

Vorlesungsskript
Einführung in die Kryptologie
Wintersemester 2011/12

Prof. Dr. Johannes Köbler
Humboldt-Universität zu Berlin
Lehrstuhl Komplexität und Kryptografie

15. November 2011

Inhaltsverzeichnis

1	Klassische Verfahren	1
1.1	Einführung	1
1.2	Kryptosysteme	1
1.3	Die affine Chiffre	3
1.4	Die Hill-Chiffre	11
1.5	Die Vigenère-Chiffre und andere Stromsysteme	13
1.6	Der One-Time-Pad	15
1.7	Klassifikation von Kryptosystemen	16
1.8	Realisierung von Blocktranspositionen und einfachen Substitutionen	24
2	Kryptoanalyse der klassischen Verfahren	26
2.1	Klassifikation von Angriffen gegen Kryptosysteme	26
2.2	Kryptoanalyse von einfachen Substitutionschiffren	27
2.3	Kryptoanalyse von Blocktranspositionen	30
2.4	Kryptoanalyse von polygrafischen Chiffren	32
2.5	Kryptoanalyse von polyalphabetischen Chiffren	33
3	Sicherheit von Kryptosystemen	40
3.1	Informationstheoretische Sicherheit	40
3.2	Weitere Sicherheitsbegriffe	49

1 Klassische Verfahren

1.1 Einführung

Kryptosysteme (Verschlüsselungsverfahren) dienen der Geheimhaltung von Nachrichten bzw. Daten. Hierzu gibt es auch andere Methoden wie z.B.

Physikalische Maßnahmen: Tresor etc.

Organisatorische Maßnahmen: einsamer Waldspaziergang etc.

Steganografische Maßnahmen: unsichtbare Tinte etc.

Andererseits können durch kryptografische Verfahren weitere **Schutzziele** realisiert werden.

- *Vertraulichkeit*
 - Geheimhaltung
 - Anonymität (z.B. Mobiltelefon)
 - Unbeobachtbarkeit (von Transaktionen)
- *Integrität*
 - von Nachrichten und Daten
- *Zurechenbarkeit*
 - Authentikation
 - Unabstreitbarkeit
 - Identifizierung
- *Verfügbarkeit*
 - von Daten
 - von Rechenressourcen
 - von Informationsdienstleistungen

In das Umfeld der Kryptografie fallen auch die folgenden Begriffe.

Kryptografie: Lehre von der Geheimhaltung von Informationen durch die Verschlüsselung von Daten. Im weiteren Sinne: Wissenschaft von der Übermittlung, Speicherung und Verarbeitung von Daten in einer von potentiellen Gegnern bedrohten Umgebung.

Kryptoanalysis: Erforschung der Methoden eines unbefugten Angriffs gegen ein Kryptoverfahren (Zweck: Vereitelung der mit seinem Einsatz verfolgten Ziele)

Kryptoanalyse: Analyse eines Kryptoverfahrens zum Zweck der Bewertung seiner kryptografischen Stärken bzw. Schwächen.

Kryptologie: Wissenschaft vom Entwurf, der Anwendung und der Analyse von kryptografischen Verfahren (umfasst Kryptografie und Kryptoanalyse).

1.2 Kryptosysteme

Es ist wichtig, Kryptosysteme von Codesystemen zu unterscheiden.

Codesysteme

- operieren auf semantischen Einheiten,
- starre Festlegung, welche Zeichenfolge wie zu ersetzen ist.

Beispiel 1 (Ausschnitt aus einem Codebuch der deutschen Luftwaffe).

xve	<i>Bis auf weiteres Wettermeldung gemäß Funkbefehl testen</i>
yde	<i>Frage</i>
sLk	<i>Befehl</i>
fin	<i>beendet</i>
eom	<i>eigene Maschinen</i>

◁

Kryptosysteme

- operieren auf syntaktischen Einheiten,
- flexibler Mechanismus durch Schlüsselvereinbarung

Definition 2 (Alphabet). Ein **Alphabet** $A = \{a_0, \dots, a_{m-1}\}$ ist eine geordnete endliche Menge von **Zeichen** a_i . Eine Folge $x = x_1 \dots x_n \in A^n$ heißt **Wort** (der **Länge** n). Die Menge aller Wörter über dem Alphabet A ist $A^* = \bigcup_{n \geq 0} A^n$.

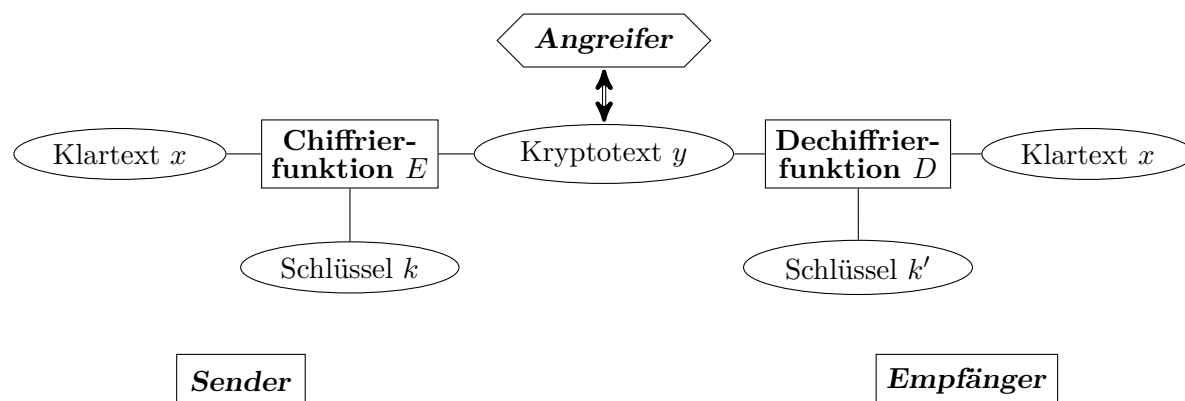
Beispiel 3. Das **lateinische Alphabet** A_{lat} enthält die 26 Buchstaben **A, ..., Z**. Bei der Abfassung von Klartexten wurde meist auf den Gebrauch von Interpunktions- und Leerzeichen sowie auf Groß- und Kleinschreibung verzichtet (\leadsto Verringerung der Redundanz im Klartext). ◁

Definition 4 (Kryptosystem). Ein **Kryptosystem** wird durch folgende Komponenten beschrieben:

- A , das **Klartextalphabet**,
- B , das **Kryptotextalphabet**,
- K , der **Schlüsselraum** (key space),
- $M \subseteq A^*$, der **Klartextraum** (message space),
- $C \subseteq B^*$, der **Kryptotextraum** (ciphertext space),
- $E : K \times M \rightarrow C$, die **Verschlüsselungsfunktion** (encryption function),
- $D : K \times C \rightarrow M$, die **Entschlüsselungsfunktion** (decryption function) und
- $S \subseteq K \times K$, eine Menge von Schlüsselpaaren (k, k') mit der Eigenschaft, dass für jeden Klartext $x \in M$ folgende Beziehung gilt:

$$D(k', E(k, x)) = x \tag{1.1}$$

Bei symmetrischen Kryptosystemen ist $S = \{(k, k) \mid k \in K\}$, weshalb wir in diesem Fall auf die Angabe von S verzichten können.



Zu jedem Schlüssel $k \in K$ korrespondiert also eine **Chiffrierfunktion** $E_k : x \mapsto E(k, x)$ und eine **Dechiffrierfunktion** $D_k : y \mapsto D(k, y)$. Die Gesamtheit dieser Abbildungen wird auch **Chiffre** (englisch *cipher*) genannt. (Daneben wird der Begriff „Chiffre“ auch als Bezeichnung für einzelne Kryptotextzeichen oder kleinere Kryptotextsequenzen verwendet.)

Lemma 5. Für jedes Paar $(k, k') \in S$ ist die Chiffrierfunktion E_k injektiv.

Beweis. Angenommen, für zwei unterschiedliche Klartexte $x_1 \neq x_2$ ist $E(k, x_1) = E(k, x_2)$. Dann folgt

$$D(k', E(k, x_1)) = D(k', E(k, x_2)) \stackrel{(1.1)}{=} x_2 \neq x_1,$$

im Widerspruch zu (1.1). □

1.3 Die affine Chiffre

Die Modularithmetik erlaubt es uns, das Klartextalphabet mit einer Addition und Multiplikation auszustatten.

Definition 6 (teilt-Relation, modulare Kongruenz). Seien a, b, m ganze Zahlen mit $m \geq 1$. Die Zahl a **teilt** b (kurz: $a|b$), falls ein $d \in \mathbb{Z}$ existiert mit $b = ad$. Teilt m die Differenz $a - b$, so schreiben wir hierfür

$$a \equiv_m b$$

(in Worten: a ist **kongruent** zu b modulo m). Weiterhin bezeichne

$$a \bmod m = \min\{a - dm \geq 0 \mid d \in \mathbb{Z}\}$$

den bei der Ganzzahldivision von a durch m auftretenden **Rest**, also diejenige ganze Zahl $r \in \{0, \dots, m-1\}$, für die eine ganze Zahl $d \in \mathbb{Z}$ existiert mit $a = dm + r$.

Die auf \mathbb{Z} definierten Operationen

$$a \oplus_m b := (a + b) \bmod m$$

und

$$a \odot_m b := ab \bmod m.$$

Tabelle 1.1: Werte der additiven Chiffrierfunktion ROT13 (Schlüssel $k = 13$).

x	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
$E(13, x)$	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

sind abgeschlossen auf $\mathbb{Z}_m = \{0, \dots, m-1\}$ und bilden auf dieser Menge einen kommutativen Ring mit Einselement, den sogenannten **Restklassenring** modulo m . Für $a \oplus_m -b$ schreiben wir auch $a \ominus_m b$.

Durch Identifikation der Buchstaben a_i mit ihren Indizes können wir die auf \mathbb{Z}_m definierten Rechenoperationen auf Buchstaben übertragen.

Definition 7 (Buchstabenrechnung). Sei $A = \{a_0, \dots, a_{m-1}\}$ ein Alphabet. Für Indizes $i, j \in \{0, \dots, m-1\}$ und eine ganze Zahl $z \in \mathbb{Z}$ ist

$$\begin{aligned} a_i + a_j &= a_{i \oplus_m j}, & a_i - a_j &= a_{i \ominus_m j}, & a_i a_j &= a_{i \odot_m j}, \\ a_i + z &= a_{i \oplus_m z}, & a_i - z &= a_{i \ominus_m z}, & z a_j &= a_{z \odot_m j}. \end{aligned}$$

Mit Hilfe dieser Notation lässt sich die Verschiebechiffre, die auch als additive Chiffre bezeichnet wird, leicht beschreiben.

Definition 8 (additive Chiffre). Bei der **additiven Chiffre** ist $A = B = M = C$ ein beliebiges Alphabet mit $m := \|A\| > 1$ und $K = \{1, \dots, m-1\}$. Für $k \in K$, $x \in M$ und $y \in C$ gilt

$$E(k, x) = x + k \quad \text{und} \quad D(c, y) = y - k.$$

Im Fall des lateinischen Alphabets führt der Schlüssel $k = 13$ auf eine interessante Chiffrierfunktion, die in UNIX-Umgebungen auch unter der Bezeichnung ROT13 bekannt ist (siehe Tabelle 1.3). Natürlich kann mit dieser Substitution nicht ernsthaft die Vertraulichkeit von Nachrichten geschützt werden. Vielmehr soll durch sie ein unbeabsichtigtes Mitlesen – etwa von Rätsellösungen – verhindert werden.

ROT13 ist eine **involutorische** – also zu sich selbst inverse – Abbildung, d.h. für alle $x \in A$ gilt

$$\text{ROT13}(\text{ROT13}(x)) = x.$$

Da ROT13 zudem keinen Buchstaben auf sich selbst abbildet, ist sie sogar eine echt involutorische Abbildung.

Die Buchstabenrechnung legt folgende Modifikation der Caesar-Chiffre nahe: Anstatt auf jeden Klartextbuchstaben den Schlüsselwert k zu addieren, können wir die Klartextbuchstaben auch mit k multiplizieren. Allerdings erhalten wir hierbei nicht für jeden Wert von k eine injektive Chiffrierfunktion. So bildet etwa die Funktion $g : A_{\text{lat}} \rightarrow A_{\text{lat}}$ mit $g(x) = 2x$ sowohl **A** als auch **N** auf den Buchstaben $g(\mathbf{A}) = g(\mathbf{N}) = \mathbf{A}$ ab. Um die vom Schlüsselwert k zu erfüllende Bedingung angeben zu können, führen wir folgende Begriffe ein.

Definition 9 (ggT, kgV, teilerfremd). Seien $a, b \in \mathbb{Z}$. Für $(a, b) \neq (0, 0)$ ist

$$\text{ggT}(a, b) = \max\{d \in \mathbb{Z} \mid d \text{ teilt die beiden Zahlen } a \text{ und } b\}$$

der **größte gemeinsame Teiler** von a und b . Für $a \neq 0, b \neq 0$ ist

$$\text{kgV}(a, b) = \min\{d \in \mathbb{Z} \mid d \geq 1 \text{ und die beiden Zahlen } a \text{ und } b \text{ teilen } d\}$$

das **kleinste gemeinsame Vielfache** von a und b . Ist $\text{ggT}(a, b) = 1$, so nennt man a und b **teilerfremd** oder man sagt, a ist **relativ prim** zu b .

Euklidischer Algorithmus: Der größte gemeinsame Teiler zweier Zahlen a und b lässt sich wie folgt bestimmen.

O. B. d. A. sei $a > b > 0$. Bestimme die natürlichen Zahlen (durch Division mit Rest):

$$r_0 = a > r_1 = b > r_2 > \dots > r_n > r_{n+1} = 0 \quad \text{und} \quad d_2, d_3, \dots, d_{n+1}$$

mit

$$r_{i-1} = d_{i+1}r_i + r_{i+1} \quad \text{für} \quad i = 1, \dots, n.*$$

Hierzu sind n Divisionsschritte erforderlich. Wegen

$$\text{ggT}(r_{i-1}, r_i) = \text{ggT}(r_i, \underbrace{r_{i-1} - d_{i+1}r_i}_{r_{i+1}})$$

folgt $\text{ggT}(a, b) = \text{ggT}(r_n, r_{n+1}) = r_n$.

Beispiel 10. Für $a = 693$ und $b = 147$ erhalten wir

i	r_{i-1}	$=$	d_{i+1}	\cdot	r_i	$+$	r_{i+1}
1	693	=	4	·	147	+	105
2	147	=	1	·	105	+	42
3	105	=	2	·	42	+	21
4	42	=	2	·	21	+	0

und damit $\text{ggT}(693, 147) = r_4 = 21$. ◁

Der Euklidische Algorithmus lässt sich sowohl iterativ als auch rekursiv implementieren.

Prozedur Euklid_{it}(a, b)

```

1  repeat
2     $r := a \bmod b$ 
3     $a := b$ 
4     $b := r$ 
5  until  $r = 0$ 
6  return  $a$ 

```

Prozedur Euklid_{rek}(a, b)

```

1  if  $b = 0$  then
2    return  $a$ 
3  else
4    return Euklidrek( $b, a \bmod b$ )

```

Zur Abschätzung von n verwenden wir die Folge der Fibonacci-Zahlen F_n :

$$F_n = \begin{cases} 0, & \text{falls } n = 0 \\ 1, & \text{falls } n = 1 \\ F_{n-1} + F_{n-2}, & \text{falls } n \geq 2 \end{cases}$$

*Also: $d_i = r_{i-2} \text{ div } r_{i-1}$ und $r_i = r_{i-2} \bmod r_{i-1}$.

Durch Induktion über $i = n, n-1, \dots, 0$ folgt $r_i \geq F_{n+1-i}$; also $a = r_0 \geq F_{n+1}$. Weiterhin lässt sich durch Induktion über $n \geq 0$ zeigen, dass $F_{n+1} \geq \phi^{n-1}$ ist, wobei $\phi = (1 + \sqrt{5})/2$ der *goldene Schnitt* ist. Der Induktionsanfang ($n = 0$ oder 1) ist klar, da $F_2 = F_1 = 1 = \phi^0 \geq \phi^{-1}$ ist. Unter der Induktionsannahme $F_{i+1} \geq \phi^{i-1}$ für $i \leq n-1$ folgt wegen $\phi^2 = \phi + 1$

$$F_{n+1} = F_n + F_{n-1} \geq \phi^{n-2} + \phi^{n-3} = \phi^{n-3}(\phi + 1) = \phi^{n-1}.$$

Somit ist $a \geq \phi^{n-1}$, d. h. $n \leq 1 + \log_\phi a$.

Satz 11. *Der Euklidische Algorithmus führt zur Berechnung von $\text{ggT}(a, b)$ (unter der Annahme $a > b > 0$) höchstens $\lfloor \log_\phi a \rfloor + 1$ Divisionsschritte durch. Dies führt auf eine Zeitkomplexität von $O(n^3)$, wobei n die Länge der Eingabe in Binärdarstellung bezeichnet und wir $O(n^2)$ Rechenschritte für eine einzelne Ganzzahldivision ansetzen.*

Erweiterter Euklidischer bzw. Berlekamp-Algorithmus: Der Euklidische Algorithmus kann so modifiziert werden, dass er eine lineare Darstellung

$$\text{ggT}(a, b) = \lambda a + \mu b \quad \text{mit} \quad \lambda, \mu \in \mathbb{Z}$$

des ggT liefert (Zeitkomplexität ebenfalls $O(n^3)$). Hierzu werden neben r_i und d_i weitere Zahlen

$$p_i = p_{i-2} - d_i p_{i-1}, \quad \text{wobei} \quad p_0 = 1 \quad \text{und} \quad p_1 = 0,$$

und

$$q_i = q_{i-2} - d_i q_{i-1}, \quad \text{wobei} \quad q_0 = 0 \quad \text{und} \quad q_1 = 1,$$

für $i = 0, \dots, n$ bestimmt. Dann gilt für $i = 0$ und $i = 1$,

$$ap_i + bq_i = r_i,$$

und durch Induktion über i ,

$$\begin{aligned} ap_{i+1} + bq_{i+1} &= a(p_{i-1} - d_{i+1}p_i) + b(q_{i-1} - d_{i+1}q_i) \\ &= ap_{i-1} + bq_{i-1} - d_{i+1}(ap_i + bq_i) \\ &= (r_{i-1} - d_{i+1}r_i) \\ &= r_{i+1} \end{aligned}$$

zeigt man, dass dies auch für $i = 2, \dots, n$ gilt. Insbesondere gilt also

$$ap_n + bq_n = r_n = \text{ggT}(a, b).$$

Korollar 12 (Lemma von Bezout). *Der größte gemeinsame Teiler von a und b ist in der Form*

$$\text{ggT}(a, b) = \lambda a + \mu b \quad \text{mit} \quad \lambda, \mu \in \mathbb{Z}$$

darstellbar.

Beispiel 13. Für $a = 693$ und $b = 147$ erhalten wir wegen

i	r_{i-1}	$=$	$d_{i+1} \cdot$	$r_i + r_{i+1}$	p_i	q_i	$p_i \cdot 693 + q_i \cdot 147 =$	r_i
0					1	0	$1 \cdot 693 + 0 \cdot 147 =$	693
1	693	$=$	4	$\cdot 147 + 105$	0	1	$0 \cdot 693 + 1 \cdot 147 =$	147
2	147	$=$	1	$\cdot 105 + 42$	1	-4	$1 \cdot 693 - 4 \cdot 147 =$	105
3	105	$=$	2	$\cdot 42 + 21$	-1	5	$-1 \cdot 693 + 5 \cdot 147 =$	42
4	42	$=$	2	$\cdot \mathbf{21} + 0$	3	-14	$3 \cdot 693 - 14 \cdot 147 =$	21

die lineare Darstellung $3 \cdot 693 - 14 \cdot 147 = 21$.

◁

Aus der linearen Darstellbarkeit des größten gemeinsamen Teilers ergeben sich eine Reihe von nützlichen Schlussfolgerungen.

Korollar 14. $\text{ggT}(a, b) = \min\{\lambda a + \mu b \geq 1 \mid \lambda, \mu \in \mathbb{Z}\}$.

