THE BINARY GOLDBACH CONJECTURE: A PROOF FOR THE EXISTENCE OF PRIME SUMS FOR ALL 2N

JOSEPH DISE

DEFINITION 1

All elements of the set of odd integers from 3 to 2N-3 – set M – are either Composite (C) or Prime (P).

DEFINITION 2

The paired sums of elements can be of three types:

1) Prime + Prime
2) Prime + Composite
3) Composite + Composite

DEFINITION 3

Let \( M \) = the number of elements in set \( M \).
Let \( Y \) = the number of P elements in set \( M \).
\[ \therefore M - Y = \text{the number of C elements in set } M. \]
Let \( X \) = the number of C elements that form C+C pairs.

ARGUMENT

Proposition A. If there are no P+P paired elements, then all paired elements form \( X \) (C+C) and \( 2Y \) (P+C).

If conjecture is false:

\[ M = X + 2Y \]

Proposition B. If there are P+P paired elements, then \( 2Y \) is greater than the number of P+P and P+C elements.

If conjecture is true:

\[ M < X + 2Y \]

\[ \therefore Proving \ X > M - 2Y \text{ for all } 2N \text{ proves the conjecture.} \]

Lemma 1: The factorization of 2N affects the rate of composite pairing – the more highly composite 2N is, the greater the proportion of C+C pairs – but a calculation can be made for a minimum \( X \) in all cases, regardless of the specific factorization of 2N.

Lemma 2: We can itemize C elements by their least prime factor; least prime factors are bounded by \( \sqrt{2N} \).
For large $2N$:

The proportion of each least prime factor composite is:

$$C_p \approx \frac{1}{p_n} \times \left( \frac{p_{n-1} - 1}{p_{n-1}} \right) \left( \frac{p_{n-2} - 1}{p_{n-2}} \right) \cdots \left( \frac{p_2 - 1}{p_2} \right) = \frac{1}{p_n} \times \prod_{2}^{n-1} \left( \frac{p_n - 1}{p_n} \right)$$

The proportion of remaining elements (primes, and composite elements that have a greater least prime factor) is:

$$R_p \approx \frac{p_n - 1}{p_n} \times \prod_{2}^{n} \left( \frac{p_n - 1}{p_n} \right) = \prod_{2}^{n} \left( \frac{p_n - 1}{p_n} \right)$$

i.e., where $C_3 \propto \frac{1}{3}$, $R_3 \propto \frac{2}{3}$; for all $C_p \propto \frac{1}{p}$, $R_p \propto \frac{p-1}{p}$

Subtracting $Y$ from the $M \times R_p$ ratio gives the approximation for $C$ elements with a least prime factor greater than a given $p$:

$$G_p \approx [M \times R_p] - Y = \left[ M \times \prod_{2}^{n} \left( \frac{p_n - 1}{p_n} \right) \right] - Y$$

**Lemma 3**: For all composites as sorted by their least prime factor, $G_p$ is the approximate number of coprime composites that are available to pair with $M \times C_p$ elements (for all $p$ coprime with $2N$). That is, $|G_p|$ elements will pair with $|M \times C_p|$ elements in a specific ratio, given by $F_p$:

$$F_p \approx 2 \times \frac{1}{p_n - 1} \times \left( \frac{p_{n-1} - 2}{p_{n-1} - 1} \right) \left( \frac{p_{n-2} - 2}{p_{n-2} - 1} \right) \cdots \left( \frac{p_2 - 2}{p_2 - 1} \right) = \frac{2}{p_n - 1} \times \prod_{2}^{n-1} \left( \frac{p_n - 2}{p_n - 1} \right)$$

Combining the $G_p$ elements with their $F_p$ pairing ratios gives:

$$X_p \geq G_p \times F_p = M[R_p \times F_p] - Y[F_p]$$

The total $X$ is the sum of all composite pairings to $p_n$.

$$X \equiv \sum_{2}^{n} X_{p_n} = M \sum_{2}^{n} R_{p_n} F_{p_n} - Y \sum_{2}^{n} F_{p_n}$$

This gives a minimum $X$ in terms of $M$ and $Y$. Since proving $X > M - 2Y$ proves the conjecture, we have a path to determining its validity.

Let the $M$ coefficient for the $X$ sum be $s$.
Let the $Y$ coefficient for the $X$ sum be $g$.

To test whether $X > M - 2Y$, then:

$$sM - gY > M - 2Y$$

$$Y(2 - g) > M(1 - s)$$
Dividing the M coefficient gives:
\[ Y(2 - g)/(1 - s) > M \]
This is identical with:
\[ Y \times \prod_{i=1}^{n} \frac{p_n}{p_n - 1} > M \]

Given by the Prime Number Theorem, \( Y \approx \frac{2M}{\ln 2M} \) for large \( M \), we can define the inequality as:
\[ \frac{2M}{\ln 2M} \times \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \gg M \]

NB: As the number of primes is always larger than the PNT estimate, using the PNT approximation on the greater side of the inequality will remain logically consistent for any size \( M \).

Thus:
\[ \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \gg \frac{\ln 2M}{2} \]

And since \( 2M \approx p_n^2 \), the inequality can be formalized as:
\[ \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \gg \ln p_n \]

This reduces the logic to a single variable. If the above inequality holds for all \( p_n \), the conjecture is true.

TO INFINITY

The slope and spacing of the functions at both average and maximum \( p_{n+1} - p_n \) gaps, or any sequence of gaps, indicate whether the \( \ln p_n \) curve could overtake the \( \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \) curve. Both the derivatives and antiderivatives of each curve indicate infinite divergence. For relatively large individual gaps between primes, the absolute increase in \( \ln p_n \) is greater than the absolute increase in \( \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \), so at some points – that is, over some \( p_{n+1} - p_n \) – the gap is reduced. However, the average growth of the curves along the average gap of primes ensures that the curves diverge. This is because:

a) the net absolute gap in the curves, at every point, requires a prime gap or sequence of primes gaps several multiples of \( p_n \) for the \( \ln p_n \) curve to increase above the \( \prod_{i=1}^{n} \frac{p_n}{p_n - 1} \) curve;

b) individual \( p_{n+1} - p_n \) gaps are significantly smaller than \( p_n \) for large \( p \); and

c) the average prime gap of \( \ln p_n \) is small relative to \( p_n \), such that it ensures a net increase in the gap between the curves over any \( p_n \) increase.

This ensures \( X > M - 2Y \) for all \( 2N \), and the conjecture is true to infinity.