# Goldbach conjecture, rewriting, contradiction

Denise Vella-Chemla

March 30, 2014

## 1 16 rewriting rules

We remind that we choosed to represent that an integer is prime by boolean 0 and the fact that it is compound by boolean 1.

We also decided to use the following conventions($3 \leqslant p \leqslant n/2$) :

- $a$ letter symbolizes an $n$ decomposition of the form $p+q$ with $p$ and $q$ primes ;
- $b$ letter symbolizes an $n$ decomposition of the form $p+q$ with $p$ compound and $q$ prime ;
- $c$ letter symbolizes an $n$ decomposition of the form $p+q$ with $p$ prime and $q$ compound ;
- $d$ letter symbolizes an $n$ decomposition of the form $p+q$ with $p$ and $q$ compound.

$a$ letter codes $\begin{pmatrix}0\\0\end{pmatrix}$ matrix, and respectively $b$ $\begin{pmatrix}0\\1\end{pmatrix}$, $c$ $\begin{pmatrix}1\\0\end{pmatrix}$ and finally $d$ $\begin{pmatrix}1\\1\end{pmatrix}$

**Example :** Hereafter the word $m_{abcd}(40)$.

| 40 | 37 | 35 | 33 | 31 | 29 | 27 | 25 | 23 | 21 |
|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 |
| $m_{abcd}(40)$ | $a$ | $c$ | $c$ | $b$ | $a$ | $c$ | $d$ | $a$ | $c$ |

In the following, we use the operation on matrices defined as :

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}.\begin{pmatrix}y_1\\y_2\end{pmatrix}=\begin{pmatrix}x_1\\y_2\end{pmatrix}$$

Our operation provides 16 rewriting rules of couples of letters, that seem relevant to study Goldbach conjecture :

| | | | |
|---|---|---|---|
| 1) $aa \to a$ | 5) $ba \to a$ | 9) $ca \to c$ | 13) $da \to c$ |
| 2) $ab \to b$ | 6) $bb \to b$ | 10) $cb \to d$ | 14) $db \to d$ |
| 3) $ac \to a$ | 7) $bc \to a$ | 11) $cc \to c$ | 15) $dc \to c$ |
| 4) $ad \to b$ | 8) $bd \to b$ | 12) $cd \to d$ | 16) $dd \to d$ |

## 2 Reminders from language theory

An alphabet is a finite set of symbols.

Alphabets used in the following are : $A = \{a, b, c, d\}$, $A_{ab} = \{a, b\}$, $A_{cd} = \{c, d\}$, $A_{ac} = \{a, c\}$ and $A_{bd} = \{b, d\}$.

A word on $X$ alphabet is a finite and ordered sequence, eventually empty, of alphabet elements. It's a letters concatenation. We note $X^*$ the set of words over $X$ alphabet.

A word is called a prefix of another one if it contains, on all its length, the same letters as it at same positions ($X$ being an alphabet and $w, u \in X^*$. $u$ is a prefix of $w$ if and only if $\exists v \in X^*$ such that $w = u.v$).

## 3 Observing words

Let us observe words associated with even numbers between 6 and 80.

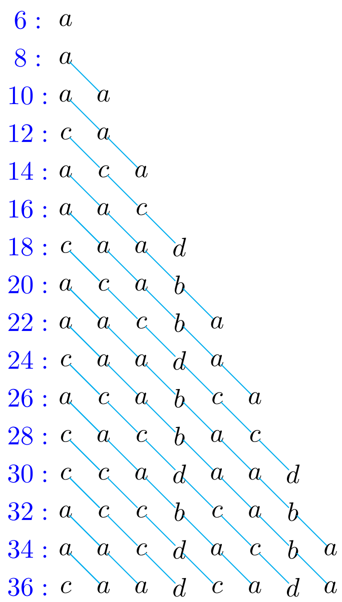


FIGURE 1 : words associated to even numbers between 6 and 36

We see that words in diagonals contain either $a$ and $b$ letters exclusively, or $c$ and $d$ letters exclusively.

# 4 Words properties

Diagonal words (diagonals) have their letters either in $A_{ab}$ alphabet or in $A_{cd}$ alphabet.

