TIPE - Courbes elliptiques et factorisation

Samuel MIMRAM

2001–2002

Une décomposition : la factorisation

Dans ce TIPE, nous allons définir des ensembles particuliers appelés courbes elliptiques et munir ces derniers d'une structure de groupe. Nous proposerons ensuite une méthode basée sur l'utilisation de ces groupes pour décomposer un "grand" entier en produit de facteurs premiers. Cette méthode est, à l'heure actuelle, la plus rapide pour factoriser les entiers compris entre 10^6 et 10^{30}.

1 Les courbes elliptiques, définitions, propriétés

1.1 Définition des courbes elliptiques

Définition 1 (Plan projectif)**.** On appelle *plan projectif* sur un corps $\mathbb{K}$ l'ensemble, noté $\mathbb{P}^2(\mathbb{K})$, des classes d'équivalence $\left(\mathbb{K}^3\backslash\{(0,0,0)\}\right)/\mathcal{R}$, où $\mathcal{R}$ est une relation d'équivalence définie par :

$$\forall\left((a,b,c),(a',b',c')\right) \in \left(\left(\mathbb{K}^3\backslash\{(0,0,0)\}\right)^3\right)^2,$$
$$(a,b,c)\mathcal{R}(a',b',c') \Leftrightarrow [\exists t \in \mathbb{K}\backslash\{0\}, (a,b,c) = t(a',b',c')]$$

Définition 2 (Courbe elliptique)**.** On appelle *courbe elliptique sur un corps* $\mathbb{K}$, notée $E(\mathbb{K})$, une courbe cubique dans le plan projectif $\mathbb{P}^2(\mathbb{K})$ i.e. définie par $F(X,Y,Z) = 0$ où F est un polynôme de degré 3, homogène en trois variables, à coefficients dans $\mathbb{K}$:

$$F(X,Y,Z) = \alpha_1X^3+\alpha_2Y^3+\alpha_3Z^3+\alpha_4X^2Y+\alpha_5X^2Z+\alpha_6Y^2X+\alpha_7Y^2Z+\alpha_8Z^2X+\alpha_9Z^2Y+\alpha_{10}XYZ = 0$$

et munie d'une origine $\mathcal{O} \in E(\mathbb{K})$.

Remarque 3. Dans la suite, nous nous intéresserons aux courbes elliptiques non singulières i.e. :

$$\forall P \in E(\mathbb{K}), \left(\frac{\partial F}{\partial X}(P), \frac{\partial F}{\partial Y}(P), \frac{\partial F}{\partial Z}(P)\right) \neq (0,0,0)$$

définies sur un corps $\mathbb{K}$ de caractéristique différente de 2 ou 3.

Lorsqu'il n'y a pas d'ambiguité sur le corps, nous noterons indifféremment les courbes $E(\mathbb{K})$ ou E.

Proposition 4. *Soit E une courbe elliptique. On peut se ramener à une équation de E, dite forme courte de Weierstrß :*

$$E : Y^2Z = X^3 + aXZ^2 + bZ^3 \tag{1}$$

On peut alors écrire cette équation en coordonnées non homogènes ($x = \frac{X}{Z}$ et $y = \frac{Y}{Z}$) :

$$E : y^2 = x^3 + ax + b \tag{2}$$

plus le point $\mathcal{O} = (0, 1, 0)$ qui est le seul point à l'infini ($Z = 0$) et que l'on choisit comme origine.

Théorème 5. *Soit E une courbe donnée par une équation de Weierstraß. Alors E est non singulière si et seulement si la quantité $\Delta = 4a^3 + 27b^2$ est non nulle.*

1.2 Structure de groupe abélien

Montrons que l'on peut munir une courbe elliptique d'une structure de groupe abélien.

