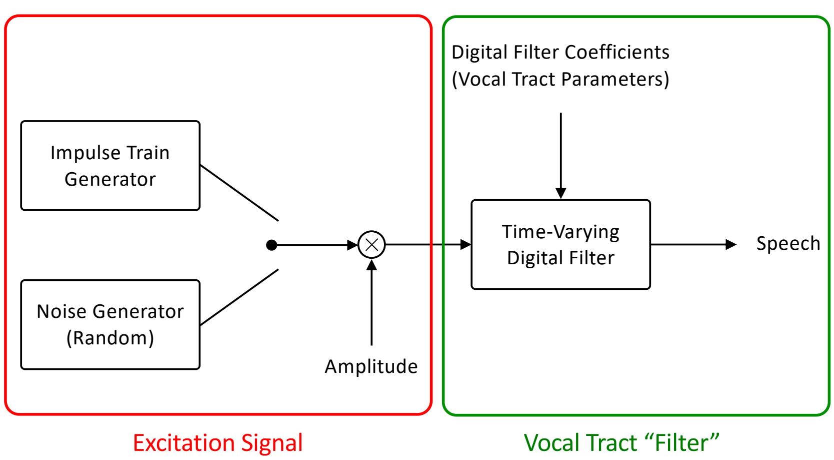
SIO 209: Signal Processing for Ocean Sciences
Class 17

Florian Meyer

Scripps Institution of Oceanography
Electrical and Computer Engineering Department
University of California San Diego

0

Recall Speech Production Model



A vocoder extracts parameters of excitation signal and the vocal tract filter; transmitting these parameters (instead of the sampled speech signal) can reduce data rate significantly

1

1

Recall Speech Transmission

- Direct digitization of speech results in a data rate of 64 kB/s since typically 8 Bits/sample are used and $f_s = 8\mathrm{kHz}$ (speech is bandlimited to $4-5\mathrm{kHz}$)

$$x[n] \longrightarrow \boxed{\hat{A}(z)} \longrightarrow \hat{w}[n]$$

speech — LPC Analysis — error

- We take segments ("frames") that are 25 ms long, i.e., there are 40 segments/sec
- For each segment we need to obtain
  - the $\hat{A}$ vector (10–14 coefficients)
  - voiced/unvoiced decision
  - pitch period
  - amplitude

$\longrightarrow$ Results in a data rate of $1.0-2.4\mathrm{kb/s}$

2

Recall High Resolution Spectral Analysis

Process Generation

$$w[n] \longrightarrow \boxed{\text{All-Poles Filter}} \longrightarrow x[n]$$

white noise — observed data

$$H(z) = \frac{1}{A(z)}$$

Auto-Regressive (AR) Model
(IIR filter with no zeros)

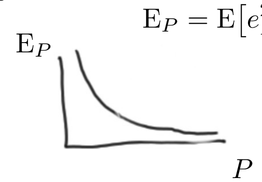
Analysis **Goal:** Estimate $P_{xx}(\omega) = |H(\omega)|^2$

$$\longrightarrow \boxed{\text{All-Zeros Filter}} \longrightarrow e_P(n) = \hat{w}(n)$$

$\hat{A}(z)$

prediction error filter
"lpc" and "levinson"

$$\mathrm{E}_P = \mathrm{E}[e_P^2[n]]$$



Terms in $A(z)$ have the form

$$\frac{(z-z_k)(z-z_k^*)}{z^2}$$

$$k = 1 \ldots K/2$$

The polynomial coefficients of $A(z)$ can be obtained from "poly"

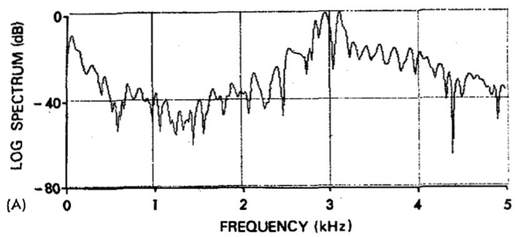
The all-pole filters can be implemented using "filter"

Note that there is also a ARMA (AR-moving-average) model which is based on a general IIR filter (poles & zeros)

