# La cryptographie

## 1 La cryptographie affine

### a/ Présentation

On associe à chaque lettre de l'alphabet numérotée par le nombre *x* de l'intervalle [ 0 ; 25 ], le nombre *y* défini par $y = ax + b$ où *a* et *b* sont deux nombres connus uniquement de l'émetteur et du récepteur.
Le couple ( *a* ; *b* ) s'appelle la clé du codage.

### b/ Création d'un « crypteur » affine

Ouvrez un tableur, créez un nouveau fichier principal qui vous suivra durant cette activité. Effectuez les manipulations suivantes dans la feuille 1 :

- Diminuez la largeur des colonnes D à AG
- Dans les cellules A2 et B2, entrez les nombres 19 et 3 qui correspondent aux coefficients a et b du codage.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | **a** | **b** | MESSAGE INITIAL | C | E | S | A | R |
| 2 | 19 | 3 | x | 2 | 4 | 18 | 0 | 17 |
| 3 | | | y = ax+b mod 26 | 15 | 1 | 7 | 3 | 14 |
| 4 | | | | | | | | |
| 5 | | | | | | | | |
| 6 | | | MESSAGE CODE | P | B | H | D | O |
| 7 | | | | | | | | |

- Dans la cellule D2, entrez =SI(D1="";"";CODE(D1)-65) et recopiez jusqu'en AG2. Expliquez la formule précédente sachant que CODE(lettre A) est égal à 65.
- Dans la cellule D3, entrez =SI(D2="";"";MOD($B$2+$A$2*D2;26)) puis recopiez jusqu'en AG3. Expliquez la formule.
- Enfin, dans la cellule D6 entrez =CAR(D3+65) puis recopiez jusqu'en AG6. Expliquez la formule sachant que CAR(65) représente la lettre A.
  Codez le message suivant: «CESAR UTILISE LE CODAGE AFFINE »

### c/ Un alphabet incomplet

Modifiez le crypteur précédent avec la clé ( 10 ; 6 ) puis avec la clé ( 12 ; 7 )
Quelle remarque peut-on faire ? Conjecturez la condition nécessaire sur le nombre *a* pour définir un crypteur correct.

Preuve de l'assertion précédente :

Supposons que $ax+b \equiv ax'+b \quad [26]$ c'est-à-dire que l'on a deux lettres identiques dans l'alphabet codé,

On peut alors écrire que alors $a(x-x') \equiv 0 \quad [26]$

et que $x-x' \equiv 0 \quad [26]$ à condition que a soit premier avec 26 (d'après le théorème de Gauss).

On en déduit alors que x = x' car x et x' sont deux entiers naturels inférieurs à 26. Si a n'est pas premier avec 26, que se passe-t-il ?

Le réflexe est alors de définir le pgcd noté *d* de *a* et 26 et les entiers *k* et *k'* tels que 26=*kd* et tel que *a* = *k'd*. On rappelle que les deux entiers *k* et *k'* sont premiers entre eux.
En s'aidant d'un crypteur incorrect, donnez des exemples de couples de nombres distincts qui donnent la même lettre codée et émettez une conjecture.

Plus généralement, si $ax+b \equiv ax'+b \ [26]$, alors $a(x-x') \equiv 0 \ [26]$ et on peut trouver un entier p tel que $a(x-x') = 26p$.

On peut donc écrire que $k'd(x-x') = kdp$ soit $k'(x-x') = kp$. Comme *k* et *k'* sont premiers entre eux, le théorème de Gauss permet d'affirmer que *x* - *x'* est divisible par *k*. Avec la clé ( 12 ; 7 ), *k* est égal à 13. Vérifiez sur votre crypteur que deux nombres congrus modulo 13 donnent la même lettre codée.

Il est inutile de décoder toutes les lettres de l'alphabet pour décoder un message. Soit la clé ( 19 ; 3 ), comment décrypter la lettre M ?

M est repéré par le nombre 12. On cherche donc *x* tel que 19*x* + 3 soit congru à 12 modulo 26. Montrer qu'alors *x* vérifie l'équation diophantienne 26*k* – 19 *x* = 17

Résoudre cette équation par la méthode habituelle :

| Solution particulière de 26*k*-19*x* = 1 par l'algorithme d'Euclide | Solution particulière de 26*k*-19*x* = 17 et solution générale par le théorème de Gauss |
|---|---|
| 26 = 19x1+7<br>19 = ..... | |

Sachant que *x* appartient à l'intervalle [ 0 ; 25 ], donnez la valeur de *x.*

Vérifiez avec votre crypteur.

