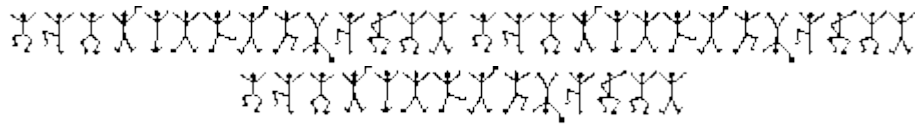


# Aspects de la cryptologie au cours de l'Histoire

*Une réflexion très actuelle sur l'art de  
briser la sécurité des systèmes*

---



**Patrick Legand**

Consultant SSI

[contactPL@patrick-legand.com](mailto:contactPL@patrick-legand.com)

<http://blog.patrick-legand.com>

- V1.0 -

## Table des matières

<b>INTRODUCTION.....</b>	<b>3</b>
<b>LES PREMIERS PAS DE LA CRYPTOLOGIE.....</b>	<b>4</b>
Les premières substitutions alphabétiques.....	5
La linguistique et les débuts de cryptanalyse.....	6
Substitutions polyalphabétiques.....	8
Les codes homophones.....	10
Contre le chiffre de Vigenère. Méthode de Babbage et Kasiski.....	12
Les avancées de la cryptologie après la Première Guerre mondiale.....	14
Techniques de chiffrement utilisées lors de la Première Guerre mondiale.....	17
<b>ENIGMA.....</b>	<b>19</b>
La cryptanalyse d'Enigma.....	21
Une bataille du chiffre pendant la Seconde Guerre mondiale.....	25
<b>CONCLUSIONS - LES PETITES LEÇONS DE L'HISTOIRE.....</b>	<b>28</b>

## Index des illustrations

Figure 2-1: Une Scytale	4
Figure 2-2: Exemples de substitutions monoalphabétiques	5
Figure 2-3: Distribution de fréquence d'apparition des lettres de l'alphabet	7
Figure 2-4 : Fréquences d'apparition de digrammes et trigrammes (langue anglaise)	7
Figure 2-5: Exemple de chiffrement par substitution polyalphabétique	9
Figure 2-6 : Codes homophones. Matrice de substitution	10
Figure 2-7 : Codes homophones bidimensionnels	11
Figure 2-8 : Méthode de décryptement du chiffre de Vigenère	13
Figure 2-9 : Codage par transposition	15
Figure 2-10: Transposition par mots-clef (1)	16
Figure 2-11: Transposition par mots-clef (2)	16
Figure 2-12: Le chiffre ADFGVX	18
Figure 2-13: Fin du télégramme de Zimmermann	18
Figure 2-14: Principes de fonctionnement d'Enigma	19
Figure 2-15: Schéma de principe de la machine complète	20
Figure 2-16: Protocole d'échanges sécurisé mis en œuvre pour Enigma	22
Figure 2-17: Cryptanalyse d'Enigma	23
Figure 2-18: Enigma. Tableau de connexions	24
Figure 2-19: Cryptanalyse d'Enigma. Méthode de Turing	26

# Introduction

Un objectif de ce document est d'offrir une petite visite guidée de la cryptologie au cours des âges, et de découvrir comment les hommes s'y prenaient pour protéger le secret de leurs messages.

Satisfaire la curiosité est un plaisir, mais il serait vain de limiter cette incursion à ce seul sujet. Le deuxième objectif, peut-être le plus important, est d'explorer quelques méthodes mises au point par d'autres hommes pour briser les codes et accéder ainsi à la correspondance de leurs adversaires.

À travers les erreurs commises par les utilisateurs et les cryptologues, mais aussi en mettant en évidence l'acharnement impitoyable des gouvernements et des cryptanalystes lorsqu'il s'agit de casser un chiffre, nous espérons donner un aperçu des enjeux qui se cachent derrière la sécurité d'un système informatique et montrer que la conception d'un volet sécurité ne doit pas être prise à la légère.

## Terminologie

<b>Cryptologie</b>	La science des documents secrets. Elle recouvre tous les aspects scientifiques, et plus particulièrement mathématiques, relatifs à la cryptologie et à la cryptanalyse.
<b>Cryptanalyse</b>	L'art et la science de décrypter les messages secrets.
<b>Chiffrement</b>	Le processus de transformation d'un message de telle manière à le rendre incompréhensible pour toute personne non autorisée
<b>Cryptogramme</b>	Le résultat du processus de chiffrement.
<b>Cryptologue</b>	Celui qui écrit un cryptogramme.
<b>Déchiffrement</b>	Le processus de reconstruction du texte clair à partir du texte chiffré, par des personnes autorisées.
<b>Décryptement</b>	Le processus de reconstruction du texte clair à partir du texte chiffré, par des personnes non autorisées (sans connaître la clef).
<b>Clef de chiffrement</b>	Un mot, un nombre ou une phrase qui est utilisé par un algorithme de cryptologie pour chiffrer ou déchiffrer un message.
<b>Stéganographie</b>	La dissimulation du message dans un ensemble de données d'apparence anodine.

# Les premiers pas de la cryptologie

Historiquement, la cryptologie fut employée par quatre catégories de personnes : les militaires, les corps diplomatiques, les journalistes et les amoureux. Si les amoureux eurent le plus souvent recours à des chiffres qui les protégeaient « *de leur petite sœur* », les militaires jouèrent un rôle prépondérant dans l'évolution et le perfectionnement de la cryptologie.

A l'origine, le secret des communications reposait sur l'utilisation de diverses formes de **stéganographie**, un procédé visant à dissimuler l'existence du message, plutôt qu'à masquer le sens du message lui-même. Bien que la stéganographie fût l'inspiratrice de ruses incroyablement ingénieuses (en Chine on écrivait les messages sur une soie très fine glissée dans une petite boule recouverte de cire, avalée ensuite par le messager ; en Italie on utilisait une encre absorbée par la coquille d'un œuf dur, ...), ce procédé se révélait d'une grande faiblesse lorsque le message était découvert.

C'est pourquoi il fallut trouver une technique destinée à préserver le secret du message, notamment dans le cas où celui-ci était intercepté : la **cryptologie** vit ainsi le jour.

Les plus anciens exemples d'écriture codée remontent au temps des conflits entre la Grèce et la Perse au V<sup>ème</sup> siècle avant J.-C. L'emploi de la « **Scytale** » des Lacédémoniens, premier procédé de chiffrement militaire connu, nous est conté par Plutarque dans un extrait de la *Vie de Lysandre* :

*« Cette scytale est une telle chose : quand les éphores envoient à la guerre un général, ou un amiral, ils font accourir deux petits bâtons ronds, et les font entièrement égaux en grandeur et en grosseur, desquels deux bâtons ils retiennent l'un par devers eux, et donnent l'autre à celui qu'ils envoient. Ils appellent ces deux petits bâtons scytales, et quand ils veulent faire secrètement entendre quelque chose de conséquence à leurs capitaines, ils prennent un bandeau de parchemin long et étroit comme une courroie qu'ils entortillent à l'entour de leur bâton rond, sans laisser rien d'espace vide entre les bords du bandeau ; puis quand ils sont ainsi bien joints, alors ils écrivent sur le parchemin ainsi roulé ce qu'ils veulent ; et quand ils ont achevé d'écrire, ils développent le parchemin et l'envoient à leur capitaine, lequel n'y saurait autrement rien lire ni connaître, parce que les lettres n'ont point de suite ni de liaison continuée, mais sont écartées, l'une çà, l'autre là, jusques à ce que prenant le petit rouleau de bois qu'on lui a baillé à son partement, il étend la courroie de parchemin qu'il a reçue tout à l'entour, tellement que, le tour et le pli du parchemin venant à se retrouver en la même couche qu'il avait été plié premièrement, les lettres aussi viennent à se rencontrer en la suite continuée qu'elles doivent être. Ce petit rouleau de parchemin s'appelle aussi bien scytale comme le rouleau de bois, ni plus ni moins que nous voyons ailleurs ordinairement, que la chose mesurée s'appelle du même nom que fait celle qui mesure. »*

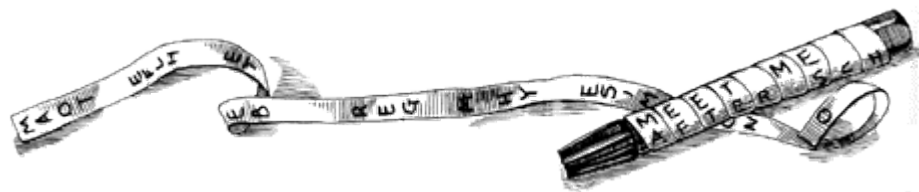


Figure 2-1: Une Scytale

# Les premières substitutions alphabétiques

Au cours de l'histoire, c'est à dire *grosso modo* des premiers temps jusqu'à l'avènement de l'informatique, la cryptologie consistait essentiellement en de savantes méthodes de manipulation de caractères.

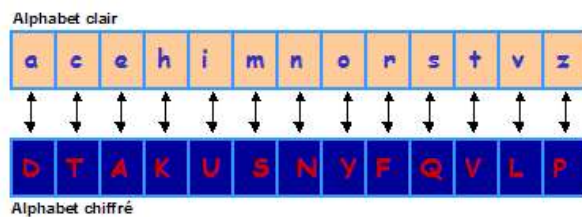
Deux méthodes étaient couramment utilisées pour chiffrer les textes : la **Transposition** et la **Substitution**. La substitution consistait à remplacer chaque caractère du texte en clair par d'autres caractères (lettres, chiffres ou autres symboles) dans le texte chiffré. L'une des premières descriptions du chiffrement par substitution apparaît d'ailleurs dans le *Kâma-sûtra*, texte fondé sur des manuscrits du IV<sup>ème</sup> siècle avant J.-C. Le Kâma-sûtra recommande aux femmes l'apprentissage du « *mlecchita-vikalpà* », l'art de l'écriture secrète, qui se révèle parfois d'un grand secours pour dissimuler une liaison !

Une substitution monoalphabétique consiste à appairer au hasard les lettres de l'alphabet et à substituer dans le message original la lettre correspondante dans l'alphabet chiffré (voir Figure 2-2).

## Appariement au hasard des lettres de l'alphabet

Message à chiffrer :  
venez chez moi ce soir

Message obtenu :  
**LANAP TKAP SYU TA QYUF**



## Alphabet de CESAR

Message à chiffrer :  
attaquons demain dix heures

Message obtenu :  
**DWWDTXRQV GHPDLQ  
GLA KH XUHV**

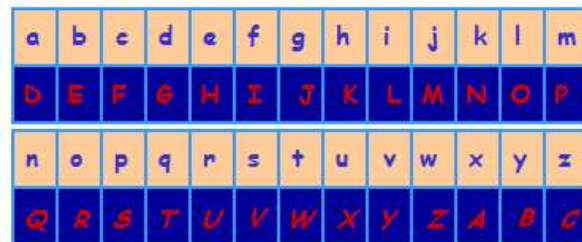


Figure 2-2: Exemples de substitutions monoalphabétiques

Selon Suétone, cette technique aurait été utilisée par **Jules César** dans la Rome Antique (« *Vies des douze Césars* ») :

*« On possède [...] de César des lettres à Cicéron, et sa correspondance avec ses amis sur ses affaires domestiques. Il y employait, pour les choses tout à fait secrètes, une espèce de chiffre (les lettres étant disposées de manière à ne pouvoir jamais former un mot), et qui consistait, je le dis pour ceux qui voudront les déchiffrer, à changer le rang des lettres, à écrire la quatrième pour la première, comme le d pour l'a, et ainsi des autres. »*

Le programme de chiffrement **ROT 13** utilisé aujourd'hui par les systèmes UNIX est un chiffre de César : chaque caractère du texte clair est remplacé par celui qui se trouve 13 places plus loin modulo 26. Le texte original peut être obtenu en chiffrant une deuxième fois le message.

La substitution monoalphabétique du premier exemple de la Figure 2-2 (appariement au hasard) offre 26! (Factorielle 26) permutations possibles, c'est à dire plus de  $4.10^{26}$  (10 avec 26 zéros derrière) façons différentes de chiffrer le même message ! Technique très prometteuse sur le plan cryptologique, mis à part une contrainte pour les utilisateurs du chiffre, qui doivent se souvenir de l'alphabet chiffré (ou l'écrire, ce qui pose évidemment le problème de sa possible découverte par l'adversaire !).

Pour contourner cette difficulté, un petit moyen mnémotechnique basé sur l'utilisation d'une **phrase clef** fut inventé : en omettant les lettres répétées et les espaces, le nom « Julius Caesar » pouvait donner par exemple la clef suivante : **JULISCAER**. Il suffisait de compléter cette clef avec les autres caractères de l'alphabet, dans l'ordre de l'alphabet, pour engendrer un alphabet chiffré :

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  
J U L I S C A E R T V W X Y Z B D F G H K M N O P Q

Pour reconstituer l'alphabet chiffré, les utilisateurs n'avaient qu'à mémoriser la clef.

A une époque sans informatique, attaquer de tels codes par une méthode exhaustive (dite aussi de force brute – c'est le cas de le dire !) aurait pris au bas mot  $10^{19}$  années (soit un milliard de fois la durée de vie de l'univers !), en essayant une combinaison par seconde. C'est probablement la raison pour laquelle ce principe élémentaire de substitution a dominé les techniques de chiffrement pendant le premier millénaire après J.-C.

Un tel procédé aurait pu exister longtemps encore, si l'homme ne n'avait pas cette désagréable manie d'échafauder des méthodes bien plus subtiles que la recherche exhaustive. En réalité, la rencontre fortuite de la linguistique, des statistiques et de la religion offrit aux érudits du moyen âge une méthode imparable pour trouver des raccourcis et briser un chiffre par substitution monoalphabétique en quelques minutes. La **cryptanalyse** était née.

