ECT306 INFORMATION THEORY & CODING

Module 2-Channels and Channel Coding

- Discrete memoryless channels. Capacity of discrete memoryless channels. Binary symmetric channels (BSC), Binary Erasure channels (BEC). Capacity of BSC and BEC. Channel code. Rate of channel code. Shannon's channel coding theorem (both achievability and converse without proof) and operational meaning of channel capacity.
- Modeling of Additive White Gaussian channels. Continuous-input channels with average power constraint. Differential entropy. Differential Entropy of Gaussian random variable. Relation between differential entropy and entropy. Shannon-Hartley theorem (with proof – mathematical subtlities regarding power constraint may be overlooked).
- Inferences from Shannon Hartley theorem – spectral efficiency versus SNR per bit, power-limited and bandwidth-limited regions, Shannon limit, Ultimate Shannon limit.

RATE OF INFORMATION TRANSMISSION OVER A DISCRETE CHANNEL

- We have the entropy of the input symbols given by

$$H(X) = \sum_{i=1}^{r} P(x_i)\, log_2\left(\frac{1}{P(x_i)}\right)$$

- Consider a discrete memoryless channel accepting symbols at the rate of , r_s message symbols/sec .
- The average rate at which information is going into the channel is given by

$$R_{in} = H(X) r_s \ \ bits/sec$$

- At the receiver, it is not possible to reconstruct the input symbol sequence with certainty by operating on the receiving sequence.
- This is due to errors introduced when the signals pass through the channel.
- Some amount of information is lost in the channel due to noise.
- This information which is lost in the channel has been called equivocation H(X/Y). Hence the net amount of information which is the mutual information I(X;Y) , is given by equation as

$$I(X;Y) = H(X) - H\left(X/_Y\right) \quad bits/message\ symbol$$

- The average rate of information transmission R_t is given by,

$$R_t = I(X;Y)\, r_s \quad bits/sec$$

$$R_t = \left[H(X) - H\left({}^{X}/_{Y}\right)\right] r_s \quad bits/sec$$

- or

$$R_t = \left[H(Y) - H\left({}^{Y}/_{X}\right)\right] r_s \quad bits/sec$$

CAPACITY OF A DISCRETE MEMORYLESS CHANNEL

- The capacity of a discrete memoryless noisy channel is defined as the maximum possible rate of information transmission over the channel. The maximum rate of transmission occurs when the source is matched to the channel.
- Therefore channel capacity C is defined as

$$C = Max(R_t)$$

$$C = Max\left[H(X) - H\left({}^{X}/_{Y}\right)\right] r_s$$

SHANNON'S THEOREM ON CHANNEL CAPACITY (SHANNON'S SECOND THEOREM)

Rate of information transmission given as

$$R_t = \left[H(X) - H\left({}^{X}/_{Y}\right)\right] r_s \quad bits/sec$$

And channel capacity equation as

$$C = Max\left[\left[H(X) - H\left({}^{X}/_{Y}\right)\right] r_s\right] bits/sec$$

Positive Statement

- Shannon's theorem on channel capacity states that "when the rate of information transmission $\boldsymbol{R_t \leq C}$, then there exists a coding technique which enables transmission over a channel with as small a probability of error as possible, even in the presence of noise in the channel."

This theorem indicates that for $R_t \leq C$, transmission of information is achieved without errors, even in the presence of noise.

Negative Statement

"If $R_t > C$, then reliable transmission of information is not possible without errors. Thus when $R_t > C$, then the errors cannot be controlled by any coding technique and the probability of error of receiving the correct message becomes close to unity."

CHANNEL EFFICIENCY AND REDUNDANCY

The channel efficiency denoted by η_{ch} is given by

$$\eta_{ch} = \frac{R_t}{C} \times 100\%$$

Substituting for $R_t and\ C$, we get

$$\eta_{ch} = \frac{\left[H(X) - H\left(X/_{Y}\right)\right] r_s}{Max\left[\left[H(X) - H\left(X/_{Y}\right)\right] r_s\right]} \times 100\%$$

or

$$\eta_{ch} = \frac{I(X;Y)}{Max[I(X;Y)]} \times 100\%$$

The channel redundancy is denoted by $R_{\eta_{ch}}$ is given by

$$R_{\eta_{ch}} = 1 - \eta_{ch}$$

COMMUNICATION CHANNELS

- In the channel diagram channel input belongs to source alphabet A with 3 symbols and channel output belongs to other alphabet B with 4 symbols.

$$P_{11} = P(b_1/a_1)$$
$$P_{12} = P(b_2/a_1)$$
$$P_{13} = P(b_3/a_1)$$
$$P_{14} = P(b_4/a_1)$$

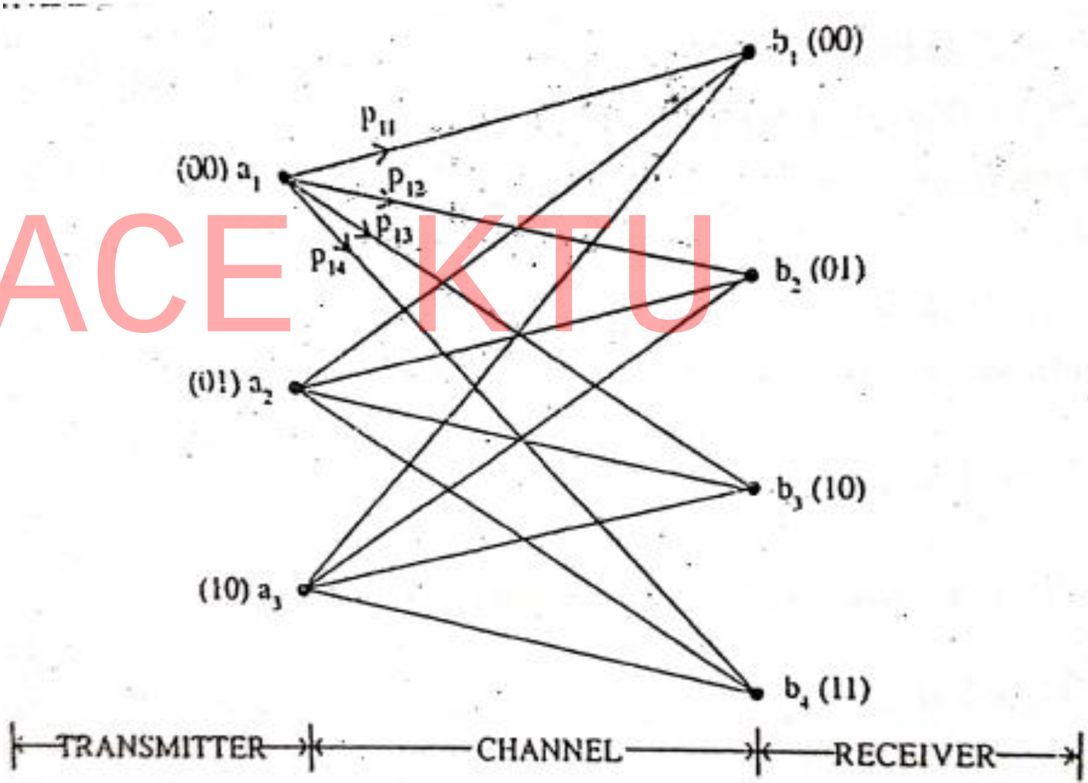


Fig. 4.3 : Illustrating channel or noise diagram

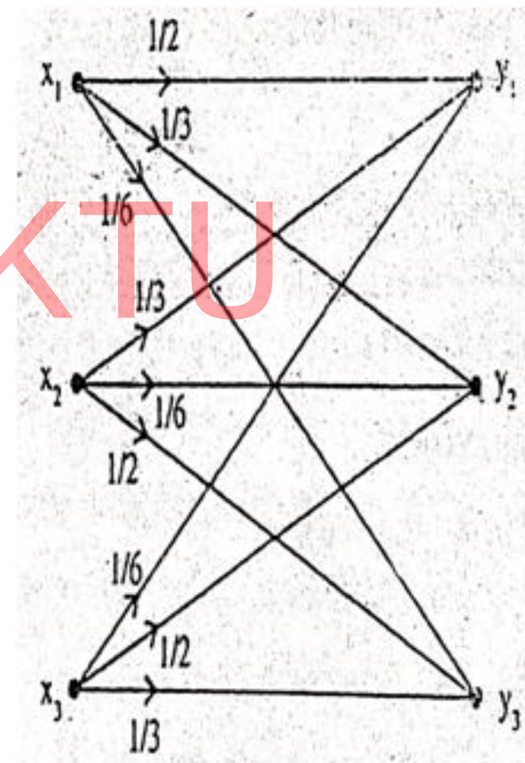
SPECIAL TYPE OF CHANNELS

- Symmetric/Uniform Channels
- Binary Symmetric Channels (BSC)
- Binary Erasure Channels (BEC)

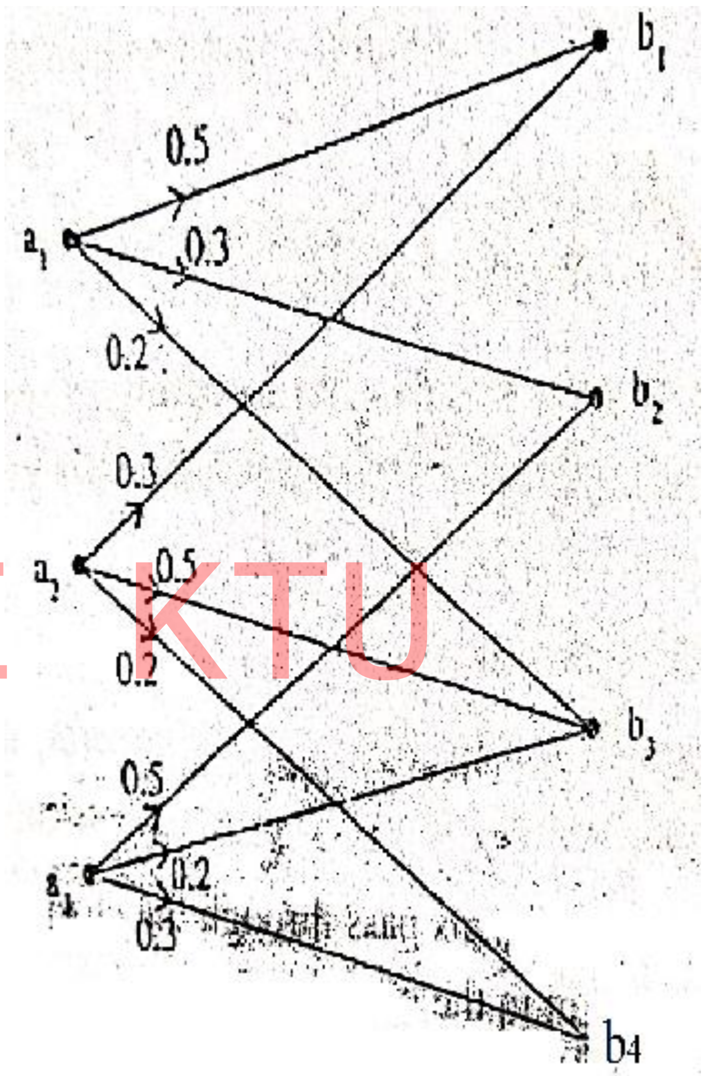
Symmetric or Uniform Channel

- A channel is said to be symmetric or uniform channel, if the second and subsequent rows of the channel matrix contain the same elements as that of first row, but in a different order.

$$P(Y/X) = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/2 \\ 1/6 & 1/2 & 1/3 \end{bmatrix} \end{array}$$



$$
P(B/A) = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array}
\begin{array}{c}
\begin{array}{cccc} b_1 & b_2 & b_3 & b_4 \end{array} \\
\left[\begin{array}{cccc} 0.5 & 0.3 & 0.2 & 0 \\ 0.3 & 0 & 0.5 & 0.2 \\ 0 & 0.5 & 0.2 & 0.3 \end{array}\right]
\end{array}
$$



Channel Capacity of Symmetric or Uniform Channel

- Consider a symmetric or uniform channel, there are ‘s’ number of output messages and ‘r’ number of input messages.
- The channel matrix or probability transition matrix is given by

$$P(Y/X) = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_r \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & \cdots & y_s \end{array} \\ \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & \cdots & P(y_s/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & \cdots & P(y_s/x_2) \\ \vdots & & & \\ P(y_1/x_r) & P(y_2/x_r) & \cdots & P(y_s/x_r) \end{bmatrix} \end{array}$$

- Since it is a symmetric or uniform channel each rows of channel matrix contain same elements but in different order.

