



東京大学 工学部 計数工学科/物理工学科

# 応用音響学：残差信号とピッチ抽出

嵯峨山 茂樹 <[sagayama@hil.t.u-tokyo.ac.jp](mailto:sagayama@hil.t.u-tokyo.ac.jp)>

東京大学 工学部 計数工学科 <http://hil.t.u-tokyo.ac.jp/>

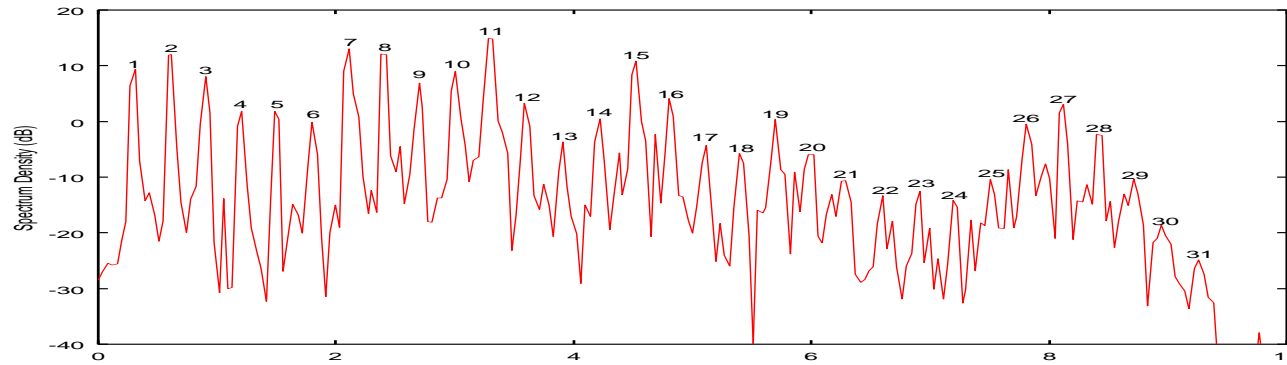
---

■ LPC 残差

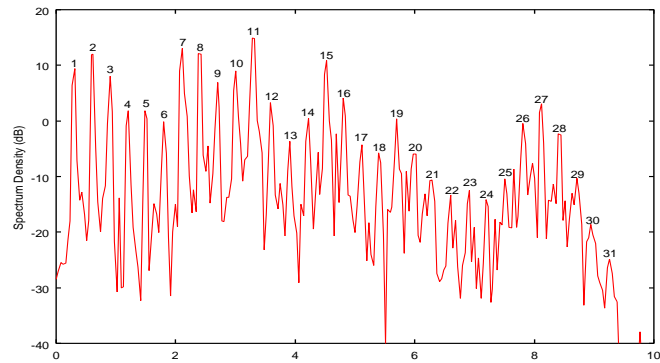
■ 零交差波



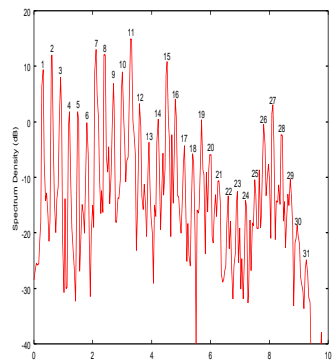
# 速回し、遅回し



20kHz サンプリングされた  
女声朗読音声「私はATR自  
動翻訳電話研究所の山田と  
申しますが。」中の母音/a/の  
スペクトル



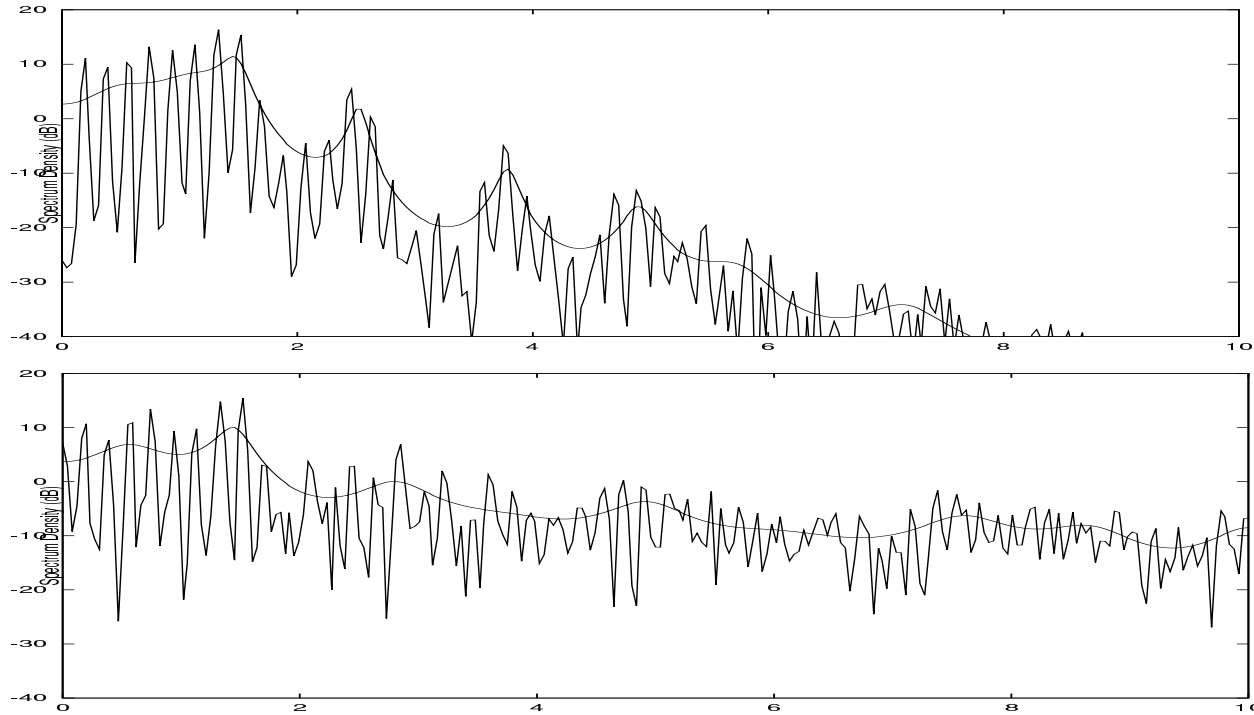
女声朗読音声の10kHz再生



女声朗読音声の5kHz再生



# 零交差波 (1bit 量子化)



男声朗読音声「来年の第一回通訳電話国際会議についてお伺いしたいんですけども、こちらでよろしいのでしょうか？」  
(mausc205:1.00s) 中の母音/a/のスペクトル

