INDIAN INSTITUTE OF SCIENCE
E9-252: MATHEMATICAL METHODS AND TECHNIQUES IN SIGNAL PROCESSING
HOME WORK #1 - SOLUTIONS, FALL 2015

INSTRUCTOR: SHAYAN G. SRINIVASA
TEACHING ASSISTANT: CHAITANYA KUMAR MATCHA

Problem 1. Provide an example of a 2D FIR filter with the following impulse response properties (a) non-causal (b) causal but unstable. Repeat this for the 2D IIR case. Justify your steps.

Solution. 2D FIR filter, non-causal:

$$\begin{aligned} h(t_1,t_2) &= (u(t_1+N)-u(t_1-N))(u(t_2+N)-u(t_2-N)) \\ &= \begin{cases} 1, & -N \le t_1 \le N \text{ and } -N \le t_2 \le N, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

for some positive integer N. The filter has finite number of non-zero coefficient and hence is a FIR filter. Since $h(-1,-1)=1$, the filter is non-causal.

2D FIR filter, causal but unstable:

$$h(t_1,t_2) = \begin{cases} 1, & -N \le t_1 \le N \text{ and } -N \le t_2 \le N, \\ \infty, & (t_1,t_2) = (N+1,N+1), \\ 0, & \text{otherwise} \end{cases}$$

is a FIR causal filter which is not BIBO stable.

All FIR filters with finite coefficients are BIBO stable. Let $x(t_1,t_2)$ be a bounded input to a FIR filter whose coefficients are $h(t_1,t_2)$. Let $|x(t_2,t_2)| < B\,\forall t_1,t_2$. Then the output is

$$\begin{aligned} y(t_1,t_2) &= \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} h(\tau_1,\tau_2)\,x(t_1-\tau,t_2-\tau) \\ \implies |y(t_1,t_2)| &= \left|\sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} h(\tau_1,\tau_2)\,x(t_1-\tau,t_2-\tau)\right| \\ &\le \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} |h(\tau_1,\tau_2)\,x(t_1-\tau,t_2-\tau)| \quad \text{(Triangle inequality)} \\ &\le B\sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} |h(\tau_1,\tau_2)| \quad (|x(t_2,t_2)| < B) \\ &< \infty \quad (h(t_1,t_2) \text{ is FIR and has finite coefficients.}) \end{aligned}$$

$$h(t_1,t_2) = = (u(t_1)-u(t_1-N))(u(t_2)-u(t_2-N)) = \begin{cases} 1, & 0 \le t_1 \le N \text{ and } 0 \le t_2 \le N, \\ 0, & \text{otherwise} \end{cases}$$

is an example of FIR causal filter which is stable.

2D IIR filter, non-causal:
$h(t_1,t_2) = a^{t_1}b^{t_2}$ for some non-zero real numbers a and b is a non-causal filter since $h(-1,-1) \ne 0$.

2D IIR filter, causal but unstable:
$h(t_1,t_2) = a^{t_1}b^{t_2}u(t_1)u(t_2)$ for real number a and b such that $|a|>1$ and $b \ne 0$.
Consider the bounded input $x(t_1,t_2) = 1\,\forall t_1,t_2$. The output at time (t_1,t_2) is

$$y(t_1,t_2) = \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty} h(\tau_1,\tau_2)\,x(t_1-\tau,t_2-\tau) = \sum_{\tau_1=0}^{\infty}\sum_{\tau_2=0}^{\infty} a^{t_1}b^{t_2} = \left(\sum_{\tau_1=0}^{\infty} a^{\tau_1}\right)\left(\sum_{\tau_2=0}^{\infty} b^{\tau_2}\right) = \infty$$

Therefore, the filter is unstable. ∎

Problem 2. Show that

$$\begin{bmatrix}\mathbf{A} & \mathbf{A}_1\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix} \qquad \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix} \qquad \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix}$$

and

$$\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{A}_1 & \mathbf{A}_2\end{bmatrix} \qquad \begin{bmatrix}\underline{b}\\ \underline{q}\end{bmatrix} \qquad \begin{bmatrix}\underline{c}^T & \underline{0}^T\end{bmatrix}$$

and $\left(\mathbf{A},\, \underline{b},\, \underline{c}^T\right)$ all have the same transfer function, for all values of $\mathbf{A}_1$, $\mathbf{A}_2$ and $\underline{q}$ that leads to valid matrix operations. Conclude that realizations can have different numbers of states.

Solution. The transfer function of $\left(\mathbf{A},\, \underline{b},\, \underline{c}^T\right)$ system is

$$H_1(z) = \underline{c}^T \left(z\mathbf{I} - \mathbf{A}\right)^{-1} \underline{b} + d. \tag{1}$$

The transfer function of $\left(\begin{bmatrix}\mathbf{A} & \mathbf{A}_1\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix},\, \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix},\, \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix}\right)$ system is

$$\begin{aligned} H_2(z) &= \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix} \left(z\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} - \begin{bmatrix}\mathbf{A} & \mathbf{A}_1\\ \mathbf{0} & \mathbf{A}_2\end{bmatrix}\right)^{-1} \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix} + d \\ &= \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix} \begin{bmatrix}z\mathbf{I} - \mathbf{A} & -\mathbf{A}_1\\ \mathbf{0} & z\mathbf{I} - \mathbf{A}_2\end{bmatrix}^{-1} \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix} + d \end{aligned} \tag{2}$$

Using block-matrix multiplication, we have

$$\begin{bmatrix}z\mathbf{I} - \mathbf{A} & -\mathbf{A}_1\\ \mathbf{0} & z\mathbf{I} - \mathbf{A}_2\end{bmatrix} \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} & \left(\mathbf{zI} - \mathbf{A}\right)^{-\mathbf{1}} \mathbf{A_1} \left(\mathbf{zI} - \mathbf{A_2}\right)^{-\mathbf{1}}\\ \mathbf{0} & \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix}z\mathbf{I} - \mathbf{A} & -\mathbf{A}_1\\ \mathbf{0} & z\mathbf{I} - \mathbf{A}_2\end{bmatrix}^{-1} = \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} & \left(z\mathbf{I} - \mathbf{A}\right)^{-1} A_1 \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\\ \mathbf{0} & \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\end{bmatrix}.$$

Therefore, (1) can be written as

$$\begin{aligned} H_2(z) &= \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix} \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} & \left(\mathbf{zI} - \mathbf{A}\right)^{-\mathbf{1}} \mathbf{A_1} \left(\mathbf{zI} - \mathbf{A_2}\right)^{-\mathbf{1}}\\ \mathbf{0} & \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\end{bmatrix} \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix} + d \\ &= \begin{bmatrix}\underline{c}^T & \underline{q}^T\end{bmatrix} \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} \underline{b}\\ \underline{0}\end{bmatrix} + d \\ H_2(z) &= \underline{c}^T \left(z\mathbf{I} - \mathbf{A}\right)^{-1} \underline{b} + d. \end{aligned} \tag{3}$$