Beweis. Sei $m = \min\{\lambda a + \mu b \geq 1 \mid \lambda, \mu \in \mathbb{Z}\}$ und $g = \text{ggT}(a, b)$. Dann folgt $g \geq m$, da g in der Menge $\{\lambda a + \mu b \geq 1 \mid \lambda, \mu \in \mathbb{Z}\}$ enthalten ist, und $g \leq m$, da g Teiler von jeder Zahl der Form $\lambda a + \mu b$ ist. \square

Korollar 15. *Der größte gemeinsame Teiler von a und b wird von allen gemeinsamen Teilern von a und b geteilt,*

$$x|a \wedge x|b \Rightarrow x|\text{ggT}(a, b).$$

Beweis. Sei $g = \text{ggT}(a, b)$. Dann existieren Zahlen $\mu, \lambda \in \mathbb{Z}$ mit $\mu a + \lambda b = g$. Da x nach Voraussetzung sowohl a als auch b teilt, teilt x auch die Zahlen μa und λb und somit auch deren Summe $\mu a + \lambda b = g$. \square

Korollar 16 (Lemma von Euklid). *Teilt a das Produkt bc und sind a, b teilerfremd, so ist a auch Teiler von c ,*

$$a|bc \wedge \text{ggT}(a, b) = 1 \Rightarrow a|c.$$

Beweis. Wegen $\text{ggT}(a, b) = 1$ existieren Zahlen $\mu, \lambda \in \mathbb{Z}$ mit $\mu a + \lambda b = 1$. Da a nach Voraussetzung das Produkt bc teilt, muss a auch $c\mu a + c\lambda b = c$ teilen. \square

Korollar 17. *Wenn sowohl a als auch b zu einer Zahl $m \in \mathbb{Z}$ teilerfremd sind, so ist auch das Produkt ab teilerfremd zu m ,*

$$\text{ggT}(a, m) = \text{ggT}(b, m) = 1 \Rightarrow \text{ggT}(ab, m) = 1.$$

Beweis. Da a und b teilerfremd zu m sind, existieren Zahlen $\mu, \lambda, \mu', \lambda' \in \mathbb{Z}$ mit $\mu a + \lambda m = \mu' b + \lambda' m = 1$. Somit ergibt sich aus der Darstellung

$$1 = (\mu a + \lambda m)(\mu' b + \lambda' m) = \underbrace{\mu\mu'}_{\mu''} ab + \underbrace{(\mu a\lambda' + \mu' b\lambda + \lambda\lambda' m)}_{\lambda''} m,$$

dass auch ab teilerfremd zu m ist. \square

Damit nun eine Abbildung $g : A \rightarrow A$ von der Bauart $g(x) = bx$ injektiv (oder gleichbedeutend, surjektiv) ist, muss es zu jedem Buchstaben $y \in A$ genau einen Buchstaben $x \in A$ mit $bx = y$ geben. Wie der folgende Satz zeigt, ist dies genau dann der Fall, wenn b und m teilerfremd sind.

Satz 18. *Sei $m \geq 1$. Die lineare Kongruenzgleichung $bx \equiv_m y$ besitzt genau dann eine eindeutige Lösung $x \in \{0, \dots, m-1\}$, wenn $\text{ggT}(b, m) = 1$ ist.*

Beweis. Angenommen, $\text{ggT}(b, m) = g > 1$. Dann ist mit x auch $x' = x + m/g$ eine Lösung von $bx \equiv_m y$ mit $x \not\equiv_m x'$. Gilt umgekehrt $\text{ggT}(b, m) = 1$, so folgt aus den Kongruenzen

$$bx_1 \equiv_m y$$

und

$$bx_2 \equiv_m y$$

sofort $b(x_1 - x_2) \equiv_m 0$, also $m|b(x_1 - x_2)$. Wegen $\text{ggT}(b, m) = 1$ folgt mit dem Lemma von Euklid $m|(x_1 - x_2)$, also $x_1 \equiv_m x_2$.

Dies zeigt, dass die Abbildung $f : \mathbb{Z}_m \rightarrow \mathbb{Z}_m$ mit $f(x) = bx \pmod m$ injektiv ist. Da jedoch Definitions- und Wertebereich von f identisch sind, muss f dann auch surjektiv sein. Dies impliziert, dass die Kongruenz $bx \equiv_m y$ für jedes $y \in \mathbb{Z}_m$ lösbar ist. \square

Korollar 19. *Im Fall $\text{ggT}(b, m) = 1$ hat die Kongruenz $bx \equiv_m 1$ genau eine Lösung, die das **multiplikative Inverse** von b modulo m genannt und mit $b^{-1} \pmod m$ (oder einfach mit b^{-1}) bezeichnet wird. Die invertierbaren Elemente von \mathbb{Z}_m werden in der Menge*

$$\mathbb{Z}_m^* = \{b \in \mathbb{Z}_m \mid \text{ggT}(b, m) = 1\}$$

*zusammengefasst (die Elemente von \mathbb{Z}_m werden auch als **Einheiten** bezeichnet).*

Korollar 17 zeigt, dass \mathbb{Z}_m^* unter der Operation \odot_m abgeschlossen ist, und mit Korollar 19 folgt, dass $(\mathbb{Z}_m^*, \odot_m)$ eine multiplikative Gruppe bildet. Allgemeiner zeigt man, dass für einen beliebigen Ring $(R, +, \cdot, 0, 1)$ mit Eins die Multiplikation auf der Menge $R^* = \{a \in R \mid \exists b \in R : ab = 1 = ba\}$ eine Gruppe $(R^*, \cdot, 1)$ (die so genannte **Einheitengruppe** von R) bildet.

Das multiplikative Inverse von b modulo m ergibt sich aus der linearen Darstellung $\lambda b + \mu m = \text{ggT}(b, m) = 1$ zu $b^{-1} = \lambda \pmod m$. Bei Kenntnis von b^{-1} kann die Kongruenz $bx \equiv_m y$ leicht zu $x = yb^{-1} \pmod m$ gelöst werden. Die folgende Tabelle zeigt die multiplikativen Inversen b^{-1} für alle $b \in \mathbb{Z}_{26}^*$.

b	1	3	5	7	9	11	15	17	19	21	23	25
b^{-1}	1	9	21	15	3	19	7	23	11	5	17	25

Nun lässt sich die additive Chiffre leicht zur affinen Chiffre erweitern.

Definition 20 (affine Chiffre). *Bei der **affinen Chiffre** ist $A = B = M = C$ ein beliebiges Alphabet mit $m := \|A\| > 1$ und $K = \mathbb{Z}_m^* \times \mathbb{Z}_m$. Für $k = (b, c) \in K$, $x \in M$ und $y \in C$ gilt*

$$E(k, x) = bx + c \quad \text{und} \quad D(k, y) = b^{-1}(y - c).$$

In diesem Fall liefert die Schlüsselkomponente $b = -1$ für jeden Wert von c eine involutorische Chiffrierfunktion $x \mapsto E(b, c; x) = c - x$ (**verschobenes komplementäres Alphabet**). Wählen wir für c ebenfalls den Wert -1 , so ergibt sich die Chiffrierfunktion $x \mapsto -x - 1$, die auch als **revertiertes Alphabet** bekannt ist. Offenbar ist diese Funktion genau dann echt involutorisch, wenn m gerade ist.

x	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
$-x$	A	Z	Y	X	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B
$-x - 1$	Z	Y	X	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B	A

Als nächstes illustrieren wir die Ver- und Entschlüsselung mit der affinen Chiffre an einem kleinen Beispiel.

Beispiel 21 (affine Chiffre). *Sei $A = \{\mathbf{A}, \dots, \mathbf{Z}\} = B$, also $m = 26$. Weiter sei $k = (9, 2)$, also $b = 9$ und $c = 2$. Um den Klartextbuchstaben $x = \mathbf{F}$ zu verschlüsseln, berechnen wir*

$$E(k, x) = bx + c = 9\mathbf{F} + 2 = \mathbf{V},$$

da der Index von \mathbf{F} gleich 5, der von \mathbf{V} gleich 21 und $9 \cdot 5 + 2 = 47 \equiv_{26} 21$ ist. Um einen Kryptotextbuchstaben wieder entschlüsseln zu können, benötigen wir das multiplikative Inverse von $b = 9$, das sich wegen

i	r_{i-1}	$=$	$d_{i+1} \cdot r_i$	$+$	r_{i+1}	$p_i \cdot 26$	$+$	$q_i \cdot 9$	$=$	r_i
0						$1 \cdot 26$	$+$	$0 \cdot 9$	$=$	26
1	26	$=$	$2 \cdot 9$	$+$	8	$0 \cdot 26$	$+$	$1 \cdot 9$	$=$	9
2	9	$=$	$1 \cdot 8$	$+$	1	$1 \cdot 26$	$+$	$(-2) \cdot 9$	$=$	8
3	8	$=$	$8 \cdot 1$	$+$	0	$(-1) \cdot 26$	$+$	$3 \cdot 9$	$=$	1

zu $b^{-1} = q_3 = 3$ ergibt. Damit erhalten wir für den Kryptotextbuchstaben $y = \mathbf{V}$ den ursprünglichen Klartextbuchstaben

$$D(k, y) = b^{-1}(y - c) = 3(\mathbf{V} - 2) = \mathbf{F}$$

zurück, da $3 \cdot 19 = 57 \equiv_{26} 5$ ist. ◁

Eine wichtige Rolle spielt die Funktion

$$\varphi : \mathcal{N} \rightarrow \mathcal{N} \quad \text{mit} \quad \varphi(n) = \|\mathbb{Z}_n^*\| = \|\{a \mid 0 \leq a \leq n - 1, \text{ggT}(a, n) = 1\}\|,$$

die sogenannte *Eulersche φ -Funktion*.

n	1	2	3	4	5	6	7	8	9
\mathbb{Z}_n^*	{0}	{1}	{1, 2}	{1, 3}	{1, 2, 3, 4}	{1, 5}	{1, ..., 6}	{1, 3, 5, 7}	{1, 2, 4, 5, 7, 8}
$\varphi(n)$	1	1	2	2	4	2	6	4	6

Wegen

$$\mathbb{Z}_{p^e} - \mathbb{Z}_{p^e}^* = \{0, p, 2p, \dots, (p^{e-1} - 1)p\}$$

folgt sofort

$$\varphi(p^e) = p^e - p^{e-1} = p^{e-1}(p - 1).$$

Um hieraus für beliebige Zahlen $m \in \mathcal{N}$ eine Formel für $\varphi(m)$ zu erhalten, genügt es, $\varphi(ab)$ im Fall $\text{ggT}(a, b) = 1$ in Abhängigkeit von $\varphi(a)$ und $\varphi(b)$ zu bestimmen. Hierzu betrachten wir die Abbildung $f : \mathbb{Z}_{ml} \rightarrow \mathbb{Z}_m \times \mathbb{Z}_l$ mit

$$f(x) := (x \bmod m, x \bmod l).$$

Beispiel 22. Sei $m = 5$ und $l = 6$. Dann erhalten wir die Funktion $f : \mathbb{Z}_{30} \rightarrow \mathbb{Z}_5 \times \mathbb{Z}_6$ mit

x	0	1	2	3	4	5	6	7	8	9
$f(x)$	(0, 0)	(1, 1)	(2, 2)	(3, 3)	(4, 4)	(0, 5)	(1, 0)	(2, 1)	(3, 2)	(4, 3)

x	10	11	12	13	14	15	16	17	18	19
$f(x)$	(0, 4)	(1, 5)	(2, 0)	(3, 1)	(4, 2)	(0, 3)	(1, 4)	(2, 5)	(3, 0)	(4, 1)

x	20	21	22	23	24	25	26	27	28	29
$f(x)$	(0, 2)	(1, 3)	(2, 4)	(3, 5)	(4, 0)	(0, 1)	(1, 2)	(2, 3)	(3, 4)	(4, 5)

Man beachte, dass f eine Bijektion zwischen \mathbb{Z}_{30} und $\mathbb{Z}_5 \times \mathbb{Z}_6$ ist. Zudem fällt auf, dass ein x -Wert genau dann in \mathbb{Z}_{30}^* liegt, wenn der Funktionswert $f(x) = (y, z)$ zu $\mathbb{Z}_5^* \times \mathbb{Z}_6^*$ gehört (die Werte $x \in \mathbb{Z}_{30}^*$, $y \in \mathbb{Z}_5^*$ und $z \in \mathbb{Z}_6^*$ sind **fett** gedruckt). Folglich bildet f die Argumente in \mathbb{Z}_{30}^* bijektiv auf die Werte in $\mathbb{Z}_5^* \times \mathbb{Z}_6^*$ ab. Für f^{-1} erhalten wir somit folgende Tabelle:

f^{-1}	0	1	2	3	4	5
0	0	25	20	15	10	5
1	6	1	26	21	16	11
2	12	7	2	27	22	17
3	18	13	8	3	28	23
4	24	19	14	9	4	29

◁

Der Chinesische Restsatz, den wir im nächsten Abschnitt beweisen, besagt, dass f im Fall $\text{ggT}(m, l) = 1$ bijektiv und damit invertierbar ist. Wegen

$$\begin{aligned} \text{ggT}(x, ml) = 1 &\Leftrightarrow \text{ggT}(x, m) = \text{ggT}(x, l) = 1 \\ &\Leftrightarrow \text{ggT}(x \bmod m, m) = \text{ggT}(x \bmod l, l) = 1 \end{aligned}$$

ist daher die Einschränkung \hat{f} von f auf den Bereich \mathbb{Z}_{ml}^* eine Bijektion zwischen \mathbb{Z}_{ml}^* und $\mathbb{Z}_m^* \times \mathbb{Z}_l^*$, d.h. es gilt

$$\varphi(ml) = \|\mathbb{Z}_{ml}^*\| = \|\mathbb{Z}_m^* \times \mathbb{Z}_l^*\| = \|\mathbb{Z}_m^*\| \cdot \|\mathbb{Z}_l^*\| = \varphi(m)\varphi(l).$$

Satz 23. Die Eulersche φ -Funktion ist multiplikativ, d. h. für teilerfremde Zahlen m und l gilt $\varphi(ml) = \varphi(m)\varphi(l)$.

Korollar 24. Sei $m = \prod_{i=1}^k p_i^{e_i}$ die Primfaktorzerlegung von m . Dann gilt

$$\varphi(m) = \prod_{i=1}^k p_i^{e_i-1}(p_i - 1) = m \prod_{i=1}^k (p_i - 1)/p_i.$$

Beweis. Es gilt

$$\varphi(\prod_{i=1}^k p_i^{e_i}) = \prod_{i=1}^k \varphi(p_i^{e_i}) = \prod_{i=1}^k (p_i^{e_i} - p_i^{e_i-1}) = \prod_{i=1}^k p_i^{e_i-1}(p_i - 1).$$

□

Der Chinesische Restsatz

Die beiden linearen Kongruenzen

$$\begin{aligned} x &\equiv_3 0 \\ x &\equiv_6 1 \end{aligned}$$

besitzen je eine Lösung, es gibt aber kein x , das beide Kongruenzen gleichzeitig erfüllt. Der nächste Satz zeigt, dass unter bestimmten Voraussetzungen gemeinsame Lösungen existieren, und wie sie berechnet werden können.

Satz 25 (Chinesischer Restsatz). *Falls m_1, \dots, m_k paarweise teilerfremd sind, dann hat das System*

$$\begin{aligned} x &\equiv_{m_1} b_1 \\ &\vdots \\ x &\equiv_{m_k} b_k \end{aligned} \tag{1.2}$$

genau eine Lösung modulo $m = \prod_{i=1}^k m_i$.

Beweis. Da die Zahl $n_i = m/m_i$ teilerfremd zu m_i ist, existieren Zahlen μ_i und λ_i mit

$$\mu_i n_i + \lambda_i m_i = \text{ggT}(n_i, m_i) = 1.$$

Dann gilt

$$\mu_i n_i \equiv_{m_i} 1$$

und

$$\mu_i n_i \equiv_{m_j} 0$$

für $j \neq i$. Folglich erfüllt $x = \sum_{j=1}^k \mu_j n_j b_j$ die Kongruenzen

$$x \equiv_{m_i} \mu_i n_i b_i \equiv_{m_i} b_i$$

für $i = 1, \dots, k$. Dies zeigt, dass (1.2) lösbar, also die Funktion

$$f : \mathbb{Z}_m \rightarrow \mathbb{Z}_{m_1} \times \dots \times \mathbb{Z}_{m_k}$$

mit $f(x) = (x \bmod m_1, \dots, x \bmod m_k)$ surjektiv ist. Da der Definitions- und der Wertebereich von f die gleiche Mächtigkeit haben, muss f jedoch auch injektiv sein, d.h. (1.2) ist sogar eindeutig lösbar. \square

Man beachte, dass der Beweis des Chinesischen Restsatzes konstruktiv ist und die Lösung x unter Verwendung des erweiterten Euklidischen Algorithmus' effizient berechenbar ist.

1.4 Die Hill-Chiffre

Die von Hill im Jahr 1929 publizierte Chiffre ist eine Erweiterung der multiplikativen Chiffre auf Buchstabenblöcke, d.h. der Klartext wird nicht zeichenweise, sondern blockweise verarbeitet. Sowohl der Klartext- als auch der Kryptotextraum enthält alle Wörter x über A einer festen Länge l . Zur Chiffrierung wird eine $(l \times l)$ -Matrix $k = (k_{ij})$ mit Koeffizienten in \mathbb{Z}_m benutzt, die einen Klartextblock $x = x_1 \dots x_l \in A^l$ in den Kryptotextblock $y_1 \dots y_l \in A^l$ transformiert, wobei

$$y_i = x_1 k_{1i} + \dots + x_l k_{li}, \quad i = 1, \dots, l$$

ist (hierbei machen wir von der Buchstabenrechnung Gebrauch). y entsteht also durch Multiplikation von x mit der Schlüsselmatrix k :

$$(x_1, \dots, x_l) \begin{pmatrix} k_{11} & \dots & k_{1l} \\ \vdots & \ddots & \vdots \\ k_{l1} & \dots & k_{ll} \end{pmatrix} = (y_1, \dots, y_l)$$

Wir bezeichnen die Menge aller $(l \times l)$ -Matrizen mit Koeffizienten in \mathbb{Z}_m mit $\mathbb{Z}_m^{l \times l}$. Als Schlüssel können nur invertierbare Matrizen k benutzt werden, da sonst der Chiffriervorgang nicht injektiv ist. k ist genau dann invertierbar, wenn die Determinante von k teilerfremd zu m ist (siehe Übungen).

Definition 26 (Determinante). Sei R ein kommutativer Ring mit Eins und sei $A = (a_{ij}) \in R^{l \times l}$. Für $1 \leq i, j \leq l$ sei A_{ij} die durch Streichen der i -ten Zeile und j -ten Spalte aus A hervorgehende Matrix. Die **Determinante** von A ist dann $\det(A) = a_{11}$, falls $l = 1$, und

$$\det(A) = \sum_{j=1}^l (-1)^{i+j} a_{i,j} \det(A_{ij}),$$

wobei $i \in \{1, \dots, l\}$ (beliebig wählbar) ist.

Die Determinantenfunktion $f : R^{l \times l} \rightarrow R$ ist durch die drei Eigenschaften multilinear, alternierend und normiert eindeutig festgelegt. Seien a_1, \dots, a_n die Spalten von A , also $A = (a_1, \dots, a_n)$. Dann heißt f **multilinear**, falls $f(a_1, \dots, a_i + b, \dots, a_n) = f(a_1, \dots, a_i, \dots, a_n) + f(a_1, \dots, b, \dots, a_n)$ und $f(a_1, \dots, ra_i, \dots, a_n) = rf(a_1, \dots, a_i, \dots, a_n)$ ist. f heißt **alternierend**, falls im Fall $a_i = a_j$ für $i \neq j$ $f(a_1, \dots, a_n) = 0$ ist. f heißt **normiert**, falls $f(E) = 1$ ist, wobei E die Einheitsmatrix ist.

Für die Dechiffrierung wird die zu k inverse Matrix k^{-1} benötigt, wofür effiziente Algorithmen bekannt sind (siehe Übungen).

Satz 27. Sei A ein Alphabet und sei $k \in \mathbb{Z}_m^{l \times l}$ ($l \geq 1$, $m = \|A\|$). Die Abbildung $f : A^l \rightarrow A^l$ mit

$$f(x) = xk,$$

ist genau dann injektiv, wenn $\text{ggT}(\det(k), m) = 1$ ist.

