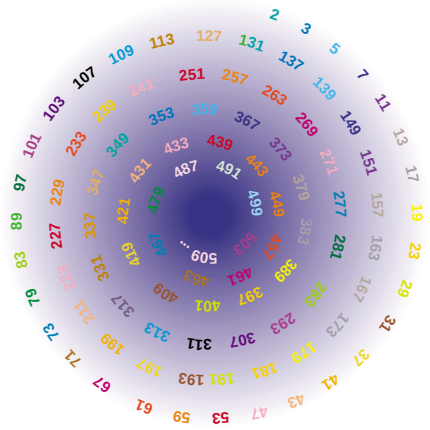

Formules pour les nombres premiers

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 :

$$[L \times 10^{n^2}] - [L \times 10^{(n-1)^2}]10^{2n-1}$$ 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 ( ).

n / m	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

$g(n, m-1)$.

n / m	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

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

n / m	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

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

n / m	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

3. UNE FORMULE POUR LES NOMBRES PREMIERS JUMEAUX

Les nombres premiers jumeaux sont les paires de nombres premiers dont la différence est 2, comme 11 – 13 ou 17 – 19. On ne sait toujours pas s'il existe une infinité de paires de nombres premiers jumeaux. En revanche, on sait que n est le premier élément d'une paire de nombres premiers jumeaux si et seulement si :

$$4((n-1)!+1) + n \text{ est un multiple de } n\,(n+2).$$

D'où l'on tire que $n + 3$ est le premier élément d'une paire de nombres premiers jumeaux si et seulement si :

$$4(n+2)! + n + 7 \text{ est un multiple de } (n+3)(n+5).$$

Il en résulte que la formule assez simple suivante engendre tous les nombres premiers jumeaux (sur le principe de la formule de Minac) :

$$j(n) = 3 + n\,[(4(n+2)! + n + 7)/(n+3)(n+5) - [(4(n+2)! + n + 6)/(n+3)(n+5)]].$$

Cette formule est amusante, car, bien qu'elle soit assez simple, personne ne sait si elle prend au total un nombre fini ou infini de valeurs différentes.

alors $[(n + 2)/p - [(n + 1)/p]]$ vaut 0. Autrement dit le sigma dans la formule compte le nombre de diviseurs de $n + 2$ compris entre 2 et $n + 1$.

De deux choses l'une :

– $n + 2$ est premier, alors le nombre de diviseurs de $n + 2$ entre 2 et $n + 1$ est 0, donc le crochet derrière $2 + n$ est [1/1], c'est-à-dire 1, et l'on a $t(n) = n + 2$, qui est bien un nombre premier.

– $n + 2$ n'est pas premier, alors le nombre de diviseurs de $n + 2$ entre 2 et $n + 1$ est supérieur à 1, donc le crochet derrière $2 + n$ vaut 0 et donc $t(n) = 2$, qui est bien un nombre premier.

La formule ne donne que des nombres premiers, les donne tous, mais très lentement et en répétant 2 trop souvent :
$t(0) = 2$, $t(1) = 3$, $t(2) = 2$, $t(3) = 5$, $t(4) = 2$, $t(5) = 7$, $t(6) = 2$, $t(7) = 2$, $t(8) = 2$, $t(9) = 11$, $t(10) = 2$, $t(11) = 13$, $t(12) = 2$, $t(13) = 2$, etc.

Le théorème de John Wilson (1741-1793), publié en 1770 par Waring, indique que $(p - 1)! + 1$ est un multiple de p si et seulement si p est premier. Cela a conduit Minác à simplifier grandement la formule de Yéléhada, laquelle devient : $t(n) = 2 + n\,[((n+1)!+1)/(n+2) - [(n+1)!/(n+2)]]$.

Ce qu'on a gagné (seulement 36 symboles pour tous les nombres premiers et la disparition de l'utilisation de la notation sigma) est malheureusement compensé par une rapidité bien inférieure : la nouvelle formule passe par des factorielles dont le calcul est long.

D'autres formules n'ont pas les défauts des formules de Yéléhada et de Minác, qui ne fournissent pas les nombres premiers dans l'ordre. Les mathématiciens Minác et Willans ont imaginé une remarquable formule dont l'explication est nettement moins simple, mais qui cette fois donne tous les nombres premiers dans l'ordre et sans répétition. Cette formule de 52 symboles repose, elle aussi, sur le théorème de Wilson.

