ECE 642 Homework #4 Solutions M.R. Bell

1. (a) (5, 1, 2) Binary Repetition Code

$$R = \frac{k}{n} = \frac{1}{5} = 0.2$$

$$P_e = 1 - \sum_{k=0}^{2} \binom{5}{k} \varepsilon^k (1-\varepsilon)^{5-k} = 1 - \left[(1-\varepsilon)^5 + 5\varepsilon(1-\varepsilon)^4 + 10\varepsilon(1-\varepsilon)^3\right]$$

$$= 10\varepsilon^3 - 15\varepsilon^4 + 6\varepsilon^5$$

(b) (7, 4, 1) Hamming Code

$$R = \frac{k}{n} = \frac{4}{7} \cong 0.5714$$

$$P_e = 1 - \sum_{k=0}^{1} \binom{7}{k} \varepsilon^k (1-\varepsilon)^{7-k} = 1 - \left[(1-\varepsilon)^7 + 7\varepsilon(1-\varepsilon)^6\right]$$

$$= 21\varepsilon^2 - 70\varepsilon^3 + 105\varepsilon^4 - 84\varepsilon^5 + 35\varepsilon^6 - 6\varepsilon^7$$

(c) (23, 12, 3) Binary Golay Code

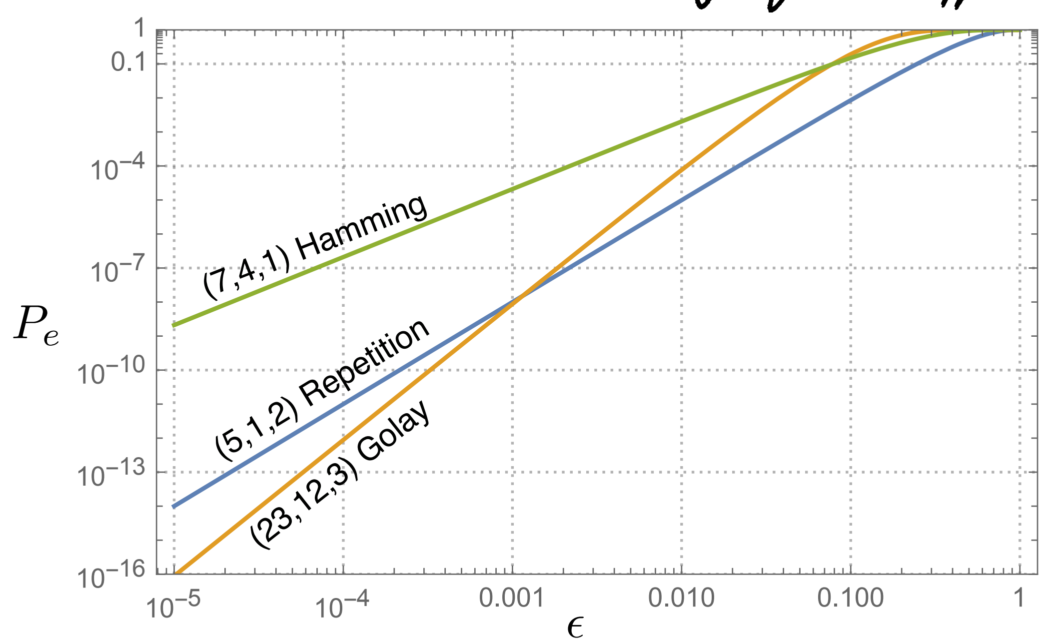
$$R = \frac{12}{23} \cong 0.5217$$

$$P_e = 1 - \sum_{k=0}^{3} \binom{23}{k} \varepsilon^k (1-\varepsilon)^{23-k} = \text{Thank You Mathematica} \ldots$$

$$= 1 - (1-\varepsilon)^{23} - 23\varepsilon(1-\varepsilon)^{22} - 253\varepsilon^2(1-\varepsilon)^{21} - 1771\varepsilon^3(1-\varepsilon)^{20}$$

$$\begin{aligned} = {} & 8855\varepsilon^4 - 134596\varepsilon^5 + 1009470\varepsilon^6 - 4903140\varepsilon^7 \\ & + 17160990\varepsilon^8 - 45762640\varepsilon^9 + 96101544\varepsilon^{10} \\ & - 162249360\varepsilon^{11} + 223092870\varepsilon^{12} - 251694520\varepsilon^{13} \\ & + 233716340\varepsilon^{14} - 178474296\varepsilon^{15} + 111546435\varepsilon^{16} \\ & + 56530320\varepsilon^{17} + 22881320\varepsilon^{18} - 7225680\varepsilon^{19} \\ & + 1716099\varepsilon^{20} - 288420\varepsilon^{21} + 30590\varepsilon^{22} - 1540\varepsilon^{23} \end{aligned}$$

