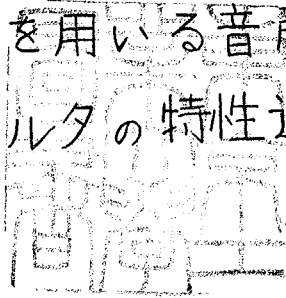


論文 / 著書情報
Article / Book Information

題目(和文)	一般化ケプストラムを用いる音声の分析・合成と合成フィルタの特性近似
Title(English)	
著者(和文)	小林隆夫
Author(English)	Takao Kobayashi
出典(和文)	学位:工学博士, 学位授与機関:東京工業大学, 報告番号:甲第1390号, 授与年月日:1982年3月26日, 学位の種別:課程博士, 審査員:今井聖
Citation(English)	Degree:Doctor of Engineering, Conferring organization: , Report number:甲第1390号, Conferred date:1982/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

一般化ケプストラムを用いる音声の分析・合成と
合成フィルタの特性近似

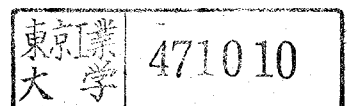


昭和 57 年 1 月

指導教官 今井 聖 教授

提出者 東京工業大学 大学院
博士課程 物理情報工学専攻

小林 隆夫



目次

第1章 序論

1.1 研究の背景と目的

1

1.2 研究の概要

5

1.3 準備

7

第2章 一般化ケプストラムを用いるスペクトル包絡の抽出

2.1 まえがき

10

2.2 一般化対数関数

12

2.3 一般化ケプストラムとその性質

2.3.1 一般化ケプストラム

14

2.3.2 一般化ケプストラムの性質

15

2.3.3 正規化一般化ケプストラム

17

2.4 一般化ケプストラムを用いるスペクトルの平滑化

2.4.1 一般化対数目盛におけるスペクトルの平滑化

20

2.4.2 スペクトル包絡の平滑化例

25

2.5 改良一般化ケプストラムを用いるスペクトル包絡の抽出

2.5.1 改良一般化ケプストラム

27

2.5.2 改良一般化ケプストラムの抽出例

30

2.6 むすび

38

第3章 2乗平均誤差評価による音声合成フィルタの対数特性近似

3.1 まえがき

39

3.2 FIRディジタルフィルタの対数特性近似法

3.2.1 近似法

41

3.2.2 係数の決定

43

3.3	FIRディジタルフィルタを用いた近似例	46
3.4	IIRディジタルフィルタの対数振幅特性近似法	
3.4.1	近似法	51
3.4.2	重み関数	53
3.4.3	係数の決定	54
3.5	安定性	
3.5.1	安定性	59
3.5.2	最小位相推移形フィルタ	60
3.6	IIRディジタルフィルタを用いた近似例	62
3.7	むすび	66

第4章 最小位相インパルス応答を用いる音声合成フィルタの特性近似

4.1	まえがき	67
4.2	対数振幅特性近似法	
4.2.1	修正最小二乗近似問題	69
4.2.2	係数の決定	71
4.2.3	最小位相インパルス応答	72
4.3	一般化対数特性の近似法	74
4.4	近似例	75
4.5	むすび	81

第5章 一般化対数特性近似ディジタルフィルタ

5.1	まえがき	82
5.2	一般化対数特性の直接近似	84
5.3	縦続接続による特性の近似	
5.3.1	2つの一般化対数特性近似ディジタルフィルタの縦続接続	
		87

第1章 序論

1.1 研究の背景と目的

デジタル信号処理技術や半導体集積回路技術の進歩に伴い、言葉（音声）を、人間相互間ばかりでなく、人間・機械間の情報伝達手段として用いようとする試みが盛んになってきている。言葉（音声）は、人間が持っている情報伝達手段の中で最も基本的なものの一つであり、それを利用するために特別な道具や訓練を必要としないことから、機械との情報交換に際しても、誰もが容易に取り扱えるように音声を利用することは当然の要求であると言える。最近では、簡単な言葉を話したり認識したりする装置や器具も実用化され、一般の人にとってもごく身近なものになりつつある。

ところで、音声を人間・機械間の情報伝達手段として利用するためには、音声波形を物理的な信号系列として取り扱い、何らかの信号処理をする必要がある。このような音声信号処理の方法としては、音声波形そのものを取り扱う方法と、適当な音声発生モデルを仮定し、音声信号をそのモデルのパラメータにより表現して取り扱う方法の2つに大別できる。後者の方法は前者に比べ、モデルのパラメータを抽出するという分析の処理が複雑になる反面、大幅な情報量の圧縮が可能であり、抽出されたパラメータは、原音声の合成や任意文章の法則合成などの合成パラメータとして使えるほか、認識に必要な特徴パラメータとしても利用できることから、非常に有用である。モデルを仮定して音声信号を取り扱う方法の代表的な例として、ケプストラム法⁽¹⁾あるいはそれを発展させたホモモルフィック信号処理の方法⁽²⁾や、線形予測法⁽³⁾⁽⁴⁾などがある。

ケプストラム法では、音声の発生モデルは、音源として平坦なスペクトルを

持つ周期的インパルス列が白色雑音が仮定され、声道の特性に相当する伝達特性を持つ時変フィルタ(声道フィルタ)の出力が音声信号となる。声道フィルタの対数振幅特性は、音声信号の対数スペクトル包絡に相当し、これはケプストラムの低次数部のみを取り出した系列のフーリエ変換で与えられる。ここでケプストラムは対数スペクトルのフーリエ係数として定義される値であり、音声信号の分析パラメータとしては、他のパラメータ、例えば線形予測係数にはみられないいくつかの良い性質を持っている。すなわち、線形予測係数が声道フィルタとして全極型モデルを仮定し、線形システムの固定理論によって得られるパラメータであるのに対し、ケプストラムは極零モデルのパラメータであるため、鼻音などのようにはっきりとした零点を持つスペクトル包絡に対して、も十分に正確な表現がとれる。また、音の大きさに対する聴覚の特性に近い評価目盛上での抽出パラメータであること、対数スペクトルとの関係が線形的であるため、法則合成を行う場合、時間方向の補間特性が良いことなどが、その長所としてあげられる。

しかし、ケプストラムを音声合成のパラメータとして考えた場合、問題点もいくつかある。一つは、線形予測係数が合成フィルタ(声道フィルタ)の係数を直接表すことから容易に合成フィルタを構成できるのに対し、ケプストラムは合成フィルタの対数周波数特性のフーリエ係数を表すにすぎないため、合成フィルタの構成が簡単でないということである。従ってケプストラム法による合成系では、ケプストラムで表された対数周波数特性を近似するフィルタをいかに実現するかということが、合成音声の品質を左右する重要な要因となる。

一方、ケプストラムが極零モデルのパラメータであるということは、全極型モデルを仮定する線形予測法に対して、利点である反面欠点にもなっている。すなわち、音韻の明瞭性や音の響きは主としてスペクトル包絡の極(フォルマント)の部分に依存し、鼻音などのようなは、はっきりとした零点を持つ場合を除いて、零点の部分は必ずしも重要ではない。ケプストラムは対数目盛上における表現パラメータであることから、急峻なスペクトル包絡の極の部分をも

分正確に表現しようとする、結果的に不要な零点の部分も極の部分と同じ重みで表されることになる。従って、合成のためのパラメータ表現として、ケプストラムは無駄があると言える。実際に、同じ明瞭度を得るために必要なパラメータの次数は、線形予測法に比べケプストラム法の方が大きくなる。また、音の大きさに対する聴覚特性は、正確にはべき法則に従うといわれていることから、音声合成のパラメータ抽出の評価目盛として対数目盛が最適かどうかという問題がある。以上のことから、ケプストラムと同様の良い性質を持ち、より効率的なスペクトル包絡の表現ができるパラメータを抽出するための評価目盛が要請されることになる。

ケプストラムを用いた合成フィルタの実現に関しては、Oppenheimらが提案した、直接たたみ込みによる方法⁽⁵⁾、全極型フィルタによる近似を用いる方法⁽⁶⁾が有名であるが、これらの方法ではあまり良好な合成音は得られないようである。これは、ケプストラムを最小位相インパルス応答という無限長の時間領域のパラメータに変換して処理することに問題があると考えられる。また、合成フィルタを構成するための計算量も多くなる。これに対して、対数振幅特性の直接近似フィルタを用いる合成系⁽⁷⁾では、これらの問題点は殆ど解決されている。対数振幅特性の直接近似フィルタ⁽⁸⁾は、フィルタ係数がケプストラムにより直接表されるため、係数の決定は非常に簡単であり、しかもかなり正確な対数特性の近似ができる。実際に得られる合成音も、原音に比べ劣化が少なく高品質であることが確かめられている。しかし、この方法では、合成フィルタの遅延素子数が通常のデジタルフィルタに比べて著しく大きくなる問題点がある。さらに、スペクトルの評価目盛を対数目盛以外に選んだ場合、そのまゝでは適用できないという問題点もある。

一方、より効率的なパラメータ表現を目的とした評価目盛としては、対数関数のかわりにべき乗 $()^{\alpha}$ を利用する方法⁽⁹⁾が提案されている。この方法は、有声音のような準同期波形に対するホモモルフィックデコンボリューション⁽¹⁰⁾による音源とフィルタの伝達関数の分離を基本としているため、無声

音のような非周期波形を含む一般の音声信号に適用するには、根拠があまり
はなっている。また、対数関数からべき乗という演算の拡張の解釈が、そのま
まの形では不自然であること、合成法がインパルス応答を求めて直接たたみ込
みにより合成音を得る方法であるため、計算量が多く、しかもそれ程良好な合
成音声の品質が期待できないなどの問題点がある。

以上のような問題点をふまえて、本研究では、より高品質な合成音を得る
ために、ケプストラム法を一般化した形での新しい音声のパラメータ表現の可
能性を探ること、およびケプストラムを用いる合成系の実現に際して最大の問題
点である、合成フィルタの特性近似法について新たな検討を加えることを目
的としている。

ケプストラムにかわる新しいパラメータを得るためには、まず対数目盛のか
わりとなる評価目盛が必要となる。この際、ケプストラムの良い性質や、従来
のケプストラム法による分析合成系との互換性などを考慮に入れ、評価目盛と
しては対数関数の自然な一般化となっていることが条件となる。また、合成フ
ィルタの特性近似については、通常のデジタルフィルタの構成で用いられる
程度の遅延素子数で良好な対数近似特性が得られること、対数関数にかわる新
しい評価目盛における特性近似法を与えることなどが要求される。

1. 2 研究の概要

前述の目的に沿って、本研究では、ケプストラム法の一般化を提案し、音声合成フィルタの近似法についていくつかの検討を行っている。

まず第2章では、対数関数の一つの自然な拡張であると考えられる関数（一般化対数関数）を定義し、この関数を評価目盛とする音声のスペクトル包絡の抽出法を示す。一般化対数関数は対数関数を含み、パラメータの値により直線から双曲線の範囲で任意の分数べき特性を実現できる関数である。ここでは、音声の一般化対数スペクトルのフーリエ係数を一般化ケプストラムと名付け、その性質、通常のケプストラムとの関係等を述べる。また、一般化ケプストラムを用いるスペクトル包絡の抽出法を示し、一般化対数関数のパラメータと得られるスペクトル包絡との関係を具体例を通して論じている。

第3章から第5章までは、分析により得られた音声の発生モデルのパラメータである、ケプストラムあるいは一般化ケプストラムを用いて、いかに合成フィルタを構成するかという問題、すなわち合成フィルタの特性近似法についてあてられている。

第3章では、一般化対数目盛の一つの特別な場合である対数目盛について、周波数領域での2乗平均誤差評価式を用いた、デジタルフィルタの特性近似法を示す。ここでは、対数目盛上での最小2乗近似問題を、適当な重みをかけた直線目盛での最小2乗近似問題に置き換えることにより、線形問題の解としてフィルタの係数を得ている。この方法によれば、フィルタ係数は簡単な計算アルゴリズムで求まり、フィルタも安定になることを示す。また具体的な近似例から、ケプストラムの次数と同一の遅延素子数の構成で、実用上十分な近似特性が得られることも示す。

第4章では、ケプストラムから計算される最小位相インパルス応答と、時間領域でのデジタルフィルタの設計法を組み合わせた、対数特性近似法を示す。

この方法は、殆ど変更なしに、一般化対数特性の近似法としても用いることが

できる。ここでは、具体的な近似法の手順を述べた後、いくつかの具体例を示し、その有効性を論じている。

第5章では、一般化対数特性を一般化対数目盛上で直接近似するデジタルフィルタの構成法を示す。ここで示す方法は、一般化対数関数の逆関数のパデ近似を利用するものであり、一般化ケプストラムの値がそのままフィルタ係数となる特長をもっている。まず、基本となるフィルタの近似構成法を示し、次に実用的な構成法について検討している。さらに、フィルタの安定性、近似誤差、パデ近似係数の修正による近似誤差の改善等について論じている。また具体的な近似例も示す。

第6章は、第2章で示した分析法と対をなす合成法を示している。合成パラメータは一般化ケプストラムであり、合成フィルタとしては、第5章で示した一般化対数特性近似デジタルフィルタを利用している。ここで述べる合成系は、従来のケプストラム法による分析合成系と完全に互換性がある。すなわち合成系への伝送パラメータは、通常のケプストラムとピッチであり、一般化ケプストラムは合成系の内部で簡単な変換式により得ている。本章では、まず合成系の全体の構成を示し、合成フィルタの実現に関して具体的な検討を加えている。次に、合成系から得られた合成音の品質評価を行った結果を示し、一般化対数関数のパラメータの値の選択について検討している。

1.3 準備

本論文で取り扱う信号はすべて離散信号系列とする。すなわち、実際の音声信号のように、時間に対して連続的な信号を $x(t)$ として、サンプリング周期 T で標本化した信号系列を

$$x(m) = x(mT), \quad m = \dots, -1, 0, 1, \dots \quad (1.1)$$

と表し、この $x(m)$ を取り扱う。

信号系列 $x(m)$ に対して、そのフーリエ変換を $X(\omega)$ と表すことにすれば、次式

$$X(\omega) = \sum_{m=-\infty}^{\infty} x(m) e^{-j\omega m} \quad (1.2)$$

で定義される。このとき、フーリエ変換 $X(\omega)$ は $\omega = 2\pi$ を周期とする周期関数となる。また $X(\omega)$ の逆変換は次式で定義される。

$$x(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega m} d\omega \quad (1.3)$$

本論文で理論的な取扱いを行うときは、特にことわらない限り、式(1.2)、(1.3)で定義されるフーリエ変換・逆変換対を用いる。しかし、現実的には式(1.2)の総和記号の範囲を無限大までとすることは不可能であり、また式(1.3)の積分も解析的に行うことが困難な場合が殆どである。従って、実際に電子計算機を用いてこれらの式を処理する場合は、次の離散的フーリエ変換・逆変換対を用いて行うことにする。すなわち、 L を適当な整数とし、 L 個の信号系列を x_0, x_1, \dots, x_{L-1} と表すことにすれば、離散的フーリエ変換 (DFT) は、

$$X_k = \sum_{m=0}^{L-1} x_m e^{-jk\Omega m} \quad (1.4)$$

で定義され、離散的逆フーリエ変換 (IDFT) は、

$$x_m = \frac{1}{L} \sum_{k=0}^{L-1} X_k e^{jm\Delta\Omega k} \quad (1.5)$$

で定義される。ここで

$$\Delta\Omega = 2\pi/L \quad (1.6)$$

である。 $X_k = X(k\Delta\Omega) = X(\omega)|_{\omega=k\Delta\Omega}$ であれば、 $X(m)$ と x_m の間には、

$$X_m = \sum_{n=-\infty}^{\infty} x(m+nL) \quad (1.7)$$

という関係が成立する。一般に、信号系列 $x(m)$ が $0 \leq m \leq L-1$ の範囲外で殆ど無視できる値であれば、式(1.7)で表される時間的エイリアスの影響は考えなくてもよい。

一方、式(1.2) ~ (1.6) は、サンプリング周期を1 [s] に正規化した場合の表現であるが、 T の値を考慮に入れたときは、次式のように表される。

$\Omega = \pi/T$ とし、式(1.2), (1.3) は

$$X(\omega) = \sum_{m=-\infty}^{\infty} x_m e^{-j\omega T m} \quad (1.8)$$

$$x(m) = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} X(\omega) e^{j\omega T m} d\omega \quad (1.9)$$

となる。このとき $X(\omega)$ は $\omega = 2\Omega$ を周期とする周期関数である。また、式(1.4), (1.5) は

$$X_k = \sum_{m=0}^{L-1} x_m e^{jk\Delta\Omega T m} \quad (1.10)$$

$$x_m = \frac{1}{L} \sum_{k=0}^{L-1} X_k e^{jm\Delta\Omega T k} \quad (1.11)$$

と表される。ここで、

$$\Delta\Omega = 2\Omega/L$$

(1.12)

とする。

本論文では、原則としてサンプリング周期を正規化した値で、理論上の取扱いを行う。

その他、離散信号系列の数学的取扱い、 z 変換、デジタルフィルタ等の詳細については、文献(10)を参照のこと。

第2章 一般化ケプストラムを用いる

スペクトル包絡の抽出

2.1 まえがき

ケプストラム法による音声の分析法では、声道の伝達特性に相当するスペクトル包絡は、ケプストラムの低次数部を取り出した系列のフーリエ変換で与えられる。ケプストラムは対数スペクトルのフーリエ係数で定義されるが、スペクトルの対数をとる演算を何らかの意味のある他の演算で置き換えることができれば、ケプストラム法の一般化を考えることができる。

このような一般化の一つとして、対数関数のかわりにべき乗 $()^x$ を用いる方法 (Spectral Root homomorphic Deconvolution System, SRDS)⁽⁹⁾が提案されている。しかし、この方法は、有声音のような準周期波形に対する、ホモモルフィックデコンボリューションによる音源とフィルタの伝達特性の分離を基本としているため、無声音のような非周期波形を含む一般の音声信号に適用するには、根拠があいまいになっている。また、対数関数 \ln とべき乗 $()^x$ という2つの演算の関係が、そのままでは不自然である。さらにこの方法では、音声合成系を構成する際、合成フィルタの伝達特性が無限長のインパルス応答の形で与えられるため、スペクトル包絡を忠実に近似するフィルタの実現が簡単でないという問題点もある。

本章では、音声信号のスペクトルの平滑化という観点から、対数関数を他の演算で置き換えることによる、SRDSより自然なケプストラム法の一般化を試みる。

対数関数のかわりとなる演算としては、任意の関数を考えることも可能であるが、対数・指数関数の性質の良さ、従来のケプストラム法との関連および音

の大きさに対する聴覚特性がべき法則に従うという事実等から、対数関数の性質をある程度保存しつつ、適当なべき乗特性が実現できることが望ましい。

そこで、ここでは、対数関数の一つの自然な拡張であると考えられる写像関数（一般化対数関数）を定義し、この関数を評価目盛とする音声のスペクトル包絡抽出法を述べる。一般化対数関数は対数関数を含み、パラメータの値を変化させることにより、直線から双曲線の範囲で連続的に任意の分数べき特性を実現できる関数である。通常のケプストラムにならって、音声スペクトルの一般化対数をとったもの（一般化対数スペクトル）のフーリエ係数を一般化ケプストラムと呼ぶことにすれば、一般化ケプストラムはケプストラムと同様な取り扱いが可能であり、スペクトル包絡抽出のための分析系も容易に構成できる。また、ケプストラムと一般化ケプストラムとの関係もSRDSと比べ、より自然となっている。

以下、まず一般化対数関数の定義を述べ、次に一般化ケプストラムに関する一般的な性質を示す。さらに、具体的なスペクトル包絡の抽出法を示し、スペクトル包絡の抽出例も示す。

2.2 一般化対数関数

W を複素変数, $\sigma \in [-1, 1]$ の範囲の実数とするとき, 次の写像関数を考える。

$$S_{\sigma}(W) = \frac{1}{\sigma} (W^{\sigma} - 1), \quad \sigma \neq 0 \quad (2.1)$$

この関数の $W=1$ のまわりにおけるテイラー展開は,

$$S_{\sigma}(W) = W - 1 + \sum_{n=2}^{\infty} \frac{(\sigma-1)(\sigma-2)\cdots(\sigma-n+1)}{n!} (W-1)^n,$$

$$|W-1| < 1 \quad (2.2)$$

と表される。式(2.2)で $\sigma \rightarrow 0$ とすれば

$$\lim_{\sigma \rightarrow 0} S_{\sigma}(W) = W - 1 - \frac{(W-1)^2}{2} + \frac{(W-1)^3}{3} - \dots$$

$$= \ln W \quad (2.3)$$

となることから, $\sigma=0$ の場合は

$$S_{\sigma}(W) = \ln W, \quad \sigma=0 \quad (2.4)$$

と定義する。

σ が 1 および -1 の場合には $S_{\sigma}(W)$ はそれぞれ

$$S_{\sigma}(W) = W - 1, \quad \sigma=1 \quad (2.5)$$

$$S_{\sigma}(W) = 1 - \frac{1}{W}, \quad \sigma=-1 \quad (2.6)$$

となる。特に $W=1$ の近傍では, 任意の ϵ の値に対して次の関係が成立する。

$$S_{\sigma}(W) \simeq W - 1 \simeq \ln W, \quad |W-1| \simeq 0 \quad (2.7)$$

図 2.1 に、実数変数 x に対する $y = S_\gamma(x)$ の概形をいくつかの γ の値について示す。

次に、 $S_\gamma(w)$ の逆関数 $S_\gamma^{-1}(w)$ は式 (2.1) の定義から次式で与えられる。

$$S_\gamma^{-1}(w) = (1 + \gamma w)^{\frac{1}{\gamma}}, \quad \gamma \neq 0 \quad (2.8)$$

$S_\gamma^{-1}(w)$ はテイラー展開により

$$S_\gamma^{-1}(w) = 1 + w + \sum_{n=2}^{\infty} \frac{(1-\gamma)(1-2\gamma)\cdots(1-(n-1)\gamma)}{n!} w^n,$$

$$|\gamma w| < 1 \quad (2.9)$$

と表される。ここで $\gamma \rightarrow 0$ とすれば

$$\begin{aligned} \lim_{\gamma \rightarrow 0} S_\gamma^{-1}(w) &= 1 + w + \frac{w^2}{2!} + \frac{w^3}{3!} + \cdots \\ &= \exp w \end{aligned} \quad (2.10)$$

となることから、 $\gamma = 0$ の場合は

$$S_0^{-1}(w) = \exp w \quad (2.11)$$

と定義する。実際、式 (2.11) は

$\gamma = 0$ の場合の $S_\gamma(w)$ の定義である $\ln w$ の逆関数になっている。

以上の様に、 $S_\gamma(w)$ は対数関数

を含み、式 (2.7) の関係を満足

しながら特性の変化する関数である

ことから、以下では一般化対数

関数と呼ぶことにする。

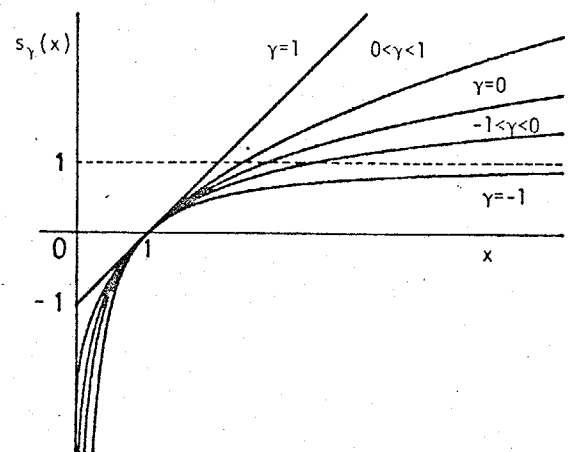


Fig. 2.1. The curves of $s_\gamma(x)$.

2.3 一般化ケプストラムとその性質

2.3.1 一般化ケプストラム

$x(m)$ を音声信号, 従って安定な実数系列とし, そのフーリエ変換を $X(\omega)$ と表すことにする。

$$X(\omega) = \sum_{m=-\infty}^{\infty} x(m) e^{-j\omega m} \quad (2.12)$$

一般にケプストラム $\hat{x}(m)$ は対数スペクトル $\hat{X}(\omega) = \ln X(\omega)$ のフーリエ係数として定義される。

$$\hat{x}(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{X}(\omega) e^{j\omega m} d\omega \quad (2.13)$$

対数関数の場合と同様にして, ここでは $X(\omega)$ の一般化対数をとったものを一般化対数スペクトル, 一般化対数スペクトルのフーリエ係数を一般化ケプストラムと呼ぶことにし, それぞれ $\check{X}(\omega)$, $\check{x}(m)$ で表すことにする。すなわち

$$\check{X}(\omega) = S_2 \{ X(\omega) \} \quad (2.14)$$

$$\check{x}(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \check{X}(\omega) e^{j\omega m} d\omega \quad (2.15)$$

とすることで, 一般化対数をとる際べき乗 $()^r$ を計算する必要があるが, 複素変数 ω に対して ω^r の値は多価関数となるため, $\check{X}(\omega)$ が一意的に定まらない。そこで, 対数関数の場合⁽¹⁰⁾と同様, べき乗をとる演算を次のように定義する。

$$X^r(\omega) = |X(\omega)|^r \cdot e^{j \arg [X(\omega)]} \quad (2.16)$$

ここで,

$$X(\omega) = X_R(\omega) + j X_I(\omega) \quad (2.17)$$

として

$$\arg [X(\omega)] = \int_0^{\omega} \{ \text{ARG} [X(\Omega)] \}' d\Omega \quad (2.18)$$

$$\begin{aligned} \{ \text{ARG} [X(\omega)] \}' &= \frac{d}{d\omega} \{ \text{ARG} [X(\omega)] \} \\ &= \frac{X_R(\omega) X_I'(\omega) - X_I(\omega) X_R'(\omega)}{X_R^2(\omega) + X_I^2(\omega)} \end{aligned} \quad (2.19)$$

であり, $\text{ARG} [X(\omega)]$ は位相の主値を表すものとする。このとき $\check{X}(\omega)$ は一意的に定まり, しかも連続関数となる。また, もとの信号系列 $x(m)$ が安定な実数系列であることから, $|X(\omega)|^2$, $\arg [X(\omega)]$ はそれぞれ偶関数および奇関数となり, 一般化ケフストラム $\check{x}(m)$ も安定な実数系列となる。

一方, $x(m)$ の z 変換を $X(z)$ とすれば

$$\check{X}(z) = s_z \{ X(z) \} \quad (2.20)$$

は収束領域に z 平面の単位円を含み, 一般化ケフストラム $\check{x}(m)$ により

$$\check{X}(z) = \sum_{m=-\infty}^{\infty} \check{x}(m) z^{-m} \quad (2.21)$$

で与えられる。

2.3.2 一般化ケフストラムの性質

一般化ケフストラムに関するいくつかの性質を以下に示す。特にことわりなし限り, $x(m)$ は信号系列, $\check{x}(m)$ は一般化ケフストラムとし, 通常のケフストラムは $\hat{x}(m)$ と表すことにする。

(1) 信号系列 $x(m)$ が2つの信号系列のたたみ込み $x(m) = h(m) * u(m)$ で与えられるとき, その z 変換は $X(z) = H(z) \cdot U(z)$ で表される。そこで, $\check{H}(z) = s_z \{ H(z) \}$, $\check{U}(z) = s_z \{ U(z) \}$ とすれば, 一般化対数関数の定義から,

$$\checkmark X(z) = \frac{1}{\delta} \left\{ H^{\delta}(z) \cdot U^{\delta}(z) - 1 \right\}$$

$$= \frac{1}{\delta} \left\{ (\delta \checkmark H(z) + 1)(\delta \checkmark U(z) + 1) - 1 \right\}$$

$$= \checkmark H(z) + \checkmark U(z) + \delta \checkmark H(z) \cdot \checkmark U(z) \quad (2.22)$$

が得られる。従って、 $x(m) = h(m) * u(m)$ であれば次の関係が成立する。

$$\checkmark x(m) = \checkmark h(m) + \checkmark u(m) + \delta \checkmark h(m) * \checkmark u(m) \quad (2.23)$$

式(2.23)で $\delta \rightarrow 0$ とすれば、通常のケプストラムの性質に他ならない。

(2) 信号 $x(m)$ の z 変換を $X(z)$ として、 $\checkmark X(z) = s_{\delta} \{X(z)\}$ を z に関して微分すれば、

$$X(z) \cdot \frac{d}{dz} \checkmark X(z) = X^{\delta}(z) \cdot \frac{d}{dz} X(z) \quad (2.24)$$

が得られる。ところで、 $X^{\delta}(z) = \delta \checkmark X(z) + 1$ であるから、式(2.24)は

$$X(z) \cdot \frac{d}{dz} \checkmark X(z) = (1 + \delta \checkmark X(z)) \cdot \frac{d}{dz} X(z) \quad (2.25)$$

と書くことができる。これを逆 z 変換すれば、信号系列 $x(m)$ と一般化ケプストラム $\checkmark x(m)$ との間の次の関係式が得られる。

$$\sum_{k=-\infty}^{\infty} k \checkmark x(k) x(m-k) = m x(m) + \delta \sum_{k=-\infty}^{\infty} k x(k) \checkmark x(m-k) \quad (2.26)$$

(3) $\hat{X}(z) = \ln X(z)$ とすると

$$\checkmark X(z) = \frac{1}{\delta} \left(e^{\delta \hat{X}(z)} - 1 \right)$$

$$= \hat{X}(z) + \frac{\delta}{2!} \hat{X}^2(z) + \frac{\delta^2}{3!} \hat{X}^3(z) + \dots \quad (2.27)$$

と表される。式(2.27)を逆Z変換すれば、次式の一般化ケプストラム $\check{x}(m)$ と通常のケプストラム $\hat{x}(m)$ の間の関係式が得られる。

$$\check{x}(m) = \hat{x}(m) + \frac{\delta}{2!} \hat{x}(m) * \hat{x}(m) + \frac{\delta^2}{3!} \hat{x}(m) * \hat{x}(m) * \hat{x}(m) + \dots \quad (2.28)$$

(4) $x(m)$ が最小位相系列であれば、ケプストラムは $\hat{x}(m) = 0, m < 0$ である。⁽¹⁰⁾ これと、式(2.28)から、一般化ケプストラムに対しても、

$$\check{x}(m) = 0, \quad m < 0 \quad (2.29)$$

が成立する。

(5) 最小位相系列では $x(m) = \check{x}(m) = 0, m < 0$ が成立することから、式(2.26)は簡単になり次式で表される。

$$\check{x}(m) = \begin{cases} 0, & m < 0 \\ S_{\delta} \{ x(0) \}, & m = 0 \\ (1 + \delta \check{x}(0)) \cdot \frac{x(1)}{x(0)}, & m = 1 \\ (1 + \delta \check{x}(0)) \cdot \frac{x(m)}{x(0)} + \sum_{k=1}^{m-1} \frac{k}{m x(0)} \{ \delta x(k) \check{x}(m-k) - \check{x}(k) x(m-k) \}, & m > 1 \end{cases} \quad (2.30)$$

2.3.3 正規化一般化ケプストラム

信号系列 $x(m)$ に対して、そのZ変換を

$$X(z) = A \cdot X_0(z) = \exp(\hat{x}(0)) \cdot X_0(z) \quad (2.31)$$

$$\hat{x}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |X(\omega)| d\omega \quad (2.32)$$

という形で表すことにする。ここで、 $X_0(z)$ はスペクトルの形状を表す項、 A は定数項で全体のパワーを表す項と考えればよい。 $\hat{x}(m) = \mathcal{Z}^{-1}(\hat{X}(z))$,
 $\hat{x}_0(m) = \mathcal{Z}^{-1}(\hat{X}_0(z))$ とすれば、対数関数の性質から明らかに

$$\hat{x}(m) = \hat{x}_0(m), \quad m \neq 0 \quad (2.33)$$

が成立する。これは、定数項 A が変化しても、スペクトルの形状を表す項 $X_0(z)$ が一定であれば、通常のケプストラム $\hat{x}(m)$ の値は $\hat{x}(0)$ の項を除いて変化しないことを意味している。ところが、一般化ケプストラムの場合は、 $X_0(z)$ が一定だとしても、 A の値が変化すれば $\hat{x}(m)$ の値も全体的に変化し、取扱いが簡単ではない。

そこで、ここでは定数項 A を除いた $X_0(z)$ の項だけにのみ値が依存する一般化ケプストラムの表現法を考える。

式(2.31)の一般化対数をとれば、式(2.22)の関係から

$$\check{X}(z) = \check{A} + (1 + \gamma \check{A}) \check{X}_0(z) \quad (2.34)$$

が得られる。ここで $\check{A} = s_0(A)$ である。式(2.34)の逆Z変換を求めれば

$$\check{x}(m) = \check{A} \delta(m) + (1 + \gamma \check{A}) \check{x}_0(m) \quad (2.35)$$

となる。ここで、 $\delta(m)$ は単位サンプル数列であり、 $\check{x}_0(z) = \mathcal{Z}^{-1}(\check{X}_0(z))$ とする。式(2.35)で $m \neq 0$ の場合は $\delta(m) = 0$ であるから、

$$\check{x}_0(m) = \frac{\check{x}(m)}{1 + \gamma \check{A}}, \quad m \neq 0 \quad (2.36)$$

が成立することになる。

特に、信号系列 $x(m)$ が最小位相系列であれば、式(2.28), (2.29)の関係

および $\hat{x}_0(0) = 0$ により

$$\check{x}_0(0) = 0 \quad (2.37)$$

が成り立つことから、式(2.35)で $m=0$ の場合は

$$\check{x}(0) = \check{A} = S_\delta(A) \quad (2.38)$$

となる。また、式(2.36)も

$$\check{x}_0(m) = \frac{\check{x}(m)}{1 + \delta \check{x}(0)} \quad (2.39)$$

と表せる。

$\check{x}_0(m)$ は、スペクトルの形状を表す項 $\check{X}_0(z)$ の逆Z変換であるから、定数項 A の値が変化しても、その値は変化しない。従って、 $X_0(z)$ が決まれば、一般化ケプストラム $\check{x}(m)$ から式(2.39)によって表現される $\check{x}_0(m)$ は一意に定まることになる。なお、式(2.39)で $\delta=0$ とすれば、式(2.33)と一致し、対数関数の場合すなわちケプストラムの関係式を表している。

式(2.39)で表現される $\check{x}_0(m)$ のことを、以下では正規化一般化ケプストラムと呼ぶことにする。

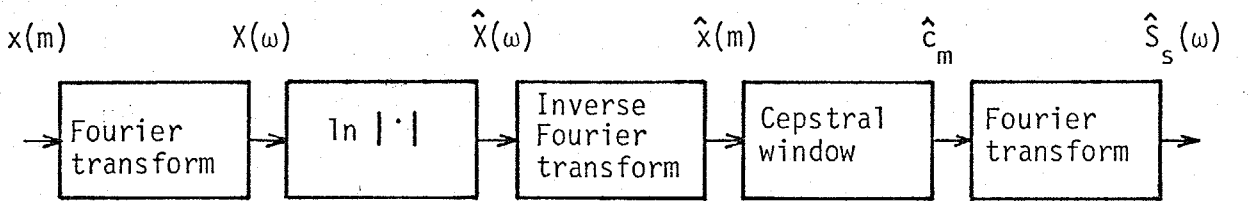
2.4 一般化ケプストラムを用いるスペクトルの平滑化

2.4.1 一般化対数目盛におけるスペクトルの平滑化

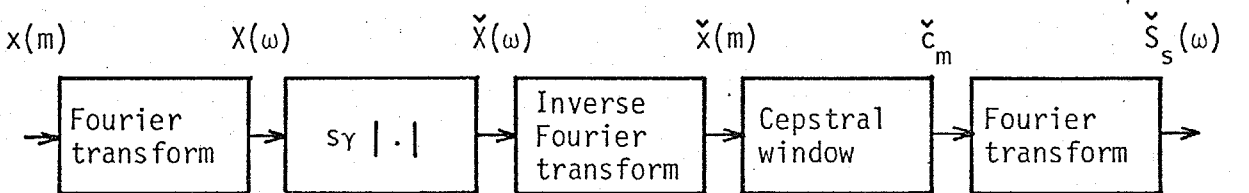
ケプストラム法における音声のスペクトル包絡(平滑化スペクトル)の抽出法は、図2.2(a)に示すとおりである。すなわち、音声信号からフーリエ変換によりスペクトルを求め、次に対数振幅スペクトルを求めて、逆フーリエ変換によりケプストラムを得る。ケプストラムに適当なケプストラム窓(ケフレンシ窓)をかけて低次数部のみを取り出し、そのフーリエ変換をとれば、求める対数スペクトル包絡(平滑化対数スペクトル)が得られる。

ケプストラム法において、スペクトルの対数をとる部分を一般化対数で置き換之れば、一般化対数目盛でのスペクトルの平滑化ができると考えられる。これを示したのが図2.2(b)である。

一般化ケプストラムの値は、定義どおりに一般化対数関数 $s_y(\cdot)$ の関数計算をし、逆フーリエ変換から求めてもよいが、あまり実用的ではない。なぜならば、べき乗の実際の計算には特別の場合を除いて、 \ln および \exp の計算を必



(a)



(b)

Fig. 2.2. (a) Spectral smoother based on the cepstral method.

