Assignment-5
CRL707 Human and Machine Speech Communication

Parveen Bajaj
2018EET2574
2019-20

October 25, 2019

Departement of Electrical Engineering
IIT Delhi, India
2019-20

Contents

List of Figures

1 Computer assignment

Extract 90 ms or four frames of speech data at 22.5 ms per frame from voiced segments of recorded speech in your voice, at 8 KHz sampling rate. For each frame of speech,

a) **Mimicking LPC-10:** Obtain the parameter values that are computed (and transmitted after quantization) by the LPC-l0 vocoder. Note that there is no need to quantize the parameters

b) **Line Spectral Frequencies (LSFs):** For each frame, form the P(z) and Q(z) polynomials from the inverse filter A(z), for each segment. On the same graph (one for each segment), plot the roots of polynomials P(z) and Q(z) in the z-plane. What are the LSFs in Hz? Plot the spectral envelope 1/A(z) for each frame and superimpose the LSFs on the plots? How are the two related?

Label all axes properly. Explain in brief the calculations for parts (a) and (c).

2 Solution

Digit one was chosen to perform this exercise. Original signal of 16 khz sampling rate was resampled to 8 khz sampling rate. Thereafter, energy plot for each frame of 22.5 msec (22.5 * 8 = 180 samples) was generated to decide which 04 frames to be taken for performing the exercise. The results achieved during the programming exercise are shown below.

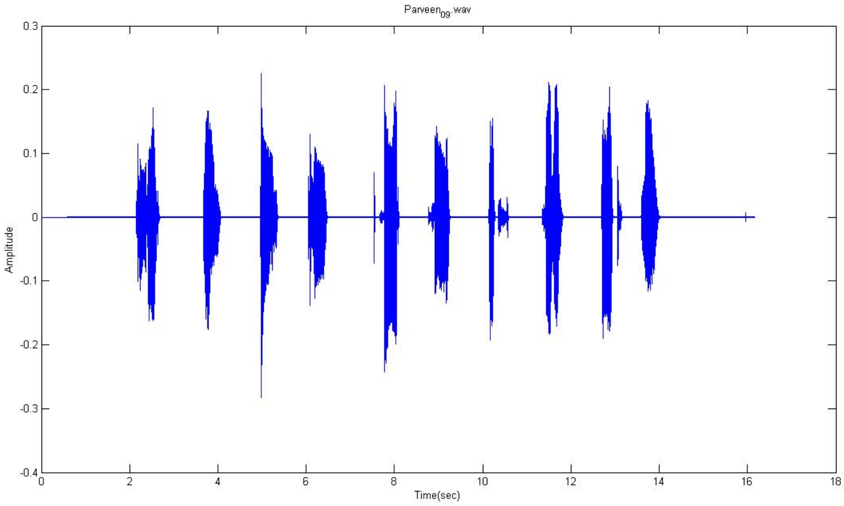

Figure 1: Original Speech Waveform at 16 KHz sampling rate

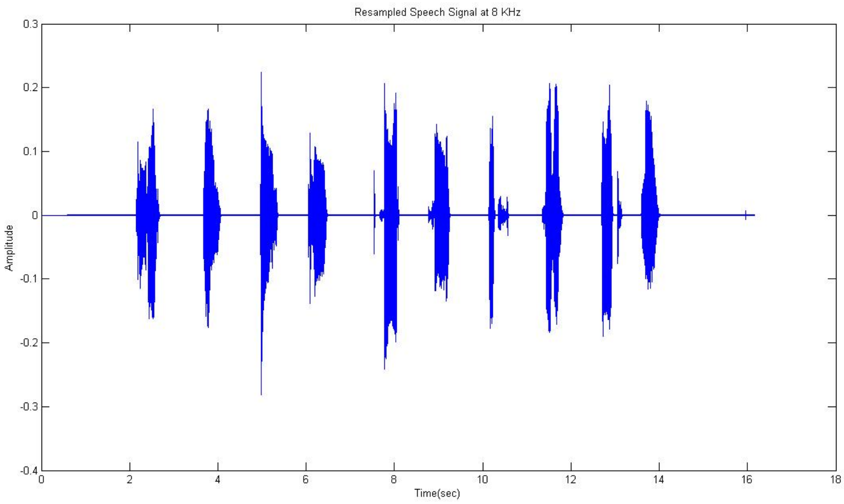

Figure 2: Resampled Speech Signal at 8 KHz sampling rate

2.1 Extract 90 msec or 04 Frames of 22.5 msec from digit-1 voiced segment

Four voiced segments were selected from digit 1 based on energy profile of each frame of digit 1. Frame numbers 4,5,6 and 7 were chosen as they were having high energy profile. The energy profile of each frame in digit - 1 along with four selected frames amplitude profiles are shown below: -

