MT362/462/5462 Cipher Systems: Sheet 1

Attempt at least questions 1 to 5. Please staple your answers to this sheet and put your name and student number at the top.

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: email `mark.wildon@rhul.ac.uk`.

To be handed in at the Monday lecture on 15th October.

☐ Tick this box if you *do not* want written feedback on your solutions.

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

The MATHEMATICA notebook `AlphabeticCiphers` on Moodle can be used to encrypt and decrypt using substitution ciphers, and compute frequencies and the Index of Coincidence. Please use it!

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 $x \in \{0, 1, \dots 99\}$ by encrypting it as the ciphertext $(x+k) \bmod 100$. Assume that the key $k \in \{0, 1, \dots 99\}$ 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?

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

3. In the first lecture you formed a *cell* 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 cell.

(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.

Write up (c), explaining your method. (An annotated printout is fine.) Did you learn the entire key?

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 at most 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.

```
12345678901234567890123456789012345678901234567890
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 `W` in position 1, 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 $R(x)_i = x_i + i \bmod 26$. For example $R(\text{bead}) =$ `CGDH` since bead $\longleftrightarrow (1, 4, 0, 3)$, $R\big((1, 4, 0, 3)\big) = (2, 6, 3, 7)$ and $(2, 6, 3, 7) \longleftrightarrow$ `CGDH`. Similarly $R^2(\text{aaaa}) =$ `CEGI`.

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 5 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.

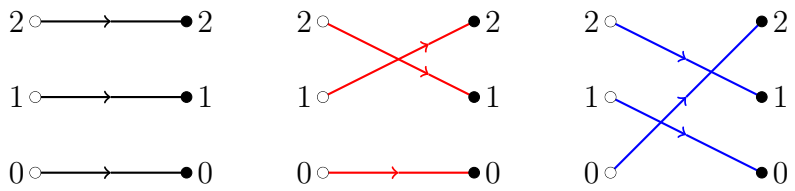
To be handed in by 5pm on Wednesday 24th October, or at the Monday 22nd October lecture.

☐ Tick this box if you *do not* want written feedback on your solutions.

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

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 $e_{(a,c)}(x) = ax+c$. 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. Let q be prime. Suppose that Alice and Bob communicate using the affine cipher on $\mathbb{Z}_q$ with keyspace $\mathcal{K} = \{(a, c) : a, c \in \mathbb{Z}_q, a \neq 0\}$, [**typo:** $\mathbb{Z}_q$**, not** $\mathbb{Z}_p$] and that Alice sends plaintext $x \in \mathbb{Z}_q$ with probability p_x.

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

(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 used with equal probability then the cryptosystem has perfect secrecy.

(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?

3. One of the encryption functions in the affine cipher on $\mathbb{Z}_{103}$ is $e_{(23,39)}$, defined by $x \mapsto 23x + 39$. What is $d_{(23,39)}(1)$?

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) 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 an appropriate 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. In this question we see what happens if either of the hypotheses is dropped.

(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 (a), (b) and (c) may fail to hold.

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

8. (a) Let $0 < c \leq 1$. By differentiating with respect to p find the maxima and minima of $-p \log p - (c - p) \log(c - p)$ for $0 \leq p \leq c$.

(b) Deduce that $-p \log p - q \log q \geq -(p + q) \log(p + q)$.

(c) Deduce that the entropy of the probability distribution $p_1, p_2, p_3, \ldots, p_n$ is at least the entropy of the probability distribution $p_1 + p_2, p_3, \ldots, p_n$.

Let $X : \Omega \to \mathcal{X}$ be a random variable taking values in a finite set $\mathcal{X}$. Let $f : \mathcal{X} \to \mathcal{X}$ be a function.

(d) Use (c) as part of a proof that $H(X) \geq H(f(X))$.

(e) Deduce that if f is bijective then $H(X) = H(f(X))$.

MT362/462/5462 Cipher Systems: Sheet 3

