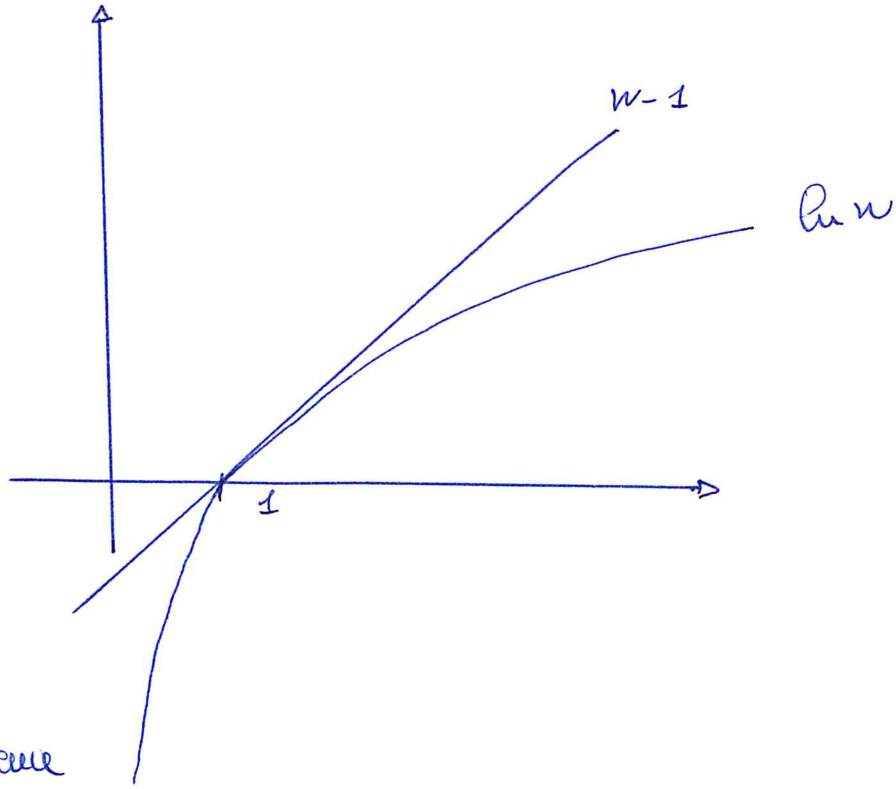
PROOFS FOR TIC 06/06/13

(1)

#1 Prove that $\ln(x) \leq x-1$, with $=$ iff (if and only if) $x=1$.

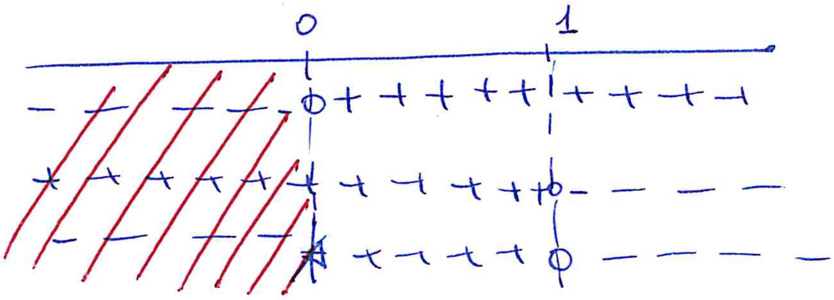


Studying the difference

$f(x) = \ln x - x + 1$

$$\frac{d f(x)}{dx} = \frac{d[\ln x - x + 1]}{dx} = \frac{1}{x} - 1 = \frac{1-x}{x}$$

$1 - x \geq 0 \qquad x \leq 1 \qquad ; \qquad x \geq 0$



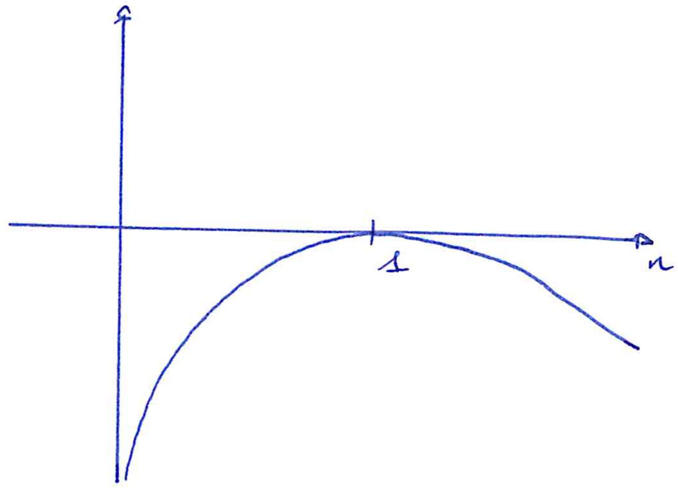
$\ln(x)$

$x \in \mathbb{R}^+$

$$\frac{d^2 f(x)}{dx} = -\frac{1}{x^2}$$

$-\frac{1}{x^2} \geq 0 \quad \nexists x \in \mathbb{R} \Rightarrow \cap$ It's a max!

$\Rightarrow$ The maximum is in $x = 1$



$\ln x - x + 1$ @ $x = 1 \Rightarrow 0 - 1 + 1 = 0 \Rightarrow \ln x = x - 1$ @ $x = 1$

Information

$$I(x) \equiv \log \frac{1}{p(x)}$$

Info given from the occurrence of the event x.

If $\log = \log_2 \Rightarrow I(x)\,[Sh]$ Shannon.

Prop.

x_i, x_j independent,

$p(x_i, x_j) = p(x_i)\,p(x_j) \Rightarrow I(x_i, x_j) = \log\left(\frac{1}{p(x_i)} \cdot \frac{1}{p(x_j)}\right) = I(x_i) + I(x$

Entropy

$$H(X) = \sum_{x} p(x) \log\left(\frac{1}{p(x)}\right) \equiv \mathbb{E}\left[I(X)\right] \quad [Sh]$$

If $p(x) \sim \quad U[\mathcal{X}]$

$$P(x_i) = \frac{1}{L}$$

$$H(P(x_i)) = \frac{1}{L}\sum_{i=1}^{N=L} \log\left(\frac{1}{L}\right) = -\frac{L}{L}\log\frac{1}{L} = +\log L$$

- $0 \leq H(X) \leq \log L$

$H(x) \geq 0$

Hp: $p(x) \geq 0$, $-\log(p(x))$ given $0 \leq p(x) \leq 1 \Rightarrow \geq 0$

$\Rightarrow H(x) \geq 0$

$H(x) \leq \log_2 L$

$$H(x) - \log_2 L = -\sum_{x} p(x)\log_2(p(x)) - \sum_{x} p(x)\log_2 L$$

$$= \sum_{x} p(x)\log_2\left(\frac{1}{p(x)L}\right)$$

Given $\ln x \leq x-1 \Rightarrow \log_2 x \leq \log_2 e\,(x-1)$

$$\leq \sum_{x} p(x)\log_2 e\left(\frac{1}{p(x)L} - 1\right) = \log_2 e \sum_{x}\left(\frac{1}{L} - p(x)\right)$$

$$= \log_2 e\,(1-1) = 0$$

$\Rightarrow H(x) \leq \log_2 L \Rightarrow 0 \leq H(x) \leq \log_2 L$

The entropy is never < 0 !

Given x, y

- $H(X) = \sum_x p(x) \log\left(\frac{1}{p(x)}\right)$
- $H(Y) = \sum_y p(y) \log\left(\frac{1}{p(y)}\right)$
- $H(X,Y) = \sum_x \sum_y p(x,y) \log \frac{1}{p(x,y)}$
- $H(X|Y=y_i) = \sum_x p(x|Y=y_i) \log \frac{1}{p(x|Y=y_i)}$
- $H(X|Y) \equiv \mathbb{E}\left[H(X|Y=y_i)\right] = \sum_y \sum_x p(y)\, p(x|Y=y_i) \log \frac{1}{p(x|Y=y_i)}$

$$= \sum_y \sum_x p(x,y) \log \frac{1}{p(x|Y=y_i)}$$

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

$$p(x,y) = p(x|Y=y_i)\, p(y)$$

$$\sum_x \sum_y p(x|Y=y_i)\, p(y) \log \frac{1}{p(x|Y=y_i)\, p(y)}$$

$$= \sum_x \sum_y p(x|Y=y_i)\, p(y) \log \frac{1}{p(x|Y=y_i)} + \underbrace{\sum_x \sum_y p(x|Y=y_i)}_{=1} p(y) \log \frac{1}{p(y)}$$

$$\sum_y \sum_x p(y)\, p(x|Y=y_i) \log \frac{1}{p(x|Y=y_i)} + \sum_y p(y) \log \frac{1}{p(y)} = H(X|Y) + H(Y)$$

• $H(X|Y) \leq H(X)$

$$H(X|Y) - H(X) = \sum_{x} \sum_{y} p(y)\, p(x|Y=y_i) \log \frac{1}{p(x|Y=y_i)} - \sum_{x} p(x) \log \frac{1}{p(x)}$$

$$= \sum_{x} \sum_{y} p(x,y) \left(\log \frac{1}{p(x|Y=y_i)} \right) - \sum_{x} \sum_{y} p(x,y) \log \frac{1}{p(x)}$$

$$= \sum_{x} \sum_{y} p(x,y) \left(\log \frac{1}{p(x|Y=y_i)} + \log p(x) \right)$$

$$= \sum_{x} \sum_{y} p(x,y) \left(\log \frac{p(y)}{p(x,y)} + \log(p(x)) \right)$$

