Goldbach's conjecture, 4 letters language, variables and invariants

Denise Vella-Chemla

10/5/2014

1 Introduction

Goldbach's conjecture states that each even integer except 2 is the sum of two prime numbers. In the following, one is interested in decompositions of an even number n as a sum of two odd integers $p+q$ with $3 \leqslant p \leqslant n/2$, $n/2 \leqslant q \leqslant n-3$ and $p \leqslant q$. We call p a *n's first range sommant* and q a *n's second range sommant.*

Notations :

We will designate by :

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

Example :

40	3	5	7	9	11	13	15	17	19
	37	35	33	31	29	27	25	23	21
l_{40}	a	c	c	b	a	c	d	a	c

2 Main array

We designate by $T = (L, C) = (\ l_{n,m}\ )$ the array containing $l_{n,m}$ elements that are one of a, b, c, d letters. n belongs to the set of even integers greater than or equal to 6. m, belonging to the set of odd integers greater than or equal to 3, is an element of list of n first range sommants.

Let us consider g function defined by :

$$g: \begin{array}{ccc} 2\mathbb{N} & \rightarrow & 2\mathbb{N}+1 \\ x & \mapsto & 2\left\lfloor \dfrac{x-2}{4} \right\rfloor + 1 \end{array}$$

$g(6) = 3, g(8) = 3, g(10) = 5, g(12) = 5, g(14) = 7, g(16) = 7, etc.$

$g(n)$ function defines the greatest of n first range sommants.

As we only consider n decompositions of the form $p+q$ where $p \leqslant q$, in T will only appear letters $l_{n,m}$ such that $m \leqslant 2\left\lfloor \dfrac{n-2}{4} \right\rfloor + 1$ in such a way that T array first letters are : $l_{6,3}, l_{8,3}, l_{10,3}, l_{10,5}, l_{12,3}, l_{12,5}, l_{14,3}, l_{14,5}, l_{14,7}, etc.$

Here are first lines of array T.

L \ C	3	5	7	9	11	13	15	17
6	a							
8	a							
10	a	a						
12	c	a						
14	a	c	a					
16	a	a	c					
18	c	a	a	d				
20	a	c	a	b				
22	a	a	c	b	a			
24	c	a	a	d	a			
26	a	c	a	b	c	a		
28	c	a	c	b	a	c		
30	c	c	a	d	a	a	d	
32	a	c	c	b	c	a	b	
34	a	a	c	d	a	c	b	a
36	c	a	a	d	c	a	d	a
...								

FIGURE 1 : words of even numbers between 6 and 36

Remarks :

1) words on array's diagonals called *diagonal words* have their letters either in $A_{ab} = \{a, b\}$ alphabet or in $A_{cd} = \{c, d\}$ alphabet.

2) a diagonal word codes decompositions that have the same second range sommant.
For instance, on Figure 4, letters *aaabaa* of the diagonal that begins at letter $l_{26,3} = a$ code decompositions $3 + 23, 5 + 23, 7 + 23, 9 + 23, 11 + 23$ and $13 + 23$.

3) let us designate by l_n the line whose elements are $l_{n,m}$. Line l_n contains $\left\lfloor \dfrac{n-2}{4} \right\rfloor$ elements.

4) n being fixed, let us call $C_{n,3}$ the column formed by $l_{k,3}$ for $6 \leqslant k \leqslant n$.

In this column $C_{n,3}$, let us distinguish two parts, the "top part" and the "bottom part" of the column.

Let us call $H_{n,3}$ column's "top part", i.e. set of $l_{k,3}$ where $6 \leqslant k \leqslant \left\lfloor \dfrac{n+4}{2} \right\rfloor$.

Let us call $B_{n,3}$ column's "bottom part", i.e. set of $l_{k,3}$ where $\left\lfloor \dfrac{n+4}{2} \right\rfloor < k \leqslant n$.

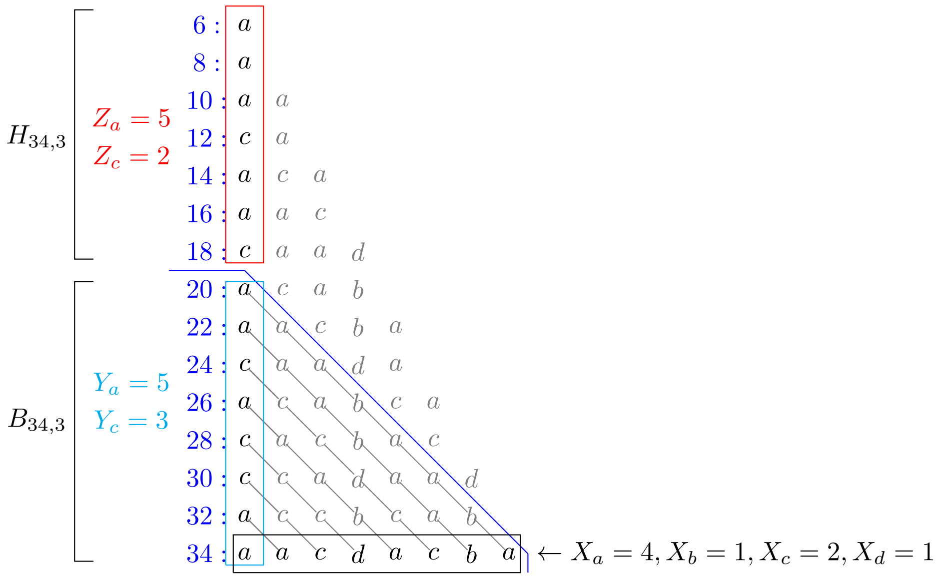


FIGURE 2 : $n = 34$

To better understand countings in next section, we will use projection P of line n on bottom part of first column $B_{n,3}$ that "associates" letters at both extremities of a diagonal. If we consider application $proj$ such that $proj(a) = proj(b) = a$ and $proj(c) = proj(d) = c$ then, since 3 is prime, $proj(l_{n,2k+1}) = l_{n-2k+2,3}$.

We can also understand the effect of this projection (that preserves second range sommant) by analyzing decompositions :

- if $p + q$ is coded by an a or a b letter, it corresponds to two possible cases in which q is prime, and so $3 + q$ decomposition, containing two prime numbers, will be coded by an a letter ;
- if $p + q$ is coded by a c or a d letter, it corresponds to two possible cases in which q is compound, and so $3 + q$ decomposition, of the form $prime + compound$ will be coded by a c letter.

We will also use in next section a projection that transforms first range sommant in a second range sommant that is combined with 3 as a first range sommant ; let us analyze the effect that such a projection will have on decompositions :

- if $p + q$ is coded by an a or a c letter, it corresponds to two possible cases in which p is prime, and so $3 + p$ decomposition, containing two prime numbers, will be coded by an a letter ;
- if $p + q$ is coded by a b or a d letter, it corresponds to two possible cases in which p is compound, and so $3 + p$ decomposition, of the form $prime + compound$ will be coded by a c letter.

