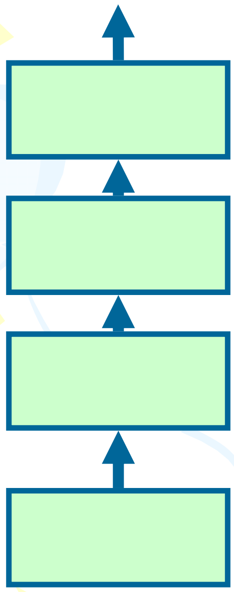
音声生成の線形モデル

音声



$s(t)$

$l(t)$

$v(t)$

$g(t)$

$e(t)$

$$S(z) = E(z)G(z)V(z)L(z)$$

「　　　　　　　　」

「　　　　　　　　」

音声生成の線形モデル (2)

- インパルス列

$$E(z) = \sigma \sum_{n=0}^{\infty} (z^{-I})^n = \frac{\sigma}{1 - z^{-I}}$$

T:

P:　　　　　　　　　　=I×T (Iは整数)

音声生成の線形モデル (3)

- 生成(声門)モデル

$$G(z)=\frac{1}{(1-e^{-cT}z^{-1})^2} \qquad (cT<<1)$$

- 放射(口唇)モデル

$$L(z)=1-z^{-1}$$

音声生成の線形モデル (4)

- 共鳴(声道)モデル

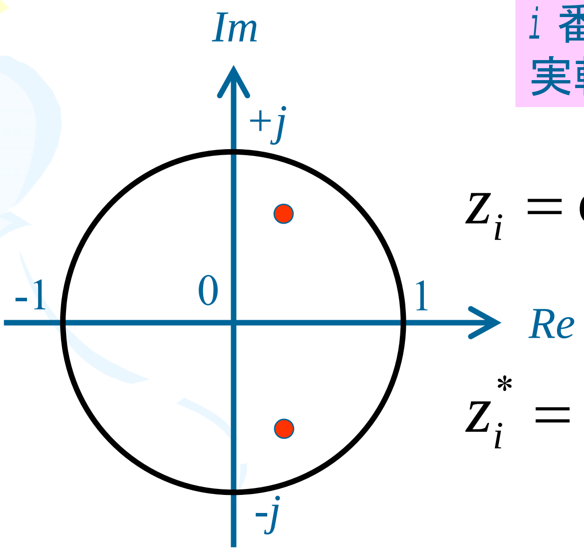
$$V(z)=\frac{1}{\prod_{i=1}^{k}\left[1-2e^{-c_iT}\cos(b_iT)z^{-1}+e^{-2c_iT}z^{-2}\right]}$$

i 番目のホルマントの
周波数: F i
帯域幅: B i

$$F_i=b_i/2\pi$$
$$B_i=c_i/2\pi$$

音声生成の線形モデル (5)

- 共鳴（声道）モデル



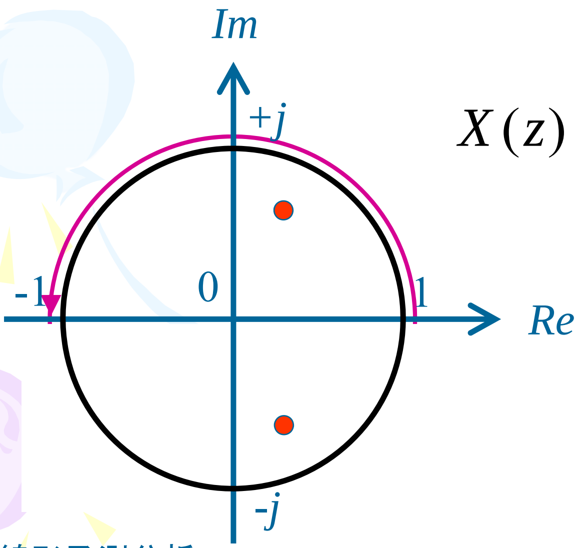
i 番目のホルマントについて
実軸に対称な2つの極が対応

$$z_i = \exp(-c_i T + j \cdot b_i T)$$

$$z_i^* = \exp(-c_i T - j \cdot b_i T)$$

極とスペクトル

- z が 1 から −1 まで半周するときの $X(z)$ の値がスペクトルに対応　⇒ 極(z_i)の近くでピーク



$$X(z) = \frac{1}{\prod_i^k (1 - z_i \cdot z^{-1})(1 - z_i^* \cdot z^{-1})}$$

自習課題

- i 番目のホルマントに、実軸に対称な2つの極

$$z_i = \exp(-c_i T + j \cdot b_i T), \quad z_i^* = \exp(-c_i T - j \cdot b_i T)$$

を対応させるとき、k 個のホルマントを持つ声道の伝達関数が下記のようになることを導出せよ。

$$V(z) = \frac{1}{\prod_{i=1}^{k}\left[1 - 2e^{-c_i T}\cos(b_i T)z^{-1} + e^{-2c_i T}z^{-2}\right]}$$