Every diagonal is a prefix of the following one that is defined on the same alphabet.

Indeed, a diagonal code decompositions that have the same second sommant and that have a first sommant that is an odd number from the list of successive odd numbers beginning at 3.

For instance, diagonal *aaabaa* beginning with first letter of 26's word on figure 1 code the following decompositions : 3+23, 5+23, 7+23, 9+23, 11+23 and 13+23.

Thus, diagonals on $A_{ab}$ alphabet "code" decompositions that have a same second sommant which is prime ; their letters code either by $a$ letters corresponding to prime numbers, or by $b$ letters corresponding to compound ones the primality characters of odd numbers (the first sommants), beginning at 3.

Diagonals on $A_{cd}$ alphabet "code" on their side decompositions that have a same second sommant which is compound ; their letters code either by $c$ letters corresponding to prime numbers, or by $d$ letters corresponding to compound ones the primality characters of odd numbers (the first sommants), beginning at 3.

Vertical words have their letters either in $A_{ac}$ alphabet or in $A_{bd}$ alphabet. A vertical word code decompositions that have same first sommant. Every vertical word is contained in a vertical word that is "on its left side" and that is defined on the same alphabet.

# 5 Some regularities

We observe some regularities easily explanables, that link together letters numbers of each kind that appear in an even number word or in a certain portion of first column of letters. Figure 2 above presents schematically variable names that will be useful to conduct reasoning :

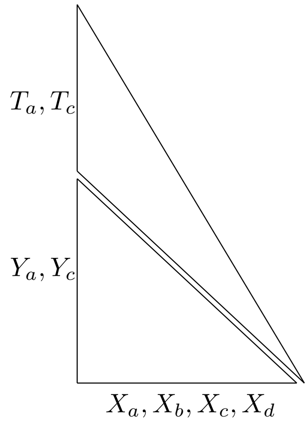


Figure 2 : linked variables

Global triangle contains words associated to even numbers from 6 to $n$.

$X_a, X_b, X_c$ et $X_d$ count $a, b, c$ or $d$ letters numbers in $n$ word.

$T_a$ and $T_c$ count numbers of $a$ or $c$ letters that are first letters of words associated with even numbers between 6 and $2\left\lceil\dfrac{n+2}{4}\right\rceil$.

$T_a$ thus corresponds to decompositions of the form $n' = 3 + p_i$, $p_i$ prime, $n' \leqslant 2\left\lceil\dfrac{n+2}{4}\right\rceil$. For instance, if $n = 34$, $T_a = \#\{3+3, 3+5, 3+7, 3+11, 3+13\}$.

$T_a$ thus corresponds to decompositions of the form $n' = p_i + p_i$, $p_i$ prime for $n' < n$. For instance, if $n = 34$, $T_a = \#\{3+3, 5+5, 7+7, 11+11, 13+13\}$.

$T_c$ corresponds to decompositions of the form $n' = 3 + c_i$, $c_i$ compound $n' \leqslant 2\left\lceil\dfrac{n+2}{4}\right\rceil$. For instance, if $n = 34$, $T_c = \#\{9+9, 15+15\}$.

$Y_a$ and $Y_c$ count numbers of $a$ or $c$ letters that are first letters of words associated to even numbers between $2\left\lceil\dfrac{n+2}{4}\right\rceil + 2$ and $n$.

The trivial one-to-one mapping on the decompositions second term permits to explain easily why $Y_a = X_a + X_b$ or $Y_c = X_c + X_d$. The simple reading of sets defined in extension suffices to convince oneself.

$$
\begin{aligned}
Y_a &= \#\{3+17, 3+19, 3+23, 3+29, 3+31\} \\
X_a &= \#\{3+31, 5+29, 11+23, 17+17\} \\
X_b &= \#\{15+19\}
\end{aligned}
$$

$$
\begin{aligned}
Y_c &= \#\{3+21, 3+25, 3+27\} \\
X_c &= \#\{7+27, 13+21\} \\
X_d &= \#\{9+25\}
\end{aligned}
$$

Hereafter two figures to “fix ideas” for even numbers $n = 32$ or $n = 34$.

