Conjecture de Goldbach, langage à 4 lettres, variables et invariants

Denise Vella-Chemla

8/5/2014

1 Introduction

La conjecture de Goldbach stipule que tout nombre pair sauf 2 est la somme de deux nombres premiers. Dans la suite, on s'intéresse aux décompositions d'un nombre pair n en somme de deux nombres impairs $p+q$ avec $3 \leqslant p \leqslant n/2$, $n/2 \leqslant q \leqslant n-3$ et $p \leqslant q$. On appelle p un *sommant de premier rang* et q un *sommant de second rang* de n.

Notations :

On désignera par :
- a : une décomposition de n de la forme $p+q$ avec p et q premiers ;
- b : une décomposition de n de la forme $p+q$ avec p composé et q premier ;
- c : une décomposition de n de la forme $p+q$ avec p premier et q composé ;
- d : une décomposition de n de la forme $p+q$ avec p et q composés.

Exemple :

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 Le tableau principal

On désigne par $T = (L, C) = (\ l_{n,m}\ )$ le tableau dont les éléments $l_{n,m}$ sont l'une des lettres a, b, c, d. L'indice n appartient à l'ensemble L des nombres pairs supérieurs ou égaux à 6. L'indice m, appartenant à l'ensemble C des nombres impairs supérieurs ou égaux à 3, est un élément de la liste des sommants de n de premier rang.

Considérons la fonction g définie ainsi :

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

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

La fonction $g(n)$ définit le plus grand des sommants de premier rang associés à n.

Comme l'on ne prend en compte que les décompositions de n de la forme $p+q$ où $p \leqslant q$, seules apparaîtront dans le tableau les lettres $l_{n,m}$ telles que $m \leqslant 2\left\lfloor \dfrac{n-2}{4} \right\rfloor + 1$ de sorte que le tableau contient les éléments suivants : $l_{6,3}, l_{8,3}, l_{10,3}, l_{10,5}, l_{12,3}, l_{12,5}, l_{14,3}, l_{14,5}, l_{14,7}, etc.$

Voici le début du tableau.

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 : mots des nombres pairs de 6 à 36

Remarques :

1) les mots situés sur les diagonales du tableau appelés *mots diagonaux* ont leurs lettres soit dans l'alphabet $A_{ab} = \{a, b\}$ soit dans l'alphabet $A_{cd} = \{c, d\}$.

2) un mot diagonal code des décompositions de même sommant de second rang.
Par exemple, sur la Figure 4, les lettres de la diagonale *aaabaa* qui commence à la lettre $l_{26,3} = a$ code les décompositions $3 + 23, 5 + 23, 7 + 23, 9 + 23, 11 + 23$ et $13 + 23$.

3) Désignons par l_n la ligne dont les éléments sont les $l_{n,m}$. La ligne l_n possède $\left\lfloor \dfrac{n-2}{4} \right\rfloor$ éléments.

4) n étant fixé, appelons $C_{n,3}$ la colonne formée des $l_{k,3}$ pour $6 \leqslant k \leqslant n$.

Dans cette colonne $C_{n,3}$, distinguons deux parties, la "partie haute" et la "partie basse" de la colonne.

Notons $H_{n,3}$ la "partie haute" de la colonne, i.e. l'ensemble des $l_{k,3}$ où $6 \leqslant k \leqslant \left\lfloor \dfrac{n+4}{2} \right\rfloor$.

Notons $B_{n,3}$ la "partie basse" de la colonne, i.e. l'ensemble des $l_{k,3}$ où $\left\lfloor \dfrac{n+4}{2} \right\rfloor < k \leqslant n$.

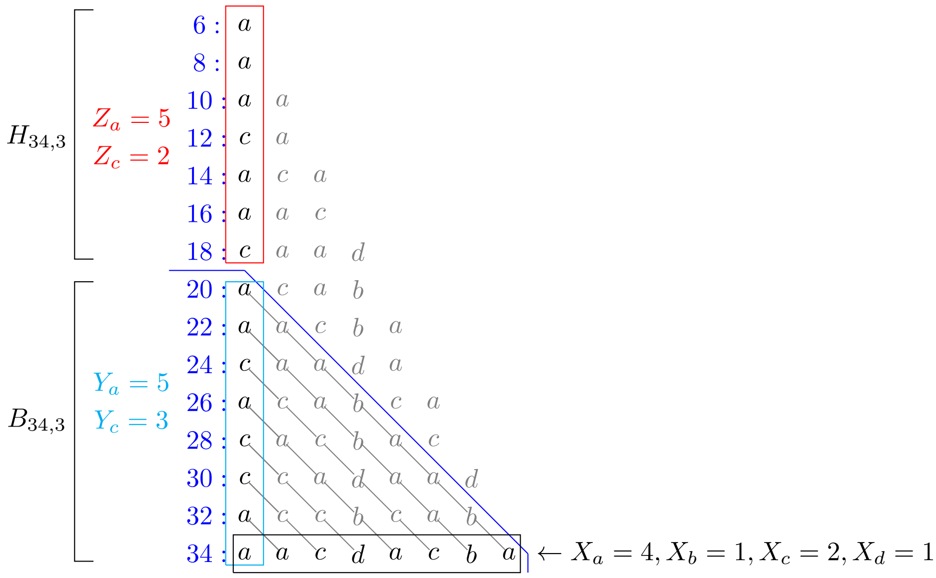


FIGURE 2 : $n = 34$

Pour mieux cerner les dénombrements de la section suivante, on utilisera la projection P de la ligne n sur la partie basse de la première colonne $B_{n,3}$ qui "associe" les lettres aux deux bouts d'une diagonale. Si on considère l'application $proj$ telle que $proj(a) = proj(b) = a$ et $proj(c) = proj(d) = c$ alors, puisque 3 est premier, $proj(l_{n,2k+1}) = l_{n-2k+2,3}$.

