

Title	位相変化の音色知覚に及ぼす影響に関する研究
Author(s)	安武, 浩二郎
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1172
Rights	
Description	Supervisor: 赤木 正人, 情報科学研究科, 修士

修 士 論 文

位相変化の音色知覚に及ぼす影響に関する研究

指導教官 赤木 正人 助教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

安武 浩二郎

1998 年 2 月 13 日

要 旨

従来の音色の研究は、人間の聴覚が位相に対して鈍感であると考えられてきたことから専ら定常的なパワースペクトルの概形との関連でなされており、位相や群遅延等の音色に対する影響の系統的検討は、ほとんど行なわれてこなかった。また、ある条件下において位相差が音色の知覚に影響するとする少数の報告例も、実験系の位相補正までは言及しておらず、やや信頼性に欠ける面がある。しかしながら、聴覚における時間情報処理、音色の知覚機構を検討する上で、位相と音色知覚との関連性を明確化することの重要性は近年増してきており、将来的には位相制御による音声合成系への応用も期待される。

そこで本研究では、まず実験系の位相を厳密に補正する逆フィルタを作成し、鼓膜面上で意図する波形に近いものが得られるようにした。そのうえで、位相特性の異なる調波複合音を用いた心理物理実験を行ない、位相に対する聴覚機構の特性を検討した。

目 次

1	はじめに	1
1.1	本研究の背景	1
1.1.1	位相に関する研究	1
1.1.2	電気音響機器および群遅延に関する研究	2
1.1.3	従来の研究の問題点	4
1.2	本研究の特色	5
2	実験系	6
2.1	電気音響機器について	6
2.1.1	ヘッドホン	6
2.1.2	カップラ	10
2.1.3	マイク、マイクロホンプリアンプ、マイクロホンパワーサプライ	11
2.1.4	D A Tリンク	12
2.1.5	D Aコンバーター	12
2.1.6	D A T	12
2.1.7	安定化電源	12
2.2	防音室	12
3	位相補正	14
3.1	位相補正手順	14
3.1.1	インパルス応答	14
3.1.2	群遅延特性	18
3.1.3	逆フィルタ	19

3.2	実験系の群遅延特性の測定	20
3.2.1	Lambda Nova 使用時の群遅延特性	20
3.2.2	HDA200 使用時の群遅延特性	24
3.2.3	安定化電源の有効性	24
3.2.4	考察	27
3.3	逆フィルタ設計	28
3.3.1	逆フィルタの通過周波数帯域 : 60Hz ~ 9.3kHz	29
3.3.2	逆フィルタの通過周波数帯域 : 100Hz ~ 9.3kHz	29
3.3.3	逆フィルタの通過周波数帯域 : 300Hz ~ 9.3kHz	34
3.3.4	逆フィルタの通過周波数帯域 : 500Hz ~ 9.3kHz	34
3.4	振幅特性の補正について	34
3.5	考察	40
4	予備実験 1	42
4.1	刺激音作成	42
4.2	実験条件	43
4.3	実験結果	43
4.4	考察	43
5	予備実験 2	46
5.1	刺激音作成	46
5.2	実験条件	48
5.3	実験結果	48
5.4	考察	48
6	本実験	52
6.1	刺激音作成	52
6.2	実験条件	54
6.3	実験結果	55
6.4	考察	57

7 おわりに	59
7.1 まとめ	59
7.2 今後の課題	59

目 次

2.1 実験系の概要	7
2.2 SR-Lambda Pro. の周波数特性	10
2.3 HDA 200 の周波数特性	11
3.1 位相補正手順	15
3.2 時間領域における TSP (N=16384)	17
3.3 時間領域における ITSP (N=16384)	17
3.4 TSP と ITSP の畳み込みの結果 (N=16384)	18
3.5 TSP と ITSP の畳み込みの結果の振幅成分拡大図 (N=16384)	19
3.6 Lambda Nova 使用時の TSP 応答の一例	21
3.7 Lambda Nova 使用時のインパルス応答: ピーク位置を 0 にシフトして 8 回 同期加算平均	21
3.8 Lambda Nova 使用時の周波数特性: 8 回同期加算平均	22
3.9 Lambda Nova 使用時の群遅延特性: 60Hz ~ 10kHz、8 回同期加算平均	22
3.10 Lambda Nova 使用時の群遅延特性: 100Hz ~ 10kHz、8 回同期加算平均	23
3.11 オーディオアナライザ BK2012 による群遅延測定結果	23
3.12 HDA200 使用時の TSP 応答の一例	24
3.13 HDA200 使用時の周波数特性: 8 回同期加算平均	25
3.14 HDA200 使用時の群遅延特性: 60Hz ~ 10kHz、8 回同期加算平均	25
3.15 HDA200 使用時の群遅延特性: 100Hz ~ 10kHz、8 回同期加算平均	26
3.16 安定化電源を用いない場合の群遅延特性の標準偏差の時間変化	26
3.17 安定化電源を用いた場合の群遅延特性の標準偏差の時間変化	27
3.18 正弦波の 10 成分調波複合音	29
3.19 カットオフ周波数 60Hz のときの逆フィルタの振幅特性の一例	30

3.20	カットオフ周波数 60Hz のときの原波形のカップラレスポンス	30
3.21	カットオフ周波数 60Hz のときの逆フィルタ処理波形	31
3.22	カットオフ周波数 60Hz のときの逆フィルタ処理波形のカップラレスポンス	31
3.23	カットオフ周波数 100Hz のときの原波形	32
3.24	カットオフ周波数 100Hz のときの原波形のカップラレスポンス	32
3.25	カットオフ周波数 100Hz のときの逆フィルタ処理波形	33
3.26	カットオフ周波数 100Hz のときの逆フィルタ処理波形のカップラレスポンス	33
3.27	カットオフ周波数 300Hz のときの原波形	34
3.28	カットオフ周波数 300Hz のときの原波形のカップラレスポンス	35
3.29	カットオフ周波数 300Hz のときの逆フィルタ処理波形	35
3.30	カットオフ周波数 300Hz のときの逆フィルタ処理波形のカップラレスポンス	36
3.31	カットオフ周波数 500Hz のときの原波形	36
3.32	カットオフ周波数 500Hz のときの原波形のカップラレスポンス	37
3.33	カットオフ周波数 500Hz のときの逆フィルタ処理波形	37
3.34	カットオフ周波数 500Hz のときの逆フィルタ処理波形のカップラレスポンス	38
3.35	カットオフ周波数 300Hz のときの原波形の周波数特性	39
3.36	カットオフ周波数 300Hz のときの原波形のカップラレスポンスの周波数特性	39
3.37	カットオフ周波数 300Hz のときの逆フィルタ処理波形のカップラレスポンスの周波数特性	40
4.1	10成分調波複合音使用時の類似度	44
5.1	31成分調波複合音使用時の類似度：全被験者	49
5.2	31成分調波複合音使用時の類似度：選抜被験者	50
6.1	位相変化のない原波形から作成した刺激音との類似度 (all:全被験者、selected: 判別率の高い被験者、obedient: 同じ刺激音の組合せの判別率が高い被験者)	55
6.2	8本の高調波の位相を π 変化させた原波形から作成した刺激音との類似度 (all:全被験者、selected: 判別率の高い被験者、obedient: 同じ刺激音の組合せの判別率が高い被験者)	56

表 目 次

5.1	原波形と逆フィルタ処理波形との判別率 [%]	51
6.1	位相を変化させる高調波の次数とその周波数 : () 内の数字はその高調波が 含まれる聴覚フィルタ内の高調波の数	54

第 1 章

はじめに

従来より「耳は位相に聾である」といわれ、位相情報は振幅情報に比べて聴覚的に重要ではないとされてきた。実際音声における音韻の知覚や、楽器の種類の知覚に位相はほとんど影響を与えない。そのため、音声の符号化、認識、合成などの研究はおもに振幅スペクトルの形状との関連でなされ、位相についてはあまり考慮されていなかった。しかしながら、近年位相が音色知覚に影響を及ぼすことが少数ながら報告されており、位相と音色知覚との相関性を明確にすることは聴覚系の時間情報処理、音色の知覚機構を検討する上で重要視されている。また、現在の音声合成系では周波数領域でパルス列を揺らすことにより音色を変化させているが、より自然で高品質な合成音声を得るためには位相による制御が不可欠であり、将来的にはこの分野への応用も期待される。

よって、本研究では、逆フィルタにより実験系の位相歪みを可能な限り取り除いて位相が音色の知覚に及ぼす影響を調べ、系統的な検討を加えて、位相に対する聴覚機構の特性を明らかにする。

1.1 本研究の背景

1.1.1 位相に関する研究

位相の変化の複合音の音色に対する影響を最初に調べたのは Helmholtz である [1]。Helmholtz は、基本周波数が 120Hz と 240Hz の 8 成分調波複合音を用いてそれぞれの成分の位相を変化させ、音色への影響を調べた。その結果、第 6～8 高調波では位相変化

の影響は見られるが、位相の変化による音色の変化は、ある母音が他の母音と知覚されるほど大きくはないと結論付けた。これが後に「耳は位相に聾である」との表現を生んだ。

Fletcher は、音色に主に関係するのは倍音構造¹であり、倍音構造は周波数、強度、位相で特徴付けられるが、強度が高い場合を除いて位相の変化が知覚に与える影響は僅かであるとしている [2]。

Licklider は、16 成分調波複合音を用いた実験 [3] で、高域成分の位相を変化させる方が、低域成分の位相を変化させるよりも音色の変化が知覚しやすいこと、基本周波数は低い方が音色を判別しやすいことを報告している。また、Plomp も 10 成分調波複合音を用いた実験で、基本周波数が 150Hz を越えない場合は音色変化の知覚がしやすいと報告している [4]。本研究では予備実験として Plomp の実験の追試を行なっている。(第 4 章)

その他にも、Schroeder は 31 成分調波複合音を用いた実験で、位相の変化によって波形の尖鋭度が変化することにより、音色が変化すると述べており [5]、また、Patterson は聴覚神経の発火パターンを模擬した位相知覚モデルを提案し、判別能力は基本周波数に依存すると報告している [6]。

位相に関する実験の刺激音として、主に正弦波または余弦波を基本波形とする調波複合音が主に用いられる理由として、音声はピッチの倍音成分を含み、また聴覚系は倍音成分だけからもピッチを知覚できること、計算機上で刺激音の合成がしやすいこと、などが挙げられる。