Attempt at least questions 1 to 4. Question 5 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 by noon on Wednesday 31st October, or at the Monday 29th October lecture.

☐ Tick this box if you *do not* want written feedback on your solutions.

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

1. 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 4.9.]

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

[*Hint:* the code used in the lecture is online at `repl.it/@mwildon/OneTimePad2`. If you do it by hand, it will be helpful to know that the plaintexts are four letter words related to cryptography.]

3. Alice and Bob communicate using the one-time pad cryptosystem of length n, in which $\mathcal{K} = \mathcal{P} = \mathcal{C} = \{\mathtt{a}, \mathtt{b}, \ldots, \mathtt{z}\}^n$. 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 $\mathbb{P}[Y_n = y | X_n = x] = \frac{1}{26^n}$ for all $y \in \mathcal{C}$.

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

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

(d) Deduce that the one-time pad has perfect secrecy.

(e) Is there a contradiction with the results of Question 2?

4. Let $\mathcal{A} = \{\mathrm{a}, \ldots, \mathrm{z}\}$

(a) Estimate the unicity distance (see Definition 4.15) 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

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

For example, if $n = 4$ then $e_{(\pi,\sigma)}(x_1, x_2, x_3, x_4) = \big(\pi(x_1), \sigma(x_2), \pi(x_3), \sigma(x_4)\big)$.

(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 14 of the printed lecture notes.]

5. (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 $j \in \{0, 1, 2, 3\}$, let $f_j : \mathbb{F}_2^4 \to \mathbb{F}_2$ be the Boolean function defined so that $f(x_3, x_2, x_1, x_0)$ is the bit in position j of $x_3x_2x_1x_0 + 5$ mod 16.

For example, since 0110 is the binary form of 6, $6 + 5 = 11$ and 11 has binary form 1011, we have $f_3\big((0,1,1,0)\big) = 1$, $f_2\big((0,1,1,0)\big) = 0$, $f_1\big((0,1,1,0)\big) = 1$ and $f_0\big((0,1,1,0)\big) = 1$.

Express each f_j as a polynomial in x_3, x_2, x_1, x_0.

6. The optional extras for Part A describe the game Shannon invented to estimate the per-character entropy of English plaintexts with letters a, ..., z and space.

(a) Play Shannon's game online at `repl.it/@mwildon/ShannonGuess2py`. Hit 'run' then type `main()` in the right-hand window to get started.

(b) Suppose the plaintext has length N. For each $i \in \mathbb{N}$, let f_i be the number of times you took i guesses to find the next letter. Let c be the number of times you asked the computer to reveal it. The output is then $h + 4.700c/N$ where h is the entropy of the probability distribution $f_1/N, f_2/N, \ldots$.

Explain why this is an estimate for the per-character entropy of English. [*Hint:* see Exercise 4.20. **Typo 4.18 corrected.**]

(c) What is your estimate for the per-character redundancy of English? Is it close to Shannon's estimate of 3.200 bits?

7. Show that $H(K|Y) \geq H(X|Y)$ in any cryptosystem. [*Hint:* use Question 8(d) on Sheet 2 with a suitable function $f : \mathcal{K} \to \mathcal{P}$.]

8. 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. Question 4 is compulsory for **Msc** 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 by noon on Wednesday 7th November, or at the Monday lecture.

☐ Tick this box if you *do not* want written feedback on your solutions.

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

The MATHEMATICA notebook used in lectures to find keystreams is available from Moodle.

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 $\{0, 1\}$.

(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 $(k_s, k_{s+1}, \ldots, k_{s+4}) = 10001$.

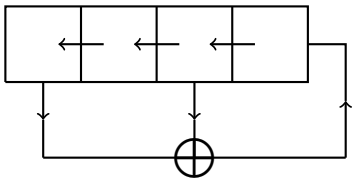
(iii) How would your answer to (i) change if the key was 10001?

(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.

(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?

2. Let F be the LFSR of width 4 with taps $\{0, 2\}$, as shown in the circuit diagram below.



(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