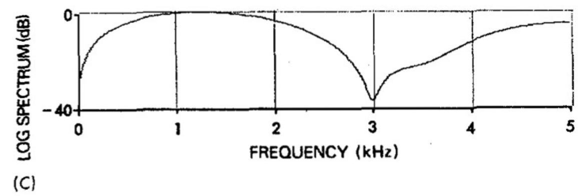
3

Unvoiced Sound

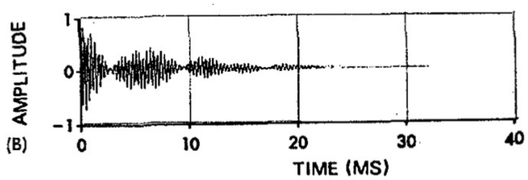
Power Spectral Density $|X(k)|^2$



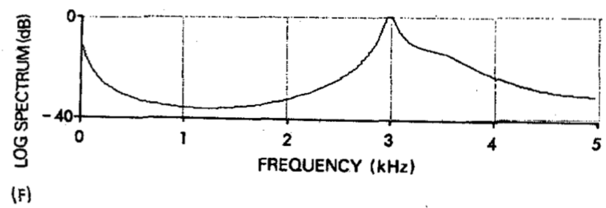
All-Zeros Filter $|A(k)|^2$



Time Domain $x[n]$



All-Poles Filter

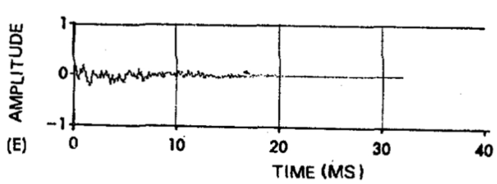


J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

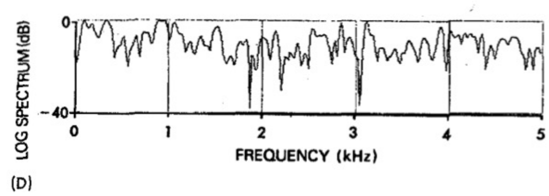
4

Unvoiced Sound

Error Signal $e_P[n]$ in Time Domain



Error Signal $e_P[n]$ in Frequency Domain

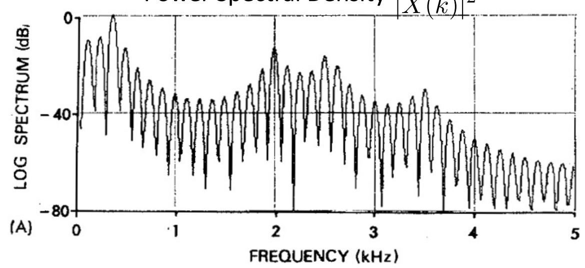


J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

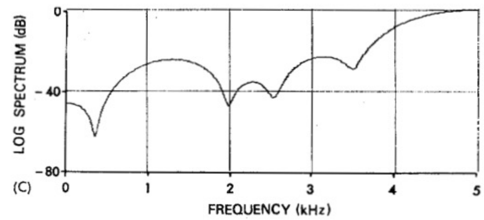
5

Voiced Sound

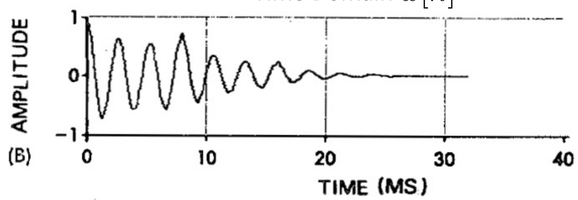
Power Spectral Density $|X(k)|^2$



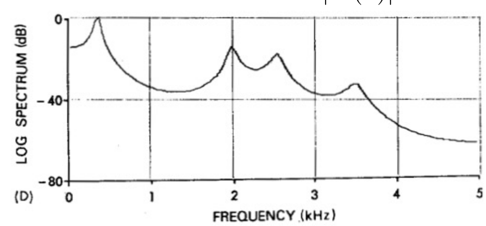
All-Zeros Filter $|A(k)|^2$



Time Domain $x[n]$



All-Poles Filter $|H(k)|^2$

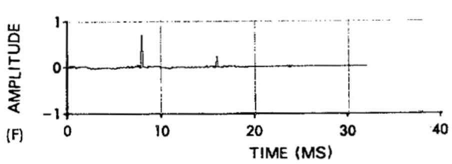


J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

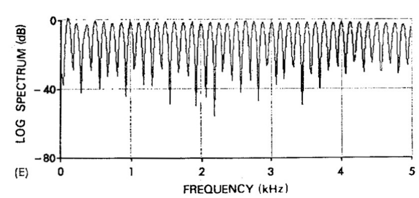
6

Voiced Sound

Error Signal $e_P[n]$ in Time Domain



Error Signal $e_P[n]$ in Frequency Domain



- The autocorrelation function of the error signal can be used to make a voiced/unvoiced decision by looking for peaks between 8 ms and 12 ms
- If a peak is detected, the sound can be declared as voiced, and the delay related to the peak is the pitch period

