La factorisation des grands nombres

JOHANNES BUCHMANN

Les systèmes modernes de cryptage de données seront sûrs tant que la décomposition des nombres de plus de 100 chiffres en facteurs premiers restera difficile. Toutefois les techniques de factorisation progressent rapidement.

Le nombre 114 381 625 757 888 867 669 235 779 976 146 612 010 218 296 721 242 362 562 561 842 935 706 935 245 733 897 830 597 123 563 958 705 058 989 075 147 599 290 026 879 543 541 est le produit de deux nombres premiers ; lesquels? Martin Gardner posa cette question aux lecteurs de *Pour la Science* en octobre 1977, dans sa rubrique de «Jeux mathématiques», mais une réponse ne fut donnée que 16 ans plus tard : en avril 1994, Paul Leyland, de l'Université d'Oxford, Michael Graff, de l'Université de l'Iowa, et Derek Atkins, de l'Institut de technologie du Massachusetts, identifièrent les deux facteurs, après avoir distribué des parties de la tâche, grâce au réseau *Internet*, à quelque 600 volontaires, qui laissèrent fonctionner sur leurs ordinateurs, pendant de nombreuses nuits, le programme écrit par Arjen Lenstra, du Centre de recherches de la Société *Bell Communications*.

La multiplication de deux nombres, même très grands, n'est pas compliquée : avec du papier et un crayon, on calcule le produit de deux nombres de 65 chiffres en une heure environ ; par ordinateur, le calcul est immédiat. En revanche, l'opération inverse, c'est-à-dire l'identification des facteurs d'un produit, est très difficile, même avec les calculateurs les plus rapides.

Les opérations mathématiques telles que la multiplication et la factorisation sont à la base des systèmes cryptographiques modernes : le cryptage est rapide, mais le décryptage est quasi impossible en pratique.

En 1978, Ronald Rivest, de l'Institut de technologie du Massachusetts, Adi Shamir, de l'Institut Weizmann à Rehovot (Israël), et Leonard Adleman, de l'Université de Californie du Sud, ont mis au point un protocole de cryptage, nommé RSA d'après leurs initiales, qui est fondé sur la factorisation : une personne voulant recevoir des messages cryptés choisit deux nombres premiers p et q, c'est-à-dire deux nombres entiers naturels qui ne sont divisibles que par 1 et par eux-mêmes ; il calcule leur produit $p \times q$ et le rend public, tout en conservant secrets les facteurs *(voir l'encadré de la page 90)*. Pour crypter un message, il suffit de connaître ce produit, nommé la clé publique, tandis que, pour décrypter un message, il faut connaître les facteurs premiers p et q : si ceux-ci ont plus de 150 chiffres, même les meilleures méthodes connues et les ordinateurs les plus puissants mettront 2 000 ans pour les trouver. Ainsi, on fabrique facilement des problèmes de factorisation, mais on ne sait pas, aujourd'hui, les résoudre en un temps raisonnable, si les facteurs premiers sont trop grands.

Trouvera-t-on un jour une méthode de factorisation rapide? Les Laboratoires RSA ont organisé un concours mondial de factorisation. Ils publient des produits de grands nombres premiers et récompensent leur factorisation *(on peut consulter leur site sur le réseau Internet : voir la bibliographie)*.

L'idée de mettre le public à contribution est judicieuse : on ignore si la factorisation est difficile par essence ou si les mathématiciens n'ont pas encore trouvé la méthode la plus habile. Aussi la seule garantie de la sécurité des procédés de cryptage est l'ignorance d'une méthode rapide de factorisation des nombres entiers.

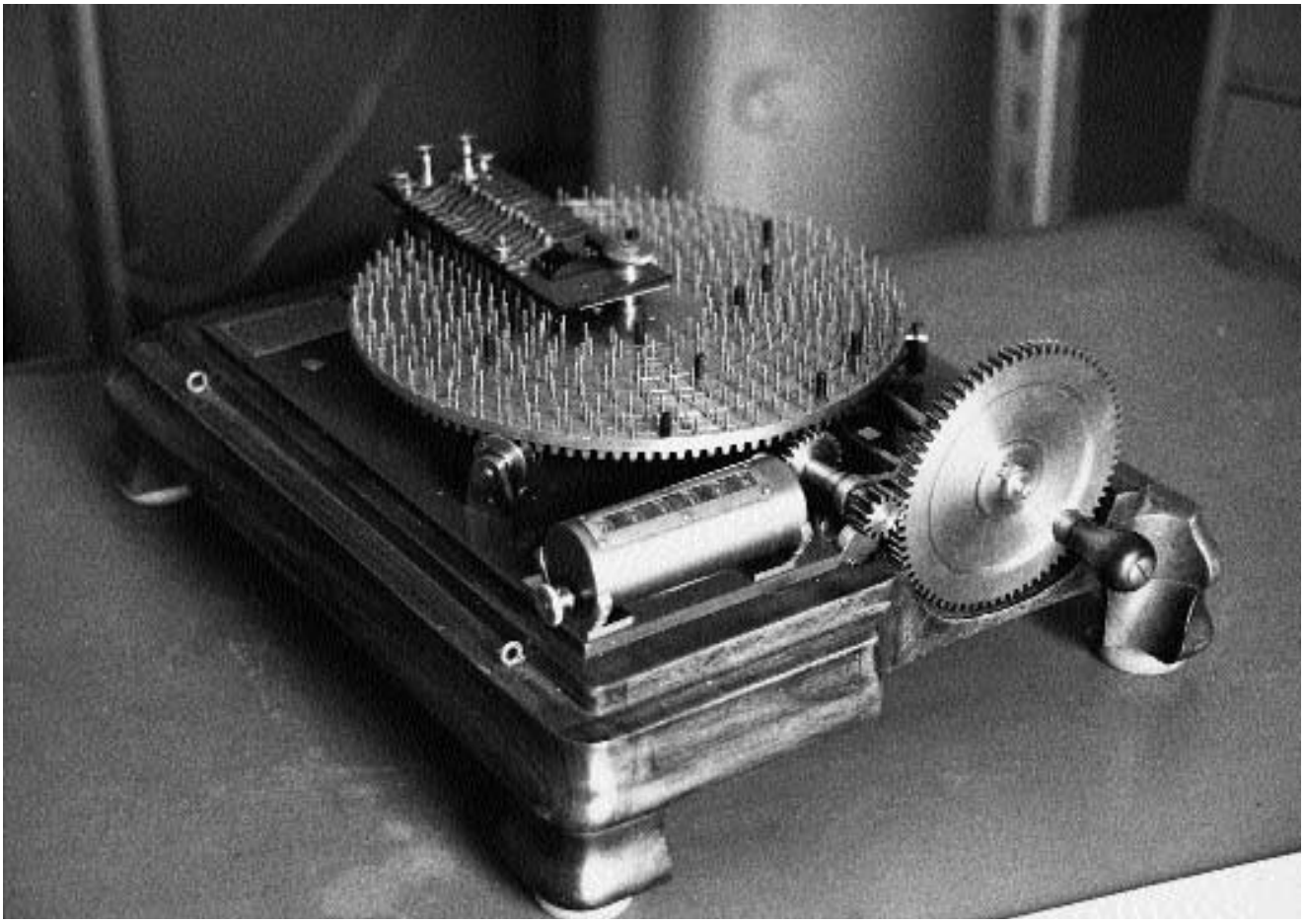
L'étude de la factorisation date de l'Antiquité : les mathématiciens d'alors savaient déjà que chaque nombre naturel est un produit de nombres premiers, et que la décomposition en facteurs premiers est unique, à l'ordre près. Par exemple, 12 se décompose seulement en $2 \times 2 \times 3$. L'étude des propriétés des nombres entiers naturels impose souvent la décomposition en facteurs premiers.

