EE 637 Final

May 5, Spring 2009

Name: Key

Instructions:

- This is a 120 minute exam containing **five** problems.
- Each problem is worth 40 points for a total score of 200 points
- You may **only** use your brain and a pencil (or pen) to complete this exam.
- You **may not** use your book, notes, or a calculator, or any other electronic or physical device for communicating or accessing stored information.

Good Luck.

Fact Sheet

- Function definitions

$$\mathrm{rect}(t) \triangleq \begin{cases} 1 & \text{for } |t| < 1/2 \\ 0 & \text{otherwise} \end{cases}$$

$$\Lambda(t) \triangleq \begin{cases} 1-|t| & \text{for } |t| < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{sinc}(t) \triangleq \frac{\sin(\pi t)}{\pi t}$$

- CTFT

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt$$
$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} df$$

- CTFT Properties

$$x(-t) \overset{CTFT}{\Leftrightarrow} X(-f)$$
$$x(t-t_0) \overset{CTFT}{\Leftrightarrow} X(f)e^{-j2\pi f t_0}$$
$$x(at) \overset{CTFT}{\Leftrightarrow} \frac{1}{|a|} X(f/a)$$
$$X(t) \overset{CTFT}{\Leftrightarrow} x(-f)$$
$$x(t)e^{j2\pi f_0 t} \overset{CTFT}{\Leftrightarrow} X(f-f_0)$$
$$x(t)y(t) \overset{CTFT}{\Leftrightarrow} X(f) * Y(f)$$
$$x(t) * y(t) \overset{CTFT}{\Leftrightarrow} X(f)Y(f)$$
$$\int_{-\infty}^{\infty} x(t)y^*(t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(f)Y^*(f)df$$

- CTFT pairs

$$\mathrm{sinc}(t) \overset{CTFT}{\Leftrightarrow} \mathrm{rect}(f)$$
$$\mathrm{rect}(t) \overset{CTFT}{\Leftrightarrow} \mathrm{sinc}(f)$$

For $a > 0$

$$\frac{1}{(n-1)!} t^{n-1} e^{-at} u(t) \overset{CTFT}{\Leftrightarrow} \frac{1}{(j2\pi f + a)^n}$$

- CSFT

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) e^{-j2\pi(ux+vy)} dxdy$$
$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v) e^{j2\pi(ux+vy)} dudv$$

- DTFT

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$$
$$x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega$$

- DTFT pairs

$$a^n u(n) \overset{DTFT}{\Leftrightarrow} \frac{1}{1 - ae^{-j\omega}}$$

- Rep and Comb relations

$$\mathrm{rep}_T[x(t)] = \sum_{k=-\infty}^{\infty} x(t-kT)$$
$$\mathrm{comb}_T[x(t)] = x(t)\sum_{k=-\infty}^{\infty} \delta(t-kT)$$
$$\mathrm{comb}_T[x(t)] \overset{CTFT}{\Leftrightarrow} \frac{1}{T}\mathrm{rep}_{\frac{1}{T}}[X(f)]$$
$$\mathrm{rep}_T[x(t)] \overset{CTFT}{\Leftrightarrow} \frac{1}{T}\mathrm{comb}_{\frac{1}{T}}[X(f)]$$

- Sampling and Reconstruction

$$y(n) = x(nT)$$
$$Y(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X\left(\frac{\omega - 2\pi k}{2\pi T}\right)$$
$$s(t) = \sum_{k=-\infty}^{\infty} y(k)\delta(t-kT)$$
$$S(e^{j\omega}) = Y\left(e^{j\omega T}\right)$$

Name: Key

Problem 1.(40pt)

Derive each of the following properties.

a) Show that if $g(t)$ has a CTFT of $G(f)$, then $g(t-a)$ has a CTFT of $e^{-2\pi jaf}G(f)$.

b) Show that if $g(t)$ has a CTFT of $G(f)$, then $g(t/a)$ has a CTFT of $|a|G(af)$.

c) Show that if x_n has a DTFT of $X(e^{j\omega})$, then $(-1)^n x_n$ has a DTFT of $X(e^{j(\omega-\pi)})$.

d) Show that if $g\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ has a CSFT of $G\left(\begin{bmatrix} u \\ v \end{bmatrix}\right)$, then $g\left(A\begin{bmatrix} x \\ y \end{bmatrix}\right)$ has a CSFT of $|A|^{-1}G\left((A^{-1})^t\begin{bmatrix} u \\ v \end{bmatrix}\right)$.