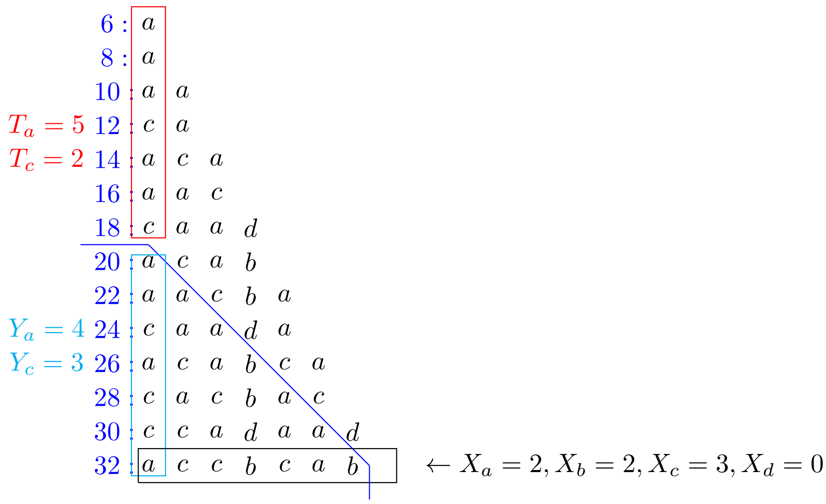


FIGURE 3 : premier exemple : $n = 32$

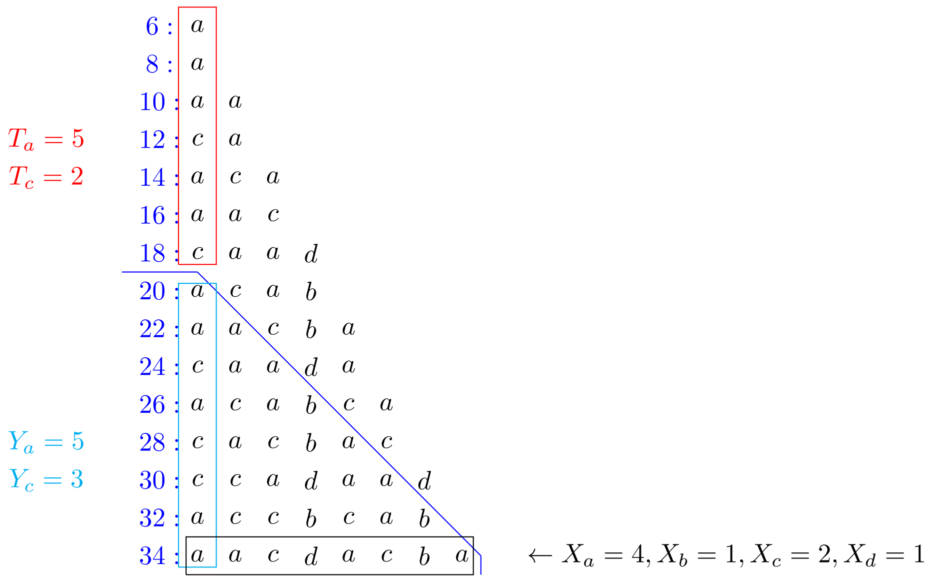


FIGURE 4 : second example : $n = 34$

Following constraints are always satisfied :

$$
\begin{aligned}
&Y_a = X_a + X_b \\
&Y_c = X_c + X_d \\
&T_a + T_c + Y_a + Y_c + \epsilon = 2(X_a + X_b + X_c + X_d)
\end{aligned}
$$

$\epsilon = 1$ if $n$ is an odd double, $\epsilon = 0$ otherwise.

We saw that constraints above can be easily understood if we come back to numbers of decompositions they represent and if we use “Cantor-like” one-to-one mappings.

Different words letters are thus very *intricated* and those intrications have as consequence that every word contains at least one $a$ letter.

We are going to prove this using a *reductio ad absurdum* reasoning in section 7.

# 6 Cantor one-to-one mappings visualization

We provide above Cantor one-to-one mappings for cases $n = 32$ and $n = 34$.

One-to-one mapping $f$ that permits to pass from line 2 to line 1 is such that $f(a) = f(c) = a$ et $f(b) = f(d) = c$.