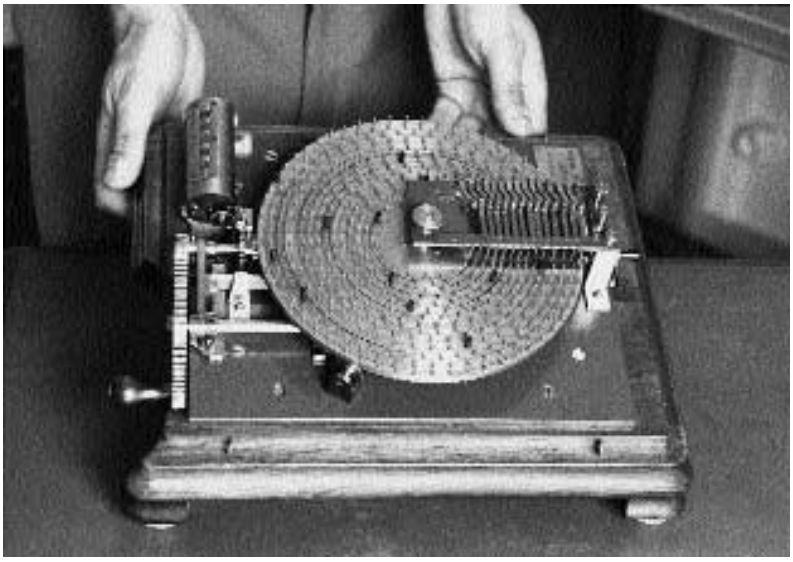
Les ordinateurs ont beaucoup apporté à la théorie des nombres. Dans cet article, nous verrons comment fonctionnent les algorithmes modernes de factorisation. Souvent les indications seront suffisamment précises pour qu'une programmation soit possible sur un ordinateur personnel ; en outre, une bibliothèque de programmes nommée LiDIA est proposée sur le réseau *Internet (voir la bibliographie)*.

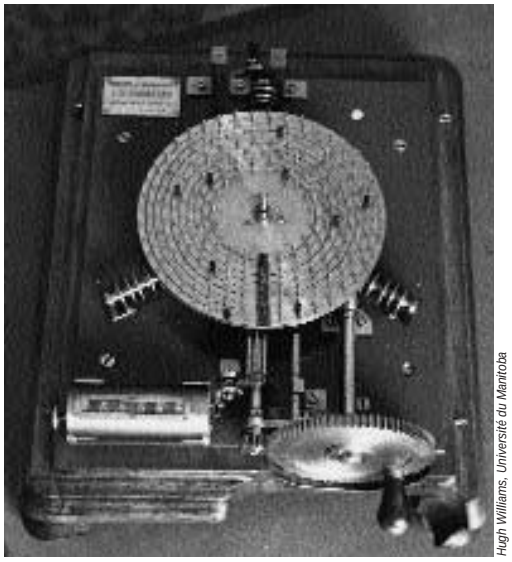
Les nombres de Fermat

Le juriste français Pierre de Fermat (1601-1665), notamment célèbre pour sa conjecture démontrée en 1997, pensait avoir trouvé une méthode pour fabriquer des nombres premiers aussi grands que l'on voulait : à partir des nombres entiers naturels i, il construisait les nombres $F_i = 2^{2^i} + 1$, nommés aujourd'hui nombres de Fermat. Les quatre premiers nombres de Fermat sont premiers : $F_0 = 2^{2^0} + 1 = 3$, $F_1 = 2^{2^1} + 1 = 5$, $F_2 = 2^{2^2} + 1 = 17$, $F_3 = 2^{2^3} + 1 = 257$ et $F_4 = 2^{2^4} + 1 = 65\,537$.

Cependant, dès 1732, le mathématicien suisse Leonhard Euler trouva que

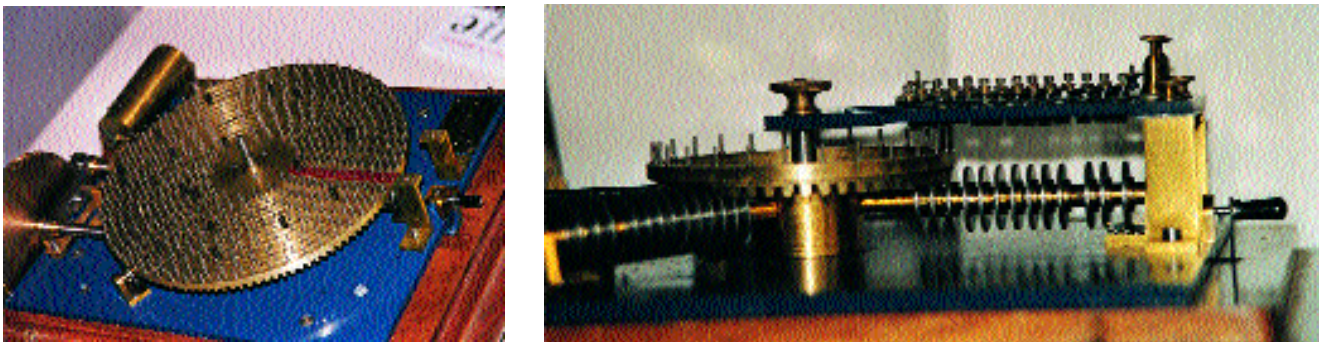






Hugh Williams, Université du Manitoba

1. LA «MACHINES À CONGRUENCES» de l'officier français Eugène Carissan (1880-1925). La manivelle fait tourner une roue dentée qui entraîne un compteur *(au premier plan sur la photographie supérieure)* et 14 couronnes indépendantes. Sur la photographie inférieure, à droite, les quatre couronnes externes ont été enlevées et le mécanisme apparaît. À chaque couronne est associé un nombre, de l'intérieur vers l'extérieur : 19, 21, 23, 26, 29, 31, 34, 37, 41, 43, 47, 55 et 59. Sur la face supérieure, chaque couronne porte un nombre de picots égal à ce nombre. Soit n un nombre à factoriser. On utilise la relation $n = x^2 - y^2$ et l'on cherche des conditions sur x pour que y^2 soit un carré parfait. Si les valeurs permises de x sont 1, 7, 8, 13, 14 et 20 modulo 21, c'est-à-dire si le reste de la division de x par 21 peut être 1, 7, 8, 13, 14, 20, alors on place un capuchon métallique sur les picots 1, 7, 8, 13, 14, 20 de la couronne 21. De même pour les autres couronnes utilisées. On tourne la manivelle et, quand les capuchons métalliques (les plots noirs) des différentes couronnes sont alignés sur la barrette, un contact électrique s'établit et un signal sonore retentit. Le nombre indiqué sur le cadran a alors des chances d'être un carré parfait. La machine est particulièrement adaptée pour la recherche de nombres convenant pour le crible quadratique, utilisé pour la factorisation des nombres. Carissan prouva avec elle, en tournant la manivelle pendant dix minutes, que 708 158 977 est un nombre premier, en montrant qu'il s'écrit d'une seule manière comme somme de deux carrés : $19\,224^2 + 18\,401^2$. La machine est aujourd'hui au Conservatoire national des arts et métiers, à Paris.



2. TOUTES LES COURONNES de la «machine à congruences» ont une position zéro (visible en rouge sur la photographie de gauche, où l'on a enlevé la barrette métallique). Les couronnes ne roulent pas seulement sur la roue dentée commune, mais aussi sur des disques, dont les axes, fixés à l'axe de la roue dentée, font des angles de 120 degrés les uns avec les autres *(à droite)*. Seules les pointes portant un capuchon produisent le contact dans la barrette métallique.

F_5, égal 4 294 967 297, était égal au produit de 641 par 6 700 417 : F_5 n'est donc pas un nombre premier (on dit que c'est un nombre composé). Puis, 150 ans plus tard, en 1880, F_6 fut factorisé et, encore 90 ans plus tard, en 1970, les deux facteurs premiers de F_7 ont été obtenus. Le nombre F_8 a été décomposé en 1980, et le nombre F_9 en 1990.

Cette histoire des nombres de Fermat montre combien la factorisation est difficile : les mathématiciens mirent 240 ans pour factoriser le nombre de Fermat F_7, à 39 chiffres, et, même équipés d'ordinateurs ils ont mis 20 ans à décomposer le nombre F_9, de 155 chiffres.

En revanche, il savaient que F_7, F_8 et F_9 n'étaient pas premiers bien avant de connaître leurs facteurs : on peut déterminer, grâce à un test spécial, si un nombre est composé, sans calculer ses facteurs. On s'assure ainsi que la factorisation est possible, avant de l'entreprendre.