(Hint: Use the notation $r = \begin{bmatrix} x \\ y \end{bmatrix}$ and $f = \begin{bmatrix} u \\ v \end{bmatrix}$, so that $G(f) = \int_{\Re^2} g(r)e^{-jr^t f}dr$.)

a) Let $\mathcal{F}\{\cdot\}$ denote the Fourier transform operator.

$$\mathcal{F}\{g(t-a)\} = \int_{-\infty}^{\infty} g(t-a)\, e^{-j2\pi f t}\, dt$$

$$\overset{t'=t-a}{=} \int_{-\infty}^{\infty} g(t')\, e^{-j2\pi f(t'+a)}\, dt'$$

$$= e^{-j2\pi f a}\int_{-\infty}^{\infty} g(t')\, e^{-j2\pi f t'}\, dt'$$

$$= e^{-j2\pi f a}\, G(f)$$

b) $$\mathcal{F}\{g(\tfrac{t}{a})\} = \int_{-\infty}^{\infty} g(\tfrac{t}{a})\, e^{-j2\pi f t}\, dt$$

1) when $a>0$, $t' = \frac{t}{a}$

$$\mathcal{F}\{g(\tfrac{t}{a})\} = \int_{-\infty}^{\infty} g(t')\, e^{-j2\pi f a t'} \cdot a\, dt'$$

$$= a\int_{-\infty}^{\infty} g(t')\, e^{-j2\pi f a t'}\, dt'$$

$$= a\, G(af)$$

2) when $a<0$, $t' = \frac{t}{a}$

$$\mathcal{F}\{g(\tfrac{t}{a})\} = \int_{\infty}^{-\infty} g(t')\, e^{-j2\pi f a t'}\, a\, dt'$$

$$= -a\int_{-\infty}^{\infty} g(t')\, e^{-j2\pi f a t'}\, dt'$$

$$= -a\, G(af)$$

So $\mathcal{F}\{g(\frac{t}{a})\} = |a|\, G(af)$

Name: Key

c)

$$\mathrm{DTFT}\{(-1)^n x_n\} = \sum_{n=-\infty}^{\infty} (-1)^n x_n e^{-j\omega n}$$

$$= \sum_{n=-\infty}^{\infty} e^{j\pi n} x_n e^{-j\omega n}$$

$$= \sum_{n=-\infty}^{\infty} x_n e^{-j(\omega-\pi)n}$$

$$= X(e^{j(\omega-\pi)})$$

$r = \begin{pmatrix} x \\ y \end{pmatrix}$ $\quad f = \begin{pmatrix} u \\ v \end{pmatrix}$

d)

$$\mathcal{F}\{g(Ar)\} = \int_{\mathbb{R}^2} g(Ar)\, e^{-jr^t f}\, dr$$

$r' = Ar \qquad r = A^{-1}r'$

The Jacobian of r is A^{-1}, and $|A^{-1}| = |A|^{-1}$.

So
$$\mathcal{F}\{g(Ar)\} = \int_{\mathbb{R}^2} g(r') e^{-j(A^{-1}r')^t f} \cdot |A|^{-1} \cdot dr'$$

$$= |A|^{-1} \int_{\mathbb{R}^2} g(r') e^{-j r'^t (A^{-1})^t f} \cdot dr'$$

$$= |A|^{-1}\, G\left((A^{-1})^t f\right)$$

Name: Key

Problem 2.(40pt)

Consider a color imaging device that takes input values of (r, g, b) and produces output (X, Y, Z) values given by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = A \begin{bmatrix} r^\alpha \\ g^\alpha \\ b^\alpha \end{bmatrix} .$$

where

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} .$$

a) Calculate the white point of the device in chromaticity coordinates.

b) What are the primaries associated with the r, g, and b components respectively? Again, use chromaticity coordinates to specify your answer.

c) What is the gamma of the device?

