N° d'ordre : 06 / PGS – 2004 / MT

République Algérienne Démocratique et Populaire

Université des Sciences et de la Technologie Houari Boumediene

جامعة هواري بومدين
للعلوم والتكنولوجيا
U S T H B

FACULTE DES MATHEMATIQUES

Mémoire présenté pour l'obtention du diplôme de PGS en :

MATHEMATIQUES

Spécialité : CRYPTOLOGIE

Par

ZEDOURI MOHAMED CHERIF

SUJET

La factorisation des entiers par la méthode des courbes elliptiques

Soutenu le 25 Mars 2004

Devant le jury composé de :

M^r .M. ABID : Maître de conférences à l'USTHB	**Président**
M^r. K.BETINA : Professeur **à l'USTHB**	**Directeur du mémoire**
M^r .M. ZITOUNI : Professeur à l'USTHB	**Examinateur**
M^r .H. HAMITI : D.G.S.C.T	**Examinateur**

Table des matières

Chapitre 1

Introduction

Les courbes elliptiques permettent l'adoption de nouveaux cryptosystèmes. Mais elles mettent aussi en danger certains cryptosystèmes existants, comme le RSA. Elles sont en effet à l'origine d'une méthode élégante de factorisation, due à H. W. Lenstra en 1985.

Une nouvelle classe de cryptographie à clé publique a récemment émergé. Cette nouvelle famille d'algorithmes est appelée cryptographie ECC (Elliptic curve cryptography). Elle est issue de recherches portant sur des formules qui permettent de déterminer la surface des ellipses. Ces recherches ont mené les mathématiciens à une famille d'équations possédant des propriétés spéciales, qui permirent d'adapter ces équations à la cryptographie.

En ce qui concerne les algorithmes de factorisation d'entiers, il y a trois algorithmes majeurs : le crible quadratique, le crible algébrique, et la méthode des courbes elliptiques.

Les algorithmes de crible quadratique et de crible algébrique recherchent pour factoriser un entier N, de nombreuses congruences, $X \equiv Y \bmod N$ où X et Y sont des carrés ou des produits des nombres premiers issus d'une base de « factorisation », quand un nombre suffisant de telles congruences est obtenu, on peut alors trouver des entiers X et Y tels que $X^2 \equiv Y^2 \bmod N$ sans que $X \equiv \pm Y \bmod N$. On a ensuite un facteur de N avec pgcd(X + Y, N) ou pgcd(X - Y, N). La façon de découvrir les congruences modulo N diffère suivant les deux méthodes, mais elles sont toutes les deux d'une complexité dépendante de la taille de N. le crible algébrique qui ne s'applique qu'aux nombres de la forme $b^n \pm 1$, avec $b \leq 12$.

La méthode des courbes elliptiques, découverte par Hendrik. W. Lenstra en 1985, a une complexité sous exponentielle en la taille de plus petit facteur p de N, elle consiste à rechercher l'ordre d'un élément du groupe fini des points d'une courbe elliptique sur GF(p) grâce à une arithmétique modulo N.

Après avoir rappelé les propriétés essentielles des courbes elliptiques sur un corps K au chapitre 2 et sur le corps $\mathbb{Z}/p\mathbb{Z}$ au chapitre 3, nous décrirons un cryptosystème utilisant les courbes elliptiques au chapitre 4 et puis l'algorithme avec ses différents développements au chapitre 5, et nous serons en mesure de détailler dans le chapitre 6 les

choix algorithmiques faits pour l'implantation MECM de l 'algorithme. Nous parlerons sur les méthodes de crible quadratique et de crible algébrique au chapitre 7. Enfin le chapitre 8 fera mention des factorisations obtenues.

Chapitre 2

Les courbes elliptiques sur un corps K

2-1 : Premières définitions

Définition 1 : *Soit K un corps, pour $m \in N$, on note $m \times 1$ l'élément $1+1+1+...+1$ où 1 figure m fois. On appelle caractéristique de K le nombre $p \in N$ tel que $p \times 1 = 0$.*

Exemple :

- K = P $p \times 1 = 0 \Leftrightarrow p = 0$ i, e la caractéristique de P est 0.
- Soit p un nombre premier, alors la caractéristique du corps Z/pZ est p, car $p \times 1 = 0$

Proposition 1 : *La caractéristique d'un corps K est soit 0, soit un nombre premier.*

Preuve

On suppose que p n'est pas premier, alors $\exists$ m, n $\in N^*$ tels que $1 < m , n < p$ et $p = m \times n$

On a donc $p \times 1 = 0$

Mais $p \times 1 = m \times n \times 1$ et $m \times n \times 1 = 1+1+1\ldots+1$ (1 figure $m \times n$ fois)

Donc $m \times x = 0$ avec $x = n \times 1 \neq 0$ car $(n < p)$

Donc $m \times x \times x^{-1} = px^{-1} = 0$ i, e $m = 0$ contradiction.

Donc p est premier. □

Définition 2 : *Un corps K est dit algébriquement clos si tout polynôme $p(x) \in K[x]$ possède une racine $a \in K$ (i, e $\exists$ $a \in K$ tel que : $p(a) = 0$).*

Exemples

1) X est algébriquement clos.

2) P n'est pas algébriquement clos car x^2+1 ne s'annule en aucun point de P.

Théorème 1: Soit K un corps, il existe un corps algébriquement clos contenant K.

Exemple :

Si K = $\mathbb{P}$, alors L = X.

2-2 : Le plan projectif

Définition 3 : *Soit K un corps, le plan projectif $P_2(K)$ est l'ensemble des points $P = (a,b,c) \neq (0,0,0)$ dans K^3 de sorte que deux points $P = (a,b,c)$ et $P' = (a',b',c')$ sont considérés comme étant des points équivalents s'il existe t dans K^* tel que $P = (a,b,c) = t(a',b',c')$. Les nombres a,b,c sont appelés les coordonnées homogènes du point P. Plus généralement, nous définissons le n-espace projectif $P_n(K)$ comme l'ensemble des classes d'équivalence des (n + 1)-uples suivants :*

$P_n(K) = \{(a_0 ,a_1, a_2, \ldots , a_n) \in K^{n+1} / a_0, a_1, a_2, \ldots , a_n$ *non tous nuls* $\}$ *avec*

$(a_0, a_1, a_2, \ldots , a_n) \approx (a_0', a_1' ,a_2', \ldots ,a_n')$ *s'il existe t dans* K^* *tel que*

$(a_0, a_1, a_2, \ldots , a_n) = t (a_0', a_1' ,a_2', \ldots , a_n')$. *[2].*

Notation

Le plan usuel sur un corps K est appelé plan affine et note $A_2(K)$ l'ensemble des points (x,y) dans K^2; si nous introduisons les coordonnées X, Y, Z tel que x = $\frac{X}{Z}$, y = $\frac{Y}{Z}$ alors à tout point (x,y) $\in$ $A_2(K)$ correspond le point (X,Y,Z) $\in$ $P_2(K)$, réciproquement si $Z \neq 0$ alors à tout point (X,Y,Z) de $P_2(K)$ correspond le point (x,y) de $A_2(K)$. Voyons à présent ce qui se passe quand Z = 0 : considérons dans $A_2(K)$ deux droites parallèles : L : ax + by + c = 0 et L′ :a′x + b′y + c′ = 0 où a′ = ta , b′ = tb, c′ = tc en coordonnées homogènes c.à.d dans $P_2(K)$, ces droites s'écrivent L : aX + bY + Z = 0 et L′ : a′X + b′Y + Z = 0. L'intersection de ces droites au lieu en un point pour lequel Z = 0 ; un tel point est appelé point à l'infini. Cela permet de donner une nouvelle définition de $P_2(K)$.

$P_2(K) = A_2(K) \bigcup$ {l'ensemble des directions dans $A_2(K)$}.[2]

Définition 4 : *Pour tout a, b ∈ K (K corps algébriquement clos de caractéristique ≠ 2,3), on définit une courbe elliptique $E_{a,b}$ comme étant l'ensemble des points (x,y) $\in K^2$ tels que $y^2 = x^3 + ax + b$ auquel on ajoute un point idéalisé θ.*

- θ est appelé l'identité ou le point à l'infini.
- Les points de $E_{a,b}$ distincts de θ sont appelés points finis.[10]

Définition5: *Si k est sous corps de K (i, e est un corps contenu dans K), les points de la courbe elliptique $E_{a,b}$ dont les coordonnées sont dans k sont appelés les points k-rationnels.[10]*

Exemple

- $k = \Theta \subset P = K$

Définition6: *Une courbe elliptique est dite non singulière si le polynôme $x^3 + ax + b$ possède des racines distinctes, sinon elle dite singulière.[10]*

Proposition2: *La courbe elliptique $E_{a,b}$ est non singulière si et seulement si $4a^3 + 27b^2 \neq 0$.[10]*

Preuve

On montre que le polynôme $x^3 + ax + b$ possède des racines distinctes si et seulement si $4a^3 + 27b^2 \neq 0$.

Si $f(x) = x^3 + ax + b = (x - \alpha)^3$ ou $(x - \alpha)^2(x - \beta)$

$f'(x) = 3(x - \alpha)^2$ ou $2(x - \alpha)(x - \beta) + (x - \alpha)^2$

Donc α est une racine double ou triple de f et alors α est une racine de f'(x), il suffit de montrer que $4a^3 + 27b^2 = 0 \Leftrightarrow x^3 + ax + b$ possède une racine double.

On suppose que $x^3 + ax + b$ possède une racine double α, alors ;

$f'(\alpha) = 0 \Leftrightarrow \alpha^2 + a = 0$

$\Leftrightarrow \alpha^2 = \frac{-a}{3}$

$\Leftrightarrow \alpha = \pm\sqrt{\frac{-a}{3}}$

$f(\alpha) = 0 \Leftrightarrow \alpha^3 + a\alpha + b = 0$

$\Leftrightarrow (\sqrt{\frac{-a}{3}})^3 + a\sqrt{\frac{-a}{3}} + b = 0$ et $(-\sqrt{\frac{-a}{3}})^3 - a\sqrt{\frac{-a}{3}} + b = 0$

$\Leftrightarrow \frac{-a}{3}(\sqrt{\frac{-a}{3}}) + a\sqrt{\frac{-a}{3}} + b = 0$ et $-(\frac{-a}{3})(\sqrt{\frac{-a}{3}}) - a\sqrt{\frac{-a}{3}} + b = 0$

$$\Leftrightarrow [\ \frac{-a}{3}(\sqrt{\frac{-a}{3}}) + a\sqrt{\frac{-a}{3}} + b\ ][-(\frac{-a}{3})(\sqrt{\frac{-a}{3}}) - a\sqrt{\frac{-a}{3}} + b\ ] = 0$$

$$\Leftrightarrow \frac{a^3}{27} - \frac{a^3}{9} - \frac{a^3}{9} + \frac{a^3}{3} + b^2 = 0$$

$$\Leftrightarrow \frac{4a^3+27b^2}{27} = 0$$

$$\Leftrightarrow 4a^3 + 27b^2 = 0 \qquad [10]$$

□

Exemple

Si p = 5 : $Z/5Z = \{0, 1, 2, 3, 4\}$

$E_{1,2} : y^2 = x^3 + x + 2$

$4a^3 + 27b^2 = 4 + 2 \times 4 = 2 \neq 0$

Donc cette courbe est non singulière.

