

KRYPTOLOGIE

1. Einleitung

1.1 Begriffe, Übersicht

Die Kryptologie ist die Wissenschaft (bzw. die Kunst), eine Nachricht mit meist geheimem Inhalt an einen Empfänger zu senden, so dass sie von Unbefugten nicht gelesen werden kann. Die Geschichte der Kryptologie verläuft fast parallel zur Entwicklung der Schrift, denn seit jeher hatten die Menschen das Bedürfnis, Nachrichten geheim zu halten. Etliche Menschenschicksale und sogar der Ausgang von Kriegen wurden direkt durch die Qualität einer guten Verschlüsselung, bzw. durch die Möglichkeit, geheime Botschaften zu entschlüsseln, beeinflusst.

In diesem Skript sollen einige grundlegende Verfahren dargestellt werden, die in der Geschichte der Kryptologie angewendet wurden und es soll aufgezeigt werden, wie es findigen Kryptanalytikern gelang, diese Codes zu knacken.

In der Kryptologie geht es um zwei Hauptprinzipien: das Verbergen von Informationen und das Entschlüsseln einer codierten Nachricht. Eine Nachricht kann dadurch verborgen werden, dass man die gesamte Nachricht „versteckt“, so dass der „Feind“ überhaupt nichts von der Nachrichtenübergabe bemerkt (Steganographie). Man kann die Nachricht jedoch auch verschlüsseln, so dass sie für einen nicht Eingeweihten nutzlos ist. In diesem Fall spricht man von Kryptographie. Als Kryptoanalyse bezeichnet man das Gebiet, welches darauf abzielt, eine geheime Botschaft zu knacken, d.h. zu entschlüsseln. Abbildung 1 zeigt einen Überblick über die Begriffe

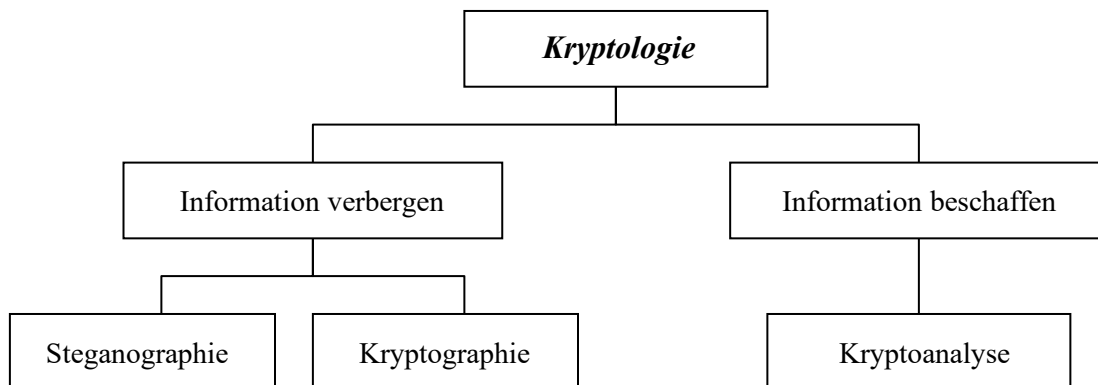


Abb.1: Überblick über die Kryptologie

1.2 Steganographie

Die Kunst, eine Nachricht zu verbergen, so dass sie nur durch den designierten Empfänger gelesen wird, würde eine eigene Abhandlung füllen. Bereits in den Schriften von Herodot finden sich Beispiele von Steganographie. So soll beispielsweise der ins Exil verbannte Grieche Demaratos um 480 v.Chr. seine Landsleute vor der Invasion der Perser gewarnt haben, indem er das Wachs einer Schreiftafel entfernte und seine Nachricht auf das darunter liegende Holz kratzte. Danach erneuerte er die Wachsschicht der Schreiftafel und sandte sie den Griechen. Dadurch, dass die Schreiftafel leer war, passierte sie die persische Zensur und die Griechen waren vom Angriff des damaligen Perserkönigs Xerxes gewarnt.

Ebenfalls bei Herodot findet man die Geschichte des Griechen Histiaeus, welcher Aristagoras von Milet zum Aufstand gegen die Perser aufrufen wollte. Er rasierte einen Boten kahl und brannte die Botschaft auf dessen Kopfhaut. Nachdem die Haare nachgewachsen waren, begab sich der Bote persönlich zu Aristagoras und rasierte seine Haare, so dass dieser die Nachricht lesen konnte. Auch die Verwendung von unsichtbarer Tinte, welche erst bei Kontakt mit einer bestimmten Flüssigkeit oder bei Erhitzung ans Tageslicht tritt, gehört zur Steganographie. Bereits im 15. Jahrhundert beschreibt

z.B. Giovanni Porta, wie man eine Nachricht in einem Ei verstecken kann. Man mische dazu Alaun und Essig und schreibe die Botschaft auf das hartgekochte Ei. Die Nachricht dringt durch die poröse Schale des Eis und ist nur sichtbar, wenn man die Schale vom Ei entfernt.

Im zweiten Weltkrieg wurde häufig eine geheime Botschaft der Länge einer Textseite per Mikrofilm auf die Grösse eines kleinen Punktes verkleinert, welcher dann in einem scheinbar harmlosen Brief als Satzzeichen auftauchte.

Die Steganographie hat den Nachteil, dass die Nachricht sofort offen liegt, wenn die Botschaft abgefangen wird. Hätten die Grenzposten z.B. dem Boten von Histiaeus die Haare rasiert, da sie Verdacht schöpfen, dann wäre die Nachricht vollkommen offen gelegen.

2. Grundlegende Kryptographietechniken

In der Kryptographie geht es jeweils darum, eine Nachricht, die in Klartext (K) vorliegt zu verschlüsseln. Es resultiert ein Geheimtext (G), der im Optimalfall nur vom Empfänger wieder in den Originalzustand (K) überführt werden kann.

Die Ausdrücke verschlüsseln, codieren und chiffrieren werden hier als Synonyme verwendet. Genau genommen bezieht sich chiffrieren auf einzelne Buchstaben, während beim codieren ganze Wörter oder Wortteile verschlüsselt werden. Grundsätzlich werden beim Verschlüsseln (genauer gesagt beim chiffrieren von Buchstaben) zwei Methoden unterschieden: die Transposition und die Substitution.

2.1 Transposition

Wird ein Text durch ein Transpositionsverfahren verschlüsselt, so werden die Buchstaben untereinander vertauscht ohne dass sie sich selbst ändern, wodurch die Nachricht unkenntlich wird. Beim knacken resp. decodieren der Botschaft geht es dann darum, die ursprüngliche Reihenfolge der Buchstaben wieder herzustellen. Transpositionsalgorithmen werden in dieser Arbeit nicht im Detail behandelt. Dennoch sollen hier zwei Beispiele angeführt werden.

2.1.1 „Gartenzaun-Transposition“

Die sogenannte „Gartenzaun-Transposition“ wird z.T. bereits von Schulkindern verwendet. Dabei wird ein Klartext K auf zwei Linien verteilt, wobei immer abwechselungsweise ein Buchstabe des Textes auf die erste und ein Buchstabe auf die zweite Linie geschrieben wird. Der Geheimtext entsteht dann dadurch, dass man die Buchstaben der zweiten Linie an die erste Linie anfügt.

Klar: DieSchattenwerdenlänger, dergraue, grameGrillenfängerstreicht
 Decatnednägrdrur, rmGilnägrect
 iShtewrelne, egaegaerlefnesrih

Geheim: Decatnednägrdrur, rmGilnägrectiShtewrelne, egaegaerlefnesrih

Es gibt natürlich unzählige Varianten dieser einfachen Transpositionsverschlüsselung. So könnte man z.B. drei oder vier Zeilen untereinander verwenden oder man könnte zusätzlich jeweils den ersten mit dem zweiten Buchstaben vertauschen, bevor man die „Gartenzaun-Transposition“ anwendet etc.

2.1.2 Die Skytale

Die Skytale ist das erste Kryptographie-Verfahren, welches militärisch Verwendung fand. Sie wurde bereits im 5. Jahrhundert n.Chr. durch die Spartaner benutzt. Für die Verschlüsselung wird lediglich ein kantiger Holzstab verwendet um welchen ein Streifen Pergament gewickelt wird (siehe Abbildung 2). Die Nachricht wird nun der Länge nach auf den Stab geschrieben, so dass die Buchstaben beim Lösen des Pergamentes anscheinend sinnlos aneinandergereiht sind. Für das Entschlüsseln der Botschaft benötigt man lediglich einen Stab mit derselben Kantenzahl, auf welchen der Pergamentstreifen aufgerollt wird.

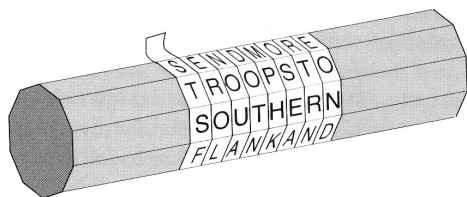


Abb.2: Skytale mit 10 Kanten

Aufgabe Wie hängen die Skytale und die „Gartenzaun-Transposition“ zusammen?
 Wie sicher ist eine Nachricht, die nach diesen Prinzipien chiffriert wurde?

