CHAPITRE IV

LA SUBSTITUTION PAR BIGRAMMES

EXPOSÉ DU PROCÉDÉ

La **substitution par bigrammes** procède du même principe que la substitution simple. Mais au lieu de chiffrer isolément chaque lettre du texte clair, on découpe celui-ci en groupes de deux lettres ou bigrammes, et l'on substitue ensuite à chacun de ces bigrammes un groupe chiffré qui lui correspond, un même bigramme du clair étant toujours représenté par le même signe conventionnel.

Par exemple, le clair « Forte colonne ennemie sur route de Laon à Soissons » sera traduit par :

Fo rt ec ol on ne en ne mi es ur ro ut ed
TX RH DO PS FU VK OZ VK LA PX IH ZL MF UC

el ao na so is so ns
JB SU KH BV YB BV QF

Dans cet exemple, on le voit, les bigrammes *ne* et *so* du clair, qui se répètent deux fois, sont respectivement chiffrés par un même bigramme VK et BV.

Le nombre des bigrammes que peuvent former les 26 lettres de l'alphabet étant de 676, on sera conduit pour les représenter à adopter : soit des groupes de 2 lettres, soit des groupes de 3 chiffres.

Tableaux de bigrammes.

On peut utiliser pour la substitution par bigrammes des « tableaux de bigrammes » qui jouent le même rôle que le tableau de correspondance (alphabet du clair - alphabet de chiffrement) employé dans la substitution simple.

Dans un tel tableau, les groupes chiffrants peuvent être ordonnés, comme dans les deux exemples ci-dessous. En pareil cas, un seul document sert à la fois au chiffrement et au déchiffrement.

	A	B	C	D	E	F	G	H	I	J
A	301	302	303	304	305	306	307	308	309	310
B	327	328	329	330	331	332	333	334	335	336
C	353	354	355	356	357	358	359	360	361	362
D	379	380	381	382	383	384	385	386	387	388
E	405	406	407	408	409	410	411	412	413	414
F	431	432	433	434	435	436	437	438	439	440

	A	B	C	D	E	F	G	H	I	J	K	L	M
A	KV	KW	KX	KY	KZ	LA	LB	LC	LD	LE	LF	LG	LH
B	LV	LW	LX	LY	LZ	MA	MB	MC	MD	ME	MF	MG	MH
C	MV	MW	MX	MY	MZ	NA	NB	NC	ND	NE	NF	NG	NH
D	NV	NW	NX	NY	NZ	OA	OB	OC	OD	OE	OF	OG	OH

Ils peuvent au contraire être disposés dans le tableau de façon tout à fait arbitraire ; deux tables sont alors nécessaires : une première table sert au chiffrement, une deuxième table pour le déchiffrement.

En somme, un tableau de bigrammes peut être considéré comme un « répertoire », ordonné ou à bâtons rompus suivant les cas (voir Chapitre v).

Certains tableaux de bigrammes sont réciproques, ce qui permet d'éviter toute loi simple dans la formation des groupes chiffrants (voir exemples ci-dessus) et d'avoir besoin de deux tables distinctes pour les opérations de chiffrement et de déchiffrement. Dans de tels tableaux, si par exemple *ad* du clair est traduit par DI du crypto, réciproquement *di* du clair sera traduit par AD du crypto.

	A	B	C	D	E	F	G	H	I	J	K	L	M
A	cg	hu	bk	di	mj	cp	sz	qt	og	rg	du	if	ev
B	df	ko	vy	zf	lk	mq	tn	le	bi	uj	ac	mb	dl
C	yg	jv	fc	te	xx	st	aa	mk	gy	bx	cz	mx	py
D	tv	nm	sp	zk	qu	ba	bx	ff	ad	cw	ei	bm	qw

Les tableaux de bigrammes sont d'un emploi peu commode, et sujet à erreurs.

Un procédé plus souvent utilisé est celui de la substitution orthogonale et diagonale par bigrammes, dit « système Play Fair ».

SUBSTITUTION PAR BIGRAMMES SYSTÈME PLAYFAIR

On se sert d'un carré de 25, pouvant être construit sur une clef, comme il a été exposé précédemment (voir Chapitre III, page 43). On évite ainsi l'emploi d'un tableau, difficile à modifier, et d'un maniement incommode.

Soit le carré de 25 ci-dessous, construit sur le mot clef **ALGÉRIE TUNISIE MAROC.**

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

Chiffrement.

Chaque bigramme du texte clair est chiffré successivement, d'après les règles suivantes :

— si les deux lettres du clair se trouvent sur **la même ligne** du tableau, prendre les deux lettres qui se trouvent **à droite** des lettres à chiffrer

la = DB it = KP ga = AB (1)

— si les deux lettres du clair se trouvent sur **une même colonne** du tableau, prendre les deux lettres qui se trouvent **en dessous** des lettres à chiffrer

dr = RP ho = TL rs = PZ

Dans les deux cas ci-dessus, il y a **substitution orthogonale par bigrammes**

— si les deux lettres du clair ne se trouvent **ni sur une même ligne, ni sur une même colonne,** prendre les deux lettres qui occupent les sommets du rectangle construit sur les deux lettres à chiffrer, la première de ces deux lettres étant celle qui se trouve sur la même ligne que la première lettre du clair.

nt = VK es = RN du = GP

Dans ce cas, il y a **substitution diagonale par bigrammes.**

(1) Dans toute cette étude nous représenterons les bigrammes du clair par des minuscules et les bigrammes du crypto par des majuscules.

