Ratio Mathematica Volume 54, 2025

Thinking over the Goldbach conjecture solution from a probabilistic point of view

Rosario D'Amico*

Abstract

This paper aims to provide a set of considerations that allow us to see a possible solution to the problematic issue of Goldbach's "strong" conjecture, which amounts to asserting that *any even natural number greater than 2 can be written as the sum of two prime numbers that are not necessarily distinct.* Specifically, we will show mathematically that a hypothetical scenario in which no even composite number exists as a sum of two primes is impossible. This will be done by adopting a probabilistic method much simpler than the arithmetical attempts already present in literature.

Keywords: Goldbach's conjecture; Prime number; Probability.[1]

* Mps Bank, Messina, Italy; damros978@gmail.com.

[1] Received on January 01, 2025. Accepted on May 31, 2025. Published on June 30, 2025. DOI: 10.23755/rm.v54i0.1662. ISSN 1592-7415; e-ISSN 2282-8214.

1. Introduction

Goldbach's conjecture is one of the oldest and most unsolved problems in that branch of mathematics that studies the properties of integers. In its formulation treated here and called "strong", it states that every even natural number greater than 2 can be written as the sum of two prime numbers that are not necessarily distinct and is therefore a hypothesis on the properties of primes and, particularly, on their distribution among positive integers.

Is this statement, which seems true, doomed to be unable to be demonstrated using existing axiomatic foundations? (See [5] pp. 330-334).

When considering whether every even integer can be expressed as the sum of two primes, it is tempting to view the puzzle as a matter of arithmetic, of number theory, while the answer might also lie in a pattern of probability. Indeed, like any conjecture, it is an assumption based on clues and probable appearances, and some mathematicians believe that this assertion is true based on probabilistic/statistical considerations[2].

This will also be done below to ensure it is easily understood even by those without high mathematical skills.

But what exactly is to be understood by a prime number? There are basically two definitions of prime number: one counts 1 among the prime numbers, the other – and it is the one most widely accepted by mathematicians – excludes it only for mere reasons of convenience and expediency [3]. We are in doubt regarding this question. Take advantage of this ambiguity as a possible lever to try to confirm Goldbach's strong conjecture is the goal of this work.

It will be achieved in the second of the next two paragraphs that make up this paper. In the first we will preliminarily present seven definitions and four useful lemmas for the purpose of establishing the basic conditions of the reasoning that will be proposed and developed in the following. In the second we will construct a particular random experiment so that we can bring the problem of verifying Goldbach's strong conjecture back into the framework of probability theory and to seek its solution by the exclusive use of probabilistic methods.

2. Definitions and preliminary results

The discussion is focused principally on the natural numbers expressed in base 10, that are the elements of the set $\mathbf{N} = \{0, 1, 2, \ldots, \nu, \ldots\}$, and

[2] In 2003, for example, it has been obtained the probability that Goldbach's conjecture fails as roughly $10^{-150.000.000.000}$ for any $n > 4*10^{14}$, SHELDON [7].

algebraic operations with such numbers.

Preliminarily, it is necessary to introduce seven basic concepts, to define both a terminology and a particular mathematical sequence and to present four fundamental results.

Definition 2.1: A natural number is called *a positive integer number*, or simply *a positive integer*, if and only if it is an element of the infinite set $\{1, 2, 3, ..., \nu, ...\} = \mathbf{N}\text{-}\{0\}$.

Definition 2.2: Let $n \in \mathbf{N}\text{-}\{0\}$. A *positive factor* of the given number n is any positive integer that divides n completely, leaving no remainder.

Definition 2.3: A *prime number*, or simply a *prime*, is any positive integer that has no positive factors other than 1 and itself. According to this definition, 1 is considered a prime number.

Definition 2.4: A *prime number* (or a *prime*) is any positive integer that has exactly two distinct positive factors: itself and the number 1. According to this definition, 1 is not a prime number.

Remark 2.1: If only positive integers greater than 1 are considered, the definitions 2.3 and 2.4 are equivalent, as it is possible to easily verify.

Definition 2.5: Let $n \in \mathbf{N}\text{-}\{0,1\}$. *A pair of 2n* is any pair (n-k, n+k), with $k \in \{0, \ldots . , n-2\}$, formed by two positive integers that sum 2n.
Hence, we have n-1 pairs of 2n.

Definition 2.6: Let $n \in \mathbf{N}\text{-}\{0,1\}$. A *non-prime pair of 2n* is any pair of 2n such that at least one of its two positive integers is not a prime number. E.g., (9, 13) and (10, 12) are two non-prime pairs of 22.

