On Goldbach's Conjecture: An Algebraic and Combinatorial Formulation

Mar Detic

May 2021 (Revised June 2025)

Abstract

We present an algebraic and combinatorial formulation of Goldbach-type representations of integers as sums of primes. Every odd prime is written in the canonical linear form $2a+1$ (with $a \in \mathbb{N}_0$) and the sum of k primes becomes a linear Diophantine constraint on the corresponding a-variables. This transforms the problem into the study of integer lattice points on hyperplanes of the form $\sum_{i=1}^{k} a_i = \frac{N-k}{2}$, together with primality filters on the linear forms $2a_i+1$. We relate this combinatorial perspective to classical analytic heuristics (Hardy–Littlewood), additive-combinatorics sumset language, and partition/stars-and-bars counting, and we provide illustrative examples and an inline visualization for the $k=2$ case.

1 Introduction

Goldbach's conjecture (strong form) asserts that every even integer greater than 2 can be expressed as the sum of two primes. The (weak) form asserts every odd integer greater than 5 is the sum of three primes (now proved by Helfgott). More generally, one may study representations of an integer N as the sum of k primes:

$$N = P_1 + P_2 + \cdots + P_k,$$

with each P_i prime and $k \geq 2$.

This paper develops and cleans up a combinatorial-algebraic framework that parameterizes candidate decompositions by integer variables a_i through the identity $P_i = 2a_i + 1$. The core observation is that the parity and linearity of that representation reduce the additive structure to a hyperplane constraint in $\mathbb{Z}^k$, with primality acting as a multiplicative filter.

2 Representation of primes and general formulation

Every odd integer (and in particular every odd prime) can be written uniquely as

$$P = 2a + 1, \qquad a \in \mathbb{N}_0,$$

where we define $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. For representations of N as a sum of k primes (all written in the form $2a_i + 1$), we have:

$$N = \sum_{i=1}^{k}(2a_i + 1) = 2\sum_{i=1}^{k} a_i + k.$$

Rearranging gives the central linear constraint

$$\sum_{i=1}^{k} a_i \ = \ \frac{N-k}{2}. \tag{1}$$

Note the parity constraint: the right-hand side of (1) must be an integer, so $N \equiv k \pmod 2$. This aligns with the simple observation that a sum of k odd numbers has parity equal to $k \bmod 2$.

Hence, for fixed N and k with matching parity, the set of *candidate* k-tuples $(a_1, \dots, a_k)$ is exactly the integer lattice points on the hyperplane $\sum a_i = (N-k)/2$ with $a_i \geq 0$. The Goldbach-type problem is then the problem of showing that at least one of those lattice points yields primes when substituted into the linear forms $2a_i + 1$.

3 Combinatorics: Stars and Bars and counting candidates

From elementary combinatorics (the stars-and-bars argument), the number of non-negative integer solutions to $\sum_{i=1}^{k} a_i = M$ is

$$\binom{M+k-1}{k-1},$$

where $M = (N-k)/2$. This counts candidate tuples without regard to primality. The primality condition significantly restricts this set: among the $\binom{M+k-1}{k-1}$ tuples, only those for which every $2a_i + 1$ is prime are valid Goldbach decompositions.

Thus the problem separates naturally into two components:

1. *Combinatorial breadth*: there are many candidate tuples (polynomial/exponential in k, M as above),
2. *Multiplicative constraint*: primality of each linear form $2a_i + 1$.

4 Specialization: the $k = 2$ (strong Goldbach) case

For $k = 2$ we obtain the simple relation

$$a_1 + a_2 = \frac{N-2}{2} = X - 1,$$

where $X = N/2$. This is the form used in the earlier drafts and in computational checks: candidate pairs (a_1, a_2) lie on the line $a_1 + a_2 = X - 1$ with $0 \leq a_1, a_2 \leq X - 1$, and the pair is valid if both $2a_1 + 1$ and $2a_2 + 1$ are prime.