C'est ainsi qu'avec le carré ci-dessus, le clair « escadrille numéro sept » sera chiffré :

es	ca	dr	il	le	nu	me	ro	se	pt
RN	MG	RP	TA	BH	XK	YF	HZ	NR	UP

S'il se présente dans le texte clair une lettre doublée, on intercalera une nulle entre les deux lettres à chiffrer (1).

Par exemple, pour chiffrer « l'ennemi est toujours... » on découpera le clair de la façon suivante :

le	nk	ne	mi	es	tk	to	uj	ou	rs	.	.	.	.
BH	YN	YK	AQ	RN	PT	VL	XU	CT	PZ	.	.	.	.

Déchiffrement.

Il procède des opérations inverses du chiffrement et n'offre aucune difficulté :

— si les deux lettres du crypto se trouvent sur une même ligne, on prendra pour lettres du clair les deux lettres qui se trouvent à leur gauche ;

— si les deux lettres du crypto se trouvent sur une même colonne, on prendra pour lettres du clair les deux lettres qui se trouvent au-dessus ;

— si les deux lettres du crypto ne se trouvent ni sur une même ligne ni sur une même colonne, on prendra pour lettres du clair celles qui occupent les sommets du rectangle construit sur les deux lettres du crypto, la première lettre du bigramme clair étant celle qui se trouve sur la même ligne que la première lettre du bigramme crypto.

MH	EH	VX	TQ	HJ	HK	DF	BR	KN
of	fe	ns	iv	er	et	ar	de	ek

DÉCRYPTEMENT DES SUBSTITUTIONS PAR BIGRAMMES

Le décryptement des substitutions par bigrammes repose sur la notion de fréquence.

Il existe, pour chaque langue, des tableaux de fréquences de bigrammes, qui laissent apparaître des bigrammes très fréquents, fréquents, rares et très rares.

(1) Un cryptogramme provenant d'une substitution par bigrammes système Playfair pourra donc comporter un nombre de lettres légèrement supérieur à celui du texte clair.

De façon générale, le pourcentage des bigrammes très rares par rapport au total des bigrammes possibles est plus élevé que celui des lettres très rares par rapport aux 26 lettres de l'alphabet. C'est ainsi qu'en français il n'y a guère que deux lettres exceptionnelles : le *k* et le *w*, soit une proportion de 2/26 ou 7,70 %. Au contraire, sur l'ensemble des 676 bigrammes qu'il est théoriquement possible de former en accouplant deux lettres de l'alphabet, il y en a près de 350 soit plus de 50 % qui ne se présenteront jamais. Par contre, les plus fortes fréquences de bigrammes n'atteignent pas en général 3 %.

La différenciation des diverses fréquences les unes par rapport aux autres sera donc beaucoup moins accusée ; et si le décrypteur ne dispose pas de textes très longs, il ne lui sera guère possible d'en tirer de conclusions utiles. Les hypothèses qu'il sera amené à faire, sur le vu du relevé de fréquences, seront beaucoup plus incertaines, et l'égareront souvent sur la mauvaise voie.

Le mot probable ne pourra être exploité avec chances de succès que s'il est possible de le situer. Il faudra donc que ce mot présente des répétitions d'un même bigramme à des intervalles pairs de lettres — et encore sous la réserve que le découpage du texte clair fasse ressortir ces répétitions. De telles conditions seront rarement remplies, et il sera difficile, en général, d'entrer dans le cryptogramme.

L'étude des répétitions relevées dans le crypto pourra orienter l'imagination du décrypteur vers telle ou telle hypothèse. De longues séquences, se terminant par un même bigramme très fréquent, pourront, en français, suggérer l'hypothèse d'un pluriel, se terminant par es. De courtes séquences, de deux ou trois bigrammes, se répétant souvent, pourront correspondre aux expressions : et la, de la, et de la.

Le décryptement des substitutions par bigrammes présentera donc des difficultés beaucoup plus considérables que la substitution simple lettre à lettre. Cependant, si le décrypteur sait à quel genre de procédé de chiffrement par bigrammes il a affaire, l'étude serrée des caractéristiques du procédé utilisé lui sera d'un grand secours. Nous en donnerons, dans ce Chapitre, un exemple, en nous plaçant dans le cas concret de la substitution par bigrammes système Playfair.

DÉCRYPTEMENT DES SUBSTITUTIONS PAR BIGRAMMES SYSTÈME PLAYFAIR

Le procédé de chiffrement est celui que nous avons indiqué plus haut. Nous en étudierons d'abord les caractéristiques, puis nous verrons quelles conséquences pratiques dégager de cette étude en vue du décryptement.

Caractéristiques du système Playfair.

Reprenons le carré de 25 :

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

En nous reportant aux règles précédemment exposées pour le chiffrement en substitution par bigrammes système Playfair, nous voyons que :

1° une lettre du clair ne peut être chiffrée que par cinq lettres du carré : les quatre autres lettres de sa ligne, et la lettre située immédiatement en dessous dans sa colonne. Elle ne sera jamais chiffrée par elle-même.

. E . . .

2° tous les bigrammes du clair qui renferment une certaine lettre H, que ce soit en première ou en deuxième lettre, seront chiffrés par des bigrammes du crypto où ne figureront que des lettres appartenant à la ligne ou à la colonne de H.

Dans l'exemple ci-dessus, l'ensemble ligne de H - colonne de H (nous l'appellerons l'équerre de H) comprend les neuf lettres :

F E R J H L T V O

Les bigrammes clairs de la forme h. ou .h seront chiffrés par des bigrammes du crypto où figureront les seules lettres F E R J H L T V O

ha = FL ho = TL he = RH etc...

3° la lettre *e* étant la plus fréquente du clair, les lettres les plus fréquentes du cryptogramme correspondront en général aux lettres qui se trouvent sur l'équerre de E.

