12

Filters: Specification, Bode Plot, and Nyquist Plot

12.1 INTRODUCTION: FILTERS AS LINEAR TIME INVARIANT (LTI) SYSTEMS

In this chapter, we will continue to analyze filters while considering the RC filter presented in Chapters 10 and 11 as an LTI system (Chapter 8). To fully characterize an LTI system we can specify the following:

1. The system's reaction to a unit impulse: the ***impulse response***, or
2. The Laplace or z-transform of the impulse response (***transfer function***), or
3. The Fourier transform of the impulse response (***frequency response***)

The impulse response is useful because ***convolution*** of the impulse response with the input provides the output. The transfer function is practical because, just as in the frequency domain, the convolution may be performed as a multiplication in the s- or z-domain (Chapter 8). The ***frequency response*** is of practical interest for the same reason but also because it relates immediately and intuitively to the filter's function and its specification into pass band, transition band, and stop band.

The frequency response $H(j\omega)$ of a filter (LTI system) can be obtained analytically by using the Fourier transform of the impulse response or by deriving the solution in the frequency domain from knowledge of the system's components (Chapter 11, Section 11.2). In addition, one can determine the frequency response either from the transfer function or the z-transform of the impulse response. To convert from the Laplace transform $H(s)$ to frequency response $H(j\omega)$, one can often simply substitute $j\omega$ for s in the transfer function expression $H(s)$. In the case of the z-transform, one can use the definition of the complex variable z (Chapter 9, Equation (9.24)): $z = e^{j\omega\Delta t}$ to convert $H(z)$ into $H(j\omega)$.

Generally we can obtain the filter/system characteristic by determining the input-output relationship. Using a combined approach, we may study the filter by providing different types of input:

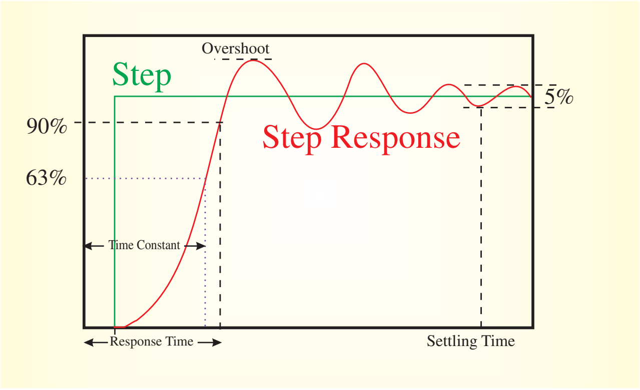

Figure 12.1 Example of a filter response to a unit step.

1. Transients such as the unit impulse δ (Dirac) or, in an experimental setting, the unit step U (Heaviside) function
2. Continuous inputs studied during a steady state (SS):
   a. Sinusoidal inputs using a range of frequencies (as we demonstrated in the example for the RC circuit in Chapter 10)
   b. White noise representing all possible frequencies

When we apply a transient such as the unit step to a filter's input, we may obtain a response as shown in Figure 12.1. The filter's step response in Figure 12.1 is typical for filters with a steep transition from pass band to stop band, and it illustrates the typical overshoot followed by a ripple (Appendix 11.1). As we have already determined, in a passive filter with R and C components, the step response is smoother (Fig. 12.2A). The response to a transient is frequently characterized by the response time (Fig. 12.1) or the RC time (the so-called time constant $\tau = RC$, Figures 12.1 and 12.2A). As will be shown in the following section, the RC value characterizes the transient response of the filter but also relates to the frequency response characterization.

12.2 TIME DOMAIN RESPONSE

In the time domain, the dynamics of the low-pass filter's output are determined by the exponential $e^{-t/RC}$ (e.g., Equations (11.2) and (11.8)). At the time equal to the time constant $t = \tau = RC$, the value of the exponential is $e^{-1} \approx 0.37$. Thus, at $t = \tau$, depending on what direction the output goes (away from 0 or toward 0), the filter output is either at ~37% or ~63% of its final amplitude. In the case of the filter considered in Chapter 10 (R = 10 kΩ, $C = 3.3$ μF), we may find this at $\tau = RC = 33$ ms (Fig. 12.2A).

