ENGR 76 Practice Problems

Siddharth Chandak

Winter 2026

Contents

Week 1

Source Coding Basics

1. Fixed Length Codes

Recall that a fixed length code is a code where each codeword has the same length k. Suppose we want to encode the English alphabet $\mathcal{X} = \{\mathrm{A}, \mathrm{B}, \ldots, \mathrm{Z}\}$ using a fixed length code.

(a). Recall that M denotes the total number of symbols in the alphabet. $M = ?$

Solution: There are 26 letters in the English alphabet, hence $M = 26$.

(b). What is the minimum possible k such that each symbol has a distinct codeword?

Solution: We need $2^k \geq M$, so $k \geq \log_2(M)$ and the smallest integer is

$$k = \lceil \log_2(M) \rceil = \lceil \log_2(26) \rceil = 5.$$

If we only had $k = 4$ bits, then there are only $2^4 = 16$ distinct codewords, which is less than 26.

2. Prefix-Free Codes

Consider the following prefix-free code for an alphabet of size 5:

Symbol	Codeword
A	0
B	10
C	110
D	1110
E	1111

(a). Encode the following message using the code above:

BADDECADE

Solution: Using the given code:

$$\begin{aligned}
B &\to 10 \\
A &\to 0 \\
D &\to 1110 \\
D &\to 1110 \\
E &\to 1111 \\
C &\to 110 \\
A &\to 0 \\
D &\to 1110 \\
E &\to 1111
\end{aligned}$$

Concatenating the codewords, we obtain:

$$\boxed{\texttt{1001110111011111100111011111}}$$

(b). Decode the following received bit stream into the corresponding sequence of symbols:

$$\texttt{1100101011111110}$$

Solution: Decoding from left to right:

$$\begin{aligned} 110 &\to C \\ 0 &\to A \\ 10 &\to B \\ 10 &\to B \\ 1111 &\to E \\ 1110 &\to D \end{aligned}$$

Thus, the decoded message is:

$$\boxed{\text{CABBED}}.$$

Since the code is prefix-free, decoding can be done instantaneously by reading bits from left to right and matching the first valid codeword.

3. Uniquely Decodable (UD)

Consider the following binary code for an alphabet of size 5 (same set of codewords as before, but with a different matching between symbols and codewords):

Symbol	Codeword
A	10
B	11
C	00
D	100
E	110

(a). Is this code prefix-free?

Solution: The code is not prefix-free because `10` (codeword for A) is a prefix of `100` (codeword for D), and `11` (codeword for B) is a prefix of `110` (codeword for E).

(b). Decode the following received bit stream into a sequence of symbols:

$$\texttt{1001100010100110}$$

Solution: Parsing the encoded bit sequence from left to right gives us:

$$\texttt{1001100010100110} = \underbrace{\texttt{100}}_{D}\ \underbrace{\texttt{110}}_{E}\ \underbrace{\texttt{00}}_{C}\ \underbrace{\texttt{10}}_{A}\ \underbrace{\texttt{100}}_{D}\ \underbrace{\texttt{110}}_{E}.$$

Thus the decoded message is:

$$\boxed{\text{DECADE}}.$$

Note that we are unable to perform instantaneous decoding. For example, at the start $\texttt{10}$ could be A or the beginning of $D = \texttt{100}$, so we must look ahead at the next bit to decide. Similarly, $\texttt{11}$ could be B or the beginning of $E = \texttt{110}$.

4. Not Uniquely Decodable Code

Consider the following binary code for an alphabet of size 5:

Symbol	Codeword
A	10
B	00
C	11
D	111
E	000

(a). Encode the following message using the code above:

BADCAB

Solution: Using the given code:

$$\begin{aligned} B &\to 00 \\ A &\to 10 \\ D &\to 111 \\ C &\to 11 \\ A &\to 10 \\ B &\to 00 \end{aligned}$$

Concatenating the codewords, we obtain:

$$\boxed{\texttt{001011111100}}$$

(b). Decode the following received bit stream:

$\texttt{001011111100}$

If more than one decoding is possible, list all valid decodings.

Solution: One valid parsing is:

$$\underbrace{00}_{B}\ \underbrace{10}_{A}\ \underbrace{111}_{D}\ \underbrace{11}_{C}\ \underbrace{10}_{A}\ \underbrace{00}_{B},$$

which yields the message:

BADCAB.

However, note that the substring `111000` can also be parsed as:

$$\underbrace{111}_{D}\ \underbrace{000}_{E}.$$

Applying this alternative parsing gives:

$$\underbrace{00}_{B}\ \underbrace{10}_{A}\ \underbrace{111}_{D}\ \underbrace{111}_{D}\ \underbrace{000}_{E},$$

which yields the message:

BADDE.

Since the same bit stream admits multiple distinct decodings, this code is not uniquely decodable.

5. Clearing Concepts!

Suppose you are working with the alphabet $\mathcal{X} = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$. You have been given the following sequence of symbols:

ABCDABCDABCABCABAAA.

(a). Recall that M denotes the total number of symbols in the alphabet. $M = ?$

Solution: $M = 4$.

(b). Recall that N denotes the total number of symbols in the sequence. $N = ?$

Solution: $N = 19$.

(c). Recall that $n(x)$ denotes the count of symbols x, i.e., the number of times symbol x appears in the sequence. What are the counts for each of the symbols?

Solution: $n(\mathrm{A}) = 8$, $n(\mathrm{B}) = 5$, $n(\mathrm{C}) = 4$, $n(\mathrm{D}) = 2$.

(d). For the code which minimizes the total length of the encoded bit sequence, what can you say about the relative lengths of the codewords?

Solution: A code that minimizes the total encoded length assigns shorter codewords to more frequent symbols. In particular, if $n(x) > n(x')$, then $\ell(x) \leq \ell(x')$. Hence here,

$$\ell(\mathrm{A}) \leq \ell(\mathrm{B}) \leq \ell(\mathrm{C}) \leq \ell(\mathrm{D}).$$

(e). Now we wish to encode the above sequence with the following code:

Symbol	Codeword
A	0
B	10
C	110
D	111

What is the length of the encoded bit sequence?

Solution:

$$\begin{aligned}\text{Length of encoded bit sequence} &= \sum_{x \in \mathcal{X}} \ell(x) n(x) \\ &= \ell(\mathrm{A}) n(\mathrm{A}) + \ell(\mathrm{B}) n(\mathrm{B}) + \ell(\mathrm{C}) n(\mathrm{C}) + \ell(\mathrm{D}) n(\mathrm{D}) \\ &= 1 \times 8 + 2 \times 5 + 3 \times 4 + 3 \times 2 \\ &= 36.\end{aligned}$$

Week 2

Basics of Probability, Entropy, Huffman Coding, and Block Coding

---

1. Probability, Surprise, and Entropy

Suppose we flip a fair coin until it either comes up heads or we have flipped it 4 times. Let X be the sequence of outcomes recorded, and let Y be the number of flips.

(a). Verify that

$$\begin{aligned} P(X = \mathrm{H}) &= \frac{1}{2} \\ P(X = \mathrm{TH}) &= \frac{1}{4} \\ P(X = \mathrm{TTH}) &= \frac{1}{8} \\ P(X = \mathrm{TTTH}) &= \frac{1}{16} \\ P(X = \mathrm{TTTT}) &= \frac{1}{16} \end{aligned}$$

Solution: We have $X = \mathrm{H}$ if and only if the first flip is heads. Since the coin is fair, this happens with probability $\frac{1}{2}$. We have $X = \mathrm{TH}$ if and only if the first flip is tails and the second flip is heads, which each happen with probability $\frac{1}{2}$. Since consecutive coin flips are independent, the probability that $X = \mathrm{TH}$ is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Similarly, the events that $X = \mathrm{TTH}$, $X = \mathrm{TTTH}$, $X = \mathrm{TTTT}$ specify the outcomes of 3, 4, 4 independent fair coin flips respectively. Therefore, their probabilities are $\left(\frac{1}{2}\right)^3, \left(\frac{1}{2}\right)^4, \left(\frac{1}{2}\right)^4$ respectively.

(b). Let A be the event that $Y > 2$. Let B be the event that $Y > 3$. Compare the surprise of events A and B.

Solution: We compute the probability.

$$\begin{aligned} P(Y > 2) &= P\left(\{X = \mathrm{TTH}\} \text{ OR } \{X = \mathrm{TTTH}\} \text{ OR } \{X = \mathrm{TTTT}\}\right) \\ &= P\left(\{X = \mathrm{TTH}\}\right) + P\left(\{X = \mathrm{TTTH}\}\right) + P\left(\{X = \mathrm{TTTT}\}\right) \\ &= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} \\ &= \frac{1}{4} \end{aligned}$$

In the second equality we have used the fact that **the probability of the union of disjoint events is the sum of the probabilities of the events**. In other words, if events are mutually exclusive, then the probability that at least one of them happens (that the first happens OR the second one happens OR…) is just the sum of their probabilities. Similarly,

$$\begin{aligned}
P(Y>3) &= P(\{X=\text{TTTH}\} \text{ OR } \{X=\text{TTTT}\}) \\
&= P(\{X=\text{TTTH}\}) + P(\{X=\text{TTTT}\}) \\
&= \frac{1}{16} + \frac{1}{16} \\
&= \frac{1}{8}
\end{aligned}$$

When we talk about the surprise of an event, we mean the surprise function of the probabilities. Plugging into the definition,

$$\begin{aligned}
S(P(A)) &= \log_2 \frac{1}{\frac{1}{4}} = 2 \\
S(P(B)) &= \log_2 \frac{1}{\frac{1}{8}} = 3
\end{aligned}$$

we conclude that B is more surprising than A.

Another solution is as follows. Note that $Y > 2$ whenever $Y > 3$. Additionally, the outcome $Y = 3$, which happens with probability $\frac{1}{8}$, is included in event A but not B. Therefore $P(A) > P(B)$. By the fact that the surprise is strictly decreasing, $S(P(A)) < S(P(B))$, i.e. A is strictly less surprising than B.

(c). Let C be the event that $Y \geq 3$. Let D be the event that $Y = 2$. Compare the surprise of events C and D.

Solution: Using the first part,

$$\begin{aligned}
P(Y \geq 3) &= P(X=\text{TTH}) + P(X=\text{TTTH}) + P(X=\text{TTTT}) \\
&= \frac{1}{8} + \frac{1}{16} + \frac{1}{16} \\
&= \frac{1}{4} \\
P(Y=2) &= P(X=\text{TH}) \\
&= \frac{1}{4}
\end{aligned}$$

Since C and D have the same probability, they have the same surprise.

(d). Compute $\mathbb{E}[Y]$, i.e., the expected number of flips.

Solution:

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{y=1}^{4} yP(Y=y) \\
&= 1P(X=\text{H}) + 2P(X=\text{TH}) + 3P(X=\text{TTH}) + 4\left(P(X=\text{TTTH}) + P(X=\text{TTTT})\right) \\
&= \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{8}{16} \\
&= \frac{15}{8}
\end{aligned}$$

Note in the second equality we have implicitly added probabilities as follows

$$\begin{aligned} P(Y=4) &= P\left(\{X=\mathrm{TTTH}\} \text{ OR } \{X=\mathrm{TTTT}\}\right) \\ &= P\left(\{X=\mathrm{TTTH}\}\right) + P\left(\{X=\mathrm{TTTT}\}\right) \end{aligned}$$

as the events $\{X = \mathrm{TTTH}\}$ and $\{X = \mathrm{TTTT}\}$ are mutually exclusive.

(e). Compute $H(Y)$, i.e, the entropy of Y.

Solution:

$$\begin{aligned} H(Y) &= \sum_{y=1}^{4} P(Y=y)\log_2\left(\frac{1}{P(Y=y)}\right) \\ &= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{8}\log_2(8) + \frac{1}{8}\log_2(8) \\ &= \frac{1}{2} + \frac{1}{4}2 + \frac{1}{8}3 + \frac{1}{8}3 \\ &= \frac{7}{4} \end{aligned}$$

(f). Compute $H(X)$, i.e, the entropy of X.

Solution:

$$\begin{aligned} &H(X) \\ &= P(X=\mathrm{H})\log_2\frac{1}{P(X=\mathrm{H})} + P(X=\mathrm{TH})\log_2\left(\frac{1}{P(X=\mathrm{TH})}\right) + P(X=\mathrm{TTH})\log_2\frac{1}{P(X=\mathrm{TTH})} \\ &\qquad + P(X=\mathrm{TTTH})\log_2\frac{1}{P(X=\mathrm{TTTH})} + P(X=\mathrm{TTTT})\log_2\frac{1}{P(X=\mathrm{TTTT})} \\ &= \frac{1}{2} + \frac{1}{4}2 + \frac{1}{8}3 + \frac{1}{16}4 + \frac{1}{16}4 \\ &= \frac{15}{8} \end{aligned}$$

(g). Are X and Y independent?

Solution: Recall that two random variables, W and Z are independent if $P(Z = z, W = w) = P(Z = z)P(W = w)$ for every z and w in the respective sets of values Z and W can take. Hence two random variables are not independent, if the above equality is not satisfied for even one pair of w and z.

In this case, we observe that this equality is not satisfied for many values of x and y. For example, $P(X = \mathrm{H}, Y = 1) = P(X = \mathrm{H}) = \frac{1}{2}$. On the other hand, $P(X = \mathrm{H}) = \frac{1}{2}$ and $P(Y = 1) = \frac{1}{2}$. So, $P(X = \mathrm{H}, Y = 1) = \frac{1}{2} \neq \frac{1}{4} = P(X = \mathrm{H})P(Y = 1)$. Hence X and Y are not independent.

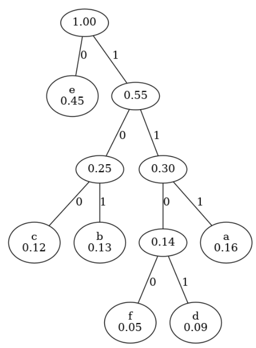
2. Huffman Code

Suppose we have the following probability distribution.

Symbol	Probability
a	0.16
b	0.13
c	0.12
d	0.09
e	0.45
f	0.05

Run the Huffman algorithm on this source. What is the average codeword length? What is the entropy for this source?

Solution: The Huffman tree is as follows:



This gives us the following code - a: 111, b: 101, c: 100, d: 1101, e: 0, f: 1100. The average codeword length is 2.24. The entropy is 2.22.

3. Joint Distribution

Suppose there are two random variables X_1, X_2, each taking values in $\{0, 1\}$. Their joint distribution is given by:

$$P(X_1 = 0, X_2 = 0) = 0.25$$
$$P(X_1 = 0, X_2 = 1) = 0.125$$
$$P(X_1 = 1, X_2 = 0) = 0.125$$
$$P(X_1 = 1, X_2 = 1) = 0.5$$

Let $Z = 2X_1 + 3X_2$ where Z takes values in $\{0, 2, 3, 5\}$.

(a). What is the probability distribution for X_1? What is the entropy of X_1?

Solution: $P(X_1 = 0) = P(X_1 = 0, X_2 = 0) + P(X_1 = 0, X_2 = 1) = 0.375$ and $P(X_1 = 1) = 0.625$.

$$H(X_1) = 0.375 \log\left(\frac{1}{0.375}\right) + 0.625 \log\left(\frac{1}{0.625}\right) = 0.95.$$

(b). What is the entropy of Z?

Solution: Let us first obtain the distribution of Z.

$$\begin{aligned}
P(Z = 0) &= P(X_1 = 0, X_2 = 0) = 0.25 \\
P(Z = 3) &= P(X_1 = 0, X_2 = 1) = 0.125 \\
P(Z = 2) &= P(X_1 = 1, X_2 = 0) = 0.125 \\
P(Z = 5) &= P(X_1 = 1, X_2 = 1) = 0.5
\end{aligned}$$

Hence the entropy of Z is

$$H(Z) = 0.25 \log(4) + 0.125 \log(8) + 0.125 \log(8) + 0.5 \log(2) = 1.75.$$

(c). What is the joint entropy of (X_1, X_2)?

Solution: Recall that joint entropy for (X_1, X_2) is given by:

$$H(X_1, X_2) = \sum_i \sum_j P(X_1 = i, X_2 = j) \log\left(\frac{1}{P(X_1 = i, X_2 = j)}\right).$$

In this case,

$$H(X_1, X_2) = 0.25 \log(4) + 0.125 \log(8) + 0.125 \log(8) + 0.5 \log(2) = 1.75.$$

(d). True or False:

i) Knowing Z is sufficient to know X_1.

ii) Knowing X_1 is sufficient to know Z.

