DES PREMIERS AUX PSEUDO-PREMIERS

Toujours aussi importants en informatique, les nombres premiers suscitent des travaux foisonnants. Leur quête par des tests probabilistes a ouvert un riche terrain de jeu.

L'AUTEUR

JEAN-PAUL DELAHAYE
professeur émérite à l'université de Lille et chercheur au laboratoire Cristal (Centre de recherche en informatique, signal et automatique de Lille)

Euclide le savait déjà : les nombres entiers se décomposent en atomes multiplicatifs irréductibles, qu'on appelle les « nombres premiers ». Par définition, un nombre premier est un nombre entier qui possède exactement deux diviseurs : 1 et lui-même. Les plus petits nombres premiers sont 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, etc. Le mathématicien grec savait démontrer qu'il y en a une infinité. Quand un nombre n'est pas premier, on dit qu'il est « composé ».

Tout entier se décompose d'une manière unique en produit de nombres premiers. Mais trouver explicitement la décomposition d'un nombre en un produit de nombres premiers est un problème de calcul qu'on étudie depuis plus de deux millénaires, de même que celui consistant à découvrir de très grands nombres premiers. On s'y intéresse d'autant plus aujourd'hui que ces questions jouent un rôle important dans la science des codes secrets et de la sécurité informatique – la cryptographie.

La méthode la plus simple pour savoir si un nombre entier p est premier est de tenter les divisions de p par a, pour tout entier a compris entre 2 et $p-1$. Si la division ne tombe jamais juste, p est premier. Sinon, on aura un début de factorisation. Aller jusqu'à $p-1$ n'est en réalité pas nécessaire : il suffit de tenter les divisions jusqu'à la racine carrée de p (car si $p=a\times b$, avec $a>1$ et $b>1$, alors a et b ne peuvent pas tous deux être strictement plus grands que la racine carrée de p).

La méthode des divisions successives est impossible à pratiquer si on s'intéresse à des entiers de plus de 40 chiffres, car le nombre de divisions à faire serait trop important, même pour nos puissants ordinateurs. Il faut donc trouver d'autres méthodes pour factoriser et repérer les nombres premiers.

Il est clair que si l'on dispose d'un algorithme permettant rapidement de factoriser tout nombre entier en produit de nombres premiers, on saura aussi rapidement tester si un nombre p quelconque est premier ou non. L'inverse n'est pas vrai, car, même si cela semble paradoxal, il est possible de savoir qu'un nombre est composé sans savoir le factoriser. Le petit théorème de Fermat, mentionné en 1640 dans une lettre du magistrat mathématicien Pierre de Fermat à Bernard Frénicle de Bressy, indique en effet que si p est un nombre premier et si a est un entier non multiple de p, alors $(a^{p-1} \bmod p)=1$. Rappelons que $(b \bmod p)$ désigne le reste de la division de b par p : $(15 \bmod 6)=3$; $(125 \bmod 10)=5$, etc.

Il résulte de ce théorème (par contraposée) que si, pour un nombre p, on trouve un entier a non multiple de p tel que $(a^{p-1} \bmod p)\neq 1$, alors on sait que p n'est pas un nombre premier, et cela sans connaître aucun de ses facteurs. Le fait que $(2^{168} \bmod 169)=38$ signifie donc que 169 n'est pas premier, mais n'indique rien de sa factorisation. Précisons que le calcul de $(a^b \bmod p)$ peut sembler difficile, mais qu'on connaît des méthodes qui l'effectuent efficacement en menant quelques multiplications et divisions entre nombres de taille raisonnable (*voir l'encadré 1*).

Jean-Paul Delahaye a notamment publié : **Au-delà du Bitcoin** (Dunod, 2022).

Savoir si un nombre est composé ou premier apparaît donc plus facile que savoir le factoriser. Des résultats théoriques assez récents le confirment.

