Digital Speech Processing— Lecture 14

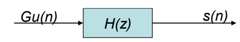
Linear Predictive Coding (LPC)-Lattice Methods, Applications

Prediction Error Signal

1. Speech Production Model

$$s(n) = \sum_{k=1}^{p} a_k s(n-k) + Gu(n)$$

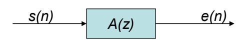
$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$



2. LPC Model:

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^{p} \alpha_k s(n-k)$$

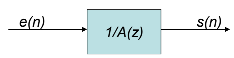
$$A(z) = \frac{E(z)}{S(z)} = 1 - \sum_{k=1}^{p} \alpha_k z^{-k}$$



3. LPC Error Model:

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

$$s(n) = e(n) + \sum_{k=1}^{p} \alpha_k s(n-k)$$



Perfect reconstruction even if a_k not equal to α_k

Lattice Formulations of LP

- both covariance and autocorrelation methods use two step solutions
  1. computation of a matrix of correlation values
  2. efficient solution of a set of linear equations
- another class of LP methods, called lattice methods, has evolved in which the two steps are combined into a recursive algorithm for determining LP parameters
- begin with Durbin algorithm--at the i^{th} stage the set of coefficients $\{\alpha_j^{(i)}, j = 1, 2, \ldots, i\}$ are coefficients of the i^{th} order optimum LP

Lattice Formulations of LP

- define the system function of the i^{th} order inverse filter (prediction error filter) as

$$A^{(i)}(z) = 1 - \sum_{k=1}^{i} \alpha_k^{(i)} z^{-k}$$

- if the input to this filter is the input segment $s_{\hat{n}}(m) = s(\hat{n} + m)w(m)$, with output $e_{\hat{n}}^{(i)}(m) = e^{(i)}(\hat{n} + m)$

$$e^{(i)}(m) = s(m) - \sum_{k=1}^{i} \alpha_k^{(i)} s(m-k)$$

- where we have dropped subscript $\hat{n}$ - the absolute location of the signal
- the z-transform gives

$$E^{(i)}(z) = A^{(i)}(z)S(z) = \left(1 - \sum_{k=1}^{i} \alpha_k^{(i)} z^{-k}\right) S(z)$$

Lattice Formulations of LP

- using the steps of the Durbin recursion

$$(\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)}, \text{ and } \alpha_i^{(i)} = k_i)$$

□we can obtain a recurrence formula for $A^{(i)}(z)$ of the form

$$A^{(i)}(z) = A^{(i-1)}(z) - k_i z^{-i} A^{(i-1)}(z^{-1})$$

- giving for the error transform the expression

$$E^{(i)}(z) = A^{(i-1)}(z)S(z) - k_i z^{-i} A^{(i-1)}(z^{-1})S(z)$$

$$e^{(i)}(m) = e^{(i-1)}(m) - k_i b^{(i-1)}(m-1)$$

Lattice Formulations of LP

- where we can interpret the first term as the z-transform of the forward prediction error for an $(i-1)^{st}$ order predictor, and the second term can be similarly interpreted based on defining a backward prediction error

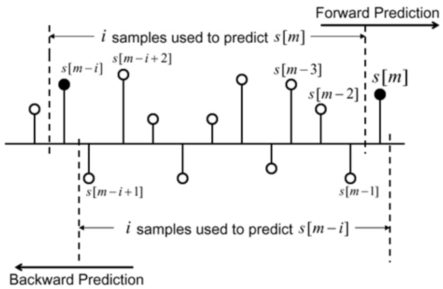
$$B^{(i)}(z) = z^{-i} A^{(i)}(z^{-1})S(z) = z^{-i} A^{(i-1)}(z^{-1})S(z) - k_i A^{(i-1)}(z)S(z)$$
$$= z^{-1} B^{(i-1)}(z) - k_i E^{(i-1)}(z)$$

- with inverse transform

$$b^{(i)}(m) = s(m-i) - \sum_{k=1}^{i} \alpha_k^{(i)} s(m+k-i) = b^{(i-1)}(m-1) - k_i e^{(i-1)}(m)$$

- with the interpretation that we are attempting to predict $s(m-i)$ from the i samples of the input that follow $s(m-i)$ => we are doing a ***backward*** prediction and $b^{(i)}(m)$ is called the ***backward prediction error sequence***

Lattice Formulations of LP



same set of samples is used to forward predict *s(m)* and backward predict *s(m-i)*

Lattice Formulations of LP

- the prediction error transform and sequence $E^{(i)}(z), e^{(i)}(m)$ can now be expressed in terms of forward and backward errors, namely

$$E^{(i)}(z) = E^{(i-1)}(z) - k_i z^{-1} B^{(i-1)}(z)$$

$$e^{(i)}(m) = e^{(i-1)}(m) - k_i b^{(i-1)}(m-1) \quad *1$$

- similarly we can derive an expression for the backward error transform and sequence at sample *m* of the form

$$B^{(i)}(z) = z^{-1} B^{(i-1)}(z) - k_i E^{(i-1)}(z)$$

$$b^{(i)}(m) = b^{(i-1)}(m-1) - k_i e^{(i-1)}(m) \quad *2$$

- these two equations define the forward and backward prediction error for an i^{th} order predictor in terms of the corresponding prediction errors of an $(i-1)^{th}$ order predictor, with the reminder that a zeroth order predictor does no prediction, so

$$e^{(0)}(m) = b^{(0)}(m) = s(m) \quad 0 \le m \le L-1 \quad *3$$

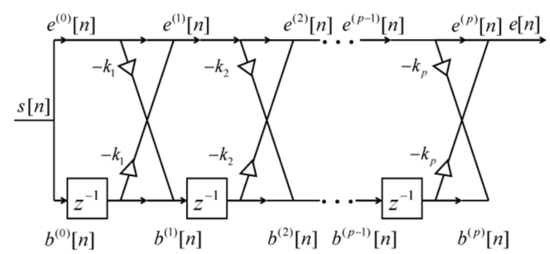
$$E(z) = E^{(p)}(z) = A(z) / S(z)$$

Lattice Formulations of LP

 Assume we know k_1 (from external computation):

1. we compute $e^{(1)}[m]$ and $b^{(1)}[m]$ from $s[m],\ 0 \le m \le L$ using Eqs. $*1$ and $*2$
2. we next compute $e^{(2)}[m]$ and $b^{(2)}[m]$ for $0 \le m \le L+1$ using Eqs. $*1$ and $*2$
3. extend solution (lattice) to p sections giving $e^{(p)}[m]$ and $b^{(p)}[m]$ for $0 \le m \le L+p-1$
4. solution is $e[n] = e^{(p)}[n]$ at the output of the p^{th} lattice section

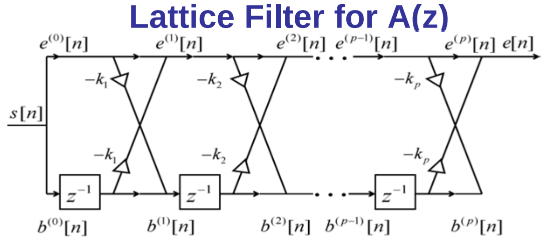
Lattice Formulations of LP



- lattice network with *p* sections—the output of which is the forward prediction error
- digital network implementation of the prediction error filter with transfer function *A(z)*

- no direct correlations
- no alphas
- *k*'s computed from forward and backward error signals

