MT362/462/5462 Cipher Systems: Sheet 1

Attempt at least questions 1 to 5. Please staple your answers together and remember to write your name or student number. You will get 1.25% of your final mark for a reasonable attempt at this sheet.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in after the second lecture on Friday 11th October.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

The Mathematica notebook `AlphabeticCiphers` on Moodle can be used to encrypt and decrypt using substitution ciphers, and compute frequencies and the Index of Coincidence. Remember 'Evaluate Notebook' to get started.

1. Decrypt `BYIKVXRYVVYGKI`, assuming it is the ciphertext output by a Caesar cipher. What is the key?

2. In Example 1.2, Alice agreed to send Bob his exam mark $x \in \{0, 1, \ldots 99\}$ by encrypting it as the ciphertext $(x+k) \bmod 100$. Assume that the key $k \in \{0, 1, \ldots 99\}$ is known to only Alice and Bob and is chosen at random. Eve, the eavesdropper, learns all messages that Alice sends to Bob.

(a) When Alice sends Bob the ciphertext 17, what, if anything, can Eve learn about (i) the plaintext x?; (ii) the key k, supposing that Bob's exam mark is equally likely to be each number between 40 and 79?

(b) Suppose later Alice sends Bob her own exam mark $x' \in \{0, 1, \ldots, 99\}$ using the same cryptoscheme, *and using the same key* k. What can Eve learn now?

(c) Malcolm, the man-in-the-middle, can modify the ciphertext. Suppose he is confident that Bob's mark is between 40 and 79. Can he trick Bob into thinking he failed?

3. In the first week you formed a *block* of four people, identified as Alice, Bob, Alice$'$, Bob$'$. The *pairs* are $\{\text{Alice}, \text{Bob}\}$ and $\{\text{Alice}', \text{Bob}'\}$. You were then emailed a substitution cipher key. Each person in a pair has the same key.

(a) Write a plaintext message x of at least 75 words on a subject of your choice, and encrypt it using your substitution cipher key π. (Keep the spaces please!) Email the ciphertext $e_\pi(x)$ to all three people in your block.

(b) Decrypt the message from the other person in your pair. [*Hint:* do **not** use frequency analysis!]

(c) Using frequency analysis, decrypt either of the messages sent to you by a person not in your pair.

(d) Write up (c), explaining your method. An annotated printout is fine. Did you learn the entire key? If you only looked at one message, why might using both (but still decrypting only one) have been easier?

4. In a *chosen plaintext attack*, the attacker **chooses** a plaintext x, and is given the corresponding ciphertext $e_k(x)$ for the key k.

Explain how to find the key by a chosen plaintext account when the cipher is (a) a substitution cipher e_π; (b) a Vigenère cipher e_k where k has length exactly 10. Make it clear which plaintexts the attacker should choose.

5. The ciphertext below is the output of a Vigenère cipher. Each line has length 50.

```
01234567890123456789012345678901234567890123456789

WKMSDBPZPQYBGLLSDBTHCBLDNBAHLECQNBOTEOCRWOCOAXRDZT
MQZFLSDBAHLECQPBVSPEGREPMEPBLCQBRNPTMDMRYKSLPCOFLS
DBNKWFLSAURHJMMREQGNJPBHBCCQEKEAUXKTHQGOHBMEPECKAK
ESDLDSDBIDUFRHOHLNSKYRGXQHOHGRPBQS
```

(a) Find all positions in which `SDB` appear in the ciphertext.

(b) Compute the Index of Coincidence on the samples of size 20 (or larger if you prefer) obtained by taking every m-th position in the ciphertext starting with the first letter `W` in position 0, for each $m \in \{2, 3, 4\}$.

For example the sample for $m = 3$ of size 20 is `WSPQGSTBNHCBERCXZQLB`. To get these samples in MATHEMATICA, evaluate `AlphabeticCiphers.nb`; then `StringTake[SplitText[Q5Ciphertext, 3][[1]], {1, 20}]`.

(c) What do (a) and (b) suggest about the key length?

(d) Determine the key: start by guessing the plaintext corresponding to each `SDB`.

(e) Why is the Index of Coincidence least for $m = 3$ and in the middle for $m = 2$?

(The idea in (a) of finding the key length by comparing the positions containing a fixed substring of the ciphertext is known as the *Kasiski test*.)

6. Let R be defined on plaintexts by $R(x)_i = x_i + i \bmod 26$, numbering positions in tuples from 0. [**Original version numbered from 1: it makes little difference.**] For example $R(\texttt{bead}) = \texttt{BFCG}$ since $\texttt{bead} \longleftrightarrow (1, 4, 0, 3)$, $R\big((1, 4, 0, 3)\big) = (1, 5, 2, 6)$ and $(1, 5, 2, 6) \longleftrightarrow \texttt{BFCG}$. Similarly $R^2(\text{aaaa}) = \texttt{ACEG}$.

Let e_π denote the substitution cipher with key π.

Propose known ciphertext attacks on the two ciphers (a) $x \mapsto R^j\big(e_\pi(x)\big)$ and (b) $x \mapsto e_\pi\big(R(x)\big)$. In (a) the key is (π, j) for some $j \in \{0, \ldots, 25\}$; in (b) the key is simply π. Assume the plaintext is an English message of about 100 words.

Which cipher has more possible keys? Which appears harder to break?

7. Let y be every m-th position in a ciphertext output by the Vigenère cipher. What statistic would you compute to perform a χ^2-test with null hypothesis that the letters in y are distributed uniformly? How is this statistic related to the Index of Coincidence?

8. Which of the ciphertexts `XXXXX` and `VWXYZ` could be the output of (a) a substitution cipher, (b) a Vigenère cipher with key of length 3? Assume that the plaintext is a single English word. (The Vigenère key need not be an English word.)

MT362/462/5462 Cipher Systems: Sheet 2

Attempt at least questions 1 to 4. Question 3 has an optional part marked ($\star$). Question 5 is compulsory for **M.Sc.** students. Please staple your answers together and remember to write your name or student number. You will get 1.25% of your final mark for a reasonable attempt at this sheet.

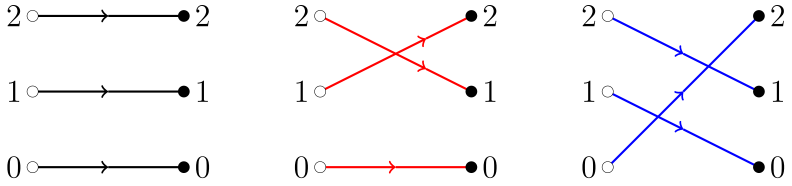
The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in after the second lecture on Friday 25th October. Note you have a fortnight to do this sheet.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

Throughout we use the notation of §3, so $\mathcal{K}$ is the keyspace, $\mathcal{P}$ the plaintexts and $\mathcal{C}$ the ciphertexts in a cryptosystem, with encryption functions $e_k : \mathcal{P} \to \mathcal{C}$ and decryption functions $d_k : \mathcal{C} \to \mathcal{P}$ indexed by keys $k \in \mathcal{K}$.

