4-9 DESIGN OF MINIMUM-PHASE FIR FILTERS

- The advantage of the Type I and Type II linear-phase FIR filters is that their delay is a constant and, thereby, they cause no phase distortion to the signal.
- The delay is equal to half the filter order. This means that the delay becomes very long for high-order filters required in cases demanding a narrow transition band. In some applications, such a long delay is not tolerable.
- In those cases, a smaller group delay can be achieved in the passband region by using minimum-phase FIR filters.
- There exist also applications where linear phase is not required and the symmetry in the coefficients of linear-phase FIR filters cannot be exploited.

- In those cases, nonlinear-phase filters meet the same amplitude criteria with a reduced number of multipliers and delay elements.
- If the passband of the filter is very wide, a saving by almost a factor of 2 can be achieved in the filter order
- This section outlines the design of nonlinear-phase filters based on the design scheme of Herrmann and Schüssler which can be used for synthesizing filters with unweighted stopband response.
- There are also more general techniques.

Nonlinear-phase FIR Filters

- Consider a nonlinear-phase FIR filter with transfer function

$$H(z) = \sum_{n=0}^{M} h[n]z^{-n}.$$

- The zeros of the transfer function

$$\widehat{H}(z) = z^{-M}H(z^{-1}) = \sum_{n=0}^{M} h(M-n)z^{-n}$$

are reciprocal to those of $H(z)$.

- This implies that the function

$$G(z) = z^{-M}H(z)H(z^{-1})$$

is a transfer function of a Type I linear-phase filter of order $2M$.

- Since $G(z)$ must be factorizable into the terms $H(z)$ and $\widehat{H}(z)$, its zeros on the unit circle have to be double.

- From the above equation, it follows that the magnitude-squared function of $H(z)$ can be expressed as

$$\left|H(e^{j\omega})\right|^2 = G(\omega).$$

- Since $G(z)$ possesses double zeros on the unit circle, $G(\omega)$ has double zeros on $[0, \pi]$, making it non-negative on $[0, \pi]$.
- These facts show that the design of a nonlinear-phase FIR filter of order M can be accomplished in terms of a Type I filter of order $2M$ and having double zeros on the unit circle.
- Based on this, Herrmann and Schüssler have proposed a simple design procedure.

