5.1 To clarify the problem, let me state it differently. The DMS has source alphabet $\mathcal{X} = \{0,1,2,3,4\}$, each of which is equally likely. At the destination, we have destination alphabet $\hat{\mathcal{X}} = \mathcal{X}$. The single letter distortion matrix is

$$D = \begin{bmatrix} 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \end{bmatrix}$$

We have a source code of length 2 with 5 codewords (0,0), (1,3), (2,1), (3,4), and (4,2). This means we encode each of the 25 possible pairs of source letters as one of these 5 codewords, and then we must represent the source pair by the codeword at the destination. In order to do this with minimum average distortion, we must find the minimum distortion codeword corresponding to each of the 25 source pairs.

Recall that if $\underline{u} = (u_1, u_2)$ and $\underline{v} = (v_1, v_2)$, then

$$d(\underline{u},\underline{v}) = d(u_1, v_1) + d(u_2, v_2)$$

Then we can construct the following minimum distortion encoding by inspection.

$\underline{u}$	$\underline{v} = f(\underline{u})$	$d(\underline{u}, f(\underline{u}))$	$\underline{u}$	$\underline{v} = f(\underline{u})$	$d(\underline{u}, f(\underline{u}))$
00	00	0	30	34	1
01	00	1	31	21	1
02	42	1	32	42	1
03	13	1	33	34	1
04	00	1	34	34	0
10	00	1	40	00	1
11	21	1	41	42	1
12	13	1	42	42	0
13	13	0	43	42	1
14	13	1	44	34	1
20	21	1			
21	21	0			
22	21	1			
23	13	1			
24	34	1			

$u_2 \backslash u_1$	0	1	2	3	4
0	x	•	•	•	•
1	•	•	x	•	•
2	•	•	•	•	x
3	•	x	•	•	•
4	•	•	•	x	•

x = 0 dist. in rep.
• = 1 dist in rep

The resulting average distortion of the code per source letter, $d(C)$ is

$$d(C) = \frac{1}{2} E\left(d(\underline{u}, f(\underline{u}))\right) = \frac{1}{2} \sum_{\underline{u}} p(\underline{u})\, d(\underline{u}, f(\underline{u}))$$

$$= \frac{1}{2} \cdot \frac{1}{25}\left[5 \cdot 0 + 20 \cdot 1\right] = \frac{20}{50} = \boxed{\frac{2}{5}}$$

2. Lets compute the minimum and maximum values of D we must consider δ_{min} & δ_{max}

$$\delta_{min} = \sum_{u \in A_u} p(u) \cdot \min_{v \in A_v} d(u,v) = 0$$

$$\delta_{max} = \min_{v \in A_v} \sum_{u \in A_u} p(u)\, d(u,v) = 1$$

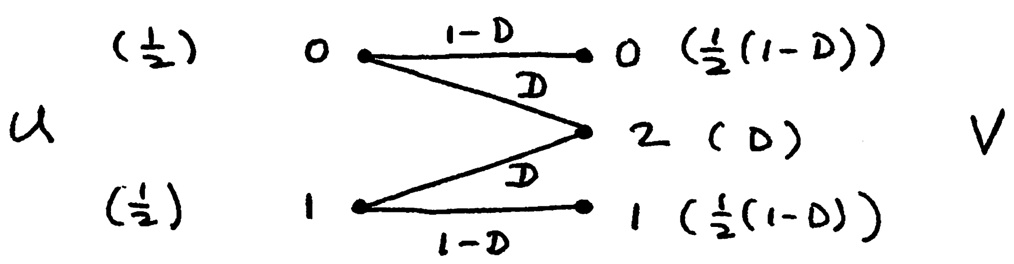
by inspection for

$$\mathbb{D} = \begin{pmatrix} 0 & \infty & 1 \\ \infty & 0 & 1 \end{pmatrix},$$

where $A_u = \{0, 1\}$ and $A_v = \{0, 1, 2\}$

In order to find $R(D)$, we must find an optimal test channel. The presence of infinite distortions in the matrix $\mathbb{D}$ forces $p(v=1/u=0) = p(v=0|u=1) = 0$. Otherwise the average distortion would be infinite.