3 Computations

1) We note in line n by :

- $X_a(n)$ the number of n decompositions of the form $prime + prime$;
- $X_b(n)$ the number of n decompositions of the form $compound + prime$;
- $X_c(n)$ the number of n decompositions of the form $prime + compound$;
- $X_d(n)$ the number of n decompositions of the form $compound + compound$.

$X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \dfrac{n-2}{4} \right\rfloor$ is the number of elements of line n.

Example $n = 34$:
$X_a(34) = \#\{3 + 31, 5 + 29, 11 + 23, 17 + 17\} = 4$
$X_b(34) = \#\{15 + 19\} = 1.$
$X_c(34) = \#\{7 + 27, 13 + 21\} = 2$
$X_d(34) = \#\{9 + 25\} = 1$

2) Let $Y_a(n)$ (resp. $Y_c(n)$) being the number of a letters (resp. c) that appear in $B_{n,3}$. We recall that there are only a and c letters in first column because it contains letters associated with decompositions of the form $3+x$ and because 3 is prime.

Example :

- $Y_a(34) = \#\{3+17, 3+19, 3+23, 3+29, 3+31\} = 5$
- $Y_c(34) = \#\{3+21, 3+25, 3+27\} = 3$

3) Because of P projection that is a bijection, and because of a, b, c, d letters definitions, $Y_a(n) = X_a(n) + X_b(n)$ and $Y_c(n) = X_c(n) + X_d(n)$. Thus, trivially, $Y_a(n) + Y_c(n) = X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \frac{n-2}{4} \right\rfloor$.

Example :

$$\begin{aligned} Y_a(34) &= \#\{3+17, 3+19, 3+23, 3+29, 3+31\} \\ X_a(34) &= \#\{3+31, 5+29, 11+23, 17+17\} \\ X_b(34) &= \#\{15+19\} \\ \\ Y_c(34) &= \#\{3+21, 3+25, 3+27\} \\ X_c(34) &= \#\{7+27, 13+21\} \\ X_d(34) &= \#\{9+25\} \end{aligned}$$

4) Let $Z_a(n)$ (resp. $Z_c(n)$) being the number of a letters (resp. c) that appear in $H_{n,3}$.

Example :

- $Z_a(34) = \#\{3+3, 3+5, 3+7, 3+11, 3+13\} = 5$
- $Z_c(34) = \#\{3+9, 3+15\} = 2$

$Z_a(n) + Z_c(n) = \left\lfloor \frac{n-4}{4} \right\rfloor$.

Reminding identified properties

$$Y_a(n) = X_a(n) + X_b(n) \tag{1}$$

$$Y_c(n) = X_c(n) + X_d(n) \tag{2}$$

$$Y_a(n) + Y_c(n) = X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \frac{n-2}{4} \right\rfloor \tag{3}$$

$$Z_a(n) + Z_c(n) = \left\lfloor \frac{n-4}{4} \right\rfloor \tag{4}$$

Let us add four new properties to those ones :

$$X_a(n) + X_c(n) = Z_a(n) + \delta_{2p}(n) \tag{5}$$

with $\delta_{2p}(n)$ equal to 1 when n is the double of a prime number and equal to 0 otherwise.

$$X_b(n) + X_d(n) = Z_c(n) + \delta_{c-imp}(n) \tag{6}$$

with $\delta_{2c-imp}(n)$ equal to 1 when n is the double of a compound odd number and equal to 0 otherwise (when there exists k such that $n = 4k$ (doubles of even numbers) or when n is the double of a prime number).

$$Z_c(n) - Y_a(n) = Y_c(n) - Z_a(n) - \delta_{4k+2}(n) \tag{7}$$

with $\delta_{4k+2}(n)$ equal to 1 when n is the double of an odd number and 0 otherwise.

$$Z_c(n) - Y_a(n) = X_d(n) - X_a(n) - \delta_{2c-imp}(n) \tag{8}$$

4 Variables evolution

In this section, let us study how different variables change.

$Z_a(n)+Z_c(n) = \left\lfloor \dfrac{n-4}{4} \right\rfloor$ is an increasing function of n, it is increased by 1 at each n that is an even double.

$Z_a(n+2) = Z_a(n)$ and $Z_c(n+2) = Z_c(n)$ are constant when n is an even double.

$Z_a(n) = Z_a(n-2) + 1$ when $\dfrac{n-2}{2}$ is prime (*ex* : $n = 24$ or $n = 28$, look to values array page 13 : we will express this abusively by *"Z_a is increasing"*) and $Z_c(n) = Z_c(n-2) + 1$ when $\dfrac{n-2}{2}$ is an odd compound number (*ex* : $n = 42$ or 50, abusively, *"Z_c is increasing"*).

$Y_a(n) + Y_c(n) = X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \dfrac{n-2}{4} \right\rfloor$ is an increasing function of n, it is increased by 1 at each n that is an odd double.

Let us see now in details how $Y_a(n)$ and $Y_c(n)$ evoluate.

In case if n is an odd double, a number more is put in $H_{n,3}$; if this number $n-3$ is prime (resp. compound), $Y_a(n) = Y_a(n-2) + 1$ (abusively *"Y_a is increasing"*, *ex* : $n = 34$) (resp. $Y_c(n) = Y_c(n-2) + 1$, abusively *"Y_c is increasing"*, *ex* : $n = 38$).

In case if n is an even double, 4 cases are to be studied. Let us study the way decompositions set $H_{n,3}$ evoluates.

- if $n-3$ and $\dfrac{n-2}{2}$ are both primes, there is a decomposition that is taken out in the bottom and another decomposition that is put in in the top of $H_{n,3}$ and those two decompositions have two letters of the same type, so $Y_a(n) = Y_a(n-2)$ and $Y_c(n) = Y_c(n-2)$ (abusively *"Y_a and Y_c constants"* (*ex* : $n = 40$) ;
- if $n-3$ is prime and $\dfrac{n-2}{2}$ is compound then $Y_a(n) = Y_a(n-2)+1$ and $Y_c(n) = Y_c(n-2)-1$ (abusively, *"Y_a is increasing and Y_c is decreasing"* (*ex* : $n = 32$) ;
- if $n-3$ is compound and $\dfrac{n-2}{2}$ is prime then $Y_c(n) = Y_c(n-2)+1$ and $Y_a(n) = Y_a(n-2)-1$ (abusively, *"Y_a is decreasing and Y_c is increasing"*) (*ex* : $n = 48$) ;
- if $n-3$ and $\dfrac{n-2}{2}$ are both compound, there is a decomposition that is taken out in the bottom and another decomposition that is put in in the top of $H_{n,3}$ and those two decompositions have two letters of the same type, so $Y_a(n) = Y_a(n-2)$ and $Y_c(n) = Y_c(n-2)$ (abusively, *"Y_a and Y_c constants"* (*ex* : $n = 52$).

In annex 1 is provided an array containing different variables values for n between 14 and 100.

5 Use gaps between variables

We are going to show in the following that $X_a(n)$ can never be equal to 0 for $n \geqslant C$, C being a constant to be defined, i.e. to show that each even integer $n \geqslant C$ can be written as a sum of two primes, or in other words verifies Goldbach's conjecture.