Where $p(x|Y=y_i) \cdot p(y) = p(x,y)$

$$\Rightarrow p(x|Y=y_i) = \frac{p(x,y)}{p(y)}$$

$$= \sum_{x} \sum_{y} p(x,y) \log \frac{p(y)\, p(y)}{p(x,y)}$$

$$\leq \sum_{x} \sum_{y} p(x,y) \log_2 e \left(\frac{p(y)\, p(x)}{p(x,y)} - 1 \right)$$

$$= \log_2 e \sum_{x} \sum_{y} \left(-p(x,y) + p(y)\, p(x) \right) = \log_2 e\, (1 - 1) = 0$$

$$H(X|Y) - H(X) \leq 0 \qquad H(X|Y) \leq H(X)$$

- Chain Rule

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

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i-1}, \dots, X_1)$$

- $H(X_1, X_2, \dots, X_n) \leq \sum_{i=1}^{n} H(X_i)$

$$H(X_k | X_{k-1} \dots X_1) \leq H(X_k)$$

$\Rightarrow$ For the chain rule $\Rightarrow H(X_1, X_2, \dots, X_n) \leq \sum_{i=1}^{n} H(X_i$

* DISCRETE MEMORY-LESS SOURCE (DMS)

$$\boxed{\text{DMS}} \xrightarrow{x_1,\, x_2,\, \dots}$$

rand. process discrete values
time discrete, ergodic
(memory less & stationary)

$$X_i \in \mathcal{X} = \{\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_n\}$$

$$P(X_{n1} = w_1, X_{n2} = w_2, X_{n3} = w_3, \dots) = \prod_{i=1}^{M} P(X_{ni} = w_i)$$

Every symbol is independent.

- Entropy of DMS

$$H(X) = H(X_n) \qquad \left[\frac{Sh}{symbol}\right]$$

The same memory lm channel has each time the same information

- Kullback-Leibler Distance

Given $\underline{a} = (a_1, a_2, a_3, \dots, a_n)$ and $\underline{b}$,
let $\sum \underline{a} = 1$ and $\sum \underline{b} = 1$

$$\sum_i a_i \log \frac{a_i}{b_i} \geq 0$$

$$\sum a_i \log \frac{a_i}{b_i} \Rightarrow -\sum a_i \log \frac{b_i}{a_i} \leq \log_2 e \sum a_i \left(\frac{b_i}{a_i} - 1\right)$$

$$= \log_2 e \sum (b_i - a_i) = \log_2 e \left(\sum b_i - \sum a_i\right) = 0$$

$$\sum a_i \log \frac{b_i}{a_i} \leq 0$$

$$\Downarrow$$

$$\sum a_i \log \frac{a_i}{b_i} \geq 0$$

It can be extended to the continuous case

$$\int a(x) \log \frac{a(x)}{b(x)} \, dx \geq 0$$

Let $X_1, X_2, \ldots$ be stationary process.

$$P_r\{X_1 = w_1, X_2 = w_2, \ldots, X_n = w_n\} = P_r\{X_{1+e} = w_1, X_{2+e} = w_2, \ldots, X_{n+e} = w_n\}$$

$$H(X_n) = H(X)$$

If it's possible to look at more symbols

$$H(X_1, X_2, X_3, X_4, \ldots, X_k)$$

$\Rightarrow$ The mean info for every symbol is

$$\frac{H(X_1, X_2, \ldots X_k)}{k}$$

and

$$\frac{H(X_1, X_2, \ldots, X_k)}{k} \le \frac{\sum_{i=1}^{k} H(X_i)}{k} = H(X)$$

The ENTROPY OF THE SOURCE or SOURCE ENTROPY is

$$H_\infty = \lim_{k \to \infty} \frac{H(X_1, X_2, \ldots, X_k)}{k}$$

If the process is stationary, this is equal to

$$H'_\infty = \lim_{k \to \infty} H(X_k | X_{k-1}, \ldots X_1)$$

- $H'_\infty(X) = H_\infty(X)$ if stationary

$$\underbrace{H(X_{K+1} | X_K, \dots, X_1)}_{a_{K+1}} \leq H(X_{K+1} | X_K, \dots, X_2) = \underbrace{H(X_K | X_{K-1}, \dots, X_1)}_{a_K}$$

$a_{K+1} \leq a_K \implies$ The series is decreasing $\implies$ It converges

$$H_\infty(X) = \lim_{K \to \infty} \frac{H(X_K, X_{K-1}, \dots, X_1)}{K} = \lim_{K \to \infty} \frac{\sum_{i=1}^{K} H(X_i | X_{i-1}, \dots, X_1)}{K}$$

Cesaro Mean

$a_n \xrightarrow{n \to \infty} a$

$\frac{1}{N} \sum_{K=1}^{N} a_K = b_n \qquad b_n \xrightarrow{n \to \infty} a$

$$= \lim_{K \to \infty} \underbrace{\frac{1}{K} \sum_{i=1}^{K} \underbrace{H(X_i | X_{i-1}, \dots, X_1)}_{a_i}}_{b_i} = H_\infty(X)$$

Given that $b_i \xrightarrow{\infty} H_\infty(X)$, for Cesaro the $a_i \xrightarrow{\infty} H_\infty(X)$.

But $a_{i \to \infty} = H'_\infty(X)$

$\implies \boxed{H'_\infty(X) = H_\infty(X)}$

Source Encoding and Data Compression

$$\boxed{S} \xrightarrow[B_s \left[\frac{sym}{sec}\right]]{H_\infty(X)} \qquad H_\infty(X) \cdot B_s = \left[\frac{Sh}{sec}\right]$$

$\dim \mathcal{X} = L$

If the symbols are equiprobable and ind $\Rightarrow \log_2 L = H_\infty(X)$

The info should be preserved, so

$$H_\infty(X) \cdot B_s = \log_2 L \cdot B_s'$$

$$\Rightarrow \frac{H_\infty(X)}{\log_2 L} = \frac{B_s'}{B_s} \leq 1 \qquad = \text{ iff ind. and uniform.}$$

- Non singular code

$$C(w_i) \neq C(w_j) \qquad \forall w_i \neq w_j$$

- Uniquely decodable code

$$C(w_1), C(w_2), \dots \longrightarrow w_1, w_2, w_3, \dots$$

each codeword brings to the original code.

- Prefix code (instant code)

There is not any codeword that is prefix to another one.

- Medium code lenght

$$L(\mathcal{C}) = \mathbb{E}\left[\ell(X)\right] = \sum_x p(x)\,\ell(x) \qquad \left[\frac{bit}{source\ sym}\right]$$

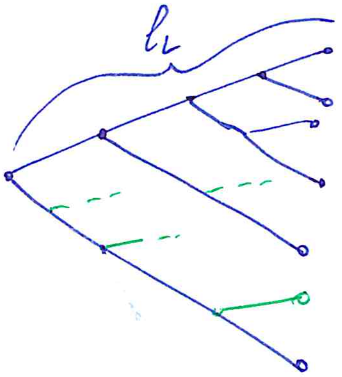
For any prefix code there is a binary tree.

- Kraft's Inequality

For any prefix code with $l_1 \leq l_2 \leq l_3 \leq \dots \leq l_L$ codeword lenght

$$\sum_{i=1}^{L} 2^{-l_i} \leq 1$$

Furthermore, for any $l_1 \leq l_2 \leq l_3 \leq \dots \leq l_L$ that satisfies Kraft there is a prefix code.



The total nodes are 2^{l_L}. If I remove from l_I, ~~nodes~~, the nodes will be $\sum_{i=1}^{L} 2^{l_L - l_i} \leq 2^{l_L}$

This means $$\sum_{i=1}^{L} 2^{-l_i} \leq 1$$

The same reasoning canbe be done from the end.

- Mc Millan: The Kraft's Inequality is valid for any uniquely decodable codes.

Source Encoding Theorem : Shannon 1948

Any code applied to a stationary source with X_n and entropy $H(X)$ has to obey

$$H(X) \leq \mathbb{E}\left[\ell(n)\right] \leq H(X) + 1$$

with more symbols, the mean entropy is

$$\frac{H(X)}{k} \leq \frac{\mathbb{E}\left[\ell(k)\right]}{k} \leq \frac{H(X)}{k} + \frac{1}{k}$$

$$\text{if } k \to \infty \quad \Rightarrow \quad \mathbb{E}\left[\ell(n)\right] = H(X)$$

- DIM

$$H(X) \leq \mathbb{E}\left[\ell(n)\right]$$

$$H(X) - \mathbb{E}\left[\ell(x)\right] = \sum_n p(n) \log \frac{1}{p(n)} - \sum_n p(n)\, \ell(n)$$

$$= \sum_n p(n) \left(\log \frac{1}{p(n)} - \log_2 2^{\ell(n)} \right) = \sum_n p(n) \log \left(\frac{1}{p(n)\, 2^{\ell(n)}} \right)$$

$$\leq \log_2 e \sum_n p(n) \left(\frac{1}{p(n)\, 2^{\ell(n)}} - 1 \right) = \log_2 e \left(\sum_n \frac{1}{2^{\ell(n)}} - \sum_n p(n) \right)$$

$$= \log_2 e \left(\underbrace{\sum_n 2^{-\ell(n)}}_{\leq 1} - \underbrace{\sum_n p(n)}_{1} \right) \leq 0$$

$$\mathbb{E}\left[\ell(n)\right] \geq H(n)$$

06/06/13 (8)

$$H(X)+1 \le E[\ell(n)]$$

To dem. is used the Shannon-Fano code

$$\ell(n) = \sup[\log_2(p(n))]$$

$$\underbrace{\log_2(p(n)) \le \ell(n)} < \log_2(p(n)) + 1$$

$$\Downarrow$$

$$2^{\log_2(p(n))} \le 2^{\ell(n)} \qquad \sum p(n) \le \sum_n 2^{\ell(n)}$$

$$1 \le \sum_n 2^{\ell(n)} \qquad \Rightarrow \sum_n 2^{-\ell(n)} \le 1$$

It satisfies McMillan, so it is a possible code.

