Goldbach's conjecture

Ralf Wüsthofen [1]

March 2, 2014 [2]

ABSTRACT.

This paper presents an elementary proof of the strong Goldbach conjecture. We show that the conjecture is the consequence of an appropriate structuring of the natural numbers. First, we choose an equivalent but more convenient form of the conjecture. Then, in view of this form, we create a structure for the natural numbers. Finally, we derive a distribution property, induced by that structure, and we realize that this has crucial impact on the conjecture. As an additional result of that property we are even able to prove a strengthened form of the conjecture.

CLAIM.

Strong Goldbach conjecture (SGB): *Every even integer greater than 2 can be expressed as the sum of two primes.*

Moreover (SSGB): *Every even integer greater than 6 can be expressed as the sum of two different primes.*

PROOF.

We replace SGB and SSGB by the following equivalent representations:

Every integer greater than 1 is prime or is the arithmetic mean of two different primes p_1 *and* p_2*.*

and

Every integer greater than 3 is the arithmetic mean of two different primes p_1 *and* p_2*.*

---

[1] rwuesthofen@gmail.com

[2] first submission to the Annals of Mathematics on March 24, 2013

SGB $\Leftrightarrow \forall n \in \mathbb{N}$, $n > 1$: (n prime $\vee \exists p_1, p_2$ prime and $\exists d \in \mathbb{N}$, $d > 0$, with: $p_1 + d = n = p_2 - d$) **(1)**

SSGB $\Leftrightarrow \forall n \in \mathbb{N}$, $n > 3$: ($\exists p_1, p_2$ prime and $\exists d \in \mathbb{N}$, $d > 0$, with: $p_1 + d = n = p_2 - d$) **(2)**

The following proof consists of a trivial Part A and a non-trivial Part B. Part A describes a structuring of the natural numbers; Part B derives a distribution property, based on the results of Part A, and proves the conjecture.

Part A

We create all triples $(t_1, t_2, t_3) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}$ of the form

$(t_1, t_2, t_3) = (p\cdot k, m\cdot k, q\cdot k)$, where:

$k \in \mathbb{N}$, $k \geq 1$; $3 \leq p, q$ primes, $p < q$, and $m \in \mathbb{N}$ defined by $m = (p + q)/2$.

Lemma. With the above notations, we have: Every $n \in \mathbb{N}$, $n \geq 3$, can be represented by at least one t_i : $\forall n \in \mathbb{N}$, $n \geq 3$, $\exists t_i$, $1 \leq i \leq 3$, such that $n = t_i$.

Proof. If n is a prime number, we have with $k = 1$: $n = t_1$ or $n = t_3$.

If n is a composite number, the following applies:

For all n different from powers of 2, $n = t_1$ or $n = t_3$ is always a possible representation using a prime decomposition $n = p\cdot k$ or $n = q\cdot k$ with prime factors $3 \leq p < q$. In this case, all infinitely many odd primes are needed.

For all powers of 2 we choose the representation $n = t_2$ with $p = 3$, $m = 4$, $q = 5$ and $k = 1, 2, 4, 8, 16, \ldots$.

□

Let $\mathbb{N}_3$ denote the natural numbers starting from 3. Based on the above result, in case of $n = p\cdot k$ we choose an arbitrary prime q with $q > p$ and in case of $n = q\cdot k$ we choose an arbitrary prime p with $3 \leq p < q$. By obtaining the triples

$(n = p\cdot k, m\cdot k, q\cdot k)$, $(p\cdot k, m\cdot k, n = q\cdot k)$, $(3\cdot k, n = 4\cdot k, 5\cdot k)$,

we say that $\mathbb{N}_3$ is structured by means of all $(t_1, t_2, t_3) = (p\cdot k, m\cdot k, q\cdot k)$. Since each element of $\mathbb{N}_3$ appears as the component of at least one triple, we say that $\mathbb{N}_3$ is entirely covered by the structure $(p\cdot k, m\cdot k, q\cdot k)$. We denote this characteristic by **(I)**.