音声生成の線形モデル (6)

- 伝達系の総合特性

$$G(z)V(z)L(z)$$

$$= \frac{(1 - z^{-1})}{(1 - e^{-cT}z^{-1})^2 \prod_{i=1}^{k}\left[1 - 2e^{-c_i T}\cos(b_i T)z^{-1} + e^{-2c_i T}z^{-2}\right]}$$

(cT<<1) より約分

全極型のフィルタになる

$$= \frac{1}{(1 - e^{-cT}z^{-1}) \prod_{i=1}^{k}\left[1 - 2e^{-c_i T}\cos(b_i T)z^{-1} + e^{-2c_i T}z^{-2}\right]}$$

音声生成の線形モデル (7)

- 全極型音声生成モデル

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

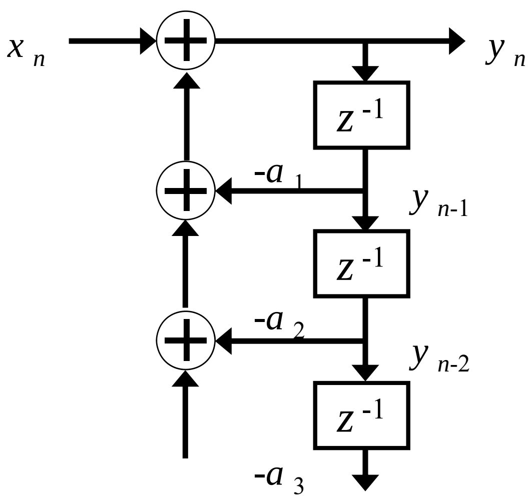
- $A(z)$ を「　　　　　　　　　　　　　　」と呼ぶ

$$A(z) \equiv \frac{1}{G(z)V(z)L(z)}$$

全極型音声生成フィルタ

- $1/A(z)$ はIIRフィルタ

Infinite
Impulse
Response



逆フィルタ

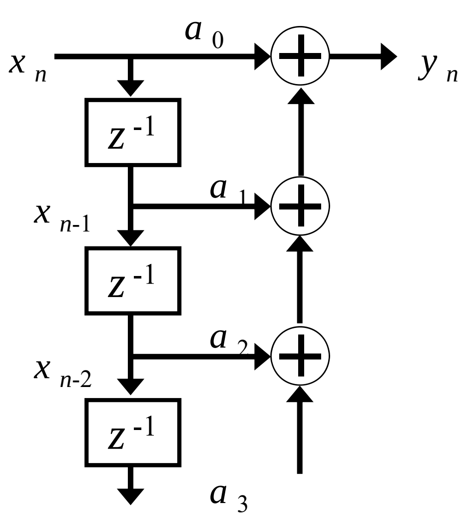
- 逆フィルタは全零型フィルタ

$$
\begin{aligned}
&A(z) \\
&= (1-e^{-cT}z^{-1})\prod_{i=1}^{k}\left[1-2e^{-c_iT}\cos(b_iT)z^{-1}+e^{-2c_iT}z^{-2}\right] \\
&= \sum_{i=0}^{M} a_i z^{-i} \qquad (a_0 = 1)
\end{aligned}
$$

逆フィルタ (2)

- $A(z)$ はインパルス応答長 M の FIR フィルタ

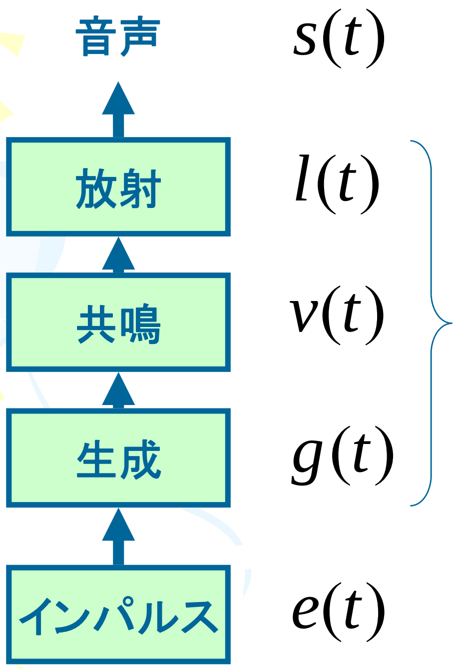
Finite Impulse Response



線形予測分析

- LPC: Linear Predictive Coding
- スペクトル解析の一手法
  -
  -
- 伝達系を全極型 IIR フィルタで表現
  -

線形予測分析 (2)



線形予測分析 (3)

