

15032

ピッチ変動に着目した擬似周期信号の  
最適フィルタに関する研究

西 一 樹

# 目次

<b>1</b>	<b>序論</b>	<b>1</b>
1.1	研究の目的	1
1.2	研究の背景と位置付け	2
1.2.1	時間周波数解析	2
1.2.2	聴覚フィルタ	3
1.2.3	聴覚情景分析	4
1.2.4	櫛形フィルタ	5
1.2.5	バイスペクトル分析	6
1.3	本研究のアプローチ	7
1.4	本論文の概要と構成	8
<b>2</b>	<b>ピッチ変動を伴う音響・音声信号のモデル</b>	<b>11</b>
2.1	はじめに	11
2.2	音声信号に見るピッチ変動	12
2.3	「振幅変動+ピッチ変動」信号モデル	13
2.3.1	物理的背景	13
2.3.2	「振幅変動+ピッチ変動」信号の定義	16
2.4	「振幅変動+ピッチ変動」信号推定の問題とは	17
2.5	本論文で扱う擬似周期信号の位置づけ	18
2.5.1	確率モデルによる分類	18
2.5.2	フィルタ設計法による分類	20
2.6	複素チャープ位相をもつ音響・音声信号モデル	20
2.7	本章のまとめ	21
<b>3</b>	<b>定常線形最適フィルタによる音響・音声信号の分離</b>	<b>23</b>
3.1	はじめに	23
3.2	定常型「振幅変動+ピッチ変動」信号の一般的性質	23

---

3.2.1	自己相関関数	23
3.2.2	スペクトル特性	25
3.3	「振幅変動+ピッチ変動」信号フィルタの基本構造	26
3.3.1	観測雑音モデル	26
3.3.2	非因果的 Wiener フィルタの導出	26
3.3.3	フィルタの周波数特性	27
3.3.4	平均二乗推定誤差	28
3.3.5	フィルタ導出の別解法	29
3.4	理想周期信号に対するフィルタ設計	30
3.5	振幅変動が定常過程をとりピッチ周波数が確率変数をとる場合	31
3.5.1	振幅変動のみを考慮した場合	32
3.5.2	ピッチ周波数の不確定性のみを考慮した場合	33
3.5.3	両者を同時に考慮した場合	34
3.6	振幅とピッチ周波数が定常白色 Gauss 雑音過程をとる場合	35
3.7	因果性を考慮した Wiener フィルタ解	36
3.8	非直線周波数スケールとの関係	38
3.9	聴覚の周波数特性を実現するフィルタ設計例	40
3.10	本章のまとめ	41
3.11	付録	42
3.11.1	Wiener 過程の諸性質	42
3.11.2	因果的 Wiener-Hopf 方程式の一般解	43
<b>4</b>	<b>ピッチ変動に対する定 Q 楕形フィルタの性質と効果</b>	<b>45</b>
4.1	はじめに	45
4.2	重み付き最小二乗法に基づく楕形フィルタの導出	46
4.3	定 BW 楕形フィルタの時間 / 周波数特性	47
4.4	定 Q 楕形フィルタの時間 / 周波数特性	48
4.4.1	定 Q 楕形関数とその Fourier 変換	48
4.4.2	定 Q 楕形関数の具体例	52
4.4.3	ピッチ誤差による波形 / スペクトル歪み	52
4.5	各種の楕形フィルタについて	56
4.6	本章のまとめ	57
4.7	付録	57
4.7.1	デルタ関数の変数置換	57

4.7.2	Poisson の和公式による導出 . . . . .	58
4.7.3	楕形関数の Fourier 変換 . . . . .	59
<b>5</b>	<b>多重擬似周期信号に対するピッチ推定と個別分離アルゴリズム</b>	<b>60</b>
5.1	はじめに . . . . .	60
5.2	Wavelet 空間とピッチ尤度関数 . . . . .	61
5.2.1	ストリーム分離問題とは . . . . .	61
5.2.2	時間 - 対数周波数分析 . . . . .	61
5.2.3	ピッチ尤度関数 . . . . .	62
5.3	Non-Parametric Kalman フィルタによるピッチ軌跡の最適推定 . . . . .	63
5.3.1	ピッチ確率密度関数 . . . . .	63
5.3.2	予測過程 . . . . .	64
5.3.3	更新過程 . . . . .	66
5.4	ストリームの分離再生 . . . . .	68
5.5	ピッチ推定誤差の影響と分析 Wavelet の Q 値の決定 . . . . .	68
5.6	実験 . . . . .	69
5.6.1	合成信号におけるストリーム分離 . . . . .	69
5.6.2	実音声におけるストリーム分離 . . . . .	70
5.7	本章のまとめ . . . . .	72
5.8	付録 . . . . .	73
5.8.1	Wavelet 分布の分解表現 . . . . .	73
5.8.2	短時間 Fourier 解析によるピッチ推定誤差の影響 . . . . .	74
<b>6</b>	<b>非定常擬似周期信号に対する線形最適フィルタ</b>	<b>75</b>
6.1	はじめに . . . . .	75
6.2	時変型「振幅変動+ピッチ変動」信号モデル . . . . .	76
6.2.1	Wiener 積分による振幅 / ピッチ変動の定式化 . . . . .	76
6.2.2	伊藤型確率微分方程式による「振幅変動+ピッチ変動」信号モデル . . . . .	77
6.3	「振幅変動+ピッチ変動」信号推定問題の動的システム表現 . . . . .	78
6.4	状態依存性雑音モデルに対する Kalman フィルタ解 . . . . .	81
6.4.1	Wiener-Hopf の積分方程式と直交原理 . . . . .	81
6.4.2	フィルタ方程式の導出 . . . . .	82
6.5	フィルタ方程式の解の性質 . . . . .	82
6.5.1	理想周期信号の場合 . . . . .	83

---

6.5.2	倍音間の干渉を無視したスカラー定常解 . . . . .	84
6.5.3	時変性を考慮したスカラー近似解 . . . . .	84
6.6	倍音間の干渉を考慮したフィルタ方程式の数値解析 . . . . .	85
6.6.1	Riccati 方程式の数値解 . . . . .	86
6.6.2	フィルタ方程式の数値解 . . . . .	86
6.7	本章のまとめ . . . . .	88
6.8	付録 . . . . .	89
6.8.1	Wiener 積分の諸性質 . . . . .	89
6.8.2	伊藤積分と伊藤の公式 . . . . .	91
6.8.3	Wiener-Hopf の積分方程式 . . . . .	94
6.8.4	フィルタ係数の関係式 . . . . .	94
6.8.5	Kalman ゲインと誤差共分散の関係式 . . . . .	96
6.8.6	Riccati 方程式 . . . . .	96
6.8.7	フィルタ方程式の解析解 . . . . .	99
<b>7</b>	<b>デジタル楕形フィルタの設計と評価</b> . . . . .	<b>101</b>
7.1	はじめに . . . . .	101
7.2	フィルタ方程式の離散化による IIR フィルタの導出 . . . . .	102
7.3	ノッチフィルタとの関係 . . . . .	105
7.4	フィルタ性能の評価 . . . . .	106
7.4.1	評価基準 . . . . .	106
7.4.2	ピッチ誤差に対する頑健性 . . . . .	107
7.4.3	ピッチ変動に対する頑健性 . . . . .	107
7.4.4	振幅 / ピッチ変動による波形 / スペクトル歪み . . . . .	107
7.4.5	実音声・楽音への適用 . . . . .	113
7.5	音声のスペクトル包絡による評価 . . . . .	117
7.5.1	LPC ケプストラムによる距離尺度 . . . . .	117
7.5.2	LPC メルケプストラムによる距離尺度 . . . . .	118
7.5.3	実音声に対する適用例 . . . . .	119
7.6	本章のまとめ . . . . .	121
7.7	付録 . . . . .	121
7.7.1	離散 Kalman フィルタによる別解法 . . . . .	121
<b>8</b>	<b>結論</b> . . . . .	<b>125</b>

---

謝辞

128

参考文献

138

## 図一覧

1.1	聴覚の周波数分析特性: 中心周波数と臨界帯域幅の関係 . . . . .	4
1.2	本論文の構成 . . . . .	9
2.1	発声音のピッチ周波数の頻度分布 . . . . .	12
2.2	弦の振動 . . . . .	14
2.3	「振幅変動+ピッチ変動」信号の記述にあたって導入する確率モデルの包含 関係 . . . . .	19
3.1	定 BW 櫛形フィルタの周波数特性 . . . . .	33
3.2	定 Q 櫛形フィルタの周波数特性 . . . . .	34
3.3	定 BW / 定 Q 複合櫛形フィルタの周波数特性 . . . . .	35
3.4	櫛形フィルタの極配置 . . . . .	38
3.5	音の周波数と高さの関係 . . . . .	39
3.6	メルスケール上でのフィルタバンクの構成 . . . . .	40
3.7	聴覚の周波数特性を実現するフィルタ設計例 . . . . .	41
4.1	定 Q 櫛形関数のフーリエ変換 . . . . .	51
4.2	ガウス核をもつ定 Q 櫛形関数の例 . . . . .	53
4.3	矩形核をもつ定 Q 櫛形関数の例 . . . . .	53
4.4	ピッチ誤差がないときの定 Q 櫛形フィルタの出力波形 / スペクトル歪み . . . . .	54
4.5	ピッチ誤差があるときの定 Q 櫛形フィルタの出力波形 / スペクトル歪み . . . . .	55
4.6	ピッチ誤差があるときの定 BW 櫛形フィルタの出力波形 / スペクトル歪み . . . . .	55
5.1	多重ピッチ推定と信号分離アルゴリズムの流れ . . . . .	67
5.2	合成信号に対する実験結果 . . . . .	71
5.3	雑音下の音声に対する実験結果 . . . . .	72
5.4	多重音声に対する実験結果 . . . . .	73
6.1	Riccati 方程式の解の比較 . . . . .	86

---

6.2	インパルス応答の比較 . . . . .	87
6.3	伝達特性の比較 (分散小) . . . . .	88
6.4	伝達特性の比較 (分散大) . . . . .	89
7.1	IIR デジタル楕形フィルタの極配置 . . . . .	104
7.2	ピッチ誤差に対する頑健性 . . . . .	109
7.3	ピッチ変動に対する頑健性 . . . . .	110
7.4	振幅変動が信号波形に及ぼす影響 . . . . .	111
7.5	振幅変動がスペクトルに及ぼす影響 . . . . .	111
7.6	ピッチ変動が信号波形に及ぼす影響 . . . . .	112
7.7	ピッチ変動がスペクトルに及ぼす影響 . . . . .	112
7.8	S/N 比を最大にするフィルタ係数と振幅/ピッチ変動量との関係 . . . . .	113
7.9	有声音「あえ」とトランペット音の混合音 . . . . .	114
7.10	抽出された有声音「あえ」のスペクトログラム . . . . .	115
7.11	抽出されたトランペット音のスペクトログラム . . . . .	115
7.12	抽出された有声音「あえ」のスペクトル断面 . . . . .	116
7.13	抽出されたトランペット音のスペクトル断面 . . . . .	116
7.14	スペクトル包絡間の距離尺度を用いたフィルタ性能比較 . . . . .	120



# 第 1 章

## 序論

### 1.1 研究の目的

我々を取り巻く自然界の秩序は、原子や分子の熱的振動から天体の周回運動に至るまで、様々な物理的振動・共鳴現象に基づく情報やエネルギーのやりとりを通して維持されている。このことは、危険回避、捕食あるいはコミュニケーション手段としての鳴き声や音声、生命維持のための鼓動など、生命活動の場においても同様である。振動現象はあらゆる情報発信の源であるから、その発信音を分析することによって振動源に関する様々な情報が得られる。実際、工学的な応用としても、機械振動の異常検知、電力波形による発電効率の監視、脳波や心音電図による生体検診など、振動体からの情報を積極的に利用した例は多い。

これら振動音を受信する側の観測系にとっては、騒音環境の中から目的とする振動音を拾い出すことが重要な使命になる。これを実現している代表例が聴覚システムである。特に生物にとって振動音を聞き分ける能力は、その善し悪しが死活問題につながる重要な意味をもつ。工学的にもこの機能を実現することは、上記の応用例を含むセンサシステムのインテリジェント化にとって不可欠な課題である。

振動音が、時間波形としては周期性をもち、スペクトルとしては離散構造をとることは、波動方程式に対する Fourier 級数解析の教えるところである。実はこの信号波形の特異性が、振動音を聞き分ける機能を実現する上での重要な鍵になる。先の聴覚システムもこの性質を巧みに利用していて、我々が騒音環境において調波音(周期音)を聞き分けられるのは、スペクトルの離散パターンがそのまま聴覚内の基底膜上に展開され、他の音のスペクトルとの明確な違いをパターンとして脳が識別するからであると考えられる [28]。これを信号処理の立場から解釈すれば、特定の調波音を抽出する機能とは、ピッチ周波数(基本周波数)間隔に細線スリットが並んだ櫛歯状の伝達特性をもつフィルタ(櫛形フィルタ)によって、その櫛歯間隔にちょうど合致した離散スペクトル成分が抜き出されることに基づいていると説明できる [29]。

「理想周期信号」<sup>†</sup>を仮定できる場合には、それを抽出するための信号処理の手続きは、上記の楕円フィルタ演算もしくは Fourier 級数解析の理論に完結する。しかしながら、この仮定が厳密に成り立つことは現実にはありえず、実際の振動音は振幅と周期が絶えず変化する「擬似周期信号」になる。この場合に上記の楕円フィルタを用いると、楕円の位置ずれが大きく影響し、騒音を抑圧する以上に元の信号波形を著しく歪ませてしまう [30]。そのため実用的には、周期性の仮定が近似的にしか成り立たない場合においても、できるだけ波形歪みが生じることなく、騒音環境下で目的信号の抽出が可能な信号処理手法あるいはフィルタ設計法の確立が不可欠になる。

そこで本論文では、楽音や音声などの擬似周期信号を聞き分ける機能が工学的アプローチによってどのように実現できるか、またその信号処理のメカニズムが聴覚のそれとどのような関係にあるのかを基本的な問題意識として、時間周波数特性やデジタルフィルタ設計、ピッチ推定などの問題と併せて、線形最適フィルタを中心とした純数理工学的な立場で議論する。

## 1.2 研究の背景と位置付け

ここでは、関連性のある既存研究を取り上げて、本研究における立場を説明する。

### 1.2.1 時間周波数解析

時変信号の解析には、通常、短時間 Fourier 変換や Wavelet 変換などの時間周波数解析 [43] が用いられる。それらに共通する考え方は、窓関数を乗じることによって適当な短い時間幅で信号を区切れば、その間は振幅や周波数などのパラメータはほぼ一定とみなせるので、窓関数により決まる時間および周波数の分解能で各時刻毎に Fourier 解析が行えるというものである。短時間 Fourier 変換と Wavelet 変換の違いは、前者が周波数に依存しない一様な時間 - 周波数分解能をとるのに対し、後者は周波数に比例して時間の分解能が上がり、逆に反比例して周波数の分解能が下がる (定 Q 構造をとる) ように、窓関数の作用が周波数により変則的に変わる点にある。

これら時間周波数空間においては、擬似周期信号を構成する各倍音成分はエネルギーの局在として表せ、その起伏の軌跡 (Ridge) から信号の時変性を視覚的に捉えることが可能になる。このことを逆に利用して、任意の信号成分に対する瞬時周波数や瞬時振幅を Ridge 上の複素振幅値から求める方法 [44] も提案されている。

---

<sup>†</sup>「理想周期信号」とは、無限の過去から未来に至るまで振幅や周期が常に一定である信号をいう。本来は単に「周期信号」とした場合もこれを指すが、後の「擬似周期信号」と区別するために、本論文では敢えてこのように呼ぶものとする。

いずれにしても、信号の挙動を二次元分布として描写することが時間周波数解析の中心目的であり、窓関数の決め方などの議論はほとんどなされていない。これに対し本論文では、特定の信号モデル下で信号抽出に最適な窓関数が決定できることを示す(3章)。その結論から、振幅変動のみを考慮した場合は短時間 Fourier 変換型が、ピッチ変動のみを考慮した場合は Wavelet 変換型が有利になることが導かれる。さらに両者を同時に考慮した場合は、低周波側では短時間 Fourier 変換型、高周波側では Wavelet 変換型をとることが最適になる。擬似周期信号解析にとっては混成型をとることが理想的であるという指摘は本論文が初めてである。

### 1.2.2 聴覚フィルタ

聴覚末梢系では音の周波数分析が行われていることは、生理学的にも心理学的にも明白な事実になっている [45]。Békésy[46] は、外耳で受けた音があぶみ骨を介して蝸牛内の基底膜を振動させ、それが進行波として伝搬していく際に、徐々に低くなる基底膜の共振周波数にちょうど一致した点で振動振幅が最大になるという仮説を立て、それを実験的にも実証した。この結果は、入力音の周波数によって基底膜上の共振位置が変わることを意味し、聴覚が周波数分析を行っていることの生理学的証拠になる。

一方 Fletcher[47] は、正弦波の純音に対して、同じ中心周波数をもつ狭帯域雑音をマスクとして加え、純音の振幅を小さくしていったとき聞こえなくなる限界点の音圧 (マスキング量) を測定する心理学実験を行っている。その結果によれば、狭帯域雑音の占有帯域幅を広げるのに比例してマスキング量は増加していくが、ある帯域幅以上になると、いくら広げてもマスキング量は変化しないことが示された。このことから、聴覚が原理的には中心周波数の異なる多数の帯域通過フィルタで構成され、入力信号の周波数に最も近い中心周波数をもつ帯域通過フィルタだけが信号分析に関与し、かつその占有帯域幅内に収まる雑音成分だけがマスクとして働くと考えられる [48]。マスクとして寄与する最大の雑音帯域幅は「臨界帯域幅」と呼ばれ、Zwicker[49] によって中心周波数と臨界帯域幅の関係が精密に測定された (図 1.1 参照)。

他にも、中心周波数と臨界帯域幅の関係を定式化したり [50]、臨界帯域幅に代わる別の帯域幅測定法として、Notched-noise 法による等価矩形幅 (Equivalent Rectangular Bandwidth: ERB)[51] やその変形版である ERB-rate[52] などが提案されている。

また、キャリア成分を Gamma 関数で振幅変調させたものをインパルス特性にもつ Gammatone フィルタ [53] やそのキャリア成分をチャープ信号に置き換えた Gammachirp フィルタ [54] が、聴覚フィルタの良い近似として提案されている。Gammatone フィルタをフィルタバンクとして構成し聴覚末梢系の処理を模擬したり [55]、基底関数として Wavelet

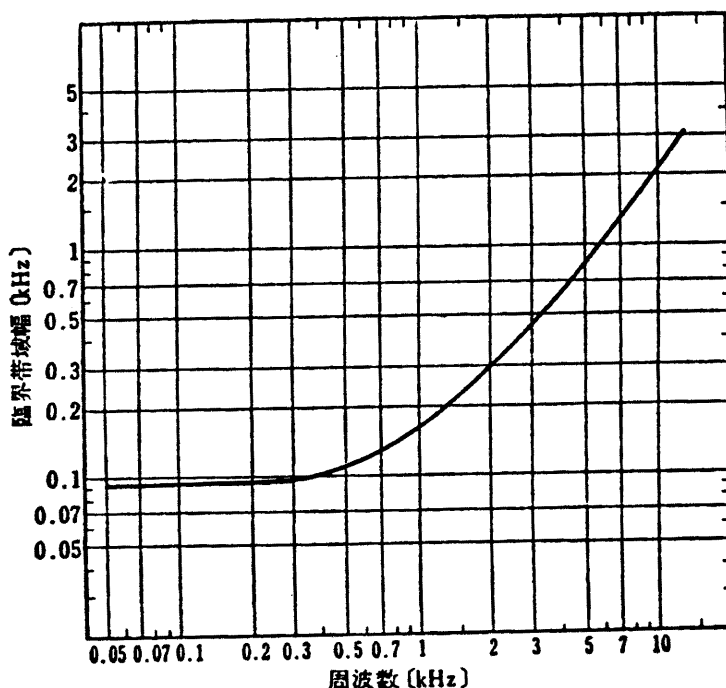


図 1.1 聴覚の周波数分析特性: 中心周波数と臨界帯域幅の関係 (境久雄 編著「聴覚と音響心理」 [45] より引用)

変換を実現し雑音除去に利用した研究 [56] などがある。特に Gammachirp フィルタは、時間 - スケール表現における最小不確定性の意味で最適な聴覚フィルタになることが示されている。

これら聴覚フィルタの研究がいずれも、聴覚特性をいかに精度良く模擬できるかという点に重点をおいているのに対し、本論文は、聴覚特性とは無関係にあくまでも純信号处理的立場で最適フィルタ設計を行うことに終始する。しかしながらその結論としては、図 1.1 に見るような、「低域側では等帯域幅特性であるのに対し、高域側では定  $Q$  特性をもつ」ことがフィルタの伝達特性から示せることは興味深い (第 3 章)。ここでは、擬似周期信号に対する線形最小二乗推定の意味で最適フィルタが求められており、Gammachirp フィルタ [54] での最適性とは立場が異なる点に注意する。

### 1.2.3 聴覚情景分析

複数の音声や楽音、騒音などが含まれる混合音中から、目的とする音を適切に聞き分けることができる人間の優れた聴覚機能を、工学的に実現しようとする研究の枠組みは「聴覚情景分析 (Auditory Scene Analysis)」 [28] と呼ばれる。上記の聴覚フィルタの議論を踏まえ、Wavelet 空間あるいは Gammatone フィルタバンク上での処理を中心とした、聴覚情

景分析に関する数多くのアルゴリズムが提案されている。

柏野ら [57] は、音の立上り立ち下がりにおける周波数間での同時性、周波数変化の同期性や調波構造の保持など、音脈形成の過程を音響心理学的知見に基づいてモデル化した。特に、音脈が調波構造をもつことを利用して、音源毎に個別に音脈追跡を行うマルチエージェントシステムが中谷ら [58, 59] によって提案されている。また Brown と Cooke [60] は、聴覚の生理学的知見に立脚し、基底膜の応答特性を模擬した Gammatone フィルタを用いて入力信号を分析し、その時変スペクトル上でのエネルギー局在化の特徴地図を用いて音脈のクラスタリングを行う音源分離システムを構築している。いずれの方法も基本的には、時変スペクトル上のピークを音脈形成の規則にしたがってグルーピングするという手続き [61] を基礎にしていて、信号の周期性すなわちスペクトル構造の調波性が最も本質的かつ不可欠な音脈分凝規則として働いている。

一方、安部ら [62, 63] は、音脈分凝要件として必ずしも調波構造に頼らない、新しい聴覚情景分析のアルゴリズムを提案している。そこでは、調波性条件を用いる代わりに、サブバンド信号に対して瞬時周波数・周波数変化率・振幅変化率などの特徴量を新たに導入し、それに基づく音脈のグルーピング戦略を Wavelet 空間と整合性よく実現している。

これら聴覚情景分析の研究は、聴覚中枢の比較的高次な機能の実現を狙ったものであるが、それが有効に働くためには、フロントエンドの時間周波数解析において、常に目的信号を最適抽出できるようなフィルタバンク設計がなされているべきものと考えられる。そこで本論文では、この最適フィルタ設計の問題と音脈分凝の問題を明確に区別するために、前者は、特定の調波音だけを目的信号としそれ以外の成分はまとめて白色 Gauss 雑音と仮定した上での線形最適フィルタ設計の問題と捉え (3,6章)、後者は、それに必要なピッチ推定値を複数のピッチ候補から並列かつ選択的に求めるための聴覚情景分析の問題として捉える (5章)。全体としては、最適な時間周波数分析機構の上に立脚した聴覚情景分析システムの実現を目指すものである。

#### 1.2.4 櫛形フィルタ

櫛形フィルタは、インパルス応答が櫛形関数 (デルタ関数が等間隔に並んだ周期関数) で表されるとともに、その Fourier 変換である伝達特性もまた櫛形関数になるという性質がある [64]。それゆえ、一定周期間隔の標本点列を同期加算する平均化フィルタであると同時に、ピッチ周波数の整数倍の倍音成分を抽出するサンプリングフィルタでもある。

目的音が完全な周期信号であり、かつそのピッチ周期が正確にわかっているならば、それを抽出するための最適フィルタは櫛形フィルタになる [65]。しかしながら、振幅やピッチ周期が変化する擬似周期信号の場合には、もはやピッチ周期を誤差なく推定することは不可能であ

り、たとえそれが僅かであっても、真の周期間隔と実際にとられる標本点列との間でずれが時間の経過にともなって拡大するため、フィルタ出力に大きな歪みが生じる。特にピッチ変動は、高調波に対して倍音次数に比例した周波数揺らぎを与え、高次倍音になるほど多くの波形歪みを生じさせる。

一般には、倍音周波数毎に等帯域幅のバンドパスフィルタを配置した櫛形状の伝達特性をもつフィルタ (定 BW(Band-Width) 櫛形フィルタ) が用いられており、その構成法として、短時間 Fourier 変換によるもの [66]、バンドパスフィルタが FIR 型のもの [67] や IIR 型のもの [99] などが提案されている。特に IIR 型は、少ない次数で急峻なフィルタ特性が得られるため最も広く用いられ、ピッチ同期法と併せて様々なタイプのものがある [68, 115]。時間軸上でのフィルタ演算としては、遠方での平均化の寄与が小さくなるように、一定ピッチ周期間隔の標本点列を重み付き加算することに対応し、これに基づいた櫛形フィルタの形式もいくつか提案されている [90, 91]。いずれも定 BW 型のものを前提にしており、その根拠についてこれまで議論されたことはなかった。

これに対し本論文では、擬似周期信号推定の問題に対する線形最適フィルタの解として櫛形フィルタが導けることを主張する。特に、振幅変動のみを考慮した場合は定 BW 櫛形フィルタ、ピッチ変動のみを考慮した場合は定 Q 櫛形フィルタがそれぞれ最適になり、その変動特性から各バンドパスフィルタの形状が一意に決定できることが示される (3章)。定 Q 櫛形フィルタとは、倍音次数に比例してバンドパスフィルタの帯域幅が広がる (つまり一定 Q 値をとる) 変則的な櫛形状の伝達特性をもつフィルタのことであり、そのインパルス応答までもが定 Q 櫛形分布になるという注目すべき性質がある。実際、この構造こそがピッチ誤差に対して有利に働くことが時間 / 周波数特性の両面から検証される (4章)。また、最適フィルタ解に対する離散化を通して、IIR 型のデジタルフィルタとして櫛形フィルタが実現できることを示す (7章)。

### 1.2.5 バイスpekトル分析

Wiener フィルタや Kalman フィルタなどの線形最適フィルタは、自己相関や相互相関あるいはパワースpekトルなどの 2 次統計量に基づいて設計される。そのため、信号が Gauss 過程をとる場合には最小二乗推定の意味で真の最適フィルタとなるが、非 Gauss 過程の信号に対してはあくまでも線形フィルタの中での最適性しか保証されないという本質的な限界がある。実際、擬似周期信号は一般に非 Gauss 過程をとるため、線形フィルタの枠組みでは真の最適推定値を求めることはできない。

この限界を越える非線形解析の一つの方法論として、3 次モーメントによる自己相関あるいは相互相関を用いたバイスペkトル分析がある [92]。特に館 [93] は、Gauss 性雑音に

埋もれた周期信号に対してバイスペクトルを求めることにより、現れた高調波のスペクトル値から各倍音の Fourier 係数を求め目的信号を再生する方法を提案している。その特徴は、3 次以上の統計量をもたない Gauss 性雑音による影響がバイスペクトルの計算過程で取り除かれること、および一定の位相関係にある高調波間だけにバイスペクトルが 0 でない値をとるため、必然的に信号の周期性に強く反応する性質をもつことにある。その結果として、周期信号に対して高い分離能力をもつことが確かめられている。このようなバイスペクトル分析は、信号に完全な周期性が仮定できる場合には相関演算の仕組みが明確になり一定の威力を発揮するものの、振幅変動やピッチ変動を考慮した場合には現象が一挙に複雑になり、擬似周期信号に対しての適切な分析手段となりえるかどうか不明な点が多い。

そこで本論文では、信号のもつ時変構造を比較的見通しよく記述できる通常の自己相関あるいは相互相関演算を基にした線形フィルタの方法論を前提に議論する。そのことで「最適性の限界」については妥協せざるをえないが、そのかわり擬似周期性の本質的な役割をできるだけ簡潔に捉えることを優先したいと考える。

### 1.3 本研究のアプローチ

擬似周期信号の推定もしくはフィルタ設計の問題を議論するにあたっての主な検討課題とそれに対する本研究のアプローチを以下に要約する。

まず第一に、「擬似周期信号」とはどのような信号であるかについて、信号処理の問題として扱う上で具体的に定式化されたものでなくてはならないという点である。そこには本質的に、マクロ的な挙動とミクロ的な挙動の二つの要因が同時に存在すると考えられる。両者とも振幅と周期の時変性を記述したものになるが、特にマクロ的な挙動は、いわば振幅と周期の非定常的な変化を記述するものであり、そのかわり緩やかな変化を仮定することにより、その軌跡は十分事前に知り得るものとする。逆にミクロ的な挙動とは、観測に供しない細かな変動であるがゆえ不規則性を有し、そのかわりミクロの限られた視野の中ではその挙動は定常的であるとする。この両者を同時に考慮した信号モデルを立てることがこの目的である。そのために本論文では、信号を Fourier 級数展開表現したときの各倍音成分において、複素振幅およびピッチ周波数などのパラメータを確率過程として扱い、その平均値と分散がマクロ的な挙動、それに従う不規則変動がミクロ的な挙動を表すものとする。議論の簡単化のため、これらの変動は互いに独立な白色 Gauss 雑音過程をとるものと仮定する。

第二は、上記の「振幅変動+ピッチ変動」信号のモデルに対して、信号推定の問題とその解を得る上での評価基準をどのように設定するかについてである。ここでは、不用意に問題を複雑化することは避け、擬似周期信号に関する基本的な性質を見出すことだけに専念するために、ごく標準的な問題設定および評価基準を前提に議論する。すなわち、観測信号には

加法性の白色 Gauss 雑音を仮定し、最小二乗誤差規範による評価基準の下で擬似周期信号に対する推定値が線形フィルタにより求められるものとする。

第三は、この最適化基準を満たす解を求める具体的な計算手続きについてである。これには、定常な信号モデルに対しては Wiener フィルタが、非定常なものに対しては Kalman フィルタの方法論が用いられる。定常信号モデルとは、振幅変動およびピッチ変動を記述する確率過程に定常性が成り立つもの、すなわちマクロ的な挙動が時不変であるものを指し、先の例では平均値と分散が時間に依存せず一定値をとる場合や、変動をとまなわず単に確率変数をとる場合に対応する。また非定常信号モデルは、一般には定常信号以外のすべてを指すが、ここでは変動の平均値と分散が既知の緩やかな変化をとまなうもの、すなわちマクロ的挙動を考慮したものに限定する。

第四は、上記のフィルタを構成する際に必要なピッチ周波数をどのように求めるかについてである。特に実際の音環境では、音声や楽音などの複数のピッチ候補から目的とするピッチ周波数を選び出し追跡できる仕組みのあることが重要である。ただし、ピッチ推定時の誤差はフィルタ設計の段階で考慮することができるので、ここではその精度はあまり要求しないものとする。そこで、ピッチ変動に対して最適性が保障された Wavelet 空間において、複数のピッチ候補を多峰的な尤度関数として表し、その分布形状の遷移としてピッチ候補を同時追跡しつつ、分布の尖鋭化により目的ピッチに絞り込むことを考える。これには、動的な最適推定の手続きをノンパラメトリックに実現できる一般化された Kalman フィルタの方法論 [69] が有用である。

## 1.4 本論文の概要と構成

本論文は、振幅変動やピッチ変動をとまなった擬似周期信号を雑音環境下で抽出するためのフィルタ設計法の確立を主たる目的とし、「振幅変動+ピッチ変動」信号のモデル化に始まり、線形最適フィルタの方法論に基づくフィルタ解の導出、その時間 / 周波数特性の間に成り立つ関係の解明、デジタルフィルタとしての構成法および混合音に対するピッチ推定法を提案するものである。本論文の構成を図 1.2 に示す。

第1章は序論である。

第2章では、一般の楽音や音声のピッチ変動を伴う擬似周期性をもつことを説明し、時変パラメータをもつ波動方程式の解から擬似周期信号が導かれることを示した上で、それに確率過程を導入することによって一般的な形で擬似周期信号を定義する。振幅変動やピッチ変動にいくつかの代表的な確率モデルを仮定し、本論文で扱う擬似周期信号の位置づけを整理する。特に、音響・音声信号モデルとして複素チャープ位相をもった確定信号を例にあげ、「振幅変動+ピッチ変動」信号のもつ基本的な性質について概観する。



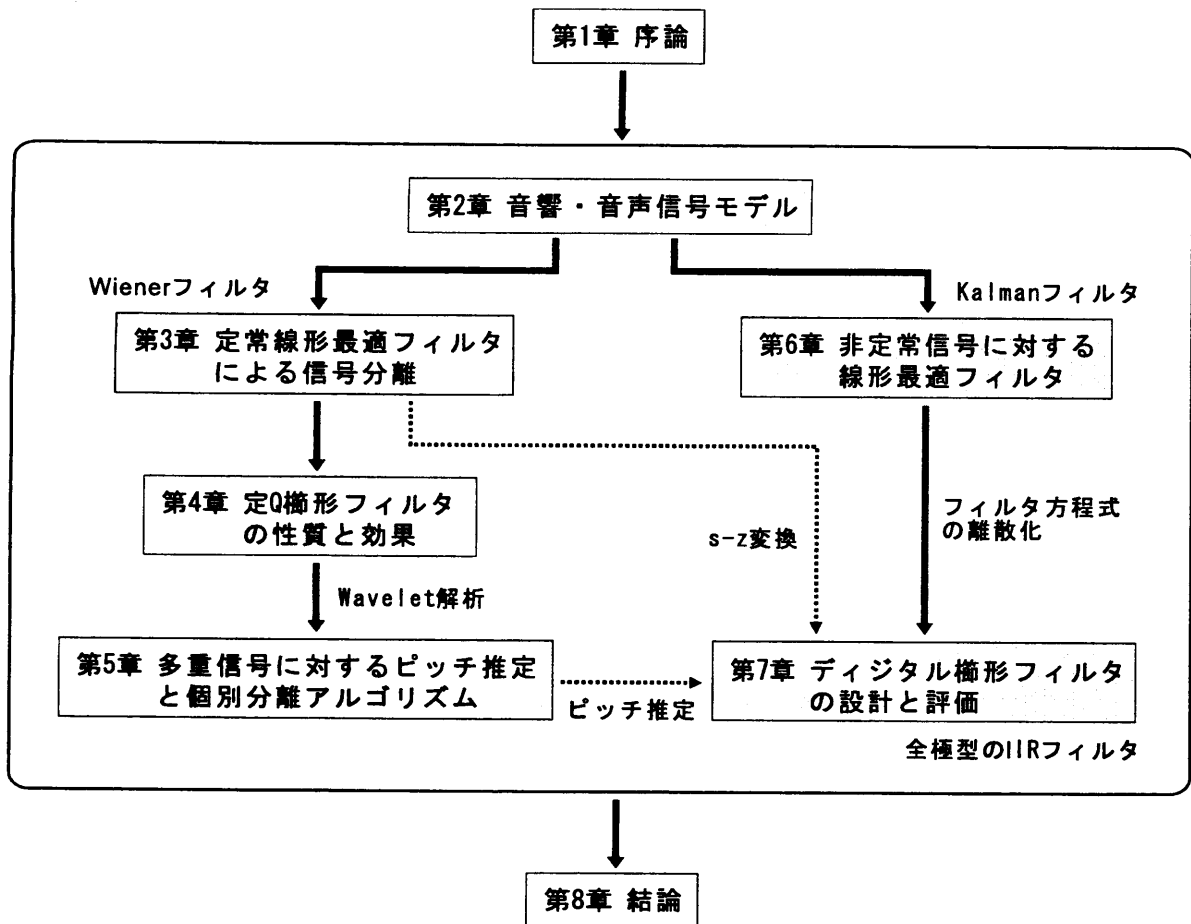


図 1.2 本論文の構成

第 3 章では、定常性を仮定した「振幅変動+ピッチ変動」信号に対するスペクトル特性、およびこのときの線形最適フィルタすなわち Wiener フィルタの伝達特性に見られる一般的性質について考察する。具体例として、振幅とピッチを一定とした理想周期信号の場合、振幅に任意の確率過程を導入した場合、ピッチに任意の確率変数を導入した場合、および振幅とピッチに白色 Gauss 雑音過程を導入した場合などのそれぞれに対するフィルタ伝達特性の詳細を論じる。特にフィルタ係数をうまく選ぶと、聴覚末梢系のもつ非線形な周波数特性によく適合することが示される。

第 4 章では、最適フィルタとして求められた楕形フィルタのどのような構造が振幅変動やピッチ変動に対して有利に働くのかについて、時間 / 周波数特性の両面から考察する。特に、定 Q 楕形フィルタの時間 / 周波数特性はともに定 Q 楕形分布になるという注目すべき性質があることを導き、この構造がピッチ誤差に有利に働くことを実例を通して示す。

第 5 章では、混合音中から各擬似周期信号成分を個別に分離抽出する方法として、Non-Parametric Kalman フィルタによる最適ピッチ追跡および Wavelet 解析を用いた擬似周期

信号再生のアルゴリズムを提案する。アルゴリズムの有用性は、多重音声に対する個別分離のシミュレーションおよび実音声実験により検証する。

第 6 章では、擬似周期信号として、特に振幅 / ピッチ変動に白色 Gauss 雑音過程を仮定し、その平均値と分散に時変性を許容した非定常信号モデルに対する線形最適フィルタとして Kalman フィルタを設計する。複素位相項にランダム性をもつ信号モデルの記述のために伊藤型の確率微分方程式による表現形式を導き、推定値を与える連続 Kalman フィルタの方程式を導出する。特にその伝達特性には、倍音間の干渉を排除するためのノッチ特性を有することが数値解により示される。

第 7 章では、櫛形フィルタをデジタルフィルタとして実現するための設計法について議論する。特に、第 6 章で導いた Kalman フィルタ方程式の離散化を通して IIR フィルタが導けることを示す。振幅 / ピッチ変動に対する定 BW および定 Q 櫛形フィルタの頑健性や音声のスペクトル包絡の再現性に関する評価実験も併せて紹介する。

第 8 章は結論であり、以上で得られた成果を総括する。

なお、第 2,3 章の一部は文献 [1]、第 4 章は文献 [2-11]、第 5 章は文献 [12-20]、第 6,7 章の一部は文献 [21-27] においてそれぞれ発表した。

## 第 2 章

# ピッチ変動を伴う音響・音声信号のモデル

### 2.1 はじめに

ギターのような弦楽器、トランペットなどの管楽器音あるいは太鼓などの打楽器は、それぞれ弦、気柱、膜という媒質の違いはあっても、振動や共鳴現象を利用している点で物理的には共通の発音機構をもっているといえる。このことは、声帯での膜の振動および声道管での共鳴を利用した音声についても、また機械振動や生体信号などについても同様である。一般に振動現象は、励振初期に見られる非定常な過渡的期間と、定在波が立つ安定な共鳴状態にある定常期間の二つに分けられるが、本論文では専ら、情報あるいはエネルギー伝達に中心的な役割をもつ後者に限定して議論する。

共鳴状態での振動は、振動体の全長に対してちょうど整数個の正弦波が入るような振動モードしか残らない安定状態にあり、このとき放射される信号波形は、それら整数倍の周波数をもつ正弦波の重ね合わせによって周期信号になる。逆に、任意の周期信号が整数倍の周波数関係にある正弦波の重ね合わせで表せることは、Fourier 級数展開理論の教えるところである。

しかしながら実際の振動現象では、振動体の変形、粘性抵抗やその他外乱により、発信音は一般に振幅、位相および周期 (ピッチ) の変動をともなった擬似周期信号になり、各倍音の離散スペクトルが占有する帯域幅に広がりが生じる。

まず本章では、音声の主としてピッチ変動を伴った擬似周期性をもつことを説明し、時変係数を考慮した波動方程式から擬似周期信号を一般的な形で定式化した上で、各種確率モデルの導入を前提とした音響・音声信号抽出問題の定義を与える。

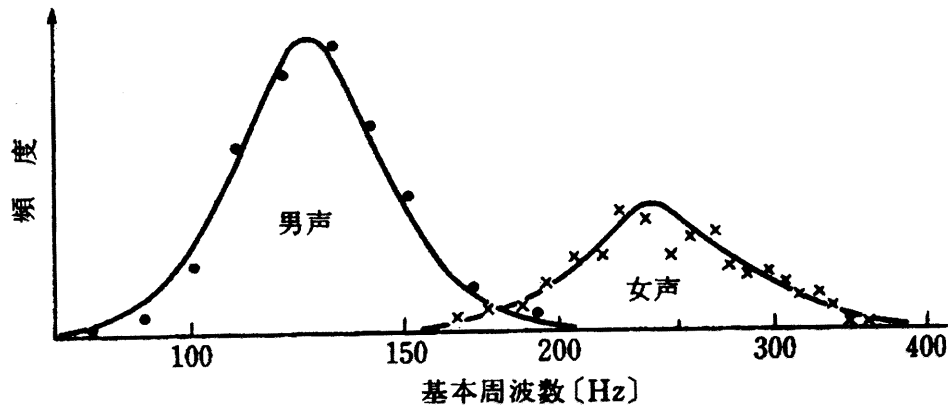


図 2.1 発声音のピッチ周波数の頻度分布 (古井 著「デジタル音声処理」 [36] より引用)

## 2.2 音声信号に見るピッチ変動

我々の母音音声には、振幅揺らぎやピッチ揺らぎが伴うことは古くから知られている [31]。揺らぎのある音声とない音声を人工的に合成し聞き比べてみると、前者の方が自然に聞こえるという実験結果も示されている [32]、音声合成の研究では、品質の向上のためにピッチ周期系列に適当な揺らぎを加えることの重要性が指摘されており [33]、白色雑音の揺らぎを与えるだけで自然性が向上することが報告されている [34]。特に、病的な音声では正常な場合より揺らぎが多いことが指摘されており、このことを利用して病的音声を定量的にスクリーニングする方法 [35] が提案されている。

ピッチ揺らぎは、音楽での旋律や声の抑揚などに関与する比較的緩やかな変化成分と、振動体の熱的擾乱や声帯の不規則振動による微細な変動成分とに分けられ、それぞれの統計的性質が詳しく調べられている。会話音声のピッチ周波数の頻度分布は、ほぼ正規分布をとり、女性は男性に比べて、平均値、標準偏差がともに2倍程度になることがわかっている (図 2.1 参照)、特にピッチ周波数の時間的推移を調べると、約 20% の上昇率に対して下降率が約 50% を占め、ピッチ周波数の下降傾向が強く現れ、その変化の時間パターンはほぼ「へ」の字で近似できる [36]。小室ら [39] は、「緩やかなピッチ変化」と「微細なピッチ変動」のどちらが知覚的に重要であるかを調べるために、ピッチ周期系列を  $n$  次多項式で最小二乗近似し、次数  $n$  を徐々に大きくするに従って原音声との類似度がどのように上昇するかを聴取実験により調べた。その結果、低い次数で表せるピッチ周期系列の概形を近似するだけで類似性が高く、「緩やかなピッチ変化」が音声の性質を保存するのに重要な働きをもつが、「微細なピッチ変動」の影響も無視できないことを示した。最近、北風ら [37] は、両者のピッチ変動を詳細に分類した上で分析合成した音声による知覚実験を通して、

「微細なピッチ変動」の不規則性が音声の自然性を保つのに大きく関与していることを明らかにした。

これまでに調べられた振幅 / ピッチ変動に関する重要な性質をまとめれば、

- (1) 揺らぎは緩やかな遅い変動 (トレンド) 成分と細かな速い変動 (ジッタあるいはシマー) 成分に分けられる [38]、
- (2) 振幅変動とピッチ変動は互いに無相関でほぼ Gauss 分布をとる [39]、
- (3) 微細変動成分は広義定常であり、揺らぎの周波数特性はほぼ  $1/f$  特性をもつ [40]、

などの性質があることがわかっている。他にも、揺らぎを考慮した音声合成システム [41] など、音声におけるピッチ変動の重要性を意識した研究は多い。

また、音声だけでなく楽音においても、信号のもつ擬似周期性は音質を決める重要な働きをもつことが示されている [42]。

## 2.3 「振幅変動 + ピッチ変動」信号モデル

以上の音声信号に見られるピッチ変動の性質を踏まえ、ここでは楽音や機械振動音を含めたより一般的な擬似周期信号のモデル化を考える。そのための物理的根拠として、弦の振動 [70] を例にあげる。

### 2.3.1 物理的背景

#### (a) 理想振動解

弦の長さを  $L$ 、張力を  $T$ 、単位長さあたりの質量すなわち線密度を  $\sigma$  とし、時刻  $t$ 、位置  $x$  での弦の微小要素  $\Delta x$  における  $x$  軸垂直方向の変位を  $u(x, t)$  とする。 $x$  軸方向への弦の運動が無視でき、振動振幅が微小であるとすると、曲線を描く弦の接線方向および  $x$  軸方向に働く張力は等しく、あらゆる位置で均一な値  $T$  をとるとしてよい (図 2.2 参照)。このとき、微小要素  $\Delta x$  に働く  $u$  軸方向の力成分は、弦の接線の傾きが  $\frac{\partial u}{\partial x}$  であることを考慮すると、位置  $x$  で  $-T(\frac{\partial u}{\partial x})_{x=x}$ 、位置  $x + \Delta x$  で  $T(\frac{\partial u}{\partial x})_{x=x+\Delta x}$  と表せるので、全体として

$$T\left\{\left(\frac{\partial u}{\partial x}\right)_{x=x+\Delta x} - \left(\frac{\partial u}{\partial x}\right)_{x=x}\right\} = T\left(\frac{\partial^2 u}{\partial x^2}\right)_{x=x}\Delta x \quad (2.1)$$

となる。これは、質量  $\sigma\Delta x$  の微小要素  $\Delta x$  に働く加速度  $\frac{\partial^2 u}{\partial t^2}$  による力とのつり合いから、

$$\sigma\Delta x \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2} \Delta x \quad (2.2)$$

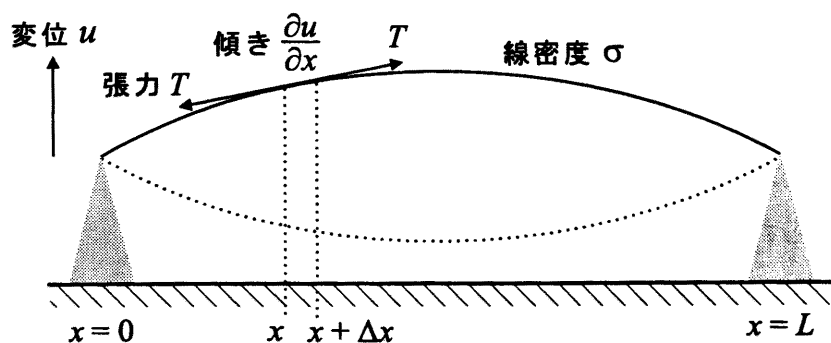


図 2.2 弦の振動

が成り立ち、結果として波動方程式

$$\frac{\partial^2 u}{\partial t^2} = \frac{T}{\sigma} \frac{\partial^2 u}{\partial x^2} \tag{2.3}$$

が導かれる。

一般に波動方程式の解は、変数分離の形

$$u(x, t) = \Phi(x)\Psi(t) \tag{2.4}$$

で表せることがわかっているのので、これを (2.3) 式に代入すると、

$$\frac{1}{\Psi(t)} \frac{d^2 \Psi(t)}{dt^2} = \frac{T}{\sigma \Phi(x)} \frac{d^2 \Phi(x)}{dx^2} \tag{2.5}$$

と整理できる。変数  $t$  だけの関数左辺と  $x$  だけの関数右辺が恒等的に等しくなるためには、定数にならなければならないから、それを  $-\lambda^2$  とおけば、

$$\frac{d^2 \Psi}{dt^2} + \lambda^2 \Psi = 0 \tag{2.6}$$

$$\frac{d^2 \Phi}{dx^2} + \lambda^2 \frac{\sigma}{T} \Phi = 0 \tag{2.7}$$

を得る。(2.7) 式の一般解

$$\Phi(x) = B \cos\left(\lambda \sqrt{\frac{\sigma}{T}} x + \varphi\right) \tag{2.8}$$

において ( $B$  は任意の実数係数)、弦が両端  $x = 0, x = L$  で固定されていることによる境界条件  $\Phi(0) = \Phi(L) = 0$  を満たすには、 $\varphi = \frac{\pi}{2}$  かつ

$$\lambda = \frac{n\pi}{L} \sqrt{\frac{T}{\sigma}} \quad (n \text{ は任意の整数}) \tag{2.9}$$

とならなければならない。つまり固定端条件により整数倍の周波数をもつ振動モードしか存在しえなくなる。これより (2.6) 式の一般解は、各振動モードの重ね合わせとして、

$$\Psi(t) = \sum_{n=1}^{\infty} A_n \cos(n\omega_0 t + \phi_n) \quad (2.10)$$

と表せる ( $A_n, \phi_n$  は任意の実数定係数)。ここで基本 (ピッチ) 周波数  $\omega_0$  は、

$$\omega_0 = \frac{\pi}{L} \sqrt{\frac{T}{\sigma}} \quad (2.11)$$

で与えられる。さらに、 $c_n = \frac{1}{2} A_n e^{j\phi_n}$ ,  $c_{-n} = c_n^*$  (上付添字 \* は複素共役の意) とおけば、(2.10) 式は複素形式で

$$\Psi(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad (2.12)$$

と表すこともできる。但し、 $c_0 = 0$  として直流成分も形式的に含めてある。これは Fourier 級数展開の式そのものであるから、 $\Psi(t)$  は

$$\Psi(t + T_0) = \Psi(t) \quad \text{where } T_0 = \frac{2\pi}{\omega_0} \quad (2.13)$$

が成り立つ任意の周期振動を表していることがわかる。

### (b) パラメータ変化を考慮した振動解

前述の周期振動解は、弦の長さの変動、粘性抵抗やその他外乱による影響を無視した理想解として導かれた。ここでは、これらの影響を考慮した場合の振動解が一般にどのように表せるかを考える。

弦の長さ  $L$  の時間的変動は、(2.9) 式を時間関数  $\lambda = n\omega_0(t)$  とすることにより反映され、また粘性抵抗による減衰効果や外乱による励振効果は 1 次微分項を付加することによって表せる。このとき、(2.6) 式は

$$\frac{d^2 \Psi_n}{dt^2} + 2\Delta\beta_n(t) \frac{d\Psi_n}{dt} + n^2 \omega_0^2(t) \Psi_n = 0 \quad (2.14)$$

と書き直せる [71]。特に、定係数  $\Delta\beta_n(t) \rightarrow \Delta\beta_n$ ,  $\omega_0^2(t) \rightarrow \omega_0^2$  としたとき、上式の解は

$$\Psi_n(t) = c_n e^{-\Delta\beta_n t} e^{\sqrt{(\Delta\beta_n)^2 - n^2 \omega_0^2} t} \quad (2.15)$$

となるので、 $\Delta\beta_n(t)$ ,  $\omega_0^2(t)$  の時間的変化が小さければ、近似的に

$$\Psi_n(t) \approx c_n e^{-\int_0^t \Delta\beta_n(s) ds} e^{\int_0^t \sqrt{(\Delta\beta_n(s))^2 - n^2 \omega_0^2(s)} ds} \quad (2.16)$$

と表せると考えてよい。さらに、 $\Delta\beta_n(t) \ll \omega_0(t)$  と仮定し、 $\sqrt{(\Delta\beta_n(t))^2 - n^2\omega_0^2(t)}$  を

$$\begin{aligned}\sqrt{(\Delta\beta_n(t))^2 - n^2\omega_0^2(t)} &= jn\omega_0(t) - j\frac{1}{2n\omega_0^2(t)}(\Delta\beta_n(t))^2 + o((\Delta\beta_n(t))^4) \\ &= jn\omega_0(t) + j\Delta\omega_n(t)\end{aligned}$$

のように、テーラー展開し 0 次項  $jn\omega_0(t)$  とそれ以外の項  $j\Delta\omega_n(t)$  の和として表せば、

$$\Psi_n(t) \approx c_n e^{\int_0^t \Delta c_n(s) ds} e^{jn \int_0^t \omega_0(s) ds} \quad \text{where } \Delta c_n(t) = -\Delta\beta_n(t) + j\Delta\omega_n(t) \quad (2.17)$$