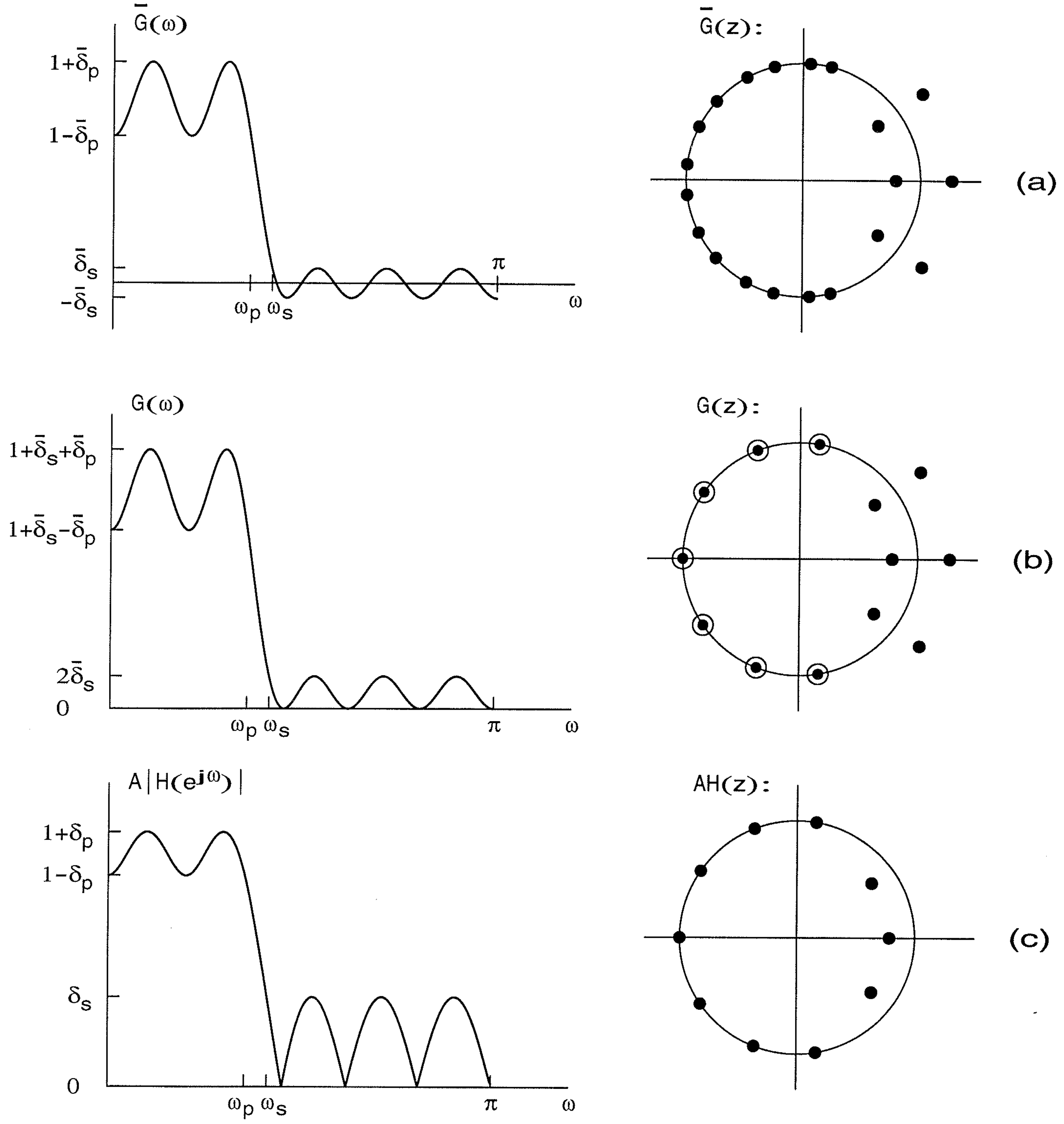
Simple Design Procedure

1. Design a linear-phase FIR filter transfer function of order $2M$ using the MPR algorithm such that $\bar{G}(\omega)$ oscillates within the limits $1 \pm \bar{\delta}_p$ in the passband $[0, \omega_p]$ and within the limits $\pm\bar{\delta}_s$ in the stopband $[\omega_s, \pi]$ [see Figure (a) on page 175]. This $\bar{G}(z)$ has single zeros on the unit circle.

2. Form $G(z) = \bar{\delta}_s z^{-M} + \bar{G}(z)$. The resulting $G(\omega) = \bar{\delta}_s + \bar{G}(\omega)$ is non-negative on $[\omega_s, \pi]$ oscillating within zero and $2\bar{\delta}_s$ [see Figure (b) on page 175]. On $[0, \omega_p]$, $G(\omega)$ oscillates within the limits $1 + \bar{\delta}_s \pm \bar{\delta}_p$. $G(\omega)$ has double zeros at those points where $\bar{G}(\omega)$ takes the stopband minimum value of $\bar{\delta}_s$. Correspondingly, $G(z)$ has double zeros on the unit circle at the frequencies where $G(\omega)$ has double zeros.

3. Perform the factorization of $G(z) = H(z)z^{-M}H(z^{-1})$ such that $H(z)$ contains the zeros inside the unit circle and one each of the double zeros on the unit circle. Scale $H(z)$ such that the passband average of the resulting filter $AH(z)$ is equal to unity [see Figure (c) on page 175].

