Piste pour une démonstration de la conjecture de Goldbach

Denise Vella-Chemla

27 Avril 2009

1 Introduction

Dans une lettre à Euler du 7 juin 1742, Goldbach énonce *"il semble que tout nombre supérieur à* 2 *soit la somme de trois nombres premiers"*. Euler reformule cette conjecture en une forme équivalente qui est *"tout nombre entier naturel pair supérieur à 2 est la somme de deux nombres premiers"*.

2 Définitions

Dans la suite de cette note, on s'intéresse à une restriction de la conjecture qui est *"tout nombre pair supérieur ou égal à 6 est la somme de deux nombres premiers impairs"*. Les nombres pairs étant trivialement composés, on ne considèrera que la divisibilité par des nombres impairs.

Le travail présenté ici se situe dans la théorie des langages. Les définitions sont empruntées à Sakarovitch ([2]).
Considérons l'alphabet binaire $\mathcal{A} = \{0, 1\}$.
Les éléments 0 et 1 de $\mathcal{A}$ sont appelés *lettres*.
Les suites finies de lettres sont appelées *mots*.
L'ensemble des mots est noté $\mathcal{A}^*$.
On appelle *langage sur* $\mathcal{A}$ ou *langage de* $\mathcal{A}^*$ tout ensemble de mots écrits sur l'alphabet $\mathcal{A}$.
L'ensemble des mots est naturellement muni d'une opération binaire appelée la *concaténation* :

$$(a_1a_2...a_m).(b_1b_2...b_n) = (a_1a_2...a_mb_1b_2...b_n)$$

Cette opération est associative, possède un élément neutre, le mot vide. On l'appellera *produit* et on appellera donc *puissance* l'itération de ce produit plusieurs fois.
Muni de la concaténation, $\mathcal{A}^*$ est un monoïde.
La concaténation n'est pas commutative puisque $\mathcal{A}$ contient deux lettres.

$$\textit{Exemple :} \quad (010)(11001) \neq (11001)(010).$$

La *longueur d'un mot* est le nombres de ses lettres.
Soient f et g des mots de $\mathcal{A}^*$; g est un *facteur gauche* ou *préfixe* de f s'il existe

h tel que $f = gh$; g est *facteur gauche propre* ou préfixe propre si h est différent du mot vide. Autrement dit, g est un préfixe de f si f "commence par" g.
Si $f = a_1a_2...a_m$ est un mot de $\mathcal{A}^*$, on appelle *image-miroir* de f le mot f^t :

$$f^t = a_m a_{m-1}...a_1$$

Un mot qui est égal à son image-miroir est appelé *palindrome.*

Exemple : (000010010000) *est un palindrome.*

Dans toute la suite, nous omettrons les parenthèses autour des mots car nous avons besoin de rechercher aisément l'existence de plusieurs 0 à des positions communes dans des puissances de mots et l'usage des parenthèses ne permettrait pas une visualisation immédiate de telles propriétés car elle décalerait les lettres des différents mots les unes par rapport aux autres.

La lettre 1 sera utilisée dans les mots pour représenter d'une manière générale le fait pour un nombre d'être composé parce que divisible par un nombre impair plus petit que lui et différent de 1. Inversement, la lettre 0 sera utilisée dans les mots pour représenter le fait pour un nombre d'être non divisible par un nombre impair plus petit que lui et différent de 1.

3 Caractère symétrique de la conjecture de Goldbach

Pour résoudre notre problème, nous allons le décomposer en sous-problèmes, résoudre les sous-problèmes puis "agréger" entre elles les résolutions des sous-problèmes. Nous allons d'abord nous préoccuper de la divisibilité par 3, puis par 5, puis successivement par tous les nombres impairs inférieurs à x ($2x$ est le nombre pair dont on cherche une décomposition de Goldbach) et nous comprendrons pourquoi un nombre pair a toujours une telle décomposition[1], conditionnée par les différents caractères de divisibilité invoqués.

Etudions le caractère symétrique de la conjecture de Goldbach (un procédé de vérification expérimental de cette conjecture est présenté dans Laisant [1]). Pour cela, voyons l'exemple du nombre pair 40. Dans le tableau suivant, on place les nombres impairs inférieurs à 20 la moitié de 40, et ceux supérieurs à 20 en vis-à-vis dans un tableau quand leur somme vaut 40 ; on place également au-dessus (resp. au-dessous) des nombres supérieurs (resp. inférieurs) à 20 une lettre 0 ou 1 indiquant le fait qu'il soit divisible par 3.