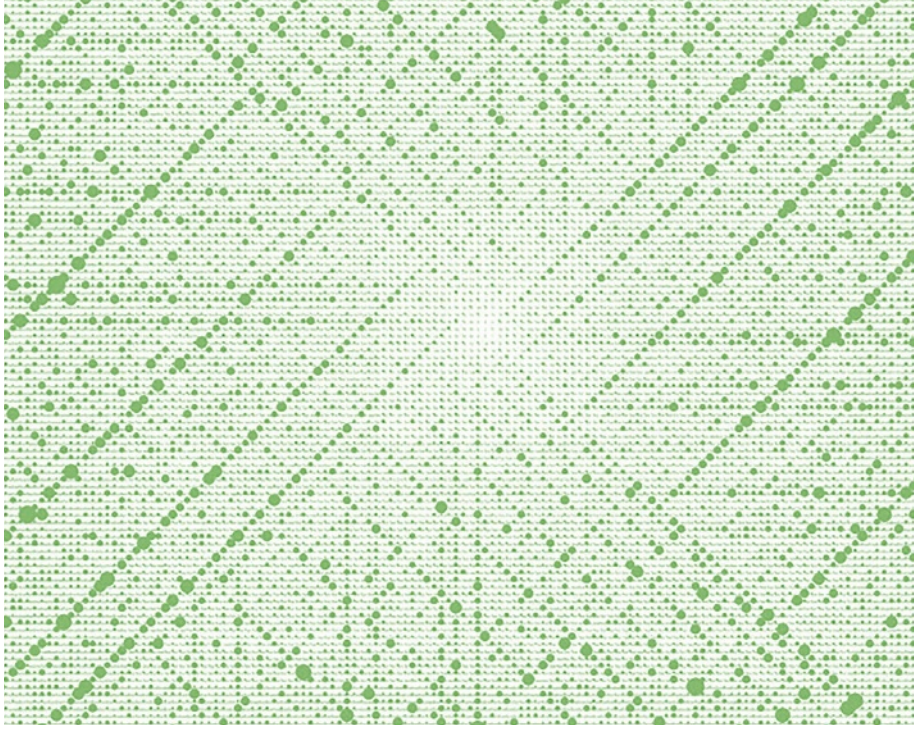
LE DÉFI DE LA FACTORISATION

En effet, en utilisant le meilleur algorithme possible, le problème de savoir si un nombre n est premier ou composé – appelé « problème de la primalité » – ne demande qu'un calcul d'une durée majorée par un polynôme de la longueur k du nombre n écrit en décimal. On dit que le problème du test de primalité est « polynomial ». Ce résultat était soupçonné depuis longtemps, mais il n'a été démontré qu'en 2002 – puis publié en 2004 – par les mathématiciens Manindra Agrawal, Neeraj Kayal et Nitin Saxena, de l'Institut indien de technologie de Kanpur, en Inde. Les chercheurs reçurent en récompense le prix Gödel en 2006.

En 2019, la preuve de leur résultat a été formalisée – c'est-à-dire contrôlée en détail à l'aide d'un système d'assistant de preuve informatique – par Hing-Lun Chan, de l'université nationale australienne, à Canberra, ce qui en garantit la validité. Le degré du polynôme mentionné ci-dessus, qu'on souhaite le plus petit possible, est égal à 12, mais si l'on accepte de considérer vraies certaines conjectures, il serait égal à 6 ou même à 3.

Il n'a en revanche pas été démontré que le problème de factoriser un nombre n est polynomial, et on pense d'ailleurs qu'il ne l'est pas. Factoriser un nombre serait donc intrinsèquement plus difficile que tester sa primalité. Cependant, si l'on réussit à fabriquer des ordinateurs quantiques assez puissants, alors on disposera d'un algorithme quantique polynomial de factorisation. C'est un résultat de 1994, dû au mathématicien américain Peter Shor. Il faut $5k+1$ bits quantiques pour factoriser un nombre dont l'écriture binaire comporte k chiffres (ce qui correspond à une écriture décimale d'environ $0{,}3\times k$ chiffres) et un temps proportionnel à k^3.

Le record actuel de factorisation a été établi en 2020 par une équipe internationale franco-américaine, et porte sur un nombre de 250 chiffres décimaux. Or on considère que si l'on sait factoriser un nombre de n chiffres obtenu comme le produit de deux nombres premiers de taille environ $n/2$ – c'est le cas pour le nombre de 250 chiffres de ce record –, alors on sait factoriser tout nombre de n chiffres. On estime donc aujourd'hui que la méthode utilisée en 2020, appelée « Number Field Sieve », permet de factoriser un nombre quelconque de 250 chiffres. Il est important de préciser « quelconque », car il est évident que si l'on se donne, par exemple, le développement décimal de 2^{1000},



La spirale d'Ulam est une manière de visualiser les nombres premiers. Elle consiste à écrire les nombres entiers en suivant une spirale tournant dans le sens trigonométrique, et à signaler les nombres premiers par un point de couleur. Le dessin *ci-dessus* est une variante où, en chaque point n, on dessine un disque ayant un rayon proportionnel au nombre de diviseurs de n.