4° DEUX BIGRAMMES INVERSES DU CLAIR SERONT TOUJOURS TRADUITS PAR DEUX BIGRAMMES INVERSES DU CRYPTO

A B L D G
F E H R J
. . . . .

ar = DF ra = FD
la = DB al = BD
br = DE rb = ED etc...

5° parmi les 24 bigrammes qu'il est possible de former en accouplant une lettre considérée avec les 24 autres lettres du carré :

huit sont chiffrés en substitution orthogonale
seize » en substitution diagonale

Dans le cas de la substitution diagonale, et dans ce cas seulement, le chiffrement est réciproque.

Le bigramme *de* du clair se chiffre en substitution diagonale par BR. Réciproquement le bigramme *br* du clair se chiffrera par DE.

Par contre, le bigramme *la* se chiffre en substitution orthogonale par DB. Le bigramme *db* du clair se chiffrera, non par LA, mais bien par GL.

6° une égalité telle que *fe* = EH implique obligatoirement la disposition FEH, soit en ligne, soit en colonne

F E H

F
E
H

7° deux égalités telles que *at* = LI et *av* = LQ indiquent que A et L sont sur une même ligne

A . . L
. . . .
I . . T
Q . . V

A L T I .
. . . . .
Q V . . .

8° deux égalités telles que *ar* = DF et *ah* = LF indiquent que A et F sont sur la même colonne

```
A . . D . L        À . . D
. . . . . .        L . . .
F . . R . H        H . . .
                   F . . R
```

9° deux égalités telles que *fr* = EJ et *fu* = JI indiquent que F et J, sont soit sur une même ligne, soit sur une même colonne, et de même pour les lettres U et I

```
F E . R J          F . . E
. . . . .          J . . R
I . . . U          . . . .
                   U . . .
                   I . . .
```

10° remarquons encore que rien n'est changé dans la correspondance bigrammes du clair bigrammes du crypto, si l'on déplace, **par permutation circulaire,** les lignes du carré, du haut vers le bas, ou les colonnes, de la gauche vers la droite. Il y a donc 25 carrés qui, à partir du carré original, peuvent servir au chiffrement, et donneront, pour un même texte clair, des cryptogrammes identiques.

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

etc....

11° remarquons enfin que si nous permutons **au hasard** les lignes ou les colonnes du carré original, les représentations des bigrammes clairs se chiffrant par substitution diagonale **ne seront pas altérées.** Par contre, les bigrammes du clair qui se chiffraient par substitution orthogonale auront de nouvelles représentations.

Dans l'un quelconque des trois carrés ci-dessus le clair « infanterie » se chiffrera par :

```
in  fa  nt  er  ie
KQ  IF  VK  HJ  KF
```

Dans le carré ci-après, où lignes et colonnes ont été permutées au hasard :

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

nous aurons pour le même clair **in fa nt er ie**
le cryptogramme..... **KQ AQ VK RJ KF**

Les bigrammes du clair *in*, *nt*, *ie*, qui se chiffrent par substitution diagonale dans les trois carrés précédents, conservent la même représentation. Mais il n'en est pas de même des bigrammes *fa*, *er*, qui provenaient d'une substitution orthogonale, et qui se traduisaient par IF et HJ.

Cette remarque nous sera très utile au cours du décryptement. Si nous sommes parvenus à reconstituer un fragment du carré, sans pouvoir encore y déterminer l'ordre relatif des lignes ou des colonnes, du moins pourrons-nous nous servir de ce fragment pour traduire, **avec certitude**, tous les bigrammes provenant d'un chiffrement **en substitution diagonale**.

Soit, par exemple, le fragment :

. . **Q** . **N** .
. . . . . .
. . **E** . **B** . **F**
. . . . . .
G . . . **A** . **R**

Nous ne pouvons rien dire des bigrammes QE, QN, NB, NA, BA, EB, EF, BF, GA, GR, AR, FR, car nous ne savons pas comment s'ordonnent les lignes et les colonnes du fragment. Mais par contre nous savons à coup sûr que :

QB = ne
BQ = en
QF = .e
BR = fa
FG = .r etc..

MÉTHODE DE DÉCRYPTEMENT

Le premier travail consiste à relever les fréquences des bigrammes du crypto, ce qui permet d'en déduire ensuite rapidement les fréquences individuelles de lettres.

On classera alors les lettres par ordre de fréquences décroissantes. Les plus fréquentes d'entre elles correspondront, dans la majorité des cas, aux lettres appartenant à l'équerre de E (voir remarque 3).

On classera de même les bigrammes les plus fréquents, en les séparant en deux catégories :

— bigrammes fréquents dont les inverses figurent dans le cryptogramme,

— bigrammes fréquents dont les inverses ne figurent pas dans le cryptogramme.

Dans la première catégorie, nous mettrons à part les bigrammes très fréquents constitués par des lettres très fréquentes Il y a de fortes chances pour qu'ils correspondent à des bigrammes en *e* du clair (forme *e.* ou *.e*) Les autres bigrammes représenteront des bigrammes clairs ne renfermant pas la lettre *e*, mais dont l'inverse est cependant fréquent.

Les bigrammes de la deuxième catégorie correspondront à des bigrammes du clair dont l'inverse est rare ou peu fréquent.

Nous avons d'autre part dressé, une fois pour toutes, un tableau des bigrammes de la langue française, en les séparant par catégories comme il est indiqué ci-dessus ; tableau auquel nous nous reporterons au cours de la recherche en vue de préciser un bigramme non encore identifié.