と書き直せる。各振動モードの重ね合わせとしては、

$$\Psi(t) \approx \sum_{n=-\infty}^{\infty} c_n e^{\int_0^t \Delta c_n(s) ds} e^{jn \int_0^t \omega_0(s) ds} \quad (2.18)$$

と表せる。ただし、 $c_{-n} = c_n^*$ ,  $\Delta c_{-n}(t) = \Delta c_n^*(t)$  である。

したがって、微小かつ緩やかなパラメータ変動をともなった場合の振動解は、大まかには周期性を保ちながら、各倍音で振幅と位相が揺らぎ、基本周波数が増減する擬似周期性をもつことになる。

### 2.3.2 「振幅変動+ピッチ変動」信号の定義

このような弦の振動は空気の疎密に変換され空气中を伝搬し、我々の耳やマイクロホンに届き音圧として観測される。音圧への変換はほぼ線形であると考えられるので、その信号波形は (2.18) 式の振動解と定数倍を除いて同じになると考えてよい。そこで本論文では、

**定義 1:** 一般に

$$x(t) = \sum_{n=1}^N c_n e^{\int_0^t \Delta c_n(s) ds} e^{jn \int_0^t \omega_0(s) ds} \quad (2.19)$$

と表せる信号、あるいは複素振幅項をまとめて

$$c_n(t) \equiv c_n e^{\int_0^t \Delta c_n(s) ds} \quad (2.20)$$

とおいた

$$x(t) = \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} \quad (2.21)$$

と表せる信号を「振幅変動+ピッチ変動」信号と定義する。断りがない限り、 $c_n$  は複素数の確率変数、 $\Delta c_n(t)$ ,  $c_n(t)$  は複素数の確率過程、 $\omega_0(t)$  は実数の確率過程をとるものとする。



物理的実現性を考慮して、最高倍音次数を有限値  $N$  とした。実数値をとる信号の場合、振幅が正負の倍音次数で互いに複素共役対称  $c_{-n}(t) = c_n^*(t)$  となるので、簡単のため  $n > 0$  の倍音成分だけで構成される解析信号を仮定した。さらに、不確定な要因をもつ実環境下では信号の挙動を決定論的に表すことはできないので、各パラメータは確率変数もしくは確率過程をとるものとした。特に本論文では、あくまで周期性を主成分とする信号を対象としているので、

**仮定 1:** 振幅やピッチ周波数の変動は十分小さいとする。

また、

**仮定 2:** すべてのパラメータは、(異なる倍音次数間、複素数の実部と虚部の間も含め) 互いに統計的に独立であるとする。ただし、振幅変動パラメータ  $\Delta c_n(t)$  の異なる次数間に関してはこの限りではない。

異なる倍音次数間の振幅変動には独立性を要求しないのは、波形包絡の変化など倍音間で共通した振幅変動をもつ可能性があるからである。さらに、各倍音振幅は複素平面の原点を中心に等方的な確率分布をもつと考えて、

**仮定 3:**

$$E[c_n] = 0, \quad \text{すなわち} \quad E[c_n(t)] = 0, \quad E[x(t)] = 0 \quad (2.22)$$

とする。

ただし、 $E[\cdot]$  は集合平均を表す。

## 2.4 「振幅変動+ピッチ変動」信号推定の問題とは

マイクロホン等を介して実際に観測される信号は、目的音源以外からの到来音や背景雑音、音 - 電気信号変換時の歪みや増幅器の熱雑音 (これらの影響をまとめて確率過程  $v(t)$ ) が加わったものになる。そこで本論文では、

**仮定 4a:** 観測信号  $y(t)$  は、「振幅変動+ピッチ変動」信号  $x(t)$  に対し加法性雑音  $v(t)$  を仮定し、

$$y(t) = x(t) + v(t) \quad (2.23)$$

と表せるものとする。ただし、 $x(t)$  と  $v(t)$  は互いに無相関であり、

$$E[v(t)] = 0 \quad (2.24)$$

とする。

**定義 2:** (2.19) 式や (2.21) 式を構成するパラメータ  $c_n, \Delta c_n(t), \omega_0(t)$  や  $c_n(t)$  に各種の確率モデルを仮定し、適切な評価基準の下で観測信号  $y(t)$  から目的信号  $x(t)$  を推定する問題を「振幅変動+ピッチ変動」信号推定の問題と定義する。

この問題の解法とそこから得られるフィルタ特性について考察することが、本論文の主たる目的になる。

## 2.5 本論文で扱う擬似周期信号の位置づけ

ここでまず、「振幅変動+ピッチ変動」信号の記述にあたってどのような確率モデルを導入するのか、および各確率モデルの下で信号推定にどのようなフィルタ設計法を用いるのかを明確にしておく。

### 2.5.1 確率モデルによる分類

定常確率過程をとるものとそれ以外の非定常確率過程をとるものに「振幅変動+ピッチ変動」信号を大別し、さらに本論文では振幅項やピッチ周波数項に以下のような確率モデルを導入する。特に、決定論的に記述された「振幅変動+ピッチ変動」信号の一例については次節で述べる。

#### 定常型「振幅変動+ピッチ変動」信号

(2.19) 式や (2.21) 式の信号モデルにおいて、振幅項  $c_n(t)$  に広義定常性、ピッチ周波数項  $\omega_0(t)$  に狭義定常性を仮定する<sup>†</sup>。その中の代表的な確率モデルとして、

- (1) 振幅  $c_n(t)$  が一定で任意の確率変数を取り、ピッチ  $\omega_0(t)$  が一定の確定値をとる理想周期信号の場合 (3.4節)
- (2) 振幅  $c_n(t)$  が任意の定常確率過程を取り、ピッチ  $\omega_0(t)$  が一定で任意の確率変数をとる場合 (3.5節)
- (3) 振幅  $\Delta c_n(t)$  とピッチ  $\omega_0(t)$  がともに定常白色 Gauss 過程をとる場合 (3.6節)

のそれぞれについて取り上げる。

---

<sup>†</sup>このとき、信号全体  $x(t)$  としては広義定常になる (3.2.1節参照)。

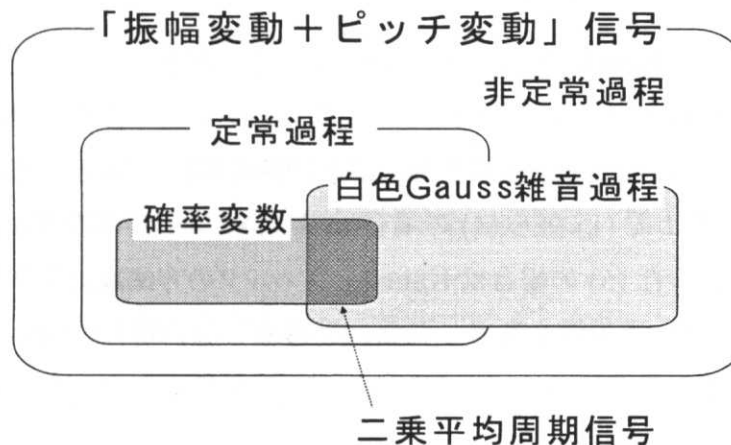


図 2.3 「振幅変動+ピッチ変動」信号の記述にあたって導入する確率モデルの包含関係

### 時変型「振幅変動+ピッチ変動」信号

定常性が成り立たない例としては特に、

#### (4) 振幅 $\Delta c_n(t)$ とピッチ $\omega_0(t)$ がともに時変白色 Gauss 過程をとる場合 (6章)

すなわち分散パラメータが時間変化することを許容した Gauss 過程をとる場合について議論する。これは、分散が時不変な定常白色 Gauss 過程を前提とした定常型「振幅変動+ピッチ変動」信号の議論 (3) の一般化に対応する。

各確率モデルの関係を表したものが図 2.3 である。

### 周期定常過程

上記 2 種類の信号とは異なり本論文の議論の対象とはならないが、名称的に混同し易いと思われる「周期定常過程」 [72] について言及しておく。これは、任意の時刻での確率分布が特定周期の整数倍の時間原点移動に対して不変であるような確率過程として定義される<sup>†</sup>。言い換えれば、周期定常過程では一定周期毎に無限の過去から無限の未来まで同一の統計的性質が繰り返し現れることが保証される。したがって、時間の経過にともなって振幅やピッチが変化することを許容した擬似周期信号とは本質的に異なるものであることに注意する。ただし (1-a) の二乗平均周期信号は、周期定常過程にも含まれる唯一のケースになる。周期定常過程の典型例としては、静止画が映っているときの TV 受像機の走査信号があげられる。

<sup>†</sup>これは狭義における定義であり、同様に相関関数に対してだけ成り立ては広義の定義となる。

### 2.5.2 フィルタ設計法による分類

各確率モデルに従う「振幅変動+ピッチ変動」信号を観測信号から推定するためのフィルタとして、本論文では線形フィルタを仮定する。このときのフィルタ設計は、最小二乗規範の下で導出される Wiener-Hopf 方程式を解くことに帰着する。Wiener-Hopf 方程式は、信号に定常性を仮定した上記 (1) から (4) の場合は Wiener フィルタの方法論により (主に第 3 章)、非定常性を許容した (5) の場合は Kalman フィルタの方法論により解くことができる (第 6 章)。特に (4) では、因果性を考慮した場合と考慮しない場合の両面からフィルタ設計を行い、因果性を考慮した Wiener フィルタの解が、定常性を仮定したときの Kalman フィルタの解に一致することを確認する。

## 2.6 複素チャープ位相をもつ音響・音声信号モデル

「振幅変動+ピッチ変動」信号推定の問題を考える上での予備調査の意味で、解析の容易な複素チャープ位相をもつ決定論的信号を例にあげ、その周波数特性が示す傾向について調べておく。そのために、(2.19) 式において、 $c_n$  を複素定数、

$$\begin{aligned}\Delta c_n(t) &= -(\Delta a_n + j\Delta b_n)t \quad (\text{定係数の 1 次式}) \\ \omega_0(t) &= \Delta\gamma_0 t + \omega_0 \quad (\text{定係数の 1 次式})\end{aligned}\quad (2.25)$$

とにおいて ( $\Delta a_n > 0$ )、すべてのパラメータを定数もしくは定係数の 1 次式とした

信号モデル 1:

$$x(t) = \sum_{n=1}^N c_n e^{-\frac{1}{2}\{\Delta a_n + j(\Delta b_n + n\Delta\gamma_0)\}t^2} e^{jn\omega_0 t} \quad (2.26)$$

を考える。ただし、**仮定 1** より、

$$\Delta a_n, |\Delta b_n|, |\Delta\gamma_0| \ll \omega_0^2 \quad (2.27)$$

が成り立つと仮定する。実際に、(2.26) 式のようなチャープ信号モデルを使って、 $\omega_0$  や  $\Delta\gamma_0$  などの未知パラメータを、観測信号との最小二乗フィッティング (非線形最小二乗法) により推定する方法が提案されている [73]。

複素 Gauss 関数の Fourier 変換が

$$\begin{aligned}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(a+jb)t^2} e^{-j\omega t} dt &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(a+jb)(t+\frac{j\omega}{a+jb})^2} dt \cdot e^{-\frac{\omega^2}{2(a+jb)}} \\ &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(a+jb)t^2} dt \cdot e^{-\frac{\omega^2}{2(a+jb)}} \\ &= \sqrt{\frac{2\pi}{a+jb}} e^{-\frac{\omega^2}{2(a+jb)}}\end{aligned}$$

と表せることを利用すれば(ただし  $a > 0$ )、(2.26) 式の Fourier 変換は、

$$X(\omega) = \sum_{n=1}^N c_n \sqrt{\frac{2\pi}{\Delta a_n + j(\Delta b_n + n\Delta\gamma_0)}} e^{-\frac{1}{2} \frac{(\omega - n\omega_0)^2}{\Delta a_n + j(\Delta b_n + n\Delta\gamma_0)}} \quad (2.28)$$

となる。したがって、(2.27) 式を考慮すれば、 $\omega = n\omega_0$  の近傍に分布が集中し異なる倍音間での干渉はほとんど無視できるため、そのスペクトル分布は

$$|X(\omega)| \approx \sum_{n=1}^N c_n \sqrt{\frac{2\pi}{\sqrt{\Delta a_n} \sigma_n}} e^{-\frac{\omega^2}{2\sigma_n^2}} \quad (2.29)$$

と表せ、分散  $\sigma_n^2$  の Gauss 分布が倍音周波数毎に立つ楕円分布になる。ただし、

$$\sigma_n^2 = \Delta a_n + \frac{1}{\Delta a_n} (\Delta b_n + n\Delta\gamma_0)^2 \quad (2.30)$$

とおいた。

特に  $\Delta a_n, \Delta b_n, \gamma_0 \rightarrow 0$  のとき、(2.26) 式は完全な周期信号になり、(2.28) 式は

$$X(\omega) = \sum_{n=1}^N 2\pi c_n \delta(\omega - n\omega_0) \quad (2.31)$$

となって、倍音スペクトル構造をもつことが示される。

(2.29) 式の意味を考えるにあたって簡単のため、 $\Delta a_1 = \dots = \Delta a_N = \Delta a, \Delta b_1 = \dots = \Delta b_N = \Delta b$  であり、低次倍音(小さい  $n$ )で  $\Delta a, |\Delta b| \gg n|\Delta\gamma_0|$ 、高次倍音(大きい  $n$ )で  $\Delta a, \Delta b \ll n|\Delta\gamma_0|$  が成り立つとする。このとき、各 Gauss 分布の標準偏差すなわち占有帯域幅は、低次倍音では  $\sigma_n \approx \sqrt{\Delta a + \frac{(\Delta b)^2}{\Delta a}}$  の一定帯域幅となり、高次倍音では  $\sigma_n \approx n\sqrt{\frac{(\Delta\gamma_0)^2}{\Delta a}}$  となって、倍音次数に比例して帯域幅が広がるいわゆる一定 Q 値をもつことになる。これは、振幅変動が各倍音で等しく帯域幅を広げる効果があるのに対して、ピッチ変動が高次倍音になるほど大きく帯域幅を広げる効果があり、低次倍音で前者が、高次倍音で後者が支配的になることを意味している。したがって、このような信号を抽出するフィルタとしても、同じような周波数特性をもつことが最適であると考えられる。実際、図 1.1 に示すように聴覚の周波数特性も同様な傾向をもつことは、このことを裏付けていると考えられる。

## 2.7 本章のまとめ

本章では、以後の議論の準備として、擬似周期信号とその推定問題とは何かを明確にするための定義を与えた。まず、母音音声の主としてピッチ変動を伴った擬似周期性をもつことを説明し(2.2節)、その信号モデルが共振長の変化や減衰増幅のある振動現象を記述した波動方程式の解から導けることを示した(2.3.1項)。振幅項やピッチ周波数項にあらため

て確率過程を導入することによって一般的な形で「振幅変動+ピッチ変動」信号を定義し(2.3.2項)、本論文で扱う「振幅変動+ピッチ変動」信号推定の問題とは何かについて明確化した(2.4節)。振幅変動やピッチ変動にいくつかの代表的な確率モデルを仮定し、本論文で扱う擬似周期信号の位置づけを整理した(2.5節)。特に、音響・音声信号モデルとして複素チャープ位相をもった確定信号を例にあげ、「振幅変動+ピッチ変動」信号のもつ基本的な性質について概観した(2.6節)。

## 第 3 章

# 定常線形最適フィルタによる音響・音声信号の分離

### 3.1 はじめに

擬似周期信号は、周期性を主成分にもちながら振幅やピッチ周波数が変化する信号として、第 2 章において一般的な形で定式化された。特にピッチ変動には、音楽での旋律や声の抑揚などの時間的に緩やかに変化する成分以外にも、振動体の熱的擾乱や声帯の不規則振動による微細な変動成分があり、これが楽音や音声の自然性を保つのに大きく関与していることを 2.2 節で述べた。本章では、この微細ピッチ変動がほぼ定常的であることに着目し、定常確率過程をとる擬似周期信号のみに限定した上で、信号分離にあたっての具備すべきフィルタ特性について考察する。時間的に緩やかに変化する成分も同時に考慮した、より一般的な時変型擬似周期信号に対しては 6 章で議論する。

定常性を仮定することによって、以下の点で見通しの良い議論が可能になる。第 1 に、時刻に不変で時間差のみの関数として自己相関あるいは相互相関関数が表せるので、Wiener-Khinchine の定理によりスペクトル密度関数が定義でき、信号の性質を周波数特性の面から解釈が可能になること、第 2 に、信号推定のためのフィルタ演算に線形性を仮定すれば、フィルタ入出力間が時不変な畳み込み演算で表せるため周波数軸上でのフィルタ設計が可能になることである。その結論として、伝達関数が信号と雑音のスペクトル密度の内分比として表せる Wiener フィルタが導かれ、振幅変動やピッチ変動がフィルタ設計に及ぼす影響について、周波数特性の面から詳細に考察することが可能になる。

### 3.2 定常型「振幅変動+ピッチ変動」信号の一般的性質

#### 3.2.1 自己相関関数

まず、本章で扱う定常型「振幅変動+ピッチ変動」信号を以下に定義する。

定義 3: (2.19) 式の「振幅変動+ピッチ変動」信号モデル

$$x(t) = \sum_{n=1}^N c_n e^{\int_0^t \Delta c_n(s) ds} e^{jn \int_0^t \omega_0(s) ds} \quad (3.1)$$

あるいは、(2.21) 式

$$x(t) = \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} \quad (3.2)$$

において、複素振幅  $c_n(t) = c_n e^{\int_0^t \Delta c_n(s) ds}$  が広義定常であり、ピッチ周波数  $\omega_0(t)$  が狭義定常であるとき、特に、定常型「振幅変動+ピッチ変動」信号と呼ぶ。

定義より、複素振幅  $c_n(t)$  の自己相関は

$$E[c_n(t+\tau)c_n^*(t)] = R_{c_n}(\tau) \quad (3.3)$$

のように、キャリア項  $\psi_n(t) \equiv e^{jn \int_0^t \omega_0(s) ds}$  の自己相関については

$$\begin{aligned} E[\psi_n(t+\tau)\psi_n^*(t)] &= E[e^{jn \int_t^{t+\tau} \omega_0(s) ds}] \\ &= E[e^{jn \int_0^\tau \omega_0(s) ds}] \\ &\equiv R_{\psi_n}(\tau) \end{aligned} \quad (3.4)$$

のように、両者とも時間差のみの関数として表せる。1 段目から 2 段目は  $\omega_0(t)$  が狭義定常であることを利用している。したがって、信号全体の自己相関関数においても

$$\begin{aligned} E[x(t+\tau)x^*(t)] &= \sum_{m=1}^N \sum_{n=1}^N E[c_m(t+\tau)c_n^*(t)] E[e^{jm \int_0^{t+\tau} \omega_0(s) ds} e^{-jn \int_0^t \omega_0(s) ds}] \\ &= \sum_{n=1}^N E[c_n(t+\tau)c_n^*(t)] E[e^{jn \int_t^{t+\tau} \omega_0(s) ds}] \\ &= \sum_{n=1}^N R_{c_n}(\tau) R_{\psi_n}(\tau) \\ &\equiv R_x(\tau) \end{aligned} \quad (3.5)$$

が成り立ち、 $x(t)$  は広義定常となる。ただし、1 段目および 2 段目は**仮定 2**を用いた。特に、(3.2) 式の各倍音成分

$$x_n(t) = c_n(t) e^{jn \int_0^t \omega_0(s) ds} \quad (3.6)$$

に対する自己相関

$$E[x_n(t+\tau)x_n^*(t)] = R_{c_n}(\tau) R_{\psi_n}(\tau) \equiv R_{x_n}(\tau) \quad (3.7)$$



が、(3.5) 式の各倍音要素に対応している。ピッチ周波数が狭義定常であることより、その期待値は

$$E[\omega_0(t)] = \bar{\omega}_0 \quad (3.8)$$

のように時間に依存しない定数となることに注意する。

### 3.2.2 スペクトル特性

このとき、信号  $x(t)$  のスペクトル密度関数 (あるいはパワースペクトル)  $S_x(\omega)$  は、自己相関関数の Fourier 変換として、

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau = \sum_{n=1}^N S_{x_n}(\omega) \quad (3.9)$$

と表せる。ただし、

$$R_{x_n}(\tau) \xleftrightarrow{\mathcal{F.T.}} S_{x_n}(\omega) = \frac{1}{2\pi} S_{c_n}(\omega) * S_{\psi_n}(\omega) \quad (3.10)$$

であり、 $S_{x_n}(\omega)$  は各倍音成分  $x_n(t)$  のスペクトル密度関数を意味する。また、 $S_{c_n}(\omega)$ 、 $S_{\psi_n}(\omega)$  はそれぞれ  $c_n(t)$ 、 $\psi_n(t)$  のスペクトル密度関数すなわち

$$R_{c_n}(\tau) \xleftrightarrow{\mathcal{F.T.}} S_{c_n}(\omega), \quad R_{\psi_n}(\tau) \xleftrightarrow{\mathcal{F.T.}} S_{\psi_n}(\omega) \quad (3.11)$$

であり、 $*$  は畳み込み演算

$$F(\omega) * G(\omega) = \int_{-\infty}^{\infty} F(\Omega) G(\omega - \Omega) d\Omega \quad (3.12)$$

を表す。

仮定 1 を考慮すると、 $S_{c_n}(\omega)$  は  $\omega = 0$  に、 $S_{\psi_n}(\omega)$  は  $\omega = n\bar{\omega}_0$  に分布が集中するインパルス状の関数になると考えられる。したがって、

**性質 1:**  $n$  次倍音のスペクトル成分  $S_{x_n}(\omega)$  は、 $\omega = n\bar{\omega}_0$  にピークをもつインパルス状の分布となる。

さらに、(3.9) 式より、

**性質 2:** 「振幅変動+ピッチ変動」信号のスペクトル密度分布  $S_x(\omega)$  は、周波数上等間隔 ( $\omega = \bar{\omega}_0, \dots, N\bar{\omega}_0$ ) にピークが並ぶ櫛形分布となる。

各櫛歯の形状は振幅変動やピッチ変動のダイナミクスによって決まることが予想され、この仕組みの解明が本論文の中心課題の一つになる。

### 3.3 「振幅変動+ピッチ変動」信号フィルタの基本構造

#### 3.3.1 観測雑音モデル

定義 2 に基づく「振幅変動+ピッチ変動」信号推定の問題を考えるにあたって、本章ではさらに (2.23) 式の観測モデル

$$y(t) = x(t) + v(t) \quad (3.13)$$

において、

仮定 5: 観測雑音  $v(t)$  は広義定常過程であるとする。

すなわち、 $v(t)$  の自己相関関数とスペクトル密度関数の間にも

$$E[v(t+\tau)v^*(t)] = R_v(\tau) \xleftrightarrow{\mathcal{F.T.}} S_v(\omega) \quad (3.14)$$

が成り立つとする。仮定 4a を考慮すれば、観測信号  $y(t)$  のスペクトル密度関数は

$$S_y(\omega) = S_x(\omega) + S_v(\omega) \quad (3.15)$$

となる。

#### 3.3.2 非因果的 Wiener フィルタの導出

このとき、以下のフィルタ設計の問題を考える。

問題設定 1: (3.13) 式に従う観測データ  $y(s)$  ( $-\infty < s < \infty$ ) を用いて、目的信号  $x(t)$  に対する任意の時刻  $t$  での推定値  $\hat{x}(t)$  を、線形時不変フィルタの出力

$$\hat{x}(t) = \int_{-\infty}^{\infty} h(s)y(t-s)ds \quad (3.16)$$

として、平均二乗誤差

$$J = E[|x(t) - \hat{x}(t)|^2] \quad (3.17)$$

が最小になるように求めたい。このときの非因果的フィルタのインパルス応答  $h(t)$  (あるいはその伝達関数  $H(\omega)$ ) を決定する問題を考える。

(3.17) 式を最小にすることは、推定誤差  $x(t) - \hat{x}(t)$  がすべての時刻の観測信号  $y(t-\tau)$  と無相関

$$E[\{x(t) - \hat{x}(t)\}y^*(t-\tau)] = 0 \quad (3.18)$$

になることと等価である (直交射影定理)[77]。これに (3.16) 式を代入し整理すると、Wiener-Hopf 方程式

$$E[x(t+\tau)y^*(t)] = \int_{-\infty}^{\infty} h(s)E[y(t-s)y^*(t-\tau)]ds \quad (3.19)$$

が得られる<sup>†</sup>。仮定 4 および (3.5),(3.14) 式より、

$$\begin{aligned} E[x(t+\tau)y^*(t)] &= R_x(\tau), \\ E[y(t-s)y^*(t-\tau)] &= R_x(\tau-s) + R_v(\tau-s) \equiv R_y(\tau-s) \end{aligned} \quad (3.20)$$

と表せることから、

$$R_x(\tau) = \int_{-\infty}^{\infty} h(s)R_y(\tau-s)ds \quad (3.21)$$

となり、両辺を Fourier 変換することによって、

$$H(\omega) = \frac{S_x(\omega)}{S_y(\omega)} = \frac{S_x(\omega)}{S_x(\omega) + S_v(\omega)} \quad (3.22)$$

と解くことができる<sup>††</sup>。このフィルタの出力、すなわち (3.16) 式を通して求められる推定値  $\hat{x}(t)$  は、(2.22), (2.24) 式より

$$\begin{aligned} E[\hat{x}(t)] &= \int_{-\infty}^{\infty} h(s)E[y(t-s)]ds = 0 \\ &= E[x(t)] \end{aligned} \quad (3.23)$$

が成り立つので、不偏推定量となることに注意する。

### 3.3.3 フィルタの周波数特性

(3.9) 式より (3.22) 式は、各倍音成分を抽出するフィルタ

$$H_n(\omega) = \frac{S_{xn}(\omega)}{S_x(\omega) + S_v(\omega)} \quad (3.24)$$

の和として

$$H(\omega) = \sum_{n=1}^N H_n(\omega) \quad (3.25)$$

と表せる。特に性質 1 を考慮すれば、近似的に

$$H_n(\omega) \approx \frac{S_{xn}(\omega)}{S_{xn}(\omega) + S_v(\omega)} \quad (3.26)$$

<sup>†</sup>(3.17) 式に (3.16) 式を代入し、 $h(t)$  の変分問題を解くことによっても同じ結果が得られる。

<sup>††</sup>これは、非因果的 Wiener フィルタとして知られる [77]。

と表せることに注意する。ここで、観測雑音にほぼ一様なスペクトル密度分布を仮定し、

仮定 6: 倍音周波数  $\omega = n\bar{\omega}_0$  の近傍では、

$$S_{xn}(\omega) \gg S_v(\omega)$$

それ以外の周波数では、

$$S_{xn}(\omega) \ll S_v(\omega)$$

が成り立つとする。

このとき、(3.24) 式より、

性質 3: 倍音周波数  $\omega = n\bar{\omega}_0$  の近傍では、

$$H_n(\omega) \approx 1$$

それ以外の周波数では、

$$H_n(\omega) \approx 0$$

となつて、 $H_n(\omega)$  は  $n$  次倍音を抽出する狭帯域バンドパスフィルタとなる。

したがって、(3.25) 式より、

性質 4:  $H(\omega)$  は、倍音周波数毎に狭帯域のバンドパスフィルタが並ぶ楕形フィルタとなる。

### 3.3.4 平均二乗推定誤差

上記の線形最適フィルタによる信号推定の平均二乗誤差は、(3.17) 式に (3.16) 式を代入し、(3.18) 式および (3.20) 式を考慮すれば、

$$\begin{aligned} J &= E[\{x(t) - \hat{x}(t)\}x^*(t)] - \int_{-\infty}^{\infty} h^*(s)E[\{x(t) - \hat{x}(t)\}y^*(t-s)]ds \\ &= E[\{x(t) - \hat{x}(t)\}x^*(t)] \\ &= R_x(0) - \int_{-\infty}^{\infty} h(s)R_x(s)ds \end{aligned} \quad (3.27)$$

で与えられ、さらに周波数軸上では、(3.22) 式より

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} \{S_x(\omega) - H^*(\omega)S_x(\omega)\}d\omega \quad (3.28)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_x(\omega)S_v(\omega)}{S_x(\omega) + S_v(\omega)}d\omega \quad (3.29)$$

と表せる。

## 3.3.5 フィルタ導出の別解法

性質 3 の結論は、(3.6) 式の各倍音成分を直接推定するフィルタの設計問題を考えることによっても導出できる。すなわち、

**問題設定 2:** (3.13) 式に従う観測データ  $y(s)$  ( $-\infty < s < \infty$ ) を用いて、倍音成分ベクトル  $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$  に対する任意の時刻  $t$  での推定ベクトル  $\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \dots, \hat{x}_N(t)]^T$  を、線形時不変フィルタの出力

$$\hat{\mathbf{x}}(t) = \int_{-\infty}^{\infty} \mathbf{h}(s)y(t-s)ds \quad (3.30)$$

として、平均二乗誤差

$$J = E[\|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|^2] \quad (3.31)$$

が最小になるように求める。このときのインパルス応答  $\mathbf{h}(t) = [h_1(t), \dots, h_N(t)]^T$  (あるいはその伝達関数  $\mathbf{H}(\omega) = [H_1(\omega), \dots, H_N(\omega)]^T$ ) を決定する問題を考える。

この場合も、(3.18) 式と同様に、直交原理

$$E[\{\mathbf{x}(t) - \hat{\mathbf{x}}(t)\}y^*(t-\tau)] = \mathbf{0} \quad (3.32)$$

から Wiener-Hopf 方程式

$$E[\mathbf{x}(t+\tau)y^*(t)] = \int_{-\infty}^{\infty} \mathbf{h}(s)E[y(t-s)y^*(t-\tau)]ds \quad (3.33)$$

が導かれ、両辺を Fourier 変換し整理すると、各倍音において (3.24) 式と同じ結果が得られる。

さらに、仮定 4a の観測信号モデルに代わって、

**仮定 4b:** 観測信号  $y(t)$  は、(3.6) 式による単一の倍音成分 (正弦波)  $x_n(t)$  を目的信号とし、

$$y(t) = x_n(t) + v(t) \quad (3.34)$$

と表せるものとする。

このとき、

**問題設定 3:** (3.34) 式に従う観測データ  $y(s)$  ( $-\infty < s < \infty$ ) を用いて、目的信号  $x_n(t)$  に対する任意の時刻  $t$  での推定値  $\hat{x}_n(t)$  を、線形時不変フィルタの出力

$$\hat{x}_n(t) = \int_{-\infty}^{\infty} h_n(s)y(t-s)ds \quad (3.35)$$

として、平均二乗誤差

$$J = E[|x_n(t) - \hat{x}_n(t)|^2] \quad (3.36)$$

が最小になるように求める。このときのインパルス応答  $h_n(t)$  (あるいはその伝達関数  $H_n(\omega)$ ) を決定する問題を考える。

このときも同様に、Wiener-Hopf 方程式の解として

$$H_n(\omega) = \frac{S_{xn}(\omega)}{S_{xn}(\omega) + S_v(\omega)} \quad (3.37)$$

と求められる。複数の倍音成分が含まれる場合には (3.26) 式のように近似式であったのに対して、倍音間の干渉がない本問題設定では上式のように厳密に表せ、**問題設定 3** の結論が「振幅変動+ピッチ変動」信号中の各倍音成分を抽出するフィルタにとっての近似解に対応していることに注意する。

### 3.4 理想周期信号に対するフィルタ設計

一般の「振幅変動+ピッチ変動」信号について議論する前に、振幅とピッチ周波数がともに一定値をとる理想周期信号について考える。すなわち、(3.1) 式の信号モデルにおいて、 $\Delta c_n(t) = 0$ ,  $\omega_0(t) = \bar{\omega}_0$  (定数) とおき、

信号モデル 2:

$$x(t) = \sum_{n=1}^N c_n e^{jn\bar{\omega}_0 t} \quad (3.38)$$

とする。このとき、上式は Fourier 級数展開式そのものになり、

$$c_n = \frac{1}{T_0} \int_0^{T_0} x(t) e^{-jn\bar{\omega}_0 t} dt, \quad \text{where } T_0 = \frac{2\pi}{\bar{\omega}_0} \quad (3.39)$$

が成り立つ。仮定 2 および仮定 3 より、

$$E[c_n] = 0, \quad E[c_n c_m^*] = \begin{cases} \bar{x}_n^2, & n = m \\ 0, & n \neq m \end{cases} \quad (3.40)$$

を満たすことに注意すれば、

$$E[x(t+\tau)x^*(t)] = R_x(\tau) = \sum_{n=1}^N \bar{x}_n^2 e^{jn\bar{\omega}_0 \tau} \quad (3.41)$$

と表せ、自己相関関数は

$$R_x(\tau + T_0) = R_x(\tau) \quad (3.42)$$

のように周期関数となり、同時に

$$E[|x(t + T_0) - x(t)|^2] = 0 \quad (3.43)$$

が成り立つ。すなわち (3.38) 式は、二乗平均の意味で周期的な信号になることがわかる。逆に上式が成り立てば、(3.38) 式による Fourier 級数展開表現が可能になり、(3.40) 式のように Fourier 係数  $c_n$  は異なる倍音間では互いに無相関になることが示せる [77]。

(3.38) 式のスペクトル密度関数は、(3.41) 式を Fourier 変換することにより、

$$S_x(\omega) = 2\pi \sum_{n=1}^N \bar{x}_n^2 \delta(\omega - n\bar{\omega}_0) \quad (3.44)$$

と表せることから、(3.22) 式によるフィルタの伝達関数は、

$$H(\omega) = \begin{cases} 1, & n\bar{\omega}_0 - \epsilon < \omega < n\bar{\omega}_0 + \epsilon \quad (n = 1, \dots, N) \\ 0, & \text{other} \end{cases} \quad (3.45)$$

と表せる。ただし、 $\epsilon$  は任意に小さな正定数である。特に、Fourier 係数  $c_n$  にガウス分布、観測雑音  $v(t)$  にガウス過程を仮定できるなら、目的信号  $x(t)$  と観測信号  $y(t)$  がともにガウス過程となるので、線形フィルタを仮定して得られた結論であっても、上式はあらゆるフィルタの中で最適になることに注意する [77]。

このときのフィルタ出力に対する平均二乗推定誤差は、(3.28) 式より、 $J \rightarrow 0$  となつて、推定値  $\hat{x}(t)$  は目的信号  $x(t)$  に一致する。すなわち理想周期信号の場合は、離散スペクトル構造にちょうど合致した楕円特性をもつフィルタによって背景雑音を完全に抑圧できることを意味する [65]。

### 3.5 振幅変動が定常過程をとりピッチ周波数が確率変数をとる場合

音程 (ピッチ周波数) を一定に保ったまま発声しても、音量 (振幅) は細かく変化し、そのときのピッチ周波数を正確に知ることにも現実には困難である。この場合を想定し、(3.2) 式に対して

信号モデル 3:

$$x(t) = \sum_{n=1}^N c_n(t) e^{jn\omega_0 t} \quad (3.46)$$

を考える。ただし、複素振幅  $c_n(t)$  は (3.3) 式で任意の  $R_{c_n}(\tau)$  をとる広義定常過程とし、ピッチ周波数  $\omega_0$  は推定時の不確定性を考慮し、任意の確率密度関数  $P(\omega_0)$  をとる確率変数とする。

†(3.27) 式を用いても同じ結果が得られる。

$P(\omega_0)$  の特性関数を  $p(\tau)$  とすれば、キャリア項  $\psi_n(t) = e^{jn\omega_0 t}$  の自己相関およびスペクトル密度関数は、 $n \neq 0$  で

$$\begin{aligned} R_{\psi_n}(\tau) &= E[e^{jn\omega_0 \tau}] = p(n\tau) \\ S_{\psi_n}(\omega) &= \frac{2\pi}{n} P\left(\frac{\omega}{n}\right) \end{aligned} \quad (3.47)$$

となるから、(3.10), (3.25) 式よりフィルタの伝達関数は

$$H(\omega) \approx \sum_{n=1}^N \frac{S_{cn}(\omega) * \frac{1}{n} P\left(\frac{\omega}{n}\right)}{S_{cn}(\omega) * \frac{1}{n} P\left(\frac{\omega}{n}\right) + S_v(\omega)} \quad (3.48)$$

と表せる。

### 3.5.1 振幅変動のみを考慮した場合

特に、ピッチ周波数が正確に決定できて  $P(\omega_0) = \delta(\omega_0 - \bar{\omega}_0)$  が成立つとき、

$$H(\omega) \approx \sum_{n=1}^N \frac{S_{cn}(\omega - n\bar{\omega}_0)}{S_{cn}(\omega - n\bar{\omega}_0) + S_v(\omega)} \quad (3.49)$$

となる。さらに簡単のため、

**仮定 7a:** すべての倍音で二乗平均振幅が等しく共通の振幅変動をもつ

と仮定し、 $S_{c1}(\omega) = \dots = S_{cN}(\omega) = S_c(\omega)$  とおき、

**仮定 8a:** 観測雑音  $v(t)$  は、

$$E[v(t)] = 0, \quad E[v(t+\tau)v(t)] = \sigma_v^2 \delta(\tau) \quad (3.50)$$

が成り立ち、そのスペクトル密度関数が

$$S_v(\omega) = \sigma_v^2 \quad (3.51)$$

と表せる定常白色雑音である

とすれば、(3.49) 式は

$$H(\omega) \approx \sum_{n=1}^N H_0(\omega - n\bar{\omega}_0), \quad \text{where } H_0(\omega) = \frac{S_c(\omega)}{S_c(\omega) + \sigma_v^2} \quad (3.52)$$

と表せ、

**性質 5:** 倍音周波数毎に等帯域幅のバンドパスフィルタが並ぶ櫛形特性を有する

ことになる。このような周波数特性をもつフィルタを特に「定 BW(Band-Width) 櫛形フィルタ」と呼ぶ(図 3.1 参照)。



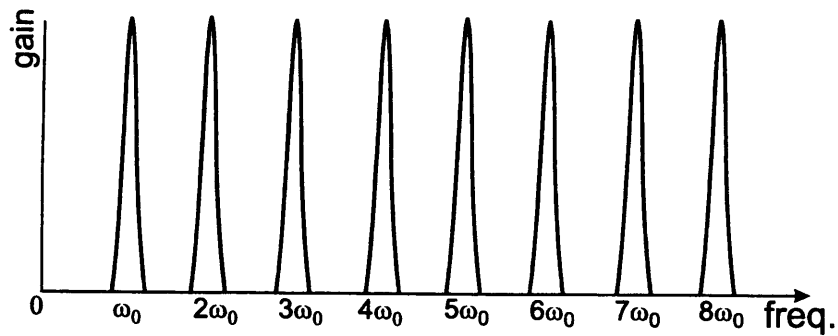


図 3.1 定 BW 櫛形フィルタの周波数特性

### 3.5.2 ピッチ周波数の不確定性のみを考慮した場合

一方、振幅変動が無視できて  $S_{cn}(\omega) = 2\pi\bar{x}_n^2\delta(\omega)$  が成り立ち、ピッチ周波数の不確定さのみを考慮した場合、(3.48) 式は

$$H(\omega) \approx \sum_{n=1}^N \frac{\frac{2\pi\bar{x}_n^2}{n} P(\frac{\omega}{n})}{\frac{2\pi\bar{x}_n^2}{n} P(\frac{\omega}{n}) + S_v(\omega)} \quad (3.53)$$

となる。このとき、仮定 7a の代わりに、

仮定 7b: 倍音次数に比例して二乗平均振幅が大きくなり、 $\bar{x}_1^2 = \bar{x}^2, \bar{x}_2^2 = 2\bar{x}^2, \dots, \bar{x}_N^2 = N\bar{x}^2$  が成り立つ

とした場合、あるいは仮定 8a の代わりに、

仮定 8b: 倍音次数に反比例して雑音成分が小さくなり  $S_v(\omega)$  が各倍音周波数  $\omega = n\bar{\omega}_0$  の近傍で一定値  $\frac{\sigma_v^2}{n}$  をとる

とした場合、(3.53) 式は

$$H(\omega) \approx \sum_{n=1}^N H_1(\frac{\omega}{n}), \quad \text{where } H_1(\omega) = \frac{2\pi\bar{x}^2 P(\omega)}{2\pi\bar{x}^2 P(\omega) + \sigma_v^2} \quad (3.54)$$

と表せる。 $H_1(\omega)$  が  $\omega = \bar{\omega}_0$  にピークをもつインパルス状の関数となるから、上式は

**性質 6:** 櫛形特性をもつと同時に、各通過帯域が倍音次数に比例して広がる定 Q 特性を有する

ことになる。このようなフィルタを特に「定 Q 櫛形フィルタ」と呼ぶ(図 3.2参照)。

定 Q 櫛形特性は、ピッチの不確定性によって生じる高次倍音でのスペクトル広がりを吸収する働きをもつ反面、その帯域での観測雑音を多く拾ってしまうという相反する性質があ

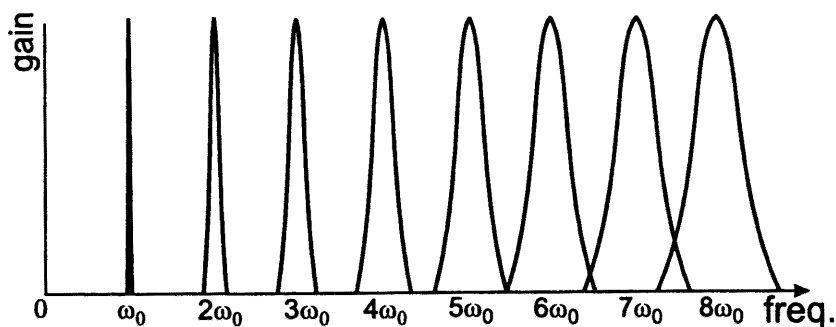


図 3.2 定 Q 櫛形フィルタの周波数特性

る。上の議論では、**仮定 8b** あるいは**仮定 8a** を仮定することによってこの問題を回避したが、この仮定が成り立たない場合でも、目的信号に比べて雑音成分の寄与が無視できる程度に小さければ、同様な結論を得ることができる。すなわち、 $P(\omega)$  が  $\omega = \bar{\omega}_0$  の近傍だけに分布が集中し、それ以外では急激に 0 になると考えて、

**仮定 9:** すべての倍音次数  $n$  において、 $\bar{\omega}_0 - \epsilon_\omega < \omega < \bar{\omega}_0 + \epsilon_\omega$  で  $\frac{2\pi\bar{x}_n^2}{n} P(\omega) \gg S_v(\omega)$ 、それ以外で  $P(\omega) = 0$  が成り立つとする。

このとき (3.53) 式は

$$H(\omega) \approx \sum_{n=1}^N \text{rect}\left(\frac{\omega - n\bar{\omega}_0}{n\epsilon_\omega}\right) \quad (3.55)$$

と表せるので ( $\text{rect}(x)$  は  $-1 < x < 1$  のとき 1、それ以外るとき 0 となる矩形関数)、この場合も**性質 6** と同じ定 Q 櫛形特性を有することになる。

### 3.5.3 両者を同時に考慮した場合

この場合は一般に次のような傾向をもつ。簡単のため、**仮定 7a** が成り立ち、 $P(\omega)$  が  $S_c(\omega)$  に比べて十分狭帯域であるとする。このとき低次倍音では、 $P(\frac{\omega}{n})$  は  $S_c(\omega)$  からみてデルタ関数としてみなせるから、(3.48) 式は (3.52) 式で近似でき、逆に高次倍音では、 $S_c(\omega)$  が  $P(\frac{\omega}{n})$  からみてデルタ関数としてみなせるようになるから (3.54) 式で近似できる。したがって、周波数全体では、

**性質 7:** 低次倍音では定 BW 櫛形特性、高次倍音では定 Q 櫛形特性、低次から高次までの中間領域では、定 BW 型から定 Q 型に徐々に遷移する複合櫛形特性を有する

ことになる。このようなフィルタを特に「定 BW/ 定 Q 複合櫛形フィルタ」と呼ぶ(図 3.3 参照)。

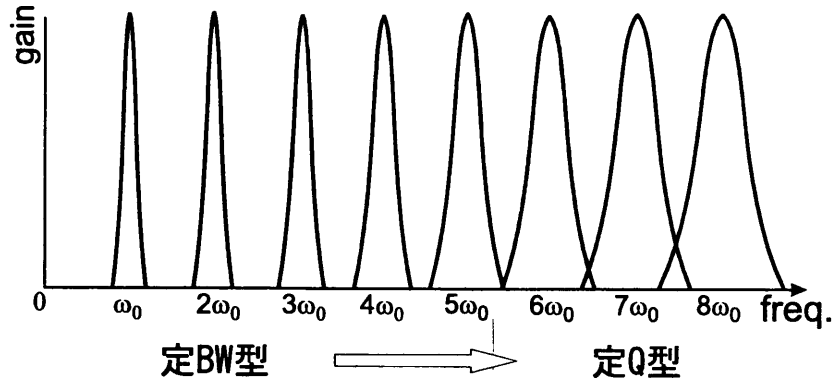


図 3.3 定 BW/ 定 Q 複合櫛形フィルタの周波数特性

### 3.6 振幅とピッチ周波数が定常白色 Gauss 雑音過程をとる場合

音声にみられる振幅とピッチ周波数の変動の多くは Gauss 過程をとること [39]、およびその周波数特性が概ね白色である [34] という知見を考慮し、(3.1) 式で

$$\begin{aligned} \Delta c_n(t) &= -\text{sign}(t)\sigma_{an}^2 + \sigma_{an}\xi_{an}(t) + j\sigma_{bn}\xi_{bn}(t) \quad (\text{複素白色 Gauss 雑音過程}), \\ \omega_0(t) &= \bar{\omega}_0 + \sigma_\omega\xi_\omega(t) \quad (\text{白色 Gauss 雑音過程}) \end{aligned} \quad (3.56)$$

とおいた場合を考える。ただし、 $\text{sign}(t)$  は  $t \geq 0$  のとき 1、 $t < 0$  のとき  $-1$  をとる符号関数であり、 $\xi_{an}(t), \xi_{bn}(t), \xi_\omega(t)$  は  $E[\xi_*(t)] = 0$ 、 $E[\xi_*(t+\tau)\xi_*(t)] = \delta(\tau)$  が成り立つ白色 Gauss 雑音、 $\sigma_{an}, \sigma_{bn}, \sigma_\omega > 0$  はそれぞれ  $\Delta c_n(t)$  の実部と虚部、 $\omega_0(t)$  に対する標準偏差である。

このとき、(3.1) 式は

信号モデル 4:

$$x(t) = \sum_{n=1}^N c_n e^{-\sigma_{an}^2|t| + \sigma_{an}\beta_{an}(t)} e^{j\{\sigma_{bn}\beta_{bn}(t) + n\sigma_\omega\beta_\omega(t)\}} e^{jn\bar{\omega}_0 t} \quad (3.57)$$

となる。ただし、 $\beta_*(t)$  は

$$\beta_*(t) \equiv \int_0^t \xi_*(s) ds \quad (3.58)$$

で定義される Brown 運動過程 (あるいは Wiener 過程) である [80]。(3.57) 式中の  $-\sigma_{an}^2|t|$  は、信号が広義定常になるための補正項を意味する。実際、各倍音成分に対する自己相関が

$$\begin{aligned} E[x_n(t+\tau)x_n^*(t)] &= E[|c_n|^2] e^{-\sigma_{an}^2\{|t+\tau|+|t|\}} E[e^{\sigma_{an}\{\beta_{an}(t+\tau)+\beta_{an}(t)\}}] E[e^{j\sigma_{bn}\{\beta_{bn}(t+\tau)-\beta_{bn}(t)\}}] \\ &\quad \cdot E[e^{jn\sigma_\omega\{\beta_\omega(t+\tau)-\beta_\omega(t)\}}] e^{jn\bar{\omega}_0\tau} \\ &= \bar{x}_n^2 e^{-\frac{1}{2}\sigma_n^2|\tau|} e^{jn\bar{\omega}_0\tau} = R_{x_n}(\tau) \end{aligned} \quad (3.59)$$

となることから確かめられる。ただし、 $\sigma_n^2 = \sigma_{an}^2 + \sigma_{bn}^2 + n^2\sigma_\omega^2$  とおいた。上式の計算には付録 3.11.1 の関係を用いた。これより、各倍音成分  $x_n(t)$  のスペクトル密度関数は

$$S_{x_n}(\omega) = \bar{x}_n^2 \frac{\sigma_n^2}{\frac{1}{4}\sigma_n^4 + (\omega - n\bar{\omega}_0)^2} \quad (3.60)$$

と表せる。

仮定 8a の下で上式を (3.25) 式に代入し、仮定 1 より  $\frac{\bar{x}_n^2}{\sigma_n^2}\sigma_n^2 \gg \frac{1}{4}\sigma_n^4$  が成り立ち、簡単のため仮定 7a により  $\bar{x}_1^2 = \dots = \bar{x}_N^2 = \bar{x}^2$ ,  $\sigma_{a1}^2 = \dots = \sigma_{aN}^2 = \sigma_a^2$ ,  $\sigma_{b1}^2 = \dots = \sigma_{bN}^2 = \sigma_b^2$  とすると、

$$H(\omega) \approx \sum_{n=1}^N \frac{\frac{\bar{x}^2}{\sigma_n^2}(\sigma_a^2 + \sigma_b^2 + n^2\sigma_\omega^2)}{\frac{\bar{x}^2}{\sigma_n^2}(\sigma_a^2 + \sigma_b^2 + n^2\sigma_\omega^2) + (\omega - n\bar{\omega}_0)^2} \quad (3.61)$$

を得る。

特に振幅変動のみを考慮した場合 ( $\sigma_\omega^2 = 0$ ) は

$$H(\omega) \approx \sum_{n=1}^N H_0(\omega - n\bar{\omega}_0), \quad \text{where } H_0(\omega) = \frac{\frac{\bar{x}^2}{\sigma_n^2}(\sigma_a^2 + \sigma_b^2)}{\frac{\bar{x}^2}{\sigma_n^2}(\sigma_a^2 + \sigma_b^2) + \omega^2} \quad (3.62)$$

となり、性質 5 の定 BW 楕形特性を有し、逆にピッチ変動のみを考慮した場合 ( $\sigma_a^2, \sigma_b^2 = 0$ ) は

$$H(\omega) \approx \sum_{n=1}^N H_1\left(\frac{\omega}{n}\right), \quad \text{where } H_1(\omega) = \frac{\frac{\bar{x}^2}{\sigma_n^2}\sigma_\omega^2}{\frac{\bar{x}^2}{\sigma_n^2}\sigma_\omega^2 + (\omega - \bar{\omega}_0)^2} \quad (3.63)$$

となって、性質 6 の定 Q 楕形特性を有することになる。したがって、もしピッチ変動が振幅変動に比べて十分小さく  $\sigma_a^2 + \sigma_b^2 \gg \sigma_\omega^2$  が成り立つとすれば、低次倍音では相対的に振幅変動の効果が、高次倍音では逆にピッチ変動の効果が支配的になるため、(3.61) 式は性質 7 の定 BW/定 Q 複合楕形特性を有することになる。

### 3.7 因果性を考慮した Wiener フィルタ解

前節の信号モデルに対し、ここでは因果性を考慮した場合の Wiener フィルタを求める。これは、(3.21) 式の定義域を限定した Wiener-Hopf 方程式

$$R_{x_n}(\tau) = \int_0^\infty h_n(s)R_y(\tau - s)ds \quad (\tau \geq 0) \quad (3.64)$$

を解くことに帰着し、付録 3.11.2 のスペクトル分解の方法 [78] により解を求めることができる。

(3.15),(3.51) および (3.60) 式より

$$S_y(\omega) = \bar{x}_n^2 \frac{\sigma_{an}^2 + \sigma_{bn}^2 + n^2 \sigma_\omega^2}{\frac{1}{4}(\sigma_{an}^2 + \sigma_{bn}^2 + n^2 \sigma_\omega^2)^2 + (\omega - n\bar{\omega}_0)^2} + \sigma_v^2 \quad (3.65)$$

と表せ、これを (3.93) 式のように因数分解すると、各項は

$$\begin{aligned} S_{y+}(\omega) &= \frac{\sqrt{\bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4 \sigma_v^2} + j(\omega - n\bar{\omega}_0) \sigma_v}{\frac{1}{2} \sigma_n^2 + j(\omega - n\bar{\omega}_0)} \\ S_{y-}(\omega) &= \frac{\sqrt{\bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4 \sigma_v^2} - j(\omega - n\bar{\omega}_0) \sigma_v}{\frac{1}{2} \sigma_n^2 - j(\omega - n\bar{\omega}_0)} \end{aligned} \quad (3.66)$$

