§ 5. Monoalphabetische Verschlüsselungsverfahren

(5.1) DEF: **a)** Ein **Alphabet** $\mathcal{B}$ ist eine beliebige nichtleere Menge, deren Elemente die **Buchstaben** des Alphabets genannt werden.

b) Ein **Wort über** $\mathcal{B}$ ist eine Aneinanderreihung von endlich vielen Buchstaben aus $\mathcal{B}$.

c) Die Menge aller Wörter über $\mathcal{B}$ wird mit $W(\mathcal{B})$ bezeichnet.

d) Die Anzahl der Buchstaben eines Wortes $w \in W(\mathcal{B})$ heißt die **Länge von** w und wird mit $|w|$ bezeichnet.

e) Zwei Wörter w_1 und w_2 über $W(\mathcal{B})$ heißen **gleich** (in Zeichen: $w_1 = w_2$), wenn sie dieselbe Länge haben und wenn die Buchstaben an den entsprechenden Stellen übereinstimmen.

(5.2) DEF: Ein **monoalphabetisches Verschlüsselungsverfahren** (oder **monoalphabetisches Kryptosystem**) ist ein 5–Tupel $(\mathcal{K}, \mathcal{G}, \mathcal{S}, E_k, D_l)$ bestehend aus

- einem **Klartextalphabet** $\mathcal{K}$
- einem **Geheimtextalphabet** $\mathcal{G}$
- einer **Schlüsselmenge** $\mathcal{S} = \{k, l\}$
- einer **Verschlüsselungsabbildung** $E_k : \mathcal{K} \longrightarrow \mathcal{G}$ (Encryption)
- einer **Entschlüsselungsabbildung** $D_l : \mathcal{G} \longrightarrow \mathcal{K}$ (Decryption)

Dabei muß die folgende **Entschlüsselungsbedingung** gelten:

$$\boxed{D_l \circ E_k = \mathrm{id}_{\mathcal{K}},}$$

d.h. $D_l(E_k(\alpha)) = \alpha$ für alle $\alpha \in \mathcal{K}$

(5.3) BEM: **a)** Ein Wort $w = \alpha_1\alpha_2 \ldots \alpha_n \in W(\mathcal{K})$ wird dann verschlüsselt, indem man die Verschlüsselungsabbildung $E_k : \mathcal{K} \longrightarrow \mathcal{G}$ auf jeden einzelnen Buchstaben von w anwendet, d.h. indem man

$$E_k(\alpha_1)E_k(\alpha_2)\ldots E_k(\alpha_n) \in W(\mathcal{G})$$

bildet. Entsprechend wird die Entschlüsselung eines Wortes vorgenommen.

b) Das Verschlüsselungsverfahren wird **monoalphabetisch** genannt, da ein Buchstabe des Klartextalphabetes immer durch denselben Buchstaben des Geheimtextalphabetes ersetzt wird. Später werden wir im Gegensatz dazu auch noch **polyalphabetische** Verschlüsselungsverfahren kennenlernen.

(5.4) BEISPIELE: **a) Additive Chiffre**:

$(\mathfrak{a}, \mathfrak{A}, \{k,l\}, V_k, V_l)$ mit $k,l \in \mathcal{R}_{26}$, $l = (-k) \bmod 26$

b) Multiplikative Chiffre:

$(\mathfrak{a}, \mathfrak{A}, \{k,l\}, M_k, M_l)$ mit $k,l \in \mathcal{R}_{26}$, $\mathrm{ggT}(k,26) = 1$, $(k \cdot l) \bmod 26 = 1$

c) Affine Chiffre: $(\mathfrak{a}, \mathfrak{A}, \{\{k,l\},\{m,n\}\}, A_{k,l}, A_{n,m})$

mit $k,l \in \mathcal{R}_{26}$, $\mathrm{ggT}(l,26) = 1$, $n = (-k) \bmod 26$, $(m \cdot l) \bmod 26 = 1$

(5.5) BEM: Ist der Schlüssel k bei einer additiven Chiffre bekannt, so läßt sich daraus der Schlüssel $l = (-k) \bmod 26$ zum Entschlüsseln berechnen. Umgekehrt läßt sich aus l auch $k = (-l) \bmod 26$ berechnen, d.h. wer verschlüsseln kann, kann auch entschlüsseln und umgekehrt.