- So channel matrix can be written as

$$
P(y/x) = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_r \end{array}
\begin{array}{c}
\begin{array}{ccccc} y_1 & y_2 & y_3 & \cdots & y_s \end{array} \\
\begin{bmatrix}
P_1 & P_2 & P_3 & \cdots & P_s \\
P_2 & P_3 & P_{s-1} & \cdots & P_1 \\
\vdots & & & & \\
P_3 & P_{s-1} & P_{s-2} & \cdots & P_2
\end{bmatrix}
\end{array}
$$

- Where $P_1, P_2, P_3 \ldots\ldots, P_{s-3}, P_{s-2}, P_{s-1}, P_s$ are the conditional probabilities $P(y_j/x_i)$.
- The sum of all the elements in any row of the channel matrix is equal to unity.

$$\sum_{j=1}^{s} P\left(\frac{y_j}{x_i}\right) = 1 \qquad\qquad \sum_{j=1}^{s} P_j = 1$$

The equivocation H(Y/X) is given by

$$H\left(Y/_X\right) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i, y_j)\, log_2 \frac{1}{P\left(y_j/_{x_i}\right)}$$

$$H\left(Y/_X\right) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i)\, P\left(y_j/_{x_i}\right) log_2 \frac{1}{P\left(y_j/_{x_i}\right)}$$

$$H\left(Y/_X\right) = \sum_{i=1}^{r}\sum_{j=1}^{s} P(x_i)\, P_j\ log_2 \frac{1}{P_j}$$

$$H\left(Y/_X\right) = \left[\sum_{i=1}^{r} P(x_i)\right] \left[\sum_{j=1}^{s} P_j\ log_2 \frac{1}{P_j}\right]$$

We know

$$\sum_{i=1}^{r} P(x_i) = 1$$

Let $$\sum_{j=1}^{s} P_j \, log_2 \frac{1}{P_j} = h = a\ constant$$

$$H\left({}^{Y}/_{X}\right) = [1]\ [h]$$

or $$H\left({}^{Y}/_{X}\right) = h = \sum_{j=1}^{s} P_j \, log_2 \frac{1}{P_j}$$

The mutual information I(X;Y) is given by

$$I(X;Y) = H(Y) - H\left({}^{Y}/_{X}\right)$$

$$I(X;Y) = H(Y) - h$$

The channel capacity is given by

$$C = Max\,[I(X;Y)]r_s$$

$$C = Max\,[H(Y) - h]r_s$$

h is a constant, therefore

$$C = [Max\,H(Y) - h]r_s$$

- The entropy of a output symbol becomes maximum if and only when all the received symbols become equiprobable.
- Since there are *s* number of output symbols, we have

$$Max\,H(Y) = log_2 s$$

So we get channel capacity of symmetric or uniform channel as,

$$C = [log_2 s - h]r_s \quad bits/sec$$

Example 1

For the channel matrix shown, find the channel capacity

$$P(b_j/a_i) = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} b_1 & b_2 & b_3 \end{array} \\ \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/2 \\ 1/6 & 1/2 & 1/3 \end{bmatrix} \end{array}$$

- **Solution**

$$H(Y/X) = h = \sum_{j=1}^{s} P_j \log_2 \frac{1}{P_j}$$

Channel capacity is given by

$$C = [\log_2 s - h] r_s \quad bits/sec$$

r_s is 1 message-symbol/sec

Example 2

For the channel matrix given below, compute the channel capacity

$$P(Y/X) = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.2 & 0.6 \end{bmatrix} \text{ with } r_s = 1000 \text{ symbols/sec}$$

- **Solution**

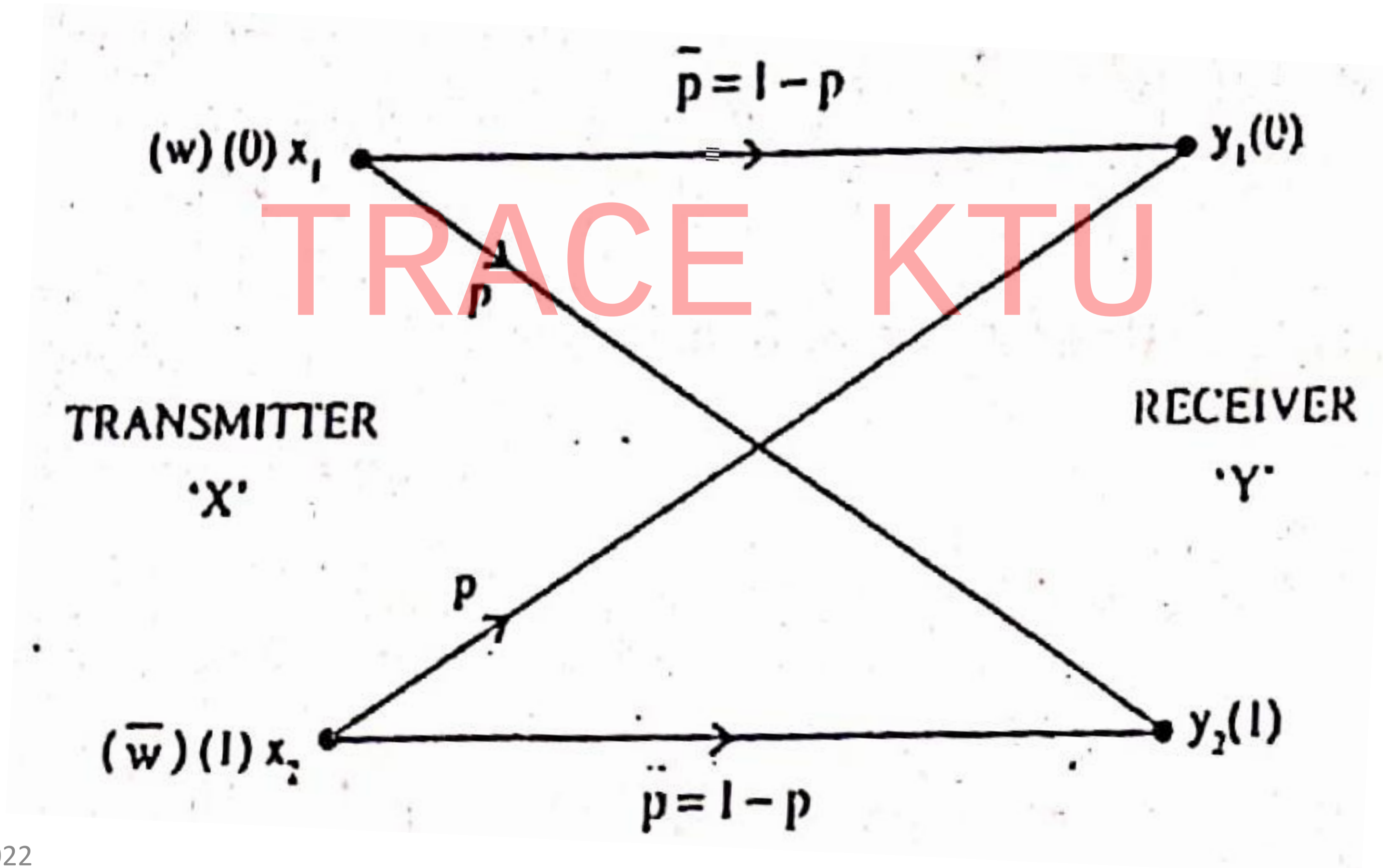
$$H(Y/X) = h = \sum_{j=1}^{s} P_j \, log_2 \frac{1}{P_j}$$

Channel capacity is given by

$$C = [log_2 s - h] r_s \quad bits/sec$$

Binary Symmetric Channel (BSC)

- A symmetric channel which has two input and two output is called Binary Symmetric Channel.
- Most commonly and widely used channel



Let $P(x1) = w$ and $P(x2) = 1 - w = \bar{w}$

Let p = probability of error

= probability of reception of ‘1’ when ‘0’ is transmitted

= probability of reception of ‘0’ when ‘1’ is transmitted

Channel matrix of a BSC can be written as

$$P(Y/X) = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix} \end{array}$$

$$P(Y/X) = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

- Since it is a symmetric channel the equivocation is given by

$$H\left({}^{Y}/_{X}\right) = h = \sum_{j=1}^{2} P_j \, log_2 \frac{1}{P_j}$$

$$H\left({}^{Y}/_{X}\right) = h = \bar{p}\log\frac{1}{\bar{p}} + p\log\frac{1}{p}$$

To Find H(Y)

Entropy of output symbol is given by

$$H(Y) = \sum_{j=1}^{2} P(y_j) \, log_2 \frac{1}{P(y_j)}$$

$$= P(y_1) \, log_2 \frac{1}{P(y_1)} + P(y_2) \, log_2 \frac{1}{P(y_2)}$$

$$P(y_1) = P\left({y_1}/{x_1}\right)P(x_1) + P\left({y_1}/{x_2}\right)P(x_2)$$

$$= \bar{p}\, w + p\, \bar{w}$$

$$P(y_2) = P\left({y_2}/{x_1}\right)P(x_1) + P\left({y_2}/{x_2}\right)P(x_2)$$

$$= p\, w + \bar{p}\, \bar{w}$$

$$H(Y) = (\bar{p}\, w + p\, \bar{w}) \log \frac{1}{(\bar{p}\, w + p\, \bar{w})} + (p\, w + \bar{p}\, \bar{w}) \log \frac{1}{(p\, w + \bar{p}\, \bar{w})}$$

- Mutual information I(X;Y) is given by

$$I(X;Y) = H(Y) - H\left({Y}/{X}\right)$$

$$I(X;Y) = (\bar{p}\,w + p\,\bar{w}\,)\log\frac{1}{(\bar{p}\,w + p\,\bar{w}\,)} + \;(p\,w + \bar{p}\,\bar{w})\log\frac{1}{(p\,w + \bar{p}\,\bar{w})}$$
$$-\left(\bar{p}\log\frac{1}{\bar{p}} + p\log\frac{1}{p}\right)$$

- Since BSC is a symmetric channel, the channel capacity is found by using the general equation of symmetric channel capacity.

$$C = [log_2 s - h]r_s$$

- In case of BSC no of output symbols s=2

$$C = [log_2 2 - h]r_s$$

$$C = [1 - h]r_s$$

$$C = \left[1 - \left(\bar{p}\log\frac{1}{\bar{p}} + p\log\frac{1}{p}\right)\right] r_s \;\; bits/sec$$

- When the input symbol probability $w = \bar{w} = \frac{1}{2}$
- Mutual information I(X;Y) becomes

$$I(X;Y) = (\bar{p}\, w + p\, \bar{w}\,) \log \frac{1}{(\bar{p}\, w + p\, \bar{w}\,)} + (p\, w + \bar{p}\, \bar{w}) \log \frac{1}{(p\, w + \bar{p}\, \bar{w})}$$

$$- \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)$$

$$= \left(\bar{p}\, \frac{1}{2} + p\, \frac{1}{2}\right) \log \frac{1}{(\bar{p}\, \frac{1}{2} + p\, \frac{1}{2})} + \left(p\, \frac{1}{2} + \bar{p}\, \frac{1}{2}\right) \log \frac{1}{(p\, \frac{1}{2} + \bar{p}\, \frac{1}{2})} - \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)$$

$$= \left(\frac{1}{2} + \frac{1}{2}\right) - \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)$$

$$I(X;Y) = 1 - \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)$$

- When the input symbols become equiprobable, the mutual information maximizes and becomes equal to channel capacity C.