Beweis. Siehe Übungen. □

Definition 28 (Hill-Chiffre). Sei $A = \{a_0, \dots, a_{m-1}\}$ ein beliebiges Alphabet und für eine natürliche Zahl $l \geq 2$ sei $M = C = A^l$. Bei der **Hill-Chiffre** ist $K = \{k \in \mathbb{Z}_m^{l \times l} \mid \text{ggT}(\det(k), m) = 1\}$ und es gilt

$$E(k, x) = xk \quad \text{und} \quad D(k, y) = yk^{-1}.$$

Beispiel 29 (Hill-Chiffre). Benutzen wir zur Chiffrierung von Klartextblöcken der Länge $l = 4$ über dem lateinischen Alphabet A_{lat} die Schlüsselmatrix

$$k = \begin{pmatrix} 11 & 13 & 8 & 21 \\ 24 & 17 & 3 & 25 \\ 18 & 12 & 23 & 17 \\ 6 & 15 & 2 & 15 \end{pmatrix},$$

so erhalten wir beispielsweise für den Klartext **HILL** wegen

$$(\text{HILL}) \begin{pmatrix} 11 & 13 & 8 & 21 \\ 24 & 17 & 3 & 25 \\ 18 & 12 & 23 & 17 \\ 6 & 15 & 2 & 15 \end{pmatrix} = (\text{NERX}) \quad \text{bzw.} \quad \begin{array}{l} 11\text{H} + 24\text{I} + 18\text{L} + 6\text{L} = \text{N} \\ 13\text{H} + 17\text{I} + 12\text{L} + 15\text{L} = \text{E} \\ 8\text{H} + 3\text{I} + 23\text{L} + 2\text{L} = \text{R} \\ 21\text{H} + 25\text{I} + 17\text{L} + 15\text{L} = \text{X} \end{array}$$

den Kryptotext $E(k, \text{HILL}) = \text{NERX}$. Für die Entschlüsselung wird die inverse Matrix k^{-1} benötigt. Diese wird in den Übungen berechnet. ◁

1.5 Die Vigenère-Chiffre und andere Stromsysteme

Bei der nach dem Franzosen Blaise de Vigenère (1523–1596) benannten Chiffre werden zwar nur einzelne Buchstaben chiffriert, aber je nach Position im Klartext unterschiedlich.

Definition 30 (Vigenère-Chiffre). Sei $A = B$ ein beliebiges Alphabet. Die **Vigenère-Chiffre** chiffriert unter einem Schlüssel $k = k_0 \dots k_{d-1} \in K = A^*$ einen Klartext $x = x_0 \dots x_{n-1}$ beliebiger Länge zu

$$E(k, x) = y_0 \dots y_{n-1}, \quad \text{wobei } y_i = x_i + k_{(i \bmod d)} \text{ ist,}$$

und dechiffriert einen Kryptotext $y = y_0 \dots y_{n-1}$ zu

$$D(k, y) = x_0 \dots x_{n-1}, \quad \text{wobei } x_i = y_i - k_{(i \bmod d)} \text{ ist.}$$

Beispiel 31 (Vigenère-Chiffre). Verwenden wir das lateinische Alphabet A_{lat} als Klartextalphabet und wählen wir als Schlüssel das Wort $k = \mathbf{WIE}$, so ergibt sich für den Klartext **VIGENERE** beispielsweise der Kryptotext

$$\begin{aligned} E(\mathbf{WIE}, \mathbf{VIGENERE}) &= \underbrace{\mathbf{V+W}}_R \underbrace{\mathbf{I+I}}_Q \underbrace{\mathbf{G+E}}_K \underbrace{\mathbf{E+W}}_A \underbrace{\mathbf{N+I}}_V \underbrace{\mathbf{E+E}}_I \underbrace{\mathbf{R+W}}_N \underbrace{\mathbf{E+I}}_M \\ &= \mathbf{RQKAVINM} \end{aligned}$$

◁

Um einen Klartext x zu verschlüsseln, wird also das Schlüsselwort $k = k_0 \dots k_{d-1}$ so oft wiederholt, bis der dabei entstehende **Schlüsselstrom** $\hat{k} = k_0, k_1, \dots, k_{d-1}, k_0 \dots$ die Länge von x erreicht. Dann werden x und \hat{k} zeichenweise addiert, um den zugehörigen Kryptotext y zu bilden. Aus diesem kann der ursprüngliche Klartext x zurückgewonnen werden, indem man den Schlüsselstrom \hat{k} wieder subtrahiert.

Beispiel 32. Vigenère-Chiffre

Chiffrierung:

$$\begin{array}{l} \mathbf{VIGENERE} \quad (\text{Klartext } x) \\ + \mathbf{WIEWIEWI} \quad (\text{Schlüsselstrom } \hat{k}) \\ \hline \mathbf{RQKAVINM} \quad (\text{Kryptotext } y) \end{array}$$

Dechiffrierung:

$$\begin{array}{l} \mathbf{RQKAVINM} \quad (\text{Kryptotext } y) \\ - \mathbf{WIEWIEWI} \quad (\text{Schlüsselstrom } \hat{k}) \\ \hline \mathbf{VIGENERE} \quad (\text{Klartext } x) \end{array}$$

◁

Die Chiffrierarbeit lässt sich durch Benutzung einer Additionstabelle erleichtern (auch als **Vigenère-Tableau** bekannt).

+	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
A	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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	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

Um eine involutorische Chiffre zu erhalten, schlug Sir Francis Beaufort, ein Admiral der britischen Marine, vor, den Schlüsselstrom nicht auf den Klartext zu addieren, sondern letzteren von ersterem zu subtrahieren.

Beispiel 33 (Beaufort-Chiffre). Verschlüsseln wir den Klartext **BEAUFORT** beispielsweise unter dem Schlüsselwort $k = \mathbf{WIE}$, so erhalten wir den Kryptotext **XMEQNSNB**. Eine erneute Verschlüsselung liefert wieder den Klartext **BEAUFORT**:

$$\begin{array}{r}
 \text{Chiffrierung:} \\
 \underline{\mathbf{WIEWIEWI}} \quad (\text{Schlüsselstrom}) \\
 - \mathbf{BEAUFORT} \quad (\text{Klartext}) \\
 \hline
 \mathbf{XMEQNSNB} \quad (\text{Kryptotext})
 \end{array}
 \qquad
 \begin{array}{r}
 \text{Dechiffrierung:} \\
 \underline{\mathbf{WIEWIEWI}} \quad (\text{Schlüsselstrom}) \\
 - \mathbf{XMEQNSNB} \quad (\text{Kryptotext}) \\
 \hline
 \mathbf{BEAUFORT} \quad (\text{Klartext})
 \end{array}$$

◁

Bei den bisher betrachteten Chiffren wird aus einem Schlüsselwort $k = k_0 \dots k_{d-1}$ ein **periodischer Schlüsselstrom** $\hat{k} = \hat{k}_0 \dots \hat{k}_{n-1}$ erzeugt, das heißt, es gilt $\hat{k}_i = \hat{k}_{i+d}$ für alle $i = 0, \dots, n - d - 1$. Da eine kleine Periode das Brechen der Chiffre erleichtert, sollte entweder ein Schlüsselstrom mit sehr großer Periode oder noch besser ein **fortlaufender Schlüsselstrom** zur Chiffrierung benutzt werden. Ein solcher nichtperiodischer Schlüsselstrom lässt sich beispielsweise ohne großen Aufwand erzeugen, indem man an das Schlüsselwort den Klartext oder den Kryptotext anhängt (sogenannte **Autokey-Chiffrierung**).[†]

[†]Die Idee, den Schlüsselstrom durch Anhängen des Klartextes an ein Schlüsselwort zu bilden, stammt

Beispiel 34 (Autokey-Chiffre). Benutzen wir wieder das Schlüsselwort **WIE**, um den Schlüsselstrom durch Anhängen des Klar- bzw. Kryptotextes zu erzeugen, so erhalten wir für den Klartext **VIGENERE** folgende Kryptotexte:

$$\begin{array}{ll}
 \text{Klartext-Schlüsselstrom:} & \text{Kryptotext-Schlüsselstrom:} \\
 \text{VIGENERE (Klartext)} & \text{VIGENERE (Klartext)} \\
 + \text{WIEVIGEN (Schlüsselstrom)} & + \text{WIERQKVD (Schlüsselstrom)} \\
 \hline
 \text{RQKZVKVR (Kryptotext)} & \text{RQKVDOMH (Kryptotext)}
 \end{array}$$

<

Auch die Dechiffrierung ist in beiden Fällen einfach. Bei der ersten Alternative kann der Empfänger durch Subtraktion des Schlüsselworts den Anfang des Klartextes bilden und gleichzeitig den Schlüsselstrom verlängern, so dass sich auf diese Weise Stück für Stück der gesamte Kryptotext entschlüsseln lässt. Noch einfacher gestaltet sich die Dechiffrierung im zweiten Fall, da sich hier der Schlüsselstrom vom Kryptotext nur durch das vorangestellte Schlüsselwort unterscheidet.

1.6 Der One-Time-Pad

Es besteht auch die Möglichkeit, eine Textstelle in einem Buch als Schlüssel zu vereinbaren und den dort beginnenden Text als Schlüsselstrom zu benutzen (Lauftextverschlüsselung). Besser ist es jedoch, aus einem relativ kurzen Schlüssel einen möglichst zufällig erscheinenden Schlüsselstrom zu erzeugen. Hierzu können beispielsweise Pseudozufallsgeneratoren eingesetzt werden. Absolute Sicherheit wird dagegen erreicht, wenn der Schlüsselstrom rein zufällig erzeugt und nach einmaliger Benutzung wieder vernichtet wird.[‡] Ein solcher „Wegwerfsschlüssel“ (*One-time-pad* oder *One-time-tape*, im Deutschen auch als **individueller Schlüssel** bezeichnet) lässt sich allerdings nur mit großem Aufwand generieren und verteilen, weshalb diese Chiffre nur wenig praktikabel ist. Dennoch wurde diese Methode beispielsweise beim „heißen Draht“, der 1963 eingerichteten, direkten Fernschreibverbindung zwischen dem Weißen Haus in Washington und dem Kreml in Moskau, angewandt.

Beispiel 35 (One-time-pad). Sei $A = \{a_0, \dots, a_{m-1}\}$ ein beliebiges Klartextalphabet. Um einen Klartext $x = x_0 \dots x_{n-1}$ zu verschlüsseln, wird auf jeden Klartextbuchstaben x_i ein neuer, zufällig generierter Schlüsselbuchstabe k_i addiert,

$$y = y_0 \dots y_{n-1}, \text{ wobei } y_i = x_i + k_i.$$

<

Der Klartext wird also wie bei einer additiven Chiffre verschlüsselt, nur dass der Schlüssel nach einmaligem Gebrauch gewechselt wird. Dies entspricht dem Gebrauch einer Vigenère-Chiffre, falls als Schlüssel ein zufällig gewähltes Wort von der Länge des Klartextes benutzt wird. Wie diese ist der One-time-pad im Binärfall also involutorisch.

von Vigenère, während er mit der Erfindung der nach ihm benannten Vigenère-Chiffre „nichts zu tun“ hatte. Diese wird vielmehr Giovan Batista Belaso (1553) zugeschrieben.

[‡] Diese Art der Schlüsselerzeugung schlug der amerikanische Major Joseph O. Mauborgne im Jahr 1918 vor, nachdem ihm ein von Gilbert S. Vernam für den Fernschreibverkehr entwickeltes Chiffriersystem vorgestellt wurde.

Von der Methode, die letzte *Zeile* nur zum Teil zu füllen, ist dagegen abzuraten. In diesem Fall würden nämlich auf dem abgewickelten Papierstreifen Lücken entstehen, aus deren Anordnung man Schlüsse auf den benutzten Schlüssel k ziehen könnte. Andererseits ist nichts dagegen einzuwenden, dass der Sender die letzte *Spalte* der Skytale nur zum Teil beschriftet.

Bevor wir weitere Beispiele für Transpositionen betrachten, wenden wir uns der Klassifikation von Substitutionschiffren zu. Ein wichtiges Unterscheidungsmerkmal ist z.B. die Länge der Klartexteinheiten, auf denen die Chiffre operiert.

Monografische Substitutionen ersetzen Einzelbuchstaben.

Polygrafische Substitutionen ersetzen dagegen aus mehreren Zeichen bestehende Klartextsegmente auf einmal.

Eine polygrafische Substitution, die auf Buchstabenpaaren operiert, wird **digrafisch** genannt. Das älteste bekannte polygrafische Chiffrierverfahren wurde von Giovanni Porta im Jahr 1563 veröffentlicht. Dabei werden je zwei aufeinanderfolgende Klartextbuchstaben durch ein einzelnes Kryptotextzeichen ersetzt.

Beispiel 37. Bei der **Porta-Chiffre** werden 400 (!) unterschiedliche von Porta für diesen Zweck entworfene Kryptotextzeichen verwendet. Diese sind in einer 20×20 -Matrix $M = (y_{ij})$ angeordnet, deren Zeilen und Spalten mit den 20 Klartextbuchstaben $A, \dots, I, L, \dots, T, V, Z$ indiziert sind. Zur Ersetzung des Buchstabenpaars $a_i a_j$ wird das in Zeile i und Spalte j befindliche Kryptotextzeichen

$$E(M, a_i a_j) = y_{ij}$$

benutzt. ◁

Eine Substitution heißt **monopartit**, falls sie die Klartextsegmente durch Einzelzeichen ersetzt, sonst **multipartit**. Wird der Kryptotext aus Buchstabenpaaren zusammengesetzt, so spricht man von einer **bipartiten** Substitution.

Ein frühes (monografisches) Beispiel einer bipartiten Chiffriermethode geht auf Polybios (circa 200 – 120 v. Chr.) zurück:

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

POLYBIOS \rightsquigarrow 30 24 21 43 01 13 24 33

Bei der **Polybios-Chiffre** dient eine 5×5 -Matrix, die aus sämtlichen Klartextbuchstaben gebildet wird, als Schlüssel.[§] Die Verschlüsselung des Klartextes erfolgt buchstabenweise, indem man einen in Zeile i und Spalte j eingetragenen Klartextbuchstaben durch das Koordinatenpaar ij ersetzt. Der Kryptotextraum besteht also aus den Ziffernpaaren $\{00, 01, \dots, 44\}$.

Die Frage, ob bei der Ersetzung der einzelnen Segmente des Klartextes eine einheitliche Strategie verfolgt wird oder ob diese von Segment zu Segment verändert wird, führt uns auf ein weiteres wichtiges Unterscheidungsmerkmal bei Substitutionen.

[§]Da nur 25 Plätze zur Verfügung stehen, muss bei Benutzung des lateinischen Alphabets entweder ein Buchstabe weggelassen oder ein Platz mit zwei Buchstaben besetzt werden.

Monoalphabetische Substitutionen ersetzen die einzelnen Klartextsegment unabhängig von ihrer Position im Klartext.

Polyalphabetische Substitutionen verwenden dagegen eine variable Ersetzungsregel, auf die sich auch die bereits verarbeiteten Klartextsegmente auswirken.

Die Bezeichnung „monoalphabetisch“ bringt zum Ausdruck, dass der Ersetzungsmechanismus auf einem einzelnen Alphabet beruht (sofern wir das Klartextalphabet als bekannt voraussetzen). Die von Caesar benutzte Chiffriermethode kann beispielsweise vollständig durch Angabe des Ersetzungsalphabets

$$\{D, E, F, G, W, \dots, Y, Z, A, B, C\}$$

beschrieben werden. Auch im Fall, dass nicht einzelne Zeichen, sondern ganze Buchstabengruppen auf einmal ersetzt werden, genügt im Prinzip ein einzelnes Alphabet zur Beschreibung. Hierzu sortiert man die Klartexteinheiten, auf denen der Ersetzungsmechanismus operiert, und bildet die Folge (sprich: das Alphabet) der zugeordneten Kryptotextsegmente.

Monoalphabetische Chiffrierverfahren ersetzen meist Texteinheiten einer festen Länge $l \geq 1$ durch Kryptotextsegmente derselben Länge.

Definition 38 (Blockchiffre). Sei A ein beliebiges Alphabet und es gelte $M = C = A^l$, $l \geq 1$. Eine **Blockchiffre** realisiert für jeden Schlüssel $k \in K$ eine bijektive Abbildung g auf A^l und es gilt

$$E(k, x) = g(x) \quad \text{und} \quad D(k, y) = g^{-1}(y)$$

für alle $x \in M$ und $y \in C$. Im Fall $l = 1$ spricht man auch von einer **einfachen Substitutionschiffre**.

Polyalphabetische Substitutionen greifen im Wechsel auf verschiedene Ersetzungsalphabete zurück, so dass unterschiedliche Vorkommen eines Zeichens (oder einer Zeichenkette) auch auf unterschiedliche Art ersetzt werden können. Welches Ersetzungsalphabet wann an der Reihe ist, wird dabei in Abhängigkeit von der Länge oder der Gestalt des bereits verarbeiteten Klartextes bestimmt.

Fast alle polyalphabetischen Chiffrierverfahren operieren – genau wie monoalphabetische Substitutionen – auf Klartextblöcken einer festen Länge l , die sie in Kryptotextblöcke einer festen Länge l' überführen, wobei meist $l = l'$ ist. Da diese Blöcke jedoch vergleichsweise kurz sind, kann der Klartext der Chiffrierfunktion ungepuffert zugeführt werden. Man nennt die einzelnen Klartextblöcke in diesem Zusammenhang auch nicht ‚Blöcke‘ sondern ‚Zeichen‘ und spricht von **sequentiellen Chiffren** oder von **Stromchiffren**.

Definition 39 (Stromchiffre). Sei A ein beliebiges Alphabet und sei $M = C = A^l$ für eine natürliche Zahl $l \geq 1$. Weiterhin seien K und \hat{K} Schlüsselräume. Eine **Stromchiffre** wird durch eine Verschlüsselungsfunktion $E : \hat{K} \times M \rightarrow C$ und einen Schlüsselstromgenerator $g : K \times A^* \rightarrow \hat{K}$ beschrieben. Der Generator g erzeugt aus einem externen Schlüssel $k \in K$ für einen Klartext $x = x_0 \dots x_{n-1}$, $x_i \in M$, eine Folge $\hat{k}_0, \dots, \hat{k}_{n-1}$ von internen Schlüsseln $\hat{k}_i = g(k, x_0 \dots x_{i-1}) \in \hat{K}$, unter denen x in den Kryptotext

$$E_g(k, x) = E(\hat{k}_0, x_0) \dots E(\hat{k}_{n-1}, x_{n-1})$$

überführt wird.

Der interne Schlüsselraum kann also wie bei der Blockchiffre eine maximale Größe von $(m^l)!$ annehmen (im häufigen Spezialfall $l = 1$ also $m!$). Die Aufgabe des Schlüsselstromgenerators g besteht darin, aus dem externen Schlüssel k und dem bereits verarbeiteten Klartext $x_0 \dots x_{i-1}$ den aktuellen internen Schlüssel \hat{k}_i zu berechnen. Die bisher betrachteten Stromchiffren benutzen z.B. die folgenden Schlüsselstromgeneratoren.

Stromchiffre	Chiffrierfunktionen	Schlüsselstromgenerator
Vigenère	$E(\hat{k}, x) = x + \hat{k}$	$g(k_0 \dots k_{d-1}, x_0 \dots x_{i-1}) = k_{(i \bmod m)}$
Beaufort	$E(\hat{k}, x) = \hat{k} - x$	$g(k_0 \dots k_{d-1}, x_0 \dots x_{i-1}) = k_{(i \bmod m)}$
Autokey mit Klartext- Schlüsselstrom	$E(\hat{k}, x) = x + \hat{k}$	$g(k_0 \dots k_{d-1}, x_0 \dots x_{i-1}) = \begin{cases} k_i, & i < d \\ x_{i-d}, & i \geq d \end{cases}$
Autokey mit Kryptotext- Schlüsselstrom	$E(\hat{k}, x) = x + \hat{k}$	$g(k_0 \dots k_{d-1}, x_0 \dots x_{i-1}) = \begin{cases} k_i, & i < d \\ y_{i-d}, & i \geq d \end{cases}$ $= k_{(i \bmod d)} + \sum_{j=1}^{\lfloor i/d \rfloor} x_{i-jd}$

Bei der Vigenère- und Beaufortchiffre hängt der Schlüsselstrom nicht vom Klartext, sondern nur vom externen Schlüssel k ab, d.h. sie sind **synchron**. Die Autokey-Chiffren sind dagegen **asynchron** (und aperiodisch).

Gespreizte Substitutionen

Bei den bisher betrachteten Substitutionen haben die einzelnen Blöcke, aus denen der Kryptotext zusammengesetzt wird, eine einheitliche Länge. Es liegt nahe, einem Gegner die unbefugte Rekonstruktion des Klartextes dadurch zu erschweren, dass man Blöcke unterschiedlicher Länge verwendet. Man spricht hierbei auch von einer **Spreizung** (*straddling*) des Kryptotextalphabets. Ein bekanntes Beispiel für diese Technik ist die sogenannte Spionage-Chiffre, die vorzugsweise von der ehemaligen sowjetischen Geheimpolizei NKWD (*Naródný Komissariát Wnutrennich Del*; zu deutsch: Volkskommissariat des Innern) benutzt wurde.