1. The cryptosystem shown below uses three keys from the affine cipher on $\mathbb{Z}_3$, each with probability $\frac{1}{3}$. Suppose that plaintext 1 is sent with probability p and plaintext 2 is sent with probability $1 - p$, so plaintext 0 is never sent.



(a) Recall that $e_{(a,c)}(x) = ax + c$. [**Sorry, this should be** $e_{(a,c)}(x) = ax + c$ **mod** 3.] Which keys (a, c) are used in this cryptosystem?

(b) Find $\mathbf{P}[Y = 1|X = 1]$. Express $\mathbf{P}[Y = 1]$, $\mathbf{P}[X = 1|Y = 1]$ in terms of p.

(c) When does the cryptosystem have perfect secrecy with respect to the probability distribution $p_0 = 0$, $p_1 = p$, $p_2 = 1 - p$ on plaintexts?

2. Alice and Bob communicate using the numeric one-time pad cryptosystem from Example 3.4, in which $\mathcal{K} = \mathcal{P} = \mathcal{C} = \{0, 1, \ldots, n-1\}$ and the encryption functions are defined by $e_k(x) = (x + k) \bmod n$. Each key $k \in \mathcal{K}$ is chosen with equal probability. Let p_x be the probability that $x \in \mathcal{P}$ is Alice's message.

(a) Show that if $x \in \mathcal{P}$ and $p_x > 0$ then $\mathbf{P}[Y_n = y|X_n = x] = \frac{1}{n}$ for all $y \in \mathcal{C}$.

(b) Find $\mathbf{P}[Y_n = y]$ for each $y \in \mathcal{C}$.

(c) Hence show that $\mathbf{P}[X_n = x|Y_n = y] = p_x$ for all $x \in \mathcal{P}$ with $p_x > 0$.

(d) What is $\mathbf{P}[X_n = x|Y_n = y]$ if $p_x = 0$? Deduce from this and (c) that the numeric one-time pad has perfect secrecy.

3. One of the encryption functions in the affine cipher (see Example 4.2) on $\mathbb{Z}_{13}$ is $e_{(2,5)}$, defined by $e_{(2,5)}(x) = 2x + 5$. Let $d_{(2,5)}$ be the decryption function, i.e. the inverse of $e_{(2,5)}$.

(a) What is $e_{(2,5)}(9)$?

(b) Find a formula for $d_{(2,5)}(y)$ and hence find $d_{(2,5)}(1)$.

(c) What encryption function is equal to $d_{(2,5)}$?

($\star$) Generalize (c) by proving that the encryption functions form a group.

4. (a) Is there a cryptosystem such that $|\mathcal{C}| < |\mathcal{P}|$?

(b) Is there a cryptosystem with perfect secrecy such that $|\mathcal{K}| < |\mathcal{C}|$?

(c) A student writes: 'since the encryption functions e_k are injective, if $k \neq k'$ then $e_k(x) \neq e_{k'}(x)$'. Is this correct? Justify your answer with a proof or counterexample, as appropriate.

(d) Give at least three different examples of cryptosystems with perfect secrecy such that $|\mathcal{P}| = |\mathcal{K}| = |\mathcal{C}| = 4$. [*Hint:* Latin squares.]

5. **(M.Sc.)** Work with the Shamir secret sharing scheme over $\mathbb{F}_{11}$ with 5 people and threshold 3 using evaluation points $c_i = i$ for $i \in \{1, 2, 3, 4, 5\}$.

(a) Find the shares for the secret $5 \in \mathbb{F}_{11}$, choosing a polynomial at random.

(b) Alice (Person 1), Bob (Person 2) and Charlie (Person 3) have the shares 7, 5, 3 respectively. The three agree to meet, simultaneously reveal their shares, and together compute the secret.

(i) What is the secret?

(ii) Show, by giving an explicit example, that if Alice lies about her share to Bob and Charlie, then she can both learn the secret and leave Bob and Charlie knowing an incorrect secret.

(iii) Suggest a way to avoid some of the problems in (ii).

6. This question proves a converse result to Shannon's Theorem (Theorem 3.10). Suppose that $|\mathcal{P}| = |\mathcal{C}| = |\mathcal{K}|$. Show that if each key is used with equal probability and for all $x \in \mathcal{P}$ and $y \in \mathcal{C}$ there is a unique key k such $e_k(x) = y$, then

(a) $\mathbf{P}[Y = y] > 0$ for all $y \in \mathcal{C}$;

(b) the cryptosystem has perfect secrecy.

7. In Theorem 3.10 we assumed that the cryptosystem was practical.

(a) Show that if the hypothesis 'for all $y \in \mathcal{C}$ there exists $x \in \mathcal{P}$ and $k \in \mathcal{K}$ such that $e_k(x) = y$' is dropped then conclusions (a), (b) and (c) of Theorem 3.10 may fail to hold.

(b) Show that if the hypothesis '$\mathbf{P}[K = k] > 0$ for each $k \in \mathcal{K}$' is dropped then again (a), (b) and (c) may fail to hold.

MT362/462/5462 Cipher Systems: Sheet 3

Attempt at least questions 1 to 5. Question 6 is compulsory for **M.Sc.** students. Please staple your answers together and remember to write your name or student number. You will get 1.25% of your final mark for a reasonable attempt at this sheet.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in after the second lecture on Friday 1st November.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

1. Consider the affine cipher (see Example 4.2) with $p = 151$.

(a) Decrypt the ciphertext 138 sent using the key $(12, 10)$.

(b) In a known plaintext and ciphertext Mark learns that $e_{(a,c)}(21) = 18$. Find all the possibilities for the key (a, c). Suppose that later he learns that $e_{(a,c)}(18) = 21$. What is the key?

2. Let q be prime. Suppose that Alice and Bob communicate using the affine cipher on $\mathbb{Z}_q$ with keyspace $\mathcal{K} = \{(a, c) : a, c \in \mathbb{Z}_q, a \neq 0\}$, and that Alice's plaintext is $x \in \mathbb{Z}_q$ with probability p_x.

(a) What is the size $|\mathcal{K}|$ of the keyspace?

(b) Show that for each x, $y \in \mathbb{Z}_q$ there are exactly $q-1$ keys k such that $e_k(x) = y$.

(c) Show that if each key is equally probable then the cryptosystem has perfect secrecy. Can Eve learn anything about the plaintext from a known ciphertext?