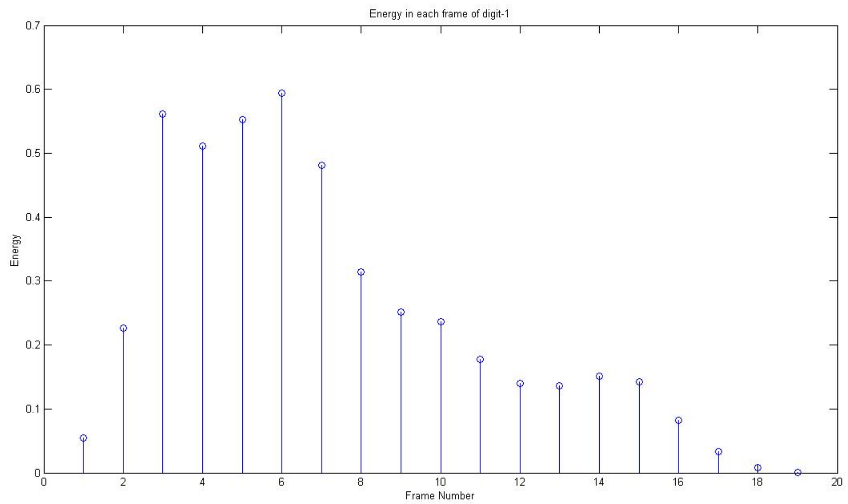

Figure 3: Energy Profile of Each Frame in digit 1

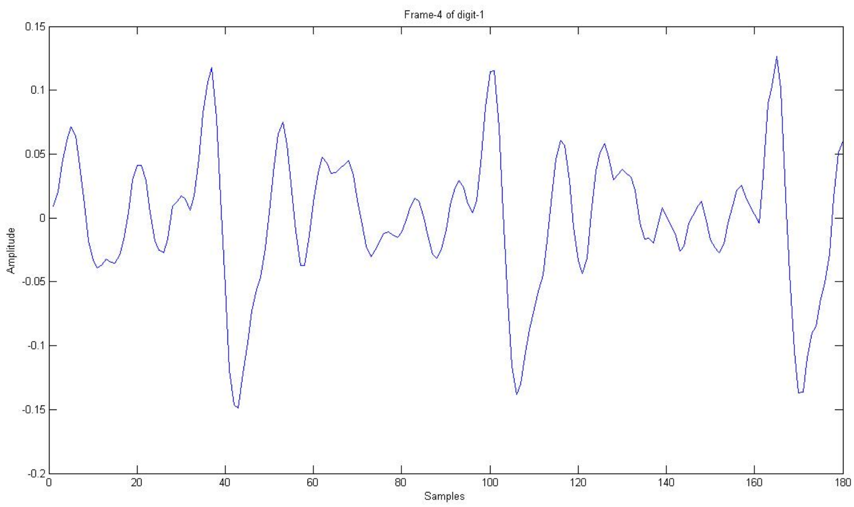

Figure 4: Amplitude profile of Frame 4 in digit 1

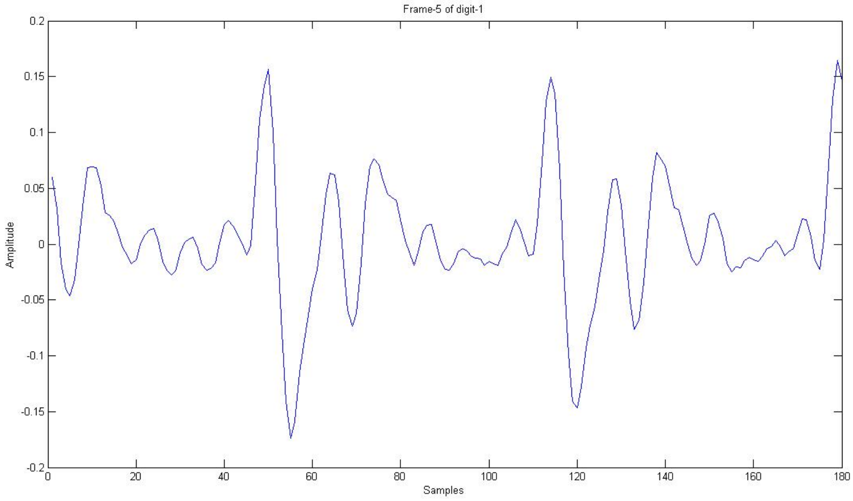

Figure 5: Amplitude profile of Frame 5 in digit 1

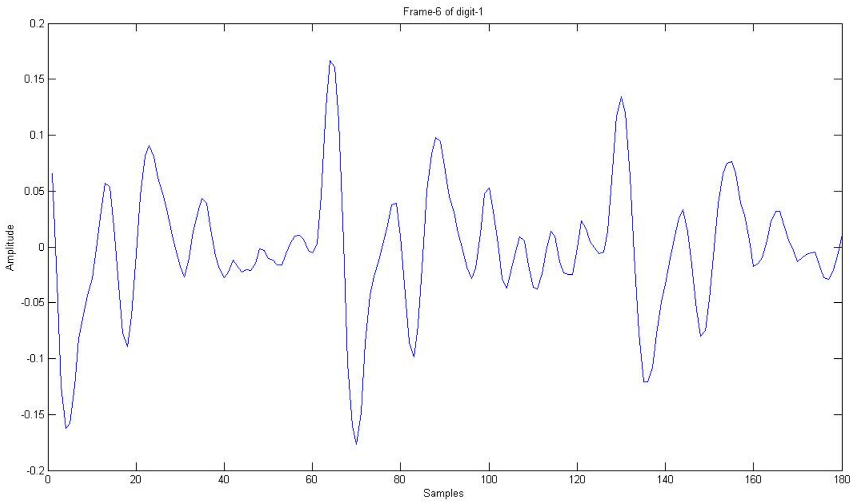

Figure 6: Amplitude profile of Frame 6 in digit 1

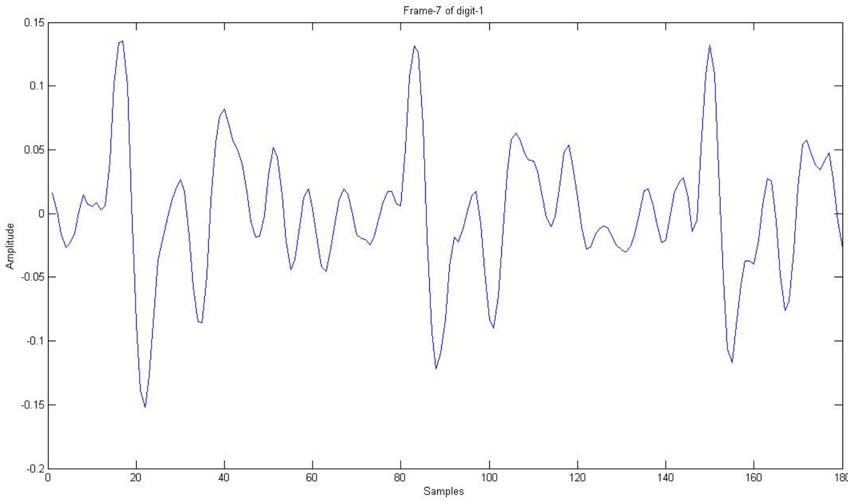

Figure 7: Amplitude profile of Frame 7 in digit 1

2.2 Mimicking LPC-10

Obtain the parameter values that are computed (and transmitted after quantization) by the LPC-l0 vocoder.
The parameters that are mentioned in LPC-10 Vocoder are:

- LPC Coefficients
- Reflection Coefficients
- Line Spectral Frequencies

2.2.1 Frame-4 of Digit 1

- LPC Coefficents
  $[1.0000, -1.7422, 0.8761, -0.0896, 0.5489, -0.5504, -0.0649, -0.0453, 0.5006, -0.2985, 0.0322]$
- Reflection Coefficients
  $[-0.8995, 0.8337, -0.1991, -0.0863, -0.1183, 0.3375, 0.2726, 0.0551, -0.2428, 0.0322]$