以下にこれらの報告結果をまとめる。

調波複合音を呈示信号として位相変化の音色に対する影響を調べる場合、

- 基本周波数は低い方がよく、150 ~ 200Hz 以下なら音色の違いを判別しやすい。
- 低域成分よりも高域成分において位相を変化させた方がよい。
- 2kHz 以上の周波数範囲では、音色の判別は困難である。

1.1.2 電気音響機器および群遅延に関する研究

音響刺激を用いる各種の聴覚実験を行なう場合には、信号源(再生系)、増幅器(伝送系)、電気音響変換器が不可欠である。電子技術の進歩に伴い、信号源としての電子計算機シス

¹ 倍音相互間の相対レベル関係

テム、D/Aコンバータあるいはテープレコーダなどの信号源と、フィルタ、アンプ、アッテネータなどの伝送系の周波数特性や歪率といった物理特性は、可聴帯域(20Hz~20kHz)を扱う限りは安定しており、ほとんど問題がないといわれている。しかし、電気音響変換器であるヘッドホンやスピーカの特性についてはまだ多くの問題が残されており、改善がなされつつある。聴覚特性を測定する心理物理実験や音声等の聴取実験を行なう場合にはヘッドホンによって刺激音を呈示することが多く、ヘッドホンの諸特性を知ることは、実験者が意図する波形を被験者に呈示するために有用である。

平原は、広い帯域幅を持つ信号音を用いる心理物理実験や音声の音韻性や音色の判断などの聴取実験に用いられるヘッドホンの諸特性を明らかにするために、5種類のヘッドホンの測定を行なった[7]。その結果どのヘッドホンについても、カップラレスポンス²と実耳レスポンスの差異が認められること、100Hz~5kHzにおける高調波歪み³は基本波に対して-60dB以下であること、200Hz以上では群遅延⁴特性の乱れが約3ms以下に小さくなることを報告している。ただし、カップラレスポンスと実耳レスポンスの差異はヘッドホンごとに異なり、そのヘッドホン間の程度差は400Hz以下の低域で顕著であった。逆に5kHz以上の高域においてはいずれのヘッドホンもカップラレスポンスと実耳レスポンスの差異は大きく、これは、音響的な反射を引き起こすアダプタプレートの影響や耳介の有無などによる音響負荷の違いによるものと考えられる。また、開放型ヘッドホンは密閉型ヘッドホンに比べて、外部輻射特性、音響遮蔽特性、音響クロストーク⁵特性で劣っていると報告している。

次に、実験系の群遅延歪みによる音色への影響について述べる。一般に電気音響伝達系は、線形で時不変(LTI)なサブシステムにより構成されると仮定される。Blauertらは、all-pa特性により生じる群遅延歪みが、通常のイヤホンやラウドスピーカでどの程度の値を持つのか測定し、またそれらの群遅延歪みが判別できるか否かを調べた[9]。この実験では、所望の群遅延曲線をアナログ回路で構築されたall-passフィルタバンクによって近似し、付加的な歪みのない信号とこのフィルタバンクを通した信号とを比較している。この報告によると音声、音楽、雑音、調波、ショートインパルスの5種類の信号のう

² 実耳を近似した音響インピーダンスを持つカップラにヘッドホンを装着して得られる

³ ヘッドホンをカップラに装着し、正弦波を加えた場合に発生する2次~5次高調波歪み成分の総和と基本波成分とのパワー

⁴ 波形の時間遅れを指す。波形がある地点まで達する時間で、各周波数成分の位相回転は周波数に比例して増加している。

⁵ 独立した伝送路間で他の伝送路の信号が、一方の伝送路に混入する現象。

ち、耳がもっとも敏感に知覚したのはショートインパルスであり、300～8kHz に帯域制限された幅 $25\mu\text{s}$ 、50dB SPL の矩形インパルスの 67 % 検知限は、4kHz にピークを持つ群遅延歪みパターンの場合で 0.4ms、7kHz の場合は 1.1ms であった。また、イヤホンの群遅延特性を近似した、楕型の群遅延歪みパターンの場合、群遅延差 (楕構造の山と谷の部分の差) 1.1ms が検知限となった。

一般のイヤホンやラウドスピーカにより付加される群遅延は、500～8kHz の帯域で最大約 0.4ms であり、ほとんどの実際的なケースで、群遅延歪みは補正される必要がないと結論づけている。

河原らは、時間的に非対象な群遅延のばらつきを持つ信号とこれに対して時間軸を反転した信号を組み合わせることで音色の検知限を測定した [10]。非対称性の大きさの制御は、群遅延の標準偏差を制御することにより行なっている。群遅延の標準偏差が適切な場合、2つの音の間には明確な音色の差異が存在し、群遅延の標準偏差が 1.0ms 以上でほぼ 100%、0.5ms でも 75%以上違いがわかると報告している。

これらの報告結果をまとめると以下のようなになる。

- 各ヘッドホンの物理特性には一長一短があり、聴覚実験の目的に沿って最適なヘッドホンを選ぶ必要がある。特にカップラレスポンスと実耳レスポンスとが大きく異なる場合には、注意が必要である。
- 刺激音としては、音声や音楽よりも定常音を用いたほうが、群遅延の影響を知覚しやすい。
- 片耳呈示、両耳呈示の違いによって、判別率に差は生じない。
- 8kHz ぐらいまでの周波数帯域で、標準偏差が 0.5ms 以上であれば位相の変化による音色の違いを知覚できる可能性がある。

1.1.3 従来の研究の問題点

電気音響変換器の位相特性は特に低域で極めて不安定であり、その特性を厳密に規律することは困難である。しかしながら、刺激音の位相を変化させて音色への影響を調べる場合、実験系の位相を補正することは、信頼性の高い実験結果を得るために不可欠である。

Helmholtz や Fletcher らの時代になされた実験は、電気音響機器が導入される以前のものであったり、音場への放射音を用いて行なわれていたりしたために、反射などによって位相の変化がかき消され、音色の差が知覚されない可能性が高い。

また、電気音響機器導入以後の実験も、実験系に生ずる群遅延歪みを無視している場合が多く、実験者によって使用するヘッドホンも異なり、使用したヘッドホンやその諸特性を明示せずに結果が示されている場合もある。このようにして得られた実験結果の頑強性は明らかでない。

そのうえ、位相差による音色の違いを対象とする報告例は少なく、聴覚における時間情報処理や音色の知覚機構の系統的な検討はほとんどなされていないため、位相と音色の厳密な相関が明確ではない。

1.2 本研究の特色

本研究では、まず最適な実験機器を選択し、実験系を構築する。次に、アンプなどの音響変換器による位相の歪みや、ヘッドホン特性、耳（耳介、外耳道）の音響特性などによる微妙な位相変化を逆フィルタで厳密に補正し、被験者が聴く刺激音の信頼性を高めた上で従来のモデルの妥当性を評価する。また、聴覚機構の位相に対する特性を系統的に検討することを試みる。

第 2 章

実験系

群遅延特性は細かいレベルで極めて不安定であり、実験系の歪みを補正するためには使用する電気音響機器の選択にも注意を払わなければならない。図 2.1 に本研究における実験系の概要を示す。

2.1 電気音響機器について

以下に本実験に用いた電気音響機器について説明する。

2.1.1 ヘッドホン

聴覚実験に用いる理想的なヘッドホンとしては、以下のような条件を満たしていることが望ましい。まず、実験者の意図する刺激音を被験者に呈示するための基本的な条件として、

1. 実耳装着状態での周波数特性が平坦であること
2. 歪み率が低いこと
3. 群遅延特性が平坦であること
4. ダイナミックレンジが広いこと
5. クロストークが少ないこと

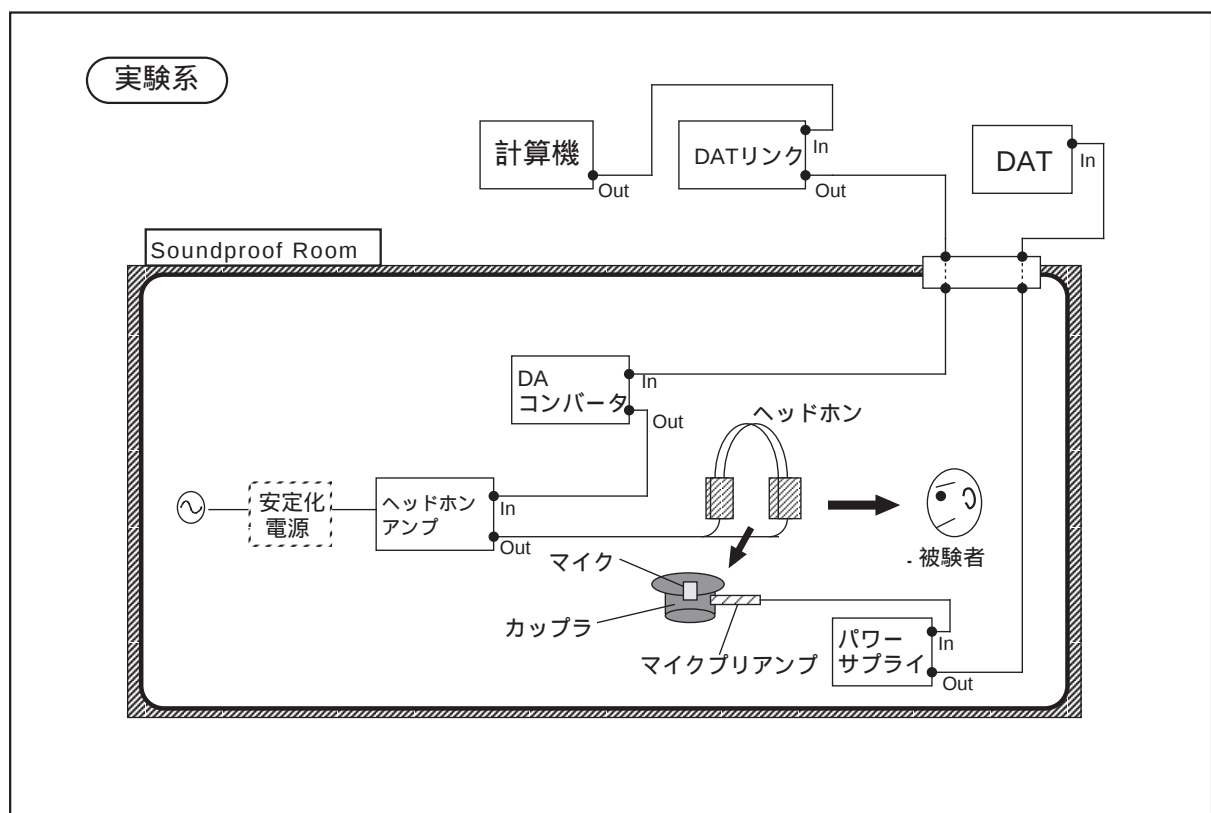


図 2.1: 実験系の概要

などが挙げられる。次に、実験の再現性を保証するための必要な条件としては、

6. 個々のヘッドホンドライバの諸特性の差異が少ないこと
7. 実耳装着状態の変動に伴う諸特性の変化が少ないこと
8. 物理的・電氣的に堅牢で、経年変化が少ないこと
9. 安定に供給されていること