Un test simple de primalité

Comment détermine-t-on qu'un nombre est premier? On peut utiliser le test de Fermat, c'est-à-dire examiner si n est un diviseur de $2^{n-1}-1$: Fermat a prouvé que tout nombre premier n (sauf 2) est diviseur de 2^{2n-1}. Par exemple, 3, qui est premier, est un diviseur de $2^{3-1}-1$, soit 3. Inversement, si n n'est pas un diviseur de $2^{n-1}-1$, n n'est pas premier. Par exemple, le nombre 6 n'est pas un diviseur de $2^{6-1}-1$, soit 31 ; il n'est pas premier.

Examinons un exemple plus compliqué, afin d'apprécier l'utilité du test de Fermat : le nombre 58 483 est-il premier? On pourrait commencer par calculer $2^{58\,482}-1$, puis chercher si le résultat est un multiple de 58 483, mais un calcul direct serait extrêmement long, car le nombre $2^{58\,482}-1$ s'écrit avec 17 604 chiffres. On utilise plutôt l'arithmétique des congruences *(voir l'encadré de la page 91)*, pour obtenir un algorithme efficace, fondé sur l'idée suivante : vérifier que 58 483 est un diviseur de $2^{58\,482}-1$ revient à chercher si le reste de la division de $2^{58\,482}$ par 58 483 est 1.

Pour calculer ce reste, on effectue une «exponentiation binaire». On commence par écrire 58 483 comme une somme de puissances de 2 : $58\,482 = 2^{15}+2^{14}+2^{13}+2^{10}+2^6+2^5+2^4+2$. Puis on utilise cette décomposition pour écrire le nombre $2^{58\,482}$ comme un produit de puissances de 2 : $2^{2^1}\times 2^{2^4}\times 2^{2^5}\times 2^{2^6}\times 2^{2^{10}}\times 2^{2^{13}}\times 2^{2^{14}}\times 2^{2^{15}}$. Ensuite on calcule successivement les restes des divisions par 58 483 de 2^{2^0}, 2^{2^1}, 2^{2^2}, 2^{2^3}... jusqu'à $2^{2^{15}}$. Le reste de la division de $2^{2^{i+1}}$ par 58 483, pour n'importe quel nombre entier i, s'obtient simplement à partir du reste précédent (celui de la division de 2^{2^i} par 58 483) : on l'élève au carré, on divise ce carré par 58 483, et on prend le reste de cette division *(voir la figure 3)*. Ainsi, de proche en proche, on détermine le reste par 58 483 de la division pour chacune des puissances qui interviennent dans le développement binaire de 58 482. On multiplie enfin tous ces restes, mais en ne gardant, après chaque multiplication, que le reste de la division par 58 483. De cette façon, les résultats intermédiaires restent de taille raisonnable. On détermine ainsi que le reste de la division de $2^{58\,482}$ par 58 483 est égal à celui du produit $4\times 7\,053\times 34\,259\times 42\,237\times 34\,763\times 53\,664\times 5010\times 10\,893$: il est égal à 11 669. Pour simplifier l'expression des résultats, les mathématiciens écrivent : $2^{58\,482}\equiv 11\,669 \bmod 58\,483$. Le signe ≡ se lit «congru à» ; il indique que les expressions à sa gauche et à sa droite ont le même reste dans la division par le nombre indiqué après l'abréviation mod, laquelle se lit «modulo».

Le système RSA

Le protocole de cryptographie RSA se fonde sur le résultat suivant :

Soit p et q deux nombres premiers. On pose $n = pq$. Si e est un nombre entier premier avec $(p-1)(q-1)$, alors il existe un nombre entier d positif, tel que, pour tout A, $A^{ed} - A$ soit divisible par n.

(X et Y sont dits premiers entre eux si leur seul diviseur commun est 1).

Le protocole est le suivant :

Alice choisit p, q, e (p et q premiers, et e premier avec $(p-1)(q-1)$.

Alice calcule $n = pq$ et d, ce qui est facile.

Alice rend publics n et e (par exemple en les publiant dans un annuaire).

Robert, qui veut communiquer une information secrète à Alice, transforme son information en un nombre entier A inférieur à n (ou en plusieurs nombres si nécessaire) avec un codage connu de tous.

Robert calcule le reste B de la division de A^e par n, envoie B à Alice par un canal qui n'a pas besoin d'être protégé (par exemple en publiant B dans un journal).

Alice, pour décoder B, va calculer B^d ; comme les restes de la division de B^d et de A^{ed} par n sont égaux à A, elle connaîtra A et donc le message de Robert.

Sécurité du système RSA

La sécurité de ce codage à clé publique repose sur la difficulté du calcul des facteurs de n : quiconque les connaît peut facilement calculer d et déchiffrer les messages envoyés par Robert à Alice.

L'application du test de Fermat, finalement, indique que le nombre 58 483 n'est pas premier, puisque le résultat n'est pas 1. Ce procédé est relativement rapide, mais il ne fournit pas la décomposition en facteurs premiers de 58 483.

La méthode des divisions successives

Quand on sait qu'un nombre n n'est pas premier, comment peut-on le factoriser? On cherche une décomposition en diviseurs propres, c'est-à-dire qui ne sont égaux ni à 1 ni à n (les nombres 1 et n sont nommés diviseurs triviaux de n) ; par exemple, 3 est un diviseur propre de 12. Puis on examine si ces diviseurs sont des nombres premiers. S'ils ne le sont pas, on les décompose à leur tour, jusqu'à ce que tous les diviseurs soient premiers.

Par exemple, $12 = 3 \times 4$. Le nombre 3 est premier, mais 4 est composé et peut être factorisé : $4 = 2 \times 2$. La décomposition complète de 12 s'écrit donc $12 = 2 \times 2 \times 3$. Ainsi, dans un algorithme de factorisation, la recherche d'un facteur propre d'un nombre composé est une opération fondamentale.

La méthode la plus simple est celle des divisions successives : on divise n par tous les nombres premiers 2, 3, 5, 7, 11, 13, 17, etc. (préalablement stockés dans une table), jusqu'à ce qu'une division tombe juste. Dans le cas de n = 58 483, on observe que le 51e nombre premier, 233, est un diviseur ($58\,483 = 233 \times 251$). On doit donc effectuer 51 divisions avec reste pour trouver ce facteur, ce qui reste possible, mais quand le plus petit facteur premier p de n devient grand, les divisions nécessaires sont de plus en plus nombreuses *(voir la figure 4)*, et la table des nombres premiers nécessaires aux calculs est de plus en plus grande. La factorisation par divisions successives ne convient que pour la recherche de petits facteurs ; pour des facteurs plus grands, une méthode fondamentalement différente s'impose.

Les courbes elliptiques

En 1985, Hendrik Lenstra a trouvé un algorithme de factorisation qui utilise des courbes elliptiques. Ces objets mathématiques, qui ont notamment été employés par le mathématicien britannique Andrew Wiles dans sa démonstration de la conjecture de Fermat, sont également utilisés dans des protocoles cryptographiques.

Dans la méthode des courbes elliptiques comme dans le test de Fermat, on effectue un calcul qui réussit si n est un nombre premier et qui échoue quand n est composé (ce que l'on sait, par exemple, grâce à un test de Fermat préliminaire) : au cours des calculs, on doit à plusieurs reprises extraire le plus grand diviseur commun (pgcd) de n et d'autres nombres, et le calcul ne se poursuit que si le pgcd est égal à 1, c'est-à-dire si les deux nombres sont premiers entre eux. Dans le cas contraire, ce pgcd est un diviseur de n, et même un diviseur propre si on a de la chance : l'objectif est alors atteint.

Du hasard intervient dans cette méthode, car on a le choix entre un grand nombre de courbes elliptiques pour mener les calculs *(voir la figure 7)*. On en choisit une au hasard. Si elle ne fournit pas de diviseur, on essaie la sui-

i	0	1	2	3	4	5	6	7
RESTE	2	4	16	256	7 063	34 259	42 237	57 220
i	8	9	10	11	12	13	14	15
RESTE	16 128	38 483	34 763	31 340	44 631	53 664	5 010	10 836

3. LES RESTES DES DIVISIONS de 2^{2^i} par 58 483 pour i compris entre 0 et 15. Pour trouver ces valeurs, il n'est pas nécessaire de calculer les nombres 2^{2^i} (qui peuvent être très grands) : on utilise les restes précédemment calculés.

