BUILDING A BOMBE

The Mathematics and Circuitry that Bested German Encryption

Jonah Weinbaum

Contents

Chapter 1

Permutations

Definition 1.1. A **permutation** σ is a bijective function from a set S to the same set

$$\sigma\colon S \to S$$

A permutation can be thought of as "swapping" elements of our set while. Some simple examples of permutations are

Example 1.2.

(a) $\sigma = \mathrm{id}_S$ known as the **identity permuation**, which maps each element to itself

(b) $S = \{a, b, c, d, \ldots, z\}$ and σ defined by $\sigma(a) = b, \ldots, \sigma(z) = a$. This shifts each letter by one place in the alphabet, wrapping around at z. We will called this the **Caesar permutation** and it will be denoted by θ_1

We will show that for a fixed rotor state the Enigma functions is a permutation on the set of the latin alphabet. It is clear that the enigma machine sends this set to itself but it is not immediately apparent that such a mapping is bijective. To see this we must shift our discription of the enigma to that of a composition of permutations. Let us first define some useful notation.

Definition 1.3. For a set S, the **symmetric group over** S is the group $(G, \circ)$ of all permutations on S, where

$$G \coloneqq \{f : S \to S \mid f \text{ is a bijection}\}$$

where the group operation is composition of functions. We denote this group $\mathrm{Sym}(S)$.

It is clear that this group is closed with respect to composition, since the composition of bijections is itself a bijection. The identity element is simply the identity permutation described in CITE, and the inverse of an element is simply its inverse as a function, since the inverse of a bijection is itself a bijection this is well-defined.

Example 1.4. The Caeser cipher is one of the simplest and earliest encryption schemes. It involves shifting the set of letters by a fixed amount to encode a message. In Caesar's case he would shift each letter in a message by three places, sending $A \mapsto D, \ldots, X \mapsto A, Y \mapsto B, Z \mapsto C$. In the context of permuations, this can be viewed as a repeated application of the Caesar permuation θ_1 described earlier CITE. For instance, to get Caesar's particular cipher we use $\theta_1 \circ \theta_1 \circ \theta_1$ (that is $\theta_1^3 \in \mathrm{Sym}(\{A, \ldots, Z\})$). For ease of notation we define

$$\theta_n \coloneqq \theta_1^n \text{ for } n \in \mathbb{N}$$

Though this continues indefinitely, due to the nature of the permuation it is clear that $\theta_n = \theta_{(n+26)}\ \forall\ n \in \mathbb{N}$ so, in particular, there are only 26 distinct Caesar ciphers – a small keyspace indeed.

Section 1.1

Enigma as a Permutation

Suppose for now that we have fixed the rotors of the enigma machine in place. We want to examine how each letter is mapped through the machine. Recall from CITE that the enigma machine maps each letter signal through the following transformations: the plugboard, three rotors, a reflector panel, back through the rotors, and back through the plugboard. Each of these components can themselves be viewed as a bijective mapping by construction. We will label the permutation associated with each as follows:

(a) $R_{(1,r_1,\ell_1,g_1)}, R_{(2,r_2,\ell_2,g_2)}, R_{(3,r_3,\ell_3,g_3)}$ will be the rotors going from right to left where $r_i \in \{\mathrm{I}, \mathrm{II}, \ldots, \mathrm{VI}\}$ define which rotors we have selected, and $\ell_i, g_i \in \{a, \ldots, z\}$ indicates the *Ringstellung* and *Grundstellung* respectively for that rotor.

(b) P will be our plugboard setting

(c) M_α will be our reflector setting where $\alpha \in \{A, B, C\}$ describes our reflector type

For the rotor and the reflector we include subscripts to describe their settings since their permutation is entirely determined by this short list of parameters. The plugboard, on the other hand, is itself (by the *Steckerverbindungen*) a description of its permutation so subscripts are unnecessary.