## La linguistique et les débuts de cryptanalyse

La **cryptanalyse** fut inventée par les Arabes. En cherchant à reconstituer la chronologie des révélations de Mahomet, les théologiens notèrent une caractéristique importante de la linguistique : un texte n'a rien d'une chaîne aléatoire de caractères. Les langues sont redondantes dans la mesure où elles emploient plus de caractères, de mots et de phrases que nécessaire, et chaque lettre apparaît selon une fréquence qui lui est propre. En particulier, le *a* et le *l* sont les lettres les plus courantes en arabe.

Le grand savant al-Kindi rédigea au IX<sup>ème</sup> siècle le « *Manuscrit sur le déchiffrement des messages cryptographiques* », le plus ancien document connu sur la cryptanalyse par l'analyse des fréquences. Le principe de cryptanalyse d'al-Kindi tient en deux paragraphes :

*Une façon d'élucider un message chiffré, si nous savons dans quelle langue il est écrit, est de nous procurer un autre texte clair dans la même langue, de la longueur d'un feuillet environ, et de compter alors les apparitions de chaque lettre. Nous appellerons la lettre apparaissant le plus souvent la « première », la suivante la « deuxième », et ainsi de suite pour chaque lettre figurant dans le texte.*

*Ensuite, nous nous reportons au texte chiffré que nous voulons éclaircir et nous relevons de même les symboles. Nous remplaçons le symbole le plus fréquent par la lettre « première » du texte clair, le suivant par la lettre « deuxième », le suivant par la « troisième », et ainsi de suite jusqu'à ce que nous soyons venus à bout de tous les symboles du cryptogramme à résoudre.*

A titre d'exemple, le tableau ci-après représente les distributions de fréquence d'apparition des 26 lettres de l'alphabet dans des textes ordinaires en français et en anglais.

FRANÇAIS	E	A	I	S	T	N	R	U	L	O	D	M	P
	15.87	9.42	8.41	7.90	7.26	7.15	6.46	6.24	5.34	5.14	3.39	3.24	2.86
	C	V	Q	G	B	F	J	H	Z	X	Y	W	K
	2.64	2.15	1.06	1.04	1.02	0.95	0.89	0.77	0.32	0.30	0.24	0.00	0.00
ANGLAIS	E	T	O	A	N	R	I	S	H	D	L	C	F
	13.05	9.02	8.21	7.81	7.28	6.77	6.64	6.46	5.85	4.11	3.60	2.93	2.88
	U	M	P	Y	W	G	B	V	K	X	J	Q	Z
	2.77	2.62	2.15	1.51	1.49	1.39	1.28	1.00	0.42	0.30	0.23	0.14	0.09

Pour un échantillonnage d'environ 10 000 caractères (joumaux ou romans)

Figure 2-3: Distribution de fréquence d'apparition des lettres de l'alphabet

L'application de la recette d'al-Kindi nécessite parfois finesse et astuce. Un ouvrage spécialisé peut présenter une disparité significative de répartition : il peut y avoir de grandes différences entre le code source d'un programme informatique, la musicologie sur l'œuvre d'un grand compositeur ou le business plan d'une entreprise sur les trois ans à venir. Par exemple, le roman de Georges Pérec *La Disparition* n'emploie pas un seul mot contenant la lettre *e*. Il fait pourtant 200 pages!

Si les lettres isolées apparaissent selon des fréquences prévisibles, il en est de même pour les paires de lettres (digrammes) et les groupes de trois lettres (trigrammes). Les organismes de cryptanalyse spécialisés ont établi de vastes banques de données sur les variations de la répartition des caractères selon les cibles potentielles, de façon à faciliter le forçage des codes employés. Ces tables comprennent des renvois et des analyses plus fines pour divers types de documents, spécialisés ou non, dont la répartition des caractères peut présenter des différences significatives, donc exploitables. Avec l'expérience acquise, **25 caractères d'un texte chiffré suffisent aujourd'hui pour qu'un bon cryptanalyste décrypte le message.**

TH	IN	ER	RE	AN	HE	AR	EN	TI	TE	AT	ON	HA
3.16	1.54	1.33	1.30	1.08	1.08	1.02	1.02	1.02	0.98	0.88	0.84	0.84
OU	IT	ES	ST	OR	NT	HI	EA	VE	CO	DE	RA	RO
0.72	0.71	0.69	0.68	0.68	0.67	0.66	0.64	0.64	0.59	0.55	0.55	0.55
THE	ING	AND	ION	ENT	FOR	TIO	ERE	HER	ATE	VER	TER	THA
4.72	1.42	1.13	1.00	0.98	0.76	0.75	0.69	0.68	0.66	0.63	0.62	0.62
ATI	HAT	ERS	HIS	RES	ILL	ARE	CON	NCE	ALL	EVE	ITH	TED
0.59	0.55	0.54	0.52	0.50	0.47	0.46	0.45	0.45	0.44	0.44	0.44	0.44

Figure 2-4 : Fréquences d'apparition de digrammes et trigrammes (langue anglaise)

L'attaque de base s'appuie donc sur les propriétés statistiques des langages naturels. Le cryptanalyste désirant casser un code monoalphabétique commence par calculer les fréquences relatives de toutes les lettres du texte chiffré. En anglais, il essaie alors de remplacer la lettre la plus fréquente par la lettre e, la seconde la plus utilisée par le T, etc. Il cherche ensuite un trigramme fréquent de la forme tXe. Il y a de fortes chances que la lettre codée par X soit h. Puis si le modèle suivant : thYt apparaît, le Y est probablement un a. Ainsi de suite, le cryptanalyste construit en tâtonnant un texte clair, lettre par lettre. A titre d'exemple, le *Scarabée d'Or* d'Edgar Poe (voir Annexe B) donne un exemple élégant de la manière dont on met en œuvre une telle technique.

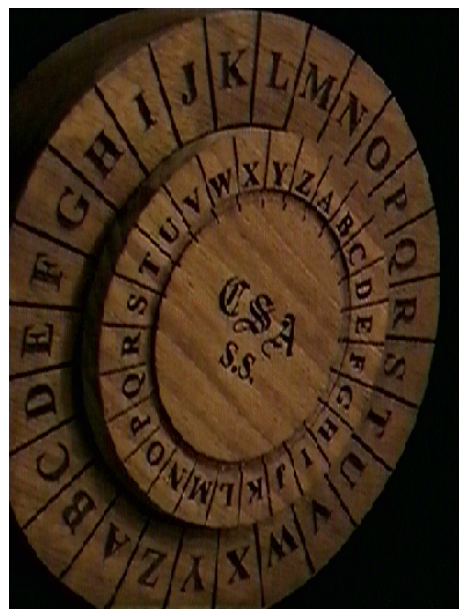
L'effervescence intellectuelle de la Renaissance encouragea l'usage massif de la cryptologie : les alchimistes souhaitaient protéger leurs découvertes, les diplomates le secret des complots politiques. Tel un contre-pouvoir, l'influence des cryptanalystes se répandit progressivement en Europe ; ils s'employèrent à casser les chiffres et réussirent à rendre caduque la technique de substitution monoalphabétique.

À la fin du XVI<sup>ème</sup> siècle, un épisode tragique montra l'importance stratégique de la cryptanalyse : le démantèlement du complot Babington, qui précipita la fin de Marie Stuart en 1587.

Le **chiffre de Marie Stuart** reposait sur une nomenclature, une sorte de substitution monoalphabétique basée sur une palette de symboles plus large que l'alphabet. Malheureusement, cette nomenclature ne résista pas longtemps à l'analyse de fréquences. Ayant percé le code, le ministre de la Reine Elisabeth, Walsingham, fit établir un faux message chiffré, prouvant ainsi le rôle actif de Marie Stuart dans un complot visant à renverser celle-ci. Babington et ses acolytes furent écartelés, et Marie Stuart exécutée sur la place publique.

## Substitutions polyalphabétiques

C'est alors que le savant florentin **Léon Battista Alberti** (XV<sup>ème</sup> siècle) eut l'idée de définir une nouvelle méthode basée sur l'utilisation de plusieurs alphabets chiffrés.



Cadran d'Alberti

Une partie du message était codée à l'aide d'un premier alphabet chiffré, une autre partie avec un second, et ainsi de suite. Il réalisait ainsi **la première avancée significative de la cryptologie depuis mille ans**.

Afin de faciliter la tâche des opérateurs, il inventa une sorte de disque composé de deux anneaux concentriques sur lesquels étaient inscrites les lettres de l'alphabet : le **cadran d'Alberti**. Le choix d'un alphabet chiffré était déterminé par la position relative des deux anneaux. Une forme élaborée du cadran d'Alberti allait être utilisée durant la Guerre de Sécession ; les Confédérés employèrent des phrases code qui leur permettaient d'utiliser jusqu'à 26 alphabets chiffrés différents par message !

Un diplomate français né en 1523, **Blaise de Vigenère**, reprit les travaux d'Alberti et proposa une nouvelle forme de chiffre, basée cette fois sur 26 alphabets.

Le principe du **chiffre à substitution polyalphabétique** était né. Il reposait sur l'utilisation d'une clef et d'une matrice d'alphabets chiffrés : le « carré de Vigenère ». La clef était inscrite plusieurs fois au-dessus du texte à chiffrer, ainsi chaque lettre du texte clair était encodée selon l'alphabet indiqué par la lettre de la clef qui la surplombe. Exemple :

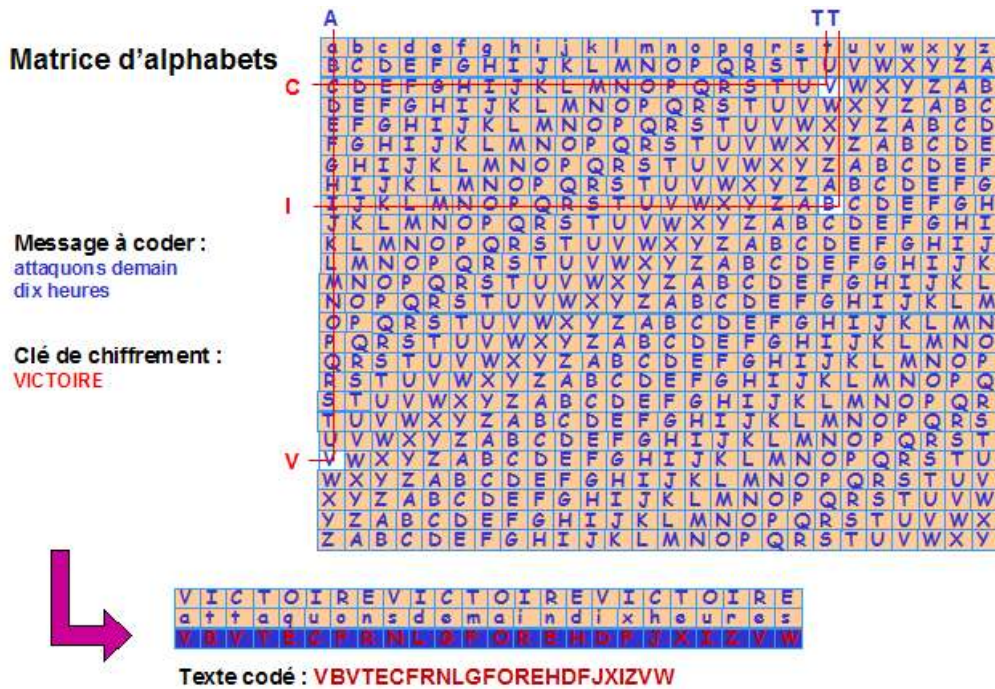


Figure 2-5: Exemple de chiffrement par substitution polyalphabétique (carré de Vigenère)

Cette technique est nettement meilleure que la simple substitution monoalphabétique car une lettre du texte clair n'est plus forcément représentée par la même lettre dans le texte chiffré. **Le grand avantage du chiffre de Vigenère est qu'il est inattaquable par l'analyse des fréquences.** Pour information, le logiciel WordPerfect utilise un chiffre à substitution polyalphabétique similaire.

Cependant, la complexité de mise en oeuvre d'un tel chiffre était incompatible avec les besoins en communication des chancelleries, qui était, lui, en forte croissance. Voilà pourquoi l'utilisation du chiffre de Vigenère demeura marginale, et que le chiffre monoalphabétique allait rester prépondérant pendant encore quelques siècles.

Toutefois, les communications militaires et gouvernementales devaient se reposer sur des dispositifs de chiffrement fiables. C'est pourquoi les cryptologues inventèrent diverses méthodes, plus simples d'utilisation que le chiffre de Vigenère mais néanmoins plus robustes que le chiffre monoalphabétique. Parmi celles-ci figure le chiffre de **substitution homophonique**.

# Les codes homophones

Dans un **chiffre de substitution homophonique** (ou **chiffre de substitution simple à représentation multiple**), un caractère du texte clair peut être associé à plusieurs caractères différents dans le texte chiffré.

Dans l'exemple de la Figure 2-6, les 26 lettres de l'alphabet sont transformées en nombres de 00 à 99. Chaque lettre a au moins une représentation numérique et la plupart en ont plusieurs. Le nombre de substituts attribués à une lettre est proportionnel à la fréquence de la lettre.

Par exemple la lettre *h* apparaît en moyenne à une fréquence de 5%, aussi se voit-elle attribuer 5 homophones : 01, 11, 24, 50 et 62. Selon une telle méthode, chaque symbole ne représente qu'un même pourcentage du texte chiffré (en l'occurrence 1%), déjouant ainsi les méthodes de cryptanalyse par analyse de fréquences.

