Goldbach conjecture, 4 letters language, variables and invariants

Denise Vella-Chemla

21/04/2014

1 Introduction

Goldbach 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 note 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 second 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 the 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.

C / L	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, diagonal letters *aaabaa* that begin 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 $\left\lfloor \frac{n-2}{4} \right\rfloor$ elements.

4) n begin 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 $6 \leqslant k \leqslant \left\lfloor \frac{n+4}{2} \right\rfloor$.

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

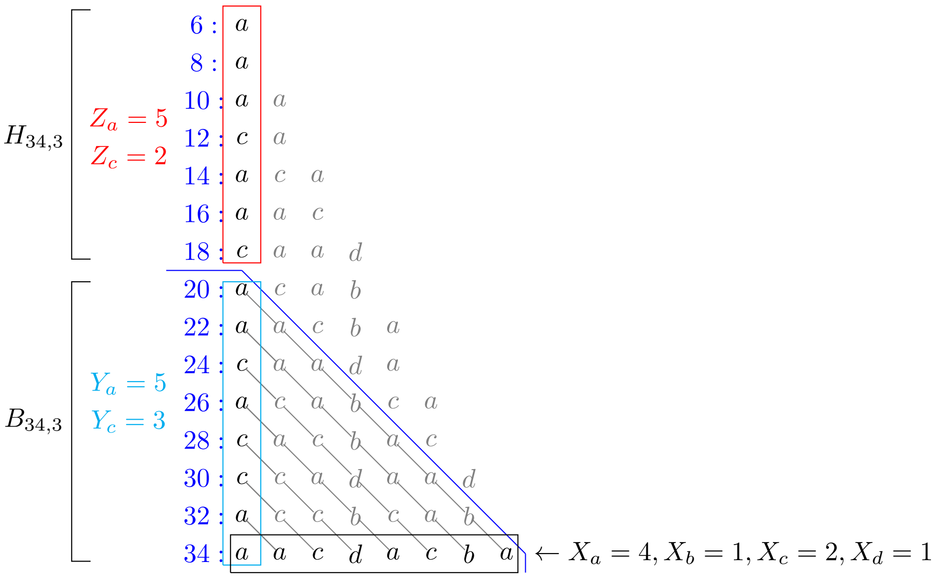


FIGURE 2 : $n = 34$

To better understand computations 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 of 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 a 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 :

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

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

$$Y_a(n) = X_a(n) + X_b(n) \quad (1)$$

$$Y_c(n) = X_c(n) + X_d(n) \quad (2)$$

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

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

Let us add two new properties to those ones :

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

with δ_{2p} equal to 1 in the case that n is the double of a prime number and equal to 0 either.

$$X_b(n) + X_d(n) = Z_c(n) + \delta_{spec} \tag{6}$$

with δ_{spec} equal to 0 in the case that there exists k such that $n = 4k$, or in the case that n is the double of a prime number, and equal to 1 either.

4 Variables evolution

In this section, let us study how different variables change, in the aim to deduce that X_a (the number of an even number decompositions that are sums of two primes) can't never be null.

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

$Z_a(n)$ is increased by 1 when $\frac{n-2}{2}$ is prime and $Z_c(n)$ is increased by 1 each time when $\frac{n-2}{2}$ is compound.

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

Let us see now in detail how $Y_a(n)$ and $Y_c(n)$ change.

Dans le cas où n est un double d'impair, on ajoute un nombre à l'intervalle $H_{n,3}$; si ce nombre $(n-3)$ est premier (resp. composé), $Y_a(n)$ (resp. $Y_c(n)$) est augmenté de 1 par rapport à $Y_a(n-2)$ (resp. $Y_c(n-2)$).

If n is an even number double, there are 4 possible cases. Let us study how top decompositions belonging to $C_{n,3}$'s top part (i.e. $H_{n,3}$) evoluate.

- if $n-3$ and $n/2-1$ are both primes, we remove at bottom and add at top of $H_{n,3}$ two letters that are of the same type, thus $Y_a(n)$ and $Y_c(n)$ remain constant ;
- if $n-3$ is prime and $n/2-1$ is compound then $Y_a(n)$ is increased by 1 and $Y_c(n)$ is decreased by 1 ;
- if $n-3$ is compound and $n/2-1$ is prime then $Y_c(n)$ is increased by 1 and $Y_a(n)$ is decreased by 1 ;
- if $n-3$ and $n/2-1$ are both compound, we remove at bottom and add at top of $H_{n,3}$ two letters that are of the same type thus $Y_a(n)$ and $Y_c(n)$ remain constants.