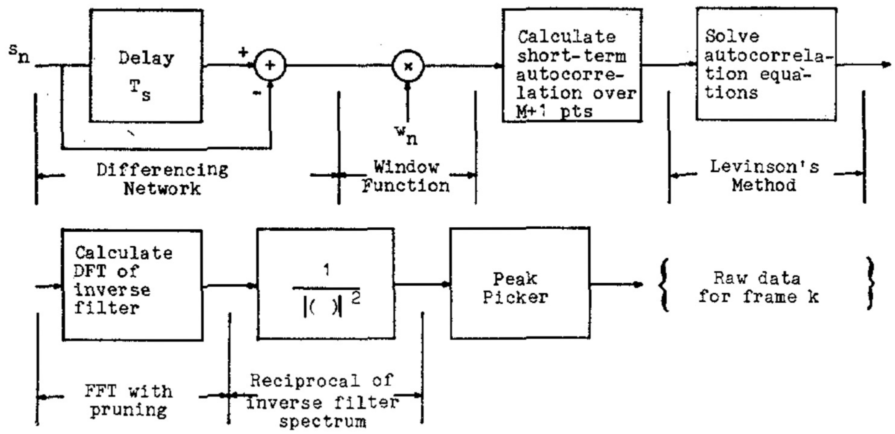
J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

7

Block Diagram of Inverse Filter Algorithm

High-Pass Filter (optional)

"lpc" in Matlab



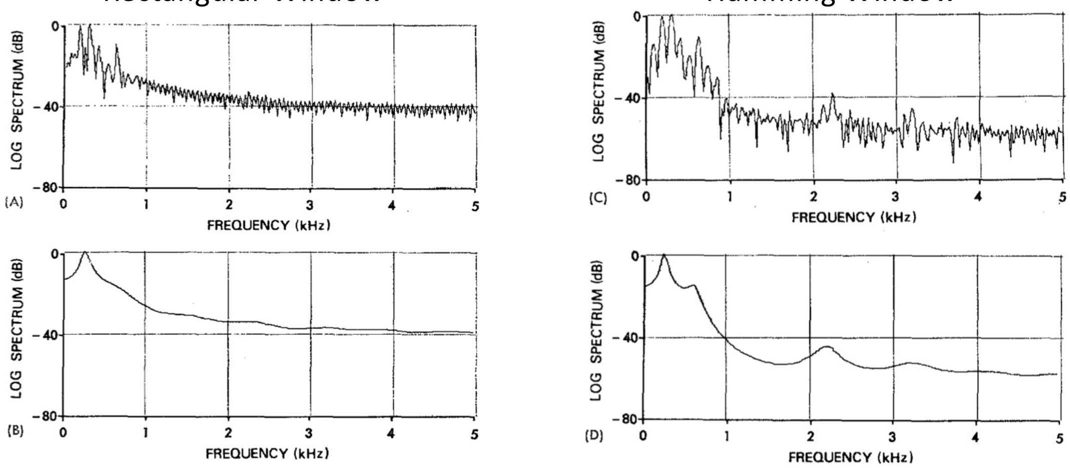
J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

8

Effects of Window Function

Rectangular Window

Hamming Window

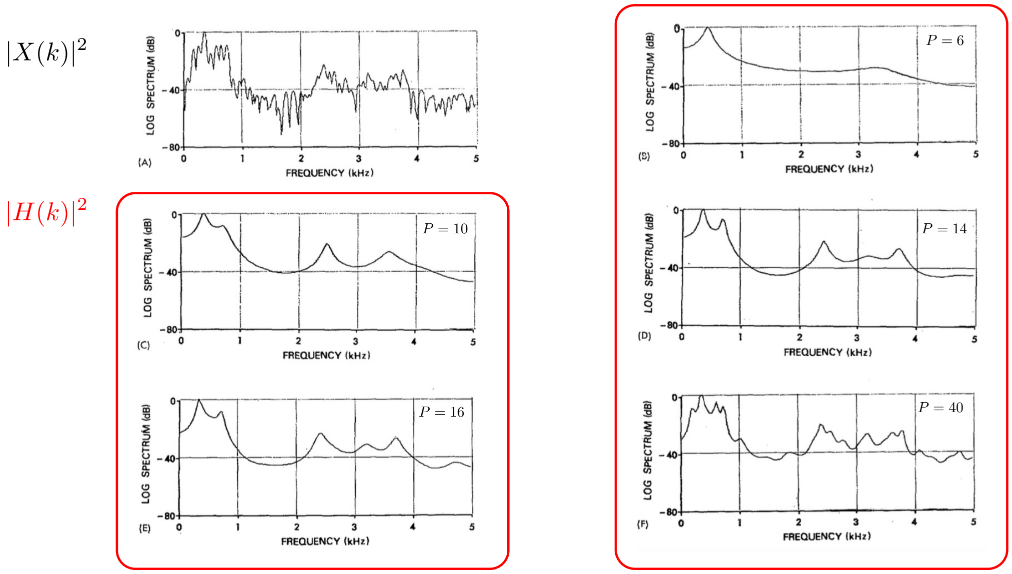


- Using a window function can help to reveal the formant structure at higher frequencies