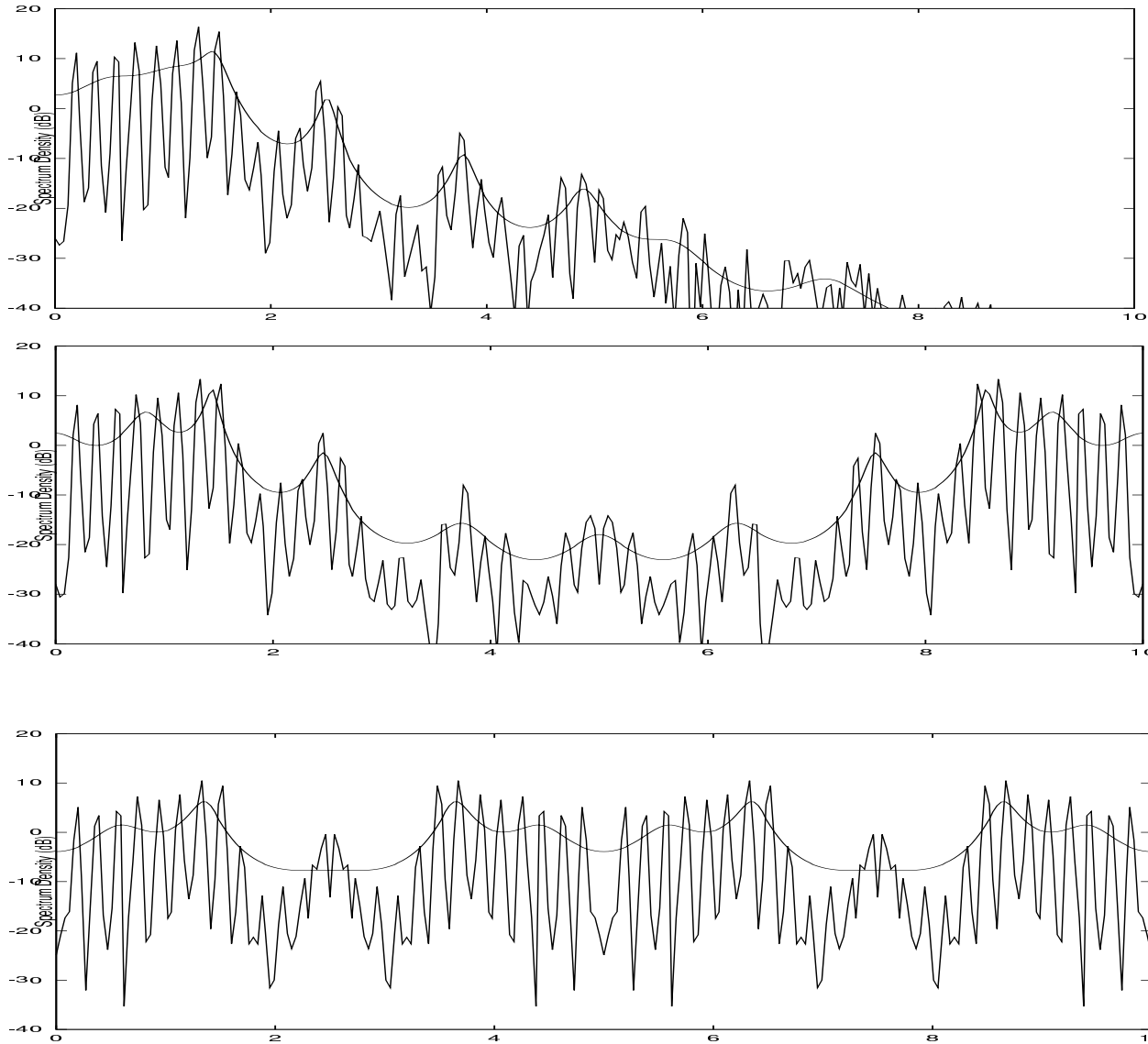
零交差波

$$y_t = \text{sgn } x_t$$

の同位置のスペクトル。



# エリアシング



男声朗読音声「来年の第一回通訳電話国際会議についてお伺いしたいんですけども、こちらでよろしいのでしょうか？」(mausc205:1.00s) 中の母音/a/のスペクトル

## 1/2ダウンサンプリング信号

$$y_t = \begin{cases} x_t, & t=0 \bmod 2 \\ 0, & \text{otherwise} \end{cases}$$

スペクトルが左右対称になる。

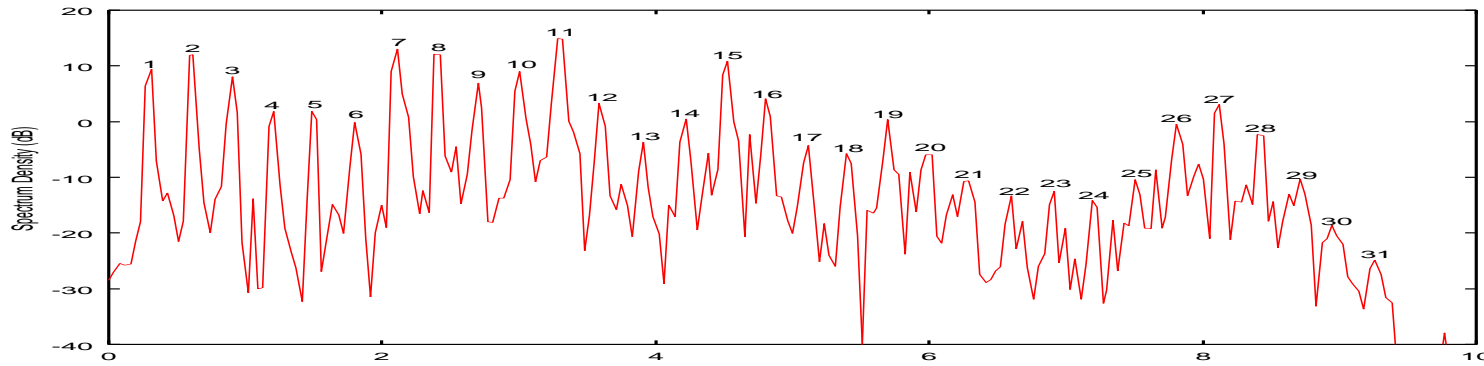