2.2 Substitution

Im Gegensatz zur Transposition, bei welcher jeder Buchstabe seine Position ändert, behält bei der Substitution jeder Buchstabe seine Stelle bei. Was die Nachricht unleserlich macht, ist die Tatsache, dass die Buchstaben selbst ihre Gestalt ändern. Bereits im *Kamasutra*, einem Text der auf den indischen Gelehrten Watsjajana (ca. 400 n.Chr.) zurückgeht und Quellen bis vier Jahrhunderte vor Christus verwendet, kommt eine Beschreibung einer Substitutionsverschlüsselung vor.

Das Kamasutra (Sanskrit für „Leitfaden der Liebe“) ist die älteste bekannte Liebeslehre und sie enthält u.a. eine Liste von 64 Künsten, die Frauen studieren und beherrschen sollten. Darunter sind Klassiker wie Kochen, Bekleidung, Massage und Zubereitung von Parfüms, aber z.B. auch Schach, Teppichweberei und Buchbinderei gehört zur Liste. Die Nummer 45 auf der Liste dieser Künste ist die Kunst der Geheimschrift, die den Frauen rät, ihre Affären geheim zu halten. Dabei wird der Vorschlag gemacht, die Buchstaben des Alphabets zufällig zu paaren und dann jeden Buchstaben in der Nachricht durch seinen Partner zu ersetzen. Dies ist eine der ersten Beschreibungen eines monoalphabetischen Algorithmus, welche im nächsten Kapitel genauer behandelt werden sollen.

3. Monoalphabetische Algorithmen

3.1 Allgemeine monoalphabetische Verschlüsselung

„Monoalphabetisch“ wird ein Verschlüsselungsalgorithmus genannt, wenn das Geheimalphabet im Verlaufe der Verschlüsselung gleich bleibt. So wird z.B. ein A immer in ein c verwandelt, ein P immer in ein d etc. Eine einfache monoalphabetische Verschlüsselung wäre z.B. durch folgende Tabelle gegeben:

Klar	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
Geheim	f	o	b	t	h	x	n	y	i	a	s	g	l	d	u	z	m	v	p	j	e	w	q	c	r	k

So würde die Nachricht aus Kapitel 2.1.1 folgendermassen verschlüsselt:

Klar: DIE SCHATTEN WERDEN LAENGER, DER GRAUE, GRAME GRILLENFAENGER STREICHT
 Geheim:tih pbyfjjhd qhvthd gfhdnhv, thv nvfeh, nvflh nvigghdxfhdnhv pjvhibyj

Anhand dieses einfachen Beispiels sollen einige grundlegende Begriffe der Kryptographie kurz erläutert werden. Die Abbildung 3 zeigt einen typischen Vorgang, wie er allgemein beim Verschlüsseln, bzw. Entschlüsseln auftritt.

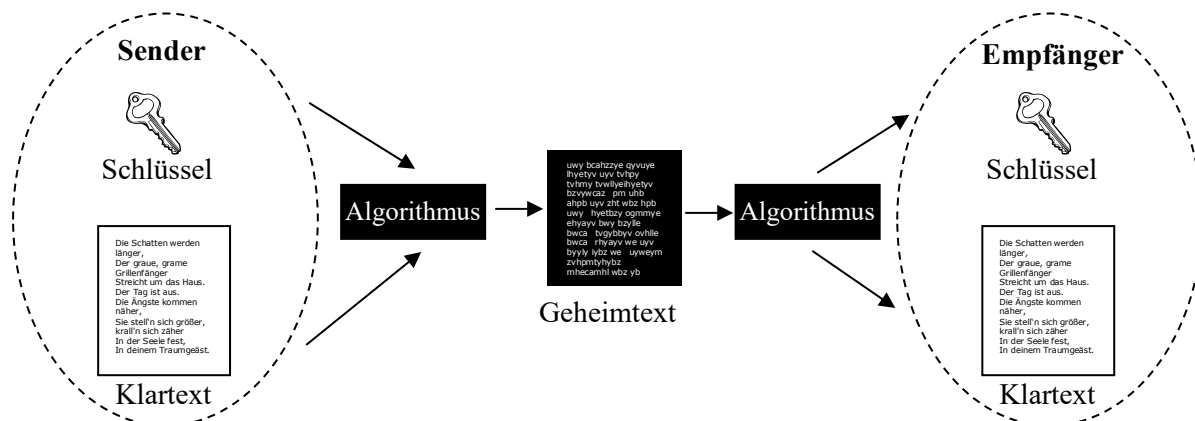


Abb. 3: Ein typischer Verschlüsselungsvorgang. Nicht der Algorithmus, sondern der Schlüssel entscheidet über die Sicherheit. Kennt ein Gegner den Algorithmus, so sollte es ihm dennoch unmöglich sein, den Klartext wieder herzustellen.

Der Algorithmus beschreibt die allgemeine Verschlüsselungsmethode (hier: bilde 26 Buchstabenpaare und verschlüssele jeweils den ersten Buchstaben durch den Zweiten). Der holländische Kryptograph Auguste Kerckhoffs stellte eine bis heute unbestrittene Maxime auf: „Die Sicherheit eines Kryptosystems darf nicht von der Geheimhaltung des Algorithmus abhängen. Die Sicherheit gründet sich nur auf die Geheimhaltung des Schlüssels“ (A. Kerckhoffs, zitiert aus Simon Singh, 2000). Der konkrete Schlüssel ist in unserem Fall durch die Tabelle von oben gegeben, durch welche die Verschlüsselung definiert ist. Dieser Schlüssel muss unbedingt geheim bleiben. Er dient zusammen mit dem Algorithmus dem Sender und dem Empfänger dazu, die Nachricht zu ver- bzw. entschlüsseln.

Die Maxime von Kerckhoffs ist recht einleuchtend: die verschiedenen Möglichkeiten, wie man eine Nachricht verschlüsseln kann sind jeweils recht begrenzt und neue Algorithmen werden sich recht schnell „herumsprechen“. Ein gutes Verschlüsselungsverfahren verdient diese Auszeichnung nur dann, wenn man die Nachricht nicht entschlüsseln kann, auch wenn man weiss nach welchem Prinzip sie verschlüsselt wurde. Für unser Beispiel scheint dies (zumindest nach unserem aktuellen Wissensstand) gewährleistet zu sein. Auch wenn man weiss, dass für jeden Buchstaben im Klartext einfach ein anderer Buchstabe hingeschrieben wurde, ist es anscheinend fast unmöglich, den Text zu entschlüsseln. Nur wenn man im Besitz des Schlüssels (der Tabelle) ist, ist die Entschlüsselung ein Kinderspiel.

Aufgabe Wie viele verschiedene Schlüssel gibt es bei der Monoalphabetischen Verschlüsselung?

Wie lange hätten 100 Codebrecher schlimmstenfalls, wenn jeder von ihnen einen Schlüssel in 20 Sekunden überprüfen könnte?

3.2 Caesar Verschlüsselung

Ein Spezialfall der monoalphabetischen Verschlüsselung für militärische Zwecke, wurde zum ersten Mal von Julius Caesar im Gallischen Krieg erwähnt. Caesar schrieb an Stelle des Klartextbuchstabens jeweils den Buchstaben, welcher im Alphabet drei Stellen weiter steht. Unten ist der Schlüssel für diese „Caesar-Verschiebung“ genannte Methode angegeben.

Klar	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
Geheim	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c

Klar: DIE SCHATTEN WERDEN LAENGER, DER GRAUE, GRAME GRILLENFAENGER STREICHT

Geheim: glh vfkdwvhq zhughq odhqjhu, ghu judxh, judph julioohqidhqjhu vwuhlfkw

Natürlich kann man die Verschiebung des Geheimalphabetes beliebig verändern. Bei der Caesar-Verschiebung bleibt jedoch die Reihenfolge der Buchstaben im Klartextalphabet und im Geheimalphabet gleich.

Aufgabe Wie viele verschiedene Schlüssel gibt es für die Caesar-Verschiebung?

Welche Schwachstelle offenbart sich bei diesem Verfahren (vergl. Kerckhoff's Maxime)?

3.3 Knacken monoalphabetischer Texte

Lange Zeit galt ein monoalphabetisch verschlüsselter Text als unknackbar, denn es war schlicht unmöglich, die riesige Anzahl von verschiedenen Schlüsseln durchzuprobieren, bis man zufällig auf den Richtigen stiess. Doch schliesslich entdeckte man in Arabien im neunten Jahrhundert eine Möglichkeit, wie man derartige Texte entschlüsseln konnte, während es in Europa noch einige hundert Jahre dauern sollte.

Al-Kindi, der als „Philosoph der Araber“ bekannte Gelehrte, welcher über 290 Bücher über jegliche Wissenschaften veröffentlichte, war der Erste, welcher die Gedankengänge aufzeichnete, die zum Entschlüsseln eines monoalphabetisch chiffrierten Textes nötig sind. Er stellte fest, dass im Arabischen die Buchstaben „a“ und „l“ sehr häufig vorkommen, während der Buchstabe „j“ zehn mal weniger häufig auftaucht. Ist ein Geheimtext genügend lang, so wird sich nach dem Gesetz der grossen Zahlen die Häufigkeit der einzelnen Buchstaben der allgemein in der jeweiligen Sprache

vorhandenen Häufigkeit annähern. So werden bei einem arabischen Text die zwei Symbole, welche am häufigsten vorkommen, mit grösster Wahrscheinlichkeit für die Buchstaben „a“ und „l“ stehen. Verfeinert man diese Überlegungen, so muss man nicht alle möglichen Schlüssel durchprobieren, sondern man kann mit Hilfe von logischen Schlussfolgerungen den Text knacken.