$E_{3,1} : y^2 = x^3 + 3x + 1$

$4a^3 + 27b^2 = 4 \times 3^3 + 2 \times 1 = 3 + 2 = 0$

Donc cette courbe est singulière.

Remarque :

Dans la suite on considère uniquement des courbes non singulières.

2-3 Loi de groupe

Tout l'intérêt d'une courbe elliptique pour la factorisation d'entiers c'est de pouvoir la munir d'une loi de groupe.

On va définir la loi de groupe sur une courbe elliptique $E_{a,b}$:

- θ agit comme l'élément neutre :

$P + \theta = \theta + P = P \quad \forall\ P \in E_{a,b}$

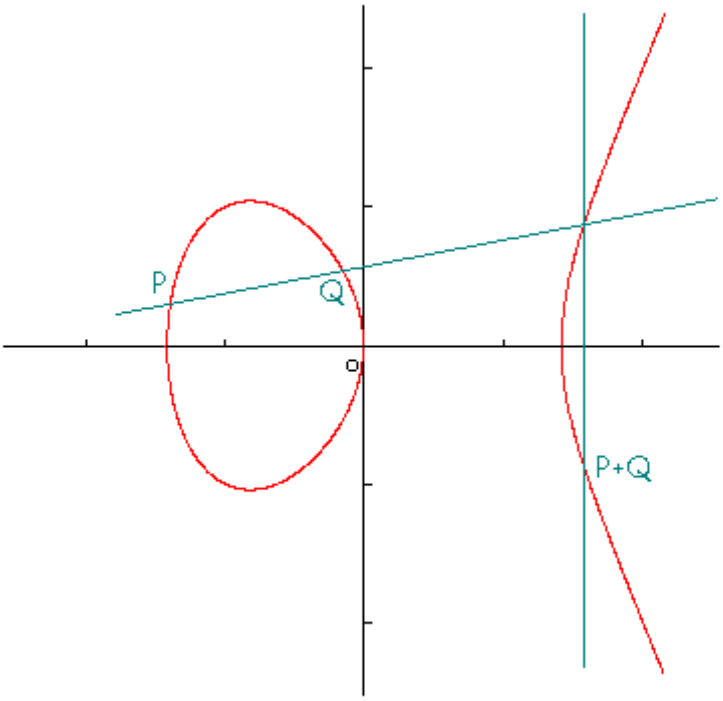
- Si $P = (x,y) \in E_{a,b}$ on pose $-P = (x, -y)$ et on a :

$-P \in E_{a,b}$ car $y^2 = x^3 + ax + b$ et $(-y)^2 = y^2 = x^3 + ax + b$

- Si P, Q et R appartiennent à une droite l : $\alpha x + \beta y + \gamma = 0$

alors $l(P) = l(Q) = l(R) = 0$.

On écrit $P + Q + R = \theta$



Courbe elliptique et opération sur cette courbe

On suppose $P \neq Q$ et $P \neq -Q$.

Soient $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$, $R = (x_R, y_R)$, la droite $l : \alpha x + \beta y + \gamma = 0$ passant par P et Q et R le troisième zéro de l.

On calcule les coordonnées de $R = P + Q$

$P \in l \Leftrightarrow l(P) = m(x_P - x_P) - (y_P - y_P) = 0$

$l : m(x - x_P) - (y - y_P) = 0$

$l : mx - mx_P - y + y_P$ (1)

$l : \alpha x + \beta y + \gamma = 0$ (2)

De (1) et (2) on tire $\alpha = m$, $\beta = -1$, $\gamma = mx_P + y_P$

$Q \in l \Leftrightarrow l(Q) = m(x_Q - x_P) - (y_Q - y_P) = 0$

$$\Leftrightarrow m = \frac{y_Q - y_P}{x_Q - x_P}$$

Si $S = (x,y) \in l$ et $S \in E_{a,b}$ on a :

$y^2 - x^3 - ax - b = 0$ (3)

$m(x - x_P) + y_P = y$ (4)

Donc $[m(x - x_P) + y_P]^2 - x^3 - ax - b = 0$

On pose $f(x) = [m(x - x_P) + y_P]^2 - x^3 - ax - b$, et soit un point

$\gamma = (\alpha, \beta) \in l \bigcap E_{a,b,}$ alors α vérifie $f(\alpha) = 0$.

On a aussi $R = (x_R, y_R) \in l \bigcap E_{a,b}$, donc x_R vérifie l'équation $f(x_R) = 0$.

D'autre part on a $f(x) = -(x - x_P)(x - x_Q)(x - x_R)$.

Le coefficient de x est m^2 dans (3) et $x_P + x_Q + x_R$ dans (4)

Donc $m^2 = x_P + x_Q + x_R$, d'où

$$\boxed{x_R = -x_P - x_Q + m^2}$$

avec $m = \frac{y_Q - y_P}{x_Q - x_P}$,

donc on a $y_R = m(x_R - x_P) + y_P$.

$$\boxed{y_R = y_P + m(x_R - x_P) = y_Q + m(x_R - x_Q)}$$

- Si P = Q, on utilise la droite tangente en P (i, e la droite avec un zéro double en P).

Soit $P = (x_P, y_P)$, x_P doit être une racine double de f(x) i, e $f(x_P) = 0$ et $f'(x_P) = 0$.

$f'(x) = 0 \Leftrightarrow 2m[m(x - x_P) + y_P] - 3x^2 - a = 0$

$\Leftrightarrow 2m(y) - 3x^2 - a = 0$ et

$f'(x_P) = 0 \Leftrightarrow 2m(y_P) - 3x^2 - a = 0$

$\Leftrightarrow m = \frac{3x_P^2 + a}{2y_P}$ [10] □

Remarque : Si on note $\omega_1, \omega_2, \omega_3$ les trois racines distinctes du polynôme $x^3 + ax + b$, les points qui vérifient $2P = \theta$ sont θ et $(\omega_1,0), (\omega_2,0), (\omega_3,0)$ car si $P = (x_P,y_P)$ alors :

$2P = \theta \Leftrightarrow P + P = \theta$

$\Leftrightarrow P = -P$

$\Leftrightarrow (x_P, y_P) = (x_P, -y_P)$

$\Leftrightarrow y_P = -y_P$

$\Leftrightarrow 2y_P = 0$

$\Leftrightarrow y_P = 0$

$\Leftrightarrow x^3 + ax + b = 0$

$\Leftrightarrow x_i = \omega_i \; i = 1, 2, 3$ [10]

Théoreme1 : La courbe elliptique $E_{a,b}$ munie de l'addition définit ci-dessus est un groupe abélien.[10] □

Exemple1 : (l'addition de deux points distincts)

Soit $E_{-7,0}$ une courbe elliptique sur le corps P , et P = (-2.35,-1.86), Q = (-0.1, 0.836), on calcule les coordonnées du point R qui est la somme de P et Q.

On a $m = \frac{y_Q - y_P}{x_Q - x_P} = \frac{0.836+1.86}{-0.1+2.35}$

$= \frac{2.696}{2.25}$

$= 1.1982$

Donc $x_R = - x_P - x_Q + m^2$

$= 2.35 + 0.1 + 1.1982^2$

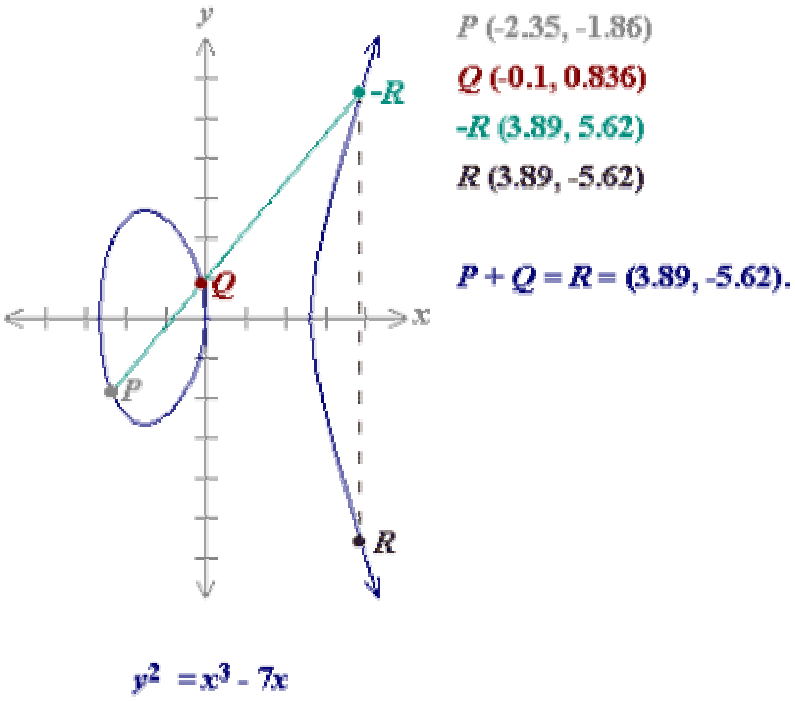
$= 3.89$

$y_R = m(x_R - x_P) + y_P$

$= 1.1982 (3.89 + 2.35) - 1.86$

$= 5.62$

Donc $P + Q = (3.89, - 5.62)$



[11-b]

Exemple2 (doublement d'un point)

Soient la courbe elliptique $E_{-3,5}$ sur le corps P et le point P = (2,2.65) de $E_{-3,5}$, on calcule les coordonnées du point R qui est la somme de P + P.