d) Calculate the values of (r, g, b) that will produce the color of an equal energy white. (Hint: You can express your solution in terms of A^{-1}.)

a) White point corresponds to $r=1$, $g=1$, $b=1$.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b+c \\ d+e+f \\ g+h+i \end{pmatrix}$$

Let $C = a+b+c+d+e+f+g+h+i$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{a+b+c}{C} \\ \frac{d+e+f}{C} \\ \frac{g+h+i}{C} \end{pmatrix}$$

b) The primary associated with r is $\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \leftarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{a}{a+d+g} \\ \frac{d}{a+d+g} \\ \frac{g}{a+d+g} \end{pmatrix}$

The primary associated with g is $\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{b}{b+e+h} \\ \frac{e}{b+e+h} \\ \frac{h}{b+e+h} \end{pmatrix}$

The primary associated with b is $\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{c}{c+f+i} \\ \frac{f}{c+f+i} \\ \frac{i}{c+f+i} \end{pmatrix}$$

Name: Key

c) The gamma of the device is α.

d) Equal energy white means $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = A \begin{pmatrix} r^{\alpha} \\ g^{\alpha} \\ b^{\alpha} \end{pmatrix}$$

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \left[A^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right]^{\frac{1}{\alpha}}.$$

Name: Key

Problem 3.(40pt)

Let $Y \in \Re^N$ be a vector containing the pixels in an image window. We can model Y as

$$Y = tS + W$$

where $t \in \Re^N$ is a deterministic column vector of length N, S is scalar valued Gaussian random variable with mean 0 and variance σ^2, and W is a independent Gaussian random vector of correlated noise with distribution $N(0, R_w)$ where R_w is an $N \times N$ positive definite covariance matrix.

Intuitively, Y is composed of a signal tS obscured by noise W. Our objective is to estimate the value of S from the observations Y. To do this, we will form a MMSE linear estimator for S given by

$$\hat{S} = Y^t \theta$$

where $\theta \in \Re^N$ is a vector of coefficients.

Furthermore, define the covariance matrix of Y given by

$$R_y = E\left[YY^t\right] ,$$

and the cross-covariance column vector of Y and S given by

$$b = E\left[YS\right] .$$

a) Calculate an expression for the MSE given by $E\left[||S - \hat{S}||^2\right]$ in terms of R_y, b, σ^2, and θ.

b) Use the expression from part a) to compute the value of θ that produces the MMSE estimate of S.

c) Calculate R_y in terms of t, σ^2, and R_w.

d) Calculate b in terms of t, σ^2, and R_w.

e) Use the above results to calculate a closed form expression for $\hat{S}$.

a)
$$\begin{aligned} E\left[||S-\hat{S}||^2\right] &= E\left[(S-Y^t\theta)^2\right] \\ &= E\left[S^2 - 2\theta^t YS + \theta^t Y \cdot Y^t \theta\right] \\ &= E[S^2] - 2\theta^t E[YS] + \theta^t E[YY^t]\,\theta \\ &= \sigma^2 - 2\theta^t b + \theta^t R_y \theta \end{aligned}$$

b)
$$\frac{\partial}{\partial\theta} E\left[||S-\hat{S}||^2\right] = 2R_y\theta - 2b = 0 \qquad \theta^* = R_y^{-1} b$$

Name: Key

c) $R_y = E[yy^t]$

$= E[(ts+W)(ts+W)^t]$

$= E[s^2tt^t + stW^t + sWt^t + WW^t]$

Since s and W are independent, $E[sWt^t] = E[s]E[W]t^t = 0$

$E[stW^t] = tE[s]E[W^t] = 0$

$R_y = \sigma^2 tt^t + R_W$

d) $b = E[Ys]$

$= E[(ts+W)s]$

$= E[ts^2 + Ws]$

$= tE[s^2] + E[W]E[s]$

$= \sigma^2 t$

e) $\hat{s} = y^t\theta = y^t R_y^{-1} b$

$= y^t (\sigma^2 tt^t + R_W)^{-1} \sigma^2 t$

Note that we can show that R_y is positive definite.
Remember $R_y = \sigma^2 tt^t + R_W$. Since tt^t is positive semi-definite, and R_W is positive definite, we have that R_y is positive definite.
This is helpful in explaining why $\theta = R_y^{-1} b$ makes $E[\|s-\hat{s}\|^2]$ minimum. ($\frac{\partial^2}{\partial\theta^2} E[\|s-\hat{s}\|^2] = 2R_y$, R_y is positive definite.)

