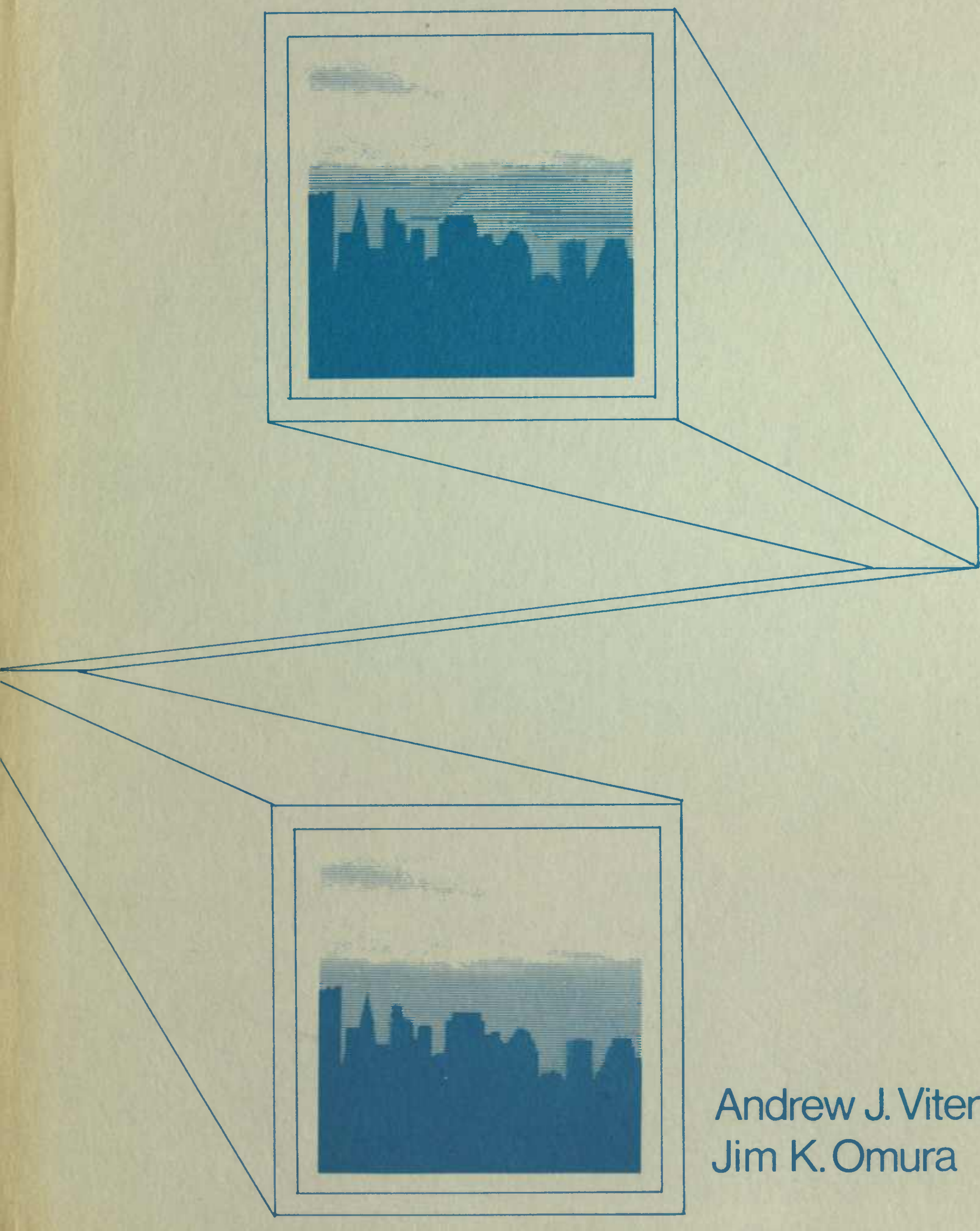
Solutions Manual to accompany

Principles of Digital Communication and Coding



Andrew J. Viterbi
Jim K. Omura

Solutions Manual to accompany

Principles of Digital Communication and Coding

Andrew J. Viterbi
LINKABIT Corporation

Jim K. Omura
University of California
Los Angeles

McGraw-Hill Book Company
New York St. Louis San Francisco Auckland Bogotá Düsseldorf
Johannesburg London Madrid Mexico Montreal New Delhi Panama
Paris São Paulo Singapore Sydney Tokyo Toronto

Solutions Manual to accompany
PRINCIPLES OF DIGITAL COMMUNICATION
AND CODING

0-07-067517-1

1 2 3 4 5 6 7 8 9 0 WHWH 7 8 3 2 1 0 9

ACKNOWLEDGEMENT

Graduate students Dariush Divsalar, Enyltho Coelho Filho, and Pil Lee have made significant contributions to this solution manual.

CHAPTER 1

1.1 (a) $\mathscr{H}(p) = -p\ln p - (1-p)\ln(1-p)$ nats.

$$\frac{d}{dp}\mathscr{H}(p) = \ln\frac{1-p}{p} = 0 \Longrightarrow p = \frac{1}{2} \text{ and } \frac{d^2}{dp^2}\mathscr{H}(p) = \frac{1}{p(1-p)} < 0.$$

(b)

$\mathscr{U}_2$	P
a_1a_1	p^2
a_1a_2	$p(1-p)$
a_2a_1	$p(1-p)$
a_2a_2	$(1-p)^2$

$$H(\mathscr{U}_2) = p^2\log\frac{1}{p^2} + p(1-p)\log\frac{1}{p(1-p)}$$
$$+ p(1-p)\log\frac{1}{p(1-p)} + (1-p)^2\log\frac{1}{(1-p)^2}$$
$$= 2\left[p\log\frac{1}{p} + (1-p)\log\frac{1}{1-p}\right]$$
$$= 2\mathscr{H}(p)$$

(c) i) $H(\mathscr{U}) = \log 52 = 5.7$ bits

ii) $H(\mathscr{U}) = \log 13 = 3.7$ bits

iii) $H(\mathscr{U}) = \frac{3}{13}\log\frac{13}{3} + \frac{10}{13}\log\frac{13}{10} = .77$ bits

(d) i) $H(\text{fair die}) = \log 6 = 2.58$ bits

ii) $P(k) = kC \qquad k = 1,2,3,4,5,6$

$$\sum_{k=1}^{6} P(k) = C\sum_{k=1}^{6} k = 1 \Longrightarrow C = \frac{1}{21}$$

$$\Longrightarrow H(\mathscr{U}) = \sum_{k=1}^{6}\frac{k}{21}\log\frac{21}{k} = 2.39 \text{ bits}$$

1.2 Inequality (1.1.8) gives

$$H(\mathscr{U}^{(1)},\ldots,\mathscr{U}^{(N)}) = \sum_{\underset{\sim}{u}} P_N(\underset{\sim}{u})\log\frac{1}{P_N(\underset{\sim}{u})}$$
$$\le \sum_{\underset{\sim}{u}} P_N(\underset{\sim}{u})\log\frac{1}{Q_N(\underset{\sim}{u})} \cdot \quad (= \text{ iff } Q_N(\cdot) = P_N(\cdot))$$

Choose $Q_N(\underset{\sim}{u}) = \prod_{n=1}^{N} P^{(n)}(u_n)$ where

$$P^{(n)}(u_n) = \sum_{i \neq n} \sum_{u_i} P_N(\underset{\sim}{u}) \qquad n = 1,2,\ldots,N.$$

are the marginal probability distributions. Then

$$H\left(\mathcal{U}^{(1)},\ldots,\mathcal{U}^{(N)}\right) \leq \sum_{\underset{\sim}{u}} P_N(u)\left(\sum_{n=1}^{N} \log \frac{1}{P^{(n)}(u_n)}\right)$$

$$= \sum_{n=1}^{N}\left(\sum_{\underset{\sim}{u}} P_N(\underset{\sim}{u}) \log \frac{1}{P^{(n)}(u_n)}\right)$$

$$= \sum_{n=1}^{N}\left(\sum_{u} P^{(n)}(u) \log \frac{1}{P^{(n)}(u)}\right)$$

$$= \sum_{n=1}^{N} H(\mathcal{U}^{(n)}).$$

with equality iff $P_N(\underset{\sim}{u}) = \prod_{n=1}^{N} P^{(n)}(u_n)$

1.3 (a) Given discrete random variables $x,y \in \mathcal{X} \times \mathcal{Y}$ with joint probability $P(x,y)$, we can define marginal probabilities $P(x) = \sum_y P(x,y)$, $P(y) = \sum_x P(x,y)$ and conditional probability $P(y/x) = P(x,y)/P(x)$. Then using inequality (1.1.8),

$$H(\mathcal{Y}\,|\,\mathcal{X}) \equiv \sum_x \sum_y P(x,y) \log \frac{1}{P(y|x)}$$

$$= \sum_x P(x)\left(\sum_y P(y|x) \log \frac{1}{P(y|x)}\right)$$

$$\leq \sum_x P(x)\left(\sum_y P(y|x) \log \frac{1}{P(y)}\right)$$

$$= \sum_y \left(\sum_x P(x)P(y|x)\right) \log \frac{1}{P(y)}$$

$$= H(\mathcal{Y}). \tag{1}$$

Now fix N and consider sequences $\underset{\sim}{u} \varepsilon \mathcal{U}_N = \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(N)}$ where $\mathcal{U}^{(k)} = \mathcal{U}$ is the alphabet of the k^{th} term in the sequence. Using the relations

$$P_N(u_1,\ldots,u_N) = P(u_N \mid u_1,\ldots,u_{N-1})\, P_{N-1}(u_1,\ldots,u_{N-1})$$
$$= \prod_{n=2}^{N} P(u_n \mid u_1,\ldots,u_{n-1}) \cdot P(u_1)$$

we have

$$H(\mathcal{U}_N) = H(\mathcal{U}_{N-1}) + H\left(\mathcal{U}^{(N)} \middle| \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(N-1)}\right) \tag{2}$$

and

$$H(\mathcal{U}_N) = H\left(\mathcal{U}^{(1)}\right) + \sum_{n=2}^{N} H\left(\mathcal{U}^{(n)} \middle| \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(n-1)}\right) \tag{3}$$

Note that from (1) we have

$$H\left(\mathcal{U}^{n)} \middle| \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(n-1)}\right) \leq H\left(\mathcal{U}^{(n)} \middle| \mathcal{U}^{(2)} x \ldots x \mathcal{U}^{(n-1)}\right)$$
$$= H\left(\mathcal{U}^{(n-1)} \middle| \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(n-2)}\right) \tag{4}$$

where the second equality comes from the stationary property. Hence

(3) is bounded by

$$H(\mathcal{U}_N) \geq NH\left(\mathcal{U}^{(N)} \middle| \mathcal{U}^{(1)} x \ldots x \mathcal{U}^{(N-1)}\right) \tag{5}$$

Using (5) in (2) gives

$$H(\mathcal{U}_N) \leq H(\mathcal{U}_{N-1}) + \frac{1}{N} H(\mathcal{U}_N).$$

or

$$\frac{1}{N} H(\mathcal{U}_N) \leq \frac{1}{N-1} H(\mathcal{U}_{N-1}) \;. \tag{6}$$

Then

$$\frac{H(\mathcal{U}_n)}{n} \leq \frac{H(\mathcal{U}_k)}{k} \quad \text{for } k \leq n.$$

(b) This follows directly from the proof of Theorem 1.1.1

when we define

$$S(N,\varepsilon) = \left\{ \underset{\sim}{u} : 2^{-N[\hat{H}+\varepsilon]} \leq P_N(\underset{\sim}{u}_n) \leq 2^{-N[\hat{H}-\varepsilon]} \right\}.$$

and use $\lim_{N\to 0} F_N = 0$ from (1.1.31) to (1.1.33).

1.4 (a) $\sigma^2 = E\{(x-m)^2\} = E\{(x-m)^2 \mid |x-m| \geq \varepsilon\} \Pr\{|x-m| \geq \varepsilon\}$

$+ E\{(x-m)^2 \mid |x-m| < \varepsilon\} \Pr\{|x-m| < \varepsilon\}$

$\geq E\{(x-m)^2 \mid |x-m| \geq \varepsilon\} \Pr\{|x-m| \geq \varepsilon\} \geq \varepsilon^2 \Pr\{|x-m| \geq \varepsilon\}.$

(b) $$\Pr\{\underset{\sim}{Z} : \frac{1}{N}\sum_{n=1}^{N} Z_n \geq \bar{Z} + \varepsilon\} \leq \Pr\{|\frac{1}{N}\sum_{n=1}^{N} Z_n - \bar{Z}| \geq \varepsilon\} \frac{\mathrm{Var}\left\{\frac{1}{N}\sum_{n=1}^{N} Z_n\right\}}{\varepsilon^2}$$

$$= \frac{\sigma^2}{N\varepsilon^2}.$$

1.5 For $\underset{\sim}{u} \varepsilon \overline{S(N,\varepsilon)_+}$ we have $1 \leq 2^{s\left[\sum_{n=1}^{N} Z_n - N(H+\varepsilon)\right]}$, $s > 0$.

Hence,

$$F_N^+ = \sum_{\underset{\sim}{u} \varepsilon \overline{S(N,\varepsilon)_+}} P_N(\underset{\sim}{u}) \leq \sum_{\underset{\sim}{u} \varepsilon \overline{S(N,\varepsilon)_+}} P_N(\underset{\sim}{u}) 2^{s\left[\sum_{n=1}^{N} Z_n - N(H+\varepsilon)\right]}$$

$$\leq \sum_{\underset{\sim}{u}} P_N(\underset{\sim}{u}) 2^{s\left[\sum_{n=1}^{N} Z_n - N(H+\varepsilon)\right]}$$

$$= E\left\{2^{s\left[\sum_{n=1}^{N} Z_n - N(H+\varepsilon)\right]}\right\} = \prod_{n=1}^{N} E\left\{2^{s\left[Z_n - (H+\varepsilon)\right]}\right\}$$

$$= 2^{-NG(\ )}$$

where $G(s) = s(H+\varepsilon) - \log E\left\{2^{sZ}\right\}$.

$$= s(H+\varepsilon) - \log\left\{\sum_{k=1}^{A} P(a_k)^{1-s}\right\}$$

and $G(0) = 0$.

Next note that

$$\frac{dG(s)}{ds} = H + \varepsilon - \frac{\sum_{k=1}^{A} P(a_k)^{1-s} \log \frac{1}{P(a_k)}}{\sum_{k=1}^{A} P(a_k)^{1-s}}$$

where

$$\left.\frac{dG(s)}{ds}\right|_{s=0} = \varepsilon > 0$$

and

$$\frac{d^2G(s)}{ds^2} = -\left[\sum_{k=1}^{A} Q_s(a_k)\left(\log\frac{1}{P(a_k)}\right)^2 - \left(\sum_{k=1}^{A} Q_s(a_k)\log\frac{1}{P(a_k)}\right)^2\right]$$

$$< 0$$

since this is negative of the variance of $\log\frac{1}{P(u)}$ with respect to distribution

$$Q_s(u) = \frac{P(u)^{1-s}}{\sum_{k=1}^{A} P(a_k)^{1-s}}$$

Thus, there is a unique maximum of $G(s)$ for some $s^{*} > 0$ where $G(s^{*}) > 0$. Then we have $F_N^{+} \leq 2^{-NG(s^{*})}$. Similarly we get $F_N^{-} \leq 2^{-NG^2(s^{**})}$ where $\tilde{G}(s^{**}) > 0$. Then

$$F_N = F_N^{+} + F_N^{-} \leq 2^{-NG(s^{*})} + 2^{-N\tilde{G}(s^{**})}$$

1.6 Multiply the inequalities

$$\log\frac{1}{P_N(\underset{\sim}{u})} \leq \ell(\underset{\sim}{u}) \leq \log\frac{1}{P_N(\underset{\sim}{u})} + 1$$

by $P_N(\underset{\sim}{u})$ and sum over all $\underset{\sim}{u} \varepsilon \mathscr{U}_N$. Then

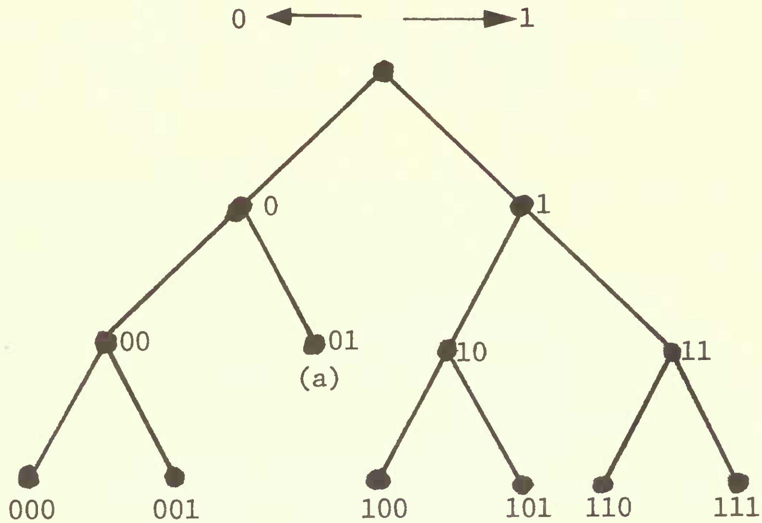
$$H(\mathscr{U}_N) \leq \langle L_N \rangle < H(\mathscr{U}_N) + 1.$$

But from (1.1.14), $H(\mathscr{U}_N) = NH(\mathscr{U})$ giving us (a) and (b). To show (c) note that in (a) the inequality $\log\frac{1}{P(u)} \leq \ell(u) \Longleftrightarrow 2^{-\ell(u)} \leq P(u)$

Hence we satisfy the Kraft-McMillan inequality,

$$\sum_{u} 2^{-\ell(u)} \leq \sum_{u} P(u) = 1.$$

The solution is best understood in terms of a tree diagram where left directed branches correspond to "0" and right to "1". Each node corresponds to a binary sequence so that code words can be represented as nodes in a tree such as



If node (a) is selected as a codeword of length 2, then in order that no other codeword have 01 as a prefix no nodes that branch out from node (a) can ever be selected as a codeword. Hence we can terminate the branching at this node. Thus uniquely decodable codes with the property that no codeword is a prefix of another codeword corresponds to nodes in a tree where no codeword node branches out from a shorter codeword node. In general, let $\ell_1, \ell_2, \ldots, \ell_A$ be the set of codeword lengths where $\ell_1 \leq \ell_2 \leq \ldots \leq \ell_A$. If $\sum_{i=1}^{A} 2^{-\ell i} = 1$ then we can easily find such a code with these lengths where all branch paths of the code tree terminate in a codeword node. If $\sum_{i=1}^{A} 2^{-\ell i} < 1$ then some branch paths can continue forever without encountering a codeword node. Let $\underline{X}_1$ be any node sequence with ℓ_1

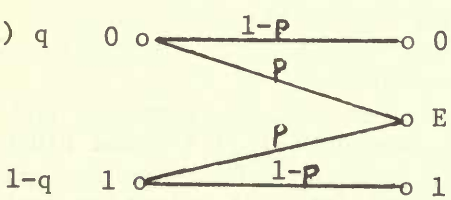
branches leading to it and denote it the first codeword. There remain $2^{\ell_1}-1$ unterminated nodes at the same level that can be a codeword or a prefix of a codeword. If $\ell_2 = \ell_1$ choose any one of these remaining nodes as the codeword node of sequence $\underline{X}_2$. If $\ell_2 > \ell_1$ then use any one of these as a prefix and proceed along any $\ell_2 - \ell_1$ additional branches to find a codeword node of $\underline{X}_2$. There now remains $(2^{\ell_1}-1)\,2^{\ell_2-\ell_1}-1$ nodes at the level ℓ_2 that can be a codeword or a prefix of a codeword. Continue in this manner until $\underline{X}_A$ is selected. If at any point this procedure cannot be completed because of no remaining nodes then $\sum_{i=1}^{A} 2^{-\ell_i} > 1$, which is a contradiction.

<u>1.7</u> (a)
$$I(\mathscr{X};\mathscr{Y}) = \sum_y \sum_x P(y,x)\log\frac{P(y,x)}{P(y)q(x)}$$
$$= \underbrace{\sum_y \sum_x P(y,x)\log\frac{1}{q(x)}}_{\substack{\shortparallel \\ H(\mathscr{X})}} - \underbrace{\sum_y \sum_x P(y,x)\log\frac{1}{q(x|y)}}_{\substack{\shortparallel \\ H(\mathscr{X}|\mathscr{Y})}}$$

But $H(\mathscr{X}|\mathscr{Y}) \geq 0$ (see 1.1.9) so $I(\mathscr{X};\mathscr{Y}) \leq H(\mathscr{X})$ and by symmetry $I(\mathscr{X};\mathscr{Y}) \leq H(\mathscr{Y})$.

(b) Use $P(x,y) = P(y)q(x|y)$ in
$$H(\mathscr{X},\mathscr{Y}) = \sum_x \sum_y P(x,y)\log\frac{1}{P(x,y)}$$
$$= \sum_x \sum_y P(x,y)\log\frac{1}{P(y)} + \sum_x \sum_y P(x,y)\log\frac{1}{q(x|y)}$$
$$= H(\mathscr{Y}) + H(\mathscr{X}|\mathscr{Y}).$$

<u>1.8</u> (a)



$P(0) = q(1-p)$

$P(E) = p$

$P(1) = (1-q)(1-p)$

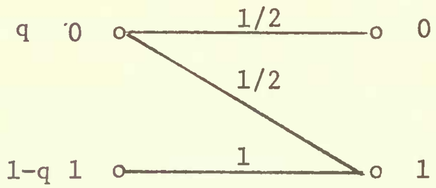
$$I(\mathscr{X};\mathscr{Y}) = q(1-p)\log\left(\frac{1-p}{q(1-p)}\right) + qp\log\left(\frac{p}{p}\right) + (1-q)p\log\left(\frac{p}{p}\right)$$

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

$$= (1-p)\mathscr{H}(q)$$

$q = \frac{1}{2}$ maximizes $\mathscr{H}(q)$ so $C = 1-p$.

(b)



$P(0) = \frac{1}{2} q$

$P(1) = 1 - \frac{1}{2} q$

$$I(\mathscr{X};\mathscr{Y}) = \frac{1}{2} q \log\left(\frac{\frac{1}{2}}{\frac{1}{2} q}\right) + \frac{1}{2} q \log\left(\frac{\frac{1}{2}}{\frac{1}{2} q}\right) + (1-q)\log\left(\frac{1}{1-\frac{1}{2} q}\right)$$

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

$$\frac{d}{dq} I(\mathscr{X};\mathscr{Y}) = \frac{1}{2}\log\left(\frac{2-q}{q}\right) - 1 = 0 \Longrightarrow q = \frac{2}{5}$$

Hence $C = \log\frac{5}{4}$.

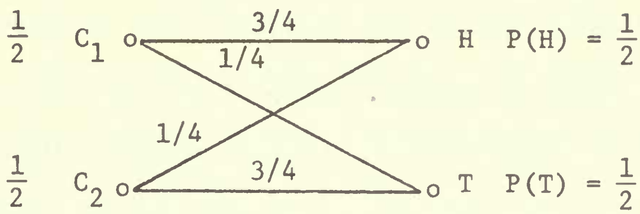
1.9 Since the encoder keeps sending the information symbol until an unerased channel output is achieved there is no error and $P_e = 0$. The probability that n channel symbols are transmitted for an information symbol is $P_n = (1-p)p^{n-1}$. The average length of a codeword is thus

$$L = \sum_{n=1}^{\infty} nP_n = \sum_{n-1}^{\infty} n(1-p)p^{n-1} = \frac{1}{1-p}$$

Thus $R = \frac{1}{L} = 1-p$ bits per channel use is the rate which also equals the channel capacity (see problem 1.8a).

1.10 There are two coins C_1 and C_2 where $P(H|C_1) = \frac{3}{4}$ and $P(H|C_2) = \frac{1}{4}$. Here $q(C_1) = q(C_2) = \frac{1}{2}$ are the probabilities of selecting each coin.

We can interpret this as a BSC with $p = \frac{1}{4}$ as follows:



(a) $I(C_1;H) = \log \frac{3/4}{1/2} = \log \frac{3}{2}$ bit

$I(C_2;H) = \log \frac{1/4}{1/2} = \log \frac{1}{2} = -1$ bit

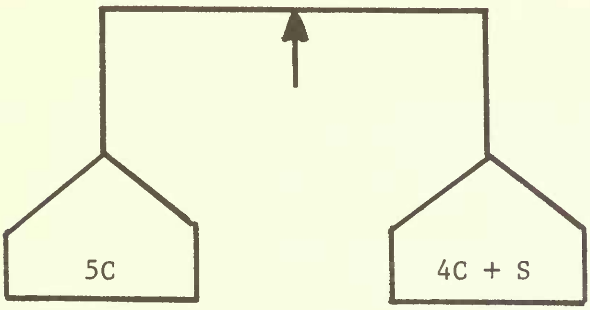
(b) $I(\mathscr{X};\mathscr{Y}) = 1 - \mathscr{H}\left(\frac{1}{4}\right)$ bits.

1.11 (a) There are 13 ways one coin may be heavier, 13 ways one coin may be lighter, and the possibility that all coins weigh the same. Thus we are attempting to determine one of 27 possible situations using a balance and a known standard coin. We are thus asked to obtain at most log 27 = 3 log 3 bits of information on the average. Each weighing has 3 possible outcomes (left, right, balance) and provides at most log 3 bits of average information. Two weighings has 9 possible outcomes and at most 2 log 3 bits of average information. Three weighings has 27 possible outcomes. Clearly two weighings cannot guarantee determining one of 27 possible situations while with three it may be possible.

(b) The maximum amount of average information from three weighings is log 27 which is achieved if all 27 weighing sequences are equally probable. This means that we must choose a weighing strategy where the outcomes of each weighing are equal probable and each weighing outcome is independent of other weighing outcomes. Clearly each weighing must reduce the number of possible cases by 1/3.

Strategy: Let S denote the standard coin, C denote a coin that can be heavy, light, or normal, h denote a coin that is heavy or normal, and ℓ denote a coin that is light or normal. We start with 13C and an S.

1st Weighing: Set aside 4C and place 5C on the left pan and 4C + S on the right pan.



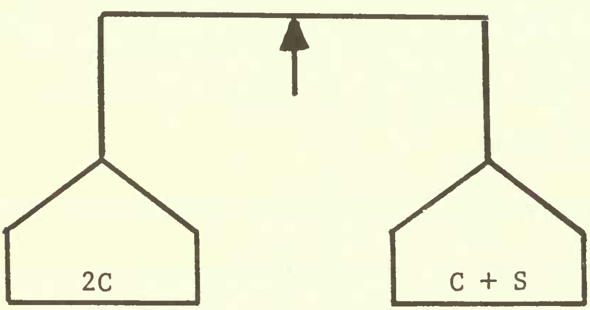
Balance ⟹ 4C remain

Left ⟹ 5h, 4ℓ remain

Right ⟹ 5ℓ, 4h remain

Note that for each of the three possible outcomes we have 9 remaining unknown possibilities to be resolved with two more weighings.

2nd Weighing When 1st Outcome is Balanced: Of the 4C set aside C and place 2C on the left and C + S on the right.



Balance ⟹ C remain

Left ⟹ 2h, ℓ remain

Right ⟹ 2ℓ, h remain

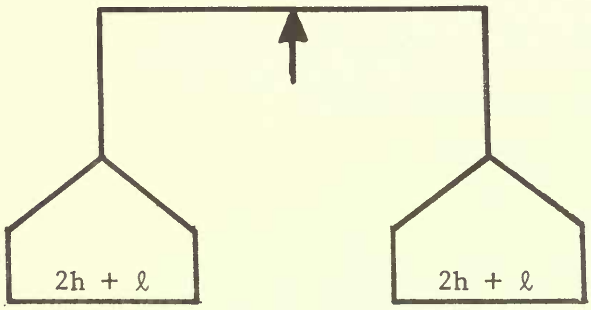
Note that here each outcome leaves only 3 remaining unknown possibilities to be resolved with one more weighing. The third weighing given the 2nd weighing outcome is as follows:

(i) Balance: Place C on the left and S on the right.

(ii) Left: Place h on the left and h on the right.

(iii) Right: Place ℓ on the left and ℓ on the right.

<u>2nd Weighing When 1st Outcome is Left</u>: Of the 5h and 4ℓ remaining set aside $h + 2\ell$ on the left and $2h + \ell$ on the right.



Balance $\Longrightarrow$ h, 2ℓ remain

Left $\Longrightarrow$ 2h, ℓ remain

Right $\Longrightarrow$ 2h, ℓ remain

One more weighing easily resolves the 3 remaining unknown possibilities by placing the same type of coin on the balance. That is, if 2h, ℓ remain then place h on the left and h on the right.

(c) Without a standard coin we cannot always reduce the number of unknown possibilities by 1/3 with each weighing. This is a necessary requirement.

<u>1.12</u> Using (1.1.8) we have

$$H_i(\mathcal{U}) = \sum_u P_i(u)\log\frac{1}{P_i(u)} \le \sum_u P_i(u)\log\frac{1}{P_\lambda(u)}, \quad i = 1,2.$$

Hence $\lambda H_1(\mathcal{U}) + (1-\lambda)\, H_2(\mathcal{U}) \le \sum_u \left[\lambda P_1(u) + (1-\lambda)\, P_2(u)\right] \log\frac{1}{P_\lambda(u)}$

$$= \sum_u P_\lambda(u)\log\frac{1}{P_\lambda(u)} = H_\lambda(\mathcal{U}).$$

<u>1.13</u> Let $\hat{p}(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}}\, e^{-\frac{y^2}{2\sigma_y^2}}$. Then (1.1.8) gives

$$\int_{-\infty}^{\infty} p(y)\log \frac{1}{p(y)}\, dy \leq \int_{-\infty}^{\infty} p(y)\log \frac{1}{\hat{p}(y)}\, dy = \int_{-\infty}^{\infty} p(y)\left[\frac{y^2}{2\sigma_y^2}\log e + \frac{1}{2}\log\left(2\pi\sigma_y^2\right)\right] dy$$

$$= \frac{1}{2}\log e + \frac{1}{2}\log\left(2\pi\sigma_y^2\right) = \frac{1}{2}\log\left(2\pi e\sigma_y^2\right)$$

with equality when $p(y) = \hat{p}(y)$ or y is a Gaussian random variable.

For the additive Gaussian noise channel we have

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

With input probability density q(x), the average mutual information is

$$I(\mathscr{X};\mathscr{Y}) = \int_{-\infty}^{\infty} p(y)\log \frac{1}{p(y)}\, dy - \int_{-\infty}^{\infty} q(x)\left\{\int_{-\infty}^{\infty} p(y|x)\log \frac{1}{p(y|x)}\, dy\right\} dx$$

But

$$\int_{-\infty}^{\infty} p(y|x)\log \frac{1}{p(y|x)}\, dy = \int_{-\infty}^{\infty} p(y|x)\left[\frac{(y-x)^2}{2\sigma^2} + \frac{1}{2}\log\left(2\pi\sigma^2\right)\right] dy$$

$$= \frac{1}{2}\log\left(2\pi e\sigma^2\right) .$$

Using this plus the above bound we have

$$I(\mathscr{X};\mathscr{Y}) \leq \frac{1}{2}\log\left(2\pi e\sigma_y^2\right) - \frac{1}{2}\log\left(2\pi e\sigma^2\right)$$

$$= \frac{1}{2}\log \frac{\sigma_y^2}{\sigma^2}$$

with equality if and only if y is a Gaussian random variable. It is Gaussian if

$$q(x) = \frac{1}{\sqrt{2\pi\mathscr{E}}}\, e^{-\frac{x^2}{2\mathscr{E}}}$$

with the result that $\sigma_y^2 = \mathscr{E} + \sigma^2$. Hence

$$\max I(\mathscr{X};\mathscr{Y}) = \frac{1}{2}\log\left(\frac{\mathscr{E}+\sigma^2}{\sigma^2}\right) = \frac{1}{2}\log\left(1+\frac{\mathscr{E}}{\sigma^2}\right)$$

which is channel capacity.

1.14 From problem 1.6(b) we have inequalities

$$-\log P_N(\underset{\sim}{u}) \leq \ell(\underset{\sim}{u}) < -\log P_N(\underset{\sim}{u}) + 1.$$

Thus

$$\frac{-\log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} \leq 1$$

and

$$\frac{-\log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} \geq \frac{\ell(\underset{\sim}{u})-1}{\ell(\underset{\sim}{u})} = 1 - \frac{1}{\ell(\underset{\sim}{u})} \geq 1 + \frac{1}{\log P_N(\underset{\sim}{u})}$$

But

$$P_N(\underset{\sim}{u}) = \prod_{n=1}^{N} P(u_n) \leq \prod_{n=1}^{N} P(u^*) = P(u^*)^N$$

and

$$\log P_N(\underset{\sim}{u}) \leq N \log P(u^*)$$

Hence

$$1 + \frac{1}{N\log P(u^*)} \leq \frac{-\log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} \leq 1$$

and

$$1 + \frac{1}{N\log P(u^*)} \leq H_N \leq 1.$$

If $H_N = 1$ for all N then $\dfrac{-\log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} = 1$ for all N.

Hence $P_N(\underset{\sim}{u}) = 2^{-\ell(\underset{\sim}{u})}$ for all N which is possible only if $P_N(\underset{\sim}{u}) = 2^{-N}$

and the source is a BSS.

1.15 For $\underset{\sim}{u} \in S(N,\varepsilon)$,

$$N[H(\mathscr{U}) + \varepsilon] \le \ell(\underset{\sim}{u}) < N[H(\mathscr{U}) + \varepsilon] + 1$$

and

$$N[H(\mathscr{U}) - \varepsilon] \le - \log P_N(\underset{\sim}{u}) \le N[H(\mathscr{U}) + \varepsilon]$$

Thus for $\underset{\sim}{u} \in S(N,\varepsilon)$,

$$\frac{N[H(\mathscr{U}) - \varepsilon]}{1 + N[H(\mathscr{U}) + \varepsilon]} \le \frac{- \log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} \le 1 \tag{1}$$

For $\underset{\sim}{u} \in S(N,\varepsilon)$,

$$N \log A \le \ell(\underset{\sim}{u}) \le N \log A + 1$$

and

$$- N \log P(u^*) \le - \log P_N(\underset{\sim}{u}) \le - N \log P(u^{**})$$

where $P(u^{**}) = \min\limits_{u} P(u) > 0$. Thus for $\underset{\sim}{u} \in \overline{S(N,\varepsilon)}$,

$$\frac{- N \log P(u^*)}{N \log A + 1} \le \frac{- \log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})} \le \frac{- N \log P(u^{**})}{N \log A} \tag{2}$$

Define $F_N = \Pr\left\{\underset{\sim}{u} \in \overline{S(N,\varepsilon)}\right\}$ and using the form

$$H_N = \sum_{\underset{\sim}{u}} P_N(\underset{\sim}{u}) \left(\frac{- \log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})}\right)$$

$$= \sum_{\underset{\sim}{u} \in S(N,\varepsilon)} P_N(\underset{\sim}{u}) \left(\frac{- \log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})}\right) + \sum_{\underset{\sim}{u} \in \overline{S(N,\varepsilon)}} P_N(\underset{\sim}{u}) \left(\frac{- \log P_N(\underset{\sim}{u})}{\ell(\underset{\sim}{u})}\right)$$

we have from inequalities (1) and (2),

$$H_N \ge \frac{N[H-\varepsilon]}{1 + N[H+\varepsilon]} (1-F_N) + \frac{- N \log P(u^*)}{N \log A + 1} F_N$$

$$= \frac{H - \varepsilon}{H + \varepsilon + \frac{1}{N}} (1-F_N) + \frac{- \log P(u^*)}{\log A + \frac{1}{N}} F_N$$

$$\xrightarrow[N \to \infty]{} \frac{H - \varepsilon}{H + \varepsilon} \quad \text{since} \quad F_N \xrightarrow[N \to \infty]{} 0 ,$$

and

$$H_N \leq (1-F_N) + \frac{-N \log P(u^{**})}{N \log A} F_N$$

$$= 1 - F_N - \frac{\log P(u^{**})}{\log A} F_N \xrightarrow[N \to \infty]{} 1.$$

Hence for any $\varepsilon > 0$,

$$\frac{H - \varepsilon}{H + \varepsilon} \leq \lim_{N \to \infty} H_N \leq 1$$

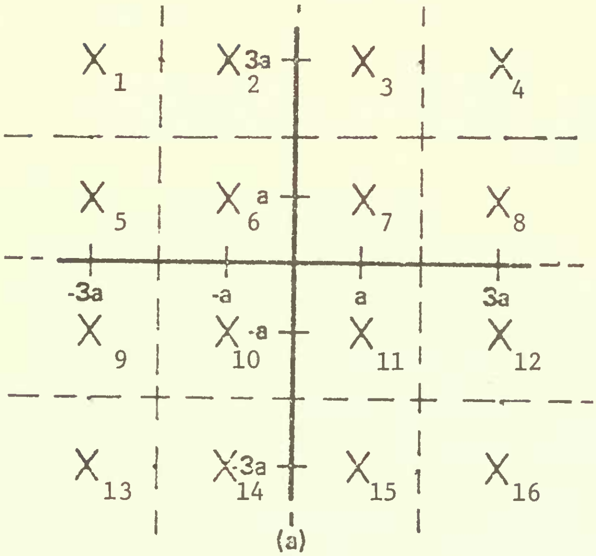
or

$$\lim_{N \to \infty} H_N = 1.$$

—//—

CHAPTER 2

2.1 (a)



The noise components in the perpendicular coordinate directions are independent with variance $N_0/2$.

Let $q = \Pr\{n>a\}$

$$= Q\left(\sqrt{\frac{2}{N_0}}\,a\right),$$

$$\mathcal{E}_{av} = \frac{1}{16}\sum_{i=1}^{16} \mathcal{E}_i$$

$$= 10a^2$$

$$P_{C_1} = P_{C_4} = P_{C_{13}} = P_{C_{16}} = \Pr\{n_1<a, n_2<a\} = (1-q)^2$$

$$P_{C_2} = P_{C_3} = P_{C_5} = P_{C_8} = P_{C_9} = P_{C_{12}} = P_{C_{14}} = P_{C_{15}} = \Pr\{n_1<a, -a<n_2<a\} = (1-q)(1-2q)$$

$$P_{C_6} = P_{C_7} = P_{C_{10}} = P_{C_{11}} = \Pr\{-a<n_1<a, -a<n_2<a\} = (1-2q)^2$$

Hence

$$P_C = \frac{1}{16}\left[4(1-q)^2 + 8(1-q)(1-2q) + 4(1-2q)^2\right]$$

$$= \frac{1}{4}\left[(1-q)^2 + 2(1-q)(1-2q) + (1-2q)^2\right]$$

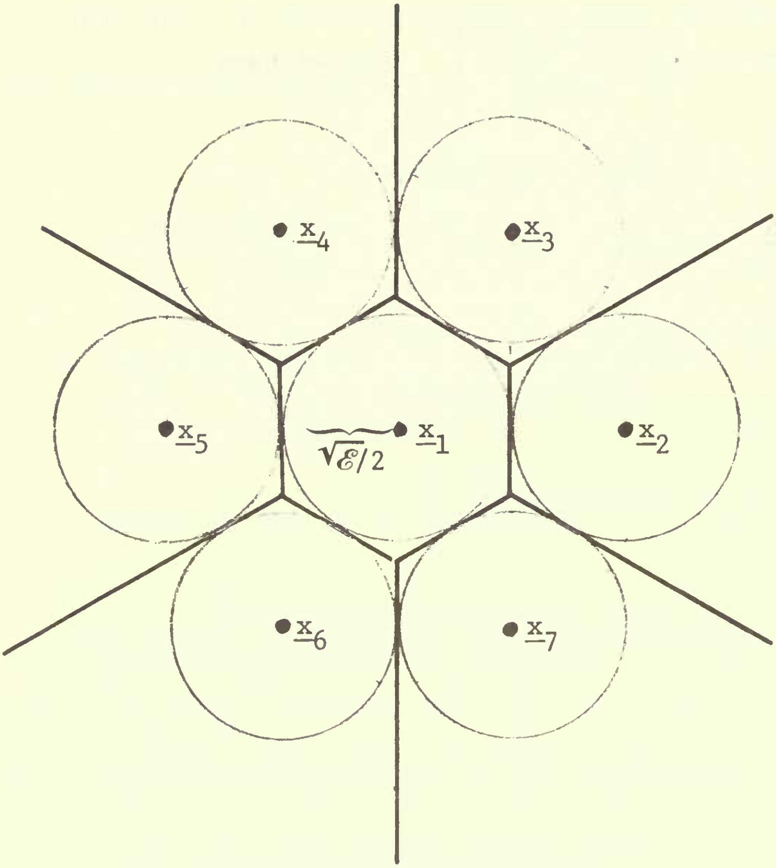
$$= \frac{1}{4}\left[(1-q) + (1-2q)\right]^2 = \left(\frac{2-3q}{2}\right)^2 = \left(1 - \frac{3}{2}q\right)^2$$

and

$$P_E = 3q - \frac{9}{4}q^2$$

(b) Rotation and translation does not change P_E.

2.2 (a)



(b) $\Lambda_m = \{\underline{y}: ||\underline{y}-\underline{x}_{m'}||^2 < ||\underline{y}-\underline{x}_m||$ for all $m' \neq m\}$ $m = 1,2,\ldots,7$ are the optimum decision boundaries. We have

$$S_m = \{\underline{y}: ||\underline{y}-\underline{x}_m||^2 \leq \sqrt{\mathscr{E}}/2\} \subset \Lambda_m \quad m = 1,2,\ldots,7.$$

Hence

$$P_{E_m} = \Pr\{\underline{y} \notin \Lambda_m | \underline{x}_m\} \leq \Pr\{\underline{y} \notin S_m | \underline{x}_m\}$$

$$= \Pr\left\{n_1^2+n_2^2 > \frac{\mathscr{E}}{4}\right\} = e^{-\frac{\mathscr{E}}{4N_0}} \quad m = 1,2,\ldots,7.$$

and

$$P_E = \frac{1}{7}\sum_{m=1} P_{E_m} \leq e^{-\frac{\mathscr{E}}{4N_0}}$$

2.3 $P_{E_m} \le \sum_{m' \ne m} P_E(m \to m')$ where

$$P_E(m \to m') = Q\left(\frac{||\underline{S}_m - \underline{S}_{m'}||}{\sqrt{2N_0}}\right).$$ Here we have

$$||\underline{S}_1 - \underline{S}_{m'}|| = \sqrt{\mathscr{E}} \qquad m' = 2,3,\ldots,7$$

$$||\underline{S}_2 - \underline{S}_{m'}|| = \sqrt{\mathscr{E}} \qquad m' = 3,7$$

$$||\underline{S}_2 - \underline{S}_{m'}|| = \sqrt{3\mathscr{E}} \qquad m' = 4,6$$

$$||\underline{S}_2 - \underline{S}_{m'}|| = 2\sqrt{\mathscr{E}} \qquad m' = 5$$

$$P_{E_1} \le \sum_{m' \ne 1} P_E(1 \to m') = 6\,Q\left(\sqrt{\frac{\mathscr{E}}{2N_0}}\right)$$

$$P_{E_2} \le \sum_{m' \ne 2} P_E(2 \to m') = 2\,Q\left(\sqrt{\frac{\mathscr{E}}{2N_0}}\right) + 2\,Q\left(\sqrt{\frac{3\mathscr{E}}{2N_0}}\right) + Q\left(\sqrt{\frac{2\mathscr{E}}{N_0}}\right)$$

By symmetry $P_{E_2} = P_{E_3} = P_{E_4} = P_{E_5} = P_{E_6} = P_{E_7} < P_{E_1}$.

Hence

$$P_E \le P_{E_1} \le 6\,Q\left(\sqrt{\frac{\mathscr{E}}{2N_0}}\right).$$

From problem 2.2 we have $P_E \le e^{-\frac{\mathscr{E}}{4N_0}}$ which is "exponentially" the same since $Q(x)$ is bounded as shown in (2.3.18).

2.4 (a) Choose $\phi_m(t) = \frac{1}{\sqrt{\mathscr{E}}} x_m(t) \quad m = 1,2,\ldots, M$ as the orthonormal basis. Then $x_{mn} = \sqrt{\mathscr{E}}\delta_{mn}$ $m,n = 1,2,\ldots, M$ and $||\underline{x}_m||^2 = \mathscr{E}$ for all m. The decision boundries become

$$\Lambda_m = \{\underline{y}: ||\underline{y} - \underline{x}_m||^2 < ||\underline{y} - \underline{x}_{m'}||^2 \text{ for all } m' \ne m\}$$

$$= \{\underline{y}: (\underline{y}, \underline{x}_m) > (\underline{y}, \underline{x}_{m'}) \text{ for all } m' \ne m\}$$

$$= \{\underline{y}: y_m > y_{m'} \text{ for all } m' \ne m\} \quad m = 1,2,\ldots,M.$$

Hence by symmetry,

$$P_E = P_{E_1} = 1 - P_{C_1} = 1 - \Pr\{y_1 > y_m \text{ for all } m \ne 1 | \underline{x}_1\}$$

(b) Given $\underset{\sim}{x}_1$ was sent we have probability density functions for the independent random variables $y_1, y_2, \ldots, y_M$ given by

$$p_{y_1}(y) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(y-\sqrt{\mathscr{E}})^2}{N_0}}$$

$$p_{y_m}(y) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{y^2}{N_0}} \qquad m = 2,3,\ldots,M$$

Hence

$$\Pr\{y_m < y_1 \text{ for all } m \neq 1 | \underset{\sim}{x}_1\}$$

$$= \int_{-\infty}^{\infty} \Pr\{y_m < \alpha \text{ for all } m \neq 1 | \underset{\sim}{x}_1, y_1 = \alpha\} \, p_{y_1}(\alpha) d\alpha$$

(c) Using $\mathscr{E} = \mathscr{E}_b \log_2 M$, the fact that as ε gets small $\ln(1+\varepsilon)$ behaves as ε, and $Q(x)$ behaves like $e^{-x^2/2}$ for large x,

$$\lim_{M\to\infty} \ln\left[1 - Q\left(x + \sqrt{2\mathscr{E}_b \log M/N_0}\right)\right]^{M-1}$$

$$= \lim_{M\to\infty} (M-1) \ln\left[1 - Q\left(x + \sqrt{2\mathscr{E}_b \log M/N_0}\right)\right]$$

$$= -\lim_{M\to\infty} M\, Q\left(x + \sqrt{2\mathscr{E}_b \log M/N_0}\right)$$

$$= -\lim_{M\to\infty} M \exp\left\{-\frac{1}{2}\left(x + \sqrt{2\mathscr{E}_b \log M/N_0}\right)^2\right\}$$

$$= -\lim_{M\to\infty} M \exp\left\{-\frac{\mathscr{E}_b \log M}{N_0}\right\}$$

$$= -\lim_{M\to\infty} M^{1 - \frac{\mathscr{E}_b}{N_0 \ln 2}}$$

$$= \begin{cases} -\infty \quad ; & \frac{\mathscr{E}_b}{N_0} < \ln 2 \\ 0 \quad ; & \frac{\mathscr{E}_b}{N_0} > \ln 2 \end{cases}$$

Hence

$$\lim_{M\to\infty}\left[1-Q\left(x+\sqrt{2\mathcal{E}/N_0}\right)\right]^{M-1}=\begin{cases}0 & ; \quad \dfrac{\mathcal{E}_b}{N_0}<\ln 2\\ 1 & ; \quad \dfrac{\mathcal{E}_b}{N_0}>\ln 2\end{cases}$$

But

$$\Pr\{y_m<\alpha \text{ for all } m\neq 1 \mid \underset{\sim}{x}_1, y_1=\alpha\}$$

$$=\prod_{m=2}^{M}\Pr\{y_m<\alpha\mid \underline{x}_1\}$$

$$=\left[1-Q\left(\frac{\alpha}{\sqrt{N_0/2}}\right)\right]^{M-1}$$

Hence

$$P_E=1-\int_{-\infty}^{\infty}\left[1-Q\left(\frac{\alpha}{\sqrt{N_0/2}}\right)\right]^{M-1}\frac{1}{\sqrt{\pi N_0}}\,e^{-\frac{(\alpha-\sqrt{\mathcal{E}})^2}{N_0}}\,d\alpha$$

$$=1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}\left[1-Q\left(x+\sqrt{\frac{2\mathcal{E}}{N_0}}\right)\right]^{M-1}dx$$

2.5 (a) $K=1 \Longrightarrow \underset{\sim}{x}_1=\sqrt{\mathcal{E}/2}\,[1\ 1],\ \underset{\sim}{x}_2=\sqrt{\mathcal{E}/2}\,[1\ -1]$

and $(\underset{\sim}{x}_1,\underset{\sim}{x}_2)=(\mathcal{E}/2)\,[1\ 1]\begin{bmatrix}1\\-1\end{bmatrix}=0.$

Let $\underset{\sim}{h}_{K-1}^{(i)}$ denote the i^{th} row of H_{K-1} and $\underset{\sim}{h}_K^{(i)}$ the i^{th} row of H_K.
The codewords of length $M=2^K$ are given by

$$\underset{\sim}{h}_K^{(i)}=\left[\underset{\sim}{h}_{K-1}^{(i)}\ \underset{\sim}{h}_{K-1}^{(i)}\right] \text{ and } \underset{\sim}{h}_K^{(i+2^{K-1})}=\left[\underset{\sim}{h}_{K-1}^{(i)}\ \ -\underset{\sim}{h}_{K-1}^{(i)}\right],\ i=1,\dots,2^{K-1}$$

Suppose $\left(\underset{\sim}{h}_{K-1}^{(i)},\ \underset{\sim}{h}_{K-1}^{(i)}\right)=0$ for all $i\neq j<2^{K-1}$

Then

$$\left(\underset{\sim}{h}_{K-1}^{(i)}\ \underset{\sim}{h}_{K-1}^{(i)},\ \underset{\sim}{h}_{K-1}^{(j)}\ \underset{\sim}{h}_{K-1}^{(j)}\right)=0 \text{ for all } i\neq j\leq 2^{K-1}$$

and

$$\left(\underset{\sim}{h}_{K-1}^{(i)}\ \underset{\sim}{h}_{K-1}^{(i)},\ \underset{\sim}{h}_{K-1}^{(j)}\ -\underset{\sim}{h}_{K-1}^{(j)}\right)=0 \text{ for all } i,\ j\leq 2^{K-1}$$

yields

$$\left(\underset{\sim}{h}_K^{(i)}, \underset{\sim}{h}_K^{(j)}\right) = 0 \quad \text{for all } i \neq j \leq 2^K .$$

(b) Note that $\left(\underset{\sim}{h}_K^{(i)}, \underset{\sim}{h}_K^{(j)}\right) = \mathcal{E}\delta_{ij}$. Hence if we subtract the 1st product term of this inner product we have

$$(\hat{\underset{\sim}{x}}_j, \hat{\underset{\sim}{x}}_k) = \mathcal{E}\delta_{jk} - \mathcal{E}/M = \begin{cases} \mathcal{E}\left(\frac{M-1}{M}\right) = \mathcal{E}' , & j = k \\ -\frac{\mathcal{E}}{M} = -\frac{\mathcal{E}'}{M-1} , & j \neq k . \end{cases}$$

(c) Let $\underset{\sim}{a} = \sqrt{\mathcal{E}/M}\,[1\ 0\ 0\ \ldots\ 0]$. Then the orthogonal signal set $\left\{\underset{\sim}{h}_K^{(i)}\right\}$ and the simplex signal set are essentially related by a simple translation of the signal set given by,

$$\hat{\underset{\sim}{x}}_i = \underset{\sim}{h}_K^{(i)} - \underset{\sim}{a} \qquad i = 1,2,\ldots,2^K .$$

(Since the 1st component of $\hat{\underset{\sim}{x}}_i$ is always "0" we can ignore it.) This means the error probability of the orthogonal signal set of energy $\mathcal{E} = \mathcal{E}'\left(\frac{M}{M-1}\right)$ is the same as the error probability of the simplex signal set of energy $\mathcal{E}'$.

(d) Let $\underset{\sim}{W} = \sum_{i=1}^{M} \underset{\sim}{x}_i$. Then,

$$\begin{aligned} 0 \leq (\underset{\sim}{W},\underset{\sim}{W}) &= \sum_i \sum_j (\underset{\sim}{x}_i, \underset{\sim}{x}_j) \\ &= \sum_i \sum_{j=i} (x_i, x_j) + \sum_{i \neq j} (\underset{\sim}{x}_i, \underset{\sim}{x}_j) \\ &= M\mathcal{E} + \sum\sum_{i \neq j} (\underset{\sim}{x}_i, \underset{\sim}{x}_j) \end{aligned}$$

or

$$\sum\sum_{i \neq j} (\underset{\sim}{x}_i \underset{\sim}{x}_j) \geq -M\mathcal{E}$$

or

$$\rho_{AV} \equiv \frac{1}{\mathscr{E} M(M-1)} \sum_{i \neq j}\sum (\underset{\sim}{x}_i, \underset{\sim}{x}_j) \geq - \frac{1}{M-1} .$$

(e) Let $\underset{\sim}{z}_1, \underset{\sim}{z}_2, \ldots, \underset{\sim}{z}_M$ be an orthogonal signal set of energy $\hat{\mathscr{E}}$. Then $(\underset{\sim}{z}_i, \underset{\sim}{z}_j) = \hat{\mathscr{E}} \delta_{ij}$.

Let

$$\underset{\sim}{a} = \alpha \frac{1}{M} \sum_{i=1}^{M} \underset{\sim}{z}_i$$

and consider the translation set formed by

$$\underset{\sim}{y}_i = \underset{\sim}{z}_i - \underset{\sim}{a}$$

Then

$$\begin{aligned}
(\underset{\sim}{y}_i, \underset{\sim}{y}_j) &= (\underset{\sim}{z}_i - \underset{\sim}{a}, \underset{\sim}{z}_j - \underset{\sim}{a}) \\
&= (\underset{\sim}{z}_i, \underset{\sim}{z}_j) - (\underset{\sim}{z}_i, \underset{\sim}{a}) - (\underset{\sim}{z}_j \underset{\sim}{a}) + (\underset{\sim}{a}, \underset{\sim}{a}) \\
&= \hat{\mathscr{E}} \delta_{ij} - \alpha \hat{\mathscr{E}}/M - \alpha \hat{\mathscr{E}}/M + \alpha^2 \hat{\mathscr{E}}/M \\
&= \begin{cases} \hat{\mathscr{E}}(1 - 2\alpha/M + \alpha^2/M) \\ \hat{\mathscr{E}}(\alpha^2 - 2\alpha)/M \end{cases}
\end{aligned}$$

Choosing

$\mathscr{E}$ and α to satisfy the equations,

$$\mathscr{E} = \hat{\mathscr{E}}(1 - 2\alpha/M + \alpha^2/M)$$

and

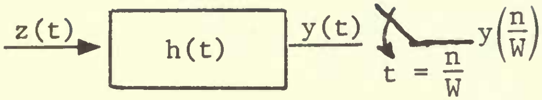
$$\mathscr{E} \rho = \hat{\mathscr{E}}(\alpha^2 - 2\alpha)/M ,$$

yields a signal set $\{\underset{\sim}{y}_i\}$ with the same energy and inner products as the signal set $\{\underset{\sim}{x}_j\}$. It then has the same error probability and

$$P_E\left(\frac{\mathscr{E}}{N_0}, \rho\right) = P_E\left(\frac{\mathscr{E}(1-\rho)}{N_0}, 0\right).$$

2.6 (a) $h(t) = \dfrac{1}{2\pi} \displaystyle\int_{-\pi W}^{\pi W} \frac{1}{W} e^{jwt} \, dw \equiv \frac{1}{2\pi W} \left[\frac{e^{j\pi W t} - e^{-j\pi W t}}{jt} \right] = \frac{\sin \pi W t}{\pi W t}$

(b)



$$y(t) = \int_0^\infty h(t-\tau)z(\tau)\,d\tau = \int_0^\infty \frac{\sin[\pi W(t-\tau)]}{\pi W(t-\tau)} z(\tau)\,d\tau$$

and
$$y\left(\frac{n}{W}\right) = \int_0^\infty \frac{\sin\left[\pi W\left(\frac{n}{W}-\tau\right)\right]}{\pi W\left(\frac{n}{W}-\tau\right)} z(\tau)\,d\tau = \int_0^\infty z(t) \frac{\sin\left[\pi W\left(t-\frac{n}{W}\right)\right]}{\pi W\left(t-\frac{n}{W}\right)}\,dt$$

(c) In Figure 2.9, $y(t)\Phi_{2n}(t)$ is integrated over $[(n-1)T/N, nT/N]$

where
$$\Phi_{2n}(t) = \frac{\sqrt{2}\,\sin\left[\pi W\left(t-\frac{n}{W}\right)\right]}{\pi W\left(t-\frac{n}{W}\right)} \sin \omega_0 T \;.$$

Thus
$$y_{2n} = \int_{(n-1)T/N}^{nT/N} y(t)\,\sqrt{2}\,\sin\omega_0 t\, \frac{\sin\left[\tilde{\pi} W\left(t-\frac{n}{W}\right)\right]}{\pi W\left(t-\frac{n}{W}\right)}\,dt \;.$$

Replace the integration over $[(n-1)T/N, nT/N$ to $[0,\infty)$ we obtain

$$y_{zn} = \int_0^\infty y(t)\,\sqrt{2}\,\sin\,\omega_0 t\, \frac{\sin\left[\pi W\left(t-\frac{n}{W}\right)\right]}{\pi W\left(t-\frac{n}{W}\right)}\,dt$$

which is the process in Figure 2.11.

2.7 (a) Here $y(t) = y_n\phi_n(t)$ where $\phi_n(t) = \sqrt{2N/T}\,\sin\,\omega_0 t$ for $(n-1)T/N \le t \le nT/N$ whereas we compute

$$y'_n = \int_{(n-1)T/N}^{nT/N} y(t)\;\phi'_n(t)\,dt$$

where

$$\phi'_n(t) = \sqrt{2N/T}\,\sin[\omega_0+\Delta\omega)(t-(n-1)T/N) + \phi],$$

$$(n-1)T/N \le t \le nT/N.$$

Using $\int_0^{T/N} \cos[(2\omega_0+\Delta\omega)t+\phi]dt \cong 0$ since $\frac{\omega_0 T}{N} >> 1$

we have
$$y_n' = y_n \int_{(n-1)T/N}^{nT/N} \phi_n(t)\ \phi_n'(t)dt$$

$$= y_n \int_{(n-1)T/N}^{nT/N} (2N/T)\sin \omega_0 t \sin[(\omega_0+\Delta\omega)(t-(n-1)T/N)+\phi]$$

$$= y_n \int_0^{T/N} (2N/T)\sin \omega_0 t \sin[(\omega_0+\Delta\omega)t+\phi]dt$$

$$= y_n(N/T) \int_0^{T/N} \{\cos(\Delta\omega t+\phi) - \cos[(2\omega_0+\Delta\omega)t+\phi]\}dt$$

$$= y_n(N/T)\int_0^{T/N} \cos(\Delta\omega t+\phi)dt$$

$$= y_n(N/T)\int_0^{T/N} \{\cos \Delta\omega t \cos\phi - \sin \Delta\omega t \sin\phi\}dt .$$

We assumed here that $\omega_0 T/N$ is a multiple of π. Also since $\Delta\omega T/N << 1$ we have $\sin \Delta\omega t \cong 0$ for $0 \leq t \leq T/N$.

Hence
$$y_n' = y_n(N/T)\cos\phi \int_0^{T/N} \cos \Delta\omega t\ dt$$

$$= y_n \cos\phi \left(\frac{\sin(\Delta\omega T/N)}{(\Delta\omega T/N)}\right)$$

(b) Here
$$y_{2n}' = \int_{(n-1)T/N}^{nT/N} y(t) \sqrt{2N/T} \sin(\omega_0 t+\phi)dt$$

$$y'_{2n+1} = \int_{(n-1)T/N}^{nT/N} y(t)\,\sqrt{2N/T}\,\cos(\omega_0 t+\phi)\,dt$$

where

$$y(t) = x_{2n}\,\phi_{2n}(t) + x_{2n+1}\,\phi_{2n+1}(t)$$

Thus

$$y'_{2n} = x_{2n}\int_{(n-1)T/N}^{nT/N} (2N/T)\sin\omega_0 t\,\sin(\omega_0+t\phi)\,dt$$

$$+\, x_{2n+1}\int_{(n-1)T/N}^{nT/N} (2N/T)\cos\omega_0 t\,\sin(\omega_0 t+\phi)\,dt$$

$$\cong x_{2n}(N/T)\int_{(n-1)T/N}^{nT/N} \cos\phi\,dt + x_{2n+1}(N/T)\int_{(n-1)T/N}^{nT/N}\sin\phi\,dt$$

$$= x_{2n}\cos\phi + x_{2n+1}\sin\phi$$

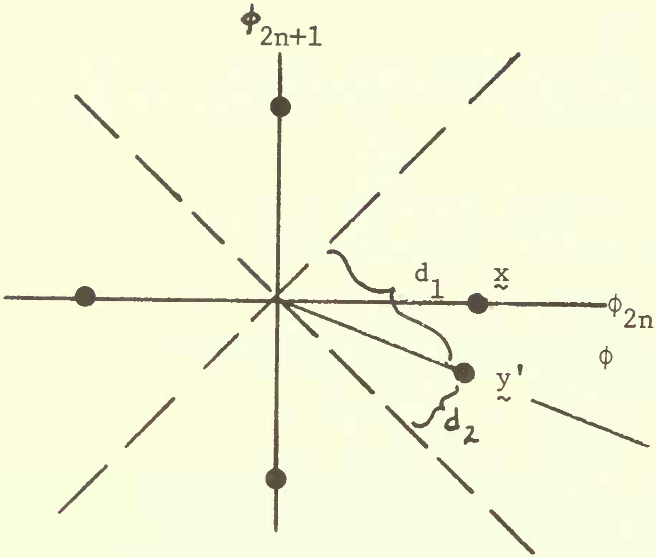
Similarly

$$y'_{2n+1} = -\,x_{2n}\sin\phi + x_{2n+1}\cos\phi\ .$$

(c) The four possible phases are given by

$$\begin{bmatrix} x_{2n} \\ x_{2n+1} \end{bmatrix} = \sqrt{\mathscr{E}}\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \sqrt{\mathscr{E}}\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \sqrt{\mathscr{E}}\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \sqrt{\mathscr{E}}\begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Suppose $x_{2n} = \sqrt{\mathscr{E}}$, $x_{2n+1} = 0$. Then the signal components in the observables become $y'_{2n} = \sqrt{\mathscr{E}_s}\cos\phi$, $y'_{2n+1} = \sqrt{\mathscr{E}_s}\sin\phi$.



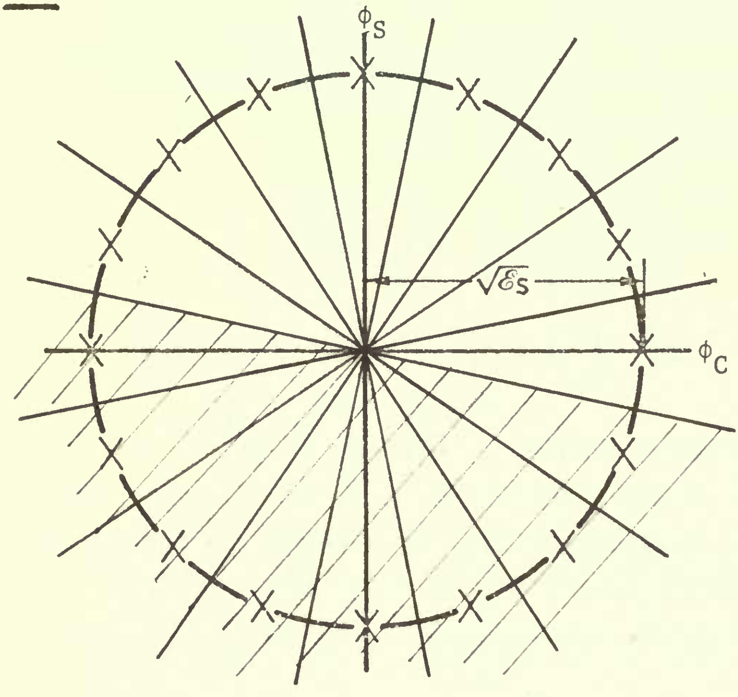
$$d_1 = \sqrt{\mathscr{E}}\cos\left(\frac{\pi}{4} - \phi\right)$$

$$d_2 = \sqrt{\mathscr{E}}\sin\left(\frac{\pi}{4} - \phi\right)$$

The error probability is the probability that the noise components cause the observable vector to lie outside the correct decision region. Hence

$$
\begin{aligned}
P_E &= 1 - P_C = 1 - Pr\,\{n_1 < d_1,\ n_2 < d_2\} \\
&= 1 - Pr\,\{n_1 < d_1\}\, Pr\,\{n_2 < d_2\} \\
&= 1 - \left[1 - Q\left(\sqrt{\frac{2}{N_0}}\, d_1\right)\right]\left[1 - Q\left(\sqrt{\frac{2}{N_0}}\, d_2\right)\right] \\
&= Q\left(\sqrt{\frac{2}{N_0}}\, d_1\right) + Q\left(\sqrt{\frac{2}{N_0}}\, d_2\right) - Q\left(\sqrt{\frac{2}{N_0}}\, d_1\right) \\
&\quad \cdot Q\left(\sqrt{\frac{2}{N_0}}\, d_2\right) \\
&= Q\left(\sqrt{\frac{2\mathcal{E}}{N_0}} \cos\left(\frac{\pi}{4} - \phi\right)\right) + Q\left(\sqrt{\frac{2\mathcal{E}}{N_0}} \sin\left(\frac{\pi}{4} - \phi\right)\right) \\
&\quad - Q\left(\sqrt{\frac{2\mathcal{E}}{N_0}} \cos\left(\frac{\pi}{4} - \phi\right)\right) \cdot Q\left(\sqrt{\frac{2\mathcal{E}}{N_0}} \sin\left(\frac{\pi}{4} - \phi\right)\right)
\end{aligned}
$$

2.8



(a) The boundries of the optimum decision regions are shown as radial lines. Each decision region is a cone of $\frac{\pi}{8}$ radians.

(b) The distance from any signal vector to the boundary is $d = \sqrt{\mathscr{E}}$ $\sin(\pi/16)$. Suppose $\underset{\sim}{x} = (\sqrt{\mathscr{E}},0)$ is the transmitted vector and n_1 and n_2 are the two noise components that are perpendicular to the two nearest boundries. Then an error occurs if the event $\{n_1 \geq d\}$ $\cup \{n_2 \geq d\}$ occurs. Hence

$$P_E = P_r\{\{n_1 \geq d\} \cup \{n_2 \geq d\}\}$$

In general n_1 and n_2 are not independent. However we have the bounds,

$$Pr\{n_1 \geq d\} \leq P_E \leq Pr\{n_1 \geq d\} + Pr\{n_2 \geq d\}$$

or

$$P \leq P_E \leq 2P$$

where

$$\begin{aligned} P &= Pr\{n_1 \geq d\} \\ &= Q\left(\sqrt{\frac{2\mathscr{E}}{N_0}} \sin(\pi/16)\right) \end{aligned}$$

is the probability that the noise forces the observation vector to lie in the shaded region above.

(c) In problem 2.1 the average energy is $\mathscr{E} = 10a^2$ or $a = \sqrt{\mathscr{E}/10}$. The error probability is

$$\begin{aligned} P_E &= 3\, Q\left(\sqrt{\frac{\mathscr{E}}{5N_0}}\right) - \frac{9}{4}\, Q\left(\sqrt{\frac{\mathscr{E}}{5N_0}}\right)^2 \\ &= 3\, Q\left(\sqrt{\frac{\mathscr{E}}{N_0}}\,(.45)\right) - \frac{9}{4}\, Q\left(\sqrt{\frac{\mathscr{E}}{N_0}}\,(.45)\right) \end{aligned}$$

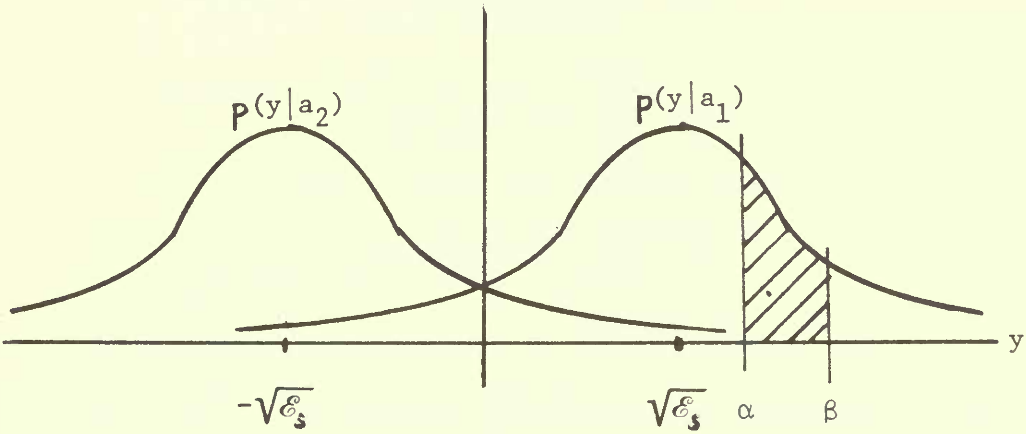
which is generally smaller than $P = Q\left(\sqrt{\frac{\mathscr{E}}{N_0}}\,(.20)\right)$.

Hence the signal set of Figure 2.12(a) is superior to the set of Figure 2.12(b).

2.9 (a) Before quantization the random variable y has conditional probability density

$$p(y|a_k) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(y-a_k)^2}{N_0}} \quad k = 1,2$$

The transition probabilities to the quantized outputs b_1, b_2, ..., b_8 are given by



$$P(b_j|a_k) = \int_{\alpha_j}^{\beta_j} (y|a_k)\,dy \quad k = 1,2;\ j = 1,2,\ \ldots,\ 8.$$

where

$$a_1 = \sqrt{\mathscr{E}_S} \qquad a_2 = -\sqrt{\mathscr{E}_S}$$

$$\alpha_j = 4a - ja \qquad j = 1,2,\ldots,7$$

$$\beta_j = 5a - ja \qquad j = 2,3,\ldots,8$$

$$\alpha_8 = -\infty,\ \beta_1 = \infty$$

(b) Let $\underset{\sim}{x} = (x_1, x_2, \ldots, x_N)$ be any sequence of N symbols from $\{a_1, a_2\}$ and $\underset{\sim}{z} = (z_1, z_2, \ldots, z_N)$ be a sequence of N quantized outputs where $z_n \in \{b_1, b_2, \ldots, b_8\}$. Then the maximum likelihood (optimum) estimate of $\underset{\sim}{x} \in \mathcal{C} = \{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_M\}$ is given by

$$\hat{\underset{\sim}{x}} = \max_{\underset{\sim}{x}}{}^{-1} P_N(\underset{\sim}{z}|\underset{\sim}{x})$$

$$= \max_{\underset{\sim}{x}}{}^{-1} \prod_{n=1}^{N} P(z_n|x_n)$$

$$= \max_{\underset{\sim}{x}}{}^{-1} \log \prod_{n=1}^{N} P(z_n|x_n)$$

$$= \max_{\underset{\sim}{x}}{}^{-1} \sum_{n=1}^{N} \log P(z_n|x_n)$$

$$= \min_{\underset{\sim}{x}}{}^{-1} \sum_{n=1}^{N} \gamma(z_n|x_n)$$

where $\gamma(z|x) = -\log P(z|x)$.

<u>2.10</u> (a) Let $u(z) = \begin{cases} 1 & ; \quad z \geq 0 \\ 0 & ; \quad z < 0 \end{cases}$

Then if $u(z) \leq f(z)$ for all z, we have

$$\Pr\{z \geq 0\} = E\{u(z)\} \leq E\{f(z)\}.$$

(b) Since $u(z) \leq e^{\rho z}$, $\rho \geq 0$ for all z

$$\Pr\{z \geq 0\} \leq E\{e^{\rho z}\} , \ \rho \geq 0$$

(c) Note that

$$P_E(m \to m') = \Pr\{z(\underset{\sim}{y}|x_m) \geq 0\}$$

$$\leq E_{\underset{\sim}{y}}\{e^{\rho z(\underset{\sim}{y}|\underset{\sim}{x}_m)}\}$$

$$= \sum_{\underset{\sim}{y}} e^{\rho z(\underset{\sim}{y}|\underset{\sim}{x}_m)} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)$$

$$= \sum_{\underset{\sim}{y}} \left[P_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})/P_N(\underset{\sim}{y}|\underset{\sim}{x}_M)\right]^{\rho} P_N(\underset{\sim}{y}\ \underset{\sim}{x}_m)$$

$$= \sum_{\mathbf{y}} P_N(\mathbf{y}|\mathbf{x}_{m'})^{\rho}\, P_N(\mathbf{y}|\mathbf{x}_m)^{1-\rho} \ , \ \rho \geq 0 \ .$$

(d) Let $\rho = \frac{1}{2} = 1 - \rho$ in (c) and we have

$$P_E(m \to m') \leq \sum_{\mathbf{y}} \sqrt{P_N(\mathbf{y}|\mathbf{x}_{m'})\, P_N(\mathbf{y}|\mathbf{x}_m)}$$

(e) For $\mathbf{x}_m = 00 \ldots 0$ and $\mathbf{x}_{m'} = 11 \ldots 1$ we have

$$P_N(\mathbf{y}|\mathbf{x}_{m'}) = \begin{cases} 1 \ , & \mathbf{y} = 11 \ldots 1 \\ 0 \ , & \text{elsewhere} \end{cases}$$

$$P_N(\mathbf{y}|\mathbf{x}_m) = p^{w(\mathbf{y})}(1-p)^{N-w(\mathbf{y})}$$

where $w(\mathbf{y})$ = # of "1" in $\mathbf{y}$. Hence the Chernoff bound is

$$P_E(m \to m') \leq p^{N(1-\rho)} \quad , \ \rho \geq 0$$

The bound is minimized by maximizing $(1-\rho)$ or letting $\rho = 0$. Then $P_E(m \to m') \leq p^N$. The maximum likelihood rule is simply: choose $\mathbf{x}_m$ if $\mathbf{y} = 11 \ldots 1$; otherwise choose $\mathbf{x}_{m'}$. Then an error occurs only when $\mathbf{x}_m$ is transmitted and $\mathbf{y} = 11 \ldots 1$ is the output which occurs with probability $P_E(m \to m') = p^N$. The Bhattacharyya bound is

$$P_E(m \to m') \leq \sqrt{p^N} = p^{\frac{N}{2}} .$$

2.11 (a) If $\mathbf{H}^T = \begin{bmatrix} \mathbf{P} \\ \mathbf{I}_{L-K} \end{bmatrix}$ then $\mathbf{G} = [\mathbf{I}_K \ \mathbf{P}]$.

Here

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \implies \mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(b) Rows of $\underset{\sim}{H}^T$ consist of all 2^{L-K} possible binary sequences except the all zero sequence. The rows of $\underset{\sim}{H}^T$ having more than a single "1" form the rows of the matrix $\underset{\sim}{P}$. Then $\underset{\sim}{G} = [\underset{\sim}{I}_K \; \underset{\sim}{P}]$ is the generator matrix.

(c) Let $\underset{\sim}{h}_i$ be the i^{th} row of $\underset{\sim}{H}^T$. Then $\underset{\sim}{v} = (v_1 \; v_2 \; \ldots, \; v_L)$ is a codeword if $\underset{\sim}{v}\underset{\sim}{H}^T = \underset{\sim}{0}$ or $\sum_{n=1}^{L} v_n \underset{\sim}{h}_i = \underset{\sim}{0}$. Since no two rows of $\underset{\sim}{H}^T$ are the same, $\underset{\sim}{v}$ must have weight of at least 3. Since the sum of any two row vectors of $\underset{\sim}{H}^T$ form a non zero vector which must also be a row vector of H^T we see that there are 3 rows that sum (modulo-2) to $\underset{\sim}{0}$. Hence $d_{min} = 3$.

2.12 (a) If $\mathcal{E}_b$ is the energy per bit then for K bits there is a total of $\mathcal{E}_b K$ units of energy. If $L = 2^K$ binary symbols per codeword is used then each binary symbol has energy $\mathcal{E}_s = \mathcal{E}_b K/L$. The BSC formed by hard quantizing an AWGN has transition probability

$$P = Q\left(\sqrt{\frac{2\mathcal{E}_s}{N_o}}\right) = Q\left(\sqrt{\frac{2\mathcal{E}_b K}{N_o L}}\right)$$

$$\text{Let } \varepsilon_i = \begin{cases} 1; & \text{if an error occurs in the } i^{th} \text{ symbol} \\ 0; & \text{if no error occurs in the } i^{th} \text{ symbol} \end{cases}$$

Then $E\{\varepsilon_i\} = p$, $E\{\varepsilon_i^2\} = p$, and $\text{Var}\{\varepsilon_i\} = p - p^2 = p(1-p)$.

Since $\{\varepsilon_i\}$ are independent for

$$\eta = \sum_{i=1}^{L} \varepsilon_i$$

we have

$$E\{\eta\} = \sum_{i=1}^{L} E\{\varepsilon_i\} = Lp$$

and

$$\text{Var}\ \{\eta\} = \sum_{i=1}^{L} \text{Var}\ \{\varepsilon_i\} = Lp(1-p)\ .$$

(b) Keeping $\mathscr{E}_b$ fixed, as K increases $\mathscr{E}_s = \mathscr{E}_b K/L$ decreases. For small x we have from

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{t^2}{2}}\, dt$$

and

$$\frac{d}{dx}\, Q(x) = -\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$$

The expansion approximation

$$Q(x) \cong Q(0) + \frac{d}{dx}\, Q(x)\Big|_{x=0} x$$

or

$$Q(x) \cong \tfrac{1}{2} - \frac{1}{\sqrt{2\pi}}\, x$$

Then

$$E\{\eta\} = LQ\left(\sqrt{\frac{2\mathscr{E}_b K}{N_o 2^K}}\right) \sim L\left(\tfrac{1}{2} - \sqrt{\frac{\mathscr{E}_b K}{\pi N_o 2^K}}\right) \sim \frac{L}{2}$$

and

$$\begin{aligned}\text{Var}\ \{\eta\} &= LQ\left(\sqrt{\frac{2\mathscr{E}_b K}{N_o 2^K}}\right)\left[1 - Q\left(\sqrt{\frac{2\mathscr{E}_b K}{N_o 2^K}}\right)\right] \\ &\sim L\left(\tfrac{1}{2} - \sqrt{\frac{\mathscr{E}_b K}{\pi N_o 2^K}}\right)\left(\tfrac{1}{2} + \sqrt{\frac{\mathscr{E}_b K}{\pi N_o 2^K}}\right) \\ &\sim \frac{L}{4}\ .\end{aligned}$$

(c) (2.10.14) can be expressed as

$$P_E \le \begin{cases} \Pr\{\eta \ge (d_{min} + 1)/2\} & , \quad d_{min} \text{ odd} \\ \Pr\{\eta \ge d_{min}/2\} & , \quad d_{min} \text{ even} \end{cases}$$

or

$$P_E \le \Pr\{\eta \ge d_{min}/2\} \quad , \quad d_{min} \text{ odd or even.}$$

Using $d_{min} = L/2$ for the binary orthogonal codes of (a) we have $P_E \le \Pr\{\eta \ge L/4\}$.

(d) The Chebyshev inequality gives

$$\Pr\{|\eta - E\{\eta\}| \ge \varepsilon\} \le \frac{\mathrm{Var}\{\eta\}}{\varepsilon^2}$$

or for large K,

$$\Pr\{|\eta - \tfrac{L}{2}| \ge L/4\} \le \frac{L/4}{(L/4)^2} = \frac{4}{L} = 2^{-(K-2)}$$

Hence

$$\begin{aligned} \Pr\{\eta \le L/4\} &= \Pr\{\eta - L/2 \le - L/4\} \\ &\le \Pr\{|\eta - \tfrac{L}{2}| \ge \tfrac{L}{4}\} \\ &\le 2^{-(K-2)} \xrightarrow[K\to\infty]{} 0 \end{aligned}$$

Hence

$$\Pr\{\eta \ge L/4\} \xrightarrow[K\to\infty]{} 1 \quad .$$

<u>2.13</u> (a) Let $x = \sqrt{\dfrac{2\mathscr{E}_s}{N_o}}$

$$P_o(-2) = \int_{-\infty}^{-a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y-\sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy$$

$$\int_{a}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y+\sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy = P_1(2)$$

$$= \int_{a+x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = Q(a + x)$$

$$P_o(-1) = \int_{-a}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y - \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy$$

$$= \int_{0}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y' + \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy' = P_1(+1)$$

$$= \int_{x}^{a+x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du -$$

$$\int_{a+x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

$$= Q(x) - Q(x + a)$$

$$P_o(+1) = \int_{0}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y - \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy$$

$$= \int_{-a}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y' + \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy' = P_1(-1)$$

$$= \int_{x-a}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \int_{x-a}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du -$$

$$\int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = Q(x - a) - Q(x)$$

$$P_o(+2) = \int_{a}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y - \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy$$

$$= \int_{-\infty}^{-a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(y' + \sqrt{\frac{2\mathscr{E}_s}{N_o}}\right)^2} dy' = P_1(-2)$$

$$= \int_{-\infty}^{x-a} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = 1 - Q(x-a)$$

(b) $d = -\ln 2[\sqrt{P_o(2)P_o(-2)} + \sqrt{P_o(1)P_o(-1)}]$

$$= -\ln 2\{\sqrt{[1-Q(x-a)]Q(x+a)} + \sqrt{[Q(x-a)-Q(x)][Q(x)-Q(x+a)]}\}$$

For $\mathcal{E}_s/N_o = 2$ we have $x = 2$. We find "a" to minimize

$$f(a) = \sqrt{[1-Q(2-a)]Q(2+a)} + \sqrt{[Q(2-a)-Q(2)][Q(2)-Q(2+a)]}$$

by solving $\frac{df(a)}{da} = 0$.

