De la fonction indicatrice d'Euler et des applications d'icelle en cryptographie

(Leonhard Euler, Bâle 1707 – Saint-Pétersbourg 1783)

Leonhard Euler est né à Bâle (Suisse) en 1707 ; son père était prêtre et mathématicien amateur. En 1720, il rentra à l'Université de Bâle où il étudia la théologie, le grec, le latin et l'hébreu (ayant l'intention de devenir prêtre), mais aussi, pendant son temps libre, la physique, l'astronomie, la médecine et les mathématiques ; ces dernières étaient enseignées par Jean I Bernoulli* (1667–1748) de trois fils duquel Euler fit la connaissance : Nicolas III (1695–1726), Daniel I (1700–1782) et Jean II (1710–1790). En 1725, Nicolas III et Daniel I partirent rejoindre l'Académie des Sciences de Saint-Pétersbourg, et Daniel y fit venir Euler en 1727. Pendant son séjour dans la capitale russe, Euler publia une quantité d'articles impressionnante malgré la perte de l'œil droit en 1735. Cependant, en raison du climat détestable qui y régnait après la mort de la Tsarevna Anna en 1740, Euler accepta l'invitation de Frédéric II de Prusse à séjourner à Berlin, où il passa 25 ans (avec sur les bras des tâches aussi intellectuelles que de vérifier le niveau d'eau du palais). En 1771, il devint aveugle. Cependant, Catherine II de Russie une fois montée sur le trône désira à tout prix récupérer le plus brillant mathématicien d'Europe ; et le prix qu'Euler demanda ne fut pas léger. En 1776, Euler retourna donc à Saint-Pétersbourg (quoique Frédéric II lui causât quelques difficultés), et malgré sa cécité, son flot de publications continua ininterrompu. Dans l'ensemble, il fut probablement le mathématicien le plus prolifique, puisque les *Leonhard Euler Opera Omnia*, qui ne regroupent pourtant pas sa correspondance, totalisent soixante-quinze volumes de 600 pages chacun...

-1. Introduction.

Dans une première partie, je vais définir puis exposer les propriétés élémentaires essentielles de la fonction indicatrice φ d'Euler. Une seconde partie, plus courte et plus informelle, présentera le système cryptographique RSA, qui utilise le théorème d'Euler pour coder et décoder les messages secrets. Je finirai avec un exemple en MAPLE illustrant le fonctionnement de la méthode.

0. Préliminaires : rappels d'arithmétique.

On dit que deux entiers a et a' sont congrus modulo m ($m \in \mathbb{N}^*$) si et seulement si $a' - a$ est divisible par m. On note $a \equiv a' \pmod{m}$. Si $a \equiv a' \pmod{m}$ et $b \equiv b' \pmod{m}$, alors $a + b \equiv a' + b' \pmod{m}$ et $ab \equiv a'b' \pmod{m}$. Si on note $\bar{a}$, ou $a + m\mathbb{Z}$, ou $a + (m)$ la classe de congruence de a modulo m, c'est-à-dire l'ensemble des entiers congrus à a modulo m, on peut alors définir une addition et une multiplication sur l'ensemble des classes de congruence modulo m par $\overline{a+b} = \bar{a} + \bar{b}$ et $\overline{ab} = \bar{a}\bar{b}$. L'ensemble des classes de congruence, muni de ces deux lois, forme un *anneau commutatif*, noté $\mathbb{Z}/m\mathbb{Z}$.

On appelle représentant d'une classe de congruence un élément de celle-ci. Ainsi, un représentant de $\bar{a}$ n'est autre qu'un entier congru à a. On peut remarquer que toute classe de congruence modulo m a un unique représentant dans $[\![0; m-1]\!]$.

Dire que k est premier avec m équivaut à affirmer l'existence d'une relation de Bézout $ku + mv = 1$, c'est-à-dire de u tel que $ku \equiv 1 \pmod{m}$. Autrement dit, k est premier avec m si et seulement si $\bar{k}$ est inversible dans l'anneau $\mathbb{Z}/m\mathbb{Z}$.

