MT5462 Advanced Cipher Systems

Mark Wildon, mark.wildon@rhul.ac.uk

Administration:

- Please take the first installment of the notes.
- All handouts will be put on Moodle marked **MSc**.
- **Lectures:** Monday 4pm (MFLEC), Friday 11am (MC201), Friday 4pm (MC336).
- **Extra lecture for MT5462:** Friday 9am (MC201).
- **Office hours in McCrea 240:** Tuesday 3.30pm, Wednesday 10am, Thursday 11am.

§1 Revision of fields and polynomials

Definition 1.1

A *field* is a set of elements $\mathbb{F}$ with two operations, $+$ (addition) and $\times$ (multiplication), and two special elements $0, 1 \in \mathbb{F}$ such that $0 \neq 1$ and

(1) $a + b = b + a$ for all $a, b \in \mathbb{F}$;

(2) $0 + a = a + 0 = a$ for all $a \in \mathbb{F}$;

(3) for all $a \in \mathbb{F}$ there exists $b \in \mathbb{F}$ such that $a + b = 0$;

(4) $a + (b + c) = (a + b) + c$ for all $a, b, c \in \mathbb{F}$;

(5) $a \times b = b \times a$ for all $a, b \in \mathbb{F}$;

(6) $1 \times a = a \times 1 = a$ for all $a \in \mathbb{F}$;

(7) for all non-zero $a \in \mathbb{F}$ there exists $b \in \mathbb{F}$ such that $a \times b = 1$;

(8) $a \times (b \times c) = (a \times b) \times c$ for all $a, b, c \in \mathbb{F}$;

(9) $a \times (b + c) = a \times b + a \times c$ for all $a, b, c \in \mathbb{F}$.

If $\mathbb{F}$ is finite, then we define its *order* to be its number of elements.

Exercise: Show, from the field axioms, that if $x \in \mathbb{F}$, then x has a unique additive inverse, and that if $x \neq 0$ then x has a unique multiplicative inverse. Show also that if $\mathbb{F}$ is a field then $a \times 0 = 0$ for all $a \in \mathbb{F}$.

Exercise: Show from the field axioms that if $\mathbb{F}$ is a field and a, $b \in \mathbb{F}$ are such that $ab = 0$, then either $a = 0$ or $b = 0$.

Theorem 1.2

Let p be a prime. The set $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ with addition and multiplication defined modulo p is a finite field of order p.

Example 1.3

The addition and multiplication tables for the finite field $\mathbb{F}_4 = \{0, 1, \alpha, 1+\alpha\}$ of order 4 are

$+$	0	1	α	$1+\alpha$
0	0	1	α	$1+\alpha$
1	1	0	$1+\alpha$	α
α	α	$1+\alpha$	0	1
$1+\alpha$	$1+\alpha$	α	1	0

$\times$	1	α	$1+\alpha$
1	1	α	$1+\alpha$
α	α	$1+\alpha$	1
$1+\alpha$	$1+\alpha$	1	α

Definition 1.4

If $f(x) = a_0 + a_1 x + a_2 + \cdots + a_m x^m$ where $a_m \neq 0$, then we say that m is the *degree* of the polynomial f, and write $\deg f = m$. We leave the degree of the zero polynomial undefined. We say that a_0 is the *constant term*.

Lemma 1.5 (Division algorithm)

Let $\mathbb{F}$ be a field, let $g(x) \in \mathbb{F}[x]$ be a non-zero polynomial and let $g(x) \in \mathbb{F}[x]$. There exist polynomials $s(x), r(x) \in \mathbb{F}[x]$ such that

$$f(x) = s(x)g(x) + r(x)$$

and either $r(x) = 0$ or $\deg r(x) < \deg g(x)$.

We say that $s(x)$ is the *quotient* and $r(x)$ is the *remainder* when $f(x)$ is divided by $g(x)$. Lemma 1.5 will not be proved in lectures. The important thing is that you can compute the quotient and remainder. In MATHEMATICA: `PolynomialQuotientRemainder`.

Lemma 1.7

Let $\mathbb{F}$ be a field.

(i) *If $f \in \mathbb{F}[x]$ has $a \in \mathbb{F}$ as a root, i.e. $f(a) = 0$, then there is a polynomial $g \in \mathbb{F}[x]$ such that $f(x) = (x - a)g(x)$.*

(ii) *If $f \in \mathbb{F}[x]$ has degree $m \in \mathbb{N}_0$ then f has at most m distinct roots in $\mathbb{F}$.*

(iii) *Suppose that $f, g \in \mathbb{F}[x]$ are non-zero polynomials such that $\deg f, \deg g < t$. If there exist distinct $c_1, \ldots, c_t \in \mathbb{F}$ such that $f(c_i) = g(c_i)$ for each $i \in \{1, \ldots, t\}$ then $f = g$.*

Lemma 1.7

Let $\mathbb{F}$ be a field.