<u>2.14</u> (a) Since $x_1, x_2, \ldots, x_M$ are orthogonal the noise in the 2M observables $y_{1c}, y_{1s}, y_{2c}, y_{2s}, \ldots, y_{Mc}, Y_{Ms}$ are independent. Since we have noncoherent detection let $y_m = \sqrt{(\frac{2}{N_o})(y_{mc}^2 + y_{ms}^2)}$ $m = 1, 2, \ldots, M$. Assuming x_1 is sent then

$$P(y_1|x_1) = y_1 \exp[-(y_1^2/2) - (\mathcal{E}/N_o)] \cdot I_o(\sqrt{2\mathcal{E}/N_o}\; y_1)$$

and

$$P(y_m|x_1) = y_m \exp(-y_m^2/2) \quad m = 2, \ldots, M$$

The maximum likelihood decision rule is simply: choose $x_{\hat{m}}$ that corresponding to the largest y_m $m = 1, 2, \ldots, M$. Hence

$$P_{E_1} = 1 - \Pr\{y_1 > y_m \quad \text{for all } m \neq 1|x_1\}$$

$$= 1 - \int_0^\infty p(y_1|x_1) \Pr\{y_m < y_1 \text{ for all } m \neq 1|x_1\} dy_1$$

$$= 1 - \int_0^\infty p(y_1|x_1) \left[\int_0^{y_1} p(y_2|x_1)\, dy_2\right]^{M-1} dy_1$$

$$= \int_0^\infty p(y_1|x_1) \left\{ 1 - \left[\int_0^{y_1} p(y_2|x_1)dy_2 \right]^{M-1} \right\} dy_1$$

(b) Here

$$\int_0^{y_1} p(y_2|x_1)dy_2 = \int_0^{y_1} y_2 e^{-y_2^2/2} dy_2 = 1 - e^{-y_1^2/2}$$

(c) Using the binomial expansion

$$\left(1 - e^{-y_1^2/2}\right)^{M-1} = \sum_{j=0}^{M-1} (-1)^j \binom{M-1}{j} e^{-jy_1^2/2} ,$$

$$1 - \left(1 - e^{-y_1^2/2}\right)^{M-1} = \sum_{j=1}^{M-1} (-1)^{j+1} \binom{M-1}{j} e^{-jy_1^2/2}$$

and the integral

$$\int_0^\infty y_1 e^{-(j+1)y_1^2/2} I_o(\sqrt{2R/N_o}\, y_1) dy_1 = \frac{1}{j+1} e^{\mathcal{E}/(j+1)N_o}$$

we get

$$p_E = p_{E_1} = e^{-\mathcal{E}/N_o} \sum_{j=1}^{M-1} (-1)^{j+1} \binom{M-1}{j} \frac{1}{j+1} e^{\mathcal{E}/(j+1)N_o}$$

$$= e^{-\mathcal{E}/N_o} \sum_{j=2}^{M} \frac{(-1)^j}{M} \binom{M}{j} e^{\mathcal{E}/jN_o}$$

2.15 (a) Define events $A_m = \{y_m \geq y\}$ $m = 2, 3, \ldots, M$.

Then

$$\Pr\{y_m \geq y \text{ for some } m \neq 1 | x_1\}$$

$$= \Pr\{\bigcup_{m=2}^{M} A_m | x_1\} = 1 - \Pr\{\bigcap_{m=2}^{M} \bar{A}_m | x_1\}$$

$$= 1 - \prod_{m=2}^{M} \Pr\{\bar{A}_m | x_1\} = 1 - \left(1 - e^{-y^2/2}\right)^{M-1}$$

Certainly $\Pr\{\bigcup_{m=2}^{M} A_m | x_1\} \leq 1$ and also

$$\Pr\{\bigcup_{m=2}^{M} A_m | x_1\} \leq \sum_{m=2}^{M} \Pr\{A_m | x_1\}$$

$$= (M - 1)\, e^{-y^2/2}$$

Hence

$$1 - \left(1 - e^{-y^2/2}\right)^{M-1} \leq \min\left[(M - 1)e^{-y^2/2}, 1\right] \leq 1$$

and for $0 \leq \rho \leq 1$

$$\min\left[(M - 1)e^{-y^2/2}, 1\right] \leq \left\{\min\left[(M - 1)e^{-y^2/2}, 1\right]\right\}^{\rho}$$

$$\leq \left\{(M - 1)e^{-y^2/2}\right\}^{\rho}$$

(b) Note that

$$P_E = \int_0^{\infty} p(y_1 | x_1) \Pr\{y_m \geq y_1 \text{ for some } m \neq 1 | x_1\} dy_1$$

$$= \int_0^{\infty} p(y_1 | x_1) \left\{1 - \left(1 - e^{-y^2/2}\right)^{M-1}\right\} dy_1$$

$$\leq \int_0^\infty p(y_1|x_1)(M-1)^\rho e^{-\rho y_1^2/2}\, dy_1$$

$$= (M-1)^\rho\, e^{-\mathcal{E}/N_o} \int_0^\infty y_1 e^{-(1+\rho)y_1^2/2}\, I_o(\sqrt{2\mathcal{E}/N_o}\, y_1)\, dy_1$$

$$= (M-1)^\rho\, e^{-\mathcal{E}/N_o}\, \frac{1}{1+\rho}\, e^{\mathcal{E}/(1+\rho)N_o}$$

$$\leq (M-1)^\rho \exp\left[-\frac{\mathcal{E}}{N_o}\left(\frac{\rho}{1+\rho}\right)\right].$$

(c) Although the signal phase is unknown, it is <u>constant</u> during the signaling time. As signaling time increases our knowledge of the unknown phase increases and it can be estimated with increased accuracy. The maximum likelihood detector incorporates this knowledge so that in the limit of large signal energy (and corresponding signal duration) the noncoherent performance approaches the coherent performance.

<u>2.16</u> (a) Add the first row of $\underset{\sim}{G}$ to the second row to get the standard form

$$\underset{\sim}{G}' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} = [\underset{\sim}{I}\ \underset{\sim}{P}]$$

Hence

$$\underset{\sim}{H}^T = \begin{bmatrix} \underset{\sim}{P} \\ \underset{\sim}{I} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is a Hamming code.

(b) $\underset{\sim}{s} = \underset{\sim}{y}\underset{\sim}{H}^T = (1\ 1\ 0)$ implies an error in the 3rd position since $\underset{\sim}{s}$ is the third row of $\underset{\sim}{H}$. Hence $\underset{\sim}{v} = (1\ 1\ 1\ 1\ 1\ 1\ 1)$.

(c) $\underset{\sim}{s} = \underset{\sim}{y}\underset{\sim}{H}^T = (0\ 0\ 1)$ implies an error in the 7th position. Hence $\underset{\sim}{v} = (1\ 0\ 0\ 1\ 1\ 0\ 0)$.

(d) $d_{min} = 3$ for Hamming codes. See problem 2.11 (c).

2.17 (a) If $x_m(t)$ is sent then $\lambda_1, \lambda_2, \ldots, \lambda_M$ are independent Gaussian random variables where λ_m has mean b and unit variance. All the other normalized observables have zero mean and unit variance. Hence

$$\Pr(\text{erasure}) = \Pr\{\lambda_{m'} \le \delta \quad \text{all } m\}$$

$$= \prod_{m'=1}^{M} \Pr\{\lambda_{m'} \le \delta\}$$

$$= \Pr\{\lambda_m \le \delta\} \prod_{m' \ne m} \Pr\{\lambda_{m'} \le \delta\}$$

$$= Q(b - \delta)[1 - Q(\delta)]^{M-1} \quad .$$

(b) By symmetry assume x_m is sent and Pr{correct decision} = $\Pr\{\lambda_m > \lambda_{m'} \text{ for all } m' \ne m,\ \lambda_m > \delta | x_m\}$.

$$= \int_{-\infty}^{\infty} \Pr\{\lambda_{m'} < \alpha \text{ for all } m' \ne m,\ \lambda_m > \delta | x_m, \lambda_m = \alpha\} p_{\lambda_m}(\alpha) d\alpha$$

$$= \int_{\delta}^{\infty} \Pr\{\lambda_{m'} < \alpha \quad \text{for all } m' \ne m | x_m\} p_{\lambda_m}(\alpha) d\alpha$$

$$= \int_{\delta}^{\infty} [1 - Q(\alpha)]^{M-1} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\alpha-b)^2}{2}} d\alpha$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\delta-b}^{\infty} [1 - Q(x + b)]^{M-1} e^{-x^2/2} dx$$

2.18 Form observable $y = \int_0^T y(t)\phi(t)dt$. Then without loss in optimality we have

$$H_o : y = n \qquad x,n \sim N(0,1)$$
$$H_1 : y = x + n$$

and

$$p_o(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} , \quad p_1(y) = \frac{1}{\sqrt{4\pi}} e^{-y^2/4} .$$

Since y is a Gaussian random variable with zero mean and unit variance if H_o and variance 2 if H_1. The decision region of H_o is given by

$$\Lambda_o = \{y : p_1(y) \le p_o(y)\}$$
$$= \{y : \frac{1}{\sqrt{2}} e^{-y^2/4} \le e^{-y^2/2}\}$$
$$= \{y : y^2 \le \ln 4\} = \{y : -\sqrt{\ln 4} \le y \le \sqrt{\ln 4}\} .$$

while $\Lambda_1 = \overline{\Lambda}_o = \{y : y < -\sqrt{\ln 4}$ and $y > \sqrt{\ln 4}\}$.

Here we have error probabilities,

$$P_{E_1} = Pr\{y \in \Lambda_o | H_1\} = 1 - Pr\{y \in \Lambda_1 | H_1\}$$

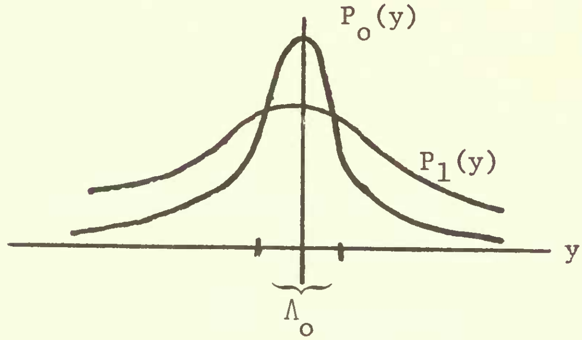
$$= 1 - 2\Pr\{y > \sqrt{\ln 4}\,|\,H_1\} = 1 - 2Q(\sqrt{\ln 2}) \quad .$$

and

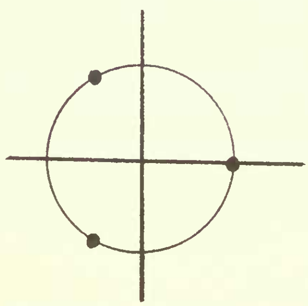
$$P_{E_o} = \Pr\{y \in \Lambda_1 | H_o\} = 2\Pr\{y \geq \sqrt{\ln 4}\,|\,H_o\}$$

$$= 2Q(\sqrt{\ln 4}) \quad .$$

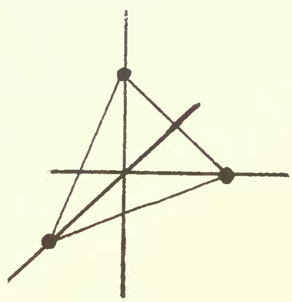
Hence $$P_E = \tfrac{1}{2}P_{E_o} + \tfrac{1}{2}P_{E_1} = \tfrac{1}{2} + Q(\sqrt{\ln 4}) - Q(\sqrt{\ln 2}) \quad .$$



2.19



Simplex



Orthogonal

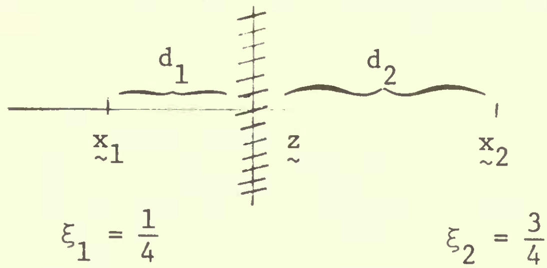
(a) The signal set is a Simplex signal set with M = 3 and energy $\mathscr{E}$. This is merely a translation of an orthogonal signal set of energy $\mathscr{E}' = \frac{3}{2}\mathscr{E}$. Hence we can use the error probability expression for orthogonal signals given by (2.11.1), to get the desired result.

(b) Since $\xi_2 = 0$, there are only two signals of equal probability. Hence the error probability is

$$P_E = Q\left(\frac{\| \underline{x}_1 - \underline{x}_o \|}{\sqrt{2N_o}}\right) = Q\left(\sqrt{\frac{3\mathscr{E}}{2N_o}}\right)$$

since $\| \underline{x}_1 - \underline{x}_o \|^2 = 3\mathscr{E}$.

(c) Here we have two signals $\underline{x}_1$ and $\underline{x}_2$ with priori probabilities $\xi_1 = \frac{1}{4}$ and $\xi_2 = \frac{3}{4}$. Consider distances between these signals and the decision boundary,



The point $\underset{\sim}{z}$ between $\underline{x}_1$ and $\underline{x}_2$ must satisfy

$$\xi_1 P_2(\underset{\sim}{z} | \underset{\sim}{x}_1) = \xi_2 P_2(\underset{\sim}{z} | \underset{\sim}{x}_2)$$

or

$$\| \underset{\sim}{z} - \underset{\sim}{x}_2 \|^2 - \| \underset{\sim}{z} - \underset{\sim}{x}_1 \|^2 = N_o \ell n 3$$

or
$$d_2^2 - d_1^2 = N_o \ell n^3$$

and also
$$d_1 + d_2 = \| \underset{\sim}{x}_2 - \underset{\sim}{x}_1 \| = \sqrt{3\mathscr{E}} \quad .$$

Thus

$$d_2 = \frac{\sqrt{3\mathscr{E}}}{2} + \frac{N_o \ell n 3}{2\sqrt{3\mathscr{E}}} , \quad d_1 = \frac{\sqrt{3\mathscr{E}}}{2} - \frac{N_o \ell n 3}{2\sqrt{3\mathscr{E}}} \quad .$$

and

$$P_E = \frac{1}{4} Q\left(\sqrt{\frac{3\mathscr{E}}{2N_o}} - \sqrt{\frac{N_o}{6\mathscr{E}}}\, \ell n 3\right) + \frac{3}{4} Q\left(\sqrt{\frac{3\mathscr{E}}{2N_o}} + \sqrt{\frac{N_o}{6\mathscr{E}}}\, \ell n 3\right) .$$

2.20 Let
$$n(t) = \sum_{k=1}^{m} n_k \psi_k(t) + n_r(t)$$

where
$$n_k = \int_0^T n(t)\psi_k(t)dt \qquad k = 1,2,\ldots, m$$

and
$$n_r(t) = n(t) - \sum_{k=1}^{m} n_k \psi_k(t) .$$

Then we see that
$$\int_0^T n_r(t)\psi_k(t)dt = 0 \qquad k = 1, \ldots, m$$

and
$$E\{n_i n_j\} = E\left\{\int_0^T\int_0^T n(\alpha)n(\beta)\psi_i(\alpha)\psi_j(\beta)\, d\alpha\, d\beta\right\}$$
$$= \int_0^T\int_0^T \phi(\alpha,\beta)\psi_i(\alpha)\psi_j(\beta)\, d\alpha\, d\beta$$
$$= \int_0^T \sigma_j^2 \psi_j(\alpha)\psi_i(\alpha)d\alpha$$
$$= \sigma_j^2 \delta_{ji}$$

Also
$$E\{n_i n_r(t)\} = E\left\{n_i \left(n(t) - \sum_{k=1}^{m} n_k \psi_k(t)\right)\right\}$$
$$= E\{n_i n(t)\} - \sum_{k=1}^{m} E\{n_i n_k\}\psi_k(t)$$

$$= E\left\{\int_0^T n(\alpha)\psi_i(\alpha)d\alpha \cdot n(t)\right\} - \sigma_i^2\psi_i(t)$$

$$= \int_0^T \phi(t,\alpha)\psi_i(\alpha)d\alpha - \sigma_i^2\psi_i(t)$$

$$= \sigma_i^2\psi_i(t) - \sigma_i^2\psi_i(t) = 0 \quad \text{all } t \quad .$$

Hence the processes $\sum_{k=1}^{m} n_k\psi_k(t)$ and $n_r(t)$ are independent of each other and there is no loss in optimality by considering only the observables,

$$y_k = \int_0^T y(t)\psi_k(t)dt$$

$$= \begin{cases} \sqrt{\mathcal{E}_k} + n_k & \text{if} \quad x_1(t) \quad \text{is sent} \\ -\sqrt{\mathcal{E}_k} + n_k & \text{if} \quad x_2(t) \quad \text{is sent} . \end{cases}$$

$$k = 1,2,\ldots, m .$$

(a) For $\underset{\sim}{y} = (y_1, y_2, \ldots, y_m)$ we have

$$P_m(\underset{\sim}{y}|\underset{\sim}{x}_1) = \prod_{k=1}^{m} \frac{1}{\sqrt{2\pi\sigma_k{}^2}} e^{-\frac{\left(y_k-\sqrt{\mathcal{E}_k}\right)^2}{2\sigma_k^2}}$$

and $$P_m(\underset{\sim}{y}|\underset{\sim}{x}_2) = \prod_{k=1}^{m} \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{\left(y+\sqrt{\mathcal{E}_k}\right)^2}{2\sigma_k^2}}$$

Hence

$$\Lambda_1 = \{\underset{\sim}{y} : P_m(\underset{\sim}{y}|\underset{\sim}{x}_1) > P_m(\underset{\sim}{y}|\underset{\sim}{x}_2)\}$$

$$= \left\{\underset{\sim}{y} : \sum_{k=1}^{m} \frac{(y_k-\sqrt{\mathcal{E}_k})^2}{2\sigma_k^2} < \sum_{k=1}^{m} \frac{(y_k+\sqrt{\mathcal{E}_k})^2}{2\sigma_k^2}\right\}$$

$$= \left\{ \underset{\sim}{y} : \sum_{k=1}^{m} y_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2} > - \sum_{k=1}^{m} y_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2} \right\}$$

$$= \left\{ \underset{\sim}{y} : \sum_{k=1}^{m} y_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2} > 0 \right\}$$

(b) By symmetry $P_E = P_{E_1} = P_{E_2}$. For $x_1(t)$ sent we have

$$\sum_{k=1}^{m} y_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2} = \sum_{k=1}^{m} \frac{\mathcal{E}_k}{\sigma_k^2} + \underbrace{\sum_{k=1}^{m} n_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2}}_{N}$$

where

$$E\{N\} = 0 \ , \quad E\{N^2\} = \sum_{k=1}^{m} \frac{\mathcal{E}_k}{\sigma_k^2} \ .$$

Hence

$$P_{E_1} = \Pr \left\{ \sum_{k=1}^{m} y_k \frac{\sqrt{\mathcal{E}_k}}{\sigma_k^2} \le 0 \middle| H_1 \right\}$$

$$= \Pr \left\{ N \le - \sum_{k=1}^{m} \frac{\mathcal{E}_k}{\sigma_k^2} \right\} = Q\left(\sqrt{\sum_{k=1}^{m} \frac{\mathcal{E}_k}{\sigma_k^2}} \right) .$$

For $\sigma_k^2 = \frac{N_o}{2}$ and $\mathcal{E} = \sum_{k=1}^{m} \mathcal{E}_k$ we have

$$P_E = Q\left(\sqrt{\frac{2\mathcal{E}}{N_o}} \right) .$$

2.21 (a) The union of events bound for any signal set over the AWGN channel is (assuming $\underset{\sim}{x}_1$ is sent)

$$P_{E_1} \le \sum_{m=2}^{M} Q\left(\frac{\| \underset{\sim}{x}_m - \underset{\sim}{x}_1 \|}{\sqrt{2N_o}} \right) .$$

But here

$$\| \underset{\sim}{x}_m - \underset{\sim}{x}_1 \|^2 = \begin{cases} 2\mathscr{E}, & k \neq 1 \text{ and } k \neq \frac{M}{2} + 1 \\ 4\mathscr{E}, & k = \frac{M}{2} + 1 \end{cases}$$

Hence

$$P_{E_1} \leq (M - 2)Q\left(\sqrt{\frac{\mathscr{E}}{N_o}}\right) + Q\left(\sqrt{\frac{2\mathscr{E}}{N_o}}\right)$$

$$\leq (M - 1)Q\left(\sqrt{\frac{\mathscr{E}}{N_o}}\right)$$

which is the "Union of events" bound for an orthogonal code.

(b) Since $\{|y_m| \geq A$ for more than one $m\}$ and $\{|y_m| < A$ for all $m\}$ are disjoint events

$$P_r = \Pr\{|y_m| \geq A \text{ for more than one } m\} + \Pr\{|y_m| < A \text{ for all } m\} .$$

Without loss in generality, suppose $x_1(t)$ is sent. Then

$y_1 = \sqrt{\mathscr{E}} + n_1$ and $y_m = n_m \quad m = 2, 3, \ldots, M/2,$

and

$$\Pr\{|y_m| \geq A \text{ for more than one } m\}$$

$$\leq \Pr\{|y_m| \geq A \text{ for some } m \neq 1\}$$

$$\leq \sum_{m=2}^{M/2} \Pr\{|y_m| \geq A\} \leq \frac{M}{2} \cdot 2Q\left(\sqrt{\frac{2}{N_o}}\, A\right) = MP_2 \quad .$$

Also

$$\Pr\{|y_m| < A \text{ for all } m\} \leq \Pr\{|y_1| < A\}$$

$$= 1 - 2Q\left(\sqrt{\frac{2}{N_o}}\,(A - \sqrt{\mathscr{E}})\right)$$

$$\leq 1 - Q\left(\sqrt{\frac{2}{N_o}}\,(A - \sqrt{\mathscr{E}})\right)$$

$$= P_1 \quad .$$

Hence $P_r \le P_1 + MP_2$.

Next define events

$$e_1 = \{y_1 < -A, \ |y_m| < A \ \text{ all } m \ne 1\}$$

$$e_k = \{|y_k| \ge A, \ |y_m| < A \ \text{ all } m \ne k\}$$

$$k = 2, 3, \ldots, M/2 .$$

Since for $k \ne 1$

$$e_k \subset \{|y_k| \ge A, \ y_1 < A\}$$

we have

$$\Pr\{e_k\} \le \Pr\{|y_k| \ge A, \ y_1 < A\}$$

$$= \Pr\{|y_k| \ge A\} \Pr\{y_1 < A\} = 2P_1P_2 .$$

Since $\| \underset{\sim}{x}_{\frac{M}{2}+1} - \underset{\sim}{x}_1 \|^2 = 4\mathscr{E} > \| \underset{\sim}{x}_k - \underset{\sim}{x}_1 \|^2 = 2\mathscr{E} \quad k \ne \frac{M}{2} + 1$

we see that the error event e_1 must be less probable than any other error event e_k $k \ne 1$.

Hence

$$\Pr\{e_1\} \le \Pr\{e_k\} \le 2P_1P_2$$

and

$$P_E = P_{E_1} = \Pr\left\{\bigcup_{k=1}^{M/2} e_k\right\} \le \sum_{k=1}^{M/2} \Pr\{e_k\} \le MP_1P_2 .$$

<u>2.22</u> (a) $$\int_{-\infty}^{\infty} \phi_{2n}(t)\phi_{2n+1}(t)dt = \int_{(n-\frac{1}{2})\tau}^{n\tau} \frac{2}{\tau} \sin \omega_0 t \cos \omega_0 t \, dt$$

$$= \frac{1}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \sin 2\omega_0 t \, dt$$

$$= -\frac{1}{2\omega_0\tau} \cos 2\omega_0 t \Big|_{(n-\frac{1}{2})\tau}^{n\tau}$$

$$= \frac{1}{2\omega_0\tau} [\cos 2\omega_0(n - \tfrac{1}{2})\tau - \cos 2\omega_0 n\tau]$$

$$= 0$$

Thus $\{\phi_k(t)\}$ are orthonormal functions. The performance would then be the same as binary modulation of any set of orthonormal functions such as those for QPSK modulation given by Table 2.1 (b). Since $\phi_k(t)$ is merely the same as QPSK signals only staggered, the spectrum is the same. It is the spectrum of the square wave f(t) shifted in frequency by $\pm\, \omega_0$.

(b) $$y(t) = \sum_k x_k\phi_k(t) = \sum_k [x_{2k}\phi_{2k}(t) + x_{2k+1}\phi_{2k+1}(t)]$$

$$y'_{2n} = \int_{(n-1)\tau}^{n\tau} y(t) \sqrt{\frac{2}{\tau}} \sin (\omega_0 t + \phi)dt$$

$$= x_{2n-1} \frac{2}{\tau} \int_{(n-1)\tau}^{(n-\frac{1}{2})\tau} \cos \omega_0 t \sin(\omega_0 t + \phi)dt$$

$$+ x_{2n} \frac{2}{\tau} \int_{(n-1)\tau}^{n\tau} \sin \omega_0 t \sin (\omega_0 t + \phi)dt$$

$$+ x_{2n+1} \frac{2}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \cos \omega_0 t \sin(\omega_0 t + \phi)dt$$

$$= \tfrac{1}{2} x_{2n-1} \sin \phi + x_{2n} \cos \phi + \tfrac{1}{2} x_{2n+1} \sin \phi$$

$$y'_{2n+1} = \int_{(n-\frac{1}{2})\tau}^{(n+\frac{1}{2})\tau} y(t) \sqrt{\frac{2}{\tau}} \cos(\omega_0 t + \phi)dt$$

$$= x_{2n} \frac{2}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \sin \omega_0 t \cos (\omega_0 t + \phi) dt$$

$$+ x_{2n+1} \frac{2}{\tau} \int_{(n-\frac{1}{2})\tau}^{(n+\frac{1}{2})\tau} \cos \omega_0 t \cos(\omega_0 t + \phi) dt$$

$$+ x_{2(n+1)} \frac{2}{\tau} \int_{n\tau}^{(n+\frac{1}{2})\tau} \sin \omega_0 t \cos(\omega_0 t + \phi) dt$$

$$= - \tfrac{1}{2} x_{2n} \sin \phi + x_{2n+1} \cos \phi - \tfrac{1}{2} x_{2(n+1)} \sin \phi$$

Note that

$$y^1_{2n} = \begin{cases} x_{2n} \cos \phi & , \text{ if } x_{2n-1} = -x_{2n+1} \\ x_{2n-1} \sin \phi + x_{2n} \cos \phi & , \text{ if } x_{2n-1} = x_{2n+1} \end{cases}$$

(c) $$\int_{-\infty}^{\infty} \phi_{2n}(t)\ \phi_{2n+1}(t) dt$$

$$= \frac{4}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \cos \left[\frac{\pi}{\tau}\left(t-(n-\tfrac{1}{2})\tau\right)\right] \cos\left[\frac{\pi}{\tau}(t-n\tau)\right]$$

$$\sin \omega_0 t \cos \omega_0 t dt$$

$$= \frac{4}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \left\{\tfrac{1}{2} \cos \frac{\pi}{2} + \tfrac{1}{2} \cos \left[\frac{\pi}{\tau}(2t-2n\tau+\tfrac{1}{2}\tau)\right]\right\}$$

$$\cdot \tfrac{1}{2} \sin 2\omega_0 t dt$$

$$= \frac{1}{\tau} \int_{(n-\frac{1}{2})\tau}^{n\tau} \cos \left[\frac{\pi}{\tau}(2t + \tfrac{1}{2}\tau)\right] \sin 2\omega_0 t dt$$

$$= -\frac{1}{\tau}\int_{(n-\frac{1}{2})\tau}^{n\tau} \sin(2\pi t/\tau)\sin 2\omega_0 t\,dt$$

$$= -\frac{1}{2\tau}\int_{(n-\frac{1}{2})\tau}^{n\tau} [\cos(2\omega_0 - 2\pi/\tau)t - \cos(2\omega_0 + 2\pi/\tau)t]dt$$

$$= \frac{1}{2\tau}\left\{-\frac{\sin(2\omega_0-2\pi/\tau)t}{(2\omega_0-2\pi/\tau)} + \frac{\sin(2\omega_0+2\pi/\tau)t}{(2\omega_0+2\pi/\tau)}\right\}\Bigg|_{(n-\frac{1}{2})\tau}^{n\tau}$$

$$= 0$$

Again the signals $\{\phi_k(t)\}$ are orthonorm̃l and binary modulation of orthonormal waveforms give the same performance over the AWGN channel.

(d) For $(n - \frac{1}{2})\tau \le t < n\tau$ we have

$$x_{2n}\phi_{2n}(t) + x_{2n+1}\phi_{2n+1}(t)$$

$$= x_{2n}\frac{2}{\sqrt{\tau}}\cos\left[\frac{\pi}{\tau}\left(t - (n - \tfrac{1}{2})\tau\right)\right]\sin\omega_0 t$$

$$+ x_{2n+1}\frac{2}{\sqrt{\tau}}\cos\left[\frac{\pi}{\tau}(t - n\tau)\right]\cos\omega_0 t$$

But

$$\cos\left[\frac{\pi}{\tau}\left(t - (n - \tfrac{1}{2})\tau\right)\right] = \cos\left[\frac{\pi}{\tau}(t - n\tau) + \frac{\pi}{2}\right]$$
$$= -\sin\left[\frac{\pi}{\tau}(t - n\tau)\right]$$

Hence

$$x_{2n}\phi_{2n}(t) + x_{2n+1}\phi_{2n+1}(t)$$

$$= -x_{2n}\frac{2}{\sqrt{\tau}}\sin\left[\frac{\pi}{\tau}(t - n\tau)\right]\sin\omega_0 t$$
$$+ x_{2n+1}\frac{2}{\sqrt{\tau}}\cos\left[\frac{\pi}{\tau}(t - n\tau)\right]\cos\omega_0 t$$

$$= \nu(t) \cos \left[\omega_0 t + \mathscr{N}(t)\right]$$

where

$$\nu(t) = \sqrt{\left(-x_{2n}\frac{2}{\sqrt{\tau}}\sin\left[\frac{\pi}{\tau}(t-n\tau)\right]\right)^2 + \left(x_{2n+1}\frac{2}{\sqrt{\tau}}\cos\left[\frac{\pi}{\tau}(t-n\tau)\right]\right)^2}$$

$$= 2/\sqrt{\tau}$$

and

$$\mathscr{N}(t) = \tan^{-1}\left(\frac{-x_{2n}\frac{2}{\sqrt{\tau}}\sin\left[\frac{\pi}{\tau}(t - n\tau)\right]}{x_{2n+1}\frac{2}{\sqrt{\tau}}\cos\left[\frac{\pi}{\tau}(t - n\tau)\right]}\right)$$

$$= \tan^{-1}\left(-\frac{x_{2n}}{x_{2n+1}}\tan\left[\frac{\pi}{\tau}(t - n\tau)\right]\right)$$

$$= \pm\ (\pi/\tau)(t - n\tau)$$

$$= \pm\ (\pi/\tau)t \mp n\pi$$

Hence

$$x_{2n}\phi_{2n}(t) + x_{2n+1}\phi_{2n+1}(t)$$

$$= (2/\sqrt{\tau})\cos\left[\omega_0 t \pm (\pi/\tau)t \mp n\pi\right]$$

$$= \pm\ (2/\sqrt{\tau})\cos\left[(\omega_0 \pm \pi/\tau)t\right] .$$

(e)

$$f(t) = \begin{cases} \sqrt{2}\sin(\pi t/\tau) & , \quad -\tau/2 \le t < \tau/2 \\ 0 & , \quad \text{otherwise} \end{cases}$$

has spectrum

$$\left(\frac{2\pi}{\tau}\right)^2 \left[\frac{\cos(\omega\tau/2)}{\omega^2 - (\pi/\tau)^2}\right]^2$$

—//—

Chapter 3

3.1 $E_0(1,\mathbf{q}) = -\ln \sum_y \left(\sum_x q(x)\sqrt{p(y|x)} \right)^2$

(a) $E_0(1,\mathbf{q}) = -\ln \left[\left(q_1\sqrt{1-p} + q_2\sqrt{p} \right)^2 + \left(q_1\sqrt{p} + q_2\sqrt{1-p} \right)^2 \right]$

$$= -\ln \left[q_1^2 + q_2^2 + 4q_1q_2\sqrt{p(1-p)} \right]$$

By symmetry, $q_1 = q_2 = \frac{1}{2}$ maximizes. Hence

$$E_0(1) = \max_{\mathbf{q}} E_0(1,\mathbf{q}) = -\ln\left(\frac{1+2\sqrt{p(1-p)}}{2}\right)$$

$$= \ln 2 - \ln\left(1+2\sqrt{p(1-p)}\right)$$

(b) $E_0(1,\mathbf{q}) = -\ln\left[\left(q_1\sqrt{1-p}\right)^2 + \left(q_1\sqrt{p} + q_2\sqrt{p}\right)^2 + \left(q_2\sqrt{1-p}\right)^2 \right]$

$$= -\ln\left[p + (1-2q_1q_2)(1-p) \right]$$

By symmetry $q_1 = q_2 = \frac{1}{2}$ maximizes and

$$\max_{\mathbf{q}} E_0(1,\mathbf{q}) = -\ln\left(\frac{1+p}{2}\right) = \ln 2 - \ln(1+p)$$

(c) $E_0(1,\mathbf{q}) = -\ln\left[\left(q_1 + q_2\sqrt{p}\right)^2 + \left(q_2\sqrt{1-p}\right)^2 \right]$

$$= -\ln\left[1 - 2q_1q_2\left(1-\sqrt{p}\right) \right]$$

Although the channel is not symmetric, $E_0(1,\mathbf{q})$ is symmetric in q_1 and q_2. Hence $q_1 = q_2 = \frac{1}{2}$ maximizes $E_0(1,\mathbf{q})$.

$$\max E_0(1,\mathbf{q}) = -\ln\left[1 - \tfrac{1}{2}\left(1-\sqrt{p}\right)\right] = -\ln\left(\frac{1+\sqrt{p}}{2}\right)$$

$$= \ln 2 - \ln\left(1+\sqrt{p}\right)$$

3.2 (a)

(i) By symmetry $q = (1/3, 1/3, 1/3)$. Hence

$$E_0(\rho) = -\ln \sum_y \left(\sum_x \frac{1}{3} \, p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} = -\ln\left[3\left(\frac{1}{3}\right)^{1+\rho}\left(2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right]$$

$$= \rho \ln \frac{3}{2} \Longrightarrow C = \left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=0} = \ln \frac{3}{2} .$$

(ii) By symmetry $q = (1/4, 1/4, 1/4, 1/4)$. Hence

$$E_0(\rho) = -\ln \sum_y \left(\sum_x \frac{1}{4} \, p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} = -\ln\left[4\left(\frac{1}{4}\right)^{1+\rho}\left(2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right]$$

$$= \rho \ln 2 \Longrightarrow C = \left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=0} = \ln 2$$

(b)

(i) $E(R) = \max\limits_{0 \le \rho \le 1} \left[\rho \ln \frac{3}{2} - \rho R \right] = \ln \frac{3}{2} - R = C - R$

(ii) $E(R) = \max\limits_{0 \le \rho \le 1} \left[\rho \ln 2 - \rho R \right] = \ln 2 - R = C - R$

3.3 (a) By symmetry $q_k = \frac{1}{Q}$ $\quad k = 1, 2, \ldots, Q.$

Then

$$E_0(\rho) = -\ln \sum_y \left(\sum_x \frac{1}{Q} \, p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} = -\ln \sum_y \left(\frac{1}{Q}\right)^{1+\rho}$$

$$= \rho \ln Q \Longrightarrow C = \ln Q$$

and

$$E(R) = \max_{0 \le \rho \le 1} \left[\rho \ln Q - \rho R \right] = \ln Q - R = C - R$$

(b) By symmetry $q_k = \frac{1}{Q}$ $\quad k = 1, 2, \ldots, Q$.

Let $p(y|x) = p + (\bar{p} - p)\,\delta_{yx}$. Then

$$E_o(\rho) = -\ln \sum_y \left(\sum_x \frac{1}{Q}\, p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

$$= -\ln \sum_y \left(\sum_x \frac{1}{Q} \left(p + (\bar{p}-p)\,\delta_{yx} \right)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

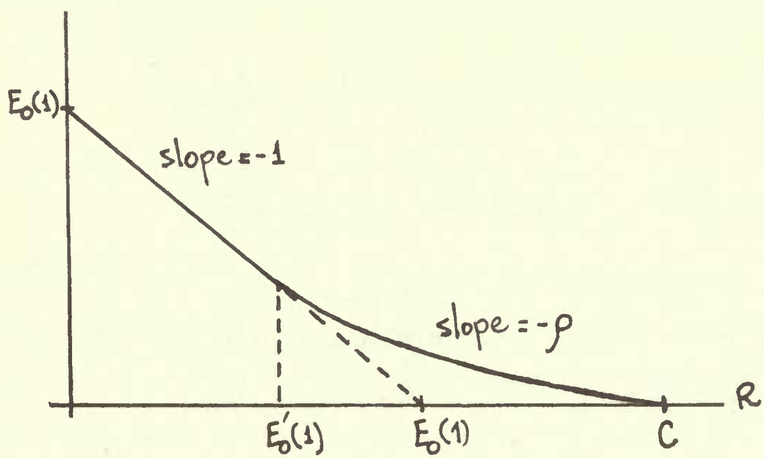
$$= -\ln Q \left(\frac{1}{Q} \right)^{1+\rho} \left[\bar{p}^{\frac{1}{1+\rho}} + (Q-1)\, p^{\frac{1}{1+\rho}} \right]^{1+\rho}$$

$$= \rho \ln Q - (1+\rho) \ln \left[(Q-1)\, p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right]$$

$$E_o'(\rho) = \ln Q - \ln \left[(Q-1)\, p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right] + \frac{(Q-1)\, p^{\frac{1}{1+\rho}} \ln p + \bar{p}^{\frac{1}{1+\rho}} \ln \bar{p}}{(1+\rho) \left[(Q-1)\, p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right]}$$

$$E_o(1) = \ln Q - 2 \ln \left[(Q-1)\sqrt{p} + \sqrt{\bar{p}} \right]$$

$$C = E_o'(0) = \ln Q + (Q-1)\, p \ln p + \bar{p} \ln \bar{p}$$



(c) By symmetry we have $q_1 = q_2 = q_3 = q_4$ and $q_5 = q_6$.

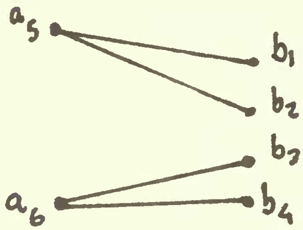
Consider (3.2.21)

$$\alpha(y, q) \equiv \sum_x q(x)\, p(y|x)^{\frac{1}{1+\rho}}$$

$$= q_1\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right] + q_5(.5)^{\frac{1}{1+\rho}}, \text{ all } y.$$

We use this to check conditions for the optimum choice of q.

(i) Let $q_1 = q_2 = q_3 = q_4 = 0$, $q_5 = q_6 = 1/2$



$$\alpha(y, q) = \frac{1}{2}(.5)^{\frac{1}{1+\rho}} = \left(\frac{1}{2}\right)^{\frac{2+\rho}{1+\rho}}$$

To check condition (3.2.23) let for $x = a_5$ and a_6

$$\sum_y p(y|x)^{\frac{1}{1+\rho}}\, \alpha(y, q)^\rho = 2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}}\left(\frac{1}{2}\right)^{\frac{(2+\rho)\rho}{1+\rho}} = \left(\frac{1}{2}\right)^\rho$$

For $x = a_1, a_2, a_3$ and a_4 we have

$$\sum_y p(y|x)^{\frac{1}{1+\rho}}\, \alpha(y, q)^\rho = \left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]\left(\frac{1}{2}\right)^{\frac{(2+\rho)\rho}{1+\rho}}$$

Checking for ρ where (3.2.23) is satisfied gives

$$\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]\left(\frac{1}{2}\right)^{\frac{(2+\rho)\rho}{1+\rho}} \geq \left(\frac{1}{2}\right)^\rho$$

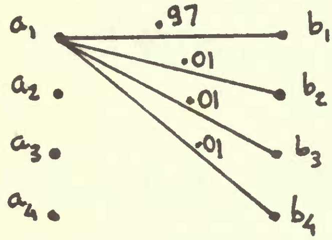
or $$(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}} \geq \left(\frac{1}{2}\right)^{\frac{-\rho}{1+\rho}}$$

which is satisfied for $\rho \geq \rho_0 = 1.9$.

Hence for $\rho \geq \rho_0$ the optimum choice of q is

$q_1 = q_2 = q_3 = q_4 = 0$, $q_5 = q_6 = 1/2$.

(ii) Let $q_1 = q_2 = q_3 = q_4 = 1/4$, $q_5 = q_6 = 0$



(similar connections from a_2, a_3 & a_4 not shown for simplicity)

$$\alpha(y,q) = \frac{1}{4}\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]$$

For $x = a_1, a_2, a_3$ and a_4 we have

$$\sum_y p(y|x)^{\frac{1}{1+\rho}}\,\alpha(y,q)^\rho = \left(\frac{1}{4}\right)^\rho \left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

and for $x = a_5, a_6$

$$\sum_y p(y|x)^{\frac{1}{1+\rho}}\,\alpha(y,q)^\rho = 2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}}\left(\frac{1}{4}\right)^\rho\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]^\rho$$

For (3.2.23) to hold we require

$$2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}}\left(\frac{1}{4}\right)^\rho\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]^\rho \geq \left(\frac{1}{4}\right)^\rho\left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}\right]^{\rho+1}$$

or

$$\left(\frac{1}{2}\right)^{\frac{-\rho}{1+\rho}} \geq (.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}}$$

which is the opposite condition from that in (i) and thus is satisfied for $\rho \leq \rho_0 = 1.9$.

Hence for $\rho \leq \rho_0$ the optimum choice of q is

$$q_1 = q_2 = q_3 = q_4 = 1/4 \quad \text{and} \quad q_5 = q_6 = 0 .$$

We conclude that for $0 \leq \rho \leq 1.9$ we use the channel in (ii) and for $1.9 \leq \rho < \infty$ we use the channel in (i). Since for $E(R)$ we maximize for $0 \leq \rho \leq 1$ we use only the channel in (ii) to get from (b),

$$E(R) = \max_{0 \le \rho \le 1} \left[E_0(\rho, q) - \rho R \right]$$

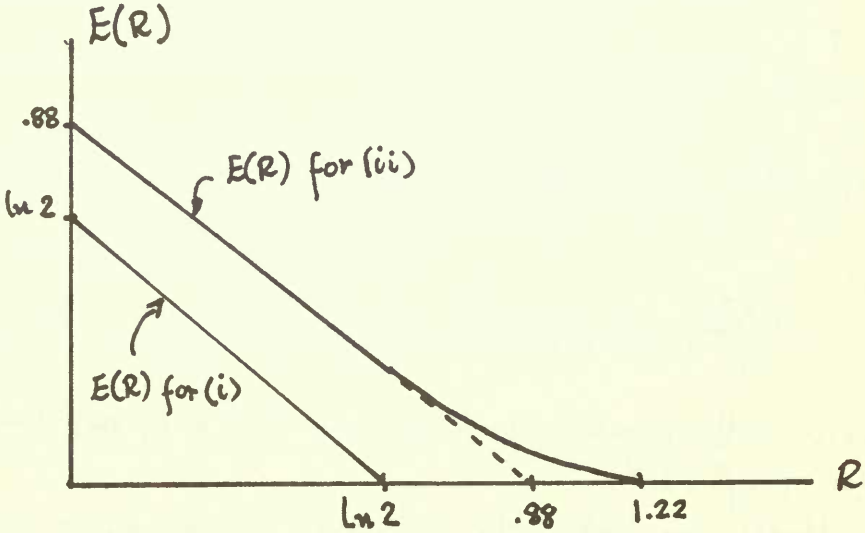
$$= \max_{0 \le \rho \le 1} \left\{ \rho \ln 4 - (1+\rho) \ln \left[(.97)^{\frac{1}{1+\rho}} + 3(.01)^{\frac{1}{1+\rho}} \right] - \rho R \right\}$$

and

$$C = \ln 4 + 3(.01) \ln(.01) + (.97) \ln(.97) = 1.22$$

Note: if $\bar{p} = .97$ is changed to a smaller value then $\rho_0 < 1$ is possible and both channels in (i) and (ii) will be used to evaluate $E(R)$ for R between 0 and C.

A sketch of $E(R)$ for cases (i) and (ii) is shown below,



<u>3.4</u> (a) From 3.2.23 we have the condition

$$\sum_y p(y|x)^{\frac{1}{1+\rho}} \alpha(y, q)^{\rho} \ge \sum_y \alpha(y, q)^{1+\rho} \quad , \text{for all } x$$

where

$$\alpha(y, q) = \sum_x q(x) \, p(y/x)^{\frac{1}{1+\rho}}$$

Equality happens when $q(x) > 0$.

We have

$$E_0(\rho, q) = -\ln \sum_y \left(\sum_x q(x) p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

$$= -\ln \sum_y \alpha(y, q)^{1+\rho}$$

and thus

$$\exp\{-E_0(\rho, q)\} = \sum_y \alpha(y, q)^{1+\rho}$$

For the maximizing value of q let $\alpha(y) = \alpha(y, q)$ and $E_0(\rho) = E_0(\rho, q)$. Then for each output

$$\alpha(b_j) = \sum_x q(x)\, p(b_j|x)^{\frac{1}{1+\rho}}$$

$$= \left(P_{j1}^{\frac{1}{1+\rho}}, P_{j2}^{\frac{1}{1+\rho}}, \dots, P_{jQ}^{\frac{1}{1+\rho}} \right) q^T$$

and

$$\begin{bmatrix} \alpha(b_1) \\ \alpha(b_2) \\ \vdots \\ \alpha(b_J) \end{bmatrix} = \begin{bmatrix} P_{11}^{\frac{1}{1+\rho}} & \cdots\cdots & P_{1Q}^{\frac{1}{1+\rho}} \\ \vdots & & \\ P_{J1}^{\frac{1}{1+\rho}} & \cdots\cdots & P_{JQ}^{\frac{1}{1+\rho}} \end{bmatrix} q^T$$

or $\quad \underline{\alpha}^T = \left[P_{jk}^{\frac{1}{1+\rho}} \right] q^T$

Now note that

$$\sum_j P_{jk}^{\frac{1}{1+\rho}} \alpha(b_j)^\rho = \underline{\alpha}^\rho \begin{bmatrix} P_{1k}^{\frac{1}{1+\rho}} \\ \vdots \\ P_{Jk}^{\frac{1}{1+\rho}} \end{bmatrix}$$

for each k which is the vector $\underline{\alpha}^\rho = (\alpha(b_1)^\rho, \dots, \alpha(b_J)^\rho)$ multiplied by the k^{th} column of $\left[P_{jk}^{1/1+\rho} \right]$.

The condition (3.2.23) then becomes

$$\underset{\sim}{\alpha}^{\rho}\left[P_{jk}^{\frac{1}{1+\rho}}\right] \geq \exp\{-E_0(\rho)\}\, \underset{\sim}{u} \qquad (1)$$

with equality for those components where $q(k)>0$.

(b) $q_k > 0$ for all k means (1) holds with equality. $\left[P_{jk}^{1/1+\rho}\right]$ is a nonsingular square matrix ($J=Q$). Hence

$$\underline{\alpha}^{\rho}\left[P_{jk}^{\frac{1}{1+\rho}}\right] = e^{-E_0(\rho)}\, \underset{\sim}{u}$$

or

$$e^{E_0(\rho)}\, \underset{\sim}{\alpha}^{\rho} = \underset{\sim}{u}\left[P_{jk}^{1/1+\rho}\right]^{-1} \triangleq \underset{\sim}{\xi}^{\rho}$$

where

$$\underset{\sim}{\xi}^{\rho} = (\xi_1^{\rho}, \dots, \xi_J^{\rho}) \quad \text{and} \quad \xi = (\xi_1, \dots, \xi_J)$$

Then $\quad e^{E_0(\rho)/\rho}\, \underset{\sim}{\alpha}^{T} = \underset{\sim}{\xi}^{T}$

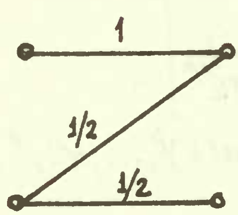
or $\quad e^{E_0(\rho)/\rho}\left[P_{jk}^{1/1+\rho}\right] \underset{\sim}{q}^{T} = \underset{\sim}{\xi}^{T}$

or $\quad \underset{\sim}{q}^{T} = e^{-E_0(\rho)/\rho}\left[P_{jk}^{1/1+\rho}\right]^{-1} \underset{\sim}{\xi}^{T}$

Next $\underset{\sim}{u}\,\underset{\sim}{q}^{T} = 1$ means $1 = e^{-E_0(\rho)/\rho}\, \underset{\sim}{u}\left[P_{jk}^{\frac{1}{1+\rho}}\right]^{-1} \underset{\sim}{\xi}^{T}$

or $\quad E_0(\rho) = \rho \ln\left\{\underline{u}\left[P_{jk}^{1/1+\rho}\right]^{-1} \underset{\sim}{\xi}^{T}\right\}$

<u>3.5</u> (a)



$$\rightarrow \quad \left[P_{jk}^{1/1+\rho}\right] = \begin{bmatrix} 1 & (1/2)^{1/1+\rho} \\ 0 & (1/2)^{1/1+\rho} \end{bmatrix}$$

$$\therefore \quad \left[P_{jk}^{\frac{1}{1+\rho}}\right]^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 2^{\frac{1}{1+\rho}} \end{bmatrix}$$

$$\underset{\sim}{\xi}^{\rho} = \underset{\sim}{u}\left[P_{jk}^{\frac{1}{1+\rho}}\right]^{-1} = [1 \ \ 1]\begin{bmatrix} 1 & -1 \\ 0 & 2^{1/1+\rho} \end{bmatrix} = \left[1 \ \ 2^{\frac{1}{1+\rho}}-1\right]$$

and

$$\underset{\sim}{\xi} = \left[1 \ \ \left(2^{\frac{1}{1+\rho}}-1\right)^{1/\rho}\right]$$

Thus, $E_o(\rho) = \rho \ln\left\{\left[1 \ \ 2^{\frac{1}{1+\rho}}-1\right]\begin{bmatrix} 1 \\ \left(2^{\frac{1}{1+\rho}}-1\right)^{1/\rho}\end{bmatrix}\right\}$

$$= \rho \ln\left\{1+\left(2^{\frac{1}{1+\rho}}-1\right)^{\frac{1+\rho}{\rho}}\right\}$$

(b) $\underset{\sim}{q}^{T} = e^{-E_o(\rho)/\rho}\left[P_{jk}^{\frac{1}{1+\rho}}\right]^{-1}\underset{\sim}{\xi}^{T}$

$$= \left[1+\left(2^{1/1+\rho}-1\right)^{\frac{1+\rho}{\rho}}\right]^{-1}\begin{bmatrix} 1 & -1 \\ 0 & 2^{\frac{1}{1+\rho}} \end{bmatrix}\begin{bmatrix} 1 \\ 2^{1/1+\rho}-1 \end{bmatrix}$$

$$= \left[1+\left(2^{\frac{1}{1+\rho}}-1\right)^{\frac{1+\rho}{\rho}}\right]^{-1}\begin{bmatrix} 1-\left(2^{1/1+\rho}-1\right)^{1/\rho} \\ \left(2^{\frac{1}{1+\rho}}-1\right)^{1/\rho}\end{bmatrix}$$

which is a function of ρ and thus varies with R.
For $\rho=0$ we have

$$\underset{\sim}{q}^{T}\Big|_{\rho=0} = \frac{1}{5/4}\begin{bmatrix} 1-1/4 \\ 2(1/4)\end{bmatrix} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$

since $\lim\limits_{\rho\to 0}\left(2^{\frac{1}{1+\rho}}-1\right)^{1/\rho} = 1/4$

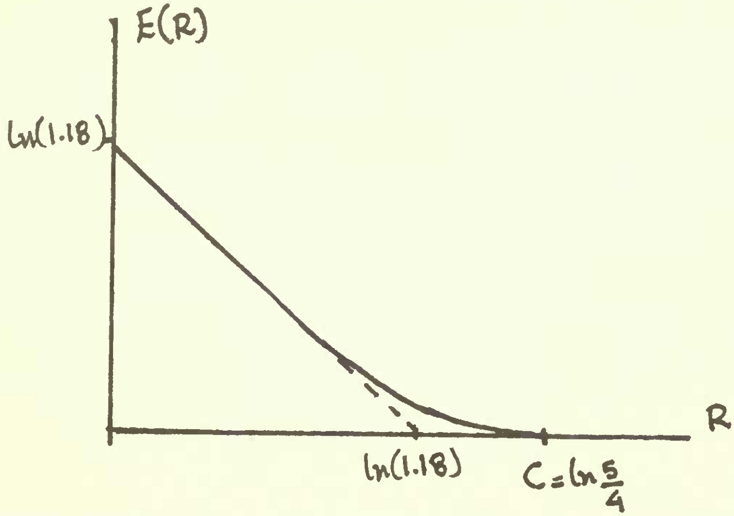
And for $\rho=1$ we get

$$q^T\Big|_{\rho=1} = \frac{1}{1+(\sqrt{2}-1)^2}\begin{bmatrix} 1-(\sqrt{2}-1) \\ \sqrt{2}(\sqrt{2}-1)\end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

(c) $$C = \lim_{\rho\to 0} E_0'(\rho) = \lim_{\rho\to 0} \frac{E_0(\rho)-E_0(0)}{\rho - 0} = \lim_{\rho\to 0}\frac{E_0(\rho)}{\rho}$$

$$= \lim_{\rho\to 0} \ln\left[1+\left(2^{\frac{1}{1+\rho}}-1\right)^{1+1/\rho}\right] = \ln\left[1+1/4\right] = \ln 5/4 .$$

(d) $E_0(1) = \ln\left[1+(\sqrt{2}-1)^2\right] = \ln(1.18)$



3.6 (a) By symmetry $q^*(x) = \dfrac{1}{Q}$ for all x .

$$\alpha(y,q) = \begin{cases} \frac{1}{Q}\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}} & y = b_1,\dots,b_Q \\ \left(\frac{1}{2}\right)^{1/1+\rho} & y = b_{Q+1}\end{cases}$$

We check now condition (3.2.23)

$$\sum_y p(y|x)^{\frac{1}{1+\rho}} \alpha(y,q)^\rho = \frac{1}{2}(Q^{-\rho}+1) \qquad \text{for all } x$$

$$\sum_y \alpha(y,q)^{1+\rho} = \frac{1}{2}(Q^{-\rho}+1)$$ and (3.2.23) holds with equality since $q(x)>0$ for all x.

(b) $E_0(\rho) = -\ln \sum_y \alpha(y)^{1+\rho} = -\ln\left(\frac{Q^{-\rho}+1}{2}\right)$ for $q^* = \left(\frac{1}{Q},\dots,\frac{1}{Q}\right)$

$= \ln 2 - \ln(1+Q^{-\rho})$ and $E_0(1) = \ln 2 - \ln(1+Q^{-1})$

$$E_0'(\rho) = \frac{Q^{-\rho}\ln Q}{1+Q^{-\rho}} = \frac{\ln Q}{1+Q^\rho} \quad ; \quad E_0'(1) = \frac{\ln Q}{1+Q}$$

$$C = E_0'(0) = \frac{\ln Q}{2}$$

$$E(R) = \begin{cases} E_0(\rho) - \rho E_0'(\rho) & \frac{\ln Q}{1+Q} < R < C \\ R = E_0'(\rho) & \\ \ln \frac{2}{1+Q^{-1}} - R & 0 < R < \frac{\ln Q}{1+Q} \end{cases}$$

we now find $E(R)$ explicitly

$$R = E_0'(\rho) = \frac{\ln Q}{1+Q^\rho} \implies 1+Q^\rho = \frac{\ln Q}{R} \implies Q^{-\rho} = \frac{1}{1-\frac{\ln Q}{R}}$$

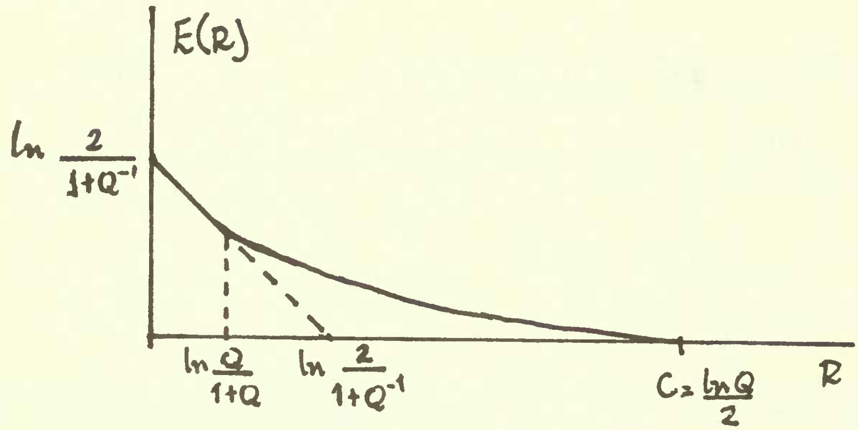
$$\implies 1+Q^{-\rho} = \frac{1}{1-R/\ln Q}$$

$$\therefore E(R) = E_0(\rho) - \rho E_0'(\rho) = \ln 2 - \ln\left(\frac{1}{1-R/\ln Q}\right) - \frac{R}{\ln Q}\ln\left(\frac{\ln Q}{R}-1\right)$$

$$= \ln 2 + \ln\left(1-\frac{R}{\ln Q}\right) - \frac{R}{\ln Q}\ln\left(\frac{1-R/\ln Q}{R/\ln Q}\right)$$

$$= \ln 2 - \mathcal{H}(R/\ln Q) \qquad \text{for } R_{cr} < R < C$$

where $\mathcal{H}(x) = -x\ln x - (1-x)\ln(1-x)$.



<u>3.7</u> (a) From problem 3.1, using (3.3.28) with $q = 1/2$

$$E_x(\rho, \mathbf{q}) = -\rho \ln\left[1 - 2q_1 q_2 (1 - z^{1/\rho})\right]$$

$$E_x(\rho) = \max_{\mathbf{q}} E_x(\rho, \mathbf{q}) = -\rho \ln\left[1 - \tfrac{1}{2}(1 - z^{1/\rho})\right]$$

$$= -\rho \ln\left[\tfrac{1}{2}(1 + z^{1/\rho})\right]$$

(note that $q = 1/2$ comes from $\max\limits_{\mathbf{q}} E_x(\rho, \mathbf{q}) \Rightarrow \max\limits_{q_1} q_1(1-q_1)$ $\Longrightarrow q_1 = q_2 = 1/2$)

where $z = \sum\limits_y \sqrt{p(y|a_1)\, p(y|a_2)}$

for channel (a) (BSC) $z = 2\sqrt{p(1-p)}$; for (b) (BEC) $z = p$

for (c) (Z channel) $z = \sqrt{p}$.

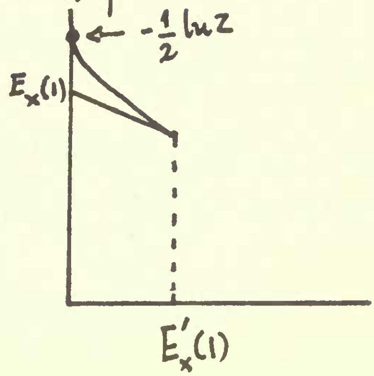
(b) $E_x'(\rho) = -\ln\left(\dfrac{1+z^{1/\rho}}{2}\right) + \dfrac{z^{1/\rho} \ln z}{\rho(1+z^{1/\rho})}$

(c) $E_{ex}(R) = E_x(\rho) - \rho \dfrac{\partial E_x(\rho)}{\partial \rho} = -\dfrac{z^{1/\rho} \ln z}{1 + z^{1/\rho}}$

$$R = -\ln\left(\frac{1+z^{1/\rho}}{2}\right) + \frac{z^{1/\rho}\ln z}{\rho(1+z^{1/\rho})} = E_x'(\rho)$$

$$E_x'(1) = -\ln\left(\frac{1+z}{2}\right) + \frac{z\ln z}{1+z}$$

$$E_{ex}(0) = \lim_{\rho\to\infty} E_{ex}(R) = -\frac{\ln z}{z}$$



3.8 (a) By symmetry choose $q(x) = \frac{1}{Q}$ for all x. Then

$$E_x(\rho) = -\rho \ln \sum_x \sum_{x'} \frac{1}{Q^2}\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^{1/\rho}$$

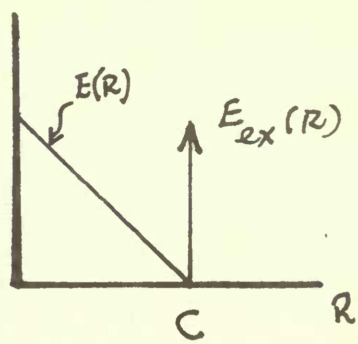
$$= -\rho \ln \frac{Q}{Q^2} = \rho \ln Q$$

$\therefore$

$$E_{ex}(R) = \sup_{1\le\rho<\infty}[\rho\ln Q - \rho R] = \sup_{1\le\rho<\infty} \rho(\ln Q - R) = \infty$$

$R < C$

$C = E_x'(0) = \ln Q$



(b) for (i), choose $\underset{\sim}{q} = (1/3, 1/3, 1/3)$

$$E_x(\rho) = E_x(\rho, \underset{\sim}{1/3}) = -\rho \ln \sum_x \sum_{x'} \frac{1}{9}\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^{1/\rho}$$

$$= -\rho\ln\left\{\frac{1}{9}\left(3 + 6\left(\frac{1}{2}\right)^{1/\rho}\right)\right\} = -\rho\ln\left\{\frac{1}{3}\left(1 + 2^{1-1/\rho}\right)\right\}$$

$\therefore$

$E_x(1) = \ln 3/2$

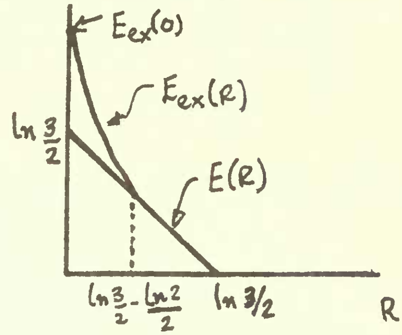
$$E_x'(\rho) = -\ln\left\{\frac{1}{3}\left(1+2^{1-1/\rho}\right)\right\} - \frac{2^{1-1/\rho}\ln 2}{1+2^{1-1/\rho}} \quad ; \quad E_x'(1) = \ln\frac{3}{2} - \frac{\ln 2}{2}$$

Now use (3.3.24) and (3.3.27)

$$E_{ex}(0) = -\sum_x \sum_{x'} \frac{1}{9} \ln\left\{\sum_y \sqrt{p(y|x)\,p(y|x')}\right\}$$

$$= -\frac{1}{9}\left(3\ln 1 + 6\ln\frac{1}{2}\right)$$

$$= \frac{2}{3}\ln 2$$



for (ii) we first try $q(x) = 1/4$ for all x. Then

$$E_x(\rho, \underset{\sim}{1/4}) = -\rho \ln \sum_x \sum_{x'} \frac{1}{16}\left(\sum_y \sqrt{p(y|x)\,p(y|x')}\right)^{1/\rho}$$

$$= -\rho\ln\left\{\frac{1}{16}\left[4(1) + 8\left(\frac{1}{2}\right)^{1/\rho} + 4(0)\right]\right\} = -\rho\ln\left(\frac{1+2^{1-1/\rho}}{4}\right)$$

$\therefore$

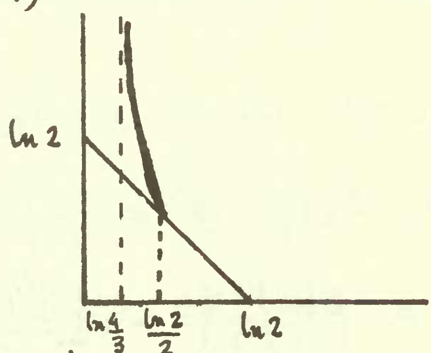
$E_x(1) = \ln 2$

$$E_x'(\rho) = -\ln\left(\frac{1+2^{1-1/\rho}}{4}\right) - \rho\left(\frac{4}{1+2^{1-1/\rho}}\right)\left(\frac{2^{1-1/\rho}}{4}\right)\frac{1}{\rho^2}\ln 2$$

$$= \ln\left(\frac{4}{1+2^{1-1/\rho}}\right) - \frac{2^{1-1/\rho}\ln 2}{\rho(1+2^{1-1/\rho})} \quad ; \quad E_x'(1) = \ln 2 - \frac{\ln 2}{2} = \frac{\ln 2}{2}$$

$$R_x(\infty) = \lim_{\rho\to\infty} E_x'(\rho)$$

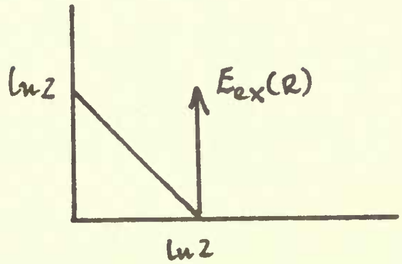
$$= \ln\frac{4}{3}$$



Now we take $q_1 = q_3 = \frac{1}{2}$ and $q_2 = q_4 = 0$. We have then

$E_x(\rho) = -\rho \ln \frac{1}{4}(2) = \rho \ln 2$ $\qquad E_x(1) = \ln 2$

$E_{ex}(R) = \sup\limits_{1 \le \rho < \infty} \rho(\ln 2 - R) = \infty \qquad R < \ln 2$



$\therefore$ the optimum distribution is

$q = (1/2, 0, 1/2, 0)$

<u>3.9</u> (a) Let $q_1 = q_2 = q_3 = q_4 = 1/4$ and check (3.2.23)

$$\alpha(y, q) = \sum_x q(x)\, p(y|x)^{\frac{1}{1+\rho}} = \frac{1}{4} \sum_x p(y|x)^{\frac{1}{1+\rho}}$$

$$\alpha(b_1) = \frac{1}{4}\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right] = \alpha(b_2) = \alpha(b_3) = \alpha(b_4)$$

$$\sum_y p(y|x)^{1/1+\rho}\, \alpha(y)^{\rho} = \left(\frac{1}{4}\right)^{\rho}\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

$$\sum_y \alpha(y)^{1+\rho} = \frac{4}{4^{1+\rho}}\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right]^{1+\rho} = \left(\frac{1}{4}\right)^{\rho}\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

and (3.2.23) checks with equality.

$$E_0(\rho) = -\ln \sum_y \alpha(y)^{1+\rho} = \rho \ln 4 - (1+\rho) \ln\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right]$$

$$E_0(1) = \ln 4 - 2 \ln\left[\sqrt{1-p} + \sqrt{p}\right]$$

$$E_0'(\rho) = \ln 4 - \ln\left[(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}\right]$$

$$- \frac{(1+\rho)\left(\frac{-1}{(1+\rho)^2}\right)\left[(1-p)^{\frac{1}{1+\rho}} \ln(1-p) + p^{\frac{1}{1+\rho}} \ln p\right]}{(1-p)^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}}$$

$C = E_0'(0) = \ln 4 - \mathcal{H}(p)$ where $\mathcal{H}(x) = -x\ln x - (1-x)\ln(1-x)$

(b) $E_x(\rho) = E_x(\rho, \underset{\sim}{\tfrac{1}{4}}) = -\rho \ln \frac{1}{16} \sum_x \sum_{x'} \left(\sum_y \sqrt{p(y|x)p(y|x')} \right)^{1/\rho}$

let $z = 2\sqrt{p(1-p)}$ then

$E_x(\rho) = -\rho \ln \frac{4}{16}(1+z^{1/\rho}) = -\rho \ln\left[\frac{1+z^{1/\rho}}{4}\right]$

$E_x(1) = \ln\left[\frac{4}{1+z}\right] = \ln 4 - \ln\left(1+2\sqrt{p(1-p)}\right) = \ln 4 - \ln\left[\sqrt{1-p}+\sqrt{p}\right]$

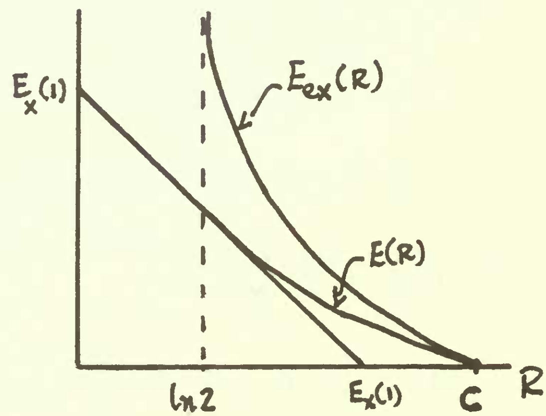
$E_x'(\rho) = -\ln\left[\frac{1+z^{1/\rho}}{4}\right] - \rho\left(\frac{4}{1+z^{1/\rho}}\right)\left(\frac{z^{1/\rho}\ln z}{4}\right)\left(-\frac{1}{\rho^2}\right)$

$= \ln\left(\frac{4}{1+z^{1/\rho}}\right) + \frac{1}{\rho}\frac{z^{1/\rho}}{1+z^{1/\rho}}\ln z$

$E_x'(1) = \ln\left(\frac{4}{1+z}\right) + \frac{z\ln z}{1+z}$

$R_x(\infty) = \lim_{\rho\to\infty} E'(\rho) = \ln 2$

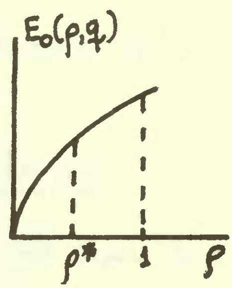
$C = \ln 4 - \mathcal{H}(p)$



3.10 (a) By the mean value theorem

$E_0(1,q) = E_0'(\rho^*, q)$ for some $\rho^* \in (0,1)$

But $E_0'(\rho^*, q) \geq E_0'(1,q)$ for $\rho^* < 1$ since $E''(\rho, q) \leq 0 \Rightarrow E_0'(\rho,q)$ decreases with ρ.



Thus $E_0(1,q) = E_0'(\rho^*,q) \geq E_0'(1,q)$

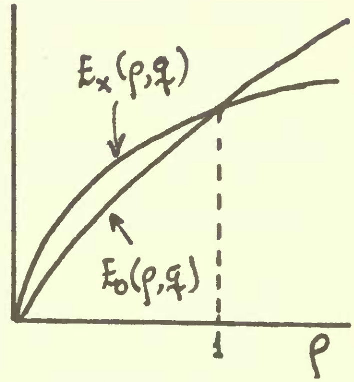
(b) Take nonpathological channel for which

$$\lim_{\rho\to\infty} E_x(\rho,q) = E_{ex}(0) < \infty \qquad \text{then}$$

$$\lim_{\rho\to\infty} E_x(\rho,q) = E_{ex}(0) \leq E_{sp}(0) = \lim_{\rho\to\infty} E_0(\rho,q)$$

then $E_x(0,q) = E_0(0,q) = 0$

and $E_x(1,q) = E_0(1,q)$



both functions are convex $\cap$ and thus cannot intersect at more than two points so

$$E_x(\rho,q) \geq E_0(\rho,q) \qquad \text{for } 0<\rho<1$$

$$E_x(\rho,q) \leq E_0(\rho,q) \qquad \text{for } 1<\rho<\infty$$

and since the functions are smooth $E_x'(1,q) \leq E_0'(1,q)$

<u>3.11</u> (a) $E_x(\rho,q) = -\rho \ln \sum_x \sum_{x'} q(x)q(x')\alpha_{xx'}$

then $f(q) = \exp\left\{-\dfrac{E_x(\rho,q)}{\rho}\right\}$

By assumption $\alpha_{xx'}$ is a nonnegative definite $Q \times Q$ matrix and so, by definition

$$\sum_x \sum_{x'} q(x)\alpha_{xx'}q(x') \geq 0$$

By appropriate linear transformation $\tilde{q}(x) = \mathcal{L}(q(x))$ we may diagonalize this quadratic form to obtain

$$f(q) = \sum_x \sum_{x'} q(x)\, \alpha_{xx'}\, q(x') = \sum_x (\tilde{q}(x))^2 \lambda_x$$

where $\lambda_1, \lambda_2, \ldots, \lambda_Q$ are the nonnegative eigenvalues of $\alpha_{xx'}$. But $f(q)$ is the sum of Q convex $\cup$ functions and so $f(q)$ is a convex $\cup$ function of $\tilde{q}$. Note that $\tilde{q}$ is a linear function of q and a convex function of a linear function of q is another convex function of q. Therefore $f(q)$ is convex $\cup$ in q.

(b) Since $f(q) = \sum_x \sum_{x'} q(x) q(x') \alpha_{xx'}$ satisfies the condition of the theorem in Appendix 3B, applying (3B.1) and (3B.2)

$$\left.\frac{\partial f(q)}{\partial q_k}\right|_{q=q^0} = \sum_{x'} q(x') \left(\sum_y \sqrt{p(y|a_k)\, p(y|x')} \right)^{1/\rho} \leq \lambda$$

for all $k = 1, 2, \ldots, Q$

with equality if $q_k > 0$.

Multiplying by $q(a_k)$ and summing over all k

$$\lambda = \sum_x \sum_{x'} q(x) q(x') \alpha_{xx'} = \exp\left\{ \frac{-E_x(\rho, q^0)}{\rho} \right\}$$

therefore we have

$$\sum_{x'} q^0(x') \left(\sum_y \sqrt{p(y|x)\, p(y|x')} \right)^{1/\rho} \leq e^{-\frac{E_x(\rho, q^0)}{\rho}}$$

with equality if $q(x) > 0$ and where q^0 is the optimizing distribution.

3.12 (a) By (3.4.16)

$$C(AWGN) = -\frac{1}{2}\ln 2\pi e - \int_{-\infty}^{\infty} w(y)\ln w(y)\,dy$$

where

$$w(y) = \frac{1}{2}\left[p_0(y) + p_0(-y)\right] \text{ and } p_0(y) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(y-\sqrt{\frac{2\mathcal{E}_s}{N_0}}\right)^2}$$

if we assume $\frac{\mathcal{E}_s}{N_0} << 1$ we have

$$w(y) = \frac{1}{2} e^{-\frac{\mathcal{E}_s}{N_0}} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\left[e^{y\sqrt{\frac{2\mathcal{E}_s}{N_0}}} + e^{-y\sqrt{\frac{2\mathcal{E}_s}{N_0}}}\right]$$

$$\cong e^{-\frac{\mathcal{E}_s}{N_0}} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\left[1 + \frac{\mathcal{E}_s}{N_0} y^2 + \dots\right]$$

and using $\ln(1+x) \cong x$ for small x,

$$\ln w(y) \cong -\frac{\mathcal{E}_s}{N_0} - \frac{y^2}{2} - \frac{1}{2}\ln 2\pi + \ln\left(1 + \frac{\mathcal{E}_s}{N_0} y\right)$$

$$\cong -\frac{\mathcal{E}_s}{N_0} - \frac{y^2}{2}\left(1 - \frac{2\mathcal{E}_s}{N_0}\right) - \frac{1}{2}\ln 2\pi$$

Then,

$$\int_{-\infty}^{\infty} w(y)\ln w(y)\,dy \cong e^{-\frac{\mathcal{E}_s}{N_0}} \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} e^{-y^2/2}\left(1 + \frac{\mathcal{E}_s}{N_0} y^2\right)\left[-\frac{\mathcal{E}_s}{N_0} - \frac{y^2}{2}\left(1 - \frac{2\mathcal{E}_s}{N_0}\right) - \frac{\ln 2\pi}{2}\right]$$

$$\cong \left(1 - \frac{\mathcal{E}_s}{N_0}\right)\int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\left[-\frac{\mathcal{E}_s}{N_0} - \frac{y^2}{2} + \frac{\mathcal{E}_s}{N_0} y^2 - \frac{\ln 2\pi}{2} - \frac{\mathcal{E}_s}{2N_0} y^4 - \frac{\ln 2\pi}{2}\frac{\mathcal{E}_s}{N_0} y^2\right]$$

$$\cong \left(1 - \frac{\mathcal{E}_s}{N_0}\right)\left[-\frac{\mathcal{E}_s}{N_0} - \frac{1}{2} + \frac{\mathcal{E}_s}{N_0} - \frac{1}{2}\ln 2\pi - \frac{3}{2}\frac{\mathcal{E}_s}{N_0} - \frac{1}{2}\frac{\mathcal{E}_s}{N_0}\ln 2\pi\right]$$

$$\cong -\frac{1}{2}\ln 2\pi e - \frac{\mathcal{E}_s}{N_0}$$

, we neglect above $\left(\frac{\mathcal{E}_s}{N_0}\right)^n$ for $n > 1$.

$\therefore$ $C(AWGN) \cong \mathcal{E}_s/N_0$ nats for $\mathcal{E}_s/N_0 << 1$ as in (3.4.19).

By (3.4.18) $C(BSC) = \ln 2 - \mathcal{H}(p)$ where $p = \text{erfc}\sqrt{\frac{2\mathcal{E}_s}{N_0}}$.

Assuming $\mathcal{E}_s/N_0 << 1$ then,

$$p = \int_{\sqrt{\frac{2\mathcal{E}_s}{N_0}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2} - \int_0^{\sqrt{\frac{2\mathcal{E}_s}{N_0}}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \cong \frac{1}{2} - \sqrt{\frac{\mathcal{E}_s}{\pi N_0}} \triangleq \frac{1}{2} - \epsilon$$

$$\therefore \; C(BSC) \cong \ln 2 - \mathcal{H}\left(\tfrac{1}{2} - \epsilon\right)$$

Now consider $\mathcal{H}(x) = -x \ln x - (1-x)\ln(1-x)$ and not that

$\mathcal{H}'(x) = -\ln x + \ln(1-x)$, $\mathcal{H}''(x) = -\frac{1}{x} - \frac{1}{1-x}$

Using a Taylor series expansion of $\mathcal{H}(x)$ around $\frac{1}{2}$ we get

$$\mathcal{H}(1/2 - \epsilon) \cong \mathcal{H}(1/2) + \mathcal{H}'(1/2)(-\epsilon) + \mathcal{H}''(1/2)\frac{(-\epsilon)^2}{2!} + \ldots$$

$$\cong \ln 2 + 0(-\epsilon) - 4\frac{\epsilon^2}{2} = \ln 2 - 2\epsilon^2$$

$$= \ln 2 - \left(\frac{2}{\pi}\right)\frac{\mathcal{E}_s}{N_0}$$

Hence

$$C(BSC) \cong \left(\frac{2}{\pi}\right)\frac{\mathcal{E}_s}{N_0} \quad \text{for } \frac{\mathcal{E}_s}{N_0} << 1 \quad \text{as in (3.4.20)}$$

(b) By (3.4.22) for $\frac{\mathcal{E}_s}{N_0} << 1$ we have

$E_{ex}(0) \cong E(0) \cong C/2$. Using the result in part (a):

for AWGN

$$\lim_{\frac{\mathcal{E}_s}{N_0} \to 0} \frac{E_{ex}(0)}{\mathcal{E}_s/N_0} \cong \frac{\frac{1}{2}\mathcal{E}_s/N_0}{\mathcal{E}_s/N_0} = \frac{1}{2} \quad \text{and} \quad \lim_{\frac{\mathcal{E}_s}{N_0} \to 0} \frac{C}{\mathcal{E}_s/N_0} \cong \frac{\mathcal{E}_s/N_0}{\mathcal{E}_s/N_0} = 1$$

for BSC

$$\lim_{\mathcal{E}_s/N_0\to 0}\frac{E_{ex}(0)}{\mathcal{E}_s/N_0}\cong\frac{\frac{1}{2}\frac{2}{\pi}\mathcal{E}_s/N_0}{\mathcal{E}_s/N_0}=\frac{1}{\pi}\quad\text{and}\quad\lim_{\mathcal{E}_s/N_0\to 0}\frac{C}{\mathcal{E}_s/N_0}\simeq\frac{\frac{2}{\pi}\mathcal{E}_s/N_0}{\mathcal{E}_s/N_0}=\frac{2}{\pi}$$

now as $\frac{\mathcal{E}_s}{N_0}\to\infty$:

since the input is binary in each case $C\le\ln 2$

$\therefore\ \frac{C}{\mathcal{E}_s/N_0}\to 0$ as $\frac{\mathcal{E}_s}{N_0}\to\infty$ for the AWGN and BSC.

