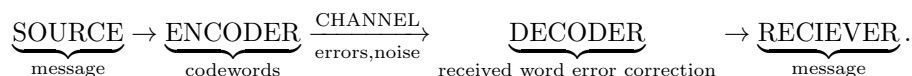


## Introduction

We model communication:



**Examples:** optical signals, electrical telegraph, SMS (compression), postcodes, CDs (error correction), zip/gz files (compression).

Given a source and a channel, modelled probabilistically, the basic problem is to design an encoder and decoder to transmit messages economically (noiseless coding; compression) and reliably (noisy coding).

**Examples:**

- Noiseless coding: Morse code: common letters are assigned shorter code-words, e.g.  $A \mapsto \bullet-$ ,  $E \mapsto \bullet$ ,  $Q \mapsto --\bullet-$ ,  $S \mapsto \bullet\bullet\bullet$ ,  $O \mapsto ---$ ,  $Z \mapsto --\bullet\bullet$ . Noiseless coding is adapted to source.
- Noisy coding: Every book has an ISBN  $a_1, a_2, \dots, a_9, a_{10}$ ,  $a_i \in \{0, 1, \dots, 9\}$  for  $1 \leq i \leq 9$  and  $a_{10} \in \{0, 1, \dots, 9, X\}$  with  $\sum_{j=1}^{10} ja_j \equiv 0 \pmod{11}$ . This detects common errors - e.g. one incorrect digit, transposition of two digits. Noisy coding is adapted to the channel.

**Plan:**

- (I) Noiseless coding - entropy
- (II) Error correcting codes - noisy channels
- (III) Information theory - Shannon's theorems
- (IV) Examples of codes
- (V) Cryptography

**Books:** [GP], [W], [CT], [TW], Buchmann, Körner. Online notes: Carne, Körner.

## Basic Definitions

**Definition** (Communication channel). A *communication channel* accepts symbols from a alphabet  $\mathcal{A} = \{a_1, \dots, a_r\}$  and it outputs symbols from alphabet  $\mathcal{B} = \{b_1, \dots, b_s\}$ . Channel modelled by the probabilities  $\mathbb{P}(y_1 \dots y_n \text{ received} | x_1 \dots x_n \text{ sent})$ . A *discrete memoryless channel* (DMC) is a channel with

$$p_{ij} = \mathbb{P}(b_j \text{ received} | a_i \text{ sent})$$

the same for each channel use and independent of all past and future uses. The channel matrix is  $P = (b_{ij})$ , a  $r \times s$  stochastic matrix.

**Definition** (Binary symmetric channel). The *binary symmetric channel* (BSC) with error probability  $p \in [0, 1)$  from  $\mathcal{A} = \mathcal{B} = \{0, 1\}$ . The channel matrix is

$$\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}.$$

A symbol is transmitted correctly with probability  $1 - p$ . Usually assume  $p < 1/2$ .

The *binary erasure channel* (BEC) has  $\mathcal{A} = \{0, 1\}$ ,  $\mathcal{B} = \{0, 1, *\}$ . The channel matrix is

$$\begin{pmatrix} 1-p & 0 & p \\ 0 & 1-p & p \end{pmatrix}.$$

So  $p = \mathbb{P}(\text{symbol can't be read})$ .

**Definition.** We model  $n$  uses of a channel by the  $n$ th extension, with input alphabet  $\mathcal{A}^n$  and output alphabet  $\mathcal{B}^n$ . A *code*  $C$  of length  $n$  is a function  $\mathcal{M} \rightarrow \mathcal{A}^n$  where  $\mathcal{M}$  is the set of possible messages. Implicitly we also have a decoding rule  $\mathcal{B}^n \rightarrow \mathcal{M}$ . The *size* of  $C$  is  $m = |\mathcal{M}|$ . The *information rate* is  $\rho(C) = \frac{1}{n} \log_2 m$ . The *error rate* is  $\hat{e}(C) = \max_{x \in \mathcal{M}} \mathbb{P}(\text{error} | x \text{ sent})$ .

**Remark.** For the remainder of the course we write  $\log$  instead of  $\log_2$ .

**Definition.** A channel can *transmit reliably at rate*  $R$  if there exists  $(C_n)_{n=1}^\infty$  with each  $C_n$  a code of length  $n$  such that

$$\lim_{n \rightarrow \infty} \rho(C_n) = R \text{ \& } \lim_{n \rightarrow \infty} \hat{e}(C_n) = 0.$$

The *capacity* is the supremum of all reliable transmission rates. We'll see in Chapter 9 that a BSC with error probability  $p < 1/2$  has non-zero capacity.

## 1 Noiseless coding

### 1.1 Prefix-free codes

For an alphabet  $\mathcal{A}$ ,  $|\mathcal{A}| < \infty$ , let  $\mathcal{A}^* = \bigcup_{n \geq 0} \mathcal{A}^n$ , the set of all finite strings from  $\mathcal{A}$ . The *concatenation* of strings  $x = x_1 \dots x_r$  and  $y = y_1 \dots y_s$  is  $xy = x_1 \dots x_r y_1 \dots y_s$ .

**Definition.** Let  $\mathcal{A}, \mathcal{B}$  be alphabets. A code is a function  $c : \mathcal{A} \rightarrow \mathcal{B}^*$ . The strings  $c(a)$  for  $a \in \mathcal{A}$  are called *codewords* or *words* (CWS).

**Example 1.1** (Greek fire code).  $\mathcal{A} = \{\alpha, \beta, \dots, \omega\}$  (greek alphabet),  $\mathcal{B} = \{1, 2, 3, 4, 5\}$ ,  $c : \alpha \mapsto 11, \beta \mapsto 12, \dots, \psi \mapsto 53, \omega \mapsto 54$ .  $xy$  means hold up  $x$  torches and another  $y$  torches nearby.

**Example 1.2.**  $\mathcal{A}$  = words in a dictionary,  $\mathcal{B} = \{A, B, \dots, Z, \omega\}$ .  $c : \mathcal{A} \rightarrow \mathcal{B}$  splits the word and follows with a space. Send message  $x_1 \dots x_n \in \mathcal{A}^*$  as  $c(x_1) \dots c(x_n) \in \mathcal{B}^*$ . So  $c$  extends to a function  $c^* : \mathcal{A}^* \rightarrow \mathcal{B}^*$ .

**Definition.**  $c$  is said to be *decipherable* if the induced map  $c^*$  (as in the previous example) is injective. In other words, each string from  $\mathcal{B}$  corresponds to at most one message.

Clearly if  $c$  is decipherable, it is necessary for  $c$  to be injective. However it is not sufficient:

**Example 1.3.**  $\mathcal{A} = \{1, 2, 3, 4\}$ ,  $\mathcal{B} = \{0, 1\}$ . Define  $c : 1 \mapsto 0, 2 \mapsto 1, 3 \mapsto 00, 4 \mapsto 01$ . Then  $c^*(114) = 0001 = c^*(312) = c^*(144)$  yet  $c$  is injective.

**Notation:**  $|\mathcal{A}| = m$ ,  $|\mathcal{B}| = a$ , call  $c$  an  $a$ -ary code of size  $m$ . For example a 2-ary code is a binary one, and a 3-ary code is a ternary code.

Our aim is to construct decipherable codes with short word lengths. Assuming  $c$  is injective, the following codes are always decipherable:

- (i) A block code has all codewords of the same length (e.g Greek fire code);
- (ii) A comma code reserves a letter from  $\mathcal{B}$  to signal the end of a word (e.g Example 1.2);
- (iii) A prefix-free code is a code where no codeword is a prefix of any other distinct word (if  $x, y \in \mathcal{B}^*$  then  $x$  is a prefix of  $y$  if  $y = xz$  for some string  $z \in \mathcal{B}^*$ ).

(i) and (ii) are special cases of (iii). As we can decode the message as it is received, prefix-free codes are sometimes called *instantaneous*.

**Exercise:** find a decipherable code which is not prefix-free.

**Definition** (Kraft's inequality).  $|\mathcal{A}| = m$ ,  $|\mathcal{B}| = a$ ,  $c : \mathcal{A} \rightarrow \mathcal{B}^*$  has word lengths  $l_1, \dots, l_m$ . Then Kraft's inequality is

$$\sum_{i=1}^m a^{-l_i} \leq 1. \quad (*)$$

**Theorem 1.1.** A prefix-free code exists if and only if Kraft's inequality  $(*)$  holds.

*Proof.* Rewrite  $(*)$  as

$$\sum_{l=1}^s n_l a^{-l} \leq 1, \quad (**)$$

where  $n_l$  is the number of codewords with length  $l$ , and  $s = \max_{1 \leq i \leq m} l_i$ .

Now if  $c : \mathcal{A} \rightarrow \mathcal{B}^*$  is prefix-free,

$$n_1 a^{s-1} + n_2 a^{s-2} + \dots + n_{s-1} a + n_s \leq a^s.$$

Indeed the LHS is the number of strings of length  $s$  in  $B$  with some codeword of  $c$  as a prefix, and the RHS is the total number of strings of length  $S$ . Dividing through by  $a^s$  we get (\*\*).

Now given  $n_1, \dots, n_s$  satisfying (\*\*), we try to construct a prefix-free code  $c$  with  $n_l$  codewords of length  $l$ ,  $\forall l \leq s$ . Proceed by induction on  $s$ ,  $s = 1$  is clear (since (\*\*) gives  $n_1 \leq a$  so can construct code).

By the induction hypothesis, there exists a prefix-code  $\hat{c}$  with  $n_l$  codewords of length  $l$  for all  $l \leq s - 1$ . Then (\*\*) implies

$$n_1 a^{s-1} + n_2 a^{s-2} + \dots + n_{s-1} a + n_s \leq a^s.$$

The first  $s - 1$  terms on the LHS sum to the number of strings of length  $s$  with a codeword of  $\hat{c}$  as a prefix and the RHS is the number of strings of length  $s$ . Hence we can add at least  $n_s$  new codewords of length  $s$  to  $\hat{c}$  and maintain the prefix-free property. □

**Remark.** This proof is constructive: just choose codewords in order of increasing length, ensuring that no previous codeword is a prefix.

**Theorem 1.2** (McMillan). *Any decipherable code satisfies Kraft's inequality.*

*Proof (Karush, 1961).* Let  $c : \mathcal{A} \rightarrow \mathcal{B}^*$  be a decipherable code with word lengths  $l_1, \dots, l_m$ . Set  $s = \max_{1 \leq i \leq m} l_i$ . For  $R \in \mathbb{N}$

$$\left( \sum_{i=1}^m a^{-l_i} \right)^R = \sum_{l=1}^{Rs} b_l a^{-l}, \quad (\dagger)$$

where  $b_l$  is the number of ways of choosing  $R$  codewords of total length  $l$ . Since  $c$  is decipherable, any string of length  $l$  formed from codewords must correspond to at most one sequence of codewords, i.e  $b_l \leq |\mathcal{B}^l| = a^l$ . Subbing this into (†)

$$\left( \sum_{i=1}^m a^{-l_i} \right)^R \leq \sum_{l=1}^{Rs} a^l a^{-l} = Rs,$$

so

$$\sum_{i=1}^m a^{-l_i} \leq (Rs)^{1/R} \rightarrow 1 \text{ as } R \rightarrow \infty.$$

Hence  $\sum_{i=1}^m a^{-l_i} \leq 1$ . □

**Corollary 1.3.** *A decipherable code with prescribed word lengths exists if and only if a prefix-free code with the same word lengths exists.*

*Proof.* Combine previous two theorems. □

Therefore we can restrict our attention to prefix-free codes.

## 2 Shannon's Noiseless Coding Theorem

*Entropy* is a measure of 'randomness' or 'uncertainty'. Suppose we have a random variable  $X$  taking a finite set of values  $x_1, \dots, x_n$  with probabilities  $p_1, \dots, p_n$  respectively. The *entropy*  $H(X)$  of  $X$  is the expected number of fair coin tosses needed to simulate  $X$  (roughly speaking).

**Example 2.1.** Suppose  $p_1 = p_2 = p_3 = p_4 = 1/4$ . Identify  $(x_1, x_2, x_3, x_4)$  with  $(HH, HT, TH, TT)$ . Then the entropy is 2.

**Example 2.2.** Suppose  $(p_1, p_2, p_3, p_4) = (1/2, 1/4, 1/8, 1/8)$ . Identify  $(x_1, x_2, x_3, x_4)$  with  $(H, TH, TTH, TTT)$ . Then the entropy is

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 3 + \frac{1}{8} \times 3 = \frac{7}{4}.$$

In a sense, the previous example (2.1) was 'more random' than this.

**Definition** (Entropy). The *entropy* of  $X$  is

$$H(X) = - \sum_{i=1}^b p_i \log p_i.$$

(Recall that  $\log =: \log_2$  here.) Note  $H(X) \geq 0$ . It is measured in *bits* (binary digits). Conventionally, we take  $0 \log 0 = 0$ .

**Example 2.3.** Take a biased coin  $\mathbb{P}(H) = p, \mathbb{P}(T) = 1 - p$ . Write  $H(p, 1 - p) := H(p)$ . Then

$$H(p) = -p \log p - (1 - p) \log(1 - p).$$

Note that  $H'(p) = \log \frac{1-p}{p}$ . Hence the entropy is maximised for  $p = 1/2$  (giving entropy 1).

**Proposition 2.1** (Gibbs' inequality). Let  $(p_1, \dots, p_n), (q_1, \dots, q_n)$  be probability distributions. Then

$$- \sum_{i=1}^n p_i \log p_i \leq - \sum_{i=1}^n p_i \log q_i.$$

(The RHS is sometimes called the *cross entropy* or *mixed entropy*) Furthermore we have equality iff  $p_i = q_i$  for all  $i$ .