## d/ Création d'un « décrypteur »

Supposons que a = 5 et b = 8. Pour coder, on utilise la fonction définie par $x \mapsto y \equiv 5x+8 \ [26]$

Pour décoder, nous allons résoudre la congruence d'inconnue *x* : $y \equiv 5x+8 \ [26]$

Pour cela, nous isolons *x* en multipliant les deux membres de la congruence par un entier relatif *u* qui doit vérifier $5u \equiv 1 \ [26]$. On obtient alors $uy \equiv 5ux+8u \ [26]$ mais $5u \equiv 1 \ [26]$ donc $x \equiv uy-8u \ [26]$.

On a montré ainsi que le décodage est effectué par une fonction affine définie par : $y \mapsto x \equiv uy+v \ [26]$

où *u* est le nombre qui vérifie $5u \equiv 1 \ [26]$ et où *v* est égal à - 8*u.*

Pour trouver l'entier relatif *u* on peut par exemple résoudre l'équation diophantienne 5 *u* – 26 *k* = 1
La résoudre et achever le décrypteur ! Test de notre décrypteur : entrez le message suivant, une lettre dans chaque cellule :

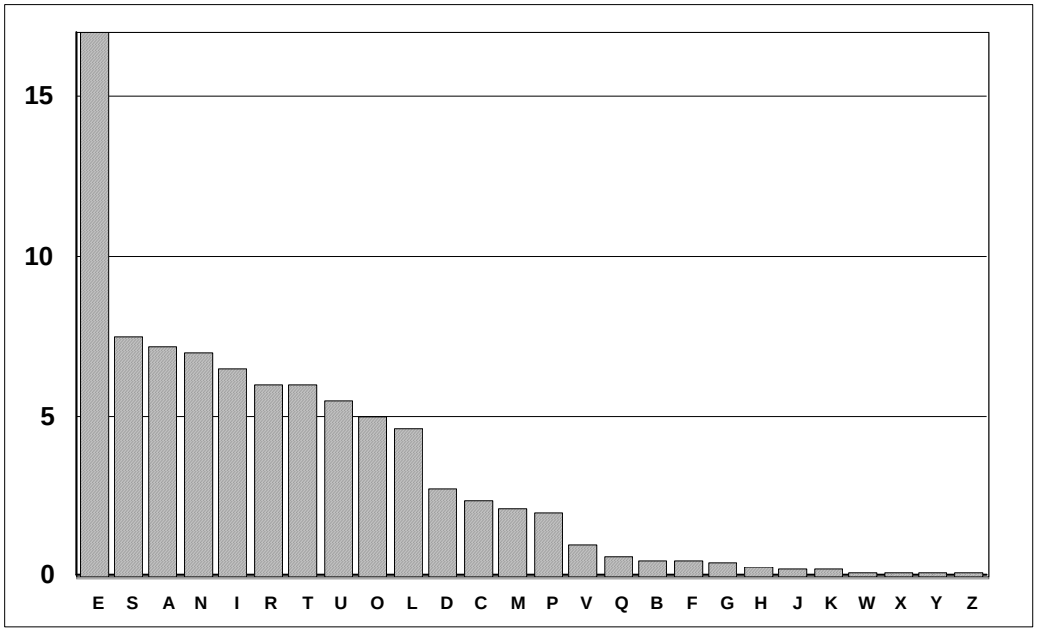
J W J C Z W L A E Q W V A E

# 2 Le système RSA

## a/ Présentation

Plusieurs inconvénients pour le système précédent (abandonné depuis longtemps bien entendu) :

- Peu de clés disponibles (intervalle [ 0 ; 25 ] ) donc possibilité de casser le code
- Répétition de lettres qui permet le repérage de certaines d'entre elles par calcul de fréquence d'apparition. Voir le diagramme des fréquences ci-dessous. A noter que même avec un alphabet désordonné et crypté avec un codage affine, il est possible de décrypter le message par l'analyse des répétitions de lettres.




- Obligation pour la personne qui veut envoyer un message codé de transmettre également au destinataire, la clé secrète.
- Obligations de changer de clé en fonction du récepteur.