The following MATLAB script simulates the response of a low-pass filter to a unit impulse and a unit step. Here we use the two different approaches (continuous time and discrete time) discussed in Chapter 11.

```
% pr12_1.m
% Filter Implementations for Impulse and Step response

clear
figure; hold;
% The basis is an analysis of a low-pass RC circuit
% we use R=10k and C=3.3uF
R=10e3;
C=3.3e-6;
RC=R*C;

%                                    UNIT IMPULSE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The analysis compares different approaches to obtain an impulse
% response

% COMPARED ARE:
% 1.   The analog continuous time approach using the Laplace
       transform
%      for the impulse response we obtain 1=RCsH(s)+ H(s)
%      after the inverse transform this is h(t)=(1/RC)*exp(-t/RC)
%      to compare with later discrete time approached, we assume
sample_rate=1000;
dt=1/sample_rate;
time=0.1;

i=1;
for t=dt:dt:time;
   yh(i)=(1/RC)*exp(-t/RC);
   i=I+1;
end;
plot(yh,'k')

% 2.   The difference equation mode
%      The difference equation: x(n*dt)=RC[(y(n*dt)-y(n*dt-1*dt))/dt] +
       y(n*dt)
%      for the algorithm we set n*dt to n and obtain x(n)=[RC/dt]*
[(y(n)-y(n-1))]+y(n)
```

```
A=RC/dt;
x=zeros(1,100);x(1)=1/dt;        % the input is an impulse at t=0 we
                                    correct the input
                                 % with 1/dt

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% To be able to directly compare the analog and discrete impulse
% response, we have to correct the amplitude of the impulse. In case
% of a sampled signal we can assume the impulse to be of duration dt
% and amplitude 1/dt. Therefore either the input (i.e., the impulse)
% or the output (i.e., the impulse response) must be corrected for the
% amplitude % of 1/dt!!
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

y_previous=0;

for n=1:100;
   y(n)=(A*y_previous+x(n))/(A+1);
   y_previous=y(n);
end;
plot(y,'r')

% 3.   The z-domain solution
%      Set the equation in (2) above to the z-domain
%       X(z)              =(A+1)Y(z)-(A/z)Y(z)
%       Y(z)              =1/[(A+1)-A/z]
%      Transformed: y(n)=A^n/(A+1)^(n+1)
for n=1:100;
   yz(n)=A^n/(A+1)^(n+1);
end;
yz=yz/dt;        % Because we calculated yz on the basis of the
                 % discrete impulse, we correct the output with 1/dt
plot(yz,'g')
title('Unit Impulse Response of a Low-Pass Filter')
xlabel('sample#')
ylabel('Amplitude')

%                                   UNIT STEP
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
figure; hold;
% Compared are
% 1.   The analog continuous time approach using the Laplace
        transform
```

```
%       for the impulse response we obtain 1=RCsH(s)+ H(s)
%       after the inverse transform this is h(t)=(1/RC)*exp(-t/RC)
%       to compare with later discrete time approaches we assume
sample_rate=1000;
dt=1/sample_rate;
time=0.1;

i=1;
for t=dt:dt:time;
   yh(i)=1-exp(-t/RC);
   i=i+1;
end;
plot(yh,'k')

% 2.    The difference equation mode
%       The difference equation: x(n*dt)=RC[(y(n*dt)-y(n*dt-1*dt))/dt] +
        y(n*dt)
%       for the algorithm we set n*dt to n and obtain x(n)=[RC/dt]*
[(y(n)-y(n-1))]+y(n)
A=RC/dt;
x=ones(1,100);
y_previous=0;

for n=1:100;
   y(n)=(A*y_previous+x(n))/(A+1);
   y_previous=y(n);
end;
plot(y,'r')
title('Unit Step Response of a Low-Pass Filter')
xlabel('sample#')
ylabel('Amplitude')
```

Note: In the preceding script, we corrected the unit impulse amplitude for the discrete time cases. For a sample interval dt, the amplitude correction is 1/dt; by applying this correction, we obtain an impulse with unit area dt × 1/dt (see also Fig. 2.4A).

12.3 THE FREQUENCY CHARACTERISTIC

The calculated amplitude ratio of the frequency characteristic of a low-pass filter is depicted in Figure 12.2B. The data in this figure are based on a filter with $\tau = 33$ ms.

