2.5 Multipartite Substitution

- $\chi : V \longrightarrow W^m, m > 1$
  - $|W|^m \geq |V|$
  - in der Regel $|W| < |V|$
- Beispiel für $W = \{1, 2, 3, 4, 5\}, m = 2$ (V enthält kein `j`):

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

  `nachricht` $\Rightarrow$ 331131322442313244
  - Nutzung als Klopfcode in Haftanstalten
- Einsatz (in Kombination mit einer Transposition) an der deutschen Westfront 1918 mit $W = \{\mathtt{A}, \mathtt{D}, \mathtt{F}, \mathtt{G}, \mathtt{V}, \mathtt{X}\}, m = 2$ und $V = \{\mathtt{a}, \ldots, \mathtt{z}, 0, \ldots, 9\}$
- falls $|W|^m > |V|$ Homophone und/oder Blender möglich

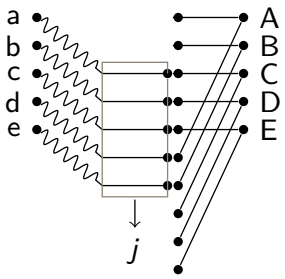
3 Monographische, polyalphabetische Substitution

- monographisch: Verschlüsselung wird auf einzelne Zeichen angewendet
- monoalphabetisch: jedes Zeichen wird gleich verschlüsselt
- polyalphabetisch: Zeichen werden unterschiedlich verschlüsselt
  - $\mathbf{X}(x_1 x_2 x_3 \ldots) = y_1 y_2 y_3 \ldots$ mit $x_i \in V, y_i = \chi_i(x_i) \in W$
- Vorteil: Angriff über Zeichenhäufigkeiten nicht (ohne weiteres) möglich
- praktisches Problem: Bestimmung der χ_i durch Schlüssel
  - sicherste Variante: jedes χ_i ist eine von $N!$ mögichen Substitutionen – aber Schlüssel schwer handhabbar
  - begleitende Alphabete als systematische Ableitung zusätzlicher Chiffrierschritte

3.1 Begleitende Alphabete

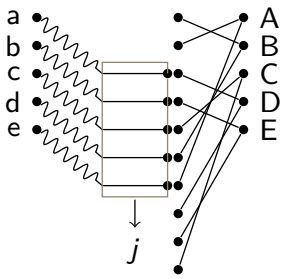
- Sei ρ der Zyklus des Standardalphabets (CAESAR-Schritt mit $t = 1$), so ist $\{\rho^j : j \in \mathbb{N}\}$ die Menge der verschobenen Standardalphabete (CAESAR-Schritte mit $t = j$, VIGENÈRE-Schritte)

j	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
0	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
1	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
2	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
3	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
⋮	⋮ ⋮
24	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
25	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



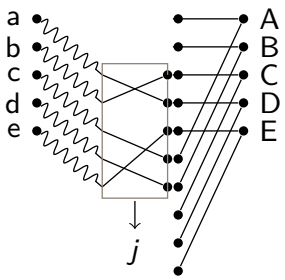
- mit einer Permutation P (P-Alphabet) ergibt sich die Menge der horizontal verschobenen P-Alphabete $\{\rho^j P : j \in \mathbb{N}\}$ (ALBERTI-Schritte)

j	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
0	N E W Y O R K C I T A B D F G H J L M P Q S U V X Z
1	E W Y O R K C I T A B D F G H J L M P Q S U V X Z N
2	W Y O R K C I T A B D F G H J L M P Q S U V X Z N E
3	Y O R K C I T A B D F G H J L M P Q S U V X Z N E W
⋮	⋮ ⋮
24	X Z N E W Y O R K C I T A B D F G H J L M P Q S U V
25	Z N E W Y O R K C I T A B D F G H J L M P Q S U V X



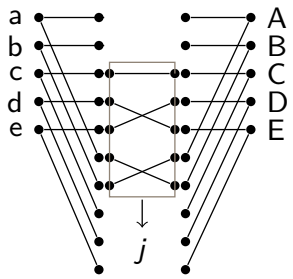
- und der vertikal verschobenen P-Alphabete $\{P\rho^j : j \in \mathbb{N}\}$

j	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
0	N E W Y O R K C I T A B D F G H J L M P Q S U V X Z
1	O F X Z P S L D J U B C E G H I K M N Q R T V W Y A
2	P G Y A Q T M E K V C D F H I J L N O R S U W X Z B
3	Q H Z B R U N F L W D E G I J K M O P S T V X Y A C
⋮	⋮ ⋮
24	L C U W M P I A G R Y Z B D E F H J K N O Q S T V X
25	M D V X N Q J B H S Z A C E F G I K L O P R T U W Y



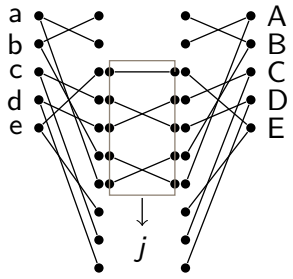
- mit einer Permutation R ergibt sich die Menge der R-rotierten Standardalphabete $\{\rho^{-j} R \rho^j : j \in \mathbb{N}\}$

j	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
0	E S W A R C H O N D U K L B F G I J M P Q T V X Y Z
1	A F T X B S D I P O E V L M C G H J K N Q R U W Y Z
2	A B G U Y C T E J Q P F W M N D H I K L O R S V X Z
3	A B C H V Z D U F K R Q G X N O E I J L M P S T W Y
⋮	⋮ ⋮
24	U Y P A F M L B S I J Z D E G H K N O R T V W X C Q
25	R V Z Q B G N M C T J K A E F H I L O P S U W X Y D



- und der R-rotierten P-Alphabete $\{P\rho^{-j} R \rho^j P^{-1} : j \in \mathbb{N}\}$ (ROTOR-Schritte)

j	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
0	L F X D N Q W C A T B V K H P E M G R O I S U J Y Z
1	S L W D H Q B J T A K N Y V M I E X R O P G U F C Z
2	S D V Y A I T O Q R K L W H J B U N C M P G E F X Z
3	Y X V C A I F H N R K L P Z M W G U O E B Q S T J D
⋮	⋮ ⋮
24	B N C H O A Q T V F W D K S R L I Z M P G E J X Y U
25	B L C D N R Q Z H T F X U O A S J G K P I E V W Y M



3.2 Periodische Schlüssel

- Festlegung, welcher Chiffrierschritt, durch periodisch wiederholten Schlüssel (Kennwort)
- Beispiel mit VIGENÈRE:

```
z u v e r s c h l u e s s e l n d e n a c h r i c h t
P A S S W O R T P A S S W O R T P A S S W O R T P A S
o u n w n g t a a u w k o s c g s e f s y v i b r h l
```

- kein (direkter) Angriff über Buchstabenhäufigkeiten
- kein Angriff über Mustersuche mit wahrscheinlichem Wort
- auch anwendbar mit anderen Chiffrierschrittsystemen
  - "Doppelter Schlüssel" etwa bei ALBERTI-Schritten (Permutation P)
  - Chiffrierschrittsysteme ohne Fixpunkte (z.B. echt involutorische) erlauben negative Mustersuche

3.3 Bestimmung der Periodenlänge nach Kullback

Sei T ein Text der Länge M über dem Alphabet V, $|V| = N$, m_i die (absolute) Häufigkeit des i-ten Zeichens des Alphabets; definiere:

$$\mathrm{Phi}(T) = \frac{1}{M(M-1)} \sum_{i=1}^{N} m_i(m_i - 1)$$

Es gilt

$$\lim_{M\to\infty} \mathrm{Phi}(T) = \sum_{i=1}^{N} p_i^2,$$

wobei p_i die Auftrittswahrscheinlichkeit des i-ten Zeichens angibt. Für eine gleichverteilte Quelle mit $N = 26$ gilt $p_i = \frac{1}{26}$, also

$$\sum_{i=1}^{N} p_i^2 = \sum_{i=1}^{26} \frac{1}{26^2} = \frac{26}{26^2} = \frac{1}{26} = \kappa_R \approx 0{,}038\,46$$

Für eine deutschen bzw. englischen Text gilt $\sum_{i=1}^{N} p_i^2 = \kappa_d \approx 0{,}076\,19$ bzw. $\sum_{i=1}^{N} p_i^2 = \kappa_e \approx 0{,}065\,77$, also $\frac{\kappa_d}{\kappa_R} \approx \frac{\kappa_e}{\kappa_R} \approx 2$.

