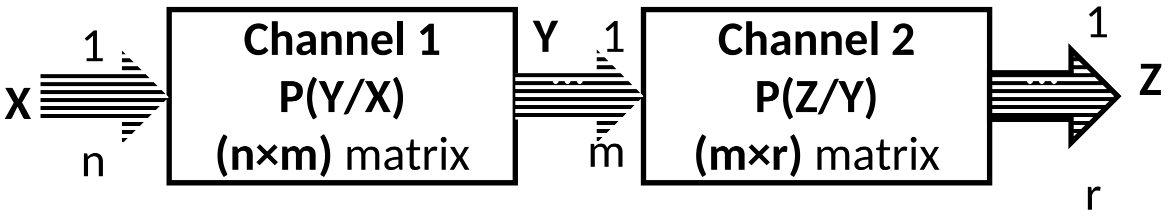
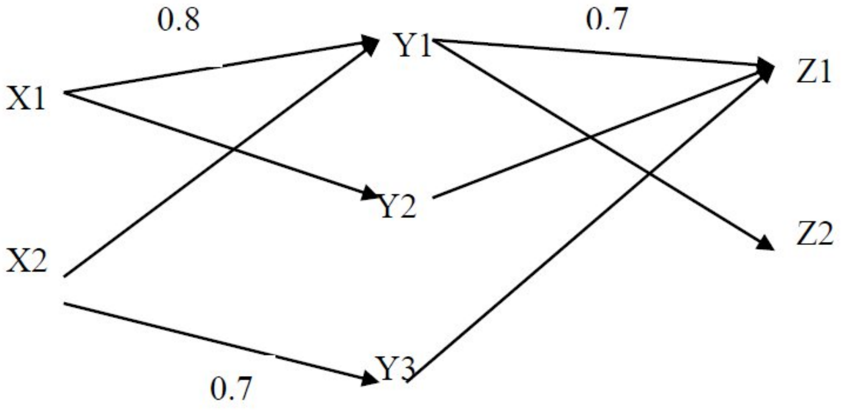
If two channels are cascaded as shown, then, the overall transition matrix of the equivalent channel is the matrix multiplication of the transitional prob of the two cascaded channels.



P(Z/X) = P(Y/X) × P(Z/Y)

Example : Find the transition matrix p(Z/X) for the cascaded channel shown:



Solution:

$$p(Y/X) = \begin{array}{c} \\ x1 \\ x2 \end{array} \begin{bmatrix} y1 & y2 & y3 \\ 0.8 & 0.2 & 0 \\ 0.3 & 0 & 0.7 \end{bmatrix} \quad and \quad p(Z/Y) = \begin{array}{c} \\ y1 \\ y2 \\ y3 \end{array} \begin{bmatrix} z1 & z2 \\ 0.7 & 0.3 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$p(Z/X) = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.3 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.76 & 0.24 \\ 0.91 & 0.09 \end{bmatrix}$$

HW: For previous example, find P(Y) and P(Z) if P(X)=[0.7 0.3]

Entropies of continuous signals:

If x and y are continuous random variables, with probability density functions p(x) and p(y), then in analogy with discrete sources the differential entropies of X & Y are given by:

$H(X) = -\int_{-\infty}^{\infty} p(x)\ log2\ p(x)dx$ in bits/sample of the random variable x

$H(Y) = -\int_{-\infty}^{\infty} p(y)\, log2\ p(y)dy$ in bits/sample of the random variable x

And other entropies are also differential entropies and are given by:

$H(X,Y) = -\iint_{-\infty}^{\infty} p(x,y)\ log2\ p(x,y)\, dxdy$ in bits/sample

$H(Y/X) = -\iint_{-\infty}^{\infty} p(x,y)\ log2\ p(y/x)\, dxdy$ in bits/sample

$H(X/Y) = -\iint_{-\infty}^{\infty} p(x,y)\ log2\ p(x/y)\, dxdy$ in bits/sample

$I(X,Y) = -\iint_{-\infty}^{\infty} p(x,y)\ log2\ \frac{p(x/y)}{p(x)} dxdy$ in bits/sample

Note that all above entropies are differential entropies and not an absolute measure of information since all prob are in fact prob. density functions.