On rappelle que si $(A, +, \star)$ et $(A', +', \star')$ sont deux anneaux commutatifs, on définit leur produit $(A \times A', \underline{+}, \underline{\star})$ de la façon suivante : $A \times A'$ est juste le produit cartésien des ensembles, et pour tous $a, b \in A$ et $a', b' \in A'$, on pose

$$(a, a') \underline{+} (b, b') = (a + b, a' +' b')$$

$$(a, a') \underline{\star} (b, b') = (a \star b, a' \star' b')$$

Autrement dit, les opérations s'effectuent composante par composante.

1. Définition de la fonction indicatrice d'Euler et propriétés élémentaires.

On définit la fonction indicatrice φ d'Euler, sur les naturels non nuls et à valeur dans les naturels non nuls, par

* Celui-là même qui posa à Isaac Newton le problème du brachistochrone en 1697.

Définition 1.1. *Si $n \in \mathbb{N}^*$, $\varphi(n)$ est le nombre de naturels inférieurs (strictement) à n et premiers avec lui.*

Exemple : les naturels inférieurs à 10 premiers avec lui sont 1, 3, 7 et 9. Par conséquent, $\varphi(10) = 4$. On remarquera aussi que $\varphi(1) = 1$ (car 0 est premier avec 1).

Si p est un nombre premier, tous les naturels $1, ..., p-1$ sont premiers avec p, et par conséquent $\varphi(p) = p - 1$.

Si n est un naturel non nul, on notera $\mathbb{Z}/n\mathbb{Z}$ l'anneau commutatif des entiers modulo n, et $U(\mathbb{Z}/n\mathbb{Z})$ (ou plus simplement U_n) le groupe des unités de cet anneau, c'est-à-dire le groupe multiplicatif des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$. Exemple : si $n = 10$, le groupe $U(\mathbb{Z}/n\mathbb{Z})$ a 4 éléments : $\bar{1}$, $\bar{3}$, $\bar{7}$ et $\bar{9}$, et sa table de multiplication est donnée par :

$\times$	$\bar{1}$	$\bar{3}$	$\bar{7}$	$\bar{9}$
$\bar{1}$	$\bar{1}$	$\bar{3}$	$\bar{7}$	$\bar{9}$
$\bar{3}$	$\bar{3}$	$\bar{9}$	$\bar{1}$	$\bar{7}$
$\bar{7}$	$\bar{7}$	$\bar{1}$	$\bar{9}$	$\bar{3}$
$\bar{9}$	$\bar{9}$	$\bar{7}$	$\bar{3}$	$\bar{1}$

Ce groupe est bien sûr abélien puisque l'anneau $\mathbb{Z}/10\mathbb{Z}$ est commutatif.

Proposition 1.2. *$\varphi(n)$ est l'ordre du groupe $U(\mathbb{Z}/n\mathbb{Z})$. De plus, $\varphi(n)$ est le nombre de générateurs du groupe cyclique à n éléments.*

Démonstration : L'application $k \mapsto \bar{k}$, qui a un naturel inférieur à n et premier avec lui, associe sa classe de congruence modulo n, est injective, et son image est précisément $U(\mathbb{Z}/n\mathbb{Z})$ (puisque $\bar{k}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si k est premier avec n). Ceci démontre la première affirmation. Soit maintenant g un générateur du groupe cyclique à n éléments. L'application $\lambda: k \mapsto g^k$, qui a un naturel k inférieur à n et premier avec lui, associe g^k, est injective. Si h est dans l'image de λ, on peut écrire $h = g^k$; il existe d'autre part une relation de Bézout $ku + nv = 1$ entre k et n ; alors $g = g^{1-nv} = g^{ku} = h^u$, et donc h est un générateur du groupe. Si h est un générateur du groupe, on peut écrire $h = g^k$ (avec $k \leq n$) et $g = h^u = g^{ku}$, donc $g^{ku-1} = 1$, donc $ku - 1$ est multiple de n, donc on peut écrire $ku + nv = 1$, et donc k est premier avec n, donc h est dans l'image de λ. Ainsi, l'image de λ est l'ensemble des générateurs du groupe cyclique à n éléments, ce qui démontre la seconde assertion.

