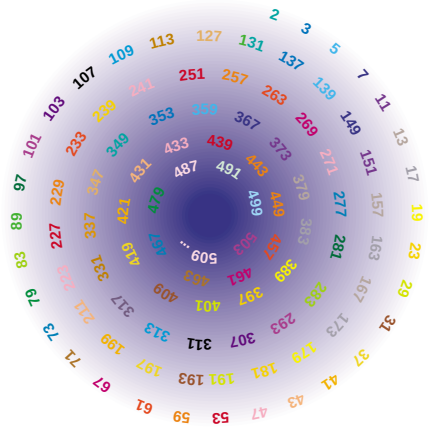

Formules pour les nombres premiers

JEAN-PAUL DELAHAYE

On sait enfermer tous les nombres premiers dans des formules simples. Est-ce utile?

Les nombres pairs sont donnés par une formule $p(n) = 2n$, qui n'a rien de mystérieux : elle indique que le double de tout entier est un nombre pair. Les nombres impairs sont donnés par la formule $I(n) = 2n+1$. Existe-t-il une formule analogue pour les nombres premiers, ces nombres qui ne sont divisibles que par 1 et eux-mêmes (notons que 1 n'est pas considéré comme premier)? Les dix plus petits nombres premiers sont 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, et Euclide a montré qu'ils sont en nombre infini.

Cette recherche d'une formule pour les nombres premiers passionne les amateurs d'arithmétique depuis longtemps et a produit d'étonnants résultats. Elle conduit à des exercices distrayants de raisonnement et de calcul, permet de préciser ce qu'on entend par formule, pour finalement nous plonger dans le monde de l'informatique.

Il est vite apparu qu'il était impossible d'engendrer tous les nombres premiers (ou même de n'engendrer que des nombres premiers, pas nécessairement tous, ni dans l'ordre) à l'aide d'une formule du type $f(n) = an + b$ pour n positif. Voyons pourquoi. Supposons que $an + b$ soit un nombre premier pour tout n supérieur à 0.

Comme la formule est exacte pour $n = 0$, $f(0) = b$, et b est un nombre premier. Mais alors $f(b) = ab + b$ est aussi un nombre premier, et comme $ab + b = (a + 1)b$ est divisible par b et $(a + 1)$, ou bien $b = 1$, ce qui est impossible puisque b est un nombre premier, ou bien $a + 1 = 1$, ce qui conduit à $a = 0$, et donc à la formule $f(n) = b$ qui donne toujours le même nombre premier. Les seules formules de type $an + b$ qui ne donnent que des nombres premiers sont celles (sans intérêt) où a est égal à 0 et b est un nombre premier. On montre par un raisonnement similaire que, si une formule (dite polynomiale) $f(n) = a_p n^p + a_{p-1} n^{p-1} + ... + a_1 n + a_0$, ne donne que des nombres premiers pour tout n positif, alors c'est que $f(n)$ est constant (et donc sans intérêt). Triste conclusion : un polynôme à une variable, aussi compliqué soit-il, ne donnera jamais de nombres premiers pour toute valeur positive de n, sauf s'il n'en donne qu'un...

Démontrons ce résultat en raisonnant par l'absurde : nous supposerons que $f(n)$ n'est pas constant et ne prend que des valeurs qui sont des nombres premiers pour n positif. On a $f(0) = a_0$, donc a_0 est premier. Par hypothèse, $f(n)$ qui n'est pas constant tend vers l'infini positif quand n tend vers l'infini (sinon $f(n)$ ne donnerait que des valeurs négatives à partir de certaines valeurs de n et donc ne conviendrait pas). Il existe donc un entier p tel que $f(pa_0)$ soit plus grand que a_0. Alors $f(pa_0)$ est un multiple de a_0 (car c'est une somme de multiples de a_0) plus grand que a_0. Ce n'est pas un nombre premier.

De multiples raffinements de ces deux résultats semblent éliminer tout espoir que nous puissions obtenir des formules intéressantes pour les nombres premiers. En voici quelques-uns.

• Si une fonction polynôme à plusieurs variables (telle que, par exemple, la fonction $f(n,m,p) = nm + p^2 + n^3p^5 + 17$) ne prend que des valeurs premières pour les valeurs entières positives de ses variables, alors c'est une fonction constante (et donc sans intérêt).