Solution:

i) True. If $Z = 0$ or 3, we know that $X_1 = 0$. If $Z = 2$ or 5, we know that $X_1 = 1$. Hence knowing Z is sufficient to know X_1.

ii) False. If $X_1 = 0$, Z can still take values 0 and 3. Hence knowing X_1 is not sufficient to know Z.

4. Block Coding

Consider the source X which has the following distribution -

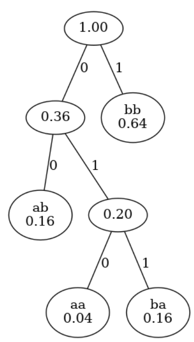
Symbol	Probability
a	0.2
b	0.8

(a). Consider (X_1, X_2) which are independent copies of X. Run the Huffman algorithm on the block (X_1, X_2) where the joint distribution is obtained using independence of X_1 and X_2.

Solution: We are given that X_1 and X_2 are independent. Hence $P(X_1 = i, X_2 = j) = P(X_1 = i)P(X_2 = j)$. Then, the distribution for X_1, X_2 is as follows.

Block X_1X_2	Probability
aa	0.04
ab	0.16
ba	0.16
bb	0.64

The Huffman tree is as follows -



Hence the codewords are aa: 010, ba: 011, ab: 00, bb: 1.

(b). What is the average codeword length? Note that this is the same as the average number of bits per block.

Solution: The average codeword length is

$$\bar{\ell_2} = 3 * 0.04 + 3 * 0.16 + 2 * 0.16 + 1 * 0.64 = 1.56.$$

Hence the average number of bits per block is 1.56 bits.

(c). What is the average number of bits per source symbol?

Solution: The average number of bits per symbol is $\bar{\ell_2}/2 = 0.78$ bits.

(d). Suppose you are given a sequence of 5000 symbols. Each symbol in this sequence is independent and identically distributed with the distribution of X. What are the average number of bits used to represent this sequence if you use the code from the first part?

Solution: A sequence of 5000 symbols means a sequence of 2500 blocks. So the average total number of bits required for 5000 symbols is $2500 * 1.56 = 3900$.

5. Better Than Entropy?

In this problem, we will consider coding for a sequence $X_1, X_2, \ldots,$. Here X_i, X_{i+1} have the following joint distribution for all odd $i = 1, 3, 5, \ldots$:

(X_i, X_{i+1})	**Probability**
a, a	0.4
a, b	0.1
b, a	0.1
b, b	0.4

Note that this is different from the setting discussed at the end of Lecture 4, where we assumed all X_i's are independent.

(a). What is the distribution of each X_i? What is $H(X_i)$? Are X_1 and X_2 independent of each other?

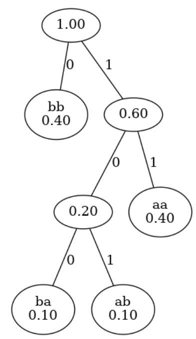
Solution: $P(X_i = a) = P(X_i = a, X_{i+1} = a) + P(X_i = a, X_{i+1} = b) = 0.5$. Similarly, $P(X_i = b) = 0.5$. $H(X_i) = 1$. Note that X_1 and X_2 are not independent as $P(X_1 = 0, X_2 = 0) = 0.4 \neq 0.25 = P(X_1 = 0)P(X_2 = 0)$.

(b). Find the Huffman code for X_i. What is the average number of bits per symbol required to represent this sequence using this code?

Solution: The Huffman code for X_i is trivially $a : 0$ and $b : 1$. The average number of bits per symbol required to represent this sequence is 1.

(c). Find the Huffman code for (X_i, X_{i+1}) for odd i. Let $\bar{\ell}_2$ denote the the average number of bits per block required to represent this sequence using this code. What is $\bar{\ell}_2$? What is the average number of bits required to represent a symbol ($\bar{\ell}_2/2$)?

Solution: The Huffman tree is as follows.



Hence the code is aa: 11, ab: 101, ba: 100, bb: 0. The average number of bits per block is $\bar{\ell_2} = 2 * 0.4 + 3 * 0.1 + 3 * 0.1 + 1 * 0.4 = 1.8$. The average number of bits required to represent a symbol is $\bar{\ell_2}/2 = 0.9$.

(d). You might have noticed that the "average number of bits required to represent a symbol" (previous part) is lower than entropy of each X_i, i.e., $\bar{\ell_2}/2 < H(X_i)$. Is this a violation of the Shannon's source coding theorem discussed in class?

Theorem 1 (Shannon's Source Coding Theorem)**.** *The entropy of a source equals the minimum number of bits per source symbol necessary on average to encode a sequence of independent and identically distributed symbols from that source. In general, this may require the use of block coding, where blocks of symbols are encoded together.*

Solution: No, this is not a violation of Shannon's theorem. Shannon's source coding theorem only applies when the sequence has independent symbols. In our case, the symbols are not independent, and hence by doing block coding we are able to exploit the dependence to perform better than entropy of individual symbols. The inequality that the code satisfies is $H(X_i, X_{i+1}) \leq \bar{\ell_2} \leq H(X_i, X_{i+1}) + 1$.

6. Huffman Coding Intuition

State whether the following statement is true or false - The Huffman algorithm can output a code where all symbols have different codeword lengths.
Solution: False. We give two intuitions for why this statement is false.

- Suppose a code had different codeword lengths for each symbol. Then there exists a codeword with the highest length. Recall that Huffman algorithm always generates a prefix free code and hence no other codeword is a prefix of this longest codeword. This implies that if we remove last bit from the longest codeword, the resulting set of codewords is still prefix free. Hence this new code is still prefix free and has a strictly better average codeword length. This is a contradiction to the fact that Huffman algorithm gives the optimal code (in terms of average codeword length) for any distribution. Hence it is not possible that all codewords have different lengths.

- For the second intuition, without loss of generality, let a and b be the symbols with the two lowest probabilities. If there are repeated probabilities, let a and b be the symbols which are merged in the first step of the Huffman algorithm. After merging they give us the supernode $\{a, b\}$. Recall that the root of the tree is the supernode with probability 1 (after merging all nodes). The length of codeword for a is one added to the distance of supernode $\{a, b\}$ from the root. Similarly, the length of codeword for b is also one added to the distance of supernode $\{a, b\}$ from the root. And hence the lengths of codewords for a and b are equal. Hence, there are at least two codewords which have the same length.

7. Block Coding using Fixed Length Codes

Recall that a fixed length code is a code where each codeword has the same length k.

(a). Let $\mathcal{X}$ be a finite alphabet with M symbols. Suppose we are encoding a single symbol from this alphabet. What is the minimum value of k such that all symbols have distinct codewords? Give a lower and upper bound on this k.

Solution: The minimum possible k such that each symbol has a distinct codeword is $k = \lceil \log_2(M) \rceil$. By definition of the ceiling operator, $y \leq \lceil y \rceil \leq y + 1$. Hence

$$\log_2(M) \leq k \leq \log_2(M) + 1.$$

(b). Consider coding blocks of length n. How many distinct blocks of length n are there? Let k_n denote the length of fixed length code with distinct codeword for each block. What is the minimum value of k_n? Give a lower and upper bound on k_n.

Solution: There are a total of M^n distinct blocks, and hence $k_n = \lceil \log_2(M^n) \rceil = \lceil n \log_2(M) \rceil$. This gives us the following bound

$$n \log_2(M) \leq k_n \leq n \log_2(M) + 1.$$

(c). What is the number of bits per symbol based on the above block coding using fixed length codes? Is there an advantage of coding blocks instead of coding individual symbols (when using fixed length codes)?

Solution: The number of bits per symbol is $\frac{k_n}{n} = \frac{\lceil n \log_2(M) \rceil}{n}$. This is bounded as follows:

$$\log_2(M) \leq \text{Number of bits per symbol } = \frac{k_n}{n} \leq \log_2(M) + \frac{1}{n}.$$

Yes, performing block coding potentially helps as the number of bits per symbol is now within $1/n$ bits of $\log_2(M)$ as compared to within 1 bits when coding individual symbols.

(d). What happens as n approaches infinity? Can block coding using fixed length codes achieve entropy for any source distribution?

Solution: As n approaches infinity, the number of bits per symbol approaches $\log_2(M)$. Recall that $H(X) \le \log_2(M)$ for source X taking values in an alphabet with M symbols. Moreover the equality is achieved if and only if X is a r.v. with uniform distribution over the alphabet. Therefore block coding using finite length codes achieves entropy only for sources with uniform distribution and not for other distributions.

8. Additional Practice for Probability and Expectation

Suppose a snake grows randomly each day by 1mm, 2mm or 3mm with probabilities 0.6, 0.2, 0.2, respectively. Their growth on any particular day is independent of their previous growths. Suppose their length before day one is 50mm. Let G_i denote the growth in mm on day i and let L_i denote the length in mm after day i.

(a). What is their expected length after day one?

Solution:

$$\begin{aligned}
\mathbb{E}[L_1] &= \sum_l lP(L_1 = l) \\
&= 51P(L_1 = 51) + 52P(L_1 = 52) + 53(L_1 = 53) \\
&= 51P(G_1 = 1) + 52P(G_1 = 2) + 53(G_1 = 3) \\
&= 51 \times 0.6 + 52 \times 0.2 + 53 \times 0.2 \\
&= 51.6
\end{aligned}$$

(b). What is their expected length after day two?

Solution: By independence we can compute

$$\begin{aligned}
P(G_1 = 1, G_2 = 1) = P(G_1 = 1)P(G_2 = 1) = 0.6 \times 0.6 = 0.36 \\
P(G_1 = 1, G_2 = 2) = P(G_1 = 1)P(G_2 = 2) = 0.6 \times 0.2 = 0.12 \\
P(G_1 = 1, G_2 = 3) = P(G_1 = 1)P(G_2 = 3) = 0.6 \times 0.2 = 0.12 \\
P(G_1 = 2, G_2 = 1) = P(G_1 = 2)P(G_2 = 1) = 0.2 \times 0.6 = 0.12 \\
P(G_1 = 2, G_2 = 2) = P(G_1 = 2)P(G_2 = 2) = 0.2 \times 0.2 = 0.04 \\
P(G_1 = 2, G_2 = 3) = P(G_1 = 2)P(G_2 = 3) = 0.2 \times 0.2 = 0.04 \\
P(G_1 = 3, G_2 = 1) = P(G_1 = 3)P(G_2 = 1) = 0.2 \times 0.6 = 0.12 \\
P(G_1 = 3, G_2 = 2) = P(G_1 = 3)P(G_2 = 2) = 0.2 \times 0.2 = 0.04 \\
P(G_1 = 3, G_2 = 3) = P(G_1 = 3)P(G_2 = 3) = 0.2 \times 0.2 = 0.04
\end{aligned}$$

We note that each of the above events are mutually exclusive. Adding the corresponding probabilities,

$$
\begin{aligned}
P(L_2 = 52) &= P(G_1 = 1, G_2 = 1) \\
&= 0.36 \\
P(L_2 = 53) &= P(\{G_1 = 1, G_2 = 2\} \text{ OR } \{G_1 = 2, G_2 = 1\}) \\
&= P(G_1 = 1, G_2 = 2) + P(G_1 = 2, G_2 = 1) \\
&= 0.12 + 0.12 \\
&= 0.24 \\
P(L_2 = 54) &= P(\{G_1 = 1, G_2 = 3\} \text{ OR } \{G_1 = 2, G_2 = 2\} \text{ OR } \{G_1 = 3, G_2 = 1\}) \\
&= P(G_1 = 1, G_2 = 3) + P(G_1 = 2, G_2 = 2) + P(G_1 = 3, G_2 = 1) \\
&= 0.12 + 0.04 + 0.12 \\
&= 0.28 \\
P(L_2 = 55) &= P(\{G_1 = 2, G_2 = 3\} \text{ OR } \{G_1 = 3, G_2 = 2\}) \\
&= P(G_1 = 2, G_2 = 3) + P(G_1 = 3, G_2 = 2) \\
&= 0.04 + 0.04 \\
&= 0.08 \\
P(L_2 = 56) &= P(G_1 = 3, G_2 = 3) \\
&= 0.04
\end{aligned}
$$

Finally, we can apply the formula for the expectation

$$
\begin{aligned}
\mathbb{E}[L_2] &= \sum_{l=52}^{56} l P(L_2 = l) \\
&= 0.36 \times 52 + 0.24 \times 53 + 0.28 \times 54 + 0.08 \times 55 + 0.04 \times 56 \\
&= 53.2
\end{aligned}
$$

You may have noticed that the expected amount of growth over two days (3.2 mm) is twice as much as the expected amount of growth over one day (1.6 mm).

This is an example of a more general phenomenon, which is not required but is useful to know.

Proposition 2 (Linearity of Expectation)**.** *For* any *two random variables* X, Y *(*not necessarily *independent or identically distributed)*

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Also, for any deterministic (i.e. not a random variable) real number c*,*

$$\mathbb{E}[cX] = c\mathbb{E}[X].$$

As we have just done above, it is possible to solve this problem without using linearity of expectation at all. However, as we will see shortly it can save some effort.

First we can compute, for any i,

$$\begin{aligned}
\mathbb{E}[G_i] &= \sum_g g \times P(G_i = g) \\
&= 1 \times P(G_i = 1) + 2 \times P(G_i = 2) + 3 \times P(G_i = 3) \\
&= 0.6 + 0.4 + 0.6 = 1.6
\end{aligned}$$

We apply this to the past two subproblems. Since $L_1 = 50 + G_1$, by linearity $\mathbb{E}[L_1] = \mathbb{E}[50] + \mathbb{E}[G_1] = 50 + 1.6 = 51.6$ (note that we have used the fact that the expectation of a non-random number is just the number itself). Similarly, $L_2 = 50 + G_1 + G_2$, so by applying linearity

$$\begin{aligned}
\mathbb{E}[L_2] &= \mathbb{E}[L_1 + G_2] \\
&= \mathbb{E}[L_1 + G_2] \\
&= \mathbb{E}[L_1] + \mathbb{E}[G_2] \\
&= 51.6 + 1.6 = 53.2
\end{aligned}$$

(c). What is their expected length after day three? Note that we do NOT want you to check all 27 possibilities.

Solution: Now using the linearity of expectation we have that $\mathbb{E}[L_3] = 50 + G_1 + G_2 + G_3$, which implies that the expected length after day two is $\mathbb{E}[L_2] = 50 + \mathbb{E}[G_1] + \mathbb{E}[G_2] + \mathbb{E}[G_3] = 54.8$.

(d). What is their expected length after day ten?

Solution: Again using linearity of expectation,

$$\begin{aligned}
\mathbb{E}[L_{10}] &= \mathbb{E}\left[50 + \sum_{i=1}^{10} L_i\right] \\
&= 50 + \sum_{i=1}^{10} \mathbb{E}[L_i] \\
&= 50 + \sum_{i=1}^{10} 1.6 \\
&= 50 + 10 \times 1.6 \\
&= 66
\end{aligned}$$

(e). Evaluate the probability that their length is greater than or equal to 55mm after day two.

Solution: We have that $L_2 \geq 55$ exactly when $G_1 + G_2 \geq 5$. This can occur if $G_1 = G_2 = 3$ or if G_1 and G_2 are 2 and 3, in either order. Then,

$$\begin{aligned}
P(L_2 \geq 55) &= P(\{G_1 = 3, G_2 = 3\} \text{ OR } \{G_1 = 2, G_2 = 3\} \text{ OR } \{G_1 = 3, G_2 = 2\}) \\
&= P(G_1 = 3, G_2 = 3) + P(G_1 = 2, G_2 = 3) + P(G_1 = 3, G_2 = 2) \\
&= P(G_1 = 3)P(G_2 = 3) + P(G_1 = 2)P(G_2 = 3) + P(G_1 = 3)P(G_2 = 2) \\
&= 0.2 \times 0.2 + 0.2 \times 0.2 + 0.2 \times 0.2 \\
&= 0.12
\end{aligned}$$

The second equality is by adding the probabilities of disjoint events and the third is using the fact that G_1, G_2 are independent.

