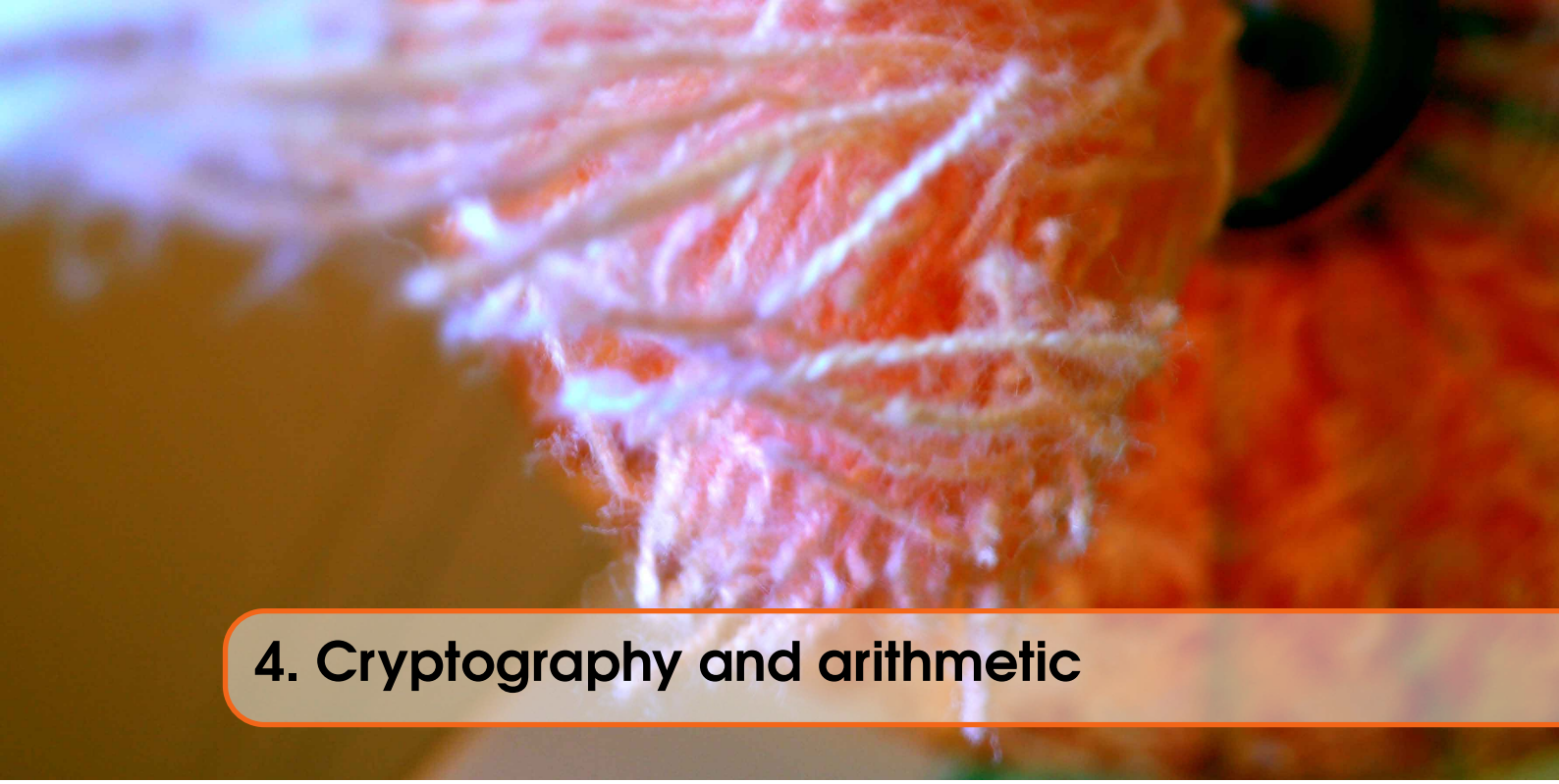

4. Cryptography and arithmetic

Suppose that we intercept the ciphertext

O H 7 F 8 6 B B 4 6 R 3 6 2 7 0 2 6 B B 9

and happen to know:

1. a 37-letter alphabet is used with integer correspondence:

$$0,1,\ldots,9,10 \leftrightarrow A, 11 \leftrightarrow 11,\ldots,35 \leftrightarrow Z, \_ \leftrightarrow 36$$

2. an *affine* cryptosysrem

$$f_E : \mathbb{Z}_{37} \to \mathbb{Z}_{37}, n \mapsto an + b$$

   is used with some unkown key $E = (a,b)$.
3. the unkown plaintext ends with the letters: 0 0 7

Let's 'break the code' by using this information to determine: i) the enciphering key; ii) and the deciphering key.