This structure is unique in the following sense:

We call triples equidistant when the successive components are equidistant. We then consider equidistant triples of the form (t_1, t_2, t_3) = (x_1·k, x_2·k, x_3·k) with integers k ≥ 1 and x_i ≥ 3, where for each i, 1 ≤ i ≤ 3, either all x_i are primes or all x_i are composites. Then, due to the initial triple (3, 4, 5), the only equidistant triples (x_1·k, x_2·k, x_3·k), which cover $\mathbb{N}_3$ completely, are those where x_1 and x_3 are odd primes and x_2 are composites. As is easily verified, all other combinations (x_1, x_2, x_3) of primes and composites are not possible. For example, in case of x_1 and x_2 primes and in case of x_2 and x_3 primes, the number 4 would be missing; in case of x_1 and x_3 composites, the number 3 would be missing.

So, the uniqueness of our structure which is represented by all triples (p·k, m·k, q·k) means the following: If we allow the coefficients x_i in (x_1·k, x_2·k, x_3·k) to be either primes or composites, the only equidistant triple form that covers $\mathbb{N}_3$ is (p·k, m·k, q·k) with primes p, q and composites m. We denote this characteristic by **(II)**.

Actually, for an entire covering of $\mathbb{N}_3$ it would be sufficient if we chose (3·k, 4·k, 5·k) together with triples (p·k, m·k, q·k) in which all other odd primes occur in any way. But for our purpose we use the structure that contains all pairs (p, q) of odd primes with p < q. We call this the maximality of the structure and denote this characteristic by **(III)**.

The following example (n = 42 = 2·3·7) illustrates the redundancies in that structure (incomplete):

(**42**, 54, 66) = (**7·6**, 9·6, 11·6)

(18, 30, **42**) = (3·6, 5·6, **7·6**)

(**42**, 70, 98) = (**3·14**, 5·14, 7·14)

(33, **42**, 51) = (11·3, **14·3**, 17·3)

(41, **42**, 43) = (41·1, **42·1**, 43·1)

(37, **42**, 47) = (37·1, **42·1**, 47·1)

Note: SGB is equivalent to saying that for composite numbers n there is always the redundant representation n = t_2 with k = 1.

Part B

After the role of k as a multiplier generating composite numbers, in this part we will identify an additional meaning of the numbers k. This dual role of k in the triple representation is a key point in the proof.

For each $k \in \mathbb{N}$, $k \geq 1$, we define

$M(k) := \{ n \in \mathbb{N} \mid n = p\cdot k \vee n = m\cdot k \vee n = q\cdot k;\ 3 \leq p < q \text{ primes},\ m = (p + q)/2 \}$.

So, in Part A we have shown that $\mathbb{N}_3 = \bigcup_{k \geq 1} M(k)$.

Note: For different numbers k_1 and k_2, the intersection $M(k_1) \cap M(k_2)$ is equal to the redundancies between the triples $(p\cdot k_1, m\cdot k_1, q\cdot k_1)$ and $(p\cdot k_2, m\cdot k_2, q\cdot k_2)$.

For the moment, we cannot say that for a fixed k ≥ 1 all multiples x·k, x ≥ 3, are contained in a single M(k). However, due to the maximality of our structure, we can say that the components in the triples (p·k, m·k, q·k) are potential candidates for this, in contrast for example to the structure in which are used only pairs of consecutive primes p' and q'. In this case, the triples (p'·k, m'·k, q'·k) also cover $\mathbb{N}_3$ completely, but infinitely many multiples of k are obviously not contained in the corresponding M'(k).

We will see now that this quality, i.e. the existence of all multiples x·k, x ≥ 3, in a single M(k), arises as an effect from the interaction of all M(k), k ≥ 1, the so-called principle of 'emergence'.

We achieve this by showing that the complete covering of $\mathbb{N}_3$ with the maximal structure, using all pairs (p, q) of odd primes with p < q in the triples (p·k, m·k, q·k), results in a forced distribution property for each number k and its corresponding prime factorization.