We saw that at $\delta_{4k+2}(n)$ and $\delta_{2c-imp}(n)$ near ($\delta_{4k+2}(n)$ and/or $\delta_{2c-imp}(n)$ being equal to 1 in certain cases), we have following equalities :

$$Z_c(n) - Y_a(n) = Y_c(n) - Z_a(n) - \delta_{4k+2}(n) \tag{7}$$

$$Z_c(n) - Y_a(n) = X_d(n) - X_a(n) - \delta_{2c-imp}(n) \tag{8}$$

We remind that

- $Y_a(n)$, counting number of primes that are between $\frac{n}{2}$ and n is equal to $\pi(n) - \pi\left(\frac{n}{2}\right)$;
- $Z_a(n)$ counting number of primes lesser than or equal to $\frac{n}{2}$ is equal to $\pi\left(\frac{n}{2}\right)$;
- $Z_c(n)$ counting number of odd compound numbers lesser than or equal to $\frac{n}{2}$ is equal to $\frac{n}{4} - \pi\left(\frac{n}{2}\right)$;
- $Y_c(n)$ counting number of odd compound numbers that are between $\frac{n}{2}$ and n is equal to $\frac{n}{4} - \pi(n) + \pi\left(\frac{n}{2}\right)$.

Rosser and Schoenfeld [1] note provides formula 3.5 of corollary 1 of theorem 2 that $\pi(x) > \frac{x}{ln\ x}$ for every $x \geqslant 17$, and as formula 3.6 of the same corollary of the same theorem that $\pi(x) < \frac{1,25506x}{ln\ x}$ for every $x > 1$.

5.1 $Z_c(n) > Z_a(n)$ inequality study

To show that $Z_c(n) > Z_a(n)$, one can simply use the fact that $Z_c(n)$ is increasing "many more times" than $Z_a(n)$ (each time $\frac{n-2}{2}$ is an odd compound number for $Z_c(n)$ and each time $\frac{n-2}{2}$ is a prime number for $Z_a(n)$ as it was shown in section 4).

To find from which value of n, $Z_c(n) > Z_a(n)$, one uses the fact that $Z_c(n) - Z_a(n)$ gap is equal to $\frac{n}{4} - 2\,\pi\left(\frac{n}{2}\right)$.

From formula 3.6 of corollary 1 of theorem 2 of [1], we have $2\,\pi\left(\frac{n}{2}\right) < 2\,\frac{1,25506\ n}{2\ (ln\ n + ln\ 0.5)}$ for every $n > 2$.

We deduce from this $-2\,\pi\left(\frac{n}{2}\right) > \frac{-1,25506\ n}{ln\ n + ln\ 0.5}$ for every $n > 2$.

$Z_c(n) - Z_a(n)$ gap is so minorable by $\frac{n\ (ln\ n + ln\ 0.5) - 5,02024n}{4\ (ln\ n + ln\ 0.5)}$. It is strictly greater than 0 for every $n \geqslant 304$ (denominator is greater or equal to 0 for every $n \geqslant 2$, numerator is strictly greater than 0 for every $n > 2e^{5.02024}$).

5.2 $Z_a(n) > Y_a(n)$ and $Y_c(n) > Z_c(n)$ inequalities study

To show that $Z_a(n) > Y_a(n)$, one can use once more the analysis of variables evolution provided in section 4 : when *"Z_a is increasing"*, *"Y_a is constant or is decreasing"* ; and when *"Y_a is increasing"* without *"Z_a is also increasing"* (when $n-3$ is prime and $\frac{n-2}{2}$ is compound), Y_a is increased only by 1 although its gap to Z_a is very quickly very greater than 1.

To show that $Y_c(n) > Z_c(n)$, one can use once more the analysis of variables evolution provided in section 4 : Y_c is increasing when $n-3$ is compound while Z_c is increasing when $\dfrac{n-2}{2}$ is compound. Z_c is an increasing function, there are times when Y_c is decreasing but not so often, and this has as consequence that over a rather small value, Z_c never catches Y_c again.

To know precisely from which values of n wished inequalities are verified, we use once more gaps values and minorations/majorations provided in [1].

To show that $Z_a(n) > Y_a(n)$ (resp. $Y_c(n) > Z_c(n)$), we show that the gap

$$Z_a(n) - Y_a(n) = Y_c(n) - Z_c(n) = 2\pi\left(\frac{n}{2}\right) - \pi(n)$$

is always strictly greater than 0.

We use formula 3.9 of corollary 1 of theorem 2 of Rosser and Schoenfeld that states that $\pi(x) - \pi\left(\dfrac{x}{2}\right) < \dfrac{7x}{5\ ln\ x}$ for every $x > 1$.

We use the fact that $2\pi\left(\dfrac{n}{2}\right) - \pi(n) = \left(\pi\left(\dfrac{n}{2}\right) - \pi(n)\right) + \pi\left(\dfrac{n}{2}\right)$

So we have

$$\begin{aligned} 2\pi\left(\frac{n}{2}\right) - \pi(n) &> \frac{-7n}{5\ ln\ n} + \pi\left(\frac{n}{2}\right) \\ &> \frac{-7n}{5\ ln\ n} + \frac{n}{2\ (ln\ n + ln\ 0.5)} \qquad \text{(because of formula 3.5 of corollary 1 of theorem 2 in [1])} \end{aligned}$$

that is strictly greater than 0

$$\frac{n(5\ ln\ n - 14\ (ln\ n + ln\ 0.5))}{10\ ln\ n(ln\ n + ln\ 0.5)} > 0$$

that is equivalent to

$$5\ ln\ n - 14\ (ln\ n + ln\ 0.5) > 0$$

that is always true when $n \geqslant 6$.

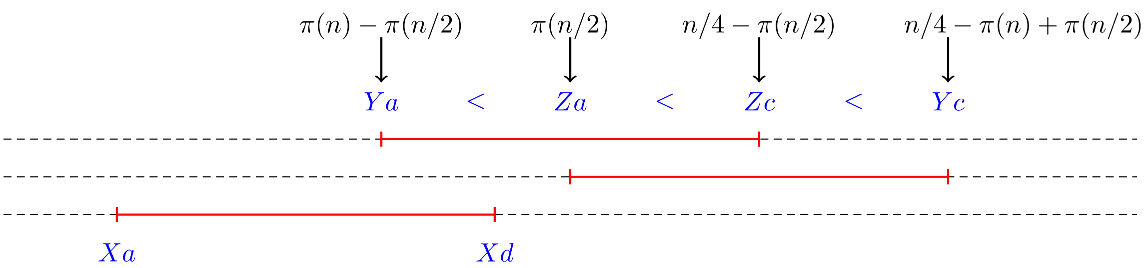
5.3 Strict order on $Y_a(n), Y_c(n), Z_a(n)$ and $Z_c(n)$ variables

$Y_a(n), Z_a(n), Z_c(n)$ and $Y_c(n)$ variables are so strictly ordered in the following way :

$$Y_a(n) < Z_a(n) < Z_c(n) < Y_c(n)$$

for every $n \geqslant 304$.