(b) Spectral smoother on the generalized logarithmic scale.

要とし、さらに一般化対数関数のパラメータの値が変わればその度毎に一般化対数の値を計算する必要があるからである。これに対して、音声信号 $x(m)$ のフーリエ変換を $X(\omega)$ とするとき、振幅特性が $|X(\omega)|$ で表される最小位相系を考えれば、ケプストラムの値から簡単な変換式により一般化ケプストラムが計算できる。

振幅特性が $|X(\omega)|$ である最小位相系のケプストラムは、 $\hat{x}(m)$ を

$$\hat{x}(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |X(\omega)| e^{j\omega m} d\omega \quad (2.40)$$

とすると、

$$\hat{c}_m = \begin{cases} 0, & m < 0 \\ \hat{x}(0), & m = 0 \\ 2\hat{x}(m), & m > 0 \end{cases} \quad (2.41)$$

で与えられる。また、この系のインパルス応答(最小位相インパルス応答)

$x_{min}(m)$ は次式で与えられる⁽¹⁰⁾。

$$x_{min}(m) = \begin{cases} 0, & m < 0 \\ \exp(\hat{c}_0), & m = 0 \\ \sum_{k=1}^m \binom{k}{m} \hat{c}_k x_{min}(m-k), & m > 0 \end{cases} \quad (2.42)$$

と(3)で、前節の一般化ケプストラムの性質(5)より、最小位相系列と一般化ケプストラムの間には式(2.30)の関係が成立することから、式(2.42)で表される $x_{min}(m)$ を用いて一般化ケプストラム \hat{c}_m は次式で与えられる。

$$\check{c}_m = \begin{cases} 0, & m < 0 \\ S_\delta \{ \chi_{\min}(0) \}, & m = 0 \\ (1 + \delta \check{c}_0) \cdot \frac{\chi_{\min}(1)}{\chi_{\min}(0)}, & m = 1 \\ (1 + \delta \check{c}_0) \cdot \frac{\chi_{\min}(m)}{\chi_{\min}(0)} \\ + \sum_{k=1}^{m-1} \frac{k}{m \chi_{\min}(0)} \{ \delta \chi_{\min}(k) \check{c}_{m-k} - \check{c}_k \chi_{\min}(m-k) \}, & m > 1 \end{cases} \quad (2.43)$$

$\delta = 0$ の場合は, 式(2.42), (2.43) の手順は不要で \hat{c}_m そのものがケプストラムの値となる。式(2.43) で $\delta \rightarrow 0$ とすれば, 式(2.42) の逆変換式と一致することから, 式(2.43) から計算された一般化ケプストラムの値は $\check{c}_m = \hat{c}_m$ となり, 通常のケプストラムと一致する。従って, 式(2.43) は $\delta = 0$ の場合を含めて有効である。

以上の手順をブロック図にして, 図 2.3 に示す。

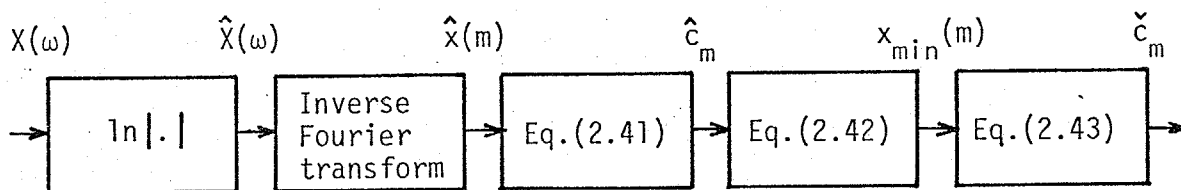


Fig. 2.3. Practical implementation of a system for obtaining the generalized cepstrum.

得られた一般化ケプストラムにケプストラム容

$$W_c(m) = \begin{cases} 1, & 0 \leq m \leq M \\ 0, & m > M \end{cases} \quad (2.44)$$

をかけて低時間部のみを取り出し、そのフーリエ変換をとれば、一般化対数目盛上で平滑化されたスペクトル

$$\check{S}_s(\omega) = \sum_{m=0}^M \check{C}_m e^{-j\omega m} \quad (2.45)$$

が得られる。ここで M はケプストラムの次数とする。

一般化ケプストラム \check{C}_m に対する正規化一般化ケプストラムを \check{v}_m で表すことにすれば、式(2.39) から \check{v}_m の値は、

$$\check{v}_m = \frac{\check{C}_m}{1 + \alpha \check{C}_0}, \quad m > 0 \quad (2.46)$$

により計算できる。ところで、一般には \check{C}_m の値より正規化一般化ケプストラムの値の方が重要であることから、実際には次のように \check{v}_m の値を直接求める方が計算量が少なく効率が良い。

まず、式(2.42)を使って最小位相インパルス応答 $x_{min}(m)$ を求める際、 $\hat{C}_0 = 0$ として計算した最小位相インパルス応答を $h_{min}(m)$ とする。すなわち

$$h_{min}(m) = \begin{cases} 1, & m=0 \\ \hat{C}_1, & m=1 \\ \sum_{k=1}^m \left(\frac{k}{m}\right) \hat{C}_k h_{min}(m-k), & m > 1 \end{cases} \quad (2.47)$$

とする。このとき、 $x_{min}(m)$ と $h_{min}(m)$ との間には、

$$x_{min}(m) = \exp(\hat{C}_0) \cdot h_{min}(m) = C_0 \cdot h_{min}(m) \quad (2.48)$$

が成立している。次に、 $h_{min}(m)$ から式(2.30)の関係を用いて計算した一

一般化ケプストラムの値を仮に \check{v}_m と表すことにすれば,

$$\check{v}_m = \begin{cases} h_{\min}(1) \quad (= \hat{c}_1), & m=1 \\ h_{\min}(m) \\ + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) \left\{ \delta h_{\min}(k) \check{v}_{m-k} = \check{v}_k h_{\min}(m-k) \right\}, & m>1 \end{cases} \quad (2.49)$$

で与えられる。式(2.43)と式(2.49)を比較すると

$$\check{c}_m = (1 + \delta \check{c}_0) \check{v}_m \quad (2.50)$$

$$\text{但し} \quad \check{c}_0 = s_0(c_0)$$

という関係があることが容易に確かめることができる。これと式(2.46)から、式(2.49)で表される \check{v}_m は正規化一般化ケプストラムに等しいことがわかる。従って、正規化一般化ケプストラムの値は \check{c}_m を求めることなく、ケプストラム \hat{c}_m ($m>0$) から式(2.47), (2.49)を使って直接計算できることになる。なお、 \check{c}_0 の値が必要であれば,

$$\check{c}_0 = s_0(c_0) = s_0\{\exp(\hat{c}_0)\} \quad (2.51)$$

で求められる。

以上のように、一般化ケプストラム \check{c}_m および正規化一般化ケプストラム \check{v}_m は最小位相ケプストラムから簡単な変換式から求められるが、 M 次までの \check{c}_m あるいは \check{v}_m を計算するには M 次までの最小位相ケプストラムしか必要としないことが、式(2.42), (2.43) あるいは式(2.47), (2.49)からわかる。従って、実際上は必要次数 M 次までのケプストラムのみが得られていれば、任意の δ に対して平滑化スペクトル $\check{S}_s(\omega)$ を求めることができる。

2.4.2 スペクトルの平滑化例

図2.4に、男声母音 /a/ のスペクトル平滑化の例を示す。図中に示してある音声波形に25.6 ms のハミング窓をかけ、DFTによりスペクトルを求め、図2.3に示した手順に従って平滑化一般化ケプストラムを求めた。ケプストラムの次数Mは20としてある。図中の $\gamma=0$ の場合は対数目盛の場合であり、通常のケプストラム法によって得られる平滑化スペクトルと一致する。比較を容易にするため、もとのスペクトルおよび平滑化スペクトルは、すべて対数振幅目盛上に描いてある。

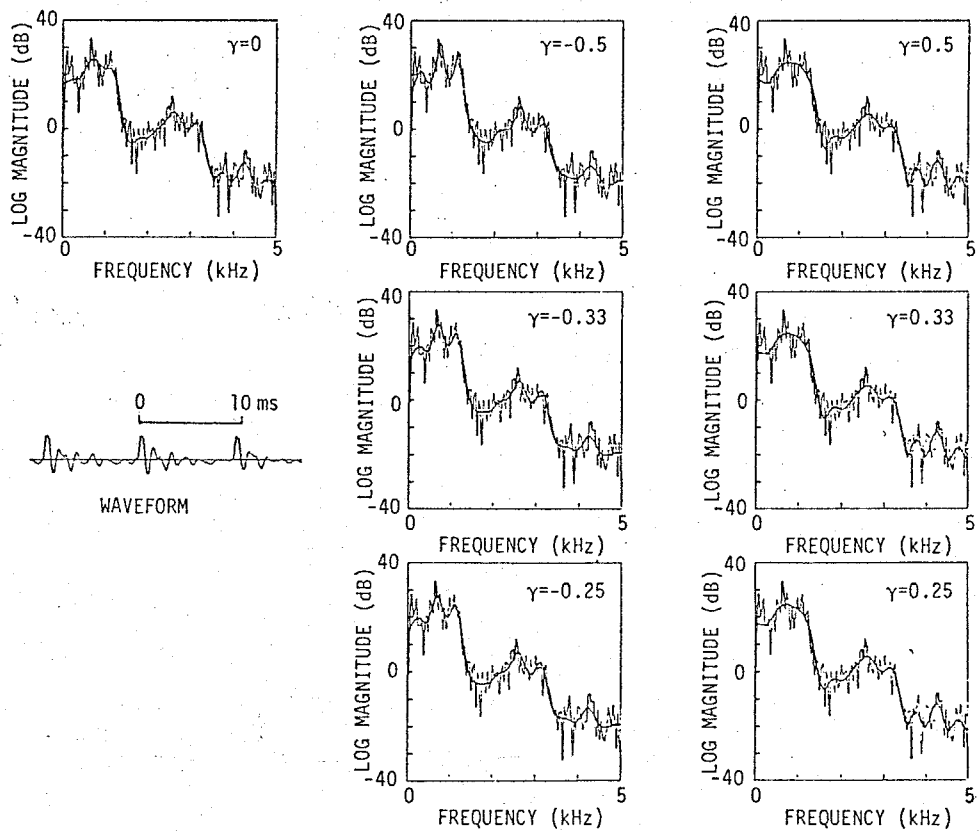


Fig. 2.4. Smoothed generalized log spectra of a vowel /a/.

図2.4から、 γ が負の場合は、スペクトルの極の部分強調した形の平滑化スペクトルが得られ、 γ が正の場合は、スペクトルの零点の部分強調した形の平滑化スペクトルが得られることがわかる。 $\gamma=0$ 、すなわち対数目盛の場合には、スペクトルの極と零点を同等に評価していると考えられる。

一般的に、ケプストラム法によって得られる合成音声は、ケプストラムの次数が低い場合、響きのないこもった感じの音になりやすいが、この原因はスペクトルの極部分を十分に表現できなためと考えられる。ところで、図2.4の平滑化例からわかるように、 γ が負の値では極部分が強調された形となることから、合成音声はより響きのある音質になることが期待できる。しかし、この点に関する議論は、実際に合成系を構成し、合成音声の品質評価を行った段階に譲ることにする。

2.5 改良一般化ケプストラムを用いるスペクトル包絡の抽出

2.5.1 改良一般化ケプストラム

ケプストラム法で得られる平滑化スペクトルは、対数スペクトル上の線形な平滑化であるため、スペクトルの微細構造の低レベル部分の影響を受けやすく、フレーム方向（時間方向）に対して不安定なものになってしまう。一般化対数目盛における平滑化では、前節の例で示したように、一般化対数関数のパラメータの値により微細構造の影響の程度は異なるが、通常のケプストラム法と同様なことが言える。

このような問題を解決する一つの手段として、対数目盛の場合は改良ケプストラム法⁽¹¹⁾が提案されている。これは、対数振幅スペクトル上の局所的ピークを結んで包絡を得る方法であり、手順は多少複雑となるが、真の包絡をフレーム方向に対して安定に抽出することができ、合成音も平滑化ケプストラムを用いた場合に比べ、より原音に近い品質が得られることが確かめられている。対数振幅スペクトルのかわりに、一般化対数振幅スペクトル上で改良ケプストラム法を適用すれば、一般化対数目盛における改良ケプストラム法を考えることができる。

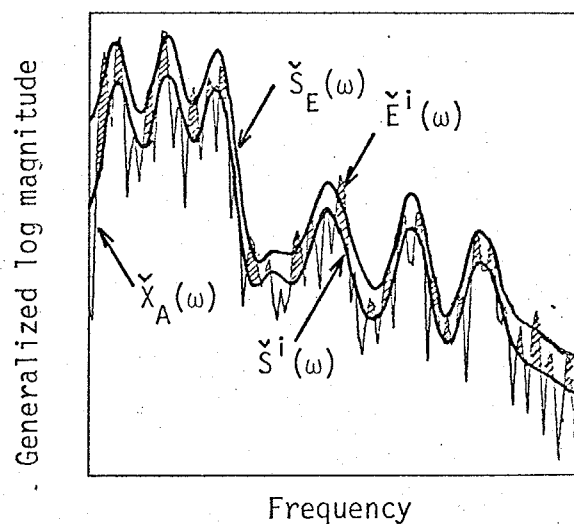


Fig. 2.5. Illustration of the improved cepstral method.

するわけ、一般化対数振幅スペクトル $\check{X}_A(\omega) = S_0 \{ |X(\omega)| \}$, ケプストラムの次数を M とするとき、まず初期値として一般化対数振幅目盛上で平滑化スペクトル $\check{S}^0(\omega)$ を求める。次に $i=0$ として、真のスペクトル包絡との残差スペクトル $\check{E}^i(\omega)$ を求める。ここで $\check{E}^i(\omega)$ は $g(x)$ を

$$g(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2.52)$$

として次式

$$\check{E}^i(\omega) = g \{ \check{X}_A(\omega) - \check{S}^i(\omega) \} \quad (2.53)$$

で定義される。さらに、この残差スペクトルから、適当な $\alpha \geq 0$ を選んで

$$\check{E}_\alpha^i(\omega) = (1 + \alpha) \check{E}^i(\omega) \quad (2.54)$$

を計算し、 $\check{E}_\alpha^i(\omega)$ に対して平滑化スペクトル $\check{S}_\alpha^i(\omega)$ を求める。これともとの平滑化スペクトル $\check{S}^i(\omega)$ を加えた

$$\check{S}^{i+1}(\omega) = \check{S}^i(\omega) + \check{S}_\alpha^i(\omega) \quad (2.55)$$

を新しく得られたスペクトル包絡とする。以上の手順を $i := i+1$ と置き直して適当な回数 J 回だけ繰り返せば、真の包絡 $\check{S}_E(\omega) = \check{S}^J(\omega)$ が得られる。

なお、式(2.54)で α は収束を速めるために加える項であり、加速係数と呼ばれる。

上記の手順で得られたスペクトル包絡は一般化対数振幅目盛上で

$$\check{S}_E(\omega) = \sum_{m=0}^M \check{C}_m \cos m\omega \quad (2.56)$$

という形で与えられる。

とすることで、上記の方法によれば、一般化ケプストラム \check{C}_m の値は、 δ が異なる

ればその度毎に一般化対数振幅スペクトル $\check{X}_A(\omega)$ を計算し、改良ケプストラム法の手順を繰り返さなければならない。これは前節で示した平滑化一般化ケプストラムが、通常のケプストラムのみが与えられていれば、任意の γ について簡単な変換式から計算できたことと比べると取扱いが簡単でない。さらに、合成への応用を考えた場合、式(2.56)で表される一般化対数周波数特性を持つフィルタを構成することが容易でないという問題点もある。

このようなことから、対数目盛での改良ケプストラムとの互換性も考慮してここでは次のような擬似的な方法を考える。まず、改良ケプストラム法により対数振幅スペクトル上で真の包絡 $\hat{S}(\omega)$ を求める。このとき、スペクトル包絡は

$$\hat{S}(\omega) = \sum_{m=0}^M \hat{C}_m \cos m\omega \quad (2.57)$$

と表されるとする。ところで得られた改良ケプストラム \hat{C}_m は、対数振幅特性が $\hat{S}(\omega)$ である最小位相系のケプストラムになっていることから、一般化ケプストラムは式(2.42), (2.43)により、また正規化一般化ケプストラムは式(2.47), (2.49)によりそれぞれ M 次までの項を計算することができ、そこで、このようにして得られた一般化ケプストラム \check{C}_m を改良一般化ケプストラム、正規化一般化ケプストラム \check{C}_m を正規化改良一般化ケプストラムとそれぞれ呼ぶことにする。

この場合、改良一般化ケプストラムによって表されるスペクトル包絡は、

$$\check{S}(\omega) = \sum_{m=0}^M \check{C}_m e^{j\omega m} \quad (2.58)$$

という形で与えられることになる。 $\check{S}(\omega)$ の一般化対数振幅特性

$$\check{S}_A(\omega) = S_2 \left\{ |S(\omega)| \right\} = S_2 \left\{ |S_2^{-1}(\check{S}(\omega))| \right\} \quad (2.59)$$

を求めてみると、

$$\check{S}_A(\omega) = \frac{1}{\delta} \left[\left\{ (1 + \delta \check{S}_R(\omega))^2 + (\delta \check{S}_I(\omega))^2 \right\}^{\frac{1}{2}} - 1 \right] \quad (2.60)$$

但し

$$\check{S}(\omega) = \check{S}_R(\omega) + j \check{S}_I(\omega) \quad (2.61)$$

となり、また、 $\check{S}_A^2(\omega) = S_\delta \{ |S(\omega)|^2 \}$ とすれば

$$\check{S}_A^2(\omega) = 2\check{S}_R(\omega) + \delta \left\{ (\check{S}_R(\omega))^2 + (\check{S}_I(\omega))^2 \right\} \quad (2.62)$$

が得られる。この $\check{S}(\omega)$ の一般化対数振幅特性 $\check{S}_A(\omega)$ と式(2.56)で表される真の一般化対数目盛上のスペクトル包絡 $\check{S}_E(\omega)$ は、当然であるが一致しない。一例として、図2.6に男声 /nambu/ の /bu/ の初めの部分のスペクトル包絡 $X(\omega)$ 、平滑化スペクトル $\check{S}_s(\omega)$ 、真のスペクトル包絡 $\check{S}_E(\omega)$ および $\check{S}(\omega)$ をいくつかの δ の値についてそれぞれ示す。比較を容易にするために、すべてのスペクトルを対数振幅目盛上に描いてある。

同図からもわかるように、 $|\delta|$ の値がそれ程大きくない場合、例えば $|\delta| \leq 0.2$ の範囲では、 $\check{S}_E(\omega)$ と $\check{S}_A(\omega)$ の差異は小さいと考えてよい。また、 $\check{S}_E(\omega)$ そのものも、改良ケプストラム法における加速係数 α や繰返し回数 J の選び方により変化することから、実用上は上記の $\check{S}(\omega)$ を一般化対数目盛での真のスペクトル包絡とみなしても大きな問題は生じない。一方、対数目盛での改良ケプストラムのみが得られていれば、 $|\delta|$ の値がそれ程大きくない範囲内で、平滑化一般化ケプストラムの場合と同様、任意の δ について一般化ケプストラムを計算できること、さらには、第5章で示すように、式(2.58)の形が合成フィルタの構成に適していることなど利点は多い。

2.5.2 改良一般化ケプストラムの抽出例

天気予報から取り出した男性音声および女性音声(それぞれ約24秒分)に対し、一般化改良ケプストラムをいくつかの δ の値について求め、その値の統計分布を調べた結果を以下に示す。分析フレーム数は全体で約9600フレー

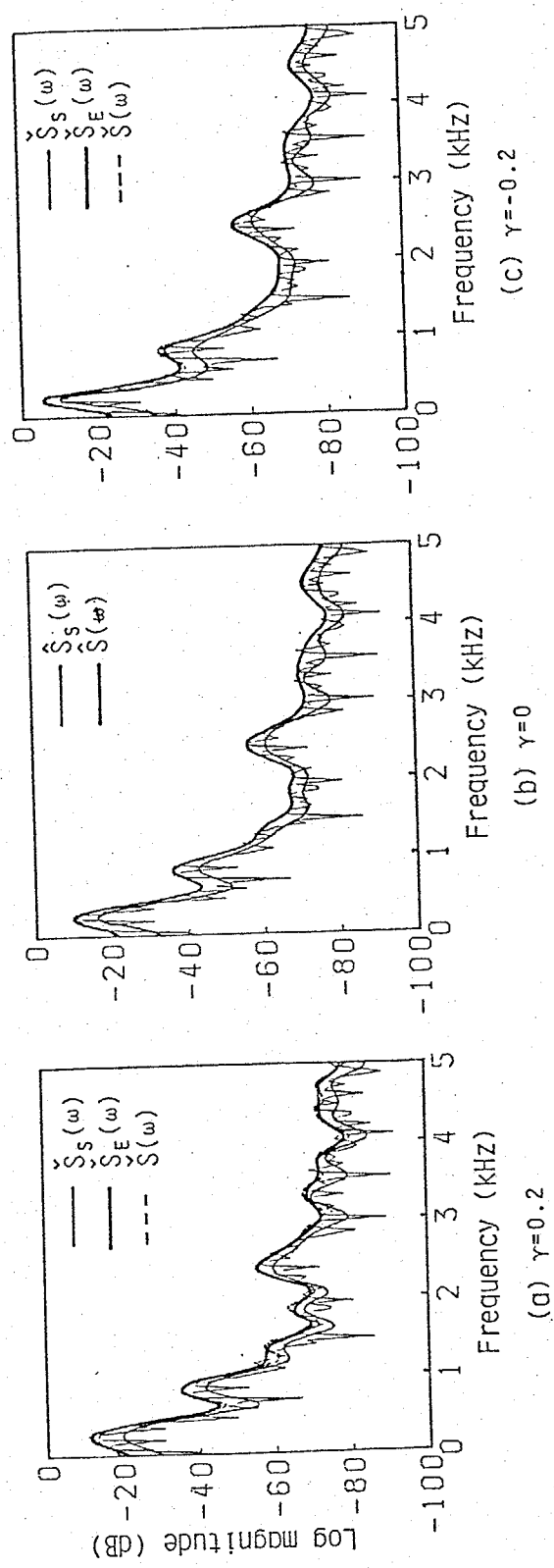


Fig. 2.6. Spectra, smoothed spectra and spectral envelopes for the beginning portion of /bu/ spoken in the phrase /nanbu/.

△である。また分析窓は 25.6 ms のブラックマン窓⁽¹⁶⁾を用いた。

図 2.7 は、正規化改良一般化ケプストラム \check{v}_2, \check{v}_3 および \check{v}_4 の分布をいくつかの γ の値について示したものである。なお \check{v}_1 については、式(2.49)により、任意の γ に対して $\check{v}_1 = \hat{c}_1$ が成立し、 γ の値による分布の違いは生じないことから、ここでは示さない。^{*} 図中で、 $\gamma=0$ は対数目盛の場合に相当し、改良ケプストラム \hat{c}_m の分布と等価である。図 2.7 からわかることは、同じ音声信号に対して γ が正の場合は、 γ が 0 の場合に比べて一般化ケプストラムの分布の分散が大きくなり、逆に γ が負の場合は小さくなるということである。また、平均値は γ が正の値に比べ負の値の方が小さくなっていることもわかる。 γ が正で 0.1 と 0.2 の場合を比べると、 $|\gamma|$ の値が大きい 0.2 の方が平均値、分散とも大きくなっている。 γ が負の場合、平均値は $|\gamma|$ の値が大きい -0.2 の方が小さくなっているが、分散は $\gamma = -0.1$ の方がわずかであるが小さくなっている。

図 2.8 には \check{v}_5, \check{v}_6 および \check{v}_7 、図 2.9 には \check{v}_8, \check{v}_9 および \check{v}_{10} 、図 2.10 には $\check{v}_{12}, \check{v}_{15}$ および \check{v}_{20} の分布をそれぞれ示す。一般化ケプストラムの次数が高くなってても、分散については低次数の項 \check{v}_2, \check{v}_3 および \check{v}_4 の分布と同様なことが言える。平均については、殆ど零の値に近く大きな差異はないが、 γ の値が正から負になるにつれて、平均値も負方向にずれていることがわかる。

図 2.11 は、 \check{v}_2 から \check{v}_{10} までの最大値及び最小値を示したものである。全体的には、最大値は γ が正の場合の方が大きくなっており γ が 0, -0.1, および -0.2 の場合はそれぞれ大きな差異はない。最小値の絶対値は、 $\gamma=0$ の場合が全体的に他の場合より小さくなっている。なお、正規化一般化ケプストラムの絶対値 $|\check{v}_m|$ の最大値を比較してみると、わずかの差であるが、 $\gamma = -0.1$ の場合が全体的に最も小さくなるという結果が得られている。

^{*} \hat{c}_1 の分布は第 6 章に示す。

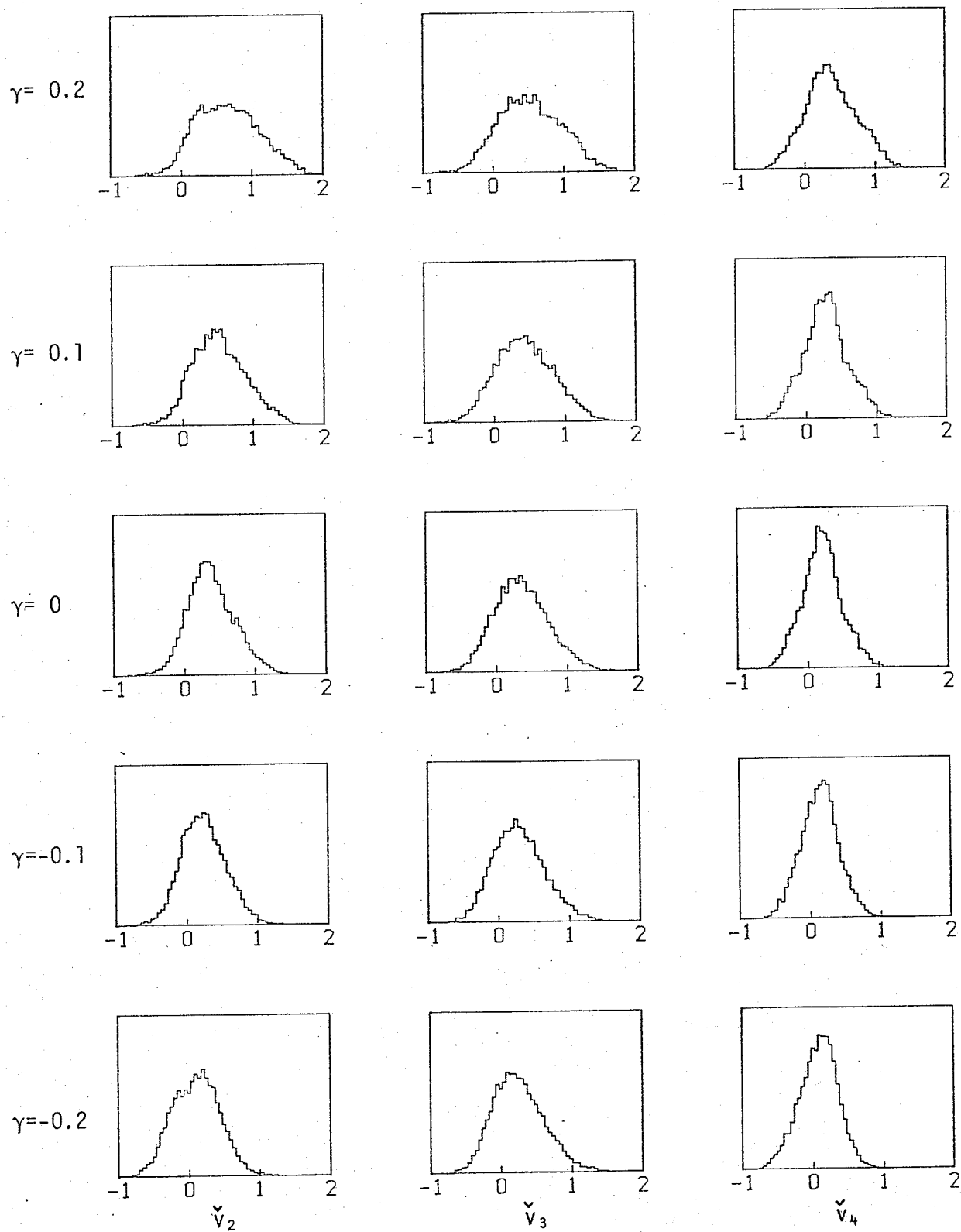
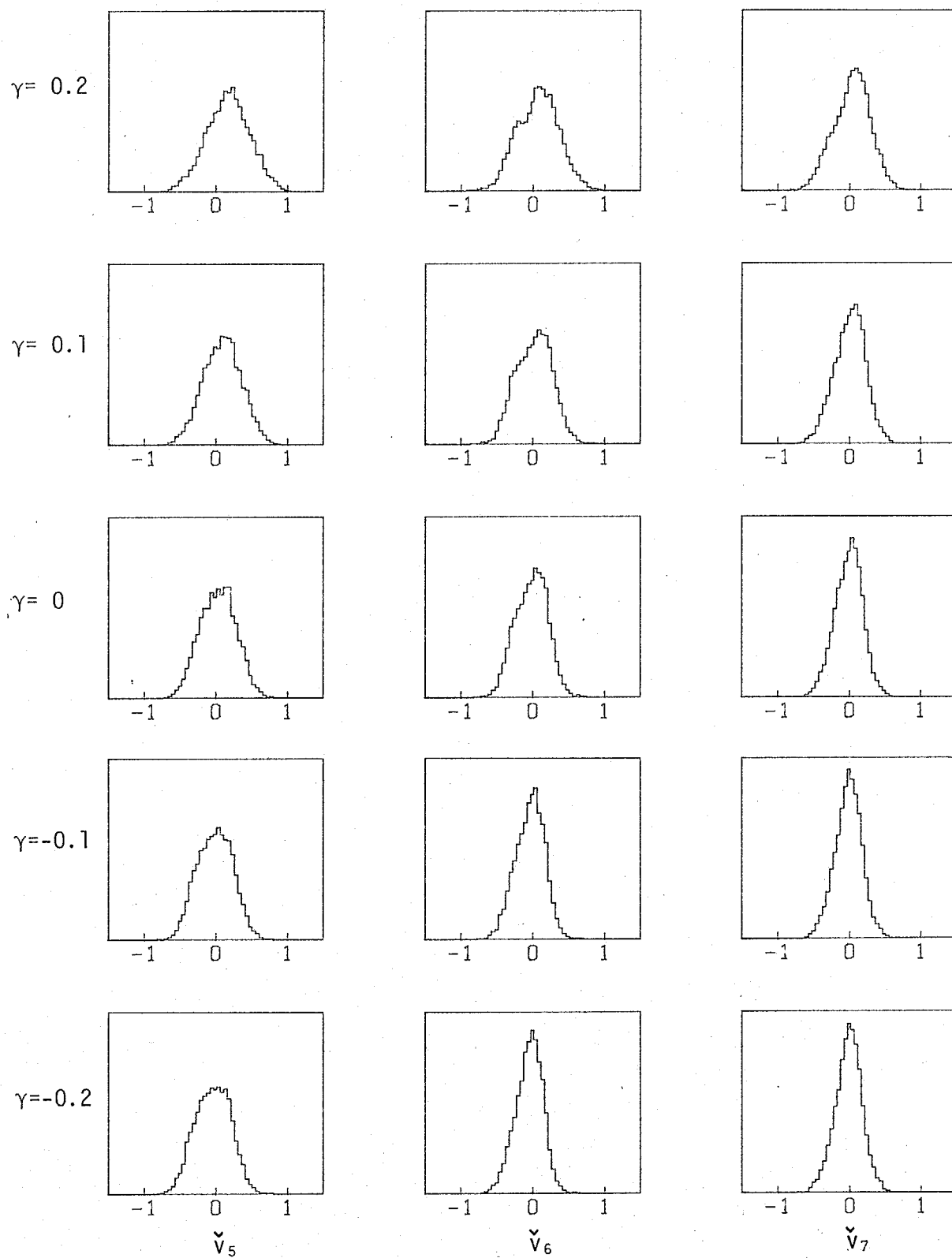


Fig. 2.7. Histograms of \check{v}_2 , \check{v}_3 and \check{v}_4 .

