Sample Problems for the 9th Quiz

1. (PSD of Line Codes)

   (a) Show that the time-averaged PSD of the line coded signal

   $$s(t) = \sum_{n=-\infty}^{\infty} a_n \cdot g(t - nT_\mathrm{b})$$

   is equal to

   $$\overline{\mathrm{PSD}}(f) = \frac{1}{T_\mathrm{b}} |G(f)|^2 \bar{S}_a(f),$$

   where

   $$\bar{S}_a(f) = \sum_{k=-\infty}^{\infty} \bar{\phi}_a(k) e^{-j2\pi f k T_\mathrm{b}}$$

   and

   $$\bar{\phi}_a(k) = \lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N-1} E[a_{m+k} a_m^*]$$

   are respectively the time-average PSD and time-average autocorrelation function of $\{a_n\}$.
   Hint: Use (cf. Slide 2-30) the formula:

   $$\overline{\mathrm{PSD}}(f) = \lim_{T\to\infty} \frac{1}{2T} E[S(f) S_{2T}^*(f)]$$

   where $s_{2T}(t) = s(t) \cdot \mathbf{1}\{|t| \le T\}$.

   (b) From (a), we can now derive the time-average PSD for a deterministic $\{a_n\}$. Show that

   $$\bar{S}_a(f) = \frac{1}{T_\mathrm{b}} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T_\mathrm{b}}\right),$$

   if all-one sequence $\{a_n\}$ is transmitted.
   Hint: Note that

   $$\sum_{k=-\infty}^{\infty} e^{-j2\pi f k T} = \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right).$$

   (c) Find the time-averaged PSDs of the unipolar NRZ line code and the polar NRZ line code (See Slide 6-54), provided that all-one sequence $\{a_n\}$ is transmitted. Do the two line codes have the same time-average PSD for this particular input sequence?
   Hint: For unipolar NRZ, $|G(f)|^2 = A^2 T_\mathrm{b}^2 \mathrm{sinc}^2(f T_\mathrm{b})$.

   (d) Show that the time-averaged PSD of the Manchester line code is given by

   $$\overline{\mathrm{PSD}}_{\mathrm{Manchester}}(f) = \frac{4A^2}{\pi^2} \sum_{k \text{ odd}} \frac{1}{k^2} \delta\left(f - \frac{k}{T}\right)$$

   provided all-one sequence $\{a_n\}$ is transmitted.
   Hint: For Manchester line code, $|G(f)|^2 = A^2 T_\mathrm{b}^2 \mathrm{sinc}^2(\frac{f T_\mathrm{b}}{2}) \sin^2\left(\frac{\pi f T_\mathrm{b}}{2}\right)$.

**Solution.**

(a) First, we derive

$$\begin{aligned} S(f) &= \mathcal{F}\{s(t)\} = \mathcal{F}\left\{\sum_{n=-\infty}^{\infty} a_n \cdot g(t-nT_{\rm b})\right\} \\ &= \sum_{n=-\infty}^{\infty} a_n \cdot \mathcal{F}\{g(t-nT_{\rm b})\} = \sum_{n=-\infty}^{\infty} a_n \cdot G(f)e^{-j2\pi f nT_{\rm b}} \end{aligned}$$

and with $T = 2NT_{\rm b}$, we further derive

$$\begin{aligned} S_{2T}(f) &= \mathcal{F}\{s(t)\} = \mathcal{F}\left\{\sum_{n=-N}^{N-1} a_n \cdot g(t-nT_{\rm b})\right\} \\ &= \sum_{n=-N}^{N-1} a_n \cdot \mathcal{F}\{g(t-nT_{\rm b})\} = \sum_{n=-N}^{N-1} a_n \cdot G(f)e^{-j2\pi f nT_{\rm b}}. \end{aligned}$$