Lattice Filter for A(z)



$$e^{(0)}[n] = b^{(0)}[n] = s[n] \quad 0 \le n \le L-1$$

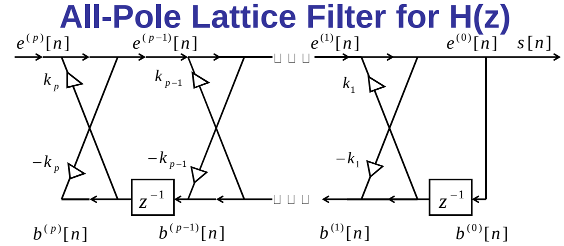
$$e^{(i)}[n] = e^{(i-1)}[n] - k_i b^{(i-1)}[n-1] \quad i = 1,2,\ldots,p,\ 0 \le n \le L-1+i$$

$$b^{(i)}[n] = -k_i e^{(i-1)}[n] + b^{(i-1)}[n-1] \quad i = 1,2,\ldots,p,\ 0 \le n \le L-1+i$$

$$e[n] = e^{(p)}[n], \quad 0 \le n \le L-1+p$$

$$E(z) = A(z)S(z)$$

All-Pole Lattice Filter for H(z)



$$e^{(p)}[n] = e[n], \quad 0 \le n \le L-1+p$$

$$e^{(i-1)}[n] = e^{(i)}[n] + k_i b^{(i-1)}[n-1]$$

$$i = p, p-1, \ldots, 1,\ 0 \le n \le L-1+i-1$$

$$b^{(i)}[n] = -k_i e^{(i-1)}[n] + b^{(i-1)}[n-1]$$

$$i = p, p-1, \ldots, 1, \quad 0 \le n \le L-1+i$$

$$s[n] = e^{(0)}[n] = b^{(0)}[n],\ 0 \le n \le L-1$$

$$S(z) = \frac{1}{A(z)} E(z)$$

All-Pole Lattice Filter for H(z)

1. since $b^{(i)}[-1]=0,\ \forall i,$ we can first solve for $e^{(i-1)}[0]$ for $i=p, p-1,...,1,$ using the relationship: $e^{(i-1)}[0]=e^{(i)}[0]$
2. since $b^{(0)}[0]=e^{(0)}[0]$ we can then solve for $b^{(i)}[0]$ for $i=1,2,...,p$ using the equation: $b^{(i)}[0]=-k_i\, e^{(i-1)}[0]$
3. we can now begin to solve for $e^{(i-1)}[1]$ as: $e^{(i-1)}[1]=e^{(i)}[1]+k_i\, b^{(i-1)}[0],\ i=p, p-1,...,1$
4. we set $b^{(0)}[1]=e^{(0)}[1]$ and we can then solve for $b^{(i)}[1]$ for $i=1,2,...,p$ using the equation: $b^{(i)}[1]=-k_i\, e^{(i-1)}[1]+b^{(i-1)}[0],\ i=1,2,...,p$
5. we iterate for $n=2,3,...,N-1$ and end up with $s[n]=e^{(0)}[n]=b^{(0)}[n]$

13

Lattice Formulations of LP

- the lattice structure comes directly out of the Durbin algorithm
- the k_i parameters are obtained from the Durbin equations
- the predictor coefficients, α_k, do not appear explicitly in the lattice structure
- can relate the k_i parameters to the forward and backward errors via

$$k_i=\frac{\sum_{m=0}^{L-1+i} e^{(i-1)}(m)b^{(i-1)}(m-1)}{\left\{\left[\sum_{m=0}^{L-1+i}[e^{(i-1)}(m)]^2\right]\left[\sum_{m=0}^{L-1+i}[b^{(i-1)}(m-1)]^2\right]\right\}^{1/2}} \quad *4$$

- where k_i is a normalized cross correlation between the forward and backward prediction error, and is therefore called a partial correlation or PARCOR coefficient
- can compute predictor coefficients recursively from the PARCOR coefficients

14

Direct Computation of *k* Parameters

- assume $s[n]$ non-zero for $0\le n\le L-1$
- assume k_i chosen to minimize total energy of the forward (or backward) prediction errors
- we can then minimize ***forward*** prediction error as:

$$E_{forward}^{(i)}=\sum_{m=0}^{L-1+i}\left[e^{(i)}(m)\right]^2=\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)-k_i b^{(i-1)}(m-1)\right]^2$$

$$\frac{\partial E_{forward}^{(i)}}{\partial k_i}=0=-2\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)-k_i b^{(i-1)}(m-1)\right]b^{(i-1)}(m-1)$$

$$k_i^{forward}=\frac{\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)\cdot b^{(i-1)}(m-1)\right]}{\sum_{m=0}^{L-1+i}\left[b^{(i-1)}(m-1)\right]^2}$$

15

Direct Computation of k Parameters

- we can also choose to minimize the ***backward*** prediction error

$$E_{backward}^{(i)}=\sum_{m=0}^{L-1+i}\left[b^{(i)}(m)\right]^2=\sum_{m=0}^{L-1+i}\left[-k_i e^{(i-1)}(m)+b^{(i-1)}(m-1)\right]^2$$

$$\frac{\partial E_{backward}^{(i)}}{\partial k_i}=0=-2\sum_{m=0}^{L-1+i}\left[-k_i e^{(i-1)}(m)+b^{(i-1)}(m-1)\right]e^{(i-1)}(m)$$

$$k_i^{backward}=\frac{\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)b^{(i-1)}(m-1)\right]}{\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)\right]^2}$$

16

Direct Computation of k Parameters

- if we window and sum over all time, then

$$\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)\right]^2=\sum_{m=0}^{L-1+i}\left[b^{(i-1)}(m-1)\right]^2$$

therefore

$$k_i^{PARCOR}=\sqrt{k_i^{forward}k_i^{backward}}=k_i^{forward}=k_i^{backward}$$

$$=\frac{\sum_{m=0}^{L-1+i} e^{(i-1)}(m)b^{(i-1)}(m-1)}{\left\{\sum_{m=0}^{L-1+i}\left[e^{(i-1)}(m)\right]^2\sum_{m=0}^{L-1+i}\left[b^{(i-1)}(m-1)\right]^2\right\}^{1/2}}$$

17

Direct Computation of *k* Parameters

- minimize ***sum*** of forward and backward prediction errors over fixed interval (covariance method)

$$E_{Burg}^{(i)}=\sum_{m=0}^{L-1}\left\{\left[e^{(i)}(m)\right]^2+\left[b^{(i)}(m)\right]^2\right\}$$

$$=\sum_{m=0}^{L-1}\left[e^{(i-1)}(m)-k_i b^{(i-1)}(m-1)\right]^2+\sum_{m=-\infty}^{\infty}\left[-k_i e^{(i-1)}(m)+b^{(i-1)}(m-1)\right]^2$$

$$\frac{\partial E_{Burg}^{(i)}}{\partial k_i}=0=-2\sum_{m=0}^{L-1}\left[e^{(i-1)}(m)-k_i b^{(i-1)}(m-1)\right]b^{(i-1)}(m-1)$$

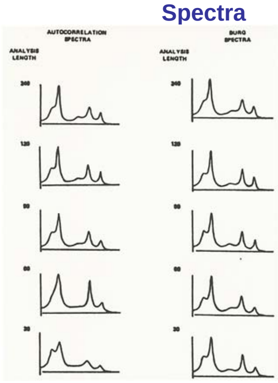
$$-2\sum_{m=0}^{L-1}\left[-k_i e^{(i-1)}(m)+b^{(i-1)}(m-1)\right]e^{(i-1)}(m)$$

$$k_i^{Burg}=\frac{2\sum_{m=0}^{L-1}\left[e^{(i-1)}(m)\cdot b^{(i-1)}(m-1)\right]}{\sum_{m=0}^{L-1}\left[e^{(i-1)}(m)\right]^2+\sum_{m=0}^{L-1}\left[b^{(i-1)}(m-1)\right]^2}$$