A graphical representation of gaps between variables can be found above, that shows their entanglement :



$Z_c(n) - Y_a(n)$, $Y_c(n) - Z_a(n)$ and $X_d(n) - X_a(n)$ gaps are strictly greater than 0 and equal to $\dfrac{n}{4} - \pi(n)$.

5.4 $X_a(n) > 0$ inequality study

To be ensured that $X_a(n)$ is never equal to 0, one has to minorate $X_d(n)$ by $\dfrac{n}{4} - \pi(n)$, i.e. the value of $X_d(n) - X_a(n)$ gap.

But $X_d(n) = Y_c(n) - X_c(n)$.

To minorate $X_d(n)$, one has to minorate $Y_c(n)$ and to majorate $X_c(n)$.

$Y_c(n)$ is the number of odd compound numbers that are between $n/2$ and n (associated to 3).

To minorate $Y_c(n)$, we use the fact that the number of odd compound numbers that are between $n/2$ and n is equal to $\dfrac{n}{4} - \pi(n) + \pi\left(\dfrac{n}{2}\right)$.

$X_c(n)$, which is the number of n's decompositions of the form *prime+compound* is majorable by the total number of odd compound numbers that are between $n/2$ and n, that is itself majorable by the number of odd compound numbers that are between n and $\dfrac{3n}{2}$.

$X_c(n)$ is so majorable by $\dfrac{n}{4} - \left(\pi\left(\dfrac{3n}{2}\right) - \pi(n)\right)$ (the number of odd compound numbers from the interval from n to $\dfrac{3n}{2}$ is $\dfrac{n}{4}$, the number of prime numbers in this interval is $\pi\left(\dfrac{3n}{2}\right) - \pi(n)$, the number of odd compound numbers in this interval is the difference of those two numbers).

$Y_c(n) - X_c(n)$ is thus always greater than the difference between $Y_c(n)$'s minoration and $X_c(n)$'s majoration, that will give

$$Y_c(n) - X_c(n) > \frac{n}{4} - \pi(n) + \pi\left(\frac{n}{2}\right) - \frac{n\,((ln\ n + ln\ 1.5)(ln\ n - 4 \times 1.25506) - 1.25506 \times 6 \times ln\ n)}{4\ (ln\ n + ln\ 1.5)\ ln\ n}.$$

$Y_c(n) - X_c(n) = X_d(n)$ is always greater than $\dfrac{n}{4} - \pi(n)$ when $n > 0$ (we also make this constatation by computing all this by program).

Indeed, $Y_c(n) - X_c(n) > \dfrac{n}{4} - \pi(n)$ when

$$\pi\left(\frac{n}{2}\right) - \frac{n\,((ln\ n + ln\ 1.5)(ln\ n - 4 \times 1.25506) - 1.25506 \times 6 \times ln\ n)}{4\ (ln\ n + ln\ 1.5)\ ln\ n} > 0.$$

We replace $\pi\left(\dfrac{n}{2}\right)$ by its minoration provided by formula 3.5 of corollary 1 of theorem 2 in [1] (that is $\dfrac{n}{2(ln\ n + ln\ 0.5)}$), we reduce to same denominator, that is always greater than 0 when $n \geqslant 2$ and that we forget, we are looking for condition that ensures that numerator is always strictly greater than 0, numerator that is equal to :

$$n[(2(ln\ n + ln\ 1.5)ln\ n) - (ln\ n + ln\ 0.5)((ln\ n + ln\ 1.5)(ln\ n - 5.02024) - 7.53036\ ln\ n)]$$

After several computations, we obtain that numerator, with unknown $ln\ n$, is equal to polynom

$$-(ln\ n)^3 + 14.6755387366(ln\ n)^2 - 2.48889541216(ln\ n) - 0.26611665186$$

The biggest root of this polynom is nearly equal to 14.502656936497 from which exponential is equal to 1988034.33365. Difference between $X_d(n)$ and $X_a(n)$ is thus always greater than $\dfrac{n}{4} - \pi(n)$ for every

$n \geqslant 1988034.33365$.

We can thus conclude that for every $n \geqslant 1988034.33365$ (necessary condition to have $X_d(n) - X_a(n) > \frac{n}{4} - \pi(n)$), $X_a(n)$ (number of n's decompositions as a sum of two primes) is strictly greater than 0.

In annex 2 are provided graphic representations of sets bijections for cases $n = 32$, 34, 98 and 100.

The file *http : //denise.vella.chemla.free.fr/annexes.pdf* contains

- an historical recall of a Laisant's note in which he presented yet in 1897 the idea of "strips" of odd numbers to be put in regard and to be colorated to *see* Goldbach decompositions ;
- a program and its execution that implement ideas presented here.

6 Demonstrations

6.1 Utilitaries

Let us demonstrate that if n is an odd number double (i.e. of the form $4k+2$), then $\left\lfloor \frac{n-2}{4} \right\rfloor = \left\lfloor \frac{n-4}{4} \right\rfloor + 1$.
Indeed, the left part of the equality is equal to $\left\lfloor \frac{(4k+2)-2}{4} \right\rfloor = \left\lfloor \frac{4k}{4} \right\rfloor = k$.
The right part of the equality is equal to $\left\lfloor \frac{(4k+2)-4}{4} \right\rfloor + 1 = \left\lfloor \frac{4k-2}{4} \right\rfloor + 1 = (k-1)+1 = k$.

Let us demonstrate that if n is an even number double (i.e. of the $4k$), then $\left\lfloor \frac{n-2}{4} \right\rfloor = \left\lfloor \frac{n-4}{4} \right\rfloor$.
$\left\lfloor \frac{4k-2}{4} \right\rfloor = k-1$ and $\left\lfloor \frac{4k-4}{4} \right\rfloor = k-1$.

We can also express this by the following way : if n is an odd number double, $\left\lfloor \frac{n-2}{4} \right\rfloor = \frac{n-2}{4} = \left\lfloor \frac{n-4}{4} \right\rfloor + 1$ although if n is an even number double, $\left\lfloor \frac{n-2}{4} \right\rfloor = \frac{n-4}{4} = \left\lfloor \frac{n-4}{4} \right\rfloor$.

6.2 5, 6 and 8 properties

5, 6 and 8 properties follows directly from variables definitions.

6.2.1 Property 5

Property 5 states that $X_a(n) + X_c(n) = Z_a(n) + \delta_{2p}(n)$ with $\delta_{2p}(n)$ that is equal to 1 in the case if n is a prime number double and that is equal to 0 otherwise.

By definition, $X_a(n)+X_c(n)$ counts number of n's decompositions of the form $prime+x$ with $prime \leqslant n/2$. But by the fact that $Z_a(n)$ counts on its side number of decompositions of the form $3 + prime$ with $prime < n/2$, adding δ_{2p} to $Z_a(n)$ permits to ensure the equality's invariance in all cases and in particular when n is a prime number double.

6.2.2 Property 6

Property 6 states that $X_b(n) + X_d(n) = Z_c(n) + \delta_{2c-imp}$ with δ_{2c-imp} that is equal to 1 in the case if n is a compound odd number, and is equal to 0 otherwise.