Fig. 2.8. Histograms of \check{V}_5 , \check{V}_6 and \check{V}_7 .

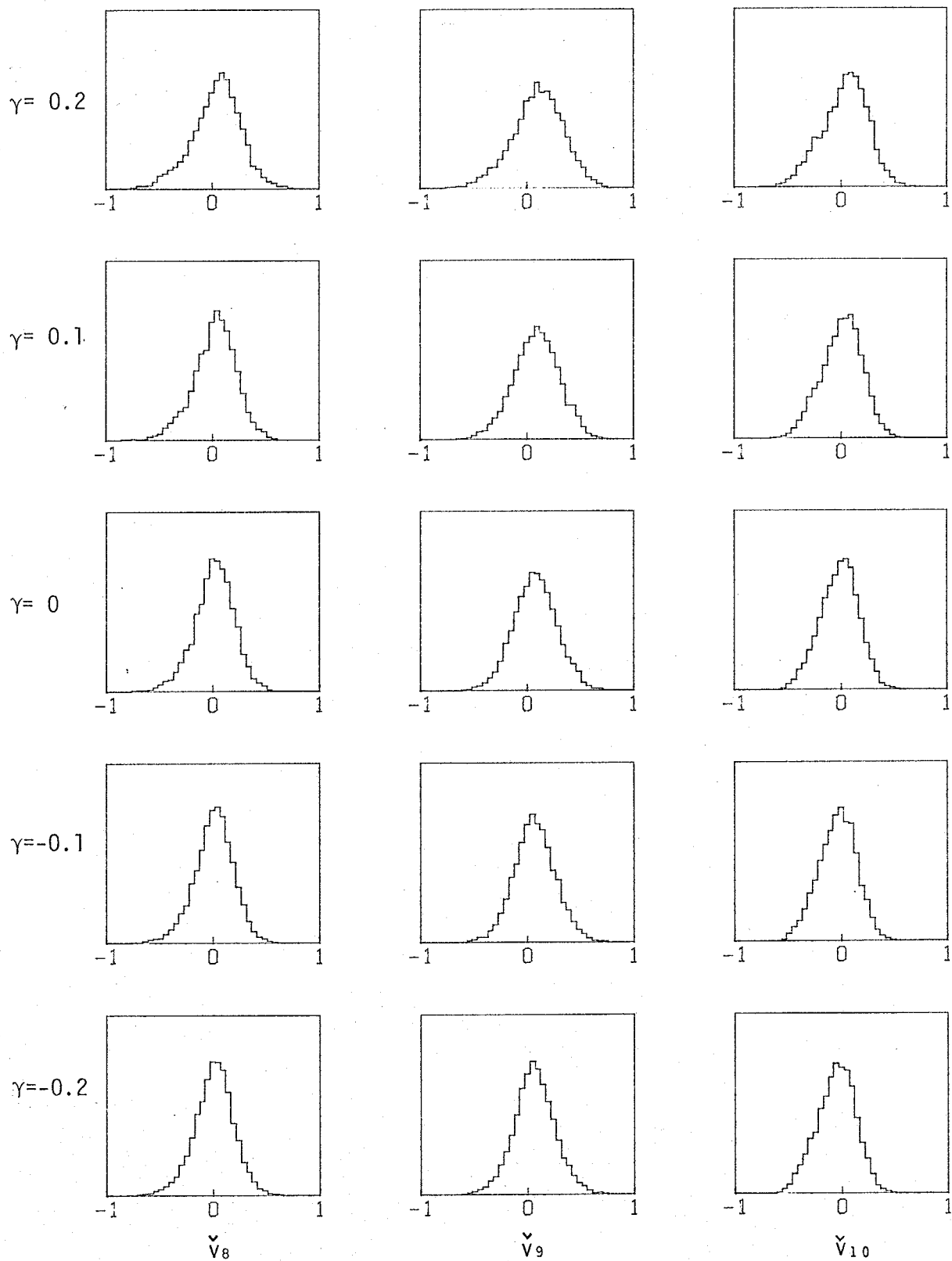
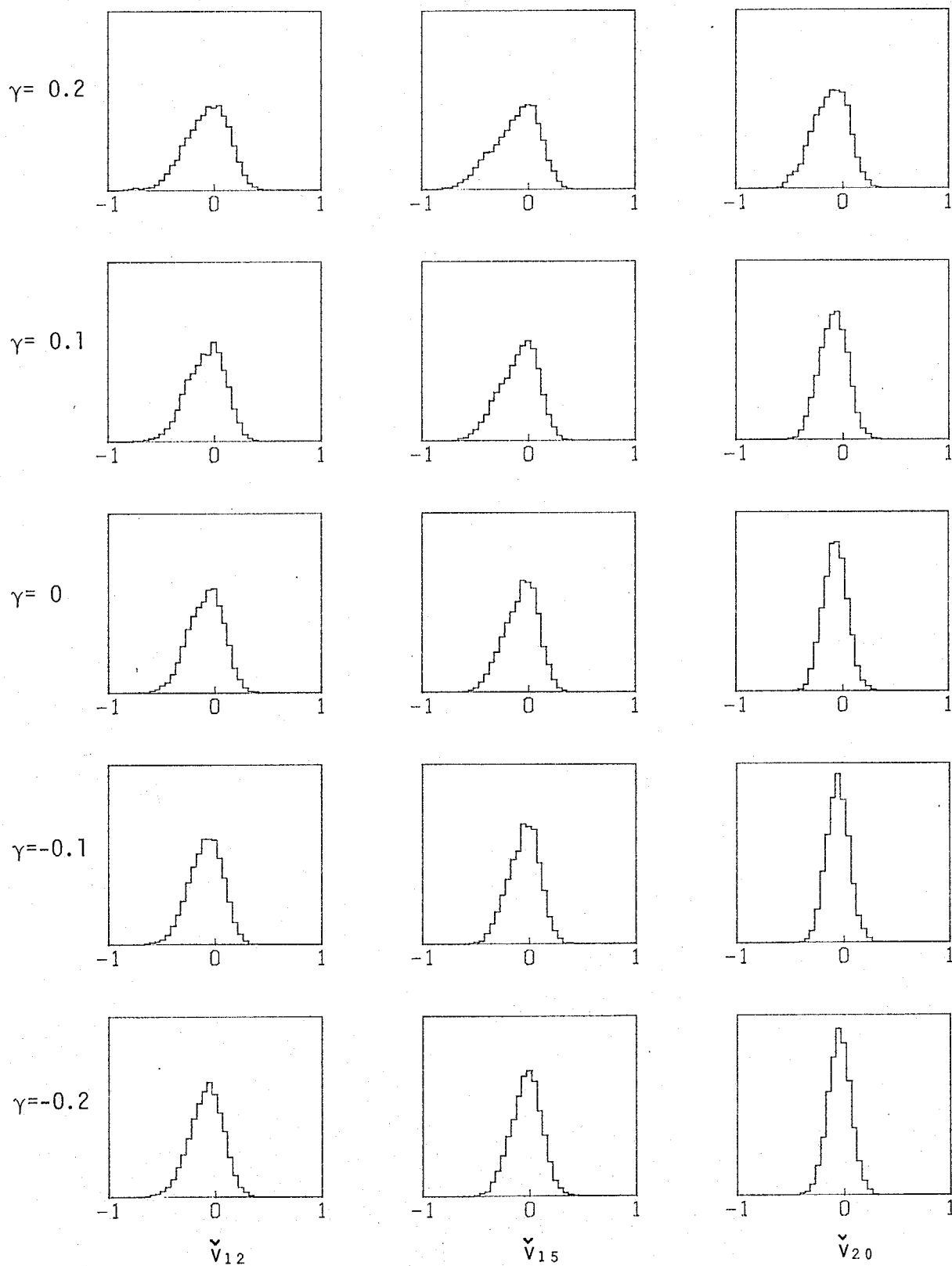


Fig. 2.9. Histograms of \check{V}_8 , \check{V}_9 and \check{V}_{10} .

Fig. 2.10. Histograms of \check{V}_{12} , \check{V}_{15} and \check{V}_{20} .

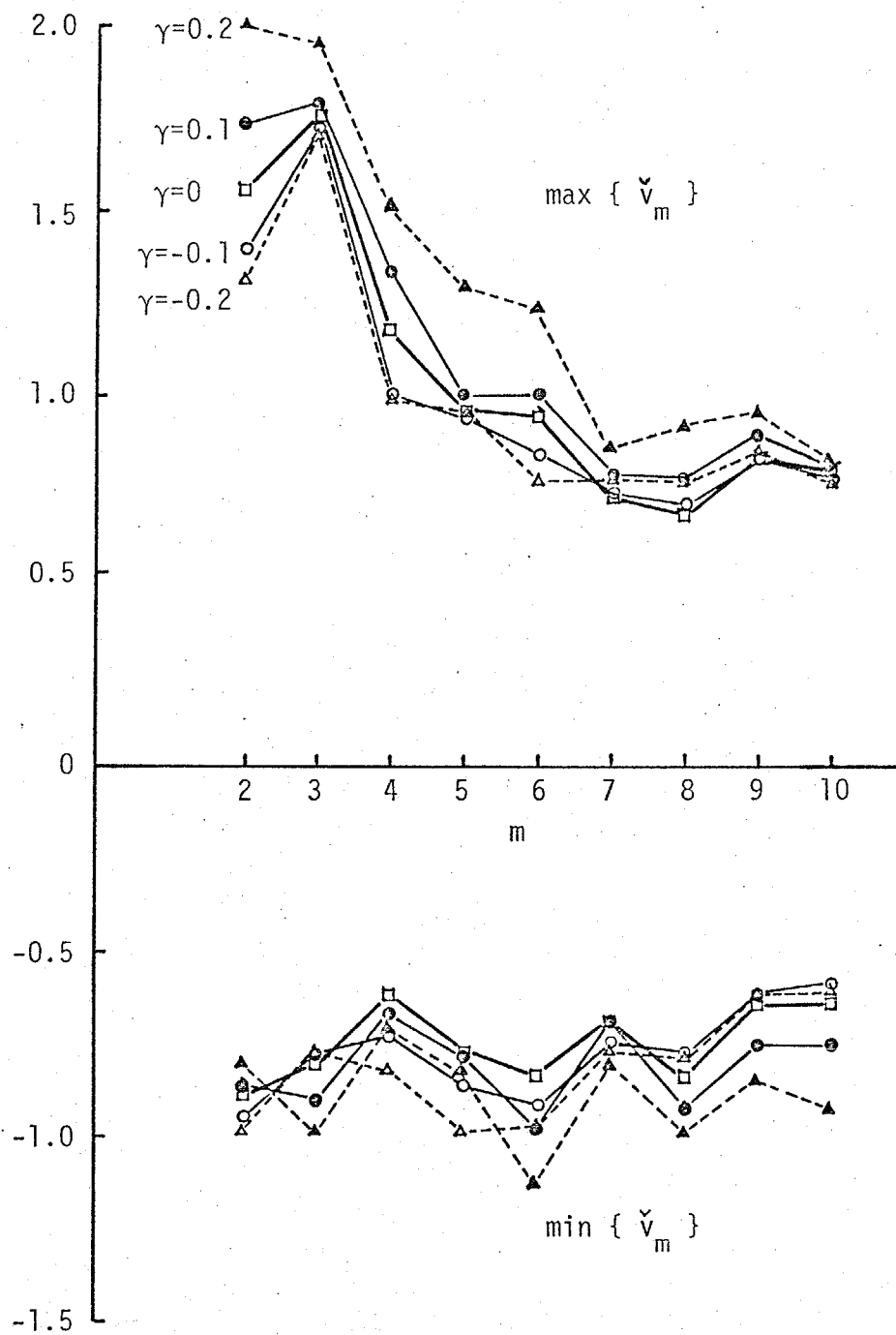


Fig. 2.11. Maximum and minimum values of generalized cepstra.

2.6 むすび

一般化対数関数を評価目盛とする音声のスペクトル包絡の抽出法を示した。一般化対数スペクトルのフーリエ係数で定義される一般化ケプストラムを用いることにより、同次数の通常のケプストラムでは表現することができなスペクトル包絡の表現が可能である。一般化ケプストラムを用いて平滑化されたスペクトルは、結果的にSRDSにより得られる平滑化スペクトルと等価となる。しかし、従来のケプストラムとの関連を考えると、一般化ケプストラムの方がより自然な一般化の表現となっている。また、改良ケプストラム法を適用して、真のスペクトル包絡を抽出するための評価目盛という観点からも、対数関数との対応がより自然である。さらに、一般化ケプストラムを用いれば、任意の α に対し、一般化対数目盛上でスペクトル包絡を直接近似する構成のフィルタを、合成フィルタとして利用できる特長がある。一般化対数特性を近似するフィルタについては、第5章で詳細に述べる。

一般化対数関数のパラメータ α をどのように選ばよいかという問題については、実際に合成音声を作成し、品質評価を行って結果を待たねばならない。

第3章 2乗平均誤差評価による音声合成フィルタ

の対数特性近似

3.1 まえがき

ケプストラムを用いる音声合成系では、ケプストラムによって表現された声道モデルの周波数特性を、いかによく近似するデジタルフィルタを実現することが、合成音声の品質を左右する重要な要因となる。ケプストラムは対数目盛におけるパラメータであることから、合成用デジタルフィルタを設計する際、その特性は直線目盛よりも対数目盛上で評価することが妥当であると考えられる。

これまで、任意の対数特性を近似するデジタルフィルタの設計法としては Johnson の手法⁽¹²⁾、対数振幅特性の直接近似の方法⁽⁸⁾、ケフレンシ領域での設計法⁽¹³⁾などが提案されている。これらの方法は、いずれも近似すべき対数特性のフーリエ級数展開を利用するものであり、フィルタ係数が比較的容易に求まる特長がある。しかし、Johnson の方法は、近似特性によっては非常に大きな誤差を生じるという問題点がある。これに対して、対数振幅特性の直接近似の方法やケフレンシ領域での設計法は、近似誤差を対数目盛上で規定できるため有効な方法である。その反面、基本フィルタをケプストラムの次数だけ連続接続するという構成のため、遅延素子数が通常の構成に比べて著しく大きくなる問題点がある。

本章では、2乗平均誤差評価を用いて、ケプストラムの次数と同一の遅延素子数の構成のデジタルフィルタの対数特性近似法を示す。

一般に、希望対数特性に対して、対数周波数特性が2乗平均誤差を最小にするようにデジタルフィルタの係数を決定することは非線形問題となる。とこ

るで、音声合成フィルタへの応用を考えた場合、フィルタ係数はできるだけ簡単に求めたいという要求があるから、非線形問題を反復計算により解くことは極力避ける必要がある。

以下に示す近似法では、対数目盛での2乗平均誤差評価式を、適当な重みをかけた直線目盛での2乗平均誤差評価式に置き換えることにより、線形問題の解としてデジタルフィルタの係数を得ている。重みは対数目盛上の誤差を直線目盛上で近似し、しかも解が簡単に求められるように選ばれているため、フィルタ係数は少ない計算量で得られる。IIRデジタルフィルタに本近似法を適用した場合は、常に安定で、位相特性が最小位相形と直線位相に近い形が同時に得られる特長がある。また逆フィルタも容易に得られる。

以下、本章では、まずFIRデジタルフィルタについて対数特性近似法を示し、具体的な計算アルゴリズムや近似例を示す。次に、IIRデジタルフィルタについて対数振幅特性近似法を示す。ここでは、フィルタの安定性やいくつかの近似例についても検討を行っている。

3.2 FIRディジタルフィルタの対数特性近似法

3.2.1 近似法

N 次のFIRディジタルフィルタの伝達関数を $H(z)$ とする。

$$H(z) = \sum_{k=0}^N h_k z^{-k} = h_0 + h_1 z^{-1} + \dots + h_N z^{-N} \quad (3.1)$$

希望対数周波数特性 $\hat{G}(\omega) = \ln G(\omega)$ が与えられたとき、対数目盛での2乗平均誤差を次式で定義する。

$$E^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{G}(\omega) - \hat{H}(e^{j\omega})|^2 d\omega \quad (3.2)$$

ここで $\hat{H}(e^{j\omega}) = \ln H(e^{j\omega})$ と表すものとする。

E^2 を最小にするようにディジタルフィルタ $H(z)$ を設計する場合、 $\hat{H}(e^{j\omega})$ が $\hat{G}(\omega)$ のフーリエ級数展開となっていればよい。 $\hat{H}(z) = \ln H(z)$ を級数展開したとき、フィルタ次数 N 次までの展開係数と、 $\hat{G}(\omega)$ のフーリエ係数すなわちケプストラムと等しくなるように h_k を決定するのがJohnsonの方法である。一方、 \hat{g}_k をケプストラムとして、対数周波数特性が

$$\hat{H}_k(e^{j\omega}) \approx \hat{g}_k \exp(-jk\omega) \quad (3.3)$$

となるように基本フィルタ $H_k(z)$ を構成し、これらを連続接続して近似するのがケフレンシ領域での設計法である。

ところで、式(3.2)の誤差評価式を用いて直接的にフィルタ係数 h_k を求めることは非線形問題である。そこで本近似法では、式(3.2)を適当な重み関数を選んで直線目盛での2乗平均誤差評価式

$$E^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega) - H(e^{j\omega})|^2 d\omega \quad (3.4)$$

で置き換えることを考える。ここで $W(\omega)$ は重み関数であり、次式の関係式を満足しているとする。

$$W(\omega) = W(-\omega), \quad 0 < W(\omega) < \infty \quad (3.5)$$

\hat{E}^2 と E^2 が等しくなるためには、式(3.2) および式(3.4)から形式的に重み関数は

$$W(\omega) = \frac{|\hat{G}(\omega) - \hat{H}(e^{j\omega})|^2}{|G(\omega) - H(e^{j\omega})|^2} \quad (3.6)$$

であればよい。しかし、このままでは重み関数に決定すべきフィルタ係数 h_k を含み、依然として非線形問題であるため適当でない。そこで、希望対数特性 $\hat{G}(\omega)$ のフーリエ級数展開を利用して次のような重みを導入する。

$$W(\omega) = \frac{|\hat{G}(\omega) - \ln \tilde{H}(\omega)|^2}{|G(\omega) - \hat{H}(\omega)|^2} \quad (3.7)$$

ここで

$$\ln \tilde{H}(\omega) = \sum_{k=0}^N \hat{g}_k e^{-j\omega k} \quad (3.8)$$

$$\hat{g}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{G}(\omega) e^{j\omega k} d\omega \quad (3.9)$$

とする。式(3.7) は、式(3.6) の $H(e^{j\omega})$ の項を $\hat{G}(\omega)$ のフーリエ級数展開の部分和で置き換えた形であり、 \hat{E}^2 と E^2 を近似的に等しくする一例と考えることができる。特に希望対数周波数特性がケプストラムで表されたスペクトル包絡

$$\hat{G}(\omega) = \hat{S}(\omega) = \sum_{m=0}^M \hat{c}_m e^{-j\omega m} \quad (3.10)$$

の場合には、 $M \leq N$ であれば式(3.7) は簡単になり、

$$W(\omega) = 1 / |G(\omega)|^2 \quad (3.11)$$

と表される。

式(3.7)は決定すべき係数 h_k によらない関数であるから、式(3.4)の誤差 E^2 を最小にする $H(z)$ を求めることは線形な最小2乗近似問題となり、容易に解くことが可能である。

3.2.2 係数の決定

式(3.7)を重み関数として、式(3.4)の E^2 を最小にするフィルタ係数 $h = (h_0, h_1, \dots, h_N)^T$ は次式の解で与えられる。

$$Ch = d \quad (3.12)$$

ここで、 $C = (C_{kl})_{k,l=0,1,\dots,N}$ は $(N+1)$ 次正方行列、 $d = (d_0, d_1, \dots, d_N)^T$ とし、 C, d の要素はそれぞれ

$$C_{kl} = C_{lk} = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) e^{j\omega(k-l)} d\omega, \quad k, l = 0, 1, \dots, N \quad (3.13)$$

$$d_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) G(\omega) e^{j\omega k} d\omega, \quad k = 0, 1, \dots, N \quad (3.14)$$

で与えられる。また、このときの E^2 の最小値 E_{min}^2 は、式(3.12)の解 h を用いて

$$E_{min}^2 = E_0^2 - h^T d \quad (3.15)$$

但し

$$E_0^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega)|^2 d\omega \quad (3.16)$$

で表される。

式(3.12)の連立方程式の係数行列 G は、式(3.13)から対称な Toeplitz 行列となることがわかる。行列 G について、実ベクトル $\mathbf{x} = (x_0, x_1, \dots, x_N)$ を変数とする二次形式 $F[\mathbf{x}]$ を作ると

$$F[\mathbf{x}] = \mathbf{x}^T G \mathbf{x} = \sum_{k=0}^N \sum_{l=0}^N C_{kl} x_k x_l$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \left| x_0 + x_1 e^{j\omega} + \dots + x_N e^{j\omega N} \right|^2 d\omega \quad (3.17)$$

が得られるが、重み関数 $W(\omega)$ に関しては式(3.5)の条件があることから、

$$\int_{-\pi}^{\pi} W(\omega) d\omega > 0 \quad (3.18)$$

が成立し、従って

$$F[\mathbf{x}] > 0, \quad \mathbf{x} \neq 0 \quad (3.19)$$

となる。これは G が正値であることを示しており、このことから式(3.12)の解が存在し、一意的にその値が定まることがわかる。

係数行列 G が対称な Toeplitz 行列であるという性質を使うと、式(3.12)の解法には、計算量が次数 $(N+1)$ の2乗程度のアルゴリズムを用いることができる⁽¹⁴⁾。行列 G の要素 C_{kl} は $|k-l|$ の値にのみ依存することから、

$$C_{|k-l|} = C_{kl} \quad (3.20)$$

と表すことにすれば、以下のアルゴリズムによりフィルタ係数 h_k が得られる。

(1) 初期化. $l := 0$ として

$$\begin{cases} \lambda_0 = c_0 \\ h_0^{(0)} = d_0 / \lambda_0 \end{cases} \quad (3.21a)$$

を代入する。

(2) $l := l + 1$ として, 次式の μ_l を計算する.

$$\mu_l = - \left(c_l + \sum_{k=1}^{l-1} \alpha_k^{(l-1)} c_{l-k} \right) / \lambda_{l-1} \quad (3.21b)$$

但し, $l=1$ の場合は $\mu_1 = -c_1 / \lambda_0$ とする.

(3) $k=1, 2, \dots, l$ について $\alpha_k^{(l)}$ を求める.

$$\begin{cases} \alpha_k^{(l)} = \alpha_k^{(l-1)} + \mu_l \alpha_{l-k}^{(l-1)}, & k=1, 2, \dots, l-1 \\ \alpha_l^{(l)} = \mu_l \end{cases} \quad (3.21c)$$

(4) λ_l を求める.

$$\lambda_l = (1 - \mu_l^2) \cdot \lambda_{l-1} \quad (3.21d)$$

(5) (3) で求めた $\alpha_k^{(l)}$ を用いて $h_k^{(l)}$ を計算する.

$$\begin{cases} h_l^{(l)} = \left(d_l + \sum_{k=0}^{l-1} \alpha_{l-k}^{(l)} d_k \right) / \lambda_l \\ h_k^{(l)} = h_k^{(l-1)} + h_l^{(l)} \alpha_{l-k}^{(l)}, & k=1, 2, \dots, l-1 \end{cases} \quad (3.21e)$$

(6) $l < N$ であれば (2) に戻り, $l = N$ ならば (7) に進む.

(7) 求める解 h_k は

$$h_k = h_k^{(N)}, \quad k=0, 1, \dots, N \quad (3.21f)$$

で与えられる.

これらの一連の計算に必要な演算回数は, 乗除算が $(2N^2 + 4N)$ 回, 加減算が $(2N^2 + N)$ 回である.

3.3 FIRデジタルフィルタを用いた近似例

任意の希望対数特性が与えられたとき、式(3.9)、(3.13)、(3.14)などの積分は解析的に行うことが困難な場合が多い。ところで、これらの式はいずれも逆フーリエ変換の定義式となっていることから、適当な周波数標本点を選んだ離散的逆フーリエ変換(IDFT)に置き換えてやれば、高速フーリエ変換(FFT)を用いて効率よく積分の計算が実行できる。IDFTを使って前述の \hat{g}_k , c_{kl} , d_k を求めることは、積分を数値積分の台形公式を用いて求めることと等価となる⁽¹⁵⁾。また、周波数標本点として

$$l\Delta\Omega = l \frac{2\pi}{L}, \quad l=0, 1, \dots, L-1 \quad (3.22)$$

と選び、式(3.4)の誤差評価式のかわりに、 $G_e = G(l\Delta\Omega)$, $W_e = W(l\Delta\Omega)$ として、

$$E_L^2 = \frac{1}{L} \sum_{l=0}^{L-1} W_e \left| G_e - H(e^{j2\pi l\Delta\Omega}) \right|^2 \quad (3.23)$$

を考え、 E_L^2 を最小にするデジタルフィルタを設計することとも等価となる⁽¹⁵⁾。周波数標本点数 L の値をフィルタの次数より十分大きく(例えば $L > 4N$)選んでおけば、連続的と離散的な場合で差異は殆ど問題とならない。

以下、FIR形デジタルフィルタを用いて、低域通過フィルタ、微分器の特性を近似した例を示す。

(1) 直線位相低域通過フィルタ

デジタルフィルタの伝達関数を

$$H(z) = z^{-N} \sum_{k=-N}^N h_{|k|} z^{-k} \quad (3.24)$$

と選んで、直線位相低域通過フィルタの特性

$$G(\omega) = |G(\omega)| e^{-jN\omega} \quad (3.25)$$

但し, $G_L = G(LA\Omega)$, $L = 256$ として

$$l_m |G_L| = \begin{cases} 0, & 0 \leq l \leq 63, 193 \leq l \leq 255 \\ -l_m 10, & l = 64, 192 \\ -2l_m 10, & 65 \leq l \leq 191 \end{cases} \quad (3.26)$$

を近似した例を図3.1に示す。フィルタ次数は $N = 10$ としてある。

フーリエ級数展開打ち切りの影響で生じるリッププルは、得られたフィルタ係数に適當な窓をかけることで軽減することかできる。ここでは、次式で示されるカイサ窓⁽¹⁶⁾

$$W(\alpha, k) = \frac{I_0 \left[\alpha \sqrt{1 - \left(\frac{k}{N}\right)^2} \right]}{I_0(\alpha)}, \quad -N \leq k \leq N \quad (3.27)$$

但し, $I_0(x)$ は零次第一種変形ベッセル関数

$$I_0(x) = \sum_{n=0}^{\infty} \left[\frac{\left(\frac{x}{2}\right)^n}{n!} \right]^2 \quad (3.28)$$

を用いて特性を改善した例を図3.2に示す。このときのフィルタ係数は、

$$h_k = W(\alpha, k) h_k, \quad -N \leq k \leq N \quad (3.29)$$

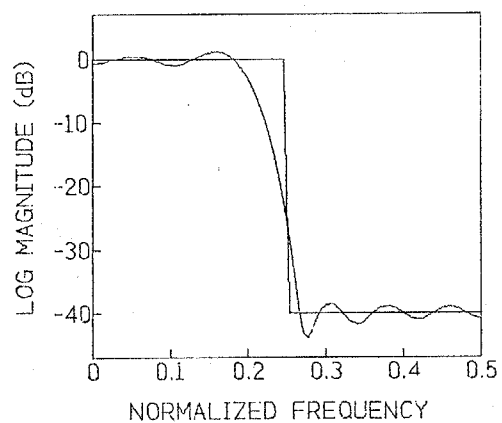
で与えられる。この例では, $\alpha = 3.0$ に選んでいる。

(2) 微分器

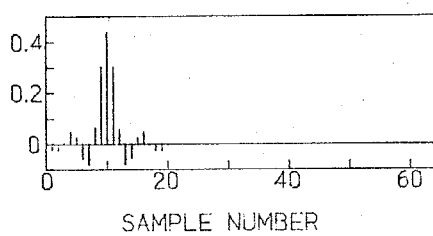
デジタルフィルタの伝達関数を

$$H(z) = z^{-N} \sum_{k=-N}^N h_k z^{-k}, \quad h_k = -h_{-k} \quad (3.30)$$

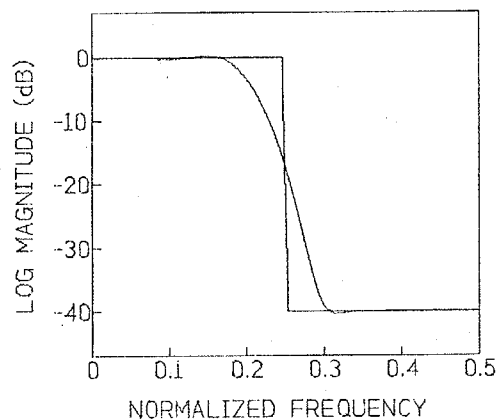
として、微分器の特性



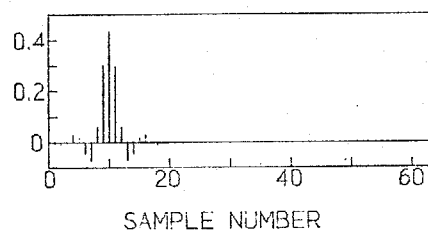
(a) Log magnitude response



(b) Impulse response



(a) Log magnitude response



(b) Impulse response

Fig. 3.1. Characteristics of the linear phase lowpass filter with $N=10$.