- $-1\le k_i^{Burg}\le 1$ ***always***

18

Comparison of Autocorrelation and Burg Spectra



• significantly less smearing of formant peaks using Burg method

Summary of Lattice Procedure

- steps involved in determining the predictor coefficients and the k parameters for the lattice method are as follows
  1. initial condition, $e^{(0)}(m) = s(m) = b^{(0)}(m)$ from Eq. *3
  2. compute $k_1 = \alpha_1^{(1)}$ from Eq. *4
  3. determine forward and backward prediciton errors $e^{(1)}(m), b^{(1)}(m)$ from Eqs. *1 and *2
  4. set $i = 2$
  5. determine $k_i = \alpha_i^{(i)}$ from Eq. *4
  6. determine $\alpha_j^{(i)}$ for $j = 1,2,\ldots,i-1$ from Durbin iteration
  7. determine $e^{(i)}(m)$ and $b^{(i)}(m)$ from Eqs. *1 and *2
  8. set $i = i+1$
  9. if i is less than or equal to p, go to step 5
  10. procedure is terminated
- predictor coefficients obtained directly from speech samples
  => without calculation of autocorrelation function
- method is guaranteed to yield stable filters ($|k_i|$<1) without using window

Summary of Lattice Procedure

Lattice Algorithms

$\mathcal{E}^{(0)} = R[0]$ (1)

$e^{(0)}[n] = b^{(0)}[n] = s[n], \quad 0 \le n \le L-1$ (2)

for $i = 1, 2, \ldots, p$

  compute k_i using either Eq. (9.125) or Eq. (9.128) (3)

  compute $e^{(i)}[n], \quad 0 \le n \le L-1+i$ using Eq. (9.117b) (4a)

  compute $b^{(i)}[n], \quad 0 \le n \le L-1+i$ using Eq. (9.117c) (4b)

  $\alpha_i^{(i)} = k_i$ (5)

compute predictor coefficients

  if $i > 1$ then for $j = 1, 2, \ldots, i-1$

    $\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)}$ (6)

  end

compute mean-squared energy

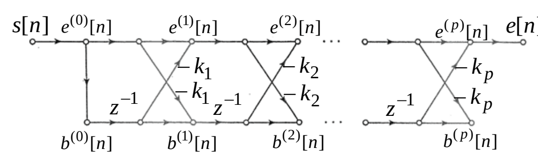
  $\mathcal{E}^{(i)} = (1 - k_i^2)\mathcal{E}^{(i-1)}$ (7)

end

$\alpha_j = \alpha_j^{(p)} \quad j = 1, 2, \ldots, p$ (8)

$e[n] = e^{(p)}[n], \quad 0 \le n \le L-1+p$ (9)

Summary of Lattice Procedure



$$e^{(0)}(m) = b^{(0)}(m) = s(m) \qquad *3$$

$$k_i = \frac{\sum_{m=0}^{L-1+i} e^{(i-1)}(m) b^{(i-1)}(m-1)}{\left\{\left[\sum_{m=0}^{L-1+i} [e^{(i-1)}(m)]^2\right]\left[\sum_{m=0}^{L-1+i} [b^{(i-1)}(m-1)]^2\right]\right\}^{1/2}} \qquad *4$$

$$e^{(i)}(m) = e^{(i-1)}(m) - k_i b^{(i-1)}(m-1) \qquad *1$$

$$b^{(i)}(m) = b^{(i-1)}(m-1) - k_i e^{(i-1)}(m) \qquad *2$$

Prediction Error Signal

1. Speech Production Model

$$s(n) = \sum_{k=1}^{p} a_k s(n-k) + Gu(n)$$

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

2. LPC Model:

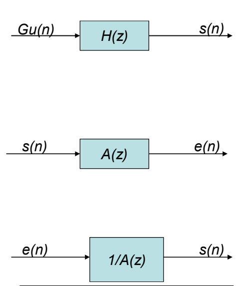
$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^{p} \alpha_k s(n-k)$$

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

3. LPC Error Model:

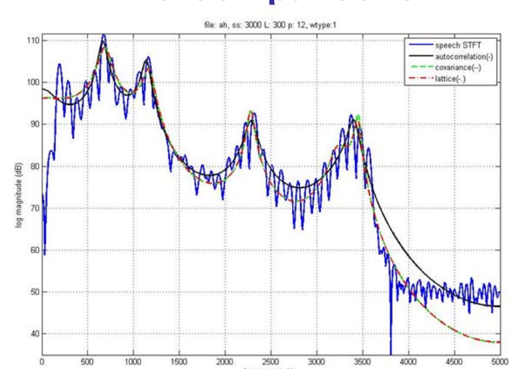
$$\frac{1}{A(z)} = \frac{S(z)}{E(z)} = \frac{1}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}}$$

$$s(n) = e(n) + \sum_{k=1}^{p} \alpha_k s(n-k)$$

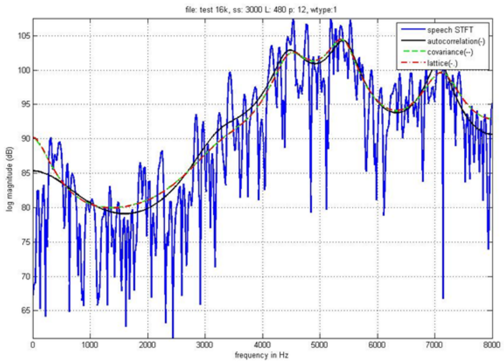


Perfect reconstruction even if a_k not equal to α_k

LPC Comparisons



LPC Comparisons



Comparisons Between LP Methods

- the various LP solution techniques can be compared in a number of ways, including the following:
  - computational issues
  - numerical issues
  - stability of solution
  - number of poles (order of predictor)
  - window/section length for analysis

26

LP Solution Computations

	Covariance Method	*Autocorrelation Method*	*Lattice Method*
	(Cholesky Decomposition)	(Durbin Method)	(Burg Method)
Storage			
Data	L_1	L_2	L_3
Matrix	$\sim p^2/2$	$\sim p$	—
Window	0	L_2	—
Computation (Multiplications)			
Windowing	0	L_2	—
Correlation	$\sim L_1 p$	$\sim L_2 p$	—
Matrix Solution	$\sim p^3$	$\sim p^2$	$5L_3 p$

- assume $L_1 \approx L_2 >> p$; choose values of $L_1=300$, $L_2=300$, $L_3=300$, $p=10$
- computation for
  - covariance method $\approx L_1 p + p^3 \approx$ 4000 *,+
  - autocorrelation method $\approx L_2 p + p^2 \approx$ 3100 *,+
  - lattice method $\approx 5L_3 p \approx$ 15000 *,+

27

LP Solution Comparisons

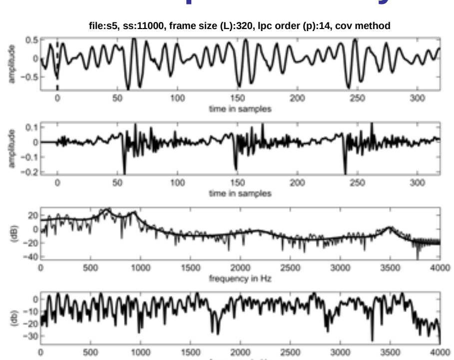
- stability
  - guaranteed for autocorrelation method
  - cannot be guaranteed for covariance method; as window size gets larger, this almost always makes the system stable
  - guaranteed for lattice method
- choice of LP analysis parameters
  - need 2 poles for each vocal tract resonance below $F_s/2$
  - need 3-4 poles to represent source shape and radiation load
  - use values of $p \approx$ 13-14

