EE 7620

Block Coding or block signalling

Given a binary source with rate $R = 1/T_b$ bits/sec.

given an AWGN channel with $S_{N_W}(f) = \frac{N_0}{2}$, and

given that power P and bandwidth W are avialable

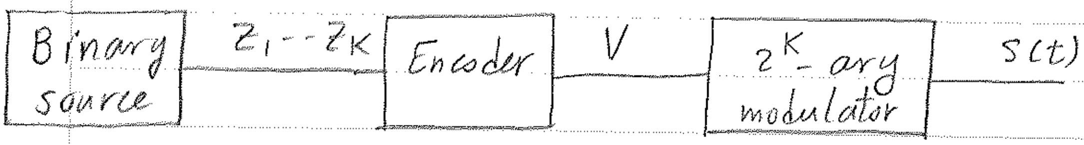
We want a block signaling scheme that gets low error probability and has a "reasonable" complexity.

In the block signaling approach we choose a block length K and 2^K signals $\{s_1(t), \ldots, s_{2^K}(t)\}$ each with duration $T = KT_b$ seconds,

energy $E_i = \|s_i(t)\|^2 \leq PT$, and

Bandwidth $\leq W$.

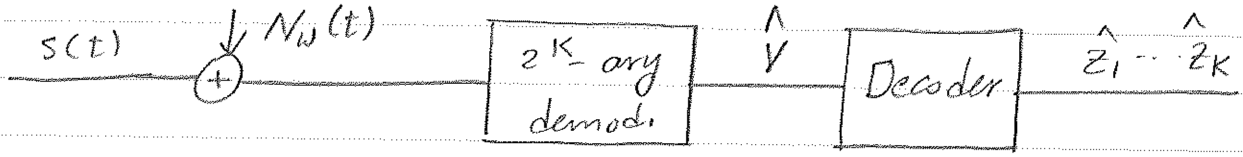
We also choose an encoder and a decoder and the system becomes



K to 1 encoder

$$\mathcal{S} = \{ s_1(t), \ldots, s_{2^K}(t) \}$$

$$\Omega_V = \{1, 2, \ldots, 2^K\}$$



Multilevel Block Signaling

In multilevel block signaling (or coding) we choose the 2^K signals in a special way so that implementation is simplified. There are two constraints:

(1) Basis functions: If a signal set with L dimensions is desired (in bandwidth W and time T), then we first choose a unit energy signal $\phi(t)$, with duration $T_L \triangleq \frac{T}{L}$ and define

$$\phi_j(t) = \phi(t - (j-1)T_L) \quad , \; j = 1, 2, \ldots, L.$$

clearly $\{\phi_1(t), \phi_2(t), \ldots \phi_L(t)\}$ is an orthonormal set of functions occupying the time interval $[0, T]$.

Notice that since $\phi_j(t)$ is time-limited to $T_L = \frac{T}{L}$ seconds, the bandwidths of the ϕ_j's is at least $\frac{1}{2T_L}$ if $\phi(t)$ is low pass and $\frac{1}{T_L}$ if $\phi(t)$ is bandpass. Since bandwidth is restricted to W, we see that we are limited to choosing

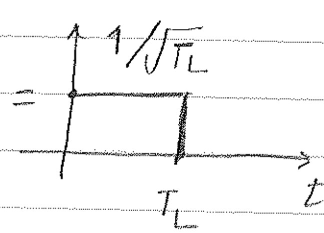
$$L \le 2WT = 2WKT_b = \frac{2W}{R}K$$

for $\phi(t)$ lowpass

and $L \le WT = \frac{W}{R}K$ if $\phi(t)$ is bandpass.

Also notice that the allowable L grows linearly with K.

Example 1: (Lowpass) Let $\phi(t) =$



Then each $\phi_j(t)$ has bandwidth $\frac{1}{T_L}$ and we can choose $L = \frac{T}{T_L} = WT = \frac{W}{R}K$ orthonormal functions

Example 2: (Bandpass) let $\phi(t) = \begin{cases} \sqrt{2/T_L}\, \cos 2\pi f_0 t, & 0 \le t \le T_L \\ 0 & , \text{elsewhere} \end{cases}$

Then each $\phi_j(t)$ has bandwidth $2/T_L$, and we can choose $L = \frac{T}{T_L} = \frac{WT}{2} = \frac{W}{2R} K$

(2) Signal vectors

We must choose 2^K L-dimensional signal vectors

$\mathcal{S} = \{ \underline{s}_1, \underline{s}_2, \dots, \underline{s}_{2^K} \}$, where

$$\underline{s}_i = (s_{i1}, s_{i2}, \dots, s_{iL})$$

To simplify implementation we choose a <u>finite</u> set $\Omega_S = \{ a_1, a_2, \dots, a_J \}$ of real numbers and <u>require</u> that $s_{ij} \in \Omega_S$ for all i, j.

The resulting signal set has signals of the form

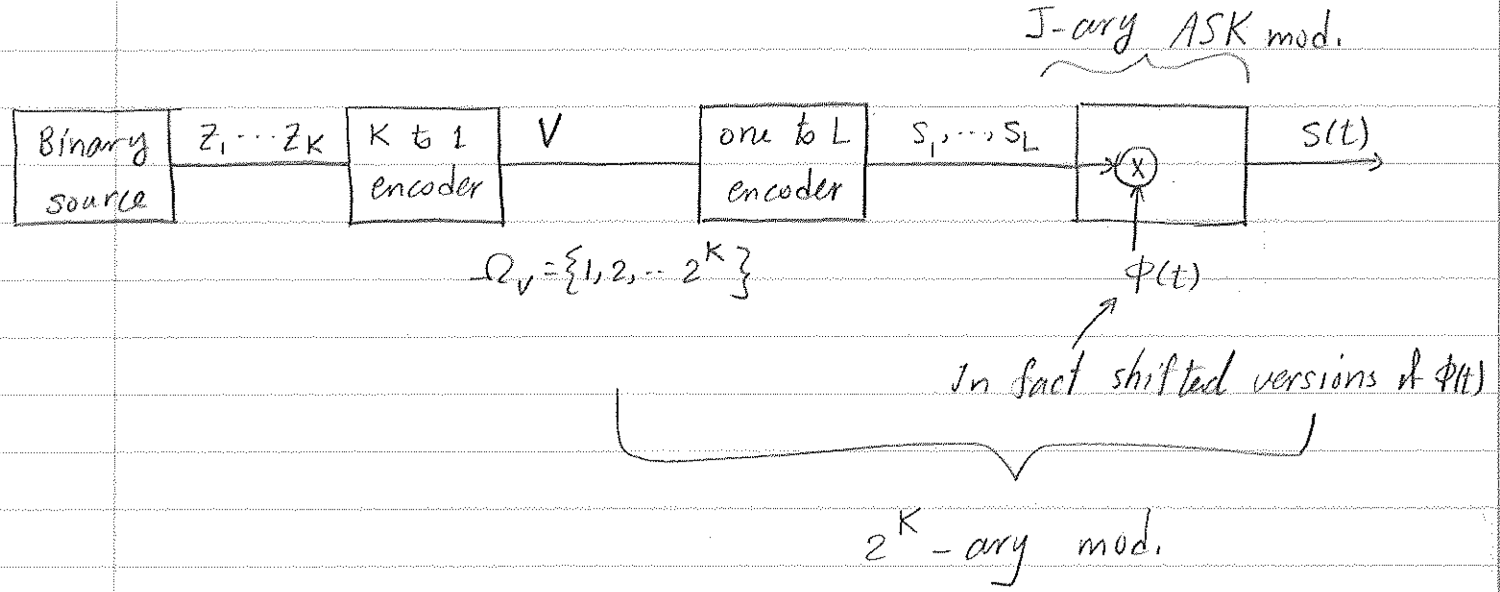
$$s_i(t) = \sum_{j=1}^{L} s_{ij}\, \phi_j(t) = \sum_{j=1}^{L} s_{ij}\, \phi(t-(j-1)T_L)$$