(d) Show that the key can be determined by a chosen plaintext attack using two plaintexts. Does this contradict perfect secrecy? Does a single plaintext suffice?

(e) Can Malcolm, the man-in-the-middle, modify a ciphertext without Bob noticing? How might this problem be reduced? [*Hint:* change $\mathcal{P}$ to a subset of $\mathbb{Z}_q$.]

3. Show that if K and X are independent random variables, taking values in sets $\mathcal{K}$ and $\mathcal{P}$ respectively, then

$$H(K, X) = -\sum_{k\in\mathcal{K}} \sum_{x\in\mathcal{P}} \mathbb{P}[K = k]\mathbb{P}[X = x](\log_2 \mathbb{P}[K = k] + \log_2 \mathbb{P}[X = x]).$$

Deduce that $H(K, X) = H(K) + H(X)$. [*Hint:* please explain your steps, taking care to use sigma notation correctly. The joint entropy $H(K, X)$ is defined in Definition 5.6.]

4. Eve intercepts the three ciphertexts `ymdg`, `smrf`, `xmom` encrypted using the same key by a one-time pad. Find all three plaintexts and the key.

[*Hint:* the code used in the lecture is online at `repl.it/@mwildon/OneTimePad2`. (Try Google Chrome if it doesn't work in your first choice of browser.) If you do it by hand, you might first guess a likely theme for the words.]

5. Let $\mathcal{A} = \{\text{a}, \ldots, \text{z}\}$, so $\mathcal{A}^n$ is all lowercase strings of length n. Assume that plaintexts are English messages in lowercase, with spaces deleted.

(a) Estimate the unicity distance (see Definition 5.14) of the Vigenère cipher using keys of length 10, chosen with equal probability from $\mathcal{A}^{10}$.

(b) Given bijections $\pi, \sigma : \mathcal{A} \to \mathcal{A}$ define $e_{(\pi,\sigma)} : \{\mathtt{a}, \ldots, \mathtt{z}\}^n \to \{\mathtt{a}, \ldots, \mathtt{z}\}^n$ by

$$e_{(\pi,\sigma)}(x)_i = \begin{cases} \pi(x_i) & \text{if } i \text{ is even} \\ \sigma(x_i) & \text{if } i \text{ is odd.} \end{cases}$$

For example, if $n = 4$ then $e_{(\pi,\sigma)}(x_0, x_1, x_2, x_3) = \bigl(\pi(x_0), \sigma(x_1), \pi(x_2), \sigma(x_3)\bigr)$.

(i) Estimate the unicity distance of the cryptosystem with keys all $e_{(\pi,\sigma)}$, supposing that keys are chosen with equal probability.

(ii) Propose a *chosen ciphertext* attack on this cryptosystem. [*Hint:* see the definition on page 19 of the printed lecture notes.]

6. **(MSc.)** Recall from the Preliminary Problem Sheet that $x_{\ell-1} \ldots x_1 x_0$ is the binary form of $2^{\ell-1}x_{\ell-1} + \cdots + 2x_1 + x_0$.

For $j \in \{0, 1, \ldots, \ell - 1\}$, let $f_j : \mathbb{F}_2^\ell \to \mathbb{F}_2$ be the Boolean function defined so that $f(x_{\ell-1}, \ldots, x_1, x_0)$ is the bit in position j of $x_{\ell-1} \ldots x_1 x_0 + 5 \bmod 16$.

For example, taking $\ell = 3$, since $6 = 0110_2$ $6 + 5 = 11$ and $11 = 1011_2$, we have $f_3(0110) = 1$, $f_2(0110) = 0$, $f_1(0110) = 1$ and $f_0(0110) = 1$.

Express f_0, f_1, f_2, f_3 as polynomials in x_3, x_2, x_1, x_0. What is the coefficient of the monomial $x_0 x_1 x_2$ in f_3?

For general j, what is the monomial with the maximum number of variables in f_j?

7. Show that in a cryptosystem with perfect secrecy $H(X|Y) = H(X)$, where as usual X is the plaintext and Y is the ciphertext.

8. Let X and Y be random variables taking values in sets $\mathcal{X}$ and $\mathcal{Y}$ respectively. Let $f : \mathcal{X} \to \mathcal{X}$ be a function. Prove the inequality $H\bigl(f(X)\,|\,Y\bigr) \leq H(X|Y)$ used in the 'extras' for Part A. [*Hint:* one proof uses the chaining rule, Lemma 5.8.]

9. Eve observes the ciphertexts `jaekbwwoswoppljoeow` and `eszxzagrhaofvquwkhj` encrypted using the same one-time pad. Decrypt them both and find the key.

MT362/462/5462 Cipher Systems: Sheet 4

Attempt at least questions 1 to 3 and 4(a), (b), (c), (d). Questions 4(e) and 5 are compulsory for **M.Sc.** students. Please staple your answers together and remember to write your name or student number.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in at the second lecture on Friday 8th November.

It is helpful if you indicate questions you did but are uncertain about, or would like to seen done in lectures.

The MATHEMATICA notebook used in lectures to find keystreams is available from Moodle. By definition, the LFSR of width ℓ with taps T, where $T \subseteq \{1, 2, \ldots, \ell\}$, has keystream $k_0k_1k_2\ldots$ such that $k_s = \sum_{t\in T} k_{s-t}$ for all $s \geq \ell$.

1. By Definition 5.8, the *period* of a keystream is its length until its first repeat. For instance $001100110011\ldots$ has period 4. Let G be the LFSR of width 5 with taps $\{4, 5\}$.

(a) (i) Let $k = 00001$. Calculate the keystream $k_0, k_1, k_2, \ldots$, defined by G for k. What is the period of this keystream?

(ii) Find s such that $(k_s, k_{s+1}, k_{s+2}, k_{s+3}, k_{s+4}) = 00100$.

(iii) How would your answer to (i) change if the key was 00100?

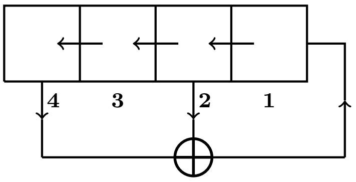
(b) Find a key k' such that the keystream defined by G for k' has period 7.

(c) Find all the periods of keystreams for G. What is the sum of the periods?

(d) By Definition 5.8, the *period* of G is the least m such that $G^m = \mathrm{id}$, the identity function. What is the period of G?

[You can use your answer to (a) in Question 3.]

2. Let F be the LFSR of width 4 with taps $\{2, 4\}$, as shown in the circuit diagram below; the numbers correspond to the possible taps.



(a) Solve the equation $F\big((x_0, x_1, x_2, x_3)\big) = (y_0, y_1, y_2, y_3)$ and hence find a formula for F^{-1}.

(b) Draw a circuit diagram for F^{-1}. Is F^{-1} an LFSR?