Plots of these block error probabilities as a function of ε for $10^{-5} \le \varepsilon \le 1$ on a log-log scale appear as follows:



2.(a) A $(2t+1, 1, t)$ binary repetition code can correct any error pattern of t or fewer errors out of $2t+1$ transmitted binary digits. Thus

$$P_e = 1 - \sum_{r=0}^{t} \binom{2t+1}{r} \varepsilon^r (1-\varepsilon)^{2t+1-r}$$

(b) A $(k(2t+1), k, s)$ extended binary repetition code is used to transmit k binary information digits using $k(2t+1)$ binary codeword digits. Each codeword is formed by concatenating k $(2t+1, 1, t)$ binary repetition code codewords, one for each of the k binary information digits to be encoded.

Note that this code can correct any pattern of t errors out of $k(2t+1)$ transmitted digits. Note, however, that it cannot correct all patterns of $t+1$ errors, since if $t+1$ errors (for example) occur in the first block of $2t+1$ binary digits, an error will be made. $\therefore$ $s = t$.

In order for a $(k(2t+1), k, t)$ extended repetition code to correct t or less errors, all k sub-blocks of $2t+1$ binary digits (each of these is a $(2t+1, 1, t)$ binary repetition code) must be correctly decoded.

Let $P_e(t,k) \triangleq$ Prob. of block error in $(k(2t+1), k, t)$ extended repetition code.

Then

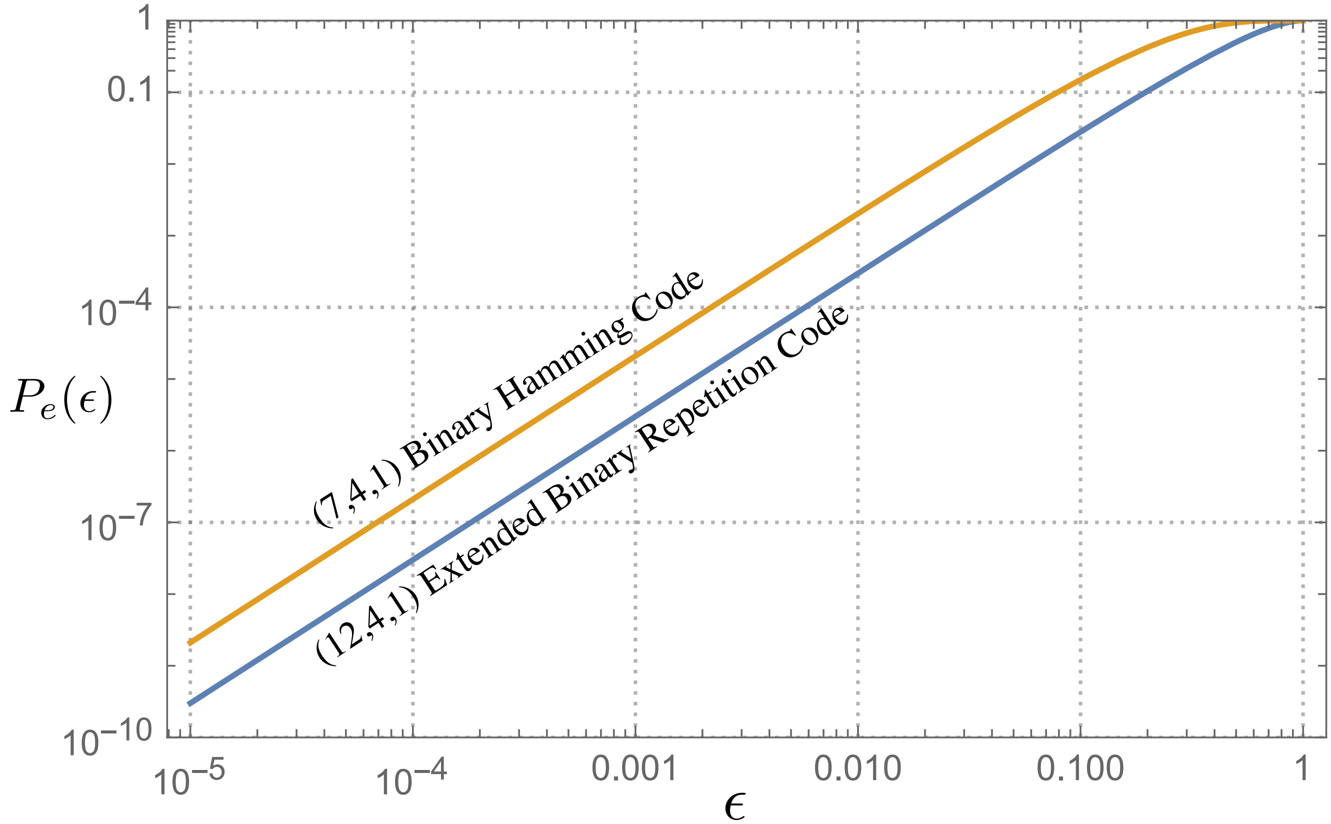
$$P_e(t,k) = 1 - [1 - P_e(t,1)]^k$$
$$= 1 - \left[\sum_{r=0}^{t} \binom{2t+1}{r} \varepsilon^r (1-\varepsilon)^{2t+1-r} \right]^k$$