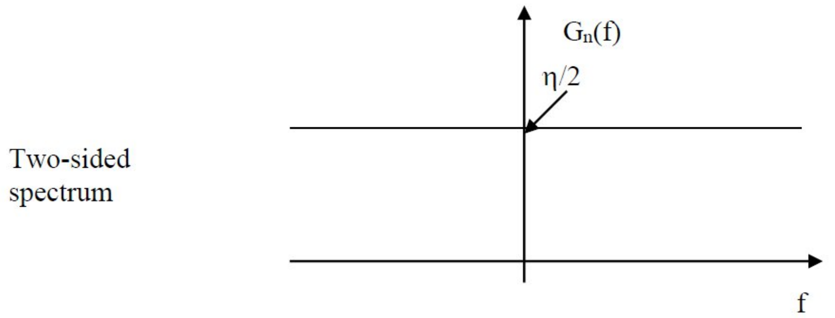
Definition: A Gaussian channel is that channel affected by the Gaussian noise.

Review of Gaussian signal:

If the noise signal n(t) is Gaussian then its PDF(prob density function):

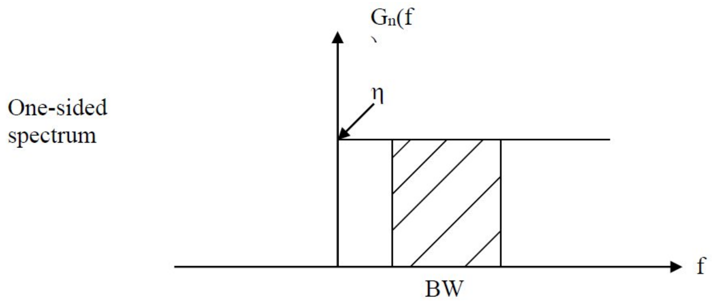
$$p(n) = \frac{1}{\sqrt{2\pi\sigma}}\ e^{-0.5\left(\frac{n-\mu}{\sigma}\right)^{2}}$$

where **μ** is mean of n(t) and $\boldsymbol{\sigma^2}$ is the variance of n(t). If n(t) is a thermal noise we can assume **μ=0** and the frequency spectrum of this noise is flat over a wide range of frequencies as shown



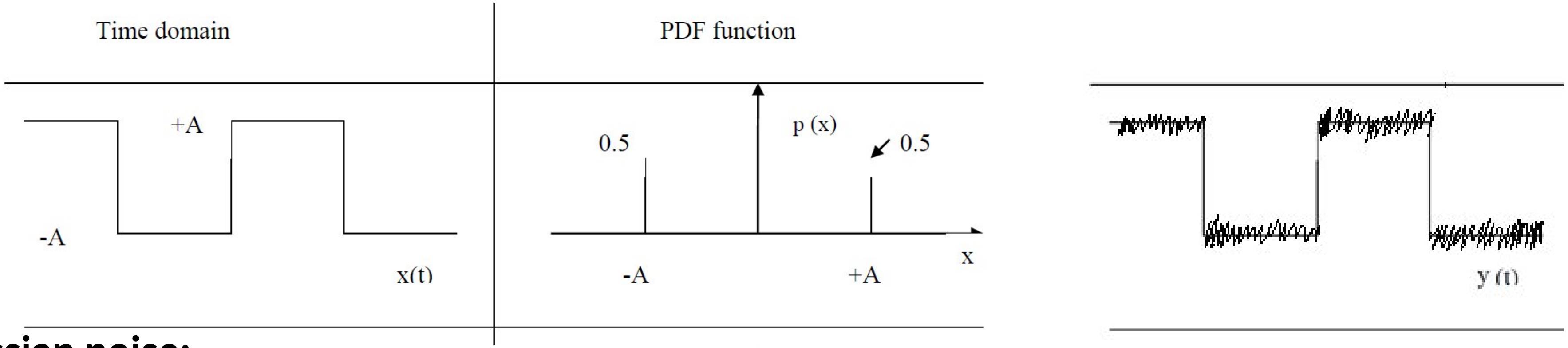
This has two sided power spectral density $G_n(f)=\eta/2$ W/Hz and one-sided power spectral density $G_n(f)= \eta$ W/Hz

From $G_n(f)$, we can find the noise power as N= η BW in watts.