But we don't succeed in deducing from all those variables entanglement that $X_a(n)$ is always strictly positive. In annex 1 are provided in an array values of different variables for n between 14 and 100.

5 Leading to a contradiction

However, let us try to reach a contradiction from the hypothesis that $X_a(n) = 0$.

If $X_a(n) = 0$, we have

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

This is equivalent to

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

and thus, because of (2), to

$$Y_c(n) = \left\lfloor \frac{n-2}{4} \right\rfloor - X_b(n) \tag{7}$$

Here, 2 cases have to be distinguished :

- *case 1* : If n is the double of an odd number (i.e. of the form $4k+2$), then

$$\left\lfloor \frac{n-2}{4} \right\rfloor = \left\lfloor \frac{n-4}{4} \right\rfloor + 1 \tag{a}$$

- *case 2* : If n is the double of an even number (i.e. of the form $4k$), then

$$\left\lfloor \frac{n-2}{4} \right\rfloor = \left\lfloor \frac{n-4}{4} \right\rfloor \tag{b}$$

We replace $\left\lfloor \frac{n-2}{4} \right\rfloor$ by those two values in equality (7) above ; we obtain :

- *case 1* : $$Y_c(n) = \left\lfloor \frac{n-4}{4} \right\rfloor + 1 - X_b(n) \tag{7a}$$
- *case 2* : $$Y_c(n) = \left\lfloor \frac{n-4}{4} \right\rfloor - X_b(n) \tag{7b}$$

On the other part, from the hypothesis $X_a(n) = 0$ and from $X_a(n) + X_c(n) = Z_a(n) + \delta_{2p}$ (5), it results that

$$X_c(n) = Z_a(n) + \delta_{2p} \tag{8}$$

We rewrite (2) in

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

By identifying $X_c(n)$ in both (2') and (8), we obtain

$$Z_a(n) + \delta_{2p} = Y_c(n) - X_d(n) \tag{9'}$$

from which results

$$Y_c(n) = Z_a(n) + \delta_{2p} + X_d(n) \tag{2'}$$

that we rewrite

$$X_d(n) = Y_c(n) - Z_a(n) - \delta_{2p} \tag{9''}$$

From two equations (9') and (2) system :

$$\begin{cases} X_d(n) &= Y_c(n) - Z_a(n) - \delta_{2p} \\ Y_c(n) &= X_c(n) + X_d(n) \end{cases}$$

results

$$X_c(n) = Z_a(n) + \delta_{2p} - Y_c(n) \tag{10}$$

Contradiction results from the fact that $Y_c(n)$ is always greater than $Z_c(n)$ (since $n \geqslant 24$), itself always greater than $Z_a(n)$, n being greater than a rather small value of n (since $n \geqslant 240$). Equation (10) that we reached under $X_a(n) = 0$ hypothesis would provide a negative value for $X_c(n)$, that is clearly impossible, $X_c(n)$ counting, let us remind it, n decompositions of the form *prime* + *compound*.

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* provides

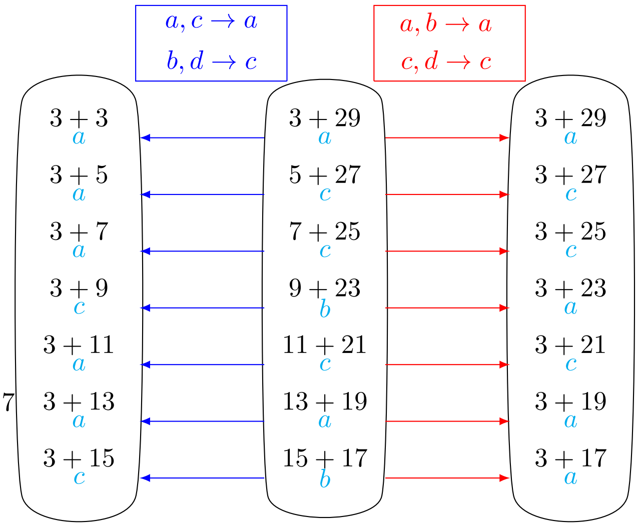
- an historical recall of a Laisant's note that presented yet in 1897 the idea of "strips" of odd numbers to be put in regard and to be colorated to *see* Goldbach decompositions ;
- a program and its execution that implements ideas presented here.