Definition 2.7: Let $n \in \mathbf{N}\text{-}\{0,1\}$. A *prime pair of 2n* is any pair of 2n such that both its positive integers are primes. E.g., (3, 19) and (5, 17).

Notation 2.1: Let $n \in \mathbf{N}\text{-}\{0\}$. Assuming the definition 2.3 to be true, ${}_{\mathrm{I}}\Pi(n)$ indicates the quantity of primes that are less than or equal to n.

Notation 2.2: Let $n \in \mathbf{N}\text{-}\{0\}$. Assuming the definition 2.4 to be true, ${}_{\mathrm{bis}}\Pi(n)$ indicates the quantity of primes that are less than or equal to n.

Notation 2.3: Let $n, \eta \in \mathbf{N}\text{-}\{0\}$, with $n \geq \eta$. Let η_n also be the positive integer η corresponding to the natural number (index) n. $\Pi(\eta_n)$ indicates the

quantity of primes in the first η of n natural numbers, i.e., fixed n, the quantity of positive integers $\eta_n = \eta$ – related to the index n – of the tuple (1_n, …, n_n) that are primes and equal to or less than η. Precisely, $\Pi(\eta_n)$ is randomly and definitively identified by one of the two terms: ${}_{I}\Pi(\eta)$ and ${}_{bis}\Pi(\eta)$.[3] If η > 1, we also can write $\eta_n \equiv \eta_\eta$[4] and $\Pi(\eta_\eta) = \Pi(\eta) > 0$.
For any n ∈ **N**-{0,1}, $\Pi(1_n)$ was thus randomly and definitively selected between the two different values: $\Pi(1_n) = {}_{I}\Pi(1) = 1$, iff $\Pi(n) = {}_{I}\Pi(n)$; otherwise, i.e., iff $\Pi(n) = {}_{bis}\Pi(n)$, we have $\Pi(1_n) = {}_{bis}\Pi(1) = 0$.

Remark 2.2: Let n ∈ **N**-{0,1}. It follows from the above that

$${}_{I}\Pi(n) \leq n,\ {}_{I}\Pi(n) = {}_{bis}\Pi(n) + 1 \text{ and } 0 < {}_{bis}\Pi(n) \leq \Pi(n) \leq {}_{I}\Pi(n).$$

Notation 2.4: Let n ∈ **N**-{0}. Later, we will use the symbol "log(n)", instead of "$\log_e n$", as an abbreviation of "logarithm of n to the base e (Euler's number)" , or "natural logarithm of n".

Finally, let n ∈ **N**-{0,1} and k ∈{0, 1, …, n-2}. Based on remark 2.2, we introduce the following sequence of positive numbers:

$${}_{(n)}S_k = [\Pi(n-k)/(n-k)]*[\Pi(n+k)-1/(n-1)] \text{ if } n > 2 \text{ and } k \in \{1, \ldots, n-2\};$$

$${}_{(n)}S_k = {}_{(n)}S_0 = \Pi(n)/n \leq 1 \text{ if } n = 2 \text{ or } k = 0. \qquad (2.1)$$

From the Rosser-Schoenfeld theorem – according to which, for any n ∈ **N** and n ≥ 67, $1/(\log(n) - 0{,}5) < {}_{bis}\Pi(n)/n < 1/(\log(n) - 1{,}5)$ BANESCU [1] – and (2.1), the following four lemmas are derived:

Lemma 2.1: Let n ∈ **N**, n ≥ $4*10^7$ and k ∈ {1, …, n-2}.[5]

Then, ${}_{(n)}S_k < 2/[\log(n) - 1{,}5]$.

Proof. By (2.1) and remark 2.2, we can write

${}_{(n)}S_k = [\Pi(n-k)/(n-k)]*[\Pi(n+k) -1 /(n-1)] \leq {}_{bis}\Pi(n+k)/(n-1)$, from which

${}_{(n)}S_k \leq [(n+k)/(n-1)]*[\,{}_{bis}\Pi(n+k)/(n+k)] < [(2n-2)/(n-1)] / [\log(n+k) -1{,}5] \leq$

$\leq 2/[\log(n) -1{,}5]$, being k ≤ n-2.

[3] Let us consider $\Pi(\eta_n)$ as a key on a computer keyboard. As soon as any two numbers η, n ∈ **N**-{0} are entered and the key $\Pi(\eta_n)$ is pressed, the computer always provides the same result: the random value between ${}_{I}\Pi(\eta)$ and ${}_{bis}\Pi(\eta)$.
[4] Then this implies here that, if η > 1, both or neither of the terms η_n and η_η are prime numbers.
[5] From now on, to choose to place n ≥ $4*10^7$ is only due to the need to have a sufficiently high integer to validate the reasoning that is conducted in this section and in the next one.

