



東京大学 工学部 計数工学科/物理工学科

応用音響学：音声分析 (5) LSP分析

嵯峨山 茂樹 <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/>

■ 概要

- 線スペクトルモデルによる音声分析
- 声道モデルと声門境界条件
- 線スペクトル対(LSP)音声分析合成定理
- 線スペクトル対(LSP)音声分析アルゴリズム
- LSP周波数の性質
- LSP音声合成フィルタ



線スペクトルモデルによる音声分析

- 目的: LPC ,PARCOR よりもさらによい音声スペクトルパラメータ

- 高い圧縮率 (安定性の保証、低いスペクトル感度、よい復元特性)

- 線形予測係数の線スペクトル表現 (板倉 1975)

- 声道モデルにおいて声門反射係数を ± 1 に設定 無損失条件

- 線スペクトル周波数 λ_i と強度 m_i

- 線スペクトル表現からの線形予測係数復元
自己相関関数の復元:

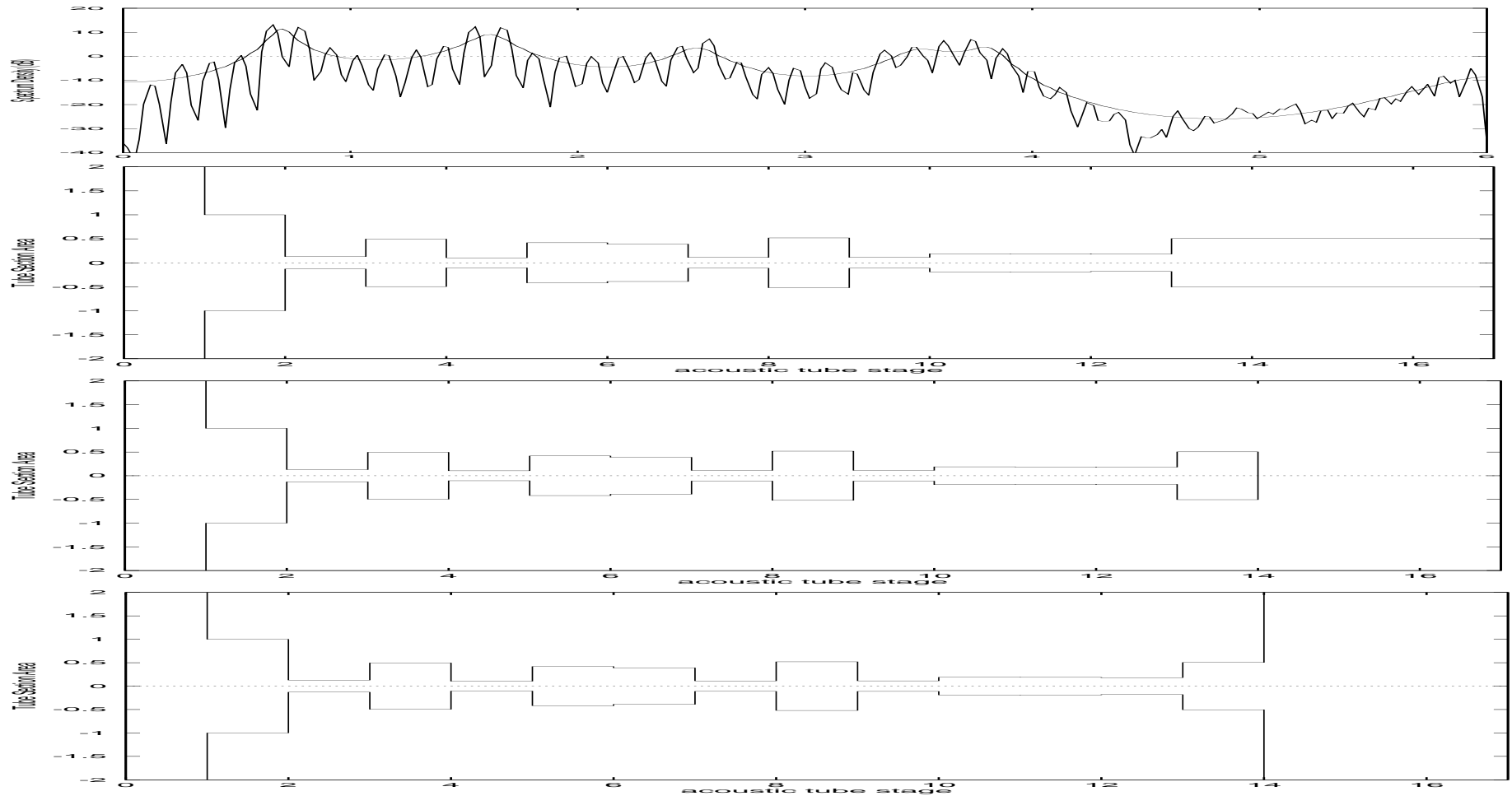
$$v_\tau = \sum_{i=1}^n m_i \cos \tau \lambda_i, \quad \tau = 0, 1, 2, \dots, 2n - 1$$

LPC or PARCOR 係数復元

- のちに複合正弦波モデル(CSM)として直交多項式理論で統一
 - 線スペクトル周波数は直交多項式の根
 - 線スペクトル強度はその Chistoffel 数



等価音響管(声道モデル)と声門の境界条件



図：音声のスペクトルと等価音響管および声門の境界条件 (男声 /a/)



無損失音響管の極周波数: LSP 分析の原理

PARCOR 漸化式:

$$A_{p+1}(z) = zA_p(z) - k_{p+1}z^p A(z^{-1})$$

において, $k_{p+1} \rightarrow \pm 1$ (無損失条件)

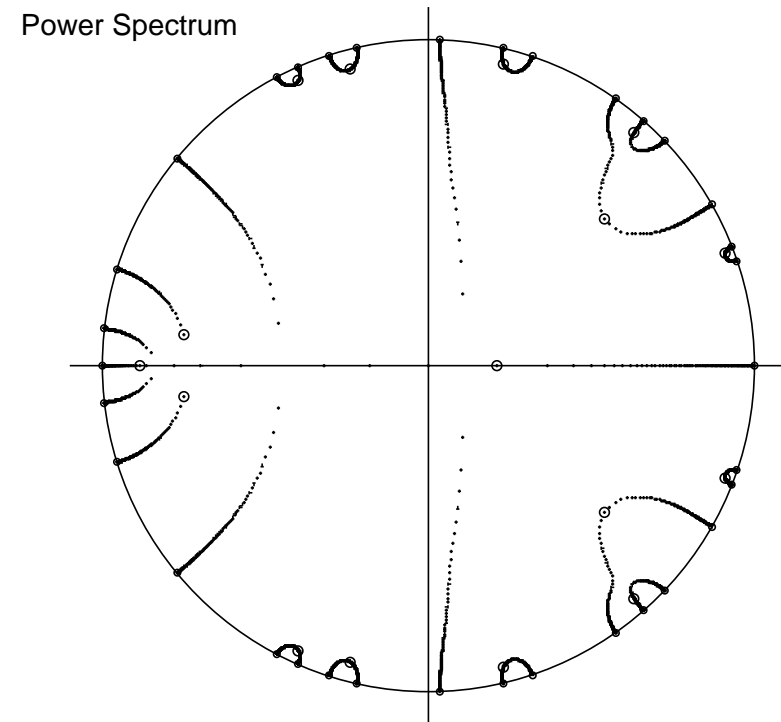
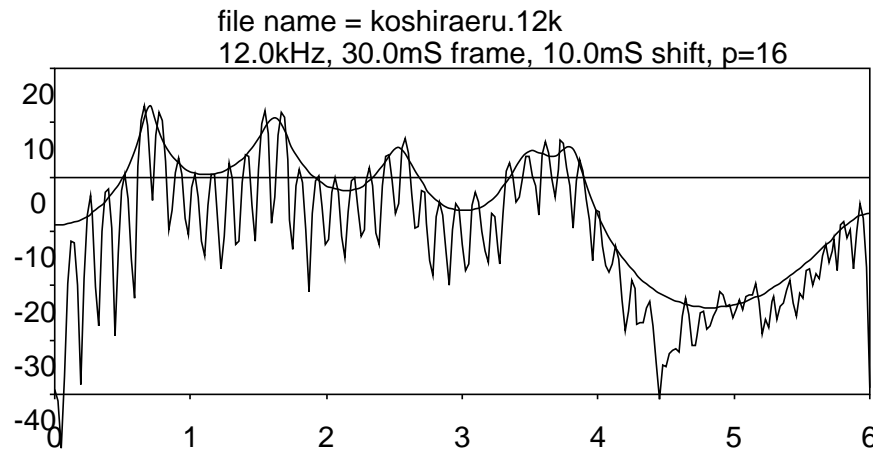