Ein solches Verschlüsselungsverfahren heißt **symmetrisch** im Gegensatz zu den **asymmetrischen** Verfahren, die wir später noch behandeln werden. Auch die multiplikative und die affine Chiffre sind symmetrisch.

(5.6) DEF: Ein Verschlüsselungsverfahren heißt **symmetrisch**, wenn sich aus dem Schlüssel für die Verschlüsselung der Schlüssel für die Entschlüsselung bestimmen läßt und umgekehrt.

(5.7) SATZ: Sei $\mathcal{K}$ ein Klartextalphabet und $\mathcal{G}$ ein Geheimtextalphabet. Dann liefert jede injektive Abbildung $\psi : \mathcal{K} \longrightarrow \mathcal{G}$ ein monoalphabetisches Verschlüsselungsverfahren

$$(\mathcal{K}, \mathcal{G}, \mathcal{S}, E, D)$$

Dabei ist $E = \psi$, und $D = \chi : \mathcal{G} \longrightarrow \mathcal{K}$ ist eine Abbildung mit $\chi \circ \psi = \mathrm{id}_{\mathcal{K}}$. Außerdem ist $\mathcal{S} = \{\psi, \chi\}$.

Bew: Da $\psi : \mathcal{K} \longrightarrow \mathcal{G}$ nach (1.9b) injektiv ist, gibt es nach (1.20) eine Abbildung $\chi : \mathcal{G} \longrightarrow \mathcal{K}$ mit $\chi \circ \psi = \mathrm{id}_{\mathcal{K}}$. Damit gilt die Entschlüsselungsbedingung.
Es sei noch angemerkt, dass es mehrere Entschlüsselungsabbildungen geben kann.

(5.8) BEM: Da jede injektive Abbildung $\mathfrak{a} \longrightarrow \mathfrak{A}$ nach (2.4a) schon bijektiv ist, gibt es also nach (2.6) genau **26!** monoalphabetische Verschlüsselungen von $\mathfrak{a}$ durch $\mathfrak{A}$. Unter ihnen gibt es die **312** affinen Verschlüsselungen, die wir formelmäßig beschreiben konnten.

(5.9) BEISPIEL: Schlüsselwortchiffre

Gegeben seien ein Schlüsselwort $\overline{w}$ aus lauter verschiedenen Buchstaben und ein Buchstabe $\alpha \in \mathfrak{a}$. (Besitzt ein Wort, das man als Schlüsselwort benutzen möchte, Buchstaben mehrfach, so streicht man diese bis jeweils auf den ersten.) Dann schreibt man die unterschiedlichen Buchstaben des Schlüsselwortes unter die des Klartextalphabets, wobei der erste Buchstabe von $\overline{w}$ unter dem angegebenen Buchstaben α steht. Die fehlenden Buchstaben werden in der natürlichen Reihenfolge aufgefüllt. Diese Zuordung liefert eine bijektive Abbildung $\psi : \mathfrak{a} \longrightarrow \mathfrak{A}$, die das monoalphabetische Kryptosystem

$$(\mathfrak{a}, \mathfrak{A}, \mathcal{S}, E, D)$$

liefert mit $E = \psi$, $D = \psi^{-1}$ und $\mathcal{S} = \{\{\alpha, \overline{w}\}, \psi^{-1}\}$.

Kryptoanalyse

Die sog. **Kryptoanalyse** beschäftigt sich mit Verfahren und Methoden, Geheimtexte zu knacken, ohne dass man den Schlüssel zum Entschlüsseln kennt.

