Groupe d'étude en théorie analytique des nombres

Maurice Margenstern

Factorisation

Groupe d'étude en théorie analytique des nombres, tome 2 (1985-1986), exp. n° 11, p. 1-21

<http://www.numdam.org/item?id=TAN_1985-1986__2__A6_0>

Groupe d'étude en
THEORIE ANALYTIQUE DES NOMBRES
3e année, 1985/86, n° 11, 21 p.

27 janvier 1986

FACTORISATION

par Maurice MARGENSTERN (*)

1. Motivations cryptographiques.

Les développements de la cryptographie ont orienté les recherches vers l'exploitation de problèmes mathématiques, dont la résolution effective soit la plus complexe possible. Plusieurs méthodes reposent sur la complexité de la décomposition en facteurs premiers d'un entier quelconque. On va montrer, sur un exemple célèbre, l'importance de la factorisation d'un nombre.

1.1. La méthode R. S. A. - Le but de la méthode (cf. [11]) est d'offrir un système de codage et décodage de messages à clés publiques le plus simple et le plus sûr possible. Elle fonctionne comme suit.

Tout abonné au système définit trois nombres n, e et d dont les deux premiers, dits clés publiques figurent dans un annuaire de tous les utilisateurs, le troisième nombre restant secret. Ces trois nombres vérifient les propriétés suivantes :

(i) $n = pq$ où p et q sont deux "grands" nombres premiers distincts.

(ii) $ed \equiv 1 \mod \varphi(n)$.

Supposons qu'un abonné A désire envoyer à un autre abonné B un message M. A découpe d'abord en mots M : $M_1, \dots, M_k$ tels que $1 \leqslant M_i \leqslant n_B$ ce qui ramène l'envoi d'un message de longueur quelconque à l'envoi d'un nombre fini de messages de longueur bornée. On peut donc supposer $1 \leqslant M \leqslant n_B$. Alors A envoie à B le texte $C \equiv M^{e_B} \mod n_B$. B peut retrouver le message M en prenant $C^{d_B} \mod n_B$ car $C^{d_B} \equiv M^{e_B d_B} \equiv M \mod n_B$ puisque $e_B d_B \equiv 1 \mod \varphi(n_B)$.

En effet, il résulte du théorème de Fermat que $M^k \equiv M \mod n$ si $k = 1 \mod \varphi(n)$ et si $\operatorname{pgcd}(M, n) = 1$. Mais ici $n = pq$ où p et q sont des nombres premiers distincts.

Soit alors $M = p^a q^b M_1$ où $\operatorname{pgcd}(M_1, n) = 1$. Alors $p^{a\varphi(n)} \equiv 1 \mod q$ et $q^{b\varphi(n)} \equiv 1 \mod p$ d'où on tire que

$$(p^a q^b)^k \equiv p^a q^b \mod pq$$

si $\varphi(pq) \mid k$. Donc pour tout M avec $1 \leqslant M \leqslant n_B$ on a bien

$$M^{e_B d_B} \equiv M \mod n_B .$$

La méthode est simple car codage et décodage sont le résultat d'une même opération fonctionnant avec des arguments différents et aussi parce que le calcul de

(*) Maurice MARGENSTERN, Mathématiques, Université Paris-Sud, Bâriment 425, 91405 ORSAY CEDEX.

$(M)^e \mod n$ peut être exécuté très rapidement en écrivant e en base 2, on ramène le calcul de $(M)^e \mod n$ à un nombre d'élévations au carré modulo n de l'ordre de $\ln e$ (voir détails dans [11]).

La méthode est sûre car décoder C, c'est trouver d ce qui est aussi difficile que trouver p et q si ces derniers nombres sont convenablement choisis comme on le verra plus loin, et factoriser n est un problème difficile comme on pourra s'en convaincre à la lecture de cet exposé.

La méthode est également sûre car elle permet de garantir l'authenticité des messages. On modifie à cet effet la procédure de la façon suivante. A calcule $S \equiv (M)^{d_A} \mod n_A$ et applique à S la procédure précédente, c'est-à-dire envoie $C \equiv (S)^{e_B} \mod n_B$. B à son tour obtiendra d'abord S en décodant C grâce à sa clé secrète, puis il obtiendra M en décodant S grâce aux clés publiques de A qui se trouvent dans l'annuaire. Donc la méthode permet à B de s'assurer que A est bien l'expéditeur du message (par exemple A aura envoyé au préalable un message en clair annonçant qu'il est l'expéditeur) et elle garantit à A que B ne pourra pas falsifier le message reçu. Cette procédure est appelée procédure du message signé, S étant en quelque sorte la signature du message M puisque A étant seul (en principe) à connaître d_A, il est le seul à pouvoir produire S à partir de M comme (en principe) seul B peut obtenir S à partir de C.

1.2. Analyse de la méthode. - Si on connait p et q on trouve aisément d car $\varphi(pq) = (p-1)(q-1)$. On peut en fait trouver une "meilleure" valeur de d car si $ed \equiv 1 \mod \text{ppcm}(p-1, q-1)$ on a alors $M^{ed} \equiv M \mod n$ où $n = pq$, p, q nombres premiers distincts, pour tout M avec $1 \leqslant M \leqslant n$. (La démonstration résulte aisément de la démonstration de la propriété correspondante pour $\varphi(n)$.)

Posons $E(M) \equiv M^e \mod n$, et $D(M) \equiv M^d \mod n$. On a donc

$$E(D(M)) \equiv D(E(M)) \equiv M \mod n$$

pour tout M, $1 \leqslant M \leqslant n$.

Cependant, comme on va le voir, il existe des valeurs de M pour lesquelles $E(M) = M$. On les appelle points fixes de E. Plus généralement, on peut chercher les points fixes de E^k qui sont appelés points fixes de E d'ordre k.

Si M est un point fixe de E d'ordre k, si on suppose $M \not\equiv 0 \mod p$ alors

$$M^{e^k - 1} \equiv 1 \mod p .$$

Cette congruence possède dans $1, \ldots, p$, f solutions en M soient $m_1, \ldots, m_f$ où $f = \text{pgcd}(p-1, e^k - 1)$. De même pour $M \not\equiv 0 \mod q$, $M^e \equiv 1 \mod q$ possède dans $1, \ldots, q$, g solutions en M soient $n_1, \ldots, n_g$ où $g = \text{pgcd}(q-1, e^k-1)$ (cf. par exemple [12]). Par application du théorème chinois on a donc fg nombres $M \mod pq$ tels que $M \equiv m_i \mod p$ et $M \equiv n_j \mod q$ d'où fg solutions $\mod n$.

Du théorème chinois, on tire également l'existence d'une unique solution en $M \bmod pq$ du système $M \equiv 0 \bmod p$, $M \equiv nj \bmod q$. Alors

$$M^{e^k} \equiv 0 \bmod p \quad \text{et} \quad M^{e^k} \equiv n_j^{e^k} \equiv n_j \bmod q .$$

Donc

$$M^{e^k} \equiv M \bmod pq$$

par unicité des solutions dans $1 , \ldots , n$ du système ci-dessus. On en déduit que E^k possède un nombre de points fixes exactement égal à

$$(1 + \text{pgcd}(p - 1 , e^k - 1)(1 + \text{pgcd}(q - 1 , e^k - 1)) .$$

En faisant $k = 1$, comme p , q et e sont necessairement impairs, on a que E^k possède au moins 9 points fixes. Trouver des points fixes de E d'ordre quelconque, c'est trouver des facteurs de $p - 1$ ou $q - 1$ ce qui peut permettre de factoriser n .

Donc, pour rendre la méthode plus sûre, il faut non seulement choisir p et q grands mais aussi s'efforcer de réaliser les conditions $p = 2p' + 1$ et $q = 2q' + 1$ avec p' , q' premiers.

D'où l'intérêt de la connaissance des méthodes de factorisation d'un entier quelconque pour tenter de casser le système R. S. A.

2. Méthodes de factorisation.

On présente ici quatre méthodes de factorisations relevant de techniques sensiblement différentes.

2.1. La méthode $p - 1$ de Pollard. - Elle s'appuie sur un résultat théorique publié dans [9]. On donne ci-dessous la démonstration du résultat puis une procédure algorithmique qui s'en déduit.

2.1.1. Principe de la méthode.

THÉORÈME (J. M. POLLARD [9]). - Soit $0 < \alpha < 1$ et $\delta > 0$. Il existe un algorithme (machine de Turing à plusieurs rubans) indiquant pour tout entier n composé, si oui ou non n possède un facteur premier p tel que $p \leqslant n^{\alpha}$ et, dans le cas où un tel facteur existe, permettant de le calculer en un nombre d'opérations $O(n^{(\alpha/2)+\delta})$.

La démonstration nécessite deux lemmes techniques permettant d'estimer le nombre d'opérations de certains calculs.

LEMME 1. - Soient M et N deux nombres naturels donnés, $a_0 , \cdots , a_{N-1}$, $b_0 , \cdots , b_{N-1}$ des entiers donnés tels que $|a_i| , |b_i| \leqslant M$.