By (3.4.10) $E(0)=\ln 2-\ln(1+Z)$ where $\begin{cases}Z_{AWGN}=e^{-\mathcal{E}_s/N_0}\\ Z_{BSC}=\sqrt{4p(1-p)}\end{cases}$

$\therefore$ for AWGN: $\lim_{\mathcal{E}_s/N_0\to\infty}E(0)=\ln 2-\ln 1=\ln 2$

for BSC: $\lim_{\frac{\mathcal{E}_s}{N_0}\to\infty}p=0\ \therefore\ \lim_{\mathcal{E}_s/N_0\to\infty}E(0)=\ln 2$

so $\frac{E(0)}{\mathcal{E}_s/N_0}\to 0$ as $\mathcal{E}_s/N_0\to\infty$ for the AWGN and BSC.

Now for the expurgated bound, by (3.4.11) we have

$E_{ex}(0)=-\frac{1}{2}\ln Z$

for AWGN $\quad E_{ex}(0)=\frac{\mathcal{E}_s}{2N_0}\Rightarrow\frac{E_{ex}(0)}{\mathcal{E}_s/N_0}=1/2$ for all $\mathcal{E}_s/N_0$

for BSC, for large $\frac{\mathcal{E}_s}{N_0}$

$$p=\operatorname{erfc}\sqrt{\frac{2\mathcal{E}_s}{N_0}}\cong\frac{e^{-\mathcal{E}_s/N_0}}{\sqrt{4\pi\mathcal{E}_s/N_0}}\left(1-\frac{1}{\mathcal{E}_s/N_0}+\dots\right)\cong\frac{e^{-\mathcal{E}_s/N_0}}{\sqrt{4\pi\mathcal{E}_s/N_0}}\Rightarrow Z\cong\sqrt{4p}$$

$$\lim_{\mathcal{E}_s/N_0\to\infty}\frac{E_{ex}(0)}{\mathcal{E}_s/N_0}\cong-\frac{1}{2}\frac{\ln\sqrt{4p}}{\mathcal{E}_s/N_0}\cong\lim_{\mathcal{E}_s/N_0\to\infty}\left(-\frac{1}{2}\frac{\ln 2-\mathcal{E}_s/2N_0-\frac{1}{4}\ln 4\pi\mathcal{E}_s/N_0}{\mathcal{E}_s/N_0}\right)=\frac{1}{4}$$

3.13 For $\mathcal{E}_s/N_0 << 1$ and by problem 3.12 we have

$$p_0(y) = \frac{1}{\sqrt{2\pi}} e^{-\left(y-\sqrt{\frac{2\mathcal{E}_s}{N_0}}\right)^2/2} \cong \frac{e^{-y^2/2}}{\sqrt{2\pi}} e^{-\mathcal{E}_s/N_0}\left(1+\sqrt{\frac{2\mathcal{E}_s}{N_0}}\,y+\dots\right)$$

$$w(y) = \frac{1}{2}\left[p_0(y)+p_0(-y)\right] \cong e^{-\mathcal{E}_s/N_0}\frac{e^{-y^2/2}}{\sqrt{2\pi}}\left(1+\frac{\mathcal{E}_s}{N_0}y^2+\dots\right)$$

Ignoring $\left(\frac{\mathcal{E}_s}{N_0}\right)^n$ for $n>1$ we have

$$p_0(y) \cong \frac{e^{-y^2/2}}{\sqrt{2\pi}}\left(1+\sqrt{\frac{2\mathcal{E}_s}{N_0}}\,y\right) \qquad \text{and}$$

$$p_1(y) \cong \frac{e^{-y^2/2}}{\sqrt{2\pi}}\left(1-\sqrt{\frac{2\mathcal{E}_s}{N_0}}\,y\right)$$

or

$$p_0(y) = w(y)\left[1+\epsilon(x,y)\right] \quad ; \quad p_1(y) = w(y)\left[1+\epsilon(x,y)\right]$$

where

$$w(y) \cong \frac{e^{-y^2/2}}{\sqrt{2\pi}} \quad \text{and} \quad \epsilon(x,y) = \begin{cases} \sqrt{\frac{2\mathcal{E}_s}{N_0}}\,y & , x=0 \\ -\sqrt{\frac{2\mathcal{E}_s}{N_0}}\,y & , x=1 \end{cases}$$

3.14 (a) Let $p_1(y|x)$ and $p_2(y'|z)$ be the transition probabilities of the two DMC's. Since the channels are independent the transition probability of the composite channel is $p_1(y|x)p_2(y'|z)$ given the pair (x,z) is sent. Therefore for the composite channel we have

$$E_0(\rho,q) = -\ln \sum_{y,y'}\left(\sum_{x,z} q(x,z)\left[p_1(y|x)p_2(y'|z)\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}$$

where $q(x,z)$ is a probability assignment on the input pairs.

If we restrict $q(x,z) = q_1(x)q_2(z)$, where $q_1(\cdot), q_2(\cdot)$ are arbitrary input probability assignments on separate channels, then

$$E_0(\rho,q) = -\ln \sum_{y,y'} \left(\sum_x q_1(x) p_1(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho} \left(\sum_x q_2(z) p(y'|z)^{\frac{1}{1+\rho}}\right)^{1+\rho}$$

$$= E_{0_1}(\rho, q_1) + E_{0_2}(\rho, q_2)$$

where $E_{0_1}(\rho,q_1) = -\ln \sum_y \left(\sum_x q_1(x)\, p(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}$

$$E_{0_2}(\rho,q_2) = -\ln \sum_{y'} \left(\sum_z q_2(z)\, p(y'|z)^{\frac{1}{1+\rho}}\right)^{1+\rho}$$

(b) If we choose $q_1^*(x)$ to maximize $E_{0_1}(\rho,q_1)$ and $q_2^*(z)$ to maximize $E_{0_2}(\rho,q_2)$, for a given ρ, then it follows easily from (3.2.23) that $E_0(\rho,q)$ is maximized by $q^*(x,z) = q_1^*(x)\, q_2^*(z)$ and thus

$$\max_q E_0(\rho,q) = \max_{q_1} E_{0_1}(\rho,q_1) + \max_{q_2} E_{0_2}(\rho,q_2)$$

<u>3.15</u> (a) We have n disjoint channels and suppose each has input distribution q^i and now we weight the i^{th} channel by $\beta(i)$. Treating as a single channel the input distribution is $\beta(1)q_1^{(1)}, \ldots \beta(1) q_{Q_1}^{(1)}, \ldots\ldots, \beta(n) q_1^{(n)}, \ldots \beta(n) q_{Q_n}^{(n)}$

and let $q^{(i)}$ be the optimum input weighting for the i^{th} channel if it were used alone. Then

$$E_0(\rho,\underline{\beta}) = -\ln \sum_{\underline{y}} \left[\sum_{\underline{x}} q(\underline{x})\, p(\underline{y}|\underline{x})^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

$$= -\ln \Big\{ \sum_{\underline{y}\in\mathcal{Y}^{(1)}} [\beta(1)]^{1+\rho} \Big[\sum_{\underline{x}\in\mathcal{X}^{(1)}} q(\underline{x})\, p(\underline{y}|\underline{x})^{1/1+\rho}\Big]^{1+\rho} + \cdots$$

$$\cdots + \sum_{\underline{y}\in\mathcal{Y}^{(n)}} [\beta(n)]^{1+\rho} \Big[\sum_{\underline{x}\in\mathcal{X}^{(n)}} q(\underline{x})\, p(\underline{y}|\underline{x})^{\frac{1}{1+\rho}}\Big]^{1+\rho} \Big\}$$

where $\mathcal{X}^{(i)}$ and $\mathcal{Y}^{(i)}$ are the input and output spaces of channel i.

$$\therefore \ E_0(\rho,\underline{\beta}) = -\ln\left\{[\beta(1)]^{1+\rho}\exp[-E_{01}(\rho)] + \ldots + [\beta(n)]^{1+\rho}\exp[-E_{0n}(\rho)]\right\}$$

since all channels are disjoint and each $q^{(i)}$ is optimum for channel i alone.

Now

$$E_0(\rho) = \max_{\underline{\beta}} E_0(\rho,\underline{\beta}) = \max_{\underline{\beta}} -\ln\left\{\sum_{i=1}^{n}[\beta(i)]^{1+\rho}\exp[-E_{0i}(\rho)]\right\}$$

$$= -\ln\left\{\min_{\underline{\beta}} \sum_{i=1}^{n}[\beta(i)]^{1+\rho}\exp[-E_{0i}(\rho)]\right\}$$

$\beta(1),\ldots\beta(n)$ is the distribution vector

But since $E_{0i}(\rho) \geq 0$ for all i and $\rho \geq 0$, $E_0(\rho,\underline{\beta})$ is a convex $\cup$ function in $\underline{\beta}$. Consequently necessary and sufficient conditions for minimum are given by (3B.1) and (3B.2)

$$\frac{\partial}{\partial\beta(k)} \sum_{i=1}^{n}[\beta(i)]^{1+\rho}\exp[-E_{0i}(\rho)] \leq \lambda \quad \text{with equality if } \beta(k)>0$$

(for all k)

or

$$[\beta(k)]^{\rho}\exp[-E_{0k}(\rho)] \leq \frac{\lambda}{1+\rho} \qquad \text{with equality if } \beta(k)>0$$

and where

$$\sum_k \beta(k)\left\{[\beta(k)]^{\rho}\exp[-E_{0k}(\rho)]\right\} = \frac{\lambda}{1+\rho}\sum_k \beta(k)$$

therefore

$$\frac{\lambda}{1+\rho} = \sum_k[\beta(k)]^{1+\rho}e^{-E_{0k}(\rho)} = \exp[-E_0(\rho)]$$

(assuming equality, which will be justified below) and the necessary and sufficient conditions become

$$[\beta(k)]^{\rho}e^{-E_{0k}(\rho)} = e^{-E_0(\rho)} = \sum_i[\beta(i)]^{1+\rho}e^{-E_{0i}(\rho)} \quad (*) \text{ for all } k$$

We try the solution $\beta(k) = \dfrac{e^{E_{0k}(\rho)/\rho}}{\sum_{i=1}^{n} \exp[E_{0i}(\rho)/\rho]}$ (>0)

and (*) becomes true equality

$$\frac{1}{\sum_{i=1}^{n} \exp[E_{0i}(\rho)/\rho]} = \frac{\sum_{i=1}^{n} \exp[E_{0i}(\rho)/\rho]}{\left\{\sum_{i=1}^{n} \exp[E_{0i}(\rho)/\rho]\right\}^{1+\rho}}$$

thus $\beta(k) = \dfrac{\exp[E_{0k}(\rho)/\rho]}{\sum_{i=1}^{n} \exp[E_{0i}(\rho)/\rho]} > 0$, for all k

and

$$E_0(\rho) = -\ln \sum_{i=1}^{n} [\beta(i)]^{1+\rho} \exp[-E_{0i}(\rho)] = -\ln\left[\sum_{i=1}^{n} \exp\left[\frac{E_{0i}(\rho)}{\rho}\right]\right]^{\rho}$$

or

$$\exp\left[\frac{E_0(\rho)}{\rho}\right] = \sum_{i=1}^{n} \exp\left[\frac{E_{0i}(\rho)}{\rho}\right]$$

(b) $$C = \lim_{\rho\to 0} \frac{\partial E_0(\rho)}{\partial \rho} = \lim_{\rho \to 0} \frac{E_0(\rho)}{\rho}$$

$$= \lim_{\rho\to 0}\left[\ln \sum_{i=1}^{n} e^{E_{0i}(\rho)/\rho}\right] = \ln\left\{\sum_{i=1}^{n} \exp \lim_{\rho\to 0}\left[\frac{E_{0i}(\rho)}{\rho}\right]\right\}$$

$$= \ln \sum_{i=1}^{n} \exp[C^{(i)}]$$

where $C^{(i)} = \lim_{\rho\to 0} \dfrac{E_{0i}(\rho)}{\rho}$

(c) see solution to problem 3.9.

3.16 (a) Since the messages $m_1, m_2, \ldots, m_L$ are L most likely messages then we have as decision rule: when $\underline{y}$ is received we choose $m_1, m_2, \ldots, m_L$ as our list if

$$P_N(\underline{y}|\underline{x}_{m_\ell}) > P_N(\underline{y}|\underline{x}_{m'}) \quad \text{for all } \ell \text{ in the list and all } m' \neq \ell$$

or
$$\prod_{\ell=1}^{L} \frac{P_N(\underline{y}|\underline{x}_{m_\ell})}{P_N(\underline{y}|\underline{x}_{m'})} > 1$$

So given that m was sent, an error occur if we receive $\underline{y}$ such that

$$\underline{y} \in \bar{\Lambda}_m \equiv \left\{ \underline{y} : \prod_{\ell=1}^{L} \frac{P_N(\underline{y}|\underline{x}_{m_\ell})}{P_N(\underline{y}|\underline{x}_{m'})} > 1, \text{ for some set of } L \text{ messages where } m_\ell \neq m \text{ for all } \ell \right\}$$

or $P_{E_m} = \Pr\{\underline{y} \in \bar{\Lambda}_m | \underline{x}_m\} = \sum_{\underline{y} \in \bar{\Lambda}_m} P_N(\underline{y}|\underline{x}_m)$

(b) if $\underline{y} \in \bar{\Lambda}_m$ then $\left[\prod_{\ell=1}^{L} \frac{P_N(\underline{y}|\underline{x}_{m_\ell})}{P_N(\underline{y}|\underline{x}_m)} \right]^{\lambda} > 1$ for any $\lambda > 0$

Summing over all possible values of $m_1, \ldots, m_L$ ($m_\ell \neq m$ for all ℓ) we must include this set and all other terms are strictly nonnegative and therefore the inequality is still satisfied and thus $\underline{y} \in \Lambda_m^b$ where

$$\Lambda_m^b \equiv \left\{ \underline{y} : \sum_{m_1 \neq m} \cdots \sum_{m_L \neq m} \left[\prod_{\ell=1}^{L} \frac{P_N(\underline{y}|\underline{x}_{m_\ell})}{P_N(\underline{y}|\underline{x}_m)} \right]^{\lambda} > 1 \right\} \quad , \lambda > 0$$

Therefore
$$P_{E_m} < \sum_{\underline{y} \in \Lambda_m^b} P_N(\underline{y}|\underline{x}_m) = \sum_{\underline{y} \in \mathcal{Y}^N} P_N(\underline{y}|\underline{x}_m) f_N(\underline{y}) \quad (*)$$

where $f_N(\underline{y})$ is the indicator function given by

$$f_N(\underline{y}) = \begin{cases} 1 & \text{if } \underline{y} \in \Lambda_m^b \\ 0 & \text{if } \underline{y} \notin \Lambda_m^b \end{cases}$$

By definition of Λ_m^b we have

$$f_N(\underline{y}) < \left\{ \sum_{m_1 \neq m} \cdots \sum_{m_L \neq m} \left[\prod_{m=1}^{L} \frac{p_N(\underline{y}|\underline{x}_{m_\ell})}{p_N(\underline{y}|\underline{x}_m)} \right]^{\lambda} \right\}^{\rho}, \rho > 0$$

Letting $\lambda = \frac{1}{1+\rho L}$ and upperbounding $f_N(\underline{y})$ in (*) as above yields the desired result.

(c) taking ensemble average

$$\overline{P_{E_1}} < \sum_{\underline{y} \in \mathcal{Y}^N} \sum_{\underline{x}_1 \in \mathcal{X}^N} q(\underline{x}_1)\, p(\underline{y}|\underline{x}_1)^{\frac{1}{1+\rho L}} \sum_{\underline{x}_2 \in \mathcal{X}^N} \cdots \sum_{\underline{x}_M \in \mathcal{X}^N} q(\underline{x}_2) \cdots q(\underline{x}_M)$$

$$\cdot \left\{ \sum_{m_1 \neq m} \cdots \sum_{m_L \neq m} \prod_{\ell=1}^{L} p(\underline{y}|\underline{x}_{m_\ell})^{\frac{1}{1+\rho L}} \right\}^{\rho}$$

Since $E(g^\rho) \leq [E(g)]^\rho$ for $0 \leq \rho \leq 1$, we have

$$\overline{P_{E_1}} < \sum_{\underline{y} \in \mathcal{Y}^N} \sum_{\underline{x} \in \mathcal{X}^N} q(\underline{x}_1)\, p(\underline{y}|\underline{x}_1)^{\frac{1}{1+\rho L}} \left\{ \sum_{\underline{x}_2 \in \mathcal{X}^N} \cdots \sum_{\underline{x}_M \in \mathcal{X}^N} q(\underline{x}_2) \cdots q(\underline{x}_M) \right.$$

$$\left. \cdot \sum_{m_1 \neq m} \cdots \sum_{m_L \neq m} \prod_{\ell=1}^{L} p(\underline{y}|\underline{x}_{m_\ell})^{\frac{1}{1+\rho L}} \right\}^{\rho}$$

The set $\{m_2 \ldots m_k \ldots m_L$ where $m_k = m$ for all $k\}$ has cardinality $\binom{M-1}{L}$.

Since now all the weightings $q(\underline{x}_k)$ are equall we get

$$\overline{P_{E_1}} < \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \sum_{\underset{\sim}{x} \in \mathcal{X}^N} q(\underset{\sim}{x})\, p(\underset{\sim}{y}|\underset{\sim}{x})^{\frac{1}{1+\rho L}} \left\{ \binom{M-1}{L} \left[\sum_{\underset{\sim}{x} \in \mathcal{X}^N} q(\underset{\sim}{x})\, p(\underset{\sim}{y}|\underset{\sim}{x})^{\frac{1}{1+\rho L}} \right]^L \right\}^\rho$$

Finally since $\binom{M-1}{L} < (M-1)^L$, we have for $0 \le \rho \le 1$

$$\overline{P_{E_1}} < \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \sum_{\underset{\sim}{x} \in \mathcal{X}^N} q(\underset{\sim}{x})\, p(\underset{\sim}{y}|\underset{\sim}{x})^{\frac{1}{1+\rho L}} \left\{ (M-1) \sum_{\underset{\sim}{x} \in \mathcal{X}^N} q(\underset{\sim}{x})\, p(\underset{\sim}{y}|\underset{\sim}{x})^{\frac{1}{1+\rho L}} \right\}^{\rho L}$$

Since this is the same for all m and upperbounding $(M-1)$ trivially by M we get

$$\overline{P_E} < M^{\tilde{\rho}} \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \left\{ \sum_{\underset{\sim}{x} \in \mathcal{X}^N} q(\underset{\sim}{x})\, p(\underset{\sim}{y}|\underset{\sim}{x})^{\frac{1}{1+\tilde{\rho}}} \right\}^{1+\tilde{\rho}}$$

where $\tilde{\rho} = \rho L$ and therefore $0 \le \tilde{\rho} \le L$.

(d) This yields $\overline{P_E} < \exp\{-N[E_0(\tilde{\rho}, q) - \tilde{\rho} R]\}$

where $R = \dfrac{\ln M}{N}$ $\qquad 0 \le \tilde{\rho} \le L$

Lower bound assuming L finite and $N \to \infty$ is

$$P_E > \exp\left\{-N\left[E_{sp}(\hat{R}) - o(N)\right]\right\} = \exp\left\{-N\left[E_{sp}(R) - o(N)\right]\right\}$$

where

$$\hat{R} = \frac{\ln(M/L)}{N} = \frac{\ln M}{N} - \frac{\ln L}{N} = R - o(N)$$

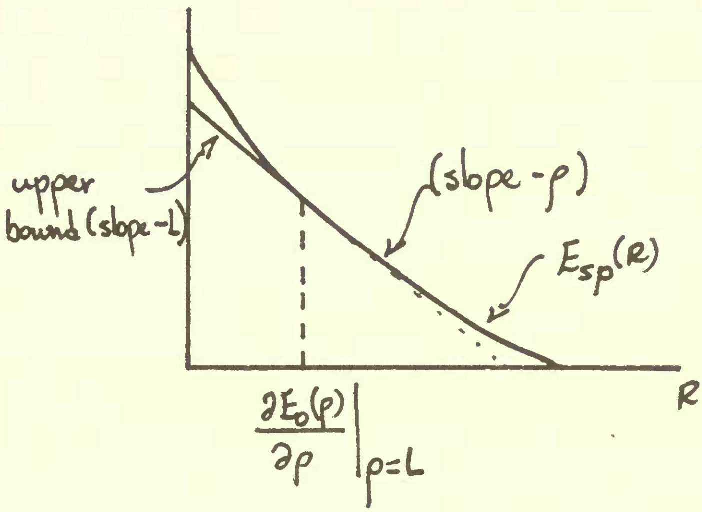
where $o(N) \to 0$ as $N \to \infty$

$$E_{sp}(R) = \sup_{0 \le \rho < \infty} \left[E_0(\rho) - \rho R\right]$$

while in the upper bound for list decoding we have

$$E_e(R) = \max_{0 \le \rho < L} \left(E_0(\rho) - \rho R \right)$$

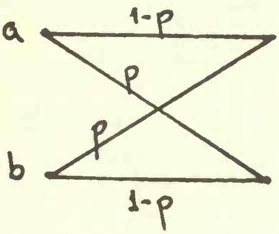
where $E_0(\rho) = \max_{q} E_0(\rho, q)$ and so we'll have



<u>3.17</u> Consider 2 messages of length N, $aa\ldots a$ and $bb\ldots b$, then from (3.5.3) we have

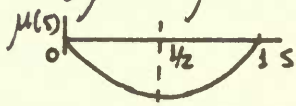
$$\mu(s) = N \ln\left\{ \sum_{y \in \mathcal{Y}} P_a(y)^{1-s} P_b(y)^s \right\} \quad , 0 \le s \le 1$$

(a) BSC



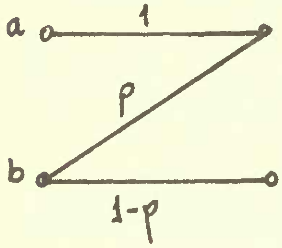
$$\mu(s) = N \ln\left[(1-p)^{1-s} p^s + (1-p)^s p^{1-s} \right]$$

$$\therefore \mu(0) = \mu(1) = \ln 1 = 0 \ ; \ \mu(s) = \mu(1-s)$$

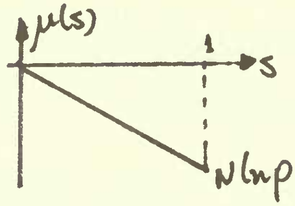


so minimum occurs at $s = 1/2$ and $\mu(1/2) = N \ln \sqrt{4p(1-p)}$

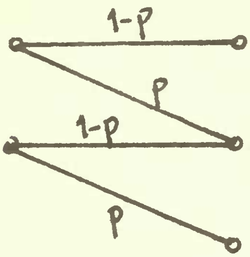
(b) Z channel



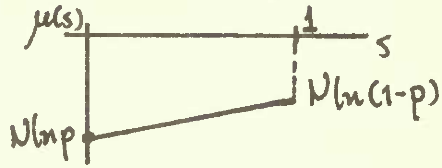
$$\mu(s) = N \ln p^s = Ns \ln p$$



(c) since only the first two inputs are used, the channel becomes effectively



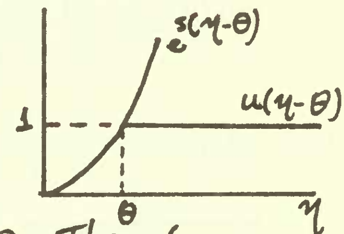
$$\mu(s) = N \ln\left[(1-p)^s p^{1-s}\right]$$
$$= N\left[(1-s)\ln p + s \ln(1-p)\right]$$



<u>3.18</u> Take η to be a continuous random variable with finite moments of all order (for a discrete r.v. the reasoning is the same with summations replacing integrals).

(a) $$P_r\{\eta > \theta\} = \int_\theta^\infty p(\eta)\,d\eta = \int_{-\infty}^{\infty} u(\eta-\theta)\,p(\eta)\,d\eta$$

where $$u(\eta-\theta) = \begin{cases} 1 & \eta \geq \theta \\ 0 & \eta < \theta \end{cases}$$



Clearly $u(\eta-\theta) \leq e^{s(\eta-\theta)}$ for $s \geq 0$. Therefore

$$P_r\{\eta > \theta\} \leq \int_{-\infty}^{\infty} e^{s(\eta-\theta)} p(\eta)\,d\eta = \left\{\int_{-\infty}^{\infty} e^{s\eta} p(\eta)\,d\eta\right\} e^{-s\theta}$$

$$= E(e^{s\eta})\,e^{-s\theta} = E\left\{e^{s(\eta-\theta)}\right\}, \quad s \geq 0$$

let $e^{\Gamma(s)} \triangleq E[e^{s\eta}]$ then $\Gamma(s) = \ln \int_{-\infty}^{\infty} e^{s\eta} p(\eta) d\eta$

and $P_r\{\eta > \theta\} \leq e^{[\Gamma(s) - s\theta]}$

(b) to obtain the tightest bound we minimize the expo nent, $\Gamma(s) - s\theta$

$$\Gamma(s) = \ln \int_{-\infty}^{\infty} e^{s\eta} p(\eta)\, d\eta \quad \text{and let} \quad \tilde{p}(\eta) \triangleq \frac{e^{s\eta} p(\eta)}{\int_{-\infty}^{\infty} e^{s\eta} p(\eta) d\eta}$$

then

$$\Gamma'(s) = \frac{\int_{-\infty}^{\infty} \eta e^{s\eta} p(\eta) d\eta}{\int_{-\infty}^{\infty} e^{s\eta} p(\ ) d\eta} = \int_{-\infty}^{\infty} \eta\, \tilde{p}(\eta) d\eta \triangleq \bar{\eta}_s$$

and

$$\Gamma''(s) = \int_{-\infty}^{\infty} \eta^2 \tilde{p}(\eta) d\eta - \bar{\eta}_s^{\,2} = Var_s(\eta) \geq 0$$

Therefore $\Gamma(s)$ (and thus $\Gamma(s) - s\theta$) is a convex ∪ function in s. This could also be shown by taking $p_a(\eta) = p(\eta)$ and $p_b(\eta) = c\,p(\eta)e^{\eta}$ in (3.5.3) and use the result of theorem (3.5.1) that $\mu(s) = \text{constant} \times \Gamma(s)$ is convex ∪ in s.

So we can minimize by setting $\Gamma'(s) - \theta = 0$

$$\therefore \; P_r\{\eta > \theta\} \leq e^{[\Gamma(s) - s\Gamma'(s)]} \qquad \text{where } \theta = \Gamma'(s)$$

(c) If $\eta = \sum_{n=1}^{N} y_n$ where the y_n's are independent and identically distributed random variables

then

$$\Gamma(s) = \ln E(e^{s\eta}) = \ln E\left(e^{s\sum_{n=1}^{N} y_n}\right) = \ln E\left(\prod_{n=1}^{N} (e^{s y_n})\right)$$

$$\therefore\ \Gamma(s) = \ln \prod_{n=1}^{N} E\left(e^{sY_n}\right) = e^{N\gamma(s)}$$

where $\gamma(s) = \ln E(e^{sY_n}) = \ln\left[\int_{-\infty}^{\infty} e^{sY_n} p(Y_n)\, dY_n\right]$ for all n

thus

$$P_r\{\eta > \theta\} \le e^{N[\gamma(s) - s\gamma'(s)]}$$

<u>3.19</u> (a) $\gamma(s) = \ln E(e^{sY_n}) = \ln\left[(1-p)e^{s.0} + p e^{s.1}\right]$

$$= \ln\left[(1-p) + pe^{s}\right]$$

$N\gamma'(s) = \dfrac{Npe^s}{(1-p)+pe^s} = \theta$ or $\gamma'(s) = \delta \triangleq \theta/N$

or $pe^s(1-\delta) = (1-p)\delta \longrightarrow pe^s = \dfrac{(1-p)\delta}{1-\delta}$

$$\therefore\ s = \ln\left[\frac{(1-p)}{p}\cdot\frac{\delta}{1-\delta}\right]$$

Therefore $\gamma(s) = \ln\left[(1-p)\left(1+\dfrac{\delta}{1-\delta}\right)\right] = \ln\left[\dfrac{1-p}{1-\delta}\right]$

$$\gamma(s) - s\gamma'(s) = \ln\left[\frac{1-p}{1-\delta}\right] - \delta\ln\left[\frac{1-p}{p}\cdot\frac{\delta}{1-\delta}\right]$$

$$= (1-\delta)\ln(1-p) + \delta\ln p - (1-\delta)\ln(1-\delta) - \delta\ln\delta$$

Then $P_r\{\eta > \theta\} \le e^{N[\gamma(s) - s\gamma'(s)]} = (1-p)^{N(1-\delta)}\, p^{N\delta}\, e^{N\mathcal{H}(\delta)}$

where $\mathcal{H}(\delta) = -\delta\ln\delta - (1-\delta)\ln(1-\delta)$ and $\delta = \theta/N$

Note: at $\delta = p$ the exponent is zero, so we need $\delta > p$.

(b) $\Gamma(s) = \ln \int_{-\infty}^{\infty} e^{s\eta} \frac{e^{-\eta^2/2}}{\sqrt{2\pi}} d\eta = \ln\{e^{s^2/2}\} = \frac{s^2}{2}$

$\Gamma'(s) = s = \theta$ and therefore $\Gamma(s) - s\Gamma'(s) = -\frac{s^2}{2} = -\frac{\theta^2}{2}$

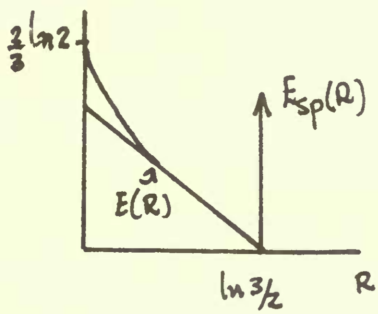
$\therefore \quad P_r\{\eta > \theta\} \leq e^{-\theta^2/2} \quad , \theta > 0$

<u>3.20</u> $E_{sp}(R) = \max_{0 \leq \rho \leq \infty} [E_0(\rho) - \rho R]$

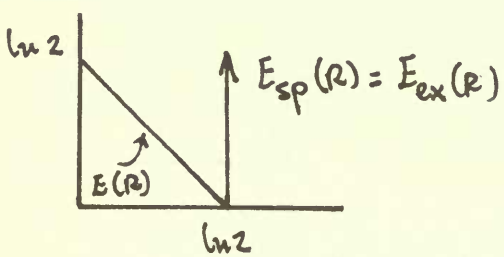
(a) shere-packing bound for problem 3.2

(i) $E_0(\rho) = \rho \ln 3/2 \quad \therefore E_{sp}(R) = \max_{0 \leq \rho \leq \infty} \rho[\ln 3/2 - R] = \infty$

for all $R < \ln \frac{3}{2}$



(ii) $E_0(\rho) = \rho \ln 2 \quad \therefore E_{sp}(R) = \infty$ for all $R < \ln 2$

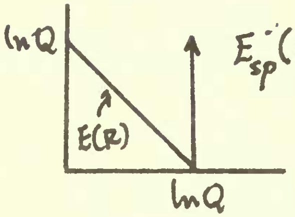


(see problems 3.2 and 3.8)

Note: in (i) $E_{sp}(R)$ is nowhere tight, while in (ii) it is everywhere tight.

(b) sphere-packing bound for problem 3.3

3.3a $E_0(\rho) = \rho \ln Q \implies E_{sp}(R) = \infty$ for all $R < \ln Q$



$E_{sp}(R) = E_{ex}(R)$ (therefore everywhere tight)

3.3b $E_0(\rho) = \rho \ln Q - (1+\rho) \ln\left[(Q-1)\, p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right]$

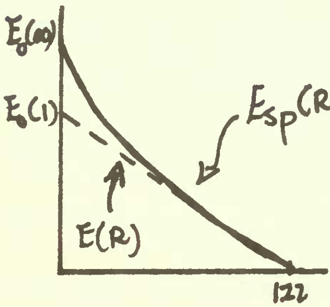
$$\lim_{\rho\to\infty} E_0(\rho) \sim \rho \ln Q - (1+\rho) \ln\left[(Q-1)\left(1 + \frac{1}{1+\rho} \ln p\right) + 1 + \frac{1}{1+\rho} \ln(1-p)\right]$$

$$= \rho \ln Q - (1+\rho) \ln\left[Q + \frac{Q-1}{1+\rho} \ln p + \frac{1}{1+\rho} \ln(1-p)\right]$$

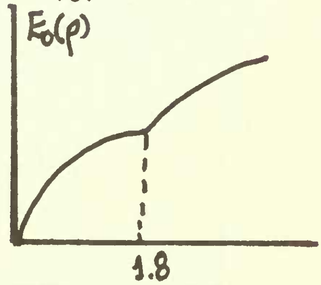
$$= \rho \ln Q - (1+\rho) \ln Q - (1+\rho) \ln\left[1 + \frac{Q-1}{Q(1+\rho)} \ln p + \frac{1}{Q(1+\rho)} \ln \bar{p}\right]$$

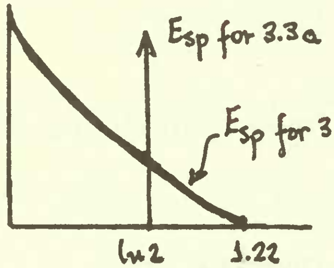
where $\bar{p} = 1-p$.

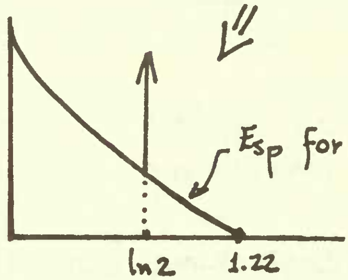
$$\therefore E_0(\infty) \cong -\ln Q + \frac{Q-1}{Q} \ln p + \frac{1}{Q} \ln \bar{p}$$



For 3.3c







(c) same process with $E_0(\rho)$ as given in problem 3.5.

(d) for problem 3.6 we have $E(R) = E_0(\rho) - \rho E_0'(\rho)$

$R = E_0'(\rho)$

Here we can solve it explicitly. Letting $\rho \to \infty$ then

$$E_{sp}(R) = \ln 2 - \mathcal{H}(R/\ln Q) \quad \text{for all } 0 < R < C = \frac{\ln Q}{2}$$

and $\lim_{R \to 0} E_{sp}(R) = \ln 2$

<u>3.21</u> (a) using inequality (g) in Appendix 3A, replacing λ by s, we have

$$\left(\sum_i a_i\right)^s \le \sum_i a_i^s \qquad \text{for } \begin{cases} s \in (0,1] \\ a_i \geq 0 \end{cases}$$

Let

$$a_{m'} = \sum_{\underset{\sim}{y}} \sqrt{P_N(\underset{\sim}{y} | \underset{\sim}{x}_m) P_N(\underset{\sim}{y} | \underset{\sim}{x}_{m'})} \qquad \text{then}$$

$$B_m^s(\mathcal{C}) \le \sum_{m' \neq m} \left[\sum_{\underset{\sim}{y}} \sqrt{P_N(\underset{\sim}{y} | \underset{\sim}{x}_m) P_N(\underset{\sim}{y} | \underset{\sim}{x}_{m'})} \right]^s \quad , m = 1, 2, \ldots, M$$

(b) averaging over the ensemble of the m^{th} codeword

$$\overline{B_m^s(\mathcal{C})}^{m} \le \sum_{\underset{\sim}{x}_m} q_N(\underset{\sim}{x}_m) \sum_{m' \neq m} \left[\prod_{n=1}^{N} \sum_{y_n} \sqrt{P(y_n | x_{mn}) P(y_n | x_{m'n})} \right]^s$$

$$= \sum_{m' \neq m} \prod_{n=1}^{N} \sum_{x} q(x) \left[\sum_{y_n \in \mathcal{Y}} \sqrt{P(y_n | x) P(y_n | x_{m'n})} \right]^s$$

Assume that all codewords in $\mathcal{C}$ satisfy $q_N(\underset{\sim}{x}) > 0$, then let

$$\gamma(s, q) \triangleq \max_{\substack{x' \\ q(x') > 0}} \left\{ \sum_x q(x) \left(\sum_y \sqrt{P(y|x) P(y|x')} \right)^s \right\}$$

Then we have

$$\overline{B_m^s(\mathcal{C})}^{\,m} \leq (M-1)\left[\gamma(s,q)\right]^N \leq M\left[\gamma(s,q)\right]^N$$

(c) the above inequality means that given a code $\mathcal{C}=\{\underset{\sim}{x}_1,\dots,\underset{\sim}{x}_M\}$ there exists at least one codeword $\hat{\underset{\sim}{x}}_m$ such that the code $\mathcal{C}_m = \{\underset{\sim}{x}_1,\dots,\underset{\sim}{x}_{m-1},\hat{\underset{\sim}{x}}_m,\underset{\sim}{x}_{m+1},\dots,\underset{\sim}{x}_M\}$ satisfies

$$P_{E_m}(\mathcal{C}_m) \leq B_m(\mathcal{C}_m) \leq M^{1/s}\left[\gamma(s,q)\right]^{N/s} \quad , s\in(0,1]$$

otherwise we contradict ensemble average.

(d) we can construct code $\hat{\mathcal{C}}_m$ from $\hat{\mathcal{C}}_{m-1}$ as follows: consider code $\hat{\mathcal{C}}_{m-1}$, fix all codewords except $\underset{\sim}{x}_m$ then as we did in part (c) there exists at least one $\hat{\underset{\sim}{x}}_m$ such that

$$B_m^s(\hat{\mathcal{C}}_m) \leq \overline{B_m^s(\hat{\mathcal{C}}_{m-1})}^{\,m}$$

choose this $\hat{\underset{\sim}{x}}_m$ and construct $\hat{\mathcal{C}}_m$ from $\hat{\mathcal{C}}_{m-1}$ by replacing $\underset{\sim}{x}_m$ by $\hat{\underset{\sim}{x}}_m$ and keeping all the other codewords from $\hat{\mathcal{C}}_{m-1}$. Then by (b) and (c) we have

$$P_{E_m}(\hat{\mathcal{C}}_m) \leq B_m(\hat{\mathcal{C}}_m) \leq M^{1/s}\left[\gamma(s,q)\right]^{N/s} \qquad s\in(0,1]$$

for all $m=1,2,\dots,M$

where
$$B_m(\hat{\mathcal{C}}_m) = \sum_{m'=1}^{m-1}\sum_{\underset{\sim}{y}}\sqrt{P_N(\underset{\sim}{y}|\hat{\underset{\sim}{x}}_m)P_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})}$$
$$+\sum_{m'=m+1}^{M}\sum_{\underset{\sim}{y}}\sqrt{P_N(\underset{\sim}{y}|\hat{\underset{\sim}{x}}_m)P_N(\underset{\sim}{y}|\hat{\underset{\sim}{x}}_{m'})}$$

(e) for code $\hat{\mathcal{C}}_M = \{\hat{\underset{\sim}{x}}_1,\hat{\underset{\sim}{x}}_2,\dots,\hat{\underset{\sim}{x}}_M\}$ we have

$$B_m(\hat{\ell}_M) = \sum_{m' \neq m} \sum_{y} \sqrt{P_N(y|x_m) P_N(y|x_{m'})}$$

but

$$B_m(\hat{\ell}_M) \leq B_m(\hat{\ell}_M) + \sum_{m'=m+1}^{M} \sum_{y} \sqrt{P_N(y|\hat{x}_m) P_N(y|x_{m'})}$$

$$= \sum_{m'=1}^{m-1} \sum_{y} P_N(y|\hat{x}_m) P_N(y|x'_m) + \sum_{m'=m+1}^{M} \sum_{y} \sqrt{P_N(y|\hat{x}_m) P_N(y|x_{m'})}$$

$$+ \sum_{m'=m+1}^{M} \sum_{y} \sqrt{P_N(y|\hat{x}_m) P_N(y|\hat{x}_{m'})}$$

$$= B_m(\hat{\ell}_m) + \sum_{m'=m+1}^{M} \sum_{y} \sqrt{P_N(y|\hat{x}_m) P_N(y|\hat{x}_{m'})}$$

for $m = 1, 2, \ldots, M$.

(f) $$P_E(\hat{\ell}_M) = \frac{1}{M} \sum_{m=1}^{M} P_{E_m}(\hat{\ell}_M) \leq \frac{1}{M} \sum_{m=1}^{M} B_m(\hat{\ell}_M)$$

$$\leq \frac{1}{M} \sum_{m=1}^{M} B_m(\hat{\ell}_m) + \frac{1}{M} \sum_{m=1}^{M} \sum_{m'=m+1}^{M} \sum_{y} \sqrt{P_N(y|x_m) P_N(y|x_{m'})}$$

changing the order of summation in the second term we have

$$P_E(\hat{\ell}_M) \leq \frac{1}{M} \sum_{m=1}^{M} B_m(\hat{\ell}_m) + \frac{1}{M} \sum_{m'=1}^{M} \sum_{m=1}^{m'-1} \sum_{y} \sqrt{P_N(y|\hat{x}_m) P_N(y|\hat{x}_{m'})}$$

$$= \frac{1}{M} \sum_{m=1}^{M} B_m(\hat{\ell}_m) + \frac{1}{M} \sum_{m=1}^{M} B_{m'}(\hat{\ell}_{m'})$$

therefore $P_E(\hat{\ell}_M) \leq \frac{2}{M} \sum_{m=1}^{M} B_m(\hat{\ell}_m)$ and using (d)

we get $P_E(\hat{\ell}_M) \leq 2 M^{1/s} [\gamma(s,q)]^{N/s}$

(g) For any q we have

$$\sum_x \sum_{x'} q(x)q(x')\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^s \leq \gamma(s,q)$$

We now find necessary conditions on q that minimizes the left hand side of the above inequality.
Let λ be a Lagrange multiplier, so we minimize

$$J(q) = \sum_x \sum_{x'} q(x)q(x')\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^s - \lambda \sum_x q(x)$$

then the necessary condition is

$$\left.\frac{\partial J}{\partial q(\hat{x})}\right|_{q=q^*} = 2\sum_x q^*(x)\left(\sum_y \sqrt{p(y|x)p(y|\hat{x})}\right)^s - \lambda \geq 0$$

with equality when $q^*(\hat{x}) > 0$. Finally for $\frac{1}{s} = \rho$, there exists a code $\mathscr{C}$ such that

$$P_E(\mathscr{C}) \leq 2M^\rho \left\{ \sum_x \sum_{x'} q(x)q(x')\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^{\frac{1}{\rho}} \right\}^{N\rho}$$

$$= 2e^{-N\{E_x(\rho,q) - \rho R\}}$$

<u>3.22</u> (a) $$P_E = \frac{1}{M}\sum_{m=1}^{M} \sum_{\underset{\sim}{y} \in \bar{\Lambda}_m} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)$$

$$\geq \frac{1}{M}\sum_{m=1}^{M} \sum_{\underset{\sim}{y} \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)$$

since $\bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m) \subset \bar{\Lambda}_m$

where $G_\gamma(\underset{\sim}{x}_m) = \left\{ \underset{\sim}{y} : \frac{1}{N} \ln \frac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)}{P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)} - J(\tilde{p}, p) < \gamma \right\}$

or $G_\gamma(\underset{\sim}{x}_m) = \left\{ y : p_N(y|\underset{\sim}{x}_m) \geq e^{-N(J(\tilde{p},p)+\gamma)} \tilde{p}_N(y|\underset{\sim}{x}_m) \right\}$

For $y \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)$ we have

$$p_N(y|\underset{\sim}{x}_m) \geq e^{-N(J(\tilde{p},p)+\gamma)} \cdot \tilde{p}_N(y|\underset{\sim}{x}_m)$$

therefore

$$P_E \geq \frac{1}{M} \sum_{m=1}^{M} \sum_{y \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)} p_N(y|\underset{\sim}{x}_m)$$

$$\geq e^{-N[J(\tilde{p},p)+\gamma]} \frac{1}{M} \sum_{m=1}^{M} \sum_{y \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m)$$

but $\bar{\Lambda}_m = \left[\bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)\right] \cup \left[\bar{\Lambda}_m \cap \bar{G}_\gamma(\underset{\sim}{x}_m)\right]$

then

$$\sum_{y \in \bar{\Lambda}_m} \tilde{p}_N(y|\underset{\sim}{x}_m) = \sum_{y \in \bar{\Lambda}_m \cap G_\gamma(x_m)} \tilde{p}_N(y|\underset{\sim}{x}_m) + \sum_{y \in \bar{\Lambda}_m \cap \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m)$$

Since $\bar{\Lambda}_m \cap \bar{G}_\gamma(\underset{\sim}{x}_m) \subset \bar{G}_\gamma(\underset{\sim}{x}_m)$ we have

$$\sum_{y \in \bar{\Lambda}_m \cap \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m) \leq \sum_{y \in \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m)$$

$$\therefore \sum_{y \in \bar{\Lambda}_m} \tilde{p}_N(y|\underset{\sim}{x}_m) \leq \sum_{y \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m) + \sum_{y \in \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m)$$

then

$$\sum_{y \in \bar{\Lambda}_m \cap G_\gamma(\underset{\sim}{x}_m)} \tilde{p}_N(y|\underset{\sim}{x}_m) \geq \sum_{y \in \bar{\Lambda}_m} \tilde{p}_N(y|\underset{\sim}{x}_m) - \sum_{y \in \bar{G}_\gamma(\underset{\sim}{x}_m)} p_N(y|\underset{\sim}{x}_m)$$

and so

$$P_E \geq e^{-N[J(\tilde{p},p)+\gamma]} \left\{ \tilde{P}_E - \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y} \in \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{P}_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \right\} \qquad (1)$$

where $\tilde{P}_E = \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y} \in \bar{\Lambda}_m} \tilde{P}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)$

(b) $$\sum_{\underset{\sim}{y} \in \bar{G}_\gamma(\underset{\sim}{x}_m)} \tilde{P}_N(\underset{\sim}{y}|\underset{\sim}{x}_m) = \tilde{P}_r \left\{ \frac{1}{N} \ln \frac{\tilde{P}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)}{P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)} - J(\tilde{p},p) \geq \gamma \right\}$$

$$= \tilde{P}_r \left\{ \frac{1}{N} \sum_{n=1}^{N} \ln \frac{\tilde{p}(y_n|x_{mn})}{p(y_n|x_{mn})} - J(\tilde{p},p) \geq \gamma \right\}$$

Let $z_n \triangleq \ln \frac{\tilde{p}(y_n|x_{mn})}{p(y_n|x_{mn})} \qquad n = 1, 2, \ldots, N$

Then the z_n are iid random variables with mean

$$\bar{z} = \tilde{E}\{z_n\} = \sum_{y} \tilde{p}(y|x) \ln \frac{\tilde{p}(y|x)}{p(y|x)} = J(\tilde{p},p) \quad ,\text{all } n.$$

By the weak law of large numbers, as $N \to \infty$ for $\gamma > 0$

$$P_r \left\{ \frac{1}{N} \sum_{m=1}^{N} z_n - \bar{z} \geq \gamma \right\} \longrightarrow 0 \qquad (2)$$

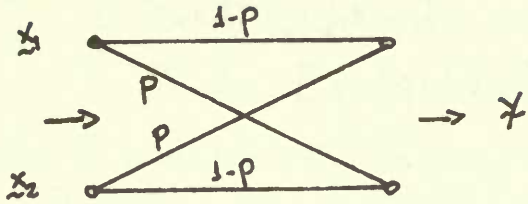
(c) the converse to the coding theorem says: there exists $\alpha > 0$ such that $\tilde{P}_E \geq \alpha$ for all N, whenever $R > \tilde{C} = \ln 2 - \mathcal{H}(\tilde{p})$.

Therefore for large N we have

$$P_E \geq \alpha\, e^{-N[J(\tilde{p},p)+\gamma]}$$

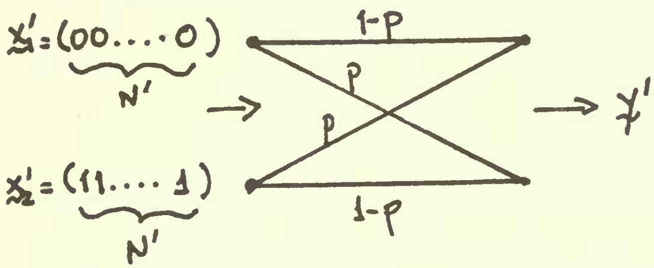
since by (2) the second term in (1) goes to zero as $N \to \infty$.

3.23



$$P_E = \frac{1}{2}\left(P_{E_1} + P_{E_2}\right) \quad \text{where } P_{E_1} = \sum_{y \in \Lambda_2} p_N(y|x_1) \text{ and } P_{E_2} = \sum_{y \in \Lambda_1} p_N(y|x_2)$$

Consider a "dummy" BSC with $\tilde{p} = 1/2$. Then $\tilde{C} = \ln 2 - \mathcal{H}(\tilde{p}) = 0$.
Let $N' = w(x_1 \oplus x_2)$ = number of places where $x_{1n} \neq x_{2n}$.
Then clearly P_E is the same if we only consider N' of x_1 and x_2 where they are different. It is equivalent to consider



then $P_E = P_E' = P_{E_1} = P_{E_2} = \sum_{y' \in \Lambda_1'} P_{N'}(y'|x_2')$

Then as in problem 3.22 for any $\gamma > 0$ let

$$G_\gamma \triangleq \left\{ y' : \frac{1}{N'} \ln \frac{\tilde{P}_{N'}(y'|x_2')}{P_{N'}(y'|x_2')} - J\left(\tfrac{1}{2}; p\right) < \gamma \right\}$$

$$P_E \geq \sum_{\Lambda_1 \cap G_\gamma(x_2)} P_{N'}(y'|x_2') \geq e^{-N'[J(\frac{1}{2};p)+\gamma]} \left\{ \tilde{P}_E - \sum_{\overline{G_\gamma(x_2)}} \tilde{P}_{N'}(y'|x_2') \right\}$$

but $\sum_{\overline{G_\gamma(x_2)}} \tilde{P}_{N'}(y'|x_2') \xrightarrow{N \to \infty} 0$

and by the converse to the coding theorem there exists $\alpha > 0$ such that $\tilde{P}_E \geq \alpha$ for all N'. From chapter 1 we find that for $\tilde{p} = 1/2$, $\tilde{c} = 0$, $\alpha = 1/2$. Then

$$P_E \geq \frac{1}{2} e^{-N'[J(\frac{1}{2};p)+\gamma]}$$

where $J(1/2;p) = \frac{1}{2}\ln\frac{1/2}{p} + \frac{1}{2}\ln\frac{1/2}{1-p} = -\ln\sqrt{4p(1-p)}$

Finally
$$P_E \geq \frac{1}{2} e^{-w(\underline{x}_1 \oplus \underline{x}_2)\left[-\ln\sqrt{4p(1-p)}+\gamma\right]}$$

<u>3.24</u> For any $\gamma > 0$ and any $\underline{x} \in \mathcal{X}_N$ let

$$G_\gamma(\underline{x}) = \left\{ \underline{y} : \frac{1}{N}\ln\frac{\tilde{p}_N(\underline{y}|\underline{x})}{p_N(\underline{y}|\underline{x})} - J(\tilde{p};p) \leq \gamma \right\}$$

thus
$$P_E = \frac{1}{M}\sum_{m=1}^{M}\int_{\bar{\Lambda}_m} p_N(\underline{y}|\underline{x}_m)\,d\underline{y} \geq \frac{1}{M}\sum_{m=1}^{M}\int_{\bar{\Lambda}_m \cap G_\gamma(\underline{x}_m)} p_N(\underline{y}|\underline{x}_m)\,d\underline{y} \quad (1)$$

(since $\bar{\Lambda}_m \cap G_\gamma(\underline{x}_m) \subset \bar{\Lambda}_m$).

But $\underline{y} \in G_\gamma(\underline{x}_m)$ implies $p_N(\underline{y}|\underline{x}_m) \geq e^{-\{J(\tilde{p};p)+\gamma\}N}\,\tilde{p}_N(\underline{y}|\underline{x}_m)$.

Then by the same reasoning used in problem 3.22 we have

$$\int_{\bar{\Lambda}_m \cap G_\gamma(\underline{x}_m)} \tilde{p}_N(\underline{y}|d)\,d\underline{y} \geq \int_{\bar{\Lambda}_m} \tilde{p}_N(\underline{y}|\underline{x}_m)\,d\underline{y} - \int_{\bar{G}_\gamma(\underline{x}_m)} \tilde{p}_N(\underline{y}|\underline{x}_m)\,d\underline{y}$$

and thus from (1)

$$P_E \geq e^{-N[J(\tilde{p};p)+\gamma]}\,\frac{1}{M}\sum_{m=1}^{M}\int_{\bar{\Lambda}_m \cap G_\gamma(\underline{x}_m)} \tilde{p}_N(\underline{y}|\underline{x}_m)\,d\underline{y}$$

$$\geq e^{-N[J(\tilde{p};p)+\gamma]}\,\frac{1}{M}\sum_{m=1}^{M}\left\{\int_{\bar{\Lambda}_m} \tilde{p}_N(\underline{y}|\underline{x}_m)\,d\underline{y} - \int_{\bar{G}_\gamma(\underline{x}_m)} \tilde{p}_N(\underline{y}|\underline{x}_m)\,d\underline{y}\right\}$$

$$= e^{-N[J(\tilde{p};p)+\gamma]}\left\{\tilde{P}_E - \tilde{P}_N(\gamma)\right\}$$

where $\tilde{P}_E = \frac{1}{M}\sum_{m=1}^{M}\int_{\bar{\Lambda}_m}\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)d\underset{\sim}{y}$ is the error probability

when code $\mathcal{C}$ is used over the "dummy" AWGN channel and

$$\tilde{P}_N(\gamma) = \frac{1}{M}\sum_{m=1}^{M}\int_{\bar{G}_\gamma(\underset{\sim}{x}_m)}\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)d\underset{\sim}{y}$$

We now want to show that $\lim_{N\to\infty}\tilde{P}_N(\gamma) = 0$

Note that

$$\int_{\bar{G}_\gamma(\underset{\sim}{x}_m)}\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)d\underset{\sim}{y} = \tilde{P}_r\left\{\frac{1}{N}\ln\frac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)}{p_N(\underset{\sim}{y}|\underset{\sim}{x}_m)} \geq J(\tilde{p};p)+\gamma\right\}$$

$$= \tilde{P}_r\left\{\frac{1}{N}\sum_{n=1}^{N}\ln\frac{\tilde{p}(y_n|x_{mn})}{p(y_n|x_{mn})} \geq J(\tilde{p};p)+\gamma\right\}$$

let $z_n \triangleq \ln\frac{\tilde{p}(y_n|x_{mn})}{p(y_n|x_{mn})}$ and as in problem 3.22 the $\{z_n\}$ are iid random variables with mean $\bar{z} = E\{z_n\} = J(\tilde{p};p)$. By the weak law of large numbers, for any $\gamma>0$

$$\lim_{N\to\infty}\int_{\bar{G}_\gamma(\underset{\sim}{x}_m)}\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)d\underset{\sim}{y} = \lim_{N\to\infty}\tilde{P}_r\left\{\frac{1}{N}\sum_{n=1}^{N}z_n - \bar{z} \geq \gamma\right\} = 0$$

for every $m = 1,2,\dots,M$ and so $\lim_{N\to\infty}\tilde{P}_N(\gamma) = 0$.

By the converse to the coding theorem, there exists an $\alpha>0$ such that $\tilde{P}_E \geq \alpha$ for $R_T > \tilde{C}_T$. Therefore for large N we have

$$P_E \geq \alpha e^{-N[J(\tilde{p};p)+\gamma]} = \alpha e^{-T\{\frac{N}{T}J(\tilde{p};p)+\gamma'\}}$$

where $\frac{N}{T}J(\tilde{p};p) = E_{sp}(R)$ and $\tilde{p}$ satisfies $\tilde{C}_T = R_T$.

Now we compute between these two channels, namely $J(\tilde{p};p)$

$$J(\tilde{p};p) = \int_{\mathcal{Y}} \tilde{p}(y_n|x_{mn}) \ln \frac{\tilde{p}(y_n|x_{mn})}{p(y_n|x_{mn})} dy_n$$

$$= \frac{1}{N} \int_{\mathcal{Y}^N} \tilde{P}_N(\underline{y}|\underline{x}_m) \ln \frac{\tilde{P}_N(\underline{y}|\underline{x}_m)}{P_N(\underline{y}|\underline{x}_m)} d\underline{y}$$

$$= \frac{1}{N} \int_{\mathcal{Y}^N} \tilde{P}_N(\underline{y}|\underline{x}_m) \left\{ \frac{\|\underline{y}-\underline{x}_m\|^2}{N_0} - \frac{\|\underline{y}-\rho\underline{x}_m\|^2}{N_0} \right\} d\underline{y}$$

$$= \frac{1}{NN_0} \int_{\mathcal{Y}^N} \tilde{P}_N(\underline{y}|\underline{x}_m) \left[\|\underline{y}-\rho\underline{x}_m+\rho\underline{x}_m-\underline{x}_m\|^2 - \|\underline{y}-\rho\underline{x}_m\|^2 \right\} d\underline{y}$$

$$= \frac{1}{NN_0} (1-\rho)^2 \|\underline{x}_m\|^2 = \frac{(1-\rho)^2}{NN_0} \mathcal{E}$$

$$\therefore \quad E_{sp}(R) = \frac{N}{T} J(\tilde{p};p) = \frac{N}{T} \frac{(1-\rho)^2}{NN_0} \mathcal{E} = (1-\rho)^2 \frac{\mathcal{E}}{N_0 T}$$

$$= (1-\rho)^2 C_T \quad ; \text{ where by (2.5.3) } \quad C_T = \frac{\mathcal{E}/N_0}{T}$$

and ρ satisfies $R_T = \tilde{C}_T = \rho^2 C_T \implies \rho = \sqrt{\frac{R_T}{C_T}}$

therefore $E_{sp}(R) = \left(1 - \sqrt{\frac{R_T}{C_T}}\right)^2 C_T = \left(\sqrt{C_T} - \sqrt{R_T}\right)^2$

<u>3.25</u> $E_x(\rho,q) = -\rho \ln \left\{ \sum_x \sum_{x'} q(x) q(x') \left[\sum_y \sqrt{p(y|x) p(y|x')} \right]^{1/\rho} \right\}$

let $z \triangleq \sum_y \sqrt{p(y|x)\, p(y|x')}$ for any $x \neq x'$. For $x' = x$ we have $\sum_y \sqrt{p(y|x)\, p(y|x)} = \sum_y p(y|x) = 1$.

$\mathcal{X} = \{x_1, x_2, \dots, x_m\}$

Define
$$F_x(\rho,q) = \sum_x \sum_{x'} q(x)\,q(x')\left[\sum_y \sqrt{p(y|x)\,p(y|x')}\right]^{1/\rho}$$
or
$$F_x(\rho,q) = \sum\sum_{x'=x} q(x)\,q(x')\left[\sum_y \sqrt{p(y|x)\,p(y|x')}\right]^{1/\rho}$$
$$+ \sum\sum_{x'\neq x} q(x)\,q(x')\left[\sum_y \sqrt{p(y|x)\,p(y|x')}\right]^{1/\rho}$$
$$= \sum_x q^2(x) + \sum\sum_{x'\neq x} q(x)\,q(x')\, z^{1/\rho}$$
The maximizing distribution is taken to be $q^* = \left[\frac{1}{m}, \ldots, \frac{1}{m}\right]$

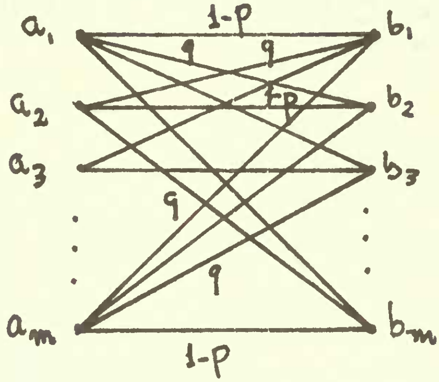
which is a reasonable choice since the equal energy signals are mutually orthogonal and, thus, no particular signal has better performance qualities than any other signal. Therefore, it is intuitively clear that the signals should be equiprobable. To carry out the maximization let

$$F_x(\rho) = \max_q F_x(\rho,q) = F_x(\rho,q)\Big|_{q=q^*} \quad \text{, then we have}$$

$$F_x(\rho) = \sum_{i=1}^{m} \frac{1}{m^2} + 2\sum_{i=1}^{m}\sum_{j=1}^{i-1} \frac{1}{m}\frac{1}{m} z^{1/\rho} = \frac{1}{m} + \frac{2z^{1/\rho}}{m^2}\sum_{i=1}^{m}(i-1)$$

$$= \frac{1}{m} + \frac{2}{m^2}\frac{m(m-1)}{2} z^{1/\rho} = \frac{1}{m} + \frac{m-1}{m} z^{1/\rho} = \frac{1+(m-1)z^{1/\rho}}{m}$$
and thus
$$E_x(\rho) = \max_q E_x(\rho,q) = E_x(\rho,q^*) = -\rho \ln F_x(\rho)$$
$$= -\rho \ln\left[\frac{1+(m-1)z^{1/\rho}}{m}\right]$$
as desired.

(a) For coherent channel with hard m-ary decision outputs we have



for $i = 1, 2, 3, \ldots, m$

$$p(b_i|a_i) = \Pr\{y_i \in \Lambda_i | x_i\}$$

where

$$\Lambda_i = \{y : y_i \geq y_j, \text{ all } j \neq i\}$$

From problem 2.4

$$p(b_i|a_i) = \Pr\{\text{correct decision} \mid x_i\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} \left[1 - \operatorname{erfc}\left(x + \sqrt{\frac{2E_s}{N_0}}\right)\right] dx$$

$$= 1 - p$$

and $p(b_i|a_j) = q$ for $j \neq i$. Thus p is the probability of an error given that x_i is sent. Clearly, the error occurs if any b_j other than b_i is chosen and, by the symmetry of the problem, the probability of choosing any output other than b_i is the same for all $j \neq i$ and is denoted by q. Thus

$$p = (m-1)q \quad \text{or} \quad q = \frac{p}{m-1}$$

$$\therefore \quad p(b_j|a_i) = \begin{cases} 1-p & \text{if } i = j \\ \dfrac{p}{m-1} & \text{if } i \neq j \end{cases} \qquad i, j = 1, 2, \ldots, m$$

Now we compute Z:

$$Z = \sum_y \sqrt{p(y|x)\, p(y|x')} = \sum_{n=1}^{m} \sqrt{p(b_n|a_i)\, p(b_n|a_j)} \qquad i \neq j$$

$$= \sqrt{p(b_i|a_i)\, p(b_i|a_j)} + \sqrt{p(b_j|a_i)\, p(b_j|a_j)} + \sum_{\substack{n=1 \\ n \neq i \\ n \neq j}}^{m} \sqrt{p(b_n|a_i)\, p(b_n|a_j)}$$

$$\therefore\ Z = \left[(1-p)\frac{p}{(m-1)}\right]^{1/2} + \left[\frac{p}{m-1}(1-p)\right]^{1/2} + (m-2)\left[\frac{p}{m-1}\cdot\frac{p}{m-1}\right]^{1/2}$$

$$= \frac{1}{m-1}\left[(m-2)p + \sqrt{4p(1-p)(m-1)}\right]$$

(b) Noncoherent channel with hard m-ary decision outputs: by problem 2.14 we have

$$P_E = e^{-\mathcal{E}/N_0} \sum_{j=2}^{M} \frac{(-i)^j}{M} \binom{M}{j} e^{\mathcal{E}/jN_0} = p$$

Everything is then the same as in part (a) and z is given by

$$Z = \frac{1}{m-1}\left[(m-2)p + \sqrt{4p(1-p)(m-1)}\right]$$

where p is as given above.

(c) Coherent channel with unquantized outputs:

$$Z = \int_{y^m} \sqrt{p(\underline{y}|x)p(\underline{y}|x')}\, d\underline{y} \qquad x' \neq x$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\Big\{(\pi N_0)^{-m/2} \exp\left[-\frac{1}{N_0}(y_i-\sqrt{\mathcal{E}})^2\right]\exp\left[-\frac{1}{N_0}\sum_{\substack{n=1\\n\neq i}}^{m} y_n^2\right]$$

$$\cdot (\pi N_0)^{-m/2} \exp\left[-\frac{1}{N_0}(y_j-\sqrt{\mathcal{E}})^2\right]\exp\left[-\frac{1}{N_0}\sum_{\substack{n=1\\n\neq j}}^{m} y_n^2\right]\Big\}\, d\underline{y}$$

$$=(\pi N_0)^{-m/2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2N_0}\left[(y_i-\sqrt{\mathcal{E}})^2+(y_j-\sqrt{\mathcal{E}})^2\right]\right\}\exp\left\{-\frac{1}{2N_0}(y_i^2+y_j^2)\right\}$$

$$\cdot \exp\left\{-\frac{1}{N_0}\sum_{\substack{n=1\\n\neq i\\n\neq j}}^{m} y_n^2\right\}\, d\underline{y}$$

$$\therefore Z = \left\{ \prod_{\substack{n=1\\ n\neq i\\ n\neq j}}^{m} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{N_0} y_n^2\right) \frac{dy_n}{\sqrt{\pi N_0}} \right\} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\pi N_0} \exp\left\{-\frac{1}{N_0}\left(y_i^2 - \sqrt{\mathcal{E}}\, y_i + \frac{\mathcal{E}}{2}\right)\right\}$$

$$\cdot \exp\left\{-\frac{1}{N_0}\left(y_j^2 - \sqrt{\mathcal{E}}\, y_j + \frac{\mathcal{E}}{2}\right)\right\} dy_i\, dy_j = \exp\left(\frac{-\mathcal{E}}{2N_0}\right)$$

(d) Noncoherent channel with unquantized outputs:

$$Z = \int_{y^{m+}} \sqrt{p(\mathbf{y}|\mathbf{x})\, p(\mathbf{y}|\mathbf{x}')} \qquad \mathbf{x}' \neq \mathbf{x}$$

$$= \int_0^{\infty} \sqrt{p(y_i|x_i)\, p(y_i|x_j)}\, dy_i \int_0^{\infty} \sqrt{p(y_j|x_i)\, p(y_j|x_j)}\, dy_j$$

$$\cdot \prod_{\substack{n=1\\ n\neq i\\ n\neq j}}^{m} \int_0^{\infty} \sqrt{p(y_n|x_i)\, p(y_n|x_j)}\, dy_n$$

When $n \neq i,\ n \neq j$ $\displaystyle\int_0^{\infty} \sqrt{p(y_n|x_i)\, p(y_n|x_j)}\, dy_n = \int_0^{\infty} p(y_n|x_i)\, dy_n = 1$

Therefore $\displaystyle Z = \left[\int_0^{\infty} \left\{\frac{2}{N_0} y_j^2 \exp\left[-y_j^2\right] e^{-\mathcal{E}/N_0}\, I_0\left(\sqrt{\frac{2\mathcal{E}}{N_0}}\, y_j\right)\right\}^{1/2} dy_j\right]^2$

$$= \frac{2}{N_0} \exp\left(\frac{-\mathcal{E}}{N_0}\right) \left[\int_0^{\infty} y \exp\left(-\frac{1}{2} y^2\right) \left[I_0\left(\sqrt{\frac{2\mathcal{E}}{N_0}}\, y\right)\right]^{1/2} dy\right]^2$$

Now we show that $D = E_{ex}(R)$ satisfies

$$R = \ln m - \mathcal{H}(D/d) - (D/d)\ln(m-1)\ ; \text{ where } d = -\ln Z.$$

$$D = E_{ex}(R) = E_x(\rho) - \rho E_x'(\rho) \quad , \text{ where } \rho \text{ satisfies}$$

$$R = E_x'(\rho) = -\ln\left[\frac{1+(m-1)z^{1/\rho}}{m}\right] - \rho\left[\frac{m}{1+(m-1)z^{1/\rho}}\right]\left[\frac{m-1}{m} z^{1/\rho} \ln z \left(-\frac{1}{\rho^2}\right)\right]$$

$$= \ln m - \ln\left[1+(m-1)z^{1/\rho}\right] + \frac{1}{\rho}\frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \ln z$$

$$= \ln m - \left[1 - \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}}\right] \ln\left[(1+m-1)z^{1/\rho}\right]$$

$$- \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \ln\left[1+(m-1)z^{1/\rho}\right] + \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \ln z^{1/\rho}$$

$$= \ln m + \left[\frac{1}{1+(m-1)z^{1/\rho}}\right] \ln\left(\frac{1}{1+(m-1)z^{1/\rho}}\right) + \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \ln\left(\frac{z^{1/\rho}}{1+(m-1)z^{1/\rho}}\right)$$

$$+ \left[\frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}}\right] \ln(m-1) - \left[\frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}}\right] \ln(m-1)$$

$$= \ln m - \mathcal{H}(\delta_\rho) - \delta_\rho \ln(m-1)$$

where

$$\delta_\rho \triangleq \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \quad \text{and} \quad \mathcal{H}(x) = -x\ln x - (1-x)\ln(1-x)$$

But $D = E_{ex}(R) = E_x(\rho) - \rho R$

$$= -\rho \ln\left[\frac{1+(m-1)z^{1/\rho}}{m}\right] - \rho \ln m$$

$$+ \rho\mathcal{H}(\delta_\rho) \quad \delta_\rho \ln(m-1)$$

or

$$\frac{D}{\rho} = \ln\left[\frac{1}{1+(m-1)z^{1/\rho}}\right] + \ln(m-1)z^{1/\rho} - \ln(m-1)z^{1/\rho} + \mathcal{H}(\delta_\rho) + \delta_\rho \ln(m-1)$$

$$\therefore \frac{D}{\rho} = \ln\left[\frac{(m-1)z^{\rho}}{1+(m-1)z^{1/\rho}}\right] - \ln z^{1/\rho} - (1-\delta_\rho)\ln(m-1) + \mathcal{H}(\delta_\rho)$$

$$= \ln \delta_\rho - \ln z^{1/\rho} - (1-\delta_\rho)\ln(m-1) - \delta_\rho \ln \delta_\rho - (1-\delta_\rho)\ln(1-\delta_\rho)$$

$$= (1-\delta_\rho)\ln\left[\frac{\delta_\rho}{(1-\delta_\rho)(m-1)}\right] - \ln z^{1/\rho}$$

Note that

$$\frac{\delta_\rho}{(m-1)(1-\delta_\rho)} = \frac{(m-1)z^{1/\rho}}{1+(m-1)z^{1/\rho}} \cdot \frac{1+(m-1)z^{1/\rho}}{(m-1)} = z^{1/\rho}$$

thus

$$\frac{D}{\rho} = (1-\delta_\rho)\ln z^{1/\rho} - \ln z^{1/\rho} = -\frac{1}{\rho}\delta_\rho \ln z$$

or $D = -\delta_\rho \ln z = \delta_\rho d \implies \delta_\rho = D/d$

therefore

$$R = \ln m - \mathcal{H}(D/d) - (D/d)\ln(m-1) .$$

<u>3.26</u> Choosing messages equiprobably we have

$$P_E = \frac{1}{M}\sum_{m=1}^{M} P_{E_m}$$

For $\underline{y} \in \bar{\Lambda}_m$ we should have $\dfrac{\tilde{p}_N(\underline{y}|\underline{x}_{m'})}{\tilde{p}_N(\underline{y}|\underline{x}_m)} \geq 1$; for some $m' \neq m$.

Define $\tilde{\Lambda}_m = \left\{\underline{y} : \sum_{m' \neq m}\left[\dfrac{\tilde{p}_N(\underline{y}|\underline{x}_{m'})}{\tilde{p}_N(\underline{y}|\underline{x}_m)}\right]^{\frac{1}{1+\rho}} \geq 1\right\}$; for $\rho > 0$.

Then

$$P_{E_m} = \sum_{\underline{y} \in \bar{\Lambda}_m} p_N(\underline{y}|\underline{x}_m) \leq \sum_{\underline{y} \in \tilde{\Lambda}_m} p_N(\underline{y}|\underline{x}_m) = \sum_{\underline{y}} f(\underline{y})\, p_N(\underline{y}|\underline{x}_m)$$

where

$$f(\underset{\sim}{y}) = \begin{cases} 1 & ; \underset{\sim}{y} \in \tilde{\Lambda}_m \\ 0 & ; \underset{\sim}{y} \notin \tilde{\Lambda}_m \end{cases}$$ and since $\bar{\Lambda}_m \subseteq \tilde{\Lambda}_m$.

Furthermore

$$f(\underset{\sim}{y}) \leq \left\{ \sum_{m' \neq m} \left[\frac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})}{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)} \right]^{\frac{1}{1+\rho}} \right\}^{\rho} \quad ; \rho > 0$$

Then

$$P_E \leq \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y}} p_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \left\{ \sum_{m' \neq m} \left[\frac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})}{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)} \right]^{\frac{1}{1+\rho}} \right\}^{\rho} \quad ; \rho > 0$$

$$\leq \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y}} p_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{-\frac{\rho}{1+\rho}} \left[\sum_{m' \neq m} \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})^{\frac{1}{1+\rho}} \right]^{\rho} \quad ; \rho > 0$$

Note that according to the probability measure of the code selection, the random variables $p_N(\underset{\sim}{y}|\underset{\sim}{x}_m)\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{-\frac{\rho}{1+\rho}}$ and $\left[\sum_{m' \neq m} \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})^{1/1+\rho}\right]$ are indenpent. Averaging P_E over the ensemble of codes we obtain, for $0 \leq \rho \leq 1$

$$\overline{P_E} \leq \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y}} \overline{p_N(\underset{\sim}{y}|\underset{\sim}{x}_m)\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{-\frac{\rho}{1+\rho}}} \; \overline{\left[\sum_{m' \neq m} \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'})^{\frac{1}{1+\rho}} \right]^{\rho}}$$

$$\leq \sum_{\underset{\sim}{y}} \sum_{\underset{\sim}{x}} q(\underset{\sim}{x}) p_N(\underset{\sim}{y}|\underset{\sim}{x}) \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x})^{-\frac{\rho}{1+\rho}} \left[(M-1) \sum_{\underset{\sim}{x}'} q(\underset{\sim}{x}') \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}')^{\frac{1}{1+\rho}} \right]^{\rho}$$

Define $$Q(\underset{\sim}{x}|\underset{\sim}{y}) \triangleq \frac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x})^{1/1+\rho} q(\underset{\sim}{x})}{\sum_{\underset{\sim}{x}'} q(\underset{\sim}{x}') \tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}')^{1/1+\rho}}$$ and note that

$Q(\underset{\sim}{x}|\underset{\sim}{y}) > 0$ and $\sum_{\underset{\sim}{x}} Q(\underset{\sim}{x}|\underset{\sim}{y}) = 1$ and then

$$\overline{P_E} \leq M^{\rho} \sum_{\underset{\sim}{y}} q(\underset{\sim}{x})^{1+\rho} \frac{p_N(\underset{\sim}{y}|\underset{\sim}{x})}{Q^{\rho}(\underset{\sim}{x}|\underset{\sim}{y})}$$

since $\sum_{\underset{\sim}{x}'} q(\underset{\sim}{x}')\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x}')^{1/1+\rho} = \dfrac{\tilde{p}_N(\underset{\sim}{y}|\underset{\sim}{x})^{1/1+\rho}q(\underset{\sim}{x})}{Q(\underset{\sim}{x}|\underset{\sim}{y})}$ is independent of $\underset{\sim}{x}$.

Then
$$\overline{P}_E \leq M^\rho \left\{ \sum_y \sum_x q(x)^{1+\rho} \frac{p(y|x)}{q^\rho(x|y)} \right\}^N$$

where
$$q(x|y) = \frac{\tilde{p}(y|x)^{1/1+\rho} q(x)}{\sum_x q(x)\tilde{p}(y|x)^{1/1+\rho}} \quad (1) \text{ (the channel is memoryless)}$$

and $Q(\underset{\sim}{x}|\underset{\sim}{y})$ is to be the n-fold product of $q(x|y)$.

$R = \dfrac{\ln M}{M}$ so $M = e^{RN}$ and thus

$$\overline{P}_E \leq e^{\rho RN} e^{-F(\rho, q, p, \tilde{p})N} = e^{-N\{F(\rho,q,p,\tilde{p}) - R\rho\}}$$

where $F(\rho, q, p, \tilde{p}) \triangleq -\ln \left\{ \sum_y \sum_x q(x)^{1+\rho} \dfrac{p(y|x)}{q^\rho(x|y)} \right\}$

and $q(x|y)$ is related to $\tilde{p}(\ |x)$ through (1).

<u>3.27</u> The conditional probabilities for this channel are given by

$$p(y_j|x_i) = \begin{cases} y_j\, e^{-y_j^2/2} & ; i \neq j \\ \dfrac{y_j}{1+E_s/N_0}\, e^{-\frac{1}{2}\frac{y_j^2}{1+E_s/N_0}} & ; i = j \end{cases}$$

The decision rule is the same as for the noncoherent channel, since the same decoder is optimum. Due to symmetry we assume, without loss of generality, that x_1 is sent.

The probability of a correct decision given x_1 is sent is then

$$P_{C_1} = P_r\{y_1 \geq y_n \text{ for all } n \neq 1 \mid x_1\}$$

$$= \int_0^\infty p(y_1|x_1)\left[\prod_{n=2}^{m} \int_0^{y_1} p(y_n|x_1)\, dy_n\right] dy_1$$

For $n \neq 1$, $\displaystyle\int_0^{y_1} p(y_n|x_1)dy_n = \int_0^{y_1} y_n\, e^{-y_n^2/2} dy_n = \left(1 - e^{-y_1^2/2}\right)$

Therefore

$$P_{C_1} = \int_0^\infty p(y_1|x_1)\left(1 - e^{-y_1^2/2}\right)^{m-1} dy_1$$

$$= \frac{1}{1+\frac{\bar{E}_s}{N_0}} \int_0^\infty y \exp\left\{-\frac{y^2}{2}\left(\frac{1}{1+\bar{E}_s/N_0}\right)\right\}\left[1 - e^{-y^2/2}\right]^{m-1} dy$$

$$\triangleq 1 - p$$

Except for the value of p, everything is the same as in problem 3.25(b) and so

$$Z = \frac{(m-2)p + \sqrt{4p(1-p)(m-1)}}{m-1}$$

For the case of unquantized output vectors we have

$$Z = \int_{y^N} \prod_{n=1}^{m} p(y_n|x_i)p(y_n|x_j)\, d\mathbf{y} \qquad i \neq j$$

$$= \int_0^\infty \sqrt{p(y_j|x_i)p(y_j|x_j)}\, dy_i \int_0^\infty \sqrt{p(y_j|x_j)\, p(y_j|x_i)}\, dy_j$$

$$\cdot \prod_{\substack{n=1 \\ n\neq i,j}}^{m} \int_0^\infty \sqrt{p(y_n|x_i)\, p(y_n|x_j)}\, dy_n$$

For $n \neq i, n \neq j$ the last product is unity therefore

$$z = \left[\int_0^\infty \sqrt{p(y_j|x_j)\, p(y_j|x_i)}\, dy_j \right]^2$$

$$= \left\{ \int_0^\infty \left\{ \frac{y^2}{1+\overline{E}_s/N_0} \exp\left[-\frac{y^2}{2}\left(1 + \frac{1}{1+\overline{E}_s/N_0} \right) \right] \right\}^{1/2} dy \right\}^2$$

$$= \frac{1}{1+\overline{E}_s/N_0} \left[\int_0^\infty y \exp\left\{ -\frac{y^2}{4}\left(\frac{2+\overline{E}_s/N_0}{1+\overline{E}_s/N_0} \right) \right\} dy \right]^2$$

$$= \frac{1+\overline{E}_s/N_0}{\left(1+\overline{E}_s/2N_0\right)^2} .$$

<u>3.28</u> $E_{ex}(R) = E_x(\rho) - \rho E_x'(\rho) = E_x(\rho) - \rho R$

for $\lim\limits_{\rho\to\infty} E_x'(\rho) < R < E_x'(\rho)\Big|_{\rho=1}$

By symmetry $q^*(x) = 1/Q$ for all x, and

$$E_x(\rho) = -\rho \ln \sum_x \sum_{x'} q^*(x) q^*(x') \left[\sum_y \sqrt{p(y|x)\, p(y|x')} \right]^{1/\rho}$$

$$= -\rho \ln \sum_x \sum_{x'} \frac{1}{Q^2} \exp\left\{ -\frac{1}{\rho} d(x, x') \right\} = -\rho \ln \frac{1}{Q} \sum_{i=1}^{Q} e^{-d_i/\rho}$$

and then

$$E_x'(\rho) = -\ln \frac{1}{Q} \sum_{i=1}^{Q} e^{-d_i/\rho} - \frac{1}{\rho} \frac{\sum_{i=1}^{Q} d_i e^{-d_i/\rho}}{\sum_{i=1}^{Q} e^{-d_i/\rho}}$$

then

$$D = E_{ex}(R) = \frac{\sum_{i=1}^{Q} d_i e^{-d_i/\rho}}{\sum_{i=1}^{Q} e^{-d_i/\rho}} = \frac{\sum_{i=1}^{Q} d_i e^{s d_i}}{\sum_{i=1}^{Q} e^{s d_i}} \; ; \; s = -1/\rho$$

$$R = E_x'(\rho) = -\frac{1}{\rho} D - \ln \frac{1}{Q} \sum_{i=1}^{Q} e^{-d_i/\rho} = sD - \ln\left(\frac{1}{Q}\sum_{i=1}^{Q} e^{s d_i}\right)$$

and for $1 \le \rho \le \infty$ we have $s = -\frac{1}{\rho} \in [-1, 0]$.