$p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$, $p_7 = 17$, etc.

$$p_n = 1 + \sum_{m=1}^{2^n} [[n/(1 + \sum_{j=2}^{m} [((j-1)! + 1)/j - [(j-1)!/j]])]^{1/n}] \quad n \geq 1$$

La formule est merveilleuse : peu de gens imaginaient avant qu'elle soit publiée en 1995 qu'une telle formule pouvait exister. Est-elle utile ? Non, d'un point de vue pratique. En effet, si l'on utilise cette formule pour programmer une méthode de calcul des nombres premiers, on obtient un programme d'une rare inefficacité. J'ai tenté l'expérience : la formule fonctionne comme prévu par la théorie, mais, bien qu'utilisant un logiciel puissant et calculant avec un grand nombre de chiffres exacts, je n'ai pas pu aller au-delà de p_6. La figure 2 explicite en détail une formule légèrement plus complexe pour p_n, mais plus transparente et plus efficace.

Sur un plan théorique, ces formules sont cependant intéressantes, car elles montrent que l'ensemble des nombres premiers n'est pas compliqué : peu de symboles permettent de le représenter entièrement et dans l'ordre. Cette conclusion n'est pas une surprise pour les informaticiens, qui utilisent pour écrire leurs programmes des langages plus riches que ceux utilisés habituellement en algèbre et en arithmétique. Écrire une formule ou un programme n'est pas très différent : d'ailleurs le nom du langage Fortran qui fut l'un des premiers langages de l'informatique scientifique, est la forme contractée de *Formula Translation,* car il était considéré simplement comme un moyen d'écrire des formules mathématiques.

Si l'on réfléchit au problème des formules pour p_n, il faut admettre que finalement ce sont les informaticiens qui ont raison : il vaut mieux enrichir un peu le vocabulaire qu'on se donne : cela permet d'écrire des formules pour les nombres premiers (ou pour d'autres choses) qui ne sont pas seulement des exercices de virtuosité d'un intérêt incertain, mais qui retranscrivent fidèlement et sans détour des procédés de calcul ou des idées naturelles. La logique mathématique a depuis longtemps compris cela et propose un système de notation qui permet des définitions de p_n par des formules à la fois courtes, claires et efficaces.

Voici, dans le langage Maple (utilisé aujourd'hui dans les classes préparatoires scientifiques aux Grandes Écoles et dans les premiers cycles des universités), une formule de 165 caractères donnant à nouveau le n-ième nombre premier p_n. Ce programme est fondé uniquement sur la fonction a mod b, qui donne le reste de la division de l'entier a par l'entier b (la fonction «mod», offerte dans tous les langages de programmation actuels peut se définir par $a \bmod b = a - b\,[a/b]$).

```
premier := proc(n)
local p,k,d;
 if n=1 then 2
 else
  p:= 1; k:= 1;
  while k<>n do
   p:= p+2; d:= 2;
   while (d*d<=p) and
        (p mod d <>  0) do d:=d+1;od;
```

4. AUTRES FORMULES ÉTONNANTES

Voici quelques formules déconcertantes concernant les nombres premiers.

(A) Formules de Willans (1964) du nombre de nombres premiers inférieurs à m :

$$Pi(m) = \sum_{j=2}^{m} \sin^2(\pi((j-1)!)^2/j)/\sin^2(\pi/j).$$

$$Pi(m) = -1 + \sum_{j=1}^{m} [\cos^2(\pi((j-1)!+1)/j)].$$

(B) Formule due à G. Hardy, qui soutenait que les mathématiques n'ont pas à être utiles, mais seulement belles (c'est bien le cas de cette formule) :

$$\lim_{r\to\infty} \lim_{m\to\infty} \lim_{n\to\infty} \sum_{i=1}^{m} (1-(\cos((i!)^r \pi/N))^{2n})$$

est le plus grand facteur premier de l'entier N.

(C) Formule donnant le plus grand commun diviseur des entiers n et m, trouvée en 1997 par Marcelo Polezzi :

$$\text{pgcd}(m,n) = 2\sum_{i=1}^{m-1} [in/m] + (m+n) - mn.$$

5. ATTENTION AUX GÉNÉRALISATIONS HÂTIVES

De nombreuses fois dans l'histoire des mathématiques, des propriétés constatées sur un petit nombre de cas ont été conjecturées «vraies dans tous les cas». Voici quelques exemples de situations qui montrent qu'il faut se méfier de telles généralisations.

(a) Pierre de Fermat considère l'expression $F_n = 2^{2^n}+1$ et, après le calcul de F_0, F_1, F_2, F_3, F_4, constate qu'il s'agit toujours de nombres premiers. Il en tire la conjecture que F_n est toujours un nombre premier. Le grand Leonhard Euler découvre plus tard que F_5 n'est pas premier. Pire : aujourd'hui, après avoir évalué et testé les valeurs suivantes de F_n, on n'a trouvé que des nombres composés. On se demande donc si, au-delà de F_5, tous les F_n ne seraient pas composés ! Si c'est le cas, ce sera l'exemple extrême de la conjecture fausse : non seulement elle possède un contre-exemple, mais tout nombre plus grand que 5 est un contre-exemple !