(i) *If $f \in \mathbb{F}[x]$ has $a \in \mathbb{F}$ as a root, i.e. $f(a) = 0$, then there is a polynomial $g \in \mathbb{F}[x]$ such that $f(x) = (x - a)g(x)$.*

(ii) *If $f \in \mathbb{F}[x]$ has degree $m \in \mathbb{N}_0$ then f has at most m distinct roots in $\mathbb{F}$.*

(iii) *Suppose that $f, g \in \mathbb{F}[x]$ are non-zero polynomials such that $\deg f, \deg g < t$. If there exist distinct $c_1, \ldots, c_t \in \mathbb{F}$ such that $f(c_i) = g(c_i)$ for each $i \in \{1, \ldots, t\}$ then $f = g$.*

Part (iii) is the critical result. It says, for instance, that a linear polynomial is determined by any two of its values: when $\mathbb{F}$ is the real numbers $\mathbb{R}$ this should be intuitive—there is a unique line through any two distinct points. Similarly a quadratic is determined by any three of its values, and so on.

Lemma 1.8 (Polynomial interpolation)

Let $\mathbb{F}$ be a field. Let

$$c_1, c_2, \ldots, c_t \in \mathbb{F}$$

be distinct and let $y_1, y_2, \ldots, y_t \in \mathbb{F}$. The unique polynomial $f(x) \in \mathbb{F}[x]$ of degree $< t$ such that $f(c_i) = y_i$ for all i is

$$f(x) = \sum_{i=1}^{t} y_i \frac{\prod_{j \neq i}(x - c_j)}{\prod_{j \neq i}(c_i - c_j)}.$$

§2 : Shamir's Secret Sharing Scheme

Example 2.1

Ten people want to know their mean salary. But none is willing to reveal her salary s_i to the others, or to a 'Trusted Third Party'. Instead Person 1 chooses a large number M. She remembers M, and whispers $M + s_1$ to Person 2. Then Person 2 whispers $M + s_1 + s_2$ to Person 3, and so on, until finally Person 10 whispers $M + s_1 + s_2 + \cdots + s_{10}$ to Person 1. Person 1 then subtracts M and can tell everyone the mean $(s_1 + s_2 + \cdots + s_{10})/10$.

Exercise 2.3

In the two person version of the scheme, Person 1 can deduce Person 2's salary from $M + s_1 + s_2$ by subtracting $M + s_1$. Is this a defect in the scheme? [**Typo in notes:** N **should be** M**.**]

Definition 2.4

Let p be a prime and let $s \in \mathbb{F}_p$. Let $n \in \mathbb{N}$, $t \in \mathbb{N}$ be such that $t \leq n < p$. Let $c_1, \ldots, c_n \in \mathbb{F}_p$ be distinct non-zero elements. In the *Shamir scheme* with n people and *threshold* t, Trevor chooses at random $a_1, \ldots, a_{t-1} \in \mathbb{F}_p$ and constructs the polynomial

$$f(x) = s + a_1 x + \cdots + a_{t-1} x^{t-1}$$

with constant term s. Trevor then issues the *share* $f(c_i)$ to Person i.

Example 2.5

Suppose that $n = 5$ and $t = 3$. Take $p = 7$ and $c_i = i$ for each $i \in \{1, 2, 3, 4, 5\}$. We suppose that $s = 5$. Trevor chooses $a_1, a_2 \in \mathbb{F}_7$ at random, getting $a_1 = 6$ **[typo in notes]** and $a_2 = 1$. Therefore $f(x) = 5 + 6x + x^2$ and the share of Person i is $f(c_i)$, for each $i \in \{1, 2, 3, 4, 5\}$, so

$$\big(f(1), f(2), f(3), f(4), f(5)\big) = (5, 0, 4, 3, 4).$$

Exercise 2.6

Suppose that Person 1, with share $f(1) = 5$, and Person 2, with share $f(2) = 0$, cooperate in an attempt to discover s. Show that for each $z \in \mathbb{F}_7$ there exists a unique polynomial $f_z(x)$ such that $\deg f \leq 2$ and $f(0) = z$, $f_z(1) = 5$ and $f_z(2) = 0$.

Theorem 2.7

In a Shamir scheme with n people, threshold t and secret s, any t people can determine s but any $t-1$ people can learn nothing about s.

Lemma 1.7

Let $\mathbb{F}$ be a field.

(i) *If $f \in \mathbb{F}[x]$ has $a \in \mathbb{F}$ as a root, i.e. $f(a) = 0$, then there is a polynomial $g \in \mathbb{F}[x]$ such that $f(x) = (x - a)g(x)$.*

(ii) *If $f \in \mathbb{F}[x]$ has degree $m \in \mathbb{N}_0$ then f has at most m distinct roots in $\mathbb{F}$.*

(iii) *Suppose that $f, g \in \mathbb{F}[x]$ are non-zero polynomials such that $\deg f, \deg g < t$. If there exist distinct $c_1, \ldots, c_t \in \mathbb{F}$ such that $f(c_i) = g(c_i)$ for each $i \in \{1, \ldots, t\}$ then $f = g$.*

Lemma 1.8 (Polynomial interpolation)

Let $\mathbb{F}$ be a field. Let $c_1, c_2, \ldots, c_t \in \mathbb{F}$ be distinct and let $y_1, y_2, \ldots, y_t \in \mathbb{F}$. The unique polynomial $f(x) \in \mathbb{F}[x]$ of degree $< t$ such that $f(c_i) = y_i$ for all i is

$$f(x) = \sum_{i=1}^{t} y_i \frac{\prod_{j \neq i}(x - c_j)}{\prod_{j \neq i}(c_i - c_j)}.$$

Example 2.8

The root key for DNSSEC, part of web of trust that guarantees an IP connection really is to the claimed end-point, and not Malcolm doing a Man-in-the-Middle attack, is protected by a secret sharing scheme with $n = 7$ and $t = 5$: search for ‘Schneier DNSSEC’.