Figure illustrating the design procedure



- The desired scaling constant at Step 3 is

$$A = \frac{2}{\sqrt{1+\bar{\delta}_p+\bar{\delta}_s}+\sqrt{1-\bar{\delta}_p+\bar{\delta}_s}}.$$

- If it is desired that the magnitude response of the scaled filter $AH(z)$ approximates unity in the passband with tolerance δ_p and zero in the stopband with tolerance δ_s, then the passband and stopband ripples of the linear-phase filter at Step 1 must satisfy

$$\bar{\delta}_p \le \frac{2\delta_p}{1+(\delta_p)^2+(\delta_s)^2/2}, \quad \bar{\delta}_s \le \frac{(\delta_p)^2/2}{1+(\delta_p)^2+(\delta_s)^2/2}.$$

- The most difficult part in the above procedure is the factorization of $G(z)$ into the terms $H(z)$ and $z^{-M}H(z^{-1})$.

- The direct approach is to simply pick up the zeros of $G(z)$.

- However, if the order of $G(z)$ is high, conventional root-finding procedures cannot be used for finding the roots.

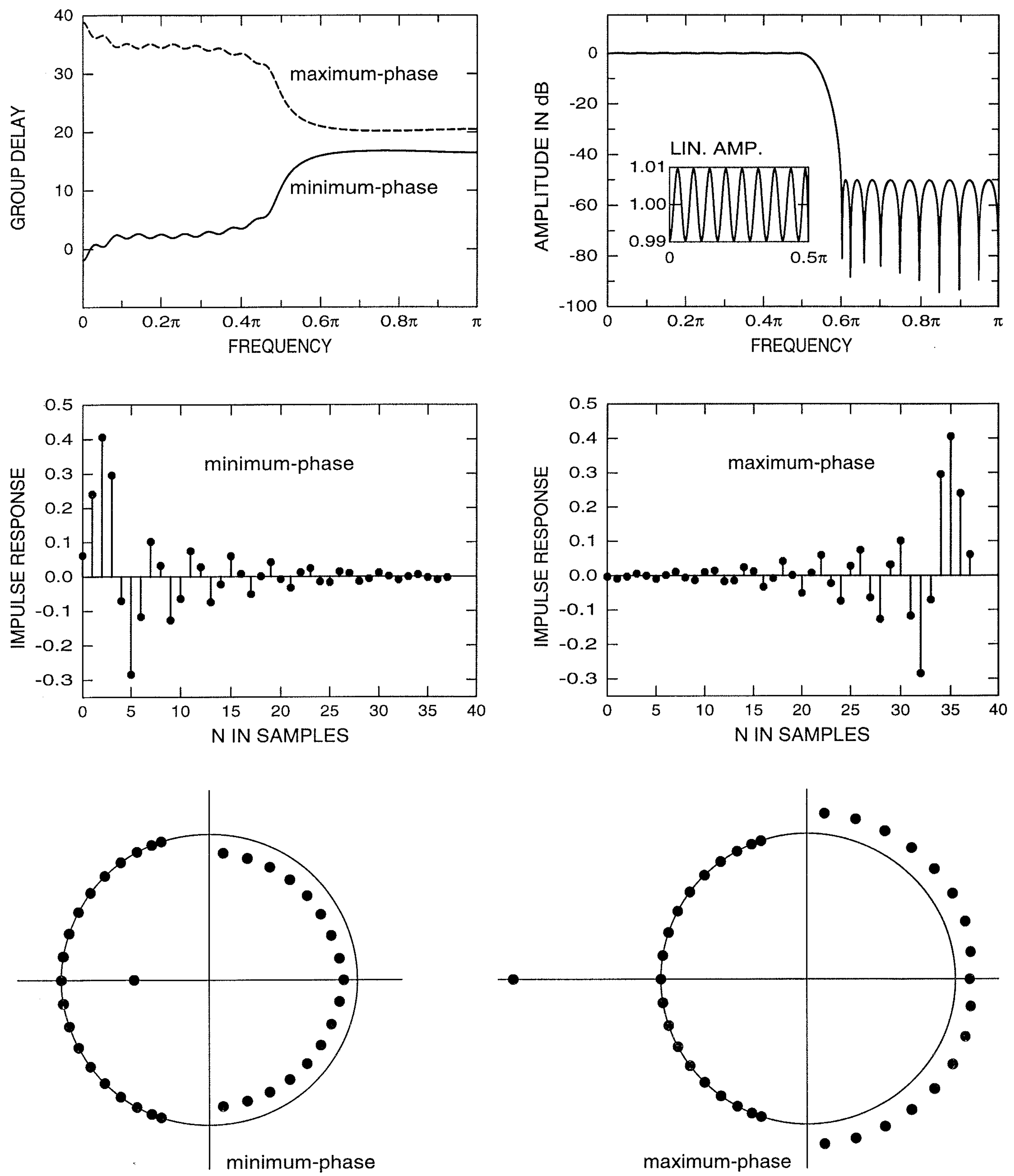
- Another approach is to perform the factorization without finding the roots of $G(z)$ (see the book chapter for references).
- The filter obtained by selecting the zeros to lie on or inside the unit circle is called a *minimum-phase* filter.
- If the zeros outside the unit circle are selected, then the resulting filter is called a *maximum-phase* design.