## 1/4ダウンサンプリング信号:

$$y_t = \begin{cases} x_t, & t=0 \bmod 4 \\ 0, & \text{otherwise} \end{cases}$$

スペクトルが左右対称反転する。

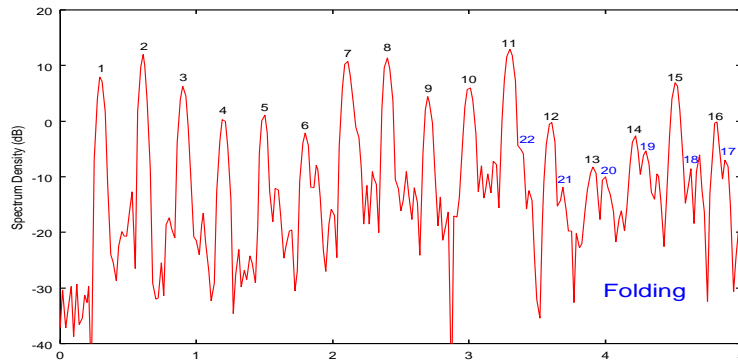


# 折り返し

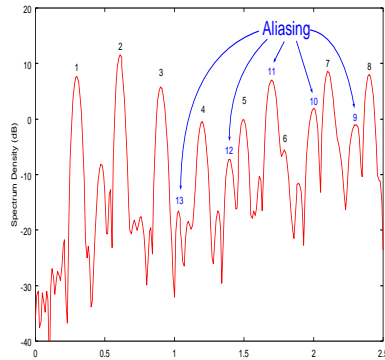


女声母音の倍音構造 (朗読音声中の/a/): 約 300Hz のピッチ周波数の整数倍の倍音が 31 個あまり観測できる。

内容: 「私は ATR 自動翻訳電話研究所の山田と申しますが。」 (fknsc202)



