

Solution Guide

for

**Theory and Applications of Digital
Speech Processing**

by

Lawrence R. Rabiner

**Rutgers University
and**

University of California at Santa Barbara

and

**Ronald W. Schafer
Hewlett-Packard Laboratories**

INTRODUCTION

This Solution Guide contains solutions to all of the problems in *Theory and Applications of Digital Speech Processing* that do not have the label “(MATLAB Exercise)”. The problems so designated are project type exercises that are designed to give students hands-on experience in programming digital speech processing algorithms and systems. We chose not to include solutions to these projects, reasoning that in most cases, it will be obvious to both the student and the instructor whether the student’s program works. Furthermore, in such problems, which are closer to real applications, there will be no single best solution. In a few cases, a problem simply requires that a plot be constructed using MATLAB. In these cases, a solution is given.

Most of the problems have been assigned in class or given on exams in courses taught by the authors. While we have made every effort to provide correct and accurate solutions, this does not mean that more elegant solutions cannot be found. Furthermore, while we have proofread the solutions very carefully, it is inevitable that we have missed some errors. We welcome suggestions for improving the solutions.

It is our firm belief that students learn a subject best if they work problems without having a solution to refer to. This is why we have made the solution guide available only to instructors in classes using *Theory and Applications of Digital Speech Processing* as the course text. We request that instructors use discretion when posting answers to class assignments on the Web.

Contents

2 Review of Fundamentals of Digital Signal Processing	3
3 Fundamentals of Human Speech Production	33
4 Hearing, Auditory Models, and Speech Perception	43
5 Sound Propagation in the Human Vocal Tract	45
6 Time-Domain Methods for Speech Processing	65
7 Frequency-Domain Representations	83
8 The Cepstrum and Homomorphic Speech Processing	111
9 Linear Predictive Analysis of Speech Signals	121
10 Algorithms for Estimating Speech Parameters	141
11 Digital Coding of Speech Signals	145
12 Frequency-Domain Coding of Speech and Audio	167
13 Text-to-Speech Synthesis Methods	183
14 Automatic Speech Recognition and Natural Language Understanding	185

Chapter 2

Review of Fundamentals of Digital Signal Processing

- 2.1** (a) This system is not linear (the constant term makes it non linear) but is shift-invariant
(b) This system is linear but not shift-invariant (since the modulation term is not shift-invariant)
(c) This system is not linear (because of the cubic power) but is shift-invariant
(d) This system is linear and shift invariant (in fact the digital system is the convolution of the input with a rectangular window of length N samples.

- 2.2** (a) The system $y[n] = x[n] + 2x[n+1] + 3$ is not linear, as seen by the following counter example.
Consider inputs $x_1[n]$ and $x_2[n]$ with outputs:

$$\begin{aligned}y_1[n] &= T[x_1[n]] = x_1[n] + 2x_1[n+1] + 3 \\y_2[n] &= T[x_2[n]] = x_2[n] + 2x_2[n+1] + 3\end{aligned}$$

If we apply the system to the input $x_3[n] = ax_1[n] + bx_2[n]$ we get an output, $y_3[n]$ of the form:

$$y_3[n] = T[ax_1[n] + bx_2[n]] = [ax_1[n] + bx_2[n]] + [2ax_1[n+1] + 2bx_2[n+1]] + 3$$

which is not equal to the linear sum $ay_1[n] + by_2[n]$ thereby showing that the system is not linear.

- (b) The system $y[n] = x[n] + 2x[n+1] + 3$ is time-invariant (shift-invariant) as seen by considering the responses to $x[n]$ and to $x[n - n_0]$, i.e.,

$$\begin{aligned}y[n] &= T[x[n]] = x[n] + 2x[n+1] + 3 \\y[n - n_o] &= T(x[n - n_0]) = x[n - n_0] + 2x[n - n_0 + 1] + 3\end{aligned}$$

- (c) The system $y[n] = x[n] + 2x[n+1] + 3$ is not causal since the output at time n depends on the output at a future time $n + 1$.

- 2.3** (a) The input sequence a^n is an eigen-function of LTI systems. Therefore if this system is LTI, the output must be of the form $A \cdot$ input or $y[n] = Aa^n$ where A is a complex constant. Since $b^n \neq Aa^n$ for any complex constant A , the system cannot be LTI.

- (b) The system is not LTI.
- (c) In this case, the input excites the system at all frequencies (since it exists only for $n \geq 0$). Therefore the system transfer function describes how the system transforms all inputs. Thus this system could be LTI and there is only one LTI system with the given transfer function, i.e.,

$$H(z) = \frac{1 - az^{-1}}{1 - bz^{-1}}, |z| > b$$

- 2.4 (a)** $x[n]$ can be written in the form:

$$x[n] = a^n u[n - n_0]$$

where

$$u[n - n_0] = \begin{cases} 1 & n \geq n_0 \\ 0 & n < n_0. \end{cases}$$

We can now solve for $X(z)$ by the following steps:

$$\begin{aligned} x[n] &= a^{n-n_0+n_0} u[n - n_0] \\ &= a^{n_0} a^{n-n_0} u[n - n_0] \\ X(z) &= a^{n_0} \sum_{n=n_0}^{\infty} a^{(n-n_0)} z^{-n} \end{aligned}$$

We can now make a change of variables to the form: $n' = n - n_0$ giving:

$$\begin{aligned} X(z) &= a^{n_0} \sum_{n'=0}^{\infty} a^{n'} z^{-n'-n_0} \\ &= a^{n_0} z^{-n_0} \sum_{n'=0}^{\infty} a^{n'} z^{-n'} \\ &= \frac{a^{n_0} z^{-n_0}}{1 - az^{-1}}; \quad |az^{-1}| < 1 \end{aligned}$$

where we have used the relation:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1 - r}; \quad |r| < 1$$

- (b) $X(e^{j\omega}) = X(z)$ evaluated on the unit circle (i.e., for $z = e^{j\omega}$). Thus we get:

$$X(e^{j\omega}) = \frac{a^{n_0} e^{-j\omega n_0}}{1 - ae^{-j\omega}}; \quad |ae^{-j\omega}| < 1,$$

or equivalently, $|a| < 1$. The Fourier transform exists when the z -transform converges in a region including the unit circle; in this case when $|a| < 1$.

- 2.5 (a)** $x[n]$ can be written in the form:

$$x[n] = x_1[n] + x_2[n]$$

where

$$x_1[n] = u[n]$$

and

$$x_2[n] = -u[n - N] = -x_1[n - N]$$

Using this form for $x[n]$, we can solve for the convolution output as the output due to $u[n]$ for the region $0 \leq n \leq N - 1$, and as the output due to $u[n] - u[n - N]$ for the region $N \leq n$. We call the output of the convolution of $x_1[n]$ with $h[n]$ as $y_1[n]$ and we solve for its value in the region $0 \leq n \leq N - 1$ using the convolution formula:

$$\begin{aligned} y_1[n] &= u_1[n] * h[n] \\ &= \sum_{m=-\infty}^{\infty} x_1[m]h[n-m] \\ &= \sum_{m=0}^n a^{n-m}(1) \\ &= a^n \sum_{m=0}^n a^{-m} \\ &= a^n \frac{1 - a^{-n-1}}{1 - a^{-1}} \\ &= \frac{1 - a^{n+1}}{1 - a} \quad 0 \leq n \leq N - 1 \end{aligned}$$

We trivially solve for $y_2[n] = -y_1[n - N]$ as

$$y_2[n] = -\frac{(1 - a^{n+1})}{(1 - a)} \quad N \leq n$$

giving, for $y[n]$, the value (for the region $N \leq n$)

$$y[n] = y_1[n] + y_2[n] = \frac{1 - a^{n+1}}{1 - a} - \frac{1 - a^{n-N+1}}{1 - a}$$

or, equivalently,

$$y[n] = a^n \frac{(a^{-N+1} - a)}{(1 - a)} \quad N \leq n$$

(b) Using z -transforms we solve for $Y(z)$ again as a sum in the form:

$$\begin{aligned} Y(z) &= X(z) \cdot H(z) \\ &= X_1(z) \cdot H(z) + X_2(z) \cdot H(z) \end{aligned}$$

where $X_1(z)$ and $X_2(z)$ are the z -transforms, respectively, of $x_1[n]$ and $x_2[n]$ of part (a) of this problem. The resulting set of z -transforms is:

$$\begin{aligned} X_1(z) &= \frac{1}{1 - z^{-1}} \\ X_2(z) &= -\frac{z^{-N}}{1 - z^{-1}} \\ H(z) &= \frac{1}{1 - az^{-1}} \end{aligned}$$

We can now solve for $Y_1(z) = X_1(z) \cdot H(z)$ giving the form:

$$Y_1(z) = \frac{1}{(1 - z^{-1})(1 - az^{-1})}$$

Using the method of partial fraction expansion we factor $Y_1(z)$ into

$$Y_1(z) = \frac{A}{1 - z^{-1}} + \frac{B}{1 - az^{-1}}$$

We can now solve for A and B using the fact that the combined numerator is 1, giving:

$$A = \frac{1}{1 - a}, \quad B = \frac{a}{1 - a}$$

Now we can invert the partial fraction expansion, giving

$$y_1[n] = \frac{1}{1 - a} u[n] - \frac{a}{1 - a} a^n u[n] = \frac{1 - a^{n+1}}{1 - a} u[n]$$

which is valid for all n but applies to the region $0 \leq n \leq N - 1$. Similarly we can trivially solve for $y_2[n] = -y_1[n - N]u[n - N]$ again giving the same total result for the region $N \leq n$.

2.6 (1) The z -transform of the exponential window is of the form:

$$W_1(z) = \sum_{n=0}^{N-1} a^n z^{-n} = \frac{(1 - a^N z^{-N})}{(1 - az^{-1})}$$

Note that zeros occur at $z_k = ae^{j2\pi k/N}$, $k = 1, 2, \dots, N - 1$.

The Fourier transform for the exponential window is just:

$$W_1(e^{j\omega}) = \frac{(1 - a^N e^{-j\omega N})}{(1 - ae^{-j\omega})}$$

(2) The rectangular window is a special case of the exponential window with $a = 1$. The zeros are now all on the unit circle at $z_k = e^{j2\pi k/N}$, $k = 1, 2, \dots, N - 1$. The Fourier transform of the rectangular window is of the form:

$$W_2(e^{j\omega}) = \frac{(1 - e^{j\omega N})}{(1 - e^{j\omega})} = e^{-j\omega(N-1)/2} \frac{\sin(\omega N/2)}{\sin(\omega/2)}$$

(3) The Hamming window can be expressed in terms of shifted sums of rectangular windows, i.e.,

$$w_3[n] = 0.54w_2[n] - 0.23w_2[n]e^{j2\pi n/(N-1)} - 0.23w_2[n]e^{-j2\pi n/(N-1)}$$

The Fourier transform of the Hamming window is thus of the form:

$$\begin{aligned} W_3(e^{j\omega}) &= 0.54W_2(e^{j\omega}) \\ &\quad - 0.23W_2(e^{j(\omega-2\pi/(N-1))}) - 0.23W_2(e^{j(\omega+2\pi/(N-1))}) \end{aligned}$$

which can be put into the form

$$\begin{aligned} W_3(e^{j\omega}) &= e^{-j\omega(N-1)/2} \left[-0.54 \frac{\sin(\omega N/2)}{\sin(\omega/2)} \right. \\ &\quad \left. + 0.23 \frac{\sin[(\omega - 2\pi/(N-1))(N/2)]}{\sin[(\omega - 2\pi/(N-1))(1/2)]} \right. \\ &\quad \left. + 0.23 \frac{\sin[(\omega + 2\pi/(N-1))(N/2)]}{\sin[(\omega + 2\pi/(N-1))(1/2)]} \right] \end{aligned}$$

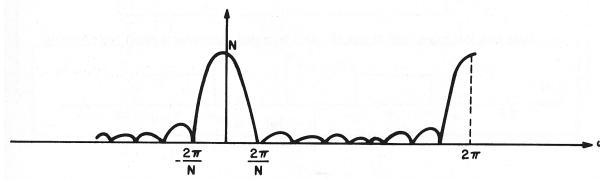


Figure P2.6.1: Magnitude response of rectangular window

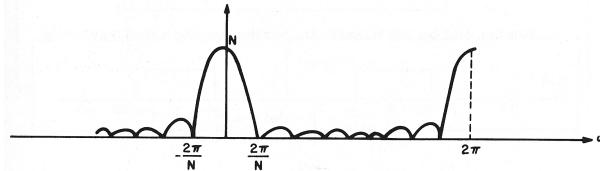


Figure P2.6.2: Magnitude response of exponential window

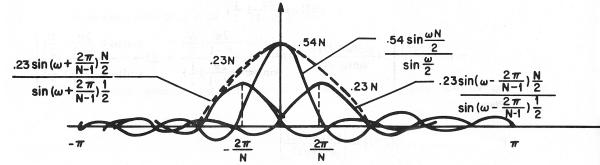


Figure P2.6.3: Magnitude response of Hamming window

A plot of $|W_2(e^{j\omega})|$ is shown in Figure P2.6.1. Notice that $W_1(z)$ has zeros on the unit circle; whereas $|W_1(z)$ has zeros on a circle of radius a as seen in Figure P2.6.2. Finally the Hamming window magnitude response is shown in Figure P2.6.3. We see how the side lobes tend to cancel and that the main lobe is about twice the width of the rectangular window response.

2.7 (a) A plot of an $N = 9$ -point triangular window is shown in Figure P2.7.1.

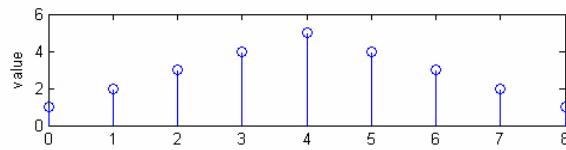


Figure P2.7.1: 9-point triangular window

(b) An N -point triangular window can be created by convolving an $(N+1)/2$ -point rectangular window with itself, i.e.,

$$w_T[n] = w_R[n] * w_R[n].$$

Since convolution in time is equivalent to multiplication in frequency, we have:

$$\begin{aligned} W_T(e^{j\omega}) &= [W_R(e^{j\omega})]^2 \\ W_R(e^{j\omega}) &= \frac{1 - e^{-j\omega(N+1)/2}}{1 - e^{-j\omega}} = e^{-j\omega(N-1)/4} \frac{\sin[\omega(N+1)/4]}{\sin(\omega/2)} \\ W_T(e^{j\omega}) &= e^{-j\omega(N-1)/2} \left[\frac{\sin[\omega(N+1)/4]}{\sin(\omega/2)} \right]^2 \end{aligned}$$

- (c) Plots of the time and frequency (log magnitude) responses of a 101-point triangular window are shown in Figure P2.7.2.

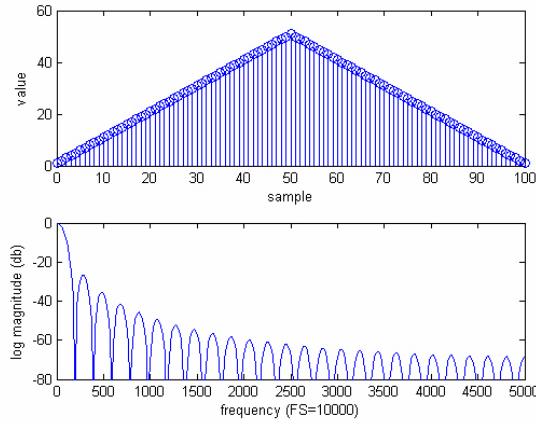


Figure P2.7.2: Time and frequency responses of 101-point triangular window

- (d) The rectangular, Hamming and triangular windows compare as follows:

1. rectangular window: cutoff frequency = $1/N$ in normalized frequency units and is F_s/N in analog frequency units, with sidelobe rejection ≥ 14 dB
2. Hamming window: cutoff frequency = $2/N$ in normalized frequency units and is $2F_s/N$ in analog frequency units, with sidelobe rejection ≥ 44 dB
3. triangular window: cutoff frequency = $2/N$ in normalized frequency units and is $2F_s/N$ in analog frequency units, with sidelobe rejection ≥ 28 dB

- 2.8 (a)** The impulse response of the ideal lowpass filter is obtained as:

$$\begin{aligned} h[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{j\omega n} d\omega \\ &= \frac{1}{2\pi} \left[\frac{e^{j\omega n}}{jn} \right]_{-\omega_c}^{\omega_c} = \frac{1}{2\pi jn} [e^{j\omega_c n} - e^{-j\omega_c n}] \\ &= \frac{\sin(\omega_c n)}{\pi n} \end{aligned}$$

- (b)** if $\omega_c = \pi/4$ then $h[n] = \frac{1}{4} \frac{\sin(\pi n/4)}{\pi n/4}$ and a plot of $h[n]$ is as shown in Figure P2.8.1.

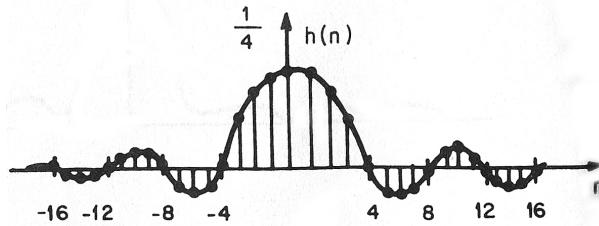


Figure P2.8.1: Impulse response of ideal lowpass filter.

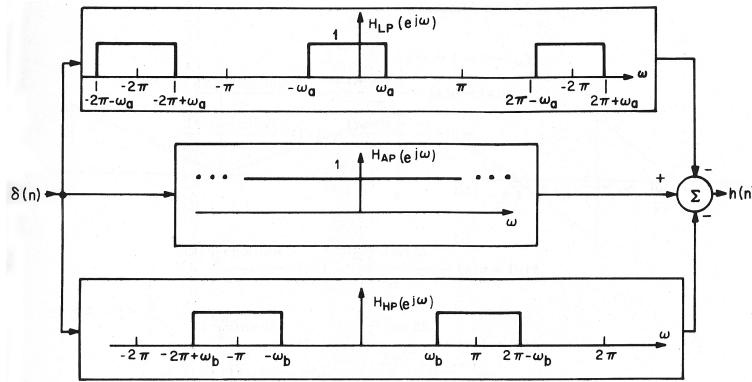


Figure P2.8.2: Parallel combination of ideal filters.

- (c) One approach to obtaining the desired impulse response is to view $H(e^{j\omega})$ as a parallel combination of lowpass, highpass, and zero phase filters, as shown in Figure P2.8.2.

We can express $H_{HP}(e^{j\omega})$ in terms of a lowpass filter with cutoff frequency $\pi - \omega_b$ with the passband shifted by π . From Part (a) we have:

$$H_{LP}(e^{j\omega}) \longleftrightarrow \frac{\sin(\omega_a n)}{\pi n}$$

where ω_a is the cutoff, giving:

$$H_{HP}(e^{j\omega}) = H_{LP}(e^{j(\omega-\pi)}) \longleftrightarrow e^{j\pi n} \frac{\sin[(\pi - \omega_b)n]}{\pi n}$$

where $\pi - \omega_b$ is the cutoff frequency.

The allpass has the property:

$$H_{AP}(e^{j\omega}) = 1 \longleftrightarrow \delta[n]$$

Putting it all together we get:

$$\begin{aligned} h[n] &= \delta[n] - \frac{\sin(\omega_a n)}{\pi n} - e^{j\pi n} \frac{\sin[(\pi - \omega_b)n]}{\pi n} \\ &= \delta[n] - \frac{\sin(\omega_a n)}{\pi n} - (-1)^n \frac{\sin[(\pi - \omega_b)n]}{\pi n} \end{aligned}$$

(d) When $\omega_a = \pi/4$ and $\omega_b = 3\pi/4$, we can express $h[n]$ for the bandpass filter as:

$$\begin{aligned} h[n] &= \delta[n] - \frac{0.25 \sin(\pi n/4)}{\pi n/4} - (-1)^n \frac{\sin[(\pi - 3\pi/4)n]}{\pi n} \\ &= \delta[n] - \frac{0.25 \sin(\pi n/4)}{\pi n/4} - (-1)^n \frac{0.25 \sin(\pi n/4)}{\pi n/4} \\ &= \delta[n] - \frac{0.25 \sin(\pi n/4)}{\pi n/4} [1 + (-1)^n] \end{aligned}$$

A plot of $h[n]$ for the bandpass filter is shown in Figure P2.8.3.

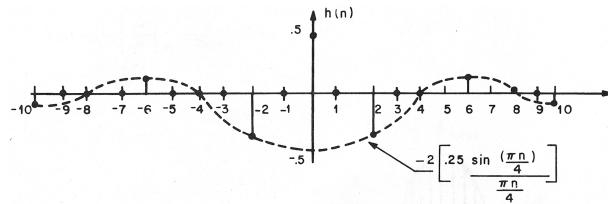


Figure P2.8.3: Impulse response of ideal bandpass filter.

2.9 (a) The magnitude response of an ideal differentiator is:

$$|H(e^{j\omega})| = |\omega|$$

and the phase response is:

$$\arg H(e^{j\omega}) = \begin{cases} -\omega\tau + \pi/2 & \omega > 0 \\ -\omega\tau - \pi/2 & \omega < 0 \end{cases}$$

A plot of the magnitude and phase is given in Figure P2.9.1.

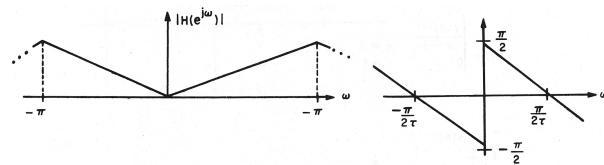


Figure P2.9.1: Magnitude and phase responses of ideal differentiator.

(b) The impulse response of the ideal differentiator is:

$$\begin{aligned}
 h[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega = \frac{j}{2\pi} \int_{-\pi}^{\pi} \omega e^{j\omega(n-\tau)} d\omega \\
 &= \frac{j}{2\pi} \left[\frac{e^{j\omega(n-\tau)}}{(j(n-\tau))^2} [j\omega(n-\tau) - 1] \right]_{-\pi}^{\pi} \\
 &= \frac{-j}{2\pi(n-\tau)^2} \left[e^{j\pi(n-\tau)}(j\pi(n-\tau) - 1) - e^{-j\pi(n-\tau)}(-j\pi(n-\tau) - 1) \right] \\
 &= \frac{-j}{2\pi(n-\tau)^2} \left[j\pi(n-\tau)[e^{j\pi(n-\tau)} + e^{-j\pi(n-\tau)}] + e^{-j\pi(n-\tau)} - e^{j\pi(n-\tau)} \right] \\
 &= \frac{\cos[\pi(n-\tau)]}{(n-\tau)} - \frac{\sin[(\pi)n-\tau]}{\pi(n-\tau)^2}
 \end{aligned}$$

(c) Using $\tau = (N-1)/2$ with N odd, we get:

$$h[n] = \frac{\cos[(\pi/2)(2n-N+1)]}{n-(N-1)/2} - \frac{\sin[(\pi/2)(2n-N+1)]}{\pi(n-(N-1)/2)^2}$$

Note that $2n$ is always even, and for n odd, $-N+1$ is even, therefore:

$$\sin[(\pi/2)(2n-N+1)] = 0, \quad n \neq (N-1)/2$$

and

$$h[n] = \frac{\cos[(\pi/2)(2n-N+1)]}{n-(N-1)/2} = \frac{(-1)^{n+1}}{n-(N-1)/2} \quad n \neq (N-1)/2$$

We see that $h[n]$ tends to decrease as $1/n$. Thus for $N = 11$ and $\tau = 5$ we get:

$$h[n] = \begin{cases} \frac{(-1)^{n+1}}{n-5} & n \neq 5 \\ 0 & n = 5 \end{cases}$$

A plot of $h[n]$ for an ideal differentiator with $N = 11$ is shown in Figure P2.9.2.

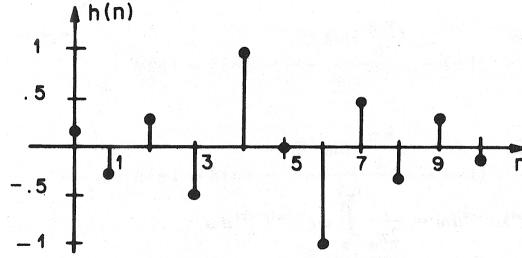


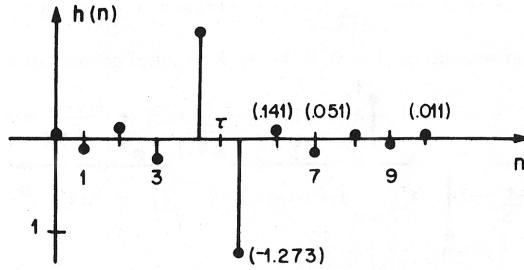
Figure P2.9.2: Impulse response of ideal $N = 11$ differentiator.

Note that the value of $h[n]$ for $n = 5$ is obtained from the equation:

$$h[n]|_{n=5} = \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} j\omega e^{j\omega(n-5)} d\omega \right]_{n=5} = \frac{j}{2\pi} \int_{-\pi}^{\pi} \omega d\omega = 0$$

(d) When $\tau = (N-1)/2$ and N is even, then $\cos[(\pi/2)(2n-N+1)] = 0$ and we get:

$$h[n] = \frac{\sin[(\pi/2)(2n-N+1)]}{\pi [n-(N-1)/2]^2} = \frac{(-1)^n}{\pi [n-(N-1)/2]^2}$$

Figure P2.9.3: Impulse response of ideal $N = 10$ differentiator.

For $N = 10$, $\tau = 9/2$ and $h[n] = \frac{(-1)^n}{\pi(n - 9/2)^2}$. A plot of $h[n]$ for an $N = 10$ point ideal differentiator is given in Figure P2.9.3.

2.10 The term $e^{-j\omega\tau}$ corresponds to a shift of τ samples in the impulse response. Therefore, start by defining the frequency response without delay as:

$$H'(e^{j\omega}) = \begin{cases} -j & 0 < \omega < \pi \\ j & -\pi < \omega < 0 \end{cases}$$

We solve for $h'[n]$ as:

$$\begin{aligned} h'[n] &= \frac{1}{2\pi} \left\{ \int_{-\pi}^0 j e^{j\omega n} d\omega - \int_0^\pi j e^{j\omega n} d\omega \right\} \\ &= \frac{1}{2\pi} \left\{ \frac{e^{j\omega n}}{n} \Big|_{-\pi}^0 - \frac{e^{j\omega n}}{n} \Big|_0^\pi \right\} \\ &= \frac{-1}{2\pi n} \{ e^{j\pi n} + e^{-j\pi n} - 2 \} \\ &= \frac{1}{\pi n} [1 - \cos(\pi n)], \quad n \neq 0 \end{aligned}$$

where we note that at $n = 0$ we get $h'[n] = 0$.

Inserting the appropriate shift of τ samples, we obtain:

$$h[n] = h'[n - \tau] = \frac{1}{\pi(n - \tau)} [1 - \cos[\pi(n - \tau)]], \quad n \neq 0$$

Using the trigonometric identity $(1/2) - (1/2)\cos(2\theta) = \sin^2(\theta)$ we can rewrite $h[n]$ as:

$$h[n] = \begin{cases} \frac{2 \sin^2(\frac{\pi(n - \tau)}{2})}{\pi(n - \tau)}, & n \neq \tau \\ 0 & n = \tau \end{cases}$$

which is plotted in Figure P2.10.1.

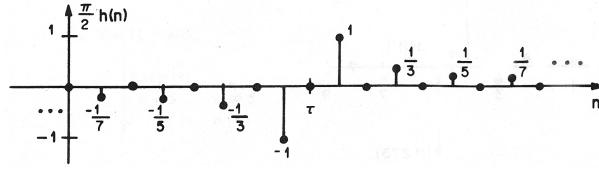


Figure P2.10.1: Impulse response of ideal hilbert transformer with delay τ .

2.11 (a) For the case where N is even we can write the convolution sum as:

$$\begin{aligned} y[n] &= \sum_{k=1}^{N-1} h[k]x[n-k] \quad \text{assuming } n \geq N-1 \\ &= \sum_{k=0}^{N/2-1} h[k]x[n-k] + \sum_{k=N/2}^{N-1} h[k]x[n-k] \end{aligned}$$

In the second term we replace k by $-k' + N - 1$ giving:

$$y[n] = \sum_{k=0}^{N/2-1} h[k]x[n-k] + \sum_{k'=-(N/2)-1}^0 h[N-1-k']x[n-N+1+k']$$

Observing that the ranges of summation for both terms are the same, we can replace the dummy variable k' by k giving:

$$y[n] = \sum_{k=0}^{N/2-1} (h[k]x[n-k] + h(N-1-k)x[n-N+1+k])$$

or since $h[k] = h[N-1-k]$ we have the result:

$$y[n] = \sum_{k=0}^{N/2-1} h[k](x[n-k] + x[n-N+1+k])$$

For the case where N is odd we can write the convolution sum (paying special attention to the term in $h[(N-1)/2]$ since it does not have a corresponding match) as:

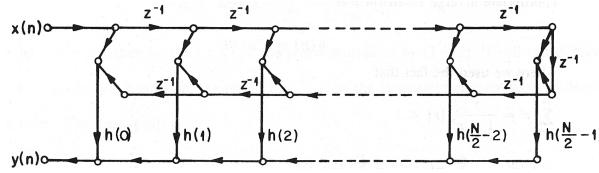
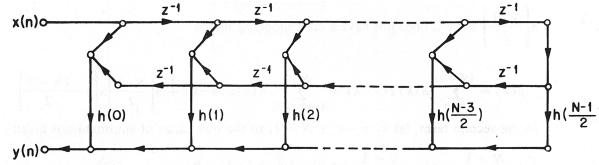
$$y[n] = \sum_{k=0}^{(N-3)/2} h[k]x[n-k] + \sum_{k=(N+1)/2}^{N-1} h[k]x[n-k] + h[(N-1)/2]x[n-(N-1)/2]$$

Again, in the second term, we let k be replaced by $-k'+N-1$ so the new range of summation is given by:

$$k = (N+1)/2 \rightarrow k' = (N-3)/2$$

and $k = N-1 \rightarrow k' = 0$, giving:

$$\begin{aligned} y[n] &= \sum_{k=0}^{(N-3)/2} h[k]x[n-k] + \sum_{k'=(N-3)/2}^{N-1} h[-k'+N-1]x[n-N+1+k'] \\ &\quad + h[(N-1)/2]x[n-(N-1)/2] \end{aligned}$$

Figure P2.11.1: Implementation of N even FIR linear phase filter.Figure P2.11.2: Implementation of N odd FIR linear phase filter.

or, equivalently:

$$y[n] = \sum_{k=0}^{(N-3)/2} h[k](x[n-k] + x[n-N+1+k]) + h[(N-1)/2]x[n-(N-1)/2]$$

- (b) For the case where N is even, we see that the first term in the expression for $y[n]$ corresponds to a “normal” FIR filter with filter length of $(N-2)/2$. The second term is implemented by delaying the samples by an amount of $N/2$ and then feeding these samples into the FIR filter from the opposite end. This is illustrated by the filter structure shown in Figure P2.11.1.

For the case where N is odd, a slight modification is required to account for the unmatched sample in the impulse response. The appropriate direct-form realization is given in Figure P2.11.2.

- 2.12 (a)** We can take the z -transform of the difference equation, giving:

$$\begin{aligned} Y(z) &= \frac{X(z) - \frac{1}{4}X(z)z^{-1}}{1 - \frac{1}{3}z^{-1}} \\ H(z) &= \frac{Y(z)}{X(z)} = \frac{1 - \frac{1}{4}z^{-1}}{1 - \frac{1}{3}z^{-1}}, \quad |z| > 1/3 \end{aligned}$$

- (b) The z -transform of $H(z)$ shows that there is a zero at $z = 1/4$ and a pole at $z = 1/3$
(c) With input $x[n] = u[n]$, we solve for $X(z)$ and then use a partial fraction expansion to solve for $Y(z)$ and then $y[n]$, as:

$$\begin{aligned}
 X(z) &= \frac{1}{1-z^{-1}} \\
 Y(z) &= \frac{1-\frac{1}{4}z^{-1}}{(1-z^{-1})(1-\frac{1}{3}z^{-1})} \\
 &= \frac{A}{1-z^{-1}} + \frac{B}{1-\frac{1}{3}z^{-1}}
 \end{aligned}$$

We can now solve for A, B by matching terms, giving:

$$\begin{aligned}
 A + B &= 1, \Rightarrow B = 1 - A \\
 -\frac{1}{3}A - B &= -\frac{1}{4} \\
 -\frac{1}{3}A - 1 + A &= -\frac{1}{4} \\
 A &= 9/8, \quad B = -1/8
 \end{aligned}$$

We now solve for $Y(z)$ and $y[n]$ as:

$$\begin{aligned}
 Y(z) &= \frac{9/8}{1-z^{-1}} - \frac{1/8}{1-\frac{1}{3}z^{-1}} \\
 y[n] &= (9/8)u[n] - (1/8)\left(\frac{1}{3}\right)^n u[n]
 \end{aligned}$$

(d) We solve for $H_i(z)$ as the inverse of $H(z)$ giving:

$$\begin{aligned}
 H_i(z) &= \frac{1}{H(z)} = \frac{1-\frac{1}{3}z^{-1}}{1-\frac{1}{4}z^{-1}} \\
 &= \frac{1}{1-\frac{1}{4}z^{-1}} - \frac{\frac{1}{3}z^{-1}}{1-\frac{1}{4}z^{-1}} \quad |z| > 1/4 \text{ stable inverse filter} \\
 h_i[n] &= \left(\frac{1}{4}\right)^n u[n] - \left(\frac{1}{3}\right) \left(\frac{1}{4}\right)^{n-1} u[n-1]
 \end{aligned}$$

2.13 (a) We can solve for $H(z)$ as:

$$\begin{aligned}
 Y(z) &= \alpha z^{-1}Y(z) + X(z) \\
 Y(z)(1-\alpha z^{-1}) &= X(z) \\
 H(z) &= \frac{Y(z)}{X(z)} = \frac{1}{(1-\alpha z^{-1})}
 \end{aligned}$$

- (b) Since the difference equation indicates that $h[n]$ is a causal (right-sided) sequence, the appropriate inverse z -transform is:

$$h[n] = \alpha^n u[n]$$

- (c) BIBO stability requires $|\alpha| < 1$.

(d) For the condition that $h[n] = \alpha^n u[n] < e^{-1}$ for $nT < 2$ msec we first solve for n in samples as:

$$n = \frac{2 \times 10^{-1}}{T} = \frac{0.002}{T}$$

We can now solve for α as:

$$\begin{aligned} (\alpha)^{(0.002/T)} &= e^{-1} \\ \ln \alpha &= \frac{-T}{2 \times 10^{-3}} \\ \alpha &= \exp[-500T] \end{aligned}$$

- 2.14 (a)** A complex zero occurs at $z = e^{\pm j\theta}$ and a complex pole occurs at $z = re^{\pm j\theta}$. The plot of the complex pole-zero locations for $r = 0.95$ and $\theta = 60^\circ$ ($\pi/3$) radians is shown in Figure P2.14.1.

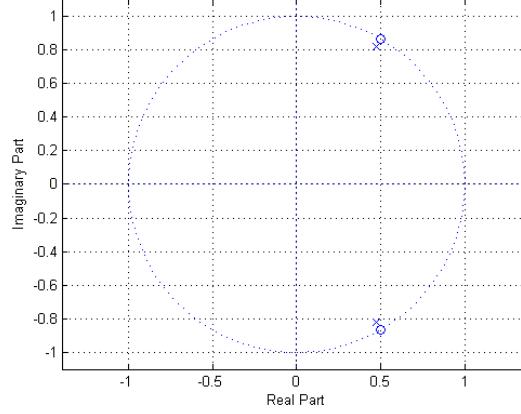


Figure P2.14.1: Pole-Zero Plot for Notch Filter.

- (b) The log magnitude plot is shown in Figure P2.14.2.
- (c) The maximum value of $|H(e^{j\omega})|$ occurs at either $\omega = 0$ or $\omega = \pi$, depending on the value of θ . The maximum value ≈ 1 (i.e., 0 dB) if r is close to 1.0.
- (d) For a notch to occur at 60 Hz, for a sampling rate of $f_S = 8000$ Hz, we need a value of $\theta = 60 * 2\pi / 8000 = 3\pi / 200$ radians.

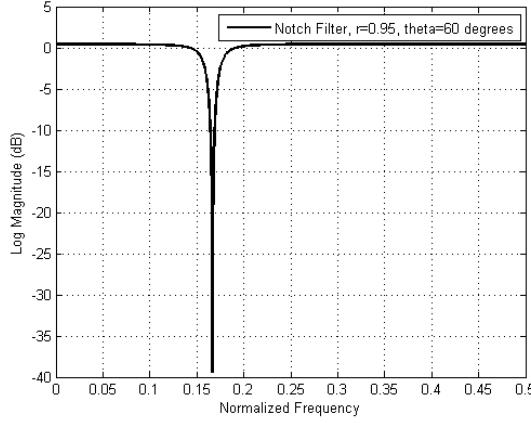


Figure P2.14.2: Log Magnitude Response for Notch Filter.

2.15 (a) We solve for $H(e^{j\omega})$ as:

$$\begin{aligned}
 H(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n} \\
 &= \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n \cos\left(\frac{\pi n}{2}\right) e^{-j\omega n} \\
 &= \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n \left[\frac{e^{j\pi n/2} + e^{-j\pi n/2}}{2}\right] e^{-j\omega n} \\
 &= \frac{1}{2} \sum_{n=0}^{\infty} \left[\frac{1}{2} e^{j(\pi/2-\omega)}\right]^n + \frac{1}{2} \sum_{n=0}^{\infty} \left[\frac{1}{2} e^{j(-\pi/2-\omega)}\right]^n \\
 &= \frac{(1/2)}{1 - (1/2)e^{j(\pi/2-\omega)}} + \frac{(1/2)}{1 - (1/2)e^{j(-\pi/2-\omega)}} \\
 &= \frac{(1/2)}{1 - (1/2)je^{-j\omega}} + \frac{(1/2)}{1 + (1/2)je^{-j\omega}} \\
 &= \frac{1}{1 + (1/4)e^{-2j\omega}}
 \end{aligned}$$

(b) The input can be written as a sum of eigenfunctions of the system, i.e.,

$$\cos\left(\frac{\pi n}{2}\right) = \frac{e^{j\pi n/2} + e^{-j\pi n/2}}{2}$$

The system response to an eigenfunction $e^{j\omega_1 n}$ is:

$$y[n] = H(e^{j\omega})|_{\omega=\omega_1} \cdot e^{j\omega_1 n}$$

Thus, from part (a), we get:

$$\begin{aligned}
 y[n] &= \left(\frac{1}{2}\right) H(e^{j\omega})|_{\omega=\pi/2} \cdot e^{j\pi n/2} + \left(\frac{1}{2}\right) H(e^{j\omega})|_{\omega=-\pi/2} \cdot e^{-j\pi n/2} \\
 &= \left(\frac{1}{2}\right) \frac{1}{1 + (1/4)e^{-j\pi}} e^{j\pi n/2} + \left(\frac{1}{2}\right) \frac{1}{1 + (1/4)e^{j\pi}} e^{-j\pi n/2} \\
 &= \left(\frac{4}{3}\right) \cos(\pi n/2)
 \end{aligned}$$

2.16 (a) The system function is of the form:

$$H(z) = \frac{A \prod_{r=1}^M (1 - c_r z^{-1})}{\prod_{k=1}^N (1 - d_k z^{-1})}, \quad A = \text{gain constant}$$

If $M < N$, we can consider this product form for $H(z)$ to be the result of establishing a common denominator for $H(z)$ given by:

$$H(z) = \sum_{k=1}^N \frac{A_k}{(1 - d_k z^{-1})}$$

The A_k 's are termed residues and should not be confused with the gain constant, denoted in this problem by A . Since, in the product form, the numerator is a polynomial in z^{-1} , and since $M < N$, we are assured that the A_k 's are just complex constants.

In order to determine a particular A_m , we multiply both sides of the expression for $H(z)$ by the corresponding denominator term so that

$$(1 - d_m z^{-1}) H(z) = \sum_{\substack{k=1 \\ k \neq m}}^N \frac{(1 - d_m z^{-1}) A_k}{(1 - d_k z^{-1})} = A_m$$

If we let $z^{-1} \rightarrow d_m$, then on the right hand side all the terms multiplied by $(1 - d_m z^{-1})$ will vanish leaving only A_m . On the left-hand side, the term $(1 - d_m z^{-1})$ in the numerator cancels with the denominator term. The final result is:

$$\frac{A \prod_{r=1}^M (1 - c_r d_m^{-1})}{(1 - d_k d_m^{-1})} = A_m, \quad m = 1, 2, \dots, N$$

In this manner, all the A_k 's are determined.

(b) Given $h_k[n] = A_k(d_k)^n u[n]$ we can compute the z -transform as:

$$\begin{aligned} H(z) &= \sum_{n=-\infty}^{\infty} h_k[n] z^{-n} = \sum_{n=0}^{\infty} A_k (d_k)^n z^{-n} \\ &= A_k \sum_{n=0}^{\infty} (d_k z^{-1})^n \\ &= \frac{A_k}{1 - d_k z^{-1}}, \quad |d_k z^{-1}| < 1 \text{ or } |z| > d_k \end{aligned}$$

2.17 (a) For the first system we have:

$$v[n] = x[n] * h_1[n] = \sum_{l=-\infty}^{\infty} x[l] h_1[n-l]$$

For the second system we have:

$$y[n] = v[n] * h_2[n] = \sum_{k=-\infty}^{\infty} v[k]h_2[n-k]$$

If we substitute for $v[k]$ we get:

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} x[l]h_1[k-l]h_2[n-k] \\ &= \sum_{l=-\infty}^{\infty} x[l] \sum_{k=-\infty}^{\infty} h_1[k-l]h_2[n-k] \end{aligned}$$

If we make the substitution of variables, $k \rightarrow k' + l$ we get:

$$\begin{aligned} y[n] &= \sum_{l=-\infty}^{\infty} x[l] \sum_{k'=-\infty}^{\infty} h_1[k']h_2[(n-l)-k'] \\ &= \sum_{l=-\infty}^{\infty} x[l](h_1[n-l] * h_2[n-l]) \\ h[n] &= h_1[n] * h_2[n] \end{aligned}$$

(b) We can express the convolution of h_1 and h_2 as:

$$h_1[n] * h_2[n] = \sum_{k=-\infty}^{\infty} h_1[k]h_2[n-k]$$

If we make the substitution of variables $k \rightarrow -k' + n$ then $k = \infty \rightarrow k' = -\infty$ and $k = -\infty \rightarrow k' = \infty$ so we can express the convolution as

$$h_1[n] * h_2[n] = \sum_{k'=\infty}^{-\infty} h_1[-k'+n]h_2[k']$$

Since k' is just a dummy variable we can write the above equation as:

$$\sum_{k=-\infty}^{\infty} h_1[n-k]h_2[k] = h_2[n] * h_1[n]$$

Therefore we get:

$$h_1[n] * h_2[n] = h_2[n] * h_1[n]$$

(c)

$$H(z) = \left[\sum_{r=0}^M b_r z^{-r} \right] \left[\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right] = H_1(z) \cdot H_2(z)$$

where $H_1(z)$ represents an all-zero (FIR) filter and $H_2(z)$ represents an all-pole IIR filter. Using the previously defined input-output notation we get:

$$H_1(z) = \frac{V(z)}{X(z)} \rightarrow V(z) = \sum_{r=0}^M b_r z^{-r} X(z)$$

$$H_2(z) = \frac{Y(z)}{V(z)} \rightarrow Y(z) = V(z) + \sum_{k=1}^N a_k z^{-k} Y(z)$$

and we obtain the difference equations by inverse transforming these relations, giving:

$$\begin{aligned} v[n] &= \sum_{r=0}^M b_r x[n-r] \\ y[n] &= v[n] - \sum_{k=1}^N a_k y[n-k] \\ \text{x}(n) \rightarrow & \boxed{h_2(n)} \xrightarrow{w(n)} \boxed{h_1(n)} \rightarrow \text{y}(n) \end{aligned}$$

Figure P2.17.1: Cascade of $h_2[n]$ with $h_1[n]$.

- (d) Now consider the two systems in the opposite order, i.e., $h_2[n]$ preceding $h_1[n]$ as shown in Figure P2.17.1. We then have:

$$\begin{aligned} W(z) &= X(z) + \sum_{k=1}^N a_k z^{-k} W(z) \\ Y(z) &= \sum_{r=0}^M b_r z^{-r} W(z) \end{aligned}$$

Inverse transformation yields:

$$\begin{aligned} w[n] &= x[n] + \sum_{k=1}^N a_k w[n-k] \\ y[n] &= \sum_{r=0}^M b_r w[n-r] \end{aligned}$$

- 2.18** The difference equation has a solution for $y[n]$ that is composed of a homogeneous and a particular solution. Since the input is zero, the total solution for this example is equal to the homogeneous solution, which is of the form:

$$y[n] = A\alpha_1^n + B\alpha_2^n$$

Substituting $y[n] = A\alpha^n$ into the difference equation, we obtain:

$$\begin{aligned} A\alpha^n &= 2\cos(bT)A\alpha^{n-1} - A\alpha^{n-2} \\ 1 &= 2\cos(bT)\alpha^{-1} - \alpha^{-2} \\ \alpha^2 - 2\cos(bT)\alpha + 1 &= 0 \\ \alpha &= \cos(bT) \pm \frac{\sqrt{4\cos^2(bT) - 4}}{2} \\ \alpha &= \cos(bT) \pm \sqrt{-\sin^2(bT)} = \cos(bT) \pm j\sin(bT) \\ y[n] &= A[\cos(bT) + j\sin(bT)]^n + B[\cos(bT) - j\sin(bT)]^n \end{aligned}$$

The initial conditions will determine the appropriate values for A and B . Alternately, we can choose A and B and then determine the corresponding initial conditions. First we rewrite $y[n]$ using Euler's identity as:

$$y[n] = Ae^{jbTn} + Be^{-jbTn}$$

(a) $y[n] = \cos(bTn)u[n] \Rightarrow A = B = 1/2$ with initial conditions:

$$\begin{aligned} y[-1] &= Ae^{-jbT} + Be^{jbT} = \frac{1}{2}[e^{-jbT} + e^{jbT}] \\ y[-1] &= \cos(bT) \\ y[-2] &= \frac{1}{2}[e^{-j2bT} + e^{j2bT}] = \cos(2bT) \\ y[n] &= \cos(bTn)u[n] \end{aligned}$$

(b) Similarly $y[n] = \sin(bTn)u[n] \Rightarrow A = -B = \frac{1}{2j}$ so we require that $y[-1] = \frac{1}{2j}[e^{-jbT} - e^{jbT}] = -\sin(bT)$ and similarly $y[-2] = -\sin(2bT)$.

2.19 (a) The network diagram for this system is shown in Figure P2.19.1.

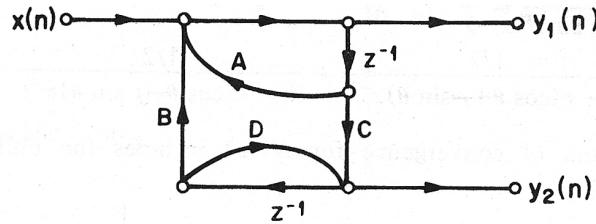


Figure P2.19.1: Network diagram of set of difference equations.

(b) Using transforms we get:

$$Y_1(z) = Az^{-1}Y_1(z) + Bz^{-1}Y_2(z) + X(z)$$

$$Y_2(z) = Cz^{-1}Y_1(z) + Dz^{-1}Y_2(z)$$

Solving the second equation for $Y_2(z)$ gives:

$$Y_2(z) = \frac{Cz^{-1}Y_1(z)}{1 - Dz^{-1}}$$

We now substitute in the first equation giving:

$$\begin{aligned} Y_1(z) &= Az^{-1}Y_1(z) + Bz^{-1} \left[\frac{Cz^{-1}Y_1(z)}{1 - Dz^{-1}} \right] + X(z) \\ &= \frac{X(z)}{1 - Az^{-1} - \frac{BCz^{-2}}{1 - Dz^{-1}}} \\ \frac{Y_1(z)}{X(z)} &= H_1(z) = \frac{1 - Dz^{-1}}{1 - (A + D)z^{-1} + (AD - BC)z^{-2}} \end{aligned}$$

From the second equation we get:

$$H_2(z) = \frac{Y_2(z)}{X(z)} = \frac{Cz^{-1}}{(1 - Dz^{-1})} \frac{Y_1(z)}{X(z)} = \frac{Cz^{-1}}{1 - (A + D)z^{-1} + (AD - BC)z^{-2}}$$

(c) If $A = D = r \cos(\theta)$ and $C = -B = r \sin(\theta)$ we get:

$$\begin{aligned} H_1(z) &= \frac{1 - r \cos(\theta)z^{-1}}{1 - 2r \cos(\theta)z^{-1} + r^2 z^{-2}} \\ &= \frac{A_1}{1 - r(\cos(\theta) + j \sin(\theta))z^{-1}} + \frac{A_1^*}{1 - r(\cos(\theta) - j \sin(\theta))z^{-1}} \end{aligned}$$

where:

$$\begin{aligned} A_1 &= \lim_{z \rightarrow r(\cos(\theta) + j \sin(\theta))} \left[\frac{z - r \cos(\theta)}{z - r(\cos(\theta) - j \sin(\theta))} \right] \\ &= \frac{j r \sin(\theta)}{2 j r \sin(\theta)} = \frac{1}{2} \\ H_1(z) &= \frac{1/2}{1 - r(\cos(\theta) + j \sin(\theta))z^{-1}} + \frac{1/2}{1 - r(\cos(\theta) - j \sin(\theta))z^{-1}} \end{aligned}$$

We assume the region of convergence for $H_1(z)$ includes the unit circle. Inverse transforming gives:

$$\begin{aligned} h_1[n] &= \frac{1}{2} [r \cos(\theta) + j \sin(\theta)]^n u[n] + \frac{1}{2} [r \cos(\theta) - j \sin(\theta)]^n u[n] \quad \text{for } |r| < 1 \\ &= \frac{1}{2} [re^{j\theta}]^n u[n] - \frac{1}{2} [re^{-j\theta}]^n u[n] \\ &= \frac{1}{2} r^n [e^{j\theta n} + e^{-j\theta n}] u[n] = r^n \cos(\theta n) u[n] \\ &= r^n \cos(\theta n) u[n] \\ H_2(z) &= \frac{A_1}{1 - r \cos(\theta) + j \sin(\theta)z^{-1}} + \frac{A_1^*}{1 - r \cos(\theta) - j \sin(\theta)z^{-1}} \end{aligned}$$

where:

$$\begin{aligned} A_1 &= \lim_{z \rightarrow r(\cos(\theta) + j \sin(\theta))} \left[\frac{r \sin(\theta)}{z - r(\cos(\theta) - j \sin(\theta))} \right] \\ &= \frac{r \sin(\theta)}{2 j r \sin(\theta)} = \frac{1}{2j} \\ H_2(z) &= \frac{1/(2j)}{1 - r(\cos(\theta) + j \sin(\theta))z^{-1}} - \frac{1/(2j)}{1 - r(\cos(\theta) - j \sin(\theta))z^{-1}} \\ h_2[n] &= r^n \sin(\theta n) u[n] \end{aligned}$$

2.20 (a) The cascade implementation is shown in Figure P2.20.1 and the direct form implementation is shown in Figure P2.20.2. The direct form system function is of the form:

$$H(z) = \frac{1 + 4z^{-1} + 6z^{-2} + 4z^{-3} + z^{-4}}{1 + (13/8)z^{-1} + (59/32)z^{-2} + z^{-3} + (35/128)z^{-4}}$$

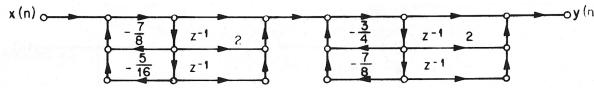


Figure P2.20.1: Cascade form implementation.

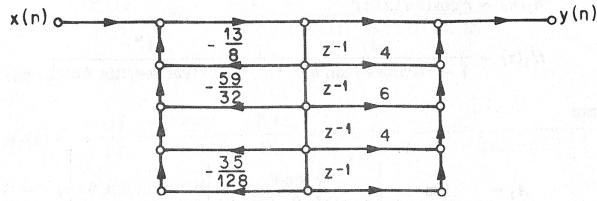


Figure P2.20.2: Direct form implementation.

- (b) The polynomial $1 + (7/8)z^{-1} + (5/16)z^{-2}$ can be written as $(1 + az^{-1})(1 + bz^{-1})$ where $-a$ and $-b$ are the roots. Since $(1 + az^{-1})(1 + bz^{-1}) = 1 + (a + b)z^{-1} + abz^{-2}$ we see that the two roots are inside the unit circle since ab and $(a + b)$ have the same sign, and $(a + b) < 1$. By the same observation, the roots of $(1 + (3/4)z^{-1} + (7/8)z^{-2})$ are also inside the unit circle and the system is therefore stable.

- 2.21 (a)** The difference equations for the network of Figure P2.21 are as follows:

$$y[n] = x[n] - v[n - 1] + a_1 y[n - 1]$$

$$v[n] = y[n] + a_2 v[n - 1]$$

- (b) The system function is derived as follows:

$$H(z) = \frac{Y(z)}{X(z)}$$

$$Y(z) = X(z) - z^{-1}V(z) + a_1 z^{-1}Y(z)$$

$$Y(z)[1 - a_1 z^{-1}] = X(z) - z^{-1}V(z)$$

$$V(z) = Y(z) + a_2 z^{-1}V(z)$$

$$V(z) = \frac{Y(z)}{1 - a_2 z^{-1}}$$

Substituting into the expression for $Y(z)$ gives:

$$Y(z) \left[1 - a_1 z^{-1} + \frac{z^{-1}}{1 - a_2 z^{-1}} \right] = X(z)$$

$$\frac{Y(z)}{X(z)} = \frac{1 - a_2 z^{-1}}{1 + (1 - a_1 - a_2)z^{-1} + a_1 a_2 z^{-2}} = H(z)$$

The system function is obtained by setting $z = e^{j\omega}$ giving:

$$H(e^{j\omega}) = \frac{1 - a_2 e^{-j\omega}}{1 + (1 - a_1 - a_2)e^{-j\omega} + a_1 a_2 e^{-j2\omega}}$$

2.22 For Network #1 we have:

$$H_1(z) = \frac{1}{1 - b_1 z^{-1}} - \frac{1}{1 - b_2 z^{-1}} = \frac{(b_1 - b_2)z^{-1}}{1 - (b_1 + b_2)z^{-1} + b_1 b_2 z^{-2}}$$

For Network #2 we have:

$$H_2(z) = \left[\frac{1}{1 - a_1 z^{-1}} \right] \left[\frac{z^{-1}}{1 - a_2 z^{-1}} \right] a_3$$

$$H_2(z) = \frac{a_3 z^{-1}}{1 - (a_1 + a_2)z^{-1} + a_1 a_2 z^{-2}}$$

We require that $a_3 = (b_1 - b_2)$, $(a_1 + a_2) = (b_1 + b_2)$ and $a_1 a_2 = b_1 b_2$ which implies that either $a_1 = b_1$ and $a_2 = b_2$ or else $a_1 = b_2$ and $a_2 = b_1$.

2.23 (a) We can factor the $H(z)$ polynomial and find the root locations from the factored form. The result of factoring is:

$$H(z) = \frac{1 - 2e^{-aT} \cos(bT) + e^{-2aT}}{1 - 2e^{-aT} \cos(bT)z^{-1} + e^{-2aT}z^{-2}}$$

$$H(z) = \frac{z^2(1 - 2e^{-aT} \cos(bT) + e^{-2aT})}{(z - e^{-(a+jb)T})(z - e^{-(a-jb)T})}$$

The roots of $H(z)$ are plotted in Figure P2.23.1.

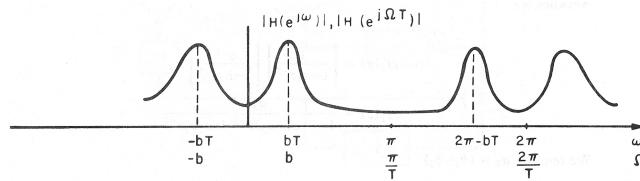


Figure P2.23.1: Location of poles and zeros of $H(z)$ in the Z -plane.

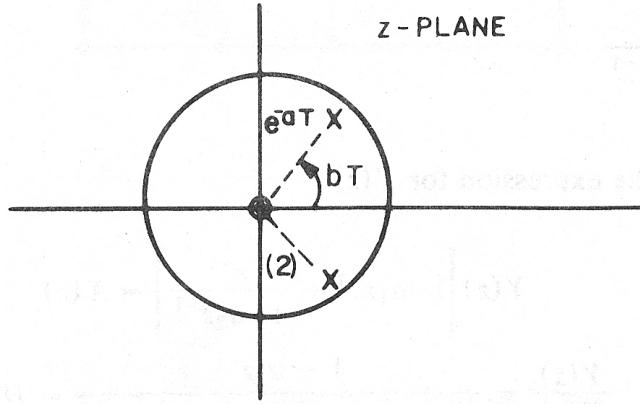


Figure P2.23.2: Frequency response of the resonator.

- (b) We can invert $H(z)$ using a partial fraction expansion giving the impulse response:

$$h[n] = \begin{cases} \frac{(1 - e^{-aT} \cos(bT) + e^{-2aT})}{\sin(bT)} e^{-aTn} \sin[(n+1)bT] & n \geq 0 \\ 0 & n < 0 \end{cases}$$

- (c) A plot of the frequency response of the resonator is given in Figure P2.23.2.

- 2.24** (a) The z -transform and Fourier transform are:

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} = 1 + 0.5z^{-5}$$

$$X(e^{j\omega}) = 1 + 0.5e^{-j5\omega}$$

- (b) The N -point DFT is obtained by evaluating $X(e^{j\omega})$ at N evenly spaced frequencies on the unit circle; i.e., at the set of values:

$$\omega = \frac{2\pi k}{N}, \quad k = 0, 1, \dots, N-1$$

Hence for the cases of $N = 50$ and $N = 10$ we get:

$$N = 50 \quad X(e^{j2\pi k/50}) = 1 + 0.5e^{-j\pi k/5}, \quad k = 0, 1, \dots, 49$$

$$N = 10 \quad X(e^{j2\pi k/10}) = 1 + 0.5e^{-j\pi k}, \quad k = 0, 1, \dots, 9$$

For $N = 5$ the DFT is not properly defined since $x[n]$ is a 6-point sequence. Thus we can either use the first five values of $x[n]$, giving:

$$X(e^{j2\pi k/5}) = 1, \quad k = 0, 1, \dots, 4$$

or we can wrap $x[n]$ around a cylinder, i.e., take the result modulo 4, giving:

$$X(e^{j2\pi k/5}) = 1.5, \quad k = 0, 1, \dots, 4$$

- (c) For $N = 5$ the DFT values can either be considered the same as the DFT for $N = 50$, evaluated every 10th point, or they can differ depending on the signal values used for the computation.
- (d) The DFT $X[k]$ is the Fourier transform evaluated at N equally spaced frequencies (points around the unit circle) if N is greater than or equal to the duration of the sequence. Otherwise $X[k]$ and $X(e^{j\omega})$ need not be directly related. For $N = 5$ the DFT values can either be considered the same as

- 2.25** (a) The time duration, T_D , of an $L = 1024$ sample sequence with a sampling rate of $F_s = 1/T = 20,000$ samples/second is:

$$T_D = \frac{L}{F_s} = \frac{1024}{20,000} = 5115 \times 10^{-4} \text{ seconds} = 51.15 \text{ msec}$$

- (b) The frequency resolution in radians, $\Delta\omega$, between the DFT values is:

$$\Delta\omega = \frac{2\pi}{NFFT} = \frac{2\pi}{1024}$$

where $NFFT$ is the size of the DFT. Using the relationship between analog and digital frequency, we can determine the analog frequency resolution as:

$$\begin{aligned}\omega &= \Omega T \\ \Delta\Omega &= \frac{\Delta\omega}{T} = \frac{2\pi}{(1024)(5 \times 10^{-5})} \\ \Delta f &= \frac{\Delta\Omega}{2\pi} = \frac{1}{(1024)(5 \times 10^{-5})} = 19.531 \text{ Hz}\end{aligned}$$

- (c) If we change the duration of the speech segment to 512 samples, the time duration would become half of the previous duration or 25.55 msec. Using a 1024-point DFT will produce the same frequency resolution of 19.531 Hz as previously since the frequency resolution depends only on the size of the FFT that is performed.

- 2.26** Assume $x_{\min}[n]$ is a minimum-phase signal with all of its poles and zeros inside the unit circle. Using z -transforms we can express $X_{\min}(z)$ as:

$$x_{\min}[n] \longleftrightarrow X_{\min}(z) = \sum_{n=-\infty}^{\infty} x_{\min}[n] z^{-n} = \frac{\prod_{i=1}^{N_z} (1 - a_i z^{-i})}{\prod_{i=1}^{N_p} (1 - b_i z^{-i})}$$

Since $x_{\min}[n]$ is a minimum-phase signal, then all the poles ($z = b_i$) and zeros ($z = a_i$) are inside the unit circle. Therefore we have the constraint $|a_i| < 1$ for all i and $|b_i| < 1$ for all i . The signal $x_{\max}[n] = x_{\min}[-n]$ has a z -transform of the form:

$$\begin{aligned}x_{\max}[n] \longleftrightarrow X_{\max}(z) &= \sum_{n=-\infty}^{\infty} x_{\max}[n] z^{-n} \\ &= \sum_{n=-\infty}^{\infty} x_{\min}[-n] z^{-n} = \sum_{n=-\infty}^{\infty} x_{\min}[n] (z^{-1})^{-n} = X_{\min}(z^{-1}) \\ &= \frac{\prod_{i=1}^{N_z} (1 - a_i z)}{\prod_{i=1}^{N_p} (1 - b_i z)}\end{aligned}$$

Thus the signal $x_{\min}[-n] = x_{\max}[n]$ has zeros at $z = 1/a_i > 1$ and poles at $z = 1/b_i > 1$. Thus all zeros and poles of $x_{\min}[-n] = x_{\max}[n]$ are outside the unit circle, so $x_{\min}[-n] = x_{\max}[n]$ is a maximum phase signal whenever $x_{\min}[n]$ is a minimum phase signal.