図1. LSP 分析の原理 — PARCOR 漸化式の $-1 < k_{p+1} < 1$ を, ± 1 に近付けることにより, $(p + 1)$ 次の LPC 多項式の根は, 単位円に近づく.



線スペクトル対 (LSP) 分析合成定理 (1979, 1982)

■ LSP 分析定理

「多項式 $A_p(z)$ のすべての根が単位円内にあるならば、

$$A_{p+1}^{\pm}(z) = zA_p(z) \mp z^p A(z^{-1})$$

の根は単位円上で交番する。」 (1976, 1979 [実係数], 1982 [複素])

LSP 周波数:

$$0 < \omega_1 < \omega_2 < \dots < \omega_p < \pi$$

をパラメータとする音声スペクトルの表現。(フィルタ安定性の必要十分条件)

■ LSP 合成定理

$A_{p+1}^+(z)$ と $A_{p+1}^-(z)$ の根が単位円上で交番するならば、

$$A_p(z) = \frac{A_{p+1}^+(z) + A_{p+1}^-(z)}{2z}$$

の根はすべて単位円内に存在。 (1976 [実係数], 1982 [複素])

$A_p(z)$ を復元。



LSP 音声分析例

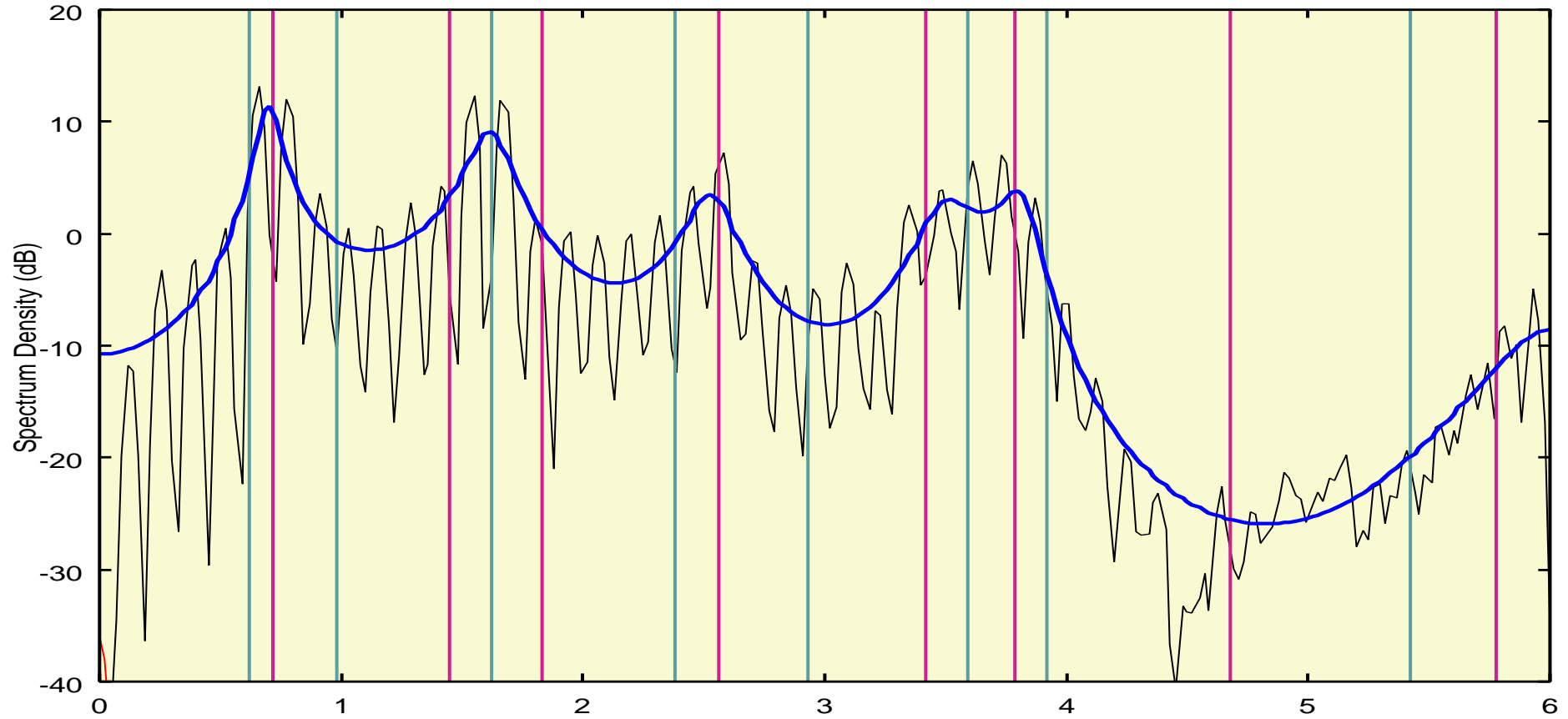
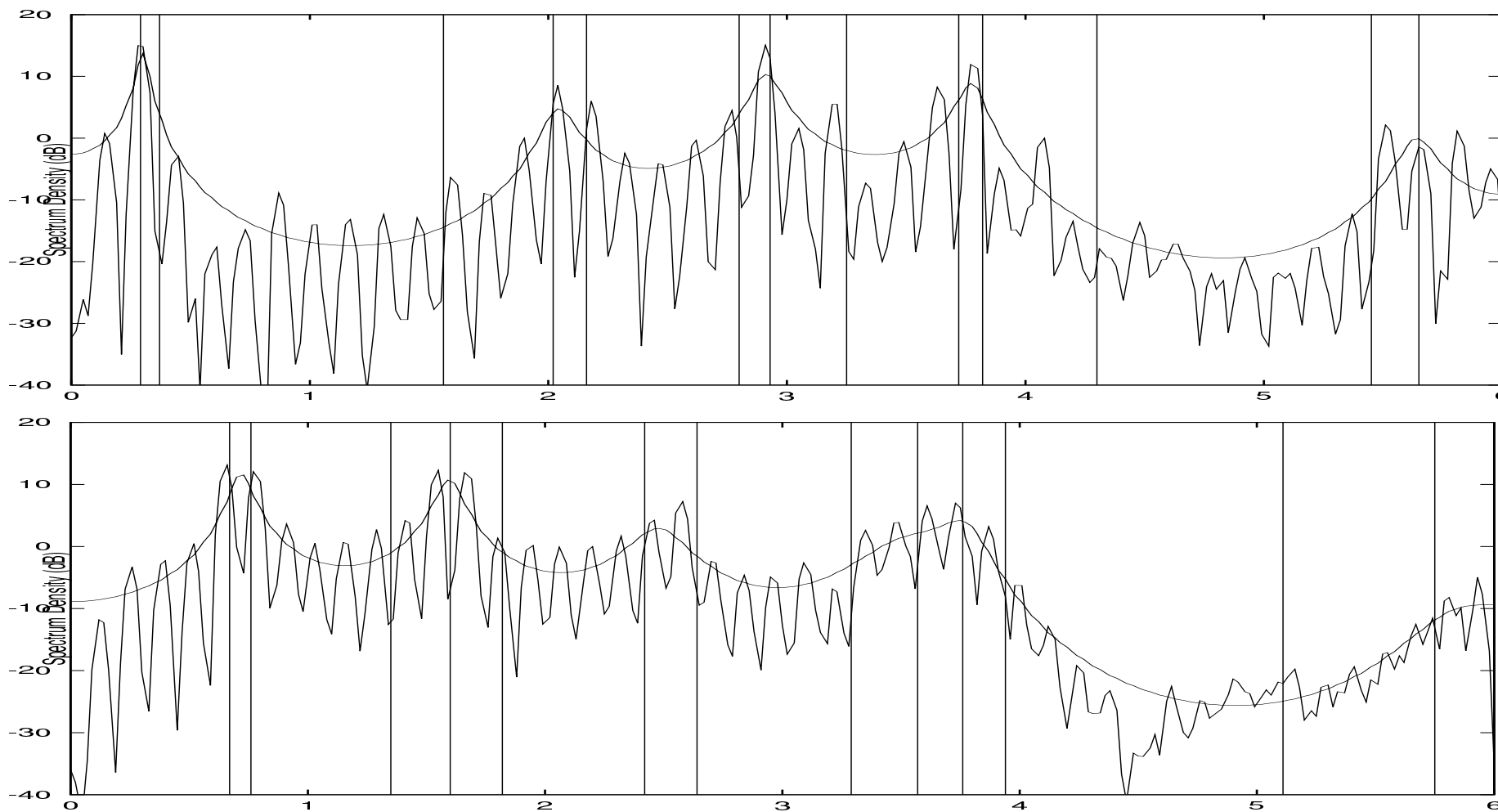