Similarly, the transfer function of $\left(\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{A}_1 & \mathbf{A}_2\end{bmatrix},\, \begin{bmatrix}\underline{b}\\ \underline{q}\end{bmatrix},\, \begin{bmatrix}\underline{c}^T & \underline{0}^T\end{bmatrix}\right)$ system is

$$\begin{aligned} H_3(z) &= \begin{bmatrix}\underline{c}^T & \underline{0}^T\end{bmatrix} \left(z\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix} - \begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{A}_1 & \mathbf{A}_2\end{bmatrix}\right)^{-1} \begin{bmatrix}\underline{b}\\ \underline{0}\end{bmatrix} + d \\ &= \begin{bmatrix}\underline{c}^T & \underline{0}^T\end{bmatrix} \begin{bmatrix}z\mathbf{I} - \mathbf{A} & \mathbf{0}\\ -\mathbf{A}_1 & z\mathbf{I} - \mathbf{A}_2\end{bmatrix}^{-1} \begin{bmatrix}\underline{b}\\ \underline{q}\end{bmatrix} + d \end{aligned} \tag{4}$$

We can also easily verify that

$$\begin{bmatrix}z\mathbf{I} - \mathbf{A} & -\mathbf{A}_1\\ \mathbf{0} & z\mathbf{I} - \mathbf{A}_2\end{bmatrix}^{-1} = \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} & \mathbf{0}\\ \left(z\mathbf{I} - \mathbf{A}\right)^{-1} A_1 \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1} & \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\end{bmatrix}.$$

Therefore, (4) can be written as

$$\begin{aligned} H_3(z) &= \begin{bmatrix}\underline{c}^T & \underline{0}^T\end{bmatrix} \begin{bmatrix}\left(z\mathbf{I} - \mathbf{A}\right)^{-1} & \mathbf{0}\\ \left(\mathbf{zI} - \mathbf{A}\right)^{-\mathbf{1}} \mathbf{A_1} \left(\mathbf{zI} - \mathbf{A_2}\right)^{-\mathbf{1}} & \left(z\mathbf{I} - \mathbf{A}_2\right)^{-1}\end{bmatrix} \begin{bmatrix}\underline{b}\\ \underline{q}^T\end{bmatrix} + d \\ &= \begin{bmatrix}\left(\underline{c}^T \left(z\mathbf{I} - \mathbf{A}\right)^{-1}\right) & \underline{0}^T\end{bmatrix} \begin{bmatrix}\underline{b}\\ \underline{q}\end{bmatrix} + d \\ H_3(z) &= \underline{c}^T \left(z\mathbf{I} - \mathbf{A}\right)^{-1} \underline{b} + d. \end{aligned} \tag{5}$$

Equations (1), (3) and (5) prove that the three systems have the same transfer function.

The number of states in a system is given by the dimension of the square matrix $\mathbf{A}$. Notice that the dimensions of the corresponding matrices for the other two systems is higher. Therefore, the same system can be represented using different numbers of states.

Remarks:

1. For block upper triangular matrix, $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$, the inverse is $\begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{C}^{-1} \end{bmatrix}$ whenever $\mathbf{A}^{-1}$ and $\mathbf{C}^{-1}$ exist. This can be easily verified by multiplying the two matrices. Also note that $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$ is non-singular if and only if $\mathbf{A}$ and $\mathbf{C}$ are non-singular. ■

Problem 3. (Modal Analysis) The following data is measured from a third-order system:

$$y = \{0.3200,\ 0.2500,\ 0.1000,\ -0.0222,\ 0.0006,\ -0.0012,\ 0.0005,\ -0.0001\}.$$

Assume that the first time index is 0, so that $y[0] = 0.32$.

(a) Determine the modes in the system, and plot them in the complex plane.

(b) The data can be written as

$$y[t] = c_1 (p_1)^t + c_2 (p_2)^t + c_3 (p_3)^t \quad t \geq 0.$$

Determine the constants c_1, c_2 and c_3.

(c) To explore the effect of noise on the system, add random Gaussian noise to each data point with variance $\sigma^2 = 0.01$, then find the modes of the noise data. Repeat several times (with different noise), and comment on how the modal estimates move.

Solution. Given that the system is 3^{rd} order system. Therefore, the AR equation for the system is

$$a_1 y[t-1] + a_2 y[t-2] + a_3 y[t-3] = y[t].$$

For the available 8 samples of data, we can write the above equation in matrix form as

$$\mathbf{Y}\underline{a} = \underline{y}$$

where

$$\mathbf{Y} = \begin{bmatrix} y[2] & y[1] & y[0] \\ y[3] & y[2] & y[1] \\ y[4] & y[3] & y[2] \\ y[5] & y[4] & y[3] \\ y[6] & y[5] & y[4] \end{bmatrix}, \quad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} y[3] \\ y[4] \\ y[5] \\ y[6] \\ y[7] \end{bmatrix}.$$

Since this is a over-determined set of equations, we consider the least squares solution given by

$$\underline{a} = \left(\mathbf{Y}^T\mathbf{Y}\right)^{-1}\mathbf{Y}^T\underline{y}.$$

Using MATLAB, we obtain the coefficients: $a_1 = -0.1755$, $a_2 = -0.0035$, $a_3 = -0.0118$

The modes of the system are the roots of the polynomial

$$a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} = 1.$$

The roots are obtained using MATLAB and the modes of the system are:

$$\begin{aligned} p_1 &= -0.2971, \\ p_2 &= 0.0608 + 0.1896j, \\ p_3 &= 0.0608 - 0.1896j. \end{aligned}$$

The data can be written as

$$y[t] = c_1 (p_1)^t + c_2 (p_2)^t + c_3 (p_3)^t \quad t \geq 0.$$

For the 8 data samples available, the above equation can be written as

$$\underline{y} = \mathbf{P}\underline{c}$$

where

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 \\ p_1 & p_2 & p_3 \\ (p_1)^2 & (p_2)^2 & (p_3)^2 \\ (p_1)^3 & (p_2)^3 & (p_3)^3 \\ (p_1)^4 & (p_2)^4 & (p_3)^4 \\ (p_1)^5 & (p_2)^5 & (p_3)^5 \\ (p_1)^6 & (p_2)^6 & (p_3)^6 \\ (p_1)^7 & (p_2)^7 & (p_3)^7 \end{bmatrix}, \quad \underline{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}, \quad \underline{y} = \begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ y[3] \\ y[4] \\ y[5] \\ y[6] \\ y[7] \end{bmatrix}.$$

The above set of equations is a overdetermined set of equations. Therefore we obtain the least squares solution given by

$$\underline{c} = \left(\mathbf{P}^T\mathbf{P}\right)^{-1}\mathbf{P}^T\underline{y}.$$

The coefficients are

$$\begin{aligned} c_1 &= 0.5015, \\ c_2 &= -0.0907 + 1.0813j, \\ c_3 &= -0.0907 + 1.0813j. \end{aligned}$$

The above procedure is repeated by adding noise to the data. Figure 1 shows the location of the system modes in the complex plane. We observe that the noise introduces a large variation in the estimation of system

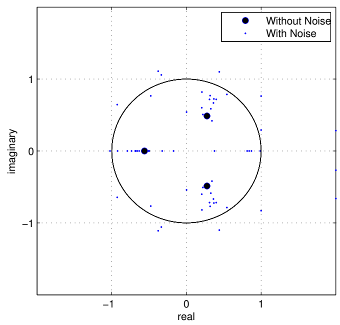

FIGURE 1. Modes of 3^{rd} order system plotted on a complex plane. The modes are estimated when the data samples do not have noise as well as when the data samples contain noise. The modes are estimated 20 times for the noisy case by adding different set of noise samples each time.

modes. This is because the variance of the Gaussian noise is much larger than some of the data samples and hence noise dominates the signal. In some cases, we even see the modes outside the unit circle i.e., the system is modeled to be unstable while the noiseless data samples indicate that the system is stable.