Théorème 1.3 ("théorème chinois"). *Si m et n sont premiers entre eux, alors l'anneau $\mathbb{Z}/mn\mathbb{Z}$ est isomorphe au produit des anneaux $\mathbb{Z}/m\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z}$. En particulier, U_{mn} est isomorphe à $U_m \times U_n$.*

Démonstration : On définit une application $\lambda: \mathbb{Z}/mn\mathbb{Z} \to (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ par

$$k + mn\mathbb{Z} \mapsto (k + m\mathbb{Z}, k + n\mathbb{Z})$$

Cette application est bien définie car si k et k' sont congrus modulo mn, ils le sont modulo m et modulo n. Il est clair que λ est un morphisme. Montrons qu'il est injectif : si $\lambda(k + mn\mathbb{Z}) = (0 + m\mathbb{Z}, 0 + n\mathbb{Z})$, alors k est multiple de m et multiple de n, et donc, comme m et n sont premiers entre eux, k est multiple de mn. Par conséquent, $k + mn\mathbb{Z} = 0 + mn\mathbb{Z}$. Ainsi, λ est injective. Comme les ensembles $\mathbb{Z}/mn\mathbb{Z}$ et $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ ont même cardinal fini mn, λ est aussi surjective. C'est donc un isomorphisme, ce qui démontre le résultat.

Exemple : l'isomorphisme entre $\mathbb{Z}/10\mathbb{Z}$ et $(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/5\mathbb{Z})$ est donné par

$$\bar{0} \mapsto (\bar{0}, \bar{0})\ ,\quad \bar{1} \mapsto (\bar{1}, \bar{1})\ ,\quad \bar{2} \mapsto (\bar{0}, \bar{2})\ ,\quad \bar{3} \mapsto (\bar{1}, \bar{3})\ ,\quad \bar{4} \mapsto (\bar{0}, \bar{4})$$

$$\bar{5} \mapsto (\bar{1}, \bar{0})\ ,\quad \bar{6} \mapsto (\bar{0}, \bar{1})\ ,\quad \bar{7} \mapsto (\bar{1}, \bar{2})\ ,\quad \bar{8} \mapsto (\bar{0}, \bar{3})\ ,\quad \bar{9} \mapsto (\bar{1}, \bar{4})$$

Corollaire 1.4. *Si m et n sont premiers entre eux, alors $\varphi(mn) = \varphi(m)\,\varphi(n)$. (On dit que φ est "multiplicative".)*

Exemple : $\varphi(2) = 1$, $\varphi(5) = 4$ et $\varphi(10) = 4$.

On peut se demander ce que le "théorème chinois" a de chinois. L'anecdote rapportée par Paulo Ribenboim, est que les généraux chinois s'en servaient pour compter leurs hommes, en leur ordonnant de se ranger 7 par 7, puis 11 par 11, 12 par 12 et 13 par 13 par exemple, puis en comptant à chaque fois combien restaient en rangée incomplète...

On termine par un résultat classique sur φ :

Théorème 1.5. *Si $n \in \mathbb{N}^*$ alors $\sum_{d|n} \varphi(d) = n$.*

Démonstration : Pour tout d' divisant n, on note $A_{n,d'}$ l'ensemble des k de $[\![0, n-1]\!]$ dont le PGCD avec n est $k \wedge n = d'$, et $N_{n,d'}$ le cardinal de cet ensemble. Il est clair que $\sum_{d'|n} N_{n,d'} = n$. La division par d' est une bijection entre $A_{n,d'}$ et $A_{n/d',1}$, donc $N_{n,d'} = N_{n/d',1}$. Mais ce dernier est le nombre d'éléments de $[\![0, n/d'-1]\!]$ premiers avec n/d', c'est-à-dire $\varphi(n/d')$. Ainsi, $\sum_{d'|n} \varphi(n/d') = n$. Mais en effectuant le changement d'indice $d = n/d'$ dans cette somme, on trouve le résultat annoncé.

Exemple : $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(5) = 4$ et $\varphi(10) = 4$, ce qui donne bien $1+1+4+4 = 10$.