図. LSP 周波数の分布例 (13次LSP分析、男声 /koshiraeru/ 中の /a/)

共振周波数付近に LSP 周波数が集中する傾向があることがわかる。



音声の短時間スペクトルとLSP周波数の例 (/i/ vs /a/)



短時間スペクトルと LSP 周波数の例 (男声 /i/ と /a/)



一般化 LSP 分析/合成定理

■ 複素信号 LPC (信号: 複素信号)

非対称正スペクトル, 正定値 Hermite 型自己相関, 複素 PARCOR 係数 $|k_i| < 1$

■ 実信号 LSP (信号: 実信号, PARCOR 係数: $k_{p+1} \rightarrow \pm 1$) [Schüssler 1976]

$A_p(z)$ の根が単位円上に移動. 共役根の偏角が LSP 周波数.

$$\omega_0 = 0, \omega_1, \dots, \omega_p, \omega_{p+1} = \pi, \omega_{p+2} = \omega_p, \dots, \omega_{2p+1} = \omega_1$$

■ 複素信号 LSP (信号: 複素, PARCOR 係数: $k_{p+1} \rightarrow \pm 1$)

$2p$ 個の LSP 周波数は非共役

$$\omega_0 = 0, \omega_1, \dots, \omega_p, \omega_{p+1} = \pi, \omega_{p+2}, \dots, \omega_{2p+1}$$

■ 一般化 LSP (信号: 複素, PARCOR 係数: k_{p+1} を任意の $|k_{p+1}| = 1$ へ)

$2p + 2$ 個の LSP 周波数は非共役. $\omega_0, \dots, \omega_{2p+1}$

■ いずれの場合も, 全極フィルタが安定であるための必要十分条件は, LSP 周波数の順序関係が保たれること. [嵯峨山 1982]



LSP 分析手順 (方法 1)

(1) p 次 LPC 分析 $A_p(z)$ を求める

(2) $A_{p+1}^{\pm}(z) = zA_p(z) \mp z^p A(z^{-1})$ の根(すべて単位円上)のうち実根を除去

(3) 残った複素共役根を実軸に射影

■ $x = \frac{z + z^{-1}}{2}$ により置き換え
(「Chebyshev 展開」)

(4) 2 個の x の n 次方程式を Newton-Raphson 法で交互に解く

■ 根は区間 $[-1,1]$ 内で交互に存在
極めて高速に求まる。

(5) 根の \arccos 値を求める。

単純化した LSP 分析プログラム例 ($p=偶数$)

```
void alflsp( int p, double alf[], double fq[]
) {
    int i,j,km,kp,n;
    double x,opm[100],opp[100],acos();
    static double tbl[2600],eps=0.0000001;
    chebtbl(tbl,(p+1)/2); /* Chebyshev 変換表作成
*/
    n=p/2;
    opp[0]=opm[0]=1.0;
    for(i=1;i<=n;i++) {
        opp[i]=alf[i]-alf[p+1-i]+opp[i-1];
        opm[i]=alf[i]+alf[p+1-i]-opm[i-1];
    }
    excheb(n,opp,opp,tbl); ex-
    cheb(n,opm,opm,tbl);
    kp=n; km=n; j=0; x=1.0;
    for(i=1;i<=p;i++) {
        if(j=1-j) { nrstep(opm,km,eps,&x); km--; }
        else { nrstep(opp,kp,eps,&x); kp--; }
        fq[i]=acos(fq[i]);
    }
}
```

chebtbl(): Chebyshev 展開係数表を作成

excheb(): Chebyshev 展開を実行

nrstep(): Newton-Raphson 法の 1step を実行

(実際は: 奇数次の場合を含める。Chebyshev 展開係数表作成は 1 回のみにする。収束しないときは条件を自動的に緩める。)



LSP 分析アルゴリズム (板倉 1979)

音声波形 $\{Y_1, Y_2, \dots, Y_N\}$

標本自己相関関数計算 $\{v_l\}$

$$v_l = \frac{1}{N} \sum_{t=l}^{N-1} Y_t Y_{t-l}, \quad v_{-l} = v_l, \quad l = 0, 1, \dots, 2n - 1$$

p 次 正規方程式求解 $\{a_i\}$

$$\begin{pmatrix} v_0 & v_1 & \cdots & v_{p-1} \\ v_1 & v_0 & \cdots & v_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p-1} & v_{p-2} & \cdots & v_0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = - \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{pmatrix}$$