Développé au MIT (Massachussets Institute of Technologie) en 1977 par Ronald Rivest, Adi Shamir et Leonard Adleman, le système RSA (initiales de ses inventeurs) est un système à clé publique très sûr.
Il est fondé sur le résultat suivant : il est très facile de savoir qu'un nombre même très grand est premier ; en revanche, on ne connaît pas d'algorithme exécutable en « temps raisonnable » capable de décomposer des grands nombres en produit de nombres premiers. Actuellement, les ordinateurs les plus puissants ne peuvent pas décomposer en nombres premiers un nombre qui comporte plus de 1 000 chiffres. Ainsi si on fait le produit de deux nombres premiers de 100 chiffres, on obtient un nombre que personne, sinon l'auteur, ne peut décomposer.
Schématiquement, le principe est très différent du codage précédent car la personne qui veut envoyer un message codé, va demander au destinataire une clé (que chacun peut connaître) et grâce à laquelle il va coder son message. Mais seul le destinataire pourra décoder le dit message car lui seul sait comment la clé a été fabriquée.

## b/ Principe

Tu veux me transmettre la valeur x
Je te transmets sans précaution la clé publique (e et n)
Tu t'en sers pour crypter le message
Tu me le transmets: y
Je suis le seul à pouvoir décrypter car j'ai la clé privée d

| | |
|---|---|
| TOI :<br>$\begin{cases} x \\ y \equiv x^e \text{ modulo } n \\ e \text{ et } n \text{ sont PUBLICS !} \end{cases}$ | ▪ **x** est la valeur que tu veux coder (tu découpes le message de façon que x < n)<br>▪ **e** et **n** sont les deux éléments de la clé publique<br>▪ Tu enfouis **x** dans un calcul qui donne **y**<br>▪ **y** est la valeur transmise |
| MOI :<br>$\begin{cases} z \equiv y^d \text{ modulo } n \\ z \equiv x^{ed} \text{ modulo } n \\ z \equiv x^1 \text{ modulo } n \end{cases}$ | ▪ Je calcule **z**<br>▪ en utilisant ma clé privée **d**<br>▪ Après ce tour de passe-passe mathématique<br>▪ Il se trouve que **z** est égal à **x** |
| LE TRUC : **$e \times d \equiv 1$ modulo f** | La justification complète est plus compliquée. Tout est dans le choix de e et f (voir plus bas). |

## c/ Exemple

| | | |
|---|---|---|
| ***Ma sauce à moi (préparation des clés)*** | | |
| ▪ Deux nombres premiers tenus secrets | **p et q** | 3 et 7 |
| ▪ Son produit **n** est diffusé | **n = p.q** | n = 21 |
| ▪ On calcule **f** | **f = (p-1) (q-1)** | f = 2x6 = 12 |
| ▪ Un nombre **e** est, lui aussi, diffusé | **e premier avec f** | e = 5 |
| ▪ Je calcule **d** , inverse de **e mod f** | **e.d = 1 mod f** | 5d = 1 mod 12<br>5x5 = 1 mod 12<br>d = 5 |
| ***Préparation du message*** | | |
| ▪ Le message est converti en chiffres<br>▪ On va coder chaque chiffre | **$x_1$ … $x_i$ …**<br>**avec $x_i$ < n** | x = 2, par exemple |
| ***Codage du message*** | | |
| ▪ Chaque **x** est codé par **y** | **$y = x^e \bmod n$** | $y = 2^5 \bmod 21$<br>y = 32 mod 21<br>y = 11 |
| ▪ Les valeurs de **y** sont transmises | **y est transmis** | 11 est transmis |
| ***Déchiffrage*** | | |
| ▪ Je calcule z | **$z = y^d \bmod n$** | $z = 11^5 \bmod 21$<br>z = 11 x121x121 mod 21<br>z = 11 x 16 x 16 mod 21<br>z = 11 x 256 mod 21<br>z = 11 x 4 mod 21<br>z = 44 mod 21 |
| ▪ Car en fait | **z = x mod n** | z = 2 mod 21 |

## d/ Théorie

| **Raisonnement** | **Quelques rappels théoriques** |
|---|---|
| ▪ Il s'agit de prouver le résultat suivant<br>**$z = y^d \bmod n$**<br>**$x^{ed} \bmod n$**<br>**x** | ▪ La fonction φ(n) est la quantité d'entiers strictement positifs inférieurs à n et premier avec n |

| *Quelques calculs préalables* | |
|---|---|
| ▪ On se souvient que | **f = (p - 1) (q - 1)** |
| ▪ C'est l'indicateur d'Euler ou fonction « phi de n » | **f = φ(n)** |
| ▪ On a également | **ed= 1 mod f** |
| | **= 1 mod φ(n)** |
| ▪ Ce qui veut dire que ed - 1 est un multiple de **φ(n)** | **ed - 1 = k . φ(n)** |
| ***Expression de z*** | |
| ▪ Si on revient à z | **$z = y^d \bmod n$** |
| | **$(x^e)^d \bmod n$** |
| | **$x^{e\,d} \bmod n$** |
| ▪ On ajoute un facteur en x que l'on retire dans l'exposant | **$z = x \cdot x^{e\,d-1} \bmod n$** |
| ▪ On connaît la valeur de cet exposant | **$z = x \cdot x^{k\,\varphi(n)} \bmod n$** |
| ▪ Il est temps d'appliquer le théorème de Fermat-Euler | **$z = x \cdot (1) \bmod n$** |
| | **$x \cdot \bmod n$** |
| ▪ Or on a choisi x < n | **z = x** |