$$\log_2(p(n)) \le \ell(n) < \log_2(p(n)) + 1$$

$$\sum_n p(n)\log_2(p(n)) \le \sum_n p(n)\log_2(p(n)) < \sum_n p(n)\log_2(p(n)) + \sum p(n)$$

$$\underline{H(x) \le E[\ell(n)] < H(x) + 1}$$

HUFFMAN (1952)

3 steps

- Order the probabilities in a decreasing order.
- Sum the two ~~least~~ least prob and create a new node
- Reorder and redo the steps

Block Encoding

$$K \to \infty \qquad \frac{H(X)}{K} \le \frac{\mathbb{E}(\ell(x))}{K} < \frac{H(X)+1}{K} \to 0$$

$$\Rightarrow \quad H(X) = \mathbb{E}[\ell(x)]$$

Tipical sequences

DMS

$\underline{X} = X_1, X_2, \dots, X_n$ iid $X_i \in \mathcal{X} = \{\tilde{w}_1, \dots, \tilde{w}_L\}$

With $n \gg 1$

$$\underline{X} = \underbrace{w_1 w_1 \dots w_1}_{np(w_1)} \underbrace{w_2 \dots w_2}_{np(w_2)} \underbrace{w_3 \dots w_3}_{np(w_3)} \dots$$

For the large number's law, $np(w)$ times that occurrence

$$p(\underline{X}) = p(w_1)^{np(w_1)} p(w_2)^{np(w_2)} \dots = \prod_{i=1}^{L} p(w_i)^{np(w_i)}$$

$$\log_2 p(\underline{X}) = \sum_i np(w_i) \log(p(w_i)) = -nH(X)$$

$\log_2 p(\underline{X}) = -n H(x)$ $\qquad P(\underline{X}) = 2^{-nH(x)}$

Number of tipical sequences $\Rightarrow \frac{1}{P(X)} = 2^{nH(X)}$

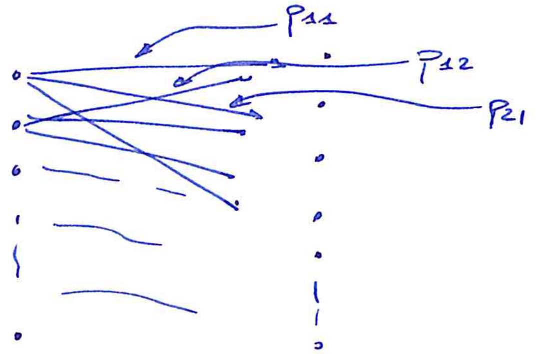
Transmission Channels and capacity

DMC , Discrete Memory-Less Channel

DMS , Discrete Memory-Less Source

The DMC channel is defined by its TRANSITION MATRIX $\underline{P}$

$$\underline{P} = \{p_{ij}\} = P_r\{Y = y_j \mid X = x_i\}$$



Let recall

$H(X)$, $H(Y)$ and $H(X|Y) \le H(X)$ = iff indip.

The <u>Mutual Information</u> is defined as

$$I(X;Y) = H(X) - H(X|Y) \quad \left[\frac{Sh}{sym}\right]$$

It is the reduction of the uncertainty after the observation of the output Y.

$$0 \le I(X,Y) \le H(X)$$

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

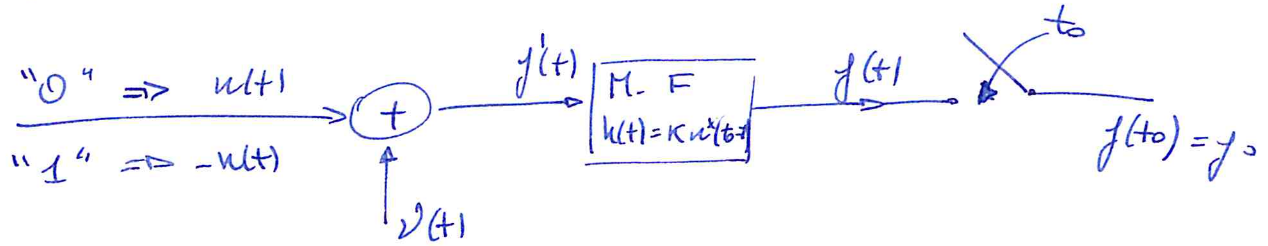
$$H(X) - H(X|Y) = \cancel{H(X)} - H(Y|X) - \cancel{H(X)} + H(Y)$$

$$= H(Y) - H(Y|X)$$

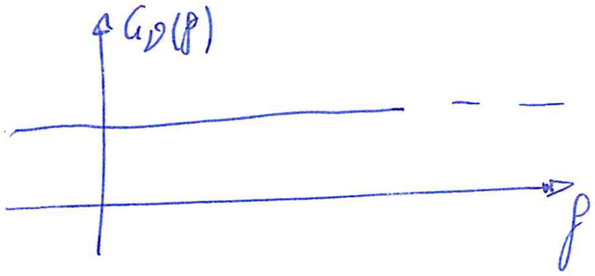
Channel Capacity.

$$C = \max I(X;Y) \qquad \left[Sh/sym \right] = \left[Sh/chann.\ use \right]$$

Antipodal - Matched Filter



$\nu(t)$ = AWGN



$h(t) = K w^*(t_0 - t)$ MATCHED FILTER

$$y(t) = \underbrace{h(t) \otimes \nu(t)}_{y_n(t)} + \underbrace{h(t) \otimes w(t)}_{y_s(t)}$$

$$y_s(t) = \int_{\mathbb{R}} w(\xi) h(t-\xi)\, d\xi = \int_{\mathbb{R}} w(\xi) K w^*(t_0 - (t-\xi))\, d\xi =$$

$$y_s(t_0) = \int_{\mathbb{R}} h(\xi)\, K\, h^*(t_0 - t_0 + \xi)\, d\xi = K \int_{\mathbb{R}} h(\xi)\, h^*(\xi)\, d\xi = K \underbrace{\int_{\mathbb{R}} |h(\xi)|^2\, d\xi}_{E_x}$$

$$= K E_x$$

$$y_n(t) = \int_{\mathbb{R}} G_v(f)\, |H(f)|^2\, df = \frac{N_0}{2} \int_{\mathbb{R}} |H(f)|^2\, df \overset{\text{Parseval}}{=} \frac{N_0}{2} \int_{\mathbb{R}} |h(t)|^2\, dt$$

$$= \frac{N_0}{2} \int_{\mathbb{R}} K^2 |h(t_0 - t)|^2\, dt = \frac{N_0 K^2}{2} \int_{\mathbb{R}} |h(t)|^2\, dt = \frac{N_0 K^2}{2} E_x$$

That is the power of the noise. At this point it is not still white but "coloured" by the filter. Anyway it is still Gaussian with $\mu = 0$ $\Rightarrow$ $y_n \sim \mathcal{N}\left(0, \sigma^2 = \frac{N_0 K^2 E_x}{2}\right)$

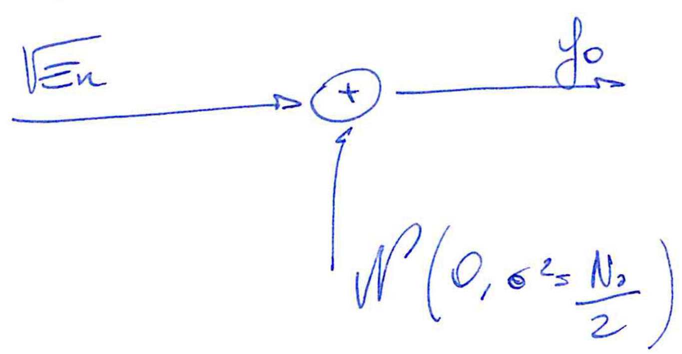
The total output signal

$$y_o = \mathcal{N}\left(\pm K E_x,\ \sigma^2 = \frac{N_0 K^2 E_x}{2}\right)$$

If we choose $K = \frac{1}{\sqrt{E_x}}$ it follows

$$y_o = \mathcal{N}\left(\pm \sqrt{E_x},\ \sigma^2 = \frac{N_0}{2}\right)$$

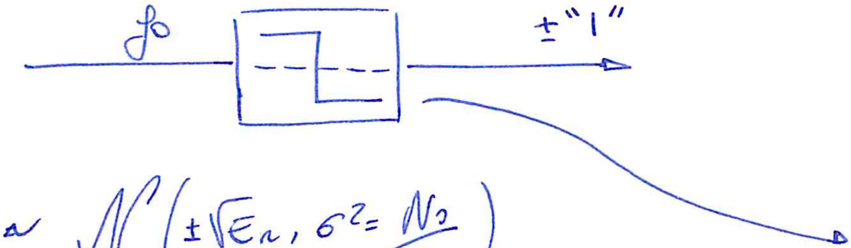
That is equivalent to write

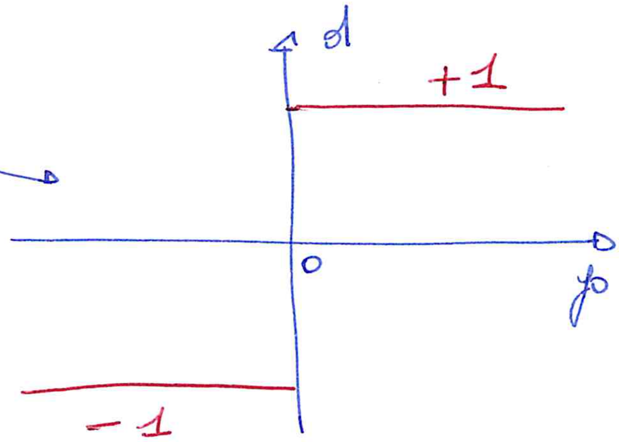


The equivalent time-discrete is referred to a bit.
That was "Antipodal Transmission over the AWGN channel, matched filter".