© Cortexd/Wikimedia commons

1 L'EXPONENTIATION RAPIDE

En calcul numérique, qu'il s'agisse de trouver des nombres premiers ou d'autres problèmes, il est fréquent de devoir calculer la puissance n-ième d'un nombre a modulo un nombre p – c'est-à-dire le reste de la division de a^n par p, noté $(a^n \bmod p)$.
Une méthode efficace, devenue essentielle, consiste à utiliser la définition récursive – c'est-à-dire qui fait appel à elle-même – selon laquelle $(a^n \bmod p)$ vaut :
- $(a \bmod p)$ si $n = 1$;
- $((a^{n/2} \bmod p)^2 \bmod p)$ si n est pair ;
- $(((a^{n//2} \bmod p)^2 \times a) \bmod p)$, où $n // 2$ est le quotient de la division euclidienne de n par 2, si n est impair.
Les divisions par deux de l'exposant permettent d'arriver très rapidement au résultat. Pour un exposant n compris entre 1 024 et 2 047, la méthode produit 10 appels en cascade, conduisant tout au plus à 20 multiplications et 20 divisions, ce qui est bien mieux que la méthode naïve utilisant au moins 1 023 multiplications.

Exemple : calcul de $(3^{19} \bmod 5)$
Pour avoir $(3^{19} \bmod 5)$, on calcule $(3^9 \bmod 5)$.
Pour calculer $(3^9 \bmod 5)$, on calcule $(3^4 \bmod 5)$.
Pour calculer $(3^4 \bmod 5)$, on calcule $(3^2 \bmod 5)$, qui vaut $(9 \bmod 5) = 4$.
Donc $(3^4 \bmod 5) = (4^2 \bmod 5) = (16 \bmod 5) = 1$.
Donc $(3^9 \bmod 5) = ((1^2 \times 3) \bmod 5) = (3 \bmod 5) = 3$.
Donc $(3^{19} \bmod 5) = ((3^2 \times 3) \bmod 5) = (27 \bmod 5) = 2$.

Cette méthode d'exponentiation rapide est également exploitable avec des matrices. Cela permet, par exemple, un calcul rapide des grands termes de la suite de Fibonacci ou de celle de Perrin *(voir l'encadré 4)* sans avoir à calculer tous ceux qui précèdent.

qui possède plus de 300 chiffres décimaux, alors ce nombre sera facile à factoriser: son écriture décimale montre qu'il est pair, on le divise donc par 2; le résultat obtenu est encore pair, on le divise donc par 2, etc.

Précisons que la factorisation record a été menée sur un réseau d'ordinateurs et a demandé au total un calcul d'une durée équivalente à plus de deux mille années d'un bon ordinateur de bureau.

Le record de factorisation quantique date quant à lui de 2022, et concerne un nombre à 15 chiffres: 261980999226229 = 15538213 × 16860433. Le calcul a été conduit par une équipe chinoise menée par Bao Yan, de l'université Tsinghua, Ziqi Tan, de l'université du Zhejiang et Shijie Wei, de l'Académie des sciences de l'information quantique à Pékin. Le résultat est toutefois discuté car il s'appuie sur un calcul non quantique pour diminuer le nombre de qubits (les unités élémentaires de calcul d'un ordinateur quantique) nécessaires – seuls dix qubits sont ainsi utilisés dans le calcul chinois. Pour l'heure, les ordinateurs quantiques, potentiellement les meilleurs, restent donc loin derrière les ordinateurs classiques quand on leur demande de factoriser des grands nombres. La question de savoir si l'on saura mettre au point des ordinateurs quantiques qui changeraient la situation ne possède pas une réponse évidente, car les progrès sont lents et il est possible qu'au-delà d'un certain seuil on rencontre une difficulté physique insurmontable.

Les records de preuve de primalité concernent quant à eux des entiers plus grands: pour des nombres premiers quelconques, ils concernent des nombres atteignant 40000 chiffres décimaux, ou même un peu plus. L'imprécision vient de la difficulté de dire ce qu'est un nombre premier «quelconque». On remarque tout de même qu'on parle ici de nombres à plusieurs dizaines de milliers de chiffres décimaux, bien au-delà des 250 chiffres du record de factorisation. Cela confirme en pratique que factoriser est plus complexe que tester la primalité.

Pour les nombres premiers ayant une forme particulière et pour lesquels des techniques spécifiques fonctionnent, on va encore plus loin: le plus grand nombre premier connu aujourd'hui est $2^{82589933}-1$, qui comporte 24862048 chiffres décimaux! Sa primalité a été démontrée en 2018 par Patrick Laroche, un informaticien de 35 ans habitant en Floride, qui participait aux calculs organisés par le Great Internet Mersenne Prime Search (GIMPS). Ce site internet confie aux volontaires des entiers à tester qui, s'ils ont de la chance, deviennent les nombres premiers record. Tout le monde peut rejoindre cette initiative.