2.27 We need to utilize the series expansion:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots = \sum_{n=0}^{\infty} x^n$$

or, equivalently,

$$1 - x = \frac{1}{1 + x + x^2 + x^3 + \dots} = \frac{1}{\sum_{n=0}^{\infty} x^n}$$

(a) Thus we can utilize the above equation to write the expression for a single zero as:

$$\begin{aligned} z - a &= (1 - az^{-1}) z = \frac{z}{1 + az^{-1} + a^2 z^{-2} + a^3 z^{-3} + \dots} \\ &= \frac{1}{z^{-1}(1 + az^{-1} + a^2 z^{-2} + a^3 z^{-3} + \dots)} \end{aligned}$$

(b) When only a finite number of terms (poles) are used, the approximation to the real zero depends on the location of the real zero. Thus, as $|a| \rightarrow 0$, the approximation converges more rapidly.

(c) The first order approximation is:

$$z - a \approx \frac{z}{1 + az^{-1}} = \frac{z^2}{z + a}$$

i.e., a system with a real pole at $z = -a$. The second order approximation is:

$$z - a \approx \frac{z}{1 + az^{-1} + a^2 z^{-2}} = \frac{z^3}{z^2 + az + a^2}$$

i.e., a second order system with a complex pole at $z = -a/2 \pm j\sqrt{3}a/2$

(d) Similarly, we can represent a single pole as an infinite number of zeros using the expression:

$$\frac{1}{z - b} = z^{-1} + bz^{-2} + b^2 z^{-3} + \dots$$

2.28 (a) The ideal solution to fractional delays is a multirate system. Thus for a delay of $D = 1/2$ sample, the system would consist of an interpolator (by a factor of 2), a lowpass filter (with cutoff of half the original sampling frequency, a unit sample delay (at the interpolated sampling rate), and a decimator (by a factor of 2) to restore the original sampling rate of the system. These operations are illustrated in Figure P2.28.1.

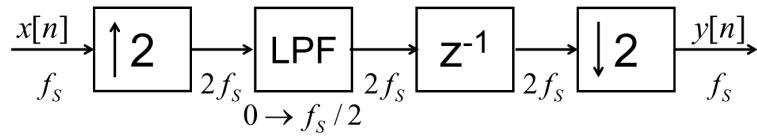


Figure P2.28.1: Ideal implementation of a delay of $D = 1/2$ sample.

- (b) The simplest approximation to a $D = 0.5$ sample delay (without interpolation and decimation) is simple linear interpolation (i.e., average the neighboring samples with equal weight), giving values of $\alpha = 0.5$ and $\beta = 0.5$.

In this case we get the resulting approximation to the ideal interpolator as:

$$\begin{aligned} y[n] &= 0.5x[n] + 0.5x[n-1] \\ H(z) &= 0.5(1+z^{-1}) \\ H(e^{j\omega}) &= 0.5(1+e^{-j\omega}) = 0.5(1+\cos(\omega)-j\sin(\omega)) \\ |H(e^{j\omega})| &= 0.5[(1+\cos(\omega))^2+\sin^2(\omega)]^{1/2} \\ &= [2+2\cos(\omega)]^{1/2} \\ &= \frac{\sqrt{2}}{2}(1+\cos(\omega))^{1/2} \end{aligned}$$

By plotting $|H(e^{j\omega})|$ versus ω we can see that the linear interpolator is basically a first order approximation to the ideal lowpass filter of the ideal interpolator which is flat in frequency until the point $\omega = \pi/2$ and then is zero from $\omega = \pi/2$ to $\omega = \pi$.

- (c) If the delay is changed to $D = 1/3$ sample, the interpolator becomes a 3-to-1 interpolator and the decimator becomes a 3-to-1 decimator, as shown in Figure P2.28.2.

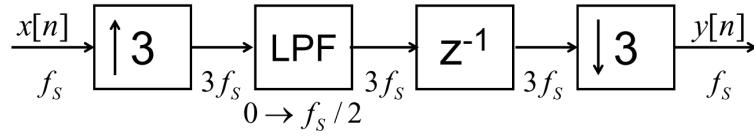


Figure P2.28.2: Ideal implementation of a delay of $D = 1/3$ sample.

- (d) The linear interpolation solution would just change the weights to reflect the desired delay of $D = 1/3$ sample, giving weights of $\alpha = 2/3$ and $\beta = 1/3$.
- (e) Since the desired delay is a rational fraction, the ideal solution becomes a multirate interpolator and a multirate decimator with a unit sample delay inserted in the middle. The signal processing operations (and the resulting sampling rates) are shown in Figure P2.28.3.

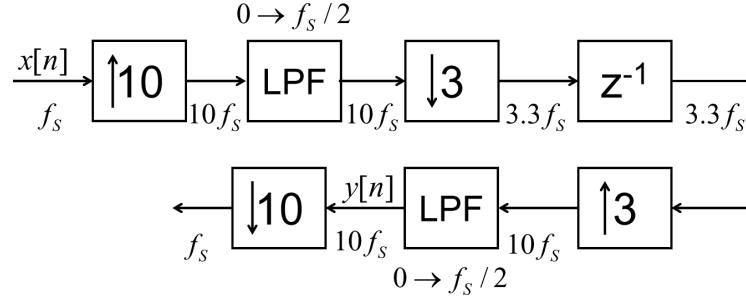


Figure P2.28.3: Ideal implementation of a delay of $D = 3/10$ sample using a single unit delay.

It is easily seen that the middle three blocks (the downsampling by a factor of 3-to-1 followed by unit sample delay followed by upsampling by a factor of 3-to-1) can be coalesced into a single block with a delay of 3 samples at the higher sampling rate. Further the serial combination of the remaining lowpass filters can be realized as a single lowpass filter, leading to the ideal implementation shown in Figure P2.28.4.

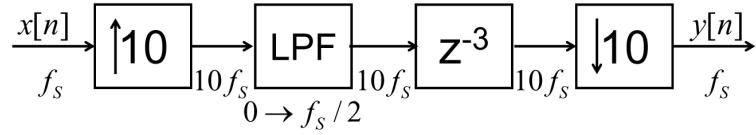


Figure P2.28.4: Ideal implementation of a delay of $D = 1/3$ sample using a three sample delay.

- 2.29 (a)** Using the trigonometric relationships we get:

$$\begin{aligned} y_1[n] &= x[n]^3 = \cos^3(\omega_0 n) \\ &= \frac{1}{4} \cos(3\omega_0 n) + \frac{3}{4} \cos(\omega_0 n) \\ &= \frac{e^{j3\omega_0 n}}{8} + \frac{e^{-j3\omega_0 n}}{8} + \frac{3e^{j\omega_0 n}}{8} + \frac{3e^{-j\omega_0 n}}{8} \\ Y_1(e^{j\omega}) &= \frac{1}{8} [\delta(\omega - 3\omega_0) + \delta(\omega + 3\omega_0)] + \frac{3}{8} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \end{aligned}$$

A plot of the magnitude response of $Y_1(e^{j\omega})$ is shown in Figure P2.29.1.

- (b)** For the input $x_2[n]$ we get an output of the form:

$$\begin{aligned} y_2[n] &= (r^n \cos(\omega_0 n) u[n])^3 = (r^{3n} u[n]) [\cos^3(\omega_0 n)] \\ &= (r^{3n} u[n]) [\frac{1}{4} [\cos(3\omega_0 n) + 3 \cos(\omega_0 n)]] \\ Y_2(e^{j\omega}) &= \frac{1}{1 - r^3 e^{-j\omega}} * Y_1(e^{j\omega}) \end{aligned}$$

The resulting log magnitude plots are shown in Figure P2.29.2, where we show log magnitude spectra for the original signal (top panel), the modulation signal, $\frac{1}{1 - r^3 e^{-j\omega}}$ (middle panel), and the cubed signal, $Y_2(e^{j\omega})$ (bottom panel).

- 2.30** When an analog signal with Nyquist frequency, F_N , of 3000 Hz (6000π radians) is sampled at a rate F_s samples per second, aliasing occurs when $F_s < 2F_N$. Hence for a sampling rate of $F_s = 10000$ there is no aliasing and the digital spectrum (on an analog frequency scale) is as shown in Figure P2.30.1(a). When the sampling rate is $F_s = 5000$ Hz, we violate the Nyquist condition and there is aliasing from $F = 2000$ Hz to $F = 3000$ Hz, as shown in Figure P2.30.1(b). Finally when the sampling rate is $F_s = 2000$ Hz, we severely violate the Nyquist condition and the entire frequency band is aliased as shown in Figure P2.30.1(c).

- 2.31 (a)** The digital signal, $x_1[n]$, resulting from sampling $x_a(t)$, with sampling rate $F_s = 10000$ Hz, can be written in the form:

$$x_1[n] = A \cdot \cos(2\pi 200n/F_s) \quad (2.1)$$

$$= \frac{A}{2} \cdot [e^{j2\pi 200n/10000} + e^{-j2\pi 200n/10000}] \quad (2.2)$$

$$= \frac{A}{2} \cdot [e^{j2\pi n/50} + e^{-j2\pi n/50}] \quad (2.3)$$

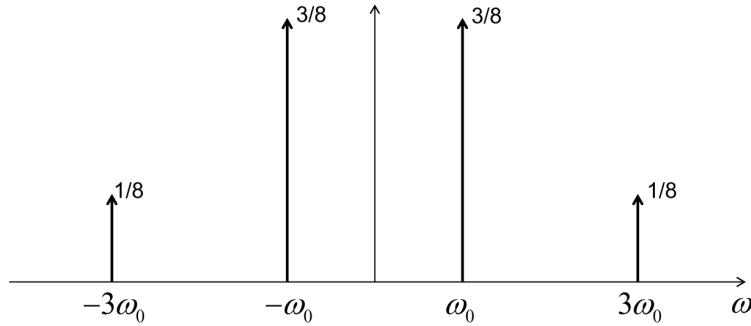


Figure P2.29.1: Magnitude spectrum of cubed cosine wave.

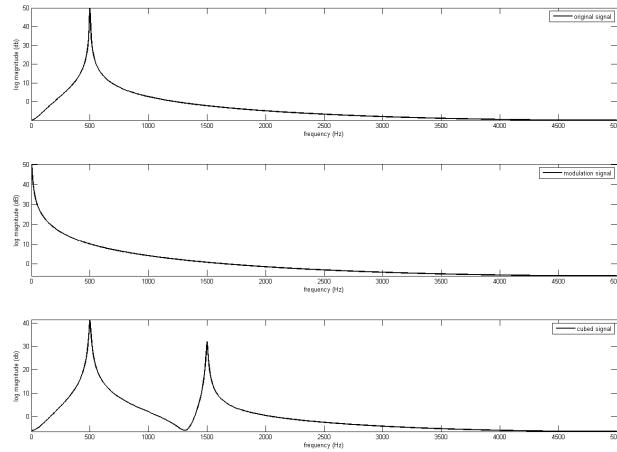


Figure P2.29.2: Log magnitude spectrum of modulated and cubed cosine wave.

which has a frequency response, $X_1(e^{j\omega})$, consisting of impulses at frequencies ± 200 Hz. A sketch of the digital frequency response is shown in the top panel of Figure P2.31.1.

- (b) In a similar manner, we can solve for the digital frequency response, $X_2(e^{j\omega})$, resulting from sampling $x_b(t)$, with sampling rate $F_s = 10000$ Hz, resulting in:

$$x_2[n] = \frac{B}{2} \cdot [e^{j2\pi 201n/F_s} + e^{-j2\pi 201n/F_s}] \quad (2.4)$$

which has a frequency response $X_2(e^{j\omega})$, consisting of impulses at frequencies ± 201 Hz. A sketch of the digital frequency response is shown in the bottom panel of Figure P2.31.1.

- (c) A digital signal, $x[n]$, is periodic with period P samples, if it obeys the relation:

$$x[n + P] = x[n] \text{ for all } n \text{ and for } P > 0 \quad (2.5)$$

It is clear that $x_1[n]$ is periodic and of period $P = 50$ samples. It is somewhat more difficult to see that the signal $x_2[n]$ is also periodic but of period $P = 10000$ samples since the cosine

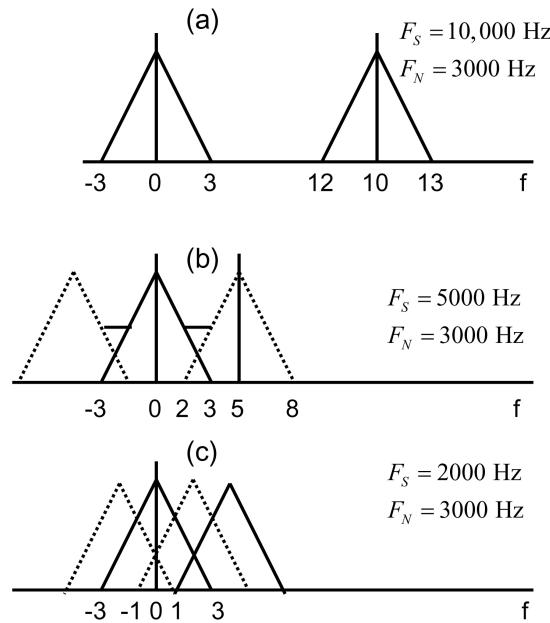


Figure P2.30.1: Digital spectra with different sampling rates.

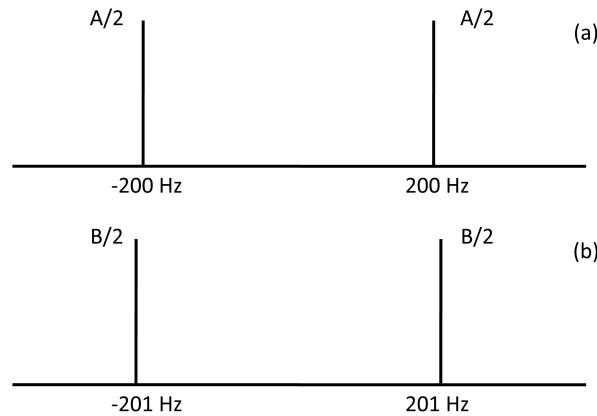


Figure P2.31.1: Frequency responses of the two signals.

function satisfies the relationship:

$$\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b) \quad (2.6)$$

$$x_2[n + 10000] = \cos(2\pi 201(n + 10000)/10000) \quad (2.7)$$

$$= \cos(2\pi 201n/10000)\cos(2\pi 201) \quad (2.8)$$

$$- \sin(2\pi 201n/10000)\sin(2\pi 201) \quad (2.9)$$

$$= \cos(2\pi 201n/10000) = x_2[n] \quad (2.10)$$

since $\cos(2\pi 201) = 1$ and $\sin(2\pi 201) = 0$ in the above equation.

Chapter 3

Fundamentals of Human Speech Production

- 3.1** (a) The regions of voiced speech, unvoiced speech and silence (background signal) are shown in Figure P3.1.1.
- (b) The pitch periods (in msec) are indicated by a bracket (in Figure P3.1.1) under the segment of speech to which the estimate corresponds.

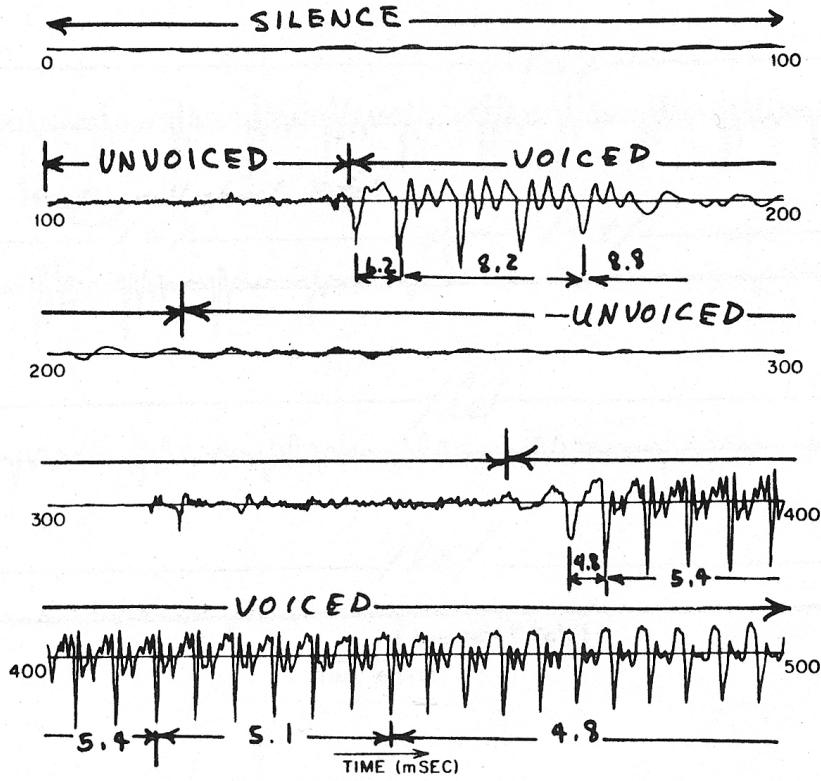


Figure P3.1.1: Locations of regions of voiced, unvoiced and silence.

- 3.2** (a) The approximate boundaries between the phonemes are marked in Figure P3.2.1.
 (b) The points of lowest and highest pitch are also marked in Figure P3.2.1.
 (c) The lowest pitch has a period of about 21.5 msec corresponding to a frequency of 46 Hz.
 This very low pitch is strongly indicative that the speaker is probably male.

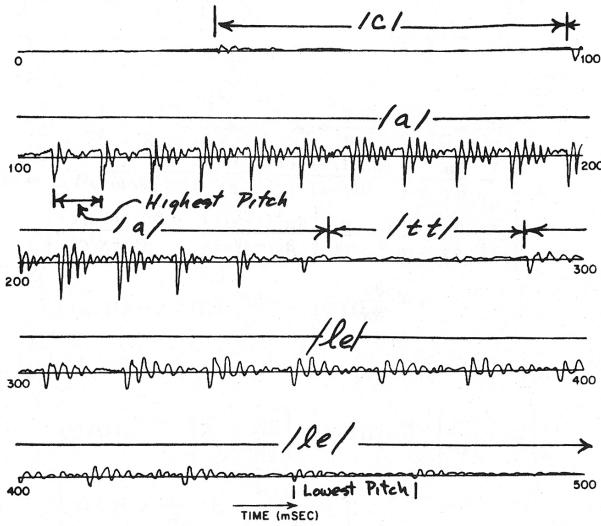


Figure P3.2.1: Time waveform of speech utterance “cattle” with phoneme boundaries and points of lowest and highest pitch marked on the waveform.

- 3.3**
1. /the/ can be pronounced as /DH/ /UH/ or /DH/ /IY/; /of/ is pronounced as /AA/ /V/; /and/ is pronounced as /AE/ /N/ /D/; /to/ is pronounced as /T/ /UW/; /a/ is pronounced either as /EY/ or /UH/; /in/ is pronounced as /IH/ /N/; /that/ is pronounced as /DH/ /AE/ /T/; /is/ is pronounced as /IH/ /Z/ (S optional); /was/ is pronounced as /W/ /AA/ /Z/ (S optional); /he/ is pronounced as /HH/ /IY/.
 2. /data/ can be pronounced as either /D/ /EY/ /T/ /AX/ or /D/ /AE/ /T/ /AX/; /lives/ can be pronounced as either /L/ /AY/ /V/ /Z/ (optional S) or /L/ /IH/ /V/ /Z/ (optional S); /record/ can be pronounced as either /R/ /IH/ /K/ /AO/ /R/ /D/ or /R/ /EH/ /K/ /ER/ /D/.
 3. /company/ is pronounced /K/ /UH/ /M/ /P/ /AX/ /N/ /IY/; /happiness/ is pronounced /HH/ /AE/ /P/ /IY/ /N/ /EH/ /S/; /willingness/ is pronounced as /W/ /IH/ /L/ /IH/ /NX/ /N/ /EH/ /S/.
 4. The sentence /I enjoy the simple life/ is pronounced as /AY/-/EH/ /N/ /JH/ /OY/-/DH/ /UH/- /S/ /IH/ /M/ /P/ /AX/ /L/-/L/ /AY/ /F/; the sentence /Good friends are hard to find/ is pronounced as /G/ /UH/ /D/-/F/ /R/ /EH/ /N/ /D/ /Z/ (optional S)-/AH/ /R/-/HH/ /AA/ /R/ /D/-/T/ /UW/- /F/ /AY/ /N/ /D/.
- *****

- 3.4** The sounds in the word /and/ are /AE/ /N/ /D/ and the (extremely) approximate locations of the sounds are:

- /AE/ samples 400-3800

- /N/ samples 3800-5400
- /D/ samples 5400-6100

The sounds in the word /that/ are /DH/ /AE/ /T/ and the (extremely) approximate locations of the sounds are:

- /TH/ samples 600-1600
- /AE/ samples 1600-3600
- /T/ samples 3600-4800

The sounds in the word /was/ are /W/ /AA/ /Z(S)/ and the (extremely) approximate locations of the sounds are:

- /W/ samples 400-1500
- /AA/ samples 1500-4800
- /Z(S)/ samples 4800-6600

The sounds in the word /by/ are /B/ /AY/ and the (extremely) approximate locations of the sounds are:

- /B/ samples 400-1700
- /AY/ samples 1700-6100

The sounds in the word /enjoy/ are /EH/ /N/ /JH/ /OY/ and the (extremely) approximate locations of the sounds are:

- /EH/ samples 600-1440
- /N/ samples 1440-2860
- /JH/ samples 2860-3500
- /OY/ samples 3500-8000

The sounds in the word /company/ are /K/ /UH/ /M/ /P/ /AX/ /N/ /IY/ and the (extremely) approximate locations of the sounds are:

- /K/ samples 500-1140
- /UH/ samples 1140-2000
- /M/ samples 2000-2440
- /P/ samples 2440-3600
- /AX/ samples 3600-4440
- /N/ samples 4440-5190
- /IY/ samples 5190-6222

The sounds in the word /simple/ are /S/ /IH/ /M/ /P/ /(AX L—EL)/ and the (extremely) approximate locations of the sounds are:

- /S/ samples 0-2590
- /IH/ samples 2590-3290
- /M/ samples 3290-3920
- /P/ samples 3920-5040

- /(AX/ samples 5040-5610
 - /L—EL)/ samples 5610-6200
- *****

- 3.5** Using the spectrogram (as shown in Figure P3.5.1) to locate the approximate centers of the three vowel regions as:

/enjoy/ - sample 5700 (0.57 sec) - center of /OY/ sound - pitch 80 samples or 125 Hz /simple/ - sample 10850 (1.085 sec) - center of /AX—EL/ sound - pitch 71 samples or 141 Hz /life/ - sample 16580 (1.658 sec) - center of /AY/ sound - pitch 88 samples or 114 Hz

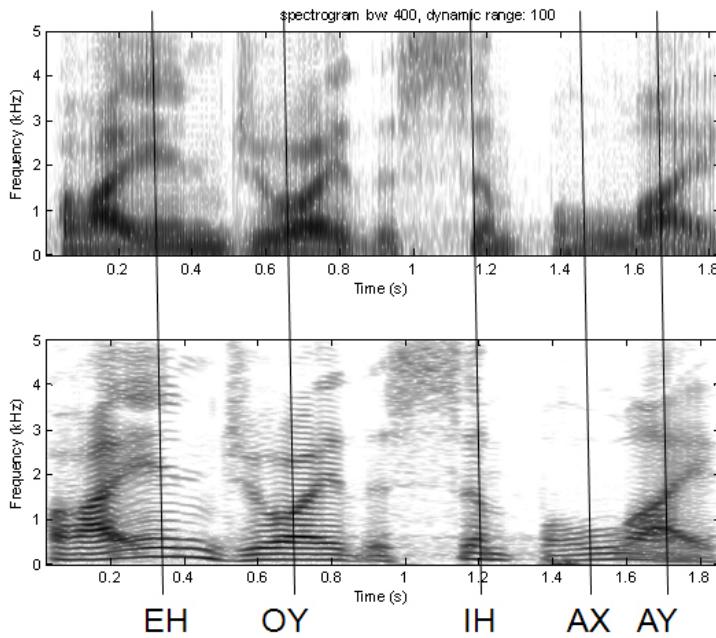


Figure P3.5.1: Locations of center of vowel sounds in utterance.

- 3.6** S samples 1-445 U samples 445-607 V samples 607-2363 U samples 2363-3828 V samples 3828-7032 U samples 7032-7450 V samples 7450-8466 U samples 8466-9144 V samples 9144-11435 U samples 11435-12599 V samples 12599-13202 U samples 13202-14222 V samples 14222-18631 U samples 18631-20357

The regions of voiced, unvoiced and silence are marked on the waveform plot of Figure P3.6.1.

- 3.7** (a) The region for the merged /D/ phonemes is approximately samples 4300-4700.
 (b) The region for the /IY/ sound in the word "each" is approximately samples 8800-9600.
 (c) The region for the /CH/ sound in the word "each" is approximately samples 10,000-10,500.

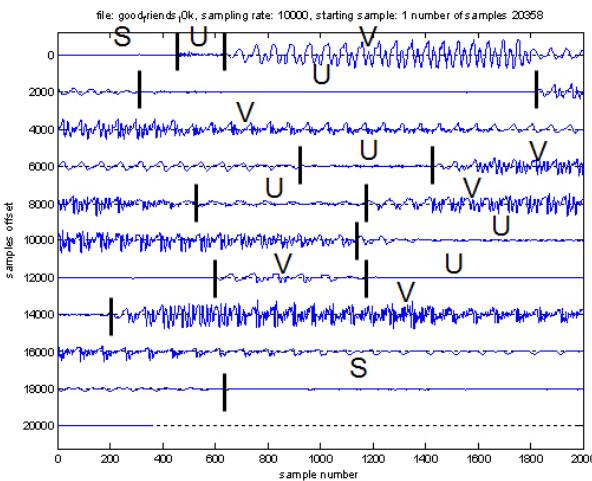


Figure P3.6.1: Waveform with marked regions for voiced, unvoiced, and silence (background) for the sentence “Good friends are hard to find”.

- (d) We estimate the fundamental frequency for the voiced segment on the first line of the waveform plot by counting the number of pitch cycles over the duration of the voiced region. We see that there are approximately 19 periods over the 1900 sample voiced region, giving a period of $1200/19=63.16$ samples. At the sampling rate of 8000 Hz, we convert the 63.16 samples to a period of 7.89 msec. Finally we convert the period to the pitch frequency giving an average fundamental frequency of 126.7 Hz.
 - (e) There are about 21 phones in the text of the spoken utterance and they take about 17,000 samples or 2.125 seconds at the sampling rate of 8000 samples/second. Assuming independent phonemes with 6 bits per phoneme, we estimate the bit rate of the utterance as the product of the number of phonemes/second (21/2.125) with the number of bits/phoneme (6) giving a total bit rate of about 60 bits/second.
- *****

- 3.8** The narrowband spectrogram is the one at the top. The narrowband spectrogram is characterized by a wide time duration with narrow frequency bandwidth; hence it is able to resolve individual pitch harmonics in frequency, leading to a series of pseudo-horizontal striations in the plot. The wideband spectrogram is characterized by a narrow time duration with wide frequency bandwidth; hence it resolves pitch periods in time, leading to a series of vertical striations in the plot.
- *****

- 3.9** (a) The bottom spectrogram is the wideband spectrogram.
 (b) The fundamental frequency at $t = 0.18$ seconds is estimated from the narrowband spectrogram where we see that at $t = 0.18$ seconds, there are about 19 harmonics in a frequency band from 0 to 2500 Hz, giving an estimate of the fundamental frequency of about 131 Hz.
 (c) From the narrowband spectrogram we see that the fundamental frequency is decreasing in the region from $t = 1.6$ to $t = 1.8$ seconds.
 (d) From the wideband spectrogram we can estimate the values for the first three formant frequencies as $F_1 = 700$ Hz, $F_2 = 1700$ Hz and $F_3 = 2400$ Hz.

- (e) The location of the merged /D/ phonemes is the region around $t = 0.6$ seconds (approximately).

- 3.10** By identifying the features of the four spectrograms and comparing them to the presumed features for the given words, it can readily be deduced that the four spectrograms correspond to the following words:

- The top left spectrogram corresponds to the word “was”
- The top right spectrogram corresponds to the word “enjoy”
- The bottom left spectrogram corresponds to the word “company”
- The bottom right spectrogram corresponds to the word “enjoy”

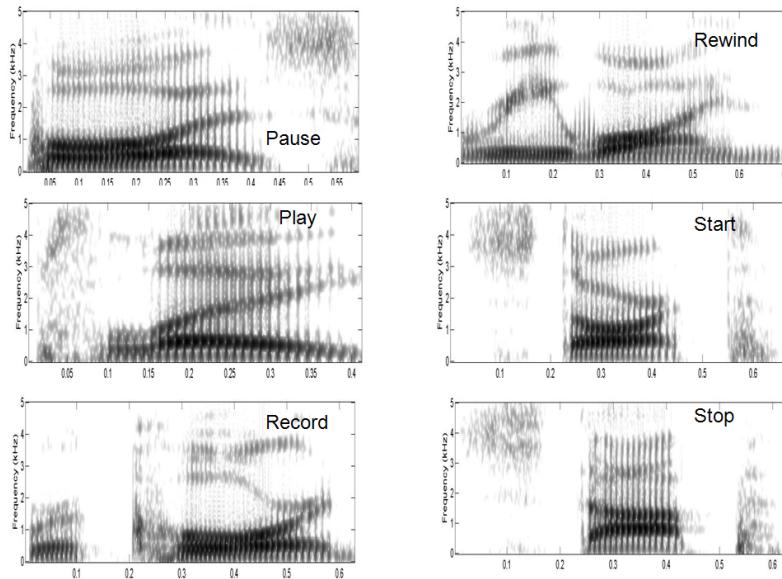


Figure P3.11.1: Spectrograms with word labels of one version of each of the control words for a voice controlled cassette tape system.

- 3.11** By examining the features of the six control words it is relatively easy to match the individual spectrograms to the spoken command words, as shown marked in Figure P3.11.1.

Thus the control word “rewind” is an all-voiced utterance with a voiced stop, /D/, at the end; hence its spectrogram is in the top row, second column. (Note that the positions of the spectrograms are different from that given in the problem statement.)

There are two words that start with strong fricatives, namely “start” and “stop”. The feature in the spectrogram that distinguishes these two control words is the falling third formant for the /R/ sound in “start”. Hence we can readily assign the spectrogram in the third row, first column to “start”, and the spectrogram in the second row, second column to “stop”.

There is only one spectrogram that ends with a strong fricative, namely “pause” and hence we can assign the spectrogram in the second row, first column to “pause”.

Of the remaining two spectrograms, corresponding to the words “record” and “play”, we can easily see that the spectrogram in the third row, second column has an initial stop consonant (/P/) and is followed by an all-voiced region; hence it represents the word “play”.

Finally by the process of elimination, the remaining spectrogram in the first row, second column corresponds to the word “record”. We see the stop gap of the /K/ in “record” which serves to verify the analysis for this word.

3.12 The sounds of these two sentences are as follows:

1. “She eats some Mexican nuts”: /SH/ /IY/ - /IY/ /T/ /S/- /S/ /UH/ /M/ - /M/ /EH/ /K/ /S/ /IH/ /K/ /IH/ /N/ - /N/ /UH/ /T/ /S/.
2. ”Where roads top providing good driving”: /WH/ /EH/ /R/ - /R/ /OW/ /D/ /S/ - /S/ /T/ /AA/ /P/ - /P/ /R/ /UH/ /V/ /AY/ /D/ /IH/ /NX/ - /G/ /UH/ /D/ - /D/ /R/ /AY/ /V/ /IH/ /NX/.

The sound at the end of each word is virtually identical to the sound at the beginning of the following word – hence there is a high degree of sound co-articulation across words, making it virtually impossible to reliably identify word boundaries in this spoken context.

3.13 It is relative easy to decode the text of the sentence without the vowels as “To give you some idea of the amount of work required in this course”. However, without the consonants, the text is virtually undecodable. The actual sentence is “Bear in mind that it accounts for half the grade in this course”.

3.14 The most commonly occurring words in this list are mono-syllabic function words which account for the top 56 word, and 93 of the top 105 words in the list. About 89% of the top 105 words from this list are mono-syllabic, so about 11% have more than one syllable.

The poly-syllabic words in this list are:

“about”, “into”, “only”, “any”, “over”, “even”, “after”, “alos”, “many”, “before”, and “because”.

3.15 The way to determine the most likely vowels is to measure the first three formant frequencies (in or around the middle of the vowels) and then compute a weighted distance between the formants of the unknown vowel sound, and the formants in the vowel chart of Chapter 3. Using this criterion (and using various MATLAB tools to measure the frequencies accurately, and to measure distances to target vowels accurately) we get the following matches:

1. row 1, column 1 – the measured formants are 703, 1113, and 2402 Hz, and the closest vowel match is /AA/ with a weighted distance of 0.05, and no other vowel being close to this minimum distance. The word that was spoken was “sob” and the vowel was /AA/.
2. row 1, column 2 – the measured formants are 557, 1299 and 1563 Hz, and the closest vowel match is /ER/ with a weighted distance of 0.16, with no other vowel being close to this minimum distance. The word that was spoken was “bird” and the vowel was /er/.

3. row 2, column 1 – the measured formants are 313, 2510 and 2939 Hz, and the closest vowel match is /IY/ with a weighted distance of 0.19, with no other vowel being close to this minimum distance. The word that was spoken was “cease” and the vowel was /iy/.
4. row 2, column 2 – the measured formants are 596, 1943 and 2930 Hz, and the closest vowel match is /EH/ with a weighted distance of 0.23 with no other vowel being close to this minimum distance. The word that was spoken was “set” and the vowel was /EH/.
5. row 3, column 1 – the measured formants are 811, 1855 and 2734 Hz, and the closest vowel match is /AE/ with a weighted distance of 0.28 with no other vowel being close to this minimum distance. The word that was spoken was “sat” and the vowel was /AE/.
6. row 3, column 1 – the measured formants are 283, 830 and 2793 Hz, and the closest vowel match is /UW/ with a weighted distance of 0.26 with no other vowel being close to this minimum distance. The word that was spoken was “boot” and the vowel was /UW/.

- 3.16** The transcription and the place and manner of articulation for the consonant sounds of the sentence “I enjoy the simple life” are as follows: /AY/ - /EH/ /N/ /JH/ /OY/ - /DH/ /UH/ - /S/ /IH/ /M/ /P/ /AX/ /L (EL)/ - /L/ /AY/ /F/

- /N/ – alveolar, voiced nasal sound
- /JH/ – alveolar-palatal, voiced fricative
- /DH/ – dental, voiced fricative
- /S/ – alveolar, unvoiced fricative
- /M/ – bilabial, voiced nasal
- /P/ – bilabial, stop
- /L/ – alveolar, voiced glide
- /F/ – labiodental, unvoiced fricative

- 3.17** Figure P3.12.1 shows a list of the word-initial consonants that occur at the beginning of English words. The word-initial consonant pairs consist of the following combinations:

1. /HH/ followed by /W/ and /Y/
2. /B/ followed by /L/, /R/ and /Y/
3. /D/ followed by /R/ and /W/
4. /G/ followed by /L/, /R/ and /W/
5. /P/ followed by /L/, /R/, /W/ and /Y/
6. /T/ followed by /R/, /W/, /Y/ and /S/
7. /K/ followed by /L/, /R/, /W/ and /Y/
8. /M/ followed by /W/ and /Y/
9. /V/ followed by /W/ and /Y/
10. /Z/ followed by /L/ and /W/
11. /F/ followed by /L/, /R/ and /Y/
12. /TH/ followed by /R/ and /W/

-	of	hy	human	sf	sphere	tr	true
b	be	j̥	just	sk	school	ts	tsunami
bl	black	k	can	skl	sclerosis	tw	twenty
br	bring	kł	class	skr	screen	ty	tuesday
by	beauty	k̥t̥	cross	skw	square	θ̥	thief
č	child	kw̥	quite	sky	skewer	θr̥	through
d	do	ky	curious	sl̥	slow	θw̥	thwart
dr	drive	l̥	like	sm̥	small	ð̥	the
dw	dwell	m̥	more	sn̥	snake	v̥	very
f	for	mw̥	moire	sp̥	special	vw̥	voyager
fl	floor	my	music	spl̥	split	vy̥	view
fr	from	n̥	not	spr̥	spring	w̥	was
fy	few	p̥	people	spy̥	spurious	y̥	you
g	good	pl̥	place	st̥	state	z̥	zero
gl	glass	pr̥	price	str̥	street	zl̥	zloty
gr	great	pw̥	pueblo	sw̥	sweet	zw̥	zweiback
gw	guava	py̥	pure	š̥	she	ž̥	genre
h	he	r̥	right	šr̥	shrewd		
hw	which	s̥	so	t̥	to		

Figure P3.12.1: List of word-initial consonants that occur at the beginning of English words.

13. /S/ followed by /L/, /W/, /M/, /N/, /P/, /T/, /K/ and /F/
14. /SH/ followed by /R/

The general rule is a consonant followed by a glide, or the fricative /S/ followed by a glide, a nasal, a voiceless stop or the fricative /F/. The place of articulation of the initial consonant and the place of articulation of the following glide sound (/W/, /L/, /R/ or /Y/) is generally unrelated as the articulators glide between initial configuration and that of the glide.

The only word-initial consonant triplets are the combinations: 1. /S/ /K/ followed by /L/, /R/, /W/ and /Y/ 2. /S/ /P/ followed by /L/, /R/ and /Y/ 3. /S/ /T/ followed by /R/ The general rule here is /S/ as the initial consonant, followed by an unvoiced stop (/K/ or /P/ or /T/) followed by a glide (/L/, /R/, /W/ or /Y/). Again, the place of articulation of the initial /S/ consonant (alveolar) is generally unrelated to the place of articulation of the stop consonant that follows (bilabial for /P/, alveolar for /T/ and velar for /K/) or the place of articulation of the glide sound (/W/, /L/, /R/ or /Y/) that follows the word-initial consonant pairs.

Chapter 4

Hearing, Auditory Models, and Speech Perception

- 4.1** The outer ear funnels sound into the ear canal utilizing the pinna structure to capture the most possible sound and to direct it to the middle ear. The middle ear converts the acoustical sound wave captured by the outer ear into mechanical vibrations that get transmitted to the inner ear. The middle ear uses a mechanical transducer consisting of the malleus (hammer), incus (anvil), and stapes (stirrup) to convert from an acoustical to a mechanical wave. The inner ear does a time-frequency analysis of the mechanical vibrations from the middle ear, along the basilar membrane, providing a temporal and spectral analysis of the incoming sound that is sent to the higher processing centers in the brain via the auditory nerve.

- 4.2** Performing a conventional Fourier series analysis we determine the frequency coefficients, a_k as:

$$\begin{aligned} a_k &= \frac{1}{T} \left[\int_{-T/2}^0 (-1)e^{-j(2\pi/T)kt} dt + \int_0^{T/2} (1)e^{-j(2\pi/T)kt} dt \right] \\ &= \frac{1}{T} \left[\frac{-1}{-j(2\pi/T)k} e^{-j(2\pi/T)kt} \Big|_{-T/2}^0 + \frac{1}{-j(2\pi/T)k} e^{-j(2\pi/T)kt} \Big|_0^{T/2} \right] \\ &= \frac{1}{j2\pi k} [1 - e^{j\pi k}] - \frac{1}{j2\pi k} [e^{-j\pi k} - 1] \\ &= \frac{1 - \cos(\pi k)}{j\pi k} = \begin{cases} \frac{2}{j\pi k} & k = 1, 3, 5, \dots \\ 0 & k = 0, 2, 4, \dots \end{cases} \end{aligned}$$

hence we see that all even harmonics are zero-valued and the signal $s(t)$ consists strictly of odd harmonics. With a period of $T = 10$ msec, we get harmonics at frequencies 100, 300, 500, ... Hz. The frequency selectivity properties of the basilar membrane process the low frequency components of the input signal at the apical end and the high frequency components at the stapes end of the membrane. Hence what we see at the apex is the sum of the low frequency components (100 Hz, 300 Hz tones), while at the stapes end we see the sum of the high frequency components (around 10 kHz). In both cases we retain the odd symmetry of the input signal as only the odd harmonic components occur.

- 4.3** Using the chart that shows the relations between the physical quantity of Intensity Level (measured in dB) and the perceived quantity of Loudness Level (measured in phons), we get the following results:

1. – a 20 dB *IL* 1000 Hz tone has an *LL* of 20 phons whereas a 20 dB *IL* tone at 500 Hz has an *LL* of 16 phons \Rightarrow the 1000 Hz tone is perceived as louder by 4 phons.
2. – a 40 dB *IL* tone at 200 Hz has an *LL* of 20 phons whereas a 30 dB *IL* tone at 2000 Hz has an *LL* of 33 phons \Rightarrow the 2000 Hz tone is perceived as louder by 13 phons
3. – a 50 dB *IL* tone at 100 Hz has an *LL* of 20 phons whereas a 50 dB *IL* tone at 1000 Hz has an *LL* of 50 phons \Rightarrow the 1000 Hz tone is perceived as louder by 30 phons

- 4.4** The perceived pitch (in Mels) is related to the physical frequency by the relation:

$$\text{Pitch (mels)} = 3323 \log_{10}(1 + f/1000)$$

and the critical bandwidth is approximately constant until about 1000 Hz and increases logarithmically above 1000 Hz. We can determine the actual critical bandwidths from the chart in the chapter.

Thus the pitch in mels and critical bandwidth (in Hz) for each of the tones is as follows:

1. – 100 Hz 137 mels perceived pitch, 60 Hz critical bandwidth
2. – 200 Hz 263 mels perceived pitch, 40 Hz critical bandwidth
3. – 400 Hz 485 mels perceived pitch, 35 Hz critical bandwidth
4. – 1000 Hz 1000 mels perceived pitch, 40 Hz critical bandwidth
5. – 2000 Hz 1541 mels perceived pitch, 60 Hz critical bandwidth
6. – 4000 Hz 2252 mels perceived pitch, 135 Hz critical bandwidth
7. – 10000 Hz 3460 mels perceived pitch, 600 Hz critical bandwidth

- 4.5** At a signal-to-noise ratio of 12 dB the confusions are mainly in place, but not in manner of production as shown in the table in the lecture. At a signal-to-noise ratio of 6 dB, the confusions are in both place and manner of production, as shown in the table in the lecture. Referring to the measured confusions from the table we get the following results for the word pairs shown above:

Word Pair	12 dB S/N confusions	-6 dB S/N confusions
pick-tick	confusions due to /P/-/K/	confusions due to /P/-/K/
peek-seek	no confusions	confusions due to /P/-/S/
take-bake	no confusions	no confusions
king-sing	no confusions	confusions due to /K/-/S/
go-doe	confusions due to /G/-/D/	confusions due to /G/-/D/
van-than	confusions due to /V/-/TH/	confusions due to /V/-/TH/
map-nap	no confusions	confusions due to /M/-/N/
go-no	no confusions	no confusions

Chapter 5

Sound Propagation in the Human Vocal Tract

5.1 Write $g_c(t)$ as follows:

$$g_c(t) = 0.5g_1(t)(1 - \cos \Omega_1 t) + g_2(t - T_1) \cos \Omega_2(t - T_1)$$

where

$$g_1(t) = \begin{cases} 1 & 0 \leq t \leq T_1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad g_2(t) = \begin{cases} 1 & 0 \leq t \leq T_2 \\ 0 & \text{otherwise} \end{cases}$$

and where $\Omega_1 = 2\pi/(2T_1)$ and $\Omega_2 = 2\pi/(4T_2)$.

- (a) From $g_c(t)$ written as above, the answer follows from the modulation and delay theorems of discrete-time Fourier transforms. That is,

$$G_c(\Omega) = 0.5G_1(\Omega) - 0.25G_1(\Omega - \Omega_1) - 0.25G_1(\Omega + \Omega_1) + 0.5[G_2(\Omega - \Omega_2) + G_2(\Omega + \Omega_2)]e^{-j\Omega T_1}$$

where

$$G_1(\Omega) = \int_0^{T_1} e^{-j\Omega t} dt = \frac{1 - e^{-j\Omega T_1}}{j\Omega} \quad \text{and} \quad G_2(\Omega) = \int_0^{T_2} e^{-j\Omega t} dt = \frac{1 - e^{-j\Omega T_2}}{j\Omega}$$

- (b) Now represent the periodic glottal excitation as

$$u_G(t) = \sum_{k=-\infty}^{\infty} g_c(t - kT_0) = g_c(t) * p(t) \quad \text{where} \quad p(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_0)$$

Therefore, $U_G(\Omega) = G_c(\Omega)P(\Omega)$ where $P(\Omega) = \frac{2\pi}{T_0} \sum_{k=-\infty}^{\infty} \delta(\Omega - 2\pi k/T_0)$, so that

$$U_G(\Omega) = \frac{2\pi}{T_0} G_c(\Omega) \sum_{k=-\infty}^{\infty} \delta(\Omega - 2\pi k/T_0) = \frac{2\pi}{T_0} \sum_{k=-\infty}^{\infty} G_c(2\pi k/T_0) \delta(\Omega - 2\pi k/T_0)$$

- (c) The plot of the spectra of glottal pulses with a range of closing times is given in Figure P5.1.1. Observe that as the closing time T_2 increases relative to the opening time T_1 , the high frequencies are more attenuated.
- (d) The comparison between the spectra of the analog and sampled glottal pulses is given in Figure P5.1.2. Note that the effect of aliasing in sampling the lowpass glottal pulse is limited to the high frequencies.

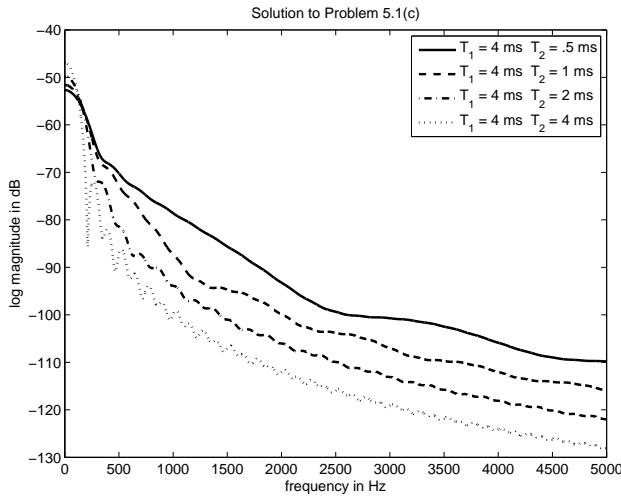


Figure P5.1.1: Spectra of glottal pulses with a range of closing times.

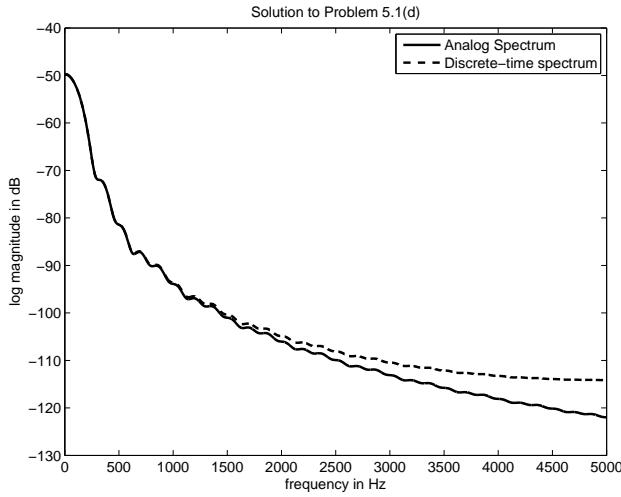


Figure P5.1.2: Comparison of spectra of analog and sampled glottal pulses.

The MATLAB program for generating the plots is

```
%solution to Problem 5.1
function problem5_1
F=0:5000;Om=2*pi*F;fsize=12;
Gc=zeros(5001,4);T1=.004; Om1=pi/T1;
for m=1:4
    T2=.0005*2^(m-1)
    Om2=.5*pi/T2;
    Gc(:,m)=.5*P(Om,T1) -.25*P((Om-Om1),T1)-.25*P((Om+Om1),T1)...
        +.5*P((Om-Om2),T2).*exp(-j*(Om)*T1)+.5*P((Om+Om2),T2).*exp(-j*(Om)*T1);
end
figure(1)
h=plot(F,20*log10(abs(Gc(:,1))),'-',F,20*log10(abs(Gc(:,2))), '--',...
```

```

F,20*log10(abs(Gc(:,3))),'-.',F,20*log10(abs(Gc(:,4))),':');
set(h,'linewidth',2)
h=legend('T_1 = 4 ms T_2 = .5 ms','T_1 = 4 ms T_2 = 1 ms',...
    'T_1 = 4 ms T_2 = 2 ms',...
    'T_1 = 4 ms T_2 = 4 ms');
set(h,'fontsize',fsiz)
axis([0,5000,-130,-40])
title('Solution to Problem 5.1(c)','fontsize',fsiz)
set(gca,'fontsize',fsiz)
ylabel('log magnitude in dB','fontsize',fsiz)
xlabel('frequency in Hz','fontsize',fsiz)
print -deps prob5_1c.eps
% solution to part d
T=1/10000;T1=.004; T2=.002;
N1=T1/T;N2=T2/T;
n1=0:N1;n2=N1+1:N1+N2;
g=[.5*(1-cos(pi*n1/N1)),cos(pi*(n2-N1)/(2*N2))];
Gd=T*freqz(g,1,0m*T);
0m1=pi/T1;0m2=.5*pi/T2;
Gc=.5*P(0m,T1) -.25*P((0m-0m1),T1) -.25*P((0m+0m1),T1)...
    +.5*P((0m-0m2),T2).*exp(-j*(0m)*T1)+.5*P((0m+0m2),T2).*exp(-j*(0m)*T1);
figure(2)
h=plot(F,20*log10(abs(Gc)),'-',F,20*log10(abs(Gd)), '--');
set(h,'linewidth',2)
h=legend('Analog Spectrum','Discrete-time spectrum');
set(h,'fontsize',fsiz)
axis([0,5000,-130,-40])
title('Solution to Problem 5.1(d)', 'fontsize',fsiz)
set(gca,'fontsize',fsiz)
ylabel('log magnitude in dB','fontsize',fsiz)
xlabel('frequency in Hz','fontsize',fsiz)
print -deps prob5_1d.eps

function G=P(0m,T)
G=(1-exp(-j*0m*T))./(j*0m);

```

5.2 From the basic equations we have:

$$u(x, t) = u^+(t - x/c) - u^-(t + x/c)$$

$$p(x, t) = \frac{\rho c}{A} [u^+(t - x/c) + u^-(t + x/c)]$$

Taking partial derivatives we get:

$$\begin{aligned}
-\frac{\partial p}{\partial x} &= -\frac{\rho c}{A} \left[\frac{\partial u^+(t - x/c)}{\partial(t - x/c)} \cdot \left[-\frac{1}{c} \right] + \frac{\partial u^-(t + x/c)}{\partial(t + x/c)} \cdot \frac{1}{c} \right] \\
&= \frac{\rho}{A} \left[\frac{\partial u^+(t - x/c)}{\partial(t - x/c)} - \frac{\partial u^-(t + x/c)}{\partial(t + x/c)} \right] \\
&= \frac{\rho}{A} \frac{\partial u}{\partial t}
\end{aligned}$$

In a similar manner we derive the relations:

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{1}{c} \cdot \frac{\partial u^+(t - x/c)}{\partial(t - x/c)} - \frac{1}{c} \cdot \frac{\partial u^-(t + x/c)}{\partial(t + x/c)} \\ \frac{A}{\rho c^2} &= \frac{A}{\rho c^2} \cdot \frac{\rho c}{A} \left[\frac{\partial u^+(t - x/c)}{\partial(t - x/c)} - \frac{\partial u^-(t + x/c)}{\partial(t + x/c)} \right] \\ &= -\frac{\partial u}{\partial x}\end{aligned}$$

5.3 (a) Since both $A_k \geq 0$ and $A_{k+1} \geq 0$, we can consider two cases, namely:

1. $A_{k+1} > A_k$, in which case we have $A_{k+1}/A_k > 1$ or, equivalently, we have $A_{k+1}/A_k = 1 + \Delta$, where $\infty > \Delta > 0$ in which case we can solve for r_k as:

$$\begin{aligned}r_k &= \frac{\frac{A_{k+1}}{A_k} - 1}{\frac{A_{k+1}}{A_k} + 1} \\ r_k &= \frac{\Delta}{2 + \Delta}\end{aligned}$$

As $\Delta \rightarrow 0$, $r_k \rightarrow 0$; similarly as $\Delta \rightarrow \infty$, $r_k \rightarrow 1$.

2. $A_k > A_{k+1}$, in which case we have $A_k/A_{k+1} > 1$ or, equivalently, we have $A_k/A_{k+1} = 1 + \Delta$, where $\infty > \Delta > 0$ in which case we can solve for r_k as:

$$\begin{aligned}r_k &= \frac{1 - \frac{A_k}{A_{k+1}}}{1 + \frac{A_k}{A_{k+1}}} \\ r_k &= \frac{-\Delta}{2 + \Delta}\end{aligned}$$

As $\Delta \rightarrow 0$, $r_k \rightarrow 0$; similarly as $\Delta \rightarrow \infty$, $r_k \rightarrow -1$.

- (b) In the limiting cases of part (a) above, namely when $A_k/A_{k+1} \rightarrow \infty$ then $r_k \rightarrow -1$; similarly when $A_{k+1}/A_k \rightarrow \infty$ then $r_k \rightarrow +1$. Between the limiting cases $-1 \leq r_k \leq +1$.

5.4 (a) A lossy tube cannot have any unattenuated resonances; i.e., resonances with infinite gain

- (b) A non-rigid uniform tube will have resonances higher than a rigid uniform tube. Broader bandwidths, as well as more attenuation, are also observed for the lower resonances in non-rigid tubes.

- (c) A tube of N concatenated uniform sections, when used to model the vocal tract, is equivalent to finite sampling of a continuous time function. As a result, repetition of the resonant modes in frequency is observed, reminiscent of continuous-time to discrete-time conversion. With a vocal tract length of 17.5 cm and a sound speed of 35,000 cm/sec, the delay for sound to propagate through each section of an N -section tube is $t_0 = 1/(2000N)$ seconds, which is equivalent to a sampling rate of $F_s = 1000N$ samples/sec. Thus, the resonance modes (resonant frequencies of the vocal tract) are symmetric with respect to nF_s where n is an integer. If the tube is not represented by a set of N uniform sections, the frequency at which the resonances would repeat is infinity (i.e., no reflection of resonances).

- (d) Consider a uniform tube as an example. The resonances are inversely proportional to the length of the tube. Hence, if the length fluctuates, the resonances will also fluctuate in an inverse manner.

5.5 (a) Starting with the parameters:

$$Z_L = \frac{j\Omega R_r L_r}{R_r + j\Omega L_r}, \quad R_r = \frac{128}{9\pi^2}, \quad L_r = \frac{8a}{3\pi c}$$

we have the relations:

$$\begin{aligned} P_N(l_n, \Omega) &= Z_L U_N(l_N, \Omega) \\ P_N(l_n, \Omega) &= \frac{\rho c}{A_N} [U_N^+(\Omega) e^{-j\Omega\tau_N} + U_N^-(\Omega) e^{j\Omega\tau_N}] \\ U_N(l_N, \Omega) &= U_N^+(\Omega) e^{-j\Omega\tau_N} - U_N^-(\Omega) e^{j\Omega\tau_N} \end{aligned}$$

Therefore we get:

$$\begin{aligned} \frac{\rho c}{A_N} [U_N^+(\Omega) e^{-j\Omega\tau_N} + U_N^-(\Omega) e^{j\Omega\tau_N}] &= Z_L [U_N^+(\Omega) e^{-j\Omega\tau_N} - U_N^-(\Omega) e^{j\Omega\tau_N}] \\ U_N^+(\Omega) e^{-j\Omega\tau_N} \left[Z_L - \frac{\rho c}{A_N} \right] &= U_N^-(\Omega) e^{j\Omega\tau_N} \left[\frac{\rho c}{A_N} + Z_L \right] \end{aligned}$$

- (b) We now have:

$$\begin{aligned} U_N^+(\Omega) e^{-j\Omega\tau_N} \left[j\Omega L_r R_r - \frac{\rho c}{A_N} R_r - \frac{-j\Omega L_r \rho c}{A_N} \right] \\ = U_N^-(\Omega) e^{j\Omega\tau_N} \left[j\Omega L_r R_r + \frac{\rho c R_r}{A_N} + \frac{j\Omega L_r \rho c}{A_N} \right] \\ L_r \left[R_r + \frac{\rho c}{A_N} \right] \frac{d}{dt} [u_N^-(t + \tau_N)] + \frac{\rho c R_r}{A_N} [u_N^-(t + \tau_N)] = \\ L_r \left[R_r - \frac{\rho c}{A_N} \right] \frac{d}{dt} [u_N^+(t - \tau_N)] - \frac{\rho c R_r}{A_N} [u_N^+(t - \tau_N)] = \end{aligned}$$

5.6 We can express r_G as:

$$r_G = \frac{Z_G - \frac{\rho c}{A_1}}{Z_G + \frac{\rho c}{A_1}}$$

$$u_1^+(t) = \frac{(1 + r_G)}{2} u_G(t) + r_G u_1^-(t)$$

We now substitute for r_G giving:

$$\begin{aligned} u_1^+(t) &= \left[\frac{1}{2} + \frac{Z_G - \frac{\rho c}{A_1}}{2(Z_G + \frac{\rho c}{A_1})} \right] u_G(t) + \left[\frac{Z_G - \frac{\rho c}{A_1}}{Z_G + \frac{\rho c}{A_1}} \right] u_1^-(t) \\ \left[Z_G + \frac{\rho c}{A_1} \right] u_1^+(t) &= \frac{1}{2} \left[Z_G + \frac{\rho c}{A_1} + Z_G - \frac{\rho c}{A_1} \right] u_G(t) + \left[Z_G - \frac{\rho c}{A_1} \right] u_1^-(t) \end{aligned}$$

$$Z_G[u_1^+(t) - u_1^-(t)] = Z_G u_G(t) - \frac{\rho c}{A_1} [u_1^+(t) + u_1^-(t)]$$

$$u_1^+(t) - u_1^-(t) = u_G(t) - \frac{\rho c}{A_1} \left[\frac{u_1^+(t) + u_1^-(t)}{Z_G} \right]$$

5.7 From the signal flow diagram of Figure P5.7.1 we obtain the frequency-domain relations:

$$(1) \quad U_1^+(\Omega) = \frac{(1+r_G)}{2} U_G(\Omega) + r_G U_1^-(\Omega)$$

$$(2) \quad U_2^+(\Omega) = (1+r_1)U_1^+(\Omega)e^{-j\Omega\tau_1} + r_1 U_2^-(\Omega)$$

$$(3) \quad U_1^-(\Omega)e^{j\Omega\tau_1} = -r_1 U_1^+(\Omega)e^{-j\Omega\tau_1} + (1-r_1) U_2^-(\Omega)$$

$$(4) \quad U_L(\Omega) = (1+r_L)U_2^+(\Omega)e^{-j\Omega\tau_2}$$

$$(5) \quad U_2^-(\Omega)e^{j\Omega\tau_2} = -r_L U_2^+(\Omega)e^{-j\Omega\tau_2}$$

From (1) we get:

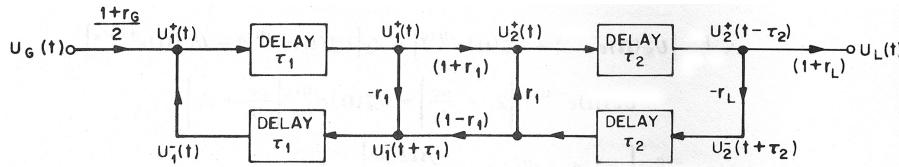


Figure P5.7.1: Model of two-tube vocal tract transmission.

$$U_G(\Omega) = \frac{2U_1^+(\Omega)}{1+r_G} - \frac{2r_G}{1+r_G} U_1^-(\Omega)$$

Next we solve for $U_1^+(\Omega)$ and $U_1^-(\Omega)$ in terms of $U_2^+(\Omega)$ and $U_2^-(\Omega)$. From (2) we get

$$U_1^+(\Omega) = \frac{U_2^+(\Omega)e^{j\Omega\tau_1} - r_1 U_2^-(\Omega)e^{j\Omega\tau_1}}{1+r_1}$$

From (3) we get

$$U_1^-(\Omega) = -r_1 U_1^+(\Omega)e^{-j2\Omega\tau_1} + (1-r_1) U_2^-(\Omega)e^{-j\Omega\tau_1}$$

Substitution for $U_1^+(\Omega)$ gives:

$$U_1^-(\Omega) = \frac{-r_1 U_2^+(\Omega)e^{-j\Omega\tau_1} + U_2^-(\Omega)e^{-j\Omega\tau_1}}{1+r_1}$$

From (4) we get

$$U_2^+(\Omega) = \frac{U_L(\Omega)e^{j\Omega\tau_2}}{(1+r_L)}$$

From (4) and (5) we get:

$$U_2^-(\Omega) = \frac{-r_L U_L(\Omega)e^{-j\Omega\tau_2}}{(1+r_L)}$$

Substituting into the expressions for $U_1^+(\Omega)$ and $U_1^-(\Omega)$ gives

$$U_1^+(\Omega) = \frac{U_L(\Omega)e^{j\Omega\tau_1}}{(1+r_L)(1+r_1)} [e^{j\Omega\tau_2} + r_1 r_L e^{-j\Omega\tau_2}]$$

$$U_1^-(\Omega) = \frac{U_L(\Omega)e^{-j\Omega\tau_1}}{(1+r_L)(1+r_1)}[-r_1e^{j\Omega\tau_2} - r_L e^{-j\Omega\tau_2}]$$

Substitution into the expression for $U_G(\Omega)$ gives:

$$U_G(\Omega) = \frac{2U_L(\Omega)e^{j\Omega(\tau_1+\tau_2)}}{(1+r_G)(1+r_L)(1+r_1)}[1 + r_1r_L e^{-j\Omega 2\tau_2} + r_1r + Ge^{-j\Omega 2\tau_1} + r_L r_G e^{-j2\Omega(\tau_1+\tau_2)}]$$

$$\frac{U_L(\Omega)}{U_G(\Omega)} = \frac{0.5(1+r_G)(1+r_L)(1+r_1)e^{-j\Omega(\tau_1+\tau_2)}}{1 + r_1r_G e^{-j\Omega 2\tau_1} + r_1r_L e^{-j\omega 2\tau_2} + r_L r_G e^{-j\Omega 2(\tau_1+\tau_2)}}$$

- 5.8 (a)** We define the parameters $\tau_1 = l_1/c$ and $\tau_2 = l_2/c$. Assuming $r_G = r_L = 1$ and $r_1 = (A_2 - A_1)/(A_2 + A_1)$, then:

$$V_a(s) = \frac{0.5(2)(2)(1+r_1)e^{-s(\tau_1+\tau_2)}}{1 + r_1(e^{-s^2\tau_1} + e^{-s^2\tau_2}) + e^{-s^2(\tau_1+\tau_2)}}$$

The poles are values of s such that:

$$1 + r_1(e^{-s^2\tau_1} + e^{-s^2\tau_2}) + e^{-s^2(\tau_1+\tau_2)} = 0$$

or, equivalently,

$$0 = e^{-s(\tau_1+\tau_2)}[e^{s(\tau_1+\tau_2)} + r_1(e^{-s\tau_1+s\tau_2} + e^{-s\tau_2+s\tau_1}) + e^{-s(\tau_1+\tau_2)}]$$

$$0 = e^{-2(\tau_1+\tau_2)}[2 \cosh[s(\tau_1 + \tau_2)] + 2r_1 \cosh[s(\tau_2 - \tau_1)]]$$

Since this is a lossless system, the poles must be on the $s = j\Omega$ axis. Then the hyperbolic functions reduce to circular functions, i.e., the poles satisfy the relationship:

$$\cos[\Omega(\tau_1 + \tau_2)] + r_1 \cos[\Omega(\tau_2 - \tau_1)] = 0$$

To get the other form, substitute for r_1 giving:

$$\cos[\Omega(\tau_1 + \tau_2)] + \frac{A_2 - A_1}{A_2 + A_1} \cos[\Omega(\tau_2 - \tau_1)] = 0$$

$$A_2(\cos[\Omega(\tau_1 + \tau_2)] + \cos[\Omega(\tau_2 - \tau_1)])$$

$$+ A_1(\cos[\Omega(\tau_1 + \tau_2)] - \cos[\Omega(\tau_2 - \tau_1)]) = 0$$

$$2A_2 \cos(\Omega\tau_1) \cos(\Omega\tau_2) - 2A_1 \sin(\Omega\tau_1) \sin(\Omega\tau_2) = 0$$

$$\frac{A_1}{A_2} \tan(\Omega\tau_2) = \cot(\Omega\tau_1); \quad \tau_1 = l_1/c; \tau_2 = l_2/c$$

- (b)** Now substituting for l_1, l_2, A_2 we get a transcendental equation to solve either graphically or iteratively, giving:

Vowel	F_1	F_2	F_3
/IY/	256	1905	2917
/AE/	646	1830	2358
/AA/	789	1276	2808
/UH/	515	1544	2574

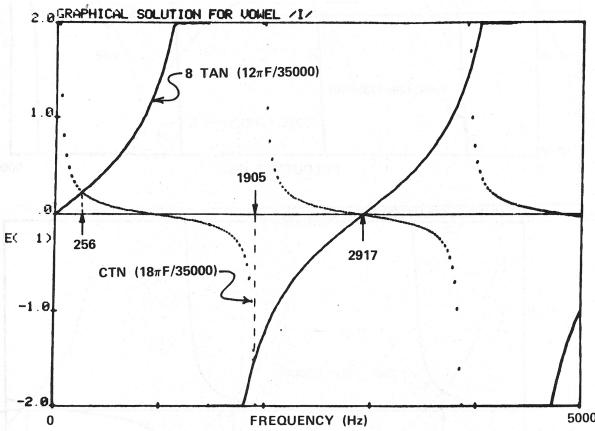


Figure P5.8.1: Graphical solution for formants of /IY/ vowel for a 2-tube approximation.