と求められる。ただし、

$$\sigma_n^2 = \sigma_{an}^2 + \sigma_{bn}^2 + n^2 \sigma_\omega^2 \quad (3.67)$$

とおいた。また、

$$\frac{S_{xn}(\omega)}{S_{y-}(\omega)} = \frac{\bar{x}_n^2 \sigma_n^2}{\{\frac{1}{2} \sigma_n^2 + j(\omega - n\bar{\omega}_0)\} \{\sqrt{\bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4 \sigma_v^2} - j(\omega - n\bar{\omega}_0) \sigma_v\}} \quad (3.68)$$

より、

$$\left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_+ = \frac{\bar{x}_n^2 \sigma_n^2}{\sqrt{\bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4 \sigma_v^2} + \frac{1}{2} \sigma_n^2 \sigma_v} \cdot \frac{1}{\frac{1}{2} \sigma_n^2 + j(\omega - n\bar{\omega}_0)} \quad (3.69)$$

となる。したがって、(3.96) 式は

$$H_{n+}(\omega) = \frac{\sqrt{\frac{1}{\sigma_v^2} \bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4} - \frac{1}{2} \sigma_n^2}{\sqrt{\frac{1}{\sigma_v^2} \bar{x}_n^2 \sigma_n^2 + \frac{1}{4} \sigma_n^4} + j(\omega - n\bar{\omega}_0)} \quad (3.70)$$

と表せる。

インパルス応答としては指数減衰の包絡特性をもつことが直ちに導ける。このことは同時に、短時間 Fourier 変換解析において窓関数を指数減衰型とした場合 [79] に対応する。

特に、仮定 1 より  $\frac{\bar{x}_n}{\sigma_v} \gg \frac{1}{2} \sigma_n$  を考慮すれば ( $\bar{x}_n = \sqrt{\bar{x}_n^2}$ )、

$$H_{n+}(\omega) \approx \frac{\frac{\bar{x}_n}{\sigma_v} \sigma_n}{\frac{\bar{x}_n}{\sigma_v} \sigma_n + j(\omega - n\bar{\omega}_0)} \quad (3.71)$$

と近似できる。 $s = j\omega$  とおいて  $s$  平面上で伝達関数  $H_n(s)$  の極を各  $n$  についてプロットしたのが図 3.4 である。特に、振幅変動だけを考慮した場合 ( $\sigma_\omega^2 \rightarrow 0$ ) は

$$H_{n+}(\omega) \approx H_0(\omega - n\bar{\omega}_0), \quad \text{where } H_0(\omega) = \frac{\frac{\bar{x}_n}{\sigma_v} \sqrt{\sigma_a^2 + \sigma_b^2}}{\frac{\bar{x}_n}{\sigma_v} \sqrt{\sigma_a^2 + \sigma_b^2} + j\omega} \quad (3.72)$$

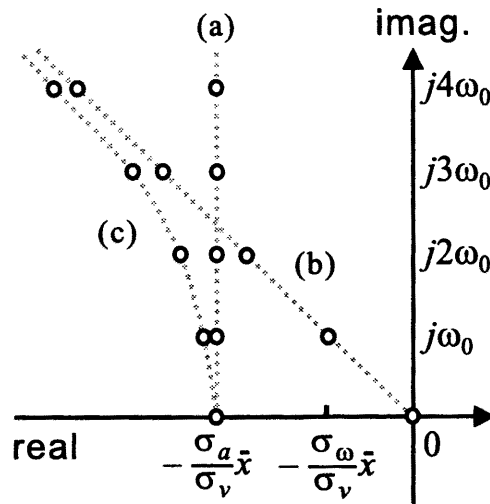


図 3.4 櫛形フィルタの極配置、(a) 定 BW 型、(b) 定 Q 型、(c) 定 BW/ 定 Q 複合型

となり、このとき極は  $s$  平面の左半面上で虚軸と並行な直線上に等間隔に並び (図 3.4 中の破線 (a))、性質 5 の定 BW 櫛形特性を有することになる。これに対し、ピッチ変動だけを考慮した場合 ( $\sigma_a^2, \sigma_b^2 \rightarrow 0$ ) は

$$H_{n+}(\omega) \approx H_1\left(\frac{\omega}{n}\right), \quad \text{where } H_1(\omega) = \frac{\frac{\bar{x}_n}{\sigma_v} \sigma_\omega}{\frac{\bar{x}_n}{\sigma_v} \sigma_\omega + j(\omega - \bar{\omega}_0)} \quad (3.73)$$

となり、 $s$  平面の左半面上で原点を通る斜め直線上に等間隔に極が並ぶことになり (図 3.4 中の破線 (b))、性質 6 の定 Q 櫛形特性を有する。さらに、振幅変動とピッチ変動をともに考慮した場合は、低次倍音側では定 BW 型特性に近く、高次倍音側では定 Q 型特性に近くなる極配置をとるようになり (図 3.4 中の破線 (c))、全体的には性質 7 の定 BW/ 定 Q 複合櫛形特性を有することが前節と同様に結論される。

### 3.8 非直線周波数スケールとの関係

ここで再び、聴覚の周波数分析機能との関係について考察する。

周波数間隔が近接した多数の帯域通過フィルタ群で構成されている聴覚末梢系では、倍音周波数関係にあるフィルタからの同時出力により特定ピッチの周期性を知覚していると考えられる。このときの感覚的な音の高さはピッチ周波数と単純な線形の関係にはなく、図 3.5 に示すように、1KHz 以下では周波数にほぼ比例して音が高く感じるのに対し、1KHz 以上では周波数の対数に比例して音の高さを知覚することが知られている [45]。このような非直線周波数の関係はメル (mel) スケールあるいはバーク (Bark) スケールと呼ばれ、全域

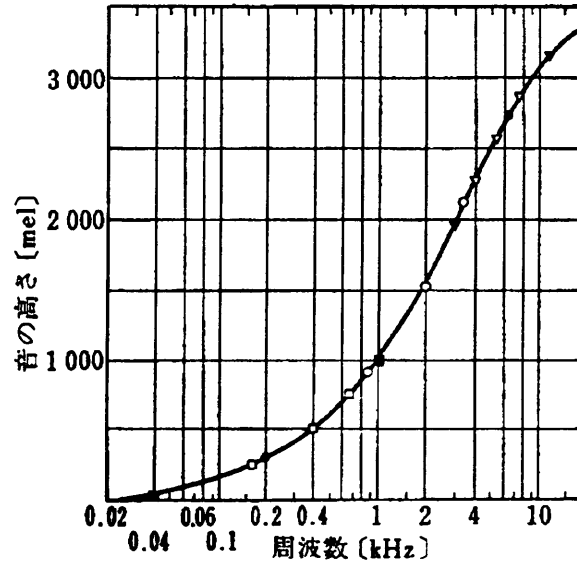


図 3.5 音の周波数と高さの関係 (境 久雄 編著「聴覚と音響心理」 [45] より引用)

通過フィルタ

$$\Psi(z, \alpha) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}} \tag{3.74}$$

$$\Psi(e^{j\omega}, \alpha) = e^{-j\tilde{\omega}(\omega, \alpha)} \tag{3.75}$$

の位相特性

$$\tilde{\omega}(\omega, \alpha) = \omega + 2 \tan^{-1} \frac{\alpha \sin \omega}{1 - \alpha \cos \omega} \tag{3.76}$$

で与えられる周波数スケール変換によって近似できる ( $\alpha$  は非線形性の度合を決めるパラメータ)[74]。この変換を施した音声のスペクトル表現を用いることによって、聴覚の特性を考慮した高性能な音声合成等が実現されている [75]。しかしながら、なぜこのような非直線周波数スケールをとることが聴覚にとって有利であるかについては、これまで明確な説明はなかった。

前節までの結果を用いれば、これは次のように説明できる。「振幅変動+ピッチ変動」信号分離にとって最適なフィルタとは、低周波側で定 BW 特性、高周波側で定 Q 特性をとる帯域通過フィルタからなることがわかっている。このような可変帯域幅のフィルタによって全帯域を一様に覆いつくすには、各帯域通過フィルタの中心周波数は、低周波側で一定間隔、高周波側で周波数に比例して間隔が広がるように配置すべきである (図 3.6 参照)。すなわち、この各フィルタの位置関係がメルスケールになる。

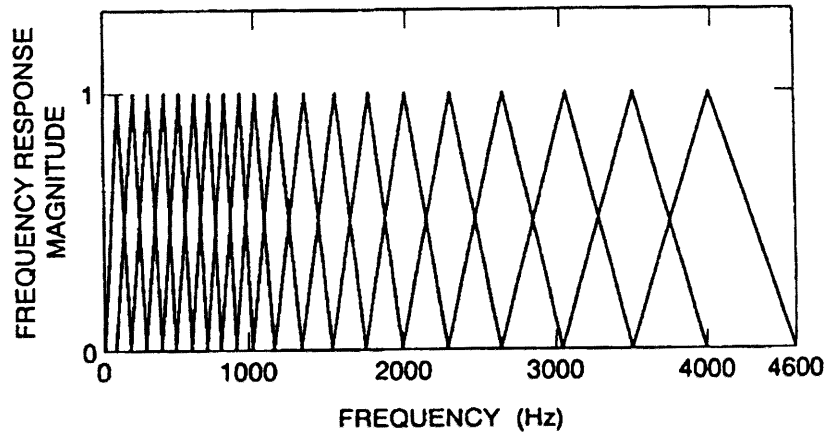


図 3.6 メルスケール上でのフィルタバンクの構成 (著, 古井 訳「音声認識の基礎 (上)」 [76] より引用)

### 3.9 聴覚の周波数特性を実現するフィルタ設計例

「振幅変動+ピッチ変動」信号モデルに基づいて導出された最適フィルタが、実際に聴覚系のもつ非線形な周波数特性によく適合することをフィルタ設計例をあげて検証する。

振幅変動とピッチ変動のモデルとしては、実音声の統計的性質を比較的良好に表す Brown 運動過程による信号モデル 4 を仮定する。このとき Wiener フィルタの厳密解は各倍音に対して (3.70) 式で表せ、また変動が十分小さいと仮定したときの近似解は (3.71) 式となり、これが定 BW/ 定 Q 複合櫛形特性を有することを前節で示した。帯域通過特性をもつこれらの伝達関数に対して適当な閾値以上の利得をとる占有帯域幅を求め、この値をもってフィルタ入力に作用する実効帯域幅と定義する。これは、(3.70) 式に対して

$$\Delta\omega_n^{ex} = 2k \frac{\bar{x}}{\sigma_v} \sqrt{\sigma_a^2 + n^2\sigma_\omega^2} \tag{3.77}$$

(3.71) 式に対して

$$\Delta\omega_n^{ap} = 2k \sqrt{\frac{\bar{x}^2}{\sigma_v^2} (\sigma_a^2 + n^2\sigma_\omega^2) + \frac{1}{4} (\sigma_a^2 + n^2\sigma_\omega^2)^2} \tag{3.78}$$

と表せる。 $k$  は利得の閾値設定から決まる帯域幅基準係数であり、聴覚のフィルタバンクの帯域幅の決定に関しても絶対的に基準を定めることができない以上、この値の選び方についての議論は本質的でないことに注意する。そこで、周波数-帯域幅特性 (図 1.1) において対数スケールをとる縦軸方向でグラフが一致するように  $k$  を決定した。倍音次数  $n$  は整数をとるため、上式によるフィルタバンクはピッチ周波数  $\omega_0$  の間隔で櫛状に配置されることを意味するが、任意のピッチ周波数に適合する櫛形フィルタが常に選択できるような構成を



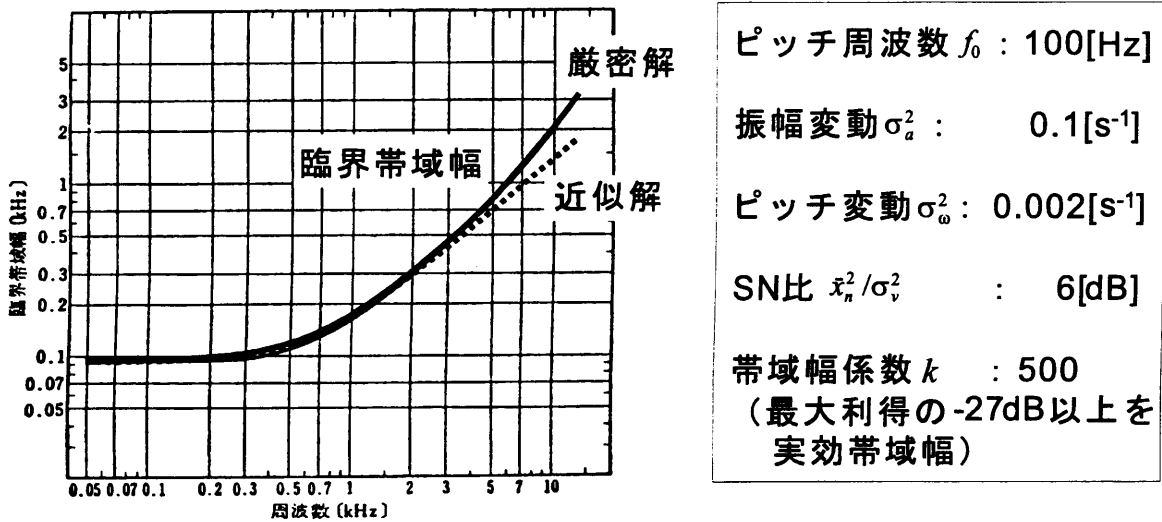


図 3.7 聴覚の周波数特性を実現するフィルタ設計例

とるために、十分密に並んだフィルタバンクが予め用意されていると考える。このときの各帯域幅は、基準となるピッチ周波数を決めた後、 $n$  に実数の連続値を代入することにより等価的に表せる。

図 3.7 は、各パラメータ設定の下で (3.77) 式および (3.78) 式を計算し、図 1.1 に重ね合わせた結果のグラフである。(3.71) 式の近似解については、低周波側では定 BW 特性によく合致しているが、高周波側ではフィルタのもつ定 Q 特性とのずれが生じている。これは、聴覚のもつ周波数 - 帯域幅特性が周波数に比例して単純に帯域幅が広がるのではなく、それよりも速い割合で帯域幅が広がっていくことを意味する。これに対し (3.70) 式の厳密解では、ほぼ全周波数領域にわたってよく合致している。厳密解において高次の帯域幅広がりをもつのは、指数部に Brown 運動過程の揺らぎをもつ信号モデルが有限な信号パワーを保持する上で不可欠な補正項  $-\sigma_n^2 t$  の効果による。すなわち、指数部における Brown 運動の揺らぎは信号波形に対して強い帯域広がり効果をもたらすためといえる。逆に、聴覚の周波数特性によく符合するという結果は、この「振幅変動+ピッチ変動」信号モデルが聴覚にとって最も標準的な分析対象となっている可能性を示唆している。

### 3.10 本章のまとめ

本章では、定常確率過程をとる「振幅変動+ピッチ変動」信号に対する線形最適フィルタである Wiener フィルタの設計を行った。まず、「振幅変動+ピッチ変動」信号のスペクトル特性や線形最適フィルタの伝達特性に見られる一般的性質について考察した (3.2,3.3 節)。

特に、振幅とピッチを一定とした理想周期信号についての基本的性質を述べた(3.4節)。具体例として、振幅変動、ピッチ変動および両者を同時に考慮した場合のそれぞれに対して、各種の確率モデルを導入することによりフィルタ設計を行い、導出されたフィルタの伝達特性の詳細について論じた(3.5, 3.6, 3.7節)。このフィルタの伝達特性が、聴覚系に見られる非直線周波数スケールと密接な関係にあることを述べた(3.8, 3.9節)。

### 3.11 付録

#### 3.11.1 Wiener 過程の諸性質

Wiener 過程 (あるいは Brown 運動過程) は、定常白色 Gauss 雑音

$$E[\xi(t)] = 0, \quad E[\xi(t+\tau)\xi(t)] = \delta(\tau) \quad (3.79)$$

の時間積分として、

$$\beta(t) \equiv \int_0^t \xi(s) ds \quad (3.80)$$

と定義される [80]。このとき、

$$\Delta\beta(t_1, t_2) \equiv \beta(t_1) - \beta(t_2) = \int_{t_1}^{t_2} \xi(s) ds \quad (3.81)$$

は Gauss 過程をとり<sup>†</sup>、その平均および分散は

$$\begin{aligned} E[\Delta\beta(t_1, t_2)] &= \int_{t_1}^{t_2} E[\xi(s)] ds = 0 \\ E[(\Delta\beta(t_1, t_2))^2] &= \int_{t_1}^{t_2} \int_{t_1}^{t_2} E[\xi(s)\xi(u)] ds du = \int_{t_1}^{t_2} \int_{t_1}^{t_2} \delta(s-u) ds du \\ &= |t_1 - t_2| \end{aligned} \quad (3.82)$$

となる。特に、 $t_1 < t_2 < t_3$  あるいは  $t_1 > t_2 > t_3$  のとき、

$$E[\Delta\beta(t_1, t_2)\Delta\beta(t_2, t_3)] = 0 \quad (3.83)$$

が成り立つから、 $\Delta\beta(t_1, t_2)$  と  $\Delta\beta(t_2, t_3)$  は互い統計的に独立となる。

特に、Wiener 過程を指数部にもつ形式

$$\phi_\sigma(t) \equiv e^{\sigma\beta(t)} \quad (3.84)$$

<sup>†</sup>特に、 $\beta(t) = \Delta\beta(t, 0)$  である。

は、幾何 Brown 運動過程と呼ばれる [81]。  $\beta(t)$  が平均 0、分散  $|t|$  の Gauss 分布をとることから、上式の期待値は

$$\begin{aligned} E[\phi_\sigma(t)] &= \frac{1}{\sqrt{2\pi|t|}} \int_{-\infty}^{\infty} e^{\sigma B} e^{-\frac{B^2}{2|t|}} dB \\ &= \frac{1}{\sqrt{2\pi|t|}} \int_{-\infty}^{\infty} e^{-\frac{1}{2|t|}\{B-\sigma|t|\}^2} dB \cdot e^{\frac{1}{2}\sigma^2|t|} \\ &= e^{\frac{1}{2}\sigma^2|t|} \end{aligned} \quad (3.85)$$

と求められる。ただし、  $\beta(t)$  の実現値を  $B$  とした。以上の関係を用いると、  $\phi_\sigma(t)$  に対する自己相関は、

1)  $t_1 > t_2 > 0$  または  $t_1 < t_2 < 0$  のとき、

$$\begin{aligned} E[\phi_\sigma(t_1)\phi_\sigma^*(t_2)] &= E[e^{\sigma\{\beta(t_1)-\beta(t_2)\}}]E[e^{2\sigma\beta(t_2)}] \\ &= e^{\frac{1}{2}\sigma^2|t_1-t_2|}e^{2\sigma^2|t_2|} \end{aligned} \quad (3.86)$$

2)  $t_1 > 0 > t_2$  または  $t_1 < 0 < t_2$  のとき、

$$\begin{aligned} E[\phi_\sigma(t_1)\phi_\sigma^*(t_2)] &= E[e^{\sigma\beta(t_1)}]E[e^{\sigma\beta(t_2)}] \\ &= e^{\frac{1}{2}\sigma^2|t_1|}e^{\frac{1}{2}\sigma^2|t_2|} \end{aligned} \quad (3.87)$$

と表せる。

また、位相変調型の幾何 Brown 運動過程

$$\phi_{j\sigma}(t) \equiv e^{j\sigma\beta(t)} \quad (3.88)$$

に対する自己相関は、

$$\frac{1}{\sqrt{2\pi|\tau|}} e^{-\frac{1}{2}\frac{\tau^2}{|\tau|}} \xleftrightarrow{\mathcal{F.T.}} e^{-\frac{1}{2}\omega^2|\tau|} \quad (3.89)$$

の関係より、

$$E[\phi_{j\sigma}(t_1)\phi_{j\sigma}^*(t_2)] = e^{-\frac{1}{2}\sigma^2|t_1-t_2|} \quad (3.90)$$

と表せる。

### 3.11.2 因果的 Wiener-Hopf 方程式の一般解

(3.64) 式の辺々を差し引いたときの残差

$$\varepsilon_{n-}(\tau) = R_{xn}(\tau) - \int_0^\infty h_n(s)R_y(\tau-s)ds \quad (3.91)$$

は、 $t \geq 0$  で  $\varepsilon_{n-}(\tau) = 0$  となるので、その Fourier 変換

$$E_{n-}(\omega) = S_{xn}(\omega) - H_{n+}(\omega)S_y(\omega) \quad (3.92)$$

は、複素平面<sup>†</sup>の上半面で正則な(すなわち上半面で零点および極をもたない)関数になる。また、 $s < 0$  のとき  $h_n(s) = 0$  であるから、その Fourier 変換  $H_{n+}(\omega)$  は、複素平面の下半面で正則な(すなわち下半面で零点および極をもたない)関数になる。ここで、 $R_y(s)$  の Fourier 変換  $S_y(\omega)$  が

$$S_y(\omega) = S_{y+}(\omega)S_{y-}(\omega) \quad (3.93)$$

のように、複素平面の下半面で正則な関数  $S_{y+}(\omega)$  と上半面で正則な関数  $S_{y-}(\omega)$  の積に因数分解でき、

$$\frac{S_{xn}(\omega)}{S_{y-}(\omega)} = \left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_- + \left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_+ \quad (3.94)$$

のように、上半面と下半面それぞれに正則な関数の和として表せるならば、(3.92)式は

$$\frac{E_{n-}(\omega)}{S_{y-}(\omega)} - \left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_- = \left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_+ - H_{n+}(\omega)S_{y+}(\omega) \quad (3.95)$$

と書き直せる。上式は辺々互いに独立な項をもつため、これが成り立つためには恒等的に0でなければならないから、その右辺部分より

$$H_{n+}(\omega) = \frac{1}{S_{y+}(\omega)} \left[ \frac{S_{xn}(\omega)}{S_{y-}(\omega)} \right]_+ \quad (3.96)$$

を得る。

<sup>†</sup> $\omega$  は通常実数値をとるが、ここでは虚数を含めた複素変数として解釈する。これは、両側 Laplace 変換の  $s$  平面を反時計周りに 90 度回転したものに对应する。

## 第 4 章

# ピッチ変動に対する定 Q 櫛形フィルタの性質と効果

### 4.1 はじめに

擬似周期信号推定にとって最適なフィルタとは、ピッチ周波数の間隔で狭帯域のバンドパスフィルタが並んだ櫛形状の伝達特性をもつフィルタであることが、これまでの議論から明らかになった。本章では、櫛形フィルタのどのような構造が振幅変動やピッチ変動に対して有利に働くのかについて、伝達関数からインパルス応答を求めることによって時間 / 周波数特性の両面から考察する。

櫛形フィルタの特性を議論する上で欠かせないのが櫛形関数の導入である。これは、デルタ関数が一定間隔で左右無限遠方まで並んだ周期関数であり、Fourier 変換が再び櫛形関数になるという重要な性質がある [64]。振幅変動に最適な定 BW 櫛形フィルタの伝達関数は、この櫛形関数と重み関数との畳み込みとして表せるので、逆 Fourier 変換が直ちに実行できてインパルス応答を解析的に求めることができる。

一方、ピッチ変動に最適な定 Q 型の櫛形フィルタに対しては、定 Q 櫛形関数を新たに導入する。この関数は、特定の周期 (櫛歯間隔) をもつ櫛形関数を中心として、その周期をわずかに伸縮させて得られる無数の櫛形関数の足し合わせとして定義される。定 Q 櫛形関数の Fourier 変換は、各櫛形関数を個別に Fourier 変換した後で足し合わせるにより求められ、再び定 Q 櫛形関数になることが示される。この関係は、櫛形関数の Fourier 変換について成り立つ性質を一般化したという点で数学的にも興味深い。定 Q 櫛形フィルタの伝達関数は、定 Q 櫛形関数を適当に重み付けすることにより表せるので、逆 Fourier 変換であるインパルス応答もこの結論から容易に求められる。

これらの結果を結合すれば、定 BW / 定 Q の複合特性をもつ櫛形フィルタについても、インパルス応答と伝達関数の相互変換が解析的に実行でき、時間 / 周波数特性の両面から見通しのよい議論が可能になる。

## 4.2 重み付き最小二乗法に基づく楕形フィルタの導出

目的信号と推定信号の二乗誤差を評価関数とするこれまでの Wiener フィルタや Kalman フィルタの方法論に代わって、本節では、目的信号と観測信号の差の二乗を評価関数とした信号推定を考える。

(2.21) 式の「振幅変動+ピッチ変動」信号モデルに対して、ここでは、時刻  $t = t_0$  の近傍で倍音振幅  $c_n$  およびピッチ周期  $T$  が一定であるとして、

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\frac{2\pi}{T}(t-t_0)} \quad (4.1)$$

を仮定する。ただし、ピッチ周期  $T$  はピッチ推定時において生じる誤差を考慮して、確率密度関数  $P(T)$  のあいまいさをもって既知であるとする。倍音次数  $n$  は形式的に無限までとり、負の倍音成分まで含む実数信号とする。目的信号  $x(t)$  に対する観測信号  $y(t)$

$$y(t) = x(t) + \varepsilon(t) \quad (4.2)$$

から複素倍音振幅の推定値  $\{\hat{c}_n\}$  を求めるために、残差  $\varepsilon(t)$  に対する  $t = t_0$  のまわりでの重み付き 2 乗平均

$$E(t_0; \{c_n\}, T) = \int_{-\infty}^{\infty} \frac{1}{T} w\left(\frac{t-t_0}{T}\right) |\varepsilon(t)|^2 dt \quad (4.3)$$

を定義する。重み  $w(t)$  は、「倍音振幅  $c_n$  やピッチ周期  $T$  が時刻  $t = t_0$  の近傍で一定とみなせる」という仮定の定式化を意味する。ただし、 $\int_{-\infty}^{\infty} w(t) dt = 1$  とし、ピッチ周期に依存せず実効フレーム長が一定となるように、窓関数  $w(t)$  を  $T$  に比例して伸縮させている。さらに、ピッチ推定値による不確定性を考慮して、上式を  $P(T)$  で平均化した

$$E(t_0; \{c_n\}) = \int_{T_1}^{T_2} E(t_0; \{c_n\}, T) P(T) dT \quad (4.4)$$

を最終的な評価関数とする最小化問題を考える。 $P(T)$  は、一般性を失うことなく  $T_b > T > T_a > 0$  でコンパクトな台をもち、 $\int_{T_a}^{T_b} P(T) dt = 1$  とする。

(4.4) 式を最小にする  $c_n$  は、 $\frac{\partial E(t_0; \{c_n\})}{\partial c_n} = 0$  の解として

$$\hat{c}_n = \int_{T_a}^{T_b} P(T) dT \int_{-\infty}^{\infty} \frac{1}{T} w\left(\frac{t}{T}\right) y(t+t_0) e^{-jn\frac{2\pi}{T}t} dt \quad (4.5)$$

と求められる。 $x(t)$  の推定値が  $t = t_0$  のまわりで、

$$\hat{x}(t) = \sum_{n=-\infty}^{\infty} \hat{c}_n e^{jn\frac{2\pi}{T}(t-t_0)} \quad (4.6)$$

と表せるから、特に  $t = t_0$  において (4.5) 式の結果を代入し、付録 4.7.3 の関係を使えば、

$$\begin{aligned}\hat{x}(t_0) &= \sum_{n=-\infty}^{\infty} \hat{c}_n \\ &= \int_{T_a}^{T_b} P(T) dT \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) y(t+t_0) \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{-jn\frac{2\pi}{T}t} dt \\ &= \int_{T_a}^{T_b} P(T) dT \int_{-\infty}^{\infty} w\left(\frac{t}{T}\right) y(t+t_0) \sum_{n=-\infty}^{\infty} \delta(t-nT) dt\end{aligned}$$

となり、さらに付録 4.7.1 で  $c(T) = t - nT$  において変数置換を行い、 $T$  に関する積分を実行すれば、

$$\begin{aligned}\hat{x}(t_0) &= \int_{-\infty}^{\infty} y(t+t_0) dt \int_{T_a}^{T_b} w\left(\frac{t}{T}\right) P(T) \left\{ \delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} \delta\left(T - \frac{t}{n}\right) \right\} dT \\ &= \int_{-\infty}^{\infty} y(t+t_0) \left\{ w(0)\delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{w(n)}{|n|} P\left(\frac{t}{n}\right) \right\} dt \\ &= \int_{-\infty}^{\infty} y(t) \left\{ w(0)\delta(t-t_0) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{w(n)}{|n|} P\left(\frac{t-t_0}{n}\right) \right\} dt\end{aligned}\tag{4.7}$$

を得る。したがって、観測値系列  $\{y(t)\}$  に対して  $\hat{x}(t_0)$  を与えるフィルタ演算

$$\hat{x}(t_0) = \int_{-\infty}^{\infty} h_{QB}(t-t_0) y(t) dt.\tag{4.8}$$

として (4.7) 式を見れば、インパルス応答  $h_{QB}(t)$  は

$$h_{QB}(t) = w(0)\delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{w(n)}{|n|} P\left(\frac{t}{n}\right)\tag{4.9}$$

と表せる。一般に、 $P(T)$  はピッチ周期の真値  $T_0$  の近傍にだけ値をもつインパルス状の分布をもつため、上式はそれを  $T_0$  の整数倍の時刻毎に並べて、窓関数  $w(t)$  で重み付けした楕形状の関数となる。

### 4.3 定BW楕形フィルタの時間 / 周波数特性

特にピッチ周期が唯一の確定値  $T_0$  に決定できるとき、すなわち

$$P(t) \rightarrow \delta(t - T_0)\tag{4.10}$$

の場合、(4.9) 式は

$$h_B(t) = \sum_{n=-\infty}^{\infty} w(n)\delta(t - nT_0)\tag{4.11}$$

となるから、その Fourier 変換である伝達関数は、

$$H_B(\omega) = \sum_{n=-\infty}^{\infty} W(T_0\omega - 2\pi n) \quad (4.12)$$

と表せる。ただし、 $W(\omega)$  は  $w(t)$  の Fourier 変換であり、櫛形関数に対する Fourier 変換の関係 (付録 4.7.3) を用いた。

この結果は、振幅変動を考慮した場合の最適フィルタである定 BW 櫛形フィルタの伝達関数に一致する。すなわち、この場合のインパルス応答は、ピッチ周期間隔で離れたサンプル点の重み付き平均をとるフィルタ演算に対応する。正確には、(3.52) 式などにおいては倍音次数を有限にとってあるので、伝達関数上で矩形窓関数が乗じられ、インパルス応答としては各櫛歯に対して sinc 関数が畳み込まれることになる。

## 4.4 定 Q 櫛形フィルタの時間 / 周波数特性

### 4.4.1 定 Q 櫛形関数とその Fourier 変換

(4.9) 式において重みを一定  $w(t) = 1$  とし、確率密度関数  $P(T)$  をコンパクト台をもつ任意関数に置き換えたものを、あらためて定 Q 櫛形関数として以下に定義する。

**定義 4:** 次式で定義される関数を定 Q 櫛形関数と呼ぶ。

$$h_Q(t) = A\delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} h_0\left(\frac{t}{n}\right) \quad \text{where } A = \int_{t_a}^{t_b} h_0(t) dt \quad (4.13)$$

但し、核関数  $h_0(t)$  は  $t < t_a$ ,  $t > t_b$  ( $0 < t_a < t_b$ ) で  $h_0(t) = 0$  となるようなコンパクトな台をもつ区分的に滑らかな任意の関数であり、不連続点で  $h_0(t) = \frac{1}{2}\{h_0(t-0) + h_0(t+0)\}$  (左側極限と右側極限の平均値) をとるものとする。

この定 Q 櫛形関数は、異なる周期をもつ複数の櫛形関数の重ね合わせとして表せることが次の補題で示される。

**補題:** 定 Q 櫛形関数の (4.13) 式は、

$$h_Q(t) = \int_{t_a}^{t_b} h_0(T) \sum_{n=-\infty}^{\infty} \delta(t - nT) dT \quad (4.14)$$

と書き直せる。

(証明) (4.14) 式から (4.13) 式を導く。

$$(4.14) \text{ 式右辺} = \int_{t_a}^{t_b} h_0(T) dT \cdot \delta(t) + \int_{t_a}^{t_b} h_0(T) \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \delta(t - nT) dT$$



において、付録 4.7.1 で  $c(T) = t - nT$  と変数置換することにより、

$$\begin{aligned}
 (4.14) \text{ 式右辺} &= A\delta(t) + \int_{t_a}^{t_b} h_0(T) \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} \delta(T - \frac{t}{n}) dT \\
 &= A\delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \int_{t_a}^{t_b} h_0(T) \frac{1}{|n|} \delta(T - \frac{t}{n}) dT \\
 &= (4.13) \text{ 式右辺}
 \end{aligned}$$

が示せる。上式中で積分と総和演算の交換が可能なのは、 $h_0(t)$  がコンパクト台をもつため  $t \neq 0$  のいたるところで (4.13) 式が有界になるからである。またデルタ関数との内積によって、 $h_0(\frac{t}{n})$  の不連続点で  $h_0(\frac{t}{n}) = \frac{1}{2}\{h_0(\frac{t}{n} - 0) + h_0(\frac{t}{n} + 0)\}$  となることに注意する [88]。

Q.E.D.

式展開の意味をイメージ化したのが図 4.1 である。(4.14) 式が周期の異なる複数の櫛形関数の重ね合わせで表せることを、積分を離散近似して示したのが同図 (b) であり、刻み幅を無限小にした極限で (4.13) 式になることを同図 (a) が表している。このとき、各櫛形関数 (同図 (b)) に対する Fourier 変換が再び櫛形関数 (同図 (c)) になることに着目すれば、刻み幅を無限小にしたときのそれらの足し合わせ (同図 (d)) で定 Q 櫛形関数の Fourier 変換が表せると考えられる。このことを正確に示したのが次の定理である。

**定理:** 定 Q 櫛形関数 (4.13) 式の Fourier 変換

$$H_Q(\omega) = \mathcal{F}\{h_Q(t)\} = \int_{-\infty}^{\infty} h_Q(t) e^{-j\omega t} dt \quad (4.15)$$

は、

$$H_Q(\omega) = B\delta(\omega) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{2\pi}{|\omega|} h_0(n \frac{2\pi}{\omega}) \quad \text{where } B = 2\pi \int_{t_a}^{t_b} \frac{h_0(t)}{t} dt \quad (4.16)$$

で与えられ、再び定 Q 櫛形関数になる。

(証明) (4.14) 式の Fourier 変換

$$H_Q(\omega) = \int_{t_a}^{t_b} h_0(T) \mathcal{F}\left\{ \sum_{n=-\infty}^{\infty} \delta(t - nT) \right\} dT$$

において、付録 4.7.3 の関係を用いると、

$$\begin{aligned}
 H_Q(\omega) &= \int_{t_a}^{t_b} h_0(T) \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta(\omega - n \frac{2\pi}{T}) dT \\
 &= \int_{t_a}^{t_b} h_0(T) \frac{2\pi}{T} dT \cdot \delta(\omega) + \int_{t_a}^{t_b} h_0(T) \frac{2\pi}{T} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \delta(\omega - n \frac{2\pi}{T}) dT \quad (4.17)
 \end{aligned}$$

を得る。

$\omega \neq 0$  のときは第2項だけを考えればよく、付録4.7.1で  $c(T) = \omega - n\frac{2\pi}{T}$  と変数置換することにより、

$$\begin{aligned} H_Q(\omega) &= \int_{t_a}^{t_b} h_0(T) \frac{2\pi}{T} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{2\pi|n|}{\omega^2} \delta(T - n\frac{2\pi}{\omega}) dT \\ &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \int_{t_a}^{t_b} h_0(T) \frac{4\pi^2|n|}{T\omega^2} \delta(T - n\frac{2\pi}{\omega}) dT \\ &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{2\pi}{|\omega|} h_0(n\frac{2\pi}{\omega}) \end{aligned} \quad (4.18)$$

と整理できる。但し、 $h_0(n\frac{2\pi}{\omega})$  の不連続点で  $h_0(n\frac{2\pi}{\omega}) = \frac{1}{2}\{h_0(n\frac{2\pi}{\omega} - 0) + h_0(n\frac{2\pi}{\omega} + 0)\}$  となる。上式と同じ結論は、Poisson の和公式 [88] によっても導かれる (付録4.7.2参照)。したがって  $-\infty < \omega < \infty$  では、(4.17) 式の第1項を含めることで (4.16) 式が成り立つ。

ここで、

$$H_0(\omega) = \frac{2\pi}{|\omega|} h_0\left(\frac{2\pi}{\omega}\right)$$

とおけば、 $H_0(\omega)$  は  $\omega < \frac{2\pi}{t_b}$ ,  $\omega > \frac{2\pi}{t_a}$  で  $H_0(\omega) = 0$  となるコンパクトな台をもった区分的に滑らかな関数になる。これを使って (4.16) 式が

$$H_Q(\omega) = B\delta(\omega) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} H_0\left(\frac{\omega}{n}\right)$$

と表せ、

$$\int_{\frac{2\pi}{t_b}}^{\frac{2\pi}{t_a}} H_0(\omega) d\omega = 2\pi \int_{t_a}^{t_b} \frac{h_0(t)}{t} dt = B$$

が成り立つので、 $H_Q(\omega) \rightarrow h_Q(t)$ ,  $H_0(\omega) \rightarrow h_0(t)$ ,  $B \rightarrow A$ ,  $\omega \rightarrow t$ ,  $\frac{2\pi}{t_b} \rightarrow t_a$ ,  $\frac{2\pi}{t_a} \rightarrow t_b$  と読み直せば、(4.16) 式は定義4に基づく定Q楕形関数になる。

Q.E.D.

定理により定Q楕形関数は、核関数の変数置換だけでFourier変換が求められ、その結果も再び定Q楕形関数になることが示された。この意味で、定Q楕形関数は楕形関数を一般化したものであるといえる。事実、 $h_0(t) = \delta(t - T_0)$  とおけば、定理は楕形関数で良く知られたフーリエ変換の関係 (付録4.7.3) そのものになる。定Q楕形関数の名は、時間および周波数領域いずれにおいても楕形状の分布をとり、伝達特性として見た場合、倍音次数に比例して帯域幅が伸長する構造をもつことによる。定Q楕形関数の各楕歯に任意の重み付けを許容した分布として伝達関数が表せる場合、そのフィルタを定Q楕形フィルタと呼ぶこととする。これは、3章の性質6の言い換えにすぎない。

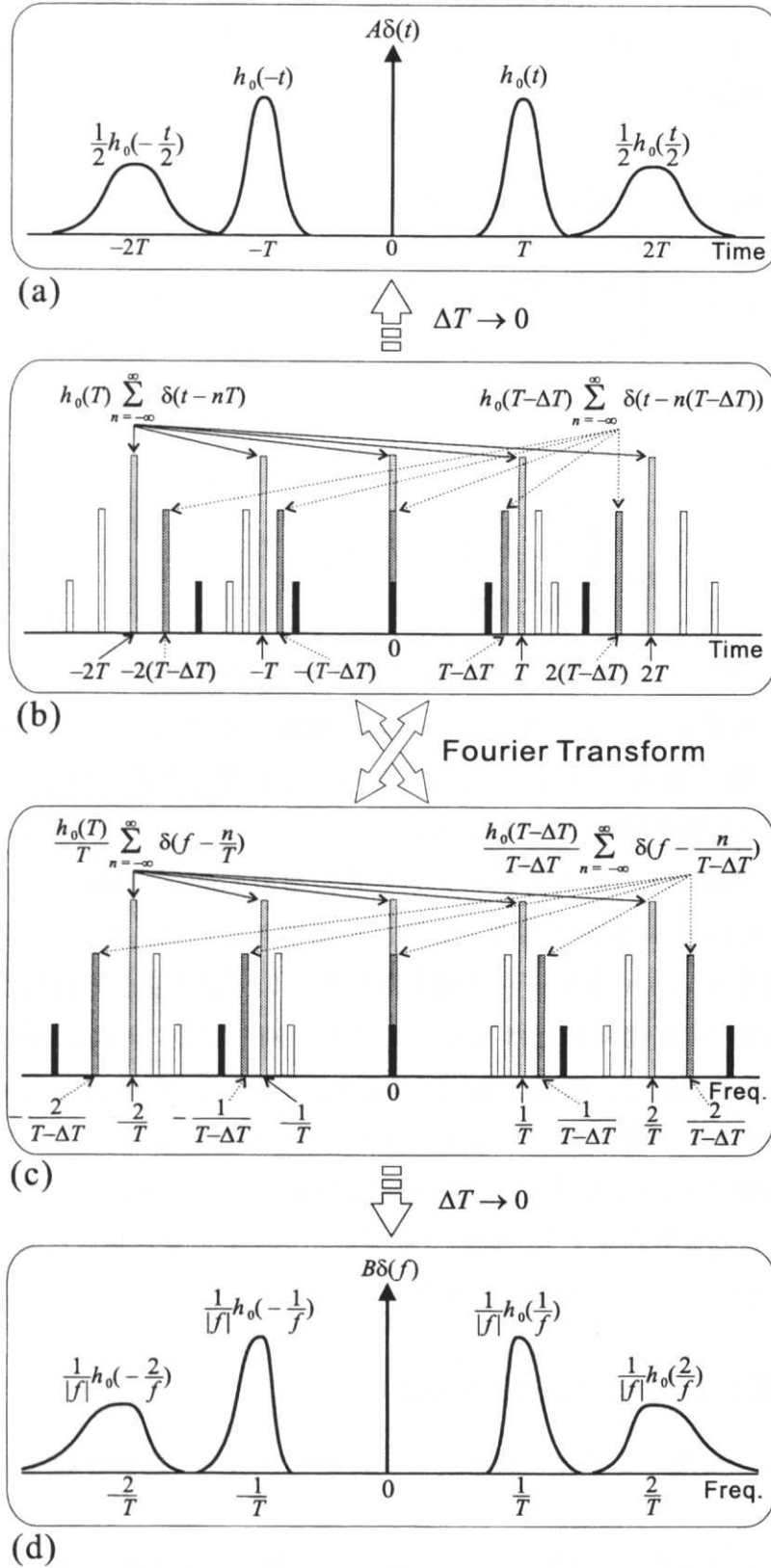


図 4.1 定 Q 楕円関数のフーリエ変換

#### 4.4.2 定 Q 楕形関数の具体例

定 Q 楕形関数の Fourier 変換が (4.16) 式によって正しく求まることを、離散 Fourier 変換との比較により検証する。

核関数  $h_0(t)$  として、

(例 1) Gauss 関数の両端を切り捨てた分布

$$h_0(t) = \begin{cases} e^{-\frac{(t-T_0)^2}{2\sigma_T^2}}, & \text{for } T_0 - T_g < t < T_0 + T_g \\ 0, & \text{other} \end{cases}$$

や、

(例 2) 矩形分布

$$h_0(t) = \begin{cases} 1, & \text{for } T_0 - T_r < t < T_0 + T_r \\ 0, & \text{other} \end{cases}$$

を考える。このとき、(4.13) 式によって作られる定 Q 楕形関数が各々図 4.2, 4.3(a) である。但し、ピッチ周期  $T_0 = 1$  に対して、Gauss 関数の標準偏差  $\sigma_T = 0.01$ 、切捨て幅  $T_g = 3\sigma_T$ 、矩形関数の幅  $T_r = 0.02$  とした。各々の定 Q 楕形関数に対する Fourier 変換を (4.16) 式により求めた結果が同図 (b) である。

比較のために、同図 (a) の離散化データから直接離散 Fourier 変換して求めた結果を同図 (c) に示す。離散 Fourier 変換の対象として、1 周期あたり 256 点、正負各軸に 128 周期分の核関数からなる計 65536 点のデータ列を用いた。いずれも正負対称なグラフとなるため、図には正軸方向の 16 周期分だけを示してある。例 1 では、各点の振幅値は最大でも 0.01% の誤差が生じるだけで、両者は見かけ上完全に一致し、例 2 においても、核関数の不連続境界で生じる Gibbs 振動を除いてほぼ一致することがわかる。

図 4.2 や図 4.3 を見てわかるように、定 Q 楕形関数は、時間領域においては周期倍数に比例して歯幅が広がる楕形分布をもち、周波数領域においても倍音次数に比例して帯域が広がる定 Q 型の楕形特性をもつことが確認できる。

#### 4.4.3 ピッチ誤差による波形 / スペクトル歪み

定 Q 楕形フィルタがピッチ誤差に対して有利に働くことを、定 BW 楕形フィルタとの比較により検証する。

母音音声「あ」の 1 周期分 (図 4.4, 4.5, 4.6(a) の点線部分) を切りだして周期化したものを入力信号として、各楕形フィルタによる応答を求めた。定 Q 楕形フィルタは (4.13) 式において核関数を (例 1) の Gauss 関数に選び、定 BW 楕形フィルタは (4.11) 式において重

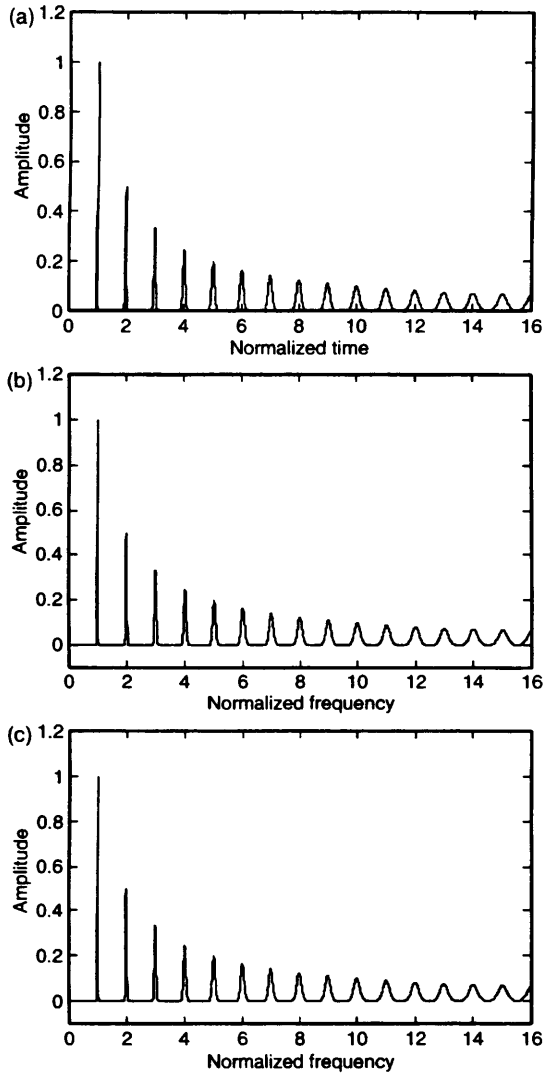


図 4.2 ガウス核をもつ定 Q 楕形関数の例とその周波数特性、(a) 時間特性、(b) 定理に基づく理論周波数特性、(c) 離散フーリエ変換による計算結果

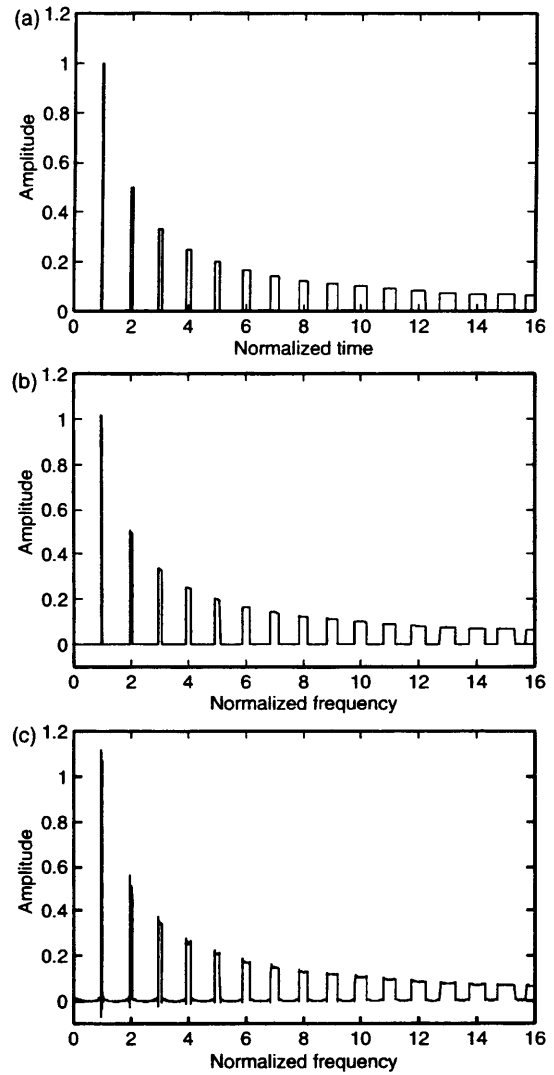


図 4.3 矩形核をもつ定 Q 楕形関数の例とその周波数特性、(a) 時間特性、(b) 定理に基づく理論周波数特性、(c) 離散フーリエ変換による計算結果

み関数を一定値とした。ただし、いずれも櫛歯の数を正負の時刻で 3 本ずつに打ち切ったものを用いた。フィルタの出力波形の 1 周期分をプロットしたのが図 4.4, 4.5, 4.6(a) の実線部分である。そのうち図 4.4(a) は、ピッチパラメータを真のピッチ周期 9.2ms に一致させたときの定 Q 櫛形フィルタの出力波形であり、図 4.5, 4.6(a) はピッチパラメータを  $T_0 = 9.5\text{ms}$  として故意に誤差を与えたときの各々定 Q、定 BW 櫛形フィルタの出力波形である。Gauss 核関数の標準偏差は  $\sigma_T = 0.3\text{ms}$  に設定した。入力信号とフィルタ出力に対するスペクトル分布が各々同図 (b) の点線と実線に対応する。破線部分は離散スペクトルの包絡特性を決めるフィルタゲインを表す。ピッチ誤差がないときの定 BW 櫛形フィルタでは、全く歪みが生じず、点線と実線が一致したため図は省略した。

定 Q 櫛形フィルタでは、図 4.2(b) のように、倍音次数に比例して通過帯域が広がるのに対して利得は反比例して下がる特性をもつことから、ピッチ誤差がないときでも高域側で減衰が生じるという欠点はあるが (図 4.4(b))、その代わりピッチ誤差がある場合でも、原音の波形およびスペクトル概形を比較的良好に保存していることがわかる (図 4.5(b))。それに対し定 BW 櫛形フィルタでは、ピッチ誤差がないときは歪みは生じないが、ピッチ誤差が少しでもあると波形に位相歪みが生じ、原音にはない凹凸がスペクトル上に現れる (図 4.6(b))。実際聴感上も、定 Q 型は元の「あ」に近い音声として聞き取れるが、定 BW 型では原音の特徴を残さない機械的な音になることが確認された。

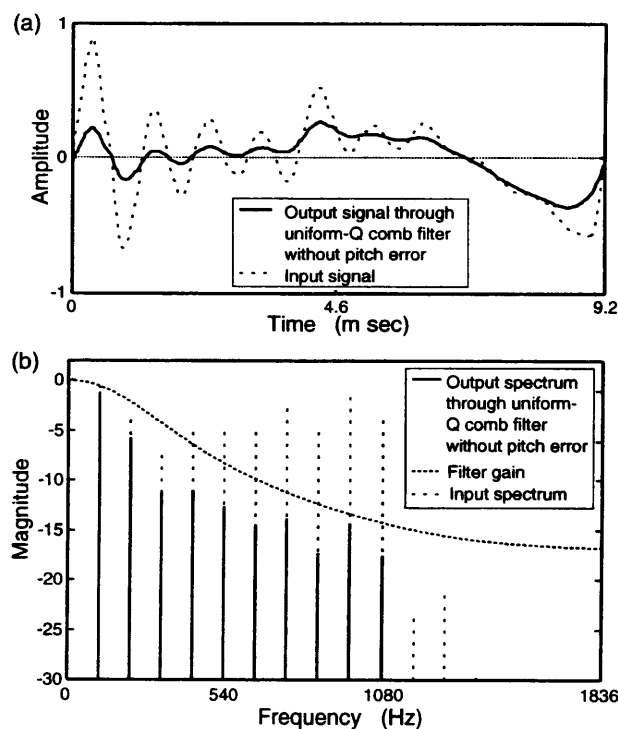


図 4.4 ピッチ誤差がないときの定 Q 櫛形フィルタの出力波形 (a)/ スペクトル歪み (b)