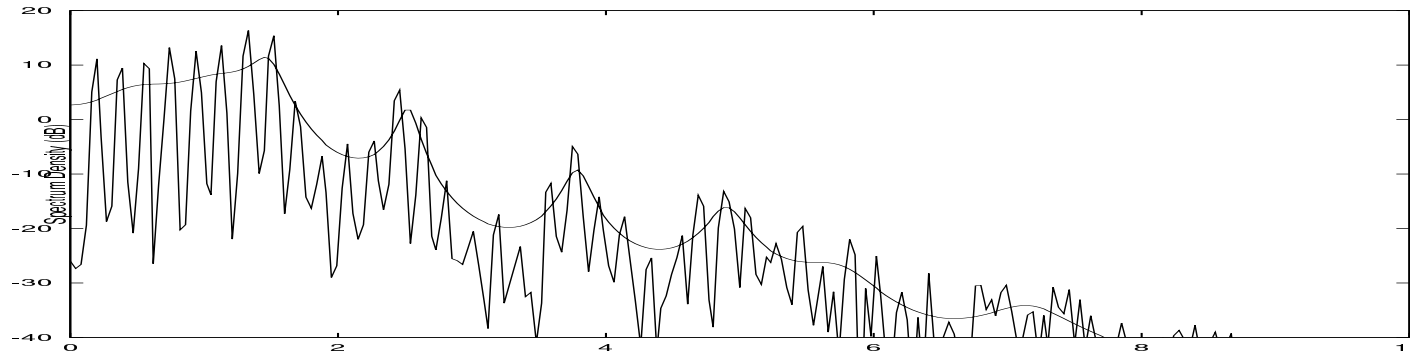
低域フィルタ無しの 10kHz ダウンサンプリング: ピッチ構造が折り返されるのが見られる。



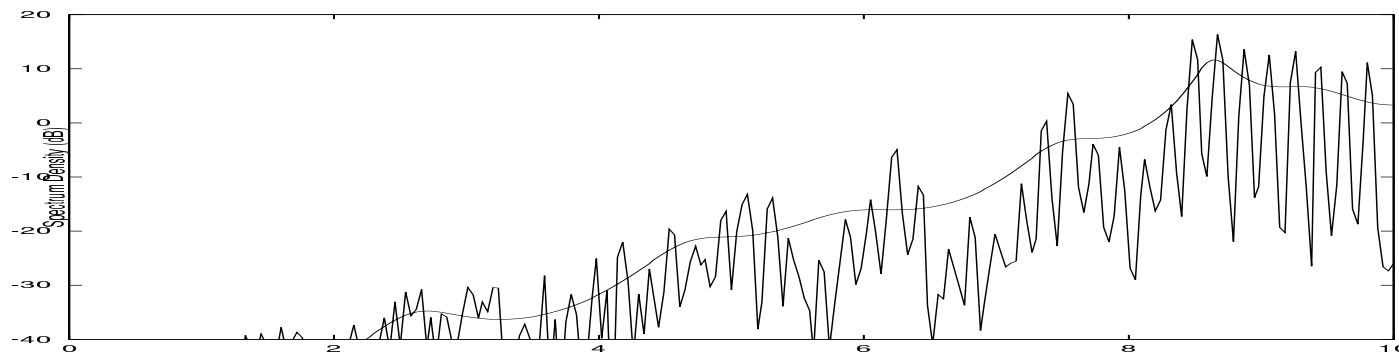
低域フィルタ無しの 5kHz ダウンサンプリング: ピッチ構造が折り返されるのが見られる。



# 例題：符号交互反転信号



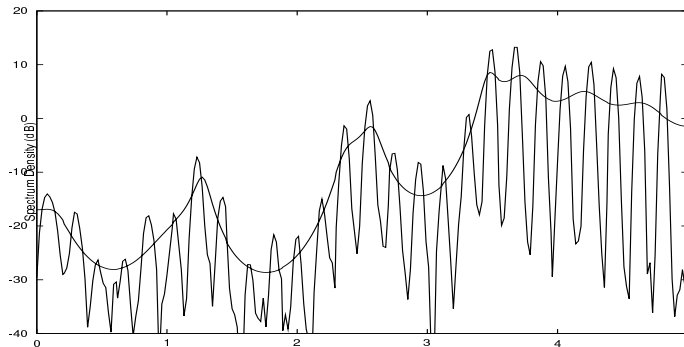
男声母音の倍音構造  
(朗読音声中の/a/)  
倍音構造が観測できる。



符号交互反転信号  
(10kHz 帯域):