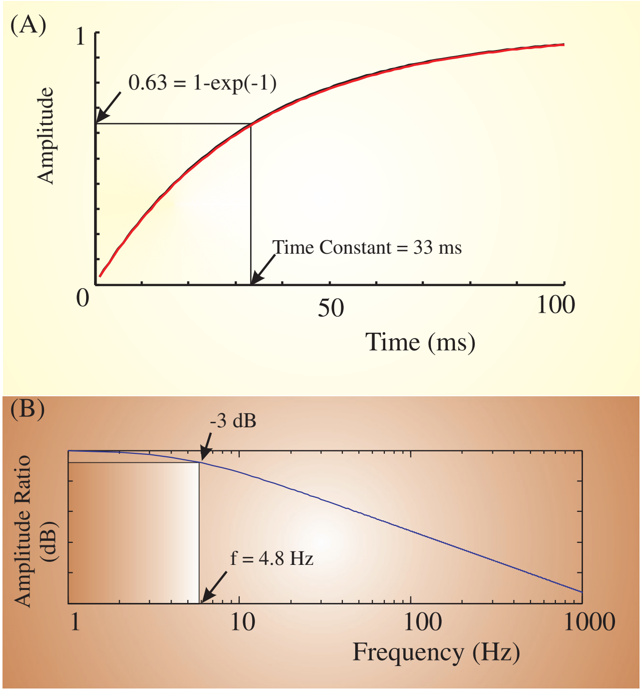

Figure 12.2 A low-pass filter's response to a unit step (A) and its frequency characteristic (B). In the time domain, the time constant τ (= RC) is determined to be [1 − exp(−1)] = 0.6321 of the final output. In the frequency domain, the cutoff frequency at the −3 dB point is at $1/2\pi\tau$ Hz.

In the frequency characteristic shown in Figure 12.2B, we can see that it would be difficult to objectively delimit the precise bands that define the filter specification (see Fig. 10.1). ***For this reason, the transition from pass band to the transition band is conventionally (though arbitrarily) taken to be the so-called −3 dB point;*** this point corresponds with the frequency where the power of the output/input ratio is equal to one half. As we saw in Chapter 3, this attenuation of $\frac{1}{2}$ in the power ratio can be expressed in decibels (Equation 3.12):

$$10\log_{10}\left(\frac{1}{2}\right) = 20\log_{10}\left(\frac{1}{\sqrt{2}}\right) \approx -3\,\mathrm{dB} \tag{12.1}$$

For the low-pass RC filter, we studied in Chapters 10 and 11, the frequency response is (Equation (11.4) or (11.6)):

$$Y(j\omega) = H(j\omega) = \frac{1}{1 + RCj\omega} \tag{12.2}$$

Equation (12.2) includes a complex number that can be split into real and imaginary components by multiplying through by an appropriately chosen fraction equal to 1:

$$\begin{aligned} H(j\omega) &= \frac{1}{1 + RCj\omega} \times \frac{1 - RCj\omega}{1 - RCj\omega} = \frac{1 - RCj\omega}{1^2 + (RC\omega)^2} \\ &= \frac{1}{1 + (RC\omega)^2} - j\frac{RC\omega}{1 + (RC\omega)^2} \end{aligned} \tag{12.3}$$

The magnitude of $H(j\omega)$ (i.e., $|H(j\omega)|$) reflects the amplitude ratio between the filter output and input in the frequency domain. Defining a and jb as the real and imaginary parts of $H(j\omega)$ (Fig. 12.3), we can calculate the power ratio of output/input as the squared amplitude ratio:

$$|H(j\omega)|^2 = H(j\omega)H(j\omega)^* = (a + jb)(a - jb) = a^2 - (jb)^2 = a^2 + b^2$$

The * indicates the complex conjugate. Combining this with Equation (12.3),

$$\begin{aligned} a^2 + b^2 &= \left[\frac{1}{1 + (RC\omega)^2}\right]^2 + \left[\frac{RC\omega}{1 + (RC\omega)^2}\right]^2 \\ &= \frac{1 + (RC\omega)^2}{\left[1 + (RC\omega)^2\right]^2} = \frac{1}{1 + (RC\omega)^2} \end{aligned} \tag{12.4}$$