On peut comprendre l'effet de cette projection (qui préserve le sommant de second rang) en analysant les décompositions :

- si $p+q$ est codée par une lettre a ou une lettre b, cela correspond aux deux cas possibles dans lesquels q est premier, alors la décomposition $3+q$, contenant deux nombres premiers sera codée par la lettre a ;
- si $p+q$ est codée par une lettre c ou une lettre d, cela correspond aux deux cas possibles dans lesquels q est composé, alors la décomposition $3+q$, de la forme *premier* + *composé* sera codée par la lettre c.

On utilisera également dans la section suivante une projection qui transforme le sommant de premier rang en sommant de second rang que l'on combine à 3 comme sommant de premier rang ; analysons l'effet qu'une telle projection aura sur les décompositions :

- si $p+q$ est codée par une lettre a ou une lettre c, cela correspond aux deux cas possibles dans lesquels p est premier, alors la décomposition $3+p$, contenant deux nombres premiers sera codée par la lettre a ;
- si $p+q$ est codée par une lettre b ou une lettre d, cela correspond aux deux cas possibles dans lesquels p est composé, alors la décomposition $3+p$, de la forme *premier* + *composé* sera codée par la lettre c.

3 Dénombrements

1) On note dans la ligne n par :

- $X_a(n)$ le nombre de décompositions de n de la forme *premier* + *premier* ;
- $X_b(n)$ le nombre de décompositions de n de la forme *composé* + *premier* ;
- $X_c(n)$ le nombre de décompositions de n de la forme *premier* + *composé* ;
- $X_d(n)$ le nombre de décompositions de n de la forme *composé* + *composé*.

$X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \frac{n-2}{4} \right\rfloor$ est le nombre d'éléments de la ligne de n.

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

2) Soit $Y_a(n)$ (resp. $Y_c(n)$) le nombre de a (resp. c) qui apparaissent dans $B_{n,3}$. On rappelle qu'il n'y a que des lettres a et c dans la première colonne car elle contient les lettres associées aux décompositions de la forme $3+x$ et que 3 est premier.

Exemple :

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

3) Compte-tenu de la projection P qui est une bijection, des définitions des lettres a, b, c, d, $Y_a(n) = X_a(n) + X_b(n)$ et $Y_c(n) = X_c(n) + X_d(n)$. Par suite, trivialement, $Y_a(n) + Y_c(n) = X_a(n) + X_b(n) + X_c(n) + X_d(n) = \left\lfloor \frac{n-2}{4} \right\rfloor$.

Exemple :

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

4) Soit $Z_a(n)$ (resp. $Z_c(n)$) le nombre de a (resp. c) qui apparaissent dans $H_{n,3}$.

Exemple :

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

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

Rappel des propriétés identifiées

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

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

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

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

Ajoutons quatre nouvelles propriétés à celles-ci :

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

avec $\delta_{2p}(n)$ qui vaut 1 dans le cas où n est le double d'un nombre premier et qui vaut 0 sinon.

$$X_b(n) + X_d(n) = Z_c(n) + \delta_{2c-imp}(n) \tag{6}$$

avec $\delta_{2c-imp}(n)$ qui vaut 1 dans le cas où n est un double d'impair composé, et qui vaut 0 sinon.

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

avec $\delta_{4k+2}(n)$ qui vaut 1 si n est un double d'impair et 0 sinon.

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

4 Evolution des variables

Dans cette section, étudions comment les différentes variables évoluent.

$Z_a(n) + Z_c(n) = \left\lfloor \dfrac{n-4}{4} \right\rfloor$ est une fonction croissante de n, elle augmente de 1 à chaque n double de pair.

$Z_a(n+2) = Z_a(n)$ et $Z_c(n+2) = Z_c(n)$ sont fixes quand n est un double de pair.

$Z_a(n) = Z_a(n-2) + 1$ quand $\dfrac{n-2}{2}$ est un nombre premier (*ex* : $n = 24$ ou $n = 28$, se reporter au tableau de valeurs page 13 : on exprime cela abusivement par *"Z_a augmente"*) et $Z_c(n) = Z_c(n-2) + 1$ quand $\dfrac{n-2}{2}$ est un nombre composé impair (*ex* : $n = 42$ ou 50, abusivement, *"Z_c augmente"*).

$Y_a(n)+Y_c(n) = X_a(n)+X_b(n)+X_c(n)+X_d(n) = \left\lfloor \dfrac{n-2}{4} \right\rfloor$ est une fonction croissante de n, elle augmente de 1 à chaque n double d'impair.

Voyons maintenant dans le détail comment évoluent $Y_a(n)$ et $Y_c(n)$.

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), $Y_a(n) = Y_a(n-2) + 1$ (abusivement *"Y_a augmente"*, *ex* : $n = 34$) (resp. $Y_c(n) = Y_c(n-2) + 1$, abusivement *"Y_c augmente"*, *ex* : $n = 38$).

Dans le cas où n est un double de pair, 4 cas se présentent. Etudions la manière dont évolue l'ensemble de décompositions $H_{n,3}$.

