BASES MATHÉMATIQUES DU SYSTÈME R.S.A.

MAURICE MIGNOTTE
UNIVERSITÉ LOUIS PASTEUR, STRASBOURG, FRANCE

Table des matières

1. Introduction

Ce texte repose sur un exposé fait à l'IREMPT de Dakar le 22 février 2006 et je remercie les collègues de Dakar pour la chaleur de leur accueil.

Le but était de donner complètement les bases mathématiques du système de cryptographie R.S.A., en utilisant des outils mathématiques aussi élémentaires que possible.

Par rapport à l'exposé cette redaction comporte quelques réductions suplémentaires.

2. Principes du système R.S.A.

Ce paragraphe sera très bref, on trouve beaucoup de présentations détaillées de ce système en naviguant sur Internet.

La méthode R.S.A. (du nom de ses inventeurs Rivest, Shamir et Adleman) est un système cryptographique à clefs publiques qui date de 1976 et qui est beaucoup utilisé aujourd'hui : cartes de crédits, puces, commerce électronique, ...

Comme dans tout système cryptographique à clefs publiques, on considère un ensemble d'individus

$$\big\{\text{Amina, Babacar, Coumba, Daouda, ..., Youssouf, Zeïnabou}\big\}$$

qui échangent des messages secrets (chiffrés), pour chacun de ces individus on a des données publiques

$$A, N_A, e_A,\ B, N_B, e_B,\ C, N_C, e_C,\ \ldots,\ Z, N_Z, e_Z,$$

où N_A et e_A sont des entiers... J'oublierai maintenant les indices. Plus précisément, chaque N est le produit pq de deux nombres premiers secrets et l'entier e est premier avec le nombre $Q = (p-1)(q-1)$.

Lorsque Amina veut envoyer un message m à Babacar, elle consulte l'annuaire et y trouve les données $N = N_B$ et $e = e_B$ relatives à Babacar. Pour simplifier (ce n'est pas une restriction sérieuse) on considère que m est un entier avec $1 < m < N_B - 1$. Le chiffrage de m est le message m' où

$$m' \equiv m^e \pmod{N}.$$

Babacar reçoit donc m', il dispose d'une clef secrète $d = d_B$, qu'il est le seul à connaître, et qui est entier positif tel que

$$ed \equiv 1 \pmod{Q}.$$

Il calcule

$$m'' \equiv m'^d \pmod{N}$$

et bien sûr on veut que $m'' = m$, le message original.

3. Validité du système

Il s'agit donc de montrer que si p et q sont deux nombres premiers distincts et si $N = pq$ alors $(m^e)^d = m^{ed} \equiv m \pmod{N}$ lorsque $ed \equiv 1 \pmod{Q}$ où, rappelons-le, $Q = (p-1)(q-1)$.

On va utiliser des résultats dont le plus récent date d'environ 1630, d'autres ayant au moins 2200 ans (présentables en terminale ?).

Proposition 0.1. Soient a et b deux entiers premiers entre eux, il existe alors u et $v \in Z$ tels que

$$au + bv = 1.$$

Première démonstration : On peut considérer a et b strictement positifs.

Considérons l'ensemble

$$I = \{ax + by;\ x,\ y \in \mathbb{Z}\}.$$

Alors $a = a \cdot 1 + b \cdot 0$ et $b = a \cdot 0 + b \cdot 1$ appartiennent à I. Comme I contient des éléments positifs, il contient un entier $d > 0$ minimal.

Montrons que tout élément x de I est un multiple de d. Par division euclidienne, $x = qd + r$ où $0 \leq r < d$. Il est clair que qd appartient à I, donc $r = x - dq \in I$ et le choix de d impose $r = 0$, ce qui montre que x est bien un multiple de d.

Récapitulons : on a $d \in I$, il existe donc u, $v \in \mathbb{Z}$ tels que $d = au + bv$; d'autre part, a et b appartiennent à I et sont donc de la forme $a = da'$ et $b = db'$, où a' et b' sont des entiers. Comme a et b sont premiers entre eux, on a $d = 1$, ce qui achève cette démonstration.