Symbol recognition

- HARD DECISION (HD)

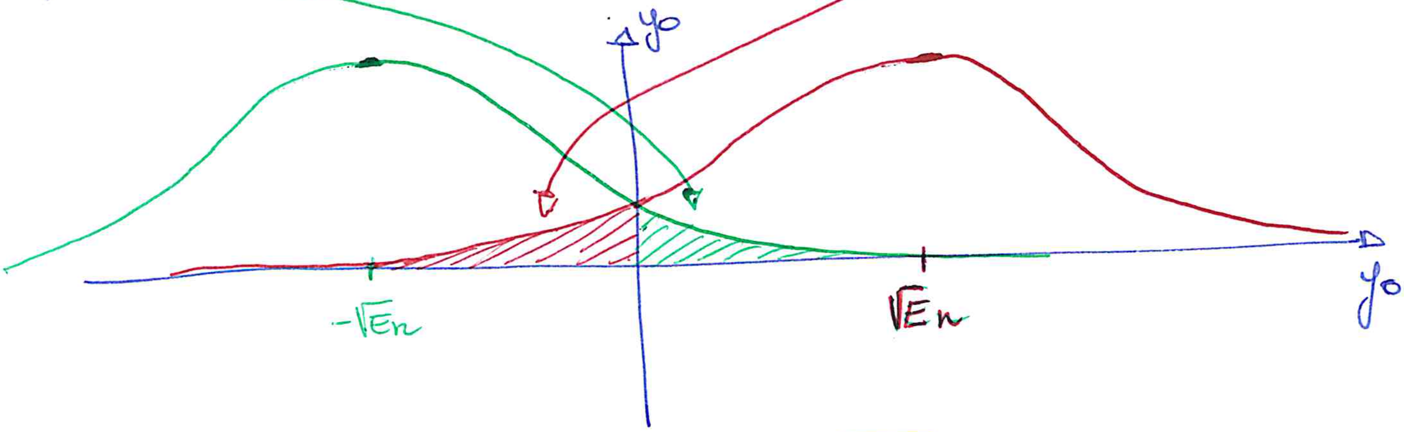




$y_0 \sim \mathcal{N}\left(\pm\sqrt{E_n}, \sigma^2 = \frac{N_0}{2}\right)$

Since the signal is symmetric

$$\underbrace{P\left(d=1 \mid -n(t)\right)}_{=\,p} = P\left(y>0 \mid -n(t)\right) = \underbrace{P\left(d=-1 \mid n(t)\right)} = P\left(y<0 \mid n(t)\right)$$



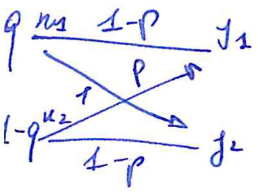
$$P_e = P_e = \frac{1}{2}\,\mathrm{erfc}\,\frac{\sqrt{E_n}}{\sqrt{2\frac{N_0}{2}}} = \underline{\frac{1}{2}\,\mathrm{erfc}\sqrt{\frac{E_n}{N_0}}} = p$$

06-06-13 (11)

Being symmetric, it is possible to consider a BSC channel with error prob. p.

Since $C_{BSC} = 1 - \mathcal{H}(p)$ where $\mathcal{H}(p)$ is the entropy of a Bernulli's Distribution with parameter "p"

BSC



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

$H(Y|X)$:

$$H(Y|X = x_1) = \mathcal{H}(p)$$
$$H(Y|X = x_2) = \mathcal{H}(p)$$
$$H(Y|X) = p(x_1)\mathcal{H}(p) + p(x_2)\mathcal{H}(p) = \mathcal{H}(p)\big(p(x_1) + p(x_2)\big)$$
$$= \mathcal{H}(p)$$

$H(Y)$:

$$p(y_1) = q(1-p) + (1-q)p = q - qp + p - qp = q + p - 2qp$$
$$p(y_2) = (1-q)(1-p) + qp = 1 + qp - q - p + qp = 1 - q - p + 2qp$$
$$\{p(y_1) + p(y_2) = 1$$

$$\Rightarrow H(Y) = \mathcal{H}(q + p - 2qp)$$

Since it is symmetric it is easy to think that the max capacity is with $H(Y) = 1$

$$\Rightarrow \underline{C = 1 - \mathcal{H}(p)}$$

· To use more time the DMC does not increase the capacity

$$I(X^{(m)}; Y^{(m)}) \le mC$$

$$I(X^{(m)}; Y^{(m)}) = H(Y^{(m)}) - H(Y^{(m)} | X^{(m)})$$

$$H(Y^{(m)}) \le \sum_{i=1}^{m} H(Y_i)$$

$$H(Y^{(m)} | X^{(m)}) = \sum_{i=1}^{m} H(Y_i | Y_{i-1}, \dots, Y_1, X_m, X_{m-1}, \dots X_1)$$

Since it is memory less:

$$= \sum_{i=1}^{m} H(Y_i | X_i)$$

$$\Rightarrow I(X^{(m)}; Y^{(m)}) \le \sum_{i=1}^{m} H(Y_i) - \sum_{i=1}^{m} H(Y_i | X_i) = \sum_{i=1}^{m} H(Y_i) - H(Y_i | X_i)$$

$$= mI(X;Y)$$

$$I(X^{(m)}; Y^{(m)}) \le mI(X;Y) = mC$$

· Channel Coding

Code Rate $R_c = \frac{\log_2 L}{m}$

Before the ch. encoder the channel cap. is

$$\frac{H_\infty(\mathcal{U})}{T}$$

After the encoder is

$$\frac{H_\infty(\mathcal{U})}{T_c}$$

This capacity has to be the same

$$\frac{H_\infty(w)}{T} = \frac{H_\infty(x)}{T_c}$$

$$\frac{H_\infty(w)}{H_\infty(x)} = \frac{T}{T_c} \qquad \frac{H_\infty(w)}{C} \leq \frac{T}{T_c} \quad \text{where } H_\infty(x) \leq C$$

$$\frac{T_c}{T} \leq \frac{C}{H_\infty(w)}$$

if $H_\infty(w) = \log_2 L$

$$\frac{T_c}{T} \leq \frac{C}{\log_2 L} \qquad \frac{1}{n} \leq \frac{C}{\log_2 L} \qquad n \geq \frac{\log_2 L}{C}$$

$$C \geq \frac{\log_2 L}{n}$$

The code rate should be less than the channel capacity.

CHANNEL CODING THEOREM, Shannon 1948

- $\forall \varepsilon > 0$, $R_c < C$
  there are codes that, using an high n, allows the err. probability ~~…~~ to be under ε
- Every code with $\varepsilon \to 0$ $(n \to \infty)$ has to have $R_c < C$.

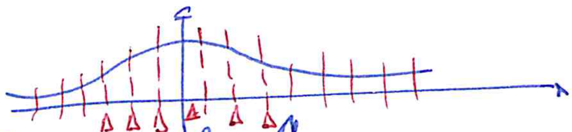
R.V.s continues

We define the Differential Entropy as

$$h(X) = \int_{\mathbb{R}} p(x) \log \frac{1}{p(x)} \, dx$$

It comes from:

Given a p.d.f.



we take some little intervals Δ

$$P(x_i) = \int_{\Delta_i} p(x)\, dx = p(x_i)\Delta$$

(there is a point inside the interval that multiplied for Δ gives $P(x_i)$)

$$H(X) = \sum_x P(x_i) \log\left(\frac{1}{P(x_i)}\right) = \sum_x p(x_i)\Delta \log\left(\frac{1}{p(x_i)\Delta}\right)$$

$$= \underbrace{\sum_x p(x_i) \log\left(\frac{1}{p(x_i)}\right)\Delta} + \underbrace{\sum_x p(x_i)}_{1} \log\left(\frac{1}{\Delta}\right)\Delta$$

$$\Delta \to 0 \Rightarrow \Delta = dx \qquad \sum_x \Rightarrow \int_{\mathbb{R}}$$

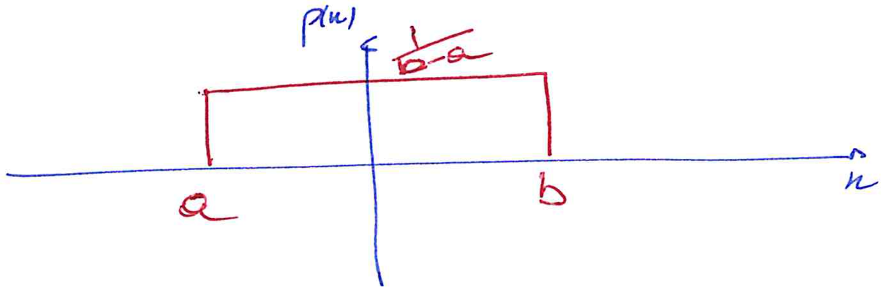
$$\int_{\mathbb{R}} p(x) \log\left(\frac{1}{p(x)}\right) dx + \log\left(\frac{1}{\Delta}\right) \Rightarrow \infty$$

This is interesting if the entropy is needed to calculate the uncertainty as something differential. It is not a measure of the uncertainty, since it should be ∞.