(f). Evaluate the probability that their length is greater than 51mm after day one but less than 55mm after day two.

Solution: Note that the length after day one and the length after day two are not independent. So computing the probabilities for separately for an event for L_1 and an event for L_2 and multiplying *would not* produce the correct answer.

Instead we compute

$$\begin{aligned}
&P(L1 > 51, L_2 < 55) \\
&= P(\{G_1 = 2, G_2 = 1\} \text{ OR } \{G_1 = 2, G_2 = 2\} \text{ OR } \{G_1 = 3, G_2 = 1\}) \\
&= P(G_1 = 2, G_2 = 1) + P(G_1 = 2, G_2 = 2) + P(G_1 = 3, G_2 = 1) \\
&= P(G_1 = 2)P(G_2 = 1) + P(G_1 = 2)P(G_2 = 2) + P(G_1 = 3)P(G_2 = 1) \\
&= 0.2 \times 0.6 + 0.2 \times 0.2 + 0.2 \times 0.6 \\
&= 0.28
\end{aligned}$$

The second equality is by adding the probabilities of disjoint events and the third is using the fact that G_1, G_2 are independent.

(g). Suppose another breed of snake has the same growth pattern, but has a maximum length of 53mm. What is the expected number of days till they reach maximum length?

Solution: Denote the number of days as D.

Since the snake grows by at least 1 mm each day, the maximum number of days is 3. The probabilities for different values of D can be computed as follows.

$$\begin{aligned}
P(D=1) &= P(G_1=3) \\
&= 0.2 \\
P(D=2) &= P\left(\{G_1=1, G_2=2\} \text{ OR } \{G_1=1, G_2=3\} \text{ OR } \{G_1=2\}\right) \\
&= P(G_1=1)P(G_2=2) + P(G_1=1)P(G_2=3) + P(G_1=2) \\
&= 0.6 \times 0.2 + 0.6 \times 0.2 + 0.2 \\
&= 0.44 \\
P(D=3) &= P\left(G_1=1, G_2=1, \right) \\
&= P(G_1=1)P(G_2=1) \\
&= 0.6^2 \\
&= 0.36
\end{aligned}$$

Now we compute the expectation

$$\begin{aligned}
\mathbb{E}[D] &= P(D=1) + 2P(D=2) + 3P(D=3) \\
&= 0.2 + 2 \times 0.44 + 3 \times 0.36 \\
&= 2.16
\end{aligned}$$

Week 3

Coding for Dependent Sources, Signals, and Fourier Representation

---

1. Joint Entropy for a Dependent Source

Consider a sequence of discrete random variables $X_1, X_2, X_3, \ldots$ generated from a source, where each symbol is a tuple $X_i = (Y, U_i)$. The random variables $\{U_i\}$ are independent and identically distributed and are independent of Y. Moreover, their entropies satisfy: $H(Y) = \frac{4}{6}$ and $H(U_i) = \frac{1}{6}$ for each i. Note that the same random variable Y is shared across all symbols in the sequence.

(a). Compute $H(X_i)$ for $i \in \{1, 2, \ldots\}$.

Solution: Using the fact that $X_i = (Y, U_i)$, we have that $H(X_i) = H(Y, U_i)$. Additionally Y and U_i are independent and hence $H(X_i) = H(Y, U_i) = H(Y) + H(U_i) = \frac{5}{6}$.

(b). Compute $H(X_1, X_2)$. Also compute $H(X_1, \ldots, X_n)$ for general $n \in \mathbb{N}$.

Solution: Note that $X_1 = (Y, U_1)$ and $X_2 = (Y, U_2)$. Since the random variable Y is shared among the two, the pair (X_1, X_2) contains the same information as (Y, U_1, U_2). This implies that $H(X_1, X_2) = H(Y, U_1, U_2)$. Using the fact that the sequence U_i is i.i.d. and independent of Y, we get $H(X_1, X_2) = H(Y) + H(U_1) + H(U_2) = H(Y) + 2H(U_1) = \frac{4}{6} + \frac{2}{6} = 1$. For general n, $H(X_1, X_2, \ldots, X_n) = H(Y, U_1, \ldots, U_n) = H(Y) + nH(U_1) = \frac{n+4}{6}$.

(c). Are the random variables $X_1, \ldots, X_n$ independent?

Solution: No, they are not independent. Intuitively, they share common information as they all share the same Y, and hence they cannot be independent. Formally, we observed that $H(X_1, X_2) = 1$. We also know that $H(X_1) + H(X_2) = 2H(X_1) = \frac{10}{6}$. Since $H(X_1, X_2) < H(X_1) + H(X_2)$, the random variables are not independent.

(d). Let $\bar{\ell}_{block,n}$ denote the average code length if Huffman coding is performed on blocks of length n. Give a lower and upper bound on $\bar{\ell}_{block,n}$.

Solution: We know that if Huffman coding is performed on source X, then the average code length satisfies $H(X) \leq \bar{\ell} \leq H(X) + 1$. In this case,

$$
\begin{aligned}
& H(X_1, \ldots, X_n) \leq \bar{\ell}_{block,n} \leq H(X_1, \ldots, X_n) + 1 \\
\implies & \frac{n+4}{6} \leq \bar{\ell}_{block,n} \leq \frac{n+4}{6} + 1.
\end{aligned}
$$

(e). What are lower and upper bounds on the average number of bits per source symbols when Huffman coding is applied on blocks of length n? As $n \uparrow \infty$, what value does the average number of bits per symbol approach?

Solution: Note that average number of bits per source symbol is just $\frac{\bar{\ell}_{block,n}}{n}$. Then,

$$\frac{1}{6} + \frac{4}{6n} \leq \frac{\bar{\ell}_{block,n}}{n} \leq \frac{1}{6} + \frac{4}{6n} + \frac{1}{n}.$$

As n increases the terms $4/(6n)$ and $1/n$ go to zero. Therefore, the average number of bits per source symbol approach $1/6$ as n increases.

2. Encoding Strategies for a Dependent Source

Let Y be a random variable with distribution

$$P(Y = 10) = 0.3, \qquad P(Y = 20) = 0.7.$$

Let $X_1, X_2, \ldots, X_N$ be independent and identically distributed random variables, each with distribution

$$P(X_i = 2) = 0.3, \qquad P(X_i = 5) = 0.7.$$

Assume that the sequence $\{X_i\}_{i=1}^N$ is independent of Y. We wish to represent the sequence $Z_1, \ldots, Z_N$ where $Z_i = Y + X_i$ for all i. Note that the random variable Y is drawn once and added to all X_i. We compare four coding strategies for representing the sequence $Z_1, \ldots, Z_N$.

(a). Compute the distribution of Z_1, and construct a Huffman code for this distribution. Suppose each Z_i is encoded separately using this same code. What is the expected total number of bits required to represent the sequence $Z_1, \ldots, Z_N$?

Solution: The distribution for Z_1 is as follows:

$$P(Z = 12) = 0.09, \quad P(Z = 15) = 0.21, \quad P(Z = 22) = 0.21, \quad P(Z = 25) = 0.49.$$

The Huffman code constructed for this distribution is as follows:

$$12 \to 000, \quad 15 \to 001, \quad 22 \to 01, \quad 25 \to 1.$$

The average code length for this code is as follows:

$$3 \times 0.09 + 3 \times 0.21 + 2 \times 0.21 + 0.49 \times 1 = 1.81.$$

Hence, the expected total number of bits required to represent the sequence $Z_1, \ldots, Z_N$ is $1.81N$.

(b). Suppose we perform block coding with block size n. That is, we compute the joint distribution of $(Z_1, \ldots, Z_n)$ and construct a Huffman code for this block distribution. The sequence is divided into N/n non-overlapping blocks of length n, and each block is encoded separately using this block code (assume that n divides N). Provide a lower and upper bound on the expected total number of bits required to represent $Z_1, \ldots, Z_N$ using this block coding approach. You can use

$$H(Z_1, \ldots, Z_n) \approx 0.88n + 0.88.$$

Solution: Let $E[\bar{\ell}_n]$ denote the average number of bits required to represent blocks of size n. Then,

$$\begin{aligned} H(Z_1, \ldots, Z_n) &\le E[\bar{\ell}_n] \le H(Z_1, \ldots, Z_n) + 1 \\ \implies 0.88n + 0.88 &\le E[\bar{\ell}_n] \le 0.88n + 1.88. \end{aligned}$$

Now, the average total number of bits required to represent the sequence $Z_1, \ldots, Z_N$ is $N/n \times E[\bar{\ell}_n]$.

$$0.88N + 0.88\frac{N}{n} \le \text{average total number of bits required} \le 0.88N + 1.88\frac{N}{n}.$$

(c). An alternative way to describe the sequence $Z_1, \ldots, Z_N$ is to separately encode the value of Y and then each of the X_i's. Construct a Huffman code for Y, and a Huffman code for X_1. Using these codes, what is the total number of bits required to represent $Y, X_1, \ldots, X_N$?

Solution: The Huffman code for Y is $10 \to 0, 20 \to 1$ and the Huffman code for X_1 is $2 \to 0, 5 \to 1$. The total number of bits required to represent $Y, X_1, \ldots, X_N$ using this approach is $N + 1$ bits.

(d). Now, suppose we use block coding to represent $X_1, \ldots, X_N$ using blocks of size n. That is, we compute the joint distribution of $(X_1, \ldots, X_n)$, construct a Huffman code for this distribution, and encode the sequence in N/n non-overlapping blocks of length n. Y is still encoded separately using the Huffman code for Y constructed in the previous part. Provide a lower and upper bound on the expected total number of bits required to represent $Y, X_1 \ldots, X_N$ using this scheme.

Solution: Let $E[\hat{\ell}_n]$ denote the average number of bits required to represent blocks of size n of X_i. Then,

$$\begin{aligned} H(X_1, \ldots, X_n) &\le E[\hat{\ell}_n] \le H(X_1, \ldots, X_n) + 1 \\ \implies 0.88n &\le E[\hat{\ell}_n] \le 0.88n + 1. \end{aligned}$$

Here we have used the fact that $H(X_1, \ldots, X_n) = nH(X_1) = 0.88n$ which follows from the fact that $X_1, \ldots, X_n$ are i.i.d. Then,

$$\begin{aligned} \frac{N}{n} \times 0.88n &\le \text{average number of bits required to represent } X_1, \ldots, X_N \le \frac{N}{n} \times (0.88n + 1) \\ \implies 0.88N &\le \text{average number of bits required to represent } X_1, \ldots, X_N \le 0.88N + \frac{N}{n}. \end{aligned}$$

One more bit is required for representing Y. Hence,

$$0.88N + 1 \le \text{expected total number of bits required } \le 0.88N + \frac{N}{n} + 1.$$

(e). Compare the four schemes above in terms of the *average number of bits per source symbol* Z_i. For each scheme, divide the expected total number of bits by N, and determine the limiting value as $n \to \infty$ and $N \to \infty$. Compare the schemes in this asymptotic sense and provide some intuition.

Solution: We compute the average number of bits per source symbol by dividing by N.

- **Scheme (a):** $\frac{1.81N}{N} = 1.81$.
- **Scheme (b):**

$$0.88 + \frac{0.88}{n} \le \frac{\mathbb{E}[L_N]}{N} \le 0.88 + \frac{1.88}{n} \;\Rightarrow\; \frac{\mathbb{E}[L_N]}{N} \to 0.88 \text{ as } n \to \infty.$$

- **Scheme (c):**

$$\frac{N+1}{N} = 1 + \frac{1}{N} \to 1 \text{ as } N \to \infty.$$

- **Scheme (d):**

$$0.88 + \frac{1}{N} \le \frac{\mathbb{E}[L_N]}{N} \le 0.88 + \frac{1}{n} + \frac{1}{N} \;\Rightarrow\; \frac{\mathbb{E}[L_N]}{N} \to 0.88 \text{ as } n, N \to \infty.$$

Thus, asymptotically,

$$\text{Scheme (a): } 1.81, \quad \text{Scheme (c): } 1, \quad \text{Schemes (b) and (d): } 0.88.$$

Schemes (b) and (d) are optimal.

Intuition: In (b), block coding over $Z_1, \ldots, Z_n$ captures the dependence induced by the common Y, spreading its cost over the block. In (d), this is done more explicitly by encoding Y once and block-coding the i.i.d. X_i's. As n grows, the one-time cost of Y becomes negligible, leading both schemes to achieve the same optimal rate. Scheme (c) also separates Y and X_i, which avoids redundancy and is much better than naive coding in (a), but it does not exploit block coding for X_i, so it remains suboptimal compared to (b) and (d).

3. Periodic Functions

Consider two signals

$$x_1(t) = \sin\left(2\pi \frac{1}{T_1} t\right) \quad \text{and} \quad x_2(t) = \sin\left(2\pi \frac{1}{T_2} t\right).$$

Here $T_1, T_2 > 0$. Note that the periods for x_1 and x_2 are T_1 and T_2 respectively.

(a). What is the period for $x_1(-t)$?

Solution: Note that $\sin(-\theta) = -\sin(\theta)$ for all θ. Hence

$$x_1(-t) = \sin\left(2\pi \frac{1}{T_1}(-t)\right) = \sin\left(-2\pi \frac{1}{T_1} t\right) = -\sin\left(2\pi \frac{1}{T_1} t\right).$$

Hence $x_1(-t)$ has the same period T_1.

(b). What is the period for $x_1(t/10)$?

Solution:

$$x_1(t/10) = \sin\left(2\pi \frac{1}{T_1} \frac{t}{10}\right) = \sin\left(2\pi \frac{1}{10T_1} t\right).$$

Hence $x_1(t/10)$ has a period of $10T_1$.

(c). Suppose $T_1 = 2$ and $T_2 = 4$. What is the period for $x_1(t) + x_2(t)$? What if T_1 and T_2 were 2 and 3, respectively?

Solution: Let $\theta_1(t) = 2\pi\frac{1}{T_1}t$ and $\theta_2(t) = 2\pi\frac{1}{T_2}t$. Both of these satisfy $\theta_1(0) = 0$ and $\theta_2(0) = 0$. We know that for a sinusoid, the signal repeats whenever its phase increases by an integer multiple of 2π. Therefore the sum repeats at the smallest $T > 0$ such that both $\theta_1(T)$ and $\theta_2(T)$ are integer multiples of 2π because this is the smallest time at which both sinusoids reset and repeat together. So, we have to find smallest $T > 0$ such that $\frac{T}{T_1}$ and $\frac{T}{T_2}$ are both integers. This is precisely the lowest common multiple of T_1 and T_2. When $T_1 = 2$ and $T_2 = 4$, the period of sum of sinusoids is 4. And when $T_1 = 2$ and $T_2 = 3$, the period of sum of sinusoids is 6.

(d). What if $T_1 = \sqrt{2}$ and $T_2 = 1$? Is the sum periodic?

Solution: Note the sum is periodic if and only if we can find a T such that $\frac{T}{T_1}$ and $\frac{T}{T_2}$ are both integers. Suppose the sum of sinusoids is periodic. This implies $T = k_1\sqrt{2}$ and $T = k_2$ where both k_1 and k_2 are integers. This would mean that $\sqrt{2} = \frac{k_2}{k_1}$ for integers k_1 and k_2, which would further imply that $\sqrt{2}$ is a rational number. This is a contradiction, and hence our premise that the sum of sinusoids is periodic is wrong. Therefore, the sum is not periodic in this case.

4. Fourier Series Representation

Consider the functions

$$f(t) = 2 + \sin\left(2\pi\frac{3}{T}t\right) + 8\cos\left(2\pi\frac{4}{T}t\right),$$

and

$$g(t) = 9\sin\left(2\pi\frac{3}{T}t\right) + \cos\left(2\pi\frac{6}{T}t\right).$$

(a). Suppose $h(t) = f(-t)$. What is the Fourier series representation for $h(t)$?

Solution:

$$\begin{aligned} f(-t) &= 2 + \sin\left(2\pi\frac{3}{T}(-t)\right) + 8\cos\left(2\pi\frac{4}{T}(-t)\right) \\ &= 2 + \sin\left(-2\pi\frac{3}{T}t\right) + 8\cos\left(-2\pi\frac{4}{T}t\right) \\ &= 2 - \sin\left(2\pi\frac{3}{T}t\right) + 8\cos\left(2\pi\frac{4}{T}t\right). \end{aligned}$$

For the last equality, we use the fact that $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$.

