Solving Goldbach's Conjecture using Gaussian Arithmetic and A Probabilistic Model

Emilio A. Diarte-Carot

ABSTRACT

This paper proves that Goldbach's conjecture is true. The proof uses Gaussian modular arithmetic to calculate the number of pairs of odd numbers, K_T, whose sum is a given even number, n, as well as, the number, K_E, of those that can potentially contain prime numbers.

Next, a probabilistic model with a binomial probability distribution is defined, which will be applied to K_E to calculate a function f(x) for the expected value, E(X), where X is the number of pairs formed by two prime numbers.

Finally, the analysis of this function, f(x), will allow us to prove that the conjecture is tru.

Keywords and phrases: goldbach conjecture. gaussian arithmetic. prime numbers. binomial distribution.

Classification: LCC Code: QA157.7, QA76.9.M35, QA273.6

Language: English

Great Britain Journals Press

LJP Copyright ID: 925614
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural & Formal

Volume 25 | Issue 12 | Compilation 1.0

Solving Goldbach's Conjecture using Gaussian Arithmetic and A Probabilistic Model

Emilio A. Diarte-Carot

ABSTRACT

This paper proves that Goldbach's conjecture is true. The proof uses Gaussian modular arithmetic to calculate the number of pairs of odd numbers, K_T, whose sum is a given even number, n, as well as, the number, K_E, of those that can potentially contain prime numbers.

Next, a probabilistic model with a binomial probability distribution is defined, which will be applied to K_E to calculate a function f(x) for the expected value, E(X), where X is the number of pairs formed by two prime numbers.

Finally, the analysis of this function, f(x), will allow us to prove that the conjecture is true.

Keywords and phrases: goldbach conjecture. gaussian arithmetic. prime numbers. binomial distribution.

I. INTRODUCTION

In 1742 the Prussian mathematician Christian Goldbach wrote a letter to Leonhard Euler, see [1], and proposed the following conjecture.

Definition 1. Goldbach's conjecture states that:

"Every even number greater than two can be expressed as the sum of two prime numbers".

Note that, Goldbach's conjecture only requires *the existence of two prime numbers whose sum is an even number greater than two.* Naturally, if there is an even number that cannot be obtained by adding two prime numbers, this even number will be a counterexample to the conjecture.

In the 283 years since its inception, many mathematicians have attempted to prove the conjecture from various angles.

Empirical verification has shown that all even numbers, up to $n \leq 4 \cdot 10^{18}$, hold the the Goldbach conjecture. See [2].

Recently, an empirical approximation of the conjecture was published. This approximation sets an upper bound on the probability that a very large number, N, is a counterexample to the conjecture. Thus, it rules out the existence of counterexamples in practice. See [3]. It would require the use of heuristic reasoning to accept this last statement because it is impossible to prove empirically.

Currently, the conjecture remains unproven.

In this paper, we will use congruences modulo 6 from Gaussian arithmetic, see [4], to accurately calculate the number of pairs of odd numbers that could potentially contain prime numbers and add up to a given even number, n.

London Journal of Research in Science: Natural & Formal

Then, we will use a binomial probability distribution to calculate a function $f(x)$ for the expected value, E(X), where X is the number of pairs formed by two prime numbers. With this, we will prove that the conjecture is true.

The paper is organized as follows:

In Section 2, *The Conjecture from the Perspective of Gauss's Modular Arithmetic*; we will use Gauss's modular arithmetic to analyze the conjecture.

In Section 3, *Probabilistic Model and Expected Value*; we will define the probabilistic model and applies its probability distribution to calculate E(X) based on the number of effective pairs, K_E.

Finally, in Section 4, *The Final Theorem*; we will use the function calculated in the previous section, to prove that the conjecture is true.