MATLAB Code:

```
y = [0.3200; 0.2500; 0.1000; -0.0222; 0.0006; -0.0012; 0.0005; -0.0001];
Y_mtx_indices = [3, 2, 1;...
                 4, 3, 2;...
                 5, 4, 3;...
                 6, 5, 4;...
                 7, 6, 5];
y_vec_indices = [4; 5; 6; 7; 8];

Y_mtx = y(Y_mtx_indices);
y_vec = y(y_vec_indices);

a = Y_mtx\y_vec
modes = roots([1; -a])

plot(modes, 'o', 'MarkerFaceColor', 'k'); hold on;
xlim([-2, 2]); ylim([-2, 2]);

P = [modes(1).^[0:7]', modes(2).^[0:7]', modes(3).^[0:7]']
c = P\y

for i=1:20
    y_noisy = y + 0.1*randn(8, 1);
    Y_mtx = y_noisy(Y_mtx_indices);
    y_vec = y_noisy(y_vec_indices);
    a = Y_mtx\y_vec;

    plot(roots([1; -a]), '.', 'MarkerSize', 5); hold on;
end

plot(cos((0:1000)*2*pi/100), sin((0:1000)*2*pi/100), 'k-')
legend('Without Noise', 'With Noise')
```

■

Problem 4. Random variables X and Y are uniformly distributed in the interval $[0, 1]$. Assuming that X and Y are independent, find the probability density function and the probability distribution function of a random variable $Z = |X - Y|$.

Solution. The p.d.fs of the random variables X and Y are

$$f_X(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} 1 & y \in [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

Since X and Y are independent, their joint p.d.f. is

$$f(x,y) = f_X(x) f_Y(y) = \begin{cases} 1 & x \in [0,1] \text{ and } y \in [0,1] \\ 0 & \text{otherwise.} \end{cases}$$

Since, $X \in [0, 1]$ and $Y \in [0, 1]$, $Z = |X - Y| \in [0, 1]$.

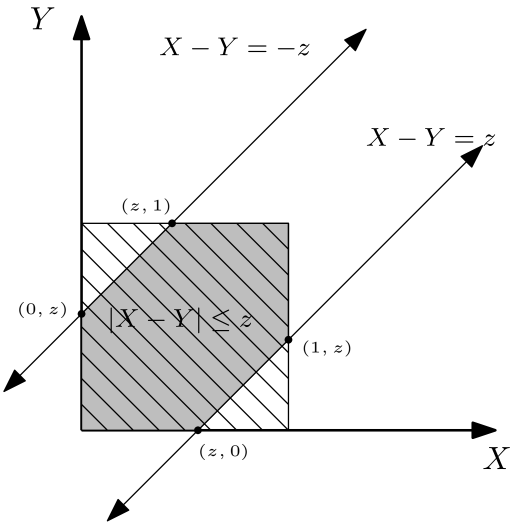

FIGURE 2. X and Y are independent and uniformly distributed on $[0, 1]$. The gray region corresponds to $Z = |X - Y| < z$.

From the figure, $Pr[Z \le z]$ is the area of the gray region. The area of each white triangle is $\frac{1}{2}(1-z)^2$. Therefore, the c.d.f of Z is

$$F_Z(z) = Pr[Z \le z] = \begin{cases} 0, & z < 0 \\ 1 - (1-z)^2 = 2z - z^2, & z \in [0,1] \\ 1, & z > 1 \end{cases}$$

We obtain the p.d.f. of Z by differentiating the c.d.f.:

$$f_Z(z) = \begin{cases} 2 - 2z, & z \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

■

Problem 5. The power spectral density of a certain sequence $x[n]$ is $\frac{1}{a+b\cos(\omega)}$ for some non-zero real constants a and b. Find the autocorrelation function. Suppose the autocorrelation function of a sequence $x[n]$ behaves as $r_{xx}(k) = \frac{1}{k}$ for time lags $k \geq 1$. What can you say about the power spectral density?

Solution. Power spectral density

$$S(\omega) = \frac{1}{a + b\cos(\omega)}.$$

The autocorrelation function can be obtained by taking inverse discrete-time Fourier transform of the power spectral density.

$$\begin{aligned} R_{xx}(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S(\omega)\, e^{j\omega k} d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{j\omega k}}{a + b\cos(\omega)} d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos(\omega k)}{a + b\cos(\omega)} d\omega + j\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(\omega k)}{a + b\cos(\omega)} d\omega \end{aligned}$$

Since $\frac{\sin(\omega k)}{a+b\cos(\omega)}$ is an odd function, the integral in the range $[-\pi, \pi]$ will be 0. Therefore,

$$R_{xx}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos(\omega k)}{a + b\cos(\omega)} d\omega.$$

Evaluating $R_{xx}(0)$:

$$R_{xx}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{a + b\cos(\omega)} d\omega$$

Substitute $t = \tan\left(\frac{\omega}{2}\right) \implies \cos(\omega) = \frac{1-t^2}{1+t^2}$ and $dt = \frac{1}{2}\sec^2\left(\frac{\omega}{2}\right) d\omega = \frac{1}{2}\left(1+t^2\right) d\omega$. We get,

$$\begin{aligned} R_{xx}(0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{2dt}{a(1+t^2) + b(1-t^2)} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dt}{(a+b) + t^2(a-b)} \end{aligned}$$

If $a = b$, $R_{xx}(0) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dt}{(a+b)} = \infty$.

For $a \neq b$,

$$\begin{aligned} R_{xx}(0) &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dt}{(a+b) + t^2(a-b)} \\ &= \frac{1}{\pi(a-b)}\int_{-\infty}^{\infty} \frac{dt}{\frac{(a+b)}{(a-b)} + t^2} \\ &= \frac{1}{\pi(a-b)}\left[\tan^{-1}\left(\sqrt{\frac{a-b}{a+b}}\, t\right)\right]_{-\infty}^{\infty} \qquad \left(\int \frac{dx}{c^2 + x^2} = \tan^{-1}\left(\frac{x}{c}\right)\right) \\ &= \frac{1}{\pi(a-b)}\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] \\ &= \frac{1}{a-b} \end{aligned} \tag{6}$$

Evaluating $R_{xx}(k)$, $k > 0, (a \neq b)$

$$R_{xx}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos(\omega k)}{a + b\cos(\omega)} d\omega$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\cos\left(\omega\left(k-1\right)\right)\cos\left(\omega\right)-\sin\left(\omega\left(k-1\right)\right)\sin\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \\
R_{xx}\left(k\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\cos\left(\omega\left(k-1\right)\right)\cos\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin\left(\omega\left(k-1\right)\right)\sin\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \qquad (7)
\end{aligned}
$$

We evaluate the two integrals separately as follows:

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\cos\left(\omega\left(k-1\right)\right)\cos\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{b}\cos\left(\omega\left(k-1\right)\right)\left(\frac{1}{a}-\frac{1}{a+b\cos\left(\omega\right)}\right)d\omega \\
&= \frac{1}{2\pi ab}\int_{-\pi}^{\pi}\cos\left(\omega\left(k-1\right)\right)d\omega-\frac{1}{2\pi a}\int_{-\pi}^{-\pi}\frac{\cos\left(\omega\left(k-1\right)\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{1}{2\pi ab}\int_{-\pi}^{\pi}\cos\left(\omega\left(k-1\right)\right)d\omega-\frac{1}{a}R_{xx}\left(k-1\right) \\
\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\cos\left(\omega\left(k-1\right)\right)\cos\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega &= \begin{cases}\frac{1}{ab}-\frac{1}{a}R_{xx}\left(k-1\right), & k=1\\ -\frac{1}{a}R_{xx}\left(k-1\right) & k>1\end{cases} \qquad (8)
\end{aligned}
$$

Let

$$
\begin{aligned}
f\left(k\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin\left(\omega\left(k-1\right)\right)\sin\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \\
\Longrightarrow f\left(1\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega=0. \quad \text{(Integrating odd function on } [-\pi,\pi]) \\
f\left(2\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin^2\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1-\cos^2\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{1}{2\pi b}\int_{-\pi}^{\pi}\frac{b-b\cos^2\left(\omega\right)-a\cos\left(\omega\right)+a\cos\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{1}{2\pi b}\int_{-\pi}^{\pi}\frac{b+a\cos\left(\omega\right)-\cos\left(\omega\right)\left(a+b\cos\left(\omega\right)\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{1}{2\pi b}\int_{-\pi}^{\pi}\frac{b+a\cos\left(\omega\right)}{a+b\cos\left(\omega\right)}d\omega-\frac{1}{2\pi b}\int_{-\pi}^{\pi}\cos\left(\omega\right)d\omega \\
&= \frac{1}{2\pi b^2}\int_{-\pi}^{\pi}\frac{b^2+ab\cos\left(\omega\right)+a^2-a^2}{a+b\cos\left(\omega\right)}d\omega-0 \\
&= \frac{1}{2\pi b^2}\int_{-\pi}^{\pi}\frac{b^2-a^2+a\left(a+b\cos\left(\omega\right)\right)}{a+b\cos\left(\omega\right)}d\omega \\
&= \frac{b^2-a^2}{b^2}\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{a+b\cos\left(\omega\right)}d\omega+\frac{a}{2\pi b^2}\int_{-\pi}^{\pi}d\omega \\
&= \frac{b^2-a^2}{b^2}\frac{1}{a-b}+\frac{a}{b^2} \\
f\left(2\right) &= -\frac{1}{b}-\frac{a}{b}+\frac{a}{b^2}
\end{aligned}
$$

For $k > 2$,

$$\begin{aligned} f(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(\omega(k-1))\sin(\omega)}{a+b\cos(\omega)}d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(\omega)}{a+b\cos(\omega)}\left(\sin(\omega(k-2))\sin(\omega)+\cos(\omega(k-2))\cos(\omega)\right)d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(\omega(k-2))\sin^2(\omega)}{a+b\cos(\omega)}d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos(\omega(k-2))\cos(\omega)\sin(\omega)}{a+b\cos(\omega)}d\omega \end{aligned}$$

$\frac{\cos(\omega(k-2))\cos(\omega)\sin(\omega)}{a+b\cos(\omega)}$ is odd function $\forall k$. Therefore it's integral in the range $[-\pi,\pi]=0$. The term $\frac{\sin(\omega(k-2))\sin^2(\omega)}{a+b\cos(\omega)}$ is odd function for $k > 2$. Therefore,

$$f(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin(\omega(k-1))\sin(\omega)}{a+b\cos(\omega)}d\omega = \begin{cases} 0, & k=1 \\ -\frac{1}{b}-\frac{a}{b}+\frac{a}{b^2}, & k=2 \\ 0, & k>2 \end{cases} \tag{9}$$

From (6), (7), (8) and (9), we have

$$\begin{aligned} R_{xx}(k) &= \begin{cases} \frac{1}{a-b}, & k=0 \\ \frac{1}{ab}-\frac{1}{a}R_{xx}(k-1), & k=1 \\ -\frac{1}{a}R_{xx}(k-1)+\frac{1}{b}+\frac{a}{b}-\frac{a}{b^2}, & k=2 \\ -\frac{1}{a}R_{xx}(k-1), & k>2 \end{cases} \\ \implies R_{xx}(1) &= \frac{1}{ab}-\frac{1}{a(a-b)} \\ R_{xx}(2) &= -\frac{1}{ab}+\frac{1}{a(a-b)}+\frac{1}{b}+\frac{a}{b}-\frac{a}{b^2} \\ R_{xx}(k) &= R_{xx}(2)\left(-\frac{1}{a}\right)^{k-2} \quad k>2. \\ R_{xx}(k) &= R_{xx}(-k), \quad k<0. \end{aligned}$$

Part 2:

If $r_{xx}(k)=\frac{1}{k}$ for $k \geq 1$, the Fourier transform of $r_{xx}(k)$ does not exist since $\sum_{i=1}^{\infty}|r_{xx}(k)|=\infty$. The system is unstable. ■

Problem 6. Suppose we are filtering a random sequence $x[n]$ through a FIR filter $1 - az^{-1}$, $|a| < 1$. Let $x[n]$ be a Bernoulli process such that $P(x[n] = 1) = p$ and $P(x[n] = 0) = 1 - p$. Examine if this is a wide sense stationary process and ergodic in mean.

Solution. $x[n]$ is a Bernoulli process. Therefore, $x[n_1]$ and $x[n_2]$ are independent for $n_1 \neq n_2$. The mean and second moments of the process are

$$\mathbb{E}[x[n]] = p \times 1 + (1-p) \times 0 = p \,\forall n,$$

$$\mathbb{E}[x[n]\,x[n]] = p \times 1^2 + (1-p) \times 0^2 = p \,\forall n.$$

For $k \neq 0$,

$$\begin{aligned} \mathbb{E}[x[n]\,x[n+k]] &= p \times p \times 1 + (1-p) \times p \times 0 + (1-p) \times p \times 0 + (1-p) \times (1-p) \times 0 \\ &= p^2. \end{aligned}$$

Therefore,

$$R_{XX}(k) = \mathbb{E}[x[n]\,x[n-k]] = \begin{cases} p, & k = 0 \\ p^2, & k \neq 0. \end{cases}$$

For the FIR filter $1 - az^{-1}$, the output is given by

$$y[n] = x[n] - ax[n-1].$$

Therefore, the mean is

$$\mu_Y(n) = \mathbb{E}[y[n]] = \mathbb{E}[x[n]] - a\mathbb{E}[x[n-1]] = (1-a)\,p.$$

The autocorrelation is

$$\begin{aligned} \mathbb{E}[y[n]\,y[n-k]] &= \mathbb{E}[x[n]\,x[n-k]] - a\mathbb{E}[x[n]\,x[n-k-1]] - a\mathbb{E}[x[n-1]\,x[n-k]] + a^2\mathbb{E}[x[n-1]\,x[n-k-1]] \\ &= \begin{cases} p^2 - ap - ap^2 + a^2p^2, & k = -1 \\ p - ap^2 - ap^2 + a^2p, & k = 0 \\ p^2 - ap^2 - ap + a^2p^2, & k = 1 \\ p^2 - ap^2 - ap^2 + a^2p^2, & \text{otherwise.} \end{cases} \end{aligned}$$

Since, $\mu_Y(n) = \mu_Y \forall n$ and $\mathbb{E}[y[n]\,y[n-k]] = R_{YY}(k) \forall n$, the process is W.S.S.