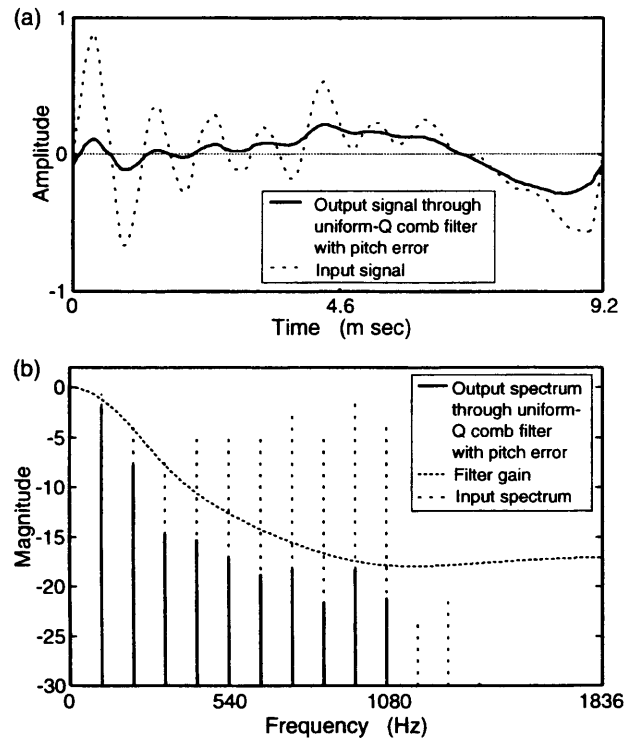


図 4.5 ピッチ誤差があるときの定 Q 櫛形フィルタの出力波形 (a)/ スペクトル歪み (b)

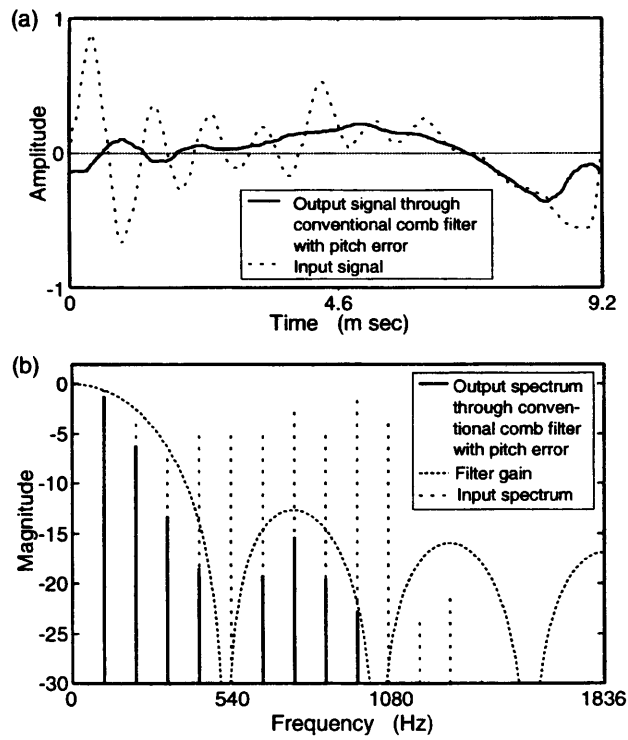


図 4.6 ピッチ誤差があるときの定 BW 櫛形フィルタの出力波形 (a)/ スペクトル歪み (b)

### 4.5 各種の楕形フィルタについて

再び(4.9)式に戻って、これまでに導かれた各種の楕形フィルタとの関係について論じる。

$P(T)$  は  $T_0$  の近傍に分布が集中するインパルス状の関数であるため、(4.9)式は

$$h_{QB}(t) \approx w\left(\frac{t}{T_0}\right) \left\{ \delta(t) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} P\left(\frac{t}{n}\right) \right\} \quad (4.19)$$

のように、窓関数と定Q楕形関数の積として近似できる。したがって、伝達関数は

$$H_{QB}(\omega) \approx T_0 W(T_0 \omega) * \left\{ B \delta(\omega) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{2\pi}{|\omega|} P\left(n \frac{2\pi}{\omega}\right) \right\}$$

$$\text{where } B = 2\pi \int_{T_a}^{T_b} \frac{P(T)}{T} dT \quad (4.20)$$

と表せ、定Q楕形関数に窓関数のスペクトル特性を畳み込んだものになる。低次倍音では、定Q楕形関数はデルタ関数とみなせるので、窓関数のスペクトル広がりによる定BW楕形特性が支配的になるのに対して、高次倍音では、逆に窓関数のスペクトル特性の方がデルタ関数とみなせるため定Q楕形特性が支配的になる。これは、第3章の性質7の結果に一致する。

また、定Q楕形関数は楕歯の広がりには反比例して高さが小さくなる構造をもっているのに対し、第3章での定Q楕形フィルタは高さ（利得）が一定であったことに注意する。これは、定Q楕形関数

$$H_Q(\omega) = A \delta(\omega) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} H_1\left(\frac{\omega}{n}\right) \quad \text{where } A = \int_{\omega_a}^{\omega_b} H_1(\omega) d\omega \quad (4.21)$$

を使って、(3.54)式が

$$H(\omega) \approx W\left(\frac{\omega}{\omega_0}\right) H_Q(\omega) \quad \text{where } W(n) = \begin{cases} n, & \frac{1}{2} < n < N + \frac{1}{2} \\ 0, & \text{other} \end{cases} \quad (4.22)$$

と表されるので<sup>†</sup>、伝達関数上で重み関数  $W(n)$  が乗じられるかどうかの違いになる。したがってインパルス応答としては、窓関数  $w(t)$  を定Q楕形関数に畳み込んだものになり、これはちょうど上の議論において周波数軸と時間軸を入れ換えたことに対応する。

<sup>†</sup>ただし、帯域通過フィルタ  $H_1(\omega)$  は、 $\omega = \omega_0$  を中心周波数として、帯域が  $\omega_a < \omega < \omega_b$  に帯域制限されたインパルス状の分布をもつものとする。



## 4.6 本章のまとめ

本章では、擬似周期信号推定に対する最適フィルタとして導かれた楕形フィルタのインパルス応答と伝達関数の間に成り立つ関係について考察した。まず議論の出発点として、楕形フィルタの時間領域における一般形式を、観測信号を「振幅変動+ピッチ変動」信号モデルに最小二乗フィッティングさせたときの複素振幅パラメータの推定値から導出した(4.2節)。このインパルス応答の形式は、定 BW/ 定 Q の複合特性をもつ楕形フィルタに対応する。特に、ピッチ周波数が唯一の確定値をとる場合には、定 BW 楕形フィルタが導かれることを示した(4.3節)。ピッチ周波数の不確定性を考慮した場合のインパルス応答から定 Q 楕形関数を定義し、その Fourier 変換に関して成り立つ性質を定理の形で与えた(4.4節)。定理の正しさは数値計算との比較により確かめられた(4.4.2項)。特に定 Q 楕形フィルタを用いれば、ピッチ誤差によって生じる歪みが小さく抑えられることを、実例を通して時間 / 周波数特性の両面から検証した(4.4.3項)。また、各種の楕形フィルタとの関係について論じた(4.5節)。

## 4.7 付録

### 4.7.1 デルタ関数の変数置換

$c(t)$  が  $t_a < t < t_b$  で単調増加もしくは単調減少の関数であるならば、 $c(t) = 0$  の根  $t = t_0$  を使って、

$$\delta(c(t)) = \frac{1}{|c'(t_0)|} \delta(t - t_0) \quad (4.23)$$

と変数置換できる。これは、任意の滑らかな関数  $\varphi(t)$  と左辺  $\delta(c(t))$  との内積が、 $s = c(t)$  が単調増加すなわち  $c'(t) > 0$  のとき、

$$\begin{aligned} \int_{t_a}^{t_b} \varphi(t) \delta(c(t)) dt &= \int_{c(t_a)}^{c(t_b)} \varphi(c^{-1}(s)) \frac{1}{c'(c^{-1}(s))} \delta(s) ds \\ &= \varphi(t_0) \frac{1}{c'(t_0)} \end{aligned}$$

となり、単調減少すなわち  $c'(t) < 0$  のとき

$$\begin{aligned} \int_{t_a}^{t_b} \varphi(t) \delta(c(t)) dt &= - \int_{c(t_b)}^{c(t_a)} \varphi(c^{-1}(s)) \frac{1}{c'(c^{-1}(s))} \delta(s) ds \\ &= -\varphi(t_0) \frac{1}{c'(t_0)} \end{aligned}$$

となつて、結果として

$$\int_{t_a}^{t_b} \varphi(t) \delta(c(t)) dt = \varphi(t_0) \frac{1}{|c'(t_0)|}$$

と表せ、また  $\varphi(t)$  と (4.23) 式の右辺との内積が、

$$\int_{t_a}^{t_b} \varphi(t) \frac{1}{|c'(t_0)|} \delta(t - t_0) dt = \varphi(t_0) \frac{1}{|c'(t_0)|}$$

となって両者が等しくなるため、超関数として成り立つことを意味する。但し、置換  $s = c(t)$  に対する逆関数を  $t = c^{-1}(s)$ 、 $c(t)$  の  $t$  に関する微分を  $c'(t)$  と置いた。

#### 4.7.2 Poisson の和公式による導出

(4.18) 式を Poisson の和公式により導く。

以下のような  $h_0(t)$  を周期化した関数を考える。

$$g(t) = \sum_{n=-\infty}^{\infty} \frac{2\pi}{|\omega|} h_0\left(t + n \frac{2\pi}{\omega}\right) \quad (4.24)$$

$g(t)$  は、定義 4 の  $h_0(t)$  の仮定により区分的に滑らかな周期  $\frac{2\pi}{|\omega|}$  ( $\omega \neq 0$ ) をもつ関数になるので、

$$g(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega t} \quad (4.25)$$

のように Fourier 級数展開できる。但し、不連続点では  $g(t) = \frac{1}{2}\{g(t-0) + g(t+0)\}$  となる。Fourier 係数  $c_n$  が

$$\begin{aligned} c_n &= \frac{|\omega|}{2\pi} \int_0^{\frac{2\pi}{|\omega|}} g(t) e^{-jn\omega t} dt \\ &= \sum_{k=-\infty}^{\infty} \int_0^{\frac{2\pi}{|\omega|}} h_0\left(t + k \frac{2\pi}{\omega}\right) e^{-jn\omega t} dt \\ &= \sum_{k=-\infty}^{\infty} \int_{k \frac{2\pi}{|\omega|}}^{(k+1) \frac{2\pi}{|\omega|}} h_0(t) e^{-jn\omega t} dt \\ &= \int_{-\infty}^{\infty} h_0(t) e^{-jn\omega t} dt \\ &= \hat{h}_0(n\omega) \end{aligned} \quad (4.26)$$

と表せることを用いて、(4.24) 式と (4.25) 式を等置し  $t = 0$  とすれば、Poisson の和公式 [88]

$$\sum_{n=-\infty}^{\infty} \frac{2\pi}{|\omega|} h_0\left(n \frac{2\pi}{\omega}\right) = \sum_{n=-\infty}^{\infty} \hat{h}_0(n\omega) \quad (4.27)$$

が導ける。但し  $\hat{h}_0(\omega) = \mathcal{F}\{h_0(t)\}$  と置いた。

一方、(4.13) 式の Fourier 変換は、

$$H_Q(\omega) = \mathcal{F}\{h_Q(t)\}$$

$$\begin{aligned}
&= A\mathcal{F}\{\delta(t)\} + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{|n|} \mathcal{F}\left\{h_0\left(\frac{t}{n}\right)\right\} \\
&= A + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \hat{h}_0(n\omega) \\
&= \sum_{n=-\infty}^{\infty} \hat{h}_0(n\omega)
\end{aligned}$$

と表せる。但し  $A = \hat{h}_0(0)$  となることを用いた。したがって、(4.27) 式の左辺で  $n = 0$  の項が  $h_0(0) = 0$  になることに注意すれば、上式より (4.18) 式が導かれる。

#### 4.7.3 楕円関数の Fourier 変換

(4.24) 式で  $h_0(t) = \frac{1}{T}\delta(t)$ ,  $\omega = \frac{2\pi}{T}$  と置いて、

$$g(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$$

とすれば、(4.25),(4.26) 式より

$$g(t) = \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{jn\frac{2\pi}{T}t}$$

と表せる。したがって、この Fourier 変換は

$$\begin{aligned}
G(\omega) &= \mathcal{F}\{g(t)\} \\
&= \frac{1}{T} \sum_{n=-\infty}^{\infty} \mathcal{F}\{e^{jn\frac{2\pi}{T}t}\} \\
&= \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta\left(\omega - n\frac{2\pi}{T}\right)
\end{aligned} \tag{4.28}$$

となる。以上は、 $h_0(t)$  が区分的に滑らかな関数ではないので形式的な証明でしかないが、「佐藤の超関数」を用いれば厳密に証明される [88]。

## 第 5 章

# 多重擬似周期信号に対するピッチ推定と個別分離アルゴリズム

### 5.1 はじめに

擬似周期信号推定の問題は、結局、楕形フィルタの設計問題に帰結することがこれまでの議論で明らかになった。本章では、この楕形フィルタを構成する際に不可欠なピッチ周波数の推定法として、特に、音声や楽音などの複数の擬似周期音が含まれるような現実の音環境においても適用可能な方法を考える。

一般に知られているピッチ推定法には、信号波形をピークピッキングしてその間隔を計る方法、適当な時間窓をかけて自己相関を計算しそのピーク間隔を計る方法 [116] や、パワースペクトルの対数を逆 Fourier 変換することによりスペクトルの包絡特性と微細構造を分離して後者を使う方法 (ケプストラム法) [117] などがある。また最近でも、狭帯域フィルタの瞬時周波数を用いる方法 [118] が提案されているが、いずれも単一の擬似周期信号を対象としていて、複数のピッチを含む混合音への対応はなされていない。

混合音に対する時間周波数分布が複数の調波スペクトル構造の重なり合いとして表せることに着目し、任意の周波数を基準としてその整数倍に位置するすべてのスペクトル値を積算した分布<sup>†</sup>をピッチ推定に用いることを考える。すなわち、混合音中の各ピッチに一致した周波数で分布が極大によるという性質を利用すれば、その極大点の位置から逆に各ピッチ候補を求めることができる [121]。

ピッチ尤度関数は時間周波数分布を介して各時刻において求められるが、その極大点の軌跡は擾乱や干渉等の影響により断続的なものになる。そのため、単なるピーク追跡では混在する各ピッチ軌跡の時間的つながりを追跡することは難しく、ピッチ軌跡の連続性に従うも

---

<sup>†</sup>これは高調波積スペクトラム [119] と呼ばれ、その対数をとったものは倍音和スペクトラム [120] として定義されている。本論文ではこれをピッチ尤度関数と呼ぶ。

のだけをピッチ候補として選別し、それを補間追跡できるような仕組みの実現が不可欠になる。ピッチ軌跡の出現・消失が起こり得る不特定数の事象を同時並行的に扱わなければならないという要請から、ピッチ軌跡の時間的統合を多峰的なピッチ尤度関数の遷移としてノンパラメトリックに実現することを考える。これには、Non-Parametric Kalman フィルタの方法論 [69] を導入する。

時間周波数領域での楕形フィルタの演算は、ピッチ周波数の整数倍に位置する複素振幅値を加算することと等価になる。その際の基準となるピッチ周波数は、離散化されたピッチ尤度関数に基づいて推定されるため、誤差は避けられずフィルタ性能の劣化を招く原因となる。この影響を最小限に抑えるためには定 Q 楕形フィルタの利用が有効であり、これは Wavelet 変換による時間 - 対数周波数領域において等価的に実現できる。そこで、時変スペクトル構造の抽出に始まって、ピッチ推定および楕形フィルタ演算までのすべてを Wavelet 空間において行うことを考える。

## 5.2 Wavelet 空間とピッチ尤度関数

### 5.2.1 ストリーム分離問題とは

ここでも楽音や音声 (特に母音部) の記述には、(2.21) 式による「振幅変動 + ピッチ変動」信号モデル

$$x(t) = \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} \quad (5.1)$$

を考える。ピッチ周波数  $\omega_0(t)$  や各倍音振幅  $c_n(t)$  の緩やかな変化がたどる軌跡をストリーム (音脈) と呼ぶことにする。音源が複数ある場合、音源毎に独立したピッチ軌跡  $\omega_0(t)$  と振幅軌跡  $c_n(t)$  をもつストリームが生成され、それらが重ね合わさって混合音となる。したがって、本論文で論じるストリーム分離の問題とは、混合音から適切な基準に従って唯一のストリームを定めるとともに、それに対応する  $\omega_0(t), c_n(t)$  をいかに推定するかという問題になる。 $\omega_0(t), c_n(t)$  が推定できれば、(5.1) 式に従って対応するストリーム波形の再構成も可能である。

### 5.2.2 時間 - 対数周波数分析

与えられた混合音から上記の時変スペクトル構造を抽出するために、Wavelet 変換

$$W(\log \omega, t) = \int_{-\infty}^{\infty} \omega \psi^*(\omega(u-t)) f(u) du \quad (5.2)$$

を用いる。分析 Wavelet  $\psi(t)$  は、局在性に優れた Gabor 関数

$$\omega\psi(\omega t) \equiv \frac{B\omega}{\sqrt{2\pi}} e^{-\frac{(B\omega)^2 t^2}{2}} e^{j\omega t} \xleftrightarrow{\mathcal{F.T.}} \Psi\left(\frac{\rho}{\omega}\right) \equiv e^{-\frac{(\rho-\omega)^2}{2(B\omega)^2}} \quad (5.3)$$

であり、 $\Psi(\rho)$  はその Fourier 変換、 $\rho$  は周波数変数、 $\omega$  は中心周波数、 $B$  は Q 値を決める係数である。こうすると、(5.2) 式は具体的に

$$W(\log \omega, t) = \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} \Psi\left(\frac{n\omega_0(t)}{\omega}\right) + \varepsilon(\omega, t) \quad (5.4)$$

と表される(付録 5.8.1 参照)。 $c_n(t)$  や  $\omega_0(t)$  が十分緩やかに変化するという仮定から、 $\varepsilon(\omega, t)$  の影響は第 1 項に比べて十分小さく、また  $\Psi\left(\frac{n\omega_0(t)}{\omega}\right)$  は、 $\omega = n\omega_0(t)$  で最大値 1、それ以外で急激に零に落ちる関数であるから、ピッチ間隔に対して Q 値を十分鋭く選ぶことにより (Q 値と  $B$  の決定法は後述)、 $\omega = n\omega_0(t)$  の近傍で

$$W(\log \omega, t) \approx c_n(t) e^{jn \int_0^t \omega_0(s) ds} \Psi\left(\frac{n\omega_0(t)}{\omega}\right) \quad (5.5)$$

と表せ、極大値をとる各倍音周波数  $\omega = n\omega_0(t)$  で

$$W(\log n\omega_0(t), t) \approx c_n(t) e^{jn \int_0^t \omega_0(s) ds} \quad (5.6)$$

となる [122]。

この式は、ピッチ周波数  $\omega_0(t)$  が決められると、Wavelet 複素振幅  $W(\log n\omega_0(t), t)$  から  $n$  次倍音成分が求まることを意味する。混合音においても、倍音軌跡の交差付近などストリームが相互干渉する部分以外は、それぞれのピッチ軌跡からそれぞれの調波構造が再生できることになる。

### 5.2.3 ピッチ尤度関数

ここで定義するピッチ尤度関数とは、各時刻の Wavelet 振幅分布だけを観測情報として、混合音に含まれるピッチ周波数とその変化率を、もっともらしさの分布として表現するものである。

Wavelet 振幅分布が倍音周波数ごとに極大値をとるという前述の性質から、ピッチ周波数の尤度関数を

$$q(\Omega_k | y_k) = \sum_{n=1}^N |W(\Omega_k + \log n, t_k)|^2 \quad (5.7)$$

と定義する [120]。  $\Omega_k = \log \omega(t_k)$  であり、添え字  $k$  は時刻  $t_k$  での値、 $y_k$  は観測情報である  $|W(\Omega_k, t_k)|$  を意味する。単一のピッチ周波数しかもたない調波音の場合、 $\Omega_k = \log \omega_0(t_k)$  のとき上式は最大値  $\sum_{n=1}^N |c_n(t_k)|^2$  をとり、ピッチ周波数の最尤推定を与える。

ピッチ周波数の変化率  $\dot{\Omega}_k = \frac{\partial \log \omega(t_k)}{\partial t}$  に関する尤度関数は、Wavelet 空間では、ピッチ周波数の変化により倍音構造全体が対数周波数軸方向へ一様にシフトするという性質を用いて定める [123]。すなわち、ピッチ周波数  $\Omega_k$  毎にピッチ変化率  $\beta(\Omega_k)$  と誤差分散  $\sigma_\beta^2(\Omega_k)$  を推定し、ピッチ変化率に関する尤度関数を、Gauss 分布

$$q(\dot{\Omega}_k | \Omega_k, y_k) = \frac{1}{\sqrt{2\pi\sigma_\beta(\Omega_k)}} e^{-\frac{(\dot{\Omega}_k - \beta(\Omega_k))^2}{2\sigma_\beta^2(\Omega_k)}} \quad (5.8)$$

として与える。両者をあわせて状態ベクトル

$$\boldsymbol{\Omega}_k = (\Omega_k, \dot{\Omega}_k)^T \quad (5.9)$$

を作ると、 $\boldsymbol{\Omega}_k$  の尤度関数は、それらの結合分布

$$q(\boldsymbol{\Omega}_k | y_k) = q(\Omega_k | y_k) q(\dot{\Omega}_k | \Omega_k, y_k) \quad (5.10)$$

で表される。以後この分布をピッチ尤度関数と呼ぶことにする。

この尤度関数は、複数ピッチの混合音に対しては、それぞれに対応した極大点をとる。特にピッチ変化率を状態のひとつに採用したことで、上昇と下降の2つのストリームが交差し、瞬時に同じピッチ周波数をもつ場合にも、ピッチ変化率の違いからピッチ状態空間上では別の極大点をとるなど、ストリーム分離に際して有効な特徴識別が可能になる。

### 5.3 Non-Parametric Kalman フィルタによるピッチ軌跡の最適推定

ここでは、複数のピッチ候補から一つを選んで追跡する問題と、交差や無音区間等によって断続したピッチ軌跡を補間追跡する問題について、動的最適推定の方法論とノンパラメトリック推定の方法論との結合によって同時に解答を与える。

#### 5.3.1 ピッチ確率密度関数

これらの問題を解く鍵は、ピッチ変化のダイナミクスに関する先験的知識と、時々刻々と変化する観測情報をいかに融合させるかにあり、動的最適推定として昔から多くの基礎的研究がなされている問題の一つと捉えられる。すなわち、各時刻で独立に求められるピッチ尤度関数をダイナミクスに従って累積合成し、全体として最適なピッチ確率密度関数  $p(\boldsymbol{\Omega}_k | Y_k)$  をいかに形成するかがここでの問題となる。 $p(\boldsymbol{\Omega}_k | Y_k)$  の表記は、過去  $t_0$  から現時点  $t_k$  までの Wavelet 分布すべて  $Y_k = \{y_k, y_{k-1}, \dots, y_0\}$  を使って推定されたピッチ状態ベクトルの確率密度関数であることを意味する。ここでの要点は、時刻  $t_k$  のピッチ確率密度関数  $p(\boldsymbol{\Omega}_k | Y_k)$  から次の時刻  $t_{k+1}$  の分布  $p(\boldsymbol{\Omega}_{k+1} | Y_{k+1})$  をいかに推定するかという点にある。そのための再帰的手続きを与えるのが Non-Parametric Kalman フィルタ [69] である。

## 5.3.2 予測過程

まず、ピッチ確率密度関数  $p(\Omega_k|Y_k)$  から  $p(\Omega_{k+1}|Y_k)$  を予測する手続きを定式化する。

$$p(\Omega_{k+1}, \Omega_k|Y_k) = p(\Omega_{k+1}|\Omega_k, Y_k)p(\Omega_k|Y_k) \quad (5.11)$$

の両辺を  $\Omega_k$  について積分し、 $\Omega_k$  の Markov 性を仮定して  $p(\Omega_{k+1}|\Omega_k, Y_k) = p(\Omega_{k+1}|\Omega_k)$  が成り立つとすると、Chappman-Kolmogorov の方程式

$$p(\Omega_{k+1}|Y_k) = \int p(\Omega_{k+1}|\Omega_k)p(\Omega_k|Y_k)d\Omega_k \quad (5.12)$$

が得られる。 $p(\Omega_{k+1}|\Omega_k)$  は遷移確率を表しており、ピッチ状態ベクトルがどのように推移していくかを記述するダイナミクスにより決定される。

ピッチ周波数が緩やかに変化することを考慮して、ピッチダイナミクスが

$$\Omega_{k+1} = A\Omega_k + B\xi \quad (5.13)$$

と表せると仮定する。ただし、 $A$  は  $2 \times 2$  の遷移行列、 $B$  は  $2 \times 1$  の入力行列、 $\xi$  は平均 0、分散 1 の Gauss 雑音の駆動項である。例えば遷移行列  $A$  には、ダイナミクスに 2 次系を仮定した

$$A = \begin{bmatrix} 1 & \Delta t \\ -\vartheta^2 \Delta t & 1 - 2\zeta\vartheta \Delta t \end{bmatrix} \quad (5.14)$$

を用いたり、さらにピッチ変化率が一定 ( $\vartheta = 0$ ) であるとした

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (5.15)$$

を用いることが考えられる。本論文では簡単のため後者を仮定する。このとき、駆動雑音  $B\xi$  の各成分に対して、分散  $\sigma_\Omega^2, \sigma_{\dot{\Omega}}^2$  の Gauss 雑音を仮定すると、

$$p(\Omega_{k+1}|\Omega_k, \dot{\Omega}_k) = \frac{1}{\sqrt{2\pi}\sigma_\Omega} e^{-\frac{(\Omega_{k+1} - \Omega_k - \dot{\Omega}_k \Delta t)^2}{2\sigma_\Omega^2}} \quad (5.16)$$

$$p(\dot{\Omega}_{k+1}|\dot{\Omega}_k) = \frac{1}{\sqrt{2\pi}\sigma_{\dot{\Omega}}} e^{-\frac{(\dot{\Omega}_{k+1} - \dot{\Omega}_k)^2}{2\sigma_{\dot{\Omega}}^2}} \quad (5.17)$$

各遷移確率が Gauss 分布で表せる。

$$\begin{aligned} p(\Omega_{k+1}|\Omega_k) &= p(\Omega_{k+1}|\Omega_k, \dot{\Omega}_k)p(\dot{\Omega}_{k+1}|\Omega_{k+1}, \Omega_k, \dot{\Omega}_k) \\ &= p(\Omega_{k+1}|\Omega_k, \dot{\Omega}_k)p(\dot{\Omega}_{k+1}|\dot{\Omega}_k) \end{aligned} \quad (5.18)$$



の関係を利用して (5.12) 式に代入することで、Non-Parametric Kalman フィルタの予測演算が具体的に記述できる。

2次元の畳み込み演算を要する (5.12) 式を直接計算する代わりに、

$$p(\Omega_{k+1}|Y_k) = p(\Omega_{k+1}|Y_k)p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_k) \quad (5.19)$$

と分解し、右辺の各項について独立に計算することを考える。このことで、予測演算の本質を損なうことなく計算量を大幅に押えられる。

ピッチ周波数の確率密度関数については、

$$\begin{aligned} p(\Omega_{k+1}|Y_k) &= \int p(\Omega_{k+1}|\Omega_k, \dot{\Omega}_k)p(\Omega_k, \dot{\Omega}_k|Y_k)d\Omega_k d\dot{\Omega}_k \\ &= \int p(\Omega_{k+1}|\Omega_k, \dot{\Omega}_k)p(\dot{\Omega}_k|\Omega_k, Y_k)p(\Omega_k|Y_k)d\Omega_k d\dot{\Omega}_k \\ &\approx \int \frac{1}{\sqrt{2\pi}\sigma_{\dot{\Omega}}} e^{-\frac{(\Omega_{k+1}-\Omega_k-\hat{\Omega}_k\Delta t)^2}{2\sigma_{\dot{\Omega}}^2}} p(\Omega_k|Y_k)d\Omega_k \end{aligned} \quad (5.20)$$

のように、 $p(\Omega_k|Y_k)$  からの予測式として表せる。ただし上式の 2 段目から 3 段目において、ピッチ変化率に関する確率密度関数が Gauss 分布

$$p(\dot{\Omega}_k|\Omega_k, Y_k) = \frac{1}{\sqrt{2\pi}\hat{\sigma}_{\dot{\Omega}}(\Omega_k)} e^{-\frac{(\dot{\Omega}_k-\hat{\Omega}(\Omega_k))^2}{2\hat{\sigma}_{\dot{\Omega}}^2(\Omega_k)}} \quad (5.21)$$

で表せると仮定し、さらに

$$p(\dot{\Omega}_k|\Omega_k, Y_k) \approx \delta(\dot{\Omega}_k - \hat{\Omega}(\Omega_k)) \quad (5.22)$$

の近似を用いている。これは、(5.16) 式での分布をシフトさせる効果をもつ  $\dot{\Omega}_k\Delta t$  項が、(5.21) 式の最大点をとるピッチ変化率  $\hat{\Omega}(\Omega_k)$  で最も強く働き、それ以外では急激に弱くなることを考慮した。(5.20) 式の予測式は、ピッチ変化に見合うように  $\hat{\Omega}_k\Delta t$  だけ分布を移動させながら、同時に分散  $\sigma_{\dot{\Omega}}^2$  の Gauss 拡散によって分布をぼかす演算で構成される。

また、ピッチ変化率の確率密度関数については、

$$\begin{aligned} p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_k) &= \int p(\dot{\Omega}_{k+1}|\Omega_{k+1}, \dot{\Omega}_k)p(\dot{\Omega}_k|\Omega_{k+1}, Y_k)d\dot{\Omega}_k \\ &= \int p(\dot{\Omega}_{k+1}|\dot{\Omega}_k)p(\dot{\Omega}_k|\Omega_{k+1}, Y_k)d\dot{\Omega}_k \end{aligned} \quad (5.23)$$

と表せることを利用する。(5.17),(5.21) 式より上式は Gauss 分布となり、平均値と分散がそれぞれ

$$\tilde{\Omega}(\Omega_{k+1}) = \hat{\Omega}(\Omega_{k+1}) \quad (5.24)$$

$$\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1}) = \hat{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1}) + \sigma_{\dot{\Omega}}^2 \quad (5.25)$$

と修正される。すなわちピッチ変化率に関する予測演算は、平均値 (推定値) を変化させずに誤差分散だけを増加させる効果をもつ。

## 5.3.3 更新過程

時刻  $t_{k+1}$  でのピッチ尤度関数  $q(\Omega_{k+1}|y_{k+1})$  を新たな観測情報として、すでに予測されたピッチ確率密度関数  $p(\Omega_{k+1}|Y_k)$  を更新する。Bayes の定理

$$p(\Omega_{k+1}|Y_{k+1}) = \frac{p(y_{k+1}|\Omega_{k+1}, Y_k)p(\Omega_{k+1}|Y_k)}{\int p(y_{k+1}|\Omega_{k+1}, Y_k)p(\Omega_{k+1}|Y_k)d\Omega_{k+1}} \quad (5.26)$$

において、 $y_{k+1}$  が観測によって確定しているとすれば、 $p(y_{k+1}|\Omega_{k+1}, Y_k)$  は  $\Omega_{k+1}$  に関する尤度関数と考えられるから、 $p(y_{k+1}|\Omega_{k+1}, Y_k)$  は  $q(\Omega_{k+1}|y_{k+1})$  と置き換えることができ

$$p(\Omega_{k+1}|Y_{k+1}) = \frac{q(\Omega_{k+1}|y_{k+1})p(\Omega_{k+1}|Y_k)}{\int q(\Omega_{k+1}|y_{k+1})p(\Omega_{k+1}|Y_k)d\Omega_{k+1}} \quad (5.27)$$

を得る。

ここでも

$$p(\Omega_{k+1}|Y_{k+1}) = p(\Omega_{k+1}|Y_{k+1})p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_{k+1}) \quad (5.28)$$

と分解し、右辺の各項について独立に計算することを考える。

ピッチ周波数の確率密度関数については、

$$p(\Omega_{k+1}|Y_{k+1}) = \frac{q(\Omega_{k+1}|y_{k+1})p(\Omega_{k+1}|Y_k)}{\int q(\Omega_{k+1}|y_{k+1})p(\Omega_{k+1}|Y_k)d\Omega_{k+1}} \quad (5.29)$$

が成り立ち、(5.7) 式の結果を使って分布が更新される。分子の乗算は、観測結果に予測結果が一致した周波数成分を強調し、予測からはずれた成分を逆に抑圧するように確率密度関数が修正されることを意味し、分母は規格化を表す。

一方、ピッチ変化率の確率密度関数に関しては、

$$p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_{k+1}) = \frac{q(\dot{\Omega}_{k+1}|\Omega_{k+1}, y_{k+1})p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_k)}{\int q(\dot{\Omega}_{k+1}|\Omega_{k+1}, y_{k+1})p(\dot{\Omega}_{k+1}|\Omega_{k+1}, Y_k)d\dot{\Omega}_{k+1}} \quad (5.30)$$

に、(5.24) 式の平均値、分散をもつ Gauss 関数と (5.8) 式を代入すると、再び Gauss 分布になることが示され、平均値と分散が

$$\begin{aligned} \hat{\Omega}(\Omega_{k+1}) &= \frac{\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1})\beta(\Omega_{k+1}) + \sigma_{\beta}^2(\Omega_{k+1})\tilde{\Omega}(\Omega_{k+1})}{\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1}) + \sigma_{\beta}^2(\Omega_{k+1})} \\ \hat{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1}) &= \frac{\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1})\sigma_{\beta}^2(\Omega_{k+1})}{\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1}) + \sigma_{\beta}^2(\Omega_{k+1})} \end{aligned} \quad (5.31)$$

と更新される。予測の信頼度  $1/\tilde{\sigma}_{\dot{\Omega}}^2(\Omega_{k+1})$  と観測の信頼度  $1/\sigma_{\beta}^2(\Omega_{k+1})$  の内分比に応じてピッチ変化率の推定値 (平均値)  $\hat{\Omega}(\Omega_{k+1})$  が決定される。(5.24) と (5.31) 式の結論は、よく知られた線形 Kalman フィルタの結果に一致する。

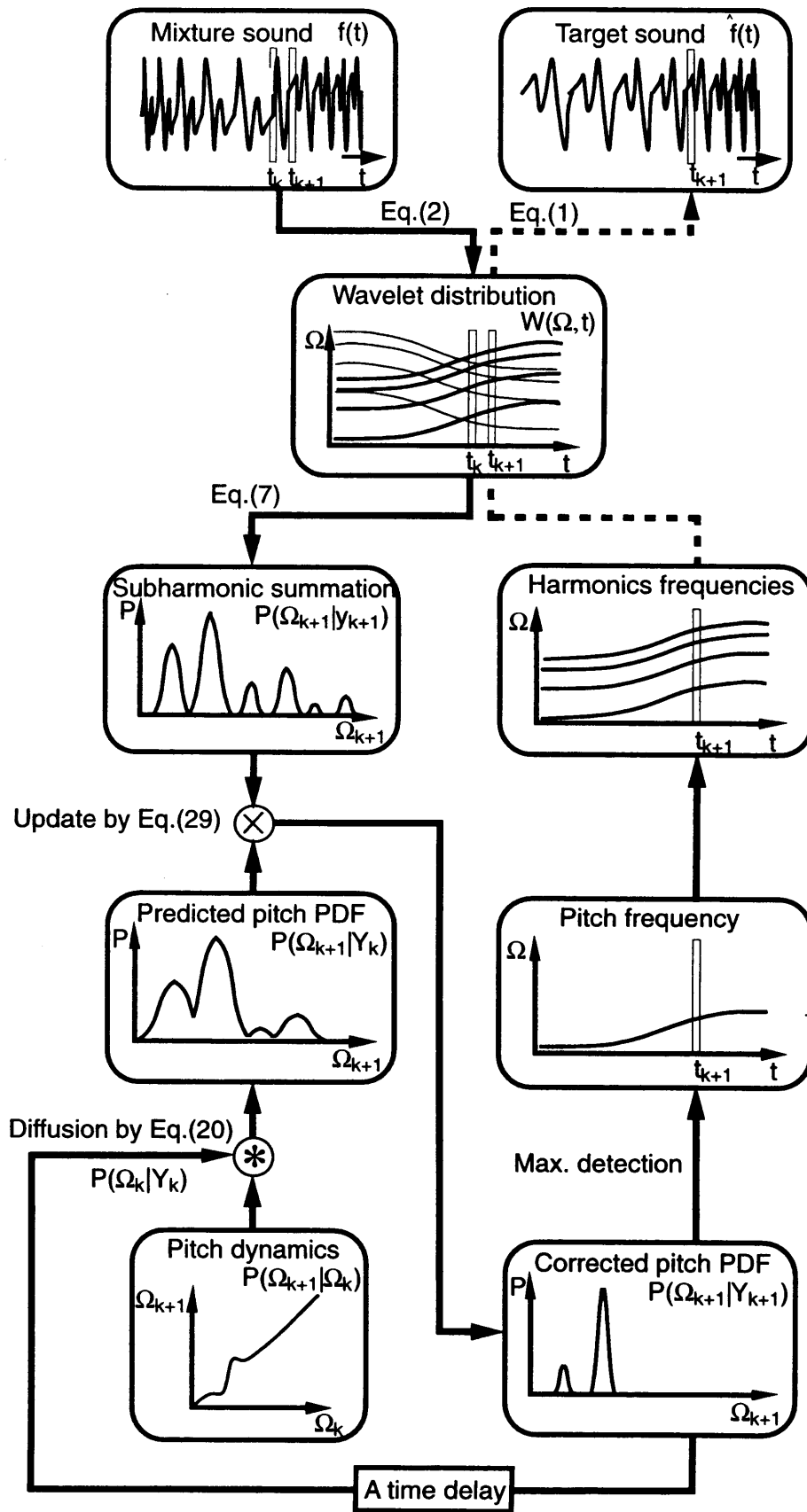


図 5.1 多重ピッチ推定と信号分離アルゴリズムの流れ

## 5.4 ストリームの分離再生

Non-Parametric Kalman フィルタによる以上の予測・更新の各手続きを繰り返すことにより、ピッチ確率密度関数が推定される。ここで初期分布  $p(\Omega_0|y_0)$  として、抽出したい目的ストリームの時刻  $t_0$  でのピッチ周波数付近に重みづけした分布を用いる。このことで、目的ストリームのピッチ周波数軌跡上でのみ確率が高くなるように  $p(\Omega_k|Y_k)$  が更新される。推定された  $p(\Omega_k|Y_k)$  の最大値の軌跡から最大事後確率 (MAP) 推定の意味で最適なピッチ周波数  $\omega_0(t_k)$  の軌跡が求められ、(5.6) 式左辺から各倍音振幅  $a_n(t)$  および (5.1) 式から目的ストリームが再生される。以上の処理の流れを図 5.1 にブロック図としてまとめた。

## 5.5 ピッチ推定誤差の影響と分析 Wavelet の Q 値の決定

ピッチ周波数  $\omega_0(t)$  の推定に誤差がともなっていると、それを基準に計算される倍音振幅  $c_n(t)$  にも誤差が波及する。そこで、ピッチ推定誤差が倍音振幅推定に及ぼす影響を調べる。

ピッチ周波数が Wavelet 空間上で対数周波数として求められることを考慮すれば、一般に対数ピッチ周波数は誤差  $\log \epsilon(t)$  を含んだ  $\log \omega_0(t) + \log \epsilon(t)$  と推定されると考えてよい。このとき、 $n$  次倍音の対数周波数  $\log n\omega_0(t)\epsilon(t)$  上の Wavelet 複素振幅値は、(5.5) 式に代わって、

$$W(\log n\omega_0(t)\epsilon(t), t) \approx c_n(t) e^{jn \int_0^t \omega_0(s) ds} e^{-\frac{(1-\epsilon(t))^2}{2(B\epsilon(t))^2}} \quad (5.32)$$

となる。ピッチ誤差がないときと比べて、係数  $e^{-\frac{(1-\epsilon(t))^2}{2(B\epsilon(t))^2}}$  だけの振幅歪みが生じる。ただし、この歪みは倍音次数  $n$  には無関係なため、全ての倍音成分に一律な振幅増減を与えるだけで、倍音間の振幅構成比には影響を与えないという有利な性質がある。短時間 Fourier 解析では、この性質が成り立たないことに注意する (付録 5.8.2 参照)。分析 Wavelet のパラメータ  $B$  を大きな値に選べば、(5.4) 式の誤差項の影響が小さくなるともに、倍音振幅の推定がピッチ誤差に鈍感になる。ところが大きくしすぎると、隣接する倍音成分との干渉が無視できなくなり、たとえピッチ周波数が正確に求められたとしても倍音振幅に正弦波状の振動成分が現れてしまう。そこで、干渉が最も起こりやすい最高次  $N$  の倍音成分において隣接する  $N-1$  次倍音との干渉を

$$e^{-\frac{\{(N-1)\omega_0(t)-N\omega_0(t)\}^2}{2(BN\omega_0(t))^2}} = e^{-\frac{1}{2(BN)^2}} = \frac{1}{2^r} \quad (5.33)$$

で許容するように  $B$  を選ぶ。分析 Wavelet の  $Q$  値には、

$$e^{-\frac{(\Delta\omega)^2}{2(B\omega)^2}} = e^{-\frac{1}{8B^2Q^2}} = \frac{1}{2} \quad (5.34)$$

の関係があるから (ただし  $\omega$  は中心周波数、 $\Delta\omega$  は帯域幅)、

$$Q = \frac{1}{\sqrt{8 \log 2} B} = \frac{\sqrt{r}}{2} N \quad (5.35)$$

が成り立つ。

## 5.6 実験

以下の実験では、サンプリング周波数:11.025kHz、サンプリング点数:8192 点、1 オクターブあたりの周波数分割数:50 点とし、 $N = 16$ 、 $r = 8$  に選び、 $Q=22.6$  の分析 Wavelet を使って計算を行った。

### 5.6.1 合成信号におけるストリーム分離

まず、合成信号に対するストリームの分離再生実験を行った。合成信号として、ピッチ周波数が低から高に変化する調波信号と逆に高から低に変化する調波信号を足し合わせたものを考える。

これを Wavelet 変換した結果が図 5.2(a) である。同図 (b) は、この Wavelet 分布に対して (5.7) 式の倍音和スペクトラムを計算した結果である。両ストリームそれぞれのピッチ周波数軌跡で極大となっていることが確かめられる。試しに、最大値をとる周波数点を各時刻で取り出しプロットした結果を同図 (c) に示した。両ストリームのピッチ軌跡が交互に断続的に現れていて、片方のピッチ軌跡だけを取り出せる状態にはなっていない。これに対して、Non-Parametric Kalman フィルタにより推定されたピッチ周波数の確率密度関数 (PDF) の時間軌跡を同図 (d) に示す。初期分布には、低い方のピッチ周波数が選ばれるように重みづけしたものをを用いている。

各倍音振幅成分がどの程度忠実に復元できているかを、相関係数

$$\text{Corr}(\mathbf{c}(t), \hat{\mathbf{c}}(t)) = \frac{\sum_{n=1}^N c_n^*(t) \hat{c}_n(t)}{\sqrt{\sum_{n=1}^N |c_n(t)|^2} \sqrt{\sum_{n=1}^N |\hat{c}_n(t)|^2}} \quad (5.36)$$

により評価する。ここで、 $\mathbf{c}(t) = [c_1(t), c_2(t), \dots, c_N(t)]^T$  は真の倍音振幅、 $\hat{\mathbf{c}}(t) = [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_N(t)]^T$  は再生された倍音振幅を表す。この値が 1 に近いほど倍音構成比すなわち音色が正確に再現できていることを示す。同図 (e) がその計算結果であり、ほぼ全時刻にわたり忠実に再生できていることがわかる。0.35sec 付近で値が特に落ちているの

は、この時刻で両ストリームのすべての倍音成分がちょうど重なっているため、倍音振幅の読み取りに大きな誤差が生じているからである。また、倍音振幅の 2 乗和 (信号エネルギー)  $\sum_{n=1}^N |c_n(t)|^2$ 、 $\sum_{n=1}^N |\hat{c}_n(t)|^2$  をそれぞれ計算した結果が同図 (f) である。この結果からも、全時刻にわたり真値がほぼ正確に再現できていることがわかる。

ピッチ確率密度関数の初期分布として、高い方のピッチ周波数が選ばれるように重みづけしたものをを用いて推定した結果が同図 (g) である。

### 5.6.2 実音声におけるストリーム分離

実音声に対する本手法の有用性を検証するために、ほぼ調波信号とみなせる有声音 (母音) を用いたストリーム分離実験を行った。混合音に含まれる外乱としては、(1) 非定常な雑音成分、(2) 他の有声音、が実環境下での主な要因となることを考慮し、この両者をそれぞれ想定した実験を行った。

#### 雑音中の音声の分離抽出

背景雑音が存在する中で「あお」と発声して得られた雑音音声信号 (図 5.3(a)) から、音声だけを分離抽出する実験を行った。ここで背景雑音には紙のがさつく音を用いた。同図 (b) が Wavelet 分布、同図 (c) が Non-Parametric Kalman フィルタによって推定されたピッチ周波数の確率密度関数である。同図 (d) はそのピッチ軌跡から音声波形を再生した結果である。聴感上、分離前は雑音に埋もれて音声がかき消えていたのに対して、雑音がほとんど取り除かれ明瞭に発音が聞こえるようになった。

#### 多重音声の分離抽出

二人が同時にそれぞれ「あお」、「いう」と発声することにより得た混合音 (図 5.4(a)) から、個々の音声を分離再生する実験を行った。Wavelet 分布 (同図 (b)) から倍音和スペクトラムを計算し、各時刻での最大点を求め (同図 (c))、それをピッチ軌跡としたときの再生音は、聴感上、両発音の前半と後半が結合しかつ歪みの伴った「あう」のような音になった。これに対して、Non-Parametric Kalman フィルタにより推定されたピッチ周波数の確率密度関数 (同図 (d)) から再構成した音声 (同図 (e)) は、聴感上も「あお」の発声だけをほぼ正確に分離再生できることが確かめられた。

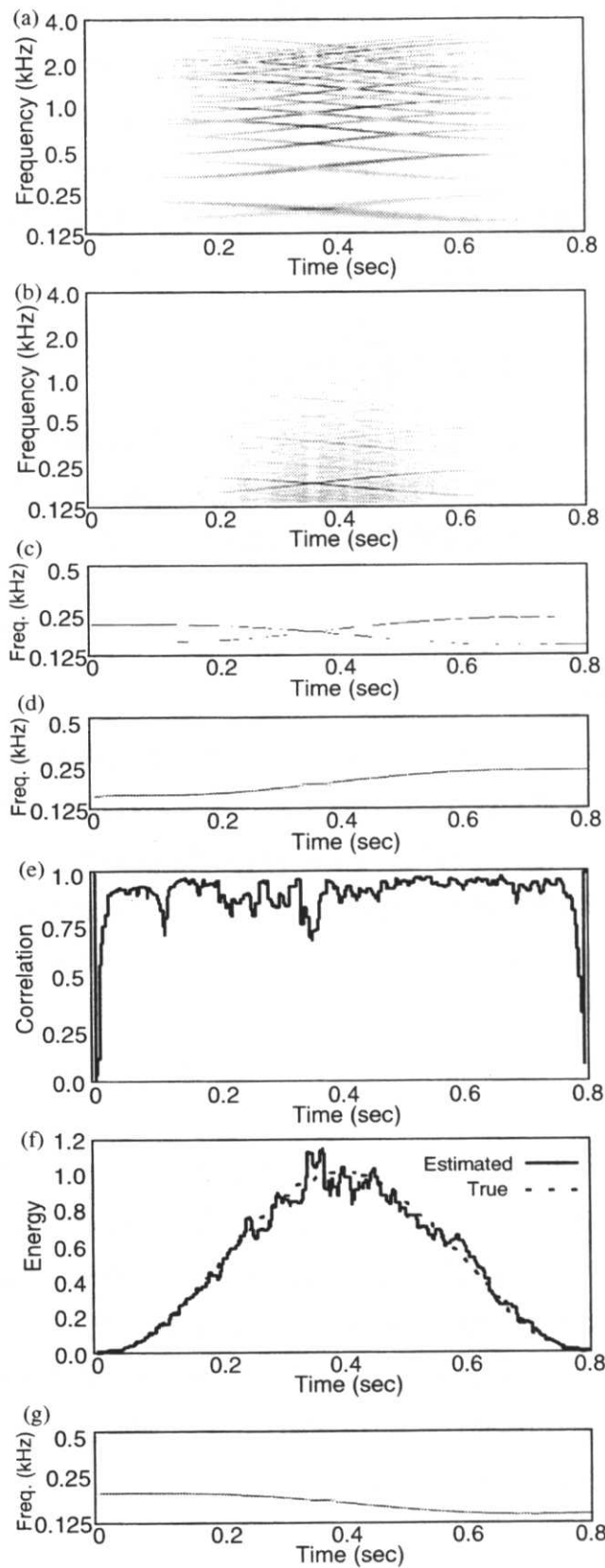


図 5.2 合成信号に対する実験結果、(a)Wavelet 分布、(b)倍音とスペクトラム、(c)最大点の軌跡、(d)上昇音に対するピッチ PDF の推定結果、(e)相関係数、(f)倍音振幅二乗和の分布、(g)下降音に対するピッチ PDF の推定結果

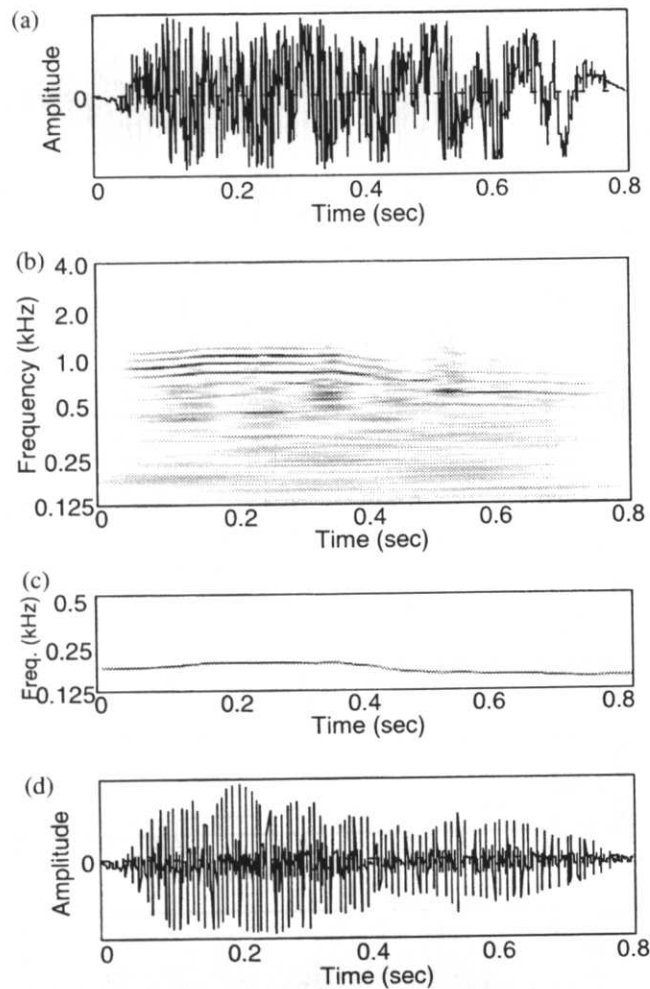


図 5.3 雑音下の音声に対する実験結果、(a) 雑音で汚された音声、(b) Wavelet 分布、(c) ピッチ PDF の推定結果、(d) 音声の抽出結果

## 5.7 本章のまとめ

本章では、異なる複数の擬似周期信号が含まれる混合音に対するピッチ推定法と個別分離アルゴリズムを提案した。まず、擬似周期信号が Wavelet 空間上でどのように表せるかを述べ、その調波スペクトル構造からピッチ尤度関数を定義した (5.2 節)。それを観測情報としつつピッチ変化のダイナミクスに従うようにピッチ尤度関数を修正する Non-Parametric Kalman フィルタの手続きを導出した (5.3 節)。ピッチ尤度関数の推定結果に基づくピッチ軌跡の決定法およびストリーム再生の手続きを説明し (5.4 節)、特に Wavelet 変換を行う際の  $Q$  値の決定法について付記した (5.5 節)。以上で導かれたアルゴリズムを用いて、計算機合成した混合音に対する個別ストリーム再生、雑音環境下での音声の抽出および複数話者から個別音声再生などの評価実験を行った (5.6 節)。