などが挙げられる。また、被験者のために考慮すべき条件としては、

10. 装着感が良いこと

が挙げられる。しかし現実的には、このような条件をすべて満たすヘッドフォンはない。特に、可聴帯域全域に渡って平坦な周波数特性を持つヘッドフォンはなく、一般的にはヘッドホンの周波数特性は帯域通過型である。そこで、本実験では入手可能なヘッドホンの中から、実用帯域における周波数特性の安定性に定評のある Lambda Nova (STAX) と HDA200 (Sennhaeiser) を使用した。

- Lambda Nova (STAX)、SRM-1/MK-2 P.P (STAX)

Lambda Nova はコンデンサ型¹・耳覆い式・開放型のヘッドホンで、元来は音楽観賞用であるが聴覚実験にも用いられている。成極電圧は 580V/DC、耳当ては実耳装着性を向上させるために平板ではない。ヘッドホンアンプとして、音量調整用ボリュームをステップアッテネータに交換した SRM 1/MK-2 P.P を使用する。

図 2. 2 に SR-Lambda Pro. (STAX) のカップラレスポンスと実耳レスポンスを示す (平原, 1997)。実線がカップラレスポンス、太線が実耳レスポンスである。カップラレスポンスは 5 組、左右で 10 個のヘッドホンドライバに対する測定値の平均値と標準偏差であり、この標準偏差はヘッドホンの周波数特性の個体差を表している。一方、実耳レスポンスは被験者 4 名が 1 個のヘッドホンドライバを時間をおいて 2 回装着して測定した 8 個のデータの平均値と標準偏差である。従って、この標準偏差は実耳装着状態の変動に伴う周波数特性の変化を表している。

この図より、周波数特性は広い帯域に渡って平坦であり、200Hz ~ 5kHz におけるカップラレスポンスと実耳レスポンスの差異も小さいことがわかる。これは、耳当

¹ 薄い振動体 (膜) に直流電圧を加えておき、それを挟んだ 2 枚の電極に音楽信号を供給することで起きる静電気力で、振動体 (膜) を駆動する。

てが平板でないために、カップラに装着すると耳当てとアダプタプレート面との間に隙間が生じ、そのときの音響漏洩の状況と実耳装着時にできる隙間による音響漏洩の状況とが類似しているためと考えられる。また、100Hz～5kHzにおけるカップラレスポンスと実耳レスポンスの標準偏差はいずれも2dB以下である。

Lambda NovaはSR-Lambdaシリーズの改良版で、振動膜の加工精度をあげるとともに、電極構造を全面変更しているため、SR-Lambda-Pro以上の再現性が期待できる。また、カップラレスポンスと実耳レスポンスの差異も耳当てが共通であることから、SR-Lambda-Proとほぼ同程度であると考えられる。

- HDA200 (Sennheiser)

動電型・耳覆い式・密閉型のヘッドホンで、元来は高騒音下での通信用であるが、聴覚実験にも用いられている。

図2.3にヘッドホンアンプとして、AU-α907MR(SANSU)を使用した場合のHDA200のカップラレスポンスと実耳レスポンスを示す(平原, 1997)。

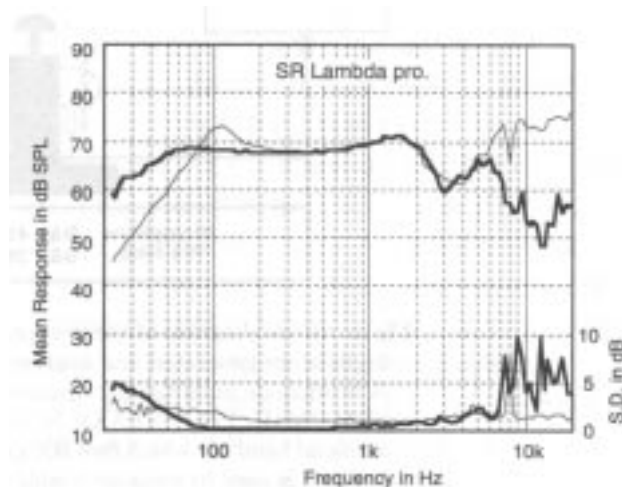


図 2.2: SR-Lambda Pro. の周波数特性

この図より、1.5kHz 以下ではカップラレスポンスと実耳レスポンスの差異が、小さいが、5kHz 以上の実耳レスポンスは、急激に低化していることがわかる。低域の差異が小さいのは、実耳装着時においても耳当てと側頭部との密着度が高く隙間がほとんど生じていないためである。また、100Hz ~ 5kHz におけるカップラレスポンスと実耳レスポンスの標準偏差はいずれも 2dB 以下である。

2.1.2 カップラ

- B&K4153

IEC318に規定される耳覆い型イヤホンの広い周波数範囲の校正用カップラで、使用周波数範囲は 20Hz ~ 10kHz。

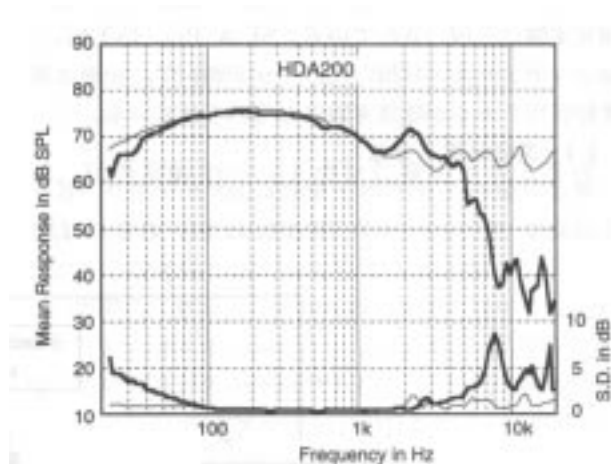


図 2.3: HDA 200 の周波数特性

2.1.3 マイク、マイクロホンプリアンプ、マイクロホンパワーサプライ

- B&K4134

1/2 インチの音圧型コンデンサマイクロホン²で、拡散音場あるいは密閉空洞およびカップラ内の測定用として設計されている。周波数範囲 4Hz ~ 20kHz (± 2 dB)、ダイナミックレンジ 21 ~ 160dB、成極電圧 200V/DC。

- B&K2669B

マイクロホン前置増幅器。1/2 インチコンデンサマイクロホンに使用。

- B&K5935

バッテリー動作 2 チャンネルマイクロホン電源。0 ~ 50dB を 10dB ステップでゲイン選択。Lin./A 荷重。

²振動膜とこれに対抗して配置された固定電極で形成され、音圧に応じて振動膜が振動することによる両者の間の静電容量の変化により電気出力信号を取り出す。構造が簡単で、安定性が高いことから音響計測に適している。

2.1.4 DATリンク

- DAT-Link+ (T.C. T.)

計算機と各種音響機器とを接続するデジタルオーディオインタフェース。デジタルデータの転送のほかに、DATデッキをフロントエンドとして用いることで、アナログデータの計算機への入力も可能である。

2.1.5 DAコンバーター

- DAC- TALENT BD (STAX)

2組のバッテリーを持ち、一方がDAプロセッサに電源を供給している間、もう一方のバッテリーが充電を行なうため、ACとDCを完全に切り離すことができる。

2.1.6 DAT

- TCD- D10PRO2 (SONY)

ポータブルタイプのデジタルオーディオレコーダー。DC駆動のため、他の機器からノイズの影響を受けない。

2.1.7 安定化電源

- HYPERSINEREGULATOR HS R- 510 (SINANO)

音響映像機器専用電源で、外来より混入するノイズや波形歪み、負荷機器から発生する高周波ノイズ等も瞬時波形制御方式の高速波形コントロールにより、常に最適なサイン波を負荷機器に供給する。また、出力高安定化回路により、電源入力に不安定な電圧や周波数が入力されても、出力には安定した電力を供給する。本研究では、予備実験2及び本実験で使用した。

2.2 防音室

本研究では、物理特性の測定および心理物理実験に、暗騒音約23dB(A)の防音室を使用した。なお室内にクリーンなAC電源は供給されておらず、電源電圧を時間をおいて測

定したところ、約 12 時間の間に $105V_{rms} \rightarrow 95V_{rms} \rightarrow 120V_{rms}$ と変動した。ヘッドホンアンプ SRM-1/MK-2 P.P(STAX) の許容電圧範囲が $100V \pm 10\%$ であることから、なんらかの電源対策を講じる必要があるため、安定化電源を使用した。

第 3 章

位相補正

3.1 位相補正手順

図 3.1に示す手順で位相補正を行なう。

3.1.1 インパルス応答

インパルス応答の導出には直接法 (パルス法) や TSP 法 (Time-stretched pulse) が知られている。直接法は、持続時間が短い矩波波などの疑似パルスを入力信号とする方法で、測定結果がそのままインパルス応答となるが、このようなパルスは、波高の割にエネルギーが小さく、S/N の良い測定を行なうためには極めて多くの同期加算が必要となり、実際の測定には適さない。そこで本研究ではインパルス応答の測定に、鈴木らが提案した OATSP (optimized Asynchronous Time-stretched pulse 法) を用いた [11][12]。これは直接法の欠点である入力信号のパワー不足を補うために提案された TSP 法を改良したもので、入力信号として時間引き延ばしパルスを用い、その応答を時間圧縮したものである。これにより、高精度高 S/N での測定が可能となった。

