Lecture 6: Linear Prediction (LPC)

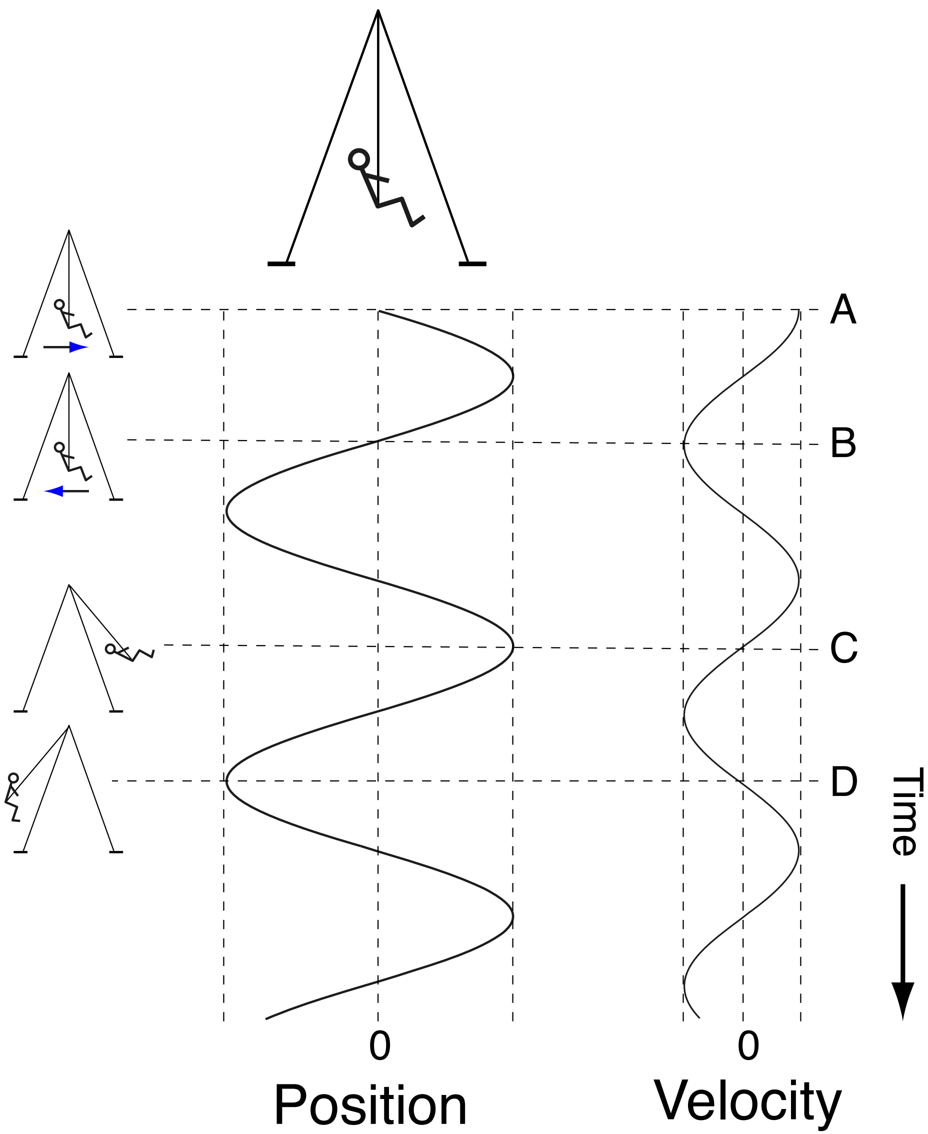
1. Resonance & The Source-Filter Model
2. Linear Prediction (LPC)
3. LP Representations
4. LP Synthesis & Modification

Dan Ellis
Dept. Electrical Engineering, Columbia University
dpwe@ee.columbia.edu http://www.ee.columbia.edu/~dpwe/e4896/

1. Resonance

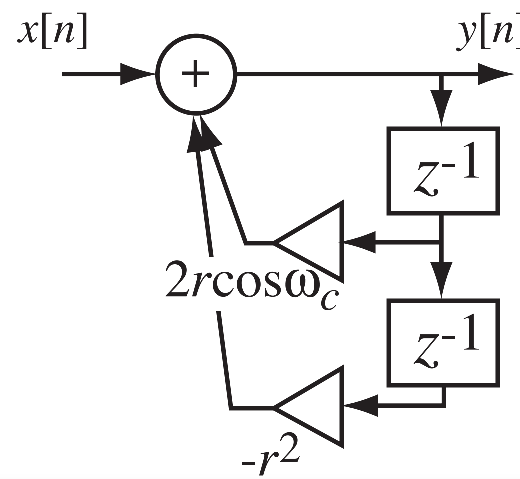
- Resonance is ubiquitous in physical systems
  - e.g. plucked/struck string, drum head
  - room "coloration"
  - vocal tract