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

- Uniform distribution

$X \sim U[a,b]$



$$h_u(X) = \int_{\mathbb{R}} p(x) \log \frac{1}{p(x)} \, dx = \int_a^b \frac{1}{b-a} \log(b-a) \, dx$$

$$= \underline{\log(b-a)}$$

It could be even < 0!

- The uniform distr. has the highest $h(X)$ for every limited interval distr.

$h_u(x) - h(x)$

$$\log(b-a) - \int_{\mathbb{R}} p(x) \log\left(\frac{1}{p(x)}\right) dx = \int_{\mathbb{R}} p(x) \log(b-a) - \int_{\mathbb{R}} p(x) \log\left(\frac{1}{p(x)}\right) dx$$

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

$$= \int_{\mathbb{R}} p(x) \log\left(\frac{p(x)}{p_u(x)}\right) \geq 0 \quad \Rightarrow \quad h_u(x) \geq h(x)$$

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$h(x) = \int_{\mathbb{R}} p(x) \log \frac{1}{p(x)} dx = \int_{\mathbb{R}} p(x) \log \sqrt{2\pi\sigma^2} dx + \int_{\mathbb{R}} p(x) \cancel{\log} \left(+ \frac{(x-\mu)^2}{2\sigma^2} \log e \right) dx$$

$$= \log \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \underbrace{\int_{\mathbb{R}} p(x) (x-\mu)^2 dx}_{\sigma^2} \cdot \log e = \log \sqrt{2\pi\sigma^2} + \frac{1}{2} \log e$$

$$= \log \sqrt{2\pi\sigma^2} + \log \sqrt{e} = \underline{\log \sqrt{2\pi\sigma^2 e}}$$

The gaussian distr. is the one with the highest $h(X)$ for the non limited intervals distr.

$$h_G(X) - h(X) = \log \sqrt{2\pi\sigma^2 e} - \int_{\mathbb{R}} p(x) \log \frac{1}{p(x)} dx$$

$$= \int_{\mathbb{R}} p(x) \log \sqrt{2\pi\sigma^2 e} + \int p(x) \log p(x) dx$$

$$= \int_{\mathbb{R}} p(x) \log \frac{1}{p_G(x)} + \int_{\mathbb{R}} p(x) \log p(x) dx$$

To demonstrate take $\int_{\mathbb{R}} p(x) \log \frac{1}{p_G(x)}$ and you will arrive to

$$= \int_{\mathbb{R}} p(x) \sqrt{2\pi\sigma^2 e} \, dx$$

$$= \int p(x) \log\left(\frac{p(x)}{p_a(x)}\right) dx \geq 0$$

$$h_u(x) \geq h(x)$$

• Exponential dn

$$p(x) = \begin{cases} a e^{-bx} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Normalization

$$\int_{\mathbb{R}} a e^{-bx}\, dx = 1 \qquad \left[\frac{a e^{-bx}}{-b}\right]_0^{+\infty} = +\frac{a}{b} = 1 \qquad a = b$$

$$p(x) = \begin{cases} a e^{-ax} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$h_{exp}(x) = \int_{\mathbb{R}} p(x) \log \frac{1}{p(x)}\, dx = a \int e^{-ax} \log(a e^{+ax})\, dx$$

$$= \underbrace{a \int e^{-ax} \log a\, dx}_{=1} + a \int e^{-ax} \log(e^{+ax})\, dx$$

$$= \log a + a^2 \int e^{-ax}\, x \log e\, dx$$

$= \log a + a \underbrace{\int a e^{-ax} \cdot x \, dx}_{\mu} \cdot \log e$

$= \log a + \mu a \log e$

The expon. has the highest $h(X)$ of ~~any~~ every dx illimitated distr.

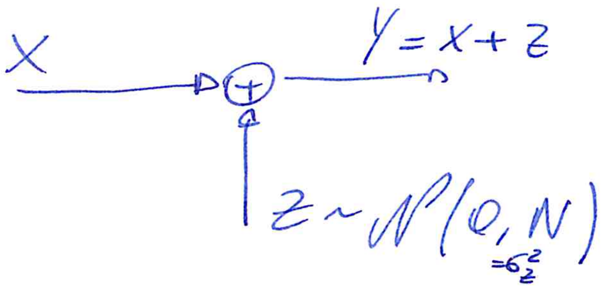
$h_{exp}(X) - h(X) = \underbrace{\log a + \mu a \log e}_{\log(a e^{\mu a}) \; \log\left(\frac{1}{p_{esp}}\right)} + \int_{\mathbb{R}} p(x) \log p(x) \, dx$

$\int_{\mathbb{R}} p(x) \left(\frac{p(x)}{p_{esp}(x)} \right) dx \geq 0$

$h_{exp}(X) \geq h(X)$

AWGN Channel capacity

- Soft Decision



Supposing $E[X^2] = S$

$h(Y|X) = \log\sqrt{2\pi\sigma_z^2 e}$ Because fixed X the only variable is z

$h(Y)$:

$I(X;Y) = h(Y) - h(Y|X)$ to maximize, since $h(Y|X)$ is maximum $h(Y)$ should be maximum and should be Gaussian

$$C = \log\sqrt{2\pi\sigma_y^2 e} - \log\sqrt{2\pi\sigma_z^2 e} = \frac{1}{2}\log\frac{\sigma_y^2}{\sigma_z^2} = \frac{1}{2}\log\frac{\sigma_u^2+\sigma_z^2}{\sigma_z^2}$$

X v.m. = 0

$$= \frac{1}{2}\log\left(1+\frac{\sigma_u^2}{\sigma_z}\right) \stackrel{!}{=} \frac{1}{2}\log\left(1+\frac{S}{N}\right)$$

- Time continuous, band-limited. AWGN

If the signal $u(t)$ has a band B, it has to be sampled with a f. $2B$.
In a time τ, there will be $2B\tau$ samples.
Let use this result and think every sample as a discrete sample.

$$C_i = \frac{1}{2}\log\left(1+\frac{S}{N}\right) \Rightarrow$$

$$C = \frac{2B\tau}{2}\log\left(1+\frac{S}{\frac{2N_0B}{2}}\right) = B\tau\log\left(1+\frac{S}{N_0B}\right)$$

If time normalized

$$\frac{C}{\tau} = B\log\left(1+\frac{S}{N_0B}\right)$$

That is the maximum theoretically admissible for an AWGN channel.

Hartley-Shannon Formula

If we define $S_r = \left[\frac{Sh}{sec}\right]$ Bit rate at the end user

$$E_{sh} = \frac{S}{S_r} \qquad \eta = \frac{S_r}{B} \qquad S_r \leq C$$

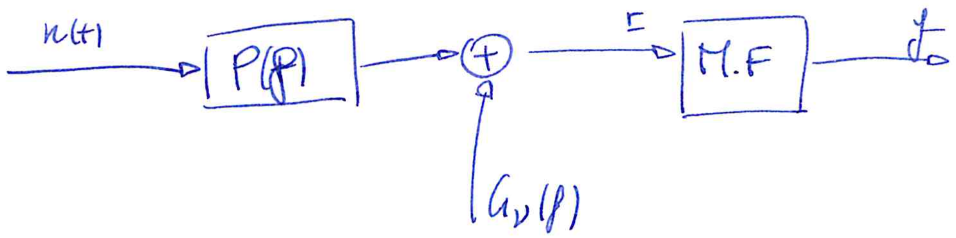
$$S_r \leq B\log\left(1+\frac{S}{N_0B}\cdot\frac{S_r}{S_r}\right) \Rightarrow \frac{S_r}{B} \leq \log\left(1+\frac{\eta E_{sh}}{N_0}\right)$$