3. Let F be an LFSR of width ℓ with taps T, so by definition

$$F\big((x_0, x_1, \ldots, x_{\ell-2}, x_{\ell-1})\big) = (x_1, x_2, \ldots, x_{\ell-1}, \sum_{t \in T} x_{\ell-t}).$$

(a) Show that if F is invertible then $\ell \in T$. [*Hint:* use the contrapositive.]

(b) Show conversely that if $\ell \in T$ then F is invertible and give a formula for F^{-1}.

4. As in Question 1, let G be the LFSR of width 5 with taps $\{4, 5\}$. Let $k_0, k_1, k_2 \ldots$ be the keystream for 00100. The corresponding power series $K(z) = k_0 + k_1 z + k_2 z^2 + \cdots$ begins $z^2 + z^6 + z^7 + z^{10} + \cdots$.

(a) Compute $K(z)$ up to the term z^{30}. [*Hint: use Question 1(a).*]

(b) Find $p \in \mathbb{N}$ such that $K(z)(1 + z^p)$ is a polynomial.

(c) Show that $K(z)(1 + z^4 + z^5)$ is a polynomial.

(d) Find a polynomial $r(z)$ of the form $1 + r_1 z + \cdots + r_6 z^6 + z^7$ such that $K(z)r(z)$ is a polynomial. Hence find an LFSR of width 7 having $k_0 k_1 k_2 \ldots$ as a keystream.

(e) ($\star$) Is there an LFSR of width at most 4 having $k_0 k_1 k_2 \ldots$ as a keystream?

5. (**M.Sc.**) Let $\boxtimes$ denote normal multiplication, performed on numbers written in binary. For example $(1, 0, 1, 0) \boxtimes (0, 1, 0, 1) = (0, 0, 1, 1, 0, 0, 1, 0)$ since $1010_2 = 10$ and $0101_2 = 5$ and $10 \times 5 = 50 = 00110010$. Define $F : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ by

$$F\big((x_7, x_6, x_5, x_4, x_3, x_2, x_1, x_0)\big) = (x_7, x_6, x_5, x_4) \boxtimes (x_3, x_2, x_1, x_0).$$

For $j \in \{7, 6, \ldots, 0\}$, define a Boolean function $f_j : \mathbb{F}_2^8 \to \mathbb{F}_2$ by $f_j(x) = F(x)_j$, numbering positions from the right. Thus $f_0(x) = x_4 x_0$ and $f_1(x) = x_4 x_1 + x_5 x_0$.

(a) Write down the truth table for f_1, with columns labelled x_5, x_4, x_1, x_0.

(b) Express f_1 in disjunctive normal form.

(c) Express f_2 in algebraic normal form.

6. Let F be an LFSR of width ℓ with taps $T \subseteq \{1, \ldots, \ell\}$ and let G be an LFSR of width ℓ with taps $U \subseteq \{1, \ldots, \ell\}$. Show that the function $H : \mathbb{F}_2^\ell \to \mathbb{F}_2^\ell$ defined by

$$H\big((x_0, x_1, \ldots, x_{\ell-1})\big) = \big(x_1, \ldots, x_{\ell-1}, \sum_{t \in T} x_{\ell-t} + \sum_{u \in U} x_{\ell-u}\big)$$

is an LFSR of width ℓ and determine the taps of H in terms of T and U.

7. A *de Bruijn sequence* of *order* ℓ is a cyclic sequence containing every element of $\mathbb{F}_2^\ell$ exactly once. Thus 00010111 is a de Bruijn sequence of order 3; for instance, to find 110, take the final two 1s and the initial 0.

(a) Use the LFSR in Example 5.2 to construct a de Bruijn sequence of order 4.

(b) Prove that there exist de Bruijn sequences of every order. (You may assume there exists an LFSR of period $2^\ell - 1$ for every $\ell \in \mathbb{N}$.)

MT362/462/5462 Cipher Systems: Sheet 5

Attempt questions 1 to 5. M.Sc. students should also do question 6. Please staple your answers together and remember to write your name or student number.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in at the lecture on Monday 18th November. Note you have a extra weekend for this sheet.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

The MATHEMATICA notebook `LFSRs.nb` used in lectures to find keystreams is on Moodle. By definition, the LFSR of width ℓ with taps T, where $T \subseteq \{1, \ldots, \ell\}$, has keystream $k_0k_1k_2\ldots$ such that $k_s = \sum_{t \in T} k_{s-t}$ for all $s \geq \ell$.

1. Let F be the LFSR of width 5 with taps $\{3, 5\}$.

(a) Show that the keystream for 00001 starts 00001 00101 and continue it until it repeats. [You can check your answer using `Keystream[{3,5},{0,0,0,0,1}]` in MATHEMATICA.]

(b) What is the period of this keystream?

(c) What is the period of F?

2. In 8-bit ASCII, '`a`' is encoded as the binary form of 97, namely 01100001, '`b`' as the binary form of 98, namely 01100010, and so on.

Fix $n \in \mathbb{N}$ and consider the cryptosystem with plaintexts $\mathcal{P} = \{\texttt{a}, \ldots, \texttt{z}\}^n$ and ciphertexts $\mathcal{C} = \mathbb{F}_2^{8n}$, in which a message of n characters is first converted to 8-bit ASCII, and then encrypted using the cryptosystem defined in Definition 6.3 with the LFSR F of width 5 with taps $\{3, 5\}$ seen in Question 1.

Your key is the first 5 bits of the binary key in your email from the lecturer.

(a) Let $k_0k_1k_2\ldots$ be the keystream for your key. Show that $k_{32m} = k_m$ for each $m \in \mathbb{N}_0$.

(b) Encrypt a message (lower-case, no spaces) of at least 25 characters. Send the sequence of bits to everyone in your block.

[*Hint:* to do this in MATHEMATICA, after loading and evaluating `LFSRs.nb` use `EncryptString[{3, 5}, {k0, k1, k2, k3, k4}, "message"]` You can also use `CharacterToASCII["x"]` to get the 8 bits for `x`, and so on.

(c) Decrypt the message from your partner.

(d) Decrypt either of the messages from the other two people in your block. [*Hint:* start by looking at bits 0 and 32 in the ciphertext. If you do not have a ciphertext to decrypt, use the one in the MATHEMATICA notebook.]

(e) What is the minimum length of ciphertext needed to determine the key?