Notation

(1) $\mathbb{N}_0 = \{k \mid k \in \mathbb{Z},\ k \geq 0\}$: Set $\mathbb{N}$ including cero.
(2) p: Prime number.
(3) P: Set of prime numbers
(4) $\overline{p}$: Non-prime number.
(5) $\overline{P}$: Set of non-prime numbers.
(6) π_n: The quantity of $p \leq n$.
(7) $A \times B = \{(a,b) | a \in A \text{ and } b \in B\}$: Cartesian Product of sets.
(8) $(p,p) \in P \times P$.
(9) $[i]_6 = \{n \mid n \equiv i \ mod.6,\ n \in \mathbb{N}\}, \quad 0 \leq i < 6$, and also,
(10) $[i]_6 = \{n \mid n = 6k + i,\ k \in \mathbb{N}_0\}, \quad 0 \leq i < 6$.
(11) π_1: The quantity of $p \leq n$, belonging to $[1]_6$.
(12) π_3: The quantity of $p \leq n$, belonging to $[3]_6$.
(13) π_5: The quantity of $p \leq n$, belonging to $[5]_6$.
(14) K_T: the total number of pairs of odd numbers, whose sum is a given even number, n.
(15) $K_E \subset K_T$: Effective pairs that can potentially contain prime numbers.

Remark 1. Simplification of subsequent calculations.

(a) For the sake of brevity in the rest of this paper, we will refer to even numbers as n, instead of writing $2n$ or $2k$, since we know that the sum of two odd numbers is always an even number.
(b) In general, all even numbers considered in this paper will be $n > 4 \cdot 10^{18}$, which will allow us to simplify expressions containing negligible numbers compared to n. Only in examples, tables and figures, used for support or reference, will use numbers with small values.
(c) In general, π_5 is slightly higher than π_1, but the difference becomes negligible when n increases. Furthermore, only the prime numbers $2 \ and \ 3 \notin \ either [1]_6 \ or \ [5]_6$, so we will consider that $\pi_i = \pi_j = \dfrac{\pi_n}{2}$, where i, j, will be 1 and 5.

II. THE GOLDBACH'S CONJECTURE FROM THE PERSPECTIVE OF GAUSS'S MODULAR ARITHMETIC

In this section, we will use congruences, modulo 6, from Gaussian arithmetic to analyze the conjecture.

First, let's review some of the key concepts and properties of modular arithmetic modulo 6, see [4] and [5].

Definition 2. General definition of congruences, modulo k:
"*m is congruent to r, modulo $k, (m \equiv r)$, if $m - r = kn$, with $k, n \in \mathbb{N}$*".

The congruence modulo k is an equivalence relation, because it has the properties *reflexive*, *symmetric*, and *transitive* and, therefore, the k residue classes form a partition of $\mathbb{N}$.

In this paper, we will make $k = 6$ and denote the residue classes as $[i]_6$, with $0 \leq i < 6$.

The choice of $k = 6$ meets the following two criteria:

(a) Keep odd and even numbers in separate sets.
(b) There are 6 classes of residues, modulo 6. $[0]_6$, $[1]_6$, $[2]_6$, $[3]_6$, $[4]_6$ and $[5]_6$. See notations (9) and (10). All primes are odd, except for 2. And $3 \in [3]_6$, the rest of the numbers in this class are multiples of 3, so all the others odd numbers are distributed between $[1]_6$ and $[5]_6$.

Regarding the proof, the relevant property is that the congruence preserves the addition, i.e. we can add the residue classes according to the following table:

Table 1: Symmetric addition table of residue classes modulo 6.

Add	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[0]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$
$[1]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$
$[2]_6$	$[2]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$
$[3]_6$	$[3]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$
$[4]_6$	$[4]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$
$[5]_6$	$[5]_6$	$[0]_6$	$[1]_6$	$[2]_6$	$[3]_6$	$[4]_6$

This table shows the following sums whose result is an even number:

(i) $[0]_6 = [1]_6 + [5]_6$ or $[0]_6 = [3]_6 + [3]_6$.
(ii) $[2]_6 = [1]_6 + [1]_6$ or $[2]_6 = [3]_6 + [5]_6$.
(iii) $[4]_6 = [5]_6 + [5]_6$ or $[4]_6 = [3]_6 + [1]_6$.

Now, to analyze the conjecture and calculate the total numbers of pairs, K_T, and effective pairs, K_E, we will use Figure 1, which is a fundamental and recurring reference here.