On pose R = P + P = 2P

On a $m = \frac{3x_P^2+a}{2y_P}$

$= \frac{3\times 2^2-3}{2\times 2.65}$

$= \frac{9}{5.3}$

$= 1.6981$

donc $x_R = -2x_P + m^2$

$= -2(2) + 1.6981^2$

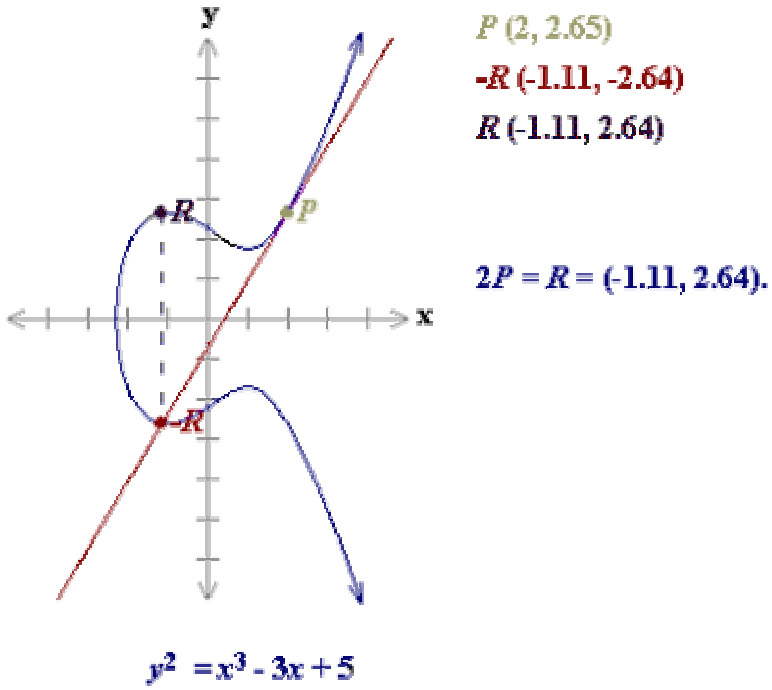
$= -1.11$

$y_R = m(x_R - x_P) + y_P$

$= 1.6981(-1.11-2) + 2.65$

$= -2.64$

Donc R = P + P = 2P = (- 1.11, - 2.64)



[11-c]

Calcul la somme de P + P si $y_P = 0$

Si P est un point tel que $y_P = 0$, alors la tangente à la courbe elliptique à P est verticale et n'intersecte pas la courbe elliptique dans n'importe quel autre point.

Par définition, 2P = θ pour un tel point P.

Si on voulait trouver 3P dans cette situation, on peut s'ajouter 2P + P, et devient P + θ = P ainsi 3P = P.

3P = P, 4P = θ, 5P = P, 7P = P, …etc

Donc $\forall \in N^*$ nP = $\begin{cases} p & \text{si n est impair} \\ 0 & \text{si n est pair} \end{cases}$

Chapitre 3

Les courbes elliptiques sur Z/pZ

3-1 Les courbes elliptiques sur Z /pZ

Définition7 : Une courbe elliptique est l'ensemble des points $(x,y) \in (Z/pZ)^2$ tel que $y^2 = x^3 + ax + b$. L'ensemble des points qui lui sont associés est alors : $E_{a,b}(Z/pZ) = \{(x:y:z) \in P_2(Z/pZ) : y^2z = x^3 + axz^2 + bz^3\}$.

Un seul point correspond à z = 0, ils 'agit de θ = (0,1,0) et chaque classe différente de θ a unique représentant (x :y :1). On considèrera donc dans la suite que $E_{a,b}(Z/pZ)$ est la réunion de θ (vu comme l'infini) avec l'ensemble $E_{a,b} = \{(x,y) \in (Z/pZ)^2, y^2 = x^3 + ax + b\}$ [2].

Exemple : (Z /23Z)

On considère la courbe elliptique $E_{a,b}$ avec a = 1, b = 0, alors l'équation de la courbe elliptique est : $y^2 = x^3 + x$.

Le point (9,5) satisfait cette équation car :

$y^2 \bmod 23 = x^3 \bmod 23 + x \bmod 23$

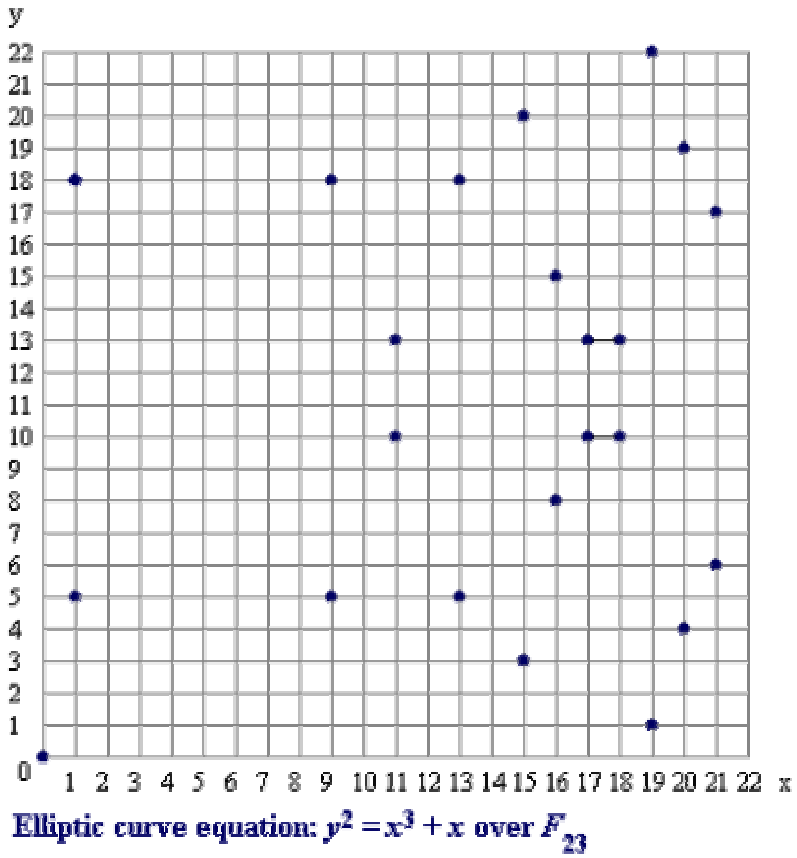
$\Leftrightarrow 25 \bmod 23 = 729 \bmod 23 + 9 \bmod 23$

$\Leftrightarrow 2 = 2$

Les points qui satisfont cette équation sont:

(0,0) (1,5) (1,18) (9,5) (9,18) (11,10) (11,13) (13,5)
(13,18) (15,3) (15,20) (16,8) (16,15) (17,10) (17,13) (18,10)
(18,13) (19,1) (19,22) (20,4) (20,19) (21,6) (21,17)

Ces points peuvent être représentes graphiquement :



[11.d]

Théoreme3 :

Le nombre de courbes elliptiques sur Z/pZ vaut p(p-1) .

Preuve

Le nombre de courbes elliptiques sur Z/pZ est le nombre de couples (a,b) $\in (Z/pZ)^2$ tels que $4a^3 + 27b^2 \neq 0 \bmod p$, le nombre de toutes les paires vaut p^2 et $4a^3 + 27b^2 \neq 0 \bmod p$ si et seulement si $a = -3c^2$, $b = 2c^3$ pour $c \in Z/pZ$, chaque c étant de plus déterminé de manière unique par $c \equiv \frac{-3b}{2a} \bmod p$ $(a \neq 0)$, le nombre de couples (a,b) tels que $4a^3 + 27b^2 \neq 0 \bmod p$ vaut p et celui des courbes elliptiques vaut $p^2 - p = p(p-1)$.

□

Théoreme4 : (Hasse)

L'ordre d'une courbe elliptique $E_{a,b}$ sur Z/pZ vaut $p + 1 - t$ avec $|t| < 2\sqrt{p}$ et réciproquement pour tout $m \in [(\sqrt{p}-1)^2, (\sqrt{p}+1)^2]$, il existe une courbe elliptique d'ordre m.

3-2- Généralisation à N non premier

Dans le cas où N est composé (non premier), on peut encore définir une loi de groupe pour une courbe elliptique $E_{a,b}$ sur Z/pZ mais vu la complexité de la théorie des courbes elliptiques définies sur anneau, on préfère généralement définir à moindre coût une *pseudo-addition*, avec les mêmes formules que les formules précédentes.

Cette *pseudo-addition* n'est cependant possible que si le nombre w à inverser lors du calcul de λ pendant la *pseudo-addition* de deux points $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$ est premier avec N.

Ainsi lorsque qu'une addition sur une courbe modulo N est impossible, les deux points à sommer sont opposés de N en calculant pgcd(w, N), c'est ce résultat qui est la base de l'algorithme de Lenstra.[1].

Chapitre 4

Cryptosystème utilisant les courbes elliptiques

Parallèlement à leur découverte de la cryptographie à clé publique, Diffie et Hellman ont proposé un protocole d'échange de clés totalement sécurisé. Alice et Bob se sont mis d'accord sur un algorithme à clé secrète à utiliser, ils veulent s'échanger une clé k, mais ils ne disposent pas de canal fiable pour cela. Diffie et Hellman suggèrent l'échange suivant :

Soient E(a, b, K) une courbe elliptique c'est à dire on choisit sur le corps K ($\mathbb{Z}/p\mathbb{Z}$ par exemple) la courbe elliptique de l'équation : $y^2 = x^3 + ax + b$, et $P \in E(a, b, K)$ et k est un nombre aléatoire dans N^*, nous calculons $k \times P$. Le résultat de cette multiplication est un autre point Q de la courbe. Si nous publions les points P et Q, personne n'a encore pu trouver k en un temps raisonnable. Si notre module est suffisamment élevé. Les superordinateurs les plus rapides au monde calculeraient durant des milliers d'années avant de trouver k .[8].

4-1 L'algorithme EC Diffie-Hellman

Une façon de faire de la cryptographie en utilisant des courbes elliptiques à appliquer le même principe à l'algorithme de Diffie-Hellman qui utilise la congruence modulaire . De fait, le nom accepté pour cette méthode est *EC Diffie-Hellman* .

Alice dispose de sa courbe elliptique E(a, b ,K) , deux points. Elle dispose aussi d'un nombre k, mais le garde secret . Son nombre k n'est qu'un nombre aléatoire.