インパルス応答導出のアルゴリズム

以下に、OATSP 法のアルゴリズムを述べる。

OATSP (以下 TSP) は周波数領域で設計され、それを IDFT することで時間関数としての信号となる。TSP は式 (3.1) のように定義される。

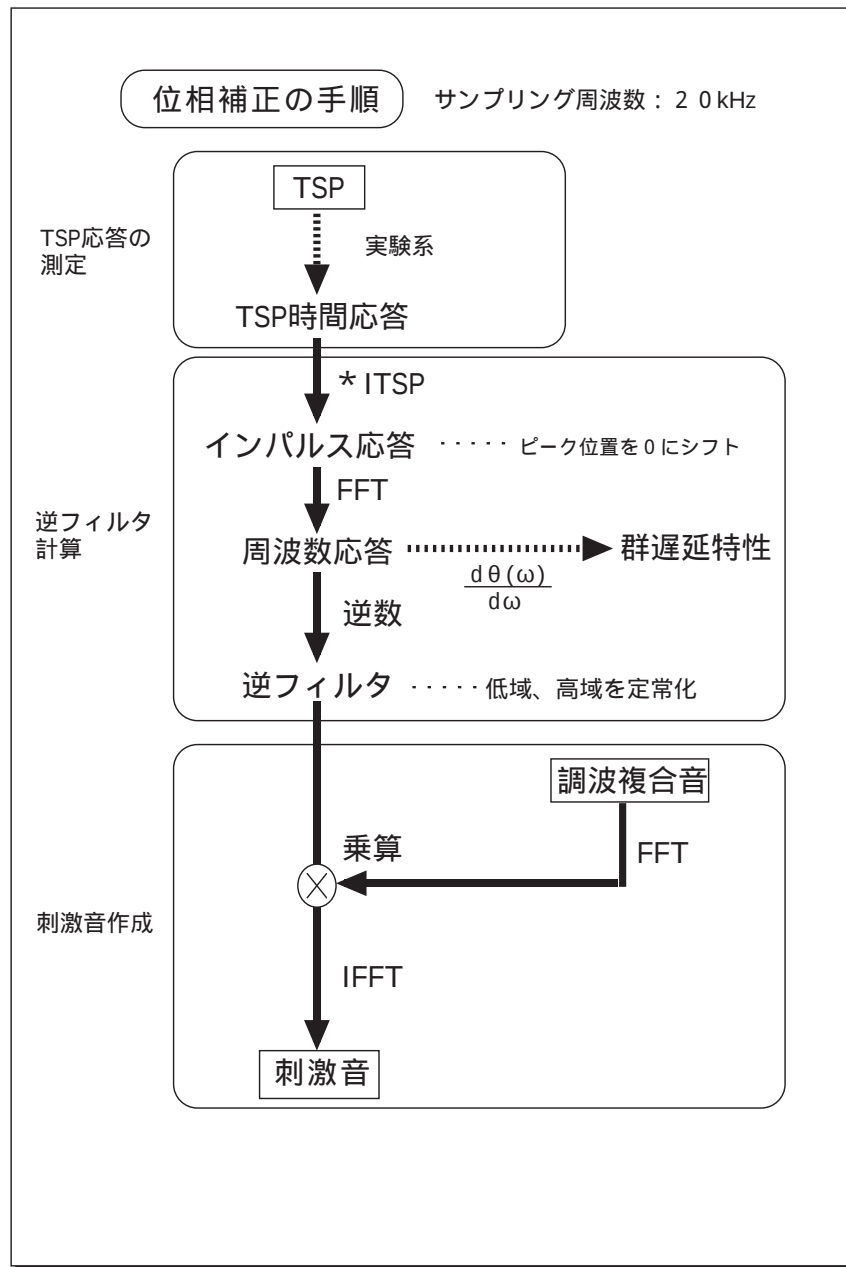


図 3.1: 位相補正手順

$$H(n) = \begin{cases} a_0 \exp(j4m\pi n^2/N^2) & (0 \leq n \leq N/2) \\ H^*(N-n) & (N/2+1 \leq n \leq N) \end{cases} \quad (3.1)$$

* : 複素共役

a_0 : 振幅

N : データ長 ($N = 2^i$)

m : 時間引き伸ばしパラメータ ($m \leq N/2$)

このとき、ITSP は式 (3.2) のようになる。

$$H^{-1}(n) = \begin{cases} a_0 \exp(-j4m\pi n^2/N^2) & (0 \leq n \leq N/2) \\ H^*(N-n) & (N/2+1 \leq n \leq N) \end{cases} \quad (3.2)$$

TSP の応答と ITSP を畳み込むことにより、時間引き延ばしが逆に圧縮されて、被測定系のインパルス応答が求められる。また、周波数応答はインパルス応答を DFT することとで求まる。

TSP 作成および評価

サンプリング周波数を 20kHz とし、TSP の各パラメータを $N = 16384$ $m = 4800$ $a_0 = 32760$ とする。このとき TSP の信号長は 682msec である。

このときの時間領域の TSP、ITSP を図 3.2 図 3.3 に示す。また、これらを畳み込んだものを図 3.4 振幅成分を拡大したものを図 3.5 に示す。ピークを除く振幅成分の最大ぶれ幅は、ピーク値の約 10^{-4} となり、TSP と ITSP を畳み込んだものが、ほぼきれいなインパルスとなることが分かった。

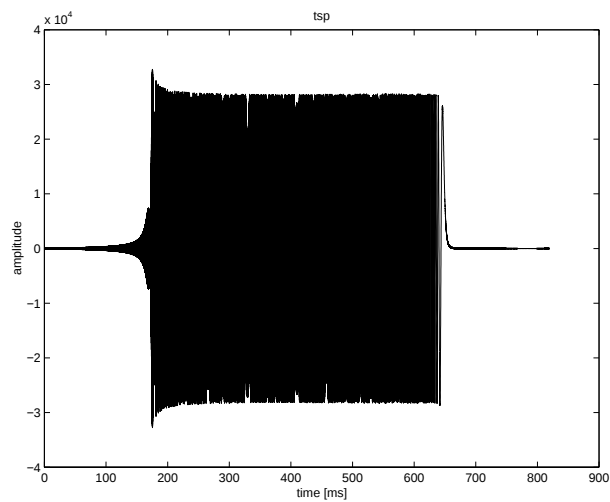


図 3.2: 時間領域における TSP (N=16384)

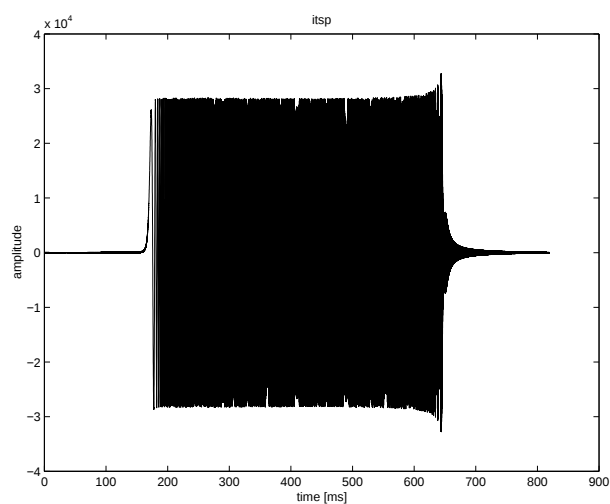


図 3.3: 時間領域における ITSP (N=16384)

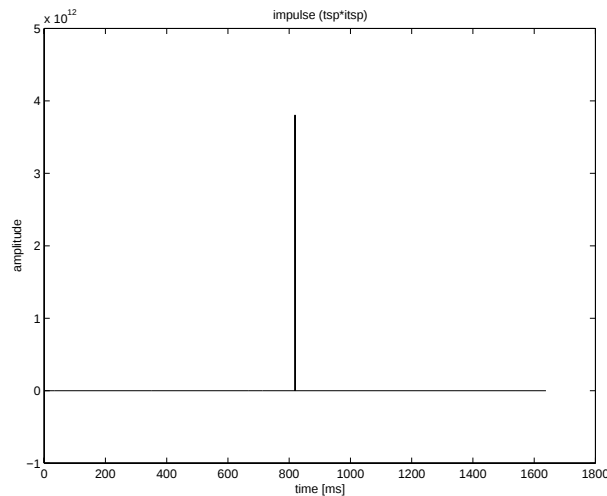


図 3.4: TSP と ITSP の畳み込みの結果 (N=16384)

インパルス応答の被測定系は以下の通りである。なお、ボード体は、アナログ部である。

Work Station → DATLINK → DA converter → Headphone Amp. →
 Headphone → Artificial Ear + Microphone → Microphone Preamp. →
 Microphone Power Supply → DAT

この被測定系のインパルス応答は残響が存在する環境とは違い、収束時間が短いので信号長を長くする必要はないが、N=4096の場合、最大ぶれ幅はピーク値の約 5×10^{-2} となり誤差が大きいため、以後、N=16384 の TSP を用いて測定を行なった。

3.1.2 群遅延特性

群遅延特性は位相特性を周波数微分したものである。位相特性でなく群遅延特性に着目した理由として、DFT により求めた位相特性は不連続な関数であるため、そのまま実験系の特性評価等に使用するのは困難であること、群遅延は時間領域での遅延量に相当するため、直観的に理解しやすいこと、評価する際に補間などの操作がしやすいことなどがあ

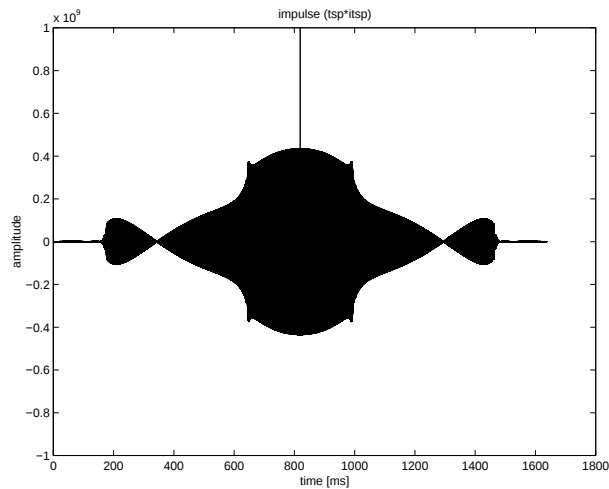


図 3.5: TSP と ITSP の畳み込みの結果の振幅成分拡大図 (N=163 84)

げられる。群遅延のピーク値やピークをとる周波数、標準偏差などが実験系の歪みの特徴をあらわす。

群遅延は位相を周波数微分したものであるが、実装上差分を用いて近似的に導出した。

3.1.3 逆フィルタ

逆フィルタは周波数応答の逆数を計算することで得られる。

逆フィルタの低域および高域部分を、定常補間することで、ハイパスフィルタ、ローパスフィルタと同じ働きをもたせることができ、逆フィルタ処理後の信号を信頼できる周波数範囲内に維持することができる。