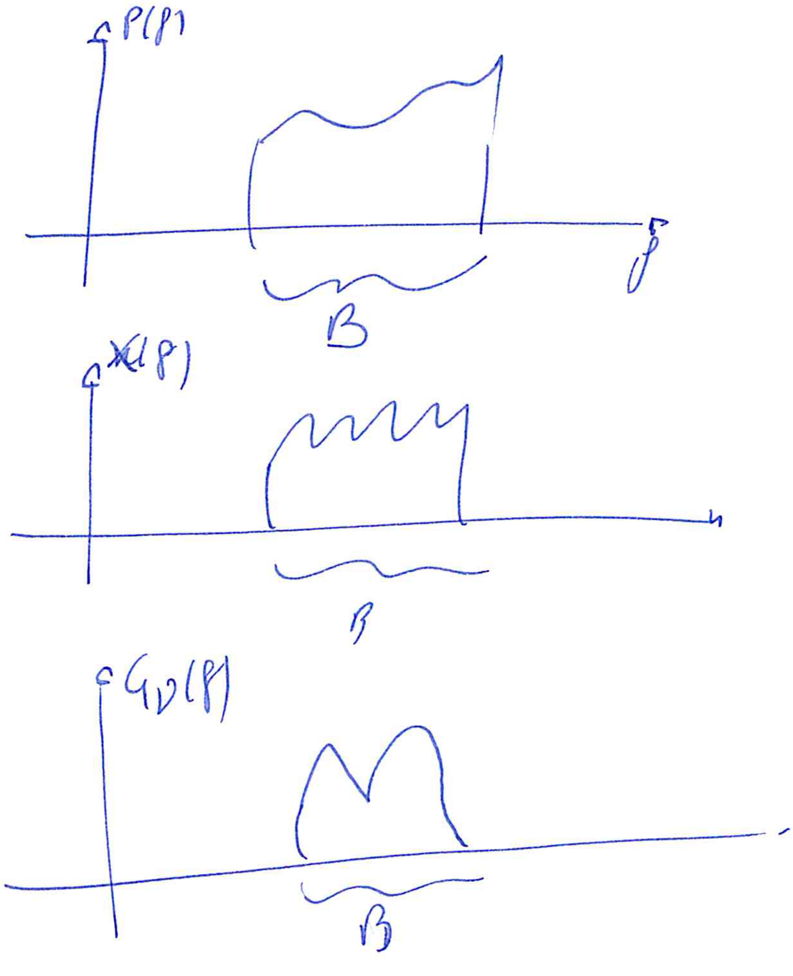
$$\eta \leq \log\left(1 + \eta \frac{E_{sh}}{N_0}\right)$$

$$2^{\eta} \leq 1 + \eta \frac{E_{sh}}{N_0} \qquad \underline{\frac{2^{\eta}-1}{\eta} \leq \frac{E_{sh}}{N_0}}$$

This has to be verified for every transmission on an AWGN channel

- Non AWGN channel





Let think using an asymptotic approach:
We divide the channel in subchannels with bandwidth Δf

$$C_i = \Delta f \log\left(1 + \frac{G_n(f_i)\cdot|P(f_i)|^2\cdot \cancel{2\Delta f}}{G_v(f_i)\cdot\cancel{2\Delta f}}\right)$$

$$= \Delta f \log\left(1 + \frac{G_n(f_i)\cdot|P(f_i)|^2}{G_v(f_i)}\right)$$

Now $\Delta f \to 0$ and sum every contribute

$$C = \int \log\left(1 + \frac{G_n(f)\,|P(f)|^2}{G_v(f)}\right) df$$

To know how should be $G_n(f)$ ~~let~~ try to maximize for $G_n(f)$.
Obviously there is a constraint.

$$S = 2\int_B G_n(f)\,df$$

To maximize with a constraint, we use the Lagrange's Multipliers.

$$\frac{d}{d\,G_n(f)}\left[\int\left(\log\left(1+\frac{G_n(f)\,|P(f)|^2}{G_v(f)}\right) - \lambda\, G_n(f)\right)df\right] = 0$$

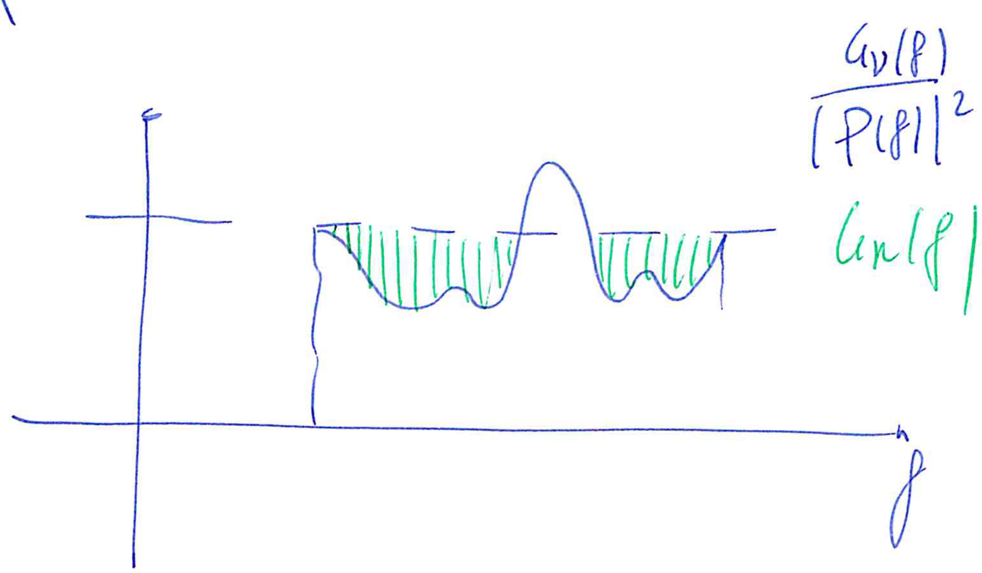
$$\frac{1}{1 + \frac{G_n(f)\,|P(f)|^2}{G_v(f)}}\cdot\frac{|P(f)|^2}{G_v(f)} - \lambda = 0$$

$$\frac{\cancel{G_v(f)}}{G_v(f) + G_n(f)\,|P(f)|^2}\cdot\frac{|P(f)|^2}{\cancel{G_v(f)}} = \lambda \qquad \lambda = \frac{|P(f)|^2}{G_v(f) + G_n(f)\,|P(f)|^2}$$

$$\frac{1}{\lambda} = \frac{G_V(f) + G_n(f)|P(f)|^2}{|P(f)|^2} = \frac{G_V(f)}{|P(f)|^2} + G_n(f)$$

$$\begin{cases} \dfrac{G_V(f)}{|P(f)|^2} + G_n(f) = \text{const.} : G_n(f) > 0 \\ 0 \quad \text{otherwise} \end{cases}$$

$$\begin{cases} G_n(f) = \text{const} - \dfrac{G_V(f)}{|P(f)|^2} \quad \text{if } G_n(f) > 0 \\ 0 \quad \text{otherwise} \end{cases}$$



- AWGN soft Decision

Mixture Distr. see notes.

ERROR CORRECTION CODES

S = avg. power in tx

B_r = bit rate [bit/sec]

B_{rc} = coded bit rate [bit coded / sec]

$$R_c = \frac{B_r}{B_{rc}} \qquad E_b = \frac{S}{B_r} \qquad E_{bc} = \frac{S}{B_{rc}}$$

The error probability with the BSC channel HD

$$p = \frac{1}{2}\,\mathrm{erfc}\sqrt{\frac{E_b}{N_0}}$$

$$p_c = \frac{1}{2}\,\mathrm{erfc}\sqrt{\frac{E_b R_c}{N_0}}$$

$\Rightarrow$ for the repetition (3,1) $p < p_c$.
It's not convenien to use!

- Weight of c_i = # of non zero elements of c_i
- Hamming distance = # of different elements between two codes c_i, c_j
- Linear code

if $\alpha c_i + \beta c_j = c_m$ where c_m is a codeword

the sum of two codewords is a codeword
A linear code has ever the $\underline{c} = \underline{0}$ element.

- For a binary code $d_H(\underline{c}_i, \underline{c}_j) = \underline{c}_i + \underline{c}_j = \underline{c}_j - \underline{c}_i$

- GENERATOR MATRIX (for linear codes).

Given $\underline{u}_m = \{u_{1m}, u_{2m}, u_{3m}, \dots, u_{mk}\}$ $\qquad \underline{c}_m = \{c_{1m}, c_{2m}, c_{3m}, \dots, c_{mn})$

$$c_m = \{g_{11}u_{1m}, g_{12}u_{2n}, \dots\ g_{1m}u_{mk}\}$$

It is possible to find a matrix $\underline{\underline{G}}$

That $\underline{c_m} = \underline{w_m}\,\underline{\underline{G}}$

$$\underline{\underline{G}} = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & & & \vdots \\ \vdots & & & \vdots \\ g_{k1} & & \cdots & g_{kn} \end{pmatrix} \qquad \underline{\underline{G}} \in M_{k \times n}$$

- Two codes are equivalents if it is possible to find the other matrix from the first one, using row and column operations.

- Systematic Codes.

$$\underline{\underline{G}} = \left[\underline{\underline{Id}} \mid \underline{\underline{P}}\right]$$

The $c_m = w_m \mid p_m$

- Codewords

Each linear combination of every $\underline{\underline{G}}$ rows is a codeword.

- Dual Code

Let take a orthogonal matrix $\underline{\underline{H}}$

$$\underline{\underline{G}}\,\underline{\underline{H}}^T = 0 \qquad \underline{\underline{H}} \text{ is orthogonal to every codeword.}$$

The code generated from $\underline{\underline{H}}$ is said Dual.

Taken a $\underline{\underline{G}}$ for a systematic code

$$\underline{\underline{G}} = \left[\underline{\underline{Id}} \mid \underline{\underline{P}}\right] \Longrightarrow \underline{\underline{H}}\left[-\underline{\underline{P}}^T \mid \underline{\underline{Id}}\right]$$

Proof

$$\underline{\underline{G}}\,\underline{\underline{H}}^T = \left[\underline{\underline{Id}} \mid \underline{\underline{P}}\right] \begin{bmatrix} -\underline{\underline{P}} \\ \underline{\underline{Id}} \end{bmatrix} = -\underline{\underline{P}} + \underline{\underline{P}} = \underline{\underline{0}}$$

- Hamming Code (7,4)

$$\underline{\underline{H}} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

In the columns there are all combinations of three bits except $\underline{0}$.