Clé publique d'Alice : E(a, b, K), P , Q_A

Clé privée d'Alice : k_A

Bob reçoit la clé publique d'Alice et calcule la sienne. Il s'agit simplement d'une valeur aléatoire. Il trouve ensuite un point Q_B en multipliant $k_B P$.

Clé publique de Bob : E(a, b, K), P, Q_B

Clé privée de Bob : k_B

Bob multiplie alors son nombre k_B par le nombre Q_A d'Alice.

Valeur secrète : $S = k_B Q_A$

Mais a quoi est égal le nombre Q_A ? Il est égal à k_AP. Ce que Bob donc effectivement fait, c'est calculer

Valeur secrète : $S = k_B(k_AP) = (k_Bk_A)P$

Il ne sait jamais à quoi est égal k_A, mais effectivement des calculs à partir de Q_A revient mathématiquement à effectuer des calculs à partir de k_AP.

Bob envoie maintenant à Alice son nombre Q_B. Et effectue à partir de ce nombre Q_B.

Valeur secrète : $S = k_AQ_B$

Mais a quoi est égal le nombre Q_B ? Il est égal à k_BP. Donc, ce qu'Alice a effectivement fait , c'est calculer

$S = k_A(k_BP) = (k_Bk_A)P$

C'est le même nombre que celui calculé par Bob.

Il s'agit d'un échange de clé à la manière de Diffie-Hellman c'est à dire sans se les communiquer directement. Alice et Bob se mettent d'accord ensemble est publiquement sur une courbe elliptique E(a, b ,K) c'est à dire qu'ils choisissent un corps fini K(Z/pZ par exemple) et une courbe elliptique $y^2 = x^3 + ax + b$. Ils se mettent aussi d'accord sur un point P situé sur la courbe.

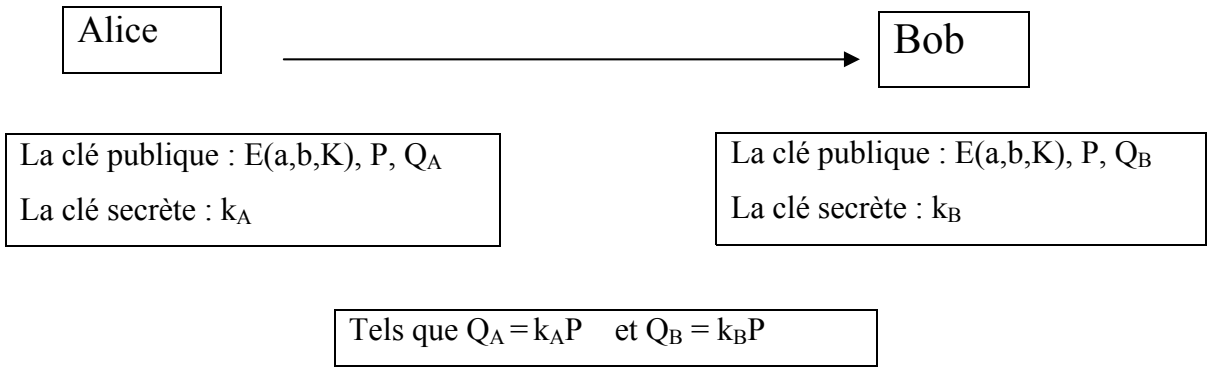
Bob peut employer la valeur secrète qu'il a calculée pour créer une clé de chiffrement. Quand Alice reçoit le message de Bob, elle emploie la valeur secrète qu'elle a calculée pour créer la clé de déchiffrement. Comme il s'agit de la même valeur secrète, ils emploient donc la même clé. Ce qui chiffre Bob, Alice peut le déchiffrer.[3]

4-2 Sécurité

Si quelqu'un a espionné leurs échanges, il connaît E(a, b, k), P, k_A, et k_B. Pour pouvoir calculer $S = k_Ak_BP$, il faut pouvoir calculer k_A. Connaissant P et k_AP. Personne n'a jamais trouvé le moyen de calculer S autrement. C'est ce que l'on appelle résoudre le logarithme discret sur la courbe elliptique(c'est le même type de problèmes, avec des notations additives, que de retrouver n dans une équation $y = x^n$ dans Z/pZ , y et x sont connus). Le logarithme discret est déjà difficile à résoudre dans les groupes bien connus $(Z/pZ)^*$. Pour les groupes plus compliqués et très différents les uns des autres, des courbes elliptiques c'est encore plus difficile. Mais Alice et Bob gardent secret leur nombre k. l'espion en est donc pour se faire. Mais il pourra peut-être déterminer quand même l'un des

nombres k. Par exemple, il sait que $Q_A=k_AP$ et il connaît Q_A et P. Il lui suffit donc de calculer l'équation pour k. Mais il s'agit de l'un des épineux problèmes dont nous avons parlé. Personne n'a pas pu le résoudre en un temps raisonnable. Soit dit en passant, on appelle ce problème *problème du logarithme discret à courbe elliptique* [3]

4-3 Transmission de messages



On suppose qu'Alice et Bob ont suivi le protocole d'échange de clé expliqué, ci dessus. Alice veut envoyer à Bob un message, ils se sont mis d'accord sur la façon de transformer un texte en suite de points de la courbe elliptique. Alice doit donc transmettre, de façon secrète, un point M de la courbe E(a, b , K). Elle choisit (secrètement) un nombre l et envoie à Bob le couple (lP, $M + lk_BP$) .

Bob, lui multiplie lP par k_B (sa clé secrète), puis retranche lk_BP à $M+lk_BP$, il retrouve M. Si quelqu'un espionne les échanges, il lui faut absolument connaître k_B pour retrouver M, c'est encore une fois le problème du logarithme discret.

Nous avons passé sous silence une des difficultés majeures, l'algorithme pour transformer un texte en points de la courbe elliptique est loin d'être trivial, il n'est pas toujours facile de trouver un point sur la courbe elliptique.[8]

Chapitre 5

Factorisation sur une courbe elliptique

5-1 Un précurseur : la méthode p-1 de pollard

Cette méthode est du à J.W.Pollard repose sur le petit théorème de Fermat.

Théoreme5 : Si p est un nombre premier alors : $\forall a \in (Z/pZ)^*$, $a^{p-1} \equiv 1 \bmod p$.

D'après ce théorème, tout facteur premier p d'un entier N, divise $pgcd(a^{p-1},N)$, donc on trouve un entier N multiple de p-1 sans être pour autant un multiple de $\varphi(N)$ (φ étant l'indicatrice d'Euler) alors, $pgcd(a^r - 1,N)$ sera un facteur non trivial de N.[1]

```
Procedure pollard (a, b, N)
# Next Prime(p) retourne le nombre premier
# Immédiatement supérieur à p.
 p := Next Prime(1).
 g := 1.
While  p < b  and    g = 1    do
 k := p
While  kp < b        do  k = kp
  a = a^k  mod N
  g = pgcd(a-1, N)
  p = NextPrime(p)
end
return(g)
end pollard
```

L'algorithme p-1 de pollard

L'algorithme p - 1 de pollard repose sur ce constat en choisissant pour r des produits de petits nombres premiers.

Exemple :

Pour factoriser N = 143, avec a = 2 et b = 5, on calcule d'abord :

$a = 2^4 \bmod 143 = 16$ et pgcd(a - 1, N) = pgcd(15,143) = 1

$a = 16^3 \bmod 143 = 92$ et pgcd(a - 1, N) = pgcd(91,143) = 13

Donc $143 = 13 \times 11$

La méthode a été ici un succès car $13 - 1 = 12 = 2^2 \times 3$.

Dans le cas général, elle réussit si au moins un facteur de p - 1 est petit (p - 1 est dit ***friable***)

5-2 L'algorithme de Lenstra

La méthode des courbes elliptiques découverte par H.W.Lenstra, on suppose que l'on doit factoriser un entier N (donné) et on note p un diviseur premier de N (que l'on ne connaît pas), on travaille sur une courbe elliptique $E_{a,b}$ (choisie au hasard sur Z/pZ).

On prend un point P sur la courbe elliptique $E_{a,b}$ aléatoire, et une borne $B_1 \in N^*$, fixé par l'utilisateur.

On choisit un entier k pas très grand mais qui est produit de petits facteurs premiers inférieurs à B_1, on cherche l'ordre du point P sur la courbe $E_{a,b}$ modulo l'un des facteurs premiers p de N .

On calcule les coordonnées du point kP en utilisant les formules classiques, les calculs s'effectuant modulo N, ils font intervenir des divisions et ceci n'est pas toujours possible modulo N. Il faut que le dénominateur d soit premier avec N mais on espère que ce n'est pas le cas, en effet si d n'est pas premier avec N, pgcd(d,N) donne un diviseur premier de N si on a pu mener les calculs jusqu'au bout. On recommence où est l'étape 1, en changeant de courbe elliptique.[1]

Procedure ECM (P, $E_{a,b}$, B_1, n)

p := NextPrime(1)

g := 1

While p < B_1 **and** g = 1 **do**

k := p

While kp < B_1 **do** kkp **end**

Ec mod Mult(k, P, $E_{a,b}$, g, n) calcul kP mod n sur $E_{a,b}$

Si une inversion est impossible g est un facteur de n.

P :=EC mod Mult(k,P, $E_{a,b}$, g, n)

P := NextPrime

End

return(g)

End ECM

L'algorithme de Lenstra

Exemple

Pour factoriser 143, considérons $B_1 = 4$, et la courbe elliptique

$E_{1,1}$:$y^2 = x^3 + x + 1$ modulo 143 et le point P = (0,1) et on applique l'algorithme ainsi ;

*** Première étape :**

On calcule R = 2P = P + P

R = (0,1) + (0,1) = (x_R, y_R)

On a : $m = \frac{3x_P^2+a}{2y_P}$ et $x_R = -2x_P + m^2$, $y_R = m(x_R - x_P) + y_P$

Donc $m = \frac{1}{2} \mod 143 = 72 \mod 143$.