意図する刺激音 (原波形) に DFT を施したものとこの逆フィルタとを周波数領域で乗算することで、入力刺激音 (逆フィルタ処理波形) が作成される。

なお、周波数応答は、インパルス応答のピーク位置を 0 に円状シフトしたものを DFT したものである。

3.2 実験系の群遅延特性の測定

実験系の群遅延特性を測定する際に考慮すべき点として、温度・湿度などの環境条件の影響、AC 電源の影響などが挙げられる。また測定に用いる電気音響機器は、研究の目的にかなったものをできる限り選択してあるが、その中でどのサブシステムがもっとも大きな影響を持つのか調べる必要がある。

測定は防音室において、空調、照明を切って行ない、そのときの暗騒音レベルは 23dB(A) であった。また、AC 電源は基準極に留意し、伝送コードはねじれが生じないように、かつ互いの電磁誘導による影響を配慮して、交差したり極端に近傍にくることのないように配置した。全ての電気音響機器は、物理特性を安定させるために 100 時間以上の通電を経て、測定を行なった。

実験系のなかで群遅延特性に対する影響が最も大きいのは、アナログ部でかつ AC 電源で駆動する Headphone Amp. → Headphone の伝送系であると考えられる。ここでは、2. 1. で説明した Lambda Nova (STAX) と HD A200 (Sennhaeiser) の二つのヘッドホンを用いて、群遅延特性の測定を行なった。また安定化電源を使用した場合と使用しない場合の群遅延特性を比較し、安定化電源の有効性について検討した。

3.2.1 Lambda Nova 使用時の群遅延特性

ヘッドホンに Lambda Nova (STAX)、ヘッドホンアンプに SRM-1 / MK- 2 P. P (STAX) を使用したときの TSP 応答の一例を図 3. 6 に、8 個の TSP 応答から同期加算平均により求めたインパルス応答、周波数特性を、図 3. 7 図 3. 8 に、群遅延特性を図 3. 9 図 3. 10 に示す。なお、音圧レベルは実験で被験者が実際に聞くことを想定して、アンプのボリュームを -24dB に固定し、安定化電源は使用していない。また、オーディオアナライザ (B&K2012) による群遅延特性の測定結果 (8 回同期加算平均) も図 3. 11 に示す。

これらより周波数特性、群遅延特性ともに 100Hz 以下の帯域では、乱れが大きいことがわかる。このヘッドホンの群遅延特性には 7. 8kHz 付近に特徴的なピークが見られる。

オーディオアナライザにより測定された群遅延特性と比較すると、ピーク位置や形状が良く一致しており、またより細かいピークも拾うことが可能で、本手法による群遅延測定が有効であることがわかった。

次に、群遅延特性の時変性を調べたところ、細かいピークの値や位置は測定ごとに異な

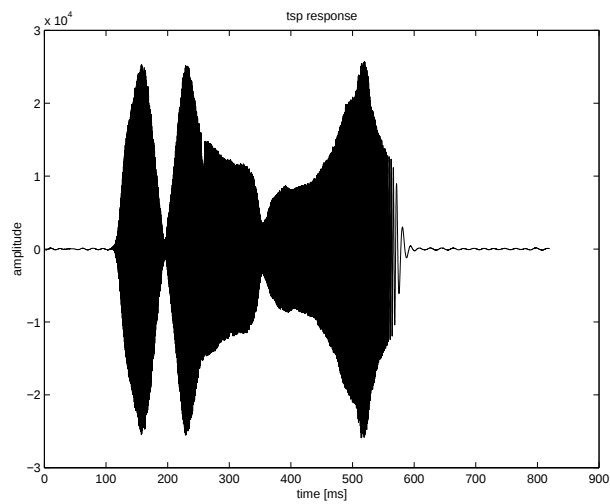


図 3.6: Lambda Nova 使用時の TSP 応答の一例

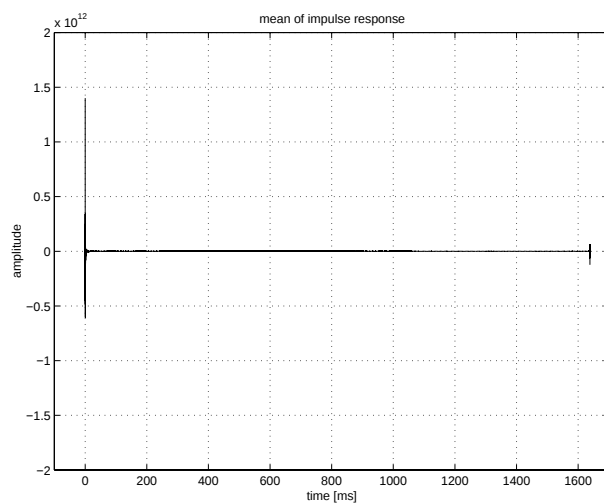


図 3.7: Lambda Nova 使用時のインパルス応答: ピーク位置を 0 にシフトして 8 回同期加算平均

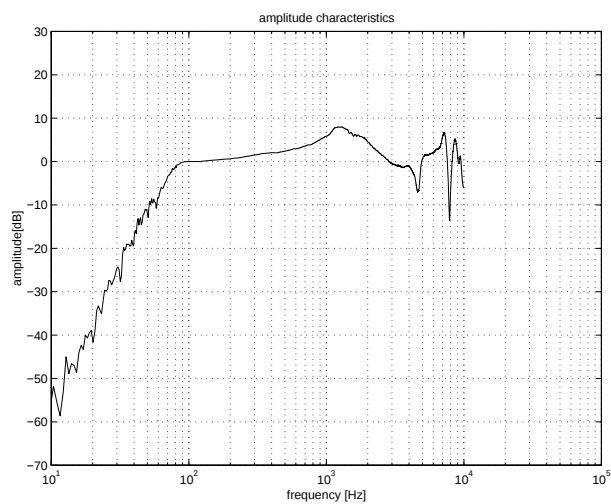


図 3.8: Lambda Nova 使用時の周波数特性: 8 回同期加算平均

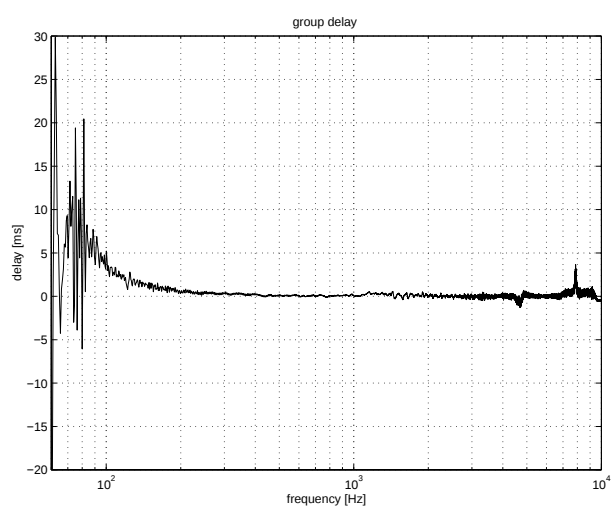


図 3.9: Lambda Nova 使用時の群遅延特性: 60Hz ~ 10kHz、8 回同期加算平均

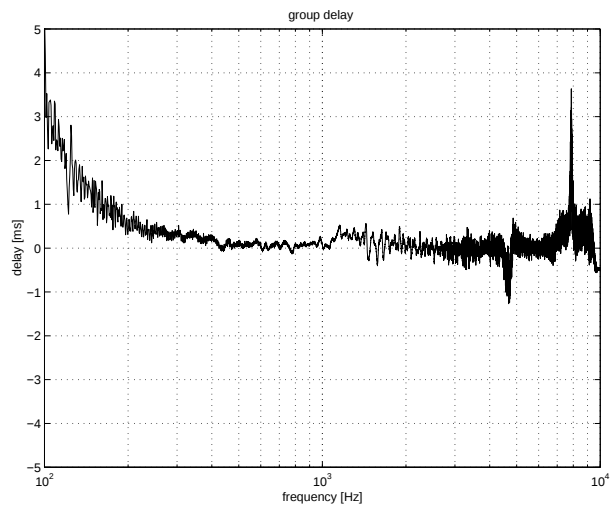


図 3.10: Lambda Nova 使用時の群遅延特性: 100Hz ~ 10kHz、8 回同期加算平均

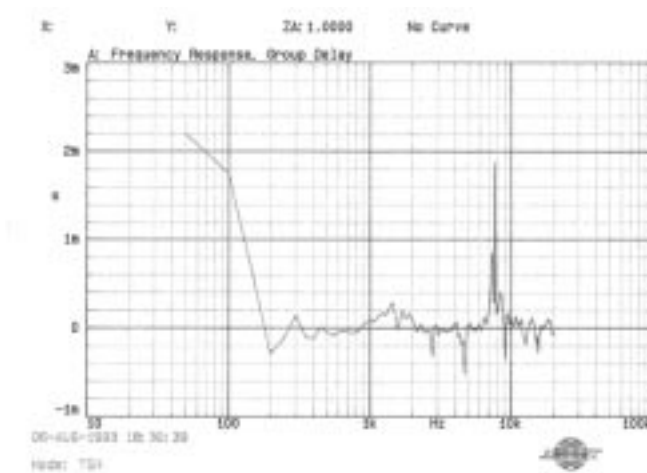


図 3.11: オーディオアナライザ BK2012 による群遅延特性：32 秒のスweep音による 8 回同期加算平均

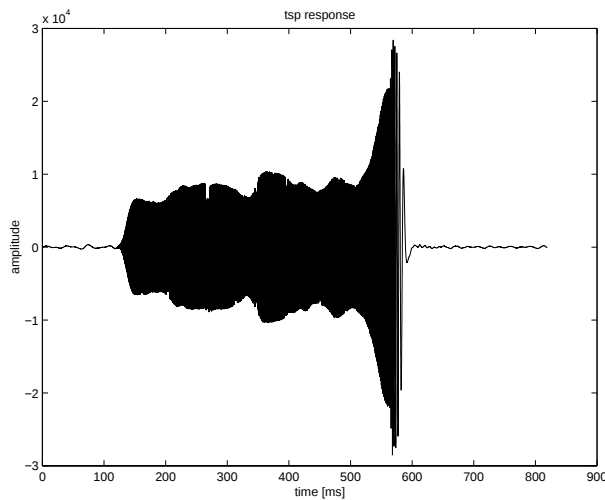


図 3.12: HDA200 使用時の TSP 応答の一例

り、特に低域では安定しなかった。

3.2.2 HDA200 使用時の群遅延特性

ヘッドホンに HDA200 (Sennheiser) ヘッドホンアンプに AU-α90 7MR (SANSUI) を使用したときの TSP 応答の一例を図 3.12 に、8 個の TSP 応答から同期加算平均によって求めた周波数特性を図 3.13 に、群遅延特性を図 3.14 図 3.15 に示す。なお、アンプのボリュームは -60dB に固定した。この測定にも安定化電源は使用していない。