Il faut savoir qu'on connaît des formules qui engendrent une infinité de nombres premiers. On pourrait donc penser qu'il suffit de les utiliser pour calculer des nombres premiers aussi grands qu'on le souhaite – plus grands encore que celui de Patrick Laroche. La formule suivante donne, par exemple, tous les nombres premiers dans l'ordre: $p_1=2, p_2=3, p_3=5, p_4=7$, etc.

$$p_n = 1+\sum_{m=1}^{2^n}\left[\left[\frac{n}{1+\sum_{j=2}^{m}\left[\frac{1+(j-1)!}{j}-\left[\frac{(j-1)!}{j}\right]\right]}\right]^{\frac{1}{n}}\right]$$

Dans cette expression, les crochets $[x]$ désignent la partie entière de x (par exemple: $[3{,}78]=3$).

Cependant de telles formules sont inefficaces en pratique, car si elles donnent bien des

2 FORMULES POUR LES NOMBRES PREMIERS

La recherche de formules donnant une infinité de nombres premiers ou les donnant tous dans l'ordre est un exercice amusant, même si en pratique les formules obtenues n'aident pas à écrire explicitement de très grands nombres premiers.
La formule suivante, due au français Roland Yéléhada, qui l'a proposée et démontrée en 1999, est simple, et pourtant elle engendre tous les nombres premiers.

$$t(n) = 2+n\left[\frac{1}{1+\sum_{p=2}^{n+1}\left[\frac{n+2}{p}-\left[\frac{n+1}{p}\right]\right]}\right]$$

Les crochets désignent la partie entière.
La somme au dénominateur de la formule compte le nombre de diviseurs de $n+2$ entre 2 et $n+1$, nombre qui vaut 0 si est seulement si $n+2$ est premier. Cela a pour effet que le grand crochet vaut 0 si n est composé – et alors $t(n)=2$, qui est premier – et vaut 1 si $n+2$ est premier – et alors $t(n)=n+2$, qui est premier.

La formule donnée dans le texte de l'article est plus compliquée, mais elle a l'avantage de produire les nombres premiers un par un, dans l'ordre. On découvre encore aujourd'hui de telles formules-jeux. Voir par exemple l'article de Jean-Christophe Pain, de l'université Paris-Saclay, « A prime sum involving Bernoulli numbers », publié en 2023.
Le célèbre mathématicien Alain Connes a utilisé en 2019 la formule ci-dessous, qui lie entre eux le nombre π, le nombre $\prod(n)$ de nombres premiers compris entre 1 et n, et la fonction Γ (qui prolonge la fonction factorielle sur l'ensemble des réels) :

$$\prod(n) = \left[\sum_{k=1}^{n} \sin^2\left(\frac{\pi\Gamma(k)}{2k}\right)\right]$$

Ici encore, les crochets désignent la partie entière. Cela lui a permis de démontrer un surprenant et quasi paradoxal résultat en théorie de la mesure.

expressions valables pour des nombres premiers sans limitation de taille, il est impossible de les exploiter pour en tirer l'écriture décimale des nombres proposés – chose exigée pour réussir à battre un record. Elles restent cependant précieuses pour démontrer d'autres résultats de théorie des nombres ou d'analyse, comme l'illustre un résultat obtenu en 2019 par le mathématicien français Alain Connes, lauréat de la médaille Fields en 1982, qui utilise une formule explicite de ce type (*voir l'encadré 2*).

ACCEPTER UN RISQUE INFIME

Tous les résultats que nous avons mentionnés jusqu'ici concernent des méthodes de factorisation ou des tests de primalité qui ne présentent pas le moindre risque d'erreur (sous réserve que les ordinateurs requis pour faire les calculs ne commettent pas d'erreur, ce qu'on considère vrai). Peut-on aller plus vite et plus loin dans la taille des entiers concernés si l'on accepte un petit risque d'erreur?

La réponse est oui pour les tests de primalité. Le petit théorème de Fermat et ses variantes produisent des preuves «probabilistes» de primalité permettant d'obtenir assez rapidement de très grands nombres «presque certainement premiers».

Le petit théorème de Fermat donne la condition nécessaire pour que p soit premier: $(a^{p-1} \bmod p) = 1$ pour tout entier a non multiple de p. Mais cette condition, que nous appellerons le «test de Fermat avec a», n'est pas suffisante en général. Si, en calculant $(a^{p-1} \bmod p)$ pour un a donné non multiple de p, on trouve quelque chose de différent de 1, on est certain que p est composé (c'est la contraposée du théorème de Fermat). Mais si l'on trouve 1, cela ne prouve pas que p est premier. Dans la majorité des cas cependant, quand on trouve 1, alors p est premier! Autrement dit, le test de Fermat est «presque» une condition nécessaire et suffisante de primalité.