$$y_t = (-1)^t x_t$$

スペクトルが左右対称反転する。



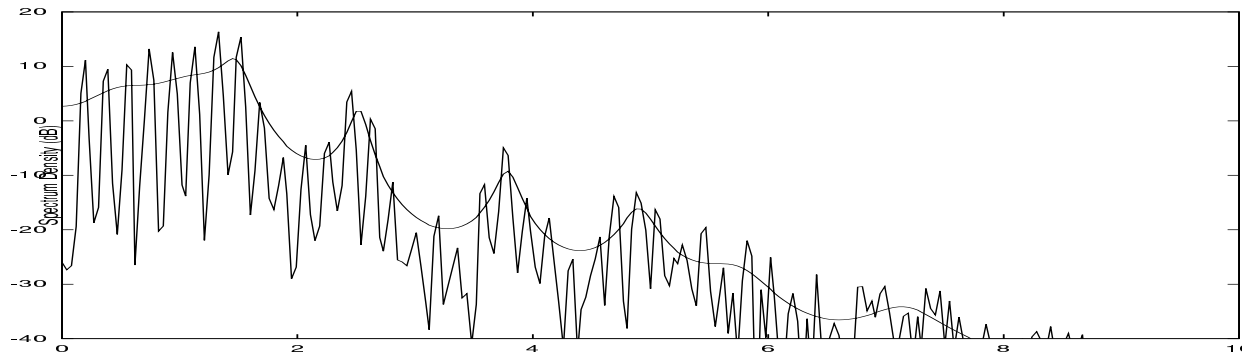
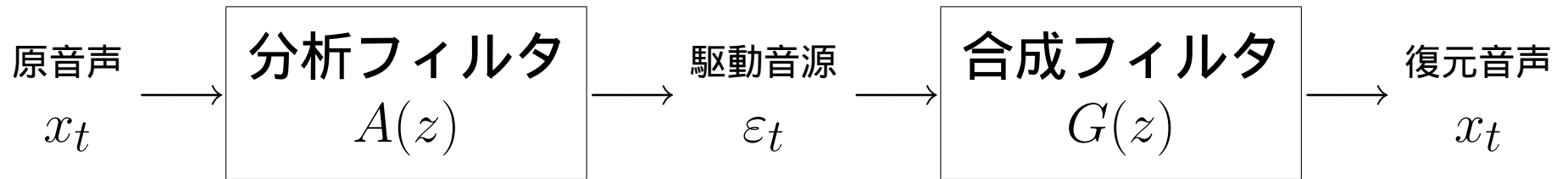
符号交互反転信号 (5kHz 帯域):

$$y_t = (-1)^t x_t$$

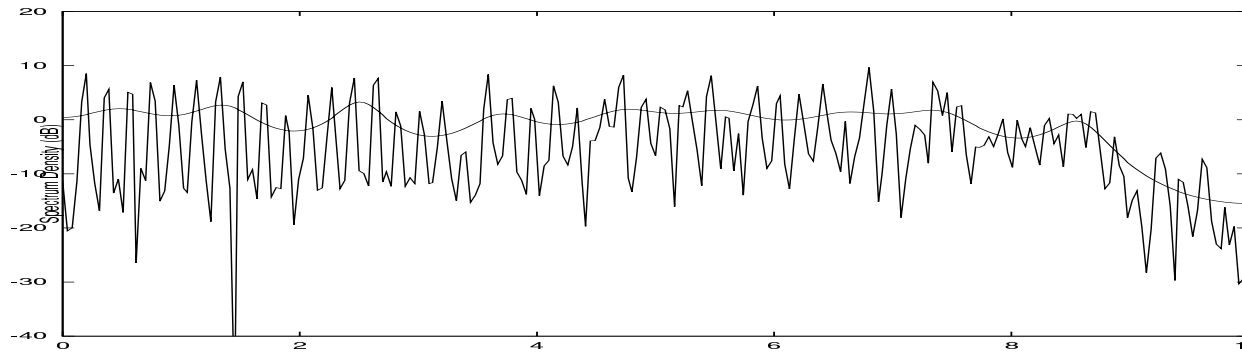
スペクトルが左右対称反転する。



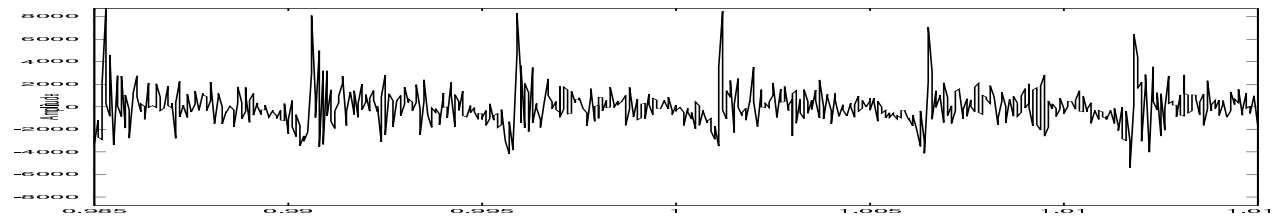
# LPC 残差信号



男声原音声「来年の第一回  
通話電話国際会議について  
お伺いしたいんですけども、  
こちらでよろしいんでしょうか？」  
(帯域 10kHz、20kHz サンプルング) 中の  
/a/(1.00S) のスペクトル