Specifically, the results from Part A

(I) $\mathbb{N}_3 = \bigcup_{k \geq 1} M(k)$ is covered by the triples (p·k, m·k, q·k)

and

(II) (p·k, m·k, q·k) is the unique equidistant triple form $(x_1\cdot k, x_2\cdot k, x_3\cdot k)$ that covers $\mathbb{N}_3$

and

(III) All pairs (p, q) of odd primes with p < q are used in the triples (p·k, m·k, q·k) for each k ≥ 1

imply the following distribution pattern for the prime factors of any fixed k within the structured $\mathbb{N}_3$:

In its unique prime factorization k is always being extended by any two additional odd prime factors p, q and by an extension m as their arithmetic mean, being either a prime extension too or a composite extension. Exclusively these two types of equidistant triples (p·k, m·k, q·k) determine the distribution of all x·k, x ≥ 3, among themselves.

Why the implication above is true:

We abandon the idea of having to find for a given number n two primes whose arithmetic mean is equal to n respectively whose sum is equal to 2·n. Instead, we identify the natural numbers starting from 3 – apart from the powers of 2 – redundantly in pairs of the form (p·k, q·k), where all combinations of odd primes p, q with p < q and integers k > 0 are used.

For each fixed k > 0, the set of all such pairs (p·k, q·k) defines a two-membered pattern for k within $\mathbb{N}_3$. This pattern, that covers $\mathbb{N}_3 \setminus$ { powers of 2 } when using all k, says that k in its unique prime factorization is always being extended by any two additional odd prime factors. The parameters which exclusively determine this two-membered pattern are the pairs (p, q) of odd primes with p < q. Hence, by applying to each such pair a fixed formula f whose result f(p, q) is an element in $\mathbb{N}_3$ we get a three-membered pattern for each k within $\mathbb{N}_3$ by means of (p·k, q·k, f(p, q)·k).

If f moreover satisfies the condition that all (p, q, f(p, q)) have the same numerical ordering and that the two distances of successive components have a constant difference, for each fixed k the triples (p·k, q·k, f(p, q)·k) distribute their components in a uniform way. In that case, the two-membered pattern for each k is being expanded into a three-membered distribution pattern for each k within $\mathbb{N}_3$ by means of (p·k, q·k, f(p, q)·k).

In our case, the formula with that condition is the arithmetic mean, so that the pairs (p·k, q·k) are expanded into triples (p·k, m·k, q·k) including also the powers of 2 by (3·k, 4·k, 5·k). Thereby, the two-membered pattern for each k is expanded into a three-membered distribution pattern that covers the whole $\mathbb{N}_3$ by using all k. This three-membered distribution pattern is the underlined above.

According to **(1)**, we now choose an integer n > 1 which is not prime. We then consider the multiple n·k for any k > 1, i.e. n is a composite extension of k. In order to fulfill the underlined distribution pattern of k in $\mathbb{N}_3$, n·k must be the middle component of a triple (p·k, m·k, q·k). This proves SGB.

According to **(2)**, we now choose a prime n > 3. Obviously, n lies between two other odd primes. We then consider the multiple n·k for any k > 1, i.e. n is a prime extension of k. In order to fulfill the underlined distribution pattern of k in $\mathbb{N}_3$, n·k must be the middle component of a triple (p·k, m·k, q·k). This, together with SGB, proves SSGB.

□

Note: If we additionally allow the values p = 1 and q = 3, we have that all multiples of any natural number k occur exclusively in the symmetric form (p·k, m·k, q·k).

EXAMPLES.

n = 14 and k = 3:

Let us assume that n is not the arithmetic mean of two primes. For n·k = 42, we find for example (p·k', m·k', q·k') = (3·6, 5·6, **7·6**), which is part of the distribution of 6 in the structure. But there is no triple (p·3, m·3, q·3) that contains n·3. Thus, n·3 violates the distribution of 3 in the structure.

This contradiction can be resolved only if n = m, that is, n must be arithmetic mean of, for example, 11 and 17.

n = 9 and k = 3:

Let us assume that n is not the arithmetic mean of two primes. For n·k = 27, we only find (p·k', m·k', q·k') = (**3·9**, m·9, q·9), which is part of the distribution of 9 in the structure. But there is no triple (p·3, m·3, q·3) that contains n·3. Thus, n·3 violates the distribution of 3 in the structure.