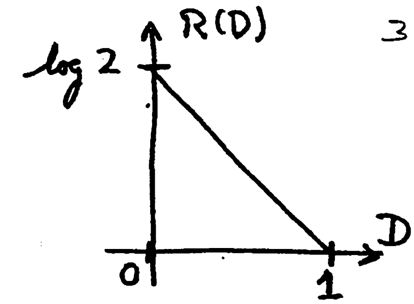
By symmetry (Obvious?, Maybe, but see McEliece prob. 3.6) we have the following test channel



Then

$$R(D) = I(U;V) = H(V) - H(V|U)$$
$$= H\left(\tfrac{1-\delta}{2}, \tfrac{1-\delta}{2}, \delta\right) - \mathcal{H}_2(\delta)$$
$$= \frac{(1-\delta)}{2}\log\left(\frac{2}{1-\delta}\right) + \frac{(1-\delta)}{2}\log\left(\frac{2}{1-\delta}\right) - \delta\log\delta + \delta\log\delta + (1-\delta)\log(1-\delta)$$
$$= (1-\delta)\log 2$$

$$\therefore \; R(D) = \begin{cases} (1-D)\log 2, & 0 \le D \le 1 \\ 0, & D > 1 \end{cases}$$



A plot of $R(D)$ is shown at the right.

3. Again, lets compute δ_{min}, the minimum possible value of D for the given source statistics ($p_u = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$) and the given distortion metric, specified by

$$\mathbb{D} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

$$\delta_{min} = \sum_{u \in A_u} p(u) \cdot \min_{v \in A_v} d(u,v) = \frac{1}{3}$$

We can find $R(D)$ using either a <u>forward test channel</u> (FTC) or a <u>backward test channel</u> (BTC).

We will use a forward test channel using the "symmetry" argument found in prob. 3.6 of McEliece. (For a BTC argument, see prob 3.7 of McEliece).

We must find row permutations π of $\mathbb{D}$ and column permutations ρ of $\mathbb{D}$ such that, taken together

$$\mathbb{D}(u,v) = \mathbb{D}(\pi(u), \rho(v))$$

Because each row has a different number of "1"s, only the identity permutation π' is valid for the rows. For this row permutation, we have only one permutation ρ' (besides the identity) of the columns that is valid, the ρ' exchanging the first two columns of $\mathbb{D}$. This combination of permutations (π', ρ') yields

$$\mathbb{D}(u,v) = \mathbb{D}(\pi'(u), \rho'(v)), \quad \forall u \in A_u, \ \forall v \in A_v$$

According to the result of problem 3.6 of McEliece, there exists a test channel achieving $R(D)$ having a transition matrix $Q(u,v) = p(v|u)$ such that for the permutations (π', ρ') determined above

$$Q(u,v) = Q(\pi'(u), \rho'(v)), \quad \forall u \in A_u, \ \forall v \in A_v$$

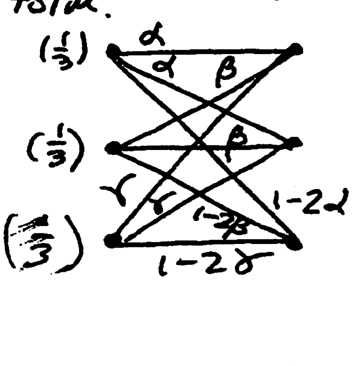
(This is just a mathematical way of saying that if relabeling the inputs and outputs doesn't change their distortion, it shouldn't change the test channel that achieves the minimum $I(U;V)$.)

This implies that the first two columns of the matrix Q can be interchanged without effecting Q.

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix} \xRightarrow{(\pi', \rho')} \begin{array}{l} q_{11} = q_{12} = \alpha \\ q_{21} = q_{22} = \beta \\ q_{31} = q_{32} = \gamma \end{array}$$

So we have restricted the test channel to the following form.

$$Q = \begin{pmatrix} \alpha, & \alpha, & 1-2\alpha \\ \beta, & \beta, & 1-2\beta \\ \gamma, & \gamma, & 1-2\gamma \end{pmatrix}$$