- monoalphabetische Verschlüsselung: Phi des Geheimtextes nahe $2\kappa_R$
- polyalphabetische Verschlüsselung: Phi des Geheimtextes nahe κ_R
- bei periodischem Schlüssel mit Periodenlänge d: wird nur jeder d-te Buchstabe des Geheimtextes betrachtet, so liegt Phi nahe bei $2\kappa_R$
- Ergebnis deutlich zuverlässiger, wenn die Phi der möglichen Startpunkte (erster, zweiter, ..., d-ter Buchstabe) gemittelt werden
- definiere $T_\rho^{(u)}$ als $t_\rho t_{\rho+u} t_{\rho+2u} \ldots$, die ρ-te Kolonne zur vermuteten Periode u (wobei $1 \le \rho \le u$) und

$$\mathrm{Phi}^{(u)}(T) = \frac{1}{u} \sum_{\rho=1}^{u} \mathrm{Phi}(T_\rho^{(u)})$$

- Berechnung von $\mathrm{Phi}^{(u)}$ für verschiedene $u \Rightarrow$ Maxima bei der Periodenlänge d und ihren Vielfachen

Beispiel:

```
GEIEI ASGDX VZIJQ LMWLA AMXZY ZMLWH FZEKE JLVDX WKWKE TXLBR ATQHL BMXAA
NUBAI VSMUK HSSPW NVLWK AGHGN UMKWD LNRWE QJNXX VVOAE GEUWB ZWMQY MOMLW
XNBXM WALPN FDCFP XHWZK EXHSS FXKIY AHULM KNUMY EXDMW BXZSB CHVWZ XPHWL
GNAMI UK
```

Dieser Geheimtext von G. W. Kulp wurde – allerdings mit Wortzwischenräumen und Interpunktion sowie etlichen Druckfehlern – in einer von Edgar Allan Poe redigierten Kolumne des *Alexander's Weekly Messenger* abgedruckt. Poe hatte monoalphabetisch chiffrierte Texte erbeten.

Mit der Häufigkeit der Buchtaben

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
12	7	2	5	10	4	6	9	6	3	10	12	14	9	2	4	4	2	7	2	7	7	16	15	4	8

ergibt sich

$$\text{Phi} = \frac{1}{187 \cdot 186}\,(12 \cdot 11 + 7 \cdot 6 + \cdots + 8 \cdot 7) \approx 0{,}0456$$

Der Wert 0,0456 von Phi ist aber auffallend gering und in der Tat liegt keine monoalphabetische Chiffre vor.

$u = 11$:

G	E	I	E	I	A	S	G	D	X	V
Z	I	J	Q	L	M	W	L	A	A	M
X	Z	Y	Z	M	L	W	H	F	Z	E
K	E	J	L	V	D	X	W	K	W	K
E	T	X	L	B	R	A	T	Q	H	L
B	M	X	A	A	N	U	B	A	I	V
S	M	U	K	H	S	S	P	W	N	V
L	W	K	A	G	H	G	N	U	M	K
W	D	L	N	R	W	E	Q	J	N	X
X	V	V	O	A	E	G	E	U	W	B
Z	W	M	Q	Y	M	O	M	L	W	X
N	B	X	M	W	A	L	P	N	F	D
C	F	P	X	H	W	Z	K	E	X	H
S	S	F	X	K	I	Y	A	H	U	L
M	K	N	U	M	Y	E	X	D	M	W
B	X	Z	S	B	C	H	V	W	Z	X
P	H	W	L	G	N	A	M	I	U	K
$\frac{8}{272}$	$\frac{6}{272}$	$\frac{8}{272}$	$\frac{12}{272}$	$\frac{10}{272}$	$\frac{8}{272}$	$\frac{10}{272}$	$\frac{4}{272}$	$\frac{8}{272}$	$\frac{16}{272}$	$\frac{20}{272}$

$$\text{Phi}^{(u)} = 0{,}0368$$

$u = 12$:

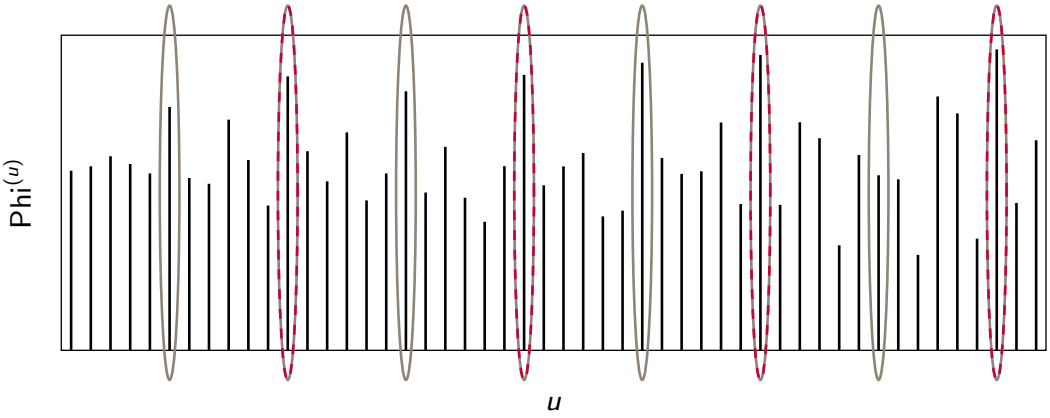
G	E	I	E	I	A	S	G	D	X	V	Z
I	J	Q	L	M	W	L	A	A	M	X	Z
Y	Z	M	L	W	H	F	Z	E	K	E	J
L	V	D	X	W	K	W	K	E	T	X	L
B	R	A	T	Q	H	L	B	M	X	A	A
N	U	B	A	I	V	S	M	U	K	H	S
S	P	W	N	V	L	W	K	A	G	H	G
N	U	M	K	W	D	L	N	R	W	E	Q
J	N	X	X	V	V	O	A	E	G	E	U
W	B	Z	W	M	Q	Y	M	O	M	L	W
X	N	B	X	M	W	A	L	P	N	F	D
C	F	P	X	H	W	Z	K	E	X	H	S
S	F	X	K	I	Y	A	H	U	L	M	K
N	U	M	Y	E	X	D	M	W	B	X	Z
S	B	C	H	V	W	Z	X	P	H	W	L
G	N	A	M	I	U	K					
$\frac{14}{240}$	$\frac{16}{240}$	$\frac{12}{240}$	$\frac{16}{240}$	$\frac{30}{240}$	$\frac{16}{240}$	$\frac{14}{240}$	$\frac{14}{210}$	$\frac{18}{210}$	$\frac{12}{210}$	$\frac{18}{210}$	$\frac{10}{210}$

$$\text{Phi}^{(u)} = 0{,}0695$$

$u = 13$:

G	E	I	E	I	A	S	G	D	X	V	Z	I
J	Q	L	M	W	L	A	A	M	X	Z	Y	Z
M	L	W	H	F	Z	E	K	E	J	L	V	D
X	W	K	W	K	E	T	X	L	B	R	A	T
Q	H	L	B	M	X	A	A	N	U	B	A	I
V	S	M	U	K	H	S	S	P	W	N	V	L
W	K	A	G	H	G	N	U	M	K	W	D	L
N	R	W	E	Q	J	N	X	X	V	V	O	A
E	G	E	U	W	B	Z	W	M	Q	Y	M	O
M	L	W	X	N	B	X	M	W	A	L	P	N
F	D	C	F	P	X	H	W	Z	K	E	X	H
S	S	F	X	K	I	Y	A	H	U	L	M	K
N	U	M	Y	E	X	D	M	W	B	X	Z	S
B	C	H	V	W	Z	X	P	H	W	L	G	N
A	M	I	U	K								
$\frac{4}{210}$	$\frac{4}{210}$	$\frac{12}{210}$	$\frac{10}{210}$	$\frac{18}{210}$	$\frac{10}{182}$	$\frac{8}{182}$	$\frac{12}{182}$	$\frac{10}{182}$	$\frac{10}{182}$	$\frac{14}{182}$	$\frac{8}{182}$	$\frac{6}{182}$

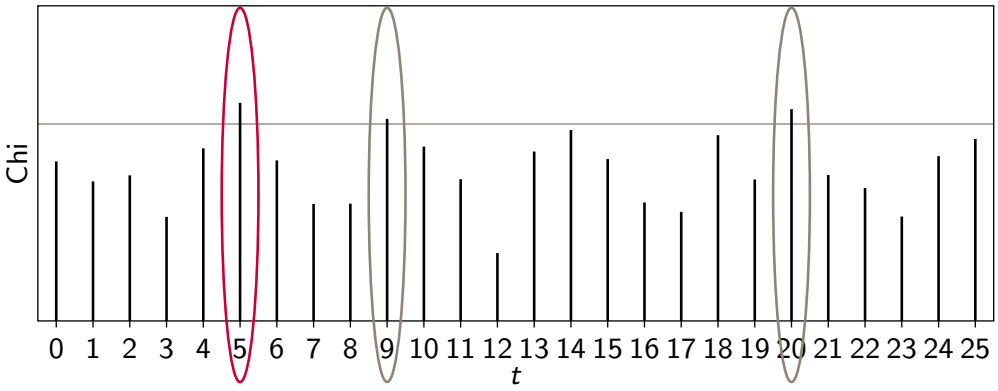
$$\text{Phi}^{(u)} = 0{,}0505$$



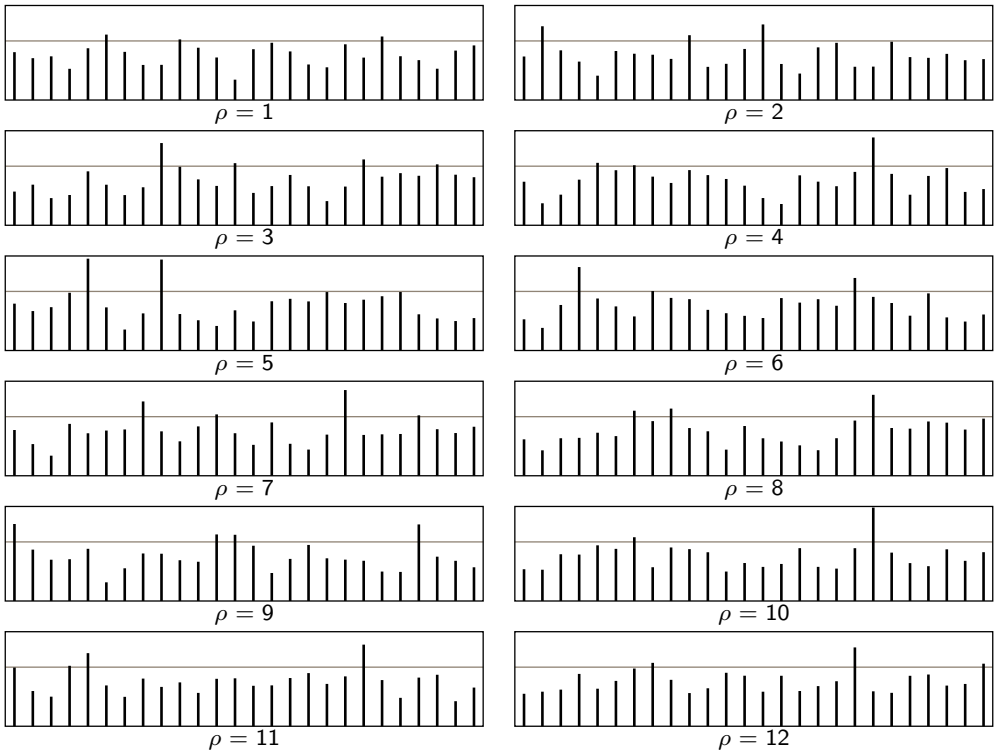
auffällig hohe Werte für Vielfache von 12 und etwas weniger stark ausgeprägt Vielfache von 6

3.4 Zurechtrücken begleitender Alphabete

- jede Kolonne $T_\rho^{(d)}$ ist durch monoalphabetische Substitution entstanden
- falls VIGENÈRE-Verschlüsselung: Bestimmung der Verschiebung mittel Chi
- für $d = 12, \rho = 1$:



Maximum bei $t = 5$ deutlich, aber auch $t = 20$ und $t = 9$ erscheinen gut möglich



- Decodierung mit den jeweiligen Maxima ergibt:
  `brale xande chdwi sitth ateht messe ngerl rgive shere aetwe samet imehi ihthe satur oanco urier anooi hersa turdl yeape rswhe nlcro rding totse satei tispu mlxsh edthr eeoan sprev ioust sihef aultw iehno uorth epodt baste rs`
  - vielversprechend, aber noch nicht ganz richtig
  - näheres Betrachten der problematischen Kolonnen würde schnell zum Ziel führen
- Betrachtung des Schlüssels:
  - Maxima bei 5, 13, 8, 19, 4, 3, 18, 19, 0, 19, 19, 18
  - zweitgrößte Werte bei 20, 1, 19, 4, 8, 18, 7, 8, 22, 6, 4, 7
  - in Buchstaben übersetzt:

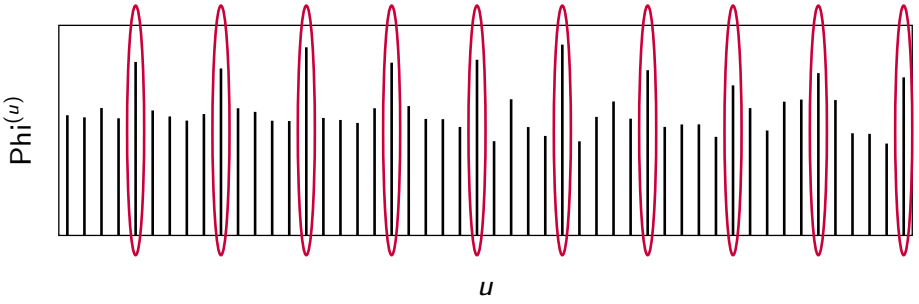
F	N	I	T	E	D	S	T	A	T	T	S
U	B	T	E	I	S	H	I	W	G	E	H

- Decodierung mit `UNITEDSTATES` ergibt:
  `mrale xande rhowi sitth atthe messe ngera rrive shere atthe samet imewi ththe satur dayco urier andot hersa turda ypape rswhe nacco rding tothe datei tispu blish edthr eeday sprev iousi sthef aultw ithyo uorth epost maste rs`