Entropy of Gaussian noise:

Since the spectrum is flat, we call this noise white noise. This white noise affects the signal x(t) as additive term, i.e., the received signal ***y(t)=x(t)+n(t).*** A very popular name of Additive, White, Gaussian, Noise (**AWGN**) is used for such thermal noise. The figure below shows how this AWGN affects equiprobable bipolar ± A signal.



Entropy of Gaussian noise:

Mathematically, we can prove that if x(t) is a random variable, then the entropy of x is maximum if x(t) has Gaussian PDF. To find this entropy, then (and assuming **μ**=0)

$$H(X) = -\int_{-\infty}^{\infty} p(x)\ln\left[\frac{1}{\sqrt{2\pi}\,\sigma}e^{-0.5\frac{x^2}{\sigma^2}}\right]dx \quad \text{nats/sample}$$

$$H(X) = \int_{-\infty}^{\infty} p(x)\ln\sqrt{2\pi}\,\sigma \; dx + \frac{1}{2\sigma^2}\int_{-\infty}^{\infty} x^2 p(x)dx$$

But: $\int_{-\infty}^{\infty} x^2 p(x)dx = mean\ square\ of\ x = \mu^2 + \sigma^2 = \sigma^2$ *and* $\int_{-\infty}^{\infty} p(x)\,dx = 1$

then $H(X) = \ln\sqrt{2\pi}\,\sigma + 0.5 = \ln\sqrt{2\pi}\,\sigma + \ln\sqrt{e}$

$$\boxed{\begin{array}{l} H(X) = \ln(\sqrt{2\pi e}\,\sigma) \ \text{nats/sample} \\ \text{0r} \quad H(X) = \log_2(\sqrt{2\pi e}\,\sigma) \ \text{bits/sample} \end{array}}$$

A Gaussian channel is a channel affected by Gaussian noise n(t). Then:
C=max[H(Y) - H(Y/X)], =max[receiver entropy – noise entropy]. It should be noted that, maximization is already included when we take the case of Gaussian noise, then

$$C = \ log2\ \left(\sqrt{2\pi e}\ \sigma_y\right) - log2\ (\sqrt{2\pi e}\ \sigma_n)$$

Using previous expression of H(X) for Gaussian signal for the signal ***y*** with variance $\sigma_y{}^2$ then for the noise ***n(t)*** with variance $\sigma_n{}^2$(noise power).

$C = \ log2\ \left(\frac{\sigma_y}{\sigma_n}\right) = \frac{1}{2}log2\ \left(\frac{\sigma_y 2}{\sigma_n 2}\right)$, but $\sigma y2$ **=** $\sigma x2$**+** $\sigma n2$ sum of powers.

$\sigma x2$= S=Signal power and $\sigma n2$= N= noise power. Then

$$C = \frac{1}{2}log2\ \left(\frac{S+N}{N}\right) = \frac{1}{2}log2\ \left(1 + \frac{S}{N}\right) \text{ bits per sample}$$

For an analogue signal sampled at the Nyquist rate, then the sampling frequency is f_s= ***2B*** samples/sec, where B is the bandwidth of the signal, hence

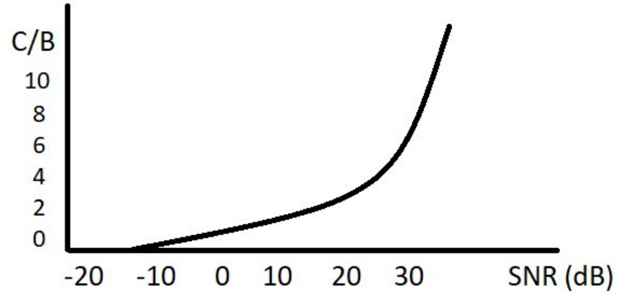
$$C = \frac{1}{2}log2\ \left(1 + \frac{S}{N}\right) \times 2B \qquad bits\ per \text{ sec} \qquad or$$

$$C = B\ log2(1 + \mathbf{SNR}) \qquad\qquad \text{bits per sec}$$

This is a very important formula known as SHANNON EQUATION named after C.E. Shannon, it is sometimes called Shannon- Hartly equation.