Dies soll anhand eines deutschen Textes exemplarisch durchgeführt werden. Dazu benötigt man eine Zusammenstellung der Häufigkeitsverteilung der einzelnen Buchstaben und der häufigsten Bigramme (Buchstabenpaare) der deutschen Sprache (Abbildung 4).

Buchstabe	Häufigkeit in %	Buchstabe	Häufigkeit in %	Buchstabenpaar	Häufigkeit in %
A	6,51	N	9,78	en	3,88
B	1,89	O	2,51	ch	2,75
C	3,06	P	0,79	de	2,00
D	5,08	Q	0,02	ei	1,88
E	17,4	R	7,00	in	1,67
F	1,66	S	7,27	er	3,75
G	3,01	T	6,15	te	2,26
H	4,76	U	4,35	nd	1,99
I	7,55	V	0,67	ie	1,79
J	0,27	W	1,89	es	1,52
K	1,21	X	0,03	:	:
L	3,44	Y	0,04		
M	2,53	Z	1,13		

Abb. 4: Relative Häufigkeit der einzelnen Buchstaben und ausgewählter Bigramme in Deutsch

Unserer Nachrichtenspionageabteilung ist es gelungen, den folgenden, verschlüsselten Text abzuhören. Wir gehen von der Vermutung aus, dass der Text in deutscher Sprache geschrieben ist und dass er monoalphabetisch verschlüsselt wurde.

„uwv bcahzye qvuyv lhyetyv uyv tvhpy tvhmy twvlliehyetyv bzvywcaz pm uhb ahpb uyv zht wbz hpb uwy hyetbzy ogmmye ehyayv bwy bzylle bwca tvgybbyv ovhille bwca rhyayv we uyv bylyly iybz we uyweym zvhpmtyhybz mhecamhl wbz yb nwb rpm heuyvyv piyv uyv ehcaz qwy ywe lwcazlgbyv zpeeyl ywe ewcaz yeuye qgllyeuyv beahcaz wca nwwet uwca upvca uwy ehcaz wca nwwet uwca upvca uwy vhpay byy wca nwwet uwca upvca uwy ehcaz wca nwwety uwca jge lpj ehca lyy wca nwe uywe lgzby wca nwe uywe mhee nwe uywey beaqybyzv lyae uwca he wca nwe uyv ivypeu uyv mwz uwv qhcaz wca nwwet uwca upvca uwy ehcaz hllyb yvbcaywez uwv beaqyvyv nyuvgalwcayv peu agiiepetblyyvyv mwz uyv upeoylaywz ogmmye hpb upeolyv rywz iyvey yvweeyvpetye uwy ehcaz qwbkyvz mwz zhpyeu rpetye bwy hlly bweu hpb up nwbz hllywe rpahpb mwz uyweyv bzpmmye jyvrywilpet peu uym oewbzyve wm khvoyzz peu hlb ywerwtym zvzgbz uhb qhvmy lwcaz uyv vhuwgb he uyweym nyzz wca nwwet uwca upvca uwy ehcaz lhß lgb jyvbpca rp bealhiye wca nwwet uwca bwcayv we uye ahiye uwv ohee ewcazb tybcayae qglibmhee peu ngyby iyve bweu epv ywe nlhyzyvvywtye jgvm iyebzyv uyv qweu we uye rgywtye wm ohbzehyehpm ywe ngybyv zvhpm uyvb ewcaz qhtz qwyuyvropgmmye nwb uyv eypy zht nytweez lhbb lgb wca ahlz uwca iybz wca oyee uye qyt hpb uym lhnvweza wca nwwet uwca upvca uwy ehcaz“

Als Erstes gehen wir den Text durch und zählen, wie oft jeder Buchstabe vorkommt. Es folgen also die absoluten Häufigkeiten der Buchstaben:

{a, 71}, {b, 63}, {c, 58}, {d, 0}, {e, 113}, {f, 0}, {g, 18}, {h, 60}, {i, 14}, {j, 5}, {k, 2}, {l, 40}, {m, 31}, {n, 23}, {o, 11}, {p, 42}, {q, 15}, {r, 10}, {s, 0}, {t, 29}, {u, 71}, {v, 77}, {w, 106}, {x, 1}, {y, 154}, {z, 58}

Da der Text aus insgesamt 1073 Buchstaben (ohne die Leerschläge) besteht, ergeben sich die folgenden prozentualen Häufigkeiten:

{a, 6.6%}, {b, 5.9%}, {c, 5.4%}, {d, 0%}, {e, 10.5%}, {f, 0%}, {g, 1.7%}, {h, 5.6%}, {i, 1.3%}, {j, 0.5%}, {k, 0.2%}, {l, 3.7%}, {m, 2.9%}, {n, 2.1%}, {o, 1.0%}, {p, 3.9%}, {q, 1.4%}, {r, 0.9%}, {s, 0%}, {t, 2.7%}, {u, 6.6%}, {v, 7.2%}, {w, 9.9%}, {x, 0.1%}, {y, 14.4%}, {z, 5.4%}

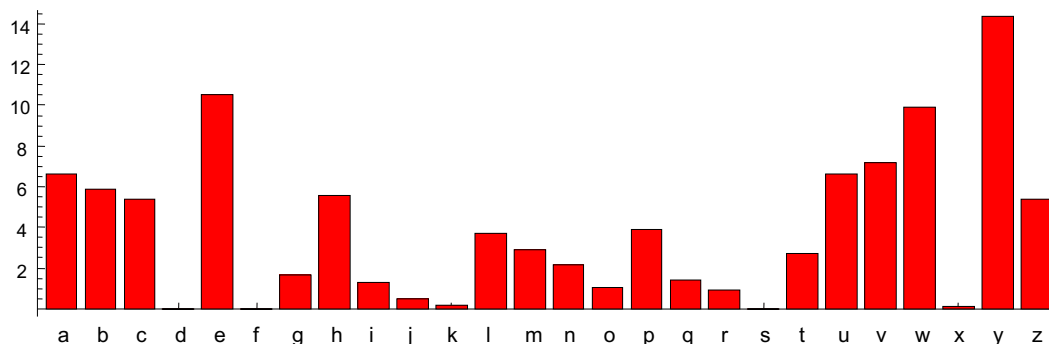


Abb. 5: Relative Häufigkeit der Buchstaben im verschlüsselten Text

In unserem Textbeispiel ist also der Buchstabe y mit Abstand am häufigsten vertreten. Vergleicht man die Häufigkeiten im Text und in der deutschen Sprache allgemein, so kann man davon ausgehen, dass der Buchstabe E durch y verschlüsselt wurde. Die nächsthäufigsten Buchstaben in unserem Geheimtext sind e (10,5%) und w (9,9%). Nach der Häufigkeitsverteilung der deutschen Sprache sind dafür N und I gute Kandidaten, aber auch R , S und T liegen dicht dahinter. Um sich beim e , welches für N stehen könnte abzusichern, kann man im Geheimtext die Häufigkeit des Bigramms ye bestimmen. Falls die Vermutung zutrifft, dass e für N steht, dann müsste ye relativ häufig vorkommen, denn das Bigramm EN ist das Häufigste der deutschen Sprache. Und tatsächlich - in unserem Text kommt das Bigramm ye 34-mal (ca 3,3%) vor – ein klares Anzeichen dafür, dass e für N steht. Wir versuchen somit, die drei häufigsten Buchstaben des Textes (y, e, w) durch die drei häufigsten Buchstaben in der deutschen Sprache zu ersetzen (E, N, I).

Betrachten wir den vierthäufigsten Buchstaben in unserem Geheimtext: v hat eine relative Häufigkeit von 7.2%. Er dürfte aus der Gruppe (A, R, S, T) kommen, da diese Buchstaben normalerweise mit einer relativen Häufigkeit von mehr als 6% auftreten. Im Text kommt ein Doppel- v vor, dies lässt den Buchstaben A als unwahrscheinlich erscheinen. Das Bigramm yv kommt insgesamt 42-Mal im Geheimtext vor (über 4%), dies deutet sehr stark auf das Bigramm ER hin. D.h. R wurde offensichtlich durch v verschlüsselt.

Wenden wir uns den nächsten Buchstaben zu: u und a haben beide eine Häufigkeit von 6.6%. Dafür kämen jetzt hauptsächlich S, T und A in Frage. Doch sowohl uu und aa kommen in unserem Text kaum vor! Dies spricht sehr gegen S und T , da SS und TT häufig sind. Hier ist es ziemlich schwierig, die beiden Buchstaben u und a zuzuordnen... offenbar ist u oder a ein Buchstabe der folgenden Gruppe (D, H, U). Hier fällt auf, dass der Klartextbuchstabe D sowohl häufig als Bigramm DE als auch als ND vorkommen müsste. Man muss also einen Geheimtextbuchstaben suchen, der häufig vor dem y (Geheimbuchstabe von E) und häufig nach dem e (Geheimtextbuchstabe von N) steht. Zudem muss der Buchstabe selbst recht häufig vorkommen, da D eine relative Häufigkeit von über 5% besitzt. Testet man die erwähnten Bigramme mit u und a ergibt sich: uy (27), eu (25), ay (8), ea (1). Ein recht eindeutiges Ergebnis: der Buchstabe D wurde offenbar durch u verschlüsselt.

Die beiden sehr häufigen Klartextbuchstaben S und T müssten jedoch im Geheimtext aufzuspüren sein. Beide kommen häufig doppelt vor. Zudem ist das Bigramm TE sehr häufig. Auf der Suche nach einem Bigramm, das oft mit y (dem Stellvertreter für E) vorkommt, stossen wir auf zy . Auch das Bigramm zz kommt recht häufig vor (5) was dafür spricht, dass T durch z verschlüsselt wurde.