- si $n-3$ et $\dfrac{n-2}{2}$ sont tous les deux premiers, on retire en bas et et on ajoute en haut de l'intervalle $H_{n,3}$ deux lettres de même sorte, alors $Y_a(n) = Y_a(n-2)$ et $Y_c(n) = Y_c(n-2)$ (abusivement *"Y_a et Y_c constants"* (*ex* : $n = 40$)) ;
- si $n-3$ est premier et $\dfrac{n-2}{2}$ est composé alors $Y_a(n) = Y_a(n-2) + 1$ et $Y_c(n) = Y_c(n-2) - 1$ (abusivement, *"Y_a augmente et Y_c diminue"* (*ex* : $n = 32$)) ;
- si $n-3$ est composé et $\dfrac{n-2}{2}$ est premier alors $Y_c(n) = Y_c(n-2) + 1$ et $Y_a(n) = Y_a(n-2) - 1$ (abusivement, *"Y_a diminue et Y_c augmente"* (*ex* : $n = 48$)) ;
- si $n-3$ et $\dfrac{n-2}{2}$ sont tous les deux composés, on retire en bas et et on ajoute en haut de l'intervalle $H_{n,3}$ deux lettres de même sorte, alors $Y_a(n) = Y_a(n-2)$ et $Y_c(n) = Y_c(n-2)$ (abusivement, *"Y_a et Y_c constants"* (*ex* : $n = 52$)).

En annexe 1 sont fournies dans une table les valeurs des différentes variables pour n compris entre 14 et 100.

5 Utiliser les écarts entre variables

On va montrer ci-dessous que $X_a(n)$ ne peut jamais être nul pour $n \geqslant C$, C étant une certaine constante à définir, i.e. montrer que tout entier pair $n \geqslant C$ peut s'écrire comme une somme de deux nombres premiers, c'est-à-dire vérifie la conjecture de Goldbach.

On a vu qu'à $\delta_{4k+2}(n)$ ou $\delta_{2c-imp}(n)$ près ($\delta_{4k+2}(n)$ et/ou $\delta_{2c-imp}(n)$ étant égaux à 1 dans certains cas), on a les égalités suivantes :

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

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

On rappelle que

- $Y_a(n)$, comptant le nombre de nombres premiers qui sont compris entre $\frac{n}{2}$ et n vaut $\pi(n) - \pi\left(\frac{n}{2}\right)$;
- $Z_a(n)$ comptant le nombre de nombres premiers inférieurs ou égaux à $\frac{n}{2}$ vaut $\pi\left(\frac{n}{2}\right)$;
- $Z_c(n)$ comptant le nombre de nombres composés impairs inférieurs ou égaux à $\frac{n}{2}$ vaut $\frac{n}{4} - \pi\left(\frac{n}{2}\right)$;
- $Y_c(n)$ comptant le nombre de nombres composés impairs qui sont compris entre $\frac{n}{2}$ et n vaut $\frac{n}{4} - \pi(n) + \pi\left(\frac{n}{2}\right)$.

L'article de Rosser et Schoenfeld [1] fournit comme formule 3.5 du corollaire 1 du théorème 2 que $\pi(x) > \frac{x}{ln\ x}$ pour tout $x \geqslant 17$, et comme formule 3.6 du même corollaire du même théorème que $\pi(x) < \frac{1,25506x}{ln\ x}$ pour tout $x > 1$.

5.1 Etude de l'inégalité $Z_c(n) > Z_a(n)$

Pour montrer que $Z_c(n) > Z_a(n)$, on peut utiliser simplement l'argument que $Z_c(n)$ augmente "bien plus souvent" que $Z_a(n)$ (à chaque $\frac{n-2}{2}$ composé impair pour $Z_c(n)$ et à chaque $\frac{n-2}{2}$ premier pour $Z_a(n)$ comme cela a été vu dans la section 4).

Pour trouver à partir de quelle valeur de n on a $Z_c(n) > Z_a(n)$, on utilise le fait que l'écart $Z_c(n) - Z_a(n)$ est égal à $\frac{n}{4} - 2\ \pi\left(\frac{n}{2}\right)$.

Par la formule (3.6), on a $2\ \pi\left(\frac{n}{2}\right) < 2\ \frac{1,25506\ n}{2\ (ln\ n + ln\ 0.5)}$ pour tout $n > 2$.

On en déduit $-2\ \pi\left(\frac{n}{2}\right) > \frac{-1,25506\ n}{ln\ n + ln\ 0.5}$ pour tout $n > 2$.

L'écart $Z_c(n) - Z_a(n)$ est donc minorable par $\frac{n\ (ln\ n + ln\ 0.5) - 5,02024n}{4\ (ln\ n + ln\ 0.5)}$. Il est strictement positif pour tout $n \geqslant 304$ (le dénominateur est positif pour $n \geqslant 2$, le numérateur est strictement positif pour tout $n > 2e^{5.02024}$).

5.2 Etude des inégalités $Z_a(n) > Y_a(n)$ et $Y_c(n) > Z_c(n)$

Pour montrer que $Z_a(n) > Y_a(n)$, on peut utiliser à nouveau l'analyse de l'évolution des variables de la section 4 : quand *"Z_a augmente"*, *"Y_a est constante ou diminue"* ; et quand *"Y_a augmente"* sans que *"Z_a n'augmente"* (lorsque $n-3$ est premier et $\frac{n-2}{2}$ est composé), Y_a n'augmente que de 1 alors que son écart à Z_a est très vite très supérieur à 1.

Pour démontrer que $Y_c(n) > Z_c(n)$, on peut utiliser à nouveau l'analyse de l'évolution des variables de la section 4 : Y_c augmente quand $n-3$ est composé tandis que Z_c augmente quand $\dfrac{n-2}{2}$ est composé. Z_c est croissante, parfois Y_c décroît mais peu souvent ce qui a pour conséquence qu'à partir d'une valeur plutôt petite, Z_c ne "rattrape plus" Y_c.

Pour connaître précisément à partir de quelles valeurs de n les inégalités souhaitées sont vérifiées, on utilise à nouveau la valeur des écarts et les minorations/majorations fournies dans [1].

Pour montrer que $Z_a(n) > Y_a(n)$ (resp. $Y_c(n) > Z_c(n)$), on montre que l'écart

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

est toujours strictement positif.