Example

- Let the specifications be: $\omega_p = 0.4\pi$, $\omega_s = 0.5\pi$, $\delta_p = 0.01$, and $\delta_s = 0.00316$ (50 dB attenuation).
- The ripples of $\bar{G}(z)$ at Step 1 become $\bar{\delta}_p \approx 0.02$ and $\bar{\delta}_s \approx 5 \cdot 10^{-6}$.
- The minimum even-order to meet these criteria is 74 so that the order of the corresponding minimum-phase filter is 37.
- The common amplitude response of the minimum-phase and maximum-phase filters as well as their group delay responses, zero locations, and impulse responses are given in the following transparency.
- The minimum order of a linear-phase filter to meet the same criteria is 46 so that the saving in the filter order provided by the minimum-phase filter is 20 percent.

- For the minimum- and maximum-phase filters, the phase and the group delay responses have the smallest and largest values, respectively, among the filters having the same amplitude response.
- It is also interesting to observe from the figure that most of the energy of the impulse response of the minimum-phase (maximum-phase) term is concentrated in the beginning (end) of the response.
- Pages 181, 182, and 183 illustrate how a minimum-phase FIR filter can be generated using a matlab code.
- In this code, the minimum-phase filter is scaled in such a way that its amplitude response satisfies $|H(e^{j\omega})| = \sqrt{G(\omega}$ so that the first equation on page 176 is not actually needed.