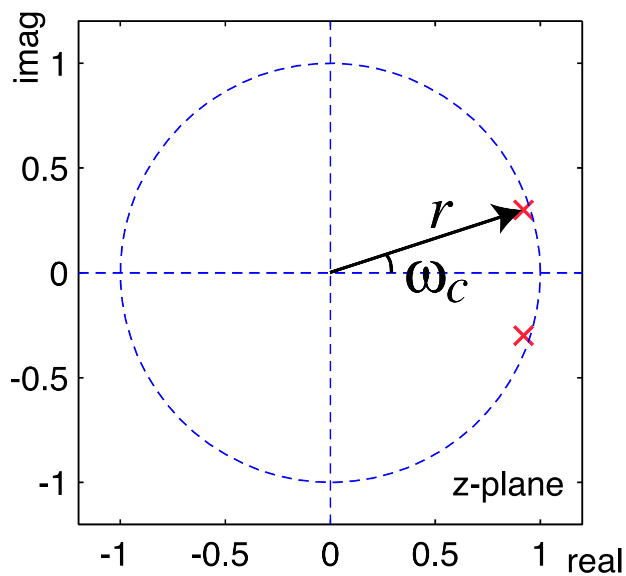
- Resonances ↔ Poles
  - easily implemented in LTI filters

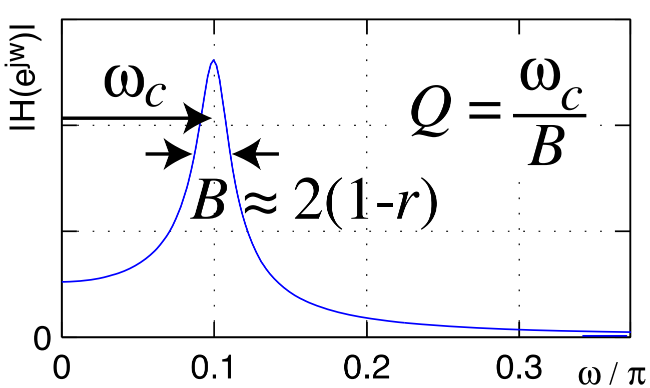


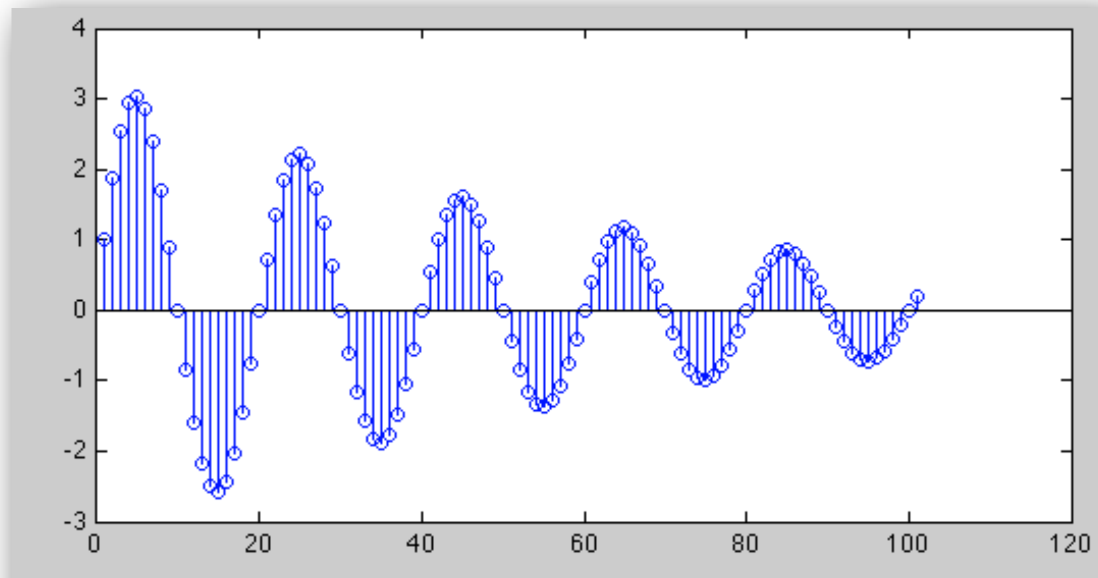
Singe Pole-Pair Resonance

- Simple resonance = Second-order IIR Band Pass Filter







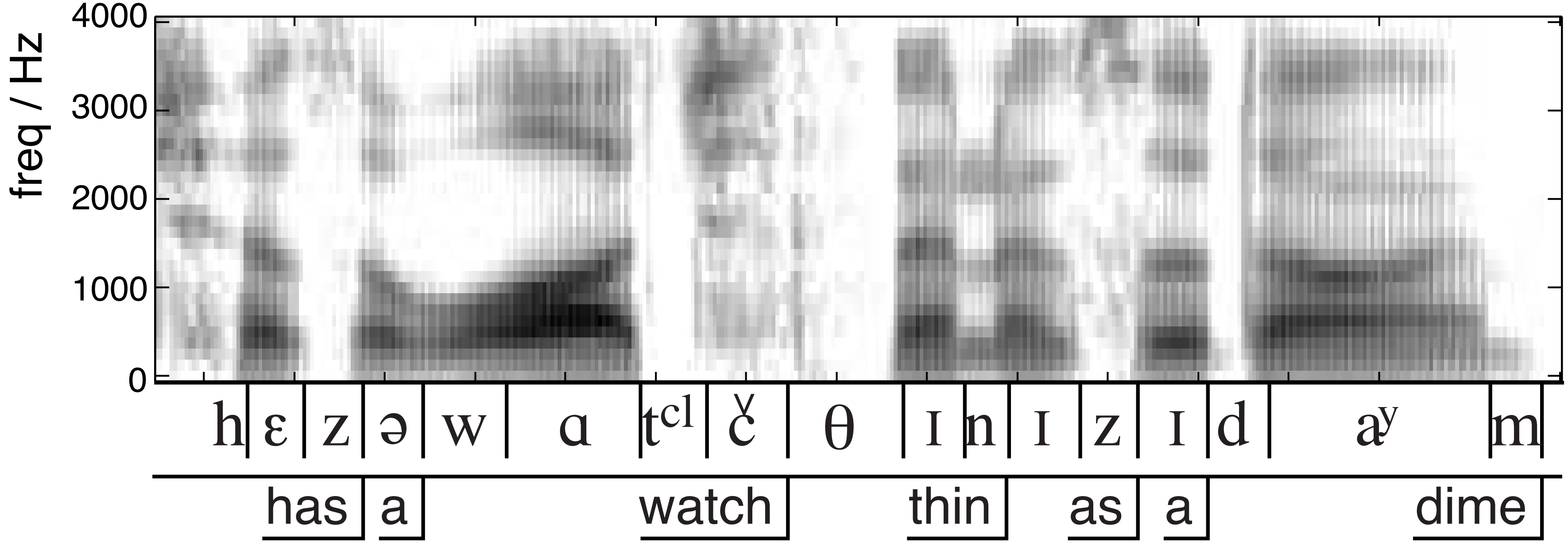



```
>>
>> Q = 10;
>> omega = pi*0.1;
>> r = 1 - omega/(2*Q);
>> a = [1 -2*r*cos(omega) r*r];
>> y = filter(1,a,[1,zeros(1,100)]);
>> stem(y)
>>
```