On utilise la formule 3.9 du corollaire 3 du théorème 2 de Rosser et Schoenfeld qui énonce que $\pi(x) - \pi\left(\dfrac{x}{2}\right) < \dfrac{7x}{5\ ln\ x}$ pour tout $x > 1$.

On utilise le fait que $2\pi\left(\dfrac{n}{2}\right) - \pi(n) = \left(\pi\left(\dfrac{n}{2}\right) - \pi(n)\right) + \pi\left(\dfrac{n}{2}\right)$

On a donc

$$\begin{aligned} 2\pi\left(\frac{n}{2}\right) - \pi(n) &> \frac{-7n}{5\ ln\ n} + \pi\left(\frac{n}{2}\right) \\ &> \frac{-7n}{5\ ln\ n} + \frac{n}{2\ (ln\ n + ln\ 0.5)} \qquad \text{(à cause de la formule 3.5 du corollaire 1 du théorème 2 de [1])} \end{aligned}$$

qui est strictement positif si

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

ce qui équivaut à

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

qui est toujours vrai pour $n \geqslant 6$.

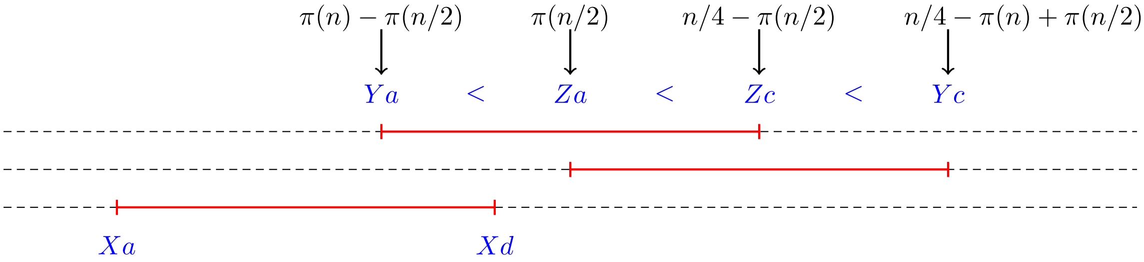
5.3 Ordre strict sur les 4 variables $Y_a(n), Y_c(n), Z_a(n)$ et $Z_c(n)$

Les variables $Y_a(n), Z_a(n), Z_c(n)$ et $Y_c(n)$ sont donc strictement ordonnées de la façon suivante :

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

pour tout $n \geqslant 304$.

Une représentation imagée des écarts entre variables est fournie ci-dessous, qui montre leur intrication :



Les écarts $Z_c(n) - Y_a(n)$, $Y_c(n) - Z_a(n)$ et $X_d(n) - X_a(n)$ sont strictement positifs et égaux à $\dfrac{n}{4} - \pi(n)$.

5.4 Etude de l'inégalité $X_a(n) > 0$

Pour être assuré que $X_a(n)$ soit toujours non nul, il faut minorer $X_d(n)$ par $\dfrac{n}{4} - \pi(n)$, i.e. la valeur de l'écart $X_d(n) - X_a(n)$.

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

Pour minorer $X_d(n)$, il faut minorer $Y_c(n)$ et majorer $X_c(n)$.

$Y_c(n)$ est le nombre de nombres composés impairs compris entre $n/2$ et n (associés à 3).

Pour minorer $Y_c(n)$, on utilise le fait que le nombre de nombres composés impairs compris entre $n/2$ et n est égal à $\dfrac{n}{4} - \pi(n) + \pi\left(\dfrac{n}{2}\right)$.

$X_c(n)$, le nombre de décompositions de n de la forme *premier + composé* est majorable par le nombre total de nombres composés impairs compris entre $n/2$ et n, lui-même majorable par le nombre de nombres composés impairs compris entre n et $\dfrac{3n}{2}$.

$X_c(n)$ est donc majorable par $\dfrac{n}{4} - \left(\pi\left(\dfrac{3n}{2}\right) - \pi(n)\right)$ (le nombre de nombres impairs de l'intervalle de n à $\dfrac{3n}{2}$ est $\dfrac{n}{4}$, le nombre de nombres premiers sur cet intervalle est $\pi\left(\dfrac{3n}{2}\right) - \pi(n)$, le nombre de nombres composés impairs sur cet intervalle est la différence de ces deux nombres).

$Y_c(n) - X_c(n)$ est donc toujours supérieur à la différence de la minoration de $Y_c(n)$ et de la majoration de $X_c(n)$, ce qui donne

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

$Y_c(n) - X_c(n) = X_d(n)$ dépasse toujours $\dfrac{n}{4} - \pi(n)$ lorsque $n > 0$ (on le constate également par programme).

En effet, $Y_c(n) - X_c(n) > \dfrac{n}{4} - \pi(n)$ lorsque

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

On remplace $\pi\left(\dfrac{n}{2}\right)$ par sa minoration fournie par la formule 3.5 du corollaire 1 du théorème 2 (qui est $\dfrac{n}{2(ln\ n + ln\ 0.5)}$), on réduit au même dénominateur, qui est toujours positif pour $n \geqslant 2$ et qu'on oublie, on cherche la condition qui assure que le numérateur est toujours strictement positif, numérateur qui est égal à :

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

Après calculs, on obtient que le numérateur, d'inconnue $ln\ n$, est égal au polynôme

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

La plus grande racine de ce polynôme est à peu près égale à 14.502656936497 d'exponentielle 1988034.33365. La différence entre $X_d(n)$ et $X_a(n)$ est donc toujours supérieure à $\dfrac{n}{4} - \pi(n)$ pour tout $n \geqslant 1988034.33365$.