Lemma 2.2: Let $n \in \mathbf{N}$, $n \geq 4*10^7$ and $k \in \{1, \ldots, n-2\}$.

Then, ${}_{bis}\Pi(n) - 1 > n/\log(n)$.

Proof. It is sufficient to prove that

$$\{n/[\log(n) - 0{,}5] - n/\log(n)\} > 1 \qquad (2.2)$$

since ${}_{bis}\Pi(n) > n/[\log(n) - 0{,}5]$, for any $n \in \mathbf{N}$ and $n \geq 67$.
It is no difficult to rewrite (2.2) in the form $n > [2\log(n) -1]*\log(n)$. But this inequality is satisfied, because, if $n \in \mathbf{N}$ and $n \geq 67$, $f(n) = n / [2\log^2(n)]$ is a function of n greater than 1 and increasing.

Lemma 2.3: Let $n \in \mathbf{N}$, $n \geq 4*10^7$ and $k \in \{1, \ldots, n-2\}$.

Then, ${}_{(n)}S_k > 1/\log^2(n+92)$.

Proof. For remark 2.2, $\Pi(n) > {}_{bis}\Pi(n) -1$ and, for lemma 2.2, we have ${}_{bis}\Pi(n) -1 > n/\log(n)$. Moreover, for direct verification, we obtain

$\Pi(68-t)/(68-t) > \Pi(94)/(94)$, for any $t \in \{0, 1, ..., 66\}$.

Applying (2.1), we find that

${}_{(n)}S_k = [\Pi(n-k)/(n-k)]*[\Pi(n+k)-1/(n-1)] > 1/[\log(n+92-k)*\log(n+k)]$, from which we get ${}_{(n)}S_k > \{1/[\log(n+92-k)*\log(n+92+k)] \geq 1/\log^2(n+92)\}$ [6].

Lemma 2.4: Let $n \in \mathbf{N}$ and $n \geq 4*10^7$. Let P(2n) be the real function of variable n, so defined:

$$P(2n) = {}_{(n)}S_0 + (1- {}_{(n)}S_0)*{}_{(n)}S_1 + (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S_1)*{}_{(n)}S_2 + \ldots + (1 - {}_{(n)}S_0)*(1- -{}_{(n)}S_1)*(1 - {}_{(n)}S_2)*\ldots*(1 - {}_{(n)}S_{n-3})*{}_{(n)}S_{n-2}. \qquad (2.4)$$

Then, P(2n) assumes increasing(decreasing) values as at least one of its constituent parameters increases (decreases): ${}_{(n)}S_0$, ${}_{(n)}S_1$, ${}_{(n)}S_2$, …, ${}_{(n)}S_{n-2}$.

Proof. From (2.4), it follows that:

---

[6] For this purpose, it is sufficient to show that

$$\log(n-k)*\log(n+k) \leq \log^2(n). \qquad (2.3)$$

Since $\log(n-k) = \{\log[(n-k)/n] + \log(n)\}$ and $\{\log(n+k) = \log[(n+k)/n] + \log(n)\}$, the inequality (2.3) can be written as $\{\log[(n-k)/n] + \log(n)\}*\{\log[(n+k)/n] + \log(n)\} \leq \log^2(n)$,

$$\{\log(n)*\log[(n-k)/n] + \log(n)*\log[(n+k)/n] + \log[(n+k)/n]*\log[(n-k)/n]\} \leq 0,$$

$$\{\log(n)*\log[(n^2-k^2)/n^2] + \log[(n+k)/n]*\log[(n-k)/n]\} \leq 0.$$

The first term of the latter inequality is less than or equal to zero, because it is the sum of two products either negative or zero, being $0 \leq k < n$ and, consequently, $[(n-k)/n] \leq 1$; so that (2.3) is verified.

$1 - P(2n) = 1 - {}_{(n)}S_0 - (1 - {}_{(n)}S_0)*{}_{(n)}S_1 - (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S_1)*{}_{(n)}S_2 - \ldots - (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S_1)*(1 - {}_{(n)}S_2)*\ldots*(1 - {}_{(n)}S_{n-3})*{}_{(n)}S_{n-2}$, from which we have

$1 - P(2n) = (1 - {}_{(n)}S_0)*[(1 - {}_{(n)}S_1) - (1 - {}_{(n)}S_1)*{}_{(n)}S_2 - \ldots - (1-{}_{(n)}S_1)*(1-{}_{(n)}S_2)**\ldots*(1-{}_{(n)}S_{n-3})*{}_{(n)}S_{n-2}]$,

$P(2n) = 1 - [(1 - {}_{(n)}S_0)*(1 - {}_{(n)}S_1)*(1-{}_{(n)}S_2)*\ldots*(1- {}_{(n)}S_{n-3})*(1- {}_{(n)}S_{n-2})]$.