In the figure 1, we must look at:
On the horizontal axis, we have:

(1) K_d = number of diagonal elements (pairs).
(2) k = positioning (k_i) of the diagonal elements.
(3) $[i]_6 = i \to \quad \to 6 \times (K_d - 1) + i$. Summands belonging to $[i]_6$

On the vertical axis, we have:

(1) K_d = number of diagonal elements (pairs).
(2) k = positioning (k_j) of the diagonal elements.
(3) $[j]_6 = i \to \quad \to 6 \times (K_d - 1) + j$. Summands belonging to $[i]_6$

Note that, Kd = the total number of pairs of the sum $[i]_6 + [j]_6$ for a given even number n.

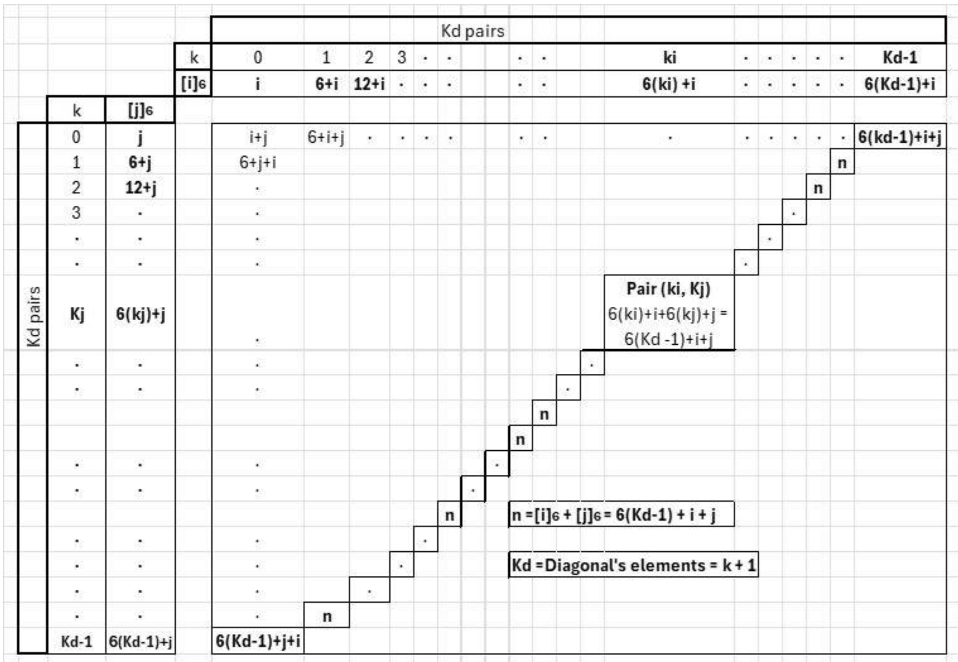

Figure 1: Template for calculating K_T, K_E and K_d, when we know n, i and j.

This figure is a template which allows us to easily calculate the number of pairs for each sum $n = [i]_6 + [j]_6$, with and i *and* $j = 1, 3$ *or* 5 depending on the elements of the sum.

Proposition 1. *The number of pairs, K_T, whose sum is a given even number n, is as follows:*

(i) *For* $n \in [0]_6$; $\quad K_T = \frac{n}{6} + \frac{n}{6} = \frac{n}{3}$.

(ii) *For* $n \in [2]_6$; $\quad K_T = \frac{n-2}{6} + \frac{n+4}{6} = \frac{n+1}{3}$.

(iii) *For* $n \in [4]_6$; $\quad K_T = \frac{n+2}{6} + \frac{n-4}{6} = \frac{n-1}{3}$.

Proof. Table 1 and Figure 1 show that.

If we set i and $j = 1, 3$ or 5 as appropriate and, take into account, that $n = 6(K_d - 1) + i + j$, we get:

(i) $[0]_6 = [1]_6 + [5]_6$ or $[0]_6 = [3]_6 + [3]_6$.
$$K_T = K_d + K_d = \frac{n}{6} + \frac{n}{6} = \frac{n}{3}$$

(ii) $[2]_6 = [1]_6 + [1]_6$ or $[2]_6 = [3]_6 + [5]_6$.
$$K_T = K_d + K_d = \frac{n+4}{6} + \frac{n-2}{6} = \frac{n+1}{3}$$

(iii) $[4]_6 = [5]_6 + [5]_6$ or $[4]_6 = [3]_6 + [1]_6$.
$$K_T = K_d + K_d = \frac{n-4}{6} + \frac{n+2}{6} = \frac{n-1}{3}$$

□

Now, let's calculate the set of effective pairs, K_E that can potentially contain two prime numbers.