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

(a) Show that if F is invertible then $0 \in T$. [*Hint:* use the contrapositive.]

(b) Show conversely that if $0 \in T$ then F is invertible and give a formula for F^{-1}.

4. (**M.Sc.**) Define a boolean function $f : \mathbb{F}_2^5 \to \mathbb{F}$ by

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{cases} 1 & \text{if at least 3 of the } x_i \text{ are 1} \\ 0 & \text{otherwise.} \end{cases}$$

Express f in (i) algebraic normal form; (ii) disjunctive normal form; (iii) conjunctive normal form.

Prove a generalization of at least one of (i), (ii), (iii) in which 3 and 5 are replaced with arbitrary t and n. What is the connection with secret sharing schemes?

5. Let F be an LFSR of width ℓ with taps $T \subseteq \{0, 1, \ldots, \ell - 1\}$ and let G be an LFSR of width ℓ with taps $U \subseteq \{0, 1, \ldots, \ell - 1\}$.

(a) [**Corrected definition.**] Show that the function $H : \mathbb{F}_2^\ell \to \mathbb{F}_2^\ell$ defined by

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

is an LFSR of width ℓ.

(b) Determine the taps of H in terms of T and U.

6. Let F be an invertible LFSR of width ℓ and let $k = (0, \ldots, 0, 1) \in \mathbb{F}_2^\ell$. Let M be the matrix representing F as in Proposition 5.9.

(a) Show that the vectors $k, kM, \ldots, kM^{\ell-1}$ form a basis for $\mathbb{F}_2^\ell$.

(b) Deduce that the keystream for k has period equal to the period of F.

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. (The proof using LFSRs needs some finite field theory.)

8. Let M be the matrix, as in Proposition 5.9, representing an LFSR of width ℓ with taps T. Show that the characteristic polynomial of M is $X^\ell + \sum_{t\in T} X^t$. [*Hint:* use the Cayley–Hamilton Theorem and Lemma 5.10. Alternatively, expand $\det(M + XI)$ on the final column as a sum of ℓ determinants of minor matrices.]

MT362/462/5462 Cipher Systems: Sheet 5

Attempt questions 1 to 4. M.Sc. students should also do 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 by noon on Wednesday 14th November, or at the Monday lecture.

☐ Tick this box if you *do not* want written feedback on your solutions.

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

A MATHEMATICA notebook for generating keystreams and doing the encryption and decryptions in Question 2 is available from Moodle.

1. Let F be the LFSR of width 5 with taps $\{0, 2\}$.

(a) Calculate the keystream for the key 00001. What is its period?

(b) Write down $F(11010)$, $F^2(11010)$, $F^3(11010)$. [*Hint:* use your answer to (a) and the Very Useful Property.]

(c) Let $k \in \mathbb{F}_2^5$ be an arbitrary non-zero word. Show that the keystream for k contains 10000. What is its entry following 10000?

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 $\mathcal{P} = \{\mathtt{a}, \ldots, \mathtt{z}\}^n$ 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 5.3 with the LFSR F of width 5 with taps $\{0, 2\}$.

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 cell. [*Hint:* in MATHEMATICA use `EncryptString[{0, 2}, {k0, k1, k2, k3, k4}, "message"]`

(c) Decrypt the message from your partner.

(d) Decrypt either of the messages from the other two people in your cell. [*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 smallest number of ciphertext bits needed to determine the key?

3. (a) Show that if $k_0k_1k_2k_3k_4k_5k_6k_7$ is a keystream of an LFSR of width 4 then the matrix equation

$$\begin{pmatrix} k_0 & k_1 & k_2 & k_3 \\ k_1 & k_2 & k_3 & k_4 \\ k_2 & k_3 & k_4 & k_5 \\ k_3 & k_4 & k_5 & k_6 \end{pmatrix} \begin{pmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \end{pmatrix} = \begin{pmatrix} k_4 \\ k_5 \\ k_6 \\ k_7 \end{pmatrix}$$

has a solution t_0, t_1, t_2, t_3. [*Hint:* use the taps to define each $t_i \in \{0, 1\}$.]

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

(c) Which of the bit sequences 0010011, 1110000, 0110111 is a keystream of an LFSR of width 4? Justify your answer. Does your answer change if the LFSR is required to be invertible?