By increasing (decreasing) at least one of the parameters ${}_{(n)}S_0$, ${}_{(n)}S_1$, ${}_{(n)}S_2$, …., ${}_{(n)}S_{n-2}$ the product $(1-{}_{(n)}S_0)*(1-{}_{(n)}S_1)*(1-{}_{(n)}S_2)*\ldots*(1-{}_{(n)}S_{n-3})*(1-{}_{(n)}S_{n-2})$ decreases(increases) and, as a consequence, P(2n) increases(decreases), since, by (2.1) and lemmas 2.1 and 2.3, $0 < {}_{(n)}S_k \leq 1$ for any $k \in \{0,\ldots, n-2\}$.

Remark 2.3: By lemma 2.4, if $n \in \mathbf{N}$ and $n \geq 4*10^7$, there exists an appropriate real number, let us call it ${}_{(n)}S$, such that

$P(2n) = {}_{(n)}S_0 + (1 - {}_{(n)}S_0)*{}_{(n)}S + (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S)*{}_{(n)}S + (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S)*(1 - {}_{(n)}S)*{}_{(n)}S + \ldots + (1 - {}_{(n)}S_0)*(1 - {}_{(n)}S)*(1 - {}_{(n)}S)*\ldots*(1 - {}_{(n)}S)*{}_{(n)}S$ – with $1/\log^2(n+92) < {}_{(n)}S < 2/[\log(n) - 1{,}5]$ –, from which, putting $1 - {}_{(n)}S = {}_nq$, we get $P(2n) = {}_{(n)}S_0 + (1 - {}_{(n)}S_0)*\{{}_{(n)}S*[(1 - {}_nq^{n-2})/(1 - {}_nq)]\}$ and, thus,

$P(2n) = {}_{(n)}S_0 + (1 - {}_{(n)}S_0)*(1 - {}_nq^{n-2})$, since ${}_{(n)}S*[(1- {}_nq^{n-2})/(1- {}_nq)]$ is the sum of the first n-2 terms of a geometric progression, having ${}_{(n)}S$ as first term and ${}_nq < 1$ as common ratio.

Remark 2.4: Let (y, z) be any pair of 2n. Note that the event A: "a given pair (y, z) of 2n is prime" cannot be considered random but is only certain or impossible, whether we use the 2.3 definition or the 2.4 definition of prime number. In fact, because of remark 2.1, the pair (y, z) – like any other pair of 2n – is prime or not, necessarily, since, for definition 2.5, y and z are both positive integers greater than 1. The probability of A thus admits only two possible values: 0 and 1.

3. Probabilistic puzzle: the heart of Goldbach's conjecture argument

In this section, we present a reasoning that can validate the following:

Goldbach's proposition: Let $M = \{1, 2, \ldots, m, \ldots\}$ be a set of positive integers, where $m \geq 4*10^7$. The definition 2.4 of prime number is also satisfied.

Then, the even number 2m can be expressed as the sum of two not necessarily distinct primes.

In this regard, suppose we set up a random mechanism, let's call it Goldbach's proof, – such as the toss of an unrigged die or the extraction with reinserting from an urn – by which a distinctive label, «non-*prime»* or «*prime»*, is equiprobably and definitively attributed to each element of the infinite set M-$\{1\} = \{2, 3, \ldots, m, \ldots\}$ and to each member $1_{\check{n}} = 1$ of the sequence $\{1_n\}_{n\in M}$ in such a way that [7]:

i. Let $\eta, x \in M$, with $\eta \geq x$. Let $F(x_\eta)$ be the relative frequency of the outcome "being a «prime» labeled number" in the first x of η trials, i.e., fixed η, the ratio between the number of positive integers $x_\eta = x$ – related to the index η – of the tuple $(1_\eta, \ldots, \eta_\eta)$ that are labeled as «prime» and equal to or less than x and the number x. If $x > 1$, $x_\eta \equiv x_x$[8] and $F(x_x) = F(x) = \Pi(x)/x$.

ii. Let $\eta \in M$ and $\eta \geq 4*10^7$. Goldbach's proposition holds if and only if 2η can be written as the sum of two as «prime» labeled positive integers not necessarily distinct.

iii. For every $\eta \in M$ -$\{1\}$, $0 \leq \Pi(\eta +1) - \Pi(\eta) \leq 1$; [9]

Within this framework, we make the following five remarks:

Remark 3.1: Let η be any member of the set **N**-$\{0, 1\}$. Notation 2.3 and Goldbach's proof, taken together, contain no reason to believe that $F(1_\eta) = \Pi(1_\eta)$. In fact, $F(1_\eta)$ and $\Pi(1_\eta)$ might take different values because they might be the results of random experiments not necessarily identical. It follows that no prime number is necessarily labeled as «prime», nor every labeled «prime» positive integer is necessarily a prime.