Following the definition of the –3 dB point, the expression in Equation (12.4) at the transition must equal $\frac{1}{2}$ — that is, the angular frequency ω or the frequency f corresponding with this –3 dB transition is

$$\begin{aligned} &\frac{1}{1 + (RC\omega)^2} = \frac{1}{2} \rightarrow 1 + (RC\omega)^2 = 2 \rightarrow RC\omega = 1 \\ &\rightarrow \omega = 2\pi f = \frac{1}{RC} \rightarrow f = \frac{1}{2\pi RC} = \frac{1}{2\pi\tau} \end{aligned} \tag{12.5}$$

Equation (12.5) relates the value of the time constant ($\tau = RC$) of the transient response with the –3 dB point of the frequency characteristic of the

RC filter. Remember again that the –3 dB point represents the frequency where the power is attenuated by a factor of 2 (Equation (12.1)); the amplitude is therefore attenuated by a factor of $\frac{1}{\sqrt{2}}$ that is, at $\frac{1}{2}\sqrt{2} \approx 0.71$ of the input amplitude. If we define $\omega_{-3dB} = (RC)^{-1}$ and using Equation (12.4), the power ratio $|H(j\omega)|^2$ can be written as

$$|H(j\omega)|^2 = \frac{1}{1 + (\omega/\omega_{-3dB})^2} \tag{12.6a}$$

or using $\omega = 2\pi f$ and $\omega_{-3dB} = 2\pi f_{-3dB}$:

$$|H(j\omega)|^2 = \frac{1}{1 + (f/f_{-3dB})^2} \tag{12.6b}$$

This shows that the simple RC circuit behaves as a first-order ***Butterworth filter*** (see Chapter 13, Sections 13.5 and 13.6). Specifically, Equation (12.6) represents a first-order filter that attenuates with a slope (roll-off) of ~6 dB per octave. An octave is a doubling of frequency; using Equation (12.6), it can be seen that, in the given low-pass filter setup, doubling of the frequency results in an increased attenuation of the output. Let's use an example with a cutoff frequency $f_{-3dB} = 10$ Hz and evaluate what happens to the attenuation factor at a series of 10 Hz, 20 Hz, 40 Hz, 80 Hz, 160 Hz, and so on. Using these values in Equation (12.6b), we get the following series of values for $|H(j\omega)|^2$: 1/2, 1/5, 1/17, 1/65, 1/257, and so on. These series show that (at the higher values for f) doubling of the frequency causes $|H(j\omega)|^2$ to change with a ratio of $\sim\frac{1}{4}$. This ratio corresponds with $10 \times \log_{10}(\frac{1}{4}) \approx -6$ dB, hence the 6 dB/octave characteristic.

In the experimental evaluation in Chapter 10, we found that filters behave as linear systems, generating a sinusoidal output of the same frequency ω as any sinusoidal input. Generally, if one determines the response of a linear system (filter) to a sine wave $A\sin(\omega t + \phi)$, the only parameters that vary between output and input are the amplitude A and the phase ϕ. This aspect of linear systems is frequently summarized in a ***Bode plot*** or a ***Nyquist plot***. Representing the frequency response $H(j\omega)$ as a complex function $a + jb$ (with a and b representing the real and imaginary parts), we have

$$\text{Gain} = \frac{A_{\text{out}}}{A_{\text{in}}} = |H(j\omega)| = \sqrt{a^2 + b^2}\text{, and} \tag{12.7a}$$

$$\text{Phase} = \phi = \tan^{-1}\left(\frac{b}{a}\right) \tag{12.7b}$$

In Equation (12.7a), gain may also represent attenuation (i.e., gain < 1); $\frac{A_{out}}{A_{in}}$ is the ratio between the amplitudes at the output and input; ϕ is the phase shift between output and input. We can also represent the complex number $H(j\omega)$ in polar form:

$$H(j\omega) = |H(j\omega)|^{e^{j\phi}} \tag{12.8}$$

A diagram of the polar representation for a single frequency ω is shown in Figure 12.3. The equations in (12.7) and the expression in Equation (12.8) are essentially equivalent in the sense that they fully specify the frequency characteristic of the RC filter with a complex value (such as the one depicted in Fig. 12.3) for each frequency ω.