Proposition 2. *The number of effective pairs, K_E is either K_d or $K_d + 1$.*

Proof. Bearing in mind that $[3]_6$ only contains one prime, 3.

In all sums that have one of the addends belonging to $[3]_6$, then there will be at most one effective pair. As follow:

(a) Sum $[3]_6 + [3]_6$: It has a single pair $(3, 3)$.
(b) Sum $[3]_6 + [5]_6$: It has a single pair $(3, n-3)$ for all $(n-3) \in [5]_6$ that is prime number.
(c) Sum $[3]_6 + [1]_6$: It has a single pair $(3, n-3)$ for all $(n-3) \in [1]_6$ that is prime number.

So, in this type of sum we will have either 1 or 0 effective pairs.

Now, for the other 3 sums, either $[1]_6 + [1]_6$, $[1]_6 + [5]_6$ or $[5]_6 + [5]_6$, all elements of the diagonal K_d are effective pairs, since, as mentioned in remark 1, all primes, except 2 and 3 belong to $[1]_6$ and $[5]_6$.

In summary, K_E *is either* K_d *or* $K_d + 1$.

Remark 2. In the next section and according to the criterion of simplifying, see remark 1, we will consider $K_E = K_d$ and $n = 6 \times K_d$.

III. PROBABILISTIC MODEL AND EXPECTED VALUE

First, we should noted that the model does not need to calculate the exact number of prime pairs, not even get close to it.

According to the Conjecture's definition, 1, we only need to prove that $X > 0$, where $X =$ *"the number of pairs of prime numbers, whose sum is equal to n"*, so we are going to find a discrete function that minimizes the lower bound of X, for any n.

This function will be the expected value, $E(X)$, of the probability distribution of the mathematical model that we are going to build.

Definition 3. The proposed model is a probabilistic model with a binomial distribution and expected value:

$$E(X) = K_d \times \frac{9}{(\ln K_d - \ln 2)^2}$$

This value is obtained from the probability distribution outlined in the following proposition.

Proposition 3. *The random variable X follows a binomial distribution, with parameters K_d and $p(x) = \frac{9}{(\ln(3 \times K_d))^2}$, whose expected value is $E(X) = K_d \times p(x) = K_d \times \frac{9}{(\ln(3 \times K_d))^2}$.*

Proof. According to [6] and [7]

An experiment, $\in$ and an event A associated with it, with probability, $P(A) = p_a$, is called a Bernoulli trial. If we repeat the experiment n times independently and define the random variable $X = number\ of\ times$ A $occurs$, with p_a fixed for all repetitions, then, X follows a binomial distribution.

Therefore, if we define a mathematical model as follows:

(a) We take, as a Bernoulli trial, an experiment consisting of adding the elements of a pair of numbers (x_i, x_j), the result of which is a number $n = 6 \times K_d$, where $x_i = 6 \times k_i + i \in [i]_6$ and $x_j = 6 \times k_j + j \in [j]_6$, see Figure 1. And, also, we define the event A as *"Pair whose elements are prime numbers"* with some probability to be defined.

(b) We repeat that trial according to the following process:
   (i) From their Cartesian product, notation (8), the first k_d elements of sets $[i]_6$ and $[j]_6$, we take the K_d elements whose sum is a given even number, n, which correspond to the diagonal (k_i, k_j) of the figure 1.
   (ii) And now, we define a random variable X = "the number of pairs containing two prime numbers" where $x = (x_i, x_j)$, is one of them with probability $p(x)$.

We can apply the binomial distribution to this model if it meets the following assumptions:

(1) Each element of the diagonal, (k_i, k_j), whose sum is n, is independent of the others, and we consider them to be so because:

(i) The actual distribution of prime numbers is unknown. While it is clearly chaotic, it is not random: $\pi(n)$ decreases with n as the primes become increasingly separated, although they sometimes appear to cluster together.