2. Procédé général de calcul de $\varphi(n)$.

Théorème 2.1. *Soit n un naturel non nul, dont les facteurs premiers deux-à-deux distincts sont $p_1, \ldots, p_k$. Alors*

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_k}\right)$$

Démonstration : Tout d'abord, si p est premier et $l \geq 1$, alors les naturels inférieurs ou égaux à p^l et multiples de p sont au nombre de p^{l-1} ; par conséquent, les naturels inférieurs ou égaux à p^l et premiers avec ce dernier sont au nombre de $p^l - p^{l-1} = p^l(1-\frac{1}{p})$. Pour passer de là au cas d'un naturel n non nul quelconque, il suffit de décomposer n en facteurs premiers et d'utiliser le corollaire 1.4 : si

$$n = p_1^{l_1} \cdots p_k^{l_k}$$

avec $p_1, \ldots, p_k$ distincts, alors $p_1^{l_1}, \ldots, p_k^{l_k}$ sont premiers entre eux deux à deux, donc

$$\begin{aligned} \varphi(n) &= \varphi(p_1^{l_1}) \cdots \varphi(p_k^{l_k}) \\ &= p_1^{l_1}\left(1 - \frac{1}{p_1}\right) \cdots p_k^{l_k}\left(1 - \frac{1}{p_k}\right) \\ &= n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_k}\right) \end{aligned}$$

c'est le résultat annoncé.

Exemple : les facteurs premiers de 10 sont 2 et 5, donc $\varphi(10) = 10 \times \frac{1}{2} \times \frac{4}{5} = 4$.

Remarquons que $\varphi(n)$ est toujours un nombre *pair* sauf si $n \in \{1, 2\}$. Ainsi, φ n'est pas surjective sur $\mathbb{N}^*$. Elle ne prend pas toutes les valeurs paires non plus, d'ailleurs (par exemple, $\varphi(n)$ ne prend jamais la valeur 14. On peut aussi montrer que φ ne prend jamais une infinité de fois la même valeur.

Profitons de l'occasion pour mentionner un problème ouvert :

Conjecture 2.2. *Pour tout naturel n non nul il existe $n' \neq n$ tel que $\varphi(n') = \varphi(n)$.*

On peut montrer par des moyens simples que la conjecture est vraie pour $n < 42\,401\,268$; après, ça se complique. (Si n est impair, $\varphi(2n) = \varphi(n)$; si n est pair mais non multiple de 4, $\varphi(\frac{1}{2}n) = \varphi(n)$; si n est multiple de 4 mais non de 3, $\varphi(\frac{3}{2}n) = \varphi(n)$; si n est multiple de 12 mais non de 9, $\varphi(\frac{2}{3}n) = \varphi(n)$; si n est multiple de 36 mais non de 7, $\varphi(\frac{7}{6}n) = \varphi(n)$; si n est multiple de 252 mais non de 49, $\varphi(\frac{6}{7}n) = \varphi(n)$; si n est multiple de 1764 mais non de 43, $\varphi(\frac{43}{42}n) = \varphi(n)$; et ainsi de suite.)

3. Le théorème d'Euler. [058′]

Théorème 3.1 ("théorème d'Euler"). *Soit n un naturel non nul, et a un entier premier avec n. Alors*

$$a^{\varphi(n)} \equiv 1 \pmod n$$

Démonstration : La classe $\bar{a}$ de a modulo n appartient à $U(\mathbb{Z}/n\mathbb{Z})$, qui est un groupe d'ordre $\varphi(n)$ (proposition 1.2). D'après le théorème de Lagrange, on a donc $\bar{a}^{\varphi(n)} = \bar{1}$ (car $\bar{1}$ est l'élément unité de $U(\mathbb{Z}/n\mathbb{Z})$). Ceci démontre le théorème.

Exemple : $7^4 = 2401 \equiv 1 \pmod{10}$.

Exercice : quel sont les quelques ($\sim 100\,000$) derniers chiffres de $3^{10^6!}$ en base 10 ?

Corollaire 3.2 ("petit théorème de Fermat"). *Si p est premier et a un entier quelconque, alors $a^p \equiv a \pmod{p}$.*

Démonstration : Si a est multiple de p, les deux membres de l'égalité sont congrus à 0 modulo p. Si a n'est pas multiple de p, il est premier avec p, donc $a^{p-1} = a^{\varphi(p)} \equiv 1 \pmod{p}$ d'après le théorème 3.1, et en multipliant par a, on trouve le résultat.

Le corollaire 3.2 donne une condition *nécessaire* pour que p soit premier. Cette condition n'est *pas suffisante* :

Définition 3.3. *On appelle nombre de Charmichaël un naturel $n > 1$ non premier tel que $a^n \equiv a \pmod{n}$ pour tout entier a.*

Exercice : montrer que $561 = 3 \times 11 \times 17$ est un nombre de Charmichaël (indication : utiliser le théorème Chinois et le petit théorème de Fermat).