On peut ainsi conclure que pour tout $n \geqslant 1988034.33365$ (la condition nécessitée pour que $X_d(n)-X_a(n) > \dfrac{n}{4} - \pi(n)$), $X_a(n)$ (le nombre de décompositions de n comme somme de deux nombres premiers) est strictement positif.

En annexe 2 sont fournies des représentations graphiques des bijections ensemblistes pour les cas $n = 32$, 34, 98 et 100.

Le fichier $http : //denise.vella.chemla.free.fr/annexes.pdf$ fournit

- un rappel historique d'un texte de Laisant qui présentait déjà en 1897 l'idée de "bandes" de nombres impairs à mettre en regard et à colorer pour voir les décompositions de Goldbach ;
- un programme et son exécution qui implémente les idées exposées dans la présente note.

6 Démonstrations

6.1 Utilitaires

Démontrons que si n est un double d'impair (i.e. de la forme $4k+2$), alors $\left\lfloor \dfrac{n-2}{4} \right\rfloor = \left\lfloor \dfrac{n-4}{4} \right\rfloor + 1$.
En effet, le membre gauche de l'égalité vaut $\left\lfloor \dfrac{(4k+2)-2}{4} \right\rfloor = \left\lfloor \dfrac{4k}{4} \right\rfloor = k$.
Le membre droit de l'égalité vaut $\left\lfloor \dfrac{(4k+2)-4}{4} \right\rfloor + 1 = \left\lfloor \dfrac{4k-2}{4} \right\rfloor + 1 = (k-1)+1 = k$.

Démontrons que si n est un double de pair (i.e. de la forme $4k$), alors $\left\lfloor \dfrac{n-2}{4} \right\rfloor = \left\lfloor \dfrac{n-4}{4} \right\rfloor$.
$\left\lfloor \dfrac{4k-2}{4} \right\rfloor = k-1$ et $\left\lfloor \dfrac{4k-4}{4} \right\rfloor = k-1$.

On peut aussi exprimer cela de la façon suivante : si n est un double d'impair, $\left\lfloor \dfrac{n-2}{4} \right\rfloor = \dfrac{n-2}{4} = \left\lfloor \dfrac{n-4}{4} \right\rfloor + 1$ tandis que si n est un double de pair, $\left\lfloor \dfrac{n-2}{4} \right\rfloor = \dfrac{n-4}{4} = \left\lfloor \dfrac{n-4}{4} \right\rfloor$.

6.2 Propriétés 5, 6 et 8

Les propriétés 5, 6 et 8 découlent directement des définitions des variables.

6.2.1 Propriété 5

La propriété 5 énonce que $X_a(n) + X_c(n) = Z_a(n) + \delta_{2p}(n)$ avec $\delta_{2p}(n)$ qui vaut 1 dans le cas où n est le double d'un nombre premier et qui vaut 0 sinon.

Par définition, $X_a(n) + X_c(n)$ compte le nombre de décompositions de n de la forme $premier + x$ avec $premier \leqslant n/2$. Dans la mesure où $Z_a(n)$ compte quant à lui le nombre de décompositions de la forme $3 + premier$ avec $premier < n/2$, l'ajout de δ_{2p} à $Z_a(n)$ permet d'assurer l'invariance de l'égalité, entre autres dans le cas où n est le double d'un nombre premier.

6.2.2 Propriété 6

La propriété 6 énonce que $X_b(n) + X_d(n) = Z_c(n) + \delta_{2c-imp}$ avec δ_{2c-imp} qui vaut 1 dans le cas où n est un double d'impair composé, et qui vaut 0 sinon.

Par définition, $X_b(n) + X_d(n)$ compte le nombre de décompositions de n de la forme $composé + x$ avec $composé \leqslant n/2$. Dans la mesure où $Z_c(n)$ compte quant à lui le nombre de décompositions de la forme $3 + composé$ avec $composé < n/2$, l'ajout de δ_{2c-imp} à $Z_c(n)$ permet d'assurer l'invariance de l'égalité, entre autres dans le cas où n est le double d'un nombre composé impair.

6.2.3 Propriété 8

La propriété 8 énonce que $Z_c(n) - Y_a(n) = X_d(n) - X_a(n) - \delta_{2c-imp}$ avec δ_{2c-imp} qui vaut 1 si n est un double de composé impair et 0 sinon.

Par définition, $Z_c(n)$ compte le nombre de nombres composés impairs inférieurs strictement à $n/2$. Il compte donc également le nombre de décompositions de n de la forme $compos\acute{e} + x$ avec $compos\acute{e} < n/2$ (appelons E cet ensemble de décompositions).

Par définition, $Y_a(n)$ compte le nombre de nombres premiers supérieurs strictement à $n/2$. Il compte donc également le nombre de décompositions de n de la forme $x + premier$ avec $premier > n/2$ (appelons F cet ensemble de décompositions).

Les décompositions de n de la forme $compos\acute{e} + premier$ sont à la fois dans E et dans F. En calculant $Z_c(n) - Y_a(n)$, on calcule donc le cardinal d'un ensemble qui s'avère être égal à $X_d(n) - X_a(n)$ par définition de ce que comptent les variables $Y_a(n), Z_c(n), X_d(n)$ et $X_a(n)$.

6.2.4 Propriété 7

Démontrons que $Z_c(n) - Y_a(n) = Y_c(n) - Z_a(n) - \delta_{4k+2}$ avec δ_{4k+2} qui vaut 1 si n est un double d'impair (il existe $k \geqslant 3$ tel que $n = 4k + 2$) et 0 sinon.

On utilise un raisonnement par récurrence :

i) On initialise les récurrences selon les 3 sortes de nombres à envisager : les doubles de pairs (de la forme $4k$, comme 16), les doubles d'impairs (donc de la forme $4k + 2$) qui sont premiers (comme 14) ou qui sont composés (comme 18).