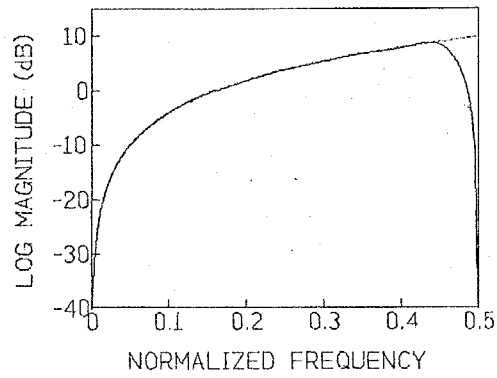
Fig. 3.2. Characteristics of the linear phase lowpass filter after windowing with $N=10$ and $\alpha=3.0$.

$$G_l = \begin{cases} j \frac{2\pi}{L} l \cdot e^{-jNl\Delta\Omega} & 0 \leq l \leq 127 \\ \pi, & l = 128 \\ -j \frac{2\pi}{L} (L-l) \cdot e^{-jNl\Delta\Omega}, & 129 \leq l \leq 255 \end{cases} \quad (3.31)$$

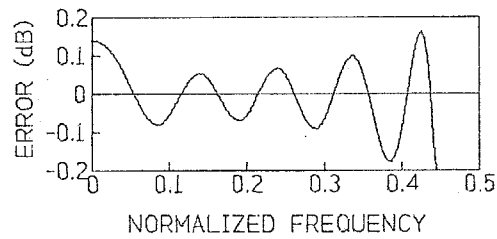
但し, $L = 256$

を近似した結果を図 3.3 に示す。なお実際の計算では, $l \rightarrow -\infty$ ($l=0$)
 を避けるため, $G_0 = \delta$ ($\delta > 0$, ここでは $\delta \approx 0.01$) として計算してい
 る。この例でも, リップル分を減少させる目的で, カイザ窓 ($\alpha = 3.0$) をかけ
 てある。

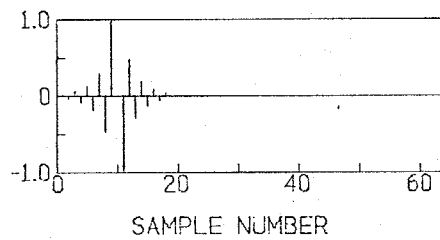
一般に, カイザ窓を FIR フィルタの特性改善のために用いる場合, α の値
 は 5.0 ~ 9.0 の範囲で選ばれることが多いが, ここでは経験的に, $\alpha = 2.0 \sim$
 4.0 程度で十分実用となる特性が得られるようである。



(a) Log magnitude response



(b) Log magnitude error



(c) Impulse response

Fig. 3.3. Characteristics of the differentiator after windowing. ($N=10$, $\alpha=3.0$)

3.4 IIRデジタルフィルタの対数振幅特性近似法

3.4.1 近似法

N 次のデジタルフィルタの伝達関数を $H(z)$ とし、その分母および分子多項式をそれぞれ $A(z)$, $B(z)$ とする。

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^N b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} \quad (3.32)$$

希望対数特性 $\hat{G}(\omega) = \ln G(\omega)$ が与えられたとき、対数目盛での2乗平均誤差 E^2 を

$$E^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{G}(\omega) - \hat{H}(e^{j\omega})|^2 d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\ln\{G(\omega) \cdot A(e^{j\omega})\} - \ln B(e^{j\omega})|^2 d\omega \quad (3.33)$$

で定義すれば、 E^2 を最小にするフィルタ係数 a_k, b_k を求めることは非線形問題となる。ここでは、FIRデジタルフィルタの近似法と同様にして、式(3.33)を適当な重みをかけた直線目盛での2乗平均誤差評価式

$$E^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega) A(e^{j\omega}) - B(e^{j\omega})|^2 d\omega \quad (3.34)$$

で置き換えることを考える。ここで $W(\omega)$ は重み関数であり、式(3.5)の関係を満たしているものとする。

重み関数 $W(\omega)$ が決定すべきフィルタ係数 a_k, b_k に依らない関数であれば、 E^2 を最小にする $A(z), B(z)$ を求めることは、線形最小2乗近似問題となり容易に解くことができる。すなわち、式(3.34)を各係数 a_k, b_k について微分し、得られた式を零とおけば、 $a = (a_1, a_2, \dots, a_N)^T$ および $b = (b_0, b_1, \dots, b_N)^T$ に対する次の方程式が得られる。

$$\begin{bmatrix} V & T \\ T^T & U \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} t \\ u \end{bmatrix} \quad (3.35)$$

ここで, $U = (u_{kl})_{k,l=1,\dots,N}$, $V = (v_{kl})_{k,l=0,1,\dots,N}$ はそれぞれ N 次および $(N+1)$ 次正方形行列, $T = (t_{kl})_{k=0,\dots,N, l=1,\dots,N}$ は $(N+1) \times N$ 次の行列, $u = -(u_1, u_2, \dots, u_N)^T$, $t = (t_0, t_1, \dots, t_N)^T$ とし, それぞれの要素は次式で与えられる。

$$u_{kl} = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega)|^2 e^{j\omega(k-l)} d\omega \quad (3.36)$$

$$v_{kl} = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) e^{j\omega(k-l)} d\omega \quad (3.37)$$

$$t_{kl} = -\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) G(\omega) e^{j\omega(k-l)} d\omega \quad (3.38)$$

$$u_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega)|^2 e^{j\omega k} d\omega \quad (3.39)$$

$$t_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) G(\omega) e^{j\omega k} d\omega \quad (3.40)$$

また, E^2 の最小値 E_{\min}^2

$$E_0^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |G(\omega)|^2 d\omega \quad (3.41)$$

とすれば, 式(3.35)の解 a, b を用いて

$$E_{\min}^2 = E_0^2 - \begin{bmatrix} b^T, a^T \end{bmatrix} \begin{bmatrix} t \\ u \end{bmatrix} \quad (3.42)$$

と書くことができる。

3.4.2 重み関数

式(3.34)に用いる重み関数は、 E^2 が \hat{E}^2 のできるだけ良い近似を与えることが望ましい。一方、係数決定の観点からは、式(3.35)が簡単に解ける形となることが望ましい。

E^2 と \hat{E}^2 が等しくなるためには、式(3.33)、(3.34)から形式的に

$$W(\omega) = \frac{|\ln G(\omega)A(e^{j\omega}) - \ln B(e^{j\omega})|^2}{|G(\omega)A(e^{j\omega}) - B(e^{j\omega})|^2} \quad (3.43)$$

であればよい。しかし、式(3.43)には決定すべき係数 a_k, b_k を含んでいるから、このままでは適当でない。そこで、FIR形の近似法と同様、希望対数振幅特性のフーリエ級数展開を利用して、次式の重み関数を導入する。

$$W(\omega) = \frac{|\ln |G(\omega)|\tilde{A}(\omega) - \ln \tilde{B}(\omega)|^2}{||G(\omega)|\tilde{A}(\omega) - \tilde{B}(\omega)|^2} \quad (3.44)$$

ここで、

$$\ln \tilde{B}(\omega) = -\ln \tilde{A}(\omega) = \frac{1}{2} \ln \tilde{G}(\omega) \quad (3.45)$$

$$\ln \tilde{G}(\omega) = \sum_{k=-N}^N \hat{g}_k e^{-j\omega k} \quad (3.46)$$

$$\hat{g}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |G(\omega)| e^{j\omega k} \quad (3.47)$$

とる。式(3.45)から

$$\ln \tilde{G}(\omega) = \ln \left\{ \tilde{B}(\omega) / \tilde{A}(\omega) \right\} \quad (3.48)$$

が成立しており、式(3.44)の重み関数は式(3.43)の $A(e^{j\omega}), B(e^{j\omega})$ の項を希望対数振幅特性 $\ln |G(\omega)|$ のフーリエ級数展開の部分項を等分した特性で置き換えた形となっている。特に希望対数振幅特性がケプストラム \hat{C}_m

により,

$$\hat{G}(\omega) = \hat{S}(\omega) = \sum_{m=0}^M \hat{C}_m \cos m\omega \quad (3.49)$$

の形で表されていて, $M \leq N$ であれば

$$\ln \tilde{G}(\omega) = \ln |G(\omega)| \quad (3.50)$$

となり, 式(3.44)は次式のように簡単な形となる。

$$W(\omega) = 1 / \tilde{G}(\omega) \quad (3.51)$$

重み関数として, 式(3.44)あるいは式(3.51)を用いることの妥当性については, 近似誤差の観点から議論の余地があるものと考えられる。しかしながら, 以下に示すように, 式(3.51)を重み関数に用いることで係数決定の手続きが単純化され, 安定性の保証, 最小位相フィルタへの変換等, 利点が多いことから, 十分目的に合ったものである。

3.4.3 係数の決定

希望対数振幅特性が与えられれば, 式(3.44)により重み関数が決まり, 式(3.35)の連立方程式を解くことにより, デジタルフィルタの係数が決定できる。式(3.35)の係数行列は対称であれば, 次数は $(2N+1)$ となるため N が大きくなった場合の計算量はかなり大きくなる。

これに対して, 希望対数振幅特性をケプストラムで表された式(3.49)の形に限定し, さらに適当な直線位相項を付加することで, 係数の決定手続きを非常に単純化することができる。

希望対数振幅特性を

$$\ln \tilde{G}(\omega) = \sum_{m=0}^N \hat{C}_m \cos m\omega = \sum_{k=-N}^N \hat{g}_k e^{j\omega k} \quad (3.52)$$

とするとき, 希望特性として直線位相を加えた

$$G(\omega) = \tilde{G}(\omega) \cdot e^{-jN\omega} \quad (3.53)$$

を考えることにする。この場合、重み関数は式(3.51)となり、これから式(3.36)~(3.40)の値を計算すると次式が得られる。

$$u_{kl} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}(\omega) e^{j\omega(k-l)} d\omega = u_{|k-l|}, \quad k, l = 1, 2, \dots, N \quad (3.54)$$

$$v_{kl} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\tilde{G}(\omega)} e^{j\omega(k-l)} d\omega = v_{|k-l|}, \quad k, l = 0, 1, \dots, N \quad (3.55)$$

$$t_{kl} = -\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega(k-l-N)} d\omega = 0, \quad k=0, 1, \dots, N, l=1, 2, \dots, N \quad (3.56)$$

$$t_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega(k-N)} d\omega = \begin{cases} 0, & k=0, 1, \dots, N-1 \\ 1, & k=N \end{cases} \quad (3.57)$$

ここで

$$u_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}(\omega) e^{j\omega k} d\omega \quad (3.58)$$

$$v_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\tilde{G}(\omega)} e^{j\omega k} d\omega \quad (3.59)$$

とする。従って、式(3.35)も簡単になり、

$$\begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} B \\ a \end{bmatrix} = \begin{bmatrix} \# \\ u \end{bmatrix} \quad (3.60)$$

となる。ここで、

$$U = \begin{pmatrix} u_0 & u_1 & \dots & u_{N-1} \\ u_1 & u_0 & & \\ \vdots & & \ddots & \\ u_{N-1} & \dots & & u_0 \end{pmatrix} \tag{3.61}$$

$$V = \begin{pmatrix} v_0 & v_1 & \dots & v_N \\ v_1 & v_0 & & \\ \vdots & & \ddots & \\ v_N & \dots & & v_0 \end{pmatrix} \tag{3.62}$$

$$\# = (0, \dots, 0, 1)^T \tag{3.63}$$

とする。式(3.60)は、 γ_i ルタ係数 a_k, b_k がそれぞれ独立に求められることを示しており、結局係数の決定方程式は

$$\begin{cases} U a = u & (3.64a) \\ V b = \# & (3.64b) \end{cases}$$

となる。
式(3.64a), (3.64b) の係数行列は、いずれも対称な Toeplitz 行列であり、希望対数振幅特性が有限な値であれば、

$$0 < \tilde{G}(\omega) < \infty \tag{3.65}$$

であることから、3.2節の式(3.17)~(3.19)と同様な関係を使って、 U, V が正値であることを証明できる。従って、式(3.64a), (3.64b) の解 a, b はそれぞれ一意的に定まる。

U, V が対称な Toeplitz 行列であるという性質を使うと、式(3.64a), (3.64b) の解法として、次のような計算量の少ないアルゴリズムを導くことができる⁽¹⁴⁾。

式(3.64)の解法

(1) 初期化. $l := 0$ とし

$$\begin{cases} \lambda_0 = u_0 \\ \sigma_0 = v_0 \end{cases} \quad (3.66a)$$

とおく.

(2) $l := l + 1$ とし, 次式の μ_l, τ_l を求める.

$$\mu_l = - \left(u_l + \sum_{k=1}^{l-1} a_k^{(l-1)} u_{l-k} \right) / \lambda_{l-1} \quad (3.66b)$$

$$\tau_l = - \left(v_l + \sum_{k=1}^{l-1} b_k^{(l-1)} v_{l-k} \right) / \sigma_{l-1} \quad (3.66c)$$

但し, $l=1$ の場合は, $\mu_1 = -u_1/\lambda_0$, $\tau_1 = -v_1/\sigma_0$

(3) μ_l, τ_l から次式により $a_k^{(l)}, b_k^{(l)}$ $k=1, 2, \dots, l-1$ について計算する.

$$\begin{cases} a_k^{(l)} = a_k^{(l-1)} + \mu_l a_{l-k}^{(l-1)}, & k=1, 2, \dots, l-1 \\ a_l^{(l)} = \mu_l, \end{cases} \quad (3.66d)$$

$$\begin{cases} b_k^{(l)} = b_k^{(l-1)} + \tau_l b_{l-k}^{(l-1)}, & k=1, 2, \dots, l-1 \\ b_l^{(l)} = \tau_l \end{cases} \quad (3.66e)$$

(4) λ_l, σ_l を求める.

$$\lambda_l = (1 - \mu_l^2) \lambda_{l-1} \quad (3.66f)$$

$$\sigma_l = (1 - \tau_l^2) \sigma_{l-1} \quad (3.66g)$$

(5) $l < N$ ならば (2) に戻り, $l = N$ ならば (6) に進む.

(6) 求める解 a_k, b_k は次式で与えられる.

$$a_k = a_k^{(N)}, \quad k = 1, 2, \dots, N \quad (3.66h)$$

$$\begin{cases} b_k = b_{N-k}^{(N)} / \sigma_N, & k = 0, 1, \dots, N-1 \\ b_N = 1 / \sigma_N \end{cases} \quad (3.66i)$$

式(3.66)の計算に必要な演算回数は、 N^2 回程度となる。

3.5 安定性

3.5.1 安定性

式(3.64)で表される決定方程式から求められた $\{a_k\}$, $\{b_k\}$ を係数とする IIR デジタルフィルタが安定であることを以下に示す。

式(3.66d)により得られる $\{a_k^{(l)}\}$ を係数とする多項式 $A_l(z)$ を次のように定義する。

$$A_l(z) = 1 + a_1^{(l)} z^{-1} + \cdots + a_l^{(l)} z^{-l} \quad (3.67)$$

さらに、この $A_l(z)$ から多項式 $\phi_l(z)$ を作る。

$$\phi_l(z) = \lambda_l^{-\frac{1}{2}} z^l A_l(z) \quad (3.68)$$

ここで、 λ_l は式(3.66f)で計算される値とする。このとき多項式系 $\{\phi_l(z)\}$ には次の性質がある。⁽¹⁷⁾

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{G}(\omega) \phi_k(e^{j\omega}) \overline{\phi_l(e^{j\omega})} d\omega = \delta_{kl} \quad (3.69)$$

ここで、 δ_{kl} はクロネッカーデルタ、 $\overline{\cdot}$ は複素共役を表すものとする。式(3.69)は、多項式系 $\{\phi_l(z)\}$ が z 平面の単位円上で、重み $\tilde{G}(\omega)$ に関して正規直交系をなしていることを示している。

ところで、直交多項式の理論によれば、 $\{\phi_l(z)\}$ の根はすべて z 平面の単位円内に存在することが知られている。⁽¹⁸⁾ 従って、 $A_l(z)$ の根も z 平面の単位円内に存在する。さらに、 $A_l(z) = A(z)$ であるから、 $A(z)$ の根すなわちデジタルフィルタの伝達関数の極はすべて z 平面の単位円内に存在することになり、安定であることがわかる。

同様にして、式(3.66e)により得られる $\{b_k^{(e)}\}$ から多項式

$$B_l(z) = 1 + b_1^{(e)} z^{-1} + \cdots + b_l^{(e)} z^{-l} \quad (3.70)$$

を作れば、 $B_1(z)$ の根はすべて z 平面の単位円内に存在することが示される。

とすると、 $B(z)$ と $B_N(z)$ の間には式(3.66)から

$$B(z) = \sigma_N^{-1} z^{-N} B_N(z^{-1}) \quad (3.71)$$

という関係が成立する。従って、 $B(z)$ の根、すなわち $H(z)$ の零点はすべて z 平面の単位円外に位置することになる。

3.5.2 最小位相推移形フィルタ

デジタルフィルタの伝達関数 $H(z)$ の位相特性は、希望特性として式

(3.53) に示した直線位相項を含んだ形を仮定したため、結果的に直線位相に近い形が得られる。フィルタの振幅特性だけを問題にする場合、同じ振幅特性であれば、遅延量の少ない最小位相形が有利なことがある。

$H(z)$ の零点がすべて z 平面の単位円外に存在するという前述の性質を利用すれば、フィルタ係数 b_k の順序を並べかえるだけで最小位相形のデジタルフィルタが構成できる。 $H(z)$ に対して、同一振幅特性をもつ最小位相形の伝達関数を $H_{min}(z)$ とすれば

$$H_{min}(z) = \frac{z^{-N} B(z^{-1})}{A(z)} = \frac{b_N + b_{N-1} z^{-1} + \dots + b_0 z^{-N}}{1 + a_1 z^{-1} + \dots + a_N z^{-N}} \quad (3.72)$$

で与えられる。

音声合成フィルタへの応用を考えた場合、この最小位相形の構成は有用であると考えられる。

伝達関数 $H(z)$ に対して $|H(z)| \cdot |H^{-1}(z)| = 1$ となるような逆フィルタ $H^{-1}(z)$ も容易に構成可能である。すなわち、次式の伝達関数

$$\begin{aligned} H^{-1}(z) &= \frac{1}{H(z^{-1})} = \frac{a_N + a_{N-1} z^{-1} + \dots + z^{-N}}{b_N + b_{N-1} z^{-1} + \dots + b_0 z^{-N}} \\ &= \sigma_N \frac{a_N + a_{N-1} z^{-1} + \dots + z^{-N}}{1 + b_1^{(N)} z^{-1} + \dots + b_N^{(N)} z^{-N}} \end{aligned} \quad (3.73)$$

を作れば、 $H^{-1}(z)$ の振幅特性は $H(z)$ の振幅特性の逆特性となる。ここで、 ω_k は式(3.66g)、 $b_k^{(n)}$ は式(3.66e)によってそれぞれ計算される値である。 $H^{-1}(z)$ の極が z 平面の単位円内に存在し、従って安定であることは容易に確かめることができる。なお、最小位相形の伝達関数 $H_{min}(z)$ に対する逆フィルタは、 $H_{min}(z)$ の分子と分母を逆にすることで得られる。最小位相形であるから当然 $H_{min}(z)$ の逆フィルタも安定となる。

3.6 IIRデジタルフィルタを用いた近似例

希望特性が与えられたとき、フィルタ係数を求めるためには式(3.58)および式(3.49)の積分を計算する必要があるが、解析的に行うことができる場合には、3.3節で述べたFIR形と同様な取扱いをすればよい。すなわち、適当な周波数標本点を式(3.22)のように選び、式(3.58)、(3.49)のフーリエ逆変換式をIDFTに置き換えればよい。 L の値を2のべき乗に選んでおけばFFTを利用して効率よく計算ができる。また、このとき得られるフィルタ係数は、次式の誤差を最小にする⁽¹⁵⁾。

$$E_L^2 = \frac{1}{L} \sum_{k=0}^{L-1} W_k |G_k A(e^{j2\pi k \Delta \Omega}) - B(e^{j2\pi k \Delta \Omega})|^2 \quad (3.74)$$

ここで、 $G_k = G(\Delta \Omega)$ 、 $W_k = W(\Delta \Omega)$ 、 $\Delta \Omega = 2\pi/L$ とする。

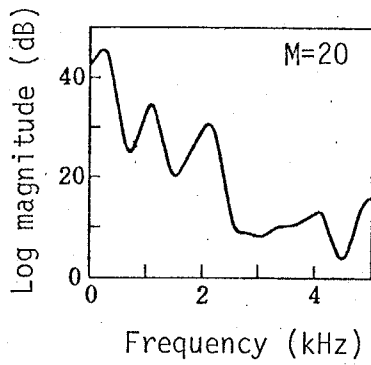
以下、具体的な近似例を示す。

(1) 音声のスペクトル包絡の近似

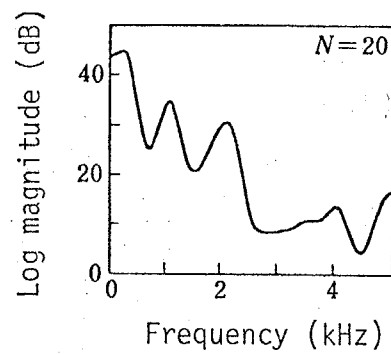
図3.4(a)は、対数スペクトル包絡の一例であり、図3.4(b)に示す /jekohama/ の /md/ の音声波形の最初の部分から、ケプストラムの次数 $M=20$ として抽出したものである。図3.5(a)は、この対数スペクトル包絡をフィルタ次数 $N=20$ 、周波数の標本点数 $L=64$ として、本近似法により得られたデジタルフィルタの対数振幅特性である。相対的に小さなフォールマントもよく近似しており、この例の対数目盛でのRMS誤差は0.39 [dB] である。

図3.4の例のように、希望対数振幅特性が滑らかな特性のときには、周波数の標本点数 L の値は64程度でも十分な近似が得られる。実際に、 $L=256$ として同じ希望特性を近似したときのRMS誤差は、 $L=64$ の場合と比べて 10^{-3} [dB] 以下の差異しか現れない。

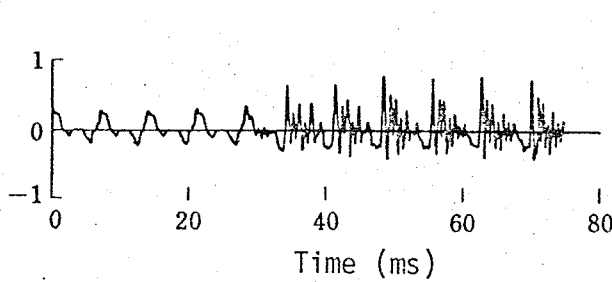
図3.5(b)は、得られたデジタルフィルタを合成フィルタに用いて、ケプストラム法により合成した音声波形である。なお、位相特性は最小位相形と



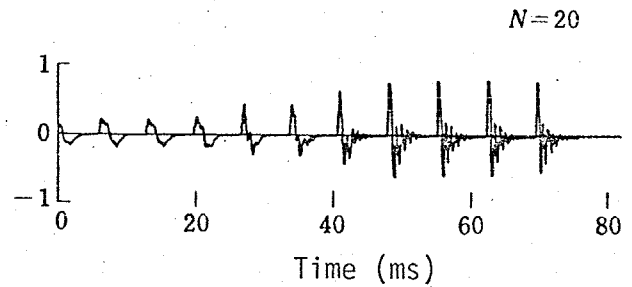
(a)



(a)



(b)



(b)

Fig. 3.4. (a) Spectral envelope for the beginning portion of /ma/, (b) waveform for the portion /ma/ in the original speech signal.

Fig. 3.5. (a) Log magnitude response of the digital filter for speech synthesis, (b) waveform for the portion /ma/ in the synthesized speech.

してある。対数振幅特性の直接近似による方法⁽⁸⁾から得られたデジタルフィルタを合成フィルタとした合成音と比較聴取した結果、殆どその区別はできなかつた。但し、このときの遅延素子数は、本近似法により得られたフィルタの方が $1/10$ 以下と大幅に少なくなっている。

(2) 低域通過フィルタ

希望対数振幅特性として

$$\ln |G_l| = \begin{cases} 0, & 0 \leq l \leq 63, 193 \leq l \leq 255 \\ -\ln 10, & l=64, 192 \\ -2 \ln 10, & 65 \leq l \leq 191 \end{cases} \quad (3.75)$$

但し、 $L=256$

を考える。本近似法を適用するために、 $\ln |G_l|$ の I D F T によりケプストラム \hat{g}_k を求め、式(3.52)から $\ln \tilde{G}_l = \ln \tilde{G}(l\Delta\Omega)$ を計算することにする。図3.6は、 $N=10$ としたとき得られるデジタルフィルタの特性を示す。FIR形の近似特性と比べるとリップル分も少なく、良好な近似特性が得られる。

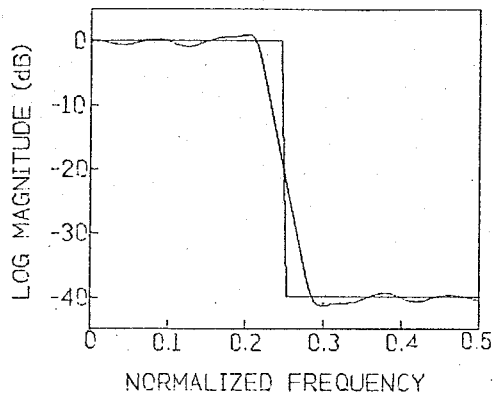
FIR形の場合と同様、デジタルフィルタの係数に適當な窓をかけることで近似特性を改善できる。図3.7は、3.3節で示したカイサ窓 $W(\alpha, k)$ を用いて、まず、ケプストラムを

$$\tilde{g}_k = W(\alpha_1, k) \cdot \hat{g}_k, \quad -N \leq k \leq N \quad (3.76)$$

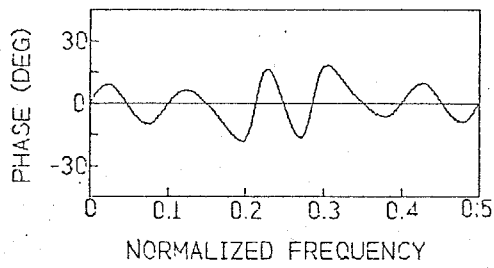
と修正し、次にこの \tilde{g}_k をもとにして得られるフィルタ係数 a_k, b_k を

$$\begin{cases} \bar{a}_k = W(\alpha_2, k) \cdot a_k, & k=1, 2, \dots, N \\ \bar{b}_k = W(\alpha_2, N-k) \cdot b_k, & k=0, 1, \dots, N \end{cases} \quad (3.77)$$

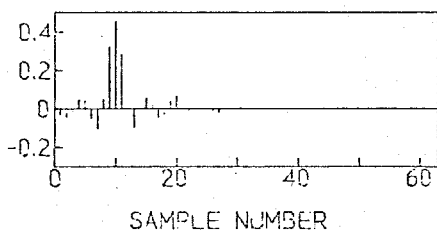
と修正したときのデジタルフィルタの特性である。ここで $\alpha_1 = 1.5, \alpha_2 = 3.0$ と選んである。遷移帯域の幅が広がるが、振幅特性は大きく改善され、位相特性も遷移帯域を除いてほぼ直線となる。



(a) Log magnitude response

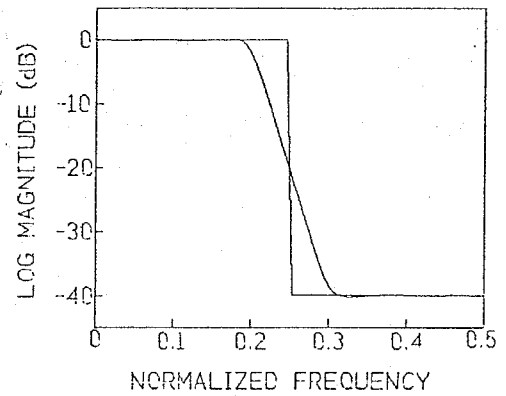


(b) Phase response

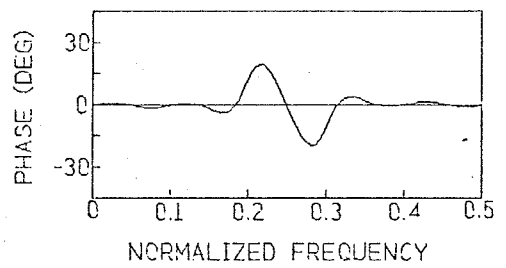


(c) Impulse response

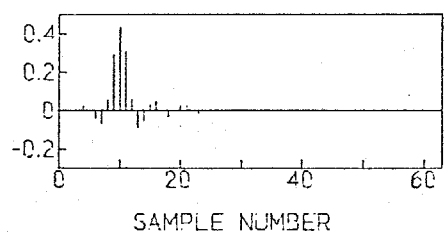
Fig. 3.6. Characteristics of the lowpass filter with $N=10$.



(a) Log magnitude response



(b) Phase response



(c) Impulse response

Fig. 3.7. Characteristics of the lowpass filter after windowing with $N=10$, $\alpha_1=1.5$ and $\alpha_2=3.0$.

3.7 むすび

2乗平均誤差評価を用いたデジタルフィルタの対数周波数特性の近似法を示した。本近似法は、FIR形およびIIR形の両方に適用することができる。いずれの場合もデジタルフィルタの係数は簡単な連立方程式の解で与えられ、その解法には計算量の少ないアルゴリズムが利用できる。特に、IIRデジタルフィルタの場合、常に安定で、位相特性は直線位相に近い形と最小位相形が同時に得られることを示した。また、安定な逆フィルタも簡単に得られる特長がある。

本近似法の問題点は、対数目盛上の誤差に対して厳密な規定ができないことである。しかし、いくつかの近似例で示したように、実用上は十分な近似特性が得られると考えてよい。実際に、IIRデジタルフィルタを音声合成フィルタとして用いた場合、ケプストラムの次数と同一の遅延素子数の構成で、精度良く対数スペクトル包絡の近似ができる。また、合成系を構成して得られた合成音声も、従来用いられている直接近似形のフィルタと比べて、殆ど差異が認められないことが確かめられている。

第4章 最小位相インパルス応答を用いる

音声合成フィルタの特性近似

4.1 まえがき

第3章では、デジタルフィルタの遅延素子数がケプストラム次数と同一であるような対数特性近似法を示した。この近似法は、対数特性の直接近似の方法⁽⁸⁾に比べ、大幅な遅延素子数の減少が可能であり、実用上十分な近似特性が得られる。しかし、一般にケプストラム法は線形予測法と比較すると、同じ程度の明瞭度・品質を得るために必要なパラメータの次数が高いため、ケプストラム次数と同一の遅延素子数の合成フィルタでも、線形予測法で用いられる合成フィルタより遅延素子数や乗算器の数が大きくなってしまふ。

このような問題に対し、Oppenheimらは、ケプストラムで表された対数振幅特性に対して最小位相インパルス応答を求め、この系列に対してLPCの手法を適用することにより、ケプストラムの次数より低い次数の音声合成モデルで表現しようとする試みを報告している。⁽¹⁹⁾ この方法では、LPCによりケプストラム次数より低次数の極を求め、次に逆LPCの方法を用いて零点を決定している。この結果得られた極-零モデルの対数振幅特性は、もとのケプストラムで表された対数振幅特性とはかなり異なるのが普通である。また、LPCを2回適用することから計算量も大きくなる。

本章では希望対数振幅特性の最小位相インパルス応答に対して、時間領域のデジタルフィルタの一設計法を適用することにより、ケプストラム次数と同じか、あるいはそれより低次数で、良好な近似特性が得られることを示す。

時間領域でのデジタルフィルタの設計法はこれまで数多く提案されているが、一般には周波数領域での誤差との関係が単純でないため、得られた近似特

性, 特に対数周波数特性は大きな誤差を生じやすい。また, 本質的に非線形問題であるため, 計算量も多くなる場合が多い。本章で利用する設計法は修正最小2乗近似問題⁽²⁰⁾であり, この方法は, 周波数領域との関係が単純で, 解を得るための計算量も設計するデジタルフィルタの次数の2乗程度と少ないなどの特長がある。

与えられた周波数特性に対して, 修正最小2乗近似問題を適用して近似特性を得る方法として Scharf⁽²¹⁾ の方法がある。しかし, この方法では, 高次のARモデルのパラメータを求める手続きが必要であり, そのための計算量が多くなるという問題点がある。

本章で示す近似法は, 希望対数振幅特性のケプストラムから計算される最小位相インパルス応答を理想インパルス応答として選び, この系列に対して修正最小2乗近似問題を適用している。このため, 高次のARモデルを構成するという煩雑な手順は必要としない。設計すべきデジタルフィルタはIIR形を仮定しているが, 特にFIR形を仮定した場合には, 本近似法は Johnson 法⁽¹²⁾と一致する。また, 希望特性が最小位相系の一般化ケプストラムで表された一般化対数特性の近似にも適用できる。

以下, 本章ではまず基本となる対数振幅特性の近似法を示し, 次に一般化対数特性の近似への応用を述べる。最後に具体的な近似例もいくつか示している。

4.2 対数振幅特性近似法

4.2.1 修正最小2乗近似問題

設計すべきデジタルフィルタの伝達関数 $H(z)$ を

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{l=0}^L b_l z^{-l}}{1 + \sum_{k=1}^K a_k z^{-k}} \quad (4.1)$$

とする。ここで K, L はそれぞれ伝達関数 $H(z)$ の分母および分子多項式の次数である。このデジタルフィルタのインパルス応答 $\{h_m\}_0^{\infty}$ は、次式によって与えられる。

$$h_m = \begin{cases} b_m - \sum_{k=1}^K a_k h_{m-k}, & 0 \leq m \leq L \\ - \sum_{k=1}^K a_k h_{m-k} & m > L \end{cases} \quad (4.2)$$

理想インパルス応答 $\{g_m\}_0^{\infty}$ が与えられたとき、時間領域での最小2乗近似問題は、次式の2乗誤差

$$E_0^2 = \sum_{m=0}^M (g_m - h_m)^2 \quad (4.3)$$

を最小にするフィルタ係数 $\{a_k\}, \{b_l\}$ を求めることである。式(4.3)は g_m のフーリエ変換を $G(\omega)$ とすれば、周波数領域では

$$E_0^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(\omega) - H(e^{j\omega})|^2 d\omega \quad (4.4)$$

と表すことができる。

式(4.3)によって表される2乗誤差 E_0^2 を最小にすることは非線形問題であり、反復的な手法を用いて直接的に解くか、または式(4.3)を適当な誤差評価式に置き換えることにより線形問題として取り扱うか、どちらかを選択しなければならない。

ここでは、計算量が少なく取り扱いが簡単な後者の方法を選び、次のような修正最小二乗近似問題を考える。すなわち、 $0 \leq m \leq L$ の範囲では、理想インパルス応答とデジタルフィルタのインパルス応答が正確に一致するようにし、 $L < m$ の範囲で次の二乗誤差

$$E^2 = \sum_{m=L+1}^M e_m^2$$

$$= \sum_{m=L+1}^M \left(g_m + \sum_{k=1}^K a_k g_{m-k} \right)^2 \quad (4.5)$$

を最小にするようにフィルタ係数を決定する。

式(4.3)を式(4.2)を使って、 $L < m$ の範囲について書き直すと

$$E_o^2 = \sum_{m=L+1}^M \left(g_m + \sum_{k=1}^K a_k h_{m-k} \right)^2 \quad (4.6)$$

となるから、式(4.5)は式(4.6)の h_m の項を g_m で置き換えたものと考えることができ。また、式(4.5)は図4.1に示すように、理想インパルス応答を、 $A(z)$ という伝達関数をもつフィルタに入力したときの出力系列 e_m の、 $L < m$ の範囲におけるパワーと解釈することもできる。

上記のようにデジタルフィルタの係数を決定することは、周波数領域では次式の二乗平均誤差

$$E^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(\omega)A(e^{j\omega}) - B(e^{j\omega})|^2 d\omega \quad (4.7)$$

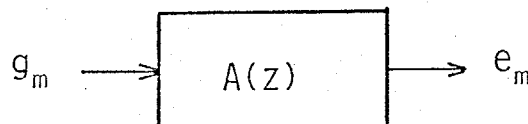


Fig. 4.1. Error for the modified least-squares problem.