Hence,

$$\begin{aligned} \overline{\rm PSD}(f) &= \lim_{T\to\infty} \frac{1}{2T} E[S(f)S_{2T}^*(f)] \\ &= \lim_{N\to\infty} \frac{1}{2NT_{\rm b}} E\left[\left(\sum_{n=-\infty}^{\infty} a_n \cdot G(f)e^{-j2\pi f nT_{\rm b}}\right)\left(\sum_{m=-N}^{N-1} a_m \cdot G(f)e^{-j2\pi f mT_{\rm b}}\right)^*\right] \\ &= \frac{1}{T_{\rm b}}|G(f)|^2 \lim_{N\to\infty} \frac{1}{2N} \sum_{m=-N}^{N-1} \sum_{n=-\infty}^{\infty} E\left[a_n a_m^*\right] e^{-j2\pi f(n-m)T_{\rm b}} \quad (\text{Let } k = n-m) \\ &= \frac{1}{T_{\rm b}}|G(f)|^2 \lim_{N\to\infty} \frac{1}{2N} \sum_{m=-N}^{N-1} \sum_{k=-\infty}^{\infty} E\left[a_{m+k} a_m^*\right] e^{-j2\pi f kT_{\rm b}} \\ &= \frac{1}{T_{\rm b}}|G(f)|^2 \sum_{k=-\infty}^{\infty} \left(\lim_{N\to\infty} \frac{1}{2N} \sum_{m=-N}^{N-1} E\left[a_{m+k} a_m^*\right]\right) e^{-j2\pi f kT_{\rm b}} \\ &= \frac{1}{T_{\rm b}}|G(f)|^2 \sum_{k=-\infty}^{\infty} \bar{\phi}_a(k) e^{-j2\pi f kT_{\rm b}} \end{aligned}$$

(b) We have $E[a_{m+k} m_m^*] = 1$ for every $m$ and $k$. Therefore,

$$\begin{aligned} \bar{S}_a(f) &= \sum_{k=-\infty}^{\infty} \left(\lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N-1} E[a_{m+k} a_m^*]\right) e^{-j2\pi f kT_{\rm b}} \\ &= \sum_{k=-\infty}^{\infty} \left(\lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N-1} 1\right) e^{-j2\pi f kT_{\rm b}} \\ &= \sum_{k=-\infty}^{\infty} e^{-j2\pi f kT_{\rm b}} = \frac{1}{T_{\rm b}} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T_{\rm b}}\right) \end{aligned}$$

Note: $a_n = 1$ for every $n$ is an extreme case, in which the "strongest" DC term is produced.

(c)

$$
\begin{aligned}
\overline{\text{PSD}}(f) &= \frac{1}{T_\text{b}}|G(f)|^2\bar{S}_a(f) \\
&= \frac{1}{T_\text{b}}A^2T_\text{b}^2\text{sinc}^2(fT_\text{b})\cdot\frac{1}{T_\text{b}}\sum_{k=-\infty}^{\infty}\delta\left(f-\frac{k}{T_\text{b}}\right) \\
&= A^2\text{sinc}^2(fT_\text{b})\sum_{k=-\infty}^{\infty}\delta\left(f-\frac{k}{T_\text{b}}\right) \\
&= A^2\delta(f)
\end{aligned}
$$

We can devise from the above formula that both line codings (i.e., unipolar and polar NRZ) have the same time-average PSD.

Note: Polar NRZ hopes to remove the "DC" (of the time-average PSD) from Unipolar NZR by assuming that the data sequence $\{a_n\}$ has zero mean in addition to i.i.d. However, when the data sequence does not fulfill this assumption, the DC remains. Particularly in this extreme example, both line codings have the same time-average PSD and no removal of "DC" can be achieved by adopting polar NRZ.

(d)

$$
\begin{aligned}
\overline{\text{PSD}}_\text{Manchester}(f) &= \frac{1}{T_\text{b}}|G(f)|^2\bar{S}_a(f) \\
&= A^2T_\text{b}\text{sinc}^2\left(\frac{fT_\text{b}}{2}\right)\sin^2\left(\frac{\pi fT_\text{b}}{2}\right)\frac{1}{T_\text{b}}\sum_{k=-\infty}^{\infty}\delta\left(f-\frac{k}{T_\text{b}}\right) \\
&= A^2\sum_{k=-\infty}^{\infty}\text{sinc}^2\left(\frac{fT_\text{b}}{2}\right)\sin^2\left(\frac{\pi fT_\text{b}}{2}\right)\delta\left(f-\frac{k}{T_\text{b}}\right) \\
&= A^2\sum_{k=-\infty}^{\infty}\text{sinc}^2\left(\frac{k}{2}\right)\sin^2\left(\frac{\pi k}{2}\right)\delta\left(f-\frac{k}{T_\text{b}}\right) \\
&= \frac{4A^2}{\pi^2}\sum_{k\text{ odd}}\frac{1}{k^2}\delta\left(f-\frac{k}{T}\right)
\end{aligned}
$$

Note: Even when an all-one sequence is transmitted, Manchester code still has no DC in its time-average PSD.

