La cryptographie est une discipline qui consiste à encoder et décoder des messages confidentiels, afin que seuls l'expéditeur et le destinataire du message puissent le lire.

Cryptographie par substitution et analyse fréquentielle.

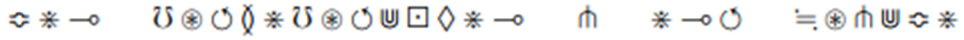
La méthode la plus simple pour crypter un message consiste à remplacer une lettre par un autre symbole : par exemple, la lettre « A » est remplacée par le symbole « ~ » , la lettre « B » est remplacée par le symbole « ./ » , etc. Cette méthode de cryptographie est peu sûre. En effet, il est relativement facile de décoder un message, même sans connaître la correspondance entre lettres et symboles, par exemple avec la méthode appelée analyse fréquentielle : le tableau suivant donne la fréquence *d* » utilisation (à 0,001 près) de chaque lettre dans la langue Française.

Lettres	A	B	C	D	E	F	G	H	I	J	K	L	M
Fréquences	0,094	0,01	0,026	0,034	0,159	0,01	0,01	0,008	0,084	0,009	0	0,053	0,032

Lettres	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Fréquences	0,072	0,051	0,028	0,011	0,065	0,079	0,073	0,062	0,022	0	0,003	0,002	0,003

1. A l'aide de ce tableau, identifier les 5 lettres les plus utilisées dans la langue Française.

2. A l'aide de ce résultat, déchiffrer le message suivant :



Une autre méthode de cryptographie :

Le masque jetable (aussi appelé Chiffre de Vernam). Cette méthode de cryptographie *a* été utilisée par Che Guevara et Fidel Castro pour communiquer, ainsi que par le KGB, puis pour le téléphone rouge entre Moscou et Washington, et est encore utilisée par des services secrets. Cette méthode repose sur trois critères :

• la clé ou masque (c'est à dire le code secret servant à encoder et décoder un message, cette clé est connue seulement de l'expéditeur et du destinataire du message) doit être aussi longue que le message à encoder.

• les caractères composant la clé doivent être choisis de façon totalement aléatoire.

• chaque clé (ou masque) ne doit être utilisé qu'une seule fois (d'où le nom de masque jetable).

Nous allons choisir le mot « MATHEMATIQUES » comme clé dans la suite de l'exercice.

Fonctionnement :

Chaque lettre est tout d'abord remplacée par un chiffre, c'est-à-dire que la lettre « A » est remplacée par « 0 », « B » est remplacé par « 1 », « C » est remplacé par « 2 » et ainsi de suite.

Méthode pour encoder :

On veut encoder le message « HELLO ».

Pour cela, on additionne le nombre associé à la première lettre du message au nombre associé à la première lettre de la clé, et on soustrait 26 si le résultat est strictement supérieur à 25. On fait de même avec la seconde lettre du message et la seconde lettre de la clé, puis avec les troisièmes lettres, les quatrièmes, etc. Enfin, on la liste des 5 lettres associées aux 5 nombres obtenus représente le message encodé.

3. Donner la liste des 5 lettres représentant le codage du mot « HELLO ».

4. Expliquer pourquoi la méthode d'analyse fréquentielle ne permet pas de décoder ce message.

Méthode pour décoder :

On associe à la première lettre du message encodé le nombre correspondant, et soustrait le nombre associé à la première lettre de la clé, en ajoutant 26 si le résultat est strictement négatif. Enfin, on retrouve la lettre correspondante. Puis on fait de même avec la seconde lettre du message encodé et la deuxième lettre de la clé, puis avec les troisièmes, les quatrièmes, etc.

5. Décoder le résultat obtenu en 3 pour vérifier que l'on obtient bien le message « HELLO ».

6. Décoder le message suivant : D E O V P G M Q E H

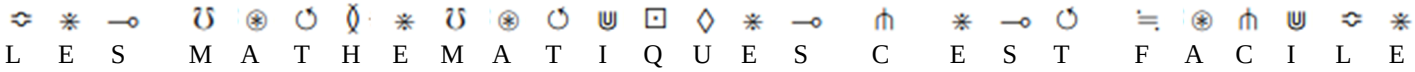
CORRECTION

1. A l'aide de ce tableau, les 5 lettres les plus utilisées dans la langue Française sont : E, A, I, S, T.

Lettres	E	A	I	S	T
Fréquences	0,159	0,094	0,084	0,079	0,073

2. Déterminons le nombre d'apparition des différents symboles

2	5	2		2	2	3	1	2	1	1	2	1
L	E	S		M	A	T	H	I	Q	U	C	F



L E S M A T H E M A T I Q U E S C E S T F A C I L E

3. La liste des 5 lettres représentant le codage du mot « HELLO » est TEESS

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Lettre	M	A	T	H	E
Nombre *n*	12	0	19	7	4

Lettre	H	E	L	L	O
Nombre *m*	7	4	11	11	14
m + *n* (à 26 près)	19	4	4	18	18
code	T	E	E	S	S

4. Les lettre A et T ont été codées par E de même pour les lettres L et O codées par S donc la méthode d'analyse fréquentielle ne permet pas de décoder ce message.

Méthode pour décoder :

5.

Lettre	M	A	T	H	E
Nombre *n*	12	0	19	7	4

Lettre	T	E	E	S	S
Nombre *m*	19	4	30	18	18
m – *n* (à 26 près)	7	4	11	11	14
Message décodé	H	E	L	L	O

6.

Lettre	M	A	T	H	E	M	T	I	Q	U
Nombre *n*	12	0	19	7	4	12	19	8	16	20

Lettre	D	E	O	V	P	G	M	Q	E	H
Nombre *m*	3	4	14	21	15	6	2	1	22	3
m – *n* (à 26 près)	17	4	21	14	11	20	9	19	6	9
Message décodé	R	E	V	O	L	U	T	I	O	N