Les représentations numériques sont lues de la ligne à la colonne ou de la colonne à la ligne, selon une convention définie au départ. En choisissant la lecture ligne / colonne, *n* a le substitut 34 car il se trouve à l'intersection de la troisième ligne et de la quatrième colonne. Le mot  *pierre* peut alors être codé en <03 27 41 51 74 17> ou <37 08 96 68 39 98>.

	0	1	2	3	4	5	6	7	8	9
0	T	H	E	P	D	A	Q	T	I	L
1	E	H	E	T	S	R	O	E	N	D
2	S	U	C	Y	H	O	A	I	A	S
3	E	O	M	B	N	I	L	P	N	R
4	S	E	L	W	T	J	I	O	Y	E
5	H	R	O	D	E	X	N	T	V	F
6	U	G	H	R	O	A	E	W	R	S
7	U	T	C	B	R	O	M	K	A	A
8	T	B	E	N	C	A	I	O	Z	M
9	N	R	A	I	S	G	E	T	E	F

Figure 2-6 : Codes homophones. Matrice de substitution

La clef du chiffre est l'assignation des homophones et le décodage se fait en cherchant le caractère correspondant au nombre approprié. Nous avons donc une correspondance un-à-plusieurs entre l'ensemble des caractères du texte clair et l'ensemble des homophones, où à chaque homophone correspond un et un seul caractère du texte clair (ce critère est essentiel si l'on veut reconstituer le message original à partir du texte chiffré).

En considérant les chiffres monoalphabétiques, les codes homophones proposent un niveau de sécurité plus élevé. Mais que dirait le cryptanalyste ? À peu près ceci : même si une lettre peut être chiffrée de différentes manières, l'alphabet chiffré reste constant, ce qui représente une différence majeure par rapport au chiffre polyalphabétique. La sécurité d'un tel code n'est pas parfaite : le chiffre homophonique n'est qu'une variante du chiffre monoalphabétique. La tâche du cryptanalyste est certes plus complexe, mais en s'appuyant sur les faiblesses de ce chiffre, le décryptement à terme du message est toujours possible (un exemple de faille : certaines lettres sont trahies par leurs relations mutuelles ; par exemple la lettre *q*, représentée par un seul symbole, est toujours suivie de la lettre *u*). Une attaque à texte clair connu est triviale, **une attaque à texte chiffré seulement ne prend que quelques secondes avec un ordinateur.**

## Les codes homophones bidimensionnels

Sans atteindre la complexité de mise en oeuvre du chiffre polyalphabétique, le Grand Chiffre de Louis XIV, élaboré en 1626 par Antoine et Bonaventure Rossignol, allait se révéler d'une efficacité rarement atteinte au cours de l'histoire : les messages secrets du Roi sont restés indéchiffrables jusqu'à la fin du XIX<sup>ème</sup> siècle ! (pourra-t-on dire la même chose dans trois cents ans à propos des mécanismes cryptologiques actuels ?). Cette variante du chiffre de substitution homophonique a notamment joué son rôle autour du mystère de l'énigme du Masque de Fer, aujourd'hui éclaircie.

Dans un premier temps, familiarisons nous à l'usage d'un code homophone bidimensionnel.

	A	B	C	D	E	G	I	L	N	O	R	T	U
A	019	155	078	120	024	011	136	060	106	046	150	072	053
B	143	036	145	035	142	066	129	116	028	188	032	122	005
C	131	093	003	119	029	104	095	034	112	100	012	033	086
D	159	037	018	105	156	064	047	154	073	002	111	071	022
E	010	130	059	149	077	128	087	027	135	054	063	115	020
G	061	114	092	089	017	023	113	144	045	096	004	085	032
I	094	025	067	042	137	169	140	097	031	070	099	107	069
L	058	127	049	161	076	101	055	048	121	038	021	123	044
N	079	146	138	090	041	103	016	152	075	062	110	083	007
O	008	160	026	158	117	126	009	141	015	050	143	013	056
R	091	118	165	068	057	030	139	132	082	148	039	108	051
T	102	167	081	157	168	098	166	153	162	124	164	084	151
U	147	043	163	001	080	040	125	052	006	133	074	109	014

Figure 2-7 : Codes homophones bidimensionnels

Le tableau ci-dessus représente un code de substitution bidimensionnel. On remplit d'abord un tableau 26 x 26 avec des nombres au hasard compris entre 1 et 676 (26x26) ou 0 et 675 ; on inscrit ensuite une lettre de l'alphabet devant chaque ligne et une au-dessus de chaque colonne. Chaque lettre dispose ainsi 26 homophones dans sa colonne, 26 homophones dans sa ligne et d'un homophone commun à l'intersection de sa ligne et de sa colonne. « tout », par exemple, pourrait être codé en <072 100 086 071> en se servant des homophones situés dans les colonnes ; il pourrait aussi être codé en <157 160 163 102> en utilisant ceux des lignes.

En outre, notons qu'il permet de coder deux messages différents avec le même jeu d'homophones. Étant donné le message codé <108 070 020 083>, quelle est la traduction correcte? En traduisant « vers le haut » (en prenant la lettre située au sommet de chaque colonne), on obtient sans ambiguïté un décodage de « tout ». En revanche en traduisant « sur le côté » (en prenant la lettre située devant chaque ligne), on obtient sans ambiguïté le message « rien ».

Dans les deux cas on a affaire à des traductions valables et des messages sensés, mais il n'y a aucun moyen de dire lequel des deux est authentique – à moins de savoir dans quelle « direction » le texte doit être interprété.

La direction, ainsi que la grille de substitution, constituent la clef de chiffrement. Aussi longtemps que ces informations restent secrètes (secret partagé uniquement par les tiers communicants), un éventuel perceur de code – à supposer qu'il ait découvert les substitutions correctes – ne peut savoir quel est le véritable sens du message. Ce système est vraiment très élégant !

Avec les codes unidimensionnels classiques, la majorité des traductions incorrectes aurait donné du charabia, indiquant de suite au cryptanalyste qu'il est sur une mauvaise voie. Mais avec ce chiffre, à moins de posséder une bonne connaissance du contexte, il est difficile de savoir quel est le vrai message. Il est même possible d'envoyer des messages contradictoires :

Faux message (Colonnes) : **abandonner tout**

Message codé : **131 160 079 162 042 062  
006 135 057 150 072 096  
069 108**

Vrai message (Lignes) : **continuer à agir**

## Le Grand Chiffre de Louis XIV

Le principe du Grand Chiffre de Louis XIV était basé sur les mêmes principes. Après avoir tenu en échec plusieurs générations de cryptanalystes, le commandant Etienne Bazeries découvrit en 1890 qu'en associant les séquences *les-en-ne-mi-s* à un groupe de chiffres qui revenait fréquemment (124-22-125-46-345), chaque nombre devait représenter une syllabe. Après plusieurs mois de travail, il finit par percer le code des Rossignol, qui renfermait des valeurs syllabiques particulièrement difficiles à trouver, et aussi des pièges (par exemple, un nombre signifiait qu'il fallait effacer le nombre précédent !). Grâce à ces travaux, on put enfin pénétrer les secrets de Louis XIV.

## Contre le chiffre de Vigenère. Méthode de Babbage et Kasiski

Au XVII<sup>ème</sup> siècle, les codes homophones avaient semble-t-il été jugés satisfaisants. Mais le XVIII<sup>ème</sup> siècle connut une évolution de la cryptanalyse. S'inspirant des méthodes de la Cour d'Autriche qui fut largement précurseur en matière d'espionnage, les chancelleries avaient compris l'intérêt stratégique de se doter d'unités de cryptanalyse efficaces. Les secrétaires au chiffre furent ainsi peu à peu contraints d'adopter le chiffre polyalphabétique de Vigenère, qualifié alors « d'indéchiffrable ». L'apparition du télégraphe (dont les origines remontent à 1753) et du code Morse en 1845, rendait nécessaire l'utilisation d'un tel chiffre.



Charles Babbage

C'est alors qu'un génie scientifique anglais, Charles Babbage, inventeur du prototype des calculateurs modernes, réussit en 1854 à faire céder le code de Vigenère. N'ayant jamais publié ses résultats, cette technique fut attribuée à Friedrich Wilhelm Kasiski, qui publia une méthode équivalente dans son livre « *Écriture secrète, et l'art du déchiffrement* » en 1863.

La méthode repose sur l'art de deviner la longueur de la clef. Connaissant cette longueur, attaquer un chiffre de Vigenère se ramène à attaquer plusieurs chiffres mono-alphabétiques. Prenons un exemple simplifié, le texte suivant :

```
XAUNMEESYIEDTLLFGSNBWQ
UFXPQTYORUTYIINUMQIEUL
SMFAFXGUTYBXXAGBHMIFII
MUMQIDEKIRIFRIRZQUHIENO
OOIGRMLYETYOVQRYSIXEOK
IYPYOIGRFBWPIYRBQURJIY
EMJIGRYKXYACPPQSPBVESTI
RZQRUFREDYJIGRYKXBLOPJ
ARNPUGEFBWMILXMZSMZYXP
NBPUMYZMEEFBUGENLRDEPB
JXONQEZTMBWOEFIIIPAHPPQ
BFLGDEMFWFAHQ
```

On s'aperçoit rapidement que les séquences UMQI et JIGRY se répètent deux fois après 30 lettres et que la séquence OIGR se répète deux fois après 25 lettres. En faisant l'hypothèse qu'une répétition provient de la même séquence du texte clair codé avec la même partie de la clef, on peut émettre une supposition : **la longueur de la clef est 5**. Partant de ce principe, il s'agit de découvrir les cinq lettres de la clef.

Si l'on découpe le texte codé en blocs de cinq en les redisant en colonne, on obtient donc pour la première colonne une suite de caractères brouillés avec un chiffre de César, décalé d'une valeur à découvrir comprise entre 1 et 26. En effectuant une analyse de fréquences sur cette première colonne, on s'aperçoit que le spectre de fréquences obtenu ressemble à celui des fréquences de l'alphabet usuel (voir Figure 2-8) à condition de faire coïncider le « I » de l'alphabet chiffré avec le « e » de l'alphabet normal. En se reportant au carré de Vigenère, on en déduit que la première lettre de la clef est probablement un **E**.

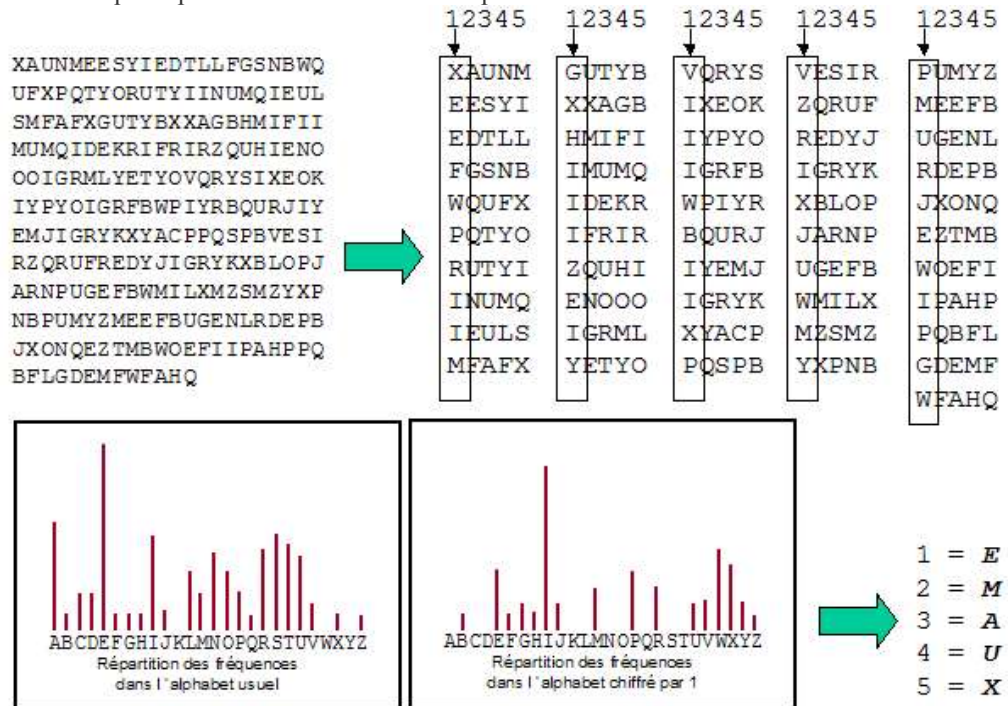


Figure 2-8 : Méthode de décryptement du chiffre de Vigenère

En répétant cette même méthode avec les quatre autres colonnes (à quelques tâtonnements près), on finit par découvrir que le mot-clef est **EMAUX** et on en déduit le texte clair suivant :

Tout passe. L'art robuste  
Seul à l'éternité,  
Le buste  
Survit à la cité.

Et la médaille austère  
Que trouve un laboureur  
Sous terre  
Révèle un empereur.

Les dieux eux-mêmes meurent,  
Mais les vers souverains  
Demeurent  
Plus forts que les airains

Sculpte, lime, cisèle ;  
Que ton rêve flottant  
Se scelle  
Dans le bloc résistant !

Ce poème s'intitule *L'Art*. Il est extrait du recueil *Émaux et Camées*, de Théophile Gautier. Survenant au début de la guerre de Crimée, cette découverte donna probablement un avantage aux Anglais sur l'ennemi russe. Procédant de la même manière, les cryptanalystes David Bates, Charles Tinker et Albert Chandler réussirent à venir à bout du chiffre des Confédérés, détruisant ainsi le mythe du disque chiffrant.

### Compléments (longueur de clef = longueur du message)

Le texte chiffré ci-dessus a été décrypté facilement parce que la longueur de la clef était trop faible. Une méthode pour renforcer la sécurité de ce mécanisme consiste à allonger la longueur de la clef, il peut même être envisageable d'employer une clef aussi longue que le message lui-même.