Alors il existe un algorithme calculant tous les

$$c_k = \sum_{i+j=k} a_i\, b_j \quad 0 \leqslant k \leqslant 2(N-1)$$

en un nombre d'opérations qui est $O(N^{1+\epsilon}\, M^{\epsilon})$ pour tout $\epsilon > 0$.

L'idée de la démonstration du lemme est la suivante. On considère les c_k comme les chiffres en base $D = M^2 N$ (ce qui évite les retenues puisque $|a_i|$, $|b_i| \leqslant M$) de nombres A et B dont les chiffres sont respectivement les a_i et les b_i dans cette même base. Or, si $C = \max(A , B)$ on sait multiplier A et B en $O((\ell n\ C)^{1+\epsilon})$ opérations (cf. par exemple [3] vol. 2). Or $\ell n\ C = O(N(\ell n\ N + \ell n\ M))$.

LEMME 2. - Soient a , n et M des naturels non nuls avec $a < n$ et $\text{pgcd}(a , n) = 1$. Alors il existe un algorithme permettant de dire s'il existe un entier m tel que $1 \leqslant m \leqslant M$ et $\text{pgcd}(a^m - 1 , n) > 1$, et au cas où un tel m existe, d'en donner le plus petit en $O(M^{1/2+\epsilon}\, n^{\epsilon})$ opérations, pour tout $\epsilon > 0$.

L'idée de la démonstration est la suivante. On peut toujours supposer que $M = L^2$ et $L = 2^k$. Alors pour m avec $1 \leqslant m \leqslant M$ on peut écrire $m = -u + vL$ où $v \leqslant u \leqslant L - 1$ et $1 \leqslant v \leqslant L$. Comme $\text{pgcd}(a , n) = 1$, alors $\text{pgcd}(a^m - 1 , n) > 1$ si, et seulement si, $\text{pgcd}(a^{vL} - a^u , n) > 1$, car si $d|n$ et $d|(a^{vL} - a^u)$, $d|(a^m - 1)$ puisque $\text{pgcd}(a , n) = 1$.

On donne un procédé de calcul rapide de $\prod_{u=0}^{L-1}(a^{vL} - a^u)$ pour $1 \leqslant v \leqslant L$. A cette fin, on calcule modulo n les coefficients de $f(x) = \prod_{u=0}^{L-1}(x - a^u)$ en regroupant les facteurs par deux dans une première étape, puis par deux les facteurs du second degré obtenus précédemment, etc. Il y a $\ell n\ L / \ell n\ 2$ étapes de calcul nécessitant chacune $O(L^{1+\epsilon}\, n^{\epsilon})$ opérations d'après le lemme 1, soit au total $O(L^{1+\epsilon}\, n^{\epsilon})$ opérations.

Soit alors $c_v = f(a^{vL}) \bmod n$. Connaissant les coefficients de f , on calcule $c_v\, r^{L+v^2}$ où $r = a^{L/2} \bmod n$ en obtenant ce nombre comme une somme donnée par le lemme 1, et d'après le lemme 1, on a tous ces nombres pour $1 \leqslant v \leqslant L$ en $O(L^{1+\epsilon}\, n^{\epsilon})$ opérations. Comme $\text{pgcd}(r , n) = 1$, on calcule les L $\text{pgcd}(c_v\, r^{L+v^2} , n)$ avec $1 \leqslant v \leqslant L$ en $O(Ln^{\epsilon})$ opérations. Ceci donne les solutions de $\text{pgcd}(f(a^{vL}) , n) > 1$. Le plus petit v , s'il en existe, et le plus grand u (pour v fixé) se calculent rapidement, d'où, si elle existe, la plus petite solution en m de $\text{pgcd}(a^m - 1 , n) > 1$ dans le temps indiqué.

C. Q. F. D.

Pour démontrer le théorème de J. M. POLLARD, il faut encore un lemme de nature arithmétique.

Soit p un nombre premier et k entier, $k \geqslant 2$. On pose

$$s_k(p) = \max_{q \geqslant k,\, q|p-1} \min\{b \ ;\ x^q \equiv b \pmod p \ \text{n'a pas de solution}\} .$$

et on définit $s_k(p) = 0$ si $x^q \equiv b \bmod p$ a des solutions pour tout $b \neq 0$ et tout $q \geq k$, $q | p - 1$.

On fixe g une racine primitive de p. Alors tout $x \in \mathbb{Z}/n\mathbb{Z}$, $n \neq 0$ s'écrit de façon unique $x = g^u$. On pose $u = \ell d\, x$. On sait alors que $x^q \equiv b \bmod p$ a des solutions si, et seulement si, $\operatorname{pgcd}(q , p - 1) | \ell d\, b$ (cf. [12] par exemple) et cette propriété ne dépend pas du choix de g. Donc

$$s_k(p) = \max_{q \geq k, q | p-1} \{b \neq 0 ; \ q \not| \ \ell d\, b\} .$$

On a le lemme suivant.

LEMME 3. - Pour tout $\delta > 0$, il existe $k \geq 2$ tel que

$$s_k(p) = O(p^\delta) .$$

En effet, la fonction $\Psi(x , y) = \sum_{n \leq x, p | n \Rightarrow p \leq y} 1$ vérifie $\Psi(x , x^\delta) \sim \rho(1/\delta)\, x$ quand $x \longrightarrow + \infty$, où ρ est la fonction de Dickman (cf. par exemple [1]).

Soit $g(p , q) = \min\{b ; \ q \not| \ \ell d\, b\}$ où $q | p - 1$. Supposons que pour tout p_i premier, $p_i < p^\delta$, on ait $q | \ell d\, p_i$. Alors $q | \ell d\, b$ pour tous les $b < p$ dont les facteurs premiers sont au plus égaux à p^δ. Il y aurait donc au moins βp nombres b de cette sorte avec $\beta > 0$. Or le nombre de $b < p$ tels que $q | \ell d\, b$ est au plus $(p - 1)/q \leq (p - 1)/k$. Donc pour k assez grand, on a une contradiction. Donc $g(p , q) \leq p^\delta$ pour tout $q \geq k$ et pour tout p assez grand, d'où le lemme 3.

Revenons à la démonstration du théorème de J. M. POLLARD. On définit un algorithme dont le travail procède comme suit :

(i) on fixe $M = [n^\alpha] + 1$,

(ii) par définition on teste si il existe $p \leq \sqrt{M}$ avec $p | n$. Si oui, on a terminé. Sinon, on passe à l'étape (iii).

(iii) $\delta > 0$ étant fixé, on fixe $k \geq 2$ et $C > 0$ tels que $s_k(p) < Cp^{\delta/2}$ pour tout nombre premier p et on pose $A = [Cn^{\delta/2}]$.

On observe que si k peut être déterminé de façon effective à l'aide de la fonction de Dickman, la détermination effective de C est beaucoup plus problématique.

Comme $\sqrt{M} \sim n^{\alpha/2}$ on peut supposer, pour n assez grand, que $A < \sqrt{M}$ (en prenant $\delta \ll \alpha$).

Pour $a = 2 , 3 , \ldots , A$ on cherche s'il existe une solution en n de

$$(*) \qquad \operatorname{pgcd}(a^n - 1 , n) > 1 \quad (1 \leq n \leq M) .$$

Si il existe un a sans solution, on a terminé, il n'y a pas de diviseur premier p de n tel que $p < n^\alpha$ (en effet, comme l'étape (ii) a été franchie et comme

$A < \sqrt{M}$ $\gcd(a , n) = 1$, si $p|n$, alors $a^{p-1} \equiv 1 \mod p$ (p premier) et si $p < n^{\alpha}$, $m = p - 1$ fournit une solution de (*) vérifiant $1 \leqslant m \leqslant M$).

Si pour tout a, (*) possède une solution, on passe à l'étape (iv).

(iv) Soient m_2 , m_3 , ... , m_A les plus petites solutions de (*) pour chaque valeur de a . Soit $y_a = \gcd(a^{m_a} - 1 , n)$.

Si il existe $a = 2 , \ldots , A$ tel que $y_a < n$. Alors y_a est un facteur non trivial de n .

Si y_a est premier, on a terminé.

Si y_a n'est pas premier, on remplace n par y_a et on revient à l'étape (iii) en conservant les mêmes valeurs pour M et A .

Si pour chaque $a = 2 , \ldots , A$, $y_a = n$ on passe à l'étape (v).

(v) Soit $H = \operatorname{ppcm}(m_2 , m_3 , \ldots , m_A)$.

On cherche si $Hz + 1$ où $z = 1 , \ldots , k - 1$ est un diviseur premier de n .

Si pour un $z = 1 , \ldots , k - 1$, $Hz + 1$ est premier et $Hz + 1|n$ on a terminé, $Hz + 1$ est un facteur premier de n .

Si pour chaque $z = 1 , \ldots , k - 1$, $Hz + 1 \nmid n$ ou n n'est pas premier alors n ne possède pas de facteur premier au plus égal à n^{α} .