• Si une fonction quotient des deux polynômes à plusieurs variables (exemple : $f(n,m,p,q,r)=(nm+p^2+q^4+r^9+19)/(p^2+q^4+r^7+1)$) ne prend que des valeurs premières pour les valeurs positives de ses variables, alors c'est une fonction constante (et donc sans intérêt).

• Si une fonction est de la forme $P2^Q + R$, où P, Q et R sont des polynômes à plusieurs variables (exemple : $f(n,p,q) = (n + 5p+q)2^{2n+p+3q} + p + 17$) ne prend que des valeurs premières pour les valeurs positives de ses variables, alors c'est une fonction constante (et donc sans intérêt).

• Même résultat avec, dans la formule précédente, 3 (ou n'importe quel entier supérieur à 2) à la place de 2, ou encore avec une somme de termes de la forme Pa^Q à la place d'un seul.

FAUT-IL RENONCER ?

Abandonner? Que nenni! Les polynômes ne sont pas tout : dans l'arsenal d'un mathématicien, il y a bien d'autres fonctions. Un résultat, démontré par W. Mills en 1947, a étonné tout le monde : il existe une constante A telle que la formule $[A^{3^n}]$ fournit un nombre premier pour tout n entier supérieur à 1.

Le crochet [ ] désigne la fonction partie entière ou «arrondi à l'entier inférieur» : exemples : [128,679]=128 ; [3,14159]=3. Si on voulait une formule valable pour tout n positif, il suffirait de prendre : $[A^{3^{n+1}}]$. La constante A, dénommée constante de Mills,

1. FORMULE AVEC DES CONSTANTES RÉELLES

Il existe un nombre réel L tel que la formule :

$$[L \times 10^{n^2}] - [L \times 10^{(n-1)^2}]10^{2n-1} \text{ pour tout } n \geq 1$$

donne le n-ième nombre premier : $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, ...
Le nombre L est 0,2003000050000007000000001100..., (le n-ième nombre premier est placé en position n^2).
Suivons les calculs pour n égal à 4 :
$[L \times 10^{16}]$ = 2003000050000007
(la multiplication par 10^{16} amène devant la virgule les décimales jusqu'au 7, et la partie entière coupe ce qui est derrière la virgule.
$[L \times 10^{9}]$ = 200300005
(même chose, mais avec le 5)
$[L \times 10^{9}]10^{7}$= 2003000050000000
(la multiplication par 10^7 égalise la longueur des deux nombres)
$[L \times 10^{16}] - [L \times 10^{9}]10^{7}$ =
2003000050000007 – 2003000050000000 = 7.

a été calculée avec une bonne précision, elle vaut : 1,3063778838630806904686144926025712916784585156713644368053759966434...

Pour $n = 1$, on obtient 2 ; pour $n = 2$, 11 ; pour $n = 3$, 1361 ; pour $n = 4$, 2521008887, nombres qui sont premiers, comme on le vérifie sans peine. Hélas, la formule $[A^{3^n}]$, amusante et curieuse, est sans intérêt pratique, car pour l'utiliser il faut connaître A avec une grande précision, ce qu'aujourd'hui on ne sait obtenir qu'en calculant... les nombres premiers eux-mêmes. Il paraît improbable qu'on puisse calculer cette constante autrement qu'à partir des nombres premiers, mais il serait merveilleux d'y parvenir. Cette formule ne donne les nombres premiers que parce que l'information sur les nombres premiers est cachée dans la constante A... Dans la même veine, un résultat étrange a été démontré par G. Wright en 1951. Il existe une constante w telle que la formule : $f(n) = [2^{2^{2^{\cdot^{\cdot^{\cdot^{w}}}}}}]$, avec n exposants, fournit un nombre premier pour tout n supérieur ou égal à 1. La constante w, dénommée bien sûr constante de Wright, vaut 1,9287800...

La formule suivante – un peu plus compliquée – va faire comprendre pourquoi ni la formule de Mills ni celle de Wright ne sont autre chose que des curiosités : il existe une constante L (appelée parfois constante de Liouville-Erdös) telle que la formule : $[L \times 10^{n^2}] - [L \times 10^{(n-1)^2}]10^{2n-1}$ donne le n-ième nombre premier pour tout $n \geq 1$: $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$, $p_7 = 17$, etc. Cette formule non seulement ne donne que des nombres premiers (comme les deux précédentes), mais elle les donne tous dans l'ordre. L'escroquerie apparaît quand on voit la constante L :

L = 0,2003000050000007000000011000000000001300..., le n-ième nombre premier est placé en position n^2. Le détail du mécanisme de calcul est expliqué en figure 1.