Exercise 2.9

Take the Shamir scheme with threshold t and evaluation points $1, \ldots, n \in \mathbb{F}_p$ where $p > n$. Trevor has shared two large numbers r and s across n cloud computers, using polynomials f and g so that the shares are $\big(f(1), \ldots, f(n)\big)$ and $\big(g(1), \ldots, g(n)\big)$.

(a) How can Trevor secret share $r + s \bmod p$?

(b) How can Trevor secret share $rs \bmod p$?

Note that all the computation has to be done on the cloud!

Exercise 2.10

Suppose Trevor shares $s \in \mathbb{F}_p$ across n computers using the Shamir scheme with threshold t. He chooses t computers and gets them to reconstruct s. Unfortunately Malcolm has compromised one of these computers. Show that Malcolm can both learn s and trick Trevor into thinking his secret is any chosen $s' \in \mathbb{F}_p$.

Remark 2.11

The Reed–Solomon code associated to the parameters p, n, t and the field elements $c_1, c_2, \ldots, c_n$ is the length n code over $\mathbb{F}_p$ with codewords all possible n-tuples

$$\{\big(f(c_1), f(c_2), \ldots, f(c_n)\big) : f \in \mathbb{F}_p[x],\ \deg f \le t-1\}.$$

It will be studied in MT5461. By Theorem 2.7, each codeword is determined by any t of its positions. Thus two codewords agreeing in $n-t+1$ positions are equal: this shows the Reed–Solomon code has minimum distance at least $n-t+1$.

Remark 2.11

The Reed–Solomon code associated to the parameters p, n, t and the field elements $c_1, c_2, \ldots, c_n$ is the length n code over $\mathbb{F}_p$ with codewords all possible n-tuples

$$\{(f(c_1), f(c_2), \ldots, f(c_n)) : f \in \mathbb{F}_p[x],\ \deg f \leq t-1\}.$$

It will be studied in MT5461. By Theorem 2.7, each codeword is determined by any t of its positions. Thus two codewords agreeing in $n-t+1$ positions are equal: this shows the Reed–Solomon code has minimum distance at least $n-t+1$.

For simplicity we have worked over a finite field of prime order in this section. Reed–Solomon codes and the Shamir secret sharing scheme generalize in the obvious way to arbitrary finite fields. For example, the Reed–Solomon codes used on compact discs have alphabet the finite field $\mathbb{F}_{2^8}$.

§3 Introduction to Boolean Functions

Definition 3.1

Let $n \in \mathbb{N}$. An n-variable *boolean function* is a function $\mathbb{F}_2^n \to \mathbb{F}_2$.

Any boolean function is uniquely determined by its *truth table*, which records the pairs $\big(x, f(x)\big)$ for each $x \in \mathbb{F}_2^n$. For example, the truth tables for and, or (denoted $\vee$) and not (denoted $\neg$) are shown below.

x	y	xy
0	0	0
0	1	0
1	0	0
1	1	1

x	y	$x \vee y$
0	0	0
0	1	1
1	0	1
1	1	1

x	$\neg x$
0	1
1	0

Exercise 3.2

Show that there are 2^{2^n} boolean function in n variables.

Boolean functions can be expressed in many different ways, not always obviously the same. In this section we look at some normal forms for Boolean functions.

Exercise 3.4

(i) Write the two variable function $f(x, y) = x \vee y$ as a polynomial in x and y.

(ii) What logical connective corresponds to $(x, y) \mapsto x + y$?

(iii) Define $\mathrm{maj}(x_1, x_2, x_3)$ to be true if at least two of x_1, x_2, x_3 are true, and otherwise false. Express maj as a polynomial.

(iv) Express $x_1x_2 \vee x_2x_3 \vee x_3x_4$ as a polynomial.

Algebraic Normal Form

Exercise 3.5

Find a simple form for the product of $f(x_1, x_2, x_3) = x_1(\neg x_2)x_3$ and $\mathrm{maj}(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$.

We define a *boolean monomial* to be a product of the form $x_{i_1} \dots x_{i_r}$ where $i_1 < \dots < i_r$. Given $I \subseteq \{1, \dots, n\}$, let

$$x_I = \prod_{i \in I} x_i.$$

By definition (or convention if you prefer), $x_\varnothing = 1$.

Lemma 3.6

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ *be a Boolean function. Then*

$$f(x_1, x_2, \dots, x_n) = x_1 g(x_2, \dots, x_n) + (1 + x_1)h(x_2, \dots, x_n)$$

where

$$g(x_2, \dots, x_n) = f(1, x_2, \dots, x_n)$$
$$h(x_2, \dots, x_n) = f(0, x_2, \dots, x_n).$$

Example 3.7

The Toffoli gate is important in quantum computation. It takes 3 input qubits and returns 3 output qubits. Its classical analogue is the 3 variable Boolean function defined in words by 'if x_1 and x_2 are both true then negate x_3, else return x_3'. Using Lemma 3.6, one gets the polynomial form $x_1x_2 + x_3$.

Theorem 3.8

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an n-variable Boolean function. There exist unique coefficients $c_I \in \{0, 1\}$, one for each $I \subseteq \{1, \ldots, n\}$, such that

$$f(x_1, \ldots, x_n) = \sum_{I \subseteq \{1,\ldots,n\}} x_I.$$

This expression for f is called the *algebraic normal form* of f.