S	10^3	10^6	10^9	10^{12}	10^{15}	10^{18}
NOMBRES PREMIERS INFÉRIEURS OU ÉGAUX À S	167	78 497	$1,8 \times 10^7$	$3,6 \times 10^{10}$	$2,8 \times 10^{13}$	$2,4 \times 10^{16}$

4. LE NOMBRE de nombres premiers inférieurs à 10^3, 10^6, etc. est aussi le nombre de divisions qui sont nécessaires, dans la méthode des divisions successives, pour obtenir un facteur premier de chaque taille respective.

Congruences et calcul du plus grand diviseur commun (pgcd)

Les mathématiciens ont étendu les opérations courantes aux restes de divisions des nombres entiers. Par exemple, quand on divise 7 par 3, le quotient est 2 et le reste 1, car $7 = 2 \times 3 + 1$. Pour les nombres négatifs, le calcul est analogue. Le quotient de –7 par 3 est –3 et le reste est 2, car $3 \times (-3) + 2 = -7$. Lors de la division de a par n, le reste est un des nombres $0, 1, 2, \ldots, n-1$. On le note $R(a,n)$.

Le plus grand diviseur commun (pgcd) de deux nombres entiers naturels a_1 et a_2 est obtenu par l'algorithme d'Euclide. On calcule la suite $a_3 = R(a_1, a_2)$, $a_4 = R(a_2, a_3)$, etc. Après un certain nombre de divisions, le reste est 0. Le pgcd de a_1 et a_2 est le dernier reste non nul. Par exemple, le pgcd de $a_1 = 631$ et de $a_2 = 405$ est obtenu de la façon suivante :

631/405 = 1 reste 226
405/226 = 1 reste 179
226/179 = 1 reste 47
179/47 = 3 reste 38
47/38 = 1 reste 9
38/9 = 4 reste 2
9/2 = 4 reste 1
2/1 = 2 reste 0.

Le dernier reste non nul est 1, donc le pgcd de 405 et de 631 est 1. Les deux nombres sont premiers entre eux. Lorsque deux nombres entiers a et b donnent le même reste dans la division par un entier naturel n, on dit qu'ils sont congrus modulo n et on écrit $a \equiv b \bmod n$. Par exemple, $10 \equiv 4 \bmod 3$, car 10 et 4 donnent le reste 1quand on les divise par 3. On démontre facilement que $a \equiv b \bmod n$ si n est un diviseur de $b - a$: le nombre $10 - 4 = 6$ est divisible par 3.

Quand on calcule le reste, pour une expression compliquée qui inclut des additions, des soustractions ou des produits, on peut remplacer tous les résultats intermédiaires par leur reste. Par exemple, si l'on veut calculer le reste de $14 \times (12 + 7)^8$ dans la division par 8, on remplace 14 par 6 (le reste de la division de 14 par 8) et 12 par 4, et l'on a : $14 \times (12 + 7)^8 \equiv 6 \times (4 + 7)^8 \bmod 8$. Ensuite, on obtient $6 \times (4 + 7)^8 \equiv 6 \times 3^8 \equiv 6 \times (3^2)^4 \equiv 6 \times (1^4) \equiv 6 \bmod 8$. On aurait aussi pu calculer $14 \times (12 + 7)^8 = 237\,769\,882\,574$, puis diviser le résultat par 8 et prendre le reste. Avec la première méthode, les résultats intermédiaires ne dépassent jamais 7, ce qui rend les calculs beaucoup plus rapides.

vante. H. Lenstra a démontré que, pour tout nombre composé n, il existe des courbes elliptiques qui fournissent un diviseur.

Comme pour la méthode par divisions successives, le temps nécessaire pour trouver un diviseur par la méthode des courbes elliptiques dépend de la taille de ce diviseur. Cependant, cette méthode identifie en un temps raisonnable les facteurs ayant jusqu'à 30 chiffres *(voir la figure 5)*.

Qu'est-ce qu'une courbe elliptique? C'est un ensemble de points d'un plan, dont les coordonnées x et y vérifient l'équation $y^2 = x^3 + ax + b$. Les paramètres a et b sont des nombres entiers choisis de telle sorte que $4a^3 + 27b^2$ soit différent de 0 *(voir la figure 6)*. Malgré leur nom, les courbes elliptiques n'ont qu'un rapport très éloigné avec les ellipses. Elles sont parfois composées de plusieurs branches.

Les courbes elliptiques sont des ensembles intéressants, car on peut définir une opération d'addition de leurs points : à deux points P_1 et P_2 d'une courbe elliptique, une construction géométrique simple associe un troisième point de la courbe, qui est nommé $P_1 + P_2$. L'opération ainsi définie a des propriétés analogues à l'addition des nombres entiers : elle est associative (le point $(P + Q) + R$ est confondu avec $P + (Q + R)$), commutative (le point $P + Q$ est le même que le point $Q + P$), elle admet un élément neutre (c'est-à-dire un point qui, ajouté à un point quelconque, ne change pas ce dernier) et chaque point a un opposé (ou inverse, tel que la somme du point et de son opposé soit égale à l'élément neutre).

L'inverse $-P$ d'un point P de la courbe est le symétrique de P par rapport à l'axe des abscisses. En général, on trouve la somme $P_1 + P_2$ de deux points P_1 et P_2 en traçant la droite qui passe par ces deux points : elle coupe la courbe en un troisième point Q, dont le symétrique $-Q$ par rapport à l'axe des abscisses est la somme cherchée, $P_1 + P_2$ *(voir la figure 6)*.

Quand P_1 et P_2 sont confondus, on obtient le point $P_1 + P_2$, ou $2P_1$, en remplaçant la droite P_1P_2 par la tangente à la courbe elliptique au point P_1. Enfin, quand P_1 est égal à $-P_2$, la droite passant par P_1 et P_2 ne coupe pas la courbe en un autre point ; on imagine alors que la droite coupe la courbe à l'infini. Pour cette raison, on ajoute aux points de la courbe un point que l'on se représente à l'infini et que l'on nomme O ; on pose $P_1 - P_1 = O$. Le point O est le zéro (l'élément neutre) de l'addition des courbes elliptiques.

L'addition de deux points peut aussi se définir algébriquement. Si P_1 et P_2 sont deux points de la courbe elliptique de coordonnées (x_1, y_1) et (x_2, y_2), alors : $-P_1 = (x_1, -y_1)$, $P_1 + P_2 = O$, si $P_1 = -P_2$; $P_1 + P_2 = P_2$, si $P_1 = O$; $P_1 + P_2 = P_1$, si $P_2 = O$; dans les autres cas, on détermine la somme $P_1 + P_2$ en calculant d'abord un nombre λ par les formules : $\lambda = (y_1 - y_2) / (x_1 - x_2)$, si P_1 est différent de P_2, et $\lambda = (3x_1^2 + a) / 2y_1$, si P_1 est égal à P_2. On détermine finalement les coordonnées (x_3, y_3) du point $P_3 = P_1 + P_2$ par : $x_3 = -x_1 - x_2 + \lambda^2$, et $y_3 = y_1 + \lambda(x_1 - x_3)$.

Notons que, lorsque P_1 est égal à $-P_2$, on pourrait étendre, d'une certaine façon, le calcul de λ : en toute rigueur, il est impossible, car la division par zéro n'est pas définie, mais on peut aussi admettre que l'élément neutre O possède des coordonnées infinies.