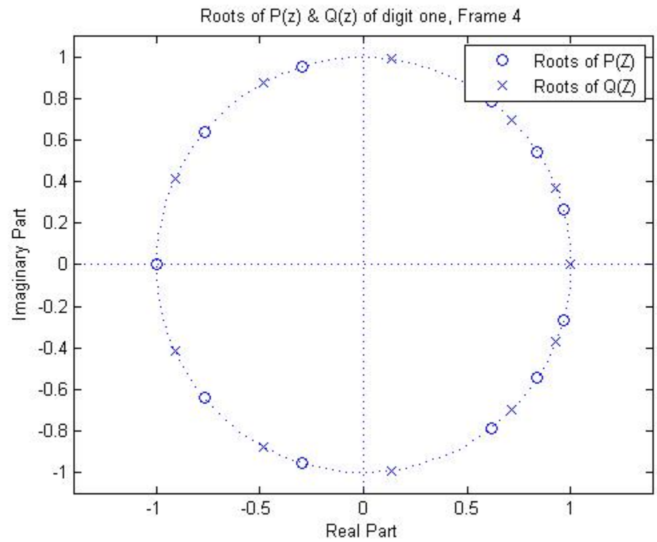

Figure 8: Plot of roots of polynominal P(z) and Q(z) for frame 4 in Digit 1

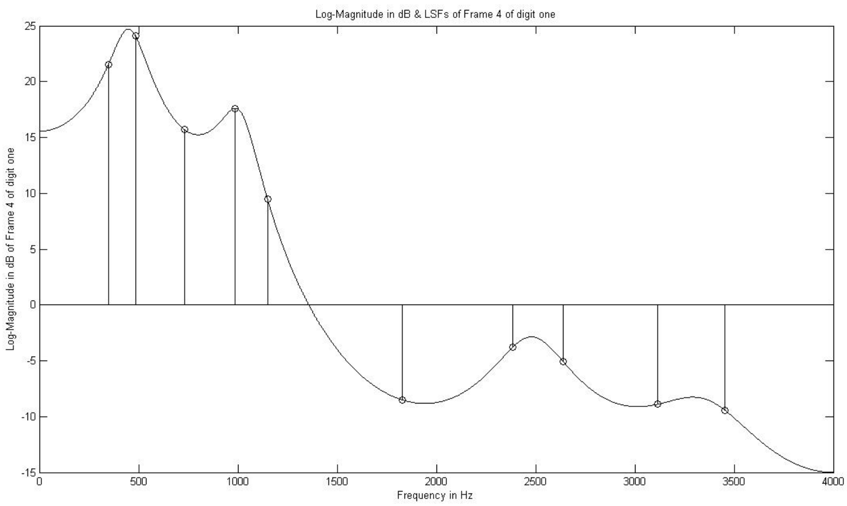

Figure 9: Plot of the Spectral Envelope of 1/A(z) with superimposed LSF's for frame 4 in Digit 1

- The LSF's (in Hz) are:
$[347.57, 484.73, 732.55, 986.29, 1150.47, 1828.06, 2385.58, 2638.41, 3114.10, 3455.30]$

2.2.2 Frame-5 of Digit 1

- LPC Coefficents
  $[1.0000, -1.5357, 0.7082, 0.0101, 0.2485, -0.2423, -0.0828, 0.0549, 0.1539, -0.0542, -0.0305]$
- Reflection Coefficients
  $[-0.8651, 0.7659, -0.0395, -0.0622, -0.1040, 0.1608, 0.1600, 0.0203, -0.1011, -0.0305]$