A plot showing the graphical solution for the /IY/ vowel is given in Figure P5.8.1.

- 5.9 (a)** The resonances at 737 and 7736 Hz indicate that the resonances repeat every 7000 Hz. We can then determine the unmarked resonances as sums and differences of the existing resonances, namely:

1. first unmarked resonance is $7000 - 5704 = 1296$ Hz
2. second unmarked resonance is $7000 - 2591 = 4409$ Hz
3. third unmarked resonance is $7000 - 736 = 6264$ Hz
4. fourth unmarked resonance is $7000 + 1296 = 8296$ Hz
5. fifth unmarked resonance is $7000 + 2591 = 9591$ Hz
6. sixth unmarked resonance is $7000 + 3500 = 10500$ Hz

- (b)** As shown in the previous problem, the uniform tube has an equivalent of 7 uniform tube sections, each of equal length of $17.5/7 = 2.5$ cm. Since the problem states that the vocal tract is, in fact, represented by 2 uniform tubes, we can infer that the 7 uniform sections are really 2 groups, one group with from $N_1 = 1 - 6$ identical uniform tubes (with the other group having a complementary number, $N_2 = 7 - N_1$ of uniform tubes). Thus the possible section groupings can be any of the following:

- 1 section for the first tube, 6 sections for the second tube
- 2 sections for the first tube, 5 sections for the second tube
- 3 sections for the first tube, 4 sections for the second tube
- 4 sections for the first tube, 3 sections for the second tube
- 5 sections for the first tube, 2 sections for the second tube
- 6 sections for the first tube, 1 section for the second tube

Thus, the minimum length difference occurs when one tube has 3 or 4 sections, and the other tube has 4 or 3 sections, and the difference in length between the two tubes is therefore 2.5 cm.

- (c)** Since all the resonant frequencies are known, we can use the equation for the 2-tube transfer function to obtain a value of $\mu = -5/7$.

5.10 We begin with the result:

$$\frac{1}{V(z)} = z^{(N/2)} \left[\frac{2}{1+r_G}, \quad \frac{-2r_G}{1+r_G} \right] \prod_{k=1}^N \hat{Q}_k \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ Eq. (5.80)}$$

where:

$$\hat{Q}_k = \begin{bmatrix} \frac{1}{1+r_k} & \frac{-r_k}{1+r_k} \\ \frac{-r_k z^{-1}}{1+r_k} & \frac{z^{-1}}{1+r_k} \end{bmatrix}$$

For $N = 2$ we have:

$$\prod_{k=1}^2 \hat{Q}_k = \begin{bmatrix} \frac{1}{1+r_1} & \frac{-r_1}{1+r_1} \\ \frac{-r_1 z^{-1}}{1+r_1} & \frac{z^{-1}}{1+r_1} \end{bmatrix} \begin{bmatrix} \frac{1}{1+r_2} & \frac{-r_2}{1+r_2} \\ \frac{-r_2 z^{-1}}{1+r_2} & \frac{z^{-1}}{1+r_2} \end{bmatrix}$$

We only need the first column since we subsequently post-multiply by the column vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, giving:

$$\prod_{k=1}^2 \hat{Q}_k = \begin{bmatrix} \frac{1+r_1 r_2 z^{-1}}{(1+r_1)(1+r_2)} & (\dots) \\ \frac{-r_1 z^{-1} - r_2 z^{-2}}{(1+r_1)(1+r_2)} & (\dots) \end{bmatrix}$$

Substituting into Eq. (5.80) gives:

$$\begin{aligned} \frac{1}{V(z)} &= z \left[\frac{2}{1+r_G} \quad \frac{-2r_G}{1+r_G} \right] \begin{bmatrix} \frac{1+r_1 r_2 z^{-1}}{(1+r_1)(1+r_2)} \\ \frac{-r_1 z^{-1} - r_2 z^{-2}}{(1+r_1)(1+r_2)} \end{bmatrix} \\ \frac{1}{V(z)} &= \frac{2z[1+r_1 r_2 z^{-1} + r_1 r_G z^{-1} + r_2 r_G z^{-2}]}{(1+r_1)(1+r_2)(1+r_G)} \\ \frac{1}{V(z)} &= \frac{0.5z^{-1}[(1+r_G)(1+r_1)(1+r_2)]}{1 + (r_1 r_2 + r_1 r_G)z^{-1} + r_2 r_G z^{-2}} \end{aligned}$$

5.11 Using the readily derived relationship:

$$\sum_{n=0}^{\infty} (az^{-1})^{-n} = \frac{1}{1-az^{-1}}, \quad |az^{-1}| < 1; \quad |z| > |a|$$

Cross-multiplying terms gives:

$$1 - az^{-1} = \frac{1}{\sum_{n=0}^{\infty} a^n z^{-n}}$$

showing that a simple zero at $z = a$ can be represented by an infinite number of poles. If we truncate the sum to a finite number of poles, the approximation will become less accurate, especially depending on how close a is to the unit circle.

- 5.12 (a)** A plot of the locations of the poles of $V_k(z)$ in the z -plane as well as the corresponding analog poles in the s -plane is given in Figure P5.12.1 and P5.12.2. It can readily be shown

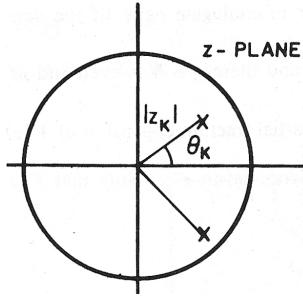


Figure P5.12.1: Location of resonator poles in the z -plane.

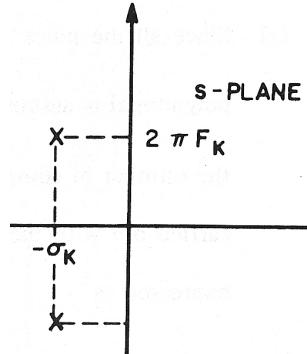


Figure P5.12.2: Location of resonator poles in the s -plane.

that the angles and magnitudes of the poles in the z - and s -planes satisfy the relations:

$$\theta_k = 2\pi F_k T, \quad |z_k| = e^{-\sigma_k T} \text{ and } \Omega_k = \theta_k / T, \quad \sigma_k = -\frac{1}{T} \ln |z_k|$$

- (b)** The difference equation relating the output, $y_k[n]$, of $V_k(z)$ to its input, $x_k[n]$ is:

$$y_k[n] = 2|z_k| \cos(\theta_k) y_k[n-1] - |z_k|^2 y_k[n-2] + (1 - 2|z_k| \cos(\theta_k) + |z_k|^2) x_k[n]$$

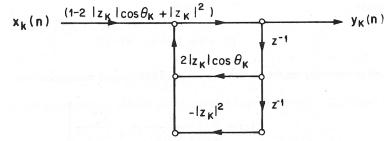


Figure P5.12.3: Digital network implementation of formant network.

- (c)** A digital network implementation of the digital formant network with three multipliers is shown in Figure P5.12.3.

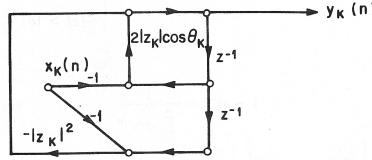


Figure P5.12.4: Rearranged digital network implementation of formant network.

(d) We can rewrite the difference equation as:

$$y_k[n] = 2|z_k| \cos(\theta_k)(y_k[n-1] - x_k[n]) - |z_k|^2(y_k[n-2] - x_k[n])$$

and the resulting network implementation is shown in Figure P5.12.4.

5.13 (a) Since all the poles are complex, they must occur in conjugate pairs (since the denominator polynomial is assumed to have real coefficients), and therefore N is even and $M = N/2$ is the number of complex conjugate pole pairs. A partial fraction expansion of $V(z)$ can be carried out with the denominator viewed as a polynomial in z^{-1} . Note that $V(z)$ can be expressed as:

$$V(z) = \frac{G}{\prod_{k=1}^M (1 - z_k z^{-1}) \sum_{k=1}^M (1 - z_k^* z^{-1})}$$

and the partial fraction expansion is therefore of the form:

$$V(z) = \sum_{k=1}^M \frac{G_k}{1 - z_k z^{-1}} + \sum_{k=1}^M \frac{F_k}{1 - z_k^* z^{-1}}$$

where:

$$G_k = \lim_{z \rightarrow z_k} [(1 - z_k z^{-1}) V(z)]$$

and where it can easily be shown that $F_k = G_k^*$. Thus we get:

$$V(z) = \sum_{k=1}^M \left[\frac{G_k}{1 - z_k z^{-1}} + \frac{G_k^*}{1 - z_k^* z^{-1}} \right]$$

(b) We can combine terms above giving:

$$\begin{aligned} \sum_{k=1}^M \left[\frac{G_k}{1 - z_k z^{-1}} + \frac{G_k^*}{1 - z_k^* z^{-1}} \right] &= \frac{G_k - z_k^* G_k z^{-1} + G_k^* - z_k G_k^* z^{-1}}{1 - (z_k + z_k^*) z^{-1} + z_k z_k^* z^{-2}} \\ &= \frac{2\Re(G_k) - 2\Re(z_k^* G_k) z^{-1}}{1 - 2\Re(z_k) z^{-1} + |z_k|^2 z^{-2}} \end{aligned}$$

where $\Re(z_k) = |z_k| \cos(\theta_k)$. For $z_k = |z_k| e^{j\theta_k}$ we get:

$$B_k = 2\Re(G_k)$$

$$C_k = 2\Re(z_k^* G_k)$$

$$V(z) = \sum_{k=1}^M \frac{B_k - C_k z^{-1}}{1 - 2|z_k| \cos(\theta_k) z^{-1} + |z_k|^2 z^{-2}}$$

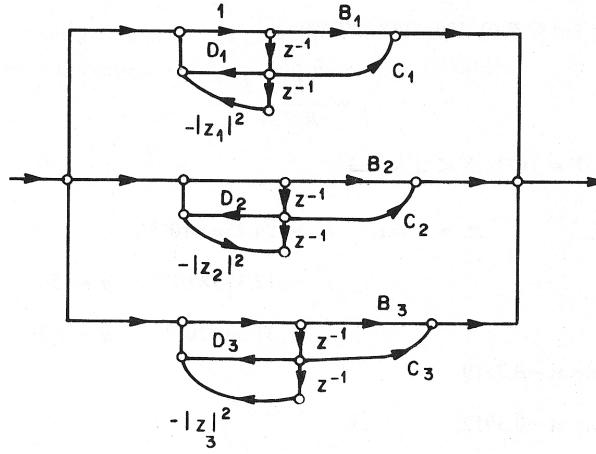


Figure P5.13.1: Digital network implementation of parallel form implementation.

- (c) For $M = 3$, and letting $D_k = 2|z_k| \cos(\theta_k)$ we have the digital network diagram for the parallel form implementation as shown in figure P5.13.1.
- (d) The cascade connection requires 2 multiplies per second-order section as compared with 4 multiplies per second-order section in the parallel implementation. Therefore the parallel form requires the most multiplications.

5.14 (a) We use the bilinear transform to give:

$$\begin{aligned}
 R(z) &= Z_L \left[\frac{2}{T} \frac{1-z^{-1}}{1+z^{-1}} \right] = \frac{P_L(z)}{U_L(z)} \\
 &= \frac{\frac{2}{T} \left[\frac{1-z^{-1}}{1+z^{-1}} \right] R_r L_r}{R_r + \frac{2}{T} \left[\frac{1-z^{-1}}{1+z^{-1}} \right] L_r} = \frac{\frac{2}{T} R_r L_r (1-z^{-1})}{R_r (1+z^{-1}) + \frac{2}{T} L_r (1-z^{-1})} \\
 &= \frac{2 R_r L_r (1-z^{-1})}{(R_r T + 2L_r) - (2L_r - R_r T) z^{-1}}
 \end{aligned}$$

(b) The difference equation is of the form:

$$p_L[n] = \left[\frac{2L_r - R_r T}{2L_r + R_r T} \right] p_L[n-1] + \frac{2R_r L_r}{2L_r + R_r T} (u_L[n] - u_L[n-1])$$

(c) There is a zero at $z = 1$, i.e., zero frequency, and a pole at:

$$z = \frac{2L_r - R_r T}{R_r T + 2L_r}$$

(d) The range of pole values is determined as:

$$R_r = 1.441, \quad L_r = 24.25a \times 10^{-6}$$

For $a = 0.5$ we get $L_r = 12.125 \times 10^{-6}$ and for $a = 1.3$ we get $L_r = 31.53 \times 10^{-6}$. When $a = 0.5$ there is a pole at -0.7119 , and when $a = 1.3$, there is a pole at -0.3912

- (e) Notice that in both cases, the pole is pretty far inside the unit circle. Therefore, as a first approximation we have $\hat{R}(z) = R_0(1 - z^{-1})$, and

$$R(-1) = \frac{2R_rL_r(2)}{R_rT + 2L_r + 2L_r - R_rT} = R_r$$

$$\hat{R}(-1) = 2R_0 \Rightarrow R_0 = R_r/2 = 0.7205$$

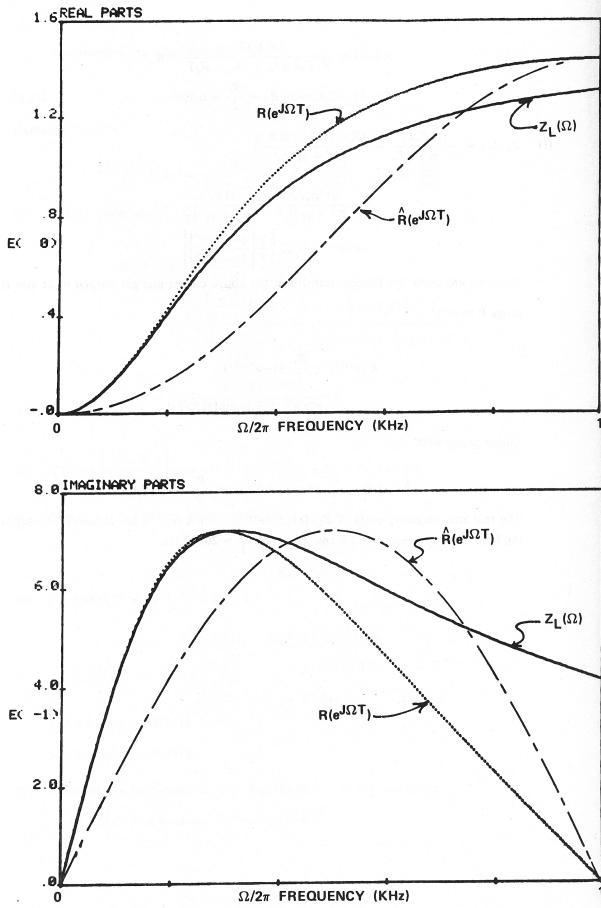


Figure P5.14.1: Real and imaginary parts of $Z_L(\Omega)$.

- (f) We have:

$$\begin{aligned} Z_L(\Omega) &= \frac{j\Omega R_r L_r}{R_r + j\Omega L_r} = \frac{(R_r - j\Omega L_r)j\Omega R_r L_r}{R_r^2 + \Omega^2 L_r^2} \\ &= \frac{\Omega^2 R_r L_r^2}{R_r^2 + \Omega^2 L_r^2} + \frac{j\Omega R_r^2 L_r}{R_r^2 + \Omega^2 L_r^2} \\ R(e^{j\Omega T}) &= Z_L \left[\frac{2}{T} \left[\frac{1 - e^{-j\Omega T}}{1 + e^{-j\Omega T}} \right] \right] \end{aligned}$$

Since we are using the bilinear transform, the above curves just get warped to fit into the range 0 to π/T .

$$\hat{R}(e^{j\Omega T}) = \frac{R_r}{2} (1 - e^{-j\Omega T})$$

$$= e^{-j\Omega T/2} R_r [j \sin(\Omega T/2)]$$

Linear phase +90 degrees gives:

$$\hat{R}(e^{j\Omega T}) = \frac{R_r}{2}(1 - \cos(\Omega T)) + j \frac{R_r}{2} \sin(\Omega T)$$

The real and imaginary parts of $Z_L(\Omega)$, $R(e^{j\Omega T})$, and $\hat{R}(e^{j\Omega T})$ are plotted in Figure P5.14.1 for the case $a = 3.0$ cm, and $F_s = 1/T = 20,000$ Hz.

5.15 (a) We can solve for $G_1(z)$ as:

$$g_1[n+1] = p[n] * p[n] = \sum_{m=-\infty}^{\infty} p[m]p[n-m]$$

$$G_1(z) = z^{-1}P(z)P(z) = P^2(z)z^{-1}$$

where

$$P(z) = \sum_{n=-\infty}^{\infty} p[n]z^{-n} = \sum_{n=0}^{N-1} z^{-n} = \frac{1 - z^{-N}}{1 - z^{-1}}$$

$$G_1(z) = P^2(z)z^{-1} = \left[\frac{1 - z^{-N}}{1 - z^{-1}} \right]^2 z^{-1}$$

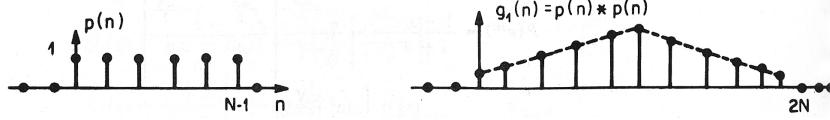


Figure P5.15.1: Waveform plots of $p[n]$ and $g_1[n]$.

where the sequences $p[n]$ and $g_1[n]$ are shown plotted in Figure P5.15.1.

(b) For $N = 10$ we have:

$$P(z) = \left[\frac{1 - z^{-10}}{1 - z^{-1}} \right] = \frac{1}{z^9} \left[\frac{z^{10} - 1}{z - 1} \right]$$

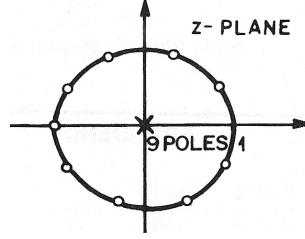


Figure P5.15.2: Plot of locations of poles and zeros of $P(z)$ in the z -plane.

The zeros of $P(z)$ are found as the roots of:

$$z^{10} - 1 = 0 \Rightarrow z = (1)^{0.1} = (e^{j2\pi k})^{0.1} = e^{j2\pi k/(10)}, \quad k = 0, 1, \dots, 9$$

Thus the numerator zeros are located on the unit circle starting at $z = 1$ and spaced at an angular distance of $\pi/5$ radians. There are 10 poles, one at $z = -1$, and 9 at the origin. Note that the pole and zero at $z = 1$ cancel. The poles and zeros of $G_1(z)$ are those of $P(z)$; however at each pole and zero location there is a double (second-order) pole or a double zero. A plot of the locations of the poles and zeros is given in Figure P5.15.2.

(c) We can solve for $P(e^{j\omega})$ as:

$$P(e^{j\omega}) = \frac{1 - e^{-j\omega N}}{1 - e^{-j\omega}} = \frac{e^{-j\omega N/2}}{e^{-j\omega/2}} \left[\frac{e^{j\omega N/2} - e^{-j\omega N/2}}{e^{j\omega/2} - e^{-j\omega/2}} \right]$$

$$P(e^{j\omega}) = e^{-j\omega(N-1)/2} \left[\frac{\sin(\omega N/2)}{\sin(\omega/2)} \right]$$

Therefore we get:

$$|G_1(e^{j\omega})| = |P(e^{j\omega})|^2 = \left[\frac{\sin(\omega N/2)}{\sin(\omega/2)} \right]^2$$

which is shown plotted in Figure P5.15.3.

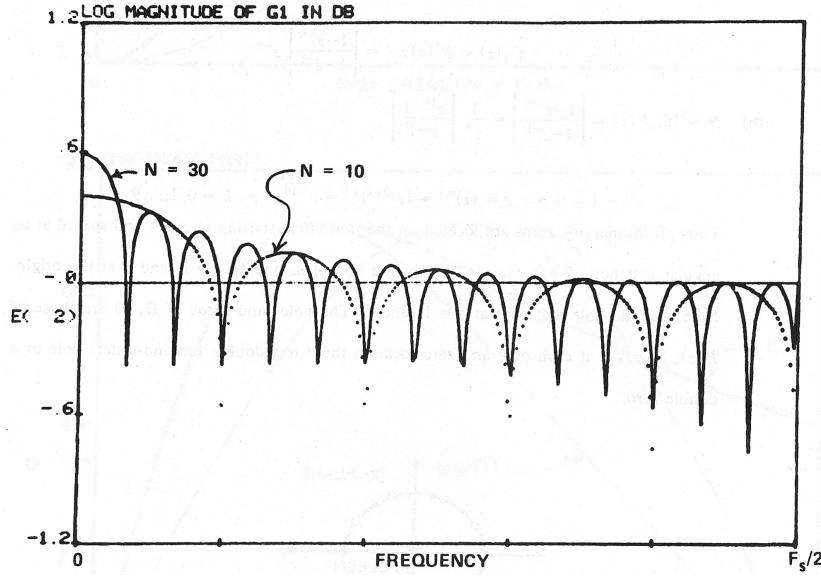


Figure P5.15.3: Plot of log magnitude of G_1 .

(d) With $g_2[n] = np[n]$ where:

$$p[n] = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases}$$

we get:

$$\begin{aligned}
 G_2(z) &= -z \frac{d}{dz} P(z) = -z \left\{ \frac{d}{dz} \left[\frac{1-z^{-N}}{1-z^{-1}} \right] \right\} \\
 &= -z \left\{ \frac{d}{dz} \left[\frac{z^N-1}{z^N-z^{N-1}} \right] \right\} \\
 &= -z \left\{ \frac{Nz^{N-1}[z^N-z^{N-1}] - [Nz^{N-1} - (N-1)z^{N-2}][z^N-1]}{(z^N-z^{N-1})^2} \right\} \\
 &= -z \left\{ \frac{Nz^{2N-1} - Nz^{2N-2} - Nz^{2N-1} + (N-1)z^{2N-2} + Nz^{N-1} - (N-1)z^{N-2}}{z^{2N}(1-z^{-1})^2} \right\} \\
 &= z^{-1}(z^{-2N+2}) \left\{ \frac{-z^{2N-2} + Nz^{N-1} - (N-1)z^{N-2}}{(1-z^{-1})^2} \right\} \\
 G_2(z) &= z^{-1} \left\{ \frac{1-Nz^{-(N-1)} + (N-1)z^{-N}}{(1-z^{-1})^2} \right\}
 \end{aligned}$$

(e) $G_2(z)$ can be rewritten in terms of powers of z as:

$$G_2(z) = z^{-(N-1)} \left[\frac{z^N - Nz + (N-1)}{(z-1)^2} \right]$$

The coefficient, $(N-1)$, of z^0 is the product of all the roots of the numerator polynomial. Thus at least one root of the numerator must have magnitude > 1 . Also note that since $g_2[n]$ has finite length, $G_2(z)$ has no poles except at $z = 0$ or $z = \infty$. Therefore the two denominator factors must cancel zeros of the numerator. For $N = 4$ we have:

$$G_2(z) = z^{-3} \frac{(z^4 - 4z + 3)}{(z-1)^2} = z^{-3}(z^2 + 2z + 3)$$

$$G_2(z) = z^{-3}(z + 1 + j\sqrt{2})(z + 1 - j\sqrt{2})$$

In this case both zeros are outside the unit circle.

5.16 (a) The z -transform of $g[n]$ can be determined using the relationship:

$$nx[n] \longleftrightarrow -z \frac{dX(z)}{dz}$$

thus giving:

$$na^n u[n] \longleftrightarrow -z \frac{d}{dz} \left\{ \frac{1}{1-az^{-1}} \right\}, \quad |z| > |a|$$

Therefore:

$$G(z) = \frac{az^{-1}}{(1-z^{-1})^2}$$

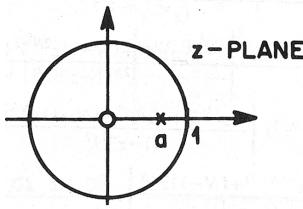
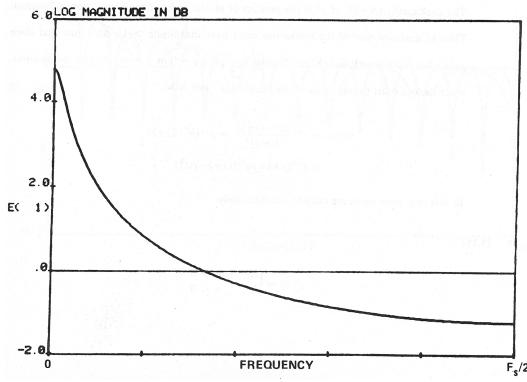
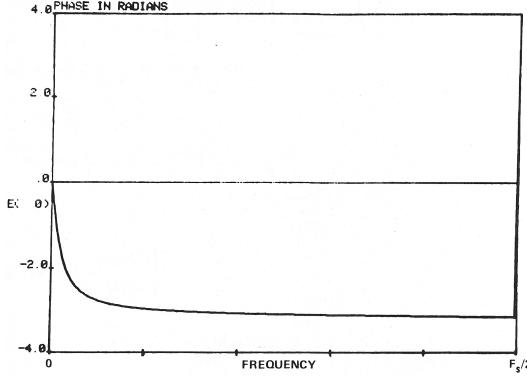
(b) The Fourier transform is:

$$G(e^{j\omega}) = \frac{ae^{-j\omega}}{(1-ae^{-j\omega})^2}$$

having a double pole at $z = a$ and a zero at $z = 0$ as shown in Figure P5.16.1.

The frequency response is obtained by evaluating $G(z)$ on the unit circle, i.e., $z = e^{j\omega}$. The effect of the poles at $z = a$ is to peak the frequency response. The phase response is of the form:

$$\arg[G(e^{j\omega})] = \arg\{-ae^{-j\omega}\} - 2\arg\{(1-ae^{-j\omega})\}$$

Figure P5.16.1: Plot pole-zero locations in z -plane.Figure P5.16.2: Plot of log magnitude of G .Figure P5.16.3: Plot of phase of G .

$$= -\omega + \pi - 2 \tan^{-1} \left\{ \frac{a \sin(\omega)}{1 - a \cos(\omega)} \right\}$$

The argument starts at $\pm\pi$ for $\omega = 0$, and has a negative linear phase component plus the $2 \tan^{-1}$ contribution.

If we assume a is positive and real, we get a log magnitude response as shown in Figure P5.16.2 and a phase response as shown in Figure P5.16.3.

(c) We have the results:

$$G(e^{j0}) = \frac{a}{(1-a)^2} \implies |G(e^{j0})| = \frac{a}{(1-a)^2} \quad a > 0$$

$$G(e^{j\pi}) = \frac{a}{(1+a)^2} = |G(e^{j\pi})|$$

$$20 \log_{10} |G(e^{j0})| - 20 \log_{10} |G(e^{j\pi})| = 60 \text{ dB}$$

$$\frac{|G(e^{j0})|}{|G(e^{j\pi})|} = 1000 = \frac{a/(1-a)^2}{a/(1+a)^2} = \left[\frac{1+a}{1-a} \right]^2$$

$$\frac{1+a}{1-a} = \sqrt{1000} \Rightarrow a = \frac{\sqrt{1000}-1}{\sqrt{1000}+1} = 0.9387$$

- 5.17 (a)** We can solve for the glottal pulse impulse response using simple z -transform properties. Thus we can write:

$$\begin{aligned} G_1(z) &= \frac{1}{(1-az^{-1})} \longleftrightarrow g_1[n] = a^n u[n] \\ G(z) &= -z^{-1} \frac{dG_1(z)}{dz} = \frac{az^{-1}}{(1-az^{-1})^2} \longleftrightarrow g[n] = na^n u[n] \end{aligned}$$

Plots of the glottal pulse model impulse response are given in Figure P5.17.1, at the top, for values of $a = 0.95$ and $a = 0.8$.

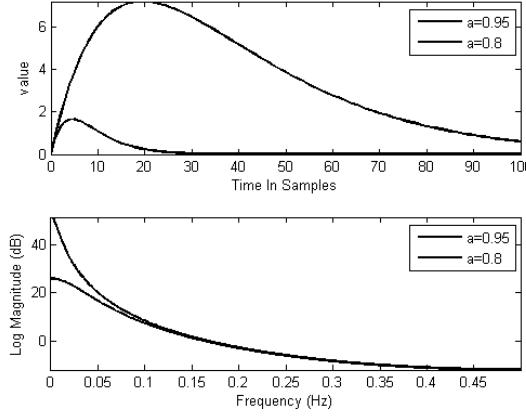


Figure P5.17.1: Plots of impulse and log magnitude frequency responses for glottal pulse model for values of $a = 0.95$ and $a = 0.8$.

- (b) Plots of the log magnitude responses of the glottal pulse model are given in Figure P5.17.1, at the bottom, for values of $a = 0.95$ and $a = 0.8$.
- (c) The transfer function for lip radiation is:

$$H(z) = 1 - z^{-1} \Rightarrow H(e^{j\omega}) = 1 - e^{-j\omega}$$

The results of including the single zero model for the radiation load are shown in Figure P5.17.2 which shows the radiation load log magnitude response in the upper panel, and the combined glottal pulse, radiation model in the lower panel for values of $a = 0.95$ and $a = 0.8$.

- (d) A flow graph representation of the system that models the combined glottal pulse and lip radiation is shown in Figure P5.17.3.

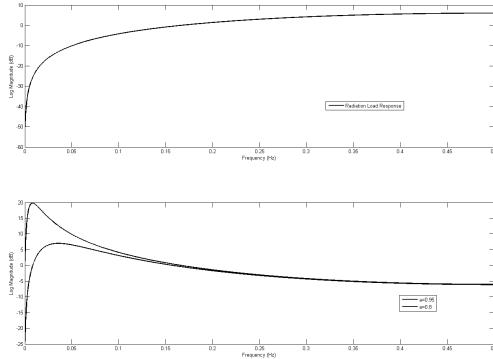


Figure P5.17.2: Plots of the log magnitude frequency responses for the radiation model (top panel) and for the two glottal pulse models with radiation for values of $a = 0.95$ and $a = 0.8$.

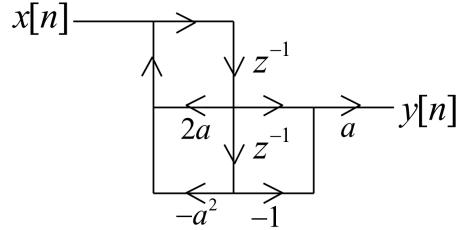


Figure P5.17.3: Flow graph representation of the system that models the combined glottal pulse and lip radiation systems.

5.18 (a) We can show the desired result via the recursion, i.e.,

$$\begin{aligned} D_0(z) &= 1 \\ D_1(z) &= D_0(z) + r_1 z^{-1} D_0(z^{-1}) = 1 + r_1 z^{-1} \\ D_2(z) &= D_1(z) + r_2 z^{-2} D_1(z^{-1}) = 1 + r_1(1 + r_2) z^{-1} + r_2 z^{-2} \\ D_3(z) &= D_2(z) + r_3 z^{-3} D_2(z^{-1}) = 1 + \dots + r_3 z^{-3} \end{aligned}$$

Thus we see that, in general, the coefficient in $D_k(z)$ of the z^{-k} term is r_k ; therefore r_N is the coefficient of the z^{-N} term in $D_N(z)$ and $r_N = -\alpha_N$.

(b) Using the recursion we get:

$$D_k(z) = D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}) \quad (5.1)$$

If we replace z by z^{-1} in Eq. (5.1) we get

$$D_k(z^{-1}) = D_{k-1}(z^{-1}) + r_k z^k D_{k-1}(z) \quad (5.2)$$

If we next multiply both sides of Eq. (5.2) by $r_k z^{-k}$ we get

$$r_k z^{-k} D_k(z^{-1}) = r_k^2 D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}) \quad (5.3)$$

Now if we subtract Eq. (5.3) from Eq. (5.1) we get

$$\begin{aligned} D_k(z) - r_k z^{-k} D_k(z^{-1}) &= (1 - r_k^2) D_{k-1}(z) \\ D_{k-1}(z) &= \frac{D_k(z) - r_k z^{-k} D_k(z^{-1})}{(1 - r_k^2)} \end{aligned} \quad (5.4)$$

which is the desired result.

- (c) r_{k-1} is the coefficient of the term in the power $z^{-(k-1)}$ of the polynomial $D_{k-1}(z)$
- (d) An algorithm for finding all the reflection coefficients and all the tube areas is as follows:
 1. From the polynomial for $D(z) = D_N(z)$ we get r_N .
 2. We then use Eq. (5.4) to iterate backwards to get r_{k-1} , $k = N - 1, N - 2, \dots, 1$.
 3. Since $r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k} \Rightarrow \frac{A_{k+1}}{A_k} = \frac{1 + r_k}{1 - r_k}$, we can find the sequence of ratios $\frac{A_{k+1}}{A_k}$. We can scale the A_k 's by any constant and still get the same set of area ratios; thus the same $D(z)$. Hence the areas are unique up to within a constant multiplier.

Chapter 6

Time-Domain Methods for Speech Processing

6.1 (a) The Fourier transform of the rectangular window is:

$$\begin{aligned} W_R(e^{j\omega}) &= \sum_{n=0}^{L-1} w_R[n] e^{-j\omega n} = \sum_{n=0}^{L-1} e^{-j\omega n} \\ W_R(e^{j\omega}) &= \frac{1 - e^{-j\omega L}}{1 - e^{-j\omega}} = \frac{e^{-j\omega L/2}}{e^{-j\omega/2}} \left[\frac{e^{j\omega L/2} - e^{-j\omega L/2}}{e^{j\omega/2} - e^{-j\omega/2}} \right] \\ W_R(e^{j\omega}) &= e^{-j\omega(L-1)/2} \left[\frac{\sin[\omega L/2]}{\sin[\omega/2]} \right] \end{aligned}$$

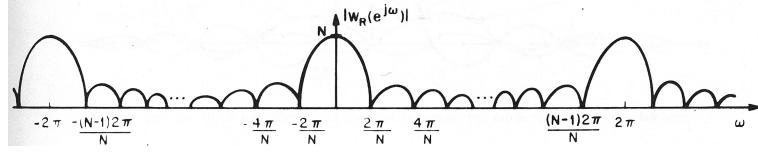


Figure P6.1.1: Plot of the log magnitude spectrum of a rectangular window.

(b) A sketch of the magnitude of $W_R(e^{j\omega})$ is shown in Figure P6.1.1.

(c) We can express the Hamming window as:

$$w_H[n] = 0.54w_R[n] - 0.46 \cos \left[\frac{2\pi n}{L-1} \right] w_R[n]$$

$$w_H[n] = 0.54w_R[n] - 0.23[e^{j2\pi n/(L-1)} + e^{-j2\pi n/(L-1)}]w_R[n]$$

Using the relationship $e^{j\theta n}x[n] \longleftrightarrow X(e^{j(\omega-\theta)})$, then

$$W_H(e^{j\omega}) = 0.54W_R(e^{j\omega}) - 0.23[W_R(e^{j(\omega-2\pi/(L-1))}) + W_R(e^{j(\omega+2\pi/(L-1))})]$$

(d) Now substitute $W_R(e^{j\omega}) = e^{-j\omega(L-1)/2} \left[\frac{\sin(\omega L/2)}{\sin(\omega/2)} \right]$, giving

$$W_H(e^{j\omega}) = 0.54e^{-j\omega(L-1)/2} \left[\frac{\sin(\omega L/2)}{\sin(\omega/2)} \right] - 0.23e^{-j(\omega-2\pi/(L-1))(L-1)/2} \frac{\sin[(\omega-2\pi/(L-1))(L/2)]}{\sin[(\omega-2\pi/(L-1))(1/2)]}$$

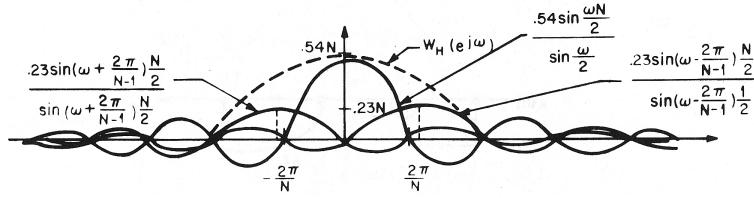


Figure P6.1.2: Plot of the three terms in the frequency response of a Hamming window.

$$-0.23e^{-j(\omega+2\pi/(L-1))(L-1)/2} \frac{\sin[(\omega+2\pi/(L-1))(L/2)]}{\sin[(\omega+2\pi/(L-1))(1/2)]}$$

Neglecting the delay term, $e^{-j\omega(L-1)/2}$, the three terms are superimposed on the plot of Figure P6.1.2. Note that the three waveforms add together in a way that tends to reduce the sidelobes, and widen the main lobe by a factor of approximately two.

6.2 (a) We are given the relation:

$$E_{\hat{n}} = \sum_{m=-\infty}^{\infty} (x[m]w[\hat{n}-m])^2$$

If we make the simple substitutions:

$$n = \hat{n}, \quad y[n] = E_n, \quad v[n] = x^2[n], \quad \tilde{w}[n] = w^2[n] = (a^2)^n u[n]$$

we get the following simple convolution:

$$y[n] = v[n] * \tilde{w}[n]$$

Transforming into the z -plane we get:

$$Y(z) = V(z) \cdot \tilde{W}(z)$$

Recognizing that we can express $\tilde{W}(z)$ as

$$\tilde{W}(z) = \frac{1}{1 - a^2 z^{-1}}$$

we get:

$$Y(z) = V(z) \cdot \frac{1}{1 - a^2 z^{-1}}$$

We can now convert to a difference equation, by cross multiplying factors, giving

$$y[n] - a^2 y[n-1] = v[n]$$

Undoing the substitutions, we can now express the difference equation as:

$$E_{\hat{n}} = a^2 E_{\hat{n}-1} + x^2[\hat{n}]$$

(b) A plot of a digital network that realizes the difference equation is given in Figure P6.2.1.



Figure P6.2.1: Plot of a network implementation of the difference equation.

- (c) In general, we require that the z -transform of the equivalent filter response, $\tilde{W}(z)$ be a rational function of z , i.e., of the form

$$\tilde{W}(z) = \frac{\sum_{r=0}^{N_z} b_r z^{-1}}{1 - \sum_{k=1}^{N_p} a_k z^{-1}}$$

with finite values of N_p and N_z .

- 6.3 (a)** We can express $E_{\hat{n}}$ in a difference equation by expressing $\tilde{w}[m]$ as a sum of two components, namely:

$$\begin{aligned}\tilde{w}_+[m] &= \begin{cases} \alpha^m & 0 \leq m \leq L \\ 0 & \text{otherwise} \end{cases} \\ \tilde{w}_-[m] &= \begin{cases} \alpha^{-m} & -L \leq m \leq -1 \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

We can now express the z -transform of $\tilde{w}[m]$ as

$$\begin{aligned}\tilde{W}_+(z) &= 1 + \alpha z^{-1} + \dots + \alpha^L z^{-L} = \frac{1 - \alpha^{L+1} z^{-(L+1)}}{1 - \alpha z^{-1}} \\ \tilde{W}_-(z) &= \alpha z + (\alpha z)^2 + \dots + (\alpha z)^L = \frac{\alpha z(1 - \alpha^L z^L)}{1 - \alpha z} + \frac{1 - \alpha^{-L} z^{-L}}{1 - \alpha^{-1} z^{-1}} \cdot z^L \alpha^L\end{aligned}$$

Now we can express $E_{\hat{n}}$ as the sum of two components, namely

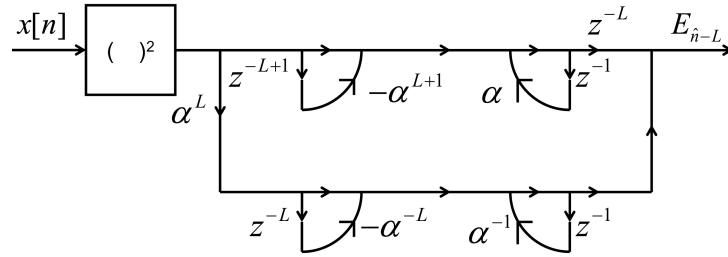
$$\begin{aligned}E_{\hat{n}}^+ &= \alpha E_{\hat{n}-1}^+ + x^2[n] - \alpha^{L+1} x^2[n-L-1] \\ E_{\hat{n}+L}^- &= \alpha^L (\alpha^{-1} E_{\hat{n}-1}^- + x^2[n] - \alpha^{-L} x^2[n-L])\end{aligned}$$

Finally we get for $\tilde{W}(z)$ the result

$$\begin{aligned}\tilde{W}(z) &= \tilde{W}_+(z) + \tilde{W}_-(z) \\ z^{-L} \tilde{W}(z) &= z^{-L} \frac{1 - \alpha^{L+1} z^{-(L+1)}}{1 - \alpha z^{-1}} + \alpha^L \frac{1 - \alpha^{-L} z^{-L}}{1 - \alpha^{-1} z^{-1}}\end{aligned}$$

- (b) In terms of computational complexity, the direct calculation requires $2L + 1$ multiply-adds, while the “recursive” network requires only 5 multiply-adds.

- (c) A digital network implementation of the recurrence relation is shown in Figure P6.3.1.

Figure P6.3.1: Digital network of recurrence formula for E_n .

6.4 (a) Using a finite length window of duration L samples, the computation for $M_{\hat{n}}$ is:

$$M_{\hat{n}} = \sum_{m=\hat{n}-L+1}^{\hat{n}} |x[m]w[\hat{n}-m]|$$

Using a recursive window we get:

$$M_{\hat{n}} = a \cdot M_{\hat{n}-1} + |x[\hat{n}]|$$

Using the finite length window, with a shift of L samples between windows, we require L multiplications and $L-1$ additions every L samples of the input.

- (b) For the recursive window we require L multiplications and L additions to obtain L samples of $M_{\hat{n}}$.
- (c) The finite length window is less efficient if we overlap adjacent frames since we are retrieving only a fraction of L samples for each overlapped frame. For no frame overlap, both methods are essentially identical computationally.

6.5 We can express $Z_{\hat{n}}$ and $Z_{\hat{n}+1}$ as:

$$\begin{aligned} Z_{\hat{n}} &= \frac{1}{2L} \sum_{m=\hat{n}-L+1}^{\hat{n}} |\operatorname{sgn}(x[m]) - \operatorname{sgn}(x[m-1])| \\ Z_{\hat{n}+1} &= \frac{1}{2L} \sum_{m=\hat{n}+1-L+1}^{\hat{n}+1} |\operatorname{sgn}(x[m]) - \operatorname{sgn}(x[m-1])| \\ Z_{\hat{n}+1} &= \frac{1}{2L} \left[\sum_{m=\hat{n}-L+1}^{\hat{n}+1} |\operatorname{sgn}(x[m]) - \operatorname{sgn}(x[m-1])| + |\operatorname{sgn}(x[\hat{n}+1]) - \operatorname{sgn}(x[\hat{n}])| \right. \\ &\quad \left. - |\operatorname{sgn}(x[\hat{n}+1-L]) - \operatorname{sgn}(x[\hat{n}-L])| \right] \end{aligned}$$

We recognize the first term as $Z_{\hat{n}}$ giving

$$Z_{\hat{n}+1} = Z_{\hat{n}} + \frac{1}{2L} [|\operatorname{sgn}(x[\hat{n}+1]) - \operatorname{sgn}(x[\hat{n}])| - |\operatorname{sgn}(x[\hat{n}+1-L]) - \operatorname{sgn}(x[\hat{n}-L])|]$$

If we let $\hat{n}+1 \rightarrow \hat{n}$ in the above expression, we obtain the desired result

$$Z_{\hat{n}} = Z_{\hat{n}-1} + \frac{1}{2L} [|\operatorname{sgn}(x[\hat{n}]) - \operatorname{sgn}(x[\hat{n}-1])| - |\operatorname{sgn}(x[\hat{n}-L]) - \operatorname{sgn}(x[\hat{n}-L-1])|]$$

6.6 (a) From the definition of the short-time autocorrelation function we have:

$$R_{\hat{n}}[-k] = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]x[m-k]w[\hat{n}+k-m]$$

If we make a change of variables of the form $m = m' + k$ we get:

$$R_{\hat{n}}[-k] = \sum_{m'=-\infty}^{\infty} x[m'+k]w[\hat{n}-m'-k]x[m']w[\hat{n}-m']$$

Rearranging terms and replacing m' by m gives:

$$R_{\hat{n}}[-k] = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]x[m+k]w[\hat{n}-k-m] = R_{\hat{n}}[k]$$

(b) Using the results of part (a) we get:

$$\begin{aligned} R_{\hat{n}}[k] &= R_{\hat{n}}[-k] = \sum_{m=-\infty}^{\infty} x[m]x[m-k]w[\hat{n}-m]w[\hat{n}+k-m] \\ &= \sum_{m=-\infty}^{\infty} x[m]x[m-k]w[\hat{n}-m]w[[\hat{n}-m]+k] \end{aligned}$$

Letting

$$\tilde{w}_k[\hat{n}] = w[\hat{n}]w[\hat{n}+k]$$

we get the result:

$$\tilde{w}_k[\hat{n}-m] = w[\hat{n}-m]w[\hat{n}-m+k]$$

$$R_{\hat{n}}[k] = \sum_{m=-\infty}^{\infty} x[m]x[m-k]\tilde{w}_k[\hat{n}-m]$$

(c) For the given $w[n]$ we get:

$$\begin{aligned} \tilde{w}_k[n] &= w[n]w[n+k] = a^n a^{n+k} \quad n \geq 0; \quad n \geq -k \\ &= \begin{cases} a^{2n+k} & n \geq 0; \quad n \geq -k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

(d)

(d)

$$\tilde{W}_k(z) = \sum_{n=-\infty}^{\infty} \tilde{w}_k[n]z^{-n} = \sum_{n=0}^{\infty} a^{2n+k} z^{-n} = a^k \sum_{n=0}^{\infty} (a^2 z^{-1})^n$$

$$\tilde{W}_k(z) = \frac{a^k}{1 - a^2 z^{-1}}; \quad |a^2 z^{-1}| < 1; \Rightarrow |z| > a^2$$

Then we get:

$$\tilde{w}_k[n] = a^k \delta[n] + a^2 \tilde{w}_k[n-1]$$

Substituting this expression into the relation for $R_n[k]$ gives:

$$\begin{aligned} R_n[k] &= \sum_{m=-\infty}^{\infty} x[m]x[m-k](a^k\delta[n-m] + a^2\tilde{w}_k[n-1-m]) \\ &= \sum_{m=-\infty}^{\infty} x[m]x[m-k]a^k\delta[n-m] + \sum_{m=-\infty}^{\infty} x[m]x[m-k]a^2\tilde{w}_k[n-1-m] \\ &= a^kx[n]x[n-k] + a^2R_{n-1}[k] \end{aligned}$$

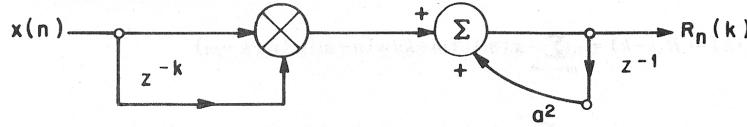


Figure P6.6.1: Digital network of recurrence formula for $R_n[k]$.

A plot of a network implementation of $R_n[k]$ is given in Figure P6.6.1.

(e) For this new choice for $w[n]$ we get:

$$\begin{aligned} \tilde{w}_k[n] &= (na^n u[n])([n+k]a^{n+k}u[n+k]) \\ &= a^k(n^2 + kn)a^{2n} \quad n \geq 0; \quad n \geq -k \\ &= a^k[n^2 a^{2n} + k n a^{2n}] \quad n \geq 0; \quad n \geq -k \end{aligned}$$

Using the relation:

$$nx[n] \longleftrightarrow -z \frac{dX(z)}{dz}$$

we get:

$$\begin{aligned} n^2 a^{2n} u[n] &\longleftrightarrow \frac{a^2 z^{-1} (1 + a^2 z^{-1})}{(1 + a^2 z^{-1})^3} \quad |z| > a^2 \\ na^{2n} u[n] &\longleftrightarrow \frac{a^2 z^{-1}}{(1 - a^2 z^{-1})^2} \quad |z| > a^2 \end{aligned}$$

Therefore we get:

$$\begin{aligned} \tilde{W}_k(z) &= a^k \left[\frac{a^2 z^{-1} (1 + a^2 z^{-1})}{(1 - a^2 z^{-1})^3} + \frac{k a^2 z^{-1} (1 - a^2 z^{-1})}{(1 - a^2 z^{-1})^3} \right] \\ &= \frac{a^{k+2} z^{-1} + a^{k+4} z^{-2} + k a^2 z^{-1} - k a^4 z^{-2}}{(1 - a^2 z^{-1})^3} \\ &= \frac{(a^{k+2} + k a^2) z^{-1} + (a^{k+4} - k a^4) z^{-2}}{(1 - a^2 z^{-1})^3} \\ &= \frac{a^{k+2} (k a^{-k} + 1) z^{-1} - a^{k+4} (k a^{-k} - 1) z^{-2}}{(1 - a^2 z^{-1})^3} \end{aligned}$$

The recursive implementation follows from the fact that:

$$\tilde{W}_k(z) = \frac{R_k(z)}{Y(z)}$$

where $y[n] = x[n]x[n-k]$, giving

$$R_k(z)(1 - a^2 z^{-1})^3 = Y(z)(a^{k+2}(k a^{-k} + 1))z^{-1} - Y(z)(a^{k+4}(k a^{-k} - 1))z^{-2}$$

where we note that $(1 - a^2 z^{-1})^3 = (1 - 3a^2 z^{-1} + 3a^4 z^{-2} - a^6 z^{-3})$. Substituting, we obtain:

$$R_{\hat{n}}(z) = Y(z)(a^{k+2}(ka^{-k}+1))z^{-1} - Y(z)(a^{k+4}(ka^{-k}-1))z^{-2} + R_k(z)(3a^2 z^{-1} - 3a^4 z^{-2} + a^6 z^{-3})$$

Inverse transformation yields:

$$\begin{aligned} R_{\hat{n}}[k] &= (a^{k+2}(ka^{-k}+1))x[n-1]x[n-1-k] - (a^{k+4}(ka^{-k}-1))x[n-2]x[n-2-k] \\ &\quad + 3a^2 R_{\hat{n}-1}[k] - 3a^4 R_{\hat{n}-2}[k] + a^6 R_{\hat{n}-3}[k] \end{aligned}$$

The corresponding digital network is shown in Figure P6.6.2.

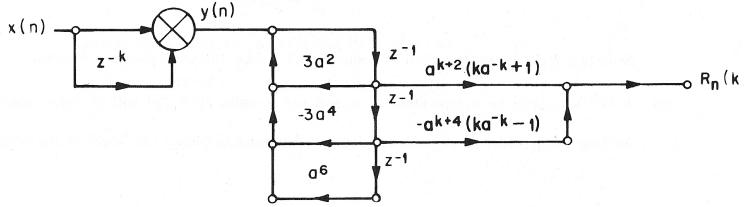


Figure P6.6.2: Digital network of recurrence formula for $R_{\hat{n}}[k]$.

6.7 (a) $R_x(\tau) = 2e^{-\tau^2}$ is a valid autocorrelation function.

(b) $R_x(\tau) = |\tau|e^{-|\tau|}$ is 0 at $\tau = 0$ and thus is smaller than $R_x(\tau)$ for $\tau \neq 0$ and therefore is not a valid autocorrelation function.

(c) $R_x(\tau) = \left(\frac{\sin(\pi\tau)}{\pi\tau}\right)^2$ is a valid autocorrelation function (and in fact actually is the autocorrelation function of an ideal lowpass signal).

(d) $R_x(\tau) = 2\frac{\tau^2 + 4}{\tau^2 + 6}$ cannot be a valid autocorrelation function since at $\tau = 0$ we get $R_x(\tau) = 4/3$, while at $\tau = \infty$ we get $R_x(\tau) = 2$ (which is larger than 4/3) so $R_x(0)$ is not the maximum value of this function so it cannot be a valid autocorrelation function.

(e) $R_x(\tau) = [0.2 \cos(3\pi\tau)]^3$ is a valid autocorrelation function as it obeys all the constraints of a valid autocorrelation function.

6.8 (a) We have:

$$w'[m] = \begin{cases} 1 & 0 \leq m \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$

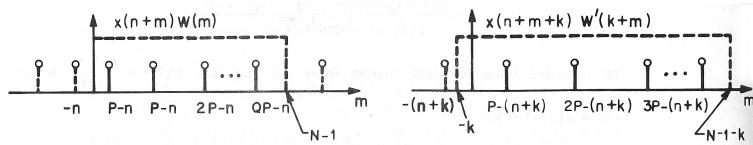
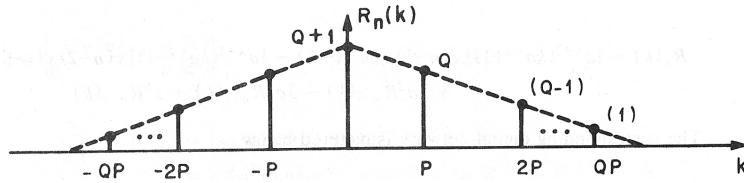
where $QN_p < L-1 < (Q+1)N_p$ for Q some integer. We then can express $R_{\hat{n}}[k]$ as:

$$R_{\hat{n}}[k] = \sum_{m=0}^{L-1-k} (x[\hat{n}+m]w'[m])(x[\hat{n}+m+k]w'[k+m])$$

Plots of the two components of $R_{\hat{n}}[k]$ are shown in Figure P6.8.1.

The two sequences inside the windows will align exactly when $\hat{n}+m = \hat{n}+m+k \implies k=0$. For other values of k , the windows will not overlap completely. The other values of k which produce an alignment of the samples are $k=lN_p$, where l is an integer in the range $-Q \leq l \leq Q$. Then $R_{\hat{n}}[k]$ appears as shown in Figure P6.8.2.

Note that $R_{\hat{n}}[k]$ is independent of \hat{n} since $x[\hat{n}]$ "looks" the same over any interval.

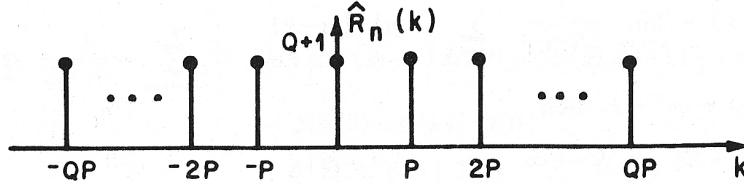
Figure P6.8.1: Plots of the two components of $R_{\hat{n}}[k]$.Figure P6.8.2: Plot of $R_{\hat{n}}[k]$.

- (b) If the window is a Hamming window of the same length, $R_{\hat{n}}[k]$ will again be independent of \hat{n} , and the samples of $R_{\hat{n}}[k]$ will be determined by forming the product of aligned samples and summing to obtain the “area” of the window overlap. Since the area of the Hamming window is smaller, $R_{\hat{n}}[0]$ is smaller. Also $R_{\hat{n}}[k]$ will fall-off more rapidly with k .

(c)

$$\hat{R}_{\hat{n}}[k] = \sum_{m=0}^{L-1} x[\hat{n} + m]x[\hat{n} + m + k]$$

Here we form the product over $L-1$ samples and sum the resulting samples. Each increment of k that shifts a sample of $x[\hat{n} + m + k]$ out of the interval $0 \leq m \leq L-1$ also shifts a new, but identical sample into the interval. As a result, $\hat{R}_{\hat{n}}[k]$ has the appearance as shown in Figure P6.8.3.

Figure P6.8.3: Plot of $\hat{R}_{\hat{n}}[k]$.

- 6.9 (a)** If $x[n] = x[n + N_p]$, $-\infty < n < \infty$, then

(i)

$$\phi[k + N_p] = \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L x[m]x[m + k + N_p]$$

Letting $m + k = n$ we see that $x[m + k + N_p] = x[m + k]$, and therefore

$$\phi[k] = \phi[k + N_p] \quad -\infty < k < \infty$$

(ii)

$$\begin{aligned}
R_{\hat{n}}[k + N_p] &= \sum_{m=0}^{L-|k|-1} x[\hat{n}+m]w'[m]x[\hat{n}+m+k+N_p]w'[m+k+N_p] \\
&= \sum_{m=0}^{L-|k|-1} x[\hat{n}+m]w'[m]x[\hat{n}+m+k]w'[m+k+N_p]
\end{aligned}$$

Therefore

$$R_{\hat{n}}[k + N_p] \neq R_{\hat{n}}[k]$$

unless $w'[n + N_p] = w'[n]$.

(iii)

$\hat{R}_{\hat{n}}[k] = \hat{R}_{\hat{n}}[k + N_p]$ since $x[\hat{n}+m+k+N_p] = x[\hat{n}+m+k]$.

(b) If $x[n] = x[n + N_p]$, $-\infty < n < \infty$, then

(i)

$$\phi[-k] = \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L x[m]x[m-k]$$

Making a change of variables so that $m = m' + k$ gives

$$\phi[-k] = \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m'=-L}^L x[m'+k]x[m'] = \phi[k]$$

Therefore $\phi[-k] = \phi[k]$.

(ii)

$$R_{\hat{n}}[-k] = \sum_{m=-\infty}^{\infty} x[\hat{n}+m]w'[m]x[\hat{n}+m-k]w'[m-k]$$

Making the substitution $m \rightarrow m' + k$ gives

$$R_{\hat{n}}[-k] = \sum_{m'=-\infty}^{\infty} x[\hat{n}+m'+k]w'[m'+k]x[\hat{n}+m']w'[m']$$

For a causal window of length $L - 1$ samples, the product is zero outside the range $0 \leq m \leq L - |k| - 1$. Therefore $R_{\hat{n}}[-k] = R_{\hat{n}}[k]$.