# Les avancées de la cryptologie après la Première Guerre mondiale

Le chiffre de Vigenère étant brisé, les cryptanalystes allaient rester invaincus pendant toute la fin du XIX<sup>ème</sup> siècle et le début du XX<sup>ème</sup>, à l'heure où, justement, l'invention et la généralisation de la TSF par les militaires demandait un chiffre robuste. La Première Guerre mondiale fut propice à l'émergence de nouveaux codes, mais ceux-ci ne résistaient bien longtemps à la perspicacité des cryptanalystes. Ces nouveaux codes alliaient généralement substitution et transposition.

## Codage par transposition

Au lieu de coder un message en remplaçant chacun de ses caractères par d'autres caractères, la technique de **transposition** consistait à permuter les caractères du texte clair. Le codage consistait à effectuer une transposition, et le décodage à remettre les caractères dans l'ordre initial.

### Note

Substitution et transposition furent les procédés les plus répandus avant l'ère informatique. Tout code se ramenait, sous une forme ou sous une autre, à des substitutions ou à des transpositions, voire les deux.

### Allergiques au maths s'abstenir

Étant donné un ensemble non ordonné  $S$  constitué par exemple de trois éléments  $a$ ,  $b$  et  $c$ ; combien d'ensembles ordonnés différents peut-on constituer à partir de  $S$ , en utilisant chacun de ses éléments une et une seule fois?

Il y a trois possibilités pour la première position puisqu'on choisit un caractère parmi trois. Il n'en reste que deux pour la seconde et, enfin, une seule solution pour la troisième puisqu'il ne reste qu'un seul caractère. Ainsi il y a  $3 \times 2 \times 1 = 6$  façons d'ordonner  $S$ .

De manière générale, un ensemble de  $n$  objets (par exemple un message comportant  $n$  caractères) peut être ordonné de  $n!$  façons différentes, où  $n!$  (Factorielle  $n$ ) est le produit de  $n$  par  $(n-1)$ , lui-même multiplié par  $(n-2)$  et ainsi de suite jusqu'à ce que le dernier terme soit 1. Il existe donc  $n!$  transpositions possibles du même message, ce qui représente un nombre important lorsque le message est long.

## Transposition des colonnes

Il est possible d'utiliser des tableaux bidimensionnels pour créer des codages par transposition simple d'un message. Par exemple, écrivons "Envahir Normandie Pas Calais X Ike X" (le caractère nul  $X$  est utilisé comme point d'arrêt mais aussi pour compléter le tableau) dans une matrice  $6 \times 5$  (voir le tableau de gauche sur la Figure 2-9).



Figure 2-9 : Codage par transposition

Le texte codé est alors : **EIMEAXNRAPLIVNNAAKAODSIEHRICSX** (un bon exercice pour le jeu de piste avec les enfants le week-end prochain, non?). Le décodage consiste à effectuer l'opération inverse: écrire en colonnes le message chiffré dans une matrice 6x5, puis à lire les lignes les unes à la suite des autres. Cette méthode de codage est appelée transposition de colonnes.

Le procédé est assez simple et prévisible. Cependant il est possible de mener la vie plus dure aux cryptanalystes en lisant les colonnes dans un ordre différent (voir tableau de droite sur la Figure 2-9).

Une lecture des colonnes de ce tableau de gauche à droite et de haut en bas donne le message chiffré suivant: **HRICSXNRAPLIAODSIEEIMEAKVNNAAK**.

### Transposition par mots-clef

L'utilisation d'une clef est un moyen destiné à faciliter la mise en pratique : la clef permet de définir le nombre de colonnes ainsi que l'ordre dans lequel permuter les colonnes. On peut décider par exemple que l'ordre de permutation sera lié à la précedence des lettres de la clef dans l'alphabet.

Prenons un exemple. Soit le texte clair suivant :

« **TRANFERER UN MILLION DE DOLLARS SUR MON COMPTE EN SUISSE** »

*(J'accepte bien sûr les dons des généreux internautes. Merci d'avance).*

Clé	G	A	S	T	O	N
Permutation	2	1	5	6	4	3
	t	r	a	n	s	f
	e	r	e	r	u	n
	m	i	l	l	i	o
	n	d	e	d	o	l
	l	a	r	s	s	u
	r	m	o	n	c	o
	m	p	t	e	e	n
	s	u	i	s	s	e

Texte en clair :

**Transférer un million de dollars sur mon compte en suisse**

Texte chiffré :

**RRIDAMPUTEMNLRMSFNOLUONE  
SUIOSCESAELEROTINRLDSNES**

Figure 2-10: Transposition par mots-clef (1)

En appliquant le principe exposé à la Figure 2-9, on obtient le texte chiffré suivant :

**RRIDAMPUTEMNLRMSFNOLUONESUIOSCESAELEROTINRLDSNES**

ce qui rend tout de suite les conversations plus claires. Une variante consiste à découper le texte clair en blocs de 6 lettres, à coder chacun des morceaux obtenus par la permutation indiquée par la clef, puis à reconcaténer les nouveaux blocs.

Clé	G	A	S	T	O	N	G	A	S	T	O	N	G	A	S	T	O	N
Permutation	2	1	5	6	4	3	2	1	5	6	4	3	2	1	5	6	4	3
Texte clair	t	r	a	n	s	f	e	r	e	r	u	n	m	i	l	l	i	o
Texte chiffré	R	T	F	S	A	N	R	E	N	U	E	R	I	M	O	I	L	L
Clé	G	A	S	T	O	N	G	A	S	T	O	N	G	A	S	T	O	N
Permutation	2	1	5	6	4	3	2	1	5	6	4	3	2	1	5	6	4	3
Texte clair	n	d	e	d	o	l	l	a	r	s	s	u	r	m	o	n	c	o
Texte chiffré	D	N	L	O	E	T	A	L	U	S	R	S	M	R	O	C	O	N
Clé	G	A	S	T	O	N	G	A	S	T	O	N	G	A	S	T	O	N
Permutation	2	1	5	6	4	3	2	1	5	6	4	3	2	1	5	6	4	3
Texte clair	m	p	t	e	e	n	s	u	i	s	s	e	m	p	t	e	e	n
Texte chiffré	P	M	N	E	T	E	U	S	E	S	I	S	P	M	N	E	T	E

Figure 2-11: Transposition par mots-clef (2)

Nous obtenons cette fois :

**RTFSANRENUERIMOILLDNLOEDALUSRSMROCONPMNETEUSESIS**

La correspondance entre la longueur du mot-clef et celle du texte clair n'est qu'une coïncidence. Pour décoder, il suffit de découper le texte en blocs de la longueur de la clef et d'appliquer la transposition inverse.

Intéressant tout cela. Mais comment le cryptanalyste voit-il les choses ?

Tout d'abord, il doit savoir déterminer la nature du système de codage (substitution ? transposition ? type de transposition, etc...). Faisons lui confiance sur ce point, un cryptanalyste entraîné connaît les différents codes et possède suffisamment de « trucs » pour les reconnaître

assez vite. Par exemple, avec une rapide analyse de fréquences, il saura qu'il ne s'agit pas d'une substitution. Par conséquent, il part du principe que ce code est une transposition.

Il va donc commencer par deviner la longueur de la clef. Il arrive que le cryptanalyste devine la présence d'un ou de plusieurs mots clefs selon le contexte (il ne faut pas s'en étonner, tous les briseurs de code recherchent frénétiquement des raccourcis ou des méthodes simples avant d'envisager de passer à des techniques plus mathématiques !). Si le message est échangé en milieu financier, pourquoi ne pas chercher des séquences de type «MILLION DE DOLLARS»? Il observe par exemple qu'à un intervalle de six lettres, apparaissent les digrammes ID, MN, OL, IO, LE, LD, qui font partie de la séquence. Il n'en faut pas plus au cryptanalyste pour formuler une hypothèse intéressante : «supposons que la clef soit de longueur 6 ». Il découpe alors le texte codé en blocs de six et les rediPOSE en colonne.

```
RTFSAN
RENUER
IMOILL
DNLOED
ALUSRS
MROCON
PMNETE
USESIS
```

Il ne lui reste plus qu'à réordonner les colonnes, ce qui est facile car il sait qu'il doit passer de IMOILL à MILLIO !

D'accord, il s'agissait d'un exemple simple. La réalité est plus complexe !

## Techniques de chiffrement utilisées lors de la Première Guerre mondiale

L'apparition de la radio et son utilisation intensive au cours de la première Guerre Mondiale fut un facteur décisif qui incitèrent les responsables gouvernementaux à considérer la cryptologie comme une discipline stratégique. Mais malgré les enjeux du conflit, aucun chiffre suffisamment robuste pour résister aux cryptanalystes n'a pu voir le jour. A peine un code était-il apparu, il était aussitôt brisé par le camp adverse.

### Le chiffre ADFGVX

C'est notamment le cas du code allemand ADFGVX, introduit juste avant la grande offensive allemande, le 21 Mars 1918. Le principe du chiffre ADFGVX consistait à répartir les 26 lettres et les dix chiffres au hasard dans une grille 6x6, ce qui offrait à peu près de 372 000 milliards de milliards de milliards de possibilités différentes (36 !) de coder un même message ! La disposition des éléments sur la grille était donc un secret que, seuls, l'émetteur et le destinataire partageaient.

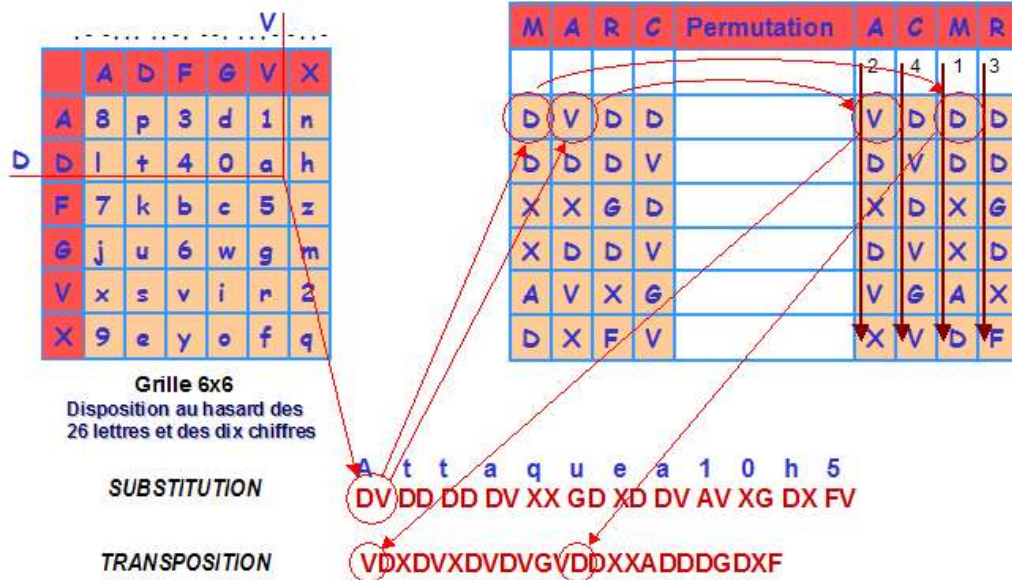


Figure 2-12: Le chiffre ADFGVX

Lors de la première étape du chiffrement, chaque lettre du message était substitué à un digramme composé des lettres indiquées par sa ligne et sa colonne (a=DV). Si le mécanisme s'était arrêté à cette première étape, nous aurions été en présence d'un simple chiffre de substitution, attaquant aisément par une analyse des fréquences.

Afin de rendre la cryptanalyse plus difficile, une seconde étape mettait en jeu l'un des mécanismes de transposition décrit ci-dessus, basé sur un mot-clef : par exemple « MARC ». En permutant les colonnes selon l'ordre alphabétique du mot-clef (ACMR) et en écrivant les lettres en descendant chaque colonne l'une après l'autre, on obtenait le message chiffré final.

La résolution de l'énigme ADFGVX en Juin 1918 par Georges Painvin permit aux Alliés de découvrir l'imminence de l'assaut final sur Paris, les Allemands étant situés à moins de 100km au nord de la capitale. L'effet de surprise étant compromis, l'ennemi fut repoussé au terme d'une bataille de cinq jours.

### Le télégramme de Zimmermann

Si les gouvernements considèrent la cryptologie comme une arme de guerre, l'épisode du télégramme de Zimmermann en illustre parfaitement la raison : son décryptement par le bureau 40 de l'Amirauté (le bureau du chiffre) changea le cours de l'Histoire.

Zimmermann, ministre des Affaires étrangères du Kaiser, projetait de lancer une guerre sous marine totale le 1<sup>er</sup> février 1917. Souhaitant neutraliser la marine américaine sans avoir à la combattre, il incita le Mexique à lancer une offensive contre les Etats-Unis, dans le but de récupérer des territoires qu'il considérait lui appartenir (Texas, Arizona, Nouveau-Mexique). Zimmerman employa un chiffre utilisé pour les communications diplomatiques de grande importance :

« ... 13486 9350 9220 76036 14219 5144 2831 17920 11347  
17142 11264 7667 7762 15099 9110 10482 97556 3569  
3670. »

Figure 2-13: Fin du télégramme de Zimmermann

Le contenu de ce télégramme décida Wilson engager les États-Unis dans le conflit, ce qui fit basculer le cours d'une guerre qui s'acheva quelques mois plus tard.

# Enigma

La cryptologie connut une véritable avancée en 1918 grâce l'invention géniale de l'Allemand Arthur Scherbius : la machine **Enigma**.