On peut utiliser ces formules sans chercher à se représenter x et y comme les coordonnées d'un point du plan. Cette abstraction devient utile quand les coordonnées des points considérés ne sont pas des nombres réels : tant qu'on peut additionner, soustraire, multiplier ou diviser ces coordonnées, on conserve les propriétés des courbes elliptiques. Nous avons notamment vu, à propos des congruences, que l'on peut additionner, soustraire et multiplier des restes de divisions. Plus précisément, le reste de la division d'un nombre entier par un nombre naturel n peut prendre les valeurs 0, 1, 2, ..., $n-1$. On calcule avec ces restes comme on le fait avec des nombres entiers, mais, quand le résultat d'un calcul sort de l'intervalle compris entre 0 et $n-1$, on le divise par n et on conserve le reste de cette division. Par exemple, dans l'ensemble des nombres modulo 12, $10 + 5 = 3$; $1 - 2 = 11$ (on utilise couramment ce calcul : 5 heures après 10 heures, il est 3 heures, et deux heures avant 1 heure, il est 11 heures). On multiplie de même : par exemple, $4 \times 5 = 8$.

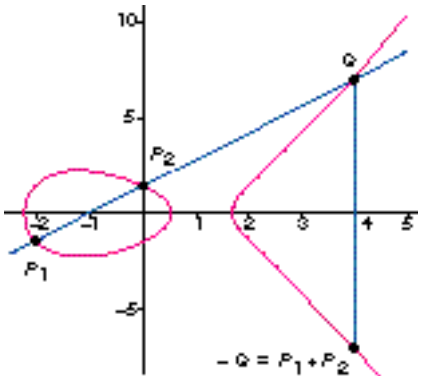
La division pose des problèmes particuliers. Une division est analogue à une multiplication par l'inverse, et l'inverse de x est l'élément y qui vérifie $xy = 1$. Par exemple, pour $n = 12$, l'inverse de 5 est 5, car $5 \times 5 = 25$ et $25 \equiv 1 \mod 12$. En revanche, 4 n'a pas d'inverse car $4y$ a plusieurs solutions, 3, 6 ou 9, modulo 12. On ne divise simplement que si n est un nombre premier. On calcule l'inverse grâce à la série d'opérations de l'algorithme classique de la division.

Avec cet ensemble de restes muni de ces opérations de bases, on définit également une courbe elliptique. Le cas le plus simple est celui où le module (le nombre par rapport auquel on prend le reste) est un nombre premier p. Les paramètres a et b de la courbe doivent être des nombres compris entre 0 et $p - 1$ tels que $4a^3 + 27b^2$ ne soit pas congru à 0 modulo p. La courbe est constituée du point O et des points associés aux couples (x, y) de nombres compris entre 0 et $p - 1$ qui vérifient la congruence $y^2 \equiv x^3 + ax + b \mod p$. Les formules pour l'addition restent inchangées, mais on remplace tous les nombres qui se présentent par le reste de leur division par p.

Par exemple, pour $p = 5$, $a = 1$ et $b = -1$ (cette écriture est une simplification de $p = 5$; $a \equiv 1 \mod 5$; $b \equiv 4 \mod 5$), les points P_1 et P_2 de coordonnées respectives $P_1 = (1, 1)$ et $P_2 = (2, 2)$ sont des points de la courbe elliptique. Pour

NOMBRE DE CHIFFRES	TEMPS DE CALCUL
6	6 secondes
9	20 secondes
12	2,5 minutes
15	16 minutes
18	95 minutes
21	7,2 heures
24	39,5 heures
27	140 heures
30	574 heures

5. TEMPS DE FACTORISATION, par la méthode des courbes elliptiques exécutée, sur un ordinateur puissant : à gauche, on lit le nombre de chiffres des nombres à factoriser et, à droite, le temps moyen nécessaire pour trouver un facteur ayant ce nombre de chiffres.

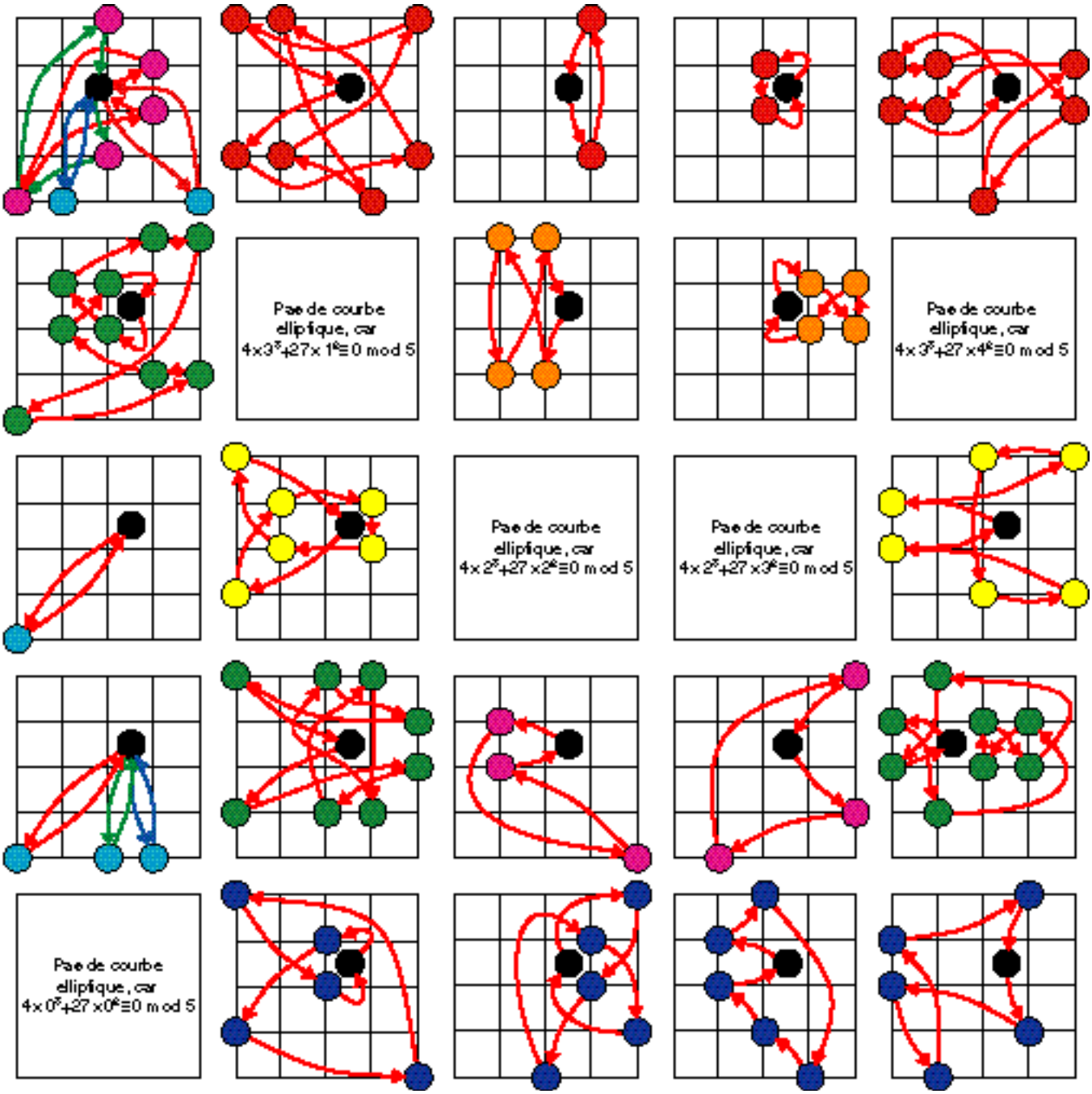


6. LA COURBE ELLIPTIQUE *(en rouge)* pour les paramètres $a = -4$ et $b = 2$. Les points P_1 et P2, de coordonnées respectives $(-2, -\sqrt{2})$ et $(0, \sqrt{2})$ sont sur la courbe elliptique, car ils vérifient son équation : $y^2 = x^3 + ax + b$.

calculer les coordonnées (x_3, y_3) du point $P_3 = P_1 + P_2$, on applique les formules indiquées précédemment, mais avec les règles de la congruence, et l'on obtient $x_3 = 3$ et $y_3 = 2$. Effectivement, (3, 2) est un point de la courbe.

L'ensemble *E(p)* des points des courbe elliptique modulo *p* n'a qu'un nombre fini d'éléments, compris entre $p - 2\sqrt{p} + 1$ et $p + 2\sqrt{p} + 1$. En essayant tous les couples de nombres possibles, on trouve que la courbe elliptique *E*(5), pour $a = 1$ et $b = 4$, est constituée des neuf éléments (3,3), (3,2), (0,3), (0,2), (1,4), (1,1), (2,3), (2,2) et O *(voir la figure 7)*.