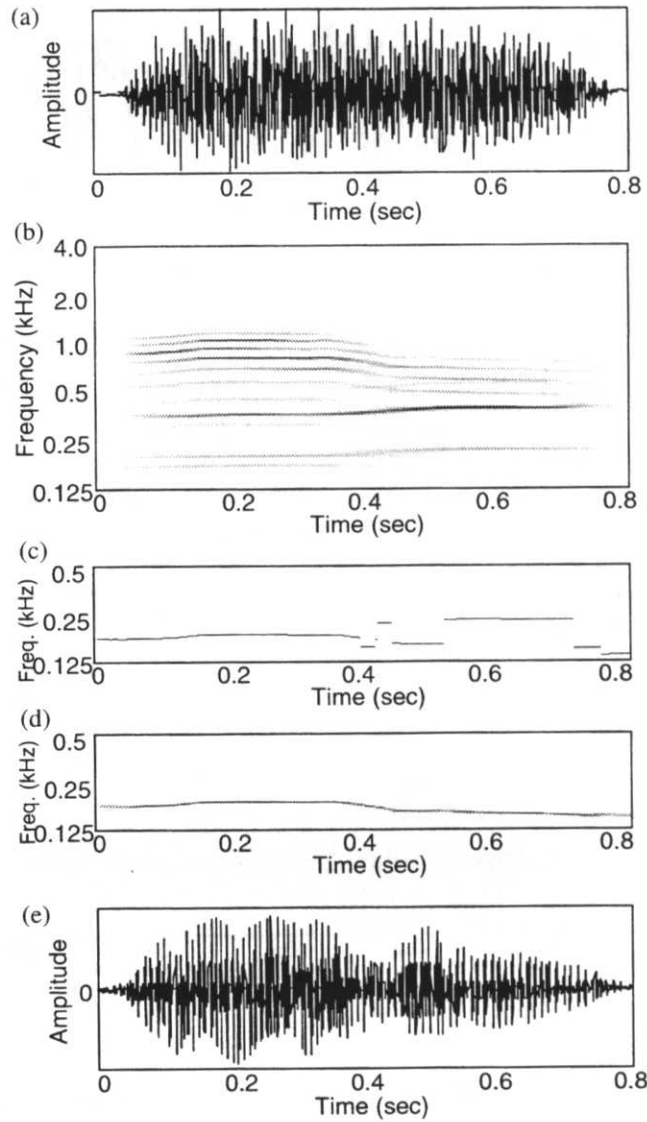


図 5.4 多重音声に対する実験結果、(a)「あお」,「いう」の混合音声、(b)Wavelet 分布、(c)倍音和スペクトラムの最大点の軌跡、(d)「あお」に対するピッチ PDF の推定結果、(e)「あお」の抽出結果

## 5.8 付録

### 5.8.1 Wavelet 分布の分解表現

以下に (5.4) 式を導く。

(5.2) 式に (5.1),(5.3) 式を代入し、 $v = u - t$  の変数置換を行う。

$$W(\log \omega, t) = \int_{-\infty}^{\infty} x(u) \frac{B\omega}{\sqrt{2\pi}} e^{-\frac{(B\omega)^2(u-t)^2}{2}} e^{-j\omega(u-t)} du$$

$$= \int_{-\infty}^{\infty} c_n(t+v) e^{jn \int_0^{t+v} \omega_0(s) ds} \frac{B\omega}{\sqrt{2\pi}} e^{-\frac{(B\omega)^2 v^2}{2}} e^{-j\omega v} dv \quad (5.37)$$

$c_n(t+v) e^{jn \int_0^{t+v} \omega_0(s) ds}$  が  $v = 0$  の近辺でしか積分に寄与せず、振幅  $c_n(t)$  やピッチ周波数  $\omega_0(t)$  が緩やかにしか変化しないことに着目し、

$$\begin{aligned} c_n(t+v) &= c_n(t) + R_c(v, t) \\ \int_0^{t+v} \omega_0(s) ds &= \int_0^t \omega_0(s) ds + \omega_0(t)v + R_\omega(v^2, t) \end{aligned} \quad (5.38)$$

と Taylor 展開したものを上式に代入する。

$$\begin{aligned} W(\log \omega, t) &= \int_{-\infty}^{\infty} \sum_{n=1}^N c_n(t) e^{jn \{ \int_0^t \omega_0(s) ds + \omega_0(t)v \}} \frac{B\omega}{\sqrt{2\pi}} e^{-\frac{(B\omega)^2 v^2}{2}} e^{-j\omega v} dv + \varepsilon(\omega, t) \\ &= \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} \int_{-\infty}^{\infty} \frac{B\omega}{\sqrt{2\pi}} e^{-\frac{(B\omega)^2 v^2}{2}} e^{-j\{\omega - n\omega_0(t)\}v} dv + \varepsilon(\omega, t) \\ &= \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} e^{-\frac{\{\omega - n\omega_0(t)\}^2}{2(B\omega)^2}} + \varepsilon(\omega, t) \end{aligned} \quad (5.39)$$

よって (5.4) 式が得られる。ただし  $\varepsilon(\omega, t)$  は、Taylor 展開の高次項  $R_c(v, t), R_\omega(v^2, t)$  に対する積分寄与をひとまとめにした残差成分を意味する。

### 5.8.2 短時間 Fourier 解析によるピッチ推定誤差の影響

Wavelet 解析の代わりに短時間 Fourier 解析を行った場合を考える。擬似周期信号  $x(t)$  の短時間 Fourier 変換  $F(\omega, t)$  は、(5.2), (5.4) 式の代わりに、

$$F(\omega, t) = \int_{-\infty}^{\infty} x(u) \frac{B}{\sqrt{2\pi}} e^{-\frac{B^2(u-t)^2}{2}} e^{-j\omega(u-t)} du \quad (5.40)$$

$$\approx \sum_{n=1}^N c_n(t) e^{jn \int_0^t \omega_0(s) ds} e^{-\frac{1}{2B^2} \{\omega - n\omega_0(t)\}^2} \quad (5.41)$$

と計算できるから、 $\omega = n\omega_0(t)$  上で極大値をとる。

5.5節の議論と同様に、ピッチ周波数  $\omega_0(t)$  の推定に際し誤差  $\varepsilon(t)$  が加わっていると考えると、 $n$  次倍音周波数  $n(\omega_0(t) + \varepsilon(t))$  上の Fourier 振幅値を求めると、

$$F(n(\omega_0(t) + \varepsilon(t)), t) \approx c_n(t) e^{jn \int_0^t \omega_0(s) ds} e^{-\frac{1}{2B^2} \{n\varepsilon(t)\}^2} \quad (5.42)$$

となる。振幅誤差  $e^{-\frac{1}{2B^2} \{n\varepsilon(t)\}^2}$  は、倍音次数  $n$  に依存し、倍音間の振幅比を変えてしまう原因となる。

## 第 6 章

# 非定常擬似周期信号に対する線形最適フィルタ

### 6.1 はじめに

第3章では、擬似周期信号の定常性、すなわち振幅変動やピッチ変動の統計的性質(期待値や分散)が時間に依存せず一定であることが仮定できれば、信号のスペクトル特性やフィルタの伝達特性について見通しの良い議論が可能になることを述べてきた。しかしながら、音声や楽音などの実信号では、この仮定は一般に成り立たず、振幅やピッチ変動の大きさが時間によって変化することが普通である。このような非定常信号の場合のフィルタ設計では、信号のスペクトル分布の計算を通して伝達関数を求めるこれまでのような Wiener フィルタの手法に代わって、信号の時変性を微分方程式で記述し、そのダイナミクスに従うように推定値を時間軸上で逐次的に求めていく Kalman フィルタの考え方が必要になる。

Kalman フィルタの理論で用いられる状態空間モデルでは、信号のランダム性を記述するのに、微分方程式を雑音で駆動することにより表しており<sup>†</sup>、通常その駆動項には信号本体に対して加法性の雑音が仮定されている。実際 Bitmead ら [82] は、加法性雑音による揺らぎを与えた周期信号モデルに対する Kalman フィルタを提案している。しかしながら、本論文で扱う「振幅変動+ピッチ変動」信号では、(2.19)式のように信号の位相項にランダム性をもつため、その状態空間モデルとしては駆動項が信号と雑音の積で与えられること<sup>††</sup>を考慮する必要がある。Park ら [83] は、正弦波信号と雑音の積を使って近似的に状態依存性雑音を表現することによって、振幅変動を模擬し離散 Kalman フィルタを導いている。厳密な状態依存性雑音を考慮した Kalman フィルタの設計法は、信号モデルを限定しない一般的な形式で McLane [84] によって研究された。しかしながら、振幅変動やピッチ変

<sup>†</sup>一般に白色 Gauss 雑音が仮定され、その時間積分である Wiener 過程(付録 3.11.1)を用いて表される。このときの微分方程式は、信号の振舞いを確率過程として表すことになるため、特に「確率微分方程式」と呼ばれる。

<sup>††</sup>このような駆動項を「状態依存性雑音」と呼ぶ。

動の時変性を考慮した擬似周期信号に対しての、状態依存性雑音を取り入れた確率微分方程式によるモデル化やそれに基づいた Kalman フィルタの設計法に関する議論は今までになかった。

状態依存性雑音により駆動される確率微分方程式を考える上で注意すべきは、信号と雑音の両者が確率過程をとるので、その積を微分方程式で表したときに、極限をとる前の差分演算をどう解釈するかという点にある。これには、雑音の差分に対して時刻の始点に信号のサンプル点をとる伊藤型と、時刻の midpoint にサンプル点をとる Stratonovich 型がある [80]。特に本論文では、信号と雑音を互いに統計的に独立にとることができて期待値演算が簡単になる伊藤型の確率微分方程式を用いて Kalman フィルタを導出する。その解の性質を調べることにより、3章では解析の容易さのため無視した倍音間の干渉の影響を詳細に議論する。

## 6.2 時変型「振幅変動＋ピッチ変動」信号モデル

### 6.2.1 Wiener 積分による振幅 / ピッチ変動の定式化

一般に振幅や基本周波数には、巨視的変動と微視的変動が同時に存在すると考えられる。巨視的変動は、非定常性を許容するかわりに緩やかな変化を前提とするため、その挙動は確定的であり変化の軌跡を事前に知り得るものとする。逆に微視的変動は、観測に供しない細かく素早い変動であり不規則性を有するが、そのかわり短時間スケールではその挙動はほぼ定常であるとする。この両者を同時に考慮した信号モデルを立てることがこの狙いである。

そこで、信号モデル 4 の振幅変動やピッチ変動の期待値や分散を時変パラメータに置き替えるために、(2.19) 式で

$$\begin{aligned}\Delta c_n(t) &= \frac{\dot{\bar{x}}_n^2(t)}{2\bar{x}_n^2(t)} - \sigma_{an}^2(t) + \sigma_{an}(t)\xi_{an}(t) + j\sigma_{bn}(t)\xi_{bn}(t) && (\text{複素定常白色 Gauss 雑音過程}) \\ \omega_0(t) &= \bar{\omega}_0(t) + \sigma_\omega(t)\xi_\omega(t) && (\text{定常白色 Gauss 雑音過程})\end{aligned}\quad (6.1)$$

とおいたもの考える。ただし、 $x(t)$  は  $t \geq 0$  で定義されるものとする。既知の時変パラメータである二乗平均振幅  $\bar{x}_n^2(t)$ 、基本周波数の期待値  $\bar{\omega}_0(t) \equiv E[\omega_0(t)]$  ( $E[\cdot]$  は集合平均の意)、および振幅や基本周波数の標準偏差  $\sigma_*(t)$  が巨視的変動を表し、白色ランダム雑音項  $\xi_*(t)$  が微視的変動を意味する。特に、 $\bar{x}_n^2(t), \sigma_*(t)$  を定数とした場合が信号モデル 4 に対応する。ただし、倍音振幅や基本周波数の駆動源  $\xi_{an}(t), \xi_{bn}(t), \xi_\omega(t)$  は平均 0、分散 1 の互いに独立な定常白色 Gauss 雑音であり、振幅係数  $c_n$  に関しては仮定 2 より  $E[c_n] = 0$ 、 $E[c_n c_m^*] = 0$  ( $n \neq m$ ) が成り立つ。また、 $\dot{\bar{x}}_n^2(t)$  は  $\bar{x}_n^2(t)$  に関する時間の一階微分を表す。

このとき、 $\Delta c_n(t)$  の第1項に関して、

$$\int_0^t \frac{\dot{\bar{x}}_n^2(s)}{2\bar{x}_n^2(s)} d\tau = \log \sqrt{\frac{\bar{x}_n^2(t)}{\bar{x}_n^2(0)}} \quad (6.2)$$

が成り立つこと、および

$$d\beta_{an}(s) = \xi_{an}(s)ds, \quad d\beta_{bn}(s) = \xi_{bn}(s)ds, \quad d\beta_\omega(s) = \xi_\omega(s)ds \quad (6.3)$$

とおくことによる、付録 6.8.1 の Wiener 積分の表記を用いることによって、(2.19) 式は

信号モデル 5:

$$x(t) = \sum_{n=1}^N x_n(t), \quad (6.4)$$

$$x_n(t) = c_n \sqrt{\frac{\bar{x}_n^2(t)}{\bar{x}_n^2(0)}} e^{-\int_0^t \sigma_{an}^2(s)ds + \int_0^t \sigma_{an}(s)d\beta_{an}(s)} e^{j\int_0^t \sigma_{bn}(s)d\beta_{bn}(s) + n\int_0^t \sigma_\omega(s)d\beta_\omega(s)} e^{jn\int_0^t \bar{\omega}_0(s)ds}$$

と表せる。したがって(6.67)式より、各倍音成分に対する自己相関は

$$\begin{aligned} R_{x_n}(t+\tau, t) &= E[x_n(t+\tau)x_n^*(t)] \\ &= \sqrt{\bar{x}_n^2(t+\tau)\bar{x}_n^2(t)} e^{-2\int_0^t \sigma_{an}^2(s)ds} E[e^{2\int_0^t \sigma_{an}(s)d\beta_{an}(s)}] e^{-\int_t^{t+\tau} \sigma_{an}^2(s)ds} \\ &\quad E[e^{j\int_t^{t+\tau} \sigma_{an}(s)d\beta_{an}(s)}] E[e^{j\int_t^{t+\tau} \sigma_{bn}(s)d\beta_{bn}(s)}] E[e^{jn\int_t^{t+\tau} \sigma_\omega(s)d\beta_\omega(s)}] e^{jn\int_t^{t+\tau} \bar{\omega}_0(s)ds} \\ &= \sqrt{\bar{x}_n^2(t+\tau)\bar{x}_n^2(t)} e^{-\frac{1}{2}\int_t^{t+\tau} \{\sigma_{an}^2(s) + \sigma_{bn}^2(s) + n^2\sigma_\omega^2(s)\}ds} e^{jn\int_t^{t+\tau} \bar{\omega}_0(s)ds} \end{aligned} \quad (6.5)$$

と表せ、特に  $\tau = 0, t = 0$  とおけば、

$$E[|x_n(t)|^2] = \bar{x}_n^2(t), \quad E[|c_n|^2] = \bar{x}_n^2(0) \quad (6.6)$$

が成り立つことに注意する。

### 6.2.2 伊藤型確率微分方程式による「振幅変動+ピッチ変動」信号モデル

時変性を有する信号の振舞いは、一般に微分方程式によって時間的推移として記述することができる。特に、複素位相項が確率的に揺らぐ上式のような信号を表すのには、伊藤確率微分方程式が有用である。これは、(6.4)式で

$$w_{an}(t) = \int_0^t \sigma_{an}(s)d\beta_{an}(s), \quad w_{bn}(t) = \int_0^t \sigma_{bn}(s)d\beta_{bn}(s), \quad w_\omega(t) = \int_0^t \sigma_\omega(s)d\beta_\omega(s) \quad (6.7)$$

とおくことにより、多変数関数

$$x_n(t) = f(t, w_{an}(t), w_{bn}(t), w_\omega(t)) \quad (6.8)$$

とみなして、(6.72) 式の「伊藤の公式」 [85] を適用することにより、

$$dx_n(t) = \left[ \frac{\dot{\bar{x}}_n^2(t)}{2\bar{x}_n^2(t)} - \frac{1}{2} \{ \sigma_{an}^2(t) + \sigma_{bn}^2(t) + n^2 \sigma_\omega^2(t) \} + jn\bar{\omega}_0(t) \right] x_n(t) dt + \sigma_{an}(t) x_n(t) d\beta_{an}(t) + j\sigma_{bn}(t) x_n(t) d\beta_{bn}(t) + jn\sigma_\omega(t) x_n(t) d\beta_\omega(t) \quad (6.9)$$

と表せる (付録 6.8.2 参照)。ただし上式は、それと等価な積分形式

$$x_n(t) = x_n(0) + \int_0^t \left[ \frac{\dot{\bar{x}}_n^2(s)}{2\bar{x}_n^2(s)} - \frac{1}{2} \{ \sigma_{an}^2(s) + \sigma_{bn}^2(s) + n^2 \sigma_\omega^2(s) \} + jn\bar{\omega}_0(s) \right] x_n(s) ds + \int_0^t \sigma_{an}(s) x_n(s) d\beta_{an}(s) + j \int_0^t \sigma_{bn}(s) x_n(s) d\beta_{bn}(s) + jn \int_0^t \sigma_\omega(s) x_n(s) d\beta_\omega(s) \quad (6.10)$$

に対して、積分範囲を  $[t, t + \delta t]$  としたとき

$$x_n(t + \delta t) - x(t) = \int_t^{t+\delta t} \left[ \frac{\dot{\bar{x}}_n^2(s)}{2\bar{x}_n^2(s)} - \frac{1}{2} \{ \sigma_{an}^2(s) + \sigma_{bn}^2(s) + n^2 \sigma_\omega^2(s) \} + jn\bar{\omega}_0(s) \right] x_n(s) ds + \int_t^{t+\delta t} \sigma_{an}(s) x_n(s) d\beta_{an}(s) + j \int_t^{t+\delta t} \sigma_{bn}(s) x_n(s) d\beta_{bn}(s) + jn \int_t^{t+\delta t} \sigma_\omega(s) x_n(s) d\beta_\omega(s) \quad (6.11)$$

の  $\delta t \rightarrow 0$  とした微分形式であると解釈する。特に、(6.10) 式の右辺第 2 項以降は状態依存性雑音を表し、伊藤積分 [80] として定義されていることに注意する (付録 6.8.2 参照)。

(6.9) 式右辺の実部と虚部により、それぞれ倍音振幅と位相が決定される。信号の確定的な挙動が第 1 項で表されるのに対して、それに加わるランダムな揺らぎは第 2 項以降により与えられる。特に、第 1 項のみを考慮し複素パラメータを一定としたときの微分方程式は、(2.14) 式による定係数の波動方程式を、互いに複素共役な 1 次系の積に因数分解したときの片方に一致することが確かめられる [86]。

### 6.3 「振幅変動 + ピッチ変動」信号推定問題の動的システム表現

(6.9) 式が各倍音で成り立つことを考慮し、これらをまとめてベクトル形式で

$$d\mathbf{x}(t) = \Omega(t)\mathbf{x}(t)dt + \left[ \Sigma_a(t)dB_a(t) + j\Sigma_b(t)dB_b(t) + j\Sigma_\omega(t)d\beta_\omega(t) \right] \mathbf{x}(t) \quad (6.12)$$

と表しておく。ただし、

$$\begin{aligned} \mathbf{x}(t) &= [x_1(t), x_2(t), \dots, x_N(t)]^T, \\ dB_a(t) &= \text{diag}\{d\beta_{a1}(t), d\beta_{a2}(t), \dots, d\beta_{aN}(t)\}, \\ dB_b(t) &= \text{diag}\{d\beta_{b1}(t), d\beta_{b2}(t), \dots, d\beta_{bN}(t)\}, \end{aligned}$$

$$\begin{aligned}\Omega(t) &= \text{diag} \left\{ \frac{\dot{\bar{x}}_1^2(t)}{2\bar{x}_1^2(t)} - \frac{1}{2} \{ \sigma_{a1}^2(t) + \sigma_{b1}^2(t) + \sigma_\omega^2(t) \} + j\bar{\omega}_0(t), \right. \\ &\quad \frac{\dot{\bar{x}}_2^2(t)}{2\bar{x}_2^2(t)} - \frac{1}{2} \{ \sigma_{a2}^2(t) + \sigma_{b2}^2(t) + 4\sigma_\omega^2(t) \} + j2\bar{\omega}_0(t), \\ &\quad \left. \dots, \frac{\dot{\bar{x}}_N^2(t)}{2\bar{x}_N^2(t)} - \frac{1}{2} \{ \sigma_{aN}^2(t) + \sigma_{bN}^2(t) + N^2\sigma_\omega^2(t) \} + jN\bar{\omega}_0(t) \right\}, \\ \Sigma_a(t) &= \text{diag} \{ \sigma_{a1}(t), \sigma_{a2}(t), \dots, \sigma_{aN}(t) \}, \\ \Sigma_b(t) &= \text{diag} \{ \sigma_{b1}(t), \sigma_{b2}(t), \dots, \sigma_{bN}(t) \}, \\ \Sigma_\omega(t) &= \text{diag} \{ \sigma_\omega(t), 2\sigma_\omega(t), \dots, N\sigma_\omega(t) \}\end{aligned}$$

であり、 $\text{diag}\{\cdot\}$  は対角行列を表す。特に、 $\mathbf{x}(t)$  は  $N$  次元の列ベクトル、 $dB_a(t), dB_b(t), \Omega(t), \Sigma_a(t), \Sigma_b(t), \Sigma_\omega(t)$  は  $N$  次元の正方行列である。

これに対し我々が観測できる信号は、**仮定 4a** を前提として

$$\mathbf{y}(t) = \mathbf{1}^T \mathbf{x}(t) + v(t) \quad (6.13)$$

と表せるものとする。ただし、 $\mathbf{1} = [1, 1, \dots, 1]^T$  ( $N$  次元列ベクトル) である。特に、

**仮定 8c:** 観測雑音に関しても時変性を考慮して、

$$v(t) = \sigma_v(t)\xi_v(t) \quad (6.14)$$

を仮定する。ただし、 $\xi_v(t)$  は

$$E[\xi_v(t)] = 0, \quad E[\xi_v(t+\tau)\xi_v(t)] = \delta(\tau) \quad (6.15)$$

が成り立つ定常白色 Gauss 雑音である。

雑音の強さを決める時変パラメータ  $\sigma_v(t)$  を時間に依存せず一定  $\sigma_v$  とした場合が**仮定 8a** に対応する。ここで、

$$\frac{dz(t)}{dt} = y(t) \quad (6.16)$$

とにおいて、(6.3) 式と同様に、雑音項に対して Wiener 積分の表記

$$d\beta_v(t) = \xi_v(t)dt \quad (6.17)$$

を用いれば、(6.13) 式は

$$dz(t) = \mathbf{1}^T \mathbf{x}(t)dt + \sigma_v(t)d\beta_v(t) \quad (6.18)$$

のように、確率微分方程式の形式で表せる。

このとき、我々が解くべき問題を以下のように設定する。

問題設定 4: (6.18) 式の観測モデル

$$dz(t) = \mathbf{1}^T \mathbf{x}(t)dt + \sigma_v(t)d\beta_v(t) \quad (6.19)$$

を通して得られた時間区間  $[0, t]$  での観測データ  $y(s)$  ( $0 \leq s \leq t$ ) を用いて、(6.12) 式の状態空間モデル

$$d\mathbf{x}(t) = \Omega(t)\mathbf{x}(t)dt + \left[ \Sigma_a(t)dB_a(t) + j\Sigma_b(t)dB_b(t) + j\Sigma_\omega(t)d\beta_\omega(t) \right] \mathbf{x}(t) \quad (6.20)$$

のダイナミクスに従う目的信号 (倍音成分)  $\mathbf{x}(t)$  の現時刻  $t$  での推定値  $\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \dots, \hat{x}_N(t)]^T$  を、線形フィルタ方程式の出力

$$d\hat{\mathbf{x}}(t) = F(t)\hat{\mathbf{x}}(t)dt + \mathbf{k}(t)dz(t) \quad (6.21)$$

として、平均二乗誤差

$$J = E[\|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|^2] \quad (6.22)$$

が最小になるように求めたい。ただし、簡単のため  $\hat{\mathbf{x}}(0) = \mathbf{0}$  とする。このときのフィルタ係数行列  $F(t)$  ( $N$  次元の正方行列)、同係数ベクトル  $\mathbf{k}(t)$  ( $N$  次元の列ベクトル) を決定する問題を考える。

(6.21) 式は、両辺を  $dt$  で割れば、通常の微分方程式

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = F(t)\hat{\mathbf{x}}(t) + \mathbf{k}(t)y(t) \quad (6.23)$$

に対応することに注意する。このとき、上式右辺第 1 項に対する遷移行列、すなわち

$$\begin{aligned} \frac{\partial \Phi_x(t, \tau)}{\partial t} &= F(t)\Phi_x(t, \tau), \\ \Phi_x(t, t) &= I \end{aligned} \quad (6.24)$$

を満たす  $\Phi_x(t, \tau)$  を用いれば ( $I$  は単位行列)、一般解は

$$\hat{\mathbf{x}}(t) = \Phi_x(t, 0)\hat{\mathbf{x}}(0) + \int_0^t \Phi_x(t, \tau)\mathbf{k}(\tau)y(\tau)d\tau \quad (6.25)$$

と表せる。 $\hat{\mathbf{x}}(0) = \mathbf{0}$  を代入し、(6.16) 式の関係を使えば、

$$\hat{\mathbf{x}}(t) = \int_0^t \mathbf{h}(t, \tau)dz(\tau) \quad (6.26)$$

のように、線形時変なフィルタ演算として (6.21) 式に等価な積分表現が得られる。ただし、 $\mathbf{h}(t, \tau) \equiv \Phi_x(t, \tau)\mathbf{k}(\tau)$  とおいた。



本論文では、あくまでも「線形フィルタ」の枠組みの中で最適なフィルタを求めることを目的としているため、非線形フィルタも含めたあらゆるフィルタの中では必ずしも最適なものにはならないことに注意する。しかしながら、振幅変動やピッチ変動が無視できる場合には、すべての信号過程は Gauss 過程になり線形フィルタが最適になることがわかっているので、振幅変動やピッチ変動が微小であるとした**仮定 1**が成り立っていれば、線形フィルタであっても準最適になると考えられる。

## 6.4 状態依存性雑音モデルに対する Kalman フィルタ解

以上のような線形動的システムの推定問題に対して、(6.21) 式のフィルタ方程式を決定する手続きは連続 Kalman フィルタ [89] として知られる。通常の連続 Kalman フィルタの理論では、信号過程が雑音入力の線形和として表される非同次形の微分方程式を状態空間モデルに仮定している。そのため、(6.12) 式の右辺第 2 項のように、雑音項が信号との積で与えられる非線形な状態依存性雑音を有する同次形の微分方程式に対しては適用できないという問題があり、このことが擬似周期信号のフィルタを設計する際の大きな妨げになっていた。Stratonovich 型の確率微分方程式に対しては状態依存性雑音を考慮した連続 Kalman フィルタが導かれているが [84]、伊藤型のものに対しての議論はなされていないため本節でこれを導出する。その導出手続きには通常の連続 Kalman フィルタ [89] と重複する部分があるが、従来にはなかった非線形な状態依存性雑音の影響を正確に扱わなければならないという必要性から詳細な式展開を行った。

### 6.4.1 Wiener-Hopf の積分方程式と直交原理

まず、線形最適解を求めるにあたって等価な条件式を以下に導く。(6.26) 式の  $\hat{\mathbf{x}}(t)$  が (6.22) 式を最小にするための必要十分条件は、 $\mathbf{h}(t, \tau)$  が Wiener-Hopf の積分方程式

$$E[\mathbf{x}(t)dz^*(\tau)] = \int_0^t \mathbf{h}(t, s)E[dz(s)dz^*(\tau)], \quad 0 \leq \tau < t \quad (6.27)$$

を満たすことである (証明は付録 6.8.3 参照)[78]。さらに、(6.27) 式の左辺から右辺を差し引き (6.26) 式の関係を使えば、

$$E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t))dz^*(\tau)] = \mathbf{0}, \quad 0 \leq \tau < t \quad (6.28)$$

が成り立つこと (直交射影定理) と等価になる。したがって**問題設定 4**を解くには、過去から現在までの観測データと推定誤差が互いに直交するようにフィルタ係数を決定すればよい。

## 6.4.2 フィルタ方程式の導出

(6.28) 式はまた、 $(\mathbf{x}(t) - \hat{\mathbf{x}}(t))$  項に対して無限小差分をとった形式

$$E[(d\mathbf{x}(t) - d\hat{\mathbf{x}}(t))dz^*(\tau)] = \mathbf{0}, \quad 0 \leq \tau < t \quad (6.29)$$

とも等価になる。これと (6.19), (6.20), (6.21) 式の関係を使うと、フィルタ係数の間に

$$F(t) = \Omega(t) - \mathbf{k}(t)\mathbf{1}^T \quad (6.30)$$

が成り立つことが導ける (付録 6.8.4 参照)。特に Kalman ゲイン  $\mathbf{k}(t)$  は、推定誤差の共分散

$$P(t) = E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t))(\mathbf{x}(t) - \hat{\mathbf{x}}(t))^*] \quad (6.31)$$

を用いて、

$$\mathbf{k}(t) = \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1} \quad (6.32)$$

と表せる (付録 6.8.5 参照)。したがって、(6.30), (6.32) 式を (6.21) 式に代入すれば、フィルタ方程式として

$$d\hat{\mathbf{x}}(t) = \left[ \Omega(t) - \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}\mathbf{1}^T \right] \hat{\mathbf{x}}(t) dt + \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1} dz(t) \quad (6.33)$$

を得る。両辺を  $dt$  で割り (6.16) 式の関係を用いれば、

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = \left[ \Omega(t) - \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}\mathbf{1}^T \right] \hat{\mathbf{x}}(t) + \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}y(t) \quad (6.34)$$

となり、通常の微分方程式の形式で表せる。 $P(t)$  に関しては、(6.20) 式から (6.33) 式を差し引き積分することによって、

$$\frac{dP(t)}{dt} = \Omega(t)P(t) + P(t)\Omega^*(t) - \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}\mathbf{1}^T P(t) + \Sigma^2(t) \quad (6.35)$$

の Riccati 方程式が得られる (付録 6.8.6 参照)。ただし、

$$\Sigma^2(t) = \text{diag} \left\{ \{ \sigma_{a1}^2(t) + \sigma_{b1}^2(t) + \sigma_\omega^2(t) \} \bar{x}_1^2(t), \{ \sigma_{a2}^2(t) + \sigma_{b2}^2(t) + 4\sigma_\omega^2(t) \} \bar{x}_2^2(t), \right. \\ \left. \dots, \{ \sigma_{aN}^2(t) + \sigma_{bN}^2(t) + N^2\sigma_\omega^2(t) \} \bar{x}_N^2(t) \right\} \quad (6.36)$$

である。

## 6.5 フィルタ方程式の解の性質

(6.34), (6.35) 式を解析的に解くことは一般に困難であるが、倍音間の干渉が無視できる程度に振幅変動とピッチ変動が微小である場合には、定常フィルタ解や時変フィルタの近似解が求められることを以下に示す。その結果からフィルタ方程式解の基本的な性質について議論する。

6.5.1 理想周期信号の場合

ここではまず、振幅変動とピッチ変動がともに無視できる理想周期信号の場合

$$\sigma_{an}(t), \sigma_{bn}(t), \sigma_{\omega}(t) \rightarrow 0 \quad (6.37)$$

を考える。このときは、(6.34) 式および (6.35) 式は解析的に解くことができ、それぞれ

$$\hat{x}(t) = P(t)e^{-\int_0^t \Omega^*(s)ds} \int_0^t \frac{1}{\sigma_v^2(\tau)} e^{\int_0^\tau \Omega^*(s)ds} \mathbf{1}y(\tau)d\tau \quad (6.38)$$

$$P(t) = e^{\int_0^t \Omega(s)ds} \left[ \int_0^t \frac{1}{\sigma_v^2(\tau)} e^{\int_0^\tau \Omega^*(s)ds} \mathbf{1}\mathbf{1}^T e^{\int_0^\tau \Omega(s)ds} d\tau \right]^{-1} e^{\int_0^t \Omega^*(s)ds} \quad (6.39)$$

となる (付録 6.8.7 参照)。ただし、簡単のため  $P(0) = \infty$  とし、

$$e^{\int_0^t \Omega(s)ds} \equiv \text{diag} \left\{ e^{\int_0^t \left\{ \frac{\dot{x}_1^2(s)}{2\bar{x}_1^2(s)} + j\bar{\omega}_0(s) \right\} ds}, \dots, e^{\int_0^t \left\{ \frac{\dot{x}_N^2(s)}{2\bar{x}_N^2(s)} + jN\bar{\omega}_0(s) \right\} ds} \right\} \quad (6.40)$$

とおいた。この場合、信号過程が Gauss 過程になるから、上式の解があらゆるフィルタの中で最適になることに注意する。

さらに、振幅やピッチ周波数の軌跡、観測雑音の分散が時間によらず一定 ( $\dot{x}_n^2(t) \rightarrow 0$ ,  $\bar{\omega}_0(t) \rightarrow \bar{\omega}_0$ ,  $\sigma_v^2(t) \rightarrow \sigma_v^2$ ) であるとし、時刻  $t$  が十分経過したときを考えれば、(6.39) 式中の積分の非対角成分は、倍音成分間の直交性により 0 に収束するため、

$$P(t) \approx \frac{\sigma_v^2}{t} I \quad (6.41)$$

となり、(6.38) 式のフィルタ解は、各倍音で近似的に

$$\hat{x}_n(t) \approx \frac{1}{t} \int_0^t e^{jn\bar{\omega}_0(t-\tau)} y(\tau)d\tau \quad (n = 1, \dots, N) \quad (6.42)$$

と表せる。このとき、観測信号  $y(t)$  から  $n$  次倍音成分の推定値  $\hat{x}_n(t)$  を与えるフィルタのインパルス応答は、時刻  $t$  において

$$h_n(\tau, t) \approx \begin{cases} \frac{1}{t} e^{jn\bar{\omega}_0\tau}, & 0 < \tau < t \\ 0, & \text{other} \end{cases} \quad (6.43)$$

と表せるから、その伝達関数は

$$H_n(\omega, t) \approx \frac{\sin\left\{\frac{1}{2}(\omega - n\bar{\omega}_0)t\right\}}{\frac{1}{2}(\omega - n\bar{\omega}_0)t} e^{-j\frac{1}{2}(\omega - n\bar{\omega}_0)t} \quad (6.44)$$

となる。常に  $H_n(n\bar{\omega}_0, t) = 1$  が成り立ち、 $t$  の増加にともなって占有帯域幅が狭くなるから、 $t \rightarrow \infty$  の極限で

$$H_n(\omega, t) \rightarrow \begin{cases} 1, & n\bar{\omega}_0 - \epsilon < \omega < n\bar{\omega}_0 + \epsilon \\ 0, & \text{other} \end{cases} \quad (6.45)$$

に収束する ( $\epsilon$  は任意に小さな正定数)。したがって、全倍音に対するフィルタの伝達関数としては (3.45) 式と同じ結論が得られる。

## 6.5.2 倍音間の干渉を無視したスカラー定常解

問題設定 3 を前提として、倍音間の干渉を無視し単一の倍音成分  $x_n(t)$  だけに着目した場合のフィルタ方程式解を考える。このとき、(6.34), (6.35) 式はそれぞれ、

$$\frac{d\hat{x}_n(t)}{dt} = \left[ \frac{\dot{\hat{x}}_n^2(t)}{2\bar{x}_n^2(t)} - \frac{1}{2}\sigma_n^2(t) + jn\bar{\omega}_0(t) - \frac{1}{\sigma_v^2(t)}p_n(t) \right] \hat{x}_n(t) + \frac{1}{\sigma_v^2(t)}p_n(t)y(t) \quad (6.46)$$

$$\frac{dp_n(t)}{dt} = \left[ \frac{\dot{\hat{x}}_n^2(t)}{\bar{x}_n^2(t)} - \sigma_n^2(t) \right] p_n(t) - \frac{1}{\sigma_v^2(t)}p_n^2(t) + \bar{x}_n^2(t)\sigma_n^2(t) \quad (6.47)$$

のスカラー方程式に縮退する。ただし、

$$\sigma_n^2(t) = \sigma_{an}^2(t) + \sigma_{bn}^2(t) + n^2\sigma_\omega^2(t) \quad (6.48)$$

とおいた。

さらに、各パラメータが一定 ( $\dot{\hat{x}}_n^2(t) \rightarrow 0$ ,  $\bar{x}_n^2(t) \rightarrow \bar{x}_n^2$ ,  $\bar{\omega}_0(t) \rightarrow \bar{\omega}_0$ ,  $\sigma_{an}^2(t) \rightarrow \sigma_{an}^2$ ,  $\sigma_{bn}^2(t) \rightarrow \sigma_{bn}^2$ ,  $\sigma_\omega^2(t) \rightarrow \sigma_\omega^2$ ,  $\sigma_v^2(t) \rightarrow \sigma_v^2$ ,  $\sigma_n^2(t) \rightarrow \sigma_n^2$ ) であるとすれば、(6.47) 式が  $\frac{dp_n(t)}{dt} = 0$  となる定常状態では、(6.47) 式の解は

$$p_n(t) = \sqrt{\bar{x}_n^2\sigma_n^2\sigma_v^2 + \frac{1}{4}\sigma_n^4\sigma_v^4 - \frac{1}{2}\sigma_n^2\sigma_v^2} \quad (6.49)$$

となるので、このとき (6.46) 式は

$$\frac{d\hat{x}_n(t)}{dt} = \left[ -\sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4} + jn\bar{\omega}_0 \right] \hat{x}_n(t) + \left[ \sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4} - \frac{1}{2}\sigma_n^2 \right] y(t) \quad (6.50)$$

と表せる。両辺を Laplace 変換し整理すると、観測値  $y(t)$  から  $n$  次倍音の推定値  $\hat{x}_n(t)$  を与えるフィルタの伝達関数として

$$H_n(s) = \frac{\hat{X}_n(s)}{Y(s)} = \frac{\sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4 - \frac{1}{2}\sigma_n^2}}{\sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4} - jn\bar{\omega}_0 + s} \quad (6.51)$$

を得る。 $\hat{X}_n(s), Y(s)$  は、それぞれ  $\hat{x}_n(t), y(t)$  の Laplace 変換を意味する。これは、3章の Wiener フィルタ解析の結論 (3.70) 式に一致する。

## 6.5.3 時変性を考慮したスカラー近似解

複数の倍音成分を含む擬似周期信号の場合でも、3.3.5項での問題設定 2 と問題設定 3 の関係で見たように、振幅変動やピッチ変動が微小であるなら、擬似周期信号を構成する倍音成分間の干渉が小さいため、近似的には各倍音を単独で推定する場合のフィルタ設計問題を考えればよく、前項の議論と同様に、(6.46), (6.47) 式のスカラー方程式を解くことに帰着

できる。また、音声や楽音などのように非定常性を無視できない信号であっても、短時間スケールでは振幅変動やピッチ変動がほぼ定常とみなせることが多い。このような場合を想定し、パラメータ  $(\bar{x}_n^2(t), \bar{\omega}_0(t), \sigma_{an}^2(t), \sigma_{bn}^2(t), \sigma_\omega^2(t), \sigma_v^2(t))$  が十分緩やかにしか変化しないと仮定する。

前項の議論と同様に準定常状態では、(6.47) 式を 0 にする解

$$p_n(t) = \sqrt{\bar{x}_n^2(t)\sigma_n^2(t)\sigma_v^2(t) + \frac{1}{4}\left\{\frac{\dot{\bar{x}}_n^2(t)}{\bar{x}_n^2(t)} - \sigma_n^2(t)\right\}^2\sigma_v^4(t) + \frac{1}{2}\left\{\frac{\dot{\bar{x}}_n^2(t)}{\bar{x}_n^2(t)} - \sigma_n^2(t)\right\}\sigma_v^2(t)} \quad (6.52)$$

で(6.46) 式が駆動されると考えられるので、このとき

$$\frac{d\hat{x}_n(t)}{dt} \approx \left[-\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t) + jn\bar{\omega}_0(t)\right]\hat{x}_n(t) + \frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)y(t) \quad (6.53)$$

となる。ただし、 $\bar{x}_n(t) = \sqrt{\bar{x}_n^2(t)}$  であり、振幅変動やピッチ変動が微小なため、 $\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t) \gg \frac{1}{2}\left|\frac{\dot{\bar{x}}_n^2(t)}{\bar{x}_n^2(t)} - \sigma_n^2(t)\right|$  となることを用いた。これは、(6.50) 式の係数が時変になったことを意味する。(3.71) 式の場合には、伝達関数の極  $s_n = -\frac{\bar{x}_n}{\sigma_v}\sigma_n + jn\bar{\omega}_0$  が時間に依存せず固定であったことから、(6.53) 式に対する伝達関数では、極が  $s_n(t) = -\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t) + jn\bar{\omega}_0(t)$  のように時刻に応じて変化していくと解釈できる。すなわち、

**性質 8:** 中心周波数が倍音周波数  $n\bar{\omega}_0(t)$  に追随し、占有帯域幅が  $\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)$  の値に比例して決まる時変特性を有する

ことがわかる。

## 6.6 倍音間の干渉を考慮したフィルタ方程式の数値解析

前節では、ベクトル形式で表されたフィルタ方程式を解析的に解くことが困難なため、単一倍音だけに着目したスカラー形式のフィルタ方程式解を求めるにとどめた。そこではパラメータの時変性を許容した解が得られてはいるが、解の基本的な性質に関しては Wiener フィルタ解析と同じ知見しか得られていないことに注意する。Kalman フィルタ方程式がベクトル形式で表現されていることの本質は、異なる倍音間どうしの干渉を含めてその挙動が厳密に記述されていることにある。この効果を詳細にみるために本節ではフィルタ方程式を数値的に解くことを試みる。倍音間の干渉を議論するにあたって、複雑な計算をできるだけ避けるため、パラメータの時変性は考慮せず定常解を調べることに限定する。

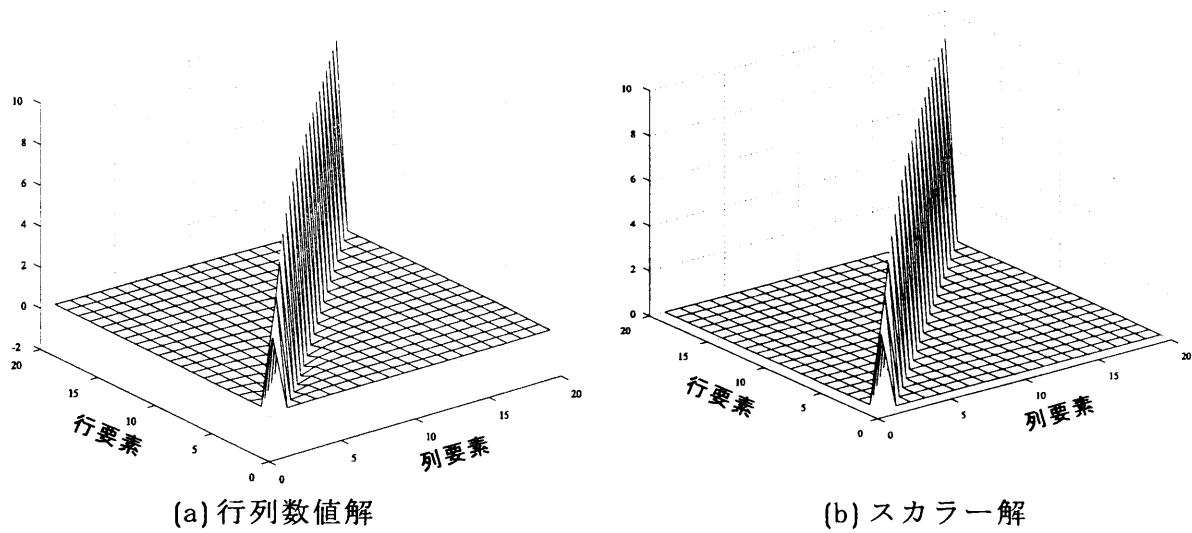


図 6.1 Riccati 方程式の解の比較

### 6.6.1 Riccati 方程式の数値解

(6.35) 式の Riccati 方程式の定常解は、 $\frac{dP(t)}{dt} = 0$  とした代数方程式

$$\Omega P + P \Omega^* - \frac{1}{\sigma_v^2} P \mathbf{1} \mathbf{1}^T P + \Sigma^2 = 0 \quad (6.54)$$

を解くことにより得られる。この場合も解析的に解くことは困難なため、数値的に求めた一例が図 6.1(a) である。一方、倍音間の干渉を無視した場合、上式は 2 次方程式となりその解は

$$p_n = \sqrt{\bar{x}_n^2 \sigma_n^2 \sigma_v^2 + \frac{1}{4} \sigma_n^4 \sigma_v^4 - \frac{1}{2} \sigma_n^2 \sigma_v^2} \quad (6.55)$$

と表せる。 $p_n$  は (6.54) 式の  $P$  の対角成分を表し、このときの非対角成分は仮定より 0 である。このスカラー解をプロットしたものが図 6.1(b) であり、行列解の結果 (同図 (a)) にほぼ一致する。他のパラメータ設定においても同様な傾向をもち、Riccati 方程式における倍音間の干渉はほとんど無視できることが確かめられる。

### 6.6.2 フィルタ方程式の数値解

ベクトル形式のフィルタ方程式 (6.34) 式に対するインパルス応答は、 $y(t) = \delta(t)$  としたときの解として形式的に

$$\mathbf{h}(t) = \frac{1}{\sigma_v^2} e^{[\Omega - \frac{1}{\sigma_v^2} P \mathbf{1} \mathbf{1}^T] t} P \mathbf{1} \quad (6.56)$$

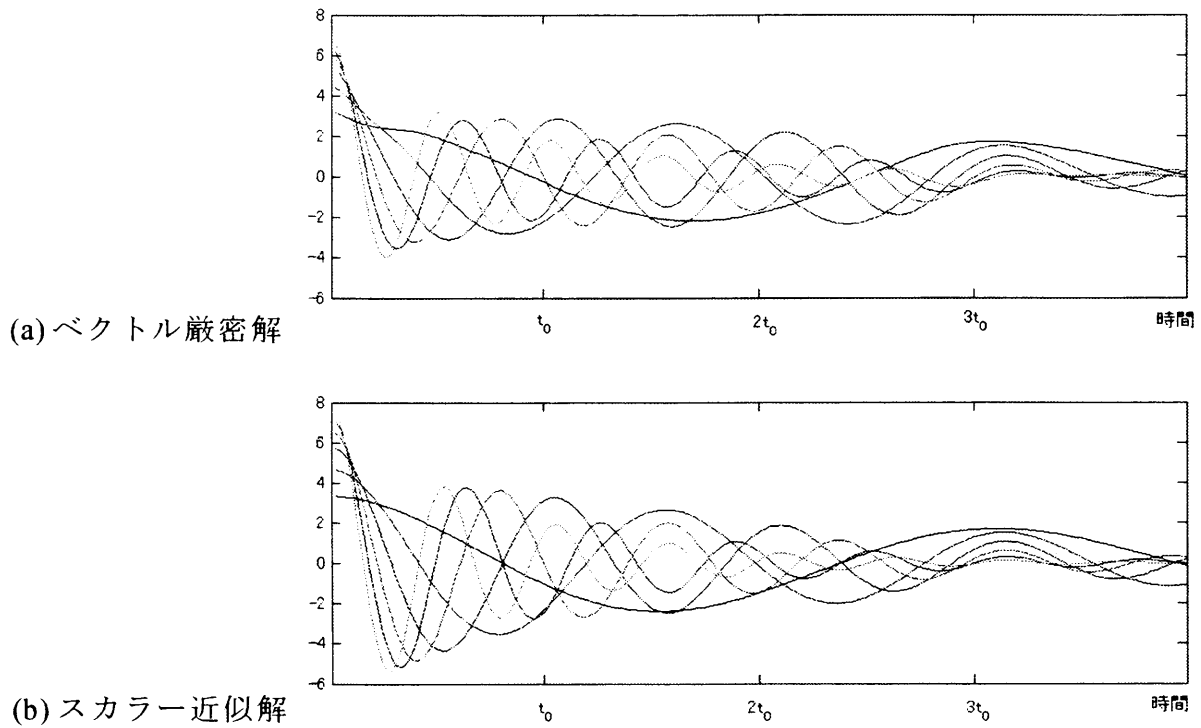


図 6.2 インパルス応答の比較

と表せる。ただし、遷移行列  $e^{[1]}$  は数値的にしか求められないことに注意する。倍音間の干渉を無視したときのスカラー方程式 (6.50) 式に対するインパルス応答は、(6.55) 式を用いて

$$h_n(t) = \frac{p_n}{\sigma_v^2} e^{(-\frac{1}{2}\sigma_n^2 - \frac{p_n}{\sigma_v^2} + jn\omega_0)t} \quad (6.57)$$

と求められる。両者の数値例をプロットしたものが図 6.2 であり、その違いはほとんどない。しかしながら、倍音間の干渉が無視できなくなるのは  $\sigma_n^2$  が大きいときであり、減衰カーブがきつくなるインパルス応答では高い周波数成分における違いを見分けることが困難になる。そこで伝達特性からの検証も行う。

ベクトル形式のフィルタ方程式 (6.34) 式に対する伝達関数は、両辺を Laplace 変換することにより、

$$\mathbf{H}(s) = \frac{1}{\sigma_v^2} [s\mathbf{I} - \mathbf{\Omega} + \frac{1}{\sigma_v^2} \mathbf{P}\mathbf{1}\mathbf{1}^T]^{-1} \mathbf{P}\mathbf{1} \quad (6.58)$$

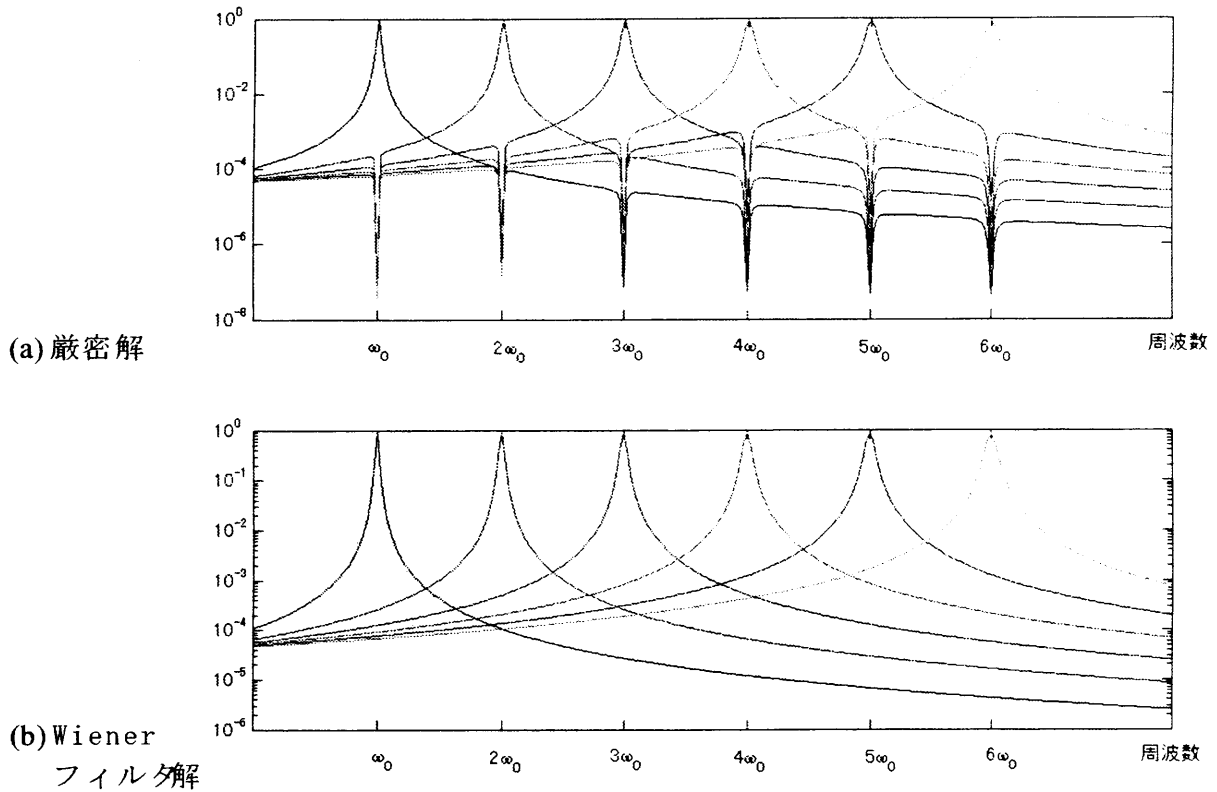


図 6.3 伝達特性の比較 (分散小)