- Analyse der Kolonnen mittels Chi immer möglich, wenn Vorschrift bekannt, nach der die Alphabete gebildet werden – also ggf. P und/oder R bekannt
- falls vertikal verschobenes Referenzalphabet: Kolonnen gegenseitig zurechtrücken $\Rightarrow$ Reduktion auf monoalphabetischen Fall
- Beispiel:

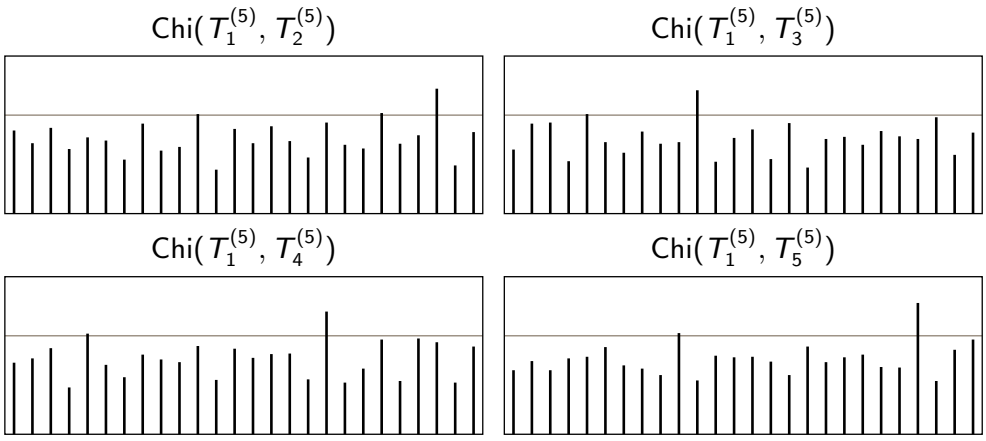
```
SWWJR GPRDN FMWJE XEWGR ZJQDN VJZRV SZXOJ VWWRO VBHRM MOFDL IPAXV
EZWUT CZOZA AQQJL UPKZZ XUMJA PCZOE BAWZR ZYKZI POFOL UOCRE NYKRI
CAMOX IOORR ZJKOL VWWJN VPKZA AFOCA MZOMR CJZDY EJXEL XRFQI ZJCMA
RJVWI DSWZX ASOTR BJBZO QPXMI PDJVZ ZXHGQ SZFDQ FJZJR BMWIC EZMWL
MECVY VWZOX TWHSR UUBMT NSJDW SSOOW CUNJY VJEWI PPFSL MOQVY CVWRI
SMMHW XMEJY NUZMV MXWCR NBRDE SNB
```

- Phi-Analyse:



$\Rightarrow d = 5$

- Verschiebungen gegenüber der ersten Kolonne:



- zurechtgerückte Alphabete auf Basis der Maxima:

1. Kolonne	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
2. Kolonne	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
3. Kolonne	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
4. Kolonne	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
5. Kolonne	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

- bei uneindeutiger Chi-Analyse: Betrachtung anderer Kolonnenpaare

- Geheimtext mit zurechtgerückten Alphabeten:

```
SZMSV GSHMR FPMSI XHMPV ZMGMR VMPAZ SCNXN VZMAS VEXAQ MRVMP ISQGZ
ECMDX CCEIE ATGSP USAID XXCSE PFPXI BDMIV ZBAIM PRVXP URSAI NBAAM
CDCXB IREAV ZMAXP VZMSR VSAIE AIELE MCEVV CMPMC EMNNP XUVZM ZMSVE
RMLFM DVMIB AVECV BMRIS QSNVM PGZED ZAXPU SCVMU FMPSV BPMRG ECCFP
MHSEC VZPXB TZXBV UXRVX NVZMA SVEXA CXDSC VMUFM PSVBP MRGEC CYMAM
SPCQA XPUSC NXPVZ MAMLV NEHMI SQR
```

- Buchstabenhäufigkeiten:

- Zuordnung der drei häufigsten Buchstaben zu denen des Englischen $\mathtt{M} = \mathtt{e}, \mathtt{V} = \mathtt{t}, \mathtt{S} = \mathtt{a}$ führt zum erfolgreichen Einstieg

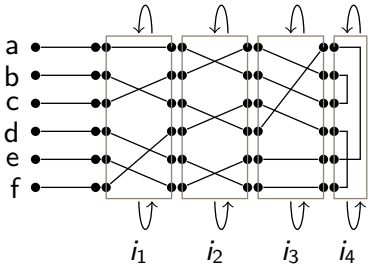
3.5 Kommerzielle Enigma

- ROTOR-Schritte gut elektromechanisch realisierbar
- nur $N = 26$ Begleitalphabete

$\Rightarrow$ Kaskadierung mehrerer Rotoren

$$\begin{aligned}\chi_{i_1,i_2,i_3} &= \rho^{-i_1} R_N \rho^{i_1} \rho^{-i_2} R_M \rho^{i_2} \rho^{-i_3} R_L \rho^{i_3} \\ &= \rho^{-i_1} R_N \rho^{i_1 - i_2} R_M \rho^{i_2 - i_3} R_L \rho^{i_3}\end{aligned}$$

- Besonderheit der Engima: Umkehrwalze



- vermeintliche Vorteile:
  - involutorisch: kein Umschalten zwischen Ver-/Entschlüsselung
  - sicherer, da jeder Rotor zweimal durchlaufen wird

- Gleichung der (kommerziellen) Enigma:

$$\begin{aligned}\chi_{i_1,i_2,i_3,i_4} &= \rho^{-i_1} R_N \rho^{i_1-i_2} R_M \rho^{i_2-i_3} R_L \rho^{i_3-i_4} U \rho^{i_4-i_3} R_L^{-1} \rho^{i_3-i_2} R_M^{-1} \rho^{i_2-i_1} R_N^{-1} \rho^{i_1} \\ &= S_{i_1,i_2,i_3,i_4} U S^{-1}_{i_1,i_2,i_3,i_4}\end{aligned}$$

- Rotorstellungen im ersten Schritt Teil des Schlüssels, danach Weiterschaltung als Zählwerk
  - R_N wird bei jedem Zeichen weiterbewegt
  - R_M wird einmal pro Umdrehung von R_N weiterbewegt
  - R_L wird einmal pro Umdrehung von R_M weiterbewegt
  - U bleibt fest

  $\Rightarrow$ Schlüsselperiode $d = 26^3 = 17576$ länger als zu verschlüsselnde Texte
- Schlüssel: Lage der drei Rotoren und Anfangstellung: $6 \cdot 26^4 = 2\,741\,856$ mögliche Schlüssel; zur damaligen Zeit zu viele für exhaustive Suche

3.6 Kerckhoffs' Superimposition

- Angriffsmöglichkeit gegen alle polyalphabetischen Substitutionen
- Voraussetzung: Mehrere mit dem gleichen Schlüssel verschlüsselte Nachrichten (Geheimtext-Geheimtext-Kompromittierung)
- Angriff über die Buchstabenhäufigkeiten an der selben Position auftretender Zeichen

```
 1  UHYBR JIMBC ...
 2  UHWPR BQLKI ...
 3  IEWHC HQKQM ...
 4  UWVRR HIKMC ...
 5  UHSHA HKSVC ...
 6  YHVHM AGQKC ...
 7  LHVHA AGRLP ...
 8  SWUIR XICJU ...
 9  UHWHV AYULC ...
10  YWXHY HBALG ...
11  WQREX BIENH ...
12  SWUHD HPJJC ...
13  GQVQR VOTQQ ...
```

vermutlich $\mathtt{H}^{(2)} = \mathtt{e}$, $\mathtt{H}^{(4)} = \mathtt{e}$, $\mathtt{R}^{(5)} = \mathtt{e}$, ...