La méthode de factorisation par les courbes elliptiques repose sur une observation du mathématicien français Joseph Louis Lagrange (1736-1813) : quand on additionne un élément d'un groupe fini à lui-même autant de fois qu'il y a d'éléments dans le groupe, on obtient l'élément neutre.



7. LES COURBES ELLIPTIQUES *E(p)*, pour $p = 5$. Chaque case correspond à un couple de paramètres *(a,b)* ; la coordonnée *a* croît de bas en haut, et *b* de gauche à droite, chacune variant de 0 à 4. Dans chaque case, les coordonnées croissent également de 0 à 4 : chaque élément du groupe est représenté par un disque de couleur à la position correspondante. L'élément neutre O *(cercle noir)* est indiqué au milieu de la case, bien que ses coordonnées soient infinies. Par exemple, le groupe pour $a = 0$ et $b = 1$ est représenté par la deuxième case en partant de la gauche dans la ligne inférieure : il contient les éléments (0,1), (0,4), (2,2), (2,3), (4,0) et O. Dans chaque groupe, une flèche rouge relie O à un autre élément *x*. Les multiples de celui-ci (*x*, *x* + *x*, *x* + *x* + *x*, etc.) sont reliés par une chaîne de flèches rouges qui finit de nouveau sur O : c'est ce que l'on nomme un cycle. Certains groupes contiennent plus d'un cycle (marqués par des couleurs différentes). La couleur de chaque cercle désigne le nombre d'éléments du cycle auquel il appartient. Sur cet exemple très simple, on voit que, pour un couple *(a,b)* choisi au hasard, la probabilité qu'il soit inutilisable (parce qu'il n'y a pas de courbe elliptique) ou improductif (parce qu'un groupe présentant la même structure à déjà été trouvé) est faible.

Dans l'exemple considéré, neuf fois n'importe quel élément de $E(5)$ est égal à O (on écrit comme d'habitude $2P$ pour $P + P$, $3P$ pour $P + P + P$, etc.).

Vérifions-le pour $P = (1, 1)$. On obtient successivement : $2P = (2, 2)$, $4P = 2 \times (2P) = (0, 2)$, $8P = 2 \times (4P) = (1, -1)$ et $9P = 8P + P = (1, -1) + (1, 1) = O$. Cet exemple montre qu'au lieu d'additionner neuf fois P, on peut employer la même astuce que pour l'exponentiation binaire : on calcule la suite $2P$, $4P$,..., 2^mP et, pour le résultat final, on additionne les puissances de deux dont on a besoin. Ainsi le temps de calcul reste raisonnable, même pour un très grand nombre à la place de 9.

Comment employer les courbes elliptiques pour trouver des diviseurs d'un nombre composé? Cherchons, par exemple, un diviseur de $n = 35$. On choisit une courbe elliptique modulo 35, c'est-à-dire deux nombres entiers a et b tels que $4a^3 + 27b^2$ ne soit pas divisible par 35 : retenons, par exemple, $a = 1$ et $b = -1$. On obtient bien une courbe elliptique, mais on ne peut pas toujours additionner ses points : dans les formules intervient un inverse modulo n, qui n'existe pas toujours, car 35 est composé. On essaie malgré tout : le calcul échoue, mais la manière dont il échoue donne l'information cherchée.

Plus précisément, on choisit un nombre entier naturel k (nous verrons comment par la suite) Pour notre exemple, prenons $k = 9$. On essaie d'additionner k fois le point $P = (2, 2)$ avec lui-même, par la méthode que nous avons considérée pour $E(5)$. On obtient $2P = (0, 22)$, $4P = (16, 19)$, $8P = (7, 13)$. Ensuite, quand on veut calculer $9P = 8P + P = (7, 13) + (2, 2)$, on doit tout d'abord évaluer λ et, pour cela, diviser par $7 - 2 = 5$ modulo 35. Cette division est impossible, car 5 n'a pas d'inverse modulo 35. On a donc finalement obtenu ce que l'on cherchait : 5 est un diviseur propre de 35! Notons que, dans le cas général, le nombre dont on ne trouve pas d'inverse n'est pas nécessairement un diviseur de n, c'est le pgcd de n et de ce nombre qui est un diviseur.

Les calculs échouent pour $k = 9$, parce que 5 est un diviseur de 35, et que 9 est le nombre d'éléments de $E(5)$. Les calculs auraient aussi échoué pour tout multiple de 9. Pour comprendre pourquoi, examinons la relation entre $E(35)$ et $E(5)$: comme on peut le vérifier sur les relations qui définissent ces courbes elliptiques, on peut «projeter» tout élément P_1 de $E(35)$ sur un élément P_2 de $E(5)$, en réduisant ses coordonnées modulo 5 ; l'élément associé à $9P_1$ est alors confondu avec $9P_2$. Faisons maintenant un raisonnement par l'absurde : si on pouvait calculer $9P_1$ dans $E(35)$, ce qui impliquerait de pouvoir calculer un λ fini, alors les coordonnées de $9P_1$ seraient finies ; en les réduisant modulo 5, on obtiendrait aussi des coordonnées finies pour $9P_2$: c'est impossible, puisque, d'après l'observation de Lagrange, $9P_2$ est l'élément neutre du groupe fini à 9 éléments $E(5)$, et que ses coordonnées sont infinies.

Comment choisir le nombre k? Si on cherche un facteur premier d'environ six chiffres, on choisit le nombre k de façon qu'il soit multiple du plus grand nombre possible de nombres de six chiffres qui n'ont pas de facteurs premiers trop gros. Alors, on a de bonnes chances pour que k soit un multiple du nombre d'éléments (nommé cardinal) de $E(p)$, que l'on ne connaît pas (p est le diviseur premier cherché). En effet, nous avons vu que le cardinal de $E(p)$ est compris entre $p - 2\sqrt{p} + 1$ et $p + 2\sqrt{p} + 1$: ce nombre est donc au plus du même ordre de grandeur que p. D'autre part, le plus souvent, il n'est pas lui-même un nombre premier, et est donc décomposable en facteurs premiers plus petits, qui, si on a de la chance, sont aussi des facteurs premiers de k. Lorsque l'on essaie de calculer k fois un élément de $E(n)$, on teste la divisibilité de n par tous ces nombres premiers d'un seul coup.

Pour trouver un nombre k approprié, on choisit un nombre B et on calcule k comme le produit de nombres premiers inférieurs à B. Par exemple, pour chercher des facteurs à au plus six chiffres, on choisit $B = 147$. On prend alors $k = 2^7 \times 3^4 \times 5^3 \times 7^2 \times 11^2 \times 13 \times 17 \times 19 \times ... \times 139$.

En général, la méthode des courbes elliptiques ne donne pas de résultat avec la première courbe elliptique choisie. On essaye plusieurs courbes, c'est-à-dire plusieurs paires de paramètres a et b. Plus B est grand, plus grand est le nombre de courbes que l'on doit essayer *(voir la figure 8)*. Ensuite, si la méthode n'a toujours pas trouvé de diviseur, on est à peu près sûr que le nombre n n'a pas de diviseur de l'ordre de grandeur que l'on s'est fixé. En effet,

NOMBRE DE CHIFFRES	6	9	12	15	18	21	24	27	30
B	147	682	2 462	8 318	23 462	63 502	162 730	395 806	945 922
NOMBRE DE COURBES	10	24	12	111	231	445	883	1 501	2 594

8. LA MÉTHODE DES COURBES ELLIPTIQUES trouve un diviseur d'un nombre n à peu près indépendamment de la taille de n. La longueur de la recherche croît avec l'ordre de grandeur du diviseur cherché. Ce tableau indique, dans la première ligne, cet ordre de grandeur en nombre de chiffres ; dans la deuxième ligne, on a indiqué le nombre B qui permet de déterminer le nombre k ; dans la troisième ligne apparaît le nombre de courbes elliptiques que l'on devra tester pour trouver un diviseur.