Par ailleurs, nous soulignerons dans le cryptogramme toutes les **suites de bigrammes caractéristiques** qui s'y trouvent, telles que : bigrammes redoublés (VK VK), bigrammes inverses successifs (TG GT), bigrammes se répétant à un rang d'intervalle (MK JP MK), bigrammes inverses espacés d'un rang (TF BJ FT), etc...

Lorsqu'un cryptologue a souvent affaire à des substitutions par bigrammes système Playfair, il aura le plus grand intérêt à se constituer une liste de mots ou expressions présentant des suites de bigrammes caractéristiques. La comparaison des suites de cette liste avec celles que présente le crypto lui donnera des indications précieuses sur les divers mots à « essayer », et pourra lui éviter des efforts d'intuition parfois pénibles. Nous donnons ci-après une liste — non limitative — des suites de bigrammes caractéristiques les plus courantes.

Nous soulignerons également les répétitions longues, portant sur une séquence de plusieurs bigrammes. Elles correspondent à des répétitions du clair et peuvent, soit par leur longueur, soit par leur répartition dans le crypto, attirer l'attention du décrypteur sur tel ou tel mot probable.

L'étude analytique du cryptogramme étant terminée, la marche à suivre sera la suivante :

On essaiera d'abord d'identifier le bigramme *es*. Ce sera en général — surtout si le texte est suffisamment long — le plus fréquent parmi les bigrammes fréquents dont les inverses figurent dans le crypto. On s'efforcera de même, en s'aidant des fréquences de bigrammes et des suites de bigrammes caractéristiques, d'identifier les autres bigrammes en *e* : *le*, *en*, *de*, *re*, *te*. Certains auteurs se basent dans cette recherche sur la différence que présentent la fréquence d'un bigramme considéré, et celle de son inverse (les bigrammes *es*. *en*, sont à peu près deux fois plus fréquents que leurs inverses ; le bigramme *te* au contraire possède à peu près la même fréquence que son inverse *et*). Si le texte étudié est très long, ce procédé a sa valeur ; mais dans beaucoup de cas, ce sera surtout l'étude des suites de bigrammes caractéristiques qui permettra, grâce aux qualités d'intuition du décrypteur, d'identifier tel ou tel groupe.

On reportera immédiatement le long du cryptogramme chacune des hypothèses envisagées, et on la matérialisera sur le papier par un essai de reconstitution du carré de 25, qui traduira — selon les règles de chiffrement du système Playfair précédemment exposées — les correspondances clair-crypto admises provisoirement comme sûres. On aura soin, dans ce travail, « d'aérer » largement le fragment du carré reconstitué, et de ne pas juxtaposer deux colonnes ou superposer deux lignes avant d'avoir la certitude qu'il doit en être ainsi. Dès que quelques lettres auront été mises en place dans le carré, on pourra traduire, en tout ou en partie, de nouveaux bigrammes du crypto. On exploitera de suite ces nouveaux résultats, et l'on poursuivra de proche en proche, jusqu'à voir apparaître un mot ou un fragment du texte clair.

La partie la plus délicate de ce travail consiste à « entrer » dans le cryptogramme pour aboutir à une correspondance clair-crypto portant sur plusieurs bigrammes successifs. Il ne faudra pas craindre, dans les hypothèses de départ, une certaine hardiesse, quitte à courir le risque de s'égarer dans des erreurs grossières, et de voir se réduire à néant une recherche longue et pénible. Le problème est de ceux qui présentent pour le cryptologue le plus de difficultés ; il nécessite à la fois une analyse très serrée du texte étudié, un gros effort d'intuition, et une connaissance parfaite de la technique du procédé de chiffrement utilisé.

Nous en donnons ci-dessous un exemple.

BIGRAMMES FRÉQUENTS DONT LES INVERSES SONT EUX-MÊMES FRÉQUENTS

Bigrammes en E :				Autres bigrammes :			
es	305	se	155	it	112	ti	98
le	246	el	141	is	103	si	64
en	242	ne	124	la	101	al	54
de	215	ed	96	ra	92	ar	86
re	209	er	163	tr	86	rt	41
te	163	et	143	ta	85	at	56
em	113	me	104	us	76	su	39
ec	100	ce	98	sa	75	as	52
eu	89	ue	85				
ep	82	pe	49				

BIGRAMMES FRÉQUENTS DONT LES INVERSES SONT RARES OU PEU FRÉQUENTS

Bigrammes en E :				Autres bigrammes :			
ie	94	ei	16	nt	197	tn	10
				on	164	no	37
				qu	134	uq	3
				an	131	na	30
				ou	118	uo	1
				ai	117	ia	14
				in	90	ni	22
				ur	88	ru	13
				co	87	oc	7
				nd	80	dn	0
				ns	79	sn	13
				pa	78	ap	26

SUITES DE BIGRAMMES CARACTÉRISTIQUES

Forme XY XY		Même, tête, préférer, point de départ, point de débarquement, poste téléphonique, etc...
Forme XY YX	ESSE	adresse, blessé, détresse, essence, presse, progresser
	SEES	croisées, fusées, refusées
	ENNE	antenne, ennemi, moyenne
	EMME	précédemment, récemment
	ERRE	guerre, terre, voie ferrée
	ETTE	cette, nette, permette
	ELLE	actuelle, celle, Excellence, elle, nouvelle, quelle, sentinelle
	EFFE	effet, effectif
	ARRA	barrage, arracher
	ANNA	dépannage
	ASSA	assaut, ambassade, assassin, massacre, passable, passage
	ATTA	attaque, attacher, rattacher
	APPA	appareil, apparaître, appartenir
	ISSI	émission, mission, omission, permission, transmission
	ILLI	millième, millier, millimètre
	IFFI	difficile, difficulté
	OPPO	opportun, opposer
	OMMO	commotion
Forme XY . . XY		qu el qu e, re ch er ch er, re nd re, re nt re r, p re pa re r
Forme XY . . YX		d is po si ti on, re gl er
Forme XY . . . . XY		re cu pe re r, u ti li sa ti on
Forme XY . . . . YX		im me di at em en t, mo me nt an em en t, r en se ig ne r, pa ra gr ap he, re si st er, si gn al is at io n
Forme XY . . . . . XY		p er te sl eg er es, ph ot og ra ph ie
Forme XY . . . . . YX		em ba rq ue me nt

EXEMPLE DE DÉCRYPTEMENT

Soit le cryptogramme ci-dessous, chiffré en substitution simple par bigrammes système Playfair. Nous savons qu'il s'agit d'un texte militaire, recueilli au cours d'une opération offensive.

