

TUTORAT 2 : ÉLÉMENTS DE CRYPTOGRAPHIE

Frédéric Chevy – chevy@lkb.ens.fr

Lire l'article ci-joint et décrypter le message suivant :

JNPFU Yy 1937, RYhKcs iMZOZMW xPOuZjNY zPFI iw aNeLeys NF wUzekZGyFYNu OeFwZus P usSPZL jlZ xGyIkehFs wBYI zslC IdeTTKsL, ZS ys LNiaekY wlaly sFkdhNiZ-wits. jlYSjNYL PyysYi HvlL kwKz, vsL UsLasyUPyML UY iw DwaEZFY, SYL GKUeFwMsNKi, GyM YFBPEZ vs DGFUY.

Ssi nEXiZIZYFi jlPFkZJNsL BesFysyM Us TPOuejlYK v'sJNeoPvsyM jlwFkZJNs UY S'P-zUekZhFFYIK zs iMZqZkg YM Sw MdsGKzs Ul aPSaNS JIPyMejNs HuGloY Jls, Le S'Gy nuGcKsiis yGutPvYDYyM, SsL GuUZFPMyluL JlwYkZjlsL LsKhyM HSIL HNZLLPFkl Jls ys HhlKKPek v'sMKs NF huzZFwkYNK ISwiIZJNY ThyIMZGyFPyM w vw oekYLis UY SP SNtesus PosI zYi NyZMYL zY DYDheus UY Sw MPZSSs zY v'PMhDs : yGIL oeBGFL vP FPZLiPyas U'lFs MYIdyGShcZs ThFzYs iNK vYi DeuwaSYL kdYGKZJNsI Ul DhFzY jNwF-MeJNY, sk jNe HuhrYkkY aslm-Ie zPyi FGkKs uYPveks JlhMeUeYFys.

rYPF HwNS zsvPdPXY vsi SGeL yGloYSSsL Us v'eFThKtPkehF JNwyMeJlY.

nGIK vw iaeYFIs y 250



Viète, inventeur de la cryptanalyse mathématique

Il y a 400 ans, les Espagnols croyaient leurs codes secrets inviolables et Viète en perçait le mystère ; aujourd'hui des cryptologues prétendent que leurs codes secrets sont inviolables.

Le 13 décembre 1603, François Viète mourrait à Paris. Quelques jours avant, sentant sa fin prochaine, ce juriste et mathématicien, considéré comme l'inventeur de l'algèbre moderne, avait décrit les méthodes qu'il utilisait pour déchiffrer les dépêches secrètes interceptées par les soldats du roi Henri IV. Ce mémoire de 1603 destiné à Sully, premier ministre du roi, est aujourd'hui perdu et l'on ne dispose que d'une transcription réalisée au XIX^e siècle par Frédéric Ritter, un historien amateur, ingénieur des Ponts et Chaussées. Celle-ci, elle-même oubliée, a été redécouverte récemment dans la Bibliothèque de l'Institut de France par l'historien de la cryptographie Peter Pesic qui en a fait une analyse approfondie.

Viète a utilisé des méthodes systématiques et formelles pour reconstituer le sens des messages chiffrés. Sa technique va au-delà de l'étude des fréquences des lettres et des recherches à tâtons mises en œuvre par ses contemporains. Viète est l'inventeur de la cryptanalyse mathématique, science qui, du fait du développement des télécommunications, concerne aujourd'hui tout le monde. À l'occasion du quatrième centenaire de sa mort, il est bon qu'un hommage soit rendu à ce génie qui croyait, grâce à sa *Règle infaillible*, percer n'importe quel message secret.

François Viète est né en 1540 à Fontenay-le-Comte à une cinquantaine de kilomètres de La Rochelle. Il étudie le droit à Poitiers avant de revenir à Fontenay-le-Comte, en 1560, occuper un siège d'avocat du roi. Il est promu avocat au Parlement de Paris en 1571, puis conseiller au Parlement de Bretagne en 1574, puis maître des requêtes ordinaire de l'Hôtel du roi en 1580. Enfin, en 1594, il devient conseiller du roi Henri IV et sera chargé jusqu'à sa mort de tout ce qui concerne l'art et la pratique des messages secrets, ce que l'on dénomme aujourd'hui la cryptologie. Son exploit public le plus remarquable est le déchiffrement de la lettre du commandeur Moreo au roi d'Espagne. Ce travail, qui dénonçait les agissements de l'Espagne, fut utile à Henri IV dont le trône était contesté par la Ligue catholique, alliée de l'Espagne.

Parallèlement à sa brillante carrière politique Viète fait œuvre de mathématicien et publie – à ses frais – un certain nombre