impulse_bpf.pd

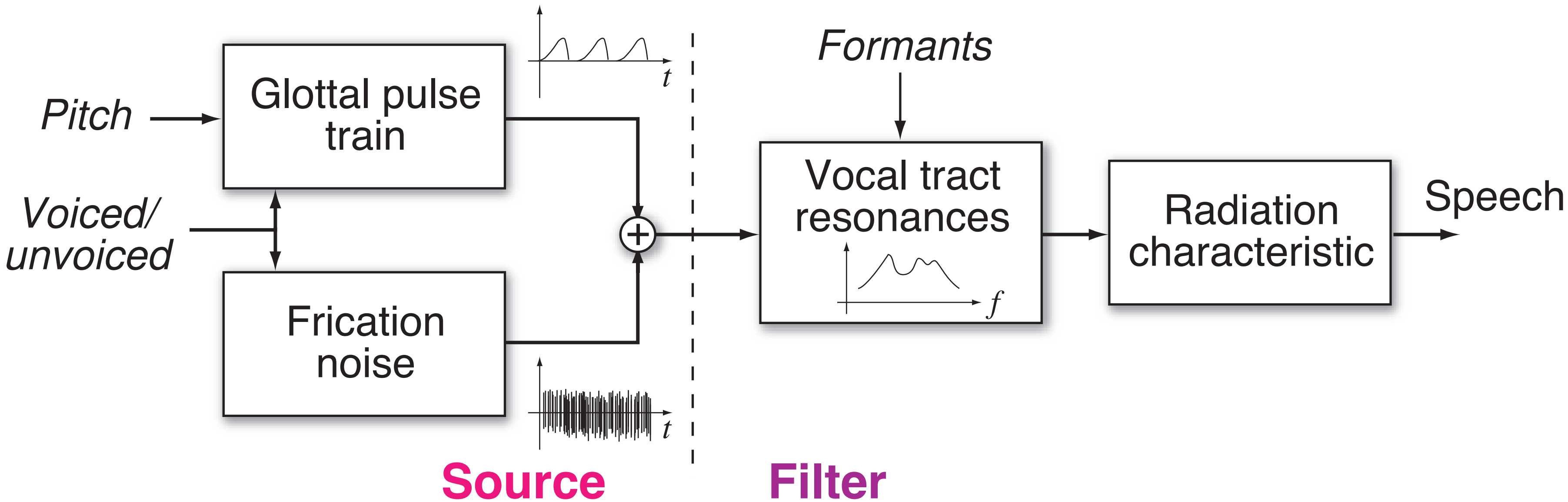
Resonances in Speech

- Vocal Tract (throat + tongue + lips) acts as variable resonator
  - resonances = "formants"



Source-Filter Model

- Separation of:
  - Source: fine structure in time/frequency
  - Filter: subsequent shaping by physical resonances



- Advantages
  - Good match to real signals
  - Salient pieces

2. Linear Prediction (LPC)

- LPC = Linear Predictive Coding
  - remove redundancy in signal
  - try to predict next point as linear combination of previous values

$$s[n] = \sum_{k=1}^{p} a_k s[n-k] + e[n]$$

  - $\{a_k\}$ are p^{th} order linear predictor coefficients
  - $e[n]$ is residual "innovation" a/k/a prediction error
- Transfer function