Pour $a=2$, par exemple, considérer que la condition $(2^{p-1} \bmod p) = 1$ est nécessaire et suffisante pour que p soit premier est correct pour tous les entiers jusqu'à $p=341$, qui est le produit de 11 et 31. Pour $a=5$, c'est aussi correct jusqu'à $p=217$, qui est le produit de 7 et 31. En exigeant de trouver 1 à la fois pour $a=2$ et $a=5$, on obtient la bonne réponse jusqu'à $p=561$, qui vaut $3\times11\times17$.

Ce nombre 561 est en fait très ennuyeux, car il met en défaut le test de Fermat pour tous les entiers a n'ayant aucun facteur premier avec 561 – en d'autres termes, pour tous ces entiers a, on obtient $(a^{560} \bmod 561) = 1$, alors que 561 est composé. De tels nombres portent le nom de «nombres de Carmichael», en référence au mathématicien américain Robert Carmichael (1879-1967). Ils font en quelque sorte semblant d'être premiers, alors qu'ils ne le sont pas. En 1994, William Alford, Andrew Granville et Carl Pomerance ont démontré qu'il y en a une infinité. Les plus petits sont 561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341, 41041, etc.

Si l'on cherche à engendrer de petits nombres premiers à l'aide de tests de Fermat, les nombres de Carmichael sont réellement gênants, car ils risquent de provoquer des erreurs. En revanche, si l'on recherche de grands nombres premiers – ayant par exemple 100 chiffres ou plus – on peut ignorer leur existence, car ils se font de plus en plus rares. Une piste explorée depuis longtemps consiste, donc, à accepter d'utiliser des tests qui se trompent parfois, mais rarement – et suffisamment rarement pour qu'en pratique ces tests ne produisent jamais d'erreur.

Un test très simple, quand on s'intéresse à un grand entier impair p et qu'on veut savoir s'il est premier, consiste à prendre au hasard un entier a entre 2 et $p-1$ et à faire confiance au test de Fermat avec a. Si l'on tire au hasard un nombre impair p de 100 chiffres ou moins, et qu'on lui applique ce test, la probabilité que le test dise que le nombre est premier, alors qu'il ne l'est pas, est inférieure à $2{,}8\times10^{-8}$. Si

3

LE MEILLEUR TEST PROBABILISTE

Le test de Rabin-Miller affine le test de Fermat – qui consiste à tester la primalité d'un entier p en vérifiant que $(a^{p-1} \bmod p) = 1$ pour un entier a non multiple de p – et le rend plus fiable. Il provient des travaux de Gary Miller, qui présente en 1975 une version du test sous une forme qui dépend de l'hypothèse de Riemann généralisée (qui n'est pas démontrée). Michael Rabin en tire ensuite un test probabiliste indépendant de cette hypothèse.
Ce test repose sur le théorème suivant. Soit p un nombre premier. On écrit $p-1$ sous la forme $2^s \times d$, avec d impair (par exemple : $57-1 = 56 = 8\times7 = 2^3\times7$). Soit a un entier entre 2 et $p-1$. Alors, de deux choses l'une : soit $(a^d \bmod p) = 1$, soit il existe un entier r entre 0 et $s-1$ tel que $((a^{d\times2^r}+1) \bmod p) = 0$.
En pratique, pour tester la primalité d'un nombre p, on divise $p-1$ par 2 autant de fois que c'est possible, et cela donne s et d. Ensuite, on calcule $b = (a^d \bmod p)$. Si $b = 1$, le test est satisfait. Sinon, on élève b au carré $(s-1)$ fois de suite, et on regarde à chaque fois si le résultat modulo p vaut -1.
Si cela ne se produit pas, le test a échoué et l'on peut être certain que p est composé. Les a qui font échouer le test et indiquent que p est composé sont appelés « témoins de non-primalité pour p ».
Un résultat donne la précieuse information que si p est un nombre impair composé, alors 3/4 au moins des a entre 2 et $p-1$ sont des témoins de non-primalité. Autrement dit, en prenant un nombre a entre 2 et $p-1$ au hasard, on a au moins 75 % de chance de repérer que p est composé, si c'est le cas.
Il en résulte que, si l'on a réussi à passer le test k fois de suite pour un nombre p, alors la probabilité que p ne soit pas premier est inférieure à $1/4^k$. Ce risque d'erreur, qu'on peut rendre aussi petit qu'on le souhaite en prenant la bonne valeur de k, permet en pratique d'obtenir des nombres premiers de cent chiffres ou plus, utilisables en cryptographie.