2. For an AT&T M12 multiplexer, 24 control bits are separated by sequences of 48 data bits (12 from each DS1 input). The frame format is given below.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $M_0$ | 48 | $C_I$ | 48 | $F_0$ | 48 | $C_I$ | 48 | $C_I$ | 48 | $F_1$ | 48 |
| $M_1$ | 48 | $C_{II}$ | 48 | $F_0$ | 48 | $C_{II}$ | 48 | $C_{II}$ | 48 | $F_1$ | 48 |
| $M_1$ | 48 | $C_{III}$ | 48 | $F_0$ | 48 | $C_{III}$ | 48 | $C_{III}$ | 48 | $F_1$ | 48 |
| $M_1$ | 48 | $C_{IV}$ | 48 | $F_0$ | 48 | $C_{IV}$ | 48 | $C_{IV}$ | 48 | $F_1$ | 48 |
| Subframe markers | | Stuffing indicators | | Frame markers | | Stuffing indicators | | Stuffing indicators | | Frame markers | |

(a) Suppose only 287 data bits arrive from DS1#1 within a duration of a DS2 frame, but DS1#2, DS1#3 and DS1#4 do have 288 data bits available in their input buffers. Give the values of $C_{\text{I}}$, $C_{\text{II}}$, $C_{\text{III}}$ and $C_{\text{IV}}$.

(b) Suppose the nominal output bit rate is 6.312 Mbps. Determine the *largest incoming bit rate* $f_{\text{in,max}}$ and the *smallest incoming bit rate* $f_{\text{in,min}}$ allowed for each DS1 in the system.

Note: Give the values of $f_{\text{in,max}}$ and $f_{\text{in,min}}$ in the format of $x.xxxxx$ Mbps, i.e,. rounding off to the 5th decimal place.

(c) Find the allowable tolerance range for DS1 inputs in terms of ppm with respect to $f_{\text{in,nominal}} = 1.544$ Mbps.

(d) Suppose that over a particular cable, decreasing one degree on the Fahrenheit scale will result in approximately 100 ppm variation. From the tolerance in (c), find the range of temperature variation allowable for this cable.

**Solution.**

(a) $C_{\text{I}} = 1$, $C_{\text{II}} = 0$, $C_{\text{III}} = 0$ and $C_{\text{IV}} = 0$.

(b) During the time period for the M12 multiplexer to send out 1176 bits, each DS1 input must provide at least 287 bits and at most 288 bits; hence,

$$\frac{288}{f_{\text{in}}} \geq \frac{1176}{6.312} \geq \frac{287}{f_{\text{in}}}$$

which implies

$$f_{\text{in,max}} = 1.54580 \approx \frac{288}{1176} 6.312 \geq f_{\text{in}} \geq \frac{287}{1176} 6.312 \approx 1.54043 = f_{\text{in,min}}.$$

(c) We shall have

$$\frac{10^6 - b_{ppm}}{f_{\text{in,min}}} = \frac{10^6}{f_{\text{in,nominal}}} = \frac{10^6 + a_{ppm}}{f_{\text{in,max}}}.$$

Thus,

$$\begin{aligned} a_{ppm} + b_{ppm} &= 10^6 \left( \frac{f_{\text{in,max}}}{f_{\text{in,nominal}}} - 1 \right) + 10^6 \left( 1 - \frac{f_{\text{in,min}}}{f_{\text{in,nominal}}} \right) \\ &= 10^6 \left( \frac{f_{\text{in,max}} - f_{\text{in,min}}}{f_{\text{in,nominal}}} \right) \\ &= 10^6 \left( \frac{\frac{6.312}{1176}}{1.544} \right) \approx 3476.26 \text{ ppm} \end{aligned}$$

(d)