$p = 2n$ (偶数) の場合

$p = 2n + 1$ (奇数) の場合

無損失条件の実軸への射影 $P_n^-(x), P_n^+(x)$

- 境界条件 $k_{p+1} = +1$: $P_n^-(x) = \frac{zA_{2n}(z) - z^{2n}A_{2n}(z^{-1})}{(2z)^n(z-1)}$
- 境界条件 $k_{p+1} = -1$: $P_n^+(x) = \frac{zA_{2n}(z) + z^{2n}A_{2n}(z^{-1})}{(2z)^n(z+1)}$

無損失条件の実軸への射影 $P_n^-(x), P_{n+1}^+(x)$

- 境界条件 $k_{p+1} = +1$: $P_n^-(x) = \frac{zA_{2n+1}(z) - z^{2n+1}A_{2n+1}(z^{-1})}{(2z)^{n+1}(z^2-1)}$
- 境界条件 $k_{p+1} = -1$: $P_{n+1}^+(x) = \frac{zA_{2n-1}(z) + z^{2n-1}A_{2n-1}(z^{-1})}{(2z)^n}$

n 次代数方程式求根 $\{x_i\}$

$P_n^-(x) = 0, P_n^+(x) = 0$
 $2n$ 個の根を $x_1 > \dots > x_p$ とする。

$n, n + 1$ 次代数方程式求根 $\{x_i\}$

$P_n^-(x) = 0, P_{n+1}^+(x) = 0$
 $(2n + 1)$ 個の根を $x_1 > \dots > x_p$ とする。

逆 cos 変換 $\{\omega_i\}$

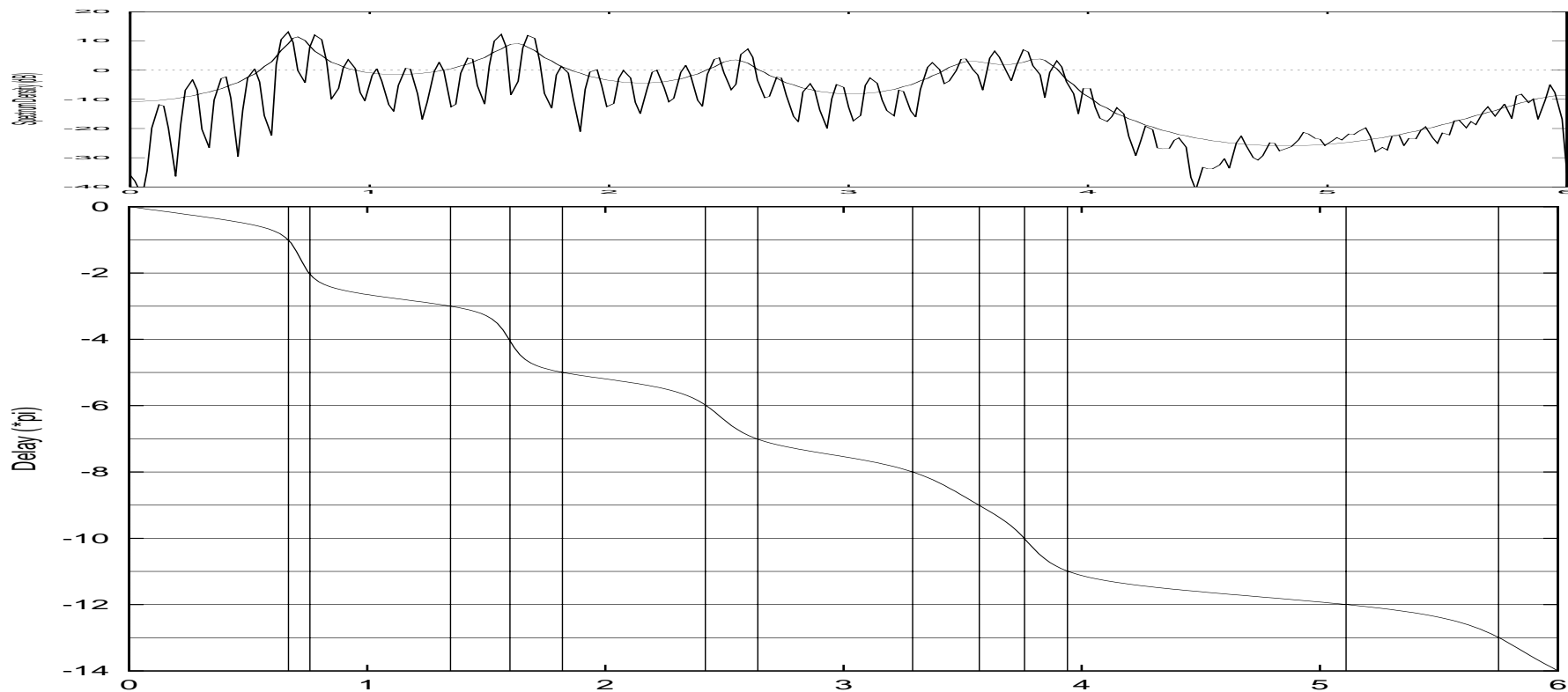
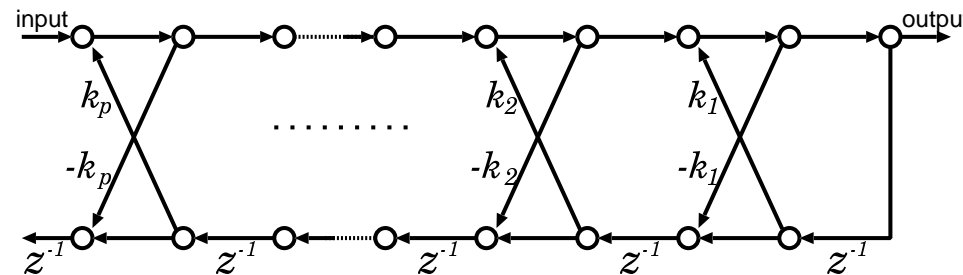
$$\omega_i = \cos^{-1} x_i, \quad i = 1, \dots, p$$

LSP 周波数 $\{\omega_1, \omega_2, \dots, \omega_p\}$



PARCOR 全帯域通過フィルタの位相遅れ特性とLSP

PARCOR 全帯域通過フィルタの位相遅れ特性が π の整数倍になる周波数はLSP
周波数。





1 次 ~ 16 次 LSP 周波数の例 (男声 / a /)