(b) Plus récemment, on a remarqué que :

3! – 2! + 1! = 5 : premier
4! – 3! + 2! – 1! = 19 : premier
5! – 4! + 3! – 2! + 1! = 101 : premier
6! – 5! + 4! – 3! + 2! – 1! = 619 : premier
7! – 6! + 5! – 4! + 3! – 2! + 1! = 4 421 : premier
8! – 7! + 6! – 5! + 4! – 3! + 2! – 1! = 35 899 : premier

En conclut-on que $n!$ – $(n–1)!$ + ... + $(–1)^n$ 2! – $(–1)^n$ 1! est toujours un nombre premier ?

Non : 9! – 8! + 7! – 6! + 5! – 4! + 3! – 2! + 1! = 326 981 = 79 × 4 139.

(c) Plus étonnant encore est l'exemple suivant. Les nombres n^{17} + 9 et $(n + 1)^{17}$ + 9 n'ont pas de facteurs communs si n = 1, n = 2, n = 3, etc. Vous pouvez vérifier cela jusqu'à 100, jusqu'à 1 000, jusqu'à un milliard, jusqu'à mille milliards de milliards, et même jusqu'à huit millions de milliards de milliards de milliards de milliards de milliards. Pourtant, un peu plus loin, pour :

n = 8 424 432 925 592 889 329 288 197 322 308 900 672 459 420 460 792 433 ≃ 8,424 10^{52}, vous rencontrerez une exception : parfois n^{17} + 9 et $(n+1)^{17}$ + 9 ont des facteurs communs.

(d) Il est des situations plus étranges encore, où seul le raisonnement permet de savoir qu'une propriété toujours testée vraie en pratique est fausse en général.

L'exemple le plus connu est celui qui concerne l'approximation de Pi(m) le nombre de nombres premiers inférieurs à m. On sait que ce nombre vaut à peu près :

$$\mathrm{Li}(m) = \int_2^m 1/\log(x)\,dx\ .$$

On constate en pratique que Pi(m) ≤ Li(m), et l'on ne connaît aucune exception à cette inégalité. Cependant le mathématicien Littlewood a prouvé en 1914 que la quantité Pi(m) – Li(m) changeait de signe une infinité de fois, et donc qu'il n'est pas vrai que Pi(m) ≤ Li(m) pour tout m. En 1933, S. Skewes a établi que le premier changement de signe se produit avant $10^{10^{10^{34}}}$ (quatre niveaux d'exposant). Cette majoration a été améliorée par H. te Riele, en 1987, par 10^{371}, mais c'est encore au-delà de ce qu'on peut atteindre numériquement : aujourd'hui, on sait qu'il existe un m tel que Pi(m) > Li(m), mais on n'en connaît aucun.

```
    if d*d>p then k:=k+1;fi;
   od;
   p;
  fi;
 end:
```

Ce qui tient sur cinq lignes :

```
premier := proc(n) local p,k,d;if n=1 then 2 else p:=1;k:=1;while k<>n do p:=p+2;d:=2;while(d*d<=p)and(p mod d <> 0)do d:=d+1;od;if d*d>p then k:=k+1;fi;od;p;fi;end;
```

Explications : proc est un mot convenu pour indiquer qu'on définit une procédure, c'est-à-dire une fonction ; local sert à préciser les variables qu'on utilisera dans le calcul ; if C then A1 else A2 fi est la structure de contrôle qui, lorsque la condition C est satisfaite, exécute l'action A1 et qui, sinon, exécute l'action A2 ; while C do A od est la structure de contrôle qui, tant que la condition C est satisfaite, exécute de manière répétée l'action A ; end indique la fin de la définition de la procédure.

Le programme, si n vaut 1, donne 2. Sinon, en parcourant les entiers de 2 en 2 (variable p), le programme recherche les nombres premiers et les compte avec la variable k, jusqu'à en avoir trouvé n. Le programme rend alors le résultat p, qui à cet instant, contient bien le n-ième nombre premier.

On peut faire encore plus court, mais cette définition possède plusieurs avantages sur la définition mathématique précédente de p_n ou sur une définition plus courte en Maple : elle se comprend immédiatement pour qui possède quelques connaissances du langage Maple, elle est fondée sur l'idée la plus naturelle qu'on puisse avoir ; elle permet de calculer en quelques secondes le centième nombre premier, et même le millième.