を最小にすることと等価である。

なお、式(4.5)の2乗誤差の総和の上限を、適当な正の有限な整数値とすれば、これは Burns らが提案した時間領域の設計法⁽²²⁾で用いる誤差評価式と等しくなる。

4.2.2 係数の決定

式(4.5)を a_k について微分して零とおくことにより、 E^2 を最小にするフィルタ係数を求めるための次式の方程式が得られる。

$$C\mathbf{a} = -\mathbf{d} \quad (4.8)$$

ここに、 $\mathbf{a} = (a_1, a_2, \dots, a_K)^T$ 、 $\mathbf{d} = (d_1, d_2, \dots, d_K)^T$ 、および $C = (C_{kl})_{k,l=1, \dots, K}$ なる K 次正方行列とし、要素 C_{kl} 、 d_k はそれぞれ $g_m = 0, m < 0$ として、次式で与えられる。

$$C_{kl} = \sum_{m=L+1}^M g_{m-k} g_{m-l}, \quad k, l = 1, 2, \dots, K \quad (4.9)$$

$$d_k = \sum_{m=L+1}^{\infty} g_m g_{m-k}, \quad k = 1, 2, \dots, K \quad (4.10)$$

式(4.9)、(4.10)は、理想インパルス応答の自己相関系列

$$\begin{aligned} r_k &= \sum_{m=0}^{\infty} g_m g_{m+k} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(\omega)|^2 e^{j\omega k} d\omega \end{aligned} \quad (4.11)$$

を使うと、

$$C_{kl} = r_{|k-l|} - \sum_{m=0}^L g_m g_{m-k}, \quad k, l = 1, 2, \dots, K \quad (4.12)$$

$$d_k = r_k - \sum_{m=0}^L g_m g_{m-k}, \quad k = 1, 2, \dots, K \quad (4.13)$$

と書くことができる。式(4.12), (4.13)を使えば, 理想インパルス応答を無限長にわたって知る必要はなく, $\{g_0, g_1, \dots, g_L\}$ および自己相関係列 $\{r_0, r_1, \dots, r_k\}$ という $(L+K+2)$ 個のデータがわかっているならば, c_{ℓ} , d_k を求めることができる。

式(4.8) を解くことにより, デジタルフィルタの分母多項式の係数 a_k が求まる。次に, $0 \leq m \leq L$ の範囲で, 理想インパルス応答とフィルタのインパルス応答が一致するよう, すなわち

$$g_m = h_m, \quad 0 \leq m \leq L \quad (4.14)$$

が成立するようにするには, 式(4.8) の解 a_k を用いて,

$$b_{\ell} = \begin{cases} g_0, & \ell = 0 \\ g_{\ell} + \sum_{k=1}^{\ell} a_k g_{\ell-k}, & \ell = 1, 2, \dots, L \end{cases} \quad (4.15)$$

とすればよい。

式(4.8) の解法としては, 計算量が方程式の次数 K の 2 乗程度である効率の良いアルゴリズムを用いることができる。^{*} また得られたフィルタの安定性も保証される。⁽²⁰⁾

4.2.3 最小位相インパルス応答

デジタルフィルタの時間領域の設計法では, 希望周波数特性が与えられた際, 理想インパルス応答をどのように設定するかが一つの問題となる。一般には, 直線目盛で希望周波数特性を逆フーリエ変換して, インパルス応答とすることが多い。

ここでは, 対数振幅特性が与えられたとき, そのケプストラムから計算される最小位相インパルス応答を理想インパルス応答に用いることにする。

希望対数振幅特性を $\hat{G}(\omega) = \ln |G(\omega)|$ とし, そのケプストラムを \hat{g}_m

*付録 A-1 参照

とする。

$$\hat{g}_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{G}(\omega) e^{j\omega m} d\omega \quad (4.16)$$

このとき、対数振幅特性が $\hat{G}(\omega)$ である最小位相系のケプストラム \hat{c}_m は、

$$\hat{c}_m = \begin{cases} 0, & m < 0 \\ \hat{g}_0, & m = 0 \\ 2\hat{g}_m, & m > 0 \end{cases} \quad (4.17)$$

で与えられ、そのインパルス応答（最小位相インパルス応答） g_m は

$$g_m = \begin{cases} 0, & m < 0 \\ \exp(\hat{c}_0), & m = 0 \\ \sum_{k=1}^m \left(\frac{k}{m}\right) \hat{c}_k g_{m-k}, & m > 0 \end{cases} \quad (4.18)$$

で与えられる。⁽¹⁰⁾

本近似法では、式(4.18)で表される系列を理想インパルス応答とし、前述の修正最小二乗近似問題を適用して係数を決定することになる。なお、設計すべきデジタルフィルタがFIR形の場合は、単に式(4.14)が成立するよりに b_e を決めればよいことから

$$b_e = g_e, \quad e = 0, 1, \dots, L \quad (4.19)$$

となる。特に、 g_m を最小位相インパルス応答に選んでいることから、これは Johnson 法⁽¹²⁾と一致する。

4.3 一般化対数特性の近似法

希望特性が対数振幅特性に限らず、最小位相系の一般化ケプストラムで表された一般化対数特性に対しても、同様の近似法を考えることができる。

一般化対数目盛上の希望特性を $\check{G}(\omega)$ とし、一般化ケプストラム \check{C}_m により

$$\check{G}(\omega) = \sum_{m=0}^{\infty} \check{C}_m e^{-j\omega m} \quad (4.20)$$

で与えられているとする。このとき、 $\check{G}(\omega)$ が最小位相系であれば、式(2.26)から、インパルス応答は \check{C}_m を用いて次式で表される。

$$g_m = \begin{cases} 0, & m < 0 \\ S_0^{-1}(\check{C}_0), & m = 0 \\ \check{C}_1 \cdot \frac{g_0}{1 + r\check{C}_0}, & m = 1 \\ \check{C}_m \cdot \frac{g_0}{1 + r\check{C}_0} \\ \quad + \frac{1}{1 + r\check{C}_0} \sum_{k=1}^{m-1} \binom{k}{m} \{ \check{C}_k g_{m-k} - r g_k \check{C}_{m-k} \}, & m > 1 \end{cases} \quad (4.21)$$

式(4.21)を理想インパルス応答として、修正最小2乗近似問題を適用すれば、式(4.5)の E^2 を最小にするフィルタ係数が得られる。

4.4 近似例

フィルタ係数 a_k を求めるために、式(4.12), (4.13) で表される c_{ke}, d_k を計算する必要がある。 c_{ke}, d_k を計算するためには、理想インパルス応答 $\{g_0, g_1, \dots, g_L\}$ と、自己相関係列 $\{r_0, r_1, \dots, r_k\}$ がわかっている必要がある。 r_k の値は式(4.11) で定義されるが、この式は逆フーリエ変換の定義式となっていることから、適当な周波数標本点を選んでIDFTに置き換えてやれば、FFTを利用して効率よく r_k の値を計算できる。このとき得られる値は、式(4.11)に数値積分の台形公式を適用して計算することと等価である。^{*}

以下に、本近似法を用いて、いくつかの特性を近似した例を示す。

(1) 音声の対数スペクトル包絡の近似

図4.2(a)に、男声 /na/ の /a/ の部分のスペクトル包絡の一例を示す。ケプストラム次数は $M=20$ である。また、図4.2(b)は、同図(a)から得られる最小位相インパルス応答である。この希望特性に対し、デジタルフィルタの次数を $L=K=18$ と選んで近似した結果を図4.3に示す。

図4.3(a)はフィルタの対数振幅特性であり、同図(b)はインパルス応答である。このときの対数振幅目盛上での最大誤差は 0.475 [dB], RMS 誤差は 0.193 [dB] である。音声のスペクトル包絡を近似する場合、通常フィルタの次数をケプストラムの次数と同じに選んでおけば、実用上十分な近似特性が得られるが、本近似例のように、ケプストラム次数より低いフィルタの次数で、良好な近似特性が得られる場合もある。一般には、最小位相インパルス応答の収束が速いような対数振幅特性程、低いフィルタ次数でも近似誤差は小さくなる。

(2) 低域通過フィルタ

ケプストラムが

* 3.3 節参照

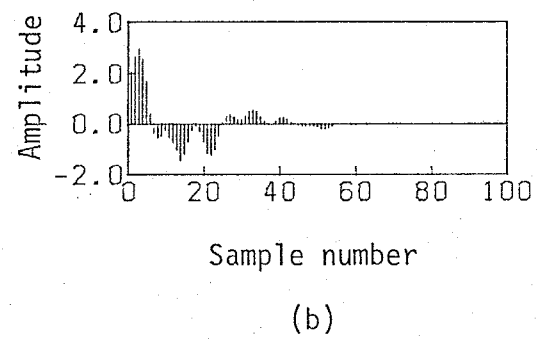
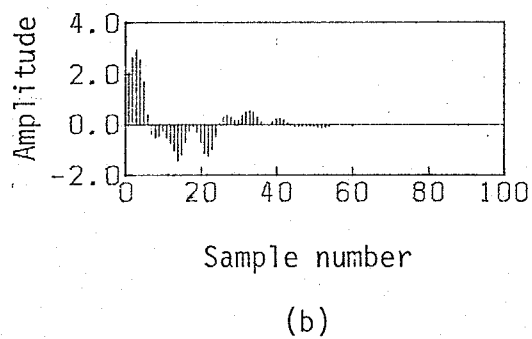
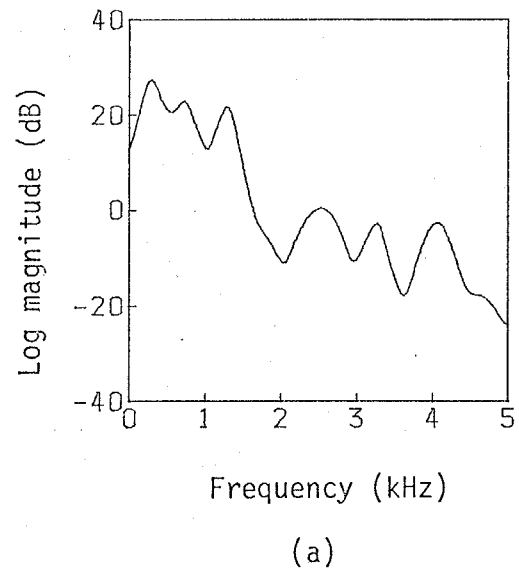
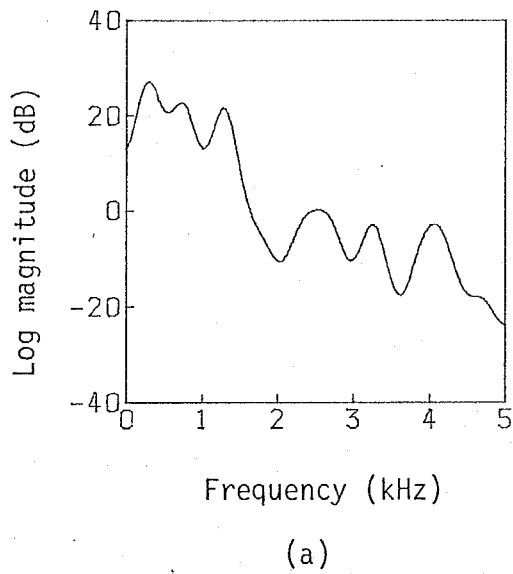


Fig.4.2. (a) Log spectral envelope of Japanese vowel /a/ ($M=20$), (b) minimum phase impulse response.

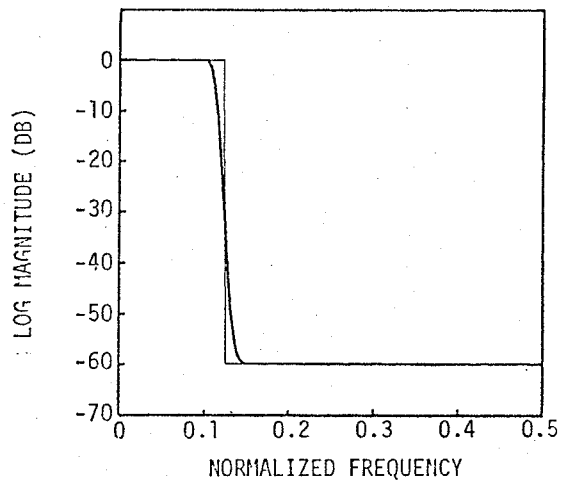
Fig.4.3. Characteristics of the obtained digital filter ($L=K=18$): (a) log magnitude response, (b) impulse response.

$$\hat{g}_m = \begin{cases} \beta \left(\frac{\omega_0}{\pi} - 1 \right), & m=0 \\ \frac{\beta}{\pi m} \cdot \sin m\omega_0 \cdot W(\alpha, m), & 1 \leq |m| \leq M \quad (4.22) \\ 0, & \text{otherwise} \end{cases}$$

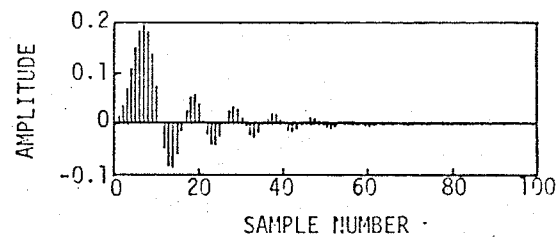
て表される低域通過フィルタの特性を考える。ここで β は減衰量、 ω_0 は遮断周波数を表す。また、 $W(\alpha, m)$ はカイザ窓であり、式(3.27)で定義される。図4.4は、 $\beta = 60$ [dB]、 $\omega_0 = 0.25\pi$ 、 $M = 50$ 、 $\alpha = 7.5$ と選んだときの、対数振幅特性および最小位相インパルス応答である。

この特性を、 $L = 50$ の FIR フィルタ、すなわち Johnson 法で近似した結果が図4.5である。フィルタ次数をケプストラム次数と同一に選んでいるにもかかわらず、良い近似特性が得られない。これに対して、図4.6は、フィルタの次数を $L = K = 10$ として本近似法によって得られた近似特性である。このときの RMS 誤差は、0.315 [dB] と、非常に良い近似を与える。

理想低域通過フィルタのまうに、



(a) Log magnitude response



(b) Minimum phase impulse response

Fig. 4.4. Desired characteristics of the lowpass filter.

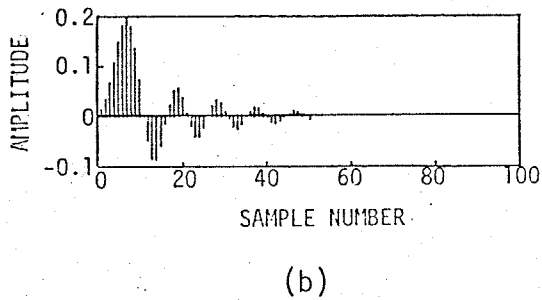
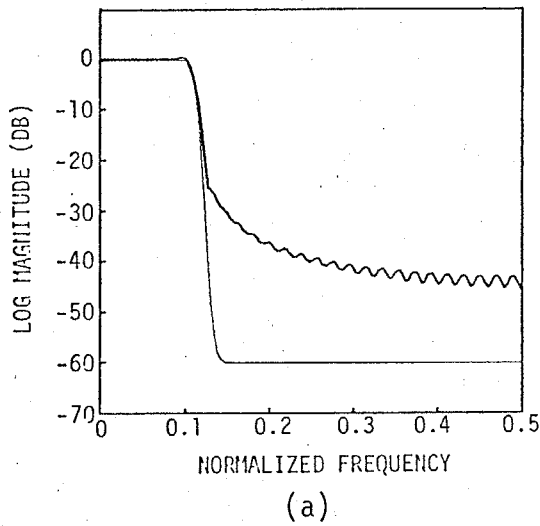


Fig.4.5. Characteristics of the obtained digital filter by Johnson method with $L=50$:
 (a) log magnitude response,
 (b) impulse response.

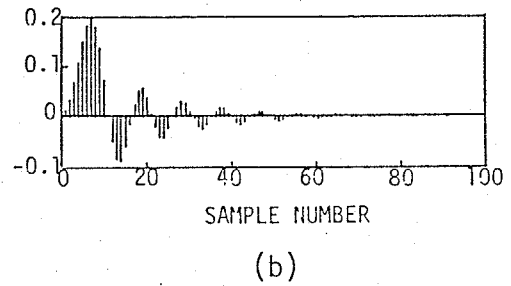
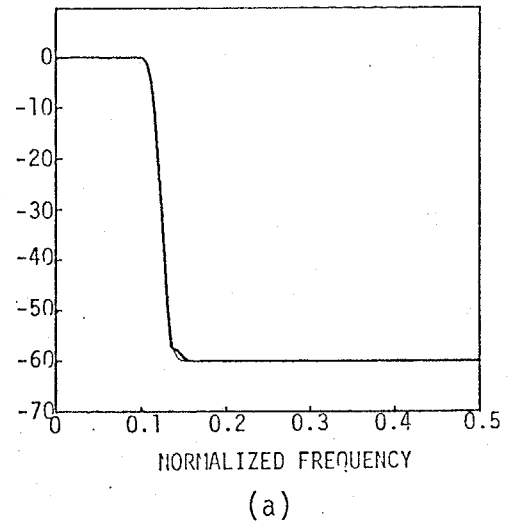
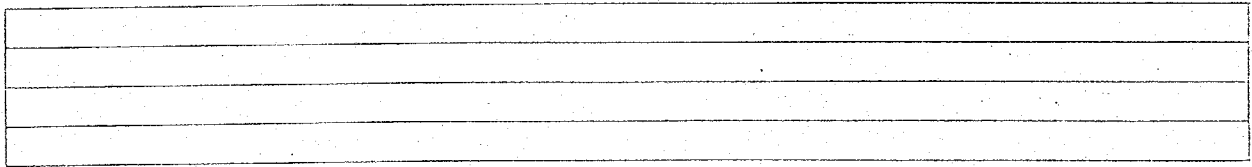


Fig.4.6. Characteristics of the obtained digital filter by this method with $K=L=10$:
 (a) log magnitude response,
 (b) impulse response.

希望対数振幅特性に不連続点がある場合、最小位相インパルス応答の収束が遅くなり、良好な近似特性が得られないことがある。このようなときは、本近似例のように、ケプストラムに窓をかけて、適当な次数で打ち切るとよい。またこのような場合、しばしば、式(4.8)の係数行列 C の行列式の値が殆ど0に近い値となることから、計算精度の取り扱いに注意を要する。なお、音声のスペクトル包絡のように、滑らかな特性では、このような問題は考慮する必要はない。

(3) 一般化対数特性の近似

図4.7は、一般化対数目盛で得られた $|a|$ のスペクトル包絡である。 α の値は、0.2, 0 および -0.2の場合で、いずれもケプストラム次数 $M=20$ に選び、対数振幅目盛上に特性を描いてある。この特性から、最小位相インパルス応答を求め、フィルタ次数 $L=K=18$ で近似したとき得られる特性を図4.8に示す。



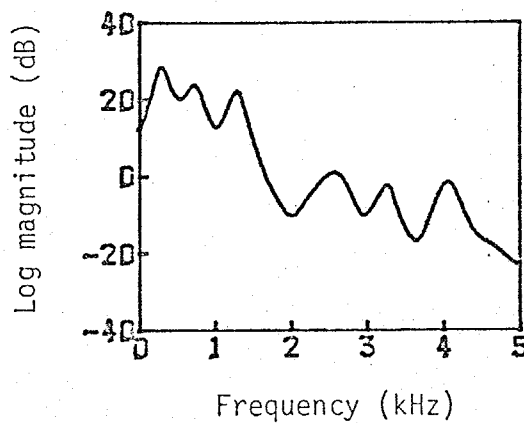
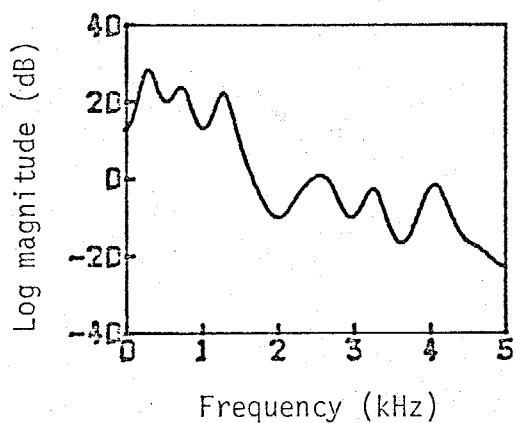
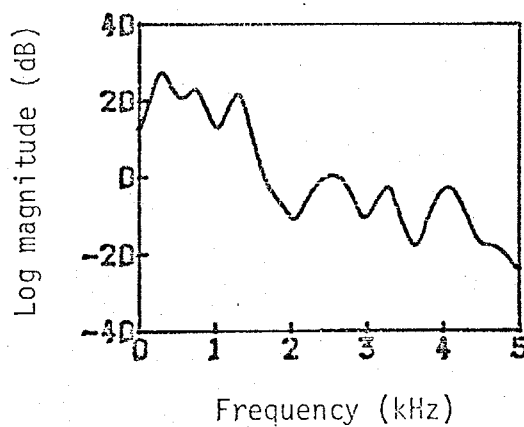
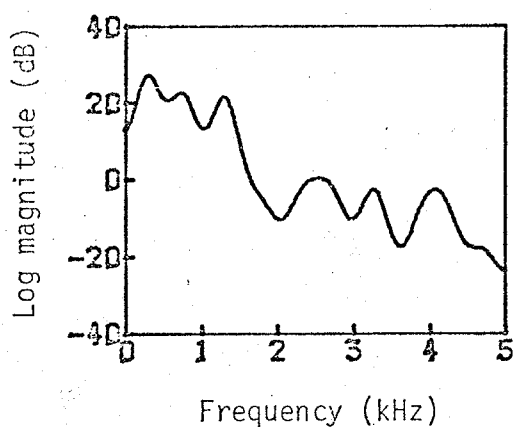
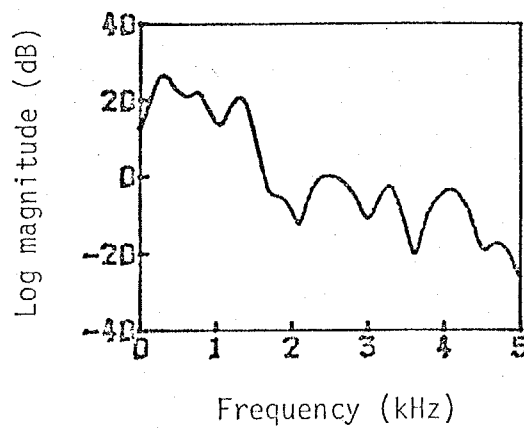
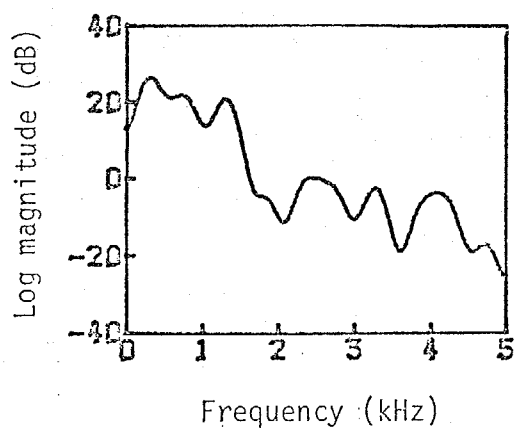


Fig.4.7. Spectral envelopes of Japanese vowel /a/. ($M=20$)

Fig.4.8. Log magnitude responses of the filters approximating the spectral envelopes depicted in Fig.4.7. ($K=L=18$)

4.5 むすび

最小位相インパルス応答を利用する、対数あるいは一般化対数特性の近似法を示した。フィルタ係数を求めるための計算量は、効率的なアルゴリズムを利用できるため、それ程大きくなるまい。また、得られたデジタルフィルタの安定性も保証される。本近似法の問題点は、時間領域の誤差評価式を用いるため、対数目盛あるいは一般化対数目盛上での誤差の規定ができないことである。しかし、対数特性の直接近似の方法や Johnson 法と比べて、遅延素子数は大幅に減少できる。さらに近似例で示したように、特性によれば、ケプストラム次数より低いフィルタ次数でも、良好な近似特性が得られる可能性があることがわかった。

対数目盛の場合について、本近似法により設計したデジタルフィルタを合成フィルタに用いて合成系を構成し、簡単な合成音声の品質の比較を行った結果、ケプストラムの次数に対してフィルタ次数を1割程度減らしても、品質の劣化は殆ど感じられないことが確かめられている。

Blank grid area for additional notes or calculations.

第5章 一般化対数特性近似

デジタルフィルタ

5.1 まえがき

一般化ケプストラムを合成パラメータとして音声合成系を構成するには、一般化ケプストラムにより表されるスペクトル包絡を伝達特性として持つ合成フィルタが必要となる。合成フィルタを設計する際、一般化ケプストラムが一般化対数目盛上のパラメータであることから、フィルタの近似特性は、一般化対数目盛上で評価することが自然であると考えられる。

デジタルフィルタの一般化対数特性の近似法としては、第4章で示した、最小位相系のケプストラムから得られるインパルス応答を用いて、時間領域で設計する方法がある。この近似法は、実用上十分な近似特性が得られるようであるが、一般化対数目盛における近似誤差を規定することができないため、一般的な使用に際して不安が残る。また、フィルタ係数は連立方程式の解で与えられ、その解法に要する計算量はフィルタ次数の2乗に比例するため、次数が大きくなると場合の計算量の増加が問題となる。

本章では、一般化ケプストラムで表された希望一般化対数特性を、一般化対数目盛上で直接近似するデジタルフィルタの構成法を示す。

ここで示す近似法は、一般化対数関数の逆関数のパテ近似を利用するものであり、一般化ケプストラムの値がそのままデジタルフィルタの係数となる特徴がある。特別な場合として、一般化対数関数のパラメータの値が0、すなわち対数関数の場合には、本近似法により得られる構成のフィルタ（一般化対数特性近似デジタルフィルタ）は、対数振幅近似（LMA）フィルタ⁽²³⁾と一致する。また、希望一般化対数特性が決まれば、フィルタの一般化

対数目盛における近似誤差の最大値を規定できる。

以下本章では、まず基本となる近似法を示し、次に具体的なデジタルフィルタの構成法を述べる。続いて、一般化対数目盛上での誤差について論じ、フィルタの安定性、近似誤差の改良のためのパテ近似の修正等について示す。また最後には具体的な近似例も示している。

5.2 一般化対数特性の直接近似

希望一般化対数特性 $\check{G}(\omega) = S_{\delta}\{G(\omega)\}$ が与えられたとき、デジタルフィルタの伝達関数を $H(z)$ とし、その一般化対数周波数特性 $\check{H}(e^{j\omega})$ が次式の関係

$$\check{H}(e^{j\omega}) = S_{\delta}\{H(e^{j\omega})\} = \check{G}(\omega) \quad (5.1)$$

を満足するように、 $H(z)$ を構成することを考える。

まず、一般化対数関数の逆関数 $S_{\delta}^{-1}(w)$ のテイラー展開である式(2.9)を用いて $S_{\delta}^{-1}(w)$ の (N, N) 次パデ近似式 $P_N^{(\delta)}(w)$ を求めると次式が得られる。^{*}

$$S_{\delta}^{-1}(w) \approx P_N^{(\delta)}(w) = \frac{1 + \sum_{k=1}^N B_{N,k}^{(\delta)} w^k}{1 + \sum_{k=1}^N A_{N,k}^{(\delta)} w^k} \quad (5.2)$$

但し

$$A_{N,k}^{(\delta)} = \frac{(-1)^k}{k!} \frac{N C_k}{2N C_k} (1-N\delta)(1-(N-1)\delta) \cdots (1-(N-k+1)\delta),$$

$$k=1, 2, \dots, N \quad (5.3)$$

$$B_{N,k}^{(\delta)} = \frac{1}{k!} \frac{N C_k}{2N C_k} (1+N\delta)(1+(N-1)\delta) \cdots (1+(N-k+1)\delta),$$

$$k=1, 2, \dots, N \quad (5.4)$$

であり、 $1/\delta$ の値が整数のときは、 $N \leq \lfloor 1/\delta \rfloor$ とする。

ところで、あるデジタルフィルタの伝達関数を $F(z)$ とし、 $S_{\delta}^{-1}(w)$ のパデ近似式 $P_N^{(\delta)}(w)$ の w を $F(z)$ で置き換えることにより得られるデジタルフィルタ

*付録A-2参照

$$H(z) = P_N^{(n)}(F(z)) \tag{5.5}$$

の一般化対数周波数特性は式(5.2), (5.5)から,

$$\begin{aligned} \check{H}(e^{j\omega}) &= s_0 \{ P_N^{(n)}(F(e^{j\omega})) \} \\ &\approx s_0 \{ s_0^{-1}(F(e^{j\omega})) \} = F(e^{j\omega}) \end{aligned} \tag{5.6}$$

となる。式(5.6)は、式(5.5)の伝達関数をもつデジタルフィルタの一般化対数周波数特性が、近似的に $F(z)$ の周波数特性と等しくなることを意味している。従って、 $F(z)$ の周波数特性を

$$F(e^{j\omega}) = \check{G}(\omega) \tag{5.7}$$

と選んでおけば、式(5.1)の関係を近似的に満足するフィルタを構成できることがわかる。

このようなことから、ここでは、伝達関数が式(5.5)で表されるフィルタを

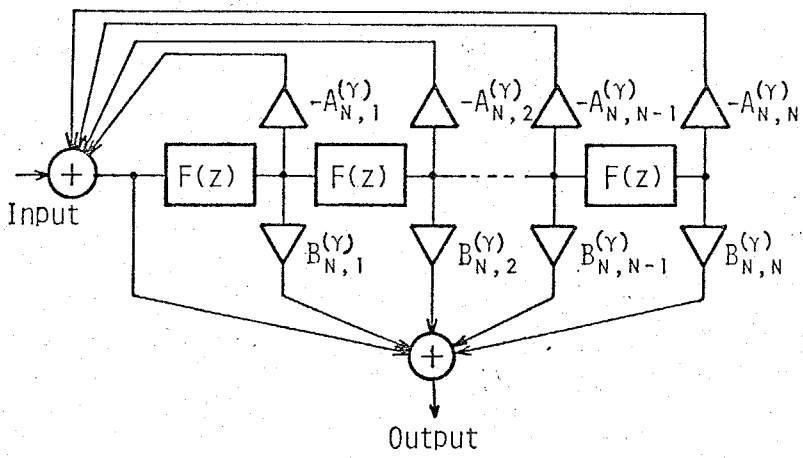


Fig. 5.1 Generalized log characteristic approximation digital filter.

一般化対数特性近似デジタルフィルタと呼び、 $F(z)$ を伝達関数とするフィルタを基礎フィルタと呼ぶことにする。図 5.1 に一般化対数特性近似デジタルフィルタの一般的な構成を示す。

特に $\gamma = 0$ の場合は、式 (5.3), (5.4) で表されるパデ近似式の係数が指数関数 $\exp w$ のパデ近似式の係数と一致することから、図 5.1 の構成のデジタルフィルタは、対数振幅近似 (LMA) フィルタ⁽²³⁾ と一致する。

希望一般化対数特性の一例として、正規化一般化ケプストラムで表されたスペクトル包絡

$$\check{G}(w) = \sum_{m=1}^M \check{v}_m e^{-j\omega m} \quad (5.8)$$

を考えれば、基礎フィルタの伝達関数 $F(z)$ は、

$$F(z) = \sum_{m=1}^M \check{v}_m z^{-m} \quad (5.9)$$

とすればよい。これは、通常の FIR デジタルフィルタで容易に構成することができ、

基礎フィルタの伝達関数は、FIR 形に限らず一般に任意でよいが、

(1) $F(z)$ 自身が安定である

(2) 入力と出力の間に遅延を含まない経路が存在しない

ことが要求される。

5.3 縦続接続による特性の近似

5.3.1 2つの一般化対数特性近似ディジタルフィルタの縦続接続

希望一般化対数特性に定数項が含まれる場合には、結果として基礎フィルタに遅延を含まない経路が存在することになり、そのままでは一般化対数特性近似ディジタルフィルタを構成できない。また、式(5.6)が成立するためにはパデ近似の誤差が十分小さいことが前提となっているが、式(5.2)で表されるパデ近似の誤差は、 $|W|$ の値がある一定の範囲を越えると急激に大きくなるため、取り扱うことができる一般化対数特性の範囲はかなり制限される。

このような場合は、与えられた希望一般化対数特性を適当に分割してそれぞれの特性を近似し、それらを縦続接続することにより、元の特性を近似するフィルタの実現が可能になることが多い。

与えられた希望一般化対数特性を $\check{G}(\omega)$ とし、これが一般化対数目盛上で2つの特性の和

$$\check{G}(\omega) = \check{G}_1(\omega) + \check{G}_2(\omega) \tag{5.10}$$

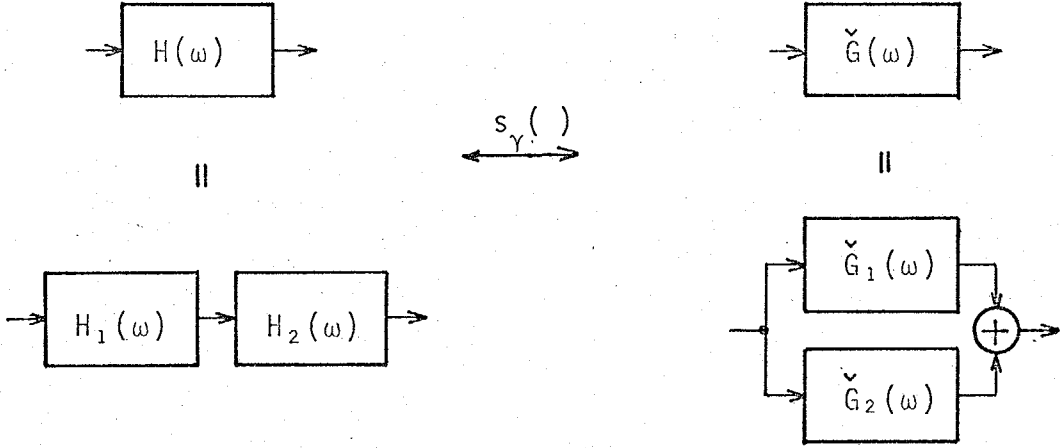


Fig. 5.2. Illustration of cascade filter approximating a generalized log frequency response.