AY	XS	EZ	MJ	JM	EL	ZL	MJ	UI	BX	KT	YD	YZ
UZ	MS	EX	FR	IK	BU	KI	BU	DL	RI	FU	TU	BF
FB	XT	EZ	UZ	UI	OD	YC	UB	KN	YJ	TM	ZH	YC
FZ	FZ	DV	IK	DQ	IK	IL	BF	KU	UI	FS	BU	BF
FB	XT	EZ	VO	TM	DV	EK	PC	ZH	FU	XS	ZY	CI
VZ	FZ	LI	XS	UZ	FZ	VC	XN	IU	RE	JN	BL	YD
KI	ZI	MJ	UI	ER	RU	IU	ZY	CI	VZ	FZ	LI	XS
DG	TB	MT	YD	BF	FB	XT	RU	UZ	ZV	MR	JM	EM
FU	KI	QJ	FR	JN	BL	YD	KI	ZL	KU	JM	EL	RU
BF	MC	IK	QJ	MY	BU	RK	BF	DO	BU	IU	QL	UB
BU	UZ	FZ	VC	XN	RF	ZG	UZ	MS	EM	BU	AH	IK
RZ	BK	TK	YA	UR	TU							

L'étude analytique du cryptogramme nous a donné les résultats suivants :

Les lettres les plus fréquentes sont :

U	Z	B	F	I	K	M	Y	T	X	D	E	R	J
31	25	22	22	19	15	15	12	11	11	10	10	9	9

Les bigrammes les plus fréquents, rangés par catégorie, sont :

Bigrammes fréquents s'inversant dans le crypto :

BU 7 UB 2 IU 3 UI 3
BF 6 FB 3
IK 5 KI 3

Bigrammes fréquents ne s'inversant pas dans le crypto :

FZ 6 UZ 6 XS 4 EZ 3 FU 3 XT 3
YD 3

Nous avons d'autre part relevé dans le crypto les suites de bigrammes caractéristiques :

MJ JM
IK BU KI BU
BF FB XT , qui se répète trois fois, et se trouve deux

fois suivie du bigramme EZ

FZ FZ
IK DQ IK
UB BU

ainsi que les répétitions :

MJ UI
ZY CI VZ FZ LI XS
UZ FZ VC XN
JM EL

Les bigrammes BU et BF sont vraisemblablement des bigrammes en *e*. Leurs inverses figurent en nombre dans le crypto et nous y trouvons également les suites caractéristiques BF FB (3 fois) et UB BU (1 fois).

Les bigrammes fréquents FZ et UZ dont les inverses ne figurent pas dans le cryptogramme peuvent être des bigrammes de deux consonnes fréquentes. Cependant la suite FZ FZ nous conduirait à avoir dans le texte clair quatre consonnes successives. Le bigramme FZ doit donc contenir une voyelle et une consonne.

Les bigrammes IU et JM s'inversent dans le cryptogramme, et nous y trouvons la suite MJ JM. Les lettres I et U étant parmi les plus fréquentes, il est possible que IU représente un bigramme en *e*. Pour MJ, la lettre J se trouve au 14e rang dans l'ordre des fréquences décroissantes de lettre. Il s'agit donc vraisemblablement d'un bigramme voyelle consonne, la voyelle étant différente de *e*.

Le bigramme BU est le plus fréquent (7). Son inverse UB apparaît deux fois dans le crypto. Nous ferons l'hypothèse que BU = *es*, ce qui nous donne dans le carré de 25 l'une des dispositions suivantes :

(1) E B . S U (2) E / B / S / U (3) E . . B / . . . . / U . . S

Les trois répétitions BF FB XT, où figurent les bigrammes inverses fréquents BF et FB nous suggèrent le mot « en ne mi ». Si cette nouvelle hypothèse est exacte, nous aurions BF = *en*.

Dans ces conditions, la disposition (2) ne pourrait convenir, et nous sommes conduits à retenir :

(4)
E B . S U
. . . . .
F N

ou (5)
E . . B
. . . .
U . . S
. . . .
F . . N

Examinons maintenant la suite MJ JM. Le bigramme MJ, nous l'avons dit plus haut, est vraisemblablement constitué par une voyelle différente de e suivie d'une consonne. Essayons le mot « at ta qu e » qui nous donne les correspondances :

MJ	JM	EL	ZL
at	ta	qu	e .

L'égalité *qu* = EL n'est pas compatible avec la disposition (4), car E et U s'y trouvant sur une même ligne, il ne pourrait s'agir que d'une substitution orthogonale dans la ligne. Nous aurions donc QEB. SUL, soit six lettres sur la même ligne, ce qui est impossible.

Par contre, l'égalité *qu* = EL peut avec la disposition (5) donner naissance aux dispositions ci-dessous :

(6)
Q
E . . B
F . . N
U . . S
L

ou (7)
Q
E . . B
U . . S
L
F . . N

Nous pouvons, à l'aide de ces fragments de carré, essayer de traduire la suite QL UB BU qui figure aux deux avant-dernières lignes du crypto.