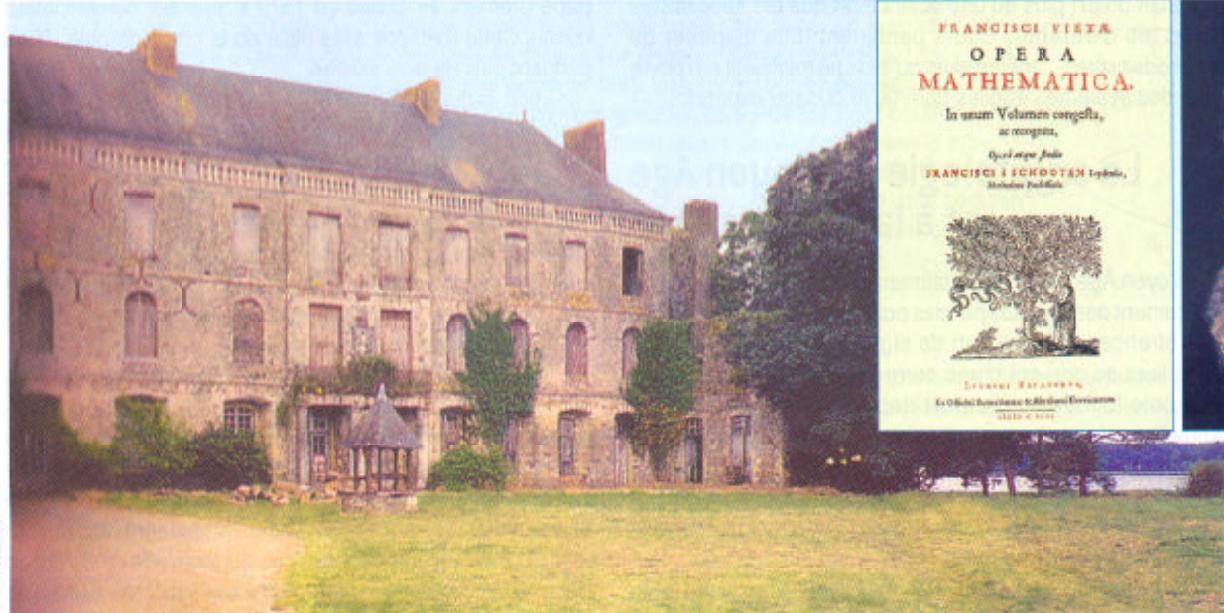
d'opuscules destinés à ses amis. En particulier, en 1591 dans *De l'analyse mathématique restaurée ou l'Algèbre nouvelle*, il introduit une innovation technique majeure : l'utilisation de lettres pour désigner des valeurs connues, lettres auparavant utilisées uniquement pour les variables. Cette nouveauté permet de formuler des solutions générales et de mettre en évidence des analogies, par exemple entre tous les problèmes où l'on doit trouver deux nombres connaissant leur somme et leur produit. Ce qu'il nomme *logistique spécieuse* est un calcul portant sur les lettres dont il énonce les règles de manipulation et dont le symbolisme sera amélioré au siècle suivant (écriture des puissances, symbole pour l'égalité) pour donner la technique de calcul qu'on apprend en algèbre élémentaire et dont on a du mal à imaginer comment on pouvait se passer.

En 1595, Viète résout brillamment un problème posé aux « mathématiciens du monde entier » par le Belge Adrien Romain (voir l'encadré page 93) et propose une élégante solution à un problème d'Apollonius : trois cercles étant donnés, construire un quatrième cercle qui leur soit tangent. Sa tentative de réformer le calendrier grégorien en 1600 reçoit un accueil défavorable : elle est condamnée par Rome. Son dernier manuscrit, dont l'importance n'a été comprise que récemment, en fait un personnage clef de l'histoire de la cryptologie.

Mathématicien du roi

La passion de Viète pour les mathématiques était telle qu'on raconte qu'il lui arrivait de rester plusieurs jours de suite sans manger ni dormir, absorbé par un problème. Le mémorialiste Gédéon Tallemant des Réaux dit de lui que jamais un homme ne fut « mieux né pour les mathématiques » et que s'il ne vécut pas plus âgé c'est qu'il se tua par son excès de travail. Pourtant les mathématiques ne sont chez Viète qu'un loisir, ainsi qu'il nous le dit dans l'opuscule où il résout le problème d'Adrien Romain : « Moi qui ne fais pas profession de mathématicien, mais que l'étude des mathématiques charme quand j'ai du temps libre. »

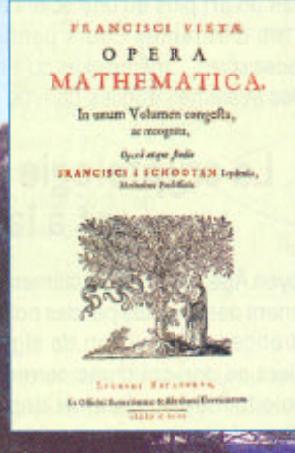
L'importance de ses exploits de briseur de code a été sous-évaluée et on a mal compris la parution de son texte pamphlet « Deschiffrement d'une lettre écrite par le commandeur



Jean-Paul Delahaye

1. Viète séjourna au château du Parc Soubise, en Vendée, notamment entre 1564 et 1570 (« Quant à moi, Poitevin de Fontenay qui habite souvent sur les bords de la Vendée, un château fort, construit jadis par la fée Mélusine »). À droite, portrait de Viète et page de titre de l'ouvrage

Opera Mundi, publié en 1646 à Leyde et où est regroupée une grande partie des textes mathématiques de Viète. En fond, les positions des olives sur les branches d'olivier correspondent aux lettres d'un message secret, procédé de stéganographie utilisé au temps de Viète.



Moreo au Roy d'Espagne son maître, du 28 octobre 1589 – où se voit que le duc de Mayne [Mayenne] s'est déclaré vouloir être Roy. » Pour saisir l'importance de l'épisode, il faut revenir aux premières années du règne de Henri IV. Du fait de sa confession protestante, le roi luttait contre la Ligue (catholique) soutenue par le roi d'Espagne Philippe II. En 1589, la Ligue tenait plusieurs grandes villes de France et Henri IV était en position précaire. Plusieurs lettres de Juan de Moreo, officier de la Ligue, écrites au roi d'Espagne et à son ambassadeur en France Don Bernardino de Mendoza, tombèrent entre les mains des soldats du roi. Viète fut chargé de les décrypter. Le travail lui prit quelque temps et n'aboutit qu'après la victoire remportée par Henri IV à la bataille d'Ivry contre la Ligue, victoire qui donna au roi un avantage décisif. Le succès de Viète ne fut pas inutile, car le contenu des lettres révélait que le chef de la Ligue en France, le Duc de Mayenne, projetait de devenir roi à la place de Henri IV. Il est probable que c'est avec l'accord d'Henri IV, voire à sa demande, que Viète publia en 1590 le contenu de la lettre de Moreo.