3.7 Isomorphie-Methode

- Geheimtext sei `FGCVT DRQDK NJHOX XQVHK NBRVX AKPZX FCGQU OCOCB YG`, wahrscheinliches Wort `machine`
- alle möglichen Lagen (viele!) mit negativer Mustersuche bestimmen

```
FGCVTDRQDKNJHOXXQVHKNBRVXAKPZXFCGQUOCOCBYG
machine
 machine
  machine
   machine
    machine
     machine
      machine
       machine
        machine
         machine
          machine
           machine
            machine
             machine
                  ⋱
```


- zerlege Verschlüsselungsvorgang gemäß

$$\chi_{i_1,i_2,i_3,i_4} = \rho^{-i_1} R_N \rho^{i_1} U'_{i_2,i_3,i_4} \rho^{-i_1} R_N^{-1} \rho^{i_1}$$

- in kurzen Abschnitten i_2, i_3, i_4 wahrscheinlich konstant, also U'_{i_2,i_3,i_4} als unbekannte Umkehrwalze auffassen
- versuche ersten Schritt der Verschlüsselung $\rho^{-i_1} R_N \rho^{i_1}$ und ersten Schritt der Entschlüsselung (ebenfalls $\rho^{-i_1} R_N \rho^{i_1}$) für alle möglichen Lagen des wahrscheinlichen Wortes, alle Rotoren und alle möglichen Initialwerte von i_1 (Automatisierung!)
- Beispiel für erste mögliche Lage (`GCVTDRQ`) und initiales $i_1 = 0$ mit dem Rotor R_N

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

  - `machine` $\Rightarrow$ `VXNCDJJ`
  - `GCVTDRQ` $\Rightarrow$ `HQVBAAW`
  - damit kann U_{i_2,i_3,i_4} keine konstante, echt involutorische Permutation sein (`V`, `J` und `A`!), diese Konfiguration "kreischt"
- wird zweiter Rotor im betrachteten Abschnitt fortgeschaltet: getrennte Betrachtung beider Teile
- anschließend nähere Analyse der möglichen Konfigurationen (wenige)

- eine mögliche Konfiguration: Lage ab dem 14. Zeichen (OXXQVHK) , initiales $i_1 = 13$, Rotor R_N (wie eben)

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

  - machine $\Rightarrow$ JJMRSAD
  - OXXQVHK $\Rightarrow$ CCTBQXU
  - damit schon (AX)(BR)(CJ)(DU)(MT)(QS) von U_{i_2,i_3,i_4} bekannt
- FGCVT DRQDK NJHOX XQVHK NBRVX AKPZX FCGQU OCOCB YG
  MNNXJ IQVQY COUCC TBQXU NCEAU OQSLQ TORVD VEDKE EQ
  T??AC ?S?S? J?DJJ MRSAD ?J?XD ?
  t??qu ?l?t? o?ama chine ?e?en ?
- Konfiguration plausibel
- noch offen: Lage der anderen beiden Rotoren und initiale Werte von i_2, i_3, i_4, also $2 \cdot 26^3 = 35152$ Möglichkeiten; Katalog oder (automatisiertes) ausprobieren möglich
- schließlich Entzifferung zu thequ ality ofama chine depen dslar gelyo nitsu se
- falls Fortschaltung des zweiten Rotors während des wahrscheinlichen Wortes: Ansatz für ersten Rotor wiederholbar

3.8 Wehrmachts-Enigma

- basiert auf kommerzieller Enigma: drei Rotoren und Umkehrwalze
- nur eine Stellung der Umkehrwalze
- später mehr als drei Rotoren, aus denen gewählt werden konnte
- verstellbare Ringe an den Rotoren kontrollierten Zusammenhang zwischen Rotorlage und dessen alphabetischer Kodierung sowie den Weiterschaltungszeitpunkt des jeweils langsameren Rotors
- zusätzlich involutorische Permutation am Eingang/Ausgang durch steckbare Kabel $\Rightarrow$ Isomorphie-Angriff unmöglich
- Gleichung der Wehrmachts-Enigma:

$$\begin{aligned}\chi_{i_1,i_2,i_3} &= T\rho^{-i_1}R_N\rho^{i_1-i_2}R_M\rho^{i_2-i_3}R_L\rho^{i_3}U\rho^{-i_3}R_L^{-1}\rho^{i_3-i_2}R_M^{-1}\rho^{i_2-i_1}R_N^{-1}\rho^{i_1}T^{-1}\\ &= TS_{i_1,i_2,i_3}US_{i_1,i_2,i_3}^{-1}T^{-1}\end{aligned}$$

  wobei U fest, R_N, R_M, R_L aus bis zu acht Rotoren ausgewählt, T beliebig, aber involutorisch, und Weiterschaltung $i_1 \Rightarrow i_2 \Rightarrow i_3$ von Ringstellung abhängig
- Rotoren und Umkehrwalze verschieden von kommerzieller Enigma, aber sukzessive aufgeklärt worden

- Schlüssel jeweils für einen Tag gültig
- zur Vermeidung eines Superimpositionsangriffs: Spruchschlüssel
- Tagesschlüssel enthält Rotorenlage, Ringstellung, Steckverbindungen und Grundstellung
- Chiffrierer wählt Spruchschlüssel: Ausgangsposition der Rotoren für eigentliche Nachricht
- *verdoppelter* Spruchschlüssel wird mit Grundstellung aus Tagesschlüssel verschlüsselt und übertragen
- eigentliche Nachricht wird mit Spruchschlüssel chiffriert
- Beispiel: Grundstellung `RTJ`, Spruchschlüssel `FQK`; Verschlüsselung von `FQKFQK` mit `RTJ` sei `WAHWIK`: Chiffrat beginnt mit `WAHWIK`, es folgt eigentliche Nachricht mit `FQK` verschlüsselt

3.9 Katalog von Charakteristiken von Rejewski

- erster und vierter Buchstabe Y_1 und Y_4 jeder Nachricht verschlüsseln den gleichen Klartextbuchstaben (erster Buchstabe Spruchschlüssel), also

$$\chi^{-1}_{i_1,i_2,i_3}(Y_1) = \chi^{-1}_{i_1+3,i_2,i_3}(Y_4),$$

  wobei i_1, i_2, i_3 Grundstellung aus Tagesschlüssel
- einsetzen und umformen:

$$\begin{aligned} Y_1 &= \chi_{i_1,i_2,i_3}\left(\chi_{i_1+3,i_2,i_3}(Y_4)\right) \\ &= Y_4 \quad TS_{i_1+3,i_2,i_3} US^{-1}_{i_1+3,i_2,i_3} T^{-1} \quad TS_{i_1,i_2,i_3} US^{-1}_{i_1,i_2,i_3} T^{-1} \\ &= Y_4 \quad TS_{i_1+3,i_2,i_3} US^{-1}_{i_1+3,i_2,i_3} S_{i_1,i_2,i_3} US^{-1}_{i_1,i_2,i_3} T^{-1} \\ &= Y_4 \quad TP_{i_1,i_2,i_3} T^{-1} \end{aligned}$$

  mit $P_{i_1,i_2,i_3} = S_{i_1+3,i_2,i_3} US^{-1}_{i_1+3,i_2,i_3} S_{i_1,i_2,i_3} US^{-1}_{i_1,i_2,i_3}$
- ausreichend aufgefangene Nachrichten vom selben Tag erlauben Rekonstruktion von $TP_{i_1,i_2,i_3} T^{-1}$
- entscheidend ist Zyklusstruktur von $TP_{i_1,i_2,i_3} T^{-1}$, da diese *von T unabhängig*
- entsprechend Kombination von Y_2/Y_5 und Y_3/Y_6 für P_{i_1+1,i_2,i_3} und P_{i_1+2,i_2,i_3}

Beispiel:
aufgefangene chiffrierte Spruchschlüssel zum gleichen Tagesschlüssel: QGALYB, RJLWPX, XRSGNM, ...
ergeben $TP_{i_1,i_2,i_3}T^{-1}$:

Y_1	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
Y_4	A	C	B	V	I	K	Z	T	J	M	X	H	U	Q	D	F	L	W	S	E	N	P	R	G	O	Y

mit den Zyklen (A)(BC)(DVPFKXGZYO)(EIJMUNQLHT)(RW)(S) der Längen 1, 1, 2, 2, 10 und 10.

