



Un mot infini irrégulier et pourtant simple

Jean-Paul Delahaye

*Professeur émérite à l'université de Lille-I,
Centre de recherche en informatique, signal et automatique
de Lille (Cristal, UMR CNRS 9189)*

Chaque matin, je prends du café ou du thé. Je trouve ennuyeux de prendre deux jours de suite la même chose, donc j'oscille : café, thé, café, thé, café, thé... Cette oscillation, trop régulière, m'ennuie aussi ! J'ai donc décidé de ne jamais faire deux fois de suite la même chose, et en particulier de ne jamais répéter deux fois de suite «café, thé». Malheureusement c'est impossible : si je ne veux pas me répéter deux fois consécutivement, alors, après un jour avec café, je dois boire du thé, et après un jour avec thé, je doit boire du café, ce qui au bout de quatre jours conduit obligatoirement à une répétition : café, thé, café, thé. Puisque ne pas se répéter deux fois de suite est impossible, je me contenterai de ne jamais me répéter trois fois de suite. Ainsi, je me comporterai de la manière la plus désordonnée possible pour ce qui est de mon petit déjeuner.

Le b.a.-ba de la construction d'un mot infini

Est-ce faisable, et ai-je raison de croire que cela m'évitera l'ennui ? Autrement dit : existe-t-il des mots composés uniquement de a et de b ne comportant jamais trois fois consécutivement la même séquence ? Si oui, peut-on les considérer comme totalement désordonnés ?

La réponse – oui à la première question, non à la seconde – est connue depuis plusieurs décennies et est donnée par ce que l'on appelle le *mot infini de Prouhet–Thue–Morse*.

Voici comment on obtient cet extraordinaire mot infini. On commence par ab , puis on remplace chaque a par ab et chaque b par ba , ce qui donne $abba$. On recommence alors la même substitution, ce qui conduit ensuite à $abbabaab$,

puis à $abbabaabbaababba$,

à $abbabaabbaabbabbaababbaabbaabbaab$ à l'étape suivante, etc.



Une partition du compositeur franco-américain Tom Johnson utilisant le mot de Prouhet–Thue–Morse. La suite est écrite avec des 0 et des 1. Les notes de la mélodie descendent quand il y a un 1 et montent quand il y a un 0.

© Tom Johnson, membre de l'Oumupo (voir article en page 15)

Un autre procédé produit la même suite: on part de a , puis on le recopie à côté de lui-même (par *concaténation*) en échangeant les a en b et les b en a ; cela donne ab , on recommence la même opération pour obtenir $abba$, et ainsi de suite.

Vous constaterez que jamais on ne trouve trois fois de suite le même sous-mot dans le mot infini obtenu. N'est-ce pas merveilleux, vu la simplicité des procédés? Le constater est bien, mais le démontrer est mieux, pour le salut de tous mes petits déjeuners à venir. Prenons donc le temps d'en proposer une démonstration.

Un raisonnement typique sur la combinatoire des mots

Pour montrer que le mot de Prouhet–Thue–Morse est *sans cube* (c'est-à-dire qu'il ne comporte jamais trois fois consécutivement le même sous-mot), il suffit de montrer que, pour tout mot infini M composé uniquement de a et de b , le fait que M est sans cube implique que $f(M)$ est sans cube, où f est l'opération qui change chaque a en ab et chaque b en ba .

Pour cela, supposons que M est sans cube et que $f(M)$ comporte un cube C . On peut alors écrire $f(M) = \dots CCC \dots$. Cherchons à obtenir une contradiction. Trois cas sont à distinguer.

Dans le premier cas, le mot C comporte un nombre pair de lettres et dans $f(M)$ le cube CCC commence à un emplacement de numéro impair (on numérote les lettres d'un mot en commençant par 1). Alors M comporte lui-même un cube, obtenu en remplaçant ab par a et ba par b dans C , ce qui contredit l'hypothèse. Faites une vérification avec, par exemple, $f(M) = ababbaabbaabbabb$, où $C = abba$; vous trouverez le cube $ababab$ dans M .

Dans le deuxième cas, le mot M comporte un nombre pair de lettres et le cube CCC commence à un emplacement de numéro pair dans $f(M)$. Par construction de $f(M)$, la lettre numéro $2n + 1$ est a (respectivement b) si, et seulement si, celle de numéro $2n + 2$ est b (respectivement a). Donc, en enlevant la dernière lettre de C et en ajoutant devant C la lettre complémentaire (a si b et b si a) de la première lettre de C , on obtient un mot C' ayant un nombre pair de lettres qui est répété trois fois dans $f(M)$, la répétition commençant un emplacement avant celle de M . On est donc ramené au cas précédent, et on conclut de la même manière.