$$i = 1, 2, \dots 2^K$$

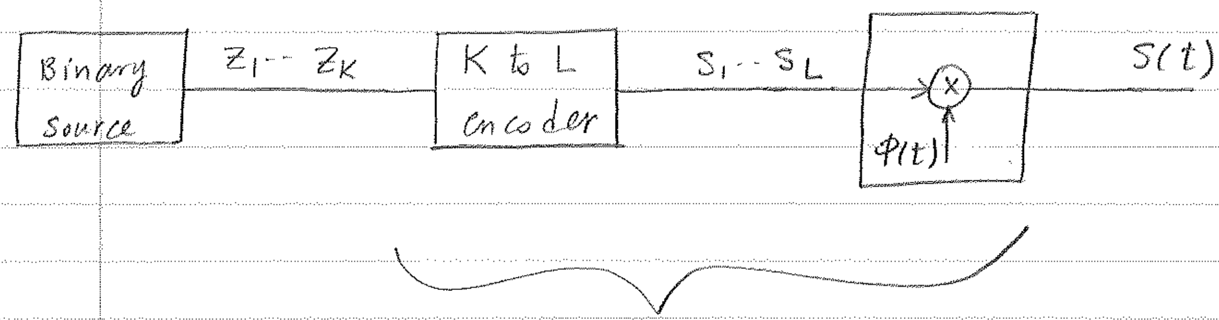
and they have the required duration T and bandwith W.

Notice that $s_i(t)$ appears to have been created by passing $s_{i1}, s_{i2}, \ldots s_{iL}$ through a J-ary modulator with signal set $\{a_1\phi(t), a_2\phi(t), \ldots, a_J\phi(t)\}$, i.e. a J-ary ASK.

This suggests the following implementation



Both the encoders shown above have discrete inputs and outputs. Hence they may be implemented by digital hardware and in fact they may be lumped together to get the following



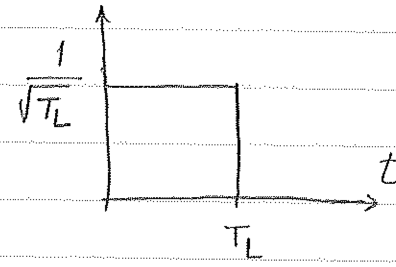
Multi-level block coded modulator

The K to L encoder assigns a distinct signal vector $\underline{s} = (s_1, \dots s_L) \in \Omega_S^L$ to each binary message sequence $(z_1 z_2 \cdots z_K) \in \{0,1\}^K$. This is called a <u>channel encoder</u>.

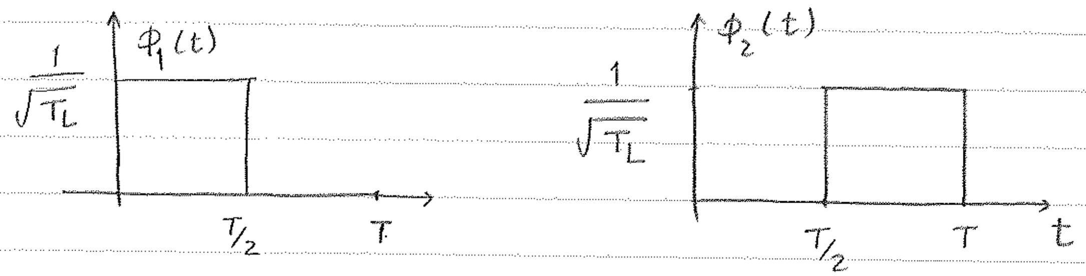
The signal vectors $\underline{s}_1, \underline{s}_2, \dots \underline{s}_{2^K}$ are commonly called <u>codewords</u>, their components are called <u>codeletters</u>, Ω_S is called the <u>code alphabet</u>, the collection of all codewords is called the <u>code book</u> or simply the <u>code</u>, K is called the <u>input block length</u>, L is called the <u>output block length</u>. K/L is defined to be the <u>code rate</u>.

(<u>Aside</u>: We could use a J-ary PSK or FSK modulator instead of the ASK modulator; in this case we would again have a block coded signalling scheme, but it would not be "multi level" block coding.)

Example 1: Let $K=3$, $L=2$, $\varphi(t) =$



Then $T = 2T_L = 3T_b$ and



Let $J=4$ and $\Omega_s = \{-3a, -a, a, 3a\}$, where a is a parameter.

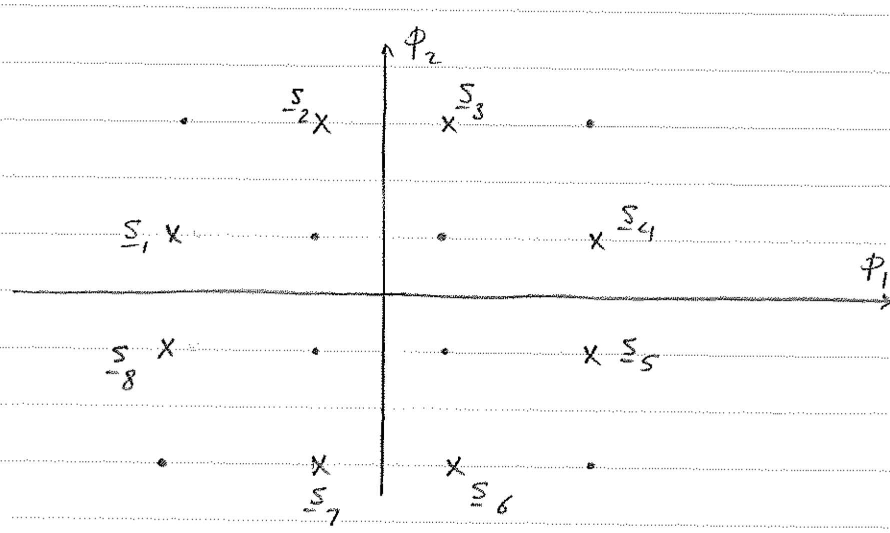
We need $2^K = 8$ codewords. One possible set of codewords is the following

$$\{\underline{s}_1, \underline{s}_2, \underline{s}_3, \underline{s}_4, \underline{s}_5, \underline{s}_6, \underline{s}_7, \underline{s}_8\} = \{(-3a, a), (-a, 3a), (a, 3a), (3a, a), (3a, -a), (a, -3a), (-a, -3a), (-3a, -a)\}$$