(iii)

$$\hat{R}_{\hat{n}}[-k] = \sum_{m=0}^{L-1} x[\hat{n}+m]x[\hat{n}+m-k]$$

Making the substitution $m \rightarrow m' + k$ gives

$$\hat{R}_{\hat{n}}[-k] = \sum_{m'=-k}^{L-1-k} x[\hat{n}+m']x[\hat{n}+m'+k]$$

The product term is of the same form as for $\hat{R}_{\hat{n}}[k]$, but the range of summation is different. Therefore

$$\hat{R}_{\hat{n}}[-k] \neq \hat{R}_{\hat{n}}[k]$$

- (c) If $x[n] = x[n + N_p]$, $-\infty < n < \infty$, then
 (i)

$$\begin{aligned} & \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L [(x[m+k] \pm x[m])^2] \geq 0 \\ & \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L [[x[m+k]x[m+k] \pm 2x[m+k]x[m] + x[m]x[m]]] \\ & = 2(\phi[0] \pm \phi[k]) \geq 0 \\ & \phi[0] \pm \phi[k] \geq 0 \end{aligned}$$

Therefore $-\phi[0] \leq \phi[k] \leq \phi[0]$ and $\phi[k]$ is maximum at the origin, i.e., for $k = 0$.

- (ii)

$$\begin{aligned} & \sum_{m=-\infty}^{\infty} (x[\hat{n}+m+k]w'[m+k] \pm x[\hat{n}+m]w'[m])^2 \geq 0 \\ & = 2(R_{\hat{n}}[0] \pm R_{\hat{n}}[k]) \geq 0 \end{aligned}$$

This relation implies:

$$R_{\hat{n}}[k] \leq R_{\hat{n}}[0]$$

- (iii)

$$\hat{R}_{\hat{n}}[k] \text{ is not } \leq \hat{R}_{\hat{n}}[0]$$

in general. This can be seen by considering the sequence:

$$x[n] = \begin{cases} 1 & n = 0 \\ 0 & 0 < n \leq L-1 \\ 2 & n = L \\ 0 & \text{otherwise} \end{cases}$$

For this sequence we have:

$$\begin{aligned} \hat{R}_0[0] &= \sum_{m=0}^{L-1} x^2[m] = 1 \\ \hat{R}_0[L] &= \sum_{m=0}^{L-1} x[m]x[m+L] = 2 \\ \hat{R}_0[L] &> \hat{R}_0[0] \implies \hat{R}_{\hat{n}}[k] \text{ not } \leq \hat{R}_{\hat{n}}[0] \quad \forall \hat{n}, k \end{aligned}$$

- (d) If $x[n] = x[n + N_p]$, $-\infty < n < \infty$, then
 (i)

$$\phi[0] = \lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L x^2[m]$$

where we have the result

$$\lim_{L \rightarrow \infty} = \sum_{m=-L}^L x^2[m]$$

which is the total energy of the signal. Hence

$$\lim_{L \rightarrow \infty} \frac{1}{2L+1} \sum_{m=-L}^L x^2[m] = \text{average energy/unit time=power}$$

Therefore we see that $\phi[0]$ = the power in the signal.

(ii)

$$R_{\hat{n}}[0] = \sum_{m=0}^{L-1} (x[\hat{n} + m]w'[m])^2 = \sum_{m=-\infty}^{\infty} (x[\hat{n} + m]w'[\hat{n}])^2$$

since $w'[m]$ is a causal window of length L . Now we substitute $w'[m] = w[-m]$ giving

$$R_{\hat{n}}[0] = \sum_{m=-\infty}^{\infty} (x[\hat{n} + m]w[-m])^2$$

Next we make the change of variables $m \rightarrow m' - \hat{n}$ giving:

$$R_{\hat{n}}[0] = \sum_{m'=-\infty}^{\infty} (x[m']w[\hat{n} - m'])^2$$

which is the short-time energy.

(iii)

$$\hat{R}_{\hat{n}}[0] = \sum_{m=0}^{L-1} (x[\hat{n} + m])^2$$

Making the change of variables $m \rightarrow m' - \hat{n}$ we get

$$\hat{R}_{\hat{n}}[0] = \sum_{m'=\hat{n}}^{L-1+\hat{n}} (x[m'])^2 = \sum_{m'=\hat{n}}^{L-1+\hat{n}} (x[m']w[\hat{n} - m'])^2$$

where

$$w[n] = \begin{cases} 1 & 0 \leq \hat{n} \leq L-1 \\ 0 & \text{otherwise} \end{cases}$$

We see that even if we restrict the window to be rectangular, $\hat{R}_{\hat{n}}[0]$ does not represent the short-time energy, since the range of summation is different.

- 6.10 (a)** The analog and digital magnitude frequency responses, $X_1(\Omega)$ and $X_1(e^{j\omega})$, are shown in Figure P6.10.1.
- (b)** The analog and digital magnitude frequency responses, $X_2(\Omega)$ and $X_2(e^{j\omega})$, are shown in Figure P6.10.2.
- (c)** The signal $x_1[n] = A \cos(2\pi 100nT) = A \cos(2\pi 100n/10000) = A \cos(2\pi n/100)$ is periodic with period 100 samples, since

$$x_1[n] = x_1[n + r100], \text{ for all } r$$

The signal $x_2[n] = B \cos(2\pi 101nT) = B \cos(2\pi 101n/10000)$ is not periodic since its natural period of $N_2 = 10,000/101 = 99.01$ samples is not an integer. Interestingly, however, $x_2[n]$

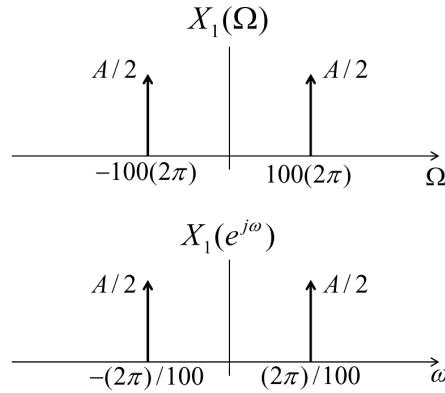


Figure P6.10.1: Analog and digital spectra for 100 Hz periodic signal.

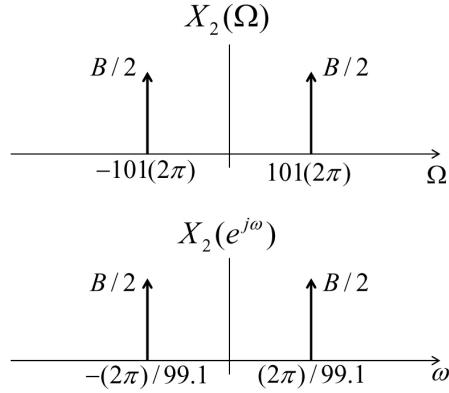


Figure P6.10.2: Analog and digital spectra for 101 Hz periodic signal.

is actually periodic at a period of $N_2 = 10,000$ samples since the function repeats every 10,000 samples.

6.11 (a) Since the signal is periodic, we can use the definition of the autocorrelation function as:

$$\phi[k] = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N x[m]x[m+k] \quad (6.1)$$

Substituting for $x[m]$ the expression $\cos(2\pi m/N_p)$ we get the following form for $\phi[k]$:

$$\phi[k] = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N \cos(2\pi m/N_p) \cos(2\pi(m+k)/N_p) \quad (6.2)$$

Now we can use the trigonometric relation:

$$\cos(a)\cos(b) = \frac{1}{2}[\cos(a+b) + \cos(a-b)] \quad (6.3)$$

giving the result:

$$\phi[k] = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{m=-N}^N \frac{1}{2}[\cos(2\pi(2m+k)/N_p) + \cos(2\pi k/N_p)] \quad (6.4)$$

We readily see that the first term in the above equation sums to 0 since the cosine function is symmetric in value, and the second terms is just the constant $\cos(2\pi k/N_p)/2$. Hence the final result is that:

$$\phi[k] = \frac{\cos(2\pi k/N_p)}{2} \quad (6.5)$$

(b) A sketch of $\phi[k]$ is given in part (a) of Figure P6.11.1.

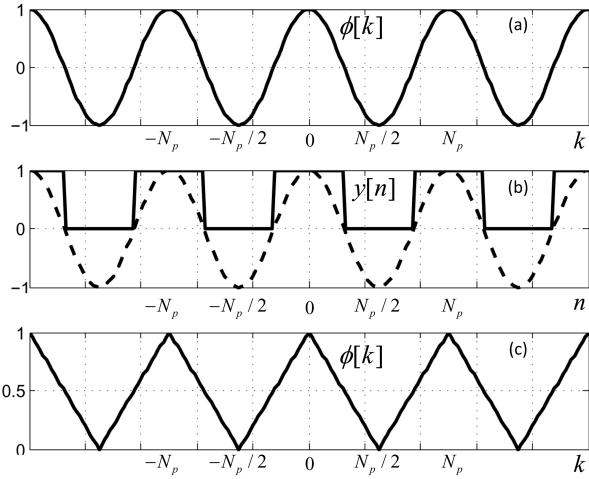


Figure P6.11.1: (a) Plot of autocorrelation function of cosine wave; (b) plot of compressed signal; (c) plot of autocorrelation of compressed signal.

(c)

$$y[n] = \begin{cases} 1 & -\frac{N_p}{4} \leq n \leq \frac{N_p}{4} \\ 0 & -\frac{N_p}{2} \leq n \leq -\frac{N_p}{4}, \quad \frac{N_p}{2} \leq n \leq \frac{N_p}{2} \end{cases} \quad (6.6)$$

We see that the resulting signal is the periodic signal shown by solid lines in part (b) of Fig. P6.11.1. Clearly we can perform the autocorrelation over a single period of the signal, and then periodically extend the resulting autocorrelation, thus giving the autocorrelation shown in part (c) of Fig. P6.11.1.

6.12 (a) We first recognize that we can write the digital signal in the form

$$x[n] = 1 + \cos\left(\frac{3\pi}{16}n\right)$$

which is a periodic signal with a period $N_p = 32$ since $\frac{3\pi}{16}N_p = 2\pi m \Rightarrow N_p = 32$. Thus a rectangular window of length $N = 64$ covers 2 complete cycles of the cosine term in $x[n]$.

Thus we can compute the modified autocorrelation as:

$$\begin{aligned}
 R[k] &= \sum_n \left[1 + \cos\left(\frac{3\pi}{16}n\right) \right] \left[1 + \cos\left(\frac{3\pi}{16}(n+k)\right) \right] \\
 &= \sum_n \left\{ 1 + \cos\left(\frac{3\pi}{16}n\right) + \cos\left(\frac{3\pi}{16}(n+k)\right) \right\} \\
 &\quad + \left\{ \cos\left(\frac{3\pi}{16}n\right) \cos\left(\frac{3\pi}{16}(n+k)\right) \right\} \\
 &= 64 + \sum_n \cos\left(\frac{3\pi}{16}n\right) + \sum_n \cos\left(\frac{3\pi}{16}(n+k)\right) \\
 &\quad + \frac{1}{2} \sum_n \cos\left(\frac{3\pi}{16}(2n+k)\right) + \frac{1}{2} \sum_n \cos\left(\frac{3\pi}{16}k\right) \\
 &= 64 + 32 \cos\left(\frac{3\pi}{16}k\right)
 \end{aligned}$$

where all sums over 1 or more cycles of the cosine sum to 0 by symmetry. A plot of $R[k]$ is shown in Figure P6.12.1.

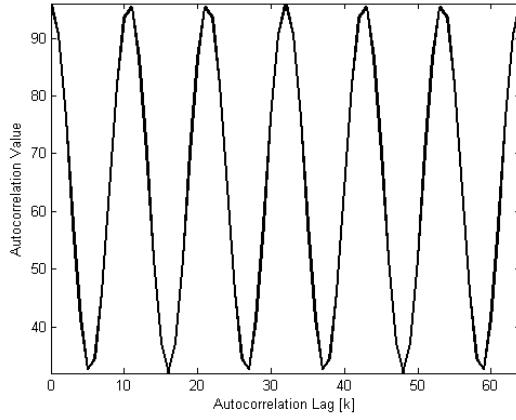


Figure P6.12.1: Plot of the modified autocorrelation.

- (b) The maximum value of $R[k]$ is attained when $\frac{3\pi}{16}k = 2\pi \cdot m$ and this occurs at $k = 32$ and 64 , showing that the digital signal has a periodicity frequency of $32/F_S = 32/1000 = 31.25$ Hz which is different from the periodicity frequency of the original analog signal which is 93.75 Hz. Notice the peaks at lags of 11 and 21 samples which are almost as large as the peak at 32 samples, but are somewhat smaller.

6.13 (a)

We solve for $X_1(e^{j\omega})$ as

$$\begin{aligned}
 X_1(e^{j\omega}) &= \sum_{n=-12}^{12} (1) e^{-j\omega n} = e^{j\omega 12} \frac{(1 - e^{-j\omega 25})}{(1 - e^{-j\omega})} = \frac{\sin(12.5\omega)}{\sin(\omega/2)} \\
 X_1(e^{j2\pi f}) &= \frac{\sin(25\pi f)}{\sin(\pi f)}
 \end{aligned}$$

A plot of the log magnitude response is given in Figure P6.13.1.

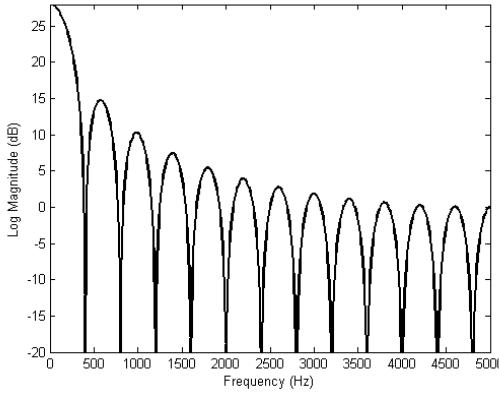


Figure P6.13.1: Plot of the log magnitude spectrum, $X_1(e^{j2\pi f/F_s})$.

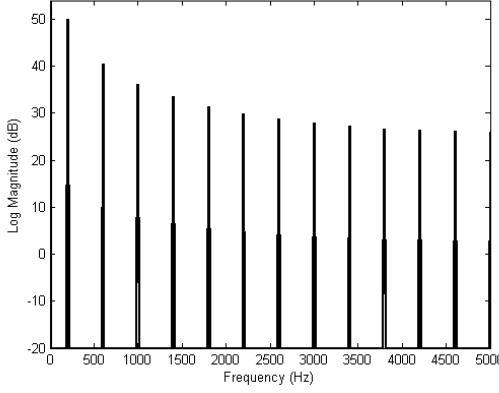


Figure P6.13.2: Plot of the log magnitude spectrum, $X_2(e^{j2\pi f/F_s})$.

- (b) $x_2[n]$ is a periodic version of $x_1[n]$ of part (a) with period $N_p = F_S/200 = 50$ samples, i.e., we can write an expression for one period of $x_2[n]$ as:

$$x_2[n] = \begin{cases} 1 & -12 \leq n \leq 12 \\ 0 & 13 \leq n \leq 37 \end{cases}$$

and the periodic extension, $\tilde{x}_2[n]$, can be written as:

$$\tilde{x}_2[n] = \sum_{r=-\infty}^{\infty} x_2[n - N_p r]$$

Hence $X_2(e^{j2\pi f})$ exists only at the frequencies $f_k = (F_S/N_p)k$, $0 \leq k \leq N_p - 1$ where its values are the sampled values of $X_1(e^{j2\pi f_k})$. A plot of the log magnitude of $X_2(e^{j2\pi f})$ is given in Figure P6.13.2.

- (c) It can be seen that, except for the component at $f = 0$, all other spectral components of $X_2(e^{j2\pi f})$ are the odd harmonics of the signal and occur at the frequencies $f = 200/10000, 600/10000, 1000/10000, \dots$ Hz.

6.14 (a) Given the relationship

$$\frac{1}{L} \sum_{m=0}^{L-1} |x[m]| \leq \left[\frac{1}{L} \sum_{m=0}^{L-1} |x[m]|^2 \right]^{1/2}$$

we can apply the inequality to the expression for $\gamma_{\hat{n}}[k]$ giving:

$$\begin{aligned} \gamma_{\hat{n}}[k] &\leq \left[\frac{1}{L} \sum_{m=0}^{L-1} ([\hat{n} + m] - x[\hat{n} + m - k])^2 \right]^{1/2} \\ \gamma_{\hat{n}}[k] &\leq \left[\frac{1}{L} \sum_{m=0}^{L-1} (x^2[\hat{n} + m] - 2x[\hat{n} + m]x[\hat{n} + m - k] + x^2[\hat{n} + m - k]) \right]^{1/2} \end{aligned}$$

Note that:

$$\sum_{m=0}^{L-1} x^2[\hat{n} + m] = \hat{R}_{\hat{n}}[0]$$

for $x[\hat{n}]$ periodic signal with period L samples. We can now write:

$$\sum_{m=0}^{L-1} x^2[\hat{n} + m - k] = \hat{R}_{\hat{n}}[0]$$

$$\sum_{m=0}^{L-1} 2x[\hat{n} + m]x[\hat{n} + m - k] = 2\hat{R}_{\hat{n}}[k]$$

Therefore:

$$\gamma_{\hat{n}}[k] \leq \left[\frac{1}{L} (2\hat{R}_{\hat{n}}[0] - 2\hat{R}_{\hat{n}}[k]) \right]^{1/2}$$

(b) $x[\hat{n}] = \cos(\omega_0 \hat{n}) = \cos\left[\frac{2\pi\hat{n}}{100}\right]$, period = 100 samples. A plot of $\hat{R}_{\hat{n}}[k]$ is shown in Figure P6.14.1.

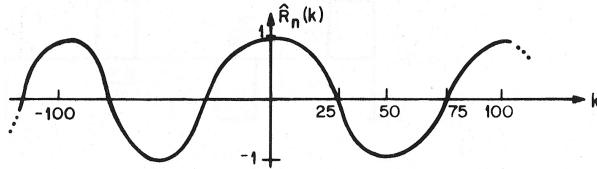


Figure P6.14.1: Plot of $\hat{R}_{\hat{n}}[k]$.

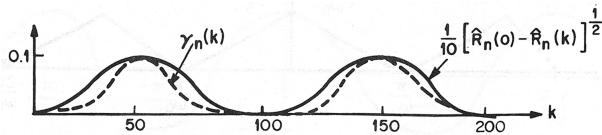
A plot of the function:

$$\left[\frac{2}{200} (\hat{R}_{\hat{n}}[0] - \hat{R}_{\hat{n}}[k]) \right]^{1/2} = \frac{1}{10} (\hat{R}_{\hat{n}}[0] - \hat{R}_{\hat{n}}[k])^{1/2}$$

is given in Figure P6.14.2. It can be seen that

$$\gamma_{\hat{n}}[k] = 0.1 (\hat{R}_{\hat{n}}[0] - \hat{R}_{\hat{n}}[k])^{1/2}$$

only at the maxima and minima of the curves.

Figure P6.14.2: Plot of $\gamma_{\hat{n}}[k]$ and its approximation.

- 6.15 (a)** A sketch of $y[n]$ as a function of n is given in Figure P6.15.1.
- (b)** A sketch of the autocorrelation for clipping levels of 0.5 and 0.75 is given in Figure P6.15.2.
- (c)** As $C_L \rightarrow 1$, the autocorrelation starts to look like an alternating pulse train (i.e., scaled unit sample pulse train). However, as $C_L \rightarrow 1$ the amplitudes of the pulses $\rightarrow 0$. For a signal with time-varying amplitude, too “high” a clipping threshold may result in a loss of autocorrelation peaks, while too “low” a threshold causes the peaks to widen, which complicates pitch detection.

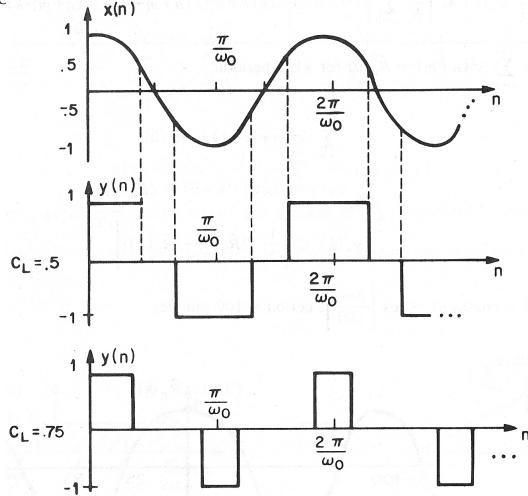


Figure P6.15.1: Plot of the output of the three-level center clipper inputs and outputs.

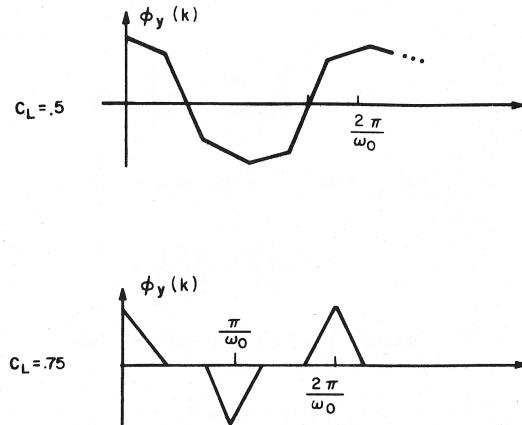


Figure P6.15.2: Sketch of center-clipped autocorrelations for clipping levels of 0.5 and 0.75.

Chapter 7

Frequency-Domain Representations

7.1 By the definition of $X_{\hat{n}}(e^{j\hat{\omega}})$ we have the relationships:

$$a_{\hat{n}}(\hat{\omega}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m] \cos(\hat{\omega}m)$$

$$b_{\hat{n}}(\hat{\omega}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m] \sin(\hat{\omega}m)$$

(a) We can readily show that:

$$a_{\hat{n}}(2\pi - \hat{\omega}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m] \cos[(2\pi - \hat{\omega}) - m]$$

Using the properties of the cos function we get:

$$\cos[(2\pi - \hat{\omega})m] = \cos(2\pi m) \cos(\hat{\omega}m) + \sin(2\pi m) \sin(\hat{\omega}m) = \cos(\hat{\omega}m)$$

Therefore we get $a_{\hat{n}}(2\pi - \hat{\omega}) = a_{\hat{n}}(\hat{\omega})$.

(b) We can readily show that:

$$b_{\hat{n}}(2\pi - \hat{\omega}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m] \sin[(2\pi - \hat{\omega}) - m]$$

Using the properties of the sin function we get:

$$\sin[(2\pi - \hat{\omega})m] = \sin(2\pi m) \cos(\hat{\omega}m) - \cos(2\pi m) \sin(\hat{\omega}m) = -\sin(\hat{\omega}m)$$

Therefore we get $-b_{\hat{n}}(2\pi - \hat{\omega}) = b_{\hat{n}}(\hat{\omega})$.

(c) and (d) both follow directly from the results in parts (a) and (b) since:

$$X_{\hat{n}}(e^{j(2\pi - \hat{\omega})}) = a_{\hat{n}}(2\pi - \hat{\omega}) - jb_{\hat{n}}(2\pi - \hat{\omega}) = a_{\hat{n}}(\hat{\omega}) + jb_{\hat{n}}(\hat{\omega}) = X_{\hat{n}}^*(e^{j\omega})$$

7.2 (a) We can readily obtain expressions of the form:

$$|X_{\hat{n}}(e^{j\hat{\omega}})| = \sqrt{a_{\hat{n}}^2(\hat{\omega}) + b_{\hat{n}}^2(\hat{\omega})}$$

$$\theta_{\hat{n}}(\hat{\omega}) = \tan^{-1} \left[-\frac{b_{\hat{n}}(\hat{\omega})}{a_{\hat{n}}(\hat{\omega})} \right]$$

(b) Similarly we get:

$$a_{\hat{n}}(\hat{\omega}) = \Re\{X_{\hat{n}}(e^{j\hat{\omega}})\} = |X_{\hat{n}}(e^{j\hat{\omega}})| \cos(\theta_{\hat{n}}\hat{\omega})$$

$$b_{\hat{n}}(\hat{\omega}) = \Im\{X_{\hat{n}}(e^{j\hat{\omega}})\} = -|X_{\hat{n}}(e^{j\hat{\omega}})| \sin(\theta_{\hat{n}}\hat{\omega})$$

7.3 (a) Linearity property:

$$\begin{aligned} V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} v[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} (x[m] + y[m])w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m} + \sum_{m=-\infty}^{\infty} y[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= X_{\hat{n}}(e^{j\hat{\omega}}) + Y_{\hat{n}}(e^{j\hat{\omega}}) \end{aligned}$$

(b) Shifting property:

$$\begin{aligned} V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} v[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} x[m-n_0]w[\hat{n}-m]e^{-j\hat{\omega}m} \end{aligned}$$

If we make the substitution $m = m' + n_0$ we get

$$\begin{aligned} V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m'=-\infty}^{\infty} x[m']w[\hat{n}-m'-n_0]e^{-j\hat{\omega}(m'+n_0)} \\ &= e^{-j\hat{\omega}n_0} \sum_{m'=-\infty}^{\infty} x[m']w[\hat{n}-n_0-m']e^{-j\hat{\omega}m'} \\ &= e^{-j\hat{\omega}n_0} X_{\hat{n}-n_0}(e^{j\hat{\omega}}) \end{aligned}$$

(c) Scaling property:

$$\begin{aligned} V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} v[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} \alpha x[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \alpha \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\ &= \alpha X_{\hat{n}}(e^{j\hat{\omega}}) \end{aligned}$$

(d) Modulation property:

$$\begin{aligned}
 V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} v[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\
 &= \sum_{m=-\infty}^{\infty} x[m]e^{j\omega_0 m}w[\hat{n}-m]e^{-j\hat{\omega}m} \\
 &= \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j(\hat{\omega}-\omega_0)m} \\
 &= X_{\hat{n}}(e^{j(\hat{\omega}-\omega_0)})
 \end{aligned}$$

(e) Conjugate symmetry:

$$\begin{aligned}
 V_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \\
 &= \left[\sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{j\hat{\omega}m} \right]^* \\
 &= [X_{\hat{n}}(e^{-j\hat{\omega}})]^* \\
 &= X_{\hat{n}}^*(e^{-j\hat{\omega}})
 \end{aligned}$$

7.4

$$X_{\hat{n}}(e^{j\hat{\omega}}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m}$$

with

$$w[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{j\theta})e^{j\theta n} d\theta$$

Hence $X_{\hat{n}}(e^{j\hat{\omega}})$ can be written as:

$$\begin{aligned}
 X_{\hat{n}}(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{j\theta})e^{j\theta(\hat{n}-m)} d\theta \right] x[m]e^{-j\hat{\omega}m} \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{j\theta})e^{j\theta\hat{n}} \left[\sum_{m=-\infty}^{\infty} x[m]e^{-j(\hat{\omega}+\theta)m} \right] d\theta \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} W(e^{j\theta})e^{j\theta\hat{n}} X(e^{j(\hat{\omega}+\theta)}) d\theta
 \end{aligned}$$

7.5

$$X_{\hat{n}}(e^{j\hat{\omega}}) = \sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m}$$

$$\begin{aligned}
|X_{\hat{n}}(e^{j\hat{\omega}})|^2 &= [X_{\hat{n}}(e^{j\hat{\omega}})] [X_{\hat{n}}(e^{j\hat{\omega}})]^* \\
&= \left[\sum_{m=-\infty}^{\infty} x[m]w[\hat{n}-m]e^{-j\hat{\omega}m} \right] \left[\sum_{r=-\infty}^{\infty} x[r]w[\hat{n}-r]e^{+j\hat{\omega}r} \right] \\
&= \sum_{r=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w[\hat{n}-m]x[m]w[\hat{n}-r]x[r]e^{-j\hat{\omega}(m-r)}
\end{aligned}$$

If we make a change of variables of the form $r = k + m$ we get:

$$\begin{aligned}
|X_{\hat{n}}(e^{j\hat{\omega}})|^2 &= \sum_{k=-\infty}^{\infty} \left[\sum_{m=-\infty}^{\infty} w[\hat{n}-m]x[m]w[\hat{n}-k-m]x[m+k] \right] e^{-j\hat{\omega}k} \\
&= \sum_{k=-\infty}^{\infty} R_{\hat{n}}[k]e^{-j\hat{\omega}k} = \phi_{\hat{n}}(e^{j\hat{\omega}})
\end{aligned}$$

Therefore showing that:

$$R_{\hat{n}}[k] \longleftrightarrow S_{\hat{n}}(e^{j\hat{\omega}})$$

- 7.6** (a) The full width of the main lobe of the Hamming window is $\Delta\omega = 8\pi/L$ in radian frequency or $\Delta f = 4/L$. If we want the full width of the main lobe to correspond to 200 Hz analog frequency, we set $\Delta f = 200$ giving $L = 4/(\Delta f) = 4/200$ seconds, or 20 msec. At a sampling rate of $F_s = 10,000$ Hz, we get $L = 0.02 * F_s = 200$ samples.
- (b) If R is chosen to compute the STFT every 10 msec, then $R = 0.01 * F_s = 100$ samples.
- (c) The spacing between sample points in the frequency domain is $F_s/N = 10,000/1024 = 9.77$ Hz.

- 7.7** (a) $w[n]$ should be stable, corresponding to all poles of $W(z)$ inside the unit circle. Also, $W(e^{j\omega})$ should approximate a “lowpass” type of frequency response, with energy concentrated around $\omega = 0$.
- (b) By definition we have:

$$X_n(e^{j\hat{\omega}}) = \sum_{m=-\infty}^{\infty} x[m]w[n-m]e^{-j\hat{\omega}m} = x[n]e^{-j\hat{\omega}n} * w[n]$$

The sequence $X_n(e^{j\hat{\omega}})$ (for fixed $\hat{\omega}$) is the output of a system whose system function is:

$$W(z) = \frac{\sum_{r=0}^{N_z} b_r z^{-r}}{1 - \sum_{k=1}^{N_p} a_k z^{-k}}$$

Therefore $X_n(e^{j\hat{\omega}})$ satisfies the difference equation:

$$X_n(e^{j\hat{\omega}}) = \sum_{k=1}^{N_p} a_k X_{n-k}(e^{j\hat{\omega}}) + \sum_{r=0}^{N_z} b_r x[n-r]e^{-j\hat{\omega}(n-r)}$$

- (c) $W(z) = 1/(1 - az^{-1})$ and $W(e^{j\omega}) = 1/(1 - ae^{-j\omega})$ where $|a| < 1$ for stability and a must be real. Then we can solve for the magnitude squared frequency response as:

$$|W(e^{j\omega})|^2 = \frac{1}{(1 - a \cos(\omega))^2 + a^2 \sin^2(\omega)} = \frac{1}{1 - 2a \cos(\omega) + a^2}$$

We define the cutoff frequency of the system as the frequency for which:

$$|W(e^{j\omega_c})|^2 = \frac{1}{2}|W(e^{j0})|^2$$

We want ω_c to be equivalent to 50 Hz at a 10 kHz sampling rate; i.e., $\omega_c = 2\pi 50 \cdot 10^{-4} = \pi/100$. Thus:

$$\frac{1}{1 + a^2 - 2a \cos(\pi/100)} = \frac{1}{2(1 - 2a + a^2)}$$

Therefore a must satisfy the quadratic equation

$$a^2 - 2a(2 - \cos(\pi/100)) + 1 = 0$$

for which the solutions are $a = 1.0319$ and $a = 0.96907$. Obviously we choose $a = 0.96907$.

- (d) As seen above, the recursive implementation of the STFT is a digital filtering procedure. Hence for narrowband systems we have to realize a digital filter with poles close to the unit circle. Such systems are notoriously sensitive numerically and require careful implementation to insure stability, low roundoff noise, etc.

7.8 (a) $X_{50}(e^{j\omega})$ uses $x[m]$ for $m \leq 49$ (since $w[0] = 0$); similarly $X_{100}(e^{j\omega})$ uses $x[m]$ for $m \leq 99$.

(b) By definition:

$$X_n(e^{j\hat{\omega}}) = \sum_{m=-\infty}^{\infty} w[n-m]x[m]e^{-j\hat{\omega}m}$$

If we let $y[n] = X_n(e^{j\hat{\omega}})$ and $v[n] = x[n]e^{-j\hat{\omega}n}$, then $y[n] = v[n] * w[n]$ as shown in Figure P7.8.1.

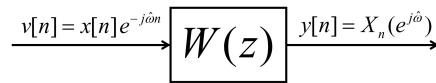


Figure P7.8.1: Block diagram of STFT signal processing.

For the given window, $w[n] = n\beta^n u[n]$ with z -transform

$$W(z) = \frac{\beta z^{-1}}{1 - 2\beta z^{-1} + \beta^2 z^{-2}} = \frac{Y(z)}{V(z)}$$

We can now derive a difference equation relating $y[n]$ to $v[n]$, giving:

$$y[n] = 2\beta y[n-1] - \beta^2 y[n-2] + \beta v[n-1]$$

Substituting for $y[n]$ and $v[n]$ we get:

$$X_n(e^{j\omega}) = 2\beta X_{n-1}(e^{j\omega}) - \beta^2 X_{n-2}(e^{j\omega}) + \beta x[n-1]e^{j\omega(n-1)}$$

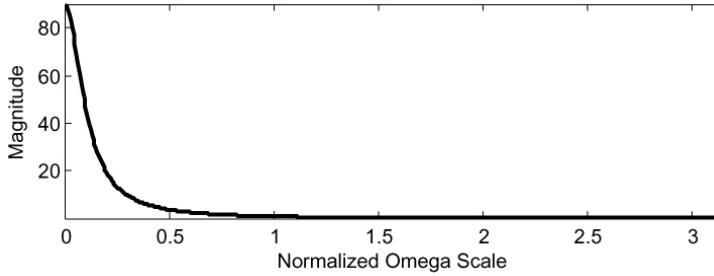


Figure P7.8.2: Plot of magnitude response for window used for STFT calculation.

(c) It is easily shown that:

$$|W(e^{j\omega})| = \frac{\beta}{|1 - \beta e^{-j\omega}|^2} = \frac{\beta}{(1 - 2\beta \cos(\omega) + \beta^2)}$$

A plot of the magnitude response for a value of $\beta = 0.9$ is shown in Figure P7.8.2.

7.9 We can use the summation formula for a geometric series to give:

$$\begin{aligned} \sum_{k=0}^{N-1} e^{j2\pi kn/N} &= \frac{1 - (e^{j2\pi n/N})^N}{1 - e^{j2\pi n/N}} \\ &= \frac{1 - e^{j2\pi n}}{1 - e^{j2\pi n/N}} \end{aligned}$$

The numerator terms are 0 for all integer values of n ; the denominator terms are non-zero except for when $n = rN$ at which point the denominator is also 0. For values of $n = rN$ we can redo the summation, giving

$$\sum_{k=0}^{N-1} (e^{j2\pi r})^k = \sum_{k=0}^{N-1} (1)^k = N$$

and since we recognize that:

$$\sum_{r=-\infty}^{\infty} \delta[n - rN] = \begin{cases} 1 & n = rN \\ 0 & \text{otherwise} \end{cases}$$

we can complete the proof giving

$$\begin{aligned} \frac{1}{N} \sum_{k=0}^{N-1} e^{j(2\pi k/N)n} &= \sum_{r=-\infty}^{\infty} \delta[n - rN] \\ &= \begin{cases} 1 & n = rN, r = 0, \pm 1, \dots \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- 7.10** When we sample the discrete Fourier transform of the sequence at N points around the unit circle, we alias the original sequence around the N -th point. Hence, if N is greater than or equal to the length of the original sequence, there is no aliasing and the sequence corresponding to the sampled discrete Fourier transform is the same as the original sequence.

This is the case (no aliasing) for $N = 40$ and for $N = 10$. For $N = 5$ and $N = 3$ we alias the original sequence around the fifth point (for $N = 5$) and around the third point (for $N = 3$) as shown in Figure P7.10.1.

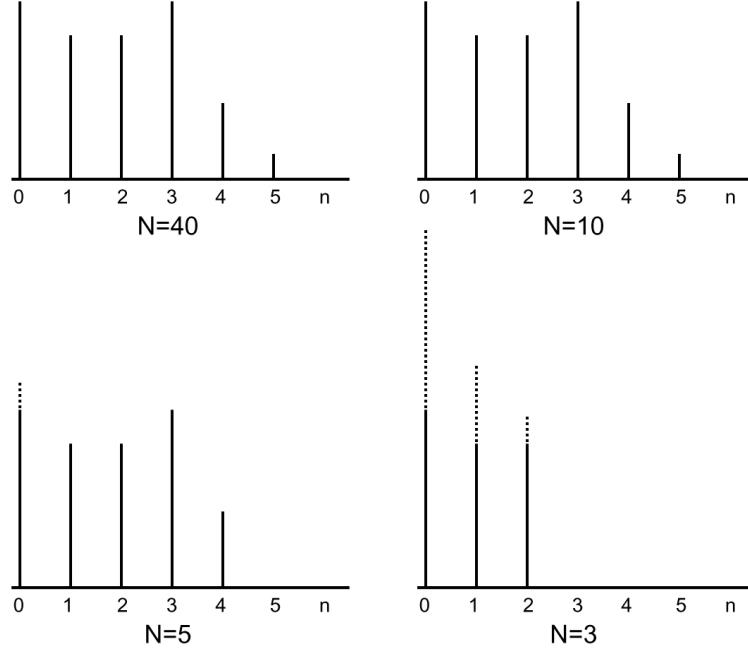


Figure P7.10.1: Plots of aliased digital sequences for $N = 40, 10, 5, 3$.

- 7.11 (a)** The Fourier transform of $x[n]$ is calculated as:

$$X(e^{j\omega}) = \sum_{n=0}^{\infty} r^n e^{-jn\omega} = \frac{1}{1 - re^{-j\omega}}$$

- (b)** We can represent $\tilde{X}[k]$ as the product of $X(e^{j\omega})$ and the sampling function $S(e^{j\omega})$ where:

$$S(e^{j\omega}) = \begin{cases} 1 & \omega = 2\pi k/N, \quad k = 0, 1, \dots, N-1 \\ 0 & \text{otherwise} \end{cases}$$

with impulse response

$$s[n] = \sum_{m=-\infty}^{\infty} \delta[n - mN]$$

We can now solve for $\tilde{x}[n]$ as:

$$\begin{aligned}\tilde{x}[n] &= x[n] * s[n] = \sum_{m=-\infty}^{\infty} x[n - mN] \\ &= \sum_{m=0}^{\infty} r^{n-mN} \\ &= \frac{r^n}{1-r^N}, \quad n = 0, 1, \dots, N-1\end{aligned}$$

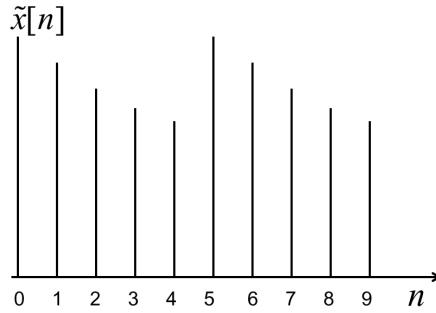


Figure P7.11.1: Plot of $\tilde{x}[n]$ for $n = 0, 1, \dots, 9$ and $r = 0.9$.

- (c) $\tilde{x}[n]$ is a periodic function of period N samples. Thus $\tilde{x}[n]$ is just a scaled r^n function with period of $N = 5$ samples. A plot of $\tilde{x}[n]$ for $n = 0, 1, \dots, 9$ is shown in Figure P7.11.1.

- 7.12 (a)** The estimates of the four formant frequencies are shown in Figure P7.12.1, and their approximate values are $F_1 = 500$ Hz, $F_2 = 1200$ Hz, $F_3 = 2500$ Hz and $F_4 = 3600$ Hz. These values are estimated from the local peaks in the log magnitude spectrum.
(b) The most reliable method of estimating F_0 is to count how many harmonics are seen in the first 1000 Hz of the log magnitude spectrum. In this case there are about 7.5 harmonics in 1000 Hz, so the fundamental frequency is approximately $1000/7.5 = 133$ Hz. The fundamental frequency estimate is also marked in Figure P7.12.1.
(c) The Hamming window needs to be about to resolve individual pitch harmonics. Hence we require that the width of the main lobe of the Hamming window log magnitude spectrum obey the constraint:

$$\left(\frac{8\pi}{L}\right) F_S \leq F_0 \cdot 2\pi$$

Using the values of $F_S = 8000$ and $F_0 = 133$ we get:

$$L \geq \frac{8\pi F_S}{2\pi F_0} = \frac{32000}{133} = 240 \text{ samples}$$

- 7.13** The characteristics features of the five spectra are shown in Figure P7.13.1.

- (a) The voiced speech spectra are A, B and D. In A and D you see the individual pitch harmonics clearly; in B you see the formant structure of a voiced speech sound.

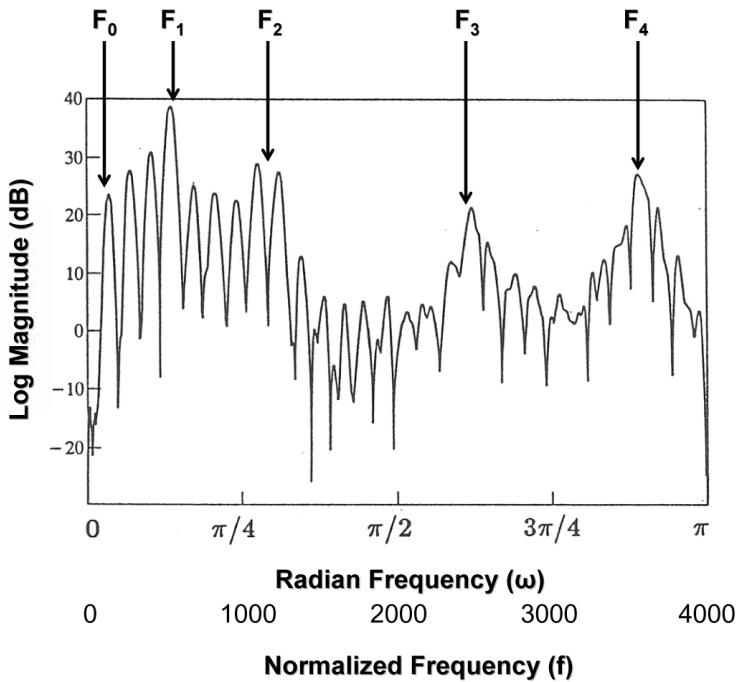


Figure P7.12.1: Log magnitude spectrum of a segment of a speech waveform from a sustained vowel with estimates of formant frequencies and fundamental frequency.

- (b) The unvoiced speech spectra are C and E. In both C and E you see the random spectral structure that is consistent with unvoiced speech sounds.
- (c) The spectra computed using the longest windows were A and C. Using long time windows leads to narrow spectral peaks as seen clearly in spectrum A and less clearly in spectrum C.
- (d)) The spectrum computed using the shortest window was B. This is clearly see in the broad peaks of the spectrum.
- (e) The spectrum which most likely corresponds to a female voice is spectrum D. The fundamental frequency for the section of speech corresponding to spectrum D is about 290 Hz.
- (f) The estimate of fundamental frequency of the section of speech corresponding to spectrum A is $F_0 = 1200/14 \approx 86$ Hz. We obtain this result by recognizing that there are 14 harmonics between 0 and 1200 Hz (assuming a sampling rate of 8000 Hz).
- (g) The estimates of the first three formant frequencies in spectrum B are:

$$\begin{aligned}F_1 &= 0.06 \cdot 8000 = 480 \text{ Hz} \\F_2 &= 0.18 \cdot 8000 = 1440 \text{ Hz} \\F_3 &= 0.275 \cdot 8000 = 2200 \text{ Hz}\end{aligned}$$

- (h) The part of the model that determines the formant frequencies is (g), the vocal tract model.
- (i) The part of the model that determines the spacing of the prominent local peaks in spectrum A is (a), the impulse train generator (controlled by the pitch period input).
- (j) The parts of the model that can shift the spectra by 20 dB are (c) and (f), the voiced and unvoiced gains. To achieve a downward shift of 20 dB we need to reduce the gains by a factor of 10.

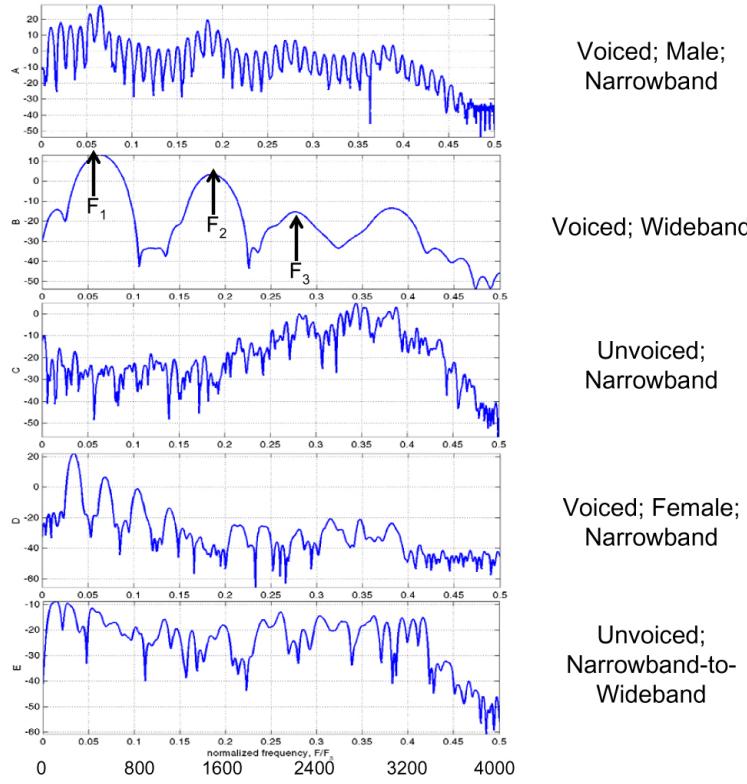


Figure P7.13.1: Annotated log magnitude spectra of different speech frames. (Note the label, A-E, on the left side of the spectral plot which identifies the spectrum under consideration).

- (k) If part (h) of the model were removed, we would need to filter the speech with the inverse of $R(z)$, i.e.,

$$R_i(z) = \frac{1}{1 - 0.99z^{-1}}$$

- 7.14 (a)** From the problem statement we have the relationships:

$$\tilde{x}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}[k] e^{j2\pi kn/N}$$

$$\tilde{X}[k] = \sum_{m=-\infty}^{\infty} x[m] e^{-j2\pi km/N}$$

We can now solve for $\tilde{x}[n]$ as:

$$\begin{aligned}
 \tilde{x}[n] &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=-\infty}^{\infty} x[m] e^{-j2\pi km/N} e^{j2\pi kn/N} \\
 &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=-\infty}^{\infty} x[m] e^{j2\pi k(n-m)/N} \\
 &= \sum_{m=-\infty}^{\infty} x[m] \left[\frac{1}{N} \sum_{k=0}^{N-1} e^{j2\pi k(n-m)/N} \right] \\
 &= \sum_{m=-\infty}^{\infty} x[m] \sum_{r=-\infty}^{\infty} \delta[m - n - r] = \sum_{r=-\infty}^{\infty} x[n - rN]
 \end{aligned}$$

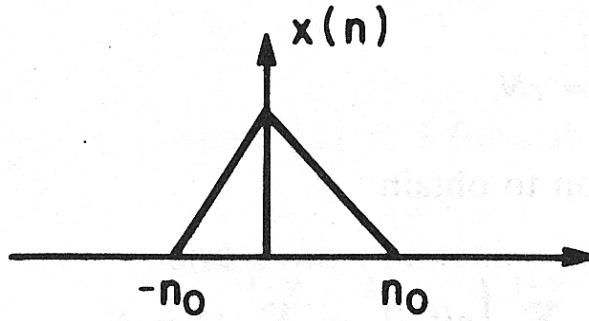


Figure P7.14.1: Base form for $x[n]$.

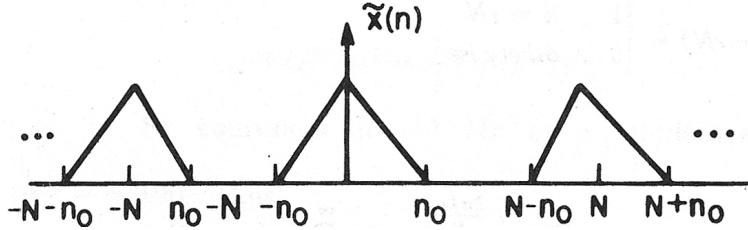


Figure P7.14.2: Form for $\tilde{x}[n]$.

- (b) $\tilde{x}[n]$ is obtained by summing replicated and shifted versions of $x[n]$. For example, if $x[n]$ has the form shown in Figure P7.14.1. To avoid aliasing requires $2n_0 \leq N$. For this example $\tilde{x}[n]$ has the form shown in Figure P7.14.2. The general requirement is $x[n] = 0$, for $n < n_1$ and $n > n_1 + N$. That is $x[n]$ must be of length N samples or shorter.
- (c) We are given that $y[n] = x[nM]$. We can define a sampling function $p[n]$ as:

$$p[n] = \sum_{r=-\infty}^{\infty} \delta[n - rM]$$

Then $w[n] = x[n]p[n]$ and $y[n] = w[nM]$. We also note that

$$p[n] = \frac{1}{M} \sum_{k=0}^{M-1} e^{j2\pi kn/M} = \begin{cases} 1 & n = rM \\ 0 & \text{elsewhere} \end{cases}$$

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{\infty} w[nM]e^{-j\omega n}$$

Since $w[n] = 0$ except for integer multiples of M , then

$$\begin{aligned} Y(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} x[n]p[n]e^{-j\omega n/M} \\ &= \sum_{n=-\infty}^{\infty} x[n] \left[\frac{1}{M} \sum_{k=0}^{M-1} e^{j2\pi kn/M} \right] e^{-j\omega n/M} \\ &= \frac{1}{M} \sum_{k=0}^{M-1} \left[\sum_{n=-\infty}^{\infty} x[n]e^{-j(\omega - 2\pi k)(n/M)} \right] \end{aligned}$$

but since

$$\sum_{n=-\infty}^{\infty} x[n]e^{-j(\omega - 2\pi k)(n/M)} = X(e^{j(\omega - 2\pi k)/M})$$

therefore:

$$Y(e^{j\omega}) = \frac{1}{M} \sum_{n=-\infty}^{\infty} X(e^{j(\omega - 2\pi k)/M})$$

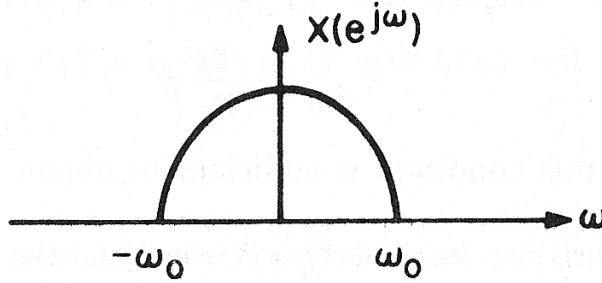


Figure P7.14.3: Base form for $X(e^{j\omega})$.

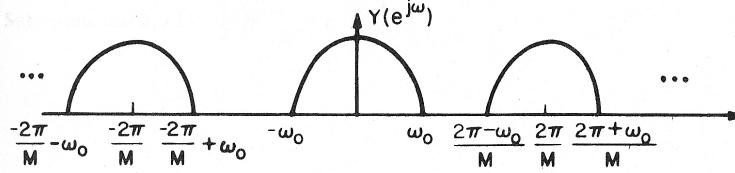


Figure P7.14.4: Form for $Y(e^{j\omega})$.

- (d) $Y(e^{j\omega})$ is obtained by summing replicated and shifted versions of $X(e^{j\omega})$. For example, if $X(e^{j\omega})$ has the form shown in Figure P7.14.3, then to avoid aliasing requires $2\omega_0 \leq (2\pi)/M$. For this condition, $Y(e^{j\omega})$ will have the form shown in Figure P7.14.4.

7.15 We are given that $w[n] \longleftrightarrow W(e^{j\Omega T}) = 0$ for $|\Omega| > \Omega_c$.

- (a) With $\hat{w}[r] = w[rR - n]$ we first express the term $\tilde{w}[m - n] = w[m - n]$ as:

$$\tilde{w}[m] \longleftrightarrow \tilde{W}(e^{j\Omega T}) = e^{-j\Omega nT} W(e^{j\Omega T})$$

$\hat{w}[r]$ is a decimated version of $\tilde{w}[m]$, and the formula for decimating yields

$$\hat{W}(e^{j\Omega T'}) = \frac{1}{R} \sum_{k=0}^{R-1} e^{-jn(\Omega T' - 2\pi k)/R} W(e^{j(\Omega T' - 2\pi k)/R})$$

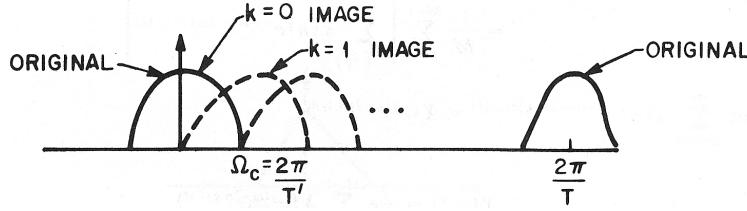


Figure P7.15.1: Typical spectrum.

- (b) A sketch of a typical spectrum is given in Figure P7.15.1. In order to avoid aliasing of the zero-th sample of the spectrum, we require

$$\frac{2\pi}{T'} > \Omega_c \Rightarrow \frac{2\pi}{RT} > \Omega_c$$

Therefore

$$R_{\max} < \frac{2\pi}{\Omega_c T}$$

Strictly speaking this condition is sufficient to obtain the desired result. However, to avoid aliasing the entire spectrum of $W(e^{j\omega})$, not just the DC value, requires

$$R_{\max} < \frac{4\pi}{\Omega_c T}$$

i.e., a 2-fold increase in sampling rate of the window.

- (c) $\sum_{r=-\infty}^{\infty} \hat{w}[r] e^{-j\Omega T' r} = \hat{W}(e^{j\Omega T'})$. Let $\Omega = 0$, then

$$\sum_{r=-\infty}^{\infty} \hat{w}[r] = \sum_{r=-\infty}^{\infty} \hat{w}[rR - n] = \hat{W}(e^{j0}) = \frac{W(e^{j0})}{R}$$

- 7.16** This problem concerns exact reconstruction with a Hann window, which is defined as

$$w_{\text{Hann}}[n] = [0.5 + 0.5 \cos(\pi n/M)] w_r[n] = \begin{cases} [0.5 + 0.5 \cos(\pi n/M)] & -M \leq n \leq M \\ 0 & \text{otherwise} \end{cases}$$

We use the rectangular window $w_r[n]$ to impose the finite range. The key to showing exact reconstruction really lies in the fact that $[0.5 + 0.5 \cos(\pi n/M)] = 0$ for $n = \pm M$. This means that the values of the rectangular window at $n = \pm M$ are irrelevant. They can be 1 or 0 (or any other value) and we still have the Hann window.

- (a) We begin by defining

$$w_{r1}[n] = \begin{cases} 1 & -M \leq n \leq M - 1 \\ 0 & \text{otherwise.} \end{cases}$$

Delay this by M to get $w'_{r1}[n] = w_{r1}[n - M]$, whose DTFT is easily shown to be

$$W'_{r1}(e^{j\omega}) = \sum_{n=0}^{2M-1} e^{-jn\omega} = \frac{1 - e^{-j\omega 2M}}{1 - e^{-j\omega}}.$$

Therefore, since $w_{r1}[n] = w'_{r1}[n + M]$, it follows that

$$W_{r1}(e^{j\omega}) = W'_{r1}(e^{j\omega})e^{j\omega M} = \frac{1 - e^{-j\omega 2M}}{1 - e^{-j\omega}}e^{j\omega M}.$$

(b) Since $w_{\text{Hann}}[n] = [0.5 + 0.5 \cos(\pi n/M)]w_{r1}[n]$, we can also write

$$w_{\text{Hann}}[n] = 0.5w_{r1}[n] + 0.25w_{r1}[n]e^{j\pi n/M} + 0.25w_{r1}[n]e^{-j\pi n/M},$$

so using the modulation theorem for DTFTs we get

$$\begin{aligned} W_{\text{Hann}}(e^{j\omega}) &= 0.5W_{r1}(e^{j\omega}) + 0.25W_{r1}(e^{j(\omega-\pi/M)}) + 0.25W_{r1}(e^{j(\omega+\pi/M)}) \\ &= 0.5\frac{1 - e^{-j\omega 2M}}{1 - e^{-j\omega}}e^{j\omega M} + 0.25\frac{1 - e^{-j(\omega-\pi/M)2M}}{1 - e^{-j(\omega-\pi/M)}}e^{j(\omega-\pi/M)M} \\ &\quad + 0.25\frac{1 - e^{-j(\omega+\pi/M)2M}}{1 - e^{-j(\omega+\pi/M)}}e^{j(\omega+\pi/M)M} \end{aligned}$$

(c) Now from (b) it follows that

$$W_{\text{Hann}}(e^{j2\pi k/M}) = 0.5W_{r1}(e^{j2\pi k/M}) + 0.25W_{r1}(e^{j(2\pi k/M - \pi/M)}) + 0.25W_{r1}(e^{j(2\pi k/M + \pi/M)}).$$

where

$$W_{r1}(e^{j2\pi k/M}) = \frac{1 - e^{-j(2\pi k/M)2M}}{1 - e^{-j(2\pi k/M)}}e^{j(2\pi k/M)M} = \frac{1 - e^{-j4\pi k}}{1 - e^{-j(2\pi k/M)}}e^{j2\pi k} = 0.$$

for $k = 1, 2, \dots, M-1$. Likewise

$$W_{r1}(e^{j(2\pi k/M \pm \pi/M)}) = \frac{1 - e^{-j(2\pi k/M \pm \pi/M)2M}}{1 - e^{-j(2\pi k/M \pm \pi/M)}}e^{j(2\pi k/M \pm \pi/M)M} = \frac{1 - e^{-j2\pi(2k \pm 1)}}{1 - e^{-j\pi(2k \pm 1)/M}}e^{j\pi(2k \pm 1)} = 0.$$

for $k = 1, 2, \dots, M-1$. Therefore the zeros of $W_{\text{Hann}}(e^{j\omega})$ align with the frequencies $2\pi k/R$ if $R = M$ and we have perfect reconstruction. If M is divisible by 2, then $R = M/2$ works as well since $2\pi k/R = 4\pi k/M$, which would be the even indexed zeros.

(d) Again using the equation derived in (b), we get

$$W_{\text{Hann}}(e^{j0}) = 0.5W_{r1}(e^{j0}) + 0.25W_{r1}(e^{j(0-\pi/M)}) + 0.25W_{r1}(e^{j(0+\pi/M)}).$$

Again, we can look at the parts.

$$W_{r1}(e^{j0}) = \sum_{n=-M}^{M-1} e^{-jn\omega} = 2M,$$

since there are $2M$ non-zero terms all one. Also we need to look at the terms

$$W_{r1}(e^{\pm j\pi/M}) = \frac{1 - e^{-j(0 \pm \pi/M)2M}}{1 - e^{-j(0 \pm \pi/M)}}e^{j(0 \pm \pi/M)M} = \frac{1 - e^{\pm j2\pi}}{1 - e^{\pm j\pi/M}}e^{\pm j\pi} = 0.$$

Therefore $W_{\text{Hann}}(e^{j0}) = 2M/2 = M$, so that the reconstruction gain with the Hann window is $C = M/R$.

(e) The three other rectangular windows that work are

$$w_{r2}[n] = \begin{cases} 1 & -M \leq n \leq M \\ 0 & \text{otherwise} \end{cases} \iff W_{r2}(e^{j\omega}) = \frac{1 - e^{-j\omega(2M+1)}}{1 - e^{-j\omega}} e^{j\omega M}.$$

and

$$w_{r3}[n] = \begin{cases} 1 & -M + 1 \leq n \leq M - 1 \\ 0 & \text{otherwise} \end{cases} \iff W_{r3}(e^{j\omega}) = \frac{1 - e^{-j\omega(2M-1)}}{1 - e^{-j\omega}} e^{j\omega(M-1)}.$$

and finally,

$$w_{r4}[n] = \begin{cases} 1 & -M + 1 \leq n \leq M \\ 0 & \text{otherwise} \end{cases} \iff W_{r4}(e^{j\omega}) = \frac{1 - e^{-j\omega(2M)}}{1 - e^{-j\omega}} e^{j\omega(M-1)}.$$

(f) The graph produced by the MATLAB program given below is shown in Figure P7.17.1. In this figure $M = 10$. Observe that all the DTFTs are zero at $\omega/\pi = .1k$ for $k = 2, 3, \dots, 18$. For our purposes, we take the even-indexed set where $\omega/\pi = .2k$ where $k = 1, 2, \dots, 9$, i.e. $\omega = 2\pi k/M$, where $M = 10$ and $k = 1, 2, \dots, M - 1$. This guarantees perfect reconstruction with $R = M = 10$. Note that the same holds for $\omega = 2\pi k/5$, for $k = 1, 2, 3, 4$, so we also have perfect reconstruction with $R = M/2 = 5$.

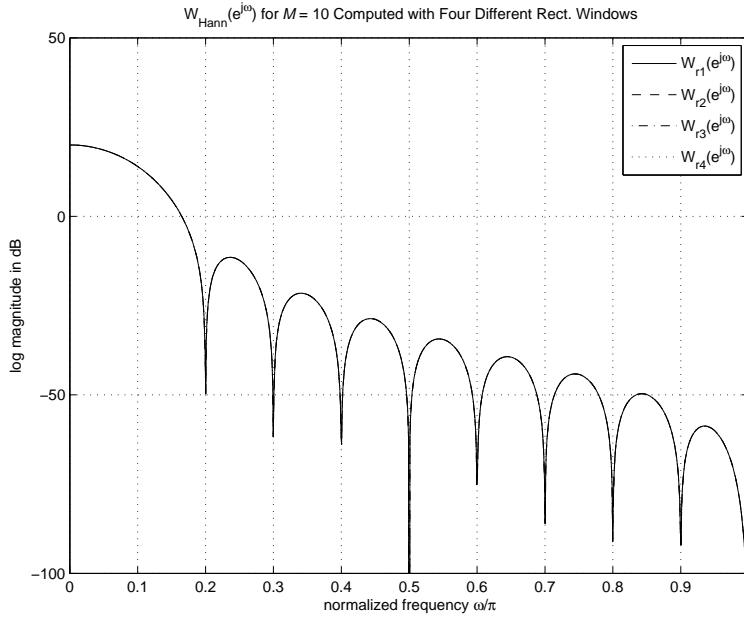


Figure P7.17.1: Plots of all four rectangular window DTFTs.