By definition, $X_b(n) + X_d(n)$ counts the number of decompositions of the form $compound + x$ with $compound \leqslant n/2$. But by the fact that $Z_c(n)$ counts on its side number of decompositions of the form $3 + compound$ with $compound < n/2$, adding δ_{2c-imp} to $Z_c(n)$ permits to ensure the equality's invariance in all cases and in particular when n is an odd compound double.

6.2.3 Property 8

Property 8 states that $Z_c(n) - Y_a(n) = X_d(n) - X_a(n) - \delta_{2c-imp}$ with δ_{2c-imp} that is equal to 1 if n is an odd compound double and is equal to 0 otherwise.

By definition, $Z_c(n)$ counts the number of odd compound numbers strictly lesser than $n/2$. It counts also the number of n's decompositions of the form $compound + x$ with $compound < n/2$ (let us call E this decompositions set).

By definition, $Y_a(n)$ counts the number of prime numbers strictly greater than $n/2$. It counts also the number of n's decompositions of the form $x + prime$ with $prime > n/2$ (let us call F this decompositions set).

n's decompositions of the form $compound + prime$ are at the same time in E and in F. By computing $Z_c(n) - Y_a(n)$, we are computing the cardinality of a set that is equal to $X_d(n) - X_a(n)$ by definition of what $Y_a(n), Z_c(n), X_d(n)$ and $X_a(n)$ variables count.

6.2.4 Property 7

Let us demonstrate that $Z_c(n) - Y_a(n) = Y_c(n) - Z_a(n) - \delta_{4k+2}$ with δ_{4k+2} that is equal to 1 if n is an odd double (there exists some $k \geqslant 3$ such that $n = 4k + 2$) and is equal to 0 otherwise.

One uses a recurrence reasoning :

i) One initialises recurrences according to the 3 sorts of numbers to be envisaged : even doubles (of the form $4k$, like 16), odd doubles (of the form $4k + 2$) that are prime (like 14) or that are compound (like 18).

Property 7 is true for $n = 14$ because $Z_c(14) = 0, Y_a(14) = 2, Y_c(14) = 1, Z_a(14) = 2$ and $\delta_{4k+2}(14) = 1$ and thus $Z_c(14) - Y_a(14) = Y_c(14) - Z_a(14) - \delta_{4k+2}(14)$;
Property 7 is true for $n = 16$ because $Z_c(16) = 0, Y_a(16) = 2, Y_c(16) = 1, Z_a(16) = 3$ and $\delta_{4k+2}(16) = 0$ and thus $Z_c(16) - Y_a(16) = Y_c(16) - Z_a(16) - \delta_{4k+2}(16)$;
Property 7 is true for $n = 18$ because $Z_c(18) = 0, Y_a(18) = 2, Y_c(18) = 2, Z_a(18) = 3$ and $\delta_{4k+2}(18) = 1$ and thus $Z_c(18) - Y_a(18) = Y_c(18) - Z_a(18) - \delta_{4k+2}(18)$;

ii) We rewrite property in the following form $Z_a(n) + Z_c(n) + \delta_{4k+2} = Y_a(n) + Y_c(n)$.
Four cases must be considered : two cases in which n is an odd double (prime or compound) and $n + 2$ is an even double and two cases in which n is an even double and $n + 2$ is an the double of an odd number (that is prime or compound).

iia) *n even double and $n + 2$ prime double (ex : $n = 56$)* :

n	δ_{2p}	δ_{2c-imp}	δ_{4k+2}
n	0	0	0
$n+2$	1	0	1

One states the hypothesis that property 7 is verified by n,

$$Z_a(n) + Z_c(n) + \delta_{4k+2}(n) = Y_a(n) + Y_c(n) \quad (H)$$

Let us demonstrate the property is true for $n + 2$,

$$Z_a(n+2) + Z_c(n+2) + \delta_{4k+2}(n+2) = Y_a(n+2) + Y_c(n+2) \quad (Ccl)$$

One has $Z_a(n + 2) = Z_a(n)$ and $Z_c(n + 2) = Z_c(n)$.

Recall property 3 concerning $Y_a(n)$ and $Y_c(n)$:

$$Y_a(n) + Y_c(n) = \left\lfloor \frac{n-2}{4} \right\rfloor \quad (3)$$

In (Ccl), we can, by recurrence hypothesis and by proprerty (3), replace the left part of the equality by $Z_a(n) + Z_c(n) + 1$ and than by $Y_a(n) + Y_c(n) + 1$ (because (H)) and then by $\left\lfloor \frac{n-2}{4} \right\rfloor + 1$ (because (3)) that is equal to $\left\lfloor \frac{n}{4} \right\rfloor$.

But in (Ccl), we can also replace the right part of the equality by $\left\lfloor \frac{n}{4} \right\rfloor$ because of property (3).

There is also, for $n+2$, equality between left and right parts of the equality, i.e. property 7 is verified by $n+2$. From the hypothesis that property is verified by n, we demonstrated that property is true for $n+2$.

iib) *n even double and* $n+2$ *odd compound double* (*ex :* $n = 48$) :

n	δ_{2p}	δ_{2c-imp}	δ_{4k+2}
n	0	0	0
$n+2$	0	1	1

One states the hypothesis that property 7 is verified by n,

$$Z_a(n) + Z_c(n) + \delta_{4k+2}(n) = Y_a(n) + Y_c(n) \quad (H)$$

Let us demonstrate the property is verified by $n+2$,

$$Z_a(n+2) + Z_c(n+2) + \delta_{4k+2}(n+2) = Y_a(n+2) + Y_c(n+2) \quad (Ccl)$$

One has $Z_a(n+2) = Z_a(n)$ and $Z_c(n+2) = Z_c(n)$.

And one has also $Y_a(n+2) = Y_a(n) + 1$ and $Y_c(n+2) = Y_c(n)$.

Let us recall property 3 concerning $Y_a(n)$ and $Y_c(n)$:

$$Y_a(n) + Y_c(n) = \left\lfloor \frac{n-2}{4} \right\rfloor \quad (3)$$

In (Ccl), we can, by recurrence hypothesis and property (3), replace the left part of the equality by $Z_a(n) + Z_c(n) + 1$ and the by $Y_a(n) + Y_c(n) + 1$ (by (H)) and then by $\left\lfloor \frac{n-2}{4} \right\rfloor + 1$ (by (3)) that is equal to $\left\lfloor \frac{n}{4} \right\rfloor$.

But in (Ccl), we can also replace the right part of the equality by $\left\lfloor \frac{n}{4} \right\rfloor$ because of evolutions of $Y_a(n)$ and $Y_c(n)$.