(ii) The sum of the elements of x, (x_i, x_j) must add up to n, i.e. $x_1 + x_5 = n$; therefore, as x_1 increases, x_5 decreases and vice versa. This compensates for the non-randomness of the distribution of prime numbers. See figure 2 and, in particular, the line of averages.

(iii) Furthermore, given the magnitude of $K_d = \frac{n}{6}$ and $n > 4 \times 10^{18}$, if we wanted to further ensure the independence of the Kd pairs, we could work with a random sample of them. For example, selecting $\frac{K_d}{2}$ pairs at random would result in the same probability, $p(x)$, as we will see later. Therefore, it is unnecessary to use a sample to calculate the probability.

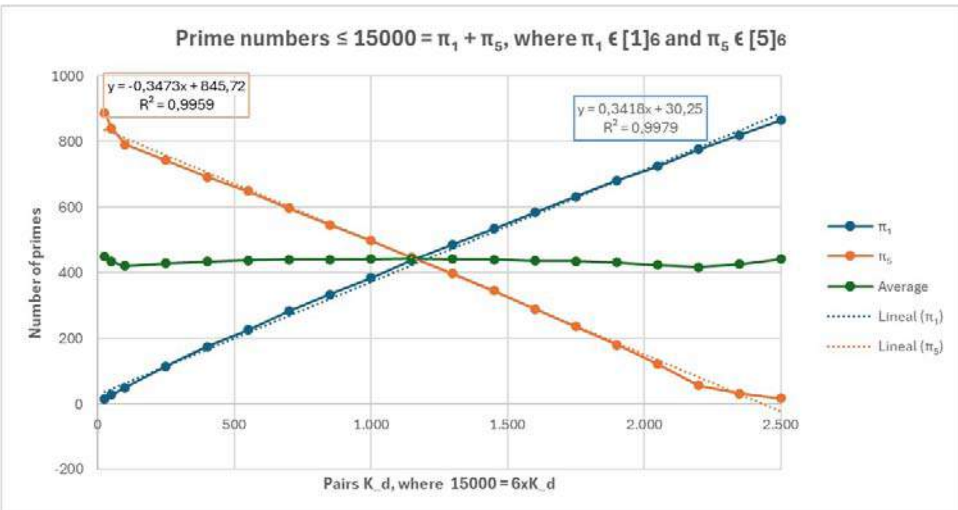

Figure 2: $n = 15,000 = 6 \times K_d$, i.e. $K_d = 2,500$ effective pairs.

London Journal of Research in Science: Natural & Formal

(2) We only have two possible outcomes: x with probability $p(x)$ and $\overline{x}$ with probability $(1 - p(x))$.

(3) The probability that the pair contains 2 prime numbers, $p(x)$, must be equal for all pairs.

As we will see below, the probability of each pair depends on its position on the diagonal, so it is different for each pair.

Fortunately, *our objective of minimizing $E(X)$ and the degree of freedom given by the definition of the conjecture 1, allows us to overcome this problem, assigning the smallest of the calculated probabilities to all pairs.*

In the following calculations, we will consider the simplifications in remark 1.

Let $x = (6 \times k_i, 6 \times k_j)$ be any pair on the diagonal (k_i, k_j), figure 1, and we want to calculate the probability that its elements are prime numbers.

As $6 \times k_i \in [i]_6$ and $6 \times k_j \in [j]_6$, we can consider that the probabilities, $p(6 \times k_i)$ and $p(6 \times k_j)$ are independent and we will use the multiplication principle of probabilities, [7], and we get:

$$p(x) = p((6 \times k_i) \cap (6 \times k_j)) = p(6 \times k_i)p(6 \times k_j).$$

Now, we will use the relative frequency to calculate $p(6 \times k_i)$ and $p(6 \times k_j)$, since, when n is extremely large, the relative frequency converges to probability, [7].

$$\text{The relative frequency} = \frac{\textit{The number of prime numbers}}{\textit{Total number considered}}.$$

Note that, the number of prime numbers is divided equally between $[i]_6$ and $[j]_6$. See Remark 1.