---

[7] For example, suppose we label as «prime» the positive integer $n+1[1_n]$ with $n \in M$, if the $2n^{th}$ $[(2n+1)^{th}]$ roll of an unrigged traditional die leads to an even number; otherwise, $n+1$ $[1_n]$ is a «non-prime» labeled number. This random experiment is possible; its introduction is thus legitimate.

[8] Then this implies here that, if $x > 1$, both or neither of the terms x_η and x_x are «prime» labeled numbers.

[9] By hypothesis (See footnote 7) all 2^{2m-1} possible (2m-1)-permutations with repetition of the two distinct labels, «prime» and «non-prime», – assigned to the first m-1 members of M-$\{1\}$ and to the first m terms of the sequence $\{1_n\}_{n\in M}$ – are equiprobable, each with probability $p^{2m-1} = (1/2)^{2m-1} \geq 0$, and one of them must occur; so the smallest labeled «non-prime» («prime») positive integer does not necessarily exist. Hence, since $F(x_\eta) = {}_{bis}\Pi(x)/x$ may occur for every $\eta, x \in M$ and $\eta \geq x$, we have that i, ii and iii are three compossible and, thus, admissible conditions.

Remark 3.2: Let η ∈ **N**-{0,1}. From remark 3.1 and hypothesis iii, it follows that, unlike necessarily primes, there is no reason to believe that η+1 is a labeled as «prime» number, if Π(η+1) -Π(η) = 1; in this case, in fact, according to hypothesis/condition i, both of the following equalities:

(η+1)*F(η+1) = Π(η+1) = $_{I}$Π(η+1) and η*F(η) = Π(η) = $_{bis}$Π(η), with η+1«non-prime» labeled number, might be satisfied.

Similarly, if Π(η+1) - Π(η) = 0, we cannot affirm that η+1 is labeled as a «non-prime» number, in this case, both the relations:
(η+1)*F(η+1) = Π(η+1) = $_{bis}$Π(η+1) and η*F(η) = Π(η) = $_{I}$Π(η), with η+1«prime» number, might be valid.

Remark 3.3: Let η, x ∈ *M*, with x ≤ η ≥ $4*10^7$. By i, condition ii is clearly equivalent to saying that any sequence $\{F(x_\eta)\}_{x\in\{1,\,2,\,\ldots,\,\eta\}}$ of relative frequencies of the event "being a «prime» labeled number" in the first x of η trials is such that the declarative sentence: "*2η can be written as the sum of two as «prime» labeled positive integers not necessarily distinct*" has the same truth value of either TRUE or FALSE as Goldbach's proposition, which, by remark 2.4, is a necessarily true o certainly false statement.

Looking closely, condition ii adds no relevant information either about the truth value of Goldbach's proposition or, therefore, about the set of all possible sequences (of relative frequencies) of the above type; hence, it does not impact in any way the probability that positive integer η is identified as «prime».

Remark 3.4: Let η ∈ *M*-{1}. Let $P(\eta_{«prime»})$ be the probability that positive integer η is identified as «prime» in the Goldbach proof.
It should also be stressed that the labels assigned to the numbers $1_n = 1$ – corresponding to the index n – and to n, when n varies over the set *M*-{1}, are equiprobably (See footnote 9) and independently chosen, in the sense that the outcome of each of them – to be a «prime» or «non-prime» labeled number – has the same probability p = 1/2 to occur and the probability (1-p) = 1/2 of not occurring, and it does not influence and is not influenced in any way by that of another or the others. Based on the remarks 3.2 and 3.3, by using the binomial distribution formula (See [8] pp. 600-601) and the famous Bayes's rule [2], we thus have:

$$P(\eta_{«prime»}) = \frac{\dfrac{p*(\eta-1)!*[p^{[\Pi(\eta)-1]}*(1-p)^{[\eta-\Pi(\eta)]}]}{[\prod(\eta)-1]!*[\eta-\prod(\eta)]!}}{\dfrac{\eta!*[p^{\Pi(\eta)}*(1-p)^{[\eta-\Pi(\eta)]}]}{[\prod(\eta)]!*[(\eta-\prod(\eta)]!}},$$

from which by simplifying we obtain: $P(\eta_{«prime»}) = F(\eta) = \Pi(\eta)/\eta$.