28

The Prediction Error Signal

29

LP Speech Analysis



Top panel: speech signal

Second panel: error signal

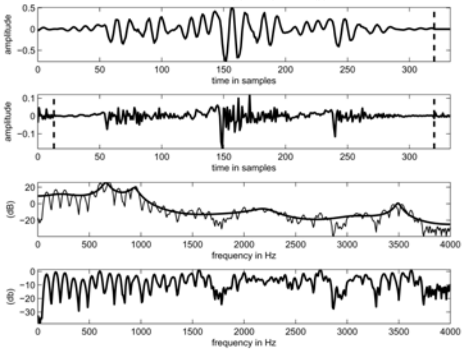
Third panel: log magnitude spectra of signal and LP model

Fourth panel: log magnitude spectrum of error signal

30

LP Speech Analysis

file:s5, ss:11000, frame size (L):320, lpc order (p):14, ac method



Top panel: speech signal

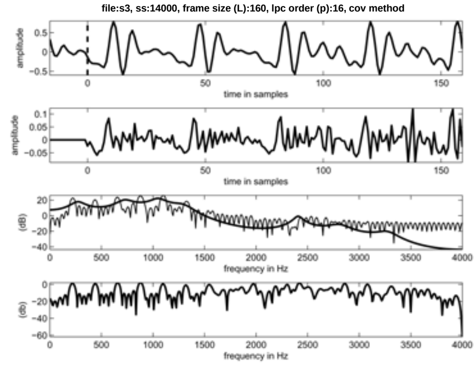
Second panel: error signal

Third panel: log magnitude spectra of signal and LP model

Fourth panel: log magnitude spectrum of error signal

LP Speech Analysis

file:s3, ss:14000, frame size (L):160, lpc order (p):16, cov method



Top panel: speech signal

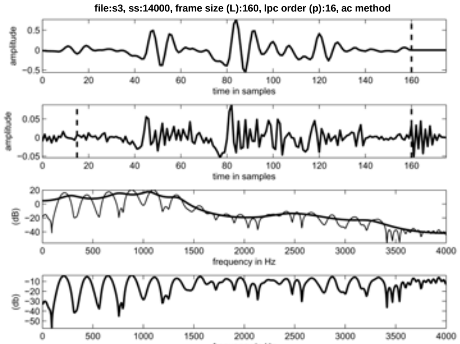
Second panel: error signal

Third panel: log magnitude spectra of signal and LP model

Fourth panel: log magnitude spectrum of error signal

LP Speech Analysis

file:s3, ss:14000, frame size (L):160, lpc order (p):16, ac method



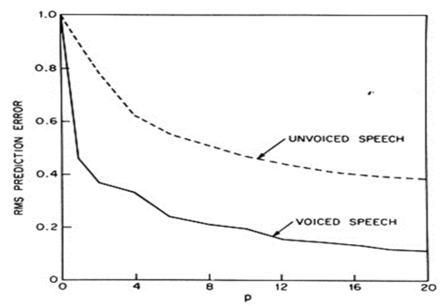
Top panel: speech signal

Second panel: error signal

Third panel: log magnitude spectra of signal and LP model

Fourth panel: log magnitude spectrum of error signal

Variation of Prediction Error with LP Order



- prediction error flattens at around *p*=13-14
- normalized prediction error for unvoiced speech larger than for voiced speech => model isn't quite as accurate

LP Solution Comparisons

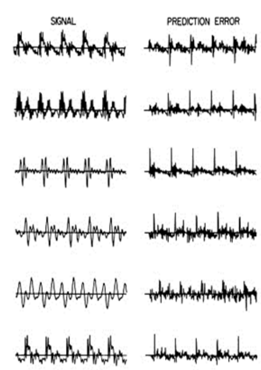
- choice of LP analysis parameters
  - use small values of *L* for reduced computation
  - use *l* order of several pitch periods for reliable prediction-especially when using tapered window
  - use *l* from 100-400 samples at 10 kHz for autocorrelation
  - for lattice and covariance methods, *L* as small as 2 *p* has been used (within pitch periods); however if pitch pulse occurs within window, bad prediction results => use much larger values of *L*

Prediction Error Signal Behavior

- the prediction error signal is computed as

$$e(n) = s(n) - \sum_{k=1}^{p} \alpha_k s(n-k) = Gu(n)$$

- $e(n)$ should be large at the beginning of each pitch period (voiced speech) => good signal for pitch detection
- can perform autocorrelation on $e(n)$ and detect largest peak
- error spectrum is approximately flat-so effects of formants on pitch detection are minimized



Normalized Mean-Squared Error

- for autocorrelation method

$$V_{\hat{n}} = \frac{\sum_{m=0}^{L+p-1} e_{\hat{n}}^2(m)}{\sum_{m=0}^{L-1} s_{\hat{n}}^2(m)}$$

- for covariance method

$$V_{\hat{n}} = \frac{\sum_{m=0}^{L-1} e_{\hat{n}}^2(m)}{\sum_{m=0}^{L-1} s_{\hat{n}}^2(m)}$$

- the prediction error sequence (defining $\alpha_0 = -1$) is

$$e_{\hat{n}}(m) = -\sum_{k=0}^{p} \alpha_k s_{\hat{n}}(m-k)$$

- giving many forms for the normalized error

$$V_{\hat{n}} = \sum_{i=0}^{p}\sum_{j=0}^{p} \alpha_i \frac{\phi_{\hat{n}}(i,j)}{\phi_{\hat{n}}(0,0)} \alpha_j$$

$$V_{\hat{n}} = -\sum_{i=0}^{p} \alpha_i \frac{\phi_{\hat{n}}(i,0)}{\phi_{\hat{n}}(0,0)}$$

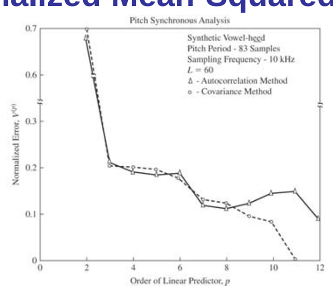
$$V_{\hat{n}} = \prod_{i=1}^{p} (1-k_i^2)$$

37

Experimental Evaluations of LPC Parameters

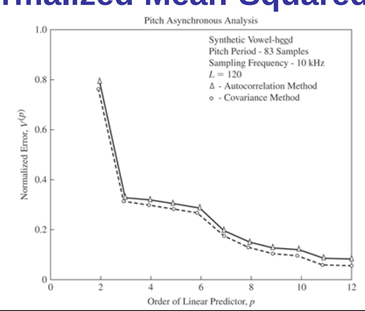
38

Normalized Mean-Squared Error



- V_n versus p for synthetic vowel, pitch period of 83 samples, L=60, pitch synchronous analysis
- covariance method-error goes to zero at p=11, the order of the synthesis filter
- autocorrelation method, V_n≈0.1 for p≥7 since error dominated by prediction at beginning of interval

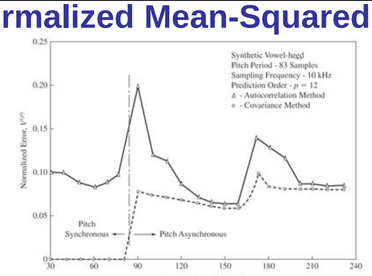
Normalized Mean-Squared Error



- normalized error versus p for pitch asynchronous analysis, L=120, normalized error falls to 0.1 near p=11 for both covariance and autocorrelation methods