Proposition 6. *Soit E une courbe elliptique et D une droite définies sur un corps $\mathbb{K}$. Si E a au moins deux points d'intersection (comptés avec leur multiplicité) avec la droite D, alors E a exactement trois points d'intersection (comptés avec leur multiplicité) avec la droite D.*

1.2.1 Approche géométrique de la loi de la sécante-tangente

Soit E une courbe elliptique définie sur $\mathbb{P}^2(\mathbb{K})$. On peut alors définir sur E une loi de composition $*$ dite *loi de composition de la sécante-tangente* (cf. figures) :

- si $(P, Q) \in E^2$ avec $P \neq Q$, on définit $P * Q$ comme étant le troisième point d'intersection de la droite D passant par P et Q avec E ;
- si $P \in E$, on définit $P * P$ comme étant le troisième point d'intersection de la droite D tangente à la courbe en P avec E (P est alors considéré comme un point double d'intersection).

1.2.2 Expression analytique de $*$

Soit E une courbe elliptique définie par :

$$E : f(x, y) = y^2 - (x^3 + ax + b) = 0 \cup \mathcal{O} = (0, 1, 0) \qquad \text{avec } 4a^3 + 27b^2 \neq 0 \tag{3}$$

Proposition 7. *Soit $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ et $P_3 = (x_3, y_3)$ trois points de $E\backslash\{\mathcal{O}\}$ tels que $P_1 \neq P_2$. Si $x_1 \neq x_2$ et si $P_3 = P_1 * P_2$, alors*

$$\begin{cases} x_3 = \lambda^2 - x_1 - x_2 \\ y_3 = \lambda(x_3 - x_1) + y_1 \end{cases}$$

avec $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$.

Remarque 8. Cette définition est encore vraie dans le cas où $x_1 = x_2$ et $y_2 = -y_1$ avec $P_1 * P_2 = \mathcal{O}$.

Proposition 9. *Soit $P_1 = (x_1, y_1)$ et $P_2 = (x_2, y_2)$ deux points de $E\backslash\{\mathcal{O}\}$. Si $P_2 = P_1 * P_1$, alors*

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2 \\ y_3 &= \lambda(x_3 - x_1) + y_1 \end{aligned}$$

avec $\lambda = \frac{3x_1^2+a}{2y_1}$.

Remarque 10. Le fait d'avoir imposé que la courbe soit non singulière nous permet d'être assurés que la tangente existera toujours.

Proposition 11. *On a de plus $\mathcal{O} * \mathcal{O} = \mathcal{O}$.*

Proposition 12. *Si $P_2 = \mathcal{O} * P_1$ ou $P_2 = P_1 * \mathcal{O}$ alors*

$$\begin{aligned} x_2 &= x_1 \\ y_2 &= -y_1 \end{aligned}$$

Remarque 13. La loi de composition interne $*$ est commutative par construction mais n'est pas, *a priori*, associative.

Corollaire 14. *Soit E est une courbe elliptique définie sur un corps $\mathbb{K}$, soit P_1 et P_2 deux points de E. Alors l'opération $+$ définie par*

$$\forall (P_1, P_2) \in E, P_1 + P_2 = \mathcal{O} * (P_1 * P_2)$$

permet de munir E d'une structure de groupe abélien (commutatif) admettant $\mathcal{O}$ comme élément neutre.
De plus, supposons $P_1 = (x_1, y_1) \neq \mathcal{O}$ et $P_2 = (x_2, y_2) \neq \mathcal{O}$. Si $x_1 = x_2$ et $y_2 = -y_1$ alors $P_1 + P_2 = \mathcal{O}$; dans les autres cas, si $P_3 = (x_3, y_3) = P_1 + P_2$, alors

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2 \\ y_3 &= \lambda(x_1 - x_3) - y_1 \end{aligned}$$

où

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & si\ P \neq Q \\ \frac{3x_1^2+a}{2y_1} & si\ P = Q \end{cases}$$

On a enfin

$$\forall P \in E, P + \mathcal{O} = \mathcal{O} + P = P$$

Remarque 15. Cette proposition est encore vraie dans les cas où $\mathbb{K}$ est un corps de caractéristique 2 ou 3.