Remark 3.5: Let $\eta \in M$, $\eta \geq 4*10^7$ and $k \in \{0, \ldots, \eta-2\}$. Let $(\eta-k, \eta+k)$ be a pair of 2η. Also, if and only if $k > 0$, we assume that the event $E_{(\eta, j)}$ has happened, i.e., that every pair $(\eta-j, \eta+j)$ of 2η, with $j \in \mathbf{N}$ varying in the range 0 to k-1, is formed by at least one labeled as «non-prime» positive integer.

According to remark 3.4, the probability ${}_{(\eta)}P_k$ that the pair $(\eta-k, \eta+k)$ of 2η consists of two positive integers both labeled as «prime» may be equal to or greater than

$${}_{(\eta)}S_k = [\Pi(\eta-k)/(\eta-k)]*[\Pi(\eta+k) -1/(\eta-1)], \text{ if } k \in \{1, \ldots, \eta-2\};$$

$${}_{(\eta)}S_k = {}_{(\eta)}P_k = {}_{(\eta)}P_0 = {}_{(\eta)}S_0 = \Pi(\eta)/\eta, \text{ if } k = 0.$$

In fact, if $k \in \{1, ..., \eta-2\}$, we have

${}_{(\eta)}P_k = [\Pi(\eta-k)/(\eta-k)]*\{[\Pi(\eta+k)-1-H_{(\eta, j)}]/(\eta+k-2(k-1) -1 -1)\} = [\Pi(\eta-k)/(\eta-k)]*$

$*\{[\Pi(\eta+k) -1 -H_{(\eta, j)}]/(\eta - k)\}$, where $H_{(\eta, j)}$ is the number of pairs of 2η, excluding the pair (η, η), that belong to the event $E_{(\eta, j)}$ and each having only one positive integer labeled as «prime» (See Appendix A).

Now we are equipped to demonstrate the following:

Theorem 3.1: Goldbach's proposition holds.

Proof. Let $m \in M$ and $m \geq 4*10^7$. Let ${}_GP(2m)$ be the probability that at least one of the m-1 pairs of 2m is formed by two positive integers both labeled as «prime» and not necessarily distinct.
Using probability calculus, we can suppose by lemma 2.4 and remark 3.5 that

$${}_GP(2m) \geq P(2m) = {}_{(m)}S_0 + (1 - {}_{(m)}S_0)*{}_{(m)}S_1 + (1 - {}_{(m)}S_0)*(1 - {}_{(m)}S_1)*{}_{(m)}S_2 +$$

$$+\ldots+ (1 - {}_{(m)}S_0)*(1 - {}_{(m)}S_1)*(1 - {}_{(m)}S_2)*\ldots.*(1 - {}_{(m)}S_{m-3})*{}_{(m)}S_{m-2}, \text{ where}$$

$${}_GP(2m) = {}_{(m)}P_0 + (1 - {}_{(m)}P_0)*{}_{(m)}P_1 + (1 - {}_{(m)}P_0)*(1 - {}_{(m)}P_1)*{}_{(m)}P_2 + \ldots.+ (1 -$$

$$- {}_{(m)}P_0)*(1 - {}_{(m)}P_1)*(1 - {}_{(m)}P_2)*\ldots.*(1 - {}_{(m)}P_{m-3})*{}_{(m)}P_{m-2}.$$

It follows directly from remark 2.3 and the Rosser-Schoenfeld theorem that

$${}_GP(2m) \geq P(2m) = {}_{(m)}S_0 + (1 - {}_{(m)}S_0)*(1 - {}_mq^{m-2}), \text{ where } 0 < {}_{(m)}S_0 < 1 \text{ and}$$

$$q = (1 - {}_{(m)}S) < 1 - [1/\log^2(m+92)]. \qquad (3.1)$$

By virtue of condition ii and remark 2.4, Goldbach's proposition necessarily holds iff ${}_GP(2m) \neq 0$, i.e., iff ${}_GP(2m) = 1$, and so, by (3.1), if and only if

$${}_GP(2m) = 1 \geq {}_{(m)}S_0 + (1 - {}_{(m)}S_0)*(1 - {}_mq^{m-2}).$$

If it were ${}_GP(2m) = 0$, then, by (3.1), we would have

$_{(m)}S_0 + (1 - {}_{(m)}S_0)*(1 - {}_mq^{m-2}) \leq 0$,

$1 - {}_mq^{m-2} \leq - {}_{(m)}S_0/(1 - {}_{(m)}S_0)$,

${}_mq^{m-2} \geq [{}_{(m)}S_0/(1 - {}_{(m)}S_0)] + 1 \geq 1$.