NOMBRE DE CHIFFRES DE n	50	60	70	80	90	100	110	120
NOMBRE D'ÉQUATIONS	3 000	4 000	7 400	15 000	30 000	51 000	120 000	245 000

9. POUR DÉTERMINER de très grands facteurs d'un nombre n, on emploie le crible quadratique. Lors de l'exécution de l'algorithme, on doit construire un système d'équations linéaires qui, selon la taille de n, peut contenir un nombre considérable d'équations.

u	−3	−2	−1	0	1	2	3
$(m+u)^2 - n$	−540	−373	−204	−33	140	315	492
CRIBLE PAR 2	−135		−51		35		123
CRIBLE PAR 3	−5		−17	−11		35	41
CRIBLE PAR 5	−1				7	7	
CRIBLE PAR 7					1	1	

10. LORS DU CRIBLE QUADRATIQUE, on choisit des nombres carrés $(m - u)^2$ proches du nombre n à décomposer (ici, $n = 7429$). Le nombre u (première ligne) numérote ces nombres carrés. On calcule la différence de chacun avec n (deuxième ligne ; $m = 86$) et on divise ces valeurs par le nombre qui constitue la base de chaque crible. Par exemple, pour $u = 1$, on divise $(m + u)^2 - n$, soit 140, par 2, puis encore par 2, puis encore par 2, pour obtenir 35 : le crible par 2 conduit à cette valeur.

lorsque la courbe $E(p)$ varie (c'est-à-dire lorsqu'on essaie différentes valeurs de a et b), les éléments de $E(p)$ sont assez régulièrement répartis entre $p - 2\sqrt{p} + 1$ et $p + 2\sqrt{p} + 1$: on a donc de bonnes chances (sans en être complètement sûr) que k soit un multiple du cardinal de certaines de ces courbes, ce qui permettra de découvrir le facteur p.

Le temps de calcul de cette méthode dépend donc principalement de la taille du diviseur premier recherché, et à peine de celle du nombre à factoriser. Si un nombre de 1 000 chiffres a un diviseur de 20 chiffres, on trouve celui-ci assez facilement. En revanche, le temps de calcul croît très vite avec la taille du plus petit facteur. La méthode des courbes elliptiques convient bien pour la recherche de facteurs qui ont jusqu'à 30 chiffres. À ce jour, le plus grand facteur premier trouvé par cette méthode, à 47 chiffres, fut découvert par Peter Montgomery, au Centre de mathématiques et d'informatique d'Amsterdam.

Le crible quadratique

Pour la recherche de diviseurs encore plus grands d'un nombre n, on emploie une autre méthode : l'idée est de trouver des nombres entiers naturels X et Y tels que n soit un diviseur de $X^2 - Y^2$.

En effet, si n divise $X^2 - Y^2$, il divise aussi $(X - Y)(X + Y)$. Si n n'est pas un diviseur de $X - Y$ ou de $X + Y$, alors un diviseur propre de n doit diviser $X - Y$, et un autre diviseur propre de n doit diviser $X + Y$. Ainsi le plus grand diviseur commun de n et de $X - Y$, par exemple, est plus grand que 1, et donc est l'un des diviseurs cherchés. Choisissons par exemple $n = 7\,429$, $X = 227$ et $Y = 210$. Le nombre 7 429 est un diviseur de $X^2 - Y^2$, puisque $X^2 - Y^2 = 7\,429$, mais il n'est pas un diviseur de $X - Y = 17$ ni un diviseur de $X + Y = 437$. Le pgcd de 17 et de 7 429 est 17, et c'est un diviseur propre de 7 429.

Comment trouve-t-on X et Y? On établit d'abord un système d'équations linéaires ; puis on détermine X et Y à partir des solutions de ce système. Le nombre d'équations dépend de la taille du nombre à factoriser ; pour un nombre de 120 chiffres il faut à peu près 245 000 équations pour autant d'inconnues *(voir la figure 9)*. Contrairement à la méthode des courbes elliptiques, c'est la taille du nombre à factoriser, et non la taille du facteur cherché, qui détermine le temps de calcul.

NOMBRE DE CHIFFRES DE n	50	60	70	80	90	100	110	120
TAILLE DE LA BASE DES FACTEURS PREMIERS (EN MILLIERS)	3	4	7,4	15	30	51	120	245
TAILLE DE L'INTERVALLE DE CRIBLE (EN MILLIONS)	0,2	2	5	6	8	14	16	26

11. POUR UNE VERSION améliorée du crible quadratique, dont les détails ne peuvent pas être donnés ici, on a besoin, selon la taille du nombre n à factoriser, de différentes tailles pour l'intervalle de crible (l'intervalle où l'on cherche des carrés convenables) et pour la base de facteurs (l'ensemble des facteurs premiers par lesquels doivent être divisibles les carrés réduits).

En premier lieu, on se donne une suite de carrés ayant deux propriétés : ils sont proches de n et leur différence avec n est, au signe près, un produit de petits nombres premiers. Pour $n = 7429$, les carrés 83^2, 87^2 et 88^2 conviennent, car les différences sont des produits des nombres premiers 2, 3, 5 et 7 :

$83^2 - 7\,429 = -540 = (-1) \times 2^2 \times 3^3 \times 5$

$87^2 - 7\,429 = 140 = 2^2 \times 5 \times 7$

$88^2 - 7\,429 = 315 = 3^2 \times 5 \times 7$.

Quand on multiplie certaines de ces lignes les unes par les autres, les exposants des décompositions s'additionnent. Si la somme des exposants est paire pour chaque facteur premier, alors le produit est un carré. Par exemple, $(87^2 - 7\,429) \times (88^2 - 7\,429) = 2^2 \times 3^2 \times 5^2 \times 7^2 = (2 \times 3 \times 5 \times 7)^2 = 210^2$.

Dès lors, pour trouver X et Y, on n'a plus qu'à appliquer une fois de plus les règles des calculs de congruences. On écrit : $(87 \times 88)^2 \equiv (87^2 - 7\,429) \times (88^2 - 7\,429) \bmod 7\,429$. Donc $(87 \times 88)^2 \equiv 210^2 \bmod 7\,429$. Dans cette équation, on peut encore remplacer 87×88 par son reste dans la division par 7 429, et on obtient $227^2 \equiv 210^2 \bmod 7\,429$, ce qui signifie que 7 429 est un diviseur de $227^2 - 210^2$. On peut donc prendre $X = 227$ et $Y = 210$. Ces valeurs pour X et Y fournissent le diviseur 17 de 7 429, comme on l'a vu précédemment.

Parmi les relations dont on dispose (elles sont très nombreuses, en général), on doit extraire celles qui se combinent en un carré de la façon décrite. À cette fin, on construit un système d'équations linéaires : à chaque relation, on associe une inconnue, qui prend la valeur 1 si la relation est employée dans la construction du carré, et 0 dans le cas contraire. Dans notre exemple, le système a trois inconnues, λ_1, λ_2 et λ_3. Le produit de ces relations peut être écrit ainsi:

$((-1) \times 2^2 \times 3^3 \times 5)^{\lambda_1} \times (2^2 \times 5 \times 7)^{\lambda_2} \times (3^2 \times 5 \times 7)^{\lambda_3}$,

ou bien, en utilisant les règles de calcul de puissances :

$(-1)^{\lambda_1} \times 2^{2\lambda_1 + 2\lambda_2} \times 3^{3\lambda_1 + 2\lambda_3} \times 5^{\lambda_1 + \lambda_2 + \lambda_3} \times 7^{\lambda_2 + \lambda_3}$.

Comme ce produit doit être un carré, tous les exposants doivent être pairs :

$$\lambda_1 \equiv 0 \bmod 2$$
$$\lambda_1 + \lambda_2 + \lambda_3 \equiv 0 \bmod 2$$
$$\lambda_2 + \lambda_3 \equiv 0 \bmod 2.$$

La première équation concerne le facteur –1, la deuxième 5 et la troisième 7. Les exposants des facteurs 2 et 3 étant pairs dans tous les cas, ils ne nécessitent pas d'équation. On peut résoudre le système d'équations par une méthode classique d'algèbre linéaire, en prenant garde de calculer modulo 2. On obtient $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 1$, et, à partir de là, on trouve X et Y comme on l'a déjà vu.