1	0	0	1	0	0	1	0	0
3	5	7	9	11	13	15	17	19
37	35	33	31	29	27	25	23	21
0	0	1	0	0	1	0	0	1

On peut associer au nombre pair 40 pour la divisibilité par 3 le mot binaire suivant : 100100100100100100.
On voit que 9 ne peut pas être un décomposant de Goldbach de 40 puisqu'il est

[1] que nous appellerons sa décomposition "centrale" parce qu'elle minimise la distance des deux sommants à x.

divisible par 3. On voit également que 33 ne peut pas être un décomposant de Goldbach de 40 car il est divisible par 3.
9 et 33 parce qu'ils ne peuvent être des décomposants de Goldbach de 40 vont "entraîner avec eux" dans leur impossibilité d'être des décomposants de Goldbach de 40 leur complémentaire à 40, en l'occurrence 31 et 7.
On va représenter cela en "symétrisant" le mot binaire initial de la façon suivante :

1	0	1	1	0	1	1	0	1
3	5	7	9	11	13	15	17	19
37	35	33	31	29	27	25	23	21
1	0	1	1	0	1	1	0	1

En algèbre de Boole, on exprime cela en disant que 1 est absorbant pour l'opération "ou booléen" (symbolisée par $\vee$).

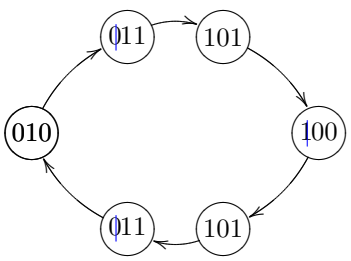
$$1 \vee 0 = 0 \vee 1 = 1 \vee 1 = 1.$$

On choisit maintenant d'associer à 40 pour la divisibilité par 3 le mot "symétrisé" suivant : 101101101101101101.

Ce mot étant un palindrome, on peut le représenter par son suffixe constitué de sa moitié droite, lue de gauche à droite, ou bien par son préfixe constitué de sa moitié gauche, lue de droite à gauche, qui sont deux mots égaux : 101101101.

On peut encore réduire le mot représentant 40 pour la divisibilité par 3 : c'est un mot périodique, la longueur de sa période la plus petite est 3. On conserve seulement 101 comme représentant de 40 en ce qui concerne la divisibilité par 3.

Lorsqu'on réitère cette analyse pour d'autres nombres pairs, on comprend que la représentation du caractère de divisibilité par 3 va associer aux nombres pairs $2x$ successifs les mots du cycle de 6 mots suivant après symétrisation, selon la classe de congruence à laquelle appartient x modulo 6 (selon le positionnement de leur "ligne de pli" sur le mot 100 : le pli peut être sur le 1, entre le 1 et le premier 0, sur le premier 0, entre les deux 0, sur le deuxième 0 ou entre le deuxième 0 et le 1 suivant du fait de la périodicité) :

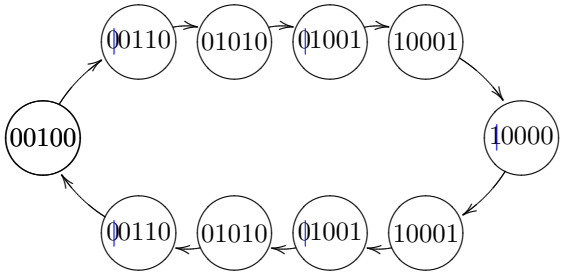


Les nombres pairs doubles de $6k$ se voient affecter le mot 010, les nombres pairs doubles de nombres de la forme $6k + 1$ se voient affecter le mot 011, et ainsi de suite jusqu'aux nombres pairs doubles de nombres de la forme $6k + 5$ qui se

voient affecter le mot 011.

Certains mots du cycle ont leur première lettre barrée : cela correspond au fait que le mot en question est associé à un nombre pair qui est le double d'un nombre impair (le nombre de nombres impairs supérieurs ou égaux à 3 et inférieurs ou égaux à $2x-3$ est impair).

Une étude similaire (et évidente) concernant le caractère de divisibilité par 5 nous fait aboutir au cycle contenant les 10 mots suivants :



Le cycle fournissant les mots d'une longueur impaire donnée *longueur* se déduit du cycle fournissant les mots de longueur le nombre impair précédent par un processus parfaitement déterminé (les mots ne commençant pas par une lettre barrée se voient concaténer un 0 au début et à la fin, les mots commençant par une lettre barrée se voient ajouter un 0 en position 2 et concaténer un 0 à la fin, et quatre sommets supplémentaires sont ajoutés au cycle à des positions intermédiaires à définir précisément ; à ces quatre sommets sont associés les deux mots suivants : 10^i1 et $010^{i-1}1$ avec i valant $longueur-2$).