$x_R = -2(0) + 72^2 \mod 143 = 36$

$y_R = -1 + 72(0 - 36) = 124$

Donc R = (36,124)

***Seconde étape :**

On calcule $R = 3 \times 2P$

R = 2(2P) + 2P

On a : 2(2P) = 2P + 2P

= (36,124) + (36,124)

= (127,71)

Donc R = (127,71) + (36,124)

On calcule R = (127,71) + (36,124)

$m = \frac{124-71}{36-127} = \frac{-53}{91}$

Mais 91 ne possède pas un inverse modulo 143

Don le calcul de ces points est impossible et alors on obtient un facteur de 143, en calculant pgcd(91,143) = 13

Donc $143 = 13 \times 11$

5-3- Complexité de l'algorithme

On reprend ici l'analyse de la complexité faite par P.R.Brend, la courbe et le point initial considérés sont choisis arbitrairement. On suppose connu les résultats suivants :

Définition 8 :

Soit M > 0, et $n_1 \geq n_2 \geq$ …..Les facteurs premiers d'un entier M, on pose :

$\rho(\alpha)$ = Lim prob ($n_1 < M^{\frac{1}{\alpha}}$) si $M \to \infty$

On peut alors montrer que ρ vérifié l'équation :

$\alpha \rho'(\alpha) + \rho(\alpha - 1) = 0$ (3.1)

et encore le résultat asymptotique :

$\log \rho(\alpha) = \alpha (\log \alpha + \log(\alpha - 1)) - \theta(\alpha)$ quand $\alpha \to \infty$ (3.2)

qui mène à :

$\frac{\rho(\alpha-1)}{\rho(\alpha)} = \alpha (\log\alpha + \theta(\log\log\alpha))$ quand $\alpha \to \infty$ (3.3)

Dans l'algorithme de lenstra le coût essentiel est du aux multiplications et aux inversions mod N, on néglige les additions modulo N et les multiplications /divisions sur un des petits entiers, on prendra donc le coût d'une multiplication, comme unité de base (ub) et K(ub), celui d'une inversion.

Les méthodes usuelles du calcul de kP, pour k entier ont pour coût c log k ub avec $c \in R^+$, sachant que le nombre d'entiers premiers inférieurs à x est asymptotiquement égal à $\pi(x) = \frac{x}{\log x}$, un essai complet testant tous les nombres premiers inférieurs à B_1 nécessite donc cB_1 (ub)

Supposons que les résultats de la définition 4 s'applique aussi pour M non infini (hypothèse raisonnable), la probabilité qu'un essai réussisse à trouver un facteur p vaut $\rho(\alpha)$ avec $\alpha = \frac{\log p}{\log B_1}$ et le travail nécessaire est alors

$$W(\alpha) \approx \frac{cp^{\frac{1}{\alpha}}}{p(\alpha)} \qquad (3.4)$$

On cherche à minimiser W en différenciant (3.4) et en annulant le résultat, On obtient alors $\log p = \frac{-\alpha^2 \rho'(\alpha)}{\rho(\alpha)}$, ou encore d'après (3.1),

$$\text{Log } p = \frac{\alpha\rho(\alpha-1)}{\rho(\alpha)} \qquad (3.5)$$

En supposant p connu, cherchons α tel que (3.5) soit satisfait, on a d'après (3.3)

$$\text{Log } p = \alpha^2(\log \alpha + o(\log\log \alpha)) \qquad (3.6)$$

D'où

$$\alpha \approx \frac{2\log p}{\sqrt{\log\log p}} \qquad (3.7)$$

Comme d'autre part

$$\log W(\alpha) \approx \log c + \frac{\rho(\alpha-1)}{\rho(\alpha)} - \log \rho(\alpha) \approx \log c + 2\alpha \log \alpha \qquad (3.8)$$

On obtient finalement

$$W(p)=c\, e^{1+\theta(1)\sqrt{2\log p \log\log p}}$$

On remarque que:

Cette complexité signifie que le temps nécessaire à l'algorithme pour trouver le facteur p de N est égal au temps mis pour réaliser W(p) multiplication modulo N, elle dépend de la taille de p et non pas de N comme celle du crible quadratique, par exemple. W(p)étant sous exponentiel, on comprend ici l'importance d'optimiser au maximum les opérations sur une courbes elliptiques afin d'améliorer la constante multiplicative c .[1]

5-4 Courbes elliptiques avec sous groupe de torsion d'ordre 16

Si on pouvait avoir des courbes d'ordre suffisamment friable, la méthode serait fortement accéléré, c'est pourquoi, H.syma ou encore A.O.L.bkin et F morain, ont proposé des courbes contenant des sous groupes de torsions d'ordre connus.

Par théorème de MAZUR, on sait que les seules groupes de torsion sur Θ sont isomorphes

avec $\begin{cases} Z/mZ & m = 1,2,3,\ldots,12 \quad \text{où} \\ (Z/2Z) \times (Z/2mZ) & m=1,2,3,4. \end{cases}$

Je ne rappellerai ici que des résultats sans démonstrations pour des courbes avec un sous groupe de torsion d'ordre 16.

Toutes les courbes de Kubert est données par :

$\varepsilon(b,c)$: $y^2 + (1 - c)xy - by = x^3 - bx^2$ avec $\varepsilon(b,c) \in Q$ tel que

$16b^5 - (-8c^2 + 20c-1)b^4 - c(1-c^3)b^3 \neq 0$.

Elle a alors pour point d'ordre fini maximal le point P = (0,0). ou a en outre le théorème suivant.

Théoreme6 : Pour b et c donnés par $d = \frac{4\alpha(\alpha+1)}{8\alpha^2-1}$, $c = \frac{(2\alpha-1)(\alpha-1)}{d}$, $b = (2d - 1)(d - 1)$,

$\varepsilon(b, c)$ possède un sous groupe de torsion d'ordre 16, si de plus $\alpha = [\frac{t+25}{(s-9)+1}]^{-1}$, où (s,t)

est un point de la courbe : $t^2 = s^3 - 8s - 32$ (3.9)

La courbe $\varepsilon(b,c)$ a un point rationnel d'abscisse $\frac{1-2d}{4}$.

Quand on sait que le point P = (12,40) est un point d'ordre infini de la courbe(3.9), chaque multiple de P permet quasiment d'obtenir une courbe $\varepsilon(b,c)$ différente.

Il suffit alors transformer ces courbes

- Soit en notation de Weierstrass :

Théoreme7 : L'équation en notation de Weierstrass provient de la normalisation de la

courbe : $y^2 = x^3 + \frac{(c-1)^2-4b}{4}x^2 + \frac{b(c-1)}{2}x + \frac{b^2}{4}$ (3.10)

Obtenue après le changement de variables :

$\begin{cases} X = x \\ Y = y + \frac{b-(1-c)x}{2} \end{cases}$

Le point d'abscisse $\frac{1-2d}{4y}$ est alors rationnel.

- Soit en notation de Montgomery (courbe de la forme $by^2 = x^3 + ax^2 + x$ avec $(a,b) \in Q^2$, $b \neq 0$, $a \neq \pm 2$)

□

Théoreme8 : On obtient en notation de Montgomery

$$y^2 = x^3 + [\frac{(2d-1)^2}{2d(d-1)} - 2]x^2 - x$$

En utilisant le changement de variables :

$X = (2d - 1)^2x + x(d - 1)$ et $Y = (d - 1)^3y - \frac{(1-c)X}{2} + \frac{b}{2}$

Le point d'abscisse $\frac{(d-1)(-2d^2+7d-1)}{4}$ est alors rationnel.

□

5-5 Ajouter une seconde phase

Peter.L.Montgomery et Richard P.Brent se sont très vite rendus compte que l'on pouvait, tout comme dans la méthode p - 1, stopper la première phase plus tôt, on obtient aussi un point $R = kP \neq 0 \bmod p$ qui va servir de base à une seconde phase.

En effet, si $k = \Pi_{pi=2}^{pi<B} p_i^{e_i}$, on peut raisonnablement penser que l'ordre de P divise kq où q est un entier premier plus grand que B_1. Les secondes phases connues à ce jour reposent sur cette hypothèse et se concentrent sur la découverte de q.

Deux secondes phases sont décrites ici :

*La seconde phase « classique ».

*La seconde phase « paradoxe des anniversaires ». [1]

5-5-1 La seconde phase « classique »

Cette seconde phase classique consiste tout simplement à calculer sR pour tous les nombres premiers s compris entre B_1 et B_2 avec $B_2 \geq B_1$, on peut cependant éviter de calculer sR par multiplication comme dans la première phase. Le calcul se fait par récurrence en remarquant que si s_j et s_{j+1} sont deux entiers premiers successifs, on a $s_{j+1}R = s_jR + (s_{j+1} - s_j)R$ étant donné que la différence $s_{j+1} - s_j$ est toujours petite.

Procedure :classical step2(k,$E_{a,b}$,B_1,B_2,N)

g := 1

MC Mod add (P, Q, $E_{a,b}$, g, N), calcul P + Q mod N sur $E_{a,b}$

#Si l'inversion est impossible, g est un facteur de N

P_1 :=EC Mod add (R, R, $E_{a,b}$, g, N)

For k = 2,3,….100 **While** g = 1 **do**

P_k = EC Mod add (P_{k-1}, P_1, $E_{a,b}$, g, N)

End

If g := 1 **Then**

p = Next Prime(B_1)

K = EC Mod Mult (p, R, $E_{a,b}$, g, N)

End

s := Next Prime(p)

While s < B_1 **and** g = 1 **do**

Ec Mod add(R, $P_{(s-p)/2}$, $E_{a,b}$, g, n)

P = s

s = Next Prime(p)

End

Return(g)

End classical step2

L'algorithme de la seconde phase classique

Cette seconde phase nécessite donc :

- $\theta(\log B_2)$ opérations pour calculer les points P_i.
- $\theta(\log B_1)$ opérations pour calculer le point NextPrime (B_1)k.

- $\theta(\frac{B_2}{\log B_2} - \frac{B_1}{\log B_1})$ opérations pour tester tous les nombres premiers entre B_1 et B_2.

On a ainsi un coût asymptotique de $\theta(\frac{B_2}{\log B_2})$. C'est cette seconde phase qui fut historiquement utilisée avec l'algorithme; le plus grand facteur découvert avec celle-ci est un entier de 40 chiffres par Ayen k. [1]

5-5-2 La seconde phase « Paradoxe des anniversaires »

A la fin de la première phase, le point R obtenu engendre un sous groupe cyclique <R> de la courbe $E_{a,b}$ supposé avoir modulo p (p; plus petit facteur de l'entier N à factoriser) un ordre q premier assez grand ($\approx 10^9$). Plutôt que de tester $o(\frac{q}{\log q})$ points s_iR comme dans la seconde phase classique, l'idée est ici de rechercher si parmi $r=o(\sqrt{q})$ points aléatoires $R_0,R_1,\ldots,R_{r-1}$ de <R> calculés modulo N, deux points seraient égaux modulo p.