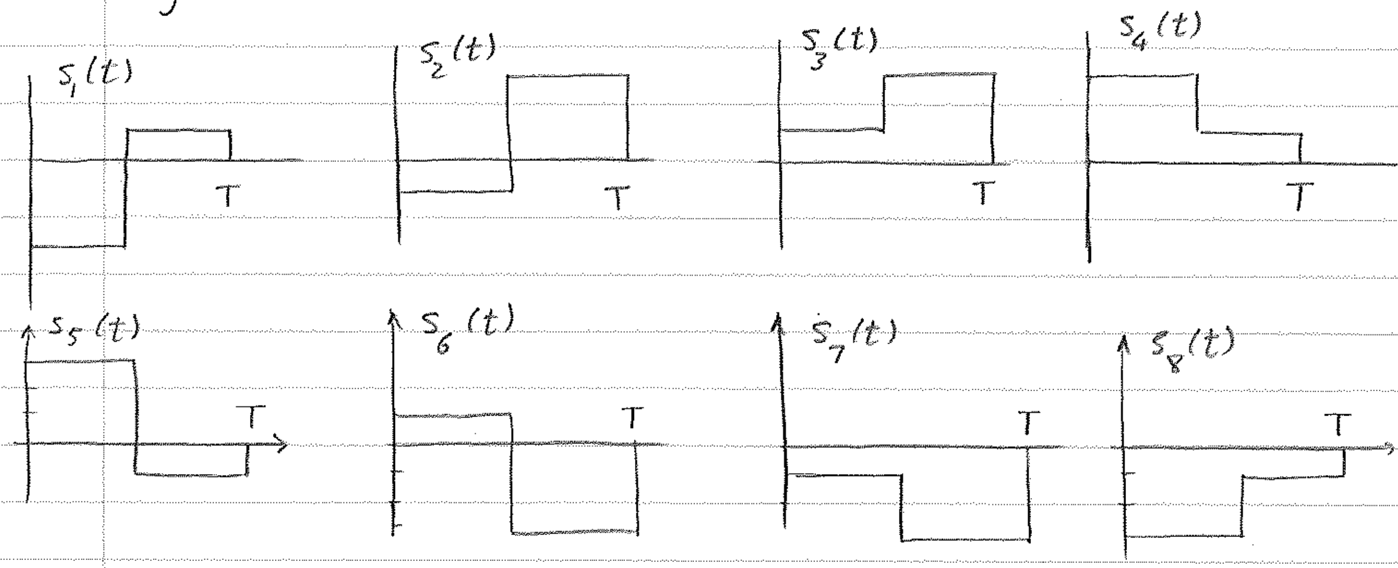
These are plotted below. Note that $\|\underline{s}_i\|^2 = 10a^2 = PT = 3PT_b$. Hence

$P = \dfrac{10a^2}{3T_b}$; parameter a can be used to adjust power.

other vectors with components in Ω_s which are not codewords are marked with •'s.



The signal set itself is



Each signal is made up of $L=2$ chips of duration $T_L = \frac{3}{2} T_b$.

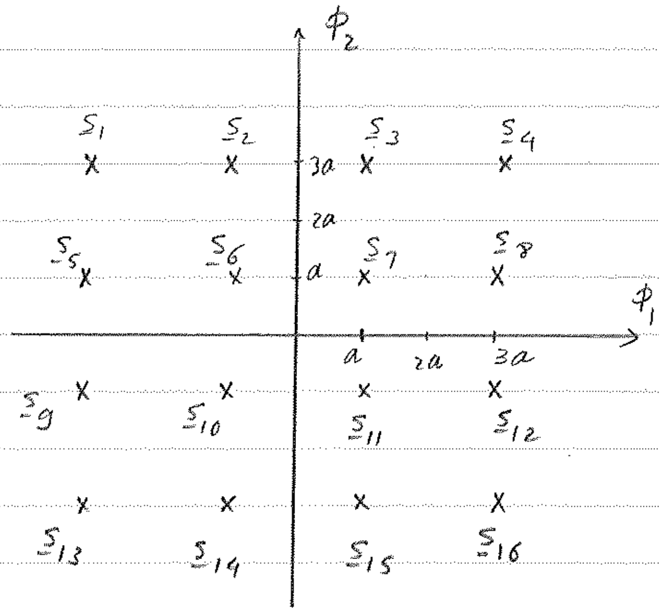
The bandwidth is approximately $W = \frac{1}{2T_L} = \frac{1}{3T_b} = R/3 \Rightarrow R/W = 3$.

Example 2

Let $K=4$, $L=2$, $J=4$, $\Omega_s = \{-3a, -a, a, 3a\}$

and $\phi(t)$ as in example 1.

We need $2^K = 16$ codewords. There are only 16 codewords with 2 components each in Ω_s. Hence the only possible code is



$$E_{ave} = \frac{1}{16} \sum_{i=1}^{16} \|\underline{s}_i\|^2$$

$$= 10\, a^2$$

$$P = \frac{E_{ave}}{T} = 10\, a^2 \frac{1}{4T_b} = \frac{10\, a^2}{4\, T_b}$$

Here the bandwidth is again $W = \frac{1}{2T_L}$. However since $T = 2T_L = KT_b = 4T_b$, $2T_L = 4/R$ and hence

$R/W = 4$.

Therefore in terms of bandwidth efficiency, example 2 is better than example 1.

After discussing the demodulator for multilevel block signalling, we will compare these two examples on the basis of $\frac{E_b}{N_0}$ required to achieve $P_b(\mathcal{E}) = 10^{-5}$.