Beispiel 40. Bei der **Spionage-Chiffre** wird in die erste Zeile einer 3×10 -Matrix ein Schlüsselwort w geschrieben, welches keinen Buchstaben mehrfach enthält und eine Länge von 6 bis 8 Zeichen hat (also zum Beispiel **SPIONAGE**). Danach werden die anderen beiden Zeilen der Matrix mit den restlichen Klartextbuchstaben (etwa in alphabetischer Reihenfolge) gefüllt.

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

◁

Man überzeugt sich leicht davon, dass sich die von der Spionage-Chiffre generierten Kryptotexte wieder eindeutig dechiffrieren lassen, da die Kryptotextsegmente 1, 2, ..., 8, 01, 02, ..., 08, 91, 92, ..., 98, die für die Klartextbuchstaben eingesetzt werden, die **Fano-Bedingung** erfüllen: Keines von ihnen bildet den Anfang eines anderen. Da die Nummern

5 und 8 der beiden letzten Spalten der Matrix auch als Zeilennummern verwendet werden, liefert dies auch eine Erklärung dafür, warum keine Schlüsselwortbuchstaben in die beiden letzten Spalten eingetragen werden dürfen.

Verwendung von Blendern und Homophonen

Die Verwendung von gespreizten Chiffren zielt offenbar darauf ab, die „Fuge“ zwischen den einzelnen Kryptotextsegmenten, die von unterschiedlichen Klartextbuchstaben herühren, zu verdecken, um dem Gegner eine unbefugte Dechiffrierung zu erschweren. Dennoch bietet die Spionage-Chiffre noch genügend Angriffsfläche, da im Klartext häufig vorkommende Wortmuster auch im Kryptotext zu Textwiederholungen führen.

Eine Möglichkeit, diese Muster aufzubrechen, besteht darin, Blender in den Klartext einzustreuen. Abgesehen davon, dass das Entfernen der Blender auch für den rechtmäßigen Empfänger mit Mühe verbunden ist, muss für den Zugewinn an Sicherheit auch mit einer Expansion des Kryptotextes bezahlt werden.

Ist man bereit, dies in Kauf zu nehmen, so gibt es auch noch eine wirksamere Methode, die Übertragung struktureller und statistischer Klartextmerkmale auf den Kryptotext abzumildern. Die Idee dabei ist, zur Chiffrierung der einzelnen Klartextzeichen a nicht nur jeweils eines, sondern eine Menge $H(a)$ von Chiffrezeichen vorzusehen, und daraus für jedes Vorkommen von a im Klartext eines auszuwählen (am besten zufällig). Da alle Zeichen in $H(a)$ für dasselbe Klartextzeichen stehen, werden sie auch **Homophone** genannt.

Definition 41 (homophonen Substitutionschiffre). Sei A ein Klartextalphabet und sei $M = A$. Weiter sei C ein Kryptotextraum der Größe $\|C\| > \|A\| = m$. In einer (einfachen) **homophonen Substitutionschiffre** beschreibt jeder Schlüssel $k \in K$ eine Zerlegung von C in m disjunkte Mengen $H(a)$, $a \in A$.

Um ein Zeichen $a \in A$ unter k zu chiffrieren, wird nach einer bestimmten Methode ein Homophon y aus der Menge $H(a)$ gewählt und für a eingesetzt.

Durch den Einsatz einer homophonen Substitution wird also erreicht, dass verschiedene Vorkommen eines Klartextzeichens auch auf unterschiedliche Weise ersetzt werden können. Damit der Empfänger den Kryptotext auch wieder eindeutig dechiffrieren kann, dürfen sich die Homophonmengen zweier verschiedener Klartextzeichen aber nicht überlappen. Daher kann es nicht vorkommen, dass zwei verschiedene Klartextbuchstaben durch dasselbe Geheimtextzeichen ersetzt werden. Man beachte, dass der Chiffriervorgang $x \mapsto E(k, x)$ nicht durch eine Funktion beschreibbar ist, da derselbe Klartext x in mehrere verschiedene Kryptotexte y übergehen kann.

Durch eine geringfügige Modifikation der Polybios-Chiffre lässt sich die folgende bipartite homophone Chiffre erhalten.

Beispiel 42 (homophone Substitution). Sei $A = \{\mathbf{A}, \dots, \mathbf{Z}\}$, $B = \{\mathbf{0}, \dots, \mathbf{9}\}$ und $C = \{\mathbf{00}, \dots, \mathbf{99}\}$.

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

HOMOPHON \rightsquigarrow 82 03 88 53 17 32 08 98

Genau wie bei Polybios wird eine 5×5 -Matrix M als Schlüssel benutzt. Die Zeilen und Spalten von M sind jedoch nicht nur mit jeweils einer, sondern mit zwei Ziffern versehen, so dass jeder Klartextbuchstabe x über vier verschiedene Koordinatenpaare ansprechbar ist. Der Kryptotextrraum wird durch M also in 25 Mengen $H(a)$, $a \in A$, mit je 4 Homophonen partitioniert. ◁

Wie wir noch sehen werden, sind homophone Chiffrierungen auch deshalb schwerer zu brechen, weil durch sie die charakteristische Häufigkeitsverteilung der Klartextbuchstaben zerstört wird. Dieser Effekt kann dadurch noch verstärkt werden, dass man für häufig vorkommende Klartextzeichen a eine entsprechend größere Menge $H(a)$ an Homophonen vorsieht. Damit lässt sich erreichen, dass die Verteilung der im Geheimtext auftretenden Zeichen weitgehend nivelliert wird.

Beispiel 43 (homophone Substitution, verbesserte Version). Ist $p(a)$ die Wahrscheinlichkeit, mit der ein Zeichen $a \in A$ in der Klartextsprache auftritt, so sollte $\|H(a)\| \approx 100 \cdot p(a)$ sein.

a	$p(a)$	$H(a)$
A	0.0647	{15, 26, 44, 59, 70, 79}
B	0.0193	{01, 84}
C	0.0268	{13, 28, 75}
D	0.0483	{02, 17, 36, 60, 95}
E	0.1748	{04, 08, 12, 30, 43, 46, 47, 53, 61, 67, 69, 72, 80, 86, 90, 92, 97}
⋮	⋮	⋮

Da der Buchstabe **A** im Deutschen beispielsweise mit einer Wahrscheinlichkeit von $p(\mathbf{A}) = 0.0647$ auftritt, sind für ihn sechs verschiedene Homophone vorgesehen. ◁

Um den Suchaufwand bei der Dechiffrierung zu reduzieren, empfiehlt es sich, eine 10×10 -Matrix anzulegen, in der jeder Klartextbuchstabe a an allen Stellen vorkommt, deren Koordinaten in $H(a)$ enthalten sind.

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

HOMOPHON \rightsquigarrow 56 98 63 34 55 29 16 68

Obiges π hat beispielsweise die Zyklendarstellung

$$\pi = (1\ 4\ 3)\ (2\ 6)\ (5) \text{ oder } \pi = (1\ 4\ 3)\ (2\ 6),$$

wenn, wie allgemein üblich, Einerzyklen weggelassen werden. Daraus erhalten wir unmittelbar π^{-1} zu

$$\pi^{-1} = (3\ 4\ 1)\ (6\ 2) \text{ oder } (1\ 3\ 4)\ (2\ 6),$$

wenn wir jeden Zyklus mit seinem kleinsten Element beginnen lassen und die Zyklen nach der Größe dieser Elemente anordnen. ◁

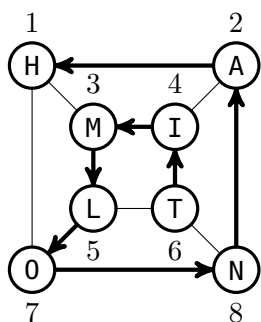
Beispiel 48. Bei der **Matrix-Transposition** wird der Klartext zeilenweise in eine $k \times m$ -Matrix eingelesen und der Kryptotext spaltenweise gemäß einer Spaltenpermutation π , die als Schlüssel dient, wieder ausgelesen. Für $\pi = (1\ 4\ 3)\ (2\ 6)$ wird also zuerst Spalte $\pi(1) = 4$, dann Spalte $\pi(2) = 6$ und danach Spalte $\pi(3) = 1$ usw. und zuletzt Spalte $\pi(6) = 2$ ausgelesen.

3	6	4	1	5	2
D	I	E	S	E	R
K	L	A	R	T	E
X	T	I	S	T	N
I	C	H	T	S	E
H	R	L	A	N	G

DIESER KLARTEXT IST NICHT SEHR LANG
 \rightsquigarrow SRSTA RENEG DKXIH EAIHL ETTSN ILTCR

◁

Beispiel 49. Bei der **Weg-Transposition** wird als Schlüssel eine Hamiltonlinie in einem Graphen mit den Knoten $1, \dots, l$ benutzt. (Eine Hamiltonlinie ist eine Anordnung aller Knoten, in der je zwei aufeinanderfolgende Knoten durch eine Kante verbunden sind.) Der Klartextblock $x_1 \dots x_l$ wird gemäß der Knotennumerierung in den Graphen eingelesen und der zugehörige Kryptotext entlang der Hamiltonlinie wieder ausgelesen.



HAMILTON \rightsquigarrow TIMLONAH

◁

Es ist leicht zu sehen, dass sich jede Blocktransposition durch eine Hamiltonlinie in einem geeigneten Graphen realisieren lässt. Der Vorteil, eine Hamiltonlinie als Schlüssel zu benutzen, besteht offenbar darin, dass man sich den Verlauf einer Hamiltonlinie bildhaft vorstellen und daher besser einprägen kann als eine Zahlenfolge.

Sehr beliebt ist auch die Methode, eine Permutationen in Form eines **Schlüsselworts** (oder einer aus mehreren Wörtern bestehenden **Schlüsselphrase**) im Gedächtnis zu behalten. Aus einem solchen Schlüsselwort lässt sich die zugehörige Permutation σ leicht rekonstruieren, indem man das Wort auf Papier schreibt und in der Zeile darunter für jeden einzelnen Buchstaben seine Position i innerhalb des Wortes vermerkt.

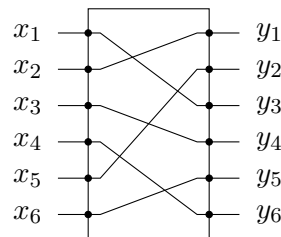
Schlüsselwort für σ	C A E S A R
i	1 2 3 4 5 6
$\sigma(i)$	3 1 4 6 2 5
Zyklendarstellung von σ	(1 3 4 6 5 2)

DIE BLOCKLAENGE IST SECHS \rightsquigarrow
EDBOIL L CANKE IGSSET EXCSYH

Die Werte $\sigma(i)$, die σ auf diesen Nummern annimmt, werden nun dadurch ermittelt, dass man die Schlüsselwort-Buchstaben in alphabetischer Reihenfolge durchzählt. Dabei werden mehrfach vorkommende Buchstaben gemäß ihrer Position im Schlüsselwort an die Reihe genommen. Alternativ kann man auch alle im Schlüsselwort wiederholt vorkommenden Buchstaben streichen, was im Fall des Schlüsselworts **CAESAR** auf eine Blocklänge von 5 führen würde.

1.8 Realisierung von Blocktranspositionen und einfachen Substitutionen

Abschließend möchten wir eine einfache elektronische Realisierungsmöglichkeit von Blocktranspositionen erwähnen, die auf binär kodierten Klartexten operieren (d.h. $A = \{0, 1\}$). Um einen Binärblock $x_1 \cdots x_l$ der Länge l zu permutieren, müssen die einzelnen Bits lediglich auf l Leitungen gelegt und diese gemäß π in einer sogenannten **Permutationsbox** (kurz **P-Box**) vertauscht werden.



Die Implementierung einer solchen P-Box kann beispielsweise auf einem VLSI-Chip erfolgen. Allerdings kann hierbei für größere Werte von l aufgrund der hohen Zahl von Überkreuzungspunkten ein hoher Flächenbedarf anfallen.

Blocktranspositionen können auch leicht durch Software als eine Folge von Zuweisungen

$$Y1 := X2; \quad Y2 := X5; \quad \dots \quad Y6 := X4;$$

implementiert werden. Bei großer Blocklänge und sequentieller Abarbeitung erfordert diese Art der Implementierung jedoch einen relativ hohen Zeitaufwand.

Von Alberti stammt die Idee, das Klartext- und Kryptotextalphabet auf zwei konzentrischen Scheiben unterschiedlichen Durchmessers anzuordnen. In Abbildung 1.1 ist gezeigt, wie sich mit einer solchen Drehscheibe beispielsweise die additive Chiffre realisieren lässt. Zur Einstellung des Schlüssels k müssen die Scheiben so gegeneinander verdreht werden, dass der Schlüsselbuchstabe a_k auf der inneren Scheibe mit dem Klartextzeichen $a_0 = A$ auf der äußeren Scheibe zur Deckung kommt. Auf der Drehscheibe in Abbildung 1.1 ist beispielsweise der Schlüssel $k = 3$ eingestellt, das heißt, $a_k = D$. Die Verschlüsselung geschieht nun durch bloßes Ablesen der zugehörigen Kryptotextzeichen auf der inneren Scheibe, so dass von der Drehfunktion der Scheiben nur bei einem Schlüsselwechsel Gebrauch gemacht wird.

Aufgrund ihrer engen Verwandtschaft mit der Klasse der Blocktranspositionen lassen sich einfache Substitutionen auch mit Hilfe einer P-Box realisieren (vergleiche Abbildung).

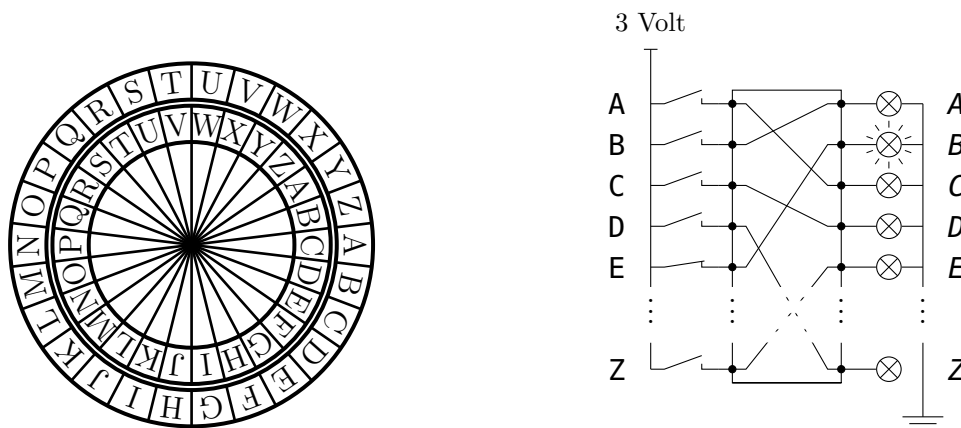


Abbildung 1.1: Realisierung von einfachen Substitutionen mit einer Drehscheibe und mit Hilfe von Steckverbindungen.

Hierfür können beispielsweise zwei Steckkontakteleisten verwendet werden. Der aktuelle Schlüssel wird in diesem Fall durch Verbinden der entsprechenden Kontakte mit elektrischen Kabeln eingestellt (siehe Abbildung 1.1). Um etwa den Klartextbuchstaben **E** zu verschlüsseln, drückt man auf die entsprechende Taste, und das zugehörige Kryptotextzeichen **B** wird im selben Moment durch ein aufleuchtendes Lämpchen signalisiert.

Schließlich lassen sich Substitutionen auch leicht durch Software realisieren. Hierzu wird ein Feld (*array*) deklariert, dessen Einträge über die Klartextzeichen $x \in A$ adressierbar sind. Das mit x indizierte Feldelemente enthält das Kryptotextzeichen, durch welches x beim Chiffriervorgang zu ersetzen ist.

Ein Nachteil hierbei ist, dass das Feld nach jedem Schlüsselwechsel neu beschrieben werden muss. Um dies zu umgehen, kann ein zweidimensionales Feld deklariert werden, dessen Einträge zusätzlich über den aktuellen Schlüsselwert k adressierbar sind. Ist genügend Speicherplatz vorhanden, um für alle $x \in A$ und alle $k \in K$ die zugehörigen Kryptotextzeichen $E(k, x)$ abspeichern zu können, so braucht das Feld nur einmal initialisiert und danach nicht mehr geändert werden.

Schlüsselwert	Klartextbuchstabe			
	A	B	...	Z
0	<i>U</i>	<i>H</i>	...	<i>C</i>
1	<i>E</i>	<i>H</i>	...	<i>A</i>
⋮	⋮	⋮	⋮	⋮
63	<i>Y</i>	<i>F</i>	...	<i>W</i>

2 Kryptoanalyse der klassischen Verfahren

2.1 Klassifikation von Angriffen gegen Kryptosysteme

Die Erfolgsaussichten eines Angriffs gegen ein Kryptosystem hängen sehr stark davon ab, wie gut die Ausgangslage ist, in der sich der Gegner befindet. Prinzipiell sollte man die Fähigkeiten des Gegners genauso wenig unterschätzen wie die Unvorsichtigkeit der Anwender von Kryptosystemen. Bereits vor mehr als einem Jahrhundert postulierte Kerckhoffs, dass die Frage der Sicherheit keinesfalls von irgendwelchen obskuren Annahmen über den Wissensstand des Gegners abhängig gemacht werden darf.

Goldene Regel für Kryptosystem-Designer (Kerckhoffs' Prinzip)

*Unterschätze niemals den Kryptoanalytiker. Gehe insbesondere immer von der Annahme aus, dass dem Gegner das angewandte System bekannt ist.**

In der folgenden Liste sind eine Reihe von Angriffsszenarien mit zunehmender Gefährlichkeit aufgeführt. Auch wenn nicht alle Eventualitäten eines Angriffs vorhersehbar sind, so vermittelt diese Aufstellung doch eine gute Vorstellung davon, welchen unterschiedlichen Bedrohungen ein Kryptosystem im praktischen Einsatz ausgesetzt sein kann.

Angriff bei bekanntem Kryptotext (*ciphertext-only attack*)

Der Gegner fängt Kryptotexte ab und versucht, allein aus ihrer Kenntnis Rückschlüsse auf die zugehörigen Klartexte oder auf die benutzten Schlüssel zu ziehen.

Angriff bei bekanntem Klartext (*known-plaintext attack*)

Der Gegner ist im Besitz von einigen zusammengehörigen Klartext-Kryptotext-Paaren. Hierdurch wird erfahrungsgemäß die Entschlüsselung weiterer Kryptotexte oder die Bestimmung der benutzten Schlüssel wesentlich erleichtert.

Angriff bei frei wählbarem Klartext (*chosen-plaintext attack*)

Der Angriff des Gegners wird zusätzlich dadurch erleichtert, dass er in der Lage ist (oder zumindest eine Zeit lang war), sich zu Klartexten seiner Wahl die zugehörigen Kryptotexte zu besorgen. Kann hierbei die Wahl der Kryptotexte in Abhängigkeit von zuvor erhaltenen Verschlüsselungsergebnissen getroffen werden, so spricht man von einem **Angriff bei adaptiv wählbarem Klartext (*adaptive chosen-plaintext attack*)**.

Angriff bei frei wählbarem Kryptotext (*chosen-ciphertext attack*)

Vor der Beobachtung des zu entschlüsselnden Kryptotextes konnte sich der Gegner zu Kryptotexten seiner Wahl die zugehörigen Klartexte besorgen, ohne dabei jedoch in den Besitz des Dechiffrierschlüssels zu kommen (**Mitternachtsattacke**). Das dabei erworbene Wissen steht ihm nun bei der Durchführung seines Angriffs zur Verfügung. Auch in diesem Fall können sich die Erfolgsaussichten des Gegners erhöhen, wenn ein **Angriff bei adaptiv wählbarem Kryptotext (*adaptive chosen-ciphertext attack*)** möglich ist, also der Kryptotext in Abhängigkeit von den zuvor erzielten Entschlüsselungsergebnissen wählbar ist.

*Diese Annahme ergibt sich meist schon aus der Tatsache, dass die Prinzipien fast aller heute im Einsatz befindlichen Kryptosysteme allgemein bekannt sind.

Angriff bei frei (oder adaptiv) wählbarem Text (*chosen-text attack*)

Sowohl Klartexte als auch Kryptotexte sind frei (oder sogar adaptiv) wählbar.