Supposing we are at ground position, the engima permutation E is as follows:

$$E = PR_1R_2R_3M_\alpha R_3^{-1}R_2^{-1}R_1^{-1}P^{-1}$$

We note that

$$\begin{aligned} E &= PR_{(1,r_1,\ell_1,g_1)}R_{(2,r_2,\ell_2,g_2)}R_{(3,r_3,\ell_3,g_3)}M_\alpha R^{-1}_{(3,r_3,\ell_3,g_3)}R^{-1}_{(2,r_2,\ell_2,g_2)}R^{-1}_{(1,r_1,\ell_1,g_1)}P^{-1} \\ &= (PR_{(1,r_1,\ell_1,g_1)}R_{(2,r_2,\ell_2,g_2)}R_{(3,r_3,\ell_3,g_3)}) \circ M_\alpha \circ (PR_{(1,r_1,\ell_1,g_1)}R_{(2,r_2,\ell_2,g_2)}R_{(3,r_3,\ell_3,g_3)})^{-1} \end{aligned}$$

Chapter 2

The UK Bombe

Section 2.1

Motivating Example

Section 2.2

Changes to Enigma

Starting in 1940, the German's enhanced the security of their key distribution. As discussed in CITE the *Grundstellung* rotor position was sent along with the daily key and an operator chose a *Spruchschlusse* to encode twice at the start of a message. Later iterations of this protocol removed the *Grundstellung* from key sheets.

These new key sheets contained the following columns columns *Tag/Datum*, *Walzenlage*, *Ringstellung*, *Steckerverbindungen*, and *Kenngruppen*

Notice the removal of the *Grundstellung* as well as the addition of the *Kenngruppen.* The *Kenngruppen* were a set of four trigrams used to identify which setting was being used to encode a message, this is particularly useful if trying to decode a message using a prior day's key. The operator would choose a trigram from the the *Kenngruppen*, append two letters to the front of the trigram, and this five letter combination (known as the *Buchstabenkenngruppe*) would preceed the message being sent. If a message was sent in multiple segments, multiple *Buchstabenkenngruppe* were used to start each segment.

When sending a message the operator was to use the following protocol

I. The time at which the message was sent is listed

II. The number of parts which the message contained is listed

III. Which message part is being sent is listed

IV. The length of the message part (not including *Buchstabenkenngruppe*) is listed

V. A *Grundstellung* rotor position is chosen and listed

VI. A *Spruchschlüssel* rotor position is chosen and encoded using the *Grundstellung*, this is listed

VII. The *Buchstabenkenngruppe* is listed

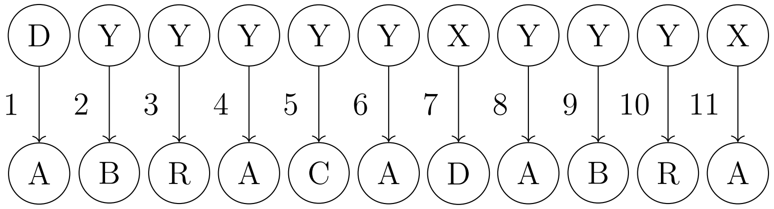
VIII. The message part encoded using the daily key and the *Spruchschlüssel* position is listed

It is clear that with this protocol, the Polish Bomba could no longer deduce the necessary details to decrypt enigma messages. All of the permutation information contained in the original key distribution protocol was removed and a new method needed to be derived for infering information about the daily key.

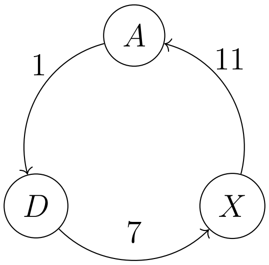
Section 2.3

Loops

The removal of the double encoded *Spruchschlüssel* does not mean that permutation information cannot be stored elsewhere in the message. For the sake of argument, let us say we knew that our encrypted message had plaintext encoding



Where the top row is the ciphertext and the bottom is the plaintext, further, the number in each mapping indicates how many steps away we are from the rotor positions when we began encoding the message.