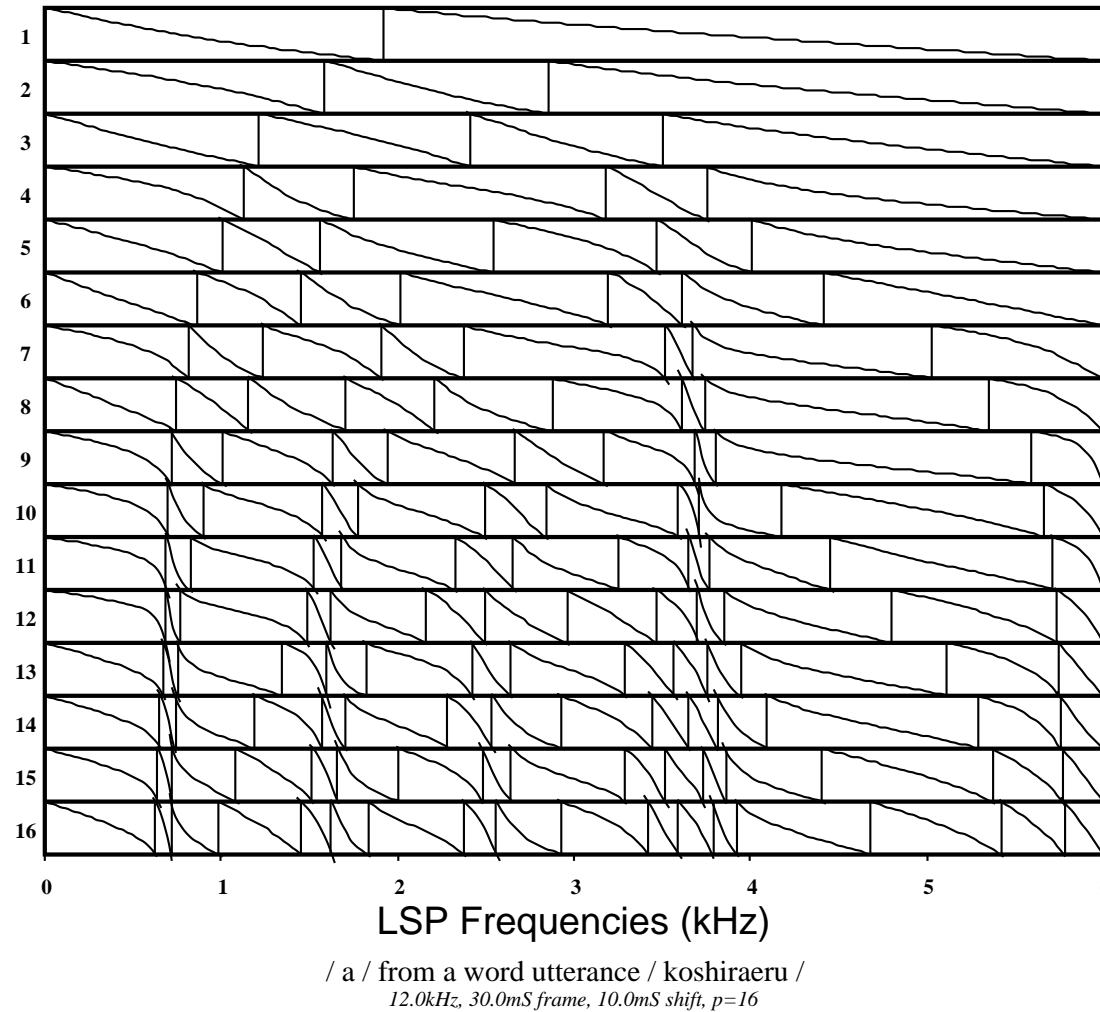


図2. LSP 分析例 — 位相遅れが π の整数倍になる周波数が LSP 周波数である .



LSP 周波数の時間パターン例

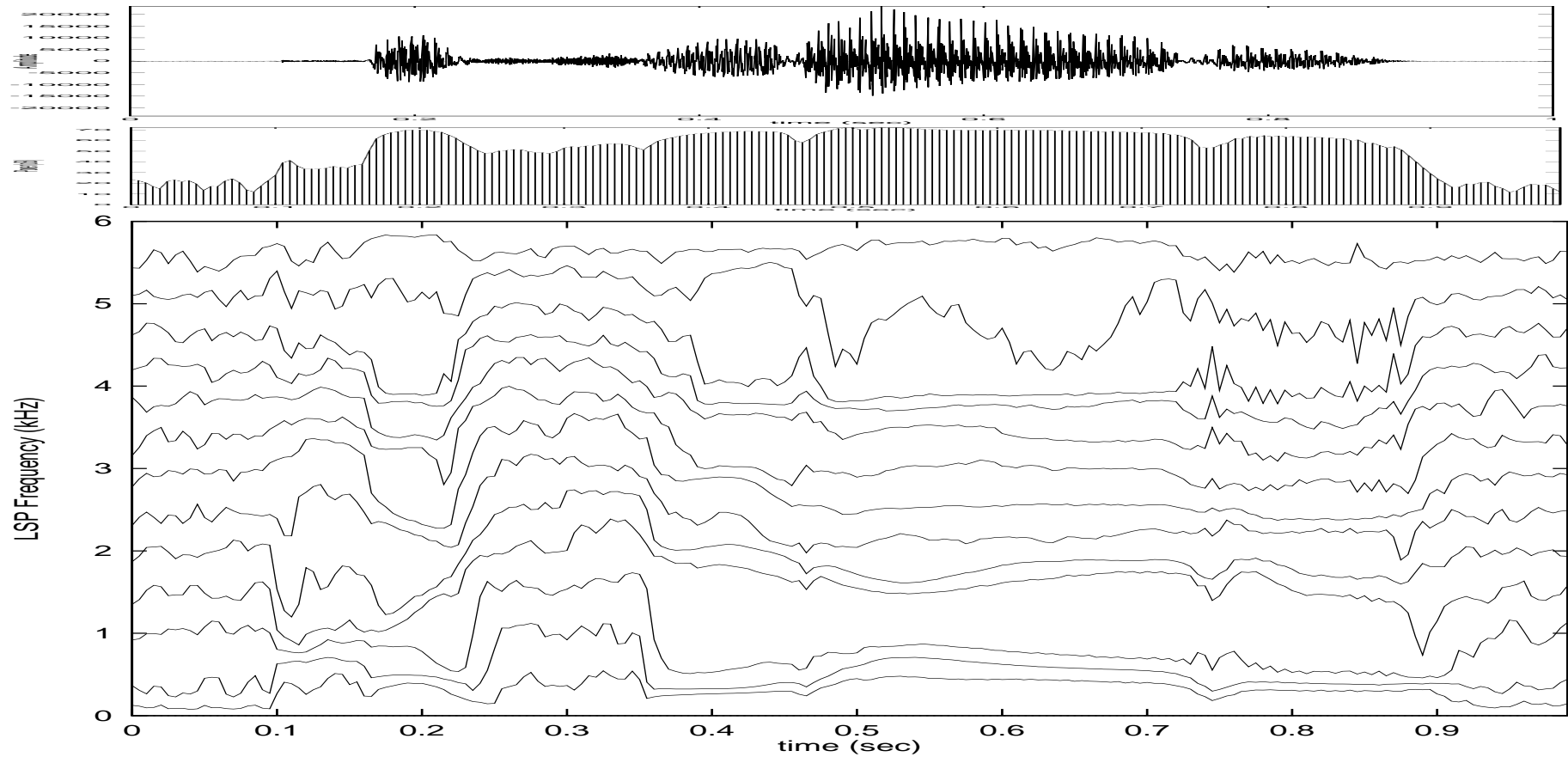


図: LSP 周波数の時間パターン例 (「こしらえる」)
時間方向に滑らかに動くため, 補間特性が優れている。



線形予測極の時間パターン例 (比較)