4. 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) Find the statistics M_{00}, M_{01}, M_{10}, M_{11} when the sequence is

$$\begin{array}{c} (0,1,0,1,1,0,0,1,0,1,0,1,0,1,0,1,0,0,1,1,0,0,1,1,0,0,1,0,1,0,1,1,0) \\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 0\ 1\ 2 \end{array}$$

(b) For $i \in \{0, 1, \ldots, n-2\}$, let

$$X_i = \begin{cases} 1 & \text{if } B_i = B_{i+1} = 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Fix i. Find $\mathbb{P}[X_i = 1]$ and hence write down $\mathbb{E}[X_i]$.

(ii) Explain why $M_{00} = X_0 + X_1 + \cdots + X_{n-2}$.

(iii) Hence show that $\mathbb{E}[M_{00}] = (n-1)/4$.

(c) What are $\mathbb{E}[M_{01}], \mathbb{E}[M_{10}], \mathbb{E}[M_{11}]$?

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

(e) Does the sequence in (a) pass the monobit test in Exercise 6.4?

5. (**M.Sc.**) Now it has been seen in lectures (and if you have not already done so) please complete the parts of Question 4 on Sheet 4 asking you to find the disjunctive and conjunctive normal form of the Boolean function $f : \mathbb{F}_2^5 \to \mathbb{F}$ defined by

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{cases} 1 & \text{if at least 3 of the } x_i \text{ are 1} \\ 0 & \text{otherwise.} \end{cases}$$

MT362/462/5462 Cipher Systems: Sheet 6

Attempt at least questions 1 and 2. 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 by noon on Wednesday 21st November, or at the Monday lecture.

☐ Tick this box if you *do not* want written feedback on your solutions.

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 $(k_0, k_1, k_2, \ldots)$ be a keystream of the LFSR F of width 2 with taps $\{0, 1\}$. 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 is 101100000101

(a) Explain why the keystreams of F and G have the form $1\star11\star\star\star\star\star1\star1$, 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 $1\star11\star00\star01\star1$. Hence calculate the taps and the key for G.

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 $\{0, 4\}$ and $\{0, 1\}$. [*Hint:* factorize 511 and use Lemma 5.12.]

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
$\widetilde{T}_n$	$\varnothing$	$\varnothing$	$\{2\}$	$\{1,2\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$	$\star$	$\{1,5,6,7\}$	$\star$	$\star$	$\star$	$\star$
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 backtaps $\widetilde{T}_9 = \{1\}$. Performing step 9 of the algorithm using $m = 4$ gives the LFSR F_{10} of length $\ell_{10} = 7$ and backtaps $\widetilde{T}_{10}$ that you are asked to find in (i). Since the length goes up, m is updated to 9.

(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 sequence $u_0, u_1, u_2, \ldots$ is generated by an LFSR of width 7, will the backtaps change in further steps of the Berlekamp–Massey algorithm? Justify your answer. [*Hint:* the method of Question 3 on Sheet 5 can be used.]

4. A keystream $k_0, k_1, k_2, \ldots$ can be encoded as a formal power series with coefficients in $\mathbb{F}_2$,

$$\sum_{s=0}^{\infty} k_s X^s = k_0 + k_1 X + k_2 X^2 + \cdots .$$

Given an LFSR F of width ℓ with taps T, define $\widetilde{g}_F(X) = 1 + \sum_{t \in T} X^{\ell - t}$.

(a) Show that the keystream $k_0, k_1, k_2, \ldots$ is the output of the LFSR F if and only if

$$\widetilde{g}_F(X) \sum_{s=0}^{\infty} k_s X^s = a(X)$$

for some polynomial $a(X)$ of degree at most $\ell - 1$.

(b) Let $k_0, k_1, k_2, \ldots$ and $k'_0, k'_1, k'_2, \ldots$ be keystreams of LFSRs F and F' of widths ℓ and ℓ' respectively. Show that $k_0 + k'_0, k_1 + k'_1, k_2 + k'_2, \cdots$ is the keystream of an LFSR G of width $\ell + \ell'$ with taps defined by

$$\widetilde{g}_G(X) = \widetilde{g}_F(x)\widetilde{g}_{F'}(x).$$

(c) Let $(u_0, u_1, u_2, \ldots)$ be the sym of the keystreams for the LFSRs of width 4 and 3 with taps $\{0, 3\}$ and $\{0, 1\}$ for keys k and k', respectively.

(i) Show, as claimed in Example 7.2 that $(u_0, u_1, u_2, \ldots)$ is a keystream of the LFSR H of width 7 with taps $\{0, 1, 5, 6\}$.