Example 3.9

Let $f : \mathbb{F}_2^3 \to \mathbb{F}_2$ be a 3-variable Boolean function

(a) Show that the coefficient $c_\varnothing$ of $x_\varnothing = 1$ in f is $f(0,0,0)$.

(b) Show that the coefficient $c_{\{3\}}$ of $x_{\{3\}} = x_3$ in f is $f(0,0,0) + f(0,0,1)$.

(c) Show that the coefficient $c_{\{1,2\}}$ of $x_{\{1,2\}} = x_1x_2$ in f is $f(0,0,0) + f(1,0,0) + f(0,1,0) + f(1,1,0)$.

For example, by (c), if $f(x_1, x_2, x_3) = x_1x_2 + x_3$ is the Toffoli function seen in Example 3.7 then

$$f(0,0,0) + f(1,0,0) + f(0,1,0) + f(1,1,0) = 0 + 0 + 0 + 1 = 1$$

is the coefficient of x_1x_2.

Exercise 3.10

What do you think is the formula for the coefficient $c_{\{2,3\}}$? Does it work for the Toffoli function? How about if $f(x_1, x_2, x_3) = x_1x_2x_3$?

Proposition 3.11

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ [**Typo:** was $\mathbb{F}_2^n$] *be an n-variable Boolean function and suppose that f has algebraic normal form*

$$f(x_1, \ldots, x_n) = \sum_{I \subseteq \{1,\ldots,n\}} c_I x_I .$$

Then

$$c_I = \sum f(z_1, \ldots, z_n)$$

where the sum is over all $z_1, \ldots, z_n \in \{0, 1\}$ *with* $\{j : z_j = 1\} \subseteq I$*.*

We reduced to this claim in the case $K = \{1, \ldots, k\}$.

Claim

If $K \subseteq \{1, 2, \ldots, n\}$ then

$$\sum_{(z_1,\ldots,z_n)} x_K(z_1, \ldots, z_n) = \begin{cases} 1 & \text{if } I = K \\ 0 & \text{otherwise} \end{cases},$$

where the sum is as in the proposition.

Disjunctive Normal Form

Definition 3.12

Fix $n \in \mathbb{N}$. Given $J \subseteq \{1, \ldots, n\}$ let

$$f_J(x_1, \ldots, x_n) = z_1 \wedge z_2 \wedge \cdots \wedge z_n$$

where

$$z_j = \begin{cases} x_j & \text{if } j \in J \\ \neg x_j & \text{if } j \notin J. \end{cases}$$

A n-variable Boolean function of the form $\bigvee_{J \in \mathcal{B}} f_J$, where $\mathcal{B}$ is a collection of subsets of $\{1, 2, \ldots, n\}$, is said to be in *disjunctive normal form*.

By definition, or convention if you prefer, the empty disjunction is false.

Example 3.13

(a) We saw in Exercise 3.4 that

$$\mathrm{maj}(x_1, x_2, x_3) = (x_1 \wedge x_2) \vee (x_2 \wedge x_3) \vee (x_1 \wedge x_3).$$

From this it is a short step to the disjunctive normal form

$$\begin{aligned}\mathrm{maj}(x_1, x_2, x_3) = (x_1 \wedge x_2 \wedge \neg x_3) \vee (x_1 \wedge \neg x_2 \wedge x_3) \\ \vee (\neg x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_3).\end{aligned}$$

(b) The truth table for $f(x_1, x_2, x_3) = x_1 x_2 + x_3$ is

x_1	x_2	x_3	f
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1

x_1	x_2	x_3	f
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

So $f(x_1, x_2, x_3)$ is true when the set of true variables is $\{3\}$, $\{2, 3\}$, $\{1, 3\}$ and $\{1, 2\}$. This easily gives the disjunctive normal form.

Theorem 3.14

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be an n-variable Boolean function. There exists a unique collection $\mathcal{B}$ of subsets of $\{1, \ldots, n\}$ such that

$$f(x_1, \ldots, x_n) = \bigvee_{J \in \mathcal{B}} f_J.$$

Definition 3.15

Fix $n \in \mathbb{N}$. Given $J \subseteq \{1, \dots, n\}$, let $g_J = z_1 \vee \cdots \vee z_n$ where, as in Definition 3.12,

$$z_j = \begin{cases} x_j & \text{if } j \in J \\ \neg x_j & \text{if } j \notin J. \end{cases}.$$

A Boolean function of the form $\bigvee_{J \in \mathcal{B}} g_J$, where $\mathcal{B}$ is a collection of subsets of $\{1, \dots, n\}$, is said to be in *conjunctive normal form*.

Example 3.16

The majority vote function maj on 3-variables is false if and only if at least two of the variables are false. Hence $\neg\mathrm{maj}(x_1, x_2, x_3) = f_\varnothing \vee f_{\{1\}} \vee f_{\{2\}} \vee f_{\{3\}}$ in disjunctive normal form and so

$$\begin{aligned} \mathrm{maj}(x_1, x_2, x_3) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee x_3) \\ \wedge (x_1 \vee \neg x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee \neg x_3) \end{aligned}$$

in conjunctive normal form.