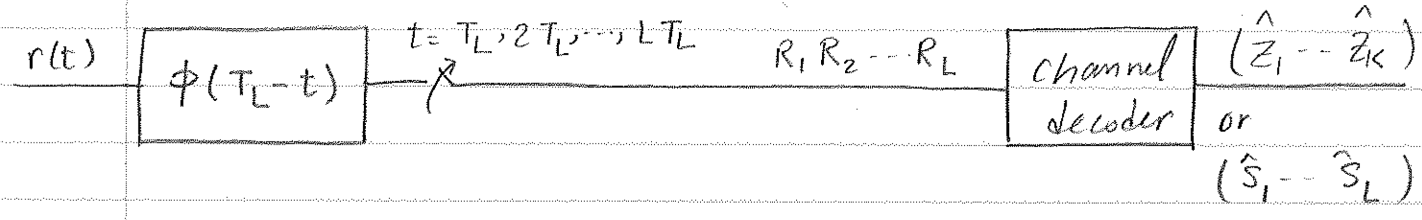
Aside; Notice that example 2 is a baseband 16 QASK system.

Demodulator

Like any optimum demodulator, the demodulator makes a decision about the transmitted sequence $(z_1, z_2 \cdots z_K)$ (or equivalently about the transmitted signal $s(t)$) based on $R(t)$, $0 \le t \le T$. The vector $\underline{R} = (R_1, \ldots, R_L)$ is a sufficient statistic, where $R_j = \int_0^T R(t)\,\phi_j(t)\,dt$. But since $\phi_j(t)$ is just a delayed version of $\phi(t)$, we can write

$$R_j = \int_0^T R(t)\,\phi(t-(j-1)T_L)\,dt$$

This suggests we can build a receiver using only one rather than L matched filters.



The <u>channel decoder</u> is the ordinary vector decision rule which, given $\underline{r}$, chooses $\hat{\underline{s}} = \underline{s}_i$ if

$\|\underline{r} - \underline{s}_i\|$ is smallest among all codewords $\underline{s}_j$, i.e., if $\underline{s}_i$ is the closest codeword to $\underline{r}$.

Error Probability

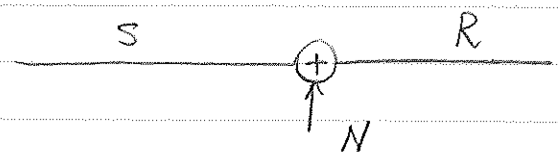
The block error probability

$$P(\mathcal{E}) = P_r(\hat{\underline{s}} \neq \underline{s}) = P_r(\hat{z}_1 \cdots \hat{z}_K \neq z_1 \cdots z_K)$$

may be computed in the usual way. However, this is likely to be very hard if there are many codewords. The union bounding technique may be used to obtain an upper bound to $P(\mathcal{E})$.

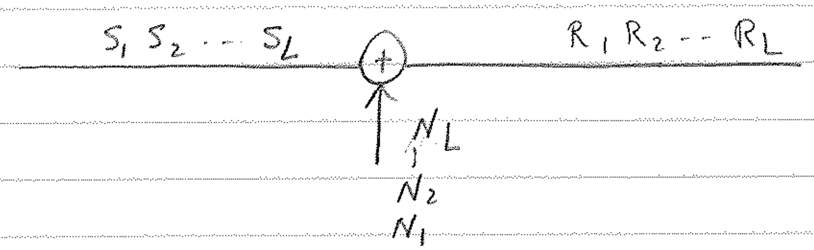
Aside:

We can model the combination of ASK modulator, AWGN channel, matched filter, and sampler as a discrete-time channel with additive Gaussian noise.



N is zero-mean Gaussian with variance $\frac{N_0}{2}$.

N is independent of S. This channel is used L times in the transmission of a codeword $(s_1, s_2, \ldots s_L)$, i.e. once every T_L seconds. The noises added in successive transmissions are independent



L uses of the channel.

Error probability for examples 1 and 2

Example 1.

Calculating error prob. for this signal set is very difficult. We use the union bound

$$P(\mathcal{E}) \simeq N_{min}\, Q\left(\frac{d_{min}}{\sqrt{2N_0}}\right) = 1.\ Q\left(\frac{d_{18}}{\sqrt{2N_0}}\right)$$

$$\Rightarrow\ P(\mathcal{E}) \simeq Q\left(\frac{2a}{\sqrt{2N_0}}\right) = Q\left(\sqrt{\frac{3}{5}\frac{PT_b}{N_0}}\right) = Q\left(\sqrt{\frac{3}{5}\frac{E_b}{N_0}}\right)$$

$$P(\mathcal{E}) = 10^{-5} \quad \Rightarrow \quad \frac{E_b}{N_0} = 14.8 \text{ dB}$$

Example 2:

For this signal set exact error probability can be calculated.

$$P(\mathcal{E}) = \frac{1}{16}\left\{4\,P(\mathcal{E}/1) + 8\,P(\mathcal{E}/2) + 4\,P(\mathcal{E}/6)\right\}$$

$$P(\mathcal{E}/1) = 2Q\left(\frac{a}{\sqrt{N_0/2}}\right) - Q^2\left(\frac{a}{\sqrt{N_0/2}}\right) \simeq 2Q\left(\frac{a}{\sqrt{N_0/2}}\right)$$

$$P(\mathcal{E}/2) = 3\,Q\left(\frac{a}{\sqrt{N_0/2}}\right) - 2Q^2\left(\frac{a}{\sqrt{N_0/2}}\right) \simeq 3Q\left(\frac{a}{\sqrt{N_0/2}}\right)$$

$$P(\mathcal{E}/6) = 4\,Q\left(\frac{a}{\sqrt{N_0/2}}\right) - 4Q^2\left(\frac{a}{\sqrt{N_0/2}}\right) \simeq 4Q\left(\frac{a}{\sqrt{N_0/2}}\right)$$

$$P(\mathcal{E}) \simeq 3\,Q\left(\frac{a}{\sqrt{N_0/2}}\right) = 3Q\left(\sqrt{\frac{4}{5}\frac{PT_b}{N_0}}\right) = 3Q\left(\sqrt{\frac{4}{5}\frac{E_b}{N_0}}\right)$$

$$P(\mathcal{E}) = 10^{-5} \Rightarrow E_b/N_0 = 14.22 \text{ dB}$$

We would like to emphasize that although not evident from the examples, the idea in multilevel block coding is to choose K and L both large (on the order of 10 to 200) in such a way that

$$2^K \ll J^L.$$

The best possible performance of multi-level block coding

First let us discuss the effects of the choice of J, Ω_S, $\underline{s}_1, \ldots \underline{s}_{2^K}$ assuming that P, R, K and L are fixed. Roughly speeking, we make $P(\mathcal{E})$ small by making the signals far apart from each other. The power P constraints them from being too far apart. If J is fixed, Ω_S should be chosen so that signals may be designed to be far apart but satisfy the power constraint. Making J large allows greater flexibility in the choice of Ω_S and the signals. Hence making J large allows us to design better signal sets. However, the complexity of implementation increases with J. We would like to be able to use $J=2$ if at all possible.

Now let's assume fixed P, W, and N_0. Shannon's positive channel coding theorem shows that it is possible to find block signal sets with the given P and W, and with 2^K signals each signal having duration $T = K/R$ (R is the source rate in bits/sec) such that if

$$R < C = W \log\left(1 + \frac{P}{WN_0}\right),$$

then arbitrarily small error probability $P(\mathcal{E})$ can be achieved. But these signal sets are not necessarily multilevel block coded signal sets.