Das Bigramm ca kommt im Geheimtext sage und schreibe 58-Mal vor. Von den häufigen Bigrammen in der Abbildung 4 ist jedoch CH das Einzige, welches keinen uns bereits bekannten Buchstaben enthält. Zudem kommt weder die cc -Doppelung (0) noch die aa -Doppelung (1) oft vor. Diese Fakten sprechen deutlich dafür, dass C in c und H in a verschlüsselt wurde.

Die bisher gefundenen Buchstabenpaare sowie der Beginn und der Schluss des teilweise entschlüsselten Textes sind somit (bereits entschlüsselte Buchstaben werden gross geschrieben):

$$y \rightarrow E, e \rightarrow N, w \rightarrow I, v \rightarrow R, u \rightarrow D, z \rightarrow T, c \rightarrow C, a \rightarrow H$$

DIEbCHhTTENqERDENlhENTERDERtRhpEtRhmEtRIllENihENTERbTREICHTpmDhbHhpbDER...
...lTDICHIEbTICHoENNNDENqEthpbDEmlhnXRINTHICHnRINTDICHDPCHDIENhCHT

Das fett markierte q deutet stark auf ein W hin, da dadurch das Wort „werden“ entsteht. Ebenso tendiert man beim fett markierten h auf ein A , da dadurch „die Nacht“ zum Vorschein kommt.

Diese Erkenntnis in den Text eingesetzt ergibt für die oben angegebenen Stellen:

DIESCHATTENWERDENlaengerDERGRAUEGRAMeGRILLENiAENGERSTREICHTUmDASHAUSDER...
...lTDICHIEStICHoENNNDENWEGAUSDEmlABxRINTHICHBRINGDICHDURCHDIENACHT

„Die Schatten werden laenger...“ und „...ich bring dich durch die Nacht“ ergeben die weiteren Ersetzungen $b \rightarrow S, t \rightarrow G, n \rightarrow B, p \rightarrow U$

Führt man diese Ersetzungen durch, so hat man den Ursprungstext praktisch wiederhergestellt, und es ist kein Problem, die restlichen Buchstaben aufzudröseln.

*Die Schatten werden länger,
Der graue, grame Grillenfänger
Streicht um das Haus.
Der Tag ist aus.
Die Ängste kommen näher,
Sie stell'n sich größer, krall'n sich zäher
In der Seele fest,
In deinem Traumgeäst.
Manchmal ist es bis zum anderen Ufer der Nacht
Wie ein lichtloser Tunnel, ein nicht enden wollender Schacht.
Ich bring dich durch die Nacht,
Ich bring dich durch die raue See
Ich bring dich durch die Nacht,
Ich bringe dich von Luv nach Lee.
Ich bin dein Lotse, ich bin dein Mann,
Bin deine Schwester, lehn dich an,
Ich bin der Freund, der mit dir wacht,
Ich bring dich durch die Nacht.
Alles erscheint dir schwerer,
Bedrohlicher und hoffnungsleerer.
Mit der Dunkelheit
Kommen aus dunkler Zeit
Ferne Erinnerungen,
Die Nacht wispert mit tausend Zungen:
"Sie alle sind aus,
Du bist allein zuhaus!"
Mit deiner stummen Verzweiflung und dem Knistern im Parkett
Und als einzigem Trost das warme Licht des Radios an deinem Bett.
Ich bring dich durch die Nacht...
Laß los, versuch zu schlafen.
Ich bring dich sicher in den Hafen.
Dir kann nichts gescheh'n,
Wolfsmann und böse Feen
Sind nur ein Blätterreigen
Vorm Fenster, der Wind in den Zweigen
Im Kastanienbaum,
Ein böser Traum,
Der's nicht wagt, wiederzukommen, bis der neue Tag beginnt.
Laß los, ich halt dich fest, ich kenn den Weg aus dem Labyrinth.
Ich bring dich durch die Nacht...*

Reinhard Mey, Ich bring' dich durch die Nacht

Die Tatsache, dass es sich um ein poetisches Lied handelt, hat die Suche nach den Ersetzungen nicht gerade vereinfacht. Durch die häufige Wiederholung von „Ich bring dich durch die..“ hat z.B. der Buchstabe *h* im Lied eine Häufigkeit von ca. 6,7%. Dies ist um etwa ein Drittel „zu viel“. Auch der Buchstabe *d* schlägt im Text mit 6,6% zu Buche, dies ist ebenfalls atypisch und hängt mit der oben erwähnten Wiederholung zusammen. Dafür sind die Buchstaben *s* (5,4%) und *t* (5,8%), die normalerweise häufiger als *d* und *h* sind, untervertreten.

Dennoch war es mit Hilfe der Auswertung der Bigramme und einigen logischen Überlegungen möglich, den Text zu entschlüsseln. Spezialisierte Kryptographen hatten für die Analyse von Meldungen viel weiterführende Tabellen zur Verfügung als wir sie hier verwendet haben. So hat man z.B. berücksichtigt, dass in militärischen Meldungen Artikel häufig weggelassen wurden und man hat für derartige Meldungen spezifische Häufigkeitstabellen erstellt.

3.4 Grenzen des Kryptoanalyseverfahrens von Al-Kindi

Das von Al-Kindi beschriebene Kryptoanalyseverfahren war ein sehr mächtiges Werkzeug, welches die Entschlüsselung von monoalphabetischen Texten ermöglichte. Doch auch dieses Verfahren hat gewisse Grenzen und Schwachstellen. Insbesondere wird davon ausgegangen, dass die Häufigkeit der Buchstaben im verschlüsselten Text in etwa die allgemeine Häufigkeitsverteilung der entsprechenden Sprache widerspiegelt. Doch dies muss nicht unbedingt der Fall sein. Ein krasses Beispiel liefert der Roman „La Disparition“ von Georges Perec, der im Jahre 1969 erschienen ist. Der Autor hat es

geschafft, den 200-seitigen Roman abzufassen, ohne ein einziges e zu verwenden. Eugen Helmlé ist es tatsächlich gelungen, das Buch ins Deutsche zu übersetzen. Es erschien unter dem Titel „Anton Voyls Fortgang“. Nachfolgend die ersten Sätze des deutschen Textes:

Kardinal, Rabbi und Admiral, als Führungstrio null und nichtig und darum völlig abhängig vom Ami-Trust, tat durch Rundfunk und Plakatanschlag kund, dass Nahrungsnot und damit Tod aufs Volk zukommt. Zunächst tat man das als Falschinformation ab. Das ist Propagandagift, sagt man. Doch bald schon ward spürbar, was man ursprünglich nicht glaubt. Das Volk griff zu Stock und zu Dolch. „Gib uns das täglich Brot“, hallts durchs Land, und „pfui auf das Patronat, auf Ordnung, Macht und Staat“. Konspiration ward ganz normal, Komplott üblich. Nachts sah man kaum noch Uniform. Angst hält Soldat und Polizist im Haus.

Die verschlüsselte Variante dieses Textes wäre nicht ganz einfach zu knacken, da der mit Abstand häufigste Buchstabe der deutschen Sprache e kein einziges Mal vorkommt. Ähnliche „Fallen“ können natürlich auch die „Verschlüssler“ anwenden. So können sie z.B. den Text absichtlich orthographisch falsch schreiben oder Füllworte, die sehr häufig vorkommen und dadurch die Entschlüsselung vereinfachen („und“, „die“ etc.), weglassen.

Dennoch würde ein erfahrener Kryptoanalysespezialist durch diesen Text nicht vor ein unüberwindliches Hindernis gestellt, denn ausser dieser extremen Abweichung von der Norm enthält der Text natürlich auch viele, für die deutsche Sprache typische Regelmässigkeiten, an welche man sich halten könnte.

3.5 Die Geheimschrift der Maria Stuart - Nomenklatoren

Im 14. Jahrhundert hielt schliesslich auch in Europa die Kryptoanalyse im Stile von Al-Kindi Einzug und monoalphabetisch verschlüsselte Texte galten nicht mehr als sicher. Man begann sich neue Tricks zu überlegen, durch welche man den Kryptoanalysten ein Bein stellen wollte. Man ersetzte beispielsweise jeden Buchstaben durch ein Symbol und fügte zusätzlich sogenannte Füller ein. Dabei handelte es sich um „leere Symbole“ die keine Bedeutung hatten, die jedoch die Häufigkeitsanalyse schwieriger gestalten sollten. Zusätzlich begann man, anstatt für jeden Buchstaben einen anderen Buchstaben oder ein Symbol zu schreiben, ganze Wörter durch ein Symbol abzukürzen. Man benutzte also eine Mischung aus *chiffrieren* (einzelne Buchstaben ersetzen) und *codieren* (ganze Wörter bzw. Sätze verschlüsseln). Dies erschwert natürlich die Kryptoanalyse, da man einem einzelnen Symbol nicht ansieht, ob es für ein reines Füllzeichen, für einen einzelnen Buchstaben, für ein ganzes Wort, oder für einen zusammenhängenden Ausdruck steht.

Solche kombinierte Verschlüsselungssysteme nennt man *Nomenklatoren*. Ein Nomenklator besteht aus einem Verschlüsselungsalphabet, evtl. einigen Füllzeichen und einer Liste von Codewörtern.