Pour achever la démonstration, il reste à établir deux points :

(a) un retour de (iv) à (iii) n'a lieu qu'un nombre fini de fois, au plus $2/\alpha$ en fait car ayant franchi l'étape (ii) on sait que tout facteur premier de n est plus grand que $n^{\alpha/2}$.

(b) L'alternative indiquée en (v) recouvre tous les cas.

En effet, supposons que n possède un facteur premier p avec $p \leqslant n^{\alpha}$ et qu'on soit arrivé à l'étape (v). On a donc pour tout $a = 2 , 3 , \ldots , A$, $y_a = n$. Donc $a^{m_a} \equiv 1 \mod p$. m_a est la plus petite solution en m de $\gcd(a^m - 1 , n) > 1$. C'est aussi la plus petite solution en m de $\gcd(a^m - 1 , p) > 1$. En effet, si $\gcd(a^m - 1 , p) = p$, alors $\gcd(a^m - 1) > 1$ et donc $m \geqslant m_a$ par définition de m_a . Donc $m_a | p - 1$, ceci pour chaque a . Donc $H | p - 1$ et donc $p = Hz + 1$. Il reste à montrer que $z \leqslant k - 1$. Supposons donc $z \geqslant k$. Soit $r = g(p , z) = \min\{b ; z \nmid \ell d\ b\}$. Donc $z \nmid \ell d\ r$ et $r \leqslant A$ car $s_k(p) \leqslant A$. Alors il existe un nombre premier q et un entier u non nul tel que $q^u || z$ et $q^u \nmid \ell d\ r$. Mais $\gcd(r , p) = 1$ car $r \leqslant A \leqslant \sqrt{M}$ (on sait que $p > n^{\alpha/2}$ puisque l'étape (ii) est franchie). Donc il existe m avec $1 \leqslant m \leqslant H$ tel que $r^m \equiv 1 \mod p$. Prenons m minimum. Alors

$$r^m \equiv 1 \mod p \Longrightarrow p - 1 | m \ \ell d\ r .$$

Soit v tel que $q^v || p - 1$. Comme $p - 1 = zH$ et $q^u || z$ on a $q^{v-u} || H$. Mais

$$p - 1 | n \text{ \textit{ssi}d } r \Longrightarrow q^{v} | n \text{ \textit{ssi}d } r .$$

Or $q^{u} \nmid \text{ssid } r$. Donc $q^{v+1-u} | n$ et comme $n | H$, $q^{v+1-u} | H$ ce qui contredit $q^{v-u} \| H$. Donc $z \leqslant k - 1$ et donc si n possède un facteur premier majoré par n^{α} et si l'on est arrivé à l'étape (v), ce facteur premier est de la forme $Hz + 1$ pour un z avec $1 \leqslant z \leqslant k - 1$.

C. Q. F. D.

Le théorème fournit donc un algorithme théorique de factorisation en $O(n^{1/4})$, ce qui est beaucoup trop pour être praticable sur de grands entiers.

2.1.2 L'algorithme $p - 1$ de Pollard. - Dans l'article [9] déjà cité, J. M. POLLARD propose un algorithme de calcul effectif basé sur le théorème précédent et appelé algorithme $p - 1$ de Pollard. Cet algorithme fonctionne comme suit. n étant un nombre naturel donné supposé composé, on choisit des entiers L et M tels que $1 < L < M < n^{1/2}$ et $M < L^{2}$. On choisit aussi un entier a , $a > 1$.

(i) On calcule $P = \prod_{p_i \leqslant L} p_i^{c_i}$, p_i suite des nombres premiers, avec $c_i \geqslant 1$ tels que, par exemple, $p_k \geqslant n$.

On calcule $b = a^{P} \bmod n$, puis $d = \text{pgcd}(b - 1 , n)$.

Si $1 < d < n$ on a terminé, d est un facteur non trivial de n .

Si $d = n$ on retourne à l'étape (i) en prenant L plus petit.

Si $d = 1$ on passe à l'étape (ii).

(ii) On calcule $F_n = b^{n} - 1 \bmod n$, où $L < n < M$ et n premier. Si on trouve un entier n tel que $1 < \text{pgcd}(F_n , n) < n$ ce pgcd est un facteur non trivial de n .

Justification. - Soit q un facteur premier de n tel que tous les diviseurs premiers de $q - 1$ soient bornés par L . Si $\text{pgcd}(a , q) = 1$ alors $q - 1 | P$ à cause du choix des c_i et donc $a^{P} \equiv 1 \bmod q$, et de ce fait $d > 1$.

Si $d = n$ on a donc $a^{P} \equiv 1 \bmod q$ pour tout facteur premier q de n . Prenons L plus petit. P est changé en un facteur du P précédent ayant même valuation pour les nombres premiers majorés par la nouvelle borne L . On peut espérer que $a^{P} \equiv 1 \bmod q$ reste vrai (si $q - 1$ se décompose toujours avec des nombres premiers bornés par le nouveau L) mais que $a^{P} \not\equiv 1 \bmod n$, ceci pour au moins un facteur q . Le nouveau d obtenu est un facteur non trivial de n .

Supposons maintenant $q - 1 = Ap$ où A a tous ses facteurs premiers bornés par L et p est un nombre premier avec $L < p \leqslant M$.

Si $d = 1$ on observe que $Ap | Pp$, $Ap = q - 1$ et $b^{p} = a^{Pp}$, et donc, si $\text{pgcd}(a , n) = 1$, alors $b^{p} \equiv 1 \bmod n$ et on peut donc espérer un facteur non trivial de n en calculant $\text{pgcd}(F_m , n)$ pour $L < m < M$.

On observe qu'à la différence de l'algorithme construit dans la démonstration du

théorème de J. M. POLLARD, cet algorithme peut ne pas aboutir.

On peut voir, heuristiquement qu'il demande souvent un temps de calcul important.

En effet, si p désigne le plus grand facteur premier de $q - 1$, on a, d'après le lemme 2 que le calcul nécessite $O(p)$ opérations. Posons $n_1 = \max\{p ; p|n\}$. On a, sucessivement, d'après la démonstration du lemme 3 dans 2.2.1 :

$$\text{card}\{n \leqslant N ; \ n_1 \leqslant n^X\} \leqslant \text{card}\{n \leqslant N ; \ n_1 \leqslant N^X\} = \Psi(N , N^X) .$$

Et donc, comme $\Psi(N , N^X) \sim \rho(1/x)\, N$,

$$\limsup_{N\to\infty} \frac{1}{N} \text{card}\{n \leqslant N ; \ n_1 \leqslant n^X\} \leqslant \rho(\tfrac{1}{x}) .$$

Donc heuristiquement

$$\underset{\sim}{P}\{n_1 \leqslant T_0\} \leqslant \rho(\frac{\ell n\ n}{\ell n\ T_0})$$

puisque $T_0 = n^{(\ell n T_0)/(\ell n n)}$ et comme $\rho(\ell n\ n/\ell n\ T_0)$ est très petit quand $T_0 \leqslant \sqrt{n}$ on a la conclusion heuristique énoncée ci-dessus.

Par cette méthode, J. M. POLLARD a trouvé les résultats suivants :

$$2^{107} + 2^{54} + 1 = 843589.8174912477117.23528569104401$$

$$121450506296081 \ \text{ divise } \ 10^{95} + 1$$

$$2670091735108484737 \ \text{ divise } \ 3^{136} + 1 .$$

2.2. <u>L'algorithme de Brillhart-Morrison</u>. - L'idée de départ remonte à LEGENDRE. Soit n un entier positif. La congruence $x^2 \equiv y^2 \mod n$ a, lorsque n est premier, exactement les solutions $x \equiv \pm y \mod n$. Si n est composé impair, il y a ces deux solutions dites triviales, mais encore deux autres au moins, puisque si $n = n_1\, n_2$ avec $n_1 , n_2 \neq 1$ les deux congruences $x + y \equiv n_1 \mod n$ et $x - y \equiv n_2 \mod n$, fournissent un couple de solutions modulo n de $x^2 \equiv y^2 \mod n$, et les deux congruences $x + y \equiv - n_1 \mod n$, et, $x - y \equiv n_2 \mod n$ en fournissent un second différent du précédent, modulo n.

Soit alors (u , v) une solution non triviale de $x^2 \equiv y^2 \mod n$ (n nombre composé impair). Alors $d = \text{pgcd}(u - v , n) \neq n$ car $u \not\equiv \pm v \mod n$ et si $d > 1$ on a un facteur non trivial de n.

Comment obtenir des solutions non triviales de $x^2 \equiv y^2 \mod n$? Une idée, remontant aussi à LEGENDRE, consiste à utiliser le développement en fraction continue de $\sqrt{n}$.

Rappelons quelques définitions et propriétés concernant les fractions continues.

