 0.260, 0.259, 0.261, 0.264, 0.259, 0.257, 0.261, 0.259, 0.263, 0.263, 0.255, 0.262, 0.259, 0.267, 0.260, 0.259, 0.266, 0.259, 0.260, 0.266, 0.254, 0.263, 0.257, 0.264. The mean value for the whole interval L_j is $x_j = 0.2618$. Due to the central symmetry, one obtains the same numbers in reverse order for the interval between $L_{c2} = \left(L_j + 3 \right) / 2 = 42544$ and $L_j - 541 = 84544$. Therefore, one can approximate the number Λ_j of remnants in a basic interval ΔM_{j+1} inside L_j as a fraction x_j of its length. One has

$$\Lambda_j \cong x_j \Delta M_{j+1} = \Delta M_{j+1} \prod_{i=3}^{j} \frac{P_i - 2}{P_i} \tag{6}$$

It is important to realize that what matters here is not the exact number of uncovered terms in the interval, but the fraction of them that can be covered by a single prime. In this case the prime is P_{j+1} and the interval is $\Delta M_{j+1} = \left(P_{j+2}^2 - P_{j+1}^2 \right) / 6$. The approximate number of terms covered by P_{j+1} is then

$$N_{j+1} \cong 2 \frac{x_j \Delta M_{j+1}}{P_{j+1}} = \frac{P_{j+2}^2 - P_{j+1}^2}{3} \prod_{i=3}^{j} \frac{P_i - 2}{P_i} \tag{7}$$

It follows that the number of terms remained uncovered, all of them twin ranks, is on the order of

$$T_j \cong \Lambda_j \left(1 - 2/P_{j+1}\right) \tag{8}$$

Theorem: The number of remnants in a basic interval is larger than the gap between the primes that determine the length of the interval.

Proof: For any prime P_i one has $P_i - 2 \geq P_{i-1}$. This allows one to write (6) as

$$\Lambda_j \geq \frac{3}{P_j} \Delta M_j \geq \frac{\left(P_{j+2} + P_{j+1}\right)\left(P_{j+2} - P_{j+1}\right)}{2P_j} \tag{9}$$

With $P_{j+2} + P_{j+1} > 2P_j$ one has $\Lambda_j > P_{j+2} - P_{j+1}$. This completes the proof.

Based on the above properties, we conclude that:

Given an arbitrary prime P_{z+1} and a super-group of length $L_z = \prod_{i=3}^{z} P_i$, the primes larger than

P_{z+1} cannot cover all remnants in the super-group larger than $M_{z+1} = \left(P_{z+1}^2 - 1\right)/6$.

4 Concluding Remarks

Many years ago Euclid gave an elementary but elegant proof that there are an infinite number of primes [5]. Now, after more than 2000 years, mathematicians, while still struggling to find a proof for the infinity of twin primes using complex analysis, do not even mention the possibility of an elementary proof. Based on the above analysis, we present here a possible framework for the construction of such a proof. The line of reasoning is as follows:

a) There is a one to one correspondence between twin primes and twin ranks;
b) Any positive integer is either a twin rank or a non-rank;
c) While the twin ranks cannot be directly calculated, the non-ranks can be easily calculated with a simple equation based on ordinary primes;
d) All positive integers that fail to satisfy that equation after using all corresponding primes and numerical coefficients are twin ranks;
e) The non-ranks can be arranged in an infinite number of groups and super-groups with a built-in symmetry a precise interval length and a well-defined number of terms;
f) Because of the built-in symmetry, if there were no twin ranks after a certain number, the number of remnants in the interval occupied by a super-group containing that number would be much smaller, and there would be no symmetry;
g) In a super-group the primes smaller than a given prime cannot cover all terms larger than a certain number on the order of the square of that prime divided by 6;

h) The depletion of twin primes in an interval is directly dependent of the ability of the ordinary primes to "cover" all positive integers in that interval;

i) The covering process goes in steps from one "basic" interval to another, with a prime P unable to cover more than about a fraction $2/P$ of the numbers in the interval;

j) If the primes smaller than a given prime cannot cover all terms in a super-group larger than a certain number, the uncovered terms cannot be entirely covered by larger primes;

k) The propositions (g) and (j) imply that there will always be twin ranks in a super-group because, regardless of its size, it is impossible to cover all the constituent terms by the available primes;

l) Since there are infinitely many super-groups, it follows that there are infinitely many twin ranks and hence infinitely many twin primes.

Competing Interests

Author has declared that no competing interests exist.

References

[1] Guy RK. Unsolved problems in number theory. New York: Springer-Verlag; 2004.

[2] Hardy GH, Wright EM. An introduction to the theory of numbers. Oxford; 1979.

[3] Hardy GH, Littlewood JE. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. Acta Mathematica 1916;41:119-196.

[4] Dinculescu A. On some infinite series related to the twin primes. The Open Mathematics Journal. 2012;5:8-14.

[5] [5] Tietze H. Famous problems of mathematics: Solved and unsolved mathematics problems from antiquity to modern times. New York: Graylock Press; 1965.

_____