3. (a) Let $k_0k_1k_2k_3k_4k_5k_6k_7$ be the keystream of an LFSR of width 4. (The taps could be anything.) Show that the the matrix equation

$$\begin{pmatrix} k_0 & k_1 & k_2 & k_3 \\ k_1 & k_2 & k_3 & k_4 \\ k_2 & k_3 & k_4 & k_5 \\ k_3 & k_4 & k_5 & k_6 \end{pmatrix} \begin{pmatrix} b_4 \\ b_3 \\ b_2 \\ b_1 \end{pmatrix} = \begin{pmatrix} k_4 \\ k_5 \\ k_6 \\ k_7 \end{pmatrix}$$

has a solution b_4, b_3, b_2, b_1. [*Hint:* remember that if T is the taps then $k_s = \sum_{t \in T} k_{s-t}$ for each $s \geq \ell$. Relate this to the four equations from the matrix.]

(b) Is the converse to (a) true? Justify your answer.

(c) Which of the bit sequences 00100110, 00100111, 11100001 and 0110111 is a keystream of an LFSR of width 4? (In the last you are only given $k_0k_1 \ldots k_6$.) Justify your answers. Do they change if the LFSR is required to be invertible?

4. Let F be the LFSR with taps $\{3,4\}$ and width 4 and let F' be the LFSR with taps $\{2,3\}$ and width 3. Let $k_0k_1k_2k_3k_4 \ldots = 00010\ldots$ and $k'_0k'_1k'_2k'_3k'_4 \ldots = 00101\ldots$ be the keystreams for 0001 and 001, respectively. Let $u_s = k_s + k'_s$ for each $s \in \mathbb{N}$.

(a) Show that $u_0u_1u_2u_3u_4 = 00111$ and find $u_0u_1u_2 \ldots u_{19}$.

(b) What is the period of $u_0u_1u_2\ldots$?

(c) Find an LFSR of width 7 that has $u_0u_1u_2\ldots$ as a keystream. [*Hint: the method from Question 3 works, but there is a better way using annihilators.*]

5. Let $B_0, B_1, \ldots, B_{n-1}$ be a sequence of bits, each 0 or 1 independently with probability $\frac{1}{2}$. For $b, b' \in \{0,1\}$, let $M_{bb'}$ be the number of $i \in \{0, \ldots, n-2\}$ such that $(B_i, B_{i+1}) = (b, b')$.

(a) Show that the expected value of M_{00} is $\mathbb{E}[M_{00}] = (n-1)/4$ and find $\mathbb{E}[M_{01}], \mathbb{E}[M_{10}], \mathbb{E}[M_{11}]$.

(b) Does the sequence below pass the monobit test in Exercise 7.4?

$$(0,1,0,1,1,0,0,1,0,1,0,1,0,1,0,1,0,0,1,1,0,0,1,1,0,0,1,0,1,0,1,1,0)$$

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2

What is n and what are the statistics M_{00}, M_{10}, M_{01}, M_{11} for this sequence?

(c) Perform a χ^2-test on $M_{00}, M_{01}, M_{10}, M_{11}$ to test the sequence in (b) for randomness on pairs of bits. [*Hint:* use $M_{00} + M_{01} + M_{10} + M_{11} = n$ to determine the degrees of freedom.]

6. (a) **(M.Sc.)** Let $f : \mathbb{F}_2^3 \to \mathbb{F}_2$ be defined by $f(x_0, x_1, x_2) = x_1x_2$. Find all the correlations $\mathrm{corr}(f, L_T)$ for $T \subseteq \{0,1,2\}$ and hence check Theorem 4.6(c) that

$$(-1)^f = \sum_{T \subseteq \{0,1,2\}} \mathrm{corr}(f, L_T)(-1)^{L_T}$$

(b) Let $S \triangle T = \{u \in S \cup T : u \notin S \cap T\}$. Show that if f is an n-variable Boolean function then $\mathrm{corr}(f + L_S, L_T) = \mathrm{corr}(f, L_{S \triangle T})$.

(c) Let $g(x_0, x_1, x_2) = x_0 + x_1x_2$. Express $(-1)^g$ in the form in (a). [*Hint:* use (b) and Theorem 4.6(c).]

MT362/462/5462 Cipher Systems: Sheet 6

Attempt at least questions 1 and 2 and either question 4 or question 5. ($\star$) denotes an optional part. Question 3 is compulsory for **M.Sc.** students. Please staple your answers together and put your name and student number on this sheet.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 10am, Thursday noon, or by appointment.

To be handed in at the lecture on Monday 25th November.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

The Mathematica notebook `LFSRs.nb` used in lectures to find keystreams is on Moodle. By definition, the LFSR of width ℓ with taps T, where $T \subseteq \{1, \ldots, \ell\}$, has keystream $k_0k_1k_2\ldots$ such that $k_s = \sum_{t \in T} k_{s-t}$ for all $s \geq \ell$.

1. Let $(k_0, k_1, k_2, \ldots)$ be a keystream of the LFSR F of width 2 with taps $\{1, 2\}$. Let $(k'_0, k'_1, k'_2, \ldots)$ be a keystream of an LFSR G of width 3 with unknown taps. The keystreams are multiplied to give $(k_0k'_0, k_1k'_1, k_2k'_2, \ldots)$. Suppose you know the product begins 101100000101.

(a) Explain why the keystreams of F and G have the form $1{\star}11{\star}{\star}{\star}{\star}{\star}1{\star}1$, where $\star$ denotes an unknown bit. By considering the possible keystreams produced by F, deduce the key for F.

(b) By considering the keystream for F explain why the keystream of G is of the form $1{\star}11{\star}00{\star}01{\star}1$. Hence find a possible set of taps and the unique key for G.

(c) Are the taps you found in (b) unique?

2. Working with polynomials with coefficients in $\mathbb{F}_2$, one can show (for instance using `Factor[X^511+1,Modulus->2]` and the same replacing `511` with `73` in Mathematica) that

- $1 + X^4 + X^9$ divides $X^{511} + 1$
- $1 + X^4 + X^9$ does not divide $X^{73} + 1$, but $1 + X + X^9$ does.

Given this, determine the periods of the LFSRs of width 9 with taps $\{4, 9\}$ and $\{1, 9\}$. [*Hint:* factorize 511 and use Corollary 6.12 and Lemma 6.13.] ($\star$) Find the periods of all the keystreams of each LFSR.

3. (**M.Sc.**) The table below shows the first 14 steps in the Berlekamp–Massey algorithm applied to the sequence

$$(u_0, u_1, \ldots, u_{14}) = (1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0)$$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ℓ_n	1	1	2	2	3	3	3	3	3	7	7	7	7	7	7
T_n	$\varnothing$	$\varnothing$	$\{2\}$	$\{1,2\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\star$	$\{1,5,6,7\}$	$\star$	$\star$	$\star$	$\star$
m	0	0	2	2	4	4	4	4	4	9	9	9	$\star$	9	

For instance, the LFSR F_9 has length $\ell_9 = 3$ and taps $T_9 = \{1\}$. Performing step 9 of the algorithm using $m = 4$ gives the LFSR F_{10} of length $\ell_{10} = 7$ and taps T_{10} that you are asked to find in (i). Since the length goes up, m is updated to 9. You should find that the final LFSR has taps $\{1, 2, 3, 4, 6, 7\}$.