Example 1 (Previous Unvi Question)

Given a binary symmetric channel with $P(Y/X) = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$ and $P(x_1) = \frac{2}{3}; P(x_2) = \frac{1}{3}$. Calculate the mutual information and channel capacity.

Solution:

Mutual information $I(X;Y) = H(Y) - H\left(\frac{Y}{X}\right)$

$$H(Y) = (\bar{p}\, w + p\, \bar{w}\,) \log \frac{1}{(\bar{p}\, w + p\, \bar{w}\,)} + (p\, w + \bar{p}\, \bar{w}) \log \frac{1}{(p\, w + \bar{p}\, \bar{w})}$$

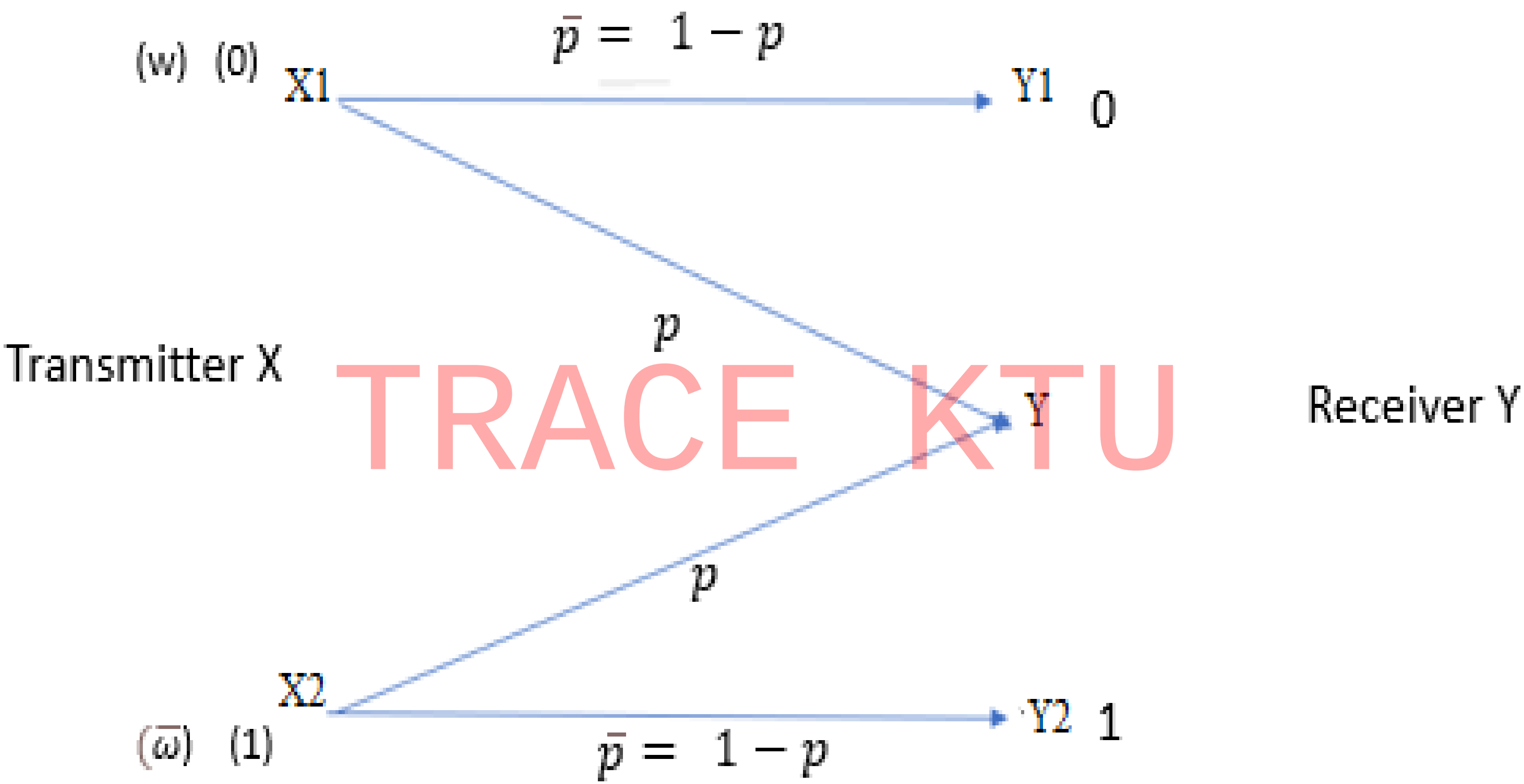
$$H\left(\frac{Y}{X}\right) = h = \bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}$$

Channel capacity $C = \left[1 - \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)\right] r_s \; bits/sec$

BINARY ERASURE CHANNEL (BEC)

- Whenever an error occurs, the symbol will be received as “y”.
- No decision will be made about the information
- An immediate request will be made through a reverse channel for retransmission (ARQ – Automatic Repeat Request) of the transmitted signal till a correct symbol is received at the output.
- Since the error is totally erased in this type of channel, it is called ‘Binary Erasure Channel’.
- The disadvantage with this is the requirement of a reverse channel.

Channel Diagram of BEC



Channel Capacity of BEC

Channel matrix of a BEC is,

$$P(Y/X) = P(y_j/x_i) = \begin{array}{c c} & \begin{array}{ccc} y_1 & y & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{pmatrix} \overline{p} & p & 0 \\ 0 & p & \overline{p} \end{pmatrix} \end{array}$$

Let $P(x_1) = w, P(x_2) = \overline{w}$

Then $w + \overline{w} = 1$ and $p + \overline{p} = 1$

$$H(Y/X) = h = \sum_{j=1}^{s} p_j \, log\frac{1}{p_j}$$

$$H(Y/X) = h = \overline{p} \, log\frac{1}{\overline{p}} + p \, log\frac{1}{p}$$

$$H(X) = \sum_{i=1}^{2} P(x_i)\, log \frac{1}{P(x_i)}$$

$$H(X) = w\, log \frac{1}{w} + \overline{w}\, log \frac{1}{\overline{w}}$$

$$P(x_i, y_j) = P(x_i) P(y_j / x_i)$$

$$P(X, Y) = P(x_i, y_j) = \begin{array}{c c} & \begin{array}{ccc} y_1 & y & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{pmatrix} \overline{p}w & p\,w & 0 \\ 0 & p\overline{w} & \overline{p}\,\overline{w} \end{pmatrix} \end{array}$$

$$P(y_1) = \overline{p}w$$

$$P(y) = pw + p\overline{w} = p(w + \overline{w}) = p$$

$$P(y_2) = \overline{p}\,\overline{w}$$

$$P(x_i/y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

$$P(X/Y) = \begin{pmatrix} (\overline{p}\, w)/(\overline{p}\, w) & pw/p & 0 \\ 0 & p\,\overline{w}/p & (\overline{p}\,\overline{w})/(\overline{p}\,\overline{w}) \end{pmatrix}$$

$$P(X/Y) = P(x_i/y_j) = \begin{pmatrix} 1 & w & 0 \\ 0 & \overline{w} & 1 \end{pmatrix}$$

$$H(X/Y) = \sum_{i=1}^{r=2} \sum_{j=1}^{s=3} P(x_i, y_j)\, log \frac{1}{P(x_i/y_j)}$$

$$= \overline{p}\, w\, log(1) + p\, w\, log \frac{1}{w} + p\, \overline{w}\, log \frac{1}{\overline{w}} + \overline{p}\, \overline{w} log(1)$$

$$H(X/Y) = p[w\, log \frac{1}{w} + \overline{w}\, log \frac{1}{\overline{w}}]$$

Mutual information of BEC is given by

$$I(X:Y) = H(X) - H(X/Y)$$

$$= w\,log\frac{1}{w} + \overline{w}\,log\frac{1}{\overline{w}} - p[w\,log\frac{1}{w} + \overline{w}\,log\frac{1}{\overline{w}}]$$

$$= [1-p][w\,log\frac{1}{w} + \overline{w}\,log\frac{1}{\overline{w}}]$$

$$I(X:Y) = \overline{p}\,H(X)$$

The channel capacity of BEC is given by,

$$C = Max[I(X:Y)]$$

$$= Max[\overline{p}\, H(X)]$$

$$= \overline{p}\, Max[H(X)] = \overline{p}\, H(X)_{max}$$

$$= \overline{p} log2 \qquad H(X)_{max} = log\, q$$

$$C = \overline{p}$$

Example 1 (Previous Unvi Question) - BSC

Given a binary symmetric channel with $P(Y/X)=\begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$ and $P(x_1)=2/3; P(x_2)=1/3$. Calculate the mutual information and channel capacity.

Solution:

$$P(Y/X) = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

$$\boldsymbol{P(x1) = w \text{ and } P(x2) = 1 - w = \bar{w}}$$

Mutual information

$$I(X;Y) = H(Y) - H\left(Y/_X\right)$$

$$H(Y) = (\bar{p}\, w + p\, \bar{w}\,) \log \frac{1}{(\bar{p}\, w + p\, \bar{w}\,)} + \; (p\, w + \bar{p}\, \bar{w}) \log \frac{1}{(p\, w + \bar{p}\, \bar{w})}$$

= 0.9799 bits/message symbol

$$H(Y/X) = h = \; \bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}$$

= 0.8113 bits/message symbol

Channel capacity

$$C = \left[1 - \left(\bar{p} \log \frac{1}{\bar{p}} + p \log \frac{1}{p}\right)\right] r_s \;\; bits/sec$$

= 0.1887 bits/message symbol

Example 2 - BSC

Q. A message source produces two independent symbols A and B with probabilities P(A)=0.4 and P(B)=0.6. Calculate the efficiency of the source and hence its redundancy. If the symbols are received in average with 4 in every 100 symbols in error, calculate the transmission rate of the system. Draw its channel diagram also.