- jede Zyklenlänge tritt doppelt auf; Charakteristik enthält sie nur einfach (im Beispiel Charakteristik=1, 2, 10)
- Kombination der Charakteristika von Y_1/Y_4, Y_2/Y_5, Y_3/Y_6 bestimmt Grundstellung und Lage der Rotoren aus Tagesschlüssel eindeutig
- Katalogisierung aller $6 \cdot 26^3 = 105456$ möglichen Charakteristika
- Bestimmung von Ringstellung und Steckerung von Hand unter Ausnutzung wahrscheinlicher Wörter gut machbar $\Rightarrow$ vollständige Rekonstruktion des Tagesschlüssels

3.10 Zygalski-Blätter

Änderung des Verfahrens im September 1938:

- Tagesschlüssel enthält nur noch Rotorenlage, Ringstellung und Steckverbindungen
- Chiffrierer wählt erste Grundstellung der drei Rotoren, diese wird *unverschlüsselt* übertragen (ohne Ringstellung wenig Aussagekraft)
- Chiffrierer wählt Spruchschlüssel: Ausgangsposition der Rotoren für eigentliche Nachricht
- *verdoppelter* Spruchschlüssel wird mit erster Grundstellung verschlüsselt und übertragen
- eigentliche Nachricht wird mit Spruchschlüssel chiffriert
- Beispiel: Grundstellung RTJ, Spruchschlüssel FQK; Verschlüsselung von FQKFQK mit RTJ sei WAHWIK: Chiffrat beginnt mit RTJWAHWIK, es folgt eigentliche Nachricht mit FQK verschlüsselt

$\Rightarrow$ keine Gewinnung der Charakteristika mehr möglich

- verdoppelter Spruchschlüssel erlaubt immer noch Angriff
- Betrachtung von Fällen, bei denen an 1. und 4., 2. und 5. oder 3. und 6. Stelle des chiffrieten Spruchschlüssels gleiches Zeichen auftritt
- Beispiel: WAHWIK
- also $\chi_{i_1,i_2,i_3}(x) = \chi_{i_1+3,i_2,i_3}(x) = \mathtt{W}$ (wobei x noch unbekanntes erstes Zeichen des Spruchschlüssels)
- damit auch $\chi_{i_1,i_2,i_3}(\mathtt{W}) = \chi_{i_1+3,i_2,i_3}(\mathtt{W}) = x$ und insbesondere $\chi_{i_1+3,i_2,i_3}\left(\chi_{i_1,i_2,i_3}(\mathtt{W})\right) = \mathtt{W}$
- W ist also Fixpunkt von $\chi_{i_1+3,i_2,i_3}\chi_{i_1,i_2,i_3}$
- aufgrund der Steckerkonfiguration ist damit nur klar, dass die Rotorenkonfiguration ebenfalls (mindestens) einen Fixpunkt aufweist, aber nicht für welchen Buchstaben
- nur für etwa 40% aller i_1, i_2, i_3 enthält $\chi_{i_1+3,i_2,i_3}\chi_{i_1,i_2,i_3}$ einen Fixpunkt
- nach Auftreten von etwa zwölf Fixpunkten i_1, i_2, i_3 eindeutig (beachte bekannte Beziehung zur Ringstellung durch bekannte Grundstellung)

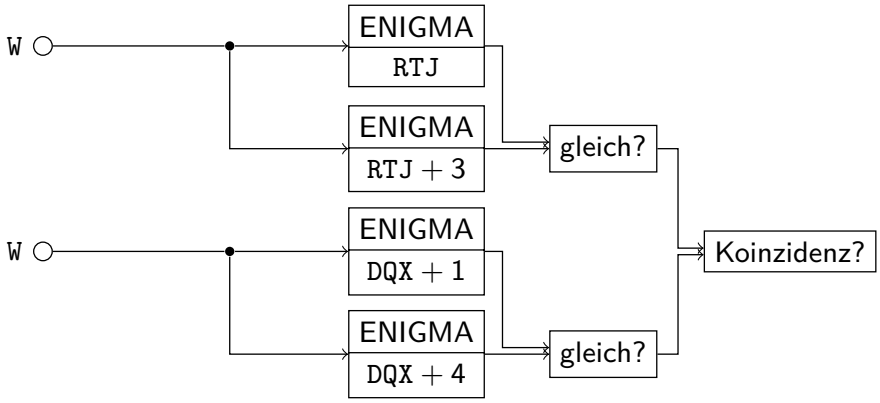
Zygalski-Blätter:

- ein Blatt pro Kombination Rotorenauswahl, Stellung langsamster Rotor und Lage des Fixpunktes (1/4, 2/5 oder 3/6)
- jedes Blatt enthält 51×51-Raster der (wiederholten) möglichen Stellungen von schnellem und mittlerem Rotor
- Loch im Raster, falls die entsprechende Konfigurationen einen Fixpunkt hat
- je ein Stapel für alle möglichen Rotorenauswahlen und Grundstellungen des langsamsten Rotors
- Blätter auf Stapel entsprechend der übertragenen Grundstellung zurechtgeschoben
- Deckungsgleiche Löcher ergeben mögliche Konfigurationen
- bei nur drei Rotoren: $6 \cdot 26 = 156$ Stapel, aber schon bei fünf Rotoren $60 \cdot 26 = 1560$ Stapel nötig

3.11 Turing-Bombe

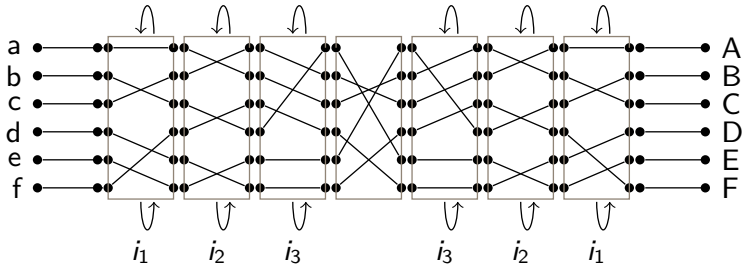
Grundlage: polnische Bombe

- nutzt ebenfalls den verdoppelten Spruchschlüssel
- funktioniert nur, wenn wiederholter Buchstabe im chiffrierten Spruchschlüssel "ungesteckert"
- benutzt parallele Enigmas (Nachbauten), um wiederholte Buchtaben sukzessive mit allen Rotorstellungen zu (de)chiffrieren; für `RTJWAHWIK` und `DQXDWJMWR`:



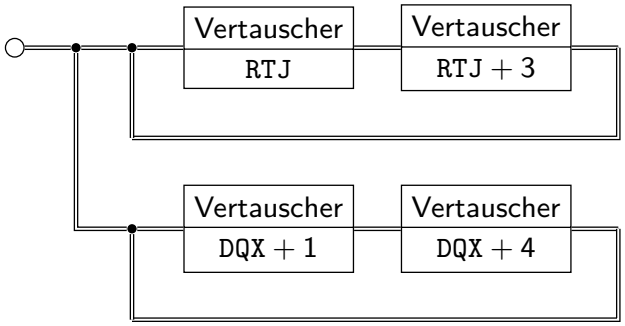
- automatischer Stopp bei passender Konfiguration