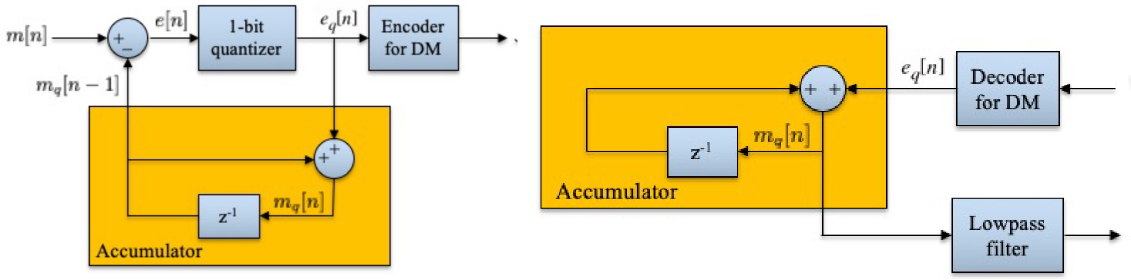
$$\frac{a_{ppm} + b_{ppm}}{100} = 34.76 \text{ degrees on the Fahrenheit scale}$$

3. The below problems demonstrate how delta modulation and delta-sigma modulation work with intuitively chosen step sizes.

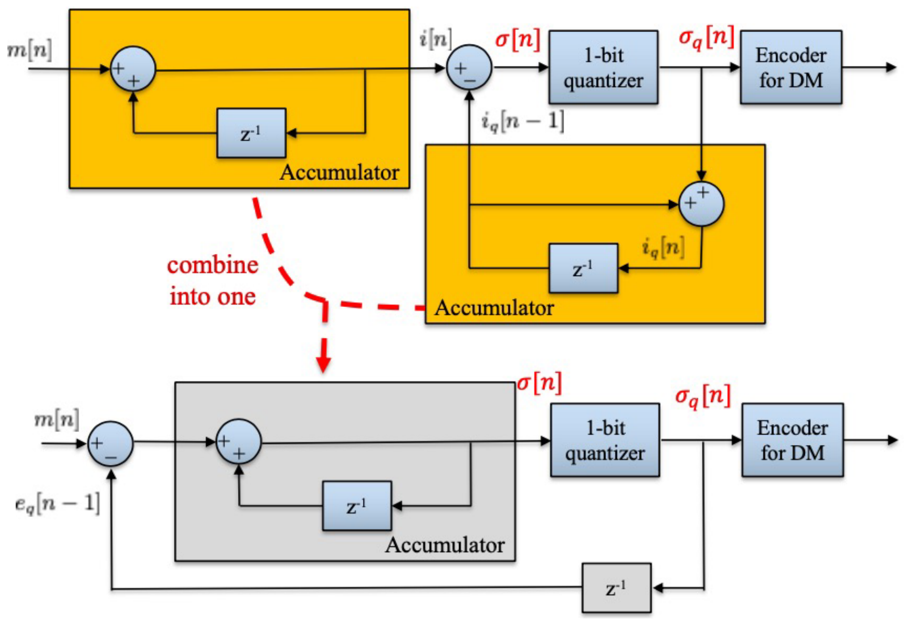
(a) Suppose in the DM modulation below,

$$\begin{aligned}&\{m[n] : 0 \le n \le 12\}\\ &= \{0.00, 0.24, 0.26, 0.48, 0.66, 0.70, 0.59, 0.32, 0.21, 0.36, 0.46, 0.21, 0.11\}.\end{aligned}$$

Find $e_q[n]$ and $m_q[n]$ with $\Delta = 0.8$.




(b) With the same $\{m[n], 0 \le n \le 12\}$ in (a) and the same $\Delta = 0.8$, find $i_q[n]$ and $\sigma_q[n]$. Explain why we need $\Delta \ge \max_n m[n]$ to prevent the slope overload distortion.




(c) Show that if the one-bit quantizer in (a) is replaced by an $\infty$-bit quantizer such that $e_q[n] = e[n]$, then $m_q[n] = m[n]$.

(d) Show that if the one-bit quantizer in (b) is replaced by an $\infty$-bit quantizer such that $\sigma_q[n] = \sigma[n]$, then $\sigma_q[n] = m[n]$.

**Solution.**

(a) Following

$$\begin{cases} e[n] = m[n] - m_q[n-1]; \\ e_q[n] = \Delta \cdot \mathrm{sgn}(e[n]); \\ m_q[n] = m_q[n-1] + e_q[n], \end{cases}$$

we obtain

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m[n]$ | 0 | 0.24 | 0.26 | 0.48 | 0.66 | 0.70 | 0.59 | 0.32 | 0.21 | 0.36 | 0.46 | 0.21 | 0.11 |
| $m_q[n]$ | 0 | 0.8 | 0 | 0.8 | 0 | 0.8 | 0 | 0.8 | 0 | 0.8 | 0 | 0.8 | 0 |
| $e[n]$ | 0 | 0.24 | $-0.54$ | 0.48 | $-0.14$ | 0.7 | $-0.21$ | 0.32 | $-0.59$ | 0.36 | $-0.34$ | 0.21 | $-0.69$ |
| $e_q[n]$ | 0 | 0.8 | $-0.8$ | 0.8 | $-0.8$ | 0.8 | $-0.8$ | 0.8 | $-0.8$ | 0.8 | $-0.8$ | 0.8 | $-0.8$ |
| code | | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |

(b) Following

$$\begin{cases} i[n] = i[n-1] + m[n]; \\ \sigma[n] = i[n] - i_q[n-1]; \\ \sigma_q[n] = \Delta \cdot \mathrm{sgn}(\sigma[n]); \\ i_q[n] = i_q[n-1] + \sigma_q[n], \end{cases}$$

we obtain

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m[n]$ | 0 | 0.24 | 0.26 | 0.48 | 0.66 | 0.70 | 0.59 | 0.32 | 0.21 | 0.36 | 0.46 | 0.21 | 0.11 |
| $i[n]$ | 0 | 0.24 | 0.50 | 0.98 | 1.64 | 2.34 | 2.93 | 3.25 | 3.46 | 3.82 | 4.28 | 4.49 | 4.60 |
| $i_q[n]$ | 0 | 0.8 | 0 | 0.8 | 1.6 | 2.4 | 3.2 | 4.0 | 3.2 | 4.0 | 4.8 | 4.0 | 4.8 |
| $\sigma[n]$ | 0 | 0.24 | $-0.3$ | 0.98 | 0.84 | 0.74 | 0.53 | 0.05 | $-0.54$ | 0.62 | 0.28 | $-0.31$ | 0.6 |
| $\sigma_q[n]$ | 0 | 0.8 | $-0.8$ | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | $-0.8$ | 0.8 | 0.8 | $-0.8$ | 0.8 |
| code | | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |

In order to prevent the slope overload distortion, we need

$$\frac{\Delta}{T_s} \geq \max_n \frac{|i[n] - i[n-1]|}{T_s} = \max_n \frac{|m[n]|}{T_s} = \frac{0.7}{T_s}.$$

Thus, $\Delta = 0.8$ satisfies the need.

Note: The circuit in (a) receives $m_q[n]$, while the circuit in (b) receives $\sigma_q[n]$. It is clear from the two tables that $m_q[n] \neq \sigma_q[n]$. Hence, moving the accumulator from the receiver to the transmitter does not yield an equivalent circuit.

(c) With

$$\begin{cases} e[n] = m[n] - m_q[n-1]; & (1) \\ \underline{e_q[n] = e[n];} & (2) \\ m_q[n] = m_q[n-1] + e_q[n]; & (3) \end{cases}$$

we obtain from (2) and (3) that $m_q[n] = m_q[n-1] + e[n]$. Then, taking (1) into $m_q[n] = m_q[n-1] + e[n]$, we obtain $m_q[n] = m_q[n-1] + (m[n] - m_q[n-1]) = m[n]$.

(d) Following

$$\begin{cases} i[n] = i[n-1] + m[n]; & (1) \\ \sigma[n] = i[n] - i_q[n-1]; & (2) \\ \underline{\sigma_q[n] = \sigma[n];} & (3) \\ i_q[n] = i_q[n-1] + \sigma_q[n]; & (4) \end{cases}$$

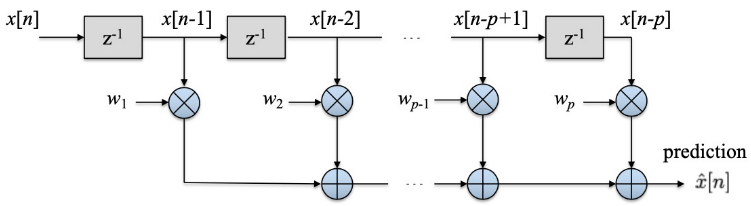
we obtain from (3) and (4) that $i_q[n] = i_q[n-1] + \sigma[n]$, which together with (2) yields $i_q[n] = i[n]$. As a result, we can rewrite (2) as $\sigma[n] = i[n] - i[n-1]$, and hence applying (1) gives $\sigma[n] = m[n]$.