(i) Verify that case (a) applies for steps $5, 6, 7, 8$ and perform step 9 to obtain the entry marked $\star$ in the column for $n = 10$.

(ii) Find the five remaining entries marked $\star$.

(iii) Given that the entire sequence $u_0, u_1, u_2, \ldots$ is generated by an LFSR of width 7, will the taps change in further steps of the Berlekamp–Massey algorithm? Justify your answer.

4. Let $k_0k_1k_2\ldots$ and $k'_0k'_1k'_2\ldots$ be keystreams of LFSRs with taps S and T and widths ℓ and ℓ', respectively. Let $u_s = k_s + k'_s$ for $s \in \mathbb{N}_0$.

(a) Generalizing Example 8.2 [**sorry, misprinted as 7.2**] and Question 4(c) on Sheet 5, show that $u_0u_1u_2\ldots$ is a keystream of an LFSR of width $\ell + \ell'$.

(b) Give an example where $u_0u_1u_2\ldots$ is also the keystream of an LFSR of strictly smaller width.

(c) ($\star$) Define the taps of the LFSR you found in (a) in terms of S and T. [*Hint:* a concise form uses the symmetric difference of sets S and T, defined by $S \triangle T = \{u \in S \cup T : u \notin S \cap T\}$.]

5. The 2-quadratic stream cipher was defined in Example 8.5. Recall that F is the LFSR of width 5 with taps $\{3, 5\}$ and F' is the LFSR of width 6 with taps $\{2, 3, 5, 6\}$. Given keys $k \in \mathbb{F}_2^5$ and $k' \in \mathbb{F}_2^6$, the keystream $u_0u_1u_2\ldots$ is defined by $u_0 = 0$ and $u_s = k_sk'_s + k_{s-1}k'_{s-1}$ for each $s \in \mathbb{N}$.

Using the attack in this example, the attacker guesses that k is $v_0v_1v_2v_3v_4$ and computes the correlation between the keystream $v_0v_1\ldots v_{1023}$ and $u_0u_1\ldots u_{1023}$. (Here $u_0u_1\ldots u_{1023}$ is obtained via a chosen plaintext attack, as in Exercise 7.1.)

The table below shows the four guessed keys $v_0v_1v_2v_3v_4$ with the highest correlations for several different k and k'. In each case the correlations for the other 32 guessed keys are close to 0.

k	k'	guessed key, correlation
00001	000001	00000, 0.223; 00001, 0.242; 10000, 0.230; 10001, 0.203
00001	000011	00000, 0.230; 00001, 0.215; 10000, 0.219; 10001, 0.211
00111	000001	00000, 0.238; 00111, 0.199; 10011, 0.199; 10100, 0.254
00111	000011	00000, 0.199; 00111, 0.219; 10011, 0.234; 10100, 0.254

Explain why in each case there are three 'fake keys', with correlation about $\frac{1}{4}$, as well as the correct key $k_0k_1k_2k_3k_4$. Predict the three fake keys when $k = 01000$ and k' is unknown.

[*Hint:* for $\frac{1}{4}$ of the positions in the F' keystream, $k'_s = 0$ and $k'_{s-1} = 1$ and so $u_s = k_{s-1}$. What keystream for F should $u_0u_1\ldots u_{1023}$ then be compared with? This should give you one 'fake' key.]

MT362/462/5462 Cipher Systems: Sheet 7

Attempt at least questions 1 to 4. Question 4(c) is optional. M.Sc. students should also attempt question 5. Please staple your answers together and put your name and student number on this sheet.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 11am, Thursday 11.30am, or by appointment.

To be handed in at the Monday lecture on 2nd December.

It is helpful if you indicate questions you did but are uncertain about, or would like seen done in lectures.

1. Let F be the Feistel Network for the function $f : \mathbb{F}_2^m \to \mathbb{F}_2^m$ so, by definition, $F\big((v,w)\big) = (w, v + f(w))$ for $(v,w) \in \mathbb{F}_2^{2m}$.

Show that if $(v', w') = F\big((v,w)\big)$ is the encryption of (v,w) then $(w,v) = F\big((w',v')\big)$ is the encryption of (w',v').

2. Consider the Q-block cipher as defined in Example 9.5, consisting of three rounds of the Feistel Network

$$F\big((v,w)\big) = (w, v + S(w + k^{(i)}))$$

where $v, w \in \mathbb{F}_2^4$ and $S\big(x_0,x_1,x_2,x_3)\big) = (x_2, x_3, x_0 + x_1x_2, x_1 + x_2x_3)$. Given a key $k \in \mathbb{F}_2^{12}$, the three round keys are $k^{(1)} = (k_0,k_1,k_2,k_3)$, $k^{(2)} = (k_4,k_5,k_6,k_7)$ and $k^{(3)} = (k_8,k_9,k_{10},k_{11})$ each in $\mathbb{F}_2^4$.

For readability below we write $v_0v_1v_2v_3\, w_0w_1w_2w_3$ for $(v,w) \in \mathbb{F}_2^8$ and similarly $k_0k_1k_2k_3\, k_4k_5k_6k_7\, k_8k_9k_{10}k_{11}$ for a key $k \in \mathbb{F}_2^{12}$.

(a) Encrypt $0000\,0000 \in \mathbb{F}_2^8$ using the key $0011\,0011\,0011$.

(b) Decrypt the ciphertext $0111\,0111$ using the key in (a)

(c) Find a key $k \in \mathbb{F}_2^{12}$ such that $e_k(0001\,0001) = 0000\,0000$.

(d) Show that given $(v,w) \in \mathbb{F}_2^8$ and $w' \in \mathbb{F}_2^4$ there is a unique round key $k_{\text{round}} \in \mathbb{F}_2^4$ such that $(w, v + S(w + k_{\text{round}})) = (w, w')$.

(e) How many keys $k \in \mathbb{F}_2^{12}$ have the property in (c)?

(f) Would your answer to (e) change if $0001\,0001$ and $0000\,0000$ were replaced with different plaintexts and ciphertexts?

3. You have a black box implementing an encryption round of a Feistel block cipher with block size $2m$. Thus, given $(v,w) \in \mathbb{F}_2^{2m}$ and a round key k_{round} both of your choice, the box will output $\big(w, v + S(w + k_{\text{round}})\big)$. You do not know the function S.

Explain how to use the box to decrypt a ciphertext $(v', w') \in \mathbb{F}_2^{2m}$ encrypted by applying the black box over 3 rounds with round keys $k^{(1)}, k^{(2)}, k^{(3)}$. [*Hint:* Question 1 is relevant.]