For the source statistics $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, this test channel, and the distortion matrix D_1 we have an average distortion

$$D = \frac{1}{3}\cdot\alpha\cdot 0 + \frac{1}{3}\cdot\alpha\cdot 0 + \frac{1}{3}(1-2\alpha)\cdot 1$$
$$+ \frac{1}{3}\beta\cdot 1 + \frac{1}{3}\cdot\beta\cdot 1 + \frac{1}{3}(1-2\beta)\cdot 0$$
$$+ \frac{1}{3}\gamma\cdot 1 + \frac{1}{3}\cdot\gamma\cdot 1 + \frac{1}{3}\cdot(1-2\gamma)\cdot 1$$
$$= \frac{1}{3}(1-2\alpha) + \frac{2}{3}\beta$$

Now for $\delta_{min} = D = \frac{1}{3} \Rightarrow \frac{1}{3} = \frac{1}{3}(1-2\alpha) + \frac{2}{3}\beta$

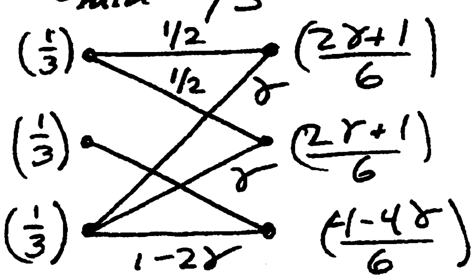
$$\Rightarrow \beta = \alpha - \frac{1}{2} \text{ for } D = \frac{1}{3}$$

Since $\beta \geq 0 \Rightarrow \alpha \geq \frac{1}{2}$, and from row 1 of Q, $\alpha + \alpha \leq 1 \Rightarrow \alpha \leq \frac{1}{2}$

$$\therefore \alpha = \frac{1}{2} \Rightarrow \beta = 0.$$

So now our channel is restricted to the following form for $D = \delta_{min} = 1/3$

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \gamma & \gamma & 1-2\gamma \end{pmatrix}$$



So the channel that achieves $R(1/3)$ is of the following form, with γ selected so that $I(U;V)$ is minimized

$$I(U;V) = H(V) - H(V/U) = H\left(\frac{2\gamma+1}{6}, \frac{2\gamma+1}{6}, \frac{4-4\gamma}{6}\right) - \frac{1}{3}\log 2$$
$$- \frac{1}{3}H(\gamma, \gamma, 1-2\gamma)$$

It can be shown that $\gamma = 1/4$ minimizes $I(U;V)$, and thus $R(\delta_{min}) = R(1/3) = I(U;V)\big|_{\gamma=1/4} = \frac{2}{3}\log 2 = 0.667$ bits.

4. Cover and Thomas, Ch. 10, Prob. 1 (Ch. 13, Prob. 1 in 1st Edition)

Let $X \sim N[0, \sigma^2]$ and the distortion measure be squared error. With one bit quantization, the decision regions are the positive and negative real axes, and the reconstruction points $\pm a$ are the centroids of the corresponding regions.

For the decision region $[0, \infty)$, the centroid a is

$$a = \int_0^\infty x \frac{2}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} dx \quad \text{let } y = \frac{x^2}{2\sigma^2}, \frac{dy}{dx} = \frac{x}{\sigma^2}.$$

change of variable

$$= \int_0^\infty \sigma\sqrt{\frac{2}{\pi}} e^{-y} dy = \sigma\sqrt{\frac{2}{\pi}}$$

The expected distortion for this quantizer is

$$D = E[d(X,\hat{X})] = \int_{-\infty}^{0} \left(x + \sigma\sqrt{\frac{2}{\pi}}\right)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

$$+ \int_0^\infty \left(x - \sigma\sqrt{\frac{2}{\pi}}\right)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

$$= 2\int_{-\infty}^{\infty} \left(x^2 + \sigma^2 \frac{2}{\pi}\right) \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

$$- 2\int_0^\infty \left(-2x\sigma\sqrt{\frac{2}{\pi}}\right) \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

$$= \sigma^2 + \frac{2}{\pi}\sigma^2 - 4\frac{1}{\sqrt{2\pi}}\sigma^2\sqrt{\frac{2}{\pi}}$$

$$= \sigma^2\left(\frac{\pi - 2}{\pi}\right)$$

5. (Cover and Thomas, prob. 13.5)

X is uniformly distributed on the set $\{1, 2, \ldots, m\}$

The distortion measure is

$$d(x, \hat{x}) = \begin{cases} 0, & x = \hat{x} \\ 1, & x \neq \hat{x} \end{cases}$$