これらより、HDA200 を使用した場合、 1000Hz 以下の低域では Lambda Nova を使用した場合よりも群遅延の乱れが大きく、それ以上の高域では Lambda Nova より乱れが小さいことがわかった。

3.2.3 安定化電源の有効性

安定化電源の有効性を調べるために、Lambda Nova を用いて群遅延特性の時間変化を測定した。なお測定時の環境条件は、室温 26°C 、湿度 57% であった。

図 3.16 に安定化電源を用いない場合の群遅延特性の時間変化を示す。

次に安定化電源を使用して測定した群遅延特性の標準偏差の時間変化を図 3.17 に示す。

安定化電源を使用しない場合、 $100\text{Hz} \sim 9.3\text{kHz}$ の周波数帯域における群遅延特性の標

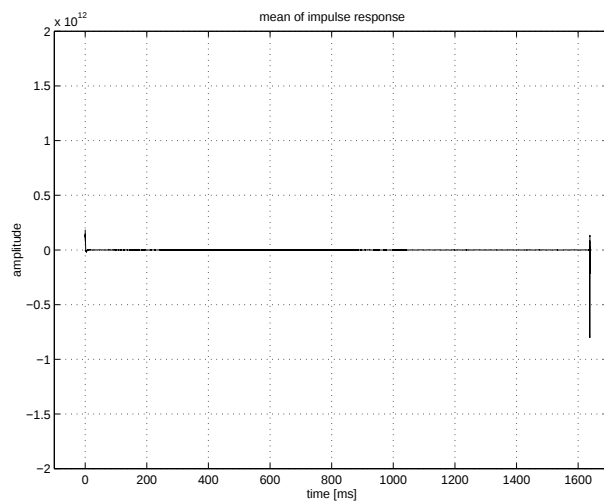


図 3.13: HDA200 使用時の周波数特性: 8 回同期加算平均

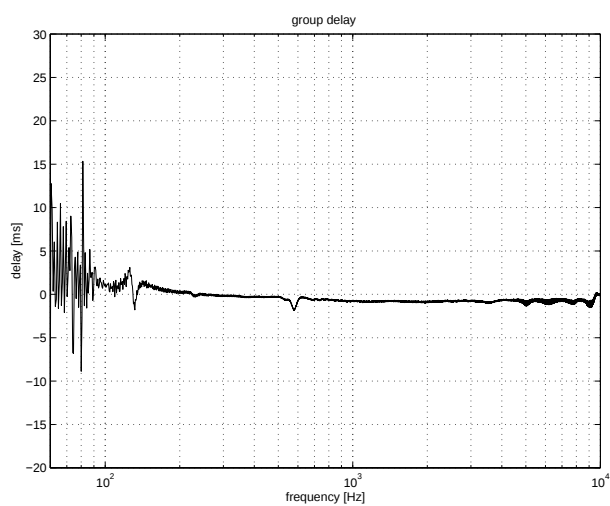


図 3.14: HD A200 使用時の群遅延特性: 60Hz ~ 10kHz、8 回同期加算平均

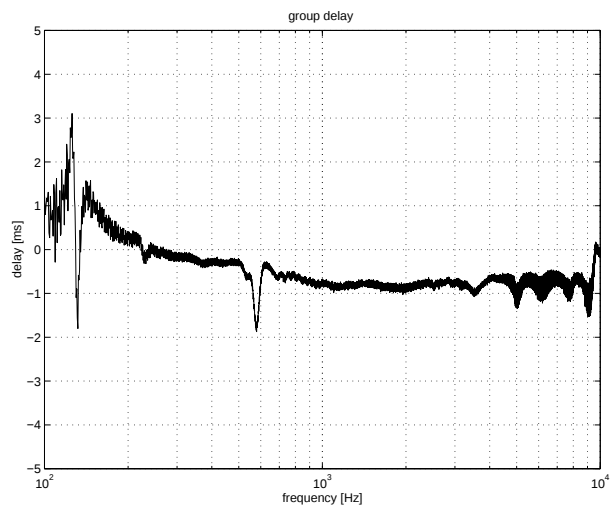


図 3.15: HDA200 使用時の群遅延特性: 10Hz ~ 10kHz、8 回同期加算平均

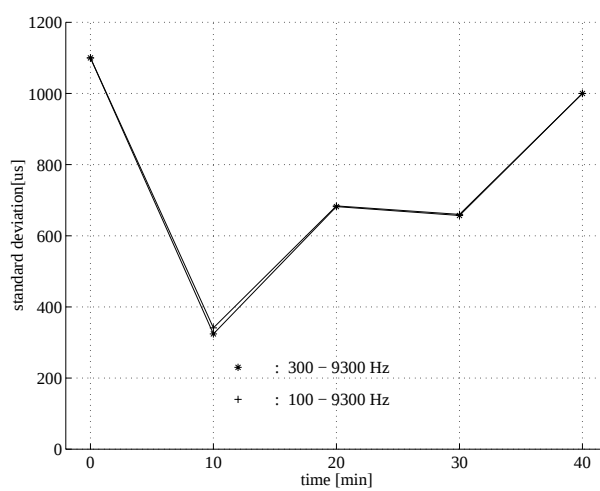


図 3.16: 安定化電源を用いない場合の群遅延特性の標準偏差の時間変化

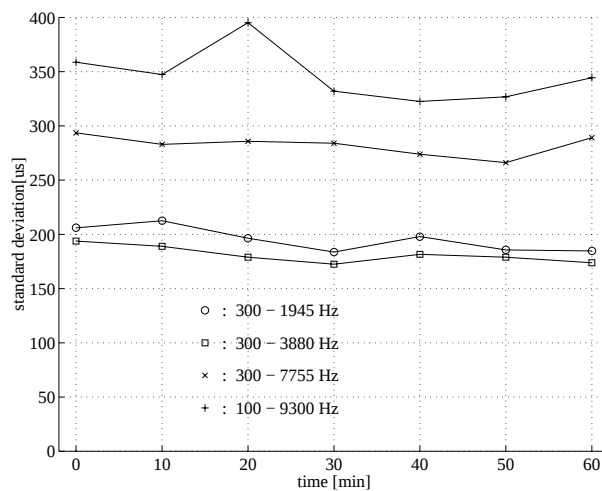


図 3.17: 安定化電源を用いた場合の群遅延特性の標準偏差の時間変化

標準偏差の平均は $757\mu s$ 、安定化電源を使用した場合の平均は $347\mu s$ であり、安定化電源を使用すると群遅延歪みはかなり低減する。また群遅延特性の時間変動も少なく、格段に特性が安定することが分かる。

これらより、安定化電源導入の意義は大きいことがわかった。

3.2.4 考察

群遅延特性は秒単位で変化してしまうほど不安定で、AC 電源の影響に比べ、温度・湿度による影響ははるかに小さかった。

ヘッドホンに Lambda Nova を使用した場合の $100\text{Hz} \sim 9.3\text{kHz}$ の帯域における標準偏差の平均は、約 $400\mu s$ であった。また TDA200 を使用した場合の群遅延の $100\text{Hz} \sim 9.3\text{kHz}$ の帯域における標準偏差の平均は、約 $300\mu s$ であった。両者に際だった性能差は見られないが、開放型の方が密閉型よりも被験者に対する負担が少ないこと、位相変化の影響が大きいと思われる低域での群遅延の標準偏差が小さいこと ($300\text{Hz} \sim 1945\text{Hz}$ の S. D は、Lambda Nova: $151\mu s$ 、HD A200: $233\mu s$)、カップラレスポンスと実耳レスポンスの差異が高域まで小さいことなどの理由から、以後の測定、実験には Lambda Nova を使用することにする。

過去の研究は、群遅延の可聴帯域の標準偏差が $500\mu s$ 以上であれば、音色の変化を知

覚できる可能性があり、また 7kHz 付近で $1.1ms$ 以上のピークがあれば知覚可能であると報告している。

これらの報告と、この実験系の群遅延特性の測定結果を比較すると、AC 電源から直接 Lambda Nova を使用した場合の群遅延の 100Hz ~ 9.3kHz の標準偏差が約 300 ~ 1000ms、7.8kHz 付近のピーク値が約 1 ~ 4ms の範囲で変動することから、Lambda Nova を使用した場合にも、実験系の位相歪みは音色知覚に影響する可能性が高いと考えられる。

安定化電源を用いた場合、特性の安定性は大幅に改善された。このことから予備実験 2 および本実験では、安定化電源を使用する。

3.3 逆フィルタ設計

実験系において、音色の知覚に影響する程度の位相歪みが確認されたので、その位相歪みを補正する逆フィルタを設計する必要がある。

補正可能帯域を検討するため、逆フィルタの通過周波数帯域を 60Hz ~ 9.3kHz、100Hz ~ 9.3kHz、300Hz ~ 9.3kHz、500Hz ~ 9.3kHz の 4 種類設定して逆フィルタを作成する。

逆フィルタを原波形にかけることで、入力刺激信号である逆フィルタ処理波形を作成する。なお原波形とは、被験者の耳元で再現させたい信号のことである。

逆フィルタ処理波形の、カップラレスポンスと原波形を比較することで、逆フィルタを評価する。

式 (3.3) であらわされる正弦波の 10 成分調波複合音の調波成分を、逆フィルタの通過帯域に応じて取り除いたものを、逆フィルタを評価するための原波形として用いた。図 3.18 に、基本周波数が 62.5Hz のときの 10 成分調波複合音を示す。

$$\sin 2\pi f_0 t + \frac{1}{2} \sin 2\pi 2f_0 t + \cdots + \frac{1}{10} \sin 2\pi 10f_0 t \quad (3.3)$$

f_0 : 基本周波数

サンプリング周波数が等しい場合、MATLAB¹ の FFT 演算処理上、計算に使う原波形のデータ長は、逆フィルタのデータ長、すなわち TSP の 2 倍のデータ長と等しくとると、逆フィルタの計算誤差が小さくなる。