Equation (12.7) is the basis for the so-called Bode plot, where the frequency characteristic is represented by two separate plots. One of the plots describes the amplitude ratio between output and input $|H(j\omega)|$ versus frequency (as in Figs. 12.2B and 12.4A). A second plot describes the phase ϕ versus frequency (Fig. 12.4B). Usually the abscissa of a Bode plot is a $\log_{10}$ axis of frequency f (= $\omega/2\pi$). Another representation of the same information is the Nyquist plot, which depicts the $H(j\omega)$ function (Equation (12.8)) in a polar plot (Fig. 12.4C). The advantage of the Nyquist plot over the Bode plot is that all information is contained in a single plot instead of two; the disadvantage is that the frequency axis is not explicitly included. In most cases, an arrow in the Nyquist plot (as in Fig. 12.4C) indicates the direction in which the frequency increases, allowing for a qualitative assessment of the frequency-related output/input relationship.

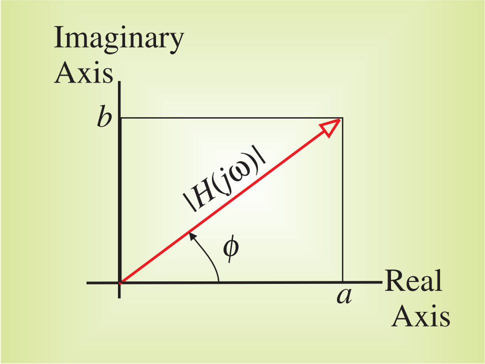

Figure 12.3 Argand diagram of a frequency response function (red arrow) at a given frequency ω. The frequency response is a complex-valued number $a + jb$, which can also be represented in polar coordinates by magnitude $|H(j\omega)|$ and phase ϕ.

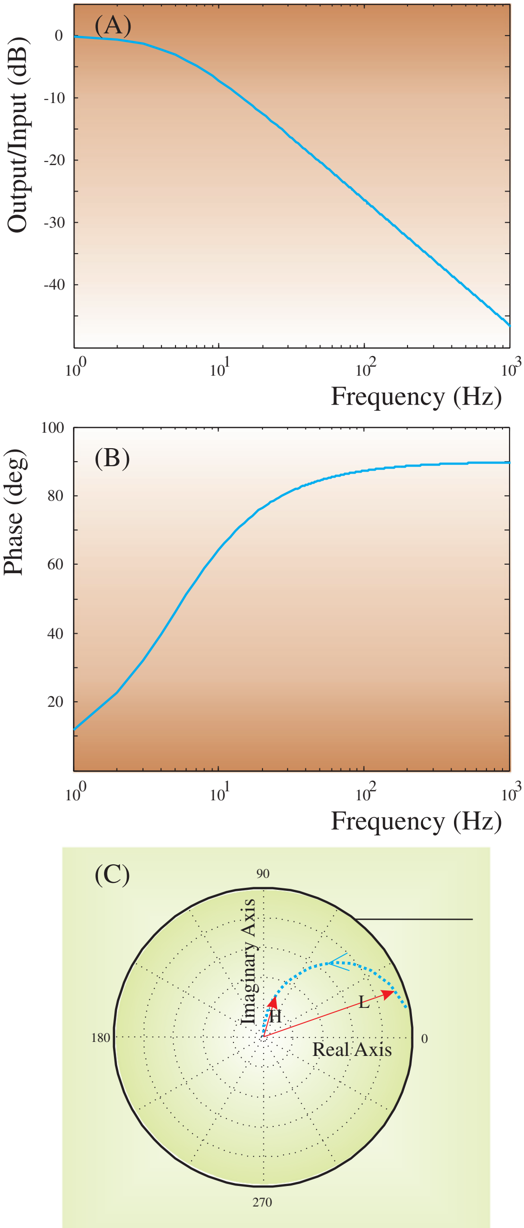

Figure 12.4 Filter characteristic of the RC-circuit low-pass filter. The Bode plot: (A) amplitude ratio and (B) the phase relationship between output and input. From the graphs in the Bode plot, the amplitude ratio and phase can be determined for each frequency. (C) An example of a Nyquist diagram that shows the output/input relationship (blue dotted line) of the same filter in a polar plot. The Nyquist diagram shows the frequency characteristic (such as the one depicted in Figure 12.3) as a complex-valued parametric function of frequency. In this type of plot, specific frequency values cannot be determined; the blue arrow indicates the direction in which the frequency increases. In this example, we indicate a low frequency (L) with an output/input ratio close to one and a small change in phase. The high frequency (H) is associated with a smaller ratio (it is a low-pass filter characteristic) and a more significant change in phase.