Revenons au problème de Goldbach et à son caractère symétrique : un nombre impair est décomposant de Goldbach d'un nombre pair $2x$ donné s'il n'en est pas "empêché", de lui-même ou par son complémentaire à $2x$, sous prétexte qu'il serait composé car divisible par 3, ou bien composé car divisible par 5, ..., ou bien composé car divisible par tout nombre impair inférieur ou égal à x.

4 Exemple du nombre pair 98

$2x$	x	*mots associés*	*position du 0 commun*	*décomposition de Goldbach*
98	49	0110110 0011000 1000000 001000010	7	$(49-14+2)+(49+14-2)$ $37+61$

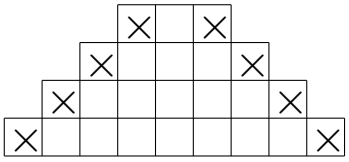
5 Affectation d'un ensemble de mots à un nombre pair

On associe au nombre pair $2x$ un ensemble de k mots de longueurs respectives $3, 5, ..., 2k+1$ où $2k+1$ est le plus grand impair inférieur ou égal à x.
On va ici distinguer deux cas selon que le nombre pair considéré est le double d'un nombre pair ou le double d'un nombre impair, i.e. considérer séparément les deux langages disjoints contenant des mots qui ont leur première lettre barrée ou pas.

- **Premier cas : nombres pairs double d'un nombre pair (tous leurs mots associés ont une initiale non barrée, sont des palindromes, et contiennent au moins une et au plus deux fois la lettre** 1)

  Trouver un décomposant de Goldbach de $2x$ consiste donc à trouver la position commune à laquelle les puissances des mots associés à $2x$, que l'on tronque pour que les mots résultant soient tous de longueur $2k+1$, contiennent tous une lettre 0. On peut parler de préfixes des puissances puisqu'on ne considère que le début du mot mais l'image de la troncature à la même longueur est plus parlante.

  Dans le pire des cas, le nombre de colonnes occupées par des 1 dans les k mots associés à $2x$ est $2k$, en positionnant les palindromes de la façon suivante :



  $2k$ étant strictement inférieur à $2k+1$, il reste forcément une colonne ne contenant que des 0 à une position inférieure ou égale à $2k+1$, le plus grand impair inférieur ou égal à x. En fait, on peut utiliser d'autres palindromes que ceux utilisés dans l'exemple de la tour de Hanoï de palindromes présentée ci-dessus et les positionner totalement autrement les uns par rapport aux autres, mais on voit difficilement comment on pourrait faire occuper un nombre de colonnes supérieur à $2k$ par des 1. Pour déduire de ce raisonnement que l'une des colonnes au moins ne contient que des 0, on utilise un principe en quelque sorte dual du principe des tiroirs[2] : ce principe énonce que si on a moins de n objets à ranger dans n tiroirs, l'un des tiroirs au moins sera vide.

[2]Le principe des tiroirs dit que si on a $n+1$ objets à ranger dans n tiroirs, l'un des tiroirs contient 2 objets.

- **Deuxième cas : nombres pairs double d'un nombre impair (tous leurs mots associés ont une initiale barrée, sont des palindromes si on les prive de leur initiale, et contiennent au moins une et au plus deux fois la lettre** 1**)**
  Si toutes les lettres barrées au début des différents mots associés à $2x$ sont des 0, on est dans le cas du double d'un nombre premier qui vérifie trivialement la conjecture de Goldbach.
  On notera donc que l'ensemble des nombres premiers est l'ensemble des nombres qui se voient affecter par notre fonction d'affectation des mots très spécifiques, commençant tous par un 0 barré.
  Si l'une des lettres barrées au début d'un des mots est un 1, on pourra peut-être utiliser le fait que les mots privés de leur première lettre sont des palindromes qui contiennent chacun soit exactement deux fois la lettre 1, soit uniquement des 0 pour à nouveau parvenir au résultat que l'une des colonnes au moins sur les $2k + 1$ colonnes ne contient que des 0.

Bibliographie

[1] C.A. Laisant, *Sur un procédé de vérification expérimentale du théorème de Goldbach*, Ed. Bulletin de la S.M.F., n°25, p.108, 1/12/1897.
[2] J. Sakarovitch, *Eléments de théorie des automates*, Ed. Vuibert informatique, 2003.