Ohne Frage ist ein Kryptosystem, das bereits bei einem Angriff mit bekanntem Kryptotext Schwächen erkennen lässt, für den praktischen Einsatz vollkommen ungeeignet. Tatsächlich müssen aber an ein praxistaugliches Kryptosystem noch weit höhere Anforderungen gestellt werden. Denn häufig unterlaufen den Anwendern sogenannte **Chiffrierfehler**, die einen Gegner leicht in eine sehr viel günstigere Ausgangsposition versetzen als dies sonst der Fall wäre. So ermöglicht beispielsweise das Auftreten stereotyper Klartext-Formulierungen einen Angriff bei bekanntem Klartext, sofern der Gegner diese Formulierungen kennt oder auch nur errät. Begünstigt durch derartige Unvorsichtigkeiten, die im praktischen Einsatz nicht vollständig vermeidbar sind, können sich selbst winzige Konstruktionsschwächen eines Kryptosystems sehr schnell zu einer ernsthaften Bedrohung der damit verfolgten Sicherheitsinteressen auswachsen. Die Geschichte der Kryptografie belegt sehr eindrucksvoll, dass es häufig die Anwender eines Kryptosystems selbst sind, die – im unerschütterlichen Glauben an seine kryptografische Stärke – dem Gegner zum Erfolg verhelfen.

Zusammenfassend lässt sich also festhalten, dass die Gefährlichkeit von Angriffen, denen ein Kryptosystem im praktischen Einsatz ausgesetzt ist, kaum zu überschätzen ist. Andererseits kann selbst das beste Kryptosystem keinen Schutz vor einer unbefugten Dechiffrierung mehr bieten, wenn es dem Gegner etwa gelingt, in den Besitz des geheimen Schlüssels zu kommen – sei es aus Unachtsamkeit der Anwender oder infolge einer Gewaltandrohung des Gegners (**kompromittierte Schlüssel**).

2.2 Kryptoanalyse von einfachen Substitutionschiffren

Durch eine Häufigkeitsanalyse können insbesondere einfache Substitutionen g leicht gebrochen werden, sofern die einzelnen Buchstaben a in der benutzten Klartextsprache mit voneinander differierenden Häufigkeiten $p(a)$ auftreten (vergleiche Tabelle 2.1). Selbst wenn, was insbesondere bei kurzen Texten zu erwarten ist, die tatsächliche Häufigkeitsverteilung nur in etwa der vom Gegner angenommenen Verteilung entspricht, reduziert sich dadurch die Zahl der in Frage kommenden einfachen Substitutionen ganz erheblich. Berechnet man die relativen Häufigkeiten h der Kryptotextbuchstaben im Kryptotext, so gilt $p(a) \approx h(g(a))$ (vorausgesetzt der Kryptotext ist genügend lang). Für die Schilderung einer nach dieser Methode durchgeführten Kryptoanalyse sei auf die Erzählung „Der Goldkäfer“ von Edgar Allan Poe verwiesen.

Tabelle 2.1: Einteilung von Buchstaben in Cliques mit vergleichbaren Häufigkeitswerten.

	Deutsch	Englisch	Französisch
sehr häufig	E	E	E
häufig	N I R S A T	T A O I N S R H	N A R S I T U
durchschnittlich	D H U L G O C M	L D C U M F	L D C M P
selten	B F W K Z P V	P G W Y B V K	V F B G Q H X
sehr selten	J Y X Q	X J Q Z	J Y Z K W

Manche der bisher betrachteten Chiffrierverfahren verwenden einen so kleinen Schlüsselraum, dass ohne großen Aufwand eine vollständige Schlüsselsuche ausgeführt werden kann.

Beispiel 50 (vollständige Schlüsselsuche). *Es sei bekannt, dass das Kryptotextstück $y = \text{SAXP}$ mit einer additiven Chiffre erzeugt wurde ($K = A = B = A_{\text{lat}}$). Entschlüsseln wir y probeweise mit allen möglichen Schlüsselwerten, so erhalten wir folgende Zeichenketten.*

k	B	C	D	E	F	G	H	I	J	K	L	M	
$D(k, y)$	RZWO	QYVN	PXUM	OWTL	NVSK	MURJ	LTQI	KSPH	JROG	IQNF	HPME	GOLD	
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	FNKC	EMJB	DLIA	CKHZ	BJGY	AIFX	ZHEW	YGDV	XFCU	WEBT	VDAS	UCZR	TBYQ

Unter diesen springen vor allem die beiden Klartextkandidaten $x = \text{GOLD}$ (Schlüsselwert $k = M$) und $x = \text{WEBT}$ ($k = W$) ins Auge. ◀

Ist $s = \|K\|$ die Größe des Schlüsselraums, so kann der Gegner bei bekanntem Kryptotext y die Suche nach dem zugehörigen Klartext x auf eine Menge von maximal s Texten x_1, \dots, x_s beschränken. Daneben hat der Gegner ein gewisses *a priori* Wissen über den Klartext, wie zum Beispiel dass er in deutscher Sprache verfasst ist, das es ihm gestattet, einen Großteil der Texte x_i auszuschließen. Ferner erscheinen aufgrund dieses Hintergrundwissens manche der übrig gebliebenen Klartextkandidaten plausibler als andere (sofern nicht nur ein einziger übrig bleibt). Mit jedem Text x_i , der nicht als Klartext in Frage kommt, kann auch mindestens ein Schlüssel ausgeschlossen werden. Sind noch mehrere Schlüsselwerte möglich, so kann weiteres Kryptotextmaterial Klarheit bringen. Manchmal hilft aber auch eine Inspektion der verbliebenen Schlüsselwerte weiter, etwa wenn der Schlüssel nicht rein zufällig erzeugt wurde, sondern aus einem einprägsamen Schlüsselwort ableitbar ist.

Meist kennt der Gegner zumindest die Sprache, in der der gesuchte Klartext abgefasst ist. Mit zunehmender Länge gleichen sich die Häufigkeitsverteilungen der Buchstaben in natürlichsprachigen Texten einer „Grenzverteilung“ an, die in erster Linie von der benutzten Sprache und nur in geringem Umfang von der Art des Textes abhängt. Diese Verteilungen weisen typischerweise eine sehr starke Ungleichmäßigkeit auf, was darauf zurückzuführen ist, dass in natürlichen Sprachen relativ viel Redundanz enthalten ist.

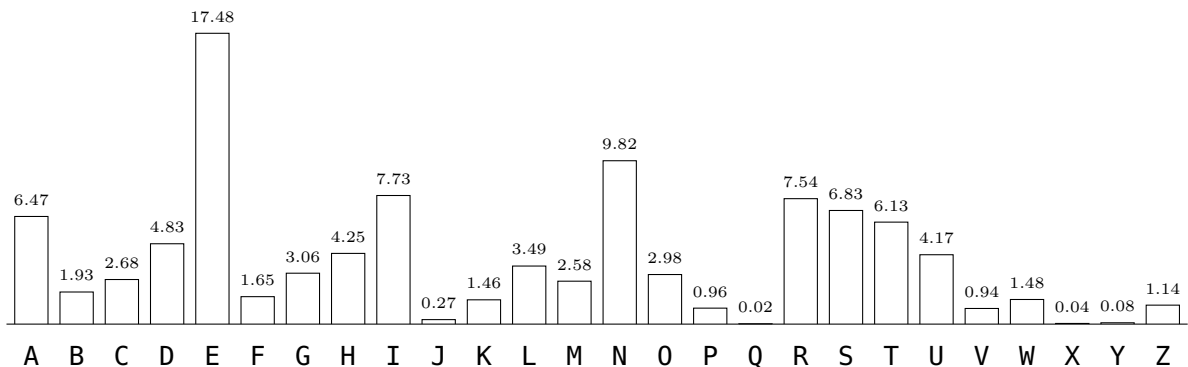


Abbildung 2.1: Häufigkeitsverteilung der Einzelbuchstaben im Deutschen (in %).

Die Abbildungen 2.1, 2.2 und 2.3, zeigen typische Verteilungen von Einzelbuchstaben in der deutschen, englischen und französischen Sprache (ohne Berücksichtigung von

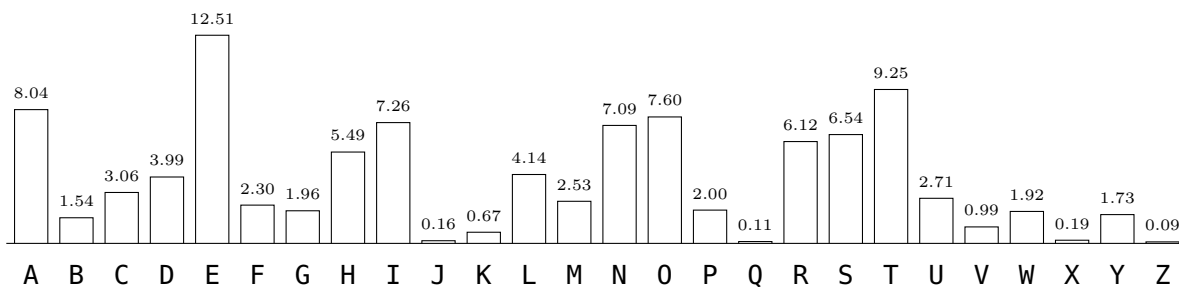


Abbildung 2.2: Häufigkeitsverteilung der Buchstaben im Englischen (in %).

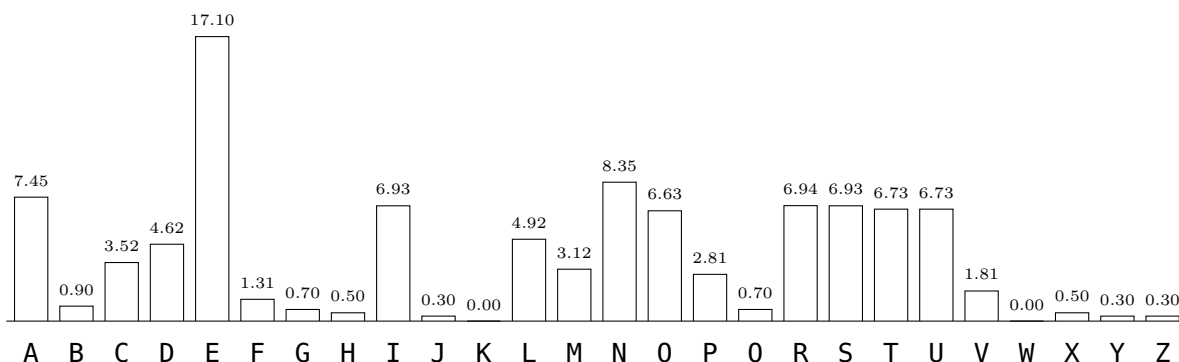


Abbildung 2.3: Häufigkeitsverteilung der Buchstaben im Französischen (in %).

Interpunktions- und Leerzeichen). Ein typischer deutscher Text besteht demnach zu 62% aus den sieben häufigsten Zeichen E, N, I, R, S, A, T (das sind nicht einmal 27% der Klartextzeichen).

Bei additiven Chiffren reicht es oftmals, den häufigsten Buchstaben im Kryptotext zu bestimmen, und davon den häufigsten Buchstaben der Klartextsprache zu subtrahieren, um den Schlüssel k zu erhalten. Bei affinen Chiffren müssen gewöhnlich nur die beiden häufigsten Buchstaben bestimmt werden; dadurch erhält man zwei Verschlüsselungsgleichungen. Dieses Gleichungssystem muss gelöst werden, und man erhält das gesuchte Schlüsselpaar.

Beispiel 51 (Analyse einer affinen Chiffre mittels Buchstabenhäufigkeiten). *Es sei bekannt, dass sich hinter dem Kryptotext*

*laoea ehoap hwvae ixobg jcbho thlob lokhe ixope vbcix ockix qoppo boapo
mohqc euogk opeho jhkpl eappj seobe ixoap opmcu*

ein deutscher Klartext verbirgt, der mit einer affinen Chiffre verschlüsselt wurde. Berechnen wir für jedes Chiffrezeichen y die (absolute) Häufigkeit $H(y)$ seines Auftretens,

y	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
$H(y)$	7	6	5	0	10	0	2	8	5	3	4	4	2	0	19	11	2	0	1	1	2	2	1	5	0	0

so liegt die Vermutung nahe, dass das am häufigsten vorkommende Chiffrezeichen O für das Klartextzeichen E und das am zweithäufigsten vorkommende P für N steht. Unter

dieser Annahme kann der gesuchte Schlüssel $k = (b, c)$ als Lösung der beiden Gleichungen

$$b \cdot \mathbf{E} + c = \mathbf{0}$$

$$b \cdot \mathbf{N} + c = \mathbf{P}$$

bestimmt werden. Subtrahieren wir nämlich die erste von der zweiten Gleichung, so erhalten wir die Kongruenz $9 \cdot b \equiv_{26} 1$, woraus sich $b = 3$ und damit $c = 2$ ergibt. Tatsächlich weist der Schlüssel $k = (3, 2)$ nicht nur für die beiden Paare $(\mathbf{E}, \mathbf{0})$ und (\mathbf{N}, \mathbf{P}) , sondern auch für alle übrigen Paare (x, y) eine gute Übereinstimmung zwischen der Häufigkeit $H(y)$, mit der $y = E(k, x)$ im Kryptotext vorkommt, und der erwarteten Häufigkeit $E(x)$ auf, mit der x in einem typischen deutschen Text der Länge 100 vorkommt (die Tabelle zeigt die Werte von $E(x)$ gerundet):

y	O	P	E	H	A	B	C	X	I	L	K	J	U	M	G	V	Q	S	T	W	R	F	N	Z	Y	D
$H(y)$	19	11	10	8	7	6	5	5	5	4	4	3	2	2	2	2	2	1	1	1	0	0	0	0	0	0
$E(x)$	17	10	7	6	8	8	6	4	3	5	4	3	3	3	1	1	1	3	0	0	2	2	1	1	0	0
x	E	N	S	T	I	R	A	H	C	D	U	L	G	M	K	P	W	O	X	Y	F	B	V	Z	Q	J

◁

2.3 Kryptoanalyse von Blocktranspositionen

Mit Hilfe von Bigrammhäufigkeiten, die manchmal auch als Kontakthäufigkeiten bezeichnet werden, lassen sich Blocktranspositionen sehr leicht brechen, sofern genügend Kryptotext vorliegt. Ist die Blocklänge l bekannt, so trägt man hierzu den Kryptotext zeilenweise in eine Matrix $S = (s_{ij})$ mit l Spalten S_1, \dots, S_l ein. Da jede Zeile dieser Matrix aus dem zugehörigen Klartextblock mit derselben Permutation π erzeugt wurde, müssen die Spalten S_j jetzt nur noch in die „richtige“ Reihenfolge gebracht werden, um den gesuchten Klartext zu erhalten. Der Nachfolger S_k von S_j (bzw. der Vorgänger S_j von S_k) kann sehr gut anhand der Werte von $\hat{p}(S_j, S_k) = \sum_i p(s_{ij}, s_{ik})$ bestimmt werden.

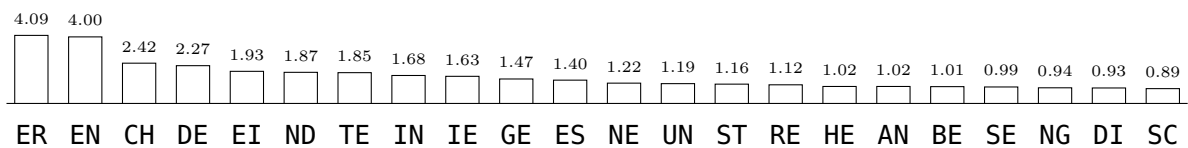


Abbildung 2.4: Die häufigsten Bigramme im Deutschen (Angaben in %).



Abbildung 2.5: Die häufigsten Bigramme im Englischen (in %; nach O.P. Meaker, 1939).

Beispiel 52 (Häufigkeitsanalyse von Bigrammen). Für den mit einer Blocktransposition (mit vermuteter Blocklänge 5) erzeugten Kryptotext

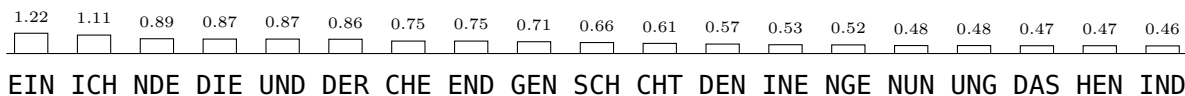


Abbildung 2.6: Die häufigsten Trigramme im Deutschen (in %).

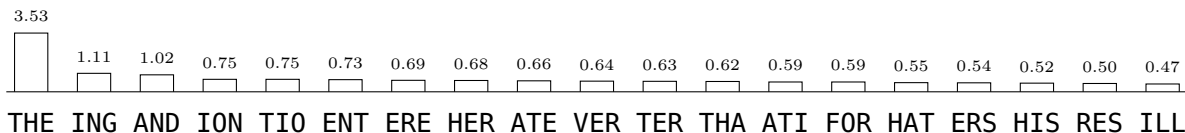


Abbildung 2.7: Die häufigsten Trigramme im Englischen (in %).

IHEHR BWEAN RNEII NRKEU ELNZK RXTAE VLOTR ENGIE

erhalten wir eine Matrix S mit den folgenden fünf Spalten.

S_1	S_2	S_3	S_4	S_5
I	H	E	H	R
B	W	E	A	N
R	N	E	I	I
N	R	K	E	U
E	L	N	Z	K
R	X	T	A	E
V	L	O	T	R
E	N	G	I	E

Um die richtige Vorgänger- oder Nachfolgerspalte von S_1 zu finden, bestimmen wir für jede potentielle Spalte S_j , $j = 2, \dots, 5$, wieviele der Bigramme $s_{ij}s_{i1}$ (bzw. $s_{i1}s_{ij}$) zu den 20 häufigsten (aus Abbildung 2.4) gehören.

					↓						↓			
S_2	S_3	S_4	S_5	S_1	S_2	S_3	S_4	S_5			S_2	S_3	S_4	S_5
H	E	H	R	I	H	E	H	R			H	E	H	R
W	E	A	N	B	W	E	A	N			W	E	A	N
N	E	I	I	R	N	E	I	I			N	E	I	I
R	K	E	U	N	R	K	E	U			R	K	E	U
L	N	Z	K	E	L	N	Z	K			L	N	Z	K
X	T	A	E	R	X	T	A	E			X	T	A	E
L	O	T	R	V	L	O	T	R			L	O	T	R
N	G	I	E	E	N	G	I	E			N	G	I	E
1	4	2	2		1	4	2	1			1	4	2	1

Da die beiden Spaltenpaare (S_3, S_1) und (S_1, S_3) jeweils vier häufige Bigramme bilden, können wir annehmen, dass im Klartext S_1 auf S_3 oder S_3 auf S_1 folgen muss. Entscheiden wir uns für die zweite Möglichkeit, so sollten wir als nächstes die Spaltenpaare (S_j, S_1) und (S_3, S_j) , $j = 2, 4, 5$ betrachten.

			↓			↓		
S_2	S_4	S_5	S_1	S_3	S_2	S_4	S_5	
H	H	R	I	E	H	H	R	
W	A	N	B	E	W	A	N	
N	I	I	R	E	N	I	I	
R	E	U	N	K	R	E	U	
L	Z	K	E	N	L	Z	K	
X	A	E	R	T	X	A	E	
L	T	R	V	O	L	T	R	
N	I	E	E	G	N	I	E	
1	2	2			1	1	5	

Aufgrund des hohen Wertes von $\hat{p}(S_3, S_5)$ können wir annehmen, dass auf S_3 die Spalte S_5 folgt. Im nächsten Schritt erhalten wir daher die folgende Tabelle.

			↓	↓			↓	↓
S_2	S_4	S_1	S_3	S_5	S_2	S_4		
H	H	I	E	R	H	H		
W	A	B	E	N	W	A		
N	I	R	E	I	N	I		
R	E	N	K	U	R	E		
L	Z	E	N	K	L	Z		
X	A	R	T	E	X	A		
L	T	V	O	R	L	T		
N	I	E	G	E	N	I		
1	2				2	1		

Diese lässt die Spaltenanordnung S_4, S_1, S_3, S_5, S_2 vermuten, welche tatsächlich auf den gesuchten Klartext führt:

S_4	S_1	S_3	S_5	S_2
H	I	E	R	H
A	B	E	N	W
I	R	E	I	N
E	N	K	U	R
Z	E	N	K	L
A	R	T	E	X
T	V	O	R	L
I	E	G	E	N

◁

2.4 Kryptoanalyse von polygrafischen Chiffren

Blocksysteme mit kleinem k (beispielsweise bigrafische Systeme) lassen sich ähnlich wie einfache Substitutionen durch Häufigkeitsanalysen brechen. Wird bei Hill-Chiffren k sehr groß gewählt, so ist eine solche statistische Analyse nicht mehr möglich. Das Hill-System kann dann zwar einem Kryptotextangriff widerstehen, jedoch kaum einem Angriff mit bekanntem Klartext und schon gar nicht einem Angriff mit gewähltem Klartext.