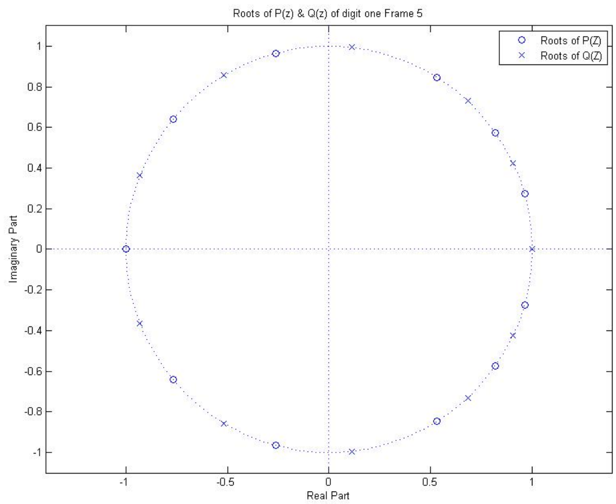

Figure 10: Plot of roots of polynominal P(z) and Q(z) for frame 5 in Digit 1

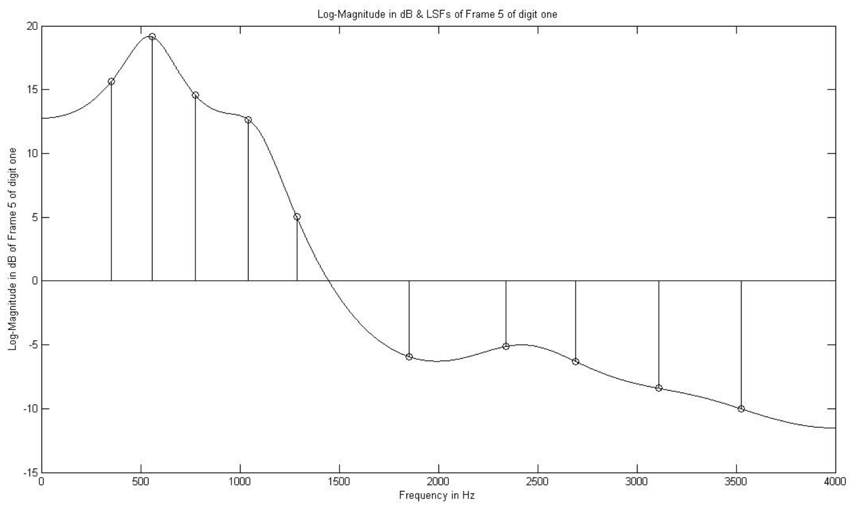

Figure 11: Plot of the Spectral Envelope of 1/A(z) with superimposed LSF's for frame 5 in Digit 1

- The LSF's (in Hz) are:
  [352.8, 557.9, 776.6, 1043.5, 1286.7, 1854.3, 2339.1, 2693.2, 3112.1, 3526.6]

2.2.3 Frame-6 of Digit 1

- LPC Coefficents
  [1.0000, −1.6707, 0.9501, −0.1163, 0.2298, −0.2361, −0.1019, 0.1567, 0.0052, 0.0510, −0.0999]
- Reflection Coefficients
  [−0.8652, 0.8150, −0.1616, −0.0530, −0.0642, 0.1718, 0.0991, −0.0971, −0.1170, −0.0999]

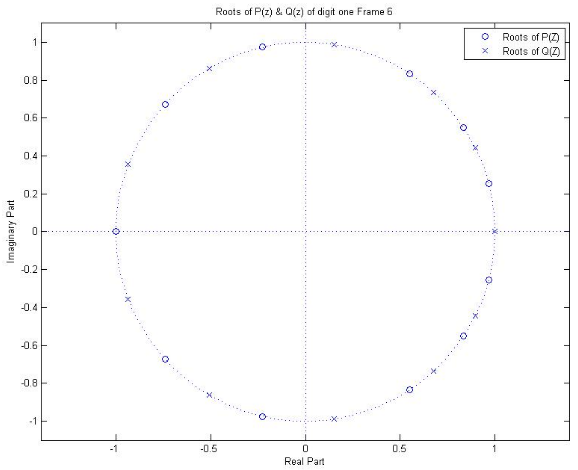

Figure 12: Plot of roots of polynominal P(z) and Q(z) for frame 6 in Digit 1

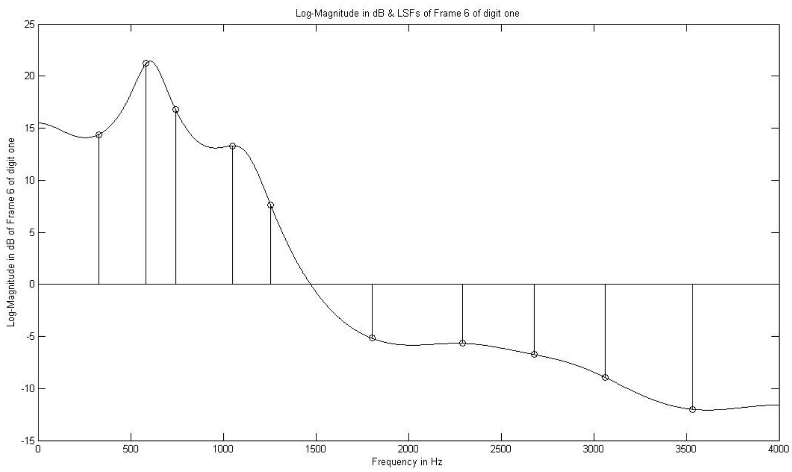

Figure 13: Plot of the Spectral Envelope of 1/A(z) with superimposed LSF's for frame 6 in Digit 1

- The LSF's (in Hz) are:
  [328.2, 584.3, 743.8, 1051.5, 1257.5, 1804.7, 2290.9, 2677.9, 3060.8, 3535.8]

2.2.4 Frame-7 of Digit 1

- LPC Coefficents
  [1.0000, −2.1468, 1.8969, −0.8372, 0.5472, −0.5033, 0.1112, −0.1337, 0.7076, −0.7423, 0.2499]
- Reflection Coefficients
  [−0.8546, 0.8660, −0.5503, 0.1599, 0.0356, 0.2930, 0.0369, −0.2209, −0.2196, 0.2499]