**Théorème de Fermat**

**Soit *p* un nombre premier et a un entier naturel (autre que 0 et 1) non divisible par p, alors $a^{p-1}$ est congru à 1 modulo *p***

- La démonstration ci-contre est valable pour où x est premier avec n ce qui est très largement le cas général !!
- Cependant on peut montrer que le procédé reste valable même dans les autres cas (voir Activité 5 page 79 2.2)

## e/ Un crypteur RSA

On appelle *a* le message qu'Annabelle désire recevoir de Béatrice. Annabelle choisit deux nombres premiers 13 et 19 et calcule le produit 13 × 19 = 247 puis elle choisit un entier *e* premier avec ( *p* -1 )( *q* -1 ), par exemple 35 qui est premier avec 12 × 18. Elle dit à Béatrice qu'elle utilise la clé publique ( *n* , *e* ) soit ( 247, 35 ) et lui demande de coder son message avec cette clé. Le message codé et envoyé par Béatrice devient $b = a^e$.

Aidez Béatrice à coder DEMAIN MIDI avec la clé (247, 35) .
Ouvrez un tableur, créez une nouvelle feuille dans le fichier déjà créé.

- Dans les cellules A2 et B2, entrez les nombres 247 et 35 (clés publiques)

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **EXPEDITEUR** | | | MESSAGE INITIAL | D | E | M | A | I | N |
| 2 | **n** | **e** | | x | 3 | 4 | 12 | 0 | 8 | 13 |
| 3 | 247 | 35 | | y = x^e mod n | 165 | 62 | 103 | 0 | 31 | 117 |
| 4 | | | | | | | | | | |
| 5 | | | | CODAGE | 165 | 62 | 103 | 0 | 31 | 117 |
| 6 | | | | | | | | | | |

- Dans la cellule E2, entrez ==SI(E1="";"";CODE(E1)-64) et recopiez cette formule jusqu'en AG2. Expliquez la formule précédente sachant que CODE(lettre A) est égal à 65.
- Dans la cellule E3, entrez =SI(D2="";"";=Puismod(E2;$B$3;$A$3)) puis recopiez jusqu'en AG3. Expliquez la formule.
- Enfin, dans la cellule E5 entrez =E3.

C'est le message que reçoit Annabelle.

On montre, grâce au théorème de Fermat que $a^{k(p-1)(q-1)+1}$ est congru à *a* modulo *n* pour tout *a* et *k*.
Il reste à montrer que l'on peut écrire le nombre $k(p-1)(q-1)+1$ sous la forme *ed* d'où

$$b^d \equiv a^{de} \left[n\right]$$
$$b^d \equiv a \quad \left[n\right]$$

ce qui fait que en élevant à la puissance *d* le message *b* envoyé par Béatrice, Annabelle le décodera.

Pourquoi peut-on dire qu'il existe deux entiers *d* et *k* tels que $ed - k(p-1)(q-1) = 1$. Conclure.

## f/ Construction d'un décrypteur simplifié RSA

Pour décrypter un message, il faut détenir la clé privée *d* c'est-à-dire connaître le produit *(p-1)(q-1)*, ce qui demande de connaître *p* et *q* et donc de savoir factoriser l'entier *n*. Il n'était pas très compliqué de factoriser 247 mais que pouvez-vous proposer pour le nombre 1088311384786973019265306446798644412580 1 ?
Actuellement, les nombres utilisés sont de l'ordre de 768 bits ce qui signifie que ces nombres s'écrivent dans le système binaire à l'aide de 768 chiffres 0 ou 1.

Par exemple, le nombre suivant écrit sur 25 bits :

1 0 0 1 0 1 0 1 0 1 0 0 0 1 0 1 0 1 0 1 1 0 1 1 1

vaut dans le système décimal :

$$1\times 2^{24}+0\times 2^{23}+0\times 2^{22}+1\times 2^{21}+0\times 2^{20}+1\times 2^{19}+0\times 2^{18}+.....+1\times 2^{1}+1\times 2^{0}$$

soit le nombre 19 565 239 qui comporte déjà 8 chiffres.