Soit x un réel strictement positif. On pose $q_0 = [x]$ et, si $x \notin \underset{\sim}{N}$,

$x = q_0 + 1/x_1$. On applique le même processus à x_1 définissant ainsi $q_1 = [x_1]$ et, si $q_1 \neq x_1$, x_2 par $x_1 = q_1 + 1/x_2$. Si $x \in \underset{\sim}{Q}$ ce processus s'arrête au bout d'un nombre fini d'étapes. Si $x \notin \underset{\sim}{Q}$, il se poursuit indéfiniment. q_i est appelé <u>quotient partiel</u> et les notations sont les suivantes :

$$[q_0 , q_1 , \dots , q_i] = q_0 + \cfrac{1}{q_1 + \cfrac{1}{\dots + \cfrac{1}{q_i}}} = q_0 + \left\lfloor \frac{1}{q_1} \right\rfloor + \dots + \left\lfloor \frac{1}{q_i} \right\rfloor .$$

Si $x \notin \underset{\sim}{Q}$ on pose

$$f_i = [q_0 , q_1 , \dots , q_i] = \frac{A_i}{B_i} .$$

On observe qu'alors $x = [q_0 , \dots , q_i + 1/x_{i+1}]$ et que les A_i , B_i sont liés par les relations de récurrence

$$\forall\, i \geqslant 1 , \quad A_{i+1} = A_{i-1} + q_{i+1} A_i$$

$$\forall\, i \geqslant 1 , \quad B_{i+1} = B_{i-1} + q_{i+1} B_i$$

$$\forall\, i \geqslant 1 , \quad A_i B_{i-1} - A_{i-1} B_i = (-1)^{i-1} .$$

La suite des $q_0 , \dots , q_n , \dots$ est appelée développement en fraction continue de x et on note encore $x = [q_0 , q_1 , \dots , q_n , \dots]$.

Si $d \in \underset{\sim}{N}$, d non carré, le développement en fraction continue de $\sqrt{d}$ présente quelques particularités. Ainsi, on démontre (cf. par exemple [2]) qu'il existe une suite d'entiers naturels non nuls, P_i , Q_i uniques, tels que $x_i = (P_i + \sqrt{d})/Q_i$ où $x = \sqrt{d}$. Les P_i , Q_i sont liés aux A_i , B_i par les relations

$$\forall\, i \geqslant 0 , \quad A_i^2 - d B_i^2 = (-1)^{i+1} Q_{i+1}$$

$$\forall\, i \geqslant 0 , \quad P_i + P_{i+1} = q_i Q_i$$

$$\forall\, i \geqslant 0 , \quad d = P_{i+1}^2 + Q_i Q_{i+1} .$$

Ces deux dernières relations permettent de calculer P_{n+1} , Q_{n+1} à partir de P_i , Q_i et q_i .

La relation $A_i^2 - d B_i^2 = (-1)^{i+1} Q_{i+1}$ fournit des entiers naturels A et Q tels que $A^2 \equiv Q \bmod n$ en prenant $d = n$ ou même kn , $k \in \underset{\sim}{N}^*$. Si Q est un carré, c'est-à-dire si $Q = u^2$ (où $u \in \underset{\sim}{N}^*$), si (A , u) est une solution non triviale de $x^2 \equiv y^2 \bmod n$ et si $\operatorname{pgcd}(A - u , n) > 1$, ce pgcd est un facteur non trivial de n .

D'où la recherche systématique des A_i , Q_{i+1} .

Une idée de KRAITCHIK [4] à nouveau étudiée par LEHMER et POWERS (cf. [5]) permet

d'accélérer la recherche d'une solution non triviale de $x^2 \equiv y^2 \mod n$. En effet, si on considère A_{i_1}, Q_{i_1+1}, ..., A_{i_r}, Q_{i_r+1} si $Q = \prod_{j=1}^{r} (-1)^{i_j+1} Q_{i_j+1}$ est un carré, soit $Q = u^2$, et si $A = \prod_{j=1}^{r} A_{i_j}$, on a encore que (A, u) est une solution de la congruence $x^2 \equiv y^2 \mod n$.

On sait que le développement en fraction continue de $\sqrt{d}$ où $d \in \mathbb{N}^*$, d non carré, est périodique, c'est-à-dire que la suite des q_i est périodique à partir d'un certain rang. Cela résulte aussitôt de ce que (cf. [2] par exemple) $\forall\, i$, $0 < P_i < \sqrt{d}$ et $0 < Q_i < 2\sqrt{d}$ et par définition des P_i, Q_i

$$\frac{P_i + \sqrt{d}}{Q_i} = q_i + \frac{Q_{i+1}}{P_{i+1} + \sqrt{d}} .$$

Les P_i, Q_i sont donc aussi périodiques à partir d'un certain rang. Soit Q_1, ..., Q_a la partie apériodique de la suite des Q_i et Q_{a+1}, ..., $Q_{a+\ell}$ la première période. Soient π_1, ..., π_s l'ensemble des facteurs premiers des Q_i pour $i \leq a + \ell$. On peut écrire chaque Q_i

$$Q_i = \prod_{j=1}^{s} \pi_j^{\alpha_{ji}} \quad \text{donc} \quad (-1)^i Q_i = (-1)^i \prod_{j=1}^{s} \pi_j^{\alpha_{ji}} ,$$

de sorte que

$$Q = \prod_{j=1}^{r} (-1)^{i_j+1} Q_{i_j+1} = (-1)^{\sum (i_j+1)} \prod_{r=1}^{s} \pi_k^{\beta_k} ,$$

où $\beta_k = \sum_{j=1}^{r} \alpha_{kij}$. On a que Q est un carré, si et seulement si

$$\sum_{j=1}^{r} (i_j + 1) \equiv 0 \mod 2 ,$$

et

$$\forall\, k , \quad \sum_{j=1}^{r} \alpha_{kij} \equiv 0 \mod 2 .$$

Si on associe à Q_i le vecteur $(\epsilon_i , \epsilon_{1i} , \ldots , \epsilon_{si})$ où $\epsilon_i , \epsilon_{ji} \in \{0 , 1\}$ et $\epsilon_i \equiv i \mod 2$ et $\epsilon_{ji} \equiv \alpha_{ji} \mod 2$, au produit $\prod_{j=1}^{r} (-1)^{i_j+1} Q_{i_j+1}$ correspond la somme modulo 2 des $(\epsilon_{i_j} , \epsilon_{1i_j} , \ldots , \epsilon_{si_j})$, c'est-à-dire une combinaison linéaire des $(\epsilon_i , \epsilon_{1i} , \ldots , \epsilon_{si})$ dans l'espace vectoriel $(\mathbb{Z}/2\mathbb{Z})^{s+1}$. Si la période est suffisamment longue, et si on rencontre assez de valeurs distinctes de Q_i, on trouvera nécessairement une telle combinaison.

Il restera alors à examiner si la solution (A , u) ainsi trouvée de la congruence $x^2 \equiv y^2 \mod n$ est une solution non triviale, puis, quand c'est le cas, à regarder si $\operatorname{pgcd}(A - u , n) > 1$.

Avant de décrire l'algorithme, on va faire quelques remarques.

La première est que le développement en fraction continue de $\sqrt{d}$ peut avoir une

période très courte. Exemple, si $v \in \mathbb{N}$, $v \geq 1$,

$$\sqrt{v^2+1} = v + \frac{1}{\lfloor 2v \rfloor} + \frac{1}{\lfloor 2v \rfloor} + \cdots$$

puisque $[\sqrt{v^2+1}] = v$, d'où on tire

$$\sqrt{v^2+1} = v + (\sqrt{v^2+1} - v) = v + (\sqrt{v^2+1} + v)^{-1}$$

d'où à nouveau

$$\sqrt{v^2+1} + v = 2v + (\sqrt{v^2+1} + v)^{-1} .$$

On peut pallier cet inconvénient en remplaçant n par kn où k est un entier sans facteur carré, petit par rapport à n. Le plus souvent, la période du développement en fraction continue de $\sqrt{d}$ est très approximativement de l'ordre de $\sqrt{d}$ et, expérimentalement, il suffit de prendre $k \leqslant 1000$.

La seconde remarque est que de l'encadrement $0 < P_i < \sqrt{d}$ et $0 < Q_i < 2\sqrt{d}$, on tire que P_i et Q_i se factorisent "au pire" sur $p_1, \ldots, p_r$ la liste des nombres premiers au plus égaux à $\sqrt{2}\sqrt[4]{d}$, d'où, a priori, des chances de succès de la méthode si la période du développement est de l'ordre de $\sqrt{d}$ (même si la partie "utile" de cette période est réduite de moitié du fait d'une symétrie des valeurs des q_i à l'intérieur d'une période). Cependant, cette valeur de r de l'ordre de $4\sqrt{2}\sqrt[4]{d}/\ell n\, d$, donc en $N^{1/4}/\ell n\, N$ est trop grande.