La propriété 7 est vraie pour $n = 14$ car $Z_c(14) = 0, Y_a(14) = 2, Y_c(14) = 1, Z_a(14) = 2$ et $\delta_{4k+2}(14) = 1$ et donc $Z_c(14) - Y_a(14) = Y_c(14) - Z_a(14) - \delta_{4k+2}(14)$;
La propriété 7 est vraie pour $n = 16$ car $Z_c(16) = 0, Y_a(16) = 2, Y_c(16) = 1, Z_a(16) = 3$ et $\delta_{4k+2}(16) = 0$ et donc $Z_c(16) - Y_a(16) = Y_c(16) - Z_a(16) - \delta_{4k+2}(16)$;
La propriété 7 est vraie pour $n = 18$ car $Z_c(18) = 0, Y_a(18) = 2, Y_c(18) = 2, Z_a(18) = 3$ et $\delta_{4k+2}(18) = 1$ et donc $Z_c(18) - Y_a(18) = Y_c(18) - Z_a(18) - \delta_{4k+2}(18)$;

ii) On réécrit la propriété 7 sous la forme $Z_a(n) + Z_c(n) + \delta_{4k+2} = Y_a(n) + Y_c(n)$.
Quatre cas sont à considérer : deux cas selon lesquels n est un double d'impair (premier ou composé) et $n + 2$ est un double de pair et deux cas selon lesquels n est un double de pair et $n + 2$ est un double d'impair (premier ou composé).

iia) *n double de pair et $n + 2$ double de premier (ex : $n = 56$)* :

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

On pose l'hypothèse que la propriété 7 est vraie pour n,

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

Démontrons qu'elle est vraie pour $n + 2$,

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

On a $Z_a(n + 2) = Z_a(n)$ et $Z_c(n + 2) = Z_c(n)$.

Rappelons la propriété 3 concernant les Y :

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

Dans (Ccl), on peut, par l'hypothèse de récurrence et par la propriété (3), remplacer le membre gauche de l'égalité par $Z_a(n) + Z_c(n) + 1$ puis par $Y_a(n) + Y_c(n) + 1$ (par (H)) puis par $\left\lfloor \frac{n-2}{4} \right\rfloor + 1$ (par (3)) qui vaut $\left\lfloor \frac{n}{4} \right\rfloor$.

Mais dans (Ccl), on peut également remplacer le membre de droite de l'égalité par $\left\lfloor \frac{n}{4} \right\rfloor$ du fait de la propriété (3).

Il y a à nouveau pour $n+2$ égalité entre les membres gauche et droit de l'égalité, i.e. la propriété 7 est vérifiée par $n+2$. De l'hypothèse que la propriété est vraie pour n, on a démontré que la propriété est vraie pour $n+2$.

iib) *n double de pair et* $n+2$ *double de composé impair* (*ex :* $n = 48$) :

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

On pose l'hypothèse que la propriété 7 est vraie pour n,

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

Démontrons qu'elle est vraie pour $n+2$,

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

On a $Z_a(n+2) = Z_a(n)$ et $Z_c(n+2) = Z_c(n)$.

On a $Y_a(n+2) = Y_a(n) + 1$ et $Y_c(n+2) = Y_c(n)$.

Rappelons la propriété 3 concernant les Y :

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

Dans (Ccl), on peut, par l'hypothèse de récurrence et la propriété (3), remplacer le membre gauche de l'égalité par $Z_a(n)+Z_c(n)+1$ puis par $Y_a(n)+Y_c(n)+1$ (par (H)) puis par $\left\lfloor \frac{n-2}{4} \right\rfloor+1$ (par (3)) qui vaut $\left\lfloor \frac{n}{4} \right\rfloor$.

Mais dans (Ccl), on peut également remplacer le membre de droite de l'égalité par $\left\lfloor \frac{n}{4} \right\rfloor$ du fait des évolutions de $Y_a(n)$ et $Y_c(n)$.

Il y a à nouveau pour $n+2$ égalité entre les membres gauche et droit de l'égalité, i.e. la propriété 7 est vérifiée par $n+2$. De l'hypothèse que la propriété est vraie pour n, on a démontré que la propriété est vraie pour $n+2$.

iic) *n double de premier et* $n+2$ *double de pair* (*ex :* $n = 74$) :

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

On pose l'hypothèse que la propriété 7 est vraie pour n,

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

Démontrons qu'elle est vraie pour $n+2$,

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

On a $Z_a(n+2) = Z_a(n) + 1$ et $Z_c(n+2) = Z_c(n)$.

Rappelons la propriété 3 concernant les Y :

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

Dans (Ccl), on peut, par l'hypothèse de récurrence et la propriété (3), remplacer le membre gauche de l'égalité par $Z_a(n) + Z_c(n) + 1$ puis par $Y_a(n) + Y_c(n)$ (par (H)) puis par $\left\lfloor \frac{n-2}{4} \right\rfloor$ (par (3)) qui vaut $\left\lfloor \frac{n}{4} \right\rfloor$.

Mais dans (Ccl), on peut également remplacer le membre de droite de l'égalité par $\left\lfloor \frac{n}{4} \right\rfloor$ du fait de la propriété (3).

Il y a à nouveau pour $n+2$ égalité entre les membres gauche et droit de l'égalité, i.e. la propriété 7 est vérifiée par $n+2$. De l'hypothèse que la propriété est vraie pour n, on a démontré que la propriété est vraie pour $n+2$.

iid) *n double de composé impair et $n+2$ double de pair (ex : $n = 70$) :*

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

On pose l'hypothèse que la propriété 7 est vraie pour n,

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

Démontrons qu'elle est vraie pour

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

On a $Z_a(n+2) = Z_a(n)$ et $Z_c(n+2) = Z_c(n) + 1$.