Solution:

$$H(X) = P(A)\log\frac{1}{P(A)} + P(B)\log\frac{1}{P(B)}$$

$$= 0.97095 \text{ bits/message-symbol}$$

$$H(X)_{max} = \log q = \log 2 = 1 \text{ bits/message-symbol}$$

$$\therefore \text{Source efficiency } \eta_s = \frac{H(X)}{H(X)_{max}} \qquad \eta_s = 97.095\%$$

$$\text{Source redundancy } R_{\eta_s} = 2.905\%$$

p = probability of error

$$= \frac{4}{100} = 0.04$$

$$\therefore \quad p = 0.04 \text{ and } \bar{p} = 0.96$$

It is given that $w = P(A) = 0.4$ and $\bar{w} = P(B) = 0.6$

$$I(X;Y) = (\bar{p}\,w + p\,\bar{w}\,)\log\frac{1}{(\bar{p}\,w + p\,\bar{w}\,)} + (p\,w + \bar{p}\,\bar{w})\log\frac{1}{(p\,w + \bar{p}\,\bar{w})}$$

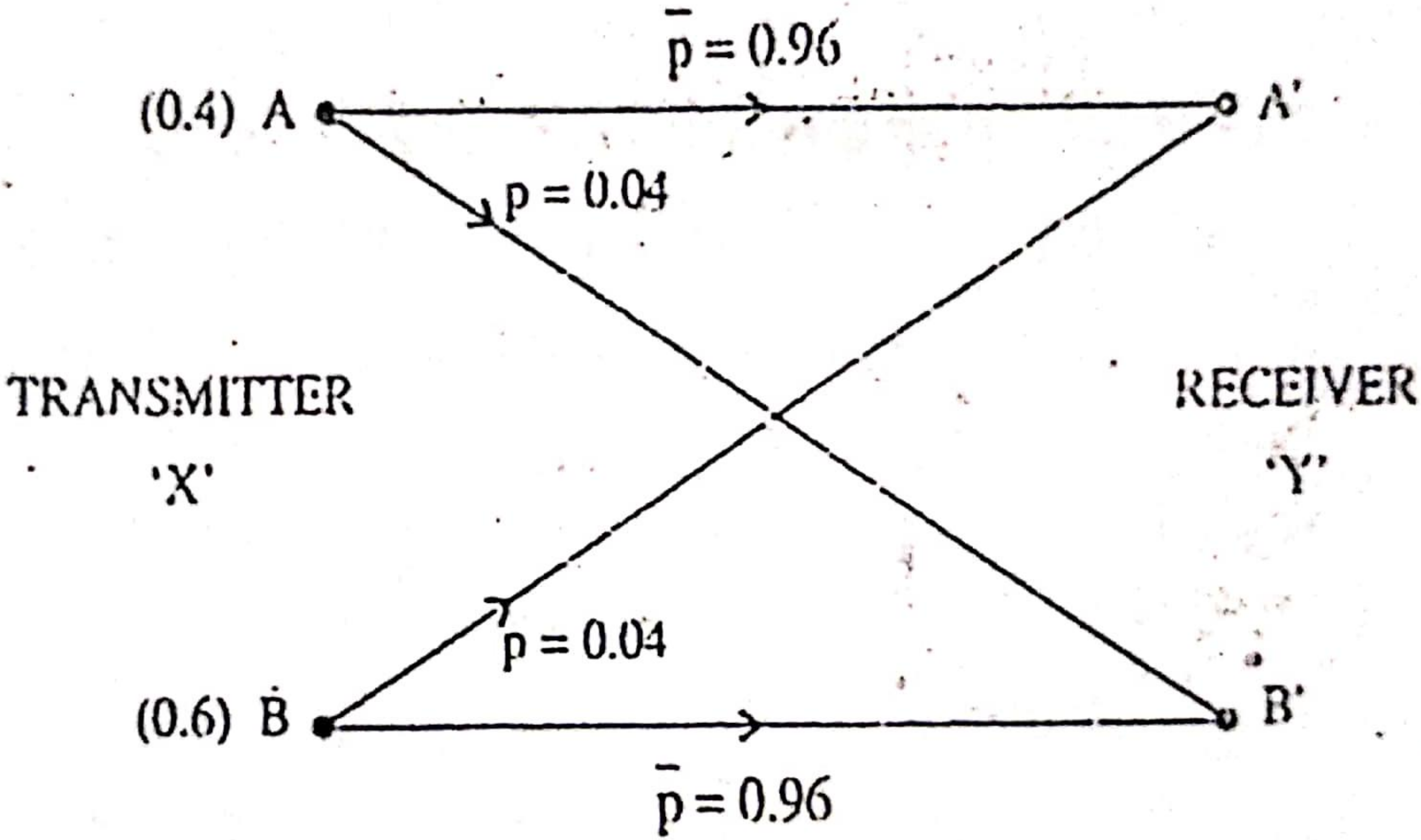
$$- \left(\bar{p}\log\frac{1}{\bar{p}} + p\log\frac{1}{p}\right)$$

= 0.7331 bits/messag e symbol

Transmission rate $\boldsymbol{R_t = I(X;Y)r_s}$ $\boldsymbol{r_s = 100\ symbols/sec}$

= 73.31 bits/sec

Channel Diagram



Continuous Sources and Channels

Differential Entropy-Entropy of continuous signals

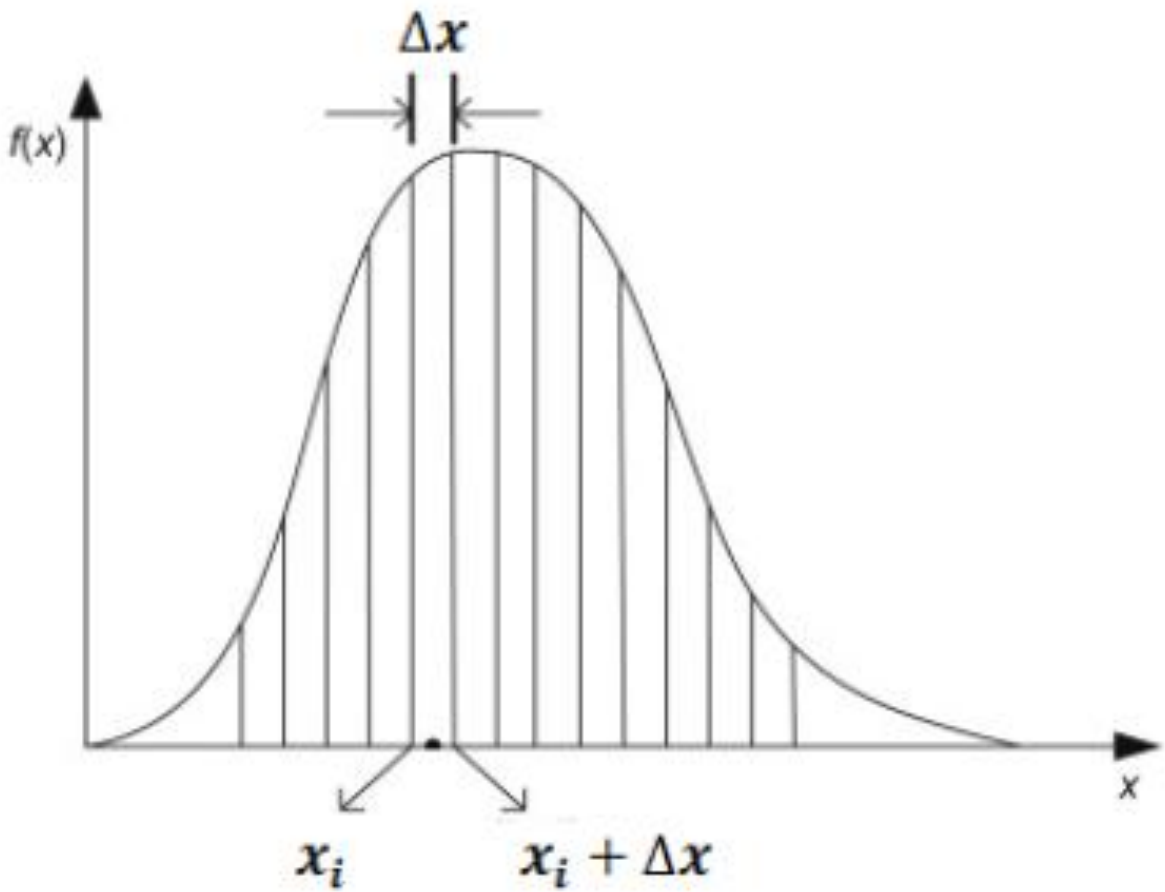
- The entropy in the case of discrete message symbols is given by

$$H(S) = \sum_{i=1}^{q} P(S_i) \log \frac{1}{P(S_i)}$$

Let X(t) be a continuous random process and X is its ensembles taken by sampling X(t) and probability density function $f_x(x)$.

Assume the continuous random variable X, to be a limiting form of a discrete random variable which assumes a value $x_i = i\Delta x$ where $i = 0, \pm 1, \pm 2, ...$ and Δx approaches zero.

By definition, the continuous random variable X assumes a value in the interval $(x_i, x_i + \Delta x)$ with probability $f_x(x_i)\Delta x$.



Δx
$f(x)$
x
x_i
$x_i + \Delta x$

Permitting Δx to approach zero, the ordinary entropy of the CRV X can be written in the limits as follows,

$$H(X) = \lim_{\Delta x \to 0} \sum_i f_x(x_i)\Delta x \, \log \frac{1}{f_x(x_i)\Delta x}$$

$$H(X) = \lim_{\Delta x \to 0} \left[\sum_i f_x(x_i)\Delta x \, \log \frac{1}{f_x(x_i)} - \sum_i f_x(x_i)\Delta x \, \log \Delta x \right]$$

As $\Delta x \to 0$, $\quad f_x(x_i)\Delta x \to f(x)dx$ and the summation can be replaced by integration.

$$H(X) = \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)} dx - \lim_{\Delta x \to 0} \log \Delta x \int_{-\infty}^{\infty} f(x) dx$$

Since $f(x)$ is a probability density function (pdf), we have $\int_{-\infty}^{\infty} f(x)dx = 1$

$$H(X) = \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)} dx - \lim_{\Delta x \to 0} \log \Delta x$$

In the above eqn, as Δx approaches zero, $\log \Delta x$ approaches infinity.

- That means the entropy of a continuous random variable is infinitely large.
- i.e, the uncertainty associated with a CRV is of the order of infinity.
- Therefore we take H(X) as '*differential entropy*', with the term $-\log \Delta x$ serving as reference.
- Therefore the differential entropy can be defined as,

$$H(X) = \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)} dx \quad bits/sample$$

Maximization of Entropy

- In the case of discrete source symbols, the entropy becomes maximum when all the symbols are equiprobable.

$$H = log_2 K$$

- In the case of practical continuous sources, there may be different constraints such as peak signal constraint, average signal constraint, average power restriction or peak power limitation etc.
- The objective is to maximize the entropy function subjected to such restrictions.
- Normalizing constraint which follows from the condition of pdf is

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Peak Signal Limitation

- When the signal is limited to ±M, then

$$\int_{-M}^{+M} f(x)dx = 1$$

- This kind of signal limitation is found in practical cases of AM,FM and pulse modulation techniques.
- The entropy becomes maximum, under peak signal constraint, if the signal is **uniformly distributed**.

i.e.

$$f(x) = \frac{1}{2M}$$

- Maximum value of Entropy is given by,

$$H(X)_{max} = \log 2M \qquad bits/sample$$

- If M is the peak signal voltage, then the peak signal power P_m is given by,

$$P_m = M^2$$

- Then, maximum entropy is given by,

$$H(X)_{max} = \frac{1}{2}\log 4P_m \qquad bits/sample$$

- If the signal is bandlimited to B Hz and is sampled at Nyquist rate r_s=2B samples/sec , then maximum entropy is given by,

$$H(X)_{max} = B\log 4P_m \qquad bits/sample$$

Average signal limitation

- Under this limitation we have

$$\int_{-\infty}^{\infty} x f(x) dx = \mu = E[X] = a\ constant$$

- This kind of limitation come across in pulse amplitude modulation and also in analog amplitude modulation, with average carrier amplitude limitation.
- If the average value equal to λ, then the ccontinuous entropy of X maximizes when it is **exponentially distributed** with parameter 1/λ .

i.e.,

$$\boldsymbol{f(x) = \frac{1}{\lambda}\, e^{-(1/\lambda)x}}$$

- The maximum entropy is given by,

$$H(X)_{max} = \log \lambda e \qquad bits/sample$$

- If the signal is bandlimited to B Hz and is sampled at Nyquist rate r_s=2B samples/sec , then

$$H(X)_{max} = 2B \log \lambda e \qquad bits/sample$$

Average power limitation

- Under this limitation we have

$$\int_{-\infty}^{\infty} x^2 f(x) dx = second\ moment = E[X^2] = a\ constant$$

- Random noise with specified variance, audio frequency telephony and other similar situations are examples of some instances where average power limitation being specified.