Doch es dauerte nicht lange und auch die Nomenklatoren konnten durch findige Kryptoanalytiker geknackt werden. Durch verfeinerte Verfahren, wie wir sie oben verwendet haben, konnte man das Verschlüsselungsalphabet rekonstruieren und die Codewörter aus dem Zusammenhang rekonstruieren. In extremer Weise kommt der Einfluss der Kryptoanalyse auf die europäische Geschichte im Schicksal von Maria Stuart zum Ausdruck.

Maria Stuart erblickte das Licht der Welt am 08.12.1542 als Tochter des schottischen Königs Jakob V. Zu dieser Zeit versuchte England, unter ihrem König Heinrich VIII, Schottland zu erobern und man hatte nur zwei Wochen vor der Geburt Maria Stuarts das schottische Heer bei Solway Moss vernichtend geschlagen. Ob dieser Niederlage erlitt Jakob einen schweren psychischen und körperlichen Zusammenbruch und er verstarb nur eine Woche nach der Geburt Marias.

Maria wurde schliesslich im Alter von lediglich neun Monaten zur Königin von Schottland gekürt. Dies liess den Eroberungsdruck der Engländer ein wenig abebben, denn es galt als unehrenhaft, ein Land, welches unter der Führung einer Kindkönigin stand anzugreifen. Stattdessen versuchte Heinrich eine Heirat zwischen Maria und seinem Sohn Edward zu arrangieren, um die zwei Länder per Ehe zusammenzuführen. Der schottische Hof erwog zunächst dieses Angebot, doch schliesslich wandte man sich nach Frankreich und entschloss, Maria mit Franz, dem Dauphin von Frankreich zu verheiraten. So wuchs Maria zu einem grossen Teil am französischen Hof auf, da man entschied, dass

4. Polyalphabetische Algorithmen

4.1 Anfänge polyalphabetischer Verschlüsselung

Das Schicksal der Maria Stuart zeigte in dramatischer Art und Weise, dass auch die zusätzlichen Sicherheitsmechanismen (Füllzeichen, Nomenklatoren, Doppelzeichen), einen gewieften Kryptoanalytiker nicht an der Entschlüsselung eines Textes hindern konnten. So suchte man nach einem neuen System der Verschlüsselung, welches ein grösseres Mass an Sicherheit bieten würde.

Die Schwachstelle jedes monoalphabetischen Systems wurde uns im Kapitel 3.3 deutlich vor Augen geführt: die Häufigkeitsverteilung der entsprechenden Sprache bleibt im Geheimtext relativ gut erhalten. Zudem ist auch die Tatsache, dass jeder Klartextbuchstabe immer durch denselben Geheimtextbuchstaben verschlüsselt wird gefährlich. Entschlüsselt man z.B. das Buchstabenpaar $a \rightarrow E$, so kann man im Geheimtext alle Vorkommen von a durch E ersetzen, da ja alle Buchstaben mit demselben Geheimalphabet verschlüsselt wurden.

Um diese Schwachstellen auszumerzen, wurden neue Algorithmen entwickelt, bei welchen nicht immer dasselbe Geheimtextalphabet für die Verschlüsselung verwendet wird. Die Pioniere auf diesem Gebiet waren Leon Battista Alberti (um 1470) und Johannes Trithemius, der in seinem Buch „Polygraphiae“ (1508) die sogenannte Trithemius-Tafel definierte („Tabula recta“, siehe Abbildung 7), welche alle 26 möglichen Verschiebungen des Alphabets untereinander auflistet.

Bemerkung

In der Zeit von Trithemius waren es üblicherweise nur 25 Buchstaben, da das j jeweils weggelassen wurde. Auf der Tafel von Trithemius fehlt ebenfalls der Buchstabe v , so dass sein Quadrat aus 24×24 Buchstaben bestand. Zudem war die Reihenfolge der Buchstaben am Schluss des Alphabets bei Trithemius „...stuxyzw“.

Eine erste Methode der polyalphabetischen Verschlüsselung besteht nun darin, bei der Chiffrierung des Klartextes in der ersten Zeile zu beginnen, um dann bei jedem Buchstaben eine Zeile nach unten zu verschieben, so dass der Reihe nach alle Geheimtextalphabete auf der Trithemius-Tafel verwendet werden.

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

Abb. 7: Trithemius-Tafel

Als Beispiel soll erneut die erste Zeile des Liedes von R. Mey (vgl. Kap. 3) verschlüsselt werden:

DIE SCHATTEN WERDEN LAENGER, DER GRAUE, GRAME GRILLENFAENGER
 djg vgmgbxn hgertd csxhba, bdr htdyj, myivo rdvzauextyicbp

Den Vorteil der polyalphabetischen Verschlüsselung erkennt man auf Anhieb. Während das erste E in ein g verschlüsselt wird, steht im Geheimtext anstelle des zweiten E 's ein n . Das dritte E ist durch q codiert worden u.s.w. Umgekehrt steht das erste x im Geheimtext für ein N , während das zweite x von einem E stammt. Die Häufigkeitsverteilung der Buchstaben wird durch dieses stete Wechseln des Geheimalphabets natürlich extrem verwischt. Dies bietet einem Kryptoanalytisten bei der Suche nach der Methode, die in Kapitel 3.3 beschrieben wird, nur wenig Anhaltspunkte.

Wie die Abbildung 8 auf der nächsten Seite zeigt, kommen alle Buchstaben im mit der Trithemius-Tafel verschlüsselten Lied von R. Mey in etwa gleich oft vor. Die relativen Häufigkeiten schwanken nur von 2.9% (q) bis 4.4% (t und u). Da man bei einer völligen Gleichverteilung der 26 Buchstaben des

Alphabets einen Erwartungswert von ca. 3.8% für jeden Buchstaben hat, kann man daraus natürlich keine Schlüsse ziehen.

Dennoch ist dieses simple Verschlüsselungsverfahren mit der Trithemiustafel nicht als sehr sicher einzustufen. Der Algorithmus besitzt keine, resp. nur einen Schlüssel. Weiss man also, mit welchem Algorithmus verschlüsselt wurde, so gestaltet sich die Dechiffrierung problemlos. Halten wir uns wieder an die Maxime von Kerckhoffs, so muss man sich für die Verschlüsselung etwas anderes einfallen lassen, denn die Sicherheit eines Kryptosystems soll ja gerade nicht dadurch beeinträchtigt werden, dass ein eventueller Gegner den Algorithmus kennt.

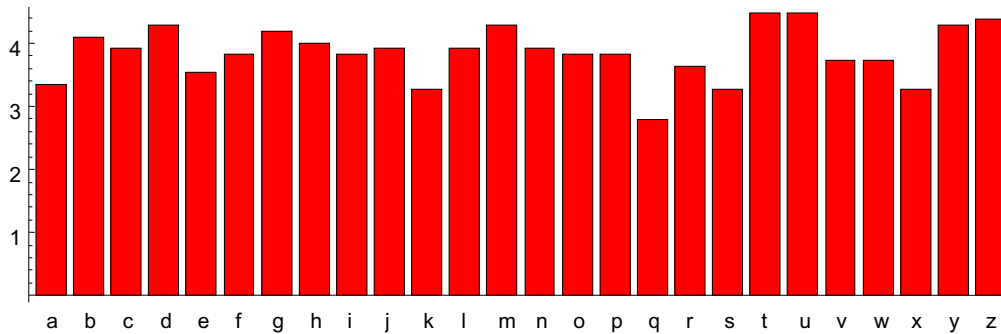


Abb. 8: Relative Häufigkeiten der Geheimbuchstaben für das Lied von R. Mey bei Verschlüsselung mit der Trithemiustafel (vergl. auch Abb.5 auf Seite 5)

Es ist nun jedoch ein Leichtes, aus den Überlegungen, die Trithemius angestellt hat ein Kryptosystem herzustellen, welches die Maxime von Kerckhoffs erfüllt. Anstatt der starren, deterministischen Trithemiustafel, kann man eine beliebige Menge von monoalphabetischen Verschlüsselungen wählen, unter welchen man stetig wechselt. Diese Menge muss natürlich auch dem Empfänger bekannt sein – sie bildet den Schlüssel des Systems. Als Beispiel nehmen wir vier verschiedene monoalphabetische Verschlüsselungen (Abb.8).

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

Abb. 8: Polyalphabetische Verschlüsselung mit vier Alphabeten

Diese sind nicht notwendigerweise aus der Familie der Caesarverschlüsselungen zu wählen, wie sie bei Trithemius vorgeschlagen werden. Die Verschlüsselung geschieht analog zur Trithemiustafel, d.h. man verschlüsselt den ersten Klartextbuchstaben mit der ersten Zeile der Geheimalphabete, den Zweiten mit der zweiten Zeile u.s.w.

DIE SCHATTEN WERDEN LAENGER, DER GRAUE, GRAME GRILLENFAENGER
 orx ejldyves zcjpgf ydgftxs, oeu onffg, ljdac tuqkyxtmfxtleu

4.2 Vigenère-Chiffre

Ein Spezialfall der polyalphabetischen Chiffrierung stellt die sogenannte Vigenère-Verschlüsselung dar. Der Name geht auf Blaise de Vigenère (1523-1596) zurück, welcher im Jahre 1580 das Werk „Traictè de Chiffres“ herausbrachte. Darin gab Vigenère unter anderem den Stand der Kryptographie seiner Zeit wieder. Er beschreibt dabei ein polyalphabetisches Verfahren, welches durch Giovanni Batista Belaso im Jahre 1553 erfunden wurde und welches als Vigenère-Chiffrierung in die Analen eingehen sollte.