On va donc chercher une valeur de r plus petite susceptible de réduire le temps de calcul. Suivant l'analyse de [6] on pose $Q = [2\sqrt{d}]$ et on suppose que la proportion des Q_i dont les facteurs premiers sont majorés par p_r est $\rho(x)$ où $p_r = Q^{1/x}$. Pour produire r nombres Q_i factorisables entièrement sur $p_1, \ldots, p_r$ il faut donc calculer environ $r/\rho(x)$ nombres Q_i. Pour chacun d'eux, la factorisation nécessite environ r divisions (on néglige les facteurs multiples) ce qui représente $(r^2/\rho(x))\,D(n)$ (où $D(n)$ est un majorant du nombre d'opérations nécessaire pour diviser un nombre majoré par n) en négligeant le calcul des Q_i eux-mêmes. La recherche d'une combinaison linéaire nulle dans $(\mathbb{Z}/2\mathbb{Z})^{r+1}$ représente environ r^3 additions binaires (par la méthode de Gauss, par exemple). Comme une division nécessite beaucoup plus d'opérations qu'une division, on a donc qu'on peut exprimer le nombre d'opérations nécessaire au calcul pour

$$T(n\ ,\ r) = r^3 + \lambda \frac{r^2}{\rho(x)} .$$

On approche $\rho(x)$ par x^{-x} et comme $x = \ell n\, Q/\ell n\, p_r$ et que $p_r \sim r\, \ell n\, r$, le calcul donne (cf. [6]) que $T(n\ ,\ r)$ passe par un minimum pour $r_0 = \exp(\frac{1}{2}\sqrt{\ell n\, Q\ \ell n \ell n Q})$.

En fait, on n'a pas besoin de prendre tous les nombres premiers inférieurs à p_r. Car, du fait que $A_i^2 - d\,B_i^2 = (-1)^{i+1} Q_{i+1}$, si $p \mid Q_{i+1}$, comme $\operatorname{pgcd}(A_i\ ,\ B_i) = 1$, $p \nmid B_i$ et donc d est résidu quadratique modulo p. Donc on ne garde que les nombres premiers p pour lesquels $(kn/p) = 1$ et les facteurs premiers de k (ce qui peut guider le choix de k). De la sorte on prendra r

nombres premiers "utiles".

D'où l'algorithme :

On prend n impair et composé. On fixe $k = 1$.

(i) On fixe r , et on prend $i_0 = [r/\rho(x)] + 1$ où $x = \ell n[2\sqrt{kn}]/\ell n(r \ \ell n \ r)$ (de sorte que $p_r \leq Q^{1/x}$ où $Q = [2\sqrt{kn}]$, (cf. ci-dessus).

(ii) On définit la "base" des nombres premiers π_1 , ... , π_r , en testant au fur et à mesure si $\pi_j | n$. Si oui, on a terminé : on a un facteur premier de n . Si non, on passe à l'étape (iii).

(iii) On développe $\sqrt{kn}$ en fraction continue. A chaque étape i avec $i \leq i_0$:

(a) on calcule A_i , q_i , P_i , Q_i . Si $Q_i = 1$, la période étant trop courte, on fixe une autre valeur de k et on reprend en (i).

(b) On regarde si Q_i se factorise entièrement sur la base des π_j . Si oui, on conserve $(\epsilon_i , \epsilon_{1i} , \ldots , \epsilon_{ri})$ et on passe à l'étape (c). Sinon, on regarde si $Q < \pi_r^2$, où Q est le quotient de Q_i par son plus grand diviseur construit sur la base des π_j . Si $Q < \pi_r^2$, Q est premier, et on retient la décomposition sous la forme $(\epsilon_i , \epsilon_{1i} , \ldots , \epsilon_{ri} , Q)$. Sinon, on passe à l'étape $(i) + 1$.

(c) On regarde si Q_i est un carré. Si oui, on passe directement à l'étape (iv), sinon, à l'étape $(i) + 1$.

(iv) Parmi les $n + 1$ vecteurs obtenus, on cherche, par la méthode de Gauss, une combinaison linéaire pour obtenir $\prod(-1)^{i_j} Q_{i_j} = u^2$ $(u \in \mathbb{N}^*)$. Soit $A \equiv \prod A_{i_j} \mod n$.

On calcule $m = \text{pgcd}(A - u , n)$.

Si $m = n$, (A , u) est une solution triviale de la congruence $x^2 \equiv y^2 \mod n$, on continue l'élimination gaussienne pour obtenir une autre combinaison linéaire donnant de nouvelles valeurs de A et u .

Si l'élimination est terminée, on reprend l'étape (iii) en fixant une nouvelle valeur de i_0 supérieure à la précédente.

Si $m = 1$ on a obtenu un facteur trivial de n , on procède comme pour le cas $m = n$.

Si $1 < m < n$ on a terminé : m est un facteur non trivial de l'entier n .

L'article [7] donne de nombreux détails sur le moyen de trier les Q_i et de les décomposer sur les π_j de la façon la plus économique, ainsi que sur un moyen d'accélérer l'élimination gaussienne.

Heuristiquement, le temps de calcul de l'algorithme est (cf. [6])
$T(n , r_0) \sim \wedge \exp(2\sqrt{\ell n \ Q \ \ell n \ell n \ Q})$, c'est-à-dire $O(e^{2\sqrt{\ell n \, n \, \ell n \ell n n}})$ soit

$$O(n^{2\sqrt{\ell n \ell n\, n / \ell n\, n}}) = O((\ell n\ n)^{2\sqrt{\ell n\, n / \ell n \ell n\, n}}) .$$

Ce temps n'est ni polynômial, ni exponentiel. Observons que pour $n \simeq 10^{50}$ et $\lambda = 1000$ on trouve $\lambda_0 = 2196$ ce qui est assez raisonnable.

L'expérience montre que le nombre d'échecs (pour des nombres d'au plus 46 chiffres), c'est-à-dire de retours à une autre valeur de k, est petit et est toujours suivi d'un succès.

Parmi les nombres cassés par cette méthode on a :

$$F_7 = 2^{128} + 1 = 59649589127497217.5704689200685129054721$$

1330000 Q_i ont été calculés dans le développement de $\sqrt{257\ F_7}$. Sur ceux-ci 2059 ont été factorisés, conduisant à la congruence non triviale

$$2335036483808358521772321436182279564762$$

$$\equiv 2518647814572804129731227193485202122232 \mod F_7 .$$

F. KLEIN a annoncé le caractère composé de F_7 en 1895. MOREHEAD et WESTERN l'ont vérifié en 1905 sans donner de factorisation. La factorisation ci-dessus a été obtenue en 1970 (2 heures de calcul et 1504 K de mémoire utilisée).

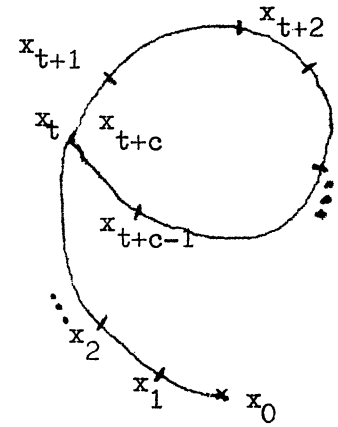
2.3. La méthode ρ de POLLARD. - Cette méthode exposée dans [10], est qualifiée de Monte Carlo par son auteur en raison des considérations probabilistes sur lesquelles elle se fonde.

2.3.1. L'algorithme ρ de Pollard. - Considérons une application f de $(\mathbb{Z}/n\mathbb{Z})$ dans lui-même, où n est un entier, $n \geqslant 2$. On définit une suite x_i d'entiers modulo n par

$$x_0 \in \mathbb{Z}/n\mathbb{Z} \quad \text{et} \quad x_{i+1} = f(x_i) \quad \text{pour} \quad i \in \mathbb{N} .$$

Comme f prend un nombre fini de valeurs et que x_{i+1} ne dépend que de x_i, la suite des x_i est périodique à partir d'un certain rang. Soit $x_0, x_1, \cdots, x_{t-1}$ la partie apériodique de la suite et c la longueur de la plus petite période (de sorte que $x_{t+c} = x_t$).

On peut représenter la suite $x_0, x_1, \cdots, x_{t-1}, x_t, \cdots, x_{t+c-1}$ par un dessin rappelant celui de la lettre grecque rho, la partie apériodique constituant le jambage de la lettre et la partie périodique s'enroulant sur le corps de la lettre.



Le rho est entièrement décrit par $x_0, x_1, \cdots, x_{t-1}, x_t, \cdots, x_{t+c-1}$, c'est-à-dire par les $t + c$ premiers éléments de la suite. On pose $\ell = t + c$ qu'on appelle longueur du rho. Comme les ℓ valeurs considérées sont deux à deux

distinctes, $\ell \leqslant n$.

Soit maintenant une application f de $\underset{\sim}{Z}$ dans $\underset{\sim}{Z}$ et un entier n composé. Soit p un diviseur premier de n . Partant de $x_0 \in \underset{\sim}{Z}$, on définit $x_{i+1} = f(x_i)$ pour $i \in \underset{\sim}{N}$ comme ci-dessus. Comme précédemment, la suite des x_i mod n est périodique à partir d'un certain rang et de même pour la suite des x_i mod p . On peut définir rho pour chacune de ces suites, et sa longueur est $\ell(n)$ dans le premier cas, $\ell(p)$ dans le second. A priori, on peut penser que $\ell(p) < \ell(n)$ et, t et c recevant la signification de tout à l'heure dans le cas des congruences modulo p on peut espérer que $x_{t+c} \not\equiv x_t \bmod n$ auquel cas $\text{pgcd}(x_{t+c} - x_t , n)$ est un facteur non trivial de n (puisque $p \mid x_{t+c} - x_t$).