Combien un nombre N écrit sur 768 bits dans le système binaire comporte-il de chiffres dans le système décimal ? On peut écrire l'inégalité suivante :$1x2^{768} \le N < 1x2^{768} + 1x2^{767} + ... + 1x2^{0}$
soit, en appliquant la formule connue des sommes de séries géométriques :

$$2^{768} \le N < 2^{769} - 1$$

En prenant le logarithme décimal de chacun des membres de la double inégalité, on obtient :

$$768\log(2) \le \log N < \log(2^{769} - 1)$$

et

$$231 \le \log N < 232$$

que l'on peut écrire sous la forme :

$$10^{231} \le N < 10^{232}$$

ce qui prouve que N possède 231 chiffres dans son écriture. Inutile donc de tenter de factoriser un tel nombre par de simples calculs.

Attention ! ici, la factorisation de *n* sera indiquée par vous-mêmes à l'ordinateur. Il n'est, bien entendu, pas dans notre propos de casser le code RSA, ce qui est illégal et, de toute façon, impossible à l'heure actuelle en raison de la taille des nombres utilisés.

A titre d'exemple, vous pouvez consulter, dans ce domaine des grands nombres premiers, l'adresse :
http://villemin.gerard.free.fr/Wwwgvmm/Premier/record.htm

Ouvrez un tableur et dans la feuille précédemment créée :

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **EXPEDITEUR** | | | MESSAGE INITIAL | D | E | M | A | I | N |
| 2 | **n** | **e** | | x | 3 | 4 | 12 | 0 | 8 | 13 |
| 3 | 247 | 35 | | y = x^e mod n | 165 | 62 | 103 | 0 | 31 | 117 |
| 4 | | | | | | | | | | |
| 5 | | | | CODAGE | 165 | 62 | 103 | 0 | 31 | 117 |
| 6 | | | | | | | | | | |
| 7 | | | | | | | | | | |
| 8 | **RECEPTEUR** | | | x = y ^d mod n | 3 | 4 | 12 | 0 | 8 | 13 |
| 9 | **p** | **q** | **d** | MESSAGE DECODE | D | E | M | A | I | N |
| 10 | 13 | 19 | 179 | | | | | | | |

- Dans la cellule E8, entrez =Puismod(E5;$C$10;$A$3) et recopiez cette formule jusqu'en AG8. Expliquez cette formule.
- Dans la cellule E9, entrez =SI(E8="";"";CAR(E8+64)) puis recopiez jusqu'en AG3. Expliquez la formule.

Exercice : Décoder :le message suivant écrit en RSA de clé publique (391 ; 5)

32 350 242 156 242 234 32 242 0 56

**COMPLEMENTS POUR LES PASSIONNES.**

## g/ Le découpage.

Dans cet exemple simplifié, les répétitions de lettres sont évidentes et nous retrouvons le problème évoqué dans le début du deuxième paragraphe et représenté par le diagramme des fréquences d'apparition des lettres de l'alphabet.
Dans la réalité, le système RSA procède par découpage du message en blocs. Voici une explication du système employé.
Nous allons utiliser la clé publique n = 5767 qui est le produit de 73 par 79 puis nous choisissons le nombre *e* égal à 9317. Ce nombre est premier avec $72 \times 78$.
Le message à transmettre est : "demain mardi 8 h 30"
Nous découpons le message en blocs de deux caractères pour obtenir :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ☚d | em | ai | n☚ | ma | rd | i☚ | 8☚ | h☚ | 30 |

Le symbole ☚ désigne l'espace.

Nous utilisons un alphabet qui comporte 37 signes à savoir les 26 lettres de l'alphabet habituel numérotées de 0 pour le "a" au 25 pour le "z", les dix chiffres numérotés de 26 pour le "0" à 35 pour le "9", 36 pour l'espace. Nous obtenons ainsi un système numérique comme le système décimal ou binaire qui comporte 37 symboles. C'est le système de base 37.
Ainsi, le bigramme "em" est chiffré "décimalement" par $4x37^1 + 12x37^0$ soit 160.
Le nombre 160 est codé par la clé publique sous la forme $160^{9317}$ modulo 5767 et on obtient alors 3718.
Revenons au système de base 37 en divisant successivement 3718 par 37 pour découvrir les puissance de 37 cachées dans ce nombre :
Nous obtenons $3718=37x100+18=37x(37x2+26)+18=2x37^2+26x37+18$ et donc le trigramme obtenu est c0s et, ceci, de manière unique ( ! ).
Pour le bigramme "ma", on a $12x37^1 + 0x37^0$ soit 444et $444^{9317}$ modulo 5767 donne 1337.
Or $1337=37 \times 36 + 5$ d'où le trigramme a☚f.
Remarquez alors qu'il n'est pas évident de repérer la répétition de la lettre "m" codée dans ces deux trigrammes. Ainsi "emma" est codé "c0sa☚f"
**Défi** : créez une feuille sur tableur qui permette de coder le message précédent.