Enigma était une version plus complexe du disque à chiffrer d'Alberty. Examinons sur un exemple simplifié son principe de fonctionnement (Figure 2-14).

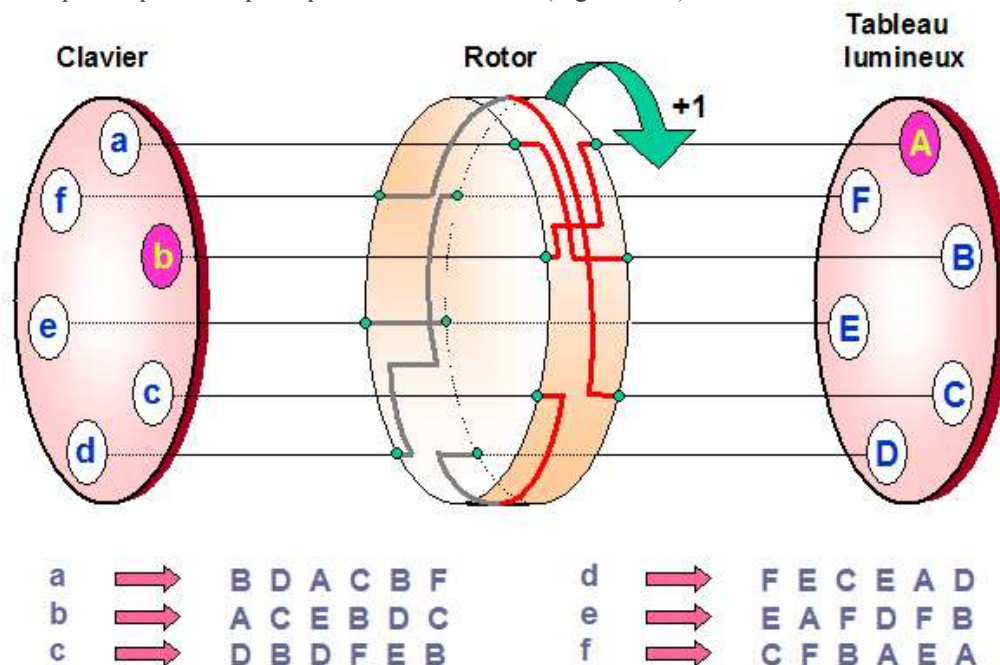


Figure 2-14: Principes de fonctionnement d'Enigma

Dans cette forme élémentaire, la machine est constituée d'un clavier destiné à la saisie du texte clair, d'un rotor, ou brouilleur, transformant chaque lettre du texte en clair en une lettre du texte chiffré, et d'un tableau lumineux pour afficher la lettre du texte chiffré. Le rotor est un tambour isolant doté de contacts électriques sur chaque face. Chaque contact est relié à un contact situé sur l'autre face par l'intermédiaire d'un fil électrique. Ce fil électrique suit un chemin particulier, défini de manière à ce que la valeur du caractère chiffré ne puisse être prévisible. Ce rotor constitue la pièce principale du système de chiffrement.

Dans l'exemple montré à la Figure 2-14, plaçons nous à l'instant « 0 », où l'on a encore rien entré au clavier. On constate par exemple que si quelqu'un saisisait la lettre « b » (texte clair), elle serait aussitôt transformée en une lettre « A » du texte chiffré. En suivant le même raisonnement, chaque lettre, a, b, c, d, e, ou f serait transformée respectivement en B, A, D, F, E, ou C, si elle était saisie à l'instant « 0 ».

Jusque là, le brouilleur met en œuvre un simple chiffre de substitution monoalphabétique, facilement attaquant par la méthode d'analyse des fréquences.

Mais ce n'est pas tout. L'idée suivante repose sur un mécanisme additionnel dont le but est de **faire pivoter d'un cran le rotor, chaque fois qu'une lettre est saisie au clavier**. Ce mécanisme revient en effet à procéder à un **changement automatique d'alphabet chiffré**, à chaque frappe clavier.

Supposons que nous nous placions maintenant à l'instant « 1 », c'est à dire immédiatement après que quelqu'un ait saisi le premier caractère. Le rotor a donc tourné d'un cran dans le sens de la flèche verte symbolisée à la Figure 2-14. En s'appuyant sur cette figure et en se représentant mentalement la nouvelle position du rotor, on constate aisément que si quelqu'un saisisait la lettre a, b, c, d, e ou f, elle serait désormais transformée respectivement en D, C, B, E, A ou F. L'alphabet chiffré a effectivement changé. En répétant le même exercice, à l'instant 2, « a, b, c, d, e, f » seraient transformées respectivement en « A, E, D, C, F, B ». Et ainsi de suite.

Revenons à l'instant 0. Si un utilisateur entrerait le message clair « bbbbbb », il produirait le message chiffré « ACEBDC » : le système de substitution polyalphabétique de Vigenère en action, et sans effort !

Certes, si l'utilisateur continue de saisir des « b » en série, il finit tôt ou tard par retomber sur la même séquence (en l'occurrence au bout de 6 frappes consécutives). Mais l'exemple de la Figure 2-14 représente une machine ultra-simplifiée, dotée seulement de 6 alphabets chiffrés.

Dans la réalité, la machine de Scherbius était autrement plus complexe. Tout d'abord un rotor de la machine Enigma comportait 26 alphabets chiffrés différents, et ensuite les premières machines comportaient 3 rotors (Figure 2-15). Lorsque le premier rotor avait effectué une rotation complète, le second rotor avançait d'un cran. Lorsque le premier rotor avait avancé de 26 fois 26 positions, le deuxième avait avancé de 26 positions et, au caractère suivant, le troisième rotor avançait d'un cran. Et ainsi de suite. Le nombre d'alphabets chiffrés était de  $26 \times 26 \times 26$  c'est à dire 17 576 ! En outre, certaines machines Enigma utilisées pendant la Seconde Guerre mondiale comportaient 4, voire 5 rotors.

Scherbius ajouta aussi un réflecteur au delà du dernier rotor, qui avait pour fonction de renvoyer le signal électrique dans les trois rotors jusqu'au tableau lumineux. Ce signal réfléchi passait par un autre chemin. Ce réflecteur n'apportait rien à la machine sur le plan cryptologique, mais la dotait d'une propriété fort intéressante qui facilitait considérablement son utilisation : saisir une lettre du texte clair permettait d'afficher la lettre correspondante du texte chiffré, saisir une lettre du texte chiffré permettait d'afficher la lettre correspondante du texte clair (à condition que les rotors eussent été initialement orientés de la même façon).

Enfin, un tableau des connexions à fiches entre le clavier et le premier brouilleur permettait d'apparier 6 couples de deux lettres, ajoutant ainsi une substitution supplémentaire à l'action des rotors.

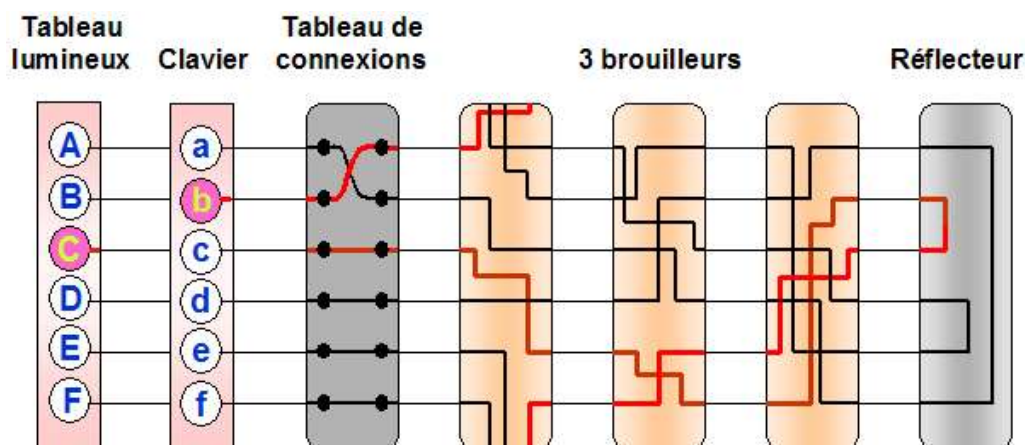


Figure 2-15: Schéma de principe de la machine complète

Au bout du compte, le nombre de façons différentes de chiffrer un même message avec Enigma était supérieur à 10 millions de milliards avec 3 rotors :

- Orientation des brouilleurs :  $26 \times 26 \times 26$  positions = 17 576 alphabets chiffrés différents
- Disposition des brouilleurs : 123, 132, 213, 231, 312, 321 = 6 positions possibles
- Tableau de connexions : 100 391 791 500 nombre de branchements possibles en appariant 6 fois 2 lettres parmi 26.

Tirant parti des leçons de la Première Guerre mondiale – et notamment du rôle important joué par les services de renseignements sur le cours de la guerre, les Allemands, détenteurs du système cryptologique le plus sûr au monde, déployèrent massivement la machine Enigma.

Dès 1925, 30 000 machines Enigma entrèrent en service chez les militaires et au sein du gouvernement, ce qui offrit un avantage incontestable à la diplomatie allemande sur l'échiquier international.



## La cryptanalyse d'Enigma

*Une vision très actuelle sur les méthodes gouvernementales destinées à briser les chiffres.*

La situation politique de la Pologne au lendemain de la Première Guerre mondiale était précaire. Elle subissait d'un côté la pression d'une Russie expansive ne songeant qu'à promouvoir son régime communiste, d'un autre la convoitise d'une Allemagne vaincue qui souhaitait reconquérir ses anciens territoires. Les perspectives étaient sombres, mais c'est en poussant l'homme dans ses retranchements qu'il réalise parfois des exploits étonnants. Ce fut le cas des cryptanalystes polonais : confrontés dès 1926 au mur des messages chiffrés d'Enigma, ils entreprirent de briser son chiffre comme s'il s'agissait de défendre leur propre survie. Ils réalisèrent une performance époustouflante dont les retombées, comme nous allons le voir, eurent une portée considérable.

La première étape de la résolution du chiffre d'Enigma fut franchie suite à la trahison d'un agent du chiffre allemand, qui fournit aux autorités polonaises les informations nécessaires à la fabrication d'une réplique exacte de la machine diplomatique. Par analogie avec les techniques actuelles, les polonais s'étaient en quelque sorte procuré les spécifications de « l'algorithme cryptologique » de l'adversaire. Ils ne leur restait plus qu'à trouver la « clef de chiffrement » d'un message donné, en l'occurrence la disposition des rotors, leur position initiale à l'instant « 0 », et les branchements du tableau des connexions.



Trouver une clef parmi des millions de milliards de possibilités confrontaient les Polonais à deux alternatives : soit ils réussissaient à se procurer un carnet de codes allemand qui donnait toutes les clefs du jour pour le mois (ce carnet était diffusé à tous les opérateurs d'Enigma), soit il fallait découvrir des raccourcis. Bien que la première méthode fut fructueuse, nous n'en parlerons pas, elle sort du cadre de cet exposé.

Nous nous focaliserons sur la deuxième méthode, un intéressant exemple de cryptanalyse. Cette attaque fut mise au point par le cryptanalyste Marian Rejewski du Biuro Szyfrow (bureau du chiffre). Pour comprendre comment il découvrit une faille dans la machine, tentons d'abord d'analyser le protocole d'échanges mis en œuvre par les Allemands.

Les Allemands disposaient d'un carnet de codes dans lequel était spécifiée la clef du jour, à savoir les branchements du tableau des connexions, l'ordre et l'orientation des rotors.

Eux mêmes fins spécialistes du chiffre, sachant notamment que toute forme de répétition est une ennemie jurée de la crypto, ils eurent l'intelligence de ne pas faire usage de la clef du jour pour chiffrer l'intégralité des messages échangés au cours d'une même journée.

La Figure 2-16 donne les éléments du protocole qu'ils avaient mis au point pour échanger un message. Ils utilisaient cette clef du jour (sur la figure appariement des lettres : P/L-N/B, ordre des rotors : 2-1-3, orientation des rotors : NAT) pour définir une autre clef de chiffrement générale, représentée par une nouvelle orientation des rotors (HAL dans l'exemple). Afin d'éviter que cette clef de chiffrement soit utilisée pour chiffrer des longueurs de texte trop importantes (encore une fois, éviter la répétition), une clef de message, elle aussi représentée par une nouvelle orientation des rotors, était sélectionnée (LIB dans l'exemple).

Par analogie avec nos systèmes actuels, cette clef de message s'apparente à la clef de session SSL. Jusque-là, le protocole est irréprochable, juste à un détail près : la clef de chiffrement générale était émise deux fois (une répétition !) pour prévenir les cas d'erreurs de l'opérateur ou les interférences radio. Erreur fatale, nous allons voir pourquoi.

Grâce à une inspiration extraordinaire, Rejewski réussit à trouver une faille dans la conception même d'Enigma, ce que nous appellerions aujourd'hui une faille dans l'algorithme.