La disposition (6) donnerait : QL UB BU
lu se es

» (7) » : QL UB BU
fu se es

Nous voyons apparaître le mot « fusées » qui nous incite à adopter la disposition (7) et à retenir les correspondances :

BU = **es** **UB** = **se**
BF = **en** **FB** = **ne**
EL = **qu**
FU = **le**
FS = **nu**
BL = **e.**
QL = **fu**

Mais nous ne pouvons encore intégrer dans la disposition (7) les égalités

MJ = **at**
JM = **ta**
XT = **mi**

Reportons ces traductions dans le cryptogramme, et essayons de déterminer le bigramme EM de la suite :

ZV	**MR**	**JM**	**EM**	**FU**	**KI**	**QJ**
		ta		**le**		

en plaçant EM dans la disposition (7).

La lettre M se trouve quelque part sur l'une des cinq lignes du carré :

1	**Q**	.	**(M)**	.	.
2	**E**	.	**(M)**	.	**B**
3	**U**	.	**(M)**	.	**S**
4	**L**	.	**(M)**	.	.
5	**F**	.	**(M)**	.	**N**

Si **M** est sur la ligne 1, nous aurons **EM** = . **q**
» 2, nous ne pouvons rien dire quant à la traduction de **EM**, l'ordre relatif des colonnes du carré restant incertain
» 3, nous aurons **EM** = . **u**
» 4, » **EM** = . **l**
» 5, » **EM** = . **f**

La traduction possible . I nous fait songer à

JM	EM	FU
ta	il	le

(bataille, ravitaillement, etc...)

Mais l'hypothèse « ravitaillement » ne résiste pas à l'examen car nous aurions alors la correspondance :

ZV	MR	JM	EM	FU	KI	QJ
ra	vi	ta	il	le	me	nt

et les correspondances *il* = EM et *me* = KI sont incompatibles entre elles.

Bornons-nous donc à retenir la correspondance *il* = EM ; notre disposition (7) devient alors :

(8)

Q	.	.	.	.	.
E	.	.	I	.	B
U	.	.	.	.	S
L	.	.	M	.	.
F	.	.	.	.	N

Nous pouvons maintenant y intégrer les correspondances XT = *mi*, MJ = *at*, JM = *ta*, ce qui nous donne :

(9)

Q	.	.	.	.	.	
E	.	.	I	.	B	
U	.	.	T	.	S	. J
L	.	.	M	.	.	. A
F	.	.	X			

L'ordre relatif des colonnes reste à déterminer, mais nous pouvons dès à présent traduire de façon certaine tous les bigrammes provenant d'une substitution diagonale ; nous avons ainsi :

IU = et UI = te
QJ = .u
TB = si
XS = .t
BX = i.
MS = .t
EX = if , etc...

de même, nous relevons dans la colonne I T M X les correspondances certaines : MT = *ti*, TM = *it*.

En reportant ces résultats le long du cryptogramme, il vient :

```
AY XS EZ MJ JM EL ZL MJ UI BX KT YD YZ UZ MS EX FR IK
   .t    at ta qu    at te i.          .t if
BU KI BU DL RI FU TU BF FB XT EZ UZ UI OD YC UB KN YJ
es    es    le    en ne mi    te
TM ZH YC FZ FZ DV IK DQ IK IL BF KU UI FS BU BF FB XT
it                        em en    te .u es en ne mi
EZ VO TM DV EK PC ZH FU XS ZY CI VZ FZ LI XS UZ FZ VC
      it          le .t          me .t
XN IU RE JN BL YD KI ZI MJ UI ER RU IU ZY CI VZ FZ LI
   et    e.          at te    et          me
XS DG TB MT YD BF FB XT RU UZ ZV MR JM EM FU KI QJ FR JN
.t    si ti    en ne mi       ta il le
BL YD KI ZL KU JM EL RU BF MC IK QJ MY BU RK BF DO BU IU
e.          ta qu    en      .u    es          es et
QL UB BU UZ FZ VC XN RF ZG UZ MS EM BU AH IK RZ BK TK YA
fu se es              .t il es
UR TU
```

Nous voyons apparaître plusieurs mots ou fragments du clair :

```
MJ UI BX    YZ UZ MS EX    DG TB MT YD
at te in .. ob je ct if    po si ti on
```

et en fin de crypto :

```
RF ZG UZ MS EM BU
?p ro je ct il es
```

La disposition (9) se complète alors comme suit :

```
Q . . . . . .
E . I . B . Z
U . T . S . J
L . M . C . A
F . X . N . .
```

La répétition ZY CI VZ FZ LI XS qui se traduit par

```
.. mb .. .e me nt
```

nous donne le mot « bombardement », d'où les correspondances

VZ = ar , FZ = de , ZY = bo .

Q	.	.	.	Y	.	O	.	
E	.	I	.	B	.	Z	.	R
U	.	T	.	S	.	J	.	
L	.	M	.	C	.	A	.	V
F	.	X	.	N	.	D		

Le début du cryptogramme se traduit alors par :

co nt EZ at ta qu ea at te in KT on ob je ct if.....
re ts

Nous y lisons EZ = *re*, ce qui nous permet de juxtaposer définitivement trois des colonnes du carré. De même, l'égalité KT = *ts* nous permet de placer les deux autres colonnes : la lettre K venant obligatoirement se placer au seul endroit restant disponible de la ligne U T S J, dans la même colonne que R et V, entre ces deux lettres.