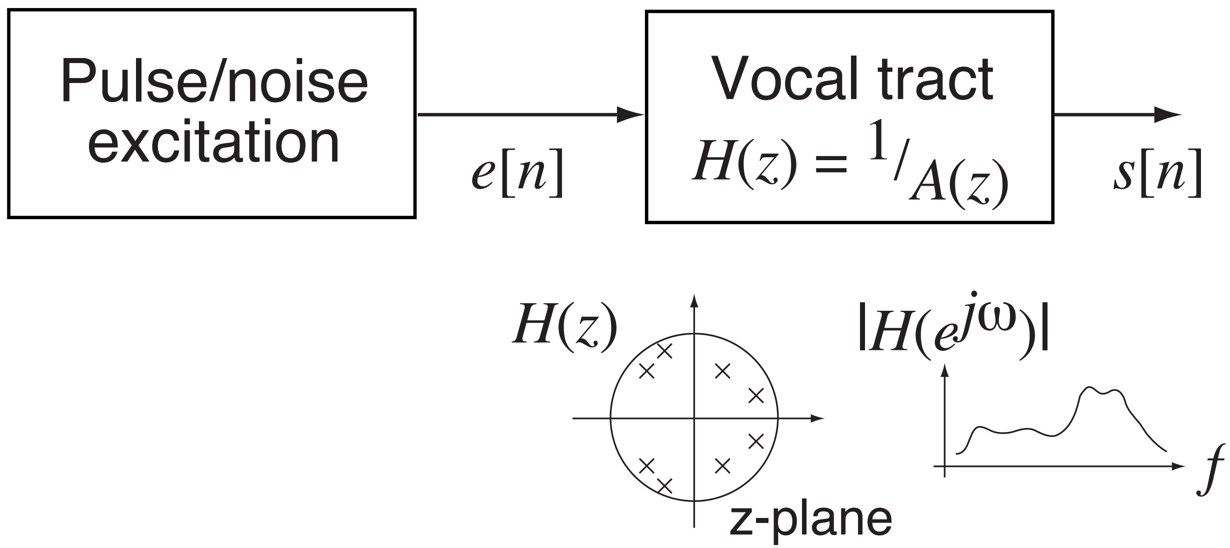
$$\frac{S(z)}{E(z)} = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}} = \frac{1}{A(z)}$$

  - all-pole "autoregressive" (AR) modeling

Voice Modeling & LPC

- Direct expression of source-filter model

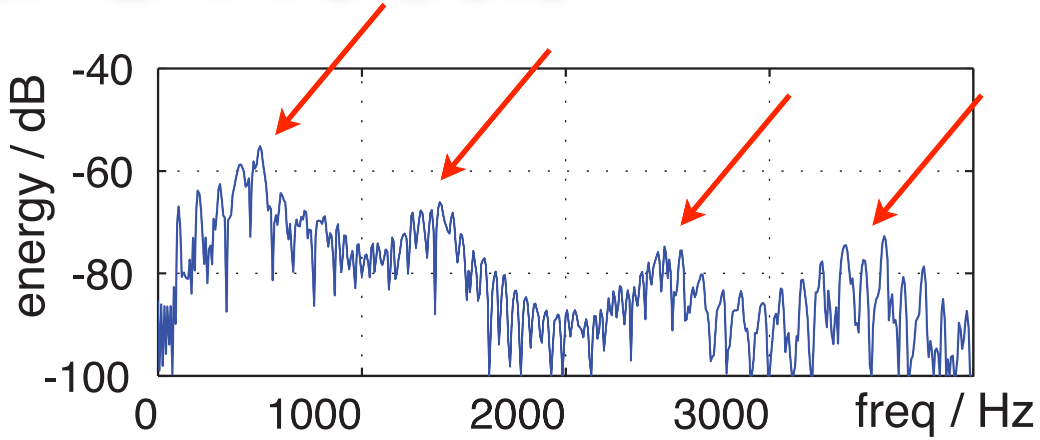
$$s[n] = \sum_{k=1}^{p} a_k s[n-k] + e[n]$$



  - acoustic tube model of vocal tract is all-pole
  - vocal tract resonances change slowly ~ 10-20ms
  - but: nasals

Estimating LPC Models

- You can "see" resonances in a spectral slice:

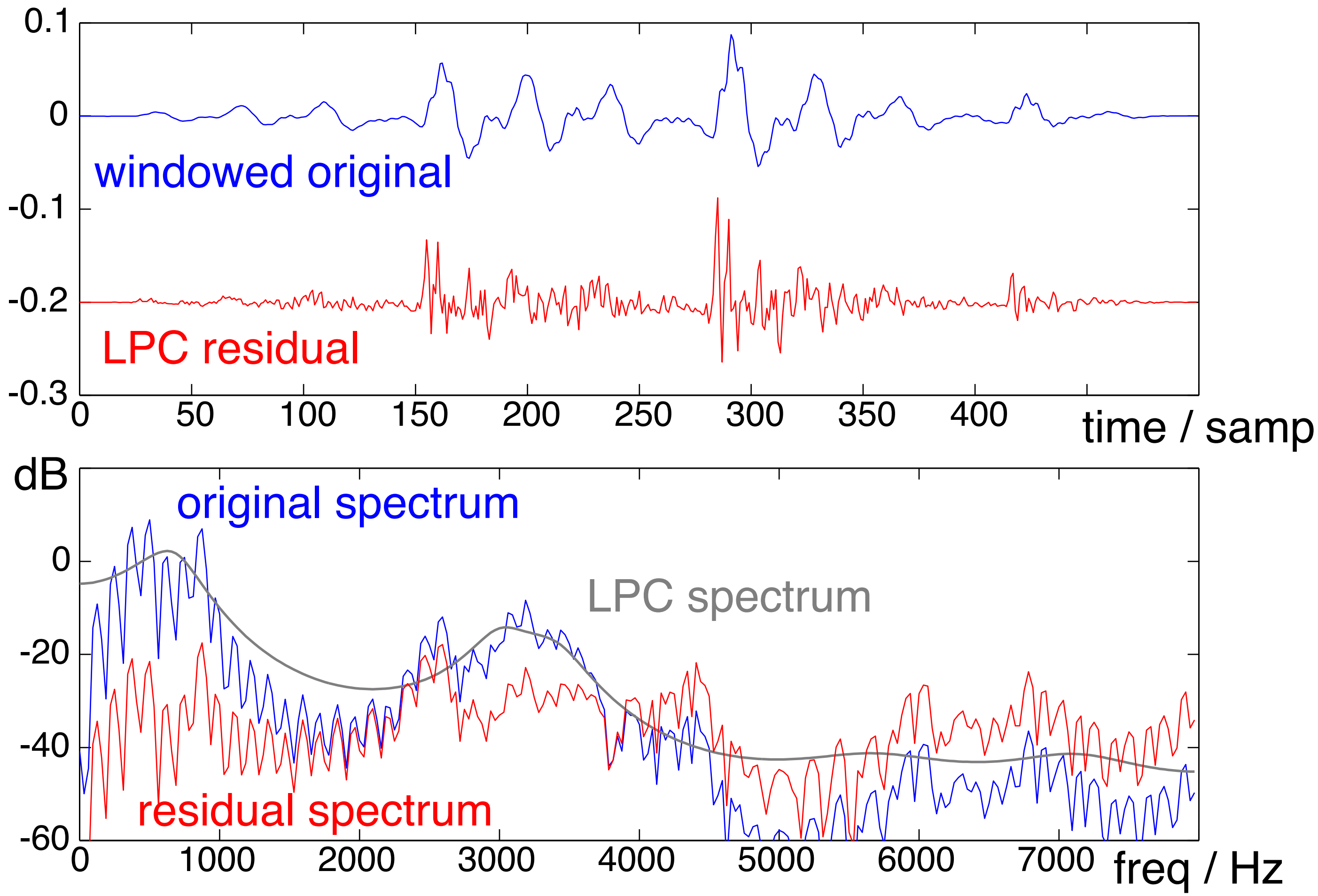


- We can find LPC coefficients $\{a_k\}$ to minimize energy of residual $e[n]$:

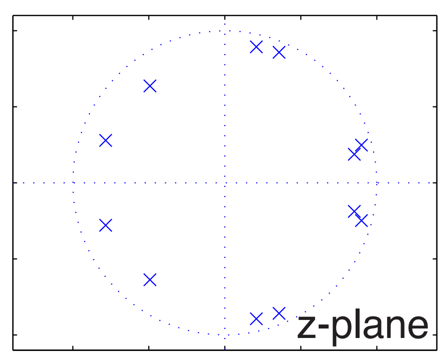
$$\sum_n e^2[n] = \sum_n \left(s[n] - \sum_{k=1}^{p} a_k s[n-k] \right)^2$$

  - differentiate w.r.t. a_k & solve
  - end up with p linear equations involving autocorrelations $r_{ss}(|j-k|) = \sum_n s[n-j]s[n-k]$

LPC Illustration



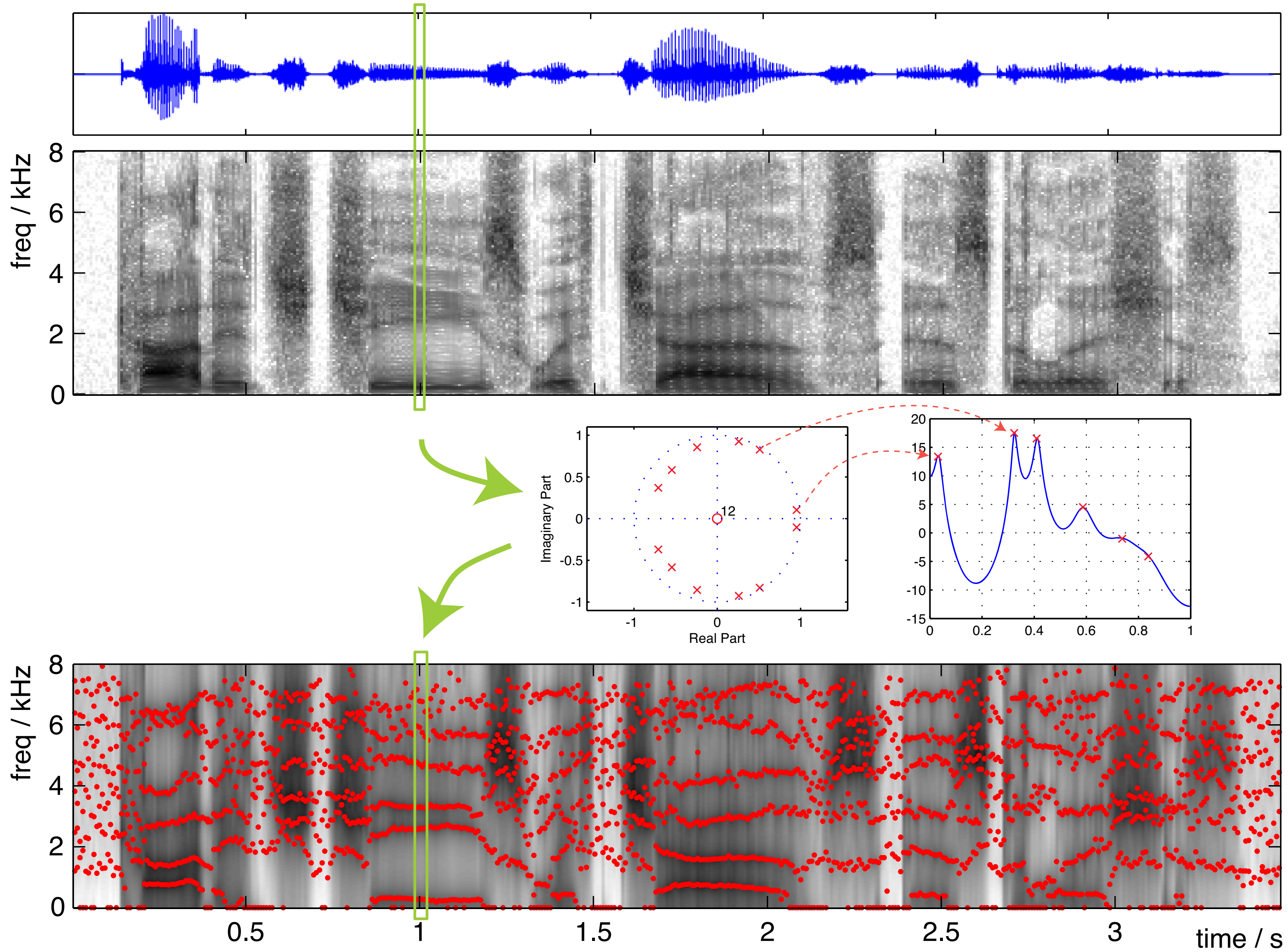
- Actual poles:



E4896_L06_lpc_diary

Short-Time LP Analysis

- Solve LPC for each ~20 ms frame



3. LP Representations

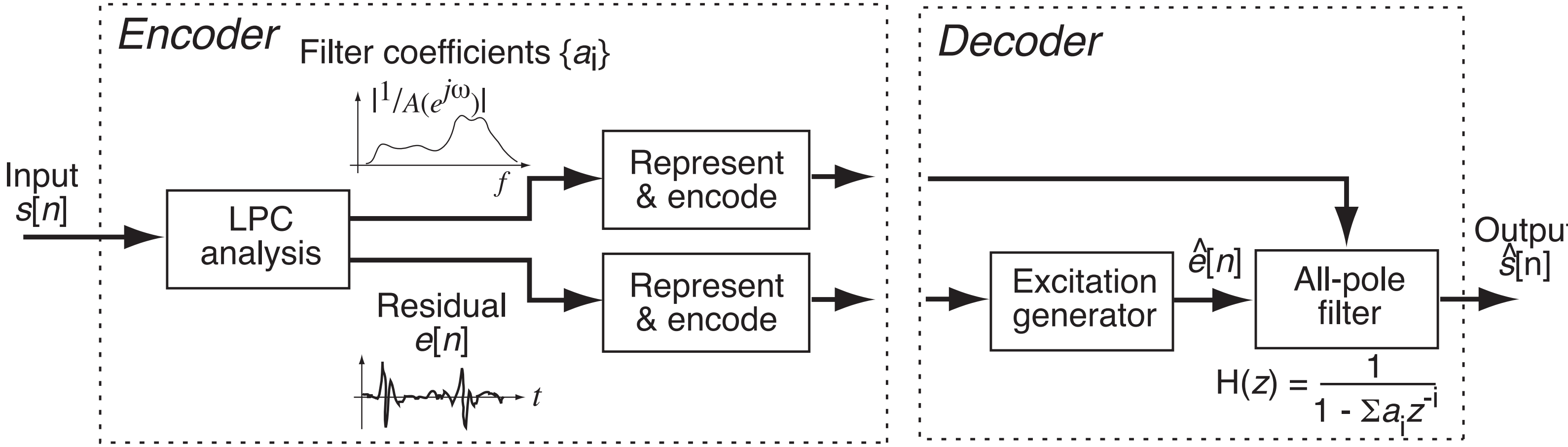
- Can interpret LPC filter fit many ways:
- Picking out resonances
  - if signal was source + resonances, should find them
- Low-order spectral approximation
  - minimizing $e^2[n]$ also minimizes $|E(e^{j\omega})|^2$
  - different from e.g. Fourier approximation...
- Finding & removing smooth spectrum
  - $\frac{1}{A(z)}$ is smooth approximation of $S(z)$
  - $\frac{S(z)}{E(z)} = \frac{1}{A(z)} \Rightarrow E(z) = S(z)A(z)$ is "unsmoothed" $S(z)$
- Signal whitening
  - removing linear dependence makes residual like white noise (iid, flat spectrum)