Note: (c) and (d) indicates that the circuit in (a) is equivalent to the circuit in (b) when the 1-bit quantizer is replaced by a "linear" operation, say the output is equal to the input. Thus, if we use a finer quantizer such as 4-bit quantizer, the (near-)equivalence of the two circuits will be better achieved.

4. The linear predictor in the figure below gives

$$\hat{x}[n] = \sum_{k=1}^{p} w_k x[n-k],$$

which uses a weighted sum of $x[n-1]$, $x[n-2]$, …, $x[n-p]$ to produce a prediction for $x[n]$.




Let the prediction error be

$$e[n] \triangleq x[n] - \hat{x}[n].$$

(a) Show that the optimal $\{w_k\}_{k=1}^{p}$ that minimizes $\langle e[n], e[n]\rangle$ satisfies:

$$\begin{bmatrix} \langle x[n-1], x[n-1]\rangle & \langle x[n-1], x[n-2]\rangle & \cdots & \langle x[n-1], x[n-p]\rangle \\ \langle x[n-2], x[n-1]\rangle & \langle x[n-2], x[n-2]\rangle & \cdots & \langle x[n-2], x[n-p]\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x[n-p], x[n-1]\rangle & \langle x[n-p], x[n-2]\rangle & \cdots & \langle x[n-p], x[n-p]\rangle \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} \langle x[n], x[n-1]\rangle \\ \langle x[n], x[n-2]\rangle \\ \vdots \\ \langle x[n], x[n-p]\rangle \end{bmatrix}$$

Hint: An inner product $\langle x[n], y[n]\rangle$ for real $x[n]$ and $y[n]$ should satisfy the following properties:

> **Definition:** A mapping from $\mathcal{V} \times \mathcal{V}$ to $\mathbb{F}$, denoted by $\langle \cdot, \cdot \rangle$, is an inner product if for every $x, y, z \in \mathcal{V}$ and $a \in \mathbb{F}$,
>
> 1. Positive-definiteness: $\langle x, x\rangle \geq 0$ with equality only when $x = 0$
> 2. Symmetry: $\langle x, y\rangle = \langle y, x\rangle$
> 3. Linearity: $\begin{cases} \langle ax, y\rangle = a\langle x, y\rangle \\ \langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \end{cases}$

(b) Show that the optimal solution satisfy $\langle x[n-i], e[n]\rangle = 0$ for $1 \leq i \leq p$. In other words, the error $e[n]$ is orthogonal to $x[n-i]$ for every $1 \leq i \leq p$.

(c) Verify that the inner product

$$\langle x[n], y[n]\rangle \triangleq E\big[x[n]y[n]\big]$$

recovers the Winer-Hopf equations for linear prediction.

(d) Verify that the slope $g_i[n]$ established in Slide 7-48 can be recovered based on the inner product:

$$\langle x[n], y[n]\rangle \triangleq x[n]y[n].$$

.

**Solution.**

(a) First we note from the properties of an inner product:

$$\begin{aligned}
J &\triangleq \langle e[n], e[n]\rangle \\
&= \left\langle x[n] - \sum_{k=1}^{p} w_k x[n-k], x[n] - \sum_{k=1}^{p} w_k x[n-k] \right\rangle \\
&= \langle x[n], x[n]\rangle - 2 \underbrace{\sum_{k=1}^{p} w_k \langle x[n], x[n-k]\rangle}_{\text{Here, we apply } \langle x[n],x[n-k]\rangle = \langle x[n-k],x[n]\rangle.} + \sum_{k=1}^{p}\sum_{j=1}^{p} w_k w_j \langle x[n-k], x[n-j]\rangle \\
&= \langle x[n], x[n]\rangle - 2\sum_{k=1}^{p} w_k \langle x[n], x[n-k]\rangle \\
&\quad + \Bigg( \underbrace{2\sum_{k=1}^{p}\sum_{j=k+1}^{p} w_k w_j \langle x[n-k], x[n-j]\rangle}_{\text{Use again } \langle x[n-k],x[n-j]\rangle = \langle x[n-j],x[n-k]\rangle} + \sum_{k=1}^{p} w_k^2 \langle x[n-k], x[n-k]\rangle \Bigg)
\end{aligned}$$