と表せ、同様にスカラー形式 (6.50) 式に対する伝達関数は

$$H_n(s) = \frac{\frac{p_n}{\sigma_v^2}}{s - jn\bar{\omega}_0 + \frac{1}{2}\sigma_n^2 + \frac{p_n}{\sigma_v^2}} \quad (6.59)$$

と表せる。振幅変動およびピッチ変動の分散  $\sigma_a^2$ ,  $\sigma_\omega^2$  が小さい場合と大きい場合のそれぞれについて上式をプロットした結果が図 6.3, 6.4 である。変動分散が小さい場合でも、他の倍音の干渉を除去するためのノッチ特性が鋭く現れていることに注意する。これは 3 章での Wiener フィルタ解析では得られなかった新たな知見である。特に変動分散が大きい場合には、ノッチ帯域が広がり、倍音間の干渉を無視したスカラー解の特性と明確な違いが現れている。

## 6.7 本章のまとめ

本章では、振幅 / ピッチ変動に白色 Gauss 雑音過程を仮定し、その平均値と分散に時変性を許容した信号モデルに対する連続 Kalman フィルタの設計を行った。まず、時変性を



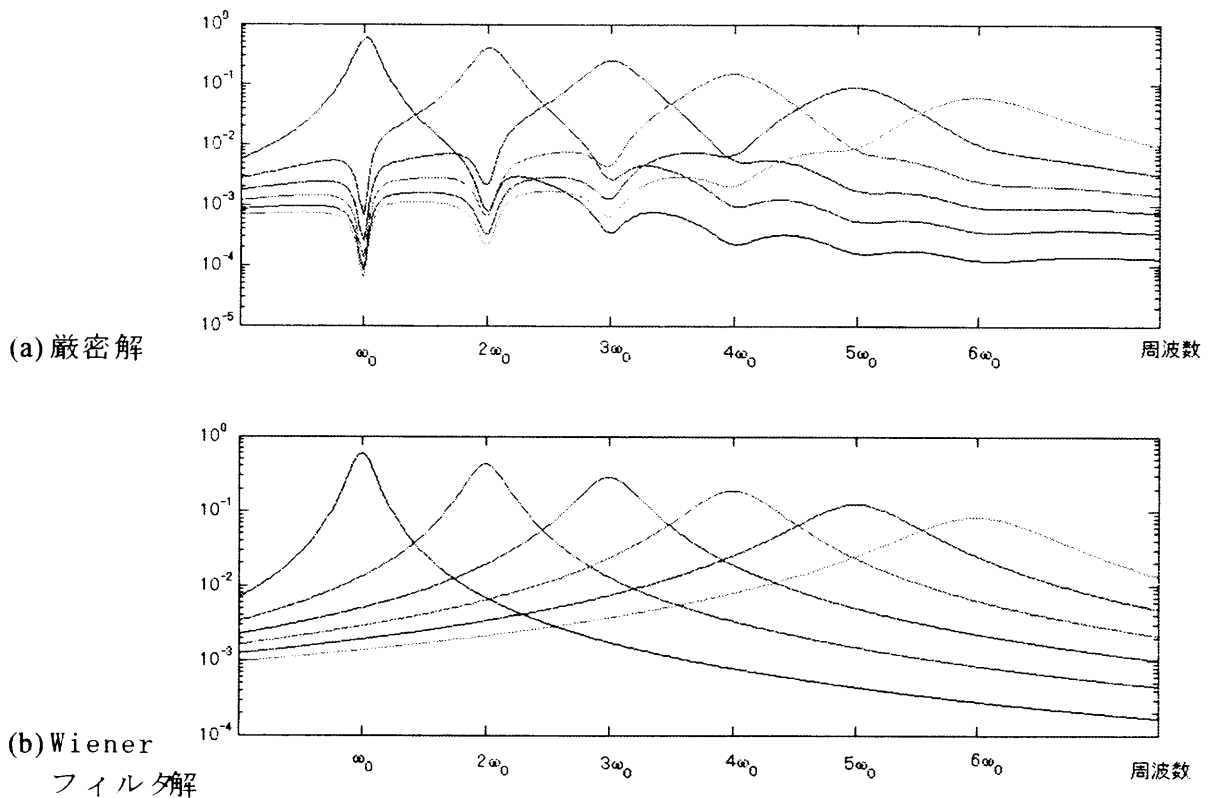


図 6.4 伝達特性の比較 (分散大)

有する擬似周期信号モデルに対して、伊藤型の確率微分方程式による表現形式 (状態空間モデル) を導出し、付加雑音を仮定した観測モデルに対する線形最適推定の問題をあらためて定義した (6.2 節)。推定値を与える連続 Kalman フィルタの方程式を導き (6.4 節)、理想周期信号や倍音間の干渉を無視したときのフィルタ方程式の定常解が Wiener フィルタ解と一致することを確かめた (6.5 節)。特に、倍音間の干渉を考慮した厳密解については、互いの干渉を排除するためのノッチ特性が現れることを新たに見出した (6.6 節)。

## 6.8 付録

### 6.8.1 Wiener 積分の諸性質

(3.80) 式の Wiener 過程を一般化して、任意の連続関数  $\sigma(t)$  による定常白色 Gauss 雑音  $\xi(t)$  の重み付き積分

$$w(t) = \int_0^t \sigma(s)\xi(s)ds \tag{6.60}$$

を考える。積分区間  $[0, t]$  を  $0 = t_0 < \dots < t_i < t_{i+1} < \dots < t_M = t$  のように  $M$  個の微小区間に分割すれば、各区間内の任意の点  $\forall \zeta_i \in [t_i, t_{i+1}]$  で  $\sigma(s)$  は一定値  $\sigma(\zeta_i)$  をとるものとみなせるから、上式は (3.81) 式の表記を使って

$$\begin{aligned} w(t) &\approx \sum_{i=0}^{M-1} \sigma(\zeta_i) \int_{t_i}^{t_{i+1}} \xi(s) ds \\ &= \sum_{i=0}^{M-1} \sigma(\zeta_i) \Delta\beta(t_i, t_{i+1}) \end{aligned} \quad (6.61)$$

と表せる。(3.83) 式より  $E[\Delta\beta(t_i, t_{i+1})\Delta\beta(t_j, t_{j+1})] = 0$  ( $i \neq j$ ) が成り立つから、上式は互いに統計的に独立な微小片  $\Delta\beta(t_i, t_{i+1})$  の重み付き加算を意味している。特に、分割幅  $\delta t = \max_{0 \leq i < M} (t_{i+1} - t_i)$  を無限小  $\delta t \rightarrow 0$  にしたときの極限として、

$$w(t) = \lim_{\delta t \rightarrow 0} \sum_{i=0}^{M-1} \sigma(\zeta_i) \Delta\beta(t_i, t_{i+1}) = \int_0^t \sigma(s) d\beta(s) \quad (6.62)$$

と書き直したものを Wiener 積分 [80] と呼ぶ。これは、(6.60) 式において

$$d\beta(s) = \xi(s) ds \quad (6.63)$$

あるいはそれと等価な表現として、 $\delta s \rightarrow 0$  の極限で

$$\Delta\beta(s, s + \delta s) \rightarrow d\beta(s) \quad (6.64)$$

とおいたことに相当する。(6.62) 式のように Wiener 積分を定義するのは、白色雑音が無限のパワーをもつため、(6.60) 式では Riemann 積分として定義できないという不都合さを解消するためである。

特に、(3.82), (3.83) 式より、Wiener 過程の微小片  $d\beta(t)$  に関して

$$\begin{aligned} E[d\beta(s)] &= 0, \\ E[d\beta(s)d\beta(u)] &= \begin{cases} ds, & s = u \\ 0, & s \neq u \end{cases} \end{aligned} \quad (6.65)$$

が成り立つ。この関係を使うと、Wiener 積分の期待値および分散は

$$\begin{aligned} E[w(t_a) - w(t_b)] &= \int_{t_b}^{t_a} \sigma(s) E[d\beta(s)] = 0, \\ E[(w(t_a) - w(t_b))^2] &= \int_{t_b}^{t_a} \int_{t_b}^{t_a} \sigma(s)\sigma(u) E[d\beta(s)d\beta(u)] = \int_{t_b}^{t_a} \sigma^2(s) ds \equiv \sigma_w^2(t_a, t_b) \end{aligned} \quad (6.66)$$

と求められる。無限小差分  $d\beta(s)$  に対して期待値演算を行うことの合理性は、(6.61) 式のように有限分割されたもの  $\Delta\beta(t_i, t_{i+1})$  に対して期待値操作をした後で、(6.62) 式のように極限操作を行うことの正当性が厳密に証明されていることにある [80]。

また、定義より Wiener 積分が Gauss 過程になることに注意すれば、これを指数部にもつ形式 (対数正規過程) に対する平均は

$$\begin{aligned} E[e^{k(w(t_a)-w(t_b))}] &= \frac{1}{\sqrt{2\pi\sigma_w^2(t_a, t_b)}} \int_{-\infty}^{\infty} e^{kW} e^{-\frac{W^2}{2\sigma_w^2(t_a, t_b)}} dW \\ &= \frac{1}{\sqrt{2\pi\sigma_w^2(t_a, t_b)}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_w^2(t_a, t_b)}\{W-k\sigma_w^2(t_a, t_b)\}^2} dW \cdot e^{\frac{1}{2}k^2\sigma_w^2(t_a, t_b)} \\ &= e^{\frac{1}{2}k^2\sigma_w^2(t_a, t_b)} = e^{\frac{1}{2}k^2 \int_{t_b}^{t_a} \sigma^2(s) ds} \end{aligned} \quad (6.67)$$

と求められる。ただし、 $(w(t_a) - w(t_b))$  の実現値を  $W$  とした。上式は、一般に係数  $k$  が複素数であっても成り立つことに注意する。Wiener 積分が Wiener 過程を一般化した概念であることは、(3.82), (3.85) 式の結論が、それぞれ (6.66), (6.67) 式で  $\sigma^2(s) \rightarrow 1, k \rightarrow \sigma$  とおいた場合に一致することからも明らかである。

### 6.8.2 伊藤積分と伊藤の公式

#### 伊藤積分

Wiener 積分の被積分項を確率過程  $f(t)$  に置き換えた積分形式

$$\eta(t) = \int_0^t f(s) d\beta(s) \quad (6.68)$$

を一般に確率積分と呼ぶ。この場合、積分区間  $[0, t]$  を  $0 = t_0 < \dots < t_i < t_{i+1} < \dots < t_M = t$  に分割したときの各微小区間  $[t_i, t_{i+1}]$  で、被積分項  $f(s)$  のサンプル点  $s = \zeta_i$  をどの時刻に選んだかによって、分割幅 ( $\delta t = \max_{0 \leq i < M} (t_{i+1} - t_i)$ ) を無限小にした極限での積分結果が異なってしまうという問題が生じる<sup>†</sup>。そのため、確率積分には複数の定義が存在する。特に本論文では、被積分項  $f(s)$  のサンプル点を各区間  $[t_i, t_{i+1}]$  の始点  $s = t_i$  に選んだ場合の

$$\eta(t) = \lim_{\delta t \rightarrow 0} \sum_{i=0}^{M-1} f(t_i) \Delta\beta(t_i, t_{i+1}) \quad (6.69)$$

と定義される「伊藤積分」[80]を採用する<sup>††</sup>。伊藤積分には、その定義から被積分項  $f(s)$  と微分子  $d\beta(s)$  が統計的に独立になり、

$$\begin{aligned} E[\eta(t)] &= \int_0^t E[f(s)] E[d\beta(s)] = 0 \\ E[|\eta(t)|^2] &= \int_0^t \int_0^t E[f(s)f^*(u) d\beta(s) d\beta(u)] = \int_0^t E[|f(s)|^2] ds \end{aligned} \quad (6.70)$$

<sup>†</sup> Wiener 積分では、(6.61) 式において、確定関数である被積分項  $\sigma(s)$  のサンプル点  $s = \zeta_i$  を微小区間内  $[t_i, t_{i+1}]$  のどの時刻に選んでも、分割幅を無限小にした極限では同じ積分値を与える。

<sup>††</sup> これに対し、サンプル点を中点  $s = \frac{t_i + t_{i+1}}{2}$  に選んだ場合は、「Stratonovich 型積分 (あるいは対称化確率積分)」と呼ばれる。

が成り立つなど、計算が簡単になるという利点がある。

### 伊藤の公式

$L$  個の互いに統計的に独立な Wiener 積分

$$w_i(t) = \int_0^t \sigma_i(s) d\beta_i(s) \quad (i = 1, \dots, L) \quad (6.71)$$

を媒介変数にもつ 2 階微分可能な汎関数  $f(t, w_1(t), \dots, w_L(t))$  に対して、

$$\begin{aligned} df(t, w_1(t), \dots, w_L(t)) &= [f_t(t, w_1(t), \dots, w_L(t)) + \frac{1}{2} \sum_{i=1}^L \sigma_i^2(t) f_{w_i w_i}(t, w_1(t), \dots, w_L(t))] dt \\ &\quad + \sum_{i=1}^L \sigma_i(t) f_{w_i}(t, w_1(t), \dots, w_L(t)) d\beta_i(t) \end{aligned} \quad (6.72)$$

あるいは、その積分形式として

$$\begin{aligned} f(t, w_1(t), \dots, w_L(t)) &= f(0, w_1(0), \dots, w_L(0)) \\ &\quad + \int_0^t [f_t(s, w_1(s), \dots, w_L(s)) + \frac{1}{2} \sum_{i=1}^L \sigma_i^2(s) f_{w_i w_i}(s, w_1(s), \dots, w_L(s))] ds \\ &\quad + \sum_{i=1}^L \int_0^t \sigma_i(s) f_{w_i}(s, w_1(s), \dots, w_L(s)) d\beta_i(s) \end{aligned} \quad (6.73)$$

が伊藤積分の意味で成り立つ。ただし、 $f$  の下付き添字はその変数に関する 1 階もしくは 2 階の偏微分を意味する。これは「伊藤の公式 (あるいは伊藤の連鎖則)」と呼ばれる<sup>†</sup>。

この証明の概略は以下の通りである。汎関数  $f(t, w_1, \dots, w_L)$  に対する微小増分が、Taylor 展開により

$$\begin{aligned} &f(t + \delta t, w_1(t) + \delta w_1(t), \dots, w_L(t) + \delta w_L(t)) - f(t, w_1(t), \dots, w_L(t)) \\ &= f_t(t, w_1(t), \dots, w_L(t)) \delta t + \sum_{i=1}^L f_{w_i}(t, w_1(t), \dots, w_L(t)) \delta w_i(t) \\ &\quad + \sum_{i=1}^L f_{tw_i}(t, w_1(t), \dots, w_L(t)) \delta t \delta w_j(t) \\ &\quad + \frac{1}{2} \sum_{i=1}^L \sum_{j=1}^L f_{w_i w_j}(t, w_1(t), \dots, w_L(t)) \delta w_i(t) \delta w_j(t) + r(t) \end{aligned} \quad (6.74)$$

と表せ、(6.71) 式の変分形式として

$$\delta w_i(t) = \sigma_i(t) \delta \beta_i(t) \quad (6.75)$$

<sup>†</sup> 本来の「伊藤の公式」[80] は、(6.71) 式を確率微分方程式とすることによる一般的な形式として与えられている。

が成り立つことに着目する。  $r(t)$  は高次の残差項を意味する。このとき、伊藤積分の定義および (6.65) 式を考慮すれば、(6.74) 式の各項の期待値および分散が、右辺第 3 項については、

$$\begin{aligned} E[f_{tw_i}(t, w_1(t), \dots, w_L(t))\delta t \delta w_j(t)] &= E[f_{tw_i}(t, w_1(t), \dots, w_L(t))\delta t]E[\delta w_i(t)] = 0 \\ E[|f_{tw_i}(t, w_1(t), \dots, w_L(t))\delta t \delta w_j(t)|^2] &= E[|f_{tw_i}(t, w_1(t), \dots, w_L(t))|^2]E[(\delta w_i(t))^2](\delta t)^2 \\ &= E[|f_{tw_i}(t, w_1(t), \dots, w_L(t))|^2](\delta t)^3 \end{aligned} \quad (6.76)$$

となり、右辺第 4 項については、  $i \neq j$  のとき

$$\begin{aligned} E[f_{w_i w_j}(t, w_1(t), \dots, w_L(t))\delta w_i(t)\delta w_j(t)] &= E[f_{w_i w_j}(t, w_1(t), \dots, w_L(t))]E[\delta w_i(t)\delta w_j(t)] = 0 \\ E[|f_{w_i w_j}(t, w_1(t), \dots, w_L(t))\delta w_i(t)\delta w_j(t)|^2] &= E[|f_{w_i w_j}(t, w_1(t), \dots, w_L(t))|^2]E[(\delta w_i(t))^2]E[(\delta w_j(t))^2] \\ &= E[|f_{w_i w_j}(t, w_1(t), \dots, w_L(t))|^2](\delta t)^2 \end{aligned} \quad (6.77)$$

$i = j$  のとき

$$\begin{aligned} E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))(\delta w_i(t))^2] &= E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))]E[(\delta w_i(t))^2] = E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))]\delta t \\ E[|f_{w_i w_i}(t, w_1(t), \dots, w_L(t))(\delta w_i(t))^2 - E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))]\delta t|^2] &= E[|f_{w_i w_i}(t, w_1(t), \dots, w_L(t))|^2]E[(\delta w_i(t))^4] - E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))]^2(\delta t)^2 \\ &= \{3E[|f_{w_i w_i}(t, w_1(t), \dots, w_L(t))|^2] - E[f_{w_i w_i}(t, w_1(t), \dots, w_L(t))]^2\}(\delta t)^2 \end{aligned} \quad (6.78)$$

となる。  $\delta t \rightarrow 0$  で、いずれも分散が 2 乗以上の速さで 0 に収束するから、各項は決定論的に

$$\begin{aligned} f_{tw_i}(t, w_1(t), \dots, w_L(t))\delta t \delta w_j(t) &\rightarrow 0 \\ f_{w_i w_j}(t, w_1(t), \dots, w_L(t))\delta w_i(t)\delta w_j(t) &\rightarrow \begin{cases} f_{w_i w_j}(t, w_1(t), \dots, w_L(t))\delta t, & i = j \\ 0, & i \neq j \end{cases} \end{aligned} \quad (6.79)$$

とみなせると考えてよい<sup>†</sup>。同様に、これ以外の高次項  $r(t)$  も無視できる。以上を (6.74) 式に代入し整理すると、  $\delta t \rightarrow 0$  の極限での微分形式として (6.72) 式が得られる。

<sup>†</sup>特に、  $(\delta w_i(t))^2 \rightarrow \delta t$  となる点が通常の微分方程式との根本的な違いになる。

## 6.8.3 Wiener-Hopfの積分方程式

(6.22) 式に (6.26) 式を代入し、これを  $\mathbf{h}(t, \tau)$  の関数

$$J(\mathbf{h}(t, \tau)) = E[\|\mathbf{x}(t) - \int_0^t \mathbf{h}(t, \tau) dz(\tau)\|^2] \quad (6.80)$$

とみなしたときの  $\mathbf{h}(t, \tau)$  に関する変分問題を考える。 $\mathbf{h}(t, \tau) = \mathbf{h}_0(t, \tau)$  のとき  $J(\mathbf{h}(t, \tau))$  が最小になるとすると、任意の関数  $\Delta \mathbf{h}(t, \tau)$  に対して

$$\left. \frac{\partial J(\mathbf{h}_0(t, \tau) + \varepsilon \Delta \mathbf{h}(t, \tau))}{\partial \varepsilon} \right|_{\varepsilon=0} = 0 \quad (6.81)$$

が成立しなければならない。これを計算すれば、

$$\Re \left[ \int_0^t \Delta \mathbf{h}^*(t, \tau) \left\{ E[\mathbf{x}(t) dz^*(\tau)] - \int_0^t \mathbf{h}_0(t, s) E[dz(s) dz^*(\tau)] \right\} \right] = 0 \quad (6.82)$$

となるので ( $\Re[\cdot]$  は実数部を抜き出す演算の意)、任意の  $\Delta \mathbf{h}(t, \tau)$  について上式が成り立つ条件として (6.27) 式が導かれる。逆に、 $\mathbf{h}(t, \tau) = \mathbf{h}_0(t, \tau)$  が (6.27) 式を満たすとき、常に

$$J(\mathbf{h}_0(t, \tau) + \varepsilon \Delta \mathbf{h}(t, \tau)) - J(\mathbf{h}_0(t, \tau)) = \varepsilon^2 E[\|\int_0^t \Delta \mathbf{h}_0(t, \tau) dz(\tau)\|^2] \geq 0 \quad (6.83)$$

が成り立つ。したがって、(6.27) 式が (6.22) 式を最小にする必要十分条件になる。

## 6.8.4 フィルタ係数の関係式

(6.29) 式に (6.20), (6.21) 式を代入すれば

$$\begin{aligned} (6.29) \text{式左辺} &= \Omega(t) E[\mathbf{x}(t) dz^*(\tau)] dt + \Sigma_a(t) E[dB_a(t) \mathbf{x}(t) dz^*(\tau)] \\ &\quad + j \Sigma_b(t) E[dB_b(t) \mathbf{x}(t) dz^*(\tau)] + j \Sigma_\omega(t) E[d\beta_\omega(t) \mathbf{x}(t) dz^*(\tau)] \\ &\quad - F(t) E[\hat{\mathbf{x}}(t) dz^*(\tau)] dt - \mathbf{k}(t) E[dz(t) dz^*(\tau)] \end{aligned}$$

となる。特に第2項については、 $t > \tau$  に注意して伊藤積分の性質を考慮すると

$$\begin{aligned} E[dB_a(t) \mathbf{x}(t) dz^*(\tau)] &= E[dB_a(t) \mathbf{x}(t) \mathbf{x}^*(\tau)] \mathbf{1} d\tau + \sigma_v(\tau) E[dB_a(t) \mathbf{x}(t) d\beta_v(\tau)] \\ &= E[dB_a(t)] E[\mathbf{x}(t) \mathbf{x}^*(\tau)] \mathbf{1} d\tau + \sigma_v(\tau) E[dB_a(t)] E[\mathbf{x}(t) d\beta_v(\tau)] \\ &= \mathbf{0} \end{aligned}$$

が成り立つ。第3,4項についても同様である。また第5項については、(6.28) 式より

$$E[\hat{\mathbf{x}}(t) dz^*(\tau)] = E[\mathbf{x}(t) dz^*(\tau)]$$

となる。第6項については、(6.19)式を代入し、 $t > \tau$ に注意して伊藤積分の性質を考慮すれば

$$\begin{aligned} E[dz(t)dz^*(\tau)] &= \mathbf{1}^T E[\mathbf{x}(t)dz^*(\tau)]dt + \sigma_v(t)E[d\beta_v(t)dz^*(\tau)] \\ &= \mathbf{1}^T E[\mathbf{x}(t)dz^*(\tau)]dt + \sigma_v(t)E[d\beta_v(t)]E[dz^*(\tau)] \\ &= \mathbf{1}^T E[\mathbf{x}(t)dz^*(\tau)]dt \end{aligned}$$

となる。以上より、

$$(6.29)\text{式左辺} = [\Omega(t) - F(t) - \mathbf{k}(t)\mathbf{1}^T]E[\mathbf{x}(t)dz^*(\tau)]dt$$

と整理できる。さらに、(6.19)式を代入し、信号過程 $\mathbf{x}(t)$ と観測雑音 $d\beta_v(\tau)$ が互いに独立であることを考慮すれば

$$\begin{aligned} (6.29)\text{式左辺} &= [\Omega(t) - F(t) - \mathbf{k}(t)\mathbf{1}^T]\{E[\mathbf{x}(t)\mathbf{x}^*(\tau)]\mathbf{1}d\tau + \sigma_v(\tau)E[\mathbf{x}(t)]E[d\beta_v(\tau)]\}dt \\ &= [\Omega(t) - F(t) - \mathbf{k}(t)\mathbf{1}^T]E[\mathbf{x}(t)\mathbf{x}^*(\tau)]\mathbf{1}d\tau dt \end{aligned} \quad (6.84)$$

となる。

したがって、

$$\Omega(t) - F(t) - \mathbf{k}(t)\mathbf{1}^T = 0 \quad (6.85)$$

が(6.29)式を満たす十分条件になる。これが必要条件になることも以下のように示せる。(6.84)式 $= 0$ が $0 \leq \tau < t$ で成り立つことから、右から $\mathbf{1}^T E[\mathbf{x}(\tau)\mathbf{x}^*(t)]$ を乗じて $\tau$ について積分すると、

$$[\Omega(t) - F(t) - \mathbf{k}(t)\mathbf{1}^T]U(t) = 0 \quad (6.86)$$

が得られる。ただし、

$$\begin{aligned} U(t) &= \int_0^t X(t, \tau)X^*(t, \tau)d\tau, \\ X(t, \tau) &= E[\mathbf{x}(t)\mathbf{x}^*(\tau)]\mathbf{1} \end{aligned} \quad (6.87)$$

とおいた。ここで、 $\mathbf{a} = [a_1, \dots, a_N]^T$ を任意の $N$ 次元列ベクトルとし、 $U(t)$ の2次形式

$$\begin{aligned} \mathbf{a}^*U(t)\mathbf{a} &= \int_0^t \mathbf{a}^*X(t, \tau)X^*(t, \tau)\mathbf{a}d\tau \\ &= \int_0^t |\mathbf{1}^T E[\mathbf{x}(\tau)\mathbf{x}^*(t)]\mathbf{a}|^2 d\tau \\ &= \int_0^t |\mathbf{1}^T \text{diag}\{R_{x_1}(\tau, t), \dots, R_{x_N}(\tau, t)\}\mathbf{a}|^2 d\tau \\ &= \int_0^t \left| \sum_{n=1}^N a_n R_{x_n}(\tau, t) \right|^2 d\tau \geq 0 \end{aligned} \quad (6.88)$$

に着目する。上式の 2 段目から 3 段目は、仮定 2 より各倍音間が無相関であることを用いている。 $|\sum_{n=1}^N a_n R_{xn}(\tau, t)|^2 \geq 0$  なので、 $\mathbf{a}^* U(t) \mathbf{a} = 0$  となるためには、

$$\sum_{n=1}^N a_n R_{xn}(\tau, t) = 0, \quad 0 \leq \tau < t \quad (6.89)$$

でなければならない。これを満たすのは、(6.5) 式より  $\{R_{x1}(\tau, t), \dots, R_{xN}(\tau, t)\}$  が線形独立なので、 $\mathbf{a} = \mathbf{0}$  のときだけに限られる。したがって、 $U(t)$  は正定値であり逆行列が存在するので、それを (6.86) 式の右から乗じることによって (6.85) 式が導かれる。

### 6.8.5 Kalman ゲインと誤差共分散の関係式

(6.26), (6.19) 式より、

$$\begin{aligned} E[\hat{\mathbf{x}}(t) d\beta_v^*(\tau)] &= \int_0^t \Phi_x(t, s) \mathbf{k}(s) E[dz(s) d\beta_v^*(\tau)] \\ &= \int_0^t \Phi_x(t, s) \mathbf{k}(s) \left\{ \mathbf{1}^T E[\mathbf{x}(s) d\beta_v^*(\tau)] dt + \sigma_v(s) E[d\beta_v(s) d\beta_v^*(\tau)] \right\} \end{aligned}$$

となり、 $\mathbf{x}(t)$  と  $d\beta_v(\tau)$  が互いに独立であること、および (6.65) 式を考慮すると、

$$E[\hat{\mathbf{x}}(t) d\beta_v^*(\tau)] = \sigma_v(\tau) \Phi_x(t, \tau) \mathbf{k}(\tau) d\tau \quad (6.90)$$

と表せる。また、(6.26), (6.28) 式より

$$E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) \hat{\mathbf{x}}^*(\tau)] = \int_0^\tau E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) dz^*(s)] \mathbf{k}^*(s) \Phi_x^*(t, s) \quad (6.91)$$

$$= 0 \quad (6.92)$$

となる。(6.28) 式に (6.19) 式を代入し、(6.90), (6.91) 式を用いると、

$$\begin{aligned} (6.28) \text{式左辺} &= E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) \mathbf{x}^*(\tau)] \mathbf{1} d\tau - \sigma_v(\tau) E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) d\beta_v^*(\tau)] \\ &= E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) (\mathbf{x}(\tau) - \hat{\mathbf{x}}(\tau))^*] \mathbf{1} d\tau - \sigma_v^2(\tau) \Phi_x(t, \tau) \mathbf{k}(\tau) d\tau \end{aligned}$$

と表せる。したがって、 $\tau \rightarrow t$  とすると

$$P(t) \mathbf{1} dt - \sigma_v^2(t) \mathbf{k}(t) dt = 0$$

となり、(6.32) 式が得られる。

### 6.8.6 Riccati 方程式

推定誤差

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$$



の無限小差分

$$d\mathbf{e}(t) = d\mathbf{x}(t) - d\hat{\mathbf{x}}(t)$$

に対して、(6.20)、(6.33) 式および (6.19) 式を代入し整理すれば、

$$d\mathbf{e}(t) = \left[ \Omega(t) - \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1} \mathbf{1}^T \right] \mathbf{e}(t) dt + d\mathbf{w}_e(t) \quad (6.93)$$

を得る。ただし、

$$d\mathbf{w}_e(t) = \left[ \Sigma_a(t) dB_a(t) + j\Sigma_b(t) dB_b(t) + j\Sigma_\omega(t) d\beta_\omega(t) \right] \mathbf{x}(t) - \frac{1}{\sigma_v(t)} P(t) \mathbf{1} d\beta_v(t)$$

とおいた。(6.93) 式右辺第1項に対する遷移行列、すなわち

$$\begin{aligned} \frac{\partial \Phi_e(t, \tau)}{\partial t} &= \left[ \Omega(t) - \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1} \mathbf{1}^T \right] \Phi_e(t, \tau), \\ \Phi_e(t, t) &= I \end{aligned} \quad (6.94)$$

を満たす  $\Phi_e(t, \tau)$  を用いれば、一般解は

$$\mathbf{e}(t) = \Phi_e(t, 0) \mathbf{e}(0) + \int_0^t \Phi_e(t, \tau) d\mathbf{w}_e(\tau)$$

と表せる。これを (6.31) 式に代入し、

$$\begin{aligned} E[\mathbf{e}(0) d\mathbf{w}_e^*(\tau)] &= E[\mathbf{e}(0) \mathbf{x}^*(t)] E \left[ \Sigma_a(t) dB_a(t) + j\Sigma_b(t) dB_b(t) + j\Sigma_\omega(t) d\beta_\omega(t) \right]^* \\ &\quad - \frac{1}{\sigma_v(t)} E[\mathbf{e}(0)] E[d\beta_v(t)] \mathbf{1}^T P(t) \\ &= 0 \end{aligned}$$

を考慮すれば、

$$P(t) = \Phi_e(t, 0) P(0) \Phi_e^*(t, 0) + \int_0^t \int_0^t \Phi_e(t, \tau) E[d\mathbf{w}_e(\tau) d\mathbf{w}_e^*(s)] \Phi_e^*(t, s) \quad (6.95)$$

を得る。

ところで、 $dB_a(t), dB_b(t), d\beta_\omega(t)$  と  $d\beta_v(t)$  が互いに無相関であることを考慮すると (仮定 2)、

$$\begin{aligned} E[d\mathbf{w}_e(\tau) d\mathbf{w}_e^*(s)] &= E \left[ [\Sigma_a(\tau) dB_a(\tau) + j\Sigma_b(\tau) dB_b(\tau) + j\Sigma_\omega(\tau) d\beta_\omega(\tau)] \mathbf{x}(\tau) \mathbf{x}^*(s) \right. \\ &\quad \left. [\Sigma_a(s) dB_a(s) - j\Sigma_b(s) dB_b(s) - j\Sigma_\omega(s) d\beta_\omega(s)] \right] \\ &\quad + \frac{1}{\sigma_v(\tau)} P(\tau) \mathbf{1} E[d\beta_v(\tau) d\beta_v(s)] \mathbf{1}^T P(s) \frac{1}{\sigma_v(s)} \end{aligned} \quad (6.96)$$

が成り立つ。特に  $\tau > s$  のとき、伊藤積分の定義より

$$\begin{aligned} E[d\mathbf{w}_e(\tau)d\mathbf{w}_e^*(s)] &= E\left[\Sigma_a(\tau)dB_a(\tau) + j\Sigma_b(\tau)dB_b(\tau) + j\Sigma_\omega(\tau)d\beta_\omega(\tau)\right]E\left[\mathbf{x}(\tau)\mathbf{x}^*(s)\right. \\ &\quad \left.[\Sigma_a(s)dB_a(s) - j\Sigma_b(s)dB_b(s) - j\Sigma_\omega(s)d\beta_\omega(s)]\right] \\ &\quad + \frac{1}{\sigma_v(\tau)}P(\tau)\mathbf{1}E[d\beta_v(\tau)]E[d\beta_v(s)]\mathbf{1}^T P(s)\frac{1}{\sigma_v(s)} \\ &= 0 \end{aligned}$$

同様に  $\tau < s$  のときも

$$E[d\mathbf{w}_e(\tau)d\mathbf{w}_e^*(s)] = 0$$

となる。一方  $\tau = s$  のときは、(6.96) 式右辺第 1 項の  $n$  行  $m$  列成分

$$\begin{aligned} (6.96) \text{ 式第 1 項 } (n, m) \text{ 成分} &= E\left[\{\sigma_{an}(\tau)d\beta_{an}(\tau) + j\sigma_{bn}(\tau)d\beta_{bn}(\tau) + jn\sigma_\omega(\tau)d\beta_\omega(\tau)\}\right. \\ &\quad \left.\cdot x_n(\tau)x_m^*(\tau)\{\sigma_{am}(\tau)d\beta_{am}(\tau) - j\sigma_{bm}(\tau)d\beta_{bm}(\tau) - jm\sigma_\omega(\tau)d\beta_\omega(\tau)\}\right] \end{aligned}$$

において、伊藤積分の定義より

$$\begin{aligned} (6.96) \text{ 式第 1 項 } (n, m) \text{ 成分} &= E\left[\{\sigma_{an}(\tau)d\beta_{an}(\tau) + j\sigma_{bn}(\tau)d\beta_{bn}(\tau) + jn\sigma_\omega(\tau)d\beta_\omega(\tau)\}\right. \\ &\quad \left.\cdot \{\sigma_{am}(\tau)d\beta_{am}(\tau) - j\sigma_{bm}(\tau)d\beta_{bm}(\tau) - jm\sigma_\omega(\tau)d\beta_\omega(\tau)\}\right]E[x_n(\tau)x_m^*(\tau)] \end{aligned}$$

が成り立つ。したがって  $n \neq m$  のとき、 $E[x_n(\tau)x_m^*(\tau)] = 0$  より (仮定 2)、

$$(6.96) \text{ 式第 1 項 } (n, m) \text{ 成分} = 0$$

となる。また、 $d\beta_{an}(\tau), \beta_{bn}(\tau), d\beta_\omega(\tau)$  が互いに無相関であること (仮定 2) および (6.65) 式より、 $n = m$  のとき

$$\begin{aligned} (6.96) \text{ 式第 1 項 } (n, m) \text{ 成分} &= \left\{\sigma_{an}^2(\tau)E[(d\beta_{an}(\tau))^2] + \sigma_{bn}^2(\tau)E[(d\beta_{bn}(\tau))^2]\right. \\ &\quad \left.+ n^2\sigma_\omega^2(\tau)E[(d\beta_\omega(\tau))^2]\right\}E[|x_n(\tau)|^2] \\ &= \{\sigma_{an}^2(\tau) + \sigma_{bn}^2(\tau) + n^2\sigma_\omega^2(\tau)\}\bar{x}_n^2(\tau)d\tau \end{aligned}$$

となる。(6.96) 式右辺第 2 項については、 $\tau = s$  のとき

$$(6.96) \text{ 式第 2 項} = \frac{1}{\sigma_v^2(\tau)}P(\tau)\mathbf{1}\mathbf{1}^T P(\tau)d\tau$$

と表せる。

以上を (6.95) 式に代入すると、

$$P(t) = \Phi_e(t, 0)P(0)\Phi_e^*(t, 0) + \int_0^t \Phi_e(t, \tau) \left[ \frac{1}{\sigma_v^2(\tau)} P(\tau) \mathbf{1} \mathbf{1}^T P(\tau) + \Sigma^2(\tau) \right] \Phi_e^*(t, \tau) d\tau$$

を得る。両辺を  $t$  に関して微分すると、

$$\begin{aligned} \frac{dP(t)}{dt} &= \frac{\partial \Phi_e(t, 0)}{\partial t} P(0) \Phi_e^*(t, 0) + \int_0^t \frac{\partial \Phi_e(t, \tau)}{\partial t} \left[ \frac{1}{\sigma_v^2(\tau)} P(\tau) \mathbf{1} \mathbf{1}^T P(\tau) + \Sigma^2(\tau) \right] \Phi_e^*(t, \tau) d\tau \\ &+ \Phi_e(t, 0) P(0) \frac{\partial \Phi_e^*(t, 0)}{\partial t} + \int_0^t \Phi_e(t, \tau) \left[ \frac{1}{\sigma_v^2(\tau)} P(\tau) \mathbf{1} \mathbf{1}^T P(\tau) + \Sigma^2(\tau) \right] \frac{\partial \Phi_e^*(t, \tau)}{\partial t} d\tau \\ &+ \Phi_e(t, t) \left[ \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1} \mathbf{1}^T P(t) + \Sigma^2(t) \right] \Phi_e^*(t, t) \end{aligned}$$

となるので、(6.94) 式の関係を使って整理すれば、(6.35) 式の Riccati 方程式が得られる。

### 6.8.7 フィルタ方程式の解析解

(6.37) 式の仮定より (6.35) 式は

$$\begin{aligned} \frac{dP(t)}{dt} &= \Omega(t)P(t) + P(t)\Omega^*(t) - \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1} \mathbf{1}^T P(t) \\ \text{where } \Omega(t) &= \text{diag} \left\{ \frac{\dot{\bar{x}}_1^2(t)}{2\bar{x}_1^2(t)} + j\bar{\omega}_0(t), \dots, \frac{\dot{\bar{x}}_N^2(t)}{2\bar{x}_N^2(t)} + jN\bar{\omega}_0(t) \right\} \end{aligned}$$

と表せる。上式を満たす  $P(t)$  を用いて

$$P_0(t) = e^{-\int_0^t \Omega(s) ds} P(t) e^{-\int_0^t \Omega^*(s) ds} \quad (6.97)$$

を新たに定義する。両辺を  $t$  で微分し、上式を繰り返し用いれば、

$$\frac{dP_0(t)}{dt} = \frac{d}{dt} \left\{ e^{-\int_0^t \Omega(s) ds} \right\} P(t) e^{-\int_0^t \Omega^*(s) ds} + e^{-\int_0^t \Omega(s) ds} P(t) \frac{d}{dt} \left\{ e^{-\int_0^t \Omega^*(s) ds} \right\} \quad (6.98)$$

$$+ e^{-\int_0^t \Omega(s) ds} \frac{dP(t)}{dt} e^{-\int_0^t \Omega^*(s) ds} \quad (6.99)$$

$$= -\Omega(t)P_0(t) - P_0(t)\Omega^*(t) + e^{-\int_0^t \Omega(s) ds} \left[ \Omega(t)P(t) + P(t)\Omega^*(t) \right. \quad (6.100)$$

$$\left. - \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1} \mathbf{1}^T P(t) \right] e^{-\int_0^t \Omega^*(s) ds} \quad (6.101)$$

$$= -\frac{1}{\sigma_v^2(t)} e^{-\int_0^t \Omega(s) ds} P(t) \mathbf{1} \mathbf{1}^T P(t) e^{-\int_0^t \Omega^*(s) ds} \quad (6.102)$$

$$= -\frac{1}{\sigma_v^2(t)} P_0(t) e^{\int_0^t \Omega^*(s) ds} \mathbf{1} \mathbf{1}^T e^{\int_0^t \Omega(s) ds} P_0(t) \quad (6.103)$$

の関係が得られる。上式の解は、微分方程式

$$\frac{d}{dt} \begin{bmatrix} E(t) \\ D(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \frac{1}{\sigma_v^2(t)} e^{\int_0^t \Omega^*(s) ds} \mathbf{1} \mathbf{1}^T e^{\int_0^t \Omega(s) ds} & 0 \end{bmatrix} \begin{bmatrix} E(t) \\ D(t) \end{bmatrix} \quad (6.104)$$

の解  $E(t), D(t)$  を使って、

$$P_0(t) = E(t)D^{-1}(t)$$

と表せる [87]。 (6.104) 式の解は

$$\begin{aligned} E(t) &= C_1, \\ D(t) &= C_2 + C_1 \int_0^t \frac{1}{\sigma_v^2(\tau)} e^{\int_0^\tau \Omega^*(s) ds} \mathbf{1}\mathbf{1}^T e^{\int_0^\tau \Omega(s) ds} \end{aligned}$$

となるから ( $C_1, C_2$  は任意の定数行列)、

$$P_0(t) = \left[ \int_0^t \frac{1}{\sigma_v^2(\tau)} e^{\int_0^\tau \Omega^*(s) ds} \mathbf{1}\mathbf{1}^T e^{\int_0^\tau \Omega(s) ds} \right]^{-1} \quad (6.105)$$

を得る。ただし、 $C_1, C_2$  を決定する際に  $P_0(0) = P(0) = \infty$  を考慮した。ゆえに、(6.97) 式の関係から (6.39) 式が得られる。

一方、

$$\hat{\mathbf{x}}_0(t) = e^{-\int_0^t \Omega(s) ds} \hat{\mathbf{x}}(t) \quad (6.106)$$

とおけば、その時間微分に関して、(6.34) 式および (6.97) 式より

$$\begin{aligned} \frac{d\hat{\mathbf{x}}_0(t)}{dt} &= \frac{d}{dt} \left\{ e^{-\int_0^t \Omega(s) ds} \right\} \hat{\mathbf{x}}(t) + e^{-\int_0^t \Omega(s) ds} \frac{d\hat{\mathbf{x}}(t)}{dt} \\ &= -\Omega(t) \hat{\mathbf{x}}_0(t) + e^{-\int_0^t \Omega(s) ds} \left\{ \left[ \Omega(t) - \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1}\mathbf{1}^T \right] \hat{\mathbf{x}}(t) + \frac{1}{\sigma_v^2(t)} P(t) \mathbf{1}y(t) \right\} \\ &= -\frac{1}{\sigma_v^2(t)} P_0(t) e^{\int_0^t \Omega^*(s) ds} \mathbf{1}\mathbf{1}^T e^{\int_0^t \Omega(s) ds} \hat{\mathbf{x}}_0(t) + \frac{1}{\sigma_v^2(t)} P_0(t) e^{\int_0^t \Omega^*(s) ds} \mathbf{1}y(t) \end{aligned}$$

の関係が成り立つ。この解は、右辺第 1 項に対する遷移行列

$$\frac{\partial \Phi_{x_0}(t, \tau)}{\partial t} = -\frac{1}{\sigma_v^2(t)} P_0(t) e^{\int_0^t \Omega^*(s) ds} \mathbf{1}\mathbf{1}^T e^{\int_0^t \Omega(s) ds} \Phi_{x_0}(t, \tau), \quad (6.107)$$

$$\Phi_{x_0}(t, t) = I \quad (6.108)$$

を用いて、

$$\hat{\mathbf{x}}_0(t) = \int_0^t \Phi_{x_0}(t, \tau) \frac{1}{\sigma_v^2(\tau)} P_0(\tau) e^{\int_0^\tau \Omega^*(s) ds} \mathbf{1}y(\tau) d\tau$$

と表せる。ただし、 $\hat{\mathbf{x}}_0(0) = \hat{\mathbf{x}}(0) = \mathbf{0}$  を考慮した。(6.107) 式と (6.98) 式を比較すれば、

$$\Phi_{x_0}(t, \tau) = P_0(t) P_0^{-1}(\tau)$$

と表せるから、

$$\hat{\mathbf{x}}_0(t) = P_0(t) \int_0^t \frac{1}{\sigma_v^2(\tau)} e^{\int_0^\tau \Omega^*(s) ds} \mathbf{1}y(\tau) d\tau$$

となる。したがって、(6.97), (6.106) 式の関係を考慮すれば (6.38) 式が得られる。

## 第 7 章

# デジタル櫛形フィルタの設計と評価

### 7.1 はじめに

これまでの議論では、擬似周期信号推定にとって最適なフィルタが櫛形フィルタになることを連続信号領域において導いてきた。本章では、これを計算機もしくは専用 DSP 上で実現するにあたってのデジタルフィルタの設計法について議論する。

デジタルフィルタには FIR 型と IIR 型がある中で、特に FIR 型で櫛形フィルタを構成する場合には次のような難しさがある。すなわち、(1) サンプリング間隔より細かい精度で、櫛歯の間隔をピッチ周期に合わせることができない、(2) ピッチ周期の変化に櫛歯間隔を同期させるのに、すべてのフィルタ係数をその都度修正する必要がある、(3) 倍音などの線スペクトルを抽出するために急峻なフィルタ特性を実現するには、フィルタ次数を多くしなければならない、などのフィルタ構造上の制約である。実際、Lim ら [90] はピッチ周期に同期して櫛歯の間隔を随時調整するタイプの櫛形フィルタを提案し、Veeneman ら [91] は櫛歯を一定間隔に固定する代わりに、二乗誤差が最小になるように適応的に櫛歯の重みづけを変える方法を提案しているが、いずれも上記の問題に対する本質的な解決策にはなっていない。ピッチ周期に同期して標本点を取り直す方法 [94, 95] を導入すれば、(1),(2) の問題を回避することはできるが、可変サンプリングおよび (3) により回路規模が大きくなり実用的ではない。さらに、 $z$  領域の単位円上で伝達関数の零点を線スペクトルの周波数に一致させてノッチ特性を得る方法 [96] や、零点を等角度間隔に並べて櫛形のノッチ特性を実現する方法 [97] などもあるが、(3) の制約から十分な分離性能が得られないという問題がある。

これに対し IIR フィルタは、極の位置を調整するだけで任意の周波数に同調させることができ、同時に急峻なフィルタ特性が得られるため、櫛形フィルタを構成する上で有利であると考えられる。Nehorai は、この性質に早くから着目し、任意の複数の周波数成分を抑圧するためのノッチフィルタ [98] や、同様な原理に基づいた周期信号に対する適応櫛

形フィルタ [99] を提案している。特に、単一あるいは複数の独立な正弦波信号 (sinusoidal signal) の抽出を目的とした ALE (Adaptive Line Enhancer) [100] や、逆に抑圧を目的とした ANF (Adaptive Notch Filter) [101] などの研究では、目的信号の周波数にフィルタが同調するように極の位置を制御する適応アルゴリズムが数多く提案されている。例えば、LMS (Least Mean Square) 法 [102]、Gauss-Newton 法 [103]、確率勾配法 [104] によるもの、周波数の時変性を考慮して追跡性能を高めた方法 [105]、極や零点の位置をある程度固定しておいた上で二乗誤差が最小になるように微調整する方法 [106]、信号の瞬時振幅や瞬時周波数を利用した MFT (Multiple Frequency Tracker) 法 [107, 108]、参照入力を用いる方法 [109] や最近では単一正弦波に対する高速適応アルゴリズム [110] などがある。

基本的には、 $z$  領域で単位円内側に沿って極が置かれ、抽出あるいは抑圧すべき個々の周波数成分に合わせて、極の配置が適応アルゴリズムを通して決定される。このときの伝達特性は、極と正の実軸とのなす角度によってフィルタの中心周波数が、極と単位円との距離によって帯域幅が決まる。最適推定の立場から見れば、フィルタの中心周波数が目的信号の周波数に一致するように実軸とのなす角度が決まることはほぼ自明であるが、帯域幅すなわち単位円との距離が信号のどのような性質に基づいて具体的に決まるかは、これまでの研究ではわかっていなかった。

このような背景の中で本章では、事前にフィルタモデルを仮定することなく、6章での Kalman フィルタ方程式を直接離散化し差分方程式に書き直した結果として、同様な IIR フィルタの形式が導けることを示す。

## 7.2 フィルタ方程式の離散化による IIR フィルタの導出

6.5節で導いた Kalman フィルタ方程式を基にして、対応するデジタルフィルタの形式を以下に求める。

微分方程式として表された (6.53) 式

$$\frac{d\hat{x}_n(t)}{dt} \approx \left[ -\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t) + jn\bar{\omega}_0(t) \right] \hat{x}_n(t) + \frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)y(t) \quad (7.1)$$

に着目し、これを離散時間刻みの差分方程式として書き直すことを考える。1 サンプルング前の推定値  $\hat{x}_n(t - T_s)$  を初期値としたときの時刻  $t$  における微分方程式の解は、右辺第1項に対する推移解

$$\phi(t, \tau) = e^{\int_{\tau}^t \left\{ -\frac{\bar{x}_n(s)}{\sigma_v(s)}\sigma_n(s) + jn\bar{\omega}_0(s) \right\} ds} \quad (7.2)$$

を用いて、

$$\hat{x}_n(t) \approx \phi(t, t - T_s)\hat{x}_n(t - T_s) + \int_{t-T_s}^t \phi(t, \tau) \frac{\bar{x}_n(\tau)}{\sigma_v(\tau)}\sigma_n(\tau)y(\tau)d\tau \quad (7.3)$$

と表せる。サンプリング周期  $T_s$  に比べて倍音周期および振幅 / ピッチ変動が十分小さく、 $\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)$ ,  $n\bar{\omega}_0(t) \ll \frac{2\pi}{T_s}$  が成り立つと仮定すれば、右辺第1項の指数部被積分項および第2項の被積分項は、積分区間  $t - T_s \leq \tau \leq t$  でそれぞれ一定値  $\{-\frac{\bar{x}_n(s)}{\sigma_v(s)}\sigma_n(s) + jn\bar{\omega}_0(s)\}|_{s=t}$ 、 $o(t, \tau)\frac{\bar{x}_n(\tau)}{\sigma_v(\tau)}\sigma_n(\tau)y(\tau)|_{\tau=t}$  をとるものとみなせるので、上式は

$$\begin{aligned}\hat{x}_n(t) &\approx e^{\{-\frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t) + jn\bar{\omega}_0(t)\}T_s}\hat{x}_n(t - T_s) + \frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)y(t)T_s \\ &\approx \left(1 - \frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)T_s\right)e^{jn\bar{\omega}_0(t)T_s}\hat{x}_n(t - T_s) + \frac{\bar{x}_n(t)}{\sigma_v(t)}\sigma_n(t)T_s y(t)\end{aligned}\quad (7.4)$$

と近似できる。下段への近似は、 $\frac{\bar{x}_n}{\sigma_v}\sigma_n T_s \ll 1$  となることを用いた。さらに、 $t = kT_s$  において離散時間刻みで表せば、

$$\hat{x}_n(kT_s) \approx \gamma_n(kT_s)e^{jn\bar{\omega}_0(kT_s)T_s}\hat{x}_n((k-1)T_s) + (1 - \gamma_n(kT_s))y(kT_s)\quad (7.5)$$

の差分方程式を得る。ただし、

$$\gamma_n(kT_s) = 1 - \frac{\bar{x}_n(kT_s)}{\sigma_v(kT_s)}\sigma_n(kT_s)T_s\quad (7.6)$$

とおいた。同じ結論は、離散時間系列で記述した信号モデルに対する Kalman フィルタ解からも導かれる(付録 7.7.1 参照)。

したがって  $z$  領域における伝達関数は、1 サンプル  $T_s$  の遅延が演算子  $z^{-1}$  を乗じることに対応するため、

$$H_n(z) = \frac{1 - \gamma_n(kT_s)}{1 - \gamma_n(kT_s)z_0^n(kT_s)z^{-1}}, \quad \text{where } z = e^{j\omega T_s}, z_0(kT_s) = e^{j\bar{\omega}_0(kT_s)T_s}\quad (7.7)$$

