# Les codes secrets et l'arithmétique

***Michel Waldschmidt***

http://www.math.jussieu.fr/~miw/articles/pdf/ArithmetiqueCryptographie.pdf

## *Messages secrets*

De tout temps les hommes et les femmes ont voulu échanger des messages secrets. Pour cela l'expéditeur doit transformer l'information de telle manière que seul le destinataire puisse déchiffrer le message. On parle alors de *cryptage*. Un principe général est que le destinataire dispose d'une information indispensable pour retrouver, à partir du message crypté, celui qui a été envoyé. Les procédés utilisés de nos jours reposent sur des résultats d'arithmétique.

## *Histoire brève*

Les méthodes anciennes de cryptographie consistaient à effectuer différentes opérations élémentaires sur les symboles du document initial, par exemple remplacer chaque lettre d'un texte par une autre lettre, selon une formule convenue à l'avance entre l'expéditeur et le destinataire. Le nouveau texte ainsi produit semble illisible, on dit que c'est un message *chiffré*, même s'il est composé de lettres (et le *chiffre* est souvent le nom donné aux organismes chargés de telles opérations). On dit aussi que le message est crypté. On attribue à Jules César la méthode de cryptographie obtenue en décalant les lettres de l'alphabet d'un nombre fixé de places. Des variantes ont été introduites, mais on sait maintenant qu'aucune n'est vraiment fiable: il est facile de retrouver le message initial à partir du message chiffré sans connaître à l'avance la recette utilisée pour un tel chiffrement. On effectue pour cela une analyse statistique des lettres qui apparaissent : la fréquence d'apparition de chaque lettre donne une information révélatrice. Cette technique permettait déjà au IXe siècle à Abu Youssouf Ya qub Ishaq Al Kindi de vérifier l'authenticité de textes saints de l'Islam : il comparaît les fréquences des différentes lettres dans les textes authentifiés et dans ceux dont il n'était pas sûr, pour déceler s'il y avait des différences significatives. Ainsi, pour déchiffrer un texte dont l'original est en français, si chaque lettre a été remplacée par une autre suivant une règle immuable, on décèlera quelle est la lettre qui remplace le *e* en cherchant celle qui apparaît le plus souvent. Bien entendu une telle méthode à des limites: elle ne fonctionnera pas avec le texte de Georges Perec *La Disparition* qui ne comporte aucun *e*!

Au XIIIe siècle, dans sa *Lettre sur les œuvres d'art et sur les nullités de la magie*, Roger Bacon donne sept méthodes pour crypter des messages.

Au XVIe siècle, le diplomate français Blaise de Vigenère était aussi un écrivain, un alchimiste et un cryptographe.

L'inventeur du principe de fonctionnement d'un ordinateur, C. Babbage (1791-1871), a mis en lumière l'intérêt des statistiques sur la fréquence des lettres pour le déchiffrage des

messages cryptés. À la même époque, on peut mentionner un bel exemple de déchiffrage, celui de l'interprétation des hiéroglyphes égyptiens par J-F. Champollion, problème qui avait défié les générations antérieures. Après avoir compris que les inscriptions sur la pierre de Rosette correspondaient au même texte en trois langues différentes, il s'est servi de ses vastes connaissances des langues anciennes pour parvenir à ses fins.

Après la guerre de 1870, le gouvernement français a pris conscience de l'influence qu'avait eue la supériorité allemande en matière de communication. Il a été alors décidé de créer des centres militaires (à Coëtquidan et Montoire) pour améliorer l'utilisation des pigeons voyageurs. Au même moment débutaient les recherches théoriques sur les ondes électromagnétiques, qui allaient donner naissance à la radio et aux moyens modernes de transmission des données.

Dans un article visionnaire intitulé *La cryptographie militaire* et publié dans le *Journal des Sciences Militaires* en 1883, A. Kerckhoffs introduit un certain nombre de principes qui sont toujours d'actualité. L'un d'eux est que toute méthode de chiffrement doit être supposée connue par l'ennemi: la sécurité du système doit dépendre uniquement du choix de *clés*, qui doivent être changées régulièrement. Les clés en question sont de nos jours des suites (finies mais assez longues) de 0 et de 1; la recette pour crypter le message fait intervenir cette clé, changer seulement la clé en gardant la même recette pour crypter le texte produit un message chiffré complètement différent. En voici un exemple.

Le *téléphone rouge* (qui était plutôt un fax) entre la Maison-Blanche et le Kremlin au temps de la guerre froide utilisait un dispositif appelé *masque jetable* inventé par G.Vernam en 1917. Ce procédé fait intervenir une clé, qui est une suite de 0 et de 1 permettant à la fois de chiffrer et de déchiffrer le message; elle est transmise séparément (par exemple par la voie diplomatique) et ne sert qu'une fois. Le message à envoyer est aussi une suite de 0 et de 1, aussi longue que la clé. Pour chiffrer, on remplace chacun des chiffres du message en clair par sa somme avec celui de la clé situé à la même position, en utilisant une loi d'addition un peu particulière, sans retenue :

$$0+0=0, \qquad 0+1=1, \qquad 1+0=1, \qquad 1+1=0.$$

Par exemple si le message en clair est

$$0,\ 1,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1$$

et la clé

$$0,\ 0,\ 1,\ 1,\ 0,\ 1,\ 0,\ 0,\ 1$$

le message chiffré sera

$$0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 1,\ 0,\ 0$$

La même opération (appelée *addition booléenne*) à partir du message chiffré et de la clé redonne le message en clair.