Das Vigenère-Verfahren verwendet die Trithemiustafel (auch Vigenèrequadrat genannt) in Zusammenhang mit einem Schlüsselwort. Wir wählen beispielsweise das Schlüsselwort „Schluessel“, um das Verfahren zu veranschaulichen (Abb. 9).

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

Abb. 9: Relevante Zeilen der Trithemiustafel für das Schlüsselwort „Schluessel“

Um einen Text zu chiffrieren, wird das

Schlüsselwort wiederholt in eine erste Zeile geschrieben unter welche der Klartext zu stehen kommt. Jeder Buchstabe des Klartextes wird nun mit derjenigen Zeile der Trithemiustafel verschlüsselt, die als ersten Buchstaben den entsprechenden Buchstaben des Schlüssels hat:

Schlüssel	SCH LUESSELS CHLUES SELSCHL UES SELSC HLUES SELSCHLUESSELS
Klartext	DIE SCHATTEN WERDEN LAENGER, DER GRAUE, GRAME GRILLENFAENGER
Geheimtext	vkl dwlslxpf ylcxif depfilc, xij yvlmg, ncuqw yvtdnlyzewfkpj

Die Vorteile liegen auf der Hand: ein Schlüsselwort kann man sich leicht merken und dennoch ist die Maxime von Kerckhoffs offenbar erfüllt, denn die Auswahl an verschiedenen Schlüsseln ist immer noch immens, da jedes beliebige Wort als Schlüsselwort in Frage kommt.

4.3 Brechen polyalphabetisch verschlüsselter Texte

Lange Zeit galten die polyalphabetisch verschlüsselten Texte als „unknackbar“, so wurde z.B. die Vigenère-Verschlüsselung auch „le code indéchiffable“ genannt. In der Tat gestaltet sich die Suche nach einer Methode zum brechen eines polyalphabetischen Textes als sehr schwierig, denn wie wir oben gesehen haben, wird die Häufigkeitsverteilung durch die Verwendung mehrerer Geheimalphabete verwischt. Zudem wird nicht nur ein Klartextbuchstabe durch verschiedene Geheimbuchstaben dargestellt, auch ein und derselbe Geheimbuchstabe kann für verschiedene Klartextbuchstaben stehen. Auf den ersten Blick hat man kaum einen Anhaltspunkt, welcher zum knacken Verwendung finden könnte. Doch Charles Babbage (Abb. 10) verblüffte 1846 die Experten, indem es ihm gelang, polyalphabetische Texte durch Phantasie, Intuition und mit endloser Geduld bei der Suche nach wahrscheinlichen Worten im Text, zu entschlüsseln.



Abb. 10 Charles Babbage

4.3.1 Länge des Schlüsselwortes Bestimmen

Zehn Jahre nach Babbage war es F. W. Kasiski, der eine systematische Beschreibung für die Entschlüsselung polyalphabetisch chiffrierter Texte gab. Kasiski hielt sich an die einzige Konstante, sozusagen den einzigen Strohalm, an welchen man sich als Codeknacker halten kann: die Länge des Schlüsselwortes. Es sollen nun zwei Methoden zur Auffindung der Schlüssellänge dargestellt werden: die Parallelstellensuche nach Kasiski und eine sogenannte Koinzidenzindexmethode nach Friedman.

4.3.1.1 Parallelstellensuche nach Kasiski

Kasiski hatte gemerkt, dass man gegebenenfalls die Länge des Schlüsselwortes herausfinden kann, wenn man im Geheimtext Stellen mit gleichen Buchstabenfolgen aufsucht. Solche Parallelstellen können durch gleiche Buchstabenfolgen im Klartext entstehen, die durch dieselben Schlüsselbuchstaben chiffriert wurden. Natürlich besteht auch die Möglichkeit, dass diese Stellen zufällig entstehen. Als Veranschaulichung soll folgendes Beispiel dienen:

```

lehmannlehmannlehmannlehmannlehmannlehmannlehmannlehmannlehmannlehmann
ICHBRINGDICHDURCHDIENACHTICHBRINGDICHDURCHDIERAUHESEEICHBRINGDICHDURCHDIENACHTICHBRINGEDICH
tgonrvarhpohqhcgopiralgofipumvpzggvnlkgrpuomldahupwlqipumvpzggvnlkgrpuomlzapuemjtbevyklpipu

```

Es gibt einige Textfragmente, die mehrmals vorkommen. In unserem überschaubaren Beispiel haben wir auch Parallelstellen der Länge 2 gesucht. Normalerweise sind nur Stellen mit der Länge 3 oder mehr aufzusuchen, da Stellen mit der Länge 2 sehr oft zufällig zustande kommen. So wäre es durchaus möglich, dass das zweimalige go durch verschiedene Buchstabenfolgen im Klartext mit verschiedenen Schlüsselteilen erzeugt wurde. Es ist aber auch denkbar, dass es durch gleiche Folgen mit demselben Teil des Schlüssels chiffriert wurde (in diesem Fall war es CH, welches durch eh chiffriert wurde). Bei dem sehr langen Textfragment ipumvpzggvnlkgrpuoml ist es extrem unwahrscheinlich, dass es zufällig mit unterschiedlichen Text- resp. Schlüsselteilen entstand.

Die Methode von Kasiski besteht nun darin, den Abstand zwischen den gefundenen Parallelstellen zu bestimmen. Sind diese tatsächlich durch das Zusammentreffen von identischen Text- und Schlüssel-

stellen entstanden, so muss der **Abstand zwischen den Parallelstellen natürlich ein Vielfaches des Schlüssels sein**. Für unser Beispiel ergeben sich folgende Abstände:

„go“ \Leftrightarrow „go“ Abstand: 7
 „ipum...“ \Leftrightarrow „ipum...“ Abstand: 28
 „ipu“ \Leftrightarrow „ipu“ Abstand: 35
 „ipu“ \Leftrightarrow „ipu“ Abstand 63

Die Länge des Schlüssels müsste also als Faktor in 7, 28, 35 und 63 vorkommen. Für unser (durch die vielen Wiederholungen etwas unrealistisches) Beispiel deutet also alles auf eine Schlüssellänge von 7 hin, was auch stimmt („lehmann“).

Die Parallelstellensuche ist extrem mühsam, da man in peinlicher Kleinstarbeit zunächst die Parallelstellen aufsuchen muss, um danach die Abstände auszählen zu können. Zudem können „falsche“ Parallelstellen (d.h. Parallelstellen, die zufällig zustande kommen) die Auswertung und die Suche nach der Länge des Schlüsselwortes stören.

4.3.1.2 Kappa-Test nach Friedman

Verblüffenderweise ist es recht einfach, bei einem monoalphabetisch verschlüsselten Text herauszufinden, ob er in einer natürlichen Sprache geschrieben wurde, oder ob es sich beim Klartext einfach um eine zufällige Buchstabenfolge handelt. Ebenso kann man von zwei verschiedenen Texten T und T' theoretisch ziemlich schnell entscheiden, ob sie in derselben Sprache geschrieben sind. Dazu kann man einen Test verwenden, der von William F. Friedman um 1920 entwickelt wurde. Zu zwei gleich langen Texten T und T' definiert man eine Kennzahl $\text{Kappa}(T, T') = \kappa(T, T')$, welche die Übereinstimmung (Koinzidenz) der zwei Texte beschreibt. Dazu legt man die Texte übereinander und zählt die Zeichen, die zusammenfallen. Kappa ist der Wert der zusammenfallenden Zeichen dividiert durch die Länge des Textes und wird auch Koinzidenzindex genannt. Es folgt die Bestimmung des Koinzidenzindex Kappa für zwei Texte der Länge 475.

ANSWÄHRENDDESSENKANMANAUFEBAYDIEGURKENERSTEIGERNDIEDERLASERBEIMCRASHVERLORENHATTEWENNICHJEMANUNRUHEENDERLETTZTENWOCHENHABENHATTIANEINENTOTENPUNKTGEBRACHDERPRÄSIDENTZEIGTSICHZWARBEREITEINE
 NDEMSCHMERZZUGEFÜGTODERVERLETZTHABEWILLUNDWERDEICHMICHINALLERFORMENTSCHULDIGENSAGTEDANIELKÜBLB
 NTEILSEINERMACHTANEINEINUNABHÄNGIGENMINISTERPRÄSIDENTENABZUGEBENUNDSORASCHWITEMÖGLICHPARLAMENTSW
 ÖCKINDERDEUTSCHENBILDDASRTL MAGAZINPUNKT12ZEIGTEEINEVIDEObOTSCHAFTDIESEINVATERAUFGENOMMENHATTEE
 AHLBENABZUHALTENWILLABERBISZUMENDESEINESMANDATS2006IMAMTBLEIBENDIEOPPOSITIONDAGEGENHÄLTETISERNAN
 RVERSICHERTEDARINDASSEIHMGTGEHEDANIELISTAMFREITAGZUMZWEITENMALOPERIERTWORDENOPERATIONUNDNARK
 DERFORDERUNGNACHDEM RÜCKTRITTARISTIDESFESTSIEHATNICHTSZUVERLIERENARISTIDEJEDOCHALLESDERWEILEROB
 OSEVERLIEFENOHNENZWISCHENFÄLLEWITSEINARZTDANACHDERINTERESSIERTENÖFFENTLICHKEITERKLÄRTEWAHRSCHEI
 ERNEBWAFFNETEUNDMARODIERENDEREBELLENGRUPPENIMNORDENEINESTADTNACHDERANDERENINPOSTAUPRINCEDESSEN

Die beiden Texte haben eine Länge von 475 Zeichen, wobei die 34 hervorgehobenen Übereinstimmungen auftreten. Dies ergibt einen Koinzidenzindex $\kappa(T, T') = 34/475 = 7,16\%$.