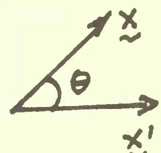
For the signal set of fig. 2.12b we have

$$d(\underline{x}, \underline{x}') = -\ln \iint \sqrt{p(\underline{y}|\underline{x})\, p(\underline{y}|\underline{x}')}\, d\underline{y} = -\ln \iint \frac{1}{\pi N_0} e^{-\frac{\|\underline{y}-\underline{x}\|^2 + \|\underline{y}-\underline{x}'\|^2}{2N_0}} d\underline{y}$$

$$= -\ln \iint \frac{1}{\pi N_0} \exp\left\{-\frac{\|\underline{y} - \frac{1}{2}(\underline{x}+\underline{x}')\|^2}{N_0}\right\} \exp\left\{-\frac{\|\underline{x}-\underline{x}'\|^2}{4N_0}\right\} d\underline{y}$$

$$= -\ln \exp\left\{-\frac{\|\underline{x}-\underline{x}'\|^2}{4N_0}\right\} = \frac{\|\underline{x}-\underline{x}'\|^2}{4N_0}$$

$$\|\underline{x}-\underline{x}'\|^2 = \|\underline{x}\|^2 + \|\underline{x}'\|^2 - 2(\underline{x}, \underline{x}') = 2\mathcal{E}_s(1-\cos\theta)$$



then

$$d_k = \frac{\mathcal{E}_s(1-\cos k\frac{\pi}{8})}{2N_0} \qquad k = 1, 2, \ldots, 16$$

<u>3.29</u> (a) The optimum decision regions are

$$\Lambda_m = \left\{\underline{y} : p_N(\underline{y}|\underline{x}_m) > p_N(\underline{y}|\underline{x}_{m'}),\ \text{all } m' \ne m\right\}$$

$$= \left\{\underline{y} : \max_{m'} p_N(\underline{y}|\underline{x}_{m'}) = p_N(\underline{y}|\underline{x}_m)\right\}$$

Hence

$$P_c = \frac{1}{M}\sum_{m=1}^{M} P_{c_m} = \frac{1}{M}\sum_{m=1}^{M} \sum_{\underline{y}\in\Lambda_m} p_N(\underline{y}|\underline{x}_m)$$

$$= \frac{1}{M}\sum_{y_N} \max_m p_N(\underline{y}|\underline{x}_m)$$

(b) for any $\beta > 0$

$$\max_m P_N(\underset{\sim}{y}|\underset{\sim}{x}_m) = \left(\max_m P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}\right)^\beta \le \left(\sum_{m=1}^{M} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}\right)^\beta$$

$$\therefore \quad P_C \le \frac{1}{M} \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \left(\sum_{m=1}^{M} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}\right)^\beta$$

(c) using Jensen's inequality we have

$$\overline{P_C} \le \overline{\frac{1}{M} \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \left(\sum_{m=1}^{M} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}\right)^\beta} \le \frac{1}{M} \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \left(\overline{\sum_{m=1}^{M} P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}}\right)^\beta$$

$$= \frac{1}{M^{1-\beta}} \sum_{\underset{\sim}{y} \in \mathcal{Y}^N} \left(\overline{P_N(\underset{\sim}{y}|\underset{\sim}{x}_m)^{1/\beta}}\right)^\beta = \frac{1}{M^{1-\beta}} \left\{\sum_{y \in \mathcal{Y}} \left(\sum_x q(x) p(y|x)^{1/\beta}\right)^\beta\right\}^N$$

$$\therefore \quad \overline{P_C} \le M^\rho \left\{\sum_y \left(\sum_x q(x) p(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right\}^N$$

$$= e^{-N\{E_0(\rho, \underset{\sim}{q}) - \rho R\}}$$

where $\rho = \beta - 1$
and $\rho \in [-1, 0]$ for $0 \le \beta \le 1$

and where

$$E_0(\rho, \underset{\sim}{q}) = -\ln \left\{\sum_y \left(\sum_x q(x)\, p(y|x)^{1/1+\rho}\right)^{1+\rho}\right\}$$

(d) for $I(\underset{\sim}{q}) > 0$ we have

$$\frac{\partial E_0(\rho, \underset{\sim}{q})}{\partial \rho} > 0 \; ; \; \left.\frac{\partial E_0(\rho, \underset{\sim}{q})}{\partial \rho}\right|_{\rho=0} = I(\underset{\sim}{q}) \; ; \; \frac{\partial^2 E_0(\rho, \underset{\sim}{q})}{\partial \rho^2} \le 0$$

and

$$E_0(\rho, \underset{\sim}{q})\Big|_{\rho=0} = 0 < E_0(\rho, \underset{\sim}{q}) \text{ for } \rho > 0$$

and $E_0(\rho,q)\Big|_{\rho=0} = 0 > E_0(\rho,q)$ for $\rho<0$

Choose q such that $I(q) = C$:
For $R > C$ we see that for some $\rho \in [-1,0]$ we have $E_0(\rho,q) - \rho R > 0$ and clearly

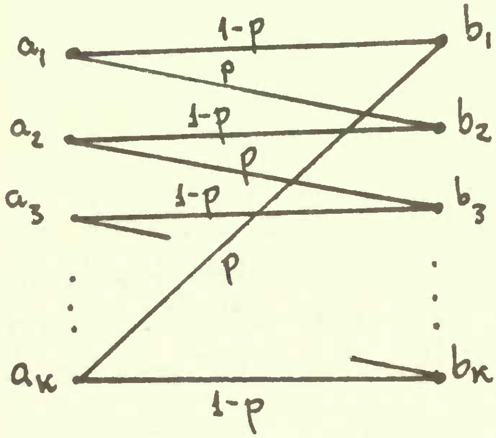
$$E_{sc}(R) = \max_{-1\le\rho\le 0}\left(E_0(\rho,q) - \rho R\right) > 0$$

(e) this result tells us there exist codes with probability of correct decoding that satisfies

$$P_c \le e^{-NE_{sc}(R)} \quad ; \text{with } E_{sc}(R) > 0 \text{ for } R > C.$$

There are many <u>bad</u> codes we can think of that do worse.
What we are interested in is a bound on P_c for <u>all</u> codes.

<u>3.30</u>



Clearly by symmetry

$$q^*(a_k) = \frac{1}{K}$$

for $k = 1, 2, \ldots, K$

Let $\bar{p} = 1-p$

$$E_0(\rho) = \max_q E_0(\rho,q) = E_0(\rho,q^*) = -\ln \sum_{j=1}^{K}\left(\sum_{k=1}^{K} q(a_k)\, p(b_j|a_k)^{\frac{1}{1+\rho}}\right)^{1+\rho}$$

$$= -\ln \sum_{j=1}^{K}\left[\frac{1}{K}\left(p^{\frac{1}{1+\rho}} + \bar{p}^{1/1+\rho}\right)\right]^{1+\rho} = -\ln K\left[\frac{1}{K}\left(p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}\right)\right]^{1+\rho}$$

$$\therefore\ E_0(\rho) = \rho \ln K - (1+\rho) \ln\left[p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right]$$

$$E_0'(\rho) = \ln K + \frac{1}{1+\rho} \frac{p^{\frac{1}{1+\rho}} \ln p + \bar{p}^{\frac{1}{1+\rho}} \ln \bar{p}}{p^{1/1+\rho} + \bar{p}^{1/1+\rho}} - \ln\left(p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right)$$

Let $\delta \triangleq \dfrac{p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \Rightarrow \dfrac{\bar{p}^{\frac{1}{1+\rho}}}{\bar{p}^{\frac{1}{1+\rho}} + p^{\frac{1}{1+\rho}}} = \dfrac{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} - p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{1/1+\rho}} = 1 - \delta$

$$\begin{aligned}
\therefore\ E_0'(\rho) &= \ln K + \frac{1}{1+\rho}\left[\delta \ln p + (1-\delta) \ln \bar{p} \right] - \ln\left[p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right] \\
&= \ln K + \delta \ln p^{\frac{1}{1+\rho}} + (1-\delta) \ln \bar{p}^{\frac{1}{1+\rho}} - \delta \ln\left[p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right] \\
&\quad - (1-\delta) \ln\left[p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right] \\
&= \ln K + \delta \ln \delta + (1-\delta) \ln (1-\delta) \\
&= \ln K - \mathcal{H}(\delta)
\end{aligned}$$

(b) $C = \left. \dfrac{\partial E_0(\rho)}{\partial \rho} \right|_{\rho=0} = \left. \left[\ln K - \mathcal{H}(\delta) \right] \right|_{\rho=0} = \ln K - \mathcal{H}(p)$

(c) Define a set of codewords $\mathcal{S}$ such that $\underset{\sim}{x}_2 \in \mathcal{S}$ results in $\underset{\sim}{y}$. For each y_n of $\underset{\sim}{y}$ there are two inputs that can result in y_n. Therefore there are 2^N distinct codewords that can result in $\underset{\sim}{y}$. But there are K^N possible distinct codewords. Thus

$$P_r\left\{ \underset{\sim}{x}_2 \in \mathcal{S} \right\} = \frac{2^N}{K^N} = \left(\frac{2}{K} \right)^N$$

(d) Let $E_k = \left\{ \underset{\sim}{x}_k \in \mathcal{S} \right\}$, $k = 2, 3, \ldots, M$

$$\Pr\left\{\bigcup_{k=2}^{M} E_k\right\} \le \sum_{k=2}^{M} \Pr\left\{\underline{x}_k \in \mathcal{S}\right\} = (M-1)\left(\frac{2}{K}\right)^N < M\left(\frac{2}{K}\right)^N$$

(e) Let $p = \bar{p} = 1/2$

$$E_0(\rho) = \rho \ln K - (1+\rho) \ln 2\left(\frac{1}{2}\right)^{\frac{1}{1+\rho}} = \rho \ln \frac{K}{2}$$

$$E(R) = \max_{0\le\rho\le 1}\left\{E_0(\rho) - \rho R\right\} = \max_{0\le\rho\le 1}\left\{\rho \ln \frac{K}{2} - \rho R\right\} = \ln \frac{K}{2} - R$$

for $R < C = \ln \frac{K}{2}$

$$\therefore \quad \overline{P_{e,m}} \le e^{-NE(R)} = e^{-N\left\{\ln\left(\frac{K}{2}\right) - R\right\}}$$

in (d) we get

$$\Pr\left\{\bigcup_{k=2}^{M} E_k\right\} < M\left(\frac{2}{K}\right)^N = e^{RN}\left(\frac{K}{2}\right)^{-N} = e^{RN} e^{-N\ln\frac{K}{2}}$$

$$= e^{-N\left\{\ln\frac{K}{2} - R\right\}}$$

which is exactly the same exponent.

(f) For $p = 0$, $E_0(\rho) = \rho \ln K$

$$E(R) = \max_{0\le\rho\le 1}\left\{\rho \ln K - \rho R\right\} = \ln K - R$$

for $R < C = \ln K$ (noiseless channel).

<u>3.31</u> Suppose $\underline{x}_1 = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1N} \end{bmatrix}$ is sent and define events

$$E_i \triangleq \left\{\|\underline{y} - \underline{x}_i\| \le \|\underline{y} - \underline{x}_1\|\right\},$$ where $\underline{y}$ is the received vector.

The probability of error given that $\underline{x}_1$ was sent is then:

$$P_{E,1} = P_r\left\{\bigcup_{i=2}^{M} E_i\right\} \le \sum_{i=2}^{M} P_r\{E_i\}$$; where $P_r\{E_i\}$ is the prob. of error for two signals only.

but

$$P_r\{E_i\} = \operatorname{erfc}\left(\frac{\|\underline{x}_i - \underline{x}_1\|}{\sqrt{2N_0}}\right) < e^{-\frac{\|\underline{x}_i - \underline{x}_1\|^2}{4N_0}}$$ (using the "hint")

$$\therefore P_{e,1} \le \sum_{i=2}^{M} \exp\left\{-\frac{\|\underline{x}_i - \underline{x}_1\|^2}{4N_0}\right\}$$ and now we average

the bound over the ensemble of codewords

$$\overline{P_{E,1}} \le (M-1)\,\overline{e^{-\frac{\|\underline{x}_i - \underline{x}_1\|^2}{4N_0}}} \le M \prod_{k=1}^{N} \overline{e^{-\frac{(x_{ik}-x_{1k})^2}{4N_0}}}$$

$$\le M\left\{\overline{e^{-\frac{(x_{ik}-x_{1k})^2}{4N_0}}}\right\}^N \le M\,2^{-C_0 N}$$

where $C_0 \triangleq -\log\left[\overline{\exp\left\{-\frac{(x_{ik}-x_{1k})^2}{4N_0}\right\}}\right]$

(a) Let $z \triangleq x_{ik} - x_{1k}$ $\longrightarrow$ z is a gaussian random variable with zero mean and variance $\sigma_z^2 = 2\mathcal{E}$.

Then

$$E\left\{e^{-\frac{z^2}{4N_0}}\right\} = \int_{-\infty}^{\infty} e^{-z^2/4N_0} \frac{1}{\sqrt{4\pi\mathcal{E}}} e^{-z^2/4\mathcal{E}}\, dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi\mathcal{E}}} e^{-\frac{z^2}{2}\left(\frac{1}{2N_0}+\frac{1}{2\mathcal{E}}\right)} dz = \left(1+\frac{\mathcal{E}}{N_0}\right)^{-1/2}$$

Hence

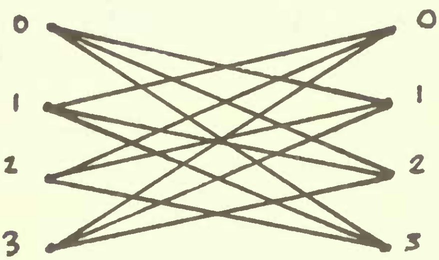
$$C_0 = \frac{1}{2}\log\left(1+\frac{\mathcal{E}}{N_0}\right)$$

(b) $E\left\{ e^{\frac{-(x_{ik}-x_{jk})^2}{4N_0}} \right\} = \frac{1}{2} + \frac{1}{2} e^{-\mathcal{E}/N_0}$

$$\therefore \quad C_0 = -\log\left(\frac{1}{2} + \frac{1}{2} e^{-\mathcal{E}/N_0}\right) = \log 2 - \log\left(1 + e^{-\mathcal{E}/N_0}\right)$$

$$= 1 - \log\left(1 + e^{-\mathcal{E}/N_0}\right)$$

3.32



$$p(y|x) = \frac{1}{3}(1 - \delta_{yx})$$

$$\delta_{yx} = \begin{cases} 1 & ; y = x \\ 0 & ; y \neq x \end{cases}$$

(a) By symmetry $q^*(x) = 1/4$ for $x = 0, 1, 2, 3$

$$C = \sum_x \sum_y p(y|x)\, q^*(x) \ln \frac{p(y|x)}{p(y)} = \sum_{x \neq y}\sum p(y|x) q^*(x) \ln \frac{p(y|x)}{p(y)}$$

$$= 12 \times \frac{1}{3} \times \frac{1}{4} \times \ln \frac{1/3}{1/4} = \ln 4/3$$

(b) $$E_0(\rho) = \max_q E_0(\rho, q) = E_0(\rho, q^*) = -\ln \sum_{y=0}^{3} \left[\sum_{x=0}^{3} \frac{1}{4} p(y|x)^{\frac{1}{1+\rho}} \right]^{1+\rho}$$

$$= -\ln 4 \left[\frac{3}{4} \left(\frac{1}{3}\right)^{\frac{1}{1+\rho}} \right]^{1+\rho} = -\ln \frac{4}{3}\left(\frac{3}{4}\right)^{1+\rho} = \rho \ln \frac{4}{3}$$

(c) $$E(R) = \max_{0 \le \rho \le 1} \left[E_0(\rho) - \rho R \right] = \ln \frac{4}{3} - R \quad ; \quad R < C = \ln \frac{4}{3}$$

$$E_x(\rho, q^*) = -\rho \ln \left\{ \sum_x \sum_{x'} q^*(x) q^*(x') \left[\sum_y \sqrt{p(y|x)\, p(y|x')} \right]^{1/\rho} \right\}$$

$$= -\rho \ln \left\{ \frac{1}{4} + \frac{12}{16} \left[1/3 + 1/3 \right]^{1/\rho} \right\} = -\rho \ln \left\{ \frac{1}{4} \left(1 + 3 \left(\frac{2}{3}\right)^{1/\rho} \right) \right\}$$

$$E_{ex}(R) = \max_{\rho \geq 1} [E_x(\rho, q^*) - \rho R]$$

$$R_{cr} = \left.\frac{dE_x(\rho, q^*)}{d\rho}\right|_{\rho=1} = -\ln\left\{\frac{1}{4}\left(1 + 3\left(\frac{2}{3}\right)^{1/\rho}\right)\right\} - \frac{\rho\left(\frac{3}{4}\right)\left(\frac{2}{3}\right)^{\frac{1}{\rho}} \ln\left(\frac{2}{3}\right)\left(-\frac{1}{\rho^2}\right)}{\frac{1}{4}\left(1 + 3\left(\frac{2}{3}\right)^{1/\rho}\right)}$$

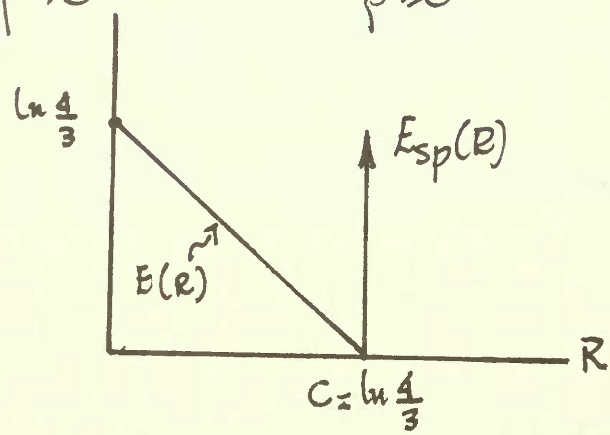
$$= \ln\frac{4}{3} - \frac{2}{3}\ln\frac{3}{2}$$

$$R_{x,\infty} = \lim_{\rho \to \infty} \frac{E_x(\rho, q^*)}{\rho} = \lim_{\rho \to \infty} \ln\left[\frac{1}{4}\left\{1 + 3\left(\frac{2}{3}\right)^{1/\rho}\right\}\right] = 0$$

$$E_{ex}(0) = -\sum_x \sum_{x'} q^*(x) q^*(x') \ln \sum_y \sqrt{p(y|x)p(y|x')}$$

$$= -\left(\frac{1}{4}\right)^2 \times 12 \ln\left[\frac{1}{3} + \frac{1}{3}\right] = \frac{3}{4}\ln\frac{3}{2}$$

$$E_{sp}(R) = \max_{\rho \geq 0} \{E_0(\rho) - \rho R\} = \max_{\rho \geq 0} \{\rho(C - R)\} = \infty \ ; \text{ for } R < C$$



3.33 (a) Consider the binary codes of 2^j codevectors of dimensionality $N-(K-j)$, as we selected. This code is clearly linear since our original code was linear. Now denote d^j_{min} as the minimum weight vector in this 2^j code. Then clearly $d_{min} \leq d^j_{min}$ (the chosen code is a subset of the original).

By lemma 3.7.1 for any binary code of 2^j codevectors of dimensionality $(N-K-j)$ we have

$$\frac{d_{min}}{N-(K-j)} \leq \frac{d^j_{min}}{N-(K-j)} \leq \frac{1}{2}\left(\frac{2^j}{2^j-1}\right) \quad \text{or } d_{min} \leq \frac{N-(K-j)}{2}\left(\frac{2^j}{2^j-1}\right)$$

(b) We know that $N \leq M$ and $\frac{d_{min}}{N} \leq \frac{1}{2}$, thus

$K = \log M \geq \log N$, therefore if we choose $0 \leq \alpha_N \leq 1$ such that $\log N - \alpha_N$ is an integer and then choose $j = \log N - \alpha_N$ where $1 \leq j \leq k$, by part (a) we have

$$d_{min} \leq \frac{N-K+\log N-\alpha_N}{2}\left(\frac{N\,2^{-\alpha_N}}{N\,2^{-\alpha_N}-1}\right) \qquad (1)$$

now $R = \frac{\ln M}{M} = \frac{K \ln 2}{N}$ or $K = \frac{NR}{\ln 2}$

Substituting for K and dividing both sides of (1) by N:

$$\frac{d_{min}}{N} \leq \left[\frac{1}{2} - \frac{R}{2\ln 2} + \frac{\log N - \alpha_N}{N}\right]\left(\frac{N}{N-2^{\alpha_N}}\right)$$

for N sufficiently large we have

$$\frac{d_{min}}{N} \leq \frac{1}{2} - \frac{R}{2\ln 2} + O\left(\frac{1}{N}\right) \qquad \text{where } O\left(\tfrac{1}{N}\right) \xrightarrow{N\to\infty} 0$$

(c) the only restriction we made for the linear code was $d_{min} \leq d^j_{min}$. But for any binary code it is true than $d_{min} \leq d_{min}$ (some subset of the code). Now consider the special subset as 2^j codevectors with first $K-j$ bits equal to zero. Then d_{min} (binary code) $\leq d_{min}$ (special subset). If we remove the first $K-j$ zeros from all vectors of this subset then d_{min} (special subset) $= d_{min}$ (subset with $K-j$ bits zero removed). This is true for any K, j so we are done here.

3.34 (a) Choosing a binary vector causes the removal of at most $\sum_{i=0}^{d-1}\binom{N}{i}$ vectors from the list.

So after choosing M' codewords if

$$M'\sum_{i=0}^{d-1}\binom{N}{i} < 2^N$$

then it is still possible to choose another code vector. Thus, the procedure terminates when

$$M\sum_{i=0}^{d-1}\binom{N}{i} \geq 2^N$$

or

$$M \geq \frac{2^N}{\sum_{i=0}^{d-1}\binom{N}{i}}$$

(b) from problem 3.19(a), for $\eta = \sum_{n=1}^{N} y_n$, where $y_n = 0$ or 1 with probabilities $(1-p)$ and p respectively, we had

$$P_r\{\eta > \theta\} \leq (1-p)^{N-\theta} p^{\theta} e^{N\mathcal{H}(\theta/N)}$$

but

$$P_r\{\eta = i\} = \binom{N}{i}(1-p)^{N-i}p^i \quad \text{therefore}$$

$$P_r\{\eta \geq d\} = \sum_{i=d}^{N}\binom{N}{i}(1-p)^{N-i}p^i \leq (1-p)^{N-d}p^d e^{N\mathcal{H}(\frac{d}{N})}$$

Note that $d \leq N/2$; thus for $\frac{1}{2} \leq p < 1$

$$\sum_{i=0}^{d-1}\binom{N}{i}(1-p)^{N-i}p^i \leq \sum_{i=d}^{N}\binom{N}{i}(1-p)^{N-i}p^i$$

Then $\sum_{i=0}^{d-1}\binom{N}{i}(1-p)^{N-i}p^i \le (1-p)^{N-d}p^d e^{N\mathcal{H}(d/N)}$

and for $p = 1/2$ we get

$$\sum_{i=0}^{d-1}\binom{N}{i}\left(\frac{1}{2}\right)^N \le \left(\frac{1}{2}\right)^N e^{N\mathcal{H}(d/N)}$$

or $\sum_{i=0}^{d-1}\binom{N}{i} \le e^{N\mathcal{H}(\frac{d}{N})} \Longrightarrow \sum_{i=0}^{d_{min}-1}\binom{N}{i} \le e^{N\mathcal{H}(d_{min}/N)}$

(c) for $R < \ln 2$, let $\delta < 1/2$ satisfy $\mathcal{H}(\delta) = \ln 2 - R$.
Choose d_{min} satisfying $\frac{d_{min}-1}{N} < \delta \le \frac{d_{min}}{N}$

Then from (a)&(b) we can find a code of minimum distance d_{min} with $M \ge e^{N[\ln 2 - \mathcal{H}(\delta)]} = e^{NR}$ codevectors.

(d) from (3.4.8) and (3.7.18)

$$-\frac{d_{min}}{N}\ln z \ge -\delta \ln z$$

for $z < 1$ then

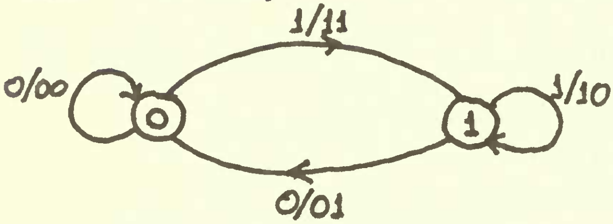
$$\frac{d_{min}}{N} \ge \delta$$

where $R = \ln 2 - \mathcal{H}(\delta)$ and $\delta = \frac{z^{1/\rho}}{1+z^{1/\rho}} < \frac{1}{2}$.

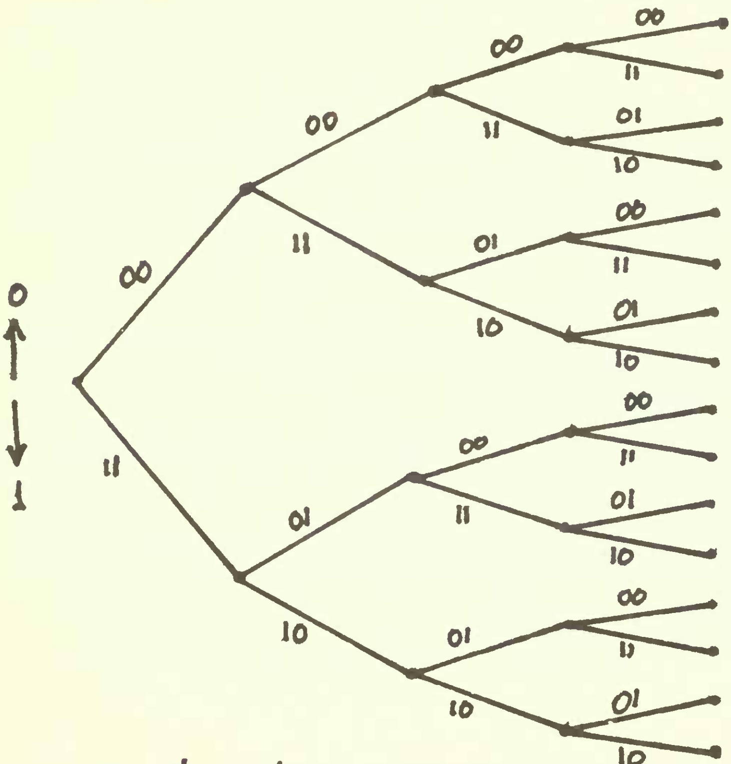
— // —

Chapter 4

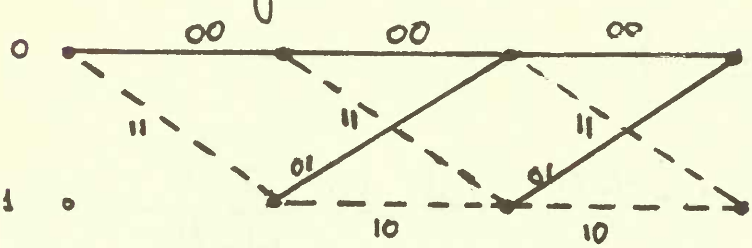
4.1 $K=2$, $r=1/2$. States: $\{0,1\}$.



state diagram

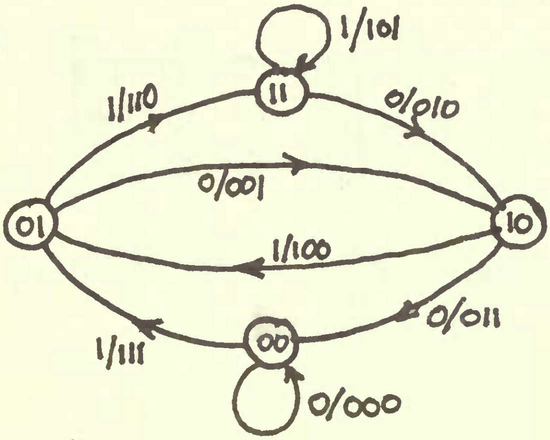


tree diagram

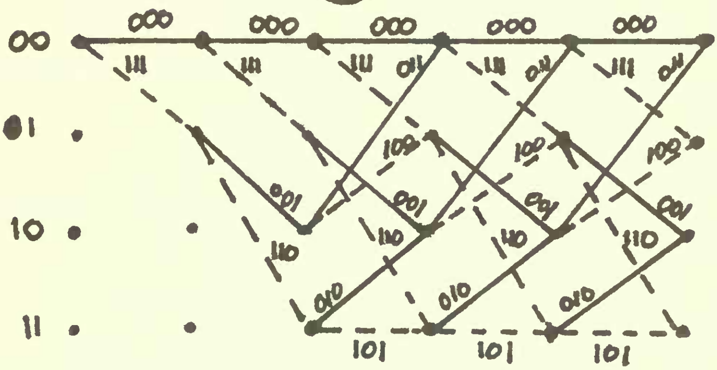


trellis diagram

4.2 $K=3$; $R=1/3$. States: {00, 01, 10, 11}

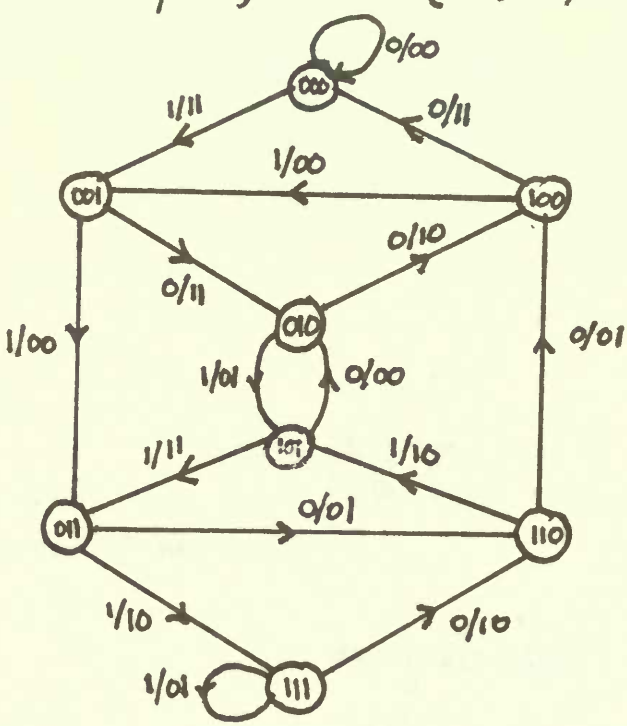


state diagram



trellis diagram

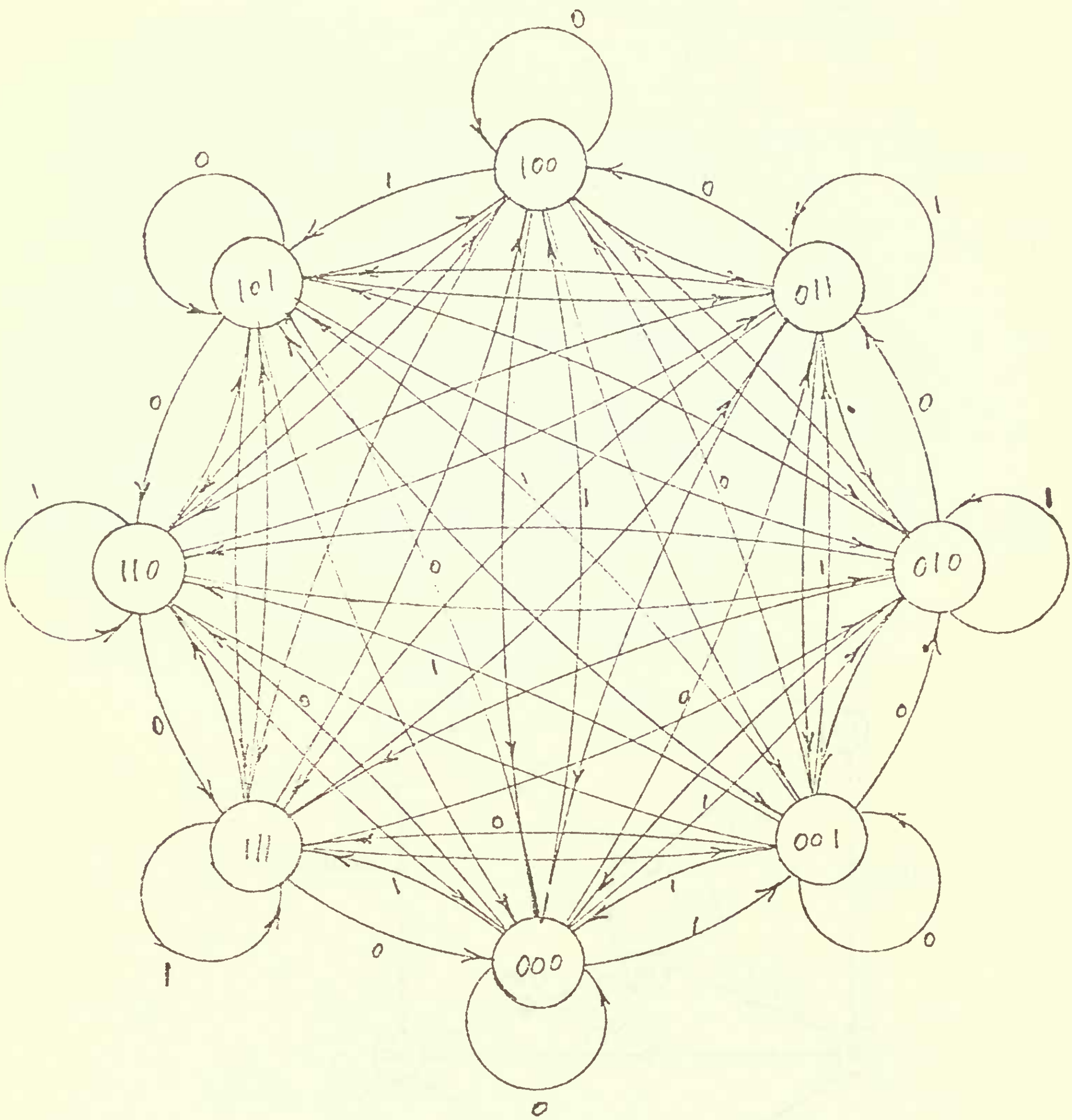
4.3 $K=4$, $R=1/2$; States: {000, 001, 010, 011, 100, 101, 110, 111}.



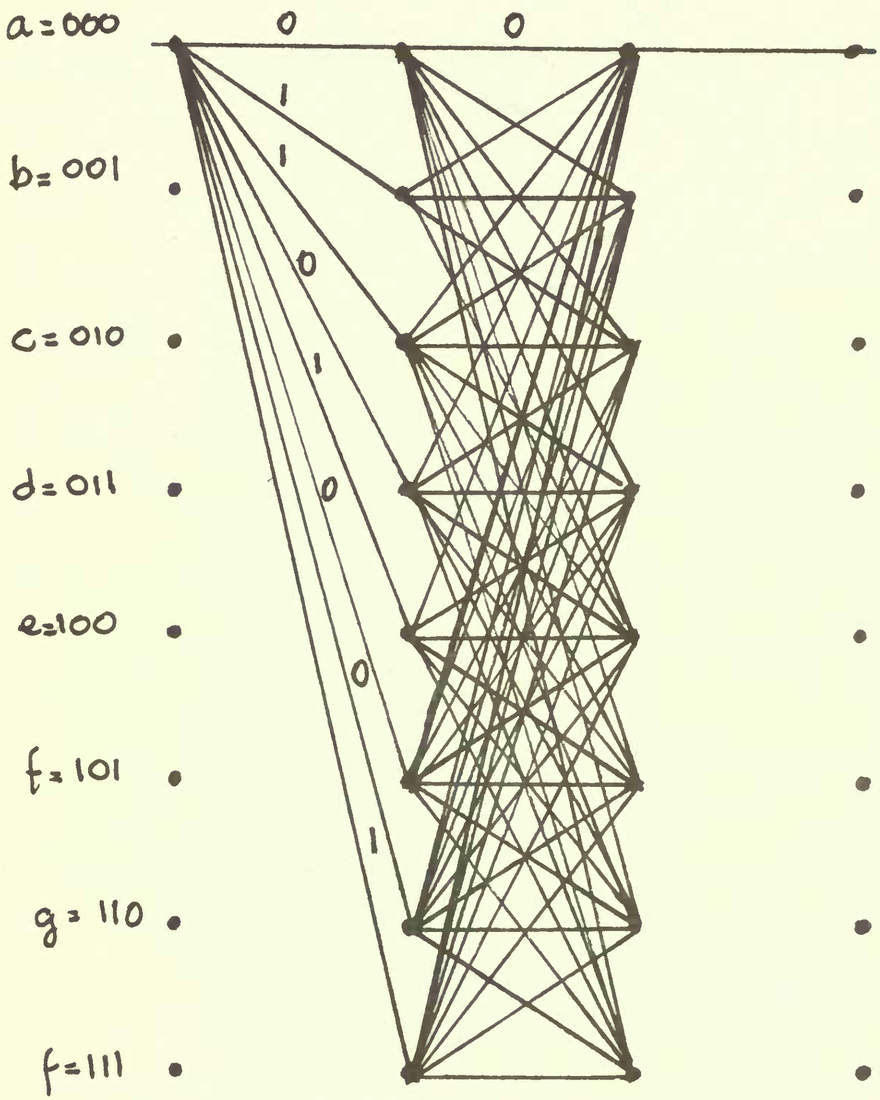
4.4 $K=2$, $R=3/4$. In fig 4.2c of the text we have the generator matrices:

$$g_0=\begin{bmatrix}1&0&0&1\\0&1&0&1\\0&0&1&0\end{bmatrix} \quad ; \quad g_1=\begin{bmatrix}0&0&0&1\\0&0&0&0\\0&0&0&1\end{bmatrix}$$

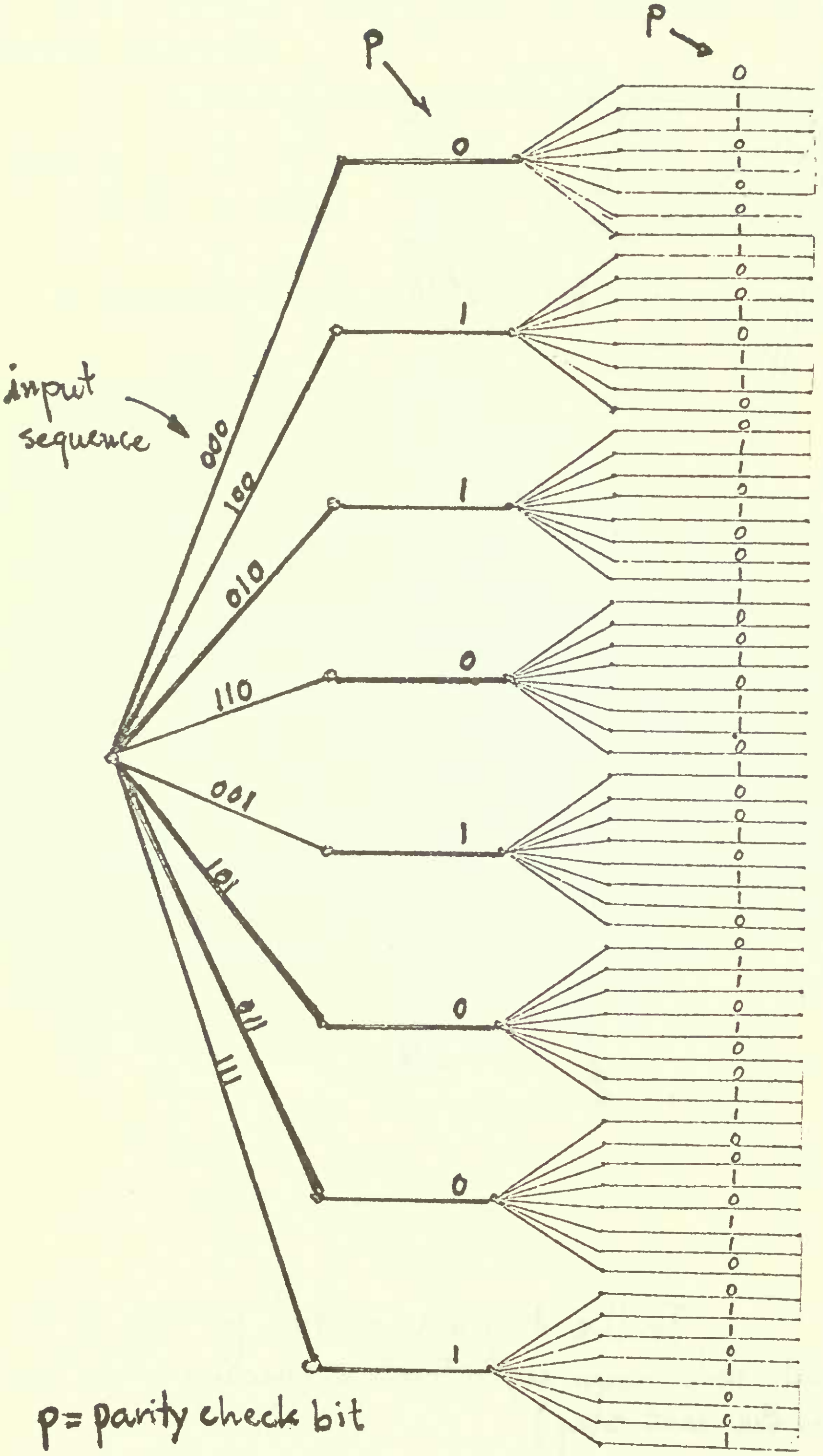
And so the output sequence has the form $u_1u_2u_3p$.



state diagram

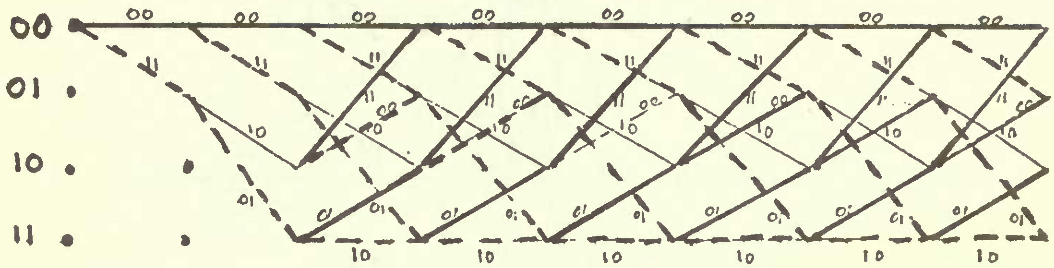


Trellis diagram (partial for less clutter - from each node leave eigth lines connecting it to the 8 nodes in the next step)

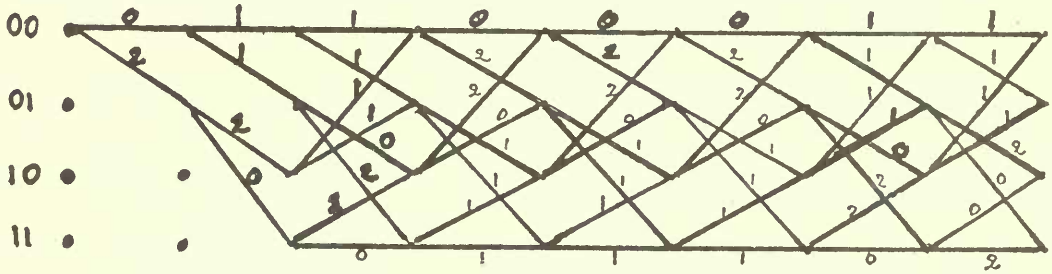


p= parity check bit

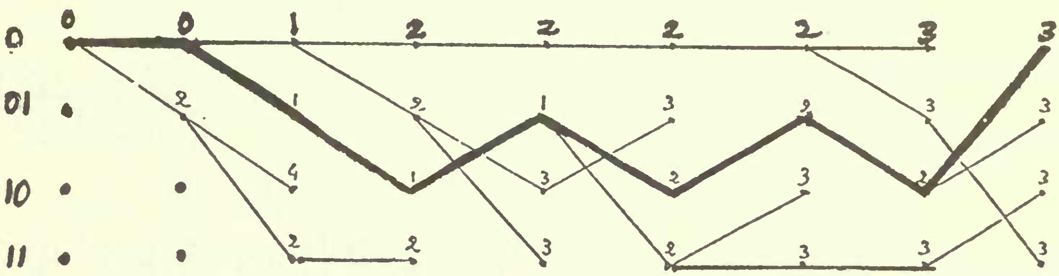
4.5



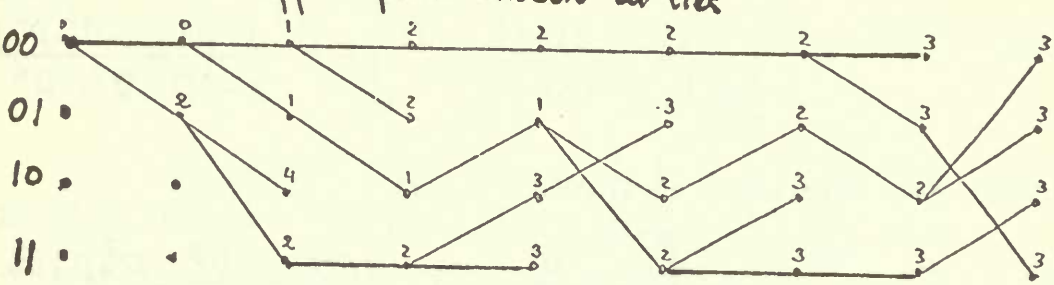
y = 00 01 10 00 00 00 10 01



weights with received sequence

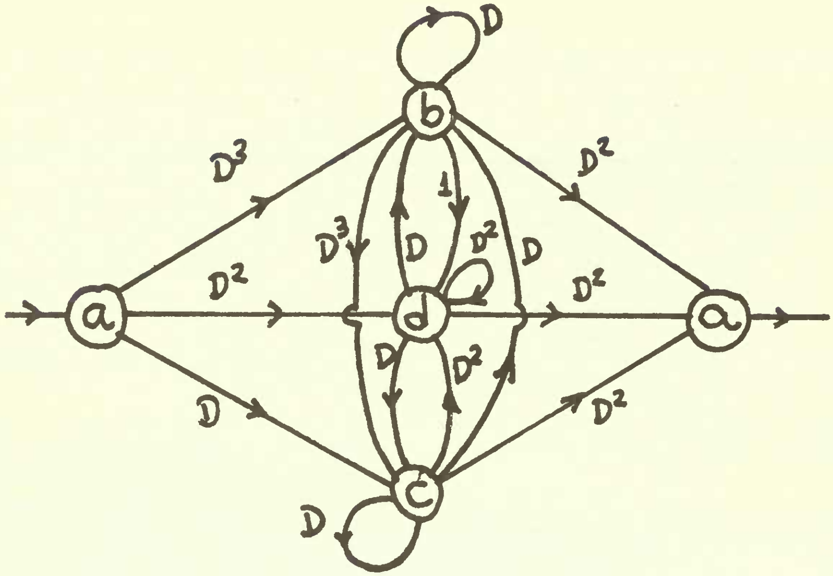


upper path chosen at ties



lower path chosen at ties

4.6



$$\begin{cases} \xi_b = D\xi_b + D\xi_c + D\xi_d + D^3 & (1)\\ \xi_c = D^3\xi_b + D\xi_c + D\xi_d + D & (2)\\ \xi_d = \xi_b + D^2\xi_c + D^2\xi_d + D^2 & (3)\end{cases}$$

$$T(D) = D^2\xi_b + D^2\xi_c + D^2\xi_d = D^2(\xi_b + \xi_c + \xi_d) \quad (4)$$

$$\left.\begin{aligned} &\text{From (1) \& (2)}: \quad \xi_c = D - D^3 + (D^3 - D + 1)\,\xi_b \\ &\text{From (2) \& (3)}: \quad \xi_c = \frac{(D^5 - D^3 - D)\xi_b - D}{D^2 + D - 1} \end{aligned}\right\} \Rightarrow \xi_b = \frac{D^5 + D^4 - 2D^3 - D^2}{D^4 - D^3 + 3D - 1}$$

From (1): $D(\xi_b + \xi_c + \xi_d) = -D^3 + \xi_b \Rightarrow D^2(\xi_b + \xi_c + \xi_c) = -D^4 + D\xi_b$

then $T(D) = D^3\left[\dfrac{D^3 + D^2 - 2D - 1}{D^4 - D^3 + 3D - 1} - D\right] = D^3\,\dfrac{1 + D + 2D - D^3 - D^4 + D^5}{1 - 3D + D^3 - D^4}$

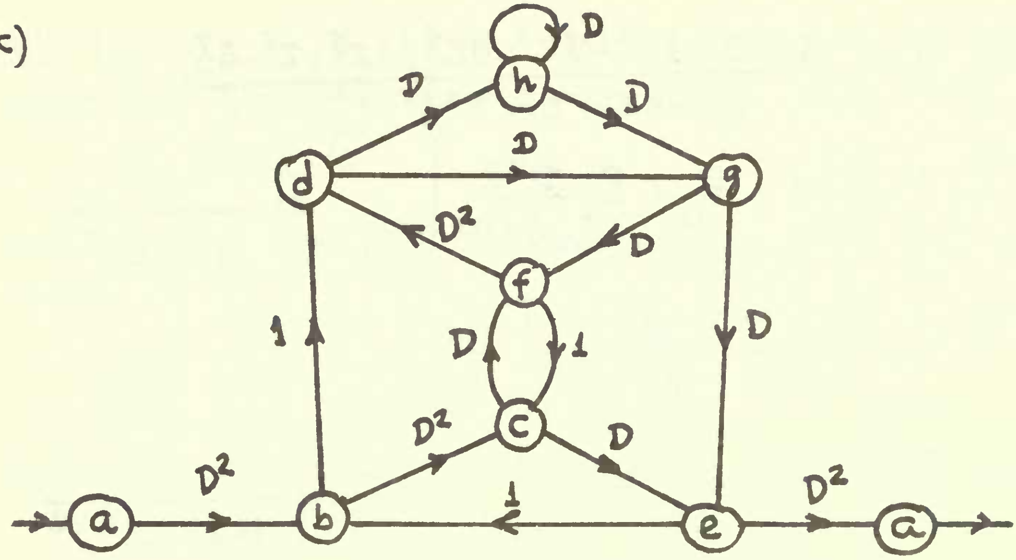
and the minimum distance is $d_{free} = 3$.

(b) 

$T(D) = \dfrac{D^3}{1-D} = D^3 + D^4 + D^5 + \ldots$

and the minimum distance among all paths is 3.

(c)



$$\begin{cases} \xi_b = D^2 + \xi_e & (1) \\ \xi_c = D^2\xi_b + \xi_f & (2) \\ \xi_d = \xi_b + D^2\xi_f & (3) \\ \xi_e = D\xi_c + D\xi_g & (4) \\ \xi_f = D\xi_c + D\xi_g & (5) \\ \xi_g = D\xi_d + D\xi_h & (6) \\ \xi_h = D\xi_h + D\xi_d & (7) \end{cases}$$

and $T(D) = D^2\xi_e$ (8)

Solving (1)-(7) we get

$$\xi_e = \frac{D^4 + D^5 - D^6}{1 - 2D - D^3}$$

and from (8) we get

$$T(D) = D^6\frac{(1 + D - D^2)}{1 - 2D - D^3}$$

and the minimum distance is $d_{free} = 6$.

4.7 For all codes we have

$$P_e < T(D)\Big|_{D=Z} \qquad \text{and} \qquad P_b < \frac{\partial T(D,I)}{\partial I}\Big|_{\substack{D=Z\\ I=1}}$$

(a) $P_e < T(D)\Big|_{D=Z} = Z^3\dfrac{1 + Z + 2Z^2 - Z^3 - Z^4 + Z^5}{1 - 3Z + Z^3 - Z^4}$

$$P_b < \frac{\partial T(D,I)}{\partial I}\Big|_{\substack{D=Z\\ I=1}} = \frac{\partial}{\partial I}\,\frac{D^3(I + DI^2 + D^2I^3 + D^2I - D^3I - D^4I^3 + D^5I^2)}{1 - 2DI - DI^3 + D^3I^2 - D^4I^2}$$

$$\therefore \quad P_b < Z^3 \frac{1+4Z+3Z^2-5Z^3-5Z^4+4Z^5+Z^6-Z^7}{(1-3Z+Z^3-Z^4)^2}$$

(b) $P_e < \dfrac{Z^3}{1-Z}$; $P_b < \dfrac{\partial}{\partial I} \dfrac{D^3 I}{1-DI}\Big|_{\substack{D=Z\\I=1}} = \dfrac{Z^3}{(1-Z)^2}$

(c) $P_e < \dfrac{Z^6(1+Z-Z^2)}{1-2Z-Z^3}$

$$P_b < \frac{\partial}{\partial I}\left(D^6 I \frac{I+D-D^2}{1-2DI-D^3}\right)\Big|_{\substack{D=Z\\I=1}} = Z^6 \frac{2-Z-2Z^2+Z^3+Z^5}{(1-2Z-Z^3)^2}$$

<u>4.8</u> $Q(x) \equiv \operatorname{erfc}(x) = \displaystyle\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\alpha^2/2} d\alpha$

We want to show that $Q(\sqrt{x+y}) \leq Q(\sqrt{x})e^{-y/2}$; $x \geq 0, y \geq 0$.

$$Q(\sqrt{x+y}) = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{x+y}}^\infty e^{-\alpha^2/2} d\alpha = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(\beta+\sqrt{x+y})^2} d\beta$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(\beta^2+x+y+2\beta\sqrt{x+y})} d\beta$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(\beta^2+x+2\beta\sqrt{x}) - \frac{y}{2} - \beta(\sqrt{x+y}-\sqrt{x})} d\beta$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(\beta+\sqrt{x})^2} e^{-y/2} d\beta = e^{-y/2} \int_{\sqrt{x}}^\infty e^{-\gamma^2/2} d\gamma = e^{-y/2} Q(\sqrt{x}).$$

Since $x, y \geq 0 \Rightarrow \sqrt{x+y} - \sqrt{x} \geq 0 \Rightarrow e^{-\beta(\sqrt{x+y}-\sqrt{x})} \leq 1$, $\beta \geq 0$.

4.9 Let $\underset{\sim}{u}_i$, $\underset{\sim}{x}_i$, G be the data (row) vector, code (row) vector and generator matrix respectively, such that

$$\underset{\sim}{x}_i = \underset{\sim}{u}_i G$$

and arrange the code in matrix form with $\underset{\sim}{x}_i$'s as rows

$$[\underset{\sim}{X}] = \begin{bmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \\ \vdots \\ \underset{\sim}{x}_{2^k} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \vdots \\ \underset{\sim}{u}_{2^k} \end{bmatrix} \begin{bmatrix} G \end{bmatrix}$$

(a) Consider any column of $[G]$. Either it consists of "zeros", in which case every $\underset{\sim}{x}_i$, $i = 1, \ldots, 2^k$, will have a "zero" in that column position, or it has at least one "1". Suppose there is only one "1" in the j^{th} column of $[G]$ and that "1" is in the ℓ^{th} row. Then the j^{th} column of $[\underset{\sim}{X}]$ will be exactly the ℓ^{th} column of $[\underset{\sim}{u}]$. If there are more "1"s in the j^{th} column of $[G]$ then the j^{th} column of $[\underset{\sim}{X}]$ will be the linear combination of the corresponding columns of $[\underset{\sim}{u}]$.

Now consider the matrix $[\underset{\sim}{u}]$. It has all possible combinations of binary sequences as rows. Then, $[\underset{\sim}{u}]$ looks like

$$\begin{bmatrix} 0\,0\,0\,0 \cdots\cdots\cdots 0 \\ 1\,0\,0 \cdots\cdots\cdots 0 \\ 0\,1\,0 \cdots\cdots\cdots 0 \\ 1\,1\,0\,0 \cdots\cdots\cdots 0 \\ 0\,0\,1\,0 \cdots\cdots\cdots 0 \\ 1\,0\,1\,0 \cdots\cdots\cdots 0 \\ \vdots \qquad\qquad\qquad \vdots \\ 0\,0 \cdots\cdots\cdots\cdots 1 \\ \vdots \qquad\qquad\qquad \vdots \\ 1\,1 \cdots\cdots\cdots\cdots 1 \end{bmatrix}$$

As it can be seen that in column j (any j) we have from the top 2^{j-1} "0"s followed by 2^{j-1} "1"s, followed again by 2^{j-1} "0"s and so on. Therefore in each column we have equal numbers of "zeros" and "ones" and any linear combination of these columns will still have an equal number of "zeros" and "ones" or it will be all "zeros".

(b) Consider $[\underset{\sim}{X}]$ as in (a), the matrix of all possible codes of length N and exclude the all-zero code vector. We have then 2^k-1 code vectors left. We know from (a) also that each column in $[\underset{\sim}{X}]$ is either all "zeros" or half "zeros", half "ones" (in any order). Assume now $[\underset{\sim}{X}]$ does not have any all "zeros" column. Then the total number of "ones" in $[\underset{\sim}{X}]$ will be $N2^{k-1}$ and so, the average number of "ones" in each code vector will be

$$w_{av}(k) = \frac{N\,2^{k-1}}{2^k-1} \quad \text{where } k = B-K+1 \text{ or } B = K+k-1.$$

and where B is the number of branches for codes with length N and in each branch we have n symbols so $N = Bn = (K+k-1)n$. If there exists any column with all zeros then we have an upper bound on $w_{av}(k)$

$$w_{av}(k) \leq \frac{2^{k-1}(K+k-1)n}{2^k-1}$$

(c) The distance between any codevector and all-zero vector is equal to its weight (number of "ones"). From (b) there exists a code whose weight is less or equal to $w_{av}(k)$. So the minimum weight code should have a free distance that satisfies $d_{free} \leq w_{av}(k)$, for all k.

(d) k varies from 1 up so $d_{free} \leq \min\limits_{k} \dfrac{2^{k-1}(K+k-1)n}{2^k-1}$.

4.10 Let $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ be any two codewords with $d = w(\underset{\sim}{x}_1 \oplus \underset{\sim}{x}_2)$. Define decision regions

$$\Lambda_1 = \{\underset{\sim}{y} : w(\underset{\sim}{y} \oplus \underset{\sim}{x}_1) \le w(\underset{\sim}{y} \oplus \underset{\sim}{x}_2)\}$$

$$\Lambda_2 = \{\underset{\sim}{y} : w(\underset{\sim}{y} \oplus \underset{\sim}{x}_2) \le w(\underset{\sim}{y} \oplus \underset{\sim}{x}_1)\}$$

and if $\underset{\sim}{y} \in \Lambda_1 \cap \Lambda_2 = \{\underset{\sim}{y} : w(\underset{\sim}{y} \oplus \underset{\sim}{x}_1) = w(\underset{\sim}{y} \oplus \underset{\sim}{x}_2)\}$, either decision can be made with no loss of optimality. If we count $\underset{\sim}{y} \in \Lambda_1 \cap \Lambda_2$ as an error (which will be true half of the time only) we upper bound the probability of error:

$$P_e \le \frac{1}{2} \sum_{\Lambda_1} P_N(\underset{\sim}{y} | \underset{\sim}{x}_2) + \frac{1}{2} \sum_{\Lambda_2} P_N(\underset{\sim}{y} | \underset{\sim}{x}_1) = \Pr\{w(\underset{\sim}{e}) \ge d/2\}$$

where $\underset{\sim}{e}$ is the binary error vector. For $d = 2t$ we have $\Pr\{w(\underset{\sim}{e}) \ge d/2\} = \Pr\{w(\underset{\sim}{e}) \ge t\}$, while for $d = 2t-1$ we have $\Pr\{w(\underset{\sim}{e}) \ge d/2\} = \Pr\{w(\underset{\sim}{e}) \ge t - 1/2\} = \Pr\{w(\underset{\sim}{e}) \ge t\}$ because t can only assume integer values. Thus we have that $\Pr\{w(\underset{\sim}{e}) \ge d/2\}$ is the same for $d = w(\underset{\sim}{x}_1 \oplus \underset{\sim}{x}_2)$ odd or even. The Bhattacharyya bound begins with this bound so we have, <u>when d is odd</u>

$$P_d \le Z^{d+1}$$

from (4.4.5) we have

$$P_e(j) \le \sum_{d=d_f}^{\infty} a(d) P_d \le \sum_{d \text{ even}} a(d) Z^d + \sum_{d \text{ odd}} a(d) Z^{d+1}$$

Note that

$$\sum_{d \text{ even}} a(d) Z^d = \frac{1}{2} \left\{ \sum_{d=d_f}^{\infty} a(d) Z^d + \sum_{d=d_f}^{\infty} a(d) (-Z)^d \right\}$$

$$= \frac{1}{2} T(Z) + \frac{1}{2} T(-Z)$$

and

$$\sum_{d \text{ odd}} a(d) Z^{d+1} = \frac{1}{2} Z \left\{ \sum_{d=d_f}^{\infty} a(d) Z^d - \sum_{d=d_f}^{\infty} a(d) (-Z)^d \right\}$$

$$\therefore \sum_{d \text{ odd}} a(d) z^{d+1} = \frac{1}{2}\left(Z\,T(Z) - Z\,T(-Z)\right)$$

and

$$P_e(j) \leq \frac{1}{2}\left\{(1+Z)\,T(Z) + (1-Z)\,T(-Z)\right\}$$

4.11

Let $u_0, u_1, u_2, \ldots$ be a binary sequence and denote the input data polynomial as $I(Z) = u_0 + u_1 Z + u_2 Z^2 + \cdots$. By definition, a convolutional code is said to be catastrophic if there exists an input sequence with infinitely many nonzero digits such that the corresponding output has only finitely many nonzero digits.

Let $T_i(Z) = g_i(Z)\,I(Z)$ be the output polynomial corresponding to input polynomial $I(Z)$. Suppose

$\gcd(g_1(Z), g_2(Z), \ldots, g_n(Z)) = h(Z)$ where $h(Z)$ has degree at least one, $h(0) \neq 0$. Then, the input sequence whose polynomial is $I(Z) = 1/h(Z)$ contains infinitely many nonzero digits and the output sequence polynomials

$T_i(Z) = g_i(Z)/h(Z)$, $i = 1, \ldots, n$ will be of finite degree and hence the corresponding output sequence will have only finitely many nonzero digits and by definition the code is catastrophic. Conversely, suppose we have a catastrophic convolutional code and assume the condition on the greatest common divider is violated ie

$$\gcd(g_1(Z), g_2(Z), \ldots, g_n(Z)) = Z^m, \text{ where } m \geq 0.$$

Then we know there exist polynomials of finite degree $p_1(Z), p_2(Z), \ldots, p_n(Z)$ such that

$$p_1(Z)\,g_1(Z) + p_2(Z)\,g_2(Z) + \cdots + p_n(Z)\,g_n(Z) = Z^m$$

Multiplying by $I(Z)$ and using $T_i(Z) = g_i(Z)\,I(Z)$, all i

we will have

$$p_1(z)T_1(z)+p_2(z)T_2(z)+\ldots+p_n(z)T_n(z)=Z^m I(z)$$

Since the code is by assumption catastrophic, for some $I(z)$ with infinitely many nonzero digits we should have $T_i(z)$, $i=1,2,\ldots,n$ with finitely many nonzero digits. Then the left side of the above equation will be a polynomial with finite degree and the right side will have infinetely many nonzero digits, which is a contradiction. Therefore a catastrophic code should have its n generator polynomials having a common factor $h(z)$ of degree at least one and $h(0)\neq 0$.

4.12 Consider a binary convolutional code with rate $1/n$ and constraint K, with code generator polynomials

$$g_1(z)=g_{0,1}+g_{1,1}Z+\ldots+g_{K-1,1}Z^{K-1}$$
$$g_2(z)=g_{0,2}+g_{1,2}Z+\cdots+g_{K-1,2}Z^{K-1}$$
$$\vdots$$
$$g_n(z)=g_{0,n}+g_{1,n}Z+\ldots+g_{K-1,n}Z^{K-1}$$

where at least one of the $g_{K-1,i}$ should be non-zero. Also, without loss of generality, assume polynomials with a nonzero constant term. Consider the binary matrix

$$\begin{bmatrix} g_{0,1} & g_{1,1} & \cdots & g_{K-1,1} \\ g_{0,2} & g_{1,2} & \cdots & g_{K-1,2} \\ \vdots & & & \\ g_{0,n} & g_{1,n} & \cdots & g_{K-1,n} \end{bmatrix}$$

where the elements of the first column are all nonzero and at least one element of the last column is also nonzero. Then there are

$$N(K)=2^{n(K-2)}(2^n-1)$$

such matrices or equivalently $N(K)$ convolutional codes with constraint K, since we have 2^n-1 possibilities for the last column and 2^n possibilities for columns j, $j=2,\dots,K-1$ and only one possibility for column 1. We know the delay of a polynomial $\phi(z)$ is the greatest l such that z^l divides $\phi(z)$ and the delay of a set of polynomials is the minimum of the delays of the polynomials in the set. We consider only sets of polynomials with the same delay for all members. If we now multiply each of the n polynomials of a code by any $\phi(z)$ of degree m then, if the delay of $\phi(z)$ is zero the result will be another code of degree $K+m-1$, while if the delay of $\phi(z)$ is greater than zero the result will not be a code in the above class. The number of sets of degree $K-1$ and delay l is equal to the number of codes of degree $K-l-1$ and delay zero, which is exactly $N(K-l)$. Therefore the number of sets of degree $K-1$ and arbitrary delay is $\Lambda(K)=\sum_{l=0}^{K-1} N(K-l)$. Using our previous result we have:

$$\Lambda(K)=\sum_{l=0}^{K-1} 2^{n(K-l-2)}(2^n-1)=2^{n(K-1)}$$

Let $\Gamma(K)$ be the number of codes of degree $K-1$ such that the greatest common divisor of their generator polynomials is one, which is the same as the number of codes that are not catastrophic. Then each set of polynomials of degree $K-1$ and arbitrary delay may be written in a unique way as the product of a code of degree $K-m-1$ and greatest common divisor of 1 and a monic polynomial of degree m, for some m. Therefore

$$\Lambda(K)=\sum_{m=0}^{K-1}\Gamma(K-m)\,\Delta(m) \qquad (*)$$

But $\Delta(m)$, number of monic polynomials of degree m is 2^m so from the above equation, by using z-transform method we can easily find that $\Gamma(K)=2^{n(K-1)}(1-2^{-(n-1)})$.

Equation (*) looks like discrete convolution, therefore the transform of $\Lambda(K)$ is the product of the transforms of $\Gamma(K)$ and Δ. The number of catastrophic codes is then $N(K) - \Gamma(K)$ and the relative frequency of catastrophic codes in the ensemble of all convolutional codes of constraint K is given by the fraction

$$F(K) = \frac{N(K) - \Gamma(K)}{N(K)} = \frac{2^{n(K-2)}(2^n - 1) - 2^{n(K-1)}(1 - 2^{-(n-1)})}{2^{n(K-2)}(2^{n-1})} = \frac{1}{2^n - 1}$$

which is independent of K.

4.13

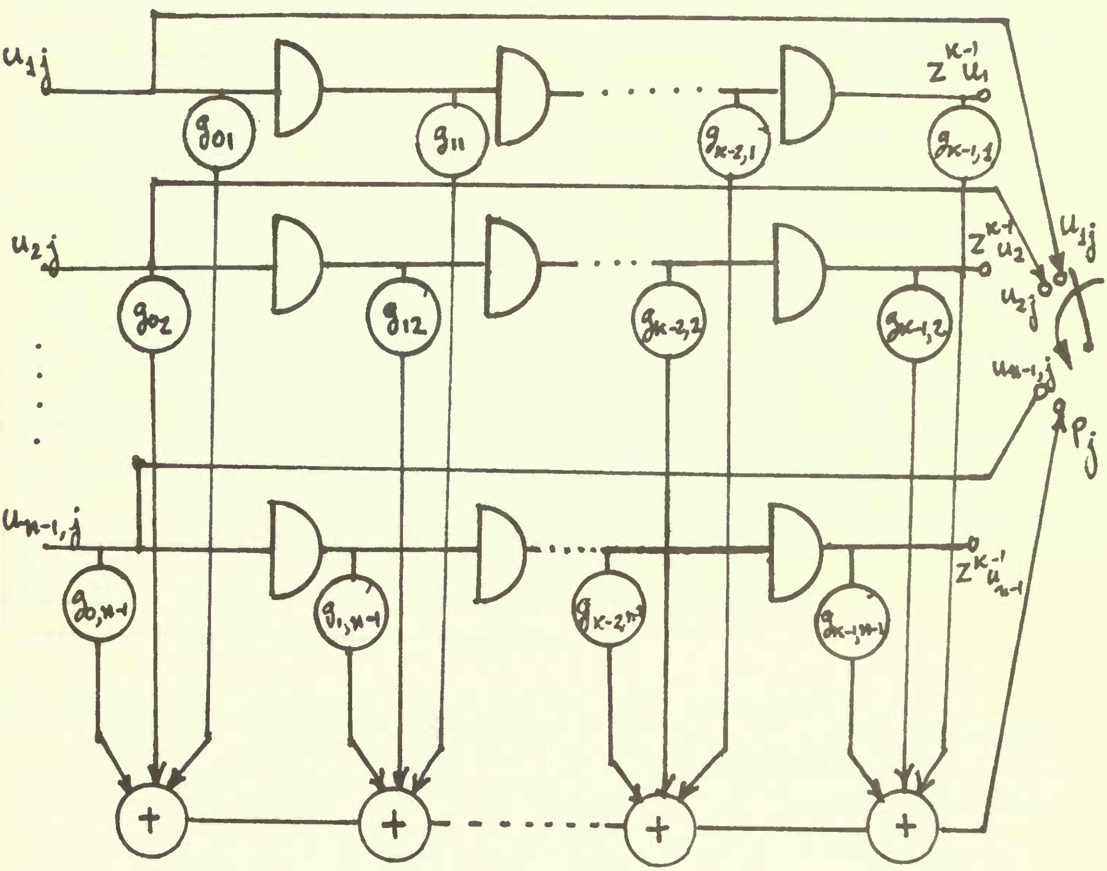
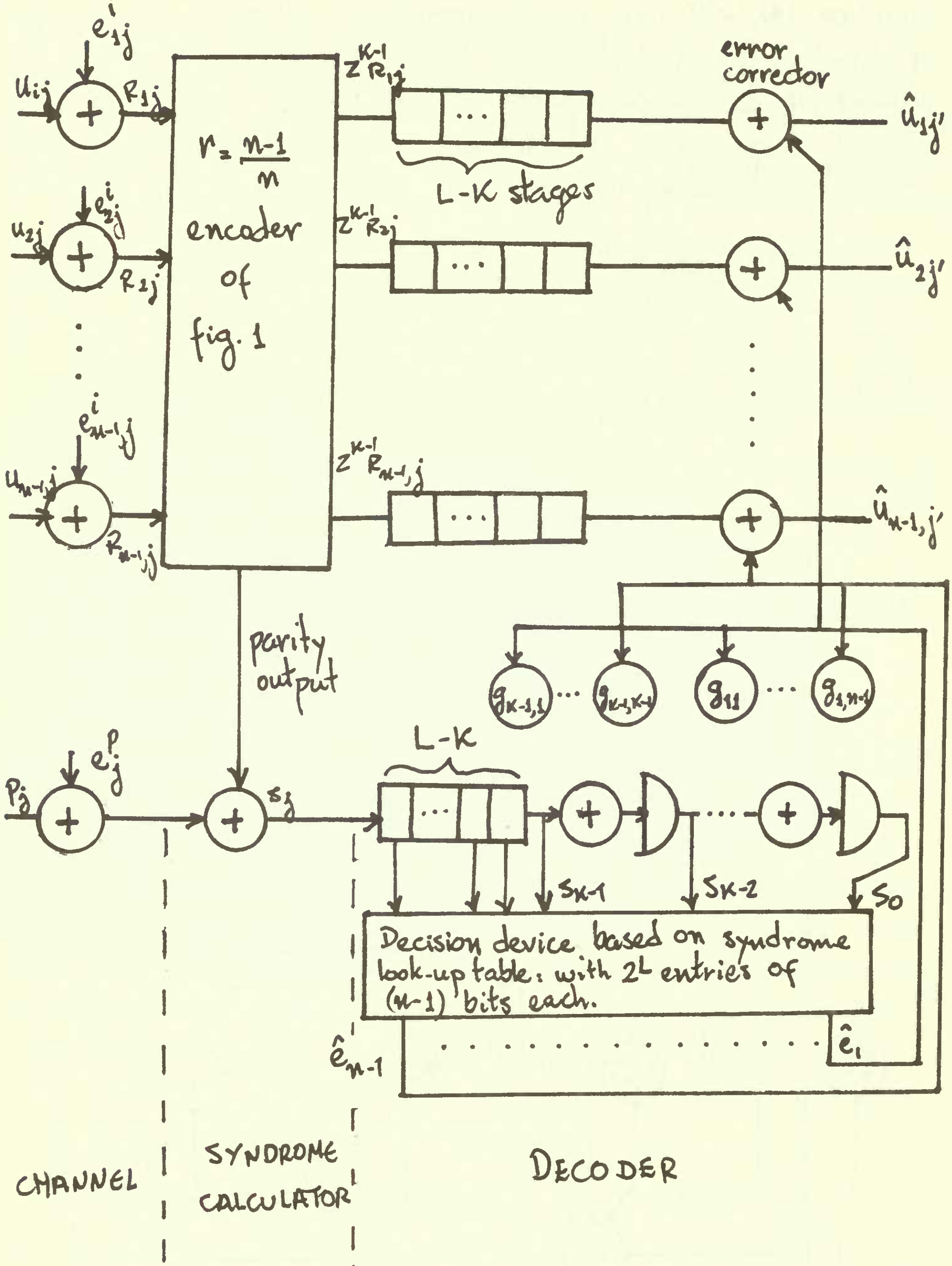


Figure 1: Encoder



error corrector
r = (n-1)/n encoder of fig. 1
L-K stages
parity output
L-K
Decision device based on syndrome look-up table: with 2^L entries of (n-1) bits each.
CHANNEL
SYNDROME CALCULATOR
DECODER

Figure 2

4.14 (a) We note that when a path diverges from the all-zero path at some node j and remerges for the first time at some node $i \geq j+K$, the corresponding data sequence should have a "1" at the first position and $K-1$ "zeros" at the tail. Then any path that diverges from the all-zero path at node j and remains unmerged for $K+\ell$ branches, must have symbols of the form $1\, u_{j+1}\, u_{j+2} \dots u_{j+\ell}\, \underbrace{00\dots0}_{K-1}$

where $u_{j+\ell}=1$ and $u_{j+1}\dots u_{j+\ell-1}$ is any binary sequence containing no strings of more than $K-2$ zeros together or the path will remerge at node $i < j+K+\ell$. When $\ell=0$ we have only one path since there is only one data sequence with length K which is of the form $100\dots0$ ("1" followed by $K-1$ "0"s). When $1 \leq \ell \leq K$ the length of the binary sequence $u_{j+1}\dots u_{j+\ell-1}$ is $\leq K-2$ and so we cannot have sequences of $K-1$ consecutive zeros. There are $2^{\ell-1}$ such sequences and so $a(\ell) = 2^{\ell-1}$ in this case. For the case $\ell \geq K$ consider the binary sequences of the form

$$1\, u_{j+1} \dots u_{j+(\ell-i)-1}\, \underbrace{100\dots0}_{i-1}\, \underbrace{1000\dots0}_{K-1}$$

By our definition the number of such binary sequences is $a(\ell-i)$. If $i = 1,\dots,K-1$ we have all combinations of sequences that diverge from all-zero path at node j and remerge for the first time at node $j+K+\ell$, $\ell \geq K$. Then

$$a(\ell) = \sum_{i=1}^{K-1} a(\ell-i)$$

(b) Using the above results we have

$$T_K(L) = \sum_{\ell=0}^{\infty} a(\ell) L^{K+\ell} = L^K + \sum_{\ell=1}^{K-1} 2^{\ell-1} L^{K+\ell} + \sum_{\ell=K}^{\infty} \sum_{i=1}^{K-1} a(\ell-i) L^{K+\ell}$$

$$= L^K + \frac{L^K}{2} \sum_{\ell=1}^{K-1} (2L)^\ell + \sum_{i=1}^{K-1} \sum_{\ell=K}^{\infty} a(\ell-i) L^{K+\ell}$$

or $$T_K(L) = L^K + \frac{L^K}{2}\,\frac{2L-(2L)^K}{1-2L} + \sum_{i=1}^{K-1}\sum_{l=K-i}^{\infty} a(l) L^{K+l+i}$$

$$= L^K + \frac{L^{K+1} - 2^{K-1}L^{2K}}{1-2L} + \sum_{i=1}^{K-1} L^i\left[T_K(L) - \sum_{l=0}^{K-i-1} a(l) L^{K+l}\right]$$

$$= L^K + \frac{L^{K+1} - 2^{K-1}L^{2K}}{1-2L} + T_K(L)\sum_{i=1}^{K-1} L^i - \sum_{i=1}^{K-1}\sum_{l=0}^{K-i-1} a(l) L^{K+l+i}$$

Interchanging summations we have

$$T_K(L) = \frac{L^K - L^{K+1} - 2^{K-1}L^{2K}}{1-2L} + T_K(L)\,\frac{L-L^K}{1-L} - \sum_{l=0}^{K-2}\sum_{i=1}^{K-l-1} a(l) L^{K+l+i}$$

$\therefore$

$$T_K(L)\left[1 - \frac{L-L^K}{1-L}\right] = \frac{L^K(1-L-2^{K-1}L^K)}{1-2L} - \sum_{i=1}^{K-1} L^{K+i} - \sum_{l=1}^{K-2}\sum_{i=1}^{K-l-1} 2^{l-1} L^{K+l+i}$$

$$T_K(L)\left[\frac{1-2L+L^K}{1-L}\right] = \frac{L^K(1-L-2^{K-1}L^K)}{1-2L} - \frac{L^{K+1}-L^{2K}}{1-L} - \sum_{l=1}^{K-2} 2^{l-1} L^{K+l}\sum_{i=1}^{K-l-1} L^i$$

$$= L^K\left[\frac{1-L-2^{K-1}L^K}{1-2L} - \frac{L-L^K}{1-L}\right] - \sum_{l=1}^{K-2} 2^{l-1} L^{K+l}\,\frac{L-L^{K-l}}{1-L}$$

$$= L^K\left[\frac{1-L-2^{K-1}L^K}{1-2L} - \frac{L-L^K}{1-L}\right] - \frac{L^{K+1}}{2(1-L)}\left\{\sum_{l=1}^{K-2}(2L)^l - L^{K-1}\sum_{l=1}^{K-2} 2^l\right\}$$

$$= L^K\left\{\frac{1-L-2^{K-1}L^K}{1-2L} - \frac{L-L^K}{1-L} - \frac{L}{2(1-L)}\,\frac{2L-(2L)^{K-1}}{1-2L} + L^{K-1}(2-2^{K-1})\right\}$$

$$= L^K\left\{\frac{1-3L+2L^2}{(1-2L)(1-L)}\right\} = L^K\left\{\frac{1-3L+2L^2}{1-3L+2L^2}\right\} = L^K$$

and finally

$$T_K(L) = \frac{L^K(1-L)}{1-2L+L^K}$$

(c) Clearly the set of binary sequences of length $l-1$ contains the set of binary sequences of length $l-1$ which do not have $K-1$ consecutive "0"s. So $2^{l-1} \geq a(l)$ for $l>0$. For $1 \leq l \leq K$ we saw that $a(l) = 2^{l-1}$. Let $l \geq K$ and consider sequences of the form

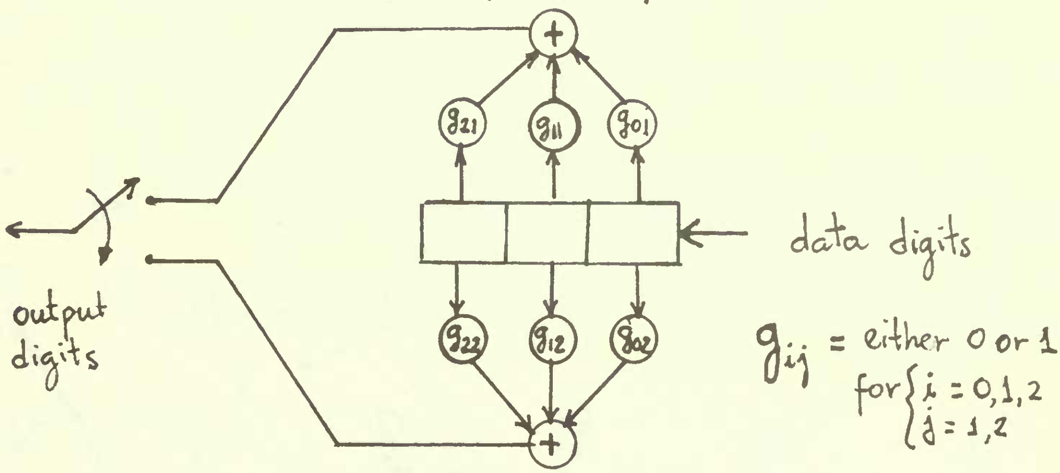
$$1\, u_{j+1} \ldots u_{j+l-1}\, 1\, \underbrace{0\,0 \ldots 0}_{K-1}$$

with at least $K-1$ consecutive "0"s in the binary vector $u_{j+1} \ldots u_{j+l-1}$. Suppose the first $K-1$ digits $u_{j+1} \ldots u_{K-1}$ are zero. There are then 2^{l-K} such vectors. So the number of vectors $u_{j+1} \ldots u_{j+l-1}$ with at least $K-1$ consecutive "0"s is less than $l\,2^{l-K}$. Then

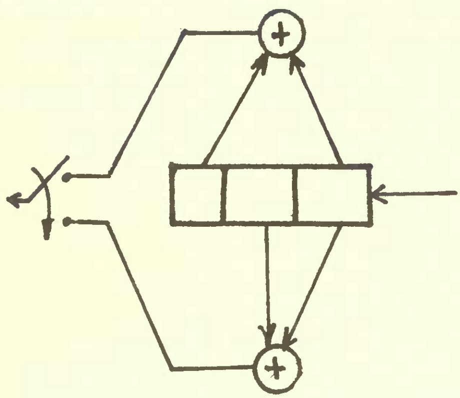
$$a(l) \geq 2^{l-1} - l\,2^{l-K} = 2^{l-1}(1 - l\,2^{-(K-1)})$$

and therefore $2^{l-1} \geq a(l) \geq 2^{l-1}(1 - l\,2^{-(K-1)})$.