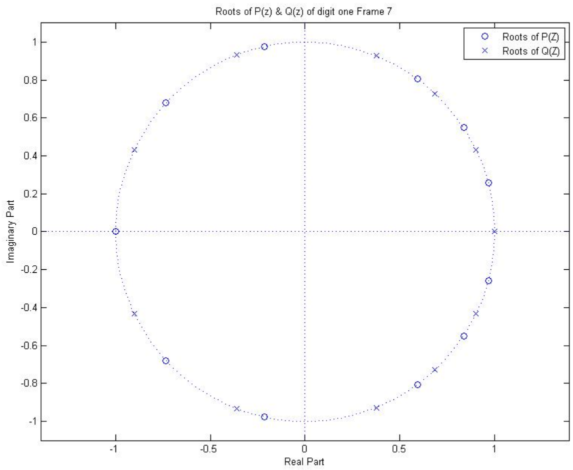

Figure 14: Plot of roots of polynominal P(z) and Q(z) for frame 7 in Digit 1

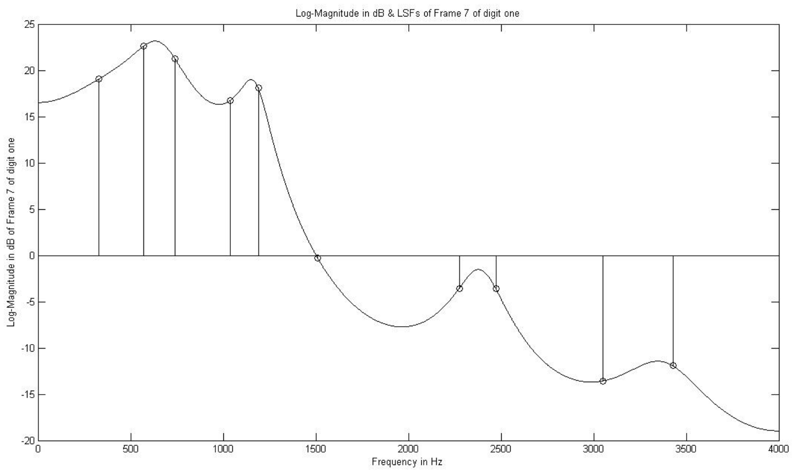

Figure 15: Plot of the Spectral Envelope of 1/A(z) with superimposed LSF's for frame 7 in Digit 1

- The LSF's (in Hz) are:
  [329.9, 570.3, 738.6, 1039.3, 1189.9, 1509.5, 2274.8, 2472.2, 3050.4, 3429.5]

3 Observations

3.1 Relation between P(z) and Q(z)

- The roots always occurs in conjugate pairs because coefficients of P and Q are real.
- The roots of polynomial P(z) and Q(z) lies on the unit circle in the complex plane.
- The roots of P and Q interlace the periphery of the unit circle.
- Because LSP's are more robust to quantization noise, stability of the vocal tract filter is maintained.
- The closer two roots are, the more resonant the filter is at the corresponding frequency.
- The LSFs collects near the spectral peaks/formant frequencies of vocal tract filter (1/A(z)).

3.2 Calculation of LPC-10 parameters

- LPC Coefficients for 10th order predictor were computed on selected frames using matlab inbuilt command (lpc).
- Reflection coefficients were computed using poly2rc function.
- LSF was computed using poly2lsf function.