Recall

$$E_1 = P^{-1}\theta_1 R_1^{-1}\theta_1^{-1}R_2^{-1}R_3^{-1}MR_3R_2\theta_1^{-1}R_1\theta_1 P$$

$$E_7 = P^{-1}\theta_7 R_1^{-1}\theta_7^{-1}R_2^{-1}R_3^{-1}MR_3R_2\theta_7^{-1}R_1\theta_7 P$$

$$E_{11} = P^{-1}\theta_{11} R_1^{-1}\theta_{11}^{-1}R_2^{-1}R_3^{-1}MR_3R_2\theta_{11}^{-1}R_1\theta_{11} P$$

Then it follows that our loop can be represented by

$$\sigma = E_{11} \circ E_7 \circ E_1$$

and we see that all the intermediate plugboard settings cancel out. Lets isolate the plugboard settings by letting $\overline{\sigma}$ represent σ without the use of the plugboard for input and output, then

$$\sigma = P^{-1}\overline{\sigma}P$$

We have that

$$\begin{aligned} &\sigma(A) = A \\ \iff &(P^{-1}\overline{\sigma}P)(A) = A \\ \iff &\overline{\sigma}(P(A)) = P(A) \end{aligned}$$

Suppose that our initial rotor position was correct, then certainly our $\overline{\sigma}$ is correct. We can make a hypothesis that A is steckered to K in the plugboard. Suppose we find that $\overline{\sigma}(K) \neq K$, then $\sigma(A) \neq A$ and our loop is broken, breaking our assumptions, thus A must not be steckered to K. But this will actually elimiate more hypotheses than just A being steckered to K. We know that $\overline{\sigma}(K)$ is some letter which is not K. So we continue with a new hypothesis that A is steckered to $\overline{\sigma}(K)$ and if we find $\overline{\sigma}(\overline{\sigma}(K)) \neq \overline{\sigma}(K)$, then we have further eliminated this possibility. Each new hypothesis suggests that A is steckered to $\overline{\sigma}^i(K)$ which will be shown to be false if $\overline{\sigma}^{i+1}(K) \neq \overline{\sigma}^i(K)$. What if we find that $\overline{\sigma}^{i+1}(K) = \overline{\sigma}^i(K)$ at some point? This cannot happen since

$$\begin{aligned} &\overline{\sigma}^{i+1}(K) = \overline{\sigma}^i(K) \\ \Rightarrow\ &\overline{\sigma}^{-i} \circ \overline{\sigma}^{i+1}(K) = \overline{\sigma}^{-i} \circ \overline{\sigma}^i(K) \\ \Rightarrow\ &\overline{\sigma}(K) = K \end{aligned}$$

which by supposition is false. Then we can continue in our hypotheses until we eventually reach a cycle where $\overline{\sigma}^i(K) = K$. Then we gather a set of impossible steckerings, that is

$$P(A) \notin \{\ \overline{\sigma}^i(K) \mid i \in \mathbb{N}\}$$

The notation we are using can be simplified significantly. The set $\{\ \overline{\sigma}^i(K) \mid i \in \mathbb{N}\}$ is equivalent to the orbit of K via the group action of the cyclic subgroup $\langle\overline{\sigma}\rangle$ which can be denoted $\langle\overline{\sigma}\rangle \cdot K$.

We then have several cases

(a) If $|\langle\overline{\sigma}\rangle \cdot K| = 26$, then A cannot be steckered to anything which is clearly impossible, thus our rotor position must be incorrect.

(b) If $|\langle\overline{\sigma}\rangle \cdot K| = 25$, then A can only be steckered to the remaining letter $\{A, \ldots, Z\} - \langle\overline{\sigma}\rangle \cdot K$

(c) If $|\langle\overline{\sigma}\rangle \cdot K| = 1$, in this case we must have intially had $\overline{\sigma}(K) = K$ so we have not eliminated any possibilities.