と表せる<sup>†</sup>。上式は、 $z = z_0^n$  に鋭いピークをもつ狭帯域フィルタとなる。

特に、すべての倍音で二乗平均振幅や振幅変動が等しく、振幅変動のみを考慮した場合 ( $\gamma(kT_s) = 1 - \frac{\bar{x}(kT_s)}{\sigma_v(kT_s)}\sqrt{\sigma_a^2(kT_s) + \sigma_b^2(kT_s)T_s}$ ) は、

$$\begin{aligned}H_n(z) &= \frac{1 - \gamma(kT_s)}{1 - \gamma(kT_s)z_0^n(kT_s)z^{-1}} \\ &= H_0(z_0^{-n}(kT_s)z)\end{aligned}\quad (7.8)$$

となり、フィルタ全体の伝達関数  $H(z) = \sum_{n=1}^N H_n(z)$  としては性質 5 の定 BW 櫛形特性をもつことになる。逆にピッチ変動のみを考慮した場合は、中心周波数  $n\bar{\omega}_0$  の近傍  $\omega =$

<sup>†</sup>s-z 変換法としてよく知られたインパルス不変法 [112] では、インパルス応答に対して直接離散化を行い、それを  $z$  変換することによって伝達関数を求めている。フィルタ方程式の離散化は結果的にはインパルス応答の離散化と等価な演算を意味するから、ここでの手続きもインパルス不変法の考え方に基づいているといえる。実際、 $s$  領域での伝達関数 ((3.71) 式) をインパルス不変法 [112] により  $z$  領域へ変換した結果は、(7.7) 式に一致することが容易に確かめられる。

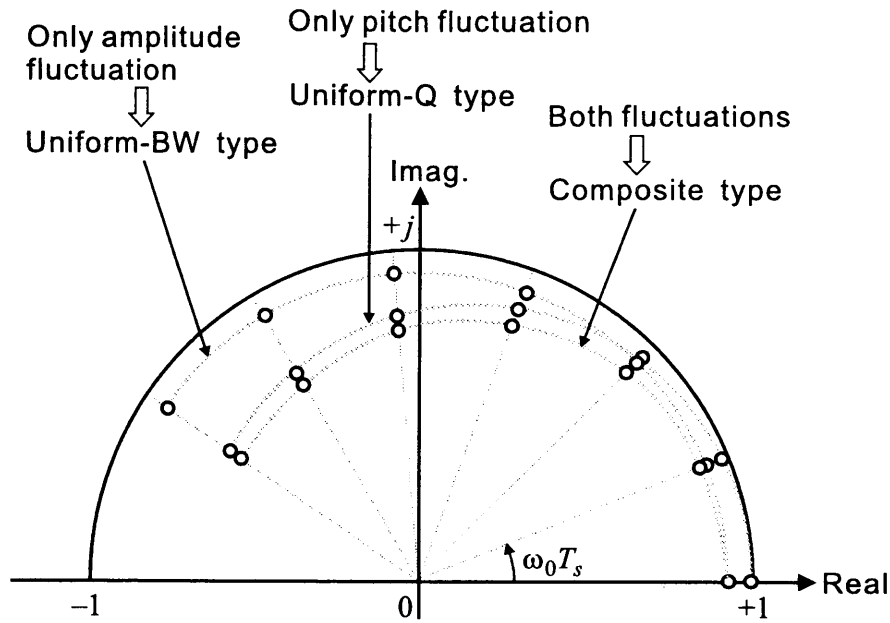


図 7.1 IIR デジタル櫛形フィルタの極配置

$n\bar{\omega}_0 + \Delta\omega$  ( $\Delta\omega \ll \bar{\omega}_0$ ) において  $e^{-\frac{n\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s} \approx 1 - \frac{n\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s$ ,  
 $e^{-\{ \frac{n\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s + j\Delta\omega T_s \}} \approx 1 - \frac{n\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s - j\Delta\omega T_s$  が成り立つため、

$$\begin{aligned}
 H_n(z) &\approx \frac{n \frac{\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s}{n \frac{\bar{x}(kT_s)}{\sigma_v(kT_s)}\sigma_\omega(kT_s)T_s + j\Delta\omega T_s} \\
 &\approx H_1(z^{\frac{1}{n}}) \tag{7.9}
 \end{aligned}$$

と表せ、性質 6 の定 Q 櫛形特性をもつことになる。これは単位円からの極の位置と帯域幅の関係 [113] から説明できる。すなわち図 7.1 に示すように、(7.8) 式の定 BW 櫛形フィルタの場合は、単位円内側に等角度間隔かつ等半径上に極が置かれるのに対して、(7.9) 式の定 Q 櫛形フィルタの場合では、倍音次数に比例して単位円から離れた位置に極が配置される。さらに、振幅変動とピッチ変動を同時に考慮した定 BW/ 定 Q 複合櫛形フィルタの場合は、倍音次数が低次から高次になるにしたがって、等半径上に極が並ぶ状態から徐々に半径が短くなる方向に極が置かれることになる。

特に実信号に対しては、極を負の周波数側 ( $z$  領域の下半平面) にも対称的に配置した

$$\begin{aligned}
 H_n(z) &= \frac{1 - \gamma_n(kT_s)}{1 - \gamma_n(kT_s)z_0^n(kT_s)z^{-1}} + \frac{1 - \gamma_n(kT_s)}{1 - \gamma_n(kT_s)z_0^{-n}(kT_s)z^{-1}} \\
 &= 2(1 - \gamma_n(kT_s)) \frac{1 - \gamma_n(kT_s) \cos(n\bar{\omega}_0(kT_s)T_s)z^{-1}}{1 - 2\gamma_n(kT_s) \cos(n\bar{\omega}_0(kT_s)T_s)z^{-1} + \gamma_n^2(kT_s)z^{-2}} \tag{7.10}
 \end{aligned}$$

と表せる実係数のフィルタを用いればよい。最終的に、全倍音成分を含む擬似周期信号に対



するフィルタの伝達関数としては

$$H(z) = \sum_{n=1}^N 2(1 - \gamma_n(kT_s)) \frac{1 - \gamma_n(kT_s) \cos(n\bar{\omega}_0(kT_s)T_s)z^{-1}}{1 - 2\gamma_n(kT_s) \cos(n\bar{\omega}_0(kT_s)T_s)z^{-1} + \gamma_n^2(kT_s)z^{-2}} \quad (7.11)$$

となる。

### 7.3 ノッチフィルタとの関係

(7.7) 式の狭帯域フィルタ  $H_n(z)$  の利得を反転させたもの

$$\begin{aligned} 1 - H_n(z) &= \gamma_n \frac{1 - z_n z^{-1}}{1 - \gamma_n z_n z^{-1}} \\ &\approx \frac{1 - z_n z^{-1}}{1 - \gamma_n z_n z^{-1}} \equiv \bar{H}_n(z), \quad \text{where } z_n = e^{j\omega_n T_s} \end{aligned} \quad (7.12)$$

は、ALE とは逆に、特定の線スペクトル成分を抑圧するノッチフィルタとなる。ただし、ノッチ周波数は任意の周波数  $\omega_n$  を選べるものとし、簡単のため定係数のフィルタとした。上式は、 $\omega_n T_s$  の角度方向に沿って単位円内側に極が、単位円上に零点が置かれることを意味する。実信号に対するノッチフィルタの形式としては

$$\begin{aligned} \bar{H}_n(z) &= \frac{1 - z_n z^{-1}}{1 - \gamma_n z_n z^{-1}} \cdot \frac{1 - z_n^* z^{-1}}{1 - \gamma_n z_n^* z^{-1}} \\ &= \frac{1 - 2 \cos(\omega_n T_s) z^{-1} + z^{-2}}{1 - 2\gamma_n \cos(\omega_n T_s) z^{-1} + \gamma_n^2 z^{-2}} \end{aligned} \quad (7.13)$$

と表せ、さらに複数の正弦波に対するノッチフィルタは

$$\bar{H}(z) = \prod_{n=1}^N \frac{1 - 2 \cos(\omega_n T_s) z^{-1} + z^{-2}}{1 - 2\gamma_n \cos(\omega_n T_s) z^{-1} + \gamma_n^2 z^{-2}} \quad (7.14)$$

で与えられる<sup>†</sup>。

特に、 $\omega_n = n\omega_0$ ,  $T_s = \frac{2\pi}{N\omega_0}$ ,  $\gamma_n = \gamma$  として、円周上均等に極と零点を配置すれば、極は  $1 - \gamma^N z^{-N} = 0$  の根、零点は  $1 - z^{-N} = 0$  の根にそれぞれ対応するから、

$$\begin{aligned} \bar{H}(z) &= \prod_{n=0}^{N-1} \frac{1 - e^{j2\pi \frac{n}{N}} z^{-1}}{1 - \gamma e^{j2\pi \frac{n}{N}} z^{-1}} \\ &= \frac{1 - z^{-N}}{1 - \gamma^N z^{-N}} \end{aligned} \quad (7.15)$$

と表せることになる [114]。このフィルタにより  $N$  点の周期をもつ信号成分が抑圧できるのは、上式が  $\frac{2\pi}{N}$  の整数倍の規格化周波数にノッチをもつ楕形特性になると同時に、漸化式では

$$\hat{x}(kT_s) = \gamma^N \hat{x}((k-N)T_s) + y(kT_s) - y((k-N)T_s) \quad (7.16)$$

<sup>†</sup>例えば、(7.12) 式は文献 [111]、(7.13)、(7.14) 式は文献 [105] などで用いられている。

となるように、 $N$ 点前のサンプル点を差し引くフィルタ演算になるためである。このようなシンプルなフィルタ構成がとれるのは、目的信号のピッチ周期がサンプリング周期の整数倍にちょうど一致するときだけに限られることに注意する。この制約を取り払う対策としては、内挿フィルタを用いてサンプリング点を分数倍の細かさに取り直す方法 [115] や、ピッチ周期がサンプリング周期の整数倍になるように予め可変サンプリングを行っておく方法 [94, 95] などがあるが、いずれも複雑なりサンプリング処理が必要になるため実用的ではない。それに対し (7.14) 式の構成を用いれば、フィルタ係数を調整するだけで任意のピッチ周期に連続的に適応させることが可能になる。

## 7.4 フィルタ性能の評価

(7.11) 式による定 BW および定 Q 櫛形フィルタを用いた擬似周期信号推定の評価実験を行った。

### 7.4.1 評価基準

フィルタ性能の定量的な評価には、以下の S/N 比と相関係数を用いる。

目的信号  $x(kT_s)$  との残差が雑音成分になるから、観測信号  $y(kT_s)$  の S/N 比を

$$\text{SNR}_{in} = 10 \log_{10} \frac{\sum_{k=0}^L |x(kT_s)|^2}{\sum_{k=0}^L |x(kT_s) - y(kT_s)|^2} \quad (7.17)$$

とし、これに対する推定信号  $\hat{x}(kT_s)$  の S/N 比を

$$\text{SNR}_{out} = 10 \log_{10} \frac{\sum_{k=0}^L |x(kT_s)|^2}{\sum_{k=0}^L |x(kT_s) - \hat{x}(kT_s)|^2} \quad (7.18)$$

と定義する。 $L$  は分析フレーム区間の信号データ点数である。フィルタ出力後の  $\text{SNR}_{out}$  が入力前の  $\text{SNR}_{in}$  に比べてどの程度向上するかによってフィルタ性能を評価する。注意すべきは、聴感上の評価として重要なのは信号波形の歪み (つまり音色の変化) の有無であるのに対して、SNR の基準では振幅の様な増減 (つまり音量の変化) だけでも悪く評価されてしまうという点である。

そこで、この欠点を補う意味で、波形全体の相似な変化には影響しない相関係数による基準もあわせて導入する。これは、観測信号に対して

$$\text{CR}_{in} = \frac{\sum_{k=0}^L x^*(kT_s)y(kT_s)}{\sum_{k=0}^L |x(kT_s)|^2 \sum_{k=0}^L |y(kT_s)|^2} \quad (7.19)$$

推定信号に対して

$$\text{CR}_{out} = \frac{\sum_{k=0}^L x^*(kT_s)\hat{x}(kT_s)}{\sum_{k=0}^L |x(kT_s)|^2 \sum_{k=0}^L |\hat{x}(kT_s)|^2} \quad (7.20)$$

と定義される。相関係数は、常に0から1の値をとり、波形の相似性が高いほど1に近い値をとる。

#### 7.4.2 ピッチ誤差に対する頑健性

振幅およびピッチが一定の理想周期信号を目的信号、それに白色 Gauss 雑音を付加したものを観測信号として、定 BW および定 Q 楕形フィルタにより目的信号を抽出する数値実験を行った。ただし、観測信号から正確なピッチ周期を知ることはできないという状況を想定し、楕形フィルタのピッチパラメータには故意に誤差を与えたものを用いた。周期信号は15次倍音で構成され、各倍音は同じ大きさの振幅をもち、最高次数の倍音周波数は Nyquist 周波数の6割になるように選んだ。適当に設定したピッチ周期の標準偏差  $\sigma_T$  に対して、ピッチパラメータとして実際に与えた誤差  $\Delta T$  の比を横軸にとって、S/N 比および相関係数を計算した結果が図 7.2 である。帯域幅または Q 値を決めるフィルタ係数は、定 BW 楕形フィルタで  $\gamma_n = 1 - \omega_0^2 \sigma_T T_s / 4\pi^2$ 、定 Q 楕形フィルタで  $\gamma_n = 1 - n\omega_0^2 \sigma_T T_s / 4\pi^2$  とし固定とした。 $\Delta T / \sigma_T = 1$  のときは、与えられたピッチ誤差にちょうど見合ったフィルタ係数が定 Q 楕形フィルタにおいて設定されていることを意味する。

(a),(b) 両者の結果とも、 $\Delta T / \sigma_T > 0.6$  で定 Q 楕形フィルタの方が勝っており、ピッチ誤差に頑健であることを示している。特に、定 BW 型ではピッチ誤差が大きくなるとフィルタ性能が極端に劣化する。逆にピッチ誤差が小さいときに定 Q 型の方が劣るのは、ピッチ誤差を緩和する働きに比べ、雑音を多く拾う効果の方が高次倍音で大きくなることによる。

#### 7.4.3 ピッチ変動に対する頑健性

前項と同じ条件の下で、誤差の代わりにランダムな変動をピッチ周期に与えたものを目的信号としてフィルタ計算を行った。楕形フィルタのピッチパラメータには、変動を加える前のピッチ周期の真値を代入した。フィルタ係数を定める際の基準  $\sigma_T$  に対する、ピッチ変動の標準偏差  $\Delta T$  の比を横軸にとって、定 BW および定 Q 楕形フィルタによる S/N 比および相関係数を計算した結果が図 7.3 である。この場合も、 $\Delta T / \sigma_T > 0.4$  において定 Q 楕形フィルタの方が勝っており、ピッチ変動に対する優位性を示している。

#### 7.4.4 振幅 / ピッチ変動による波形 / スペクトル歪み

母音音声「ア」の1周期分を切り出し周期化した信号に対して、信号モデル 4 にしたがって振幅変動やピッチ変動を与えたものを目的信号、それに白色 Gauss 雑音を付加したものを観測信号として、楕形フィルタによって観測信号から目的信号を抽出するシミュ

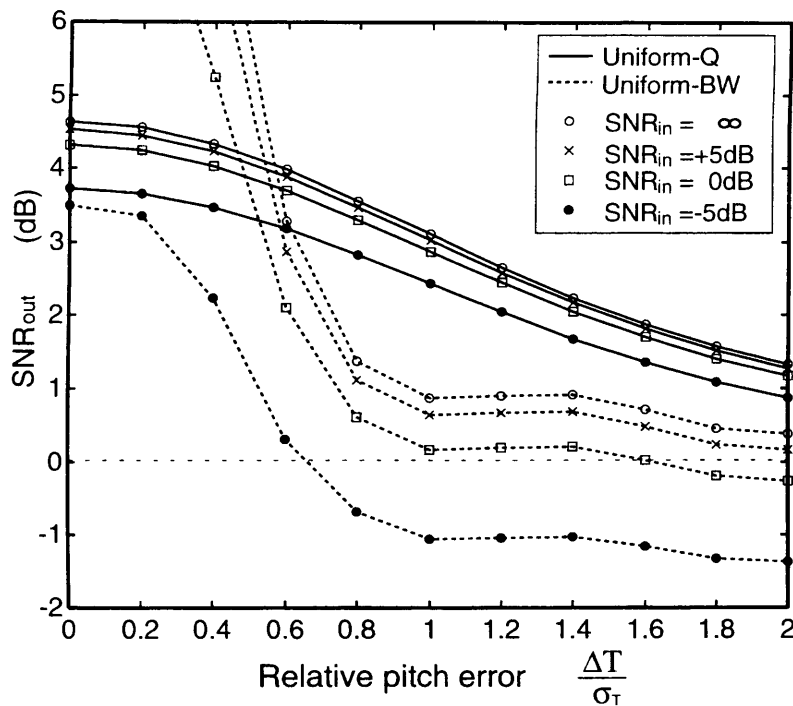
レーション実験を行った。以下共通に、変動を与える前の合成周期信号のピッチ周期：9.2ms、最高倍音次数：10次、観測信号のSN比：-9dB、サンプリング周波数：44.1kHzとした。

### 振幅変動に対する影響

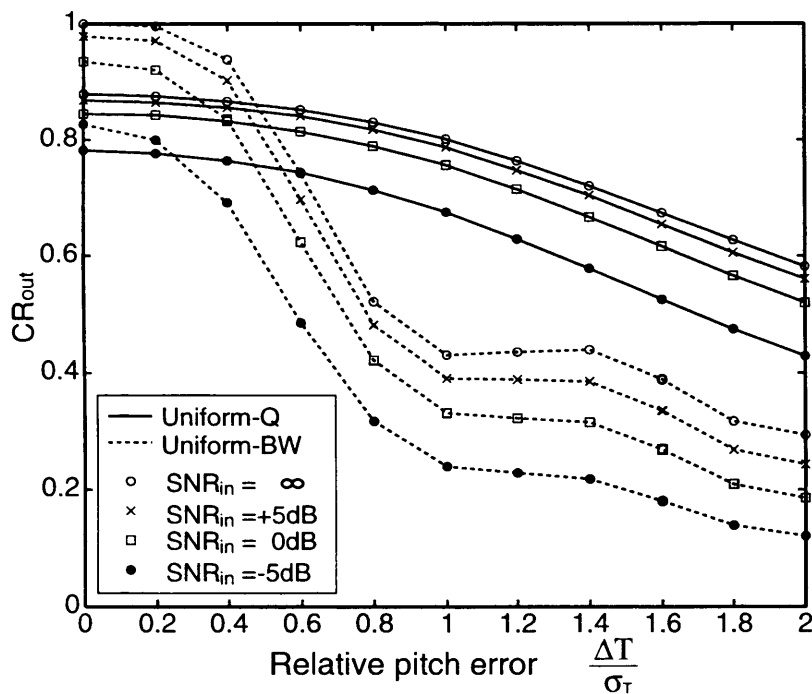
振幅変動（各倍音で均一に分散  $\sigma_a^2 = 10\text{s}^{-1}$ ）だけを与えた場合の抽出結果が図7.4(c),(d)である。信号波形の一部を切り出して示した。振幅変動に合わせて定BW型では  $\sigma_a^2 = 10\text{s}^{-1}, \sigma_b^2 = \sigma_\omega = 0$  とし、比較対象の定Q櫛形フィルタでは基本周波数の帯域幅が定BW型と同じになるように  $\sigma_a = \sigma_b = 0, \sigma_\omega^2 = 10\text{s}^{-1}$  と設定し、各フィルタの係数  $\gamma_n$  を計算した。定BW型ではS/N比が13.8dBに改善され目的信号（同図(a)）に見かけ上ほぼ一致しているのに対して（同図(c)）、定Q型では9.0dBにしか改善されず波形上も高調波雑音が見られる（同図(d)）。これは、スペクトル分布（図7.5）からわかるように、定Q型では帯域が広がる高域側で雑音を多く拾ってしまうためである（同図(d)）。このような振幅変動だけの場合、すべての倍音成分で均一の帯域幅をもつ定BW型特性が有利であるといえる。定Q櫛形フィルタではたとえどのようにフィルタ係数  $\gamma_n$  を調整しても、最大で12.1dBまでしか改善されないことが確かめられた。

### ピッチ変動に対する影響

逆に、ピッチ変動（分散  $\sigma_\omega^2 = 10\text{s}^{-1}$ ）だけがある場合の抽出結果が図7.6(c),(d)である。各フィルタの係数  $\gamma_n$  は上記と同一の設定を用いた。この場合、定BW型では4.9dBにしか改善されず波形全体に歪みが現れているのに対して（同図(c)）、定Q型では7.2dBまで改善され原信号波形（同図(a)）を比較的良好に復元していることが確認できる（同図(d)）。原信号スペクトル（図7.7(a)）に見るように、ピッチ変動は倍音次数に比例して帯域幅を拡大させる効果をもつため、それを抽出するフィルタとしても定Q型特性の方が整合性がよく、原信号のスペクトル構造を保持するのに都合がよいといえる（同図(d)）。これに対し定BW型では、雑音成分を取り除く以上に原信号スペクトルの形状を変えてしまうことになる（同図(c)）。たとえ定BW櫛形フィルタの帯域幅を連続的に変化させても、最大で5.9dBまでしか改善されないことが確かめられた。

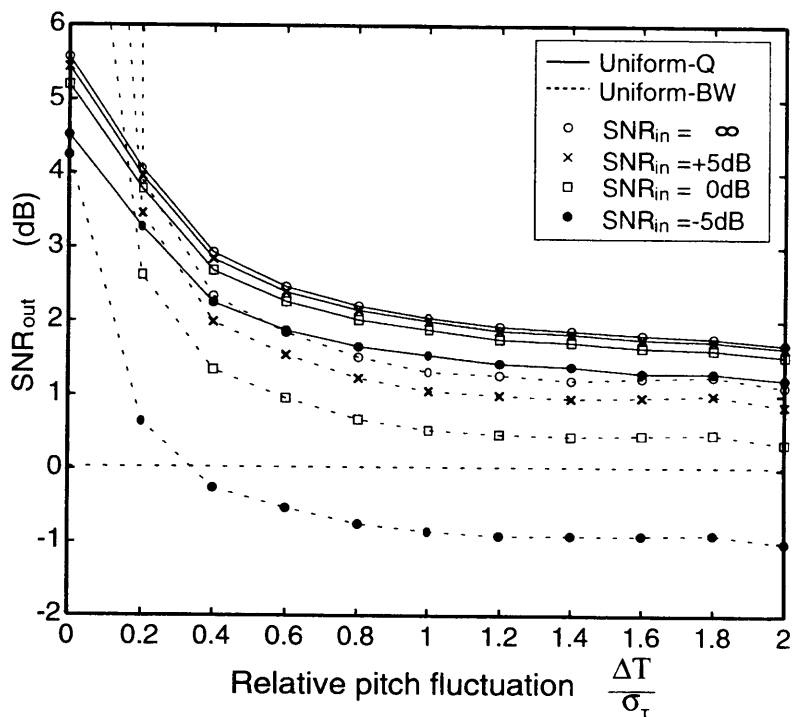


(a) S/N 比による評価

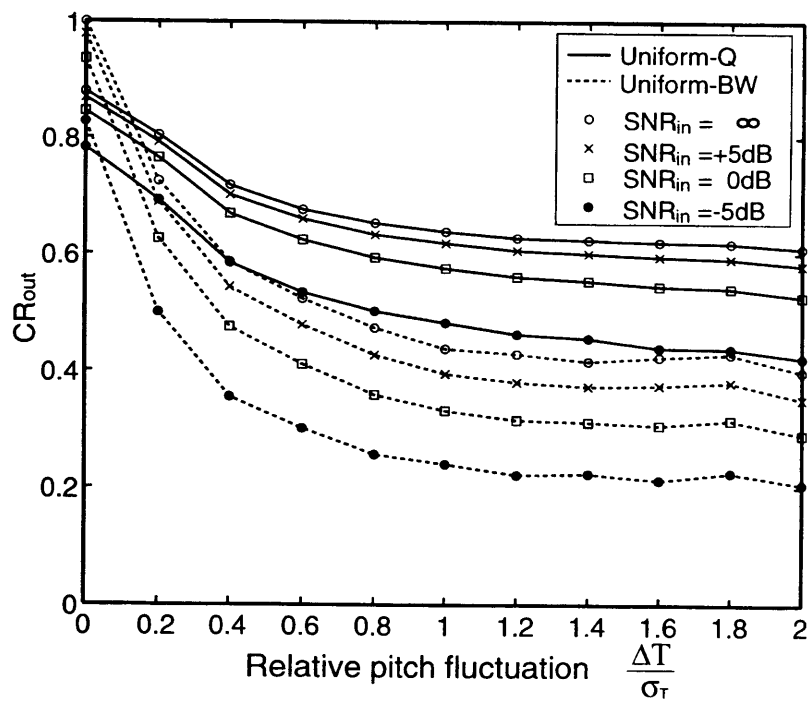


(b) 相関係数による評価

図 7.2 ピッチ誤差に対する頑健性



(a) S/N 比による評価



(b) 相関係数による評価

図 7.3 ピッチ変動に対する頑健性

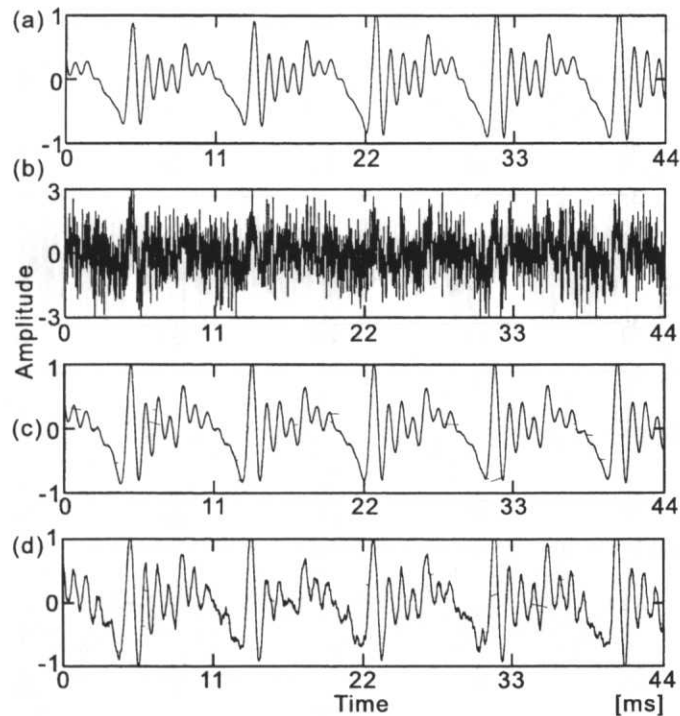


図 7.4 振幅変動が信号波形に及ぼす影響、(a) 振幅変動をともなった信号、(b) 雑音で汚された信号、(c) 定 BW 楕形フィルタの出力、(d) 定 Q 楕形フィルタの出力

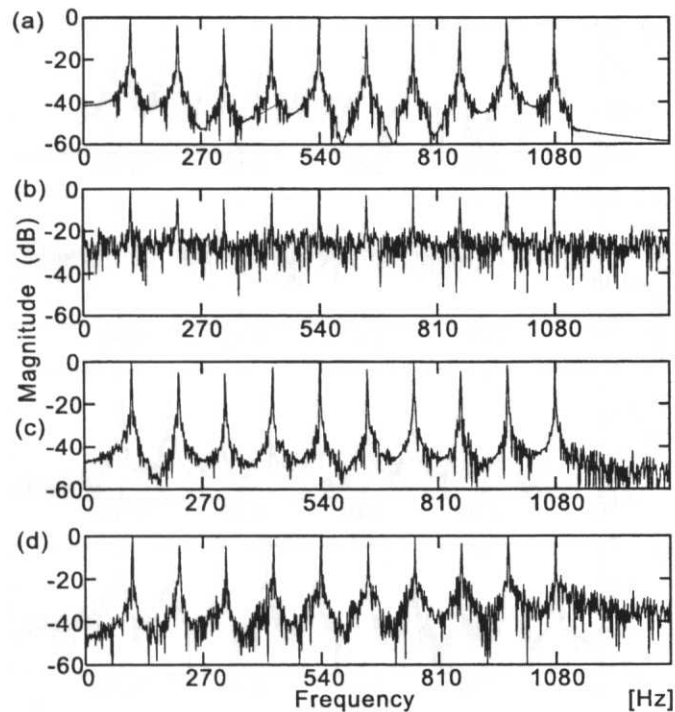


図 7.5 振幅変動がスペクトルに及ぼす影響、(a) 振幅変動をともなった信号、(b) 雑音で汚された信号、(c) 定 BW 楕形フィルタの出力、(d) 定 Q 楕形フィルタの出力

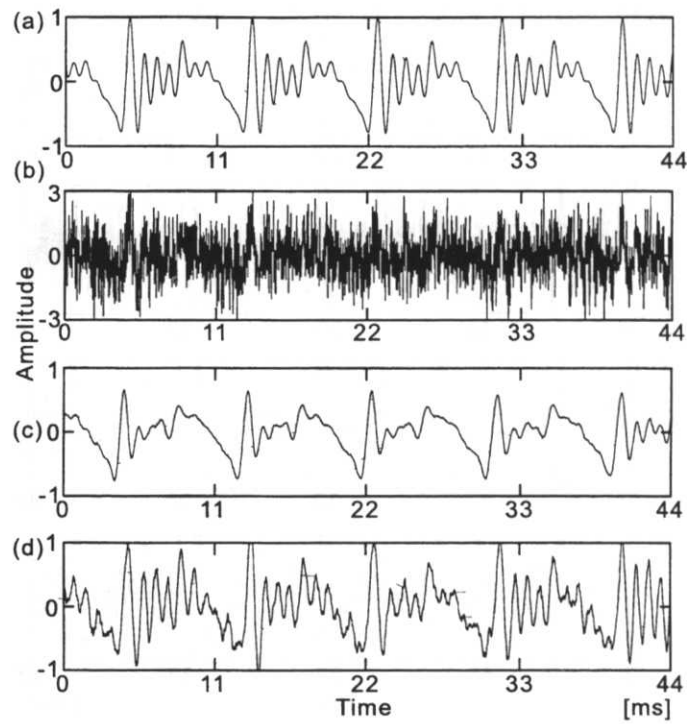


図 7.6 ピッチ変動が信号波形に及ぼす影響、(a) ピッチ変動をともなった信号、(b) 雑音で汚された信号、(c) 定 BW 楕形フィルタの出力、(d) 定 Q 楕形フィルタの出力

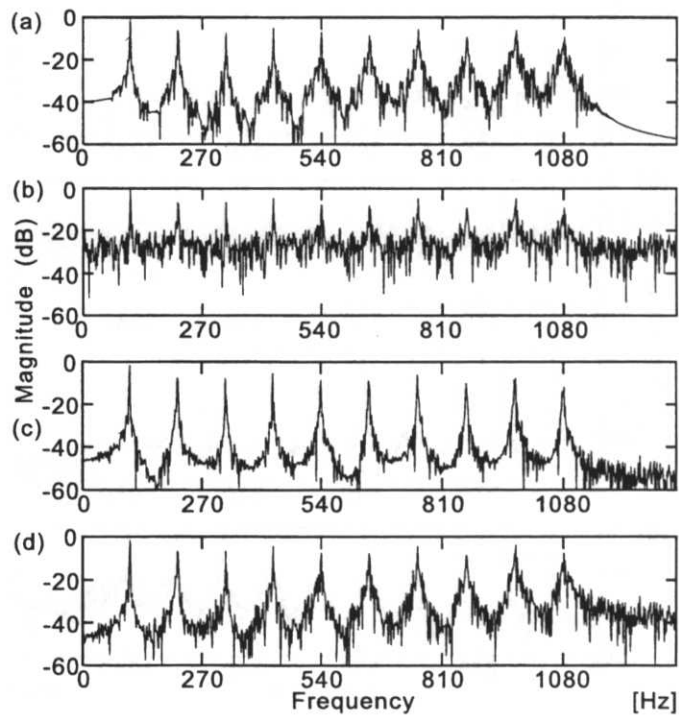


図 7.7 ピッチ変動がスペクトルに及ぼす影響、(a) ピッチ変動をともなった信号、(b) 雑音で汚された信号、(c) 定 BW 楕形フィルタの出力、(d) 定 Q 楕形フィルタの出力



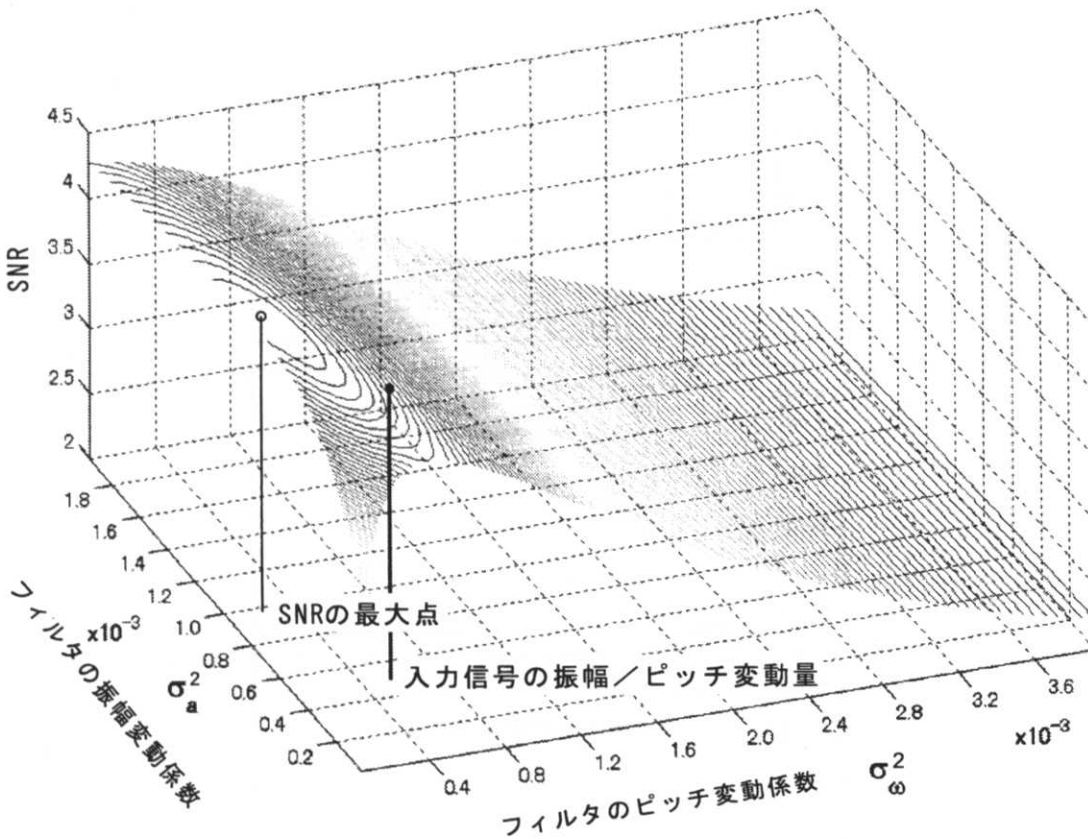


図 7.8 S/N 比を最大にするフィルタ係数と振幅/ピッチ変動量との関係

最適なフィルタ係数と振幅/ピッチ変動量の関係

入力信号の振幅/ピッチ変動量を  $\sigma_a^2 = \sigma_\omega^2 = 0.5 \times 44.1s^{-1}$  に固定し、定 BW/ 定 Q 複合楕円フィルタの振幅/ピッチ変動パラメータをいろいろ変えて S/N 比を計算した結果が図 7.8 である。振幅/ピッチ変動量に見合ったフィルタ係数を設定すれば、ほぼ最大の S/N 比で信号分離可能なことが確かめられた。

7.4.5 実音声・楽音への適用

擬似周期信号モデルに基づいて設計された本フィルタが実際の音声や楽音にも適用できることをみるために、混合音からの個別音の分離再生実験を行った。使用した音は、有声音「あえ」と僅かにビブラートのかかったトランペット音（音階：E<sub>2</sub>）との混合音である。図 7.9(a) はそのスペクトログラムであり、各調波スペクトル分布が重なり合ったものになっている。多峰性のあるピッチ存在確率の極大を個別に追跡する方法（5章）によって、ピッチ周波数を推定した結果が図 7.9(b),(c) である。自然音のため細かなピッチ変動がある

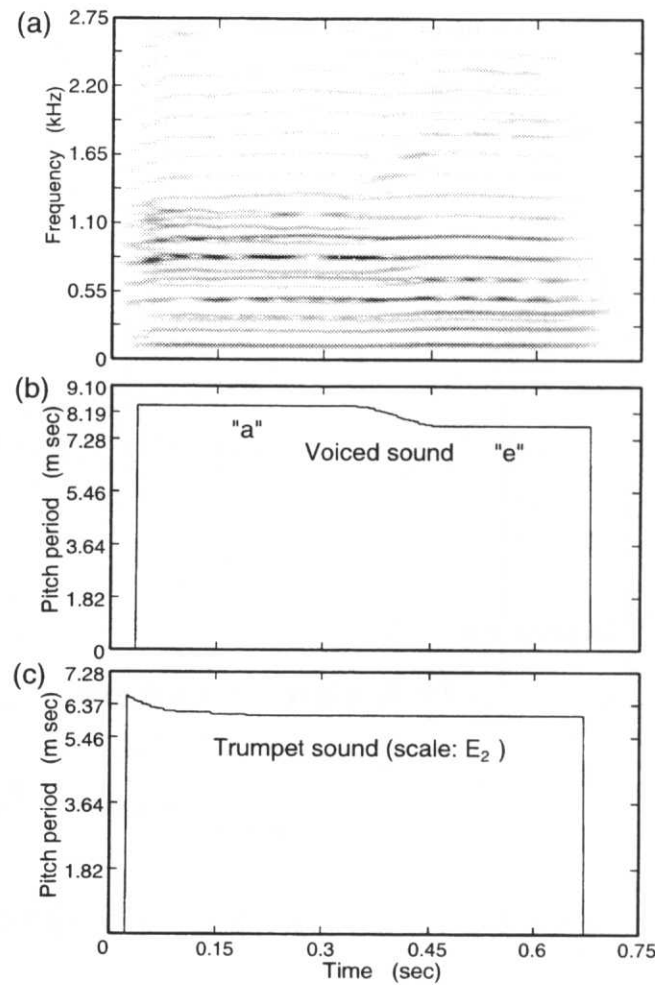


図 7.9 有声音「あえ」とビブラートのかかったトランペット音 (音階:  $E_2$ ) との混合音、(a) スペクトログラム、(b) 「あえ」に対するピッチ軌跡の推定結果、(c) トランペット音に対するピッチ軌跡の推定結果、

と同時に、そのピッチ推定値には、検出の際の周波数分解能が有限であることや、互いの音が外乱として作用し合うことなどにより、必然的に誤差が生じることに注意する。定 BW および定 Q 楕形フィルタにより抽出された有声音「あえ」およびトランペット音に対するスペクトログラムが図 7.10, 7.11 である。特に、0.15sec および 0.5sec 付近の断面分布をプロットしたのがそれぞれ図 7.12, 7.13 である。いずれも定 BW 楕形フィルタでは低い周波数側でスペクトル歪みがあり高い周波数側でスプリアスが生じているのに対して、定 Q 型ではそれらが比較的少ないことがわかる。

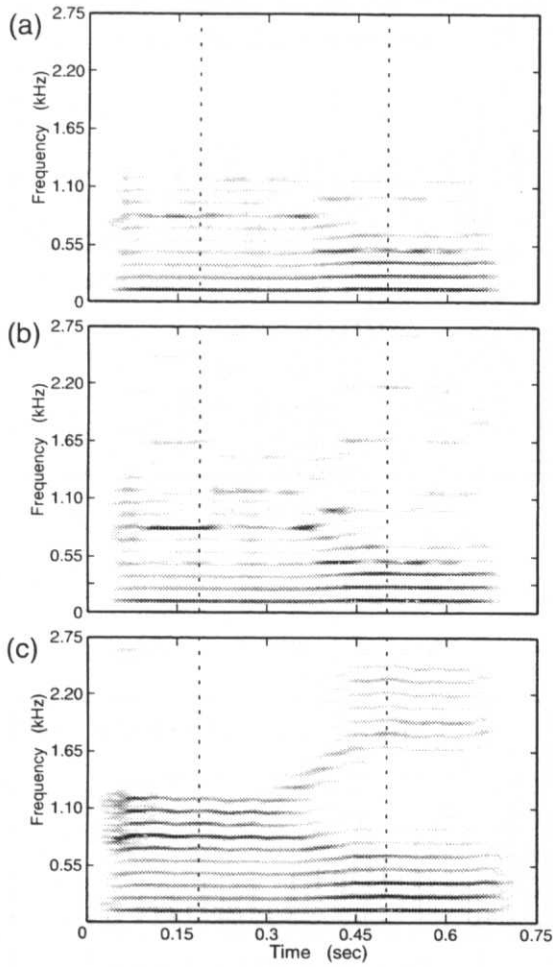


図 7.10 抽出された有声音「あえ」のスペクトログラム、(a) 定 Q 楕形フィルタの出力結果、(b) 定 BW 楕形フィルタの出力結果、(c) 原信号

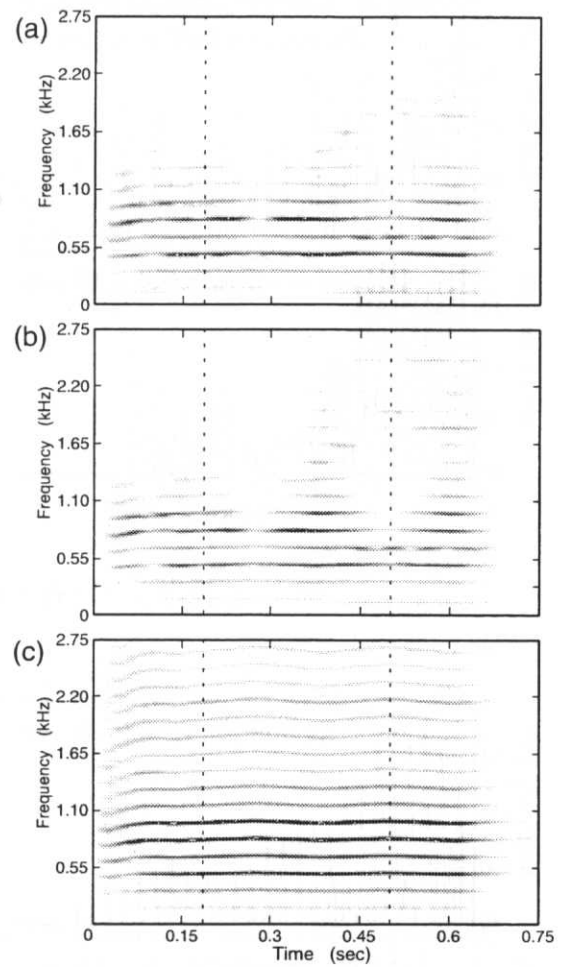


図 7.11 抽出されたトランペット音のスペクトログラム、(a) 定 Q 楕形フィルタの出力結果、(b) 定 BW 楕形フィルタの出力結果、(c) 原信号

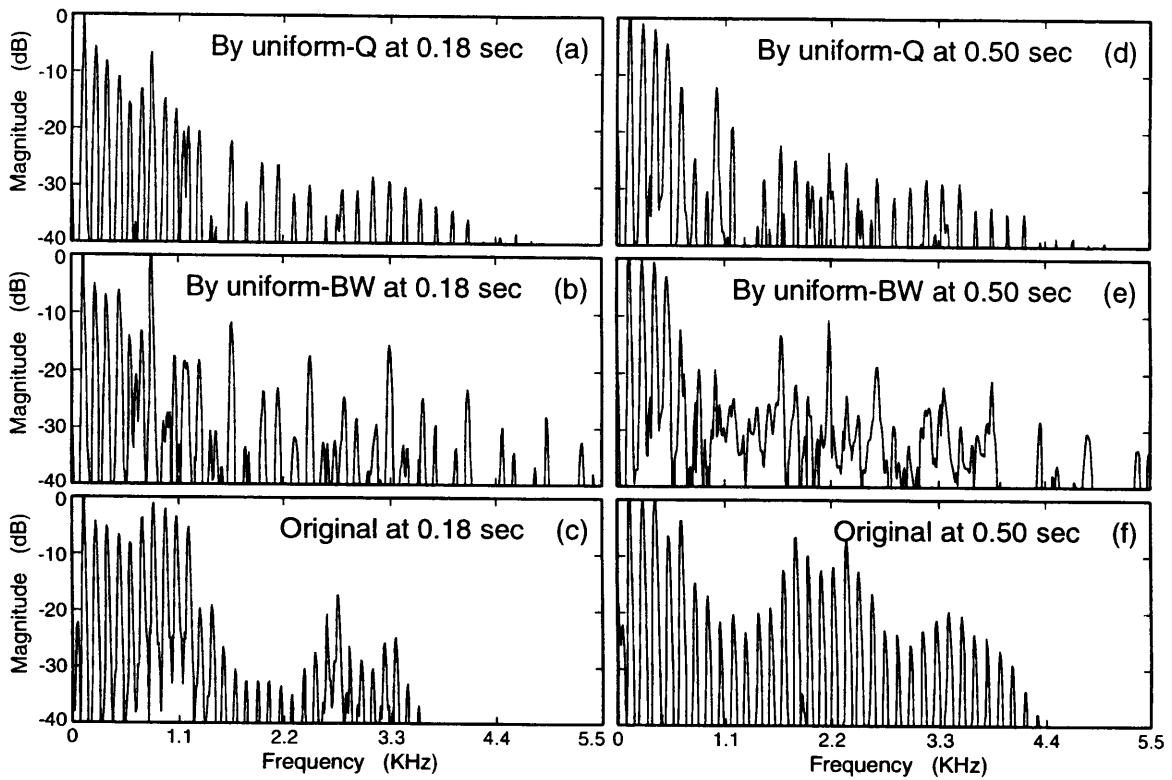


図 7.12 抽出された有声音「あえ」のスペクトル断面

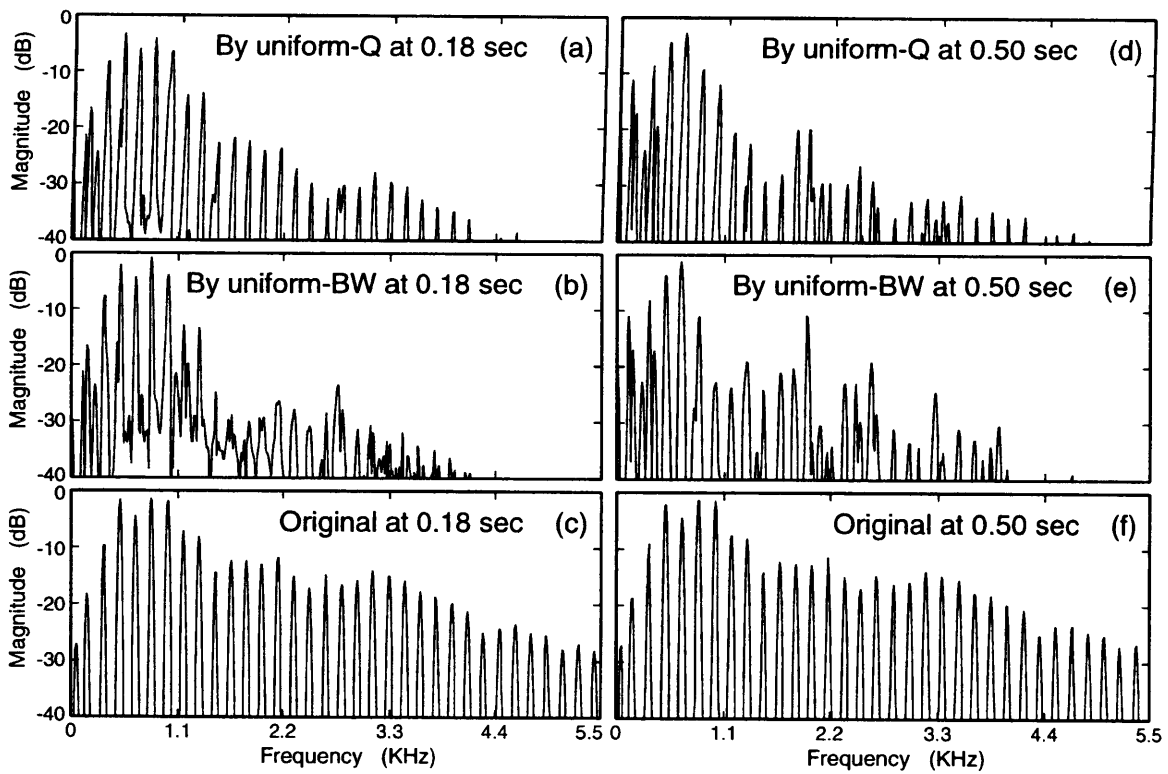


図 7.13 抽出されたトランペット音のスペクトル断面

## 7.5 音声のスペクトル包絡による評価

これまでのシミュレーションでは、SN比や相関係数の計算結果にみるように、主に信号波形どうしの直接比較を評価基準として用いてきた。我々の聴覚系では、音声認識を代表例としてその処理対象は信号波形自体よりもむしろそのスペクトル特性、さらに高次の会話認識では調波性の微細構造を取り去ったスペクトル包絡を音声の特徴量として主に利用していると考えられる。実際、音声認識の工学的な実現に際しては、そのフロントエンドとしてスペクトル包絡を抜き出す信号処理の手続きが不可欠になっている。そこで、今まで設計してきた擬似周期信号分離のためのフィルタが、音声のスペクトル包絡の比較においても騒音環境下で有利に働くことを以下で検証する。

ここではスペクトル包絡の計算手続きとして、実際の音声認識システムへの特徴量として用いられているLPCケプストラムとLPCメルケプストラムを導入し、それぞれについてスペクトル包絡の比較のための距離尺度を考える。以下にその概略を説明する。

### 7.5.1 LPCケプストラムによる距離尺度

LPC法では、スペクトル包絡を与えるフィルタモデルとして、伝達関数が

$$H_{lpc}(z) = \frac{b_0}{1 + \sum_{m=1}^M a_m z^{-m}} \quad (7.21)$$

と表せる全極型フィルタを仮定している。線形予測係数  $a_m$  は、対象とする音声データからLPC法によって事前に求めておくことができる [74]。LPCケプストラムとは、上式を指数関数型の伝達関数

$$H_{lpc}(z) = e^{\frac{C_0}{2}} + \sum_{m=1}^{\infty} C_m z^{-m} \quad (7.22)$$

に置き換えたことによる係数  $C_m$  を意味し、このときパラメータ  $a_m, C_m$  間には、

$$C_m = \begin{cases} 2 \log b_0, & m = 0 \\ -a_1, & m = 1 \\ -a_m - \frac{1}{m} \sum_{k=1}^{m-1} k C_k a_{m-k}, & m \geq 2 \end{cases} \quad (7.23)$$

のような漸化式の関係があることが容易に示され [74]、これにより線形予測係数  $a_m$  を使ってLPCケプストラム係数  $C_m$  が求められる。

聴覚における音の大きさに関する心理物理学的特性を考慮すれば、スペクトル包絡はフィルタの対数二乗振幅すなわち  $\log |H_{lpc}(z)|^2$  によって表現するのが適当である。そこで、与えられた2つの音声データに対するスペクトル包絡をそれぞれ  $\log |H_{lpc}(e^{j\Omega})|^2, \log |H'_{lpc}(e^{j\Omega})|^2$  とし、両者の距離を

$$d(\mathbf{a}, \mathbf{a}') = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \{ \log |H_{lpc}(e^{j\Omega})|^2 - \log |H'_{lpc}(e^{j\Omega})|^2 \} d\Omega \right]^{\frac{1}{2}} \quad (7.24)$$

と定義する。これに (7.22) 式を代入すれば、

$$d(\mathbf{a}, \mathbf{a}') = \left[ (C_0 - C'_0)^2 + 2 \sum_{m=1}^{\infty} (C_m - C'_m)^2 \right]^{\frac{1}{2}} \quad (7.25)$$

と表せ、LPC ケプストラム係数  $C_m, C'_m$  間のユークリッド距離に直接対応した形でスペクトル包絡間の距離が計算でき、しかも (7.23) 式より間接的には各線形予測係数  $a_m, a'_m$  を用いて求められる。

### 7.5.2 LPC メルケプストラムによる距離尺度

3.8節で述べたように、人の聴感覚は感覚的な音の高さとピッチ周波数との間は単純な線形関係にはなく、1KHz 以下では周波数にほぼ比例して音が高く感じるのに対し、1KHz 以上では周波数の対数に比例して音の高さを知覚するといった、非直線周波数目盛すなわちメルスケールの関係にある [45]。したがって、聴感覚に忠実な音声認識を実現しようとする場合には、上記の議論に加え周波数軸の伸縮を施した新たな距離尺度を定義しておく必要がある。そこで、単位円周上で周波数を刻む  $z$  変数を

$$\tilde{z}^{-1} = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}} \quad ((3.74) \text{ 式の再掲}) \quad (7.26)$$