The MATLAB program that produced the plot of the different window DTFTs is

```
function problem7_16
%solution to problem 7.16
omega=(0:1024)*pi/1024;
M=10;

%
```

```

WH1=.5*Wr1(omega,M) + .25*Wr1(omega-pi/M,M) + .25*Wr1(omega+pi/M,M);
WH2=.5*Wr2(omega,M) + .25*Wr2(omega-pi/M,M) + .25*Wr2(omega+pi/M,M);
WH3=.5*Wr3(omega,M) + .25*Wr3(omega-pi/M,M) + .25*Wr3(omega+pi/M,M);
WH4=.5*Wr4(omega,M) + .25*Wr4(omega-pi/M,M) + .25*Wr4(omega+pi/M,M);

plot(omega/pi,20*log10(abs(WH1)),'-',omega/pi,20*log10(abs(WH2)), '--',...
      omega/pi,20*log10(abs(WH3)), '-.', omega/pi,20*log10(abs(WH4)), ':')
axis([0,1,-100,50]);grid
xlabel('normalized frequency \omega/\pi')
ylabel('log magnitude in dB')
title(['W_{Hann}(e^{j\omega}) for \it{M} = ',num2str(M), ' Computed with Four Different Rect. W...
legend('W_{r1}(e^{j\omega})','W_{r2}(e^{j\omega})',...
      'W_{r3}(e^{j\omega})','W_{r4}(e^{j\omega})')

print -deps prob7_16.eps

function W=Wr1(omega,M)
W =( (1-exp(-j*omega*2*M))./(1-exp(-j*omega)) ).*exp(j*omega*M);

function W=Wr2(omega,M)
W =( (1-exp(-j*omega*(2*M+1)))./(1-exp(-j*omega)) ).*exp(j*omega*M);

function W=Wr3(omega,M)
W=( (1-exp(-j*omega*(2*M-1)))./(1-exp(-j*omega)) ).*exp(j*omega*(M-1));

function W=Wr4(omega,M)
W =( (1-exp(-j*omega*2*M))./(1-exp(-j*omega)) ).*exp(j*omega*(M-1));
*****

```

7.17 This problem deals with back-to-back Hann windows for both analysis and synthesis.

- (a) The effective window is $w_{\text{eff}}[n] = (w_{\text{Hann}}[n])^2$ so

$$\begin{aligned}
 w_{\text{eff}}[n] &= [0.5 + 0.5 \cos(\pi n/M)]^2 (w_r[n])^2 \\
 &= [0.25 + 0.5 \cos(\pi n/M) + 0.25 \cos^2(\pi n/M)] w_r[n] \quad (\text{since } w_r^2[n] = w_r[n]) \\
 &= [0.25 + 0.5 \cos(\pi n/M) + 0.125 + 0.125 \cos(2\pi n/M)] w_r[n] \\
 &= [0.375 + 0.5 \cos(\pi n/M) + 0.125 \cos(2\pi n/M)] w_r[n]
 \end{aligned}$$

Note that there was an error in the problem statement in the first printing of the book. The coefficient of $\cos(\pi n/M)$ was given as 1 rather than 0.5 as determined above.

- (b) Using the modulation theorem for DTFTs

$$\begin{aligned}
 W_{\text{eff}}(e^{j\omega}) &= 0.375W_r(e^{j\omega}) + 0.25W_r(e^{j(\omega-\pi/M)}) + 0.25W_r(e^{j(\omega+\pi/M)}) \\
 &\quad + 0.0625W_r(e^{j(\omega-2\pi/M)}) + 0.0625W_r(e^{j(\omega+2\pi/M)}).
 \end{aligned}$$

- (c) Now assuming that $w_r[n] = w_{r1}[n]$ from Problem 7.16 we find that

$$W_r(e^{j\omega}) = \frac{1 - e^{-j\omega 2M}}{1 - e^{-j\omega}} e^{j\omega M}.$$

We need to look at $W_r(e^{j4\pi k/M})$, $W_r(e^{j(4\pi k/M \pm \pi/M)})$, and $W_r(e^{j(4\pi k/M \pm 2\pi/M)})$ for $k = 1, 2, \dots, M/2 - 1$ since the range of k covers the frequency range between 0 and 2π . (Note

that we assume that $M/2$ is an integer.)

$$W_r(e^{j4\pi k/M}) = \frac{1 - e^{-j(4\pi k/M)2M}}{1 - e^{-j4\pi k/M}} e^{j(4\pi k/M)M} = \frac{1 - e^{-j8\pi k}}{1 - e^{-j4\pi k/M}} e^{j4\pi k} = 0$$

for $k = 1, 2, \dots, M/2 - 1$, and

$$W_r(e^{j(4\pi k/M \pm \pi/M)}) = \frac{1 - e^{-j(4\pi k/M \pm \pi/M)2M}}{1 - e^{-j(4\pi k/M \pm \pi/M)}} e^{j(4\pi k/M \pm \pi/M)M} = \frac{1 - e^{-j2\pi(4k \pm 1)}}{1 - e^{-j(4\pi k/M \pm \pi/M)}} e^{j\pi(4k \pm 1)} = 0$$

for $k = 1, 2, \dots, M/2 - 1$, and

$$W_r(e^{j(4\pi k/M \pm 2\pi/M)}) = \frac{1 - e^{-j(4\pi k/M \pm 2\pi/M)2M}}{1 - e^{-j(4\pi k/M \pm 2\pi/M)}} e^{j(4\pi k/M \pm 2\pi/M)M} = \frac{1 - e^{-j4\pi(2k \pm 1)}}{1 - e^{-j(4\pi k/M \pm 2\pi/M)}} e^{j4\pi(2k \pm 1)} = 0,$$

for $k = 1, 2, \dots, M/2 - 1$. Therefore, $W_{\text{eff}}(e^{j4\pi k/M}) = 0$ for $k = 1, 2, \dots, M - 1$, and we can have perfect reconstruction with back-to-back Hann windows if $R = M/2$.

Now we need to determine the reconstruction gain by finding $W_{\text{eff}}(e^{j0})$.

$$\begin{aligned} W_{\text{eff}}(e^{j0}) &= 0.375W_r(e^{j0}) + 0.25W_r(e^{j(0-\pi/M)}) + 0.25W_r(e^{j(0+\pi/M)}) \\ &\quad + 0.0625W_r(e^{j(0-2\pi/M)}) + 0.0625W_r(e^{j(0+2\pi/M)}) \end{aligned}$$

This gets a little tedious, but it is easily shown that $W_r(e^{j0}) = 2M$, $W_r(e^{\pm j\pi/M}) = 0$, and $W_r(e^{\pm j2\pi/M}) = 0$ so

$$W_{\text{eff}}(e^{j0}) = 0.375(2M) = (2M)3/8 = 3M/4$$

So the reconstruction gain of back-to-back Hann windows with $R = M/2$ is

$$C = W_{\text{eff}}(e^{j0})/R = (3M/4)/(M/2) = 3/2.$$

7.18 The figure from the problem is repeated below.

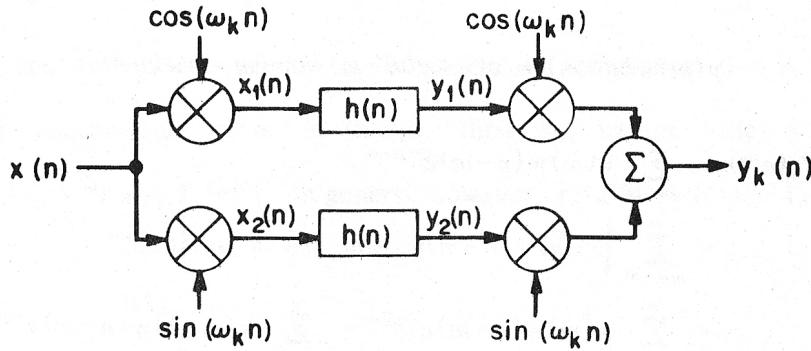


Figure P7.18.1:

- (a) The impulse response is obtained by letting $x[n] = \delta[n]$. From Figure P7.18.1 we get:

$$x_1[n] = x[n] \cos(\omega_k n) = \delta[n] \cos(\omega_k n) \equiv \delta[n]$$

$$x_2[n] = x[n] \sin(\omega_k n) = \delta[n] \sin(\omega_k n) \equiv 0$$

Then

$$y_1[n] = x_1[n] * h[n] = \delta[n] * h[n] = h[n]$$

$$y_2[n] = x_2[n] * h[n] = 0 * h[n] = 0$$

Finally

$$y_k[n] = h_k[n] = y_1[n] \cos(\omega_k n) = h[n] \cos(\omega_k n)$$

(b) Let $H_k(e^{j\omega})$ denote the overall system function, i.e.,

$$H_k(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h_k[n] e^{-j\omega n} = \sum_{n=-\infty}^{\infty} h[n] \cos(\omega_k n) e^{-j\omega n}$$

Substitute $\cos(\omega_k n) = (1/2)(e^{j\omega_k n} + e^{-j\omega_k n})$ giving:

$$\begin{aligned} H_k(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} \frac{h[n]}{2} (e^{j\omega_k n} + e^{-j\omega_k n}) e^{-j\omega n} \\ &= \frac{1}{2} \sum_{n=-\infty}^{\infty} h[n] e^{-j(\omega - \omega_k)n} + \frac{1}{2} \sum_{n=-\infty}^{\infty} h[n] e^{-j(\omega + \omega_k)n} \\ &= \frac{1}{2} H(e^{j(\omega - \omega_k)}) + \frac{1}{2} H(e^{j(\omega + \omega_k)}) \end{aligned}$$

7.19 (a) Setting $y[n] = x[n] - x[n - 1]$ gives the windowed version of $y[n]$ as:

$$y[m]w[n - m] = x[m]w[n - m] - x[m - 1]w[n - m]$$

We can now write the STFT of $y[n]$ as:

$$\begin{aligned} Y_n(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} y[m]w[n - m] e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} [x[m]w[n - m] - x[m - 1]w[n - m]] e^{-j\hat{\omega}m} \\ &= \sum_{m=-\infty}^{\infty} x[m]w[n - m] e^{-j\hat{\omega}m} - \sum_{m=-\infty}^{\infty} x[m - 1]w[n - m] e^{-j\hat{\omega}m} \end{aligned}$$

If we make the change of variables $m = k + 1$ in the second term of the equation above we get

$$\begin{aligned} Y_n(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} x[m]w[n - m] e^{-j\hat{\omega}m} - e^{-j\hat{\omega}} \sum_{k=-\infty}^{\infty} x[k]w[n - 1 - k] e^{-j\hat{\omega}k} \\ &= X_n(e^{j\hat{\omega}}) - e^{-j\hat{\omega}} X_{n-1}(e^{j\hat{\omega}}) \end{aligned}$$

(b) If $w[n - m] \approx w[n - 1 - m]$, then $X_n(e^{j\hat{\omega}}) \approx X_{n-1}(e^{j\hat{\omega}})$. This implies that $w[n]$ is a smooth window with no discontinuities or rapid variations.

- (c) Linear filtering of $y[n]$ can be expressed as a convolution of the form:

$$y[n] = \sum_{k=0}^{N-1} h[k]x[n-k]$$

Using the STFT formulation we get:

$$\begin{aligned} y[m]w[n-m] &= \sum_{k=0}^{N-1} h[k]x[m-k]w[n-m] \\ Y_n(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} y[m]w[n-m]e^{-j\hat{\omega}m} \\ Y_n(e^{j\hat{\omega}}) &= \sum_{m=-\infty}^{\infty} \sum_{k=0}^{N-1} h[k]x[m-k]w[n-m]e^{-j\hat{\omega}m} \end{aligned}$$

If we let $m = m' + k$ we get:

$$\begin{aligned} Y_n(e^{j\hat{\omega}}) &= \sum_{m'=-\infty}^{\infty} \sum_{k=0}^{N-1} h[k]x[m']w[n-k-m']e^{-j\hat{\omega}(m'+k)} \\ &= \sum_{k=0}^{N-1} h[k]e^{-j\hat{\omega}k} \left[\sum_{m'=-\infty}^{\infty} x[m']w[n-k-m']e^{-j\hat{\omega}m'} \right] \\ &= \sum_{k=0}^{N-1} h[k]e^{-j\hat{\omega}k} X_{n-k}(e^{j\hat{\omega}}) \\ &= X_n(e^{j\hat{\omega}}) * h_{\hat{\omega}}[n] \end{aligned}$$

where $h_{\hat{\omega}}[n] = h[n]e^{-j\hat{\omega}n}$.

- (d) If the spectrum of the window is “impulsive” in appearance, this is with most of the energy concentrated close to $\hat{\omega} = 0$, then we expect little difference between $H(e^{j\hat{\omega}})X_n(e^{j\hat{\omega}})$ and $Y_n(e^{j\hat{\omega}})$. In general, however, $Y_n(e^{j\hat{\omega}}) \neq H(e^{j\hat{\omega}})X_n(e^{j\hat{\omega}})$.

- 7.20 (a)** Figure P7.20.1 shows a sketch of the locations of the N filter bands for both N even and N odd.

- (b) For N even we have:

$$\begin{aligned} \tilde{h}[n] &= \sum_{k=0}^{N-1} P_k h_k[n] = \sum_{k=0}^{N-1} |P_k| w_k[n] e^{j(\omega_k n + \phi_k)} \\ &= \sum_{k=0}^{N-1} P_k w_k[n] e^{j\omega_k n} \\ \tilde{h}[n] &= P_0 w_0[n] + P_{N/2} w_{N/2}[n] e^{-j\pi n} + \sum_{k=0}^{N/2-1} P_k w_k[n] e^{j\omega_k n} + \sum_{k=N/2+1}^{N-1} P_k w_k[n] e^{j\omega_k n} \\ \tilde{h}[n] &= P_0 w_0[n] + (-1)^n P_{N/2} w_{N/2}[n] + \sum_{k=0}^{N/2-1} P_k w_k[n] e^{j\omega_k n} + \sum_{k=1}^{N/2-1} P_{N-k} w_{N-k}[n] e^{j\omega_{N-k} n} \end{aligned}$$

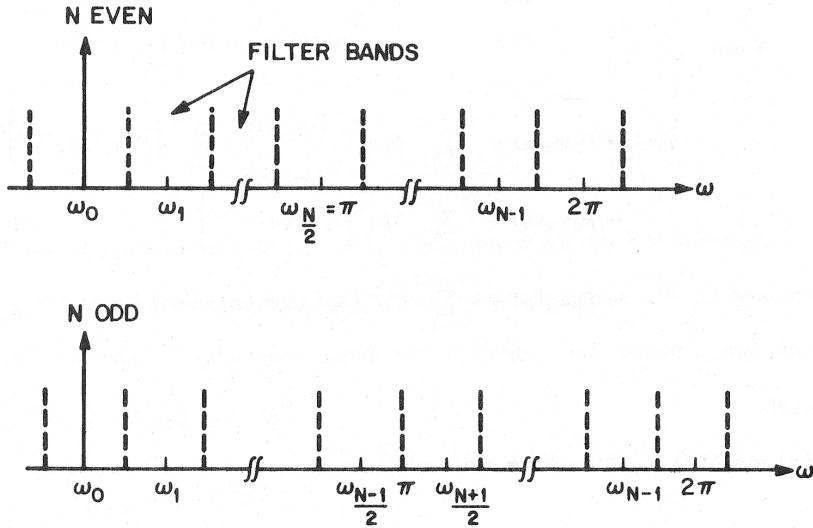


Figure P7.20.1:

We substitute $P_{N-k} = P_k^*$, $\omega_{N-k} = 2\pi - \omega_k$, $w_{N-k}[n] = w_k[n]$ giving

$$\begin{aligned}\tilde{h}[n] &= P_0 w_0[n] + (-1)^n P_{N/2} w_{N/2}[n] + \sum_{k=1}^{N/2-1} \{P_k w_k[n] e^{j\omega_k n} + (P_k w_k[n] e^{j\omega_k n})^*\} \\ \tilde{h}[n] &= P_0 w_0[n] + \sum_{k=1}^{N/2-1} 2\Re \{P_k w_k[n] e^{j\omega_k n}\} + (-1)^n P_{N/2} w_{N/2}[n] \\ \tilde{h}[n] &= P_0 w_0[n] + \sum_{k=1}^{N/2-1} 2|P_k| w_k[n] \cos(\omega_k n + \phi_k) + (-1)^n P_{N/2} w_{N/2}[n]\end{aligned}$$

For N odd we have:

$$\begin{aligned}\tilde{h}[n] &= P_0 w_0[n] + \sum_{k=1}^{(N-1)/2} P_k w_k[n] e^{j\omega_k n} + \sum_{k=1}^{(N-1)/2} [P_k w_k[n] e^{j\omega_k n}]^* \\ \tilde{h}[n] &= P_0 w_0[n] + \sum_{k=1}^{(N-1)/2} 2\Re \{P_k w_k[n] e^{j\omega_k n}\} \\ \tilde{h}[n] &= P_0 w_0[n] + \sum_{k=1}^{(N-1)/2} 2|P_k| w_k[n] \cos(\omega_k n + \phi_k)\end{aligned}$$

7.21 (a)

$$\begin{aligned}
H(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} \tilde{h}[n]e^{-j\omega n} \\
&= \alpha_1 + \alpha_2 e^{-j\omega N} + \alpha_3 e^{-2j\omega N} = e^{-j\omega N}[\alpha_1 e^{j\omega N} + \alpha_2 + \alpha_3 e^{-j\omega N}] \\
&= e^{-j\omega N}[\alpha_2 + (\alpha_1 + \alpha_3) \cos(\omega N) + j(\alpha_1 - \alpha_3) \sin(\omega N)] \\
|H(e^{j\omega})|^2 &= [\alpha_2 + (\alpha_1 + \alpha_3) \cos(\omega N)]^2 + [(\alpha_1 - \alpha_3) \sin(\omega N)]^2
\end{aligned}$$

(b) From part (a) we get:

$$\theta(\omega) = -\omega N + \tan^{-1} \left[\frac{(\alpha_1 - \alpha_3) \sin(\omega N)}{\alpha_2 + (\alpha_1 + \alpha_3) \cos(\omega N)} \right]$$

(c)

$$\begin{aligned}
|H(e^{j\omega})|^2 &= [\alpha_2 + (\alpha_1 + \alpha_3) \cos(\omega N)]^2 + [(\alpha_1 - \alpha_3) \sin(\omega N)]^2 \\
\frac{\partial}{\partial \omega} |H(e^{j\omega})|^2 &= -2[\alpha_2 + (\alpha_1 + \alpha_3) \cos(\omega N)]N(\alpha_1 + \alpha_3) \sin(\omega N) \\
&\quad + 2[(\alpha_1 - \alpha_3) \sin(\omega N)]N(\alpha_1 - \alpha_3) \cos(\omega N) = 0 \\
-\alpha_2(\alpha_1 + \alpha_3) \sin(\omega N) &+ [(\alpha_1 - \alpha_3)^2 - (\alpha_1 + \alpha_3)^2] \sin(\omega N) \cos(\omega N) = 0 \\
\frac{\partial}{\partial \omega} |H(e^{j\omega})|^2 &= -\alpha_2(\alpha_1 + \alpha_3) \sin(\omega N) - 4\alpha_1\alpha_3 \sin(\omega N) \cos(\omega N) = 0
\end{aligned}$$

The minima and maxima occur for the conditions:

$$\begin{aligned}
\sin(\omega N) &= 0 \\
\cos(\omega N) &= -\alpha_2 \frac{(\alpha_1 + \alpha_3)}{4\alpha_1\alpha_3}
\end{aligned}$$

The second equation is satisfied by real ω if and only if $4|\alpha_1||\alpha_3| > |\alpha_1 + \alpha_3||\alpha_2|$. Since $|\alpha_1 + \alpha_3| << |\alpha_2|$ with α_1 and α_3 positive, the previous inequality will not hold and only the first equation is valid, corresponding to minima and maxima alternating for $\omega = \pm k\pi/N$, $k = 0, 1, 2, \dots$

(d) Substituting $\omega = \pm k\pi/N$ into the magnitude expression gives:

$$|H(e^{j\omega})|_{\omega=\pm k\pi/N}^2 = [\alpha_2 \pm (\alpha_1 + \alpha_3)]^2$$

The maxima correspond to $[\alpha_2 + \alpha_1 + \alpha_3]^2$ and the minima correspond to $[\alpha_2 - \alpha_1 - \alpha_3]^2$. Therefore the peak-to-peak ripple is

$$R_A = 10 \log_{10} \left\{ \frac{[\alpha_2 + \alpha_1 + \alpha_3]^2}{[\alpha_2 - \alpha_1 - \alpha_3]^2} \right\} = 20 \log_{10} \left\{ \frac{[\alpha_2 + \alpha_1 + \alpha_3]}{[\alpha_2 - \alpha_1 - \alpha_3]} \right\}$$

(e) (i)

$$\alpha_1 = 0.1, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.2 \quad R_A = 20 \log \left\{ \frac{1.3}{0.7} \right\} = 5.38$$

(ii)

$$\alpha_1 = 0.15, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.15 \quad R_A = 20 \log \left\{ \frac{1.3}{0.7} \right\} = 5.38$$

(iii)

$$\alpha_1 = 0.1, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.1 \quad R_A = 20 \log \left\{ \frac{1.2}{0.8} \right\} = 3.52$$

(f) $\cos(\omega_0 N) = -(\alpha_1 + \alpha_3)/\alpha_2$ for $\omega = \pm\omega_0$. Note

$$\sin^2(\omega_0 N) = 1 - \cos^2(\omega_0 N) = 1 - \frac{(\alpha_1 + \alpha_3)^2}{\alpha_2^2}$$

$$\sin(\omega_0 N) = \pm \left[\frac{\alpha_2^2 - (\alpha_1 + \alpha_3)^2}{\alpha_2^2} \right]^{1/2}$$

Substitute for $\sin(\omega N)$ and $\cos(\omega N)$ in the expression for the phase to obtain:

$$\begin{aligned}\theta(\omega_0) &= -\omega_0 N \pm \tan^{-1} \left\{ \frac{(\alpha_1 - \alpha_3) \frac{(\alpha_2^2 - (\alpha_1 + \alpha_3)^2)^{1/2}}{\alpha_2}}{\alpha_2 - \frac{(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_3)}{\alpha_2}} \right\} \\ &= -\omega_0 N \pm \tan^{-1} \left\{ \frac{(\alpha_1 - \alpha_3) (\alpha_2^2 - (\alpha_1 + \alpha_3)^2)^{1/2}}{\alpha_2^2 - (\alpha_1 + \alpha_3)^2} \right\} \\ &= -\omega_0 N \pm \tan^{-1} \left\{ \frac{(\alpha_1 - \alpha_3)}{(\alpha_2^2 - (\alpha_1 + \alpha_3)^2)^{1/2}} \right\}\end{aligned}$$

The minima and maxima occur at $\omega = \pm\omega_0$ and the difference in phase between maxima and minima is therefore

$$R_p = 2 \tan^{-1} \left\{ \frac{\alpha_1 - \alpha_3}{(\alpha_2^2 - (\alpha_1 + \alpha_3)^2)^{1/2}} \right\}$$

(g) (i)

$$\alpha_1 = 0.1, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.2 \quad R_p = 2 \tan^{-1} \left[\frac{0}{(1 - 0.09)^{1/2}} \right] = 0 \text{ deg}$$

(ii)

$$\alpha_1 = 0.15, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.15 \quad R_p = 2 \tan^{-1} \left[\frac{-0.1}{(1 - 0.09)^{1/2}} \right] = -13.8 \text{ deg}$$

(iii)

$$\alpha_1 = 0.1, \quad \alpha_2 = 1.0, \quad \alpha_3 = 0.1 \quad R_p = 2 \tan^{-1} \left[\frac{0}{(1 - 0.04)^{1/2}} \right] = 0 \text{ deg}$$

The ratio of $\alpha_2/(\alpha_1 + \alpha_3)$ is the important parameter in determining the amplitude ripple and therefore only the sum of $\alpha_1 + \alpha_3$ is important. On the other hand, the difference between α_1 and α_3 can be chosen to provide a specified amplitude ripple and zero phase ripple.

7.22 (a) The coefficients, $\tilde{X}[k]$ must be samples of the Fourier transform of $h_v[n]$, i.e.,

$$\tilde{X}[k] = H_v(e^{j2\pi k/N_p})$$

where

$$H_v(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h_v[n] e^{-j\omega n}$$

(b)

$$\tilde{X}[n](e^{j\omega}) = \sum_{m=-\infty}^{\infty} \tilde{x}[m]w[n-m]e^{-j\omega m}$$

but

$$\begin{aligned}\tilde{x}[n] &= \frac{1}{N_p} \sum_{k=0}^{N_p-1} \tilde{X}[k] e^{j2\pi kn/N_p} \\ &= \frac{1}{N_p} \sum_{k=0}^{N_p-1} H_v(e^{j2\pi k/N_p}) e^{j2\pi kn/N_p}\end{aligned}$$

Substituting for $\tilde{x}[n]$ we get

$$\begin{aligned}\tilde{X}[n] &= \sum_{m=-\infty}^{\infty} \left[\frac{1}{N_p} \sum_{k=0}^{N_p-1} H_v(e^{j2\pi k/N_p}) e^{j2\pi kn/N_p} \right] w[n-m] e^{-j\omega m} \\ &= \frac{1}{N_p} \sum_{k=0}^{N_p-1} H_v(e^{j2\pi k/N_p}) \sum_{m=-\infty}^{\infty} w[n-m] e^{-j(\omega - 2\pi k/N_p)m} \\ &= \frac{1}{N_p} \sum_{k=0}^{N_p-1} H_v(e^{j2\pi k/N_p}) W_n(e^{j(\omega - 2\pi k/N_p)})\end{aligned}$$

(c) The sequence $\tilde{x}[n]$ is periodic with period N_p . Therefore $X_n(e^{j\omega})$ will be periodic in n with period N_p , and there will be N_p different values for a given ω .

(d)

$$\begin{aligned}w[n] &= \begin{cases} 1 & 0 \leq n \leq N_p - 1 \\ 0 & \text{otherwise} \end{cases} \\ W_n(e^{j\omega}) &= \sum_{m=-\infty}^{\infty} w[n-m] e^{-j\omega m} = \sum_{m=n-N_p+1}^n e^{-j\omega m}\end{aligned}$$

Let $m = -m' + n$ giving

$$\begin{aligned}W_n(e^{j\omega}) &= \sum_{m'=N_p-1}^0 e^{j\omega m'} e^{-j\omega n} \\ &= \sum_{m'=0}^{N_p-1} e^{j\omega m'} e^{-j\omega n} \\ &= e^{-j\omega n} \sum_{m'=0}^{N_p-1} e^{j\omega m'}\end{aligned}$$

$$\begin{aligned}W_n(e^{j\omega}) &= e^{-j\omega n} \left[\frac{1 - e^{j\omega N_p}}{1 - e^{j\omega}} \right] \\ &= e^{-j\omega n} \frac{e^{j\omega N_p/2}}{e^{j\omega/2}} \left[\frac{e^{j\omega N_p/2} - e^{-j\omega N_p/2}}{e^{j\omega/2} - e^{-j\omega/2}} \right] \\ &= e^{-j\omega n} e^{j\omega(N_p-1)/2} \frac{\sin(\omega N_p/2)}{\sin(\omega/2)}\end{aligned}$$

The magnitude $|W_n(e^{j\omega})|$ has the form shown in Figure P7.22.1.

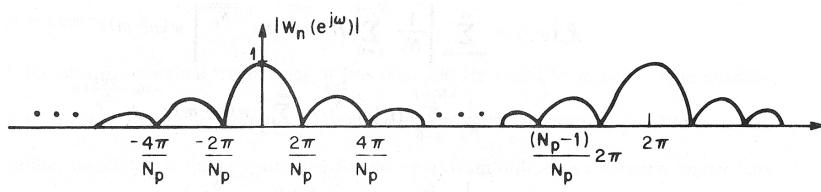


Figure P7.22.1:

- (e) The expression obtained in part (b) shows that $\tilde{X}(e^{j2\pi k/N_p})$ is obtained from the circular convolution of $H_v(e^{j2\pi k/N_p})$ and $W_n(e^{j2\pi k/N_p})$. From the sketch of Figure P7.22.1 we see that $W_n(e^{j2\pi k/N_p})$ will appear as an impulse in the convolution provided the phase of $W_n(e^{j\omega})$ is zero; i.e.,

$$e^{-j\omega n} e^{j\omega(N_p-1)/2} = 1$$

or

$$-n + \frac{N_p - 1}{2} = 0$$

$$n = \frac{N_p - 1}{2} \implies N_p \text{ odd}$$

7.23 (a)

$$\begin{aligned} a_n(\hat{\omega}_k) &= [x[n] \cos(\hat{\omega}_k n)] * h[n] \\ &= \sum_{m=-\infty}^{\infty} x[m] \cos(\hat{\omega}_k m) h[n-m] \\ b_n(\hat{\omega}_k) &= [x[n] \sin(\hat{\omega}_k n)] * h[n] \\ &= \sum_{m=-\infty}^{\infty} x[m] \sin(\hat{\omega}_k m) h[n-m] \end{aligned}$$

- (b) $H(e^{j\omega})$ represents a lowpass filter. Let $A_{\hat{\omega}_k}(e^{j\omega})$ and $B_{\hat{\omega}_k}(e^{j\omega})$ denote the Fourier transforms of $a_n(\hat{\omega}_k)$ and $b_n(\hat{\omega}_k)$. Then

$$A_{\hat{\omega}_k}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x[m] \left[\frac{e^{j\hat{\omega}_k m} + e^{-j\hat{\omega}_k m}}{2} \right] h[n-m] e^{-j\omega n}$$

Let $n = n' + m$ giving:

$$A_{\hat{\omega}_k}(e^{j\omega}) = \sum_{n'=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x[m] \left[\frac{e^{j\hat{\omega}_k m} + e^{-j\hat{\omega}_k m}}{2} \right] h[n'] e^{-j\omega n'} e^{-j\omega m}$$

$$\begin{aligned} A_{\hat{\omega}_k}(e^{j\omega}) &= \sum_{m=-\infty}^{\infty} \left[x[m] \frac{e^{-j(\omega-\hat{\omega}_k)m}}{2} + x[m] \frac{e^{-j(\omega+\hat{\omega}_k)m}}{2} \right] \sum_{n'=-\infty}^{\infty} h[n'] e^{-j\omega n'} \\ &= \frac{1}{2} \left[X(e^{j(\omega-\hat{\omega}_k)}) + X(e^{j(\omega+\hat{\omega}_k)}) \right] H(e^{j\omega}) \end{aligned}$$

But $x[n] = \cos(\omega_0 n) \longleftrightarrow X(e^{j\omega}) = (1/2)[\delta(\omega - \omega_0 T) + \delta(\omega + \omega_0 T)]$

$$A_{\hat{\omega}_k}(e^{j\omega}) = \frac{1}{2} \left[\frac{1}{2}\delta[\omega - \omega_0 \hat{\omega}_k] + \frac{1}{2}\delta[\omega + \omega_0 - \hat{\omega}_k] + \frac{1}{2}\delta[\omega - \omega_0 + \hat{\omega}_k] + \frac{1}{2}\delta[\omega + \omega_0 + \hat{\omega}_k] \right] H(e^{j\omega})$$

Similarly

$$B_{\hat{\omega}_k}(e^{j\omega}) = \frac{1}{2j} \left[X(e^{j(\omega - \hat{\omega}_k)}) - X(e^{j(\omega + \hat{\omega}_k)}) \right] H(e^{j\omega})$$

$$B_{\hat{\omega}_k}(e^{j\omega}) = \frac{1}{4} [\delta[\omega - \omega_0 - \hat{\omega}_k] + \delta[\omega + \omega_0 - \hat{\omega}_k] - \delta[\omega - \omega_0 + \hat{\omega}_k] - \delta[\omega + \omega_0 + \hat{\omega}_k]] H(e^{j\omega})$$

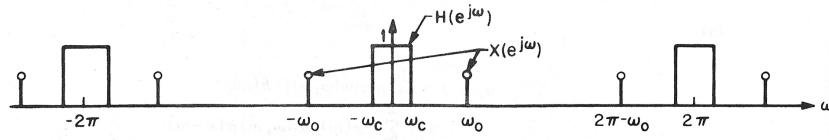


Figure P7.23.1:

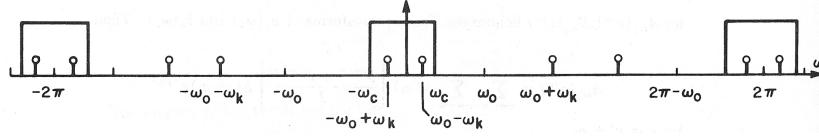


Figure P7.23.2:

A graphical interpretation can be given to the equation for $A_{\hat{\omega}_k}(e^{j\omega})$ as shown in Figure P7.23.1 and P7.23.2. Fig. P7.23.1 shows plots of $H(e^{j\omega})$ and $X(e^{j\omega})$ on the same coordinate system. Modulating $\cos(\omega_0 n)$ by $\cos(\hat{\omega}_k n)$ splits each of the line spectra for $X(e^{j\omega})$ and shifts them to $\omega_0 + \hat{\omega}_k$ and $\omega_0 - \hat{\omega}_k$ as shown in Fig. P7.23.2. The product of $H(e^{j\omega})$ with the line spectra represents $A_{\hat{\omega}_k}(e^{j\omega})$. Therefore

$$A_{\hat{\omega}_k}(e^{j\omega}) == \frac{1}{4} [\delta[\omega + \omega_0 - \hat{\omega}_k] + \delta[\omega - \omega_0 + \hat{\omega}_k]]$$

Similarly

$$B_{\hat{\omega}_k}(e^{j\omega}) = \frac{1}{4j} [\delta[\omega + \omega_0 - \hat{\omega}_k] - \delta[\omega - \omega_0 + \hat{\omega}_k]]$$

where we have assumed that no aliasing occurs. Inverse transforming gives

$$a_n(\hat{\omega}_k) = \frac{1}{2} \cos[(\omega_0 - \hat{\omega}_k)n]$$

$$b_n(\hat{\omega}_k) = -\frac{1}{2} \sin[(\omega_0 - \hat{\omega}_k)n]$$

(c)

$$M_n(\hat{\omega}_k) = [a_n^2(\hat{\omega}_k) + b_n^2(\hat{\omega}_k)]^{1/2} = \left[\frac{1}{4} \cos^2[(\omega_0 - \hat{\omega}_k)n] + \frac{1}{4} \sin^2[(\omega_0 - \hat{\omega}_k)n] \right]^{1/2} = \frac{1}{2}$$

The phase derivative, $\dot{\phi}_n(\hat{\omega}_k)$ is obtained by sampling the continuous-time phase derivative given by:

$$\begin{aligned}\dot{\phi}(t, \hat{\Omega}_k) &= \frac{d}{dt} \left[-\tan^{-1} \left\{ \frac{-\sin(\Omega_0 - \hat{\Omega}_k)t}{\cos(\Omega_0 - \hat{\Omega}_k)t} \right\} \right] \\ &= \frac{d}{dt} [(\Omega_0 - \hat{\Omega}_k)t] = \Omega_0 - \hat{\Omega}_k \\ \therefore \dot{\phi}_n(\hat{\omega}_k) &= \omega_0 - \hat{\omega}_k\end{aligned}$$

(d)

$$\begin{aligned}y_k[n] &= a_n(\hat{\omega}_k) \cos(\hat{\omega}_k n) + b_n(\hat{\omega}_k) \sin(\hat{\omega}_k n) \\ &= \frac{1}{2} \cos[(\omega_0 - \hat{\omega}_k)n] \cos(\hat{\omega}_k n) - \frac{1}{2} \sin[(\omega_0 - \hat{\omega}_k)n] \sin(\hat{\omega}_k n)\end{aligned}$$

Using the identities:

$$\cos[(\omega_0 \hat{\omega}_k)n] = \cos(\omega_0 n) \cos(\hat{\omega}_k n) + \sin(\omega_0 n) \sin(\hat{\omega}_k n)$$

$$\sin[(\omega_0 - \hat{\omega}_k)n] = \sin(\omega_0 n) \cos(\hat{\omega}_k n) + \sin(\hat{\omega}_k n) \cos(\omega_0 n)$$

we substitute them into the equations giving

$$\begin{aligned}y_k[n] &= \frac{1}{2} [\cos(\omega_0 n) \cos^2(\hat{\omega}_k n) + \sin(\omega_0 n) \sin(\hat{\omega}_k n) \cos(\hat{\omega}_k n)] \\ &\quad - \frac{1}{2} [\sin(\omega_0 n) \sin(\hat{\omega}_k n) \cos(\hat{\omega}_k n) - \cos(\omega_0 n) \sin^2(\hat{\omega}_k n)] \\ &= \frac{1}{2} \cos(\omega_0 n) [\cos^2(\hat{\omega}_k n) + \sin^2(\hat{\omega}_k n)] \\ &= \frac{1}{2} \cos(\omega_0 n)\end{aligned}$$

(e)

$$\dot{\phi}_n(\hat{\omega}_k) = \frac{b_n(\hat{\omega}_k) \dot{a}_n(\hat{\omega}_k) - a_n(\hat{\omega}_k) \dot{b}_n(\hat{\omega}_k)}{[a_n(\hat{\omega}_k)]^2 + [b_n(\hat{\omega}_k)]^2}$$

From part (b), assuming $T = 1$, we substitute $\omega \rightarrow \Omega$, $n \rightarrow t$ to obtain:

$$a_n(\hat{\Omega}_k) = \frac{1}{2} \cos[(\Omega_0 - \hat{\Omega}_k)t] \rightarrow \dot{a}_n(\hat{\Omega}_k) = \frac{-(\Omega_0 - \hat{\Omega}_k)}{2} \sin[(\Omega_0 - \hat{\Omega}_k)t]$$

$$b_n(\hat{\Omega}_k) = -\frac{1}{2} \sin[(\Omega_0 - \hat{\Omega}_k)t] \rightarrow \dot{b}_n(\hat{\Omega}_k) = \frac{-(\Omega_0 - \hat{\Omega}_k)}{2} \cos[(\Omega_0 - \hat{\Omega}_k)t]$$

Sampling these and substituting into the expression for $\dot{\phi}_n(\hat{\omega}_k)$ gives

$$\begin{aligned}\dot{\phi}_n(\hat{\omega}_k) &= \frac{\frac{1}{4}(\omega_0 - \hat{\omega}_k) \{ \sin^2[(\omega_0 - \hat{\omega}_k)n] + \cos^2[(\omega_0 - \hat{\omega}_k)n] \}}{\frac{1}{4} \{ \cos^2[(\omega_0 - \hat{\omega}_k)n] + \sin^2[(\omega_0 - \hat{\omega}_k)n] \}} \\ &= \dot{\phi}_n(\hat{\omega}_k) = (\omega_0 - \hat{\omega}_k)\end{aligned}$$

(f)

$$\dot{a}_n(\hat{\omega}_k) \approx \frac{1}{T} (a_n(\hat{\omega}_k) - a_{n-1}(\hat{\omega}_k))$$

Assume $T = 1$ for convenience, giving

$$\begin{aligned} b_n(\hat{\omega}_k) \dot{a}_n(\hat{\omega}_k) &= b_n(\hat{\omega}_k) a_n(\hat{\omega}_k) - b_n(\hat{\omega}_k) a_{n-1}(\hat{\omega}_k) \\ a_n(\hat{\omega}_k) \dot{b}_n(\hat{\omega}_k) &= a_n(\hat{\omega}_k) b_n(\hat{\omega}_k) - a_n(\hat{\omega}_k) b_{n-1}(\hat{\omega}_k) \\ \dot{\phi}_n(\hat{\omega}_k) &= \frac{-b_n(\hat{\omega}_k) a_{n-1}(\hat{\omega}_k) + a_n(\hat{\omega}_k) b_{n-1}(\hat{\omega}_k)}{[a_n(\hat{\omega}_k)]^2 + [b_n(\hat{\omega}_k)]^2} \end{aligned}$$

Note that

$$a_{n-1}(\hat{\omega}_k) = \frac{1}{2} \cos[(\omega_0 - \hat{\omega}_k)(n-1)] = \frac{1}{2} (\cos[(\omega_0 - \hat{\omega}_k)n] \cos[(\omega_0 - \hat{\omega}_k)] + \sin[(\omega_0 - \hat{\omega}_k)n] \sin[(\omega_0 - \hat{\omega}_k)])$$

and

$$b_{n-1}(\hat{\omega}_k) = -\frac{1}{2} (\sin[(\omega_0 - \hat{\omega}_k)n] \cos[(\omega_0 - \hat{\omega}_k)] - \sin[(\omega_0 - \hat{\omega}_k)] \cos[(\omega_0 - \hat{\omega}_k)n])$$

Substitution and cancelling yields:

$$\begin{aligned} \dot{\phi}_n(\hat{\omega}_k) &= \sin(\omega_0 - \hat{\omega}_k) [\cos^2[(\omega_0 - \hat{\omega}_k)n] + \sin^2[(\omega_0 - \hat{\omega}_k)n]] \\ \dot{\phi}_n(\hat{\omega}_k) &= \sin(\omega_0 - \hat{\omega}_k) \end{aligned}$$

This approximation is valid for $|\omega_0 - \hat{\omega}_k|$ small, since $\sin(\omega_0 - \hat{\omega}_k) \approx (\omega_0 - \hat{\omega}_k)$ for small arguments.

Chapter 8

The Cepstrum and Homomorphic Speech Processing

8.1

$$\begin{aligned}\hat{X}(e^{j\omega}) &= \log |X(e^{j\omega})| + j \arg[X(e^{j\omega})] \\ c[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |X(e^{j\omega})| e^{j\omega n} d\omega \\ c[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |X(e^{j\omega})| [\cos(\omega n) + j \sin(\omega n)] d\omega\end{aligned}$$

Recall that for $x[n]$ real, $|X(e^{j\omega})|$ is an even function; therefore

$$\int_{-\pi}^{\pi} \log |X(e^{j\omega})| (j \sin(\omega n)) d\omega = 0$$

with the result

$$c[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |X(e^{j\omega})| \cos(\omega n) d\omega$$

Inverse transforming $\hat{X}(e^{j\omega})$ we obtain

$$\hat{x}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\log |X(e^{j\omega})| + j \arg\{X(e^{j\omega})\}] [\cos(\omega n) + j \sin(\omega n)] d\omega$$

Recall that for $x[n]$ real, $\arg\{X(e^{j\omega})\}$ is an odd function, therefore

$$\int_{-\pi}^{\pi} j \arg\{X(e^{j\omega})\} \cos(\omega n) d\omega = 0$$

with the result

$$\hat{x}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |X(e^{j\omega})| \cos(\omega n) d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \arg\{X(e^{j\omega})\} \sin(\omega n) d\omega$$

and

$$\begin{aligned}\hat{x}[-n] &= \int_{-\pi}^{\pi} \log |X(e^{j\omega})| \cos(\omega n) d\omega + \int_{-\pi}^{\pi} \arg\{X(e^{j\omega})\} \sin(\omega n) d\omega \\ \therefore \frac{\hat{x}[-n] + \hat{x}[n]}{2} &= \int_{-\pi}^{\pi} \log |X(e^{j\omega})| \cos(\omega n) d\omega = c[n]\end{aligned}$$

8.2 (a) We can write $H(z)$ in the normalized format:

$$H(z) = 8 \left[\frac{1 - 4z^{-1}}{1 - \frac{1}{6}z^{-1}} \right] = 32z^{-1} \left[\frac{1 - z/4}{1 - z^{-1}/6} \right]$$

We now can form the log of $H(z)$ as:

$$\hat{H}(z) = \log |H(z)| = \log(32) + \log[-z^{-1}] + \log(1 - z/4) - \log(1 - z^{-1}/6)$$

where the log terms for the minus sign and the z^{-1} term have no effect and are omitted from subsequent computation. Recognizing the log series:

$$\log(1 - x) = \sum_{n=1}^{\infty} \frac{x^n}{n}, \quad |x| < 1$$

we get:

$$\hat{H}(z) = \log(32) + \sum_{n=-\infty}^{-1} \frac{(1/4)^{-n}z^n}{n} - \sum_{n=1}^{\infty} \frac{(1/6)^n z^{-n}}{n}$$

giving:

$$\hat{h}[n] = \begin{cases} \log(32) & n = 0 \\ \frac{(1/6)^n}{n} & n > 0 \\ \frac{(1/4)^{-n}}{n} & n < 0 \end{cases}$$

(b) Figure P8.2.1 (at the top) shows a plot of $\hat{h}[n]$. By setting $\hat{h}[n]$ to a value of zero, we can better see the behavior of $\hat{h}[n]$ for the range $-10 \leq n \leq n$ as shown in Figure P8.2.1 at the bottom.

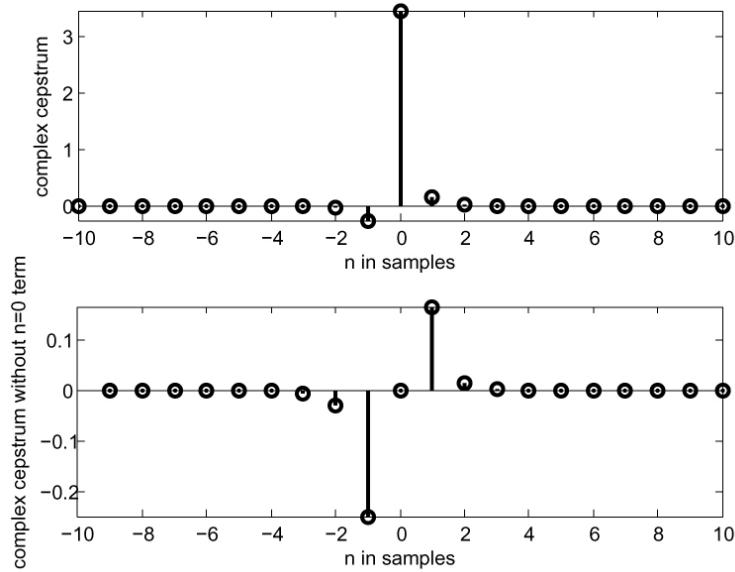


Figure P8.2.1: Plot of complex cepstral sequence; top panel with all cepstral values, bottom panel without cepstral value at $n=0$.

(c) We can solve for $c[n]$ as:

$$\begin{aligned} c[n] &= \frac{\hat{h}[n] + \hat{h}[-n]}{2} \\ &= \begin{cases} \log(32) & n = 0 \\ \frac{(1/6)^n}{2n} - \frac{(1/4)^n}{2n} & n > 0 \\ -\frac{(1/6)^{-n}}{2n} + \frac{(1/4)^{-n}}{2n} & n < 0 \end{cases} \end{aligned}$$

8.3 From the log power series expansion of each complex pole we get

$$\hat{v}[n] = \begin{cases} 0 & n = 0 \\ \sum_{k=1}^q \left[\frac{c_k^n}{n} + \frac{(c_k^*)^n}{n} \right] & n > 0 \\ 0 & n < 0 \end{cases}$$

$$\begin{aligned} \hat{v}[n] &= \sum_{k=1}^q \left[\frac{(r_k e^{j\theta_k})^n}{n} + \frac{(r_k e^{-j\theta_k})^n}{n} \right] & n > 0 \\ &= 2 \sum_{k=1}^q \frac{r_k^n}{n} \cos(\theta_k n) & n > 0 \end{aligned}$$

8.4 The $z-$ transform of the complex cepstrum corresponding to $H^{-1}(z) = 1/H(z)$ is:

$$\begin{aligned} H^{-1}(z) &= \log \left\{ \frac{1 - \sum_{k=1}^p \alpha_k z^{-k}}{G} \right\} \\ &= \log \left[1 - \sum_{k=1}^p \alpha_k z^{-k} \right] - \log[G] \end{aligned}$$

and the $z-$ transform of the complex cepstrum $\hat{h}[n]$ is:

$$\hat{H}(z) = -\log[1 - \sum_{k=1}^p \alpha_k z^{-k}] + \log[G] = -H^{-1}(z)$$

Since $1/H(z)$ is minimum phase, the recursion relation can be applied to obtain:

$$-\hat{h}[n] = \begin{cases} -\log[G] & n = 0 \\ G\alpha_n - \sum_{k=0}^{n-1} \left(\frac{k}{n} \right) \hat{h}[k]G\alpha_{n-k} & n > 0 \end{cases}$$

or, equivalently:

$$\hat{h}[n] = \begin{cases} \log[G] & n = 0 \\ \sum_{k=0}^{n-1} \left(\frac{k}{n}\right) \hat{h}[k] G \alpha_{n-k} - G \alpha_n & n > 0 \end{cases}$$

8.5 (a)

$$Y(z) = \sum_{n=0}^{N-1} \alpha^n x[n] z^{-n} = \sum_{n=0}^{N-1} x[n] (\alpha z^{-1})^n$$

$$Y(z) = X\left(\frac{z}{\alpha}\right)$$

If the zeros of $X(z)$ occur at z_0, z_1, \dots, z_{N-1} then the zeros of $Y(z)$ occur at $\alpha z_0, \alpha z_1, \dots, \alpha z_{N-1}$. Therefore $\hat{y}[n] = \alpha \hat{x}[n]$.

- (b) Let $\max_i |z_i|$ denote the magnitude of the maximum magnitude zero of $X(z)$. $y[n]$ is not minimum phase if any zero of $Y(z)$ lies on or outside the unit circle. Therefore choose α so that:

$$|\alpha| \max_i |z_i| \geq 1 \implies |\alpha| \geq \frac{1}{\max_i |z_i|}$$

- (c) Let $\min_i |z_i|$ denote the magnitude of the maximum magnitude zero of $X(z)$. $y[n]$ is not maximum phase if all the zeros of $Y(z)$ lie outside the unit circle. Therefore choose α so that:

$$|\alpha| \min_i |z_i| \geq 1 \implies |\alpha| > \frac{1}{\min_i |z_i|}$$

8.6

The z -transform of $x[n]$ is:

$$X(z) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}$$

The z -transform of $x[-n]$ is:

$$\sum_{n=-\infty}^{\infty} x[-n] z^{-n} = \sum_{n=\infty}^{-\infty} x[n] z^n = X(z^{-1})$$

$x[n]$ is minimum phase which means that all poles and zeros are inside the unit circle and we can express $X(z)$ as:

$$X(z) = \frac{\prod_{k=1}^M (1 - a_k z^{-1})}{\prod_{k=1}^N (1 - c_k z^{-1})}$$

where $|a_k| < 1$ and $|c_k| < 1$. Then the z -transform of $x[-n]$ has the form:

$$X(z^{-1}) = \frac{\prod_{k=1}^M (1 - a_k z)}{\prod_{k=1}^N (1 - c_k z)}$$

The zeros occur at $1/a_k$ and the poles occur at $1/c_k$; therefore $x[-n]$ is maximum phase.

8.7 (a)

$$\tilde{X}[k] = \hat{X}(e^{j\frac{2\pi}{N}k}) = \sum_{m=-\infty}^{\infty} \hat{x}[m] e^{-j2\pi km/N} \quad 0 \leq k \leq N-1$$

(b)

$$\begin{aligned} \tilde{x}[n] &= \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}[k] e^{j2\pi kn/N} \quad 0 \leq n \leq N-1 \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left[\sum_{m=-\infty}^{\infty} \hat{x}[m] e^{-j2\pi km/N} \right] e^{j2\pi kn/N} \\ &= \sum_{m=-\infty}^{\infty} \frac{\hat{x}[m]}{N} \sum_{k=0}^{N-1} e^{j2\pi k(n-m)} \end{aligned}$$

The geometric series in the equation above can be summed to give

$$\frac{1}{N} \sum_{k=0}^{N-1} e^{j2\pi k(n-m)} = \begin{cases} 1 & (n-m) = rN, \quad r \text{ integer} \\ 0 & \text{otherwise} \end{cases}$$

Recognizing that $n - m = rN \implies m = n + rN$ and that $m = -\infty \implies r = -\infty$ and $m = \infty \implies r = \infty$, gives us:

$$\tilde{x}[n] = \sum_{r=-\infty}^{\infty} \hat{x}[n + rN]$$

8.8 (a)

$$X(z) = 1 + \alpha z^{-N_p}$$

$$\hat{X}(z) = \log[X(z)] = \log[1 + \alpha z^{-N_p}]$$

Using the log power series ($\log(1+u) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} u^n$, $|u| < 1$) we get:

$$\begin{aligned} \hat{X}(z) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\alpha z^{-N_p})^n; \quad |\alpha z^{-N_p}| \leq 1 \\ &= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\alpha^n}{n} z^{-nN_p}, \quad |\alpha|^{1/N_p} < |z| \end{aligned}$$

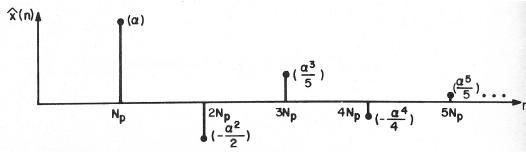


Figure P8.8.1: Plot of complex cepstrum for impulse plus delayed impulse signal.

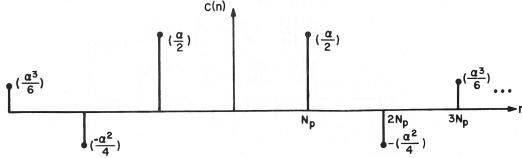


Figure P8.8.2: Plot of (real) cepstrum for impulse plus delayed impulse signal.

By looking at the coefficients in this power series, it is easily seen that

$$\hat{x}[n] = \sum_{r=1}^{\infty} \frac{(-1)^{r+1} \alpha^r}{r} \delta[n - rN_p]$$

and the sequence of cepstral values is plotted in Figure P8.8.1.

- (b)** The formula for deriving the cepstrum from the complex cepstrum is:

$$c[n] = \frac{\hat{x}[n] + \hat{x}[-n]}{2}$$

$$c[n] = \sum_{r=-\infty, r \neq 0}^{\infty} (-1)^{r+1} \frac{\alpha^{|r|}}{1|r|} \delta[n - rN_p]$$

which is shown plotted in Figure P8.8.2.

- (c)** The aliased cepstrum can be represented as

$$\tilde{x}[n] = \sum_{k=-\infty}^{\infty} \hat{x}[n + kN]$$

Based on the results of part (a) we can substitute for $\hat{x}[n + kN]$ giving

$$\tilde{x}[n] = \sum_{k=-\infty}^{\infty} \left[-\sum_{r=1}^{\infty} \frac{(-\alpha)^r \delta[n + kN - rN_p]}{r} \right]$$

Note that $\hat{x}[n] \neq 0$ only if $\delta[n + kN - rN_p] \neq 0$ or equivalently $kN - rN_p = n$. For the case $N_p = N/6$ we get

$$6k - r = \frac{n}{N_p} = \text{integer}$$

The approximation (aliased) complex cepstrum, $\tilde{x}[n]$ is periodic with a period of N samples. The approximation will have the same appearance as $\hat{x}[n]$ for the case $N_p = N/6$. However the amplitudes of the samples will be increased due to aliasing, and, as a result, $\tilde{x}[0]$ will be

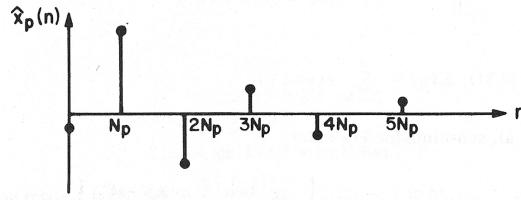


Figure P8.8.3: Plot of aliased complex cepstrum for impulse plus delayed impulse signal.

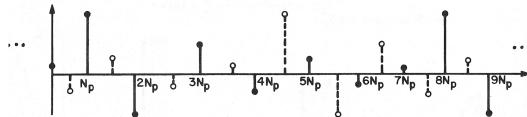


Figure P8.8.4: Plot of extraneous signal values due to aliasing.

non-zero. Specifically, for $N = 6N_p$ we get:

$$\begin{aligned}\tilde{x}[n] &= -[(\alpha^6/6 + \alpha^{12}/(12) + \dots)] \quad n = 0 \\ &= [\alpha + \alpha^7/7 + \dots] \quad n = N_p \\ &= -[\alpha^2/2 + \alpha^8/8 + \dots] \quad n = 2N_p \\ &= [\alpha^3/3 + \alpha^9/9 + \dots] \quad \alpha = 3N_p \\ &= -[\alpha^4/4 + \alpha^{10}/(10) + \dots] \quad n = 4N_p \\ &= [\alpha^5/5 + \alpha^{11}/(11) + \dots] \quad n = 5N_p\end{aligned}$$

where $\tilde{x}[n]$ is shown plotted in Figure P8.8.3. If N is not divisible by N_p , then as a result of aliasing, additional samples will appear with separation less than N_p . For example, for $N = (7/2)N_p$ we get the aliased complex cepstrum shown in Figure P8.8.4.

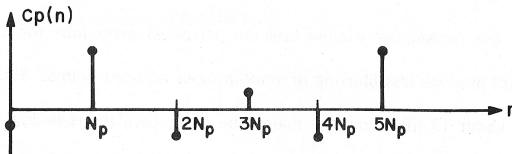


Figure P8.8.5: Plot of aliased complex cepstrum for impulse plus delayed impulse signal.

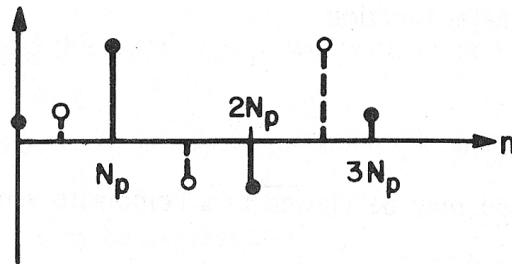


Figure P8.8.6: Plot of extraneous signal values due to aliasing.

- (d) The (real) cepstrum approximation (due to aliasing) is given by:

$$\tilde{c}[n] = \frac{\tilde{x}[n] + \tilde{x}[-n]}{2}$$

Therefore, from part (c), the cepstrum can be sketched and is shown in Figure P8.8.5 for the case $N = 6N_p$. When N is not divisible by N_p , then additional samples will again appear with separation less than N_p . Figure P8.8.6 shows the aliased real cepstrum for the case $N = (7/2)N_p$.

- (e) From part(d) it is easily seen that if $N < 2N_p$, the largest impulse will be to the right of its negative time image at $n - N - B_p$, and therefore confusion would result.

- 8.9** (a) Multiplication in the cepstral domain results in convolution of the Fourier transforms, giving:

$$C^{(y)}(e^{j\omega}) = \frac{1}{2\pi j} \int_{-\pi}^{\pi} C^{(x)}(e^{j\theta}) L(e^{j(\omega-\theta)}) d\theta$$

- (b) A low-time (quefrency) window is required to smooth $\log |X(e^{j\omega})|$.
- (c) The Fourier transform of the cepstral window is convolved with $\log |X(e^{j\omega})|$. The Fourier transform of the rectangular window has the narrowest main lobe for a given window length and will produce less blurring or smoothing of adjacent spectral points. Since the first side lobe is only about 13 dB below the main lobe peak, oscillation in $\log |X(e^{j\omega})|$ will be introduced at discontinuities of $\log |X(e^{j\omega})|$. The Hamming window has a wide main lobe resulting in greater smoothing of adjacent spectral points. The largest side lobe is approximately 43 dB below the main lobe peak.
- (d) The cepstral window should be shorter than the expected pitch period. The objective of the smoothing is to eliminate the excitation information and recover an approximation to the overall vocal tract transfer function.

- 8.10** The windowed excitation impulse train is

$$p_w[n] = w[n - n_0] \sum_{k=-\infty}^{\infty} \delta[n - kN_p] = \sum_{k=-\infty}^{\infty} w[kN_p - n_0] \delta[n - kN_p]$$

- (a) Figure P10.2.1 shows a MATLAB plot of $p[n]$ with $N_p = 100$ and shifted windows of length $L = 2N_p + 1 = 201$ for values of $n_0 = 0, N_p/4, N_p/2, 3N_p/4$. The windowed sequence $p_w[n]$ is the product of $p[n]$ and the shifted window for each case, and the picture would simply

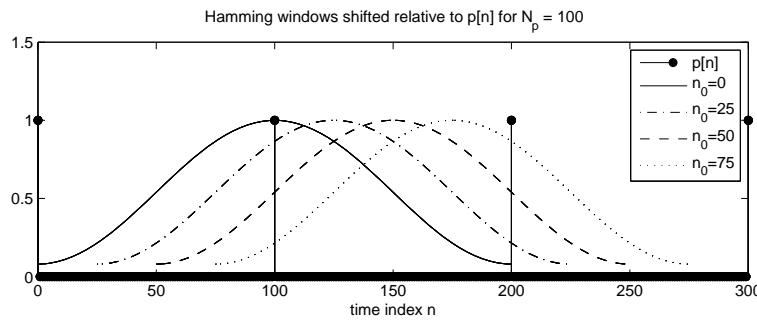


Figure P10.2.1: Plot of $p[n]$ with $N_p = 100$ and shifted Hamming windows

scale with N_p since for convenience we have made the window length proportional to N_p .