This contradiction can be resolved only if n = m, that is, n must be arithmetic mean of, for example, 7 and 11.

n = 19 and k = 3:

Let us assume that n is not the arithmetic mean of two primes. For n·k = 57, we find for example (p'·k, m'·k, q'·k) = (17·3, 18·3, **19·3**), which is part of the distribution of 3 in the structure. But there is no triple (p·3, m·3, q·3) with p < 19 < q that contains n·3. Thus, n·3 violates the distribution of 3 in the structure.

This contradiction can be resolved only if n = m, that is, n must be arithmetic mean of 7 and 31.

REMARKS.

a. The statement in the binary Goldbach conjecture is actually nothing more than the symmetric structure (p·k, m·k, q·k) used in the proof. As we have shown, in fact it is a specific case of a general distribution principle within the natural numbers.

In order to get rid of the usual interpretation of the conjecture that focuses on the sums of primes and thus opposes their multiplicative character, we have tackled the problem differently after reformulating to the triple form: Instead of searching for primes which build the needed arithmetic mean equal to a given n, we have made an access from the reverse direction. Based on the multiplicative prime decomposition, we realize n·k as the result of a fixed formula (i.e. the arithmetic mean) applied to the pattern (p·k, q·k).

A key point in the proof is the dual role of the numbers k: on the one hand as multiplier they generate composite numbers, on the other hand their own occurrence within $\mathbb{N}$ is strictly set by the used triple.

While in a single M(k) the existence of all multiples x·k, x ≥ 3, for k is not guaranteed, the union of all M(k), k ≥ 1, determines these multiples for any fixed k by means of (p·k, m·k, q·k), because that union is equal to the whole $\mathbb{N}_3$. In other subjects, this effect of the formation of new properties after the transition of single items to a whole system is called emergence. ("*The whole is more than the sum of its parts.*")

b. Due to the unpredictable nature of the distribution of the primes – especially their incidence within certain intervals – all studies on the representation of natural numbers as the sum of primes are problematic when they use approaches based on that distribution.

   Despite tremendous efforts over the centuries, the best result so far was five summands. I was always convinced that the solution must lie in the constructive characteristics of the prime numbers and not in their distribution (see more on this in the next section).

c. Using an appropriate model for the environment in which the problem is given, also other arithmetic questions in number theory can possibly be solved in an elementary way (unless there is no solution on the basis of the underlying axiom system).

GOLDBACH AND THE UNEXPECTED SIMPLICITY.

In the course of the attempts to solve the strong and the weak Goldbach conjecture – both formulated by Goldbach and Euler in their correspondence in 1742 – was taken in fact a gigantic wrong route by focusing exclusively on the additive character of the statements.

After a proof has not even been achieved for three prime summands (the weak conjecture for odd numbers) without additional assumptions, in the twenties of the last century mathematicians began to search for a maximum number of primes which is necessary to represent any natural number greater than 1 as their sum. At the beginning, there were proofs that required hundreds of thousands (!) of primes (L. Schnirelmann). The weak conjecture could then be proved in 1937 (I. Vinogradov), however only above a very large constant so that enough primes are available as summands.

Then it took almost another century until being able to reduce the sum representation for all integers > 1 to five respectively six summands of primes (T. Tao). Ironically, in May 2013, was then claimed the closing of the huge gap of numbers for the weak Goldbach version, using numerical verification combined with an extremely complex estimative proof (H. Helfgott).

The so-called Hardy-Littlewood circle method, which was employed and constantly improved in those approaches, has ignored the actual role of the primes in the context of the originally formulated problem of Goldbach and Euler, by examining permanently "how many" prime numbers are available as summands. As this method does not work for the binary Goldbach conjecture, the concern for this original problem has been pushed into the background up to the present time. And this in spite of the fact, that a solution would have concluded the issue of sum representations by primes in a fundamental way.

In the eighteenth century it was quite normal that a brilliant mathematician answered also less brilliant ones and so Euler wrote in his response to Goldbach on 30 June 1742, he was convinced of the correctness of the conjecture. ("*Dass … ein jeder numerus par eine summa duorum primorum sey, halte ich für ein ganz gewisses theorema, ungeachtet ich dasselbe nicht demonstriren kann.*")

It is possible that his thoughts were not far away from the structural interpretation of the conjecture as shown in the present paper – certainly they were very far away from the methods initiated two hundred years later.