(b). Suppose $h(t) = 5f(t) - g(t)$. What is the Fourier series representation for $h(t)$?

Solution:

$$\begin{aligned}
5f(t) - g(t) &= 10 + (5-9)\sin\left(2\pi\frac{3}{T}t\right) + 40\cos\left(2\pi\frac{4}{T}t\right) - \cos\left(2\pi\frac{6}{T}t\right) \\
&= 10 - 4\sin\left(2\pi\frac{3}{T}t\right) + 40\cos\left(2\pi\frac{4}{T}t\right) - \cos\left(2\pi\frac{6}{T}t\right)
\end{aligned}$$

(c). Suppose $h(t) = f(t+T/2)$. What is the Fourier series representation for $h(t)$?

Solution:

$$\begin{aligned}
f(t+T/2) &= 2 + \sin\left(2\pi\frac{3}{T}(t+T/2)\right) + 8\cos\left(2\pi\frac{4}{T}(t+T/2)\right) \\
&= 2 + \sin\left(3\pi + 2\pi\frac{3}{T}t\right) + 8\cos\left(4\pi + 2\pi\frac{4}{T}t\right) \\
&= 2 + \sin\left(\pi + 2\pi\frac{3}{T}t\right) + 8\cos\left(2\pi\frac{4}{T}t\right) \\
&= 2 - \sin\left(2\pi\frac{3}{T}t\right) + 8\cos\left(2\pi\frac{4}{T}t\right).
\end{aligned}$$

Here the second last equality follows from the fact that $\sin(2\pi n + x) = \sin(x)$ and $\cos(2\pi n + x) = \cos(x)$, where n is an integer. The last equality follows from the fact that $\sin(\pi + x) = -\sin(x)$. This can be observed using the plot of $\sin(x)$ or using the expression for $\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$.

(d). (OPTIONAL) Repeat the above parts for general f and g, i.e., if

$$f(t) = b_0 + \sum_{j=1}^{\infty} a_j \sin\left(2\pi\frac{j}{T}t\right) + \sum_{j=1}^{\infty} b_j \cos\left(2\pi\frac{j}{T}t\right),$$

and

$$g(t) = \tilde{b}_0 + \sum_{j=1}^{\infty} \tilde{a}_j \sin\left(2\pi\frac{j}{T}t\right) + \sum_{j=1}^{\infty} \tilde{b}_j \cos\left(2\pi\frac{j}{T}t\right).$$

Solution: For the first part, note that

$$\begin{aligned}
f(-t) &= b_0 + \sum_{j=1}^{\infty} a_j \sin\left(2\pi\frac{j}{T}(-t)\right) + \sum_{j=1}^{\infty} b_j \cos\left(2\pi\frac{j}{T}(-t)\right) \\
&= b_0 - \sum_{j=1}^{\infty} a_j \sin\left(2\pi\frac{j}{T}t\right) + \sum_{j=1}^{\infty} b_j \cos\left(2\pi\frac{j}{T}t\right),
\end{aligned}$$

where we use $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$.

For the second part, note that

$$5f(t) - g(t) = (5b_0 - \tilde{b}_0) + \sum_{j=1}^{\infty}(5a_j - \tilde{a}_j)\sin\left(2\pi\frac{j}{T}t\right) + \sum_{j=1}^{\infty}(5b_j - \tilde{b}_j)\cos\left(2\pi\frac{j}{T}t\right).$$

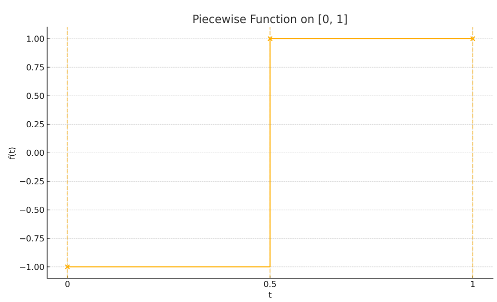
Finally, for the third part, note that

$$\begin{aligned}
f(t+T/2) &= b_0 + \sum_{j=1}^{\infty} a_j \sin\left(2\pi\frac{j}{T}(t+T/2)\right) + \sum_{j=1}^{\infty} b_j \cos\left(2\pi\frac{j}{T}(t+T/2)\right) \\
&= b_0 + \sum_{j=1}^{\infty} a_j \sin\left(2\pi\frac{j}{T}t + \pi j\right) + \sum_{j=1}^{\infty} b_j \cos\left(2\pi\frac{j}{T}t + \pi j\right) \\
&= b_0 + \sum_{j=1}^{\infty}(-1)^j a_j \sin\left(2\pi\frac{j}{T}t\right) + \sum_{j=1}^{\infty}(-1)^j b_j \cos\left(2\pi\frac{j}{T}t\right).
\end{aligned}$$

Here we use $\sin(x+\pi j) = (-1)^j \sin(x)$ and $\cos(x+\pi j) = (-1)^j \cos(x)$.

5. Periodic and Even Periodic Extension

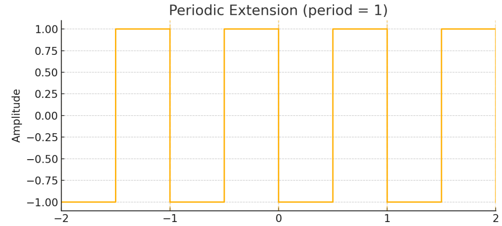
Consider the signal $x(t)$ given in the plot below.

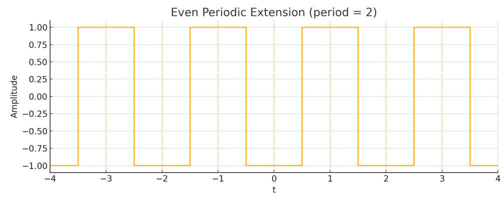


Let $x_p(t)$ denote its periodic extension and $x_{ep}(t)$ denote its even periodic extension.

(a). What is the period for $x_p(t)$ and the period for $x_{ep}(t)$?

Solution: Consider the periodic and the even periodic extensions.





The period for $x_p(t)$ is 1 and the period for $x_{ep}(t)$ is 2.

(b). What is the value of $x_p(-0.8)$? What is the value of $x_{ep}(-0.8)$?

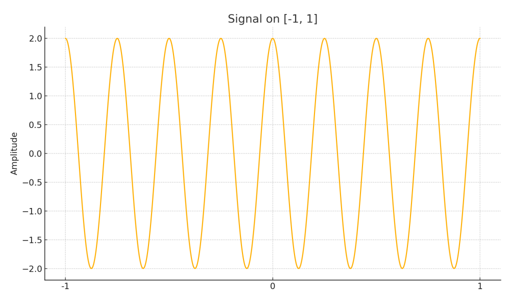
Solution: Looking at the plots, $x_p(-0.8) = -1$ and $x_{ep}(-0.8) = 1$.

(c). What is the value of $x_p(3.8)$? What is the value of $x_{ep}(3.8)$?

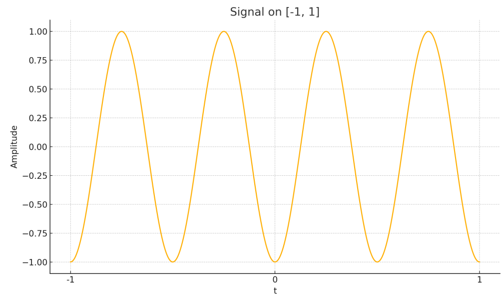
Solution: We observe that x_p is periodic with period 1, and hence $x_p(3.8) = x_p(3 \times 1 + 0.8) = x_p(0.8) = 1$. Similarly, x_{ep} is periodic with period 2, and hence $x_{ep}(3.8) = x_{ep}(1 \times 2 + 1.8) = x_{ep}(1.8) = -1$. The key here is that a periodic function repeats its values after integer multiples of its time period.

6. Identifying Fourier Cosine Series Coefficients

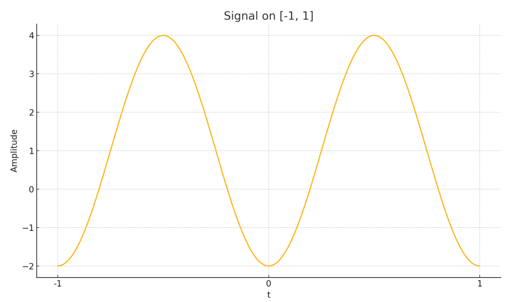
Below are plots of some even periodic signals. These plots only show the values of the function over the domain $[-1, 1]$, but all functions repeat their values outside of this interval. For each function, you need to write its Fourier series representation.



(a). **Solution**: Notice that the minima and maxima of the function are -2 and 2, respectively. Hence the total range is 4, and the amplitude is 2. There are a total of 8 peaks in $[-1, 1]$, and hence the frequency is $8/2 = 4$. Hence the representation is $2\cos(2\pi 4t)$.



(b). Solution: Notice that the minima and maxima of the function are -1 and 1, respectively. Hence the total range is 2. But notice that there is a minima at 0 (i.e., a 'valley' instead of a peak), and hence the amplitude is -1. There are 4 peaks in $[-1, 1]$, and hence the frequency is $4/2 = 2$. Hence the representation is $-\cos(2\pi 2t)$.



(c). Solution: Notice that the minima and maxima of the function are -2 and 4, respectively. Hence the total range is 6. But notice that there is a minima at 0, and hence the amplitude is -3. There is a constant shift by $+1$ (due to which the range is -2 to 4 instead of -3 to 3). There are 2 peaks in $[-1, 1]$, and hence the frequency is $2/2 = 1$. Hence the representation is $1 - 3\cos(2\pi t)$.

Week 4

Discrete Cosine Transform and Quantization

1. Understanding DCT Coefficients

Recall the Discrete Cosine Transform (DCT) for a signal of length L allows us to express it in the following form:

$$y[n] = b_0 + \sum_{j=1}^{L-1} b_j \cos\left(\pi \frac{j}{L}(n + 0.5)\right).$$

Recall that the coefficient b_0 equals the average of the signal, i.e.,

$$b_0 = \frac{1}{L} \sum_{n=0}^{L-1} y[n].$$

(a). Suppose y is a signal of length 10 such that $y[n] = n$ for $n \in \{0, \ldots, 9\}$. What is the value of b_0?

Solution: The coefficient b_0 is the average of the signal. The sum of the signal is given by

$$\sum_{n=0}^{L-1} y[n] = \sum_{n=0}^{9} n = \frac{9 \times 10}{2} = 45.$$

Here we use the fact that the sum of first N natural numbers is $N \times (N+1)/2$. Hence the coefficient b_0 is

$$b_0 = \frac{1}{L} \sum_{n=0}^{L-1} y[n] = \frac{1}{10} \times 45 = 4.5.$$

(b). Let $y \in \mathbb{R}^L$ be a signal with DCT coefficients $b_0, b_1, \ldots, b_{L-1}$. We define a low-pass filtered version y^{LP} by applying a low-pass filter in the frequency domain with cutoff $k_c > 0$, i.e.,

$$b_k^{(\mathrm{LP})} = \begin{cases} b_k, & 0 \leq k \leq k_c, \\ 0, & k > k_c. \end{cases}$$

Let $b_0^{(\mathrm{LP})}, b_1^{(\mathrm{LP})}, \ldots, b_{L-1}^{(\mathrm{LP})}$ denote the DCT coefficients of y^{LP}. What can we say about the average of the signal? Extend your interpretation to low-pass filters in images as well.

Solution: Since $k_c > 0$, the coefficient b_0 is preserved, so $b_0^{(\mathrm{LP})} = b_0$ and the average of the signal remains unchanged. In images, retaining the $(0, 0)$ DCT coefficient similarly preserves the average brightness, while higher-frequency content is smoothed out.

(c). Let $y \in \mathbb{R}^L$ be a signal with DCT coefficients $b_0, b_1, \ldots, b_{L-1}$. We define a high-pass filtered version y^{HP} by applying a high-pass filter in the frequency domain with cutoff $k_c > 0$, i.e.,

$$b_k^{(\mathrm{HP})} = \begin{cases} 0, & 0 \le k \le k_c, \\ b_k, & k > k_c. \end{cases}$$

Let $b_0^{(\mathrm{HP})}, b_1^{(\mathrm{HP})}, \ldots, b_{L-1}^{(\mathrm{HP})}$ denote the DCT coefficients of y^{HP}. What can we say about the average of the signal? Extend your interpretation to high-pass filters in images as well.

Solution: Since $k_c > 0$, the coefficient b_0 is removed, so $b_0^{(\mathrm{HP})} = 0$ and the average of the signal is zero. In images, removing the $(0,0)$ DCT coefficient eliminates the average brightness, leaving only spatial variations such as edges and textures.

(d). Suppose y is a signal of length 6. You know the values of $y[n]$ for $n = 0, 2, 3, 4, 5$, but the value of $y[1]$ is unknown. Let $y[0] = 8, y[2] = 0, y[3] = -4, y[4] = 2, y[5] = 9$. Suppose further that $b_0 = 3$. Then what is the value of $y[1]$?

Solution: We know that b_0 is the average of the signal, and hence $L \times b_0 = \sum_{n=0}^{L-1} y[n]$. In this case, $L \times b_0 = 18 = y[0] + y[1] + \ldots + y[5]$. This gives us that

$$8 + y[1] + 0 - 4 + 2 + 9 = 18,$$

which implies that $y[1] = 3$.

2. Discrete Cosine Transform

Suppose we observe $L = 3$ samples $y[0], y[1], y[2]$. DCT can be used to represent the signal as follows:

$$y[n] = \sum_{j=0}^{2} b_j \cos\left(\frac{\pi}{3}\left(n + \frac{1}{2}\right) j\right), n = 0, 1, 2.$$

The coefficients b_j can be found using

$$b_0 = \frac{1}{3} \sum_{n=0}^{2} y[n]$$

$$b_j = \frac{2}{3} \sum_{n=0}^{2} y[n] \cos\left(\frac{\pi}{3}\left(n + \frac{1}{2}\right) j\right), j = 1, 2.$$

You have been given that $\cos(\pi/6) = \sqrt{3}/2$ and $\cos(\pi/3) = 1/2$.

(a). Create the 3×3 matrix C such that

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix} = C \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}.$$

Solution: We first write the three equations, each corresponding to different values of n.

$$y[0] = b_0 + b_1 \cos\left(\frac{\pi}{3} \times \frac{1}{2}\right) + b_2 \cos\left(\frac{\pi}{3} \times \frac{1}{2} \times 2\right)$$
$$y[1] = b_0 + b_1 \cos\left(\frac{\pi}{3} \times \frac{3}{2}\right) + b_2 \cos\left(\frac{\pi}{3} \times \frac{3}{2} \times 2\right)$$
$$y[2] = b_0 + b_1 \cos\left(\frac{\pi}{3} \times \frac{5}{2}\right) + b_2 \cos\left(\frac{\pi}{3} \times \frac{5}{2} \times 2\right).$$

Stacking these equations and writing in a vector form gives us the following:

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix} = \begin{bmatrix} 1 & \cos\left(\frac{\pi}{3}\left(\frac{1}{2}\right)\right) & \cos\left(\frac{\pi}{3}\left(\frac{1}{2}\right)2\right) \\ 1 & \cos\left(\frac{\pi}{3}\left(\frac{3}{2}\right)\right) & \cos\left(\frac{\pi}{3}\left(\frac{3}{2}\right)2\right) \\ 1 & \cos\left(\frac{\pi}{3}\left(\frac{5}{2}\right)\right) & \cos\left(\frac{\pi}{3}\left(\frac{5}{2}\right)2\right) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 & \cos\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{3}\right) \\ 1 & \cos\left(\frac{\pi}{2}\right) & \cos(\pi) \\ 1 & \cos\left(\frac{5\pi}{6}\right) & \cos\left(\frac{5\pi}{3}\right) \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}.$$

Hence the matrix C is equal to

$$C = \begin{bmatrix} 1 & \sqrt{3}/2 & 1/2 \\ 1 & 0 & -1 \\ 1 & -\sqrt{3}/2 & 1/2 \end{bmatrix}.$$

Here we use the property that $\cos(\pi - x) = -\cos(x)$ and hence $\cos(5\pi/6) = \cos(\pi - \pi/6) = -\cos(\pi/6)$. Similarly, $\cos(2\pi - x) = \cos(x)$ and hence $\cos(5\pi/3) = \cos(2\pi - \pi/3) = \cos(\pi/3)$.