Is it possible to get arbitrarily small $P(\mathcal{E})$ with J-level block coded signalling? It can be shown that the answer is yes, provided $R < C_J$, where C_J is called the <u>J-level</u> <u>channel</u> <u>capacity</u> in bits per second.

C_J depends on J, P, W and N_0. Moreover $C_J < C$, and C_J increases monotonically to C as $J \to \infty$. Equivalently for any given R/W, there is a minimum amount of E_b/N_0 that is necessary in order to be able to obtain J-level block coded signals with arbitrarily small $P(\mathcal{E})$.

The formula for C_2 is given below

$$C_2 = \frac{2P}{N_0 \ln 2} - 2W \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \log \cosh\left[\frac{P}{WN_0} + \sqrt{\frac{P}{WN_0}}\, x\right] dx$$

and it can be shown that

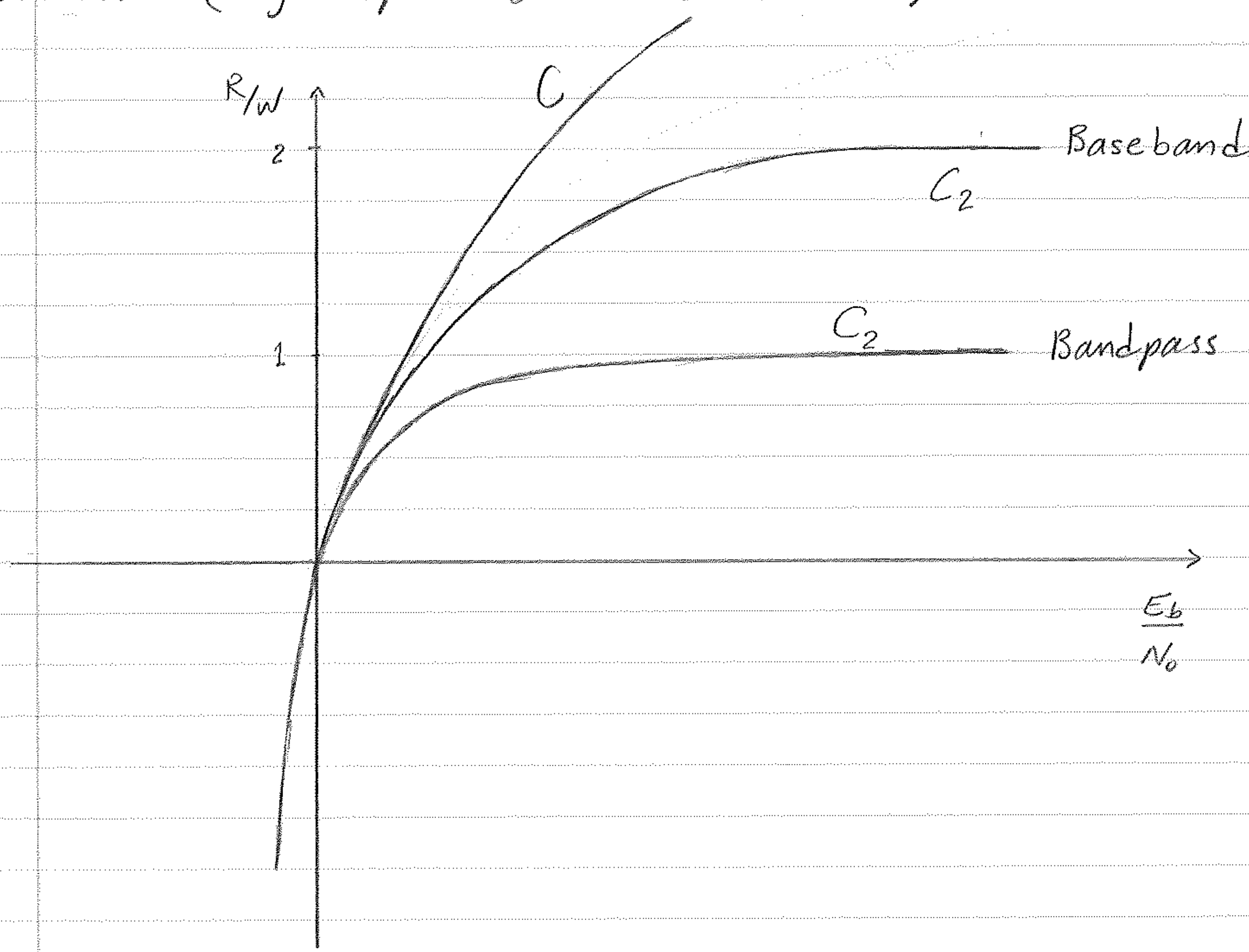
$$C_2 \longrightarrow \frac{P}{N_0 \ln 2} \quad \text{as } W \longrightarrow \infty$$

Letting $P = \frac{E_b}{T_b} = E_b R$ and $R < C_2$ we get

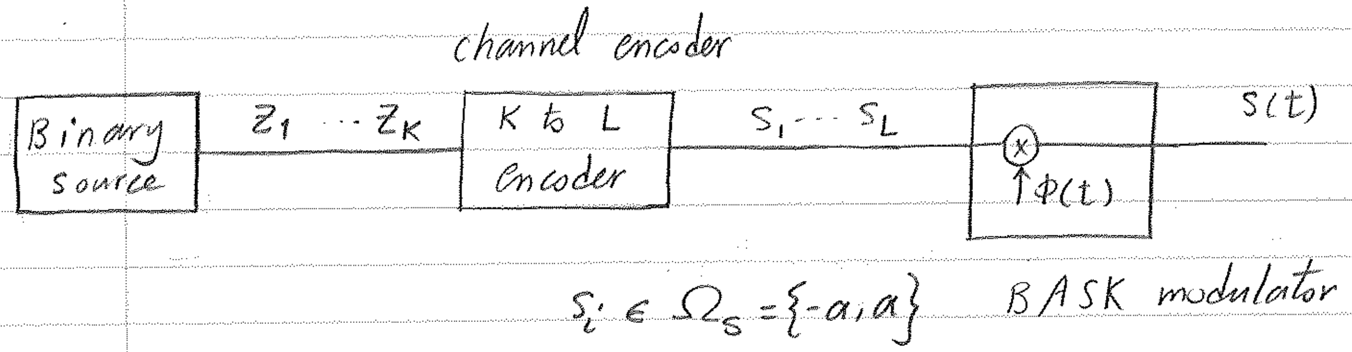
$$R/W < \frac{2}{\ln 2} \frac{E_b}{N_0} R/W - 2\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \log \cosh\left[\frac{E_b}{N_0} R/W + \sqrt{\frac{E_b}{N_0} R/W}\, x\right] dx$$