Es zeigt sich, dass Texte aus verschiedenen Sprachen verschiedene Kappawerte besitzen. So gilt z.B. dass für zwei genug lange deutsche Texte das Kappa um den Wert $\kappa_d \approx 7,6\%$ liegen wird, während bei englischen Texten der Kappawert $\kappa_e \approx 6,6\%$ beträgt.

Fragen: In welchem Bereich liegt Kappa? Wann nimmt Kappa die Extremwerte an?

Wie gross ist der Erwartungswert κ_r von κ , wenn man zwei Texte nimmt, die aus einer Reihe von zufällig gewählten Buchstaben bestehen?

Wie kann man den Erwartungswert für κ der deutschen Sprache mit Hilfe der Abbildung 5 auf Seite 5 bestimmen?

Der Koinzidenzindex Kappa wird durch die Häufigkeitsverteilung der einzelnen Buchstaben der jeweiligen Sprache bestimmt. Wird für zwei gegebene Texte T und T' $\text{Kappa}(T, T')$ berechnet und werden die Texte dann mit demselben monoalphabetischen Verfahren verschlüsselt, so bleibt das Kappa erhalten, denn die Buchstaben der beiden Texte wurden ja entsprechend vertauscht, so dass die Übereinstimmungen immer noch an denselben Stellen auftreten.

Nimmt man einen polyalphabetisch verschlüsselten Text, der mit einem Schlüssel der Länge d chiffriert wurde, so wird der erste Buchstabe des Textes gleich verschlüsselt wie der $(d+1)$ -ste, ebenso der $(2d+1)$ -ste, der $(3d+1)$ -ste u.s.w. Genauso wurde der zweite Buchstabe gleich verschlüsselt wie der $(d+2)$ -te, der $(2d+2)$ -te u.s.w. Dieser Sachverhalt wurde bereits bei der Parallelstellensuche nach Kasiski verwendet.

Verschiebt man also den Geheimtext um d Stellen nach rechts und schreibt ihn unter den Geheimtext, so stehen untereinander Buchstaben, die durch dasselbe Alphabet (also monoalphabetisch) verschlüsselt wurden. Bei genug langen Texten wird also der Kappawert des Geheimtextes und des um d Stellen verschobenen Geheimtextes in etwa dem Kappawert entsprechen, den man für zwei beliebige Texte dieser Sprache erwarten kann. Genauer formuliert gilt folgender Sachverhalt:

Sei C ein polyalphabetisch verschlüsselter deutscher Geheimtext der Länge l , der mit einem Schlüssel der Länge d chiffriert wurde und $C^{(p)}$ der um p Stellen nach rechts verschobene Geheimtext. Dann gilt für den Erwartungswert von κ :

$$E(\kappa(C, C^{(d \cdot k)})) = E(\kappa(T, T')) = \kappa_d \quad \text{für zwei deutsche Texte } T \text{ und } T'$$

Ist jedoch die Verschiebung kein Vielfaches der Schlüssellänge, dann stehen untereinander Buchstaben, die durch unterschiedliche Alphabete verschlüsselt wurden, d.h. die Häufigkeitsverteilungen wurden komplett durchmischt. Deshalb ist für Kappa auch nicht ein Wert zu erwarten, der nahe an κ_d liegt. Viel mehr wird der Kappawert bei einer Verschiebung, die kein Vielfaches der Schlüssellänge ist, um den bei zufälligen (gleichverteilten) Buchstabenfolgen zu erwartenden Wert von $\kappa_r = 1/26 \approx 3,8\%$ liegen.

Somit sind die Grundlagen für den Kappatest von Friedman geschaffen. Man berechnet $\kappa(C, C^{(p)})$ für verschiedene Verschiebungen $p=1,2,3,4,5,6,\dots$ des Geheimtextes und vergleicht diese Werte mit dem zu erwartenden Kappawert κ_d von ca. 7,6%. Hat das Schlüsselwort die Länge 7 (wie bei „lehmann“), so müssten $\kappa(C, C^{(7)})$, $\kappa(C, C^{(14)})$, $\kappa(C, C^{(21)})$ u.s.w. um den Wert 7,6% schwanken, während alle anderen Kappawerte sich um den Wert 3,8% gruppieren sollten. Dies soll an einem Beispiel verdeutlicht werden.

Gegeben sei der nachfolgende Geheimtext C der Länge $l=1073$:

```
niwauridtwvoozneftsovqejlwbobammybiweyzavtonxiwxoorkbjjoqmhlcenichscknmbtsoacbk
uklaoniyalosymemfxiohwzksmctwtwdxasczojymcswzcbivlfaampjawpwbqxdwzkonvexmkdqxd
wgfoudrscegmkekbekvmheidsadekjachemsvozomnwnblorfiurbgiwmaxtsczbdaorlcfxmvea
vfskrtwvovogodtwlorkkzkkrtakzlsnylampnujknqonskzqdmhtzaxoniupvezmhvqwbiehwa
woqmhtzaxoniupvezmhvqwximhlqurjbi fowngmhnwfvcfnskzvmoiuptsvneavdybceakzlxqdwqf
wixntqfnmsnwaureoslmjvmrnvqurixiuptsvnejnjoexdvmjwqddazokkrtakzlsnylampnujkn
qonskzdivlwawbamhwqfdlsrkkzgmbejjwnzyhdqurmbuflzypnnyvctoemjwqddwzveuedpwsb
uoewxiesvcfutorrmadnorfmwbqxnwzmxoonvqwximhleacxorluadbkukmfhnyfmcqoadtwcqx
dscknclikbsvtoifhmrieseqlnmsnwnzkdcmwvnozjwwqvxvcgmvvnmwkwfkdmbnauhkzuelbmxlkl
kmaxhsgwulbwctvikgibmwtampddwajklsokifnmsnwutobdiuptbqxgvqurlerupvsmxauplviglg
anozcuupreamhdixovsczjjsvqdakzcgmhwxaxlonzixovnijssxviuplcoosupwrvgodnkwxnmv
vlwoswnwovcflfezoifjdkmdtwzjoqqefdgbupefaloznejeaxlsnvmfjeoiymfsuuaakbsxqontim
wmsntwcmvtjimwlorkvampdwsolgqodwzresyemflqcdwzfoctotsotoosnfbdkaclgaampradbvs
krfwalskrkwfnmxwwoseaneetslgbifbzskrbjqqfqlsczlbmkrdamfkrt
```

Zunächst geht es darum, herauszufinden nach welchem Algorithmus der Text verschlüsselt wurde. Nehmen wir an der Text sei monoalphabetisch verschlüsselt. Dann müsste sich die Häufigkeitsverteilung der deutschen Sprache in den Häufigkeiten der Buchstaben einigermassen wieder spiegeln. Dies ist jedoch nicht der Fall - die Häufigkeiten der Buchstaben tendieren eher zu einer Gleichverteilung (Abb. 11).

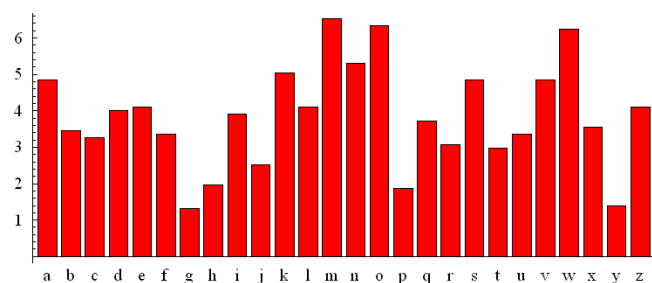


Abb. 11: Häufigkeitsverteilung des Geheimtextes

Gehen wir also davon aus, dass der Text durch ein Vigenère-Verfahren verschlüsselt wurde. Der Friedman-Test besteht darin, $\kappa(C, C^{(p)})$ für $p=1,2,3,4,5,6,7,\dots$ zu bestimmen.

Die Abbildung 12 zeigt die Kappawerte für $p \in \{1,2,3...60\}$. Während sich die meisten Kappawerte um $\kappa_r \approx 3,8\%$ gruppieren, liegen einige nahe bei $\kappa_d \approx 7,6\%$. Auffällig ist auch, dass der Abstand dieser Werte konstant bleibt: es handelt sich um die Kappawerte für $C^{(7)}$, $C^{(14)}$, $C^{(21)}$, $C^{(28)}$, $C^{(35)}$ u.s.w. Es ist somit offensichtlich, dass der Text polyalphabetisch mit einem Schlüssel der Länge 7 chiffriert wurde. Diese Information ist für das Knacken des Geheimtextes von essenzieller Wichtigkeit, denn nun kann der Text wie erwähnt in sieben monoalphabetisch chiffrierte Gruppen aufgeteilt und nach dem Verfahren, welches in Kapitel 3.3 beschrieben wird, entziffert werden.