(b). Create the 3×3 matrix D such that

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = D \begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix}.$$

Note that this part is just for your understanding and will not be tested in the exam.

Solution: We follow the same steps as the previous part.

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ (2/3)\cos\left(\frac{\pi}{3}\left(\frac{1}{2}\right)\right) & (2/3)\cos\left(\frac{\pi}{3}\left(\frac{3}{2}\right)\right) & (2/3)\cos\left(\frac{\pi}{3}\left(\frac{5}{2}\right)\right) \\ (2/3)\cos\left(\frac{\pi}{3}\left(\frac{1}{2}\right)2\right) & (2/3)\cos\left(\frac{\pi}{3}\left(\frac{3}{2}\right)2\right) & (2/3)\cos\left(\frac{\pi}{3}\left(\frac{5}{2}\right)2\right) \end{bmatrix} \begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix}.$$

This implies that

$$D = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/\sqrt{3} & 0 & -1/\sqrt{3} \\ 1/3 & -2/3 & 1/3 \end{bmatrix}.$$

(c). Verify that

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix} = CD \begin{bmatrix} y[0] \\ y[1] \\ y[2] \end{bmatrix}.$$

Solution: It can be verified that

$$CD = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. Some Fun with DCT

Let the DCT coefficients for signal $y[0], y[1], y[2]$ be given by b_0, b_1, b_2. Suppose we define $\tilde{b}_0 = b_0, \tilde{b}_1 = -b_1$ and $\tilde{b}_2 = b_2$. Let the corresponding signal be $\tilde{y}$, i.e,

$$\tilde{y}[n] = \sum_{j=0}^{2} \tilde{b}_j \cos\left(\frac{\pi}{3}\left(n + \frac{1}{2}\right) j\right), \quad n = 0, 1, 2.$$

(a). What are the values of $\tilde{y}[0], \tilde{y}[1], \tilde{y}[2]$ in terms of $y[0], y[1], y[2]$?

Solution: We use the computations from the last question here. For the original signal, note that

$$\begin{aligned} y[0] &= b_0 + (\sqrt{3}/2)b_1 + (1/2)b_2 \\ y[1] &= b_0 - b_2 \\ y[2] &= b_0 - (\sqrt{3}/2)b_1 + (1/2)b_2. \end{aligned}$$

Now, for the modified signal

$$\begin{aligned} \tilde{y}[0] &= \tilde{b}_0 + (\sqrt{3}/2)\tilde{b}_1 + (1/2)\tilde{b}_2 = b_0 - (\sqrt{3}/2)b_1 + (1/2)b_2 \\ \tilde{y}[1] &= \tilde{b}_0 - \tilde{b}_2 \\ \tilde{y}[2] &= \tilde{b}_0 - (\sqrt{3}/2)\tilde{b}_1 + (1/2)\tilde{b}_2 = b_0 + (\sqrt{3}/2)b_1 + (1/2)b_2, \end{aligned}$$

where we have substituted $\tilde{b}_0 = b_0, \tilde{b}_1 = -b_1, \tilde{b}_2 = b_2$. Finally this implies that

$$\tilde{y}[0] = y[2],\ \ \tilde{y}[1] = y[1] \text{ and } \tilde{y}[2] = y[0].$$

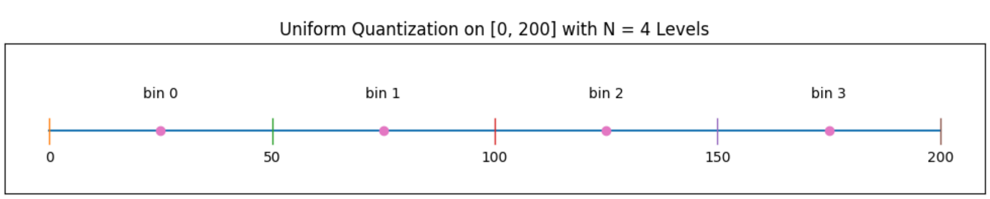
This means that the original signal gets flipped in time when the coefficients with odd indices are multiplied by -1.

(b). Generalize this for signals of length L where $\tilde{b}_j = (-1)^j b_j$. What is the corresponding signal $\tilde{y}$? We just want you to observe the pattern and not try to formally prove this statement.

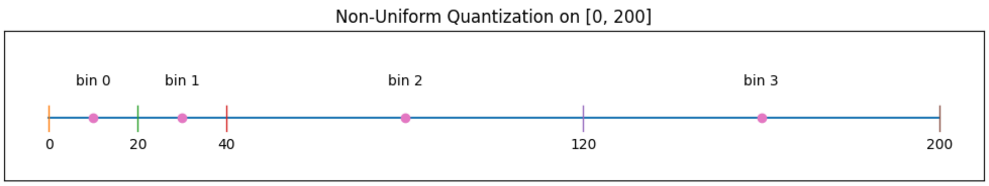
Solution: The signal gets flipped in time. For example, consider the case with $L = 5$. Suppose the original signal is $y[0], y[1], y[2], y[3], y[4]$ and the corresponding DCT coefficients are b_0, b_1, b_2, b_3, b_4. Then the signal corresponding to the coefficients $b_0, -b_1, b_2, -b_3, b_4$ is $y[4], y[3], y[2], y[1], y[0]$.

4. Uniform and Non-Uniform Quantization

In the lecture, we studied uniform quantization, where the total range of values that the variable can take is divided into equal-width intervals. Consider a setting where $X \in [0, 200]$, and we want to quantize the variable using $N = 4$ possible values.



In uniform quantization, the interval $[0, 200]$ is partitioned into four equal-width bins. However, uniform quantization is not the only possible choice. In many applications, it is preferable to use non-uniform quantization. Consider the following non-uniform quantization.



In both cases, dequantization is performed by mapping each quantized value to the midpoint of the corresponding bin, which serves as the reconstruction value. For example, suppose $X = 60.5$. Then

- Under uniform quantization, it will be mapped to bin 1, and after the dequantization, the reconstruction would be $\hat{X} = 75$.
- Under non-uniform quantization, it will be mapped to bin 2, and after the dequantization, the reconstruction would be $\hat{X} = 80$.

(a). Suppose $X = 140$. What bin will this be mapped to under uniform quantization, and what will be the value of the reconstructed $\hat{X}$?

Solution: Note that $140 \in [100, 150]$ and hence $X = 140$ lies in bin 2. The reconstructed value is the midpoint of the bin which is equal to $\hat{X} = 125$.

(b). Suppose $X = 140$. What bin will this be mapped to under non-uniform quantization, and what will be the value of the reconstructed $\hat{X}$?

Solution: Note that $140 \in [120, 200]$ and hence $X = 140$ lies in bin 3. The reconstructed value is the midpoint of the bin which is equal to $\hat{X} = 160$.

(c). We define the reconstruction error for a given symbol X as $(X - \hat{X})^2$ where $\hat{X}$ is the reconstruction after dequantization. What is the reconstruction error when $X = 140$ for uniform and non-uniform quantization?

Solution: For uniform quantization, $\hat{X} = 125$, and hence $(X - \hat{X})^2 = (140 - 125)^2 = 225$. For non-uniform quantization, $\hat{X} = 160$, and hence $(X - \hat{X})^2 = (140 - 160)^2 = 400$.

(d). Suppose we define the mean squared error (MSE) for a sequence of symbols $X_1, X_2, \ldots, X_M$ as $\frac{1}{M}\sum_{i=1}^{M}(X_i - \hat{X}_i)^2$, i.e., the mean of reconstruction error for each symbol in the sequence. Consider the following sequence of length 10: 7, 34, 18, 140, 11, 27, 3, 38, 14, 23. Which of the quantization methods gives a smaller MSE in this case? Can you give an intuitive reason for the same?

Solution: The MSE under uniform quantization is 165.7 and the MSE under non-uniform quantization is 67.7. So, the non-uniform quantization has a lower MSE in this case. This is because the sequence is heavily concentrated in $[0, 40]$, and the non-uniform quantizer assigns smaller intervals here. As a result, typical values incur smaller reconstruction errors, while the few large values have limited impact on the average error.

We make the following remarks regarding the quantizer used in JPEG and how it differs from what we implemented in our image compressor (Project 1c). This discussion is not required for this course, and is included purely for curiosity.

Remark 3 (JPEG.)**.** *In Project 1c, the quantization used for images applies a uniform quantizer to each 2D DCT coefficient. However, JPEG does not use a single quantizer across all coefficients. Instead, different DCT components are quantized using different step sizes. This effectively results in a non-uniform quantization across frequency components. This is done because different DCT coefficients have different statistical distributions, and because the human visual system is less sensitive to errors in higher-frequency components.*

Remark 4 (Simple example.)**.** *Consider a* 2×2 *block of DCT coefficients*

$$\begin{bmatrix} A[0,0] & A[0,1] \\ A[1,0] & A[1,1] \end{bmatrix} = \begin{bmatrix} 120 & 18 \\ 15 & 3 \end{bmatrix}.$$

Suppose we quantize $A[0,0]$ *using a step size of* 10*, and the remaining coefficients using a step size of* 40*. Although each coefficient is quantized uniformly, the quantization is* ***non-uniform across components****. This reflects the JPEG strategy of using different quantizers for different DCT coefficients, rather than a single global quantizer.*

Week 5

Sampling, Interpolation, and Reconstruction

1. Number of Samples Required

Suppose we know that signal $x(t)$ is periodic with period $T = 11$.

(a). Suppose we know that the Fourier series representation of $x(t)$ is limited to the frequency band $\left[0, \frac{3}{11}\right]$. What is the Fourier series representation for $x(t)$? How many samples do we need to accurately represent $x(t)$?

Solution: The signal $x(t)$ can be represented as:

$$x(t) = b_0 + a_1 \sin\left(2\pi\frac{1}{11}t\right) + b_1 \cos\left(2\pi\frac{1}{11}t\right) + \ldots + a_3 \sin\left(2\pi\frac{3}{11}t\right) + b_3 \cos\left(2\pi\frac{3}{11}t\right).$$

The total number of coefficients in this representation is 7, and hence we require 7 samples to accurately represent the signal.

(b). Suppose we know that the Fourier series representation of $x(t)$ is limited to the frequency band $\left[\frac{5}{11}, \frac{10}{11}\right]$. What is the Fourier series representation for $x(t)$? How many samples do we need to accurately represent $x(t)$?

Solution: The signal $x(t)$ can be represented as:

$$x(t) = a_5 \sin\left(2\pi\frac{5}{11}t\right) + b_5 \cos\left(2\pi\frac{5}{11}t\right) + \ldots + a_{10} \sin\left(2\pi\frac{10}{11}t\right) + b_{10} \cos\left(2\pi\frac{10}{11}t\right).$$

The total number of coefficients in this representation is 12, and hence we require 12 samples to accurately represent the signal.

2. Frequency Components and Bandwidth

For each of the following functions, what are the frequency components present in that signal, what is the frequency band of the signal, what is its bandwidth and is it passband or baseband?

(a). $x(t) = 1 - \cos(2\pi t) + 18\sin(2\pi \times 4t)$

Solution:

- Frequency components: 0, 1, 4 Hz
- Frequency Band: $[0, 4]$ Hz
- Bandwidth = $4 - 0 = 4$ Hz
- Baseband Signal

(b). $x(t) = 15\cos(2\pi \times 25t) - \cos(2\pi \times 200t)$

Solution:

- Frequency components: 25, 200 Hz
- Frequency Band: $[25, 200]$ Hz
- Bandwidth = $200 - 25 = 175$ Hz
- Passband Signal

(c). $x(t) = \cos(\pi t) - \sin\left(2\pi \times \frac{2}{5}t\right)$

Solution:

- Frequency components: $\frac{1}{2}, \frac{2}{5}$ Hz
- Frequency Band: $\left[\frac{2}{5}, \frac{1}{2}\right]$ Hz
- Bandwidth = $0.5 - 0.4 = 0.1$ Hz
- Passband Signal

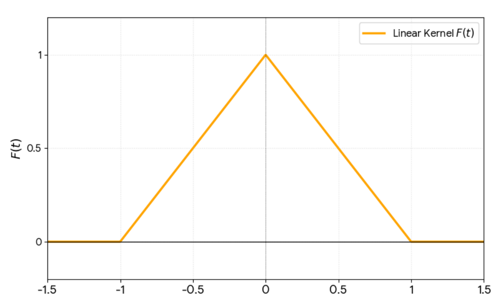
(d). $x(t) = 100 + \cos(2\pi \times 1000t) + 2\sin(2\pi \times 500t)$

Solution:

- Frequency components: 0, 500, 1000 Hz
- Frequency Band: $[0, 1000]$ Hz
- Bandwidth = $1000 - 0 = 1000$ Hz
- Baseband Signal

3. Linear Interpolation

Define the linear interpolation function $F_{lin}(t) = 1 - |t|$ for $|t| < 1$ and 0 otherwise. The plot below represents this function graphically.



Suppose we have sample signal $x(t)$ at multiples of T_s, i.e., you have the samples $x(0), x(T_s), x(2T_s), \ldots$, and reconstruct the signal using linear interpolation, i.e.,

$$\hat{x}(t) = \sum_{m \in \mathbb{Z}} x(mT_s) F_{lin}\left(\frac{t - mT_s}{T_s}\right).$$

(a). Which samples does the value of the reconstructed signal at $5T_s$ depend on?

Solution: Note that

$$\begin{aligned}
\hat{x}(5T_s) &= \sum_{m \in \mathbb{Z}} x(mT_s) F_{lin}\left(\frac{(5-m)T_s}{T_s}\right) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{lin}\left(5 - m\right) \\
&= \ldots + x(4T_s)F_{lin}(1) + x(5T_s)F_{lin}(0) + x(6T_s)F_{lin}(-1) + \ldots \\
&= x(5T_s).
\end{aligned}$$

For the last equality, note that the only value of integer n for which $F_{lin}(n)$ is non-zero is $n = 0$, and hence the only value of m at which $F_{lin}(5 - m)$ is non-zero is at $m = 5$.

The key insight here is that any valid interpolation function preserves the values at the sampling points. Since the original signal was sampled st $5T_s$, the reconstructed signal will have the same value at $5T_s$, i.e., $\hat{x}(5T_s) = x(5T_s)$.

(b). Which samples does the value of the reconstructed signal at $\frac{22}{7}T_s$ depend on?

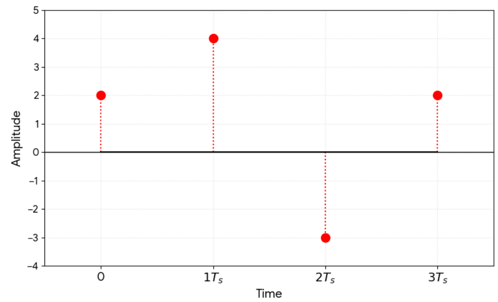
Solution: Note that $\frac{22}{7}T_s = 3.143T_s$. Then,

$$\begin{aligned}
&\hat{x}(3.14T_s) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{lin}\left(\frac{(3.14-m)T_s}{T_s}\right) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{lin}\left(3.14 - m\right) \\
&= \ldots + x(2T_s)F_{lin}(1.14) + x(3T_s)F_{lin}(0.14) + x(4T_s)F_{lin}(-0.86) + x(5T_s)F_{lin}(-1.86) + \ldots \\
&= x(3T_s)F_{lin}(0.14) + x(4T_s)F_{lin}(-0.86).
\end{aligned}$$

The last equality here follows from the fact that $F_{lin}(t)$ is equal to zero for all t less than -1 or greater than 1. Hence, $\hat{x}(3.14T_s)$ depends on the samples at $3T_s$ and $4T_s$.

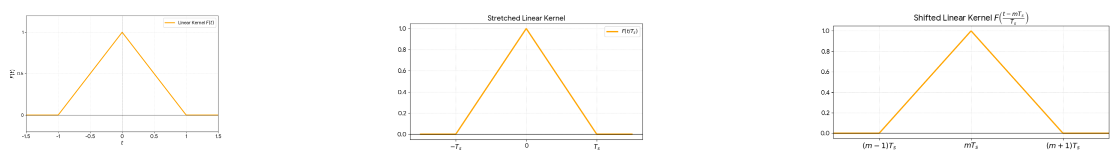
The key insight here is that the value of the reconstructed signal at a non-sample t depends only on the samples on either side of t. So, if time τ lies between mT_s and $(m+1)T_s$, then $\hat{x}(\tau)$ depends on $x(mT_s)$ and $x((m+1)T_s)$.