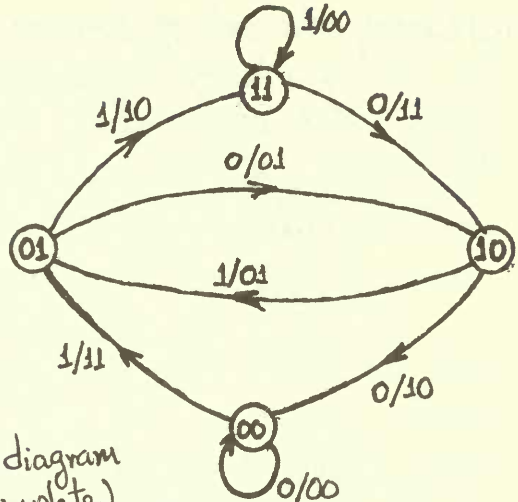
4.15 Since $K=3$, $R=1/2$ the form of the encoder is



Suppose initially all registers are zero. If we insert a "1" then we go from state 00 to state 01 and by the given partial state diagram our output is 11. This implies $g_{01} = g_{02} = 1$. If we go now to state 10 from state 01 our output must be 01. This implies $g_{11}=0$ and $g_{12} = 1$. If we go from 11 to itself output should be 00 so this implies $g_{21} = 1$ and $g_{22}=0$. Our encoder and state diagram will be

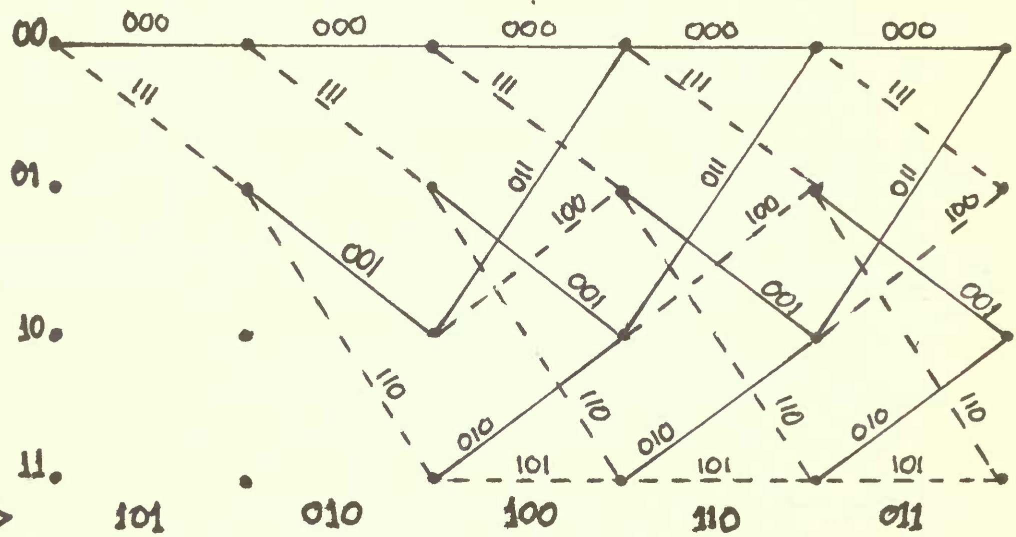


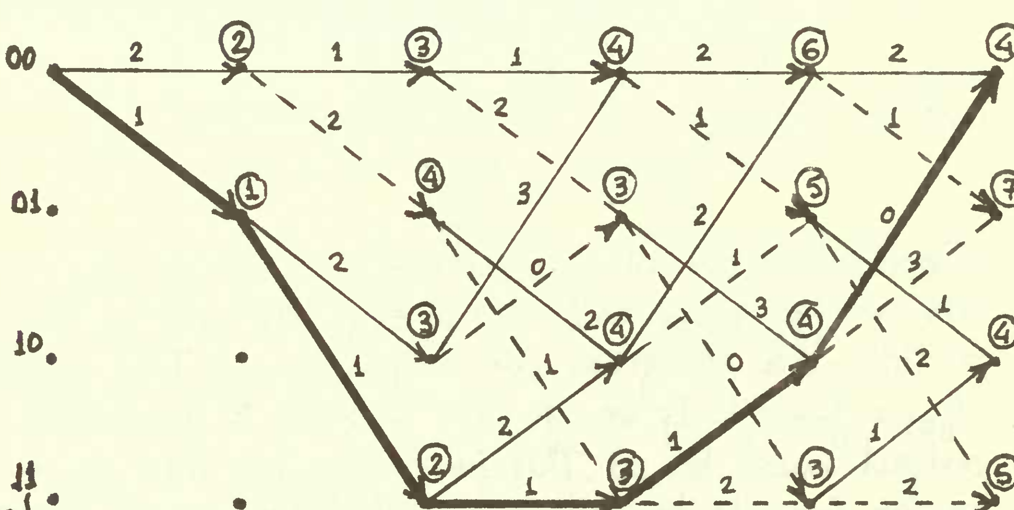
encoder



state diagram (complete)

4.16





Decoded data $\longleftrightarrow$ 1 1 1 0 0

$\underline{4.17}$

$$\xi_b = LI\,\xi_c + D^2LI$$
$$\xi_c = DL\,\xi_b + DL\,\xi_d$$
$$\xi_d = DLI\,\xi_b + DLI\,\xi_d$$

$\therefore$

$$\underset{\sim}{\xi} = \begin{bmatrix} 0 & LI & 0 \\ DL & 0 & DL \\ DLI & 0 & DLI \end{bmatrix} \underset{\sim}{\xi} + \begin{bmatrix} D^2LI \\ 0 \\ 0 \end{bmatrix} = A\,\underset{\sim}{\xi} + \underset{\sim}{b} \qquad (1)$$

Therefore

$$A = \begin{bmatrix} 0 & LI & 0 \\ DL & 0 & DL \\ DLI & 0 & DLI \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} D^2LI \\ 0 \\ 0 \end{bmatrix}$$

From (1) $\underset{\sim}{\xi} = (I-A)^{-1}\underset{\sim}{b}$ thus

$$\underset{\sim}{\xi} = \begin{bmatrix} 1 & -LI & 0 \\ -DL & 1 & -DL \\ -DLI & 0 & 1-DLI \end{bmatrix}^{-1} \begin{bmatrix} D^2LI \\ 0 \\ 0 \end{bmatrix}$$

$$\det(I-A) = 1 - DLI + LI\,(D^2L^2I - DL - D^2L^2I) = 1 - DLI - DL^2I$$
$$= 1 - DLI(1+L)$$

$\therefore$

$$\underset{\sim}{\xi} = \frac{1}{1-DLI(1+L)} \begin{bmatrix} 1-DLI & LI-DL^2I^2 & DL^2I \\ DL & 1-DLI & DL \\ DLI & DL^2I^2 & 1-DL^2I \end{bmatrix} \begin{bmatrix} D^2LI \\ 0 \\ 0 \end{bmatrix}$$

and finally

$$\xi = \frac{D^2LI}{1-DLI(1+L)} \begin{bmatrix} 1-DLI \\ DL \\ DLI \end{bmatrix}$$

$$T(D,L,I) = D^2L\,\xi_c = D^2L\,\frac{D^3L^2I}{1-DLI(1+L)}$$

$$= \frac{D^5L^3I}{1-DLI(1+L)}$$, which is exactly (4.33).

<u>4.18</u> From problem 4.17

$$A = \begin{bmatrix} 0 & LI & 0 \\ DL & 0 & DL \\ DLI & 0 & DLI \end{bmatrix} = L \begin{bmatrix} 0 & I & 0 \\ D & 0 & D \\ DI & 0 & DI \end{bmatrix} \equiv LA'$$

Then

$A^K = L^K A'^K$. We now find the eigenvalues of A':

$$\det\left[A' - \lambda I_{3\times 3}\right] = 0 \quad \Rightarrow \quad \det \begin{bmatrix} -\lambda & I & 0 \\ D & -\lambda & D \\ DI & 0 & DI-\lambda \end{bmatrix} = 0$$

$\therefore \; \lambda^3 - DI\lambda^2 - DI\lambda = 0 \;\Rightarrow\; \lambda_1 = 0$; $\lambda_2 = \dfrac{DI + \sqrt{D^2I^2 + 4DI}}{2}$ and

$\lambda_3 = \dfrac{DI - \sqrt{D^2I^2 + 4DI}}{2}$;

We can write A' as

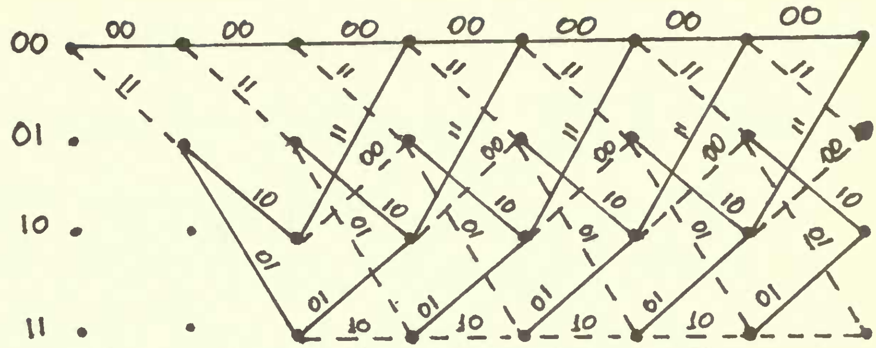
$$A' = M\Lambda M^{-1} = M \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} M^{-1}$$

where M is the matrix of eigenvectors of A'. A^K is then

$$A^K = L^K A'^K = M \begin{bmatrix} \lambda_1^K & 0 & 0 \\ 0 & \lambda_2^K & 0 \\ 0 & 0 & \lambda_3^K \end{bmatrix} M^{-1} L^K$$

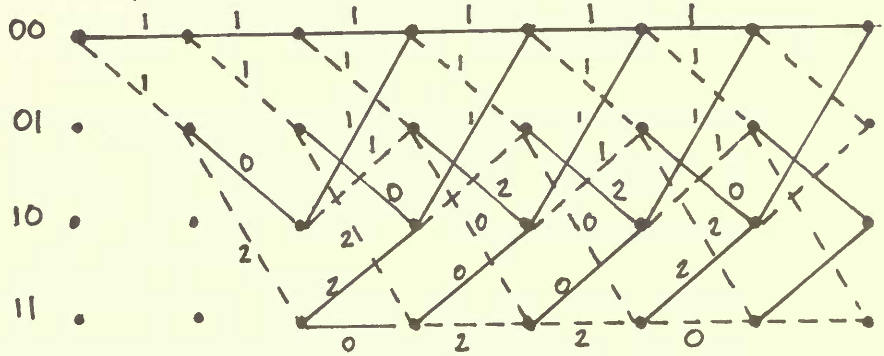
In matrix Λ, λ_2^K or λ_3^K should have term with lowest degree proportional to $D^{K/2}$ and so the elements of A^K decrease at least as fast as $D^{K/2}$.

4.19

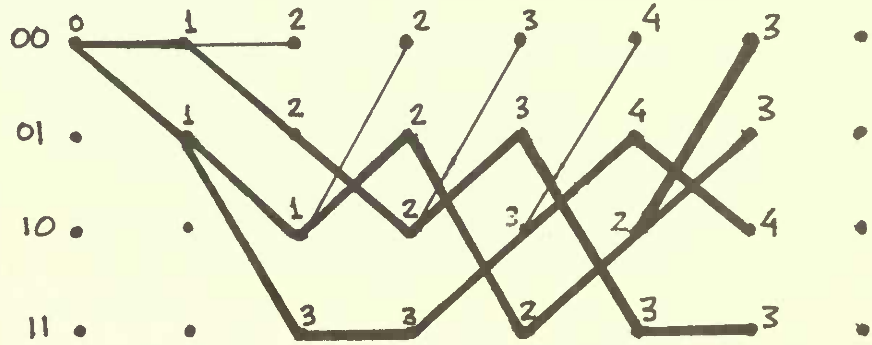


received sequence : 10 10 10 01 01 10

(in darken lines: path for the correct sequence transmitted)



Weights of branches with received sequence



Trellis with survivors' Hamming distances (lower path chosen in case of tie).

4.20

(a) Suppose all-zero inputs. Suppose $\max j = i_{J/2}$

From the fact that $P_0(j) \geq P_0(-j)$ we know when input is "0" there is more possibility that the output is encoded as that of positive metric.

$$[\Pi(z)]^d = \Big[\sum_{j=-i_J/2}^{i_J/2} P_0(j)\, z^j\Big]^d$$

$$= \sum_{k=-d\cdot i_J/2}^{d\cdot i_J/2} a_k \cdot P_0(j_1)^{\alpha_1} \cdot P_0(j_2)^{\alpha_2} \cdots P_0(j_l)^{\alpha_l} \cdot Z^k$$

where $\alpha_1 + \alpha_2 + \cdots + \alpha_l = d$, α : nonnegative integer

$j_1\alpha_1 + j_2\alpha_2 + \cdots + j_l\alpha_l = k$

a_k = the number of such event

P_d = Pairwise error probability for an incorrect path at Hamming distance d from the correct path upon remerging.

= Total sum of all possible products (d times) of $P_0(j)$ whose sum-metrics are negative plus half of all possible products (d times) of $P_0(j')$ whose sum-metrics are zero.

$$= \sum_{k=-d\cdot i_J/2}^{-1} a_k \cdot P_0(j_1)^{\alpha_1} P_0(j_2)^{\alpha_2} \cdots P_0(j_l)^{\alpha_l}$$

$$+ \frac{1}{2} a_0 P_0(j_1')^{\alpha_1'} P_0(j_2')^{\alpha_2'} \cdots P_0(j_n)^{\alpha_n'}$$

$$= \{[\Pi(z)]^d\}_-$$

//

(b)

$$T(D,I) = \sum_{d=d_f}^{\infty} \sum_{i=1}^{\infty} a(d,i)\, D^d I^i \qquad (4.4.9)$$

$$\frac{\partial}{\partial I} T(D,I)\Big|_{I=1} = \sum_{d=d_f}^{\infty}\Big(\sum_{i=1}^{\infty} i\, a(d,i)\Big) D^d \qquad (4.4.10)$$

$$= \sum_{d=d_f}^{\infty} b(d)\, D^d$$

$$P_b = E\{n_b\} \le \sum_{d=d_f}^{\infty} \Big(\sum_{i=1}^{\infty} i\, a(d,i)\Big) P_d \qquad (4.4.8)$$

$$= \sum_{d=d_f}^{\infty} b(d)\, P_d$$

//

(d) $J=2$. Let $i_{J/2} = i_1 = 1$.

$$\Pi(z) = P_0(-1)\, z^{-1} + P_0(1)\, z$$

$$\frac{d}{dz}\Pi(z) = -P_0(-1)\, z^{-2} + P_0(1)$$

at $z^* = \sqrt{\frac{P_0(-1)}{P_0(1)}} < 1$, $\left.\frac{d}{dz}\Pi(z)\right|_{z=z^*} = 0$.

$$\frac{d^2}{dz^2}\Pi(z) = 2P_0(-1)\, z^{-3} > 0 \quad \text{for } 0 < z \le 1.$$

$$\therefore \tilde{Z} = \min_{0 \le z \le 1} \Pi(z) = \Pi(z^*) = P_0(-1)\cdot\sqrt{\frac{P_0(1)}{P_0(-1)}} + P_0(1)\sqrt{\frac{P_0(-1)}{P_0(1)}}$$

$$= 2\sqrt{P_0(1)\, P_0(-1)} = 2\sqrt{p(1-p)} = Z$$

where p is crossover probability.

Here $p = P_0(-1) < \frac{1}{2}$

(e) Proof of the first part for $J \ge 4$ is omitted.

$J=2$:

$$[\Pi(z)]^d = \sum_{i=0}^{d} \binom{d}{i} [P_0(-1)\, z^{-1}]^{d-i} [P_0(1)\, z]^i$$

$$= \sum_{i=0}^{d} \binom{d}{i} P_0(-1)^{d-i} P_0(1)^i z^{-d+2i}$$

$$\frac{1}{2}\tilde{Z}^d = \frac{1}{2}\left[2\sqrt{P_0(-1)P_0(1)}\right]^d = 2^{d-1} P_0(-1)^{\frac{d}{2}} P_0(1)^{\frac{d}{2}}$$

When d is odd;

$$\{[\Pi(z)]^d\}_- \Big/ \frac{1}{2}\tilde{Z}^d$$

$$= 2^{-d+1} P_0(-1)^{-\frac{d}{2}} P_0(1)^{-\frac{d}{2}} \cdot \left\{ \sum_{i=0}^{\frac{d-1}{2}} \binom{d}{i} P_0(-1)^{d-i} P_0(1)^i \right\}$$

$$= 2^{-d+1} \left\{ \sum_{i=0}^{\frac{d-1}{2}} \binom{d}{i} [P_0(-1)/P_0(1)]^{\frac{d}{2}-i} \right\}$$

$$\le 2^{-d+1} \cdot \sum_{i=0}^{\frac{d-1}{2}} \binom{d}{i} = 2^{-d+1} \cdot 2^{d-1} = 1. \qquad \left(\because \frac{P_0(-1)}{P_0(1)} \le 1\right)$$

When d is even;

$$\frac{\{\Pi(z)^d\}_-}{\frac{1}{2}\tilde{Z}^d} = \frac{\frac{1}{2}\binom{d}{d/2} P_0(-1)^{\frac{d}{2}} P_0(1)^{\frac{d}{2}} + \sum_{i=0}^{\frac{d}{2}-1} \binom{d}{i} P_0(-1)^{d-i} P_0(1)^{i}}{2^{d-1} P_0(-1)^{\frac{d}{2}} P_0(1)^{\frac{d}{2}}}$$

$$= 2^{-d+1} \cdot \left\{ \frac{1}{2}\binom{d}{d/2} + \sum_{i=0}^{\frac{d}{2}-1} \binom{d}{i} \left[\frac{P_0(-1)}{P_0(1)}\right]^{\frac{d}{2}-i} \right\}$$

$$\leq 2^{-d+1} \left\{ \frac{1}{2}\binom{d}{d/2} + \sum_{i=0}^{\frac{d}{2}-1} \binom{d}{i} \right\} = 1$$

$$\therefore \; P_d = \{\Pi(z)^d\} \leq \frac{1}{2}\tilde{Z}^d .$$

Therefore

$$P_b \leq \sum_{d=d_f}^{\infty} b(d) P_d \leq \frac{1}{2} \sum_{d=d_f}^{\infty} b(d) \tilde{Z}^d = \frac{1}{2} \frac{\partial}{\partial I} T(D,I) \Big|_{\substack{I=1 \\ D=\tilde{Z}}} . \;//$$

4.21.

$$d = d(\underline{x}, \underline{x}') = \sum_{n=1}^{N} d(x_n, x_n') = -\sum_{n=1}^{N} \ln \sum_{y_n} \sqrt{P(y_n|x_n) P(y_n|x_n')}$$

$$= -\ln \prod_{n=1}^{N} \sum_{y_n} \sqrt{P(y_n|x_n) P(y_n|x_n')}$$

$$= -\ln \sum_{y_1} \sum_{y_2} \cdots \sum_{y_N} \sqrt{\prod_{n=1}^{N} P(y_n|x_n) \prod_{n=1}^{N} P(y_n|x_n')}$$

$$= -\ln \sum_{\underline{y}} \sqrt{P_N(\underline{y}|\underline{x}) P_N(\underline{y}|\underline{x}')}$$

$$P_d = P_E(\underline{x} \to \underline{x}') \leq \sum_{\underline{y}} \sqrt{P_N(\underline{y}|\underline{x}) P_N(\underline{y}|\underline{x}')} \qquad (2.3.15)$$

$$= \exp\left[-\left(-\ln \sum_{\underline{y}} \sqrt{P_N(\underline{y}|\underline{x}) P_N(\underline{y}|\underline{x}')}\right)\right] = e^{-d}$$

$P_e(j) \leq \sum_{\text{all } d} \Pr\{$ error caused by any one of $a(d)$ incorrect path at Bhattacharyya distance $d\}$

$$\leq \sum_{d} a(d) P_d$$

$$\leq \sum_{d} a(d) e^{-d}$$

To define generating function ;

① The Bhattacharyya distance between any two different inputs $x, x' \in \mathcal{X}$ should be same for all $x, x' \in \mathcal{X}$.

② For one input binary data, one signal should be transmitted to channel.

Let k = # of different simbols from correct path.

Then $P_k < \left[\sum_y \sqrt{p(y|x)\,p(y|x')} \right]^k = D_0^{\,k}$

where $D_0 = \sum_y \sqrt{p(y|x)\,p(y|x')}$

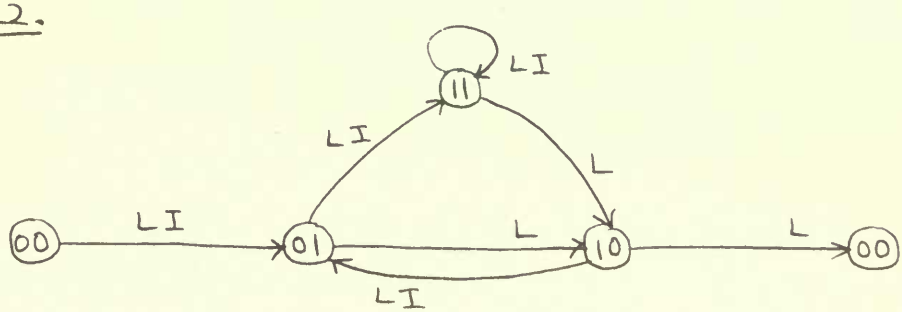
and $P_e(j) \le \sum_k a(k) P_k \le \sum_k a(k) D_0^{\,k}$

Since each branch ha equal distance,

(4.4.7) $\qquad P_e(j) \le T(L)\big|_{L=D_0}$

(4.4.13) $\qquad P_b(j) \le \frac{1}{b} \frac{\partial}{\partial I} T(L,I)\big|_{I=1,\, L=D_0}$. "

4.22.



$$T(L,I) = \frac{L^2 I \{ L(1-LI) + L^2 I \}}{1 - LI - L^2 I - LI \cdot L \cdot LI + LI \cdot L^2 I} = \frac{L^3 I}{1-(L+L^2)I}$$

$$\frac{\partial}{\partial I} T(L,I) = \frac{L^3}{1-(L+L^2)I} + \frac{L^3 I (L+L^2)}{[1-(L+L^2)I]^2} = \frac{L^3}{[1-(L+L^2)I]^2}$$

$$\therefore\ P_b(j) < \frac{\partial}{\partial I} T(L,I)\Big|_{I=1,\,L=D_0} = \frac{D_0^3}{[1-(D_0+D_0^2)]^2}$$

where D_0's are given as follows :

1) Hard decision;

$$D_0 = \sum_y \sqrt{p(y|x)\,p(y|x')} = 2\sqrt{(1-p)\cdot\frac{p}{7}} + 6\sqrt{\frac{p}{7}\cdot\frac{p}{7}}$$

$$= \frac{2}{\sqrt{7}}\sqrt{p(1-p)} + \frac{6}{7}p$$

2) Soft decision ;

when each x has equal energy $\mathcal{E}$,

from (2.3.17),

$$D_0 = \int_{-\infty}^{\infty} \sqrt{p(y|x)\,p(y|x')}\,dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left\{-\frac{1}{2N_0}\left[(y-x)^2+(y-x')^2\right]\right\} dy$$

$$= \exp\left\{-\|x-x'\|^2/4N_0\right\} = \exp\left\{-\mathcal{E}/2N_0\right\}$$ //

4.23.

$$T_K(L) = \frac{L^K(1-L)}{1-2L+L^K}$$

$$\le \frac{L^K(1-L)}{1-2L} = \frac{1}{2}L^K\cdot\frac{2-2L}{1-2L} = \frac{1}{2}L^K\left(1+\frac{1}{1-2L}\right)$$

$$= \frac{1}{2}L^K\left(1+\sum_{k=0}^{\infty} 2^k L^k\right)$$

//

4.24 (a) $\xi_b(D,t+1) = \xi_c(D,t)$

$\xi_c(D,t+1) = D\xi_b(D,t) + D\xi_d(D,t)$

$\xi_d(D,t+1) = D\xi_b(D,t) + D\xi_d(D,t)$

$$\underline{\xi}(D,t+1) = \begin{bmatrix} 0 & 1 & 0 \\ D & 0 & D \\ D & 0 & D \end{bmatrix} \underline{\xi}(D,t) \equiv A\,\underline{\xi}(D,t)$$

$$\underline{\xi}(D,1) = \begin{bmatrix} D^2 \\ 0 \\ 0 \end{bmatrix} ;$$

(b) $\underline{\xi}(D,t+1) = A\,\underline{\xi}(D,t) = A^2\,\underline{\xi}(D,t-1) = A^t\,\underline{\xi}(D,1)$

$$\sum_{t=0}^{\infty} \underline{\xi}(D,t+1) = \sum_{t=0}^{\infty} A^t\,\underline{\xi}(D,1) = (I-A)^{-1}\,\underline{\xi}(D,1)$$

$$(I-A)^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ -D & 1 & -D \\ -D & 0 & 1-D \end{bmatrix}^{-1} = \frac{1}{1-2D} \begin{bmatrix} 1-D & 1-D & D \\ D & 1-D & D \\ D & D & 1-D \end{bmatrix}$$

$$T(D) = D^2 \sum_{t=2}^{\infty} \xi_c(D,t) = [0 \;\; D^2 \;\; 0] \sum_{t=1}^{\infty} \underline{\xi}(D,t+1) = [0 \;\; D^2 \;\; 0](I-A)^{-1}\underline{\xi}(D,1)$$

$$= [0 \;\; D^2 \;\; 0] \begin{bmatrix} 1-D & 1-D & D \\ D & 1-D & D \\ D & D & 1-D \end{bmatrix} \frac{1}{1-2D} = \frac{D^5}{1-2D}$$

(c) $[0 \;\; D^2 \;\; 0]\,\underline{\xi}(D,t)$ corresponds to all paths that return to the all-zero state at $t+1$ branches after diverging. Hence $[0 \;\; D^2 \;\; 0] \sum_{t=1}^{\tau-1} \underline{\xi}(D,t)$ represents all paths that return to the all-zero state by τ branches. If a decision is to be based on metrics computed for only τ branches then paths not yet remerged at τ can also cause decoding errors. These paths have transfer functions $\xi_b(D,\tau)$, $\xi_c(D,\tau)$ and $\xi_d(D,\tau)$.

Hence

$$P_e(j,\tau) \le [0\ D^2\ 0] \sum_{t=1}^{\tau-1} \underset{\sim}{\xi}(D,t) + [1\ 1\ 1]\, \underset{\sim}{\xi}(D,\tau)$$

(d) from 4.18(b) we have

$$\begin{aligned} A^3 &= DA^2 + DA \\ A^4 &= DA^3 + DA^2 = D^2A^2 + DA^2 + D^2A \\ &= (D^2+D)A^2 + D^2A \\ A^5 &= (D^2+D)[DA^2+DA] + D^2A^2 \\ &= (D^3+2D^2)A^2 + (D^3+D^2)A \\ A^6 &= (D^3+2D^2)[DA^2+DA] + (D^3+D^2)A^2 \\ &= (D^4+3D^3+D^2)A^2 + (D^4+2D^3)A \end{aligned}$$

For $\tau = 7$:

$$\sum_{t=1}^{\tau-1} \underset{\sim}{\xi}(D,t) = \sum_{t=0}^{5} A^t \underset{\sim}{\xi}(D,1) = (I + A + A^2 + A^3 + A^4 + A^5)\underset{\sim}{\xi}(D,1)$$

$$= [I + A + A^2 + DA^2 + DA + (D^2+D)A^2 + D^2A + (D^3+2D^2)A^2 + (D^3+D^2)A]\,\underset{\sim}{\xi}(D,1)$$

$$= \{I + (1+D+2D^2+D^3)A + (1+2D+3D^2+D^3)A^2\}\underset{\sim}{\xi}(D,1)$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ D & 0 & D \\ D & 0 & D \end{bmatrix};\quad A^2 = \begin{bmatrix} D & 0 & D \\ D^2 & D & D^2 \\ D^2 & D & D^2 \end{bmatrix};\quad \underset{\sim}{\xi}(D,1) = \begin{bmatrix} D^2 \\ 0 \\ 0 \end{bmatrix}$$

$$A\underset{\sim}{\xi}(D,1) = \begin{bmatrix} 0 \\ D^3 \\ D^3 \end{bmatrix};\quad A^2\underset{\sim}{\xi}(D,1) = \begin{bmatrix} D^3 \\ D^4 \\ D^4 \end{bmatrix}$$

Hence $[0\ D^2\ 0] \sum_{t=1}^{6} \underset{\sim}{\xi}(D,t) = [0\ D^2\ 0] \sum_{t=0}^{5} A^t \underset{\sim}{\xi}(D,1)$

$$= D^5 + 2D^6 + 4D^7 + 4D^8 + D^9$$

Next note that $\underset{\sim}{\xi}(D,7) = A^6\underset{\sim}{\xi}(D,1)$

$$= (2D^3 + D^4)\begin{bmatrix} 0 \\ D^3 \\ D^3 \end{bmatrix} + (D^2 + 3D^3 + D^4)\begin{bmatrix} D^3 \\ D^4 \\ D^4 \end{bmatrix}$$

$[1\ 1\ 1]\underset{\sim}{\xi}(D,7) = D^5 + 9D^6 + 9D^7 + D^8$ and finally

$$P_b(j,7) \le 2D^5 + 11D^6 + 13D^7 + 5D^8 + D^9 \quad \blacksquare$$

4.25

(a) Without loss in generality, let correct path be all-zero path.

$$P_e(j,L) = \Pr\{\text{the decoded path leaves the correct path at node } j\}$$

$$= \Pr\{\text{the paths that initially leave the correct path at node } j \text{ and go to some state } \underline{x} \text{ in } L \text{ branches}\}$$

$$= \Pr\{\bigcup_{\text{all } \underline{x}} [\text{the paths that go to state } \underline{x} \text{ in } L \text{ branches}]\}$$

$$\le \sum_{\text{all } \underline{x}} \Pr\{\text{the paths that go to state } \underline{x} \text{ in } L \text{ branches}\}$$

$$= \sum_{\text{all } \underline{x}} \{\text{generating function with } D = \sqrt{4p(1-p)} \text{ that go to state } \underline{x} \text{ in } L \text{ branches}\}$$

$$= \sum_{\text{all } \underline{x}} \xi_x(D,L)\Big|_{D=\sqrt{4p(1-p)}} = [1\,1 \cdots 1]\,\underline{\xi}(D,L)\Big|_{D=\sqrt{4p(1-p)}}$$

(b)
$$\begin{aligned}
\xi_a(D,t+1) &= \xi_a(D,t) + D^2\xi_c(D,t)\\
\xi_b(D,t+1) &= D^2\xi_a(D,t) + \xi_c(D,t)\\
\xi_c(D,t+1) &= D\xi_b(D,t) + D\xi_d(D,t)\\
\xi_d(D,t+1) &= D\xi_b(D,t) + D\xi_d(D,t)
\end{aligned}$$

$$\therefore\ \underline{\xi}(D,t+1) = \underline{A}\,\underline{\xi}(D,t)$$

where
$$\underline{A} = \begin{bmatrix} 1 & 0 & D^2 & 0 \\ D^2 & 0 & 1 & 0 \\ 0 & D & 0 & D \\ 0 & D & 0 & D \end{bmatrix}.$$

$$\sum_x \xi_x(D,6) = [1\,1\,1\,1]\,\underline{\xi}(D,6) = [1\,1\,1\,1]\,\underline{A}\,\underline{\xi}(D,5)$$

$$= \cdots = [1\,1\,1\,1]\,\underline{A}^5\,\underline{\xi}(D,1)$$

Using the Cayley-Hamilton theorem;

$$|s\underline{I}-\underline{A}| = \begin{vmatrix} s-1 & 0 & -D^2 & 0 \\ -D^2 & s & -1 & 0 \\ 0 & -D & s & -D \\ 0 & -D & 0 & s-D \end{vmatrix}$$

$$= (s-1)\begin{vmatrix} s & -1 & 0 \\ -D & s & -D \\ -D & 0 & s-D \end{vmatrix} + D^2 \begin{vmatrix} 0 & -D^2 & 0 \\ -D & s & -D \\ -D & 0 & s-D \end{vmatrix}$$

$$= (s-1)(s^3 - Ds^2 - Ds) + D^4(-sD + D^2 - D^2)$$

$$= s^4 - s^3(1+D) - s(D^5 - D)$$

$$\therefore \quad \underline{A}^4 = (1+D)\underline{A}^3 + (D^5-D)\underline{A}$$

$$\underline{A}^5 = (1+D)^2\underline{A}^3 + (D^5-D)\underline{A}^2 + (1+D)(D^5-D)\underline{A}$$

$$A^2 = \begin{bmatrix} 1 & D^3 & D^2 & D^3 \\ D^2 & D & D^4 & D \\ D^3 & D^2 & D & D^2 \\ D^3 & D^2 & D & D^2 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 1+D^5 & D^3+D^4 & D^2+D^3 & D^3+D^4 \\ D^2+D^3 & D^5+D^2 & D^4+D' & D^5+D^2 \\ D^3+D^4 & D^2+D^3 & D^5+D^2 & D^2+D^3 \\ D^3+D^4 & D^2+D^3 & D^5+D^2 & D^2+D^3 \end{bmatrix}$$

$$\sum_x \xi_x(D,6) = [1111]\, A^5\, \underline{\xi}(D,1)$$

$$= (1+D)^2 [1111]\, A^3\, \underline{\xi}(D,1) + (D^5-D)[1111]\, A^2\, \underline{\xi}(D,1)$$

$$+ (1+D)(D^5-D)\, A\, \underline{\xi}(D,1)$$

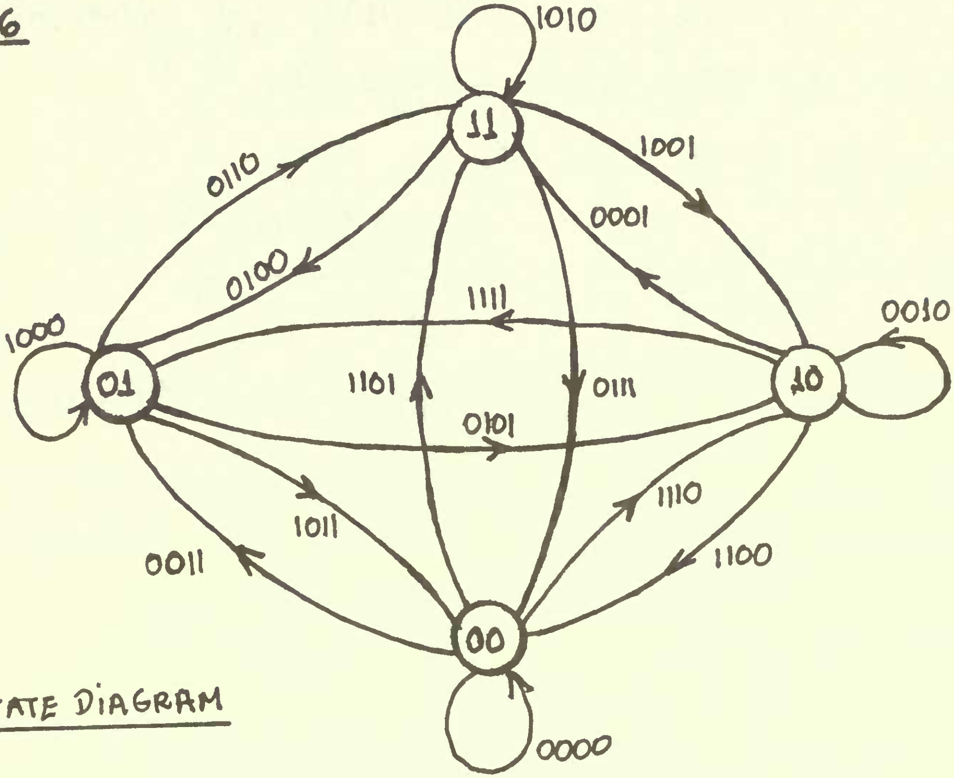
$$= (1+D)^2(3D^2 + 3D^2 + D^4 + D^5)D^2 + (D^5-D)(D+2D^2+D^3)D^2$$

$$+ (1+D)(D^5-D)(2D)D^2$$

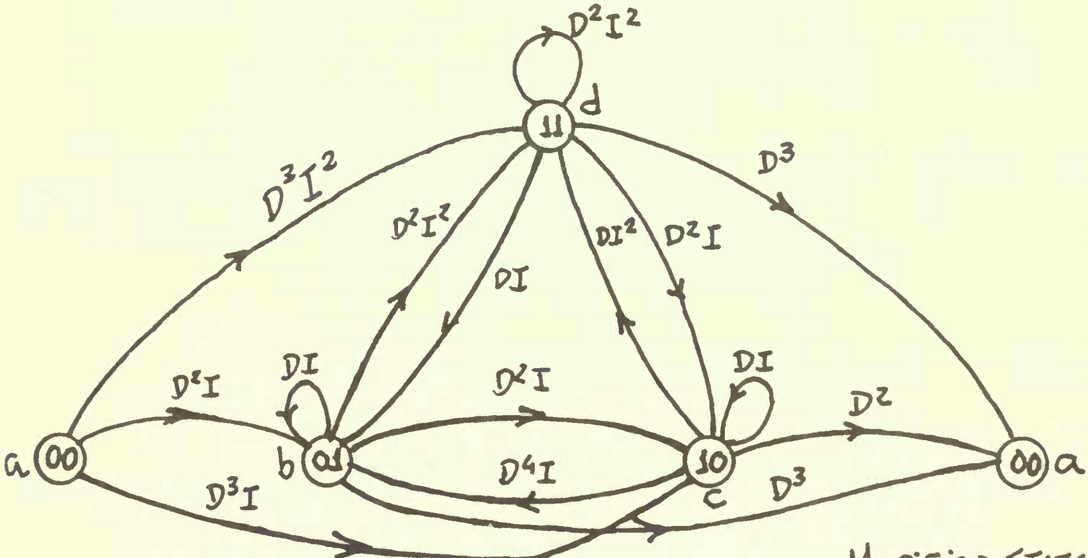
$$= 5D^5 + 9D^6 + 6D^7 + 6D^8 + 5D^9 + D^{10}$$

//

4.26



STATE DIAGRAM



MODIFIED STATE DIAGRAM

$$\xi_b = DI\,\xi_b + D^4 I\,\xi_c + DI\,\xi_d + D^2 I \qquad (1)$$

$$\xi_c = D^2 I\,\xi_b + DI\,\xi_c + D^2 I\,\xi_d + D^3 I \qquad (2)$$

$$\xi_d = D^2 I^2 \xi_b + DI^2 \xi_c + D^2 I^2 \xi_d + D^3 I^2 \qquad (3)$$

$$T(D,I) = D^3 \xi_b + D^2 \xi_c + D^3 \xi_d \qquad (4)$$

$((1)\times D-(2) \Rightarrow)\quad D\xi_b - \xi_c = (D^5 I - DI)\xi_c \Rightarrow \xi_b = \frac{1}{D}(D^5 I - DI + 1)\xi_c$

$((2)\times I-(3) \Rightarrow)\quad I\xi_c - \xi_d = 0 \Rightarrow \xi_d = I\xi_c$

Solving for ξ_c we get $\xi_c = \dfrac{D^3 I}{1-2DI-D^6I^2}$

$\therefore\ \xi_b = \dfrac{(D^5I-DI+1)D^2I}{1-2DI-D^6I^2}\ ;\ \xi_d = \dfrac{D^3I^2}{1-2DI-D^6I^2}$

and

$$T(D,I) = \frac{1}{1-2DI-D^6I^2}\left[D^5I\left(D^5I-DI+1\right)+D^5I + D^6I^2\right]$$

finally

$$T(D,I) = \frac{D^5I\left(2+D^5I\right)}{1-2DI-D^6I^2}$$

From (4.4.13) we have

$$P_b \le \frac{1}{b}\left.\frac{\partial T(D,I)}{\partial I}\right|_{\substack{I=1\\D=Z}}$$

$$\frac{\partial T(D,I)}{\partial I} = \frac{2D^5(1+D^5I)}{(1-2DI-D^6I^2)^2}$$

so finally $\displaystyle P_b \le \frac{1}{2}\left.\frac{2D^5(1+D^5I)}{(1-2DI-D^6I^2)^2}\right|_{\substack{I=1\\D=Z}} = \frac{Z^5+Z^{10}}{(1-2Z-Z^6)^2}$

4.27 $\quad n_k = \int_{-\infty}^{\infty} n(t)h(t-kT)dt$

$$E[n_kn_j] = E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n(\alpha)n(\beta)h(\alpha-kT)h(\beta-jT)\,d\alpha\, d\beta\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E[n(\alpha)n(\beta)]\,h(\alpha-kT)h(\beta-jT)\,d\alpha\, d\beta$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{N_0}{2}\delta(\alpha-\beta)\,h(\alpha-kT)\,h(\beta-jT)\,d\alpha\, d\beta$$

$$= \frac{N_0}{2}\int_{-\infty}^{\infty} h(\alpha-kT)\,h(\alpha-jT)\,d\alpha \equiv \frac{N_0}{2}h_{k-j}$$

Choose $\tilde{y}_k = \int_{-\infty}^{\infty} y(t)\, p(t-kT)\, dt$

$$= \int_{-\infty}^{\infty} x(t)\, p(t-kT)\, dt + \int_{-\infty}^{\infty} n(t)\, p(t-kT)\, dt$$

Let $\tilde{n}_k \equiv \int_{-\infty}^{\infty} n(t)\, p(t-kT)\, dt$

Then

$$E\left[\tilde{n}_k \tilde{n}_j\right] = \frac{N_0}{2} \int_{-\infty}^{\infty} p(t-kT)\, p(t-jT)\, dt = \frac{N_0}{2} \delta_{kj}$$

Thus the noise components of $y_{-N}, y_{-N+1}, \ldots, y_{N-1}$ are all independent. Next note that

$$x(t) = \sum_{k=-N}^{N-1} u_k h(t-kT)$$

hence

$$\int_{-\infty}^{\infty} x(t)\, p(t-kT)\, dt = \sum_{k=-N}^{N-1} u_j \int_{-\infty}^{\infty} h(t-jT)\, p(t-kT)\, dt$$

$$= \sum_{k=-N}^{N-1} u_j \tilde{h}_{k-j}$$

$$= \sum_{i=k-N+1}^{k+N} u_{k-i} \tilde{h}_i$$

$$= \sum_{i=0}^{L-1} u_{k-i} \tilde{h}_i$$

since we assume that $\begin{cases} h(t) = 0 & \text{for } t < 0 \\ h(t) \cong 0 & \text{for } t \geq LT \end{cases}$

This gives the "integrate and dump" filter output as

$$\tilde{y}_k = \sum_{i=0}^{L-1} \tilde{h}_i u_{k-i} + \tilde{n}_k$$

Define $z_k \equiv \sum_{i=0}^{L-1} \tilde{h}_i u_{k-i}$

We have the conditional probability

$$P_{2N}(\underset{\sim}{y}|\underset{\sim}{u}) = \prod_{n=-N}^{N-1} \left(\frac{1}{\pi N_0}\right)^{1/2} e^{-(y_k - z_k)^2/N_0}$$

For two sequences $\underset{\sim}{u}$ and $\underset{\sim}{u}'$ with error $\underset{\sim}{\varepsilon} = \frac{1}{2}(\underset{\sim}{u} - \underset{\sim}{u}')$ we have the pairwise error probability

$$P_E(\underset{\sim}{\varepsilon}) = Q\left(\|\underset{\sim}{z} - \underset{\sim}{z}'\| / \sqrt{2N_0}\right)$$

$$= Q\left(\sqrt{\frac{2}{N_0} \sum_{k=-N}^{N-1} \left(\sum_{i=0}^{L-1} \tilde{h}_i \varepsilon_{k-i}\right)^2}\right)$$

$$< \prod_{k=-N}^{N-1} e^{-\frac{1}{N_0}\left(\sum_{i=0}^{L-1} \tilde{h}_i \varepsilon_{k-i}\right)^2}$$

The maximum likelihood detector chooses $\underset{\sim}{u}$ that maximizes $P_{2N}(\underset{\sim}{y}|\underset{\sim}{u})$ when $\underset{\sim}{y}$ is the observable. Hence

$$\hat{\underset{\sim}{u}} = \max_{\underset{\sim}{u}}{}^{-1} P_{2N}(\underset{\sim}{y}|\underset{\sim}{u}) = \min_{\underset{\sim}{u}}{}^{-1} \|\underset{\sim}{y} - \underset{\sim}{z}(\underset{\sim}{u})\|^2$$

$$= \max_{\underset{\sim}{u}}{}^{-1} \frac{1}{N_0}\left[2(\underset{\sim}{y}, \underset{\sim}{z}(\underset{\sim}{u})) - \|\underset{\sim}{z}(\underset{\sim}{u})\|^2\right]$$

Let

$$\lambda = \frac{2}{N_0}(\underset{\sim}{y}, \underset{\sim}{z}(\underset{\sim}{u})) - \frac{1}{N_0}\|\underset{\sim}{z}(\underset{\sim}{u})\|^2 = \frac{2}{N_0}\sum_{k=-N}^{N-1} y_k z_k(\underset{\sim}{u}) - \frac{1}{N_0}\sum_{k=-N}^{N-1} z_k^2(\underset{\sim}{u})$$

$$= \frac{2}{N_0}\sum_{k=-N}^{N-1}\left[y_k \sum_{i=0}^{L-1} \tilde{h}_i u_{k-i} - \frac{1}{N_0}\left(\sum_{i=0}^{L-1} \tilde{h}_i u_{k-i}\right)^2\right]$$

$$= \frac{1}{N_0}\sum_{k=-N}^{N-1} \lambda_k(y_k; u_k, u_{k-1}, \ldots, u_{k-(L-1)})$$

This metric shows us that the maximum likelihood demodulator can be realized with the Viterbi algorithm. Next we have the bit error bound of eq. (4.9.21)

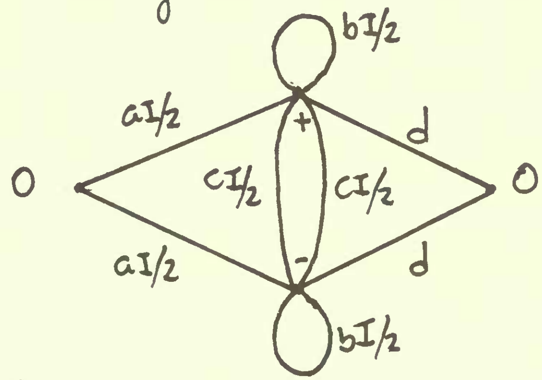
$$P_b \le \sum_{\underset{\sim}{\varepsilon}} w(\underset{\sim}{\varepsilon}) \prod_{k=-N}^{N-1} \frac{1}{2^{w(\underset{\sim}{\varepsilon})}} e^{-\frac{1}{N_0}\left(\sum_{i=0}^{L-1} \tilde{h}_i \varepsilon_{k-i}\right)^2}$$

For $L=2$ we have $2^{L-1}=2$ states with branch values defined by

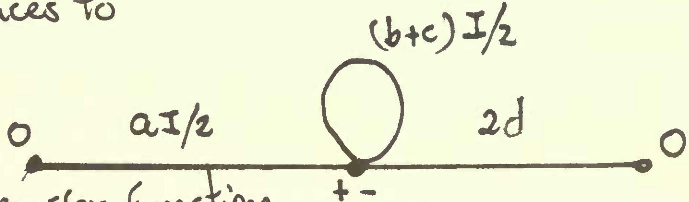
$$a = e^{-\frac{1}{N_0}(\tilde{h}_0)^2} ; \quad b = e^{-\frac{1}{N_0}(\tilde{h}_0+\tilde{h}_1)^2} ;$$

$$c = e^{-\frac{1}{N_0}(\tilde{h}_0-\tilde{h}_1)^2} ; \quad d = e^{-\frac{1}{N_0}(\tilde{h}_1)^2}$$

The error state diagram is



This reduces to



yielding transfer function

$$T(a,b,c,d;I) = \frac{adI}{1-(b+c)I/2}$$

and the bound

$$P_b \le \left.\frac{\partial T}{\partial I}\right|_{I=1} = \frac{ad}{[1-(b+c)/2]^2} = \frac{e^{-\frac{1}{N_0}(\tilde{h}_0^2+\tilde{h}_1^2)}}{\left[1-\frac{1}{2}e^{-\frac{1}{N_0}(\tilde{h}_0^2+\tilde{h}_1^2)}\left(e^{-\frac{2}{N_0}\tilde{h}_0\tilde{h}_1}+e^{+\frac{2}{N_0}\tilde{h}_0\tilde{h}_1}\right)\right]^2}$$

4.28 Assume $\tilde{y}_k = f_0 u_k + f_1 u_{k-1} + \tilde{n}_k$, where $E[\tilde{n}_k \tilde{n}_j] = \frac{N_0}{2}\delta_{kj}$.

(a)

Then $y_k = f_0 \tilde{y}_k + f_1 \tilde{y}_{k+1}$

$$= f_0[f_0 u_k + f_1 u_{k-1} + \tilde{n}_k] + f_1[f_0 u_{k+1} + f_1 u_k + \tilde{n}_{k+1}]$$

$$= (f_0^2 + f_1^2) u_k + f_0 f_1 (u_{k-1} + u_{k+1}) + f_0 \tilde{n}_k + f_1 \tilde{n}_{k+1}$$

$$= h_0 u_k + h_1 (u_{k-1} + u_{k+1}) + f_0 \tilde{n}_k + f_1 \tilde{n}_{k+1}$$

$$E[(f_0 \tilde{n}_k + f_1 \tilde{n}_{k+1})(f_0 \tilde{n}_j + f_1 \tilde{n}_{j+1})]$$

$$= \frac{N_0}{2}\left[f_0^2 \delta_{kj} + f_0 f_1 (\delta_{k,j+1} + \delta_{k+1,j}) + f_1^2 \delta_{kj}\right]$$

$$= \frac{N_0}{2}\left[h_0 \delta_{kj} + h_1 (\delta_{k,j+1} + \delta_{k+1,j})\right] = \frac{N_0}{2} h_{k-j}$$

Hence $y_k = h_0 u_k + h_1(u_{k-1} + u_{k+1}) + n_k$, where $E[n_k n_j] = \frac{N_0}{2} h_{k-j}$

(b) and (c) follow as in problem 4.27 with $\tilde{h}_0 = f_0$ and $\tilde{h}_1 = f_1$.

Then

$$P_b \le \frac{e^{-\frac{1}{N_0}(f_0^2 + f_1^2)}}{\left[1 - \frac{1}{2} e^{-\frac{1}{N_0}(f_0^2 + f_1^2)}\left(e^{-\frac{2}{N_0} f_0 f_1} + e^{+\frac{2}{N_0} f_0 f_1}\right)\right]^2}$$

$$= \frac{e^{-\frac{1}{N_0} h_0}}{\left[1 - \frac{1}{2} e^{-\frac{1}{N_0} h_0}\left(e^{-\frac{2}{N_0} h_1} + e^{+\frac{2}{N_0} h_1}\right)\right]^2}$$

(d) Assume $\tilde{y}_k = \sum_{i=0}^{L-1} f_i u_{k-i} + \tilde{n}_k$. Then

$$y_k = \sum_{i=0}^{L-1} f_i \tilde{y}_{k+i} = \sum_{i=0}^{L-1} f_i \left[\sum_{j=0}^{L-1} f_j u_{k+i-j} + \tilde{n}_{k+i}\right]$$

Now assuming $\begin{cases} h_i = 0 & \text{for } |i| \geq \mathcal{L} \\ f_i = 0 & \text{for } i<0 \text{ and } i \geq \mathcal{L} \end{cases}$

we have

$$H(D) = \sum_i h_i D^i = f(D)\, f(D^{-1})$$

$$= \Big(\sum_j f_j D^j\Big)\Big(\sum_k f_k D^{-k}\Big) = \sum_j \sum_k f_j f_k D^{j-k}$$

$$= \sum_j \sum_i f_j f_{j-i} D^i = \sum_i \Big(\sum_j f_j f_{j-i}\Big) D^i$$

or

$$h_i = \sum_j f_j f_{j-i} \quad ;$$

then $$y_k = \sum_i f_i \Big[\sum_j f_j u_{k+i-j} + \tilde{n}_{k+i}\Big]$$

$$= \sum_i f_i \Big[\sum_\ell f_{i-\ell} u_{k+\ell} + \tilde{n}_{k+i}\Big]$$

$$= \sum_\ell \Big(\sum_i f_i f_{i-\ell}\Big) u_{k+\ell} + \sum_i f_i \tilde{n}_{k+i}$$

$$= \sum_\ell h_\ell u_{k+\ell} + \sum_i f_i \tilde{n}_{k+i} = \sum_{\ell=-\mathcal{L}+1}^{\mathcal{L}-1} h_\ell u_{k+\ell} + \sum_i f_i \tilde{n}_{k+i}$$

$$E\Big[\Big(\sum_i f_i \tilde{n}_{k+i}\Big)\Big(\sum_j f_j \tilde{n}_{\ell+j}\Big)\Big] = \sum_i \sum_j f_i f_j \frac{N_0}{2} \delta_{k+i,\ell+j}$$

$$= \frac{N_0}{2} \sum_i f_i f_{i+k-\ell} = \frac{N_0}{2} h_{k-\ell}$$

and this is consistent.

(e) Same as in problem 4.27 with $\tilde{h}_i = f_i$ for $i = 0, \ldots, \mathcal{L}-1$.

(f) $$\sum_n \left(\sum_{i=0}^{L-1} f_i \varepsilon_{n-i}\right)^2 = \sum_n \left(\sum_i \sum_j f_i f_j \varepsilon_{n-i} \varepsilon_{n-j}\right)$$

$$= \sum_n \left(\sum_i \sum_j f_{n-i} f_{n-j} \varepsilon_i \varepsilon_j\right)$$

$$= \sum_i \sum_j \left(\sum_n f_{n-i} f_{n-j}\right) \varepsilon_i \varepsilon_j = \sum_i \sum_j h_{i-j} \varepsilon_i \varepsilon_j$$

$$= \sum_n \sum_k h_k \varepsilon_n \varepsilon_{n-k}$$

$$= \sum_n h_0 \varepsilon_n^2 + 2 \sum_n \sum_{k=1}^{L-1} h_k \varepsilon_n \varepsilon_{n-k}$$

$$= \sum_n \left(\varepsilon_n^2 h_0 + 2 \sum_{k=1}^{L-1} \varepsilon_n \varepsilon_{n-k} h_k \right)$$

4.29.

(a) For a non-coherent MFSK, squared matched filter - envelope detector outputs for each chip (or equivalently sums of the squares of the in-phase and quadrature correlator outputs) are summed to form the decision statistics.

Let $z_j = x_j^2 + y_j^2$ where x_j is the in-phase correlator output and y_j the quadrature correlator output for the j^{th} chip.

Using Chernhoff bound; $\Pr\{z>0\} \le E[e^{\rho z}]$, $\rho>0$ (see prob. 2.10. b)

$$P_J = \Pr\left\{\sum_{j=1}^{J} z_j' > \sum_{j=1}^{J} z_j\right\} = \Pr\left\{\sum_{j=1}^{J} (z_j' - z_j) > 0\right\}$$

$$\le E\left[\exp\left\{\rho \sum_{j=1}^{J} (z_j' - z_j)\right\}\right] = \prod_{j=1}^{J} E[\exp(\rho(z_j' - z_j))]$$

$$= \prod_{j=1}^{J} E[\exp\{\rho(x_j'^2 + y_j'^2)\}] \cdot E[\exp\{-\rho(x_j^2 + y_j^2)\}]$$

and $E(x_j^2)+E(y_j^2) = 2\mathcal{E}_b/N_0J$

$$E[\exp\{-\rho(x_j^2+y_j^2)\}] = \exp\left(-\frac{\rho}{1+2\rho}\frac{\mathcal{E}_b}{N_0}\right)/(1+2\rho)$$

$$E[\exp\{\rho(x_j^2+y_j^2)\}] = 1/_{1-2\rho} \qquad , \quad 0<\rho<\tfrac{1}{2}$$

Letting $2\rho=\lambda$, $0<\lambda<1$.

$$P_J < \prod_{j=1}^{J} \frac{1}{1-\lambda^2}\exp\left[-\frac{\lambda}{1+\lambda}\frac{\mathcal{E}_b}{N_0}\right]$$

$$< \max_{0<\rho<1}\left[\frac{1}{1-\rho^2}\exp\left[-\frac{\rho}{1+\rho}\frac{\mathcal{E}_b}{N_0}\right]\right]^J = Z^J$$

where $\quad Z = \max_{0<\rho<1}\frac{1}{1-\rho^2}\exp\left[-\frac{\rho}{1+\rho}\frac{\mathcal{E}_b}{N_0}\right]$ //

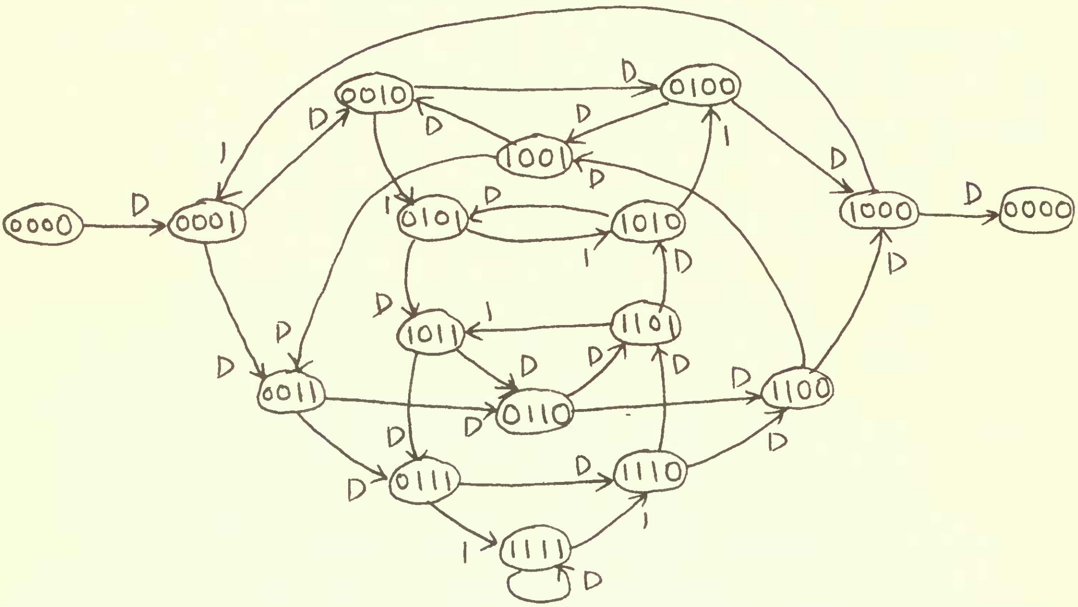
(b) Since no more than 2^j paths merge after $K+j$ branches and a merger error can cause no more than $j+1$ bit errors, the bit error probability, averaged over the ensemble of K code, is bounded by

$$P_b < \sum_{j=0}^{\infty}(j+1)2^j P_{K+j} < \sum_{j=0}^{\infty}(j+1)2^j Z^{K+j} = \frac{Z^K}{(1-2Z)^2}$$ //

4.30.

(a) Suppose the all-zero signal goes into the encoder. In the Trellis diagram, suppose the decoded path diverges from the correct path and remerges after J branches. Since the distances between any two different signals x, x' are same for all $x, x' \in \mathcal{X}$, the labels on the diagram are either D (when either one of modulo-2 adder output is "1") or 1 (when both of adder outputs are "0").

Then its state diagram labeled with distances from all zero path is shown below:



Using Mason's rule (see S.J. Mason "Feedback Theory – Some Properties of Signal Flow Graphs." Proc. IRE, vol. 41, no. 9, pp. 1144–1159, Sept. 1953), we have

$$T = \frac{1}{\Delta} \sum_{j=1}^{L} T_j \Delta_j$$

with $\Delta = 1 - \sum_{i=1}^{N} P_i^{(1)} + (-1)^2 \sum_{i=1}^{m_2} P_i^{(2)} + \cdots + (-1)^h \sum_{i=1}^{m_h} P_i^{(h)} + \cdots + (-1)^H \sum_{i=1}^{m_H} P_i^{(H)}$

where L = # of all possible forward paths. N = # of feedback loops

m_h = # of possible combinations of h nontouching loops.

H = largest possible # of non touching loops.

T_j = Distance of j th forward path.

$P_i^{(h)}$ = product of distances of i th combination of h nontouching loops.

Δ_j = value of Δ for that part of signal flow diagram not touching j th forward path.

Example of Mason's rule; For figure 4.10.

$$T(D,L,I) = \frac{D^2LI \cdot DLI \cdot DL \cdot D^2L + D^2LI \cdot DL \cdot D^2L(1-DLI)}{1 - DLI - DL \cdot LI - DLI \cdot DL \cdot LI + DLI \cdot (DL \cdot LI)}$$

$$= \frac{D^4L^2I\{D^2L^2I + DL(1-DLI)\}}{1-DL(1+L)I} = \frac{D^5L^3I}{1-DL(1+L)I} \qquad (4.3.3)$$

Therefore we have for our state diagram,

$$T(D) = \frac{1}{\Delta} \cdot \left\{ D^4 \cdot \tilde{\Delta}_1 + D^5 \cdot \tilde{\Delta}_2 + D^6 \tilde{\Delta}_3 + D^6 \tilde{\Delta}_4 + \cdots + D^{12} \tilde{\Delta}_\alpha \right\}$$

where $\tilde{\Delta}_\beta$ has the form;

$$\tilde{\Delta}_\beta = 1 + \sum_{l=1} c_l D^l \quad , \quad c_l \text{ is non-negative integer.}$$

since all loop has the distance D^m where m is a positive integer. For the same reason, Δ has the same form as $\tilde{\Delta}_\beta$.

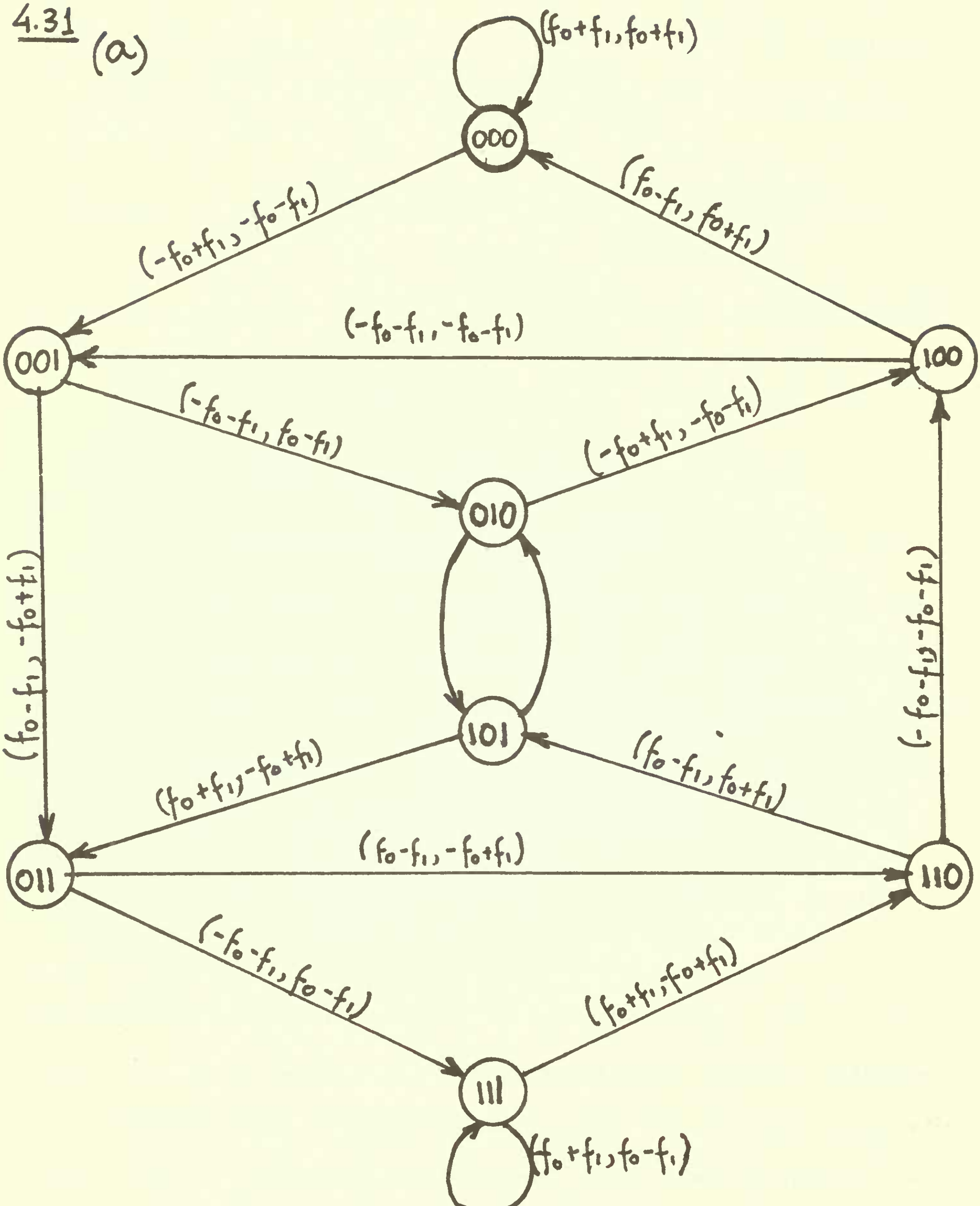
$$\therefore\ T(D) = D^4 \cdot \frac{1 + D \cdot a(D)}{1 + D \cdot b(D)}$$

$$\therefore\ P_b < T(D)\Big|_{D=Z} = Z^4 \frac{1 + Z \cdot a(Z)}{1 + Z \cdot b(Z)}$$ //

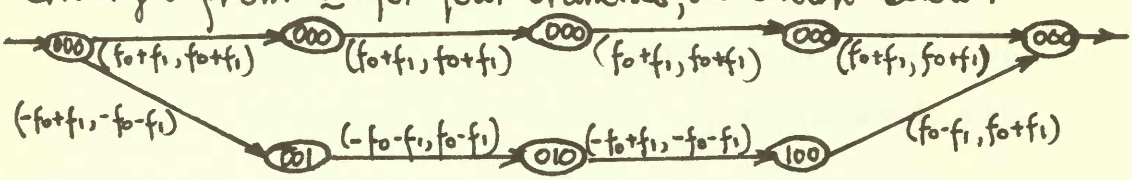
(b) For the general class of semi-orthogonal convolutional encoder with $k \geq 3$ or $K \geq 7$, the smallest distance path in the state diagram has distance D^K when only one "1" is decoded. Therefor $T(D)$ has the following form: $T(D) = D^K\{1 + D\,a(D)\}/\{1 + D\,b(D)\}$

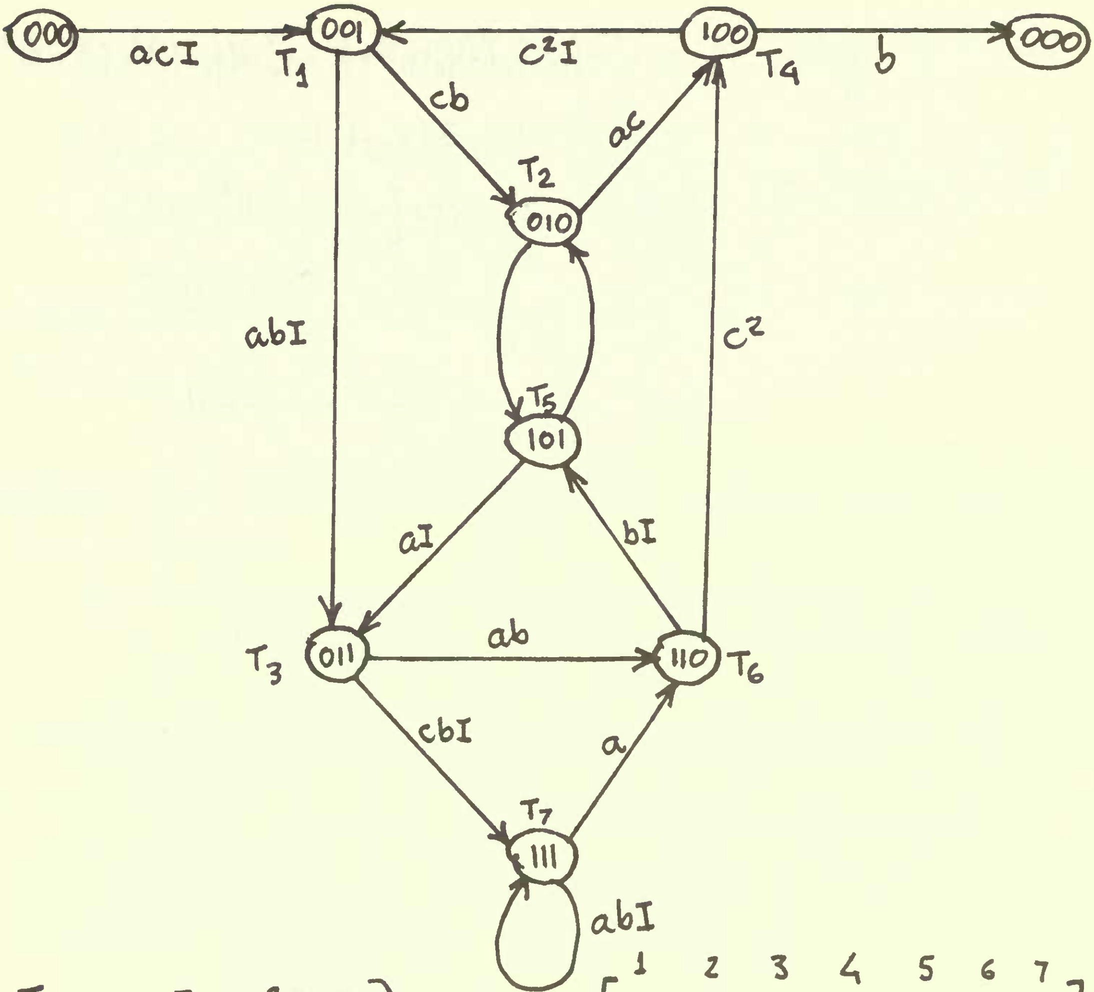
$$\therefore\ P_b \leq T(D)\Big|_{D=Z} = Z^K \frac{1 + Z \cdot a(Z)}{1 + Z \cdot b(Z)}$$ //

4.31 (a)



(b) $\underset{\sim}{u} = \underset{\sim}{0}$ yields the output sequence $\underset{\sim}{z}$, where $z_k = f_0 + f_1$ for all k. For $\underset{\sim}{u}'$, where $u'_k = \delta_{k0}$, the path corresponding to $\underset{\sim}{u}'$ diverges from $\underset{\sim}{u}$ for four branches, as shown below:





$$\left.\begin{aligned}
T_1 &= acI + c^2 I\, T_4\\
T_2 &= bc\,T_1 + ab\,T_5\\
T_3 &= abI\,T_1 + aI\,T_5\\
T_4 &= ac\,T_2 + c^2\,T_6\\
T_5 &= I\,T_2 + bI\,T_6\\
T_6 &= ab\,T_3 + a\,T_7\\
T_7 &= bcI\,T_3 + abI\,T_7
\end{aligned}\right\} \Rightarrow A = \begin{array}{c|ccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7\\ \hline
1 & 0 & 0 & 0 & c^2I & 0 & 0 & 0\\
2 & bc & 0 & 0 & 0 & ab & 0 & 0\\
3 & abI & 0 & 0 & 0 & aI & 0 & 0\\
4 & 0 & ac & 0 & 0 & 0 & c^2 & 0\\
5 & 0 & I & 0 & 0 & 0 & bI & 0\\
6 & 0 & 0 & ab & 0 & 0 & 0 & a\\
7 & 0 & 0 & bcI & 0 & 0 & 0 & abI
\end{array}$$

$\therefore$

$$\left.\begin{aligned}
\underset{\sim}{T} &= A\,\underset{\sim}{T} + \underset{\sim}{B}\\
\underset{\sim}{T} &= (I_{7\times 7} - A)^{-1}\,\underset{\sim}{B}
\end{aligned}\right. \Leftarrow \left\{\; B^T = \begin{bmatrix} acI & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right.$$

Note that $\frac{1}{4}\|\underset{\sim}{z}-\underset{\sim}{z}'\|^2 = f_0^2+(f_0+f_1)^2+(f_0+f_1)^2+f_1^2+f_0^2+(f_0+f_1)^2+f_1^2+0^2$

$$= 2f_0^2+2f_1^2+3(f_0+f_1)^2$$

$$P_E(\underset{\sim}{u}\to\underset{\sim}{u}') = Q\left(\frac{\|\underset{\sim}{z}-\underset{\sim}{z}'\|}{\sqrt{2N_0}}\right) < \exp\left\{-\|\underset{\sim}{z}-\underset{\sim}{z}'\|^2/4N_0\right\}$$

$$= e^{-\frac{2f_0^2+2f_1^2+3(f_0+f_1)^2}{N_0}}$$

$$= e^{-\frac{5f_0^2+5f_1^2+6f_0f_1}{}}$$

$$= e^{-\frac{5h_0+6h_1}{N_0}}$$

where we used $h_0 = f_0^2+f_1^2$ and $h_1 = f_0f_1$.

(c) For any sequence $\underset{\sim}{u}'$ with $\underset{\sim}{z}'$, the pairwise error bound is

$$P_E(\underset{\sim}{u}\to\underset{\sim}{u}') = Q\left(\|\underset{\sim}{z}'-\underset{\sim}{z}\|/\sqrt{2N_0}\right) < e^{-\frac{\|\underset{\sim}{z}'-\underset{\sim}{z}\|^2}{4N_0}}$$

$$= \prod_k e^{-\frac{(z_k'-z_k)^2}{4N_0}}$$

Thus we assign branch values $e^{-\frac{(z_k'-f_0-f_1)^2}{4N_0}}e^{-\frac{(z_{k+1}'-f_0-f_1)^2}{4N_0}}$ to each of the branches with outputs z_k', z_{k+1}'. (Here we have $z_k = z_{k+1} = f_0+f_1$) Also we use the parameter I for those branches corresponding to "1" data symbols. Define

$$a \equiv e^{-f_0^2/N_0} \; ; \; b \equiv e^{-f_1^2/N_0} \; ; \; c \equiv e^{-(f_0+f_1)^2/N_0}$$

Then we have the modified state diagram on the following page

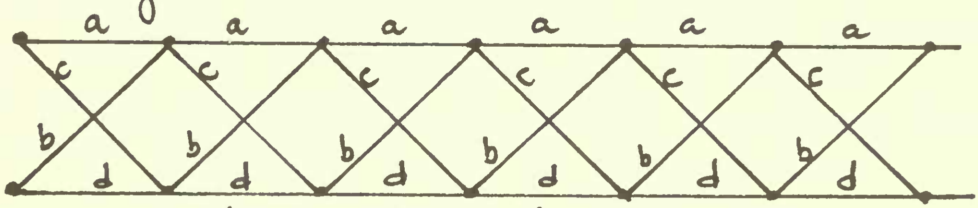
We're only interested in $T(a,b,c;I) = bT_4$ so

$$T(a,b,c;I) = [0\ 0\ 0\ b\ 0\ 0\ 0]\,\underset{\sim}{T} = \underset{\sim}{C}^T(I_{7\times 7} - A)^{-1}\underset{\sim}{B}$$

$$= abc\,I\,M_{14}$$

where M_{14} is the first column, fourth row component of $(I_{7\times 7} - A)^{-1}$.

<u>4.32</u>(a) Let $s_k = x_{k-1}$ be the state and we have the trellis diagram



$\underset{\sim}{y} =$ a b c b a a

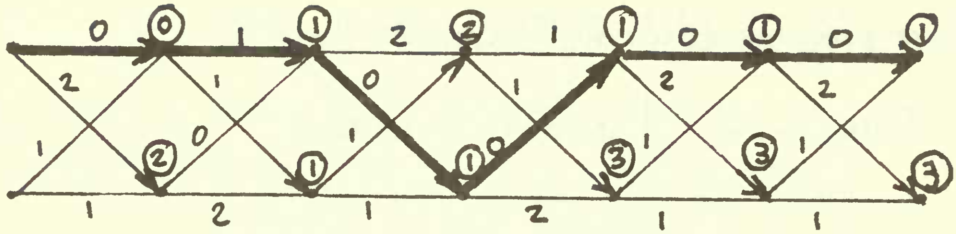
We want to find the path with outputs $\underset{\sim}{z}$ that maximizes $p(\underset{\sim}{y}|\underset{\sim}{z})$ or $\log p(\underset{\sim}{y}|\underset{\sim}{z}) = \sum_k \log p(y_k|\ _k)$

or minimizes $m(\underset{\sim}{y},\underset{\sim}{z}) = \sum_k m(y_k, z_k)$ where

$m(y,z) = -\left[\log\left(\frac{p(y|z)}{q^2}\right)\right] / \log\left(\frac{q}{p}\right)$, is given by

$$\{m(y,z)\} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{bmatrix} \end{array}$$

The maximum likelihood data sequence is then
$\hat{x}_1 = 0$, $\hat{x}_2 = 0$, $\hat{x}_3 = 1$, $\hat{x}_4 = 0$, $\hat{x}_k = 0$ for $k \geq 5$
and branch metrics are shown in the trellis diagram



(b) $P_E(\underset{\sim}{x} \to \underset{\sim}{x}') \le \sum_{\underset{\sim}{y}} \sqrt{p(\underset{\sim}{y}|\underset{\sim}{x})\,p(\underset{\sim}{y}|\underset{\sim}{x}')}$

$$= \prod_k \sum_y \sqrt{p(y|z_k)\,p(y|z_k')}$$

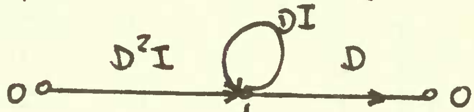
where $z_k = f(x_k, x_{k-1})$ and $z_k' = f(x_k', x_{k-1}')$.

But $\underset{\sim}{x} = \underset{\sim}{0}$ so $z_k = a$ for all k. Hence

$$\sum_y \sqrt{p(y|a)\,p(y|z')} = \begin{cases} 1 & z' = a \\ 2\sqrt{pq} & z' = b \text{ or } z' = d \\ 4pq & z' = c \end{cases}$$

$$= \begin{cases} D^0 & z' = a \\ D^1 & z' = b \text{ or } z' = d \\ D^2 & z' = c \end{cases}$$

where $D = \sqrt{4pq}$. The state diagram with D and I is



yielding transfer function $T(D,I) = D^3 I/(1-DI)$ and hence the bound

$$P_b \le \partial T(D,I)/\partial I \Big|_{I=1} = D^3/(1-D)^2$$

(c) This DMC is the same as two uses of a BSC with crossover p. Hence $E_o(\rho)$, R_o, $E_x(\rho)$ and C are all two times the corresponding values for the BSC.