For the purpose of comparison with a $(7,4,1)$ Hamming code, if we want our extended repetition code to send 4 information bits, and correct any 1 error, then we should take $k=4$ and $t=1$. So we are comparing a $(12,4,1)$ extended binary repetition code to a $(7,4,1)$ Hamming code.

<u>(7,4,1) Hamming:</u> $R = \frac{k}{n} = \frac{4}{7}$, $P_e = 21\varepsilon^2 - 70\varepsilon^3 + 105\varepsilon^4 - 84\varepsilon^5 + 35\varepsilon^6 - 6\varepsilon^7$

<u>(12,4,1) Repetition:</u> $R = \frac{4}{12} = \frac{1}{3}$, $P_e = 1 - (1-\varepsilon)^3 - 3(1-\varepsilon)^2\varepsilon$
$= 3\varepsilon^2 - 2\varepsilon^3$

Plots of the error performance of the two codes are given on the following page.



1
0.1
10^{-4}
10^{-7}
10^{-10}
$P_e(\epsilon)$
10^{-5}
10^{-4}
0.001
0.010
0.100
1
ϵ
(7,4,1) Binary Hamming Code
(12,4,1) Extended Binary Repetition Code

3. (a) $h(X,Y) - h(X) - h(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{1}{f(x,y)}\, dx\, dy$

$$- \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)}\, dx - \int_{-\infty}^{\infty} f(y) \log \frac{1}{f(y)}\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{f(x)f(y)}{f(x,y)}\, dx\, dy \overset{J}{\leq} \log \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \frac{f(x)f(y)}{f(x,y)}\, dx\, dy \right)$$

$$= \log \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)\, dx\, dy \right) = \log(1) = 0.$$

Since $\log(\cdot)$ is strictly convex $\cap$, equality occurs iff $\frac{f(x)f(y)}{f(x,y)} = \text{const.} = 1 \Leftrightarrow f(x,y) = f(x)f(y)$

$\Leftrightarrow$ X and Y are statistically independent

$\therefore h(X,Y) \leq h(X) + h(Y)$, with equality $\Leftrightarrow X \perp\!\!\!\perp Y$.