40

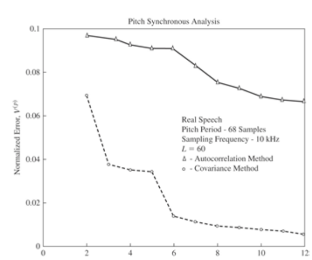
Normalized Mean-Squared Error

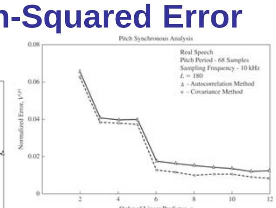


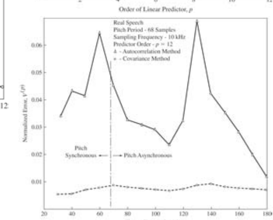
- normalized error versus L, p=12
- for L < pitch period (83 samples), covariance method gives smaller normalized error than autocorrelation method
- for values of L at or near multiples of pitch period, normalized error jumps due to large prediction error in vicinity of pitch pulse
- when L > 2 * pitch period, normalized error same for both autocorrelation and covariance methods

41

Normalized Mean-Squared Error

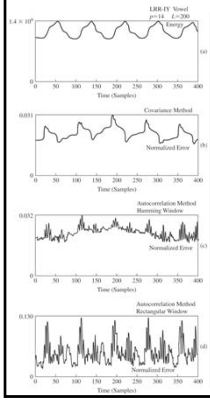






- results for natural speech

Normalized Mean-Squared Error



- variability of normalized error with positioning of frame
- sample-by-sample LP analysis of 40 msec of vowel /i/
- signal energy in part a; normalized error in part b for *p*=14 pole analysis using 20 msec frame size for covariance method; normalized error in part c for autocorrelation method using HW; normalized error in part d for autocorrelation method using RW
- average pitch period of 84 samples => 2.5 pitch periods in 20 msec window
- substantial variations in normalized error for covariance method-especially peaked when pitch pulses enter window => longer windows give larger normalized errors
- substantial high frequency variations in normalized error for autocorrelation method with some pitch modulations

Properties of the LPC Polynomial

44

Minimum-Phase Property of $A(z)$

$A(z)$ has all its zeros inside the unit circle

Proof: Assume that z_o ($|z_o|^2 > 1$) is a zero (root) of $A(z)$

$$A(z) = (1 - z_o z^{-1})A'(z)$$

The minimum mean-squared error is

$$E_{\hat{n}} = \sum_{m=-\infty}^{\infty} e_{\hat{n}}[m]^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |A(e^{j\omega})|^2 |S_{\hat{n}}(e^{j\omega})|^2 \, d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|1 - z_o e^{-j\omega}\right|^2 \left|A'(e^{j\omega})\right|^2 \left|S_{\hat{n}}(e^{j\omega})\right|^2 d\omega > 0$$

$$\left|1 - z_o e^{-j\omega}\right|^2 = \left|z_o\right|^2 \left|1 - (1/z_o^*)e^{-j\omega}\right|^2$$

Thus, $A(z)$ could not be the optimum filter because we could replace z_o by $(1/z_o^*)$ and decrease the error.

45

PARCORs and Stability

- prove that $|k_i| \geq 1 \Rightarrow |z_j^{(i)}| \geq 1$ for some j

$$A^{(i)}(z) = A^{(i-1)}(z) - k_i z^{-i} A^{(i-1)}(z^{-1}) = \prod_{j=1}^{i}(1 - z_j^{(i)} z^{-1})$$

It is easily shown that $-k_i$ is the coefficient of z^{-i} in $A^{(i)}(z)$, i.e., $\alpha_i^{(i)} = k_i$. Therefore,

$$|k_i| = \prod_{j=1}^{p} z_j^{(i)}$$

If $|k_i| \geq 1$, then either all the roots must be **on** the unit circle or at least one of them must be **outside** the unit circle.

46

PARCORs and Stability

- if $|k_i| \geq 1$, then either all the roots must be **on** the unit circle or at least one of them must be **outside** the unit circle. Since this is true for all $A^{(i)}(z)$, $i = 1, 2, \ldots, p$, a necessary condition for the roots of $A^{(p)}(z)$ to be inside the unit circle is:

$$|k_i| < 1, \quad i = 1, 2, \ldots, p$$

- for the i^{th}-order optimum linear predictor,

$$E^{(i)} = (1 - k_i^2)E^{(i-1)} = \prod_{j=1}^{i}(1 - k_j^2)E^{(0)} > 0$$

so $|k_i| < 1$ and therefore $A^{(p)}(z)$ has all its roots inside the unit circle.

47

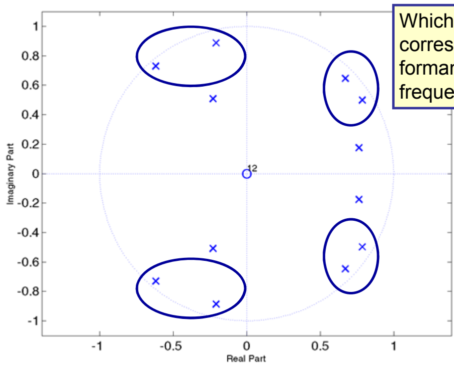
Root Locations for Optimum LP Model

$$\tilde{H}(z) = \frac{G}{A(z)} = \frac{G}{1 - \sum_{i=1}^{p} \alpha_i z^{-i}}$$

$$= \frac{G}{\prod_{i=1}^{p}(1 - z_i z^{-1})} = \frac{G z^p}{\prod_{i=1}^{p}(z - z_i)}$$

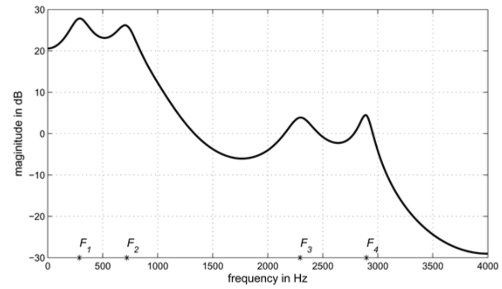
48

Pole-Zero Plot for Model



Which poles correspond to formant frequencies?

Pole Locations



Pole Locations (F_S=10,000 Hz)

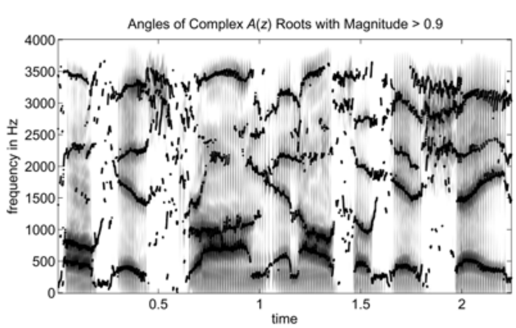
root magnitude	*θ root angle(degrees)*	*F root angle (Hz)*	*formant*
0.9308	10.36	288	F_1
0.9308	-10.36	-288	F_1
0.9317	25.88	719	F_2
0.9317	-25.88	-719	F_2
0.7837	35.13	976	
0.7837	-35.13	-976	
0.9109	82.58	2294	F_3
0.9109	-82.58	-2294	F_3
0.5579	91.44	2540	
0.5579	-91,44	-2540	
0.9571	104.29	2897	F_4
0.9571	-104.29	-2897	F_4

$$F = (\theta / 180) \cdot (F_S / 2)$$

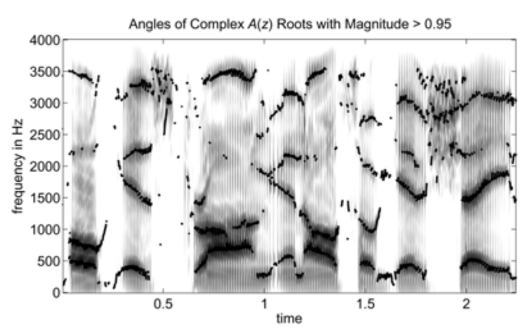
Estimating Formant Frequencies

- compute *A(z)* and factor it.
- find roots that are close to the unit circle.
- compute equivalent analog frequencies from the angles of the roots.
- plot formant frequencies as a function of time.