4.33 (a) Using the same reasoning as in problem 4.27 we have

$$\text{actual: } \tilde{y}_k = \sum_i u_i \tilde{h}_{k-i} + \tilde{n}_k \equiv s_k + \tilde{n}_k$$

$$\text{assumed } \hat{y}_k = \sum_i u_i \hat{h}_{k-i} + \tilde{n}_k \equiv \hat{s}_k + \tilde{n}_k$$

where
$$\begin{cases} \tilde{h}_{k-i} = \int_{-\infty}^{\infty} h(t-iT)\, p(t-kT)\, dt \\ h_{k-i} = \int_{-\infty}^{\infty} \hat{h}(t-iT)\, p(t-kT)\, dt \end{cases}$$

Assuming $\underset{\sim}{u}$ is sent, we conclude $\hat{\underset{\sim}{s}}$ is sent. Similarly for $\underset{\sim}{u}'$ we suppose $\hat{\underset{\sim}{s}}'$ is sent. The actual received sequence is $\underset{\sim}{y} = \underset{\sim}{s} + \tilde{\underset{\sim}{n}}$ but in the decision rule we compare what is supposed to be sent ie, $\hat{\underset{\sim}{y}} = \hat{\underset{\sim}{s}} + \tilde{\underset{\sim}{n}}$. The pairwise error probability $P(\underset{\sim}{u} \to \underset{\sim}{u}')$ is, for the AWGN channel,

$$P_E(\underset{\sim}{u} \to \underset{\sim}{u}') = P\{ \|\hat{\underset{\sim}{s}} - \underset{\sim}{y}\| \geq \|\hat{\underset{\sim}{s}} - \underset{\sim}{y}\| \mid \underset{\sim}{u} \text{ is sent} \}$$

$$= \sum_{\hat{\Lambda}} P(\underset{\sim}{y} \mid \underset{\sim}{u}) = \sum_{\hat{\Lambda}} P(\underset{\sim}{y} \mid \underset{\sim}{s})$$

where $\hat{\Lambda} = \{ \underset{\sim}{y} : P(\underset{\sim}{y} \mid \hat{\underset{\sim}{s}}') \geq P(\underset{\sim}{y} \mid \hat{\underset{\sim}{s}}) \}$

$$P(\underset{\sim}{y} \mid \underset{\sim}{u}) = \left(\frac{1}{\pi N_0}\right) e^{-\|\underset{\sim}{y} - \underset{\sim}{s}\|^2 / N_0}$$

and
$$p(\underset{\sim}{y} \mid \hat{\underset{\sim}{s}}) = \left(\frac{1}{\pi N_0}\right)^N e^{-\|\underset{\sim}{y} - \hat{\underset{\sim}{s}}\|^2 / N_0}$$

Define an indicator function $f(\underset{\sim}{y})$ as

$$f(\underset{\sim}{y}) = \begin{cases} 1 & \underset{\sim}{y} \in \hat{\Lambda} \\ 0 & \underset{\sim}{y} \notin \hat{\Lambda} \end{cases}$$

Then

$$P_E(\underset{\sim}{u} \to \underset{\sim}{u}') = \sum_{\underset{\sim}{y}} f(\underset{\sim}{y})\, p(\underset{\sim}{y}|\underset{\sim}{s})$$

$f(\underset{\sim}{y})$ can be bounded as

$$f(\underset{\sim}{y}) = \begin{cases} 1 \le \sqrt{\dfrac{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}}')}{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}})}} & \underset{\sim}{y} \in \hat{\Lambda} \\[2ex] 0 < \dfrac{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}}')}{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}})} & \underset{\sim}{y} \notin \hat{\Lambda} \end{cases}$$

Hence $P_E(\underset{\sim}{u} \to \underset{\sim}{u}') \le \sum_{\underset{\sim}{y}} \sqrt{\dfrac{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}}')}{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}})}}\, p(\underset{\sim}{y}|\underset{\sim}{s})$

Now

$$p(\underset{\sim}{y}|\underset{\sim}{s}) = \prod_{i=-N}^{N-1} \frac{1}{\sqrt{\pi N_0}}\, e^{-(y_i - s_i)^2/N_0}$$

and $$\sqrt{\frac{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}}')}{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}})}} = \prod_{n=-N}^{N-1} e^{-\frac{1}{2N_0}\left\{(y_n - \hat{s}_n')^2 - (y_n - \hat{s}_n)^2\right\}}$$

so that

$$\sqrt{\frac{p(\underset{\sim}{y}|\hat{\underset{\sim}{s}}')}{p(\underset{\sim}{y}|\underset{\sim}{s})}}\, p(\underset{\sim}{y}|\underset{\sim}{s}) = \prod_{i=-N}^{N-1} \frac{1}{\sqrt{\pi N_0}}\, e^{-\frac{1}{2N_0}\left\{(y_i - \hat{s}_i')^2 - (y_i - \hat{s}_i)^2 + 2(y_i - s_i)^2\right\}}$$

Hence

$$P_E(\underset{\sim}{u} \to \underset{\sim}{u}') \le \prod_{i=-N}^{N-1} e^{-R_i/4N_0}$$

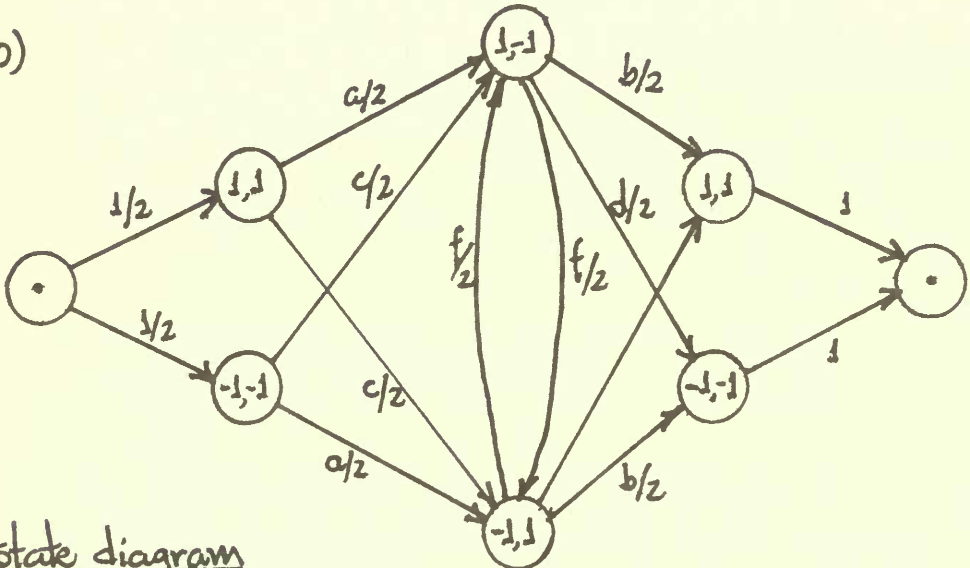
where

$$\begin{aligned} R_i &= 2(\hat{s}_i'^2 - \hat{s}_i^2 + 2s_i^2) - (\hat{s}_i' - \hat{s}_i + 2s_i)^2 \\ &= (\hat{s}_i' - \hat{s}_i)(\hat{s}_i' + 3\hat{s}_i - 4s_i) \\ &= \left\{\sum_{j=0}^{L-1} \hat{h}_j (u_{i-j}' - u_{i-j})\right\}\left\{\sum_{k=0}^{L-1} \left[\hat{h}_k (u_{i-k}' + 3u_{i-k}) - 4\tilde{h}_k u_{i-k}\right]\right\} \end{aligned}$$

Given $\underset{\sim}{u}$ and $\underset{\sim}{u}'$ there are $w(\underset{\sim}{u}, \underset{\sim}{u}')$ places where there may be a bit error. If we sum over all $\underset{\sim}{u}$ & $\underset{\sim}{u}'$ (possibly infinite sequences) we can bound the bit error probability as

$$P_b \le \sum_{\underset{\sim}{u}} \sum_{\underset{\sim}{u}' \neq \underset{\sim}{u}} w(\underset{\sim}{u}, \underset{\sim}{u}') \prod_{k=1}^{\infty} e^{-R_k/4N_0}$$

(b)



state diagram

where

$$a = e^{-\frac{1}{N_0}\left(-\hat{h}_0^2 + 2\hat{h}_0\tilde{h}_0 + 2\hat{h}_0\hat{h}_1 - 2\hat{h}_0\tilde{h}_1\right)}$$

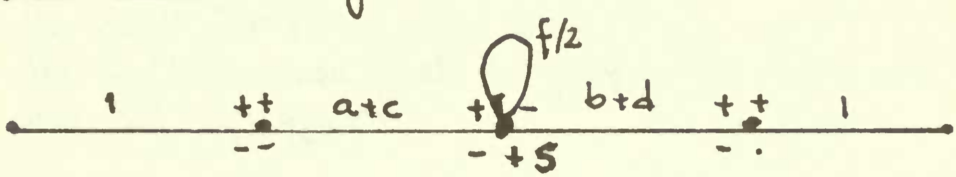
$$b = e^{-\frac{1}{N_0}\left(-\hat{h}_1^2 + 2\hat{h}_1\tilde{h}_1 + 2\hat{h}_0\hat{h}_1 - 2\tilde{h}_0\hat{h}_1\right)}$$

$$c = e^{-\frac{1}{N_0}\left(-\hat{h}_0^2 + 2\hat{h}_0\tilde{h}_0 - 2\hat{h}_0\hat{h}_1 + 2\hat{h}_0\tilde{h}_1\right)}$$

$$d = e^{-\frac{1}{N_0}\left(-\hat{h}_1^2 + 2\hat{h}_1\tilde{h}_1 - 2\hat{h}_0\hat{h}_1 + 2\tilde{h}_0\hat{h}_1\right)}$$

$$f = e^{-\frac{1}{N_0}\left(-\hat{h}_0 + \hat{h}_1\right)\left(\hat{h}_0 - 2\tilde{h}_0 - \hat{h}_1 + 2\tilde{h}_1\right)}$$

reduced state diagram



state equations:

$$s = a + c + f/2\, s$$

$$= \frac{a+c}{1-f/2}$$

$$T(a,b,c,d,f) = (b+d)\,s$$

$$= \frac{(a+c)(b+d)}{1-f/2}$$

—//—

Chapter 5

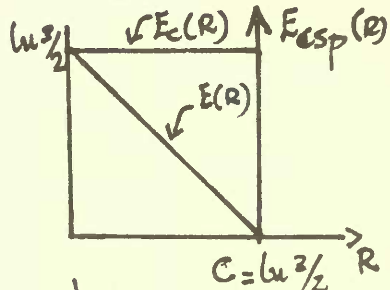
5.1 (a) In problem 3.2, channel (i), we had

$E_o(\rho) = \rho \ln 3/2$ and $E(R) = \ln 3/2 - R = C - R$; $0<R<C$.

$$\left.\begin{array}{l} R_o = E_o(1) = \ln 3/2 \\ C = E_o'(0) = \ln 3/2 \end{array}\right\} \Rightarrow E_c(R) = E_o(1) = \ln 3/2 \quad ; 0<R<C=R_o.$$

$E_{sp}(R) = E_o(\rho) = \rho \ln 3/2$

$$R = \frac{E_o(\rho)}{\rho} = \ln 3/2 \qquad 0<\rho<\infty$$



For channel (ii) we had $E_o(\rho) = \rho \ln 2$ and so everything is the same as above with $\ln 3/2$ replaced by $\ln 2$.

(b) In problem 3.3 (a) we had $E_o(\rho) = \rho \ln Q$ so again everything is as in part (a) with $\ln Q$ replacing $\ln 3/2$.

In 3.3(b) we had:

$$E_o(\rho) = \rho \ln Q - (1+\rho) \ln\left[(Q-1)\, p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}\right]$$

$$E_o(1) = \ln Q - 2 \ln\left[(Q-1)\sqrt{p} + \sqrt{\bar{p}}\right]$$

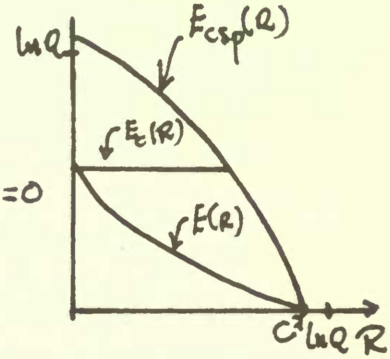
$$C = \ln Q + (Q-1)p \ln p + \bar{p} \ln \bar{p}$$

$E_c(R) = E_o(\rho)$; $R_o < R < C = E_o'(\rho)\big|_{\rho=0}$

$R = E_o(\rho)/\rho$

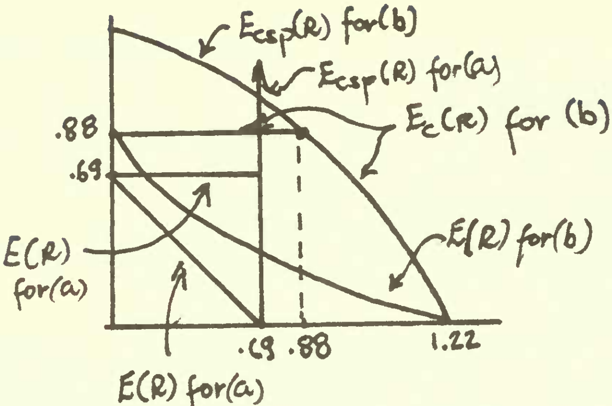
$R_o = E_o(1)$; $0<R<R_o$

$E_{csp}(R) = E_o(\rho)$ $\quad 0<\rho<\infty$

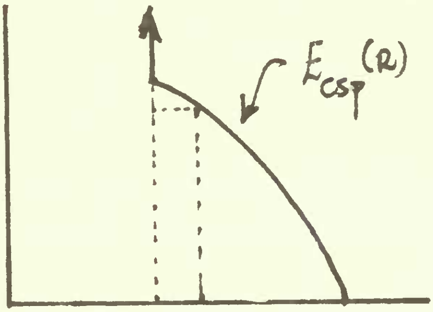


For 3.3(c) use the result for 3.3(a) with $Q=2$ and the result for 3.3(b) with $Q=4$.

For $E_c(R)$ we use just the four inputs a_1, a_2, a_3 & a_4 and set $q_5 = q_6 = 0$.



for $E_{csp}(R)$ we take the upper envelope so we will have



(c) from problem 3.5 we have

$$E_o(\rho) = \rho \ln\left[1 + \left(2^{1/1+\rho} - 1\right)^{\frac{1+\rho}{\rho}}\right]$$

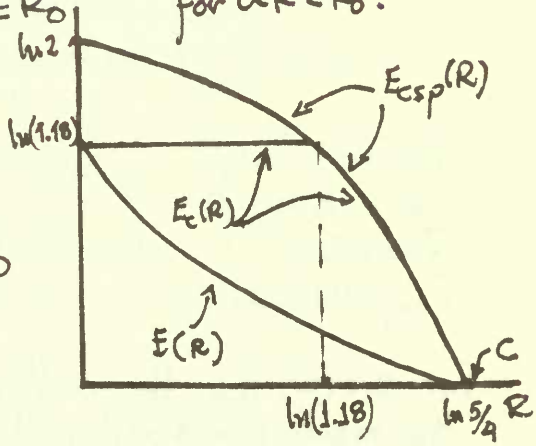
$$E_o(1) = \ln(1.18) \quad ; \quad C = \ln 5/4$$

$$R_o = \ln(1.18) \quad ; \quad E_c(R) = R_o \quad \text{for } 0 < R < R_o .$$

$$\left.\begin{aligned} E_c(R) &= E_o(\rho) \\ R &= E_o(\rho)/\rho \end{aligned}\right\} ; R_o < R < C$$

$$\left.\begin{aligned} E_{csp}(R) &= E_o(\rho) \\ R &= E_o(\rho)/\rho \end{aligned}\right\} \quad 0 < \rho < \infty$$

$$\lim_{\rho \to \infty} E_o(\rho) = \ln 2$$



(d) from problem 3.6 we have $E_o(\rho) = \ln 2 - \ln\left(1 + Q^{-\rho}\right)$

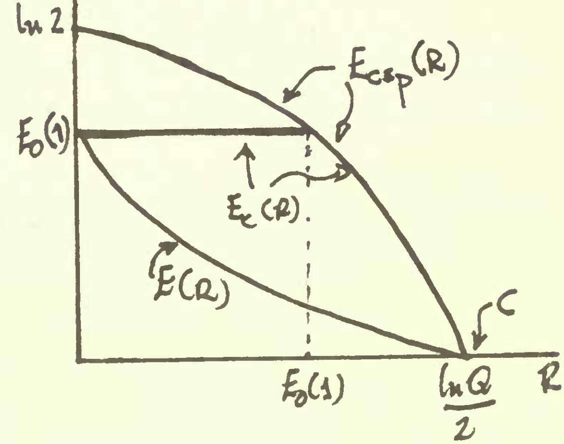
$E_0(1) = \ln 2 - \ln(1+Q^{-1})$; $C = (\ln Q)/2$

$E_c(R) = E_0(1) = R_0$, $0 < R < R_0$;

$$\left.\begin{aligned} E_c(R) &= \ln 2 - \ln(1+Q^{-\rho}) \\ R &= \ln 2/\rho - (\ln(1+Q^{-\rho}))/\rho \end{aligned}\right\} , R_0 < R < C ;$$

$$\left.\begin{aligned} E_{csp}(R) &= E_0(\rho) \\ R &= E_0(\rho)/\rho \end{aligned}\right. ; 0 < \rho < \infty$$

$$\lim_{\rho \to \infty} E_0(\rho) = \ln 2$$



5.2 $t/R \geq E_c(R)/E(R)$

$$E(R) = \begin{cases} C/2 - R & \text{for } 0 \leq R \leq C/4 \\ (\sqrt{C} - \sqrt{R})^2 & \text{for } C/4 < R \leq C \end{cases}$$

$$E_c(R) = \begin{cases} C/2 & \text{for } 0 \leq R \leq C/2 \\ C - R & \text{for } C/2 \leq R \leq C \end{cases}$$

therefore

$$\frac{t}{R} \geq \frac{E_c(R)}{E(R)} = \begin{cases} \dfrac{C/2}{C/2 - R} = \dfrac{1}{1 - 2R/C} & ; 0 \leq R < C/4 \\[2ex] \dfrac{C/2}{(\sqrt{C} - \sqrt{R})^2} = \dfrac{1}{2(1 - \sqrt{R/C})^2} & ; C/4 \leq R \leq C/2 \\[2ex] \dfrac{C - R}{(\sqrt{C} - \sqrt{R})^2} = \dfrac{1 - R/C}{(1 - \sqrt{R/C})^2} & ; C/2 \leq R \leq C \end{cases}$$

5.3 (a) $M = 2^{bL}$; $R_b = \dfrac{\ln M}{N_b} = \dfrac{\ln 2^{bL}}{(L+K-1)n} = \dfrac{Lb\ln 2}{(L+K-1)n} = \dfrac{L}{L+K-1}R$

(b)
$$P_E \le L P_e < L \,\frac{(2^b-1)}{1-2^{-b[E_0(\rho)/R-\rho]}}\, 2^{-bKE_0(\rho)} \quad ; 0<\rho<1$$

but $\ln 2\,\dfrac{bK}{R} = nK \implies 2^{-bKE_0(\rho)/R} = e^{-nKE_0(\rho)} = e^{-\frac{N_bK}{L+K-1}E_0(\rho)}$

so
$$P_E < \frac{L(2^b-1)}{1-2^{-b[E_0(\rho)/R-\rho]}}\, e^{-\frac{N_bKE_0(\rho)}{L+K-1}} \qquad \text{for } 0<\rho<1 .$$

(c) Let $\mathcal{K} = \dfrac{L(2^b-1)}{1-2^{-b[E_0(\rho)/R-\rho]}}$ and $\theta = \dfrac{L}{L+K-1}$ $(\Rightarrow R_b = \theta R)$

so $\dfrac{K-1}{L+K-1} = 1-\theta$ and therefore $\left(\dfrac{K-1}{L+K-1} < \dfrac{K}{L+K-1}\right)$

$$P_E < \mathcal{K}\exp\left[-N_b(1-\theta)E_c(R)\right] \quad \text{where}$$

$$\left.\begin{aligned} E_c(R) &= E_0(\rho) \\ R_b/\theta = R &= \frac{E_0(\rho)(1-\epsilon)}{\rho} \end{aligned}\right. \qquad \text{for } 0\le\rho\le 1$$

(d) Minimizing with respect to θ above we get

$$P_E < \mathcal{K}\exp\left[-N_b \max_{\theta;\, R\theta = R_b}(1-\theta)E_c(R)\right] \quad ; (\theta R = R_b)$$

(e) from (c) $R_b/\theta = \dfrac{E_0(\rho)(1-\epsilon)}{\rho} \implies \dfrac{R_b\rho}{1-\epsilon} = \theta E_0(\rho)$

or $(1-\theta)E_0(\rho) = (1-\theta)E_c(R) = E_0(\rho) - \dfrac{R_b}{1-\epsilon}\rho$;

where maximizing $E_b(R_b)$ over θ is the same as maximizing over ρ, so $E_b(R_b) = \max\limits_{0\le\rho\le1}\left[E_0(\rho) - R_b\rho/(1-\epsilon)\right]$.

5.4 Let $L = k_{crit} = \dfrac{K\rho E_o'(\rho)}{E_o(\rho) - \rho E_o'(\rho)}$; $0 < \rho\ 1$

$$\therefore\ R_b = \frac{k_{crit}}{k_{crit} + K - 1} R \approx \frac{k_{crit}}{k_{crit} + K} R = \frac{K\rho E_o'(\rho)}{K E_o(\rho)} \frac{E_o(\rho)}{\rho} = E_o'(\rho)$$

$$E_b(R_b) = \frac{(K-1)\, E_o(\rho)}{k_{crit} + K - 1} \approx \frac{K\, E_o(\rho)}{k_{crit} + K} = \frac{K[E_o(\rho) - \rho E_o'(\rho)]\, E_o(\rho)}{K\, E_o(\rho)}$$

$$= E_o(\rho) - \rho E_o'(\rho)$$

5.5 (a) Note that the terminated code is a special case of block codes. Thus

$$P_E > \exp\{-N_b[E_{sp}(R_b) + o(N_b)]\} \quad \text{where} \begin{cases} N_b = (L+K-1)n \\ R_b = \dfrac{L}{L+K-1} R \end{cases}$$

(b) $$P_E > \exp\{-n(L+K-1)[E_{sp}(R_b) + o(K)]\}$$

$$= 2^{-\frac{bK}{R}\frac{L+K}{K}[E_{sp}(R_b) + o(K)]} > 2^{-\frac{bK}{R(1-\theta)}[E_{sp}(R_b) + o(K)]}$$

$$\text{where} \begin{cases} E_{sp}(R_b) = E_o(\rho) - \rho E_o'(\rho) & 0 \le \rho < \infty \\ R_b = E_o'(\rho) = \theta R & 0 \le \theta \le 1 \end{cases}$$

$$\text{since} \quad \frac{L+K}{K} < \frac{L+K-1}{K-1} = \frac{1}{1-\theta}$$

(c) $P_E > 2^{-bK[E_{csp}(R) + o(K)]/R}$ which is true for all θ

and where $E_{csp}(R) = \min\limits_{\theta:\ R_b = \theta R} [E_{sp}(R)/(1-\theta)] = \min\limits_{0<\rho<\infty} \dfrac{E_o(\rho) - \rho E_o'(\rho)}{1 - E_o'(\rho)/R}$

The minimum occurs at

$$\theta = \left(1 - \frac{E_o'(\rho)}{R}\right)(-\rho E_o''(\rho)) - [E_o(\rho) - \rho E_o'(\rho)]\left(-\frac{E_o''(\rho)}{R}\right)$$

or $-\rho + \frac{\rho E_0'(\rho)}{R} + \frac{E_0(\rho)}{R} - \frac{\rho E_0'(\rho)}{R} = 0 \Rightarrow \frac{E_0(\rho)}{\rho} = R$

$$\Rightarrow \frac{E_0(\rho) - \rho E_0'(\rho)}{1 - E_0'(\rho)/R} = E_0(\rho)$$

5.6 (a) d_{free} is the minimum distance between closest paths thus, for the terminated convolutional coded treated as a block code, we have

$$d_{min}(\text{block}) = d_{free}(\text{convolutional})$$

(b) $$\frac{d_{free}}{(K-1)n} = \frac{d_{min}}{\frac{(K-1)(L+K-1)n}{L+K-1}} = \frac{d_{min}}{(1-\theta)N_b} \leq \frac{D(R_b)}{N_b(1-\theta)}$$

where $R_b = \theta R$. Minimizing with respect to L and θ

$$\frac{d_{free}}{(K-1)n} \leq \min_{0<\theta<1} \frac{D(\theta R)}{N_b(1-\theta)}$$

(c) using the Plotkin bound $\frac{D}{N_b} \sim \frac{1}{2}\left(1 - \frac{R_b}{\ln 2}\right)$ we will have

$$\frac{d_{free}}{(K-1)n} \leq \min_{0<\theta<1} \frac{1}{2}\left(\frac{1-\theta R/\ln 2}{1-\theta}\right) = \frac{1}{2} \quad ; R < \ln 2.$$

5.7 (a) From sec. 5.4 we have $P_b \geq \min_j P_b(j) \geq \min_j \pi_k(j)$. But $\pi_k(j)$ can be lower bounded by the average probability of pairwise errors for one incorrect path unmerged for the minimum length which is just K branches. Using the same method in finding (3.7.17) we will have

$$P_b \geq e^{\left[d_{free} \ln\left\{\sum_y \sqrt{p_0(y)p_1(y)} + o(K^{-1/2})\right\}\right]}$$

noting that here $d_{free} = d_{min}$ for the terminated code.

By assumption $d_{free} \le \delta(R)$ and so $P_b \geq Z^{\delta(R)}$.

(b) If $\frac{\delta(R)}{Kn} \le \frac{-R}{\ln(2e^{-R}-1)}$, then

$$P_b \geq Z^{\left[\frac{-RKn}{\ln(2e^{-R}-1)}\right]} = 2^{-\left[\frac{\ln Z}{\ln 2}\frac{RKn}{\ln(2e^{-R}-1)}\right]}$$

but $\frac{R \ln Z}{\ln(2e^{-R}-1)} = E_{cex}(R)$ and $R = \frac{b}{n}\ln 2$, therefore

$$P_b \geq 2^{-\frac{Kb}{R}E_{cex}(R)}$$

(c) from theorem 5.3.1 we have $P_b \le \gamma\, 2^{-\frac{Kb}{R}E_{cex}(R)}$ and our lower bound in (b) contradicts this result. We must then have at least one convolutional code such that

$$\frac{\delta(R)}{Kn} \geq \frac{-R}{\ln(2e^{-R}-1)}$$

(d) Since the terminated convolutional code is a special case of block code, we have for input-binary channels

$$P_b < LP_E < L\exp\left[-N_b E_{ex}(R)\right]$$

where $N_b = (L+K-1)n$ and $R = \frac{L}{L+K-1}\frac{b}{n}\ln 2$

$$E_{ex}(R) = -\delta_b \ln Z$$

by (3.4.8) $R = \ln 2 - \mathcal{H}(\delta_b) \qquad 0 \le R \le \ln 2 - \mathcal{H}\left(\frac{Z}{1+Z}\right)$

where $\delta_b = \frac{Z^{1/\rho}}{1+Z^{1/\rho}}$ and therefore we have

$$e^{d_{free}\ln Z} < P_b < L e^{N_b \delta_b \ln Z}$$

from part (a) in problem 5.6 $d'_{min}(\text{block}) = d_{free}(\text{convolutional})$

thus $d_{min} = \frac{d'_{min}}{N_b} > \delta_b$ where $R = \ln 2 - \mathcal{H}(\delta_b)$, which is the Gilbert bound.

5.8 By the argument used in sec. 5.1, using the Gallager bound we have

$$\overline{P_e(j)} \le \sum_{k=0}^{L-1} \overline{\Pi_k(j)} \le L \max_k \overline{\Pi_k(j)}$$

but $\overline{\Pi_k(j)} < \left[(2^b-1)2^{bk}\right]^\rho e^{-(K+k)n E_0(\rho)}$

$$< 2^{b(k+1)\rho}\, 2^{-(K+k)n E_0(\rho)/\ln 2}$$

and $R = b\ln 2/n$ so

$$\overline{\Pi_k(j)} < 2^{(k+1)b\rho - (K+k)bE_0(\rho)/R}$$

$$= 2^{-\frac{Kb}{R}\left[\frac{K+k}{K}E_0(\rho) - \frac{k+1}{K}\rho R\right]}$$

$$= 2^{-\frac{Kb}{R}(1+\lambda)\left[E_0(\rho) - \rho\tilde{R}\right]} \qquad ;\ 0 \le \rho \le 1 .$$

where $\lambda = k/K$ and $\tilde{R} = \dfrac{\lambda + 1/K}{1+\lambda} R$

Since the exponent is identical to that of the block coding bound in (3.7.17), minimizing with respect to ρ we obtain

$$\overline{\Pi_k(j)} < 2^{-\frac{Kb}{R}(1+\lambda)E(R,\lambda)}$$

where $E(R,\lambda) = E_0(\rho) - \rho E_0'(\rho)$, $0 < \rho \le 1$
and $\tilde{R} = E_0'(\rho)$ $\qquad E_0'(1) \le \tilde{R} < C$
and where $E(R,\lambda) = E_0(1) - \tilde{R}$ for $\tilde{R} \le E_0'(1)$.

Now
$$\max_k \overline{\Pi_k(j)} = 2^{-\frac{Kb}{R}\min_\lambda (1+\lambda)E(R,\lambda)}$$

and for large K we have

$$\frac{d}{d\lambda}(1+\lambda)E(R,\lambda) = E_0(\rho) - \rho E_0'(\rho) + (1+\lambda)\left[-\rho E_0''(\rho)\right]\frac{d\rho}{d\lambda} = 0$$

but from $\tilde{R} = E_0'(\rho) = \left[(\lambda + 1/K)/(1+\lambda)\right] R$, we have

$$\frac{1-1/K}{(1+\lambda)^2} R = E_0''(\rho) \frac{d\rho}{d\lambda}$$ and with this we get

$$E_0(\rho) - \rho E_0'(\rho) - \rho R \, \frac{1-1/K}{1+\lambda} = 0$$

which gives us

$$\lambda = \frac{\rho R(1-1/K)}{E_0(\rho) - \rho E_0'(\rho)} - 1$$

and then

$$R = \frac{(1+\lambda)\, E_0'(\rho)}{\lambda + 1/K} = \frac{\dfrac{\rho R(1-1/K)}{E_0(\rho)-\rho E_0'(\rho)}}{\dfrac{\rho R(1-1/K)}{E_0(\rho)-\rho E_0'(\rho)} - (1-1/K)} E_0'(\rho)$$

finally we get

$$R = \frac{E_0(\rho)}{\rho} \quad \text{when} \quad \frac{1+\lambda}{\lambda+1/K} E_0'(1) < R < C \quad \text{or} \quad E_0(1) < R < C$$

To check that this value of λ gives a minimum we have

$$\frac{d^2\left[(1+\lambda)\, E(R,\lambda)\right]}{d\lambda^2} = E_0'(\rho)\frac{d\rho}{d\lambda} - E_0'(\rho)\frac{d\rho}{d\lambda} - \rho E_0''(\rho)\frac{d\rho}{d\lambda} - \frac{d\rho}{d\lambda}\left[\frac{(1-1/K)R}{1+\lambda}\right]$$

$$+ \frac{\rho(1-1/K)R}{(1+\lambda)^2} = - \frac{R^2(1-1/K)^2}{(1+\lambda^3)\, E_0''(\rho)} \geqslant 0$$

And so $\min\limits_{\lambda} (1+\lambda) E(R,\lambda) = \rho R (1-1/K) = E_0(\rho)(1-1/K)$

So $$\overline{P_e(j)} < L\, 2^{-\frac{Kb}{R}\left[E_0(\rho) + O(K^{-1})\right]}$$

where $\rho = \dfrac{E_0(\rho)}{R}$ for $E_0(1) < R < C$

and $\rho = 1$ for $R \leq E_0(1)$.

5.9 We want to show that $\lim_{K\to\infty} \dfrac{\overline{Pr\{k \le k_{crit} - \epsilon K\}}}{\overline{P_e}} = 0$.

Using the same argument as in the proof of theorem 5.5.1 we have

$$\frac{\overline{Pr\{k \le k_{crit} - \epsilon K\}}}{\overline{P_e}} \le \frac{\sum_{k=0}^{k_{crit}-\epsilon K} \overline{\Pi_k(j)}}{2^{-bK[E_0(\rho^*)+o(K^{-1})]/R}}$$

$$\le \max_{0 \le k \le k_{crit}-\epsilon K} \overline{\Pi_k(j)}$$

We note that k_{crit} maximizes $\overline{\Pi_k(j)}$ as seen previously in the text and in problem 5.8. For small ϵ, $k_{crit} - \epsilon K$ will maximize $\overline{\Pi_k(j)}$ in the given range of k. But, for this value of $k = k_{crit} - \epsilon K$ the exponent of $\max \overline{\Pi_k(j)}$ which was found in problem 5.8 will be a little larger. We can say that there exists an $\alpha > 0$ such that

$$\max_{0 \le k \le k_{crit}-\epsilon K} \overline{\Pi_k(j)} = \overline{\Pi_{k_{crit}-\epsilon K}(j)} \le 2^{-\frac{Kb}{R}[E_0(\rho)+\alpha+o(K^{-1})]}$$

Then $$\frac{\overline{Pr\{k \le k_{crit} - \epsilon K\}}}{\overline{P_e}} \le \frac{(k_{crit}+1)\, 2^{-\frac{Kb}{R}[E_0(\rho)+\alpha+o(K^{-1})]}}{2^{-\frac{Kb}{R}[E_0(\rho^*)+o(K^{-1})]}}$$

$$= (k_{crit}+1)\, 2^{-\frac{Kb}{R}\left[E_0(\rho)-E_0(\rho^*)+\alpha+o(K^{-1})\right]}$$

And, as $\rho \to \rho^*$ we have

$$\lim_{K\to\infty} \frac{\overline{Pr\{k \le k_{crit} - \epsilon K\}}}{\overline{P_e}} \le (k_{crit}+1)\, 2^{-\frac{Kb}{R}\alpha} \xrightarrow[K\to\infty]{} 0$$

5.10 (a) $\xi_b(D,I;j,i+1) = D^{Z_5^{(i+1)}} I\, \xi_c(D,I;j,i)$

$\xi_c(D,I;j,i+1) = D^{Z_6^{(i+1)}} \xi_b(D,I;j,i) + D^{Z_4^{(i+1)}} \xi_d(D,I;j,i)$

$\xi_d(D,I;j,i+1) = D^{Z_2^{(i+1)}} \xi_b(D,I;j,i) + D^{Z_3^{(i+1)}} \xi_d(D,I;j,i)$

$\therefore$

$$A(i+1) = \begin{bmatrix} 0 & D^{Z_5^{(i+1)}} I & 0 \\ D^{Z_6^{(i+1)}} & 0 & D^{Z_4^{(i+1)}} \\ D^{Z_2^{(i+1)}} & 0 & D^{Z_3^{(i+1)}} \end{bmatrix}$$

(b) without loss of generality we can assume tht the all-zero data sequence is transmitted. A node error occurs at node j when some path leaving the zero state at time j and remerging later causes an error. The union bound for all such error paths is then

$$P_e(j) \leq \sum_{i=j}^{\infty} D^{Z_7^{(i+1)}} \xi_c(D,I;j,i) \Big|_{\substack{I=1 \\ D=Z}}$$

(c) note that we can express the bound in (b) as

$$P_e(j) \leq \sum_{i=j}^{\infty} [0 \;\; D^{Z_7^{(i+1)}} \;\; 0]\, \xi(D,I;j,i) \Big|_{\substack{I=1 \\ D=Z}}$$

Consider $\displaystyle\sum_{i=j+1}^{\infty} [0 \;\; D^{Z_7^{(i+1)}} \;\; 0]\, \xi(D,I;j,i)$

$$= \sum_{i=j+1}^{\infty} [0 \;\; D^{Z_7^{(i+1)}} \;\; 0]\, A(i)A(i-1)\dots A(j+1)\, \xi(D,I;j,j)$$

$$= \sum_{i=j+1}^{\infty} [0 \;\; D^{Z_7^{(i+1)}} \;\; 0]\, A(i)A(i-1)\cdots A(j+1) \begin{bmatrix} D^{Z_1^{(j)}} \\ 0 \\ 0 \end{bmatrix}$$

Over the time-varying code ensemble, the $\{Z_k^{(i)}\}$ are independent random variables for all $k=1,\ldots,7$ and all i. Hence

$$\overline{\sum_{i=j+1}^{\infty}[0\ D^{z_7^{(i+1)}}\ 0]\,\xi_7(D,I;j,i)} = \sum_{i=j+1}^{\infty}\overline{[0\ D^{z_7^{(i+1)}}\ 0]}\,\bar{A}\cdots\bar{A}\ \overline{\begin{bmatrix} D^{z_1^{(j)}} \\ 0 \\ 0 \end{bmatrix}}$$

But $\bar{D} \equiv \overline{D^{z}} = \frac{1}{4}D^0 + \frac{1}{2}D + \frac{1}{4}D^2 = \frac{1}{4}(1+2D+D^2)$

$$= \left(\frac{1+D}{2}\right)^2$$

and

$$\bar{A} = \begin{bmatrix} 0 & \bar{D} & 0 \\ \bar{D} & 0 & \bar{D} \\ \bar{D}I & 0 & \bar{D}I \end{bmatrix}$$

which corresponds to the transmitted matrix where all branches are labeled with $\bar{D}$ and all "one" data branches are labeled also with I. The transfer function for this "averaged" state diagram is obtained from (4.6.1) as

$$T_3(L,I)\Big|_{L=\bar{D}} = \frac{\bar{D}^3 I}{1-\bar{D}(1+\bar{D})I}$$

Note that $\overline{\sum_{i=j+1}^{\infty}[0\ D^{z_7^{(i+1)}}\ 0]\,\xi_7(D,I;j,i)} = \sum_{i=j}^{\infty}[0\ \bar{D}\ 0]\,\bar{A}^{i-j}\begin{bmatrix} \bar{D} \\ 0 \\ 0 \end{bmatrix}$

$$= [0\ \bar{D}\ 0]\sum_{i=1}^{\infty}\bar{A}^{i}\begin{bmatrix} \bar{D} \\ 0 \\ 0 \end{bmatrix} = [0\ \bar{D}\ 0]\sum_{i=0}^{\infty}\bar{A}^{i}\begin{bmatrix} \bar{D} \\ 0 \\ 0 \end{bmatrix}$$

$$= [0\ \bar{D}\ 0][I-\bar{A}]^{-1}\begin{bmatrix} \bar{D} \\ 0 \\ 0 \end{bmatrix} = T_3(L,I)\Big|_{L=\bar{D}}$$

hence

$$P_b(j) \le T_3(D,I)\Big|_{\substack{I=1 \\ D=\bar{D}}}\ ;\ \bar{P}_b \le \frac{\partial T_3(D,I)}{\partial I}\Big|_{\substack{I=1 \\ D=\bar{D}}}$$

(d) this follows in exactly the same way.

<u>5.11</u> For $\mathcal{L}$ 3 we have the following matrix A

$$A = \begin{bmatrix}
f(0,\Delta_0) & & & & & & & f(0,\Delta_7) & f(0,\Delta_8) \\
f(-1,\Delta_0) & & & & & & & f(-1,\Delta_7) & f(-1,\Delta_8) \\
f(1,\Delta_0) & & & & & & & f(1,\Delta_7) & f(1,\Delta_8) \\
 & & f(1,\Delta_2) & f(1,\Delta_3) & & f(1,\Delta_5) & & & \\
 & & f(-1,\Delta_2) & f(-1,\Delta_3) & & f(-1,\Delta_5) & & & \\
 & f(1,\Delta_1) & & & f(1,\Delta_4) & & f(1,\Delta_6) & & \\
 & f(-1,\Delta_1) & & & f(-1,\Delta_4) & & f(-1,\Delta_6) & & \\
 & & f(0,\Delta_2) & f(0,\Delta_3) & & f(0,\Delta_5) & & & \\
 & f(0,\Delta_1) & & & f(0,\Delta_4) & & f(0,\Delta_6) & &
\end{bmatrix}$$

where

$$\begin{aligned}
\Delta_0 &= (0,0) & \Delta_5 &= (-1,1) \\
\Delta_1 &= (0,-1) & \Delta_6 &= (-1,-1) \\
\Delta_2 &= (0,1) & \Delta_7 &= (1,0) \\
\Delta_3 &= (1,1) & \Delta_8 &= (-1,0) \\
\Delta_4 &= (1,-1) & &
\end{aligned}$$

For a, b, c, d, e, f, g, h and j defined in figure 4.23, we get the following for A

$$\begin{bmatrix}
1/2 & & & & & & & 1/2 & 1/2 \\
1/4\,a & & & & & & & 1/4\,j & 1/4\,h \\
1/4\,a & & & & & & & 1/4\,h & 1/4\,j \\
 & & 1/4\,b & 1/4\,d & & 1/4\,g & & & \\
 & & 1/4\,c & 1/4\,c & & 1/4\,f & & & \\
 & 1/4\,c & & & 1/4\,f & & 1/4\,e & & \\
 & 1/4\,b & & & 1/4\,g & & 1/4\,d & & \\
 & & 1/2 & 1/2 & & & & & \\
 & 1/2 & & & & & 1/2 & &
\end{bmatrix}$$

We have:

all paths ending in $(1,1)$ and $(-1;1)$ are the same.

all paths ending in $(1,-1)$ and $(-1,1)$ are the same.

all paths ending in $(0,1)$ and $(0,-1)$ are the same.

all paths ending in $(1,0)$ and $(-1,0)$ are the same

Hence A reduces to $\tilde{A}$ below

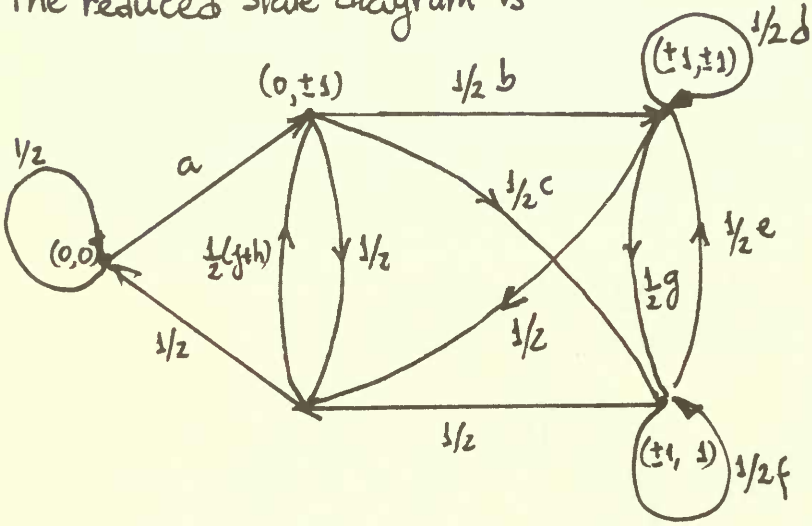
$$\tilde{A} = \begin{bmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ a & 0 & 0 & 0 & 1/2(j+h) \\ 0 & 1/2\,b & 1/2\,d & 1/2\,g & 0 \\ 0 & 1/2\,c & 1/2\,e & 1/2\,f & 0 \\ 0 & 1/2 & 1/2 & 1/2 & 0 \end{bmatrix}$$

where $\Delta_0 = (0,0)$ $\quad\Delta_3 = (\pm 1, \mp 1)$
$\Delta_1 = (0, \pm 1)$ $\quad\Delta_4 = (\pm 1, 0)$
$\Delta_2 = (\pm 1, \pm 1)$

and now

$$\overline{P_E(\underline{\varepsilon})} \leq \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tilde{A}^N \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The reduced state diagram is



5.12.

Let $\underline{v}$ be the row eigenvector associated with λ.

$$\underline{v}\,\underline{A} = \lambda \underline{v} \quad , \quad \underline{v}\,\underline{A}^N = \lambda^N \underline{v}$$

$$\alpha = \frac{V_{max}}{V_{min}} > 1$$

$$[1\,1\cdots 1] \le \frac{1}{V_{min}} \cdot \underline{v}$$

$$[1\,1\cdots 1]\,\underline{A}^N \le \frac{1}{V_{min}}\,\underline{v}\,\underline{A}^N = \frac{1}{V_{min}}\lambda^N \underline{v}$$

$$[1\,1\cdots 1]\,\underline{A}^N \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \le \frac{1}{V_{min}}\lambda^N \underline{v} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{1}{V_{min}}\lambda^N \cdot v_j \le \frac{V_{max}}{V_{min}}\lambda^N.$$

$$[1\,1\cdots 1] \ge \frac{1}{V_{max}} \cdot \underline{v}$$

$$[1\,1\cdots 1]\,A^N \ge \frac{1}{V_{max}}\,\underline{v}\,\underline{A}^N = \frac{1}{V_{max}}\lambda^N \underline{v}$$

$$[1\,1\cdots 1]\,A^N \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ge \frac{1}{V_{max}}\lambda^N \underline{v} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{1}{V_{max}}\lambda^N \cdot v_j \ge \frac{V_{min}}{V_{max}} \cdot \lambda^N$$

$\therefore$ We have

$$\frac{1}{\alpha}\lambda^N \le [1\,1\cdots 1]\,\underline{A}^N \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \le \alpha \lambda^N$$

where $\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ means any column vector with one "1" and "0" elsewhere.

5.13.

When $l=1$, $\quad (k_1+1) = 1\cdot(k_1+1)$

When $l=2$, the inequality holds:

$$(k_1+1)+(k_2+1) \le 2(k_1+1)(k_2+1)$$

$$\Rightarrow \quad k_1+k_2+2 \le 2k_1k_2+2k_1+2k_2+2$$

$$\Rightarrow \quad 0 \le 2k_1k_2+k_1+k_2$$

the equality holds iff $k_1=k_2=0$.

Suppose the inequality hold for $l=n\ge 2$.

Then for $l=n+1$;

$$\sum_{i=1}^{n+1}(k_i+1) = k_{n+1}+1+\sum_{i=1}^{n}(k_i+1)$$

$$(n+1)\prod_{i=1}^{n+1}(k_i+1) = (n+1)(k_{n+1}+1)\prod_{i=1}^{n}(k_i+1)$$

$$= (n\,k_{n+1}+k_{n+1}+1)\prod_{i=1}^{n}(k_i+1) + n\prod_{i=1}^{n}(k_i+1)$$

$$\ge (n\,k_{n+1}+k_{n+1}+1)\frac{1}{n}\sum_{i=1}^{n}(k_i+1) + \sum_{i=1}^{n}(k_i+1)$$

$$\ge n\,k_{n+1}+k_{n+1}+1 + \sum_{i=1}^{n}(k_i+1)$$

$$\ge k_{n+1}+1+\sum_{i=1}^{n}(k_i+1) = \sum_{i=1}^{n+1}(k_i+1)$$

Since $\sum_{i=1}^{n}(k_i+1)\ge n$ with equality when all $k_i=0$.

$$\therefore \quad \sum_{i=1}^{l}(k_i+1) \le l\prod_{i=1}^{l}(k_i+1) \qquad \text{for all } l\ge 2$$

with equality when all $k_i=0$, $i=1,2,\cdots,l$.

5.14 $$P_N(\underset{\sim}{y}|\underset{\sim}{x},\underset{\sim}{s}_1)=\prod_{n=1}^{N} p(y_n/x_n,s_{1n})$$

$$\underset{\sim}{s}_n=(x_{n-(L_0-1)},\ldots,x_{n-2},x_{n-1})$$

(a) Assume $x=(x_1,\ldots,x_N)$

$x'=(x_1',\ldots,x_N')$

$$P_E(\underset{\sim}{x}\to\underset{\sim}{x}')=P_E(x,x'|\underset{\sim}{s}_1,\underset{\sim}{s}_1')$$

$$=P_r\left\{\ln\frac{p_N(\underset{\sim}{y}|x',\underset{\sim}{s}_1')}{P_N(\underset{\sim}{y}|x,\underset{\sim}{s})}\geq 0\ \middle|\ x,\underset{\sim}{s}_1\right\}$$

$$=\sum_{\Lambda_{xx'}} p(\underset{\sim}{y}/x,\underset{\sim}{s}_1)$$

where

$$\Lambda_{xx'}=\left\{\underset{\sim}{y}:\ \ln\frac{p_N(\underset{\sim}{y}/x',\underset{\sim}{s}_1')}{p_N(\underset{\sim}{y}/x,\underset{\sim}{s}_1)}\geq 0\right\}$$

Define an indicator function $f(\underset{\sim}{y})$ as

$$f(\underset{\sim}{y})=\begin{cases}1, & \underset{\sim}{y}\in\Lambda_{xx'}\\ 0, & \underset{\sim}{y}\notin\Lambda_{xx'}\end{cases}$$

Then $P_E(x,x'|\underset{\sim}{s}_1,\underset{\sim}{s}_1')=\sum_{\underset{\sim}{y}} f(\underset{\sim}{y})\,p(\underset{\sim}{y}/x,\underset{\sim}{s}_1)$

$f(\underset{\sim}{y})$ can be bounded as

$$f(\underset{\sim}{y})\leq\sqrt{\frac{p_N(\underset{\sim}{y}|x',\underset{\sim}{s}_1')}{P_N(\underset{\sim}{y}|x,\underset{\sim}{s})}}\qquad\text{for any } \underset{\sim}{y}$$

$$\therefore\ P_E(\underset{\sim}{x},\underset{\sim}{x}'|\underset{\sim}{s}_1,\underset{\sim}{s}_1')\leq\sum_{\underset{\sim}{y}}\sqrt{\frac{P_N(\underset{\sim}{y}/x',\underset{\sim}{s}_1')}{p_N(\underset{\sim}{y}/x,\underset{\sim}{s}_1)}}\,P_N(\underset{\sim}{y}/x,\underset{\sim}{s}_1)$$

so

$$P_E(\underline{x},\underline{x}'|\underline{s}_1,\underline{s}_1') \leq \sum_{\underline{y}} \sqrt{P_N(\underline{y}|\underline{x}',\underline{s}_1')\,P_N(\underline{y}|\underline{x},\underline{s}_1)}$$

But $P_N(\underline{y}|\underline{x},\underline{s}_1) = \prod_{n=1}^{N} p(y_n|x_n,s_n)$, hence

$$P_E(x,x'|\underline{s}_1,\underline{s}_1') \leq \sum_{y} \prod_{n=1}^{N} \sqrt{p(y_n|x_n,s_n)\,p(y_n|x_n',s_n')}$$

$$= \prod_{n=1}^{N} \sum_{y} \sqrt{p(y_n|x_n,s_n)\,p(y_n|x_n',s_n')}$$

(b) Let x and x' have independent components selected according to $q(x)$, then

$$\overline{P_E(x,x'|\underline{s}_1,\underline{s}_1')} = \mathop{E}_{x,x'}\left\{P_E(x,x'|\underline{s}_1,\underline{s}_1')\right\}$$

$$= \sum_{x,x'} q(x)\,q(x')\,P_E(x,x'|\underline{s}_1,\underline{s}_1')$$

$$\leq \sum_{x_1}\sum_{x_1'}\sum_{x_2}\sum_{x_2'}\cdots\sum_{x_N}\sum_{x_N'} \prod_{n=1}^{N} q(x_n)\,q(x_n')$$

$$\times \sum_{\underline{y}} \sqrt{p(y_n|x_n,s_n)\,p(y_n|x_n',s_n')}$$

(c) We have $\underline{x} = (x,x')$, $q(\underline{x}) = q(x)\,q(x')$ and $s_{k+1} = g(\underline{x}_k,\underline{s}_k))$ and let

$$\underline{s}_n = (s_n,s_n') \in \{\Delta_1,\Delta_2,\ldots,\Delta_{K^{2(L-1)}}\} = \Lambda$$

then

$$\overline{P_E(\underline{x},\underline{x}',\underline{s})} \leq \sum_{x_1}\cdots\sum_{x_N} \prod_{n=1}^{N} q(x_n) \sum_{y} \sqrt{p(y|x_n,s_n)\,p(y|x_n',s_n')}$$

Let

$$a_{ij} = \begin{cases} f(\underset{\sim}{x},\ j) & \text{if } \Delta i = g(\underline{x}, s_i) \\ 0 & \text{otherwise} \end{cases}$$

where

$$A = \{a_{ij}\}$$

$$f(\underset{\sim}{x}, \underline{s}) = q(\underline{x})\left(\sum_{y} \sqrt{p(y|x,s)\,p(y|x',s')}\right)$$

Then

$$\overline{P_E(\underset{\sim}{x}, \underset{\sim}{x}'|\underline{s}_1)} = [1\ 1\ 1 \ldots 1]\, A^N\, \underset{\sim}{i}(\underline{s}_1)$$

where $\underset{\sim}{i}(\underline{s})$ is the $K^{2(b-1)}$-dimensional column vector with "1" in the position corresponding to state $\underset{\sim}{s}_1$ and "0" otherwise.

(d) From (5.8.15), the average bit error probability bound due to single code merger error events for the general finite memory channel is

$$\overline{P}_{b_1} \leq \frac{2^b-1}{b} \sum_{k=0}^{\infty} b(K+1)\, 2^{bK}\, \overline{P}_k$$

where $\overline{P}_k = \overline{P_{E_1}(\underline{x}, \underline{x}'|\underline{s}_1)}$, with $N = n(K+k)$ such that

$$\overline{P}_{b_1} \leq \frac{2^b-1}{2} \sum_{k=0}^{\infty} b(K+1)\, 2^{bK} [1\ 1 \ldots 1]\, A^{n(K+k)}\, \underset{\sim}{i}(\underline{s}_1)$$

by the same arguments as in section 5.8

$$\frac{1}{\alpha}\lambda^N \leq [1\ 1 \ldots 1]\, A^N\, \underset{\sim}{i}(\underset{\sim}{s}_1) \leq \alpha\, \lambda^N$$

where λ is a real maximum eigenvalue of A and $\alpha > 0$ is the largest component of the left eigenvector associated with λ divided by its smallest component.

Hence

$$\overline{P}_{b_1} \leq \frac{\alpha(2^b-1)}{(1-2^b\lambda^n)^2}\,\lambda^{nK} \quad , \; 2^b\lambda^n < 1 \; ;$$

all other steps follow in a similar manner with the result of theorem 5.8.1 which applies for this general finite memory case.

—//—

Chapter 6

6.1 In equation (6.2.11) let

$$a(y) = \sum_x q(x)\,p(y|x)^{1-\alpha\rho} \quad ; \quad b(y) = w(y)^{\alpha\rho}$$

then, with $\theta = 1/(1-\alpha\rho)$ we have:

$$\exp\left[-E_c(\alpha,\rho)\right] = \sum_y \sum_x q(x)\,p(y|x)\left[\frac{p(y|x)}{w(y)}\,e^{-R}\right]^{-\alpha\rho}$$

$$\leq e^{\alpha\rho R}\left[\sum_y\left(\sum_x q(x)\,p(y|x)^{1-\alpha\rho}\right)^{\frac{1}{1-\alpha\rho}}\right]^{1-\alpha\rho}\left(\sum_y w(y)\right)^{\alpha\rho}$$

$$= \exp\left\{\alpha\rho R + (1-\alpha\rho)\ln\left[\sum_y\left(\sum_x q(x)\,p(y|x)^{\frac{1}{1+\frac{\alpha\rho}{1-\alpha\rho}}}\right)^{1+\frac{\alpha\rho}{1-\alpha\rho}}\right]\right\}$$

$$= \exp\left\{\alpha\rho R - (1-\alpha\rho)\,E_o\!\left(\frac{\alpha\rho}{1-\alpha\rho}\right)\right\}$$, which is (6.2.15).

In (6.2.12) let $a(y) = \sum_x q(x)\,p(y|x)^{1-\alpha\rho}$; $\theta = 1-\alpha\rho$

and $b(y) = \left[\sum_{x'} q(x')\,p(y|x')^{\alpha}\right]^{\rho}$;

Then

$$\exp\left[-E_{cs}(\alpha,\rho)\right] = \sum_y\left(\sum_x q(x)\,p(y|x)^{1-\alpha\rho}\right)\left(\sum_{x'} q(x')\,p(y|x')^{\alpha}\right)^{\rho}$$

$$\leq \left[\sum_y\left(\sum_x q(x)\,p(y|x)^{1-\alpha\rho}\right)^{\frac{1}{1-\alpha\rho}}\right]^{1-\alpha\rho}\left[\sum_y\left(\sum_{x'} q(x')\,p(y|x')^{\alpha}\right)^{\frac{1}{\alpha}}\right]^{\alpha\rho}$$

$$= \exp\left\{(-1-\alpha\rho)\,E_o\!\left(\frac{\alpha\rho}{1-\alpha\rho}\right) - \alpha\rho\,E_o\!\left(\frac{1-\alpha}{\alpha}\right)\right\} \quad ;$$

Finally for (6.2.13) let (with $\theta = \alpha\rho$)

$$a(y) = \sum_x q(x)\, p(y|x)^\alpha \quad \text{and} \quad b(y) = w(y)^{1-\alpha\rho}.$$

Then $\exp[-E_I(\alpha,\rho)] = \sum_y \left(\sum_x q(x) p(y|x)^\alpha\right)^\rho w(y)^{1-\alpha\rho} e^{-\alpha\rho R}$

$$\leq e^{-\alpha\rho R}\left[\sum_y \left(\sum_x q(x) p(y|x)^\alpha\right)^{1/\alpha}\right]^{\alpha\rho}\left[\sum_y w(y)\right]^{1-\alpha\rho}$$

$$= \exp\left\{-\alpha\rho R - \alpha\rho\, E_0\left(\frac{1-\alpha}{\alpha}\right)\right\} \quad ; \quad [\div[6.2.17)]$$

6.2 (a) Using the same method as in sec. 6.1 we get

$E[m(x_u)] = C - \beta$, for the correct path.

$E[m(x'_u)] < -\beta$, for the incorrect paths.

(b) (6.2.11) becomes

$$\exp\{-E_c(\alpha,\rho)\} = \sum_y \sum_x q(x)\, p(y|x)\left[\frac{p(y|x)}{w(y)} e^{-\beta}\right]^{-\alpha\rho}$$

(6.2.12) is unchanged and (6.2.13) becomes

$$\exp\{-E_I(\alpha,\rho)\} = \sum_y w(y)\left\{\sum_{x'} q(x')\left[\frac{p(y|x')}{w(y)} e^{-\beta}\right]^\alpha\right\}^\rho$$

and (6.2.15)–(6.2.17) become

$$\exp\{-E_c(\alpha,\rho)\} \leq \exp\left\{\alpha\rho\beta - (1-\alpha\rho)\, E_0\left(\frac{\alpha\rho}{1-\alpha\rho}\right)\right\} \equiv \delta'_c$$

$$\exp\{-[E_I(\alpha,\rho) - \rho R]\} \leq \exp\left\{\rho(R-\alpha\beta) - \alpha\rho\, E_0\left(\frac{1-\alpha}{\alpha}\right)\right\} \equiv \delta'_I$$

$$\exp\{-[E_{cI}(\alpha,\beta) - \rho R]\} \leq \exp\left\{\rho R - (1-\alpha\rho)\, E_0\left(\frac{\alpha\rho}{1-\alpha\rho}\right) - \alpha\rho\, E_0\left(\frac{1-\alpha}{\alpha}\right)\right\}$$

$$\equiv \delta'_c\, \delta'_I .$$

(c) $T(t,\tau) \le \delta_C'^{\,nt}\,\delta_I'^{\,n\tau}$

For $\delta_C' < 1$, we need $\beta < \frac{1-\alpha\rho}{\alpha\rho} E_o\left(\frac{\alpha\rho}{1-\alpha\rho}\right) = \frac{E_o(\gamma)}{\gamma}$;

where $\gamma = \frac{\alpha\rho}{1-\alpha\rho}$.

For $\delta_I' < 1$, we need $R - \alpha\beta < \alpha E_o\left(\frac{1-\alpha}{\alpha}\right)$

or $R < \alpha\left[\beta + E_o\left(\frac{1-\alpha}{\alpha}\right)\right]$;

Choosing $\alpha = 1/(1+\rho)$, this results in

(1) $\beta < \frac{E_o(\rho)}{\rho}$ and $R < \frac{\beta + E_o(\rho)}{1+\rho}$, and thus

$P_r(C > L) < AL^{-\rho}$ provided (1) is satisfied.

(d) Similarly $\bar{P}_b < D\,2^{-bK\rho}$, $0 \le \rho \le 1$

provided (1) is satisfied.

<u>6.3</u> Assume $P_r(\tilde{C} > L) = K L^{-\rho}$; $R = E_o(\rho)/\rho$.

Then

(a) $E(\tilde{C}) = \sum_{L=0}^{\infty} L\left[P_r(\tilde{C} > L-1) - P_r(\tilde{C} > L)\right] = \sum_{L=0}^{\infty} P_r(\tilde{C} > L)$

$$= K \sum_{L=0}^{\infty} L^{-\rho} < \infty \text{ provided } \rho > 1;$$

This implies $R < E_o(1) = R_o$.

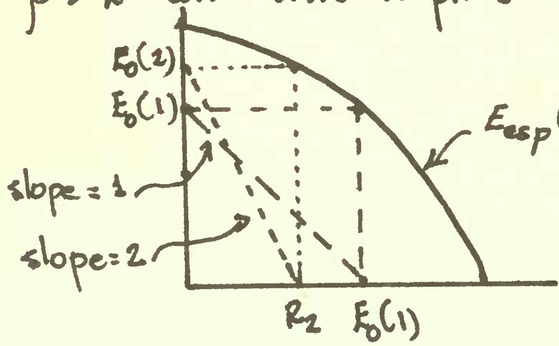
(b) $E(\tilde{C}^2) = \sum_{L=0}^{\infty} L^2\left\{P_r(\tilde{C} > L-1) - P_r(\tilde{C} > L)\right\} = \sum_{L=0}^{\infty} L\,P_r(\tilde{C} > L)$

$$= K \sum_{L=0}^{\infty} L^{1-\rho} < \infty \text{ if } \rho > 2;$$

This implies $R < E_0(\rho)/2 \equiv R_2$.

(c) $$E(\tilde{C}^k) = \sum_{L=0}^{\infty} L^{k-1} P_r(\tilde{C} > L) = K \sum_{L=0}^{\infty} L^{k-1-\rho} < \infty$$

if $\rho > k$ and this implies $R < E_0(k)/k$.



6.4 We know from problem 5.3 that the block error probability for terminated convolutional code of D branches and constraint length k is

$$P_E < D \frac{(2^b - 1)\, 2^{-bk E_c(R)/R}}{1 - 2^{-b\varepsilon E_c(R)/R}} \quad , 0 < \rho < 1$$

Since at each step we have to increase the number of computations per branch by 2^b, the number of computations required at the end of $k-1$ steps is

$$2^b + 2^{2b} + 2^{3b} + \dots + 2^{b(k-1)} = \frac{2^{bk} - 2^b}{2^b - 1} < 2^{bk}$$

then, denoting the total number of computations per branch by $\tilde{C}$ we will have

$$P_r\{\tilde{C} \geq 2^{bk}\} < D \frac{(2^b - 1)\, 2^{-bk E_c(R)/R}}{1 - 2^{-b\varepsilon E_c(R)/R}} \quad , 0 < \rho < 1 .$$

Now let $L = 2^{bk}$ and $R = E_0(\rho)/\rho$, so we have

$$P\{\tilde{C} > L\} < D' 2^{-bk\rho} = D' L^{-\rho} \quad , 0 < \rho < 1 \quad \text{and}$$

where

$$D' = D/(1 - 2^{-b\varepsilon\rho}) .$$

6.5 (a) $\frac{k_{crit}}{K} = \frac{\rho E_0'(\rho)}{E_0(\rho) - \rho E_0'(\rho)}$ which implies $\frac{k_{crit}}{K} + 1 = \frac{E_0(\rho)}{E_0(\rho) - \rho E_0'(\rho)}$

$\therefore \frac{k_{crit}}{K} + 1 = \frac{\rho R}{E_0(\rho) - \rho E_0'(\rho)}$

$$\frac{l_{crit}}{K} = \frac{\rho \ln L / nK}{E_0(\rho) - \rho E_0'(\rho)} = \frac{\rho R \left(\ln L / \ln 2^{bk}\right)}{E_0(\rho) - \rho E_0'(\rho)}$$

then $$\frac{l_{crit}}{K} = \left(\frac{\ln L}{\ln 2^{bk}}\right)\left[\frac{k_{crit}}{K} + 1\right]$$

where L takes the place of 2^{bk} in sequential decoding.

(b) We must add one constraint length to k_{crit} since k_{crit} represents the length of error events (first to last error) and we require $K-1$ branches to remerge.

6.6 $P(E_i)$ is the block coding error probability, where there are 2^{i-1} diverging paths of block length $n(i+T)$ over the ensemble. The correct sequence is independent of each of the diverging sequences. Hence

$$\overline{P(E_i)} \le (2^{i-1})^{\rho} \left\{ \sum_y \left(\sum_x q(x) p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right\}^{n(i+T)}$$

for any $\rho \in [0,1]$.
Now let
$$\sum_y \left(\sum_x q(x) p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} = e^{-E_0(\rho, q)}$$

Thus
$$\overline{P_E} \le \sum_{i=1}^{L} \overline{P(E_i)} \le \sum_{i=1}^{L} 2^{(i-1)\rho} e^{-n(i+T)E_0(\rho,q)}$$
$$= \sum_{i=1}^{L} 2^{(i-1)\rho} e^{-n(T+1)E_0(\rho,q)} e^{-n(i-1)E_0(\rho,q)}$$

$$\therefore \quad \overline{P_E} = e^{-n(T+1)E_o(\rho,q)} \sum_{i=1}^{L} \left[2^\rho e^{-nE_o(\rho,q)}\right]^{i-1}$$
$$\leq e^{-n(T+1)E_o(\rho,q)} \sum_{i=0}^{\infty} \left[2^\rho e^{-nE_o(\rho,q)}\right]^{i}$$

and thus

$$\overline{P_E} \leq \frac{e^{-n(T+1)E_o(\rho,q)}}{1-2^\rho e^{-nE_o(\rho,q)}}, \text{ where we require } 2^\rho e^{-nE_o(\rho,q)} < 1.$$

This requirement is equivalent to

$$R = \frac{\ln 2}{n} < \frac{E_o(\rho,q)}{\rho} = \frac{E_o(\rho,q) - E_o(0,q)}{\rho - 0}$$

As long as $R < C = \lim_{\rho \to 0} \max_q \frac{E_o(\rho)}{\rho}$ then there exists a $\rho \in (0,1)$ and q such that $R < E_o(\rho,q)/\rho$ and $E_o(\rho,q) > 0$. This means that for R in this range the bound decreases exponentially with T.

6.7 (a) We know that the minimum probability of error decision rule is to choose m that maximizes $p(m|\underline{y})$, but

$$p(m|\underline{y}) = \frac{p(m,\underline{y})}{p(\underline{y})} = \frac{\sum_{\underline{t}_m} p(m,\underline{y},\underline{t}_m)}{p(\underline{y})}$$

$$= \frac{\sum_{\underline{t}_m} p(\underline{y}|m,\underline{t}_m)\, p(\underline{t}_m|m)\, \pi_m}{\prod_{n=1}^{N} w(y_n)}$$

and $p(\underline{t}_m|m) = q_{N-N_m}(\underline{t}_m) = \prod_{k=1}^{N-N_m} q(t_k)$.

$\underline{t}_m$ is independent of $\underline{x}_m$ and we have a DMC so

$$p(\underline{y}|m,\underline{t}_m) = p(\underline{y}|\underline{z}_m) = \prod_{n=1}^{N_m} p(y_n|x_{mn}) \prod_{k=1}^{N-N_m} p(y_{N_m+k}|t_k)$$

Therefore we can write

$$p(m|\underline{y}) = \frac{\sum_{\underline{t}_m} \prod_{n=1}^{N_m} p(y_n|x_{mn}) \prod_{k'=1}^{N-N_m} p(y_{N_m+k'}|t_{k'}) \prod_{k=1}^{N-N_m} q(t_k)\, \pi_m}{\prod_{n=1}^{N} w(y_n)}$$

$$= \frac{\pi_m \prod_{n=1}^{N_m} p(y_n|x_{mn}) \prod_{k=1}^{N-N_m} \sum_{t_k} p(y_{N_m+k}, t_k)}{\prod_{n=1}^{N} w(y_n)}$$

$$= \frac{\pi_m \prod_{n=1}^{N_m} p(y_n|x_{mn}) \prod_{k=1}^{N-N_m} w(y_{N_m+k})}{\prod_{n=1}^{N} w(y_n)}$$

$$= \frac{\pi_m \prod_{n=1}^{N_m} p(y_n|x_{mn})}{\prod_{n=1}^{N_m} w(y_n)} = \pi_m \prod_{n=1}^{N_m} \frac{p(y_n|x_{mn})}{w(y_n)}$$

Maximizing $p(m|\underline{y})$ is equivalent to maximizing $\log p(m|\underline{y})$. Therefore

$$\log p(m|\underline{y}) = \log \pi_m + \sum_{n=1}^{N_m} \log \frac{p(y_n|x_{mn})}{w(y_n)}$$

$$= \sum_{n=1}^{N_m} \left[\log \frac{p(y_n|x_{mn})}{w(y_n)} + \frac{1}{N_m} \log \pi_m \right]$$

$$= L(m, \underline{y})$$

(b) Suppose the number of bits required to produce x_m is l_m. Then $\pi_m = 2^{-l_m} = 2^{-(l_m/N_m)N_m} = 2^{-RN_m}$ where the rate R is defined as $R = l_m/N_m$. So we have

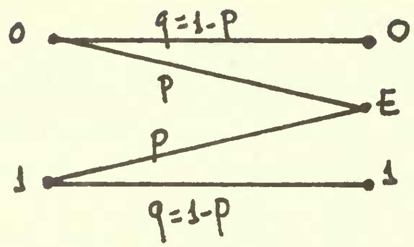
$$L(m, \underline{y}) = \sum_{n=1}^{N_m} \left[\log \frac{p(y_n|x_{mn})}{w(y_n)} + \frac{1}{N_m} \log 2^{-RN_m} \right]$$

or

$$L(m, \underline{y}) = \sum_{n=1}^{N_m} \left[\log \frac{p(y_n|x_{mn})}{w(y_n)} - R \right]$$, which is exactly (6.1.1).

6.8

BEC



(a) Clearly the set of paths that agree with y in all unerased positions contains paths which are extended by the algorithm. Therefore

$$\Pr\{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), y) = 1\} \le \Pr\{\text{path } x'_{ji}(k) \text{ agrees with } y, \text{ all unerased positions}\}$$

We now condition the right hand side on the number of unerased positions to get

$$\Pr\{\text{path } x'_{ji}(k) \text{ agrees with } y \text{ in all unerased positions}\}$$
$$= \sum_{l=0}^{2(k-j)} \Pr\{\text{path } x'_{ji}(k) \text{ agrees with } y \text{ in all unerased positions} \mid \text{there are } l \text{ unerased positions}\} \cdot \Pr\{\text{there are } l \text{ unerased positions}\}$$
$$= \sum_{l=0}^{2(k-j)} \left(\tfrac{1}{2}\right)^{l} \binom{2(k-j)}{l} q^{l} (1-q)^{2(k-j)-l}$$
$$= \sum_{l=0}^{2(k-j)} \binom{2(k-j)}{l} \left(\tfrac{q}{2}\right)^{l} (1-q)^{2(k-j)-l} = \left[\tfrac{q}{2} + (1-q)\right]^{2(k-j)}$$
$$= (1-q/2)^{2(k-j)} = 2^{2(k-j)\log(1-q/2)}$$

Let $$R_0 \equiv -\log \sum_{y}\Big(\sum_{x} q(x)\sqrt{p(y|x)}\Big)^2$$
$$= -\log \sum_{y}\Big(\tfrac{1}{2}\sqrt{p(y|0)} + \tfrac{1}{2}\sqrt{p(y|1)}\Big)^2$$
$$= -\log\Big(\tfrac{1}{4}q + \tfrac{1}{4} + 1 - q\Big) = -\log(1-q/2)$$

Finally we have

$$\Pr\{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y}) = 1\} \le 2^{-2(k-j)R_0}$$

But

$$\bar{\tau}_j = \sum_{k=j+1}^{\infty} \sum_{i=1}^{2^{k-j-1}} \overline{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y})}$$

where

$$\overline{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y})} = 0 \cdot \Pr\{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y}) = 0\} + 1 \cdot \Pr\{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y}) = 1\}$$

$$\le 2^{-2(k-j)R_0}$$

So

$$\bar{\tau}_j \le \sum_{k=j+1}^{\infty} \sum_{i=1}^{2^{k-j-1}} 2^{-2(k-j)R_0} = \sum_{k=j+1}^{\infty} 2^{k-j-1} 2^{-2(k-j)R_0}$$

$$= \sum_{k=j+1}^{\infty} 2^{(2jR_0-j-1)} 2^{(k-2kR_0)} = 2^{(2jR_0-j-1)} \sum_{k=j+1}^{\infty} \left[2^{1-2R_0}\right]^k$$

$$= 2^{(2jR_0-j-1)} \frac{2^{(1-2R_0)(j+1)}}{1-2^{(1-2R_0)}} = \frac{2^{-2R_0}}{1-2^{(1-2R_0)}} \ ;$$

For convergence we need $2^{(1-2R_0)} < 1$ or $R = \frac{1}{2} < R_0$.

(b) Clearly the set of paths that reach the top of the stack before the correct path and agree with $\underset{\sim}{y}$ in all unerased positions is a subset of all incorrect paths which are extended by the algorithm. So we can write

$$\Pr\{e(\underset{\sim}{x}, \underset{\sim}{x}'_{ji}(k), \underset{\sim}{y}) = 1\} \ge \Pr\{\underset{\sim}{x}'_{ji}(k) \text{ reaches the top of the stack before } \underset{\sim}{x}(k) \text{ and agrees with } \underset{\sim}{y} \text{ in all unerased positions}\}$$

$$= \Pr\{\underset{\sim}{x}'_{ji}(k) \text{ reaches top before } \underset{\sim}{x}(k) \mid \underset{\sim}{x}'_{ji}(k) \text{ agrees with } \underset{\sim}{y} \text{ in all unerased positions}\}$$

$$\cdot \Pr\{\underset{\sim}{x}'_{ji}(k) \text{ agrees with } \underset{\sim}{y} \text{ in all unerased positions}\}$$

The probability that $\underset{\sim}{x}'_{ji}(k)$ reaches the top of the stack before $\underset{\sim}{x}(k)$ given it agrees with $\underset{\sim}{y}$ in all unerased positions is 1/2, because $\underset{\sim}{x}'_{ji}(k)$ has the same metric in all nodes as the correct path $\underset{\sim}{x}(k)$. Using the result in part (a) and this fact we get

$$\Pr\{e(\underset{\sim}{x},\underset{\sim}{x}'_{ji}(k),\underset{\sim}{y}) = 1\} \geq \frac{1}{2}\, 2^{-2(k-j)R_0}$$

and as in part (a) we will have

$$\overline{\tilde{C}_j} \geq \frac{1}{2}\,\frac{2^{-2R_0}}{1-2^{(1-2R_0)}}$$

and for convergence of the series requires:

$$R = \frac{1}{2} < R_0$$

—//—

Chapter 7.

7.1. Let $P_N^*(\underline{v}|\underline{u}) = \prod_{n=1}^{N} p^{(n)}(v_n|u_n)$ and $P_N^*(\underline{v}) = \prod_{n=1}^{N} p^{(n)}(v_n)$.

Then
$$I(\underline{P}_N^*) = \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N^*(\underline{v}|\underline{u}) \ln \frac{P_N^*(\underline{v}|\underline{u})}{P_N^*(\underline{v})}$$
$$= \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N^*(\underline{v}|\underline{u}) \left\{ \sum_{n=1}^{N} \ln \frac{p^{(n)}(v_n|u_n)}{p^{(n)}(v_n)} \right\}$$
$$= \sum_{n=1}^{N} \left\{ \sum_{u}\sum_{v} Q(u) p^{(n)}(v|u) \ln \frac{p^{(n)}(v|u)}{p^{(n)}(v)} \right\}$$
$$= \sum_{n=1}^{N} I(\underline{p}^{(n)}) = \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) \ln \frac{P_N^*(\underline{v}|\underline{u})}{P_N^*(\underline{v})}$$

Now, since $\ln x \leq x - 1$,
$$\sum_{n=1}^{N} I(\underline{p}^{(n)}) - I(\underline{P}_N) = I(\underline{P}_N^*) - I(\underline{P}_N)$$
$$= \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) \ln \frac{P_N^*(\underline{v}|\underline{u}) \cdot P_N(\underline{v})}{P_N^*(\underline{v}) P_N(\underline{v}|\underline{u})}$$
$$\leq \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) \left\{ \frac{P_N^*(\underline{v}|\underline{u}) P_N(\underline{v})}{P_N^*(\underline{v}) P_N(\underline{v}|\underline{u})} - 1 \right\}$$
$$= \sum_{\underline{u}}\sum_{\underline{v}} \frac{Q_N(\underline{u}) P_N^*(\underline{v}|\underline{u}) P_N(\underline{v})}{P_N^*(\underline{v})} - 1 = 0$$
$$\therefore \sum_{n=1}^{N} I(\underline{p}^{(n)}) \leq I(\underline{P}_N) \qquad //$$

7.2. For any $P_N(\underline{v}|\underline{u})$, let $p^{(n)}(v_n|u_n)$ be the marginal conditional distribution for the n^{th} pair (v_n, u_n) derived from this distribution.

Note that:

$$\sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u})\, P_N(\underline{v}|\underline{u})\, d_N(\underline{u},\underline{v})$$

$$= \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u})\, P_N(\underline{v}|\underline{u}) \left\{ \frac{1}{N} \sum_{n=1}^{N} d(u_n, v_n) \right\}$$

$$= \frac{1}{N} \sum_{n=1}^{N} \left\{ \sum_{u}\sum_{v} Q(u)\, P^{(n)}(v|u)\, d(u,v) \right\}$$

$$= \sum_{u}\sum_{v} Q(u) \left\{ \frac{1}{N} \sum_{n=1}^{N} P^{(n)}(v|u) \right\} d(u,v)$$

Hence if $P_N(\underline{v}|\underline{u}) \in \mathcal{P}_{D,N}$, then

$$\hat{P}(v|u) = \frac{1}{N} \sum_{n=1}^{N} P^{(n)}(v|u) \in \mathcal{P}_{D,1} .$$

Also from (7.2.46) and (7.2.47), we have

$$\frac{1}{N} I(\underline{P}_N) \geq \frac{1}{N} \sum_{n=1}^{N} I(\underline{P}^{(n)}) \geq I\left(\frac{1}{N} \sum_{n=1}^{N} \underline{P}^{(n)} \right) = I(\hat{\underline{P}})$$

with equality if and only if $P_N(\underline{v}|\underline{u}) = \prod_{n=1}^{N} P^{(n)}(v_n|u_n)$.