On the other hand, since, if $m \in M$ and $m \geq 67$, $a_m = [1-1/\log^2(m)]^{\log(m)*\log(m)}$ and $b_m = (m-2)/[\log(m+92)*\log(m+92)]$ are respectively the general terms of the two strictly increasing sequences $\{a_m\}$ and $\{b_m\}$ of positive real numbers, one can easily check that:

${}_mq^{m-2} \leq [1-1/\log^2(m+92)]^{m-2} \leq e^{-(m-2)/[\log(m+92)*\log(m+92)]} < 0{,}1 < 1$.

This would lead to a clear contradiction. Hence, ${}_GP(2m) = 1$. Goldbach's proposition is thus proved.

4 Conclusions

The goal was achieved by combining number theory and probability calculus. Goldbach's strong conjecture – i.e., the assertion: "*Every even natural number greater than 2 can be written as the sum of two prime numbers that are not necessarily distinct*" – has already been verified by computer methods for all even natural numbers up to $4*10^{14}$ Bonavoglia [3], Richstein [6] pp.1745-1749. In view of Goldbach's proposition, this statement is thus true for any even integer greater than 2.

However, it is almost impossible to carry out the random experiment "Goldbach's proof " in practice, so much so that if it did occur it would be believed that this process occurred under conditions other than the assumptions under which it is considered theoretically. In fact, with a normal coin that has a fifty percent probability of giving HEAD, the ratio between the number of times you get HEAD in n tosses, with n "large", and the number n is equal to 1/2 with probability close to 1. But with the coin of numbers labeled as «prime» the relative frequency decreases with each toss and is generally much less than 1/2. In short, Goldbach's proof is a sort of thought experiment whose only purpose is to explore by means of probability the sequence of primes, highlighting surprising or paradoxical consequences.

Finally, it is still opened a numerical problem: if an even number is the sum of two prime numbers, what are these two numbers? The greater the starting even number, the more difficult it is to identify the two primes that make up its sum.

This essay offers us a different perspective on a long-standing question, anyway, providing a good complement to existing literature.

Appendix A

In the assumptions of remark 3.5, we first show that it is possible to have

$H_{(\eta, j)} \lessapprox (k-1)/\log(\eta-1)$, where $\eta \in \mathbf{N}$, $\eta \geq 4*10^7$, $k \in \{1, ..., \eta-2\}$ and $j \in \{0, ..., k-1\}$.

In this regard, suppose that the event $E_{(\eta,j)}$ has occurred and that we know which of the numbers forming the η-1 pairs of 2η of $E_{(\eta, j)}$ are those labeled as «prime» – e.g., all those of the type η-j (or η+j) and anyway at most one for each pair of 2η of $E_{(\eta,j)}$ –.

Therefore, since $[\Pi(\eta +k -1) -\Pi(\eta) +\Pi(\eta) -\Pi(\eta -k +1) \pm 1]$ is the number of positive integers labeled as «prime» and included in the interval $[n -k +1, n +k -1]$, we can reasonably expect that $H_{(\eta, j)} \approx [\Pi(\eta +k -1) - \Pi(\eta -k +1)]/2$ and, by remark 2.2 and the Rosser-Schoenfeld theorem, that

$2H_{(\eta, j)} \approx [\Pi(\eta +k -1) -\Pi(\eta -k +1)] \lessapprox [(\eta +k -1)/\log(\eta +k -1)] - [(\eta -k +1)/\log(\eta +k -1)] \leq$

$\leq (2k-2)/\log(\eta+k-1) < (2k-2)/\log(\eta-1)$, because, as said earlier, we have

$\Pi(68-t)/(68-t) > \Pi(94)/(94)$, for any $t \in \{0, 1, ..., 66\}$. Hence, we can assume that

$H_{(\eta, j)} \lessapprox (k-1)/\log(\eta-1)$, with $\eta \in \mathbf{N}$, $\eta \geq 4*10^7$, $k \in \{1, ..., \eta-2\}$ and $j \in \{0, ..., k-1\}$. (A.1)

Then, since by definition $H_{(\eta, j)} = 0$, if $k = 1$, establish the truth of

${}_{(\eta)}P_k = [\Pi(\eta-k)/(\eta-k)]*\{[\Pi(\eta+k) -1 -H_{(\eta, j)}]/(\eta -k)\} \geq {}_{(\eta)}S_k$ means to prove (for $k > 1$) that