Viète, dans son introduction du texte de 1603, suggère aussi que d'autres lettres montrant que le roi d'Espagne n'avait pas que des intentions loyales, furent transmises au duc de Mayenne et contribuèrent à la mise au point du compromis permettant à Henry IV de conserver son trône. Cette interprétation de l'historien Pesic expliquerait aussi pourquoi Viète publia une lettre en 1590 qui annonçait que

les codes secrets du roi d'Espagne et de ses alliés étaient connus du roi.

Voici quelques lignes de cette introduction : « Je n'ai point caché la voie que j'y ai tenue [pour le déchiffrement des lettres], mais j'en ai toujours ouvert la lumière à ceux qui se sont adressés à moi de la part du Roy. Et si ce service a profité ou non, nul ne le sait mieux que M. de Mayne [le Duc de Mayenne] auquel par le commandement de sa Majesté, plusieurs paquets furent fait voir afin qu'il connaît la conspiration que ses partisans mêmes faisaient contre lui. »

Faire savoir qu'on a cassé un code ennemi est maladroit et l'on tente le plus souvent de masquer que l'on a percé les méthodes de chiffrage de l'adversaire, de manière à éviter que celui-ci en adopte d'autres. Il est donc raisonnable de penser que le sacrifice stratégique de la publication de la lettre de 1590 ne fut consenti que parce qu'il plaçait le duc de Mayenne en position gênante : les lettres déchiffrées l'accusaient de vouloir s'emparer du trône. Viète dans sa lettre indique qu'il garde les originaux : « J'en tiens et garde soigneusement les originaux que je reconnaiss en bonne forme et bien sellez et signez, lesquels je représenterai toujours avec mes traductions, et les Alphabets et Dictionnaires que j'ai compris pour y parvenir et quand de par vous il me sera ordonné. »

Assez étrangement, malgré les révélations de Viète, les Espagnols ne changèrent pas leur système de chiffrage ! Le roi d'Espagne, en revanche, formula une plainte auprès

du pape accusant Viète d'avoir utilisé la magie pour lire les lettres chiffrées. Le pape ne réagit pas : il disposait de gens sachant aussi lire les lettres du roi d'Espagne et savait qu'aucune aide diabolique n'était nécessaire pour réaliser le travail de décryptage opéré par Viète.

Si le roi d'Espagne n'imaginait pas possible la lecture de ses lettres chiffrées c'est qu'à cette époque la cryptologie était un art plus qu'une science, et que les spécialistes dans les différentes cours pensaient tous disposer de méthodes sûres... même alors qu'elles ne mettaient en œuvre que des systèmes fragiles comme le codage espagnol.

La cryptologie au Moyen Âge et à la Renaissance

Au Moyen Âge, les cryptosystèmes utilisés sont faibles : remplacement des voyelles par des points, utilisation d'un alphabet étranger, introduction de signes spéciaux. Ainsi, les templiers se servent d'une correspondance simple lettre-symbole fondée sur la croix des huit bénédicteuses qui est l'emblème de l'ordre, mais qui n'assure qu'un secret illusoire.

Chiffre de SULLY (1599)																					
A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T	U	X	Y	Z
3	20	8	11	7	2	15	4	12	16	1	13	17	5	19	14	18	6	9	22	1	10
J	?	f	b	d	x	y	j	c	g	b	a	n	m	v	w	c	q	o	l	t	
Y	?	z	3	+	?	+	of	t	h	x	#	ff	..	h	s	+	2	?	?	h	

le Roy	4	ayant	a,	il	9
le Pape	3	ans	b,	le	8
le Roy d'Espagne	2	argent	c,	la	7
l'Empereur	5	attendu	d,	lettre	6
le Grand Seigneur	6	attendant	e,	mois	5
le Royne d'Angleterre	7	apres	f,	ment	4
le Roy d'Ecosse	8	bay	g,	moys	3
l'Archiduc d'Autriche	9	ban	h,	nous	2
l'Infante d'Espagne	10	bau	j,	nostre	1
les Etats des Pays-Bas	11	balli	k,	nest	b
la Seigneurie de Venise	12	car	l,	non	c
le Roy du Danemarck	13	convient	m,	ouverture	d
le Roy du Suede	14	cn	n,	occasion	e
les Cantons Suisses	15	contenant	o,	outre	f
le duc de Savoye	16	donne	p,	obligation	g
le duc de Lorraine	17	dire	q,	pour	h
le duc de Guise	18	dont	r,	par	i
le prince Maurice	19	despeches	s,	pro	k
le comte d'Ecosse	20	dequoy	t,	parquet	l
le secretaire GAYL	21	ent	u,	que	m
le secretaire LEVISTON	22	encores	x,	qui	n
le sieur de BOISSIZE	23	et	y,	quoy	o
le sieur de BUZENVAL	24	entre	z,	quand	p
l'evêque de Glasco	25	feut	?	quelle	u
France	26	fois	b	requ	r
Ecosse	27	foy	c,	reception	s
Flandres	28	grans	d	reste	t
Hollande	29	gens	e	sans	q
Angleterre	30	garde	f,	sainct	x
Suede	31	guenee	g,	action	y
Danemark	32	hon	h	S.M., V.M.	z
Lettres nulles :	Y,	hommes	i,	tout	2,
Doulement :		hautes	l,	tant	3,
veru	8,	heures	m,	toutefois	4,
venant	9,	je	n,	tost	5,
vérifiable	10	intention	o,	vous	6,
viva	11,	jay	p,	vostre	7,

2. Le nomenclateur de Sully. Jusqu'au XIV^e siècle, un procédé de chiffrement mixte combinant substitutions des lettres et substitutions de syllabes ou de mots entiers domina la cryptographie. Le nomenclateur est composé d'un alphabet de chiffrement avec homophones (une même lettre est représentable de plusieurs façons), lettres nulles (lettres comptant pour rien) et d'une table de chiffrement ou répertoire, c'est-à-dire d'une liste de noms propres et de syllabes avec leurs codages. Le nomenclateur de Sully fut vraisemblablement composé par Viète.