- 逆フィルタ $A(z)$

周波数域領域:積
時間領域:畳込

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

$$e(n) = \sum_{i=0}^{M} a_i \cdot s(n-i)$$

a_i は IIR フィルタの係数:
「線形予測係数」と呼ぶ

$$= s(n) + \sum_{i=1}^{M} a_i \cdot s(n-i) \qquad (a_0 = 1)$$

$$= s(n) - \hat{s}(n) \quad \left(\hat{s}(n) = -\sum_{i=1}^{M} a_i \cdot s(n-i) \right)$$

線形予測分析 (4)

- $e(n)$ はインパルス列 ⇒ ほとんどの n で 0
  - 2乗誤差αを最小化する係数 a_i を求める

$$\alpha = \sum_{n=n_0}^{n_1} e^2(n)$$

$$= \sum_{n=n_0}^{n_1} \sum_{i=0}^{M} \sum_{j=0}^{M} a_i s(n-i) s(n-j) a_j$$

$$= \sum_{i=0}^{M} \sum_{j=0}^{M} a_i c_{i\,j} a_j \quad \left(c_{i\,j} = \sum_{n=n_0}^{n_1} s(n-i) s(n-j) \right)$$

線形予測分析 (5)

- αの最小化 ⇒ αを a_k で偏微分

$$0 = 2\sum_{i=0}^{M} a_i c_{i\,k}$$

- M元連立一次方程式がM個得られる

$$\sum_{i=1}^{M} a_i c_{i\,k} = -c_{0\,k} \quad (k = 1, 2, \cdots, M)$$

$a_0 = 1$ の分を右辺に移動

線形予測分析 (6)

- M元連立一次方程式がM個
  - 「共分散法」: そのまま代数的に解く

$$c_{i\,j} = \sum_{n=M}^{N-1} s(n-i)s(n-j)$$

N は観測点の数

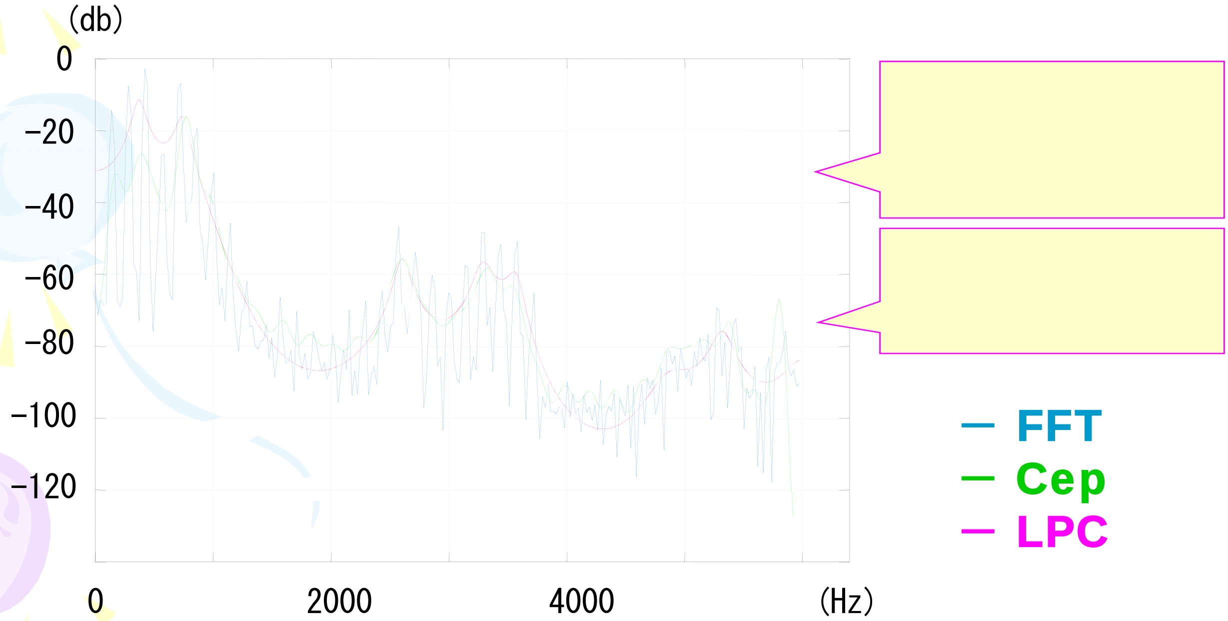
  - 「自己相関法」:
    近似を入れて安定性を確保

$$r(l) = \sum_{n=0}^{N-1-l} s(n)s(n+l) \quad (l \geq 0)$$

$r(l)$ を自己相関という

線形予測によるスペクトル概形分析

- 母音 /o/ のLPC スペクトル



ここまでのポイント

1. 声道の伝達関数
2. 線形予測分析における逆フィルタとは何か？
3. 観測された音声波形 $s(t)$ から、線形予測係数 a_i を計算する手順を説明せよ。