4. 3DES is the block cipher of block size 64 and keyspace $\mathbb{F}_2^{56} \times \mathbb{F}_2^{56} \times \mathbb{F}_2^{56}$ with encryption functions defined by

$$e_{(k,k',k'')}(x) = e_{k''}\bigl(d_{k'}\bigl(e_k(x)\bigr)\bigr)$$

where e_k and d_k are the encryption and decryption functions for DES.

(a) Show that there is a meet-in-the-middle attack using multiple chosen plaintexts that finds the key using about 2^{112} encryptions/decryptions.

(b) Assume no attack better than (a) exists. Is 3DES secure?

(c) ($\star$) Suggest why the middle map is decryption rather than encryption.

5. (**M.Sc.**) The LFSR of width 4 with taps $\{1, 4\}$ [**Corrected**] has keystream 10001111010110**0** with period 15.

(a) What is the minimum possible width of an LFSR with keystream 10001111010110**1**? [*Hint:* use the theoretical results in §5 of the M.Sc. course.]

(b) Find an LFSR generating the keystream in (a) by applying the Berlekamp–Massey algorithm.

6. The *affine block cipher* of block size n has keyspace all pairs (A, b), where A is an invertible $n \times n$ matrix with entries in $\mathbb{F}_2$ and $b \in \mathbb{F}_2^n$. The encryption functions $e_{(A,b)} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ are defined by

$$e_{(A,b)}(x) = xA + b.$$

(a) Define the decryption functions $d_{(A,b)} : \mathbb{F}_2^n \to \mathbb{F}_2^n$.

(b) How can the key be recovered in a chosen plaintext attack? How many plaintext / ciphertext pairs are required?

(c) Does repeating the cipher, as in the example of 2DES, so a plaintext is first encrypted with one key (A, b), and then another key (A', b'), make this cipher any more secure?

(d) Does this cipher have the ‘confusion’ property?

(e) Does this cipher have the ‘diffusion’ property?

MT362/462/5462 Cipher Systems: Sheet 8

Attempt at least questions 1 to 4. M.Sc. students should also attempt question 5. Please staple your answers together and put your name and student number on this sheet.

The lecturer will be happy to discuss any of the questions in drop-in sessions: Tuesday 3.30pm, Wednesday 10am, Thursday 11am, or by appointment.

To be handed in at the lecture on Monday 9th December. [Corrected deadline.]

Your feedback to the lecturer: what question, if any, do you most want solved in lectures? What was easy, hard, interesting this week?

1. Let $S : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^4$ be the S-box in the Q-block cipher, defined by $S\big((x_0, x_1, x_2, x_3)\big) = (x_2, x_3, x_0 + x_1x_2, x_1 + x_2x_3)$. Recall from Example 9.5 that the Feistel network in round i of this cipher is

$$(v^{(i-1)}, v^{(i)}) \mapsto \big(v^{(i)}, v^{(i-1)} + S(v^{(i)} + k^{(i)})\big)$$

where $k^{(i)} \in \mathbb{F}_2^4$ is the round key.

(a) Let $\Delta \in \mathbb{F}_2^4$. Show that if $\Delta_2 = 0$, i.e. Δ is of the form $(\star, \star, 0, \star)$ then

$$S(x + \Delta) + S(x) = \begin{cases} (0, \Delta_3, \Delta_0, \Delta_1) & \text{if } x_2 = 0 \\ (0, \Delta_3, \Delta_0 + \Delta_1, \Delta_1 + \Delta_3) & \text{if } x_2 = 1. \end{cases}$$

(b) Deduce Lemma 10.1(i), that $S(x + 1000) = S(x) + 0010$ for all $x \in \mathbb{F}_2^4$.

(c) Find all possibilities for $S(x + 0010) + S(x)$ where $x \in \mathbb{F}_2^4$.

(d) Let $\Gamma = 0000\,1000$. Let $(v, w) \in \mathbb{F}_2^8$ be chosen uniformly at random. Let (v', w') and (v'_Γ, w'_Γ) be the encryptions of (v, w) and $(v, w) + \Gamma$, respectively **over the first two rounds** of the Q-block cipher.

Show that no matter what the key is, $(v', w') + (v'_\Gamma, w'_\Gamma)$ is equally likely to be each of the four differences $\{0010\,0000, 1010\,0001, 1010\,0010, 0010\,0011\}$. [**Corrected: bit in position** 4 **should be** 0**, as changed above, not** 1.]

(e) Suggest a subexhaustive attack on the Q-block cipher in which the attacker first guesses $k^{(3)}$, and then $(k^{(1)}, k^{(2)})$.

2. Let e_k for $k \in \mathbb{F}_2^{12}$ be the encryption maps in the Q-block cipher. Show using Lemma 10.1(i) that $e_k(x) = e_{k+1000\,0010\,1000}(x)$ for all $x \in \mathbb{F}_2^8$.

3. Which functions in the Q-block cipher are responsible for

(a) confusion (non-linearity between nearby bits in the input and output),

(b) diffusion (spreading non-linearity across all output bits)?

4. Take the affine block cipher from Example 10.4, so the key is a pair (A, b) where A is an $n \times n$ matrix with entries from $\mathbb{F}_2$ and $b \in \mathbb{F}_2^n$.

Give a chosen plaintext attack that finds the key using $n+1$ encryptions of plaintexts of your choice. (Specify the chosen plaintexts precisely.)

($\star$) Can the key be determined using n or fewer encryptions?

5. Let $\mathcal{P} = \mathcal{C} = \mathbb{F}_2^8$. Consider the cryptosystem with keys $(k, k') \in \mathbb{F}_2^8 \times \mathbb{F}_2^8$ and encryption functions defined by $e_{(k,k')}(x) = P(x+k) + k'$, where P is the pseudo-inversion function from AES.

(a) Find $e^{-1}_{(k,k')}(z)$ for $z \in \mathbb{F}_2^8$.

(b) In a difference attack on this cryptosystem, the attacker takes $\mathbf{\Delta} = 1000\,0000$ corresponding to $1 \in \mathbb{F}_{2^8}$ and chooses $x \in \mathbb{F}_2^8$. She uses her black box to calculate $z = e_{(k,k')}(x)$ and $z_{\mathbf{\Delta}} = e_{(k,k')}(x_\Delta)$, and finds $\mathbf{\Gamma} = z + z_{\mathbf{\Delta}}$. Suppose that $\mathbf{\Gamma} \neq 1000\,0000$. Show, using Lemma 10.8, that she can find $\{k, k + \mathbf{\Delta}\}$.

(c) Find all possible keys (k, k') in terms of $\mathbf{\Gamma}$.