で表せるものとする。このとき、基礎フィルタがそれぞれ $F_1(z)$ および $F_2(z)$ で与えられる2つの一般化対数特性近似デジタルフィルタ $H_1(z)$ および $H_2(z)$ の連続接続

$$H(z) = H_1(z) \cdot H_2(z) \quad (5.11)$$

により、元の特性 $\check{G}(\omega)$ を近似したとき、その一般化対数周波数特性 $\check{H}(e^{j\omega})$ が次式

$$\check{H}(e^{j\omega}) = \check{G}(\omega) \quad (5.12)$$

を満足するための条件を求めてみる。式(2.22)より

$$\begin{aligned} \check{H}(z) &= S_\delta \{ H_1(z) \cdot H_2(z) \} \\ &= \check{H}_1(z) + \check{H}_2(z) + \delta \check{H}_1(z) \check{H}_2(z) \end{aligned} \quad (5.13)$$

となることから、基礎フィルタ $F_1(z)$ の周波数特性

$$F_1(e^{j\omega}) = \check{H}_1(e^{j\omega}) = \check{G}_1(\omega) \quad (5.14)$$

と選ぶと、式(5.12)が成立するためには

$$F_2(e^{j\omega}) = \check{H}_2(e^{j\omega}) = \frac{\check{G}_2(\omega)}{1 + \delta \check{G}_1(\omega)} \quad (5.15)$$

であればよい。従って、基礎フィルタの周波数特性として、それぞれ式(5.14)および式(5.15)に選んだ一般化対数特性近似デジタルフィルタを連続接続すれば、与えられた一般化対数特性 $\check{G}(\omega)$ の近似ができることになる。

特に $\delta = 0$ の場合には、式(5.15)は、 $F_2(e^{j\omega}) = \check{G}_2(\omega) = \hat{G}_2(\omega)$ となり、これは対数目盛上において和の形で表された特性は、直線目盛上ではそれぞれの特性を連続接続すればよいという事実と一致する。

5.3.2 縦続接続による構成例

ここでは、希望一般化対数特性を一般化ケプストラムで表されたスペクトル包絡

$$\check{S}(\omega) = \sum_{m=0}^M \check{C}_m e^{-j\omega m} \quad (5.16)$$

に限定して、いくつかの縦続接続による構成の具体例を示す。

(1) \check{C}_0 の分離

式(5.16)で表される周波数特性をもつデジタルフィルタを構成することは容易であり、

$$F(z) = \sum_{m=0}^M \check{C}_m z^{-m} \quad (5.17)$$

なるFIRデジタルフィルタの伝達関数を考えれば、 $F(e^{j\omega}) = \check{S}(\omega)$ となることは明らかである。しかし、 $F(z)$ は遅延を含まない項 \check{C}_0 を含むため、このままでは基礎フィルタとして用いることはできない。そこで、 $\check{S}(\omega)$ を次のように、 \check{C}_0 の項と遅延を含む項の2つに分割する。

$$\begin{aligned} \check{G}(\omega) = \check{S}(\omega) &= \check{C}_0 + \sum_{m=1}^M \check{C}_m e^{-j\omega m} \\ &= \check{G}_1(\omega) + \check{G}_2(\omega) \end{aligned} \quad (5.18)$$

但し

$$\check{G}_1(\omega) = \check{C}_0 \quad (5.19)$$

$$\check{G}_2(\omega) = \sum_{m=1}^M \check{C}_m e^{-j\omega m} \quad (5.20)$$

式(5.14)、(5.15)から、基礎フィルタの周波数特性として

$$F_1(e^{j\omega}) = \check{C}_0 \quad (5.21)$$

$$F_2(e^{j\omega}) = \sum_{m=1}^M \frac{\check{C}_m}{1 + \delta \check{C}_0} e^{-j\omega m}$$

$$= \sum_{m=1}^M \check{v}_m e^{-j\omega m} \quad (5.22)$$

が得られる。このうち、 $F_1(e^{j\omega})$ は定数であるから、定数倍

$$H_1(z) = s_z^{-1}(\check{C}_0) = \exp(\hat{C}_0) = c_0 \quad (5.23)$$

で特性の実現ができる。一方、 $F_2(z)$ としては、伝達関数を

$$F_2(z) = \sum_{m=1}^M \check{v}_m z^{-m} \quad (5.24)$$

と選び、一般化対数特性近似ディジタルフィルタ

$$H_2(z) = P_N^{(\omega)}(F_2(z)) \quad (5.25)$$

を構成すればよい。式(5.23)で表される $H_1(z)$ と式(5.25)で表される

$H_2(z)$ を連続接続すれば、一般化対数周波数特性 $\check{S}(\omega)$ を実現できる。

式(5.22)の \check{v}_m は、正規化一般化ケプストラムに等しい。従って、実際

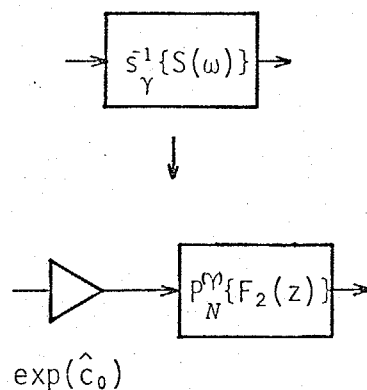


Fig. 5.3. Cascade filter approximating a generalized log spectral envelope of speech.

に一般化対数スペクトル包絡を近似するフィルタを構成する場合、 \hat{c}_0 の値と正規化一般化ケプストラム \check{v}_m がわかっているだけでよいことになる。

(2) \check{v}_2 の分離

正規化一般化ケプストラムにより表された、定数項を含まないスペクトル包絡である式 (5.22) の特性をそのまま使って基礎フィルタを構成すると、パデ近似の誤差を小さくするために、近似の次数を大きくとらなければならない場合が多い。これは、式 (5.22) で $|F_2(e^{j\omega})|$ の値が大きくなり、低次数のパデ近似式では十分な近似が得られないためである。ところで、第2章で示した正規化一般化ケプストラムの分布からわかるように、 $|\check{v}_m|$ の値が大きくなるのは殆どが低次の項である。従って、低次数の一般化ケプストラムの項を分離するだけでも、近似誤差を小さくするために必要なパデ近似の次数を、低くすることができる。

ここでは、希望一般化対数特性を

$$\check{G}(\omega) = \sum_{m=1}^M \check{v}_m e^{-j\omega m} \quad (5.26)$$

とし、 \check{v}_2 の項を分離して近似することを考える。 $\check{G}(\omega)$ を2つの特性

$$\begin{aligned} \check{G}(\omega) &= \check{v}_2 e^{-j\omega} + \sum_{m=2}^M \check{v}_m e^{-j\omega m} \\ &= \check{G}_1(\omega) + \check{G}_2(\omega) \end{aligned} \quad (5.27)$$

但し

$$\check{G}_1(\omega) = \check{v}_2 e^{-j\omega} \quad (5.28)$$

$$\check{G}_2(\omega) = \sum_{m=2}^M \check{v}_m e^{-j\omega m} \quad (5.29)$$

に分割すれば、基礎フィルタの周波数特性は、式 (5.14)、(5.15) から

$$F_1(e^{j\omega}) = \check{v}_2 e^{-j\omega} \quad (5.30)$$

$$F_2(e^{j\omega}) = \sum_{m=2}^M \check{v}_m e^{-j\omega m} / (1 + \sigma \check{v}_1 e^{-j\omega}) \quad (5.31)$$

が得られる。従って、基礎フィルタの伝達関数は

$$F_1(z) = \check{v}_1 z^{-1} \quad (5.32)$$

$$F_2(z) = \sum_{m=2}^M \check{v}_m z^{-m} / (1 + \sigma \check{v}_1 z^{-1}) \quad (5.33)$$

とすればよい。基礎フィルタ $F_2(z)$ は IIR 形構成となるため、それ自身の安定性が問題となる。 $F_2(z)$ が安定であるための条件は

$$|\sigma \check{v}_1| = |\sigma \hat{c}_1| < 1 \quad (5.34)$$

で表される。

(3) \check{v}_1, \check{v}_2 の分離

ここでは、式(5.26)の希望一般化対数特性に対して、式(5.27)のように分割するかわりに、次式

$$\begin{aligned} \check{G}(\omega) &= \left\{ \check{v}_1 e^{-j\omega} + \check{v}_2 e^{-j2\omega} \right\} + \sum_{m=3}^M \check{v}_m e^{-j\omega m} \\ &= \check{G}_1(\omega) + \check{G}_2(\omega) \end{aligned} \quad (5.35)$$

但し

$$\check{G}_1(\omega) = \check{v}_1 e^{-j\omega} + \check{v}_2 e^{-j2\omega} \quad (5.36)$$

$$\check{G}_2(\omega) = \sum_{m=3}^M \check{v}_m e^{-j\omega m} \quad (5.37)$$

とすることを考える。

(2)の場合と同様にして、基礎フィルタの伝達関数を求めると次式が得られる。

$$F_1(z) = \check{v}_1 z^{-1} + \check{v}_2 z^{-2} \quad (5.38)$$

$$F_2(z) = \sum_{m=3}^M \check{v}_m z^{-m} / (1 + \gamma \check{v}_1 z^{-1} + \gamma \check{v}_2 z^{-2}) \quad (5.39)$$

このときの基礎フィルタ $F_2(z)$ が安定となるための条件は、 $\check{v}_1 (= \hat{c}_1)$ と \check{v}_2 が図 5.4 に示す三角形の内部の領域の値をとることである。

上記 (1) ~ (3) 以外にもいろいろ特性の分割方法を考えることができるが、基礎フィルタの構成、安定性の条件等は簡単でなくなる。

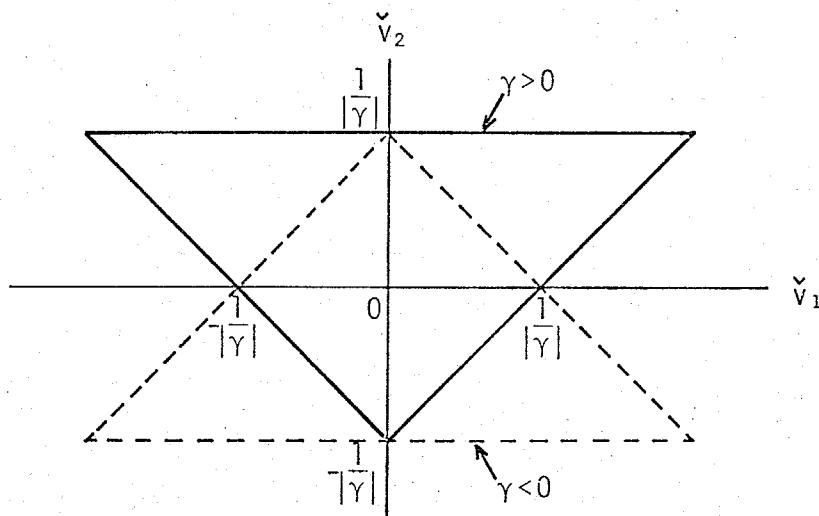


Fig. 5.4. Region of stability for the basic filter $F_2(z)$.

5.4 一般化対数目盛上での誤差

対数目盛で近似特性を評価する場合には、実数部は振幅特性、虚数部は位相特性というように、はっきりとした意味があった。一般化対数目盛上では対数目盛の場合と違って、このまうなはっきりとした意味づけはできない。しかし近似誤差が十分小さな範囲では、対数目盛上での誤差と一般化対数目盛上での誤差との間に、簡単な関係を導くことができる。

一般化対数目盛上での誤差 $\check{E}_r(\omega)$ を

$$\check{E}_r(\omega) = \check{H}(e^{j\omega}) - \check{G}(\omega) \quad (5.40)$$

で定義し、対数目盛上での誤差を

$$\hat{E}_r(\omega) = \hat{H}(e^{j\omega}) - \hat{G}(\omega) \quad (5.41)$$

で表すことにする。ここで、 $\check{H}(e^{j\omega})$ は $G(\omega)$ の近似特性とし、 $\check{\cdot} = S_\sigma(\cdot)$ 、 $\hat{\cdot} = Q_m \cdot$ とする。

一般に、 $A(z)$ 、 $B(z)$ を z の関数として

$$\begin{aligned} S_\sigma \left\{ \frac{B(z)}{A(z)} \right\} &= \frac{1}{\sigma} \left\{ \frac{B^\sigma(z)}{A^\sigma(z)} - 1 \right\} = \frac{1}{\sigma} \left\{ \frac{1 + \sigma \check{B}(z)}{1 + \sigma \check{A}(z)} - 1 \right\} \\ &= \frac{\check{B}(z) - \check{A}(z)}{1 + \sigma \check{A}(z)} \end{aligned} \quad (5.42)$$

が成立する。このことから、直線目盛上での相対誤差

$$E_r(\omega) = \frac{H(e^{j\omega}) - G(\omega)}{G(\omega)} = \frac{H(e^{j\omega})}{G(\omega)} - 1 \quad (5.43)$$

を用いて、式(5.40)は次のように表すことができる。

$$\begin{aligned} \check{E}_r(\omega) &= \{1 + \sigma \check{G}(\omega)\} \cdot S_\sigma \left\{ \frac{H(e^{j\omega})}{G(\omega)} \right\} \\ &= \{1 + \sigma \check{G}(\omega)\} \cdot S_\sigma \{1 + E_r(\omega)\} \end{aligned} \quad (5.44)$$

一方式(5.41)は

$$\begin{aligned}\hat{E}(\omega) &= \ln H(e^{j\omega}) - \ln G(\omega) = \ln \frac{H(e^{j\omega})}{G(\omega)} \\ &= \ln \{1 + E_r(\omega)\}\end{aligned}\quad (5.45)$$

と表せる。ところで, $|E_r(\omega)| \ll 1$ であれば, 式(2.7)により

$$S_r\{1 + E_r(\omega)\} \approx E_r(\omega) \approx \ln \{1 + E_r(\omega)\} \quad (5.46)$$

が成立する。従って, 式(5.44), (5.45)から, $|E_r(\omega)| \ll 1$ の場合には, $\check{E}_r(\omega)$ と $\hat{E}(\omega)$ の間に

$$\check{E}_r(\omega) \approx \{1 + \delta \check{G}(\omega)\} E_r(\omega) \approx \{1 + \delta \check{G}(\omega)\} \hat{E}(\omega),$$

$$|E_r(\omega)| \ll 1 \quad (5.47)$$

という関係が得られる。

特に, 希望一般化対数特性 $\check{G}(\omega)$ が

$$\check{G}(\omega) = r e^{j\omega}, \quad r > 0 \quad (5.48)$$

ならば, 式(5.47)は

$$\hat{E}(\omega) \approx \frac{\check{E}_r(\omega)}{1 + \delta r e^{j\omega}}, \quad |E_r(\omega)| \ll 1 \quad (5.49)$$

となり, 最大誤差に関して近似的に

$$\max_{0 \leq \omega \leq 2\pi} |\hat{E}(\omega)| \leq \max_{0 \leq \omega \leq 2\pi} |\check{E}_r(\omega)| / (1 - \delta |r|) \quad (5.50)$$

で評価することができる。

また, 一般化対数関数 $S_\delta(w)$ の定義および式(5.2) ~ (5.4) から

$$S_\delta(w) = -\frac{1}{\delta} (w^{-\delta} - 1) \quad (5.51)$$

$$P_N^{(\delta)}(-\omega) = \frac{1}{P_N^{(\delta)}(\omega)} \quad (5.52)$$

と表せることを使うと, $\check{G}(\omega) = re^{j\omega}$, $H(e^{j\omega}) = P_N^{(\delta)}(\check{G}(\omega))$ の場合には,

$$\check{E}_{-\delta}(\omega + \pi) = -re^{j\omega} - s_{-\delta} \{ P_N^{(\delta)}(-re^{j\omega}) \}$$

$$= -re^{j\omega} - s_{-\delta} \{ 1 / P_N^{(\delta)}(re^{j\omega}) \}$$

$$= s_{\delta} \{ P_N^{(\delta)}(re^{j\omega}) \} - re^{j\omega}$$

となり, 結局

$$\check{E}_{\delta}(\omega) = -\check{E}_{-\delta}(\omega + \pi) \quad (5.53)$$

が得られる。従って

$$\max_{0 \leq \omega \leq 2\pi} |\check{E}_{\delta}(\omega)| = \max_{0 \leq \omega \leq 2\pi} |\check{E}_{-\delta}(\omega)| \quad (5.54)$$

が成立し, 式(5.50)は

$$\max_{0 \leq \omega \leq 2\pi} |\hat{E}(\omega)| \leq \max_{0 \leq \omega \leq 2\pi} |\check{E}_{101}(\omega)| / (1 - |\delta|r) \quad (5.55)$$

と書くことができる。

5.5 安定性とパデ近似の修正

5.5.1 安定性

まず、基礎フィルタの伝達関数を

$$F(z) = r z^{-m}, \quad r > 0 \quad (5.56)$$

としたときの、一般化対数特性近似ディジタルフィルタ $P_N^{(a)}(F(z))$ が安定であるための条件を考える。

$S_0^{-1}(W)$ の (N, N) 次パデ近似 $P_N^{(a)}(W)$ の極および零点をそれぞれ $\{w_k^p\}$, $\{w_k^z\}$, $k=1, 2, \dots, N$ とすれば、一般化対数特性近似ディジタルフィルタ $H(z) = P_N^{(a)}(r z^{-m})$ の極 $\{z_k^p\}$ および零点 $\{z_k^z\}$ は容易に求まり、

$$z_k^p = \left(\frac{w_k^p}{r} \right)^{-\frac{1}{m}} \quad (5.57)$$

$$z_k^z = \left(\frac{w_k^z}{r} \right)^{-\frac{1}{m}} \quad (5.58)$$

となる。従って、 $H(z) = P_N^{(a)}(r z^{-m})$ が安定であるためには、 $|z_k^p| < 1$ かつ

$$r < R_s = \min_k \{ |w_k^p| \} \quad (5.59)$$

であればよい。また、最小位相であるためには、 $|z_0^p| < 1$, $|z_0^z| < 1$ かつ

$$r < R_M = \min_k \{ |w_k^p|, |w_k^z| \} \quad (5.60)$$

を満足すれば良い。式(5.59), (5.60) から明らかに $R_M \leq R_s$ となる。

表5.1に、パデ近似の次数 $N=3, 4$ の場合について、いくつかの r の値に対する R_s および R_M の値を示す。

一般にパデ近似の誤差は、パデ近似式の極および零点で極大値をとり、

Table 5.1. Region of minimum phase and stability for $P_N^{(p)}(r\bar{z}^m)$,
 R_M and R_S .

γ	N=3		N=4	
	R_M	R_S	R_M	R_S
0.2	3.195	8.401	3.889	13.401
0.1	3.775	6.066	4.703	8.571
0	4.644	4.644	6.046	6.046
-0.1	3.775	3.775	4.703	4.703
-0.2	3.195	3.195	3.889	3.889

$r \geq R_M$ の領域では誤差が大きくて使用にならないことから、実際上は式 (5.60) の条件、すなわち最小位相となる条件が重要となる。

次に、基礎フィルタの伝達関数 $F(z)$ が一般的な場合を考える。 $F(z)$ はそれ自身が安定であることを仮定していることから、 z 平面の $|z| \geq 1$ の領域で解析的である。一方、最大値の原理⁽²⁴⁾ から、 $|z| \geq 1$ の領域で $|F(z)|$ の最大値は境界すなわち単位円上でとり、その値は

$$W = \max_{0 \leq \omega \leq 2\pi} |F(e^{i\omega})| \quad (5.61)$$

となる。このことから

$$W < R_M \quad (5.62)$$

であれば、 $H(z) = P_N^{(p)}(F(z))$ の極および零点は z 平面の単位円内に存在することが容易に確かめられる。従って、式 (5.62) の条件があれば、一般化対数特性近似デジタルフィルタは安定かつ最小位相となる。

実際上は、パデ近似の誤差が十分小さな範囲、例えば $\sqrt{\varepsilon_r(\omega)} \leq 0.02$ とな

るような範囲であれば安定と考えてよい。

5.5.2 パラ近似の修正

一般にパラ近似 $P_N^{(\sigma)}(w)$ は、 $|w|$ の値がある範囲を越えると急激に近似誤差が大きくなる。図5.5は、いくつかの σ の値について、基礎フィルタの伝達関数を $F(z) = rz^{-1}$ としたときの一般化対数目盛上での誤差

$$\check{E}_r(w) = \sigma_r \{ P_N^{(\sigma)}(re^{-jw}) \} - re^{-jw} \quad (5.63)$$

の最大値 $\max_{\omega} |\check{E}_r(w)|$ を r に対して描いたものである。なお、式(5.54)の関係から、 $\max_{\omega} |\check{E}_r(w)|$ の値は r が同じであれば、 σ の符号には関係なく絶対値だけで決まる。

図5.5から、誤差が実用上十分小さくなる r の範囲は、 R_M の値に比べてかなり小さくなることがわかる。

そこで、パラ近似式の係数 $A_{N,k}^{(\sigma)}$ 、 $B_{N,k}^{(\sigma)}$ に対して、実数 $\delta_{N,k}^{(\sigma)}$ 、 $\varepsilon_{N,k}^{(\sigma)}$ を用いて次式で表される修正

$$\overline{A}_{N,k}^{(\sigma)} = (1 - \delta_{N,k}^{(\sigma)}) \cdot A_{N,k}^{(\sigma)} \quad (5.64)$$

$$\overline{B}_{N,k}^{(\sigma)} = (1 - \varepsilon_{N,k}^{(\sigma)}) \cdot B_{N,k}^{(\sigma)} \quad (5.65)$$

を用いて、 $\overline{A}_{N,k}^{(\sigma)}$ 、 $\overline{B}_{N,k}^{(\sigma)}$ を

$$\left\{ \overline{A}_{N,k}^{(\sigma)}, \overline{B}_{N,k}^{(\sigma)} \mid \min_{\substack{\delta_{N,k}^{(\sigma)} \\ \varepsilon_{N,k}^{(\sigma)}}} \max_{\substack{\omega \\ 0 \leq \omega \leq \pi \\ r=R}} |\check{E}_r(w)| \right\} \quad (5.67)$$

と選ぶことにより、 $r < R$ の範囲で近似誤差を小さくすることを考える。

式(5.67)は非線形最適化問題であり簡単ではないが、一例として、パラ近似の次数を $N=3$ 、最適化の範囲 $R=3.0$ とし、Davidon-Fletcher-Powell法⁽²⁵⁾により最適化を行った結果を示す。

表5.2は、 $\delta_{N,k}^{(\sigma)}$ 、 $\varepsilon_{N,k}^{(\sigma)}$ をいくつかの σ の値に対して求めた結果であり、

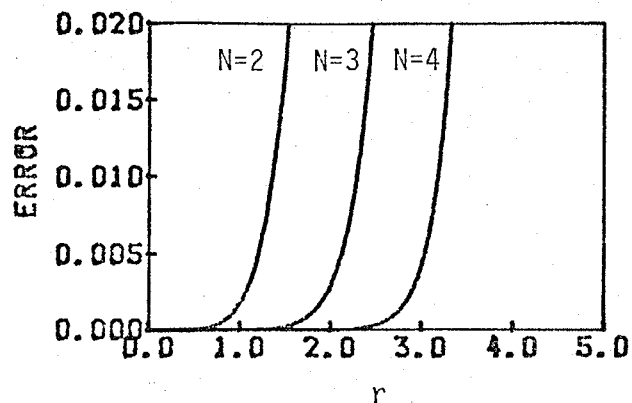
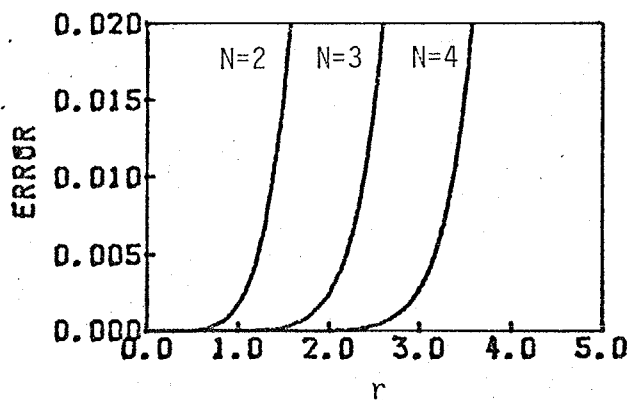
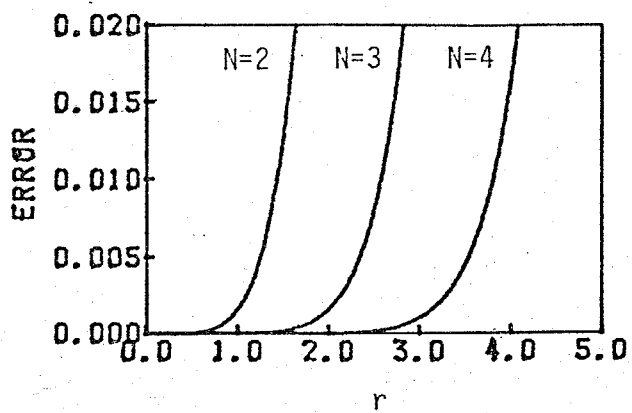
(a) $|\gamma|=0.2$ (b) $|\gamma|=0.1$ (c) $\gamma=0$ Fig.5.5. Approximation errors $\max_{\omega} |\check{E}_{\gamma}(\omega)|$ versus r .

表5.3は、このときの誤差の最大値 $\max_{r \leq 3} |\check{E}_\gamma(\omega)|$, $P_N^{(r)}(r \neq 1)$ が最小位相および安定となる範囲, R_M と R_S である。式(5.67)の最適化をしない場合、パデ近似の次数 $N=3$ では $r \leq 3$ の範囲で最大誤差がかなり大きくなるのに対し、最適化後は、最大誤差が 0.02 程度と、十分実用になる近似

Table 5.2. The coefficients of the best Padé approximants with $N=3$ and $R=3.0$.

γ	$\delta_{3,1}^{(\gamma)}$	$\delta_{3,2}^{(\gamma)}$	$\delta_{3,3}^{(\gamma)}$	$\epsilon_{3,1}^{(\gamma)}$	$\epsilon_{3,2}^{(\gamma)}$	$\epsilon_{3,3}^{(\gamma)}$
0.2	-0.131221	-0.345640	-0.748244	0.032956	0.080234	0.151227
0.1	-0.045646	-0.110642	-0.204502	0.024799	0.062998	0.126543
0	0.000157	0.004738	0.023498	0.000157	0.004738	0.023498
-0.1	0.024799	0.062998	0.126543	-0.045646	-0.110642	-0.204502
-0.2	0.032956	0.080234	0.151227	-0.131221	-0.345640	-0.748244

Table 5.3. $\max |\check{E}_\gamma(\omega)|$, R_M and R_S of the best Padé approximants.

γ	$\max \check{E}_\gamma $	R_M	R_S
0.2	0.01819	3.531	7.043
0.1	0.02034	4.059	5.685
0	0.02095	4.738	4.738
-0.1	0.02034	4.059	4.059
-0.2	0.01819	3.531	3.531

特性が得られている。また、表5.1と表5.3の比較から、最小位相となる範囲も広がることがわかる。

図5.6に、最適化前の ω に対する近似誤差の絶対値 $|\check{E}_r(\omega)|$ を、 $r < 0$ の r の値について計算した結果を、図5.7に、同じ誤差について最適化後の結果を示す。なお、 $r > 0$ の場合は、式(5.53)の関係から、図5.6、図5.7のグラフを ω に対して π だけ移動させた形となるだけであるため、ここでは示していない。この図から、パデ近似式の極または零点の付近で急激に大きくなる誤差が、最適化により平均化され、結果として最大誤差が小さくなる様子が見える。

式(5.63)の誤差の場合、 $r_2 < r_1 < R_M$ であれば、 $\max_{r=r_2} |\check{E}_r(\omega)| < \max_{r=r_1} |\check{E}_r(\omega)|$ となることから、基礎フィルタの伝達関数 $F(z)$ が一般的な形でも、式(5.61)で定義される \mathcal{W} の値が変わっていれば、一般化対数特性近似デジタルフィルタ $H(z) = P_N^{(a)}(F(z))$ の誤差は

$$\max_{0 \leq \omega \leq 2\pi} \left| \mathcal{S}_r \left\{ P_N^{(a)}(F(e^{j\omega})) \right\} - F(e^{j\omega}) \right| < \max_{\substack{0 \leq \omega \leq 2\pi \\ r = \mathcal{W}}} |\check{E}_r(\omega)| \quad (5.68)$$

で規定できる。ここで $\check{E}_r(\omega)$ は、式(5.63)とする。従って、 $N=3$ の場合、表5.2の $\check{E}_{N,r}^{(a)}$ 、 $\check{\sigma}_{N,r}^{(a)}$ で修正した係数を用いれば、 $\mathcal{W} < 3.0$ となるような一般化対数特性に対して、近似誤差の最大値は、表5.3に示した $\max |\check{E}_r|$ の値を越えることはない。

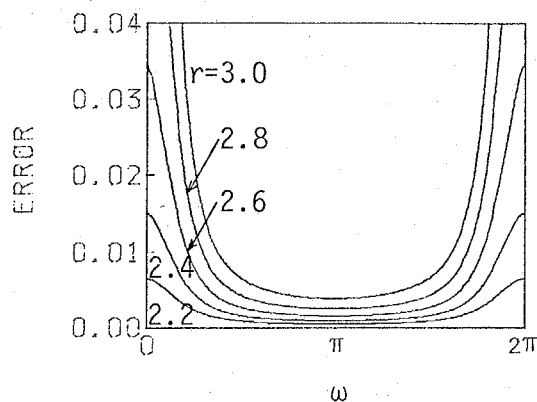
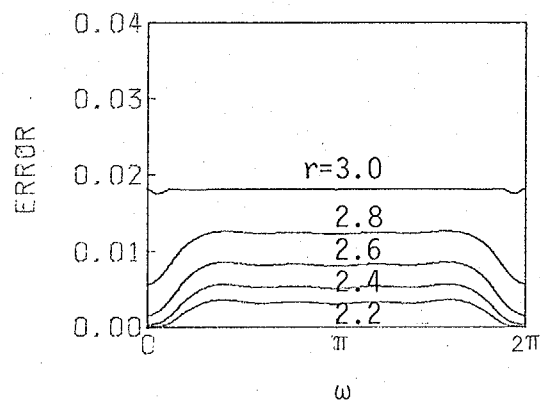
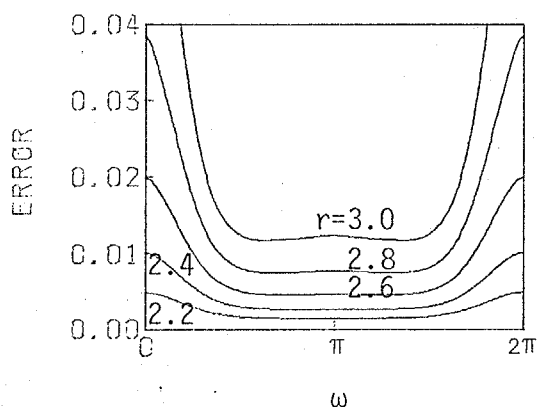
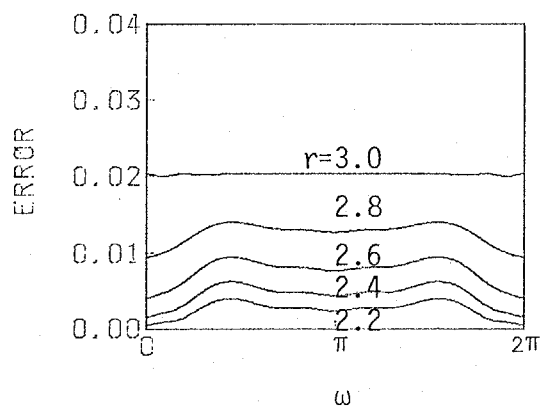
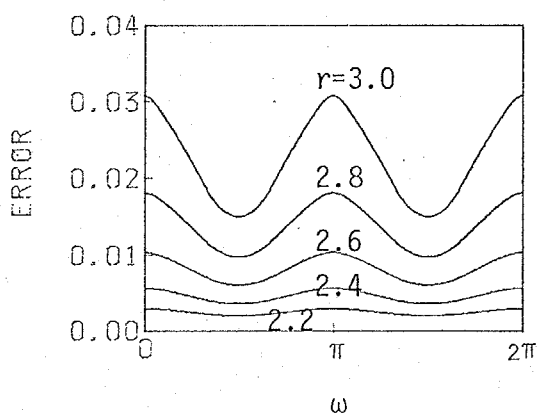
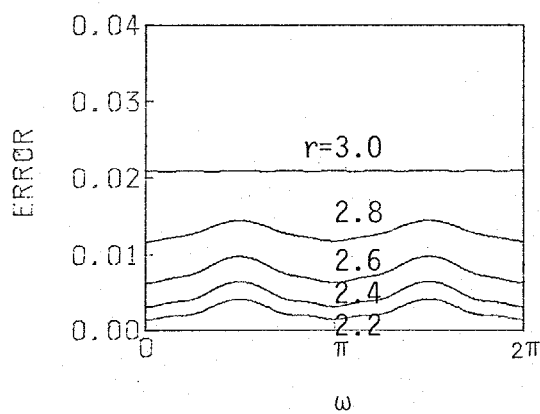
(a) $\gamma = -0.2$ (a) $\gamma = -0.2$ (b) $\gamma = -0.1$ (b) $\gamma = -0.1$ (c) $\gamma = 0$ (c) $\gamma = 0$

Fig.5.6. Approximation errors $|\check{E}_\gamma(\omega)|$ versus ω . ($N=3$)

Fig.5.7. Approximation errors $|\check{E}_\gamma(\omega)|$ versus ω after optimization. ($N=3, R=3.0$)

5.6 近似例

一般化対数特性近似ディジタルフィルタを用いて、一般化ケプストラムにより表された音声のスペクトル包絡を近似した例を図5.8に示す。

図5.8(a)は、図2.6に示した、男声/nambu/の/bu/の初めの部分のスペクトルから、 $\gamma = -0.1$ として改良一般化ケプストラムにより得られたスペクトル包絡である。同図(b)は、縦続接続による構成例(2)で示した \hat{c}_0 および \check{v}_2 を分離した構成で、パデ近似の次数を $N=4$ と選んだ場合の一般化対数特性近似ディジタルフィルタの振幅特性である。なお、ケプストラム次数は $M=20$ であり、図2.6との比較を容易にするため、対数振幅目盛上に描いてある。図5.8(c)は誤差を示したもので、一般化対数目盛上での誤差 $\check{E}_A(\omega)$ の絶対値と、対数振幅目盛上での誤差

$$\hat{E}_A(\omega) = \ln |H(e^{j\omega})| - \ln |S(\omega)| \quad (5.69)$$

を描いてある。ここで $S(\omega) = s_2^{-1} \{ \check{S}(\omega) \}$ でありスペクトル包絡を表す。

誤差特性からわかるように、この場合の近似誤差は殆ど問題とされない程度に小さい。フィルタの遅延素子数は、 $M=20$ 、 $N=4$ としていることから、全体で84となる。また、このフィルタは安定である。

一般化対数関数のパラメータ γ の値が0のときは、一般化対数特性近似ディジタルフィルタはLMAフィルタに一致するが、LMAフィルタの近似例については文献(23)を参照されたい。

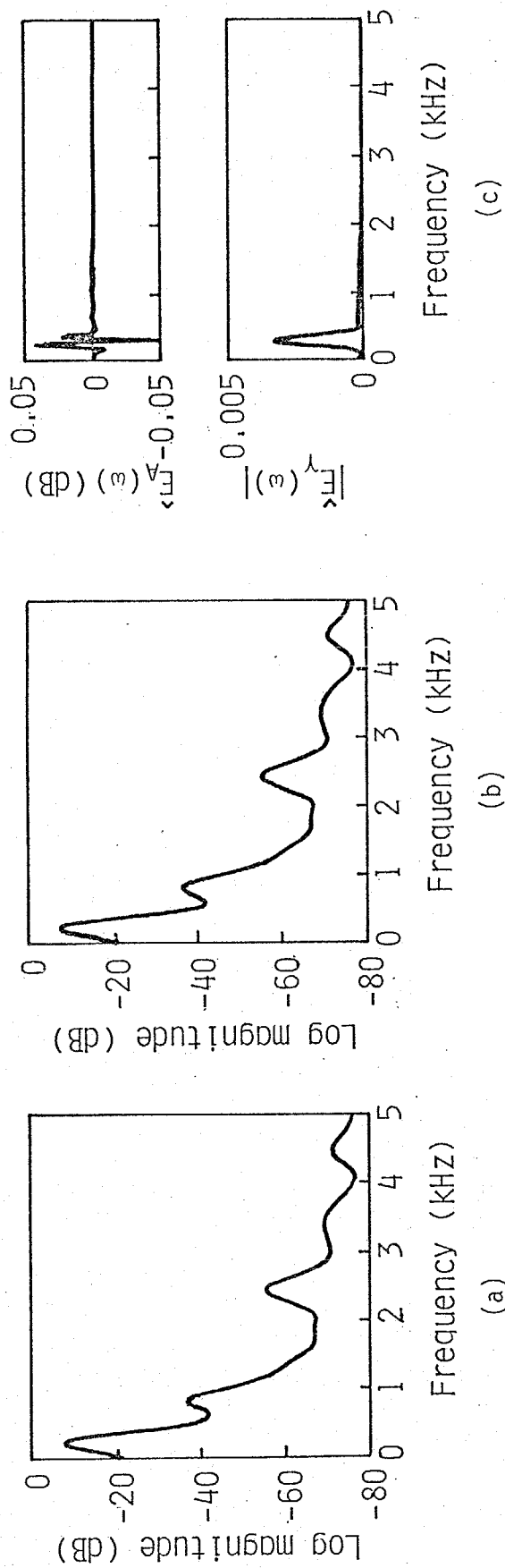


Fig. 5.8. An example of the characteristics of the speech synthesis filter: (a) desired response, spectral envelope for the beginning portion of /bu/ with $M=20$ and $\gamma=-0.1$, (b) log magnitude response of the digital filter with $N=4$, (c) approximation errors.