Nous avons ainsi le carré :

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

Le décryptement s'achève sans difficulté. Le texte clair est :

« Contre attaque a atteint son objectif Pertes très faibles K Ennemi rejeté dans ses positions de départ fortement K tenues Ennemi réagit par un violent bombardement. Je demande tirs de contre batterie K et bombardement position ennemie K. Je ravitaille troupes de contre atKtaque K en cartouches grenades et fusées Je demande projectiles mortiers Stockes K »

et le carré de 25 est maintenant reconstitué dans sa totalité :

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

Recherche du mot clef.

Tel qu'il est reconstitué ci-dessus, le carré de 25 a de très grandes chances de ne pas être le même que celui du chiffreur. Nous avons vu en effet qu'il existait 25 carrés différents, se déduisant les uns des autres par permutation circulaire des lignes ou des colonnes, et donnant, pour un même clair, le même cryptogramme.

Toutefois, si le carré de chiffrement a été construit sur un mot clef, il sera possible de retrouver le carré de 25 original, tel qu'il a été établi en vue du chiffrement.

Nous partirons des lettres rares V X Y Z. et nous chercherons comme il a été dit au Chapitre III page 49 à reconstituer le tableau de construction.

Proposons-nous de rechercher quelles sont les lettres qui, dans tous les cas possibles, se trouvent au-dessus de Z dans le tableau dont s'est servi le chiffreur pour établir le carré. Toutes les fois que Z ne sera pas à l'extrême gauche du carré (E Z B I R, R E Z B I, I R E Z B, B I R E Z) soit dans 20 cas sur 25, ce sera la lettre E, située immédiatement à sa gauche. Par contre, lorsque Z se trouvera à l'extrême gauche du carré (Z B I R E) soit dans 5 cas sur 25, nous aurons une disposition de la forme ci-dessous.

. . . . .
O Y H G Q
Z B I R E
. . . . .

et ce sera la lettre Q qui se trouvera au-dessus de Z dans le tableau de construction.

On voit qu'il n'existe, pour chaque lettre étudiée, que deux hypothèses : ou la lettre située immédiatement à gauche, ou la lettre qui se trouve immédiatement au-dessus de cette dernière :

. Z

Reprenons le carré de 25 reconstitué en fin de décryptement. Au-dessus de V X Y Z, nous pouvons avoir :

T C D Q
M N O E
V X Y Z

C'est la suite M N O Q qui s'ordonne le mieux. Les lettres M N O sont situées immédiatement à gauche de V X Y ; par contre, pour que Q se trouve au-dessus de Z dans le tableau de construction, il est nécessaire que Q se trouve à l'extrême droite du carré et Z à l'extrême gauche, ce qui nous oblige à permuter les colonnes de notre carré pour aboutir à la disposition :

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

où les lettres V X Y Z sont espacées de quatre rangs.

Le tableau de construction comporte donc des colonnes de 4 et des colonnes de 3 lettres. En « remontant » les lettres V X Y Z, on a :

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

où apparaît le mot clef « ALPHABET » ; le carré original tel qu'il a été utilisé par le chiffreur étant :

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

CAS DU MOT PROBABLE A SUITE DE BIGRAMMES CARACTÉRISTIQUES

Un cas beaucoup plus favorable — mais qui se présentera rarement dans la pratique — est celui où l'on dispose d'un mot probable offrant de telles particularités qu'il soit facile de le situer immédiatement dans le cryptogramme. On dispose alors de plusieurs correspondances bigrammes clairs-bigrammes crypto qui permettent de reconstituer tout de suite un important fragment du carré de chiffrement.

Le télégramme ci-dessous émane d'un poste de Corps d'Armée. Nous soupçonnons qu'il est adressé à la 53e Division, et nous chercherons à y situer le mot probable « cinquante troisième division ».

Ce mot probable peut avoir été découpé de deux manières :

ci nq ua nt et ro is ie me di vi si on
c in qu an te tr oi si em ed iv is io n

Dans le premier cas, les deux bigrammes inverses *is*, *si* donneront naissance dans le crypto à deux bigrammes inverses espacés de 5 rangs. Dans le second cas, aux deux bigrammes *oi*, *io* du clair correspondront dans le crypto deux bigrammes inverses espacés de 6 rangs, et aux deux bigrammes *si*, *is* du clair deux bigrammes inverses espacés de 4 rangs.

ZR DR ZR DF QY ML CZ AT RD EV ZL UA QG CI QY GV HL
LB ZP DF BZ UZ BL YL DR RJ GH GE YL VQ FD LA LQ AJ
DR GU FD SP RD HP ZL ZA TQ GV GE TQ ZP HS ZP DK GZ
EV JB ZO

En parcourant le cryptogramme, nous remarquons qu'il présente deux bigrammes inverses LB et BL espacés de 5 rangs, ce qui implique la correspondance :

ci nq ua nt et ro is ie me di vi si on
UA QG CI QY GV HL LB ZP DF BZ UZ BL YL

Partant de cette correspondance, et laissant de côté le cryptogramme, nous allons, en nous basant uniquement sur le raisonnement analytique, reconstituer dans sa plus grande partie le carré de chiffrement.

L'égalité *nq* = QG implique l'une des deux dispositions

(1) N Q G

(2)
N
Q
G

Mais l'égalité *nt* = QY, pour s'adjoindre à la précédente, entraîne l'une des dispositions :

(3) N Q G T Y

(4)
N
Q
G
T
Y

(5)
N Q G
. . .
Y T

Comment intégrer dans ces trois dispositions l'égalité *et* = GV. Aucune solution n'est possible dans (3) et (5). Par contre nous pouvons avoir dans (4) :

(4)

N	.	
Q	.	
G	.	E
T	.	V
Y	.	