The time average of $y[n]$ is

$$\begin{aligned} \hat{y} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} y[i] \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (x[i] - ax[i-1]) \\ &= \lim_{N\to\infty} \frac{1}{N} (x[N] - ax[0]) + \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N-1} (x[i] - ax[i]) \\ &= 0 + (1-a) \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N-1} x[i] \\ &= (1-a) \left(\lim_{N\to\infty} \frac{N-1}{N} \right) \lim_{N\to\infty} \frac{1}{N-1} \sum_{i=1}^{N-1} x[i] \\ &= (1-a) \lim_{N\to\infty} \frac{1}{N-1} \sum_{i=1}^{N-1} x[i] \\ \hat{y} &= (1-a)\,\hat{x} \end{aligned}$$

where $\hat{x}$ is the time average of the Bernoulli process $x[n]$. Since the Bernoulli process is ergodic, $\hat{x} = \mu_X = p$. Therefore, $\hat{y} = (1-a)\,\hat{x} = (1-a)\,p$. Hence, the process $y[n]$ is ergodic in mean. ■

Problem 7. Consider a collection of all $n \times n$ matrices with real entries i.e., $M_n(\mathbb{R})$. Is this a vector space? Justify.

(1) Suppose we consider $\mathcal{S} := \{\mathbf{X} \in M_n(\mathbb{R}) : \det(\mathbf{X}) = 0\}$, examine if $\mathcal{S}$ is a subspace.

(2) Suppose $\mathbf{P}$, $\mathbf{X} \in M_n(\mathbb{R})$. Let T be an operator such that $T(\mathbf{X}) = \mathbf{P}^T\mathbf{X}\mathbf{P}$ for a fixed matrix $\mathbf{P}$. Examine if T is linear.

Solution. Yes, $M_n(\mathbb{R})$ is a vector space:

1) Closed under addition: Adding two real matrices results in a matrix with real entries.

2) Identity element: Defining $\mathbf{0}_n$ as $n \times n$ matrix with all elements as 0, $\mathbf{X} + \mathbf{0}_n = \mathbf{X} \forall \mathbf{X} \in M_n(\mathbb{R})$

3) Inverse element: $\mathbf{X} \in M_n(\mathbb{R})$, we can define a matrix $\mathbf{Y} = -\mathbf{X}$ by flipping the signs of the elements of $\mathbf{Y}$. In this case, $\mathbf{X} + \mathbf{Y} = \mathbf{0}_n$.

4) Associative: Since the matrix addition is achieved by addition of individual elements, associativity of $M_n(\mathbb{R})$ follows from the associativity of real numbers over addition.

5) Closed under scalar multiplication: Multiplying a real matrix by a scalar will result in a real matrix.

6) Following three properties follow from the associative and distributive laws of real numbers when applied to individual elements of the matrices $\mathbf{X}, \mathbf{Y} \in M_n(\mathbb{R})$ (a) $a(b\mathbf{X}) = (ab)\mathbf{X}$ (b) $(a+b)\mathbf{X} = a\mathbf{X} + b\mathbf{X}$ (c) $a(\mathbf{X} + \mathbf{Y}) = a\mathbf{X} + a\mathbf{Y}$ $\forall a, b \in \mathbb{R}$.

7) Multiplicative element $1 \in \mathbb{R}$ satisfies $1.\mathbf{X} = \mathbf{X} \forall \mathbf{X} \in M_n(\mathbb{R})$.

Part 1:

$\mathcal{S} := \{\mathbf{X} \in M_n(\mathbb{R}) : \det(\mathbf{X}) = 0\}$

For $n = 1$, $\mathcal{S} = \{[0]\}$ is a trivial subspace.

For $n > 1$, S is not a subspace because $\mathcal{S}$ is not closed under addition. Example: $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{S}$, $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \in \mathcal{S}$. But $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \notin \mathcal{S}$.

Part 2:

Using the left and right distributive laws of matrix multiplication,

$$T(\mathbf{X} + \mathbf{Y}) = \mathbf{P}^T(\mathbf{X} + \mathbf{Y})\mathbf{P} = \mathbf{P}^T(\mathbf{X}\mathbf{P} + \mathbf{Y}\mathbf{P}) = \mathbf{P}^T\mathbf{X}\mathbf{P} + \mathbf{P}^T\mathbf{Y}\mathbf{P} = T(\mathbf{X}) + T(\mathbf{Y}).$$

Also, using commutative law matrix multiplication by scalar ($a\mathbf{X} = \mathbf{X}a$), we have

$$T(a\mathbf{X}) = \mathbf{P}^T(a\mathbf{X})\mathbf{P} = a\mathbf{P}^T\mathbf{X}\mathbf{P} = aT(\mathbf{X}).$$

Therefore, the operator $T(\mathbf{X})$ is linear. ■

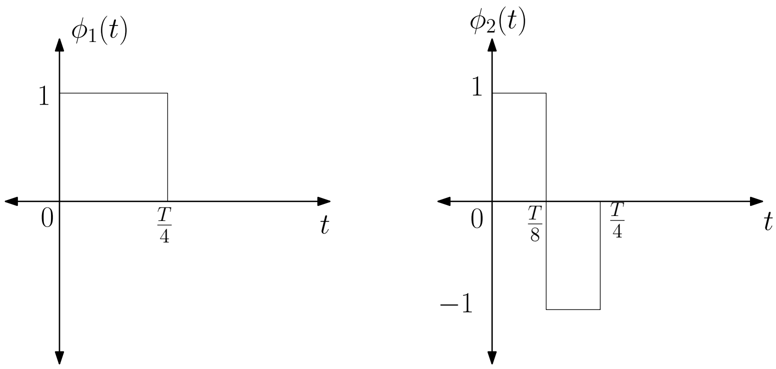

FIGURE 3. Signals $\phi_1(t)$ and $\phi_2(t)$

Problem 8. Consider the continuous time signal $f(t) = \sum_{i=1}^{N-1} a_i u\left(t - i\frac{T}{8}\right)$ where a_i is any real number, T is a time unit and N is a positive integer. Let $\phi_1(t) = u(t) - t\left(t - \frac{T}{4}\right)$ and $\phi_2(t) = u(t) - 2u\left(t - \frac{T}{8}\right) + u\left(t - \frac{T}{4}\right)$.

(1) Are $\phi_1(t)$ and $\phi_2(t)$ linearly independent?

(2) Expand the signal $f(t)$ in the $\phi_1(t) - \phi_2(t)$ plane after normalizing $\phi_1(t)$ and $\phi_2(t)$. Interpret your results graphically.