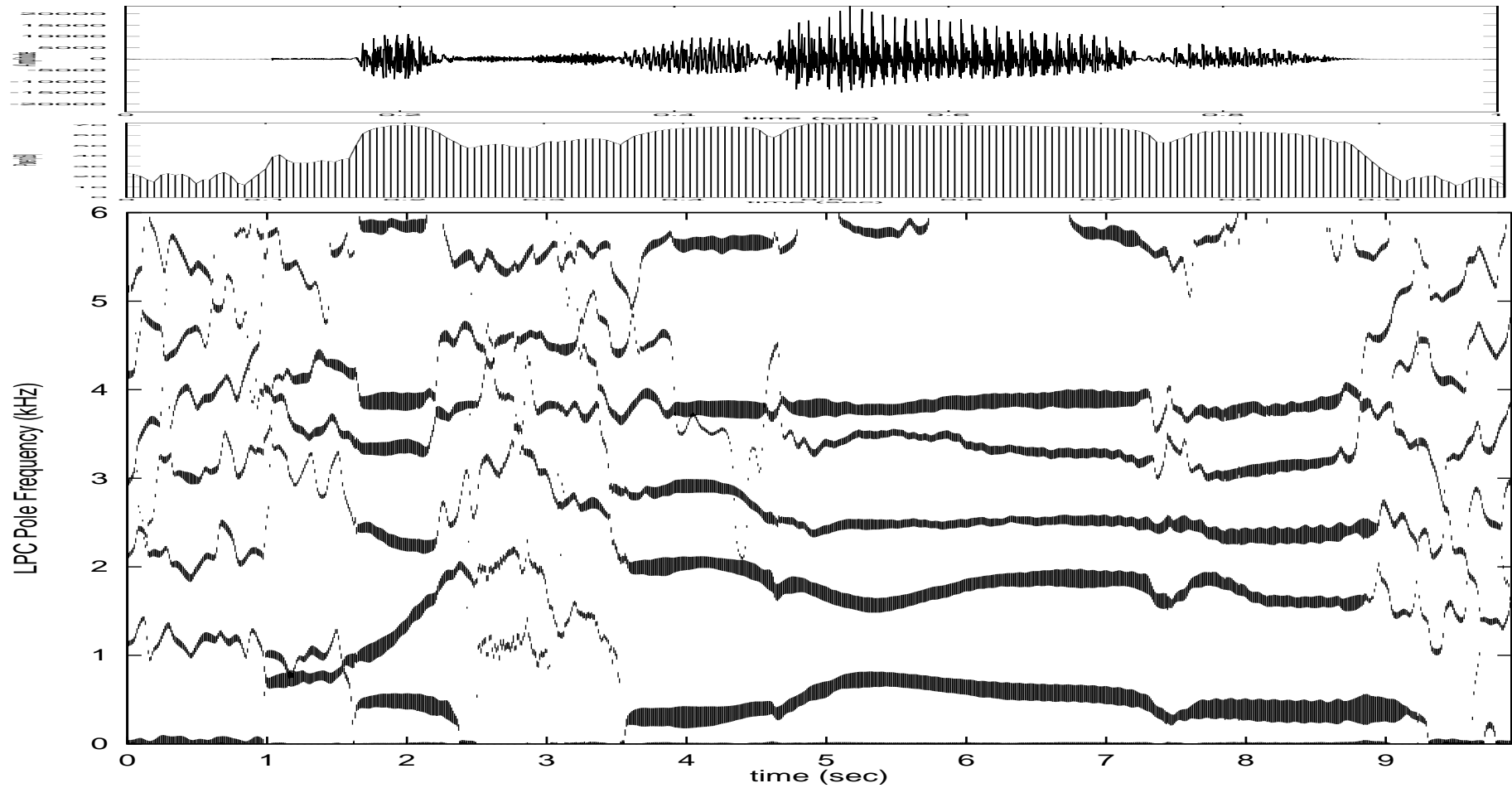
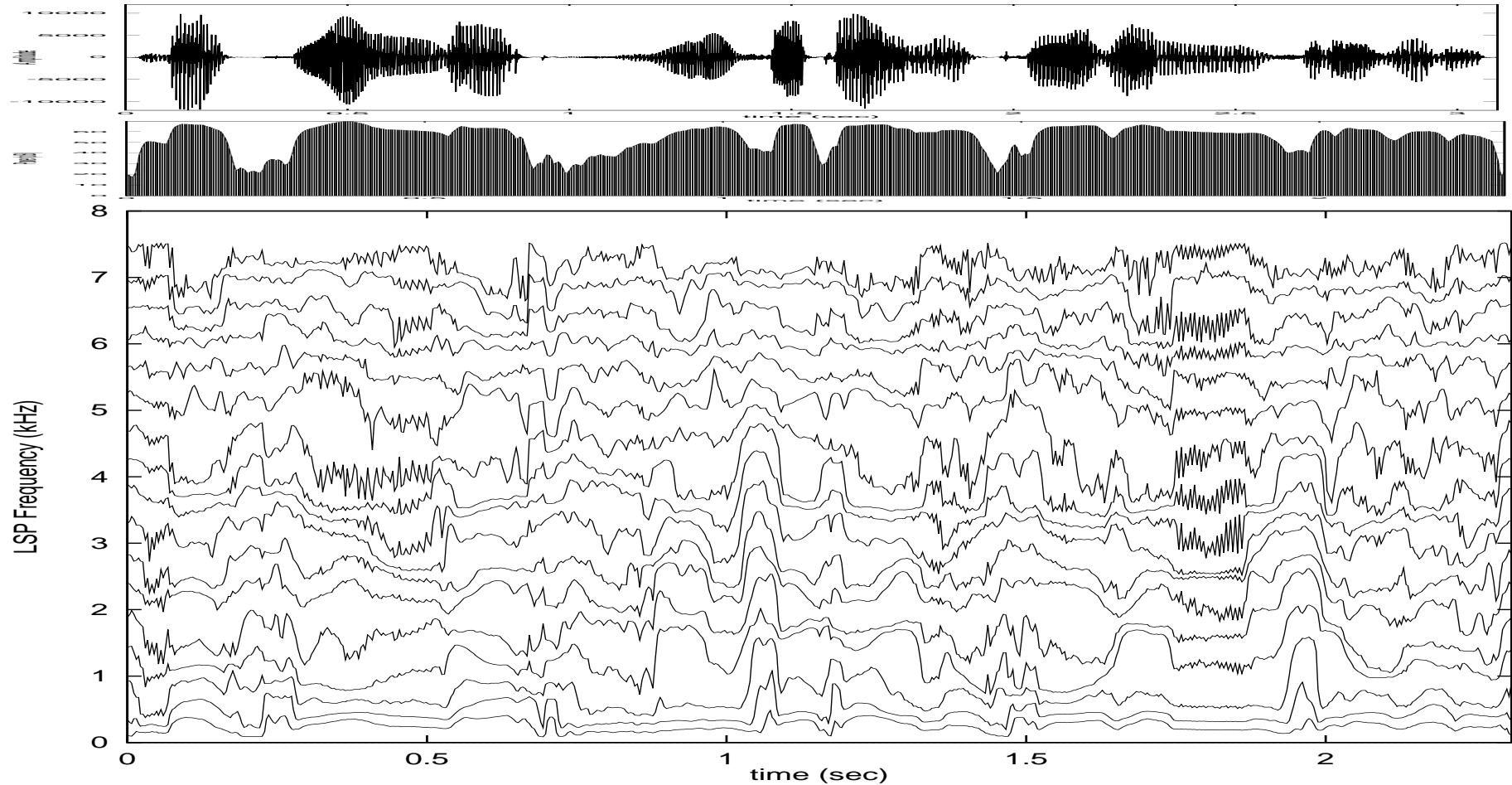