¹ 行列演算ツールで高度な数値計算機能と多彩なグラフィック機能を備えている。

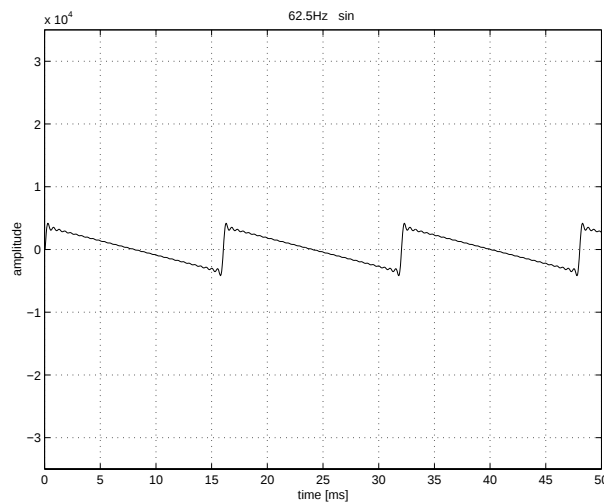


図 3.18: 正弦波の 10 成分調波複合音

3.3.1 逆フィルタの通過周波数帯域：60Hz～9.3kHz

通過周波数帯域が 60Hz～9.3kHz のときの逆フィルタの振幅特性の一例を図 3.19 に示す。

図 3.18 の 10 成分調波複合音を原波形として入力したときに、カップラに装着したマイクロホンで收音されるカップラレスポンスを図 3.20 に示す。また計算機上で原波形に 60Hz～9.3kHz に帯域制限した逆フィルタをかけた逆フィルタ処理波形を図 3.21 に、その波形を入力信号としたときに、カップラに装着したマイクロホンで收音されるカップラレスポンスを図 3.22 に示す。

3.3.2 逆フィルタの通過周波数帯域：100Hz～9.3kHz

逆フィルタ評価には式 (3.3) の 100Hz 以下の調波成分を取り除いたものを原波形として用いた。この原波形を図 3.23 に、原波形のカップラレスポンスを図 3.24 に示す。また逆フィルタ処理波形を図 3.25 に、逆フィルタ処理波形のカップラレスポンスを図 3.26 に示す。

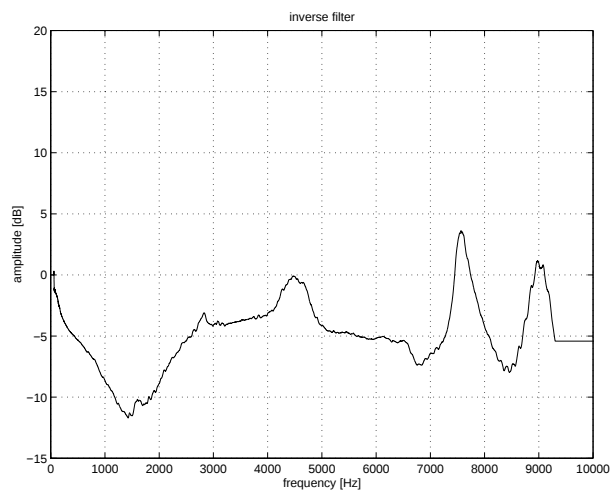


図 3.19: カットオフ周波数 60Hz のときの逆フィルタの振幅特性の一例

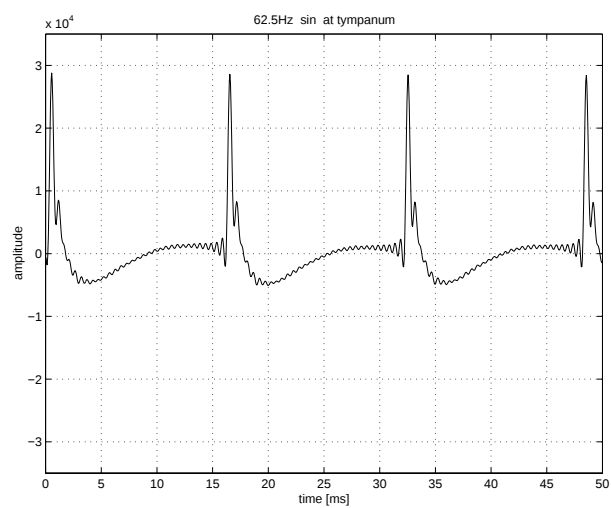


図 3.20: カットオフ周波数 60Hz のときの原波形のカップラレスポンス

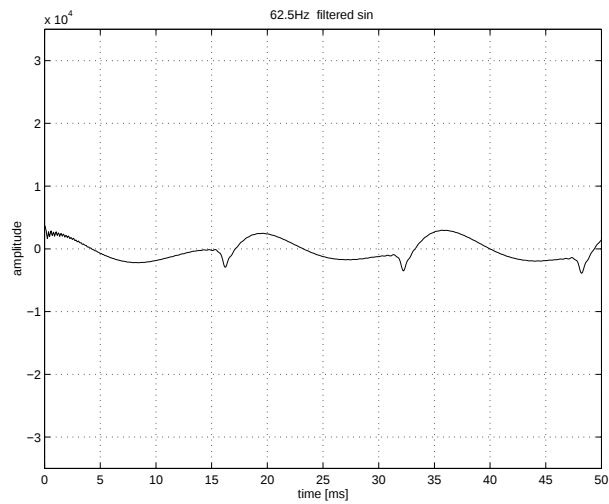


図 3.21: カットオフ周波数 60Hz のときの逆フィルタ処理波形

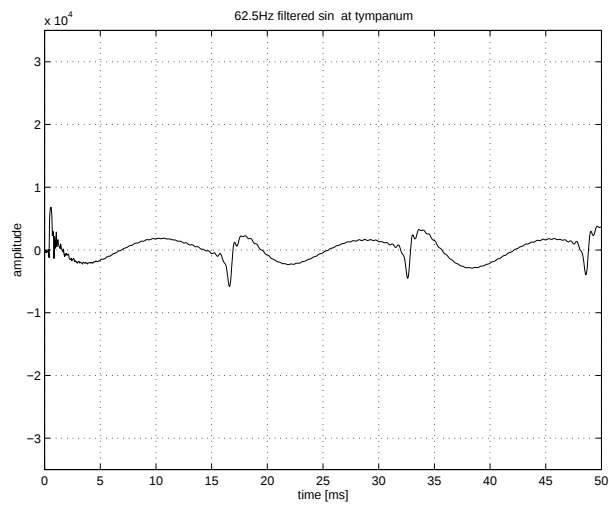


図 3.22: カットオフ周波数 60Hz のときの逆フィルタ処理波形のカップラレスポンス

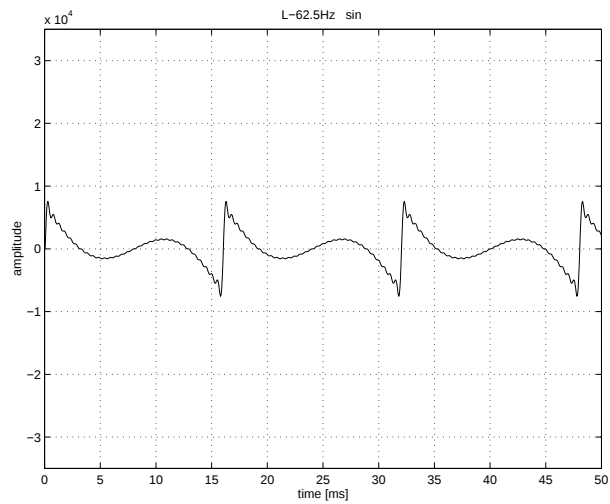


図 3.23: カットオフ周波数 100Hz のときの原波形

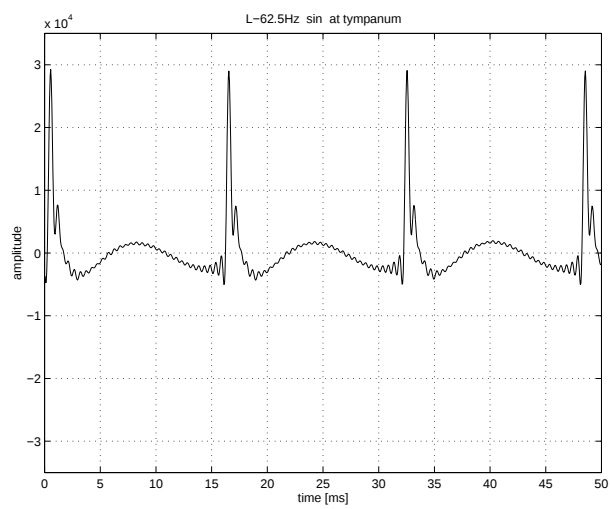


図 3.24: カットオフ周波数 100Hz のときの原波形のカップラレスポンス

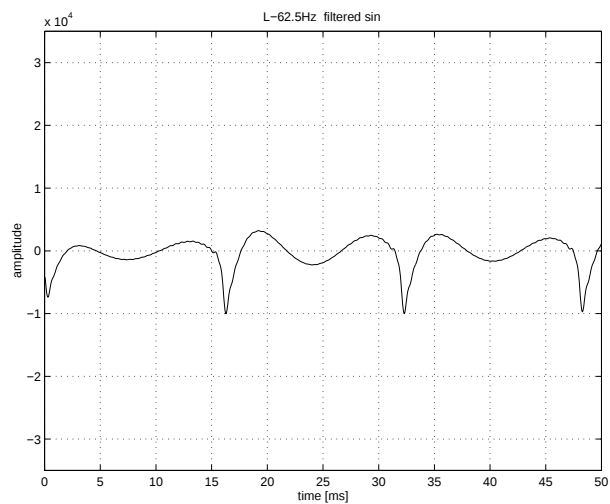


図 3.25: カットオフ周波数 100Hz のときの逆フィルタ処理波形

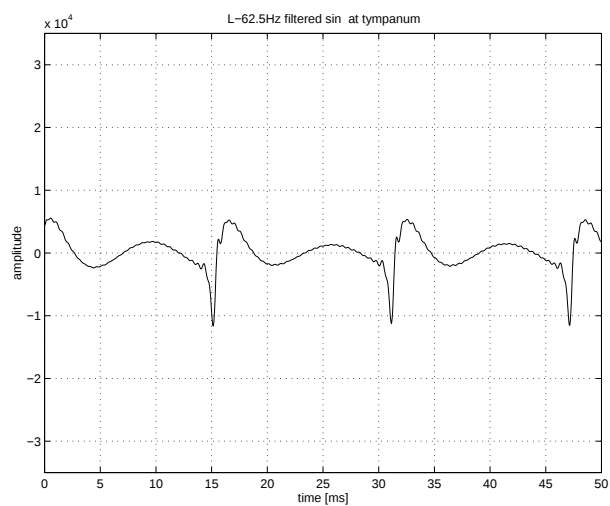


図 3.26: カットオフ周波数 100Hz のときの逆フィルタ処理波形のカップラレスポンス

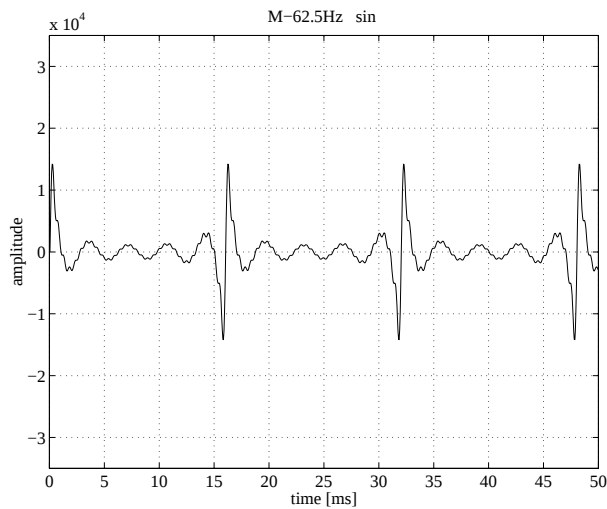


図 3.27: カットオフ周波数 300Hz のときの原波形

3.3.3 逆フィルタの通過周波数帯域：300Hz～9.3kHz

逆フィルタ評価には式 (3.3) の 300Hz 以下の調波成分を取り除いたものを原波形として用いた。この原波形を図 3.27 に、原波形のカップラレスポンスを図 3.28 に示す。また逆フィルタ処理波形を図 3.29 に、逆フィルタ処理波形のカップラレスポンスを図 3.30 に示す。