The following MATLAB program can be used to produce the graphs shown in Figure 12.4.

```
% pr12_2.m
% Bode_Nyquist.m
% Bode Plot and Nyquist Plot for a low-pass filter

% Filter Components
R=10e3;
C=3.3e-6;

% Formula for amplitude (A) = 1/sqrt[1 + (RCw)^2] with w=2 x pi x f
for i=1:5000;
   f(i)=i;
   A(i)=1/(sqrt(1+(R*C*2*pi*f(i))^2));      % formula derived for the
absolute part
   H(i)=1/(1+R*C*2*pi*f(i)*j);              % frequency response
   rl=real(H(i));                           % real part of H
   im=abs(imag(H(i)));                      % magnitude of the imaginary
                                              part of H
   PHI(i)=atan(im/rl);                      % phase
end;

% for w=1/RC there is A=1/sqrt[1/2] ~ 0.7
% 20 x log10{[1/sqrt(2)]} ~ -3.0 (The -3dB point)
F_3db=1/(2*pi*R*C);    % Here we use frequency F (=w/(2 x pi))

figure
subplot(3,1,1), semilogx(f,20*log10(A))
xlabel('Frequency(Hz)')
ylabel('Amplitude Ratio (dB)')
axis([0 1000 -50 0]);

t=['BODE PLOT Low Pass Filter: R = 'num2str(R)' Ohm;
C = 'num2str(C)' F; and -3dB frequency = 'num2str(F_3db)' Hz'];
title(t)
subplot(3,1,2),semilogx(f,(PHI*360)/(2*pi))
xlabel('Frequency(Hz)')
ylabel('Phase (degrees)')
axis([0 1000 0 100]);

subplot(3,1,3),polar(PHI,A)
xlabel('Real')
ylabel('Imaginary')
title('Nyquist Plot')
```

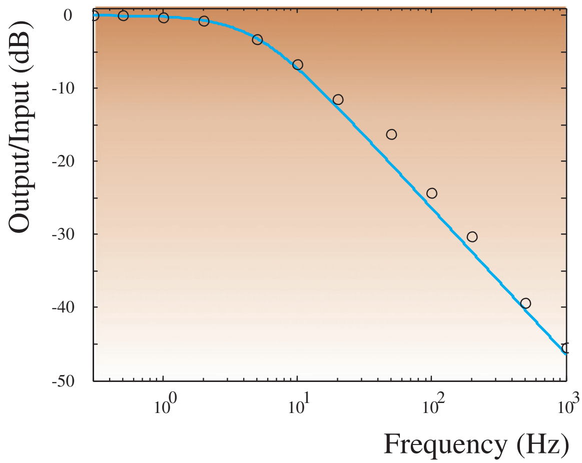

Figure 12.5 The amplitude ratio part of the Bode plot for a low-pass filter. The open circles are the measured values from the experiment described in Chapter 10 (Table 10.1); the blue line is the theoretically derived response $|H(j\omega)|$, Equation (12.7a). The discrepancy between measured and calculated data is due to measurement error made by the author.

Finally we can relate the experimental findings we obtained in Chapter 10 in which we measured the response of the RC filter to sinusoidal inputs. The ratios of output amplitude to input amplitude we calculated there represent the output/input ratio graph $|H(j\omega)|$ of the Bode plot. These measurements can be compared with the theoretical expectation based on the known resistance and capacitance values. The MATLAB program pr12_3.m shows our experimental results superimposed on the theoretical curve; you can use pr12_3.m to plug in your own recorded values; your results should look similar to Figure 12.5.