*Proof.* Since  $\log x = \frac{\ln x}{\ln 2}$ , we may replace  $\log$  with  $\ln$ . Put  $I = \{1 \leq i \leq n : p_i \neq 0\}$ . Now  $\ln x = x - 1$  for all  $x > 0$  with equality iff  $x = 1$ . Hence  $\ln \frac{q_i}{p_i} \leq \frac{q_i}{p_i} - 1$  for all  $i \in I$ . So

$$\begin{aligned} \sum_{i \in I} p_i \ln \frac{q_i}{p_i} &\leq \underbrace{\sum_{i \in I} q_i}_{\leq 1} - \underbrace{\sum_{i \in I} p_i}_{=1} \leq 0 \\ \implies - \sum_{i \in I} p_i \ln p_i &\leq - \sum_{i \in I} p_i \ln q_i \end{aligned}$$

$$\implies -\sum_{i=1}^n p_i \ln p_i \leq -\sum_{i=1}^n p_i \ln q_i.$$

If equality holds, then  $\sum_{i \in I} q_i = 1$  and  $\frac{p_i}{q_i} = 1$  for all  $i \in I$ . So  $q_i = p_i$  for all  $1 \leq i \leq n$ .  $\square$

**Corollary 2.2.**  $H(p_1, p_2, \dots, p_n) \leq \log n$  with equality iff  $p_1 = p_2 = \dots = p_n = 1/n$ .

*Proof.* Take  $q_1 = q_2 = \dots = q_n = 1/n$  in Gibbs' inequality.  $\square$

Let  $\mathcal{A} = \{\mu_1, \dots, \mu_m\}$ ,  $|\mathcal{B}| = a$  ( $m, n \geq 2$ ). The random variable  $X$  takes values  $\mu_1, \dots, \mu_m$  with probabilities  $p_1, \dots, p_m$ .

**Definition.** If  $c : \mathcal{A} \rightarrow \mathcal{B}^*$  is a code, we say it is *optimal* if it has the smallest possible expected word length. i.e.  $\mathbb{E}S := \sum_{i=1}^m p_i l_i$  is minimal amongst all decipherable codes.

**Theorem 2.3** (Shannon's Noiseless Coding Theorem). *The expected word length  $\mathbb{E}S$  of an optimal code satisfies*

$$\frac{H(X)}{\log a} \leq \mathbb{E}S < \frac{H(X)}{\log a} + 1.$$

**Remark.** The lower bound is actually true for any decipherable code.

*Proof.* We first get the lower bound. Let  $c : \mathcal{A} \rightarrow \mathcal{B}^*$  be decipherable with word lengths  $l_1, \dots, l_m$ . Let  $q_i = \frac{a^{-l_i}}{D}$  where  $D = \sum_{i=1}^m a^{-l_i}$ . Note  $\sum_{i=1}^m q_i = 1$ . By Gibbs' inequality

$$\begin{aligned} H(X) &\leq -\sum_{i=1}^m p_i \log q_i \\ &= -\sum_{i=1}^m p_i (-l_i \log a - \log D) \\ &= \left( \sum_{i=1}^m p_i l_i \right) \log a + \log D. \end{aligned}$$

By McMillan,  $D \leq 1$  so  $\log D \leq 0$ . Hence

$$H(X) \leq \left( \sum_{i=1}^m p_i l_i \right) \log a \implies \frac{H(X)}{\log a} \leq \mathbb{E}S.$$

And we have equality iff  $p_i = a^{-l_i}$  for some integers  $l_1, \dots, l_m$ . Note we have only used decipherability so far.

Now we get the upper bound. Take  $l_i = \lceil -\log_a p_i \rceil$ . Then

$$-\log_a p_i \leq l_i < -\log_a p_i + 1.$$

Hence  $\log_a p_i \geq -l_i$ , so  $p_i \geq a^{-l_i}$ . Therefore  $\sum_{i=1}^m a^{-l_i} \leq \sum_{i=1}^m p_i = 1$ . By Kraft's inequality, there exists a prefix-free code  $c$  with word lengths  $l_1, \dots, l_m$ .  $c$  has expected word length

$$\mathbb{E}S = \sum_{i=1}^m p_i l_i < \sum_{i=1}^m p_i (-\log_a p_i + 1) = \frac{H(X)}{\log a} + 1.$$

□

**Example 2.4** (Shannon-Fano Coding). We mimic the above proof: given  $p_1, \dots, p_m$ , set  $l_i = \lceil -\log_a p_i \rceil$ . Construct a prefix-free code with word lengths  $l_i$  by choosing codewords in order of increasing length, ensuring any new codeword has no previous codeword as a prefix (Kraft's inequality ensures we can do this).

**Example 2.5.** Take  $a = 2, m = 5$ .

$i$	$p_i$	$\lceil -\log_2 p_i \rceil$	code
1	0.4	2	00
2	0.2	3	010
3	0.2	3	011
4	0.1	4	1000
5	0.1	4	1001

Then  $\mathbb{E}S = \sum_{i=1}^m p_i l_i = 2.8$ ,  $H = H/\log a = 2.12$ . [See also Carne p13.]



### 3 Huffman Coding

How to construct an optimal code? Take  $\mathcal{A} = \{\mu_1, \dots, \mu_m\}$ ,  $p_i = \mathbb{P}(X = \mu_i)$ . For simplicity take  $|\mathcal{B}| = a = 2$ . Without loss of generality  $p_1 \geq p_2 \geq \dots \geq p_m$ . Huffman gave an inductive definition of codes that we can prove are optimal. If  $m = 2$ , we take codewords 0, 1. If  $m > 2$ , first take the Huffman code for messages  $\mu_1, \dots, \mu_{m-2}, \nu$  with probabilities  $p_1, \dots, p_{m-2}, p_{m-1} + p_m$ . Then append 0 (respectively 1) to the codeword for  $\nu$  to give a codeword for  $\mu_{m-1}$  (respectively  $\mu_m$ ).

**Notes.**

- Huffman codes are prefix-free;
- Huffman codes are not unique: choice is needed if some of the  $p_i$  are equal.

**Example 3.1.** Revisit Example 2.5. We have

$i$	$p_i$	$c^{(1)}$	$p_i^{(2)}$	$c^{(2)}$	$p_i^{(3)}$	$c^{(3)}$	$p_i^{(4)}$	$c^{(4)}$
1	0.4	1	0.4	1	0.4	1	0.6	0
2	0.2	01	0.2	01	0.4	00	0.4	1
3	0.2	000	0.2	000	0.2	01		
4	0.1	0010	0.2	001				
5	0.1	0011						

**Theorem 3.1.** *Huffman codes are optimal (Huffman, 1952).*

*Proof.* We show by induction on  $m$  that Huffman codes of size  $m = |\mathcal{A}|$  are optimal.

$m = 2$ : codewords are 0, 1 - clearly optimal.

$m > 2$ : let  $c_m$  be a Huffman code for  $X_m$ , which takes values  $\mu_1, \dots, \mu_m$  with probabilities  $p_1 \geq p_2 \geq \dots \geq p_m$ ; each  $c_m$  is constructed from Huffman code  $c_{m-1}$  for  $X_{m-1}$  which takes values  $\mu_1, \dots, \mu_{m-2}, \nu$  with probabilities  $p_1, \dots, p_{m-2}, p_{m-1} + p_m$ . Then the expected word length is

$$\mathbb{E}S_m = \mathbb{E}S_{m-1} + p_{m-1} + p_m. \quad (*)$$

Let  $c'_m$  be an optimal code for  $X_m$ . Wlog  $c'_m$  is still prefix-free. Wlog the last two codewords of  $c'_m$  have maximal length and differ only in the final position (see next lemma). Say

$$c'_m(\mu_{m-1}) = y0, \quad c'_m(\mu_m) = y1 \text{ for some } y \in \{0, 1\}^*.$$

Let  $c'_{m-1}$  be some prefix-free code for  $X_{m-1}$ , given by

$$c'_{m-1}(\mu_i) = \begin{cases} c'_m(\mu_i) & 1 \leq i \leq m-2 \\ c'_{m-1}(\nu) = y & \end{cases}.$$

Then the expected word length satisfies

$$\mathbb{E}S'_m = \mathbb{E}S'_{m-1} + p_{m-1} + p_m. \quad (**)$$

By the inductive hypothesis,  $c_{m-1}$  is optimal, so  $\mathbb{E}S_{m-1} \leq \mathbb{E}S'_{m-1}$ . By (\*) and (\*\*) this implies  $\mathbb{E}S_m \leq \mathbb{E}S'_m$ .  $\square$

**Lemma 3.2.** Suppose letters  $\mu_1, \dots, \mu_m$  in  $\mathcal{A}$  are sent with probabilities  $p_1, p_2, \dots, p_m$ . Let  $c$  be an optimal (prefix-free) code with word lengths  $l_1, \dots, l_m$ . Then

- (i) If  $p_i > p + j$ , then  $l_i \leq l_j$ ;
- (ii) Amongst all codewords of maximal length there exist two that differ only in the final digit.

*Proof.* (i) is obvious. For (ii), could otherwise just delete the final digit of the codeword of maximal length (since prefix-free).  $\square$

**Remark.** Note not all optimal codes are Huffman (look at the case  $m = 4$ ).

Our main result says that if we have a prefix-free optimal code with word lengths  $l_1, \dots, l_m$  and associated probabilities  $p_1, \dots, p_m$ , then there is a Huffman code with these word lengths.

## 4 Joint Entropy

If  $X, Y$  are random variables with values in  $\mathcal{A}$  and  $\mathcal{B}$  respectively, then  $(X, Y)$  is a random variable with values in  $\mathcal{A} \times \mathcal{B}$ , and the *entropy*  $H(X, Y)$  is called the joint entropy, given by

$$H(X, Y) = - \sum_{x \in \mathcal{A}} \sum_{y \in \mathcal{B}} \mathbb{P}(X = x, Y = y) \log \mathbb{P}(X = x, Y = y).$$

This generalises to any finite number of random variables.

**Lemma 4.1.** Let  $X, Y$  be random variables taking values in  $\mathcal{A}$  and  $\mathcal{B}$  respectively. Then

$$H(X, Y) \leq H(X) + H(Y),$$

with equality if and only if  $X$  and  $Y$  are independent.

*Proof.* Write  $\mathcal{A} = \{x_1, \dots, x_m\}$ ,  $\mathcal{B} = \{y_1, \dots, y_n\}$ . Let

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j), \quad p_i = \mathbb{P}(X = x_i), \quad q_j = \mathbb{P}(Y = y_j).$$

Apply Gibbs' inequality to the probability distributions  $\{p_{ij}\}$  and  $\{p_i q_j\}$  to obtain

$$\begin{aligned} - \sum_{i,j} p_{ij} \log p_{ij} &\leq - \sum_{i,j} p_{ij} \log(p_i q_j) \\ &= - \sum_i \left( \sum_j p_{ij} \right) \log p_i - \sum_j \left( \sum_i p_{ij} \right) \log q_j \\ &= - \sum_i p_i \log p_i - \sum_j q_j \log q_j \\ &= H(X) + H(Y). \end{aligned}$$

With equality if and only if  $p_{ij} = p_i q_j$  for all  $i, j$ . □

## Error-correcting codes

### 5 Noisy channels and Hamming's code

**Definition.** A *binary*  $[m, n]$ -code is a subset  $C$  of  $\{0, 1\}^n$  of size  $m = |C|$ .  $n$  is the length of the code and the elements of  $C$  are called codewords.

We use an  $[n, m]$ -code to send one of  $m$  messages through a BSC (binary symmetric channel) making  $n$  uses of the channel. Clearly  $1 \leq m \leq 2^n$ , so  $0 \leq \frac{1}{n} \log m \leq 1$ .

**Definition.** For any  $x, y \in \{0, 1\}^n$  the *Hamming distance* is

$$d(x, y) = |\{i : 1 \leq i \leq n, x_i \neq y_i\}|.$$

**Definition.**

- (i) The *ideal observer* decoding rule decodes  $x \in \{0, 1\}^n$  as  $c \in C$  maximising  $\mathbb{P}(c \text{ sent} | x \text{ received})$ .
- (ii) The *maximum likelihood* decoding rule decodes  $x \in \{0, 1\}^n$  as  $c \in C$  maximising  $\mathbb{P}(x \text{ received} | c \text{ sent})$ .
- (iii) The *minimum distance* decoding rule decodes  $x \in \{0, 1\}^n$  as  $c \in C$  minimising  $d(x, C)$ .

**Lemma 5.1.**

- (a) If all the messages are equally likely, then (i) and (ii) above are equivalent.
- (b) If  $p < 1/2$  (error probability) then (ii) and (iii) are equivalent.

**Remark.** If  $p = 1/2$  the code is called *useless*. If  $p = 0$  the code is called *lossless*.

*Proof.*

- (a) We have

$$\mathbb{P}(c \text{ sent} | x \text{ received}) = \frac{\mathbb{P}(c \text{ sent}, x \text{ received})}{\mathbb{P}(x \text{ received})} = \frac{\mathbb{P}(c \text{ sent})}{\mathbb{P}(x \text{ received})} \mathbb{P}(x \text{ received} | c \text{ sent}).$$

So by hypothesis,  $\mathbb{P}(c \text{ sent})$  is independent of  $c \in C$ . So for fixed  $x$ , maximising  $\mathbb{P}(c \text{ sent} | x \text{ received})$  is the same as maximising  $\mathbb{P}(x \text{ received} | c \text{ sent})$ .

- (b) Let  $r = d(x, c)$ . Then  $\mathbb{P}(x \text{ received} | c \text{ sent}) = p^r (1-p)^{n-r} = (1-p)^n \left(\frac{p}{1-p}\right)^r$ . Since  $p < 1/2$ ,  $\frac{p}{1-p} < 1$ . So maximising  $\mathbb{P}(x \text{ received} | c \text{ sent})$  is the same as minimising  $r$ .

□

We choose to use minimum distance decoding from now on.