Consider any joint distribution $p(x, \hat{x})$ that satisfies

$$E[d(X, \hat{X})] \leq D$$

Because $D = P(\{X \neq \hat{X}\})$, we can apply Fano's Inequality and get

$$H(X|\hat{X}) \leq \mathcal{H}(D) + D \log(m-1)$$

and hence

$$\begin{aligned} I(X; \hat{X}) &= H(X) - H(X|\hat{X}) \\ &\geq \log m - H(D) - D \log(m-1) \end{aligned}$$

We can achieve this lower bound by using a test channel (Forward Test Channel) with

$$p(\hat{x}|x) = \begin{cases} 1 - D, & \hat{x} = x \\ D/(m-1), & \hat{x} \neq x \end{cases}$$

It is easy to verify that this satisfies the bound with equality for $D < 1 - \frac{1}{m}$. Hence

$$R(D) = \begin{cases} \log m - \mathcal{H}(D) - D \log(m-1), & 0 \leq D < 1 - \frac{1}{m} \\ 0, & D > 1 - \frac{1}{m}. \end{cases}$$

6. We assume that X has zero-mean and variance σ^2. To prove the lower bound, we use the same approach used in class to compute the Gaussian rate distortion function. Let $(X,\hat{X})$ be RVs such that $E[(X-\hat{X})^2] \le D$. Then

$$
\begin{aligned}
I(X;\hat{X}) &= h(X) - h(X|\hat{X}) \\
&= h(X) - h(X-\hat{X}|\hat{X}) \\
&\ge h(X) - h(X-\hat{X}) \\
&\ge h(X) - h(N[0, E[(X-\hat{X})^2]]) \\
&= h(X) - \tfrac{1}{2}\log\left[2\pi e\, E[(X-\hat{X})^2]\right] \\
&\ge h(X) - \tfrac{1}{2}\log\left[2\pi e D\right].
\end{aligned}
$$

To prove the upper bound, we consider the joint distribution described by the following test channel

$$
\begin{tikzcd}
X \arrow[r] & \oplus \arrow[r] & \odot \arrow[r] & \hat{X} = \frac{\sigma^2-D}{\sigma^2}(X+Z) \\
& Z \sim \frac{\sigma^2-D}{\sigma^2} N\left[0, \frac{D\sigma^2}{\sigma^2-D}\right] \arrow[u] & \frac{\sigma^2-D}{\sigma^2} \arrow[u]
\end{tikzcd}
$$

$$
\begin{aligned}
E[(X-\hat{X})^2] &= E\left[\left(\frac{D}{\sigma^2}X - \frac{\sigma^2-D}{\sigma^2}Z\right)^2\right] \\
&= \left(\frac{D}{\sigma^2}\right)^2 EX^2 + \left(\frac{\sigma^2-D}{\sigma^2}\right)^2 E[Z^2] \\
&= \left(\frac{D}{\sigma^2}\right)\sigma^2 + \left(\frac{\sigma^2-D}{\sigma^2}\right)\frac{D\sigma^2}{\sigma^2-D} \\
&= D,
\end{aligned}
$$

because X and Z are both zero-mean and statistically independent.

Furthermore, for this case

$$I(X;\hat{X}) = h(\hat{X}) - h(\hat{X}|X)$$
$$= h(\hat{X}) - h\left(\frac{\sigma^2 - D}{\sigma^2} Z\right).$$

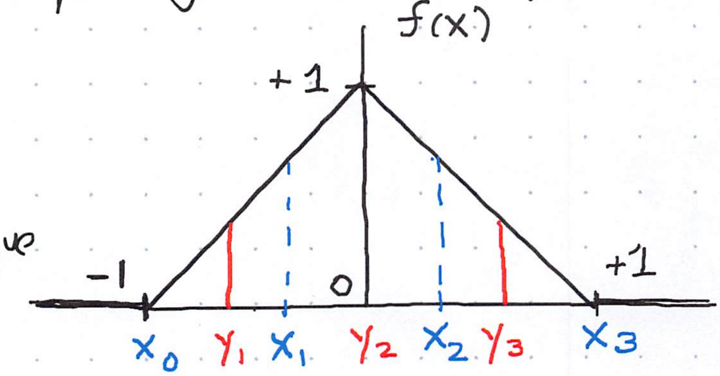
Now

$$E[\hat{X}^2] = \left(\frac{\sigma^2 - D}{\sigma^2}\right)^2 E[(X+Z)^2]$$
$$= \left(\frac{\sigma^2 - D}{\sigma^2}\right)^2 (E[X^2] + E[Z^2])$$
$$= \left(\frac{\sigma^2 - D}{\sigma^2}\right)^2 \left(\sigma^2 + \frac{D\sigma^2}{\sigma^2 - D}\right)$$
$$= \sigma^2 - D.$$

