ISSN: 2994-9459

Research Article

Current Research in Statistics & Mathematics

A Mathematical Proof of the Strong Goldbach's Conjecture

Vassilly Voinov*

Independent scholar, Almaty, Republic of Kazakhstan

***Corresponding Author**
Vassilly Voinov, Independent scholar, Almaty, Republic of Kazakhstan.

Submitted: 2025, Dec 08; **Accepted:** 2026, Jan 08; **Published:** 2026, Jan 19

Citation: Voinov, V. (2026). A Mathematical Proof of the Strong Goldbach's Conjecture. *Curr Res Stat Math, 5*(1), 01-06.

Abstract

A proof of Goldbach's conjecture based on the additive number theory and the theory of B-trees is presented. A core idea of the proof is the partitioning of a set of primes into parts with exactly two elements that are bounded subsets summed to a given even integer n *> 2. An algorithm realizing the idea is suggested. For illustration, the R-script of the algorithm is appended. Results of this research also confirm the fact that the proposed approach for constructing Goldbach's pairs is in complexity class P.*

Keywords: Additive Number Theory, B-Trees, Linear Diophantine Equations, Partitions, Strong Goldbach's Conjecture, Complexity Class P

1. Introduction

The Goldbach's conjecture was first suggested by German mathematician Christian Goldbach in his letter to Leonhard Euler in 1742. In recent terms, it states that **every even integer greater than 2 can be expressed as the sum of two prime numbers**. This conjecture is also known as the Strong Goldbach's conjecture. Using the sieve of Eratosthenes and a computer, it was shown that the conjecture is true for all even natural numbers n less than $4 \cdot 10^{18}$. There were numerous attempts to solve the problem mathematically (see, e.g. [1]), but up to now the conjecture is still considered as being unsolved. Many heuristic approaches are also known. The last one dated 2025, e.g., uses empirically found symmetries in sets of prime numbers [2].

This paper is organized as follows: Section 2 provides a mathematical background for the proposed approach and considers solutions of a linear Diophantine equation used for constructing Goldbach's pairs of primes, Section 3 discusses the infinitness of even integer numbers for which the conjecture is valid, Section 4 provides results of a computer experiment, Section 5 contains a discussion and conclusions, Appendix gives the R-script that provides Goldbach's pairs for a given even integer n.

2. Mathematical Background

Consider a set $A = \{a_1, a_2, \dots, a_l\}$, $l \in \mathbb{N}$, of ordered primes needed for presenting a given even integer n as a sum of exactly two possibly identical elements of A. To obtain those presentations, we may to solve the linear Diophantine equation

$$a_1 s_1 + a_2 s_2 + \cdots + a_l s_l = n, \tag{1}$$

where variables s_i, $i = 1, \dots, l$, are nonnegative integers. Adjusting formulas (11.7) and (11.8) of Voinov and Nikulin [3] for our task, we can find the number of solutions and solutions themselves using the sequence of embedding sums

$$R_a = \sum_{s_l=0}^{\left[\frac{n}{a_l}\right]} \sum_{s_{l-1}=0}^{\left[\frac{n-a_l s_l}{a_{l-1}}\right]} \cdots \sum_{s_2=0}^{\left[\frac{n-a_l s_l-a_{l-1}s_{l-1}-\cdots-a_3 s_3}{a_2}\right]} 1 \tag{2}$$

that actually presents a self-balanced B-tree. The sign $[x]$ in (2) means the integer part of x. Taking into account that $s_1 = (n - a_l s_l - \cdots - a_2 s_2)/a_1$ and $s_1 + s_2 + \cdots + s_l = 2$, where $\{s_2, s_3, \dots, s_l\}$ are sets of summation indices in (2), the solutions (partitions of A on exactly 2 parts) defined by a search in B-tree, are written as

$$\{a_1^{s_1}, a_2^{s_2}, \dots, a_l^{s_l}\}. \tag{3}$$

The notation (3) means that a particular solution will have s_1 terms a_1, s_2 terms a_2, and so on. Traversing through all nodes of the B-tree and using necessary conditions on s_i, $i = 1,2, \dots, l$ the solutions as sets $\{s_1, s_2, \dots, s_l\}$ can be obtained explicitly.

From the theory of B-trees (see, e.g.,[4]). it follows that the complexity of one search is $O(logn)$. For $O(n)$ sequential searches, we have to perform, the final complexity will be $O(nlogn)$. From this it follows that the proposed approach for constructing Goldbach's pairs is in P.