Rappelons la propriété 3 concernant les Y :

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

Dans (Ccl), on peut, par l'hypothèse de récurrence et la propriété (3), remplacer le membre gauche de l'égalité par $Z_a(n) + Z_c(n) + 1$ puis par $Y_a(n) + Y_c(n)$ (par (H)) puis par $\left\lfloor \frac{n-2}{4} \right\rfloor$ (par (3)) qui vaut $\left\lfloor \frac{n}{4} \right\rfloor$.

Mais dans (Ccl), on peut également remplacer le membre de droite de l'égalité par $\left\lfloor \frac{n}{4} \right\rfloor$ du fait de la propriété (3).

Il y a à nouveau pour $n+2$ égalité entre les membres gauche et droit de l'égalité, i.e. la propriété 7 est vérifiée par $n+2$. De l'hypothèse que la propriété est vraie pour n, on a démontré que la propriété est vraie pour $n+2$.

Annexe 1 : tableau de valeurs des variables pour n compris entre 14 et 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

Annexe 2 : bijections ensemblistes

- Cas $n = 32$

$a, c \to a$

$b, d \to c$

$a, b \to a$

$c, d \to c$

$3+3$ a	$3+29$ a	$3+29$ a
$3+5$ a	$5+27$ c	$3+27$ c
$3+7$ a	$7+25$ c	$3+25$ c
$3+9$ c	$9+23$ b	$3+23$ a
$3+11$ a	$11+21$ c	$3+21$ c
$3+13$ a	$13+19$ a	$3+19$ a
$3+15$ c	$15+17$ b	$3+17$ a