**Example 5.1.** Suppose 000, 111 are sent with probabilities  $\alpha = 9/10$ ,  $\beta = 1/10$  respectively through a BSC with error probability  $p = 1/4$ . Suppose 110 is received. Then

$$\mathbb{P}(000 \text{ sent} | 110 \text{ received}) = \frac{\alpha p^2(1-p)}{\alpha p^2(1-p) + (1-\alpha)p(1-p)^2} = \frac{3}{4},$$

$$\text{similarly } \mathbb{P}(111 \text{ sent} | 110 \text{ received}) = \frac{1}{4}.$$

So the ideal observer decodes as 000. But the maximum likelihood/minimum distance rules decode as 111.

**Remarks.**

- Minimum distance decoding may be expensive in terms of time and storage if  $|C|$  is large.
- Need to specify a convention in case there is no unique maximiser (e.g make a random choice, or request the message is sent again).

We aim to detect, or even correct errors.

**Definition.** A code  $C$  is

- *d-error* detecting if changing up to  $d$  digits in each codeword can never produce another codeword. In other words, each codeword is of Hamming distance greater than  $d$  from every other codeword.
- *e-error* correcting if knowing that  $x \in \{0,1\}^n$  differs from a codeword in at most  $e$  places we can deduce the codeword.

**Examples.**

- (a) A *repetition code* of length  $n$  has codewords  $\underbrace{00 \dots 0}_{n \text{ times}}, \underbrace{11 \dots 1}_{n \text{ times}}$ . This is a  $[n, 2]$ -code. It is  $(n-1)$ -error detecting and  $\lfloor \frac{n-1}{2} \rfloor$ -error correcting. But the information rate is only  $1/n$ .
- (b) A *simple parity check code* or *paper tape code*: identify  $\{0,1\}$  with  $\mathbb{F}_2$  and let  $C = \{(x_1, \dots, x_n) \in \{0,1\}^n : \sum_{i=1}^n x_i = 0\}$ . This is a  $[n, 2^{n-1}]$ -code, 1-error detecting but cannot correct errors. The information rate is  $\frac{n-1}{n}$ .
- (c) Hamming's original code (1950): a 1-error correcting binary  $[7, 16]$ -code. Take  $C \subseteq \mathbb{F}_2^7$  where

$$C = \{c \in \mathbb{F}_2^7 : c_1 + c_3 + c_5 + c_7 = 0, c_2 + c_3 + c_6 + c_7 = 0, c_4 + c_5 + c_6 + c_7 = 0\}.$$

The bits  $c_3, c_5, c_6, c_7$  are arbitrary and  $c_1, c_2, c_4$  are forced (called the check digits) so  $|C| = 2^4$ . To decode: suppose we receive  $x \in \mathbb{F}_2^7$ . We form the *syndrome*:  $z = z_x = (z_1, z_2, z_4) \in \mathbb{F}_2^3$  where

$$z_1 = x_1 + x_3 + x_5 + x_7$$

$$z_2 = x_2 + x_3 + x_6 + x_7$$

$$z_4 = x_4 + x_5 + x_6 + x_7.$$

If  $x \in C$ , then  $z_x = (0, 0, 0)$ . If  $d(x, c) = 1$  for some  $c \in C$ , then place where  $x$  and  $c$  differ is given by  $z_1 + 2z_2 + 4z_4$  (not mod 2). Check: if  $x = c + e_i$  where  $e_i$  has all 0's except a 1 in the  $i$ th position, then  $z_x = z_{e_i}$ , so check for each  $1 \leq i \leq 7$ .

**Lemma 5.2.** *The Hamming distance is a metric on  $\mathbb{F}_2^n$ .*

*Proof.* Trivial. □

**Definition.** The *minimum distance of a code* is the minimum of  $d(c_1, c_2)$  for all codewords  $c_1, c_2$  with  $c_1 \neq c_2$ .

**Lemma 5.3.** *Let  $C$  be a code with minimum distance  $d > 0$ . Then*

- (i)  *$C$  is  $(d - 1)$ -error detecting, but cannot detect all sets of  $d$  errors.*
- (ii)  *$C$  is  $\lfloor \frac{d-1}{2} \rfloor$ -error correcting, but cannot correct all sets of  $\lfloor \frac{d-1}{2} \rfloor + 1$  errors.*

*Proof.*

- (i) If  $x \in \mathbb{F}_2^n$  and  $c \in C$  are such that  $0 < d(x, c) \leq d - 1$ , then we know that  $x \notin C$  so this is  $(d - 1)$ -error detecting. However there must exist  $c_1, c_2 \in C$  such that  $d(c_1, c_2) = d$ , so we cannot say if there's an error if  $c_1$  is 'corrupted' to  $c_2$  in  $d$  errors.
- (ii) Take  $e = \lfloor \frac{d-1}{2} \rfloor$ . If  $x \in \mathbb{F}_2^n$  and  $c_1 \in C$  are such that  $d(x, c_1) \leq e$  then for any  $c_1 \neq c_2 \in C$  we have  $d(x, c_2) \geq d(c_1, c_2) - d(c_1, x) \geq d - e > e$ . So  $C$  is  $e$ -error correcting. Now take  $c_1, c_2 \in C$  with  $d(c_1, c_2) = d$ . Then take  $x \in \mathbb{F}_2^n$  such that  $x$  differs from  $c_1$  is precisely  $e + 1$  places where  $c_1$  and  $c_2$  differ. Then  $d(c_1, x) = e + 1$  and  $d(x, c_2) = d - (e + 1) \leq e + 1$ . So  $C$  cannot be  $(e + 1)$ -error correcting. □

**Definition.** A  $[n, m]$ -code with minimum distance  $d$  is called a  $[n, m, d]$ -code.

**Notes.**

- $m \leq 2^n$  with equality if and only if  $C = \mathbb{F}_2^n$  (trivial code)
- $d \leq n$ , with equality in case of the repetition code.

**Example 5.2.**

- (i) Repetition code of length  $n$  is a  $[n, 2, n]$ -code,  $(n - 1)$ -error detecting and  $\lfloor \frac{n-1}{2} \rfloor$ -error correcting.
- (ii) Simple parity check code is a  $[n, 2^{n-1}, 2]$ -code, 1-error detecting and 0-error correcting.
- (iii) Hamming's original code is 1-error correcting, implying  $d \geq 3$ . Also 0000000, 1110000 are distance 3 apart, so  $d = 3$ . So this is a  $[7, 16, 3]$ -code and is 2-error detecting.

## 6 Covering estimates

Take  $x \in \mathbb{F}_2^n$ ,  $r \geq 0$ . Then  $\overline{B}(x, r) = \{y \in \mathbb{F}_2^n : d(x, y) \leq r\}$  is the *closed Hamming ball*. Denote  $V(n, r) = |\overline{B}(x, r)| = \sum_{i=0}^r \binom{n}{i}$ , the *volume*.

**Lemma 6.1** (Hamming's bound). *An  $e$ -error correcting code  $C$  of length  $n$  has*

$$|C| \leq \frac{2^n}{V(n, e)}.$$

*Proof.*  $C$  is  $e$ -error correcting so  $\{B(c, e)\}_{c \in C}$  are pairwise disjoint balls, so  $\sum_{c \in C} |B(c, e)| = |C|V(n, e) \leq |\mathbb{F}_2^n| = 2^n$ .  $\square$

**Lemma 6.2.** *A code  $C$  of length  $n$  that can correct  $e$  errors is perfect if  $|C| = \frac{2^n}{V(n, e)}$ . Equivalently, for all  $x \in \mathbb{F}_2^n$  there exists a unique  $c \in C$  such that  $d(x, c) \leq e$ . In this case, any  $e + 1$  errors will make you decode incorrectly.*

**Example 6.1.**

(a) Hamming  $[7, 16, 3]$ -code is 1-error correcting and

$$\frac{2^n}{V(n, e)} = \frac{2^7}{V(7, 1)} = \frac{2^7}{1 + 7} = 2^4 = |C|.$$

(b) Binary repetition code of length  $n$  (for  $n$  odd) is perfect.

**Remark.** If  $\frac{2^n}{V(n, e)} \notin \mathbb{Z}$  then there does not exist a perfect  $e$ -error correcting code of length  $n$ . Converse is also false ( $n = 90, e = 2$  on Example Sheet 2).

**Definition.** Define  $A(n, d) = \max\{m : \exists [n, m, d]\text{-code}\}$ .

The  $A(n, d)$  are unknown in general. But we have some special cases:

**Examples.**

- $A(n, 1) = 2^n$  (trivial code)
- $A(n, n) = 2$  (repetition code)
- $A(n, 2) = 2^{n-1}$  (simple parity check code)

**Lemma 6.3.**  $A(n, d + 1) \leq A(n, d)$ .

*Proof.* Let  $m = A(n, d + 1)$  and take a  $[n, m, d + 1]$ -code  $C$ . Let  $c_1, c_2 \in C$  have  $d(c_1, c_2) = d + 1$ . Let  $c'_1$  differ from  $c_1$  in a single place where  $c_1$  and  $c_2$  differ. Hence  $d(c'_1, c_2) = d$ . If  $c \in C \setminus \{c_1\}$ , then  $d(c, c_1) \leq d(c, c'_1) + d(c'_1, c_1)$  so  $d + 1 \leq d(c, c'_1) + 1$ . Hence  $d(c, c'_1) \geq d$ . So replacing  $c_1$  with  $c'_1$ , we get an  $[n, m, d]$ -code.  $\square$

**Corollary 6.4.**  $A(n, d) = \max\{m : \exists [n, m, d']\text{-code for some } d' \geq d\}$ .

**Theorem 6.5.**

$$\frac{2^n}{V(n, d-1)} \underbrace{\leq}_{\text{GSV bound}} A(n, d) \underbrace{\leq}_{\text{Hamming bound}} \frac{2^n}{V(n, \lfloor \frac{d-1}{2} \rfloor)}.$$



*Proof.* We have already proved the Hamming bound. So let  $m = A(n, d)$ . Let  $C$  be a  $[n, m, d]$ -code. Then there does not exist  $x \in \mathbb{F}_2^n$  with  $d(x, c) \geq d$  for all  $c \in C$  (otherwise could replace  $C$  with  $C \cup \{x\}$ , contradicting maximality of  $m$ ). Hence

$$\mathbb{F}_2^n \subseteq \bigcup_{c \in C} \overline{B}(c, d-1) \implies 2^n \leq \sum_{c \in C} |\overline{B}(c, d-1)| = mV(n, d-1).$$

□

**Example 6.2.**  $n = 10, d = 3$ , have  $V(n, 1) = 11, V(n, 2) = 56$ . The above theorem gives  $19 \leq \frac{2^{10}}{56} \leq A(10, 3) \leq \frac{2^{10}}{11} \leq 93$ . It was known that  $72 \leq A(10, 3) \leq 93$ , but the exact value of  $A(10, 3)$  was only found in 1999.

### Asymptotics of $V(n, r)$

We study  $\frac{\log A(n, \lfloor n\delta \rfloor)}{n}$  as  $n \rightarrow \infty$  to see how large the information rate can be for a given error rate.

**Proposition 6.6.** *Let  $\delta \in (0, 1/2)$ . Then*

$$(i) \log V(n, \lfloor n\delta \rfloor) \leq nH(\delta);$$

$$(ii) \frac{1}{n} \log A(n, \lfloor n\delta \rfloor) \geq 1 - H(\delta).$$

Where  $H(\delta) = -\delta \log \delta - (1 - \delta) \log(1 - \delta)$ .

*Proof.* First we show (i)  $\Rightarrow$  (ii): by the GSV bound,

$$A(n, \lfloor n\delta \rfloor) \geq \frac{2^n}{V(n, \lfloor n\delta \rfloor - 1)} \geq \frac{2^n}{V(n, \lfloor n\delta \rfloor)}$$

and so

$$\frac{\log A(n, \lfloor n\delta \rfloor)}{n} \geq 1 - \frac{\log V(n, \lfloor n\delta \rfloor)}{n} \geq 1 - H(\delta).$$

Now we prove (i):  $H(\delta)$  is increasing for  $\delta < 1/2$ , so wlog we may assume  $n\delta \in \mathbb{Z}$ . Now

$$\begin{aligned} 1 &= (\delta + (1 - \delta))^n = \sum_{i=0}^n \binom{n}{i} \delta^i (1 - \delta)^{n-i} \geq \sum_{i=0}^{n\delta} \binom{n}{i} \delta^i (1 - \delta)^{n-i} \\ &= (1 - \delta)^n \sum_{i=0}^{n\delta} \binom{n}{i} \left( \frac{\delta}{1 - \delta} \right)^i \\ &\geq (1 - \delta)^n \sum_{i=0}^{n\delta} \binom{n}{i} \left( \frac{\delta}{1 - \delta} \right)^{n\delta} \\ &= \delta^{n\delta} (1 - \delta)^{n(1-\delta)} V(n, n\delta). \end{aligned}$$

Now taking logs:

$$0 \geq n(\delta \log \delta + (1 - \delta) \log(1 - \delta)) + \log V(n, n\delta).$$

□

The constant  $H(\delta)$  in the above bound best possible:

**Lemma 6.7.** *We have*

$$\lim_{n \rightarrow \infty} \frac{\log V(n, \lfloor n\delta \rfloor)}{n} = H(\delta).$$

*Proof.* Exercise.

□

## 7 Constructing new codes from old

We're given  $C$ , a  $[n, m, d]$ -code. Can check the details in the following:

**Examples.**

1. The *parity check extension*  $C^+$  is

$$\left\{ \left( c_1, \dots, c_n, \sum_{i=1}^n c_i \right) : (c_1, \dots, c_n) \in C \right\}$$

Is a  $[n+1, m, d']$ -code with  $d \leq d' \leq d+1$ , depending on whether  $d$  is odd or even.

2. Fix  $1 \leq i \leq n$ . Deleting the  $i$ th digit from each codeword gives the *punctured code*  $C^-$

$$\{(c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_n) : (c_1, \dots, c_n) \in C\}.$$

If  $d \geq 2$ , then it is a  $[n-1, m, d']$ -code with  $d-1 \leq d' \leq d$ .