とおくことによって、新たな周波数スケール  $\tilde{z}$  で (7.22) 式を置き換えた

$$H_\alpha(\tilde{z}) = e^{\frac{\tilde{c}_0}{2} + \sum_{m=1}^{\infty} \tilde{c}_m \tilde{z}^{-m}} \quad (7.27)$$

を考える [74]。上式と (7.21) 式の関数値が一致しなければならないという制約から両者を等置すれば、線形予測係数  $a_m$  と LPC メルケプストラム係数  $\tilde{C}_m$  の関係が求められる。具体的には、一旦 (7.21) 式のメルスケール上での表現

$$H_\alpha(\tilde{z}) = \frac{\tilde{b}_0}{1 + \sum_{m=1}^{\tilde{M}} \tilde{a}_m \tilde{z}^{-m}} \quad (7.28)$$

と元の (7.28) 式を等置し、 $a_m$  からメルスケール上の予測係数  $\tilde{a}_m$  を求めておいた上で (7.23) 式に対応する関係を導く。その結果として、

$$a_0 = 1 \quad (7.29)$$

$$i = -M, -(M-1), \dots, -1, 0;$$

$$\tilde{a}_m^{(i)} = \begin{cases} a_{-i} + \alpha \tilde{a}_0^{(i-1)}, & m = 0 \\ (1 - \alpha^2) \tilde{a}_0^{(i-1)} + \alpha \tilde{a}_1^{(i-1)}, & m = 1 \\ \tilde{a}_{m-1}^{(i-1)} + \alpha (\tilde{a}_m^{(i-1)} - \tilde{a}_{m-1}^{(i)}), & m = 2, \dots, M \end{cases} \quad (7.30)$$

$$\tilde{b}_0 = b_0 / \tilde{a}_0^{(0)}, \tilde{a}_m = \tilde{a}_m^{(0)} / \tilde{a}_0^{(0)}, \quad 1 \leq m \leq \tilde{M} \quad (7.31)$$

$$\tilde{C}_m = \begin{cases} \log \tilde{b}_0, & m = 0 \\ \tilde{a}_1, & m = 1 \\ -\tilde{a}_m - \frac{1}{m} \sum_{k=1}^{m-1} k \tilde{C}_k \tilde{a}_{m-k}, & m \geq 2 \end{cases} \quad (7.32)$$

の漸化式が導ける [74]。

したがって、2つの音声データ間のメルスケール上のスペクトル包絡距離は

$$\begin{aligned} d(\tilde{\mathbf{a}}, \tilde{\mathbf{a}}') &= \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \{ \log |H_{lpc}(e^{j\tilde{\Omega}})|^2 - \log |H'_{lpc}(e^{j\tilde{\Omega}})|^2 \} d\tilde{\Omega} \right]^{\frac{1}{2}} \\ &= \left[ (\tilde{C}_0 - \tilde{C}'_0)^2 + 2 \sum_{m=1}^{\infty} (\tilde{C}_m - \tilde{C}'_m)^2 \right]^{\frac{1}{2}} \end{aligned} \quad (7.33)$$

と表せ、LPCメルスケール係数  $\tilde{C}_m, \tilde{C}'_m$  間のユークリッド距離に等しく、さらにこれは (7.29) ~ (7.32) 式を通して各線形予測係数  $a_m, a'_m$  から求められる。

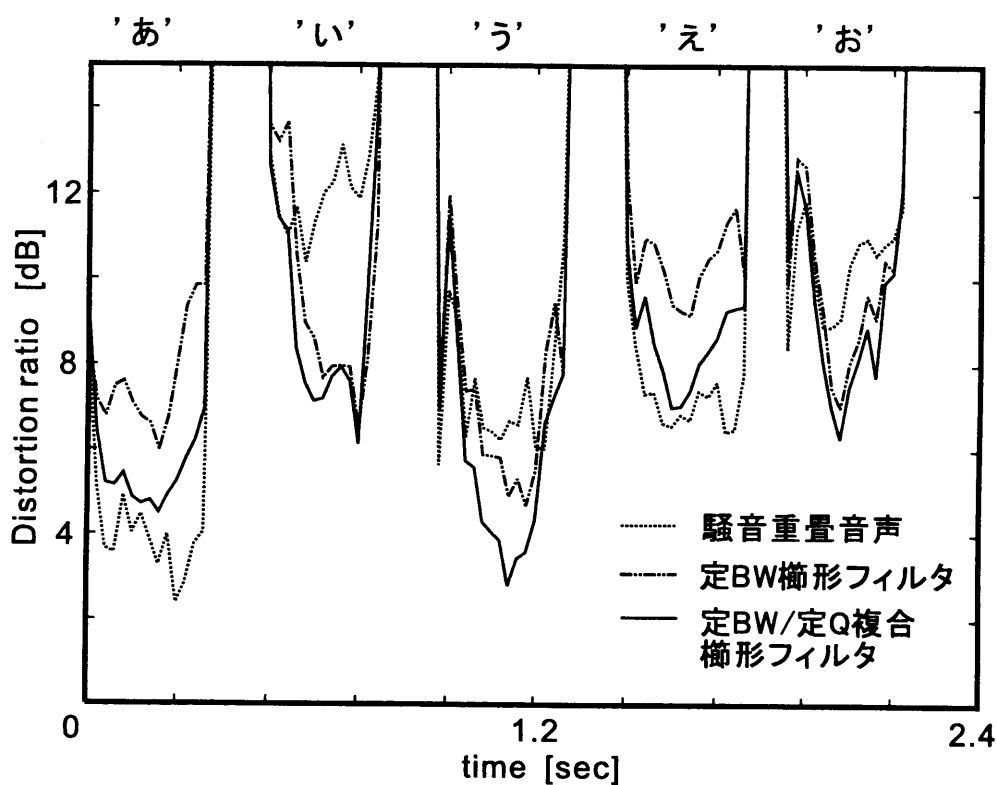
### 7.5.3 実音声に対する適用例

振幅変動やピッチ変動をともなった実音声に定BW/定Q楕形フィルタを適用した場合の性能評価を、上記2種類の距離尺度を用いて行う。

「あいうえお」の母音音声に白色 Gauss 雑音を付加したものを観測音として、振幅変動のみを考慮した定BW型および振幅/ピッチ変動をともに考慮した定BW/定Q複合型の2種類の楕形フィルタを使って雑音下の音声を抽出する。(7.11)式によって楕形フィルタを構成する際のピッチパラメータの軌跡  $\bar{\omega}_0(kT_s)$  は、事前に観測音から第5章の手法を用いて推定しておく。また各楕形フィルタの帯域幅を決めるパラメータ  $\gamma_n(kT_s)$  は、簡単のため時間に依存せず一定とし、さらに観測音から直接推定することは困難なので、原音声を使って調べた振幅変動とピッチ変動の分散値により設定した。サンプリング周波数  $1/T_s$  を 16kHz、フィルタの倍音次数  $N$  を 30、人の聴覚特性に合うようにメルスケールパラメータ  $\alpha$  を 0.31 に選び、定BWおよび定BW/定Q複合楕形フィルタのそれぞれの出力結果を原音声データと比較した。

具体的には、まず各音声データからLPC法により線形予測係数  $a_m$  を推定し、それを使って(7.23)式および(7.29)~(7.32)式からそれぞれLPCケプストラム係数  $C_m$ 、LPCメルケプストラム係数  $\tilde{C}_m$  を求める。これを原音声、観測音、定BW楕形フィルタの出力、定BW/定Q複合楕形フィルタの出力のそれぞれに対して求め、原音声との距離を(7.25)、(7.33)式によって計算した結果が図7.14(a),(b)である。

いずれも複合型の方が定BW型のものに比べ全体的に良好な結果を示している。これは、ピッチ推定時の誤差やピッチ変動による影響をフィルタ特性に反映させたか否かの差であると解釈できる。特にピッチ変動を考慮しない定BW型では、部分的には極端な性能劣



(a) LPC ケプストラムによる比較



(b) LPC メルケプストラムによる比較

図 7.14 スペクトル包絡間の距離尺度を用いたフィルタ性能比較



化が見られる。このことから、比較的安定した性能をもつ定BW/定Q複合楕形フィルタを音声認識の前処理として用いることによって、騒音環境下での認識率を全体的に向上できる可能性が期待される。

「あ」や「え」の部分では、何もフィルタ処理しない方がむしろ良いという結果が現れている。これは定BW/定Q複合楕形フィルタであっても吸収しきれないピッチ誤差や変動があるためと考えられ、楕形フィルタがさらに有効に働くためには、ピッチ推定などのパラメータ決定の精度を高めるなどの改善策を検討する必要がある。

## 7.6 本章のまとめ

本章では、これまで連続信号領域で導かれてきた楕形フィルタをデジタルフィルタとして実現するための設計法について考察した。特に、6章でのKalmanフィルタ方程式を直接離散化し差分方程式に書き直した結果として、全極型のIIRフィルタの形式が自然に導けることを示した(7.2節)。同じ結論は、予め離散表現しておいた信号の状態空間モデルや観測モデルに対するKalmanフィルタ解からも導くことができた(付録7.7.1)。このIIRフィルタの形式は、ALEやANFなどで用いられているノッチフィルタとも密接な関係がある(7.3節)。以上で得られたIIR型の定BWおよび定Q楕形フィルタを用いて、振幅変動やピッチ変動に対する頑健性やそれらが波形/スペクトルに及ぼす影響、実音声と楽音の混合音に適用した場合などの評価実験を行った(7.4節)。特に、変動量に見合った極配置をとることによって最良のSN比で目的信号の分離が可能なが確かめられた。最後に、LPCケプストラムとLPCメルケプストラムによるスペクトル包絡としての再現性の比較を行い、ピッチ変動の影響も考慮した定BW/定Q複合楕形フィルタの優位性を検証した(7.5節)。

## 7.7 付録

### 7.7.1 離散Kalmanフィルタによる別解法

簡単のため単一の倍音信号を対象とし、(6.4)式を離散化した信号モデル

$$x_n(kT_s) = c_n \frac{\bar{x}_n(kT_s)}{\bar{x}_n(0)} e^{-\sum_{i=1}^k \sigma_{an}^2(iT_s)T_s + \sum_{i=1}^k \sigma_{an}(iT_s)\Delta\beta_{an}(iT_s)} \cdot e^{j\{\sum_{i=1}^k \sigma_{bn}(iT_s)\Delta\beta_{bn}(iT_s) + n \sum_{i=1}^k \sigma_{\omega}(iT_s)\Delta\beta_{\omega}(iT_s)\}} e^{jn \sum_{i=1}^k \bar{\omega}_0(iT_s)T_s} \quad (7.34)$$

を出発点に考える。ただし、各 $\Delta\beta_{an}(iT_s)$ 、 $\Delta\beta_{bn}(iT_s)$ 、 $\Delta\beta_{\omega}(iT_s)$ は、

$$E[\Delta\beta(iT_s)] = 0,$$

$$E[\Delta\beta(iT_s)\Delta\beta(jT_s)] = \begin{cases} T_s, & i = j \\ 0, & i \neq j \end{cases} \quad (7.35)$$

が成り立つ互いに無相関な白色 Gauss 雑音である。このとき、(3.85) 式より

$$E[e^{2\sigma_{an}(iT_s)\Delta\beta_{an}(iT_s)}] = e^{2\sigma_{an}^2(iT_s)T_s} \quad (7.36)$$

となるので、(6.6) 式と同様に、

$$\begin{aligned} E[|x_n(kT_s)|^2] &= \bar{x}_n^2(kT_s) e^{-\sum_{i=1}^k 2\sigma_{an}^2(iT_s)T_s} \prod_{i=1}^k E[e^{2\sigma_{an}(iT_s)\Delta\beta_{an}(iT_s)}] \\ &= \bar{x}_n^2(kT_s) \end{aligned} \quad (7.37)$$

が成り立つ。

上式は漸化式として

$$\begin{aligned} x_n(kT_s) &= \frac{\bar{x}_n(kT_s)}{\bar{x}_n((k-1)T_s)} e^{-\sigma_{an}^2(kT_s)T_s + \sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)} \\ &\quad \cdot e^{j\{\sigma_{bn}(kT_s)\Delta\beta_{bn}(kT_s) + n\sigma_\omega(kT_s)\Delta\beta_\omega(kT_s)\}} e^{jn\bar{\omega}_0(kT_s)T_s} x_n((k-1)T_s) \end{aligned} \quad (7.38)$$

と表せることに着目し、振幅やピッチ周波数のランダム成分を第2項に分離して、

$$x_n(kT_s) = \alpha_n(kT_s) e^{jn\bar{\omega}_0(kT_s)T_s} x_n((k-1)T_s) + u(kT_s), \quad (7.39)$$

$$\text{where } u(kT_s) = \Delta C_n(kT_s) \alpha_n(kT_s) e^{jn\bar{\omega}_0(kT_s)T_s} x_n((k-1)T_s),$$

$$\alpha_n(kT_s) = \frac{\bar{x}_n(kT_s)}{\bar{x}_n((k-1)T_s)} e^{-\frac{1}{2}\{\sigma_{an}^2(kT_s) + \sigma_{bn}^2(kT_s) + n^2\sigma_\omega^2(kT_s)\}T_s},$$

$$\begin{aligned} \Delta C_n(kT_s) &= e^{-\frac{1}{2}\sigma_{an}^2(kT_s)T_s + \sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)} e^{\frac{1}{2}\sigma_{bn}^2(kT_s)T_s + j\sigma_{bn}(kT_s)\Delta\beta_{bn}(kT_s)} \\ &\quad \cdot e^{\frac{1}{2}n^2\sigma_\omega^2(kT_s)T_s + jn\sigma_\omega(kT_s)\Delta\beta_\omega(kT_s)} - 1 \end{aligned}$$

と書き直したものをあらためて信号の状態空間モデルとする。右辺第2項の  $u(kT_s)$  は状態依存性雑音を意味し、特に伊藤積分の定義(6.8.2項)に見習って、雑音源  $\Delta\beta((kT_s))$  は信号  $x_n((k-1)T_s)$  に対して時間的に進んだ独立なランダム変動をもつものとする<sup>†</sup>。一方、観測モデルは、(6.13) 式から直ちに

$$y(kT_s) = x_n(kT_s) + v(kT_s), \quad \text{where } v(kT_s) = \frac{1}{T_s} \sigma_v(kT_s) \Delta\beta_v(kT_s) \quad (7.40)$$

と表せる。

以上の離散信号モデルを基にした Kalman フィルタの問題を考える。特に、連続 Kalman フィルタでの結論(6.4節)によれば、状態依存性を厳密に考慮して導いた解が、状態依存で

<sup>†</sup>このような定義に基づく差分方程式は、伊藤型確率微分方程式の名に対応させて「伊藤型確率差分方程式」と呼ぶにふさわしい。

あることを特に意識することなく、通常通り雑音項を扱って求めた解に一致することが確かめられている。同様な考察は離散 Kalman フィルタにおいても同様に成り立つと考えられるので、状態依存性雑音  $u(kT_s)$  を通常の雑音項として扱ってフィルタ解を求めればよい。

(3.85),(3.90) 式より、

$$\begin{aligned} E[e^{\sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)}] &= e^{\frac{1}{2}\sigma_{an}^2(kT_s)T_s}, & E[e^{j\sigma_{bn}(kT_s)\Delta\beta_{bn}(kT_s)}] &= e^{-\frac{1}{2}\sigma_{bn}^2(kT_s)T_s}, \\ E[e^{jn\sigma_{\omega}(kT_s)\Delta\beta_{\omega}(kT_s)}] &= e^{-\frac{1}{2}n^2\sigma_{\omega}^2(kT_s)T_s}, & E[e^{2\sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)}] &= e^{2\sigma_{an}^2(kT_s)T_s} \end{aligned} \quad (7.41)$$

となるので、

$$\begin{aligned} E[\Delta C_n(kT_s)] &= e^{-\frac{1}{2}\sigma_{an}^2(kT_s)T_s} E[e^{\sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)}] e^{\frac{1}{2}\sigma_{bn}^2(kT_s)T_s} E[e^{j\sigma_{bn}(kT_s)\Delta\beta_{bn}(kT_s)}] \\ &\quad \cdot e^{\frac{1}{2}n^2\sigma_{\omega}^2(kT_s)T_s} E[e^{jn\sigma_{\omega}(kT_s)\Delta\beta_{\omega}(kT_s)}] - 1 \\ &= 0 \\ E[|\Delta C_n(kT_s)|^2] &= e^{-\sigma_{an}^2(kT_s)T_s} E[e^{2\sigma_{an}(kT_s)\Delta\beta_{an}(kT_s)}] e^{\sigma_{bn}^2(kT_s)T_s} e^{n^2\sigma_{\omega}^2(kT_s)T_s} - 1 \\ &= e^{\{\sigma_{an}^2(kT_s)+\sigma_{bn}^2(kT_s)+n^2\sigma_{\omega}^2(kT_s)\}T_s} - 1 \end{aligned} \quad (7.42)$$

が成り立つ。この関係を用いると、

$$\begin{aligned} E[u(kT_s)] &= E[\Delta C_n(kT_s)]\alpha_n(kT_s)e^{jn\bar{\omega}_0(kT_s)T_s} E[x_n((k-1)T_s)] \\ &= 0 \end{aligned} \quad (7.43)$$

となり、 $i > j$  のとき、

$$\begin{aligned} E[u(iT_s)u^*(jT_s)] &= E[\Delta C_n(iT_s)]\alpha_n(iT_s)\alpha_n(jT_s)e^{jn\{\bar{\omega}_0(iT_s)-\bar{\omega}_0(jT_s)\}T_s} \\ &\quad \cdot E[x_n((i-1)T_s)x_n^*((j-1)T_s)\Delta C_n^*(jT_s)] \\ &= 0 \end{aligned} \quad (7.44)$$

同様に  $i < j$  のとき、

$$E[u(iT_s)u^*(jT_s)] = 0 \quad (7.45)$$

さらに  $i = j$  のとき、

$$\begin{aligned} E[|u(iT_s)|^2] &= E[|\Delta C_n(iT_s)|^2]\alpha_n^2(iT_s)E[|x_n((i-1)T_s)|^2] \\ &= \bar{x}_n^2(iT_s)(1 - e^{-\{\sigma_{an}^2(iT_s)+\sigma_{bn}^2(iT_s)+n^2\sigma_{\omega}^2(iT_s)\}T_s}) \equiv \sigma_u^2(iT_s) \end{aligned} \quad (7.46)$$

となるので、

$$E[u(iT_s)u^*(jT_s)] = \begin{cases} \sigma_u^2(iT_s), & i = j \\ 0, & i \neq j \end{cases} \quad (7.47)$$

が成り立つ。また、

$$\begin{aligned} E[v(iT_s)] &= 0, \\ E[v(iT_s)v^*(jT_s)] &= \begin{cases} \frac{1}{T_s}\sigma_v^2(iT_s), & i = j \\ 0, & i \neq j \end{cases} \end{aligned} \quad (7.48)$$

も容易に示せる。

Kalman フィルタの結論 [89] によれば、雑音項に関するこれらの条件の下で、(7.39), (7.40) 式に対する推定値は、漸化式

$$\hat{x}_n(kT_s) = \left\{ 1 - \frac{p_n(kT_s)}{\sigma_v^2(kT_s)} T_s \right\} A_n(kT_s) \hat{x}_n((k-1)T_s) + \frac{p_n(kT_s)}{\sigma_v^2(kT_s)} T_s y(kT_s) \quad (7.49)$$

で与えられ、誤差分散  $p_n(kT_s) = E[|x_n(kT_s) - \hat{x}_n(kT_s)|^2]$  に関して

$$p_n(kT_s) = \left[ \{A_n(kT_s)p_n((k-1)T_s)A_n^*(kT_s) + \sigma_u^2(kT_s)\}^{-1} + \frac{T_s}{\sigma_v^2(kT_s)} \right]^{-1} \quad (7.50)$$

が成り立つ。ただし、 $A_n(kT_s) = \alpha_n(kT_s)e^{jn\bar{\omega}_0(kT_s)T_s}$  とおいた。特に、 $p_n(kT_s) = p_n((k-1)T_s)$  となる準定常状態の場合を考えれば、上式の解は

$$p_n(kT_s) = \frac{\sigma_v^2(kT_s)}{2|A_n(kT_s)|^2 T_s} \left\{ -B_n(kT_s) + \sqrt{B_n^2(kT_s) + 4|A_n(kT_s)|^2 \frac{\sigma_u^2(kT_s)}{\sigma_v^2(kT_s)}} \right\}, \quad (7.51)$$

$$\text{where } B_n(kT_s) = 1 + \frac{\sigma_u^2(kT_s)}{\sigma_v^2(kT_s)} T_s - |A_n(kT_s)|^2 \quad (7.52)$$

となる。振幅 / ピッチ変動が十分小さく、 $\frac{\sigma_u^2(kT_s)}{\sigma_v^2(kT_s)} T_s \ll 1$ ,  $|A_n(kT_s)| \approx 1$  を仮定すれば、

$$\begin{aligned} p_n(kT_s) &\approx \sigma_v(kT_s)\sigma_u(kT_s)\frac{1}{\sqrt{T_s}} \\ &\approx \bar{x}_n(kT_s)\sigma_v(kT_s)\sigma_n(kT_s) \end{aligned} \quad (7.53)$$

と近似できる。これを (7.49) 式に代入することによって (7.5) 式が得られる。

## 第 8 章

### 結論

本論文では、振幅変動やピッチ変動をともなった周期性が近似的にしか成り立たないような信号を雑音環境下で適切に抽出するためのフィルタ設計法について論じた。具体的には、

- (1) 「振幅変動+ピッチ変動」信号とその推定問題に対する定義、
- (2) その定常モデルと非定常モデルのそれぞれに対する線形最適フィルタの設計、
- (3) 最適解として得られた楕円フィルタの時間 / 周波数特性に成り立つ関係の解明、
- (4) デジタルフィルタによる実現、
- (5) 混合音に対するピッチ推定と各擬似周期信号の個別分離アルゴリズムの開発

を主な論点とした。

以下に、各章で得られた結果を要約する。

まず第 1 章では、時間周波数解析、聴覚フィルタ、聴覚情景分析、楕円フィルタ、バイスペクトル分析などの既存研究との関係について論じ、擬似周期信号推定の問題に対する本研究の立場は、あくまでも線形最適フィルタ設計を中心とした純信号処理的な議論にあることなど、本論文の方向性を明確化した。

第 2 章では、「振幅変動+ピッチ変動」の物理的根拠として時変係数をとる波動方程式を取り上げ、その近似解から信号モデルを定式化し、各種確率モデルの導入を前提としたフィルタ設計問題の定義を与えた。特に、議論の対象となる「振幅変動+ピッチ変動」信号モデルが、音響・音声信号の性質を比較的良好に反映したものになっていることの根拠についても言及した。

第 3 章では、「振幅変動+ピッチ変動」に微細な不規則変動を許容するかわりに定常性を仮定した場合のフィルタ設計を行った。1) 振幅とピッチを一定とした理想周期信号の場合にはインパルス列からなる楕円フィルタ、2) 振幅変動のみを考慮した場合には定

BW(Band-Width) 型の楕形フィルタ、3) ピッチ変動のみを考慮した場合には定 Q 型の楕形フィルタ、さらに 4) 両者を同時に考慮した場合には低次倍音から高次倍音になるに従って定 BW 型から定 Q 型に徐々に遷移する複合楕形フィルタがそれぞれ導かれることを示した。特に、振幅変動とピッチ変動の分散値を適切に設定すると、各バンドパスフィルタの占有帯域幅が聴覚末梢系のもつ非線形な周波数特性によく適合することを確かめた。

第 4 章では、ピッチ変動のみを考慮した場合の最適フィルタである定 Q 楕形フィルタの時間/周波数特性に成り立つ関係について論じた。まず、観測信号を擬似周期信号モデルに最小二乗フィッティングさせたときの複素振幅パラメータの推定値から、楕形フィルタの時間領域における一般形式が導かれることを示した。特に、ピッチ周波数の不確定性を考慮した場合のインパルス応答から定 Q 楕形関数を定義し、その Fourier 変換が再び定 Q 楕形関数になることを定理の形で与えた。その応用例として、ピッチ誤差によって生じる歪みが定 Q 楕形フィルタを用いることで小さく抑えられることを時間 / 周波数特性の両面から検証した。

第 5 章では、混合音中の複数のピッチ候補から目的ピッチを選択的に追跡しつつ、その軌跡上の調波成分を分離抽出するアルゴリズムを提案した。ここでは、複数のピッチ候補を尤度関数の多峰性により表現し、その中でピッチ変化のダイナミクスに従うものだけが最終的な目的ピッチとして選別される仕組みを Non-Parametric Kalman フィルタにより実現した。目的信号抽出のための定 Q 楕形フィルタ演算は、Wavelet 空間においてピッチ軌跡から決まる各倍音周波数上の複素振幅を読みとり合成することによって等価的に実現できる。本アルゴリズムによって混合音中のピッチ軌跡が選択的に補間追跡でき、目的ストリームが分離再生できることを、シミュレーションおよび実音声実験により検証した。

第 6 章では、第 3 章の信号モデルに対して振幅やピッチに既知の緩やかな変化を許容した時変モデルに対するフィルタ設計を行った。ここでは、信号の状態空間モデルを伊藤型の確率微分方程式で記述することによって、連続 Kalman フィルタの方程式を導出した。特に倍音間の干渉を無視し、振幅変動やピッチ変動のパラメータを時不変としたときのフィルタ方程式の定常解は、第 3 章の Wiener フィルタ解に一致し、倍音間の干渉を考慮した厳密解では、倍音間の干渉を互いに排除するためのノッチ特性が現れることを新たに見出した。

第 7 章では、前章で導いた Kalman フィルタ方程式が離散化近似によって全極型のデジタルフィルタとして実現可能なことを示した。計算機上で合成した各種の「振幅変動 + ピッチ変動」信号に対して、その変動量に見合った極配置をとることによって最良の SN 比で目的信号の分離が可能なことを検証した。特に、極と正の実軸とのなす角度はピッチ周波数によって、極と単位円との距離は振幅 / ピッチ変動および背景雑音の分散、平均二乗振幅によってそれぞれ決定される。SN 比および波形 / スペクトル歪みの比較実験において、ピッチ誤差 / 変動に対しては定 Q 楕形フィルタが、振幅変動に対しては定 BW 楕形フィル

タが有利になることが確認できた。また、混合音から音声や楽音を個別抽出する実験では、ピッチ誤差 / 変動の影響が支配的であり、定 Q 櫛形フィルタの方が少ないスペクトル歪みで抽出できることが確かめられた。さらに、LPC ケプストラムやメルケプストラムを用いて計算したスペクトル包絡の比較実験では、振幅変動とピッチ変動の両者を同時に考慮した定 BW / 定 Q 複合櫛形フィルタが原音声を最も忠実に再現できることを確かめた。

以上で得られた成果を要約すれば、

- (1) 擬似周期信号推定にとっての線形最適フィルタとは、低次倍音では定 BW 櫛形特性、高次倍音では定 Q 櫛形特性、低次から高次までの中間領域では、定 BW 型から定 Q 型に徐々に遷移する複合櫛形特性を有すること、
  - (2) そのフィルタ特性には、聴覚系のもつ非線形な周波数特性と密接な関係があること、
  - (3) 特に定 Q 櫛形特性のものは、インパルス応答においても定 Q 櫛形分布をもつという際だった性質があること、
  - (4) それらは全極型の IIR フィルタとして容易に実現できること、
  - (5) ピッチ推定との統合化が Wavelet 解析を通して実現できること、
  - (6) 実用的には、騒音環境下における音声認識の前処理としての利用が期待できること、
- が本論文において明らかとなった。

今後の課題としては、

- (1) フィルタ設計において既知としていた信号と雑音のエネルギー比や振幅 / ピッチ変動の分散などのパラメータの決定法、もしくはフィルタ係数を自動調整する適応アルゴリズムの開発、
  - (2) 倍音間の干渉を排除するノッチ特性を含めたデジタルフィルタの構成法、
  - (3) 音声の子音部などの擬似周期性が成り立たない信号成分に対するフィルタ処理を含めた音声強調統合システムへの発展、
- などがあげられる。

## 謝辞

本論文は、安藤教授のご指導の下で実に13年間の年月をかけて完成させたもので、実質的な内容としてはこの5年間の研究成果の集大成であります。

安藤先生から直接ご指導を受けたのは、さらに遡って筆者が電気通信大学大学院修士課程に在学中の時からになるため、筆者の研究者としての形成のほとんどは、先生からの影響によるものであると言っても過言ではありません。とりわけ混沌としがちな思考パターンを整理し、簡潔明解な論理へと導くことの重要性を繰り返し具体的に示して下さいたことは、研究だけでなく生きてゆく糧として大いに役立つものと思います。思考力、洞察力、知識量、努力量などそのどれをとっても先生が期待されるものには遠く及ばなかった筆者を、この長い年月に渡って見捨てずに見守って下さったことは、何にもまして有難く感じております。

計数工学科の木村教授、舘教授、嵯峨山教授、篠田助教授には、審査を通じて有益なご助言を頂きました。

安藤研究室の来海助手、博士課程の小野さんをはじめとする研究室の新旧メンバーの方々とは、輪講等を通して多くの議論をして下さった上に、他大学の筆者に対していつも暖かく接していただきました。特に現ソニーの安部さんには、音声処理の研究に興味をもつきっかけを与えていただきました。

筆者が所属する電通大でも多くの先生方、学生の方々の援助がありました。特に研究室の小菅先生には、本来果たすべき日頃の雑務を極力軽減し、研究に専念できる環境を維持していただきました。武田先生や岩倉先生には、多くの相談に乗っていただいた上、なかなか成果のあがらない筆者を根気よく待ちつづけて下さり、また三橋先生には、気が滅入りがちな筆者に多くの勇気を与えていただきました。研究室の学生の方々には、数値計算等の手助けをしてもらいました。

審査発表会や申請手続きなどを支えて下さった事務の方々にも感謝申し上げます。

最後に、妻をはじめとして家族の支えがなければ、決して本論文を完成するには至らなかったことを強調しておきます。



## 参考文献

- [1] 西 一樹, “擬似周期信号の線形最適フィルタ理論 (その1) – 定常信号モデルと Wiener フィルタ解析 –”, 第15回デジタル信号処理シンポジウム講演論文集, pp.179-184, 2000.
- [2] 西 一樹, 安藤 繁, “重み付き最小二乗法に基づいた調波音推定”, 電子情報通信学会技術研究報告, SP96-87, pp.1-6, 1997.
- [3] 西 一樹, 安藤 繁, “定 Q 櫛形フィルタ: ピッチ推定誤差を許容する調波信号抽出”, 電子情報通信学会技術研究報告, EA97-4, pp.23-28, 1997.
- [4] 西 一樹, 安藤 繁, “調波信号抽出のための定 Q 櫛形フィルタ: そのインパルス応答と伝達特性”, 電子情報通信学会技術研究報告, EA97-20, pp.1-6, 1997.
- [5] 西 一樹, 安藤 繁, “定 Q 櫛形フィルタ: ピッチ推定誤差に頑健な倍音抽出法”, 計測自動制御学会第36回学術講演会予稿集, pp.545-546, 1997.
- [6] 西 一樹, 安藤 繁, “定 Q 櫛形フィルタによる調波信号抽出”, 電子情報通信学会基礎・境界ソサイエティ大会講演論文集, pp.74, 1997.
- [7] 西 一樹, 安藤 繁, “調波音抽出のための定 Q 櫛形フィルタ – その最適性とインパルス応答 –”, 日本音響学会秋季研究発表会講演論文集, pp.577-578, 1997.
- [8] 西 一樹, 安藤 繁, “櫛形ウェーブレット変換による調波信号解析”, 第12回デジタル信号処理シンポジウム講演論文集, pp.321-326, 1997.
- [9] 西 一樹, 安藤 繁, “調波信号センシングのための最適フィルタ設計とそのウェーブレット解析への応用”, 電気学会センサシステム応用研究会資料, SSA97-16, pp.91-96, 1997.
- [10] K.Nishi and S.Ando, “An Optimum Comb Filter for Time-Varying Harmonics Extraction”, *IEICE Trans. Fundamentals*, vol.E81-A, no.8, pp.1622-1627, Aug. 1998.

- [11] 西 一樹, 安藤 繁, “定 Q 櫛形フィルタとその時間/周波数特性 - ピッチ変動誤差に頑健な調波信号推定のためのフィルタ特性 -”, 電子情報通信学会論文誌, vol.J83-A, no.2, pp.152-160, Feb. 2000.
- [12] 西 一樹, 安藤 繁, “時間周波数領域における楽音の分離抽出-確率分布観測型カルマンフィルタによる周期成分追跡機構を利用して-”, 計測自動制御学会第 34 回学術講演会予稿集, pp.949-950, 1995.
- [13] 西 一樹, 安部 素嗣, 安藤 繁, “Wavelet 空間における特定倍音成分の分離抽出”, 電子情報通信学会技術研究報告, EA95-39, pp.29-35, 1995.
- [14] 西 一樹, 安部 素嗣, 安藤 繁, “楽音中の特定倍音成分を分離抽出・追跡する適応フィルタ”, 第 10 回デジタル信号処理シンポジウム講演論文集, pp.135-140, 1995.
- [15] 西 一樹, 安部 素嗣, 安藤 繁, “最適倍音追跡フィルタによる多重音声の分離抽出”, 電子情報通信学会技術研究報告, SP95-148, pp.45-52, 1996.
- [16] 西 一樹, 安部 素嗣, 安藤 繁, “最適倍音追跡フィルタによる音響ストリーム分離”, 電子情報通信学会総合大会講演論文集, D-685, Mar. 1996.
- [17] K.Nishi, S.Ando and S.Aida, “Optimum Harmonics Tracking Filter for Auditory Scene Analysis,” *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Proc.*, pp.573-576, Atlanta, May 1996.
- [18] K.Nishi, M.Abe and S.Ando, “Optimum Sound Tracking Systems for Auditory Scene Analysis,” *Technical Digest of the 14th Sensor Symposium*, pp.93-96, Kawasaki, Jun. 1996.
- [19] 西 一樹, 安部 素嗣, 安藤 繁, “聴覚情景分析のための最適倍音追跡システム”, 計測自動制御学会第 35 回学術講演会予稿集, pp.673-674, 1996.
- [20] 西 一樹, 安部 素嗣, 安藤 繁, “聴覚情景分析のための多重ピッチ追跡と調波分離アルゴリズム”, 計測自動制御学会論文集, vol.34, no.6, pp.483-490, June 1998.
- [21] 西 一樹, 安藤 繁, “ピッチ変動と振幅変動に頑健性を有する調波信号抽出フィルタ”, 電気学会論文誌, vol.119-E, no.4, pp.215-220, Apr. 1999.
- [22] K.Nishi, “Time-Varying Harmonics Enhancement Based on Kalman-Bucy Filter,” *Proc. Int. Conf. on Speech Processing*, pp.433-438, Seoul, Aug. 1999.

- [23] 西 一樹, 安藤 繁, “振幅変動とピッチ変動を同時に考慮した擬似周期信号抽出のための IIR 櫛形デジタルフィルタ”, 電子情報通信学会技術研究報告, DSP99-78, pp.29-34, 1999.
- [24] K.Nishi and S.Ando, “IIR Comb Filters for Time-Varying Harmonics Extraction Based on Kalman-Bucy Filter Analysis,” *Proc. ISCIE Int. Symp. on Stochastic Systems Theory and Its Applications*, pp.235-240, Yokohama, Nov. 1999.
- [25] 西 一樹, 安藤 繁, “Kalman-Bucy フィルタ解析に基づく IIR 櫛形デジタルフィルタ”, 第 14 回デジタル信号処理シンポジウム講演論文集, pp.91-96, 1999.
- [26] 西 一樹, “擬似周期信号の線形最適フィルタ理論 (その 2) – 伊藤確率微分方程式による時変信号モデルと Kalman フィルタ解析 –”, 第 15 回デジタル信号処理シンポジウム講演論文集, pp.185-190, 2000.
- [27] K.Nishi, “Kalman Filter Analysis for Quasi-Periodic Signals,” *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Proc.*, accepted, Salt Lake City, May 2001.
- [28] A.S.Bregman, *Auditory Scene Analysis -the perceptual organization of sound-*, MIT Press, 1990.
- [29] A. de Cheveigne, “Separation of concurrent harmonic sounds: fundamental frequency estimation and a time-domain cancellation model of auditory processing,” *J. Acoust. Soc. Am.*, vol.93, no.6, pp.3271–3290, June 1993.
- [30] 前田 恵介, 菊地 義信, 粕谷 英樹, “病的音声に含まれる声門雑音を検出するためのくし形フィルタに関する検討”, 電子情報通信学会論文誌, vol.J74-A, no.4, pp.731–733, Apr. 1991.
- [31] P.Lieberman, “Perturbations in vocal pitch,” *J. Acoust. Soc. Am.*, vol.33, no.5, pp.597–603, 1961.
- [32] A.J.Rozsypal and B.F.Millar, “Perception of jitter and shimmer in synthetic vowels,” *J. Phonetics*, vol.7, pp.343–355, 1979.
- [33] S.Hiki, K.Sugawara and J.Oizumi, “On the rapid fluctuation of voice pitch,” 音響学会誌, vol.22, pp.290–291, 1966.
- [34] 坂本 勤, 粕谷 英樹, “合成音の自然性に寄与する基本周期の変動特性”, 音響学会秋季研究発表会講演論文集, 2-3-8, pp.169–170, Sep. 1985.

- [35] H.Kasuya et al., "An acoustic analysis of pathological voice and its application to the evaluation of laryngeal pathology," *Speech Commun.*, vol.5, pp.171-181, 1986.
- [36] 古井, デジタル音声処理, 東海大学出版会, 1985.
- [37] 北風 裕教, 赤木 正人, "基本周波数の微細変動成分に対する知覚", 電子情報通信学会技術研究報告, SP98-168, pp.9-16, Mar. 2000.
- [38] 皆川 知也, 赤木 正人, "連続発話母音の基本周波数変動とその知覚", 音響学会聴覚研究会資料, H-98-29, 1998.
- [39] 小室 修, 粕谷 英樹, "基本周期のゆらぎの性質とそのモデル化に関する検討", 音響学会誌, vol.47, no.12, pp.928-934, 1991.
- [40] 青木 直史, 伊福部 達, "持続発声母音における振幅ゆらぎ及びピッチゆらぎの周波数特性とその音響心理的效果", 電子情報通信学会論文誌, vol.J82-A, no.5, pp.649-657, May 1999.
- [41] 遠藤 康男, 粕谷 英樹, "周期ごとのゆらぎを考慮した音声の分析・変換・合成システム", 電子情報通信学会論文誌, vol.J81-A, no.7, pp.1031-1041, July 1998.
- [42] 楽器の科学, コロナ社, 1992.
- [43] F.Hlawatsch and G.F.Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Process. Mag.*, pp.21-66, Apr. 1992.
- [44] P.Guillemain and R.Kronland-Martinet, "Characterization of acoustic signals through continuous linear time-frequency representations," *Proc. IEEE*, vol.84, no.4, pp.561-585, Apr. 1996.
- [45] 境 久雄, 聴覚と音響心理, コロナ社, 1992.
- [46] G.von Békésy, *Experiments in Hearing*, McGraw-Hill, New York, 1960.
- [47] H.Fletcher, "Auditory patterns," *Review of Modern Physics*, vol.12, pp.47-65, 1940.
- [48] 赤木 正人, "聴覚フィルタとそのモデル", 電子情報通信学会誌, vol.77, no.9, pp.948-956, 1994.
- [49] E.Zwicker, "Subdivision of the audible frequency range into critical bands," *J. Acoust. Soc. Am.*, vol.33, pp.248, 1961.

- [50] E.Zwicker and E.Terhardt, "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency," *J. Acoust. Soc. Am.*, vol.68, no.5, pp.1523-1525, 1980.
- [51] B.C.J.Moore and B.R.Glasberg, "Suggested formulae for calculating auditory-filter band-widths and excitation patterns," *J. Acoust. Soc. Am.*, vol.74, no.3, pp.750-753, 1983.
- [52] B.R.Glasberg and B.C.J.Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research*, vol.47, pp.103-138, 1990.
- [53] E.de Boer and H.R.de Jongh, "On cochlear encoding: Potentialities and limitations of the reverse correlation technique," *J. Acoust. Soc. Am.*, vol.63, no.1, pp.115-135, 1978.
- [54] T.Irino and R.Patterson, "A time-domain level-dependent auditory filter: The gam-macharp," *J. Acoust. Soc. Am.*, vol.101, pp.412-419, 1997.
- [55] R.Patterson, "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," *J. Acoust. Soc. Am.*, vol.72, no.6, pp.1788-1803, 1982.
- [56] 鶴木 祐史, 赤木 正人, "雑音が付加された波形からの信号波形の一抽出法", 電子情報通信学会論文誌, vol.J80-A, no.3, pp.444-453, Mar. 1997.
- [57] 柏野 邦夫, 田中 英彦, "二つの周波数成分の分離知覚に関する工学的モデル - 複数の要因の評価と統合 -", 電子情報通信学会論文誌, vol.J77-A, no.5, pp.731-740, May 1994.
- [58] 中谷 智広, 奥乃 博他, "音環境理解のためのマルチエージェントによる調波構造ストリームの分離", 人工知能学会誌, vol.10, no.2, pp.232-241, 1995.
- [59] T.Nakatani et al., "Localization by harmonic structure and its application to harmonic sound segregation," *Proc. IEEE ICASSP'96*, pp.653-656, Atlanta, 1996.
- [60] G.J.Brown and M.Cooke, "Computational auditory scene analysis," *Computer Speech and Language*, vol.8, pp.297-336, 1994.
- [61] R.J.McAulay and T.F.Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. Acoust., Speech and Signal Process.*, vol.34, pp.744-754, 1986.

- [62] 安部 素嗣, 安藤 繁, “共有 FM-AM の時間周波数統合に基づく聴覚情景解析 (I) – Lagrange 微分特微量とその周波数軸統合”, 電子情報通信学会論文誌, vol.J83-D-II, no.2, pp.458–467, Feb. 2000.
- [63] 安部 素嗣, 安藤 繁, “共有 FM-AM の時間周波数統合に基づく聴覚情景解析 (II) – 最適な時間軸統合とストリーム音の再合成”, 電子情報通信学会論文誌, vol.J83-D-II, no.2, pp.468–477, Feb. 2000.
- [64] M.J. ライトヒル著, 高見 穎部訳, フーリエ解析と超関数, ダイヤモンド社, 1975.
- [65] 小畑 秀文, “最尤推定法に基づく音声波形の復元”, 音響学会誌, vol.50, no.10, pp.809–815, Oct. 1994.
- [66] 長淵 裕実, “コームフィルタリング雑音抑圧法の基本性能と PCM 音声の品質改善への応用”, 電子通信学会論文誌, vol.J65-A, no.6, pp.572–578, 1982.
- [67] 苜木 禎史 他, “可変コムフィルタを用いた音声の抽出”, 電子情報通信学会技術研究報告, EA94-68, pp.25-32, 1994.
- [68] Y.K.Jang and J.F.Chicharo, “Adaptive IIR comb filter for harmonic signal cancellation,” *Int. J. Electronics*, vol.75, no.2, pp.241–250, 1993.
- [69] 安藤 繁, “確率分布観測型カルマンフィルタ”, 第 33 回 SICE 学術講演会予稿集, 108E-2, pp.225–226, 1994.
- [70] 鹿児島 誠一, 振動・波動入門 (新物理学ライブラリ 5), サイエンス社, 1995.
- [71] P.M.Morse, *Vibration and Sound*, 2nd ed., McGraw-Hill, 1948.
- [72] 情報理論とその応用学会編, 確率過程 (応答と話題), 培風館, 1994.
- [73] O.Besson et al., “Parameter estimation for random amplitude chirp signals,” *IEEE Trans. Signal Process.*, vol.47, no.12, pp.3208–3219, Dec. 1999.
- [74] 今井 聖, 音声信号処理, 森北出版, 1996.
- [75] 今井 聖, 住田 一男, 古市 千枝子, “音声合成のためのメル対数スペクトル近似 (MLSA) フィルタ”, 電子通信学会論文誌, vol.J66-A, no.2, pp.122–129, 1983.
- [76] L.Rabiner, B.H.Juang 著, 古井 監訳, 音声認識の基礎 (上), NTT アドバンステクノロジー, 1995.

- [77] A.Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.
- [78] 砂原 善文, 確率システム理論 II,III 巻, 朝倉書店, 1982.
- [79] S.Tomažic, "On short-time Fourier transform with single-sided exponential window," *Signal Processing*, vol.55, pp.141-148, 1996.
- [80] 小倉 久直, 確率過程論 I,II 巻, コロナ社, 1985.
- [81] B. エクセンドール著, 谷口 説男訳, 確率微分方程式, シュプリンガー・フェアラーク東京, 1999.
- [82] R.R.Bitmead et al., "A Kalman filtering approach to short-time Fourier analysis," *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-34, no.6, pp.1493-1501, Dec. 1986.
- [83] S.H.Park et al., "Short-time Fourier analysis via optimal harmonic FIR filters," *IEEE Trans. Signal Process.*, vol.45, no.6, pp.1535-1542, June 1997.
- [84] P.J.McLane, "Optimal linear filtering for linear systems with state-dependent noise," *Int. J. Control*, vol.10, no.1, pp.41-51, 1969.
- [85] L.Arnold, *Stochastic Differential Equations: Theory and Applications*, Krieger Publishing Company, 1992.
- [86] 青島 伸治, "複素 1 次系とその応用", 計測自動制御学会論文集, vol.26, no.7, pp.811-817, July 1990.
- [87] 西村 敏充, 狩野 弘之, 制御のためのマトリクス・リカッチ方程式, 朝倉書店, 1996.
- [88] 今井 功, 応用超関数論 I,II 巻, サイエンス社, 1991.
- [89] 有本 卓, カルマン・フィルター, 産業図書, 1992.
- [90] J.S.Lim, A.V.Oppenheim and L.D.Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-26, no.4, pp.354-365, Aug. 1978.
- [91] D.E.Veeneman and B.Mazor, "A fully adaptive comb filter for enhancing block-coded speech," *IEEE Trans. Acoust., Speech and Signal Process.*, vol.37, no.6, pp.955-957, June 1989.

- [92] 上野 敏行, 舘, 山田 一郎, 藤村 貞夫, “バースペクトルの測定とその応用”, 応用物理, vol.45, no.5, pp.384–396, 1976.
- [93] 舘, “バースペクトル分析による雑音に埋もれた周期信号の分離”, 計測自動制御学会論文誌, vol.9, no.6, pp.729–738, Dec. 1973.
- [94] 植田 光, 太田 智之, 西 一樹, 安藤 繁, “調波信号センシングのためのピッチ同期定 Q 楕形デジタルフィルタ”, 電気学会センサシステム応用研究会資料, SSA98-5, pp.27-32, 1998.
- [95] 西 一樹, 太田 智之, 植田 光, 安藤 繁, “マルチレートサンプリングに基づく定 Q 楕形フィルタの実現”, 電子情報通信学会技術研究報告, EA97-105, pp.17-23, 1998.
- [96] Y.Tadokoro and K.Abe, “Notch Fourier Transform,” *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-35, no.9, pp.1282–1288, Sep. 1987.
- [97] S.D.Hendry, “Computation of harmonic comb filter weights,” *IEEE Trans. Signal Process.*, vol.41, no.4, pp.1677–1680, April 1993.
- [98] A.Nehorai, “A minimal parameter adaptive notch filter with constrained poles and zeros,” *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-33, no.4, pp.983–996, Aug. 1985.
- [99] A.Nehorai and B.Porat, “Adaptive comb filtering for harmonic signal enhancement,” *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-34, no.5, pp.1124–1138, Oct. 1986.
- [100] B.Widrow et al., “Adaptive noise cancelling principles and applications,” *Proc. IEEE*, vol.63, pp.1692–1716, Dec. 1975.
- [101] D.V.B.Rao and S.Y.Kung, “Adaptive notch filtering for the retrieval of sinusoids in noise,” *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-32, pp.791–802, Aug. 1984.
- [102] V.U.Reddy, B.Egardt and T.Kailath, “Optimized lattice-form adaptive line enhancer for a sinusoidal signal in broad-band noise,” *IEEE Trans. Acoust., Speech and Signal Process.*, vol.ASSP-29, no.3, pp.702–709, June 1981.
- [103] S.C.Pei and C.C.Tseng, “Complex adaptive IIR notch filter algorithm and its application,” *IEEE Trans. Circuits and Systems II*, vol.41, no.2, pp.158–163, Feb. 1994.



- [104] L.Ljung and S.Gunnarsson, "Adaptation and tracking in system identification – a survey," *Automatica*, vol.26, pp.7-21, 1990.
- [105] M.V.Dragosevic and S.S.Stankovic, "An adaptive notch filter with improved tracking properties," *IEEE Trans. Signal Process.*, vol.43, no.9, pp.2068–2077, Sep. 1995.
- [106] P.Handel and A.Nehorai, "Tracking analysis of an adaptive notch filter with constrained poles and zeros," *IEEE Trans. Signal Process.*, vol.42, no.2, pp.281–291, Feb. 1994.
- [107] P.Tichavsky and P.Handel, "Two algorithms for adaptive retrieval of slowly time-varying multiple sinusoids in noise," *IEEE Trans. Signal Process.*, vol.43, no.5, pp.1116–1127, May 1995.
- [108] P.Tichavsky and A.Nehorai, "Comparative study of four adaptive frequency trackers," *IEEE Trans. Signal Process.*, vol.45, no.6, pp.1473–1484, June 1997.
- [109] J.F.Chicharo and T.S.Ng, "Gradient-based adaptive IIR notch filtering for frequency estimation," *IEEE Trans. Acoust., Speech and Signal Process.*, vol.38, no.5, pp.769–777, May 1990.
- [110] K. Matsuura et al., "Adaptive line enhancers on the basis of least-squares algorithm for a single sinusoid detection," *IEICE Trans. Fundamentals*, vol.E82-A, no.8, pp.1536–1543, Aug. 1999.
- [111] 蒋海云 他, "单一正弦波检出用複素係数適応 IIR ノッチフィルタの定常特性の解析", *電子情報通信学会論文誌*, vol.J83-A, no.7, pp.850–858, Jul. 2000.
- [112] A.V.Oppenheim and R.W.Schafer, *Digital Signal Processing*, Prentice-Hall, 1975.
- [113] B.F.Boroujeny, "An IIR adaptive line enhancer with controlled bandwidth," *IEEE Trans. Signal Process.*, vol.45, no.2, pp.477–481, Feb. 1997.
- [114] M.T.Kilani and J.F.Chicharo, "A Constrained Notch Fourier Transform," *IEEE Trans. Signal Process.*, vol.43, no.9, pp.2058–2067, Sept. 1995.
- [115] S.C.Pei and C.C.Tseng, "A comb filter design using fractional-sample delay," *IEEE Trans. Circuits and Systems II*, vol.45, no.6, pp.649–653, June 1998.
- [116] W.Hess, *Pitch determination of speech signals*, Springer-Verlag, 1983.

- 
- [117] A.M.Noll, "Cepstrum pitch determination," *J. Acoust. Soc. Am.*, vol.41, no.2, pp.293-309, Feb. 1967.
- [118] 阿部 敏彦, 小林 隆夫, 今井 聖, "瞬時周波数に基づく雑音環境下でのピッチ推定", 電子情報通信学会論文誌, vol.J79-D-II, no.11, pp.1771-1781, Nov. 1996.
- [119] M.R.Schroeder, "Period histogram and product spectrum: New methods for fundamental frequency measurement," *J. Acoust. Soc. Am.*, vol.43, pp.829-834, June 1968.
- [120] D.N.Hermes, "Measurement of Pitch by Subharmonics Summation," *J. Acoust. Soc. Am.*, vol.83, pp.257-264, 1988.
- [121] 阿部 林治, 神林 紀嘉, "調和性スペクトラムモデルによる基本周波数推定アルゴリズム", 音響学会誌, vol.54, no.4, pp.296-304, 1998.
- [122] N.Delprat et al., "Asymptotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequency," *IEEE Trans. Information Theory*, vol.38, pp.644-664, 1992.
- [123] M.Abe and S.Ando, "Application of Loudness/Pitch/Timbre Decomposition Operators to Auditory Scene Analysis," *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Proc.*, pp.2646-2649, Atlanta, May 1996.