Cette idée était déjà celle sous-jacente à la méthode de factorisation rho-pollard. Dans cet algorithme on calcule à partir de $a \in (Z/pZ)^* \setminus \{1\}$ la suite définir par :

$x_0=\alpha$ et $\forall i \in N^*$, $x_{i+1}=x_i^! + 1 \bmod N$.

Cette suite considérée sur Z/pZ étant pseudo aléatoire, on recherche l'indice i tel, que $x_{2i} \equiv x_i \bmod p$, on peut alors montrer que l'ordre de grandeur de i est alors $\sqrt{p}$.

Dans le cas des courbes elliptiques, un tel schéma n'est pas réalisable tout simplement parce que seules l'addition de points et la multiplication par des entiers sont définies. La multiplication de deux entiers n'y pas de sens. En outre on ne connaît pas de fonction déterministe pseudo aléatoire de <R> dans <R>.

Un moindre mal peut alors être la génération proposée par R.P.Brent :

$R_0 =$ et $\forall j \in \{1, 2, 3, \ldots, r-1\}$ $R_{j+1} = \begin{cases} 2R_j & \text{avec probabilité } \frac{1}{2} \\ 2R_j+R & \text{avec probabilité } \frac{1}{2} \end{cases}$

P.L .Montgomery propose pour sa part cette autre génération de points R_i : $R_0 = R$ et

$\forall$ j $\in$ {1, 2, 3, …, r-1 }, $R_j = p(m_j)R$. Où $\{m_j\}_{j=1}^{r-1}$ est un sous ensemble de N et p(X) est un polynôme à coefficients entiers tel que p(X) ± p(Y) ait beaucoup de petits diviseurs. Un tel polynôme peut être par exemple $p(X) = X^k$ avec k très friable car $(p(X) - p(Y))(p(X) + p(Y)) = X^{2k} – Y^{2k}$ a autant de facteurs polynomiaux irréductibles que 2k possède de diviseurs entiers. Dans ces deux suites de points, rechercher s'il existe un indice i tel que $R_{2i} = R_i \bmod p$ est inutile car ici l'existence de deux indices i et j tels que $R_i = R_j$ n'implique plus $R_{i+1} = R_{j+1}$. On doit donc comparer deux à deux les points générés au prix de r(r-1)/2 opérations. En comparant seulement les abscisses de ces points on augmente alors sensiblement la probabilité de succès car non seulement on recherche deux indices i et j tels que $R_i = R_j$, mais aussi tels que $R_i = - R_j$ (l'inverse d'un point P = (x,y) sur $E_{a,b}$ et le point – P=(x,-y)). La difficulté est alors de trouver si deux valeurs parmi les abscisses $x_0,x_1,\ldots,x_{r-1}$ modulo N sont égal modulo p. Les techniques usuelles de hachages ou de tris sont ici caduques et on doit se résoudre à calculer la quantité :

$$d = \prod_{0 \le i \le j < r} (x_i - x_j) \bmod N \qquad (3.12)$$

Puis pgcd(d, N) en espérant retrouver p. Avec cette méthode, on teste non seulement pendant leurs générations si les r points R_i sont nuls modulo p mais aussi s'il en est de même pour les r(r-1)/2 autres points $R_i \pm R_j$. Au total, plus de multiples de R sont calculés ici que pendant la seconde phase classique même si chaque point n'est pas forcément le produit de R par un nombre premier.

```
Procedure Brith Day Step 2(R, Ea,b, r ,N)
  g := 1
  Q := R
  # Absc (Q) retourne l'abscisse du point Q
  x0 := Absc (Q)
for  k:=1,2,3.....r-1    While    g=1     do
   Q := Ec Mod Add(Q, Q, Ea,b, g, N)
   # Random () retourne aléatoirement 0ou1.
   If  Random() = 1    and        g = 1     then
        Q := Ec Mod Add (Q, R, Ea,b, g, N)
    end
     xk := Absc(Q)
   end
if     g = 1   then
  d := 1
    for k := 0, 1, …, r-2     do
      for l := k, k+1, …, r-1  do
          d := d(xk - xl) mod N
    end
   end
g := pgcd(d, N)
end
return(g)
end Brith day Step 2
```

Algorithme:Le paradoxe des anniversaires.

Dans une variante de cet algorithme, on considère deux familles de points $\{R_i\}_{i=0}^{r_1-1}$ et $\{S_j\}_{j=0}^{r_2-1}$ et on recherche s'il existe deux indices i et j tels que $R_i = \pm S_j$ ($0 \le i < r_1, 0 \le j < r_2$)

Pour se faire, on calcule une quantité analogue à (3.12) :

$$\text{Pgcd}\Big(\prod_{0\le i<r_1}\prod_{0\le j<r_2}(x_i^r - x_j^s), N\Big) \qquad (3.13)$$

Où x_i^r est l'abscisse du point R_i et x_j^s est l'abscisse du oint S_j .

Dans la version principale, la probabilité de succès est alors la même que celle de deux personnes parmi r aient la même date anniversaire sur une planète q jours par an.

En général pour r<<q, cette probabilité vaut : $P_r = 1 - \prod_{j=1}^{r-1}(1-\frac{j}{q}) \approx 1 - e^{\frac{-r^2}{2q}}$.

Lorsqu'on considère que les abscisses, la probabilité de succès devient alors :

$P_r \approx 1 - e^{\frac{-r^2}{q}}$. et celle-ci est supérieure à $\frac{1}{2}$ pour $r > \sqrt{\log(1)q}$

Ainsi la réalisation naïve d'une seconde phase coûte :

1- $0(\sqrt{q})$ opérations pour générer les points R_i.

2- $0(\sqrt{q})$ opérations pour comparer les $\sqrt{q}$ abscisses obtenues.

Il est cependant possible grâce à une arithmétique polynomiale asymptotiquement rapide de ramener ce deuxième coûte à $0(\sqrt{q}\,(\log q)^2)$ opérations.

C'est avec cette seconde phase que le facteur de 43 chiffres fut découvert.[1]

Chapitre 6

Implantation de l'algorithme

6-1 Implantation de la première phase

Deux points essentiels sont optimisé dans une implantation efficace de l'algorithme de Lenstra : l'addition de deux points sur une courbe elliptique et la multiplication d'un point par un entier k. La première est proposée par P.L.Montgomery, elle consiste à utiliser des courbes de la forme $By^2 = x^3 + Ax^2 + x$ pour $(A, B) \in (Z/pZ)^2$, $B \neq 0$, $A \neq \pm 2$, les additions y sont alors moins coûteuses car l'inversion y est remplacée par des multiplications. La seconde consiste à utiliser l courbes et à effectuer l additions en même temps, remplaçant par la même l inversions (dont le coût est prohibitif) par une seule et quelques multiplications.

Quand à la multiplication, la méthode binaire pour la réaliser est ici d'autant plus efficace qu'il est possible de compacter les nombres premiers 2 par 2 afin d'en diminuer le poids, face à cette astuce, des méthodes de coût théoriquement moindre comme la multiplication par blocs, ou par chaînes d'additions perdent leur attrait et ne seront pas décrites.[1]

6-1-1 Choix de l'addition sur les courbes elliptiques

L'addition naïve

Pour des courbes données avec une parametrisation de Weierstrass , les formules de l'addition se programment sans difficulté sur Z/pZ en tenant compte de la découverte d'éventuels facteurs de N lors de l'inversion, la procédure ECModAdd() est un exemple d'une telle pseudo-addition.

```
Procedure   Ec Mod Add ( P, Q, E_{a,b}, g, N )
 g := 1
If  P = θ then return (Q) end
If  Q = θ then return(P) end
If  P = Q then
   #  ord (P) retourne l'ordonnée du point P.
   if  ord (P) = 0 then return (θ) end
   g := pgcd( 2ord (P) mod N, N )
   if g = 1 then
     λ := ( [ 3 Absc (P)^2 + 1 ]/ 2 ord (P) ) mod N
   end
else
     if P = - Q then return (θ)  end
     g := pgcd ( Absc(Q) - Absc(P) mod N, N )
     if  g = 1    then
     λ := [(ord(Q) - ord(P)) / Absc(Q) - Absc(P)] mod N

end
end
if g = 1 then
     x := λ^2 - Absc(P) - Absc(Q) mod N
     y := λ (Absc (P) – x) - ord (P) mod N
   #  Point(x,y) retourne le point d'abscisse x
   #  et d'ordonnée y
    return (Point(x,y))
 else
  return(θ)
end
end EcModAdd
```

Algorithme de l'addition de deux points sur courbe elliptique

Pour sommer deux points identiques, 4 multiplications et une inversion sont nécessaires, alors que pour une addition simple; il faut 3 multiplications, 5 additions et une inversion.[1]

La parametrisation de Montgomery

Pour éviter l'inversion de l'addition naïve, on peut chercher à garder la trace du numérateur et dénominateur des coordonnées d'un point P. Ainsi au lieu de travailler avec P =(x,y) sur la courbe $y^2 = x^3+ax +b$, on utilise P = (X,Y,Z) sur la courbe $Y^2Z = X^3 + aYZ^2 + Z^3$ où $x = \frac{X}{Z}$ et $y = \frac{Y}{Z}$. Les expressions de X, Y et Z sont cependant si coûteuses lors d'une addition ($\approx$ 10 multiplications pour chaque coordonnée) qu'elles ont poussé P. L. Montgomery à considérer des courbes particulières et une parametrisation adaptée.