(b) Show $h(X) \geq h(X|Y) \Leftrightarrow h(X|Y) - h(X) \leq 0$.

$$h(X|Y) - h(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{1}{f(x|y)}\, dx\, dy - \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)}\, dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{f(x)}{f(x|y)}\, dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \log \frac{f(x)f(y)}{f(x,y)}\, dx\, dy$$

$$\overset{J}{\leq} \log \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \frac{f(x)f(y)}{f(x,y)}\, dx\, dy \right) = \log \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)\, dx\, dy \right)$$

$= \log 1 = 0$, with equality iff $\frac{f(x)f(y)}{f(x,y)} = \text{const} = 1$

$\Leftrightarrow$ X and Y are statistically independent

$\therefore h(X) \geq h(X|Y)$, with equality iff $X \perp\!\!\!\perp Y$.

(c) $h(Y,Z|X) - h(Y|X) - h(Z|X) = \iiint_{\mathbb{R}^3} f(x,y,z) \log \frac{f(y|x)f(z|x)}{f(y,z|x)}\, dx\, dy\, dz$

$$\overset{J}{\leq} \log \left(\iiint_{\mathbb{R}^3} f(x,y,z) \frac{f(y|x)f(z|x)f(x)}{f(x,y,z)}\, dx\, dy\, dz \right)$$

$$= \log \left(\iiint_{\mathbb{R}^3} f(y|x)f(z|x)f(x)\, dx\, dy\, dz \right) = \log \left(\int_{-\infty}^{\infty} f(x) \left[\int_{-\infty}^{\infty} f(y|x)\, dy \right] \left[\int_{-\infty}^{\infty} f(z|x)\, dz \right] dx \right)$$

$= \log \left(\int_{-\infty}^{\infty} f(x) \cdot 1 \cdot 1\, dx \right) = \log(1) = 0$, with equality iff

$\frac{f(y,z|x)}{f(y|x)f(z|x)} = 1 \Leftrightarrow f(y,z|x) = f(y|x)f(z|x) \Leftrightarrow$ X and Y are stat. indep conditioned on Z

$(X|Z \perp\!\!\!\perp Y|Z)$.

4. (a) $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

$$h(X) = -\int_0^\infty \lambda e^{-\lambda x} \log(\lambda e^{-\lambda x})\,dx \overset{\text{Integration by parts}}{=} \cdots = \log e - \log\lambda = \log\frac{e}{\lambda}$$
$$= 1 - \ln\lambda \text{ (nats)}$$

(b) $f(x) = \begin{cases} \dfrac{1}{\pi\sqrt{x(1-x)}}, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases} \Rightarrow h(X) = \int_0^1 \frac{1}{\pi\sqrt{x(1-x)}} \log\left[\pi\sqrt{x(1-x)}\right]dx$

This is most easily evaluated using a trigonometric substitution

$\sqrt{ax^2+bx+c} = \sqrt{x(1-x)} = \sqrt{x-x^2} \Rightarrow a=-1,\ b=1,\ c=0$

From calculus, if you have an expression of the form $\sqrt{-a}\sqrt{\beta^2-(x-\alpha)^2}$ (which we have with $\alpha = 1/2$ and $\beta = 1/2$), we can make the substitution

$$x-\alpha = \beta\sin\theta \Rightarrow \sqrt{\beta^2-(x-\alpha)^2} = \sqrt{\beta^2-\beta^2\sin^2\theta} = \beta\cos\theta$$

and $\dfrac{dx}{d\theta} = \beta\cos\theta$, so we can write $h(X)$ as

$$h(X) = \int_{-\pi/2}^{\pi/2} \frac{1}{\pi\beta\cos\theta}\log[\pi\beta\cos\theta]\,\beta\cos\theta\,d\theta = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\pi/2}\log[\pi\beta\cos\theta]\,d\theta$$
$$= \frac{2}{\pi}\int_0^{\pi/2}\log[\pi\beta\cos\theta]\,d\theta = \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\log[\pi\beta]\,d\theta + \frac{2}{\pi}\int_0^{\pi/2}\log[\cos\theta]\,d\theta$$
$$= \log\pi\beta + \frac{2}{\pi}\int_0^{\pi/2}\log[\cos\theta]\,d\theta$$
$$= \log_e\frac{\pi}{2} + \frac{2}{\pi}\left(-\frac{\pi}{2}\ln 2\right)$$
$$= \ln\frac{\pi}{2} - \ln 2 = \ln\frac{\pi}{4} \text{ (nats)}.$$