6. The University of Erewhon has, at fabulous expense, purchased an examination database from TTTT (Totally Trusted Transmission Technologies) in which the grades, which must be numbers between 0 and 99, are encrypted using 3DES with a fixed secret key $k \in \mathbb{F}_2^{168}$. A typical table is a set of ordered pairs

$$\big\{\big(\text{Alice}, e_k(75)\big), \big(\text{Bob}, e_k(40)\big), \big(\text{Charlie}, e_k(65)\big), \ldots\big\}.$$

Criticize the security of this system. How could it be improved?

7. In a variation on the Q-block cipher, the round keys $k^{(i)}$ are added using addition modulo 16, denoted $\boxplus$, rather than the usual addition in $\mathbb{F}_2^4$, denoted $+$. For example $0011 \boxplus 1001 = 1100$. Suggest a difference attack on the modified cipher.

8. Fix a block cipher with encryption maps $e_k : \mathbb{F}_2^n \to \mathbb{F}_2^n$ for $k \in \mathbb{F}_2^m$. Show that there exists $Q \in \mathbb{N}$ such that $e_k^Q = e_k^{-1}$ for all keys k. Does this mean all block ciphers can be broken?

MT362/462/5462 Cipher Systems: Sheet 9

Attempt questions 1 to 4 and 6. M.Sc. students should also attempt question 7: the final part ($\star$) is optional.

This sheet need not be handed in. Model answers will be posted on Moodle as usual. You are welcome to email the lecturer `mark.wildon@rhul.ac.uk` with any questions.

Private keys, and other private information, are written in red.

1. (a) Compute 2^{131} mod 3023. [*Hint:* to do this by hand, first compute 2^2, 2^4, 2^8, 2^{16} $\ldots 2^{128}$ by repeated squaring: note that $(2^m)^2 = 2^{2m}$.]

(b) Find x such that $2^x \equiv 35$ mod 37.

2. Suppose that Bob's RSA public key is $(2279, 17)$. As Eve you observe the RSA ciphertext 37 sent to Bob. Find Bob's private key and hence find the plaintext.

3. Generate an RSA public key (n, a) with $n > 2^{128}$ and private key (n, r). Use the MATHEMATICA notebook `PKC.nb` on Moodle and the `PowerMod` function.

(a) Email your public key to your partner in your cell.

(b) Email a message x of your choice, using the RSA Cryptosystem, to your partner in your block. [*Hint:* you know their public key when you receive their email from (a). Your message can be a number between 0 and $n-1$, or if you use the functions in the notebook, an English string.]

(c) Decrypt the message from your partner. [If your partner is uncooperative, you may use the lecturer as a substitute in (a) and (c).]

(d) Suppose all emails are observed by Eve. What, if anything, can she learn?

(e) Suppose all emails can be modified by Malcolm. What, if anything, can he learn?

4. Consider the cryptoscheme in which English plaintexts are converted to 8-bit ASCII (‘`a`’ ↔ 01100001, ‘`b`’ ↔ 01100010, and so on, as on Problem Sheet 5) and then encrypted using RSA with the appropriate public key.

For example ‘hi’ becomes 1101000 1101001 which is the binary form of 13409. If Alice's public key is (n, a) then she is sent 13409^a mod n. Assume that $n \approx 2^{2048}$.

(a) Alice is expected an important message ‘yes’ or ‘no’ from Bob. Show that Eve can decrypt Bob's ciphertext without knowing Alice's private key.

(b) Can the problem in (a) occur if Alice and Bob use a symmetric cipher such as AES where the key is entirely private? How can it be avoided while still using the RSA cryptosystem?

5. Let (n, a) be Alice's RSA public key. Suppose that $n = pq$. Let $t = (p-1)(q-1)$. Show that an attacker who knows n and t can easily find p and q. [*Hint:* find a quadratic equation for p with coefficients expressed in terms of n and t.]

6. In Diffie–Hellman Key Exchange, we saw that the eavesdropper Eve knows the prime p, the base g and g^a mod p. Only Alice knows her exponent a. (We write g^a mod p entirely in black because although a is private, g^a mod p is public.)

Bob wants to send a message $x \in \{1, \ldots, p-1\}$ to Alice.

(a) Suppose Bob sends xg^a mod p. Show that Eve can find x.

(b) Explain why Bob can send $x(g^a)^r$ mod p for any private r of his choice. (This is not entirely obvious because Bob knows g^a mod p but not g^a.) Can Alice find x?

(c) Suppose Bob sends $x(g^a)^r$ mod p and then sends r. Can Alice find x? Can Eve find x?

(d) Suppose Bob sends $x(g^a)^r$ mod p and then sends g^r mod p. Can Alice find x? Can Eve find x?

Remark: (d) is the ElGamal cryptoscheme: Alice publishes (g, g^a, p) as her public key, and keeps (g, a, p) as her private key.

7. **(M.Sc.)** Let $e_k : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ for $k \in \mathbb{F}_2^{12}$ be the encryption maps in the Q-block cipher. Find $\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,5\}})$ and $\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,6\}})$. Assuming you have good estimates for these statistics, and for $\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2\}}) = \frac{1}{2}(-1)^{k_0+k_6}$, how many possibilities are there for k? ($\star$) Find some further high correlations that give more information about the key.

8. **(M.Sc.)** Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$. Suppose that $\operatorname{corr}(L_U \circ F, L_T) = c > 0$. Let $k \in \mathbb{F}_2^n$ and define $G : \mathbb{F}_2^n \to \mathbb{F}_2^n$ by $G(x) = F(x + k)$.

(a) Show that $\operatorname{corr}(L_U \circ G, L_T) = (-1)^{L_T(k)} c$.

An attacker has a collection $\{(v^{(j)}, v'^{(j)}) : 1 \leq j \leq q\}$ of chosen plaintext/ciphertext pairs. She estimates the correlation in (a) by computing $S_j = L_U(v'^{(j)}) + L_T(v^{(j)})$ for each j, and taking $C = \frac{1}{q}\sum_{j=1}^{q}(-1)^{S_j}$.

(b) Find $\mathbb{P}[S_j = 0]$ and $\mathbb{P}[S_j = 1]$. [**Corrected from** $\mathbb{P}[Z_j = 1]$ **and** $\mathbb{P}[Z_j = -1]$**.**]

(c) Show that if q is large then the distribution of C is approximately normal with mean c and variance $\frac{1-c^2}{q}$. [*Hint:* use the Central Limit Theorem.]

(d) How large must q be for the attacker to be confident of learning $L_T(k)$?

9. Harry the Horrible Hacker creates an RSA public key (n, a).

(a) What is to stop him from publishing (n, a) on the web and claiming that it is Alice the Angelic's public key?

(b) Can he upload data to the web that appears to have been encrypted by Alice?

(c) What are the implications for laws that require key disclosure, for example a Section 49 notice under the Regulation of Investigatory Powers Act 2000?