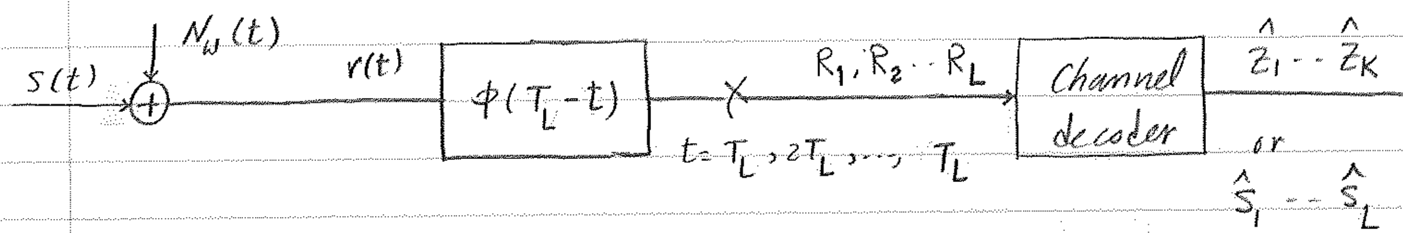
This together with capacity C are "approximately" plotted on the next page. From this figure we see that if $R/W < 1$, then the minimum required $\frac{E_b}{N_0}$ for binary coded signalling is negligibly greater than that required by the best of all possible signal sets. This can also be seen from the fact that both C and C_2 converge to $\frac{P}{N_0 \ln 2}$ as $W \to \infty$.

We conclude that the restriction to binary coded signalling does not limit potential performance if $R/W < 1$. Since binary coded signals are simplest to implement and since they are as good as any others when $R/W < 1$, they are used predominently in coded systems when $R/W < 1$, for example in channels where power is in short supply but bandwidth is plentiful (e.g. space communication link).



The following is the block diagram of a binary block coded signalling scheme.





Notice from the figure for C_2 that there are no binary coded systems having $R/W > 2$. In general for a J-level block coded system, $R/W \leq 2 \log_2 J$. This is because for J-level coded system with input and output blocklengths K and L we must have

the number of codewords $= 2^K \leq J^L =$ the number of J-level vectors

But we must also have $L < 2WT = 2W\frac{K}{R}$, which implies

$$K \leq L \log_2 J \leq 2W\frac{K}{R} \log_2 J$$

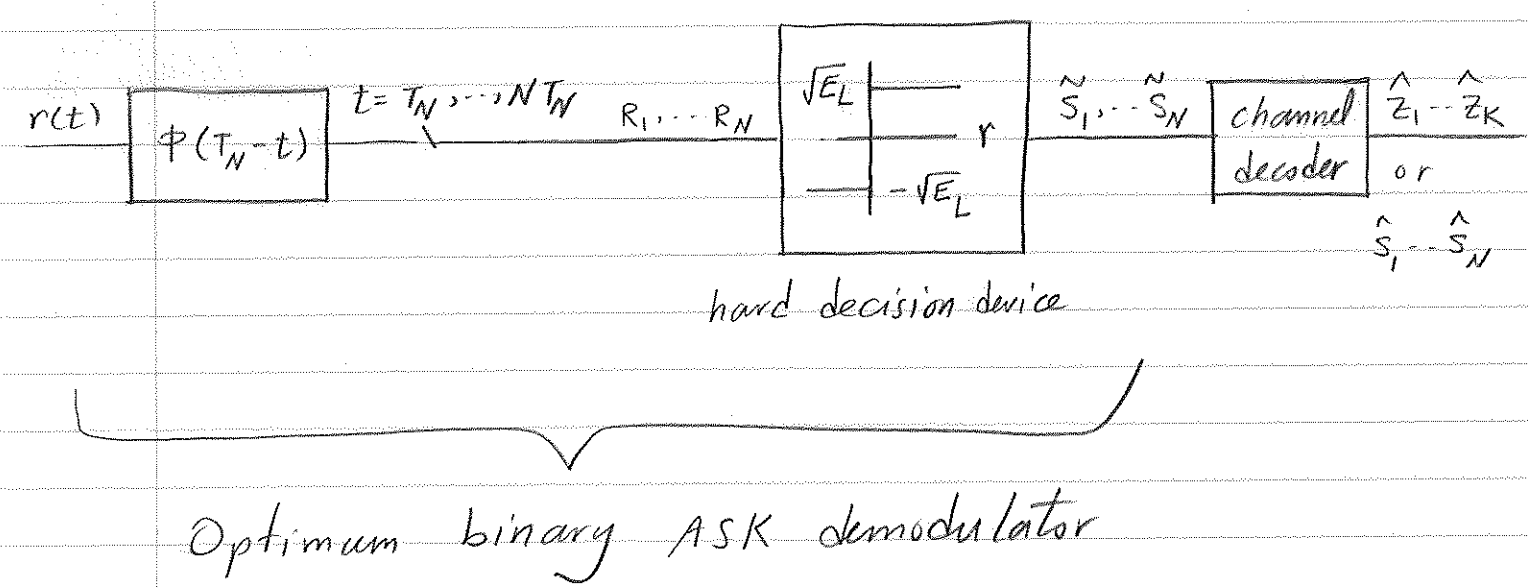
$$\Longrightarrow \quad R/W \le 2 \log_2 J$$

If the signals are to be bandpass, we must have

$$R/W \le \log_2 J$$

Hard decision demodulator

Assume $\Omega_S = \{-\sqrt{E_C}, \sqrt{E_C}\}$. The following is a hard decision receiver for the modulator described on the previous page



In a hard decision receiver, the receiver first makes a separate decision about each s_j, $j = 1, 2, \cdots, L$. $\tilde{s}_j$ is the best estimate of s_j based on $r(t)$, $(j-1)T_L \le t < jT_L$. Then the channel decoder makes the best possible decision about which codeword was transmitted based on $\tilde{s}_1 \cdots \tilde{s}_L$. (Note that $\tilde{s}_1 \cdots \tilde{s}_L$ is not necessarily a codeword so some sort of channel decoding is needed.) This receiver is definitely suboptimum, but it is simpler to implement than the optimum receiver because the decoding rule can be implemented as a digital machine because the inputs $\tilde{s}_j$ are binary.

Channel decoder

In order to minimize $P(\mathcal{E})$, the channel decoder should use the decoding rule

$g(\underline{\tilde{s}}) = \underline{s}_i$ if and only if

$p_{\underline{\tilde{S}} | \underline{S}}(\underline{\tilde{s}} | \underline{s}_i)$ is largest among all codewords.