On sait depuis très récemment qu'il existe une infinité de nombres de Charmichaël (et on possède même des estimations sur leur répartition). On peut montrer sans trop de mal que tout nombre de Charmichaël a tous ses facteurs premiers distincts, et en a au moins trois.

Le petit théorème de Fermat ne permet donc pas de tester le caractère premier d'un nombre. On peut néanmoins déduire du théorème d'Euler une condition nécessaire et suffisante pour qu'un nombre soit premier, appelé *test de primalité de Lucas-Lehmer* :

Test 3.5. *Supposons que p soit tel que pour tout facteur premier q de $p-1$ il existe un entier a (pouvant dépendre de q) tel que $a^{p-1} \equiv 1 \pmod{p}$ et que $a^{(p-1)/q} \not\equiv 1 \pmod{p}$; alors p est premier.*

Démonstration (de la condition suffisante) : Il suffit de montrer que $\varphi(p) = p-1$. Si ce n'est pas le cas, comme néanmoins $\varphi(p) \leq p-1$, il existe un premier q et $r \geq 1$ tels que q^r divise $p-1$ mais pas $\varphi(p)$. Si s est l'ordre de $\bar{a}$ dans U_p, alors par hypothèse s divise $p-1$, mais ne divise pas $\frac{p-1}{q}$; par conséquent, q^r divise s, et donc q^r divise $\varphi(p)$ (théorème de Lagrange), une contradiction.

Remarque : On peut montrer que la condition décrite ci-dessus est aussi *nécessaire* pour que p soit premier. Plus précisément, si p est premier, il existe un a tel que $a^{p-1} \equiv 1 \pmod{p}$ et $a^t \not\equiv 1 \pmod{p}$ pour tout $t < p-1$.

Venons-en à présent à un résultat sans intérêt théorique, mais qui sera important dans la méthode cryptographique que nous allons décrire :

Résultat 3.6. *Soit $r \geq 2$ un naturel, s un naturel premier avec $\varphi(r)$, et a un naturel premier avec r. Si $b \equiv a^s \pmod{r}$ alors $a \equiv b^t \pmod{r}$ où t est un naturel congru à $s^{\varphi(\varphi(r))-1}$ modulo $\varphi(r)$.*

Démonstration : On a $t \equiv s^{\varphi(\varphi(r))-1} \pmod{\varphi(r)}$ donc $st \equiv s^{\varphi(\varphi(r))} \equiv 1 \pmod{\varphi(r)}$ (d'après le théorème d'Euler). Par conséquent, $st = 1 + u\varphi(r)$ pour un certain entier u. Dans ces conditions,

$$b^t \equiv a^{st} = a^{1+u\varphi(r)} = a(a^{\varphi(r)})^u \equiv a \pmod{r}$$

d'après le théorème d'Euler. Ce qui démontre le résultat.

On peut enfin décrire sommairement le principe du test de primalité de Miller-Rabin. Pour cela, on constate que si $a^2 \equiv 1 \pmod{p}$ avec p premier, alors $a \equiv \pm 1 \pmod{p}$, car le polynôme $X^2 - 1$ admet deux racines (éventuellement confondues) dans tout corps, à savoir 1 et -1. Si $2 \leq b \leq p-2$, alors $b^{p-1} \equiv 1 \pmod{p}$, donc $b^{(p-1)/2} \equiv \pm 1 \pmod{p}$; si c'est $+1$, on a $b^{(p-1)/4} \equiv \pm 1 \pmod{p}$, et ainsi de suite. On a ainsi le

Test 3.7. *Soit $p > 2$ un naturel. Ecrivons $p - 1 = 2^z t$ avec t impair, et soit $2 \leq b \leq p-2$. On calcule successivement les résidus de $b^t, b^{2t}, \ldots, b^{p-1}$ modulo p. Si p est premier, le dernier nombre ainsi obtenu avant d'obtenir 1 doit être -1. Ainsi, si on obtient un nombre différent de -1, suivi de 1, ou si on n'obtient jamais 1, alors on sait avec certitude que p est composé. Sinon, p est "probablement premier". Un nombre composé qui passe le test en étant déclaré "probablement premier" s'appelle un "pseudo-premier fort dans la base b".*

Rabin a montré qu'un nombre p composé est pseudo-premier dans au plus $\lfloor \frac{p}{4} \rfloor$ bases. Par conséquent, un nombre composé a moins d'une chance sur quatre de passer le test en étant déclaré "probablement premier", et en faisant prendre différentes valeurs à b, on peut réduire arbitrairement la probabilité d'erreur.