4 LA SUITE DE RAOUL PERRIN

La suite de Raoul Perrin est remarquable à plusieurs titres. Elle est définie par des formules qui ressemblent à celles appliquées pour la suite de Fibonacci : $u(0) = 3$; $u(1) = 0$; $u(2) = 2$ et $u(n) = u(n-2) + u(n-3)$ si $n > 2$. Cela donne les valeurs successives : 3 ; 0 ; 2 ; 3 ; 2 ; 5 ; 5 ; 7 ; 10 ; 12 ; 17 ; 22 ; 29 ; 39 ; 51 ; 68 ; 90 ; 119 ; 158 ; 209…

Notons r l'unique nombre réel qui vérifie $r^3 = 1 + r$. Ce nombre, appelé « nombre plastique », vaut :

$$\psi = \sqrt[3]{\frac{1}{2} + \frac{\sqrt{69}}{18}} + \sqrt[3]{\frac{1}{2} - \frac{\sqrt{69}}{18}}$$

soit environ : 1,3247.

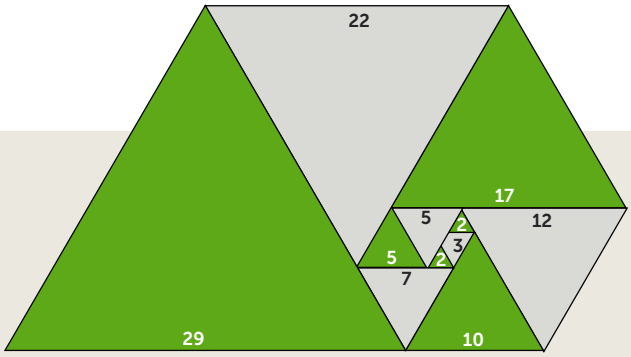
Il joue le même rôle que le nombre d'or pour la suite de Fibonacci, puisque le rapport $u(n+1) / u(n)$ tend vers r lorsque n tend vers l'infini. Il permet de calculer $u(n)$ simplement, car $u(n) = [r^n + 1/2]$ pour $n > 9$ (où le crochet désigne la partie entière).

La suite de Perrin possède l'étrange propriété que si p est premier, alors p divise $u(p)$, et que la réciproque est presque toujours vraie. Le premier nombre p divisant $u(p)$ sans être premier est $p = 271\,441 = 521^2$. On dit que c'est un « pseudo-premier de Perrin ».

Si $n > 4$, on a :

$$\begin{aligned} u(n) &= u(n-3) + u(n-2) \\ &= u(n-3) + (u(n-4) + u(n-5)) \\ &= (u(n-3) + u(n-4)) + u(n-5) \\ &= u(n-1) + u(n-5) \end{aligned}$$

De la formule $u(n) = u(n-1) + u(n-5)$ pour $n > 4$, on tire le magnifique pavage du plan ci-dessus, qui utilise uniquement un tout petit trapèze et des triangles équilatéraux de côté $u(n)$ pour $n \geq 2$.



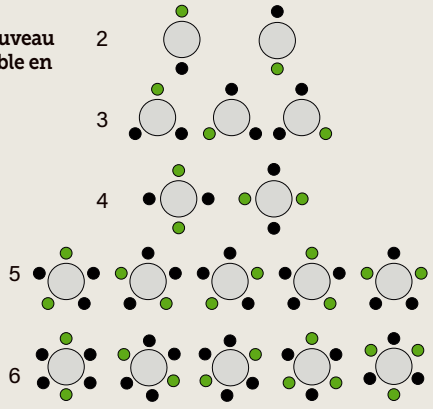
Chose étrange : imaginons qu'à cause d'une mauvaise épidémie, des convives cherchent à s'asseoir à une table ronde autour de laquelle sont disposées n chaises en respectant les deux consignes ci-dessous.

(a) Deux convives ne doivent pas occuper des chaises adjacentes ;

(b) Il ne doit plus y avoir aucun nouveau convive qui puisse s'ajouter à la table en respectant la première consigne.

Alors le nombre de façons d'occuper la table est égal à $u(n)$ dès que $n \geq 2$. Le dessin ci-contre l'illustre en présentant les configurations possibles pour 2 ; 3 ; 4 ; 5 et 6 chaises.