Example 1

Let $A = \{2,3,5,7\}$, $n = 10$, then

$$R_a = \sum_{s_4=0}^{1} \sum_{s_3=0}^{\left[\frac{10-7s_4}{5}\right]} \sum_{s_2=0}^{\left[\frac{10-7s_4-5s_3}{3}\right]} 1, \; s_1 = (10 - 7s_4 - 5s_3 - 3s_2)/2. \tag{4}$$

Calculations of R_a and s_1 can be presented as the following B-tree:

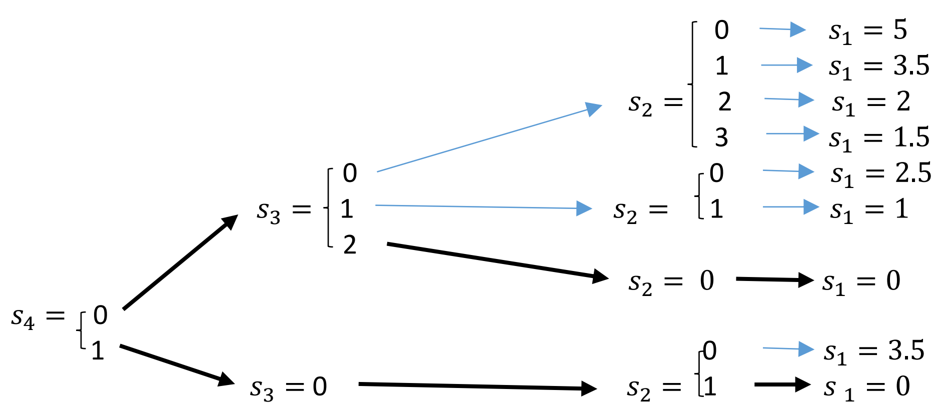

Figure 1: The B-tree generated by the algorithm in (4)

From Fig. 1 we see 9 traverses starting from the root s_4, passing through all 4 nodes and finishing at leaves s_1. Two of them, marked by bold-faced arrows, satisfy the conditions that s_i, $i = 1, \dots, 4$, are nonnegative integers and $s_1 + s_2 + s_3 + s4 = 2$: $\{0,0,2,0\}$ and $\{0,1,0,1\}$. Note that the variables are listed in the backward direction. Thus, there are 2 solutions of the equation $2s_1 + 3s_2 + 5s_3 + 7s_4 = 10$ that correspond to two Goldbach's representations: 5+5 = 10 and 3+7 = 10.

This example clearly clarifies a given below algorithm that can be used for constructing Goldbach's pairs.

Alg. 1

- Generate a set $A = \{a_1, a_2, \dots, a_l\}$ of ascending primes less than n,
- Solve in bounded nonnegative integers a linear Diophantine equation $a_1 s_1 + a_2 s_2 + \cdots + a_l s_l = n$ using the R-command "get.subsetsum" from the R-package "nilde" [5],
- Construct and traverse a corresponding B-tree,
- Select traverses that satisfy the condition $s_1 + s_2 + \cdots + s_l = 2$ and list the pairs.

3. A Note on Existing Goldbach's Pairs for Every Even Integer

The question of existing Goldbach's pairs for every even integer greater than 2 can be answered by considering a lower bound on the number N of linear Diophantine equation solutions. This problem has been thoroughly considered in [6]. A brief summary of it reads as follows. Having introduced the parallelepiped $B(y_1, \dots, y_l)$ as the set of points $y_1, \dots, y_l$ such that $a_i s_i \leq y_i \leq a_i(s_i + 1)$, $i = 1, \dots, l$, and $\sum_{i=1}^{l} y_i \leq n$ that has l-dimensional volume $\prod_{i=1}^{l} a_i$, Beget-Dov suggested the bounds defined by points s_i, $i = 1, \dots, l$, that belong to a B, if those points are in the pyramid with volume $n^l/l!$. Padberg modified the Beget-Dov's bounds for a stronger result [7,8]!. Lambe introduced even better bounds [9].

$$\binom{n+l}{l} / \prod_{i=1}^{l} a_i \leq N \leq \binom{n+\sum_{i=1}^{l} a_i}{l} / \prod_{i=1}^{l} a_i. \tag{5}$$

These bounds possess an important property that the ratio of an upper bound to a lower one tends to one when l and n become large. Moreover, it follows that even for extremely large values of even numbers n, the lower bound $\binom{n+l}{l} / \prod_{i=1}^{l} a_i$ being very small, is still positive. In other words, any linear Diophantine equation with positive integer coefficients possesses at least one solution, thus conforming validity of the strong Goldbach's conjecture.