Remarque 16. Le calcul de λ fait appel à un inverse (dans $\mathbb{K}$) ce qui justifie la nécessité pour $\mathbb{K}$ d'être un corps.

1.3 Cardinalité

Remarque 17. Si p est un nombre premier, on notera dans la suite $\mathbb{F}_p$ un corps fini de cardinal p.

Théorème 18 (Théorème de Hasse)**.** *Soit p un nombre premier. Si $E(\mathbb{F}_p)$ est une courbe elliptique définie sur le corps fini $\mathbb{F}_p$ de cardinalité p alors la cardinalité $\#E(\mathbb{F}_p)$ de $E(\mathbb{F}_p)$ vérifie*

$$|p + 1 - \#E(\mathbb{F}_p)| \leq 2\sqrt{p} \tag{4}$$

Remarque 19. Deuring a montré que pour tout entier premier p, pour tout entier e compris entre $p+1-2\sqrt{p}$ et $p+1+2\sqrt{p}$, il existe une courbe elliptique $E(\mathbb{F}_p)$ telle que $\#E = e$. C'est donc le meilleur encadrement possible.

2 Une méthode de factorisation

L'algorithme de factorisation utilisant les courbes elliptiques est une variante de la méthode $p-1$ de Pollard dans $\mathbb{F}_p$.

Définition 20 (B-lissité)**.** Soit $N \in \mathbb{N}$ et soit $N = \prod_{i=1}^{m} p_i^{\alpha_i}$ la décomposition de N en facteurs premiers.

N est dit *B-superlisse* si et seulement si :

$$\forall i \in [\![1, m]\!], p_i^{\alpha_i} \leq B$$

Méthode 21 (de Lenstra)**.** Soit $N \in \mathbb{N}$ un entier, supposé non premier, admettant un facteur premier p (que l'on veut trouver). Voici les étapes du déroulement de la méthode de Lenstra pour la factorisation de N à l'aide de la courbe elliptique $E_{a,b}$ définie par :

$$E_{a,b} : y^2 = x^3 + ax + b \cup \mathcal{O} = (0, 1, 0)$$

où a et b sont des entiers. On supposera la cardinalité de $\#E_{a,b}(\mathbb{F}_p)$ B-superlisse.

- On vérifie que N **n'est divisible ni par** 2**, ni par** 3 (sinon, on a trouvé un facteur de N) pour nous assurer que $\mathbb{Z}/p\mathbb{Z}$ sera de caractéristique différente de 2 ou 3 et donc le groupe $(E_{a,b}(\mathbb{Z}/p\mathbb{Z}), +)$ bien défini.
- En notant $\Delta = 4a^3 + 27$, on vérifie que $\mathbf{\Delta \wedge N = 1}$ pour nous assurer que $E(\mathbb{Z}/N\mathbb{Z})$ est non singulière. Si $1 < \Delta \wedge N < N$ alors nous avons trouvé un facteur non trivial de N et si $\Delta \wedge N = N$ alors nous choisissons une autre courbe elliptique (d'autres valeurs de a, b et B).
- On choisit un point $P \in E_{a,b}(\mathbb{Z}/N\mathbb{Z})\backslash\{\mathcal{O}\}$. $\#E_{a,b}(\mathbb{F}_p)$ étant supposée B-superlisse, $\#E(\mathbb{F}_p)$ divise $k = \text{PPCM}(1, 2, \ldots, B)$. D'où, d'après le théorème de Lagrange, l'ordre de P divise k donc $kP = \mathcal{O}$.
  Si l'on calculait $kP = P + P + \ldots + P$ dans $E(\mathbb{F}_p)$, on serait alors amené à calculer $Q + P = \mathcal{O}$ avec $Q = k'P$ et $k' < k$. En pratique, on va effectuer le **calcul de kP** non pas dans $E_{a,b}(\mathbb{F}_p)$ mais dans $E_{a,b}(\mathbb{Z}/N\mathbb{Z})$.
  Rappelons les formules d'addition pour $P_3 = P_1 + P_2$:
  - si $x_1 = x_2$ et $y_2 = -y_1$ alors $P + Q = \mathcal{O}$