§4 Berlekamp–Massey Algorithm

If F is an LFSR width ℓ with taps $T \subseteq \{0, 1, \ldots, \ell - 1\}$ then, for each position $s \in \mathbb{N}$,

$$F(k_s, \ldots, k_{s+\ell-1}) = (k_{s+1}, \ldots, k_{s+\ell-1}, \sum_{t \in T} k_{s+t}).$$

Hence (as seen in Question 1 of Sheet 5), $k_{s+\ell} = \sum_{t \in T} k_{s+t}$. Setting $r = s + \ell$ this becomes

$$k_r = \sum_{t \in T} k_{r-\ell+t} \quad \text{for } r \geq \ell. \tag{$\dagger$}$$

Proposition 4.1

Let $n \geq \ell$. If an LFSR F of width ℓ generates the keystream $(k_0, k_1, \ldots, k_{n-1}, c)$ of length $n + 1$ then any LFSR F' generating the keystream $(k_0, k_1, \ldots, k_{n-1}, \neg c)$ has width ℓ' where

$$\ell' \geq n + 1 - \ell.$$

As a final preliminary, we need the *symmetric difference* of sets T and U defined by

$$T \mathbin{\triangle} U = \{v \in T \cup U : v \notin T \cap U\}.$$

The following lemma shows how symmetric differences arise when we combine LFSRs.

Lemma 4.2 (**corrected**)

Let F and G be LFSRs of width ℓ with taps T and U respectively. The function H defined by

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

is an LFSR with taps $T \mathbin{\triangle} U$.

Example 4.3

The keystream of the LFSR F of width 5 with taps $\{0, 1, 2\}$ for the key $(0, 1, 1, 0, 0)$ has period 14.

$$\underset{0}{(0},\underset{1}{1},\underset{2}{1},\underset{3}{0},\underset{4}{0},\underset{5}{0},\underset{6}{0},\underset{7}{1},\underset{8}{0},\underset{9}{0},\underset{10}{1},\underset{11}{1},\underset{12}{1},\underset{13}{1},\ldots)$$

The set $\widetilde{T}$ is $\{5-0, 5-1, 5-2\} = \{3, 4, 5\}$ and, as claimed by (‡), $k_n = k_{n-3} + k_{n-4} + k_{n-5}$ for all $n \in \mathbb{N}$ with $n \geq 5$.

We use the following notation in the algorithm.

- $k_0, k_1, k_2, \ldots$ is the keystream;
- for $n \in \mathbb{N}$, ℓ_n is the minimal width of an LFSR F_n with taps T_n generating the first n positions $k_0, k_1, \ldots, k_{n-1}$ of the keystream;
- $\ell_0 = 0$ and $T_0 = \varnothing$;
- $\widetilde{T}_n = \{\ell_n - t : t \in T_n\}$ (the set of *backwards taps*)

By convention, the LFSR of width 0, necessarily with taps $\varnothing$, generates $(0, 0, \ldots)$. It is the unique minimal width LFSR generating this keystream.

Theorem 4.4

Let $n \in \mathbb{N}_0$*.*

(a) *If the LFSR* F_n *generates* $(k_0, k_1, \ldots, k_{n-1}, k_n)$ *then* $\ell_{n+1} = \ell_n$ *and we may take* $\widetilde{T}_{n+1} = \widetilde{T}_n$*.*

(b) *Suppose the LFSR* F_n *generates* $(k_0, k_1, \ldots, k_{n-1}, \neg k_n)$*. If* $\ell_n = 0$ *then let* $m = 0$*, else let* m *be maximal such that* $\ell_m < \ell_{m+1}$*. Let* $U = \{\widetilde{t} + n - m : \widetilde{t} \in \widetilde{T}_m\}$*. Then setting*

$$\widetilde{T}_{n+1} = \widetilde{T}_n \bigtriangleup (U \cup \{n - m\})$$

defines an LFSR F_{n+1} *of minimal width*

$$\ell_{n+1} = \max(\ell_n, n + 1 - \ell_n)$$

that generates $(k_0, k_1, \ldots, k_{n-1}, k_n)$*.*

Results Used in Proof

- If $k_0, k_1, \ldots, k_{n-1}$ is generated by the LFSR with taps $\widetilde{T}$ then

$$k_r = \sum_{\widetilde{t} \in \widetilde{T}} k_{r-\widetilde{t}} \quad \text{for } r < n. \tag{‡}$$

- Proposition 4.1
  Let $n \geq \ell$. If an LFSR F of width ℓ generates the keystream $(k_0, k_1, \ldots, k_{n-1}, c)$ of length $n+1$ then any LFSR F' generating the keystream $(k_0, k_1, \ldots, k_{n-1}, \neg c)$ has width ℓ' where

$$\ell' \geq n + 1 - \ell.$$

- For any sets $\widetilde{S}$ and $\widetilde{T}$,

$$\sum_{\widetilde{t} \in S \,\triangle\, T} k_{r-\widetilde{t}} = \sum_{\widetilde{t} \in \widetilde{S}} k_{r-\widetilde{t}} + \sum_{\widetilde{t} \in \widetilde{T}} k_{r-\widetilde{t}}.$$

Example 4.1

We take the first 12 positions of the keystream generated by the LFSR in Example 4.3 but change the final 1 to 0.