So, Due to the symmetry with respect to the main diagonal, (k_i, k_i), in Figure 1, we get:

$$p(6 \times k_i) = \frac{\pi(6 \times K_d) - \pi(6 \times k_i)}{2 \times (K_d - k_i)} = \frac{\pi(6 \times k_j)}{2 \times (k_j)} \quad \text{and,}$$

$$p(6 \times k_j) = \frac{\pi(6 \times K_d) - \pi(6 \times k_j)}{2 \times (K_d - k_j)} = \frac{\pi(6 \times k_i)}{2 \times (k_i)}$$

Therefore,
$$p(x) = \frac{\pi(6 \times k_j)}{2 \times k_j} \times \frac{(\pi 6 \times k_i)}{2 \times k_i}.$$

The Prime Numbers Theorem, [8], states that $\frac{n}{\ln n} \leq \pi(n)$, and $\pi(n)$ becomes greater than $\frac{n}{\ln n}$ as n increases. So, applying it to $p(x)$ we get:

$$p(x) = \frac{6 \times k_i}{2 \times k_i \times \ln(6 \times k_i)} \times \frac{6 \times k_j}{2 \times k_j \times \ln(6 \times k_j)},$$

and simplifying; we get: $p(x) = \dfrac{3}{\ln(6 \times k_i)} \times \dfrac{3}{\ln(6 \times k_j)}$.

As we can see, p(x) depends on the pair's position on the diagonal and it is variable.

Due to the degree of freedom mentioned above, and since we want to minimize $E(X)$, *let's assign the minimum value of* $p(x)$ *to all pairs*, thereby ensuring that they all have the same value, as required by the binomial distribution, which will also be minimal.

To calculate this minimum value we will use the Fermat's Theorem for differential calculus, [9].

Setting $6 \times k_i = z$ and taking into account that $k_j = K_d - k_i$, and $n = 6 \times K_d$ we can write:

$$f(z) = \frac{3}{\ln(z)} \times \frac{3}{\ln(n-z)}.$$

So, calculating the first derivative $\dfrac{df(z)}{dz}$ and setting it equal to zero we get the minimum of this function at $z = \dfrac{n}{2}$, which corresponds to $k_i = k_j = \dfrac{K_d}{2}$. Value that could be expected due to symmetry and the function is positive within the specified range, $(1, n-1)$. See figure 3

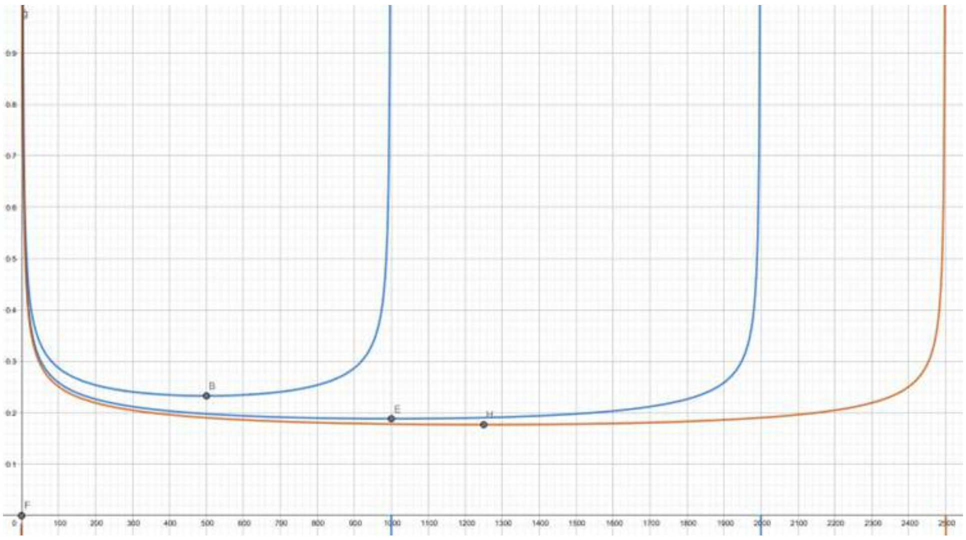

Figure 3: $f(z) = \dfrac{3}{\ln(z)} \times \dfrac{3}{\ln(n-z)}$, $n = 1,000;\ 2,000$ *and* $2,500$.