- ersetze Anordnung mit Umkehrwalze durch lineare Anordung, in der R_N, R_M, R_L je zweifach (einmal gespiegelt, also invers) vorkommen (Vertauscher, "double-ended scrambler", "commutator")



- Vorteil: unabhängige Wirkung auf alle 26 Buchstaben

- baue rückgekoppeltes Netzwerk aus Vertauschern (wieder für `RTJWAHWIK` und `DQXDWJMWR`):

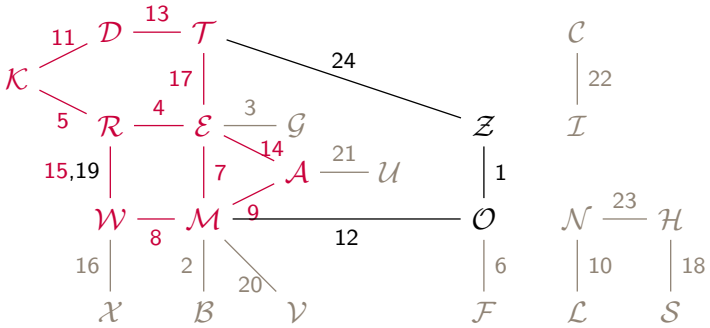


- an Leitung für `W` wird eine Spannung angelegt
  - ist die Konfiguration falsch und enthält keinen Fixpunkt, sind alle Leitungen verbunden, alle stehen unter Spannung
  - enthält die Konfiguration einen Fixpunkt bei `W`, so bleiben alle anderen Leitungen Spannungsfrei (galvanische Trennung der `W`-Leitung)
  - enthält die Konfiguration einen Fixpunkt bei x (Stecker!), so bleibt *nur* die x-Leitung spannungsfrei (galvanische Trennung der x-Leitung)
- simultanes Weiterschalten aller Vertauscher, bis mögliche Konfiguration gefunden; dann auch ggf. nötiger Stecker gefunden
- mehrere parallele Netzwerke für verschiedene verdoppelte Buchstaben möglich

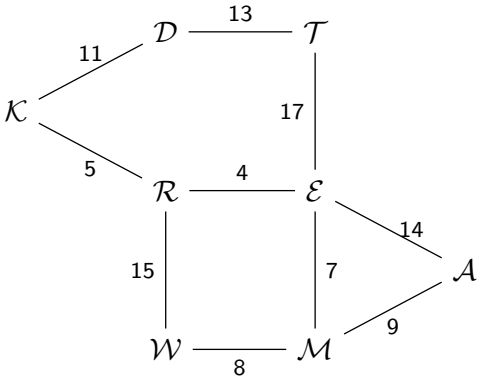
- Prototyp der Turing-Bombe am 18. März 1940 einsatzbereit, reguläre Inbetriebnahme am 26. Mai 1940
- Deutsche beendeten die Verdoppelung der Spruchschlüssel zum 1. Mai 1940
- mit der Turing-Bombe war aber auch ein Angriff über ein wahrscheinliches Wort möglich
- vermutete Übereinstimmung:

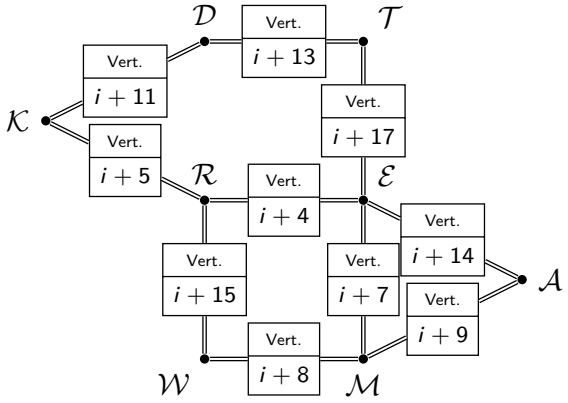
```
oberkommandoderwehrmacht
ZMGERFEWMLKMTAWXTSWVUINZ
```

- Darstellung als Graph:



- Auswahl eines eng vermaschten Teilgraphen mit höchstens zehn Kanten
- Aufbau eines entsprechenden Rückkopplungsgraphen aus Vertauschern als Kanten bildet das *Menü*





- am Knoten $\mathcal{E}$ wird an Leitung für e eine Spannung angelegt
  - ist die Konfiguration falsch und enthält keinen Fixpunkt, sind alle Leitungen verbunden, alle stehen unter Spannung
  - enthält die Konfiguration einen Fixpunkt bei e, so bleiben alle anderen Leitungen Spannungsfrei (galvanische Trennung der e-Leitung)
  - enthält die Konfiguration einen Fixpunkt bei x (Stecker!), so bleibt *nur* die x-Leitung spannungsfrei (galvanische Trennung der x-Leitung)
- simultanes Weiterschalten aller Vertauscher, bis mögliche Konfiguration gefunden; dann auch ggf. nötige Stecker gefunden

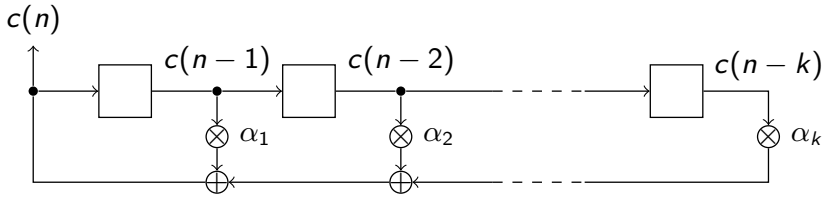
- ist eine Konfiguration mit Fixpunkt gefunden, kann jedem Knoten der entsprechende Fixpunkt-Buchstabe (entsprechend Steckerung) zugeordnet werden, beispielsweise

$\mathcal{A}$	$\mathcal{D}$	$\mathcal{E}$	$\mathcal{K}$	$\mathcal{M}$	$\mathcal{R}$	$\mathcal{T}$	$\mathcal{W}$
a	t	e	f	x	z	d	m

- Zuordnung muss involutorisch sein, hier verletzt
- wünschenswert: in einem solchen Fall Bombe nicht stoppen
- Lösung: Welchmans “diagonal board”
  - verbinde Leitung d von Knoten $\mathcal{A}$ mit Leitung a von Knoten $\mathcal{D}$, Leitung e von Knoten $\mathcal{A}$ mit Leitung a von Knoten $\mathcal{E}$, . . .
  - . . . und insbesondere auch Leitung w von Knoten $\mathcal{M}$ mit Leitung m von Knoten $\mathcal{W}$
  - letzteres führt zu einer Verbindung des Fixpunkt-Stromkreises mit der restlichen Schaltung, die Bombe stoppt nicht
  - die Verbindung von Leitung t von Knoten $\mathcal{D}$ mit Leitung d von Knoten $\mathcal{T}$ hat dagegen (richtigerweise) keinen Einfluss bei dieser Konfiguration

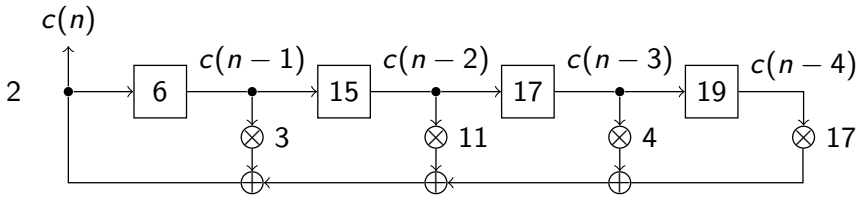
3.12 Quasiperiodische Schlüssel mit Schieberegistern

- selbst einfache polyalphabetische Verfahren (z.B. VIGENÈRE) sind sicher, wenn Periode des Schlüssels mindestens ähnlich der Textlänge
- Idee: systematische Erzeugung einer Zeichenfolge für z.B. VIGENÈRE; diese Systematik parametrierbar, Parameter sind Schlüssel
- Ansatz: linear rückgekoppeltes Schieberegister



  alle Rechenoperationen in $\mathbb{Z}_N$
- Schlüssel: initiale Belegung der Register und Wahl der Koeffizienten α
- Vorsicht bei Wahl der Koeffizienten α, um zu kurze Perioden zu vermeiden

Beispiel:



$$3 \cdot 6 + 11 \cdot 15 + 4 \cdot 17 + 17 \cdot 19 = 574 \simeq 2$$

Klartext	a	b	r	o	...
Schlüsselfolge	2	5	4	21	...
Geheimtext	C	G	V	J	...