Les besoins de la diplomatie ont donné naissance à la cryptologie d'État, car les ambassadeurs souhaitent protéger leurs correspondances. Les « secrétaires-chiffreurs » ont pour tâche de concevoir les méthodes de chiffrements et de décrypter les messages interceptés chez leurs ennemis.

Parmi les cryptologues dont le nom nous est parvenu, plusieurs méritent attention. Gabrieli de Lavinde, secrétaire du pape Clément VII, utilise en 1379 le premier nomenclateur connu. Cette méthode sera celle de la cryptographie d'État pendant plus de trois siècles.

Leon Batista Alberti (1404-1472) architecte et mathématicien florentin invente un cadran chiffrant dont le principe sera repris jusqu'au XIX^e siècle. Il nous laisse aussi le premier traité manuscrit de cryptologie. L'abbé Trithème (1462-1516) décrit un procédé de chiffrement fondé sur un tableau carré qui, une fois perfectionné, conduira à un système de chiffrage relativement sûr. Son *Polygraphiae libri sex*, premier ouvrage de cryptologie imprimé, servira d'ouvrage de référence pendant plusieurs siècles.

Le physicien napolitain Giovani Battista Porta (1535-1615) expose dans son ouvrage de 1563 la méthode de la substitution bigrammique : le texte clair est chiffré deux lettres par deux lettres. Le mathématicien milanais Gerolamo Cardano (1501-1576), à qui on doit le... cardan et une méthode de résolution des équations du troisième degré, est l'inventeur du mécanisme de l'autoclé qui consiste à utiliser le texte qu'on doit chiffrer comme clé de chiffrement (en plus d'une clef courte pour le début).

Le diplomate français Blaise de Vigenère (1523-1596) nous laisse un *Traicté des chiffres ou secrètes manières d'escrire* où une méthode appelée aujourd'hui « Carré de Vigenère » résistera à toutes les attaques de la cryptanalyse jusqu'au XIX^e siècle.

À la cour de France, l'usage du chiffre fut vraisemblablement introduit par Louis XII, mais ce n'est que sous François I^e que l'on est certain de tels usages. Dans une lettre de 1546, le roi interdit en effet aux gens de la cour « d'escrire aucune lettre et advertisement en chiffres et caractères ou noms supposés ou autrement » exception étant accordée aux ambassadeurs. François I^e utilisait les services d'un secrétaire chiffrer Philibert Babou (1500-1569) qui excellait à déchiffrer les messages dans les langues de l'époque. Babou continua d'exercer son talent au service de Henri II, fils de François I^e. Viète prit plus tard ce rôle.

Les nomenclateurs

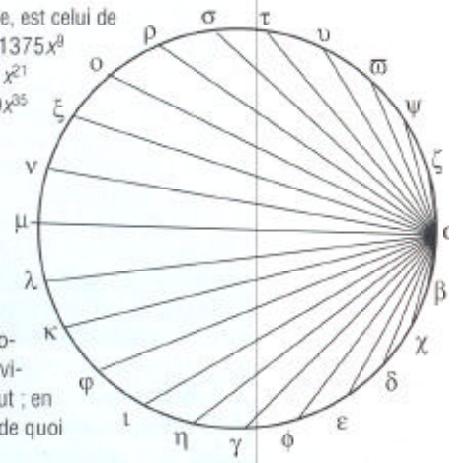
Malgré des progrès théoriques et l'invention d'une grande variété de procédés du XIV^e au XVI^e siècle, dont certains réellement difficiles à casser, la principale méthode mise en œuvre par les ambassadeurs et pour les courriers militaires est le codage par substitution monoalphabétique et ses perfectionnements élémentaires (par exemple, les chiffrements différents d'une même lettre, et les nomenclateurs où des mots et des syllabes chiffrés d'un seul coup).

La substitution monoalphabétique (une unité de code par lettre) est assez facile à élucider, car les lettres les plus fréquentes se repèrent facilement et à partir d'elles les autres sont identifiées de proche en proche (pourvu que le mes-

L'exploit mathématique

François Viète a relevé le défi posé par le Belge Adrien Romain. Le problème, traduit en notation actuelle, est celui de la résolution d'une équation de degré 45 : $45x - 3795x^3 + 95634x^5 - 1138500x^7 + 7811375x^9 - 34512075x^{11} + 105306075x^{13} - 232676280x^{15} + 38942375x^{17} - 488494125x^{19} + 483841800x^{21} - 378658800x^{23} + 236030652x^{25} - 117679100x^{27} + 46955700x^{29} - 14945040x^{31} + 3764565x^{33} - 740259x^{35} + 111150x^{37} - 12300x^{39} + 945x^{41} - 45x^{43} + x^{45} = a$, où a est une constante s'exprimant avec des radicaux carrés.

Au XVI^e siècle, on ne savait résoudre que des équations de degré bien inférieur, un, deux ou trois et exceptionnellement un peu plus. Cependant Viète en utilisant sa nouvelle algèbre avait remarqué que diviser un angle en 3 produisait une équation de degré 3 et plus généralement que diviser un angle en n conduisait à l'écriture d'une équation de degré n . Il en déduisit que l'équation devait provenir de la division d'un angle en 45 parties égales, ce qui était bien le cas. En 1595, Viète fournit non seulement la solution qu'attendait Adrien Romain, mais aussi 22 autres solutions (pour lui, seules les solutions positives comptent vraiment), chacune avec huit décimales exactes. La méthode utilisée par Viète pour traiter le problème d'Adrien Romain rappelle celles mises en œuvre pour déchiffrer des codes inconnus, car c'est en devinant d'où pouvait provenir l'équation (d'un problème géométrique) et ce qu'elle signifiait qu'il en vint à bout ; en cryptanalyse de même on réussit d'autant mieux qu'on connaît la langue du message et qu'on a une idée de quoi il parle. À droite, la représentation géométrique des 23 solutions de l'équation d'Adrien Romain.



sage soit assez long). Cette méthode de cryptanalyse fondée sur les fréquences des lettres a été inventée un très grand nombre de fois, mais la plus ancienne mention de cette idée est due au savant arabe Qalqashandi en 1412. C'est pour la contrer qu'on associe plusieurs unités de code à la même lettre (les lettres fréquentes sont impossibles à repérer). L'utilisation de lettres nulles (ne chiffrant rien) masque les lettres répétées facilement repérables (on en insère une ou plusieurs entre les lettres répétées). Les mots ou syllabes codés d'un bloc rendent le tout encore plus difficile car on chiffre ainsi de plusieurs façons un même mot : soit lettre par lettre, soit syllabe par syllabe, soit globalement.