Therefore we need to compute $p_{\tilde{\underline{S}}|\underline{S}}(\tilde{\underline{s}}/\underline{s}_i)$. Let us first find $p(\tilde{s}_j/s_j)$ for some j. The transformation from s_j to $\tilde{s}_j$ includes the effects of the binary modulator, AWGN channel, matched filter, sampler, and hard decision device. It may be modeled by a discrete-time channel called a <u>binary symmetric channel</u> (BSC).

If $S_j = +\sqrt{E_c}$, then $R_j = \sqrt{E_c} + N_j$

where $N_j = N_w(t) * \phi(T_L - t)\Big|_{t = jT_L}$

Hence

$$\tilde{S}_j = \begin{cases} +\sqrt{E_c} & \text{if } R_j \geq 0 \text{ or equivalently if } N_j \geq -\sqrt{E_c} \\ -\sqrt{E_c} & \text{if } R_j < 0 \text{ or equivalently if } N_j < -\sqrt{E_c} \end{cases}$$

Therefore

$$P_{\tilde{S}_j|S_j}(+\sqrt{E_c}/+\sqrt{E_c}) = P_r(N_j \geq -\sqrt{E_c}) = 1 - Q\left(\sqrt{\frac{2E_c}{N_0}}\right)$$

$$P_{\tilde{S}_j|S_j}(-\sqrt{E_c}/+\sqrt{E_c}) = P_r(N_j < -\sqrt{E_c}) = Q\left(\sqrt{\frac{2E_c}{N_0}}\right)$$

Let $\quad p \triangleq Q\left(\sqrt{\frac{2E_c}{N_0}}\right)$, then

$$p_{\tilde{S}_j|S_j}\left(+\sqrt{E_c}\,|+\sqrt{E_c}\right) = 1-p$$

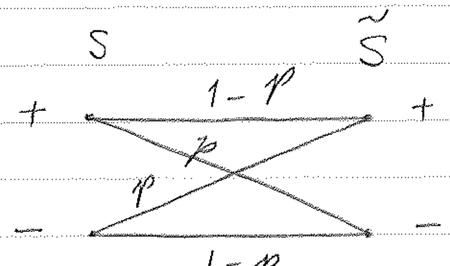
$$p_{\tilde{S}_j|S_j}\left(-\sqrt{E_c}\,|+\sqrt{E_c}\right) = p$$

Similarly if $\quad S_j = -\sqrt{E_c}$, $R_j = -\sqrt{E_c} + N_j$ and

$$p_{\tilde{S}_j|S_j}\left(-\sqrt{E_c}\,|-\sqrt{E_c}\right) = P_r(R_j \leq 0) = P_r(N_j < \sqrt{E_c}) = 1-p$$

$$p_{\tilde{S}_j|S_j}\left(+\sqrt{E_c}\,|-\sqrt{E_c}\right) = p$$

We can summarize these transition probabilities with a transition diagram for the BSC:



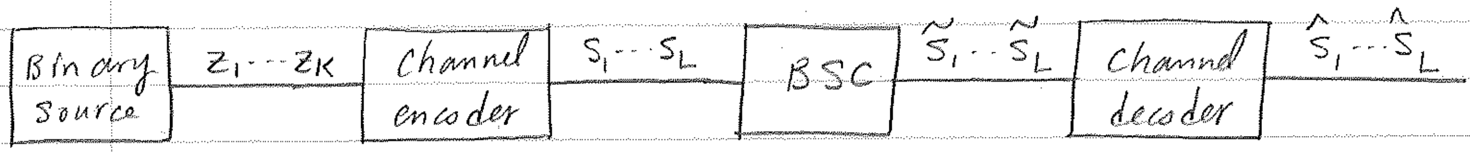
$+ = +\sqrt{E_c}$

$- = -\sqrt{E_c}$

$p = Q\left(\sqrt{\frac{2E_c}{N_0}}\right)$

The diagram indicates that the possible inputs and outputs are $\pm\sqrt{E_c}$. A line is drawn between every input and output and labelled with the conditional probability of the corresponding transition.

Our system with hard decision receiver now looks like:



$$p = Q\left(\sqrt{\frac{2E_c}{N_0}}\right)$$

Since the N_j's affecting $\tilde{s}_j$'s are independent and since $\tilde{s}_j$ is a function only of s_j and N_j, it follows that $\tilde{s}_j$ is conditionally independent of all s_k, $\tilde{s}_k$, $k \neq j$, given s_j.

This implies

$$P_{\underline{\tilde{s}}/\underline{s}}(\underline{\tilde{s}}/\underline{s}) = p(\tilde{s}_1/s_1)\, p(\tilde{s}_2/s_2) \cdots p(\tilde{s}_N/s_N)$$

Each term $p(\tilde{s}_j/s_j)$ is either p or $1-p$.

Let $d_H(\underline{\tilde{s}}, \underline{s}) \triangleq$ number of places that $\underline{s}$ and $\underline{\tilde{s}}$ differ

$d_H(\underline{\tilde{s}}, \underline{s})$ is called the Hamming distance between $\underline{\tilde{s}}$ and $\underline{s}$.

We have $p(\tilde{\underline{s}} \mid \underline{s}) = p^{d_H(\tilde{\underline{s}}, \underline{s})} (1-p)^{L - d_H(\tilde{\underline{s}}, \underline{s})}$

The best channel decoder is now:

$g(\tilde{\underline{s}}) = \underline{s}_i$ iff $p^{d_H(\tilde{\underline{s}}, \underline{s}_i)} (1-p)^{L - d_H(\tilde{\underline{s}}, \underline{s}_i)}$ is largest

iff $\left(\frac{p}{1-p}\right)^{d_H(\tilde{\underline{s}}, \underline{s}_i)} (1-p)^L$ is largest

$g(\tilde{\underline{s}}) = \underline{s}_i$ iff $d_H(\tilde{\underline{s}}, \underline{s}_i)$ is smallest among all codewords.

(Note: We have used the fact that $p < 1/2$ to imply that $\frac{p}{1-p} < 1$) The decoding rule is the minimum <u>Hamming</u> distance rule. It is the intuitive rule.

Let us agree to say that if $\tilde{s}_j \neq s_j$ then a BSC channel error has occurred. A BSC channel error does not necessarily cause a decoding error. Indeed it is common to say that the decoder has "corrected" a BSC channel error if no decoding error occurs. Note that if $\underline{s}_i$ is transmitted, then $d_H(\underline{s}_i, \tilde{\underline{s}})$ = number of BSC channel errors that have occurred.