Ein über ein lineares Schieberegister der Länge k erzeugter Schlüssel lässt sich mittels eines wahrscheinlichen Wortes der Mindestlänge $2k$ leicht zurückgewinnen.
Beispiel (Voraussetzung: $k \leq 4$):

Geheimtext	C	G	V	J	F	M	C	I	H	T	X	U	F	...
angen. Klartext	b	r	o	a	d	c	a	s	t					
Schlüsselfolge	1	15	7	9	2	10	2	16	14					

Es muss gelten:

$$c(n) = \alpha_1 c(n-1) + \alpha_2 c(n-2) + \alpha_3 c(n-3) + \alpha_4 c(n-4)$$

Also:

$$\begin{pmatrix} 2 \\ 10 \\ 2 \\ 16 \\ 14 \end{pmatrix} = \begin{pmatrix} 9 & 7 & 15 & 1 \\ 2 & 9 & 7 & 15 \\ 10 & 2 & 9 & 7 \\ 2 & 10 & 2 & 9 \\ 16 & 2 & 10 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}$$

Dieses überbestimmte Gleichungssystem besitzt keine Lösung $\Rightarrow$ (mindestens) eine Annahme ist falsch

Versuche nächste Lage des wahrscheinlichen Wortes:

Geheimtext	C	G	V	J	F	M	C	I	H	T	X	U	F	...
angen. Klartext	a	b	r	o	a	d	c	a	s	t	i	n	g	...
Schlüsselfolge	2	5	4	21	5	9	0	8	15	0	15	7	25	...

Also:

$$\begin{pmatrix} 9 \\ 0 \\ 8 \\ 15 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 4 & 5 \\ 9 & 5 & 21 & 4 \\ 0 & 9 & 5 & 21 \\ 8 & 0 & 9 & 5 \\ 15 & 8 & 0 & 9 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}$$

Dieses überbestimmte Gleichungssystem besitzt die Lösung $\alpha_1 = 3$, $\alpha_2 = 11$, $\alpha_3 = 4$, $\alpha_4 = 17$. Damit lässt sich die Schlüsselfolge in beide Richtungen fortsetzen und der Text entschlüsseln.

Randbemerkung: Lösung von Gleichungssystemen über $\mathbb{Z}_N$

- Zu lösen:

$$Ax \equiv b \mod N, \quad A \in \mathbb{Z}_N^{m \times k}, x \in \mathbb{Z}_N^k, b \in \mathbb{Z}_N^m, \quad N = p_1 p_2 \cdots p_n$$

- Äquivalent:

$$Ax \equiv b \mod p_1 \ \wedge \ Ax \equiv b \mod p_2 \ \wedge \ldots \wedge \ Ax \equiv b \mod p_n$$

- jeweils Lösung über Gaußsches Eliminationsverfahren
- Multiplikation mit dem Inversen statt Division
- Bestimmung des Inversen mit erweitertem euklidschen Algorithmus; Beispiel zur Invertierung von 9 in $\mathbb{Z}_{13}$:

$$13 = 1 \cdot 9 + 4$$
$$9 = 2 \cdot 4 + 1$$

$$\begin{aligned} 1 &= 9 - 2 \cdot 4 \\ &= 9 - 2 \cdot (13 - 1 \cdot 9) = 3 \cdot 9 - 2 \cdot 13 \end{aligned}$$

also $3 \cdot 9 \equiv 1$ und damit ist 3 das Inverse zu 9

Beispiel:

$$\begin{pmatrix} 17 & 1 \\ 8 & 14 \\ 7 & 16 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \equiv \begin{pmatrix} 8 \\ 12 \\ 6 \end{pmatrix} \mod 26$$

Teillösung in $\mathbb{Z}_2$:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \mod 2 \quad \Rightarrow \quad \alpha_1 \equiv 0, \alpha_2 \equiv 0 \mod 2$$

Teillösung in $\mathbb{Z}_{13}$:

$$\begin{pmatrix} 4 & 1 \\ 8 & 1 \\ 7 & 3 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \equiv \begin{pmatrix} 8 \\ 12 \\ 6 \end{pmatrix} \mod 13 \qquad 4^{-1} \equiv 10 \mod 13$$

$$\begin{pmatrix} 1 & 10 \\ 0 & 12 \\ 0 & 11 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \equiv \begin{pmatrix} 2 \\ 9 \\ 5 \end{pmatrix} \mod 13 \qquad 12^{-1} \equiv 12 \mod 13$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \equiv \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} \mod 13 \quad \Rightarrow \quad \alpha_1 \equiv 1, \alpha_2 \equiv 4 \mod 13$$

Zusammen: $\alpha_1 \equiv 14, \alpha_2 \equiv 4 \mod 26$

3.14 Nicht-periodische Schlüssel mittels Auto-Key

- komplett nicht-periodische Schlüssel lassen sich erreichen, wenn der Text in die Schlüsselerzeugung einbezogen wird
- grundsätzlich problematisch: Verschleppung von Chiffrierfehlern
- einfache Auto-Key-Verfahren oft leicht angreifbar
- Beispiel:
  - VIGENÈRE-Verschlüsselung
  - erste d Klartextzeichen $x_1 \dots x_d$ werden mit geheimem Schlüsselkeim $c_1 \dots c_d$ chiffriert
  - danach wird i-tes Zeichen mit $c_i = x_{i-d}$ chiffriert
  - es gilt jedoch

$$\begin{aligned} y_i &= x_i + c_i && \text{für } 1 \leq i \leq d \\ y_{i+d} - y_i &= x_{i+d} - c_i \\ y_{i+2d} - y_{i+d} + y_i &= x_{i+2d} + c_i \\ y_{i+3d} - y_{i+2d} + y_{i+d} - y_i &= x_{i+3d} - c_i \\ &\vdots \end{aligned}$$

  - effektiv also lediglich Verdoppelung der Periode, mit abhängigen Schlüsseln der beiden Hälften
- ähnliche Ansätze als Stromchiffren aber durchaus effektiv

4 Polygraphische Substitution

- polygraphisch: mehrere Klartextzeichen werden auf einmal verschlüsselt
- Anzahl möglicher Substitutionen steigt rasant; bei Verschlüsselung von n Zeichen aus einem Alphabet der Mächtigkeit $N = 26$:

n	1	2	3
$(N^n)!$	$4{,}03 \cdot 10^{26}$	$1{,}88 \cdot 10^{1621}$	$1{,}19 \cdot 10^{66\,978}$

- allgemeine Substitution (keine Systematik) nur bis etwa $n = 3$ handhabbar (Tabelle mit $26^3 = 17576$ Einträgen)