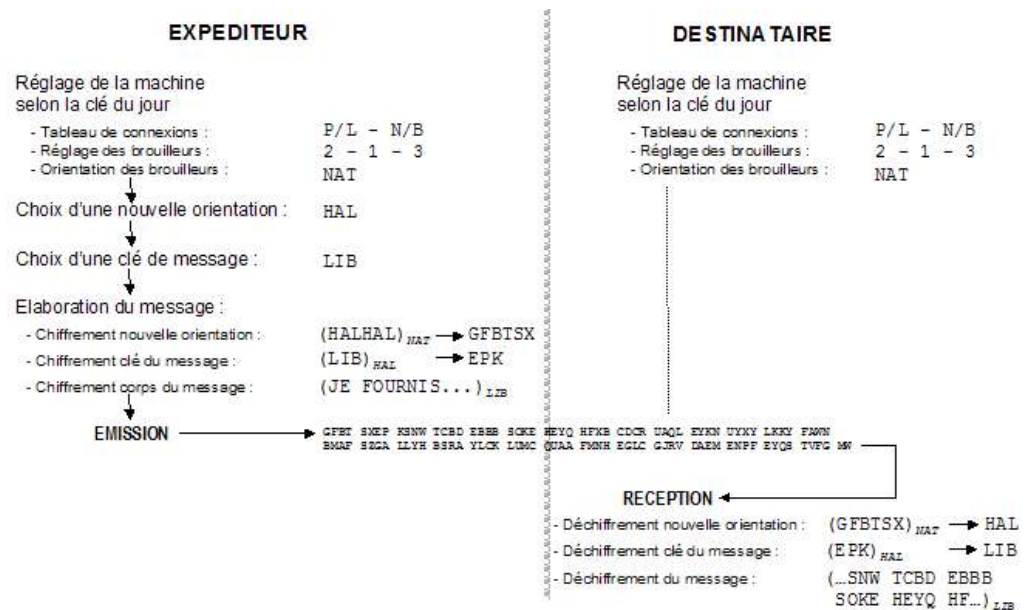


Figure 2-16: Protocole d'échanges sécurisé mis en œuvre pour Enigma

Rejewski se concentra sur les entêtes de chaque message qui, rappelons le, étaient constitués d'une double séquence des 3 mêmes lettres, chiffrées par une même clef du jour (voir texte chiffré sur la figure). Du lien qui existait entre la 1<sup>ère</sup> et la 4<sup>ème</sup> lettre, entre la 2<sup>ème</sup> et la 5<sup>ème</sup> lettre, entre la 3<sup>ème</sup> et la 6<sup>ème</sup>, Rejewski pouvait établir chaque jour un alphabet de relations pour un réglage donné des rotors et du tableau de connexions. Un exemple est fourni Figure 2-17.

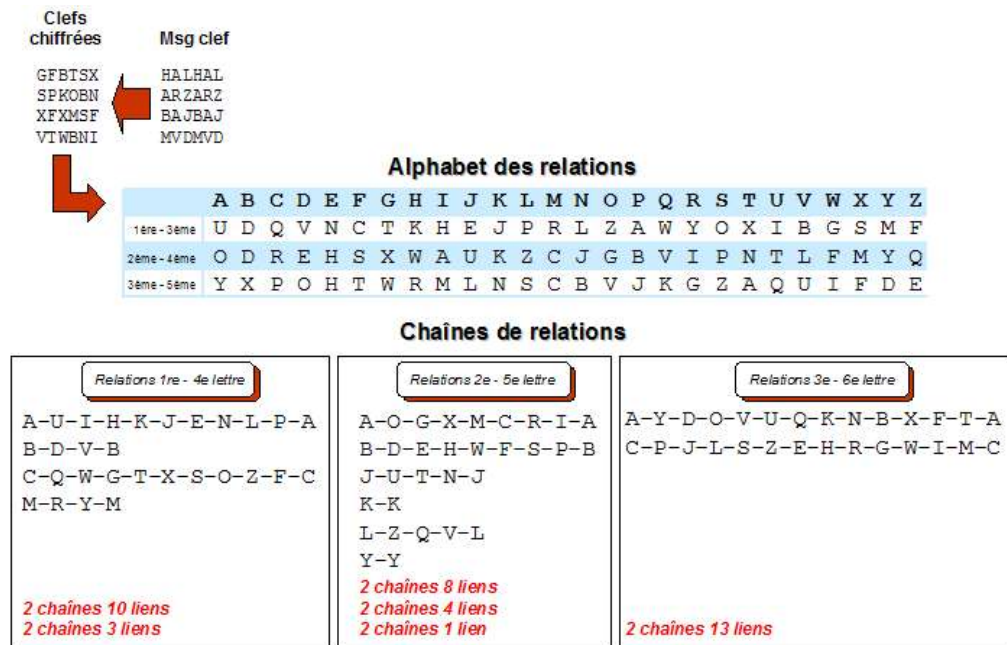


Figure 2-17: Cryptanalyse d'Enigma

A partir de l'alphabet des relations, il construisit ensuite tous les chaînages possibles entre les lettres, comme le montre l'exemple de la Figure 2-17.

C'est à ce point précis que Rejewski découvrit une faille : il s'aperçut qu'indépendamment des lettres qui interviennent dans chaque chaîne, le nombre de chaînes et le nombre de liens dans une chaîne étaient fonction du réglage des rotors, comme si ces éléments constituaient une « signature » de l'orientation initiale des rotors (les branchements du tableau des connexions n'intervenaient pas) ! Trouver la bonne orientation se ramenait à chercher une occurrence parmi 105 456 cas possibles (17 576 alphabets x 6 positions possibles des brouilleurs) ; si ce chiffre est élevé, il l'est bien moins que les 10 millions de milliards de possibilités théoriques de chiffrer un même message !

Rejewski et son équipe s'attelèrent donc à la tâche fastidieuse de consigner dans un carnet les longueurs de chaîne et le nombre de liens engendrées par les 105 456 positions des rotors. Incroyable : au bout d'un an de ce travail de fourmi, l'objectif était atteint !

La suite en découle : lorsque Rejewski recevait un message chiffré, il lui suffisait d'effectuer le relevé des chaînages, de comparer le nombre de chaînes et de liens obtenus avec les valeurs répertoriées dans son carnet et il en déduisait la position initiale des rotors ! Il n'avait plus qu'à orienter les rotors à la bonne position et, saisissant le message chiffré, il lisait le texte clair ! Ou presque... lorsque le texte clair affichait un message du type « **DETOP DE MUNIIONS TOINP DELPA** », cela voulait dire que les lettres « t » et « p » avaient été appariées. En procédant à quelques essais, il était facile de rétablir les bonnes connexions. Le tour était joué !

Les travaux du Biuro Sczyfrow ne s'arrêtèrent pas en si bon chemin. Avec Enigma, les Allemands avaient su tirer parti des progrès de la mécanique pour inventer un chiffre robuste. En réponse, Rejewski fit de même, il se servit de la mécanique pour automatiser le processus de cryptanalyse et inventa une machine à casser les chiffres. Il conçut ce qui allait devenir la fameuse « Bombe », redoutable machine à décrypter composée de six rotors bâtis sur le principe d'Enigma et fonctionnant en parallèle. Cette méthode permettait de trouver la clef du jour en deux heures environ !

Grâce au travail remarquable de Rejewski et de son équipe, la Pologne décrypta rondement – et à l'insu de tous, confidentialité oblige ! – la correspondance secrète de la diplomatie allemande, et eut accès à toute la stratégie du *blitzkrieg* d'Hitler jusqu'en 1938 !

Malheureusement, c'était de bonne guerre, l'Allemagne décida de renforcer la sécurité d'Enigma. Elle fit fabriquer deux rotors supplémentaires et fit passer le nombre de connexions possibles du tableau de 6 à 10. Les possibilités d'agencement des rotors passèrent ainsi de 6 à 60 (3 rotors parmi 5), et le nombre de clef possibles s'éleva désormais à 159 milliards de milliards !

À ce stade, la Pologne ne disposait plus du budget suffisant pour construire une version plus sophistiquée de la machine de Rejewski. Elle confia alors le résultat des travaux du Biuro Szyfrow aux Britanniques, plus exactement à la légendaire Government Code and Cypher School (GC&CS) de Bletchley Park, qui allait jouer un rôle très important en matière de décryptement au cours de la Seconde Guerre mondiale.

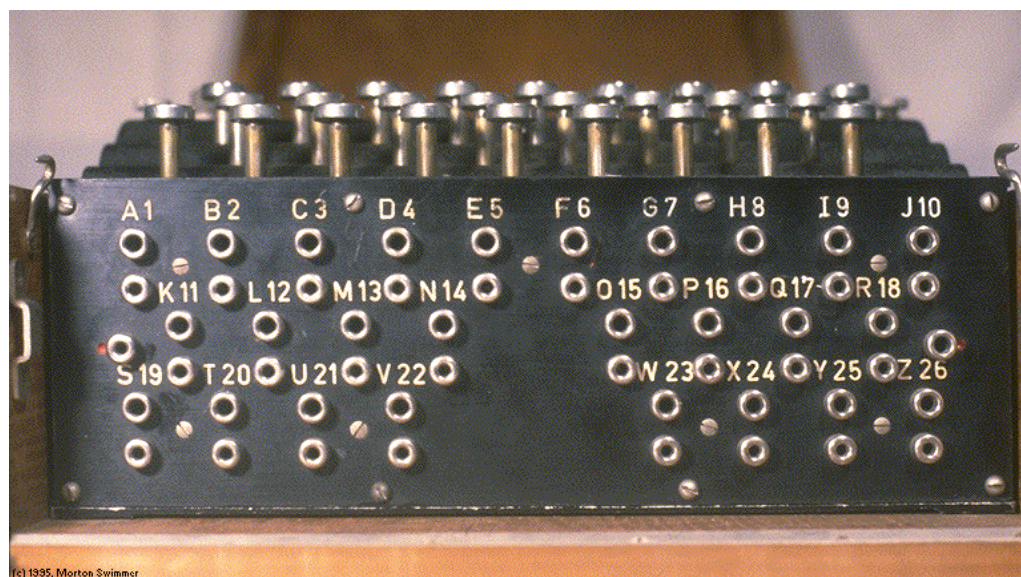


Figure 2-18: Enigma. Tableau de connexions

# Une bataille du chiffre pendant la Seconde Guerre mondiale

Durant le conflit, la sécurité des communications allemandes allait donc en grande partie reposer sur la version sophistiquée d'Enigma, où la Bombe construite par les Polonais devenait peu à peu obsolète.

Le personnel de Bletchley Park, bien plus nombreux que celui du Biuro Szcifrow, continua toutefois d'appliquer les méthodes de Rejewski, ce qui offrit sur un plateau au gouvernement britannique les plans stratégiques de l'invasion du Danemark et de la Norvège en 1940.

Les cryptanalystes remarquèrent ensuite **d'importantes faiblesses dans l'utilisation d'Enigma** : ils constatèrent notamment que les opérateurs définissaient des **clefs évidentes et prévisibles** (trois lettres qui se suivent sur le clavier (QWE), initiales d'une petite amie, ...), portant ainsi gravement atteinte à la solidité du chiffre d'Enigma.

Mais les responsables de Bletchley Park prévoyaient que les Allemands modifieraient un jour le protocole d'échange de clefs, ce qui rendrait obsolète la méthode polonaise. Aussi, confièrent-ils à l'un des plus brillants mathématiciens issus de Cambridge, Alan Turing, la mission de trouver une méthode générale pour briser le chiffre d'Enigma.

## La méthode d'Alan Turing



Turing réalisa ce que l'on appelle aujourd'hui une **attaque à texte clair connu**. En étudiant une foison d'anciens messages décryptés, il s'était aperçu que les Allemands envoyaient notamment un bulletin météo tous les jours peu après 6 heures du matin, et que le format de ce message respectait un cadre strict.

Un premier exercice consistait donc à localiser un mot probable dans le flux chiffré, par exemple « **wetter** », et de noter le cryptogramme associé, par exemple « **TERAMW** ».

Partant de cela, une méthode simpliste consistait régler la machine dans une position donnée, saisir ensuite « **wetter** » et vérifier si le texte chiffré obtenu était « **TERAMW** ».

Si ce n'était pas le cas, il fallait réitérer la même opération, avec une autre configuration, et renouveler celle-ci jusqu'à tomber sur le bon texte chiffré (ce qui s'appelle une recherche exhaustive). Le problème, c'est qu'avec 159 milliards de milliards de possibilités, il n'était pas réaliste d'espérer trouver la clef du message en un temps raisonnable. Là encore, il fallait trouver des raccourcis.

S'inspirant de la démarche de Rejewski, Turing rechercha d'abord une méthode permettant de s'affranchir du tableau des connexions. En effet, sans les maudites substitutions du tableau des connexions, le nombre de clefs possibles retombait à 1 054 560 (17 576 orientations possibles des rotors, multipliées par 60 positions potentielles des rotors).

Turing travailla alors à l'identification d'une série de boucles reliant certaines lettres du texte clair aux lettres du texte chiffré. La Figure 2-19 (à gauche) montre que dans la position réelle du rotor au moment où le message est chiffré – et que nous cherchons justement à déterminer, appelons la « P » – la machine chiffre le « w » en « T ». Le premier rotor tourne alors d'un cran, atteignant ainsi la position P+1. Elle chiffre ensuite le « e » en « E », atteint la position P+2, où nous constatons qu'elle chiffre le « t » en « R ». Et ainsi de suite. Lorsqu'elle atteint la position P+5, nous remarquons qu'elle chiffre le « r » en « W ». Une boucle a été identifiée.