On retrouve bien : $u(2) = 2$; $u(3) = 3$; $u(4) = 2$; $u(5) = 5$ et $u(6) = 5$. On remarque que dans les configurations à 6 chaises possibles, certaines permettent d'asseoir 2 convives, d'autres 3 convives.



l'on tire au hasard un nombre impair p quelconque, cette fois de 200 chiffres ou moins, cette probabilité de faux positif est inférieure à $3{,}9 \times 10^{-27}$. Pour un nombre impair p de 1000 chiffres ou moins tiré au hasard, le risque est inférieur à $1{,}2 \times 10^{-123}$. Cela signifie que si vous trouvez des nombres de 200 chiffres avec cette méthode à raison de un par seconde pendant un million d'années, le risque que l'un de ces nombres ne soit pas premier est inférieur à 10^{-12} (une chance sur mille milliards) !

Puisqu'aujourd'hui on peut, avec un peu de soin, manipuler des nombres de plusieurs milliards de chiffres et leur appliquer le test de Fermat, il est possible de connaître des nombres « presque certainement premiers » très grands, bien au-delà des 24 millions de chiffres du record détenu par $2^{82589933} - 1$. Personne, cependant, ne s'amuse à cela, car le point de vue du mathématicien est qu'on ne peut pas accepter de dire qu'un nombre est premier quand c'est un test probabiliste qui l'affirme. Dans le domaine de la cryptographie, cependant, on n'a pas cette fausse pudeur et, lorsqu'on a besoin de grands nombres premiers explicites, on n'hésite pas à mettre en œuvre des tests de primalité probabilistes.

LE TEST DE MILLER-RABIN

La méthode qu'on préfère en pratique n'est pas celle utilisant le test de Fermat, car le risque pour des nombres de 100 chiffres ou moins est jugé trop grand. On applique une méthode fondée sur un raffinement du théorème de Fermat, appelée « test de Miller-Rabin » (*voir l'encadré 3*). Cette variante du test de Fermat est aujourd'hui encore étudiée. Des résultats de Matt Kownacki, mis en ligne sous la forme de *preprint* en 2018 mais non publiés dans une revue à comité de lecture, ont confirmé son bon fonctionnement, et donc le caractère infime du risque qu'on prend en l'exploitant. Ce résultat indique que quand un entier p est composé, alors les témoins de non-primalité – qui jouent un rôle équivalent dans le test de Miller-Rabin aux nombres a tels que $(a^{p-1} \bmod p) \neq 1$ dans le test de Fermat – sont non seulement nombreux – il y en a au moins 75% parmi les entiers impairs entre 2 et $p-1$ –, mais qu'ils sont uniformément répartis. Cela rend presque impossible, en

pratique, de ne pas tomber sur l'un d'eux rapidement en essayant quelques entiers a pris au hasard. Autrement dit: il n'y a pas de piège sournois à craindre avec le test de Miller-Rabin.

En réalité, le choix du nombre a conduisant à un risque d'erreur aussi faible que possible peut être fait mieux qu'en le tirant au hasard: en 2022, une étude menée par Nikolai Antonov et Shamil Ishmukhametov, de l'université de Kazan, en Russie, a permis d'optimiser ce choix en mettant en œuvre une méthode d'apprentissage par renforcement.

Il a par ailleurs été démontré depuis longtemps que la probabilité que l'algorithme déclare faussement premier un nombre composé, quand il a réussi à passer le test n fois, est inférieure à 4^{-n}. Cela signifie qu'en appliquant le test de Miller-Rabin avec 20 nombres a choisis au hasard, le risque d'erreur quand le test indique que p est premier est inférieur à 1/1000000000000 (un sur mille milliards)!

Lorsqu'on ne s'intéresse qu'à des nombres de taille bornée, les tests probabilistes peuvent même redonner des tests sans erreurs (donc non probabilistes). Par exemple, pour les entiers p inférieurs à 2^{64} (environ $1{,}8\times10^{19}$), si le test de Miller-Rabin est positif avec les nombres a = 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, alors il est certain que p est premier.

PSEUDO-PREMIERS DE PERRIN

Les nombres composés qui trompent les tests, comme 561 trompe le test de Fermat, sont appelés «pseudo-premiers». Ils sont intéressants à étudier et il y en a une multitude de catégories puisqu'il existe un grand nombre de tests de primalité probabilistes. Il existe, par exemple, les nombres pseudo-premiers de Catalan, d'Euler, de Frobenius, de Lucas, etc. Les nombres pseudo-premiers de Perrin sont particulièrement étonnants. Ils ont suggéré à Robert Dougherty-Bliss et Doron Zeilberger, de l'université Rutgers, aux États-Unis, de nouvelles variantes de tests de primalité, présentées dans un article de 2023, qui sont le résultat d'une série de calculs informatiques massifs.