$$(0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, \ldots)$$

0 1 2 3 4 5 6 7 8 9 10 11

The table below shows ℓ_n and the set $\widetilde{T}_n$ for each n. Where case (ii) applies, the relevant m is shown. The final row indicates whether the set of taps is unique. (This is not given by the algorithm, but can be determined using the linear algebra method.)

n	1	2	3	4	5	6	7	8	9	10	11	12
ℓ_n	0	2	2	2	3	3	3	5	5	5	5	7
$\widetilde{T}_n$	$\varnothing$	$\{1\}$	$\{1\}$	$\{1,2\}$	$\{1,2,3\}$	$\varnothing$	$\varnothing$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,5\}$
m	0		1	1	4		4				7	
unique?	✓	×	×	✓	×	×	✓	×	×	✓	✓	×

Example 4.5 [continued]

n	1	2	3	4	5	6	7	8	9	10	11	12
ℓ_n	0	2	2	2	3	3	3	5	5	5	5	7
$\widetilde{T}_n$	$\varnothing$	$\{1\}$	$\{1\}$	$\{1,2\}$	$\{1,2,3\}$	$\varnothing$	$\varnothing$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,5\}$
m	0		1	1	4		4				7	
unique?	✓	×	×	✓	×	×	✓	×	×	✓	✓	×

- Initialization: $\ell_0 = 0,\ T_0 = \widetilde{T}_0 = \varnothing$.
- Choose F_1: since $k_0 = 0$, the minimal width LFSR generating (k_0) is the unique LFSR of width 0, with taps $T_1 = \widetilde{T}_1 = \varnothing$.
- Step $n = 1$: case (b), we got $\ell_2 = 2,\ \widetilde{T}_2 = \{1\}$.
- Step $n = 2$: case (a), we got $\ell_3 = \ell_2 = 2,\ \widetilde{T}_3 = \widetilde{T}_2 = \{1\}$.
- Step $n = 3$: since F_3 generates $(0,1,1,1)$ which is wrong in position 3, case (b) applies. The length last increased at step 1, so $m = 1$. We have

$$U = \{\widetilde{t} + 3 - 1 : \widetilde{t} \in \widetilde{T}_1\} = \varnothing$$

and $\widetilde{T}_4 = \widetilde{T}_3 \triangle (U \cup \{3-1\}) = \{1\} \triangle \{2\} = \{1,2\}$. We take $\ell_4 = \max(\ell_3, 3+1-\ell_3) = \max(2,2) = 2$.

Example 4.5 [continued]

n	1	2	3	4	5	6	7	8	9	10	11	12
ℓ_n	0	2	2	2	3	3	3	5	5	5	5	7
$\widetilde{T}_n$	$\varnothing$	$\{1\}$	$\{1\}$	$\{1,2\}$	$\{1,2,3\}$	$\varnothing$	$\varnothing$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,4,5\}$	$\{3,5\}$
m	0		1	1	4		4				7	
unique?	✓	×	×	✓	×	×	✓	×	×	✓	✓	×

- Step $n = 4$: since F_4 generates $(0,1,1,0,1)$ which is wrong in position 4, case (b) applies. Again $m = 1$. We have

$$U = \{\widetilde{t} + 4 - 1 : \widetilde{t} \in \widetilde{T}_1\} = \varnothing$$

and $\widetilde{T}_5 = \widetilde{T}_4 \triangle (U \cup \{4-1\}) = \{1,2\} \triangle \{3\} = \{1,2,3\}$. We take $\ell_5 = \max(\ell_4, 4+1-\ell_4) = \max(2,3) = 3$.

- Step $n = 5$: since F_5 generates $(0,1,1,0,0,1)$, which is wrong in position 5, case (b) applies. The length increased at step 4, so $m = 4$. We have

$$U = \{\widetilde{t} + 5 - 4 : \widetilde{t} \in \widetilde{T}_4\} = \{2,3\}$$

and $\widetilde{T}_5 = \widetilde{T}_4 \triangle (U \cup \{5-4\}) = \{1,2,3\} \triangle \{2,3,1\} = \varnothing$. We take $\ell_6 = \max(\ell_5, 5+1-\ell_5) = \max(3,3) = 3$.

§5 Discrete Fourier Transform

Given $x \in \mathbb{F}_2$ we define $(-1)^x$ by regarding x as an ordinary integer. Thus $(-1)^0 = 1$ and $(-1)^1 = -1$. Given $f : \mathbb{F}_2^n \to \mathbb{F}_2$ we define $(-1)^f : \mathbb{F}_2^n \to \{-1, 1\}$ by $(-1)^f(x) = (-1)^{f(x)}$.

Definition 5.1

Let $f, g : \mathbb{F}_2^n \to \mathbb{F}$ be Boolean functions. We define the *correlation* between f and g by

$$\mathrm{corr}(f, g) = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)}(-1)^{g(x)}.$$

Lemma 5.2

Let $f, g : \mathbb{F}_2^n \to \mathbb{F}$ be Boolean functions. Let

$$\begin{aligned} m_{\mathrm{same}} &= \left|\{x \in \mathbb{F}_2^n : f(x) = g(x)\}\right| \\ m_{\mathrm{diff}} &= \left|\{x \in \mathbb{F}_2^n : f(x) \neq g(x)\}\right|. \end{aligned}$$

Then $\mathrm{corr}(f, g) = (m_{\mathrm{same}} - m_{\mathrm{diff}})/2^n$.