Spectrogram with LPC Roots



Spectrogram with LPC Roots



Comparison to ABS Methods

- error measure for ABS methods is log ratio of power spectra, i.e.,

$$E' = \int_{-\pi}^{\pi} \left\{ \log \left[\frac{|S_n(e^{j\omega})|^2}{|H(e^{j\omega})|^2} \right] \right\}^2 d\omega$$

- thus for ABS minimization of E' is equivalent to minimizing mean squared error between two log spectra
- comparing E_{LPC} and E_{ABS} we see the following:
  - both error measures related to ratio of signal to model spectra
  - both tend to perform uniformly over whole frequency range
  - both are suitable to selective error minimization over specified frequency ranges
  - error criterion for LP modeling places higher weight on frequency regions

where $|S_n(e^{j\omega})|^2 < |H(e^{j\omega})|^2$, whereas the ABS error criterion places equal weight on both regions

$\Rightarrow$ for unsmoothed signal spectra (as obtained by STFT methods), LP error criterion yields better spectral matches than ABS error criterion

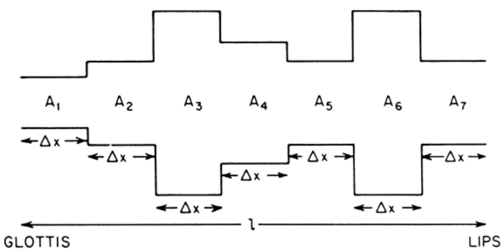
$\Rightarrow$ for smoothed signal spectra (as obtained at output of filter banks), both error criteria will perform well

55

Relation of LP Analysis to Lossless Tube Models

56

Discrete-Time Model - I



- Make all sections the same length with delay

$$\tau = \Delta x / c \quad \text{where} \quad \ell = N\Delta x$$

57

Discrete-Time Model - II

- N-section lossless tube model corresponds to discrete-time system when:

$$F_S = \frac{cN}{2\ell}$$

where c is the velocity of sound, N is the number of tube sections, F_S is the sampling frequency, and ℓ is the total length of the vocal tract.

- The reflection coefficients $\{r_k,\ 1 \le k \le N-1\}$ are related to the areas of the lossless tubes by:

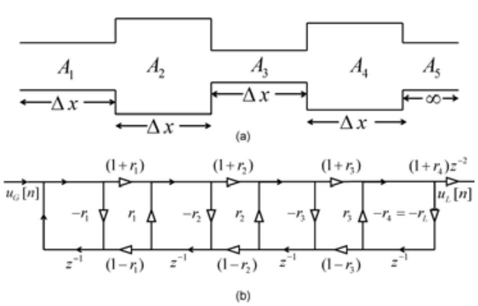
$$r_k = \frac{A_{k+1} - A_k}{A_{k+1} + A_k}$$

- Can find transfer function of digital system subject to constraints of the form $r_G = 1$,

$$r_N = r_L = \frac{\rho c / A_N - Z_L}{\rho c / A_N + Z_L}$$

58

Discrete-Time Model - III



59

Discrete-Time Model - IV

Given the system function:

$$V(z) = \frac{U_L(z)}{U_G(z)} = \frac{0.5(1+r_G)\prod_{k=1}^{N}(1+r_k)z^{-N/2}}{D(z)}$$

$$-1 \le r_k = \left(\frac{A_{k+1} - A_k}{A_{k+1} + A_k} \right) \le 1 \quad \text{reflection coefficient}$$

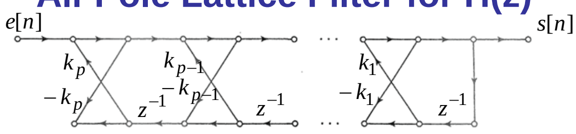
- if $r_G = 1$ (i.e., $R_G = \infty$) and $r_N = r_L = \dfrac{\rho c / A_N - Z_L}{\rho c / A_N + Z_L}$

then $D(z)$ satisties the recursion

$$\begin{aligned} D_0(z) &= 1 \\ D_k(z) &= D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}) \quad k = 1,2,\dots,N \\ D(z) &= D_N(z) \end{aligned}$$

60

All-Pole Lattice Filter for H(z)



$$\begin{aligned} A^{(0)}(z) &= 1 \\ A^{(i)}(z) &= A^{(i-1)}(z) - k_i z^{-i} A^{(i-1)}(z^{-1}) \quad i = 1,2,\ldots,p \\ A(z) &= A^{(p)}(z) \end{aligned}$$

$$S(z) = \frac{1}{A(z)} E(z)$$

If $r_i = -k_i$ and $N = p$, then $D(z) = A(z)$.

61

Tube Areas from PARCORS

- Relation to areas:

$$-1 \le -k_i = r_i = \left(\frac{A_{i+1} - A_i}{A_{i+1} + A_i} \right) = \left(\frac{(A_{i+1}/A_i) - 1}{(A_{i+1}/A_i) + 1} \right) \le 1$$

- Solve for A_{i+1} in terms of A_i

$$A_{i+1} = \left(\frac{1-k_i}{1+k_i} \right) A_i > 0 \qquad \frac{A_{i+1}}{A_i} = \left(\frac{1-k_i}{1+k_i} \right) > 0$$

- Log area ratios (good for quantization)

$$g_i = \log\left(\frac{A_{i+1}}{A_i} \right) = \log\left(\frac{1-k_i}{1+k_i} \right)$$

Minimizes spectral sensitivity under uniform quantization

62

Estimating Tube Areas from Speech (Wakita)

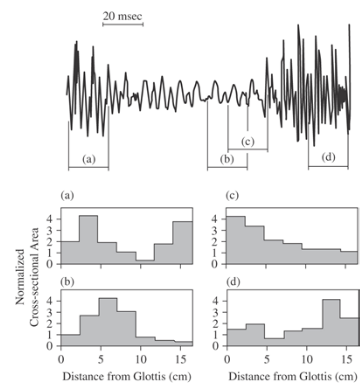
1. Sample speech $[s[n] = s_a(nT)]$ at sampling rate $F_s = 1/T = pc/(2\ell)$.
2. Remove effects of glottal source and radiation by pre-emphasis $x[n] = s[n] - s[n-1]$.
3. Compute the PARCOR coefficients on a short-time basis: k_i, $i = 1,2,\ldots,p$.
4. Assuming $A_1 = 1$ (arbitrary), compute

$$A_{i+1} = \left(\frac{1-k_i}{1+k_i} \right) A_i, \quad i = 1,2,\ldots,p-1$$

H. Wakita, IEEE Trans. Audio and Electroacoustics, October, 1973.

63

Estimation of Tube Areas

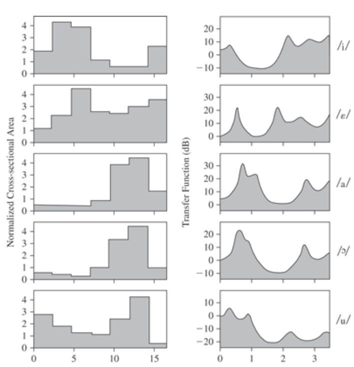


Analysis of syllable /IY/ /B/ /AA/
- /IY/ sound in part (a)
- /B/ sound in parts (b) and (c)
- /AA/ sound in part (d)

Using Wakita method to estimate tube areas from the speech waveform as shown in lower plot.