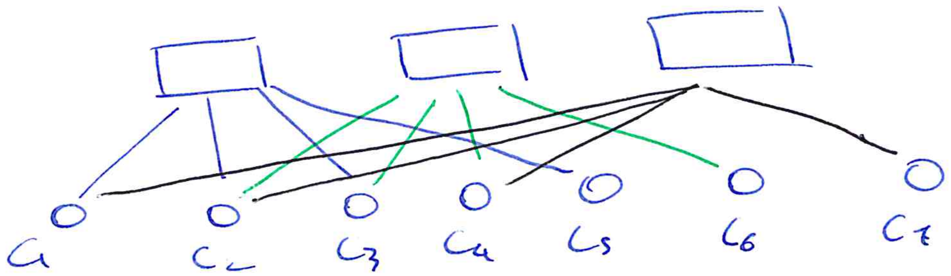
$$\Downarrow$$

$$\underline{\underline{G}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

- Parity Equations

$$\underline{\underline{H}}\,\underline{c}^T = \underline{0} \longrightarrow \begin{cases} c_1 + c_2 + c_3 + c_5 = 0 \\ c_2 + c_3 + c_4 + c_6 = 0 \\ c_1 + c_2 + c_4 + c_7 = 0 \end{cases}$$

This could be even represented with a <u>bipartite graph</u>



It is eq. to $\underline{\underline{H}}\,c^T = 0$

- Detection theory

$$\underline{c}_m \longrightarrow \boxed{\text{CHANN}} \xrightarrow{\ \underline{y}\ } \boxed{\text{DETECTOR}} \longrightarrow \hat{\underline{c}}_m$$

I want to obtain $\underline{c}_m$ from $\underline{y}$

$P_e = P_r\left\{ \hat{\underline{c}}_m \neq \underline{c}_m \right\}$

$P(\underline{c}_m)$ = A-PRIORI PROBABILITY

$P(\underline{c}_m | \underline{y})$ = A-POSTERIORI PROBABILITY

The perfect detection is done choosing the Maximum A Posteriori

- MAP

$$\hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\operatorname{argmax}} \left\{ P(\underline{c}_m | \underline{y}) \right\}$$

This could be even written as:

$$P(\underline{c}_m | \underline{y}) = \frac{P(\underline{y} | \underline{c}_m) P(\underline{c}_m)}{P(\underline{y})}$$

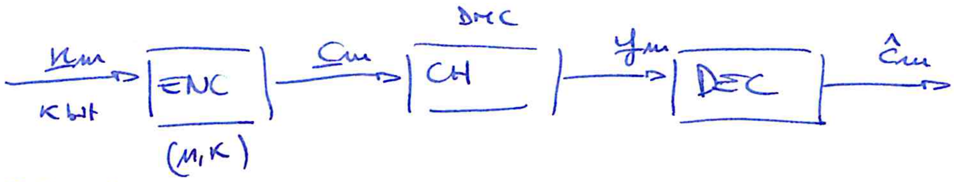
then

$$\hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\operatorname{argmax}} \left\{ \frac{P(\underline{y} | \underline{c}_m) P(\underline{c}_m)}{P(\underline{y})} \right\} = \underset{\underline{c}_m \in \mathcal{C}}{\operatorname{argmax}} \left\{ P(\underline{y} | \underline{c}_m) P(\underline{c}_m) \right\}$$

If $P(\underline{c}_m)$ is equal for each codeword, so the code has the same A-PRIORI PROB.

$$\text{MAP} \Leftrightarrow \text{ML} \Rightarrow \hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\operatorname{argmax}} \left\{ P(\underline{y} | \underline{c}_m) \right\}$$

OPTIMUM DECODING FOR BLOCK CODES OVER BSC



Hyp: $P(\underline{u}_m) = \frac{1}{2^k} \ \forall m \Rightarrow P(\underline{c}_m) = \frac{1}{2^k} \ \forall m$

It is possible to use the ML

$$\hat{c}_m = \arg\max_{c_m \in \mathcal{C}} P(\underline{y}_m|\underline{c}_m)$$

Since it is DMC

$$P(y_m|\underline{c}_m) = \prod_{i=1}^{n} P(y_{mi}|c_{mi})$$

es — — — — — — — — — es

$$P(y_{mi}|c_{mi}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_{mi}-c_{mi})^2}{2\sigma}}$$

es — — — —

Since we want a BSC, the prop. (es) would be

$$P(c_m = \hat{c}_m) = 1-p \qquad P(c_m \neq \hat{c}_m) = p$$

$$P(y_{mi}|c_{mi}) = (1-p)\left(\frac{p}{1-p}\right)^{d_H(y_{mi},c_{mi})}$$

$$\Rightarrow P(y_m|c_m) = (1-p)^n \left(\frac{p}{1-p}\right)^{\sum_{i=1}^{n} d_H(y_{mi},c_{mi})} \qquad 0 \le p \le 1$$

$$\hat{c}_m = \arg\min_{c \in \mathcal{C}} \left\{ d_H(y_m, c_m) \right\}$$

MINIMUM HAMMING DISTANCE DECODER

AWGN Channel : Soft Decision

$$\xrightarrow[k\text{ bit}]{u_m} \underset{(n,k)}{\boxed{\text{ENC}}} \xrightarrow{c_m} \underset{\text{DMC}}{\boxed{\text{CH}}} \xrightarrow{y_m} \boxed{\text{DEC}} \xrightarrow{\hat{c}_m}$$

$$P(\underline{u}_m) = \frac{1}{2^k} \Rightarrow P(\underline{c}_m) = \frac{1}{2^k}$$

ML $\quad \hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\arg\max}\ P(\underline{y}_m | \underline{c}_m)$

Memory-less $\Rightarrow \quad P(\underline{y}_m | \underline{c}_m) = \prod_{i=1}^{n} P(y_{mi} | c_{mi})$

$$P(y_{mi} | c_{mi}) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(y_{mi}-c_{mi})^2}{2\sigma^2}}$$

$$P(\underline{y}_m | \underline{c}_m) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n} e^{-\frac{1}{2\sigma^2} \cdot \sum (y_{mi} - c_i)^2}$$

$$\hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\arg\max}\ e^{-\sum_i^n (y_{mi} - c_{mi})^2}$$

$$\hat{\underline{c}}_m = \underset{\underline{c}_m \in \mathcal{C}}{\arg\min} \sum_i^n (y_{mi} - c_{mi})^2 = \underset{\underline{c}_m \in \mathcal{C}}{\arg\min}\ d_E^2(\underline{y}_m, \underline{c}_m) = \underset{\underline{c}_m \in \mathcal{C}}{\arg\min}\ d_E(\underline{y}_m$$

MINIMUM EUCLIDEAN DISTANCE DECODER

This could be even seen as:

$$\sum_i^n (y_{mi} - c_{mi})^2 = \underbrace{\sum_i^n (y_{mi})^2}_{\text{const.}} + \underbrace{\sum_i^n (c_{mi})^2}_{\text{code sym energy}} - 2\sum_i^n (y_{mi} c_{mi})$$

It is equal to minimise $\quad E_c - 2 \underbrace{\sum_i^n (y_{mi} c_{mi})}_{\text{correlation}}$

SYNDROME AND STANDARD ARRAY DECODING

$\underline{c}\underline{\underline{H}}^T = 0 \qquad \underline{y}\underline{\underline{H}}^T = (\underline{c}+\underline{e})\underline{\underline{H}}^T$

$\underline{e}$ = vector with 1 in the place where the error has occurred.

$(\underline{c}+\underline{e})\underline{\underline{H}}^T = \underbrace{\cancel{\underline{c}\underline{\underline{H}}^T}}_{0} + \underline{e}\underline{\underline{H}}^T = \underline{e}\underline{\underline{H}}^T = \underline{s}$ SYNDROME

Then we construct a table with a syndrome for each error pattern. It is kept only the minimum weight error word for each syndrome.

The correction is done by the sum of $\underline{e}$ to y.

The std. array impl. has a pattern of each error with each codeword (coset), so to decode it is only a look-up in the std. array.

$BER = \dfrac{\#B.ERROR}{\#BITS}$ (NUMBER OF WRONG BITS ON THE TOTAL)

$CER = \dfrac{\#CODE\ ERRORS}{\#CODES}$ (NUMBER OF WRONG WORDS OVER THE TOTAL)

- CYCLIC LINER BLOCK CODES

given $\underline{c} = (c_{n-1}, c_{n-2}, \dots, c_1, c_0)$ as a codeword

if the code is cyclic $\underline{c}^{(1)} = (c_{n-2}, c_{n-3}, \dots, c_1, c_0, c_{n-1})$ is also a codeword

They are interesting because they could be associated to a polynome

$$C(D) = c_{n-1}D^{n-1} + c_{n-2}D^{n-2} + \dots + c_1 D^1 + c_0$$

• $C^{(1)}(D) = D \cdot C(D) \bmod (D^n - 1)$

$$C(D) = c_{n-1}D^{n-1} + c_{n-2}D^{n-2} + \dots + c_1 D + c_0$$

$$DC(D) = c_{n-1}D^{n} + c_{n-2}D^{n-1} + \dots + c_1 D^2 + c_0 D$$

$$C^{(1)}(D) = c_{n-2}D^{n-1} + \dots + c_1 D^2 + c_0 D + c_{n-1}$$

$$DC(D) = c_{n-1}D^{n} + c_{n-2}D^{n-1} + \dots + c_1 D^2 + c_0 D + c_{n-1} - c_{n-1}$$

$$= c_{n-1}(D^n - 1) + C^{(1)}(D)$$

$$DC(D) = \underbrace{c_{n-1}}_{Q}\underbrace{(D^n - 1)}_{D} + \underbrace{C^{(1)}(D)}_{R} \leftarrow !$$

$$DC(D) = Q(D)D(D) + R(D)$$

generally

$$C^{(m)}(D) = D^m C(D) \bmod (D^n - 1)$$

• Minimum grade Codeword

$$g(D) = D^{n-k} + \dots + 1$$

• It is unique : if I sum a similar word, I could get something shorter ⇒ impossible.
• It has to have the 1 in the end, otherwise I can shift and make it shorter.