4 Program Code

```
%% Assignment - 5
% Question: 06 Computer Assignment
% Prepared By: Parveen Bajaj, EET182574

%*****************************************************************************************
clc, clear all, close('all');

%% Load the voice recorded file
speech_file = input('Enter speech file:', 's');
%[speech,fs] = audioread(speech_file);
[speech,fs,bits] = wavread(speech_file);
%% Display the signal
figure(1),plot([1:size(speech)]/fs,speech);
title(speech_file);
ylabel('Amplitude');
xlabel('Time(sec)');

%% Resample the signal from 16 khz to 8 khz
resampled_speech = speech(1:2:end);
new_fs = fs/2;
figure,plot([1:size(resampled_speech)]/new_fs,resampled_speech);
title('Resampled Speech Signal at 8 KHz');
ylabel('Amplitude');
xlabel('Time(sec)');

%% Extract digit 1 segment from speech signal
digit_1 = resampled_speech(29280:32670);
figure, plot(digit_1);title('Digit\_1 chosen for experiment'),xlabel('Samples'),ylab

%% Get Frames of 22.5 msec = 22.5 * 8 = 180 samples
FrameDurInTime = 22.5;
NoOfSamples = FrameDurInTime * 10^-3 * new_fs; % 10^-3 is for msec
Frames = buffer(digit_1,NoOfSamples);

%% Energy in each frame
Energy =sum(Frames.^2,1);
figure,stem(Energy);title('Energy in each frame of digit-1'),xlabel('Frame Number'),

%% Select 04 frames of 22.5 msec (22.5 * 8 = 180 samples) from Voiced segments of re
% Will chose frame number 4,5,6,7
SelectedFrames = Frames(:,[4,5,6,7]);

figure, plot(SelectedFrames(:,1)), title('Frame-4 of digit-1');xlabel('Samples'),yla
```

```
figure, plot(SelectedFrames(:,2)), title('Frame-5 of digit-1');xlabel('Samples'),yla
figure, plot(SelectedFrames(:,3)), title('Frame-6 of digit-1');xlabel('Samples'),yla
figure, plot(SelectedFrames(:,4)), title('Frame-7 of digit-1');xlabel('Samples'),yla

%% Mimicking LPC-10: Obtain the parameter values that are computed (and transmitted
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For Frame-4
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
temp = digit_1(4*180+1:8*180); % For frames-4,5,6,7

% Delays for Avg Mag Diff Fun
t1=linspace(20,39,20);
t2=linspace(40,78,20);
t3=linspace(80,156,20);
t=[t1 t2 t3];

AvgMagDiffFun = zeros(1,156);
Pitch = [];
for i=1:length(t)
    prt=0;
    for n=1:130
        prt= prt+abs(temp(n)-temp(n+t(i)));
    end
    AvgMagDiffFun(t(i))=prt;
    Pitch = [Pitch prt];
end

%% Computation of Linear Predictive Coding , LPC-10, coefficients and Reflection
%% coefficients, Gain for each Frame

P = 10;
RMS = sqrt(mean(sum(SelectedFrames(:,1).^2)));
[Coefficents,Var] = lpc(SelectedFrames(:,1),P); % LPC Coefficients
k = poly2rc(Coefficents); % Reflection Coefficients

% Get P(z) and Q(z) using A(z) which is having coefficents in variable
% Coefficents.

LSF = poly2lsf(Coefficents);
LSF_hz= (LSF*new_fs)/(2*pi);% LSF's in Hz

bz = Coefficents(end:-1:1);

az = [Coefficents 0];
bz = [0 bz];
```

```
Pz = az+bz;
Qz = az-bz;

figure,zplane(Pz,Qz);legend('Roots of P(Z)','Roots of Q(Z)');title('Roots of P(z) &

b = [1];
[h,freq]=freqz(b,az,4000,new_fs);
h = 20*log10(h);
value_lsf=h(ceil(LSF_hz));

figure;plot(freq,h,'k');hold on;
stem(LSF_hz,value_lsf,'k');
xlabel('Frequency in Hz');
ylabel('Log-Magnitude in dB of Frame 4 of digit one');
title('Log-Magnitude in dB & LSFs of Frame 4 of digit one');

%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For Frame-5
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
temp = digit_1(5*180+1:9*180); % For frames-4,5,6,7
% Delays for Avg Mag Diff Fun
t1=linspace(20,39,20);
t2=linspace(40,78,20);
t3=linspace(80,156,20);
t=[t1 t2 t3];

AvgMagDiffFun = zeros(1,156);
Pitch = [];
for i=1:length(t)
    prt=0;
    for n=1:130
        prt= prt+abs(temp(n)-temp(n+t(i)));
    end
    AvgMagDiffFun(t(i))=prt;
    Pitch = [Pitch prt];
end