Angriff mit gewähltem Klartext O. B. d. A. sei $A = \{0, 1, \dots, m-1\}$. Bei einem GK-Angriff verschafft sich der Gegner den Kryptotext zu $100\dots 0, 010\dots 0, \dots, 0\dots 001 \in A^l$:

$$\begin{aligned} g(100\dots 0) &= k_{11} k_{12} \dots k_{1l} \\ g(010\dots 0) &= k_{21} k_{22} \dots k_{2l} \\ &\vdots \\ g(0\dots 001) &= k_{l1} k_{l2} \dots k_{ll} \end{aligned}$$

und erhält damit die Schlüsselmatrix k .

BK-Angriff (bekannter Klartext). Sind bei einem BK-Angriff ausreichend geeignete Klartext-Kryptotextpaare bekannt, so kann das Hill-System folgendermaßen gebrochen werden: Sind x_i, y_i ($i = 1, \dots, \mu$) Paare mit $x_i k = y_i$ und gilt $\text{ggT}(\det X, m) = 1$ für eine aus l Blöcken $x_i, i \in I$, als Zeilen gebildete Matrix X , so lässt sich die Schlüsselmatrix k zu $k = YX^{-1}$ bestimmen (Y ist die aus den Blöcken $y_i, i \in I$, gebildete Matrix).

2.5 Kryptoanalyse von polyalphabetischen Chiffren

Die Vigenère-Chiffre galt bis ins 19. Jahrhundert als sicher. Da der Schlüsselstrom bei der Vigenère-Chiffre periodisch ist, lassen sie sich mit statistischen Methoden ebenfalls leicht brechen, insbesondere wenn der Kryptotext im Verhältnis zur Periode d (Länge des Schlüsselwortes) genügend lang ist.

Bestimmung der Schlüsselwortlänge

Es gibt mehrere Methoden, eine Vigenère-Chiffre zu brechen, sobald die Länge des Schlüsselwortes bekannt ist. So kann man beispielsweise den Kryptotext zeilenweise in eine d -spaltige Matrix schreiben. Verfahrensbedingt wurden dann die einzelnen Spalten y_1, \dots, y_d durch eine monoalphabetische Substitution (genauer: durch eine Verschiebechiffre) verschlüsselt. Sie können daher einzeln wie eine additive Chiffre durch eine Häufigkeitsanalyse gebrochen werden. Hierbei liefert jede Spalte y_i einen Buchstaben k_i des Schlüsselwortes der Vigenère-Chiffre.

Zur Bestimmung der Schlüsselwortlänge betrachten wir zwei Vorgehensweisen: den Kasiski-Test und die Koinzidenzindex-Untersuchung.

Der Kasiski-Test. Die früheste generelle Methode zur Bestimmung der Periode bei der Vigenère-Chiffre stammt von Friedrich W. Kasiski (1860). Kommt ein Wort an zwei verschiedenen Stellen im Kryptotext vor, so kann es sein, dass die gleiche Klartextsequenz zweimal auf die gleiche Weise, d. h. mit der gleichen Schlüsselsequenz, verschlüsselt wurde. In diesem Fall ist die Entfernung δ der beiden Vorkommen ein Vielfaches der Periode d . Werden mehrere Paare mit verschiedenen Entfernungen δ_i gefunden, so liegt die Vermutung nahe, dass d gemeinsamer Teiler aller (oder zumindest vieler) δ_i ist, was die Anzahl der noch in Frage kommenden Werte für d stark einschränkt.

Beispiel 53 (Kasiski-Test).

$$\begin{array}{r} \text{DERERSTEUNDLETZTEVERS...} \quad (\text{Klartext } x) \\ + \text{KASKASKASKASKASKAS...} \quad (\text{Schlüsselstrom } \hat{k}) \\ \hline \text{NEJORKDEM\dot{X}DDOTR\dot{D}ENORK...} \quad (\text{Kryptotext } y) \end{array}$$

Da die Textstücke **ORK**, bzw. **DE** im Kryptotext in den Entfernungen $\delta_1 = 15$ und $\delta_2 = 9$ vorkommen, liegt die Vermutung nahe, dass die Periode $d = \text{ggT}(9, 15) = 3$ ist. \triangleleft

Koinzidenzindex-Untersuchungen. Zur Bestimmung der Periode d gibt es neben heuristischen Methoden auch folgenden statistischen Ansatz, der erstmals von William Frederick Friedman im Jahr 1920 beschrieben wurde. Er basiert auf der Beobachtung, dass eine längere Periode eine zunehmende *Glättung* der Buchstabenhäufigkeiten im Kryptotext bewirkt.

Definition 54 (Koinzidenzindex). Der **Koinzidenzindex** (engl. *index of coincidence*) eines Textes y der Länge n über dem Alphabet \mathcal{B} ist definiert als

$$IC(y) = \frac{1}{n \cdot (n - 1)} \cdot \sum_{a \in \mathcal{B}} H(a) \cdot (H(a) - 1).$$

Hierbei ist $H(a)$ die absolute Häufigkeit des Buchstabens a im Text y .

$IC(y)$ gibt also die Wahrscheinlichkeit an, mit der man im Text y an zwei zufällig gewählten Positionen den gleichen Buchstaben vorfindet. Er ist umso größer, je ungleichmäßiger die Häufigkeiten $H(a)$ sind (siehe unten).

Um die Periode d einer Vigenère-Chiffre zu bestimmen, schreibt man den Kryptotext y für $d = 1, 2, 3, \dots$ in eine Matrix mit d Spalten und berechnet für jede Spalte y_i den Koinzidenzindex $IC(y_i)$. Für genügend lange Kryptotexte ist dasjenige d , welches das maximale arithmetische Mittel der Spaltenindizes $IC(y_i)$ liefert mit hoher Wahrscheinlichkeit die gesuchte Periode. Enthält eine Spalte nämlich nur Kryptozeichen, die alle mit demselben Schlüsselbuchstaben k erzeugt wurden, so stimmt der Koinzidenzindex dieser Spalte mit dem Koinzidenzindex des zugehörigen Klartextes überein, nimmt also einen relativ großen Wert an. Wurden dagegen die Kryptozeichen einer Spalte mit unterschiedlichen Schlüsselbuchstaben generiert, so wird hierdurch eine Glättung der Häufigkeitsverteilung bewirkt, weshalb der Spaltenindex kleiner ausfällt.

Ist die Einzelbuchstabenverteilung $p : A \rightarrow [0, 1]$ der Klartextsprache bekannt, so kann der Suchraum für den Wert der Periode d erheblich eingeschränkt werden. Hierzu berechnet man den erwarteten Koinzidenzindex

$$E_{d,n}(IC) = E(IC(Y)),$$

wobei Y ein mittels einer Vigenère-Chiffre mit einem zufälligen Schlüsselwort der Länge d aus einem zufälligen Klartext der Länge n generierter Kryptotext ist. Im Fall $d = 1$ gilt $IC(y) = IC(x)$. Zudem können wir bei längeren Texten von den gegenseitigen Abhängigkeiten der Zeichen im Text absehen und erhalten

$$E_{1,\infty}(IC) = \sum_{a \in A} p(a)^2.$$

Dieser Wert wird auch als Koinzidenzindex der zugrunde liegenden Sprache bezeichnet.

Definition 55 (Koinzidenzindex einer Sprache). Der **Koinzidenzindex** IC_L einer Sprache mit Buchstabenverteilung $p : A \rightarrow [0, 1]$ ist definiert als

$$IC_L = \sum_{a \in A} p(a)^2.$$

IC_L ist zudem ein Maß für die Rauheit der Verteilung p :

Definition 56 (Rauheitsgrad; Measure of Roughness). Der **Rauheitsgrad** MR_L einer Sprache L mit Einzelbuchstabenverteilung p ist

$$MR_L = \sum_{a \in A} (p(a) - 1/m)^2 = \sum_{a \in A} p(a)^2 - 1/m = IC_L - 1/m,$$

wobei $m = \|A\|$ ist.

Beispiel 57. Für die englische Sprache ($m = 26$) gilt beispielsweise $IC_{\text{Englisch}} \approx 0.0687$ und $MR_{\text{Englisch}} \approx 0.0302$. \triangleleft

Übersteigt dagegen die Periode d die Klartextlänge n , so ist der Kryptotext bei zufälliger Wahl des Schlüsselwortes ebenfalls rein zufällig, was auf einen erwarteten Koinzidenzindex von

$$E_{d,n}(IC) = \sum_{a \in A} \|A\|^{-2} = \|A\|^{-1}, \quad d \geq n \geq 2$$

führt. Allgemein gilt

$$E_{d,n}(IC) = \frac{n-d}{d \cdot (n-1)} \cdot IC_L + \frac{n \cdot (d-1)}{d \cdot (n-1)} \cdot \|A\|^{-1}, \quad n \leq d,$$

da von den $\binom{n}{2} = n(n-1)/2$ möglichen Positionspaaren ungefähr $d \cdot \binom{n/d}{2} = n(n-d)/2d$ Paare nur eine Spalte und $\binom{d}{2} (n/d)^2 = n^2(d-1)/2d$ Paare zwei unterschiedliche Spalten betreffen.

Untenstehende Tabelle gibt den Erwartungswert $E_{d,n}(IC)$ des Koinzidenzindex für Kryptotexte der Länge $n = 100$ in Abhängigkeit von der Periodenlänge d einer Vigenère-Chiffre wieder (in Promille; Klartext ist ein zufällig gewählter Text der englischen Sprache mit 100 Buchstaben).

d	1	2	3	4	5	6	8	10	100
$E_{d,100}(IC)$	69	54	48	46	44	43	42	41	39

Beispiel 58. Berechnet sich der Koinzidenzindex eines Vigenère-Kryptotextes der Länge 100 zu 0.045, so liegt die Vermutung nahe, dass das verwendete Schlüsselwort die Länge vier oder fünf hat, falls y aus einem Klartext der englischen Sprache erzeugt wurde. \triangleleft

Der Koinzidenzindex kann auch Hinweise dafür liefern, mit welchem Kryptoverfahren ein vorliegender Kryptotext erzeugt wurde. Bei Transpositionschiffren sowie bei einfachen Substitutionen bleibt nämlich der Koinzidenzindex im Gegensatz zu polyalphabetischen und polygrafischen Verfahren erhalten. Erstere lassen sich von letzteren zudem dadurch unterscheiden, dass bei ihnen sogar die Buchstabenhäufigkeiten unverändert bleiben.

Zur Bestimmung des Schlüsselwortes bei bekannter Periode d kann auch wie folgt vorgegangen werden. Man schreibt den Kryptotext y in Spalten y_i auf und berechnet für $a \in A$ und $i = 1, \dots, d$ die relativen Häufigkeiten $h_i(a)$ von a in y_i . Da y_i aus dem Klartext durch Addition von k_i entstanden ist, kommt die Verteilung

$$h_i(a+k), \quad a \in A$$

für $k = k_i$ der Klartextverteilung $p(a)$, $a \in A$ näher als für $k \neq k_i$. Da

$$\alpha_i(k) := \sum_{a \in A} p(a) h_i(a+k)$$

ein Maß für die Ähnlichkeit der beiden Verteilungen $p(a)$ und $h_i(a+k)$ ist (siehe Übungen), wird der Wert von $\alpha_i(k)$ wahrscheinlich für $k = k_i$ maximal werden.

Beispiel 59. *Der folgende Kryptotext y*

HUDS KUAE ZGXR AVTF PGWS WGWS ZHTP PBIL LRTZ PZHW LOIJ VFIC
 VBTH LUGI LGPR KHWM YHTI UAXR BHTW UCGX OSPW AOCH IMCS YHWQ
 HWCY YOCG OGTZ LBIL SWBF LOHX ZWSI ZVDS ATGS THWI SSUX LMST
 MHWI KSPX OGWI HRPF LSAM USUV VAIL LHGI LHWV VIVL AVTW OCIJ
 PTIC MSTX VII

der Länge 203 wurde von einer Vigenère-Chiffre mit Schlüssellänge $d = 4$ aus englischem Klartext erzeugt. Schreiben wir den Kryptotext in vier Spalten y_1, \dots, y_4 der Länge $|y_1| = |y_2| = |y_3| = 51$ und $|y_4| = 50$, so ergeben sich folgende Werte für $\alpha_i(k)$ (in Promille):

k	0	1	2	3	4	5	6	7	8	9	10	11	12
$\alpha_1(k)$	36	31	31	45	38	26	42	73	44	26	36	47	30
$\alpha_2(k)$	44	41	40	51	41	31	37	43	34	28	36	26	28
$\alpha_3(k)$	47	41	48	37	49	40	35	30	48	32	25	42	31
$\alpha_4(k)$	38	40	27	41	65	47	28	34	39	33	35	36	30

k	13	14	15	16	17	18	19	20	21	22	23	24	25
$\alpha_1(k)$	32	36	29	28	39	48	42	42	39	42	42	35	31
$\alpha_2(k)$	43	68	45	35	27	42	43	40	35	30	24	31	45
$\alpha_3(k)$	26	43	76	37	31	39	45	35	34	37	26	30	25
$\alpha_4(k)$	30	48	44	35	42	47	38	39	34	27	38	36	37

Da $\alpha_1(k)$ für $k = 7 = H$, $\alpha_2(k)$ für $k = 14 = O$, $\alpha_3(k)$ für $k = 15 = P$ und $\alpha_4(k)$ für $k = 4 = E$ einen Maximalwert annimmt, lautet das Schlüsselwort **HOPE**. Damit ergibt sich folgender Klartext (aus der Erzählung „Der Goldkäfer“ von Edgar Allan Poe).

A GOOD GLASS IN THE BISHOPS HOSTEL IN THE DEVILS SEAT FORTYONE
 DEGREES AND THIRTEEN MINUTES NORTH EAST AND BY NORTH MAIN
 BRANCH SEVENTH LIMB EAST SIDE SHOOT FROM THE LEFT EYE OF THE
 DEATHS HEAD A BEE LINE FROM THE TREE THROUGH THE SHOT FIFTY
 FEET OUT

◁

Zur Bestimmung des Schlüsselwortes kann man auch die Methode des *gegenseitigen Koinzidenzindex* verwenden. Dabei ist die verwendete Klartextsprache (und somit deren Häufigkeitsverteilung) irrelevant, da die Spalten – wie der Name schon sagt – gegenseitig in Relation gesetzt werden. Aber zuerst die Definition.

Definition 60 (Gegenseitiger Koinzidenzindex). *Der gegenseitige Koinzidenzindex von zwei Texten y und y' mit den Längen n und n' über dem Alphabet \mathcal{B} ist definiert als*

$$IC(y, y') = \frac{1}{n \cdot n'} \cdot \sum_{a \in \mathcal{B}} H(a) \cdot H'(a).$$

Hierbei ist $H(a)$ bzw. $H'(a)$ die absolute Häufigkeit des Buchstabens a im Text y bzw. y' .

$IC(y, y')$ ist also die Wahrscheinlichkeit, dass bei zufälliger Wahl einer Position in y und einer Position in y' der gleiche Buchstabe vorgefunden wird. $IC(y, y')$ ist umso größer, je besser die Häufigkeitsverteilung von y und y' (d. h. H und H') übereinstimmen.

Ist nun y ein Kryptotext, der mit einem Schlüsselwort bekannter Länge d erzeugt wurde, und sind $y_i, i = 1, \dots, d$ die zugehörigen Spalten, so gibt der gegenseitige Koinzidenzindex der Spalten y_i und $y_j + \delta$ (für $1 \leq i < j \leq d$) die Wahrscheinlichkeit an, dass man bei zufälliger Wahl einer Position in y_i und in $y_j + \delta$ denselben Buchstaben vorfindet, wobei δ eine Verschiebung von Spalte y_j relativ zur Spalte y_i ist (mit $0 \leq \delta \leq 25$). Mit großer Wahrscheinlichkeit nimmt also $IC(y_i + \delta, y_j)$ für $\delta = \delta_{ij} = k_j - k_i$ einen relativ großen Wert an, während für $\delta \neq \delta_{ij}$ mit kleinen Werten zu rechnen ist.

Beispiel 61. *Betrachten wir den Kryptotext aus vorigem Beispiel, so ergeben sich für $IC(y_i, y_j + \delta)$ die folgenden Werte (in Promille):*

δ	0	1	2	3	4	5	6	7	8	9	10	11	12
$IC(y_1 + \delta, y_2)$	40	31	25	38	25	21	46	74	50	33	31	44	43
$IC(y_1 + \delta, y_3)$	26	47	25	21	47	32	18	49	91	42	27	51	45
$IC(y_1 + \delta, y_4)$	38	40	29	31	35	24	32	58	42	32	44	50	43
$IC(y_2 + \delta, y_3)$	50	85	49	21	28	35	24	34	46	25	24	27	59
$IC(y_2 + \delta, y_4)$	46	53	40	37	51	42	29	23	24	32	40	55	38
$IC(y_3 + \delta, y_4)$	49	36	38	60	36	25	34	19	29	42	41	33	54

δ	13	14	15	16	17	18	19	20	21	22	23	24	25
$IC(y_1 + \delta, y_2)$	34	31	28	24	31	44	45	37	48	64	44	25	31
$IC(y_1 + \delta, y_3)$	31	29	32	23	29	27	39	45	46	39	58	44	24
$IC(y_1 + \delta, y_4)$	39	31	20	34	36	30	40	45	24	42	78	47	22
$IC(y_2 + \delta, y_3)$	50	50	53	51	24	22	26	43	36	35	32	24	34
$IC(y_2 + \delta, y_4)$	31	32	45	67	49	25	27	29	29	34	37	38	35
$IC(y_3 + \delta, y_4)$	27	36	78	47	25	29	33	27	28	47	32	27	54

Also ist (mit großer Wahrscheinlichkeit)

$$\delta_{12} = 7, \delta_{13} = 8, \delta_{14} = 23, \delta_{23} = 1, \delta_{24} = 16, \delta_{34} = 15.$$

Wir können nun alle Spalten relativ zur ersten Spalte so verschieben, dass der ganze Text eine einheitliche Verschiebung δ hat, also die zweite Spalte um -7 , die dritte um -8 und die vierte um -23 . Für die Bestimmung von δ , muss man nur den häufigsten Buchstaben in dem auf diese Weise erzeugten Text bestimmen (oder eine vollständige Suche durchführen). Dieser ist **L** (16, 3%). Also ist $\delta = L - E = H = 7$ und das Schlüsselwort lautet **HOPE** ($H + 7 = O, H + 8 = P, H + 23 = E$). ◁

Analyse der Lauftextverschlüsselung

Zum Brechen einer Stromchiffre mit Klartextschlüsselstrom kann man so vorgehen: Man geht zunächst davon aus, dass jeder Kryptotextbuchstabe durch Summation eines Klartext- und Schlüsselstrombuchstabens mit jeweils mittlerer bis hoher Wahrscheinlichkeit entstanden ist. Dies sind beispielsweise im Englischen die Buchstaben **E, T, A, O, I,**

N, S, R, H. Zu einem Teilwort w des Kryptotextes bestimmt man dann alle Paare von Wörtern (w_1, w_2) mit $w_1 + w_2 = w$ und $w_1, w_2 \in \{E, T, A, O, I, N, S, R, H\}$. In der Regel ergeben sich nur sehr wenige sinnvolle Paare, aus denen durch Kontextbetrachtungen und Erweitern von w nach links und rechts der Kryptotext entschlüsselt werden kann. Wird die Analyse durch ein Computerprogramm durchgeführt, kann an die Stelle der Kontextbetrachtungen auch die Häufigkeitsverteilung von n -Grammen der Sprache treten. Das Programm wählt dann solche Wortpaare (w_1, w_2) , die eine hohe Wahrscheinlichkeit haben.

