8.3 Klassische Verfahren

Klassisch heißen Verfahren, die ohne Computer auskamen. Sie arbeiten mit zwei Grundoperationen, die – in ganz anderem Gewand – bis heute in jeder modernen Blockchiffre stecken.

- **Transposition:** Die Buchstaben bleiben, ihre Reihenfolge ändert sich.
- **Substitution:** Die Reihenfolge bleibt, die Buchstaben werden ersetzt.

8.3.1 Transposition: die Skytala

Die älteste überlieferte Chiffre ist rund 2400 Jahre alt. Die Spartaner wickelten einen schmalen Lederstreifen um einen Stab bestimmter Dicke, die **Skytala**, und schrieben die Nachricht längs des Stabes. Abgewickelt ergab der Streifen eine sinnlose Buchstabenfolge; nur ein Stab desselben Durchmessers machte sie wieder lesbar. Der Schlüssel ist also die Dicke des Stabes.



Eine Skytala. Auf dem aufgewickelten Streifen steht die Nachricht längs des Stabes und ist lesbar. Wickelt man den Streifen ab, stehen die Buchstaben in vertauschter Reihenfolge – der Text ist erst wieder zu entziffern, wenn der Empfänger einen Stab desselben Durchmessers benutzt.

Informatisch ist die Skytala nichts anderes als eine Tabelle: Man schreibt den Text zeilenweise in eine Tabelle mit s Spalten und liest ihn spaltenweise wieder aus.

```
Klartext: ANGRIFFBEIMORGENGRAUEN,  s = 4

A N G R        Auslesen spaltenweise:
I F F B        AIERGE NFIGRN GFMEA RBONU
E I M O
R G E N
G R A U
E N
```

Der Angriff ist einfach: Es gibt nur wenige sinnvolle Spaltenzahlen, man probiert sie durch. Die Skytala scheitert schon im Ciphertext-only-Modell.

8.3.2 Substitution: die Cäsar-Chiffre

Bei der **Cäsar-Chiffre** wird das Alphabet zyklisch um k Stellen verschoben. Nach Z folgt wieder A. Der Schlüssel ist die Zahl k.

```
Klartext:   A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Geheim (k=4): 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
```

Rechnet man die Buchstaben in Zahlen um (A = 0, B = 1, ..., Z = 25), wird daraus eine Addition mit Rest:

$$c = (m + k) \bmod 26 \qquad \text{und} \qquad m = (c - k) \bmod 26$$

Dabei ist m der Klartext- und c der Geheimtextbuchstabe. Das „mod 26“ erledigt genau den Übergang von Z zurück zu A.

Weil es nur 25 brauchbare Schlüssel gibt, ist die Cäsar-Chiffre mit Bleistift und Papier in Minuten zu brechen: Man probiert alle durch (**Brute-Force-Angriff**). Ein Computer erledigt das in Mikrosekunden.

8.3.3 Allgemeine monoalphabetische Substitution

Man kann jedem Buchstaben statt der Verschiebung eine beliebige Zuordnung geben – dann ist der Schlüssel eine vollständige Vertauschung des Alphabets. Davon gibt es

$$26! = 403\,291\,461\,126\,605\,635\,584\,000\,000 \approx 4 \cdot 10^{26}$$

Stück. Eine Maschine, die eine Milliarde Schlüssel pro Sekunde prüft, wäre damit rund 13 Milliarden Jahre beschäftigt — ungefähr so lange, wie das Universum alt ist. Trotzdem ist das Verfahren unsicher, und diese Diskrepanz ist die wichtigste Lehre dieses Abschnitts.

Merke

Ein großer Schlüsselraum ist notwendig, aber nicht hinreichend für Sicherheit. Er schützt nur vor Brute Force, nicht vor einem klügeren Angriff.

8.3.4 Kryptoanalyse durch Häufigkeitsanalyse

Der klügere Angriff nutzt die Statistik der Sprache. In deutschen Texten kommen die Buchstaben höchst ungleich vor:

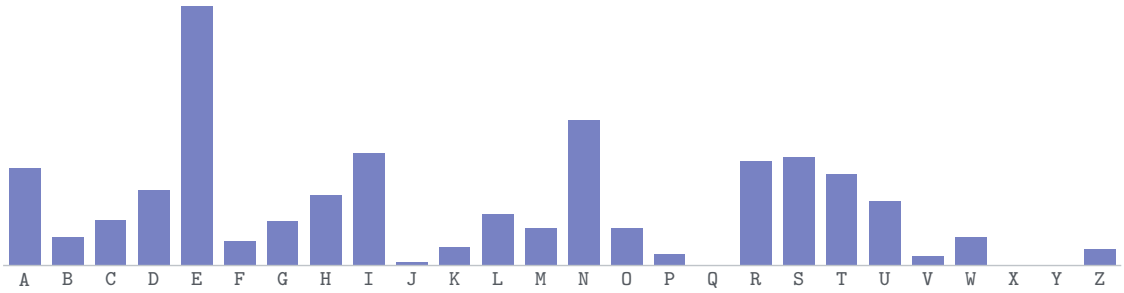
Buchstabe	E	N	I	S	R	A	T	D	H	U	L	C
Anteil %	17,4	9,8	7,6	7,3	7,0	6,5	6,2	5,1	4,8	4,4	3,4	3,1