Exercise 5.3

Let $X \in \mathbb{F}_2^n$ be a random variable distributed uniformly at random, so $\mathbb{P}[X = x] = 1/2^n$ for each $x \in \mathbb{F}_2^n$. Show that

$$\mathrm{corr}(f, g) = \mathbb{P}[f(X) = g(X)] - \mathbb{P}[f(X) \neq g(X)]$$

and

$$\begin{aligned} \mathbb{P}[f(X) = g(X)] &= \tfrac{1}{2}(1 + \mathrm{corr}(f, g)), \\ \mathbb{P}[f(X) \neq g(X)] &= \tfrac{1}{2}(1 - \mathrm{corr}(f, g)). \end{aligned}$$

Given $T \subseteq \{1, \ldots, n\}$, define $L_T : \mathbb{F}_2^n \to \mathbb{F}_2$ by

$$L_T(x) = \sum_{t \in T} x_t.$$

We think of L_T as 'tapping' (like an LFSR) the positions in T. For example, $L_{\{i\}}(x_1, \ldots, x_n) = x_i$ returns the entry in position i and $L_\varnothing(x) = 0$ is the zero function.

Exercise 5.4

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. Show that $\mathrm{corr}(f, L_\varnothing) = 0$ if and only if $\mathbb{P}[f(X) = 0] = \mathbb{P}[f(X) = 1] = \frac{1}{2}$.

Lemma 5.5

The linear functions $\mathbb{F}_2^n \to \mathbb{F}$ are precisely the $L_T : \mathbb{F}_2^n \to \mathbb{F}_2$ for $T \subseteq \{1, \ldots, n\}$. If $S, T \subseteq \{1, \ldots, n\}$ then

$$\mathrm{corr}(L_S, L_T) = \begin{cases} 1 & \text{if } S = T \\ 0 & \text{otherwise.} \end{cases}$$

Example 5.6

Let $\mathrm{maj} : \mathbb{F}_2^3 \to \mathbb{F}_2$ be the majority vote function from Exercise 3.4(ii). Then

$$\mathrm{corr}(\mathrm{maj}, L_T) = \begin{cases} \frac{1}{2} & \text{if } T = \{1\}, \{2\}, \{3\} \\ -\frac{1}{2} & \text{if } T = \{1, 2, 3\} \\ 0 & \text{otherwise.} \end{cases}$$

Moreover

$$(-1)^{\mathrm{maj}} = \tfrac{1}{2}(-1)^{L_{\{1\}}} + \tfrac{1}{2}(-1)^{L_{\{2\}}} + \tfrac{1}{2}(-1)^{L_{\{3\}}} - \tfrac{1}{2}(-1)^{L_{\{1,2,3\}}}.$$

To generalize the previous example, we define an inner product on the vector space of functions $\mathbb{F}_2^n \to \mathbb{R}$ by

$$\langle \theta, \phi \rangle = \frac{1}{2^n} \sum_{x \in 2^n} \theta(x)\phi(x).$$

Exercise: check that, as required for an inner product, $\langle \theta, \theta \rangle \geq 0$ and that $\langle \theta, \theta \rangle = 0$ if and only if $\theta(x) = 0$ for all $x \in \mathbb{F}_2^n$.

Lemma 5.7

Let $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ be Boolean functions. Then

$$\langle (-1)^f, (-1)^g \rangle = \operatorname{corr}(f, g).$$

To generalize the previous example, we define an inner product on the vector space of functions $\mathbb{F}_2^n \to \mathbb{R}$ by

$$\langle \theta, \phi \rangle = \frac{1}{2^n} \sum_{x \in 2^n} \theta(x)\phi(x).$$

Exercise: check that, as required for an inner product, $\langle \theta, \theta \rangle \geq 0$ and that $\langle \theta, \theta \rangle = 0$ if and only if $\theta(x) = 0$ for all $x \in \mathbb{F}_2^n$.

Theorem 5.8 (Discrete Fourier Transform)

(a) *The functions $(-1)^{L_T}$ for $T \subseteq \{1, \ldots, n\}$ are an orthonormal basis for the vector space of functions $\mathbb{F}_2^n \to \mathbb{R}$.*

(b) *Let $\theta : \mathbb{F}_2^n \to \mathbb{R}$. Then*

$$\theta = \sum_{T \subseteq \{1,\ldots,n\}} \langle \theta, (-1)^{L_T} \rangle (-1)^{L_T}.$$

(c) *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. Then*

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

§6 Linear Cryptanalysis

Example 6.1

Let $S : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ be the S-box in the Q-block cipher (see Example 8.4 in the main notes), defined by

$$S\big((x_0, x_1, x_2, x_3)\big) = (x_2, x_3, x_0 + x_1x_2, x_1 + x_2x_3).$$

(a) Suppose we look at position 0 of the output by considering $L_{\{0\}} \circ S : \mathbb{F}_2^4 \to \mathbb{F}_2$. We have

$$(L_{\{0\}} \circ S)\big((x_0, x_1, x_2, x_3)\big) = x_2 = L_{\{2\}}\big((x_0, x_1, x_2, x_3)\big).$$

Hence $L_{\{0\}} \circ S = L_{\{2\}}$. By Lemma 5.5,

$$\operatorname{corr}(L_{\{0\}} \circ S, L_T) = \begin{cases} 1 & \text{if } T = \{2\} \\ 0 & \text{otherwise.} \end{cases}.$$