(c). Recall the discrete samples studied in class for nearest neighbor interpolation.

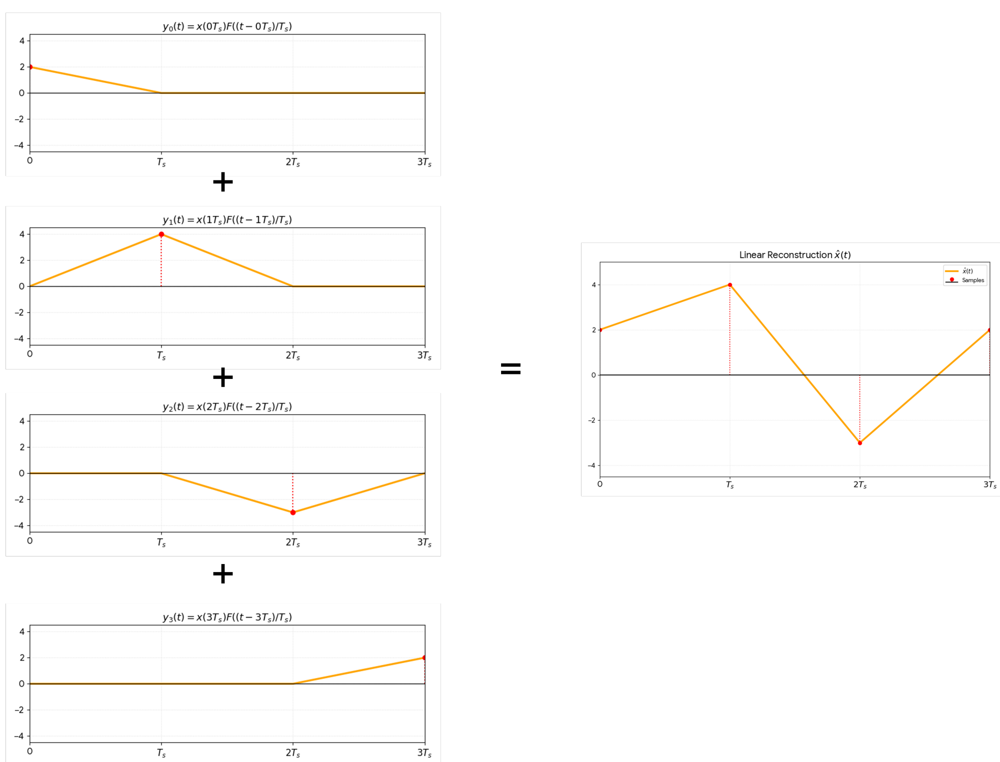


Construct the reconstructed signal for these samples using linear interpolation. Plot all the intermediate steps - $F_{lin}(t/T_s)$, $F_{lin}((t - mT_s)/T_s)$, $x(mT_s)F_{lin}((t - mT_s)/T_s)$ and show how they add up to give the reconstructed signal.

Solution: We first plot the signals $F_{lin}(t/T_s)$ and $F_{lin}((t - mT_s)/T_s)$, which are just the stretched and shifted versions of the linear interpolation function.

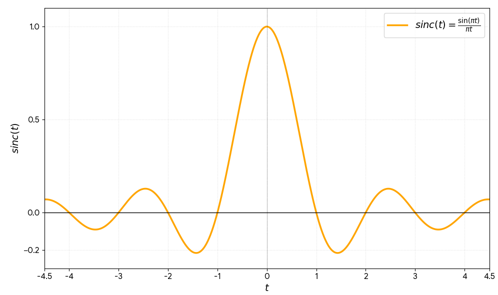


Next we show the influence of each of the samples, and how adding them up gives us them us the reconstructed signal.



4. Sinc Interpolation

Define the linear interpolation function $F_{sinc}(t) = \text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$. The plot below represents this function graphically.



Suppose we have sample signal $x(t)$ at multiples of T_s, i.e., you have the samples $x(0), x(T_s), x(2T_s), \ldots$, and reconstruct the signal using sinc interpolation, i.e.,

$$\hat{x}(t) = \sum_{m \in \mathbb{Z}} x(mT_s) F_{sinc}\left(\frac{t - mT_s}{T_s}\right).$$

(a). Which samples does the value of the reconstructed signal at $5T_s$ depend on?

Solution: Note that

$$\begin{aligned}
\hat{x}(5T_s) &= \sum_{m \in \mathbb{Z}} x(mT_s) F_{sinc}\left(\frac{(5-m)T_s}{T_s}\right) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{sinc}\left(5 - m\right) \\
&= \ldots + x(4T_s)F_{sinc}(1) + x(5T_s)F_{sinc}(0) + x(6T_s)F_{sinc}(-1) + \ldots \\
&= x(5T_s).
\end{aligned}$$

For the last equality, note that the only value of integer n for which $F_{sinc}(n)$ is non-zero is $n = 0$, and hence the only value of m at which $F_{sinc}(5 - m)$ is non-zero is at $m = 5$.

The key insight here is that any valid interpolation function preserves the values at the sampling points. Since the original signal was sampled st $5T_s$, the reconstructed signal will have the same value at $5T_s$, i.e., $\hat{x}(5T_s) = x(5T_s)$.

(b). Which samples does the value of the reconstructed signal at $\frac{22}{7}T_s$ depend on?

Solution: Note that $\frac{22}{7}T_s = 3.143T_s$. Then,

$$\begin{aligned}
&\hat{x}(3.14T_s) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{sinc}\left(\frac{(3.14-m)T_s}{T_s}\right) \\
&= \sum_{m \in \mathbb{Z}} x(mT_s) F_{sinc}\left(3.14 - m\right) \\
&= \ldots + x(2T_s)F_{sinc}(1.14) + x(3T_s)F_{sinc}(0.14) + x(4T_s)F_{sinc}(-0.86) + x(5T_s)F_{sinc}(-1.86) + \ldots.
\end{aligned}$$

Note that for non-integer t, $F_{sinc}(t)$ is always non-zero. And hence you need all samples (past and future) to reconstruct the signal at a non-sampled time.

5. Reconstruction-I

Let $x(t) = 10 - \cos(2\pi 50t) + 15\sin(2\pi 120t)$.

(a). What is the frequency band of the signal?

Solution: The signal lies in the band $[0, 120]$ Hz.

(b). Suppose samples for $x(t)$ are taken with sampling frequency $f_s = 110$ Hz. These samples are then used to reconstruct the signal $\hat{x}(t)$ using sinc interpolation. True or False:

a) $\hat{x}(t) = x(t)$
b) $\hat{x}(t)$ has bandwidth less than or equal to 55 Hz.
c) $\hat{x}(t) = 10 - \cos(2\pi 50t)$.

Solution:

a) False. The sampling frequency is less than $2B$, as $110 < 240$, and hence we do not achieve a perfect reconstruction.
b) True. The reconstructed signal has bandwidth less than $f_s/2$.
c) False. Sinc interpolation does not simply low-pass filter the original signal. Instead, it reconstructs a bandlimited signal whose samples exactly match the given samples. Since the 120 Hz component lies above the Nyquist frequency (55 Hz), it aliases to a lower frequency in the samples. This aliased component must therefore appear in the reconstructed signal; otherwise, the reconstructed signal would not match the samples. Hence, the high-frequency term does not vanish and changes $\hat{x}(t)$.

(c). Suppose samples for $x(t)$ are taken with sampling frequency $f_s = 250$ Hz. These samples are then used to reconstruct the signal $\hat{x}(t)$ using sinc interpolation. True or False:

a) $\hat{x}(t) = x(t)$
b) $\hat{x}(t)$ has bandwidth less than or equal to 125 Hz.

Solution:

a) True. The sampling frequency is greater than $2B$.
b) True. The reconstructed signal has bandwidth less than $f_s/2$. To be more precise, the reconstructed signal is the same as the original signal, and hence the bandwidth of the reconstructed signal is just 120 Hz.

(d). Suppose samples for $x(t)$ are taken with sampling frequency $f_s = 1000$ Hz. These samples are then used to reconstruct the signal $\hat{x}(t)$ using sinc interpolation. True or False:

a) $\hat{x}(t)$ has bandwidth less than or equal to 500 Hz.

b) $\hat{x}(t)$ has components at frequency higher than 120 Hz.

c) $\hat{x}(t) = x(t)$

Solution:

a) True. The reconstructed signal has bandwidth less than $f_s/2$.

b) False. For any $f_s > 2B$, $\hat{x}(t) = x(t)$. And hence the frequency components are only at $0, 50, 120$.

c) True. The sampling frequency is greater than $2B$.

6. Reconstruction-II

(a). Let $x(t)$ be a signal bandlimited in $[0, 6]$ Hz. Let $\hat{x}(t)$ be the signal reconstructed from samples taken at sampling frequency $f_s = 4Hz$ using sinc interpolation.

True or False:

a) $\hat{x}(t)$ is a signal with bandwidth less than or equal to 2 Hz.

b) $\hat{x}(t)$ can have components with frequency between 2 Hz and 6 Hz.

c) $\hat{x}(t) = x(t)$

d) $\hat{x}(t)$ *can* be the signal $\cos(2\pi 4t)$

e) $\hat{x}(t)$ *can* be the signal $\cos(2\pi t)$

Solution:

a) True. By the Nyquist sampling theorem, the reconstructed signal is bandlimited to $[0, f_s/2]$ Hz. Here $f_s = 4$ Hz.

b) False. Since the reconstructed signal is bandlimited in $[0, 2]$ Hz, there can be no components with frequency above 2 Hz.

c) False. Since $f_s/2 < B$, we cannot guarantee perfect reconstruction.

d) False. We know $\hat{x}(t)$ is bandlimited to $[0, 2]$ Hz, and hence it cannot be $\cos(2\pi 4t)$.

e) True. The function $\hat{x}(t)$ can be $\cos(2\pi t)$.

(b). Let $y(t)$ be a signal bandlimited in $[0, 6]$ Hz. Let $\hat{y}(t)$ be the signal reconstructed from samples taken at sampling frequency $f_s = 14Hz$ using sinc interpolation.

True or False:

a) $\hat{y}(t)$ is a signal with bandwidth less than or equal to 6 Hz.

b) $\hat{y}(t)$ can have components with frequency between 6 Hz and 7 Hz.

c) $\hat{y}(t) = y(t)$

d) $\hat{y}(t)$ *can* be the signal $\cos(2\pi 4t)$

e) $\hat{y}(t)$ *can* be the signal $\cos(2\pi t)$

Solution:

a) True. Since $f_s/2 > B$, we have perfect reconstruction and hence $\hat{y}(t) = y(t)$ for all t. Since $y(t)$ is bandlimited in $[0, 6]$ Hz, $\hat{y}(t)$ is also bandlimited in $[0, 6]$ Hz.

b) False. Since the reconstructed signal is the same as the original signal, we know that the reconstructed signal only has components in the band $[0, 6]$ Hz.

c) True. Since $f_s/2 > B$, we can guarantee perfect reconstruction.

d) True. The function $\hat{y}(t)$ can be $\cos(2\pi 4t)$.

e) True. The function $\hat{y}(t)$ can be $\cos(2\pi t)$.

Week 6

Modulation and Demodulation Strategies

1. On-Off Keying

Let b_m for $m = 1, 2, \ldots$ be the bits we wish to communicate. We use on-off keying to create signal $x(t)$. Suppose the mth bit is 0, then the signal $x(t) = 0$ from $(m-1)T_b$ to mT_b, and if the mth bit is 1, then the signal $x(t) = V$ from $(m-1)T_b$ to mT_b. Recall that we consider the signal $x(t)$ to be approximately band-limited to $[0, 1/T_b]$.

(a). What is the bit rate in terms of V and T_b? What is the band of signal $x(t)$ in terms of the data rate?

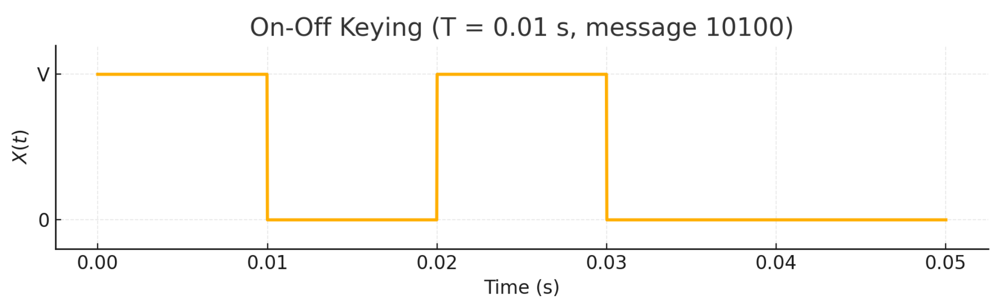
Solution: The bit rate is $1/T_b$. It does not depend on V. The band of signal $x(t)$ is $[0, 1/T_b] = [0, R_b]$ Hz, where $R_b = 1/T_b$ is the bit rate.

(b). Suppose you have been allotted the band $[0, 100]$ Hz. What is the maximum bit rate that you can achieve? What is the value of T_b corresponding to this rate? Does this impose any limitations on V?

Solution: The maximum bit rate will be achieved when $1/T_b = 100$ and hence $R_b = 100$. The value of T_b corresponding to this rate of 100 bits/sec is 0.01 seconds. This does not impose any limitations on V.

(c). Suppose the first 5 bits are 1,0,1,0,0. For the value of T_b calculated in the previous part, plot the signal $x(t)$.

Solution:



2. Passband Signal

You generate a baseband signal $x(t)$ as done in the previous question. You then multiply this function with $\sin(2\pi f_c t)$ to obtain a signal $s(t) = x(t) \times \sin(2\pi f_c t)$. Here we assume that $f_c \gg 1/T_b$.

(a). What is the band for $s(t)$ in terms of V, T_b and f_c? What is the bit rate in terms of V, T_b and f_c? What is the bandwidth of the signal in terms of the rate?

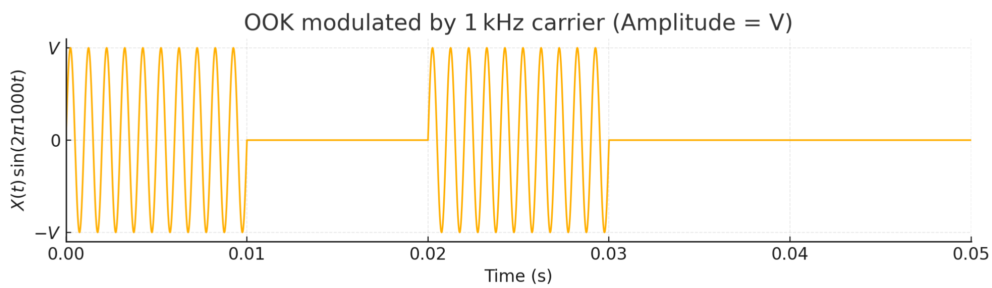
Solution: The band for $s(t)$ is $[f_c - 1/T_b, f_c + 1/T_b]$. The bit rate is still $1/T_b$. The bandwidth of the signal is $2/T_b$ or $2R_b$ where R_b is the bit rate.

(b). Suppose you have been allotted the band $[900, 1100]$. What is the optimal choice of f_c and T_b to obtain the maximum rate? What is the maximum rate?

Solution: We want to maximize the bit rate, and hence we want to maximize $1/T_b$. This is possible only if we occupy the complete allocated band. Given the band $[900, 1100]$, we can maximize the bit rate by setting $f_c - 1/T_b = 900$ and $f_c + 1/T_b = 1100$. This gives us the optimal carrier frequency as 1000 Hz, and optimal T_b as $1/100$ seconds. The maximum rate obtained is 100 bits/sec.

(c). Suppose the first 5 bits are 1,0,1,0,0. For the value of T_b and f_c obtained in the previous part, plot the signal $s(t)$.

Solution:



(d). Suppose you are still allotted the band $[900, 1100]$. Unfortunately, the transmitter you have can only work with a maximum bit rate of 50 bits/sec. What is the range of f_c that you can use to work with this data rate of 50 bits/sec?