Ne connaissant pas p , on ne connait pas a priori la période des x_i . La détecter en testant $x_i \equiv x_j \bmod n$ à chaque pas du calcul des x_i demande $\ell(\ell - 1)/2$ tests (où ℓ est la longueur du rho associé à n) et demande surtout de garder en mémoire tous les x_i pour $i \leqslant \ell - 1$. Si n est de l'ordre de 10^{30} c'est tout-à-fait exclus. On peut avoir une estimation de la longueur du rho grâce au lemme suivant.

LEMME. - Soit f une application de $\underset{\sim}{Z}$ dans $\underset{\sim}{Z}$ et $x_0 \in \underset{\sim}{Z}$. La suite x_i définie par $x_{i+1} = f(x_i)$ pour $i \in \underset{\sim}{N}$ est périodique à partir d'un certain rang si et seulement si, il existe un i tel que $x_{2i} = x_i$.

Démonstration. - La condition est évidemment suffisante et c'est une période, multiple de la plus petite. Supposons la suite des x_i périodique à partir d'un certain rang. Soit $x_0 , \ldots , x_{t-1}$ sa partie apériodique. Si $i < t$, $x_{2i} \neq x_i$ par définition de la partie apériodique de la suite. Soit donc $i = t + u$, où $u \in \underset{\sim}{N}$. Alors $x_{2i} = x_i$, si et seulement si $2i - i \equiv 0 \bmod c$, c'est-à-dire $c \mid i$. Il suffit donc de prendre pour i le plus petit multiple de c au moins égal à t.

Soit alors $r(n)$ le plus petit indice i non nul tel que $x_{2i} \equiv x_i \bmod n$. Alors $t(n) \leqslant r(n) \leqslant t(n) + c(n)$.

D'où l'algorithme :

(i) on fixe une application f de $\underset{\sim}{Z}$ dans $\underset{\sim}{Z}$,

(ii) on choisit x_0 ; on pose $x = x_0$ et $y = x_0$,

(iii) on remplace x par $f(x)$ et y par $f(f(y))$,

(iv) on calcule $d = \text{pgcd}(x - y , n)$.

Si $1 < d < n$, on a terminé : d est un diviseur non trivial de n .

Si $d = 1$, on retourne en (iii).

Si $d = n$, on a terminé : on n'a pas trouvé de diviseur non trivial de n . On recommence, soit en revenant à (ii) (on prend un autre x_0), soit en revenant à (i) (on prend une autre fonction).

2.3.2. Temps de calcul de la méthode ; applications. - Le temps de calcul de la méthode est r et demande une mémoire quasiment négligeable.

On va considérer ℓ comme une variable aléatoire en faisant l'hypothèse que n étant fixé, toutes les applications de $\mathbb{Z}/n\mathbb{Z}$ dans lui-même, sont équiprobables. k étant fixé, on a donc

$$\mathrm{P}\{\ell = k\} = \sum_{\iota=0}^{k-1} \mathrm{P}\{t = \iota \ ; \ c = k - \iota\} \ .$$

Or,

$$\mathrm{P}\{t = \iota \ ; \ c = k - \iota\} = \frac{1}{n} \prod_{1 \leq j < k} \left(1 - \frac{j}{n}\right)$$

puisque x_k est fixé, et que x_0 , ... , x_{k-1} sont deux à deux distincts, compris entre 0 et $n - 1$ et distincts de x_k . Donc

$$\mathrm{P}\{\ell = k\} = \frac{k}{n} \prod_{1 \leq j < k} \left(1 - \frac{j}{n}\right) \ .$$

D'où

$$\ell n \ \mathrm{P}\{\ell = k\} = \ell n \frac{k}{n} + \sum_{1 \leq j < k} \ell n\left(1 - \frac{j}{n}\right)$$

$$= \ell n \frac{k}{n} - \frac{k^2}{2n} + O\left(\frac{k}{n}\right) \ .$$

D'où (cf. [6])

$$\mathrm{P}\{\ell = k\} = \frac{k}{n} e^{-k^2/2n} \left(1 + O\left(\frac{k}{n}\right)\right) \ .$$

On peut donc approcher $E(\ell)$ par $\sum_{1 \leq k \leq n} (k^2/n) \, e^{-k^2/2n}$ qu'on approche par $\int_0^n (x^2/n) \, e^{-x^2/2n} \, dx \sim \sqrt{(\pi/2)\, n}$ (calcul facile) d'où on déduit $E(\ell) \sim \sqrt{(\pi/2)\, n}$.

Par ailleurs, comme $\ell = t + c$ et que t et c jouent des rôles symétriques en prenant pratiquement les mêmes valeurs, on en déduit que

$$E(t) \sim E(c) \sim \sqrt{\frac{\pi}{8}\, n} \ .$$

Un calcul précis (cf. [3], vol. 1, p. 117 et vol. 2, p. 454) donne

$$E(\ell) = \sqrt{\frac{\pi}{2}\, n} - \frac{1}{3} + O\left(\frac{1}{\sqrt{n}}\right) \ .$$

Comme $r \leq \ell \leq n$ on a un temps de calcul $O(p^{1/2})$ où p est le plus petit facteur premier de n .

On donne dans [6] un calcul approché de $E(r)$ permettant de retrouver le résultat annoncé dans [10] à savoir $E(r) \sim r^2/12 \sqrt{(\pi/2)\, n}$.

En effet, r est entièrement déterminé par ℓ et c puisque si

$$\lceil x \rceil = \min\{k \in \mathbb{Z} \ ; \ k \geq x\} \ , \quad r = c \left\lceil \frac{\ell - c}{c} \right\rceil \ .$$

Donc

$$E(r) = \sum_{0<\gamma\leq\lambda\leq n} r(\lambda , \gamma)\ \underset{\sim}{P}\{\ell = \lambda\ ,\ c = \gamma\} .$$

Mais $\underset{\sim}{P}\{\ell = \lambda\ ,\ c = \gamma\} = \underset{\sim}{P}\{t = \lambda - \gamma\ ;\ c = \gamma\} = \frac{1}{\lambda}\ \underset{\sim}{P}\{\ell = \lambda\}$. (cf. ci-dessus). Donc

$$E(r) = \sum_{0<\lambda\leq n} \underset{\sim}{P}\{\ell = \lambda\}\left(\frac{1}{\lambda} \sum_{0<\gamma\leq\lambda} \gamma \left\lceil \frac{\lambda - \gamma}{\gamma} \right\rceil\right) .$$

Soit $a(\lambda) = \frac{1}{\lambda} \sum_{0<\gamma\leq\lambda} \gamma \lceil(\lambda - \gamma)/\gamma\rceil$. Comme $\lceil(\lambda - \gamma)/\gamma\rceil = \lceil\lambda/\gamma\rceil - 1$ on a

$$a(\lambda) = -\frac{1 + \lambda}{2} + \frac{1}{\lambda} \sum_{0<\gamma\leq\lambda} \gamma \left\lceil \frac{\lambda}{\gamma} \right\rceil .$$

Or, $\frac{1}{\lambda} \sum_{0<\gamma\leq\lambda} \gamma \left\lceil \frac{\lambda}{\gamma} \right\rceil = \lambda \sum_{0<\gamma\leq\lambda} \frac{\gamma}{\lambda} \left\lceil \frac{\lambda}{\gamma} \right\rceil \frac{1}{\lambda}$ qu'on approche donc par $\lambda \int_0^1 x \left\lceil \frac{1}{x} \right\rceil dx$. Mais

$$\int_0^1 x \left\lceil \frac{1}{x} \right\rceil dx = \sum_{n=1}^{+\infty} \int_{1/n+1}^{1/n} x \left\lceil \frac{1}{x} \right\rceil dx = \sum_{n=1}^{+\infty} (n + 1) \int_{1/n+1}^{1/n} x\, dx$$

$$= \sum_{n=1}^{+\infty} (n + 1)\ \frac{1}{2}\left(\frac{1}{n^2} - \frac{1}{(n + 1)^2}\right) = \sum_{n=1}^{+\infty} \frac{1}{2}\left(\frac{1}{n^2} + \frac{1}{n} - \frac{1}{n + 1}\right)$$

$$= \frac{1}{2} \sum_{n=1}^{+\infty} \frac{1}{n^2} + \frac{1}{2} \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n + 1}\right) = \frac{\pi^2}{12} + \frac{1}{2} .$$

D'où $a(\lambda) = \lambda \frac{\pi^2}{12} - \frac{1}{2}$ et donc

$$E(r) \sim \frac{\pi^2}{12} \sum_{0<\lambda\leq n} \lambda\ \underset{\sim}{P}\{\ell = \lambda\} - \frac{1}{2} \sum_{0<\lambda\leq n} \underset{\sim}{P}\{\ell = \lambda\} .$$

Soit $E(r) \sim (\pi^2/12)\sqrt{(\pi/2)\ n}$.