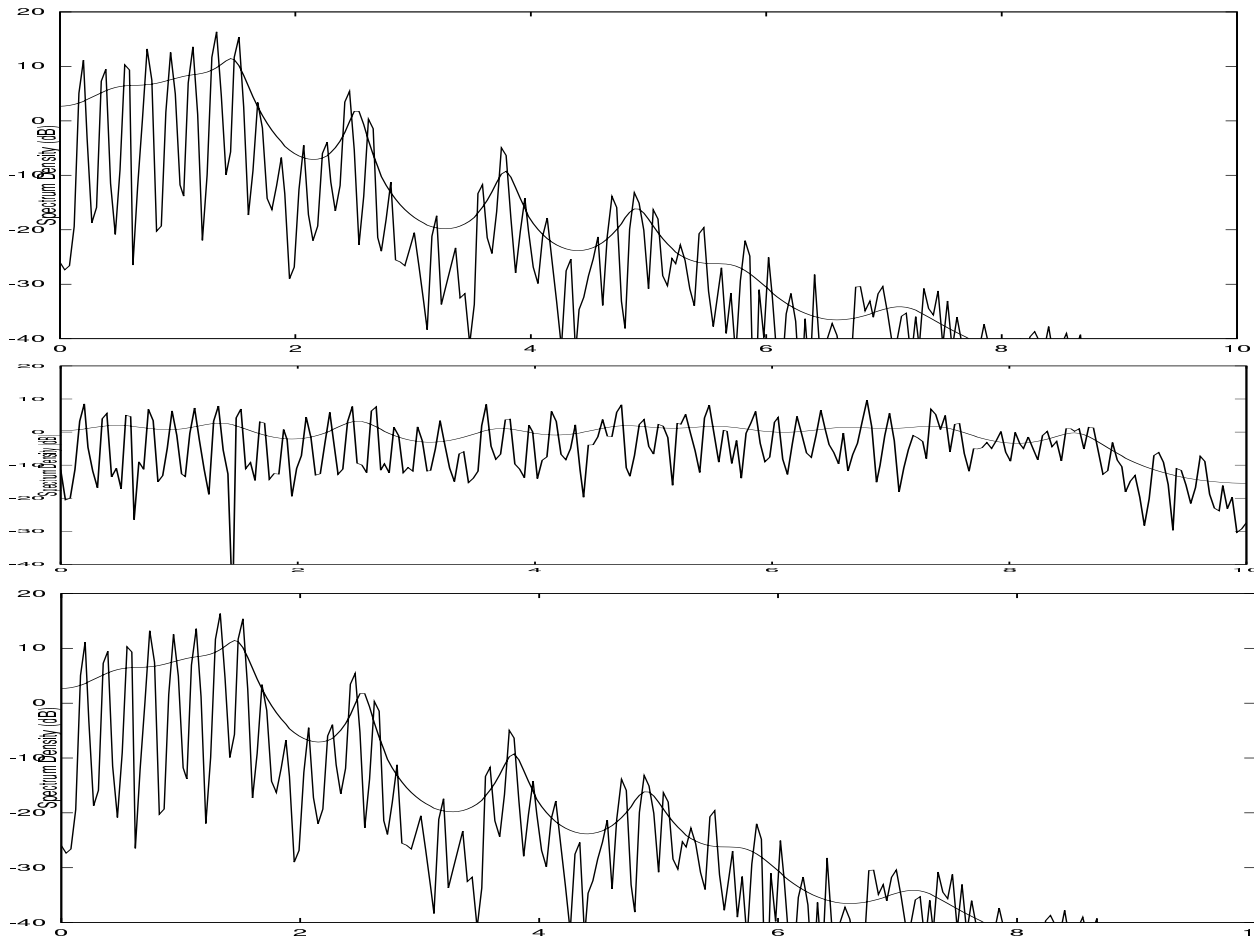
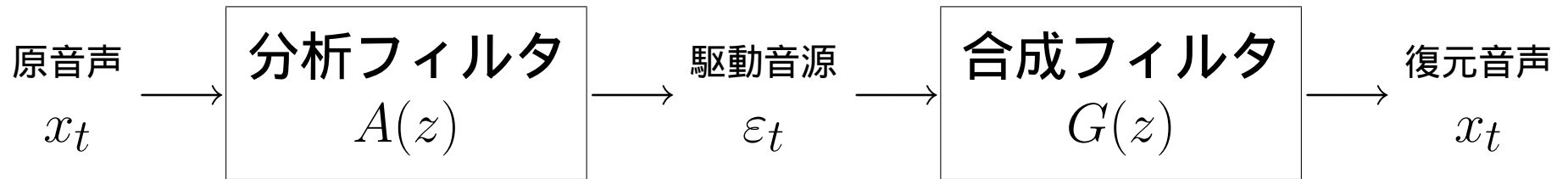
/a/ の 20 次 LPC 分析  
予測残差信号  $\varepsilon_t$  のスペクトル



/a/ の 20 次 LPC 分析予測  
残差信号  $\varepsilon_t$  の波形 (t=1.00S)



# LPC分析 残差駆動LPC音声合成



男声原音声「来年の第一回  
通話電話国際会議について  
お伺いしたいんですけども、  
こちらでよろしいんでしょうか？」  
(帯域 10kHz、20kHz サンプリング) 中の  
/a/(1.00S)のスペクトル