$[\Pi(\eta-k)/(\eta-k)]*\{[\Pi(\eta+k) -1 -H_{(\eta, j)}]/(\eta-k)\} \geq [\Pi(\eta-k)/(\eta-k)]*\{[\Pi(\eta+k) -1]/(\eta -1)\}$, which takes the form

$[\Pi(\eta+k) -1 - H_{(\eta, j)}]/(\eta -k) \geq [\Pi(\eta+k) -1]/(\eta -1)$. By doing the calculations, we obtain

$1 - \{H_{(\eta, j)} / [\Pi(\eta+k) -1]\} \geq (\eta -k)/(\eta -1)$,

$(k -1)/(\eta-1) \geq H_{(\eta, j)} / [\Pi(\eta+k) -1]$. Thus, applying (A.1), it is sufficient to show that

$(k -1)/(\eta-1) \gg (k -1) /\{[\Pi(\eta+k) -1]*\log(\eta-1)\}$, and so that

$1 \gg (\eta-1) /\{[\Pi(\eta+k) -1]*\log(\eta-1)\}$, where $k \in \{2, ..., \eta-2\}$. (A.2)

But (A.2) holds if $\{(\eta-1) / [\log(\eta-1) - (1/2)]\} -1 \gg (\eta-1)/\log(\eta-1)$.

In fact, by remark 2.2 and using the above-mentioned Rosser-Schoenfeld theorem, we have $(\Pi(\eta+k) -1) < (\Pi(\eta+k) \geq {}_{bis}\Pi(\eta+k) \geq (\eta+k) / [\log(\eta+k) - (1/2)]$, and therefore $(\Pi(\eta+k) -1) < (\Pi(\eta+k) \geq {}_{bis}\Pi(\eta+k) \geq (\eta+k) / [\log(\eta+k) - (1/2)] > (\eta-1) / [\log(\eta-1) - (1/2)]$,

since, if $\eta \in \mathbf{N}$ and $\eta \geq 67$, $g(\eta) = (\eta -1)/[\log(\eta -1) - (1/2)]$ is an increasing function of η greater than 1.
The inequality $\{(\eta -1) / [\log(\eta -1) - (1/2)]\} -1 \gg (\eta -1) / \log(\eta -1)$ is satisfied if

$-\log(\eta -1)*[\log(\eta -1) - (1/2)] \gg - (1/2)*(\eta -1)$,

$\log^2(\eta -1) - (1/2)*[\log(\eta -1)] \ll (1/2)*(\eta -1)]$, and even more so, if

$\log^2(\eta -1) \ll (1/2)*(\eta -1)$.

But this last relation is true, because, if $\eta \in \mathbf{N}$ and $\eta \geq 4*10^7$, $q(\eta) = (\eta -1)/\log^2(\eta -1)$ is an increasing function of η much greater than 2. (See [4] pp. 62-64). Therefore, (A.2) is demonstrated.

References

[1] M. BANESCU, *On the function π(x)*, Analele ştiinţifice ale Universităţii "Ovidius" Constanţa. Seria Matematică, vol. 22(1), 2014, pp. 25-33. DOI:10.2478/auom-2014-0002.

[2] T. BAYES & R. PRICE (1763). "An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S." *Philosophical Transactions of the Royal Society of London*. **53**: 370 –18 DOI:10.1098/rstl.1763.0053.

[3] P. BONAVOGLIA, *1 è un numero primo*, in http://www.crittologia.eu/mate/1_primo.html.

[4] R. D'AMICO (2023), *Probabilistic puzzle: toward a solution of Goldbach's conjecture?* in Journal of Mathematical Economics and Finance, [S.l.], vol. 9, n .2, pp. 51-65, dec. 2023. ISSN 2458-0813. DOI: https://doi.org/10.14505/jmef.v9.2(17).04, Available also at: https://journals.aserspublishing.eu/jmef/article/view/8311.

[5] M. DU SAUTOY (2003), *The Music of the Primes*, tr.it. *L'enigma dei numeri primi. L'ipotesi di Riemann, il più grande mistero della matematica*, Biblioteca Univ. Rizzoli, Milano, 2005.

[6] J. RICHSTEIN, *Verifying the Goldbach conjecture up to 4.10^{14}*, Mathematics of Computation, vol. 70, no. 236, July 2000.

[7] N. SHELDON, *A Statistician's Approach to Goldbach's Conjecture,* Teaching Statistics, vol.25, n.1, pp. 12-13, Spring 2003, https://doi.org/10.1111/1467-9639.00108.

[8] G. ZWIRNER, L. SCAGLIANTI, *Itinerari nella matematica,* vol. 2, CEDAM, Padova, 1990.