There is also, for $n+2$, equality between left and right parts of the equality, i.e. property 7 is verified by $n+2$. From the hypothesis that property is verified by n, we demonstrated that property is verified by $n+2$.

iic) *n prime double and* $n+2$ *even double* (*ex :* $n = 74$) :

n	δ_{2p}	δ_{2c-imp}	δ_{4k+2}
n	1	0	1
$n+2$	0	0	0

One states the hypothesis that property 7 is verified by n,

$$Z_a(n) + Z_c(n) + \delta_{4k+2}(n) = Y_a(n) + Y_c(n) \quad (H)$$

Let us demonstrate that property is verified by $n+2$,

$$Z_a(n+2) + Z_c(n+2) + \delta_{4k+2}(n+2) = Y_a(n+2) + Y_c(n+2) \qquad (Ccl)$$

One has $Z_a(n+2) = Z_a(n) + 1$ and $Z_c(n+2) = Z_c(n)$.

Let us recall property 3 concerning $Y_a(n)$ and $Y_c(n)$:

$$Y_a(n) + Y_c(n) = \left\lfloor \frac{n-2}{4} \right\rfloor \qquad (3)$$

In (Ccl), we can, by recurrence hypothesis and property (3), replace the left part of the equality by $Z_a(n) + Z_c(n) + 1$ and then by $Y_a(n) + Y_c(n)$ (because (H)) and then by $\left\lfloor \frac{n-2}{4} \right\rfloor$ (by (3)) that is equal to $\left\lfloor \frac{n}{4} \right\rfloor$.

But in (Ccl), one can also replace the right part of the equality by $\left\lfloor \frac{n}{4} \right\rfloor$ because of property (3).

There is once more, for $n+2$, equality between left and right part of the equality, i.e. property 7 is verified by $n+2$. From the hypothesis that property is true for n, we demonstrated that property is verified by $n+2$.

iid) *n odd compound double and $n+2$ even double* (ex : $n = 70$) :

n	δ_{2p}	δ_{2c-imp}	δ_{4k+2}
n	0	1	1
$n+2$	0	0	0

We state the hypothesis that property 7 is true for n,

$$Z_a(n) + Z_c(n) + \delta_{4k+2}(n) = Y_a(n) + Y_c(n) \qquad (H)$$

Let us demonstrate that it is true for

$$Z_a(n+2) + Z_c(n+2) + \delta_{4k+2}(n+2) = Y_a(n+2) + Y_c(n+2) \qquad (Ccl)$$

One has $Z_a(n+2) = Z_a(n)$ and $Z_c(n+2) = Z_c(n) + 1$.

Let us recall property 3 concerning $Y_a(n)$ and $Y_c(n)$:

$$Y_a(n) + Y_c(n) = \left\lfloor \frac{n-2}{4} \right\rfloor \qquad (3)$$

In (Ccl), we can, by recurrence hypothesis and property (3), replace the left part of the equality by $Z_a(n) + Z_c(n) + 1$ and then by $Y_a(n) + Y_c(n)$ (by (H)) and then by $\left\lfloor \frac{n-2}{4} \right\rfloor$ (by (3)) that is equal to $\left\lfloor \frac{n}{4} \right\rfloor$.

But in (Ccl), one can also replace the right part of the equality by $\left\lfloor \frac{n}{4} \right\rfloor$ because of property (3).

There is once more, for $n+2$, equality between left and right parts of the equality, i.e. property 7 is verified by $n+2$. From the hypothesis that property is verified by n, we demonstrated property is verified by $n+2$.

Annex 1 : variables values array for n between 14 and 100

n	$X_a(n)$	$X_b(n)$	$X_c(n)$	$X_d(n)$	$Y_a(n)$	$Y_c(n)$	$\left\lfloor \frac{n-2}{4} \right\rfloor$	$Z_a(n)$	$Z_c(n)$	$\left\lfloor \frac{n-4}{4} \right\rfloor$	$\delta_{2p}(n)$	$\delta_{2c-imp}(n)$	$\delta_{4k+2}(n)$
14	2	0	1	0	2	1	3	2	0	2	1	0	1
16	2	0	1	0	2	1	3	3	0	3	0	0	0
18	2	0	1	1	2	2	4	3	0	3	0	1	1
20	2	1	1	0	3	1	4	3	1	4	0	0	0
22	3	1	1	0	4	1	5	3	1	4	1	0	1
24	3	0	1	1	3	2	5	4	1	5	0	0	0
26	3	1	2	0	4	2	6	4	1	5	1	0	1
28	2	1	3	0	3	3	6	5	1	6	0	0	0
30	3	0	2	2	3	4	7	5	1	6	0	1	1
32	2	2	3	0	4	3	7	5	2	7	0	0	0
34	4	1	2	1	5	3	8	5	2	7	1	0	1
36	4	0	2	2	4	4	8	6	2	8	0	0	0
38	2	2	5	0	4	5	9	6	2	8	1	0	1
40	3	1	4	1	4	5	9	7	2	9	0	0	0
42	4	0	3	3	4	6	10	7	2	9	0	1	1
44	3	2	4	1	5	5	10	7	3	10	0	0	0
46	4	2	4	1	6	5	11	7	3	10	1	0	1
48	5	0	3	3	5	6	11	8	3	11	0	0	0
50	4	2	4	2	6	6	12	8	3	11	0	1	1
52	3	3	5	1	6	6	12	8	4	12	0	0	0
54	5	1	3	4	6	7	13	8	4	12	0	1	1
56	3	4	5	1	7	6	13	8	5	13	0	0	0
58	4	3	5	2	7	7	14	8	5	13	1	0	1
60	6	0	3	5	6	8	14	9	5	14	0	0	0
62	3	4	7	1	7	8	15	9	5	14	1	0	1
64	5	2	5	3	7	8	15	10	5	15	0	0	0
66	6	1	4	5	7	9	16	10	5	15	0	1	1
68	2	5	8	1	7	9	16	10	6	16	0	0	0
70	5	3	5	4	8	9	17	10	6	16	0	1	1
72	6	2	4	5	8	9	17	10	7	17	0	0	0
74	5	4	6	3	9	9	18	10	7	17	1	0	1
76	5	4	6	3	9	9	18	11	7	18	0	0	0
78	7	2	4	6	9	10	19	11	7	18	0	1	1
80	4	5	7	3	9	10	19	11	8	19	0	0	0
82	5	5	7	3	10	10	20	11	8	19	1	0	1
84	8	1	4	7	9	11	20	12	8	20	0	0	0
86	5	5	8	3	10	11	21	12	8	20	1	0	1
88	4	5	9	3	9	12	21	13	8	21	0	0	0
90	9	0	4	9	9	13	22	13	8	21	0	1	1
92	4	6	9	3	10	12	22	13	9	22	0	0	0
94	5	5	9	4	10	13	23	13	9	22	1	0	1
96	7	2	7	7	9	14	23	14	9	23	0	0	0
98	3	6	11	4	9	15	24	14	9	23	0	1	1
100	6	4	8	6	10	14	24	14	10	24	0	0	0

n	$X_a(n)$	$X_b(n)$	$X_c(n)$	$X_d(n)$	$Y_a(n)$	$Y_c(n)$	$\left\lfloor \frac{n-2}{4} \right\rfloor$	$Z_a(n)$	$Z_c(n)$	$\left\lfloor \frac{n-4}{4} \right\rfloor$	δ_{2p}	δ_{2c-imp}	δ_{4k+2}
999 998	4 206	32 754	37 331	175 708	36 960	213 039	249 999	41 537	208 461	249 998	0	1	1
1 000 000	5 402	31 558	36 135	176 904	36 960	213 039	249 999	41 537	208 462	249 999	0	1	0
9 999 998	28 983	287 084	319 529	1 864 403	316 067	2 183 932	2 499 999	348 511	2 151 487	2 499 998	1	0	1
10 000 000	38 807	277 259	309 705	1 874 228	316 066	2 183 933	2 499 999	348 512	2 151 487	2 499 999	0	1	0