Deuxième démonstration : Cette démonstration est un peu plus lourde que la précédente, mais elle possède l'avantage de fournir un moyen de calcul (efficace)

des entiers u et v. On suppose $a > b$, sans perte de généralité. On travaille avec des vecteurs

$$w_i = \begin{pmatrix} t_i \\ u_i \\ v_i \end{pmatrix}, \quad \text{avec } w_{-1} = \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix}, \ w_0 = \begin{pmatrix} b \\ 0 \\ 1 \end{pmatrix}.$$

Puis, si t_{i-1} n'est pas nul, on définit

$$w_i = w_{i-2} - q_{i-1} w_{i-1},$$

où q_{i-1} est le quotient de t_{i-2} par t_{i-1}, c'est-à-dire que

$$t_{i-2} = q_{i-1} t_{i-1} + t_i, \quad \text{avec } \ 0 \leq t_i < t_{i-1}.$$

Il est clair que la suite t_i est strictement décroissante. Il existe donc un entier j tel que $t_{j+1} = 0$. On s'arrête alors et on pose $d = t_j$.

La relation $t_{j-1} = t_j q_j = d q_j$ montre que d divise t_{j-1} et t_j. Puis la relation $t_{i-2} = q_{i-1} t_{i-1} + t_i$ montre que d divise t_{j-2} et t_{j-1}. Et ainsi de suite... On en conclut que d divise à la fois $a = t_{-1}$ et $b = t_0$. Sous l'hypothèse que a et b sont premiers entre eux, on voit que $d = 1$. D'autre part, on a par définition

$$a u_{-1} + b v_{-1} = t_{-1} = a \ \text{ et } \ a u_0 + b v_0 = t_0 = b.$$

Montrons, par récurrence sur i, que l'on a

$$a u_{-1} + b v_{-1} = t_{-1} = a, \ \text{ pour } \ -1 \leq i \leq j.$$

On vient de voir que ceci est vrai pour pour $i = -1$ et $i = 0$. La relation

$$w_i = w_{i-2} - q_{i-1} w_{i-1}$$

permet facilement de montrer que si cette propriété a lieu pour les indices $i-2$ et $i-1$ alors elle a encore lieu pour les indices $i-1$ et i. D'où le résultat. En particulier, pour l'indice j, il vient $a u_j + v_j b = t_j = d$.

Fin de la démonstration.

Exemple. Prenons $a = 11$ et $b = 46$. Le tableau de calculs se présente comme suit :

i	-1	0	1	2	3	4
t_i	71	46	25	21	4	1
u_i	1	0	1	-1	2	-11
v_i	0	1	-1	2	-3	17
q_i		1	1	1	1	5

Et on obtient la relation :

$$(-11) \times 71 + 17 \times 46 = 1.$$

Remarque 1. La relation de la propositiion 1 est connue sous le nom de relation de Bézout. En fait, Bézout a démontré ce résultat pour des polynômes coefficients dans un corps. Pour les entiers, ce résultat était connu bien avant, il est dû à Bachet de Méziriac.

Remarque 2. Il est clair que si une relation de la forme $au + bv = 1$ a lieu avec u, $v \in \mathbb{Z}$ alors les entiers a et b sont premiers entre eux.

Proposition 0.2. (Euclide-Gauss) Si deux entiers b et c, premiers entre eux, divisent un entier a, alors a est divisible par le produit bc.

$$\boxed{\left.\begin{array}{rcl} a & \in & b\mathbb{Z} \\ a & \in & c\mathbb{Z} \\ b \wedge c & = & 1 \end{array}\right\} \implies a \in bc\mathbb{Z}}$$

Démonstration. Par la proposition 1, il existe u et v dans $\mathbb{Z}$ tels que

$$bu + cv = 1.$$

D'autre part, par hypothèse, $a = ba' = ca''$, où a' et a'' sont dans $\mathbb{Z}$. Si on multiplie la relation de Bézout ci-dessus par a, il vient :

$$a = abu + acv = a'bcu + a''bcv.$$

Ce qui prouve le résultat.

Proposition 0.3. (Euclide-Gauss) Si a et b sont deux entiers premiers entre eux et si a divise le produit bc, alors a divise l'entier c.

$$\boxed{\left.\begin{array}{rcl} bc & \in & a\mathbb{Z} \\ a \wedge b & = & 1 \end{array}\right\} \implies c \in a\mathbb{Z}}$$

Démonstration. On a cette fois

$$au + bv = 1, \quad \text{avec } u \text{ et } v \ \in \mathbb{Z}.$$

Par multiplication par c,

$$acu + bcv = c,$$

et comme a divise bc, on voit que a divise bien c.