One-to-one mapping $g$ that permits to pass from line 2 to line 3 is such that $g(a) = g(b) = a$ et $g(c) = g(d) = c$.

It's the double decomposition $3 + n/2$ in the case where $n$ is an odd's double that necessitates introduction of the variable $\epsilon$ that is equal to 1 in that case and equal to 0 otherwise.

– One-to-one mappings if $n = 32$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 1 | $a$ | $a$ | $a$ | $c$ | $a$ | $a$ | $c$ |
| | 3 | 5 | 7 | 9 | 11 | 13 | 15 |
| | 3 | 5 | 7 | 9 | 11 | 13 | 15 |
| 2 | $a$ | $c$ | $c$ | $b$ | $c$ | $a$ | $b$ |
| | 29 | 27 | 25 | 23 | 21 | 19 | 17 |
| | 29 | 27 | 25 | 23 | 21 | 19 | 17 |
| 3 | $a$ | $c$ | $c$ | $a$ | $c$ | $a$ | $a$ |
| | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

– One-to-one mappings if $n = 34$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 1 | $a$ | $a$ | $a$ | $c$ | $a$ | $a$ | $c$ | $a$ |
| | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 |
| | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 |
| 2 | $a$ | $a$ | $c$ | $d$ | $a$ | $c$ | $b$ | $a$ |
| | 31 | 29 | 27 | 25 | 23 | 21 | 19 | 17 |
| | 31 | 29 | 27 | 25 | 23 | 21 | 19 | 17 |
| 3 | $a$ | $a$ | $c$ | $c$ | $a$ | $c$ | $a$ | $a$ |
| | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

## 7 Looking for a contradiction

Let us imagine that a word $m_n$ is associated to an even number $n$ that contradicts Goldbach conjecture, i.e. $m_n$ doesn't contain any $a$ letter (we remind that $a$ letter symbolizes sum of two primes).

$m_n$ containing no $a$, we have $X_a = 0$. But since $Y_a = X_a + X_b$, we deduce $Y_a = X_b$. Identifying $Y_a$ to $X_b$ and $Y_c$ to $X_c + X_d$ in the last constraint always satisfied provided in the paragraph above, one obtains the following equalities :

$$T_a + T_c + Y_a + Y_c + \epsilon = 2(X_a + X_b + X_c + X_d)$$
$$T_a + T_c + X_b + X_c + X_d + \epsilon = 2X_a + 2X_b + 2X_c + 2X_d$$
$$T_a + T_c + \epsilon = X_b + X_c + X_d$$
$$T_a + T_c + \epsilon = X_b + Y_c$$

We must now remember what those variables represent :

- $T_a + T_c = (n-4)/4$ ;
- $X_b$ counts the number of $n$ decompositions of the form of a sum of two odd numbers $p+q$ with $p \leqslant n/2$ compound and $q$ prime ;
- $Y_c$ counts the number of compound odd numbers between $n/2$ and $n-3$.

Number $X_b$ of $n$ decompositions of the form of a sum of two odd numbers $p+q$ with $p \leqslant n/2$ compound and $q$ prime being necessarily lesser than the number of primes between $n/2$ and $n-3$, we have $X_b < Y_a$ (we used here a sort of inverted "pigeonhole principle" : if one puts 0 or 1 object in $k$ holes, there can't be more objects than holes, i.e. more than $k$ objects). But the number of prime numbers contained in an interval $[2k+3, 4k+1]$ is always lesser than the number of compound odd numbers contained in this interval for $k > 25$. In those cases, $Y_a < Y_c$ and $X_b + Y_c < Y_a + Y_c < 2Y_c$.

But, for all integers greater than a certain small integer (such that 100), $(n-4)/4$ is greater than $2Y_c$. This ensures we never have $T_a + T_c + \epsilon = X_b + Y_c$ that should result from the absence of an $a$ letter in a word.

We reached a contradiction that would be a consequence of the absence of an $a$ letter in a word. This brings the impossibility that an even number contradict Goldbach conjecture. Rewriting rules intricate totally words letters in such a manner that their numbers must mandatory respect certain constraints. This work can be localized in a lexical theory of numbers, according to which numbers are words. This theory relies on the fact that letters order in words is primordial.