Alternative Forms

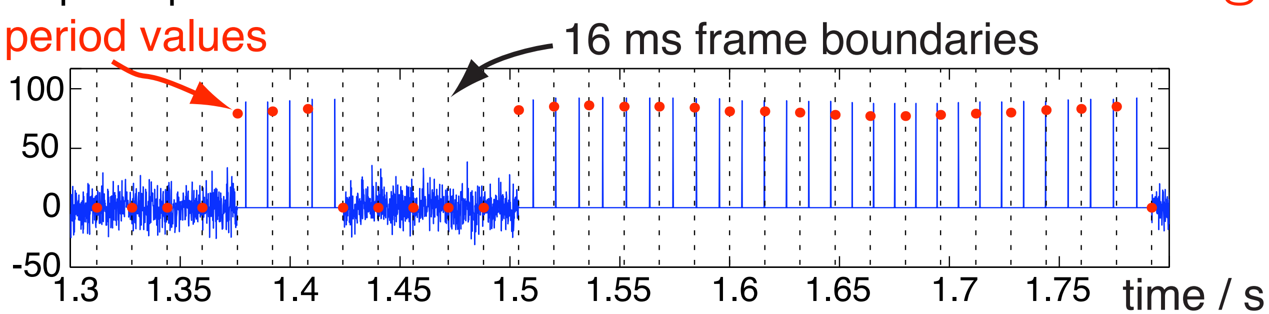
- **Many formulations for p^{th} order all-pole IIR:**
  - predictor coefficients $\{a_k\}$ / polynomial $A(z)$
  - roots $\{\lambda_i = r_i e^{j\omega_i}\}$ of $A(z)$
  - reflection coefficients (for lattice filter structure)
  - Line Spectral Frequencies (LSF)
- **Choice depends on:**
  - mathematical convenience
  - numerical stability
  - statistical properties (e.g. for coding)
  - opportunities for modification

4. LPC Synthesis

- LP analysis on ~20ms frames gives prediction filter $A(z)$ and residual $e[n]$
  - recombining them should yield perfect $s[n]$
  - coding applications further compress $e[n]$

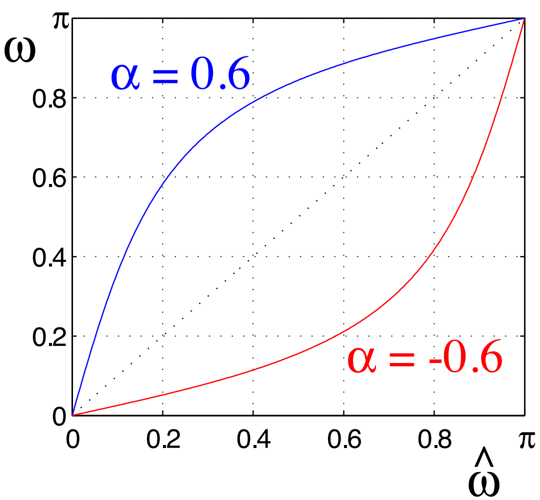


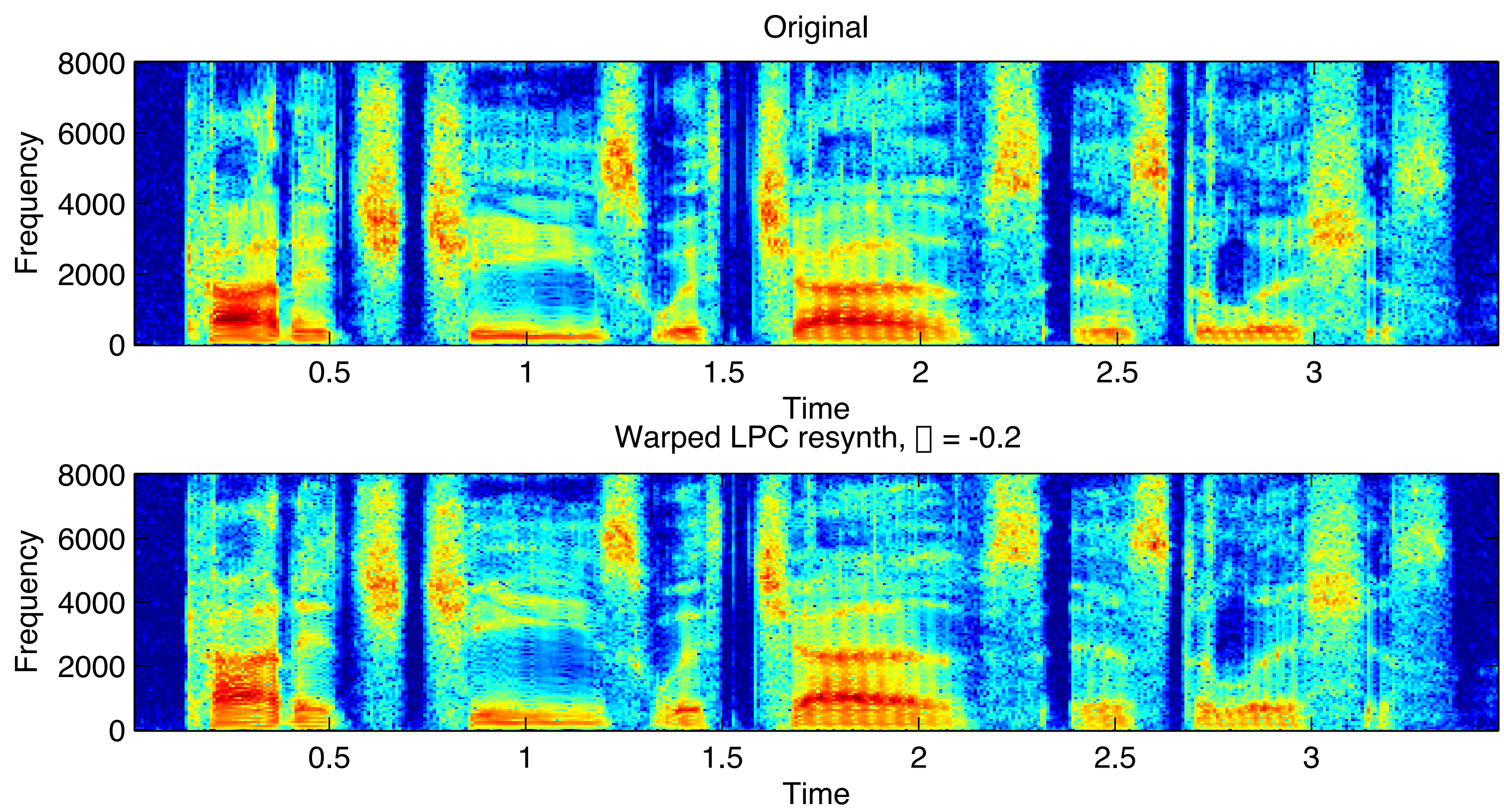
  - e.g. simple pitch tracker → "buzz-hiss" encoding