3. Fix  $1 \leq i \leq n$ , and  $\alpha \in \mathbb{F}_2$ . The *shortened code*  $C'$  is

$$\{(c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_n) : (c_1, \dots, c_{i-1}, \alpha, c_{i+1}, \dots, c_n) \in C\}.$$

It has parameters  $[n, m', d']$  with  $d' \geq d$  and  $m' \geq \frac{m}{2}$  for a suitable choice of  $\alpha$ .

## Shannon's Theorems

### 8 AEP and Shannon's first coding theorem

**Definition.** A *source* is a sequence of random variables  $X_1, X_2, \dots$  taking values in some alphabet  $\mathcal{A}$ . A source is *Bernoulli (memoryless)* if  $X_1, X_2, \dots$  are iid: write  $(X, X_n)$ . A source  $X_1, X_2, \dots$  is *reliably encodable at rate  $r$*  if there exists a sequence of subsets  $(A_n)_{n \geq 1}$  with  $A_n \subseteq \mathcal{A}^n$  such that:

1.  $\lim_{n \rightarrow \infty} \frac{\log A_n}{n} = r;$
2.  $\lim_{n \rightarrow \infty} \mathbb{P}((X_1, \dots, X_n) \in A_n) = 1.$

The *information rate*  $H$  of a source is the infimum of all reliable encoding rates.  
Exercise:  $0 \leq H \leq \log |\mathcal{A}|$  with both bounds attainable.

Shannon's first coding theorem computes the information rate of certain sources, including Bernoulli sources.

**Reminders from IA Probability:**

We have a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . A discrete random variable  $X$  is a function  $X : \Omega \rightarrow \mathcal{A}$ . The probability mass function  $p_x : \mathcal{A} \rightarrow [0, 1]$  is defined by  $x \mapsto \mathbb{P}(X = x)$ . Can consider  $p \circ X = p(X) : \Omega \rightarrow [0, 1]$ , a random variable taking values in  $[0, 1]$ .

Given a source  $X_1, X_2, \dots$  of random variables with values in  $\mathcal{A}$ , the probability mass function of  $X^{(n)} = (X_1, \dots, X_n)$  is  $p_{X^{(n)}}$  given by  $(x_1, \dots, x_n) \mapsto \mathbb{P}((X_1, \dots, X_n) = (x_1, \dots, x_n))$ . Since  $p_{X^{(n)}} : \mathcal{A}^n \rightarrow [0, 1]$  and  $X^{(n)} : \Omega \rightarrow \mathcal{A}^n$ , you can form  $p(X^{(n)}) = p_{X^{(n)}} \circ X^{(n)} : \Omega \rightarrow [0, 1]$ .

**Example 8.1.** Let  $\mathcal{A} = \{A, B, C\}$ . Suppose

$$X^{(2)} = \begin{cases} AB & \text{with probability 0.3} \\ AC & \text{with probability 0.1} \\ BC & \text{with probability 0.1} \\ BA & \text{with probability 0.2} \\ CA & \text{with probability 0.25} \\ CB & \text{with probability 0.05} \end{cases}.$$

Then

$$p(X^{(2)}) = \begin{cases} 0.3 & \text{with probability 0.3} \\ 0.1 & \text{with probability 0.2} \\ 0.2 & \text{with probability 0.2} \\ 0.25 & \text{with probability 0.25} \\ 0.05 & \text{with probability 0.05} \end{cases}.$$

So some points are “lumped together”.

Given a source  $X_1, X_2, \dots$  *converges in probability* to a random variable  $L$  (possibly constant) if for all  $\varepsilon > 0$ ,  $\mathbb{P}(|X_n - L| > \varepsilon) \rightarrow 0$  as  $n \rightarrow \infty$ . We write  $X_n \xrightarrow{\mathbb{P}} L$ .

The *Weak Law of Large Numbers* (WLLN) says that if  $(X; X_n)$  are iid real-valued random variables with finite expectation  $\mathbb{E}X$ , we have

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\mathbb{P}} \mathbb{E}X.$$

**Example 8.2.** If  $X_1, X_2, \dots$  are iid Bernoulli, then  $p(X_1), p(X_2), \dots$  are iid random variables and  $p(X_1, \dots, X_n) = p(X_1) \dots p(X_n)$ . Note

$$-\frac{1}{n} \log p(X_1, \dots, X_n) = -\frac{1}{n} \sum_{i=1}^n \log p(X_i) \xrightarrow{\mathbb{P}} -\mathbb{E}(-\log p(X_1)) = H(X_1) \text{ as } n \rightarrow \infty.$$

**Lemma 8.1.** *The information rate of a Bernoulli source  $X_1, X_2, \dots$  is at most the expected word length of an optimal code  $c : \mathcal{A} \rightarrow \{0, 1\}^*$  for  $X_1$ .*

*Proof.* Let  $l_1, l_2, \dots$  be the lengths of codewords when we encode  $X_1, X_2, \dots$  using  $c$ . Let  $\varepsilon > 0$ . Set  $A_n = \{x \in \mathcal{A}^n : c^*(x) \text{ has length at less than } n(\mathbb{E}l_1 + \varepsilon)\}$ . Then

$$\begin{aligned} \mathbb{P}((X_1, \dots, X_n) \in A_n) &= \mathbb{P}\left(\sum_{i=1}^n l_i < n(\mathbb{E}l_1 + \varepsilon)\right) \\ &= \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n l_i - \mathbb{E}l_1\right| < \varepsilon\right) \\ &\rightarrow 1 \text{ as } n \rightarrow \infty. \end{aligned}$$

Now,  $c$  is decipherable so  $c^*$  is injective. Hence  $|A_n| \leq 2^{n(\mathbb{E}l_1 + \varepsilon)}$ . Making  $A_n$  larger if necessary,  $|A_n| = \lfloor e^{n(\mathbb{E}l_1 + \varepsilon)} \rfloor$  so

$$\frac{\log(A_n)}{n} \rightarrow \mathbb{E}l_1 + \varepsilon.$$

Hence  $X_1, X_2, \dots$  is reliably encodable at rate  $r = \mathbb{E}l_1 + \varepsilon$  for all  $\varepsilon > 0$ . Hence the information rate is at most  $\mathbb{E}l_1$ .  $\square$

**Corollary 8.2.** *A Bernouilli source has information rate less than  $H(X_1) + 1$ .*

*Proof.* Combine the above with the Noiseless Coding Theorem.  $\square$

We encode  $X_1, X_2, \dots$  in blocks

$$\underbrace{X_1, \dots, X_N}_{Y_1}, \underbrace{X_{n+1}, \dots, X_{2N}, \dots}_{Y_2}, \dots$$

so  $Y_1, Y_2, \dots$  take values in  $\mathcal{A}^N$ . Exercise: show that if  $X_1, X_2, \dots$  has information rate  $H$  then  $Y_1, Y_2, \dots$  has information rate  $NH$ .

**Proposition 8.3.** *The information rate  $H$  of a Bernouilli source  $X_1, X_2, \dots$  is at most  $H(X_1)$ .*

*Proof.* Apply the previous corollary to  $Y_1, Y_2, \dots$  and obtain

$$NH < H(Y_1) + 1 = H(X_1, \dots, X_N) + 1 = \sum_{i=1}^N H(X_i) + 1 = NH(X_1, \dots, X_n) + 1.$$

Hence  $H < H(X_1) + \frac{1}{N}$ . Since  $N$  is arbitrary,  $H \leq H(X_1)$ .  $\square$

**Definition.** A source  $X_1, X_2, \dots$  satisfies the *Asymptotic Equipartition Property* (AEP) for some constant  $H \geq 0$  if

$$-\frac{1}{n} \log p(X_1, X_2, \dots) \xrightarrow{\mathbb{P}} H \text{ as } n \rightarrow \infty.$$

**Example 8.3.** Tossing a biased coin,  $\mathbb{P}(H) = p$ . Let  $(X; X_n)$  be the results of independent coin tosses. After a large number  $N$  of tosses, expect on average  $pN$  heads and  $(1-p)N$  tails. The probability of any particular sequence of  $pN$  heads and  $(1-p)N$  tails is  $p^{pN}(1-p)^{(1-p)N} = 2^{N(p \log p + (1-p) \log (1-p))} = 2^{-NH(X)}$ . Not every sequence of tosses will be like this, but there is only a small probability of “atypical” sequences. With high probability we get a “typical” sequence and its probability will be close to  $2^{-NH(X)}$ .

**Lemma 8.4.** *The AEP for a source  $X_1, X_2, \dots$  is equivalent to the following property*

$\forall \varepsilon > 0 \exists n_0(\varepsilon)$  such that  $\forall n \geq n_0(\varepsilon) \exists$  a “typical set”  $T_n \subseteq \mathcal{A}^n$  such that

(i)  $\mathbb{P}((X_1, \dots, X_n) \in T_n) > 1 - \varepsilon$ ;

(ii)  $2^{-n(H+\varepsilon)} \leq p(x_1, \dots, x_n) \leq 2^{-n(H-\varepsilon)}$  for all  $(x_1, \dots, x_n) \in T_n$ .

*Proof.* Obvious and non-examinable. □

**Theorem 8.5** (Shannon's First Coding Theorem (FCT)). *If a source  $X_1, X_2, \dots$  satisfies the AEP with constant  $H$ , then the source has information rate  $H$ .*

*Proof.* Let  $\varepsilon > 0$  and let  $T_n \subseteq \mathcal{A}^n$  be typical sets. Then for some  $n_0(\varepsilon)$  and all  $n \geq n_0(\varepsilon)$

$$p(x_1, \dots, x_n) \geq 2^{-n(H+\varepsilon)} \text{ for all } (x_1, \dots, x_n) \in T_n$$

$$\implies \mathbb{P}(T_n) \geq 2^{-n(H+\varepsilon)} |T_n| \implies 1 \geq 2^{-n(H+\varepsilon)} |T_n| \implies \frac{\log |T_n|}{n} \leq H + \varepsilon.$$

Taking  $A_n = T_n$ , shows the source is reliably encodable at rate  $H + \varepsilon$ . Conversely, if  $H = 0$ , we're done. Otherwise pick  $0 < \varepsilon < H/2$  and suppose for contradiction that the source is reliably encodable at rate  $H - 2\varepsilon$ , say with sets  $A_n \subseteq \mathcal{A}^n$ . Let  $T_n \subseteq \mathcal{A}^n$  be typical sets. Then for all  $(x_1, \dots, x_n) \in T_n$

$$p(x_1, \dots, x_n) \leq 2^{-n(H-\varepsilon)}$$

$$\implies \mathbb{P}(A_n \cap T_n) \leq 2^{-n(H-\varepsilon)} |A_n|$$

$$\implies \frac{\log \mathbb{P}(A_n \cap T_n)}{n} \leq H - \varepsilon + \frac{\log |A_n|}{n} \rightarrow -(H - \varepsilon) + H - 2\varepsilon = -\varepsilon.$$

Hence  $\log \mathbb{P}(A_n \cap T_n) \rightarrow -\infty$  and  $\mathbb{P}(A_n \cap T_n) \rightarrow 0$ . But  $\mathbb{P}(T_n) \leq \mathbb{P}(A_n \cap T_n) + \mathbb{P}(\mathcal{A}^n \setminus A_n) \rightarrow 0$ , a contradiction to  $\mathbb{P}(T_n) \rightarrow 1$ . Thus the information rate is exactly  $H$ .  $\square$

**Corollary 8.6.** *A Bernoulli source  $X_1, X_2, \dots$  has information rate  $H(X_1)$ .*

*Proof.* We've already seen that  $-\frac{1}{n} \log p(X_1, \dots, X_n) \xrightarrow{\mathbb{P}} H(X_1)$ , so done by Shannon's First Coding Theorem.  $\square$

**Remarks.**

- The AEP is useful for noiseless coding. We can
  - encode the typical sequences using a block code;
  - encode the atypical sequences arbitrarily.
- Many sources, which are not necessarily Bernoulli satisfy the AEP. Under suitable hypotheses the sequence  $\frac{1}{n} H(X_1, \dots, X_n)$  is decreasing and the AEP is satisfied.

## 9 Capacity & Shannon's second coding theorem

Recall:

**Definition.** We model  $n$  uses of a channel by the  $n$ th extension, with input alphabet  $\mathcal{A}^n$  and output alphabet  $\mathcal{B}^n$ . A code  $C$  of length  $n$  is a function  $\mathcal{M} \rightarrow \mathcal{A}^n$  where  $\mathcal{M}$  is the set of possible messages. Implicitly we also have a decoding rule  $\mathcal{B}^n \rightarrow \mathcal{M}$ . The size of  $C$  is  $m = |\mathcal{M}|$ . The information rate is  $\rho(C) = \frac{1}{n} \log_2 m$ . The error rate is  $\hat{e}(C) = \max_{x \in \mathcal{M}} \mathbb{P}(\text{error} | x \text{ sent})$ .

**Definition.** A channel can *transmit reliably at rate  $R$*  if there exists  $(C_n)_{n=1}^\infty$  with each  $C_n$  a code of length  $n$  such that

$$\lim_{n \rightarrow \infty} \rho(C_n) = R \text{ \& } \lim_{n \rightarrow \infty} \hat{e}(C_n) = 0.$$

The *capacity* is the supremum of all reliable transmission rates.

Suppose we are given a source where

- it has information rate  $r$  bits per second;
- it emits symbols at  $s$  symbols per second.

Suppose we are also given a channel where

- it has capacity  $R$  bits per transmission;
- it transmits symbols at  $S$  transmissions per second.

Usually, information theorists take  $S = s = 1$ . If  $rs \leq RS$  then you can encode and transmit reliably, and if  $rs > RS$  you cannot.

We'll compute the capacity of a BSC with error probability  $p$ .