4. Computer Experiment

The algorithm suggested in Section 2 for constructing Goldbach's pairs of prime numbers has been realized as the R-script titled "conjecture" (see Appendix). This R-script was tested for 14 even integers n: 100, 120, …, 360. Number N of solutions and computing times of the PC Intel ® Core ™ i7-2600 CPU @ 3.40 GHz, RAM of 24.00 GB are given in Table 1 for every n.

n	N	Time, sec.
100	6	0.397
120	12	1.054
140	7	2.642
160	8	5.798
180	14	12.509
200	8	26.098
220	9	51.439
240	18	99.363
260	10	187.836
280	14	340.734
300	21	612.189
320	11	1063.721
340	13	1850.295
360	22	3103.453

Table 1: *N* is the number of prime pairs representing n. Computing times are given in the third column.

Section 2 states that the complexity of the proposed approach for proving Goldbach's conjecture is $O(nlogn)$. The statistical fit of computing times in the third column of this table shown in Fig. 2 does not contradict this conclusion.

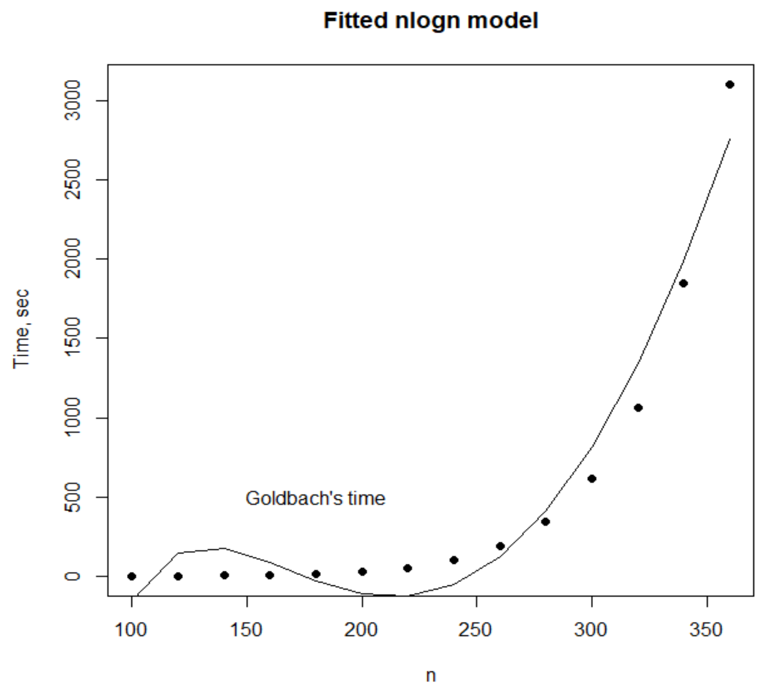

Figure 2: Fit of times by *nlogn* model

Statistical criteria of the fit are: residual standard error equals 205.7 on 10 degrees of freedom, multiple R-squared is 0.961, adjusted R^2 = 0.9494, F-statistic is 82.24 on 3 and 10 degrees of freedom, *p*-value equals 2.388e-07, AIC = 194.1556.

Note that the proposed algorithm has no limit on the maximal possible values of *n*. Only the power and memory capacity of a computer set a bound on those values. To illustrate the usefulness of the R-script "conjecture", two sets of Goldbach's pairs for *n* = 122 and *n* = 242 are shown below:

1. For n = 122: 61+61, 43+79, 19+103, 13+109,
2. For n = 242: 103+139, 79+163, 61+181, 43+199, 31+211, 19+223, 13+229, 3+239.

5. A Discussion and Conclusions.

Additive number theory proves to be very powerful for solving numerous, even very hard mathematical problems like, e.g., the strong Goldbach's conjecture. The theory permitted, e.g., to find polynomial-time solutions for subset sums, one-dimensional bin-packing, traveling salesman and other problems [10]. It is of importance to mention the difference between the approach used for analysis of these problems and that of Goldbach's conjecture. In the first case, the analysis was based on searching for balanced **binary** trees with no more than two keys at every node. In the second one, we used searching the self-balanced ***B*-tree** that may have more than 2 keys at a node. Nevertheless, the one *O*(n) completeness for both cases is *O*(*nlogn*). Note also that the function *nlogn* is well approximated by a polynomial. It is worth also mentioning the decision tree problem [11], p. 282. In 1976, based on the NP- completeness of the three-dimensional matching (3DM) problem, Hyafil and Rivest [12] decided that the decision tree problem is NP-complete. Today we know that 3DM is actually in P (see [10], p. 9) and, hence, so is the decision tree problem. Summarizing all the above, we may conclude that this research provides a firm mathematical proof of the strong Goldbach's conjecture.