La cryptologie politique du temps de Viète se réduit aux nomenclateurs et, bien sûr, ceux comportant de nombreuses lettres aux multiples codages et de nombreux mots codés d'un bloc sont les plus difficiles à déchiffrer. Le souci d'accroître la sécurité conduira naturellement à une surenchère dans la taille des nomenclateurs amenant en quelques siècles des nomenclateurs comportant plusieurs dizaines de pages.

La sous-estimation des possibilités de l'analyse

Les premiers nomenclateurs étaient ordonnés : les mots et unités codiques y sont classés en parallèle dans l'ordre naturel, c'est-à-dire l'ordre alphabétique pour les lettres ou groupes de lettres et par ordre croissant pour les chiffres ou nombres. Il est remarquable que d'autres systèmes de chiffrage plus sûrs sont connus et décrits dans les traités (par exemple dans le traité de Blaise de Vigenère), mais ne semblent pas avoir été employés dans les ambassades et pour les courriers militaires à cette époque. Le fait qu'aucune méthode avec clef (qu'on peut changer et choisir spécifiquement pour chacun de ses correspondants) n'ait été mise en œuvre est étonnant et fragilisait beaucoup les systèmes.

Viète n'est pas l'inventeur de la cryptanalyse qui a existé dès les premiers messages chiffrés, mais, comme nous allons le voir, il a proposé une méthode rigoureuse et mathématique de cryptanalyse qui constitue une avancée sensible.

Si l'on recherche les premières traces certaines de la cryptanalyse, il semble qu'elles se trouvent chez les Arabes où, dans l'encyclopédie *Subh al-a sha* achevée en 1412, une section du texte est consacrée à l'art de la dissimulation. Son auteur, Qalqashandi, qui connaît les méthodes de substitution et de transposition propose le premier petit texte de cryptanalyse de l'histoire. Il se fonde sur l'analyse linguistique de la langue arabe. Sont en particulier indiquées des listes de lettres qu'on ne trouve jamais associées dans un même mot ou qui y voisinent rarement, ainsi que des combinaisons impossibles. On y lit aussi l'ordre de fréquence des lettres en arabe telles qu'on les trouve dans le Coran. L'idée de l'analyse des fréquences est donc connue à cette époque. On ignore quels usages furent faits de ce savoir cryptanalytique et s'il joua un rôle dans l'histoire de l'Islam de cette période. Ce premier texte semble avoir été ignoré et rien ne suggère que Viète ni aucun de ses contemporains n'en eurent connaissance.

Pour faire progresser cette technique, il fallait un esprit ouvert aux raisonnements et calculs organisés. Ce n'est donc pas un hasard si Viète, inventeur de l'algèbre moderne, comprit le premier l'intérêt de procéder systématiquement et logiquement. Le mathématicien Adrian Albert, en 1941, lors d'une réunion de la Société américaine de mathématiques affirma que « La cryptographie est plus qu'un sujet s'exprimant sous forme mathématique, [...] la cryptographie abstraite est identique aux mathématiques abstraites » : les mathématiques, grâce à l'impulsion de Viète, feront des miracles.

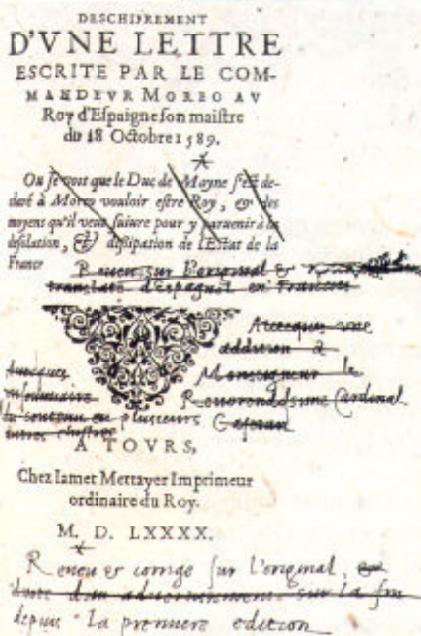
Il est amusant de constater que Vigenère ne croyait pas à la possibilité d'une cryptanalyse systématique. Dans son *Traicté des chiffres* de 1586, il décrit la cryptanalyse comme un « exercice certes d'un inestimable rompement de cerveau ». Pourtant son traité est surtout consacré à la description et à la conception de chiffres robustes.

Wallis un peu plus tard lui aussi pensait que seuls la chance et l'acharnement permettent de déchiffrer un message. Pour lui, il n'existe pas de méthode définie pour lire des messages chiffrés et ceux qui s'y essayent ne peuvent qu'espérer faire de bonnes conjectures et voir apparaître quelques éléments d'où ils pourront déduire qu'ils sont sur le chemin du vrai.

La règle infaillible de Viète

Viète propose dans son mémoire de 1603 une méthode (différente de la méthode des fréquences qui s'applique mal lorsqu'un texte est court) pour repérer les voyelles dans un message chiffré par substitution monoalphabétique. Cette méthode procède uniquement par raisonnement et est équivalente à la résolution d'un système d'équations : elle marque la naissance de la cryptanalyse mathématique.