(3) Suppose a source emits $\phi_1(t)$ and $\phi_2(t)$ randomly with source probabilities p and $1-p$ respectively. Imagine a cloud of uncorrelated Gaussian noise $\mathcal{N}\left(0, \sigma^2\right)$ acting in the $\phi_1(t) - \phi_2(t)$ plane. Determine the optimal linear decision boundaries to minimize the probability of misclassifying the signals $\phi_1(t)$ and $\phi_2(t)$. Explicitly evaluate the probability of misclassification.

Solution. Given

$$\begin{aligned}
\phi_1(t) &= \begin{cases} 1, & 0 \le t \le \frac{T}{4} \\ 0, & \text{otherwise,} \end{cases} \\
\phi_2(t) &= \begin{cases} 1, & 0 \le t \le \frac{T}{8} \\ -1, & \frac{T}{8} \le t \le \frac{T}{4} \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Part 1:

$$\begin{aligned}
\langle \phi_1(t), \phi_2(t) \rangle &= \int_{-\infty}^{\infty} \phi_1(t)\, \phi_2(t)\, dt \\
&= \int_0^{\frac{T}{8}} dt - \int_{\frac{T}{8}}^{\frac{T}{4}} dt \\
&= 0.
\end{aligned}$$

The signals $\phi_1(t)$ and $\phi_2(t)$ are orthogonal. Therefore, the signals are linearly independent.

Part 2:

Since $\phi_1(t)$ and $\phi_2(t)$ are orthogonal, we obtain the orthonormal basis of the signal space by normalizing the signals $\phi_1(t)$ and $\phi_2(t)$.

$$\begin{aligned}
\langle \phi_1(t), \phi_1(t) \rangle &= \int_0^{\frac{T}{4}} dt = \frac{T}{4} \\
\langle \phi_2(t), \phi_2(t) \rangle &= \int_0^{\frac{T}{4}} dt = \frac{T}{4}.
\end{aligned}$$

The orthonormal basis is given by

$$\hat{\phi}_1(t) = \frac{\phi_1(t)}{\sqrt{\langle \phi_1(t), \phi_1(t) \rangle}} = \frac{2}{\sqrt{T}} \phi_1(t) = \begin{cases} \frac{2}{\sqrt{T}}, & 0 \le t \le \frac{T}{4} \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{\phi}_2(t) = \frac{\phi_2(t)}{\sqrt{\langle \phi_2(t), \phi_2(t)\rangle}} = \frac{2}{\sqrt{T}}\phi_2(t) = \begin{cases} \frac{2}{\sqrt{T}}, & 0 \le t \le \frac{T}{8} \\ -\frac{2}{\sqrt{T}}, & \frac{T}{8} \le t \le \frac{T}{4} \\ 0, & \text{otherwise.} \end{cases}$$

We can write $f(t)$ in terms of the shifted versions of the orthogonal basis. We shift the orthogonal bases by integer multiples of $\frac{T}{8}$. We have

$$\hat{\phi}_1\left(t - k\frac{T}{8}\right) = \begin{cases} \frac{2}{\sqrt{T}}, & k\frac{T}{8} \le t \le (k+2)\frac{T}{8} \\ 0, & \text{otherwise,} \end{cases}$$

$$\hat{\phi}_2\left(t - k\frac{T}{8}\right) = \begin{cases} \frac{2}{\sqrt{T}}, & k\frac{T}{8} \le t \le (k+1)\frac{T}{8} \\ -\frac{2}{\sqrt{T}}, & (k+1)\frac{T}{8} \le t \le (k+2)\frac{T}{4} \\ 0, & \text{otherwise.} \end{cases}$$

We also have

$$f(t) = \sum_{i=1}^{N-1} a_i u\left(t - i\frac{T}{8}\right) = \begin{cases} 0, & t < \frac{T}{8} \\ \sum_{i=1}^{n} a_i, & n\frac{T}{8} \le t < (n+1)\frac{T}{8}, n = 1, 2, \cdots, N-1 \\ \sum_{i=1}^{N-1} a_i & k > (N-1) \end{cases}$$

We now identify the projections $f(t)$ on $\hat{\phi}_1\left(t - k\frac{T}{8}\right), \hat{\phi}_2\left(t - k\frac{T}{8}\right)$

$$\begin{aligned} \left\langle f(t), \hat{\phi}_1\left(t - k\frac{T}{8}\right)\right\rangle &= \frac{2}{\sqrt{T}} \int_{k\frac{T}{8}}^{(k+2)\frac{T}{8}} f(t)\,dt \\ &= \frac{2}{\sqrt{T}} \int_{k\frac{T}{8}}^{(k+1)\frac{T}{8}} f(t)\,dt + \frac{2}{\sqrt{T}} \int_{(k+1)\frac{T}{8}}^{(k+2)\frac{T}{8}} f(t)\,dt \\ &= \frac{2}{\sqrt{T}}\left(f\left(k\frac{T}{8}\right) \int_{k\frac{T}{8}}^{(k+1)\frac{T}{8}} dt + f\left((k+1)\frac{T}{8}\right) \int_{(k+1)\frac{T}{8}}^{(k+2)\frac{T}{8}} dt \right) \\ &= \frac{2}{\sqrt{T}}\left(f\left(k\frac{T}{8}\right)\frac{T}{8} + f\left((k+1)\frac{T}{8}\right)\frac{T}{8}\right) \\ &= \frac{\sqrt{T}}{4}\left(f\left(k\frac{T}{8}\right) + f\left((k+1)\frac{T}{8}\right)\right) \\ &= \begin{cases} 0, & k < 0 \\ \frac{\sqrt{T}}{4}a_1, & k = 0 \\ \frac{\sqrt{T}}{4}\left(\sum_{i=1}^{k} a_i + \sum_{i=1}^{k+1} a_i\right), & 1 \le k \le N-2 \\ \frac{\sqrt{T}}{4}\left(\sum_{i=1}^{N-1} a_i + \sum_{i=1}^{N-1} a_i\right) & k \ge N-1 \end{cases} \\ \left\langle f(t), \hat{\phi}_1\left(t - k\frac{T}{8}\right)\right\rangle &= \begin{cases} 0, & k < 0 \\ \frac{\sqrt{T}}{4}a_1, & k = 0 \\ \frac{\sqrt{T}}{4}\left(a_{k+1} + 2\sum_{i=1}^{k} a_i\right), & 1 \le k \le N-2 \\ \frac{\sqrt{T}}{2}\sum_{i=1}^{N-1} a_i & k \ge N-1 \end{cases} \end{aligned}$$

Similarly,

$$\begin{aligned} \left\langle f(t), \hat{\phi}_2\left(t - k\frac{T}{8}\right)\right\rangle &= \frac{2}{\sqrt{T}} \int_{k\frac{T}{8}}^{(k+1)\frac{T}{8}} f(t)\,dt - \frac{2}{\sqrt{T}} \int_{(k+1)\frac{T}{8}}^{(k+2)\frac{T}{8}} f(t)\,dt \\ &= \frac{2}{\sqrt{T}}\left(f\left(k\frac{T}{8}\right) \int_{k\frac{T}{8}}^{(k+1)\frac{T}{8}} dt - f\left((k+1)\frac{T}{8}\right) \int_{(k+1)\frac{T}{8}}^{(k+2)\frac{T}{8}} dt \right) \\ &= \frac{2}{\sqrt{T}}\left(f\left(k\frac{T}{8}\right)\frac{T}{8} - f\left((k+1)\frac{T}{8}\right)\frac{T}{8}\right) \end{aligned}$$