J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

9

Effects of Filter Order



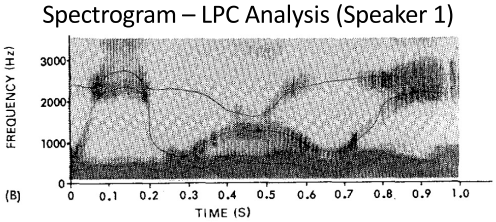
J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

10

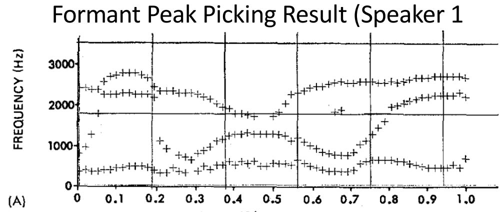
Speech Analysis

- Analysis of phrase "we were away"

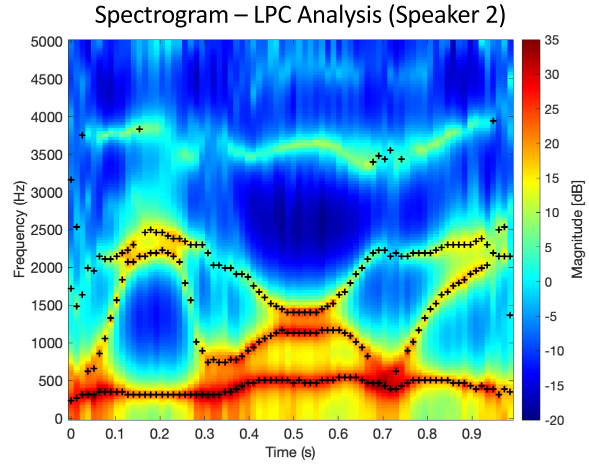
Spectrogram – LPC Analysis (Speaker 1)



Formant Peak Picking Result (Speaker 1



Spectrogram – LPC Analysis (Speaker 2)



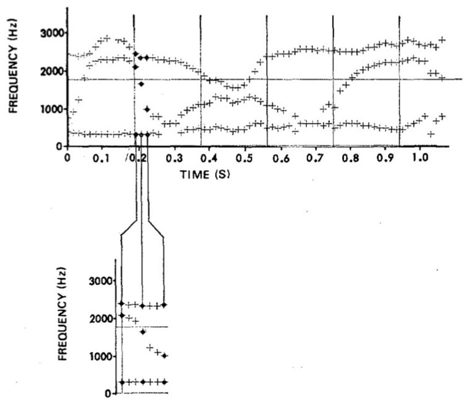
J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

11

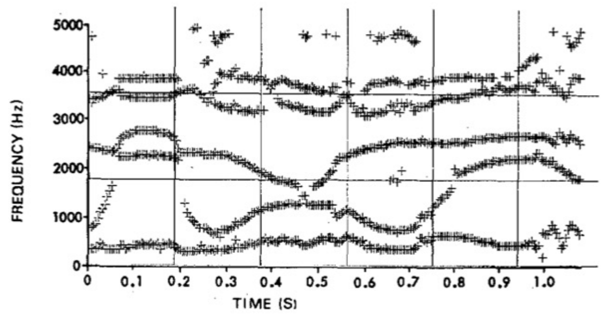
Speech Analysis

- Analysis of phrase "we were away"

Formant Peak Picking Result
(Zoom in with 80% Overlap)



Formant Peak Picking Result
(80% Overlap)

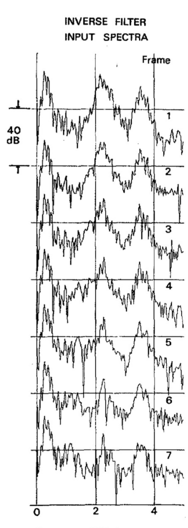


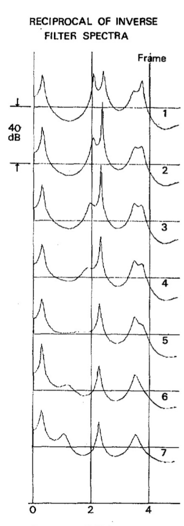
J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

12

Speech Analysis

- High-resolution spectral analysis can reveal small changes in formant structure





J. Markel, "Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation." IEEE Trans. Audio and Electroacoustics, 1972

13