C. Q. F. D.

Dans [10], J. M. POLLARD présente une version un peu modifiée de cet algorithme.

(a) Il fixe $f(x) = x^2 + b$ où $b \neq 0\ ,\ -2$ afin d'éviter des valeurs de x_0 pour lesquelles ℓ serait de l'ordre de p .

(b) Le calcul du pgcd étant très légèrement supérieur à celui d'un produit, J. M. POLLARD remplace l'étape (iv) par le calcul de $Q_i \equiv \prod_{j=1}^{i} (x_{2j} - x_j) \bmod n$, où i est un entier fixé, et il calcule ensuite $\mathrm{pgcd}(Q_i\ ,\ n)$.

La modification (b) entraîne un gain sur le temps de calcul, mais le risque d'échec de la méthode est plus important. En effet, si $n = pq$, p et q premiers et proches on peut avoir $x_{2i_1} - x_{i_1} \equiv 0 \bmod p$ et $x_{2i_2} - x_{i_2} \equiv 0 \bmod q$ avec $0 < |i_2 - i_1| < i$ de sorte que $\mathrm{pgcd}(Q_i\ ,\ n) = n$. Il convient alors de changer de fonction f . L'expérience montre qu'en général et pour n pas trop grand, deux fonctions au plus permettent de factoriser n .

Enfin, il convient de remarquer que f ne décrit pas l'ensemble des fonctions de $\underset{\sim}{Z}/n\underset{\sim}{Z}$ dans lui-même, mais un ensemble notablement restreint. Les simulations

numériques rapportées dans [6] permettent de penser que λ et r ont asymptotiquement la même distribution sur cet ensemble restreint que sur l'ensemble plein.

Par cette méthode, J. M. POLLARD a obtenu les factorisations suivantes :

$$2^{77} - 3 = 1291.99432527.1177212722617$$

$$2^{79} - 3 = 5.3414023.146481287.241741417 .$$

2.4. La méthode de LENSTRA. - Cette méthode (cf. [8]) est basée sur la propriété de l'ensemble des points d'une courbe elliptique de posséder une structure de groupe (en fait une infinité de telles structures).

2.4.1. Rappels sur les courbes elliptiques. - Une courbe elliptique est l'ensemble des points du plan projectif $P_2(\mathbb{C})$ dont les coordonnées sont racines d'un polynôme homogène en X , Y , T , de degré trois, irréductible et tel qu'en tout point qui annule le polynôme, sa dérivée première ne s'annule pas.

On démontre alors (cf. par exemple [13], chap. III) que toute courbe elliptique admet une équation de la forme

$$Y^2 T = 4X^3 - g_2 XT^2 - g_3 T^3 \qquad \text{(i)}$$

avec $g_2^3 + 27g_3^2 \neq 0$ puisque on suppose l'irréductibilité du polynôme. On désigne par E la courbe d'équation (i).

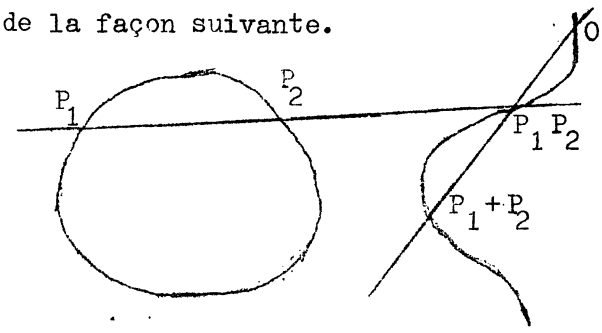
Il est facile de voir que E admet un seul point à l'infini $(0 , 1 , 0)$ qui est un point d'inflexion. Si F est défini par (i), le calcul de $\partial F/\partial X$, $\partial F/\partial Y$, $\partial F/\partial T$ en $(0 , 1 , 0)$ montre que la tangente en ce point à E est la droite à l'infini (d'équation $T = 0$).

Pour les points du plan affine qu'on peut prendre de coordonnées $(x , y , 1)$, l'équation (i) devient

$$y^2 = 4x^3 - g_2 x - g_3 . \qquad \text{(ii)}$$

On définit sur E une loi de groupe de la façon suivante.

On fixe un point 0 sur la courbe qui sera, par définition le neutre de la loi de groupe.



Si $P_1 P_2 \in E$, la droite passant par ces points (si $P_1 = P_2$, la tangente passant par ce point) recoupe E en un point noté $P_1 P_2$. Alors la droite joignant 0 à $P_1 P_2$ coupe à nouveau E en un point que l'on note $P_1 + P_2$, (cf. figure ci-dessus).

Cette loi est évidemment commutative. On peut montrer, par le calcul, qu'elle est associative. Le calcul des coordonnées de $P_1 + P_2$ est relativement simple

quand, pour la courbe E d'équation (i), on prend pour neutre le point à l'infini. On voit alors immédiatement que le symétrique de (x , y) est $(x , - y)$ car $(x , y , 1)$, $(x , - y , 1)$ et $(0 , 1 , 0)$ sont alignés.

Soient $P_1 , P_2 \in E$ de coordonnées respectives $(x_1 , y_1 , 1)$ et $(x_2 , y_2 , 1)$ avec $P_1 \neq P_2$. On suppose P_1 et P_2 à distance finie, car $P + 0 = P$ pour tout $P \in E$ et 0 est le seul point à l'infini de E . Soient $(x_3 , y_3 , 1)$ les coordonnées de $P_1 + P_2$. Celles de $P_1 P_2$ sont donc $(x_3 , - y_3 , 1)$. L'équation de la droite joignant P_1 à P_2 est (si $x_1 \neq x_2$)

$$y = a(x - x_1) + y_2 \quad \text{avec} \quad a = (y_2 - y_1)/(x_2 - x_1) .$$

Par définition, x_1 , x_2 , x_3 sont les solutions de (ii) quand on y remplace y par $ax + b$. Dans le polynôme en x obtenu après cette substitution, la somme des racines est $a^2/4$. Donc

$$\text{(iii)} \qquad \begin{cases} x_3 = - x_1 - x_2 + \frac{1}{4} \left(\frac{y_2 - y_1}{x_2 - x_1} \right)^2 \\ y_3 = - y_1 - \frac{y_2 - y_1}{x_2 - x_1} (x_3 - x_1) . \end{cases}$$

Si $x_1 = x_2$, comme on suppose $P_1 \neq P_2$, on a alors $P_1 + P_2 = 0$. Si $P_1 = P_2$, on définit $P_1 + P_2 = 2P$, par passage à la limite lorsque P_2 tend vers P_1 sur E (on retrouve la définition donnée plus tôt). La droite joignant P_1 à P_2 devient la tangente en P_1 à E d'équation

$$y = y_1 + \frac{12 x_1^2 - g_2}{2 y_1} (x - x_1) \quad \text{(on suppose } y_1 \neq 0 \text{)} .$$

Ce qui donne

$$\text{(iv)} \qquad \begin{cases} x_3 = - 2x_1 + \frac{1}{4} \left(\frac{12 x_1^2 - g_2}{2 y_1} \right)^2 \\ y_3 = - y_1 - \frac{12 x_1^2 - g_2}{2 y_1} (x_3 - x_1) . \end{cases}$$

On observe que si $y_1 = 0$, la tangente a pour équation $x = x_1$ et dans ce cas $2P_1 = 0$ (Nota Bene : ce qui n'entraîne pas $P_1 = 0$).

Les formules (iii) et (iv) se transportent en coordonnées projectives. Il suffit de remplacer, formellement, x_i et y_i par, respectivement, x_i/t_i et y_i/t_i . Dans le cas où $P_1 \neq P_2$ et P_1 , P_2 ne sont pas alignés avec 0 (c'est-à-dire non symétriques par rapport à Ox), le calcul donne

$$x_3/t_3 = - \frac{x_1 t_2 + x_2 t_1}{t_1 t_2} + \frac{1}{4} \left(\frac{y_2 t_1 - y_1 t_2}{x_2 t_1 - x_1 t_2} \right)^2$$

en observant que P_1 , P_2, étant à distance finie, $t_1 , t_2 \neq 0$ et comme P_1 , P_2 ne sont pas alignés avec 0 , $x_2/t_2 \neq x_1/t_1$ soit $x_2 t_1 - x_1 t_2 \neq 0$. On trouve

donc

$$x_3/t_3 = \frac{(y_2\ t_1 - y_1\ t_2)^2\ t_1\ t_2 - 4(x_1\ t_2 + x_2\ t_1)(x_2\ t_1 - x_1\ t_2)^2}{4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^2}$$

d'où

$$y_3/t_3 = -\ y_1/t_1 - \frac{y_2\ t_1 - y_1\ t_2}{x_2\ t_1 - x_1\ t_2}\ (x_3/t_3 - x_1/t_1)\ .$$

On pose

$$X = (y_2\ t_1 - y_1\ t_2)^2\ t_1\ t_2 - 4(x_1\ t_2 + x_2\ t_1)(x_2\ t_1 - x_1\ t_2)^2\ .$$