Shifting that word I can obtain another codeword

$$C^{(m)} = D^m g(D) \quad \text{if } m < k-1$$

For the others I can:

Multiply for a greater D^m $m>k-1$ and calculate the remainder, otherwise I can combine linearly two words obtained with $m<n-k$

$$\underset{gr\ n-1}{C(D)} = \underset{gr<k-1}{Q(D)}\ \underset{gr\ n-k}{g(D)}$$

a) If $\mathcal{C}$ is a cyclic code, its generator $g(D)$ divides D^n-1

$$C(D) = Q(D)\,g(D)$$
$$C^{(1)}(D) = \tilde{Q}(D)\,g(D)$$

$$C^{(1)}(D) = D\,C(D) + c_{n-1}(D^n-1)$$

$$\tilde{Q}(D)\,g(D) = D\,Q(D)\,g(D) + c_{n-1}(D^n-1)$$

$$\left(\tilde{Q}(D) - D\,Q(D)\right)g(D) = c_{n-1}(D^n-1)$$

with $c_{n-1}=1$

$\underline{D^n-1 = \bar{Q}(D)\,g(D)}$ so it divides D^n-1

b) If $g(D)$ divides D^n-1, it can generate a cyclic code

$$D^n-1 = \tilde{Q}(D)\,g(D) \qquad C(D) = Q(D)\,g(D)$$

$$C^{(1)}(D) = c_{n-1}(D^n-1) + C(D)\,D$$

$$C^{(1)}(D) = D\cdot Q(D)\,g(D) + \tilde{Q}(D)\,g(D)\,c_{n-1}$$

$$C^{(1)}(D) = g(D)\left(D\,Q(D) + \tilde{Q}(D)\,c_{n-1}\right) = g(D)\,\bar{Q}(D)$$

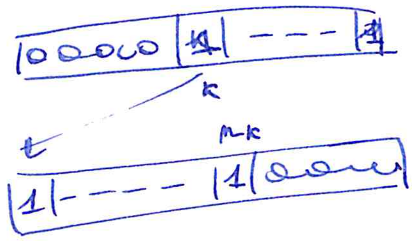
GENERATOR MATRIX

It is composed by every traslation of $g(D)$ with $m<k-1$ and these are put as the rows.

• Systematic cyclic codes

Note

$D^{n-k} x(D)$ is $x(D)$ translated ~~tra~~ of $n-k$ positions



$$D^{n-k} x(D) : g(D) = Q(D) \quad + R(D)$$

$$D^{n-k} x(D) = Q(D)\, g(D) + \underbrace{R(D)}_{k-1}$$

$$\underbrace{\left(\underbrace{D^{n-k} x(D)}_{gr\, n-1} + \underbrace{R(D)}_{gr\, k-1}\right)} = \underbrace{Q(D)\, g(D)}_{gr.\, n-1}$$

This is code!

And it is systematic!

To obtain systematic code

- Shift $n-k$ positions the input word
- Calculate the $D^{n-k} x(D) \bmod g(D)$
- Add this to the previous shifted code.

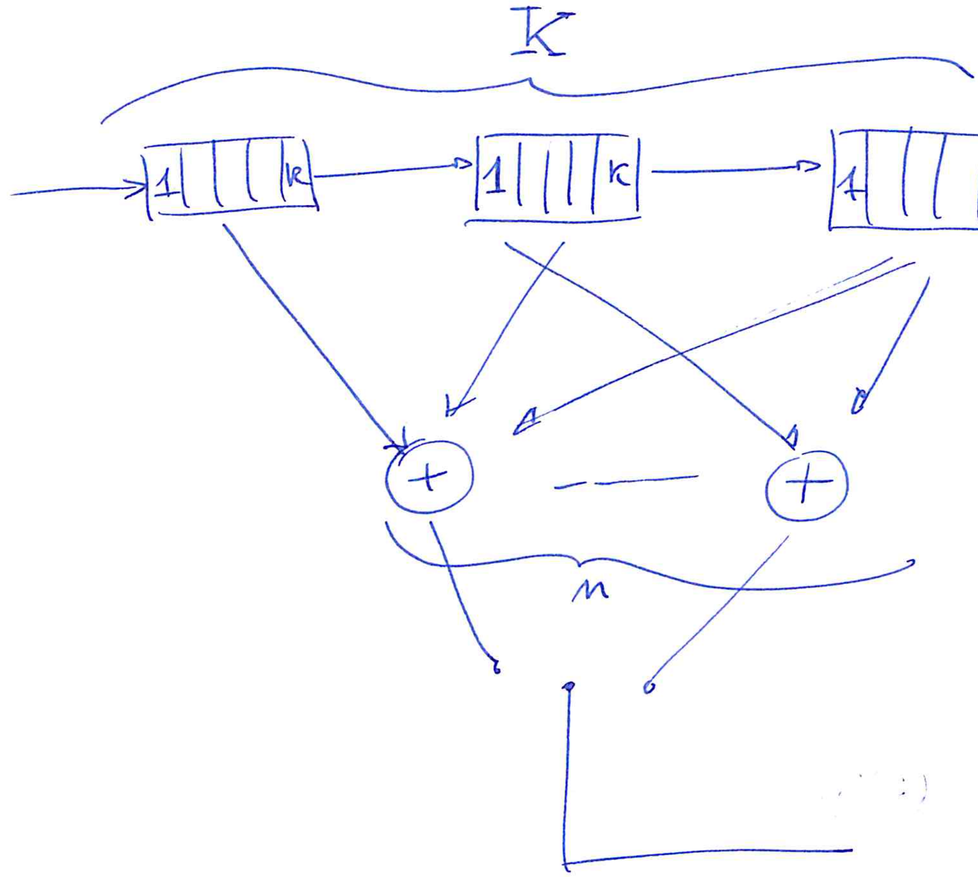
• Syndrome correction

$$r(D) = c(D) + e(D)$$

$$r(D) \bmod g(D) = \underline{\cancel{c(D) \bmod g(D)}} + e(D) \bmod$$

$$= e(D) \bmod g(D)$$

Convolutional codes:



It can be defined by a generator g, that could be a polynom or defined in octal.

It is 1 where there is a connection to K_{th} register.
K has to be 1.

These encoder/decoder has a memory, so it has a STATE.

It is possible to create a STATE DIAGRAM or the EQUIVALENT TRELLIS DIAGRAM. Each line in the TRELLIS is a state.
On the connection of the states there are usually the output:
if $R_c = 2$ two output and the line (different colors etc) should show the input bit.

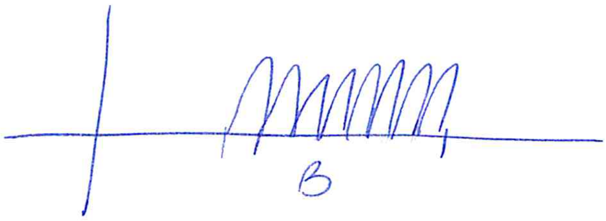
To decode:
- create the trellis and then calculate the lenght of every branch.
- Viterbi Algorithm.
  - Calculate the lenght of ~~every~~ branch to every state and keep only the shortest for each ~~branch~~ state (survivor)
  - Go ahead of one bit.

The original algorithm ends only when the trellis ends.
It is possible to have a really long code and to say that every bit that is 5-6 bit before the current node in the Trellis is right.

This is usually right.
(I mean 5-6 $\underline{K}$) Truncated V.A.
The distances could be calculated using d_H or d_E.

d_{FREE} is the min distance to return to the initial state. ~~after~~
The conv. codes makes burst errors, so it could be great to concatenate a Reed-Solomon or BCH to this one.

OFDM



We define a symbol as (as complex envelope)

$$a_0 = \sum_{i} a_i g(t - nT_s)$$

Then the OFDM shifts the freq.

$$i(t) = \sum_{n} a_n e^{j2\pi f_n t}$$

$f_n = \Delta f \cdot n$

It is orthogonal if

$$\frac{1}{T_s}\int_0^{T_s} e^{j2\pi f_n t} e^{-j2\pi f_l t} df = \frac{1}{T_s}\int_0^{T_s} e^{j2\pi f (n-l) t} df \Rightarrow \text{if } T_s = \frac{1}{\Delta f} \Rightarrow = \delta_{n-l}$$

If the freq. is shifted by $\frac{1}{T_s}$ $\Rightarrow$ the symbols are orthogonal.