1- Care must be taken regarding the units, here B is in Hz., SNR=signal to noise power ratio is in absolute, then, C is in bits/sec. If SNR is given in dB, then: **SNR(absolute)=$10^{0.1(\text{SNR in dB})}$**.

2-The ratio [C/B]=$\log_2(1+SNR)$ gives what is called channel utilization ratio (bps per Hz) that increases with SNR as shown.



3-The equation C=B $\log_2(1+SNR)$ gives the maximum theoretical performance in terms of maximum bit rate that can be transmitted over a channel having a bandwidth B and SNR ratio.

Example: Find the maximum theoretical information rate that can be transmitted over a telephone channel having 3.5KHz bandwidth and 15dB SNR.

Solution:

C is the maximum theoretical information rate, using Shannon eq, then:

C=B $\log_2(1+SNR)$

where, SNR=15dB, changing into absolute SNR=$10^{0.1\times15}$=31., then:

C =3500 $\log_2(1+31)$=17500bps.

Example: A source produces 16 equiprobable symbols at a rate of 500 symbols/sec, check the possibility of transmitting this rate over the telephone channel of previous example

Solution:

First, we find the rate of information from the source, which is the source entropy rate R(X):

R(X)= H(X)× rate of symbols.

$H(X)=H(X)|max=\log_2 16=4$ bits/symbol (equiprobable case)

Then: R(X)=4 × 500= 2000 bps. Now since R(X) < 17500, then yes it is possible to transmit source output over this channel.

Example:

Find the minimum theoretical SNR required to transmit a compressed video information at a rate of 27Mbps over a channel having 5MHz bandwidth.

Solution:

For the minimum theoretical SNR, then put C=source bit rate =27Mbps, then:

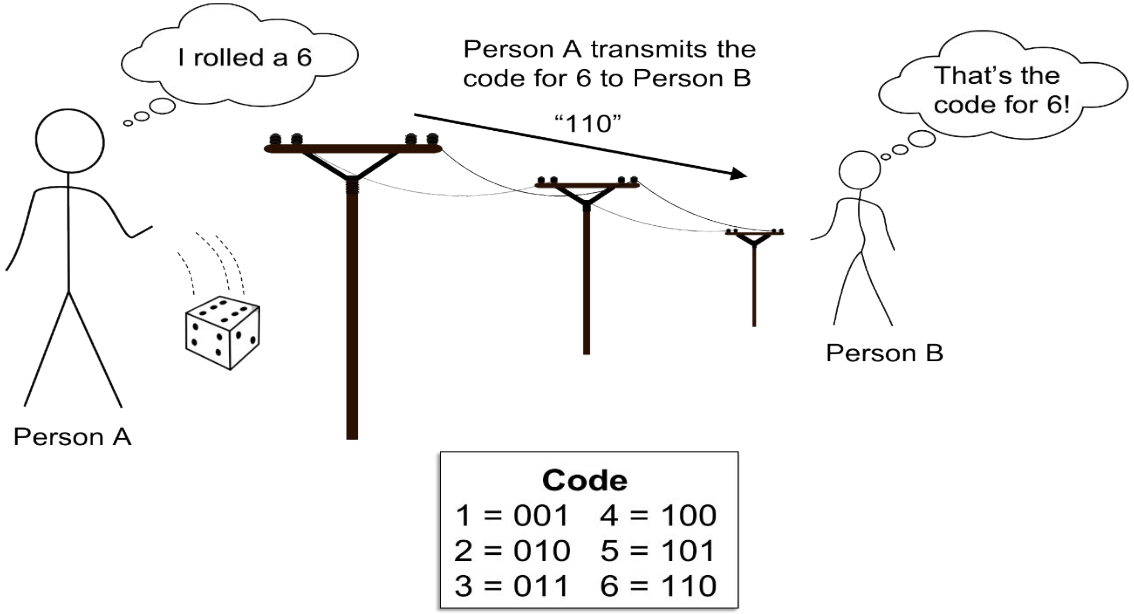
$C = B \log_2(1+SNR)$

$27\times 10^6 = 5\times 10^6 \log_2(1+SNR)$,or

$1+SNR = 2^{5.4}$ Ⓟ SNR=41.2 absolute or SNR=16.1 dB

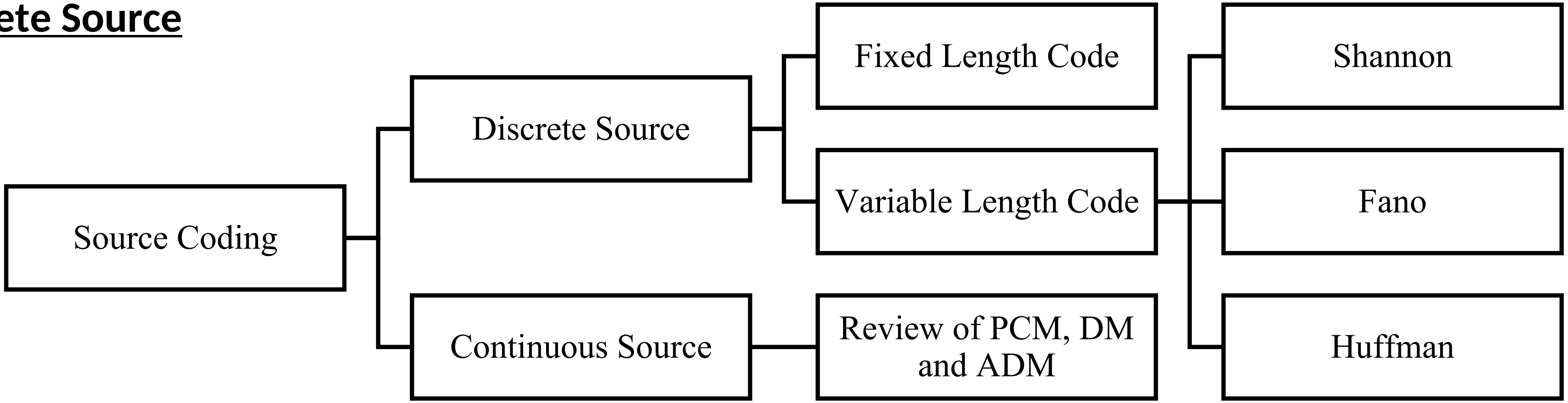
Source Coding of Discrete Source :.

An important point in communications is the efficient representation of data generated by a discrete source. The process by which this representation is accomplished is called source encoding. The device that performs the representation is called a source encoder. · For the source encoder to be efficient, we require knowledge of the statistics of the source. In particular, if some source symbols are known to be more probable than others, then we may exploit this feature in the generation of a source code by assigning short code words to frequent source symbols, and long code words to rare source symbols. ·



Source coding does not change or alter the source entropy, i.e. the average number of information bits per source symbol. In this sense source entropy is a fundamental property of the source. Thus the aim of source coding is to represent information as accurately as possible using as few bits as possible and in order to do so **redundancy from the source needs to be removed**. This will reduce the number of symbols in a message to the minimum necessary to represent the information in the message.

Source Coding of Discrete Source



A discrete Source is that source produces a finite set of messages x_1, x_2, ...x_n with prob. $p(x_1)$, $p(x_2)$, ...$p(x_n)$. A source coder will transform each message into a finite sequence of digits called "codeword" of the message. If binary digits are used in this codeword, then we obtain what is called "**binary source coding**". Ternary source coding also possible. The selection of codewords for different messages is *done* according to the following considerations:-

1- The average code length "***Lc*** " must be as minimum as possible. This ***Lc*** is given by:-

$$Lc = \bar{li} = \sum_{i=1}^{n} li\, P(xi) \quad \frac{bits}{symbol} \text{ where } li \text{ is the length of codeword of message } xi$$

2- The codewords at Rx must be **uniquely** decodable. **.** So that the original source sequence can be reconstructed perfectly.