Enfin, comment trouve-t-on les relations elles-mêmes? On utilise un procédé de crible, auquel l'algorithme du «crible quadratique» doit son nom : on cherche des carrés de nombres dont la différence avec n se décompose en petits facteurs premiers.

On fixe d'abord les nombres premiers qui peuvent intervenir dans les relations. Dans l'exemple considéré précédemment, on a pris 2, 3, 5 et 7. Pour traiter le signe, on ajoute le nombre –1. L'ensemble de ces nombres premiers forme ce que l'on appelle la base des facteurs.

Ensuite on calcule des carrés de nombres qui sont proches de n. Pour cela, on prend le plus grand nombre entier m inférieur à $\sqrt{n}$. Dans l'exemple, il s'agit de $m = 86$. Les carrés de nombres au voisinage de n sont alors $(m-3)^2 = 83^2$, $(m-2)^2 = 84^2$, $(m-1)^2 = 85^2$, $m^2 = 86^2$, $(m+1)^2 = 87^2$, $(m+2)^2 = 88^2$,..., et en général tous les $(m + u)^2$, où u est un nombre entier positif ou négatif petit par rapport à m.

Enfin on se fixe l'intervalle de crible, c'est-à-dire le domaine des nombres u avec lesquels on veut travailler. On

écrit les différences des carrés à n (nommés les carrés réduits) dans une liste *(voir la figure 10)*.

On détermine maintenant les carrés réduits dont tous les facteurs premiers sont dans la base de facteurs. On pourrait utiliser la méthode des divisions successives, mais un procédé de crible est plus rapide : pour trouver quels carrés réduits sont divisibles par un nombre premier p, on détermine tous les nombres u compris entre 0 et $p - 1$ pour lesquels $(m + u)^2 - n$ est divisible par p. À partir d'une valeur trouvée pour u, on trouve d'autres valeurs en ajoutant ou en soustrayant des multiples de p à cette valeur de u.

Dans l'exemple, le crible par 2 s'applique de la façon suivante : comme $(m + 1)^2 - n$ (égal à 140) est divisible par 2, il en est de même pour $(m - 1)^2 - n$, $(m - 3)^2 - n$ et $(m + 3)^2 - n$. On divise ces nombres par 2 jusqu'à ce que l'on obtienne un nombre impair (troisième ligne de la figure 10).

Le crible par 3 fonctionne de la même manière : on constate que $m^2 - n$ et $(m + 2)^2 - n$ sont divisibles par 3. En partant de $u = 0$ et $u = 2$, on se déplace par pas de longueur 3 vers la droite et vers la gauche, et on divise tant qu'on le peut par 3 les nombres ainsi obtenus. Par la même méthode, on crible par les autres nombres premiers de la base de facteurs. Partout où il y a un 1 à la fin de la liste, on peut décomposer le carré réduit sur la base de facteurs. On détermine la décomposition à l'aide de la méthode des divisions successives.

Ici, nous avons décrit la plus simple version du crible quadratique. La technique est plus vieille que l'ordinateur : l'officier français Eugène Olivier Carissan (1880-1925) construisit une machine mécanique avec laquelle on peut effectuer le crible pour plusieurs nombres premiers à la fois *(voir la figure 1)*.

Pour décomposer un nombre de 100 chiffres, on doit encore apporter de nombreuses améliorations. La base de facteurs et l'intervalle du crible deviennent gigantesques. Pour avoir une idée de leur ordre de grandeur, examinons les paramètres de la factorisation de RSA-120, un nombre de 120 chiffres que les Laboratoires *RSA* avaient donné à factoriser. La base de facteurs contenait 245 810 éléments, et le système d'équations à résoudre 245 810 inconnues et 252 222 équations. Le calcul aurait duré environ 50 ans sur une seule machine. Pour factoriser le nombre de 129 chiffres qui ouvre l'article, c'est aussi le crible quadratique qui fut employé.

Parallélisation

En dehors de l'amélioration de l'algorithme, l'emploi simultané d'un grand nombre d'ordinateurs accélère la factorisation. On peut utiliser soit des calculateurs parallèles (mais ils sont très chers), soit des systèmes partagés, par exemple des réseaux de machines qui sont très peu utilisées pendant la nuit ou le week-end.

À Sarrebrück, nous utilisons un réseau de 250 terminaux répartis sur le campus. Le système LiPS *(Library for Parallel Systems)*, que nous avons mis au point, reconnaît automatiquement quand une machine n'est pas sollicitée par son utilisateur principal, et il lance alors, par exemple, une factorisation. Quand l'utilisateur principal veut de nouveau travailler sur sa machine, LiPS arrête le programme de factorisation et le réveille de nouveau dès que la machine est libre. LiPS expédie enfin les résultats des calculs à l'ordinateur qui les centralise.

Même sur notre système réparti, le crible quadratique demande plusieurs semaines de calculs pour la factorisation d'un nombre de 130 chiffres. Le temps de calcul est doublé chaque fois que l'on ajoute trois chiffres, car les carrés réduits que l'on crible sont à peu près de la taille de $\sqrt{n}$.

Les algorithmes plus rapides, mais reposant sur le même principe que le crible quadratique doivent obtenir des relations grâce à la décomposition de nombres plus petits (de l'ordre non plus de $\sqrt{n}$, mais de n à la puissance 1/3 ou 1/4, par exemple). Le seul algorithme clairement meilleur, sur ce point, que le crible quadratique est le crible algébrique, qui a décomposé le nombre de Fermat F_9.

Avec le crible algébrique, un ensemble d'équipes dont nous faisons partie a décomposé un nombre RSA de 130 chiffres, en avril 1996. Les spécialistes s'attendent à ce qu'un crible algébrique amélioré puisse, dans un petit nombre d'années, factoriser un nombre quelconque de 160 chiffres. Ce serait un pas important, car les nombres composés qui sont employés dans la plupart des systèmes RSA ont moins de 160 chiffres : l'utilisation de tels systèmes RSA ne serait plus sûre.

La factorisation des nombres naturels est-elle un problème difficile? Pour le moment, oui : aucun algorithme au monde ne peut retrouver les facteurs d'un produit de deux nombres premiers de 150 chiffres. De tels produits constituent donc une bonne base pour la sécurité de protocoles cryptographiques, mais pour combien de temps?

Les mathématiciens, malgré des siècles de recherche, n'ont pas encore trouvé un algorithme de factorisation réellement rapide : est-ce suffisant pour affirmer que ce problème est difficile en soi? Non : grâce au développement des ordinateurs, des algorithmes ont déjà été découverts, et rien ne laisse supposer que les progrès considérables des 20 dernières années touchent déjà à leur fin.

Johannes BUCHMANN est professeur d'informatique à l'École d'enseignement supérieur technique de Darmstadt.

Martin HELLMAN, *Les mathématiques de la cryptographie à clef révélée*, in *Pour La Science*, octobre 1979.

Carl POMERANCE, *La recherche des nombres premiers*, in *Pour La Science*, février 1983.

D.M. BRESSOUD, *Factorizations and Primality Testing*, Springer, Heidelberg, 1989.

P. RIBENBOIM, *The Book of Prime Number Records*, Springer, Heidelberg, 1989.

H. COHEN, *A Course in Computational Algebraic Number Theory*, Springer, Heidelberg, 1993.

Arjen K. LENSTRA et Hendrik W. LENSTRA, *The Development of the Number Field Sieve*, Springer Lecture Notes in Mathematics, Band 1554, Springer, Heidelberg, 1993.

Hans RIESEL, *Prime Numbers and Computer Methods for Factorization*, 2, Auflage, Birkhäuser, Basel, 1994.

Yves HELLEGOUARCH, *Fermat enfin démontré*, in *Pour La Science*, février 1996.

François MORAIN, *La machine des frères Carissan*, in *Pour La Science*, janvier 1998.

Les bibliothèques de programmes LiDIA et LiPS sont disponibles à l'adresse Internet : / / www.informatik.th-darmstadt.de / TI / Forschung

Le site Internet des laboratoires RSA a pour adresse http: / / www.rsa.com / rsalabs /