London Journal of Research in Science: Natural & Formal

Where $B = (500; 0.233)$; $E = (1,000; 0.189)$ and $H = (1,250; 0.177)$.

Then, the value for all pairs is: $p(x) = \frac{9}{(ln(3 \times K_d))^2}$.

And the espected value will be: $E(x) = K_d \times \frac{9}{(ln(3 \times K_d))^2}$.

□

Remark 3. As mentioned in point (iii) of the previous proposition, to increase the independence of pairs K_d, we could randomly eliminate $\alpha \times K_d$, then, the sample of $(1-\alpha) \times K_d$ would contain $(1-\alpha) \times \pi(6 \times K_d)$ prime numbers. However, the odds remain unchanged, since:

(i) Random deletion preserves the symmetries of Figure 1, because it only reduces its dimensions.

(ii) So, to calculate probabilities, we can use the formulas from Proposition 3 for relative frequency, but weighted by the factor $(1 - \alpha)$ and we get:

$$p(6 \times k_i) = \frac{(1-\alpha) \times (\pi(6 \times K_d) - \pi(6 \times k_i))}{2 \times (1-\alpha) \times (K_d - k_i)} \quad \text{and,}$$

$$p(6 \times k_j) = \frac{(1-\alpha) \times (\pi(6 \times K_p) - \pi(6 \times k_j))}{2 \times (1-\alpha) \times (K_d - k_j)}.$$

And the factor $(1 - \alpha)$ can be eliminated, so the result is the same as that of the aforementioned proposition. For this reason, it is not necessary to use the random sampling process.

IV. THE FINAL THEOREM

Theorem 1. *Goldbach's conjecture is true.*

Proof. In proposition 3, we obtained that: $E(X) = K_d \times \frac{9}{(ln(3 \times K_d))^2}$.

This discrete function minimizes the lower bound of X, since, in that proposition, we have simplified some values and, in addition, we take the minimum probability value, $p(x)$, for all pairs. Therefore, we can state that: $X > E(X)$.

in figure 4, we can see that, not only is $X(E)$ below X, but the difference becomes greater as K_d increases.

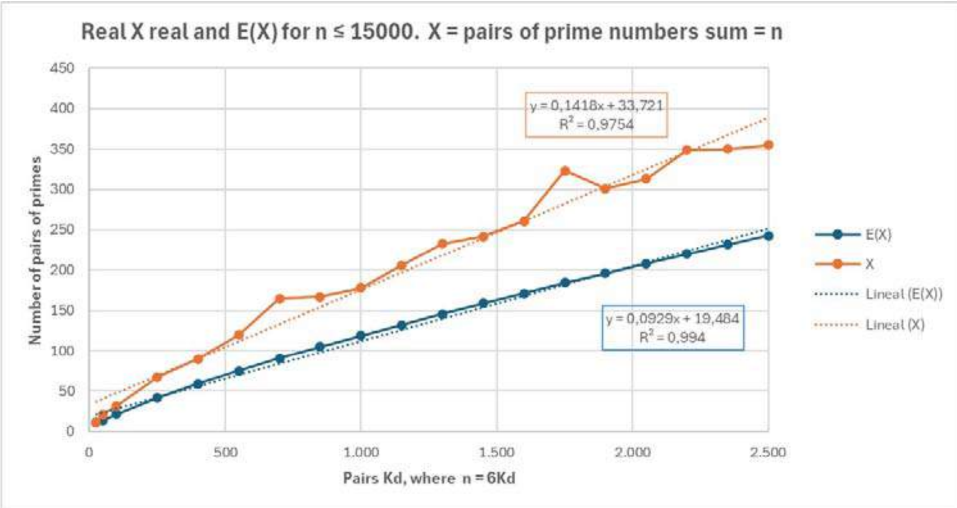

Figure 4: Real X vs $E(X)$, with $n = 15,000$ and $K_d = 2,500$

Finally, we just need to prove that, within the range of interest, $E(X) > 0$ and always increases as K_d increases.. To do this, we could proceed either graphically or analytically.