## h/ La signature électronique

Un des intérêts du système RSA est qu'il permet de signer le message envoyé autrement dit lorsque vous recevrez le message de Béatrice, vous serez sûr que celui-ci provient bien de Béatrice. Voici la procédure mise en place :

Béatrice reçoit un message d'Annabelle. Est-ce bien un message d'Annabelle ?

Annabelle code son message à l'aide de sa clé privée *d* qu'elle seule connaît puis une seconde fois à l'aide de la clé publique *e'* que lui a fourni Béatrice. Béatrice reçoit le message d'Annabelle. Elle décode le message d'abord à l'aide de sa clé privée *d'* (ce qui lui assure la confidentialité du message) puis, ensuite, à l'aide de la clé publique *e* fournie par Annabelle (ce qui lui assure l'origine du message).

Exercice :
Annabelle propose la clé publique ( 5989 ; 625 )
Béatrice propose la clé publique ( 7031 ; 10625 )
Béatrice reçoit le message suivant d'Annabelle :

1 5975 3805 942 5975 893 3489

Ce message provient-il d'Annabelle ?

Et celui-ci ?

2118 4076 2118 6937 6937 0 5652 2118 3805 2118 5316 5054 2118
3805 4823 5544 0 6937 5448 0 3805 3805 0 1 2118 4365 4365 2118

## 3 Ajouts théoriques

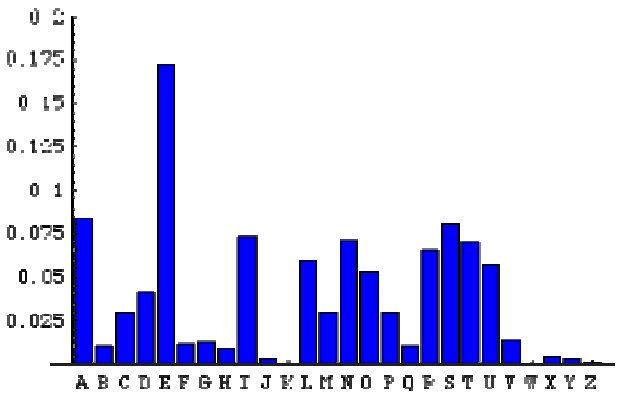
- Le produit (*p*-1)(*q*-1) est une valeur particulière de ce qu'on appelle la fonction indicatrice d'Euler. Cette fonction généralement désignée par φ est la fonction qui, à tout entier naturel *n* non nul, associe le nombre d'entiers inférieurs à *n* et premiers avec *n.*
  Ainsi φ( 20 ) = 8 car 1, 3, 7, 9, 11, 13, 17 et 19 sont premiers avec 20.
  Il est immédiat que φ (*p* ) = *p* – 1 si *p* est premier.

- Le théorème de Fermat-Euler (cas général du petit théorème de Fermat cité) précise que si *a* est un nombre premier avec *n,* alors

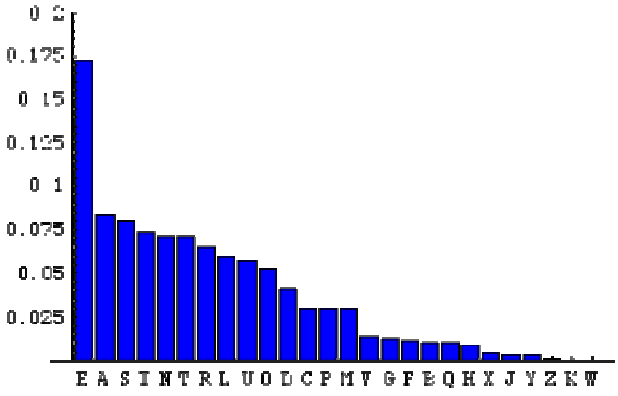
$$a^{\varphi(n)} \equiv 1 \quad [n]$$

| équences d'apparition des lettres | | | |
|---|---|---|---|
| Lettre | Fréquence | Lettre | Fréquence |
| A | 8.40 % | N | 7.13 % |
| B | 1.06 % | O | 5.26 % |
| C | 3.03 % | P | 3.01 % |
| D | 4.18 % | Q | 0.99 % |
| E | 17.26 % | R | 6.55 % |
| F | 1.12 % | S | 8.08 % |
| G | 1.27 % | T | 7.07 % |
| H | 0.92 % | U | 5.74 % |
| I | 7.34 % | V | 1.32 % |
| J | 0.31 % | W | 0.04 % |
| K | 0.05 % | X | 0.45 % |
| L | 6.01 % | Y | 0.30 % |
| M | 2.96 % | Z | 0.12 % |