**Proposition 9.1.** *A binary symmetric channel with error probability  $p < 1/4$  has non-zero capacity.*

*Proof.* We use the GSV bound. Pick  $\delta \in (2p, 1/2)$ . We claim reliable transmission at rate  $R = 1 - H(\delta) > 0$ . Let  $C_n$  be a code of length  $n$ , and suppose it has minimum distance  $\lfloor n\delta \rfloor$  of maximal size. Then (by Proposition 6.6(ii))

$$|C_n| = A(n, \lfloor n\delta \rfloor) \geq 2^{n(1-H(\delta))}.$$

Replacing  $C_n$  by a subcode, we can assume  $|C_n| = \lfloor 2^{nR} \rfloor$  and still minimum distance  $\geq \lfloor n\delta \rfloor$ . Using minimum distance decoding

$$\begin{aligned} \hat{e}(C_n) &\leq \mathbb{P} \left( \text{in } n \text{ uses, BSC makes } \geq \left\lfloor \frac{\lfloor n\delta \rfloor - 1}{2} \right\rfloor \text{ errors} \right) \\ &\leq \mathbb{P} \left( \text{in } n \text{ uses, BSC makes } \geq \left\lfloor \frac{n\delta - 1}{2} \right\rfloor \text{ errors} \right). \end{aligned}$$

Pick  $\varepsilon > 0$  with  $p + \varepsilon < \frac{\delta}{2}$ . For  $n$  sufficiently large,  $\frac{n\delta - 1}{2} = n \left( \frac{\delta}{2} - \frac{1}{2n} \right) > n(p + \varepsilon)$ . Hence  $\hat{e}(C_n) \leq \mathbb{P}(\text{BSC makes } \geq n(p + \varepsilon) \text{ errors}) \rightarrow 0$ , using the next lemma.  $\square$



**Lemma 9.2.** Let  $\varepsilon > 0$ . A BSC with error probability  $p$  is used to transmit  $n$  digits. Then

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{BSC makes } \geq n(p + \varepsilon) \text{ errors}) = 0.$$

*Proof.* Consider random variables  $U_i$  taking value 1 if the  $i$ th digit is mistransmitted, and value 0 otherwise. Then  $\mathbb{E}U_i = p$ , so  $\mathbb{P}(\text{BSC makes } \geq n(p + \varepsilon) \text{ errors}) \leq \mathbb{P}(|\frac{1}{n} \sum_{i=1}^n U_i - p| \geq \varepsilon) \rightarrow 0$ .  $\square$

**Definition** (Conditional Entropy). Let  $X, Y$  be random variables taking values in alphabets  $\mathcal{A}, \mathcal{B}$  respectively. Then we define the *conditional entropy*

$$H(X|Y = y) = - \sum_{x \in \mathcal{A}} \mathbb{P}(X = x|Y = y) \log \mathbb{P}(X = x|Y = y)$$

$$H(X|Y) = \sum_{y \in \mathcal{B}} \mathbb{P}(Y = y) H(X|Y = y).$$

Clearly  $H(X|Y) \geq 0$ .

**Lemma 9.3.** We have

$$H(X, Y) = H(X|Y) + H(Y).$$

*Proof.*

$$\begin{aligned} H(X|Y) &= - \sum_{y \in \mathcal{B}} \sum_{x \in \mathcal{A}} \mathbb{P}(X = x|Y = y) \mathbb{P}(Y = y) \log \mathbb{P}(X = x|Y = y) \\ &= - \sum_{y \in \mathcal{B}} \sum_{x \in \mathcal{A}} \mathbb{P}(X = x, Y = y) \log \left( \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \right) \\ &= - \sum_{y \in \mathcal{B}} \sum_{x \in \mathcal{A}} \mathbb{P}(X = x, Y = y) \log \mathbb{P}(X = x, Y = y) \\ &\quad + \sum_{y \in \mathcal{B}} \sum_{x \in \mathcal{A}} \mathbb{P}(X = x, Y = y) \log \mathbb{P}(Y = y) \\ &= H(X, Y) - \sum_{y \in \mathcal{B}} \mathbb{P}(Y = y) \log \mathbb{P}(Y = y) \\ &= H(X, Y) - H(Y). \end{aligned}$$

$\square$

**Example 9.1.** We roll fair die. Let  $X$  be the value defined by the roll, let  $Y$  be equal to 0 if  $X$  is even and 1 if  $X$  is odd. Then  $H(X, Y) = H(X) = \log 6$ ,  $H(Y) = \log 2 = 1$ ,  $H(X|Y) = \log 3 = H(X, Y) - H(Y)$ . Similarly  $H(Y|X) = 0 = H(X, Y) - H(X)$ .

**Corollary 9.4.**  $H(X|Y) \leq H(X)$ .

*Proof.* Combine the previous with Lemma 4.1.  $\square$

Now replace  $X, Y$  with random vectors  $X^{(r)} = (X_1, \dots, X_r), Y^{(s)} = (Y_1, \dots, Y_s)$ . Similarly can define  $H(X_1, \dots, X_r | Y_1, \dots, Y_s) = H(X^{(r)} | Y^{(s)})$ .

**Remark.**  $H(X, Y | Z)$  is the entropy of  $(X, Y)$  given  $Z$ , not the entropy of  $X$  and  $Y | Z$ .

**Lemma 9.5.** *Let  $X, Y, Z$  be random variables. Then  $H(X, Y) \leq H(X | Y, Z) + H(Z)$ .*

*Proof.* We expand  $H(X, Y, Z)$  in two different ways.

$$H(X, Y, Z) = H(Z | X, Y) + H(X | Y) + H(Y),$$

$$H(X, Y, Z) = H(X | Y, Z) + H(Z | Y) + H(Y).$$

Since  $H(Z | X, Y) \geq 0$ , we have

$$\begin{aligned} H(X | Y) &\leq H(X | Y, Z) + H(Z | Y) \\ &\leq H(X | Y, Z) + H(Z). \end{aligned}$$

□

**Proposition 9.6** (Fano's inequality). *Let  $X, Y$  be random variables taking values in  $\mathcal{A}$ ,  $|\mathcal{A}| = m$ . Let  $p = \mathbb{P}(X \neq Y)$ . Then*

$$H(X | Y) \leq H(p) + p \log(m - 1).$$

*Proof.* Define  $Z = 0$  if  $X = Y$  and  $Z = 1$  if  $X \neq Y$ . Then  $\mathbb{P}(Z = 0) = 1 - p$ ,  $\mathbb{P}(Z = 1) = p$ . So  $H(Z) = H(p)$ . By the previous lemma,

$$H(X | Y) \leq H(p) + H(X | Y, Z). \quad (*)$$

Since  $Z = 0$  implies  $X = Y$ ,  $H(X | Y = y, Z = 0) = 0$ . Also there are  $m - 1$  remaining possibilities for  $X$ , so  $H(X | Y = y, Z = 1) \leq \log(m - 1)$ . Therefore

$$\begin{aligned} H(X | Y) &= \sum_{y, z} \mathbb{P}(Y = y, Z = z) H(X | Y = y, Z = z) \\ &\leq \sum_{y \in \mathcal{A}} \mathbb{P}(Y = y, Z = 1) \log(m - 1) \\ &= \mathbb{P}(Z = 1) \log(m - 1) \\ &= p \log(m - 1). \end{aligned}$$

So by  $(*)$ ,  $H(X | Y) \leq H(p) + p \log(m - 1)$ . □

**Remark.**  $H(p)$  represents the information needed to decide whether or not there is an error, and  $p \log(m - 1)$  represents the information needed to resolve the error assuming the worst possible case.

**Definition.** Let  $X, Y$  be random variables. Then the *mutual information* is  $I(X; Y) := H(X) - H(X | Y)$ . By Corollary 9.4,  $I(X; Y) \geq 0$  with equality if and only if  $X, Y$  are independent. Clearly  $I(X; Y) = I(Y; X)$ .

Suppose we're given a discrete memoryless channel (DMC) with input alphabet  $\mathcal{A}$ ,  $|\mathcal{A}| = m$  and output alphabet  $\mathcal{B}$ . Let  $X$  be a random variable taking values in  $\mathcal{A}$  and used as input to the channel. Let  $Y$  be a random variable output, depending on  $X$  and the channel matrix.

**Definition.** The (*information*) *capacity* is  $\max_X I(X; Y)$ .

**Remarks.**

- The maximum is over all probability distributions  $(p_1, \dots, p_m)$  for  $X$  on  $\mathcal{A}$ .
- The maximum is attained since  $(p_1, \dots, p_m) \mapsto I((p_1, \dots, p_m); Y)$  is continuous on the compact set

$$\{(p_1, \dots, p_m) \in [0, 1]^m : p_1 + \dots + p_m = 1\}.$$

- The information capacity depends only on the channel matrix.

**Theorem 9.7** (Shannon's Second Coding Theorem). *For a DMC, the operational capacity is equal to the information capacity.*

**Remark.** We'll show one inequality in general and the other for the BSC only.

Assuming Shannon's Second Coding Theorem, let us compute the capacity of certain channels.

**Example 9.2.** BSC, error probability  $p$ . Input  $X$ :  $\mathbb{P}(X = 0) = \alpha$ ,  $\mathbb{P}(X = 1) = 1 - \alpha$ . Output  $Y$ :  $\mathbb{P}(Y = 0) = \alpha(1 - p) + (1 - \alpha)p$ ,  $\mathbb{P}(Y = 1) = (1 - \alpha)(1 - p) + \alpha p$ . Then  $C$  is

$$\begin{aligned} \max_{\alpha} I(X; Y) &= \max_{\alpha} (H(Y) - H(Y|X)) \\ &= \max_{\alpha} (H(\alpha(1 - p) + (1 - \alpha)p) - H(p)) \\ &= 1 - H(p). \end{aligned}$$

Where the maximum is attained for  $\alpha = 1/2$ . Hence  $C = 1 + p \log p + (1 - p) \log(1 - p)$ .

**Remark.** We can choose to calculate either  $H(Y) - H(Y|X)$  or  $H(X) - H(X|Y)$  depending on which is easier.

**Example 9.3.** BEC, erasure probability  $p$ . Input  $X$ :  $\mathbb{P}(X = 0) = \alpha$ ,  $\mathbb{P}(X = 1) = 1 - \alpha$ . Output  $Y$ :  $\mathbb{P}(Y = 0) = \alpha(1 - p)$ ,  $\mathbb{P}(Y = *) = p$ ,  $\mathbb{P}(Y = 1) = (1 - \alpha)(1 - p)$ . Can calculate  $H(X|Y = 0) = 0$ ,  $H(X|Y = 1) = 0$ ,  $H(X|Y = *) = H(\alpha)$ , which gives  $H(X|Y) = pH(\alpha)$ . So

$$\begin{aligned} C = \max_{\alpha} I(X; Y) &= \max_{\alpha} (H(X) - H(X|Y)) = \max_{\alpha} (H(\alpha) - pH(\alpha)) \\ &= (1 - p) \max_{\alpha} H(\alpha) \\ &= 1 - p. \end{aligned}$$

Where the max is attained for  $\alpha = 1/2$ .

Now model using a channel  $n$  times as the  $n$ th extension, i.e replace alphabets  $\mathcal{A}, \mathcal{B}$  with  $\mathcal{A}^n, \mathcal{B}^n$ ,

$$\mathbb{P}(y_1, \dots, y_n \text{ received} | x_1, \dots, x_n \text{ sent}) = \prod_{i=1}^n \mathbb{P}(y_i | x_i).$$

**Lemma 9.8.** The  $n$ th extension of a DMC with information capacity  $C$  has information capacity  $nC$ .

*Proof.* Take random variable input  $X_1, \dots, X_n$  producing output  $Y_1, \dots, Y_n$ . Since the channel is memoryless,

$$H(Y_1, \dots, Y_n | X_1, \dots, X_n) = \sum_{i=1}^n H(Y_i | X_1, \dots, X_n) = \sum_{i=1}^n H(Y_i | X_i).$$

Hence

$$\begin{aligned}
 I(X_1, \dots, X_n; Y_1, \dots, Y_n) &= H(Y_1, \dots, Y_n) - H(Y_1, \dots, Y_n | X_1, \dots, X_n) \\
 &= H(Y_1, \dots, Y_n) - \sum_{i=1}^n H(Y_i | X_i) \\
 &\leq \sum_{i=1}^n (H(Y_i) - H(Y_i | X_i)) \\
 &= \sum_{i=1}^n I(X_i; Y_i) \leq nC.
 \end{aligned}$$

To finish, we find a distribution for  $X_1, \dots, X_n$  giving equality. Equality is attained by taking  $X_1, \dots, X_n$  independent, each of the same distribution such that  $I(X_i; Y_i) = C$ . Indeed, if  $X_1, \dots, X_n$  are independent, so are  $Y_1, \dots, Y_n$ .  $\square$

**Proposition 9.9.** *For a DMC the (operational) capacity is at most the information capacity. Let  $C$  be the information capacity. Suppose reliable transmission is possible at some rate  $R > C$ , i.e there exist a sequence of codes  $(C_n)_{n \geq 1}$  with  $C_n$  having length  $n$  and size  $\lceil 2^{nR} \rceil$  for all  $n$  such that*

$$\lim_{n \rightarrow \infty} \rho(C_n) = R \text{ and } \lim_{n \rightarrow \infty} \hat{e}(C_n) = 0.$$

(Recall  $\hat{e}(C_n) = \max_{c \in C_n} \mathbb{P}(\text{error} | c \text{ sent})$ .)

*Proof.* Define the average error rate  $e(C_n) = \frac{1}{|C_n|} \sum_{c \in C_n} \mathbb{P}(\text{error} | c \text{ sent})$ . Note  $e(C_n) \leq \hat{e}(C_n)$  and so  $e(C_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Take input random variable  $X$  equidistributed over  $C_n$  (i.e takes all values with equal probability). Let  $Y$  be the random variable output when  $X$  is transmitted and decoded. So  $e(C_n) = \mathbb{P}(X \neq Y)$ , define  $p := e(C_n)$ . Now we have

$$H(X) = \log |C_n| = \log \lceil 2^{nR} \rceil \geq nR - 1 \text{ for sufficiently large } n.$$

And

$$\begin{aligned}
 H(X|Y) &\leq H(p) + p \log(|C_n| - 1) && \text{(Fano's inequality)} \\
 &\leq 1 + pnR.
 \end{aligned}$$

Recall  $I(X; Y) = H(X) - H(X|Y)$ . By the previous lemma  $nC \geq I(X; Y)$  so

$$\begin{aligned}
 nC &\geq nR - 1 - (1 + pnR) \\
 \implies pnR &\geq n(R - C) - 2 \implies p \geq \frac{n(R - C) - 2}{nR} \rightarrow \frac{R - C}{R} \neq 0.
 \end{aligned}$$

Since  $R > C$ , which contradicts  $p \rightarrow 0$ . Hence we conclude that we cannot transmit reliably at any rate exceeding  $C$ .  $\square$

To complete the proof of Shannon's SCT for a BSC with error probability  $p$ , we need to show that the operational capacity is at least  $1 - H(p)$  (i.e the information capacity).

