# Chapitre 6

# Les bases de la cryptographie et le code RSA

## 6.1 Introduction au principe de cryptographie

Le but de la cryptographie est de cacher le contenu d'un message. Il y a pour cela différentes possibilités. Voici deux méthodes utilisées (parmi tant d'autres) :
- Cacher le message (dans une image, par exemple) : stéganographie.
- Rendre le message incompréhensible en transformant un texte clair en un cryptogramme : chiffrement.

**Deux exemples de chiffrement**

1. Le carré de Polybe est la clé qui a permis de créer le premier système de chiffrement polygraphique connu.

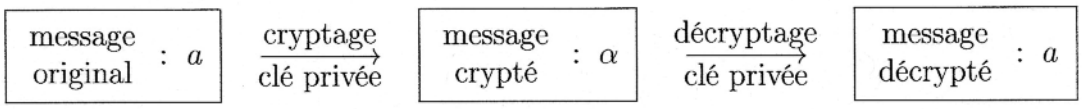
| ↗ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | a | b | c | d | e |
| 2 | f | g | h | ij | k |
| 3 | l | m | n | o | p |
| 4 | q | r | s | t | u |
| 5 | v | w | x | y | z |

   Ainsi, le mot SECRET sera codé par 43 15 13 42 15 44. Pour coder et décoder, il faut que celui qui envoie le message et que celui qui le reçoit aient tous les deux la même clé (ici, le carré de Polybe).

2. La machine Enigma était la clé du cryptage allemand durant la deuxième guerre mondiale (le film «U571» s'est inspiré du fait qu'une machine Enigma a été dérobée aux Allemands durant l'assaut d'un de leur sous-marin, fournissant ainsi aux Alliés la clé pour décoder les messages allemands).

Les deux systèmes de cryptographie ci-dessus utilisent une clé unique qui appartient au codeur et au décodeur. Si une tierce personne parvient à s'emparer de la clé, elle serait en mesure d'intercepter et de décoder les messages, ou même d'usurper l'identité d'une des deux autres personnes.

Ces méthodes sont dites à clé privée. Elles sont symétriques, car celui qui reçoit le message utilise la même clé que celui qui l'envoie.

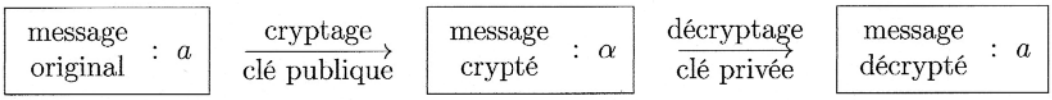



On voit qu'en utilisant la même clé, celui qui reçoit le message peut envoyer une réponse !

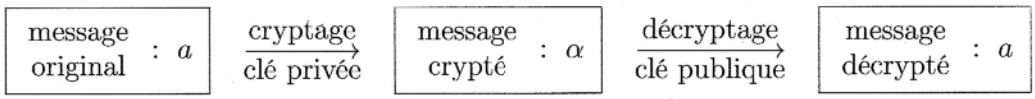
Vers la fin des années 1960, on a découvert des systèmes cryptographiques à clé publique. Ce système utilise deux clés : l'une privée, l'autre publique. La clé publique peut être connue de tout le monde, tandis que la clé privée n'est connue que d'une personne (donc elle est plus facile à protéger que dans les méthodes à clé privée vue précédemment).

**Deux utilités des systèmes à clé publique**

1. Tout le monde peut envoyer un message qui ne pourra être lu que par celui qui a la clé privée.




2. Tout le monde peut recevoir un message qui n'a pu être écrit que d'une seule personne (authentification de l'auteur, signature électronique).




Bien évidemment, on peut combiner ces deux utilités pour avoir un message qui ne peut être lu que par une seule personne et qui n'a pu être écrit que par une unique personne !

## 6.2 Le système de cryptographie RSA

Le plus célèbre et le premier des systèmes de cryptographie à clé publique est le système RSA (Ronald Rivest, Adi Shamir et Leonard Adleman). Entre autres, ce système est à la base des méthodes de paiements par Internet !

### 6.2.1 Mise en place

1. On choisit deux nombres premiers distincts $p$ et $q$ suffisamment grands. On calcule le produit $n = pq$. Généralement, on utilise des nombres premiers d'environ 300 chiffres (en 1024 bits, on forme des chaînes de longueur $2^{1024} \cong 1.79 \cdot 10^{308}$).
2. On choisit un nombre $e$ premier à $\varphi(n)$, c'est-à-dire que $\text{pgcd}(e, \varphi(n)) = 1$.
3. On cherche un nombre $d \in \mathbb{N}$ qui correspond à l'inverse de $e$ modulo $\varphi(n)$. Autrement dit, on cherche $d \in \mathbb{N}$ tel que $d \cdot e \equiv 1 \pmod{\varphi(n)}$.

   Le nombre $d$ existe, car $e$ est premier à $\varphi(n)$.