We know that $f_E(0) = B$ and $f_E(7) = 9$ or, using integers in place of letters,

$$f_E(0) \equiv 11 \mod 37 \tag{4.1}$$
$$0a + b \equiv 11 \mod 37 \tag{4.2}$$

and

$$7a + b \equiv 9 \mod 37\,. \tag{4.3}$$

From (4.2) we deduce $b \equiv 11$. Then from (4.3) we find:

$$7a \equiv -2 \mod 37 \tag{4.4}$$
$$a \equiv -2 \times 7^{-1} \mod 37 \tag{4.5}$$

To find 7^{-1} we can use the Euclidean algorithm

$$37 = 5.7 + 2 \tag{4.6}$$
$$7 = 3.2 + 1 \tag{4.7}$$
$$2 = 2.1 + 0 \tag{4.8}$$

and Bezout's Identity

$$1 = 7 - 3.2 \tag{4.9}$$
$$= 7 - 3(37 - 5.7) \tag{4.10}$$
$$= 16.7 - 3.37 \tag{4.11}$$

to deduce

$$7 \times 16 \equiv 1 \mod 37 \tag{4.12}$$

and finally:

$$7^{-1} \equiv 16 \mod 37 \tag{4.13}$$

Now from equation (4.5) we find:

$$a \equiv -2 \times 16 \equiv -16 \equiv 5 \mod 37 \tag{4.14}$$

The enicphering key is thus $E = (5, 11)$. Equations (3.13) and (3.14) then yield the deciphering key $D = (15, 12)$.

We could use the deciphering key to decipher the cipertext. We simply need to apply the function

$$f_d(n) \colon \mathbb{Z}_{37} \to \mathbb{Z}_{37}, n \mapsto 15n + 12 \tag{4.15}$$

to each letter of ciphertext. Applying f_D to the first letter of ciphertext

$$f_D(O) \leftrightarrow f_D(24) \tag{4.16}$$
$$\equiv 15 \times 24 + 12 \mod 37 \tag{4.17}$$
$$\equiv 27 + 20 \mod 37 \tag{4.18}$$
$$\equiv 10 \mod 37 \tag{4.19}$$
$$\leftrightarrow A \tag{4.20}$$

we find that the first letter of plaintext is 'A'. Subsequent letters of plaintext can be found in a similar manner.

4.1 A lesson to be learned

One lesson to learn from the above example is that when using an affine cryptosystem we should not always sign off with our name since this can help an adversary determine the enciphering key (which enables the adversary to send us 'fake news') and determine the deciphering key (from which the adversary can decipher our enrypted message).

More generally, and for the same reason, we should not send an encrypted message containing any portions whose corresponding plaintext can be identified by an adversary. This applies to many other cryptosystems too, including the Enigma machine. The German Navy had not learned

this lesson! During the battle of the North Atlantic, while lying in wait for convoys of Allied ships, U-boats were collecting weather data and sending it back home. The Allies recognized these patterns and used them to decipher important German Navy messages.

If we send any large amount of ciphertext using an affine cryptosystem then there will likely be one pattern that an adversary can easily spot. If the plaintext was written in, say, English using a 26-letter alphabet then the most frequent letter in the plaintext will most likely be 'E' and the second most frequent letter will likely by 'T'. The inventor of Morse code, Samuel Morse (1791–1872), needed to know frequecies of letters in English text and estimated that 12% of characters are 'E' and 9% are 'T'. If the ciphertext is long enough then the most frequent letter of ciphertext will be deciphered as 'E' and the second most frequent letter of ciphertect will be deciphered as 'T'. As in the above example, the adversary can solve two simultaneous equations to establish the enciphering and deciphering keys. This is one of several reasons why no affine enciphering function

$$f_E : \mathbb{Z}_N \to \mathbb{Z}_N, n \mapsto an + b \tag{4.21}$$

should ever be used in practice.