5.7 むすび

一般化対数目盛上で与えられた希望特性を、直接近似するデジタルフィルタの構成法を示した。本章で示した一般化対数特性近似デジタルフィルタは、希望特性が一般化ケプストラムで表された一般化対数スペクトル包絡の場合には、一般化ケプストラムの値がそのままフィルタ係数になるという特長があった。また、近似例で示したように、パデ近似の次数を適当に選ぶことにより、非常に良い近似特性が得られる。

一般化対数目盛上で与えられた特性は、対数目盛の場合と比べ、特性を分割して近似することがやや複雑になる。しかし、音声のスペクトル包絡を近似する場合には、一般化ケプストラムの低次の項 ($\hat{c}_0, \hat{v}_1 \sim \hat{v}_3$) に大きな値が集中していることから、本章中で示したような簡単な特性の分割による近似で十分であると考えられる。パデ近似の次数が4次上になると、実際に一般化対数特性近似デジタルフィルタを構成した際、パデ近似の係数感度が大きくなるという問題が生じる。この点については、4次のパデ近似であれば、あらかじめパデ近似式を2次-2次に因数分解することにより、係数感度を低くすることができるといえる。

一般化対数特性近似デジタルフィルタは、通常の構成に比べて遅延素子数がかかなり大きくなる欠点はあるが、音声のスペクトル包絡の近似では、フィルタ係数の決定が不要であること、近似誤差をある程度規定できること、さらに対数目盛の場合のLMAフィルタと互換性があることなどの理由から、音声合成フィルタとして利用することにより、良好な結果が期待できる。音声合成系への応用については、次章で述べる。

第6章 一般化ケプストラムによる音声の合成

6.1 まえがき

ケプストラムを用いる音声の合成法としては、Oppenheim らが提案した時間領域での直接たたみ込みの方法⁽⁵⁾がよく知られている。これは、ケプストラムから計算される最小位相インパルス応答と駆動音源波形（周期的インパルス列か白色雑音）を直接たたみ込んで合成音声を得るものである。しかし、一般にこの方法は、本来無限長の最小位相インパルス応答を有限項で打ち切ることの影響で、期待されるほど高品質な合成音声は得られないようである。また、ケプストラムから最小位相インパルス応答という時間領域のパラメータへの変換が必要であるため、計算量も多くなる。

これに対して、対数振幅特性の直接近似フィルタを合成フィルタに用いた場合、ケプストラムで表現されたスペクトル包絡を正確に近似することができ、その結果として高品質な合成音声を得られることが確かめられている。⁽⁷⁾また、合成フィルタの係数がケプストラムにより直接表されるため、係数の決定は簡単である。

本章では、一般化ケプストラムを用いる音声の合成法を示す。

音声合成フィルタとしては、第5章で示した一般化対数特性近似デジタルフィルタを用いる。このフィルタを合成フィルタに用いる理由としては、一般化ケプストラムにより表されたスペクトル包絡を一般化対数目盛上で直接近似できること、一般化ケプストラムの値がそのままフィルタ係数となるため、フィルタ係数決定のための計算が不要であること、任意の一般化対数のパラメータ γ の値に対する構成が可能であり、特に $\gamma=0$ の場合はLMAフィルタ⁽²³⁾と一致することなどの点があげられる。

本章で示す合成系は、対数振幅特性の直接近似フィルタあるいはLMAフィルタを用いる従来のケプストラム法による合成系と、構成が基本的に同じであり、一般化ケプストラムと一般化対数特性近似デジタルフィルタを用いる点異なる。但し、合成系に伝送するパラメータは、通常のケプストラムかあるいは改良ケプストラムおよびビット率であるため、従来の合成系と完全に互換性のある合成系となっている。合成に必要な一般化ケプストラムの値は、伝送されたケプストラムから簡単な変換式により、合成系内部で得ている。従って、分析系は従来のケプストラム法で用いられる系を変更することなくそのまま利用できる。また、本合成系では、一般化ケプストラムによる合成法の一つの特別な場合として、通常のケプストラムによる合成もできる。この場合は、LMAフィルタを合成フィルタに用いた合成系と一致する。

以下本章では、まず合成系の全体的な構成を示した後、合成フィルタの実現に関していくつかの検討を加える。次に、実際に合成音声を作成し、品質評価を行った結果を述べる。またこの結果から、一般化対数関数のパラメータの選択についても示唆している。

6.2 合成系の構成

6.2.1 全体構成

合成フィルタとして一般化対数特性近似ディジタルフィルタを利用する、一般化ケプストラムによる合成系の構成を図6.1に示す。

合成系への分析系からの入力は、通常のケプストラムがあるいは改良ケプストラム、およびピッチである。

ピッチは有声区間ではピッチ周期（単位ms）を、無声区間では零を伝送する。合成系では、伝送されたピッチの値が零かどうかで有声/無声の区別を行い、有声区間ではピッチ周期に対応する周期のインパルス列を、無声区間ではM-系列をそれぞれ合成フィルタの駆動音源として発生する。なお、インパルス列の振幅は、合成フィルタの入力系列のパワーを一定に保つために、ピッチ周期の値を I_v としたとき $\sqrt{I_v}$ で表される。ここで、 ν はフレーム番号である。（以下、特にことわらな限り、 ν はフレーム番号を表すものとする。）

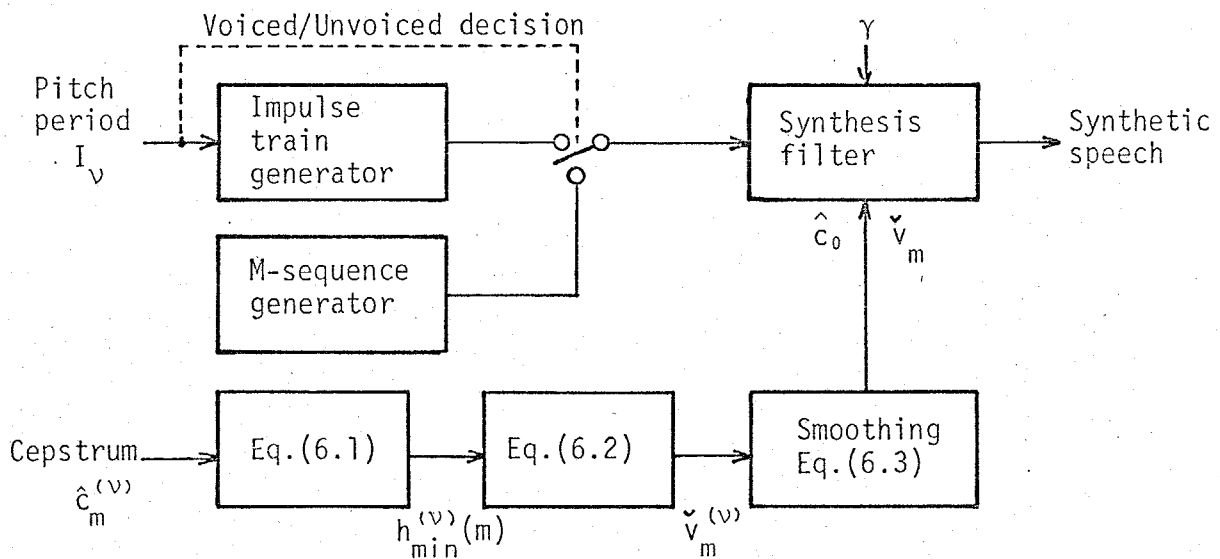


Fig. 6.1. Block diagram of the synthesis system.

一般化ケプストラムの値は、伝送されたケプストラムから、第2章で示したケプストラムと一般化ケプストラムの関係式を利用して求める。第5章で示したように、一般化対数特性近似ディジタルフィルタを構成する場合、実際に必要とするのは \hat{c}_0 及び正規化一般化ケプストラムの値である。伝送されたフレーム番号 ν におけるケプストラムを $\{\hat{c}_m^{(\nu)}\}$ とすれば、正規化一般化ケプストラム $\{\check{v}_m^{(\nu)}\}$ は、式(2.47)、(2.49)の関係により次式で求められる。

$$h_{min}^{(\nu)}(m) = \begin{cases} 1, & m=0 \\ \hat{c}_1^{(\nu)}, & m=1 \\ \sum_{k=1}^m \binom{k}{m} \hat{c}_k^{(\nu)} h_{min}^{(\nu)}(m-k), & 1 < m \leq M \end{cases} \quad (6.1)$$

$$\check{v}_m^{(\nu)} = \begin{cases} h_{min}^{(\nu)}(1) (= \hat{c}_1^{(\nu)}), & m=1 \\ h_{min}^{(\nu)}(m) + \sum_{k=1}^{m-1} \binom{k}{m} \{ \delta h_{min}^{(\nu)}(k) \check{v}_{m-k}^{(\nu)} - \check{v}_k^{(\nu)} h_{min}^{(\nu)}(m-k) \}, & 1 < m \leq M \end{cases} \quad (6.2)$$

ここで、 M は合成時のケプストラム次数とし、伝送するケプストラム $\hat{c}_m^{(\nu)}$ は最小位相系を仮定している。

一般化対数特性近似ディジタルフィルタで用いる一般化ケプストラムの値は、式(6.2)で得られた値をさらにフレーム方向に平滑化した

$$\begin{cases} \check{v}_m = \frac{1}{4} (\check{v}_m^{(\nu-1)} + 2\check{v}_m^{(\nu)} + \check{v}_m^{(\nu+1)}) \\ \hat{c}_0 = \frac{1}{4} (\hat{c}_0^{(\nu-1)} + 2\hat{c}_0^{(\nu)} + \hat{c}_0^{(\nu+1)}) \end{cases}, \quad 1 \leq m \leq M \quad (6.3)$$

とする。

6.2.2 合成フィルタ

合成フィルタは、一般化対数特性近似デジタルフィルタを用い、近似誤差を小さくするために、適当に特性を分割して縦続接続により近似する。図

6.2に合成フィルタの構成を示す。このフィルタの伝達関数は

$$\begin{aligned}
 H(z) &= H_0(z) \cdot H_1(z) \cdot H_2(z) \\
 &= A \cdot P_{N_1}^{(*)}(F_1(z)) \cdot P_{N_2}^{(*)}(F_2(z)) \quad (6.4)
 \end{aligned}$$

で表される。ここで、 $H_0(z) = A$ は定数倍であり、その値は

$$A = S_z^{-1}(\hat{c}_0) = \exp(\hat{c}_0) = c_0 \quad (6.5)$$

である。 $H_1(z)$, $H_2(z)$ はそれぞれ $F_1(z)$, $F_2(z)$ を基礎フィルタとする一般化対数特性近似デジタルフィルタである。このうち、 $F_1(z)$ は一般化ケプストラムの低次数部を表し、 $F_2(z)$ は一般化ケプストラムの低次数部を分離した残りの部分である。

基礎フィルタの伝達関数 $F_1(z)$, $F_2(z)$ の選び方によりいろいろ構成が考えられるが、ここでは5.3節で示した2通りの構成の実現可能性を考える。

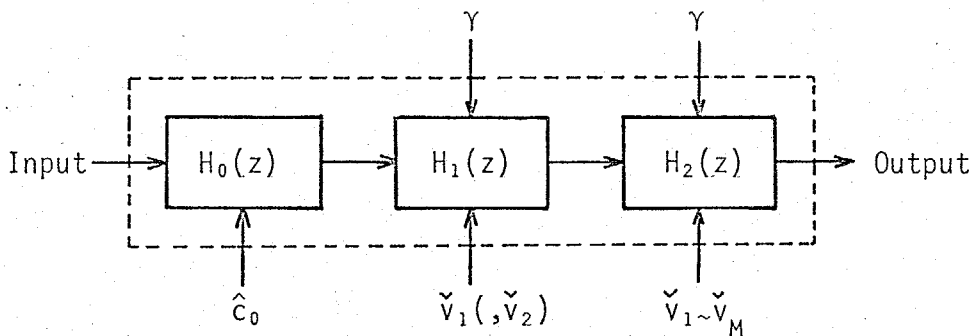


Fig. 6.2. Speech synthesis filter.

(構成 I)

一般化対数スペクトル包絡を \check{v}_1 の項と残りの項に分割して近似する。基礎フィルタの伝達関数はそれぞれ

$$F_1(z) = \check{v}_1 z^{-1} \quad (6.6)$$

$$F_2(z) = \sum_{m=2}^M \check{v}_m z^{-m} / (1 + \delta \check{v}_1 z^{-1}) \quad (6.7)$$

となる。

(構成 II)

一般化対数スペクトル包絡を \check{v}_1, \check{v}_2 を分離して近似する。基礎フィルタの伝達関数はそれぞれ

$$F_1(z) = \check{v}_1 z^{-1} + \check{v}_2 z^{-2} \quad (6.8)$$

$$F_2(z) = \sum_{m=3}^M \check{v}_m z^{-m} / (1 + \delta \check{v}_1 z^{-1} + \delta \check{v}_2 z^{-2}) \quad (6.9)$$

で与えられる。

構成 I, II いずれの場合も、基礎フィルタの安定性が問題となるが、これは \check{v}_1, \check{v}_2 および δ の値に依存する。構成 I の場合、 $|\check{v}_1| = |\hat{c}_1|$ の最大値を $\hat{c}_{1\max}$ とすれば、 $F_2(z)$ が安定であるためには式(5.34)から

$$\hat{c}_{1\max} < \frac{1}{|\delta|} \quad (6.10)$$

であればよい。また構成 II の場合には、 \check{v}_2 の正の最大値を $\check{v}_{2\max}$ 、負の最小値を $-\check{v}_{2\min}$ ($\check{v}_{2\min} > 0$) とすれば、 $F_2(z)$ が安定であるためには、図 5.4 から

$$\begin{cases} \hat{c}_{1\max} + \check{v}_{2\min} < \frac{1}{\delta}, & \delta > 0 \\ \check{v}_{2\max} < \frac{1}{\delta}, & \delta > 0 \end{cases} \quad (6.11)$$

$$\left\{ \begin{array}{l} \hat{c}_{1 \max} + \check{v}_{2 \max} < -\frac{1}{\delta}, \quad \delta < 0 \\ \check{v}_{2 \min} < -\frac{1}{\delta}, \quad \delta < 0 \end{array} \right. \quad (6.12)$$

であればよい。なお $\delta = 0$ の場合、すべて安定であることは明らかである。

図6.3は、天気予報からとった音声(約48秒, 9600Hz, 2.5節で用いた音声資料と同一)について、 $\check{v}_2 = \hat{c}_1$ の分布を示したものである。これから $\hat{c}_{1 \max}$ は3以下であることがわかる。他の音声でも、 $\hat{c}_{1 \max}$ が3を取えることは殆どないことが経験的に確かめられている。従って、構成Iの場合、 $|\delta| \leq 1/3$ では常に実現可能と考えてよい。

構成IIの場合、 δ の値により $\check{v}_{2 \max}$, $\check{v}_{2 \min}$ の値が変化するため、それぞれ δ に対して安定性の検討が必要となる。ここでは、2.5節で \check{v}_2 の分布を調べた δ の範囲内で安定性を調べておく。 $\delta = 0.2$ の場合、 $\check{v}_{2 \min} < 0.9$, $\check{v}_{2 \max} < 2.0$ および前述のように $\hat{c}_{1 \max} < 3.0$ であるから、式(6.11)を満足し安定である。 $0 < \delta < 0.2$ の場合も同様にして安定であることが確かめられる。 $\delta = -0.2$ の場合、 $\check{v}_{2 \max} < 1.5$, $\check{v}_{2 \min} < 1.0$ であり、これと $\hat{c}_{1 \max} < 3.0$ から式(6.12)の条件を満足している。 $-0.2 < \delta < 0$

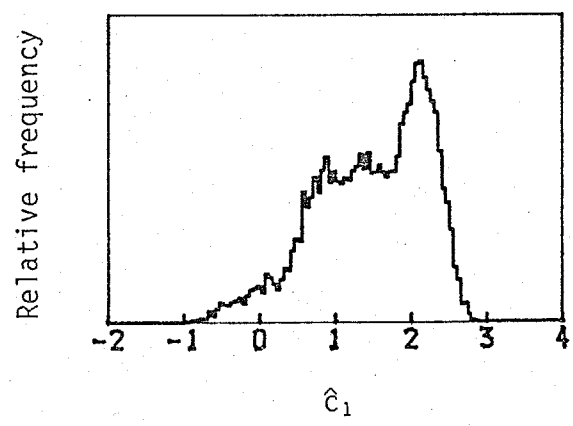


Fig. 6.3. Histogram of improved cepstrum \hat{c}_1 .

の場合も同様であり、結局 $|\delta| \leq 0.2$ で構成Ⅱのデジタルフィルタの実現が可能である。なお、同一の音声資料については、 $|\delta| \leq 1/3$ の範囲で安定となることが確かめられている。

構成Ⅱの場合、 $\check{\nu}_1, \check{\nu}_2$ の値の組合せによっては、パデ近似の誤差が十分小さくできない可能性が考えられる。このようなときは、式(6.8)の特性をさらに分割し、

$$F_{11}(z) = \check{\nu}_1 z^{-1} \quad (6.13)$$

$$F_{12}(z) = \check{\nu}_2 z^{-2} / (1 + \delta \check{\nu}_1 z^{-1}) \quad (6.14)$$

をそれぞれ基礎フィルタの伝達関数とする一般化対数特性近似デジタルフィルタ $H_{11}(z), H_{12}(z)$ を縦続接続して

$$\begin{aligned} H(z) &= H_{11}(z) \cdot H_{12}(z) \\ &= P_{N_1}^{(a)}(F_{11}(z)) \cdot P_{N_1}^{(a)}(F_{12}(z)) \end{aligned} \quad (6.15)$$

とすればよい。 $F_{12}(z)$ の安定条件は構成Ⅰの $F_2(z)$ の安定条件と同じになる。従って、 $|\delta| \leq 1/3$ の範囲で常に実現できる。

6.3 一般化対数関数のパラメータと合成音声の品質

6.3.1 一般化ケプストラムを用いる音声の合成

音声の合成に一般化ケプストラムを用いる効果を調べるために、6.2節で示した合成系を用いて合成を行った。ここでは、合成系に伝送する分析パラメータの抽出法および合成フィルタの近似誤差について述べる。

合成に用いた音声サンプルは、2.5節で用いた音声資料(天気予報)の中から取り出した次の3種類である。

(1) 「南部では東の風」 男声, 約 1.5 [s]

(2) 「関東地方の」 男声, 約 1.0 [s]

(3) 「明日は北の風」 女声, 約 1.2 [s]

原音声は遮断周波数 5 kHz のローパスフィルタを通した後, 10 [kHz] でサンプリングし 12 ビットに量子化しておく。これを分析長 25.6 [ms] のブロックマン窓をかけたケプストラム法あるいは改良ケプストラム法を用いてケプストラムを抽出する。改良ケプストラム法における加速係数は $\alpha = 1.0$, 繰り返し回数は $J = 3$ と選んだ。

ブロック抽出は, まず有聲/無聲の判断を改良ケプストラムを用いた低周波領域の平均レベルにより行う⁽²⁶⁾。すなわち, 改良ケプストラム法により得られたスペクトル包絡を $\hat{S}_v(\omega)$ として, 低周波領域での平均レベル B_v を

$$B_v = \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \hat{S}_v(\omega) d\omega}{\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) d\omega} \quad (6.16)$$

但し,

$$W(\omega) = \begin{cases} 1, & \omega_1 < |\omega| < \omega_2 \\ 0, & \text{otherwise} \end{cases} \quad (6.17)$$

を定義し、 B_p と適当な閾値 B と比較して、 $B_p \geq B$ ならば有聲、 $B_p < B$ ならば無聲と判断する。これは、有聲区間ではピッチ周期によって決まる比較的パワーの大きい基本波成分の存在により、低周波領域でのスペクトル包絡の値が相対的に大きくなるのに対し、無聲区間ではパワーの小さい乱流雑音が音源となっているため、低周波領域でのスペクトル包絡の値は相対的に小さくなるという事実に基づいている。重み関数の形は、 ω_1, ω_2 の値によって決まるが、これらの値は有聲区間で基本波成分によるスペクトル包絡の値の上昇分的確にとらえるために、通常の音声で基本周波数が存在する範囲に設定する必要がある。ここでは、 $f_1 = \omega_1 / 2\pi \approx 80$ [Hz] および $f_2 = \omega_2 / 2\pi \approx 320$ [Hz] となるように選ぶ。また、閾値 B の値は経験的に決めた。上記の方法によれば、殆ど確実に有聲/無聲の判断を行うことができる。

次に、有聲であると判断された場合には、窓長 40 [ms] のフラックマン窓で音声信号を取り出し、その前後両側にそれぞれ 56 点の 0 を付け加え、512 点の DFT を用いてケプストラム $\hat{C}_m^{(p)}$ を求める。さらに次式の重みつきケプストラム

$$\tilde{C}_m^{(p)} = \left(1 + \frac{m}{64}\right) C_m^{(p)}, \quad 29 \leq m \leq 151 \quad (6.18)$$

の値を計算し

$$I_p = m \max_{30 \leq m \leq 150} \left\{ \tilde{C}_{m-1}^{(p)} + 2\tilde{C}_m^{(p)} + \tilde{C}_{m+1}^{(p)} \right\} \quad (6.19)$$

をピッチ周期とする。ここで p は分析フレーム番号である。

無聲と判断された区間では、 $I_p = 0$ とする。

以上のようにして得られた分析パラメータを、伝送周期 5 [ms] で合成系へ伝送する。ここでは、伝送パラメータの量子化は行わないことにする。

図 6.4 は、品質評価に用いた音声サンプル「南部では東の風」の一部分の原音声波形、およびこれに対応するスペクトル包絡をいくつかの m の値につい

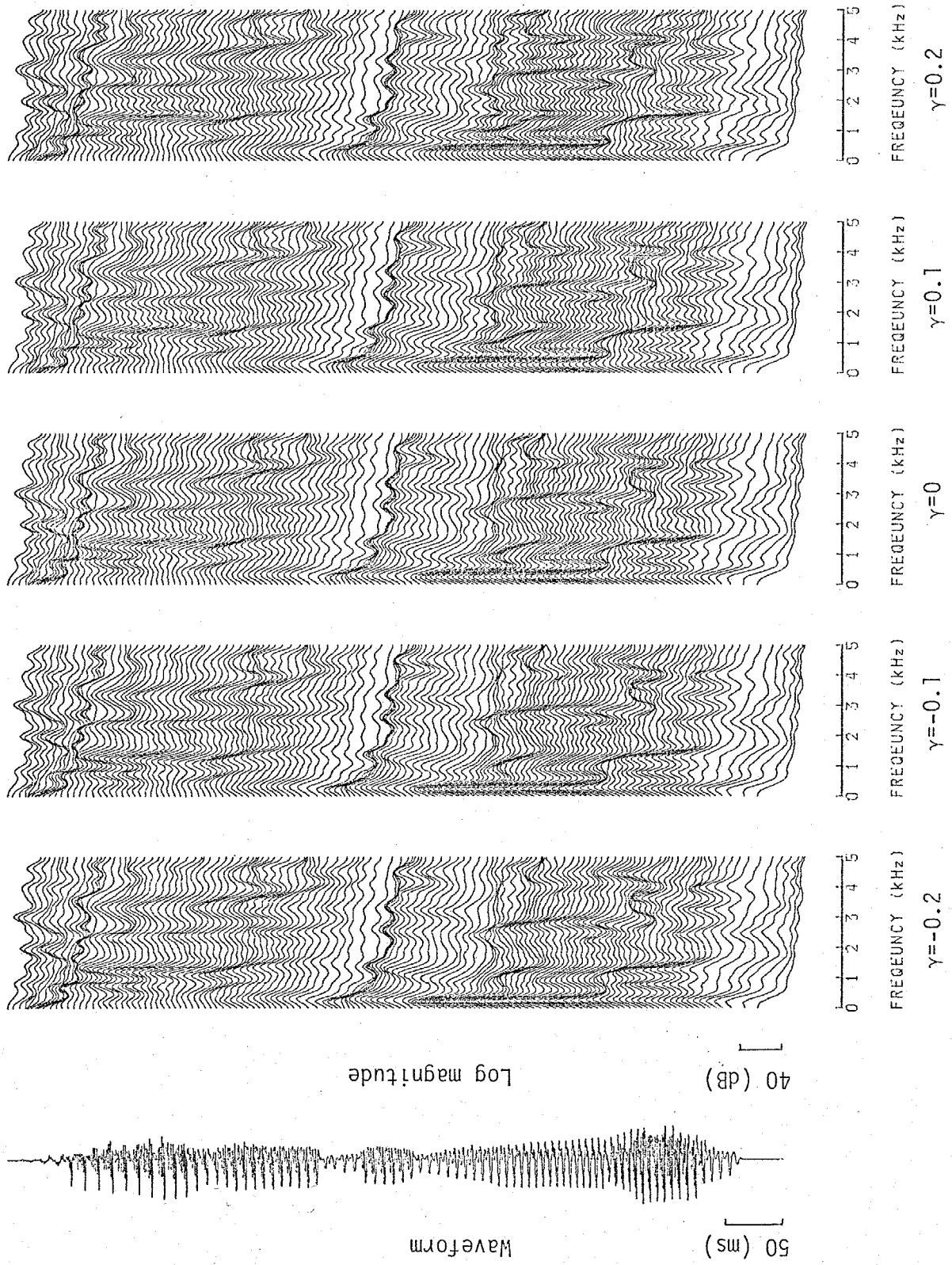


Fig. 6.4. Waveform and spectral envelopes for the male speech /nanbudewa/.

て描いたものである。ケプストラムは改良ケプストラム法を用いて抽出し、次数は $M=20$ である。なお、スペクトル包絡はすべて対数振幅目盛上に描いてある。第2章のスペクトル包絡抽出例でも述べたように、 γ が負の値になるにつれて、スペクトルの極に相当する部分が強調されていく様子が見られる。

合成に用いる一般化対数特性近似デジタルフィルタとして、ここでは、構造が簡単な構成Iを用いることにする。これは、基礎フィルタの伝達関数として、式(6.6)、(6.7)を用いるものである。

図6.5に、パデ近似の次数 N_1, N_2 をそれぞれ4として、音声サンプル「南部ては東の風」のスペクトル包絡を近似したときの、一般化対数目盛における誤差

$$\check{E}_2(\omega) = \check{H}(e^{j\omega}) - \check{S}(\omega) \quad (6.20)$$

の最大値 $\max |\check{E}_2(\omega)|$ および、対数振幅目盛における誤差

$$\hat{E}_A(\omega) = |\ln |H(e^{j\omega})| - \ln |S(\omega)|| \quad (6.21)$$

の最大値 $\max \hat{E}_A(\omega)$ を示す。ここで、 $H(z)$ は、 $F_1(z), F_2(z)$ をそれぞれ式(6.6)、(6.7)として、式(6.4)で表される。

対数振幅目盛での近似誤差を見ると、いずれの場合も、誤差は0.15 [dB] を越えることはなく、実用上殆ど無視できる値であることがわかる。なお、 $|\gamma| = 1/4$ とすれば、一般化対数関数は4次の有理式となり、4次のパデ近似は正確にもとの一般化対数関数の逆関数を表すことから、誤差は0となる。

近似誤差が非常に小さいことからわかるように、いずれの場合も、フィルタは安定である。

ここで用いた音声サンプルに限らず、通常の音声に対しても、構成Iの合成フィルタで、 N_1, N_2 を4に選んでおけば、近似誤差は殆ど無視することができ、フィルタも安定であると考えてよい。

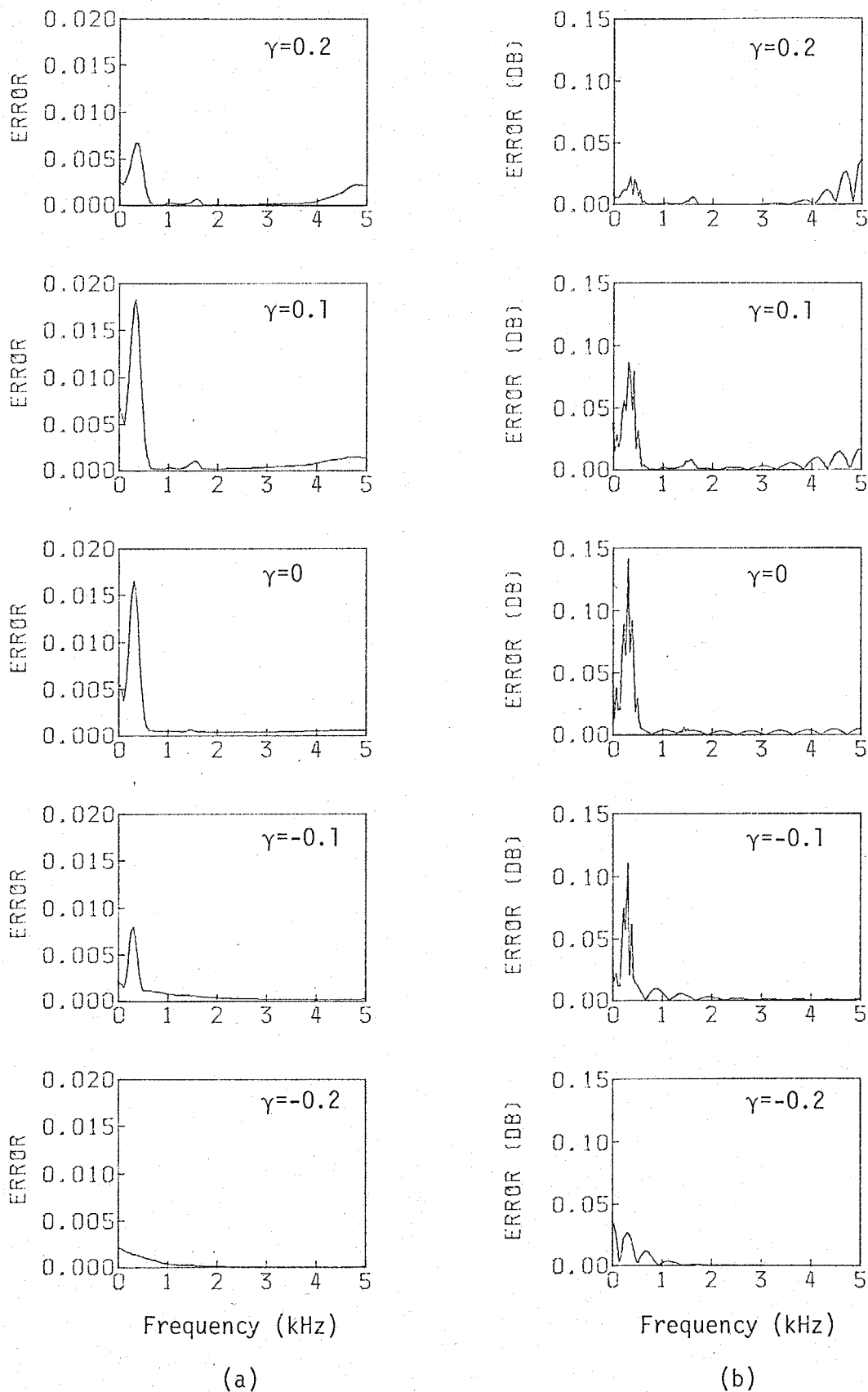


Fig. 6.5. Curves of approximation error for the synthesis filters:
 (a) $\max_{\nu} |\check{E}_{\gamma}(\omega)|$ versus ω , (b) $\max_{\nu} \hat{E}_A(\omega)$ versus ω .