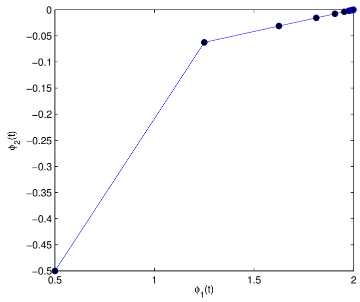

FIGURE 4. Trajectory of $f(t)$ in $\hat{\phi}_1(t) - \hat{\phi}_2(t)$ plane. The trajectory is obtained by projecting $f(t)$ on to $\hat{\phi}_1(t-\tau)$ and $\hat{\phi}_2(t-\tau)$ i.e., the shifted versions of the orthonormal basis.

$$= \frac{\sqrt{T}}{4}\left(f\left(k\frac{T}{8}\right) - f\left((k+1)\frac{T}{8}\right)\right)$$

$$= \begin{cases} 0, & k<0 \\ -\frac{\sqrt{T}}{4}a_1, & k=0 \\ \frac{\sqrt{T}}{4}\left(\sum_{i=1}^{k} a_i - \sum_{i=1}^{k+1} a_i\right), & 1 \le k \le N-2 \\ \frac{\sqrt{T}}{4}\left(\sum_{i=1}^{N-1} a_i - \sum_{i=1}^{N-1} a_i\right) & k \ge N-1 \end{cases}$$

$$\left\langle f(t), \hat{\phi}_2\left(t - k\frac{T}{8}\right)\right\rangle = \begin{cases} 0, & k<0 \\ -\frac{\sqrt{T}}{4}a_1, & k=0 \\ -\frac{\sqrt{T}}{4}a_{k+1}, & 1 \le k \le N-2 \\ 0 & k \ge N-1 \end{cases}$$

Therefore, in the $\phi_1(t) - \phi_2(t)$, the signal follows the points (x_k, y_k) given by

$$(x_k, y_k) = \begin{cases} (0,0), & k<0 \\ \left(\frac{\sqrt{T}}{4}a_1, -\frac{\sqrt{T}}{4}a_1\right), & k=0 \\ \left(\frac{\sqrt{T}}{4}\left(a_{k+1} + 2\sum_{i=1}^{k-1} a_i\right), -\frac{\sqrt{T}}{4}a_{k+1}\right), & 1 \le k \le N-2 \\ \left(\frac{\sqrt{T}}{2}\sum_{i=1}^{N-1} a_i, 0\right) & k \ge N-1 \end{cases}$$

Since, we have integrated over step functions, the trajectory between the points is obtained using linear interpolation.

Figure 4 shows an example trajectory considering $a_i = \frac{1}{2^i}$ and $T = 16$ units. The path follows the points given by,

$$(x_k, y_k) = \begin{cases} (0,0), & k<0 \\ \left(\frac{1}{2}, -\frac{1}{2}\right), & k=0 \\ \left(\left(\frac{1}{2^{k+1}} + 2\sum_{i=1}^{k} \frac{1}{2^i}\right), -\frac{1}{2^{k+1}}\right), & 1 \le k \le N-2 \\ \left(2\sum_{i=1}^{N-1} \frac{1}{2^i}, 0\right) & k \ge N-1 \end{cases}$$

$$= \begin{cases} (0,0), & k<0 \\ \left(\frac{1}{2}, -\frac{1}{2}\right), & k=0 \\ \left(\left(\frac{1}{2^{k+1}} + 2\left(1 - \frac{1}{2^k}\right)\right), -\frac{1}{2^{k+1}}\right), & 1 \le k \le N-2 \\ \left(2\left(1 - \frac{1}{2^{N-1}}\right), 0\right) & k \ge N-1 \end{cases}$$

$$(x_k, y_k) = \begin{cases} (0,0), & k<0 \\ \left(\frac{1}{2}, -\frac{1}{2}\right), & k=0 \\ \left(2 - \frac{3}{2^{k+1}}, -\frac{1}{2^{k+1}}\right), & 1 \le k \le N-2 \\ \left(2 - \frac{1}{2^{N-2}}, 0\right) & k \ge N-1 \end{cases}$$

Part 3:

The signals $\phi_1(t)$ and $\phi_2(t)$ correspond to the points $\underline{s}_1 = \left(\frac{\sqrt{T}}{2}, 0\right)$ and $\underline{s}_2 = \left(0, \frac{\sqrt{T}}{2}\right)$ on the $\hat{\phi}_1(t) - \hat{\phi}_2(t)$ respectively. Let $\underline{S} = (S_x, S_y)$ indicate the signal transmitted by source. It is given that

$$Pr\left[\underline{S} = \underline{s}_1\right] = p$$

$$Pr\left[\underline{S}=\underline{s}_2\right] = 1-p$$

Let Z_1 and Z_2 be the noise values for the $\phi_1(t)$ and $\phi_1(t)$ coordinates. Given that $Z_1 \sim \mathcal{N}\left(0,\sigma^2\right)$ and $Z_2 \sim \mathcal{N}\left(0,\sigma^2\right)$. Z_1 and Z_2 are also independent.

Let the received noisy signal be represented by the vector

$$\underline{R} = (X,Y) = (S_x+Z_1, S_y+Z_2)$$

Therefore, the conditional p.d.f.s of X and Y given the transmitted signal are

$$\begin{aligned} f_X\left(x \mid \underline{S}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-S_x)^2/2\sigma^2}, \\ f_Y\left(y \mid \underline{S}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-S_y)^2/2\sigma^2}. \end{aligned}$$

Since Z_1 and Z_2 are independent, the joint p.d.f. is given by

$$\begin{aligned} f_{X,Y}\left(x,y \mid \underline{S}\right) &= \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left((x-S_x)^2+(y-S_y)^2\right)} \\ \Longrightarrow f_{X,Y}\left(x,y \mid \underline{S}=\underline{s}_1\right) &= \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(\left(x-\frac{\sqrt{T}}{2}\right)^2+y^2\right)} \\ \Longrightarrow f_{X,Y}\left(x,y \mid \underline{S}=\underline{s}_2\right) &= \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(x^2+\left(y-\frac{\sqrt{T}}{2}\right)^2\right)} \end{aligned}$$

Using Bayes theorem, the conditional probability (*a-posteriori* probability) of $\underline{s}_i, i=1,2$ being transmitted given that we have received the vector $\underline{R}=(x,y)$ is

$$T_i(x,y) = Pr\left(\underline{S}=\underline{s}_i \mid x,y\right) = \frac{f\left(x,y \mid \underline{S}=\underline{s}_i\right) Pr\left(\underline{S}=\underline{s}_i\right)}{f(x,y)},\ i=1,2. \tag{10}$$

The optimal decision is to maximize the above *a-posteriori* probability i.e., we make a decision that $\underline{s}_1$ is transmitted is $T_1(x,y) \geq T_2(x,y)$ for the received vector (x,y). Therefore, the optimal decision regions for $\underline{s}_1$ and $\underline{s}_2$ are R_1 and R_2 defined as

$$\begin{aligned} R_1 &= \left\{(x,y) \mid T_1(x,y) \geq T_2(x,y)\right\}, \\ R_2 &= \left\{(x,y) \mid T_1(x,y) < T_2(x,y)\right\}. \end{aligned}$$