– sinon

$$x_3 = \lambda^2 - x_1 - x_2$$
$$y_3 = \lambda(x_1 - x_3) - y_1$$

où

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{si } P \neq Q \\ \frac{3x_1^2 + a}{2y_1} & \text{si } P = Q \end{cases}$$

On aura $P + Q = \mathcal{O}$ donc $x_1 \equiv x_2 \pmod{p}$, soit $p | x_1 - x_2$. Mais il se peut que l'on ait : $x_1 \not\equiv x_2 \pmod{N}$. Lors du calcul λ pour additionner Q et P, $(x_1 - x_2)$ ne sera alors pas inversible (dans $\mathbb{Z}/N\mathbb{Z}$) car $p | \text{PGCD}(x_1 - x_2, N)$ donc $(x_1 - x_2) \wedge N \neq 1$. Dans ce cas, si $1 < (x_1 - x_2) \wedge N < N$ alors nous avons trouvé un facteur non trivial de N et si $(x_1 - x_2) \wedge N = N$.

– Si le calcul de kP dans $E(\mathbb{Z}/N\mathbb{Z})$ aboutit alors nous choisissons une autre courbe elliptique (d'autres valeurs de a, b et B).

Pour le choix de la borne de lissité, on peut tenir le raisonnement suivant : si p est un facteur premier de N alors $p < \sqrt{N}$; or, d'après le théorème de Hasse (théorème 18), on a $\#E(\mathbb{F}_p) < (\sqrt{p} + 1)^2$. On peut donc prendre $B \geq \lceil (N^{1/4} + 1)^2 \rceil$.

Remarque 22. Dans la pratique, pour des raisons évidentes de rapidité de calcul, on ne prend pas $k = \text{PPCM}(1, 2, \ldots, B)$ mais simplement $k = \text{PPCM}(B, B - 1)$. Cela diminue la probabilité de trouver un facteur premier mais améliore de beaucoup la rapidité de l'algorithme.

Remarque 23. De plus, plus au lieu de l'addition naïve, le calcul de kP est effectué avec l'exponentiation rapide ; le principe reste cependant valable.

Remarque 24. Pour le choix de P dans $E_{a,b}$, on peut imposer $b = -a$, ce qui permet d'être assuré que $P = (1, 1)$ appartient à la courbe elliptique.

Il a été montré que le temps moyen d'aboutissement de cette méthode est en $O(\mathrm{e}^{(1+\varepsilon) \ln N \ln \ln N})$.

L'inconvénient de cette méthode est qu'elle n'est pas, stricto sensu, un algorithme car elle n'aboutit pas forcément à un résultat. Il faut lui adjoindre des tests de primalité (il en existe utilisant les courbes elliptiques).

3 Implémentation

L'implémentation a été réalisée à l'aide de Visual C++ 6.0 et testée sur un Pentium 4 1,5 GHz avec 256 Mo de RAM. Elle a en particulier permis de trouver, en 9 minutes seulement, un facteur premier d'un entier produit de trois facteurs premiers de 8 chiffres.

Une classe supportant les opérations dans les courbes elliptiques a été conçue. Le programme utilise entre autres l'algorithme d'Euclide étendu avec obtention des cœfficients de Bezout (pour l'inversion dans $\mathbb{Z}/n\mathbb{Z}$), l'exponentiation rapide et la gestion d'entiers de taille arbitraire (supérieure à 32 bits).

Bibliographie

[1] Thomas Cormen, Charles Leiserson, Ronald Rivest. *Introduction à l'algorithmique*. Dunod, 1994.

[2] Johannes Buchmann. La factorisation des grands nombres. *Pour la science*, (251) : pages 88–96, septembre 1998.

[3] François Arnault. *Théorie des nombres et cryptographie*. Cours de DEA, Université de Limoges, mars 2000.
`http://www.unilim.fr/laco/perso/francois.arnault`.