4.1 Worked examples

Example (even): $N = 18$**.** Here $X = 9$, so $a_1 + a_2 = 8$. The integer pairs (a_1, a_2) with $0 \le a_1 \le a_2 \le 8$ are:

$$(0, 8), (1, 7), (2, 6), (3, 5), (4, 4).$$

Compute $2a + 1$ values:

$$(1, 17), (3, 15), (5, 13), (7, 11), (9, 9).$$

Only $(2, 6)$ gives $(5, 13)$ (both prime) and $(3, 5)$ gives $(7, 11)$ (both prime). So $G(18) = 2$ (counting ordered or unordered appropriately).

Example (odd): $N = 19$**.** For $k = 3$ (weak Goldbach), one can set $k = 3$ and use (1). For $k = 3$, $\sum a_i = \frac{19-3}{2} = 8$. Many compositions exist; checking primality of the three linear forms yields actual decompositions such as $19 = 3 + 7 + 9$ (but 9 is not prime) — valid prime decompositions include $19 = 3 + 5 + 11$ which corresponds to a-tuple $(1, 2, 5)$. (This illustrates the method; the classical weak Goldbach for sufficiently large N is known to hold.)

5 Geometric interpretation and visualization

Equation (1) defines an affine hyperplane in $\mathbb{R}^k$, and the candidate tuples are the integer lattice points on that hyperplane in the non-negative orthant. For $k = 2$, this is the line $a_1 + a_2 = X - 1$ in the integer lattice $\mathbb{Z}^2$. For larger k it is a $(k-1)$-dimensional simplex slice.

We provide an inline visualization for the $k = 2$ case with the concrete example $N = 18$ (so $X = 9$ and $a_1 + a_2 = 8$). The plotted points are the integer lattice points on the line, and valid prime pairs are highlighted.

Remarks. For $k > 2$ the visualization becomes higher-dimensional: one may visualize 2D projections or heatmaps of the density of prime-valid tuples on slices of the simplex. The hyperplane approach gives a direct geometric interpretation of where candidate decompositions live.

6 Analytic heuristic and probabilistic density

Writing $P = 2a + 1$ gives an approximate density of the a-values corresponding to primes:

$$\Pr(2a + 1 \text{ is prime}) \approx \frac{1}{\log(2a + 1)}.$$

Under a heuristic independence assumption for the primality of the different linear forms (which must be treated carefully because of correlations), the expected number of valid k-tuples on the hyperplane $\sum a_i = M$ is heuristically

$$E\big[G_k(N)\big] \approx \sum_{\substack{a_1+\cdots+a_k=M \\ a_i \ge 0}} \prod_{i=1}^{k} \frac{1}{\log(2a_i + 1)}.$$