Solution: The bit rate has been fixed at 50 bits/sec in this case, and hence $1/T_b = 50$. Now, we need the band $[f_c - 50, f_c + 50]$ to be within our allocated band of $[900, 1100]$. Hence f_c can take values in $[950, 1050]$ Hz.

(e). Suppose you have chosen the carrier frequency to be $f_c = 1075$ Hz. What is the maximum rate achievable in this scenario using the above modulation strategy?

Solution: We require the approximate band $[f_c - 1/T_b, f_c + 1/T_b]$ to be inside the allocated band $[900, 1100]$ Hz. This gives us the constraints $f_c - 1/T_b \geq 900$ and $f_c + 1/T_b \leq 1100$, which give us the constraints $1/T_b \leq 125$ and $1/T_b \leq 25$, respectively. The second constraint here is more restrictive and will dictate the maximum rate. So the maximum rate achievable using our modulation scheme is $R_b = 1/T_b = 25$ bits/sec.

3. Decoding Schemes for Baseband Signals

Let b_m for $m = 1, 2, \ldots$ be the bits we wish to communicate. We use on-off keying to create signal $x(t)$ with $V = 1$ and $T_b = 1$. Suppose the m-th bit is 0, then the signal $x(t) = 0$ from $(m-1)$ to m, and if the

m-th bit is 1, then the signal $x(t) = 1$ from $(m-1)$ to m.

We receive $y(t)$ at the receiver, which is a noisy version of the transmitted signal. Suppose $y(t) = x(t) + \eta(t)$, where

$$P(\eta(t) = 0.6) = 0.5$$
$$P(\eta(t) = -0.6) = 0.5$$

for all t, and $\{\eta(t)\}_{t \geq 0}$ are a collection of independent random variables.

Note: While this model is mathematically consistent, it is not physically realistic. Independence at every time t means the noise can change sign arbitrarily fast, even over extremely small time intervals. Real physical noise cannot fluctuate infinitely fast and typically has some temporal correlation (it varies smoothly over very small time scales).

We use this model purely for analytical simplicity.

(a). Suppose we define $u_m = y(m - 1/2)$. If $b_m = 1$, what is the distribution for u_m? If $b_m = 0$, what is the distribution for u_m?

Solution: If $b_m = 1$, then for any given t, $y(t) = 1 + 0.6 = 1.6$ w.p. 0.5 and $y(t) = 1 - 0.6 = 0.4$ w.p. 0.5. Hence, if $b_m = 1$, then $P(u_m = 1.6) = 0.5$ and $P(u_m = 0.4) = 0.5$.

If $b_m = 0$, then for any given t, $y(t) = 0 + 0.6 = 0.6$ w.p. 0.5 and $y(t) = 0 - 0.6 = -0.6$ w.p. 0.5. Hence, if $b_m = 0$, then $P(u_m = 0.6) = 0.5$ and $P(u_m = -0.6) = 0.5$.

(b). If $u_m \geq 0.5$, then you output decoded bit m as $\hat{b}_m = 1$ and if $u_m < 0.5$, then you output decoded bit m as $\hat{b}_m = 0$. What is the probability that the actual bit was 1, i.e, $b_m = 1$, but you output the decoded bit as 0, i.e, $\hat{b}_m = 0$?

Solution: Suppose $b_m = 1$. Then our observation can be $u_m = 1.6$ with probability 0.5, and $u_m = 0.4$ with probability 0.5. Among these two, we decode $\hat{b}_m = 1$ only when $u_m = 1.6$ if we use the thresholding strategy. And hence we decode incorrectly ($\hat{b}_m = 0$) with probability 0.5 when the actual bit b_m was 1.

(c). Suppose we define

$$v_m = \frac{1}{2}\Big(y(m - 1/4) + y(m - 3/4)\Big).$$

If $b_m = 1$, what is the distribution for v_m? If $b_m = 0$, what is the distribution for v_m?

Solution: First consider the case of $b_m = 1$. Using independence,

$$P(y(m-1/4) = 1.6, y(m-3/4) = 1.6) = 0.25$$
$$P(y(m-1/4) = 1.6, y(m-3/4) = 0.4) = 0.25$$
$$P(y(m-1/4) = 0.4, y(m-3/4) = 1.6) = 0.25$$
$$P(y(m-1/4) = 0.4, y(m-3/4) = 0.4) = 0.25$$

This implies that when $b_m = 1$,

$$P(v_m = 1.6) = 0.25$$
$$P(v_m = 1) = 0.5$$
$$P(v_m = 0.4) = 0.25.$$

Similarly, when $b_m = 0$, we have

$$P(v_m = 0.6) = 0.25$$
$$P(v_m = 0) = 0.5$$
$$P(v_m = -0.6) = 0.25.$$

(d). If $v_m \geq 0.5$, then you output decoded bit m as $\hat{b}_m = 1$ and if $v_m < 0.5$, then you output decoded bit m as $\hat{b}_m = 0$. What is the probability that the actual bit was 1, i.e, $b_m = 1$, but you output the decoded bit as 0, i.e, $\hat{b}_m = 0$?

Solution: Suppose $b_m = 1$. Then our observation can be $v_m = 1.6$ with probability 0.25, $v_m = 1$ with probability 0.5, and $v_m = 0.4$ with probability 0.25. Among these three, we decode $\hat{b}_m = 1$ only when $v_m = 1.6$ or $v_m = 1$ if we use the thresholding strategy. And hence we decode incorrectly ($\hat{b}_m = 0$) when $v_m = 0.4$, which happens with probability 0.25.

(e). Did the averaging help decrease the probability of error? Compare the two schemes and explain intuitively why averaging changes the error probability.

Solution: Yes. In the first scheme, a single "bad" noise realization (i.e., -0.6) causes an error, which happens with probability 0.5. In the second scheme, we average two independent samples. An error occurs only if *both* noise values are -0.6, which has probability $0.5 \times 0.5 = 0.25$.

Intuitively, averaging makes it less likely that random noise pushes the observation below the threshold, because multiple independent "bad" outcomes must occur simultaneously.

4. Some Decoding Strategies for Passband Signals

Let b_m for $m = 1, 2, \ldots$ be the sequence of bits we wish to communicate. You first decide to generate a baseband signal using the on-off keying model. You set $T_b = 1$ and $V = 3$. Let us call this function $x(t)$. You then multiply this function with $\sin(2\pi 20t)$ to obtain a signal $s(t)$, i.e., $s(t) = x(t) \times \sin(2\pi 20t)$. Suppose you are working with a noiseless channel, i.e., the receiver receives $y(t) = s(t)$ exactly.

(a). Suppose we use the following decoding strategy: we first compute

$$w_m = \int_{m-1}^{m} (y(t))^4 dt$$

for decoding bit m. We set the threshold α and if $w_m \geq \alpha$, then the decoded bit is $\hat{b}_m = 1$ and if $w_m < \alpha$, then the decoded bit is $\hat{b}_m = 0$. What is the range of thresholds α that allow for perfect decoding, i.e., $\hat{b}_m = b_m$ for all m?

Note that the integral of $\sin^4(2\pi 20t)$ from 0 to 0.05 is $3/160$.

Solution: Suppose bit $b_m = 1$. Then $x(t) = 3$ between $m-1$ and m, and hence $y(t) = 3\sin(2\pi 20t)$ between $m-1$ and m. Hence,

$$w_m = \int_{m-1}^{m} (3\sin(2\pi 20t))^4 \, dt = 3^4 \times \frac{3}{160} \times 20 = \frac{243}{8}.$$

The final step here follows from the fact that $\sin(2\pi 20t)$ is periodic with period 0.05 and hence we can just multiply the given integral of $\sin^4(2\pi 20t)$ from 0 to 0.05 by $1/0.05$ to get the integral from 0 to 1. Similarly, when bit $b_m = 0$, then $y(t) = 0$ from $m-1$ to m, and hence w_m is also zero.

So, we need to set the threshold α such that $w_m \geq \alpha$ when $b_m = 1$ and $w_m < \alpha$ when $b_m = 0$. When $b_m = 1$, $w_m = 243/8 = 30.375$. So, the range of thresholds α that work is $(0, 30.375]$.

(b). Suppose we use the following decoding strategy: we first compute

$$q_m = \int_{m-1}^{m} |y(t)| dt$$

for decoding bit m. Here $|\cdot|$ denotes the absolute value. We set the threshold α and if $q_m \geq \alpha$, then the decoded bit is $\hat{b}_m = 1$ and if $q_m < \alpha$, then the decoded bit is $\hat{b}_m = 0$. What is the range of thresholds α that allow for perfect decoding, i.e., $\hat{b}_m = b_m$ for all m?

Note that the integral of $|\sin(2\pi 20t)|$ from 0 to 0.05 is $1/(10\pi)$.

Solution: Suppose bit $b_m = 1$. Then $x(t) = 3$ between $m-1$ and m, and hence $y(t) = 3\sin(2\pi 20t)$ between $m-1$ and m. Hence,

$$q_m = \int_{m-1}^{m} |3\sin(2\pi 20t)| dt = 3 \times \frac{1}{10\pi} \times 20 = \frac{6}{\pi}.$$

The final step here follows from the fact that $\sin(2\pi 20t)$ is periodic with period 0.05 and hence we can just multiply the given integral of $|\sin(2\pi 20t)|$ from 0 to 0.05 by $1/0.05$ to get the integral from 0 to 1. Similarly, when bit $b_m = 0$, then $y(t) = 0$ from $m-1$ to m, and hence q_m is also zero.

So, we need to set the threshold α such that $q_m \geq \alpha$ when $b_m = 1$ and $q_m < \alpha$ when $b_m = 0$. When $b_m = 1$, $q_m = 6/\pi = 1.909$. So, the range of thresholds α that work is $(0, 1.909]$.

5. Decoding in Presence of Noise

Let b_m for $m = 1, 2, \ldots$ be the sequence of bits we wish to communicate. You first decide to generate a baseband signal using the on-off keying model. You set $T_b = 0.1$ and $V = 3$. Let us call this function $x(t)$. You then multiply this function with $\sin(2\pi 200t)$ to obtain a signal $s(t)$, i.e., $s(t) = x(t) \times \sin(2\pi 200t)$. The signal $s(t)$ is transmitted and the receiver receives signal $y(t)$. We use the same decoding scheme in every case: we first compute

$$p_m = \frac{1}{T_b} \int_{(m-1)T_b}^{mT_b} (y(t))^2 dt,$$

and for threshold p_{thresh}, the decoded bit $\hat{b}_m$ is 1 if $p_m \geq p_{thresh}$ and 0 otherwise.

(a). Suppose the receiver receives the transmitted signal, i.e., $y(t) = s(t)$. Then what is the range of thresholds that allow for perfect decoding, i.e., $\hat{b}_m = b_m$ for all m? Note that the integral of $\sin^2(2\pi 200t)$ from 0 to 0.005 is $1/400$.

Solution: Suppose bit $b_m = 1$. Then $x(t) = 3$ between $(m-1)T_b$ and mT_b, and hence $y(t) = 3\sin(2\pi 200t)$ between $(m-1)T_b$ and mT_b. Hence,

$$p_m = \frac{1}{0.1} \int_{0.1(m-1)}^{0.1m} (3\sin(2\pi 200t))^2 dt = \frac{1}{0.1} \times 3^2 \times \frac{1}{400} \times 20 = \frac{9}{2}.$$

The final step follows from the fact that $\sin(2\pi 200t)$ has period 0.005, and the interval length 0.1 equals 20 full periods, so we multiply the given integral over $[0, 0.005]$ by $0.1/0.005$. Similarly, when bit $b_m = 0$, then $y(t) = 0$ from $(m-1)T_b$ to mT_b, and hence $p_m = 0$.

Thus, when bit 1 is transmitted, the average energy over the bit interval is 4.5, while it is 0 when bit 0 is transmitted. So, we need to set the threshold p_{thresh} such that $p_m \geq p_{thresh}$ when $b_m = 1$ and $p_m < p_{thresh}$ when $b_m = 0$. Since $p_m = 4.5$ when $b_m = 1$, the range of thresholds that work is $(0, 4.5]$.

(b). Now suppose the channel attenuates the signal and then adds a constant shift, i.e., $y(t) = \alpha s(t) + \beta$, where $0 < \alpha < 1$ is the attenuation factor. Is it always possible to perform error-free decoding? For given α and β, what values of p_{thresh} allow for perfect decoding.

Solution: Suppose bit $b_m = 1$. Then $y(t) = 3\alpha \sin(2\pi 200t) + \beta$ in the interval $[(m-1)T_b, mT_b]$.

$$\begin{aligned}
p_m &= \frac{1}{0.1} \int_{0.1(m-1)}^{0.1m} (3\alpha \sin(2\pi 200t) + \beta)^2 dt \\
&= 10 \left[\int_{0.1(m-1)}^{0.1m} 9\alpha^2 \sin^2(2\pi 200t) dt + \int_{0.1(m-1)}^{0.1m} 3\alpha\beta \sin(2\pi 200t) dt + \int_{0.1(m-1)}^{0.1m} \beta^2 dt \right] \\
&= \frac{9\alpha^2}{2} + 0 + \beta^2
\end{aligned}$$

In the final equality, the first term is obtained using the same integral as in the previous part. The second term is zero because the interval contains an integer number of carrier periods, so the integral of $\sin(2\pi 200t)$ over the interval is zero. The final term follows from $\int_{0.1(m-1)}^{0.1m} \beta^2 dt = 0.1\beta^2$.

Suppose bit $b_m = 0$. Then $y(t) = \beta$ in the interval $[(m-1)T_b, mT_b]$ and hence

$$p_m = \frac{1}{0.1} \int_{0.1(m-1)}^{0.1m} \beta^2 dt = \beta^2.$$

Perfect decoding is possible as long as

$$\beta^2 < p_{thresh} \leq \frac{9\alpha^2}{2} + \beta^2.$$

(c). Suppose $y(t) = s(t) + \eta(t)$. Here $\eta(t)$ denotes the noise. We have been given following information regarding the noise sequence:

$$\int_{0.1(m-1)}^{0.1m} \sin(2\pi 200t)\eta(t) dt = 0, \quad \text{for all } m > 0,$$

and

$$\int_{0.1(m-1)}^{0.1m} \eta(t)^2 dt \in [0, 0.1\gamma^2], \quad \text{for all } m > 0.$$

For what values of $\gamma > 0$ is error-free decoding possible? What range of thresholds enables this?

Solution: Suppose bit $b_m = 1$. Then $y(t) = 3\sin(2\pi 200t) + \eta(t)$ in the interval $[(m-1)T_b, mT_b]$.

$$\begin{aligned}
p_m &= \frac{1}{0.1}\int_{0.1(m-1)}^{0.1m} (3\sin(2\pi 200t) + \eta(t))^2 dt \\
&= 10\left[\int_{0.1(m-1)}^{0.1m} 9\sin^2(2\pi 200t)dt + \int_{0.1(m-1)}^{0.1m} 3\sin(2\pi 200t)\eta(t)dt + \int_{0.1(m-1)}^{0.1m} \eta(t)^2 dt\right] \\
&= \frac{9}{2} + 0 + 10\int_{0.1(m-1)}^{0.1m} \eta(t)^2 dt
\end{aligned}$$

We know that $\int_{0.1(m-1)}^{0.1m} \eta(t)^2 dt \in [0, 0.1\gamma^2]$ and hence $10\int_{0.1(m-1)}^{0.1m} \eta(t)^2 dt$ lies between 0 and γ^2. So, when $b_m = 1$, p_m lies in the range $[4.5, 4.5 + \gamma^2]$.

Suppose bit $b_m = 0$. Then $y(t) = \eta(t)$ in the interval $[(m-1)T_b, mT_b]$, and hence

$$p_m = \frac{1}{0.1}\int_{0.1(m-1)}^{0.1m} \eta^2 dt,$$

so p_m lies in the range $[0, \gamma^2]$.

Perfect decoding is possible only if the intervals $[0, \gamma^2]$ and $[4.5, 4.5 + \gamma^2]$ do not overlap, which requires $\gamma^2 < 4.5$. For $\gamma < \sqrt{4.5}$, the threshold p_{thresh} can be in the range $(\gamma^2, 4.5]$.