Declarations

Funding: Not applicable

Conflict of Interest/ Competing Interests: The author declares that he has no competing interests.

Availability of Data and Material: All data generated or analyzed during this study are included in the manuscript.

Code Availability: R version 4.4.1, the R-script used is provided in Appendix.

Authors' Contributions: Not applicable.

Acknowledgements: The author is grateful to Hazell Mitchell for her help in this article publishing.

References

1. Goldbach's conjecture. https://en.wikipedia.org/ goldbach's conjecture, retrieved on November 01, 2025.
2. Symmetries of prime numbers. https://habr.com/ симметрия простых чисел/хабр, retrieved on November 01, 2025.
3. Voinov, V. G., & Nikulin, M. S. (1997). On a subset sum algorithm and its probabilistic and other applications. In *Advances in combinatorial methods and applications to probability and statistics* (pp. 153-163). Boston, MA: Birkhäuser Boston.
4. B-tree. https://en.wikipedia.org/b-tree, retrieved on November 01, 2025.
5. Pya Arnqvist, N., Voinov, V., Makarov, R., Voinov, Y.: "nilde": Nonnegative integer solutions of linear Diophantine equations with applications (2021). R-package version 1.1-7. https://CRAN, Rproject.org/package-nilde, https://doi.org/10.13140/RG.2.2.26198.19523.P
6. Mahmoudvand, R., Hassani, H., Farzaneh, A., & Howell, G. (2010). On the Number of Nonnegative Solutions to the Inequality $a_1x_1+\cdots+a_r x_r \leq n$ IAENG International Journal of Applied Mathematics, 40(1), 1-5.
7. Beged-Dov, A. G. (1972). Lower and upper bounds for the number of lattice points in a simplex. SIAM Journal on Applied Mathematics, 22(1), 106-108.
8. Padberg, M.W. (1971). A remark on "an inequality for the number of lattice points in a simplex". SIAM Journal on Applied Mathematics 20(4) 638-641.
9. Lambe, T. A. (1974). Bounds on the number of feasible solutions to a knapsack problem. SIAM Journal on Applied Mathematics, 26(2), 302-305.
10. Voinov, V. (2025). Partitions, the P versus NP problem and applications. Curr Res Stat Math, 4(2), 01-13. https://doi.org/10.33140/CRSM.04.02.02
11. Garey, M.R. &Johnson, D.S. (1979). Computers and intractability: A guide to the theory of NP-completeness. WH Freeman and Co, New York.
12. Hyafil, L. & Rivest, R.L.(1976). Constructing optimal binary decision trees is NP-complete. Information Processing Letters 5(1), 15-17

Appendix: R-script "conjecture"

```
######################################################################################

#### To use the script one has to: 1) install R, 2) install R-packages: "primes", "nilde"
#### and "microbenchmark", 3) type the desired value of n in the second line of the script,
#### 4) copy the script and put it into R's editor, 5) click Ctrl A and Ctrl R.
library(primes);library(nilde);library(microbenchmark)
n=120 ## type it here !!!
a<-generate_primes(max=n)
Tnat<-rep(NA,1)
a1<-length(a)
a2<-microbenchmark(a4<-get.subsetsum(a=a,M=2,n=n,problem="bsubsetsum",
bounds<-rep(2,times=a1)),times=1L)
T1<-a2$time/1000000000
a5<-microbenchmark(a6<-as.matrix(a4$solutions[,colSums(a4$solutions)==2]),
times=1L)
T2<-a5$time/1000000000
a4$p.n
for(i in (1:a4$p.n)){
a1i<-a[a6[,i]>0]
a2i<-a1i[!is.na(a1i)]
a3i<-length(a2i)
if(a3i==1){
a2i<-c(a2i,a2i)}
print(a2i)};Tnat<-T1+T2;message("Tnat = ",Tnat)
Lb<-choose(n+a1,a1)/prod(a);
message("Lb = " ,Lb)
```

For n=120 the output must be as follows:
[1] 12 ## the number of decisions
Solutions: 59+61, 53+67, 47+73, 41+79, 37+83, 31+89, 23+97, 19+101, 17+103, 13+107, 11+109, 7+113
Computing time in sec. Tnat = 1.0351034
Lower bound for N Lb = 1.01862479223064e-15.