Proposition 0.4. (Fermat) Si p est un nombre premier et $x \in Z$ alors

$$x^p \equiv x \pmod{p}.$$

Nous allons donner deux démonstrations — radicalement diférentes — de ce résultat appelé petit théorème de Fermat, l'une multiplicative et l'autre "additive".

Première démonstration. Si le nombre premier p divise x alors la relation à démontrer s'écrit : $0^p = 0 \pmod{p}$, ce qui est trivialement vrai.

Si p ne divise pas x, considérons l'application

$$f : \mathbb{Z}/p\mathbb{Z} \xrightarrow{y \mapsto xy} \mathbb{Z}/p\mathbb{Z} \ .$$

Montrons que f est injective : si y, $y' \in \mathbb{Z}$ vérifient

$$xy \equiv xy' \pmod{p}$$

alors p divise le produit $x(y-y')$. Comme p ne divise pas x et que p est premier, x et p sont premiers entre eux et la proposition 3 montre que p divise $y-y'$, d'où cette affirmation.

Mais comme $\mathbb{Z}/p\mathbb{Z}$ est fini, l'application f est en fait bijective. En d'autres termes, lorsque y parcourt $\mathbb{Z}/p\mathbb{Z}$, les yx parcourent aussi $\mathbb{Z}/p\mathbb{Z}$.

En particulier, lorsque y parcout $(\mathbb{Z}/p\mathbb{Z})^*$ alors il en est de même pour les yx. Donc

$$\prod_{y\in(\mathbb{Z}/p\mathbb{Z})^*} y = \prod_{y\in(\mathbb{Z}/p\mathbb{Z})^*} yx = x^{p-1} \prod_{y\in(\mathbb{Z}/p\mathbb{Z})^*} y.$$

Une nouvelle application de la proposition 3 montre que ceci implique :

$$x^{p-1} \equiv 1 \pmod{p} \quad \text{(le "vrai" petit théorème de Fermat)}$$

et donc

$$x^p \equiv x \pmod{p}.$$

Ce qu'il fallait démontrer.

Deuxième démonstration. Un instant de réflexion montre qu'il suffit de démontrer le résultat pour tout $x \geq 0$. On va alors procéder par récurrence sur x. Pour $x = 0$, tout simplement,

$$0^p \equiv 0 \pmod{p}.$$

Supposons maintenant le résultat vrai pour x, alors

$$(x+1)^p = x^p + C_p^1 x^{p-1} + C_p^2 x^{p-2} + \cdots + C_p^{p-1} x + 1$$

$$\equiv x^p + 1 \equiv x + 1 \pmod{p},$$

la première congruence provenant du fait que p divise les coefficients du binôme C_p^k pour $1 \leq k \leq p-1$, et la seconde de l'hypothèse de récurrence. D'où le résultat.

Corollaire 0.1. Si p est un nombre premier et si $t \equiv 1 \pmod{p}$ et $t > 0$, alors pour tout $x \in \mathbb{Z}$ on a

$$x^t \equiv x \pmod{p}.$$

Démonstration. Posons $t = k(p-1)+1$. Alors $k \geq 0$. Raisonnons par récurrence sur k. Si $k = 0$, c'est clair puisque dans ce cas $t = 1$. Si le résultat est vrai pour un entier $k \geq 0$ alors

$$x^{(k+1)(p-1)+1} = x^{k(p-1)} \cdot x^p \equiv x^{k(p-1)} \cdot x = x^{k(p-1)+1} \equiv x \pmod{p},$$

où la première congruence résulte du petit théorème de Fermat et la dernière de l'hypothèse de récurrence. Ceci achève la démonstration.

Théorème 0.1. Le système R.S.A. est valide :

$$m^{ed} \equiv m \pmod{N} \text{ si } ed \equiv 1 \pmod{(p-1)(q-1)}.$$

Démonstration. Nous avons maintenant tous les outils pour conclure.