(ii) Show that the map $(k, k') \to (u_0, u_1, u_2, u_3, u_4, u_5, u_6)$ is injective. [*Hint:* show that if $(k, k') \to (0, 0, 0, 0, 0, 0, 0)$ then $(k_0, k_1, k_2, k_3, k_4, k_5, k_6)$ is a keystream of both the LFSRs F and G, and use (a).]

(iii) Suppose you know s and $(u_s, u_{s+1}, \ldots, u_{s+6})$. Explain how to find the keys k and k'.

(d) How is $\widetilde{g}_F(X)$ related to the minimal polynomial $g_F(X)$ of F?

MT362/462/5462 Cipher Systems: Sheet 7

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 or by appointment.

To be handed in at the Friday lecture on 30th November.

☐ Tick this box if you *do not* want written feedback on your solutions.

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 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 $(v,w) \in \mathbb{F}_2^{2m}$.

Let $(v',w') = (w, v+f(w))$. Noting the order w', v' carefully, express $F\big((w',v')\big)$ in terms of v and w.

2. Consider the Q-block cipher as defined in Example 8.5, 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 $k^{(3)} = (k_8,k_9,k_{10},k_{11})$.

(a) Encrypt $(0,0,0,0,0,0,0,0) \in \mathbb{F}_2^8$ using the key $(0,0,1,1,0,0,1,1,0,0,1,1)$.

(b) Decrypt the ciphertext $(0,1,1,1,0,1,1,1)$ using the key in (a).

(c) Find a key $k \in \mathbb{F}_2^{12}$ such that $e_k\big((0,0,0,1,0,0,0,1)\big) = (0,0,0,0,0,0,0,0)$.

(d) Show that given $(v,w) \in \mathbb{F}_2^8$ and $w' \in \mathbb{F}_2^8$ 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 $(0,0,0,1,0,0,0,1)$ and $(0,0,0,0,0,0,0,0)$ 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 $(y, z) \in \mathbb{F}_2^{2m}$ 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

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

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 in about 2^{112} operations.

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

(c) (Optional.) Why is the middle map decryption rather than encryption? [*Hint:* think 'backward compatability'.]

5. (**M.Sc.**) The LFSR of width 4 with taps $\{0, 3\}$ has keystream 100011110101100 with period 15. What is the minimum width of an LFSR with keystream 100011110101101? [*Hint:* use the theoretical results in §4 of the M.Sc. course. You can check your answer by applying the Berlekamp–Massey algorithm.]

6. The 2-quadratic stream cipher was defined in Example 7.5. Recall that F is the LFSR of width 5 with taps $\{0, 2\}$ and F' is the LFSR of width 6 with taps $\{0, 1, 3, 4\}$. 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 $k = 01000$.

[*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?]

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 either of the Friday 7th December lectures.

☐ Tick this box if you *do not* want written feedback on your solutions.

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 $S : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ 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 8.5 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 9.7(i), namely 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 $\Gamma = 0000\,0010$. 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. Show that no matter what the key is, $(v', w') + (v'_\Gamma, w'_\Gamma)$ is equally likely to be each of the four differences $\{0010\,1000, 0010\,1001, 0010\,1010, 0010\,1011\}$. [**Corrected 5th December.**]

2. Let e_k for $k \in \mathbb{F}_2^{12}$ be the encryption maps in the Q-block cipher. Show using Lemma 9.7(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. For $k \in \mathbb{F}_2^8$ define $e_k : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ by $e_k(x) = P(x) + k$ where $P : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ is the pseudo-inverse function defined by identifying $\mathbb{F}_2^8$ with the finite field $\mathbb{F}_{2^8}$. In Example 9.4 we attacked the cipher with key space $\mathbb{F}_2^8 \times \mathbb{F}_2^8$ and encryption functions defined by

$$e_{(k_{\text{otp}},k)}(x) = e_k(x + k_{\text{otp}}).$$

By Exercise 9.6(a), the attack typically finds k and one false key $k + \Gamma$ using 2^9 decryptions. Please finish this exercise:

(b) How many encryptions are needed to test all the pairs (k_{otpguess}, k) and $(k_{\text{otpguess}}, k + \Gamma)$ for $k_{\text{otpguess}} \in \mathbb{F}_2^8$?

(c) Deduce that the attack finds the key (k_{otp}, k) using at most 2^{10} encryptions/decryptions. Why is this sub-exhaustive?

5. (M.Sc.) By Theorem 5.8(c) in the M.Sc. notes, if $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is a Boolean function then

$$(-1)^f = \sum_{T \subseteq \{0,\ldots,n-1\}} \text{corr}(f, L_T)(-1)^{L_T}$$

where $L_T(x_0, \ldots, x_{n-1}) = \sum_{t \in T} x_t$.

(a) Define $\overline{f} : \mathbb{F}_2^n \to \mathbb{F}$ by $\overline{f}(x) = \overline{f(x)}$. Show that

$$\text{corr}(\overline{f}, L_T) = -\text{corr}(f, L_T).$$

(b) Show that $\sum_{T \subseteq \{0,\ldots,n-1\}} \text{corr}(f, L_T)^2 = 1$.

(c) The Toffoli function is the 3-variable Boolean function defined by

$$g(x_0, x_1, x_2) = \begin{cases} \overline{x_2} & \text{if } x_0 = x_1 = 1 \\ x_2 & \text{otherwise} \end{cases}.$$

Compute the eight correlations $\text{corr}(g, L_T)$ for $T \subseteq \{0, 1, 2\}$ and so check that part (b) and Theorem 5.8(c) holds for g.

(d) Observe that $(\frac{3}{4})^2 + 7(\frac{1}{4})^2 = \frac{9}{16} + 7\frac{1}{16} = 1$. Hence it is consistent with (b) that there is a 3-variable Boolean function whose eight correlations are $\frac{3}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}$ in some order. Is there such a function?

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

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

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

7. Find a difference attack on the one-time-pad / AES cipher in which the attacker first guesses k_{otp} and after that k. [**Correction: not the Q-block cipher**]

MT362/462/5462 Cipher Systems: Sheet 9

Attempt questions 1 to 4 and 6. **M.Sc.** students should also attempt question 7.

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} ... 2^{128} by repeated squaring: note that $(2^m)^2 = 2^{2m}$.]

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

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

3. Using the MATHEMATICA notebook `PKC.nb` on Moodle generate an RSA public key (n, c) with $n > 2^{128}$ and private key (p, q, r).

(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 cell. [*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 the 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 non-public key cipher such as AES? 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) Suppose Bob sends $x(g^a)^r$ mod p for some private r of his choice. 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 $\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2,5\}})$ and $\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2,6\}})$. Assuming you have good estimates for these statistics, and for $\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2\}}) = \frac{1}{2}(-1)^{k_0+k_6}$, how many possibilities are there for k?

8. (**M.Sc.**) Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$. Suppose that $\mathrm{corr}(L_U \circ F, L_T) = c > 0$. 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 $\mathrm{corr}(L_U \circ G, L_T) = (-1)^{L_T(k)} c$.

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

(b) Find $\mathbb{P}[Z_j = 1]$ 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. The SHA-256 hash function h takes values in $\mathbb{F}_2^{256}$. To validate a block of bitcoin transactions a miner must find $x \in \mathbb{F}_2^{256}$ such that $h(x)$ begins with 72 zeros. A modern general purpose microprocessor requires about 128 cycles to calculate a single $h(x)$, and runs at 4GHz, so executes 4×10^9 cycles per second. A special purpose device advertised on the web for \$1790 claims 44×10^{12} hashes per second.

(a) Estimate the time required to validate a block using both systems.

(b) The reward for the miner is 12.5 bitcoins (the reward is halved every 210,000 blocks) plus the transaction fees for all the transactions in the newly validated block. What are the implication for the bitcoin economy?