- The entropy will be maximum when it is Gaussian distribution with mean zero and variance σ^2.

i.e., $$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

- The maximum entropy is given by

$$H(X)_{max} = \frac{1}{2}\log 2\pi e\sigma^2 \qquad bits/sample$$

- If the signal is bandlimited to B Hz and is sampled at Nyquist rate r_s=2B samples/sec , then

$$H(X)_{max} = \mathrm{B}\log 2\pi e\sigma^2 \qquad bits/sample$$

- If X represents Gaussian noise (AWGN) with an average power σ^2 =N, then

$$H(X)_{max} = \mathrm{B}\log 2\pi eN \qquad bits/sample$$

Average power limitation with unidirectional distribution (causal systems)

- Under this limitation we have

$$\int_0^\infty x^2 f(x)dx = P_0 = a\ constant$$

- AM with average carrier power constraint is an example.
- The maximum entropy is given by

$$H(X)_{max} = \frac{1}{2}\log\left[\frac{\pi e P_0}{2}\right] \qquad bits/sample$$

- If the signal is bandlimited to B Hz and is sampled at Nyquist rate r_s=2B samples/sec , then

$$H(X)_{max} = \mathrm{B}\log\left[\frac{\pi e P_0}{2}\right] \qquad bits/sample$$

Numerical Problem

Q. A continuous random variable, X is uniformly distributed in the interval (0, 4). Find the differential entropy H(X). Suppose that X is a voltage which is applied to an amplifier whose gain is 8. Find the differential entropy of the output of the amplifier.

- **Solution**

The PDF of X is given by,

$$f(x) = \frac{1}{b-a}$$

$$=1/4$$

Differential entropy

$$H(X) = \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)} dx$$

H(X)=2 bits/sample

Let Y be the output of amplifier

$$Y = 8\,X \qquad (1)$$

By theorem of probability density function

$$f(y) = f(x)\frac{dx}{dy}$$

Differentiating eqn (1)

$$dy = 8\,dx$$

$$dx/dy = 1/8$$

$$f(y) = 1/32$$

X varies from 0 to 4, Y varies from 0 to 32

Differential entropy of Y is,

$$H(Y) = \int_{-\infty}^{\infty} f(y)\log\frac{1}{f(y)}\,dy$$

$$= 5 \text{ bits/sample}$$

Joint, Conditional Entropy

- Consider a pair of continuous random variable *(X, Y)* distributed according to the joint p.d.f. *f(x, y).* The joint entropy is given by

$$H(X,Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log f(x,y)\, dx\, dy$$

- Conditional entropy is given by

$$H(X/Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log f(x/y)\, dx\, dy$$

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log f(y/x)\, dx\, dy$$

Mutual Information

- Mutual information I(X;Y) is given by

$$I(X;Y) = H(X) - H\left({}^{X}/_{Y}\right)$$

$$I(X;Y) = H(Y) - H\left({}^{Y}/_{X}\right)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

- The mutual information between two continuous random variables X, Y with joint p.d.f *f(x, y)* is given by

$$\boldsymbol{I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) log \frac{f(x,y)}{f(x)f(y)} dxdy}$$

Rate of Transmission and Channel Capacity

The rate of transmission of continuous signal is given by

$$R_t = I(X;Y)r_s \quad bits/sec$$

$$R_t = \left[H(Y) - H\left({}^{Y}/_{X}\right)\right]r_s$$

$$R_t = [H(Y) - H(N)]r_s$$

Channel capacity is given by

$$C = Max(R_t)$$

$$C = Max\left[[H(Y) - H(N)]r_s\right]$$

Consider a continuous random variable Y defined by,

$$Y=X+N$$

Where X & N are statistically independent. Show that the conditional entropy of Y, given X is, $H(Y/X)=H(N)$.

Where $H(N)$ is the differential entropy of N.

Proof

The equivocation for continuous signal is given by,

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log f(y/x)\,dx\,dy \qquad (1)$$

Since $f(x,y) = f(y/x)f(x)$

We have,

$$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y/x)f(x)\log f(y/x)\,dx\,dy \qquad (2)$$

$$H(Y/X) = -\int_{-\infty}^{\infty} f(x)dx \int_{-\infty}^{\infty} f(y/x)\log f(y/x)dy \qquad (3)$$

Given that $Y = X + N$

When N represents the channel noise and X the transmitted signal, then $f(y/x)$ depends on $(y - x)$ and not on x or y. Let $f_N(n)$ denote the PDF of N. Then,

$$f(y/x) = f_N(y - x)$$

$$-\int_{-\infty}^{\infty} f(y/x)\log f(y/x)dy = -\int_{-\infty}^{\infty} f_N(y-x) log f_N(y-x)dy$$

Since $y = x+n$

$dy=dn$ with x constant

Also $y\text{-}x=n$

Substitute these in above equation

$$= - \int_{-\infty}^{\infty} f_N(n) log f_N(n) dn$$

$$= H(N) \qquad (4)$$

Substitute eqn(4) in (3) we get,

$$H(Y/X) = \int_{-\infty}^{\infty} f(x)dx\, H(N)$$

Since $\int_{-\infty}^{\infty} f(x)dx = 1$

$$H(Y/X) = H(N)$$

Properties of Mutual Information

$$I(X;Y) = \int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(x,y) \log \left[\frac{f(x/y)}{f(x)}\right] dxdy \qquad (1)$$

$$I(Y;X) = \int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(x,y) \log \left[\frac{f(y/x)}{f(y)}\right] dxdy \qquad (2)$$

Property 1 $I(X;Y) = I(Y;X)$

From Baye's Rule, We have

$$\frac{f(x/y)}{f(x)} = \frac{f(y/x)}{f(y)}$$

Equation (1) and (2) are identical

$$I(X;Y) = I(Y;X)$$

Property 2 $I\,(X;Y) \geq 0$

Proof

We know, $f(x,y) = f(x/y)f(y)$

$$f(x/y) = \frac{f(x,y)}{f(y)} \qquad (3)$$

Using eqn (3) in (1), we get

$$I\,(X;Y) = \int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\log\left[\frac{f(x,y)}{f(x)f(y)}\right]dxdy$$

$$I\,(X;Y) = \frac{1}{\ln 2}\int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\ln\left[\frac{f(x,y)}{f(x)f(y)}\right]dxdy \qquad (4)$$

We know the inequality,

$$\ln\frac{1}{x} \geq 1 - x$$

Put

$$x = \frac{f(x)f(y)}{f(x,y)}$$

$$\ln\left[\frac{f(x,y)}{f(x)f(y)}\right] \geq 1 - \frac{f(x)f(y)}{f(x,y)}$$

Multiplying both sides by $f(x,y)$ and then double integrating, we get

$$\int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\ln\left[\frac{f(x,y)}{f(x)f(y)}\right]dxdy \;\geq\; \int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\left[1 - \frac{f(x)f(y)}{f(x,y)}\right]dxdy$$

Dividing both sides by ln 2, we get

$$\frac{1}{ln2}\int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\ln\left[\frac{f(x,y)}{f(x)f(y)}\right]dxdy \;\geq \frac{1}{ln2}\int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto}[f(x,y)-f(x)f(y)]\; dxdy$$

But LHS of above equation is $I(X;Y)$) from equation (4)

$$I(X;Y) \geq \frac{1}{ln2}\left[\int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)dxdy - \int_{-\propto}^{+\propto} f(x)\,dx \int_{-\propto}^{+\propto} f(y)\,dy\right]$$

Since $f(x,y), f(x) and\ f(y) are\ all\ pdfs, their\ integral\ values\ are 1$

$$I(X;Y) \geq \frac{1}{ln2}[1-1]$$

$$I(X;Y) \geq \; 0$$

Property 3 $\boldsymbol{I(X;Y) = H(X) - H(X/Y)}$

Proof

$$H(X/Y) = \int_{-\propto}^{+\propto}\int_{-\propto}^{+\propto} f(x,y)\log\frac{1}{f(x/y)}\,dxdy \qquad (1)$$

Also

$$H(X) = \int_{-\propto}^{+\propto} f(x) log\frac{1}{f(x)}\,dx \qquad (2)$$

We have

$$\int_{-\propto}^{+\propto} f(y/x)\,dy = 1$$

Similar to $\sum_{j=1}^{s} P(bj/ai)$ ***= 1***

$$H(X) = \int_{-\propto}^{+\propto} f(x) log \frac{1}{f(x)} dx \int_{-\propto}^{+\propto} f(y/x) dy$$

$$\int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(y/x) f(x) \log \frac{1}{f(x)} dx\, dy$$

$$f(y/x) f(x) = f(x,y)$$

$$H(X) = \int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(x,y) \log \frac{1}{f(x)} dx\, dy \tag{3}$$

Subracting eqn (1) from (3)

$$H(X) - H(X/Y) = \int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(x,y) \left[\log \frac{1}{f(x)} - log \frac{1}{f(x/y)} \right] dxdy$$

$$= \int_{-\propto}^{+\propto} \int_{-\propto}^{+\propto} f(x,y) log\left[\frac{f(x/y)}{f(x)}\right] dxdy$$

$$= I(X;Y)$$

$$I(X;Y) = H(X) - H(X/Y)$$

Similarly, we can prove

$$I(X;Y) = H(Y) - H(Y/X)$$

Property 4 $\boldsymbol{I(X;Y) = H(X) + H(Y) - H(X,Y)}$

Proof

$$H(X,Y) = H(X/Y) + H(Y) \qquad (1)$$

$$I(X;Y) = H(X) - H\left(\frac{X}{Y}\right) \qquad (2)$$

From (1)

$$H(X/Y) = H(X,Y) - H(Y) \qquad (3)$$

Substitute (3) in (2)

$$I(X;Y) = H(X) - [H(X,Y) - H(Y)]$$

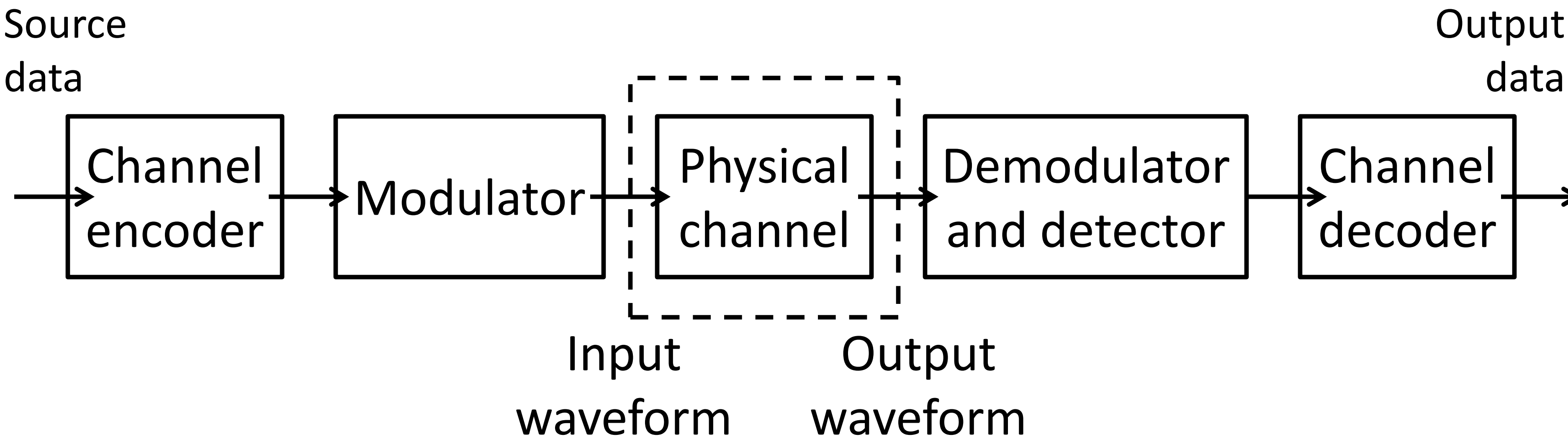
$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