Remarquons d'abord que la proposition 4 implique

$$m^{ed} \equiv m \pmod{p} \quad \text{et} \quad m^{ed} \equiv m \pmod{q}.$$

Comme p et q sont premiers entre eux, la proposition 2 implique que le produit pq divise la différence $m^{ed} - m$, autrement dit on a bien

$$m^{ed} \equiv m \pmod{N}, \quad N = pq.$$

Remarque 3. Le lecteur pourra noter que le théorème ci-dessus a lieu sous l'hypothèse plus faible

$$ed \equiv 1 \pmod{Q'}, \quad \text{où} \quad Q' = \operatorname{ppcm}(p-1, q-1).$$

4. Sécurité du système R.S.A.

Il est facile de voir que si on connait les facteurs p et q de N alors on lève le secret. Il est généralement admis que la seule façon de casser le système R.S.A. est de factoriser N, mais ce fait n'est pas démontré.

Notons que la connaissance de l'entier

$$Q = (p-1)(q-1) = pq - (p+q) + 1 = N - (p+q) + 1$$

permet de factoriser N : On connaît le produit $N = pq$ et la somme $p+q$, équation du second degré...

La théorie actuelle de la factorisation des grands entiers est extrêment élaborée. Je ne peux que donner un historique grossier des limites des méthode fatorisation depuis 1976 :

$$\begin{array}{lll} \text{en 1976,} & \text{Limite} & \simeq \quad 10^{70}, \\ \text{en 2000,} & \text{Limite} & \simeq \quad 10^{100}, \\ \text{en 2006,} & \text{Limite} & \simeq \quad 10^{160}. \end{array}$$

Donc aujourd'hui il est conseillé de prendre

$$p \ \text{ et } \ q \geq 10^{100}, \quad \text{donc } N \geq 10^{200}.$$

5. Aspect pratique

Pour le codage on doit donc calculer m^e modulo N, avec $N \geq 10^{200}$ et $e \simeq 10^{200}$. Le décodage correspond à un calcul similaire.

Ce n'est pas une mince affaire !

Bien sûr le calcul de $m^2,\ m^3,\ \ldots, m^e$ est inenvisageable ! On utilise la méthode d'exponentiation rapide connue des chinois depuis un millénaire.

On écrit e en binaire :

$$e = \sum_{i=0}^{k} \varepsilon_i 2^i, \quad \varepsilon_i \in \{0,1\},\ \varepsilon_k = 1.$$

Donc

$$m^e \equiv m^{\sum_{i=0}^{k} \varepsilon_i 2^i} = \prod_{i=0}^{k} m^{\varepsilon_i 2^i} = \prod_{\substack{0 \leq i \leq k \\ \varepsilon_i \neq 0}} m^{2^i} \pmod N,$$

ce qui revient à calculer $m^2 = m \cdot m$, $m^4 = m^2 \cdot m^2$, ..., $m^{2^i} \pmod N$, puis le produit de droite. Soit en tout au plus $2k$ produits modulo N. On a $2^k \leq e$, donc

$$k \leq \frac{\mathrm{Log}e}{\mathrm{Log}2}.$$

Pour e voisin de 10^{200} on $2k \leq 1340$ (noter que $2^{10} = 1024$). Avec un ordinateur assez puissant, on peut effectuer un tel calcul en moins d'une seconde.

6. Authentification-signature

Revenons à l'échange du message m entre Amina et Babacar. Pour que Babacar puisse être certain que c'est bien Amina qui lui écrit, on peut procéder ainsi pour l'envoi du message crypté :

1) Amina calcule $m' = m^{d_A}$, où d_A est sa clef secrète, puis $m'' \equiv m'^{e_B}$, où e_B est la clef publique de Babacar, enfin elle envoie m''.

2) Babacar reçoit m'', il calcule d'abord $m' = m''^{d_B}$, où d_B est sa propre clef secrête, puis $m \equiv m'^{e_A}$, où e_A est la clef publique d'Amina. Il a ainsi récupéré le message m et possède la certitude que l'expéditeur ne peut être qu'Amina.

Remarque 4. Cette présentation comporte une tricherie. Certaines congruences sont modulo N_A, d'autres modulo N_B. La méthode indiquée ne fonctionne que si N_A et N_B sont dans un ordre convenable.

Exercice 0.1. Corriger tout celà !

7. Réalisation d'un système R.S.A.

Il s'agit d'abord de trouver deux nombres premiers p et q assez grands pour obtenir $N = pq$. Heureusement il y a assez de nombres premiers. En effet :

Théorème 0.2. (Euclide) Il y a une infinité de nombres premiers.