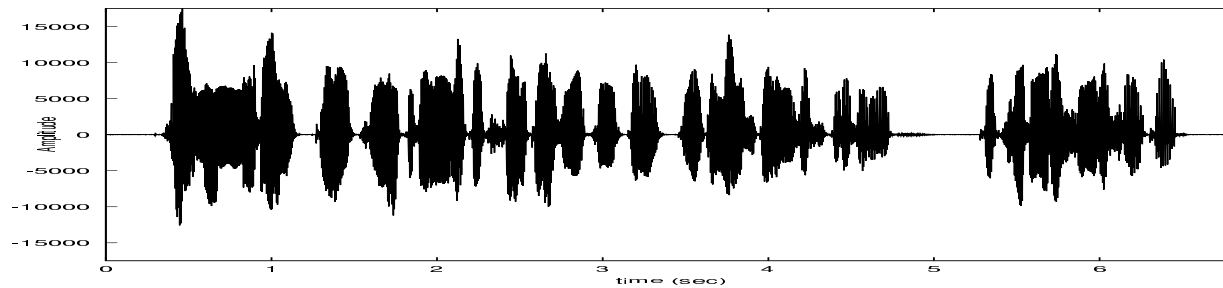
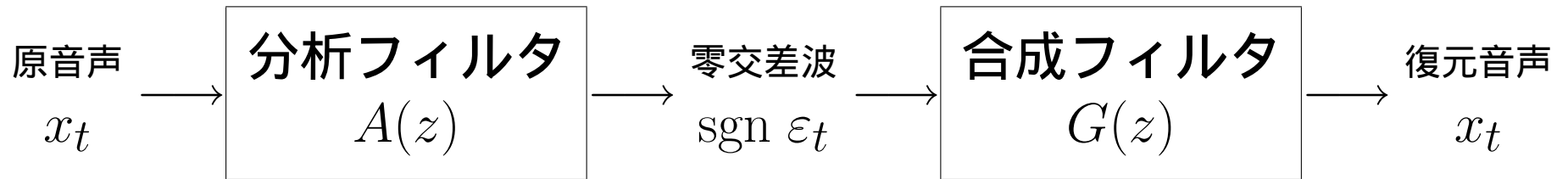
/a/の20次LPC分析  
予測残差信号  $\epsilon_t$  のスペクトル

復元音声の同箇所  
のスペクトル  
復元音声  $\epsilon_t$  のスペクトル

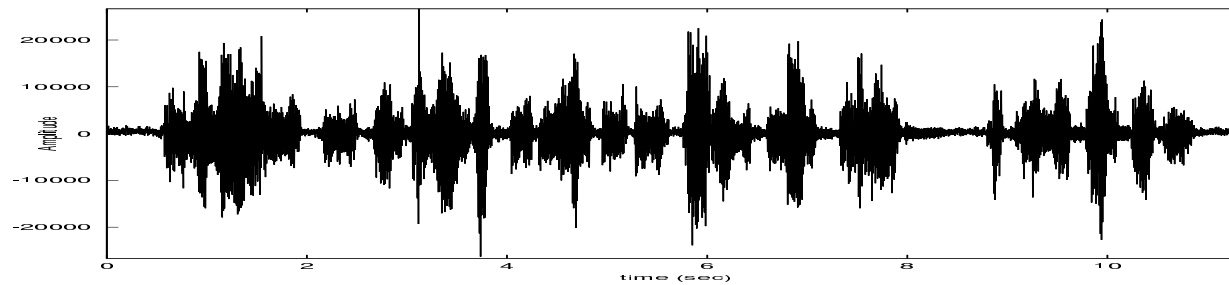




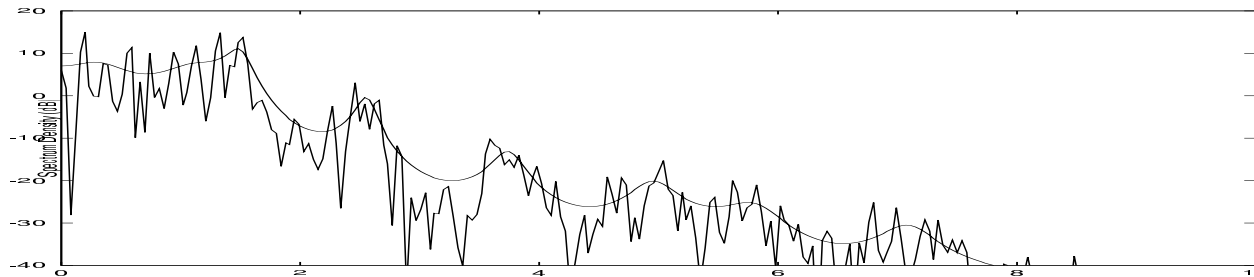
# LPC分析 零交差波駆動LPC音声合成



男声原音声「来年の第一回  
通話電話国際会議について  
お伺いしたいんですけども、  
こちらでよろしいんでしょうか？」  
(帯域 10kHz、20kHz サンプルング) の信号波形