La recherche de formules n'utilisant qu'un symbolisme mathématique limité est un jeu un peu gratuit que les mathématiciens professionnels regardent le plus souvent avec amusement sans le prendre au sérieux (ce qui n'empêche pas que d'éminents mathématiciens comme G. Hardy ou E. Wright s'y soient adonnés).

6. FORMULES SIMPLES DONNANT DE GRANDES QUANTITÉS DE NOMBRES PREMIERS

Certaines formules, sans donner tous les nombres premiers, ni même ne donner que des nombres premiers, en donnent une grande quantité en suivant. C'est le cas de la formule d'Euler $f(n) = n^2 – n + 41$, qui donne des nombres premiers pour toutes les valeurs de n allant de 0 à 40.

Les polynômes suivants battent le record d'Euler :

$103n^2 – 3945n + 34381$ est premier pour n = 0, 1, ..., 42 (R. Ruby).

$47n^2 – 1701n + 10181$ est premier pour n = 0, 1, ..., 42 (G. Fung).

$36n^2 – 810n + 2753$ est premier pour n = 0, 1, ...,44 (R. Ruby).

Une conjecture très vraisemblable (car liée à une autre bien testée) est qu'aussi grand que soit A on peut trouver un polynôme de la forme $n^2 + n + B$ qui donne des nombres premiers pour n = 0, ..., A. On sait cependant que B sera nécessairement très grand : la valeur de B correspondant à A = 41 est plus grande que 10^{18}, mais reste inconnue.

7. a^b EST DIOPHANTIEN

La découverte de Matiiassevitch en 1970, qui résout le dixième problème de Hilbert, fut le couronnement d'une longue série de progrès commencée par la formulation, dans les années 1930, d'une définition de la notion générale d'algorithmes par A. Church et A. Turing : pour démontrer qu'aucun algorithme ne peut traiter le problème des équations diophantiennes, il faut d'abord disposer d'une définition satisfaisante de la notion d'algorithme !

De nombreuses avancées conduisirent à une situation en 1970 où, pour établir l'indécidabilité du dixième problème de Hilbert, seule restait à prouver que a^b est diophantien, c'est-à-dire qu'il existe une équation $P(a, b, c, x_1, x_2,..., x_n) = 0$, P polynôme, possédant des solutions entières $x_1, x_2,..., x_n$, si et seulement si $a^b = c$.

Ce problème en apparence élémentaire est en fait extrêmement difficile. Comme le caractère diophantien de a^b entraîne l'existence d'un polynôme dont les valeurs positives sont les nombres premiers – chose jugée invraisemblable à l'époque – nombreux furent les mathématiciens (dont le grand logicien polonais Alfred Tarski) qui concluaient à tort que a^b n'était pas diophantien.

Le jeune mathématicien Youri Matiiassevitch, alors chercheur à l'Institut Sketlov de Leningrad, travailla avec obstination plusieurs années de suite sur la preuve du caractère diophantien de l'exponentielle. Il crut avoir trouvé une solution, mais, au moment même où il en faisait l'exposé devant ses collègues, il découvrit une erreur. Enfin il trouva seul une solution, au début du mois de janvier 1970. Quelques jours avant sa découverte, absorbé dans des pensées qui allaient résoudre un problème vieux de 70 ans, il avait quitté la soirée de fête du Nouvel An en enfilant par distraction le manteau de son oncle, d'une taille franchement inadéquate...

LE DIXIÈME PROBLÈME DE HILBERT

Pourtant ce sont les mathématiciens professionnels, à l'occasion d'un défi lancé par David Hilbert en 1900, qui ont obtenu le plus surprenant résultat concernant les formules simples définissant l'ensemble des nombres premiers.

Le problème posé par David Hilbert en 1900 (parmi 23 problèmes) était de trouver une méthode générale et systématique pour étudier les équations diophantiennes, c'est-à-dire les équations polynomiales à coefficients entiers, comme l'équation $X^2 + Y^3 = Z^5$. On savait que la recherche des solutions entières de ce type d'équations était difficile ; Hilbert pensait d'ailleurs qu'aucune méthode systématique n'existait. La démonstration de cette conjecture sera terminée en 1970 par le mathématicien Youri Matiiassevitch, qui prouva qu'une certaine équation polynomiale paramétrée ne pouvait pas être résolue pour toutes les valeurs de ses paramètres à l'aide d'une seule méthode (autrement dit, à l'aide d'un seul algorithme).