Specifically, the sequences are

$$\begin{array}{ll} n_0 = 0 & p_w[n] = .08\delta[n] + \delta[n - N_p] + .08\delta[n - 2N_p] \\ n_0 = N_p/4 & p_w[n] = .8654\delta[n - N_p] + .2147\delta[n - 2N_p] \\ n_0 = N_p/2 & p_w[n] = .54\delta[n - N_p] + .54\delta[n - 2N_p] \\ n_0 = 3N_p/4 & p_w[n] = .2147\delta[n - N_p] + .8654\delta[n - 2N_p] \end{array}$$

(b) The z -transforms of the above windowed impulse trains are:

$$\begin{array}{ll} n_0 = 0 & P_w(z) = .08z^{-N_p}(z^{N_p} + 12.5 + z^{-N_p}) \\ n_0 = N_p/4 & P_w(z) = .8654z^{-N_p}(1 + .2481z^{-N_p}) \\ n_0 = N_p/2 & P_w(z) = .54z^{-N_p}(1 + z^{-N_p}) \\ n_0 = 3N_p/4 & P_w(z) = .2147z^{-N_p}(1 + 4.0307z^{-N_p}) = .8654z^{-2N_p}(1 + .2481z^{N_p}) \end{array}$$

These are all polynomials in the variable z^{-N_p} so we can use the formulas for the cepstrum in Eqs. (8.24) to compute the complex cepstrum by assuming that the same polynomial structure with z^{-N_p} replaced by z^{-1} , finding the cepstrum, and then spacing the cepstrum values at corresponding multiples of N_p .

(c) To computed the complex cepstrum, we need to factor the polynomials in part (b).

(i) For $n_0 = 0$, the polynomial is really of second order and it has symmetric coefficients so it is linear phase. Thus, there will be one basic root of the polynomial inside the unit circle and the other one will be at the reciprocal location. This means that we can write it as:

$$P_w(z) = .08z^{-N_p}(z^{N_p} + 12.5 + z^{-N_p}) = -(0.08/\alpha)z^{-N_p}(1 - \alpha z^{N_p})(1 - \alpha z^{-N_p})$$

where we could use MATLAB to factor the polynomial [1,12.5,1] to get one root at $\alpha = -.0805186$ and the other one at $1/\alpha$. With this representation (ignoring the minus sign and the factor z^{-N_p} outside) we can use Eq. (8.24) to write

$$\hat{p}_w[n] = \begin{cases} -\sum_{k=1}^{\infty} \frac{\alpha^k}{k} \delta[n - kN_p] & n > 0 \\ \log|.08/\alpha| = -0.006462328 & n = 0 \\ -\sum_{k=1}^{\infty} \frac{\alpha^k}{k} \delta[n + kN_p] & n < 0 \end{cases}$$

where $\alpha = -.0805186$. Thus, this position of the window gives a mixed-phase complex cepstrum with non-zero values at multiples of N_p . The cepstrum peak will be $.0805186$ at $\pm N_p$. The cepstrum peak value is very small due to the small ratio between the end impulses and the central impulse.

(ii) For $n_0 = N_p/4$, the polynomial clearly is minimum-phase and the complex cepstrum will have impulses spaced by N_p . Specifically,

$$\hat{p}_w[n] = \begin{cases} -\sum_{k=1}^{\infty} \frac{(-.2481)^k}{k} \delta[n - kN_p] & n > 0 \\ \log(.8654) = -0.144679 & n = 0 \\ 0 & n < 0 \end{cases}$$

In this case, the first cepstrum peak at N_p has value $-(-.2481)^1/1 = .2481$.

(iii) For $n_0 = 3N_p/4$, the polynomial clearly is maximum-phase. In fact, this case is the time-reversed version of (ii). The complex cepstrum will have impulses spaced by N_p ,

but now on the left side of the plot of the complex cepstrum. Specifically,

$$\hat{p}_w[n] = \begin{cases} -\sum_{k=1}^{\infty} \frac{(-.2481)^k}{k} \delta[n + kN_p] & n < 0 \\ \log(.8654) = -0.144679 & n = 0 \\ 0 & n > 0 \end{cases}$$

In this case, the first cepstrum peak at $-N_p$ has value $-(-.2481)^1/1 = .2481$.

- (iv) For $n_0 = N_p/2$, the polynomial clearly has its roots on the unit circle. The basic first-order polynomial is $1 + z^{-1}$ which has its root at $z = -1$. Thus, Eq. (8.24) does not apply. However, if the second impulse were just slightly smaller than the first one (e.g., when $n_0 = N_p/2 - 1$), we would have minimum-phase and the cepstrum peak at N_p would be close to 1. Alternatively, if the window was positioned so that the second impulse in $p_w[n]$ were slightly larger than the first (e.g., when $n_0 = N_p/2 + 1$), we would have maximum-phase and the cepstrum peak at $-N_p$ would be close to 1.
- (d) (i) $p_w[n]$ will be minimum-phase for $0 < n_0 < N_p/2$.
(ii) $p_w[n]$ will be maximum-phase for $N_p/2 < n_0 < N_p$.
(iii) As n_0 starts at 0 and approaches $N_p/2$ the signal remains minimum-phase. Around $n_0 = 0$ the cepstrum peak is small and it grows to close to 1 when n_0 is close to $N_p/2$. Similarly as n_0 moves past $N_p/2$, the signal flips to maximum-phase and the cepstrum peak is about 1 but on the negative quefrency side. As n_0 approaches $3N_p/4$ the peak gets smaller and as it approaches N_p we are back to the condition $n_0 = 0$ where the cepstrum peak is the smallest.
(iv) The cepstrum peak is smallest for $n_0 = 0$.
- (e) With a longer window, more pitch impulses would be included and the cepstrum peak height would be less sensitive to the window starting point. If less than 2 pitch periods are included in the window, i.e., $L < 2N_p + 1$, the peak will be very sensitive to the window location. For example, for any shift other than $n_0 = 0$, there would be only one pitch impulse within the window, so no periodicity would be shown as a cepstrum peak.

Chapter 9

Linear Predictive Analysis of Speech Signals

9.1 (a) By definition the autocorrelation can be written as a convolution of the form:

$$\tilde{R}[m] = \sum_{n=0}^{\infty} h[n]h[n+m] = h[m] * h[-m]$$

Thus we can replace m by $-m$ giving

$$\tilde{R}[-m] = h[-m] * h[m]$$

Since convolution is commutative, we trivially have the result

$$\tilde{R}[m] = h[m] * h[-m] = h[-m] * h[m] = \tilde{R}[-m]$$

(b)

$$\tilde{R}[m] = \sum_{n=0}^{\infty} h[n]h[n-m]$$

$$\begin{aligned}\tilde{R}[m] &= \sum_{n=0}^{\infty} \left\{ \left[\sum_{k=1}^p \alpha_k h[n-k] + G\delta[n] \right] \left[\sum_{l=1}^p \alpha_l h[n-m-l] + G\delta[n-m] \right] \right\} \\ &= \sum_{n=0}^{\infty} \left\{ \sum_{k=1}^p \sum_{l=1}^p \alpha_k \alpha_l h[n-k] h[n-m-l] + G \sum_{k=1}^p \alpha_k h[n-k] \delta[n-m] \right. \\ &\quad \left. + G \sum_{l=1}^p \alpha_l h[n-m-l] \delta[n] + G^2 \delta[n] \delta[n-m] \right\}\end{aligned}$$

Assume $m \geq 0$ giving

$$\begin{aligned}\tilde{R}[m] &= \sum_{n=0}^{\infty} \sum_{k=1}^p \sum_{l=1}^p \alpha_k \alpha_l h[n-k] h[n-m-l] + G \sum_{k=1}^p \alpha_k h[n-k] \\ &\quad + G \sum_{l=1}^p \alpha_l h[n-m-l]\end{aligned}$$

Since, for $m \geq 0$, $h[-m-l] = 0$ for $l = 1, 2, \dots, p$, then

$$\begin{aligned}\tilde{R}[m] &= \sum_{n=0}^{\infty} \sum_{k=1}^p \alpha_k h[n-k] \left[\sum_{l=1}^p \alpha_l h[n-m-l] + G\delta[n-m] \right] \\ &= \sum_{k=1}^p \alpha_k \sum_{n=0}^{\infty} h[n-k] h[n-m]\end{aligned}$$

If we change variables to $n = n' + m$ we get

$$\tilde{R}[m] = \sum_{k=1}^p \alpha_k \sum_{n'=-m}^{\infty} h[n'+m-k] h[n']$$

Since $m \geq 0$ and $h[n']$ is causal, we change the lower limit of summation to zero to obtain

$$\tilde{R}[m] = \sum_{k=1}^p \alpha_k \tilde{R}[m-k]$$

For $m \leq 0$ we obtain, in similar fashion:

$$\tilde{R}[m] = \sum_{k=1}^p \alpha_k \tilde{R}[k+m]$$

where m is negative. Therefore

$$\tilde{R}[m] = \sum_{k=1}^p \alpha_k \tilde{R}[|m-k|], \quad m = 1, 2, \dots, p.$$

9.2 Define

$$A(e^{j2\pi k/N}) = 1 - \sum_{n=1}^p \alpha_n e^{-j2\pi kn/N}, \quad 0 \leq k \leq N-1$$

where $A(e^{j2\pi k/N})$ is the N -point FFT of the augmented sequence, $-\alpha_n$, of length $p+1$ samples, where $p+1 \leq N$ and $\alpha_0 = -1$. The FFT samples of $H(e^{j\omega})$ are obtained from the relation:

$$H(e^{j2\pi k/N}) = \frac{G}{A(e^{j2\pi k/N})} \quad 0 \leq k \leq N-1$$

9.3 (a) The mean and variance are calculated as:

$$\begin{aligned}\overline{x[n]} &= \overline{\epsilon[n] + \beta\epsilon[n-1]} = 0 \text{ (since } \epsilon[n] \text{ is a zero-mean process)} \\ \overline{x^2[n]} &= \overline{(\epsilon[n] + \beta\epsilon[n-1])^2} = \overline{\epsilon^2[n]} + 2\beta\overline{\epsilon[n] \cdot \epsilon[n-1]} + \beta^2\overline{\epsilon^2[n-1]} \\ &= \sigma_{\epsilon}^2 + \beta^2\sigma_{\epsilon}^2 = (1 + \beta^2)\sigma_{\epsilon}^2 = (1 + \beta^2)\end{aligned}$$

(b) The transfer function from $\epsilon[n]$ to $x[n]$ is:

$$H_1(z) = \frac{X(z)}{\mathcal{E}(z)} = 1 + \beta z^{-1}$$

To recover $\epsilon[n]$ we need to send $x[n]$ through the inverse system, of the form:

$$H_2(z) = \frac{\mathcal{E}(z)}{X(z)} = \frac{1}{1 + \beta z^{-1}}$$

(c) The correlation of $x[n]$ is:

$$\begin{aligned} R_x[k] &= E\{x[n]x[n+k]\} \\ &= E\{(\epsilon[n] + \beta\epsilon[n-1]) \cdot (\epsilon[n+k] + \beta\epsilon[n+k-1])\} \\ &= E\{\epsilon[n]\epsilon[n+k]\} + \beta E\{\epsilon[n-1]\epsilon[n+k] + \epsilon[n]\epsilon[n+k-1]\} \\ &\quad + \beta^2 E\{\epsilon[n-1]\epsilon[n+k-1]\} \\ &= \delta[k] + \beta\delta[k-1] + \beta\delta[k+1] + \beta^2\delta[k] \\ &= (1 + \beta^2)\delta[k] + \beta\delta[k-1] + \beta\delta[k+1] \end{aligned}$$

giving the final result:

$$r_x[1] = \frac{R_x[1]}{R_x[0]} = \frac{\beta}{1 + \beta^2}$$

9.4 (a)

$$\begin{aligned} \phi_n[i, k] &= \sum_{l=0}^{L-1} s_n[l-i]s_n[l-k] \\ \phi_n[i+1, k+1] &= \sum_{l=0}^{L-1} s_n[l-(i+1)]s_n[l-(k+1)] \\ \phi_n[i+1, k+1] &= \sum_{l=0}^{L-1} s_n[l-1-i]s_n[l-1-k] \end{aligned}$$

If we make the change of variables $l = l' + 1$ we get:

$$\begin{aligned} \phi_n(i+1, k+1) &= \sum_{l'=-1}^{L-2} s_n[l'-i]s_n[l'-k] \\ &= \sum_{l'=0}^{L-1} s_n[l'-i]s_n[l'-k] + s_n[-1-i]s_n[-1-k] \\ &\quad - s_n[L-1-i]s_n[L-1-k] \\ &= \phi_n[i, k] + s_n[-1-i]s_n[-1-k] - s_n[L-1-i]s_n[L-1-k] \end{aligned}$$

(b) From part (a) we have:

$$\begin{aligned} \phi_n[i+1, i+1] &= \phi_n[i, i] + s_n[-i-1]s_n[-i-1] - s_n(L-1-i)s_n[L-1-i] \\ \phi_n[1, 1] &= \phi_n[0, 0] + s_n^2[-1] - s_n^2[L-1] \\ \phi_n[2, 2] &= \phi_n[1, 1] + s_n^2[-2] - s_n^2[L-2] \\ &\quad \cdot \\ &\quad \cdot \\ &\quad \cdot \\ \phi_n[p, p] &= \phi_n[p-1, p-1] + s_n^2[-p-1] - s_n^2[L-p-1] \end{aligned}$$

- (c) Again from part (a) the elements along the $(i+1)^{th}$ lower diagonal are obtained recursively as:

$$\begin{aligned}\phi_n[i+j+1, j+1] &= \phi_n[i+j, j] + s_n[-i-j-1]s_n[-j-1] \\ -s_n[L-1-i-j]s_n[L-1-k] &\quad i = 0, \dots, p-1; \quad 0 \leq j \leq p-i-1\end{aligned}$$

- (d) Since the covariance matrix is symmetric, the upper diagonals are obtain from:

$$\phi_n[k, i] - \phi_n[i, k], \quad i > k$$

- 9.5** If we go through the math that we used to derive the standard LPC analysis, we see that we are basically partitioning the matrix equation for solving for the LPC coefficients into a sub-matrix equation using the set of rows and columns corresponding to $k_0 \leq i, j \leq k_1$. Hence, for the autocorrelation method we get the matrix equation:

$$\begin{bmatrix} R_n[0] & R_n[1] & \cdots & R_n[k_1 - k_0] \\ R_n[1] & R_n[0] & \cdots & R_n[k_1 - k_0 - 1] \\ \vdots & \vdots & \ddots & \vdots \\ R_n[k_1 - k_0] & R_n[k_1 - k_0 - 1] & \cdots & R_n[0] \end{bmatrix} \begin{bmatrix} \alpha_{k_0} \\ \alpha_{k_0+1} \\ \vdots \\ \alpha_{k_1} \end{bmatrix} = \begin{bmatrix} R_n[k_0] \\ R_n[k_0 + 1] \\ \vdots \\ R_n[k_1] \end{bmatrix}$$

For the covariance method we get the matrix equation:

$$\begin{bmatrix} \phi_n[k_0, k_0] & \phi_n[k_0, k_0 + 1] & \cdots & \phi_n[k_0, k_1] \\ \phi_n[k_0 + 1, k_0] & \phi_n[k_0 + 1, k_0 + 1] & \cdots & \phi_n[k_0 + 1, k_1] \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n[k_1, k_0] & \phi_n[k_1, k_0 + 1] & \cdots & \phi_n[k_1, k_1] \end{bmatrix} \begin{bmatrix} \alpha_{k_0} \\ \alpha_{k_0+1} \\ \vdots \\ \alpha_{k_1} \end{bmatrix} = \begin{bmatrix} \phi_n[k_0, 0] \\ \phi_n[k_0 + 1, 0] \\ \vdots \\ \phi_n[k_1, 0] \end{bmatrix}$$

- 9.6 (a)** From the 4th row of the table

$$A^{(4)}(z) = 1 - .8047z^{-1} + .0414z^{-2} - .494z^{-3} + .4337z^{-4}.$$

- (b) The k_i parameter is equal to the coefficient of z^{-i} in the polynomial $A^{(i)}(z)$, i.e., $k_i = \alpha_i^{(i)}$. Therefore, $k_1 = .8328$, $k_2 = .1044$, $k_3 = .1786$, and $k_4 = -.4337$.
- (c) Verify the Durbin algorithm step in polynomial form.

$$\begin{aligned}A^{(3)}(z) &= 1 - .7273z^{-1} + .0289z^{-2} - .1786z^{-3} \\ -k_4z^{-4}A^{(3)}(z^{-1}) &= -0.0775z^{-1} + .0125z^{-2} - .3254z^{-3} + .4337z^{-4} \\ \text{adding} \\ A^{(4)}(z) &= 1 - .8048z^{-1} + .0414z^{-2} - .4940z^{-3} + .4337z^{-4}\end{aligned}$$

- (d) Use the result $E^{(i)} = (1 - k_i^2)E^{(i-1)}$ to work backwards and forwards given that $E^{(2)} = .5803$. To get $E^{(4)}$, write it as

$$\begin{aligned}E^{(4)} &= (1 - k_4^2)E^{(3)} = (1 - k_4^2)(1 - k_3^2)E^{(2)} \\ &= (1 - .4337^2)(1 - .1786^2).5803 = .4561\end{aligned}$$

To determine $R[0]$ recall that $R[0] = E^{(0)}$. Therefore $E^{(2)} = (1 - k_2^2)(1 - k_1^2)R[0]$ so

$$R[0] = \frac{E^{(2)}}{(1 - k_2^2)(1 - k_1^2)} = \frac{.5803}{(1 - .1044^2)(1 - .8328^2)} = 1.9145$$

To determine $R[1]$, recall that $k_1 = R[1]/R[0]$ so

$$R[1] = k_1 R[0] = .8328(1.9145) = 1.5944.$$

- (e) Because $E^{(i)} = (1 - k_i^2)E^{(i-1)}$, we will see a big reduction in $E^{(i)}$ whenever $|k_i|$ is close to 1. Note that $k_1 = .8328$, so we get the first big reduction in the first step. The next relatively large jump occurs at $k_6 = -.7505$ and another reduction occurs for $k_9 = .5605$.
- (f) Since these roots are complex and the coefficients of the predictor polynomial are real, five more are at the complex conjugate locations. Since the degree of the polynomial is 11, there is one more real root. It has to be inside the unit circle, but we can't tell whether it is + or -.
- (g) $G^2 = E^{(11)} = R[0] \prod_{k=1}^{11} (1 - k_i^2)$. Using the value $R[0] = 1.9145$ determined in (d) and the k_i s from the table, we get $G = .3634$.
- (h) The formants are approximately represented by the poles that are closest to the unit circle. The formant frequency would be determined from the angle of z_i by the equation $F = (\angle z_i)F_s/(2\pi)$, where F_s is the sampling rate. The poles and their magnitudes and angles are given in Table P9.6.1. The first formant has the smallest bandwidth because its pole is

i	$ z_i $	$\angle z_i$	F (in Hz)	Formant
1	.2567	2.0677	2632	no
2	.9681	1.4402	1834	F_2
3	.9850	.2750	350	F_1
4	.8647	2.0036	2551	F_3 ??
5	.9590	2.4162	3076	F_3 or F_4 ???

Table P9.6.1: Values of roots of prediction error filter.

closest to the unit circle.

- (i) A MATLAB plot of the frequency response of the all-pole model filter is given in Figure P9.6.1

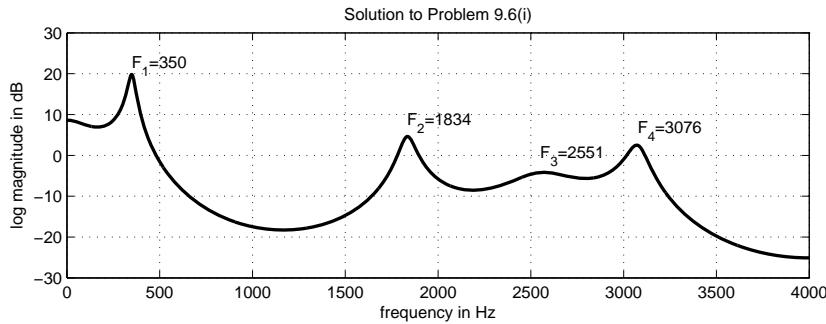


Figure P9.6.1: Plot of frequency response of all-pole model.

9.7 We can write the error signal, $e[n]$, in the form:

$$e[n] = s[n] - \tilde{s}[n] = s[n] - \beta s[n - n_0]$$

with transfer function:

$$\frac{E(z)}{S(z)} = 1 - \beta z^{-n_0}$$

and with mean-squared error, E_n of the form:

$$\begin{aligned} E_n &= \sum_m e_n^2[m] = \sum_m (s_n[m] - \tilde{s}_n[m])^2 \\ &= \sum_m (s_n[m] - \beta s_n[m - n_0])^2 \end{aligned}$$

We now can differentiate E_n with respect to β and set the result to 0, giving:

$$\frac{\partial E_n}{\partial \beta} = 0 = 2 \sum_m (s_n[m] - \beta s_n[m - n_0])(-s_n[m - n_0])$$

We can now solve for β as:

$$\sum_m s_n[m] s_n[m - n_0] = \beta \sum_m s_n^2[m - n_0]$$

giving, for β , the result:

$$\beta = \frac{\sum_m s_n[m] s_n[m - n_0]}{\sum_m s_n^2[m - n_0]} = \frac{\phi_n[0, n_0]}{\phi_n[n_0, n_0]}$$

with mean-squared prediction error, E_n of the form:

$$\begin{aligned} E_n &= \sum_m s_n^2[m] - \beta \sum_m s_n[m] s_n[m - n_0] \\ &= \phi_n[0, 0] - \beta \phi_n[0, n_0] \\ &= \phi_n[0, 0] - \frac{\phi_n[0, n_0]}{\phi_n[n_0, n_0]} \cdot \phi_n[0, n_0] \\ &= \frac{\phi_n[0, 0] \cdot \phi_n[n_0, n_0] - \phi_n^2[0, n_0]}{\phi_n[n_0, n_0]} \end{aligned}$$

For the autocorrelation method we get the results:

$$\begin{aligned} \phi_n[i, k] &= R_n[|i - k|] \\ \beta &= \frac{R_n[n_0]}{R_n[0]} \\ E_n &= \frac{R_n^2[0] - R_n^2[n_0]}{R_n[0]} \end{aligned}$$

and for the covariance method we get the results:

$$\begin{aligned} \phi_n[i, k] &= \phi_n[k, i] \\ \beta &= \frac{\phi_n[0, n_0]}{\phi_n[n_0, n_0]} \\ E_n &= \frac{\phi_n[0, 0] \cdot \phi_n[n_0, n_0] - \phi_n^2[0, n_0]}{\phi_n[n_0, n_0]} \end{aligned}$$

9.8 (a) (i) The solution for $R_n[k]$ is as follows:

$$\begin{aligned} R_n[k] &= \sum_{m=0}^{L-1-k} \beta^{m+n} \beta^{m+n+k} \\ &= \beta^{2n+k} \sum_{m=0}^{L-1-k} \beta^{2m} \\ &= \begin{cases} \beta^{2n+k} \frac{(1 - \beta^{2(L-k)})}{(1 - \beta^2)} & 0 \leq k \leq L-1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

(ii) The coefficient that minimizes the prediction error for a first order predictor is:

$$\begin{aligned} \alpha_1 &= \frac{R_n[1]}{R_n[0]} \\ &= \frac{\beta^{2n+1} \frac{(1 - \beta^{2(L-1)})}{(1 - \beta^2)}}{\beta^{2n} \frac{(1 - \beta^{2L})}{(1 - \beta^2)}} \\ &= \beta \frac{(1 - \beta^{2(L-1)})}{(1 - \beta^{2L})} \end{aligned}$$

(iii) as $L \rightarrow \infty$, $\alpha_1 \rightarrow \beta$.

(iv) We can write the optimum prediction error as:

$$\begin{aligned} \mathcal{E}_n^{\text{auto}} &= R_n[0] - \alpha_1 R_n[1] \\ &= \beta^{2n} \frac{(1 - \beta^{2L})}{(1 - \beta^2)} - \beta \frac{(1 - \beta^{2(L-1)})}{(1 - \beta^{2L})} \beta^{2n+1} \frac{(1 - \beta^{2(L-1)})}{(1 - \beta^2)} \\ &= \frac{\beta^{2n}}{(1 - \beta^2)} (\dots) - \frac{\beta^{2n+2}}{(1 - \beta^2)} (\dots) \text{ where } (\dots) \rightarrow 1 \text{ as } L \rightarrow \infty \\ &\approx \frac{\beta^{2n}}{(1 - \beta^2)} \cdot (1 - \beta^2) = \beta^{2n} \neq 0 \text{ as } L \rightarrow \infty \end{aligned}$$

(b) (i) We write the covariance as:

$$\begin{aligned} \phi_n[i, k] &= \sum_{m=0}^{L-1} \beta^{n+m-i} \beta^{m-k+n} \\ &= \beta^{2n-i-k} \sum_{m=0}^{L-1} \beta^{2m} \\ &= \beta^{2n-i-k} \left(\frac{1 - \beta^{2L}}{1 - \beta^2} \right) \end{aligned}$$

(ii) For the optimum predictor we get:

$$\begin{aligned}\phi_n[1, 1]\alpha_1 &= \phi_n[1, 0] \\ \alpha_1 &= \frac{\phi_n[1, 0]}{\phi_n[1, 1]} \\ &= \frac{\beta^{2n-1} \left(\frac{1 - \beta^{2L}}{1 - \beta^2} \right)}{\beta^{2n-2} \left(\frac{1 - \beta^{2L}}{1 - \beta^2} \right)} \\ &= \beta\end{aligned}$$

(iii) α_1 is independent of L in this case.

(iv) We can compute the optimum prediction error for the covariance method as:

$$\begin{aligned}\mathcal{E}_n^{\text{cov}} &= \phi_n[0, 0] - \alpha_1 \phi_n[0, 1] \\ &= \beta^{2n} \left(\frac{1 - \beta^{2L}}{1 - \beta^2} \right) - \beta \cdot \beta^{2n-1} \left(\frac{1 - \beta^{2L}}{1 - \beta^2} \right) \\ &= 0 \text{ independent of } L\end{aligned}$$

9.9 (a) From Figure P9.9 in the text and the problem statement we have:

$$\begin{aligned}w[n] &= x[n] * h[n] \\ E\{\epsilon[n]\epsilon[n+m]\} &= \begin{cases} \sigma_\epsilon^2 & m = 0 \\ 0 & m \neq 0 \end{cases} \\ y[n] &= w[n] + \epsilon[n]\end{aligned}$$

Thus we get:

$$\begin{aligned}\mathcal{E} &= E\{|y[n] - \hat{h}[n] * x[n]|^2\} \\ &= E\{y^2[n]\} - 2E\{y[n] \cdot \hat{h}[n] * x[n]\} + E\{\hat{h}[n] * x[n]\}^2\end{aligned}$$

If we examine each term separately we get:

$$\begin{aligned}E\{y^2[n]\} &= E\{w^2[n]\} + 2E\{w[n]\epsilon[n]\} + E\{\epsilon^2[n]\} \\ &= E\{[h[n] * x[n]]^2\} + \sigma_\epsilon^2 \\ &= E\left\{\sum_{k=0}^{\infty} \sum_{r=0}^{\infty} h[k]h[r]x[n-k]x[n-r]\right\} + \sigma_\epsilon^2 \\ &= \sum_{k=0}^{\infty} h[k] \sum_{r=0}^{\infty} h[r] E\{x[n-k]x[n-r]\} + \sigma_\epsilon^2 \\ &= \sum_{k=0}^{\infty} h[k] \sum_{r=0}^{\infty} h[r] \phi_{xx}[n-k, n-r] + \sigma_\epsilon^2\end{aligned}$$

For $x[n]$ stationary:

$$\begin{aligned}\phi_{xx}[n-k, n-r] &= R_x[k-r] = R_x[r-k] \\ E\{y^2[n]\} &= \sum_{k=0}^{\infty} h[j] \sum_{r=0}^{\infty} h[r] R_x[k-r] + \sigma_\epsilon^2 \\ &= \sum_{k=0}^{\infty} h[k] \sum_{r=0}^{\infty} h[r-k] R_x[r] + \sigma_\epsilon^2\end{aligned}$$

$$\begin{aligned}
E\{y[n] \cdot \hat{h}[n] * x[n]\} &= E\left\{\sum_{k=0}^{M-1} \hat{h}[k] y[n] x[n-k]\right\} \\
&= \sum_{k=0}^{M-1} \hat{h}[k] \phi_{xy}[n-k, n] \\
E\{\hat{h}[n] * x[n]\}^2 &= \sum_{r=0}^{M-1} \sum_{k=0}^{M-1} \hat{h}[k] \hat{h}[r] R_x[k-r]
\end{aligned}$$

Now we substitute these 3 terms into the expression for \mathcal{E} giving:

$$\begin{aligned}
\mathcal{E} = &\left[\sum_{r=0}^{\infty} \sum_{k=0}^{\infty} h[k] h[r] R_x[k-r] + \sigma_{\epsilon}^2 \right] - 2 \left[\sum_{k=0}^{M-1} \hat{h}[k] \phi_{xy}[n-k, n] \right] \\
&+ \left[\sum_{r=0}^{M-1} \sum_{k=0}^{M-1} \hat{h}[k] \hat{h}[r] R_x[k-r] \right]
\end{aligned}$$

We require that at the minimum we have:

$$\frac{\partial \epsilon}{\partial \hat{h}[i]} = 0$$

Thus:

$$\frac{\partial \epsilon}{\partial \hat{h}[i]} = -2\phi_{xy}[n-i, n] + \sum_{r=0}^{M-1} \hat{h}[r] R_x[i-r] + \sum_{k=0}^{M-1} \hat{h}[k] R_x[k-i] = 0$$

Since:

$$\sum_{k=0}^{M-1} \hat{h}[k] R_x[k-i] = \sum_{r=0}^{M-1} \hat{h}[r] R_x[i-r]$$

we can write

$$\phi_{xx}[n-i, n] = \sum_{k=0}^{M-1} \hat{h}[k] R_x[|k-i|], \quad i = 0, 1, \dots, M-1$$

Note also that:

$$\phi_{xy}[n-i, n] = R_{xy}[i]$$

then

$$\sum_{k=0}^{M-1} \hat{h}[k] R_x[|k-i|] = R_{xy}[i] \quad i = 0, 1, \dots, M-1$$

- (b) The matrix equation obtained in part (a) for solving for the predictor coefficients has the same form as in the LPC method, except for the additional term involving the cross-correlation between input and output. Since $R_x[|k-i|]$ represents a Toeplitz matrix, the Levinson-Durbin recursion can be applied to obtain a solution.
- (c) The mean-squared error is given by:

$$\mathcal{E} = E\{y^2[n]\} - 2 \sum_{k=0}^{M-1} \hat{h}[k] R_{xy}[k] + \sum_{k=0}^{M-1} \hat{h}[k] \sum_{r=0}^{M-1} \hat{h}[r] R_x[|r-k|]$$

where $E\{y^2[n]\} = \phi_{yy}[0, 0]$ for $x[n]$ stationary. The minimum mean-squared error is obtained when $\hat{h}[k]$ satisfies the equation obtained in part (a), which implies that

$$R_{xy}[k] = \sum_{r=0}^{M-1} \hat{h}[r] R_x[|r-k|]$$

$$\mathcal{E}_{\min} = \phi_{yy}[0, 0] - \sum_{k=0}^{M-1} \hat{h}[k] R_{xy}[k]$$

The minimum mean-squared error is seen to consist of a fixed term, $\phi_{yy}[0, 0]$ that depends on $x[n]$, $h[n]$, and $\epsilon[n]$, minus a terms that depends on the predictor coefficients, $\hat{h}[k]$.

9.10

$$A^{(i)}(z) = 1 - \sum_{j=1}^i \alpha_j^{(i)} z^{-j}$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1$$

$$\alpha_i^{(i)} = k_i$$

Substituting for $\alpha_i^{(i)}$ gives:

$$A^{(i)}(z) = 1 - \sum_{j=1}^{i-1} \alpha_j^{(i)} z^{-j} - k_i z^{-i}$$

Substituting for $\alpha_j^{(i)}$ gives:

$$A^{(i)}(z) = 1 - \sum_{j=1}^{i-1} \left[\alpha_j^{(i-1)} z^{-j} - k_i \alpha_{i-j}^{(i-1)} z^{-j} \right] - k_i z^{-i}$$

$$= \left[1 - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} z^{-j} \right] + k_i \sum_{j=1}^{i-1} \alpha_{i-j}^{(i-1)} z^{-j} - k_i z^{-i}$$

In the second term, we do the variable substitution $j \rightarrow j' + i$, giving:

$$A^{(i)}(z) = \left[1 - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} z^{-j} \right] + k_i \sum_{j'=i-1}^1 \alpha_{j'}^{(i-1)} z^{j-i} - k_i z^{-i}$$

$$= \left[1 - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} z^{-j} \right] - k_i z^{-i} \left[\sum_{j'=1}^{i-1} \alpha_{j'}^{(i-1)} (z^{-1})^{-j'} \right]$$

$$= A^{(i-1)}(z) - k_i z^{-i} A^{(i-1)}(z^{-1})$$

- 9.11** We form the expression consisting of the denominator minus the numerator terms and recognize that it can be written as a squared term, i.e.,

$$\begin{aligned} & \sum_{m=0}^{N-1} (e^{(i-1)}[m])^2 + \sum_{m=0}^{N-1} (b^{(i-1)}[m-1])^2 - 2 \sum_{m=0}^{N-1} e^{(i-1)}[m] b^{(i-1)}[m-1] \\ &= \sum_{m=0}^{N-1} (e^{(i-1)}[m] - b^{(i-1)}[m-1])^2 \geq 0 \end{aligned}$$

If the numerator is positive, then the denominator is greater than or equal to the numerator and the result is proved. If the numerator is negative, we can repeat the construction by forming the expression consisting of the denominator plus the numerator, which leads to the result:

$$\sum_{m=0}^{N-1} (e^{(i-1)}[m] + b^{(i-1)}[m-1])^2 \geq 0$$

Thus the denominator is again greater than or equal to the numerator and the result is proved.

9.12 (a) We desire a spectrally flattened version of the signal of the form:

$$y[n] = \sum_{k=1}^M \left[e^{j2\pi kn/N_p} + e^{-j2\pi kn/N_p} \right]$$

Let $\Delta = 2\pi n/N_p$; then we can sum from $-Mm$ to M , and subtract the $k = 0$ term giving:

$$\begin{aligned} y[n] &= \sum_{k=-M}^M e^{j\Delta k} - 1 \\ &= e^{-j\Delta M} \frac{(1 - e^{j\Delta(2M+1)})}{(1 - e^{j\Delta})} - 1 \\ &= \frac{e^{-j\Delta(2M+1)/2} - e^{j\Delta(2M+1)/2}}{e^{-j\Delta/2} - e^{j\Delta/2}} - 1 \\ &= \frac{\sin(2M+1)\Delta/2}{\sin(\Delta/2)} - 1 \\ &= \frac{\sin(\pi(2M+1)n/N_p)}{\sin(\pi n/N_p)} - 1 \end{aligned}$$

(b) Since:

$$s[n] = \sum_{k=1}^M \left[\beta_k e^{j2\pi kn/N_p} + \beta_k^* e^{-j2\pi kn/N_p} \right]$$

we can express $s[n] - \sum_{l=1}^{2M} \alpha_l s[n-l]$ as:

$$\begin{aligned} s[n] - \sum_{l=1}^{2M} \alpha_l s[n-l] &= \sum_{k=1}^M \left[\beta_k e^{j2\pi kn/N_p} + \beta_k^* e^{-j2\pi kn/N_p} \right] \\ &\quad - \sum_{l=1}^{2M} \alpha_l \sum_{k=1}^M \left[\beta_k e^{j2\pi k(n-l)/N_p} + \beta_k^* e^{-j2\pi k(n-l)/N_p} \right] \\ &= \sum_{k=1}^M \beta_k e^{j2\pi kn/N_p} \left[1 - \sum_{l=1}^{2M} \alpha_l e^{-j2\pi kl/N_p} \right] \\ &\quad + \sum_{k=1}^M \beta_k^* e^{-j2\pi kn/N_p} \left[1 - \sum_{l=1}^{2M} \alpha_l e^{j2\pi kl/N_p} \right] \end{aligned}$$

The prediction error is minimized (in this case it vanishes) by choosing α_l so that:

$$1 - \sum_{l=1}^{2M} \alpha_l e^{-j2\pi kl/N_p} = 0 \quad k = \pm 1, \pm 2, \dots, \pm M$$

Expressing $A(z)$, the polynomial for the LPC coefficients, as

$$A(z) = 1 - \sum_{l=1}^{2M} \alpha_l z^{-l}$$

we see that the roots of $A(z)$ occur at $z_k = e^{-j2\pi k/N_p}$, $k = 1, 2, \dots, M$.

(c) We can write the system function, $H(z)$ in terms of the roots as:

$$H(z) = \frac{1}{\prod_{k=1}^p (1 - z_k z^{-1})}$$

with $|z_k| \leq 1$, giving:

$$\hat{h}[n] = \begin{cases} 0 & n = 0 \\ \sum_{k=1}^p \frac{(z_k)^n}{n} & n > 0 \\ 0 & n < 0 \end{cases}$$

(d) If we set $y[n] = n\hat{h}[n]$, we get

$$\begin{aligned} y[n] &= \sum_{k=1}^p (z_k)^n & n > 0 \\ &= \sum_{k=1}^p e^{\pm j2\pi nk/N_p} \\ &= \sum_{k=1}^M \left[e^{j2\pi nk/N_p} + e^{-j2\pi nk/N_p} \right] \end{aligned}$$

which is a spectrally flattened signal.

9.13 (a) The inverse filter for the speech model is given by:

$$A(z) = \sum_{k=0}^{25} \alpha_k z^{-k} = \frac{1}{H(z)}$$

where $H(z)$ represents the combined vocal tract, glottal shaping and radiation impedance transfer function. The output of the inverse filter, $y[n]$, approximates the original glottal excitation waveform, which for voiced speech is a quasi-periodic pulse train. The original sampled speech, $x[n]$, can be represented as:

$$x[n] = y[n] * h[n]$$

and since $H(e^{j\omega}) = 1/A(e^{j\omega})$ we have:

$$\log |X(e^{j\omega})| = \log \left\{ \frac{|Y(e^{j\omega})|}{|A(e^{j\omega})|} \right\} = \log |Y(e^{j\omega})| - \log |A(e^{j\omega})|$$

The required spectral correction network to achieve this relation is shown in Figure P9.13.1.

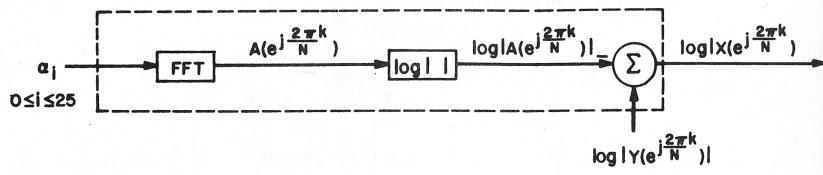


Figure P9.13.1: Spectral correction network.

- (b) The standard method of windowing the speech waveform results in a spectrum that is the convolution of the spectrum of the sampled speech with the spectrum of the window. As a result, spectral zeros can be lost due to the “averaging” effect of the convolution. The alternate approach suggested in this problem does not require explicit windowing of the speech samples if the covariance method is used for the LPC analysis. The 25th-order predictor yields a very good approximation to the overall transfer function, $H(e^{j\omega})$, including the presence of zeros. The final short-time spectrum, $X(e^{j\omega})$, is the product of $H(e^{j\omega})$ and the whitened excitation, $Y(e^{j\omega})$, and any zeros will be well represented.

- 9.14 (a)** To determine stability, we need to find out if the denominator polynomial has all of its roots inside the unit circle. Consider the denominator polynomial:

$$A(z) = 1 - 0.25z^{-1} - 0.75z^{-2} - 0.875z^{-3} = 1 - \alpha_1^{(3)}z^{-1} - \alpha_2^{(3)}z^{-2} - \alpha_3^{(3)}z^{-3}$$

to be a prediction error filter; therefore we can convert from a_i to k_i using the backward iteration:

$$\begin{aligned} k_i &= a_i^{(i)} \quad i = p, p-1, \dots, 1 \\ \alpha_j^{(i-1)} &= [\alpha_j^{(i)} + \alpha_i^{(i)}\alpha_{i-j}^{(i)}]/(1 - k_i^2) \quad 1 \leq j \leq i-1 \end{aligned}$$

Thus we compute the values of k_i using the backward iteration:

$$\begin{aligned} k_3 &= \alpha_3^{(3)} = 0.875 \\ \alpha_1^{(2)} &= [\alpha_1^{(3)} + \alpha_3^{(3)}\alpha_2^{(3)}]/(1 - k_3^2) = \frac{0.25 + (0.875)(0.75)}{1 - (0.875)^2} = 3.84 \\ \alpha_2^{(2)} &= [\alpha_2^{(3)} + \alpha_3^{(3)}\alpha_1^{(3)}]/(1 - k_3^2) = \frac{0.75 + (0.875)(0.25)}{1 - (0.875)^2} = 4.13 \\ k_2 &= \alpha_2^{(2)} = 4.13 \\ \alpha_1^{(1)} &= [\alpha_1^{(2)} + \alpha_2^{(2)}\alpha_1^{(2)}]/(1 - k_2^2) = \frac{3.84 + (4.13)(3.84)}{1 - (4.13)^2} = -1.23 \\ k_1 &= \alpha_1^{(1)} = -1.23 \end{aligned}$$

Since $|k_i| > 1$ for k_2 and k_1 , the system is unstable.

- (b) The system is not minimum phase since there is a zero at $z = 4$, which is outside the unit circle, as well as poles outside the unit circle.

- 9.15 (a)** We can solve for the PARCOR coefficients from the LPC coefficients using the backward iteration formula, giving:

$$\begin{aligned}
 \alpha_1^{(3)} &= -0.5 \\
 \alpha_2^{(3)} &= -0.25 \\
 \alpha_3^{(3)} &= -0.5 \\
 k_3 &= \alpha_3^{(3)} = -0.5 \\
 \alpha_1^{(2)} &= \frac{\alpha_1^{(3)} + \alpha_3^{(3)}\alpha_2^{(3)}}{1 - k_3^2} = \frac{-0.5 + (-0.5)(-0.25)}{1 - (0.5)^2} = -0.5 \\
 \alpha_2^{(2)} &= \frac{\alpha_2^{(3)} + \alpha_3^{(3)}\alpha_1^{(3)}}{1 - k_3^2} = \frac{-0.25 + (-0.5)(-0.5)}{1 - (0.5)^2} = 0 \\
 k_2 &= \alpha_2^{(2)} = 0 \\
 \alpha_1^{(1)} &= \frac{\alpha_1^{(2)} + \alpha_2^{(2)}\alpha_1^{(2)}}{1 - k_2^2} = \frac{-0.5 + (0)(-0.5)}{1} = -0.5 \\
 k_1 &= \alpha_1^{(1)} = -0.5
 \end{aligned}$$

- (b)** The polynomial is stable since all PARCOR coefficients are of magnitude < 1 .

- 9.16** The energy of the linear prediction residual using a third order predictor is:

$$\begin{aligned}
 E_n^{(3)} &= E_n^{(0)}(1 - k_1^2)(1 - k_2^2)(1 - k_3^2) \\
 &= (2000)(1 - 0.25)(1 - 0.25)(1 - 0.04) \\
 &= 1080
 \end{aligned}$$

- 9.17** Recall the property of the LPC solutions that the mean-squared errors can be represented as:

$$E_p = E_0 \prod_{i=1}^p (1 - k_i^2)$$

We can use the equation above to solve for k_1, k_2, k_3 giving:

$$\begin{aligned}
 E_1 &= E_9(1 - k_1^2) \text{ or } 5.1 = 10 \cdot (1 - k_1^2), k_1^2 = 0.49, k_1 = \pm 0.7 \\
 E_2 &= E_1(1 - k_2^2) \text{ or } 4.284 = 5.1 \cdot (1 - k_2^2), k_2^2 = 0.16, k_2 = \pm 0.4 \\
 E_3 &= E_2(1 - k_3^2) \text{ or } 4.11264 = 4.284 \cdot (1 - k_3^2), k_3^2 = 0.04, k_3 = \pm 0.2
 \end{aligned}$$

Using the Durbin recursion formula we can solve for the α parameters from the k parameters. For simplicity we assume all k parameters are positive.

$$\begin{aligned}
 \alpha_1^{(1)} &= 0.7 \\
 \alpha_2^{(2)} &= k_2 = 0.4 \\
 \alpha_1^{(2)} &= \alpha_1^{(1)} - k_2 \alpha_1^{(1)} = 0.7 - 0.4 \cdot 0.7 = 0.42 \\
 \alpha_3^{(3)} &= k_3 = 0.2 \\
 \alpha_1^{(3)} &= \alpha_1^{(2)} - k_3 \alpha_2^{(2)} = 0.42 - 0.2 \cdot 0.4 = 0.34 \\
 \alpha_2^{(3)} &= \alpha_2^{(2)} - k_3 \alpha_1^{(2)} = 0.4 - 0.2 \cdot 0.42 = 0.316
 \end{aligned}$$

The resulting third order all-pole model is:

$$H(z) = \frac{G}{1 - 0.34z^{-1} - 0.316z^{-2} - 0.2z^{-3}}$$

where G is the LPC gain.

9.18 (a) The value of $E_n(\beta, P)$ can be expressed as:

$$\begin{aligned} \mathcal{E}_n(\beta, N_p) &= \sum_m (s_n[m] - \beta s_n[m - N_p])^2 \\ &= \sum_m s_n^2[m] - 2\beta \sum_m s_n[m]s_n[m - N_p] + \beta^2 \sum_m s_n^2[m - N_p] \\ \frac{\partial \mathcal{E}_n(\beta, N_p)}{\partial \beta} &= -2 \sum_m s_n[m]s_n[m - N_p] + 2\beta \sum_m s_n^2[m - N_p] = 0 \\ \beta &= \frac{\sum_m s_n[m]s_n[m - N_p]}{\sum_m s_n^2[m - N_p]} \end{aligned}$$

(b) To solve for the optimum value of N_p , we substitute the value of β from the solution above and find the maximum value over a range of values for N_p , i.e.,

$$N_p^{\text{opt}} = \max_{N_p} \frac{\sum_m s_n[m]s_n[m - N_p]}{\sum_m s_n^2[m - N_p]}$$

(c) This method of finding the pitch period takes more computation than the autocorrelation method, but it provides more accuracy in the value of N_p for reasonably large pitch periods. For small pitch periods, it is difficult to distinguish the pitch peak in the autocorrelation function from the peaks due to the formant structure.

9.19 (a) The LPC prediction error filter can be expressed as:

$$A(z) = 1 - \sum_{k=1}^p \alpha_k z^{-k}$$

(a) The LSP polynomials are of the form:

$$\begin{aligned} P(z) &= A(z) + z^{-(p+1)}A(z^{-1}) \\ Q(z) &= A(z) - z^{-(p+1)}A(z^{-1}) \end{aligned}$$

If we assume that p is even, then we get:

$$\begin{aligned} P(-1) &= A(-1) + (-1)^{p+1}A(-1) \\ &= 1 - \sum_{k=1}^p \alpha_k (-1)^{-k} - \left(1 - \sum_{k=1}^p \alpha_k (-1)^{-k} \right) = 0 \\ Q(1) &= A(1) - (1)^{p+1}A(1) = 0 \end{aligned}$$

(b) The zeros of $P(z)$ and $Q(z)$ occur at:

$$A(z) \pm z^{-(p+1)} A(z^{-1}) = 0$$

$$H_{\text{allpass}}(z) = \frac{z^{-(p+1)} A(z^{-1})}{A(z)} = \pm 1$$

where $H_{\text{allpass}}(z)$ is an all-pass filter since the zeros of $A(z)$ are inside the unit circle. On the unit circle we have $|z| = 1 \Rightarrow |H_{\text{allpass}}(z)| = 1$; thus the zeros of $P(z)$ and $Q(z)$ are at angles such that the phase of $H_{\text{allpass}}(z)$ is a multiple of π . Any other values of z , off the unit circle, give $|H_{\text{allpass}}(z)| \neq 1$, so the roots have to lie on the unit circle.

9.20 One LSP polynomial is:

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1}) = 1 + z^{-(p+1)}$$

The roots of $P(z)$ thus occur at $z_p^k = e^{j(2k+1)\pi/(p+1)}$, $k = 0, 1, 2, \dots, p$.

Similarly, the LSP $Q(z)$ polynomial is of the form:

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1}) = 1 - z^{-(p+1)}$$

and the roots of $Q(z)$ occur at $z_q^k = e^{j2k\pi/(p+1)}$, $k = 0, 1, 2, \dots, p$. Thus it is seen that the LSF frequencies are equally spaced over the interval $[0, \pi]$ with $\Delta\omega = \pi/(p+1)$.

9.21 We first convert the LSF frequencies to normalized radian frequencies to form the polynomials $Q(z)$ and $P(z)$. We recall that, for even order polynomials, $P(z)$ has a root at $z = -1$ and $Q(z)$ has a root at $Z = 1$. The LSF frequencies of $2000/3 = 6666.67$ and 2000 Hz correspond to normalized radian frequencies of:

$$\omega_1 = \frac{2\pi \cdot (2000/3)}{8000} = \pi/6 \text{ radians}$$

$$\omega_2 = \frac{2\pi \cdot 2000}{8000} = \pi/2 \text{ radians}$$

with complex root magnitudes of 1, and real roots at $\omega_0 = 0$ (for $Q(z)$) and $\omega_3 = \pi$ (for $P(z)$). Thus we can now form the polynomials $Q(z)$ and $P(z)$ as:

$$\begin{aligned} Q(z) &= (1 - z^{-1})(1 - 2z^{-1} \cos(\omega_2) + z^{-2}) \\ &= 1 - z^{-1} + z^{-2} - z^{-3} \quad (\omega_2 = \pi/2, \cos(\pi/2) = 0) \\ P(z) &= (1 + z^{-1})(1 - 2z^{-1} \cos(\omega_1) + z^{-2}) \\ &= 1 - (\sqrt{3} - 1)z^{-1} - (\sqrt{3} - 1)z^{-2} + z^{-3} \quad (\omega_1 = \pi/6, \cos(\pi/6) = \sqrt{3}/2) \end{aligned}$$

Now we can solve for the linear prediction inverse filter, $A(z)$ as:

$$A(z) = \frac{P(z) + Q(z)}{2} = 1 - \frac{\sqrt{3}}{2}z^{-1} + \left(1 - \frac{\sqrt{3}}{2}\right)z^{-2}$$

9.22 (a) We can write the following expression for $e[n]$:

$$e[n] = w[n] - \sum_{i=1}^p \alpha_i w[n-i] = -\sum_{i=1}^p \alpha_i w[n-i]$$

$$\begin{aligned} R_e[m] &= \sum_{n=-\infty}^{\infty} e[n]e[n+m] \\ &= \sum_{n=-\infty}^{\infty} \sum_{i=0}^p \sum_{j=0}^p \alpha_i \alpha_j w[n-i]w[n+m-j] \end{aligned}$$

We make a change of variables and let $n' = n + i$, giving:

$$R_e[m] = \sum_{n'=-\infty}^{\infty} \sum_{i=0}^p \sum_{j=0}^p \alpha_i \alpha_j w[n']w[n'+m+i-j]$$

Make another changes of variables and let $j = l + i$ giving:

$$R_e[m] = \sum_{i=0}^p \sum_{l=-i}^{p-i} \alpha_i \alpha_{i+l} \sum_{n'=\infty}^{\infty} w[n']w[n'+m-l]$$

We note that:

$$\alpha_{i+l} = \begin{cases} 0 & i+l < 0 \implies l < -i \\ 0 & i+l > p \implies l > p-i \end{cases}$$

$$\begin{aligned} R_e[m] &= \sum_{l=-\infty}^{\infty} \left[\sum_{i=0}^p \alpha_i \alpha_{i+l} \right] \left[\sum_{n'=-\infty}^{\infty} w[n']w[n'+m-l] \right] \\ &= \sum_{l=-\infty}^{\infty} R_a[l] R_w[m-l] \end{aligned}$$

- (b)** For a 10 kHz sampling rate, a total of 121 values of m are needed to give $R_e[m]$ from $m = 30$ (3 msec) to $m = 150$ (15 msec). $R_a[l]$ is a symmetric sequence of length $(2p+1)$ samples and $R_w[l]$ is a symmetric sequence of $(2L-1)$ samples. If we assume $R_a[l]$ and $R_w[l]$ are available (i.e., no additional computation is required), then for each value of m , a total of $(2p+1)$ multiplies and adds are required to implement the convolution of R_a and R_w . Thus the total computation for 121 values of m is:

$$T_c = 121 * (2p+1) = 242p + 121$$

For $p = 10$, a total of 2541 multiplies and adds are required to give $R_e[n]$. If $R_a[l]$ and $R_w[l]$ are not available, the amount of computation grows linearly with L (instead of p).

9.23 (a)

$$e[n] = -\sum_{i=0}^p \alpha_i x[n-i]$$

$$\begin{aligned}\mathcal{E}^{(p)} &= \sum_{n=0}^{L-1+p} e^2[n] = \sum_{n=0}^{L-1+p} \left[-\sum_{i=0}^p \alpha_i x[n-i] \right] \left[-\sum_{j=0}^p \alpha_j x[n-j] \right] \\ &= \sum_{i=0}^p \alpha_i \sum_{j=0}^p \alpha_j \sum_{n=0}^{L-1+p} x[n-i]x[n-j]\end{aligned}$$

But:

$$\sum_{n=0}^{L-1+p} x[n-i]x[n-j] = \sum_{n=0}^{L-1+p} x[n]x[n-j+i] = R[i-j]$$

So:

$$\mathcal{E}^{(p)} = \sum_{i=0}^p \alpha_i \sum_{j=0}^p \alpha_j R[i-j] = \alpha R_a \alpha^t$$

where R_a is the matrix:

$$R_a = \begin{bmatrix} R[0] & R[1] & \dots & R[p] \\ R[1] & R[0] & \dots & R[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ R[p] & R[p-1] & \dots & R[0] \end{bmatrix}$$

(b)

$$\tilde{e}[n] = -\sum_{i=0}^p \alpha_i \hat{x}[n-i]$$

Repeating the derivation of part (a) we get:

$$\mathcal{F}^{(p)} = \sum_{i=0}^p \alpha_i \sum_{j=0}^p \alpha_j \hat{R}[i-j] = \alpha R_{\hat{\alpha}} \alpha^t$$

where $R_{\hat{\alpha}}$ is the matrix:

$$R_{\hat{\alpha}} = \begin{bmatrix} R[0] & R[1] & \dots & R[p] \\ R[1] & R[0] & \dots & R[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ R[p] & R[p-1] & \dots & R[0] \end{bmatrix}$$

(c)

$$D = \frac{\mathcal{F}^{(p)}}{\mathcal{E}^{(p)}} = \frac{\alpha R_{\hat{\alpha}} \alpha^t}{\alpha R_a \alpha^t}$$

Since D is a ratio of prediction residuals, and since $\mathcal{F}^{(p)}$ must be greater than (or equal to) $\mathcal{E}^{(p)}$; therefore

$$D \geq 1.0$$

9.24 (a) We begin with the relations (from the previous problem):

$$D(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}) = \frac{\mathcal{F}^{(p)}}{\mathcal{E}^{(p)}} = \frac{\boldsymbol{\alpha} R_{\hat{\alpha}} \boldsymbol{\alpha}^t}{\hat{\boldsymbol{\alpha}} R_{\hat{\alpha}} \hat{\boldsymbol{\alpha}}^t}$$

Also, from the previous problem, we have:

$$\mathcal{F}^{(p)} = \sum_{i=0}^p \alpha_i \sum_{j=0}^p \alpha_j \hat{R}[j-i]$$

We make a change of variables and let $k = j - i$, $j = k + i$, giving:

$$\mathcal{F}^{(p)} = \sum_{i=0}^p \alpha_i \sum_{k=-i}^{p-i} \alpha_{k+i} \hat{R}[k]$$

We can rearrange the summations on i and k by recognizing that

$$\alpha_l = \begin{cases} 0 & l < 0 \\ 0 & l > p \end{cases}$$

We can now complete the square by summing on k from $-p$ (the smallest value of k) to $+p$ (the largest value of k) giving:

$$\mathcal{F}^{(p)} = \sum_{k=-p}^p \left[\sum_{i=0}^p \alpha_i \alpha_{k+i} \right] \hat{R}[k]$$

We recognize that the inner summation is $b[k]$, thus giving

$$\mathcal{F}^{(p)} = \sum_{k=-p}^p b[k] \hat{R}[k]$$

Since $b[k] = b[-k]$ and $\hat{R}[k] = \hat{R}[-k]$, we get:

$$\mathcal{F}^{(p)} = \hat{R}[0]b[0] + 2 \sum_{k=1}^p b[k] \hat{R}[k]$$

(b) Since all individual quantities are pre-computed, to evaluate D as a ratio of residuals, i.e.,

$$F = \frac{\alpha R_{\hat{\alpha}} \alpha^t}{\hat{\alpha} R_{\hat{\alpha}} \hat{\alpha}^t}$$

requires:

1. $(p+1) \times (p+2)$ multiplies and adds to multiply (α by $R_{\hat{\alpha}}$) and then multiply the result by α^t .
2. 1 divide to give D since $\hat{\alpha} R_{\hat{\alpha}} \hat{\alpha}^t$ is precomputed.

For the alternative method of evaluating D , as discussed in part (a), we require:

1. $(p+1)$ multiplies and adds to give the product $\hat{R}[k] \cdot b[k] + b[0] \hat{R}[0]$.
2. 1 divide

Thus, neglecting the divide, the second form requires a factor of $(p+2)$ less computation, i.e., if R is the ratio of computation, then we get:

$$R = \frac{(p+1)(p+2)}{(p+1)} = (p+2)$$

so the second method is much more efficient than the first method.

Chapter 10

Algorithms for Estimating Speech Parameters

10.1 (MATLAB Exercise)

10.2 (a) For each of the 7 statistically independent pitch detectors we have:

$$P\{\text{correct}\} = p$$

$$P\{\text{incorrect}\} = 1 - p$$

The probability of n correct decisions and $7 - n$ incorrect decisions in 7 Bernoulli trials is:

$$P\{n \text{ correct}\} = \binom{7}{n} p^n (1-p)^{7-n},$$

where:

$$\binom{7}{n} = \frac{7!}{n!(7-n)!}$$

The pitch estimate is correct if $n = 4, 5, 6, 7$. A table of the combinations is as follows:

n	$\binom{7}{n}$	$P\{n \text{ correct}\}$
4	35	$35p^4(1-p)^3 = 35p^4(1-3p+3p^2-p^3) = 35(-p^7 + 3p^6 - 3p^5 + p^4)$
5	21	$p^5(1-p)^2 = p^5(1-2p+p^2) = 21(p^7 - 2p^6 + p^5)$
6	7	$p^6(1-p) = 7(-p^7 + p^6)$
7	1	p^7

The probability of a correct overall pitch estimate is therefore:

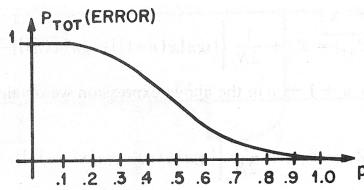
$$\begin{aligned} P_{\text{TOT}}(\text{no error}) &= P\{4 \text{ correct}\} + P\{5 \text{ correct}\} + P\{6 \text{ correct}\} + P\{7 \text{ correct}\} \\ &= -20p^7 + 70p^6 - 84p^5 + 35p^4 \\ &= p^4(35 - 84p + 70p^2 - 20p^3) \end{aligned}$$

The probability of error is then $1 - P_{\text{TOT}}(\text{no error})$ or:

$$P_{\text{TOT}}(\text{error}) = 1 - p^4(35 - 84p + 70p^2 - 20p^3)$$

(b) A table of the values of $P_{\text{TOT}}(\text{error})$ is as follows:

p	$P_{TOT}(\text{error})$
0	1
0.2	0.967
0.3	0.874
0.4	0.710
0.5	0.500
0.6	0.290
0.7	0.126
1	0

Figure P10.2.1: Plot of total probability of error versus p .

A plot of $P_{TOT}(\text{error})$ is given in Figure P10.2.1.

- (c) We require p such that:

$$p^4(35 - 84p + 70p^2 - 20p^3) \geq 0.95$$

From a cumulative binomial distribution table, the probability of observing 4 or more successful estimates is:

$$P\{4 \text{ or more}\} = \begin{cases} 0.967 & \text{for } p = 0.8 \\ 0.929 & \text{for } p = 0.75 \end{cases}$$

Iterating with $p = 0.78$ as an initial guess we obtain:

$$p = 0.775$$

10.3 (a) We are given that:

$$F_k = 2^{k-1} F_1 \quad k = 1, 2, \dots, M \quad \text{bandpass lower cutoff frequency}$$

$$F_{k+1} = 2^k F_1 \quad k = 1, 2, \dots, M \quad \text{bandpass upper cutoff frequency}$$

Given $50 < F_0 < 800$ a simple solution is to let $F_1 = 50$, and $M = 4$ giving the set of filters shown in the table:

k	lower cutoff	upper cutoff
1	50	100
2	100	200
3	200	400
4	400	800

- (b) The ideal bandpass filters are shown in Figure P10.3.1 for positive frequency where $M = 4$.
(c) Tone detection can be accomplished using a short-time zero crossing detector. If the zero-crossing count exceeds a given threshold for the filter, the tone is detected.

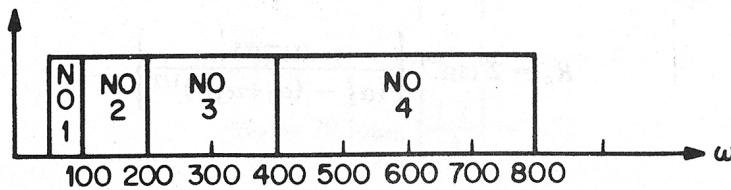


Figure P10.3.1: Plot of locations of ideal filters.