IL utilise des courbes de la forme :

$By^2 = x^3 +Ax + x$ pour $(A,B) \in (Z/NZ)^2$, $B \neq 0$, $A \neq \pm 2$ (5.1)

En considérant $P_1 = (x_1,y_1)$ et $P_2 = (x_2,y_2)$ avec $x_1 \neq x_2$ et $x_1x_2 \neq 0$, l'abscisse de $P_3 = (x_3,y_3) = P_1+P_2$ vérifie :

$$x_3 = B\left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2 - A - x_1 - x_2,$$

qui peut s'écrire encore

$$x_3(x_1-x_2)^2 = B\frac{(x_2y_1 - x_1y_2)^2}{x_1x_2} \quad (5.2)$$

En utilisant (5.2), on obtient de même pour $P_4 = (x_4,y_4) = P_1 - P_2$

$$x_4(x_1-x_2)^2 = B\frac{(x_2y_1 - x_1y_2)^2}{x_1x_2} \quad (5.3)$$

Ces deux équations conduisent alors à une expression pour $P_1 \neq P_2$ encore valable quand $x_1x_2 = 0$:

$$x_4x_3(x_1 - x_2)^2 = (x_1x_2 - 1)^2 \quad (5.4)$$

De même, on aurait pour $P_1 = P_2$:

$$4x_1x_3(x_1^2 + Ax_1 + 1) = (x_1 - 1)^2 \quad (5.5)$$

Ces formules permettent alors de calculer l'abscisse de la somme de deux points connaissant les abscisses de ces points ainsi que celle de leur différence. Pour la méthode de Lenstra, on doit calculer mP avec $m \in N$. Si on note (X_m, Y_m, Z_m) et (X_n, Y_n, Z_n), les

coordonnées homogènes de mP et nP, on obtient alors à partir de (5.4) et (5.5) connaissant (m - n)P de coordonnées (X_{m-n}, Y_{m-n}, Z_{m-n}) les formules suivantes pour mP $\neq$ nP :

$X_{m+n} = Z_{m+n}[(X_m - Z_m)(X_n + Z_n) + (X_m + Z_m)(X_n - Z_n)]$,

$Z_{m+n} = X_{m-n}[(X_m - Z_m)(X_n + Z_n) – (X_m + Z_m)(X_n - Z_n)]$,

Et pour mP = nP :

$4X_nZ_n = (X_n + Z_n)^2(X_n - Z_n)^2$,

$X_{2n} = 4X_nZ_n((X_n – Z_n)^2 + \frac{A+2}{4} 4X_nZ_n)$.

Cette façon de faire nécessite pour une somme simple 6 multiplications et 4 additions et pour une somme double 5 multiplications et 4 additions si on connaît la différence des 2 points. Il faut cependant remarquer que les courbes de la forme (5.1) ne représentent pas toutes les courbes sur Z/pZ . En effet à partir d'une courbe elliptique en notation de Montgomery

$$Y^2 = BX^3 + AX^2 + BX \ ,$$

Le changement de variables linéaire

$$\begin{cases} x = BX + \frac{A}{3} \\ y = BY \end{cases}$$

permet d'obtenir la courbe

$$y^2 = x^3 + axb$$

On a alors

$$\begin{cases} a = -3(\frac{A}{3})^2 + B^2 \\ b = 2(\frac{A}{3})^2 - (\frac{A}{3})B^2 \end{cases}$$

Mais la factorisation inverse n'est possible que si une solution $\frac{A}{3}$ de

$$(\frac{A}{3})^3 + b(\frac{A}{3}) + a = 0,$$

est telle que la quantité

$$B^2 = b + 3(\frac{A}{3})^2$$

soit un résidu quadratique.[1]

Additionner 2l points sur l courbes en même temps

Au lieu de calculer en coordonnées affines $P + Q$ sur une courbe E, on calcule en même temps $P_i + Q_i$ sur l courbes E_i ($0 \leq i < l$), Si on calculait successivement ces l additions, on aurait alors l entiers ω_i à inverser et un coût prohibitif.

L'optimisation est basé sur la remarque suivante ; si on a deux entiers x et y à inverser, on a leurs inverses respectif par $\frac{1}{x} = \frac{y}{xy}$ et par $\frac{1}{y} = \frac{x}{xy}$. Pour une unique inversion $\frac{1}{xy}$, on obtient donc deux inverses $\frac{1}{x}$ et $\frac{1}{y}$.

Cette remarque est généralisée pour l inversions ω_i^{-1} dans l'algorithme ManyInversions() :

Procedure ManyInversions(ω, l, g, N)

$v_0 := \omega_0$

for $k := 1,2,\ldots,l-1$ **do**

$v_k := v_{k-1}\,\omega_k \bmod N$

end

$g := p\gcd(v_{l-1}, N)$

if $g := 1$ **then**

$u := v_{l-1}^{-1} \bmod N$

for $k := l-1,l-2,...,1$ **do**

$t_k := u_{k-1}u \bmod N$

$u := \omega_k u \bmod N$

end

$t_0 := u$

end

return(t)

end ManyInversions

l Inversions modulaires avec une seule.

On remplace ainsi l inversions par une inversion et $3(l-1)$ multiplications. Le prix d'une somme double sur une courbe vaut alors

$$7 + \frac{k-3}{l}$$

et celui d'une somme simple

$$6 + \frac{k-3}{l}$$

où k est le coût d'une inversion.[1]

6-1-2 La multiplication de points par des entiers

La méthode binaire

Pour calculer kP où k est un entier et P un point, une méthode naïve consisterait à calculer $2P,3P,...,kP$ avec k–1 additions.

Il est possible de réduire ce nombre en remarquant que si k s'écrit en base 2 : $k = b_0 + 2(b_1 + 2(\ldots(b_{[Log_2^k]-1} + 2\,b_{[Log_2^k]})\ldots))$,

Alors : $kP = b_0\,P + 2(b_1\,P) + \ldots + 2(b_{[Log_2^k]-1})\,P + 2\,b_{[Log_2^k]}\,P)\ldots))$,

Ainsi, à titre d'exemple, $11\,P$ s'écrit :

$$11\,P = P + 2(P + 2(2\,P))$$

En utilisant 5 additions au lieu de 10.

La procédure EcModMult() est une implantation de la remarque précédente

```
Procedure EcModMult(k, P, E_{a,b}, g, N)
g := 1
e := [log_2^k]
R := P
for l = e - 1, e - 2, ..., 0 while g = 1 do
        R := EcModMult(R, R, E_{a,b}, g, N)
if [k / 2^e] mod 2 = 1 and g = 1 then
        R := EcModMult(R, R, E_{a,b}, g, N)
end
end
return(R)
end EcModMult
```

Méthode binaire de calcul de kP

Dans le cadre de la parametrisation de Montgomery, il est nécessaire de connaître la différence de deux points avant d'en faire la somme. Cela n'est pas gênant si dans la méthode binaire, au lieu d'obtenir $2mP$ ou $(2m+1)P$ à partir de mP, on obtient le couple

$(2mP,\ (2m+1)P)$ ou $(mP,(m+1)P)$ à partir de $(mP,(m+1)P)$. Cependant, chaque étape nécessite alors 11 multiplications pour un coût total de $11\log_2^k$ multiplications.

En revanche, lorsque l'on effectue plusieurs inversions ensemble en notation de Weierstrass, on effectue à chaque étape une somme double et parfois une somme simple. Statistiquement, un nombre a environ autant de 0 que de 1dans sa décomposition binaire, on a donc besoin pour calculer kP, de $[\log_2^k]$ additions doubles et de $\frac{1}{2}[\log_2^k]$ additions simples en moyennes. Ce qui conduit à un coup total de :

$$(10+\frac{3(k-1)}{2l})\log_2^k \ .[1]$$

Chapitre 7

Autres méthodes de la factorisation des entiers

7-1 La méthode de crible quadratique

C'est en fait l'une des méthodes les plus efficaces de factorisation, elle est basée sur l'idée suivante :

On essaye de factoriser un nombre composite impair N ou, plus précisément, d'en trouver un diviseur propre de N. La méthode du crible quadratique trouve deux entiers x et y tels que :

$x^2 \equiv y^2 \bmod N$ (1)

$x \not\equiv \pm y \bmod N$ (2)

Alors N divise (x - y)(x + y) mais ne divise ni x + y ni x – y. Par conséquent pgcd(x – y, N) est un diviseur propre de N.[4]

La méthode de trouver x et y

On va maintenant décrire une méthode pour trouver x et y satisfaisant les équations (1) et (2).

Soient $m = [\sqrt{N}\,]$ et $P(X) = (X - m)^2 - N$.

Exemple : Soit N =7429 l'entier à factoriser

On a donc $m = [\sqrt{N}\,] = 86$ et $P(X) = (X + 86)^2 - 7429$.

$P(-3) = 83^2 - 7429 = -540 = -1.2^2.3^2.5$

$P(1) = 87^2 - 7429 = 140 = 2^2.5.7$

$P(2) = 88^2 - 7429 = 315 = 3^2.5.7$

D'où :

$83^2 \equiv -1.2^2.3^2.5 \bmod 7429$

$87^2 \equiv 2^2.5.7 \bmod 7429$

$88^2 \equiv 3^2.5.7 \bmod 7429$

En multipliant les deux dernières congruences on obtient ;

$(87.88)^2 \equiv (2.3.5.7)^2 \bmod 7429$.

On peut donc choisit x = 87.88 mod 7429 = 227 et y =2.3.5.7 mod 7429 = 210.

Don pgcd(x –y, N) = pgcd(17,7429) donne un diviseur de 7429.

Alors 7429 =17.437

En fait le point majeur de cet exemple est que pour certains nombres s, p(s) n'a que de petits facteurs premiers. On peut alors utiliser la congruence :

$$(s + m)^2 \equiv p(s) \bmod N$$

en choisissant bien les nombres s pour trouver x et y.

La méthode générale de construction de x et y.

Définition 9 : *(B-lissité) : Soit* $n \in N$ *et soit* $n = \prod_{i=1}^{m} P_i^{\alpha_i}$ *la décomposition de n en facteurs premiers. N est dit B-lisse si et seulement si* $\forall i \in [1,m] \quad P_i \leq B$.

N est dit B-superlisse si et seulement si $\forall i \in [1,m] \quad P_i^{\alpha_i} \leq B$ *.[9]*

On peut présenter la méthode générale de construction de x et y. On choisit un entier B positif et on cherche des entiers s tels que p(s) n'ait que des facteurs premiers appartenant à une base de factorisation :

$$F(B)=\{p \text{ premier}; p \leq B\} \cup \{-1\}$$

De telles valeurs de p(s) sont appelées B-lisse, il faut alors trouver autant des valeurs de s que F(B) n'a d'éléments $s_1,s_2,\ldots,s_l$ tels que $p(s_1)$, $p(s_2),\ldots,p(s_l)$ soient B-lisses et $F(B)=\{p_1,p_2,\ldots,p_l\}$, pour tout $i \in \{1,\ldots,l\}$, comme $p(s_i)$ est B-lisse, les facteurs premiers de $p(s_i)$ sont les éléments de F(B) dont $p(s_i)$ s'écrit :

$$P(s_i) = \prod_{j=1}^{l} P_j^{\alpha_j^i}$$

On va chercher un produit de ces éléments qui soit un carré . Il s'écrit :

$$P(s_1)^{\lambda_1} \ldots\ldots P(s_l)^{\lambda_l} = \prod_{j=1}^{l} P_j^{\lambda_1 \alpha_j^1 + \ldots + \lambda_l \alpha_j^l}$$

pour, $\lambda_1, \lambda_2, \ldots, \lambda_l \in N^*$. Ce produit est un carré si et seulement si $\lambda_1, \lambda_2, \ldots, \lambda_l$ est une solution du système linéaire suivant :

$$\begin{pmatrix} \alpha_1^1 \alpha_1^2 \ldots\ldots\ldots\ldots\ldots\ldots \alpha_1^l \\ \alpha_2^1 \alpha_2^2 \ldots\ldots\ldots\ldots\ldots\ldots \alpha_2^l \\ \\ \\ \alpha_l^1 \alpha_l^2 \ldots\ldots\ldots\ldots\ldots\ldots \alpha_l^l \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \\ \\ \lambda_l \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \\ \\ \\ 0 \end{pmatrix} \text{ mod}[2]$$

Ce système est un système linéaire sur $\mathbb{Z}/2\mathbb{Z}$ et peut donc être résolu par le pivot de gauss.[4]

Le criblage.