**Proposition 9.10.** *Consider BSC with error probability  $p$ . Suppose  $R < 1 - H(p)$ . Then there is a sequence of codes  $(C_n)_{n \geq 1}$  with  $C_n$  of length  $n$  and size  $\lceil 2^{nR} \rceil$  for all  $n$  such that*

$$\lim_{n \rightarrow \infty} \rho(C_n) = R \text{ and } \lim_{n \rightarrow \infty} e(C_n) = 0.$$

**Remark.** Note that the above proposition deals with the average error rate  $e$ , not the error rate  $\hat{e}$ .

*Proof.* We'll use the method of 'random coding'. Wlog  $p < 1/2$ . Take  $\varepsilon > 0$  such that

$$p + \varepsilon < 1/2 \text{ and } R < 1 - H(p + \varepsilon).$$

(This is possible by continuity of  $H$ ). We use minimum distance decoding (in case of a tie, make arbitrary choice). Let  $m = \lceil 2^{nR} \rceil$ . Pick  $C = \{c_1, \dots, c_m\}$  at random from  $\mathcal{C} = \{[n, m]\text{-codes}\}$ , a set of size  $\binom{2^n}{m}$ . Choose  $1 \leq i \leq m$  at random (i.e each with probability  $1/m$ ). We sent  $c_i$  through the channel and get output  $Y$ . Then

$$\begin{aligned} \mathbb{P}(Y \text{ not decoded as } c_i) &= \text{average value of } e(C) \text{ as } C \text{ runs through } \mathcal{C} \\ &= \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} e(C). \end{aligned}$$

We can pick  $C_n$ , an  $[n, m]$ -code with  $e(C_n)$  at most this average. So it is enough to show

$$\mathbb{P}(Y \text{ is not decoded as } c_i) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Let  $r = \lfloor n(p + \varepsilon) \rfloor$ . Then if  $B(Y, r) \cap C = \{c_i\}$ ,  $Y$  is correctly decoded as  $c_i$ . So

$$\mathbb{P}(Y \text{ is not decoded as } c_i) \leq \mathbb{P}(c_i \notin B(Y, r)) + \mathbb{P}(B(Y, r) \cap C \supsetneq \{c_i\}).$$

Consider the two cases separately, according as  $d(c_i, Y) > r$  (case (i)) or  $d(c_i, Y) \leq r$  (case (ii)). For (i) we have

$$\begin{aligned} \mathbb{P}(d(c_i, Y) > r) &= \mathbb{P}(\text{BSC makes } > r \text{ errors}) \\ &= \mathbb{P}(\text{BSC makes } > n(p + \varepsilon) \text{ errors}) \\ &\rightarrow 0 \text{ as } n \rightarrow \infty \text{ by a previous theorem.} \end{aligned}$$

For (ii), if  $j \neq i$ ,  $\mathbb{P}(c_j \in B(Y, r) | c_i \in B(Y, r)) = \frac{V(n, r) - 1}{2^n - 1} \leq \frac{V(n, r)}{2^n}$  (since

$V(n, r) \leq 2^n$ . Hence

$$\begin{aligned}
 \mathbb{P}(B(Y, r) \cap C \supsetneq \{c_i\}) &\leq \sum_{j \neq i} \mathbb{P}(c_j \in B(Y, r) \text{ and } c_i \in B(Y, r)) \\
 &\leq \sum_{j \neq i} \mathbb{P}(c_j \in B(Y, r) | c_i \in B(Y, r)) \\
 &\leq (m-1) \frac{V(n, r)}{2^n} \\
 &\leq \frac{mV(n, r)}{2^n} \\
 &\leq 2^{nR} 2^{nH(p+\varepsilon)2^{-n}} \\
 &= 2^{n(R-(1-H(p+\varepsilon)))} \xrightarrow{n \rightarrow \infty} 0
 \end{aligned}$$

since  $R < 1 - H(p + \varepsilon)$ .  $\square$

**Proposition 9.11.** *We can replace  $e$  by  $\hat{e}$  in the previous result.*

*Proof.* Pick  $R'$  with  $R < R' < 1 - H(p)$ . The previous result constructs a sequence  $(C'_n)_{n \geq 1}$  of codes,  $C'_n$  of length  $n$  and size  $\lfloor 2^{nR'} \rfloor$  for all  $n$ , and  $e(C'_n) \rightarrow 0$  as  $n \rightarrow \infty$ .

For each  $n$ , order the codewords of  $C'_n$  by  $\mathbb{P}(\text{error} | c \text{ sent})$  and delete the worst half, to get a code  $C_n$  with  $\hat{e}(C_n) \leq 2e(C'_n)$ . So  $\hat{e}(C_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Since  $C_n$  has length  $n$  and size  $\lfloor 2^{nR'} \rfloor / 2 = \text{floor } 2^{nR'-1}$ . But  $2^{nR'-1} = 2^{n(R'-1/n)} \geq 2^{nR}$  for all  $n$  sufficiently large. So can replace  $C'_n$  by a code of smaller size  $\lfloor 2^{nR} \rfloor$  and still get  $\hat{e}(C_n) \rightarrow 0$  and  $\rho(C_n) \rightarrow R$ .  $\square$

**Remarks.**

1. A BSC with error probability  $p$  has operational capacity  $1 - H(p)$ . Indeed the above proposition says we can transmit reliably at any rate  $R < 1 - H(p)$ , whereas 9.9 says the capacity is at most  $1 - H(p)$ .
2. Shannon's SCT shows that good codes exist, but the proof does not say how to construct them.

## 10 The Kelly criterion

Let  $1 > p > 0$ ,  $u > 0$ ,  $1 > w \geq 0$ . Suppose a coin is tossed  $n$  times in succession,  $\mathbb{P}(H) = p$ . If I pay  $k$  ahead of a particular throw, then get back  $ku$  if the throw is heads, or get back 0 if throw is tails.

Say we have initial bankroll  $X_0 = 1$ . If we have bankroll of  $X_n$  after  $n$ th toss, bet  $wX_n$ , retaining  $(1 - w)X_n$ . Then bankroll  $X_{n+1}$  after  $(n + 1)$ th throw is

$$X_{n+1} = \begin{cases} X_n(wu + (1 - w)) & \text{if } n\text{th toss is heads} \\ X_n(1 - w) & \text{if } n\text{th toss is tails} \end{cases}$$

Let  $Y_{n+1} = X_{n+1}/X_n$ . Then  $Y_1, Y_2, \dots$ , is a sequence of iid random variables. Hence  $\log Y_1, \log Y_2, \dots$  is a sequence of iid random variables. Note  $\log X_n = \sum_{i=1}^n \log Y_i$ .

**Lemma 10.1.** *Let  $\mathbb{E} \log Y_1 = \mu$ ,  $\text{Var}(\log Y_1) = \sigma^2$ . Then if  $a > 0$*

$$(i) \quad \mathbb{P} \left( \left| \frac{1}{n} \sum_{i=1}^n \log Y_i - \mu \right| \geq a \right) \leq \frac{\sigma^2}{na^2}$$

$$(ii) \quad \mathbb{P} \left( \left| \frac{\log X_n}{n} - \mu \right| \geq a \right) \leq \frac{\sigma^2}{na^2}$$

$$(iii) \quad \text{Given } \varepsilon > 0 \text{ and } \delta > 0 \text{ there exists } N \in \mathbb{N} \text{ such that } \mathbb{P} \left( \left| \frac{\log X_n}{n} - \mu \right| \geq \delta \right) \leq \varepsilon \text{ for all } n \geq N.$$

*Proof.* Obvious by Chebyshev.  $\square$

**Lemma 10.2.** *Consider a single toss of a coin,  $\mathbb{P}(H) = p < 1$ . Suppose a bet of  $k$  on  $H$  has payout ratio  $u > 0$ . Suppose we have initial bankroll 1, and we bet  $w$  on  $H$ , retaining  $1 - w$  ( $0 \leq w < 1$ ). If  $Y =$  fortune after toss then  $\mathbb{E} \log Y = p \log(1 + (u - 1)w) + (1 - p) \log(1 - w)$ . The value of  $\mathbb{E} \log Y$  is maximised by taking  $w = 0$  if  $up \leq 1$ , and otherwise taking  $w = \frac{up-1}{u-1}$ .*

*Proof.* Not given (see Körner's book).  $\square$

**Remark.** Put  $q = 1 - p$ . If  $up > 1$ , at the optimum  $w$  have

$$\mathbb{E} \log Y = p \log p + q \log q + \log u - q \log(u - 1)$$

and  $-(p \log p + q \log q)$  is the entropy of a simple probabilistic system.

The suggestion that those making a long sequence of bets should maximise the expectation of the log is called Kelly's criterion.

**Exercise:** show that if you bet less than the optimum, your bankroll will tend to increase, but more slowly; and if you bet more than some critical proportion, your bankroll will tend to decrease.

## 11 Algebraic Coding Theory

### 11.1 Linear codes

**Definition.** A code  $C \subseteq \mathbb{F}_2^n$  is *linear* if

- $0 \in C$ ;
- whenever  $x, y \in C$ , we have  $x + y \in C$ .

i.e  $C$  is a  $\mathbb{F}_2$ -vector subspace of  $\mathbb{F}_2^n$ .

**Definition.** The *rank* of a linear code  $C$  is its dimension as a  $\mathbb{F}_2$ -vector space. A linear code of length  $n$ , rank  $k$  is called a  $(n, k)$ -code. If the minimum distance of the code is  $d$ , it is a  $(n, k, d)$ -code.



Let  $v_1, \dots, v_k$  be a basis of  $C$ . Then  $C = \{\sum_{i=1}^k \lambda_i v_i : \lambda_1, \dots, \lambda_k \in \mathbb{F}_2\}$ . So  $|C| = 2^k$ . So an  $(n, k)$ -code is a  $[n, 2^k]$ -code. The information rate is  $k/n$ .

**Definition.** The *weight* of  $x \in \mathbb{F}_2^n$  is  $w(x) = d(x, 0)$ .

**Lemma 11.1.** *The minimum distance of a linear code is the minimum weight of a non-zero code word.*

*Proof.* Obvious. □

**Definition.** For  $x, y \in \mathbb{F}_2^n$ , let  $x \cdot y = \sum_{i=1}^n x_i y_i \in \mathbb{F}_2$ . This operation is symmetric and bilinear, but not positive definite.

**Definition.** Take some  $P \subseteq \mathbb{F}_2^n$ . The *parity check code* defined by  $P$  is

$$C = \{x \in \mathbb{F}_2^n : p \cdot x = 0, \forall p \in P\}.$$

**Examples.**

- $P = \{(1, \dots, 1)\}$  gives the simple parity check code;
- $P = \{(1, 0, 1, 0, 1, 0, 1), (0, 1, 1, 0, 0, 1, 1), (0, 0, 0, 1, 1, 1, 1)\}$  gives Hamming's  $[7, 4, 3]$ -code;
- Show that  $C^+$  and  $C^-$  are linear if  $C$  is linear. When is  $C'$  linear?

**Lemma 11.2.** *Every parity check code is linear.*

*Proof.* Obvious. □

**Definition.** Take  $C \subseteq \mathbb{F}_2^n$  a linear code. The *dual code*

$$C^\perp = \{x \in \mathbb{F}_2^n : x \cdot y = 0, \forall y \in C\}.$$

This is a parity-check code, so is linear. Note  $C \cap C^\perp$  is not necessarily  $\{0\}$ .

**Lemma 11.3.**  $\dim(C) + \dim(C^\perp) = n$ .

*Proof.* See later. □

**Corollary 11.4.** *Let  $C$  be linear. Then  $(C^\perp)^\perp = C$ . In particular,  $C$  is a parity check code.*

*Proof.* Clearly  $C \subseteq (C^\perp)^\perp$ . Equality follows from the previous lemma. □

**Definition.** Let  $C$  be a  $(n, k)$ -code.

- (i) A *generator matrix*  $G$  for  $C$  is a  $k \times n$  matrix with rows a basis of  $C$ ;
- (ii) A *parity check matrix*  $H$  for  $C$  is a generator matrix for  $C^\perp$ . It is a  $(n - k) \times n$  matrix.

The codewords of  $C$  can be viewed either as:

- Linear combinations of rows of  $G$ ;
- Linear dependence relations between the columns of  $H$ , i.e  $C = \{x \in \mathbb{F}_2^n : Hx = 0\}$ .

## Syndrome decoding

**Definition.** Let  $C$  be a  $(n, k)$ -code. The *syndrome* of  $x \in \mathbb{F}_2^n$  is  $Hx$ . If we receive  $x = c + z$ , where  $c$  is a codeword and  $z$  is the ‘error pattern’, the syndrome  $Hx = H(c + z) = Hc + Hz = Hz$ . If  $C$  is  $e$ -error correcting, we precompute a list of all possible  $Hx$  for all  $z$  with  $w(z) \leq e$ . On receiving  $x \in \mathbb{F}_2^n$  we look for  $Hx$  in our list. If successful,  $Hx = Hz$  so  $H(x - z) = 0$  and we decode  $x$  as  $c = x - z \in C$  with  $d(x, c) = w(z) \leq e$ .