[4] *Site de Certicom*.
`http://www.certicom.com/resources/ecc/ecc.html`.

[5] Marc Joye. *Une introduction élémentaire à la théorie des courbes elliptiques*. Technical report, Université catholique de Louvain, juin 1995.
`http://www.dice.ucl.ac.be/crypto`.

[6] Reynald Lercier. *Algorithmique des courbes elliptiques dans les corps finis*. Thèse, LIX – CNRS, juin 1997.
`http://www.medicis.polytechnique.fr/~lercier/preprints`.

[7] Douglas Stinson. *Cryptographie – Théorie et pratique*, chapître 5, pages 159–170. International Thomson publishing, 1995.

Figures

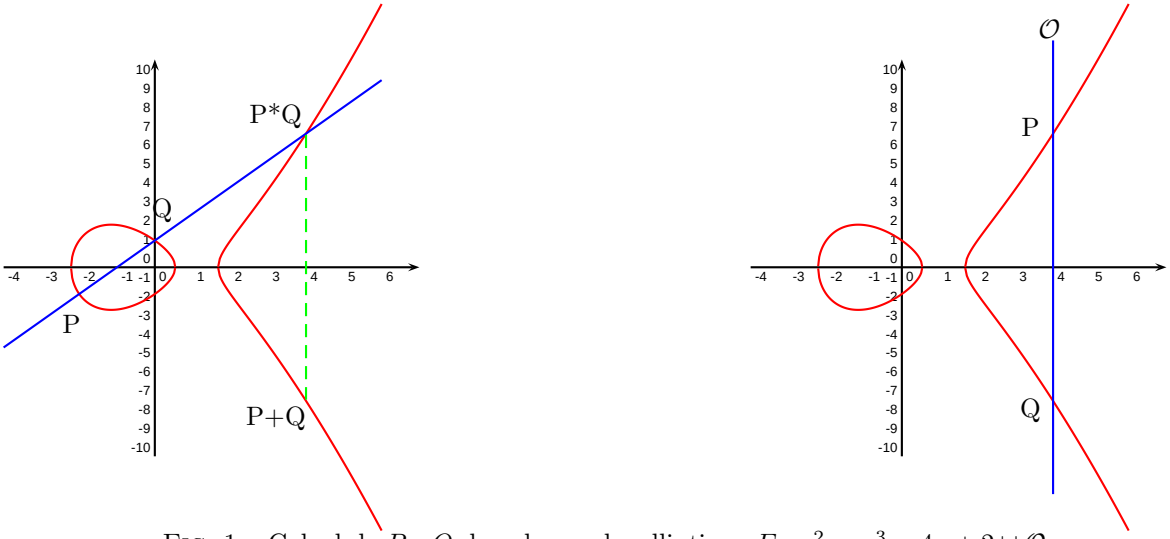

FIG. 1 – Calcul de $P * Q$ dans la courbe elliptique $E : y^2 = x^3 - 4x + 2 \cup \mathcal{O}$

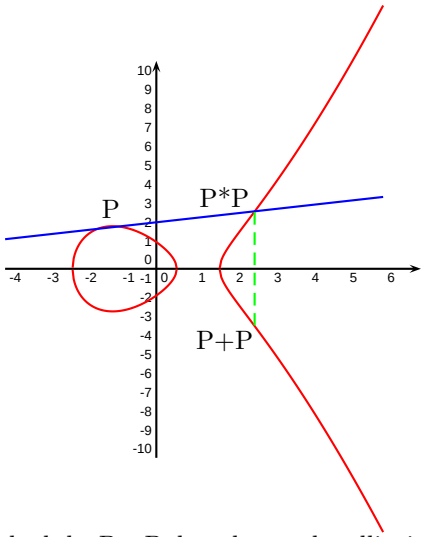

FIG. 2 – Calcul de $P * P$ dans la courbe elliptique $E : y^2 = x^3 - 4x + 2 \cup \mathcal{O}$