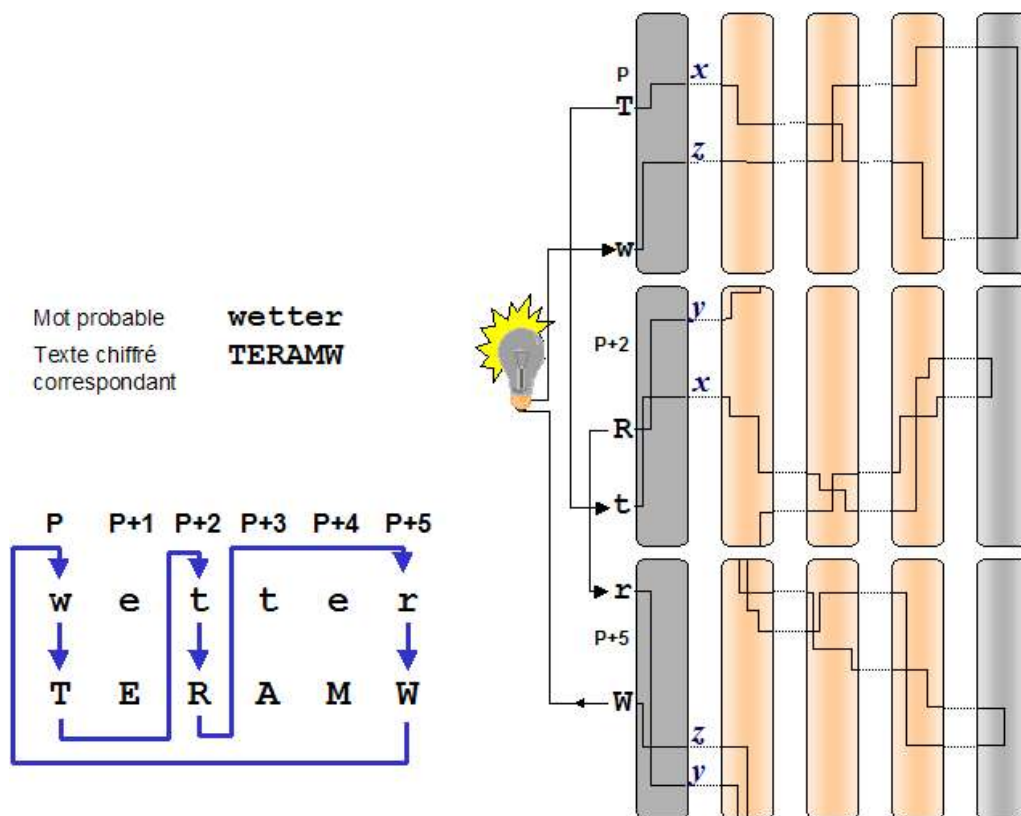


Figure 2-19: Cryptanalyse d'Enigma. Méthode de Turing

Quel est l'intérêt de cette démarche ?

Imaginons que le cryptanalyste – qui a du temps et des moyens, ne l'oublions pas – construise une machine spéciale composée de trois machines Enigma fonctionnant chacune en parallèle. Supposons qu'il établisse un circuit électrique reliant les lettres w, T (1<sup>ère</sup> machine), t, R (2<sup>ème</sup> machine), r, W (3<sup>ème</sup> machine) et une ampoule (Figure 2-19 à droite).

On imagine la suite: cette machine est conçue pour dérouler automatiquement, à vitesse rapide, et de façon synchrone (les rotors des trois machines avancent en même temps), toutes les orientations possibles des rotors à partir d'une position prédéfinie. En réglant initialement les rotors des trois machines à une position arbitraire, mais en respectant impérativement le décalage  $P_x$ ,  $P_{x+2}$  et  $P_{x+5}$ , on comprend aisément que l'on va finir par tomber, tôt ou tard, sur une position où le w est transformé en T, le t réinjecté est transformé en R, et le r en W. Dans ce cas, il s'agit de la bonne combinaison, elle est signalée par le fait que la lampe s'allume. La clef est donc trouvée!

Certes, ce raisonnement est un peu rapide, car il ne faut pas oublier le tableau des connexions. En effet, la différence majeure entre l'opérateur et notre cryptanalyste, c'est que le cryptanalyste ne connaît pas la combinaison du tableau des connexions de l'opérateur. Par exemple lorsque l'opérateur réel a saisi le « w », c'est en réalité la valeur substituée du « w » qui entre dans le circuit des rotors ; lorsque le cryptanalyste entre un « w », c'est une autre valeur qui entre dans le circuit.

Est-ce vraiment gênant ? À vrai dire, non. Observons la Figure 2-19. Lorsque le « w » est saisi, c'est en réalité une valeur « x » qui sort des rotors, elle même substituée en « T » par le tableau des connexions. Que se passe-t-il ensuite ? Le « t » est réinjecté dans la deuxième machine Enigma, passe à nouveau à travers le tableau des connexions qui est configuré de la même manière. Ce qui veut dire que le « t » est reconverti en « x » avant d'attaquer les rotors. Que s'est-il passé ? Les effets des deux tableaux de connexions s'annulent mutuellement ! Le raisonnement est exactement le même avec les tableaux des 2<sup>ème</sup> et 3<sup>ème</sup> machines. En passant sur quelques détails, on montre que le montage d'Alan Turing permettait de trouver la bonne combinaison des rotors, en faisant totalement abstraction du problème des tableaux de connexions ! La complexité du problème était donc diminué d'un facteur de l'ordre de centaines de centaines de millions ! Astucieux, non ?

Une fois la position des rotors identifiée, il suffisait de comparer les textes chiffrés obtenus aux originaux pour trouver les bonnes substitutions de la table des connexions. C'était une affaire de quelques minutes.

La suite devient on ne peut plus simple : en construisant 60 machines de ce type (correspondant aux 60 possibilités de choisir 3 rotors parmi 5), chaque machine devait s'acquitter d'une tâche beaucoup plus accessible : trouver une combinaison parmi 17 576. En effectuant un test par seconde, la clef du jour était trouvée en cinq heures.

Voilà comment donner un aperçu, en se basant sur un cas réel, de l'ingéniosité, de la haute technicité, de la motivation acharnée, des moyens colossaux... mis en oeuvre pour briser des chiffres. En matière de finance, de défense ou d'intelligence économique, rien ne résiste à un adversaire motivé. Mieux vaut le savoir !

# Conclusions - Les petites leçons de l'Histoire

L'histoire de la cryptologie est étonnante, passionnante et riche en enseignements. Étudier la cryptologie à travers l'histoire n'est pas un exercice anodin, pratiqué dans le seul but de satisfaire la curiosité personnelle. Les erreurs commises par les utilisateurs, la démarche des cryptanalystes ou l'attitude des gouvernements face à la cryptologie en disent long sur la sécurité effective de nos systèmes actuels. En bref, il serait naïf de croire que la mise en oeuvre de n'importe quel produit de sécurité offre systématiquement une sécurité de confiance. Les gouvernements cherchent à garder le pouvoir sur l'information. Tous les moyens sont bons, y compris celui qui consiste à injecter des fonctions permettant d'affaiblir un chiffre ou d'accéder, par des moyens cachés, aux clefs de chiffrement. Pour qu'un système d'information soit réellement sécurisé, il ne faut pas choisir n'importe quelle architecture. De plus, des contraintes doivent être respectées en ce qui concerne la mise en oeuvre et, surtout, l'utilisation. Sur ce dernier point, l'utilisateur représente indiscutablement l'un des maillons les plus faibles.

Quelques réflexions en vrac à propos des systèmes cryptologiques et de leur utilisation :

- Les comportements imprudents des utilisateurs minent la sécurité, bien plus que les faiblesses éventuelles des composants de sécurité.
- Ne jamais sous-estimer la puissance de travail de l'adversaire. Quand le jeu en vaut la chandelle (flux financiers, communications militaires...), l'adversaire saura déplacer des montagnes pour briser une sécurité.
- Corrolaire : si une information est secrète, elle a de bonnes chances d'être lue par l'adversaire, par un moyen ou un autre...
- Partir du principe que tout algorithme de chiffrement est connu de l'adversaire, y compris – et surtout – si cet algorithme est tenu secret ! Un principe de base en sécurité : les cryptosystèmes les plus sûrs reposent aujourd'hui sur des algorithmes **publics**. **La sécurité repose uniquement sur le secret de la clef**
- Les attaques sérieuses contre les systèmes cryptologiques se basent sur des techniques beaucoup plus subtiles que la simple « force brute » (recherche exhaustive). Les cryptanalystes trouvent des raccourcis insoupçonnés qui diminuent d'un facteur abyssal la robustesse théorique des cryptosystèmes.
- Une clef de chiffrement est un élément primordial de la sécurité du système. Si la clef utilisée est prévisible, l'adversaire pourra toujours la deviner et déchiffrer le message. Une clef de chiffrement doit être **aléatoire**.
- Toute forme de répétition donne au cryptanalyste une prise pour casser le système. Il est impératif de changer régulièrement toute clef de chiffrement (ou mot de passe...).

## Une remarque à propos des « petites » sécurités

« **Rendre plus compliqué** », la belle expression ! Le subterfuge providentiel par lequel on s'extirpe triomphalement d'une fuite peu glorieuse devant un problème de sécurité. L'expression si souvent entendue dans la conception des systèmes informatiques actuels ! Personne ne veut traiter la sécurité, on ne veut même pas en entendre parler, cela coûte trop cher, alors on s'achète une bonne conscience avec des petites solutions, en se disant que cela rendra toujours l'attaque « plus compliquée ».

Quelle sottise ! Comme si les petites sécurités impressionnaient nos adversaires ! En matière de sécurité, ingénieurs, concepteurs de systèmes, bâtisseurs sommes tous des amateurs. Les attaquants, seuls, sont des professionnels.

L'expérience de la cryptanalyse d'Enigma en est une éclatante illustration. Il est vrai que le tableau des connexions avait de quoi forcer l'admiration : faire passer un nombre de clefs possibles de 1 054 560 à 159 milliards de milliards, il y avait matière à en imposer au

néophyte ! Mais derrière ces apparences ronflantes, se cachait une banale substitution monoalphabétique, technique que l'homme savait briser depuis mille ans. Certes, cela rendait l'attaque plus compliquée. Seulement les cryptanalystes s'en sont débarrassé comme si elle n'existait pas !

Une leçon à méditer...

# ANNEXE A : CHIFFRES DÉRIVÉS

## SYSTÈME DE GRONFELD

Au XVII<sup>ème</sup> siècle, le belge Gronfeld proposait une variante du chiffre de César mettant en jeu la notion de « clef additive ». Le principe consistait à décaler les lettres d'une quantité variable, selon une valeur indiquée par la clef. Ce système était en fait une sorte de chiffre de Vigenère.

## LE CODE REBECCA

"The Key to Rebecca" est le titre d'un roman d'aventure écrit par l'auteur britannique Ken Follett. Publié sous le titre français "Le Code Rebecca", ce roman retrace une aventure d'espionnage imaginaire se déroulant au Caire en 1942, au moment où Rommel, venant de prendre Tobrouk, pensait que la conquête de l'Égypte n'était plus qu'une question de jours. La faille ? Elle est cette fois matérielle : chiffreur et déchiffreur devaient garder le code à portée de la main... donc à portée de "l'ennemi" !

Ce code livresque se retrouve également dans un ouvrage "Le Cercle de la Croix" de Ian Pears (roman policier se déroulant dans l'Angleterre du XVII<sup>ème</sup> siècle).

## JEAN TRITHÈME (1499)

Petite litanie du dimanche à la messe :

*Dans la félicité à perpétuité,  
Dans son royaume à perpétuité,  
En Paradis à perpétuité,  
Ainsi qu'en toute éternité;*

*Dans la gloire à perpétuité,  
Mais dans son règne;  
Sempiternel, toujours dans la félicité,  
Tant dans la lumière que dans la  
béatitude,  
Et toujours dans la gloire à  
perpétuité,  
Mais dans son règne;*

*En une infinité encore à perpétuité,  
Comme dans la gloire autant que dans les  
Cieux,  
A tout jamais, oui ! à tout jamais à  
perpétuité;*

*Dans son royaume et dans la félicité,  
Irrévocablement, dans son royaume,  
Et sans cesse qu'il soit à perpétuité dans la  
lumière,  
Et encore à perpétuité !*

Pseudo-litanie ou petit chiffre par substitution ? Remplaçons chaque expression par la lettre associée :

A = dans les cieux  
B = à tout jamais  
C = un monde sans fin  
D = en une infinité  
E = à perpétuité  
F = sempiternel  
G = durable  
H = sans cesse  
I-J = irrévocablement  
K = éternellement  
L = dans la gloire  
M = dans la lumière

N = en paradis  
O = toujours  
P = dans la divinité  
Q = dans la déité  
R = dans la félicité  
S = dans son règne  
T = dans son royaume  
U-V-W = dans la béatitude  
X = dans la magnificence  
Y = au trône  
Z = en toute éternité

Voici donc le message déchiffré : « Retenez les formules de l'Abbé Trithème ».

Jean Trithème ("Ioannis Trithemius" en latin), abbé de son état, est considéré comme un père de la cryptologie. C'est lui qui le premier publia en latin un ouvrage traitant de cryptologie : "Polygraphiae" était le nom de son livre, paru en 1499 en Germanie (à Spanheim plus précisément), pour le compte de l'empereur Maximilien. La lithanie ci-dessus est basée sur l'un des 14 alphabets du système de chiffre de Trithème. Ainsi, chaque lettre est remplacée par un mot ou petit groupe de mots; des mots inutiles sont ajoutés pour articuler le texte.

Trithème a notamment défini des chiffres qui ne trahissaient quasiment pas leur présence (en tant qu'abbé, quoi de plus normal de tomber sur une litanie en fouillant dans ses archives !).

## LE CHIFFRE DES NIHILISTES

Emprisonnés dans les geôles du Tsar, les Nihilistes russes avaient imaginé un système pour communiquer entre eux : en tapant sur les murs ou sur la tuyauterie ! Les matons pouvaient eux aussi entendre les coups sourds frappés sur les murs, il fallait donc que le système soit ingénieux. Ce fut la cas, et les services du Tsar finirent par l'adopter. Le chiffre des Nihilistes a même été utilisé au delà de la seconde guerre mondiale, avec cependant quelques améliorations importantes.

### 1<sup>er</sup> système

Une grille 6 x 6, où chaque case correspond à une lettre de l'alphabet cyrillique. En langues romanes / germaniques, cela peut donner ceci :

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I - J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	X	X	Y	Z

Par exemple, "BONJOUR" pouvait être transmis comme suit :

1/2 - 3/4 - 3/3 - 2/4 - 3/4 - 4/2.