$\boxed{\text{recall: } \beta = 1/2, \text{ and from integral tables } \int_0^{\pi/2} \ln[\cos x]\,dx = -\frac{\pi}{2}\ln 2}$

(c) We can write the triangle density as $f(x) = \begin{cases} \frac{t}{\alpha}x, & x \leq 0 \leq \alpha \\ \dfrac{2-tx}{2t^{-1}-\alpha}, & \alpha \leq x \leq 2t^{-1} \\ \text{(here } \alpha \in [0, 2t^{-1}]\text{)} \end{cases}$.

$$h(X) = -\int_0^\alpha \frac{t}{\alpha}x\log\left(\frac{t}{\alpha}x\right)dx - \int_\alpha^{2t^{-1}}\left(\frac{2-tx}{2t^{-1}-\alpha}\right)\log\left(\frac{2-tx}{2t^{-1}-\alpha}\right)dx$$

Now $\int_0^\alpha \frac{t}{\alpha}x\log\left(\frac{t}{\alpha}x\right)dx = \frac{\alpha}{t}\int_0^t u\log u\,du$ and $\int_\alpha^{2t^{-1}}\left(\frac{2-tx}{2t^{-1}-\alpha}\right)\log\left(\frac{2-tx}{2t^{-1}-\alpha}\right)\frac{dx}{t}$

$$= \cdots = \left(\frac{2t^{-1}-\alpha}{t}\right)\int_0^t u\log u\,du$$

Thus

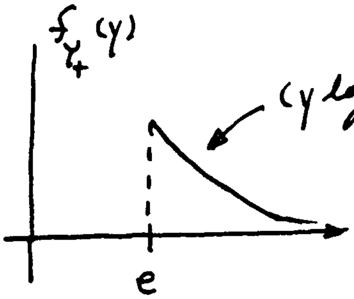
$$h(X) = \left[-\frac{\alpha}{t} + \frac{\alpha}{t} - \frac{2t^{-1}}{t}\right]\int_0^t u\log u\,du = \cdots = \frac{1}{2} - \ln t \text{ (nats)}$$

5. There are many possible solutions. Here's one.

Let X be an r.v. with density $f(x) = \begin{cases} x^{-2}, & x \geq 1 \\ 0, & \text{elsewhere} \end{cases}$

A simple calculation yields $h(X) = 2$ nats.

(a) Define an r.v. $Y_+ = e^X$. This has a pdf that appears as follows

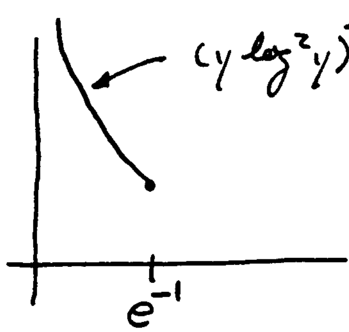


and a quick calculation shows

$$h(Y_+) = \int_e^\infty \frac{1}{y \log^2 y} \log\left[y \log^2 y\right] dy = +\infty$$

i.e. the integral diverges to $+\infty$

(b) Define the r.v. $Y_- = e^{-X}$. The pdf of Y_- appears as follows



and the differential entropy in this case is

$$h(Y_-) = \int_0^{e^{-1}} \frac{1}{y \log^2 y} \log\left[y \log^2 y\right] dy = -\infty$$

i.e. the integral diverges to $-\infty$.

6. The easiest way to solve this problem is to use the following Lemma and guess* the appropriate solution

Lemma: Given any two pdf's $f(x)$ and $g(x)$ defined on a subset $A \subseteq \mathbb{R}$ (the real line),

$$-\int_A f(x) \log f(x)\, dx \leq -\int_A f(x) \log g(x)\, dx, \quad \text{with equality iff } f(x) = g(x).$$

Proof: $-\int_A f(x) \log f(x)\, dx + \int_A f(x) \log g(x)\, dx = \int_A f(x) \log \frac{g(x)}{f(x)}\, dx \overset{J}{\leq} \log \int_A f(x) \frac{g(x)}{f(x)}\, dx$

$= \log \int_A g(x)\, dx = \log 1 = 0$, with equality iff $f(x) = g(x)$, by Jensen's inequality and the strict convexity of $\log(\cdot)$.