図: 音声の 13 次線形予測極の時間パターン例 (「こしらえる」)



LSP 周波数の時間パターン例



図：音声の LSP 周波数の時間パターン例 (男声「爆音が銀世界の高原に広がる」)



LSP 周波数の時間パターン例

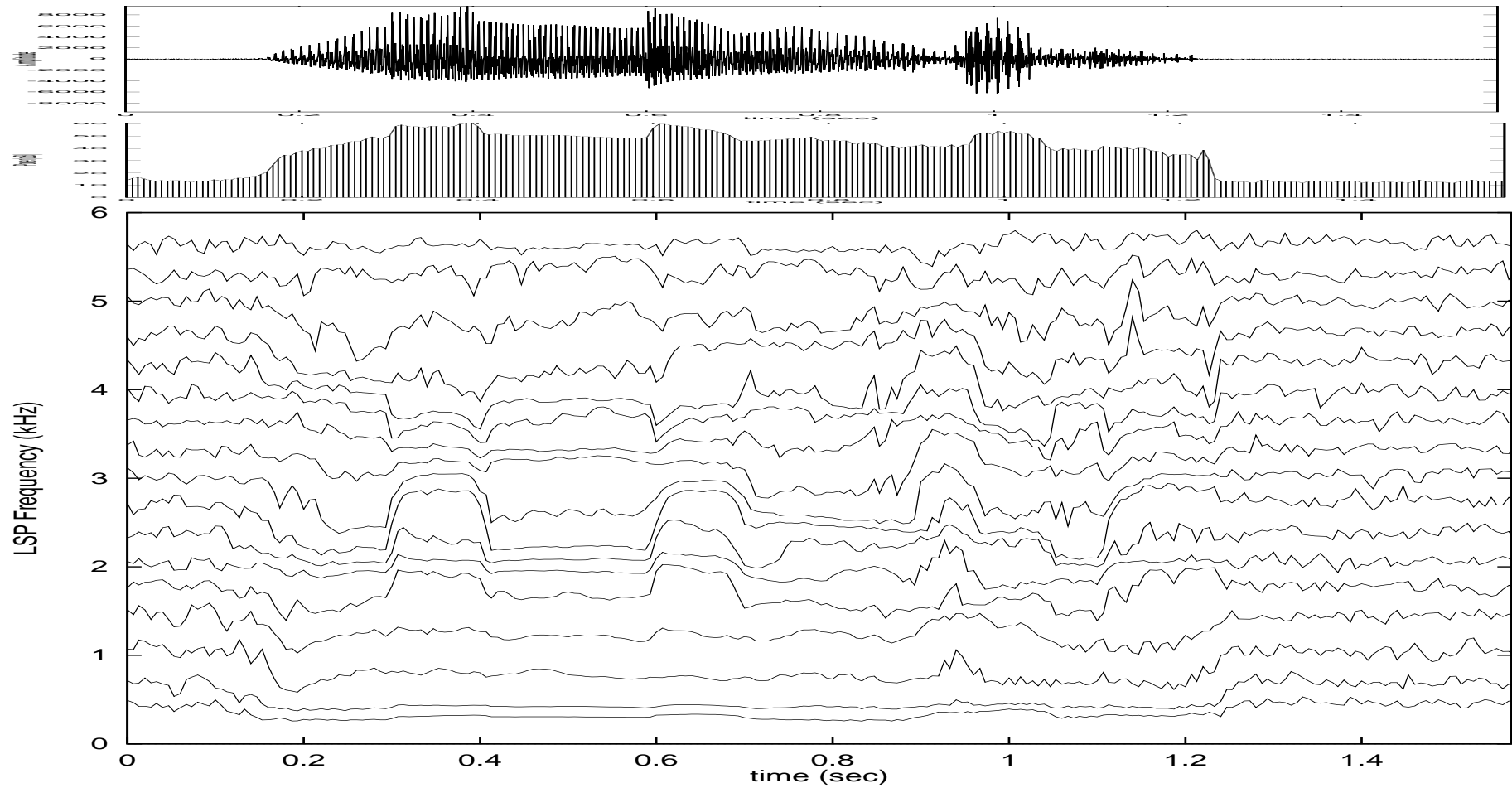


図: LSP 周波数の時間パターン例 (「みんなぜみ」)



LSP 周波数の特性

- 周波数領域のパラメータ 扱いやすい
- 安定性の必要十分条件: $0 < \omega_1 < \omega_2 < \dots < \omega_p < \pi$
- 時間方向に滑らか (直線補間が有効): 更新周波数を **PARCOR** の 75% にできる
- 量子化特性がよい (4 bit/parameter): **PARCOR** の場合の約 80% の情報量
- 総合的に **PARCOR** の 60% の情報量で同等の音声品質
- 多くの音声符号化に利用



LSP 音声合成フィルタの導出

p =偶数の場合、

$$c_i = -2 \cos \omega_i, \quad i = 1, 2, \dots, p$$

$$c_{-1} = c_0 = -z^{-1}$$

と書けば、

$$A_p(z^{-1}) - 1$$

$$= \frac{1}{2} \{A_p^+(z^{-1}) + A_p^-(z^{-1})\}$$

$$= \frac{z^{-1}}{2} \left[\sum_{i=2,4,\dots,p} (c_i + z^{-1}) \prod_{j=0,2,\dots,i-2} (1 + c_j z^{-1} + z^{-2}) - \prod_{j=2,4,\dots,p} (1 + c_j z^{-1} + z^{-2}) \right.$$

$$\left. + \sum_{i=1,3,\dots,p-1} (c_i + z^{-1}) \prod_{j=-1,1,\dots,i-2} (1 + c_j z^{-1} + z^{-2}) - \prod_{j=1,3,\dots,p-1} (1 + c_j z^{-1} + z^{-2}) \right]$$

により、**LSP 全極型フィルタ**が実現できる。



LSP 音声合成フィルタ (偶数次)