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

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

Annex 2 : sets bijections

- case $n = 32$



$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

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

$Y_a = X_a + X_b = 4$

$Y_c = X_c + X_d = 3$

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

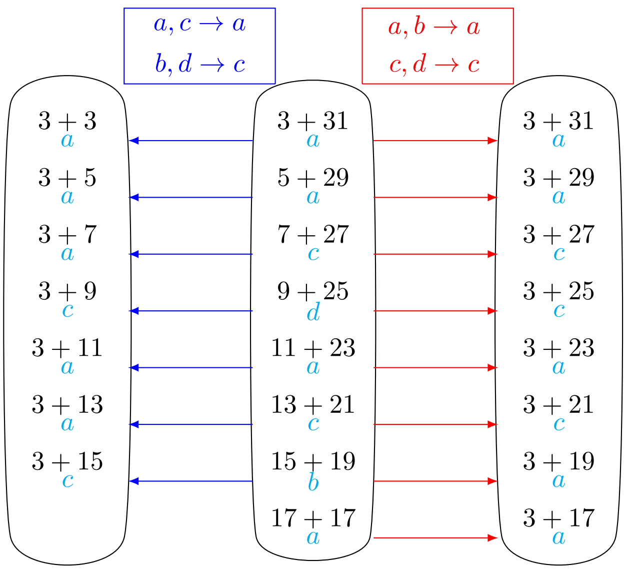
$X_a = 2$

$X_b = 2$

$X_c = 3$

$X_d = 0$

- case $n = 34$



$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

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

$Y_a = X_a + X_b = 5$

$Y_c = X_c + X_d = 3$

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

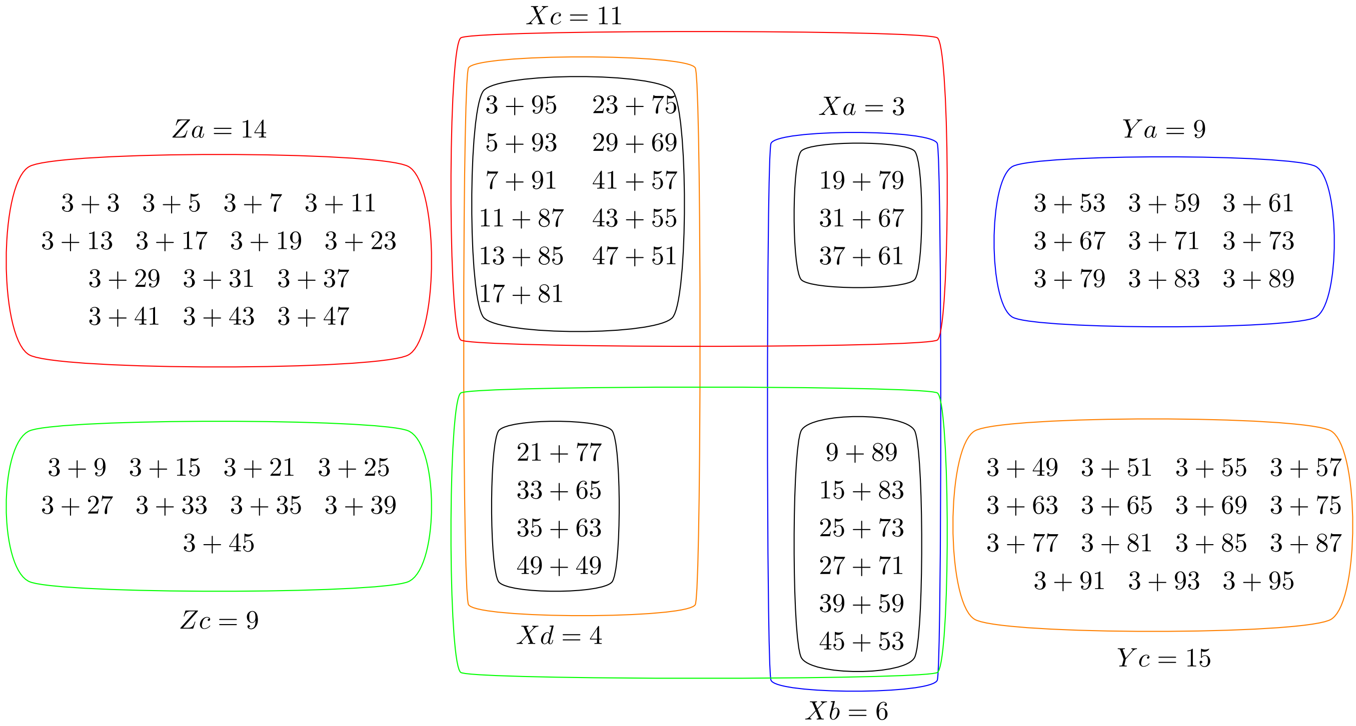
$X_a = 4$

$X_b = 1$

$X_c = 2$

$X_d = 1$

- case $n = 98$



- case $n = 100$