Démonstration. Soient p_1, ..., p_n des nombres premiers. Considérons alors le nombre $K = p_1 \cdots p_n + 1$. Soit p un diviseur premier de K (éventuellement égal à K), alors $p \notin \{p_1, \ldots, p_n\}$ sinon il diviserait $1 = K - p_1 \cdots p_n$, ce qui est absurde.

Ainsi la collection des nombre premiers n'est jamais épuisée. Pas mal pour une démonstration qui a plus de 2200 ans.

Remarque 5. On peut montrer (Hadamard – de la Vallée-Poussin) que la fréquence des nombres premiers voisins de x est équivalente à $1/\mathrm{Log}x$. Résultat difficile.

Maintenant comment tester si un nombre $n \geq 10^{100}$ est premier ? La méthode scolaire qui consiste à chercher si n a un diviseur > 1 et $\leq \sqrt{n}$ demande un nombre de divisions de l'ordre de $\sqrt{n}$. Prenons $n \geq 10^{40}$, il faut alors au moins

$2 \cdot 10^{19}$ divisions. Considérons une machine imaginaire qui effectue 10^{10} divisions à la seconde (machine qui n'existe pas encore !), il faut alors plus de $2 \cdot 10^9$ secondes, soit plus de $2 \cdot 10^4$ jours, environ 60 ans... Il faut donc trouver une autre méthode.

On pourrait espérer que le petit théorème de Fermat admet une réciproque, soit :

Si $u \in \mathbb{Z}$ est tel que

$$\forall x \in \mathbb{Z}, \quad x \wedge u = 1 \Rightarrow x^{u-1} \equiv 1 \pmod{u}$$

alors u est premier.

Il n'en est rien, il existe une infinité d'entiers composés qui vérifient cette condition, le plus petit est $561 = 3 \cdot 11 \cdot 17$. En effet si x est un entier premier avec 561, alors le petit théorème de Fermat implique

$$x^{560} = (x^2)^{560} \equiv 1 \pmod{3}, \quad x^{560} = (x^{10})^{56} \equiv 1 \pmod{11}$$

ainsi que

$$x^{560} = (x^{16})^{35} \equiv 1 \pmod{17}$$

et la proposition 2 montre que $x^{560} \equiv 1 \pmod{561}$.

En fait on utilise des tests qui correspondent à des raffinement sophistiqués du petit théorème de Fermat et on ne sait pas montrer — sans admettre l'hypothèse de Riemann généralisée — que ces tests permettent de conclure en un temps raisonnable que n est premier.

On a aussi inventé d'autres tests extêmement savants. C'est seulement en 2003 que trois chercheurs indiens ont apporté un progrès théorique capital : on peut déterminer que n est premier en un temps qui croît de façcon polynômiale en fonction du nombre de chiffres de n ; mais ce résultat n'est pas encore utilisable en pratique (pour plus de détails, faire "AKS primalité" sur votre moteur de recherche favori).

Remarque 6. On a la situation un peu paradoxale suivante, si pour un nombre n on trouve un entier x tel que

$$x^n \not\equiv x \pmod{n}$$

(calcul qui peut se faire rapidement d'après ce qu'on vu) on sait que n est composé (sans avoir trouvé un facteur de n...!!!). Par contre si

$$x^n \equiv x \pmod{n},$$

on ne peut rien conclure.

Supposons désormais p et q choisis. On calcule alors $N = pq$ que l'on rend public et le produit $Q = (p-1)(q-1)$ que l'on garde secret. On choisit e premier avec Q, que l'on publie. Grâce à l'algorithme d'Euclide, on calcule d tel que $ed \equiv 1 \pmod{Q}$ et on le garde secret.

C'est fini.

8. Les Anciens cités

Euclide : mathématicien grec, vers 325 av JC – vers 265 av JC.
Claude-Gaspard Bachet de Méziriac, sieur de Meyseria, mathématicien français, 1581–1638.
Pierre de Fermat, mathématicien français, 1601–1665.
Etienne Bézout, mathématicien français, 1730–1783.
Carl Friedrich Gauss, mathématicien allemand, 1777–1855.

E-mail address: mignotte@math.u-strasbg.fr