**Definition.** We say codes  $C_1, C_2 \subseteq \mathbb{F}_2^n$  are *equivalent* if reordering each codeword of  $C_1$  using the same permutation gives the codewords of  $C_2$ . i.e  $\sigma C_1 = C_2$  for some  $\sigma \in S_n$ . Usually we only consider codes up to equivalence.

**Lemma 11.5.** Every  $(n, k)$ -linear code is equivalent to one with generator  $G = (I_k | B)$  for some  $k \times (n - k)$  matrix  $B$ .

*Proof.* Given a  $k \times n$  generator matrix  $G$  for  $C$ , using row operations, we can transform  $G$  into row-echelon form, i.e

$$G_{ij} = \begin{cases} 0 & j < l(i) \\ 1 & j = l(i) \end{cases}$$

for some  $l(1) < l(2) < \dots < l(k)$ . Permuting the columns replaces code by an equivalent code, so wlog we can assume  $l(i) = i$  for all  $1 \leq i \leq k$  which gives the required form upon applying further row operations.  $\square$

**Remark.** A message  $y \in \mathbb{F}_2^k$ , viewed as a row vector is encoded as  $yG$ . So if  $G = (I_k | B)$ , then  $yG = (y | yB)$ , where  $y$  is the message and  $yB$  are check digits.

Now we will prove

**Lemma 11.6.**  $\dim(C) + \dim(C^\perp) = n$ .

*Proof.* Wlog  $C$  has generator matrix  $G = (I_k | B)$ .  $G$  has  $k$  linearly independent columns so the linear map  $\gamma : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^k$  defined by  $x \mapsto Gx$  is surjective and  $\text{Ker}(\gamma) = C^\perp$ . By the rank-nullity theorem,  $\dim \mathbb{F}_2^n = \dim \text{Ker}(\gamma) + \dim \text{Im}(\gamma) = n$ , i.e  $n = \dim(C) + \dim(C^\perp)$ .  $\square$

**Lemma 11.7.** An  $(n, k)$ -linear code with generator matrix  $G = (I_k, B)$  has parity check matrix  $H = (B^T | I_{n-k})$ .

*Proof.*  $GH^T = (I_k | B) \begin{pmatrix} B \\ I_{n-k} \end{pmatrix} = B + B = 0$ . So the rows of  $H$  generate a subcode of  $C^\perp$ . But  $\text{rank}(H) = n - k$  and  $n - k = \dim C^\perp$ . Hence  $C^\perp$  has generator matrix  $H$ , and  $H$  is a parity check matrix for  $C$ .  $\square$

**Lemma 11.8.** Let  $C$  be a linear code with parity check matrix  $H$ . Then  $d(C) = d$  if and only if

- (i) any  $(d - 1)$  columns of  $H$  are linearly independent;
- (ii) some set of  $d$  columns of  $H$  are linearly dependent.

*Proof.* Exercise.  $\square$

## 12 Examples of codes

### Hamming codes

**Definition.** For  $d \geq 1$ , let  $n = 2^d - 1$ . Let  $H$  be the  $d \times n$  matrix whose columns are the non-zero elements of  $\mathbb{F}_2^d$ . The *Hamming*  $(n, n-d)$ -linear code is the code with parity matrix  $H$  (original code is  $d = 3$ ).

**Lemma 12.1.** *The Hamming  $(n, n-d)$ -code  $C$  has minimum distance  $d(C) = 3$ , is perfect and is 1-error correcting.*

*Proof.* Any two columns of  $H$  are linearly independent, but there are 3 columns that are linearly dependent, hence  $d(C) = 3$  by the previous lemma. By a previous result,  $C$  is  $\lfloor \frac{3-1}{2} \rfloor = 1$ -error correcting.

Note that  $n = 2^d - 1$ ,  $e = 1$  so  $V(n, e) = 1 + 2^d - 1 = 2^d$  and  $|C| = 2^{n-d} = \frac{2^n}{2^d} = \frac{2^n}{V(n, e)}$  and  $C$  is perfect.  $\square$

### Reed-Muller codes

Take  $X = \{p_1, \dots, p_n\}$ , set of size  $n$ . There is a correspondence between  $\mathcal{P}(X)$  and  $\mathbb{F}_2^n$  via  $\mathcal{P}(X) \rightarrow \{f : X \rightarrow \mathbb{F}_2\} \rightarrow \mathbb{F}_2^n$  the composition of  $A \mapsto \mathbb{1}_A$  and  $f \mapsto (f(p_1), \dots, f(p_n))$ . The symmetric difference then corresponds to vector addition, and the intersection  $A \cap B$  to the wedge product  $x \wedge y = (x_1 y_1, \dots, x_n y_n)$ .

Take  $X = \mathbb{F}_2^d$ , so  $n = 2^d = |X|$ . Let  $v_0 = (1, 1, 1, \dots, 1)$ . Let  $v_i = \mathbb{1}_{H_i}$  where  $H_i = \{p \in X : p_i = 0\}$  for  $1 \leq i \leq d$ .

**Definition.** Let  $0 \leq r \leq d$ . The *Reed-Muller code*  $\text{RM}(d, r)$  of order  $r$  and length  $2^d$  is the linear code spanned by  $v_0$  and all wedge products of  $r$  or fewer of the  $v_i$ . By convention, the empty wedge product is  $v_0$ .

**Example 12.1.**  $d = 3$ ,  $X = \mathbb{F}_2^3 = \{p_1, \dots, p_8\}$  in binary order.

$X$	000	001	010	011	100	101	110	111
$v_0$	1	1	1	1	1	1	1	1
$v_1$	1	1	1	1	0	0	0	0
$v_2$	1	1	0	0	1	1	0	0
$v_3$	1	0	1	0	1	0	1	0
$v_1 \wedge v_2$	1	1	0	0	0	0	0	0
$v_2 \wedge v_3$	1	0	0	0	1	0	0	0
$v_1 \wedge v_3$	1	0	1	0	0	0	0	0
$v_1 \wedge v_2 \wedge v_3$	1	0	0	0	0	0	0	0

**Examples.**

- $\text{RM}(3, 0)$  - span of  $v_0$ , a repetition code of length 8;
- $\text{RM}(3, 1)$  - span of  $v_0, v_1, v_2, v_3$ , equivalent to a parity check extension of Hamming's  $(7, 4)$ -code;
- $\text{RM}(3, 2)$  - an  $(8, 7)$ -code, actually simple parity check code of length 8;
- $\text{RM}(3, 3) = \mathbb{F}_2^8$ , trivial code.

**Theorem 12.2.**

(i) The vectors  $v_{i_1} \wedge \dots \wedge v_{i_s}$  where  $1 \leq i_1 < i_2 < \dots, i_s \leq d$  and  $0 \leq s \leq d$  form a basis for  $\mathbb{F}_2^n$ ;

(ii)  $\text{rank}(\text{RM}(d, r)) = \sum_{s=0}^r \binom{d}{s}$ .

*Proof.* In (i) we've listed  $\sum_{s=0}^d \binom{d}{s} = (1+1)^d = 2^d = n$  vectors, so it is enough to show spanning, i.e.  $\text{RM}(d, d) = \mathbb{F}_2^n$ . Let  $p \in X$ . Put

$$y_i = \begin{cases} v_i & p_i = 0 \\ v_0 + v_i & p_i = 1 \end{cases}.$$

Then  $\mathbb{1}_{\{p\}} = y_1 \wedge \dots \wedge y_d$ . Expand this using the distributive law to see that  $\mathbb{1}_{\{p\}} \in \text{RM}(d, d)$ . Since  $\{\mathbb{1}_{\{p\}}\}_{p \in X}$  clearly spans, we are done.

For (ii),  $\text{RM}(d, r)$  is spanned by  $v_{i_1} \wedge \dots \wedge v_{i_s}$  for  $1 \leq i_1 < \dots < i_s \leq d$  with  $0 \leq s \leq r$ . These vectors are linearly independent by (i), so a basis. Hence  $\text{rank}(\text{RM}(d, r)) = \sum_{s=0}^r \binom{d}{s}$ .  $\square$

**Definition.** Let  $C_1, C_2$  be linear codes of length  $n$  with  $C_2 \subseteq C_1$ . The *bar product* is  $C_1 | C_2 = \{(x | x+y) : x \in C_1, y \in C_2\}$ . It is a linear code of length  $2n$ .

**Lemma 12.3.**

(i)  $\text{rank}(C_1 | C_2) = \text{rank}(C_1) + \text{rank}(C_2)$ ;

(ii)  $d(C_1 | C_2) = \min\{2d(C_1), d(C_2)\}$ .

*Proof.* (i) If  $C_1$  has basis  $x_1, \dots, x_k$ ,  $C_2$  has basis  $y_1, \dots, y_l$ , then  $C_1 | C_2$  has basis  $\{(x_i | x_i) : 1 \leq i \leq k\} \cup \{(0 | y_i) : 1 \leq i \leq l\}$ . So  $\text{rank}(C_1 | C_2) = k + l = \text{rank}(C_1) + \text{rank}(C_2)$ ;

(ii) Let  $0 \neq (x | x+y) \in C_1 | C_2$ . If  $y \neq 0$  then  $w(x | x+y) = w(x) + w(x+y) \geq w(y) \geq d(C_2)$ . If  $y = 0$  then  $w(x | x+y) = 2w(x) \geq 2d(C_1)$ . So  $d(C_1 | C_2) \geq \min\{2d(C_1), d(C_2)\}$ . But there exists  $0 \neq x \in C_1$  with  $w(x) = d(C_1)$ , so  $d(C_1 | C_2) \leq w(x | x) = 2d(C_1)$ . There exists  $0 \neq y \in C_2$  with  $w(y) = d(C_2)$ , so  $d(C_1 | C_2) \leq w(0 | y) = w(y) = d(C_2)$ .  $\square$

**Theorem 12.4.**

- (i)  $\text{RM}(d, r) = \text{RM}(d-1, r) \mid \text{RM}(r-1, r-1)$ ;
- (ii)  $\text{RM}(d, r)$  has minimum distance  $2^{d-r}$ .

*Proof.*

- (i) Exercise, noting  $\text{RM}(d-1, r-1) \subseteq \text{RM}(d-1, r)$  so the bar product is defined;
- (ii) If  $r = 0$  then  $\text{RM}(d, 0)$  is a repetition code of length  $n = 2^d$ , having minimum distance  $2^{d-0}$ . If  $r = d$ , then  $\text{RM}(d, d) = \mathbb{F}_2^n$  having minimum distance  $1 = 2^{d-d}$ . We prove the  $0 < r < d$  case by induction on  $d$ . Recall from (i) that  $\text{RM}(d, r) = \text{RM}(d-1, r) \mid \text{RM}(d-1, r-1)$ . By induction, the minimum distance of  $\text{RM}(d-1, r)$  is  $2^{d-1-r}$  and the minimum distance of  $\text{RM}(d-1, r-1)$  is  $2^{d-1-r}$ . Hence the minimum distance of  $\text{RM}(d, r)$  is  $\min\{2 \cdot 2^{d-1-r}, 2^{d-r}\} = 2^{d-r}$ .

□

## 13 Cyclic codes

### GRM Recap

**Fact.** Let  $F$  be a field. The rings  $\mathbb{Z}$  and  $F[X]$  both have a division algorithm: if  $a, b \in \mathbb{Z}$ ,  $b \neq 0$ , there exist  $q, r \in \mathbb{Z}$  such that  $a = qb + r$  and  $0 \leq r < |b|$ . If  $f, g \in F[X]$ ,  $g \neq 0$ , there exist  $q, r \in F[X]$  such that  $f = qg + r$  with  $\deg r < \deg g$ .

Recall ideals  $I$ : subsets of a ring  $R$  such that whenever  $r \in R$ ,  $x \in I$  we have  $rx \in I$  (e.g.  $2\mathbb{Z} \trianglelefteq \mathbb{Z}$ ). Principle ideals generated by  $r \in R$  are  $(x) = \{rx : r \in R\}$ . By the division algorithm, every ideal in  $\mathbb{Z}$  or  $F[X]$  is principle.

If  $I \trianglelefteq R$  then the set of cosets  $R/I = \{x + I : x \in R\}$  is a ring, called the *quotient ring*. In particular, we identify  $\mathbb{Z}/n\mathbb{Z}$  and  $\{0, 1, \dots, n-1\}$  and reduce modulo  $n$  after each operation.

Similarly,  $F[X]/(f(X)) \leftrightarrow \left\{ \sum_{i=0}^{n-1} a_i X^i : a_i \in F \right\} \leftrightarrow F^n$ , reducing ‘modulo  $f$ ’ after each operation using the division algorithm.

**Definition.** A linear code  $C \subseteq \mathbb{F}_2^n$  is *cyclic* if

$$(a_0, a_1, \dots, a_{n-1}) \in C \implies (a_{n-1}, a_0, a_1, \dots, a_{n-2}) \in C.$$

Identify  $\mathbb{F}_2[X]/(X^n - 1)$  with  $\mathbb{F}_2^n$  via

$$a_0 + a_1X + \dots + a_{n-1}X^{n-1} + (X^n - 1) \leftrightarrow^\pi (a_0, \dots, a_{n-1}).$$

**Lemma 13.1.** A code  $C \subseteq \mathbb{F}_2^n$  is cyclic if and only if  $\pi(C) = \mathcal{C}$  satisfies

- (i)  $0 \in \mathcal{C}$ ;
- (ii)  $f, g \in \mathcal{C} \Rightarrow f + g \in \mathcal{C}$ ;
- (iii)  $f \in \mathbb{F}_2[X], g \in \mathcal{C} \Rightarrow fg \in \mathcal{C}$ .

In other words,  $\mathcal{C}$  is an ideal of  $\mathbb{F}_2[X]/(X^n - 1)$ .