- (d) In using non-ideal filters, the resulting overlap between bands will result in “cross-talk”. If the energy due to cross-talk is sufficient, the zero-crossing detector may not produce the correct count.
 - (e) In this case fundamental frequencies of less than 300 Hz would be missing. One possibility is to pass the signal through a nonlinearity in order to restore the fundamental. Another possibility is to determine the fundamental from differences between higher harmonics.
-

Chapter 11

Digital Coding of Speech Signals

11.1

$$p(x) = \begin{cases} 1/\Delta, & |x| < \frac{\Delta}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} E\{x\} &= \int_{-\infty}^{\infty} xp(x)dx = \int_{-\Delta/2}^{\Delta/2} \left(\frac{x}{\Delta}\right) dx = \frac{1}{2\Delta} \left[x^2\right]_{-\Delta/2}^{\Delta/2} \\ &= \frac{1}{\Delta} \left[\frac{\Delta^2}{4} - \frac{\Delta^2}{4} \right] = 0 \end{aligned}$$

$$\begin{aligned} Var\{x\} &= E\{(x - \bar{x})^2\} = E\{x^2\} = \int_{-\infty}^{\infty} x^2 p(x)dx \\ &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} x^2 dx = \frac{1}{\Delta} \left[x^3\right]_{\Delta/2}^{\Delta/2} \\ &= \frac{1}{3\Delta} \left[\frac{\Delta^3}{4}\right] = \frac{\Delta^2}{12} \end{aligned}$$

11.2 The probability that $|x| > 4\sigma_x$ can be written in the form:

$$\begin{aligned} P\{|x| > 4\sigma_x\} &= 1 - P\{-4\sigma_x \leq x \leq 4\sigma_x\} \\ &= 1 - \int_{-4\sigma_x}^{4\sigma_x} p(x)dx \\ &= 1 - 2 \int_0^{4\sigma_x} p(x)dx \\ &= 1 - \frac{2}{\sqrt{2}\sigma_x} \int_0^{4\sigma_x} e^{-\sqrt{2}x/\sigma_x} dx \\ &= 1 - \frac{\sqrt{2}}{\sigma_x} \left[\frac{-\sigma_x}{\sqrt{2}} e^{-\sqrt{2}x/\sigma_x} \right]_0^{4\sigma_x} \\ &= 1 + e^{-4\sqrt{2}} - 1 = e^{-4\sqrt{2}} \\ P\{|x| > 4\sigma_x\} &= 3.5 \times 10^{-3} \end{aligned}$$

- 11.3** Let $y[n]$ be the output of a linear, shift-invariant system with impulse response $h[n]$, and input, $x[n]$, where $x[n]$ is a stationary, white noise process, and

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$

The autocorrelation function of $y[n]$ is given by:

$$\begin{aligned}\phi_{yy}[n, n+m] &= E\{y[n]y[n+m]\} \\ &= E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n-k] \sum_{r=-\infty}^{\infty} h[r]x[n+m-r]\right\} \\ &= \sum_{k=-\infty}^{\infty} h[k] \sum_{r=-\infty}^{\infty} h[r]E\{x[n-k]x[n+m-r]\}\end{aligned}$$

For $x[n]$ stationary, $E\{x[n-k]x[n+m-r]\} = E\{x[-k]x[m-r]\}$ and for $x[n]$ white, $E\{x[-k]x[m-r]\} = \begin{cases} \sigma_x^2 & m-r = -k \\ 0 & \text{otherwise} \end{cases}$

$$\therefore \phi_{yy}[n, n+m] = \sum_{k=-\infty}^{\infty} h[k]h[k+m]\sigma_x^2 = \sigma_x^2 \sum_{k=-\infty}^{\infty} h[k]h[k+m]$$

- 11.4** The mean and variance of $e_3[n]$ can be determined as:

$$\begin{aligned}e_3[n] &= ae_1[n] + (1-a)e_2[n] \\ E\{e_3[n]\} &= aE\{e_1[n]\} + (1-a)E\{e_2[n]\} \\ &= 0 \\ \sigma_{e_3}^2 &= E\{e_3[n]^2\} \\ &= E\{(ae_1[n] + (1-a)e_2[n])^2\} \\ &= a^2E\{e_1[n]^2\} + (1-a)^2E\{e_2[n]^2\} \\ &= a^2\sigma_{e_1}^2 + (1-a)^2\sigma_{e_2}^2\end{aligned}$$

We differentiate the expression for $\sigma_{e_3}^2$ with respect to the multiplier a , giving:

$$\begin{aligned}\frac{\partial \sigma_{e_3}^2}{\partial a} &= 2a\sigma_{e_1}^2 + 2(1-a)(-1)\sigma_{e_2}^2 = 0 \\ 2a\sigma_{e_1}^2 + 2a\sigma_{e_2}^2 &= 2\sigma_{e_2}^2 \\ a &= \frac{\sigma_{e_2}^2}{\sigma_{e_1}^2 + \sigma_{e_2}^2}\end{aligned}$$

The resulting variance of $e_3[n]$ is:

$$\begin{aligned}\sigma_{e_3}^2 &= \left(\frac{\sigma_{e_2}^2}{\sigma_{e_1}^2 + \sigma_{e_2}^2}\right)^2 \cdot \sigma_{e_1}^2 + \left(\frac{\sigma_{e_1}^2}{\sigma_{e_1}^2 + \sigma_{e_2}^2}\right)^2 \cdot \sigma_{e_2}^2 \\ &= \frac{\sigma_{e_1}^2 \cdot \sigma_{e_2}^2}{\sigma_{e_1}^2 + \sigma_{e_2}^2}\end{aligned}$$

11.5 (a) From the block diagram we can derive the relations:

$$\begin{aligned} x[n] &= x_c(nT/M) \\ \hat{x}[n] &= x[n] + e[n] = x_c(nT/M) + e[n] \\ \hat{y}[n] &= x_c(nT/M) * h[n] + e[n] * h[n] = x_c(nT/M) + f[n] \\ \hat{w}[n] &= \hat{y}[nM] = x_c(nT) + f[nM] = x_c(nT) + g[n] \end{aligned}$$

Thus we see that we get $\hat{w}[n] = x_c(nT)$ with noise component $g[n]$.

(b) We get, for the noise powers, the following:

$$\begin{aligned} \sigma_e^2 &= \frac{\Delta^2}{12} = \frac{X_{\max}^2}{3} \cdot 2^{-2B} \\ \sigma_f^2 &= \frac{1}{2\pi} \int_{-\pi/M}^{\pi/M} \sigma_e^2 |H(e^{j\omega})|^2 d\omega = \frac{\sigma_e^2(2\pi/M)}{2\pi} = \frac{\sigma_e^2}{M} = \sigma_g^2 \\ \sigma_g^2 &= \sigma_f^2 \quad \text{since down sampling does not change the long term average power} \end{aligned}$$

(c) We get the result:

$$\sigma_f^2 = \frac{\sigma_e^2}{M} = \left(\frac{X_{\max}^2}{3} \cdot 2^{-2B} \right) / M$$

which equates to:

$$2^{-2(B+1)} = \frac{2^{-2B}}{M} \Rightarrow M = 4$$

Thus a value of $M = 4$ gives a 1-bit improvement in SNR by spreading the quantization noise outside the signal band, and then filtering it out.

11.6 (a) A block diagram of the basic components needed for A/D and D/A conversion is given in Figure P11.6.1.



Figure P11.6.1: Block diagram of audio processing system.

(b) In order to accomodate a range of inputs of 100:1 in peak level and still maintain 60 dB SNR we require:

$$SNR = 60 + 20 \log_{10} 100 = 100 \text{dB}$$

at peak level.

$$SNR = 6B + 4.77 = 100$$

$$B = \frac{100 - 4.77}{6} = 15.87$$

Therefore we require $B = 16$ bits.

- (c) The sampling rate should be greater than twice the bandwidth of the analog lowpass filter in order to avoid aliasing. A lowpass analog filter is used at the input and output of the system. The stop-band frequency is chosen as one-half the sampling rate. The stop-band attenuation must be chosen to produce negligible aliasing. A typical value can be 40 dB, and the filter specification might therefore be given as shown in Figure P11.6.2.

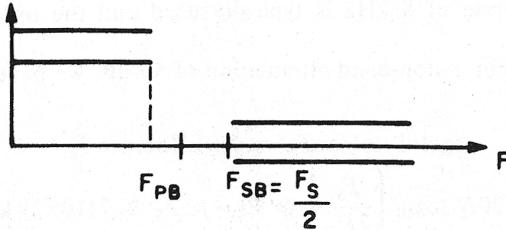


Figure P11.6.2: Lowpass filter specifications.

As an example, a Butterworth filter is 3 dB down at the nominal cutoff frequency with a falloff of $6N$ dB/octave, where N is the filter order. For $F_{c0} = 8$ kHz and 40 dB attenuation at $F = F_s/2$ we require:

$$20N \log_{10} \left[\frac{F_s}{2F_{c0}} \right] \geq 40; \quad 2F_{c0} = 16 \text{ kHz}$$

$$\log_{10} \left[\frac{F_s}{16} \right] \geq \frac{2}{N} \implies F_s \geq 16(10)^{2/N} \text{ kHz}$$

Choosing N and solving for the sampling frequency, F_s we get the results shown in the table below:

N	F_s
6	34.5 kHz
8	28.5 kHz
10	25.4 kHz
20	20.1 kHz

Elliptic filters offer a sharper cutoff for the same order filter, but possess poorer phase characteristics. Thus we see that for practical order filters, a sampling rate of 20 kHz or more is required.

- (d) Telephone quality speech corresponds to a bandwidth of about 3.5 kHz, and typical signal-to-noise ratio of 30-40 dB, with a range of signal levels of 100 to 1. If we maintain a 30 dB SNR , then

$$SNR_{\max} = 30 + 20 \log_{10}(100) = 30 + 40 = 70 \text{ (dB)}$$

and

$$SNR = 6B + 4.77 = 70$$

$$B = \frac{70 - 4.77}{6} = 11 \text{ bits}$$

A sampling rate of 8 kHz is typically used and the order of the required lowpass filter is reduced. For a stop-band attenuation of 40 dB, we have:

$$F_{c0} = 3.5 \text{ kHz}$$

$$20N \log_{10} \left[\frac{F_s}{2F_{c0}} \right] \geq 40 \implies F_s \geq 7(10^{2/N}) \text{ kHz}$$

Again choosing N and solving for the sampling frequency, F_s we get the results shown in the table below:

N	F_s
4	22.13 kHz
6	15.08 kHz
8	12.45 kHz
35	7.98 kHz

Note that since the transition region is narrow, a high order filter is required.

- 11.7 (a)** The sampling rate has no effect on the SNR as this depends strictly on the signal levels and the number of bits allocated for the quantizer in the A-to-D converter. Hence the SNR remains at 89 dB.
- (b)** If the signal level is changed by a factor of 10, then the signal power is reduced by 20 dB, so the overall SNR is reduced by 20 dB to a level of 69 dB.
- (c)** If the A-to-D converter is changed from 16-bits to 12-bits, then the noise is increased by 4-bits, or 24 dB, changing the overall SNR to $89-24=65$ dB.

- 11.8 (a)** From the problem statement we have:

$$w[n] = y'[n] + e_2[n] = x[n] + f[n]$$

and

$$y'[n] = x[n] + e_1[n]$$

$$\therefore w[n] = x[n] + e_1[n] + e_2[n]$$

If $e_1[n]$ and $e_2[n]$ are uncorrelated (i.e., the sampling offset, ϵ , is not zero) then

$$\sigma_f^2 = \sigma_{e_1}^2 + \sigma_{e_2}^2$$

and since the quantizers are the same we get:

$$\sigma_{e_1}^2 = \sigma_{e_2}^2 \implies \sigma_f^2 = 2\sigma_{e_1}^2$$

$$\therefore SNR = \frac{\sigma_x^2}{\sigma_f^2} = \frac{\sigma_x^2}{2\sigma_{e_1}^2} = \frac{SNR_1}{2}$$

- (b)** Since the individual noise variances are uncorrelated, the overall noise variance is merely the sum. Therefore, in general, we get:

$$SNR_N = \frac{1}{N} SNR_1$$

11.9 (a) From the problem statement we have:

$$\begin{aligned} e[n] &= x[n] - \hat{x}[n] \\ \hat{x}[n] &= \left[\frac{x[n]}{\Delta} \right] \cdot \Delta + \frac{\Delta}{2} \text{ where } [\cdot] \text{ denotes greatest integer} \\ x[n] &= \left[\frac{x[n]}{\Delta} \right] \cdot \Delta + x_f[n] = x_i[n] + x_f[n] \\ \therefore e[n] &= x_f[n] - \frac{\Delta}{2} \end{aligned}$$

In the absence of noise, $e[n]$ is a deterministic function of the input waveform $x[n]$. The strict definition of statistical independence states that two random variables or processes are independent if their joint pdf is separable. This implies that the conditional pdf produces no new information, i.e., $P(x|e) = P(x)$, or equivalently $P(e|x) = P(e)$.

We view $x[n]$ as a sample sequence from a random process. We can map the pdf for $x[n]$ into the pdf for $e[n]$ according to the quantization rule. Clearly, given $x[n]$, this information will definitely affect $P(e)$. In fact, given $x[n]$, we can determine $e[n]$ exactly; therefore $x[n]$ and $e[n]$ are not statistically independent. (Note that the range of $e[n]$ is $-\Delta/2 < e[n] < \Delta/2$).

- (b)** From the previous discussion, $e[n]$ and $x[n]$ will never be statistically independent. However, $e[n]$ can take on the appearance of a uniformly distributed white noise process if we let $\Delta \rightarrow 0$ and $n \rightarrow \infty$. This can be seen as follows. If the input, $x[n]$, is not grossly oversampled, we expect modest amplitude variations between adjacent samples. Consequently, for large Δ , the adjacent samples will often fall in the same quantization bin, and the error will have the same variations as the input signal. As $\Delta \rightarrow 0$, the adjacent samples seldom fall in the same bin (assuming $x[n]$ is not constant) and, as a result, $e[n]$ takes on the appearance of a white noise process. Note, however, that given $x[n]$, we can compute $e[n]$ and the two signals are not statistically independent in the true sense.

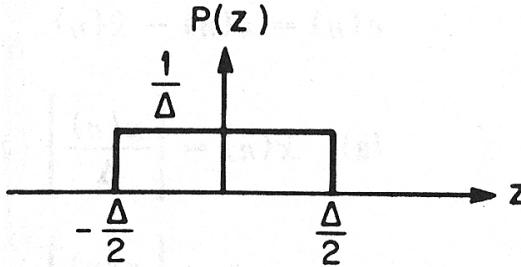


Figure P11.9.1: Probability density of $z[n]$.

- (c)** Figure P11.9.1 shows the probability density of $z[n]$ which is of the form:

$$P(z) = \frac{1}{\Delta}, \quad -\Delta/2 \leq z \leq \Delta/2$$

We now have:

$$\begin{aligned} y[n] &= x[n] + z[n] \\ \hat{y}[n] &= Q(y[n]) = Q(x[n] + z[n]) \end{aligned}$$

The error can now be written as:

$$e_1[n] = x[n] - \hat{y}[n] = x[n] - \left[\frac{x[n] + z[n]}{\Delta} \right] \cdot \Delta - \frac{\Delta}{2}$$

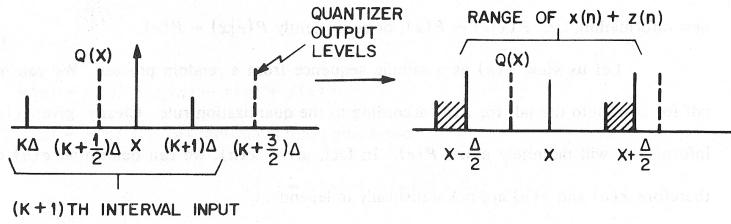


Figure P11.9.2: Range of quantizer input.

Consider an input sample, $x[n] = x$. The quantizer input has the range shown in Figure P11.9.2. In computing the quantization error, the shaded regions are equivalent. Therefore since $z[n]$ is uniformly distributed for $-\Delta/2 \leq z \leq \Delta/2$, the fractional part of $x[n] + z[n]$ can have any value between 0 and Δ , independent of the particular value of $x[n]$. Therefore we can write:

$$P(e_1|x) = P(e_1) = \frac{1}{\Delta}, \quad -\Delta/2 \leq e_1 \leq \Delta/2; \quad \text{any } x$$

and $e_1[n]$ is independent of the input $x[n]$.

(d)

$$\begin{aligned} e_1[n] &= x[n] - \hat{y}[n] \\ z[n] &= y[n] - x[n] \end{aligned}$$

We define:

$$\begin{aligned} e_2[n] &= y[n] - \hat{y}[n] \\ &= (y[n] - x[n]) + (x[n] - \hat{y}[n]) \\ &= x[n] + e_1[n] \end{aligned}$$

Since $z[n]$ is independent of $e_2[n]$, the variances add, giving:

$$\sigma_{e_1}^2 = \sigma_{e_2}^2 + \sigma_z^2 = \frac{\Delta^2}{12} + \frac{\Delta^2}{12} = \frac{\Delta^2}{6}$$

The variance of $e[n]$ is $\sigma_e^2 = \frac{\Delta^2}{12}$, so the variance of e_1 is twice the variance of e .

(e)

$$e_2[n] = y[n] - \hat{y}[n]$$

which is the quantization noise of a normal B -bit quantizer, so we get:

$$\sigma_{e_2}^2 = \frac{\Delta^2}{12}$$

11.10 (a)

$$\sigma^2[n] = \sum_{m=-\infty}^{\infty} x^2[m]h[n-m]$$

$$\text{Var}(x[n]) = E\{(x[n] - \bar{x}[n])^2\} = \sigma_x^2$$

Given $\bar{x}[n] = 0 \rightarrow \sigma_x^2 = E\{x^2[n]\}$ we get:

$$\begin{aligned}
E\{\sigma^2[n]\} &= E\left\{\sum_{m=-\infty}^{\infty} x^2[m]h[n-m]\right\} \\
&= \sum_{m=-\infty}^{\infty} E\{x^2[m]h[n-m]\} \\
&= \sum_{m=-\infty}^{\infty} \sigma_x^2 h[n-m] \\
&= \sigma_x^2 \sum_{m=-\infty}^{\infty} h[n-m] = \sigma_x^2 \sum_{m=-\infty}^{\infty} h[m]
\end{aligned}$$

(b)

$$h[n] = \begin{cases} \alpha^n & n \geq 0; \quad |\alpha| < 1 \\ 0 & n < 0 \end{cases}$$

$$E\{x^2[m]x^2[l]\} = \begin{cases} B & m = l \\ 0 & m \neq l \end{cases}$$

$$\begin{aligned}
Var(\sigma^2[n]) &= E\{(\sigma^2[n])\} - E\{(\sigma^2[n])^2\} \\
&= E\{(\sigma^2[n])^2\} - E\{\sigma^2[n]\}^2 \\
&= E\left\{\sum_{m=-\infty}^{\infty} x^2[m]h[n-m] \sum_{l=-\infty}^{\infty} x^2[l]h[n-l]\right\} \\
&\quad - \left[\sigma_x^2 \sum_{m=-\infty}^{\infty} h[m]\right] \left[\sigma_x^2 \sum_{l=-\infty}^{\infty} h[l]\right]
\end{aligned}$$

$$\begin{aligned}
Var(\sigma^2[n]) &= \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[n-m]h[n-l]E\{x^2[m]x^2[l]\} - \sigma_x^4 \sum_{m=0}^{\infty} \alpha^m \sum_{l=0}^{\infty} \alpha^l \\
&= \sum_{m=-\infty}^{\infty} h^2[n-m]B - \sigma_x^4 \sum_{m=0}^{\infty} \alpha^m \sum_{l=0}^{\infty} \alpha^l \\
&= B \sum_{m=-\infty}^{\infty} h^2[m] - \sigma_x^4 \sum_{m=0}^{\infty} \alpha^m \sum_{l=0}^{\infty} \alpha^l \\
&= B \sum_{m=0}^{\infty} \alpha^{2m} - \sigma_x^4 \sum_{m=0}^{\infty} \alpha^m \sum_{l=0}^{\infty} \alpha^l \\
&= \frac{B}{1-\alpha^2} - \frac{\sigma_x^4}{(1-\alpha)^2}
\end{aligned}$$

- (c) From part (a), as α approaches 1, $h[n]$ approaches a unit step and the summation produces an infinite value for the “short-time” energy $\sigma^2[n]$. In part (b), the expression for the variance of $\sigma^2[n]$ also approaches ∞ as $\alpha \rightarrow 1$. For $\alpha = 0$, the variance of $\sigma^2[n]$ in part (b) is a constant, equal to the true variance of $x^2[n]$. We can make the following interpretation. For $\alpha = 0$

$$h[n] = \begin{cases} 0^0 = 1 & n = 0 \\ 0^n = 0 & n \neq 0 \end{cases}$$

and we obtain the true variance. For $\alpha \neq 1$, $h[n]$ behaves nearly like a finite length window, with the window length increasing as α increases. As the window becomes longer, we are summing the weighted variances of each value of m , and in the limit as the window becomes infinitely long, the summation becomes infinite. This implies that we should use a short window, or perhaps include a normalization term for longer windows.

11.11

$$\Delta[n] = M\Delta[n-1], \quad \Delta_{\min} \leq \Delta[n] \leq \Delta_{\max}$$

$$M = \begin{cases} P & \text{if } c[n-1] = 01 \text{ or } 11 \\ 1/P & \text{if } c[n-1] = 00 \text{ or } 10 \end{cases}$$

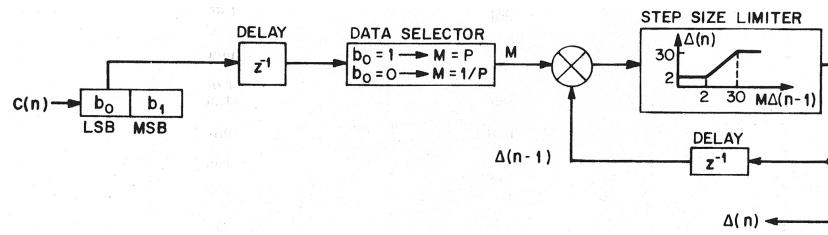


Figure P11.11.1: Block diagram of adaptation decision logic.

(a) A block diagram of the adaptation decision logic is shown in Figure P11.11.1.

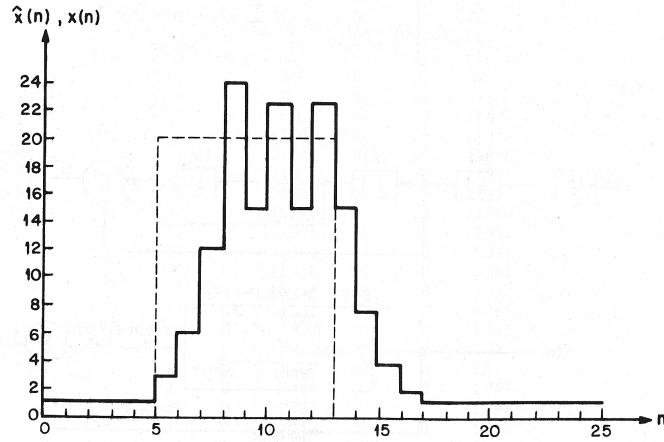


Figure P11.11.2: Waveform plot of $x[n]$ and $\hat{x}[n]$.

(b) Given $x[n]$ of the form:

$$x[n] = \begin{cases} 0 & n < 5 \\ 20 & 5 \leq n \leq 13 \\ 0 & 13 < n \end{cases}$$

and using the parameters $\Delta_{\min} = 2$, $\Delta_{\max} = 30$, $\Delta[0] = \Delta_{\min} = 2$, $c[0] = 00$, and $P = 2$ we derive the coding parameters shown in the table below:

n	M	$\Delta[n]$	$c[n]$	$x[n]$	$\hat{x}[n]$
0	0.5	2	00	0	1
1	0.5	2	00	0	1
2	0.5	2	00	0	1
3	0.5	2	00	0	1
4	0.5	2	00	0	1
5	0.5	2	01	20	3
6	2	4	01	20	6
7	2	8	01	20	12
8	2	16	01	20	24
9	2	30	00	20	15
10	0.5	15	01	20	22.5
11	2	30	00	20	15
12	0.5	15	01	20	22.5
13	2	30	00	20	15
14	0.5	15	00	0	7.5
15	0.5	7.5	00	0	3.75
16	0.5	3.75	00	0	1.875
17	0.5	2	00	0	1
18-25	0.5	2	00	0	1

A plot of the variation of $x[n]$ and $\hat{x}[n]$ is shown in Figure P11.11.2.

11.12 (a) A block diagram of the step-size adaptation algorithm is shown in Figure P11.12.1.

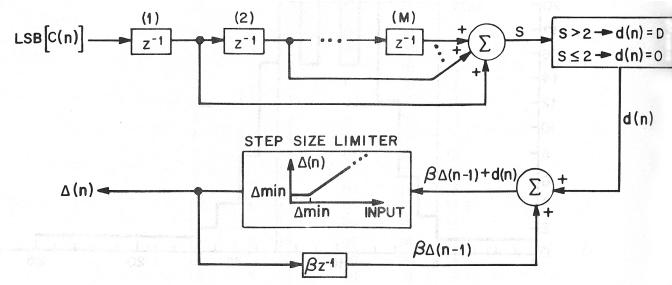


Figure P11.12.1: Block diagram of the step-size adaptation algorithm.

(b) The maximum step size is reached when the LSB is always 1. Then:

$$\Delta[n] = \beta\Delta[n-1] + D \quad \forall n$$

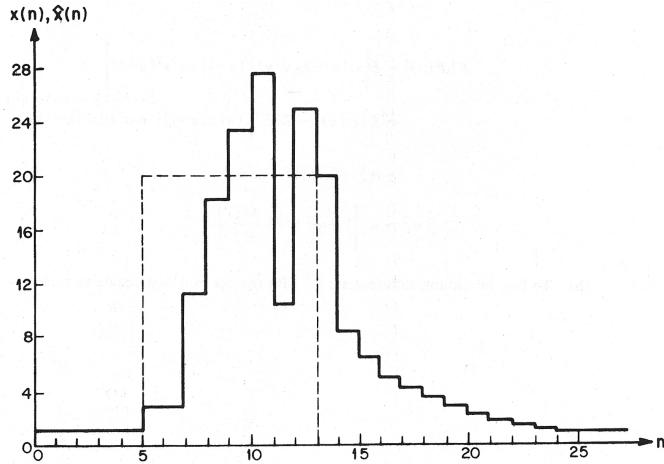
$$\Delta_{\max} = \beta\Delta_{\max} + D$$

$$\Delta_{\max} = \frac{D}{1-\beta}$$

(c) For the input:

$$x[n] = \begin{cases} 0 & n < 5 \\ 20 & 5 \leq n \leq 13 \\ 0 & 13 < n \end{cases}$$

with parameters $\beta = 0.8$ and $D = 6$, with initial conditions ($n = 0$) of $\Delta[0] = 0$ and $c[0] = 00$ we get the results shown in the table below, and plotted in Figure P11.12.2.

Figure P11.12.2: Plot of $x[n]$ and $\hat{x}[n]$ over the range $0 \leq n \leq 25$.

n	$d[n]$	$\Delta[n]$	$c[n]$	$x[n]$	$\hat{x}[n]$
0	0	2	00	0	1
1	0	2	00	0	1
2	0	2	00	0	1
3	0	2	00	0	1
4	0	2	00	0	1
5	0	2	01	20	3
6	0	2	01	20	3
7	6	7.6	01	20	11.4
8	6	12.08	01	20	18.2
9	6	15.664	01	20	23.496
10	6	18.531	01	20	27.797
11	6	20.825	00	20	10.412
12	0	16.66	01	20	24.99
13	0	13.328	01	20	19.992
14	6	16.662	00	0	8.331
15	0	13.33	00	0	6.665
16	0	10.664	00	0	5.332
17	0	8.531	00	0	4.265
18	0	6.82	00	0	3.412
19	0	5.46	00	0	2.73
20	0	4.368	00	0	2.18
21	0	3.494	00	0	1.747
22	0	2.795	00	0	1.397
23	0	2.236	00	0	1.118
24	0	2	00	0	1
25	0	2	00	0	1

- (d) For a sampling rate of $F_s = 1/T = 10$ kHz, 10 msec corresponds to $NT = 10^{-2}$, or $N = \frac{10^{-2}}{T}$ samples. The required value for β is determined from:

$$\beta^N = e^{-1} = \beta^{10^{-2}/T}$$

or

$$\beta = e^{-T \cdot 100}$$

11.13

$$\tilde{x}[n] = \alpha x[n - 1]$$

We assume $x[n]$ is zero mean and stationary.

(a)

$$d[n] = x[n] - \tilde{x}[n]$$

$$\begin{aligned} \text{Var}\{d[n]\} &= E\{(d[n] - E\{d[n]\})^2\} \\ &= E\{d^2[n]\} - E^2\{d[n]\} \end{aligned}$$

where

$$\begin{aligned} E\{d^2[n]\} &= E\{x^2[n] - 2\alpha x[n]x[n - 1] + \alpha^2 x^2[n - 1]\} \\ &= E\{x^2[n]\} - 2\alpha E\{x[n]x[n - 1]\} + \alpha^2 E\{x^2[n - 1]\} \\ &= \sigma_x^2 - 2\alpha\phi[1] + \alpha^2\sigma_x^2 \end{aligned}$$

$$\therefore \sigma_d^2 = \sigma_x^2 \left[1 + \alpha^2 - 2\alpha \frac{\phi[1]}{\sigma_x^2} \right]$$

(b) To find minimum, differentiate σ_d^2 with respect to α and equate to zero, giving:

$$\begin{aligned} \frac{\partial \sigma_d^2}{\partial \alpha} &= \sigma_x^2 \left[2\alpha - 2 \frac{\phi[1]}{\sigma_x^2} \right] = 0 \\ \alpha &= \frac{\phi[1]}{\sigma_x^2} \end{aligned}$$

Note that this represents a minimum since the second derivative is positive.

(c) $\sigma_{d_{\min}}^2$ is obtained by substituting the value of α from part (b), giving

$$\begin{aligned} \sigma_{d_{\min}}^2 &= \sigma_x^2 \left[1 + \frac{\phi^2[1]}{\sigma_x^4} - 2 \frac{\phi^2[1]}{\sigma_x^4} \right] \\ &= \sigma_x^2 \left[1 - \frac{\phi^2[1]}{\sigma_x^4} \right] = \sigma_x^2(1 - \rho^2[1]) \end{aligned}$$

(d)

$$\begin{aligned} \sigma_d^2 \leq \sigma_x^2 &\implies \sigma_x^2 \left[1 + \alpha^2 - 2\alpha \frac{\phi[1]}{\sigma_x^2} \right] \leq \sigma_x^2 \\ 1 + \alpha^2 - 2\alpha \frac{\phi[1]}{\sigma_x^2} &\leq 1 \\ \alpha^2 - \frac{2\alpha\phi[1]}{\sigma_x^2} &\leq 0 \\ |\alpha| &\leq 2 \frac{\phi[1]}{\sigma_x^2} \end{aligned}$$

11.14

$$d[n] = x[n] = x[n - n_0]$$

$$\begin{aligned} \text{Var}\{d[n]\} &= E\{(x[n] - x[n - n_0])^2\} \\ &= E\{x^2[n] - 2x[n]x[n - n_0] + x^2[n - n_0]\} \\ &= E\{x^2[n]\} - 2E\{x[n]x[n - n_0]\} + E\{x^2[n - n_0]\} \\ &= \phi[0] - 2\phi[n_0] + \phi[0] \\ &= 2(\phi[0] - \phi[n_0]) \end{aligned}$$

(a) The condition that must be met is:

$$\begin{aligned} \text{Var}\{d[n]\} < \text{Var}\{x[n]\} &\implies 2(\phi[0] - \phi[n_0]) < \phi[0] \\ \phi[0] - \phi[n_0] &< \frac{\phi[0]}{2} \\ \phi[n_0] &> \frac{\phi[0]}{2} \end{aligned}$$

(b)

$$d[n] = x[n] - \alpha x[n - n_0]; \quad \alpha = \frac{\phi[n_0]}{\phi[0]}$$

$$\begin{aligned} \text{Var}\{d[n]\} &= \phi[0] - 2\alpha\phi[n_0] + \alpha^2\phi[0] \\ &= \phi[0] - 2\frac{\phi^2[n_0]}{\phi[0]} + \frac{\phi^2[n_0]}{\phi[0]} \end{aligned}$$

$$\text{Var}\{d[n]\} < \text{Var}\{x[n]\} \implies \phi[0] - \frac{\phi^2[n_0]}{\phi[0]} < \phi[0]$$

$$\phi^2[0] - \phi^2[n_0] < \phi^2[0]$$

$$\phi^2[n_0] > 0$$

11.15 In Chapter 11 we have the following relations:

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

$$E\{(x[n] - \tilde{x}[n])\tilde{x}[n - j]\} = E\{d[n]\tilde{x}[n - j]\} \quad 1 \leq j \leq N$$

We want to show that:

$$E\{(x[n] - \tilde{x}[n])\tilde{x}[n]\} = 0$$

For the second occurrence of $\tilde{x}[n]$ in the above equation, substitute the expression:

$$E\{(x[n] - \tilde{x}[n]) \sum_{k=1}^p \alpha_k \hat{x}[n-k]\} = \sum_{k=1}^p \alpha_k E\{(x[n] - \tilde{x}[n]) \hat{x}[n-k]\}$$

But $E\{(x[n] - \tilde{x}[n]) \hat{x}[n-k]\} = 0$:

$$\therefore E\{(x[n] - \tilde{x}[n]) \tilde{x}[n]\} = E\{d[n] \tilde{x}[n]\} = 0$$

11.16 (a)

$$d[n] = x[n] - \alpha_1 \hat{x}[n-1]$$

$$\sigma_d^2 = E\{d^2[n]\} - E^2\{d[n]\} = E\{d^2[n]\}$$

where $E\{d^2[n]\} = E\{x^2[n] - 2\alpha_1 x[n] \hat{x}[n-1] + \alpha_1^2 \hat{x}[n-1]^2\}$ and $\hat{x}[n-1] = x[n-1] + e[n-1]$. Thus we have:

$$\sigma_d^2 = E\{x^2[n]\} - 2\alpha_1 E\{x[n](x[n-1] + e[n-1])\} + \alpha_1^2 E\{(x[n-1] + e[n-1])^2\}$$

We assume that $x[n]$ and $e[n-1]$ are uncorrelated so $E\{x[n]e[n-1]\} = 0$. Thus we have:

$$\begin{aligned} \sigma_d^2 &= \sigma_x^2 - 2\alpha_1 \rho_x[1] \sigma_x^2 + \alpha_1^2 \sigma_x^2 + \alpha_1^2 \sigma_e^2 \\ &= \sigma_x^2 (1 - 2\alpha_1 \rho_x[1] + \alpha_1^2) + \alpha_1^2 \sigma_e^2 \end{aligned}$$

(b) Solve for σ_x^2 from the results of part (a) as:

$$\sigma_x^2 = \frac{\sigma_d^2 - \alpha_1^2 \sigma_e^2}{(1 - 2\alpha_1 \rho_x[1] + \alpha_1^2)}$$

Therefore:

$$\begin{aligned} G_p = \frac{\sigma_x^2}{\sigma_d^2} &= \frac{\sigma_d^2 - \alpha_1^2 \sigma_e^2}{(1 - 2\alpha_1 \rho_x[1] + \alpha_1^2) \sigma_d^2} \\ &= \frac{1 - \alpha_1^2 \sigma_e^2 / \sigma_d^2}{(1 - 2\alpha_1 \rho_x[1] + \alpha_1^2)} = \frac{1 - \frac{\alpha_1^2}{SNR_Q}}{(1 - 2\alpha_1 \rho_x[1] + \alpha_1^2)} \end{aligned}$$

11.17 The fundamental equations of the differential quantizer are the following:

$$\begin{aligned} \hat{d}[n] &= d[n] + e[n] \quad \text{quantization of } d[n] \\ \tilde{s}[n] &= \sum_{k=1}^p \alpha_k \hat{s}[n-k] \quad \text{linear prediction on } \hat{s}[n] \\ \hat{s}[n] &= \tilde{s}[n] + \hat{d}[n] \quad \text{feedback loop to linear predictor} \\ d[n] &= s[n] - \tilde{s}[n] \quad \text{definition of differential signal} \end{aligned}$$

(a) Using the above equations we get:

$$\begin{aligned}\hat{s}[n] &= \tilde{s}[n] + \hat{d}[n] \\ &= \tilde{s}[n] + d[n] + e[n] \\ &= \tilde{s}[n] + s[n] - \tilde{s}[n] + e[n] \\ &= s[n] + e[n]\end{aligned}$$

(b) We derive the relationship from the basic equations as:

$$\begin{aligned}s[n] &= \hat{s}[n] - e[n] \\ &= \tilde{s}[n] + \hat{d}[n] - e[n] \\ &= \sum_{k=1}^p \alpha_k(s[n-k] + e[n-k]) + d[n] + e[n] - e[n] \\ &= d[n] + \sum_{k=1}^p \alpha_k(s[n-k] + e[n-k])\end{aligned}$$

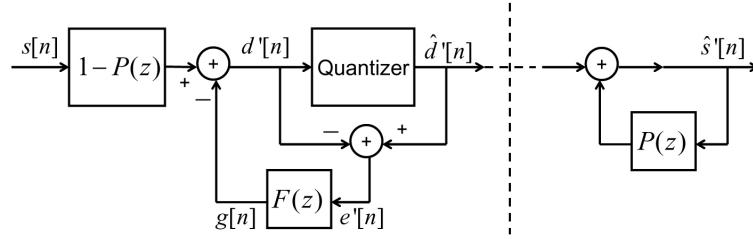


Figure P11.17.1: Block diagram of differential quantizer with noise shaping.

(c) From the block diagram of Figure P11.17.1 we get the relations:

$$\begin{aligned}d'[n] &= s[n] - \sum_{k=1}^p \alpha_k s[n-k] - \sum_{k=1}^p \beta_k e'[n-k] \\ e'[n] &= \hat{d}'[n] - d'[n]\end{aligned}$$

(d) To determine how to choose $F(z)$ so that $d'[n] = d[n]$ we solve for $d'[n]$ and $d[n]$ as:

$$\begin{aligned}d[n] &= s[n] - \sum_{k=1}^p \alpha_k(s[n-k] + e[n-k]) \\ d'[n] &= s[n] - \sum_{k=1}^p \alpha_k s[n-k] - \sum_{k=1}^p \beta_k e'[n-k]\end{aligned}$$

Recognizing that $e'[n] = e[n]$ since $e'[n]$ is the difference between the output and the input to the quantizer, which is, by definition, the quantity $e[n]$, we can see that when we have the condition $\beta_k = \alpha_k \forall k$ then $d'[n] = d[n]$. Similarly we see that $\hat{d}'[n] = \hat{d}[n]$ and $\hat{s}'[n] = \hat{s}[n]$.

(e) From the block diagram of Figure P11.17.1 we can write the relationship for $d'[n]$ as:

$$d'[n] = s[n] - \sum_{k=1}^p \alpha_k s[n-k] - \sum_{k=1}^p \beta_k e'[n-k]$$

and from the receiver side of the quantization system we can show that:

$$\hat{s}'[n] = \hat{d}'[n] + \sum_{k=1}^p \alpha_k \hat{s}'[n-k]$$

and by the definition of the quantizer we get:

$$\hat{d}'[n] = d'[n] + e'[n]$$

and by combining the results of the above equations we get:

$$\begin{aligned}\hat{s}'[n] &= d'[n] + e'[n] + \sum_{k=1}^p \alpha_k (\hat{e}'[n-k] + s[n-k]) \\ &= s[n] - \sum_{k=1}^p \alpha_k s[n-k] - \sum_{k=1}^p \beta_k e'[n-k] + e'[n] \\ &\quad + \sum_{k=1}^p \alpha_k \hat{e}'[n-k] + \sum_{k=1}^p \alpha_k s[n-k] \\ &= s[n] + \sum_{k=1}^p \alpha_k \hat{e}'[n-k] + e'[n] - \sum_{k=1}^p \beta_k e'[n-k]\end{aligned}$$

(f) We can rewrite the results in part (e) giving:

$$\begin{aligned}\hat{s}'[n] - s[n] &= \hat{e}'[n] = \sum_{k=1}^p \alpha_k \hat{e}'[n-k] + e'[n] - \sum_{k=1}^p \beta_k e'[n-k] \\ \hat{E}'(z) &= \sum_{k=1}^p \alpha_k z^{-k} E'(z) + E'(z) - \sum_{k=1}^p \beta_k z^{-k} E'(z) \\ \frac{\hat{E}'(z)}{E'(z)} &= \frac{1 - \sum_{k=1}^p \beta_k z^{-k}}{1 - \sum_{k=1}^p \alpha_k z^{-k}} = \text{Noise Shaping Filter}\end{aligned}$$

(g) The quantization noise of the quantizer, $e'[n]$, does not appear at the receiver output. Rather we get the result:

$$\hat{s}'[n] = s[n] + \hat{e}'[n]$$

where $\hat{e}'[n]$ is related to $e'[n]$ by the transfer function of part (f). Thus the noise can be spectrally shaped to take advantage of the ear's sensitivity (masking ability). The goal is to hide the noise under the formant resonances.

11.18

$$\Delta[n] = \begin{cases} \beta \Delta[n-1] + D_2 & \text{if } c[n] = c[n-1] = c[n-2] \\ \beta \Delta[n-1] + D_1 & \text{otherwise} \end{cases}$$

where $0 < \beta$ and $0 < D_1 \ll D_2$.

- (a) In the “steady-state” condition, $\Delta[n] = \Delta[n - 1] = \Delta_{\max}$

$$\Delta_{\max} = \beta\Delta_{\max} + D_2$$

$$\Delta_{\max} = \frac{D_2}{1 - \beta}$$

- (b) In the “idle channel” mode $\Delta[n] = \Delta[n - 1] = \Delta_{\min}$

$$\Delta_{\min} = \beta\Delta_{\min} + D_1$$

$$\Delta_{\min} = \frac{D_1}{1 - \beta}$$

11.19

$$\Delta[n] = M\Delta[n - 1]; \quad \Delta_{\min} \leq \Delta[n] \leq \Delta_{\max}$$

$$M = \begin{cases} P & \text{if } c[n] = c[n - 1] \\ 1/P & \text{if } c[n] \neq c[n - 1] \end{cases}$$

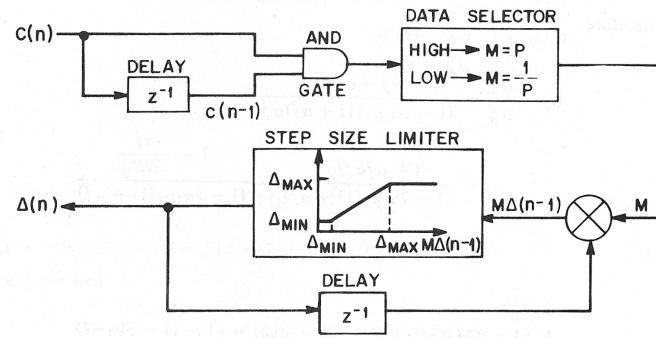


Figure P11.19.1: Block diagram of the step-size logic.

- (a) A block diagram of the step-size logic is given in Figure P11.19.1.

- (b) Given input of the form:

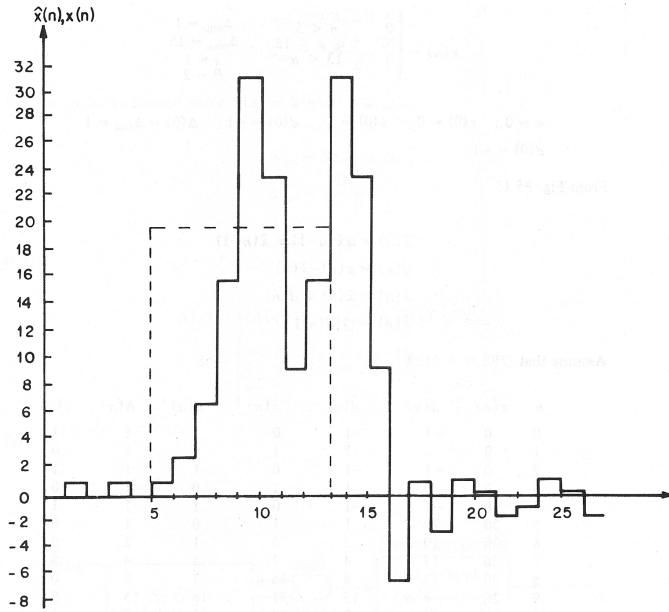
$$x[n] = \begin{cases} 0 & n < 5 \\ 20 & 5 \leq n \leq 13 \\ 0 & 13 < n \end{cases}$$

with parameters $\Delta_{\min} = 1$, $\Delta_{\max} = 15$, $\alpha = 1$, $P = 2$ and initial conditions $n = 0$, $x[0] = 0$, $\hat{x}[0] = 1$, $d[0] = -1$, $\Delta[0] = \Delta_{\min} = 1$ and $\hat{d}[0] = -1$.

The basic equations for the adaptive delta modulator are:

$$\begin{aligned} \tilde{x}[n] &= \alpha\hat{x}[n - 1] = \hat{x}[n - 1] \\ d[n] &= x[n] - \tilde{x}[n] \\ \hat{x}[n] &= \tilde{x}[n] + \hat{d}[n] \\ \hat{d}[n] &= Q(d[n]) \end{aligned}$$

Assuming that $Q[0] = +\Delta[n]$ we get the waveform samples given in the table below and plotted in Figure P11.19.2.

Figure P11.19.2: Plots of $x[n]$ and $\hat{x}[n]$ for $0 \leq n \leq 25$.

n	$x[n]$	$d[n]$	$\hat{d}[n]$	$\hat{x}[n]$	$\tilde{x}[n]$	$\Delta[n]$	$c[n]$
0	0	-1	-1	0	1	1	1
1	0	0	1	1	0	1	0
2	0	-1	-1	0	1	1	1
3	0	0	1	1	0	1	0
4	0	-1	-1	0	1	1	1
5	20	20	1	1	0	1	0
6	20	19	2	3	1	2	0
7	20	17	4	7	3	4	0
8	20	13	8	16	7	8	0
9	20	4	15	31	16	15	0
10	20	-11	-7.5	23.5	31	7.5	1
11	20	-3.5	-15	8.5	23.5	15	1
12	20	11.5	7.5	15	8.5	7.5	0
13	20	4	15	31	16	15	0
14	0	-31	-7.5	23.5	31	7.5	1
15	0	-23.5	-15	8.5	23.5	15	1
16	0	-8.5	-15	-6.5	8.5	15	1
17	0	6.5	7.5	1	-6.5	7.5	0
18	0	-1	-3.75	-2.75	1	3.75	1
19	0	2.75	1.875	1.125	-2.75	1.875	0
20	0	-1.125	-1	0.125	1.125	1	1
21	0	-0.125	-2	-1.875	0.125	2	1
22	0	1.875	1	-0.875	-1.875	1	0
23	0	0.875	2	1.125	-0.875	2	0
24	0	-1.125	-1	0.125	1.125	1	1
25	0	-0.125	-2	-1.1875	0.125	2	1

11.20

$$x[n] = 0.1 \cos\left(\frac{\pi n}{4}\right)$$

Let \hat{x}_1 denote coder number 1 and $\hat{x}[n]$ denote coder number 2.

- (a) We assume a zero input to the quantizer produces an output of 1. The resulting coder signals for coders number 1 and number 2 are given in the table below.

n	$x[n]$	$\hat{x}_1[n]$	\tilde{x}	$d[n]$	$\hat{d}[n]$	$\hat{x}_2[n]$
0	0.1	1	0	0.1	1	1
1	0.0707	1	1	-0.929	-1	0
2	0	1	0	0	1	1
3	-0.0707	-1	1	-1.071	-1	0
4	-0.1	-1	0	-0.1	-1	-1
5	-0.0707	-1	-1	0.929	1	0
6	0	1	0	0	1	1
7	0.0707	1	1	-0.929	-1	0
8	0.1	1	0	0.1	1	1
9	0.0707	1	1	-0.929	-1	0
10	0	1	0	0	1	1
11	-0.0707	-1	1	-1.071	-1	0
12	-0.1	-1	0	-0.1	-1	-1
13	-0.0707	-1	-1	0.929	1	0
14	0	1	0	0	1	1
15	0.0707	1	1	-0.929	-1	0
16	0.1	1	0	0.1	1	1
17	0.0707	1	1	-0.929	-1	0
18	0	1	0	0.	1	1
19	-0.0707	-1	1	-1.071	-1	0
20	-0.1	-1	0	-0.1	-1	-1

- (b) The idle channel noise in the first coder would probably be judged more objectional. The noise in the second coder includes more variation in the waveshape and is higher in frequency in the sense that the output changes each sample. The sampling rate will also tend to effect which of the noise waveforms is more objectionable.

11.21

$$y[n] = x[n] + e_1[n]$$

$$\begin{aligned} \hat{y}[n] &= y[n] + e_2[n] = x[n] + e_1[n] + e_2[n] \\ &= \text{signal} + \text{noise} \end{aligned}$$

- (a) The noise variance is merely the sum of the individual noise variances if they are uncorrelated, i.e.,

$$\begin{aligned} \text{Var}\{e_1[n] + e_2[n]\} &= \sigma_{e_1}^2 + \sigma_{e_2}^2 \\ \therefore \text{SNR} &= \frac{\text{Var}\{x[n]\}}{\text{Var}\{\text{noise}\}} = \frac{\sigma_x^2}{\sigma_{e_1}^2 + \sigma_{e_2}^2} \end{aligned}$$

(b) Note that:

$$\sigma_y^2 = \sigma_x^2 + \sigma_{e_1}^2; \quad SNR_1 = \frac{\sigma_x^2}{\sigma_{e_1}^2}; \quad SNR_2 = \frac{\sigma_y^2}{\sigma_{e_2}^2} = \frac{\sigma_x^2 + \sigma_{e_1}^2}{\sigma_{e_2}^2}$$

$$SNR = \frac{\frac{\sigma_x^2}{\sigma_{e_1}^2}}{1 + \frac{\sigma_{e_2}^2}{\sigma_{e_1}^2}}$$

We note that:

$$\frac{\sigma_{e_2}^2}{\sigma_{e_1}^2} = \frac{(\sigma_{e_1}^2 + \sigma_x^2)\sigma_{e_2}^2}{(\sigma_{e_1}^2 + \sigma_x^2)\sigma_{e_1}^2} = \frac{\sigma_{e_1}^2 + \sigma_x^2}{\sigma_{e_1}^2} \cdot \frac{1 + SNR_1}{SNR_2}$$

Substituting into the expression for SNR gives:

$$SNR = \frac{\frac{\sigma_x^2}{\sigma_{e_1}^2}}{1 + \left[\frac{1 + SNR_1}{SNR_2} \right]} = \frac{SNR_1}{1 + \left[\frac{1 + SNR_1}{SNR_2} \right]}$$

11.22 (a) Since it is postulated that $x[n]$ and $\epsilon[n]$ are independent, we get the following result for the mean of $f[n] = x[n] \cdot \epsilon[n]$:

$$E\{f[n]\} = E\{x[n] \cdot \epsilon[n]\} = E\{x[n]\} \cdot E\{\epsilon[n]\} = 0$$

since the mean of both $x[n]$ and $\epsilon[n]$ is zero.

(b) The variance of $f[n]$ can be calculated as:

$$E\{f^2[n]\} = E\{x^2[n] \cdot \epsilon^2[n]\} = E\{x^2[n]\} \cdot E\{\epsilon^2[n]\} = \sigma_x^2 \cdot \sigma_\epsilon^2$$

11.23 Theoretically speaking, the approximate ordering of the vocoders from highest to lowest quality would be:

1. Phase Vocoder
2. LPC Vocoder
3. Homomorphic Vocoder
4. Channel Vocoder
5. Parallel Formant Vocoder
6. Serial Formant Vocoder

The phase vocoder would have the highest quality since it does not require a pitch detector to obtain the excitation waveform. Thus, theoretically, a phase vocoder can perfectly reproduce any bandlimited waveform.

The LPC vocoder would have the highest quality of the remaining vocoders since the LPC model explicitly minimizes a squared-error criterion and therefore produces an “optimal” fit to the speech waveform. The assumed all-pole model is the main theoretical limitation of the LPC vocoder since some sounds require zeros as well as poles.

The homomorphic vocoder and the channel vocoder are next highest in quality since in these vocoders the speech spectral information is preserved in the coding of the cepstrum (the low time part) in the homomorphic vocoder, and the smoothed spectrum for the channel vocoder. The degradations in these models are due to finite analysis windows and assumptions about the separation and interactions of the source (pitch) and system signals.

The lowest quality systems are the formant vocoders which assume the time-varying speech spectra can be coded in terms of smoothly varying resonances or formants. Thus, these vocoders are totally reliant on all aspects of the speech model. The parallel formant vocoder gives better quality than the serial formant vocoder since more information about spectral levels at the formants is preserved in the parallel formant vocoder than in the series formant vocoder.

11.24 Using Fourier transforms we solve for the various components of the two-band subband coder as:

$$\begin{aligned} Y_0(e^{j\omega}) &= \frac{1}{2} \left[H_0(e^{j\omega/2})X(e^{j\omega/2}) + H_0(e^{j(\omega-2\pi)/2})X(e^{j(\omega-2\pi)/2}) \right] \\ Y_1(e^{j\omega}) &= \frac{1}{2} \left[H_1(e^{j\omega/2})X(e^{j\omega/2}) + H_1(e^{j(\omega-2\pi)/2})X(e^{j(\omega-2\pi)/2}) \right] \\ \hat{Y}(e^{j\omega}) &= \frac{G_0(e^{j\omega})}{2} \left[H_0(e^{j\omega})X(e^{j\omega}) + H_0(e^{j(\omega-\pi)})X(e^{j(\omega-\pi)}) \right] \\ &\quad + \frac{G_1(e^{j\omega})}{2} \left[H_1(e^{j\omega})X(e^{j\omega}) + H_1(e^{j(\omega-\pi)})X(e^{j(\omega-\pi)}) \right] \\ &= \frac{X(e^{j\omega})}{2} \left[G_0(e^{j\omega})H_0(e^{j\omega}) + G_1(e^{j\omega})H_1(e^{j\omega}) \right] \\ &\quad + \frac{X(e^{j(\omega-\pi)})}{2} \left[G_0(e^{j\omega})H_0(e^{j(\omega-\pi)}) + G_1(e^{j\omega})H_1(e^{j(\omega-\pi)}) \right] \end{aligned}$$

The conditions for $\hat{Y}(e^{j\omega}) = X(e^{j\omega})$ are:

$$\begin{aligned} \frac{1}{2} \left[G_0(e^{j\omega})H_0(e^{j\omega}) + G_1(e^{j\omega})H_1(e^{j\omega}) \right] &= 1 \quad \forall \omega \\ \frac{1}{2} \left[G_0(e^{j\omega})H_0(e^{j(\omega-\pi)}) + G_1(e^{j\omega})H_1(e^{j(\omega-\pi)}) \right] &= 0 \quad \forall \omega \end{aligned}$$

11.25 (a) To perform 8-bit quantization on the speech file, we first normalize the peak level to 1.0, then we scale the speech file by a factor of $2^8 = 256$, round to the nearest integer and divide by a factor of $2^8 = 256$. The error signal is defined as the difference between the unquantized speech and the quantized speech samples. A plot of the error signal versus sample index is given in Figure P11.25.1.

Using 8-bit quantization, we expect a long-term SNR of about $6B - 7.2$ or about 40.8 dB. The short-time SNR can't be much more than the long-term average. The variability in SNR measurements comes from using a short window duration of $N = 100$ samples. The plot of short-time SNR versus sample index is given in the upper curve in Figure P11.25.2.

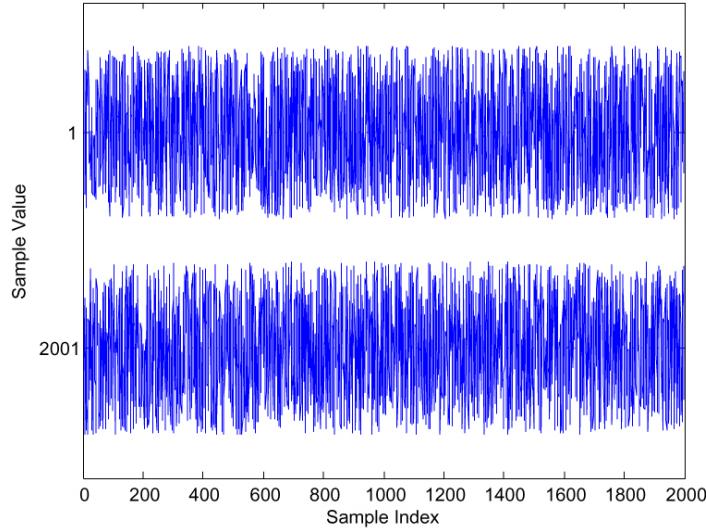


Figure P11.25.1: Plot of first 4000 samples of error signal for 8-bit quantization of the speech samples in the file “test_8k”.

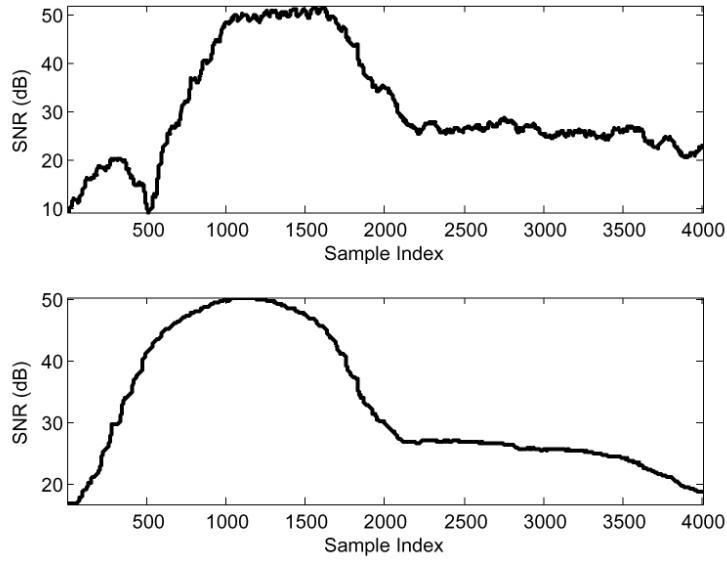


Figure P11.25.2: Plots of short-time SNR values using windows of duration $N = 100$ and $N = 600$ samples.

- (b) The plot of short-time SNR versus sample index using a window of duration $N = 600$ is shown in the lower curve in Figure P11.25.2. The resulting curve is a lot smoother than the curve for a value of $N = 100$.
- (c) For a μ -law quantizer there will be less variation in SNR level using a window of duration $N = 100$, and a lower maximum of the short-time SNR curve.

Chapter 12

Frequency-Domain Coding of Speech and Audio

12.1 (a) We begin with simple complex exponential modulation for which the DTFT is:

$$x_k[n] = x_0[n]e^{j\omega_k n} \iff X_k(e^{j\omega}) = X_0(e^{j(\omega-\omega_k)}).$$

Assuming $\omega_k = 2\pi k/N$, and that $|X_0(e^{j\omega})| = 0$ for $|\omega| \geq \pi/N$, we would have a plot like Figure P12.1.1 below.

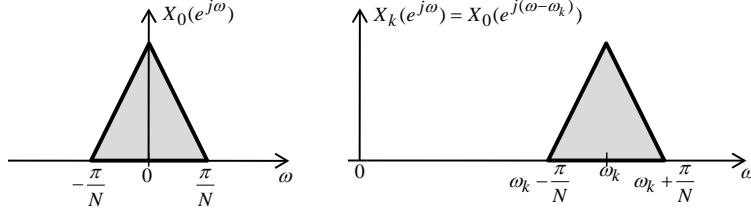


Figure P12.1.1: Illustration of spectral shift for $x_k[n] = x_0[n]e^{j\omega_k n}$.

(b) For decimation by N ,

$$v_k[n] = x_k[nN] = x_0[nN]e^{j(2\pi k/N)nN} = x_0[nN]e^{j2\pi nk} = x_0[nN].$$

Thus the decimation removes the exponential factor—in effect, demodulating and decimating at the same time. Therefore, we have

$$v_k[n] = x_k[nN] = x_0[nN] \iff V_k(e^{j\omega}) = \frac{1}{N} \sum_{m=0}^{N-1} X_0(e^{j(\omega-2\pi m)/N}).$$

The function $X_0(e^{j\omega/N})$ is a frequency-expanded by N version of $X_0(e^{j\omega})$ with the original spectrum spread out over $-\pi < \omega < \pi$, followed by a wide gap of zero from π to $2\pi N - \pi$, and periodic with period $2\pi N$. This is depicted in Figure P12.1.2. Then $N - 1$ copies $X_0(e^{j(\omega-2\pi m)/N})$ are shifted by multiples of 2π to fill in the zero gap as shown in Figure P12.1.3. The end result is therefore

$$V_k(e^{j\omega}) = \frac{1}{N} X_0(e^{j\omega/N}) \quad -\pi \leq \omega \leq \pi, \quad \text{and with period } 2\pi.$$



Figure P12.1.2: Illustration of frequency-expansion of spectrum.

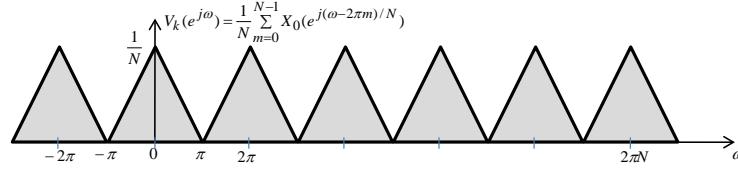


Figure P12.1.3: Spectrum after downsampling a complex bandpass signal.

- (c) Now consider the real bandpass signal $x_k[n] = x_0[n] \cos(\omega_k n + \phi_k)$. The DTFT is

$$x_k[n] = x_0[n] \cos(\omega_k n + \phi_k) \iff X_k(e^{j\omega}) = \frac{e^{j\phi_k}}{2} X_0(e^{j(\omega-\omega_k)}) + \frac{e^{-j\phi_k}}{2} X_0(e^{j(\omega+\omega_k)}).$$

Figure P12.1.4 depicts this for the same conditions as in part (a).