Hence

$$R_N(D) \geqslant R(D) .$$

Since for any $P(v|u) \in \mathcal{P}_{D,1}$ we can choose $P_N(\underline{v}|\underline{u}) = \prod_{n=1}^{N} P^{(n)}(v_n|u_n)$ where $\frac{1}{N} I(\underline{P}_N) = I(\underline{P})$, we also have

$$R_N(D) = R(D) , \quad N = 1, 2, \cdots .$$

<u>7.3.</u>

$$-E_0(\rho, \underline{P}) = \ln \sum_u \left\{ \sum_v P(v)\, Q(u|v)^{1/1+\rho} \right\}^{1+\rho} = \ln \sum_u \alpha(u)^{1+\rho}$$

$$\left[Q(u|v)^{1/1+\rho} \right]' = Q(u|v)^{1/1+\rho} \cdot \frac{1}{(1+\rho)^2} \ln Q(u|v)$$

$$\alpha'(u) = \sum_v P(v) \left[Q(u|v)^{1/1+\rho} \right]' = \frac{1}{1+\rho} \sum_v P(v)\, Q(u|v)^{1/1+\rho} \ln Q(u|v)^{-1/1+\rho}$$

$$\left[\alpha(u)^{1+\rho} \right]' = \alpha(u)^{1+\rho} \ln \alpha(u) + (1+\rho)\, \alpha(u)^{\rho}\, \alpha'(u)$$

$$= \alpha(u)^{1+\rho} \left\{ \ln \alpha(u) + \sum_v \frac{P(v)\, Q(u|v)^{1/1+\rho}}{\alpha(u)} \ln \frac{1}{Q(u|v)^{1/1+\rho}} \right\}$$

Since $\sum_v \beta(v|u) = 1$, $\sum_v \beta'(v|u) = \left[\sum_v \beta(v|u)\right]' = 0$, $\forall \rho$.

$$\left[\alpha(u)^{1+\rho}\right]' = \alpha(u)^{1+\rho} \cdot \left\{ \sum_v \beta(v|u) \ln \frac{\alpha(u)}{Q(u|v)^{1/1+\rho}} \right\}$$

$$-E_0'(\rho, \underline{P}) = \left[\ln \sum_u \alpha(u)^{1+\rho}\right]' = \frac{\sum_u \left[\alpha(u)^{1+\rho}\right]'}{\sum_z \alpha(z)^{1+\rho}}$$

$$= \sum_u \frac{\alpha(u)^{1+\rho}}{\sum_z \alpha(z)^{1+\rho}} \cdot \sum_v \beta(v|u) \ln \frac{P(v)}{\beta(v|u)}$$

$$= \sum_u \sum_v \omega(u) \beta(v|u) \ln \frac{P(v)}{\beta(v|u)}$$

$$\omega'(u) = \frac{\left[\alpha(u)^{1+\rho}\right]'}{\sum_z \alpha(z)^{1+\rho}} - \frac{\alpha(u)^{1+\rho}}{\sum_z \alpha(z)^{1+\rho}} \cdot \sum_y \frac{\left[\alpha(y)^{1+\rho}\right]'}{\sum_z \alpha(z)^{1+\rho}}$$

$$= \omega(u) \left\{ \sum_v \beta(v|u) \ln \frac{P(v)}{\beta(v|u)} - \sum_y \omega(y) \sum_v \beta(v|y) \ln \frac{P(v)}{\beta(v|y)} \right\}$$

$$= \omega(u) \left\{ \sum_v \beta(v|u) \ln \frac{P(v)}{\beta(v|u)} + E_0'(\rho, \underline{P}) \right\}$$

$$\beta'(v|u) = P(v) \cdot \left\{ \frac{1}{\alpha(u)} \left[Q(u|v)^{1/1+\rho}\right]' - \frac{1}{\alpha^2(u)} \alpha'(u)\, Q(u|v)^{1/1+\rho} \right\}$$

$$= \frac{1}{1+\rho} \frac{P(v)\, Q(u|v)^{1/1+\rho}}{\alpha(u)} \cdot \left\{ \ln \frac{P(v)}{\beta(v|u)\,\alpha(u)} - \sum_v \beta(v|u) \ln \frac{P(v)}{\beta(v|u)\,\alpha(u)} \right\}$$

$$= \frac{1}{1+\rho} \beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} - \sum_v \beta(v|u) \ln \frac{P(v)}{\beta(v|u)} \right\}$$

$$-E_0''(\rho, \underline{P}) = \sum_u \sum_v \left[\omega(u)\beta(v|u)\right]' \ln \frac{P(v)}{\beta(v|u)} - \sum_u \sum_v \omega(u) \cancel{\beta(v|u)} \cdot \frac{\left[\beta(v|u)\right]'}{\cancel{\beta(v|u)}}$$

$$= \sum_u \sum_v \left\{ \omega'(u)\beta(v|u) + \omega(u)\beta'(v|u) \right\} \ln \frac{P(v)}{\beta(v|u)}$$

$$- \sum_u \omega(u) \sum_v \beta'(v|u)$$

$$= \sum_u \sum_v \omega(u)\beta(v|u) \left\{ \sum_y \beta(y|u) \ln \frac{P(y)}{\beta(y|u)} + E_0'(\rho, \underline{P}) \right.$$

$$\left. + \frac{1}{1+\rho} \ln \frac{P(v)}{\beta(v|u)} - \frac{1}{1+\rho} \sum_y \beta(y|u) \ln \frac{P(y)}{\beta(y|u)} \right\} \ln \frac{P(v)}{\beta(v|u)} - 0$$

$$= \frac{1}{1+\rho} \sum_u \sum_v \omega(u)\beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} \right\}^2 + \frac{\rho}{1+\rho} \sum_u \sum_v \omega(u) \left\{ \beta(v|u) \ln \frac{P(v)}{\beta(v|u)} \right\}^2$$

$$- \left[E_0'(\rho, \underline{P})\right]^2$$

Since $-1 < \rho < 0$ the second term is negative,

$$-E_0''(\rho, \underline{P}) \le \frac{1}{1+\rho} \sum_u \sum_v \omega(u)\beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} \right\}^2 .$$

For $x \ge 1$ we have the bound; $(\ln x)^2 \le \frac{4}{e^2} x \le x$,

so that for $\Delta = \{(u,v) \; ; \; P(v) \ge \beta(v|u)\}$ we have

$$-E_0''(\rho, \underline{P}) \le \frac{1}{1+\rho} \sum\sum_{(u,v)\in\Delta} \omega(u)\beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} \right\}^2$$

$$+ \frac{1}{1+\rho} \sum\sum_{(u,v)\notin\Delta} \omega(u)\beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} \right\}^2$$

$$\le \frac{1}{1+\rho} + \frac{1}{1+\rho} \sum\sum_{(u,v)\notin\Delta} \omega(u)\beta(v|u) \left\{ \ln \frac{P(v)}{\beta(v|u)} \right\}^2$$

For $(u,v) \notin \Delta$ we have the inequality

$$\ln \alpha(u) \le \ln \alpha(u) - \ln Q(u|v)^{1/1+\rho} = \ln \frac{\alpha(u)}{Q(u|v)^{1/1+\rho}}$$

$$= \ln \frac{P(v)}{\beta(v|u)} < 0$$

or $\left(\ln \frac{P(v)}{\beta(v|u)}\right)^2 \le \left(\ln \alpha(u)\right)^2$

Now

$$\ln \alpha(u) = \frac{1}{1+\rho} \ln \alpha(u)^{1+\rho} = \frac{1}{1+\rho} \ln \left[\omega(u) \sum_z \alpha(z)^{1+\rho} \right]$$

$$= \frac{1}{1+\rho} \left\{ \ln \omega(u) - E_0(\rho, \underline{P}) \right\}$$

and $\left(\ln \alpha(u)\right)^2 = \left(\frac{1}{1+\rho}\right)^2 \left\{ \left(\ln \omega(u)\right)^2 + E_0^2(\rho, \underline{P}) - 2 E_0(\rho, \underline{P}) \ln \omega(u) \right\}$

$$\le \left(\frac{1}{1+\rho}\right)^2 \left\{ \left(\ln \omega(u)\right)^2 + E_0^2(\rho, \underline{P}) \right\}$$

Therefore we have

$$-E_0''(\rho, \underline{P}) \le \frac{1}{1+\rho} + \left(\frac{1}{1+\rho}\right)^3 \left\{ \sum_u \omega(u) \left(\ln \omega(u)\right)^2 + E_0^2(\rho, \underline{P}) \right\}$$

It is easy to show that the uniform distribution maximizes the first term in the bracket so that

$$\sum_u \omega(u) \left(\ln \omega(u)\right)^2 \le \left(\ln \frac{1}{A}\right)^2 = \left(\ln A\right)^2$$

Also

$$-E_0(\rho,\underline{P}) = \ln \sum_u \left(\sum_v P(v) Q(u|v)^{1/1+\rho} \right)^{1+\rho}$$
$$\leq \ln \sum_u \left(\sum_v P(v) Q(u|v) \right)^{1+\rho} = \ln \left(\sum_u Q(u)^{1+\rho} \right)$$
$$\leq \ln \left(\sum_u 1 \right) = \ln A$$

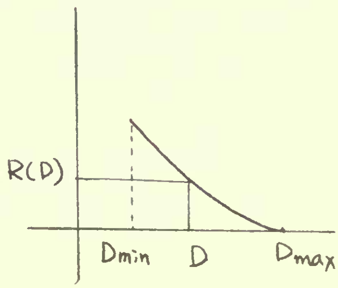
Finally we then have

$$-E_0''(\rho,\underline{P}) \leq \frac{1}{1+\rho} + \left(\frac{1}{1+\rho}\right)^3 2 (\ln A)^2$$

Since $-\frac{1}{2} \leq \rho \leq 0$ choose $\rho = -\frac{1}{2}$. Then

$$-E_0''(\rho,\underline{P}) \leq 2 + 16 [\ln A]^2 , \quad -\frac{1}{2} \leq \rho \leq 0 \quad //$$

7.4.



Corollary 7.2.2.: Given $\epsilon > 0$, there exists a block code $\mathcal{B}$ of rate $R \leq R(D) + \epsilon$ with average distiortion $d(\mathcal{B}) \leq D + \epsilon$, provided that $D > D_{min}$.

(a) Define $D_\delta = D - \delta$, $\delta > 0$. Then There exists a code $\mathcal{B}$ of rate $R \leq R(D_\delta) + \delta$ with $d(\mathcal{B}) \leq D_\delta + \delta = D$. For any $\epsilon > 0$ choose δ small enough so that $R(D-\delta) + \delta \leq R(D) + \epsilon$.
This is possible since $R(D)$ is continuous in D for $D > D_{min}$.

(b) We can always find $\delta_1, \delta_2 > 0$ as small as we please that satisfies $R(D) = R(D+\delta_1) + \delta_2$
There exists a code $\mathcal{B}$ of rate $R \leq R(D+\delta_1) + \delta_2$ such that $d(\mathcal{B}) \leq (D+\delta_1) + \delta_2 = D + \delta_1 + \delta_2$.
Thuse for any $\epsilon > 0$, there exists a code $\mathcal{B}$ of rate $R \leq R(D)$ with average distortion $d(\mathcal{B}) \leq D + \epsilon$.

7.5. Suppose in Fig. 7.6 we have $d(\mathcal{B},\mathcal{C}) \le D$. Let $P_N(\underline{v}_{\hat{m}}|\underline{u})$ be the conditional probability of $\underline{v}_{\hat{m}}$ given $\underline{u}$ out of the source. Then

$$d(\mathcal{B},\mathcal{C}) = E\{d_N(\underline{u},\underline{v}_{\hat{m}})|\mathcal{B},\mathcal{C}\} = \sum_{\underline{u}}\sum_{\underline{v}} Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, d_N(\underline{u},\underline{v}) \le D$$

Hence $\underline{P}_N \in \mathcal{P}_{D,N}$ and since $R(D) = R_N(D)$ (see problem 7.2) we have

$$R(D) = R_N(D) \le \frac{1}{N} I(\underline{P}_N)$$

The data processing theorem (Theorem 1.2.1) gives

$$I(P_N) = I(\mathcal{U}_N;\mathcal{V}_N) \le I(\mathcal{X}_{\tilde{N}};\mathcal{Y}_{\tilde{N}})$$

$$\le \tilde{N}\tilde{C} = NC \qquad \left(C \triangleq \frac{\tilde{N}}{N}\tilde{C}\right)$$

Hence $\quad R(D) \le C$.

Thus if $R(D) > C$ then $d(\mathcal{B},\mathcal{C}) > D$ for all $\mathcal{B},\mathcal{C}$.

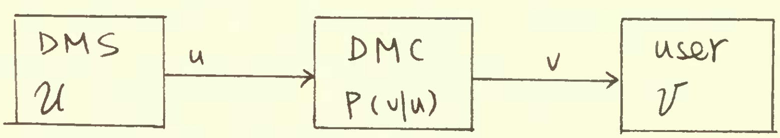
7.6. Suppose $\underline{P} \in \mathcal{P}_D$ satisfies

$$R(D) = I(\underline{P}) = \sum_u \sum_v Q(u) P(v|u) \ln \frac{P(v|u)}{\sum_{u'} Q(u') P(v|u')}$$

Also suppose that

$$C = \max_{\underline{Q}} \sum_u \sum_v Q(u) P(v|u) \ln \frac{P(v|u)}{\sum_{u'} Q(u') P(v|u')} = I(\underline{P})$$

Now consider a DMC with input alphabet $\mathcal{U}$, output alphabet $\mathcal{V}$, and transition probability $P(v|u)$ as follows,



Then the average distortion is

$$E\{d(u,v)\} = \sum_u \sum_v Q(u) P(v|u)\, d(u,v) \le D$$

since $\underline{P} \in \mathcal{P}_D$. Hence without any encoding and decoding we achieve average distortion D where $R(D) = C = R$.

For the equiprobable binary source with error distortion at fidelity D and the binary symetric channel with crossover probability ϵ where $\epsilon = D$,

$$Q(0) = Q(1) = \frac{1}{2}, \quad P(0|0) = P(1|1) = 1-D, \quad P(0|1) = P(1|0) = D.$$

$$d(u,v) = 1 - \delta_{uv}$$

$$I(\underline{P}) = \sum_u \sum_v Q(u) P(v|u) \ln \frac{P(v|u)}{\sum_{u'} Q(u') P(v|u')}$$

$$= \frac{1}{2}\left\{ 2\cdot(1-D) \ln \frac{1-D}{1/2} + 2D \ln \frac{D}{1/2} \right\} = (1-D)\ln(1-D) + D\ln D + \ln 2$$

$$= \ln 2 - \mathcal{H}(D) = C = R(D)$$

$$E\{d(u,v)\} = \sum_u \sum_v Q(u) P(v|u) d(u,v) = \frac{1}{2} \times 2 \times D \times 1 = D$$

"

7.7 Define $P_N(\underline{v}|\underline{u}) = \begin{cases} 1 ; & \underline{v} = \underline{v}(\underline{u}) \\ 0 ; & \underline{v} \neq \underline{v}(\underline{u}) \end{cases}$

where $\underline{v}(\underline{u}) \in \mathcal{B}$ is the minimum distortion codeword selected for $\underline{u}$. Here $d(\mathcal{B}) = \sum_{\underline{u}} \sum_{\underline{v}} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) d_N(\underline{u},\underline{v}) \le D$ so that $P_N \in \mathcal{P}_{D,N}$.

Since $R(D) = R_N(D)$ for all N (see problem 7.2) then

$$R(D) = R_N(D) \le \frac{1}{N} I(\underline{P}_N)$$

But in chapter 1 we have shown that

$$I(P_N) \le H(\mathcal{V}_N) = \sum_{m=1}^{M} P_m \ln \frac{1}{P_m} \le \ln M$$

Hence

$$R(D) \le \frac{1}{N} \sum_{m=1}^{M} P_m \ln \frac{1}{P_m} \le \frac{\ln M}{N} = R.$$

"

7.8. From problem 3.21. we have

$$P_E \le 2e^{-NE_x(R)} = 2 \min_{s\in[-1,0)} \min_{\underline{q}} M^{-1/s}[\gamma(s,\underline{q})]^{-N/s}$$

where
$$\gamma(s,\underline{q}) = \max_{\substack{x' \\ q(x')\ge 0}} \sum_x q(x)\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^{-s}$$

We now make two assumptions;

$$\min_{-1\le s\le 0} \min_{\underline{q}} M^{-\frac{1}{s}}[\gamma(s,\underline{q})]^{-N/s} = M^{-1/s^*}[\gamma(s^*,\underline{q}^*)]^{-N/s^*}$$

where $s^* \in [-1,0)$ and $q^*(x)>0$ for all $x\in\mathcal{X}$.

The optimizing s^* is in $[-1,0)$ for low rates R which is the case of interest. Most channels of interest also have $q^*(x)>0$ for all $x\in\mathcal{X}$.

Next define

$$f(s,\underline{q}) = \sum_x\sum_{x'} q(x)q(x')\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)^{-s}$$
$$= \sum_x\sum_{x'} q(x)q(x')e^{sd(x,x')}$$

where
$$d(x,x') = -\ln\left(\sum_y \sqrt{p(y|x)p(y|x')}\right)$$

Note that in general $f(s,\underline{q}) \le \gamma(s,\underline{q})$. Let $2c$ be a Lagrange multiplier for the constraint $\sum_x q(x)=1$ and minimize $J(\underline{q}) = f(s^*,\underline{q}) - 2c\sum_x q(x)$

subject to $q(x)\ge 0$ for all $x\in\mathcal{X}$ and fixed s^*.

Then
$$\left.\frac{\partial J(\underline{q})}{\partial q(x')}\right|_{q=q^*} = 2\sum_x q^*(x)e^{s^*d(x,x')} - 2c \ge 0$$

with equality for x' such that $q^*(x')>0$.

Hence

$$f(s^*, \underset{\sim}{q}^*) = \sum_x \sum_{x'} q^*(x) q^*(x') e^{s^* d(x,x')}$$

$$= \sum_x q^*(x) e^{s^* d(x,x')} \qquad \text{for } q^*(x') > 0$$

$$= \max_{\substack{x' \\ q(x')>0}} \sum_x q^*(x) e^{s^* d(x,x')} = \gamma(s^*, \underset{\sim}{q}^*)$$

and $q^*(x)$ is also the probability distribution that minimizes $\gamma(s^*, \underset{\sim}{q})$. By assumption $q^*(x) > 0$ for all $x \in \mathcal{X}$. Thus

$$\gamma(s^*, \underset{\sim}{q}^*) = \sum_x q^*(x) e^{s^* d(x,x')} \qquad \text{all } x' \in \mathcal{X}.$$

Now consider a source with alphabet $\mathcal{U} = \mathcal{X}$, probability $\{q^*(u); u \in \mathcal{U}\}$, representation alphabet $\mathcal{V} = \mathcal{X}$. and distortion $d(u,v)$.

Define $\quad p^*(v) = q^*(v) > 0 \quad$ for all $v \in \mathcal{V}$

and $\quad p^*(v|u) = \lambda(u) p^*(v) e^{s^* d(u,v)}$

where $\quad \lambda(u) = 1 / \sum_v p^*(v) e^{s^* d(u,v)} = 1/\gamma(s^*, \underset{\sim}{q}^*)$.

This choice of $p^*(v|u)$ satisfies the conditions of Theorem 7.6.2 where $s = s^*$ and

$$D_{s^*} = \sum_x \sum_{x'} \frac{q^*(x) q^*(x') e^{s^* d(x,x')}}{\gamma(s^*, \underset{\sim}{q}^*)} d(x,x')$$

$$R(D_{s^*} ; \underset{\sim}{q}^*) = s^* D_{s^*} - \ln \gamma(s^*, q^*)$$

Now recall that

$$E_x(R) = \max_{s \le 0} \max_{\underset{\sim}{q}} \frac{1}{s}\left[R + \ln \gamma(s, \underset{\sim}{q})\right] = \frac{1}{s^*}\left[R + \ln \gamma(s^*, \underset{\sim}{q}^*)\right]$$

or $\quad R = s^* E_x(R) - \ln \gamma(s^*, \underset{\sim}{q}^*)$

Hence $E_x(R) = D_{s^*}$ where $R = R(D_{s^*} ; \underset{\sim}{q}^*)$. Here by symmetry $R(D) = \max_{\underset{\sim}{q}} R(D; \underset{\sim}{q}) = R(D; \underset{\sim}{q}^*)$ at $D = E_x(R)$.

In general if we relax the condition $q^*(x) > 0$ for all $x \in \mathcal{X}$, then

$$\gamma(s^*, \underset{\sim}{q}^*) \le \sum_x q^*(x) e^{s^* d(x,x')}$$

with equality when $q^*(x') > 0$.

Then defining $p^*(v)$ and $P^*(v|u)$ as before with

$$\lambda(u) = 1 \Big/ \sum_v p^*(v) e^{s^* d(u,v)} \le 1/\gamma(s^*, \underset{\sim}{q}^*)$$

rusults in

$$R(D_{s^*}, \underset{\sim}{q}^*) = s^* D_{s^*} + \sum_u q^*(u) \ln \lambda(u) \le s^* D_{s^*} - \ln \gamma(s^*, \underset{\sim}{q}^*)$$

Then using $E_x(R) = \frac{1}{s^*}[R + \ln \gamma(s^*, \underset{\sim}{q}^*)]$

where

$$R = R(D_{s^*}, \underset{\sim}{q}^*)$$

yields

$$E_x(R) = \frac{1}{s^*}[R(D_{s^*}, \underset{\sim}{q}^*) + \ln \gamma(s^*, \underset{\sim}{q}^*)]$$

$$\ge \frac{1}{s^*}[s^* D_{s^*} - \ln \gamma(s^*, \underset{\sim}{q}^*) + \ln \gamma(s^*, \underset{\sim}{q}^*)] = D_{s^*}$$

Hence $E_x(R) \ge D$ where

$$R = R(D) = \max_{\underset{\sim}{q}} R(D; \underset{\sim}{q}) = R(D, \underset{\sim}{q}^*)$$

In general, necessary and sufficient conditions for equality are

$$\gamma(s^*, \underset{\sim}{q}^*) = \sum_x q^*(x) e^{s^* d(x,x')} \quad \text{for all } x' \in \mathcal{X}.$$

7.9. With $d(x,x') = -\ln\left(\sum_y \sqrt{p(y|x)\,p(y|x')}\right)$ define for code $\mathcal{C} = \{\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_M\}$ of blocklength N and rate $R = \frac{\ln M}{N}$,

$$d_{\min}(\mathcal{C}) = \min_{\substack{\underline{x}, \underline{x}' \in \mathcal{C} \\ \underline{x}' \ne \underline{x}}} d_N(\underline{x}, \underline{x}')$$

where

$$d_N(\underline{x}, \underline{x}') = \frac{1}{N} \sum_{n=1}^{N} d(x_n, x_n')$$

For all such codes let

$$d(N,R) = \max_{\mathcal{C}} d_{min}(\mathcal{C})$$

Let $\mathcal{C}^*$ be a code of blocklength N and rate R that achieves the maximum minimum distance with the fewest codeword pairs. Hence

$$d(N,R) = d_{min}(\mathcal{C}^*) \geq d(\underline{x}|\mathcal{C}^*) \quad \text{for all } \underline{x} \in \mathcal{X}_N$$

where

$$d(\underline{x}|\mathcal{C}^*) = \min_{\underline{x}' \in \mathcal{C}} d_N(\underline{x},\underline{x}')$$

This follows from the fact that if there exist an $\underline{x}^* \in \mathcal{X}_N$ such that $d(\underline{x}^*|\mathcal{C}^*) > d_{min}(\mathcal{C}^*)$ then by interchanging $\underline{x}^*$ with a codeword in $\mathcal{C}^*$ that achieves the minimum distance when paired with another codeword, there would result a new code with fewer pair of codewords that achieve the minimum distance. This contradicts the definition of $\mathcal{C}^*$. Now we have the

Gilbert Bound: $d(N,R) \geq D$ where $R = R(D)$.

(Proof): $\mathcal{C}^*$ is a code of rate R with average distortion

$$d(\mathcal{C}^*) = \sum_{\underline{x}} q_N(\underline{x})\, d(\underline{x}|\mathcal{C}^*) \leq d_{min}(\mathcal{C}) = d(N,R)$$

The converse source coding theorem states that any code with average distortion $d(\mathcal{C}^*)$ must have rate

$$R \geq R(d(\mathcal{C}^*))$$

For D satisfying $R = R(D)$ we then have

$$R = R(D) \geq R(d(\mathcal{C}^*))$$

or $\quad d(\mathcal{C}^*) \geq D$

Therefore $\quad d(N,R) \geq d(\mathcal{C}^*) \geq D$.

7.10 (a) Define $D^* = \liminf_{l \to \infty} E\{d(\underline{u}|T_l)\}$ such a limit exists for all l $\quad 0 \le E\{d(\underline{u}|T_l)\} \le d_0 < \infty$.

Let $\epsilon > 0$ arbitrary. Then there exists a value l_0 such that $\quad E\{d(\underline{u}|T_{l_0})\} \le D^* + \epsilon$.

Let $f(l)$ be a number satisfying

$f(l) = l - t l_0, \quad 0 \le f(l) < l_0, \quad t$: an integer.

Then for larger l,

$$E\{d(\underline{u}|T_l)\} \le \sum_{i=1}^{t} E\{d(\underline{u}|T_{l_0})\} + E\{d(\underline{u}|T_{f(l)})\} \qquad (*)$$

$$\le \frac{t \cdot l_0}{l}(D^* + \epsilon) + \frac{f(l)}{l} \cdot d_0 \le D^* + \epsilon + \frac{l_0}{l} d_0$$

Hence $\limsup_{l \to \infty} E\{d(\underline{u}|T_l)\} \le D^* + \epsilon$.

Since ϵ is arbitrary, we can choose ϵ very small, Therefore $\lim_{l \to \infty} E\{d(\underline{u}|T_l)\} = D^*$

exists uniquely and is independent of source output sequence.

For a fixed l, define $\quad D^* = \liminf_{m \to \infty} E\{d(\underline{u}|T_{ml})\}$.

Since $m \to \infty$ for T_{ml} is equivalent to $l \to \infty$ for T_l, both limits should be same. For arbitrary $\epsilon' > 0$, there exists a value m_0 such that $\quad E\{d(\underline{u}|T_{m_0 l})\} \le D^* + \epsilon$.

Let $f'(m) = m - t' m_0, \quad 0 \le f'(m) < m_0, \quad t'$: an integer.

Then rest of arguments are same and finally we have

$$D^* = \lim_{l \to \infty} E\{d(\underline{u}|T_l)\} = \lim_{m \to \infty} E\{d(\underline{u}|T_{ml})\} \;\;//$$

(*) $\Pr\{d(\underline{u}|T_{l_1+l_2}) \ge x\} \le \Pr\{d(\underline{u}|T_{l_1}) + d(\underline{u}|T_{l_2}) \ge x\}, \ \forall x > 0$.

$$E\{d(\underline{u}|T_{l_1+l_2})\} = \int_0^\infty \Pr\{d(\underline{u}|T_{l_1+l_2}) \ge x\}\, dx$$

$$\le \int_0^\infty \Pr\{d(\underline{u}|T_{l_1}) + d(\underline{u}|T_{l_2}) \ge x\}\, dx = E\{d(\underline{u}|T_{l_1}) + d(\underline{u}|T_{l_2})\} \;//$$

(b) $\Pr\{d_\ell(\underline{u}|\underline{v}) \le D+\delta \mid \underline{u}\} = \Pr\{\frac{1}{n\ell}\sum_{i=1}^{n\ell} d(u_i,v_i) \le D+\delta \mid \underline{u}\}$

Let $d(u_i,v_i) = x_i$. Then x_i's are i.i.d. r.v.

and $\Pr\{x=0\} = 1-D$, $\Pr\{x=1\} = D$ and $\overline{x} = D$.

Hence we can use Chernoff bound. Assume $0<\delta \ll D < \frac{1}{2}$.

$$\Pr\{\frac{1}{N}\sum_{i=1}^{N} x_i \le \bar{x}+\delta\} \gtrsim [\,\overline{e^{\lambda_0 x}}\; e^{-\lambda_0(\bar{x}+\delta)}]^N$$

where λ_0 is given by $(\bar{x}+\delta)\overline{e^{\lambda_0 x}} = \overline{x e^{\lambda_0 x}}$

So,
$$D+\delta = \frac{\Pr\{x=0\}\cdot 0\cdot e^{\lambda_0\cdot 0} + \Pr\{x=1\}\cdot 1\cdot e^{\lambda_0\cdot 1}}{\Pr\{x=0\}\cdot e^{\lambda_0\cdot 0} + \Pr\{x=1\}\cdot e^{\lambda_0\cdot 1}} = \frac{D\cdot e^{\lambda_0}}{1-D+De^{\lambda_0}}$$

$$\therefore\; e^{\lambda_0} = \frac{(D+\delta)(1-D)}{D(1-D-\delta)} \quad \text{or} \quad \lambda_0 = \ln\frac{(D+\delta)(1-D)}{D(1-D-\delta)}$$

$$\overline{e^{\lambda_0 x}} = 1-D+D\cdot e^{\lambda_0} = 1-D+\frac{(D+\delta)(1-D)}{1-D-\delta} = \frac{1-D}{1-D-\delta}$$
$$= \frac{1}{1-\frac{\delta}{1-D}} \ge \frac{1}{1-\delta^2}$$

$$\left[\overline{e^{\lambda_0 x}}\right]^{n\ell} \ge \frac{1}{(1-\delta^2)^{n\ell}} \ge \frac{1}{1-n\ell\delta^2} \ge 1-\frac{1}{n\ell\delta^2}$$

$$\lambda_0(D+\delta) = D\cdot\ln\frac{(D+\delta)(1-D)}{D(1-D-\delta)} + \delta\ln\frac{(D+\delta)(1-D)}{D(1-D-\delta)}$$

$$\delta\ln\frac{(D+\delta)(1-D)}{D(1-D-\delta)} = \delta\ln\frac{1-D}{D} - \delta\ln\frac{1-D-\delta}{D+\delta} \le \delta\ln\frac{1-D}{D} = -\delta R'(D)$$

since $R(D) = \ln 2 - \mathcal{H}(D) = \ln 2 + D\ln D + (1-D)\ln(1-D)$

and $R'(D) = \ln D + 1 - \ln(1-D) - 1 = \ln\frac{D}{1-D}$

$$D\ln\frac{(D+\delta)(1-D)}{D(1-D-\delta)} = D\ln\frac{(D+\delta)(1-D)^2}{D^2(1-D-\delta)} + D\ln D - D\ln(1-D)$$

$$= \ln\left(\frac{D+\delta}{D^2}\right)^D + \ln\left(\frac{(1-D)^2}{1-D-\delta}\right)^D + D\ln D - D\ln(1-D)$$

$$\le \ln 2 + \ln(1-D) + D\ln D - D\ln(1-D) = R(D)$$

$$\therefore\; [e^{-\lambda_0(D+\delta)}]^{n\ell} \ge e^{-n\ell[R(D)-R'(D)\delta]}$$

$$\therefore\; \Pr\{d_\ell(\underline{u},\underline{v}) \le D+\delta \mid \underline{u}\} \ge \left(1-\frac{1}{n\ell\delta^2}\right)e^{-n\ell[R(D)-R'(D)\delta]}$$

$$E[z_1] = E[\text{\# of paths with distortion } D+\delta \text{ or less from } \underline{u}]$$
$$= (\text{\# of total tree paths}) \times Pr\{\text{a path has distortion} \le D+\delta \mid \underline{u}\}$$
$$= 2^{l} \cdot Pr\{d_l(\underline{u},\underline{v}) \le D+\delta \mid \underline{u}\}$$
$$\ge \left(1-\frac{1}{nl\delta^2}\right) e^{-nl[R-R(D)+\delta R'(D)]}$$

$\therefore$ For small δ such that $R > R(D) - \delta R'(D)$, we can find l large enough to have $E[z_1] > 1$.

(c) From the branching process extinction theorem, we know when $E[z_1] > 1$, $Pr\{z_k > 0\} \le Pr\{z_{k+1} > 0\}$ and

$$\lim_{m\to\infty} Pr\{z_m > 0\} = \eta > 0.$$

Since $z_m > 0$ implies that there exists at least one codeword $\underline{v} = (\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_m)$ in the tree such that

$$d(\underline{u}_j, \underline{v}_j) \le D+\delta \qquad \text{for all } j = 1, 2, \cdots, m.$$

We have $Pr\{z_m > 0\} \le Pr\{d(\underline{u}_j, \underline{v}_j) \le D+\delta,\ \forall j = 1,2,\cdots,m\}$

or $Pr\{z_m > 0\} \le Pr\{\frac{1}{m}\sum_{j=1}^{m} d(\underline{u}_j, \underline{v}_j) \le D+\delta\} \le Pr\{d(\underline{u}|T_{ml}) \le D+\delta\}$

$$\therefore \lim_{m\to\infty} Pr\{d(\underline{u}|T_{ml}) \le D+\delta\} \ge \lim_{m\to\infty} Pr\{z_m > 0\} = \eta > 0 \quad //$$

(d) For arbitrary $\delta > 0$, choose l large enough so that

$$D^* - \delta < E\{d(\underline{u}|T_l)\} < D^* + \delta.$$

The expectation taken over the tree code ensemble for any source output sequence $\underline{u}$ also can be considered as

$$E\{d(\underline{u}|T_l)\} = \lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} d_l(\underline{u}_i | T_l^{i})$$

Therefore

$$\lim_{m\to\infty} Pr\{d(\underline{u}|T_{ml}) \le D^*-\delta\} \le \lim_{m\to\infty} Pr\{\frac{1}{m}\sum_{i=1}^{m} d_l(\underline{u}_i|T_l^{i}) \le D^*-\delta\}$$
$$= Pr\{\lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} d_l(\underline{u}_i|T_l^{i}) \le D^*-\delta\}$$
$$= Pr\{E[d(\underline{u}|T_l) \le D^*-\delta\} = 0$$

(e) From (c), $\lim_{m\to\infty} \Pr\{d(\underline{u}|T_{ml}) \le D+\delta\} > 0$, for any $\delta > 0$,

and from (d), $\lim_{m\to\infty} \Pr\{d(\underline{u}|T_{ml}) \le D^*-\delta'\} = 0$, for any $\delta' > 0$.

Therefore, for all $\delta > 0,\ \delta' > 0$, $\quad D+\delta > D^* - \delta'$

or for any $\epsilon = \delta+\delta' > 0$, $\quad D^* \le D+\epsilon$

$\therefore$ If $R = \frac{1}{n}\ln 2$ satisfies $R > R(D)$, there exists a sequence of tree codes such that

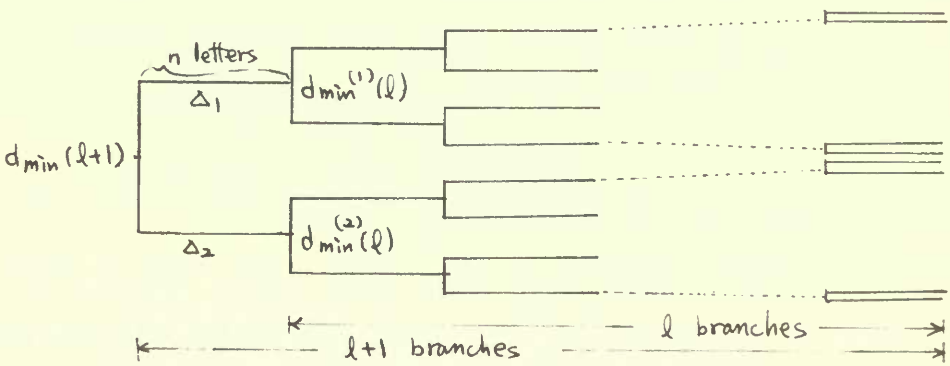
$$\lim_{l\to\infty} E\{d(\underline{u}|T_l)\} \le D \quad \text{or} \quad \lim_{m\to\infty} E\{d(\underline{u}|T_{ml})\} \le D.$$

<u>7.11</u> $G(t|l) = \Pr\{d(\underline{u}|T_l) \ge \frac{t}{nl}\} = \Pr\{nl\cdot[\min_{\underline{v}\in\mathcal{V}_{nl}} d(\underline{u},\underline{v})] \ge t\}$

Since $G(t|l)$ does not depend on source ensemble,

$$G(t|l) = \Pr\{\min_{\underline{v}\in\mathcal{V}_{nl}} [nl\cdot d(\underline{0},\underline{v})] \ge t\} = \Pr\{d_{min}(l) \ge t\}$$

where $\quad d_{min}(l) \equiv \min_{\underline{v}\in\mathcal{V}_{nl}} \{nl\cdot d(\underline{0},\underline{v})\}$



Define $\Delta_j \equiv$ # of "1"'s in j th path, $j=1$ or 2,

$$P_{\Delta_j}(k) \equiv \Pr\{\Delta_j = k\} = (\tfrac{1}{2})^n \binom{n}{k}$$

Then $d_{min}(l+1) = \min_{j=1\text{ or }2} \{\Delta_j + d_{min}^{(j)}(l)\}$. And, since Δ_j and $d_{min}^{(j)}(l)$ are i.i.d. r.v.'s respectively,

$$G(t|l+1) = \Pr\{d_{min}(l+1) \ge t\} = \Pr\{\Delta_j + d_{min}^{(j)}(l) \ge t,\ j=1\text{ or }2\}$$

$$G(t|l+1) = \left[Pr\{ \Delta + d_{min}(l) \ge t \} \right]^2 = \left[Pr\{ d_{min}(l) \ge t - \Delta \} \right]^2$$

$$= \left[\sum_{k=0}^{n} P_\Delta(k)\, Pr\{ d_{min}(l) \ge t - \Delta \mid \Delta = k \} \right]^2$$

$$= \left[\sum_{k=0}^{n} \left(\tfrac{1}{2}\right)^n \binom{n}{k} Pr\{ d_{min}(l) \ge t - k \} \right]^2$$

$$= \left[\sum_{k=0}^{n} \binom{n}{k} \left(\tfrac{1}{2}\right)^n G(t-k|l) \right]^2 \quad //$$

And, for $l=0$,

$$Pr\{ d(\underline{u}|T_l) = \tfrac{t}{n} \} = \left(\tfrac{1}{2}\right)^{2n} \left\{ \binom{n}{t}^2 + 2\binom{n}{t}\binom{n}{t+1} + 2\binom{n}{t}\binom{n}{t+2} + \cdots + 2\binom{n}{t}\binom{n}{n} \right\}$$

$$G(t|1) = Pr\{ d(\underline{u}|T_l) \ge \tfrac{t}{n} \} = Pr\{ d(\underline{u}|T_l) = \tfrac{t}{n} \} + Pr\{ d(\underline{u}|T_l) = \tfrac{t+1}{n} \}$$

$$+ \cdots + Pr\{ d(\underline{u}|T_l) = \tfrac{n}{n} \} = \left(\tfrac{1}{2}\right)^{2n} \cdot \left\{ \binom{n}{t}^2 + \binom{n}{t+1}^2 + \cdots + \binom{n}{n}^2 \right.$$

$$\left. + 2\left[\binom{n}{t}\binom{n}{t+1} + \cdots + \binom{n}{t}\binom{n}{n} + \binom{n}{t+1}\binom{n}{t+2} + \cdots + \binom{n}{t+1}\binom{n}{n} + \cdots + \cdots + \binom{n}{n-1}\binom{n}{n} \right] \right\}$$

$$= \left(\tfrac{1}{2}\right)^{2n} \left\{ \binom{n}{t} + \binom{n}{t+1} + \cdots + \binom{n}{n} \right\}^2 = \left[\sum_{k=t}^{n} \binom{n}{k} \left(\tfrac{1}{2}\right)^n \right]^2$$

$$= \left[\sum_{k=0}^{n} \binom{n}{k} \left(\tfrac{1}{2}\right)^n G(t-k|0) \right]^2$$

where $G(t|0) = \begin{cases} 0 & : \quad t \ge 1 \\ 1 & : \quad -n \le t \le 0 \end{cases}$ //

7.12

condition on the distortion: $\int_{-\infty}^{\infty} Q(u)\, d^\alpha(u,0)\, du \le d_0^\alpha$, $\alpha > 0$

With the same argument on page 424, we have (Eq. (7.5.8))

$$d(\mathcal{B}) \le D(\underline{P}) + \int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, d(\underline{u}|\mathcal{B})\, \Phi(\underline{u},\underline{v};\mathcal{B})\, d\underline{u}\, d\underline{v}.$$

Using Hölder's inequality and $\Phi^\beta = \Phi$, $\beta \ne 0$

$$\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, d(\underline{u}|\mathcal{B})\, \Phi(\underline{u},\underline{v};\mathcal{B})\, d\underline{u}\, d\underline{v}$$

$$\le \left[\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) [d(\underline{u}|\mathcal{B})]^\alpha\, d\underline{u}\, d\underline{v} \right]^{1/\alpha}$$

$$\times \left[\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, \Phi(\underline{u},\underline{v};\mathcal{B})\, d\underline{u}\, d\underline{v} \right]^{\frac{\alpha-1}{\alpha}}$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u}) \left[d(\underline{u}|\mathcal{B})\right]^\alpha d\underline{u}\, d\underline{v}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) \left[d(\underline{u}|\mathcal{B})\right]^\alpha d\underline{u} \le \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) \left[d(\underline{u},\underline{0})\right]^\alpha d\underline{u}$$

$$= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) \left[\frac{1}{N}\sum_{i=1}^{N} d(u_i,0)\right]^\alpha d\underline{u} \le \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) \left[\frac{1}{N}\sum_{i=1}^{N} d^\alpha(u_i,0)\right] d\underline{u}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[\int_{-\infty}^{\infty} Q(u_i)\, d^\alpha(u_i,0)\, du_i\right] \le d_0^\alpha$$

Hence $d(\mathcal{B}) \le D(\underline{P}) + (d_0^\alpha)^{1/\alpha}\cdot\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, \Phi(\underline{u},\underline{v}:\mathcal{B})\, d\underline{u}\, d\underline{v}\right]^{\frac{\alpha-1}{\alpha}}$

Now averaging this over the code and using Jensen's inequality yields

$$\overline{d(\mathcal{B})} \le D(\underline{P}) + d_0\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \overline{Q_N(\underline{u}) P_N(\underline{v}|\underline{u})\, \Phi(\underline{u},\underline{v}:\mathcal{B})\, d\underline{u}\, d\underline{v}}\right]^{\frac{\alpha-1}{\alpha}}$$

$$\le D(\underline{P}) + d_0\, e^{\frac{\alpha-1}{\alpha}\left[-N\, E(R;\rho,\underline{P})\right]}$$

Define $E(R,D) \equiv \sup_{\underline{P}\in\mathcal{P}_D} \max_{-1\le\rho\le 0} E(R;\rho,\underline{P})$, and since $D(\underline{P}) \le D$

$$\overline{d(\mathcal{B})} \le D + d_0\, e^{-[(\alpha-1)/\alpha]\, N E(R,D)}$$

$\therefore$ There exists a block code with average distortion $d(\mathcal{B})$ satisfying $\quad d(\mathcal{B}) \le D + d_0 e^{-[(\alpha-1)/\alpha] N E(R,D)}$

where $\quad E(R,D) > 0 \quad$ for $R > R(D)$ //

<u>7.13</u> We already know that (a) $R(D)$ is a nonincreasing function of D, and (b) $R(D)$ is a convex $\cup$ function of D for $D_{min} \le D \le D_{max}$.

Want to show that for $D \in (D_{min}, D_{max})$

1) $R(D)$ is continuous and 2) strictly decreasing.

1) Suppose $R(D)$ is discontinuous at $D_1 \in (D_{min}, D_{max})$. Then for some $\theta \in [0,1]$ and for some small $\alpha > 0$, such that $\quad D_{min} \le D_1 - \alpha < D_1 + \alpha \le D_{max}$,

the following inequality does not hold:

$$R[\theta(D_1-\alpha)+(1-\theta)(D_1+\alpha)] \le \theta R[D_1-\alpha]+(1-\theta)R[D_1+\alpha]$$

This contradicts to convexity condition (b).

2) Suppose $R(D)$ is not strictly decreasing. Then there exist some D_2 and D_3 such that $R(D_2)=R(D_3)$, $D_{min} < D_2 < D_3 < D_{max}$. But this again contradicts to convexity condition (b). Since for some $\theta \in [0,1]$ and for some small $\beta > 0$, $D_{min} \le D_2-\beta < D_2 < D_3 < D_3+\beta \le D_{max}$, the following inequality does not hold:

$$R[\theta(D_2-\beta)+(1-\theta)(D_3+\beta)] \le \theta R(D_2-\beta)+(1-\theta)R(D_3+\beta) .$$

Therfore the convexity of $R(D)$, together with the property of non-increasing, implies that $R(D)$ is a continuous strictly decreasing function of D for $D_{min} < D < D_{max}$.

Let $R(D) = I(\underline{P}^*) = \min\limits_{\underline{P}\in\mathcal{P}_D} I(\underline{P})$. Suppose $D(\underline{P}^*) = D' < D$

Then from the property of strictly decreasing $R(D') > R(D)$.

But when we set $\mathcal{P}_{D'} = \{P(v|u) ; D(\underline{P}) \le D'\}$, $\underline{P}^* \in \mathcal{P}_{D'}$, and $R(D') = \min\limits_{\underline{P}\in\mathcal{P}_{D'}} I(\underline{P}) \le I(\underline{P}^*) = R(D)$.

Therefore we have contradiction.

$\therefore$ If $\underline{P} \in \mathcal{P}_D$ yields $R(D) = I(\underline{P})$, then $D(\underline{P}) = D$.

<u>7.14</u> $R(D) = I(\underline{P}^*) = \min\limits_{\underline{P}\in\mathcal{P}_D} I(\underline{P})$, $D(\underline{P}^*) = D$ for $d(u,v)$

where $\mathcal{P}_D = \{P(v|u) ; D(\underline{P}) \le D\}$.

Define another distortion measure $\tilde{d}(u,v) = d(u,v) - \min\limits_{v\in\mathcal{V}} d(u,v)$.

Then
$$\tilde{D}(\underline{P}) = \sum_u\sum_v Q(u)P(v|u)\tilde{d}(u,v)$$
$$= \sum_u\sum_v Q(u)P(v|u)d(u,v) - \sum_u\sum_v Q(u)P(v|u)\min_{v\in\mathcal{V}} d(u,v)$$
$$= D(\underline{P}) - \sum_u Q(u)\min_{v\in\mathcal{V}} d(u,v) = D(\underline{P}) - D_{min}.$$

$$\tilde{R}(D) = I(\underline{P}^{**}) = \min_{\underline{P} \in \tilde{\mathcal{P}}_D} I(\underline{P})$$

where $\tilde{\mathcal{P}}_D = \{ P(v|u) \; ; \; \tilde{D}(\underline{P}) \le D \}$

$$= \{ P(v|u) \; ; \; D(\underline{P}) \le D + D_{min} \} = \mathcal{P}_{D+D_{min}}$$

$$\therefore \tilde{R}(D) = \min_{\underline{P} \in \tilde{\mathcal{P}}_D} I(\underline{P}) = \min_{\underline{P} \in \mathcal{P}_{D+D_{min}}} I(\underline{P}) = R(D+D_{min}) \quad //$$

7.15 $D_{min} = 0$, $\quad D_{max} = \min_{v \in \mathcal{V}} \frac{1}{2}\{ d(0,v) + d(1,v) \} = \min\{ \alpha ; \frac{1}{2} \}$

By the symmetry of source probability and distortion measure, we have $P(0) = P(1)$.

Suppose $P(0) = P(1) = 0$. Then $P(2) = 1$ and

$$D(\underline{P}) = \sum_u \sum_v Q(u) P(v|u) d(u,v) = \alpha .$$

This means $P(0) = P(1) > 0$ for $0 < D < \alpha$.

Suppose $P(0) = P(1) = \frac{1}{2}$, $P(2) = 0$. Then from

$$P(v|u) = \lambda(u) P(v) e^{s d(u,v)}$$

we have

$$2P(0) = \{ \lambda(0) + \lambda(1) e^s \} P(0)$$
$$2P(1) = \{ \lambda(0) e^s + \lambda(1) \} P(1)$$

and

$$\lambda(0) = \lambda(1) = \frac{2}{1+e^s}, \qquad D_s = \frac{e^s}{1+e^s}$$

$$\therefore \underline{R(D)} = sD + 2 \cdot \frac{1}{2} \cdot \ln\left(\frac{2}{1+e^s}\right) = D \ln \frac{D}{1-D} + \ln 2 + \ln(1-D)$$

$$= \ln 2 + D \ln D + (1-D)\ln(1-D) = \underline{\ln 2 - \mathcal{H}(D)}$$

for $\underline{0 \le D \le D^*}$. where D^* is the point at which $P(2)$ becomes greater than 0.

But when $\alpha \ge \frac{1}{2}$, there is no such a point in the region $[0, \frac{1}{2} = D_{max}]$.

$$\therefore \underline{R(D) = \ln 2 - \mathcal{H}(D), \quad 0 \le D \le \tfrac{1}{2}, \quad \text{when } \alpha \ge \tfrac{1}{2}.}$$

On the other hand, when $0 < \alpha < \frac{1}{2}$, there exists such a point in the region $(0, \alpha = D_{max})$.

For the region $D^* < D < \alpha$, $P(0) = P(1) > 0$ and $P(2) > 0$, and we have another equation for λ;

$$2P(2) = \{\lambda(0) + \lambda(1)\} e^{s\alpha} \cdot P(2) \Rightarrow 1 = \lambda e^{s\alpha}$$

Together with previous result; $\qquad 2 = \lambda(1+e^{s})$

we can find unique solution of s and λ.

$$\therefore \underline{R(D)} = sD + 2 \cdot \tfrac{1}{2} \cdot \ln\lambda = sD - s\alpha = \underline{s(D-\alpha)}, \ \underline{D^* \le D \le \alpha < \tfrac{1}{2}}.$$

where s can be calculated from following equation

$$1 + e^{s} - 2e^{s\alpha} = 0$$

and D^* can be calculated from the continuity condition of slope of $R(D)$:

$$s = \ln \frac{D^*}{1-D^*} \quad \text{or} \quad D^* = \frac{e^{s}}{1+e^{s}} \qquad //$$

<u>7.16</u> From the result of the example in page 442, we have lower bound to $R(D)$

$$R(D) \ge H(\mathcal{U}) - \mathcal{H}(D) - D\ln(A-1)$$

for all D, $\quad 0 = D_{min} \le D \le D_{max} = 1 - \max_u Q(u)$.

And we know when $Q(u) = 1/A \quad \forall u \in \mathcal{U}$,

$$R(D) = H(\mathcal{U}) - \mathcal{H}(D) - D\ln(A-1)$$

and $\quad 1 - \max_u Q(u) = 1 - \frac{1}{A} = (A-1)\frac{1}{A} = (A-1)\min_u Q(u)$

But in case that $Q(u)$ is not uniform, this lower bound gives the exact expression for $R(D)$ only when for all $v \in \mathcal{V}$, $P(v) > 0$, and there exist $s \le 0$ and

$\lambda(u)$, $u=1,2,\cdots,A$ which satisfy the following inequality with equality : $\sum_{u=1}^{A} \lambda(u) Q(u) e^{sd(u,v)} \le 1$

The choice

$$s = \ln \frac{D}{(A-1)(1-D)}, \quad \lambda(u) = \frac{1-D}{Q(u)}, \quad u=1,2,\cdots,A$$

satisfies the inequality with equality. And from (7.6.31)

$$\frac{1}{\lambda(u)} = \frac{Q(u)}{1-D} = \sum_{v=1}^{A} P(v) e^{sd(u,v)} = \sum_{v\ne u} P(v)e^{s} + \sum_{v=u} P(v)$$

$$= \sum_{v=1}^{A} P(v)e^{s} + P(u)(1-e^{s}) = e^{s} + P(u)(1-e^{s}), \quad \forall u \in \mathcal{U}.$$

or

$$P(u) = \frac{\frac{1}{\lambda(u)} - e^{s}}{1-e^{s}} = \frac{\frac{Q(u)}{1-D} - \frac{D}{(A-1)(1-D)}}{1 - \frac{D}{(A-1)(1-D)}} = \frac{(A-1)Q(u)-D}{A(1-D)-1}, \quad \forall u \in \mathcal{U}.$$

For $A \ge 2$, $D < (A-1) \min_u Q(u)$, the numerator is always positive. Since $\min_u Q(u) < \frac{1}{A}$,

$$D < (A-1) \min_u Q(u) < (A-1)\frac{1}{A} = 1 - \frac{1}{A}$$

$$\therefore \ A(1-D)-1 > A(1-1+\tfrac{1}{A}) - 1 = 0.$$

So that $P(u) > 0 \ \ \forall u$, $D < (A-1)Q(u) \ \ \forall u$, or

$$D < (A-1) \min_u Q(u).$$

Therefore for $0 \le D \le (A-1) \min_u Q(u)$,

$$R(D) = H(\mathcal{U}) - \mathcal{H}(D) - D\ln(A-1).$$

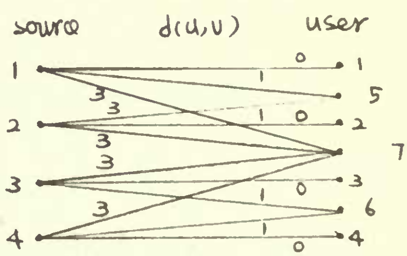
Because of continuity of $R(D)$ for $0 \le D \le D_{max}$, at $D = (A-1)\min_u Q(u)$, the equality also satisfies.

7.17 $Q(1)=Q(2)=Q(3)=Q(4) \triangleq Q = \frac{1}{4}$

From the symmetry of source and distortion measure, we know

$P(1)=P(2)=P(3)=P(4) \triangleq P_a$

$P(5)=P(6) \triangleq P_b$.



otherwise $d(u,v)=\infty$

$D_{min} = 0$, $\quad D_{max} = \min_v \sum_{u=1}^{4} Q(u)\, d(u,v) = \sum_{u=1}^{4} \frac{1}{4}\, d(u,7) = 3$.

Let's think about some special cases first.

(a) $P_a = 0$, $P_b = 0$, $P(7) = 1$.

Then $P(1|u) = P(2|u) = \cdots = P(6|u) = 0$, $P(7|u) = 1$, $\forall\, u \in \{1.2.3.4\}$

$\therefore\ D = \sum_u \sum_v Q(u)\, P(v|u)\, d(u,v) = \sum_u \frac{1}{4} \cdot P(7|u)\, d(u,7) = 3 = D_{max}$.

$R(D_{max}) = R(3) = 0$.

(b) $P_a = \frac{1}{4}$, $P_b = 0$, $P(7) = 0$. Then $P(u|u) = 1$. $P(v|u) = 0, \forall\, v \neq u$.

$\therefore\ D = 0 = D_{min}$. , $\quad R(0) = R(D_{min}) = h(\mathcal{U}) = \ln 4 = 2\ln 2$.

(c) $P_a = 0$, $P_b = \frac{1}{2}$, $P(7) = 0$.

Then $P(5|1) = P(5|2) = P(6|3) = P(6|4) = 1$, otherwise $P(v|u) = 0$.

$\therefore\ D = 4 \times \frac{1}{4} \times 1 \times 1 = 1$.

$R(1) = I(\underset{\sim}{P}) = \frac{1}{4} \sum_u \sum_v P(v|u) \ln \frac{P(v|u)}{P(v)} = \frac{1}{4} \times 4 \times 1 \times \ln \frac{1}{1/2} = \ln 2$.

(d) $P(7) = 0$, $0 < P_a < \frac{1}{4}$, and $0 < P_b < \frac{1}{2}$

Again by symmetry , $\lambda(1) = \lambda(2) = \lambda(3) = \lambda(4) = \lambda$.

From $P(v|u) = \lambda(u)\, P(v)\, e^{s d(u,v)}$, $P(v) \neq 0$,

$$P(v) = \sum_u Q(u) P(v|u) = \frac{1}{4} \lambda P(v) \sum_u e^{s d(u,v)}$$

Hence , we have

$$P_a = \frac{1}{4} \lambda \cdot P_a \cdot 1 \quad , \quad P_b = \frac{1}{4} \lambda \cdot P_b \cdot 2e^s$$

$\therefore\ \lambda = 4$, $\quad e^s = \frac{1}{2}$ or $s = -\ln 2$.

$\therefore\ D_s = \sum_u \sum_v \lambda(u)\, Q(u)\, P(v)\, e^{s d(u,v)}\, d(u,v) = P(5) + P(6) = 2P_b$

$$R(D_s) = s D_s + \sum_u Q(u) \ln \lambda(u) = -\ln 2 \cdot D_s + 4 \cdot \frac{1}{4} \cdot \ln 4$$

$$= 2\ln 2 - D_s \cdot \ln 2$$

$\therefore\ R(D) = 2\ln 2 - D\ln 2$, $\quad 0 < D < 1$.

(e) $P_a = 0$, $0 < P_b < \frac{1}{2}$, $0 < P(7) < 1$.

From $P(v) = \frac{1}{4} \lambda P(v) \sum_u e^{s d(u,v)}$, $P(v) \neq 0$, we have

$$P_b = \frac{1}{4}\lambda P_b\, 2e^{s}, \qquad P(7) = \frac{1}{4}\lambda P(7)\cdot 4e^{3s}$$

$$\therefore\ \lambda = 2\sqrt{2}, \qquad e^{s} = 1/\sqrt{2} \quad \text{or} \quad s = -\frac{1}{2}\ln 2.$$

$$\therefore\ D_s = \frac{1}{4}\cdot 2\sqrt{2}\sum_u\sum_v P(v)e^{sd(u,v)}d(u,v) = \frac{1}{\sqrt{2}}\{4P_b e^{s} + 4P(7)\cdot 3\cdot e^{3s}\}$$

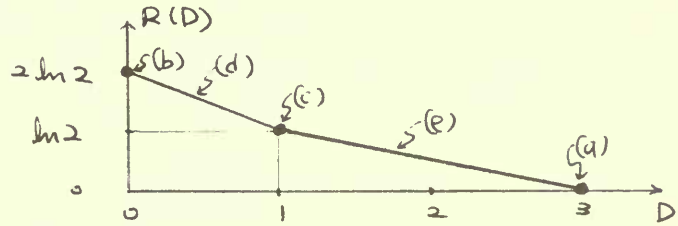
$$= 2P_b + 3P(7) = 1 - P(7) + 3P(7) = 1 + 2P(7)$$

$$R(D_s) = sD_s + 4\cdot\frac{1}{4}\ln 2\sqrt{2} = -\frac{1}{2}\ln 2\cdot D_s + \frac{3}{2}\ln 2$$

$$= (3 - D_s)\frac{1}{2}\ln 2.$$

$$\therefore\ R(D) = (3-D)\frac{1}{2}\ln 2, \qquad 1 < D < 3.$$

Therefore the rate distortion function is given as below:



7.18

For some $s \le 0$, suppose $\{P_s(v); -\infty < v < \infty\}$ satisfies

$$Q(u) = \frac{\int_{-\infty}^{\infty} P_s(v)\, e^{sd(u-v)}\, dv}{\int_{-\infty}^{\infty} e^{sd(z)}\, dz}$$

and choose λ such that

$$[\lambda(u)]^{-1} = Q(u)\int_{-\infty}^{\infty} e^{sd(z)}\, dz.$$

Then this choice of $P_s(\cdot)$ and λ satisfies the necessary and sufficient conditions of Theorem 7.7.1.

$$[\lambda(u)]^{-1} = \frac{\int_{-\infty}^{\infty} P_s(v)e^{sd(u-v)}\, dv}{\int_{-\infty}^{\infty} e^{sd(z)}\, dz}\cdot\int_{-\infty}^{\infty} e^{sd(z')}\, dz' = \int_{-\infty}^{\infty} P_s(v)e^{sd(u-v)}\, dv$$

if $P_s(v) > 0$, $\quad P_s(v|u) = \lambda(u)P_s(v)e^{sd(u-v)}$

$$\int_{-\infty}^{\infty} P_s(v|u)\, dv = 1 = \lambda(u)\int_{-\infty}^{\infty} P_s(v)\, e^{sd(u-v)}\, dv$$

if $P_s(v) = 0$,

$$\int_{-\infty}^{\infty} \lambda(u)\, Q(u)\, e^{sd(u-v)}\, du = \frac{\int_{-\infty}^{\infty} e^{sd(u-v)}\, du}{\int_{-\infty}^{\infty} e^{sd(z)}\, dz} = 1.$$

Hence, $$D_s = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda(u) Q(u) P_s(v) e^{s d(u-v)} d(u-v)\, du\, dv = \frac{\int_{-\infty}^{\infty} d(z) e^{sd(z)} dz}{\int_{-\infty}^{\infty} e^{sd(w)} dw}$$

and $$R(D_s) = sD_s + \int_{-\infty}^{\infty} Q(u) \ln \lambda(u)\, du$$
$$= sD_s + h(\mathcal{U}) - \ln \int_{-\infty}^{\infty} e^{sd(z)} dz = R_{LB}(D_s)$$

$\therefore$ Shannon's lower bound is tight.

<u>7.19</u>

(a) $$R(D) \geq h(\mathcal{U}) + sD - \ln \int_{-\infty}^{\infty} e^{sd(z)} dz \triangleq R_{LB}(D, s)$$

$$\frac{\partial}{\partial s} R_{LB}(D,s) = D - \frac{\int_{-\infty}^{\infty} d(z) e^{sd(z)} dz}{\int_{-\infty}^{\infty} e^{sd(\alpha)} d\alpha} = D - \int_{-\infty}^{\infty} d(z) G_s(z)\, dz$$

where $$G_s(z) \triangleq \frac{e^{sd(z)}}{\int_{-\infty}^{\infty} e^{sd(\alpha)} d\alpha}$$

$$\frac{\partial}{\partial s} G_s(z) = \frac{d(z) e^{sd(z)}}{\int_{-\infty}^{\infty} e^{sd(\alpha)} d\alpha} - \frac{e^{sd(z)} \int_{-\infty}^{\infty} d(\alpha) e^{sd(\alpha)} d\alpha}{\left[\int_{-\infty}^{\infty} e^{sd(\alpha)} d\alpha\right]^2}$$

$$= d(z)\, G_s(z) - G_s(z) \cdot \int_{-\infty}^{\infty} d(\beta)\, G_s(\beta)\, d\beta .$$

$$\frac{\partial^2}{\partial s^2} R_{LB}(D,s) = -\int_{-\infty}^{\infty} d(z) \frac{\partial}{\partial s} G_s(z)\, dz$$

$$= -\int_{-\infty}^{\infty} [d(z)]^2 G_s(z)\, dz + \left[\int_{-\infty}^{\infty} d(z) G_s(z)\, dz\right]^2 \leq 0$$

$\therefore$ $R_{LB}(D,s)$ is convex $\cap$ function of s, and possesses a unique maximum at the value of s that satisfies

$$\frac{\partial}{\partial s} R_{LB}(D,s) = 0 \qquad \text{i.e.} \qquad D = \int_{-\infty}^{\infty} d(z) G_s(z)\, dz .$$

(b) $\mathcal{G}_D \triangleq \{ \underline{G} \,;\, \int d(z) G(z)\, dz \leq D \}$. We want to show

$$R_{LB}(D) = h(\mathcal{U}) - \max_{\underline{G} \in \mathcal{G}_D} h(\underline{G}) = h(\mathcal{U}) + \min_{\underline{G} \in \mathcal{G}_D} [-h(\underline{G})] :$$

Want to minimize ; $-h(\underline{G})$

subject to ; $G(z) \geq 0$, $\int G(z)\, dz = 1$, $\int d(z) G(z)\, dz = D$.

$$J(\underline{G};\alpha,s) = -h(G) - \ln\alpha \int G(z)\,dz - s\int d(z)G(z)\,dz$$

$$= \int G(z)\cdot\{\ln G(z) - \ln\alpha - s\,d(z)\}\,dz = \int G(z)\cdot\ln\frac{G(z)}{\alpha e^{sd(z)}}\,dz$$

Take a variation $\epsilon\eta(z)$ about $G(z)$ such that

$G(z)+\epsilon\cdot\eta(z) \geq 0 \quad \forall z,$ for which $\int \eta(z)\,dz = 0$.

Then $J(\underline{G}+\epsilon\underline{\eta};\alpha,s) = \int [G(z)+\epsilon\eta(z)]\ln\frac{G(z)+\epsilon\eta(z)}{\alpha e^{sd(z)}}\,dz$

$$\frac{\partial}{\partial\epsilon}J(\underline{G}+\epsilon\underline{\eta};\alpha,s)\Big|_{\epsilon=0} = \int \eta(z)\ln\frac{G(z)}{\alpha e^{sd(z)}}\,dz + \int G(z)\cdot\frac{\eta(z)}{G(z)}\,dz$$

$$= \int \eta(z)\ln\frac{G(z)}{\alpha e^{sd(z)}}\,dz + 0 = 0$$

If $K = \dfrac{G(z)}{\alpha e^{sd(z)}}$ is not a function of z, then

$\frac{\partial}{\partial\epsilon}J(\underline{G}+\epsilon\underline{\eta};\alpha,s)\Big|_{\epsilon=0} = 0$. and $G(z) = K\alpha\, e^{sd(z)}$

From $1 = \int G(z)\,dz = K\alpha\int e^{sd(z)}\,dz$, we have $\alpha = \dfrac{1}{K\int e^{sd(z)}dz}$

Therefore

$$\min_{\underline{G}\in\mathcal{G}_D}[-h(\underline{G})] = \min_{\underline{G}\in\mathcal{G}_D}\left[\int G(z)\ln G(z)\,dz\right]$$

$$= \min_{\underline{G}\in\mathcal{G}_D}\int G(z)[\ln K\alpha + sd(z)]\,dz$$

$$= \min_{\underline{G}\in\mathcal{G}_D}\left[s\int d(z)G(z)\,dz + \left(\ln\frac{K}{K\int e^{sd(z)}dz}\right)\cdot\int G(z)\,dz\right]$$

$$= sD - \ln\int e^{sd(z)}\,dz$$

$$\therefore\ R_{LB}(D) = h(\mathcal{U}) + \min_{\underline{G}\in\mathcal{G}_D}[-h(G)] = h(\mathcal{U}) + sD - \ln\int e^{sd(z)}\,dz \;//$$

<u>7.20</u> $d(u,v) = d(u-v) = |u-v|$, $d(z) = |z|$, $s\leq 0$

(a) $\displaystyle\int_{-\infty}^{\infty} e^{sd(z)}\,dz = \int_{-\infty}^{\infty} e^{s|z|}\,dz = 2\int_0^{\infty} e^{sz}\,dz = \frac{2}{-s} = \frac{2}{|s|}$

$$G_s(z) = \frac{e^{sd(z)}}{\int e^{sd(\alpha)}\,d\alpha} = \frac{|s|}{2} e^{s|z|}$$

$$D = \int d(z) G_s(z)\,dz = \int_{-\infty}^{\infty} |z| \cdot \frac{|s|}{2} e^{s|z|}\,dz = \int_0^{\infty} -sz\,e^{sz}\,dz = \frac{1}{|s|}$$

$$R_{LB}(D) = h(\mathcal{U}) + sD - \ln \int_{-\infty}^{\infty} e^{sd(z)}\,dz$$

$$= h(\mathcal{U}) - 1 - \ln(2D) = h(\mathcal{U}) - \ln(2eD)$$

(b) When $P(v)$ satisfy (7.7.46) and $G_s(z) = e^{sd(z)} / \int e^{sd(\alpha)}\,d\alpha$,

$$Q(u) = \frac{\int P(v) e^{sd(u-v)}\,dv}{\int e^{sd(\alpha)}\,d\alpha} = \int P(v) \frac{e^{sd(u-v)}}{\int e^{sd(\alpha)}\,d\alpha}\,dv = \int P(v) G_s(u-v)\,dv$$

Let $\tilde{H}(\omega)$ be the Fourier transformation of $H(\cdot)$.
Then we have $\tilde{Q}(\omega) = \tilde{P}(\omega)\cdot\tilde{G}_s(\omega)$ or $\tilde{P}(\omega) = \tilde{Q}(\omega)/\tilde{G}_s(\omega)$.
And $\tilde{G}_s(\omega) = \int_{-\infty}^{\infty} G_s(z) e^{-j\omega z}\,dz = \frac{|s|}{2} \int_{-\infty}^{\infty} e^{s|z| - j\omega z}\,dz = \frac{s^2}{s^2+\omega^2}$

$$\therefore\ \tilde{P}(\omega) = \left(\frac{s^2+\omega^2}{s^2}\right) \tilde{Q}(\omega) = \tilde{Q}(\omega) + D^2 \omega^2 \tilde{Q}(\omega)$$

Taking inverse Fourier transform, we have

$$P(v) = Q(v) - D^2 Q''(v), \quad -\infty < v < \infty.$$

(c) $Q(u) = \frac{\alpha}{2} e^{-\alpha|u|}$, $-\infty < u < \infty$. $Q''(u) = \frac{\alpha^3}{2} e^{-\alpha|u|} + k\delta(u)$.

Since $Q(v)$ does not have its derivatives at $v=0$, we can put it as a delta function. But $\int P(v)\,dv = 1$, $\int Q(v)\,dv = 0$, $D \neq 0$,

$$\int_{-\infty}^{\infty} Q''(v)\,dv = 0 = \alpha^2 + k \qquad \therefore \text{ We have } k = -\alpha^2.$$

Therefore $P(v) = Q(v) - D^2 Q''(v) = \frac{\alpha}{2} e^{-\alpha|u|} - D^2 \left\{ \frac{\alpha^3}{2} e^{-\alpha|u|} - \alpha^2 \delta(u) \right\}$

Hence for all D, $0 \le D \le D_{max} = \frac{1}{\alpha}$, $P(v) \ge 0$. And since

$$h(\mathcal{U}) = -\int_{-\infty}^{\infty} Q(u) \ln\left[\frac{\alpha}{2} e^{-\alpha|u|}\right] du = \ln\left(\frac{2}{\alpha}\right) + \alpha \int_{-\infty}^{\infty} |u|\,Q(u)\,du$$

$$= \ln\left(\frac{2}{\alpha}\right) + 1 = \ln\left(\frac{2e}{\alpha}\right)$$

$$\therefore\ R(D) = R_{LB}(D) = h(\mathcal{U}) - \ln(2eD) = -\ln(\alpha D) \quad \text{for } 0 \le D \le \frac{1}{\alpha}$$

(d) $Q(v) = \frac{2}{\pi}(1+v^2)^{-2}$, $Q'(v) = \frac{2}{\pi}(-2)(2v)(1+v^2)^{-3}$

$$Q''(v) = \frac{2}{\pi}(-4)\cdot\left\{(1+v^2)^{-3} - 3\cdot 2\cdot v^2(1+v^2)^{-4}\right\}$$

$$= \frac{2}{\pi}(-4)(1+v^2)^{-4}\cdot\left\{1+v^2-6v^2\right\} = \frac{2}{\pi}(-4)(1+v^2)^{-4}(1-5v^2)$$

$$P(v) = Q(v) - D^2 Q''(v) = \frac{2}{\pi}(1+v^2)^{-4}\cdot\left\{(1+v^2)^2 + 4D^2(1-5v^2)\right\}$$

$$= \frac{2}{\pi}(1+v^2)^{-4}\cdot\left\{v^4 + 2v^2(1-10D^2) + 1 + 4D^2\right\}$$

$$= \frac{2}{\pi}(1+v^2)^{-4}\left\{[v^2+(1-10D^2)]^2 + 24D^2\left[1-\frac{100}{24}D^2\right]\right\} > 0$$

$$\Rightarrow D^2 < \frac{24}{100} = \frac{6}{25} \Rightarrow D < \frac{\sqrt{6}}{5}$$

$$h(\mathcal{U}) = -\int_{-\infty}^{\infty} \frac{2}{\pi}(1+u^2)^{-2} \ln\left[\frac{2}{\pi}(1+u^2)^{-2}\right] du = \ln(8\pi) - 2$$

$$\therefore R_{LB}(D) = R(D) = \ln(8\pi) - 2 - \ln(2eD) = \ln\left(\frac{4\pi}{D}\right) - 3, \quad \text{for } 0 < D < \frac{\sqrt{6}}{5}.$$

7.21

Shannon's lower bound is tight if and only if for some $s \le 0$, a probability density $P_s(\cdot)$ satisfies (7.7.46)

$$Q(u) = \int_{-\infty}^{\infty} P_s(v)\frac{e^{sd(u-v)}}{\int_{-\infty}^{\infty} e^{sd(z)}dz}dv = \int_{-\infty}^{\infty} P_s(v)G_s(u-v)dv$$

where $G_s(z) = e^{sd(z)} / \int_{-\infty}^{\infty} e^{sd(\alpha)} d\alpha$

Cramer's theorem says that $Q(\cdot)$ is a Gaussian density if and only if $P_s(\cdot)$ and $G_s(\cdot)$ are both Gaussian density. But if $d(u,v) = d(u-v) \neq (u-v)^2$, then $G_s(\cdot)$ can not be a Gaussian density, and so there does not exist a $P_s(\cdot)$ that satisfies (7.7.46).

Therefore, for a memoryless Gaussian source and $d(u,v) \neq (u-v)^2$, $R(D) > R_{LB}(D)$ for all D.

7.22 $d(u,v) = d(u-v) = |u-v|^\alpha, \qquad d(z) = |z|^\alpha$

$$\int_{-\infty}^{\infty} e^{sd(z)}\,dz = 2\int_0^\infty e^{s|z|^\alpha}\,dz = 2\int_0^\infty e^{-(rz)^\alpha}\,dz \qquad \left(\text{where } r = |s|^{\frac{1}{\alpha}}\right)$$

$$= \frac{2}{\alpha r}\Gamma\left(\frac{1}{\alpha}\right) = \frac{2}{\alpha |s|^{\frac{1}{\alpha}}}\Gamma\left(\frac{1}{\alpha}\right)$$

$$G_s(z) = \frac{e^{sd(z)}}{\int_{-\infty}^{\infty} e^{sd(\omega)}\,d\omega} = \frac{\alpha|s|^{\frac{1}{\alpha}}}{2\Gamma\left(\frac{1}{\alpha}\right)} e^{s|z|^\alpha}$$

$$D_s = \int_{-\infty}^{\infty} d(z)G_s(z)\,dz = 2\cdot\frac{\alpha|s|^{\frac{1}{\alpha}}}{2\Gamma\left(\frac{1}{\alpha}\right)}\int_0^\infty z^\alpha e^{sz^\alpha}\,dz$$

$$= \frac{\alpha|s|^{\frac{1}{\alpha}}}{\Gamma\left(\frac{1}{\alpha}\right)}\cdot\frac{\Gamma\left(\frac{\alpha+1}{\alpha}\right)}{\alpha|s|^{\frac{\alpha+1}{\alpha}}} = \frac{1}{|s|}\cdot\frac{\Gamma\left(\frac{\alpha+1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \;//$$

$$R_{LB}(D_s) = h(\mathcal{U}) + sD_s - \ln\left[\int_{-\infty}^{\infty} e^{sd(z)}\,dz\right]$$

$$= h(\mathcal{U}) - \frac{\Gamma\left(\frac{\alpha+1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} - \ln\left[\frac{2}{\alpha|s|^{\frac{1}{\alpha}}}\Gamma\left(\frac{1}{\alpha}\right)\right] \;//$$

$\alpha=1$; $\quad \Gamma\left(\frac{1}{\alpha}\right) = \Gamma(1) = 1. \quad \Gamma\left(\frac{\alpha+1}{\alpha}\right) = \Gamma(2) = 1.$

$$D_s = \frac{1}{|s|}, \quad R_{LB}(D_s) = h(\mathcal{U}) - 1 - \ln\left[\frac{2}{|s|}\cdot 1\right] = h(\mathcal{U}) - \ln(2eD).$$

$\alpha=2$: $\quad \Gamma\left(\frac{1}{\alpha}\right) = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma\left(\frac{\alpha+1}{\alpha}\right) = \Gamma\left(\frac{3}{2}\right) = \sqrt{\pi}/2$

$$D_s = \frac{1}{2|s|}, \quad R_{LB}(D_s) = h(\mathcal{U}) - \frac{1}{2} - \ln\left[\frac{2}{2\sqrt{|s|}}\sqrt{\pi}\right] = h(\mathcal{U}) - \frac{1}{2}\ln(2\pi eD) \;//$$

7.23

Lemma: A lower bound to the rate distortion function for a continuous memoryless source with differential entropy $h(\mathcal{U})$ is given by

$$R(D) \geq R_{LB}(D) = h(\mathcal{U}) + sD - \ln\int_{-\infty}^{\infty} e^{sd(u,v^*)}\,du$$

where v^* satisfies

$$\int_{-\infty}^{\infty} e^{sd(u,v^*)}\,du = \max_{v\in\mathcal{V}}\int_{-\infty}^{\infty} e^{sd(u,v)}\,du$$

and s satisfies the constraint ;

$$D = \int_{-\infty}^{\infty} d(u,v^*)\, e^{sd(u,v^*)}\, du \Big/ \int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du .$$

(Proof) Choose $\lambda(u)$ such that

$$[\lambda(u)]^{-1} = Q(u) \int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du .$$

Then
$$\int_{-\infty}^{\infty} \lambda(u)\, Q(u)\, e^{sd(u,v)}\, du = \frac{\int_{-\infty}^{\infty} e^{sd(u,v)}\, du}{\int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du} \le 1 .$$

Hence $\lambda \in \Lambda_s$. where Λ_s is defined in (7.7.14).

And for any $s \le 0$ we have

$$R(D) = \sup_{s \le 0} \left[sD + \int_{-\infty}^{\infty} Q(u) \ln \lambda(u)\, du \right] \qquad (7.7.13)$$

or
$$\begin{aligned} R(D) &\ge sD + \int_{-\infty}^{\infty} Q(u) \ln \lambda(u)\, du \\ &= sD - \int_{-\infty}^{\infty} Q(u) \left[\ln Q(u) + \ln \int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du \right] du \\ &= sD + h(\mathcal{U}) - \ln \int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du = R_{LB}(D) \end{aligned}$$

$$\frac{\partial}{\partial s} R(D) = D - \frac{\int_{-\infty}^{\infty} d(u,v^*)\, e^{sd(u,v^*)}\, du}{\int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du} = 0 .$$

Therefor s satisfies the constraint :

$$D = \int_{-\infty}^{\infty} d(u,v^*)\, e^{sd(u,v^*)}\, du \Big/ \int_{-\infty}^{\infty} e^{sd(u,v^*)}\, du .$$ Q.E.D.

If $d(u,v) = d(u-v)$, then

$$\int_{-\infty}^{\infty} e^{sd(u,v)}\, du = \int_{-\infty}^{\infty} e^{sd(u-v)}\, du = \int_{-\infty}^{\infty} e^{sd(z)}\, dz$$

∴ We have Shannon's lower bound :

$$R(D) \ge R_{LB}(D) = sD + h(\mathcal{U}) - \ln \int_{-\infty}^{\infty} e^{sd(z)}\, dz$$

7.24. We can prove Theorem 7.7.1 by proving similar theorem to Theorem 7.6.1 first:

$$J(\underline{P}:\lambda,s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Q(u)P(v|u)\ln\frac{P(v|u)}{\lambda(u)P(v)e^{sd(u,v)}}\,du\,dv \quad (7.6.18)'$$

Theorem 7.6.1$'$: A necessary and sufficient condition for $\{P(v|u)\}$ to minimize $J(\underline{P}:\lambda,s)$, subject to only inequality constraint (7.7.4), is that it satisfies (7.7.7) and (7.7.8).

The proof of Theorem 7.6.1$'$ is exactly same as that of Theorem 7.6.1 with only exception that summation is changed to integral.

To find the necessary and sufficient conditions for $\underline{P}\in\mathcal{P}_D$ that yield $R(D)$, from (7.7.5) and (7.7.7),

$$1 = \int_{-\infty}^{\infty} P(v|u)\,dv = \lambda(u)\int_{-\infty}^{\infty} P(v)e^{sd(u,v)}\,dv .$$

Hence we have (7.7.9).

$\therefore$ By substituting (7.7.7) into (7.7.6) and using the fact that $R(D)=I(\underline{P})$ when $D=D(\underline{P})$, we obtain (7.7.10) and (7.7.11).//

7.25

(a) Similar to Theorem 7.7.2:

$$R(D) = \sup_{s\le 0,\,\lambda\in\Lambda_s}\left[sD+\sum_{u=0}^{\infty}Q(u)\ln\lambda(u)\right]$$

where $\Lambda_s = \left\{\lambda(u) : \sum_{u=0}^{\infty}\lambda(u)Q(u)e^{s|u-v|}\le 1,\quad v\in\mathcal{V}\right\}$

choose λ such that

$$\lambda(0)Q(0) = \frac{1}{1+e^s}, \quad \lambda(u)Q(u) = \frac{1-e^s}{1+e^s}, \quad u\ge 1.$$

Then

$$\sum_{u=0}^{\infty} \lambda(u) Q(u) e^{s|u-v|} = \lambda(0)Q(0)e^{sv} + \sum_{u=1}^{\infty} \lambda(u)Q(u)e^{s|u-v|}$$

$$= \frac{e^{sv}}{1+e^s} + \frac{1-e^s}{1+e^s}\left\{ \sum_{u=1}^{v-1} e^{s(v-u)} + \sum_{u=v}^{\infty} e^{s(u-v)} \right\}$$

$$= \frac{e^{sv}}{1+e^s} + \frac{1-e^s}{1+e^s}\cdot\left\{ e^s \cdot \frac{1-e^{s(v-1)}}{1-e^s} + \frac{1}{1-e^s} \right\} = \frac{e^{sv}+e^s-e^{sv}+1}{1+e^s} = 1$$

Hence such choice of λ belongs to Λ_s. Then

$$\sum_{u=0}^{\infty} Q(u) \ln \lambda(u) = Q(0) \ln \frac{1}{Q(0)(1+e^s)} + \sum_{u=1}^{\infty} Q(u) \ln \frac{1-e^s}{Q(u)(1+e^s)}$$

$$= \sum_{u=0}^{\infty} Q(u) \ln \frac{1}{Q(u)} + \left(\ln \frac{1}{1+e^s}\right)\cdot \sum_{u=0}^{\infty} Q(u) + \left(\ln(1-e^s)\right) \sum_{u=1}^{\infty} Q(u)$$

$$= H(\mathcal{U}) - \ln(1+e^s) + [1-Q(0)]\ln(1-e^s) \;//$$

$$\therefore\ R(D) \ge R_{LB}(D) = \sup_{s\le 0}\left[sD + H(\mathcal{U}) - \ln(1+e^s) + [1-Q(0)]\ln(1-e^s)\right]$$

Since $m(s) = sD + H(\mathcal{U}) - \ln(1+e^s) + [1-Q(0)]\ln(1-e^s)$ is strictly convex $\cap$ of s,

$$\frac{\partial}{\partial s} m(s) = D - \frac{e^s}{1+e^s} + [1-Q(0)]\frac{-e^s}{1-e^s} = 0,$$

and we get lower bound of $R(D)$ parametrically by

$$D_s = \frac{e^s}{1+e^s} + [1-Q(0)]\frac{e^s}{1-e^s}$$

and

$$R_{LB}(D_s) = H(\mathcal{U}) + sD_s - \ln(1+e^s) + [1-Q(0)]\ln(1-e^s) \;//$$

(b) It suffice to show that parameter s is the slope of $R_{LB}(D)$ at $D=D_s$.

$$R_{LB}'(D_s) = \left.\frac{dR_{LB}(D)}{dD}\right|_{D=D_s} = \frac{dR_{LB}(D)}{ds}\cdot\frac{ds}{dD}$$

$$= \left\{ D_s + s\cdot\frac{dD_s}{ds} - \left[\frac{e^s}{1+e^s} + [1-Q(0)]\frac{e^s}{1-e^s}\right]\right\}\frac{ds}{dD}$$

$$= s\cdot\frac{dD_s}{ds}\cdot\frac{ds}{dD} = s \;//$$

With similar argument to theorem 7.7.1 and Lemma 7.7.1,

$$R(D_s) = sD_s + \sum_u Q(u)\ln\lambda(u) \quad , \quad D_s = \sum_u\sum_v \lambda(u)Q(u)P(v)e^{s|u-v|}|u-v| .$$

where $[\lambda(u)]^{-1} = \sum_{v=0}^{\infty} P(v)e^{s|u-v|}$, $\forall u \in \mathcal{J}$

and $R'(D_s) = \left.\frac{dR(D)}{dD}\right|_{D=D_s} = s$ //

$\therefore \quad R_{LB}(D) = R(D)$

<u>7.26</u>

See Rubin [1973].

<u>7.27</u>

See Tan and Yao [1975].

<u>7.28</u>

(a) Given $\underline{u} \in \mathcal{U}_N$ an error is made if for some $\underline{\tilde{u}} \in \mathcal{U}_N$ where $\underline{\tilde{u}} \neq \underline{u}$, we have $f(\underline{\tilde{u}}) = f(\underline{u})$ and $Q_N(\underline{\tilde{u}}) \geq Q_N(\underline{u})$.