4. La cryptographie à clefs publiques : motivations et réalisation.

Il est relativement facile de concevoir un code secret dont le décryptage est *rigoureusement impossible* par qui ne possède pas la clef : il suffit pour cela de choisir une base r, et de tirer au hasard une suite (k_i) d'entiers entre 0 et $r-1$. Le message sera d'abord transformé en une suite (M_i) d'entiers entre 0 et $r-1$; pour coder, il suffit d'appeler (C_i) l'unique entier entre 0 et $r-1$ qui est congru à $M_i + k_i$ modulo r. Pour déchiffrer, connaissant C_i et k_i, il suffit de remarquer que M_i est l'unique entier entre 0 et $r-1$ qui est congru à $C_i - k_i$ modulo r. Si les k_i ont été choisi complètement au hasard, les C_i sont eux-aussi complètement aléatoires, c'est-à-dire pour ceux qui ne connaissent pas les k_i, et il est donc impossible de décrypter le message sans connaître la clef (k_i). L'inconvénient de ce système est la transmission de la clef : celle-ci est aussi longue que le message, et ne peut pas être utilisée deux fois (sans quoi elle n'est plus inviolable). Si N utilisateurs doivent pouvoir communiquer les uns avec les autres en secret, il faut alors disposer de $\frac{N(N-1)}{2}$ clefs, ce qui est bien souvent inacceptable. Du reste, si on peut transmettre la clef en sécurité, on ne voit pas pourquoi on ne pourrait pas transmettre le message lui-même par le même canal !

Le principe de la *cryptographie à clefs publiques* est d'utiliser, non pas un code pour chaque paire d'utilisateurs, mais un pour chaque utilisateur : chacun doit pouvoir *chiffrer* un message, mais seul son destinateur peut le *déchiffrer*. C'est-à-dire que la clef de chiffrement est publique, et la clef de déchiffrement est gardée secrète, c'est-à-dire gardée uniquement par le destinataire du message, qui ne la communique à personne. Il n'y a donc plus de canal protégé de transmission des clefs à maintenir : la clef de chiffrement est divulguée à tous, et celle de déchiffrement ne l'est à personne. On remarquera ainsi que même l'émetteur du message n'est pas capable de décoder ce que lui-même envoie, seul le récepteur en étant capable. De cette manière, tous utiliseront la même clef de codage pour le même déstinataire. Mathématiquement, cela se traduit en disant que pour chaque destinataire λ il y a une clef de codage U_λ publique et une clef de décodage V_λ privée, qui sont des bijections réciproques (entre l'ensemble des messages et l'ensemble des codes), et que tous connaissent U_λ tandis que seul λ connaît V_λ. Pour produire cet effet, le système que nous décrivons, appelé "système RSA" en l'honneur de Ronald Rivest, Adi Shamir et Léonard Adleman, utilise le fait suivant : il est facile de produire des grands nombres ("grand" signifiant de l'ordre de grandeur de 10^{100}) dont on connaît la décomposition en facteurs premiers, tandis qu'il est très difficile de décomposer en facteurs premiers un grand nombre donné. A titre d'exemple, la complexité des meilleurs algorithmes de factorisation des entiers à n chiffres (c'est-à-dire le temps qu'il faut, exprimé en unités quelconques, pour factoriser un entier à n chiffres) est $O(e^{c\sqrt{n \ln n}})$ pour une certaine constante c. En comparaison, la complexité d'un algorithme probabiliste pour déterminer le caractère premier d'un entier de n chiffres, tel que l'algorithme de Miller-Rabin, est $O(n^3)$ ou même mieux si un bon algorithme de multiplication est utilisé. On peut ainsi déterminer le caractère premier ou composé de nombres allant jusqu'à mille cinq cents chiffres, mais on ne parvient à factoriser que des nombres d'environ cent vingt chiffres (en se donnant beaucoup de mal).