Gaussian Channel

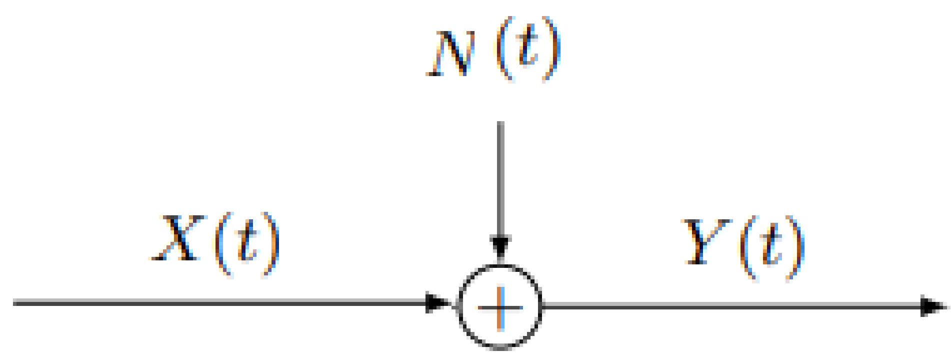
- So far we have studied limits on the maximum rate at which information can be sent over a channel reliably in terms of the channel capacity.
- We next formulate the information capacity theorem for band limited, power limited Gaussian channels.

Gaussian Channel

- An important and useful channel is the Gaussian channel.

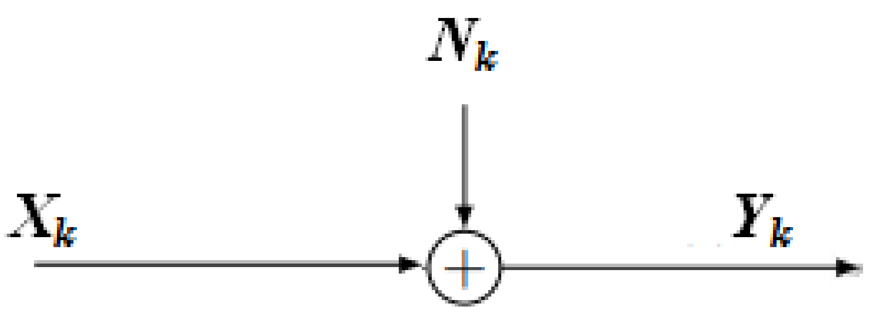


$$y(t) = x(t) + n(t)$$



Gaussian Channel

- This is a time discrete channel with output Y_k at time k.
- This output is the result of the sum of the input X_k and the noise N_k.
- This noise is drawn from a Gaussian distribution with mean zero and variance σ^2.
- Thus, $Y_k=X_k+N_k$
- The noise N_k is independent of the input X_k.



Capacity of a Gaussian Channel

- Since the transmitter is usually power-limited (Let S is the transmitted signal power in Watts), let us put a constraint on the average power in X_k .

$$E[X_k^{\;2}] = S$$
$$where k = 1,2,\ldots K$$

- Thus the information capacity of the channel is given by

$$C = \underset{f_{x_k}(x)}{Max}\{I(X;Y)/E[X_k^2] = S\}$$

Capacity of a Gaussian Channel

$$I(X;Y) = H(Y) - H(Y/X)$$

$$I(X_k;Y_k) = H(Y_k) - H(Y_k/X_k)$$

- X_k and N_k are independent random variables
- Therefore

$$H(Y_k/X_k) = H(N_k)$$

- Hence we can write,

$$I(X_k;Y_k) = H(Y_k) - H(N_k)$$

Capacity of a Gaussian Channel

- If we assume Y_k to be Gaussian, and N_k is Gaussian by definition, then X_k is also Gaussian.
- In order to maximize the mutual information between the channel input X_k and channel output Y_k the transmitted signal should be Gaussian.
- Therefore we can write

$$C = I(X_k;Y_k) \big/ E[X_k^2] = S$$

and also X_k is Gaussian

Capacity of a Gaussian Channel

- We know that if two independent Gaussian random variables are added, the variance of the resulting Gaussian random variable should be the sum of the variances.
- Let N_0B is the variance of noise random variable.
- Where $N_0/2$ (Also represented as **η/2**)is the two sided power spectral density of noise signal.
- Therefore, the variance of the received sample Y_k equals

$$\sigma_{Y_k}{}^2 = S + N_0B$$

Capacity of a Gaussian Channel

- It can be shown that the maximum differential entropy of a Gaussian random variable with variance σ^2 is

$$H(Y_k)_{\max} = \frac{1}{2}\log_2(2\pi e\sigma^2)$$

$$H(Y_k)_{\max} = \frac{1}{2}\log_2[2\pi e(S + N_0B)]$$

$$H(N_k)_{\max} = \frac{1}{2}\log_2[2\pi e(N_0B)]$$

- Channel capacity

$$C = \left[H(Y_k) - H(N_k)\right]_{\max}$$

$$C = \left[H(Y_k)_{\max} - H(N_k)_{\max}\right]$$

Capacity of a Gaussian Channel

$$C = H(Y_k)_{\max} - H(N_k)_{\max}$$

$$C = \frac{1}{2}\log_2[2\pi e(S + N_0 B)] - \frac{1}{2}\log_2[2\pi e(N_0 B)]$$

$$C = \frac{1}{2}\log_2\left[\frac{2\pi e(S + N_0 B)}{2\pi e(N_0 B)}\right]$$

$$C = \frac{1}{2}\log_2\left[1 + \frac{S}{N_0 B}\right] bits/channeluse$$

Capacity of a Gaussian Channel

$$C = \frac{1}{2}\log_2\left[1+\frac{S}{N_0 B}\right] bits/channeluse$$

- We are transmitting 2B samples per second, i.e.,the channel is being 2B times in one second.
- Therefore, the information capacity can be expressed as

$$C = 2B\frac{1}{2}\log_2\left[1+\frac{S}{N_0 B}\right]$$

$$C = B\log_2\left[1+\frac{S}{N_0 B}\right] bits/\sec$$

Capacity of a Gaussian Channel

$$C = B\log_2\left[1+\frac{S}{N_0 B}\right] bits/\sec$$

- Let $N=N_0B$ is the noise power, then channel capacity can be written as

$$C = B\log_2\left[1+\frac{S}{N}\right] bits/\sec$$

- This basic formula for the capacity of the band-limited , AWGN waveform channel with band-limited and average power – limited input was first derived by Shannon in 1948.
- It is known as Shannon’s third theorem, or the Information Capacity Theorem or Shannon- Hartley theorem.

Capacity of a Gaussian Channel

$$C = B\log_2\left[1+\frac{S}{N}\right] bits/\sec$$

$$C = B\log_2\left[1+SNR\right] bits/\sec$$

- SNR is the signal to noise ratio.
- Capacity varies linearly with Bandwidth, B and logarithmically with SNR.

Information Capacity Theorem

- The Information Capacity Theorem is one of the important result in Information Theory.
- In a single formula one can see the trade off between the channel bandwidth, the average transmitted power and the noise power spectral density.
- Given the channel BW and the SNR the channel capacity can be computed.
- The channel capacity is the fundamental limit on the rate of reliable communication for a power limited and bandlimited Gaussian channel.

Problems

- Given an AWGN channel with 5 K Hz bandwidth and the noise power spectral density $\eta/2=10^{-9}$ W/Hz. The signal power required at the receiver is l mW. Calculate the capacity of this channel.

Solution:

Given $\eta/2=N_0/2=10^{-9}$

S= l mW

B=5 KHz

C=?

- A telephone channel has a BW of 3000Hz and the SNR=20dB. Determine the channel capacity. If the SNR is increased to 25dB, determine the capacity.

Solution:

$$10log_{10}\frac{S}{N} = 20$$

S/N=?

B=3000 Hz

C=?

Information Capacity Theorem or Shannon – Hartley theorem

It states that the capacity of a band limited Gaussian channel with AWGN is given by,

$$C = B\log_2\left(1+\frac{S}{N}\right)\, bits/\sec$$

where

B=channel bandwidth in Hz

S=Signal power in Watts

N=Noise power in watts=N_0B or ηB

where the two sided power spectral density of noise is $(N_0/2)$ watts/Hz or $(\eta/2)$

Implications of Shannon – Hartley theorem

1st Implication or Capacity of a channel with infinite bandwidth

From Shannon Hartley Law we have,

$$C = B\log_2(1 + \frac{S}{N})\ bits/\sec \qquad (1)$$

- When B is increased, channel capacity C also increases and , the maximum rate of information transmission can be enhanced to any value as we need.
- However, the channel capacity does not become infinite.
- This is because, B increases, the noise power N which is dependent on B, also increases thereby reducing (S/N) .

- The product of B and $\log_2(1+\frac{S}{N})$ will increase only up to a certain value and becomes constant with increasing B.
- This value is denoted as C_∞

Substitute N = ηB in eqn (1)

$$C = B\log_2\left(1+\frac{S}{\eta B}\right)$$

$$= \frac{S}{\eta}\left(\frac{\eta B}{S}\right)\log_2\left(1+\frac{S}{\eta B}\right)$$

$$= \frac{S}{\eta}\log_2\left(1+\frac{S}{\eta B}\right)^{\left(\frac{\eta B}{S}\right)}$$

Let x= $\frac{S}{\eta B}$

$$C = \frac{S}{\eta}\log_2(1+x)^{\left(\frac{1}{x}\right)}$$

Accordingly, when B→∞ , x →0

$$\lim_{\substack{B\to\infty \\ x\to 0}} C = \lim_{\substack{B\to\infty \\ x\to 0}} \frac{S}{\eta}\log_2(1+x)^{\left(\frac{1}{x}\right)}$$

$$C_\infty = \frac{S}{\eta}\log_2 \lim_{x\to 0}(1+x)^{\left(\frac{1}{x}\right)}$$

$$C_\infty = \frac{S}{\eta}\log_2 e$$

$$C_\infty = \frac{S}{\eta}\ 1.44 \qquad\qquad bits/sec$$

Shannon’s Limit

- An ideal system is defined as one that transmits data at a bit rate R_t equal to the channel capacity C.
- Then the average transmitted power can be expressed as,

$$S = E_b C$$

- Where E_b = transmitted energy per bit in joules

Using $N = \eta B$ and $S = E_b C$ in eqn (1) we get for an ideal system

$$C = B \log_2 \left(1 + \frac{E_b C}{\eta B}\right)$$

or

$$\frac{C}{B} = \log_2 \left(1 + \frac{E_b C}{\eta B}\right)$$

The quantity C/B is called “ Bandwidth efficiency”