From (10),

$$\begin{aligned} T_1(x,y) &\geq T_2(x,y) \\ \Longrightarrow f\left(x,y \mid \underline{S}=\underline{s}_1\right) Pr\left(\underline{S}=\underline{s}_1\right) &\geq f\left(x,y \mid \underline{S}=\underline{s}_2\right) Pr\left(\underline{S}=\underline{s}_2\right) \\ \Longrightarrow \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(\left(x-\frac{\sqrt{T}}{2}\right)^2+y^2\right)} \times p &\geq \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(x^2+\left(y-\frac{\sqrt{T}}{2}\right)^2\right)} \times (1-p) \\ \Longrightarrow -\frac{1}{2\sigma^2}\left(\left(x-\frac{\sqrt{T}}{2}\right)^2+y^2\right) + \ln\left(\frac{p}{1-p}\right) &\geq -\frac{1}{2\sigma^2}\left(x^2+\left(y-\frac{\sqrt{T}}{2}\right)^2\right) \\ \Longrightarrow 2\sigma^2 \ln\left(\frac{p}{1-p}\right) &\geq \left(x-\frac{\sqrt{T}}{2}\right)^2 - x^2 + y^2 - \left(y-\frac{\sqrt{T}}{2}\right)^2 \\ \Longrightarrow 2\sigma^2 \ln\left(\frac{p}{1-p}\right) &\geq -\sqrt{T}x + \sqrt{T}y \\ \Longrightarrow x-y &\geq -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right) \end{aligned}$$

Therefore, the decision boundary is the line $x-y = -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)$ and the decision regions are

$$\begin{aligned} R_1 &= \left\{(x,y) \mid x-y \geq -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right\} \\ R_2 &= \left\{(x,y) \mid x-y < -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right\}. \end{aligned}$$

The misclassification occurs when $\underline{s}_1$ is transmitted but the received vector $(x,y) \in R_2$ or when $\underline{S}=\underline{s}_2$ and $(x,y) \in R_1$.

Figure 5 shows the Gaussian cloud and the linear decision boundary for signal classification.

If $\hat{\underline{S}}$ denotes the decision made, the probability of misclassification is

$$Pr\left[\hat{\underline{S}} \neq \underline{S}\right] = Pr\left[\hat{\underline{S}} \neq \underline{s}_1 \mid \underline{S}=\underline{s}_1\right] Pr\left[\underline{S}=\underline{s}_1\right]$$

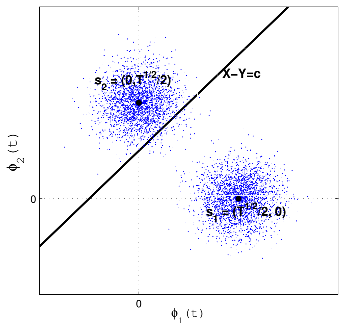

FIGURE 5. The signals $\underline{s}_1 = \phi_1(t)$ and $\underline{s}_2 = \phi_2(t)$ are represented in the $\hat{\phi}_1(t) - \hat{\phi}_2(t)$ plane. The Gaussian cloud around the points $\underline{s}_1$ and $\underline{s}_2$ is due to the noise. The optimal linear decision boundary is of the form $X - Y = c$.

$$\begin{aligned} & +Pr\left[\hat{\underline{S}} \neq \underline{s}_2 \mid \underline{S} = \underline{s}_2\right] Pr\left[\underline{S} = \underline{s}_2\right] \\ = \; & Pr\left[(x,y) \in R_2 \mid \underline{S} = \underline{s}_1\right] \times p \\ & +Pr\left[(x,y) \in R_1 \mid \underline{S} = \underline{s}_2\right] \times (1-p) \\ = \; & Pr\left[X - Y < -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right) \mid \underline{S} = \underline{s}_1\right] \times p \\ & +Pr\left[X - Y \geq -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right) \mid \underline{S} = \underline{s}_1\right] \times (1-p) \\ = \; & Pr\left[Z_1 + \frac{\sqrt{T}}{2} - Z_2 < -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right] \times p \\ & +Pr\left[Z_1 - Z_2 - \frac{\sqrt{T}}{2} \geq -\frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right] \times (1-p) \\ Pr\left[\hat{\underline{S}} \neq \underline{S}\right] = \; & Pr\left[Z_1 - Z_2 < -\frac{\sqrt{T}}{2} - \frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right] \times p \\ & +Pr\left[Z_1 - Z_2 \geq \frac{\sqrt{T}}{2} - \frac{2\sigma^2}{\sqrt{T}} \ln\left(\frac{p}{1-p}\right)\right] \times (1-p) \qquad (11) \end{aligned}$$

Since Z_1 and Z_2 are Gaussian random variables, their linear combination $Z_1 - Z_2$ is also a Gaussian random variable with mean and variance given by

$$\begin{aligned} \mathbb{E}\left[Z_1 - Z_2\right] &= 0 \\ \mathrm{Var}\left(Z_1 - Z_2\right) &= \mathbb{E}\left[\left(Z_1 - Z_2 - 0\right)^2\right] \\ &= \mathbb{E}\left[Z_1^2\right] + \mathbb{E}\left[Z_2^2\right] - 2\mathbb{E}\left[Z_1 Z_2\right] \\ &= \sigma^2 + \sigma^2 - 2\mathbb{E}\left[Z_1\right]\mathbb{E}\left[Z_2\right] \\ \mathrm{Var}\left(Z_1 - Z_2\right) &= 2\sigma^2. \end{aligned}$$

Therefore, $Z_1 - Z_2 \sim \mathcal{N}\left(0, 2\sigma^2\right)$ and

$$\begin{aligned} Pr\left[Z_1 - Z_2 \geq c\right] &= \int_c^\infty \frac{1}{\sqrt{4\pi\sigma^2}} e^{-x^2/4\sigma^2} dx \\ &= \int_{\frac{c}{\sqrt{2}\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \qquad \left(\text{Using } t = \frac{x}{\sqrt{2}\sigma}\right) \\ &= Q\left(\frac{c}{\sqrt{2}\sigma}\right) \end{aligned}$$

where $Q(\cdot)$ is the probability in the tail of Gaussian function given by

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

We can also show that $Z_2 - Z_1 \sim \mathcal{N}\left(0, 2\sigma^2\right)$ and hence $Pr\left[Z_1 - Z_2 < -c\right] = Pr\left[Z_2 - Z_1 > c\right] = Q\left(\frac{c}{\sqrt{2}\sigma}\right)$. Therefore, from (11), the probability of misclassification is

$$\begin{aligned} P_e &= p \times Q\left(\frac{1}{\sqrt{2}\sigma}\left(\frac{\sqrt{T}}{2} + \frac{2\sigma^2}{\sqrt{T}}\ln\left(\frac{p}{1-p}\right)\right)\right) \\ &\quad + (1-p) \times Q\left(\frac{1}{\sqrt{2}\sigma}\left(\frac{\sqrt{T}}{2} - \frac{2\sigma^2}{\sqrt{T}}\ln\left(\frac{p}{1-p}\right)\right)\right) \end{aligned}$$

■