12.4 NOISE AND THE FILTER FREQUENCY RESPONSE

In the previous sections we analyzed an RC filter's response to either a transient signal (such as δ or U) or a sine wave. The autospectrum (= power spectrum) of the response to white noise input can also be used to obtain the frequency characteristic. In the frequency domain, truly white noise represents all frequencies. Because the noise is random, subsequent samples are unrelated and its autocorrelation function is a delta function: (i.e., correlation equal to 1 at zero lag and 0 elsewhere). The Fourier transform of this autocorrelation function represents the power spectrum

(Chapter 8, Section 8.4.2), and the transform of a delta function is a constant (Equation (6.9)). Therefore a sufficiently long epoch of white noise provides all the frequencies to the filter's input, similar to feeding it sinusoidal signals for a range of frequencies as we discussed earlier. The output of the filter will be "colored" noise, meaning that frequency components will be attenuated in the transition and stop bands of the filter. At first sight the noise approach may seem a bit sloppy, but the results from this approach can easily be compared with other mathematical techniques: (1) the Fourier transform of the impulse response in continuous time systems or (2) by substituting $e^{j\omega t}$ for z in the z-transform of the impulse response in discrete time systems (see also Chapter 13, Section 13.4). You may run script pr12_4.m to compare the different techniques. Note that the result from the noise input technique may slightly vary each time that you run the script.

The following is a MATLAB script that demonstrates the different techniques for obtaining the output/input ratio for a digital filter:

```
% pr12_4.m
% Two-point Smoothing Filter's Frequency Response
% filter equation for the digital (FIR) filter:
%       y(n)=(x(n)+x(n-1))/2

clear;
wT=0:.1:2*pi;

% 1.   Calculate the abs component of the fft of the impulse response
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The impulse response y for an impulse at time=0
% equals to pulse of .5 at time=0 and at time=T,
% i.e.
y=zeros(1,63);
y(1)=.5;y(2)=.5;

Y=fft(y);                          % fft of the impulse response y
                                   % –> frequency response Y
figure
subplot(2,1,1),plot(wT,abs(Y))
title('Frequency Response = fft of the Impulse Response')
axis([0 max(wT) 0 max(abs(Y))]);
ylabel('Amplitude Ratio');

% 2.   The second method is to use the z-transform and replace z by
       exp(jwT)
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The z-transform of y(n)=(x(n)+x(n-1))/2 is:
%                    Y(z)=.5* X(z)*[1+1/z]
%          –>        H(z)=Y(z)/X(z)=.5 + .5*(1/z) = .5 + .5*exp(-jwT)
YY=(.5+.5*(exp(-j*wT)));
subplot(2,1,2),plot(wT,abs(YY))
title('Frequency Response = based on z-transform')
axis([0 max(wT) 0 max(abs(YY))]);
ylabel('Amplitude Ratio');
xlabel('Frequency (wT: Scale 0-2pi)');    % NOTE: Normally one would
                                            show 0-pi
                                          % with pi=the Nyquist
                                            frequency

% 3.  The third method is to use white noise and compare the power
      spectra of in- and output
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Because white noise represents all frequencies at the input and
% the output shows what is transferred. The output over the input
% power spectra represent a frequency response estimate.

x=randn(10000,1);                          % create white noise
wT=1:length(x);wT=(wT/length(x))*2*pi;     % New Frequency Scale

for n=2:length(x);                         % Calculate the output of
                                             the 2-point
                                           % smoothing
    y(n)=(x(n)+x(n-1))/2;
end;

figure  % plot the input x
subplot(3,1,1),plot(x)
hold
subplot(3,1,1),plot(y,'k')
title('Input Noise (blue) and Output Noise (black)')
xlabel('sample #')
ylabel('amplitude')
X=fft(x);        % Calculate the power spectra
Y=fft(y);        % NOTE: The power spectrum is the
                 % fft of the autocorrelation
Px=X.*conj(X)/length(x);
Py=Y.*conj(Y)/length(y);
```

```
subplot(3,1,2), plot(wT,Px)
hold

subplot(3,1,2),plot(wT,Py,'k')
title('POWER SPECTRA Input Noise (blue) and Output Noise (black)')
xlabel('Frequency Scale (0-2pi)')
ylabel('power')

for i=1:length(x);
   h_square(i)=Py(i)/Px(i);
end;
subplot(3,1,3),plot(wT,sqrt(h_square), 'k')
title('Frequency Response = based on Input-Output white Noise')
xlabel('Frequency Scale(0-2pi)')
ylabel('Amplitude Ratio')
```