LPC Warping

- Replacing delays z^{-1} with allpass elements $\dfrac{z+\alpha}{\alpha z+1}$ warps frequencies but not magnitudes

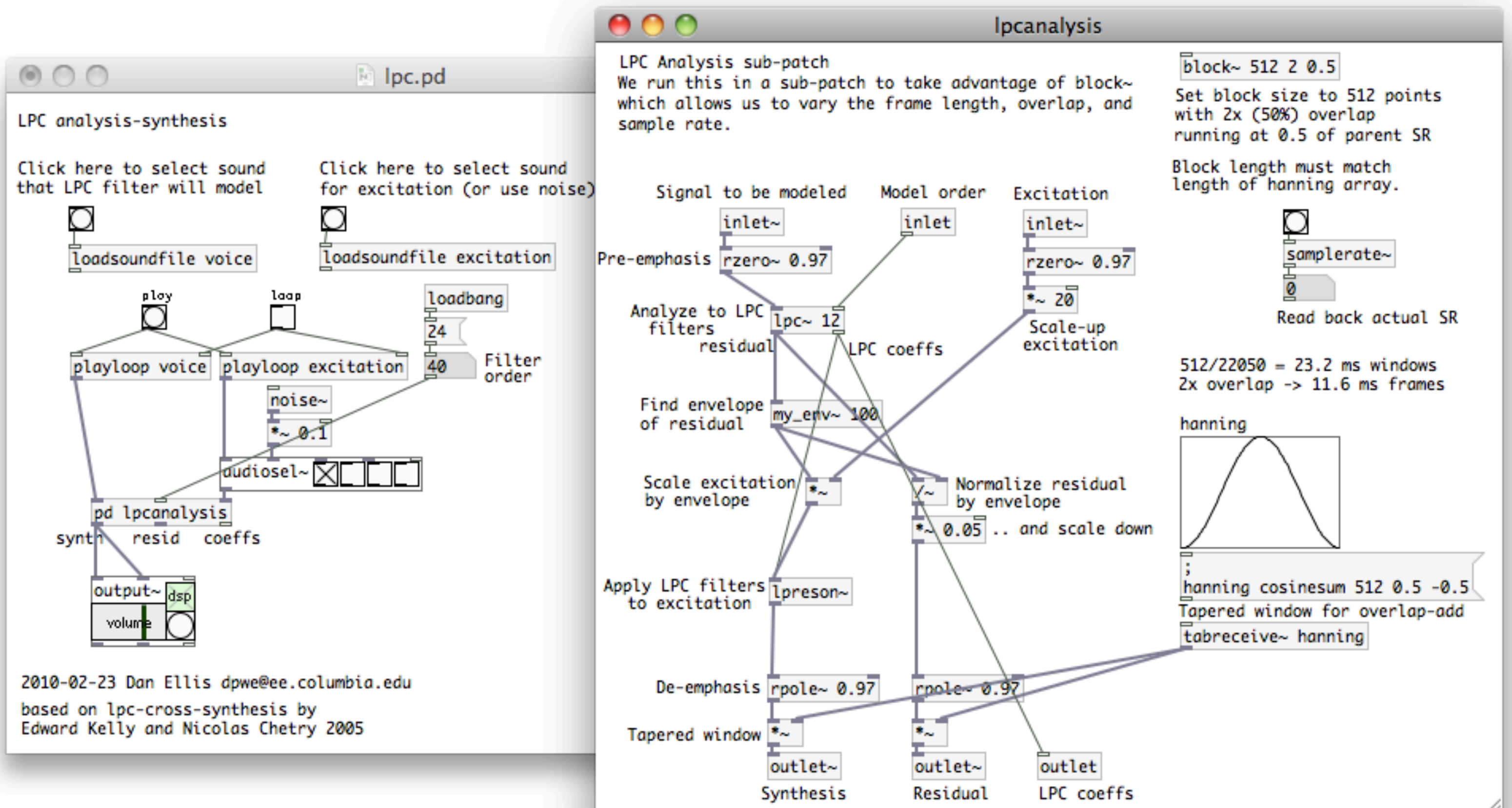




  - http://www.ee.columbia.edu/~dpwe/resources/matlab/polewarp/

Cross-Synthesis

- Mix residual (source) of one signal with resonances (filter) of another
  - or: just use white noise as excitation
  - formants carry phonemes → vocoder



Summary

- Resonances (poles) color sound
- Source + Filter model
  decouples excitation and resonances
- Linear Prediction is a simple way to model and implement resonances (filter)
- Many interpretations, representations, modifications