Illustrons sur un exemple le procédé de Viète. Prenons le texte suivant, chiffré par substitution monoalphabétique (une substitution monoalphabétique échange l'ordre des lettres de l'alphabet). Exemple : A → f ; B → r ; C → k ; ... ; Z → j. :



t y e n p y e n l q w q y f y k l m
l q t y h g m j w n k k y j m f o g
w g x y k y w j y l k q a f y z j n f
w g q k u q y l y

Il y a $26! = 4,03 \cdot 10^{26}$ substitutions possibles et bien sûr nous ignorons laquelle a été utilisée. Il n'est pas envisageable – même aujourd'hui – de rechercher la substitution par énumération exhaustive des divers cas, car ceux-ci sont trop nombreux.

Pour découvrir la substitution, nous allons, en suivant les idées de Viète, partir à la recherche des voyelles principales A E I O U en faisant l'hypothèse que parmi trois lettres consécutives du texte, on trouve toujours une ou plusieurs de ces cinq lettres. En français, cette hypothèse n'est pas toujours satisfait (elle l'est plus souvent en espagnol).

Plus étonnant peut-être, la cryptanalyse à cette époque et pendant longtemps encore sera considérée comme une supercherie. Sauf dans les cas élémentaires, on juge le déchiffrement impossible et donc ceux qui se vantent d'y réussir sont des menteurs. Voltaire traitera de charlatans les gens qui prétendent lire les lettres chiffrées, car il juge impossible de comprendre une langue que l'on n'a pas apprise. Peut-être cette sous-estimation du pouvoir du raisonnement et des méthodes systématiques était-elle feinte et destinée à donner confiance à ceux qui pensaient naïvement ne prendre aucun risque en utilisant des systèmes de chiffrement élémentaires.

Que contient le mémoire de Viète de 1603 ? D'abord il remarque que le roi d'Espagne utilise le même chiffre vers tous ses vices-rois et ambassadeurs et que ceux-ci utilisent le même chiffre pour lui répondre. C'est, évidemment, un point faible !

À chaque lettre de l'alphabet du code espagnol, indique Viète, sont associés trois ou quatre symboles, et même un peu plus pour les voyelles. Aux syllabes sont associés un ou deux symboles. Deux jargons (listes de mots avec leurs équivalents codiques) sont ajoutés, l'un pour les mots fréquents, l'autre pour les noms propres. La diversité des syllabes permet de décomposer un même mot de plusieurs façons. Viète donne l'exemple de al-do-bran-di-no et al-do-b-ra-n-din-o.

gnol), mais pour un court texte, elle a une bonne chance de l'être et, de toutes les façons, si cette tentative ne marche pas, on pourra reprendre le raisonnement avec des paquets de 4 lettres consécutives au lieu de 3.

Ainsi, le triplet t y e du début comporte au moins une voyelle. Le triplet n / p qui n'a aucune lettre commune avec le premier triplet aussi. Le triplet p y e contient aussi une voyelle, mais on ne le garde pas, car les lettres y et e sont déjà dans le premier triplet sélectionné. En avançant un peu plus loin, on trouve le triplet q w q, puis plus loin enfin h g m. Tous les autres triplets utilisent une des 11 lettres ainsi repérées : t y e n l p q w h g m. Cette procédure assure que parmi ces 11 lettres se trouvent les cinq voyelles A E I O U (sous réserve qu'elles soient toutes utilisées au moins une fois, ce que nous supposerons).

Maintenant considérons f y k qui est un triplet de lettres consécutives du texte chiffré. Il contient par hypothèse une voyelle. Ce ne peut pas être f n i k (qui ne sont pas parmi les 11 lettres repérées au-dessus). Donc y est le représentant d'une des lettres A E I O U. Le triplet j m f conduit de même à repérer que m est le représentant d'une des lettres A E I O U ; j n f donne que n représente une des lettres A E I O U ; enfin les triplets k u q et o g f donnent que q et g aussi sont des lettres parmi A E I O U.

Nous savons donc précisément où sont les voyelles A E I O U. À partir de là, le décryptement est devenu plus facile. Nous en donnons la solution en fin d'article après la bibliographie, page 95. Viète, en 1590, publie un texte donnant le contenu en clair d'une lettre chiffrée du commandeur Moreo et dans laquelle il apparaît que le duc de Mayenne souhaite s'emparer du trône occupé par Henri IV. Cette publication, montrant que les lettres chiffrées du roi d'Espagne étaient lues par Viète, est un sacrifice sans doute volontairement consenti par le roi Henri IV pour mettre le duc de Mayenne en difficulté. Cet imprimé n'existe aujourd'hui qu'en deux exemplaires annotés, de la main de Viète pense-t-on.

Cette richesse du système se retourne contre les Espagnols, car les messages sont envoyés plusieurs fois (jusqu'à quatre fois) et, à chaque envoi, le message est chiffré différemment. Grave erreur : dès que l'on commence à découvrir quelques associations, en passant d'un chiffrage du message à un autre on tire l'écheveau des associations jusqu'à connaître complètement le nomenclature si le message est assez long.

Tâtonnements et méthode

Viète insiste sur ce qu'il appelle les chiffres essentiels, c'est-à-dire les nombres qui sont présents dans le texte chiffré sans être eux-mêmes chiffrés et qu'on repère, car ce sont des dates ou des nombres arrondis comme 500, 4 000, 10 000 etc. Parfois les chiffres essentiels sont identifiables à cause de marques particulières comme des accents ou des points. Derrière de tels nombres, les mots possibles sont faciles à deviner. Derrière 4 000 le mot fantassin (fanti) est probable. Derrière 100 000 ce sera plutôt ducados (ducats). Proche d'un nombre correspondant à une année, on trouvera sans doute le nom d'un mois ou d'un jour. Viète indique aussi d'autres pistes pour découvrir des mots en début ou en fin de message.

Il précise qu'il faut travailler systématiquement en décomptant tous les symboles, combien de fois chacun

apparaît, comment ils se suivent et « n'épargner ni le labeur ni le papier ».