Ist ein Geheimtext durch eine additive, multiplikative bzw. affine Chiffre verschlüsselt worden, so kann man die wenigen möglichen Schlüssel ausprobieren und evtl. auch mit Hilfe eines Computers entschlüsseln. Dies ist aber kaum noch möglich bei einer beliebigen monoalphabetischen Chiffre des deutschen Alphabets, von denen es bekanntermaßen 26! Stück gibt (26! ist eine Zahl mit 27 Stellen).

Eine Methode zum Entschlüsseln geht von folgenden Feststellungen aus:

- Bei einem monoalphabetischen Verfahren kommt ein Buchstabe im Geheimtext genauso oft vor wie der durch ihn verschlüsselte Buchstabe im Klartext
- In "normalen" deutschen Texten kommen die Buchstaben mit unterschiedlicher Häufigkeit vor. Jeder Buchstabe hat eine charakteristische Häufigkeit, die man durch Auszählen verschiedener Texte feststellen kann.

(5.10) Häufigkeiten der Buchstaben in deutschen Texten

(aus Beutelspacher: Kryptologie)

Buchstabe	**Häufigkeit**	**Buchstabe**	**Häufigkeit**
a	6,51 %	n	9,78 %
b	1,89 %	o	2,51 %
c	3,06 %	p	0,79 %
d	5,08 %	q	0,02 %
e	17,40 %	r	7,00 %
f	1,66 %	s	7,27 %
g	3,01 %	t	6,15 %
h	4,76 %	u	4,35 %
i	7,55 %	v	0,67 %
j	0,27%	w	1,89 %
k	1,21%	x	0,03 %
l	3,44 %	y	0,04 %
m	2,53%	z	1.13%

In speziellen Texten können natürlich auch andere Häufigkeiten vorkommen, daher fasst man die Buchstaben zu Gruppen zusammen.

(5.11) Buchstabengruppen (aus Beutelspacher: Kryptologie)

Gruppe	**Gesamthäufigkeit dieser Buchstaben**
e,n	27,18%
i,s,r,a,t	34,48%
d,h,u,l,c,g,m,o,b,w,f,k,z	36,52%
p,v,j,y,x,q	1,82%

(5.12) Die häufigsten Bigramme der deutschen Sprache

(aus Beutelspacher: Kryptologie)

Bigramm	**Häufigkeit**	**Bigramm**	**Häufigkeit**
en	3,88 %	nd	1,99 %
er	3,75 %	ei	1,88 %
ch	2,75 %	ie	1,79 %
te	2,26 %	in	1,67 %
de	2,00 %	es	1,52 %

(5.13) BEM: **a)** Für die **Kryptoanalyse** eines monoalphabetisch verschüsselten Textes des deutschen Alphabets nimmt man zunächst eine **Häufigkeitsanalyse** vor, um etwa die Verschlüsselungen von e und n und anderer Buchstaben herauszufinden. e und n kommen normalerweise in einem deutschen Text am häufigsten vor (s. (5.10)) . Damit versucht man, sinnvolle Textteile herauszufinden. Durch Häufigkeitsanalyse von Bigrammen lassen sich u.U. weitere Buchstaben finden. Den Rest muss man irgendwie raten.

b) Eine monoalphabetische Verschlüsselung des deutschen Alphabets ist nicht sehr sicher. Als Alternative kann man polyalphabetische Verschlüsselungen wählen, die eine Kryptoanalyse durch eine Häufigkeitsanalyse sehr erschweren oder sogar unmöglich machen. Damit werden wir uns im folgenden beschäftigen.

c) Es ist natürlich auch noch nicht die Frage geklärt, wie man bei einem vorliegenden Geheimtext überhaupt feststellen kann, ob er durch eine monoalphabetische Verschlüsselung entstanden ist oder nicht.