Allein E und N machen gut ein Viertel jedes Textes aus, die zwölf häufigsten Buchstaben rund 80 %. Eine monoalphabetische Substitution verschiebt diese Verteilung nur, sie zerstört sie nicht: Der häufigste Geheimtextbuchstabe ist mit großer Wahrscheinlichkeit das E.

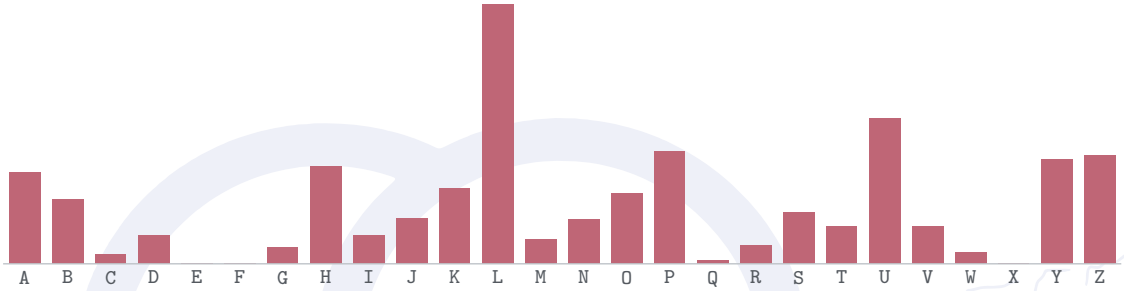
> In der Webfassung steht hier ein Bedienfeld zum Ausprobieren: **Verschlüsselungswerkzeug**.

Probieren Sie es im Bedienfeld aus: Bei Cäsar wandert die Verteilung geschlossen um k Stellen nach rechts — die Form bleibt erkennbar. Schalten Sie dann auf Vigenère um, und die Verteilung flacht ab. Genau deshalb braucht der Angriff auf Vigenère einen zweiten Gedanken, den Kasiski-Test aus dem nächsten Abschnitt.

Deutscher Klartext



Geheimtext nach Cäsar mit $k = 7$



Dieselben Balken, nur um sieben Stellen gewandert. Die Substitution verschiebt die Verteilung — sie zerstört sie nicht. Der höchste Balken ist immer noch das E, es heißt jetzt nur L.

Buchstabenhäufigkeit im Deutschen, darunter dieselbe Verteilung nach einer Cäsar-Chiffre mit $k = 7$.

Das untere Diagramm ist der ganze Angriff: Man sucht den höchsten Balken und weiß, dass er das E sein muss. Daraus ergibt sich der Schlüssel, und der Rest fällt von allein.

Hilfreich sind außerdem **Bigramme** (Buchstabenpaare). Im Deutschen sind ER, EN, CH, DE und EI besonders häufig; CH und SCH verraten sich schnell. Bei kurzen oder ungewöhnlichen Texten versagt die Methode allerdings – ein Text „über den Einfluss von Ozon auf die Zebras im Zentrum von Zaire“ hat eine ganz andere Statistik als die Zeitung von gestern.

Arbeitsauftrag 3

6. Verschlüsseln Sie `INFORMATIK` mit der Cäsar-Chiffre und $k = 7$.
7. Entschlüsseln Sie den Geheimtext `OBXE LITLL UXBF XGMLVAENXLLXEG`. Notieren Sie Ihren Weg.
8. Schreiben Sie ein Programm, das zu einem Text und einem Schlüssel k den Geheimtext liefert, und ein zweites, das zurückrechnet. Buchstaben außerhalb von A–Z sollen unverändert bleiben.
9. Erweitern Sie Ihr Programm so, dass es alle 25 Möglichkeiten ausgibt – der vollständige Brute-Force-Angriff.
10. Schreiben Sie ein Programm, das die Häufigkeit aller Zeichen eines Textes zählt und als Balkendiagramm ausgibt. Testen Sie es an einem längeren deutschen Text.
11. Begründen Sie, warum ein Brute-Force-Angriff auf die allgemeine monoalphabetische Substitution aussichtslos, das Verfahren aber trotzdem unsicher ist.