Les messages chiffrés d'Italie n'usent guère que des dix caractères 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, nous dit Viète. Mais ils les accouplent admirablement pour représenter les lettres, les syllabes et les noms propres et les mots les plus fréquents. » Viète propose une méthode qu'il prétend infaillible et dont il explique le principe sur l'exemple d'un chiffre simple (monoalphabétique : une lettre – un symbole).

Il propose d'examiner les groupes de deux symboles consécutifs, ou de trois symboles consécutifs qu'il appelle dyades et triades. On en collecte un certain nombre dans le message et on utilise la propriété que, dans chacun, se trouve une voyelle. C'est une propriété de l'italien : sauf en de rares exceptions, jamais trois consonnes ne se suivent. Un examen soigneux conduit à découvrir quelles sont les voyelles. Ce travail de raisonnement proposé par Viète est équivalent à la résolution d'un système d'équations et si Viète ne ramène pas explicitement à un problème algébrique le traitement logique qu'il fait sur les dyades et les triades pour identifier les voyelles, il est évident que l'expérience qu'il a des situations analogues en mathématiques, comme par exemple les systèmes d'équations, est sans doute à l'origine de son idée.

Viète précise ensuite que ce travail qui lui permet de repérer les voyelles dans le cas d'un chiffre simple se généralise à des « chiffres composés » (avec codages différents de la même voyelle) à condition de prendre un nombre plus grand de dyades et de triades. Cette méthode « algébrique » ne repose pas sur l'intuition, mais exploite une propriété linguistique du message clair. Sa technique, qu'il nomme règle infaillible, est une déduction quasi mécanique, algorithmique dirait-on aujourd'hui, de quelques éléments (les voyelles) qui ensuite permettent l'élucidation complète du système de chiffrage. Viète insiste sur l'aspect très général de sa méthode et ses possibilités d'extension. C'est la première trace connue de ce que nous nommons aujourd'hui la cryptanalyse mathématique.

Viète tire de la nature automatique de l'attaque du message l'assurance que sa règle est infaillible et croit pouvoir affronter tout chiffre qu'on lui proposera. À dire vrai, sa méthode s'applique difficilement aux codes polyalphabétiques (divers alphabets de substitution sont utilisés à tour de rôle et non plus un seul) ou aux méthodes avec transpositions (l'ordre des lettres du message clair n'est pas conservé). Quant aux nombreuses méthodes que le XX^e siècle a imaginées, fondées sur l'arithmétique des nombres premiers ou d'autres idées ignorées du temps de Viète, elles donneront lieu à des développements de la cryptanalyse mathématique qui iront bien au-delà de ce que le XVII^e siècle pouvait imaginer.

Tout chiffre est inviolable... pour celui qui le conçoit

Une constante de l'histoire de la cryptologie est l'assurance des inventeurs que leur code est supérieur à tous ceux qui ont précédé et que cela rend parfaitement sûr son usage. Cent exemples attestent qu'il s'agit souvent d'une illusion. La connaissance de l'histoire de la cryptologie ne semble pas un vaccin suffisant pour s'opposer à la suffisance des cryptologues qui

apparaît incorrigible (peut-être faudrait-il rechercher une explication psychologique à cet état de choses ?).

L'idée que les codes actuels sont pratiquement inviolables est souvent défendue aujourd'hui : il est admis et même proclamé que nous sommes dans une situation où le bouclier a pris l'avantage sur l'épée. Les arguments avancés sont parfois généraux, du type : « Les progrès de la science cryptologique permettent de concevoir des codes sûrs. » Parfois aussi on se fonde sur des idées plus subtiles comme la suivante. La puissance de l'informatique permet d'utiliser des systèmes de chiffrage dont la robustesse croît plus vite que la capacité à les briser. En effet, remarque-t-on, lorsque l'on augmente la longueur des clés de dix digits, on ralentit de manière modérée la vitesse du chiffrage, alors qu'en revanche, on augmente le temps de décryptage (pour quelqu'un ne connaissant pas la clé et travaillant à l'aide d'une méthode exhaustive) d'un facteur $1024 (2^{10}=1024)$ puisque le nombre de clés est multiplié par 1024. Cet argument n'est pourtant guère recevable, car on y raisonne « à algorithmes constants », c'est-à-dire en supposant qu'on ne découvrira pas de nouvelles méthodes pour contourner l'énumération exhaustive des clés. L'histoire, à nouveau, montre que de nouvelles idées apparaissent sans cesse et réussissent à faire ce qu'on imaginait impossible : raisonner « à algorithmes constants » est absurde et le moyen assuré pour un cryptologue d'avoir de mauvaises surprises.

Viète en introduisant les raisonnements mathématiques purs dans l'art des briseurs de code a prouvé que certains de ses prédecesseurs étaient naïfs en croyant à l'inviolabilité des nomenclateurs – même lorsqu'ils sont volontaires. Pourtant assez étrangement, il fut lui-même victime de la myopie commune des cryptologues en n'imaginant pas qu'on pourrait aller beaucoup plus loin dans la conception de systèmes de chiffrage perfectionnés et dans la conception des nouvelles méthodes de cryptanalyse mathématique, science qu'il venait d'inventer.

Jean-Paul DELAHAYE est professeur d'informatique à l'Université de Lille. delahaye@lifi.fr

Dossier hors-série Pour la science sur la cryptographie, *L'art du secret*, Juillet-Octobre 2002.

Patrick HEBRARD, *La cryptologie dans l'histoire*, vol. 1, *Essai sur l'histoire secrète des chiffres de l'antiquité à la Première guerre mondiale*. As des Réservistes du Chiffre et de la Sécurité de l'Information, mars 2001.

Peter PESIC et François VIÈTE, *Father of Modern Cryptanalysis – Two New Manuscripts*, Cryptologia, vol. 21, n° 1, 1997.