* Lagrange Multipliers may be useful in determining the initial guess, but are difficult to use for complete solution

We will now "guess" that the solution is $g(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma_X^2} \right\}$ ~~n~~oting that this gaussian density has mean μ and variance σ_X^2. Substituting this $g(x)$ into the Lemma above, we have that for <u>any</u> pdf $f(x)$

$$h(X) = -\int_{-\infty}^{\infty} f(x) \log f(x)\, dx \le -\int_{-\infty}^{\infty} f(x) \log g(x)\, dx$$

$$= -\int_{-\infty}^{\infty} f(x) \left[-\log \sqrt{2\pi}\,\sigma_X - \frac{(x-\mu)^2}{2\sigma_X^2} \right] dx$$

$$= \log \sqrt{2\pi}\,\sigma_X + \frac{\sigma_X^2}{2\sigma_X^2} = \frac{1}{2} \log 2\pi\sigma_X^2 + \frac{1}{2}\log e$$

$$= \frac{1}{2} \log 2\pi e \sigma_X^2 .$$

$\therefore \quad h(X) \le \frac{1}{2} \log 2\pi e \sigma_X^2$ if X has mean μ and variance σ_X^2, with equality iff $f(x) = g(x)$

$\therefore$ The pdf $f(x)$ with mean μ and variance σ_X^2 that has maximum differential entropy $h(X)$ is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma_X^2} \right\}$$

and the resulting diff. entropy is

$$h_{max}(X) = \frac{1}{2} \log 2\pi e \sigma_X^2 .$$

<u>n.b.</u> One way to guess the answer is Gaussian is to use Lagrange multipliers:

$$\Phi(f(\cdot)) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx - \lambda_0 \left[\int_{-\infty}^{\infty} f(x)\, dx - 1 \right] - \lambda_1 \left[\int_{-\infty}^{\infty} x f(x)\, dx - \mu \right] + \lambda_2 \left[\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx - \sigma_X^2 \right]$$

$$\Rightarrow \phi(f(x)) = -f(x) \ln f(x) - \lambda_0 f(x) - \lambda_1 x f(x) - \lambda_2 f(x) (x-\mu)^2$$

$$\frac{\partial \phi}{\partial f(x)} = -\ln f(x) - 1 - \lambda_0 - \lambda_1 x - \lambda_2 (x-\mu)^2 = 0, \quad \forall f(x)$$

$\Longrightarrow f(x) = A \exp\left\{ -\lambda_1 x - \lambda_2 (x-\mu)^2 \right\}$. In principle, one could substitute back into the constraints to find A, λ_1, and λ_2, but this becomes very messy. But for $A = \frac{1}{\sqrt{2\pi}\,\sigma_X}$, $\lambda_1 = 0$ and $\lambda_2 = \frac{1}{2\sigma_X^2}$, this is the gaussian with mean μ and variance σ_X^2, so this is a good (and the correct) guess.

7. Using Lagrange multipliers, we can set up the following objective function or Lagrangian for the $f(x)$ maximizing $h(X)$ subject to the constraints:

$$\Phi(f(\cdot)) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx - \sum_{i=1}^{n} \lambda_i \left[\int_{-\infty}^{\infty} g_i(x) f(x)\, dx - \eta_i \right]$$

From this, we have the integrand objective function

$$\phi(f(x)) = -f(x) \ln f(x) - \sum_{i=1}^{n} \lambda_i g_i(x) f(x).$$

In order for $f(x)$ to be the maximizing pdf, from the Euler-Lagrange equation

$$\frac{\partial \phi(f(x))}{\partial f(x)} = 0 \,, \quad \forall x \in \mathbb{R}.$$

So we have

$$\frac{\partial \phi(f(x))}{\partial f(x)} = -\ln f(x) + 1 - \sum_{i=1}^{n} \lambda_i g_i(x) = 0$$

$$\Rightarrow f(x) = \exp\left\{1 - \sum_{i=1}^{n} \lambda_i g_i(x)\right\} = A \exp\{-\lambda_1 g_1(x) - \cdots - \lambda_n g_n(x)\}$$

where A is a constant and $\lambda_1, \ldots, \lambda_n$ (the Lagrange multipliers) are constants that can be found (in principle) by substitution back into the constraint equations

$$\int_{-\infty}^{\infty} f(x)\, g_i(x)\, dx = \eta_i \,, \quad i = 1, \ldots, n.$$