Hence we have

$$I(X;\hat{X}) = h(\hat{X}) - h\left(\frac{\sigma^2 - D}{\sigma^2} Z\right)$$
$$= h(\hat{X}) - h(Z) - \log\left(\frac{\sigma^2 - D}{\sigma^2}\right)$$
$$\leq h(\mathcal{N}[0, \sigma^2 - D]) - \frac{1}{2}\log\left[(2\pi e)\frac{D\sigma^2}{\sigma^2 - D}\right] - \log\left(\frac{\sigma^2 - D}{\sigma^2}\right)$$
$$= \frac{1}{2}\log\left[2\pi e(\sigma^2 - D)\right] - \frac{1}{2}\log\left[(2\pi e)\frac{D\sigma^2}{\sigma^2 - D}\right] - \frac{1}{2}\log\left(\frac{\sigma^2 - D}{\sigma^2}\right)^2$$
$$= \frac{1}{2}\log\frac{\sigma^2}{D}.$$

7. Here we have three reconstruction levels $\{y_1, y_2, y_3\}$ three decision regions, and four decision levels $\{x_0, x_1, x_2, x_3\}$. Superimposing these on a plot of the pdf, we have:



Now from the two rules for the Lloyd-Max quantizer algorithm, we have

$$y_j = \frac{\int_{x_{j-1}}^{x_j} x\, f(x)\, dx}{\int_{x_{j-1}}^{x_j} f(x)\, dx} \quad \text{---} \quad (1)$$

and

$$x_j = \frac{y_{j+1} + y_j}{2} \quad \text{---} \quad (2)$$

$$\Rightarrow \quad y_{j+1} = 2x_j - y_j. \quad \text{---} \quad (2')$$

From the symmetry and form of the pdf $f(x)$, it is apparent that:

$$\begin{cases} x_0 = -1 \\ x_4 = +1 \end{cases} \quad \begin{cases} x_1 = -a \\ x_2 = +a \end{cases} \quad \begin{cases} y_1 = -b \\ y_3 = +b \end{cases} \quad \text{and } y_2 = 0.$$

So we must use the above relations, (1) and (2') to find the values of a and b.

From (2′), we have

$$y_3 = 2x_2 - y_2 \Rightarrow b = 2a - 0 \Rightarrow \boxed{b = 2a},$$

and from (1), we have

$$y_3 = \frac{\int_{x_2}^{x_3} x f(x)\,dx}{\int_{x_2}^{x_3} f(x)\,dx} = \frac{\int_a^1 x(1-x)\,dx}{\int_a^1 (1-x)\,dx} = \frac{2a+1}{3}$$

$$\therefore \quad \left.\begin{aligned} b &= 2a \\ 3b &= 2a+1 \end{aligned}\right\} \Rightarrow \begin{aligned} &6a - 2a = 1 \Rightarrow 4a = 1 \Rightarrow \boxed{a = 1/4} \\ &\text{and} \\ &b = 2a \Rightarrow \boxed{b = 1/2} \end{aligned}$$

Thus we have

$$x_0 = -1, \quad x_1 = -\frac{1}{4}, \quad x_2 = \frac{1}{4}, \quad x_3 = 1$$

$$y_1 = -\frac{1}{2}, \quad y_2 = 0, \quad y_3 = \frac{1}{2}$$

The resulting mean-square error is (using symmetry)

$$D = 2\int_0^{1/4} f(x)(x-0)^2\,dx + 2\int_{1/4}^1 f(x)\left(x - \tfrac{1}{2}\right)^2 dx$$

$$= 2\int_0^{1/4} (1-x)x^2\,dx + 2\int_{1/4}^1 (1-x)\left(x-\tfrac{1}{2}\right)^2 dx$$

$$= \cdots = \frac{5}{192} = 0.0260417$$