Considérons la suite numérique $u(n)$ définie par l'ingénieur français Raoul Perrin (1841-1910) par les formules: $u(0)=3$; $u(1)=0$; $u(2)=2$ et $u(n)=u(n-2)+u(n-3)$ si $n>2$. Ses vingt premiers termes sont: 3; 0; 2; 3; 2; 5; 5; 7; 10; 12; 17; 22; 29; 39; 51; 68; 90; 119; 158; 209.

Perrin remarque que si p est premier, alors $u(p)$ est multiple de p, et on peut le démontrer. Il lui semble même que la condition est suffisante. On peut en effet vérifier que:
- $u(2)=2$ est multiple de 2, qui est premier;
- $u(3)=3$ est multiple de 3, qui est premier;
- $u(4)=2$ n'est pas multiple de 4, qui n'est pas premier;
- $u(5)=5$ est multiple de 5, qui est premier;
- $u(6)=5$ n'est pas multiple de 6, qui n'est pas premier; etc.

L'ingénieur français ne trouva aucun contre-exemple à cette règle: il a fallu attendre 1982 pour que William Adams et Daniel Shanks découvrent que le nombre $p=271441$ est composé (car $271441=521^2$), mais vérifie bien que $u(271441)$ est un multiple de p. C'est le plus petit nombre pseudo-premier de Perrin. On remarque qu'il est beaucoup plus grand que le plus petit pseudo-premier de Fermat, qui est 561. Le deuxième nombre pseudo-premier de Perrin est $904631=7\times13\times9941$. Il n'y en a que dix-sept inférieurs à 1 milliard, mais on sait aussi, grâce à un résultat de 2010 de Jon Grantham, qu'il y a une infinité de pseudo-premiers de Perrin. Il existe des méthodes efficaces pour effectuer le calcul de $u(p)$ en ramenant le calcul à l'évaluation d'une matrice à une certaine puissance, calcul qu'on mène en appliquant la méthode expliquée dans l'encadré 1. La recherche de nombres premiers en utilisant la suite de Perrin est donc à la fois peu risquée et d'un coût raisonnable.

Lionel Fourquaux, de l'université de Rennes, a proposé en 2015 des résultats permettant de connaître des familles infinies de nombres pseudo-premiers de Perrin. En exploitant ces méthodes pour ne pas se faire piéger, on améliore encore la méthode probabiliste de Perrin pour calculer de grands nombres premiers. La méthode de Miller-Rabin est toujours celle préférée en pratique, car quand on la programme elle reste meilleure que les autres: on n'a pas encore de résultats probabilistes aussi bons pour les autres tests que ceux dont on dispose pour celui de Miller-Rabin. Mais peut-être finira-t-il par être concurrencé par une méthode tirée d'une des nombreuses catégories de pseudo-premiers, qu'il faut continuer à explorer avec soin.

Pendant l'épidémie de Covid-19, en 2021, Vincent Vatter, professeur à l'université de Floride, a proposé une caractérisation amusante des termes de la suite de Perrin. Elle consiste à compter le nombre de façons de placer des convives autour d'une table disposant de n chaises, en s'imposant qu'il n'y ait jamais deux convives occupant des sièges adjacents (*voir l'encadré 4*).

Puisqu'il semble qu'on continue de progresser dans notre compréhension de l'arithmétique des nombres premiers et que, en même temps, on dispose de machines de plus en plus puissantes, il est probable que les records de calcul seront rapidement battus. Peut-être connaîtra-t-on bientôt un nombre premier de 100 millions de chiffres? Cela rapporterait 150000 dollars américains à ses découvreurs. Mieux encore: un nombre premier de 1 milliard de chiffres serait récompensé de 250000 dollars. Ces prix sont offerts par l'Electronic Frontier Foundation (EFF), une ONG californienne de protection des libertés sur internet. ■

BIBLIOGRAPHIE

R. Dougherty-Bliss et D. Zeilberger, Lots and lots of Perrin-type primality tests and their pseudo-primes, *arXiv preprint*, 2023.

L. Fourquaux, Construction de nombres pseudo-premiers de Perrin, 2023.

V. Vatter, Social distancing, primes, and Perrin numbers, *Math Horizons*, 2022.

N. Antonov et Sh. Ishmukhametov, An intelligent choice of witnesses in the Miller–Rabin primality test. Reinforcement learning approach, *Lobachevskii Journal of Mathematics,* 2022.

P. Zimmermann, Factorization of RSA-250, *Archive Cado-nsf-discuss,* Inria, 2020.

A. Connes, Around Wilson's Theorem, *Journal of Number Theory,* 2019.

M. Kownacki, On the distribution of witnesses in the Miller-Rabin test, *arXiv preprint,* 2016.

R. Perrin, Question 1484, *L'Intermédiaire des Mathématiciens*, 1899.