**La clé privée**

La clé privée est composée des nombres $p$, $q$ et $d$.

**La clé publique**

Les nombres $n$ et $e$ sont mis à disposition dans un annuaire. Les nombres $p$, $q$ et $d$ sont gardés secrets.

### 6.2.2 Sûreté du système RSA

Les informations données dans la clé publique ne permettent pas de retrouver la clé privée, car il est actuellement impossible de trouver $p$ et $q$ si on connaît le nombre $n$ en un temps raisonnable (pour autant que $n$ soit suffisamment grand). Or, pour trouver $d$, il faut connaître $\varphi(n)$ qui ne peut être connu qu'à l'aide de $p$ et de $q$.

En 2005, il y avait une récompense de 200'000\$ pour celui qui réussissait à factoriser un nombre $n$ relativement grand. En 2009, les récompenses ne sont plus que de 10'000\$ pour un nombre de 32-bits. La dernière factorisation a été réussie en 2007 (pour un nombre à peine plus petit) et a nécessité environ 1757 jours de calculs. Pour plus de précision (et aussi des nouvelles plus fraîches), le lecteur consultera le site `http://www.rsa.com` sous l'onglet «historical», puis sous la rubrique «cryptographic challenges».

### 6.2.3 Théorème RSA

Soit $p$ et $q$ deux nombres premiers distincts et $n = pq$. Si $e$ est un nombre premier à $\varphi(n)$ et si $d$ est son inverse modulo $\varphi(n)$, alors pour tout entier $a$ ($a < n$), on a :

$$(a^e)^d \equiv (a^d)^e \equiv a \pmod{n}$$

**Preuve**

Puisque $p$ et $q$ sont des nombres premiers distincts, on sait que $n = \text{ppcm}(p, q)$. Ainsi, pour montrer que $a^{de} \equiv a \pmod{pq}$, il suffit de montrer que $a^{de}$ est solution du système chinois suivant

$$\begin{cases} x \equiv a \pmod{p} \\ x \equiv a \pmod{q} \end{cases}$$

En effet, dans ce cas $a^{de}$ serait solution du système, au même titre que $a$. Par unicité de la solution modulo $pq$, on saurait que $a^{de} \equiv a \pmod{pq}$.

On ne va démontrer que $a^{de} \equiv a \pmod{p}$, car pour l'autre, on reprend les mêmes arguments en échangeant les rôles de $p$ et de $q$.

Par hypothèse, on sait que $de \equiv 1 \pmod{\varphi(n)}$. Ainsi, il existe $k \in \mathbb{Z}$ tel que

$$de = 1 + k\,\varphi(n) = 1 + k(p-1)(q-1)$$

Donc, comme $a^{p-1} \equiv 1 \pmod{p}$ grâce au théorème de Fermat, on obtient

$$a^{de} = a^{1+k(p-1)(q-1)} = a^1 \cdot a^{k(p-1)(q-1)} = a \cdot (a^{p-1})^{k(q-1)} \equiv a \cdot 1 \equiv a \pmod{p}$$

□

### 6.2.4 Méthode de codage et de décodage

Le processus de codage et de décodage se déroule en plusieurs étapes.

1. Préparation du message à encoder.

   Le message doit être éventuellement décomposé en blocs. À chaque bloc, on va associer un nombre strictement compris entre 1 et $n$. Ce sont ces nombres que l'on va encoder à l'aide du système RSA.

   La façon dont on associe des nombres à chaque bloc est très variable. Cela peut suivre une démarche assez simple (comme on le verra en exercice) ou un procédé bien plus complexe où un autre cryptage pourrait être utilisé.

2. Encodage avec RSA

   Le message est maintenant une suite de nombres strictement compris entre 1 et $n$. Pour coder ce message on élève chacun des nombres le composant à la puissance $d$ ou à la puissance $e$ selon si on veut encoder avec la clé privée ou publique.

3. Décodage avec RSA

   Le message est maintenant une suite de nombres strictement compris entre 1 et $n$. Pour décoder ce message on élève chacun des nombres le composant à la puissance $e$ ou à la puissance $d$ selon si on veut décoder avec la clé privée ou publique.

   Bien sûr, si le message a été encodé avec la clé privée, il faut le décoder avec la clé publique, et vice-versa.

   On retombe bien sur les nombres qu'on avait avant l'encodage, puisque le théorème nous dit que $(a^e)^d \equiv (a^d)^e \equiv a \pmod{n}$.

4. Reconstitution du message original.

   Il faut effectuer la démarche inverse de celle effectuée lors de la préparation du message à encoder.

#### Remarque banale mais importante

Lors de la préparation du message, il ne faut pas associer chaque lettre à un nombre (en code `ASCII`, entre 1 et 255), car une simple analyse de fréquences permettra de casser le code (bien sûr, si le message est suffisamment long pour qu'une telle analyse soit pertinente).