5. Description de la méthode.

Le destinataire doit commencer par créer la clef de codage, qu'il rendra publique, et la clef de déchiffrement, qu'il gardera secrète. Pour cela, il détermine deux nombres premiers p et q grands (c'est-à-dire de l'ordre de grandeur de 10^{50}, ou même plus de nos jours) tels que la décomposition en facteur premiers de $p-1$ et $q-1$ soit connue. Pour des raisons de sécurité, on aura intérêt à ce que cette décomposition comporte au moins un grand facteur premier. Ceci se fait par exemple en tirant des nombres p_1, p_2 au hasards jusqu'à ce que p_1, p_2 et $p := 2p_1p_2 + 1$ soient premiers, et de même pour q. (Notons que la probabilité qu'un nombre impair aléatoire entre 1 et 10^{50} soit premier est environ 1 chance sur 58. Il existe un théorème appelé "théorème des nombres premiers" qui affirme que la probabilité qu'un entier tiré au hasard entre 1 et n soit premier est asymptotiquement équivalente à $\frac{1}{\ln n}$ lorsque n est grand.)

Une fois p et q trouvés, le destinataire calcule $r = pq$ et le rend public. r est alors un entier d'une centaine de chiffres, dont la décomposition en facteurs premiers est connue du destinataire, mais pas du public. D'autre part, le destinataire connaît $\varphi(r) = (p-1)(q-1)$, et connaît aussi la décomposition en facteurs premiers de $\varphi(r)$, donc aussi $\varphi(\varphi(r))$.

Le destinataire choisit alors s grand, premier avec $\varphi(r)$ et le rend public. Et en secret, il calcule $t \equiv s^{\varphi(\varphi(r))-1} \pmod{\varphi(r)}$.

Pour envoyer un message au destinataire, on représente ce message sous la forme d'une suite finie (M_i) d'entiers compris entre 0 et $r-1$. Les M_i ont une probabilité d'à peu près 10^{-50} d'être multiples de p ou de q, et on peut donc raisonnablement supposer qu'ils sont premiers avec r. L'auteur du message utilise alors la clef de codage (r, s) rendue publique, et calcule C_i, qui est l'unique entier entre 0 et $r-1$, congru à M_i^s modulo r. Pour décoder le message, le destinataire utilise t, qui est secret : M_i est l'unique entier entre 0 et $r-1$ congru à C_i^t modulo r, en vertu du résultat 3.6.

On voit qu'ainsi pour chiffrer un message il suffit de connaître r et s, qui constituent la clef publique, tandis que pour le déchiffrer, il faut connaître t, et pour le calculer, un hypothétique individu mal intentionné devrait au préalable calculer $\varphi(r)$, ce qui nécessite la décomposition en facteurs premiers de r, laquelle est très difficile à obtenir avec les moyens actuels si r est assez grand.

On pourrait penser que pour calculer $C_i \equiv M_i^s \pmod r$ il faudrait effectuer s multiplications, ce qui avec s de l'ordre de grandeur de 10^{100} est totalement inacceptable. Ce n'est pas le cas : si on écrit s en binaire sous la forme

$$s = s_0 + s_1 \times 2^1 + \cdots + s_\nu 2^\nu$$

(avec $\nu \simeq 300$, $s_j \in \{0, 1\}$), on va calculer successivement $M_i^{2^j}$ modulo r pour $0 \leq j \leq \nu$, ce qui ne demande que ν multiplications, puis en multipliant entre eux ceux des $M_i^{2^j}$ pour lesquels $s_j = 1$, on obtient M_i^s. En tout, on a effectué au plus 2ν multiplications (chacune accompagnée de la réduction modulo r).

6. Authentification de la provenance du message.

L'inconvénient de la méthode de cryptographie à clefs publiques est que chaque émetteur possible du message possède la même clef de chiffrement, et que par conséquent il est facile de se faire passer pour un autre. Pour remédier à cette difficulté, le système RSA propose une solution simple.

On constate en effet que pour chaque utilisateur λ, on dispose de deux bijections réciproques, U_λ l'application "codante", connue de tous, qui transforme un message en son code, et V_λ l'application "décodante", connue de λ seul, et qui reconstruit le message à partir du code. L'idée, pour fournir une preuve d'authenticité d'un message, est d'utiliser ces applications *à l'envers* : en effet, V_λ peut servir de codage, et U_λ de décodage, auquel cas seul λ peut effectuer le codage, tandis que tous peuvent décoder.