Plaçons encore l'égalité *on* = YL. Nous pouvons avoir :

(5)

N	.	.	L
Q	.	.	.
G	.	.	E
T	.	.	V
Y	.	.	O

ou bien (6)

N	.	.	L
Q	.	.	.
G	.	E	.
T	.	V	.
Y	.	.	O

sans que l'ordre relatif des colonnes puisse encore être fixé.

Mais O et L se trouvant dans les deux cas sur une même colonne, l'égalité *ro* = HL ne peut provenir que d'une substitution orthogonale en colonne. C'est-à-dire que R et H se trouvent certainement sur la même colonne que L et O. Ceci n'est possible qu'avec la disposition (6) et nous avons :

(7)

N	.	.	L
Q	.	.	R
G	.	E	H
T	.	V	.
Y	.	.	O

ou bien (8)

N	.	.	L
Q	.	.	.
G	.	E	R
T	.	V	H
Y	.	.	O

Essayons de grouper ensemble d'autres lettres du cryptogramme. La réciprocité des égalités *ci* = UA, *ua* = CI nous indique que ces deux bigrammes ont été chiffrés par substitution diagonale :

(9)

C	.	U
.	.	.
A	.	I

Exploitons d'autre part les deux égalités $di = BZ$, et $vi = UZ$. La disposition (9) devient :

(10)
. . B . D
. . . . .
C . U . V
. . . . .
A . I . Z ;

ou bien (11)
C . . . U V
. . . . . .
A . D B I Z

Mais l'égalité $si = BL$ n'est pas compatible avec la disposition (11). Par contre, nous pouvons avoir avec la disposition (10) les deux solutions :

(12)
S . . . B . D
. . . . . .
. . C . U . V
. . . . . .
L . A . I . Z

(13)
. . S . .
. . B . D
. . . . .
C . U . V
. . . . .
A . I . Z
. . L . .

En rapprochant les dispositions (7) et (8) des dispositions (12) et (13) nous constatons que la disposition (13) est à rejeter. En effet, la colonne de L qui contient déjà certainement les lettres R H O, ne peut encore contenir les lettres S B U I.

La disposition (12) est donc la seule qui soit compatible avec l'une des deux dispositions (7) et (8), et dans ces deux cas possibles :

les lettres T C U V sont certainement sur une même ligne
— E D V Z — sur une même colonne
— L R H O S — sur une même colonne
— N L A I Z — sur une même ligne

les lettres S B D ne peuvent être que sur la ligne de Q ou sur la ligne de Y qui seules comportent encore trois places disponibles.

La lettre S doit donc se trouver à la fois sur la colonne de L et sur l'une des lignes Q ou Y. Or, elle ne peut se trouver sur la ligne de Y puisque, dans les deux dispositions seules possibles (7) et (8) c'est la lettre O qui figure certainement à l'intersection de la ligne de Y et de la colonne de L.

La lettre S figure donc obligatoirement à l'intersection de la colonne de L et de la ligne de Q ; ce qui nous entraîne à rejeter définitivement la disposition (7) et à adopter le fragment de tableau carré :

N L
Q . . . . S
G . E . . R
T . V . . H
Y . . . . O

Nous pouvons alors placer à coup sûr la lettre D qui appartient à la colonne de E et à la ligne de Q ; ainsi que la lettre Z qui appartient à la colonne de E et à la ligne de N.

N . Z . . L
Q . D . . S
G . E . . R
T . V . . H
Y . . . . O

En intégrant dans ce nouveau fragment la disposition (12) nous avons

N . Z . A . I . L
Q . D . . . B . S
G . E . . . . . R
T . V . C . U . H
Y . . . . . O

où nous pouvons placer les égalités : *ie* = Z P, *me* = DF.

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

L'ordre relatif des colonnes reste encore incertain, mais nous pouvons néanmoins à l'aide de ce fragment traduire tous les bigrammes du crypto qui proviennent d'une substitution **diagonale.**

On a alors :

ZR	DR	ZR	DF	QY	ML	CZ	AT	RD	EV	ZL	UA	QG	CI	QY	GV	HL
le	se	le	me	nt	sa	va	nc	es	de		ci	nq	ua	nt	et	ro
LB	ZP	DF	BZ	UZ	BL	YL	DR	RJ	GH	GE	YL	VQ	FD	LA	QL	AJ
is	ie	me	di	vi	si	on	se		rt		on	td	em		ns	
DR	GU	FD	SP	RD	HP	ZL	ZA	TQ	GV	GE	TQ	ZP	HS	ZP	DK	GZ
se	pt	em	br	es	ur			gn	et		gn	ie	rl	ie		en
EV	JB	ZO														
de		l.														

Nous lisons facilement ZL = *la*. LA = *ai*. Ceci nous permet d'ordonner définitivement les colonnes du carré :

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

Le décryptement se complète facilement :

« Les éléments avancés de la cinquante troisième division se porteront demain six septembre sur la ligne Tergnier Liez Vendeuil »

et le carré de 25 est maintenant reconstitué dans son entier :

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

Recherche du mot clef.

Partons des lettres rares V X Y Z. Au-dessus de ces quatre lettres, nous pouvons avoir, dans le tableau de construction :

R	U	V	O
H	J	K	L
V	X	Y	Z

Nous avons certainement dans le tableau de construction la disposition :

H J K L
V X Y Z

Mais ceci n'est possible que si la colonne contenant K se trouve à la gauche de celle qui contient Y. Nous sommes donc amenés à étudier la forme :

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

où les lettres H J K L sont immédiatement à gauche des lettres V X Y Z.

En remontant le tableau à partir des lettres V X Y Z, nous avons :

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

qui se redresse aisément comme ci-dessous :

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

et nous donne le mot clef « INSPECTION ».