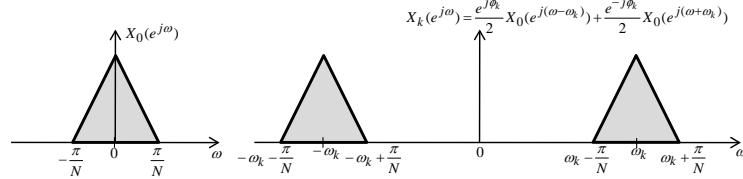


Figure P12.1.4: Spectrum after cosine modulation.

- (d) Now consider downsampling the real bandpass signal.

$$v_k[n] = x_k[nN] = x_0[nN] \cos[(2\pi k/N)nN + \phi_k] = x_0[nN] \cos(2\pi kn + \phi_k) = x_0[nN] \cos(\phi_k).$$

Thus the decimation demodulates and decimates at the same time, but with a constant scale factor $\cos(\phi_k)$. This is identical to what would happen in the following process: (1) multiply $x_k[n]$ by $\cos(\omega_k n)$ (rather than $\cos(\omega_k n + \phi_k)$ as required for synchronous demodulation); (2) ideal lowpass filter with cutoff π/N ; and (3) decimate the resulting output by N . There is no phase synchrony in either case, which leads to the $\cos(\phi_k)$ multiplicative factor. Therefore, we have

$$v_k[n] = \cos(\phi_k) x_0[nN] \iff V_k(e^{j\omega}) = \frac{\cos(\phi_k)}{N} \sum_{m=0}^{N-1} X_0(e^{j(\omega-2\pi m)/N}).$$

This is the same result as in part (b) except for the scale factor $\cos(\phi_k)$, so it follows that

$$V_k(e^{j\omega}) = \frac{\cos(\phi_k)}{N} X_0(e^{j\omega/N}) \quad -\pi \leq \omega \leq \pi, \quad \text{and with period } 2\pi.$$

Clearly, $\phi_k = \pi/2$ gives $V_k(e^{j\omega}) = 0$.

- (e) If $X_0(e^{j\omega})$ is only approximately bandlimited to the band $|\omega| < \pi/N$, then there will be aliasing in the overlap regions of the shifted copies in

$$\sum_{m=0}^{N-1} X_0(e^{j(\omega-2\pi m)/N}).$$

This aliasing is canceled in the MPEG filterbank.

- 12.2** The MPEG analysis/synthesis system is shown in Figure P12.2a. The analysis filter bank has $N = 32$ filters whose outputs are computed once for every N samples at the input rate; i.e, the outputs are maximally-decimated, meaning that N samples of input are represented by one sample at each of the N filter outputs. The synthesis filter bank does interpolation so that the N (quantized) analysis filter bank output samples are transformed into N samples of the overall output at the same sampling rate as the input. The MPEG standard document simply gives a weighting function (that we call $c[m]$) and the flow chart in Figure 12.2c for implementing the analysis filter bank (and another flow chart for implementing the synthesis filter bank). This problem is a guided tour through the MPEG system. It takes a detailed look at the way that the filters provide nearly perfect reconstruction (from unquantized samples) and also at how the flowchart of Figure 12.2c implements the analysis filter bank.

- (a) The impulse responses of the MPEG filters are

$$h_k[n] = 2h[n] \cos(\omega_k n - \phi_k) = h[n]e^{j(\omega_k n - \phi_k)} + h[n]e^{-j(\omega_k n - \phi_k)} \quad k = 0, 1, \dots, N-1,$$

where $\omega_k = 2\pi(2k+1)/(4N)$ and $\phi_k = \omega_k N/2$. From the modulation theorem for the DTFT, it follows that

$$H_k(e^{j\omega}) = e^{-j\phi_k} H(e^{j(\omega-\omega_k)}) + e^{j\phi_k} H(e^{j(\omega+\omega_k)}).$$

- (b) Since we assume that the filter impulse response length is $L = 513$, but $h[512] = 0$, the filter outputs are

$$u_k[n] = \sum_{m=0}^{511} x[n-m]h_k[m] \quad k = 0, 1, \dots, N-1,$$

so the downsampled filter outputs are

$$s_k[n] = u_k[nN] = \sum_{m=0}^{511} x[nN-m]h_k[m] \quad k = 0, 1, \dots, N-1.$$

Substituting for $h_k[m]$ and $N = 32$, we get

$$s_k[n] = \sum_{m=0}^{511} x[n32-m]2h[m] \cos[\omega_k(m-16)] \quad k = 0, 1, \dots, N-1.$$

- (c) By the method of part (a) it follows that

$$\begin{aligned} G_k(e^{j\omega}) &= Ne^{-j\psi_k} H(e^{j(\omega-\omega_k)}) + Ne^{j\psi_k} H(e^{j(\omega+\omega_k)}) \\ &= Ne^{j\phi_k} H(e^{j(\omega-\omega_k)}) + Ne^{-j\phi_k} H(e^{j(\omega+\omega_k)}). \end{aligned}$$

- (d) The upsampled (quantized) channel signals are

$$v_k[n] = \sum_{r=-\infty}^{\infty} \hat{s}_k[r] \delta[n - rN]$$

Therefore, the outputs of the synthesis filters are

$$\hat{u}_k[n] = \sum_{r=-\infty}^{\infty} \hat{s}_k[r] g_k[n - rN] \quad k = 0, 1, \dots, N-1,$$

where $g_k[n] = 2Nh[n] \cos[\omega_k(n + N/2)]$. Therefore, the total summed output of the synthesis filter bank is (with $N = 32$)

$$\hat{x}[n] = \sum_{k=0}^{31} \left(\sum_{r=-\infty}^{\infty} \hat{s}_k[r] (64h[n - r32]) \cos[\omega_k(n - r32 + 16)] \right).$$

- (e) Assuming that we wish to compute the ($N = 32$) samples of the r_0^{th} block of output samples, we would set $r = r_0$ and limit n to the range $r_032 \leq n \leq (r_0 + 1)32$. Equivalently, we can set $n = r_032 + m$, where $0 \leq m \leq 31$. Therefore, the equation for computing the r_0^{th} block of 32 analysis filter bank output samples is

$$\begin{aligned} \hat{x}[r_032 + m] &= \sum_{k=0}^{31} \left(\sum_{r=-\infty}^{\infty} \hat{s}_k[r] (64h[r_032 + m - r32]) \cos[\omega_k(r_032 + m - r32 + 16)] \right) \\ &= \sum_{k=0}^{31} \left(\sum_{r=-\infty}^{\infty} \hat{s}_k[r] (64h[(r_0 - r)32 + m]) \cos[\omega_k((r_0 - r)32 + m + 16)] \right). \end{aligned}$$

Now set $r' = r_0 - r$ and we get for $m = 0, 1, \dots, 31$,

$$\hat{x}[r_032 + m] = \sum_{k=0}^{31} \left(\sum_{r'=0}^{15} \hat{s}_k[r_0 - r'] (64h[r'32 + m]) \cos[\omega_k(r'32 + m + 16)] \right).$$

The upper limit of 15 on the sum on r' results from the fact that the length of the impulse response is 16×32 ; i.e., the impulse response spans 16 blocks of 32 samples.

- (f) This simply requires that the impulse response be loaded into MATLAB and examined. the first and last ten samples and the five samples around the central sample $h[256]$ should be found to have the values below.

$$\begin{aligned} h[0] &= 0.0000000000 = h[512] \\ h[1] &= -0.2385 \times 10^{-6} = h[511] \\ h[2] &= -0.2385 \times 10^{-6} = h[510] \\ h[3] &= -0.2385 \times 10^{-6} = h[509] \\ h[4] &= -0.2385 \times 10^{-6} = h[508] \\ h[5] &= -0.2385 \times 10^{-6} = h[507] \\ h[6] &= -0.2385 \times 10^{-6} = h[506] \\ h[7] &= -0.4770 \times 10^{-6} = h[505] \\ h[8] &= -0.4770 \times 10^{-6} = h[504] \\ h[9] &= -0.4770 \times 10^{-6} = h[503] \\ &\vdots && \vdots \\ h[254] &= 0.0178470610 = h[258] \\ h[255] &= 0.0178794860 = h[257] \\ h[256] &= 0.0178904535 = h[256] \end{aligned}$$

That is $h[512 - n] = h[n]$ for $0 \leq n \leq 512$. This implies that the frequency response of the prototype lowpass filter will have the form

$$H(e^{j\omega}) = A(e^{j\omega})e^{-j\omega 256}$$

where $A(e^{j\omega})$ is purely real, and the factor $e^{-j\omega 256}$ corresponds to a time delay of 256 samples.

- (g) A program for making the requested plot is the following:

```
% Create plot of MPEG window, impulse response and lowpass frequency
% response.
load mpegwindow
h=h/2;
subplot(311)
han1=plot(0:512,h1);
set(han1,'linewidth',2)
title(' (a) MPEG Analysis Filter Bank Window')
ylabel('\itc\rm[\itn\rm]')
xlabel('sample time \itn')
hold on
han1=plot([256,256],[-0.04,.04],'r:',0,0,'or',512,0,'or');
set(han1,'markersize',4)
plot([0,512],[0,0],'k')
axis([0,512,-.04,.04])
subplot(312)
han2=plot(0:512,h);
set(han2,'linewidth',2)
hold on
han1=plot([256,256],[-0.04,.04],'r:',0,0,'or',512,0,'or');
set(han1,'markersize',4)
plot([0,512],[0,0],'k')
axis([0,512,-.01,.02])
title(' (b) Impulse Response of Prototype Lowpass Filter')
ylabel('\ith\rm[\itn\rm]')
xlabel('sample time \itn')
omega=(0:8192)*pi/8192;
H=20*log10(abs(freqz(h,1,omega)));
Hp=H(129) % measure gain at omega=pi/64
Hs=max(H(257:8193)) % measure maximum gain in stopband
subplot(313)
han3=plot(omega/pi,H);grid
axis([0,1/8,-150,50])
set(han3,'linewidth',2)
title(' (c) Frequency Response of Prototype Lowpass Filter')
ylabel('20log_{10}|\itH\rm(\ite^{j\omega}\rm)|')
xlabel('Normalized frequency \omega/\pi')
%
```

- (h) Note that in the above program, ω takes values $(0:8192)*\pi/8192$. Thus, index 129 corresponds to $\omega = 128\pi/8192 = \pi/64$ and therefore, the statement $Hp=H(129)$ with no ; prints the gain in dB at $\omega = \pi/64$. Since index 257 corresponds to $\omega = 256\pi/8192 = \pi/32$, the statement $Hs=max(H(257:8193))$ in the above program prints the maximum gain in dB over the range $\pi/32 \leq \omega \leq \pi$. When the program runs, it prints

```
Hp =
-3.010020712550408
```

```
Hs =
-95.557324861119497
```

- (i) A program for making the requested plot is as follows:

```
% Plot first four frequency responses of MPEG analysis bank
load mpegwindow
h=h(:)/2;
n=(0:512);
num=4;
k=[0:num-1]';
hk=2*[ones(num,1)*h'].*cos(pi*[(2*k+1)]*[(n-16)]/64);
omega=(-128:1024)*pi/8192';
Hk=zeros(1025+128,num);
for k=1:num
    Hk(:,k)=freqz(hk(k,:),1,omega)';
end
han=plot(omega/pi,20*log10(abs(Hk(:,1))),'-',omega/pi,20*log10(abs(Hk(:,2))),'--',...
    omega/pi,20*log10(abs(Hk(:,3))),':',omega/pi,20*log10(abs(Hk(:,4))),'-.');
title('First Four Channels of MPEG Analysis Filter Bank')
xlabel('\omega /pi');ylabel('log magnitude in dB')
text(1/64,7,'k=0');text(3/64,7,'k=1');text(5/64,7,'k=2');text(7/64,7,'k=3')
axis([-1/64,1/8,-150,20])
set(han,'linewidth',2);
```

- (j) Here we derive the frequency-domain representation of the output of the synthesis filter bank assuming no quantization of the channel signals.

First the DTFTs of the analysis filter bank outputs are

$$\begin{aligned} U_k(e^{j\omega}) &= H_k(e^{j\omega})X(e^{j\omega}) \\ S_k(e^{j\omega}) &= \frac{1}{N} \sum_{r=0}^{N-1} U_k(e^{j(\omega-2\pi r)/N}) = \frac{1}{N} \sum_{r=0}^{N-1} H_k(e^{j(\omega-2\pi r)/N})X(e^{j(\omega-2\pi r)/N}) \end{aligned}$$

Now assume no quantization so that $\hat{S}_k(e^{j\omega}) = S_k(e^{j\omega})$. Then we have the following for the outputs of the individual synthesis channels:

$$\begin{aligned} V_k(e^{j\omega}) &= S_k(e^{j\omega N}) = \frac{1}{N} \sum_{r=0}^{N-1} H_k(e^{j(\omega-2\pi r/N)})X(e^{j(\omega-2\pi r/N)}) \\ \hat{U}_k(e^{j\omega}) &= G_k(e^{j\omega})V_k(e^{j\omega}) \\ &= \frac{1}{N} \sum_{r=0}^{N-1} G_k(e^{j\omega})H_k(e^{j(\omega-2\pi r/N)})X(e^{j(\omega-2\pi r/N)}) \end{aligned}$$

Now, the overall output is the sum of all the channel signals, so

$$\begin{aligned}
 \hat{X}(e^{j\omega}) &= \sum_{k=0}^{N-1} \hat{U}_k(e^{j\omega}) = \sum_{k=0}^{N-1} \left(\frac{1}{N} \sum_{r=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j(\omega-2\pi r/N)}) X(e^{j(\omega-2\pi r/N)}) \right) \\
 &= \sum_{r=0}^{N-1} \left(\frac{1}{N} \sum_{k=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j(\omega-2\pi r/N)}) \right) X(e^{j(\omega-2\pi r/N)}) \\
 &= \sum_{r=0}^{N-1} \tilde{H}_r(e^{j\omega}) X(e^{j(\omega-2\pi r/N)}) \\
 &= \tilde{H}_0(e^{j\omega}) X(e^{j\omega}) + \sum_{r=1}^{N-1} \tilde{H}_r(e^{j\omega}) X(e^{j(\omega-2\pi r/N)}),
 \end{aligned}$$

where

$$\tilde{H}_r(e^{j\omega}) = \frac{1}{N} \sum_{k=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j(\omega-2\pi r/N)}) \quad r = 0, 1, \dots, N-1.$$

(k) Perfect reconstruction with delay ($\hat{x}[n] = x[n - n_d]$) requires

$$\tilde{H}_0(e^{j\omega}) = \frac{1}{N} \sum_{k=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j\omega}) = e^{-j\omega n_d},$$

and also that the aliasing component gains be zero; i.e.,

$$\tilde{H}_r(e^{j\omega}) = \frac{1}{N} \sum_{k=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j(\omega-2\pi r/N)}) = 0 \quad r = 1, 2, \dots, N-1.$$

First consider the flat gain condition. Using the equations derived for $H_k(e^{j\omega})$ and $G_k(e^{j\omega})$ and recalling that $\phi_k = \omega_k N/2 = -\psi_k$, we can write $\tilde{H}_0(e^{j\omega})$ as

$$\begin{aligned}
 \tilde{H}_0(e^{j\omega}) &= \frac{1}{N} \sum_{k=0}^{N-1} \left[e^{-j\phi_k} H(e^{j(\omega-\omega_k)}) + e^{j\phi_k} H(e^{j(\omega+\omega_k)}) \right] \left[N e^{j\phi_k} H(e^{j(\omega-\omega_k)}) + N e^{-j\phi_k} H(e^{j(\omega+\omega_k)}) \right] \\
 &= \sum_{k=0}^{N-1} \left[H^2(e^{j(\omega-\omega_k)}) + H^2(e^{j(\omega+\omega_k)}) + 2 \cos(2\phi_k) H(e^{j(\omega-\omega_k)}) H(e^{j(\omega+\omega_k)}) \right]
 \end{aligned}$$

Now $\phi_k = \omega_k N/2 = \frac{2\pi}{4N}(2k+1)N/2 = \frac{\pi}{4}(2k+1)$ so $2\phi_k = \frac{\pi}{2}(2k+1)$. Therefore, since $2\phi_k$ is an odd multiple of $\pi/2$ for all k , it follows that $\cos(2\phi_k) = 0$ for all k and therefore,

$$\tilde{H}_0(e^{j\omega}) = \sum_{k=0}^{N-1} \left[H^2(e^{j(\omega-\omega_k)}) + H^2(e^{j(\omega+\omega_k)}) \right]$$

Now consider the nature of the prototype lowpass filter. Its impulse response has the property that $h[L-1-n] = h[n]$ for $n = 0, 1, \dots, L-1$ where $L = 513 = 2pN + 1$, with $p = 8$ and $N = 32$. As shown in part (f), the lowpass prototype filter is a generalized linear phase system with time delay $(L-1)/2 = pN = 256$ samples. This means that the frequency response of the prototype lowpass filter can be expressed as

$$H(e^{j\omega}) = A(e^{j\omega}) e^{-j\omega pN},$$

where $A(e^{j\omega})$ is a real periodic function of ω . If we substitute this form into the above equation for $H_0(e^{j\omega})$, we obtain

$$\begin{aligned}\tilde{H}_0(e^{j\omega}) &= \sum_{k=0}^{N-1} \left[A^2(e^{j(\omega-\omega_k)}) e^{-j(\omega-\omega_k)2pN} + A^2(e^{j(\omega+\omega_k)}) e^{-j(\omega+\omega_k)2pN} \right] \\ &= \left[A^2(e^{j(\omega-\omega_k)}) e^{j\omega_k 2pN} + A^2(e^{j(\omega+\omega_k)}) e^{-j\omega_k 2pN} \right] e^{-j\omega 2pN}\end{aligned}$$

Now consider the quantity $\omega_k 2pN = \frac{2\pi}{4N}(2k+1)2pN = \pi(2k+1)p$. Since the quantity $(2k+1)$ is always odd, it follows that

$$e^{\pm j\omega_k 2pN} = \begin{cases} +1 & \text{if } p \text{ is even} \\ -1 & \text{if } p \text{ is odd,} \end{cases}$$

and therefore, since $p = 8$, it follows that

$$\tilde{H}_0(e^{j\omega}) = \sum_{k=0}^{N-1} \left[A^2(e^{j(\omega-\omega_k)}) + A^2(e^{j(\omega+\omega_k)}) \right] e^{-j\omega 2pN} = \tilde{A}_0(e^{j\omega}) e^{-j\omega 2pN}.$$

That is, the overall composite system function from the input $X(e^{j\omega})$ to the output $\hat{X}(e^{j\omega})$ has the form of a generalized linear phase system with gain

$$\tilde{A}_0(e^{j\omega}) = \sum_{k=0}^{N-1} \left[A^2(e^{j(\omega-\omega_k)}) + A^2(e^{j(\omega+\omega_k)}) \right]$$

and time delay $n_d = 2pN = 512$ samples. The prototype lowpass must be chosen so that $\tilde{A}_0(e^{j\omega}) \approx 1$ for all ω . We will see in part (l) that the filter in the MPEG standard satisfies this condition very well.

Now we must consider the alias terms. The equation for the gain of the r^{th} alias component is

$$\begin{aligned}\tilde{H}_r(e^{j\omega}) &= \sum_{k=0}^{N-1} G_k(e^{j\omega}) H_k(e^{j(\omega-2\pi r/N)}) \quad r = 1, 2, \dots, N-1 \\ &= \sum_{k=0}^{N-1} \left[H(e^{j(\omega-\omega_k)}) H(e^{j(\omega-\omega_k-2\pi r/N)}) + H(e^{j(\omega+\omega_k)}) H(e^{j(\omega+\omega_k-2\pi r/N)}) \right. \\ &\quad \left. + e^{-j2\phi_k} H(e^{j(\omega+\omega_k)}) H(e^{j(\omega-\omega_k-2\pi r/N)}) + e^{j2\phi_k} H(e^{j(\omega-\omega_k)}) H(e^{j(\omega+\omega_k-2\pi r/N)}) \right],\end{aligned}$$

where $r = 1, 2, \dots, N-1$. Since we assume that $H(e^{j\omega}) = 0$ outside the interval $|\omega| \leq \pi/N$ as shown in Figure 12.32(b) and Figure P12.2.1 part (a), the region of support of all the functions is of length $2\pi/N$, and therefore $H(e^{j(\omega-\omega_k)})$ and $H(e^{j(\omega+\omega_k)})$ do not overlap with their corresponding copies shifted by $2\pi r/N$. Therefore, the first line of the above equation disappears leaving only

$$\begin{aligned}\tilde{H}_r(e^{j\omega}) &= \sum_{k=0}^{N-1} \left[e^{-j2\phi_k} H(e^{j(\omega+\omega_k)}) H(e^{j(\omega-\omega_k-2\pi r/N)}) + e^{j2\phi_k} H(e^{j(\omega-\omega_k)}) H(e^{j(\omega+\omega_k-2\pi r/N)}) \right] \\ &r = 1, 2, \dots, N-1.\end{aligned}$$

Now, not all $N-1$ terms are non-zero in this equation. Only when the product terms overlap with each other is there any contribution, and these contributions will be in the overlap regions between a filter band and adjacent bands on either side. In fact, for a given

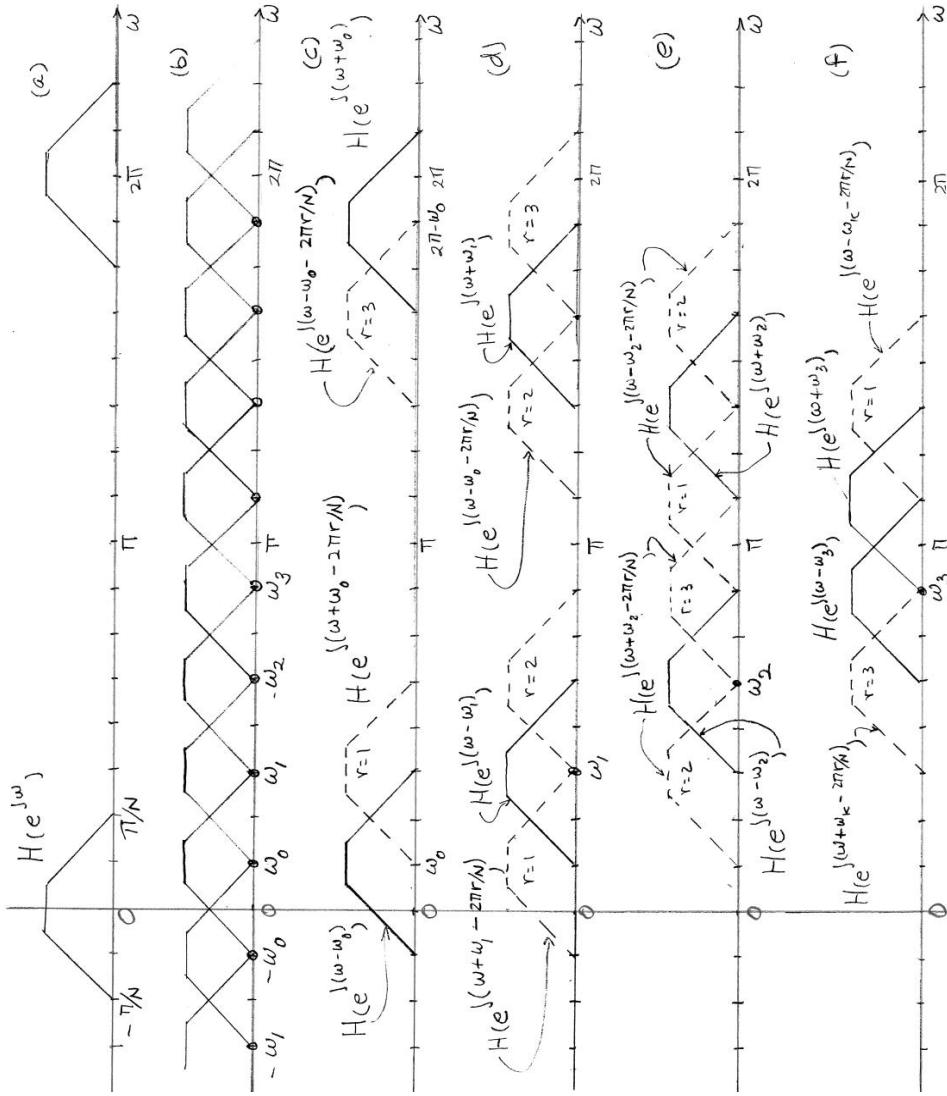


Figure P12.2.1: Stylized illustration of DTFTs in alias cancellation: (a) Lowpass prototype showing total width of $2\pi/N$; (b) Illustration of center frequencies for an $N = 4$ band system; (c) Illustration of the terms $H(e^{j(\omega-\omega_0)})H(e^{j(\omega+\omega_k-2\pi r/N)})$ for $r = 1$ and $H(e^{j(\omega+\omega_0)})H(e^{j(\omega-\omega_k-2\pi r/N)})$ for $r = 3$; (d), (e), and (f) Illustrations as in part (c) for terms corresponding to ω_1 , ω_2 and ω_3 .

value of r , only four terms will be non-zero. This is evident in the simple example for $N = 4$ shown in Figure P12.2.1. Specifically, generalizing from this example, it follows that for a given r , the sum over all channels reduces to the four terms

$$\begin{aligned}\tilde{H}_r(e^{j\omega}) &= e^{j2\phi_{r-1}} H(e^{j(\omega-\omega_{r-1})})H(e^{j(\omega+\omega_{r-1}-2\pi r/N)}) \\ &\quad + e^{j2\phi_r} H(e^{j(\omega-\omega_r)})H(e^{j(\omega+\omega_r-2\pi r/N)}) \\ &\quad + e^{-j2\phi_{N-1-r}} H(e^{j(\omega+\omega_{N-1-r})})H(e^{j(\omega-\omega_{N-1-r}-2\pi r/N)}) \\ &\quad + e^{-j2\phi_{N-r}} H(e^{j(\omega+\omega_{N-r})})H(e^{j(\omega-\omega_{N-r}-2\pi r/N)}).\end{aligned}$$

As illustrated in Figure P12.2.1 these terms come in pairs. For example,

$$\omega_{r-1} - 2\pi r/N = \frac{2\pi}{4N}(2(r-1)+1) - \frac{2\pi}{N} = \frac{2\pi}{4N}(2r-2+1-4r) = \frac{2\pi}{4N}(-2r-1) = -\omega_r$$

so that

$$H(e^{j(\omega+\omega_{r-1}-2\pi r/N)}) = H(e^{j(\omega-\omega_r)}).$$

Similarly, it can be shown that

$$\omega_r - 2\pi r/N = -\omega_{r-1}$$

so that

$$H(e^{j(\omega+\omega_r-2\pi r/N)}) = H(e^{j(\omega-\omega_{r-1})}).$$

This means that

$$\begin{aligned} H(e^{j(\omega-\omega_{r-1})})H(e^{j(\omega+\omega_{r-1}-2\pi r/N)}) &= H(e^{j(\omega-\omega_r)})H(e^{j(\omega+\omega_r-2\pi r/N)}) \\ &= H(e^{j(\omega-\omega_{r-1})})H(e^{j(\omega-\omega_r)}). \end{aligned}$$

Similarly, it can be shown that

$$-\omega_{N-r} - 2\pi r/N = \omega_{N-1-r} \quad \text{and} \quad -\omega_{N-1-r} - 2\pi r/N = \omega_{N-r},$$

so that

$$\begin{aligned} H(e^{j(\omega+\omega_{N-1-r})})H(e^{j(\omega-\omega_{N-1-r}-2\pi r/N)}) &= H(e^{j(\omega+\omega_{N-r})})H(e^{j(\omega-\omega_{N-r}-2\pi r/N)}) \\ &= H(e^{j(\omega-\omega_{N-1-r})})H(e^{j(\omega-\omega_{N-r})}). \end{aligned}$$

Therefore, we have at last,

$$\begin{aligned} \tilde{H}_r(e^{j\omega}) &= (e^{j2\phi_{r-1}} + e^{j2\phi_r})H(e^{j(\omega-\omega_{r-1})})H(e^{j(\omega-\omega_r)}) \\ &\quad + (e^{j2\phi_{N-1-r}} + e^{j2\phi_{N-r}})H(e^{j(\omega-\omega_{N-1-r})})H(e^{j(\omega-\omega_{N-r})}). \end{aligned}$$

Now we are finally in a position to show that $\tilde{H}_r(e^{j\omega}) = 0$ for $r = 1, 2, \dots, N-1$ under the assumption that $|H(e^{j\omega})| = 0$ for $\pi/N \leq |\omega| \leq \pi$.

To do this recall that $2\phi_k = \omega_k N/2 = 2\frac{2\pi}{4N}(2k+1)N/2 = \frac{\pi}{2}(2k+1)$. This means that

$$(e^{\pm j2\phi_{k-1}} + e^{\pm j2\phi_k}) = [(\pm j)^{2k-1} + (\pm j)^{2k+1}] = (\pm j)^{2k}[(\pm j)^{-1} + (\pm j)] = 0.$$

Thus, because the phase shifts between successive channels differ by $\pi/2$ radians, the factors in parentheses in the final equation for $\tilde{H}(e^{j\omega})$ are both zero. That is, the gains on the aliasing terms $X(e^{j(\omega-2\pi r/N)})$ are all zero, so if the lowpass filter is designed so that $\tilde{A}_0(e^{j\omega}) = 1$, the MPEG analysis synthesis filter banks yield nearly perfect reconstruction. We will verify this in the next part of this problem.

- (1) The program for making the requested plots is as follows:

```
load mpeg_window
%   make Hk filters
figure(1)
n=(0:512)';
k=[0:31]';
hk=2*[ones(32,1)*h'].*cos(pi*[(2*k+1)*[(n-16)]/64];
omega=(0:8000)*pi/8000';
Hk=zeros(8001,32);
for k=1:32
    Hk(:,k)=freqz(hk(k,:),1,omega)';
end
%   make Gk filters
n=(0:512);
```

```

k=[0:31]';
gk=64*[ones(32,1)*h'].*cos(pi*[(2*k+1)]*[(n+16)]/64);
omega=(0:8000)*pi/8000';
Gk=zeros(8001,32);
for k=1:32
    Gk(:,k)=freqz(gk(k,:),1,omega)';
end
% make overall frequency response
H0=zeros(8001,1);
for m=1:32
    H0=H0+Gk(:,m).*Hk(:,m);
end
plot(omega/pi,20*log10(abs(H0/32)),'k')
title('Composite Response of MPEG Analysis/Synthesis Filter Bank')
xlabel('normalized frequency \omega/pi')
ylabel('log magnitude in dB')
hold on
h=plot([0,1],[0,0],'k--');
set(h,'linewidth',2)
A=axis;
axis([0,1,A(3),A(4)])
print -deps mpeg_composite_response.eps
%
% now plot alias term for r=4
%
figure(2)
Hk=zeros(8001,32);
r=4;
for k=1:32
    Hk(:,k)=freqz(hk(k,:),1,omega-2*pi*r/32)';
end
H4=zeros(8001,1);
for m=1:32
    H4=H4+Gk(:,m).*Hk(:,m);
end
maximum_aliasing_attenuation=-20*log10(max(abs(H4/32)))
plot(omega/pi,20*log10(abs(H4/32)),'k')
title(['Alias Term (r=',num2str(r),') for MPEG Analysis/Synthesis Filter Bank'])
xlabel('normalized frequency \omega/pi')
ylabel('log magnitude in dB')
print -deps mpeg_alias_term4.eps

```

The plot of the overall composite frequency response $20 \log_{10} |\tilde{H}_0(e^{j\omega})|$ is shown in Figure P12.2.2. The maximum error with respect to flat gain of 0 dB is 0.00078730 dB.

The plot of the aliasing component gain $20 \log_{10} |\tilde{H}_r(e^{j\omega})|$ for $r = 4$ is shown in Figure P12.2.3. The maximum aliasing component gain for $r = 4$ is -111.1771 dB.

- (m) The basic equation for the analysis filter bank outputs is

$$s_k[n] = \sum_{m=0}^{511} x[n32 - m] 2h[m] \cos[\omega_k(m - 16)], \quad k = 0, 1, \dots, N - 1.$$

We define $w_m[n]$ as a sequence indexed by m , which is associated with analysis frame n ;

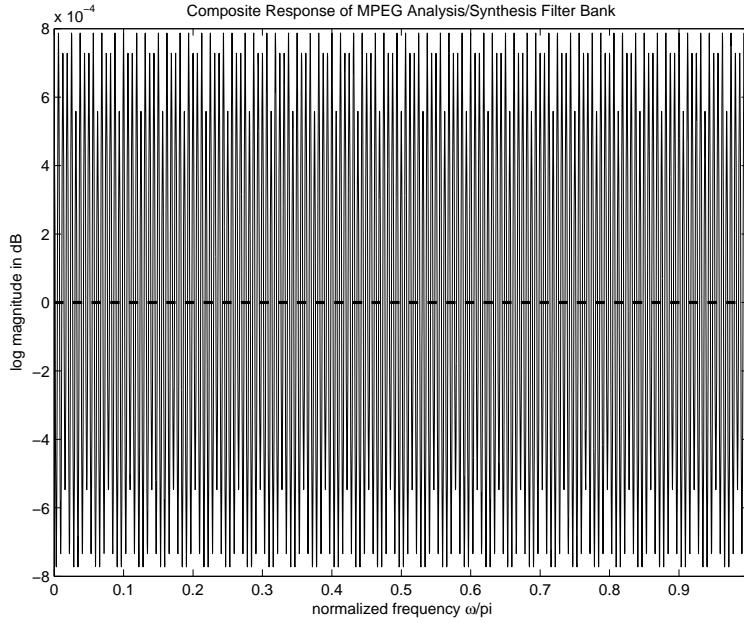


Figure P12.2.2: Overall frequency response $20 \log_{10} |\tilde{H}_0(e^{j\omega})|$ of MPEG filter bank.

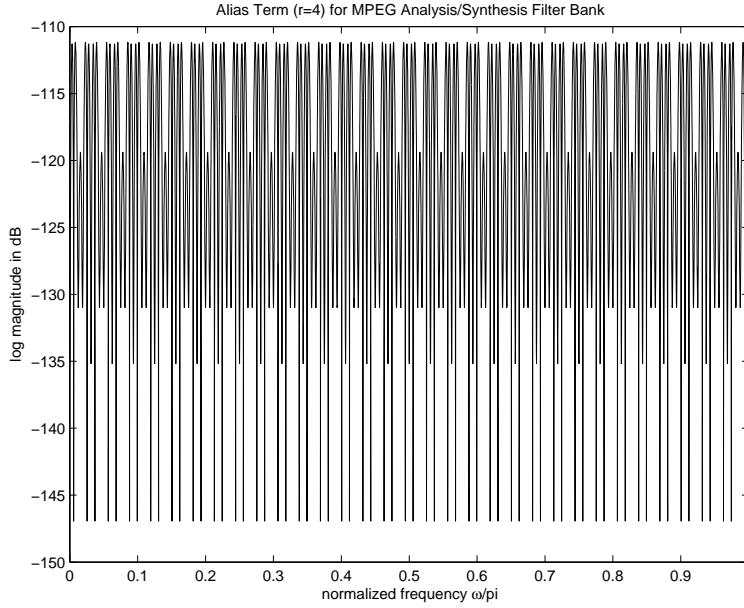


Figure P12.2.3: Aliasing component gain $20 \log_{10} |\tilde{H}_r(e^{j\omega})|$ for $r = 4$.

i.e.,

$$w_m[n] = x[n32 - m]2h[m], \quad m = 0, 1, \dots, 511.$$

Then we only need to compute $w_m[n]$ once and use it N times in

$$s_k[n] = \sum_{m=0}^{511} w_m[n] \cos[\omega_k(m - 16)], \quad k = 0, 1, \dots, N - 1$$

to compute all the channel signals at analysis time n .

- (n) Define $m = q + r64$ for $q = 0, 1, \dots, 63$ and $r = 0, 1, \dots, 7$. This is just counting modulo 64.

Clearly, $r = 0, q = 0, 1, \dots, 63$ indexes the first 64 samples, etc. If we make this substitution in the expression for $s_k[n]$ and recall that $\omega_k = \frac{2\pi}{4N}(2k+1) = \frac{2\pi}{128}(2k+1)$, we get

$$s_k[n] = \sum_{q=0}^{63} \sum_{r=0}^7 w_{q+r64}[n] \cos \left[\frac{2\pi}{128}(2k+1)(q+r64-16) \right], \quad k = 0, 1, \dots, N-1.$$

Now, note that

$$\begin{aligned} \cos \left[\frac{2\pi}{128}(2k+1)(q+r64-16) \right] &= \cos \left[\frac{2\pi}{128}(2k+1)r64 \right] \cos \left[\frac{2\pi}{128}(2k+1)(q-16) \right] \\ &\quad - \sin \left[\frac{2\pi}{128}(2k+1)r64 \right] \sin \left[\frac{2\pi}{128}(2k+1)(q-16) \right] \\ &= \cos [\pi(2k+1)r] \cos \left[\frac{2\pi}{128}(2k+1)(q-16) \right] \\ &\quad - \sin [\pi(2k+1)r] \sin \left[\frac{2\pi}{128}(2k+1)(q-16) \right] \\ &= (-1)^r \cos \left[\frac{2\pi}{128}(2k+1)(q-16) \right]. \end{aligned}$$

Therefore, grouping the factors dependent on r together, the outputs of the analysis filter bank are

$$s_k[n] = \sum_{q=0}^{63} \left(\sum_{r=0}^7 w_{q+r64}[n] (-1)^r \right) \cos \left[\frac{2\pi}{128}(2k+1)(q-16) \right], \quad k = 0, 1, \dots, N-1.$$

- (o) Define

$$\begin{aligned} z_{q+r64}[n] &= w_{q+r64}[n] (-1)^r \\ &= x[n32 - q - r64] 2h[q+r64] (-1)^r \\ &= x[n32 - q - r64] c[q+r64] \end{aligned}$$

where

$$c_{q+r64}[n] = 2h[q+r64] (-1)^r \quad q = 0, 1, \dots, 63 \quad \text{and} \quad r = 0, 1, \dots, 7.$$

That is, c is the lowpass impulse response with every other group of 64 samples multiplied by (-1) . This was shown in Figure P12.2b part (a).

- (p) Now, first we multiply the 512 samples preceding and including the sample at $n32$ by the analysis window, and then time-alias with period 64 to obtain

$$y_q[pn] = \sum_{r=0}^7 z_{q+r64}[n], \quad q = 0, 1, \dots, 63.$$

(This requires that the input be buffered in a 512 sample buffer.) Therefore, the channel outputs at the n^{th} block of the input is

$$s_k[n] = \sum_{q=0}^{63} y_q[n] \cos \left[\frac{2\pi}{128}(2k+1)(q-16) \right], \quad k = 0, 1, \dots, N-1.$$

If we move along the input time scale in jumps of $N = 32$ samples, then the above equation computes the decimated-by- N output of the analysis filter bank.

- (q) We refer to the numbers on the right in Figure P12.2c.
- (1) This represents moving the window in jumps of $N = 32$ samples so that that decimated output is obtained.
 - (2) This represents inserting a new block of $N = 32$ samples into the buffer.
 - (3) This corresponds to Eq. (P12.20).
 - (4) This corresponds to Eq. (P12.19).
 - (5) \mathbf{Y} is a column vector of length 64. \mathbf{M} is a matrix with 32 rows and 64 columns

$$\mathbf{M}[i, k] = \cos \left[\frac{2\pi}{128} (2k+1)(q-16) \right] \quad \begin{cases} i = 0, 1, \dots, 31 \\ k = 0, 1, \dots, 63 \end{cases}$$

The matrix product $\mathbf{S} = \mathbf{M}\mathbf{Y}$ gives a column vector \mathbf{S} with the 32 channel signals at time n .

- 12.3** This problem shows how the FFT can be used to implement the “matrixing” operation of block (5) in Figure P12.2c.

- (a) We want to implement

$$s_k[n] = \sum_{q=0}^{63} y_q[n] \cos \left[\frac{2\pi}{128} (2k+1)(q-16) \right], \quad k = 0, 1, \dots, N-1,$$

which, since $y_q[n]$ is real, can also be represented as

$$s_k[n] = \mathcal{R}e \left\{ \sum_{q=0}^{63} y_q[n] e^{-j2\pi(2k+1)(q-16)/128} \right\}, \quad k = 0, 1, \dots, N-1.$$

- (b) Breaking up the complex exponential, we obtain

$$\begin{aligned} s_k[n] &= \mathcal{R}e \left\{ \sum_{q=0}^{63} y_q[n] e^{-j2\pi(2k)q/128} e^{-j2\pi q/128} e^{j2\pi(2k+1)16/128} \right\} \\ &= \mathcal{R}e \left\{ e^{j\pi(2k+1)/4} \sum_{q=0}^{63} \left(y_q[n] e^{-j\pi q/64} \right) e^{-j2\pi kq/64} \right\}, \quad k = 0, 1, \dots, N-1. \end{aligned}$$

- (c) Now define

$$Y_k[n] = \sum_{q=0}^{63} \left(y_q[n] e^{-j\pi q/64} \right) e^{-j2\pi kq/64} \quad k = 0, 1, \dots, N-1.$$

This is the 64-point DFT of the sequence

$$y_q[n] e^{-j\pi q/64} \quad q = 0, 1, \dots, 31$$

We can compute $Y_k[n]$ as the 64-point FFT of the 64-point sequence $y_q[n] e^{-j\pi q/64}$. Then multiply $Y_k[n]$ by $e^{j\pi(2k+1)/4}$ for $k = 0, 1, \dots, 31$. Finally, we take the real part to get $s_k[n]$ for $k = 0, 1, \dots, 31$. This implements the matrix multiplication as defined in the MPEG standard.

12.4 Here we plot all 32 impulse responses of the analysis filter bank. A MATLAB program to make this plot is as follows:

```

subplot('position',[.1,.1,.6,.8])
H=zeros(513,32);
omegak=pi*(2*(0:31)+1)/64 ;
for m=1:32;
    H(:,m)=2*h.*cos(omegak(m).*n' - omegak(m)*16);
end
strips(H(:,),513)
set(gca,'yticklabel',...
{'k=31','k=30','k=29','k=28','k=27','k=26','k=25',...
'k=24','k=23','k=22','k=21','k=20','k=19','k=18',...
'k=17','k=16','k=15','k=14','k=13','k=12','k=11',...
'k=10','k=9','k=8','k=7','k=6','k=5','k=4','k=3',...
'k=2','k=1','k=0'})
xlabel('time in samples')
title('Impulse Responses of MPEG Bandpass Filters')
print -deps impulse_responses_0-31.eps

```

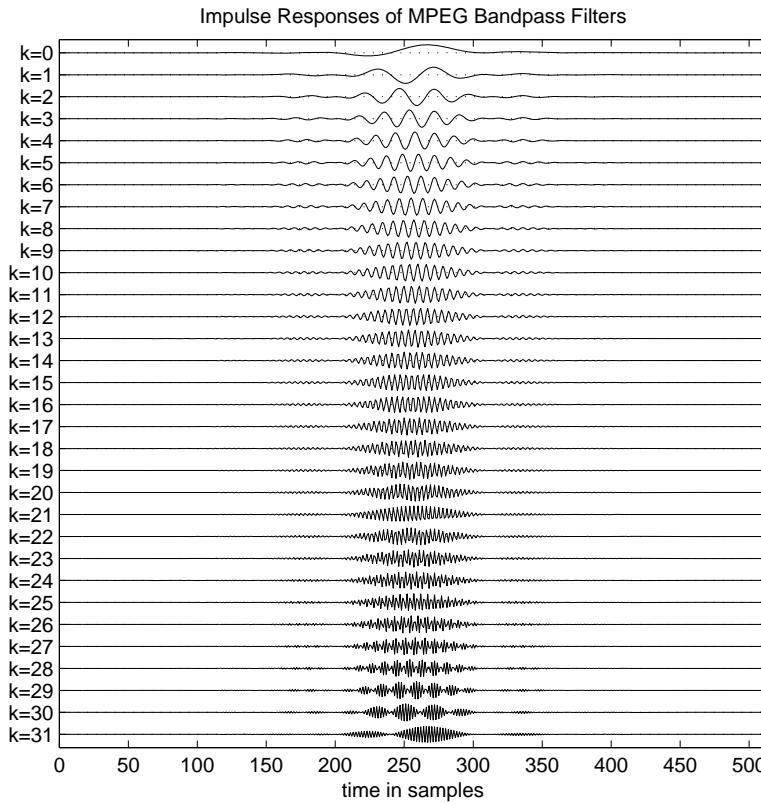


Figure P12.4.1: Impulse responses of MPEG analysis filter bank.

This figure illustrates that, in contrast to a wavelet transform, the time resolution is essentially the same for all channels in the analysis filter bank.

Chapter 13

Text-to-Speech Synthesis Methods

All the problems in this chapter are MATLAB Exercises, so there are no solutions.

Chapter 14

Automatic Speech Recognition and Natural Language Understanding

14.1 The correct matching strings for the eleven digits (/zero/-/nine/ plus /ph/) are shown in Figure P14.1.1.

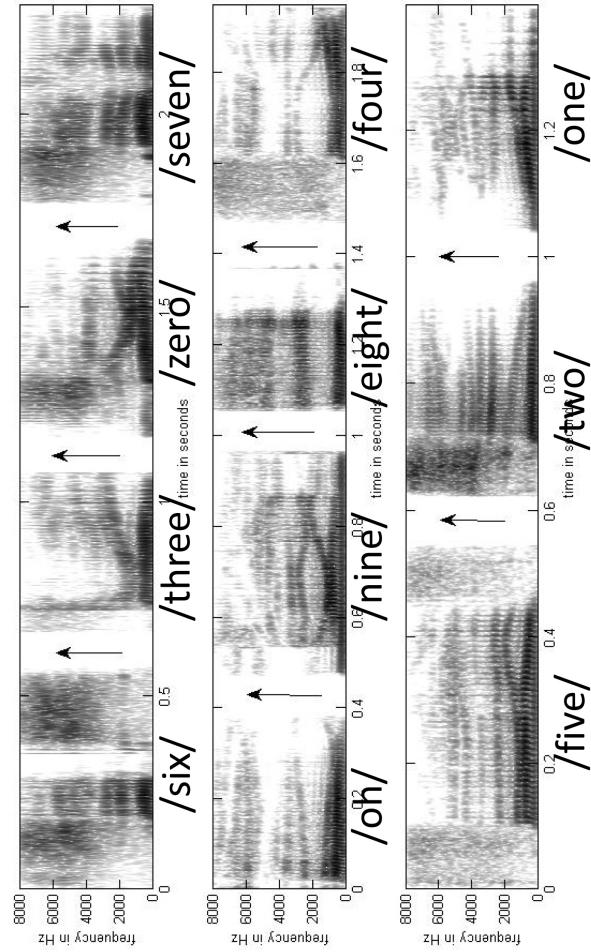


Figure P14.1.1: Labeled spectrograms for the eleven digits.

We can readily determine the best matches as follows:

- the first digit in the first row begins with a strong fricative, followed by a short internal vowel followed by a stop gap and ending with a strong fricative, thereby strongly indicating the digit /six/.
- the second digit in the first row begins with a weak fricative, followed by an /R/ sound (lowered second and third formants) followed by a clear transition into an /IY/ vowel, thereby indicating the digit /three/.
- the third digit in the first row begins with a voiced fricative (note the periodicity striations and the presence of formants during the fricative region), followed by an /IY/ vowel (high second and third formants), followed by an /R/ (lowered second and third formants) and ending with an /OW/ sound, thereby indicating the digit /zero/.
- the fourth digit in the first row begins with a strong fricative, followed by an /EH/, then a weak voiced fricative, a short /EH/ vowel, and ending with a clear nasal sound, thereby indicating the digit /seven/.
- the first digit in the second row is pretty much a steady vowel corresponding clearly to the digit /oh/.
- the second digit in the second row begins with a clear nasal sound, followed by the diphthong /AY/ and ending in another clear nasal, thereby indicating the digit /nine/.
- the third digit in the second row begins with the diphthong /EY/ followed by a clear stop gap followed by the release from a stop consonant, thereby indicating the digit /eight/.
- the fourth digit in the second row begins with a weak fricative, followed by the vowel /AO/ and ending with an /R/ (lowered third formant), thereby indicating the digit /four/.
- the first digit in the third row begins with a weak fricative, followed by the diphthong /AY/ and ending with a weak voiced fricative, thereby indicating the digit /five/.
- the second digit in the third row begins with a burst release from a stop consonant, followed by an /UW/ vowel (notice the very low second formant), thereby indicating the digit /two/.
- the third digit in the third row begins with a /W/ sound (low second formant) followed by the vowel /UH/ and ending with a clear nasal, thereby indicating the digit /one/.

14.2 (a) From the Bayesian analysis we have the following:

$$\begin{aligned} P_e &= P(\omega_1)P(x \in R_2|\omega_1) + P(\omega_2)P(x \in R_1|\omega_2) \\ &= P(\omega_1)\int_{R_2} p(x|\omega_1)dx + P(\omega_2)\int_{R_1} p(x|\omega_2)dx \end{aligned}$$

For minimum classification error we require that:

$$P(\omega_1|x) = P(\omega_1) \frac{p(x|\omega_1)}{p(x)} > P(\omega_2|x) = P(\omega_2) \frac{p(x|\omega_2)}{p(x)}$$

The threshold value, x_0 , occurs when:

$$P(\omega_1|x) = P(\omega_2|x)$$

Plugging in the values given in the problem statement ($\alpha = 0.5$, $m_1 = 3$, $m_2 = -3$, $\sigma_1 = \sigma_2 = 1$) we get:

$$\frac{\alpha}{\sigma_1} \exp\left(-\frac{(x_0 - m_1)^2}{2\sigma_1^2}\right) = \frac{(1 - \alpha)}{\sigma_2} \exp\left(-\frac{(x_0 - m_2)^2}{2\sigma_2^2}\right)$$

Taking natural logs we get:

$$\ln\left(\frac{\alpha}{\sigma_1}\right) - \frac{(x_0 - m_1)^2}{2\sigma_1^2} = \ln\left(\frac{1-\alpha}{\sigma_2}\right) - \frac{(x_0 - m_2)^2}{2\sigma_2^2}$$

Solving the quadratic equation for x_0 gives:

$$x_0 = \frac{\left(\frac{m_1}{\sigma_1^2} - \frac{m_2}{\sigma_2^2}\right) \pm \left[\left(\frac{m_1}{\sigma_1^2} - \frac{m_2}{\sigma_2^2}\right)^2 - 2\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) + \left[\ln\left(\frac{1-\alpha}{\sigma_2}\right) - \ln\left(\frac{\alpha}{\sigma_1}\right) + \frac{m_1^2}{2\sigma_1^2} - \frac{m_2^2}{2\sigma_2^2}\right]\right]^{1/2}}{\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)}$$

(b) Substituting the values for α , m_1 , m_2 , σ_1 and σ_2 we get:

$$x_0 = 0$$

$$P_2 = \text{erf}(3) = 0.0013$$

14.3 (a) The threshold value, \hat{x}_0 , is obtained when:

$$\lambda_{21}P(\omega_2)p(x|\omega_2) = \lambda_{12}P(\omega_1)p(x|\omega_1)$$

$$\lambda_{21}(1-\alpha)\exp[-(\hat{x}_0 - m_2)^2/2] = \lambda_{12}\exp[-(\hat{x}_0 - m_1)^2]$$

Taking natural logs we get:

$$\begin{aligned} \ln[\lambda_{21}(1-\alpha)] - (\hat{x}_0 - m_2)^2/2 &= \ln[\lambda_{12}\alpha] - (\hat{x}_0 - m_1)^2/2 \\ (\hat{x}_0 - m_1)^2 - (\hat{x}_0 - m_2)^2 &= 2(\ln[\lambda_{12}\alpha] - \ln[\lambda_{21}(1-\alpha)]) \\ \hat{x}_0(m_2 - m_1) &= \left(\ln[\lambda_{12}\alpha] - \ln[\lambda_{21}(1-\alpha)] + \frac{(m_2^2 - m_1^2)}{2}\right) \\ \hat{x}_0 &= \left(\ln[\lambda_{12}\alpha] - \ln[\lambda_{21}(1-\alpha)] + \frac{(m_2^2 - m_1^2)}{2}\right) / (m_2 - m_1) \\ \hat{x}_0 &= \frac{1}{(m_1 - m_2)} \ln\left[\frac{1-\alpha}{\alpha} \frac{\lambda_{21}}{\lambda_{12}}\right] + \frac{m_1 + m_2}{2} \end{aligned}$$

(b) The minimum risk error rate for $\alpha = 0.5$, $m_1 = 3$, $m_2 = -3$, $\lambda_{12} = 4$, and $\lambda_{21} = 1$ is:

$$\hat{x}_0 = (\ln(2) - \ln(0.5)) / (-6) = -0.23$$

The resulting minimum risk error rate is thus:

$$P_e = 0.5\text{erf}(2.77) + 0.5\text{erf}(3.23) = 0.5(0.0028) + 0.5(0.0006) = 0.0017$$

which is larger than the minimum error rate (the one that is independent of risk).

14.4 1. Since:

$$\begin{aligned} |c_1[n] - c_2[n]| &\geq 0, \quad \forall n, \text{ then} \\ d(c_1, c_2) &= \left[\sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])^2 \right]^{1/2} \geq 0 \\ d(c_1, c_2) &= 0, \text{ if and only if } c_1[n] = c_2[n] \quad \forall n \end{aligned}$$

2. Since:

$$(c_1[n] - c_2[n])^2 = (c_2[n] - c_1[n])^2, \quad d(c_1, c_2) \text{ is symmetric}$$

3. We define:

$$A = d(c_1, c_2) + d(c_2, c_3)$$

where

$$d(c_1, c_2) = \left[\sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])^2 \right]^{1/2}$$

Thus we have:

$$A^2 = d^2(c_1, c_2) + d^2(c_2, c_3) + 2d(c_1, c_2)d(c_2, c_3) \quad (14.1)$$

We define B as:

$$B = \left[\sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])^2 \right]^{1/2} \left[\sum_{n=-\infty}^{\infty} (c_2[n] - c_3[n])^2 \right]^{1/2}$$

The Cauchy-Schwartz inequality states that for $a_i, b_i \in \mathcal{R}$ (the set of real numbers), we have the inequality:

$$\begin{aligned} \left(\sum_i a_i^2 \right)^{1/2} \left(\sum_i b_i^2 \right)^{1/2} &\leq \sum_i a_i b_i \\ \therefore B &\geq \sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])(c_2[n] - c_3[n]) \end{aligned}$$

Therefore, according to Equation (14.1), we have:

$$\begin{aligned} A^2 &\geq \sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])^2 + \sum_{n=-\infty}^{\infty} (c_2[n] - c_3[n])^2 + 2 \sum_{n=-\infty}^{\infty} (c_1[n] - c_2[n])(c_2[n] - c_3[n]) \\ &= \sum_{n=-\infty}^{\infty} [(c_1[n] - c_2[n]) + (c_2[n] - c_3[n])]^2 \\ &= \sum_{n=-\infty}^{\infty} [(c_1[n] - c_3[n])^2] = d^2(c_1, c_3) \end{aligned}$$

Therefore we get:

$$A = d(c_1, c_2) + d(c_2, c_3) \geq \sqrt{d^2(c_1, c_3)} = d(c_1, c_3)$$

14.5 The first step is to label each of the nodes (from 1 to 14) as shown in Figure P14.5.1. We can now make a table of predecessor nodes and a table of successor nodes so that we can perform the dynamic path computation in the most efficient manner.

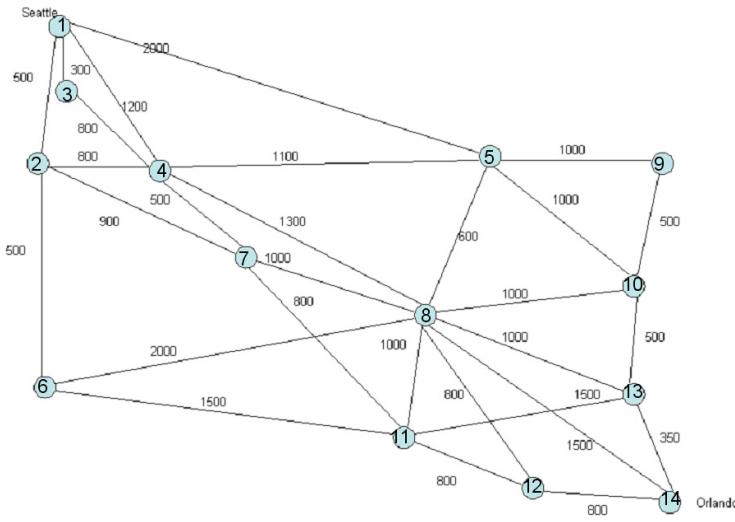


Figure P14.5.1: Node labeling of grid.

Node	Predecessor Node(s)	Successor Node(s)
1	Initial	2,3,4,5
2	1	4,6,7
3	1	4
4	1,2,3	5,7,8
5	1,4	8,9,10
6	2	8,11
7	2,4	8,11
8	4,5,7	11,12,13,14
9	5	10
10	5,8,9	13
11	6,7,8	12,13
12	8,11	14
13	8,10,11	14
14	8,12,13	Terminal

Next we determine the smallest distance path to each node by performing the computation of each node sequentially. The results set of node distances and predecessor nodes is as shown in Figure P14.5.2.

The minimum total distance and backpointer(s) to each node is shown in bold. Hence the minimum total distance to node 8 is 2400, and there is a tie between nodes 4 and 7 as the best predecessor node. The smallest total path distance is 3750 and the backpointer at node 14 points to node 13. By following the path backpointers, we get the trace (with two equal distance paths) shown in Figure P14.5.3.

The two equivalent best paths through the network are:

Path 1 : node 1 (distance 0) – node 2 (distance 500) – node 7 (distance 1400) – node 8 (distance 2400) – node 13 (distance 3400) – node 14 (distance 3750)

Path 2 : node 1 (distance 0) – node 3 (distance 300) – node 4 (distance 1100) – node 8 (distance 2400) – node 13 (distance 3400) – node 14 (distance 3750)

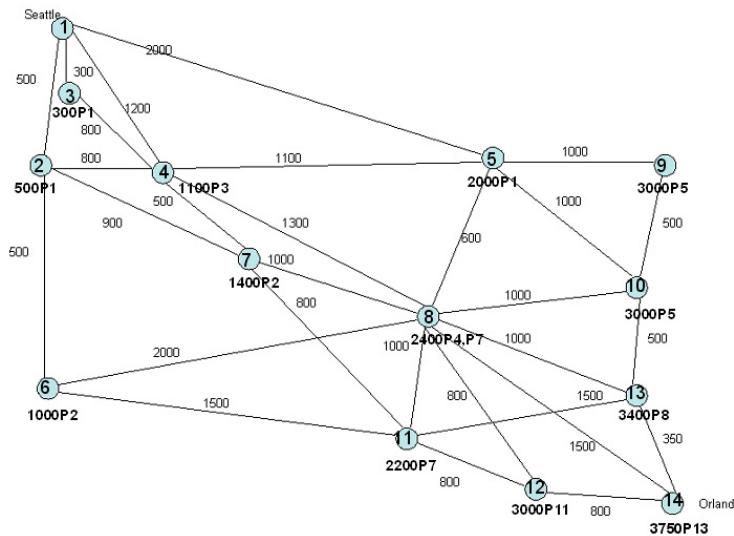


Figure P14.5.2: Node distance labeling of grid.

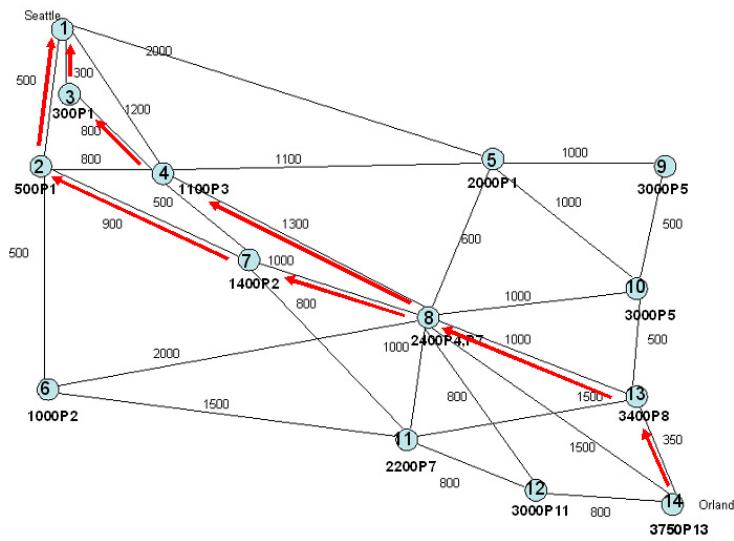


Figure P14.5.3: Total distance labeling of grid.

14.6 (a) The probability of the state sequence, $Q = [1 \ 2 \ 3 \ 4]$ is:

$$p(Q) = \pi_1 a_{12} a_{23} a_{34} = (0.25)(0.3)(0.1)(0.25) = 1.875 \times 10^{-3} = 0.001875$$

(b) The probability of the combined state and observation sequence, $p(Q, O)$ is:

$$p(Q, O) = \pi_1 b_1(5) a_{12} b_2(3) a_{23} b_3(2) a_{34} b_4(1)$$

$$p(Q, O) = (0.25)(0.1)(0.3)(0.1)(0.2)(0.25)(0.0) = 0$$

(c) Using the forward-backward recursion, we get:

$$p(O)_{fb} = 0.001933; \quad \log(p(O)_{fb}) = -6.249$$

and there is no optimum state sequence, while using the Viterbi recursion we get:

$$\log(p(O)_V = -7.3698, \quad p(O)_V = 0.000630$$

with the optimal state sequence being: $Q_{opt} = [4 \ 3 \ 2 \ 1]$

(d) The average duration of each state is:

$$\bar{d}_i = \frac{1}{1 - a_{ii}}$$

$$\bar{d}_1 = \frac{1}{0.9} = 1.1 = \bar{d}_2 = \bar{d}_4$$

$$\bar{d}_3 = \frac{1}{0.75} = 1.33$$