6.3.2 合成音声の品質評価

ここでは、一般化対数関数のパラメータ δ の値と、得られた合成音声の品質の関係について調べた結果を述べる。

まず、任意の δ について、一般化ケプストラムと改良一般化ケプストラムを比較すると、改良一般化ケプストラムを用いて得られた合成音の方が、明らかに原音声に近かった。これは、通常のケプストラム法の場合でも同じである。

δ の値を正に選んだ場合、得られる合成音声は、 $\delta=0$ すなわち通常のケプストラム法に比べ、響きのないつぶれたような感じの音になる。これは、図

6.4 あるいは第2章で示したスペクトル包絡の形からわかるように、スペクトル包絡の極に相当する部分が、抑えられた形になるためと考えられる。特に

$0.2 < \delta$ の範囲では、明瞭性が下がるのが、きりとわかる。ただし、 δ の値がそれ程大きくないところでは、改良ケプストラム法でしばしば生じる、スペクトル包絡の著しく急峻なピークが抑えられることが、フレーム方向に対するスペクトルの平滑化と同じ結果となるためか、むしろ聞きやすい音になる場合もあった。

逆に、 δ の値を負に選んだ場合、通常のケプストラム法と比べて得られる合成音声は、やや響きがあり明瞭性も良くなったように聞こえる。これは、スペクトル包絡の極に相当する部分が強調されるためであると考えられる。しかし、 δ の値を負方向に大きくしていくと、しだいに合成音声は不安定な感じが増してきて、特に $\delta < -0.2$ では、かえって聞きにくくなる部分が出てくる。

δ の値に対する合成音声の品質の変化を、男声と女声について比較すると、男声の方が大きかった。

次に、合成音声の品質の定量的な評価の一つとして、対比較試験を行った結果を示す。

合成音声のサンプルは、前述の3つの原音声(男声2, 女声1)に対して、 δ の値をそれぞれ $-0.2, -0.1, 0, 0.1, 0.2$ に選び、改良一般化ケプストラムを用いて合成したものを用いた。合成の際の条件は δ の値を除いてす

べて同じにし、ケプストラム次数は、男声が20次、女声が15次とした。

試験の内容は、 α の値を変えた5種類の合成音声から2つを選んで対を作り、これを、順序を入れたすべての組合せ（ここでは20通り）について、一定時間間隔で被験者に聞かせ、対のどちらがより「はっきりとした (sharp)」か、あるいは、より「好ましい (preferable)」かを判断させるものである。特に後者の評価基準による対比較試験は、フレファレンステストとして良く知られている。

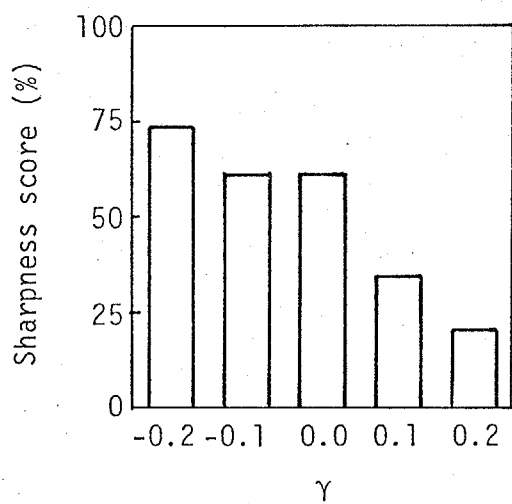
被験者は成人男性8人であり、ヘッドフォンによる両耳聴取で行った。

図6.6に、男声「南部では東の風」について、また図6.7に、男声「関東地方の」について、対比較試験を行った結果を示す。評価基準は、どちらの図でも、(a)が「はっきりした」で、(b)が「好ましい」である。図の縦軸のスコアは、一つのサンプルと他のサンプルのすべての組合せについて、同じサンプルだけを良いと判断したとき、そのサンプルのスコアが100%となるように算出してある。

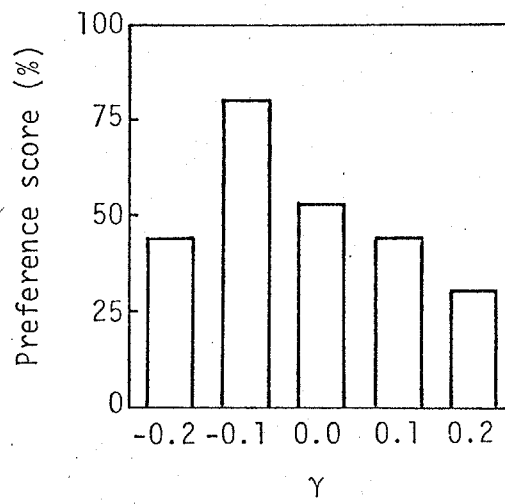
この図から、 α が負の値ほど、合成音声の品質が「はっきりしている」と判断されていることがわかる。これは、スペクトル包絡の山の部分を強調した結果であると考えられる。一方、音の「好ましさ」からすると、 α の値が-0.1で、最も高いスコアを示している。これは、被験者の好みにも関係するが、一般に α の値が負の方向に大きくなると、響きのある感じは増えるが、それと共に不安定な感じも出てくることによると考えられる。つまり、適度にはっきりとしていて響きがあり、不安定な感じもそれ程しない、というのが $\alpha = -0.1$ の場合ということになる。また、フレファレンススコアの値は、原音声のサンプルの違いによっても差がでている。この場合は、「南部では東の風」の方が α の値による合成音声の品質の差が大きいようである。

図6.8は、女声「明日は北の風」についての対比較試験の結果である。

図6.8(a)は、評価基準が「はっきりとした」の場合である。必ずかの差異はあるが、殆どのスコアが50%に近い値であることから、品質の差は男声

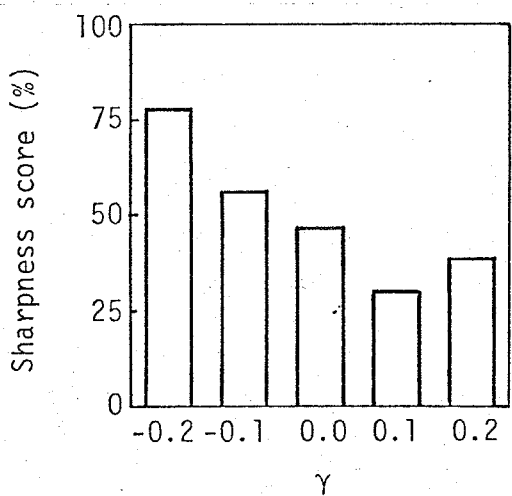


(a)

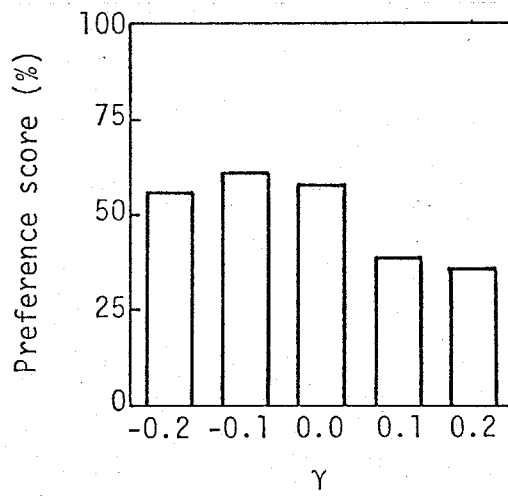


(b)

Fig. 6.6. (a) Sharpness score, (b) preference score of each synthesized speech "nanbu dewa higashi no kaze".



(a)



(b)

Fig. 6.7. (a) Sharpness score, (b) preference score of each synthesized speech "kantō-chihō no".

程大きく出てないことがわかる。図6.8(b)はフリツァレンススコアである。この場合は、 $\gamma = 0$ と -0.1 でスコアが高く、少しの違いではあるが、 $\gamma = 0$ となれば通常のセプストラムを用いた合成音声か、最も高いスコアを得ている。

以上の結果から判断すると、音声の分析合成のための評価目盛としては、対数目盛があるいは、 γ の値をやや負にした一般化対数目盛が良いことになる。また、 $|\gamma|$ の値が大きい極端なべき乗目盛、例えば $|\gamma| = 1/3$ 等は好ましくないことになる。

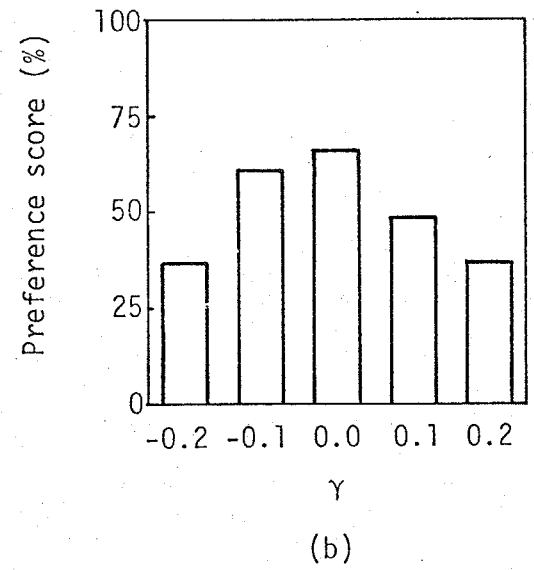
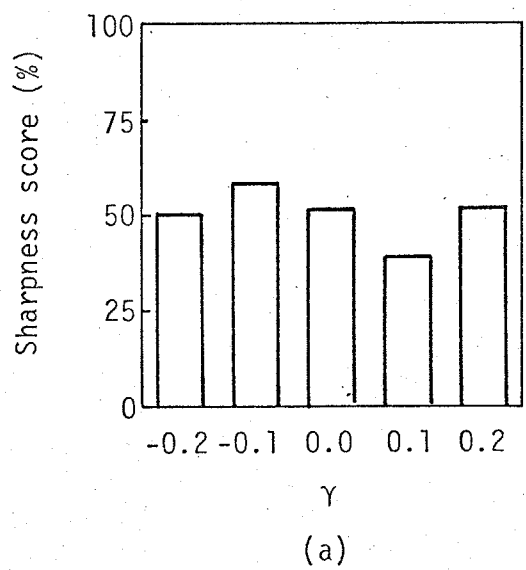


Fig. 6.8. (a) Sharpness score, (b) preference score of each synthesized speech "asu wa kita no kaze".

6.4 むすび

一般化対数特性近似ディジタルフィルタを合成フィルタに用いた、一般化ケプストラムによる音声の合成法を示した。本章に示した合成系は、従来のケプストラム法による合成系と完全な互換性があり、一般化ケプストラムによる合成法の一つの特別な場合として、通常のケプストラムによる合成もできる。一般化対数のパラメータを変化させることにより、合成音声の品質を変化させることができ、プレファレンスラストの結果から、 γ の値を適当に選べば、従来のケプストラム法に比べ、より品質の良い合成音を得られる可能性があることがわかった。

本章で示した合成系は、ケプストラム法との互換性、 γ の可変性を考慮した構成である。たが、 γ の値を適当な値、例えばプレファレンススコアが高かった $\gamma = -0.1$ に固定し、伝送パラメータも正規化一般化ケプストラムとすれば、合成系の構成はより簡単に行うことができる。

最適な γ の値の選択に関するより詳細な検討、伝送パラメータの量子化を行った場合の、 γ の値と合成音声の品質の関係等については、今後の課題である。

第7章 結論

本研究では、対数関数のかわりに、その自然な拡張である一般化対数関数という新たな評価目盛を導入することにより、ケプストラム法を一般化することを試みた。理論的に優れた音声の分析・合成法でも、実際に際して複雑な処理が要求されたり、合成系の実現が困難であっては意味がない。その点、本研究で示した分析・合成法は、従来のケプストラム法を含みながら自然に一般化した形となっており、分析・合成における処理はケプストラム法と比べて殆どかわらない特長がある。さらに、実際に合成系を構成し、得られた合成音声の品質の検討を行った結果、一般化対数関数のパラメータの値により、品質を変化させることができ、また、従来のケプストラム法と比べて品質が向上する可能性があることがわかった。

まず第2章では一般化対数関数を定義し、この関数を評価目盛とする音声のスペクトル包絡の抽出法を示した。一般化対数目盛上の表現パラメータである一般化ケプストラムは、通常のケプストラムと似た性質を持ち、これを用いることにより、同一次数のケプストラムでは表現することができないスペクトル包絡の表現が可能であることを示した。一般化ケプストラムは、通常のケプストラムから簡単な変換式により計算することができ、また一つの特別な場合としてケプストラムとも一致することから、音声のスペクトル包絡を表現する際の有用なパラメータであると考えられる。

第3章から第5章までは、一般化ケプストラムを用いる合成系を構成する際最大の問題となる、合成フィルタの特性近似法について述べた。

第3章では、一般化対数目盛の一つの特別な場合である対数目盛について、周波数領域での2乗平均誤差評価式を用いた、デジタルフィルタの特性近似法を示した。この方法によれば、フィルタの係数は簡単な計算アルゴリズムで求まり、安定なフィルタが得られる。また、具体的な近似例から、ケプストラ

ムの次数と同一の遅延素子数のフィルタの構成で、十分実用となる対数近似特性が得られることを示した。

第4章では、最小位相インパルス応答と時間領域の誤差評価式を用いた、デジタルフィルタの対数または一般化対数特性の近似法を示した。この方法でも、得られるフィルタは安定であり、希望特性のケプストラムの次数と同じがそれ以下の遅延素子数の構成で、良好な近似特性が得られる。第3章で示した近似法およびここで示した近似法は、音声合成フィルタを構成する際、遅延素子数を少ない構成で実現したい目的には有効な方法であると考えられる。

第5章では、一般化対数特性を一般化対数目盛上で直接近似するデジタルフィルタの構成法を示した。この方法は、一般化ケプストラムの値がそのままフィルタ係数となるため、フィルタ係数を求めるための計算が不要であり、一般化対数目盛上での近似誤差の最大値を規定できるものであった。従って、実時間で動作する分析合成系を構成する場合には、非常に有用となる近似法である。

第6章では、第2章で述べた一般化ケプストラムを合成パラメータとし、合成フィルタとして、第5章で示した一般化対数特性近似デジタルフィルタを利用する合成系を示した。ここで示した合成系は、従来のケプストラム法による合成系と完全に互換性があり、分析系は、ケプストラム法で用いられる系を変更することなしにそのまま利用できる。ここでは、合成フィルタの実現に関する検討や、実際に得られた合成音の品質評価を行った。その有効性については、本章の最初に述べた通りである。

最後に、本研究ではとりあげることができなかった問題点をあげ、本論文のむすびとしたい。その一つは、本研究では分析合成系を実現する際、すべて電子計算機を用いてソフトウェア的なシミュレーションによって行ったが、これをハードウェアで実現するための諸問題点の検討がある。この点は、主として合成フィルタとして用いる、一般化対数特性近似デジタルフィルタのハードウェア実現に関する問題点（係数の量子化、演算語長、感度等）が占める。

一方、一般化対数関数のパラメータの最適値の選択に関して、いろいろな角度から、より詳細な検討を加えることも重要な問題点である。例えば、原音声サンプルの種類を多くするだけでなく、一般化ケプストラムを伝送パラメータとして、それを量子化したときの情報伝送量と合成音声の品質の関係から調べることなどが考えられる。また、本研究で得られた成果の応用として、一般化ケプストラムの認識パラメータへの利用、周波数軸に非線形周波数軸を用いた、マルチ一般化ケプストラムの合成パラメータへの利用等の問題もあげることができ

謝 辞

本研究を進めるにあたり、終始適切な御助言と温かい御指導をいただいた
今井聖教授に心から感謝いたします。また、有益な御助言・御討論をいただいた
北村正博士に深く感謝いたします。さらに、日頃、音質評価試験等で御協力
下さった今井研究室の皆様にも心から感謝いたします。特に、住田一男、麻生寛
両氏には論文の校正の助力をいただいたことも合わせて感謝いたします。

参考文献

(1) Bogert, B. P., Healy, M. and Tukey, J.: "The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and shape cracking", Proc. Symp. on Time Series Analysis, M. Rosenblatt, Ed., Ch. 15, p.209, J. Wiley, New York (1963).

(2) Oppenheim, A. V. and Schafer, R. W.: "Homomorphic analysis of speech", IEEE Trans. Audio & Electroacoustics, AU-16, p.221 (June 1968).

(3) Markel, J. D. and Gray, A. H.: "Linear Prediction of Speech", Springer-Verlag, New York (1976).

(4) Makhoul, J.: "Linear prediction: A tutorial review", Proc. IEEE, vol. 63, p.561 (1975).

(5) Oppenheim, A. V.: "A speech analysis-synthesis system based on homomorphic filtering", J. Acoust. Soc. Am., vol. 45, p.458 (Feb. 1969).

(6) Weinstein, C. J. and Oppenheim, A. V.: "Predictive coding in a homomorphic vocoder", IEEE Trans. Audio & Electroacoustics, AU-19, p.243 (Sept. 1971).

(7) 今井, 北村: "対数振幅特性近似フィルタを用いた音声の分析合成系", 信学論(A), J61-A, 6, p.527 (昭53-06)

(8) 今井, 北村, 竹谷: "デジタルフィルタの対数伝達特性の直接近似", 信学論(A), J59-A, 2, p.157 (昭51-02)

(9) Lim, J. S.: "Spectral root homomorphic deconvolution system", IEEE Trans. Acoust., Speech & Signal Process., ASSP-27, p.223 (June 1979).

(10) Oppenheim, A. V. and Schafer, R. W.: "Digital Signal Processing", Prentice-Hall, Englewood Cliffs, N.J. (1975).

(11) 今井, 阿部: "改良ケプストラム法によるスペクトル包絡の抽出", 信学論(A), J62-A, 4, p.217 (昭54-04)

- (12) Johnson, A. T.: "Magnitude equalization using digital filters", IEEE Trans. Circuit Theory, CT-20, p.308 (May, 1973).
- (13) 今井聖: "ケフレンシ領域におけるデジタルフィルタ設計", 信学論 (A), J60-A, 10, p. 944 (昭52-10).
- (14) Zohar, S.: "The solution of a Toeplitz set of linear equations", J. Assoc. Comput. Mach., 21, 2, p.272 (April, 1974).
- (15) 小林隆夫: "2乗平均誤差評価によるデジタルフィルタの対数特性近似", 修士論文, 東京工業大学 (1978).
- (16) Harris, F. J.: "On the use of windows for harmonic analysis with the discrete Fourier transform", Proc. IEEE, vol. 66, p.51 (Jan. 1978)
- (17) Tretter, S. A.: "The all-pole model is theoretically stable", IEEE Trans. Audio & Electroacoustics, AU-20, p.316 (Oct. 1972)
- (18) Szego, G.: "Orthogonal Polynomials", CH. 11, American Mathematical Society Publications, New York (1939).
- (19) Kopec, G.E., Oppenheim, A. V. and Tribolet, J. M.: "Speech analysis by homomorphic prediction", IEEE Trans. Acoust., Speech & Signal Process., ASSP-25, p.40 (Feb. 1977).
- (20) Mullis, C. T. and Roberts, R. A.: "The use of second-order information in the approximation of discrete-time linear systems", IEEE Trans. Acoust., Speech & Signal Process., ASSP-24, p.226 (June 1976).
- (21) Scharf, L. L. and Luby, J. C.: "Statistical design of autoregressive-moving average digital filters", IEEE Trans. Acoust., Speech & Signal Process., ASSP-27, p.240 (June 1979).
- (22) Burrus, C. S. and Parks, T. W.: "Time domain design of recursive digital filters", IEEE Trans. Audio & Electroacoust., AU-18, p.137 (June 1970).

- (23) 今井聖: "対数振幅近似(LMA)フィルタ", 信学論(A), J63-A, 12, p.886 (昭55-12).
- (24) 森正武: "数値解析と複素関数論", 数理科学シリーズ7, 筑摩書房 (1975).
- (25) S.L.S. ジャコビ, J.S. コワック, J.T. ケーソ著, 関根訳: "非線形最適化問題の反復解法", 培風館 (1976).
- (26) 阿部, 今井: "CV音節のケフストラムパラメータからの音声合成", 信学論(D), J64-D, 9, p.861 (昭56-09).
- (27) E.W. チェーニ著, 一松, 松島訳: "近似理論入門", 共立出版 (1977)

本研究に関する発表文献

- (1) 小林, 今井: "2乗平均誤差評価による非再帰形ディジタルフィルタの対数特性近似", 信学論(A)技術談話室, J62-A, 3, p. 203 (昭54-03)
- (2) 小林, 今井: "2乗平均誤差評価による再帰形ディジタルフィルタの対数振幅特性近似法", 信学論(A), J62-A, 9, p. 576, (昭54-09)
- (3) 小林, 今井: "最小位相インパルス応答を用いたIIRディジタルフィルタの対数振幅特性近似", 電子通信学会総合全国大会, 14, p. 1-14 (昭55)
- (4) Kobayashi, T. and Imai, S.: "Log magnitude approximation technique for IIR digital filters using minimum phase impulse response", Bull. P.M.E.(T.I.T.), No.46, p.33 (1980).
- (5) 小林, 今井: "一般化対数目盛における音声スペクトルの平滑化法", 信学論(A)技術談話室, J64-A, 6, p. 473 (昭56-06)
- (6) Kobayashi, T. and Imai, S.: "Spectral smoothing technique for speech analysis on a generalized logarithmic scale", Bull. P.M.E.(T.I.T.), No. 48, p.31 (1981).
- (7) 小林, 今井: "一般化ケフストラムによる音声の合成", 信学論(A) 投稿中(採録決定済)

付録

A-1 式(4.8)の解法

式(4.8)の解法については文献(20.)で詳細に述べられているが、ここでは、簡単にそのアルゴリズムを紹介しておく。

($K=L$ の場合)

初期値として

$$\left\{ \begin{array}{l} \alpha_0 = r_0 - g_0^2 \\ a_0^{(0)} = 1 \\ u_0^{(0)} = 1/\alpha_0 \\ v_0^{(0)} = g_0/\alpha_0 \\ \delta_0 = 1 + g_0^2/\alpha_0 \end{array} \right. \quad (\text{A.1})$$

と置く。 $a_{n+1}^{(m)} = u_{n+1}^{(m)} = v_{n+1}^{(m)} = 0$ として、次の一連の計算を $n=0, 1, \dots, K-1$ について繰り返す。

$$\lambda_n = \sum_{k=0}^n a_k^{(m)} r_{n+1-k} \quad (\text{A.2a})$$

$$\mu_n = - \sum_{k=0}^n a_k^{(m)} g_{n+1-k} \quad (\text{A.2b})$$

$$\theta_n = \lambda_n u_0^{(m)} + \mu_n v_0^{(m)} \quad (\text{A.2c})$$

$$\phi_n = \lambda_n v_0^{(m)} + \mu_n \delta_n \quad (\text{A.2d})$$

$$\alpha_{n+1} = \alpha_n - \lambda_n \theta_n + \mu_n \phi_n \quad (\text{A.2e})$$

$$\delta_{n+1} = \delta_n + \phi_n^2 / \alpha_{n+1} \quad (\text{A.2f})$$

$$a_k^{(n+1)} = a_k^{(n)} - \lambda_n u_{n+1-k}^{(n)} - \mu_n v_{n+1-k}^{(n)}, \quad 0 \leq k \leq n+1 \quad (\text{A.2g})$$

$$u_k^{(n+1)} = u_k^{(n)} - \frac{\theta_n}{\alpha_{n+1}} \cdot a_{n+1-k}^{(n+1)}, \quad 0 \leq k \leq n+1 \quad (\text{A.2h})$$

$$v_k^{(n+1)} = v_k^{(n)} - \frac{\theta_n}{\alpha_{n+1}} \cdot a_{n+1-k}^{(n+1)}, \quad 0 \leq k \leq n+1 \quad (\text{A.2i})$$

解は,

$$a_k = a_k^{(K)}, \quad k=1, 2, \dots, K \quad (\text{A.3})$$

で与えられる。

($K < L$ の場合)

$i = L - K$ とし、式 (A.2a), (A.2b) 中の g_m, r_m をそれぞれ

$$g'_m = \begin{cases} 0, & m < 0 \\ g_{m+i}, & m \geq 0 \end{cases} \quad (\text{A.4})$$

$$r'_m = r_m - \sum_{k=0}^{i-1} g_k g_{k+m} \quad (\text{A.5})$$

と置き換えて、式 (A.1) ~ (A.3) により a_k を求める。

($K > L$ の場合)

$i = K - L$ とし、式 (A.2b) 中の g_m を

$$g'_m = \begin{cases} 0, & m < i \\ g_{m-i}, & m \geq i \end{cases} \quad (\text{A.6})$$

で置き換えて、式 (A.1) ~ (A.3) により a_k を求める。

A.2 一般化対数関数の逆関数のパデ近似

ここでは、一般化対数関数の逆関数のパデ近似である式(5.2)~(5.4)の導出を示す。

一般に近似すべき関数 $f(w)$ が「原点のまわりで」

$$f(w) = \sum_{n=0}^{\infty} c_n w^n \quad (A.7)$$

と表されるとき、 $f(w)$ の次数 (N, N) のパデ近似

$$P_N(w) = \frac{B(w)}{A(w)} = \frac{\sum_{k=0}^N B_k w^k}{1 + \sum_{k=1}^N A_k w^k} \quad (A.8)$$

の係数 A_k, B_k は次式の解で与えられる。(27)

$$\begin{pmatrix} c_N & c_{N-1} & \dots & c_1 \\ c_{N+1} & c_N & \dots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_{2N+1} & \dots & c_N \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix} = \begin{pmatrix} c_{N+1} \\ c_{N+2} \\ \vdots \\ c_{2N} \end{pmatrix} \quad (A.9)$$

$$\begin{pmatrix} B_0 \\ B_1 \\ \vdots \\ B_N \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_N \end{pmatrix} + \begin{pmatrix} 0 & \dots & \dots & 0 \\ c_0 & & & \\ c_1 & c_0 & & \\ \vdots & \vdots & \ddots & \\ c_{N-1} & \dots & c_2 & c_0 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix} \quad (A.10)$$

よって、 $S_r^{-1}(w)$ の次数 (N, N) のパデ近似を

$$P_N^{(\sigma)}(w) = \frac{B_N(\sigma, w)}{A_N(\sigma, w)} = \frac{\sum_{k=0}^N B_{N/k}^{(\sigma)} w^k}{1 + \sum_{k=1}^N A_{N/k}^{(\sigma)} w^k} \quad (A.11)$$

と表すことにすれば、 $S_r^{-1}(w)$ のテイラー展開である式(2.9)

$$S_{\delta}^{-1}(W) = 1 + W + \sum_{n=2}^{\infty} \frac{(1-\delta)(1-2\delta)\dots(1-(n-1)\delta)}{n!} W^n$$

$$= \sum_{n=0}^{\infty} \frac{f_{\delta}(n)}{n!} W^n \tag{A.12}$$

但し

$$f_{\delta}(n) = \begin{cases} 1, & n=0, 1 \\ (1-\delta)(1-2\delta)\dots(1-(n-1)\delta), & n > 1 \end{cases} \tag{A.13}$$

および式(A.9)から, $A_{N,k}^{(\delta)}$ は次式の解として得られる。

$$K(N)\alpha = -d \tag{A.14}$$

ここで, $\alpha = (A_{N,1}^{(\delta)}, \dots, A_{N,N}^{(\delta)})^T$ であらう。

$$K(N) = \begin{pmatrix} \frac{f_{\delta}(N)}{N!} & \dots & \frac{f_{\delta}(2)}{2!} & 1 \\ \frac{f_{\delta}(N+1)}{(N+1)!} & \dots & \frac{f_{\delta}(2)}{2!} & \\ \vdots & \dots & \vdots & \\ \frac{f_{\delta}(2N-1)}{(2N-1)!} & \dots & \frac{f_{\delta}(N+1)}{(N+1)!} & \frac{f_{\delta}(N)}{N!} \end{pmatrix} \tag{A.15}$$

$$d = \left(\frac{f_{\delta}(N+1)}{(N+1)!}, \frac{f_{\delta}(N+2)}{(N+2)!}, \dots, \frac{f_{\delta}(2N)}{(2N)!} \right)^T \tag{A.16}$$

とする。 $K(N)$ の行列式を求めると, 簡単な変形により

$$\det K(N) = (1+\delta)^{N-1} \cdot \frac{(N-1)!}{N!} \cdot f_{\delta}(1)$$

$\frac{f_{\delta}(N)}{(N+1)!}$	\dots	$\frac{f_{\delta}(3)}{4!}$	$\frac{f_{\delta}(2)}{3!}$
$\frac{f_{\delta}(N+1)}{(N+2)!}$	\dots	$\frac{f_{\delta}(4)}{5!}$	$\frac{f_{\delta}(3)}{4!}$
\vdots	\dots	\vdots	\vdots
$\frac{f_{\delta}(2N-2)}{(2N-1)!}$	\dots	$\frac{f_{\delta}(N+1)}{(N+2)!}$	$\frac{f_{\delta}(N)}{N!}$

$$= (1+\delta)^{N-1} (1+2\delta)^{N-2} \frac{(N-1)!}{N!} \cdot \frac{(N-2)!}{(N+1)!} \cdot f_\delta(1) \cdot f_\delta(2) \cdot$$

$f_\delta(N)$	$f_\delta(4)$	$f_\delta(3)$
$(N+2)!$	$6!$	$5!$
$f_\delta(N+4)$	$f_\delta(5)$	$f_\delta(4)$
$(N+3)!$	$7!$	$6!$
\vdots	\vdots	\vdots
$f_\delta(2N-3)$	$f_\delta(N+1)$	$f_\delta(N)$
$(2N-4)!$	$(N+3)!$	$(N+2)!$

$$= (1+\delta)^{N-1} (1+2\delta)^{N-2} \cdots (1+(N-1)\delta) \cdot$$

$$\frac{(N-1)!}{N!} \frac{(N-2)!}{(N+1)!} \cdots \frac{0!}{(2N-1)!} \cdot f_\delta(1) \cdot f_\delta(2) \cdots f_\delta(N)$$

であるから

$$\det K(N) = \prod_{n=1}^N \frac{(N-n)!}{(2N-n)!} (1+n\delta)^{N-n} f_\delta(n) \quad (A.17)$$

が得られる。

ここで $K(N) = (k_1, k_2, \dots, k_N)$ と表すことにし、次の行列

$$K_k(N) = (k_1, \dots, k_{k-1}, d, k_{k+1}, \dots, k_N) \quad (A.18)$$

の行列式を求めると、 $\det K(N)$ の場合と同様にして、

$$\det K_k(N) = (-1)^k \det (d, k_1, \dots, k_{k-1}, k_{k+1}, \dots, k_N)$$

$$= (-1)^k$$

$\frac{f_\delta(N+1)}{(N+1)!}$	\cdots	$\frac{f_\delta(N-k+2)}{(N-k+2)!}$	$\frac{f_\delta(N-k)}{(N-k)!}$	\cdots	1
$f_\delta(N+2)$	\vdots	\vdots	\vdots	\vdots	$\frac{f_\delta(2)}{2!}$
$(N+2)!$	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$f_\delta(2N)$	$\frac{f_\delta(2N-k+1)}{(2N-k+1)!}$	$\frac{f_\delta(2N-k-1)}{(2N-k-1)!}$	\cdots	$\frac{f_\delta(N)}{N!}$	
$(2N)!$	$(2N-k+1)!$	$(2N-k-1)!$	\cdots	$N!$	

k
 $N-k$

$$= (-1)^k \left[\frac{1}{(1+\delta)^{N-1}} \frac{1}{N-k} \frac{N!}{N!} f_\delta(1) \right] \left[\frac{1}{(1+2\delta)^{N-2}} \frac{1}{N-k-1} \frac{(N-1)!}{(N+1)!} f_\delta(2) \right] \cdots \left[\frac{1}{(1+(N-k)\delta)^k} \frac{1}{1} \frac{(k+1)!}{(2N-k-1)!} f_\delta(N-k) \right]$$

$\frac{f_\delta(N+1)}{(2N-k+1)!}$	$\frac{f_\delta(N-k+2)}{(2N-2k+2)!}$
\vdots	\vdots
$\frac{f_\delta(N+k)}{(2N)!}$	$\frac{f_\delta(N+1)}{(2N-k+1)!}$

$$= (-1)^k \frac{1}{(N-k)!} \frac{(N-1)! \cdots (k+1)!}{(N+1)! \cdots (2N-k-1)!}$$

$$(1+\delta)^{N-1} \cdots (1+(N-k)\delta)^k \cdot f_\delta(1) \cdots f_\delta(N-k)$$

$$\left[\frac{(k-1)!}{(2N-k+1)!} (1+(N-k+1)\delta)^{k-1} \cdot f_\delta(N-k+2) \right]$$

$$\cdots \left[\frac{1}{(2N-1)!} (1-(N-1)\delta) \cdot f_\delta(N) \right]$$

$$\frac{1}{(2N)!} f_\delta(N+1)$$

$$= (-1)^k \frac{1}{(N-k)!} \frac{(N-1)! \cdots (k+1)! \cdot (k-1)! \cdots 1}{(N+1)! \cdots (2N-k-1)! \cdot (2N-k+1)! \cdots (2N)!}$$

$$(1+\delta)^{N-1} (1+2\delta)^{N-2} \cdots (1-(N-1)\delta)$$

$$f_\delta(1) \cdots f_\delta(N-k) \cdot f_\delta(N-k+2) \cdots f_\delta(N+1)$$

と変るから、結局

$$\det K_k(N) = (-1)^k \frac{1}{k!} \frac{N! (2N-k)!}{(N-k)! (2N)!} \cdot \frac{f_\delta(N+1)}{f_\delta(N-k+1)} \cdot \det K(N) \quad (A.19)$$

ラプラスの公式と、式(A.17), (A.19) から

$$\begin{aligned} A_{N,k}^{(\sigma)} &= \frac{\det K_k(N)}{\det K(N)} \\ &= \frac{(-1)^k}{k!} \frac{N! (2N-k)!}{(N-k)! (2N)!} \frac{f_\sigma(N+1)}{f_\sigma(N-k+1)} \\ &= \frac{(-1)^k}{k!} \frac{{}_N C_k}{{}_2N C_k} (1-N\sigma) \cdots (1-(N-k+1)\sigma) \quad (\text{A.20}) \end{aligned}$$

が得られ、これが式(5.3)となる。

$B_{N,k}^{(\sigma)}$ は、式(A.10) から求めてもよいが、次の関係

$$S_\sigma^{-1}(-W) = (1+\sigma W)^{-\frac{1}{\sigma}} = \frac{1}{S_\sigma^{-1}(W)} \quad (\text{A.21})$$

を利用すれば簡単に得られる。もし $P_N^{(\sigma)}(W)$ が $S_\sigma^{-1}(W)$ のパデ近似であれば、式(A.21) から

$$\frac{1}{P_N^{(\sigma)}(-W)} = \frac{A(-\sigma, -W)}{B(-\sigma, -W)} \quad (\text{A.22})$$

もやはり $S_\sigma^{-1}(W)$ のパデ近似となる。パデ近似の一意性⁽²⁷⁾ から

$$P_N^{(\sigma)}(W) = \frac{1}{P_N^{(\sigma)}(-W)} \quad (\text{A.23})$$

が成立し、

$$B(\sigma, W) = A(-\sigma, -W) \quad (\text{A.24})$$

が得られる。式(A.20)と式(A.24)から

$$B_{N,k}^{(\sigma)} = \frac{1}{k!} \frac{{}_N C_k}{{}_2N C_k} (1+N\sigma) \cdots (1+(N-k+1)\sigma) \quad (\text{A.25})$$

となり、式(5.4)が得られた。