Il reste un dernier cas : le mot M comporte un nombre impair de lettres. Il résulte de l'équivalence notée pour le deuxième cas que le mot M , qui apparaît à la fois en commençant à un rang pair dans $f(M)$ et en commençant à un rang impair, est nécessairement composé d'une alternance de a et de b (comme $ababa$, par exemple). Le mot M commence et finit donc par la même lettre. On en déduit que dans $f(M)$ il y aura deux lettres, de rangs $2n + 1$ et $2n + 2$, qui seront identiques (soit à la jonction entre le premier M et le deuxième M du cube, soit à la jonction entre le deuxième et le troisième). Mais cela est impossible, toujours à cause de l'équivalence notée dans le deuxième cas. Nous avons ainsi épuisé toutes les possibilités découlant de l'hypothèse que $f(M)$ comporte un cube, et avons à chaque fois abouti à une contradiction. Si M ne comporte pas de cube, alors $f(M)$ non plus.

De la musique aux échecs en passant par l'écriture en binaire

Dois-je être vraiment satisfait de ce mot qui ne se répète jamais trois fois consécutivement et qui me permet de prendre mon petit déjeuner de manière plus apaisée ? Le mot infini de Prouhet–Thue–Morse est-il vraiment désordonné ? Non, bien sûr ! La double définition qui en a été donnée montre que le mot infini de Prouhet–Thue–Morse n'est pas du tout désordonné. En voici d'ailleurs une troisième définition, qui n'utilise que quelques mots : la $n^{\text{ème}}$ lettre est un a si, et seulement si, le nombre de 1 dans l'écriture binaire de n est pair.

D'où vient le nom de ce formidable mot infini ? Il se nomme « mot de Prouhet–Thue–Morse » car il a été présenté et étudié en 1906 par Axel Thue (1863–1922) dans un obscur journal norvégien. Cet objet combinatoire ne retint vraiment l'attention de la communauté mathématique que lorsque le mathématicien américain Harold Calvin Marston Morse (1892–1977) le redécouvrit quinze ans plus tard, en 1921. Il sera retrouvé encore une fois par le grand maître d'échecs

Machgielis (Max) Euwe (1901–1981, champion du monde de 1935 à 1937). Euwe démontra que la règle du jeu d'échecs qui impose de ne jamais jouer trois fois de suite la même séquence de coups autorisait des parties infinies. À chaque a du mot infini, on associe la séquence de coups suivante: $g1-f3$, $g8-f6$, $f3-g1$, $f6-g8$ (les blancs, puis les noirs, sortent leur cavalier gauche, puis le rentrent); à chaque b de la suite, on associe la séquence de coups $b1-c3$, $b8-c6$, $c3-b1$, $c6-b8$ (les blancs, puis les noirs, sortent leur cavalier droit, puis le rentrent). En suivant le mot de Prouhet–Thue–Morse, on ne rejouera jamais trois fois de suite les mêmes coups, et donc la partie durera indéfiniment.

Et Prouhet? Le mot infini, à vrai dire, avait déjà été remarqué avant l'article de Thue pour une de ses autres propriétés extraordinaires. En effet, le mathématicien français Eugène Prouhet (1817–1867), dans un article antérieur (*Mémoire sur quelques relations entre puissances d'un nombre*, *Comptes-rendus de l'Académie des Sciences de Paris*, série I, 33, 1851), démontrait une jolie propriété arithmétique.

MÉMOIRES PRÉSENTÉS.

THÉORIE DES NOMBRES. — *Mémoire sur quelques relations entre les puissances des nombres*; par **M. E. PROUHET**. (Extrait par l'auteur.)

(Commissaires, MM. Sturm, Lamé, Binet.)

« n et m étant deux nombres entiers quelconques, il existe une infinité de suites de n^m nombres, susceptibles de se partager en n groupes de n^{m-1} termes chacun et tels que la somme des puissances k des termes soit la même pour tous les groupes, k étant un nombre entier inférieur à m .

» n^m nombres en progression arithmétique jouissent de la propriété précédente. Pour opérer le partage de ces nombres en groupes, on écrira en cercle les indices $0, 1, 2, \dots, n-1$; on lira ces indices en suivant le cercle et en ayant soin d'en passer un à chaque tour; deux, tous les n tours; trois, tous les n^2 tours, et ainsi de suite. Ces indices, écrits à mesure qu'on les lit sous les termes de la progression, apprendront à quel groupe appartient chaque terme.

» Si l'on applique la règle et le théorème précédents aux 27 premiers nombres de la suite naturelle, on arrive aux identités suivantes :

$$\begin{aligned} 1+6+8+12+14+16+20+22+27 &= 2+4+9+10+15+17+21+23+25=3+5+7+11+13+18+19+24+26 \\ 1^1+6^1+8^1+\dots &= 2^1+4^1+9^1+\dots = 3^1+5^1+7^1+\dots \end{aligned}$$

» Lorsque $n=10$ et que la progression commence à 0, tous les nombres dont la somme des chiffres, divisée par 10, laisse le même reste, appartiennent à la même classe. »

Le début de la publication de Prouhet.