Name: Key

Problem 4.(40pt)

Consider a non-linear prediction problem for which we are trying to predict the value of a scalar Y_n from a vector of observations Z_n. Our assumption is that we can estimate Y_n using the non-linear predictor given by

$$\hat{Y}_n = f(Z_n, \theta)$$

where $\theta \in \Re^p$ is a p dimensional parameter vector that controls the behavior of the nonlinear predictor.

Fortunately, we are given some training data pairs with the form (Y_n, Z_n).[1] The data is partitioned into two sets. The first set, $n \in S_1$, contains $N = |S_1|$ pairs, and is used for training purposes. The second set, $n \in S_2$, contains $M = |S_2|$ pairs, and is used for testing purposes.

Using these data, we can define the training MSE as

$$MSE_1(\theta) = \frac{1}{N} \sum_{n \in S_1} \| Y_n - f(Z_n, \theta) \|^2 ,$$

the testing MSE as

$$MSE_2(\theta) = \frac{1}{M} \sum_{n \in S_2} \| Y_n - f(Z_n, \theta) \|^2 ,$$

and the expected MSE as

$$MSE_3(\theta) = E\left[\| Y_n - f(Z_n, \theta) \|^2\right] .$$

Based on these error measures, we can define the following two estimates for the parameter vector.

$$\hat{\theta} = \arg\min_{\theta} MSE_1(\theta)$$

$$\theta^* = \arg\min_{\theta} MSE_3(\theta)$$

a) Which of the two quantities would you expect to be smaller, $MSE_2(\hat{\theta})$ or $MSE_2(\theta^*)$? Why?

b) What is the disadvantage of using $MSE_2(\theta^*)$?

c) Approximately how large should N be in order for $\hat{\theta}$ to be useful?

d) Sketch the plots of $MSE_1(\hat{\theta})$, $MSE_2(\hat{\theta})$, and $MSE_2(\theta^*)$ as a function of the amount of training data N.

e) Which value would you expect to be smaller, $MSE_1(\hat{\theta})$ or $MSE_2(\hat{\theta})$. Why?

f) If you are reporting results of your experiment, which value should you report, $MSE_1(\hat{\theta})$ or $MSE_2(\hat{\theta})$. Why?

[1] Assume that each training data pair is independent, and each pairs has the same distribution.

Name: Key

a) $MSE_2(\theta^*)$ is expected to be smaller.

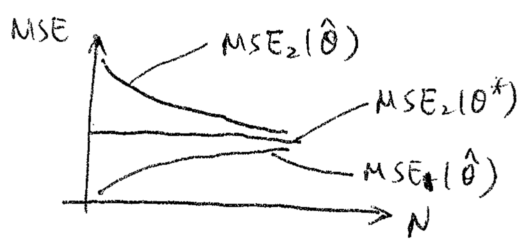
$$E[\|Y_n - f(Z_n, \theta^*)\|^2] \leq E[\|Y_n - f(Z_n, \hat{\theta})\|^2]$$

$$\begin{aligned}
E[MSE_2(\theta^*)] &= E\left[\frac{1}{M}\sum_{n\in S_2} \|Y_n - f(Z_n, \theta^*)\|^2\right] \\
&= \frac{1}{M}\sum_{n\in S_2} E[\|Y_n - f(Z_n, \theta^*)\|^2] \\
&= E[\|Y_n - f(Z_n, \theta^*)\|^2] \\
&\leq E[\|Y_n - f(Z_n, \hat{\theta})\|^2] \\
&= \frac{1}{M}\sum_{n\in S_2} E[\|Y_n - f(Z_n, \hat{\theta})\|^2] \\
&= E[MSE_2(\hat{\theta})]
\end{aligned}$$

Therefore, $E[MSE_2(\theta^*)] \leq E[MSE_2(\hat{\theta})]$

b) In order to obtain θ^*, we have to know the distribution of Y_n and Z_n, which is usually not available.

c) N should be at least larger than p. But the larger, the better.

d) 

e) $MSE_1(\hat{\theta})$ is smaller, because it is testing on the same dataset as the training data. Whereas $MSE_2(\hat{\theta})$ is testing on the testing data using the estimate from the training data.

f) Report $MSE_2(\hat{\theta})$, because in real cases it's meaningless to test on training data. Our goal is to use the estimated $\hat{\theta}$ along with new data to estimate Y_n. The new data won't be the same as the training data.