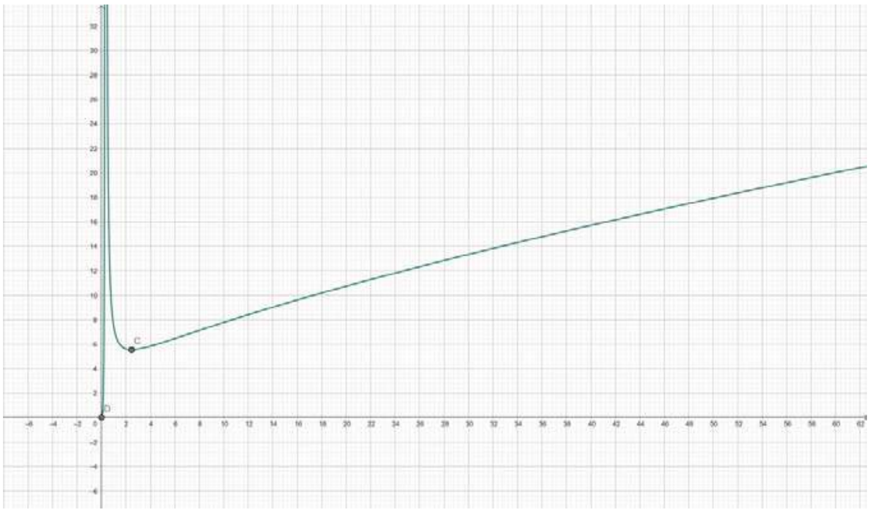

Figure 5: $E(X) = K_d \times \frac{9}{(ln(3 \times K_d))^2}$

Graphically, Figure 5 shows the function and prove that it only has a minimum at C, $2 < K_d < 3$. And, then, the function increases "ad infinitum".

Analytically, we define:

$$f(z) = z \times \frac{9}{(ln(3 \times z))^2}.$$

So, calculating the first derivative $\frac{df(z)}{dz}$ and setting it equal to zero, we get a minimum at $z = \frac{e^2}{3} \approxeq 2.463$, from which the function increases.

In both cases, the conclusion is the same: *Goldbach's conjecture is true.*

$\square$

ACKNOWLEDGMENTS

I thank God for giving me the ability to enjoy mathematical challenges. Above all, I thank Him for giving me a family that always encourages me to persevere. My wife, Marisa; my son Emilio and my daughter Pilar; their respective spouses, Michelle and Carlos; my grandchildren Gabriel, Carlos, Pablo and Pilar.

A very special thank you to Ramón Carbó-Dorca, Institute of Computational Chemistry and Catalysis, Universitat de Girona, Girona, Spain, for their valuable comments and reviews on this paper.

REFERENCES

1. Chr. Goldbach, Letter to Leonhard Euler, Moscow, June 7, 1742
2. T. Oliveira e Silva, S. Herzog and S. Pardi, Silvio, Empirical Verication of the Even Goldbach Conjecture and Computation of Prime Gaps up to 4×1018, Mathematics of Computation, 83 (2014), 288, 2033-2060.
3. A. Farhadian, Goldbach Conjecture: Violation Probability and Generalization to Prime-like Distributions. arXiv:2504.14353v1 [math.NT] 19 Apr 2025.
4. C. F. Gauss, Disquisitiones Arithmeticae, Yale University Press, New Haven, 1965.
5. E. A. Diarte-Carot, Solving the Collatz Conjecture, Using Gaussian Arith-metic., Journal of Applied Mathematics and Physics, 13, 1960-1968.https://doi.org/10.4236/jamp.2025.135109
6. J. Glyn, D. Burley, D. Clements, P. Dyke, J. Searl and J. Wright, Modern Engeneering Mathematics 2nd Edition, Addison-Wesly Publishing Company, Harlow. England 1996, ISBN 0-201-87761-9, pp. 873-874
7. P. L. Meyer, Introductory Probability and Statistical Aplications, translated from the 1st English edition by CF. Prado and G. Ardila. Reading, Mas-sachussets: Addison-Wesly Publishing Company 1970, ISBN 0-201-04715-2, pp. 64 and 12
8. C. F. Gauss, Werke, vol. 2 (1st ed.), Göttingen: Teubner, 1863. pp. 444447
9. J. Stewart, Calculus: Early Transcendentals (6a edición), Brooks/Cole. ISBN 0-495-01166-5 (2008), 273.