Alors

$$y_3/t_3 = \frac{-\ 4t_2\ y_1(x_2t_1-x_1t_2)^3 - X(y_2t_1-y_1t_2) + 4t_2\ x_1(y_2t_1-y_1t_2)(x_2t_1-x_1t_2)^2}{4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^3}$$

d'où, après simplifications

$$y_3/t_3 = \frac{4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^2\ (x_1\ y_2 - y_1\ x_2) - X(y_2\ t_1 - y_1\ t_2)}{4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^3}$$

ce qui permet d'écrire

$$\text{(v)}\qquad \begin{cases} x_3 = X(x_2\ t_1 - x_1\ t_2) \\ y_3 = 4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^2\ (x_1\ y_2 - y_1\ x_2) - X(y_2\ t_1 - y_1\ t_2) \\ t_3 = 4t_1\ t_2(x_2\ t_1 - x_1\ t_2)^3\ . \end{cases}$$

Dans le cas où P_1 et P_2 sont alignés avec 0, $P_1 + P_2 = 0$. Si $P_1 = P_2$ et $y_1 \neq 0$ les formules (iv) deviennent, après calcul :

$$\text{(vi)}\qquad \begin{cases} X = (12x_1^2 - g_2\ t_1^2)^2 - 32\ x_1\ y_1^2\ t_1 \\ x_3 = 2X\ y_1\ t_1 \\ y_3 = -(12x_1^2 - g_2\ t_1^2)(X - 16x_1\ y_1^2\ t_1^2) - 32y_1^4\ t_1^2 \\ t_3 = 32y_1^3\ t_1^3\ . \end{cases}$$

2.4.2. <u>L'algorithme de LENSTRA</u>. - L'idée de la méthode est la suivante. Considérons pour p, nombre premier fixé, l'ensemble E_p des solutions, dans $P_2(\mathbb{Z}/p\mathbb{Z})$ de l'équation (i). Les formules (v) et (vi) du paragraphe 2.4.1 restent valables et on obtient un groupe fini. En particulier, si $n_p = \text{card}(E_p)$, on aura pour tout $P \in E_p$: $n_p\ P = 0$ ce qui se traduit, si $nP = (x_n\ ,\ y_n\ ,\ t_n)$ par $t_{n_p} \equiv 0 \mod p$. Si donc $p \mid n$, on peut espérer que $\text{pgcd}(t_{n_p}\ ,\ n)$ est un diviseur non trivial de n.

Mais on ne connait pas n. LENSTRA propose la stratégie suivante :

(i) n est donné, composé et impair.

(ii) on se place dans $(\mathbb{Z}/n\mathbb{Z})^3$ et on choisit $P = (x_0 , y_0 , 1)$ et on choisit $g_2 , g_3 \in \mathbb{Z}$ tels que

$$y_0^2 \equiv 4x_0^3 - g_2 x_0 - g_3 \mod n .$$

(iii) On fixe un entier k (voir ci-dessous comment).

(iv) Pour j variant de 1 à k, pour chaque pas :

(v-a) on définit $P_{j!} = j! P = (x_{j!} , y_{j!} , t_{j!}) \mod n$, grâce aux formules (v) et (vi) du § 2.4.1.

(v-b) On calcule $d = \text{pgcd}(t_{j!} , n)$.

Si $1 < d < n$, on a terminé, d est un facteur non trivial de n .

Sinon, on passe à l'étape $j + 1$.

(vi) Si $\text{pgcd}(t_{k!} , n) = 1$ ou n l'algorithme a échoué.

Analysons, heuristiquement, cet algorithme (cf. [8]). Un théorème de Hasse dit que $p + 1 - 2\sqrt{p} < n_p < p + 1 + 2\sqrt{p}$ de sorte que, heuristiquement, quand on parcourt l'ensemble des courbes elliptiques modulo p on puisse considérer n_p comme un nombre au hasard voisin de p . En outre, il suffit que $n_p | j!$ pour que $t_{j!} \equiv 0 \mod p$. Heuristiquement, la probabilité pour que $n_p | k!$ est de l'ordre de $1/p\ \Psi(p , k)$ (si les facteurs premiers de n_p ne dépassent pas k, on a de fortes chances que $n_p | k!$). Donc il convient d'essayer $p/(\Psi(p , k))$ courbes elliptiques.

Soit $p = n^{\alpha}$ avec $0 < \alpha < 1$ et $k = L^{\beta}$ où $L = \exp\sqrt{\ell n\ n\ \ell n \ell n\ n}$. Alors

$$\frac{p}{\Psi(p , k)} = \exp[(\frac{\alpha}{\beta} \frac{\ell n\ n}{\sqrt{\ell n\ n\ \ell n \ell n\ n}} \ell n(\frac{\alpha}{\beta} \frac{\ell n\ n}{\sqrt{\ell n\ n\ \ell n \ell n\ n}}))(1 + O(1))]$$

Or

$$\ell n \frac{\alpha}{\beta} \frac{\ell n\ n}{\sqrt{\ell n\ n\ \ell n \ell n\ n}} = \ell n \frac{\alpha}{\beta} + \ell n \sqrt{\frac{\ell n\ n}{\ell n \ell n\ n}} = \frac{1}{2}(\ell n \ell n\ n)(1 + O(1)) .$$

Donc

$$\frac{p}{\Psi(p , k)} = \exp[(\frac{\alpha}{\beta} (\sqrt{\frac{\ell n\ n}{\ell n\ n\ \ell n \ell n\ n}}) \frac{1}{2} \ell n \ell n\ n)(1 + O(1))]$$

$$= \exp[\frac{\alpha}{2\beta}(\sqrt{\ell n\ n\ \ell n \ell n\ n})(1 + O(1))]$$

$$= L^{(\alpha/2\beta)+O(1)} .$$

Pour une courbe elliptique, le calcul de $k!\ P$ demande $\ell n\ k!$ opérations (à un facteur constant près). Il est donc en $O(k\ \ell n\ k)$. Donc pour $p/(\Psi(p , k))$ courbes, le calcul demande $L^{(\alpha/2\beta)+\beta+O(1)}$ opérations.

α étant fixé, le minimum de $(\alpha/2\beta) + \beta$ est atteint pour $\beta = \sqrt{\alpha/2}$ et vaut $\sqrt{2\alpha}$. Donc, si on prend $\beta = \sqrt{\alpha/2}$ le calcul demande $L^{\sqrt{2\alpha}+O(1)}$ opérations.

Le calcul de $j_1 P$ demande très peu de mémoire. On peut donc organiser le calcul des $p/\Psi(p, k)$ courbes en parallèle, k étant fixé d'après le calcul précédent.

Pour $n = 10^{50}$ et $\alpha = 1/4$, on trouve $L^{\sqrt{\alpha/2}} \sim 3884$. Il convient donc, pour de telles données, de tester la méthode sur près de 4000 courbes elliptiques, ce qui peut effectivement s'organiser en parallèle (puisque pour $\beta = \sqrt{\alpha/2}$, $p/\Psi(p, k) = L^{\sqrt{\alpha/2}+O(1)}$).

BIBLIOGRAPHIE

[1] DICKMAN (Karl). - On the frequency of numbers containing prime factors of a certain relative magnitude, Arkiv för Mat., Astron. och Fys., Series A, t. 22, 1930, fasc. 10, p. 1-14.

[2] HUA (Loo Keng). - Introduction in number theory. - Berlin, Heidelberg, New York, Springer-Verlag, 1982.

[3] KNUTH (Donald E.). - The art of computer programming, vol. 1 and 2. - Reading, Addison-Wesley publishing Company, 1969.

[4] KRAITCHIK (M.). - Théorie des nombres. Vol. 2 : Analyse indéterminée du 2e. degré et factorisation. - Paris, Gauthier-Villars, 1926.

[5] LEHMER (D. H.) and POWERS (R. E.). - On factoring large numbers, Bull. Amer. math. Soc., t. 37, 1931, p. 770-776.

[6] MONIER (L.). - Algorithmes de factorisation d'entiers, Thèse de 3e cycle, Orsay, 1980.

[7] MORRISON (Michael A.) and BRILLHART (John). - A method of factoring and the factorization of F_7, Math. of Comput., t. 29, 1975, p. 183-205.

[8] NICOLAS (J.-L.). - Tests de primalité et méthodes de factorisation (à paraître).

[9] POLLARD (J. M.). - Theorems on factorization and primality testing, Proc. Cambridge phil. Soc., t. 76, 1974, p. 521-528.

[10] POLLARD (J. M.). - A Monte Carlo method for factorization, Nordisk Tidskr. Informationsbehandling (BIT), t. 15, 1975, p. 331-334.

[11] RIVEST (R. L.), SHAMIR (A.) and ADLEMAN (L.). - A method for obtainning digital signatures and public key cryptosystems, Comm. Assoc. comput. Mach., t. 21, 1978, p. 120-126.

[12] VINOGRADOV (I. M.). - Elements of number theory. - New York, Dover Publications, 1954.

[13] WALKER (R. J.). - Algebraic curves. - New York, Dover Publications, 1962.