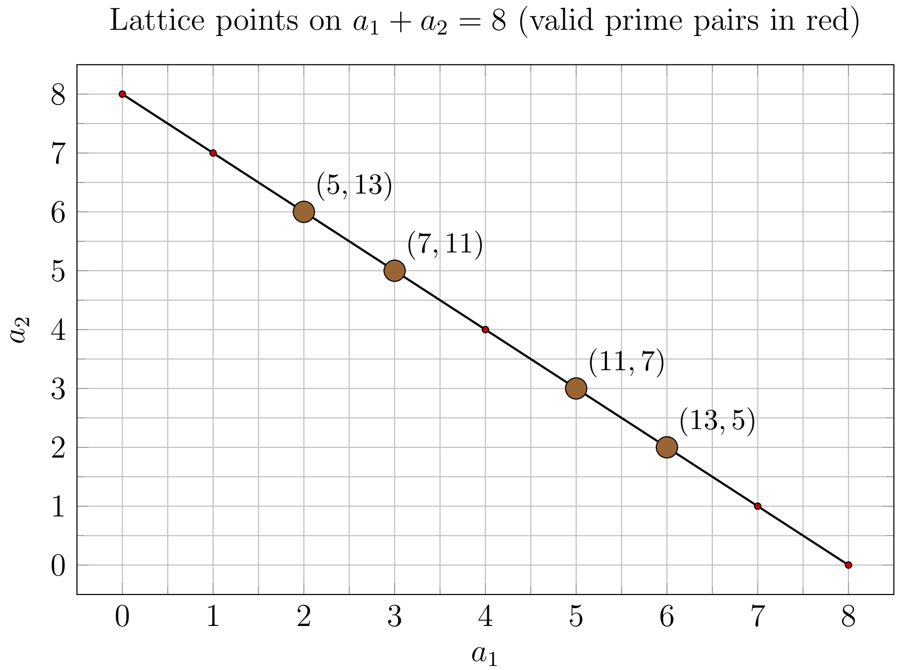

Figure 1: Integer lattice points for $a_1 + a_2 = 8$ (corresponding to $N = 18$). Red markers indicate pairs where both $2a_i + 1$ are prime.

For $k = 2$ this recovers (up to constants) the Hardy–Littlewood heuristic:

$$E[G_2(N)] \asymp \frac{N}{(\log N)^2},$$

with the precise constant given by the Hardy–Littlewood singular series.

This analytic viewpoint explains why combinatorially abundant candidate tuples are expected (on heuristic grounds) to include prime-satisfying tuples for large N — yet converting such heuristics into a rigorous proof remains hard because of local obstructions and correlation effects.

7 Primality filter and core difficulty

The algebraic/combinatorial transformation reduces the additive structure to a simple linear Diophantine constraint, but the multiplicative primality filter on each linear form $2a_i + 1$ is the central, non-linear obstacle. In other words:

- There are many candidate lattice points (stars-and-bars), but
- Primality is a global multiplicative condition that cannot be reduced to simple additive combinatorics.

Bridging this gap is exactly where analytic number theory (sieves, Hardy–Littlewood methods, exponential sums) and computational verification operate.

8 Future directions

1. **Computational exploration.** Enumerate k-tuples for large N and empirically study $G_k(N)$ (counts of prime-valid tuples), comparing to Hardy–Littlewood style predictions.

2. **Sieve-theoretic translation.** Translate Chen-type and Brun-type results into the a-variable framework to obtain intermediate results (e.g., decompositions with almost-primes).

3. **Modular obstruction analysis.** Study congruence classes of a_i modulo small primes to classify impossible tuples and refine counting heuristics.

4. **Higher-dimensional visualizations.** Produce density heatmaps and random slices of the simplex $\sum a_i = M$ colored by the number of linear forms that are prime or almost-prime.

5. **Rigorous asymptotics.** Attempt to bound $G_k(N)$ from below for large N under plausible hypotheses (e.g., GRH or variants) or by leveraging modern sieve machinery.

9 Conclusion

The $P = 2a + 1$ parameterization gives a clean algebraic picture of Goldbach-type problems: candidate decompositions correspond to integer lattice points on a simple hyperplane, and primality acts as a multiplicative sieve selecting a sparse subset. This reframing clarifies the dual nature of the problem (combinatorial abundance vs. multiplicative scarcity) and supplies a convenient coordinate system for both computational experiments and analytic approximations. Integrating partition/stars-and-bars counting with sieve theory and additive combinatorics appears a promising way to organize further work.

References

[1] T. Oliveira e Silva, S. Herzog, and S. Pardi, *Empirical verification of the Goldbach Conjecture and computation of Goldbach partitions up to* 4×10^{18}, Mathematics of Computation, vol. 83, no. 287, pp. 1521–1551, 2014.

[2] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum'; III: On the expression of a number as a sum of primes*, Acta Mathematica, vol. 44, pp. 1–70, 1923.

[3] J.-R. Chen, *On the representation of a large even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica, vol. 16, pp. 157–176, 1973.