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 /