Le décodage des messages envoyés par la machine *Enigma* de l'armée allemande pendant la seconde guerre mondiale a été réalisé par une équipe comportant notamment le mathématicien A.Turing, grâce au premier ordinateur programmable électronique mis au point par M. Newman.

Ce sont les travaux précurseurs de C. Shannon vers 1950 qui fondent la théorie mathématique moderne des communications et de la transmission des données. Le principe du chiffrement à l'aide de ce qu'ils ont appelé une *clé publique* a été proposé par W. Diffie et M.E. Hellman en 1976. Sa réalisation en 1978 par R.L. Rivest, A. Shamir et L.M. Adleman a

donné naissance au système RSA pour encrypter les messages, qui est le plus utilisé de nos jours.

## *Le protocole d'échange de valises*

Un exemple élémentaire de transfert de données sécurisé est le suivant: Alice veut envoyer une valise à Bob sans que Charlie, qui va la transporter, puisse l'ouvrir. Elle possède un cadenas et une clé, Bob aussi a un cadenas et une clé, mais ils ne sont pas compatibles (sinon ce serait trop simple). Comment faire? Avec ces données minimales, il n'est pas difficile de trouver la solution suivante. Alice veut envoyer la valise dont le contenu est confidentiel, elle doit donc la fermer à clé avant de la donner à Charlie. Quand Bob la reçoit, il ne peut pas l'ouvrir; il pourrait la renvoyer telle quelle à Alice mais cela ne servirait à rien. Alors il met son cadenas en plus de celui d'Alice et confie à Charlie la valise fermée par deux cadenas. Quand Alice la reçoit de nouveau, avec les deux cadenas, elle enlève le sien et renvoie finalement la valise à Bob qui peut utiliser sa clé et accéder au contenu (VOIR LES DESSINS).

Ce protocole peut se traduire en termes arithmétiques. On remplace Charlie par des connections électroniques et la valise par un message numérique, donc une suite de 0 et de 1. On coupe ce message en petits morceaux de même longueur, donc ce que l'on veut envoyer est un nombre (représenté par son développement binaire) inférieur à une borne fixée. On remplace chaque cadenas et chaque clé par des nombres. Mettre le cadenas ou l'enlever avec la clé consiste à faire des opérations sur le nombre qui représente le message. Ces opérations font intervenir les clés et les cadenas. Quand on applique successivement à un nombre un cadenas et la clé correspondante, on doit retrouver le nombre initial. Si on ne connaît pas la clé, on ne doit pas pouvoir ouvrir le cadenas.

En cryptographie, le cadenas et la clé jouent le plus souvent le même rôle, comme des boutons servant d'interrupteur: quand le courant passe et qu'on appuie, on coupe le courant, quand il est coupé et qu'on appuie, le courant passe. On l'a déjà vu avec le masque jetable: quand on fait la somme booléenne du message en clair et de la clé, on obtient le message chiffré, quand on ajoute au message chiffré la clé, on retrouve le message en clair. Ajouter deux fois la clé est une opération blanche, parce que 0+0=0 et 1+1=0.

## *Vérification de l'identité*

Dans de multiples circonstances, une personne doit prouver son identité, ou tout au moins prouver qu'elle connaît un mot de passe ou un code secret qu'elle est supposée être la seule à savoir. Prenons l'exemple de l'usager qui veut retirer de l'argent dans un distributeur automatique d'une banque en utilisant sa carte à puce. Il est le seul à connaître son code personnel: dans une transaction sécurisée la banque ne connaît pas le code secret. Il n'y a pas de fichier informatique avec tous les codes secrets, ce serait trop risqué. Comment la banque peut-elle vérifier que l'usager connaît ce code? Pour assurer la confidentialité, il faut prévoir que des tiers peuvent intercepter toutes les communications qui vont être échangées. L'usager doit donc faire savoir à la banque qu'il connaît son code secret sans le lui révéler: à la fin de la transaction, la banque ne connaît toujours pas le code secret, elle s’est contentée de s'assurer que l'utilisateur le connaît.

Les développements récents du sujet permettent d'apporter des réponses à cette question. Le premier pas du processus est le suivant : la banque va poser une question à l'usager, et cette question sera différente pour chaque transaction. Le premier message envoyé par la banque est choisi au hasard (on dit qu'il est *aléatoire*). Ici encore il est représenté par un nombre entre 0 et une borne fixée; la puce va faire un calcul à partir de ce nombre en utilisant le code secret, et c'est le résultat qui va être envoyé à la banque. Il faut donc trouver une formule qui permette, connaissant le message initial et le message de retour, d'en déduire que le code secret est bon.