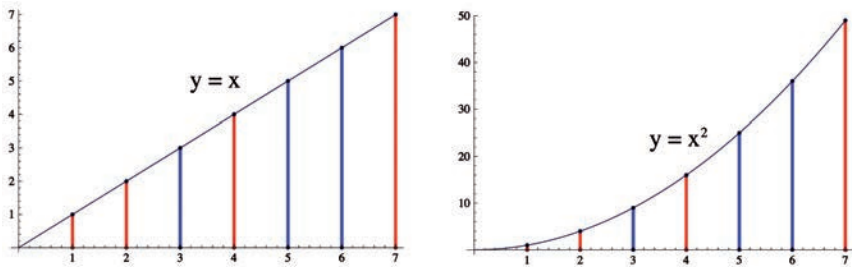
© Institut de France

Des applications arithmétiques élégantes et inattendues

Voici la propriété mise à jour par Prouhet. Si l'on place à droite d'une égalité les quatre entiers correspondant aux numéros des quatre premiers emplacements des a (ce sont les entiers 0, 3, 5, 6) et à gauche les quatre entiers correspondant aux quatre premiers emplacements des b (ce sont 1, 2, 4, 7), alors on obtient une série de deux égalités remarquables simultanées :

$$\begin{aligned} 0 + 3 + 5 + 6 &= 1 + 2 + 4 + 7, \\ 0^2 + 3^2 + 5^2 + 6^2 &= 1^2 + 2^2 + 4^2 + 7^2. \end{aligned}$$

$$A = \{0, 3, 5, 6\} \quad B = \{1, 2, 4, 7\}$$



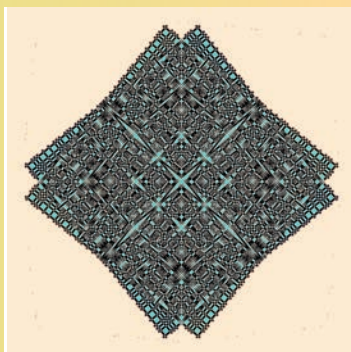
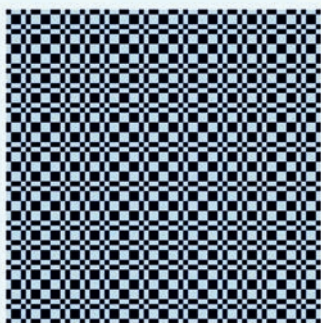
La somme des longueurs des segments en bleu et en rouge sont les mêmes.

© Dessin emprunté à Michel Rigo (voir bibliographie)

Il y a mieux encore. Si l'on met à droite d'une égalité les huit entiers correspondant aux numéros des huit premiers emplacements des a (ce sont les entiers 0, 3, 5, 6, 9, 10, 12, 15) et à gauche les huit entiers correspondant aux huit premiers emplacements des b (ce sont 1, 2, 4, 7, 8, 11, 13, 14), alors on obtient une série de trois égalités remarquables simultanées :

$$\begin{aligned} 0 + 3 + 5 + 6 + 9 + 10 + 12 + 15 &= 1 + 2 + 4 + 7 + 8 + 11 + 13 + 14, \\ 0^2 + 3^2 + 5^2 + 6^2 + 9^2 + 10^2 + 12^2 + 15^2 &= 1^2 + 2^2 + 4^2 + 7^2 + 8^2 + 11^2 + 13^2 + 14^2, \\ 0^3 + 3^3 + 5^3 + 6^3 + 9^3 + 10^3 + 12^3 + 15^3 &= 1^3 + 2^3 + 4^3 + 7^3 + 8^3 + 11^3 + 13^3 + 14^3. \end{aligned}$$

En exploitant le mot de Prouhet–Thue–Morse, ces résultats se généralisent et produisent en fait quatre égalités entre deux paquets de seize nombres, cinq égalités entre deux paquets de trente-deux nombres, *etc.* Le mot de Prouhet–Thue–Morse conduit donc à l'affirmation très surprenante suivante : pour tout nombre n , on peut trouver deux ensembles de nombres entiers distincts A et B , tels que la somme des nombres de A est la même que la somme des nombres de B , de même pour les carrés, les cubes... et les puissances $n^{\text{èmes}}$ de ces mêmes nombres.



Un schéma tiré directement du mot de Prouhet–Thue–Morse (à gauche) et un dessin qui en est issu de manière moins directe (à droite).

© J.-P. Delahaye

Même la plus simple des questions sur la répétition des lettres d'un mot conduit à des découvertes mathématiques et oblige à des raisonnements intéressants ! Le mot de Prouhet–Thue–Morse est en fait immensément plus riche de propriétés, toutes plus élégantes les unes que les autres. De bien belles mathématiques en combinatoire des mots résultent de l'étude des mots infinis et des transformations (comme la fonction f) qui les changent les uns en les autres.

J. – P. D.

Pour en savoir (un peu) plus :

The Ubiquitous Prouhet–Thue–Morse Sequence. Jean-Paul Allouche et Jeffrey Shallit, in *Sequences And Their Applications*, Springer-Verlag, 1999.

Des mots magiques infinis. Jean-Paul Delahaye, in *Mathématiques pour le plaisir*, Belin–Pour La Science, 2010.

Automatic Sequences, Theory, Applications, Generalizations. Jean-Paul Allouche et Jeffrey Shallit, Cambridge University Press, 2003.

Le problème de Prouhet et la suite de Thue–Morse. Michel Rigo, exposé disponible en ligne (orbi.ulg.ac.be/bitstream/2268/129326/1/RigoSBPM.pdf).