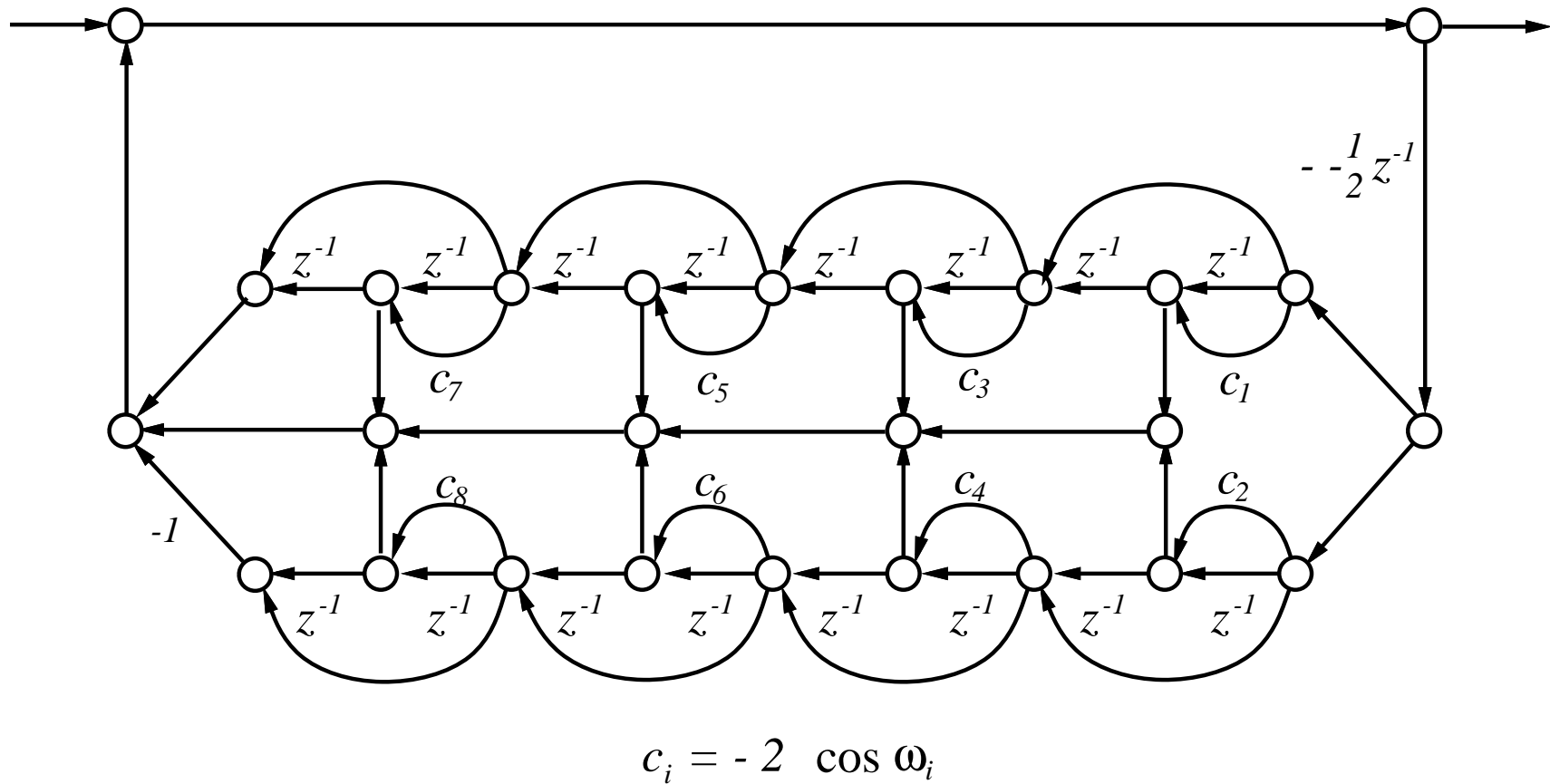
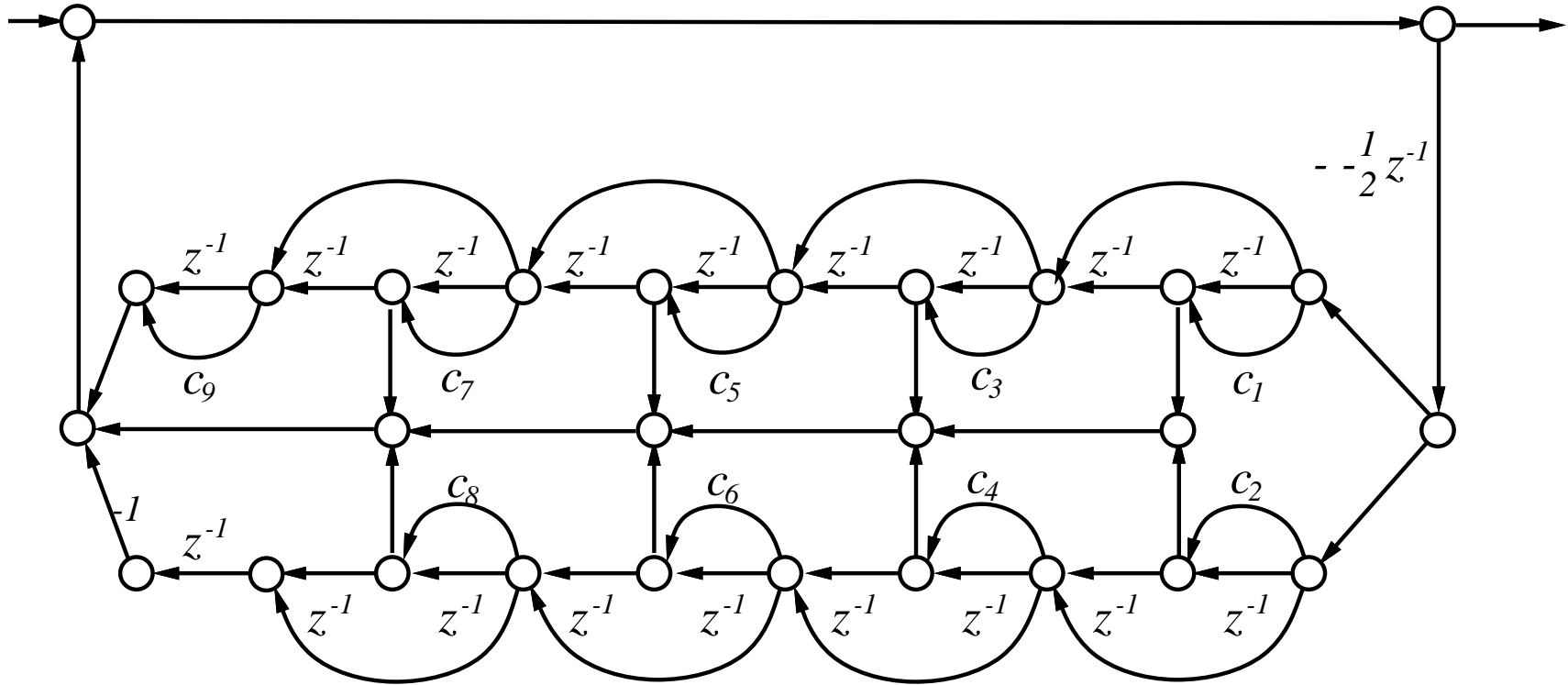


図3. LSP 周波数の \cos 値を係数に持つ全極型フィルタ.



LSP 音声合成フィルタ (奇数次)



$$c_i = -2 \cos \omega_i$$

図4. LSP 周波数の \cos 値を係数に持つ全極型フィルタ.



LSP 音声合成 LSI (1980)

■ 初の LSP 音声合成 LSI

- 1976年 TI 社から PARCOR 音声合成 LSI . “Speak & Spell” 用 . p-MOS 「TIショック」

■ 初の C-MOS による音声合成 LSI (ECL1565)

- 8mm × 8mm . 4000ゲート . 3.6 μ m ゲート CMOS . 10mW (3V 時)以下 .
- 直列パイプライン演算回路
- 144クロックで1サンプル出力 . マスタクロック 921.6kHz.
- アナログ混載は難 . 全デジタル . (パルス密度変調出力)
- マイクロプロセッサ向けのインタフェース . (UART をモデルに)

■ 世界最小の音声合成器(?)

- LSI + PROM (32k) + CMOS IC 2 + Tr 1 + スピーカー + 単三 3本