Beispiel 62. Gegeben ist der Kryptotext *MOQKTHCBLMWXF...* Wir beginnen die Untersuchung mit einer Wortlänge von vier Buchstaben, also $w = \text{MOQK}$. Der erste Buchstabe M kann nur auf eine der folgenden Arten zustande gekommen sein:

$$\begin{array}{r} ABCDE\dots I\dots T\dots Z \quad (\text{Klartextzeichen}) \\ + \quad MLKJI\dots E\dots T\dots N \quad (\text{Schlüsselzeichen}) \\ \hline = \quad MMMM\dots M\dots M\dots M \quad (\text{Kryptotextzeichen}) \end{array}$$

Es ergeben sich folgende wahrscheinliche Paare für die Einzelbuchstaben von w :

$$\begin{array}{lll} M: & (E,I) & O: & (A,O) & Q: & (I,I) & K: & (R,T) \\ & (I,E) & & (H,H) & & & & (S,S) \\ & (T,T) & & (O,A) & & & & (T,R) \end{array}$$

Diese führen auf folgende $3 \cdot 3 \cdot 1 \cdot 3 = 27$ Wortpaare (w_1, w_2) :

w_1	EAIR	EAIS	EAIT	EHIR	...	THIS	...	TOIT
w_2	IOIT	IOIS	IOIR	IHIT	...	THIS	...	TAIR

Als sinnvoll stellt sich aber nur die Wahl $w_1 = w_2 = \text{THIS}$ heraus. ◀

Autokey Chiffren

Kryptotextschlüsselstrom. Diese Systeme bieten eigentlich keinen großen kryptografischen Schutz, da sie ohne Kenntnis des Schlüsselwortes sehr leicht entschlüsselt werden können (falls die Länge des Schlüsselwortes im Verhältnis zur Länge des Kryptotextes relativ kurz ist). Man subtrahiert dazu den Kryptotext y für $\delta = 1, 2, \dots$ von dem um δ Positionen verschobenen Kryptotext – also $y_{0+\delta} y_{1+\delta} y_{2+\delta} y_{3+\delta} \dots$ minus $y_0 y_1 y_2 y_3 \dots$ –, bis sinnvoller (Klar-) Text erscheint:

$$\begin{array}{r} DUMSQMOZKFN\dots \quad (\text{Kryptotext } y) \\ - \quad \quad \quad DUMSQMO\dots \quad (\text{„Kryptotextschlüsselstrom“}) \\ \hline = \quad \quad \quad \dots\text{NSCHUTZ}\dots \quad (\text{Klartext } x) \end{array}$$

Klartextschlüsselstrom. Neben der oben beschriebenen Analyse der Lauftextverschlüsselung kann das Brechen der Autokey-Systeme mit Klartextschlüsselstrom auch analog zur Kasiski-Methode erfolgen: Sei d die Länge des Schlüsselwortes $k_0 \dots k_{d-1}$. Falls im Klartext die gleiche Buchstabenfolge $x_i \dots x_{i+l-1}$ im Abstand $2d$ auftritt (beispielsweise $d = 3$ und $l = 2$),

3 Sicherheit von Kryptosystemen

3.1 Informationstheoretische Sicherheit

Claude E. Shannon untersuchte die Sicherheit kryptografischer Systeme auf informationstheoretischer Basis (1945, freigegeben 1949). Seinen Untersuchungen liegt das Modell einer Nachrichtenquelle zugrunde, die einzelne Nachrichten unter einer bestimmten Wahrscheinlichkeitsverteilung aussendet.

Bei der Betrachtung der informationstheoretischen Eigenschaften von Kryptosystemen gehen wir von einer Wahrscheinlichkeitsverteilung auf den Paaren $(k, x) \in K \times M$ aus, d. h. $p(k, x)$ gibt die Wahrscheinlichkeit an, dass der Klartext x mit dem Schlüssel k verschlüsselt wird. Dabei setzen wir voraus, dass nach jeder Verschlüsselung einer Nachricht x ein neuer Schlüssel gewählt wird. Dies bedeutet, dass beispielsweise bei der additiven Chiffre für $M = A^n$ zu setzen ist (und nicht $M = A$ wie in der formalen Definition eines Kryptosystems), falls mit dem selben Schlüssel eine Folge von n Buchstaben chiffriert wird, bevor er gewechselt wird.

Weiterhin nehmen wir an, dass der Schlüssel unabhängig vom Klartext gewählt wird. D.h. es ist $p(k, x) = p(k)p(x)$, wobei

$$p(k) = \sum_{x \in M} p(k, x)$$

die Wahrscheinlichkeit für den Schlüssel k und

$$p(x) = \sum_{k \in K} p(k, x)$$

die Wahrscheinlichkeit für den Klartext x ist. Für einen Kryptotext y berechnet sich die Wahrscheinlichkeit zu

$$p(y) = \sum_{k, x: E(k, x) = y} p(k, x)$$

und für einen fest vorgegebenen Kryptotext y mit $p(y) > 0$ ist

$$p(x|y) = \frac{p(x, y)}{p(y)} = \sum_{k: E(k, x) = y} \frac{p(k, x)}{p(y)}$$

die (bedingte) Wahrscheinlichkeit dafür, dass y aus dem Klartext x erzeugt wurde.

Definition 63 (informationstheoretisch sicher). Ein Kryptosystem heißt unter einer Schlüsselverteilung $p(k)$ **absolut sicher (informationstheoretisch sicher)**, falls für jede Klartextverteilung $p(x)$ gilt:

$$p(x) = p(x|y) \text{ für alle } x \in M \text{ und alle } y \in C \text{ mit } p(y) > 0.$$

Bei einem absolut sicheren Kryptosystem ist demnach die *a posteriori* Wahrscheinlichkeit $p(x|y)$ einer Klartextnachricht x gleich der *a priori* Wahrscheinlichkeit $p(x)$, d.h. die

Wahrscheinlichkeit von x ist unabhängig davon, ob der Kryptotext y bekannt ist oder nicht. Die Kenntnis von y erlaubt somit keinerlei Rückschlüsse auf die gesendete Nachricht x . Dies bedeutet, dass es dem Gegner nicht möglich ist – auch nicht mit unbegrenzten Rechenressourcen – das System zu brechen. Wie wir sehen werden, lässt sich dieses Maß an Sicherheit nur mit einem extrem hohen Aufwand realisieren.

Ist $p(x, y) > 0$, so folgt wegen $p(x|y)p(y) = p(x, y) = p(y|x)p(x)$

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

(Satz von Bayes) und daher ist die Bedingung $p(x) = p(x|y)$ gleichbedeutend mit $p(y) = p(y|x)$.

Beispiel 64. Sei (M, C, E, D, K) ein Kryptosystem mit $M = \{x_1, \dots, x_4\}$, $K = \{k_1, \dots, k_4\}$, $C = \{y_1, \dots, y_4\}$ und

E	x_1	x_2	x_3	x_4
k_1	y_1	y_4	y_3	y_2
k_2	y_2	y_1	y_4	y_3
k_3	y_3	y_2	y_1	y_4
k_4	y_4	y_3	y_2	y_1

Weiter sei $p(k_1) = 1/2$, $p(k_2) = 1/4$ und $p(k_3) = p(k_4) = 1/8$. Unter der Klartextverteilung $p(x_1) = 1/2$, $p(x_2) = p(x_3) = p(x_4) = 1/6$ ergibt sich dann folgende Verteilung der Kryptotexte:

$$\begin{aligned} p(y_1) &= 1/2 \cdot 1/2 + (1/4 + 1/8 + 1/8) \cdot 1/6 = 1/3 \\ p(y_2) &= 1/4 \cdot 1/2 + (1/8 + 1/8 + 1/2) \cdot 1/6 = 1/4 \\ p(y_3) &= 1/8 \cdot 1/2 + (1/8 + 1/2 + 1/4) \cdot 1/6 = 5/24 \\ p(y_4) &= 1/8 \cdot 1/2 + (1/2 + 1/4 + 1/8) \cdot 1/6 = 5/24 \end{aligned}$$

Die bedingten Wahrscheinlichkeiten $p(x|y_1)$ berechnen sich wie folgt:

$$\begin{aligned} p(x_1|y_1) &= p(k_1, x_1)/p(y_1) = (1/2)(1/2)/(1/3) = 3/4 \\ p(x_2|y_1) &= p(k_2, x_2)/p(y_1) = (1/4)(1/6)/(1/3) = 1/8 \\ p(x_3|y_1) &= p(k_3, x_3)/p(y_1) = (1/8)(1/6)/(1/3) = 1/16 \\ p(x_4|y_1) &= p(k_4, x_4)/p(y_1) = (1/8)(1/6)/(1/3) = 1/16 \end{aligned}$$

Wegen $p(x_1) = 1/2 \neq 3/4 = p(x_1|y_1)$ ist das Kryptosystem nicht absolut sicher, zumindest nicht unter der gegebenen Schlüsselverteilung.

Die Bedingung $p(x) = p(x|y)$ ist nach dem Satz von Bayes genau dann erfüllt, wenn $p(y) = p(y|x)$ ist. Da jedoch für jedes Paar (x, y) genau ein Schlüssel $k = k_{x,y} \in K$ mit $E(k, x) = y$ existiert, also $p(y|x) = p(k_{x,y})$ ist, ist dies äquivalent zu $p(y) = p(k_{x,y})$. Für $y = y_1$ bedeutet dies, dass alle Schlüssel $k_i = k_{x_i, y_1}$ die gleiche Wahrscheinlichkeit $p(k_i) = 1/4$ haben müssen. Eine leichte Rechnung zeigt, dass dann auch $p(y_i) = 1/4$ für $i = 1, \dots, 4$ ist. Somit ist das betrachtete Kryptosystem genau dann absolut sicher, wenn der Schlüssel unter Gleichverteilung gewählt wird. \triangleleft

Wie in diesem Beispiel lässt sich allgemein folgende hinreichende Bedingung für die absolute Sicherheit von Kryptosystemen zeigen.

Satz 65. Ein Kryptosystem mit $\|M\| = \|C\| = \|K\|$, in dem es für jeden Klartext x und jeden Kryptotext y genau einen Schlüssel k mit $E(k, x) = y$ gibt, ist absolut sicher, wenn die Schlüssel unter Gleichverteilung gewählt werden.

Beweis. Bezeichne $k_{x,y}$ den eindeutig bestimmten Schlüssel, der den Klartext x auf den Kryptotext y abbildet. Wegen $p(k_{x,y}) = \|K\|^{-1}$ für alle x, y folgt zunächst

$$p(y) = \sum_{k,x:E(k,x)=y} p(k, x) = \sum_x p(x) \cdot p(k_{x,y}) = \|K\|^{-1} \sum_x p(x) = \|K\|^{-1}$$

und damit

$$p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(x) \cdot p(y|x)}{p(y)} = \frac{p(x) \cdot p(k_{x,y})}{p(y)} = \frac{p(x) \cdot \|K\|^{-1}}{\|K\|^{-1}} = p(x).$$

□

In den Übungen wird gezeigt, dass auch die Umkehrung dieses Satzes gilt.

Verwendet man beim One-time-pad nur Klartexte einer festen Länge n , d. h. $M \subseteq A^n$, so ist dieser nach obigem Satz absolut sicher (vorausgesetzt, der Schlüssel wird rein zufällig, also unter Gleichverteilung gewählt). Variiert die Klartextlänge, so kann ein Gegner aus y nur die Länge des zugehörigen Klartextes x ableiten. Wird jedoch derselbe Schlüssel k zweimal verwendet, so kann aus den Kryptotexten die Differenz der zugehörigen Klartexte ermittelt werden:

$$\left. \begin{array}{l} y_1 = E(x_1, k) = x_1 + k \\ y_2 = E(x_2, k) = x_2 + k \end{array} \right\} \rightsquigarrow y_1 - y_2 = x_1 - x_2$$

Sind die Klartexte natürlichsprachig, so können aus $y_1 - y_2$ die beiden Nachrichten x_1 und x_2 ähnlich wie bei der Analyse einer Lauftextverschlüsselung (siehe Abschnitt 2.5) rekonstruiert werden.

Da in einem absolut sicheren Kryptosystem der Schlüsselraum K mindestens die Größe des Klartextraumes M haben muss (siehe Übungen), ist der Aufwand extrem hoch. Vor der Kommunikation muss ein Schlüssel, dessen Länge der des zu übertragenden Klartextes entspricht, zufällig generiert und zwischen den Partnern auf einem sicheren Kanal ausgetauscht werden. Wird hingegen keine absolute Sicherheit angestrebt, so kann der Schlüsselstrom auch von einem Pseudo-Zufallsgenerator erzeugt werden. Dieser erhält als Eingabe eine Zufallsfolge s_0 (den sogenannten *Keim*) und erzeugt daraus eine lange Folge $v_0 v_1 \dots$ von Pseudo-Zufallszahlen. Als Schlüssel muss jetzt nur noch das Wort s_0 ausgetauscht werden.

In der Informationstheorie wird die Unsicherheit, mit der eine durch X beschriebene Quelle ihre Nachrichten aussendet, nach ihrer Entropie bemessen. Das heißt, die Unsicherheit über X entspricht genau dem Informationsgewinn, der sich aus der Beobachtung der Quelle X ziehen lässt. Dabei wird die in einer einzelnen Nachricht x steckende Information um so höher bemessen, je seltener x auftritt. Tritt eine Nachricht x mit einer positiven Wahrscheinlichkeit $p(x) = \Pr[X = x] > 0$ auf, dann ist

$$Inf_X(x) = \log_2(1/p(x))$$

der **Informationsgehalt** von x . Ist dagegen $p(x) = 0$, so sei $Inf_X(x) = 0$. Dieser Wert des Informationsgehalts ergibt sich zwangsläufig aus den beiden folgenden Forderungen:

- Der gemeinsame Informationsgehalt $Inf_{X,Y}(x, y)$ von zwei Nachrichten x und y , die aus stochastisch unabhängigen Quellen X und Y stammen, sollte gleich $Inf_X(x) + Inf_Y(y)$ sein;
- der Informationsgehalt einer Nachricht, die mit Wahrscheinlichkeit $1/2$ auftritt, soll genau 1 (bit) betragen.

Die Einheit, in der der Informationsgehalt gemessen wird, ist bit (basic indissoluble information unit). Die Entropie von X ist nun der erwartete Informationsgehalt einer von X stammenden Nachricht.

Definition 66 (Entropie). Sei X eine Zufallsvariable mit Wertebereich $W(X) = \{x_1, \dots, x_n\}$ und sei $p_i = \Pr[X = x_i]$. Dann ist die **Entropie** von X definiert als

$$\mathcal{H}(X) = \sum_{i=1}^n p_i Inf_X(x_i) = \sum_{i=1}^n p_i \log_2(1/p_i).$$

Beispiel 67. Sei X eine Zufallsvariable mit der Verteilung

x_i	sonnig	leicht bewölkt	bewölkt	stark bewölkt	Regen	Schnee	Nebel
p_i	1/4	1/4	1/8	1/8	1/8	1/16	1/16

Dann ergibt sich die Entropie von X zu

$$\mathcal{H}(X) = 1/4 \cdot (2 + 2) + 1/8 \cdot (3 + 3 + 3) + 1/16 \cdot (4 + 4) = 2.625.$$

◁

Die Entropie nimmt im Fall $p_1 = \dots = p_n = 1/n$ den Wert $\log_2(n)$ an. Für jede andere Verteilung p_1, \dots, p_n gilt dagegen $\mathcal{H}(X) < \log_2(n)$ (Beweis unten). Generell ist die Unsicherheit über X um so kleiner, je ungleichmäßiger X verteilt ist. Bringt X nur einen einzigen Wert mit positiver Wahrscheinlichkeit hervor, dann (und nur dann) nimmt $\mathcal{H}(X)$ den Wert 0 an. Für den Nachweis von oberen Schranken für die Entropie benutzen wir folgende Hilfsmittel aus der Analysis.

Definition 68 (konkav). Eine reellwertige Funktion f ist **konkav** auf einem Intervall I , falls für alle $x \neq y \in I$ gilt:

$$f\left(\frac{x+y}{2}\right) \geq \frac{f(x) + f(y)}{2}$$

Gilt sogar „>“ anstelle von „≥“, so heißt f **streng konkav** auf I .

Beispiel 69. Die Funktion $f(x) = \log_2(x)$ ist streng konkav auf $(0, \infty)$.

◁

Für den Beweis des nächsten Satzes benötigen wir die Jensensche Ungleichung, die wir ohne Beweis angeben.

Satz 70 (Jensensche Ungleichung). Sei f eine streng konkave Funktion auf I und seien $0 < a_1, \dots, a_n < 1$ reelle Zahlen mit $\sum_{i=1}^n a_i = 1$. Dann gilt für alle $x_1, \dots, x_n \in I$,

$$f\left(\sum_{i=1}^n a_i x_i\right) \geq \sum_{i=1}^n a_i f(x_i).$$

Hierbei tritt Gleichheit genau dann ein, wenn alle x_i den gleichen Wert haben.

Satz 71. Sei X eine Zufallsvariable mit endlichem Wertebereich $W(X) = \{x_1, \dots, x_n\}$ und Verteilung $\Pr[X = x_i] = p_i$, $i = 1, \dots, n$. Dann ist $H(X) \leq \log_2(n)$, wobei Gleichheit genau im Fall $p_i = 1/n$ für $i = 1, \dots, n$ eintritt.

Beweis. Es gilt

$$\begin{aligned} \mathcal{H}(X) &= \sum_{i=1}^n p_i \log_2(1/p_i) \\ &\leq \log_2 \sum_{i=1}^n p_i/p_i \\ &= \log_2 n. \end{aligned}$$

Nach obigem Satz tritt Gleichheit genau im Fall $1/p_1 = \dots = 1/p_n$ ein, was mit $p_i = 1/n$ für $i = 1, \dots, n$ gleichbedeutend ist. \square

Eine wichtige Eigenschaft der Entropie ist, dass sie eine untere Schranke für die mittlere Codewortlänge von Binärcodes bildet. Ein **Binärcode** für X ist eine (geordnete) Menge $C = \{y_1, \dots, y_n\}$ von binären Codewörtern y_i für die Nachrichten x_i mit der Eigenschaft, dass die Abbildung $c : M^* \rightarrow \{0, 1\}^*$ mit $c(x_{i_1} \dots x_{i_k}) = y_{i_1} \dots y_{i_k}$ injektiv ist. Die Injektivität von c stellt sicher, dass jede Folge $y_{i_1} \dots y_{i_k}$ von Codewörtern eindeutig decodierbar ist.

Die **mittlere Codewortlänge** von C unter X ist

$$L(C) = \sum_{i=1}^n p_i \cdot |y_i|.$$

C heißt **optimal**, wenn kein anderer Binärcode für X eine kürzere mittlere Codewortlänge besitzt. Für einen optimalen Binärcode C für X gilt (ohne Beweis)

$$\mathcal{H}(X) \leq L(C) < \mathcal{H}(X) + 1.$$

Beispiel 72. Sei X die Zufallsvariable aus dem letzten Beispiel. Betrachten wir die beiden Codes $C_1 = \{001, 010, 011, 100, 101, 110, 111\}$ und $C_2 = \{00, 01, 100, 101, 110, 1110, 1111\}$, so erhalten wir für die mittlere Codewortlänge von C_1 den Wert $L(C_1) = 3$, während C_2 wegen $|y_i| = \log_2(1/p_i)$ den Wert $L(C_2) = \mathcal{H}(X)$ erreicht und somit optimal ist. \triangleleft

Die Redundanz eines Codes für eine Zufallsvariable X ist um so höher, je größer seine mittlere Codewortlänge im Vergleich zur Entropie von X ist. Um auch Codes über unterschiedlichen Alphabeten miteinander vergleichen zu können, ist es notwendig, die Codewortlänge in einer festen Einheit anzugeben. Hierzu berechnet man die **Bitlänge** eines Wortes x über einem Alphabet A mit $m > 2$ Buchstaben zu $|x|_2 = |x| \log_2(m)$. Beispielsweise ist die Bitlänge von **GOLD** (über dem lateinischen Alphabet) $|\mathbf{GOLD}|_2 = 4 \log_2(26) = 18,8$. Entsprechend berechnet sich für einen Code $C = \{y_1, \dots, y_n\}$ unter einer Verteilung p_1, \dots, p_n die mittlere Codewortlänge (in bit) zu

$$L_2(C) = \sum_{i=1}^n p_i \cdot |y_i|_2.$$

Damit können wir die Redundanz eines Codes als den mittleren Anteil der Codewortbuchstaben definieren, die keine Information tragen.

Definition 73 (Redundanz). Die (*relative*) **Redundanz** eines Codes C für X ist definiert als

$$\mathcal{R}(C) = \frac{L_2(C) - \mathcal{H}(X)}{L_2(C)}.$$

Beispiel 74. Während eine von X generierte Nachricht im Durchschnitt $\mathcal{H}(X) = 2.625$ bit an Information enthält, haben die Codewörter von C_1 eine Bitlänge von 3. Der Anteil an „überflüssigen“ Zeichen pro Codewort beträgt also

$$\mathcal{R}(C_1) = \frac{3 - 2.625}{3} = 12,5\%,$$

wogegen C_2 keine Redundanz besitzt. ◁

Auch Schriftsprachen wie Deutsch oder Englisch und Programmiersprachen wie C oder PASCAL können als eine Art Code aufgefasst werden. Um die statistischen Eigenschaften einer solchen Sprache L zu erforschen, erweist es sich als zweckmäßig, die Textstücke der Länge n (n -Gramme) von L für unterschiedliche n getrennt voneinander zu betrachten. Sei also L_n die Zufallsvariable, die die Verteilung aller n -Gramme in L beschreibt. Interpretieren wir diese n -Gramme als Codewörter einer einheitlichen Codewortlänge n , so ist

$$\mathcal{R}(L_n) = \frac{n \log_2 m - \mathcal{H}(L_n)}{n \log_2 m}$$

die Redundanz dieses Codes. Es ist zu erwarten, dass eine Sprache umso mehr Redundanz aufweist, je restriktiver die Gesetzmäßigkeiten sind, unter denen in ihr Worte und Sätze gebildet werden.