Evaluating the differential entropy corresponding to this $f(x)$, we have

$$h(X) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx = -\int_{-\infty}^{\infty} f(x) \ln\left[A \exp\{-\lambda_1 g_1(x) - \cdots - \lambda_n g_n(x)\}\right] dx$$

$$= \int_{-\infty}^{\infty} f(x)\left[-\ln A + \lambda_1 g_1(x) + \cdots + \lambda_n g_n(x)\right] dx$$

$$= -\ln A + \lambda_1 \int_{-\infty}^{\infty} f(x) g_1(x)\, dx + \cdots + \lambda_n \int_{-\infty}^{\infty} f(x) g_n(x)\, dx$$

$$= -\ln A + \lambda_1 \eta_1 + \cdots + \lambda_n \eta_n .$$

8 (a)

$$-I(X;Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(x,y)}{f(x)f(y)}\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(x)f(y)}{f(x,y)}\,dx\,dy \overset{J}{\leq} \log\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{f(x)f(y)}{f(x,y)}\,dx\,dy$$

$$= \log\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)f(y)\,dx\,dy = \log\int_{-\infty}^{\infty} f(x)\,dx\int_{-\infty}^{\infty} f(y)\,dy = \log 1 = 0$$

with equality iff $\frac{f(x,y)}{f(x)f(y)} = \text{const} = 1 \iff f(x,y) = f(x)f(y) \iff X \perp\!\!\!\perp Y$

$\therefore$ $I(X;Y) \geq 0$, with equality iff X and Y are statistically independent.

(b) $$I(X;Y) = h(X) - h(X|Y) = \int_{-\infty}^{\infty} f(x)\log\frac{1}{f(x)}\,dx - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{1}{f(x|y)}\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(x|y)}{f(x)}\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(x,y)}{f(x)f(y)}\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{f(y|x)}{f(y)}\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{1}{f(y)}\,dx\,dy - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log\frac{1}{f(y|x)}\,dx\,dy$$

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

$$\therefore I(X;Y) = I(Y;X)$$

just as in the discrete case.

9. The cost-capacity function of the additive Gaussian noise channel with a mean power constraint is given by

$$C(\beta) = \sup_{f(x) \in \mathcal{F}(\beta)} \{ I(X;Y) \}$$

where $\mathcal{F}(\beta)$ is the set of all pdfs such that $E[X^2] = \int_{\mathbb{R}} x^2 f(x)\,dx \le \beta$.

We start by noting that

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

Now

$$h(Y/X) = E\left[\log \frac{1}{f(Y|X)}\right] = \iint_{\mathbb{R}^2} f_{XY}(x,y) \log \frac{1}{f(y|x)}\,dy\,dx$$

$$= \int_{\mathbb{R}} f_X(x) \int_{\mathbb{R}} f(y|x) \log \frac{1}{f(y|x)}\,dy\,dx \quad \ldots \quad (*)$$

Now $f(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$

and

$$f_{XY}(x,y) = f_{XZ}(x(x,y), z(x,y)) \left| \frac{\partial(x,z)}{\partial(x,y)} \right|$$

where $x(x,y) = x$ and $z(x,y) = y - x$. Thus

$$\frac{\partial(x,z)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

$$\therefore f_{XY}(x,y) = f_{XZ}(x, y-x)\,|1| = f_X(x) f_Z(y-x)$$

from which it follows that

$$f(y|x) = \frac{f_X(x) f_Z(y-x)}{f_X(x)} = f_Z(y-x)$$

Now substituting into $(*)$, we have

$$h(Y|X) = \int_{\mathbb{R}} f_X(x) \int_{\mathbb{R}} f_Z(y-x) \log \frac{1}{f_Z(y-x)}\,dy\,dx = \ldots$$

$$= \int_{\mathbb{R}} f_X(x) \cdot h(Z)\, dx \;= h(Z)$$

because for any fixed x,

$$\int_{\mathbb{R}} f_Z(y-x) \log \frac{1}{f_Z(y-x)}\, dy = \int_{\mathbb{R}} f_Z(r) \log \frac{1}{f_Z(r)}\, dr = h(Z)$$