Le travail fait pour élaborer cette équation aboutissait à une autre conclusion remarquable : à tout ensemble de nombres entiers dont les éléments peuvent être énumérés par un programme (l'ensemble des nombres premiers en est un) correspond une fonction polynôme dont les valeurs positives, lorsque les variables prennent des valeurs positives, constituent exactement les éléments de cet ensemble. Il existe donc une fonction polynôme (à plusieurs variables) qui, lorsque les variables prennent des valeurs positives, prend des valeurs positives et négatives, l'ensemble de celles qui sont positives étant exactement l'ensemble des nombres premiers.

Cela ne contredit pas les énoncés cités au début, qui indiquaient seulement que ne pouvait exister une fonction polynôme à plusieurs variables qui, lorsque ses variables prennent des valeurs positives, ne prend que des valeurs positives, l'ensemble des valeurs prises étant exactement l'ensemble des nombres premiers.

L'écart entre le résultat positif et le résultat négatif est mince comme une feuille de papier à cigarette et, à vrai dire, a profondément étonné les mathématiciens. Il fallut sept ans pour que le polynôme dont l'existence résulte du résultat de Matiiassevitch soit élaboré explicitement et publié en 1976 par J. Jones, D. Sato, H. Wada et D. Wiens. Le voici :

$$(1-(wz+h+j-q)^2-((gk+2g+k+1)(h+j)+h-z)^2-(2n+p+q+z-e)^2-(16(k+1)^3(k+2)(n+1)^2+1-f^2)^2-(e^3(e+2)(a+1)^2+1-o^2)^2-((a^2-1)y^2+1-x^2)^2-(16r^2y^4(a^2-1)+1-u^2)^2-(((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x+cu)^2)^2-(n+l+v-y)^2-((a^2-1)l^2+1-m^2)^2-(ai+k+1-l-i)^2-(p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m)^2-(q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x)^2-(z+pl(a-p)+t(2ap-p^2-1)-pm)^2)(k+2).$$

Quand les 26 variables $a,b,c,...,z$ prennent toutes les valeurs de l'ensemble des nombres entiers positifs, l'expression prend des valeurs positives et négatives. L'ensemble des valeurs positives prises est l'ensemble des nombres premiers.

L'éminent mathématicien soviétique You Linnik, à qui un collègue apprenait l'existence de ce polynôme inattendu, s'écria : «C'est merveilleux, très bientôt nous allons sans doute apprendre une quantité de choses nouvelles sur les nombres premiers.» On lui précisa alors qu'il existait un polynôme analogue pour tout ensemble de nombres énumérables par programme. «C'est lamentable, dit alors Linnik. Très probablement nous n'allons rien apprendre de nouveau sur les nombres premiers.»

Il est vrai que le polynôme est très décevant, car il ne prend que très rarement des valeurs positives (les seules intéressantes !) et, à moins d'étudier soigneusement la façon dont il a été construit, on ne réussit à tirer aucun nombre premier de ce polynôme qui, en théorie, les donne tous ! Il y a quelques années, je l'avais mentionné dans cette chronique, et plusieurs lecteurs m'avaient fait part de leur désillusion et même de leur doute concernant les propriétés annoncées. Je veux les rassurer : il n'y a aucune erreur, le polynôme possède bien les propriétés annoncées, mais il n'est pas plus utile pour calculer les nombres premiers que la formule de Willans. Les formules vraiment utiles, répétons-le, correspondent à des programmes et, parce qu'on ne leur impose pas un carcan artificiel par limitation du symbolisme, sont efficaces.

Jean-Paul DELAHAYE est professeur à l'Université de Lille. e-mail : delahaye@lifl.fr

Une feuille de calcul *Maple* associée à cet article, rédigée par Eric Wegrzynowski, est mise à la disposition des lecteurs à l'adresse : http: / / www.lifl.fr / ~wegrzyno / FormPrem.html.

C. BOXA, *A Note on Diophantine Representation*, in *The American Mathematical Monthly*, pp. 138-143, février 1993.

C. CALDWELL, *The Prime Pages*. http: / / www.utm.edu / research / primes / .

G. HARDY et E. WRIGHT, *An Introduction to the Theory of Numbers*, Oxford Science Publications, Clarendon Press, Oxford, cinquième édition, 1979.

Y. MATIIASSEVITCH, *Le dixième problème de Hilbert. Son indécidabilité*, Masson, Paris, 1995.

P. RIBENBOIM, *Nombres premiers : mystères et records*, Presses universitaires de France, 1994.

***Le secret des nombres* (arithmétique pour l'enseignement de spécialité de terminale S), ouvrage collectif des Éditions Archimède, ISSN 0987-0806, *Tangente*, hors série n° 6, 1998.**