Derive

$$\begin{aligned}
\frac{\partial}{\partial w_i} J &= -2\langle x[n], x[n-i]\rangle + 2 \underbrace{\sum_{j=i+1}^{p} w_i w_j \langle x[n-i], x[n-j]\rangle}_{\text{The case of } k=i} \\
&\quad + 2 \underbrace{\sum_{k=1}^{i-1} w_k w_i \langle x[n-k], x[n-i]\rangle}_{\text{The case of } j=i} + 2 w_i \langle x[n-i], x[n-i]\rangle \\
&= -2\langle x[n], x[n-i]\rangle + 2\sum_{j=1}^{p} w_j \langle x[n-i], x[n-j]\rangle
\end{aligned}$$

Thus, the optimal solution should satisfy

$$\begin{bmatrix}
\langle x[n-1], x[n-1]\rangle & \langle x[n-1], x[n-2]\rangle & \cdots & \langle x[n-1], x[n-p]\rangle \\
\langle x[n-2], x[n-1]\rangle & \langle x[n-2], x[n-2]\rangle & \cdots & \langle x[n-2], x[n-p]\rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle x[n-p], x[n-1]\rangle & \langle x[n-p], x[n-2]\rangle & \cdots & \langle x[n-p], x[n-p]\rangle
\end{bmatrix}
\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix}
=
\begin{bmatrix} \langle x[n], x[n-1]\rangle \\ \langle x[n], x[n-2]\rangle \\ \vdots \\ \langle x[n], x[n-p]\rangle \end{bmatrix}$$

(b) We can rewrite $\partial J/\partial w_i$ as

$$\begin{aligned}
\frac{\partial}{\partial w_i} J &= -2\langle x[n-i], x[n]\rangle + 2\sum_{j=1}^{p} w_j \langle x[n-i], x[n-j]\rangle \\
&= 2\left\langle x[n-i], -x[n] + \sum_{j=1}^{p} w_j x[n-j] \right\rangle \\
&= -2\,\langle x[n-i], e[n]\rangle\,.
\end{aligned}$$

Hence, with the optimal $\{w_i\}$, the error $e[n]$ is orthogonal to $x[n-i]$ for every $1 \le i \le p$.
Note: As a result, what the DPCM quantizes is the remaining message $e[n] = m[n] - \hat{m}[n]$ that is orthogonal to $m[n-i]$ for every $1 \le i \le p$.

(c) It is straightforward. Hence, we omit the solution.

(d) From (b), we obtain

$$\frac{\partial}{\partial w_i} J = -2 \langle x[n-i], e[n] \rangle = -2x[n-i]e[n].$$

Note: If, subject to the inner product $\langle x[n], y[n] \rangle = x[n]y[n]$, we wish to find the optimal $\{w_i\}$ that minimizes $e^2[n]$ based on $\{x[n-k]\}_{k=1}^{p}$, the solution in principle must satisfy

$$\begin{bmatrix} x[n-1]x[n-1] & x[n-1]x[n-2] & \cdots & x[n-1]x[n-p] \\ x[n-2]x[n-1] & x[n-2],x[n-2] & \cdots & x[n-2]x[n-p] \\ \vdots & \vdots & \ddots & \vdots \\ x[n-p]x[n-1] & x[n-p]x[n-2] & \cdots & x[n-p]x[n-p] \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} x[n]x[n-1] \\ x[n]x[n-2] \\ \vdots \\ x[n]x[n-p] \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} x[n-1]x[n-1] & x[n-1]x[n-2] & \cdots & x[n-1]x[n-p] \\ x[n-2]x[n-1] & x[n-2]x[n-2] & \cdots & x[n-2]x[n-p] \\ \vdots & \vdots & \ddots & \vdots \\ x[n-p]x[n-1] & x[n-p]x[n-2] & \cdots & x[n-p]x[n-p] \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = x[n] \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix} \begin{bmatrix} x[n-1] & x[n-2] & \cdots & x[n-p] \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = x[n] \begin{bmatrix} x[n-1] \\ x[n-2] \\ \vdots \\ x[n-p] \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} x[n-1] & x[n-2] & \cdots & x[n-p] \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = x[n] \quad (\text{provided } x[n-i] \neq 0 \ \forall i)$$

However, at the time we produce the prediction $\hat{x}[n]$ of $x[n]$, we do not know $x[n]$. Thus, the above equation does not help us to derive $\{w_k\}$. This is different from the Wiener-Hopf equations, where we do know $\langle x[n], x[n-k] \rangle = R_X[k]$ even if we do not know $x[n]$. Therefore, a linear adaptive predictor that adjusts $w_i$ based on $g_i[n]$ is used instead.