**Histogramme par ordre alphabétique**




**Histogramme par ordre décroissant des fréquences**




Tous les tableaux de cette page ont été construits en comptant les fréquences dans un texte français de 100'000 lettres composé de textes de Gustave Flaubert (20'600 lettres), de Jules Verne (19'438) et de trois articles de l'Encyclopedia Universalis, le premier consacré à Bruges (8'182), le deuxième à l'artillerie (25'078) et le dernier à la population (26'702). On a utilisé le programme *Mathematica* ci-dessous:

stat.nb (12 Ko)

| Fréquences des bigrammes | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | _A | _B | _C | _D | _E | _F | _G | _H | _I | _J | _K | _L | _M | _N | _O | _P | _Q | _R | _S | _T | _U | _V | _W | _X | _Y | _Z |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A_ | 31 | 242 | 392 | 208 | 48 | 135 | 232 | 37 | 1255 | 32 | 7 | 663 | 350 | 1378 | 17 | 412 | 44 | 905 | 409 | 613 | 599 | 301 | 2 | 6 | 69 | 12 |
| B_ | 158 | 2 | 1 | 2 | 130 | 1 | 2 | 0 | 132 | 4 | 10 | 181 | 1 | 1 | 146 | 1 | 3 | 187 | 29 | 16 | 44 | 3 | 0 | 0 | 4 | 0 |
| C_ | 312 | 0 | 73 | 19 | 765 | 2 | 2 | 411 | 209 | 3 | 5 | 124 | 5 | 1 | 677 | 11 | 7 | 100 | 14 | 142 | 132 | 2 | 0 | 0 | 11 | 0 |
| D_ | 427 | 1 | 8 | 24 | 2409 | 2 | 5 | 25 | 378 | 3 | 0 | 14 | 21 | 5 | 231 | 4 | 6 | 134 | 64 | 3 | 406 | 4 | 1 | 0 | 5 | 0 |
| E_ | 616 | 176 | 917 | 998 | 782 | 258 | 209 | 67 | 179 | 96 | 8 | 1382 | 1056 | 2121 | 136 | 699 | 190 | 1514 | 3318 | 1307 | 761 | 258 | 11 | 125 | 15 | 60 |
| F_ | 181 | 1 | 1 | 8 | 180 | 118 | 1 | 1 | 190 | 0 | 0 | 43 | 1 | 1 | 213 | 1 | 2 | 106 | 12 | 1 | 61 | 0 | 0 | 0 | 1 | 0 |
| G_ | 135 | 1 | 10 | 9 | 408 | 4 | 63 | 3 | 69 | 6 | 4 | 74 | 10 | 103 | 47 | 5 | 1 | 197 | 12 | 23 | 81 | 1 | 0 | 0 | 2 | 0 |
| H_ | 267 | 5 | 4 | 1 | 285 | 0 | 0 | 0 | 149 | 3 | 0 | 3 | 4 | 17 | 107 | 0 | 3 | 18 | 5 | 0 | 42 | 0 | 1 | 0 | 7 | 0 |
| I_ | 176 | 85 | 203 | 172 | 1030 | 114 | 115 | 6 | 49 | 14 | 0 | 798 | 181 | 797 | 524 | 75 | 215 | 400 | 897 | 1243 | 11 | 190 | 1 | 40 | 0 | 4 |
| J_ | 76 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 91 | 0 | 0 | 0 | 0 | 0 | 42 | 0 | 0 | 0 | 2 | 0 |
| K_ | 8 | 0 | 0 | 0 | 6 | 0 | 3 | 0 | 6 | 0 | 0 | 0 | 10 | 3 | 9 | 0 | 0 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |
| L_ | 1270 | 14 | 22 | 58 | 2366 | 25 | 14 | 39 | 512 | 4 | 1 | 647 | 18 | 41 | 281 | 69 | 47 | 16 | 126 | 42 | 369 | 14 | 0 | 0 | 15 | 1 |
| M_ | 510 | 152 | 11 | 11 | 1099 | 0 | 1 | 1 | 302 | 0 | 0 | 7 | 243 | 4 | 334 | 201 | 2 | 10 | 10 | 8 | 52 | 1 | 0 | 0 | 3 | 0 |