$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

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

$Y_a = X_a + X_b = 4$

$Y_c = X_c + X_d = 3$

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

$X_a = 2$

$X_b = 2$

$X_c = 3$

$X_d = 0$

- Cas $n = 34$

$a, c \to a$

$b, d \to c$

$a, b \to a$

$c, d \to c$

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

$Z_a = X_a + X_c = 5$

$Z_c = X_b + X_d = 2$

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

$Y_a = X_a + X_b = 5$

$Y_c = X_c + X_d = 3$

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

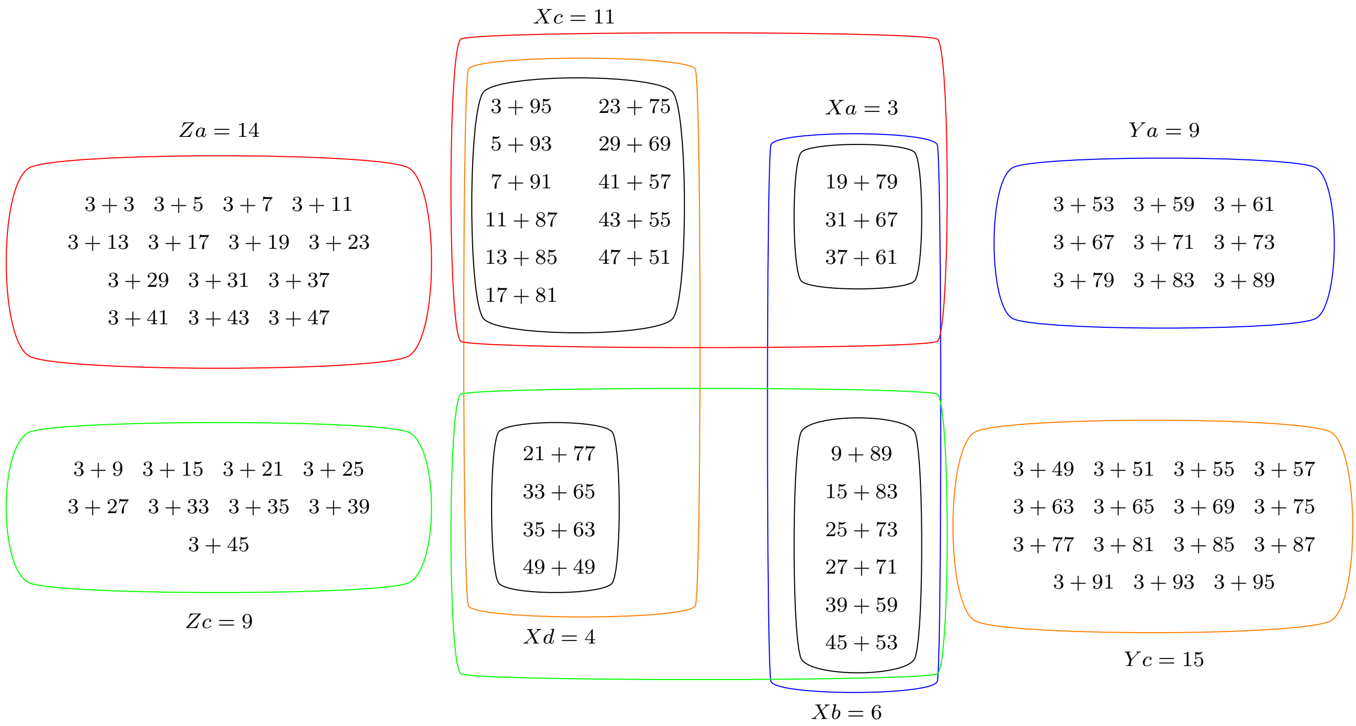
$X_a = 4$

$X_b = 1$

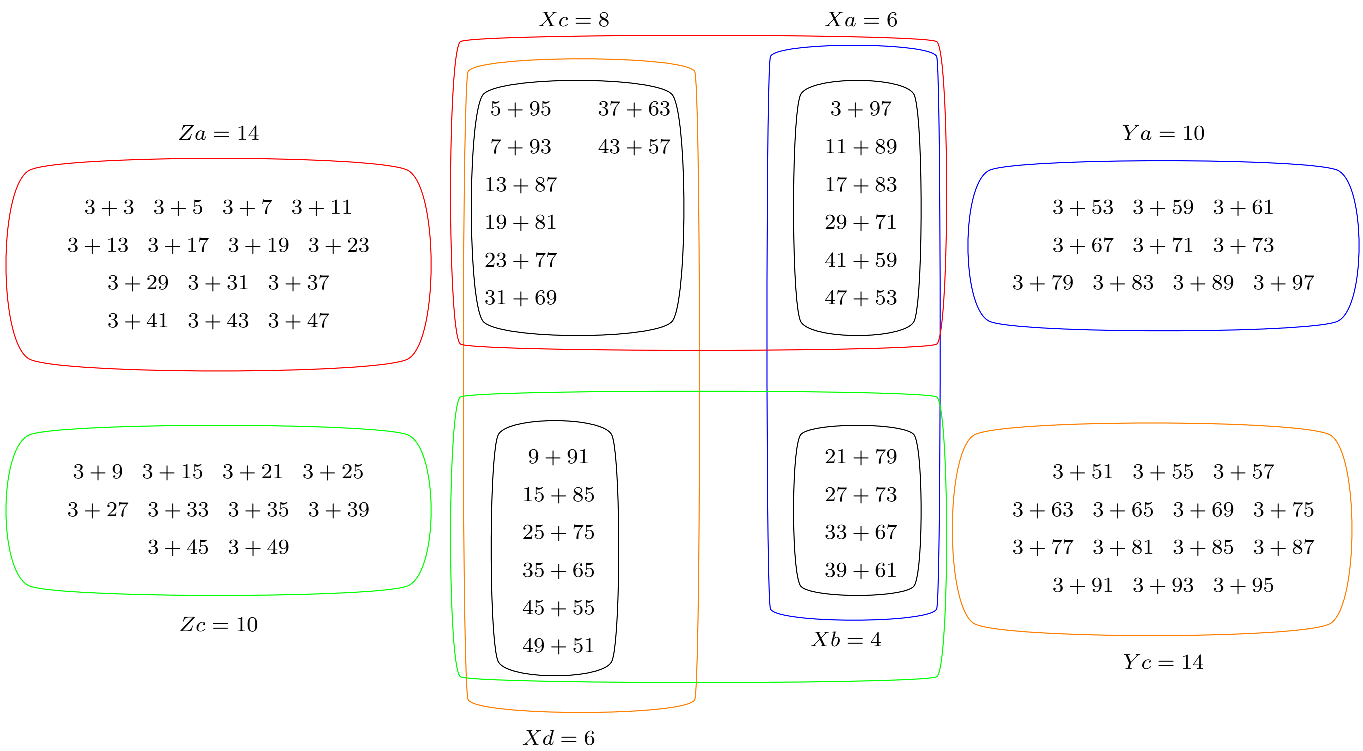
$X_c = 2$

$X_d = 1$

- Cas $n = 98$



- Cas $n = 100$



Annexe 3 : Règles de réécriture et théorie des automates

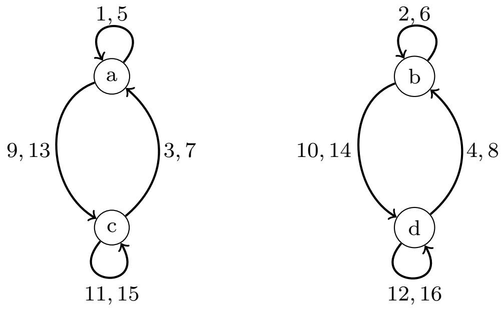
Cette annexe étudie l'évolution des variables $X_a(n), X_b(n), X_c(n), X_d(n)$ que l'on déduit de l'analyse des règles de réécriture des mots, présentées du point de vue de la théorie des automates.

Si on considère chaque ligne du tableau comme un mot sur l'alphabet $A = \{a, b, c, d\}$, le mot du nombre pair $n+2$ s'obtient de la façon suivante à partir du mot de n :

- la première lettre du mot de $n+2$ est a si $n-3$ est premier et c sinon (cette première lettre est la seule qui introduit de l'indéterminisme car elle n'appartient pas au mot de n ou ne se déduit pas des lettres de ce dernier) ;
- les lettres suivantes du mot de $n+2$ sont obtenues par réécriture du mot de n selon les règles ci-dessous :

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

  On peut représenter ces règles de réécriture par les deux automates déterministes suivants, dont les arcs sont étiquetés par les règles applicables à une lettre donnée du mot de n :



- enfin, la concaténation d'une lettre en fin de mot, dans le cas où n est un double de pair (i.e. de la forme $4k$) obéit à la règle suivante :
  - si le mot de n a a ou b comme dernière lettre, après avoir obtenu le mot de $n+2$ en appliquant les règles de réécriture, on lui concatène la lettre a ;
  - si le mot de n a c ou d comme dernière lettre, après avoir obtenu le mot de $n+2$ en appliquant les règles de réécriture, on lui concatène la lettre d.

Si on prend comme convention de noter $X_{xy}(n)$ le nombre d'occurences de la séquence de lettres xy dans le mot de n, les égalités suivantes fournissent l'évolution des nombres de lettres a, b, c ou d lors du passage du mot de n au mot de $n+2$.

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

avec $\delta_a(n)$ qui vaut 1 si n est un double de pair (i.e. de la forme $4k$) et si la dernière lettre du mot de n est une lettre a ou b, $\delta_d(n)$ qui vaut 1 si n est un double de pair (i.e. de la forme $4k$) et si la dernière lettre du mot de n est une lettre c ou d et enfin $\delta_{n-3}(n)$ qui vaut 1 si $n-3$ est premier et 0 sinon.

Bibliographie

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