§6 Linear Cryptanalysis

Example 6.1

Let $S : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ be the S-box in the Q-block cipher (see Example 8.4 in the main notes), defined by

$$S\big((x_0, x_1, x_2, x_3)\big) = (x_2, x_3, x_0 + x_1x_2, x_1 + x_2x_3).$$

(b) Instead if we look at position 2, the relevant Boolean function is $L_{\{2\}} \circ S$, for which $L_{\{2\}} \circ S\big((x_0, x_1, x_2, x_3)\big) = x_0 + x_1x_2$.
Exercise: show that

$$\operatorname{corr}(L_{\{2\}} \circ S, L_T) = \begin{cases} \frac{1}{2} & \text{if } T = \{0\}, \{0,1\}, \{0,2\} \\ -\frac{1}{2} & \text{if } T = \{0,1,2\} \\ 0 & \text{otherwise} \end{cases}.$$

(This generalizes the correlations computed in Example 7.2 in the main course.)

Example 6.2

For $k \in \mathbb{F}_2^{12}$ let $e_k : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ be the Q-block cipher, as defined in Example 8.4. Then $e_k\big((v, w)\big) = (v', w')$ where

$$v' = w + S\big(v + S(w + k^{(1)}) + k^{(2)}\big).$$

Recall that $k^{(1)} = (k_0, k_1, k_2, k_3)$ and $k^{(2)} = (k_4, k_5, k_6, k_7)$. Example 6.1 suggests looking at $\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2\}})$. (See the optional question on Problem Sheet 9 for the theoretical justification for this.) We have

$$\begin{aligned}(L_{\{0\}} \circ e_k)\big((v, w)\big) = L_{\{0\}}\big((v', w')\big) &= v'_0 \\ L_{\{2\}}\big((v, w)\big) &= v_2.\end{aligned}$$

Exercise: using that $k^{(1)}_0 = k_0$, $k^{(1)}_1 = k_1$, $k^{(1)}_2 = k_2$ and $k^{(2)}_2 = k_6$, check that

$$v'_0 = v_2 + (w_1 + k_1)(w_2 + k_2) + k_0 + k_6.$$

Example 6.2 [continued]

When we compute $\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2\}})$ by averaging over all $(v, w) \in \mathbb{F}_2^8$, the values of k_1 and k_2 are irrelevant. For instance, if both are 0 we average $(-1)^{w_1 w_2}$ over all four $(w_1, w_2) \in \mathbb{F}_2^2$ to get $\frac{1}{2}$; if both are 1 we average $(-1)^{(w_1+1)(w_2+1)}$, seeing the same summands in a different order, and still getting $\frac{1}{2}$. Hence

$$
\begin{aligned}
\mathrm{corr}(L_{\{0\}} \circ e_k, L_{\{2\}}) &= \frac{1}{2^8} \sum_{(v,w)\in\mathbb{F}_2^8} (-1)^{v_2+w_1 w_2+k_0+k_6}(-1)^{v_2} \\
&= \frac{1}{2^8} \sum_{(v,w)\in\mathbb{F}_2^8} (-1)^{w_1 w_2+k_0+k_6} \\
&= (-1)^{k_0+k_6} \frac{1}{4} \sum_{w_1,w_2\in\{0,1\}} (-1)^{w_1 w_2} \\
&= \tfrac{1}{2}(-1)^{k_0+k_6}.
\end{aligned}
$$

We can estimate this correlation from a collection of plaintext/ciphertext pairs $(v, w), (v', w')$ by computing $(-1)^{v'_0+v_2}$ for each pair. The average is $\frac{1}{2}(-1)^{k_0+k_6}$ which tells us $k_0 + k_6$.

Attack on the Q-block cipher

Using our collection of plaintext/ciphertext pairs we can also estimate

$$\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,5\}}) = \tfrac{1}{2}(-1)^{k_0+k_6+k_1}$$
$$\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,6\}}) = \tfrac{1}{2}(-1)^{k_0+k_6+k_2}$$

and so learn k_1 and k_2 as well as $k_0 + k_6$. There are similar high correlations of $\frac{1}{2}$ for output bit 1. Using these one learns k_2 and k_3 as well as $k_1 + k_7$.

Exercise 6.3

Given $k_0 + k_6, k_1 + k_7, k_1, k_2, k_3$, how many possibilities are there for the key in the Q-block cipher?

Attack on the Q-block cipher

Using our collection of plaintext/ciphertext pairs we can also estimate

$$\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,5\}}) = \tfrac{1}{2}(-1)^{k_0+k_6+k_1}$$
$$\operatorname{corr}(L_{\{0\}} \circ e_k, L_{\{2,6\}}) = \tfrac{1}{2}(-1)^{k_0+k_6+k_2}$$

and so learn k_1 and k_2 as well as $k_0 + k_6$. There are similar high correlations of $\frac{1}{2}$ for output bit 1. Using these one learns k_2 and k_3 as well as $k_1 + k_7$.

Exercise 6.3

Given $k_0 + k_6, k_1 + k_7, k_1, k_2, k_3$, how many possibilities are there for the key in the Q-block cipher?

The attack by differential cryptanalysis required chosen plaintexts. The attack by linear cryptanalysis works with any observed collection of plaintext/ciphertext pairs. It is therefore more widely applicable, as well as more powerful.