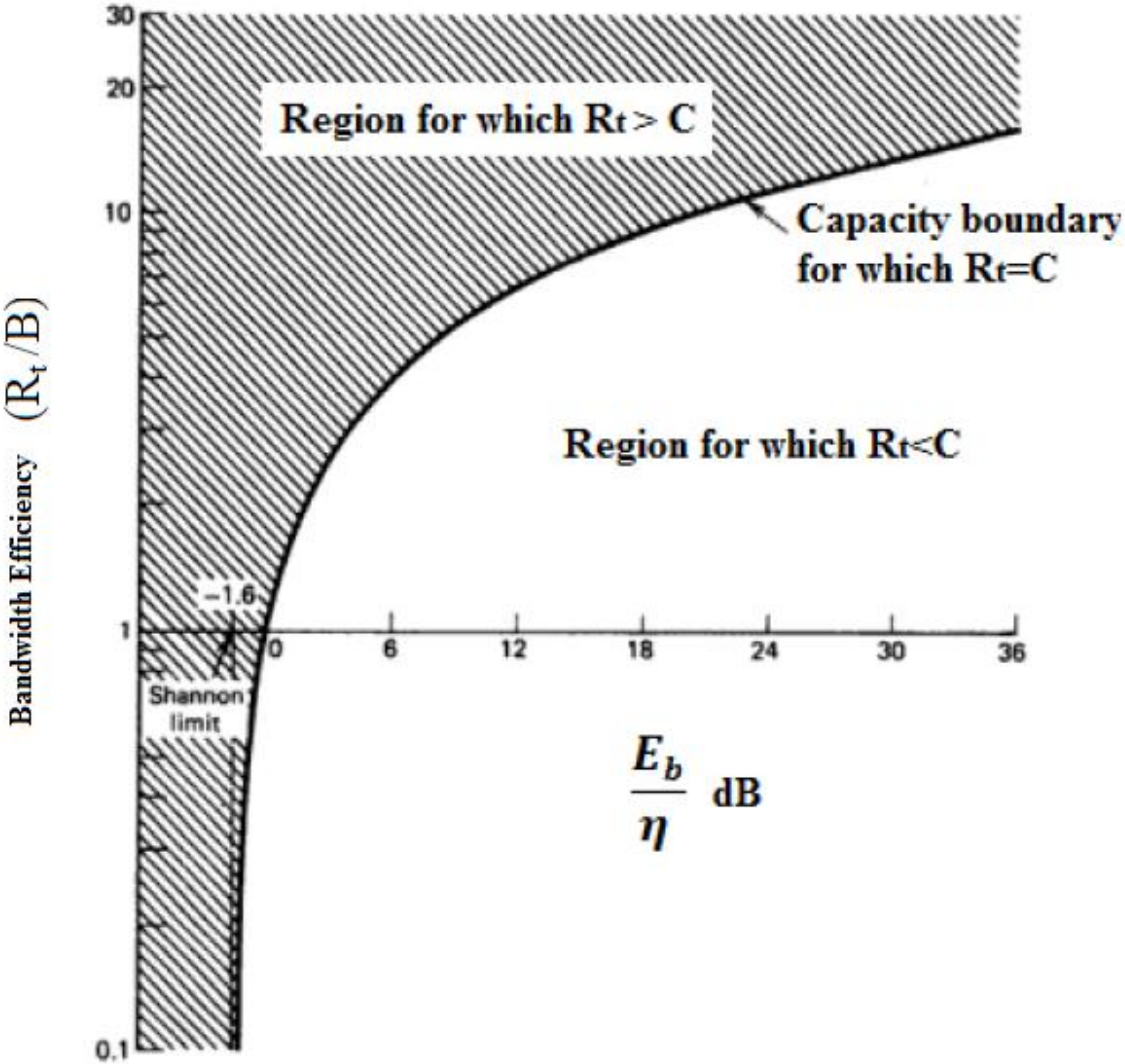
- The quantity E_b/η is given by

$$\frac{E_b}{\eta} = \frac{2^{C/B} - 1}{C/B}$$

- When (R_t/B) is plotted as a function of (E_b/η), we get the bandwidth efficiency diagram.

- The resulting curve represents the capacity boundary for which $R_t = C$.

Bandwidth-Efficiency Diagram



Based on the diagram following observations are made:

1. **For infinite bandwidth, the signal -to - noise ratio E_b/η approaches the limiting value.**

$$\left(\frac{E_b}{\eta}\right)_{\infty} = \lim_{B\to\infty}\left(\frac{E_b}{\eta}\right) = \lim_{B\to\infty}\left[\frac{2^{C/B}-1}{(C/B)}\right]$$

$$\frac{C}{B} = x \quad \text{As B} \to \infty\text{ , } x \to 0$$

$$\left(\frac{E_b}{\eta}\right)_{\infty} = \lim_{x\to 0}\left(\frac{2^x-1}{x}\right) \qquad (2)$$

Using L' Hospital Rule, the above limit can be evaluated as below:

$$\text{Let } y = 2^x$$

- Taking *ln* on both sides,

$$\ln y = x \ln 2$$

- Differentiating,

$$\frac{1}{y}\, dy = \ln 2\, dx$$

$$\frac{dy}{dx} = y\,(\ln 2) \quad = 2^x \ln 2$$

- Differentiating both numerator and denominator of the RHS of eqn (2) with respect to ‘x’, we get

$$\left(\frac{E_b}{\eta}\right)_{\infty} = \lim_{x \to 0}\left[\frac{\frac{d}{dx}(2^x - 1)}{\frac{d}{dx}(x)}\right]$$

$$\left(\frac{E_b}{\eta}\right)_\infty = \lim_{x\to 0}\left[\frac{\frac{d}{dx}(2^x - 1)}{\frac{d}{dx}(x)}\right]$$

$$= \lim_{x\to 0}\left[\frac{2^x \ln 2}{1}\right] \quad = 2^0 \ln 2$$

$$\left(\frac{E_b}{\eta}\right)_\infty = ln2 = 0.693$$

$$\left(\frac{E_b}{\eta}\right)_\infty in\ dB = 10\ \log_{10} 0.693$$

$$\left(\frac{E_b}{\eta}\right)_\infty in\ dB \cong -1.6\ dB$$

This value 0.693 or -1.6 dB is called **“Shannon’s Limit”..**

$$\left(\frac{E_b}{\eta}\right)_{\infty} = ln2 = 0.693$$

- The Shannon's limit is a **fraction**.
- This implies that for very large bandwidths, reliable communication is possible even for the case when the signal power is **less than the noise power.**
- The channel capacity corresponding to this **limiting value** is given by,

$$C_{\infty} = \frac{S}{\eta}\log_2 e$$

2. The capacity boundary, defined by the curve for critical bit rate $R_t=C$.

   It separates combinations of system parameters that have the potential for supporting error free transmission ($R_t<C$) from those for which error-free transmission is not possible ($R_t>C$).

3. The Bandwidth-Efficiency diagram highlights trade-off between (E_b/η) and (R_t/B).

- This is given by 2nd implication of Shannon-Hartley law.

Information Capacity Theorem or Shannon – Hartley theorem

It states that the capacity of a band limited Gaussian channel with AWGN is given by,

$$C = B\log_2\left(1 + \frac{S}{N}\right)\ bits/\sec$$

where

B=channel bandwidth in Hz

S=Signal power in Watts

N=Noise power in watts=N_0B or ηB

where the two sided power spectral density of noise is $(N_0/2)$ watts/Hz or $(\eta/2)$

Implications of Shannon – Hartley theorem

1st Implication or Capacity of a channel with infinite bandwidth

From Shannon Hartley Law we have,

$$C = B\log_2\left(1 + \frac{S}{N}\right)\ bits/\sec \qquad (1)$$

- When B is increased, channel capacity C also increases and , the maximum rate of information transmission can be enhanced to any value as we need.
- However, the channel capacity does not become infinite.
- This is because, B increases, the noise power N which is dependent on B, also increases thereby reducing (S/N) .

- The product of B and $\log_2(1+\frac{S}{N})$ will increase only up to a certain value and becomes constant with increasing B.
- This value is denoted as C_∞

Substitute $N = \eta B$ in eqn (1)

$$C = B\log_2\left(1+\frac{S}{\eta B}\right)$$

$$= \frac{S}{\eta}\left(\frac{\eta B}{S}\right)\log_2\left(1+\frac{S}{\eta B}\right)$$

$$= \frac{S}{\eta}\log_2\left(1+\frac{S}{\eta B}\right)^{\left(\frac{\eta B}{S}\right)}$$

Let $x = \frac{S}{\eta B}$

$$C = \frac{S}{\eta} \log_2 (1 + x)^{\left(\frac{1}{x}\right)}$$

Accordingly, when B→∞ , x →0

$$\lim_{\substack{B \to \infty \\ x \to 0}} C = \lim_{\substack{B \to \infty \\ x \to 0}} \frac{S}{\eta} \log_2 (1 + x)^{\left(\frac{1}{x}\right)}$$

$$C_\infty = \frac{S}{\eta} \log_2 \lim_{x \to 0} (1 + x)^{\left(\frac{1}{x}\right)}$$

$$C_\infty = \frac{S}{\eta} \log_2 e$$

$$C_\infty = \frac{S}{\eta}\ 1.44 \qquad bits/sec$$

Shannon's Limit

- An ideal system is defined as one that transmits data at a bit rate R_t equal to the channel capacity C.
- Then the average transmitted power can be expressed as,

$$S = E_b C$$

- Where E_b = transmitted energy per bit in joules

Using $N=\eta B$ and $S= E_b C$ in eqn (1) we get for an ideal system

$$C = B\log_2\left(1 + \frac{E_b C}{\eta\, B}\right)$$

or

$$\frac{C}{B} = \log_2\left(1 + \frac{E_b C}{\eta\, B}\right)$$

The quantity C/B is called " Bandwidth efficiency"