Calculation of $P(\mathcal{E})$ for hard decision receiver

Method 1: Union upper bound to $P(\mathcal{E})$

$$P(\mathcal{E}) = \frac{1}{2^K} \sum_{i=1}^{2^K} P(\mathcal{E}/\underline{s}_i)$$

$$P(\mathcal{E}/\underline{s}_i) = P_r\left(d_H(\tilde{\underline{s}}, \underline{s}_k) \leq d_H(\tilde{\underline{s}}, \underline{s}_i) \text{ for som } k \neq i / \underline{s}_i \right)$$

$$\leq \sum_{k \neq i} P_r\left(d_H(\tilde{\underline{s}}, \underline{s}_k) \leq d_H(\tilde{\underline{s}}, \underline{s}_i) / \underline{s}_i \right)$$

$$= \sum_{k \neq i} P_r\left(d_H(\tilde{\underline{s}}, \underline{s}_i) \geq \frac{d_H(\underline{s}_i, \underline{s}_k)}{2} / \underline{s}_i \right)$$

$$= \sum_{k \neq i} \sum_{l \geq \frac{d_H(\underline{s}_i, \underline{s}_k)}{2}} \binom{N}{l} p^l (1-p)^{L-l}$$

If p is small, then when we upper bound each $P(\mathcal{E}|\underline{s}_i)$, the upper bound is dominated by the terms having smallest l.

Let
$$d_{min} \triangleq \min_{i \neq k} d_H(\underline{s}_k, \underline{s}_i)$$

Then

$$P(\mathcal{E}) \stackrel{\sim}{\leq} \frac{N_1}{2^K} \binom{N}{\frac{d_{min}}{2}} p^{\frac{d_{min}}{2}} (1-p)^{L - \frac{d_{min}}{2}}$$

where N_1 = # of terms for which $d_H(\underline{s}_i, \underline{s}_k) = d_{min}$

We conclude that $P(\mathcal{E})$ depends critically on d_{min}. d_{min} is an important measure of code performance large d_{min} causes low $P(\mathcal{E})$. Furthermore let

$$t \triangleq \left\lfloor \frac{d_{min}-1}{2} \right\rfloor \quad \text{where } \lfloor x \rfloor = \text{integer part of } x.$$

A code will correct any sequence of t or fewer BSC channel errors, i.e., if t or fewer BSC channel errors occur, the decoder will still make the correct decision.

Proof: If $\underline{s}_i$ is transmitted and t or fewer errors occur, then $d_H(\underline{s}_i, \tilde{\underline{s}}) \leq t$. But for any $\underline{s}_k$, $d(\underline{s}_i, \underline{s}_k) \geq d_{min} > 2t$. This implies that

$$d_H(\tilde{\underline{s}}, \underline{s}_k) > t \geq d_H(\underline{s}_i, \tilde{\underline{s}})$$

(since $d_H(\underline{s}_i, \underline{s}_k) \leq d_H(\underline{s}_i, \tilde{\underline{s}}) + d_H(\tilde{\underline{s}}, \underline{s}_k)$ called triangular inequality. This implies that

$$d_H(\tilde{\underline{s}}, \underline{s}_k) \geq d_H(\underline{s}_i, \underline{s}_k) - d_H(\underline{s}_i, \tilde{\underline{s}}) > 2t - t = t.)$$

Therefore, $\tilde{\underline{s}}$ is closer to $\underline{s}_i$ than any $\underline{s}_k$, $k \neq i$'s, and a correct decision will be made

Method 2:

$$P(\mathcal{E}) \leq \Pr(\text{BSC makes } t+1 \text{ or more errors})$$

$$= \sum_{e=t+1}^{L} \binom{L}{e} p^e (1-p)^{L-e} \qquad t \triangleq \left\lfloor \frac{d_{min}-1}{2} \right\rfloor$$

$$\simeq \binom{L}{t+1} p^{t+1} \underbrace{(1-p)^{L-t-1}}_{\simeq 1} \qquad \text{if } p \text{ small}$$

$$\tilde{\leq} \binom{L}{t+1} p^{t+1}$$

Best possible performance with binary coding and hard-decision receiver

It is possible to find a binary coded signal set with arbitrarily small $P(\mathcal{E})$ when a hard decision receiver is used, provided $R < C_2^h = 2W\left[1 - \mathcal{H}(p)\right]$ bits/sec.

where $p = Q\left(\sqrt{\frac{P}{WN_0}}\right)$ and (baseband)

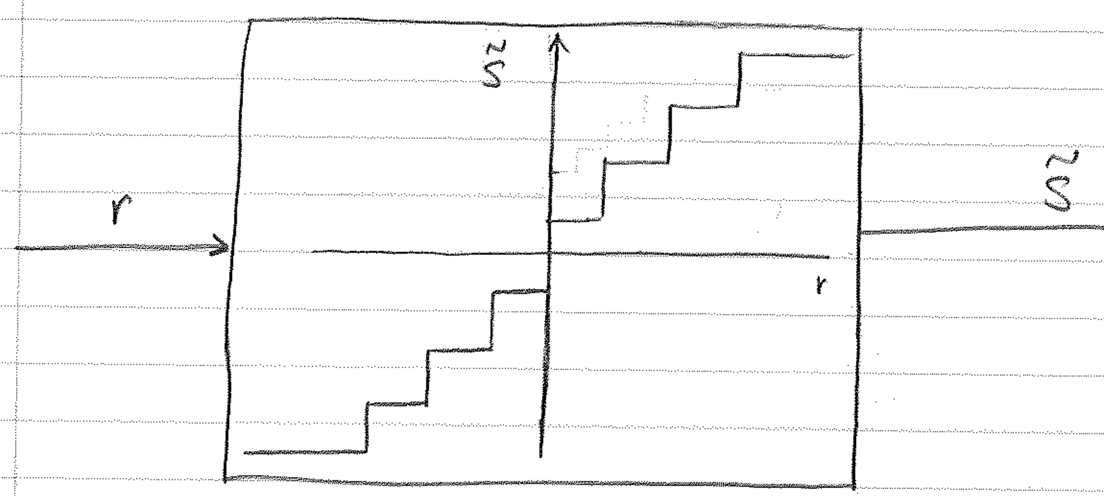
$$\mathcal{H}(p) \triangleq -p \log_2 p - (1-p)\log(1-p).$$

C_2^h is called the capacity of the binary hard-decision channel. $C_2^h < C_2 < C$. C_2^h is significantly less than C_2. In the Figure we have plotted $\frac{E_b}{N_0}$ vs. R/W needed for arbitrarily small $P(\mathcal{E})$ with binary hard-decision receiver. This is given by

$$R/W < 2\left[1 - \mathcal{H}\left(Q\left(\sqrt{\frac{E_b}{N_0} R/W}\right)\right)\right]$$

From the Figure we see that if $R/W < 1/2$, then $\frac{E_b}{N_0}$ must be 2 dB greater with hard decision than without. Hard-decision loses 2 dB.

It is possible to find a compromise between the optimum receiver and hard-decision receiver. If the hard-decision device is replaced by the following "soft-decision" device, then much of the 2 dB can be regained with an increase in decoder complexity.



8-level soft-decision device

This device is simply a quantizer. Typically 8 to 16 levels suffice to recover most of the lost 2 dB.