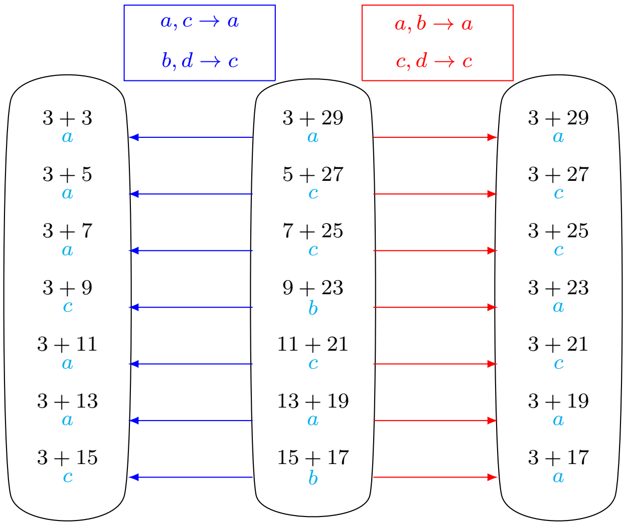
Annex 2 : sets bijections

- Case $n = 32$

$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

$Z_a + Z_c = \left\lfloor \frac{n-4}{4} \right\rfloor = 7$



$Y_a = X_a + X_b = 4$

$Y_c = X_c + X_d = 3$

$Y_a + Y_c = \left\lfloor \frac{n-2}{4} \right\rfloor = 7$

$X_a = 2$

$X_b = 2$

$X_c = 3$

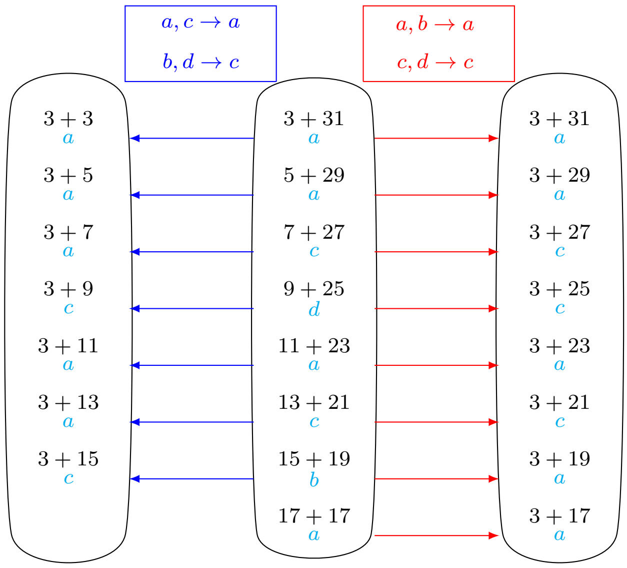
$X_d = 0$

- Case $n = 34$

$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

$Z_a + Z_c = \left\lfloor \frac{n-4}{4} \right\rfloor = 7$



$Y_a = X_a + X_b = 5$

$Y_c = X_c + X_d = 3$

$Y_a + Y_c = \left\lfloor \frac{n-2}{4} \right\rfloor = 8$

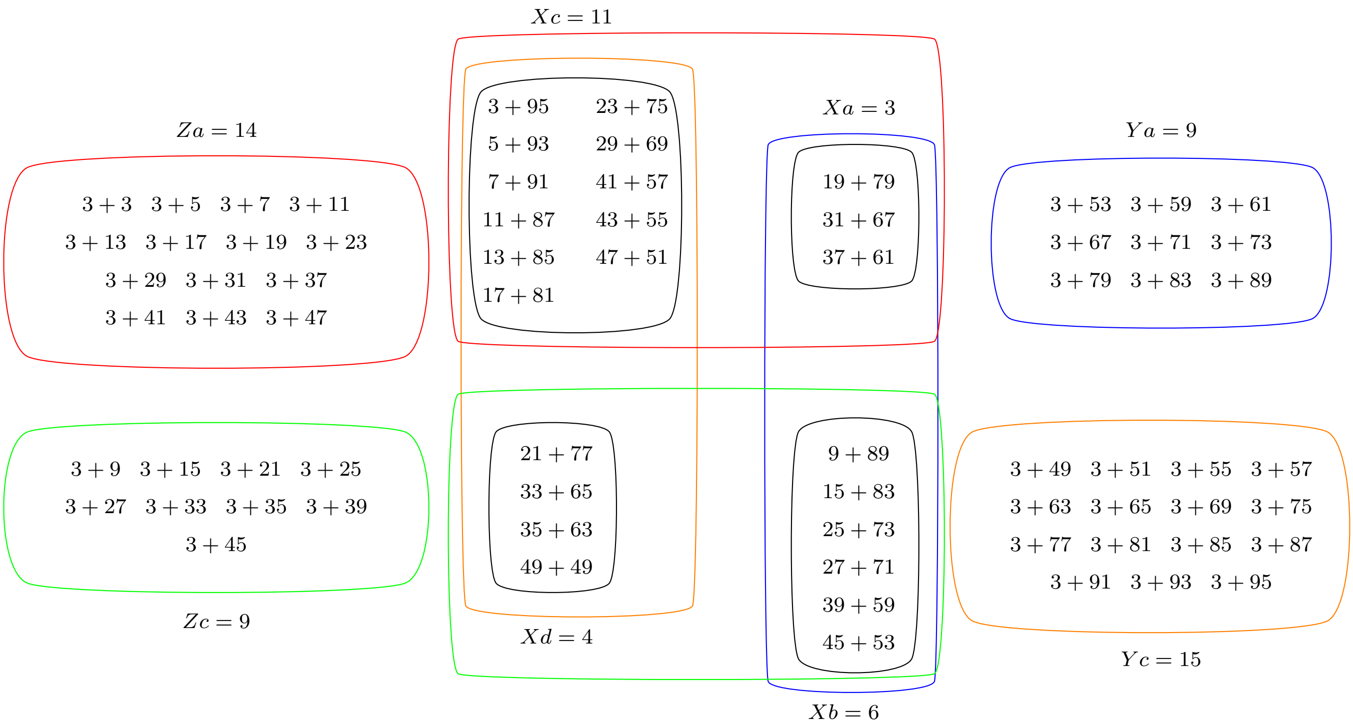
$X_a = 4$

$X_b = 1$

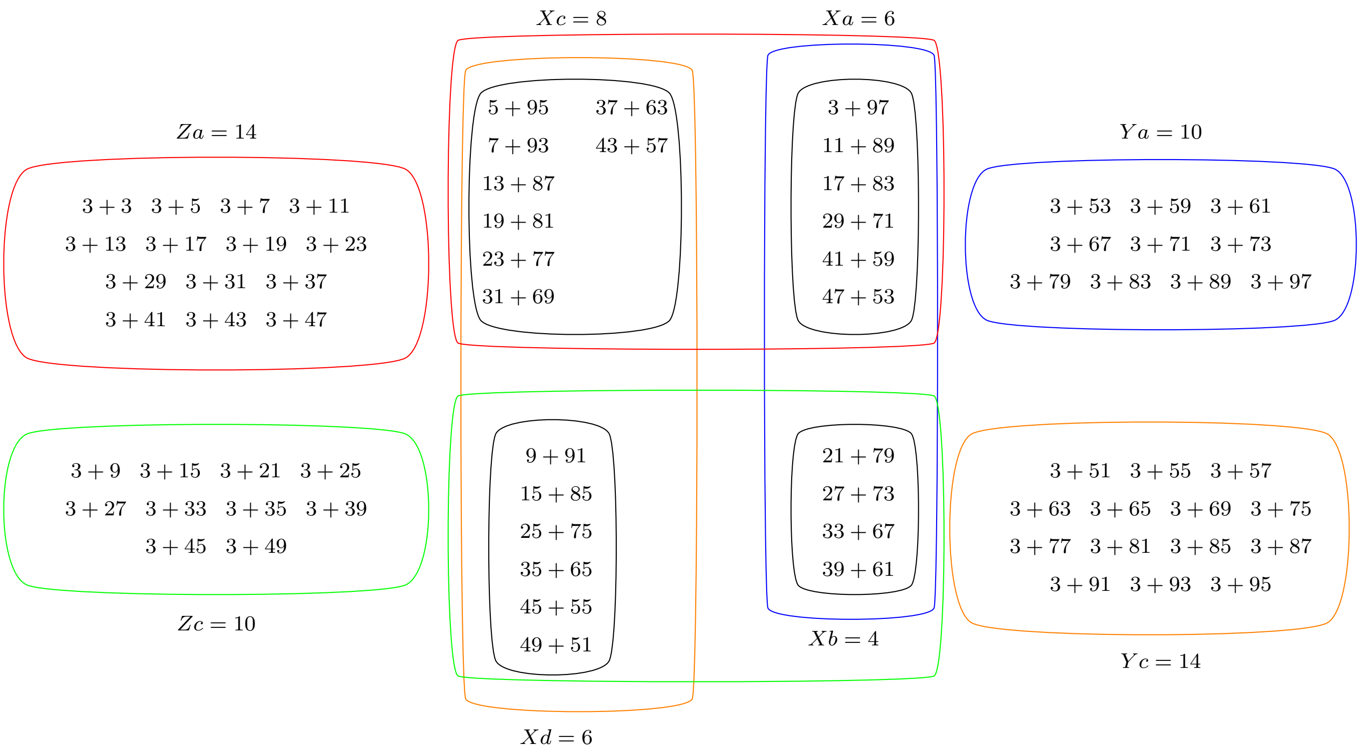
$X_c = 2$

$X_d = 1$

- **Case** $n = 98$



- **Case** $n = 100$



Annex 3 : rewriting rules and automata theory

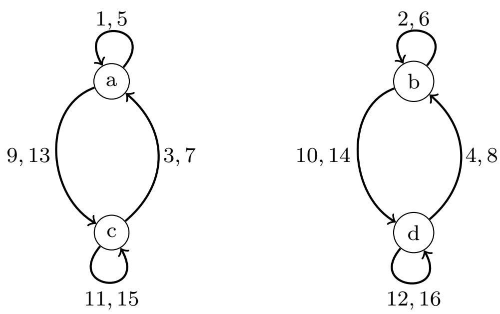
This annex studies variables $X_a(n), X_b(n), X_c(n), X_d(n)$ evolution that we deduce from analyzing words rewriting rules, presented from automata theory point of view.

If we consider each line of the global array as a word on alphabet $A = \{a, b, c, d\}$, $n+2$ even number's word is obtained by the following way from even number n's word :

- first letter of $n+2$'s word is a if $n-3$ is prime and c otherwise (this first letter is the only one that introduces indeterminism because it doesn't belong to n's word and it can't be deduced from n's word letters ;
- following letters of $n+2$'s word are obtained by applying parallely following rewriting rules to n's word :

$$\begin{aligned}
aa &\rightarrow a && (1)\\
ab &\rightarrow b && (2)\\
ac &\rightarrow a && (3)\\
ad &\rightarrow b && (4)\\
ba &\rightarrow a && (5)\\
bb &\rightarrow b && (6)\\
bc &\rightarrow a && (7)\\
bd &\rightarrow b && (8)\\
ca &\rightarrow c && (9)\\
cb &\rightarrow d && (10)\\
cc &\rightarrow c && (11)\\
cd &\rightarrow d && (12)\\