64

Estimations of Tube Areas



Use of Wakita method to estimate tube areas for the five vowels, /IY/, /EH/, /AA/, /AO/, /UW/

65

Alternative Representations of the LP Parameters

66

LP Parameter Sets

Parameter Set	*Representation*
LP Coefficients and Gain	$\{\alpha_k, 1 \le k \le p\}, G$
PARCOR Coefficients	$\{k_i, 1 \le i \le p\}$
Log Area Ratio Coefficients	$\{g_i, 1 \le i \le p\}$
Roots of Predictor Polynomial	$\{z_k, 1 \le k \le p\}$
Impulse Response of $H(z)$	$\{h[n], 0 \le n \le \infty\}$
LP Cepstrum	$\{\hat{h}[n], -\infty \le n \le \infty\}$
Autocorrelation of Impulse Response	$\{\tilde{R}(i), -\infty \le i \le \infty\}$
Autocorrelation of Predictor Polynomial	$\{R_a[i], -p \le i \le p\}$
Line Spectral Pair Parameters	$P(z), Q(z)$

PARCORs to Prediction Coefficients

- assume that $k_i, i = 1,2,\ldots,p$ are given. Then we can skip the computation of k_i in the Levinson recursion.

for $i = 1,2,\ldots,p$
  $\alpha_i^{(i)} = k_i$
  if $i > 1$, then for $j = 1,2,\ldots,i-1$
    $\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)}$
  end
end
$\alpha_j = \alpha_j^{(p)} \quad j = 1,2,\ldots,p$

Prediction Coefficients to PARCORs

- assume that $\alpha_j, j = 1,2,\ldots,p$ are given. Then we can work backwards through the Levinson Recursion.

$\alpha_j^{(p)} = \alpha_j \quad$ for $j = 1,2,\ldots,p$
$k_p = \alpha_p^{(p)}$
for $i = p, p-1,\ldots,2$
  for $j = 1,2,\ldots,i-1$
    $$\alpha_j^{(i-1)} = \frac{\alpha_j^{(i)} + k_i \alpha_{i-j}^{(i)}}{1 - k_i^2}$$
  end
  $k_{i-1} = \alpha_{i-1}^{(i-1)}$
end

LP Parameter Transformations

- roots of the predictor polynomial

$$A(z) = 1 - \sum_{k=1}^{p} \alpha_k z^{-k} = \prod_{k=1}^{p} \left(1 - z_k z^{-1}\right)$$

- where each root can be expressed as a z-plane or s-plane root, i.e.,

$$z_k = z_{kr} + j\, z_{ki}; \qquad s_k = \sigma_k + j\Omega_k$$
$$z_k = e^{s_k T}$$

giving

$$\Omega_k = \frac{1}{T} \tan^{-1}\left[\frac{z_{ki}}{z_{kr}}\right]; \qquad \sigma_k = \frac{1}{2T} \log\left(z_{kr}^2 + z_{ki}^2\right)$$

- important for formant estimation

LP Parameter Transformations

- cepstrum of IR of overall LP system from predictor coefficients

$$\hat{h}(n) = \alpha_n + \sum_{k=1}^{n-1} \left(\frac{k}{n}\right) \hat{h}(k) \alpha_{n-k} \qquad 1 \le n$$

- predictor coefficients from cepstrum of IR

$$\alpha_n = \hat{h}(n) - \sum_{k=1}^{n-1} \left(\frac{k}{n}\right) \hat{h}(k) \alpha_{n-k} \qquad 1 \le n$$

where

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n} = \frac{G}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}}$$

LP Parameter Transformations

- IR of all pole system

$$h(n) = \sum_{k=1}^{p} \alpha_k h(n-k) + G\delta(n) \qquad 0 \le n$$

- autocorrelation of IR

$$\tilde{R}(i) = \sum_{n=0}^{\infty} h(n) h(n-i) = \tilde{R}(-i)$$

$$\tilde{R}(i) = \sum_{k=1}^{p} \alpha_k \tilde{R}(|i-k|) \qquad 1 \le i$$

$$\tilde{R}(0) = \sum_{k=1}^{p} \alpha_k \tilde{R}(k) + G^2$$

LP Parameter Transformations

- autocorrelation of the predictor polynomial

$$A(z) = 1 - \sum_{k=1}^{p} \alpha_k z^{-k}$$

with IR of the inverse filter

$$a(n) = \delta(n) - \sum_{k=1}^{p} \alpha_k \delta(n-k)$$

with autocorrelation

$$R_a(i) = \sum_{k=0}^{p-i} a(k)a(k+i) \qquad 0 \le i \le p$$

73

LP Parameter Transformations

- log area ratio coefficients from PARCOR coefficients

$$g_i = \log\left[\frac{A_{i+1}}{A_i}\right] = \log\left[\frac{1-k_i}{1+k_i}\right] \qquad 1 \le i \le p$$

with inverse relation

$$k_i = \frac{1-e^{g_i}}{1+e^{g_i}} \qquad 1 \le i \le p$$

74

Quantization of LP Parameters

- consider the magnitude-squared of the model frequency response

$$\left|H(e^{j\omega})\right|^2 = \frac{1}{|A(e^{j\omega})|^2} = P(\omega, g)$$

where g is a parameter that affects P.

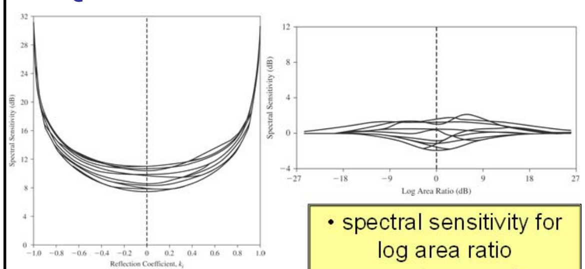
- spectral sensitivity can be defined as

$$\frac{\partial S}{\partial g_i} = \lim_{\Delta g_i \to 0} \left| \frac{1}{\Delta g_i} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \log \frac{P(\omega, g_i)}{P(\omega, g_i + \Delta g_i)} \right| d\omega \right] \right|$$

which measures sensitivity to errors in the g_i parameters

75

Quantization of LP Parameters



- spectral sensitivity for k_i parameters; low sensitivity around 0; high sensitivity around 1

- spectral sensitivity for log area ratio parameters, g_i – low sensitivity for virtually entire range is seen

76

Line Spectral Pair Parameters

$A(z) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \ldots + \alpha_p z^{-p}$

$=$ all-zero prediction filter with all zeros, z_k, inside the unit circle

$\tilde{A}(z) = z^{-(p+1)} A(z^{-1}) = \alpha_p z^{-1} + \ldots + \alpha_2 z^{-p+1} + \alpha_1 z^{-p} + z^{-(p+1)}$

$=$ reciprocal polynomial with inverse zeros, $1/z_k$

- consider the following:

$L(z) = \frac{\tilde{A}(z)}{A(z)} =$ allpass system $\Rightarrow \left|L(e^{j\omega})\right| = 1,$ all ω

- form the symmetric polynomial $P(z)$ as:

$P(z) = A(z) + \tilde{A}(z) = A(z) + z^{-(p+1)} A(z^{-1}) \Rightarrow P(z)$ has zeros for $L(z) = -1; \; (A(z) = -\tilde{A}(z))$

$\Rightarrow \arg\{L(e^{j\omega_k})\} = (k + 1/2) \cdot 2\pi, \; k = 0,1,\ldots,p-1$

- form the anti-symmetric polynomial $Q(z)$ as:

$Q(z) = A(z) - \tilde{A}(z) = A(z) - z^{-(p+1)} A(z^{-1}) \Rightarrow Q(z)$ has zeros for $L(z) = +1; \; (A(z) = \tilde{A}(z))$

$\Rightarrow \arg\{L(e^{j\omega_k})\} = k \cdot 2\pi, \; k = 0,1,\ldots,p-1$

77

Line Spectral Pair Parameters

- zeros of $P(z)$ and $Q(z)$ fall on unit circle and are interleaved with each other $\Rightarrow$ set of $\{\omega_k\}$ called Line Spectral Frequencies (LSF)
- LSFs are in ascending order
- stability of $H(z)$ guaranteed by quantizing LSF parameters

$$A(e^{j\omega}) = \frac{P(e^{j\omega}) + Q(e^{j\omega})}{2}$$

$$\left|A(e^{j\omega})\right|^2 = \frac{\left|P(e^{j\omega})\right|^2 + \left|Q(e^{j\omega})\right|^2}{4}$$

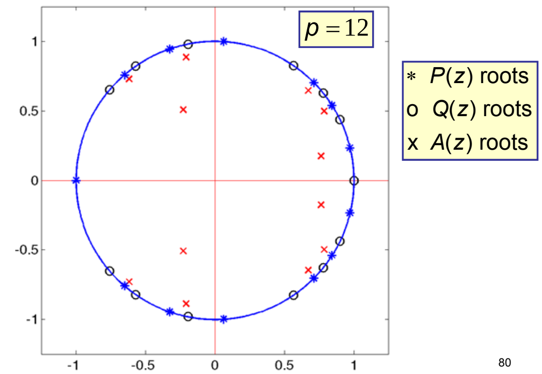
78

Line Spectrum Pair (LSP) Parameters

- properties of LSP parameters
  1. $P(z)$ corresponds to a lossless tube, open at the lips and open ($k_{p+1}=1$) at the glottis
  2. $Q(z)$ corresponds to a lossless tube, open at the lips and closed ($k_{p+1}=-1$) at the glottis
  3. all the roots of $P(z)$ and $Q(z)$ are on the unit circle
  4. if p is an even integer, then $P(z)$ has a root at $z=+1$ and $Q(z)$ has a root at $z=-1$
  5. a necessary and sufficient condition for $|k_i|<1,\ i=1,2,...,p$ is that the roots of $P(z)$ and $Q(z)$ alternate on the unit circle
  6. the LSP frequencies get close together when roots of $A(z)$ are close to the unit circle
  7. the roots of $P(z)$ are approximately equal to the formant frequencies

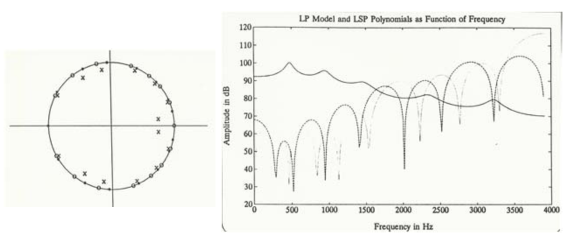
79

LSP Example



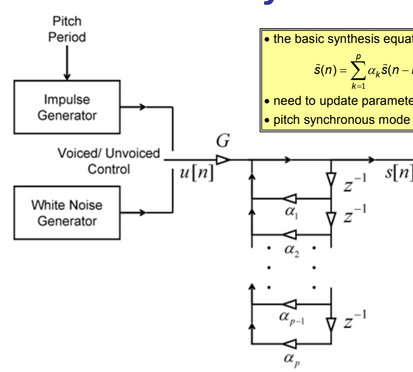
80

Line Spectrum Pair (LSP) Parameters



81

LPC Synthesis



- the basic synthesis equation is

$$\tilde{s}(n)=\sum_{k=1}^{p}\alpha_k\tilde{s}(n-k)+Gu(n)$$

- need to update parameters every 10 msec or so
- pitch synchronous mode works best

82

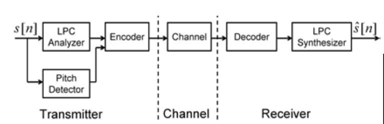
LPC Analysis-Synthesis

1. Extract α_k parameters properly
2. Quantize α_k parameters properly so that there is little quantization error
   - Small number of bits go into coding the α_k coefficients
3. Represent *e(n)* via:
   - Pitch pulses and noise—LPC Coding
   - Multiple pulses per 10 msec interval—MPLPC Coding
   - Codebook vectors—CELP
     - Almost all of the coding bits go into coding of *e(n)*

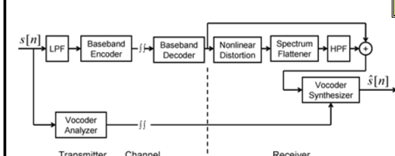
83

LPC Vocoder

header



- bit rates of 72 F_S where F_S=100, 67, and 33 for bit rates of 7200 bps, 4800 bps and 2400 bps



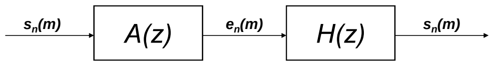
3600 bps

2400 bps

VEV LPC

84

LPC Basics



$$A(z) = 1 - \sum_{k=1}^{p} \alpha_k z^{-k} = 1 - P(z) = \frac{E(z)}{S(z)};\ \text{prediction error filter}$$

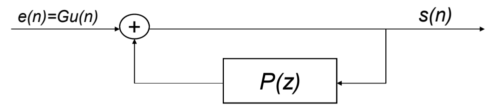
$$e_{\hat{n}}(m) = s_{\hat{n}}(m) - \sum_{k=1}^{p} \alpha_k s_{\hat{n}}(m-k);\ \text{prediction error}$$

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}} = \frac{1}{1 - P(z)};\ \text{all pole model}$$

$$E_{\hat{n}} = \sum_{m=-\infty}^{\infty} \left[e_{\hat{n}}(m)\right]^2 = \sum_{m=-\infty}^{\infty} \left(s_{\hat{n}}(m) - \sum_{k=1}^{p} \alpha_k s_{\hat{n}}(m-k) \right)^2$$

85

LPC Basics-Speech Model



$$H(z) = \frac{G}{1 - \sum_{k=1}^{p} \alpha_k z^{-k}} = \frac{G}{1 - P(z)} = \frac{S(z)}{U(z)}$$

$$H(e^{j\omega}) = \frac{G}{1 - \sum_{k=1}^{p} \alpha_k e^{-j\omega k}}$$

86

Summary

- the LP model has many interesting and useful properties that follow from the structure of the Levinson-Durbin algorithms
- the different equivalent representations have different properties under quantization
  - polynomial coefficients (bad)
  - polynomial roots (okay)
  - PARCOR coefficients (okay)
  - lossless tube areas (good)
  - LSP root angles (good)
- almost all LPC representations can be used with a range of compression schemes and are all good candidates for the technique of Vector Quantization

87