Toute la question est maintenant : comment trouver les entiers s tels que p(s) soit B-lisse ? Eh bien : en criblant !

Essayer tous les entiers s un par un et vérifier si l'on peut les factoriser en utilisant les éléments de F(B) est clairement très coûteux . Une technique plus efficace consiste à utiliser des techniques des criblage. On se fixe un intervalle de criblage :

$$S = \{-C,-C+1,\ldots,0,\ldots,C-1,C\}$$

C est appelé la borne de l'intervalle du criblage. Pour tout $s \in S$, on calcule la valeur p(s), puis successivement, pour chaque $p \in F(B)$, on divise les valeurs P(s) par la plus grande puissance possible de p. Les nombres B-lisses sont ceux pour lesquels il reste 1 ou -1 à la fin de ce processus.

Cependant, si l'on doit tester la divisibilité de chaque P(s) par chaque p, la tâche reste très coûteuse. Supposons que l'on connaisse les zéros de P(X) modulo un nombre premier $p \in F(B)$, c'est-à-dire que l'on ait déterminé tous les $s \in \{1,\ldots,n\}$ tels que P(s) soit divisible par p. Alors les nombres $s \in S$ tels que p divise P(s) sont exactement les raciner de P(X) modulo p plus un multiple de p(on n'a pas à effectuer de divisions inutiles). Cette technique est appelée le criblage avec p.[4]

Efficacité de l'algorithme

Dans cette partie, nous allons brièvement nous intéresser à la question de l'efficacité et de la complexité du crible quadratique et des autres de cryptage . Un fait surprenant : jusqu'à ce jour , personne n'a été capable de donner une analyse complète et rigoureuse de la méthode du crible quadratique. Sous certaines hypothèses (que nous

verrons plus tard), il a été possible d'estimer sa complexité, mais ce résultat n'est pas valable en toute généralité.

Commençons par définir une fonction qui nous servira à mesurer la complexité des différents algorithmes. Etant donnés n,u,v des nombres , on définit :

$$L_n(u,v) = e^{v(\log n)^u (\log\log n)^{v-u}}$$

On a, en particulier :

$$L_n(0,v) = (\log n)^v \qquad \text{et}$$
$$L_n(1,v) = e^{v\log n}$$

Ainsi, si un algorithme qui factorise l'entier n (dont la décomposition binaire est de longueur$[\log_2 n]$ s'exécute en temps $L_n(0,v)$, alors cette algorithme est polynomial en la taille de l'entrée. S'in exécute en temps $L_n(1,v)$, alors il est exponentiel en la taille de son entrée. Enfin ,s'il exécute en temps $L_n(u,v)$ avec $0 <u<1$, alors il est sous exponentiel.

On choisit les bornes B et C de sorte que le nombre du nombre d'éléments B-lisse et le nombre d'éléments de F(B) soient à prés égaux. L'hypothèse d'analyse (qui n'est pas trouvée, mais qui semble expérimentalement satisfaite) et que la proportion d'entiers B-lisse $s \in S$ est la même que la proportion d'entiers B-lisse inférieurs à $\sqrt{n}$. On montre alors que le crible quadratique s'exécute en temps

$$L_n(\frac{1}{2}, \quad 1+o(1))$$

Sous l'hypothèse que nous avons faite, le crible quadratique est sous exponentiel. [4]

7-2 La méthode du crible algébrique

la méthode du crible algébrique a été développée par Pollard en 1988. Elle est assez similaire à la méthode du crible quadratique: on cherche des entiers x et y tels que $x^2 \equiv y^2 \bmod N \; et \; x \neq y \bmod N$. Pour cela, deux bases de factorisation sont utilisées, la première constituée de tous les nombres premiers inférieurs à une certaine borne, la deuxième constituée de tous les idéaux premiers de norme inférieure à une certaine borne dans l'anneau des entiers d'un corps de nombre bien choisi. La complexité heuristique de la méthode du crible algébrique est donnée par:

$\exp((c+\theta(1)).(\log N)^{\frac{1}{3}}(\log\log N)^{\frac{2}{3}})$ où $c = (\frac{64}{9})^{\frac{1}{3}} \approx 1.923$. [7]

7-3 La méthode des corps de nombres :

Cette méthode a été introduit par pollard, c'est la méthode la plus rapide, elle se divise en fait deux algorithmes distincts :

Le NFS Spécialement adapté à des nombres du type $n = a.r^t + b.s^u$ (alias SNFS Spécial Number Field Sieve) .

Le NFS applicable à des nombres arbitraires (alias GNFS Général Number Field Sieve). [5]

Chapitre 8

Résultats

RSA data sécurity a mis en place le **RSA** Factoring challenge en Mars 1991. Cet organisme proposa alors 2 listes de nombres à factoriser. La première est composée des clefs potentielles pour le système de chiffrement à clefs publiques **RSA** de 100, 110, 120, …, jusqu à 500 chiffres.

En 1992 les deux premiers nombres de cette liste ont été factorisé par la crible quadratique (100 chiffres par Arjen Lenstra et Mark Manasse, 110 chiffres par Arjen Lenstra), en juin 1993 Thomas Denny, Bruce Dodson, Arjen Lenstrz, Walter Lion, Mark Mannasse et Herman te Riel avec des machines de saarbruecken university, Lehigh university, Bellcore, CWIAmesterddam, et DEC Systems Research Center ont aussi factorisée le nombre de 120 chiffres par la même méthode. Les trois quarts des calculs ont été faits sur stations de travail et le reste sur la Mas Par de bellcore.

Le 2 avril 1994, un challenge, proposé par Martin Gardner en 1976 dans le Scientific American, fut relevé: non, il ne fallait pas plus d'un milliard d'années pour factoriser RSA129, un nombre de 129 chiffres, qui ne sut résister aux affres du crible quadratique. Puis en 1996, le crible sur corps de nombres, découvert par Pollard, donnait les mêmes résultats, en six fois moins de temps. Le 12 avril 1996, Arjen K Lenstra et son équipe annoncèrent par Courier électronique la factorisation de RSA 130. Il utilisa le crible sur corps de nombres avec le polynôme .

A la fin de 1998 (S.cavallar, B.Dodson, H.Lenstra, P.C.Leyaud, P.L.Montgomery, B.A.Morphy, Zemmermann) ont factorisé RSA 140 par la méthode du crible quadratique.

Pour la factorisation de RSA 155, les meilleurs spécialistes du domaine A.K.Lenstra et P.L.Montgomery se sont attelés à l'écriture des programmes nécessaires, les calculs ont été repartis entre 300 ordinateurs de 11 équipes triées sur le volet, représentant 6 pays (Pays bas, usa, Australie, Royaume uni, France). Les calculs ont

commencé le 27 avril 1999 et la première phase s'est terminée le 13 juillet, la seconde s'est achevée le 22 août par la découverte de deux facteurs premiers de nombre de RSA 155.

Les facteurs premiers en temps raisonnable a déposé les centaines de chiffres dans les dernières années du siècle (le record actuel 158 chiffres) au janvier 2002**.[**6**]**

Chapitre 9

Conclusion

Les méthodes de factorisation d'entiers se répartissent en deux classes. La première est formée des algorithmes *probabilistes*, qui trouvent un facteur de l'entier qu'on chercher à factoriser, en un temps difficile à estimer. Citons l'algorithme *p-1* de Pollard, la méthode des courbes elliptiques de Lenstra. La seconde comprend les algorithmes dits *tous-usages*, qui peuvent factoriser un entier en un temps donné. Les algorithmes les plus performants à ce jour sont ceux du *crible quadratique* et du *crible algébrique* .

Pour la méthode des courbes elliptiques, la première phase de l'algorithme de Lenstra, avec la seconde phase classique est bien rodée. La seconde phase du « paradoxe des anniversaires » a été encore peu utilisée. Une utilisation intensive devrait néanmoins permettre de mieux concerner quelle est l'efficacité de cette dernière en pratique avec la clef découverte de nouvelles factorisations.

Les routines développées ici pour obtenir un algorithmique efficace ne sont pas toutes spécifiques à la factorisation. Ainsi , l'arithmétique liée aux courbes elliptiques et aux polynômes est d'un usage général et pourra certainement être réemployée dans l'autre domaines; calcul du nombre de points d'une courbe elliptique sur **GF**(p), l'algorithme discret sur courbes elliptiques.

Bibliographie

[1] Mémoire de DEA. Factoriser des entiers par la méthode des courbes elliptiques. Reynald Lercier.

[2] Introduction à la théorie des courbes elliptiques. Marc Joye .

[3] La cryptographie décryptée. H.X.Mel § Doris Baker.

[4] Factorisation. Alexandra Bruasse-Bac.

[5] Crible quadratique, fractions continuées et consort. Cyril Banderier (1996/1997).

[6] Revue d'explosion des Mathématique (juin 2002).

[7] Le cryptosystème RSA. Jean-Sebestien Coron.

[8] Le protocole de Diffie et Hellman. Sommaire de la cryptographie expliquée.

[9] Courbes elliptiques et factorisation. Samuel Mimran (2001-2002).

[10] Cours de PGS (2002-2003). Betina Kamel.

[11] Elliptic curve groupe over Φ_p .www.certicom.com/research/online.html:

a) Elliptic curve addition.
b) Adding the points O and –P.
c) Doubling the point P.
d) Elliptic curve groupe over F_p.