%% Computation of Linear Predictive Coding , LPC-10, coefficients and Reflection
%% coefficients, Gain for each Frame

% For Frame-5
P=10;
RMS=sqrt(mean(sum(SelectedFrames(:,2).^2)));
[Coefficents,Var]=lpc(SelectedFrames(:,2),P); % LPC Coefficients
```

```
k = poly2rc(Coefficents); % Reflection Coefficients

% Get P(z) and Q(z) using A(z) which is having coefficents in variable
% Coefficents.

LSF = poly2lsf(Coefficents);
LSF_hz= (LSF*new_fs)/(2*pi);% LSF's in Hz

bz = Coefficents(end:-1:1);

az = [Coefficents 0];
bz = [0 bz];
Pz = az+bz;
Qz = az-bz;

figure,zplane(Pz,Qz);legend('Roots of P(Z)','Roots of Q(Z)');title('Roots of P(z) &

b = [1];
[h,freq]=freqz(b,az,4000,new_fs);
h = 20*log10(h);
value_lsf=h(ceil(LSF_hz));

figure;plot(freq,h,'k');hold on;
stem(LSF_hz,value_lsf,'k');
xlabel('Frequency in Hz');
ylabel('Log-Magnitude in dB of Frame 5 of digit one');
title('Log-Magnitude in dB & LSFs of Frame 5 of digit one');

%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For Frame-6
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
temp = digit_1(6*180+1:10*180); % For frames-4,5,6,7
% Delays for Avg Mag Diff Fun
t1=linspace(20,39,20);
t2=linspace(40,78,20);
t3=linspace(80,156,20);
t=[t1 t2 t3];

AvgMagDiffFun = zeros(1,156);
Pitch = [];
for i=1:length(t)
    prt=0;
    for n=1:130
        prt= prt+abs(temp(n)-temp(n+t(i)));
```

```
    end
    AvgMagDiffFun(t(i))=prt;
    Pitch = [Pitch prt];
end

%% Computation of Linear Predictive Coding , LPC-10, coefficients and Reflection
%% coefficients, Gain for each Frame

% For Frame-6
P=10;
RMS=sqrt(mean(sum(SelectedFrames(:,3).^2)));
[Coefficents,Var]=lpc(SelectedFrames(:,3),P); % LPC Coefficients
k = poly2rc(Coefficents); % Reflection Coefficients

% Get P(z) and Q(z) using A(z) which is having coefficents in variable
% Coefficents.

LSF = poly2lsf(Coefficents);
LSF_hz= (LSF*new_fs)/(2*pi);% LSF's in Hz

bz = Coefficents(end:-1:1);

az = [Coefficents 0];
bz = [0 bz];
Pz = az+bz;
Qz = az-bz;

figure,zplane(Pz,Qz);legend('Roots of P(Z)','Roots of Q(Z)');title('Roots of P(z) &

b = [1];
[h,freq]=freqz(b,az,4000,new_fs);
h = 20*log10(h);
value_lsf=h(ceil(LSF_hz));

figure;plot(freq,h,'k');hold on;
stem(LSF_hz,value_lsf,'k');
xlabel('Frequency in Hz');
ylabel('Log-Magnitude in dB of Frame 6 of digit one');
title('Log-Magnitude in dB & LSFs of Frame 6 of digit one');

%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For Frame-7
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
temp = digit_1(7*180+1:11*180); % For frames-4,5,6,7
```

```
% Delays for Avg Mag Diff Fun
t1=linspace(20,39,20);
t2=linspace(40,78,20);
t3=linspace(80,156,20);
t=[t1 t2 t3];

AvgMagDiffFun = zeros(1,156);
Pitch = [];
for i=1:length(t)
    prt=0;
    for n=1:130
        prt= prt+abs(temp(n)-temp(n+t(i)));
    end
    AvgMagDiffFun(t(i))=prt;
    Pitch = [Pitch prt];
end

%% Computation of Linear Predictive Coding , LPC-10, coefficients and Reflection
%% coefficients, Gain for each Frame

% For Frame-7
P=10;
RMS=sqrt(mean(sum(SelectedFrames(:,4).^2)));
[Coefficents,Var]=lpc(SelectedFrames(:,4),P); % LPC Coefficients
k = poly2rc(Coefficents); % Reflection Coefficients

% Get P(z) and Q(z) using A(z) which is having coefficents in variable
% Coefficents.

LSF = poly2lsf(Coefficents);
LSF_hz= (LSF*new_fs)/(2*pi);% LSF’s in Hz

bz = Coefficents(end:-1:1);

az = [Coefficents 0];
bz = [0 bz];
Pz = az+bz;
Qz = az-bz;

figure,zplane(Pz,Qz);legend(’Roots of P(Z)’,’Roots of Q(Z)’);title(’Roots of P(z) &

b = [1];
[h,freq]=freqz(b,az,4000,new_fs);
h = 20*log10(h);
value_lsf=h(ceil(LSF_hz));
```

```
figure;plot(freq,h,'k');hold on;
stem(LSF_hz,value_lsf,'k');
xlabel('Frequency in Hz');
ylabel('Log-Magnitude in dB of Frame 7 of digit one');
title('Log-Magnitude in dB & LSFs of Frame 7 of digit one');
```