6 Autokey Ciphers

The first one to propose autokey ciphers was BELLASO in 1564. Also this cipher is often attributed to VIGENÈRE.

Encryption and Decryption

The alphabet Σ is equipped with a group operation $*$. As key chose a string $k \in \Sigma^l$ of length l. For encrypting a plaintext $a \in \Sigma^r$ one concatenates k and a and truncates this string to r letters. This truncated string then serves as keytext for a running-key encryption:

Plaintext:	a_0	a_1	$\ldots$	a_{l-1}	a_l	$\ldots$	a_{r-1}
Keytext:	k_0	k_1	$\ldots$	k_{l-1}	a_0	$\ldots$	a_{r-l-1}
Ciphertext:	c_0	c_1	$\ldots$	c_{l-1}	c_l	$\ldots$	c_{r-1}

The formula for encryption is

$$c_i = \begin{cases} a_i * k_i & \text{for } i = 0, \ldots l-1, \\ a_i * a_{i-l} & \text{for } i = l, \ldots r-1. \end{cases}$$

Example, $\Sigma = \{\mathtt{A}, \ldots, \mathtt{Z}\}$, $l = 2$, $k = \mathtt{XY}$:

```
P L A I N T E X T
X Y P L A I N T E
-----------------
M J P T N B R Q X
```

Remark: Instead of the standard alphabet (or the TRITHEMIUS table) one could also use a permuted primary alphabet.

Here is the formula for decryption

$$a_i = \begin{cases} c_i * k_i^{-1} & \text{for } i = 0, \ldots l-1, \\ c_i * a_{i-l}^{-1} & \text{for } i = l, \ldots r-1. \end{cases}$$

A Perl program is `autokey.pl` in the web directory `http://www.staff.uni-mainz.de/pommeren/Cryptology/Classic/Perl/`.

Approaches to Cryptanalysis

The four most promising approaches are:

- Exhaustion for small l.
- Interpretation as running-key cipher from position l, in case of a key word or phrase from the plaintext language even from the beginning of the ciphertext:

   - Probable word and zigzag exhaustion
   - Frequent word fragments
   - Frequency analysis
   - Frequent letter combinations

  The repetition of the plaintext in the key makes the task considerably easier.

- Similarity with the TRITHEMIUS-BELLASO cipher, see Section 8 below

- Algebraic cryptanalysis (for known plaintext): Solving equations. We describe this for a commutative group, the group operation written as addition, that is, we consider Σ, Σ^r, and Σ^{r+l} as $\mathbb{Z}$-modules.

We interpret the encryption formula as a system of linear equations with an $r \times (r+l)$ coefficient matrix:

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{l-1} \\ c_l \\ \vdots \\ c_{r-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 1 & & & \\ & 1 & 0 & \dots & 1 & & \\ & & \ddots & \ddots & & \ddots & \\ & & & 1 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} k_0 \\ k_1 \\ \vdots \\ k_{l-1} \\ a_0 \\ \vdots \\ a_{r-1} \end{pmatrix}$$

This is a system of r linear equations with the $r+l$ unknowns (the components of) $k \in \Sigma^l$ and $a \in \Sigma^r$. "In general" such a system is solvable as soon as l of the unknowns are guessed, that means known plaintext of length l (not necessarily connected). Since the involved $\mathbb{Z}$-modules are (in most interesting cases) not vector spaces, solving linear equations is a bit intricate but feasible. This is comprehensively treated in the next chapter.

Ciphertext Autokey

Using ciphertext instead of plaintext as extension of the l-letter key is a useless variant, but also proposed by VIGENÈRE. We only describe it by an example:

Example, $\Sigma = \{\texttt{A}, \dots, \texttt{Z}\}$, $l = 2$, $k = \texttt{XY}$:

```
P L A I N T E X T
X Y M J M R Z K D
-----------------
M J M R Z K D H W
```

Exercise. Give a formal description of this cipher. Why is cryptanalysis almost trivial? Work out an algorithm for cryptanalysis.

Exercise. Apply your algorithm to the cryptogram

```
IHTYE VNQEW KOGIV MZVPM WRIXD OSDIX FKJRM HZBVR TLKMS FEUKE
VSIVK GZNUX KMWEP OQEDV RARBX NUJJX BTMQB ZT
```

Remark: Using a nonstandard alphabet makes this cipher a bit stronger.