La barre (/) était représentée par un bref temps d'arrêt, les tirets (-) par des temps d'arrêt plus longs.

Pour compliquer le processus, les prisonniers convenaient d'un mot de passe servant à remplir la grille. Par exemple : « DIFFICILE »; le mot de passe était « DIFCLE » (les doublons étaient éliminés). La grille était complétée exactement comme dans le cas du chiffre de César étudié plus haut :

	1	2	3	4	5
1	D	E	K	Q	V
2	I	A	M	R	W
3	F	B	N	S	X
4	C	G	O	T	Y
5	L	H	P	U	Z

"BONJOUR" devenait ainsi :

3/2 - 4/3 - 3/3 - 2/1 - 4/3 - 5/4 - 2/4.

Cette première méthode n'est qu'une substitution classique très facile à décrypter. Les Nihilistes définirent une variante plus complexe.

### 2<sup>ème</sup> système

Ils transformaient une clef littérale en clef numérique. Sur base du même carré ils transformaient un texte clair en antigramme; ils y ajoutaient la clef numérique pour former le cryptogramme définitif.

Un petit exemple historique : les Nihilistes s'étaient mis d'accord à l'avance sur la clef littérale : "CHEMIN DE FER", et il fallait transmettre le message suivant : " FAITES SAUTER LE PALAIS D'HIVER ". Avec la première grille, cela donnait ce qui suit.

clef littérale : C H E M I N D E F E R C H E M I N D E F E

clef numérique : 13 23 15 32 24 33 14 15 21 15 42 13 23 15 32 24 33  
14 15 21 15

Texte clair : F A I T E S S A U T E R L E P A L A I S D

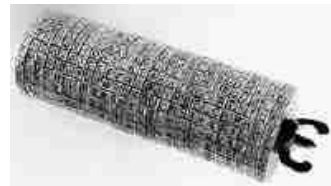
Antigramme : 21 11 24 44 15 43 11 45 44 15 42 31 15 35 11 31 11  
24 43 14 23

Cryptogramme 34 34 39 76 39 76 57 26 66 59 57 55 54 30 67 35 64  
25 39 64 29

## LE DISQUE CHIFFRANT DE JEFFERSON



Thomas Jefferson, troisième président des États Unis et principal auteur de la Déclaration d'indépendance, fut aussi l'inventeur d'un disque chiffant composé de 26 cylindres en bois disposés sur le même axe



## LE PROCÉDÉ PLAYFAIR

Le procédé Playfair fut inventé en 1854 et utilisé par les Britanniques pendant la Première Guerre mondiale. Il s'agit d'un chiffre à substitution simple par polygramme où les lettres sont chiffrées par paires.

# ANNEXE B

## Un exemple de cryptanalyse

Voici un extrait du *Scarabée d'Or* d'Edgar Poe, œuvre considérée par les cryptologues professionnels comme l'une des plus brillantes démonstrations de la cryptanalyse d'un chiffre monoalphabétique.

« (...) Ici, Legrand, ayant de nouveau chauffé le vélin, le soumit à mon examen. Les caractères suivants apparaissaient en rouge, grossièrement tracés entre la tête de mort et le chevreau :

53 ++++ + 305))6\* ;4826)4++).4++) ;806\*;48 + 8 P 60))85 ;1 ++(  
; : ++\*8  
+ 83(88)5\* + ;46( ;88\*96\*? ;8)\* ++( ;485) ;5\* + 2:\* ++( ;4956\*2  
(5\*-4)8 P 8\* ;4069285) ;)6+8)4 +++++ ; 1( ++9 ;48081 ;8:8 ++1;48  
+ 85 ;  
4)485 + 528806\*81(++9 ;48 ;(88 ;4( ++?34 ;48)4++ ;161 ;:188 ;  
++ ? ;

- Mais, dis-je, en lui tendant la bande de vélin, - je n'y vois pas plus clair. Si tous les trésors de Golconde devaient être pour moi le prix de la solution de cette énigme, je serais parfaitement sûr de ne pas les gagner.

- Et cependant, dit Legrand, la solution n'est certainement pas aussi difficile qu'on se l'imaginait au premier coup d'oeil. Ces caractères, comme chacun pourrait le deviner facilement, forment un chiffre, c'est-à-dire qu'ils présentent un sens ; mais, d'après ce que nous savons de Kidd, je ne devais pas le supposer capable de fabriquer un échantillon de cryptologie bien abstruse. Je jugeai donc tout d'abord que celui-ci était d'une espèce simple, - tel cependant qu'à l'intelligence grossière du marin il dût paraître absolument insoluble sans la clef.

- Et vous l'avez résolu, vraiment ?

- Très-aisément ; j'en ai résolu d'autres dix mille fois plus compliqués. Les circonstances et une certaine inclination d'esprit m'ont amené à prendre intérêt à ces sortes d'énigmes, et il est vraiment douteux que l'ingéniosité humaine puisse créer une énigme de ce genre dont l'ingéniosité humaine ne vienne à bout par une application suffisante. Aussi, une fois que j'eus réussi à établir une série de caractères lisibles, je daignai à peine songer à la difficulté d'en dégager la signification.

" Dans le cas actuel, - et, en somme, dans tous les cas d'écriture secrète, - la première question à vider, c'est la langue du chiffre : car les principes de solution, particulièrement quand il s'agit des chiffres les plus simples, dépendent du génie de chaque idiome, et peuvent être modifiés. En général, il n'y a pas d'autre moyen que d'essayer successivement, en se dirigeant suivant les probabilités, toutes les langues qui vous sont connues jusqu'à ce que vous ayez trouvé la bonne. Mais, dans le chiffre qui nous occupe, toute difficulté à cet égard était résolue par la signature. Le rébus sur le mot Kidd n'est possible que dans la langue anglaise. Sans cette circonstance, j'aurais commencé mes essais par l'espagnol et le français, comme étant les langues dans lesquelles un pirate des mers espagnoles aurait dû le plus naturellement enfermer un secret de cette nature. Mais, dans le cas actuel, je présuimai que le cryptogramme était anglais.

" Vous remarquez qu'il n'y a pas d'espaces entre les mots. S'il y avait eu des espaces, la tâche eût été singulièrement plus facile. Dans ce cas, j'aurais commencé par faire une collation et une analyse des mots les plus courts, et, si j'avais trouvé, comme cela est toujours probable, un mot d'une seule lettre, a ou I (un, je) par exemple, j'aurais considéré la solution comme assurée. Mais, puisqu'il n'y avait pas d'espaces, mon premier devoir était de relever les lettres prédominantes, ainsi que celles qui se rencontraient le plus rarement. Je les comptai toutes, et je dressai la table que voici :

Le caractère 8 se trouve 33 fois.    Le caractère + et l se trouve 8 fois.

Le caractère ; se trouve 26 fois.	Le caractère 0 se trouve 6 fois.
Le caractère 4 se trouve 19 fois.	Le caractère 9 et 2 se trouve 5 fois.
Le caractère ++ et ) se trouve 16 fois.	Le caractère : et 3 se trouve 4 fois.
Le caractère * se trouve 13 fois.	Le caractère ? se trouve 3 fois.
Le caractère 5 se trouve 12 fois.	Le caractère P se trouve 2 fois.
Le caractère 6 se trouve 11 fois.	Le caractère - et . se trouve 1 fois.

" Or, la lettre qui se rencontre le plus fréquemment en anglais est e. Les autres lettres se succèdent dans cet ordre : a o i d h n r s t u y c f g l m w b k p q x z. E prédomine si singulièrement, qu'il est très-rare de trouver une phrase d'une certaine longueur dont il ne soit pas le caractère principal.

"Nous avons donc, tout en commençant, une base d'opérations qui donne quelque chose de mieux qu'une conjecture. L'usage général qu'on peut faire de cette table est évident ; mais, pour ce chiffre particulier, nous ne nous en servons que très médiocrement. Puisque notre caractère dominant est 8, nous commencerons par le prendre pour l'e de l'alphabet naturel. Pour vérifier cette supposition, voyons si le 8 se rencontre souvent double ; car l'e se redouble très fréquemment en anglais, comme par exemple dans les mots : meet, fleet, speed, seen, been, agree, etc. Or, dans le cas présent, nous voyons qu'il n'est pas redoublé moins de cinq fois, bien que le cryptogramme soit très court.

" Donc 8 représentera e. Maintenant, de tous les mots de la langue, the est le plus utilisé ; conséquemment, il nous faut voir si nous ne trouverons pas répétée plusieurs fois la même combinaison de trois caractères, ce 8 étant le dernier des trois. Si nous trouvons des répétitions de ce genre, elles représenteront très probablement le mot the. Vérification faite, nous n'en trouvons pas moins de 7 ; et les caractères sont ;48. Nous pouvons donc supposer que ; représente t, que 4 représente h, et que 8 représente e, - la valeur du dernier se trouvant ainsi confirmée de nouveau. Il y a maintenant un grand pas de fait.

"Nous n'avons déterminé qu'un mot, mais ce seul mot nous permet d'établir un point beaucoup plus important, c'est-à-dire les commencements et les terminaisons d'autres mots. Voyons, par exemple, l'avant-dernier cas où se présente la combinaison ;48, presque à la fin du chiffre. Nous savons que le ; qui vient immédiatement après est le commencement d'un mot, et des six caractères qui suivent ce the, nous n'en connaissons pas moins de cinq. Remplaçons donc ces caractères par les lettres qu'ils représentent, en laissant un espace pour l'inconnu :

*t eeth.*

" Nous devons tout d'abord écarter le th comme ne pouvant pas faire partie du mot qui commence par le premier t, puisque nous voyons, en essayant successivement toutes les lettres de l'alphabet pour combler la lacune, qu'il est impossible de former un mot dont ce th puisse faire partie. Réduisons donc nos caractères à :

*t ee,*

et reprenant de nouveau tout l'alphabet, s'il le faut, nous concluons au mot tree (arbre), comme à la seule version possible. Nous gagnons ainsi une nouvelle lettre, r, représentée par (, plus deux mots juxtaposés, the tree (l'arbre). Un peu plus loin, nous retrouvons la combinaison ;48, et nous nous en servons comme de terminaison à ce qui précède immédiatement. Cela nous donne l'arrangement suivant :

*the tree ;4(t?34 the,*

ou, en substituant les lettres naturelles aux caractères que nous connaissons,

*the tree thr t?3h the.*

"Maintenant, si aux caractères inconnus nous substituons des blancs ou des points, nous aurons :

*the three thr... h the,*

et le mot through (par, à travers) se dégage pour ainsi dire de lui-même. Mais cette découverte nous donne trois lettres de plus, o, u et g, représentées par t, ? et 3. "Maintenant, cherchons attentivement dans le cryptogramme des combinaisons de caractères connus, et nous trouverons, non loin du commencement, l'arrangement suivant :

*83(88, ou egree,*

qui est évidemment la terminaison du mot degree (degré), et qui nous livre encore une lettre d, représentée par +. Quatre lettres plus loin que ce mot degree, nous trouvons la combinaison :

*o 46( 088\*),*

dont nous traduisons les caractères connus et représentons l'inconnu par un point; cela nous donne:

*th.rtee\*,*

arrangement qui nous suggère immédiatement le mot thirteen (treize), et nous fournit deux lettres nouvelles, l, et n, représentées par 6 et \*. Reportons-nous maintenant au commencement du cryptogramme, nous trouvons la combinaison :

*53++++ +*

"Traduisant comme nous avons déjà fait, nous obtenons

*.good,*

ce qui nous montre que la première lettre est un a, et que les deux premiers mots sont a good (un bon, une bonne). Il serait temps maintenant, pour éviter toute confusion, de disposer toutes nos découvertes sous forme de table. Cela nous fera un commencement de clef :

5 représente	a	6 représente	i
+ représente	d	* représente	n
8 représente	e	++ représente	o
3 représente	g	( représente	r
4 représente	h	; représente	t
		? représente	u

" Ainsi, nous n'avons pas moins de onze des lettres les plus importantes, et il est inutile que nous poursuivions la solution à travers tous ses détails.

Je vous en ai dit assez pour vous convaincre que des chiffres de cette nature sont faciles à résoudre, et pour vous donner un aperçu de l'analyse raisonnée qui sert à les débrouiller. Mais tenez pour certain que le spécimen que nous avons sous les yeux appartient à la catégorie la plus simple de la cryptologie. Il ne me reste plus qu'à vous donner la traduction complète du document, comme si nous avions déchiffré successivement tous les caractères. La voici :

*A good glass in the bishop's hostel in the devil's seat forty-one degrees and thirteen minutes northeast and by north main branch seventh limb east side shoot from the left eye of the death's-head a bee-line from the tree through the shot fifty feet out.*

*(Un bon verre dans l'hostel de l'évêque dans la chaise du diable quarante et un degrés et treize minutes nord-est quart de nord principale tige septième branche côté est lâchez de l'oeil gauche de la tête de mort une ligne d'abeille de l'arbre à travers la balle cinquante pieds au large.)*

- Mais, dis-je, l'énigme me paraît d'une qualité tout aussi désagréable qu'auparavant. Comment peut-on tirer un sens quelconque de tout ce jargon de chaise du diable, de tête de mort et d'hostel de l'évêque ?

- Je conviens, répliqua Legrand, que l'affaire a l'air encore passablement sérieux, quand on y jette un simple coup d'oeil.

(...). »

*C'est à peu près suffisant pour ce qui concerne le sujet de notre exposé. Si des passionnés désirent connaître la signification réelle de ce cryptogramme, la meilleure solution est de lire ce conte dans son intégralité.*