Definition 75 (Entropie einer Sprache). Für eine Sprache L über einem Alphabet A mit $\|A\| = m$ und n -Gramm-Verteilung L_n ist $\mathcal{H}(L_n)/n$ die **Entropie von L_n** (pro Buchstabe). Falls dieser Wert für n gegen ∞ gegen einen Grenzwert

$$\mathcal{H}(L) = \lim_{n \rightarrow \infty} \mathcal{H}(L_n)/n$$

konvergiert, so wird dieser Grenzwert als die **Entropie von L** bezeichnet. In diesem Fall konvergiert $\mathcal{R}(L_n)$ gegen den Grenzwert

$$\mathcal{R}(L) = \lim_{n \rightarrow \infty} \mathcal{R}(L_n) = \frac{\log_2 m - \mathcal{H}(L)}{\log_2 m},$$

der als die (*relative*) **Redundanz** von L bezeichnet wird. Der Zähler $\mathcal{R}_{\text{abs}}(L) = \log_2 m - \mathcal{H}(L)$ in diesem Ausdruck wird auch als die **absolute Redundanz** der Klartextsprache (gemessen in bit/Buchstabe) bezeichnet.

Für eine Reihe von natürlichen Sprachen wurden die Redundanzen $\mathcal{R}(L_n)$ der n -Gramme (für nicht allzu große Werte von n) empirisch bestimmt, woraus sich $\mathcal{R}(L)$ näherungsweise bestimmen lässt.

Beispiel 76. Im Deutschen hat die Einzelbuchstabenverteilung L_1 eine Entropie von $\mathcal{H}(L_1) = 4,1$ bit, während eine auf A_{lat} gleichverteilte Zufallsvariable U einen Entropiewert von $\mathcal{H}(U) = \log(26) = 4,76$ hat. Für die Bi- und Trigramme ergeben sich

Entropiewerte von $\mathcal{H}(L_2)/2 = 3,86$ und $\mathcal{H}(L_3)/3 = 3,61$ bit pro Buchstabe. Mit wachsender Länge sinkt die Entropie von deutschsprachigen Texten weiter ab und strebt gegen einen Grenzwert $\mathcal{H}(L)$ von 1,56 bit pro Buchstabe.

n	$\mathcal{H}(L_n)$	$\mathcal{H}(L_n)/n$	$\mathcal{R}_{abs}(L_n)/n$	$\mathcal{R}(L_n)$
1	4,10	4,10	0,66	14%
2	7,72	3,86	0,90	19%
3	10,82	3,61	1,15	24%
\vdots	\vdots	\vdots	\vdots	\vdots
∞	∞	$\mathcal{H}(L) = 1,56$	$\mathcal{R}_{abs}(L) = 3,20$	$\mathcal{R}(L) = 67\%$

Ein durchschnittlicher deutscher Text hinreichender Länge enthält also einen Redundanzanteil von ca. 67%, so dass er sich bei optimaler Kodierung auf circa 1/3 seiner ursprünglichen Länge komprimieren lässt. \triangleleft

Wir betrachten nun den Fall, dass mit einem Kryptosystem Klartexte der Länge n verschlüsselt werden, ohne dass dabei der Schlüssel gewechselt wird. D. h. die Chiffrierfunktion hat die Form

$$E_n : K \times A^n \rightarrow C_n,$$

wobei wir die Klartextlänge n variabel halten und der Einfachheit halber annehmen, dass die Menge C_n der zugehörigen Kryptotexte die gleiche Kardinalität $\|C_n\| = \|A^n\| = m^n$ wie der Klartextrraum hat. Ist y ein abgefangener Kryptotext, so ist

$$K(y) = \{k \in K \mid \exists x \in A^n : E_n(k, x) = y \wedge p(x) > 0\}$$

die Menge aller in Frage kommenden Schlüssel für y . $K(y)$ besteht aus einem „echten“ (d. h. dem zur Generierung von y tatsächlich benutzten) und $\|K(y)\| - 1$ so genannten „unechten“ Schlüsseln. Aus informationstheoretischer Sicht ist das Kryptosystem desto unsicherer, je kleiner die erwartete Anzahl

$$\bar{s}_n = \sum_{y \in C_n} p(y) \cdot (\|K(y)\| - 1) = \sum_{y \in C_n} p(y) \cdot \|K(y)\| - 1$$

der unechten Schlüssel ist. Ist \bar{s}_n gleich 0, so liefert der abgefangene Kryptotext y dem Gegner genügend Information, um den benutzten Schlüssel und somit den zu y gehörigen Klartext eindeutig bestimmen zu können (sofern er über unbegrenzte Ressourcen an Rechenkraft und Zeit verfügt).

Definition 77 (Eindeutigkeitsdistanz). Die **Eindeutigkeitsdistanz** n_0 eines Kryptosystems ist der kleinste Wert von n , für den $\bar{s}_n = 0$ wird.

Als nächstes wollen wir eine untere Schranke für \bar{s}_n (und damit für n_0) herleiten. Hierzu benötigen wir den Begriff der bedingten Entropie $\mathcal{H}(X|Y)$ von X , wenn Y bereits bekannt ist.

Definition 78 (bedingte Entropie). Seien X, Y Zufallsvariablen. Dann ist die **bedingte Entropie** von X unter Y definiert als

$$\mathcal{H}(X|Y) = \sum_{y \in W(Y)} p(y) \cdot \mathcal{H}(X|y),$$

wobei $X|y$ die Zufallsvariable mit der Verteilung $\Pr[X|y = x] = p(x|y) = \Pr[X = x \mid Y = y]$ ist (d. h. $\mathcal{H}(X|y) = \sum_{x \in W(X)} p(x|y) \cdot \log_2(1/p(x|y))$).

Satz 79.

$$(i) \mathcal{H}(X, Y) = \mathcal{H}(Y) + \mathcal{H}(X|Y).$$

(ii) $\mathcal{H}(X, Y) \leq \mathcal{H}(X) + \mathcal{H}(Y)$, wobei Gleichheit genau dann eintritt, wenn X und Y stochastisch unabhängig sind.

Beweis. s. Übungen. □

Korollar 80. $\mathcal{H}(X|Y) \leq \mathcal{H}(X)$, wobei Gleichheit genau dann eintritt, wenn X und Y stochastisch unabhängig sind.

Satz 81. In jedem Kryptosystem gilt für die Klartextentropie $\mathcal{H}(X)$, die Schlüssellentropie $\mathcal{H}(K)$ und die Kryptotextentropie $\mathcal{H}(Y)$

$$\mathcal{H}(K|Y) = \mathcal{H}(K) + \mathcal{H}(X) - \mathcal{H}(Y).$$

Beweis. Zunächst ist $\mathcal{H}(K|Y) = \mathcal{H}(K, Y) - \mathcal{H}(Y)$. Es reicht also zu zeigen, dass

$$\mathcal{H}(K, Y) = \mathcal{H}(K) + \mathcal{H}(X)$$

ist. Da bei Kenntnis des Schlüssels der Wert von X bereits eindeutig durch Y und der Wert von Y eindeutig durch X festgelegt ist, folgt unter Berücksichtigung der gemachten Annahme, dass X und K unabhängig sind,

$$\mathcal{H}(K, Y) = \mathcal{H}(K, X, Y) = \mathcal{H}(K, X) + \underbrace{\mathcal{H}(Y|K, X)}_{=0} = \mathcal{H}(K) + \mathcal{H}(X).$$

□

Jetzt verfügen wir über alle Hilfsmittel, um die erwartete Anzahl

$$\bar{s}_n = \sum_{y \in C_n} p(y) \cdot \|K(y)\| - 1$$

der unechten Schlüssel nach unten abschätzen zu können. Seien X_n und Y_n die Zufallsvariablen, die die Verteilungen der n -Gramme der Klartextsprache und der zugehörigen Kryptotexte beschreiben.

Lemma 82.

$$(i) \mathcal{H}(K|Y_n) \leq \log_2(\bar{s}_n + 1),$$

$$(ii) \mathcal{H}(K|Y_n) \geq \mathcal{H}(K) - n\mathcal{R}(L) \log_2 m.$$

Beweis.

(i) Unter Verwendung der Jensenschen Ungleichung folgt

$$\begin{aligned} \mathcal{H}(K|Y_n) &= \sum_{y \in C_n} p(y) \cdot \mathcal{H}(K|y) \\ &\leq \sum_{y \in C_n} p(y) \cdot \log_2 \|K(y)\| \\ &\leq \log_2 \sum_{y \in C_n} p(y) \cdot \|K(y)\| \\ &= \log_2(\bar{s}_n + 1). \end{aligned}$$

(ii) Mit Satz 81 folgt

$$\mathcal{H}(K|Y_n) = \mathcal{H}(K) + \mathcal{H}(X_n) - \mathcal{H}(Y_n).$$

Die Klartextentropie $\mathcal{H}(X_n)$ lässt sich durch

$$\mathcal{H}(X_n) = \mathcal{H}(L_n) \geq n\mathcal{H}(L) = n(1 - \mathcal{R}(L)) \log_2 m$$

abschätzen, wobei $m = \|A\|$ ist. Zudem lässt sich die Kryptotextentropie $\mathcal{H}(Y_n)$ wegen $W(Y_n) = C_n$ und $\|C_n\| = m^n$ durch

$$\mathcal{H}(Y_n) \leq n \log_2 m$$

abschätzen. Somit ist

$$\mathcal{H}(K|Y_n) = \mathcal{H}(K) + \underbrace{\mathcal{H}(X_n) - \mathcal{H}(Y_n)}_{\geq -n\mathcal{R}(L) \log_2 m}.$$

□

Zusammen ergibt sich also

$$\log_2(\bar{s}_n + 1) \geq \mathcal{H}(K) - n\mathcal{R}(L) \log_2 m.$$

Im Fall, dass der Schlüssel unter Gleichverteilung gezogen wird, erreicht $\mathcal{H}(K)$ den maximalen Wert $\log_2 \|K\|$, was auf die gesuchte Abschätzung für \bar{s}_n führt. Wir fassen zusammen.

Satz 83. *Werden mit einem Kryptosystem Klartexte $x \in A^n$ der Länge n mit einem unter Gleichverteilung gezogenen Schlüssel $k \in K$ verschlüsselt, und ist $\|C_n\| = \|A^n\| = m^n$ für den zugehörigen Kryptotextraum $C_n = \{E(k, x) \mid k \in K, x \in A^n\}$, so gilt für die erwartete Anzahl \bar{s}_n der unechten Schlüssel,*

$$\bar{s}_n \geq \frac{\|K\|}{m^{n\mathcal{R}(L)}} - 1.$$

Setzen wir in obiger Abschätzung $\bar{s}_n = 0$, so erhalten wir folgende untere Schranke für die Eindeutigkeitsdistanz n_0 des Kryptosystems.

Korollar 84. *Unter den Bedingungen des obigen Satzes gilt*

$$n_0 \geq \frac{\log_2 \|K\|}{\mathcal{R}(L) \log_2 m} = \frac{\log_2 \|K\|}{\log_2 m - \mathcal{H}(L)} = \frac{\log_2 \|K\|}{\mathcal{R}_{abs}(L)}.$$

Man beachte, dass wir nur die Mindestmenge an Kryptotext zur eindeutigen Bestimmung des *Schlüssels* abgeschätzt haben. Natürlich erlaubt die eindeutige Bestimmung des Schlüssels auch die eindeutige Bestimmung des Klartexts. Unter Umständen kann jedoch der *Klartext* auch schon bei Kenntnis von wesentlich weniger Kryptotext eindeutig bestimmbar sein.

Beispiel 85. *Für Substitutionen bei deutschsprachigem Klartext ergeben sich folgende Werte $\log_2 \|K\|/\mathcal{R}_{abs}(L)$ als untere Schranke für die Eindeutigkeitsdistanz n_0 (wobei wir von einer absoluten Redundanz von $\mathcal{R}_{abs}(L) = 3.2$ bit/Zeichen ausgehen, was einer relativen Redundanz von $\mathcal{R}(L) = 3,2/4,76 \approx 67\%$ entspricht):*

Kryptosystem	Schlüssellanzahl $\ K\ $	$\log_2 \ K\ $	$\log_2 \ K\ / \mathcal{R}_{abs}(L)$
additive Chiffre	26	4.7	$\frac{4.7}{3.2} \approx 1.5$
affine Chiffre	$12 \cdot 26 = 312$	8.3	2.6
einfache Substitution	$26!$	88.4	27.6
Vigenère-Chiffre	26^d	$4.7 \cdot d$	$1.5 \cdot d$
DES-Algorithmus	2^{56}	56	17.5

Dagegen erhalten wir für Blocktranspositionen folgende unteren Schranken für die Mindestmenge an Kryptotext, die zur eindeutigen Bestimmung des Schlüssels benötigt wird:

Analyse auf der Basis von	$\mathcal{R}_{abs}(L)$	Blocklänge l				
		10	20	50	100	1 000
Einzelzeichen	0,66	59	165	578	1415	22 986
Bigrammen	0,90	40	111	390	954	15 502
Trigrammen	1,15	24	65	226	553	9 473
n -Grammen, $n \rightarrow \infty$	3,20	7	19	67	164	2 665

Auch wenn die unteren Schranken für n_0 bei der Analyse auf der Basis von Einzelzeichen endlich sind, ist in diesem Fall $n_0 = \infty$, da eine solche Analyse nicht zum Ziel führen kann, unabhängig davon, über wie viel Kryptotext der Gegner verfügt. \triangleleft

3.2 Weitere Sicherheitsbegriffe

Wie wir gesehen haben, muss für die Benutzung eines informationstheoretisch sicheren Kryptosystems ein immenser Aufwand betrieben werden. Daher begnügt man sich in der Praxis meist mit schwächeren Sicherheitsanforderungen.

- Ein Kryptosystem gilt als **komplexitätstheoretisch sicher** oder als **berechnungssicher** (**computationally secure**), falls es dem Gegner nicht möglich ist, das System mit einem für ihn lohnenswerten Aufwand zu brechen. Das heißt, der Zeitaufwand und die Kosten für einen erfolgreichen Angriff (sofern er überhaupt möglich ist) übersteigen den potentiellen Nutzen bei weitem.
- Ein Kryptosystem gilt als **nachweisbar sicher** (**provably secure**), wenn seine Sicherheit mit bekannten komplexitätstheoretischen Hypothesen verknüpft werden kann, deren Gültigkeit gemeinhin akzeptiert wird.
- Als **praktisch sicher** (**practically secure**) werden dagegen Kryptosysteme eingestuft, die über mehrere Jahre hinweg jedem Versuch einer erfolgreichen Kryptoanalyse widerstehen konnten, obwohl sie bereits eine weite Verbereitung gefunden haben und allein schon deshalb ein lohnenswertes Ziel für einen Angriff darstellen.

Die komplexitätstheoretische Analyse eines Kryptosystems ist äußerst schwierig. Dies hängt damit zusammen, daß der Aufwand eines erfolgreichen Angriffs unabhängig von der vom Gegner angewandten Strategie abgeschätzt werden muss. Das heißt, es müssen nicht nur alle derzeit bekannten kryptoanalytischen Ansätze, sondern alle *möglichen* in Betracht gezogen werden. Dabei darf sich die Aufwandsanalyse nicht ausschließlich an einer vollständigen Rekonstruktion des Klartextes orientieren, da bereits ein geringfügiger

Unterschied zwischen dem a posteriori und dem a priori Wissen für den Gegner einen Vorteil bedeuten kann.

Aus den genannten Gründen ist es bis heute noch für kein praktikables Kryptosystem gelungen, seine komplexitätstheoretische Sicherheit mathematisch zu beweisen. Damit ist auch nicht so schnell zu rechnen, zumindest nicht solange der Status fundamentaler komplexitätstheoretischer Fragen wie etwa des berühmten $P \stackrel{?}{=} NP$ -Problems offen ist. Dagegen gibt es eine ganze Reihe praktikabler Kryptosysteme, die als nachweisbar sicher oder praktisch sicher gelten.

Wir schließen diesen Abschnitt mit einer Präzisierung des komplexitätstheoretischen Sicherheitsbegriffs. Hierzu ist es erforderlich, die Verletzung der Vertraulichkeit als ein algorithmisches Problem für den Gegner zu formulieren.

Definition 86 (Vorteil eines Gegners). Sei $S = (M, C, E, D, K)$ ein Kryptosystem mit Schlüsselverteilung K . Ein Gegner ist ein Paar (G, V) von probabilistischen Algorithmen, wobei $G = (X_0, X_1)$ zwei Klartexte $x_0 \neq x_1 \in M$ generiert und V bei Eingabe zweier Klartexte $x_0, x_1 \in M$ und eines Kryptotextes $y \in C$ ein Bit ausgibt. Der Vorteil von (G, V) ist

$$\alpha(G, V) = \Pr[V(X_0, X_1, E(K, X_B)) = B] - 1/2,$$

wobei B eine unabhängig von G und V auf $\{0, 1\}$ gleichverteilte Zufallsvariable ist, d.h. $\Pr[B = 0] = \Pr[B = 1] = 1/2$.

Satz 87. Bei einem absolut sicheren Kryptosystem S kann kein Gegner einen Vorteil größer als 0 erzielen.

Beweis. Bei einem absolut sicheren Kryptosystem hängt die Kryptotextverteilung $Y = E(K, X)$ nicht von der zugrunde liegenden Klartextverteilung X ab. Daher sind auch die ZVen $V(X_0, X_1, E(K, X_B))$ und B stochastisch unabhängig und es folgt

$$\begin{aligned} & \Pr[V(X_0, X_1, E(K, X_B)) = B] \\ &= \Pr[V(X_0, X_1, E(K, X_{1-B})) = 0] \cdot \Pr[B = 0 \mid V(X_0, X_1, E(K, X_{1-B})) = 0] + \\ & \quad \Pr[V(X_0, X_1, E(K, X_{1-B})) = 1] \cdot \Pr[B = 1 \mid V(X_0, X_1, E(K, X_{1-B})) = 1] \\ &= 1/2. \end{aligned}$$

□

In den Übungen wird auch die umgekehrte Implikation bewiesen. Ein Kryptosystem ist somit genau dann absolut sicher, wenn kein Gegner einen Vorteil größer 0 erzielt. Für die Präzisierung des komplexitätstheoretischen Sicherheitsbegriffs sind nun die beiden folgenden Fragen von entscheidender Bedeutung:

- Über welche Rechenressourcen verfügt ein Gegner realistischweise?
- Wie groß darf der vom Gegner erzielte Vorteil höchstens sein, damit die Vertraulichkeit der Nachricht noch gewahrt bleibt?

Die Antwort auf diese Fragen liefert Definition 88. Dabei gehen wir davon aus, dass das gewünschte Maß an Sicherheit eines Kryptosystems S durch die Wahl eines Sicherheitsparameters $s \in \mathbb{N}$ regulierbar ist. Natürlich hängt dann auch der Vorteil $\alpha_s(G, V)$ eines Gegners vom Parameterwert s ab. Zudem sollten alle legalen Operationen (wie die Chiffrierung oder die Schlüsselgenerierung) effizient (d.h. in Zeit $s^{O(1)}$) durchführbar sein. Typischerweise werden Kryptosysteme nach der Schlüssellänge $s = |k|$ parameterisiert.

Definition 88 (komplexitätstheoretisch sicher). Sei S ein Kryptosystem mit variablem Sicherheitsparameter s .

- Eine Funktion $\varepsilon : \mathbb{N} \rightarrow \mathbb{R}$ heißt **vernachlässigbar**, wenn für jedes Polynom p eine Zahl n_0 existiert, so dass $\varepsilon(n) < 1/p(n)$ für alle $n \geq n_0$ ist.
- Ein Gegner (G, V) für S heißt **effizient**, wenn sowohl G als auch V durch probabilistische Schaltkreise der Größe $s^{O(1)}$ berechenbar sind.
- S heißt **komplexitätstheoretisch sicher**, wenn jeder effiziente Gegner (G, V) nur einen vernachlässigbaren Vorteil $\alpha_s(G, V)$ (als Funktion von s) erzielen kann.