| N_ | 405 | 30 | 438 | 785 | 985 | 124 | 222 | 24 | 316 | 17 | 7 | 89 | 68 | 249 | 303 | 130 | 82 | 55 | 846 | 1694 | 114 | 109 | 0 | 1 | 19 | 20 |
| O_ | 6 | 83 | 88 | 101 | 46 | 32 | 115 | 7 | 452 | 14 | 3 | 184 | 391 | 1646 | 8 | 175 | 19 | 491 | 126 | 109 | 1086 | 28 | 9 | 4 | 62 | 4 |
| P_ | 671 | 1 | 3 | 21 | 441 | 5 | 1 | 136 | 119 | 0 | 0 | 377 | 2 | 4 | 505 | 125 | 1 | 363 | 31 | 65 | 140 | 1 | 0 | 0 | 1 | 0 |
| Q_ | 2 | 0 | 3 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 3 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 975 | 0 | 0 | 0 | 0 | 0 |
| R_ | 896 | 53 | 168 | 302 | 1885 | 46 | 96 | 5 | 583 | 11 | 3 | 292 | 181 | 88 | 520 | 82 | 51 | 176 | 386 | 445 | 183 | 77 | 1 | 1 | 21 | 5 |
| S_ | 809 | 85 | 306 | 735 | 1377 | 151 | 73 | 83 | 565 | 36 | 0 | 453 | 192 | 107 | 521 | 496 | 191 | 137 | 702 | 578 | 343 | 92 | 1 | 6 | 30 | 10 |
| T_ | 881 | 25 | 166 | 515 | 1484 | 52 | 19 | 64 | 984 | 28 | 3 | 331 | 70 | 40 | 363 | 268 | 96 | 668 | 404 | 269 | 270 | 41 | 4 | 6 | 18 | 3 |
| U_ | 168 | 87 | 165 | 162 | 781 | 40 | 83 | 4 | 534 | 41 | 3 | 302 | 128 | 516 | 19 | 184 | 15 | 980 | 591 | 469 | 14 | 177 | 1 | 264 | 8 | 4 |
| V_ | 277 | 0 | 1 | 0 | 502 | 0 | 0 | 0 | 288 | 0 | 0 | 1 | 0 | 0 | 167 | 0 | 0 | 81 | 0 | 0 | 11 | 0 | 0 | 0 | 0 | 0 |
| W_ | 11 | 1 | 1 | 0 | 3 | 0 | 0 | 2 | 8 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| X_ | 35 | 14 | 37 | 36 | 68 | 8 | 7 | 5 | 57 | 0 | 0 | 21 | 15 | 3 | 7 | 56 | 11 | 3 | 15 | 35 | 2 | 18 | 0 | 4 | 0 | 0 |
| Y_ | 63 | 0 | 7 | 7 | 59 | 3 | 4 | 0 | 0 | 0 | 0 | 13 | 8 | 5 | 15 | 14 | 0 | 10 | 75 | 9 | 2 | 4 | 0 | 0 | 0 | 0 |
| Z_ | 8 | 0 | 2 | 6 | 49 | 3 | 1 | 0 | 1 | 1 | 0 | 11 | 4 | 2 | 15 | 4 | 1 | 0 | 3 | 1 | 0 | 7 | 4 | 0 | 0 | 2 |

| Les 20 bigrammes les plus fréquents | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bigrammes | ES | DE | LE | EN | RE | NT | ON | ER | TE | EL | AN | SE | ET | LA | AI | IT | ME | OU | EM | IE |
| Nombres | 3318 | 2409 | 2366 | 2121 | 1885 | 1694 | 1646 | 1514 | 1484 | 1382 | 1378 | 1377 | 1307 | 1270 | 1255 | 1243 | 1099 | 1086 | 1056 | 1030 |

| Les 20 trigrammes les plus fréquents | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Trigrammes | ENT | LES | EDE | DES | QUE | AIT | LLE | SDE | ION | EME | ELA | RES | MEN | ESE | DEL | ANT | TIO | PAR | ESD | TDE |
| Nombres | 900 | 801 | 630 | 609 | 607 | 542 | 509 | 508 | 477 | 472 | 437 | 432 | 425 | 416 | 404 | 397 | 383 | 360 | 351 | 350 |