*Proof.* If  $g(X) = a_0 + a_1X + \dots + a_{n-1}X^{n-1} + (X^n - 1)$  then  $Xg(X) = a_{n-1} + a_0X + \dots + a_{n-2}X^{n-1} + (X^n - 1)$ . So  $\mathcal{C}$  is cyclic iff (i) and (ii) hold (linearity) and  $g(X) \in \mathcal{C} \Rightarrow Xg(X) \in \mathcal{C}$ . Note that closure under multiplication by  $X$  implies closure under multiplication by any  $f(X)$ , so we have the result.  $\square$

From now on, we identify  $\pi(C)$  and  $\mathcal{C}$ .

Basic problem:

$$\begin{aligned} \text{cyclic codes of length } n &\leftrightarrow \text{ideals in } \mathbb{F}_2[X]/(X^n - 1) \\ &\leftrightarrow \text{ideals in } \mathbb{F}_2[X] \text{ containing } (X^n - 1) \\ &\leftrightarrow \text{polynomials } g(X) \in \mathbb{F}_2[X] \text{ with } g(X) \mid (X^n - 1) \end{aligned}$$

**Theorem 13.2.** Let  $C \leq \mathbb{F}_2[X]/(X^n - 1)$  be a cyclic code. Then there exists a unique polynomial  $g(X)$  such that

- (i)  $C = \{f(X)g(X) + (X^n - 1) : f(X) \in \mathbb{F}_2[X]\} = (g)$ ;
- (ii)  $g(X) \mid X^n - 1$ .

In particular,  $p(X) \in \mathbb{F}_2[X]$  represents a codeword  $g(X) \mid p(X)$ . We say  $g(X)$  is the generator polynomial of  $C$ .

*Proof.*

- (i) Let  $g(X) \in \mathbb{F}_2[X]$  be of least degree representing a non-zero codeword. Now  $\deg g \leq n - 1$ . Since  $C$  is cyclic, we have  $(g) \subseteq C$ .

Let  $p(X) \in \mathbb{F}_2[X]$  represent a codeword. By the division algorithm,  $p(X) = q(X)g(X) + r(X)$  where  $\deg r < \deg g$ . So  $r(X) = p(X) - q(X)g(X) \in C$ , contradicting minimality of  $\deg g$  unless  $r = 0$ . Hence  $g(X) \mid p(X)$  and so  $C \subseteq (g)$ , i.e.  $C = (g)$ .

(ii) Take  $p(X) = X^n - 1$ , giving (ii).

Uniqueness: suppose  $g_1(X), g_2(X)$  both satisfy (i) and (ii). Then  $g_1(X) \mid g_2(X)$  and  $g_2(X) \mid g_1(X)$ , i.e.  $g_1(X), g_2(X)$  are equal up to a unit. But the only unit in  $\mathbb{F}_2[X]$  is 1, so  $g_1(X) = g_2(X)$ .  $\square$

**Lemma 13.3.** *Let  $C$  be a cyclic code of length  $n$  with generator polynomial  $g(X) = a_0 + a_1X + \dots + a_kX^k$  ( $a_k \neq 0$ ). Then  $C$  has a basis given by  $g(X), Xg(X), \dots, X^{n-k-1}g(X)$ . In particular,  $\text{rank}(C) = n - k$ .*

*Proof.* Trivial check.  $\square$

**Corollary 13.4.** *Given generator polynomial of degree  $k$  as above, a generator matrix for  $C$  is given by*

$$G = \begin{pmatrix} a_0 & a_1 & \dots & a_k & 0 & \dots & 0 \\ 0 & a_0 & \dots & a_{k-1} & a_k & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & a_0 & a_1 & \dots & a_k \end{pmatrix}$$

a  $(n - k) \times n$  matrix.

**Corollary 13.5.** *If  $g(X)h(X) = X^n - 1$  where  $g(X)$  is the generator polynomial, then writing  $h(X) = b_0 + b_1X + \dots + b_{n-k}X^{n-k}$  ( $b_{n-k} \neq 0$ ),  $h$  is called the parity check polynomial and the parity check matrix is*

$$H = \begin{pmatrix} b_{n-k} & b_{n-k-1} & \dots & b_0 & 0 & \dots & 0 \\ 0 & b_{n-k} & \dots & b_1 & b_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & b_{n-k} & b_{n-k-1} & \dots & b_0 \end{pmatrix}$$

a  $k \times n$  matrix.

*Proof.* Check that the “dot product” of the  $i$ th row of  $G$  and the  $j$ th row of  $H$  is the coefficient of  $X^{(n-k-i)+j}$  in  $g(X)h(X) = X^n - 1$ . But  $1 \leq i \leq n - k$  and  $1 \leq j \leq k$  so  $0 < (n - k - i) + j < n$  so all these coefficients are 0. Hence the row of  $G$  and  $H$  are “orthogonal”. Also as  $b_{n-k} \neq 0$ ,  $\text{rank}(H) = k = \text{rank}(C^\perp)$ , so  $H$  is the parity check matrix.  $\square$

**Remarks.** The cyclic code given by  $h(X)$  is the dual code  $C^\perp$  with coefficients ‘reversed’. (Given a polynomial  $f(X) = \sum_{i=0}^m f_iX^i$  of degree  $m$  in  $\mathbb{F}_2[X]$ , the reverse polynomial is  $\check{f}(X) = \sum_{i=0}^m f_{m-i}X^i = X^{\deg f} f(1/X)$ .)

**Lemma 13.6.** *If  $n$  is odd, then  $X^n - 1 = f_1(X) \dots f_t(X)$  with  $f_i(X)$  distinct irreducibles in  $\mathbb{F}_2[X]$  (false if  $n$  is even, e.g.  $X^2 - 1 = (X - 1)^2$ ). Thus there exist  $2^t$  cyclic codes of length  $n$ .*

*Proof.* Not given (see II Galois Theory).  $\square$

## 14 BCH codes

These are a certain type of cyclic code.

**Fact (A).** Suppose  $p$  prime,  $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$  is a field. If  $f(X) \in \mathbb{F}_p[X]$  is irreducible, then  $K = \mathbb{F}_p[X]/(f)$  is a field of order  $p^{\deg f}$ .

**Fact (B).** Let  $q = p^\alpha$  for  $p$  prime,  $\alpha \in \mathbb{N}$ . Then there exists a field  $\mathbb{F}_q$  of order  $q$  and it is unique up to isomorphism.

**Fact (C).** The multiplicative group  $\mathbb{F}_q^\times = \mathbb{F}_q \setminus \{0\}$  is cyclic. A generator of this group is called a *primitive element*.



Notation:  $n$  an odd integer. Pick  $r \geq 1$  such that  $2^r \equiv 1 \pmod{n}$  (exists since  $(n, 2) = 1$ ). Let  $K = \mathbb{F}_{2^r}$  and let  $\mu_n(K) = \{x \in K : x^n = 1\} \leq K^\times$ . Then  $\mu_n(K)$  is a cyclic group of order  $n$  (as  $K^\times$  is cyclic and  $n \mid (2^r - 1)$ ). So  $\mu_n(K)$  is generated by some  $\alpha \in K$ , called a *primitive  $n$ th root of unity*.

**Definition.** The *cyclic code of length  $n$  with defining set  $A \subseteq \mu_n(K)$*  is

$$C = \{f(x) + (X^n - 1) : f(x) \in \mathbb{F}_2[X], f(a) = 0 \forall a \in A\}.$$

The *generator polynomial*  $g(x)$  is the non-zero polynomial of least degree such that  $g(a) = 0$  for all  $a \in A$ . Equivalently, the lcm of the minimal polynomials of the elements  $a \in A$ .

**Definition.** The cyclic code with defining set  $\{\alpha, \alpha^2, \dots, \alpha^{\delta-1}\}$  for  $\alpha \in \mu_n(K)$  is called a BCH (Bose, Ray-Chandhari, Hocquenghem) with *design distance*  $\delta$ .

**Theorem 14.1.** A BCH code  $C$  with design distance  $\delta$  has  $d(C) \geq \delta$ .

We first prove

**Lemma 14.2** (Vandermonde).

$$\det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \dots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq j < i \leq n} (x_i - x_j).$$

*Proof.* Exercise. □

*Proof of theorem.* Consider

$$\begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{\delta-1} & \alpha^{2(\delta-1)} & \dots & \alpha^{(\delta-1)(n-1)} \end{pmatrix}.$$

By the previous lemma, any  $\delta - 1$  columns of  $H$  are linearly independent. But any codeword of  $C$  is a dependence relation between the columns of  $H$ . Hence any non-zero codeword has weight  $\geq \delta$ , and  $d(C) \geq \delta$ . □

**Remark.**  $H$  in the above proof is *not* a parity-check matrix: its entries are not in  $\mathbb{F}_2$ .

## Decoding BCH codes

Let  $C$  be a cyclic code with defining set  $\{\alpha, \alpha^2, \dots, \alpha^{\delta-1}\}$  where  $\alpha \in K$  is a primitive  $n$ th root of unity. By the above theorem, we should be able to correct  $t := \lfloor \frac{\delta-1}{2} \rfloor$  errors.

We send  $c \in C$  and receive  $r = c + e$ , where  $e$  is the error pattern with  $\leq t$  non-zero errors. Note

$$\mathbb{F}_2^n \leftrightarrow \mathbb{F}_2[X]/(X^n - 1)$$

$$r, c, e \leftrightarrow r(X), c(X), e(X).$$

Note also  $c(\alpha^j) = 0$  for each  $1 \leq j \leq \delta - 1$ . Hence  $r(\alpha^j) = e(\alpha^j)$ .

**Definition.** The *error locator polynomial* is  $\sigma(X) = \prod_{i \in \mathcal{E}} (1 - \alpha^i X) \in K[X]$  where  $\mathcal{E} = \{i : e_i = 1, 0 \leq i \leq n - 1\}$ .

Problem: assuming  $\deg \sigma = |\mathcal{E}|$ , where  $2t + 1 \leq \delta$ , recover  $\sigma(X)$  from  $r(X)$ .

**Theorem 14.3.** Assume  $\deg \sigma = |\mathcal{E}| \leq t$ . Then  $\sigma(X)$  is the unique polynomial in  $K[X]$  of least degree such that

$$(i) \quad \sigma(0) = 1;$$

$$(ii) \quad \sigma(X) \sum_{j=1}^{2t} r(\alpha^j) X^j \equiv \omega(X) \pmod{X^{2t+1}} \text{ for some } \omega(X) \in K[X] \text{ of degree at most } t.$$

*Proof.* Define  $\omega(X) = -X\sigma'(X)$ , the error co-locator. Then

$$\omega(X) = \sum_{i \in \mathcal{E}} \alpha^i X \prod_{j \in \mathcal{E} \setminus \{i\}} (1 - \alpha^j X).$$

Work in the ring of formal power series  $K[[X]] = \{\sum_{i=0}^{\infty} \beta_i X^i : \beta_i \in K\}$ . Then

$$\begin{aligned} \frac{\omega(X)}{\sigma(X)} &= \sum_{i \in \mathcal{E}} \frac{\alpha^i X}{1 - \alpha^i X} = \sum_{i \in \mathcal{E}} \sum_{j=1}^{\infty} (\alpha^i X)^j \\ &= \sum_{j=1}^{\infty} \left( \sum_{i \in \mathcal{E}} (\alpha^i)^j \right) X^j \\ &= \sum_{j=1}^{\infty} e(\alpha^j) X^j \end{aligned}$$

so  $\sigma(X) \sum_{j=1}^{\infty} e(\alpha^j) X^j = \omega(X)$ . By the definition of  $c$ ,  $c(\alpha^j) = 0$  for all  $1 \leq j \leq \delta - 1$  and so for all  $1 \leq j \leq 2t$ . As  $r = c + e$ ,  $r(\alpha^j) = e(\alpha^j)$  for all  $1 \leq j \leq 2t$ . Hence

$$\sigma(X) \sum_{j=1}^{2t} r(\alpha^j) X^j \equiv \omega(X) \pmod{X^{2t+1}}.$$

This verifies (i),(ii) with  $\omega(X) = -X\sigma'(X)$ , so  $\deg \omega = \deg \sigma = |\mathcal{E}| \leq 2t$ .

Now we show uniqueness: suppose we have  $\tilde{\sigma}, \tilde{\omega} \in K[X]$  which also satisfy (i) and (ii). Without loss of generality  $\deg \tilde{\sigma} \leq \deg \sigma$ . Now,  $\sigma(X)$  has distinct non-zero roots, so  $\sigma(X)$  and  $\omega(X) = -X\sigma'(X)$  are coprime. By (ii)  $\tilde{\sigma}\omega(X) = \sigma(X)\tilde{\omega}(X) \pmod{X^{2t+1}}$ , so  $\sigma(X)\tilde{\omega}(X) = \tilde{\sigma}(X)\omega(X)$  since  $\sigma, \tilde{\sigma}, \omega, \tilde{\omega}$  all have degree  $\leq t$ . But since  $\sigma, \omega$  are coprime we must have  $\sigma(X) \mid \tilde{\sigma}(X)$ . Since  $\deg \tilde{\sigma} \leq \deg \sigma$ , we have  $\tilde{\sigma}(X) = \lambda\sigma(X)$  for some  $\lambda \in K$ . But (i) says  $\lambda = 1$  so  $\sigma = \tilde{\sigma}$ .  $\square$

## Decoding BCH codes

We receive  $r(X)$

- Compute  $\sum_{j=1}^{2t} r(\alpha^j)X^k$ ;
- Set  $\sigma(X) = 1 + \sigma_1X + \dots + \sigma_tX^t$  and compute coefficients of  $X^i$  for  $t+1 \leq i \leq 2t$  to obtain linear equations for  $\sigma_1, \dots, \sigma_t$  (i.e.  $\sum_{j=0}^t \sigma_j r(\alpha^{i-j}) = 0$ );
- Solve these over  $K$ , keeping solutions of least degree;
- Compute  $\mathcal{E} = \{0 \leq i \leq n-1 : \sigma(\alpha^{-i}) = 0\}$  and check  $|\mathcal{E}| = \deg \sigma$ ;
- Set  $e(X) = \sum_{i \in \mathcal{E}} X^i$ . Then  $c(X) = r(X) + e(X)$ .