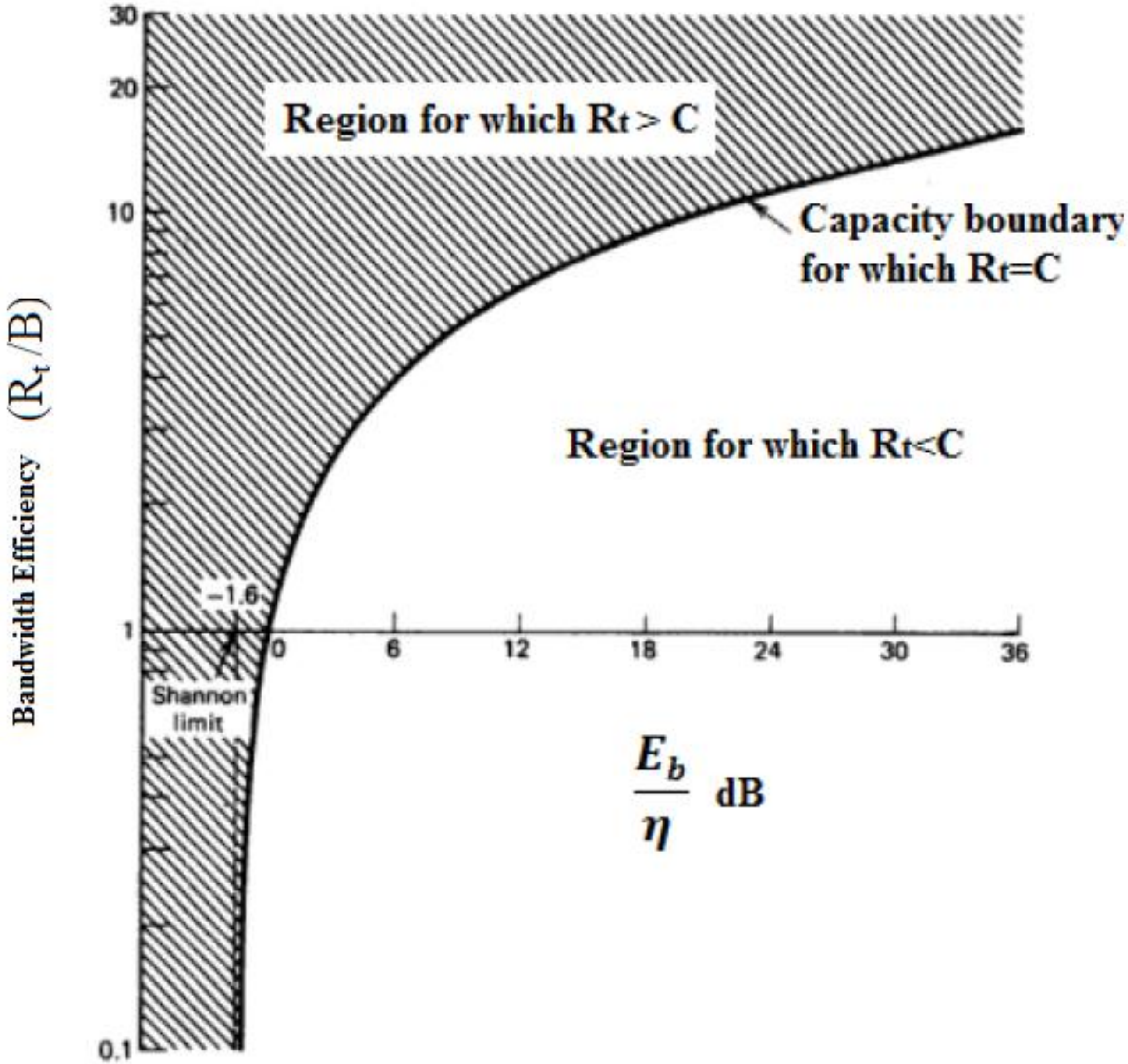
- The quantity E_b/η is given by

$$\frac{E_b}{\eta} = \frac{2^{C/B} - 1}{C/B}$$

- When (R_t/B) is plotted as a function of (E_b/η), we get the bandwidth efficiency diagram.

- The resulting curve represents the capacity boundary for which $R_t = C$.

Bandwidth-Efficiency Diagram



Based on the diagram following observations are made:

1. **For infinite bandwidth, the signal -to - noise ratio E_b/η approaches the limiting value.**

$$\left(\frac{E_b}{\eta}\right)_{\infty} = \lim_{B\to\infty}\left(\frac{E_b}{\eta}\right) = \lim_{B\to\infty}\left[\frac{2^{C/B}-1}{(C/B)}\right]$$

$$\frac{C}{B} = x \quad \text{As } B \to \infty,\ x \to 0$$

$$\left(\frac{E_b}{\eta}\right)_{\infty} = \lim_{x\to 0}\left(\frac{2^x-1}{x}\right) \qquad (2)$$

Using L' Hospital Rule, the above limit can be evaluated as below:

$$\text{Let } y = 2^x$$

- Taking *ln* on both sides,

$$\ln y = x \ln 2$$

- Differentiating,

$$\frac{1}{y}\, dy = \ln 2\, dx$$

$$\frac{dy}{dx} = y\,(\ln 2) \quad = 2^x \ln 2$$

- Differentiating both numerator and denominator of the RHS of eqn (2) with respect to ‘x’, we get

$$\left(\frac{E_b}{\eta}\right)_\infty = \lim_{x \to 0}\left[\frac{\frac{d}{dx}(2^x - 1)}{\frac{d}{dx}(x)}\right]$$

$$\left(\frac{E_b}{\eta}\right)_\infty = \lim_{x\to 0}\left[\frac{\frac{d}{dx}(2^x - 1)}{\frac{d}{dx}(x)}\right]$$

$$= \lim_{x\to 0}\left[\frac{2^x \ln 2}{1}\right] \quad = 2^0 \ln 2$$

$$\left(\frac{E_b}{\eta}\right)_\infty = ln2 = 0.693$$

$$\left(\frac{E_b}{\eta}\right)_\infty \; in\; dB = 10\ \log_{10} 0.693$$

$$\left(\frac{E_b}{\eta}\right)_\infty \; in\; dB \cong -1.6\; dB$$

This value 0.693 or -1.6 dB is called **“Shannon’s Limit”.**.

$$\left(\frac{E_b}{\eta}\right)_{\infty} = ln2 = 0.693$$

- The Shannon's limit is a **fraction**.
- This implies that for very large bandwidths, reliable communication is possible even for the case when the signal power is **less than the noise power.**
- The channel capacity corresponding to this **limiting value** is given by,

$$C_{\infty} = \frac{S}{\eta}\log_2 e$$

2. The capacity boundary, defined by the curve for critical bit rate $R_t=C$.

   It separates combinations of system parameters that have the potential for supporting error free transmission ($R_t<C$) from those for which error-free transmission is not possible ($R_t>C$).

3. The Bandwidth-Efficiency diagram highlights trade-off between (E_b/η) and (R_t/B).

- This is given by 2nd implication of Shannon-Hartley law.

Implications of Shannon – Hartley theorem

2nd Implication - Bandwidth-SNR Trade Off

- An important implication of Shannon-Hartley law is the exchange of bandwidth with, signal to noise ratio and vice versa.

Suppose $\frac{S_1}{N_1} = 7$ and Bandwidth B1=4 KHz

- Therefore Channel Capacity,

$$C_1 = B_1 \log\left(1 + \frac{S_1}{N_1}\right)$$

$$= 4 \times 10^3 \log(1 + 7)$$

$$= 12 \times 10^3 \; bits/sec$$

- Keeping the channel capacity C_2 same as C_1 and if signal-to-noise ratio is increased to 15, then

$$C_2 = C_1 = 12 \times 10^3 = B_2 \log\left(1 + \frac{S_2}{N_2}\right)$$

$$12 \times 10^3 = B_2 \log(1 + 15)$$

- We get B_2=3 KHz
- Since the noise power $N=\eta B$, as the bandwidth gets reduced from 4 to 3 KHz, the noise also decreases indicating an increase in signal power as shown below.
- We have $N_1 = \eta B_1 = (\eta)$ (4KHz)
- And $N_2 = \eta B_2 = (\eta)$ (3KHz)

$$\text{Consider} \quad \frac{S_2/N_2}{S_1/N_1} = \frac{15}{7}$$

Therefore $\frac{S_2}{S_1} = \frac{15N_2}{7N_1} = \frac{(15)(\eta)(3\ KHz)}{(7)(\eta)(4\ KHz)} = \frac{45}{28} = 1.6$

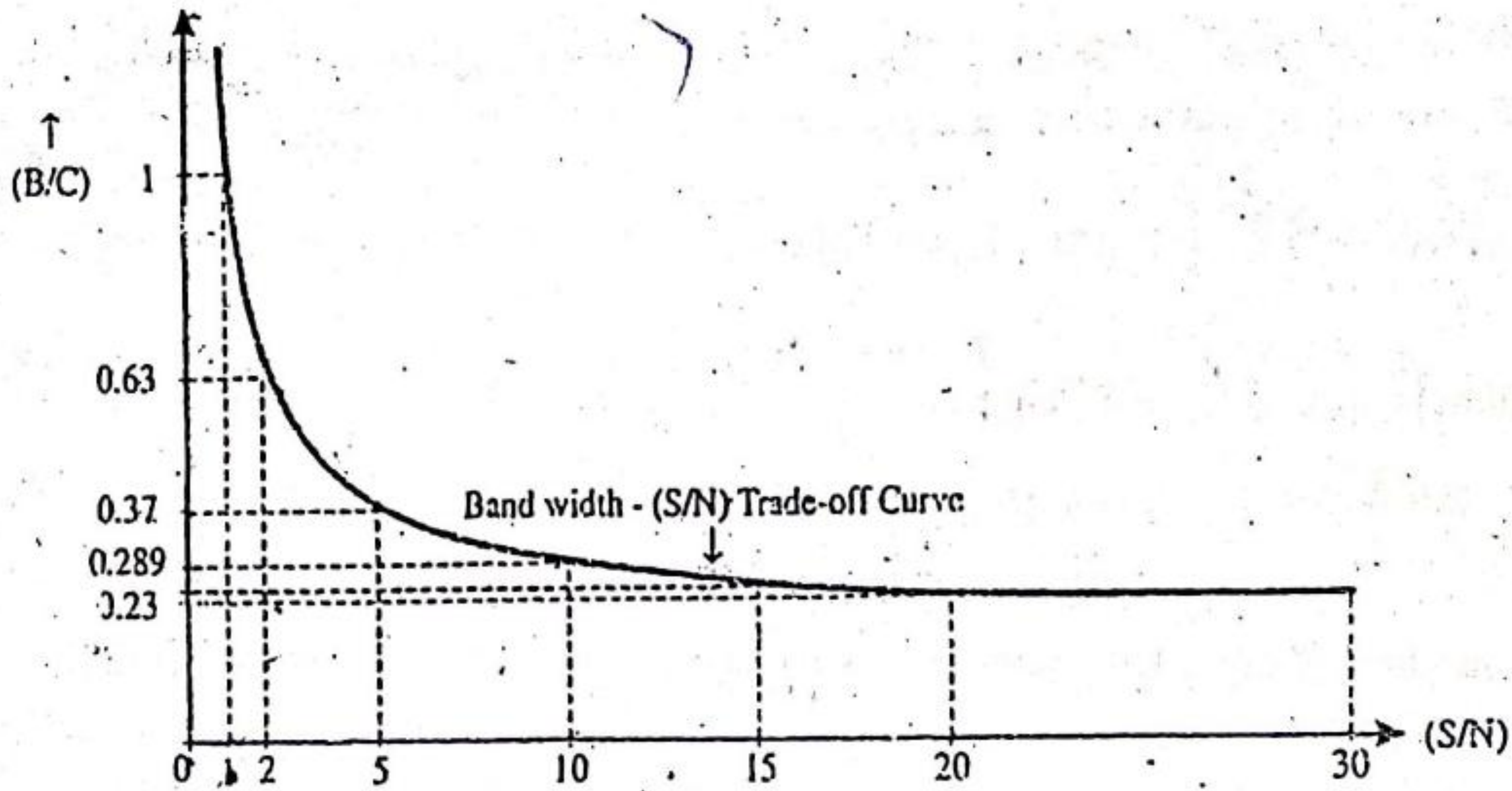
- Thus a 25% reduction in Bandwidth from 4 KHz to 3 KHz requires a 60 % approximate increase in signal power for maintaining the same channel capacity.
- Let us look into the exact significance by drawing the "trade - off curve".
- From Shannon-Hartley law

$$\frac{B}{C} = \frac{1}{\log_2\left(1 + \frac{S}{N}\right)}$$

The values of (B/C) for different values of (S/N) are listed below.

S/N	0.5	1	2	5	10	15	20	30
B/C	1.71	1	0.63	0.37	0.289	0.25	0.23	0.2

Bandwidth to (S/N) Trade-Off Curve



- It shows a plot of (B/C) as a function of (S/N).
- Using this trade off curve the same channel capacity can be obtained by increasing bandwidth if (S/N) is small.
- Furthermore, the curve also indicates that there exists a threshold point at around (S/N)=10 up to which the exchange rate of bandwidth with (S/N) is advantageous.
- Beyond (S/N)=10, the reduction in B with increasing (S/N) is very poor.
- FM, PM and PCM systems including DM and ADM systems require larger bandwidths with reasonably good (S/N) ratio.