$$\therefore\; h(Y/X) = h(Z).$$

Let $\mathcal{F}(\beta)$ = set of all pdfs having mean 0 and variance less than or equal to β. Then

$$C(\beta) = \max_{f_X(x) \in \mathcal{F}(\beta)} \{ I(X;Y) \} = \max_{f_X(x) \in \mathcal{F}(\beta)} \{ h(Y) - h(Y|X) \}$$

$$= \max_{f_X(x) \in \mathcal{F}(\beta)} \{ h(Y) \} - h(Z)$$

$$= \max_{f_X(x) \in \mathcal{F}(\beta)} \{ h(Y) \} - \tfrac{1}{2} \log 2\pi e\, \sigma_Z^2 \qquad --- (**)$$

We now use the following Lemma to maximize $I(X;Y)$ and hence find $C(\beta)$:

<u>Lemma:</u> Let $f(x)$ and $g(x)$ be any two valid pdfs. Then

$$\int_{\mathbb{R}} f(x) \log \frac{1}{g(x)}\, dx \geq \int_{\mathbb{R}} f(x) \log \frac{1}{f(x)}\, dx,$$

with equality iff $f(x) = g(x)$ (a.e.)

<u>Proof:</u> $\displaystyle\int_{\mathbb{R}} f(x) \log \frac{1}{f(x)}\, dx - \int_{\mathbb{R}} f(x) \log \frac{1}{g(x)}\, dx = \int_{\mathbb{R}} f(x) \log \frac{g(x)}{f(x)}\, dx$

$$\overset{(J)}{\leq} \log\left(\int_{\mathbb{R}} f(x) \frac{g(x)}{f(x)}\, dx \right) = \log\left(\int_{\mathbb{R}} g(x)\, dx \right) = \log 1 = 0$$

Furthermore, because $\log(\cdot)$ is strictly convex $\cap$, equality occurs iff $f(x) = g(x)$ a.e.

~~Claim~~

In order to find $C(\beta)$ as given in $(**)$, we must find the $f_X(x)$ having variance $\sigma_X^2 \leq \beta$ that maximizes $h(Y)$.

<u>Claim:</u> If X is a Gaussian with mean 0 and variance β, then Y is Gaussian with mean 0 and variance $\sigma_Y^2 = \beta + \sigma_Z^2$, and $I(X;Y)$ is maximized for all X with $E[X^2] \leq \beta$.

<u>Proof:</u> Because $X \perp\!\!\!\perp Z$, Y will have variance $\sigma_Y^2 = \beta + \sigma_Z^2$ regardless of the distribution of X (under the constraint that $E[X^2] \leq \beta$).

Now consider any $f_Y(y)$ with mean 0 and variance $\sigma_Y^2 = \beta + \sigma_Z^2$, and let

$g(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left(\frac{-y^2}{2\sigma_Y^2}\right)$. Then from the lemma above

$$h(Y) = \int_{\mathbb{R}} f(y) \log \frac{1}{f(y)}\, dy \leq \int_{\mathbb{R}} f(y) \log \frac{1}{g(y)}\, dy$$

$$= \int_{\mathbb{R}} f(y) \left[\log \sigma_Y\sqrt{2\pi} + \frac{y^2}{2\sigma_Y^2}\right] dy = \log \sigma_Y\sqrt{2\pi} + \frac{\sigma_Y^2}{2\sigma_Y^2}$$

$$= \frac{1}{2} \log 2\pi e \sigma_Y^2 \quad \text{with equality iff}$$

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left(\frac{-y^2}{2\sigma_Y^2}\right) \text{ (a.e.)}$$

Thus it follows that

$$C(\beta) = \frac{1}{2} \log 2\pi e \sigma_Y^2 - \frac{1}{2} \log 2\pi e \sigma_Z^2$$

$$= \frac{1}{2} \log 2\pi e (\sigma_Z^2 + \beta) - \frac{1}{2} \log 2\pi e \sigma_Z^2$$

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

and this is achieved by an X which is Gaussian with mean 0 and variance $\sigma_X^2 = \beta$.