PLUS DE TRICHE!

Ce type de formules étant une forme de tricherie, interdisons-nous d'utiliser des nombres réels qui peuvent cacher une infinité d'informations dans leurs décimales, infinité où l'on peut loger insidieusement tous les nombres premiers !

Peut-on, en respectant cette nouvelle contrainte, trouver une formule qui donne tous les nombres premiers ? Si on ne peut pas les obtenir tous dans l'ordre, on se contenterait de les avoir dans le désordre, avec des répétitions, ou même d'en avoir une infinité.

Ce qu'on a vu au début montre qu'on ne réussira qu'avec un symbolisme un peu plus riche que celui des polynômes. Proposons d'accepter le symbolisme naturel suivant considéré comme banal en algèbre et en arithmétique :

– [x] partie entière de x, définie comme l'arrondi à l'entier inférieur de x (déjà utilisée plus haut) ;

– $\sum$, pour la somme généralisée ; exemple $\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + ... + n^2$

– !, la notation factorielle définie par $n! = n(n-1)(n-2)...4.3.2.1$;

– | | la valeur absolue, définie par $|n| = n$ si $n \geq 0$ et $|n| = -n$ si $n \leq 0$.

Il est maintenant possible de trouver des formules qui ne «trichent pas». Le premier exemple est une formule qui donne tous les nombres premiers et qui pourtant comporte 37 symboles :

$$t(n) = 2 + n[1/(1 + \sum_{p=2}^{n+1} [(n+2)/p - [(n+1)/p]])] \qquad n \geq 0$$

Le fonctionnement de cette formule, due à Roland Yélé-hada, s'explique en quelques mots. Si $n + 2$ est un multiple de p, alors $(n + 2)/p$ est un entier q, et donc $(n + 1)/p = q - 1/p$. Il en résulte que $[(n + 1)/p] = q - 1$ et que $[(n + 2)/p - [(n + 1)/p]]$ est égal à 1. En revanche, si $n + 2$ n'est pas un multiple de p,

2. JUSTIFICATION D'UNE FORMULE POUR P_n

$P(i) = [((i-1)! + 1) / i - [(i - 1)! / i]]$; $P(i)$ vaut 1 si i est premier.

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	0	1	0	1	0	0	0	1	0	1	0	0	0	1	0

$P_i(m) = \sum_{i=1}^{m} P(i)$; $P_i(m)$ est le nombre de nombres premiers ≤ m

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	2	2	3	3	4	4	4	4	5	5	6	6	6	6	7	7

$P_i(m) / n$

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	2	2	3	3	4	4	4	4	5	5	6	6	6	6	7	7
2	0,5	1	1	1,5	1,5	2	2	2	2	2,5	2,5	3	3	3	3	3,5	3,5
3	0,33	0,66	0,66	1	1	1,33	1,33	1,33	1,33	1,66	1,66	2	2	2	2	2,33	2,33
4	0,25	0,5	0,5	0,75	0,75	1	1	9	1	1,25	1,25	1,5	1,5	1,5	1,5	1,75	1,75
5	0,2	0,4	0,4	0,6	0,6	0,8	0,8	0,8	0,8	1	1	1,2	1,2	1,2	1,2	1,4	1,4
6	0,16	0,33	0,33	0,5	0,5	0,66	0,66	0,66	0,66	0,83	0,83	1	1	1	1	1,16	1,16

$[P_i(m) / n]$; [a] donne la valeur entière du nombre a

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	2	2	3	3	4	4	4	4	5	5	6	6	6	6	7	7
2	0	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3
3	0	0	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2
4	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

min ($[P_i(m)/n]$, 1) = $g(n, m)$; min $(x,y) = (x + y - |x - y|)/2$
min donne le minimum des nombres entre ( ).

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

$g(n, m-1)$.

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	1	1	1	1	1	1	1	1	1	1	**1**	1	1
4	0	0	0	0	0	0	1	1	1	1	1	1	1	1	**1**	1	1
5	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1

$g(n, m) - g(n, m-1)$.

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0

$m(g(n, m) - g(n, m-1))$. $\quad p_n = \sum_{m=2}^{n^2+1} m(g(n, m) - g(n, m-1))$.

m / n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	**5**	0	0	0	0	0	0	0	0	0	0		0	0
4	0	0	0	0	0	**7**	0	0	0	0	0	0	0	0		0	0
5	0	0	0	0	0	0	0	0	0	**11**	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	**13**	0	0	0	0	0