Ajoutons que l'utilisateur possède non seulement un code secret (une clé) mais aussi un code public (un cadenas) que tout le monde connaît. Donc ce que va faire la banque est une opération à partir du message renvoyé en utilisant le code public, et vérifier que l'on retrouve bien le message initial (comme dans l'échange de valises entre Alice et Bob: quand on utilise la clé et la serrure successivement on retrouve l'état initial - ici tout se passe comme si on utilisait d'abord la clé et ensuite la serrure!).

Le problème ainsi posé est celui de la *cryptographie à clé publique*, concept dont nous avons vu qu'il avait été introduit en 1976 par Diffie et Hellmann.

## *Une fonction trappe*

Si une bille qui roule tombe dans un trou, elle ne remontera pas toute seule. Le principe de la trappe est que certains trajets sont plus faciles à faire dans un sens que dans l'autre: il ne suffit pas de savoir comment on a parcouru un trajet pour être capable de le faire dans l'autre sens. Dans le même ordre d'idées, un labyrinthe peut représenter un problème très difficile ; si on vous donne une solution, il est très facile de vérifier si elle est bonne.

Une situation similaire se présente en arithmétique: si je vous donne deux nombres à multiplier, vous n'aurez pas trop de mal à trouver leur produit (dans la mesure où il n'y a pas trop de chiffres); en revanche, si je vous fournis le résultat, trouver les deux nombres initiaux peut être nettement plus difficile. Par exemple si je vous dis que j'ai obtenu 2 047 en multipliant deux nombres plus petits (mais supérieurs à 1), il vous faudra un peu de temps pour trouver que ces deux nombres sont 23 et 89.

Les méthodes de cryptage modernes reposent sur des questions de mathématique pour lesquelles aucune méthode de résolution efficace n'est connue. À la question de décomposer un nombre en produit d'entiers comme dans l'exemple de 2 047, les ordinateurs actuels sont capables de donner la réponse en un temps raisonnable pour des nombres ne dépassant pas 150 ou 200 chiffres. Au-delà, le temps de calcul nécessaire est prohibitif.

Le système RSA utilise aussi une sorte de trappe qui permet d'inverser l'opération d'encryption, c'est-à-dire qui permet de décrypter ce qui a été crypté au moyen d'un entier N si on connaît l'écriture de N en facteurs premiers. La mise au point de ces méthodes repose sur des outils d'arithmétique (théorie des nombres premiers) qui font actuellement l'objet de nombreuses recherches de pointe dans divers laboratoires, notamment en France.

**Conclusion.** Les méthodes qui, comme le protocole RSA, font intervenir des outils de la théorie des nombres élaborée, apportent une grande leçon: des recherches mathématiques (sur les nombres premiers notamment) tout à fait désintéressées peuvent se révéler, des années ou

des décennies plus tard, cruciales pour telle ou telle application; et ce de manière imprévisible. Dans son livre *L'Apologie d'un Mathématicien*, le grand théoricien des nombres britannique G. H. Hardy (1877-1947), qui était un fervent pacifiste, se targuait de travailler dans un domaine parfaitement pur, l'arithmétique, et de n'avoir rien fait qui puisse être considéré comme «utile». Ses travaux étaient peut-être «inutiles» à son époque. De nos jours, l'arithmétique dont il était un éminent spécialiste, est présente dans de nombreux objets électroniques de la vie courante : cartes à puce, lecteurs de disques compacts ou de DVD par exemple.

**Références.** Pour en savoir beaucoup plus sur le sujet, on pourra lire le passionnant livre de Jacques Stern: *La science du secret*, Éd Odile Jacob, 1998. L'aspect historique est plus développé dans la traduction française de *The Code Book* de Simon Singh : *Histoire des codes secrets, de l'Égypte des Pharaons à l'ordinateur quantique* Le Livre de Poche 1999.

*Michel WALDSCHMIDT*
*8 rue Berlioz*
*91470 Limours*
e-mail: miw@math.jussieu.fr
URL http://www.math.jussieu.fr/~miw

DESSINS

Pierre de Rosette avec une légende qui indique les 3 langues et la façon dont Young et Champollion s'y sont pris.

Dessins de A B C
Légendes :

1. Alice veut envoyer une valise à Bob en la confiant à Charlie. Elle dispose d'un cadenas et d'une clé. Bob a aussi un cadenas et une clé, mais qui ne sont pas compatibles avec ceux d'Alice.
2. Alice ferme la valise avec le cadenas dont elle a la clé et la confie à Charlie qui la transmet à Bob.
3. Bob met son cadenas sur la valise qui est ainsi fermée à double tour ! Il sollicite encore Charlie qui retourne voir Alice.
4. Alice utilise sa clé pour enlever son cadenas et demande une dernière fois à Charlie de transmettre la valise à Bob.
5. Quand Bob reçoit la valise, elle n'est fermée qu'avec la clé dont il a le cadenas : il peut donc l'ouvrir.

Voir aussi le fascicule « L'Explosion des Mathématiques »
<http://smf.emath.fr/en/Publications/ExplosionDesMathematiques/?fr>