Peter PESIC, *Secret Symbols, and Systems, Parallels Between Cryptanalysis and Algebra, 1580-1700*, Isis, Vol. 88, pp. 674-692, 1997.

Jean-Paul GUICHARD et Jean-Pierre SICRE, *François Viète, un juriste mathématicien*, les éditions de l'Actualité Poitou-Charentes, 1995.

Frédéric RITTER. Les références des documents retrouvés par Pesic sont : Bibliothèque de l'Institut de France, Ms. 2004-2012, *Manuscrits de Frédéric Ritter relatif à François Viète* ; la transcription du mémoire de 1603 est en MS. 2009, 182-189.

Pages internet sur Viète : <http://www.district-parthenay.fr/partenay/cre-parth/GUICHARD.Jp/VIETEaccueil.html>

Le message en clair de l'article est :

t y e n l p y e n l q w q y f y k l m l q t y h g m j w n k k y j
L E M A T H E M A T I C I E N E S T U T I L E P O U R C A S S E R
m f o g f w g x y k y w j y l k q a f y z j n f w g q k u q y l y
U N B O N C O D E S E C R E T S I G N E F R A N C O I S V I È T E

Auteur
Bibliographie



DÉCRYPTEMENT DES SUBSTITUTIONS

Dès la Renaissance, les cryptologues les plus talentueux s'étaient rendu compte qu'une substitution simple, dite monoalphabétique, ne masquait pas la fréquence d'apparition des symboles du cryptogramme. En effet, dans toutes les langues, certaines lettres (E, S, A, N...) ou groupe de lettres (ET, SS, LL...) reviennent statistiquement plus souvent que d'autres dans les textes usuels. En fran-

çais par exemple, la fréquence d'apparition statistique des lettres (dans un texte suffisamment long) est indiquée dans le tableau en bas de la page.

Imaginons le message suivant chiffré avec un «Jules César» où chaque lettre est remplacée par la lettre décalée de

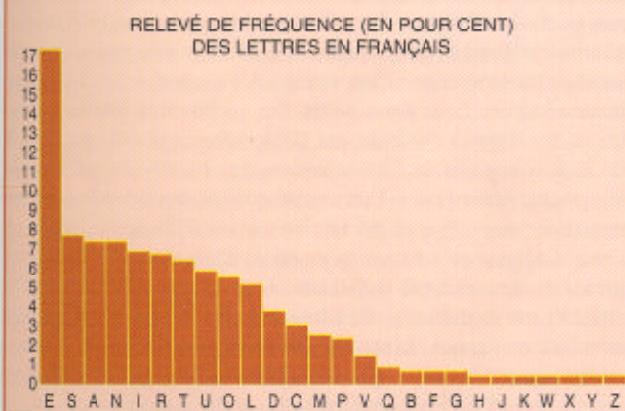
trois rangs dans l'alphabet (ainsi E est remplacée par H) :

clair : ENVOYEZ DU RENFORT AVEC DES ARMES

crypto : HQYRBHCGXUHQIRWDYHFGHVDPHV

On peut bien sûr, pour déchiffrer le message, essayer les 25 combinaisons de décalage (26 décalages moins le décalage identique, c'est-à-dire 0) pour retrouver le «clair», mais il est globalement plus rapide d'utiliser les propriétés statistiques de la langue. On constate qu'il y a 6 «H» dans le cryptogramme et aucune lettre de fréquence supérieure, de sorte qu'il faut essayer $H = E$, c'est à dire un décalage de trois lettres. Pour des substitutions monoalphabétiques plus complexes, il convient d'identifier les autres lettres et notamment on rechercherait le chiffre des lettres S, A et N, qui sont les lettres les plus fréquentes ; avec de la persévérance et de l'intuition, on arrivera au déchiffrement. Bien sûr, plus le cryptogramme est long, plus le relevé de fréquence sera significatif.

Pour éviter ce type d'analyse cryptographique conduisant au décryptement, les cryptologues ont mis au point les substitutions simples avec homophones, c'est-à-dire plusieurs symboles pour chiffrer les lettres d'occurrence les plus fréquentes (le «E», le «S», le «A», etc.). Le chiffre de Philibert Babou de 1558 en



est un bon exemple. Pour dérouter le décrypteur, Babou insère également des lettres nulles, des symboles spéciaux pour chiffrer les bigrammes («EE», «FF», «MM», ...) et un petit nomenclateur compressif qui complexifie encore le décryptement.

Cependant, ces codes avec homophones restent des substitutions simples. Bien plus efficace est la substitution poly-alphabétique mise en œuvre par Vigenère qui utilise un alphabet de chiffrement différent (fonction de la clé) à chaque chiffrement de lettre :

ENVOYEZ L'ACADEMIE
CLEF CLEF CLEF CLEF CLEF CLEF
GYZ TAP DQC NEAG WIWK P

Ici la clé est le mot «CLEF» et l'on procède comme indiqué sur la figure 9. Le principe de décryptement d'une substitution polyalphabétique revient toujours à se ramener à une situation de décryptement d'une substitution simple. Pour cela, il convient tout d'abord de déterminer la longueur de la clé. On peut utiliser, par exemple, la méthode de Kasiski (1863) qui consiste à observer des séquences de lettres identiques dans le crypto indiquant vraisemblablement que c'est le même mot qui a été chiffré avec la même position relative de la clé. On en déduit alors que la distance qui sépare ces deux séquences de lettres identiques dans le cryptogramme est un multiple de la taille de la clé. En répétant ces corrélations et par tâtonnements, on finit par retrouver la longueur de la clé, quatre dans notre exemple. On réarrange ensuite le cryptogramme par groupe de 4 lettres comme suit :

GYZT
APDQ
CNEA
CWIW
KP

Pour chaque colonne, on procède à un décryptement d'une substitution simple de type Jules César. Dans notre exemple, on essayerait $C = E$, etc, et on reconstituerait la clé pour terminer le décryptement.