Si λ veut envoyer un message à μ, il procède alors ainsi : il code, en utilisant V_λ, un petit message, comportant entre autres son nom, le nom de μ, et la date, et il rajoute ce message codé, appelé signature, à la fin du message principal, qu'il code alors avec U_μ. Le code ainsi obtenu est envoyé à μ. Celui-ci décode alors le message avec V_μ ; pour authentifier le message, il ne reste qu'à décoder la signature avec U_λ : si les informations contenues dans celle-ci coïncident avec les informations du message, alors le message est authentique, puisque seul λ connaît V_λ, et donc aurait pu produire la signature. De plus, μ ne peut pas "copier" la signature de λ ainsi obtenue, car celle-ci contient le nom de μ !

7. Un exemple en MAPLE.

Les quelques lignes de MAPLE qui suivent ont pour rôle de tirer au hasard une clef de codage/ déchiffrement, et de s'en servir pour coder un petit message :

```
proc() global p1;
  p1 := 0; while not isprime(p1) do p1 := 2*rand(10^30)()+1; od;
end():
p1;
proc() global p,p2;
  p2 := 0; p := 0; while not (isprime(p) and isprime(p2))
  do p2 := 2*rand(10^30)()+1; p := 2*p1*p2+1; od;
end():
p2;
p;
proc() global q1;
  q1 := 0; while not isprime(q1) do q1 := 2*rand(10^30)()+1; od;
end():
```

```
q1;
proc() global q,q2;
  q2 := 0; q := 0; while not (isprime(q) and isprime(q2))
  do q2 := 2*rand(10^30)()+1; q := 2*q1*q2+1; od;
end():
q2;
q;
r := p*q; phir := (p-1)*(q-1); phiphir := 2*(p1-1)*(p2-1)*(q1-1)*(q2-1);
proc() global s;
  s := 0; while igcd(s,phir)<>1 do s := 2*rand(10^120)()+1; od;
end():
s;
t := Power(s,phiphir-1) mod phir;
encode := proc(MSG) Power(MSG,s) mod r; end:
decode := proc(COD) Power(COD,t) mod r; end:
encode(11112222333344445555666677778888899990000);
decode(");
```

8. Conclusion en quelques remarques.

Le système RSA de cryptographie à clefs publiques, que nous venons de décrire, n'est certes pas le seul possible ; c'est toutefois de loin le meilleur connu. Jusqu'en 1982, un autre système, le codage du *sac à dos* paraissait plus prometteur ; cependant, en 1982, ce dernier a subi une attaque importante de la part d'Adi Shamir, et en 1984 il apparaissait clairement que ce système n'était plus viable (Ralph Merkle a d'ailleurs dû payer $1000 à la suite d'un pari sur la sécurité du système du sac à dos). Le système RSA a lui-aussi subi des attaques ; celles-ci concernent principalement la possibilité de factoriser des grands entiers ; en raison du développement des algorithmes de factorisation rapide et de la puissance des gros systèmes informatiques, il apparaît maintenant qu'il faille choisir les nombres premiers intervenant dans le codage comme ayant au moins 200 chiffres, sinon plus, et vérifiant des conditions supplémentaires (portant notamment sur la décomposition en facteurs premiers de $p-1$ et $p+1$). Toutefois, aucun coup fatal n'a encore véritablement été porté au système, comme le serait par exemple l'existence d'un moyen de décodage en temps polynomial.

La morale à tirer de l'existence des systèmes de cryptographie à clef publiques basées sur des théorèmes d'arithmétique est probablement une belle leçon d'humilité pour les mathématiciens. En effet, Karl Friedrich Gauss avait affirmé que les mathématiques étaient la reine des science, et la théorie des nombres la reine des mathématiques puisqu'elle n'avait aucune application pratique. Il est ironique qu'à peine un siècle après la mort de Gauss la théorie des nombres se soit avérée aussi utile justement.

Naturellement, l'arithmétique ne s'arrête pas avec la fonction d'Euler, loin de là. Ceux qui veulent aller plus loin dans le sujet peuvent consulter

⋆ Paulo Ribenboim : “Nombres premiers : mystères et records.”

Et pour les plus audacieux, qui savent lire l'anglais, on peut signaler

⋆ Alan Baker : “A concise introduction to the theory of numbers.”