Hence $Pr\{\underline{\hat{u}} \neq \underline{u} \,|\, \underline{u}\} \leq \sum_{\underline{\tilde{u}} \neq \underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f}) \cdot \Phi(\underline{u},\underline{\tilde{u}}|\underline{Q})$

and for $0 \leq \rho \leq 1$

$$Pr\{\underline{\hat{u}} \neq \underline{u} \,|\, \underline{u}\} \leq \left[Pr\{\underline{\hat{u}} \neq \underline{u}|\underline{u}\}\right]^{\rho} \leq \left[\sum_{\underline{\tilde{u}} \neq \underline{u}} \psi(\underline{u},\underline{\tilde{u}}|f)\Phi(\underline{u},\underline{\tilde{u}}|\underline{Q})\right]^{\rho}$$

Hence

$$Pr\{\underline{\hat{u}} \neq \underline{u}\} \leq \sum_{\underline{u}} Q_N(\underline{u})\left[\sum_{\underline{\tilde{u}} \neq \underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\,\Phi(\underline{u},\underline{\tilde{u}}|\underline{Q})\right]^{\rho}$$

(b) For any $0 \leq \rho \leq 1$, $\Phi(\underline{u},\underline{\tilde{u}}|\underline{Q}) \leq \left(\frac{Q_N(\underline{\tilde{u}})}{Q_N(\underline{u})}\right)^{\frac{1}{1+\rho}}$ Thus

$$Pr\{\underline{\hat{u}} \neq \underline{u}\} \leq \sum_{\underline{u}} Q_N(\underline{u})^{1-\frac{\rho}{1+\rho}}\left[\sum_{\underline{\tilde{u}} \neq \underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\, Q_N(\underline{\tilde{u}})^{1/1+\rho}\right]^{\rho}$$

$$= \sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho}\left[\sum_{\underline{\tilde{u}} \neq \underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\, Q_N(\underline{\tilde{u}})^{1/1+\rho}\right]^{\rho} \quad //$$

(c) For $0 \le \rho \le 1$, using Jensen's inequality, we have

$$\overline{\Pr\{\hat{\underline{u}} \ne \underline{u}\}} \le \sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho} \Big[\sum_{\tilde{\underline{u}} \ne \underline{u}} \overline{\psi(\underline{u}, \tilde{\underline{u}} | \underline{f})}\, Q_N(\tilde{\underline{u}})^{1/1+\rho} \Big]^{\rho}.$$

But $\overline{\psi(\underline{u}, \tilde{\underline{u}} | \underline{f})} = \Pr\{f(\tilde{\underline{u}}) = f(\underline{u})\} = \frac{1}{M}$. Therefore

$$\overline{\Pr\{\hat{\underline{u}} \ne \underline{u}\}} \le M^{-\rho} \sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho} \Big[\sum_{\tilde{\underline{u}} \ne \underline{u}} Q_N(\tilde{\underline{u}})^{1/1+\rho} \Big]^{\rho}$$

$$\le M^{-\rho} \Big(\sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho} \Big) \cdot \Big(\sum_{\underline{u}'} Q_N(\underline{u}')^{1/1+\rho} \Big)^{\rho}$$

$$= M^{\rho} \Big(\sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho} \Big)^{1+\rho} \quad ''$$

(d) Since $M = e^{RN}$ and

$$\sum_{\underline{u}} Q_N(\underline{u})^{1/1+\rho} = \sum_{\underline{u}} \prod_{n=1}^{N} Q(u_n)^{1/1+\rho} = \prod_{n=1}^{N} \Big(\sum_{u} Q(u)^{1/1+\rho} \Big) = \Big(\sum_{u} Q(u)^{\frac{1}{1+\rho}} \Big)^{N}$$

We get from (c)

$$\overline{\Pr\{\hat{\underline{u}} \ne \underline{u}\}} \le e^{-N E_s(R:\rho)}$$

where $E_s(R:\rho) = \rho R - E_0(\rho)$, $E_0(\rho) = (1+\rho) \ln \Big(\sum_u Q(u)^{\frac{1}{1+\rho}} \Big)$

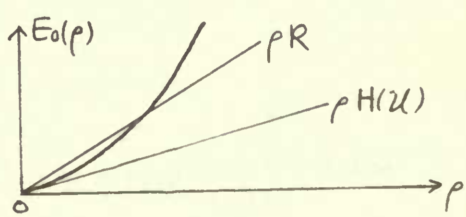
Note that $$E_0'(\rho) = - \sum_u Q(u;\rho) \ln Q(u:\rho)$$

where $$Q(u:\rho) = \frac{Q(u)^{1/1+\rho}}{\sum_{u'} Q(u')^{1/1+\rho}}$$

Then $E_0'(\rho)\Big|_{\rho=0} = H(\mathcal{U})$, and

$$E_0''(\rho) = \frac{1}{1+\rho} \sum_u Q(u:\rho) \Big[\ln Q(u:\rho) - \sum_{u'} Q(u:\rho) \ln Q(u':\rho) \Big]^2 > 0.$$

Thus we have



Hence $E_s(R) \triangleq \max_{0\le\rho\le 1} E_s(R:\rho) > 0$ for $R > H(\mathcal{U})$.

$\therefore$ There exists at least one encoder for each N that satisfies this averaged bound.

<u>7.29</u>

Here redefin $\Phi(\cdot|\cdot)$ as $\Phi(\underline{u},\underline{\tilde{u}}|Q,\underline{v}) = \begin{cases} 1 & : Q_N(\underline{\tilde{u}}|\underline{v}) \ge Q_N(\underline{u}|\underline{v}) \\ 0 & : Q_N(\underline{\tilde{u}}|\underline{v}) < Q_N(\underline{u}|\underline{v}) \end{cases}$

Then following the previous problem we have

$$\Pr\{\underline{\hat{u}} \ne \underline{u} | \underline{u},\underline{v}\} \le \left[\sum_{\underline{\tilde{u}}\ne\underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\,\Phi(\underline{u},\underline{\tilde{u}}|Q,\underline{v})\right]^\rho$$

and

$$\Pr\{\underline{\hat{u}} \ne \underline{u}\} \le \sum_{\underline{v}} \sum_{\underline{u}} Q_N(\underline{u},\underline{v}) \left[\sum_{\underline{\tilde{u}}\ne\underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\,\Phi(\underline{u},\underline{\tilde{u}}|Q,\underline{v})\right]^\rho$$

$$\le \sum_{\underline{v}} Q_N(\underline{v}) \sum_{\underline{u}} Q_N(\underline{u}|\underline{v})^{\frac{1}{1+\rho}} \left[\sum_{\underline{\tilde{u}}\ne\underline{u}} \psi(\underline{u},\underline{\tilde{u}}|\underline{f})\, Q_N(\underline{\tilde{u}}|\underline{v})^{\frac{1}{1+\rho}}\right]^\rho$$

and

$$\overline{\Pr\{\underline{\hat{u}} \ne \underline{u}\}} \le \sum_{\underline{v}} Q_N(\underline{v}) \sum_{\underline{u}} Q_N(\underline{u}|\underline{v})^{\frac{1}{1+\rho}} \left[\sum_{\underline{\tilde{u}}\ne\underline{u}} \overline{\psi(\underline{u},\underline{\tilde{u}}|\underline{f})}\, Q_N(\underline{\tilde{u}}|\underline{v})^{\frac{1}{1+\rho}}\right]^\rho$$

$$\le M^{-\rho} \sum_{\underline{v}} Q_N(\underline{v}) \left[\sum_{\underline{u}} Q_N(u|v)^{\frac{1}{1+\rho}}\right]^{1+\rho}$$

$$= e^{-\rho R N} \left[\sum_{v} Q(v) \left\{\sum_{u} Q(u|v)^{\frac{1}{1+\rho}}\right\}^{1+\rho}\right]^N$$

$$= e^{-N E_s(R;\rho)}$$

where

$$E_s(R;\rho) = \rho R - E_0(\rho), \quad E_0(\rho) = \ln\left\{\sum_v Q(v)\left[\sum_u Q(u|v)^{\frac{1}{1+\rho}}\right]^{1+\rho}\right\}.$$

Note that

$$E_0'(\rho)\big|_{\rho=0} = -\sum_v Q(v)\left\{\sum_u Q(u|v)\ln Q(u|v)\right\} = H(\mathcal{U}|\mathcal{V})$$

Hence defining

$$E_s(R) \triangleq \max_{0\le\rho\le1} E_s(R;\rho)$$

we have $\overline{\Pr\{\underline{\hat{u}} \neq \underline{u}\}} \leq e^{-NE_s(R)}$

where $E_s(R) > 0$ for $R > H(\mathcal{U}|\mathcal{V})$.

Therefore for any N there exists an encoder function f where $\Pr\{\underline{\hat{u}} \neq \underline{u}\} \leq \overline{\Pr\{\underline{\hat{u}} \neq \underline{u}\}} \leq e^{-NE_s(R)}$

7.30.

(a) The minimum average error decision rule is defined by the regions: for $m = 1, 2, \cdots, M$

$$\Lambda_m = \{\underline{y} : Q_m P_N(\underline{y}|\underline{x}_m) > Q_{m'} P_N(\underline{y}|\underline{x}_{m'}) \text{ for all } m' \neq m\}.$$

Assume $\underline{x}_m$ is sent. Then

$$P_{E_m} = \sum_{\underline{y} \notin \Lambda_m} P_N(\underline{y}|\underline{x}_m)$$

But for $\underline{y} \in \Lambda_m$, we have

$$1 \leq \left[\frac{\sum_{m' \neq m} Q_{m'}^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_{m'})^{\frac{1}{1+\rho}}}{Q_m^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_m)^{\frac{1}{1+\rho}}}\right]^\rho$$

and thus

$$Q_m P_{E_m} \leq \sum_{\underline{y}} Q_m^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_m)^{\frac{1}{1+\rho}} \left[\sum_{m' \neq m} Q_{m'}^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_{m'})^{\frac{1}{1+\rho}}\right]^\rho$$

Hence we have

$$P_E = \sum_{m=1}^{M} Q_m P_{E_m}$$

$$\leq \sum_{\underline{y}} \sum_{m=1}^{M} Q_m^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_m)^{\frac{1}{1+\rho}} \left[\sum_{m' \neq m} Q_{m'}^{\frac{1}{1+\rho}} P_N(\underline{y}|\underline{x}_m)^{\frac{1}{1+\rho}}\right]^\rho$$

Averaging this over the code ensemble and using the Jenson's inequality gives the bound:

$$\overline{P_E} \leq \sum_{\underline{y}} \sum_{m=1}^{M} Q_m^{\frac{1}{1+\rho}} \overline{P_N(\underline{y}|\underline{x})^{\frac{1}{1+\rho}}} \left[\left(\sum_{m' \neq m} Q_{m'}^{\frac{1}{1+\rho}}\right) \overline{P_N(\underline{y}|\underline{x})^{\frac{1}{1+\rho}}}\right]^\rho$$

Since $\sum_{m' \neq m} Q_{m'}^{\frac{1}{1+\rho}} \leq \sum_{m'=1}^{M} Q_{m'}^{\frac{1}{1+\rho}}$,

$$\overline{P_E} \leq \left(\sum_{m=1}^{M} Q_m^{\frac{1}{1+\rho}} \right)^{1+\rho} \sum_{\underline{y}} \left(\overline{P_N(\underline{y}|\underline{x})^{\frac{1}{1+\rho}}} \right)^{1+\rho}$$

$$= \left(\sum_{m=1}^{M} Q_m^{\frac{1}{1+\rho}} \right)^{1+\rho} \sum_{\underline{y}} \left(\sum_{\underline{x}} q_N(\underline{x}) P_N(\underline{y}|\underline{x})^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

(b) $$\sum_{\underline{y}} \left(\sum_{\underline{x}} q_N(\underline{x}) P_N(\underline{y}|\underline{x})^{\frac{1}{1+\rho}} \right)^{1+\rho} = \left[\sum_{y} \left(\sum_{x} q(x) P(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right]^N$$

$$= e^{-N E_0(\rho, \underline{q})}, \qquad \text{where } E_0(\rho, \underline{q}) = -\ln \sum_y \left(\sum_x q(x) P(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

Since $Q_m = \prod_{\ell=1}^{L} Q(a_{m\ell})$,

$$\left(\sum_{m=1}^{M} Q_m^{\frac{1}{1+\rho}} \right)^{1+\rho} = \left(\sum_{\underline{a}} Q(\underline{a})^{\frac{1}{1+\rho}} \right)^{1+\rho} = \left(\sum_{\underline{a}} \prod_{\ell=1}^{L} Q(a_\ell)^{\frac{1}{1+\rho}} \right)^{1+\rho}$$

$$= \left[\sum_{i=1}^{A} Q(i)^{\frac{1}{1+\rho}} \right]^{L(1+\rho)} = e^{+L E_s(\rho)}$$

where $E_s(\rho) = (1+\rho) \ln \left\{ \sum_{i=1}^{A} Q(i)^{\frac{1}{1+\rho}} \right\}$

$$\therefore \ \overline{P_E} \leq e^{-N E_0(\rho, \underline{q}) + L E_s(\rho)}$$

(c) $$E_s(0) = \ln \left(\sum_{i=1}^{A} Q(i) \right) = \ln 1 = 0.$$

$$E_s'(\rho) = \ln \left(\sum_{i=1}^{A} Q(i)^{\frac{1}{1+\rho}} \right) - \frac{\sum_{i=1}^{A} Q(i)^{\frac{1}{1+\rho}} \ln Q(i)^{\frac{1}{1+\rho}}}{\sum_{i=1}^{A} Q(i)^{\frac{1}{1+\rho}}}$$

$$= -\sum_{i=1}^{A} Q(i:\rho) \ln Q(i:\rho)$$

where $$Q(i:\rho) = \frac{Q(i)^{\frac{1}{1+\rho}}}{\sum_{j=1}^{A} Q(j)^{\frac{1}{1+\rho}}}, \quad i = 1, 2, \cdots, A.$$

If no $Q(i) = 1$, then $\frac{\partial E_s(\rho)}{\partial \rho} > 0$ for all $\rho > 0$.

Also $Q(i:\rho)\big|_{\rho=0} = Q(i)$ and so $E_s'(\rho)\big|_{\rho=0} = H(\mathcal{U})$.

(d) $\lambda = L/N$. Then $L = \lambda N$ and

$$\overline{P_E} \leq e^{-N[E_0(\rho,\underline{q}) - \lambda E_s(\rho)]}$$

Note that from problem 7.28 we have $E_s''(\rho) > 0$ while $E_0''(\rho,\underline{q}) < 0$ is well known.

Then the exponent $E(\rho,\underline{q}) = E_0(\rho,\underline{q}) - \lambda E_s(\rho)$ has the property:

$$E'(\rho,\underline{q})\Big|_{\rho=0} = E_0'(\rho,\underline{q})\Big|_{\rho=0} - \lambda E_s'(\rho)\Big|_{\rho=0} = I(\underline{q}) - \lambda H(\mathcal{U})$$

By choosing $\underline{q}$ that yields $I(\underline{q}) = C$ we see that

$$E'(\rho,\underline{q}) > 0$$

for some ρ and $\underline{q}$ as long as $\lambda H(\mathcal{U}) < C$.

Hence $E(R) \triangleq \max_{\underline{q}} \max_{0 \leq \rho \leq 1} \{E_0(\rho,\underline{q}) - \lambda E_s(\rho)\} > 0$

for $\lambda H(\mathcal{U}) < C$ and finally we have

$$\overline{P_E} \leq e^{-NE(R)} \xrightarrow[N \to \infty]{} 0$$

— // —

Chapter 8.

8.1

For l^{th} source-user pair: $-\infty < u < \infty, -\infty < v < \infty$

$$Q^{(l)}(u) = \frac{1}{\sqrt{2\pi\sigma_l^2}} e^{-u^2/2\sigma_l^2}, \quad d^{(l)}(u,v) = w_l (u-v)^2 .$$

Choose $p^{(l)}(v)$, for some β_l; $p^{(l)}(v) = \frac{1}{\sqrt{2\pi\beta_l^2}} e^{-v^2/2\beta_l^2}$.

Then $[\lambda^{(l)}(u)]^{-1} = \int_{-\infty}^{\infty} p^{(l)}(v) e^{sd^{(l)}(u,v)} dv = \sqrt{\frac{\alpha_l^2}{\alpha_l^2+\beta_l^2}}\, e^{-u^2/2(\alpha_l^2+\beta_l^2)}$

where $\alpha_l^2 = -1/2w_l s$.

Choose β_l^2 to satisfy $\alpha_l^2+\beta_l^2 = \sigma_l^2$. Then

$$D_s^{(l)} = w_l\left[\frac{\alpha_l^2\beta_l^2}{\alpha_l^2+\beta_l^2} + \left(\frac{\alpha_l^2}{\alpha_l^2+\beta_l^2}\right)^2\sigma_l^2\right] = w_l\sigma_l^2 = \frac{-1}{2s}$$

$$\begin{aligned} R^{(l)}(D_s^{(l)}) &= s D_s^{(l)} + \int_{-\infty}^{\infty} Q^{(l)}(u) \ln \lambda^{(l)}(u)\, du \\ &= -\frac{1}{2} - \int_{-\infty}^{\infty} Q^{(l)}(u) \cdot \left\{ \frac{1}{2} \ln \frac{D_s^{(l)}/w_l}{\sigma_l^2} - \frac{u^2}{2\sigma_l^2} \right\} du \\ &= \frac{1}{2} \ln \frac{w_l\sigma_l^2}{D_s^{(l)}}, \qquad 0 < D_s^{(l)} \le w_l\sigma_l^2 \end{aligned}$$

$$\therefore R^{(l)}(D_s^{(l)}) = \begin{cases} \frac{1}{2} \ln \frac{w_l\sigma_l^2}{D_s^{(l)}} & ; \ 0 < D_s^{(l)} \le w_l\sigma_l^2 \\ 0 & ; \ D_s^{(l)} > w_l\sigma_l^2 \end{cases}$$

with slope

$$s = \frac{d}{dD_s^{(l)}} R^{(l)}(D_s^{(l)}) = \begin{cases} -1/2D_s^{(l)} & ; \ 0 \le D_s^{(l)} \le w_l\sigma_l^2 \\ 0 & ; \ D_s^{(l)} > w_l\sigma_l^2 \end{cases}$$

Hence $D_s^{(l)} = \begin{cases} -1/2s & ; \ -\infty < s \le -1/2w_l\sigma_l^2 \\ w_l\sigma_l^2 & ; \ -1/2w_l\sigma_l^2 < s \le 0 \end{cases}$

or $D_\theta^{(l)} = \min(\theta, w_l\sigma_l^2)$ where $\theta = -1/2s$

and $R^{(l)}(D_\theta^{(l)}) = \max\left(0, \frac{1}{2} \ln \frac{w_l\sigma_l^2}{\theta}\right)$

Therefore for sum distortion measure, the rate distortion function is given in terms of parameter $\theta \ge 0$ as

$$D_\theta = \frac{1}{L}\sum_{l=1}^{L} \min(\theta, w_l\sigma_l^2), \quad R(D_\theta) = \frac{1}{L}\sum_{l=1}^{L} \max\left(0, \frac{1}{2}\ln\frac{w_l\sigma_l^2}{\theta}\right) //$$

8.2

$$D_\theta = \frac{1}{L}\sum_{\ell=1}^{L} \min(\theta, \lambda_\ell), \quad R_L(D_\theta) = \frac{1}{L}\sum_{\ell=1}^{L} \max\left(0, \frac{1}{2}\ln\frac{\lambda_\ell}{\theta}\right)$$

(a) $D < \min\{\lambda_1, \lambda_2, \cdots, \lambda_L\}$. Then

$$\min(\theta, \lambda_\ell) = \theta, \quad \max\left(0, \frac{1}{2}\ln\frac{\lambda_\ell}{\theta}\right) = \frac{1}{2}\ln\frac{\lambda_\ell}{\theta}, \quad \forall \ell = 1, 2, \cdots, L.$$

Therefore $D_\theta = \frac{1}{L}\sum_{\ell=1}^{L}\theta = \theta$, and

$$R_L(D_\theta) = \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{2}\ln\frac{\lambda_\ell}{\theta} = \frac{1}{2L}\ln\prod_{\ell=1}^{L}\frac{\lambda_\ell}{D} = \frac{1}{2L}\ln\frac{|\Phi|}{D^L}$$

(b) $\theta = \max\{\lambda_1, \lambda_2, \cdots, \lambda_L\}$ then

$$\min(\theta, \lambda_\ell) = \lambda_\ell, \quad \max\left(0, \frac{1}{2}\ln\frac{\lambda_\ell}{\theta}\right) = 0, \quad \forall \ell = 1, 2, \cdots, L.$$

Therefore $D_\theta = \frac{1}{L}\sum_{\ell=1}^{L}\lambda_\ell$, and $R_L(D_\theta) = 0$.

And $\frac{1}{L}\sum_{\ell=1}^{L}\lambda_\ell$ is the least value of R for which $R_L(D) = 0$

$$\therefore D_\theta = D_{max}.$$

8.3

$$d(\mathcal{B}) = \int\cdots\int Q_L(\underline{\hat u})\, d(\underline{\hat u}|\mathcal{B})\, d\underline{\hat u} = \int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\, d(\underline{\hat u}|\mathcal{B})\, d\underline{\hat u}\, d\underline{\hat v}$$

since $\int\cdots\int P_L(\underline{\hat v}|\underline{\hat u})\, d\underline{\hat v} = 1$. Define $\Phi(\underline{\hat u}, \underline{\hat v}; \mathcal{B}) = \begin{cases} 1 & : d_L(\underline{\hat u}, \underline{\hat v}) \le d(\underline{\hat u}|\mathcal{B}) \\ 0 & : d_L(\underline{\hat u}, \underline{\hat v}) > d(\underline{\hat u}|\mathcal{B}) \end{cases}$

$$d(\mathcal{B}) = \int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\, d(\underline{\hat u}|\mathcal{B})[1 - \Phi(\underline{\hat u}, \underline{\hat v}; \mathcal{B})]\, d\underline{\hat u}\, d\underline{\hat v}$$
$$+ \int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\, d(\underline{\hat u}|\mathcal{B})\, \Phi(\underline{\hat u}, \underline{\hat v}; \mathcal{B})\, d\underline{\hat u}\, d\underline{\hat v}$$

By definition $d(\underline{\hat u}|\mathcal{B})[1 - \Phi(\underline{\hat u}, \underline{\hat v}; \mathcal{B})] \le d_L(\underline{\hat u}, \underline{\hat v})$,

$$\int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\, d(\underline{\hat u}|\mathcal{B})[1 - \Phi(\underline{\hat u}, \underline{\hat v}; \mathcal{B})]\, d\underline{\hat u}\, d\underline{\hat v}$$
$$\le \int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\, d_L(\underline{\hat u}, \underline{\hat v})\, d\underline{u}\, d\underline{v} = \int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})\left[\frac{1}{L}\sum_{\ell=1}^{L}(\hat u_\ell - \hat v_\ell)^2\right] d\underline{\hat u}\, d\underline{\hat v}$$
$$= \frac{1}{L}\sum_{\ell=1}^{L}\left[\iint Q^{(\ell)}(\hat u)\, p^{(\ell)}(\hat v|\hat u)(\hat v - \hat u)^2\, d\hat u\, d\hat v\right] = \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)}$$

and $\int\cdots\int Q_L(\underline{\hat u}) P_L(\underline{\hat v}|\underline{\hat u})[d(\underline{\hat u}|\mathcal{B})]^2\, d\underline{\hat u}\, d\underline{\hat v} = \int\cdots\int Q_L(\underline{\hat u})[d(\underline{\hat u}|\mathcal{B})]^2\, d\underline{\hat u}$

$$\le \int\cdots\int Q_L(\underline{\hat u})[d_L(\underline{\hat u}, \underline{0})]^2\, d\underline{u} = \int\cdots\int Q_L(\underline{\hat u})\left[\frac{1}{L}\sum_{\ell=1}^{L}\hat u_\ell^2\right]^2 d\underline{\hat u}$$
$$\le \int\cdots\int Q_L(\underline{\hat u})\left[\frac{1}{L}\sum_{\ell=1}^{L}\hat u_\ell^4\right] d\underline{\hat u} = \frac{1}{L}\sum_{\ell=1}^{L}\left[\int \hat u^4 Q^{(\ell)}(\hat u)\, d\hat u\right]$$
$$= \frac{1}{L}\sum_{\ell=1}^{L} 3\lambda_\ell^2 \le 3\left[\max_{1\le\ell\le L}\lambda_\ell\right]^2 = \left[\sqrt{3}\max_{1\le\ell\le L}\lambda_\ell\right]^2 \triangleq d_0^2$$

$$d(\mathcal{B}) \le \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)} + \int\cdots\int Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, d(\hat{\underline{u}}|\mathcal{B})\, \Phi(\hat{\underline{u}},\hat{\underline{v}};\mathcal{B})\, d\hat{\underline{u}}\, d\hat{\underline{v}}$$

$$\le \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)} + \left[\int\cdots\int Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, d^2(\hat{\underline{u}}|\mathcal{B})\, d\hat{\underline{u}}\, d\hat{\underline{v}}\right]^{\frac{1}{2}} \left[\int\cdots\int Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, \Phi(\hat{\underline{u}},\hat{\underline{v}};\mathcal{B})\, d\hat{\underline{u}}\, d\hat{\underline{v}}\right]^{\frac{1}{2}}$$

$$\le \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)} + (d_0^2)^{\frac{1}{2}} \cdot \left[\int\cdots\int Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, \Phi(\hat{\underline{u}},\hat{\underline{v}};\mathcal{B})\, d\hat{\underline{u}}\, d\hat{\underline{v}}\right]^{\frac{1}{2}}$$

Since $\int\cdots\int \overline{Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, \Phi(\hat{\underline{u}},\hat{\underline{v}};\mathcal{B})}\, d\hat{\underline{u}}\, d\hat{\underline{v}} \le e^{-L\, E_L(R,\rho,\underline{P}_L)}$

where $E_L(R,\rho,\underline{P}_L) = -\rho R - \frac{1}{L}\ln \int\cdots\int \left[\int\cdots\int P_L(v)\, Q_L(u|v)^{1/1+\rho}\, d\underline{v}\right]^{1+\rho} d\underline{u}$.

$$\therefore\ \overline{d(\mathcal{B})} \le \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)} + d_0\left[\iint \overline{Q_L(\hat{\underline{u}}) P_L(\hat{\underline{v}}|\hat{\underline{u}})\, \Phi(\hat{\underline{u}},\hat{\underline{v}};\mathcal{B})}\, d\hat{\underline{u}}\, d\hat{\underline{v}}\right]^{1/2}$$

$$\le \frac{1}{L}\sum_{\ell=1}^{L} D^{(\ell)} + d_0\, e^{-(L/2)E_L(R,\rho,\underline{P}_L)}$$

//

8.4

$$\phi_k = E\{u_t u_{t+k}\} = \rho^{|k|}, \qquad 0<\rho<1, \quad k = 0, \pm 1, \pm 2, \cdots$$

$$\Phi(\omega) = \sum_{k=-\infty}^{\infty} \phi_k e^{-jk\omega} = \sum_{k=-\infty}^{\infty} \rho^{|k|} e^{-jk\omega} = \sum_{k=0}^{\infty} \rho^k e^{-jk\omega} + \sum_{k=-\infty}^{0} \rho^{-k} e^{-jk\omega} - 1$$

$$= \frac{1}{1-\rho e^{-j\omega}} + \frac{1}{1-\rho e^{j\omega}} - 1 = \frac{1-\rho^2}{1-2\rho\cos\omega+\rho^2}$$

For small $\theta \le \delta$, i.e. $\min[\theta, \Phi(\omega)] = \theta$, $\forall\, \theta \in [-\pi, \pi]$

$$D_\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min[\theta, \Phi(\omega)]\, d\omega = \theta$$

$$\min_{\omega} \Phi(\omega) = \Phi(\omega)\Big|_{\omega=\pm\pi} = \frac{1-\rho^2}{1+2\rho+\rho^2} = \frac{1-\rho}{1+\rho}$$

$$R(D_\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \max\left[0, \frac{1}{2}\ln\frac{\Phi(\omega)}{\theta}\right] d\omega = \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\frac{\Phi(\omega)}{D_\theta}\, d\omega$$

$$= \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\left[\frac{1-\rho^2}{D_\theta}\cdot\frac{1}{1-2\rho\cos\omega+\rho^2}\right] d\omega$$

$$= \frac{1}{2}\ln\frac{1-\rho^2}{D_\theta} - \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln[1-2\rho\cos\omega+\rho^2]\, d\omega = \frac{1}{2}\ln\frac{1-\rho^2}{D_\theta}$$

$$\therefore\ R(D) = \frac{1}{2}\ln\frac{1-\rho^2}{D}, \quad \text{for } D \le \frac{1-\rho}{1+\rho}$$

//

<u>8-5</u> For any discrete-time zero-mean stationary ergodic source with spectral density $\Phi(\omega)$ and the squared-error distortion measure

$$R(D) = \inf_L R_L(D), \qquad R_L(D) \le \hat{R}_L(D)$$

where $\hat{R}(D)$ is the rate distortion function obtained with the transformation $\hat{\underline{u}} = \underline{u}\Gamma$ where coordinates are independent.

From page 490 we know that $I(\hat{\underline{P}}_L) = I(\underline{P}_L)$, $D(\hat{\underline{P}}_L) = D(\underline{P}_L)$ and $d_L(\hat{\underline{u}},\hat{\underline{v}}) = d_L(\underline{u},\underline{v}) = \frac{1}{L}\sum_{\ell=1}^{L}(\hat{u}_\ell - \hat{v}_\ell)^2$.

For $\hat{\underline{P}}_L \in \mathcal{P}_{D,L} = \{\hat{\underline{P}}_L ;\ D(\hat{\underline{P}}_L) \le D\}$, we have $\hat{R}_L(D) \le \frac{1}{L} I(\hat{\underline{P}}_L)$.

Choose $\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}}) = \prod_{\ell=1}^{L} \frac{1}{\sqrt{2\pi\beta D}} e^{-(\hat{v}_\ell - \beta\hat{u}_\ell)^2/2\beta D}$

where $\beta = 1 - \frac{D}{\sigma^2}$, $\quad \sigma^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi(\omega)d\omega$.

Then $\int\cdots\int d_L(\hat{\underline{u}},\hat{\underline{v}})\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})d\hat{\underline{v}} = \beta D + \frac{D^2}{\sigma^4}\cdot\frac{1}{L}\sum_{\ell=1}^{L}\hat{u}_\ell^2$

and $D(\hat{\underline{P}}_L) = \int\cdots\int \hat{Q}_L(\hat{\underline{u}})\left[\int\cdots\int d_L(\hat{\underline{u}},\hat{\underline{v}})\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})d\hat{\underline{v}}\right]d\hat{\underline{u}}$

$$= \beta D + \frac{D^2}{\sigma^4}\cdot\frac{1}{L}\sum_{\ell=1}^{L}\sigma^2 = D.$$

$\therefore\ \hat{\underline{P}}_L \in \mathcal{P}_{DL}$ and $R(D) \le R_L(D) \le \hat{R}_L(D) \le \frac{1}{L} I(\hat{\underline{P}}_L)$.

But $\frac{1}{L}I(\hat{\underline{P}}) = \frac{1}{L}\int\cdots\int \hat{Q}_L(\hat{\underline{u}})\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})\ln\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})d\hat{\underline{u}}d\hat{\underline{v}} - \frac{1}{L}\int\cdots\int \hat{P}_L(\hat{\underline{v}})\ln\hat{P}_L(\hat{\underline{v}})d\hat{\underline{v}}$

$$= K(D) + h(\hat{\mathcal{V}}_L).$$

where
$$K(D) \triangleq \frac{1}{L}\int\cdots\int \hat{Q}_L(\hat{\underline{u}})\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})\ln\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})d\hat{\underline{v}}d\hat{\underline{u}}$$

$$= \int\cdots\int \hat{Q}_L(\hat{\underline{u}})\hat{P}_L(\hat{\underline{v}}|\hat{\underline{u}})\frac{1}{L}\sum_{\ell=1}^{L}\left[-\frac{1}{2}\ln 2\pi\beta D - \frac{(\hat{v}_\ell - \beta\hat{u}_\ell)^2}{2\beta D}\right]d\hat{\underline{v}}d\hat{\underline{u}}$$

$$= -\frac{1}{2}\ln 2\pi\beta D - \frac{1}{2\beta D}\int\cdots\int Q_L(\hat{\underline{u}})\left\{\frac{1}{L}\sum_{\ell=1}^{L}\beta D\right\}d\hat{\underline{u}}$$

$$= -\frac{1}{2}\ln 2\pi\beta D - \frac{1}{2} = -\frac{1}{2}\ln 2\pi e\beta D.$$

and $h(\hat{\mathcal{V}}_L) \triangleq -\frac{1}{L}\int\cdots\int \hat{P}_L(\hat{\underline{v}})\ln\hat{P}_L(\hat{\underline{v}})d\hat{\underline{v}}$.

Since $\int\cdots\int \hat{\underline{v}}\hat{P}_L(\hat{\underline{v}})d\hat{\underline{v}} = 0$, and

$$\int\cdots\int \hat{\underline{v}}^2 \hat{P}_L(\hat{\underline{v}})\, d\hat{\underline{v}} = \int\cdots\int \hat{Q}_L(\hat{\underline{u}}) \left[\int\cdots\int \hat{\underline{v}}^2 \hat{P}_L(\hat{\underline{v}} | \hat{\underline{u}})\, d\hat{\underline{v}} \right] d\hat{\underline{u}}$$

$$= \int\cdots\int \hat{Q}_L(\hat{\underline{u}}) \left[\beta D + \frac{\beta^2}{L} \sum_{\ell=1}^{L} \hat{u}_\ell^2 \right] d\underline{u} = \beta D + \beta^2 \sigma^2 = \sigma^2 - D \triangleq \alpha^2 .$$

Let $\overline{P}_L(\underline{\hat{v}}) = \prod_{\ell=1}^{L} \frac{1}{\sqrt{2\pi\alpha^2}} e^{-\hat{v}_\ell^2/2\alpha^2}$, then

$$h(\hat{\mathcal{V}}_L) \le h(\overline{\mathcal{V}}_L) = \frac{1}{2} \ln 2\pi e \alpha^2 = \frac{1}{2} \ln 2\pi e [\sigma^2 - D] .$$

$$\therefore R(D) \le \frac{1}{L} I(\underline{P}) = K(D) + h(\hat{\mathcal{V}}_L) \le K(D) + h(\overline{\mathcal{V}}_L)$$

$$= -\frac{1}{2} \ln 2\pi e \beta D + \frac{1}{2} \ln 2\pi e [\sigma^2 - D] = \frac{1}{2} \ln \frac{\sigma^2}{D} .$$

Therefore for discrete-time stationary source,

$$R(D) \le \frac{1}{2} \ln \frac{\sigma^2}{D} \quad //$$

For continuous-time stationary sources where

$$\Phi(\omega) = 0 \quad \text{for } |\omega| > \omega_0 = \frac{\pi}{T_0} = 2\pi B$$

We know $R(D) \le R_G(D)$ where $R_G(D)$ is the rate distortion function for the special case of Gaussian source with same $\Phi(\omega)$ given parametrically.

$$D_\theta = \frac{1}{2\pi} \int_{-\infty}^{\infty} \min[\theta, \Phi(\omega)]\, d\omega = \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} \min[\theta, \Phi(\omega)]\, d\omega$$

and $R(D_\theta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \max\left[0, \ln \frac{\Phi(\omega)}{\theta}\right] d\omega = \frac{1}{4\pi} \int_{-\omega_0}^{\omega_0} \max\left[0, \ln \frac{\Phi(\omega)}{\theta}\right] d\omega$

And when $\Phi(\omega)$ is flat over $|\omega| \le \omega_0$, it gives maximum rate distortion function. (see Berger (1971) p. 139)

Then $\Phi(\omega) = K, \quad |\omega| \le \omega_0$

$$\sigma^2 = \frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} \Phi(\omega)\, d\omega = \frac{2\omega_0 K}{2\pi} = \frac{K}{T_0} = 2KB .$$

For $\theta \le K$, $\quad D_\theta = \frac{2\omega_0}{2\pi} \theta = \frac{1}{T_0} \theta . \qquad \theta = T_0 D_\theta$

and $\quad R(D_\theta) = \frac{2\omega_0}{4\pi} \ln \frac{K}{T_0 D_\theta} = B \ln \frac{\sigma^2}{D_\theta}$

Therefore for band-limited continuous-time stationary source $R(D) \le R_G(D) \le B \ln \frac{\sigma^2}{D}$ //

<u>8.6</u>

$$R_L(D) = \sup_{s\le 0,\, \underline{\lambda}_L \in \Lambda_{s,L}} \left[sD + \frac{1}{L} \int \cdots \int Q_L(\underline{u}) \ln \lambda_L(\underline{u})\, d\underline{u} \right] \qquad (8.3.10)$$

where $\Lambda_{s,L} = \left\{ \lambda_L(\underline{u}) : \int \cdots \int \lambda_L(\underline{u}) Q_L(\underline{u}) e^{s d_L(\underline{u},\underline{v})} d\underline{u} \le 1,\ \underline{v} \in \mathcal{V}_L \right\}$ (8.3.11)

Extend L to NL where $N \ge 1$, integer.

$$R_{NL}(D) = \sup_{s\le 0,\, \underline{\lambda}_{NL} \in \Lambda_{s,NL}} \left[sD + \frac{1}{NL} \int \cdots \int Q_{NL}(\underline{u}) \ln \lambda_{NL}(\underline{u})\, d\underline{u} \right]$$

where

$$\Lambda_{s,NL} = \left\{ \lambda_{NL}(\underline{u}) ; \int \cdots \int \lambda_{NL}(\underline{u}) Q_{NL}(\underline{u}) e^{s d_{NL}(\underline{u},\underline{v})} d\underline{u} \le 1,\ \underline{v} \in \mathcal{V}_{NL} \right\}$$

Now choose $\lambda_{NL}(\underline{u}) = \dfrac{1}{Q_{NL}(\underline{u})} \prod_{n=1}^{N} \lambda_L(\underline{u}_n) Q_L^{(n)}(\underline{u}_n)$

where $\lambda_L(\underline{u}) \in \Lambda_{s,L}$. Then

$$R_{NL}(D) \ge \sup_{s\le 0,\, \underline{\lambda}_L \in \Lambda_{s,L}} \Big[sD + \frac{1}{NL} \int \cdots \int Q_{NL}(\underline{u}) \Big\{ -\ln Q_{NL}(\underline{u}) + \sum_{n=1}^{N} \ln \lambda_L(\underline{u}_n) + \sum_{n=1}^{N} \ln Q_L^{(n)}(\underline{u}_n) \Big\} d\underline{u} \Big]$$

$$= \frac{1}{NL} h(\mathcal{U}_{NL}) - \frac{1}{NL} \cdot N \cdot h(\mathcal{U}_L) + \sup_{s\le 0,\, \underline{\lambda}_L \in \Lambda_{s,L}} \left[sD + \frac{N}{NL} \int \cdots \int Q_L(\underline{u}) \ln \lambda_L(\underline{u})\, d\underline{u} \right]$$

$$= \frac{1}{NL} h(\mathcal{U}_{NL}) - \frac{1}{L} h(\mathcal{U}_L) + R_L(D)$$

$$R(D) = \lim_{N\to\infty} R_{LN}(D) \ge \left\{ \lim_{N\to\infty} \frac{1}{NL} h(\mathcal{U}_{NL}) \right\} - \frac{1}{L} h(\mathcal{U}_L) + R_L(D)$$

$$= h - \frac{1}{L} h(\mathcal{U}_L) + R_L(D) \qquad //$$

<u>8.7</u> From (7.7.41) and (8.3.15)

$$R(D_s) \ge h - h(\mathcal{U}_1) + R_1(D_s)$$

$$\ge h - h(\mathcal{U}_1) + h(\mathcal{U}_1) + sD_s - \ln \int_{-\infty}^{\infty} e^{s d(z)} dz$$

$$= h + sD_s - \ln \int_{-\infty}^{\infty} e^{s d(z)} dz \triangleq R_{LB}(D_s) \qquad //$$

8.8 $Q_L(\underline{u}) = (2\pi|\Phi_L|^{1/L})^{-L/2}\cdot \exp[-\frac{1}{2}\underline{u}\,\Phi_L^{-1}\underline{u}^T]$

Let Γ be the unitary matrix such that $\Phi_L = \Gamma\Lambda\Gamma^T$ with Λ diagonal and $\hat{\underline{u}} = \underline{u}\Gamma$. Then

$$Q_L(\hat{\underline{u}}) = \prod_{\ell=1}^{L}(2\pi\lambda_\ell)^{-1/2}\exp(-\hat{u}_\ell^2/2\lambda_\ell) = \prod_{\ell=1}^{L} Q_\ell(\hat{u}_\ell)$$

where λ_ℓ is the ℓ^{th} eigenvalue of Φ_L. Hence

$$\frac{1}{L}h(\mathcal{U}_L) = \frac{1}{L}h(\hat{\mathcal{U}}_L) = -\frac{1}{L}\int\!\cdots\!\int Q_L(\hat{\underline{u}})\ln Q_L(\hat{\underline{u}})\,d\hat{\underline{u}}$$

$$= -\frac{1}{L}\sum_{\ell=1}^{L}\int Q_\ell(\hat{u}_\ell)\ln Q_\ell(\hat{u}_\ell)\,d\hat{u}_\ell = \frac{1}{L}\sum_{\ell=1}^{L} h(\hat{\mathcal{U}}^{(\ell)})$$

$$= \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{2}\ln(2\pi e\lambda_\ell) = \frac{1}{2}\ln 2\pi e + \frac{1}{2L}\sum_{\ell=1}^{L}\ln\lambda_\ell .$$

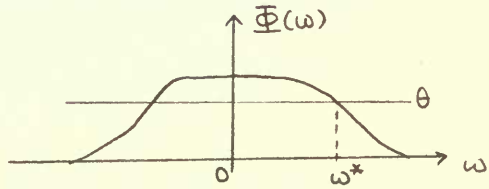
By definition of h, and applying Toeplitz theorem, we have

$$h \triangleq \lim_{L\to\infty}\frac{1}{L}h(\mathcal{U}_L) = \lim_{L\to\infty}\left[\frac{1}{2}\ln 2\pi e + \frac{1}{2L}\sum_{\ell=1}^{L}\ln\lambda_\ell\right]$$

$$= \frac{1}{2}\ln 2\pi e + \frac{1}{2}\cdot\lim_{L\to\infty}\left[\frac{1}{L}\sum_{\ell=1}^{L}\ln\lambda_\ell\right] = \frac{1}{2}\ln 2\pi e + \frac{1}{2}\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\Phi(\omega)\,d\omega$$

$$= \frac{1}{2}\ln 2\pi e E. \quad \text{where } E \triangleq \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln[\Phi(\omega)]\,d\omega\right\} \quad //$$

8.9



$$\sigma^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(\omega)\,d\omega = \frac{A}{\pi}\int_0^{\infty}\frac{d\omega}{1+(\frac{\omega}{\omega_0})^2}$$

$$= \frac{A}{\pi}\int_0^{\pi/2}\frac{\omega_0\sec^2\alpha\,d\alpha}{\sec^2\alpha} = \frac{A\omega_0}{2}$$

where $\alpha = \tan^{-1}\frac{\omega}{\omega_0}$. $\quad \therefore A = \frac{2\sigma^2}{\omega_0}$ //

$$\Phi(\omega^*) = \theta = \frac{A}{1+(\frac{\omega^*}{\omega_0})^2} \Rightarrow 1+(\frac{\omega^*}{\omega_0})^2 = \frac{A}{\theta} = \frac{2\sigma^2}{\omega_0\theta} = 1+\beta^2 \Rightarrow \beta = \frac{\omega^*}{\omega_0}$$

$$D_\theta = \frac{1}{2\pi}\int_{-\infty}^{\infty}\min[\theta,\Phi(\omega)]\,d\omega = \frac{1}{\pi}\left[\theta\omega^* + \int_{\omega^*}^{\infty}\Phi(\omega)\,d\omega\right] \quad \left(\omega=\omega^* \Rightarrow \tan\alpha = \frac{\omega^*}{\omega_0} = \beta\right)$$

$$= \frac{\theta\omega_0\beta}{\pi} + \frac{1}{\pi}\cdot\frac{2\sigma^2}{\omega_0}\int_{\tan^{-1}\beta}^{\pi/2}\frac{\omega_0\sec^2\alpha\,d\alpha}{\sec^2\alpha} = \frac{2\sigma^2}{1+\beta^2}\cdot\frac{\beta}{\pi} + \frac{2\sigma^2}{\pi}\left[\frac{\pi}{2} - \tan^{-1}\beta\right]$$

$$= \sigma^2\left[1 + \frac{2}{\pi}\left(\frac{\beta}{1+\beta^2} - \tan^{-1}\beta\right)\right] \quad //$$

$$\int_0^{\omega^*} \ln\left[1+\left(\frac{\omega}{\omega_0}\right)^2\right] d\omega = \omega_0 \int_0^{\frac{\omega^*}{\omega_0}=\beta} \ln[1+z^2]\, dz$$

$$= \omega_0 \left[z\ln(1+z^2) - 2z + 2\tan^{-1} z \right]\Big|_0^\beta = \omega_0\left\{\beta \ln(1+\beta^2) - 2\beta + 2\tan^{-1}\beta\right\}$$

$$R(D_\theta) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \max\left[0, \ln\frac{\Phi(\omega)}{\theta}\right] d\omega = \frac{1}{2\pi}\int_0^{\omega^*} \ln\left[\frac{2\sigma^2}{\omega_0\theta}\cdot\frac{1}{1+\left(\frac{\omega}{\omega_0}\right)^2}\right] d\omega$$

$$= \frac{1}{2\pi}\left\{\omega^* \ln(1+\beta^2) - \int_0^{\omega^*} \ln\left[1+\left(\frac{\omega}{\omega_0}\right)^2\right] d\omega\right\}$$

$$= \frac{1}{2\pi}\left\{\omega^* \ln(1+\beta^2) - \omega_0\beta\ln(1+\beta^2) + 2\omega_0\beta - 2\omega_0 \tan^{-1}\beta\right\}$$

$$= \frac{\omega_0}{\pi}(\beta - \tan^{-1}\beta) \quad //$$

8.10

$$\underline{u} = (u^{(1)}, u^{(2)}, \cdots, u^{(k)}, \cdots), \quad \underline{v} = (v^{(1)}, v^{(2)}, \cdots, v^{(k)}, \cdots)$$

$$u(t) = \sum_{k=1}^{\infty} u^{(k)} f_k(t), \qquad v(t) = \sum_{k=1}^{\infty} v^{(k)} f_k(t), \qquad 0 \le t \le T.$$

$$\int_0^T f_k(t)\cdot f_j(t) = \delta_{kj}, \qquad j,k = 1,2,\cdots.$$

$$\phi(t,s) = E\{u(t)u(s)\}, \quad 0 \le t,s \le T. \qquad \int_0^T \phi(t,s) f(s)\, ds = \lambda f(t).$$

Gaussian source with squared error criterion:

$$E\{u^{(k)}u^{(j)}\} = \lambda_k \delta_{kj}, \quad j,k = 1,2,\cdots. \qquad d_T(\underline{u},\underline{v}) = \frac{1}{T}\sum_{k=1}^{\infty}(u^{(k)}-v^{(k)})^2.$$

$$Q_T(\underline{u}) = \prod_{k=1}^{\infty} Q^{(k)}(u), \qquad Q^{(k)}(u) = \frac{1}{\sqrt{2\pi\lambda_k}} e^{-u^2/2\lambda_k}$$

$$\mathcal{B} = \{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_M\}. \qquad \Phi(\underline{u},\underline{v};\mathcal{B}) \triangleq \begin{cases} 1 : d_T(\underline{u},\underline{v}) \le d_T(\underline{u}|\mathcal{B}) \\ 0 : d_T(\underline{u},\underline{v}) > d_T(\underline{u}|\mathcal{B}) \end{cases}$$

$$d_T(\mathcal{B}) = \int\cdots\int Q_T(\underline{u}) d_T(\underline{u}|\mathcal{B})\, d\underline{u} = \int\cdots\int Q_T(\underline{u}) P_T(\underline{v}|\underline{u})\, d_T(\underline{u}|\mathcal{B})\, d\underline{u}\, d\underline{v}$$

$$= \int\cdots\int Q_T(\underline{u}) P_T(\underline{v}|\underline{u})\, d_T(\underline{u}|\mathcal{B})\left[1 - \Phi(\underline{u},\underline{v};\mathcal{B}) + \Phi(\underline{u},\underline{v};\mathcal{B})\right] d\underline{u}\, d\underline{v}.$$

$$\int\cdots\int Q_T(\underline{u}) P_T(\underline{v}|\underline{u}) d_T(\underline{u}|\mathcal{B})\left[1 - \Phi(\underline{u},\underline{v};\mathcal{B})\right] d\underline{u}\, d\underline{v}$$

$$\le \int\cdots\int Q_T(\underline{u}) P_T(\underline{v}|\underline{u})\, d_T(\underline{u},\underline{v})\, d\underline{u}\, d\underline{v} = \int\cdots\int Q_T(\underline{u}) P_T(\underline{v}|\underline{u})\left[\frac{1}{T}\sum_{k=1}^{\infty}(u^{(k)}-v^{(k)})^2\right] d\underline{u}\, d\underline{v}$$

$$= \frac{1}{T}\sum_{k=1}^{\infty}\left[\iint Q^{(k)}(u) p^{(k)}(v|u)(u-v)^2\, du\, dv\right] \triangleq \frac{1}{T}\sum_{k=1}^{\infty} D^{(k)}$$

$$\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})[d_T(\underline{u}|\mathcal{B})]^2 d\underline{u}d\underline{v} = \int\cdots\int Q_T(\underline{u})[d_T(\underline{u}|\mathcal{B})]^2 d\underline{u}$$

$$\le \int\cdots\int Q_T(\underline{u}) d_T^2(\underline{u},\underline{0})d\underline{u} = \int\cdots\int Q_T(\underline{u})\left[\frac{1}{T}\sum_{k=1}^{\infty} u^{(k)^2}\right]^2 d\underline{u} \le \int\cdots\int Q_T(\underline{u})\left(\frac{1}{T}\sum_{k=1}^{\infty} u^{(k)^4}\right)d\underline{u}$$

$$= \frac{1}{T}\sum_{k=1}^{\infty}\left[\int u^4 Q^{(k)}(u)\,du\right] = \frac{1}{T}\sum_{k=1}^{\infty} 3\lambda_k^2 \triangleq d_{0,T}^2$$

$$\overline{\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})\,\Phi(\underline{u},\underline{v};\mathcal{B})\,d\underline{u}d\underline{v}}$$

$$\le \int\cdots\int\left[\overline{\int\cdots\int P_T(\underline{v})\,\Phi(\underline{u},\underline{v};\mathcal{B})d\underline{v}}\right]^{-\rho}\cdot\left[\int\cdots\int P_T(\underline{v})Q_T(\underline{u}|\underline{v})^{1/1+\rho}d\underline{v}\right]^{1+\rho}d\underline{u}$$

$$\le M^{\rho}\int\cdots\int\left[\int\cdots\int P_T(\underline{v})Q_T(\underline{u}|\underline{v})^{1/1+\rho}d\underline{v}\right]^{1+\rho}d\underline{u}$$

$$= \exp\left\{-T\left[-\rho R - \frac{1}{T}\ln\left\{\int\cdots\int\left[\int\cdots\int P_T(\underline{v})Q_T(\underline{u}|\underline{v})^{1/1+\rho}d\underline{v}\right]d\underline{u}\right\}\right]\right\}$$

$$\triangleq \exp\{-T\cdot E_T(R,\rho,\underline{P}_T)\}$$

$$d_T(\mathcal{B}) \le \frac{1}{T}\sum_{k=1}^{\infty} D^{(k)} + \left[\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})d_T^2(\underline{u}|\mathcal{B})d\underline{u}d\underline{v}\right]^{\frac{1}{2}}\left[\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})\,\Phi(\underline{u},\underline{v};\mathcal{B})d\underline{u}d\underline{v}\right]^{\frac{1}{2}}$$

$$\le \frac{1}{T}\sum_{k=1}^{\infty} D^{(k)} + d_{0,T}\cdot\left[\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})\,\Phi(\underline{u},\underline{v};\mathcal{B})d\underline{u}d\underline{v}\right]^{\frac{1}{2}}$$

$$\overline{d_T(\mathcal{B})} \le \frac{1}{T}\sum_{k=1}^{\infty} D^{(k)} + d_{0,T}\left[\overline{\int\cdots\int Q_T(\underline{u})P_T(\underline{v}|\underline{u})\,\Phi(\underline{u},\underline{v};\mathcal{B})d\underline{u}d\underline{v}}\right]^{\frac{1}{2}}$$

$$\le \frac{1}{T}\sum_{k=1}^{\infty} D^{(k)} + d_{0,T}\cdot\exp\left\{-\frac{T}{2}\cdot E_T(R,\rho,\underline{P}_T)\right\}$$

Using same arguement as (8.2.78) ~ (8.2.82), we have

$$\overline{d_T(\mathcal{B})} \le \frac{1}{T}\sum_{k=1}^{\infty}\min(\theta,\lambda_k) + d_{0,T}\cdot\exp\left\{-\frac{T}{2}\left(-\rho R - \frac{1}{T}\sum_{k=1}^{\infty}\max\left(0,\frac{\rho}{2}\ln\left[\frac{\theta(1+\rho)}{\lambda_k+\rho\theta}\right]\right)\right)\right\}$$

$T\to\infty$, and using Toeplitz theorem.

$$\lim_{T\to\infty}\frac{1}{T}\sum_{k=1}^{\infty}\min(\theta,\lambda_k) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\min(\theta,\Phi(\omega))d\omega = D_\theta.$$

$$\lim_{T\to\infty}\frac{-1}{T}\sum_{k=1}^{\infty}\max\left(0,\frac{\rho}{2}\ln\left[\frac{\theta(1+\rho)}{\lambda_k+\rho\theta}\right]\right) = \frac{-1}{4\pi}\int_{-\infty}^{\infty}\max\left\{0,\rho\ln\left[\frac{\theta(1+\rho)}{\Phi(\omega)+\rho\theta}\right]\right\}d\omega \triangleq E_\infty(\rho,\theta)$$

$$E(R,D_\theta) \triangleq \max_{-1\le\rho\le 0}\left[-\rho R + E_\infty(\rho,\theta)\right]. \qquad d_0 = \left[\frac{3}{2\pi}\int_{-\infty}^{\infty}[\Phi(\omega)]^2 d\omega\right]^{\frac{1}{2}}$$

$\therefore$ We have that for each $\epsilon_1 > 0$, $\epsilon_2 > 0$, there exists an into $N_0(\theta, R, \epsilon_1, \epsilon_2)$ such that for each $T \geq T_0$ there exists a block code $\mathcal{B}$ of rate R and duration T such that

$$d(\mathcal{B}) \leq D_\theta + \epsilon_1 + d_0 \exp\left\{-\frac{T}{2}[E(R, D_\theta) - \epsilon_2]\right\}$$

where $E(R, D_\theta) > 0$ for $R > R(D_\theta)$.

8.11.

$$d(\mathcal{B}) = \sum_{\underline{u}} Q_N(\underline{u}) \, d(\underline{u}|\mathcal{B})$$

Define indicator function $\Phi(\underline{u}; \mathcal{B}) = \begin{cases} 1 & d(\underline{u}|\mathcal{B}) > D \\ 0 & d(\underline{u}|\mathcal{B}) \leq D \end{cases}$

Using the facts: $(1-\Phi) + \Phi = 1$, $d(\underline{u}|\mathcal{B}) \leq d_0$, $d(\underline{u}|\mathcal{B})[1-\Phi] \leq D$;

$$d(\mathcal{B}) = \sum_{\underline{u}} Q_N(\underline{u}) d(\underline{u}|\mathcal{B})[1 - \Phi(\underline{u}; \mathcal{B})] + \sum_{\underline{u}} Q_N(\underline{u}) d(\underline{u}|\mathcal{B}) \Phi(\underline{u}; \mathcal{B})$$

$$\leq D \sum_{\underline{u}} Q_N(\underline{u}) + d_0 \cdot \sum_{\underline{u}} Q_N(\underline{u}) \Phi(u; \mathcal{B})$$

Averaging this over ensemble of codes yields

$$\overline{d(\mathcal{B})} \leq D + d_0 \overline{\sum_{u} Q_N(\underline{u}) \Phi(\underline{u}; \mathcal{B})} \leq D + d_0 \sum_{\underline{u}} Q_N(\underline{u}) \sum_{\mathcal{B}} P(\mathcal{B}) \Phi(u; \mathcal{B})$$

$$= D + d_0 \sum_{\underline{u}} Q_N(\underline{u}) \Pr\{d(\underline{u}|\mathcal{B}) > D \,|\, \underline{u}\} \leq D + d_0 e^{-\exp N[R - R(D) + o(N)]}$$

//

8.12

Define: $F(\epsilon) \triangleq \mathcal{H}(\epsilon) + \epsilon \ln \frac{\epsilon}{2} = -(1-\epsilon)\ln(1-\epsilon) - \epsilon \ln 2$

$$F'(\epsilon) = +\ln(1-\epsilon) + \frac{1-\epsilon}{1-\epsilon} - \ln 2 = \ln(1-\epsilon) + 1 - \ln 2$$

$$F''(\epsilon) = \frac{-1}{1-\epsilon} < 0 \qquad \forall \epsilon: 0 < \epsilon < 1.$$

Since $F''(\epsilon) < 0$, $F(\epsilon)$ is convex $\cap$.

$F(0) = 0$, $F(.3) = -(.7)\ln(.7) - (.3)\ln 2 \simeq 0.042$

$\therefore$ We know from convexity that $F(\epsilon) > 0$ for $0 < \epsilon < .3$

i.e. $\mathcal{H}(\epsilon) > -\epsilon \ln \frac{\epsilon}{2}$ for $0 < \epsilon < .3$ //

8.13.

We know $\lim_{N\to\infty} \Pr\{ | \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} - Nq | \le N\gamma \} = 1$ for any $\gamma > 0$.

Equivalently $\lim_{N\to\infty} \sum_{|\ell - Nq| \le N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} = 1$

and $\lim_{N\to\infty} \sum_{|\ell - Nq| > N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} = 0$.

Since $D_\ell \le 1$ for all $\ell = 0, 1, \cdots, N$,

$$\lim_{N\to\infty} \sum_{|\ell-Nq|>N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} D_\ell \le \lim_{N\to\infty} \sum_{|\ell-Nq|>N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} = 0.$$

$\therefore \lim_{N\to\infty} \sum_{|\ell-Nq|>N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} D_\ell = 0$ because no term is negative.

And since $D_\ell \le D_{N(q+\gamma)}$ for all $\ell : |\ell - Nq| \le N\gamma$

$$\lim_{N\to\infty} \sum_{|\ell-Nq|\le N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} D_\ell \le D_{N(q+\gamma)} \cdot \lim_{N\to\infty} \sum_{|\ell-Nq|\le N\gamma} \binom{N}{q} q^{\ell}(1-q)^{N-\ell} = D_{N(q+\gamma)}$$

But we can choose $\gamma > 0$ as small as we please when N goes infinity. Hence

$$\lim_{N\to\infty} \sum_{|\ell-Nq|\le N\gamma} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} D_\ell = D_{Nq} = D.$$

$$\therefore \lim_{N\to\infty} \sum_{\ell=0}^{N} \binom{N}{\ell} q^{\ell}(1-q)^{N-\ell} D_\ell = D \quad //$$

8.14.

(a) L_N^A = # of the distinct composition of output sequences of length N from source alphabet of size A.

$= L_{N-1}^{A} + L_{N}^{A-1}$. $\quad \forall A, N = 2, 3, \cdots$

$L_N^1 = 1 = (N+1)^{1-1}$, $\quad L_1^A = A \le (1+1)^{A-1}$

$$L_N^2 = L_{N-1}^2 + L_N^1 = L_{N-2}^2 + L_{N-1}^1 + L_N^1 = \cdots = L_1^2 + L_2^1 + \cdots + L_{N-1}^1 + L_N^1$$
$$= 2 + (N-1) = N+1 = (N+1)^{2-1}$$

$$L_2^A = L_1^A + L_2^{A-1} = L_1^A + L_1^{A-1} + L_2^{A-2} = \cdots = L_1^A + L_1^{A-1} + \cdots + L_1^2 + L_2^1$$

$$= A + (A-1) + \cdots + 2 + 1 = \frac{1}{2}A(A+1) \leq (2+1)^{A-1}$$

Suppose $L_N^{A-1} \leq (N+1)^{A-2}$, $L_{N-1}^A \leq N^{A-1}$, then

$$L_N^A = L_N^{A-1} + L_{N-1}^A \leq (N+1)^{A-2} + N^{A-1}.$$

Want to show $(N+1)^{A-2} + N^{A-1} \leq (N+1)^{A-1}$.

$$(N+1)^{A-1} - N^{A-1} - (N+1)^{A-2} = (N+1)(N+1)^{A-2} - (N+1)^{A-2} - N \cdot N^{A-2}$$

$$= N \cdot (N+1)^{A-2} - N(N)^{A-2} = N\{(N+1)^{A-2} - N^{A-2}\} \geq 0.$$

$$\therefore \quad L_N^A \leq (N+1)^{A-1}.$$

(b) For each $\mathcal{C}_N(l)$ consider an ensemble of block codes $\mathcal{B}_l = \{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_M\}$ of block length N and rate $R = \frac{1}{N}\ln M$ chosen with probability

$$P(\mathcal{B}_l) = \prod_{m=1}^{M} P_N(\underline{v}_m) = \prod_{m=1}^{M}\prod_{n=1}^{N} p^{(l)}(v_{mn})$$

where $p^{(l)}(v) = \sum_u p^{(l)}(v|u)Q^{(l)}(u)$

and $p^{(l)}(v|u)$ is the conditional probability yielding the rate distortion function $R(D_l : Q^{(l)}) = I(\underline{P}^{(l)})$.

<u>Lemma 8.5.2′</u>

Let an ensemble of codes be selected according to $p^{(l)}(v)$. Then there exists an integer N^* such that for any $N \geq N^*$

$$Pr\{d(\underline{u}|\mathcal{B}_l) > D_l \mid \underline{u} \in \mathcal{C}_N(l)\} \leq \exp\{-\exp[N(\delta - o(N))]\}$$

where $o(N)$ goes to zero as fast as $\sqrt{N^{-1}\ln N}$.

<u>Proof</u>: Same as (8.5.39) we have

$$Pr\{d(\underline{u}|\mathcal{B}_l) > D_l \mid \underline{u} \in \mathcal{C}_N(l)\} \leq \exp[-M\, Pr\{d(\underline{u}|\mathcal{B}) \leq D_l \mid \underline{u} \in \mathcal{C}_N(l)\}]$$

But $Pr\{d(\underline{u}|\mathcal{B}_l) \leq D_l \mid \underline{u} \in \mathcal{C}_N(l)\} \geq \exp[-N[R(D_l : Q^{(l)}) + o(N)]]$

(see Martin [1976], Appendix B) for every $N \geq N^*$.

$\therefore \Pr\{ d(\underline{u}|\mathcal{B}_\ell) > D_\ell \mid \underline{u} \in \mathcal{C}_N(\ell)\}$

$$\leq \exp\{-\exp(NR)\cdot \exp[-N\cdot\{R(D_\ell : Q^{(\ell)}) + o(N)\}]\}$$

$$= \exp[-\exp\{N[R - R(D_\ell : Q^{(\ell)}) - o(N)]\}]$$

$$= \exp(-\exp\{N[\delta - o(N)]\})$$

Q.E.D.

Theorem 8.5.2'

There exists an integer N^* such that for any $N \geq N^*$ and any composition $C_N(\ell)$ there exists a code $\mathcal{B}_\ell$ with rate R such that $d(\underline{u}|\mathcal{B}_\ell) \leq D_\ell$ for all $\underline{u} \in \mathcal{C}_N(\ell)$ where D_ℓ is defined to be $R = R(D_\ell : Q^{(\ell)}) + \delta$.

proof:

Consider the function $\Phi(\underline{u}|\mathcal{B}_\ell) \triangleq \begin{cases} 1 & d(\underline{u}|\mathcal{B}_\ell) > D_\ell \\ 0 & d(\underline{u}|\mathcal{B}_\ell) \leq D_\ell \end{cases}$

Then

$$\overline{\sum_{\underline{u}\in\mathcal{C}_N(\ell)} \Phi(\underline{u}|\mathcal{B}_\ell)} = \sum_{\underline{u}\in\mathcal{C}_N(\ell)} \sum_{\mathcal{B}} P(\mathcal{B})\, \Phi(\underline{u}|\mathcal{B})$$

$$= \sum_{\underline{u}\in\mathcal{C}_N(\ell)} \Pr\{d(\underline{u}|\mathcal{B}_\ell) > D_\ell \mid \underline{u}\in\mathcal{C}_N(\ell)\} \leq |\mathcal{C}_N(\ell)|\cdot \exp(-\exp\{N[\delta - o(N)]\})$$

The last inequality comes from above lemma. Since $|\mathcal{C}_N(\ell)|$ is upperbounded by A^N, the # of elements in $\mathcal{U}_N$,

$$\overline{\sum_{\underline{u}\in\mathcal{C}_N(\ell)} \Phi(\underline{u}|\mathcal{B}_\ell)} \leq A^N \exp(-\exp\{N[\delta - o(N)]\})$$

$$= \exp(N \ln A - \exp\{N[\delta - o(N)]\})$$

Choose N^* such that $\exp\{N^*[\delta - o(N^*)]\} > N^* \ln A$

then $\overline{\sum_{\underline{u}\in\mathcal{C}_N(\ell)} \Phi(\underline{u}|\mathcal{B}_\ell)} < 1$ for all $N \geq N^*$.

$\therefore$ There is at least one code $\mathcal{B}_\ell$ for which $\sum_{\underline{u}\in\mathcal{C}_N(\ell)} \Phi(\underline{u}|\mathcal{B}_\ell) < 1$.

And this implies that $\Phi(\underline{u}|\mathcal{B}_\ell) = 0$ for all $\underline{u}\in\mathcal{C}_N(\ell)$, or equivalently $d(\underline{u}|\mathcal{B}_\ell) \leq D_\ell$ for all $\underline{u}\in\mathcal{C}_N(\ell)$.

Q.E.D.

(c) For $N \geq N^*$, a code satisfying $d(\underline{u}|\mathcal{B}_\ell) \leq D_\ell$ can be found for each composition class $C_N(\ell)$. Using these L_N codes, a composite code $\mathcal{B}_c$ can be defined by

$$\mathcal{B}_c \triangleq \bigcup_{\ell=1}^{L_N} \mathcal{B}_\ell$$

This is a code of block length N with $L_N e^{NR}$ code words and rate $\quad R_c = R + \frac{1}{N} \ln L_N$

From (a) $\quad \ln L_N \leq (A-1)\ln(N+1) \qquad \therefore R_c \leq R + (A-1)\cdot \frac{\ln(N+1)}{N}$

Thus by choosing N large enough, the rate of composite code R_c can be made arbitrary close to R.

8.15.

(a)
$$P_N(R,D) = \min_{\mathcal{B}} \Pr\{d(\underline{u}|\mathcal{B}) > D \,|\, \mathcal{B}\} \leq \Pr\{d(\underline{u}|\mathcal{B}) > D\}$$

$$= \sum_{\underline{u}} Q'(\underline{u}) \Pr\{d(\underline{u}|\mathcal{B}) > D \,|\, \underline{u}\}.$$

From Lemma 8.5.1. we have

$$\Pr\{d(\underline{u}|\mathcal{B}) > D \,|\, \underline{u}\} \leq e^{-\exp. N[R - R(D;\underline{Q}') + o(N)]}$$

Therefore

$$P_N(R,D) \leq \sum_{\underline{u}} Q'(\underline{u})\, e^{-\exp N[R - R(D;\underline{Q}') + o(N)]}$$

$$\leq e^{-\exp N[R - \max_{\underline{Q}'} R(D;\underline{Q}') + o(N)]}$$

when $R > \max_{\underline{Q}'} R(D;\underline{Q}')$

$$F(R,D) = -\lim_{N\to\infty} \frac{1}{N} \ln P_N(R,D)$$

$$\leq \lim_{N\to\infty} \frac{1}{N} \exp N[R - \max_{\underline{Q}'} R(D;\underline{Q}') + o(N)] = \infty$$

$$\therefore F(R,D) = \infty.$$

(b) From Prob. 8.14 we know for $R(D) < R < \max_{\underline{Q}'} R(D;\underline{Q}')$

$$\{\underline{u} \in \mathcal{U}_N ; d(\underline{u}|\mathcal{B}) > D\} \subset \bigcup_{l=1}^{L_N} \{\underline{u} \in \mathcal{C}_N(l) : R(D) < R(D:\underline{Q}^{(l)})\}$$

$$P_N(R,D) = \min_{\mathcal{B}} Pr\{d(\underline{u}|\mathcal{B}) > D|\mathcal{B}\} \leq Pr\{\underline{u} \in \mathcal{U}_N ; d(\underline{u}|\mathcal{B}) > D\}$$

$$\leq Pr\{\underline{u} \in \mathcal{C}_N(l) : R(D) < R(D:\underline{Q}^{(l)}),\ l = 1, 2, \cdots, L_N\}$$

$$\leq \sum_{l:\ R(D) < R(D:\underline{Q}^{(l)})} Pr\{\underline{u} \in \mathcal{C}_N(l)\}$$

$$\leq \sum_{l:\ R(D) < R(D:\underline{Q}^{(l)})} \exp\left[-N J(\underline{Q}^{(l)}, \underline{Q}) + o(N)\right]$$

$$F(R,D) \geq -\lim_{N\to\infty} \frac{1}{N} \ln \left[\sum_{l:\ R(D) < R(D:\underline{Q}^{(l)})} e^{-N J(\underline{Q}^{(l)}, \underline{Q}) + o(N)}\right]$$

$$\geq -\lim_{N\to\infty} \frac{1}{N} \ln \left[L_N \cdot e^{-N \min_{\tilde{\underline{Q}}} J(\tilde{\underline{Q}}, \underline{Q}) + o(N)}\right]$$

$$\geq \lim_{N\to\infty} \left\{-(A-1)\frac{\ln(N+1)}{N} + \min_{\tilde{\underline{Q}}} J(\tilde{\underline{Q}}, \underline{Q}) + o(N)\right\}$$

$$= \min_{\tilde{\underline{Q}}} J(\tilde{\underline{Q}}, \underline{Q})$$

(c) Converse coding theorem: $R < R(D, \tilde{Q}) \Rightarrow \tilde{d}(\mathcal{B}) > D$.

$$\tilde{d}(\mathcal{B}) = \sum_{\underline{u}} \tilde{Q}(\underline{u})\, d(\underline{u}|\mathcal{B})$$

$$= \sum_{\underline{u};\ d(\underline{u}|\mathcal{B}) \leq D} \tilde{Q}(\underline{u})\, d(\underline{u}|\mathcal{B}) + \sum_{u:\ d(\underline{u}|\mathcal{B}) > D} \tilde{Q}(\underline{u})\, d(\underline{u}|\mathcal{B})$$

$$\leq D \cdot \left[1 - \tilde{Pr}\{d(\underline{u}|\mathcal{B}) > D|\mathcal{B}\}\right] + d_0 \cdot \tilde{Pr}\{d(\underline{u}|\mathcal{B}) > D|\mathcal{B}\}$$

$$= D + (d_0 - D) \cdot \tilde{Pr}\{d(\underline{u}|\mathcal{B}) > D|\mathcal{B}\}$$

$$\therefore\ \tilde{Pr}\{d(\underline{u}|\mathcal{B}) > D|\mathcal{B}\} \geq \frac{\tilde{d}(\mathcal{B}) - D}{d_0 - D} \triangleq \alpha > 0.$$

where $d_0 = \max_{\underline{u} \in \mathcal{U}_N,\ \underline{v} \in \mathcal{V}_N} d_N(\underline{u}, \underline{v}) > D$. and $\tilde{d}(\mathcal{B}) > D$.

(d) Define $X(\mathcal{B}, D) = \{\underline{u} \in \mathcal{U}_N : d(\underline{u}|\mathcal{B}) > D\}$

$$\Pr\{d(\underline{u}|\mathcal{B}) > D \mid \mathcal{B}\} = \sum_{X(\mathcal{B},D)} Q_N(\underline{u}) \geq \sum_{X(\mathcal{B},D)\cap G_\gamma} Q_N(\underline{u})$$

$$\geq e^{-N[J(\tilde{\underline{Q}},\underline{Q})+\gamma]} \sum_{X(\mathcal{B},D)\cap G_\gamma} \tilde{Q}_N(\underline{u})$$

$$= e^{-N[J(\tilde{\underline{Q}},\underline{Q})+\gamma]} \cdot \left\{ \sum_{X(\mathcal{B},D)} \tilde{Q}_N(\underline{u}) - \sum_{X(\mathcal{B},D)\cap G_\gamma^c} \tilde{Q}_N(\underline{u}) \right\}$$

$$\geq e^{-N[J(\tilde{\underline{Q}},\underline{Q})+\gamma]} \cdot \left\{ \sum_{X(\mathcal{B},D)} \tilde{Q}_N(\underline{u}) - \sum_{G_\gamma^c} \tilde{Q}_N(\underline{u}) \right\}$$

$$\geq e^{-N[J(\tilde{\underline{Q}},\underline{Q})+\gamma]} \left\{ \alpha - \frac{\sigma^2}{N\gamma^2} \right\}$$

(e) For any $\gamma > 0$ and any $\tilde{\underline{Q}}$ that satisfies $R < R(D;\tilde{\underline{Q}})$ and $R(D) < R < \max_{Q'} R(D;\underline{Q}')$,

$$F(R,D) = -\lim_{N\to\infty} \frac{1}{N} \ln P_N(R,D)$$

$$\leq -\lim_{N\to\infty} \frac{1}{N} \left\{ \ln\left(\alpha - \frac{\sigma^2}{N\gamma^2}\right) - N[J(\tilde{\underline{Q}},\underline{Q}) + \gamma] \right\}$$

$$= J(\tilde{\underline{Q}},\underline{Q}) + \gamma$$

Since $\gamma > 0$ can be choose arbitrary small and $\tilde{\underline{Q}}$ to minimize $J(\tilde{\underline{Q}},\underline{Q})$, we have upper bound

$$F(R,D) \leq \min_{\tilde{\underline{Q}}} J(\tilde{\underline{Q}},\underline{Q}).$$

Together with lower bound in (b), finally we have

$$F(R,D) = \min_{\tilde{\underline{Q}}} J(\tilde{\underline{Q}},\underline{Q}).$$

<u>8.16.</u>

This is a Hamming 1 perfect $(2^m, 2^m-1-m)$ code with $m=3$. (See Prob. 2.11 and Fig. 2.17 for generator matrix G and parity-check matrix H).

$$|S_{\frac{1}{n}}| = \binom{n}{0} + \binom{n}{1} = 1+n$$

$$\therefore \quad nD = \frac{1\cdot\binom{n}{1}}{|S_{\frac{1}{n}}|} = \frac{n}{n+1} \qquad \text{or} \qquad D = \frac{1}{n+1}$$

In this case $\quad n=7 \qquad$ or $\qquad D = \frac{1}{8} = 0.125.$

$$R(D) = \ln 2 - \mathcal{H}(D) = \ln 2 + \frac{1}{8}\ln\frac{1}{8} + \frac{7}{8}\ln\frac{7}{8} = 0.316$$

$$R = \frac{4}{7}\ln 2 = 0.396.$$

8.17

$$p = Q(1,0) + Q(1,1) \quad \Rightarrow \quad Q(1,0) = p - Q(1,1)$$

$$\hat{p} = Q(1,1) + Q(0,1) \quad \Rightarrow \quad Q(0,1) = \hat{p} - Q(1,1)$$

$$\sum_u \sum_{\hat{u}} Q(u,\hat{u})\,d(u,\hat{u}) = Q(1,0) + Q(0,1) = p + \hat{p} - 2Q(1,1)$$

$$\begin{aligned}
\therefore \bar{d}(p,\hat{p}) &= \min_Q \sum_u \sum_{\hat{u}} Q(u,\hat{u})\,d(u,\hat{u}) \\
&= \min_{0 \le Q(1,1) \le \min(p,\hat{p})} \{p + \hat{p} - 2Q(1,1)\} \\
&= p + \hat{p} - 2\cdot\min(p,\hat{p}) \\
&= |p - \hat{p}|
\end{aligned}$$

8.18

For all joint distributions $\{Q(u,\hat{u}) : u,\hat{u} \in \mathcal{U}\}$ where

$$Q(u) = \sum_{\hat{u}} Q(u,\hat{u}) \qquad \forall\, u \in \mathcal{U}$$

and

$$\hat{Q}(\hat{u}) = \sum_u Q(u,\hat{u}) \qquad \forall\, \hat{u} \in \mathcal{U}$$

define the distance between the two sources as

$$\bar{d}(\mathcal{S},\hat{\mathcal{S}}) = \min_Q \sum_u \sum_{\hat{u}} Q(u,\hat{u})\,d(u,\hat{u})$$

Let $\mathcal{B}$ be any block code of length N and let $\underline{u} \in \mathcal{U}_N$ be an output sequence of length N from $\mathcal{S}$ and $\underline{\hat{u}} \in \mathcal{U}_N$ from $\hat{\mathcal{S}}$. Let $\underline{v}(\underline{\hat{u}}) \in \mathcal{B}$ satisfies

$$d_N(\underline{\hat{u}}, \underline{v}(\underline{\hat{u}})) = \min_{\underline{v} \in \mathcal{B}} d_N(\underline{\hat{u}}, \underline{v})$$

Then $\min_{\underline{v} \in \mathcal{B}} d_N(\underline{u}, \underline{v}) \le d_N(\underline{u}, \underline{v}(\underline{\hat{u}})) \le d_N(\underline{u}, \underline{\hat{u}}) + d_N(\underline{\hat{u}}, \underline{v}(\underline{\hat{u}}))$

where the second inequality is the triangular inequality.

Therefore we have

$$\min_{\underline{v}} d_N(\underline{u}, \underline{v}) \le d_N(\underline{u}, \underline{\hat{u}}) + \min_{\underline{v}} d_N(\underline{\hat{u}}, \underline{v})$$

and by symmetry

$$\min_{\underline{v}} d_N(\underline{\hat{u}}, \underline{v}) \le d_N(\underline{u}, \underline{\hat{u}}) + \min_{\underline{v}} d_N(\underline{u}, \underline{v})$$

$$\therefore \left| \min_{\underline{v}} d_N(\underline{u}, \underline{v}) - \min_{\underline{v}} d_N(\underline{\hat{u}}, \underline{v}) \right| \le d_N(\underline{u}, \underline{\hat{u}})$$

Averaging this with respect to $Q(u, \hat{u})$, we have

$$|d(\mathcal{B}|\mathcal{S}) - d(\mathcal{B}|\hat{\mathcal{S}})| \le \bar{d}(\mathcal{S}, \hat{\mathcal{S}}) \qquad //$$

Since $R(D; \cdot)$ is a strictly decreasing function of D,

$$d(\mathcal{B}|\hat{\mathcal{S}}) \le d(\mathcal{B}|\mathcal{S}) + \bar{d} \Rightarrow R(D : \underline{\hat{Q}}) \ge R(D + \bar{d} : \underline{Q})$$

$$d(\mathcal{B}|\hat{\mathcal{S}}) \ge d(\mathcal{B}|\mathcal{S}) - \bar{d} \Rightarrow R(D; \underline{\hat{Q}}) \le R(D - \bar{d} : \underline{Q})$$

$$\therefore R(D + \bar{d}(\mathcal{S}, \hat{\mathcal{S}}) ; \underline{Q}) \le R(D; \underline{\hat{Q}}) \le R(D - \bar{d}(\mathcal{S}, \hat{\mathcal{S}}) ; \underline{Q}) \qquad //$$

— // —

NOTES

NOTES

NOTES

NOTES

ASSIGNMENT OF RIGHTS

The McGraw-Hill Companies, Inc. ("**McGraw-Hill**"), with offices at 1221 Avenue of the Americas, New York, New York 10020, for good and valuable consideration, receipt of which is hereby acknowledged, does hereby assign and transfer to Andrew J. Viterbi and Jim K. Omura (the "**Assignees**"), without warranty, all rights, title and interest of McGraw-Hill, in and to the copyright of the work published by McGraw-Hill under the title

***Principles of Digital Communication and Coding* by Andrew J. Viterbi and Jim K. Omura**
Copyright Registration Number TX 274-294 (1979)

pursuant to an agreement between the Assignees and McGraw-Hill, dated July 22, 1976. McGraw-Hill retains the right to sell (including by remainder) any existing inventory of the work.

Further, McGraw-Hill assigns to the Assignees all rights of renewal it may possess and all McGraw-Hill's interest, if any, in renewal copyrights that may be secured under the laws now or hereafter in force and effect in the United States or in any other country or countries of the world for the above work. This assignment is subject to any outstanding third party licenses.

In Witness Whereof, this assignment has been duly executed on this 28th day of September, 2005.

The McGraw-Hill Companies, Inc.
By:

Gerald W. Saykes
Director of Editorial Operations
McGraw-Hill Higher Education

0-07-067517-1