da &\rightarrow c && (13)\\
db &\rightarrow d && (14)\\
dc &\rightarrow c && (15)\\
dd &\rightarrow d && (16)
\end{aligned}$$

We can represent those rewriting rules by the two above deterministic automata, from which edges are labelled by applicable rules to one given letter of n's word :



- finally, one letter concatenation at the end of the word, in case if n is an even double (i.e. of the form $4k$) obeys to following rule :
  - if n's word has an a or b letter as last letter, after having obtained $n+2$'s word by applying rewriting rules, we concatenate a letter a to it (at last position) ;
  - if n's word has a c or d letter as last letter, after having obtained $n+2$'s word by applying rewriting rules, we concatenate a letter d to it (at last position).

If we take as convention to notate $X_{xy}(n)$ occurences number of xy letters sequences in n's word, the following equalities provide a, b, c or d letters numbers evolution when passing from n's word to $n+2$'s word.

$$\begin{aligned}
X_a(n+2) &= X_a(n) - X_{ca}(n) - X_{da}(n) + X_{ac}(n) + X_{bc}(n) + \delta_{n-3_is_prime}(n) + \delta_a(n)\\
X_b(n+2) &= X_b(n) - X_{cb}(n) - X_{db}(n) + X_{ad}(n) + X_{bd}(n) + \delta_{n-3_is_prime}(n)\\
X_c(n+2) &= X_c(n) - X_{ac}(n) - X_{bc}(n) + X_{ca}(n) + X_{da}(n) + \delta_{n-3_is_prime}(n)\\
X_d(n+2) &= X_d(n) - X_{ad}(n) - X_{bd}(n) + X_{cb}(n) + X_{db}(n) + \delta_{n-3_is_prime}(n) + \delta_d(n)
\end{aligned}$$

with $\delta_a(n)$ that is equal to 1 if n is an even number (i.e. of the form $4k$) and if n's word last letter is an a or b letter, $\delta_d(n)$ that is equal to 1 if n is an even number (i.e. of the form $4k$) and if n's word last letter is a c or d letter and finally with $\delta_{n-3}(n)$ that is equal to 1 if $n-3$ is prime and equal to 0 otherwise.

Bibliographie

[1] J. Barkley Rosser, Lowell Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois Journal of Mathematics, 1962.