**/a/の20次LPC分析**  
零交差波による合成信号の  
波形



零交差波による復元音声の  
**/a/のスペクトル**



# LPC 残差信号の自己相関

- 白色化された(周期)信号の自己相関関数 インパルス(列)
- LPC 残差  $\varepsilon_t$  の自己相関関数  $u_\tau$ :

$$\begin{aligned}
 u_\tau &= \mathcal{E}[\varepsilon_t \varepsilon_{t-\tau}] = \mathcal{E}\left[\sum_{i=0}^p a_i x_{t-i} \sum_{j=0}^p a_j x_{t-j-\tau}\right] \\
 &= \sum_{i=0}^p \sum_{j=0}^p a_i a_j \mathcal{E}[x_{t-i} x_{t-j-\tau}] = \sum_{i=0}^p \sum_{j=0}^p a_i a_j v_{i-j-\tau} \\
 &= \sum_{k=-p}^p A_k v_{k-\tau} \quad \text{但し} \quad A_{i-j} = \sum_{i=0}^p \sum_{j=0}^p a_i a_j
 \end{aligned}$$

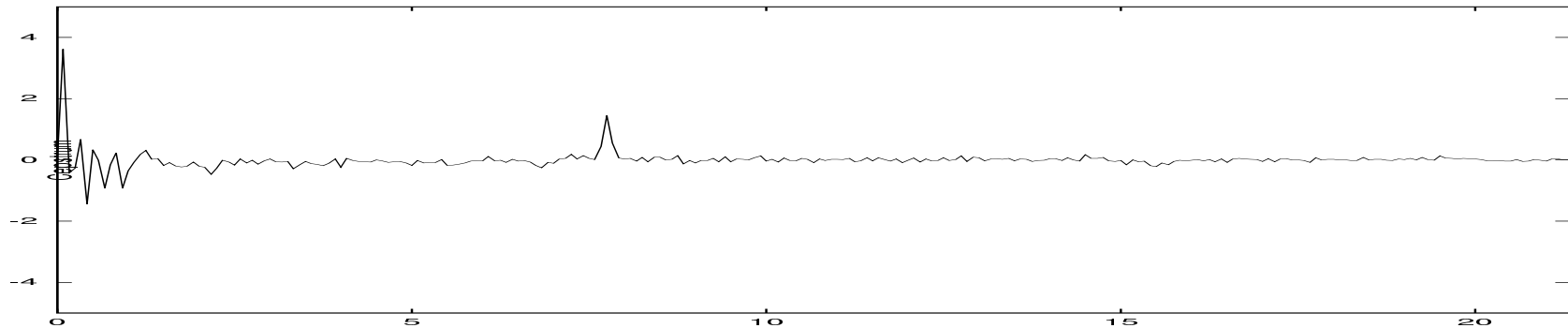
$v_\tau$  は信号の自己相関、 $A_k$  は線形予測係数の自己相関

$$\varepsilon = a * x \quad \Rightarrow \quad u = A * v$$



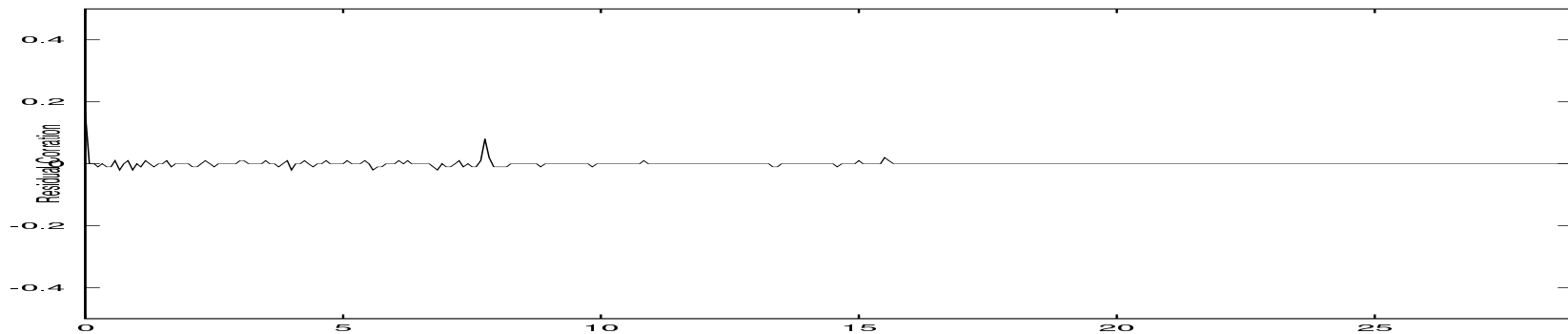
# ピッチ抽出

## ■ ケプストラム法



音声短時間ケプストラム /a/ (男声(MAU)「こしらえる」520mS付近)

## ■ 残差相関法 (変形相関法)



音声LPC残差相関 /a/ (男声(MAU)「こしらえる」520mS付近)