Name: Key

Problem 5.(40pt)

Consider the 1-D error diffusion algorithm specified by the equations

$$\begin{aligned} b_n &= Q(y_n) \\ e_n &= y_n - b_n \\ y_n &= x_n + e_{n-1} \end{aligned}$$

where x_n is the input, b_n is the output, and $Q(\cdot)$ is a binary quantizer with the form

$$Q(y) = \begin{cases} 1 & \text{if } y > 0.5 \\ 0 & \text{if } y \le 0.5 \end{cases} .$$

where we assume that $e_0 = 0$ and the algorithm is run for $n \ge 1$.

Furthermore, define $d_n = x_n - b_n$.

a) Draw a flow diagram for this algorithm. Make sure to label all the signals in the flow diagram using the notation defined above.

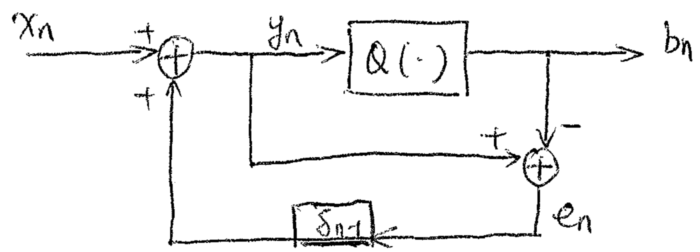
b) Calculate b_n for $n = 1$ to 10 when $x_n = 0.25$ and $e_0 = 0$.

c) Calculate an expression for d_n in terms of the quantization error e_n.

d) Calculate an expression for $\sum_{n=1}^{N} d_n$ in terms of the quantization error e_n.

e) What does the result of part d) above tell you about the output of error diffusion?

a)



b)

n	1	2	3	4	5	6	7	8	9	10
x_n	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
y_n	0.25	0.5	0.75	0	0.25	0.5	0.75	0	0.25	0.5
b_n	0	0	1	0	0	0	1	0	0	0
e_n	0.25	0.5	−0.25	0	0.25	0.5	−0.25	0	0.25	0.5

c) $d_n = x_n - b_n = x_n - (y_n - e_n) = x_n - (x_n + e_{n-1} - e_n) = e_n - e_{n-1}$

d) $\sum_{n=1}^{N} d_n = \sum_{n=1}^{N} (e_n - e_{n-1}) = e_N - e_0 = e_N$.

Name: Key

e) ① It tells us that the accumulative error is bounded.

Now we prove that e_N is bounded.

1) When $N=0$, $e_0 = 0$ is bounded.

2) We assume that e_k is bounded, then

$y_{k+1} = x_{k+1} + e_k$ is bounded since x_{k+1} is bounded as well.

Then $e_{k+1} = y_{k+1} - b_{k+1}$ is bounded, since y_{k+1} is bounded

and $b_{k+1} = Q(y_{k+1})$ is bounded.

Thus e_k is bounded $\Rightarrow$ e_{k+1} is bounded

3) By induction, e_N is bounded for $N \in \mathbb{Z}^+$.

② It tells us that the local average of the signal is approximately maintained.

$$\sum_{n=1}^{N} d_n = \sum_{n=1}^{N} (x_n - b_n) = e_N$$

$$\Rightarrow \sum_{n=1}^{N} b_n = \sum_{n=1}^{N} x_n - e_N$$

$$\Rightarrow \frac{1}{N}\sum_{n=1}^{N} b_n = \frac{1}{N}\sum_{n=1}^{N} x_n - \frac{1}{N} e_N$$

So the local average of the output is approximately the same as the local average of the input.