6. Decoding with Imperfect Synchronization

Let b_m for $m = 1, 2, \ldots$ be the sequence of bits we wish to communicate. You first decide to generate a baseband signal using the on-off keying model. You set $T_b = 1$ and $V = 2$. Let us call this function $x(t)$. You then multiply this function with $\sin(2\pi 10t)$ to obtain a signal $s(t)$, i.e., $s(t) = x(t) \times \sin(2\pi 10t)$. The signal $s(t)$ is transmitted and the receiver receives signal $y(t)$. Suppose the receiver wishes to use the power detection and thresholding scheme, but there was an error in synchronization. So to decode each bit m, the receiver computes

$$w_m = \int_{(m-1)+0.1}^{m+0.1} (y(t))^2 dt.$$

(a). The value of w_m now depends on the bits b_m and b_{m+1}. Let $w(b_m, b_{m+1})$ denote the value of w_m obtained for different values of bits b_m and b_{m+1}. Compute $w(0,0)$, $w(0,1)$, $w(1,0)$, and $w(1,1)$. Note that the integral of $\sin^2(2\pi 10t)$ from 0 to 0.1 is 0.05.

Solution: Note that

$$\begin{aligned}
w_m &= \int_{(m-1)+0.1}^{m+0.1} y(t)^2\, dt = \int_{(m-1)+0.1}^{m} y(t)^2\, dt + \int_{m}^{m+0.1} y(t)^2\, dt \\
&= \int_{(m-1)+0.1}^{m} x(t)^2 \sin^2(2\pi 10t)\, dt + \int_{m}^{m+0.1} x(t)^2 \sin^2(2\pi 10t)\, dt \\
&= \int_{(m-1)+0.1}^{m} 4b_m^2 \sin^2(2\pi 10t)\, dt + \int_{m}^{m+0.1} 4b_{m+1}^2 \sin^2(2\pi 10t)\, dt \\
&= 4b_m \int_{0}^{0.9} \sin^2(2\pi 10t)\, dt + 4b_{m+1} \int_{0}^{0.1} \sin^2(2\pi 10t)\, dt.
\end{aligned}$$

When $b_m = b_{m+1} = 0$, both terms are zero, and hence $w(0,0) = 0$. When $b_m = b_{m+1} = 1$, this calculation simplifies to the standard case and we receive $w(1,1) = 2$. When $b_m = 0$ and $b_{m+1} = 1$, we get $w(0,1) = 0.2$. When $b_m = 1$ and $b_{m+1} = 0$, we get $w(1,0) = 1.8$.

(b). Now, to decode the bit $\hat{b}_m$, the receiver uses the thresholding scheme, such that $\hat{b}_m = 1$ if $w_m \geq \alpha$ and 0 otherwise. What property must α satisfy in terms of $w(0,0), w(0,1), w(1,0)$, and $w(1,1)$ to ensure perfect decoding, i.e., $\hat{b}_m = b_m$ for all m. Note that this question can be solved independently of the previous part.

Solution: We want the threshold α such that $w(b_m, b_{m+1}) < \alpha$ when $b_m = 0$ and $w(b_m, b_{m+1}) \geq \alpha$ when $b_m = 1$. Hence, we require

$$\max\{w(0,0), w(0,1)\} < \alpha \leq \min\{w(1,0), w(1,1)\}.$$

Week 7

Error Correction Capabilities and Hamming Code

1. Repetition Code

Consider the repetition code of size $M = 4$ where the bits 00, 01, 10 and 11 are repeated w times.

(a). Suppose $w = 5$. What are the 4 codewords?

Solution: The 4 codewords are 0000000000, 0101010101, 1010101010, 1111111111.

(b). What is the length of each codeword n and the minimum distance d_{min} of the code in terms of w?

Solution: The length of each codeword is $n = 2w$ bits. The minimum distance of the code is $d_{min} = w$ (e.g., consider the codewords corresponding to 00 and 01).

(c). If we want to correct t bit flips, what is the minimum w and n required?

Solution: The minimum distance needs to be at least $2t + 1$ in order to correct t bit flips. And hence a minimum of $w = 2t + 1$ and $n = 2(2t + 1)$ is required to correct t bits.

(d). What is the rate of the code required to correct t bit flips?

Solution: Note that the number of information bits k is independent of the value of w and is fixed at 2. To correct t bit flips, we require $n = 2(2t + 1)$. Hence the rate of the code required to correct t bit flips is $\frac{2}{2(2t+1)} = \frac{1}{2t+1}$. Hence the rate of the code decreases as we increase the code correcting requirements of the code.

(e). If we want to detect t bit flips, what is the minimum w and n required?

Solution: The minimum distance needs to be at least $t + 1$ to detect t flips. And hence a minimum of $w = t + 1$ and $n = 2(t + 1)$ is required to correct t bits.

2. One-Sided Error Channel or Z-channel

You are working with an on-off keying communication system where the power can sometimes fail. You model this as a binary channel that can flip a 1 to a 0, but can never flip a 0 to a 1. In particular, if the transmitter transmits a 0, it is always received perfectly as a 0. If the transmitter transmits a 1, it can be received as a 1 or flipped to a 0. You wish to use an error correcting code to achieve reliable recovery of the transmitted information bits despite these errors. Note that in the decoding you will take advantage of this one-sided nature of the channel, i.e., that 0 is always received correctly.

(a). Suppose the codewords are 111 and 000. How many errors can we correct using this code?

Solution: Note that 000 will always be transmitted as 000. 111 may be received as 000 only if three or more errors are possible (in which case it cannot be decoded correctly). Hence the code can correct a maximum of 2 errors.

(b). Suppose the codewords are 101 and 010. How many errors can we correct using this code?

Solution: If a maximum of 1 error can occur, then 010 can be received as 000 or 010 and 101 can be received as 101 or 100 or 001. Since, the two sets are disjoint, 1 error can be corrected. If a maximum of 2 errors can occur, then both 101 and 010 can be received as 000, and hence decoding is not possible. So, a maximum of 1 errors can be corrected using this code.

(c). Recall the definition of Hamming distance covered in class. Even thought, the Hamming distance of codewords was 3 in both part (a) and part (b), we were able to correct 2 errors in one case and 1 in the other. Suggest a new distance function and give bounds on error correction capabilities based on this distance.

Solution: Since the channel only allows errors of type $1 \to 0$, the usual Hamming distance is not the most relevant measure of separation. What matters is how many 1's must be flipped to 0's to transform one codeword into another.

For $x, y \in \{0,1\}^n$, define the *Z-distance*

$$d_Z(x,y) = \max\Big(|\{i : x_i = 1,\ y_i = 0\}|,\ |\{i : x_i = 0,\ y_i = 1\}|\Big).$$

The first term counts the number of $1 \to 0$ flips needed to turn x into y, and the second counts those needed to turn y into x.

Two distinct codewords x and y are confusable under t errors if both can be reduced (via $1 \to 0$ flips) to a common received word using at most t errors. Since only $1 \to 0$ errors are allowed, any common received word must have 1's only in positions where both x and y have 1's. The closest such word is obtained by keeping exactly those common 1's and setting all other positions to 0.

From x, this requires flipping the positions where $x_i = 1$ and $y_i = 0$, and from y, flipping the positions where $y_i = 1$ and $x_i = 0$. The minimum number of errors needed so that both codewords can reach a common word is therefore

$$\max\Big(|\{i : x_i = 1,\ y_i = 0\}|,\ |\{i : x_i = 0,\ y_i = 1\}|\Big) = d_Z(x,y).$$

Therefore, a code $\mathcal{C}$ can correct up to t errors on the Z-channel if

$$d_Z(x,y) \geq t+1 \quad \text{for all distinct } x, y \in \mathcal{C}.$$

If the minimum Z-distance of the code is $d_Z^{\min}$, then the maximum number of correctable errors is

$$t_{\max} = d_Z^{\min} - 1.$$

For comparison:

- In part (a), $d_Z(111, 000) = 3$, so $t_{\max} = 2$.
- In part (b), $d_Z(101, 010) = 2$, so $t_{\max} = 1$.

This explains why the two codes, despite having the same Hamming distance, have different error-correction capabilities on the Z-channel.

3. Binary Erasure Channel

You are working with a communication system where bits may be erased during transmission. You model this as a *binary erasure channel* (BEC): each transmitted bit is either received correctly or replaced by an erasure symbol "?". Erasures are detectable (i.e., the receiver knows which positions were erased), and no bit is ever flipped.

(a). Suppose the codewords are 111 and 000. How many erasures can be corrected using this code?

Solution: The two codewords differ in all three positions, so $d_{\min} = 3$. If at most two erasures occur, at least one position remains unerased, and that position determines uniquely whether the transmitted word was 111 or 000. If all three positions are erased, the received word is ???, and decoding is impossible. Hence, the code can correct a maximum of 2 erasures.

(b). Suppose the codewords are 101 and 010. How many erasures can be corrected using this code?

Solution: Again, the Hamming distance between the two codewords is 3. If at most two erasures occur, at least one position remains unerased, and since the two codewords differ in every coordinate, that position uniquely determines the transmitted word. If three erasures occur, decoding is impossible. Thus, this code can also correct a maximum of 2 erasures.

(c). Suppose the codewords are 110 and 011. How many erasures can be corrected?

Solution: The Hamming distance is 2. If one erasure occurs, at least two positions remain known, and the codewords differ in at least one of those positions, so decoding is unique. If two erasures occur, it is possible that the only known position is one where both codewords agree, making them indistinguishable. Hence, this code can correct at most 1 erasure.

(d). Recall the definition of Hamming distance. Using Hamming distance, give a general bound on the erasure correction capability of a code over the binary erasure channel.

Solution: Let

$$d_{\min} = \min_{x \neq y \in \mathcal{C}} d_H(x, y)$$

be the minimum Hamming distance of the code. If two distinct codewords differ in $d_{\min}$ positions, then they become indistinguishable only if all those differing positions are erased. Therefore, as long as the number of erasures satisfies

$$t < d_{\min},$$

unique decoding is guaranteed. Hence, the maximum guaranteed number of correctable erasures is

$$t_{\max} = d_{\min} - 1.$$

4. Hamming Code

Consider the Hamming code studied in class with information bits b_1, b_2, b_3, b_4 and parity bits p_1, p_2, p_3, where

$$p_1 = b_1 \oplus b_2 \oplus b_3,$$

$$p_2 = b_1 \oplus b_2 \oplus b_4,$$

and

$$p_3 = b_2 \oplus b_3 \oplus b_4.$$

The codeword is of the form $b_1 b_2 b_3 b_4 p_1 p_2 p_3$.

(a). Perform encoding for the $b_1 b_2 b_3 b_4 = 0101$ to obtain the codeword.

Solution:

$$\begin{aligned} p_1 &= b_1 \oplus b_2 \oplus b_3 = 0 \oplus 1 \oplus 0 = 1 \\ p_2 &= b_1 \oplus b_2 \oplus b_4 = 0 \oplus 1 \oplus 1 = 0 \\ p_3 &= b_2 \oplus b_3 \oplus b_4 = 1 \oplus 0 \oplus 1 = 0. \end{aligned}$$

So the codeword is 0101100.

(b). You are communicating over a noisy channel where at most one bit flip can occur during transmission. At the receiver you receive 1010101. What was the original message, i.e., the 4 information bits?

Solution: We have received 1010101. We will compute the XORs of bits corresponding to the three circles given in the lecture notes.

$$\begin{aligned} p_1 \oplus b_1 \oplus b_2 \oplus b_3 &= 1 \oplus 0 \oplus 1 \oplus 1 = 1 \\ p_2 \oplus b_1 \oplus b_2 \oplus b_4 &= 1 \oplus 0 \oplus 0 \oplus 0 = 1 \\ p_3 \oplus b_2 \oplus b_3 \oplus b_4 &= 0 \oplus 1 \oplus 0 \oplus 1 = 0. \end{aligned}$$

This means there is a bit flip in the first two circles. As we know that a maximum of one bit flip can occur, the bit flipped is b_1. Hence b_1 was actually 0 and was received as 1. So the actual message was 0010. The corresponding codeword is 0010101. We can verify that the received code is at a distance of 1 from the original codeword.

(c). You are communicating over a noisy channel where at most one bit flip can occur during transmission. At the receiver you receive 0010100. What was the original message, i.e., the 4 information bits?

Solution: We have received 0010100. We will compute the XORs of bits corresponding to the three circles given in the lecture notes.

$$\begin{aligned} p_1 \oplus b_1 \oplus b_2 \oplus b_3 &= 0 \oplus 0 \oplus 1 \oplus 1 = 0 \\ p_2 \oplus b_1 \oplus b_2 \oplus b_4 &= 0 \oplus 0 \oplus 0 \oplus 0 = 0 \\ p_3 \oplus b_2 \oplus b_3 \oplus b_4 &= 0 \oplus 1 \oplus 0 \oplus 0 = 1. \end{aligned}$$

This means there is a bit flip in the third circle. As we know that a maximum of one bit flip can occur, the bit flipped is p_3. Hence p_3 was actually 1 and was received as 0. So the actual message was 0010. The corresponding codeword is 0010101. We can verify that the received code is at a distance of 1 from the original codeword.

(d). You are working with an erasure channel. This channel can erase some of the transmitted bits. Note that every bit is either received perfectly or gets erased but cannot be flipped. At the receiver, you receive ???0101. What was the original message, i.e., just the 4 information bits?

Solution: We receive ???0101. This gives us the information that

$$\begin{aligned}
b_4 &= 0 \\
p_1 &= 1 = b_1 \oplus b_2 \oplus b_3 \\
p_2 &= 0 = b_1 \oplus b_2 \oplus b_4 \\
p_3 &= 1 = b_2 \oplus b_3 \oplus b_4.
\end{aligned}$$

Substituting $b_4 = 0$ in the last two equations, we get $b_1 \oplus b_2 = 0$ and $b_2 \oplus b_3 = 1$. We know that $b_1 \oplus b_2 = 0$ and that $b_1 \oplus b_2 \oplus b_3 = 1$. This implies that $b_3 = 1$. Substituting this in $b_2 \oplus b_3 = 1$ gives us that $b_2 = 0$. Substituting this in $b_1 \oplus b_2 = 0$ gives us $b_1 = 0$. So the original message was 0010, and the corresponding codeword was 0010101.

5. Hamming Bound

Recall from class that we showed the Hamming bound on M using a ball packing argument. This bound is given by:

$$M \leq \frac{\text{Total number of binary sequences of length } n}{\text{Number of sequences in the Hamming ball of radius } t \text{ around a codeword}}.$$

Here $t = \lfloor \frac{d_{min}-1}{2} \rfloor$ is the number of bit flips a code with minimum distance d_{min} can correct. In this problem, we will work with $d_{min} = 5$.

(a). Recall that a Hamming ball of radius t consists of all sequences with Hamming distance $\leq t$ from a given codeword. For $n = 8$, find the number of sequences in a Hamming ball of radius 2 around a codeword.

Solution: Sequences in a Hamming ball of radius 2 around a codeword include sequences which are obtained after 0 bit flips, sequences which are obtained after 1 bit flip, and sequences which are obtained after 2 bit flips. There is just one sequence obtained after 0 bit flips (the original codeword). There are 8 sequences obtained after 1 bit flip (there are a total of 8 bits, and hence we obtain a new sequence at distance 1 by flipping each bit). There are $\binom{8}{2} = \frac{8!}{2!(8-2)!} = 28$ sequences obtained after 2 bit flips (there are a total of 8 bits, and we obtain a new sequence at distance 2 for each 2 bits that we choose to flip). So, the total number of sequences in a Hamming ball of radius 2 around a codeword is $1 + 8 + \binom{8}{2} = 1 + 8 + 28 = 37$.

(b). Use the Hamming bound given above to determine the upper bound on the **rate** of any code for $n = 50$ and $d_{min} = 5$.

Solution: Total number of binary sequences of length 50 are 2^{50}. Now, $d_{min} = 5$ corresponds to $t = 2$. Number of sequences in a Hamming ball of radius 2 around a codeword are $1+50+\binom{50}{2} = 1+50+1225 = 1276$. So the above bound gives us $M \leq \frac{2^{50}}{1276}$. The rate is defined as $\log_2(M)/n$. Now, $\log_2(M) = \log_2\left(\frac{2^{50}}{1276}\right) = 50-\log_2(1276)$. So the upper bound on the rate is $\frac{50-\log_2(1276)}{50} = 0.793$.