Responses for minimum- and maximum-phase FIR filters of order 37



```
% Matlab m-file (minfir.m) for designing minimum-phase FIR
%lowpass filters. A Remez-routine, called reme.m, containing
%more grid points is used together with the Herrmann-Schussler-
%synthesis technique.
%Tapio Saram"aki 11.2.1996
%Can be found in Sun's: ~ts/matlab/dsp
disp('Hi there')
disp('I am an program for designing minimum-phase')
disp('lowpass FIR filters')
omp=input('Passband edge as a fraction of pi= ');
oms=input('Stopband edge as a fraction of pi= ');
dp=input('Passband ripple= ');
ds=input('Stopband ripple= ');
den=1+dp^2-ds^2/2;
Dp=2*dp/den;Ds=(ds^2/2)/den;
%Initial order for the prototype linear-phase
%FIR filter
[n,f0,m0,w]=remezord([omp,oms],[1,0],[Dp,Ds]);
%Order must be even.
if n > 2*floor(n/2) n=n+1;end
h=reme(n,f0,m0,w);
%Test whether the criteria are met
isu=0;
[a,z]=zeroam(h,oms,oms,1);
if a > Ds isu=1;end
increase=1;
if isu==0 increase=0;hs=h;ns=n;end
if increase==1
ll=0;
   while ll<1
   n=n+2;
   h=reme(n,f0,m0,w);
   isu=0;
   [a,z]=zeroam(h,oms,oms,1);
   if a > Ds isu=1;end
   if isu==0 ll=1;hs=h;ns=n;end
   end
end
if increase==0
ll=0;
   while ll<1
   n=n-2;
   h=reme(n,f0,m0,w);
   isu=0;
   [a,z]=zeroam(h,oms,oms,1);
   if a > Ds isu=1;end
   ll=1;
   if isu==0 ll=0;hs=h;ns=n;end
   end
end
```

```
%
h=hs;n=ns;
%
AS=20*log10(Ds);
[H,f]=freqz(h,1,2048,2);
figure(1);subplot(211)
plot(f,20*log10(abs(H)));grid;axis([0 1 3*AS/2 10 ]);
title(['Amplitude response for the prototype filter of order ', num2str(n)]);
xlabel('Frequency as a fraction of pi'); ylabel('Amplitude in dB')
subplot(212);zplane(h);title('Zero plot')
%Actual stopband ripple; find the minimum of
%the zero-phase frequency response
[H,f]=zeroam(h,oms,1.,30000);
stop=min(H);
stop=abs(stop)
%Because the Remez in Matlab is using too few grid points
%and the stopband response is not exactly equiripple!!
%Therefore, the zeros of the separable filter are not
%exactly double on the unit circle
%Actual passband ripple
[a,z]=zeroam(h,omp,omp,1);
pass=1-a;
%Increase the value of the center tap by stop
h(n/2+1)=h(n/2+1)+stop;
%Resulting linear-phase filter
figure(2),zplane(h);
title('Zeros for the separable filter');
[H,f]=freqz(h,1,2048,2);
%scale the filter such that the passband average
%is 1
a1=1+pass+stop;
a2=1-pass+stop;
a=(a1+a2)/2;
h=h/a;
[H,f]=freqz(h,1,2048,2);
figure(3);subplot(211);
plot(f,20*log10(abs(H)));grid;axis([0 1 3*AS/2 10 ]);
title('Amplitude response for the scaled separable filter');
xlabel('Frequency as a fraction of pi'); ylabel('Amplitude in dB')
subplot(212);plot(f,(abs(H))),grid;axis([0. omp 1-1.1*Dp 1+1.1*Dp]);
title('Passband amplitude response');
xlabel('Frequency as a fraction of pi'); ylabel('Amplitude')
%
%desired value for the minimum phase filter at DC; used later for
%scaling the minimum phase filter
[H,f]=freqz(h,1,100,2);
dc=sqrt(abs(H(1)))
%Roots
r=sort(roots(h));
rm1=[ ]; yy= [ ];
```

```
for i=1:length(r)
 if abs(r(i))< 0.98, % roots inside the unit circle
  rm1(length(rm1)+1)=r(i);
 end
 if abs(r(i))>0.98 & abs(r(i))< 1/.98,
  yy(length(yy)+1)=r(i); % roots on the unit circle
 end                 % to yy
end
%
rm2=yy(1:length(yy)/2);   %one out of the double zeros
rm=[rm1 rm2];            %to the minimum phase filter
Hmin=poly(rm);
figure(4)
zplane(Hmin);
title(['Zeros for the minimum-phase filter of order ', num2str(n/2)]);
%scaling
Hmin=dc*Hmin/sum(Hmin);
%Overall response
figure(5);subplot(211)
[H,f]=freqz(Hmin,1,2048,2);
plot(f,20*log10(abs(H))),grid;axis([0 1 3*AS/4 10]);
title('Response for the minimum-phase filter');
xlabel('Frequnecy as a fraction of pi'); ylabel('Amplitude in dB')
%Passband details
subplot(212);plot(f,(abs(H))),grid;axis([0 omp 1-dp 1+dp]);
title('Passband details for the minimum-phase filter');
xlabel('Frequency as a fraction of pi'); ylabel('Amplitude')
figure(6)
subplot(211)
grpdelay(Hmin,1,2048); axis([0 1 -n/2 n/2]);
title('Group delay response for the minimum-phase filter');
xlabel('Frequnecy as a fraction of pi'); ylabel('Delay in samples')
subplot(212);impz(real(Hmin));
title('Impulse response for the minimum-phase filter');
xlabel('n in samples'); ylabel('Impulse response')
```