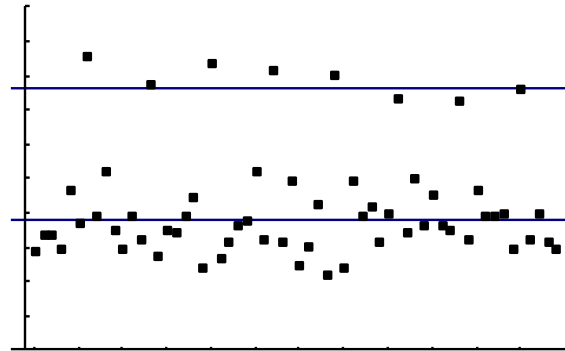


Abb.12: $Kappa(C,C(p))$ für verschiedene p -Werte

4.3.2 Knacken des Codes

Nachdem man die Länge des Schlüssels herausgefunden hat, kann man den Text in Gruppen einteilen, welche (jede für sich genommen) monoalphabetisch chiffriert wurden. Da unser Beispieltext einen Schlüssel der Länge $l=7$ hatte, erhalten wir die folgenden sieben Gruppen (b_i steht für den i -ten Buchstaben des Textes) $G_1=(b_1,b_8,b_{15},b_{22},b_{29},...)$, $G_2=(b_2,b_9,b_{16},b_{23},b_{30},...)$,..., $G_7=(b_7,b_{14},b_{21},b_{28},b_{35},...)$.

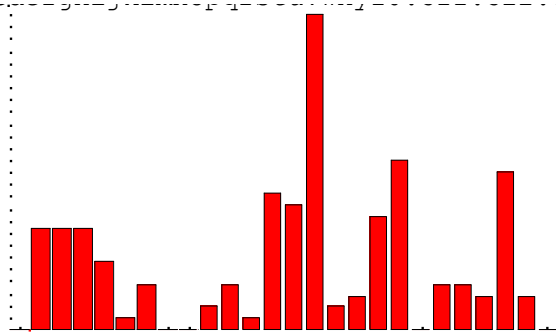


Abb.13: Häufigkeitsverteilung von G_1

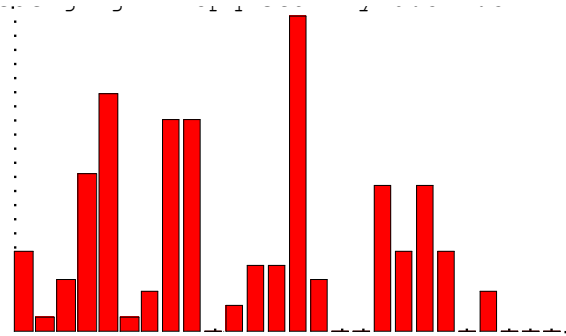


Abb.14: Häufigkeitsverteilung von G_2

Da es sich bei den einzelnen Gruppen um Caesar-Verschlüsselungen handelt, müssen sich die „Berge und Täler“ der Häufigkeitsverteilung der deutschen Sprache in verschobener Form in den Häufigkeitsverteilungen der Gruppen (siehe Abb. 13 und 14) wiederfinden. Charakteristisch für die Häufigkeitsverteilung der deutschen Sprache ist nicht nur, dass der Buchstabe e klar am Häufigsten vorkommt, sondern es gibt auch ein ziemlich „langes Tal“ (v, w, x, y, z kommen alle sehr selten vor), welches nach einem „breiten Berg“ (r, s, t, u sind recht häufig) steht (vergleiche Abb. 4). Dabei ist die Spitze mit dem e fünf Stellen nach dem langen Tal (z) zu erwarten.

Vergleicht man das eben Gesagte mit der Abbildung 13, so scheint die Spitze beim Buchstaben o tatsächlich dem e zu entsprechen, denn fünf bis neun Buchstaben vor dem o ($f-j$) befindet sich ein Tal, welches den Buchstaben und wieder davor (b, c, d, e) ein „breiter Berg“. Wenn das o dem e entspricht, dann entspricht das a dem k , denn a steht vier Buchstaben vor e . *Das Schlüsselwort beginnt also mit k .*

Dieselben Überlegungen kann man bei G_2 anstellen. Erster Kandidat für e wäre n , aber dies ist nicht sehr wahrscheinlich, denn fünf bis neun Buchstaben vor dem n ist kein Tal sichtbar. Zweiter Kandidat für e ist e selbst, da er am zweithäufigsten vorkommt. Und tatsächlich, von v bis z ist tatsächlich das zu erwartende Tal sichtbar, und bei r, s, t, u zeichnet sich der breite Berg ab. *Der zweite Buchstabe des Schlüsselwortes ist somit a .*

Dieselben Überlegungen kann man noch für G_3, G_4, G_5, G_6 und G_7 anstellen und man findet das Schlüsselwort „Kasiski“. Dadurch kann der Text problemlos entschlüsselt werden. Es erscheint wieder unser bereits bekanntes Lied von Reinhard Mey (siehe Seite 7).

5. Schlüsseltausch & Asymmetrische Verschlüsselung

Die bisher vorgestellten Verschlüsselungsverfahren waren alle symmetrisch. Die zwei Parteien, welche ein Geheimnis übertragen möchten (in der Kryptographie nennt man sie oft Alice und Bob), einigen sich auf *einen* geheimen Schlüssel, welcher für das Verschlüsseln *und* das Entschlüsseln zuständig ist. Dieses Verfahren bringt jedoch einige Probleme mit sich. Insbesondere muss dieser Schlüssel unter den Parteien ausgetauscht werden. Doch wie ist dieser Austausch durchzuführen? Überträgt man den Schlüssel über denselben Kanal wie die Nachricht selbst und vertraut darauf, dass der Kanal nicht abgehört wird, so könnte man ebenso gut die Nachricht direkt übermitteln. Auch verschlüsseln lässt sich der Schlüssel nicht ohne weiteres, denn man müsste ja für die Verschlüsselung des Schlüssels einen Schlüssel vereinbaren u.s.w. An dieser Stelle erreicht man ein typisches mathematisches Paradoxon: es scheint unmöglich zu sein, dass Alice und Bob ein Geheimnis (einen Schlüssel) erzeugen können, so dass es einem Feind (der jedes Wort mithört) nicht möglich ist, den Schlüssel herauszukriegen.

Über zweitausend Jahre lang galt das oben beschriebene Problem der sicheren Schlüsselverteilung als unlösbar. Man dachte, dass es unmöglich ist, einen Schlüssel von Alice zu Bob zu „transportieren“, ohne dass Eve, die jedes Wort mithört, den Schlüssel auch erfahren kann. Doch in den Siebzigerjahren gab es einige verwegene Kryptographen, die daran glaubten, dass es eine Lösung für das Problem geben müsste. Zwei von ihnen waren Whitfield Diffie und Martin Hellman. Ihre leidenschaftliche Suche nach einer Lösung für das Schlüsselproblem führte die beiden zusammen und die Ergebnisse ihrer gemeinsamen Bemühungen sollten schliesslich in die Geschichte der modernen Kryptographie eingehen.

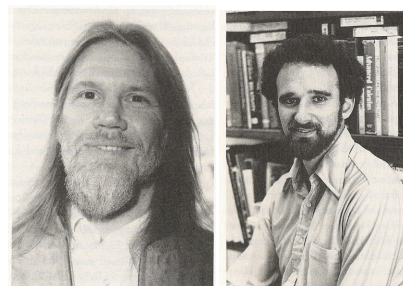


Abb. 15: Whitfield Diffie und Martin Hellman

Es waren nicht verwegene, komplexe mathematische Berechnungen, welche die zwei glauben liess, dass es eine Lösung für das Problem geben könnte, sondern zwei schlichte aber interessante Gedankenexperimente. Angenommen Alice will mit Bob ein Geheimnis erzeugen (einen Schlüssel), doch beide wissen, dass jegliche Kommunikation durch Eve abgehört wird. Sie könnten prinzipiell folgendermassen vorgehen:

1. Alice steckt das Geheimnis in eine Truhe, schliesst diese mit einem Schloss ab und schickt sie an Bob. Eve sieht die Truhe zwar, kann sie jedoch nicht öffnen. Auch Bob ist es nicht möglich die Truhe zu öffnen. Er nimmt sein persönliches Schloss, verschliesst die Truhe damit noch einmal ab und schickt die Truhe zurück an Alice. Unterwegs sieht Eve die Truhe mit den zwei Schlössern, doch sie kann sie natürlich nicht öffnen. Bei Alice angekommen entnimmt diese ihr Schloss von der Truhe und schickt sie erneut an Bob. Dieser kann nun sein Schloss entfernen und den Inhalt bestaunen.
2. Angenommen der Schlüssel bestehe aus einer bestimmten Farbmischung. Wie können Alice und Bob eine gemeinsame Farbe mischen, ohne dass Eve dieselbe Mischung herstellen kann?

Man geht davon aus, dass alle drei einen Dreiliterkanister mit einem Liter gelber Farbe haben. Nun fügt Bob seinem Kanister einen Liter einer beliebigen Farbe hinzu (z.B. Königsblau) und Alice tut dasselbe mit ihrem Kanister (z.B. Purpurrot). Nun schicken sich die beiden ihre Kanister zu.

Eve sieht die zwei Kanister mit den Farbmischungen, doch sie kann die Mischungen nicht „trennen“, d.h. sie kann nicht herausfinden, welchen Farbton Bob resp. Alice hinzugefügt haben. Auch Alice und Bob können nach Erhalt der Kanister nicht herausfinden, welche Farbe die jeweils andere Partei hinzugefügt hat, doch wenn sie nun den Kanister, den sie erhalten haben mit ihrer persönlichen Farbe auffüllen (Königsblau für Bob und Purpurrot für Alice), dann erhalten beide dieselbe Farbe, denn in beiden Kanistern sind nun jeweils ein Liter gelbe, königsblaue und purpurrote Farbe drin. Alice und Bob haben also eine gemeinsame Farbe generiert, ohne dass es Eve möglich ist, diese Farbe herauszufinden.