3.3.4 逆フィルタの通過周波数帯域：500Hz～9.3kHz

逆フィルタ評価には式 (3.3) の 500Hz 以下の調波成分を取り除いたものを原波形として用いた。この原波形を図 3.31 に、原波形のカップラレスポンスを図 3.32 に示す。また逆フィルタ処理波形を図 3.33 に、逆フィルタ処理波形のカップラレスポンスを図 3.34 に示す。

3.4 振幅特性の補正について

通過周波数帯域が 300Hz～9.3kHz の逆フィルタを使用したときの振幅の補正結果について考える。

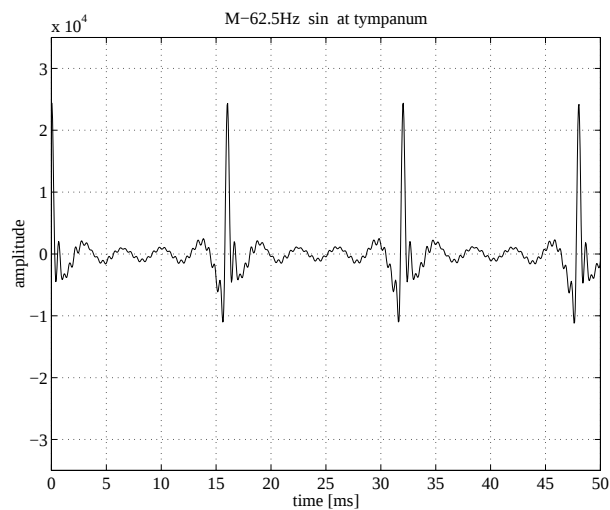


図 3.28: カットオフ周波数 300Hz のときの原波形のカップラレスポンス

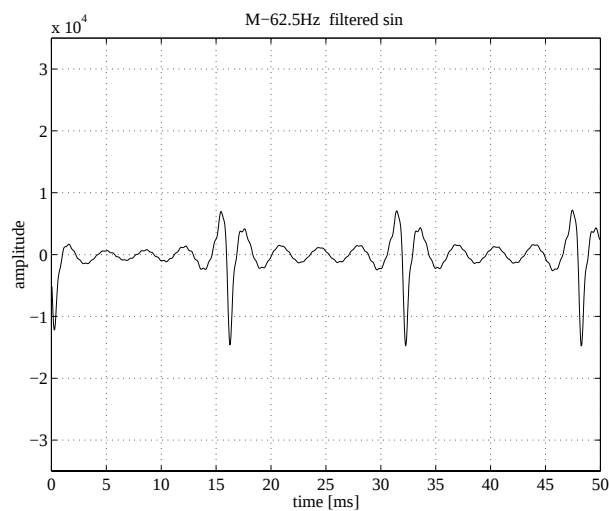


図 3.29: カットオフ周波数 300Hz のときの逆フィルタ処理波形

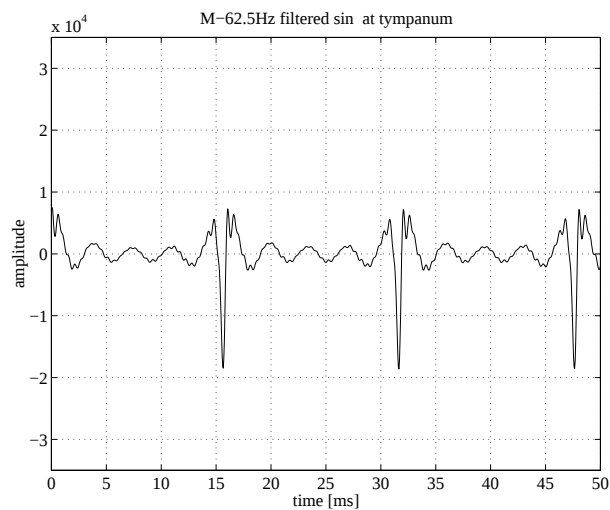


図 3.30: カットオフ周波数 300Hz のときの逆フィルタ処理波形のカップラレスポンス

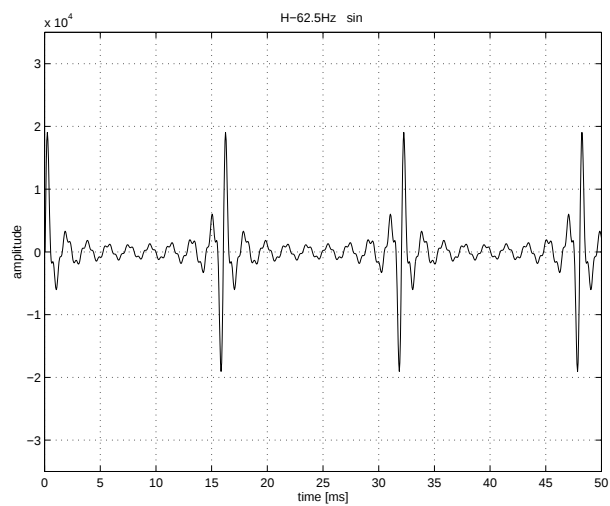


図 3.31: カットオフ周波数 500Hz のときの原波形

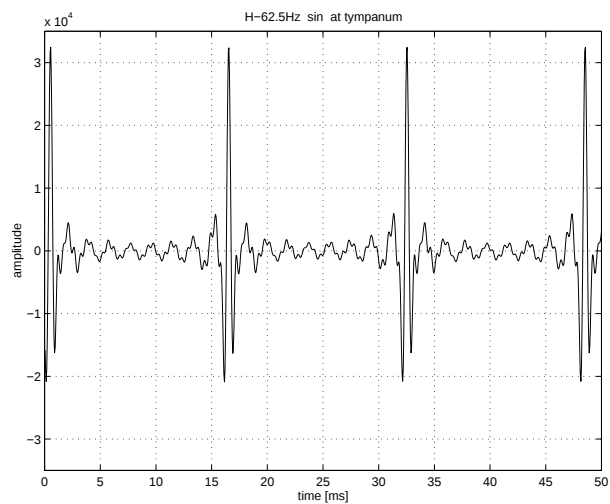


図 3.32: カットオフ周波数 500Hz のときの原波形のカップラレスポンス

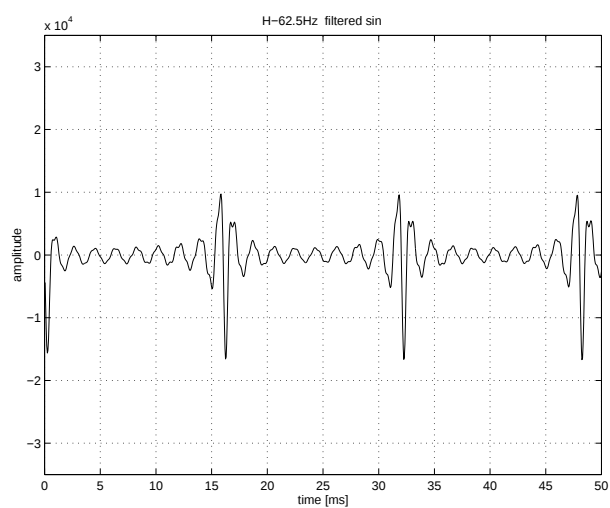


図 3.33: カットオフ周波数 500Hz のときの逆フィルタ処理波形

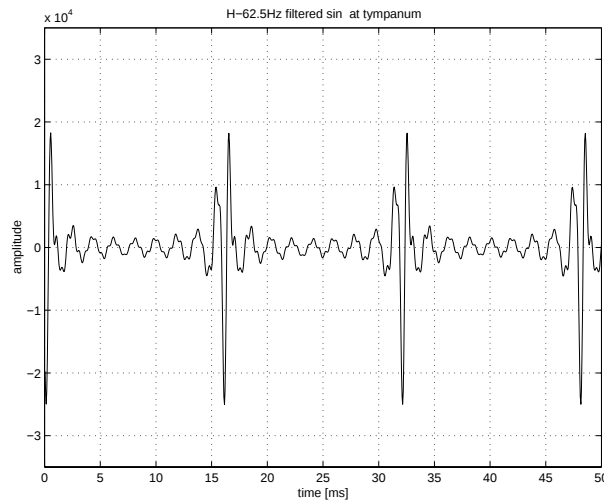


図 3.34: カットオフ周波数 500Hz のときの逆フィルタ処理波形のカップラレスポンス

$$\sin - \sin : \quad \sin 2\pi f_0 t + \frac{1}{2} \sin 2\pi 2 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \quad (3.4)$$

f_0 : 基本周波数

基本周波数が 250Hz のとき、式 (3.4) で示される原波形を DFT したものを図 3.35 に示す。また原波形のカップラレスポンスと逆フィルタ処理波形のカップラレスポンスを DFT したものをそれぞれ図 3.36 に図 3.37 に示す。

式 (3.4) より原波形の調波成分の振幅は、調波の次数と反比例の関係にあり、FFT したときに生じるピークを結んだものは、理論上反比例曲線となる。図 3.35 の原波形の周波数特性をみると、調波成分に対応するピークを結んだものは、反比例曲線となっている。そこで、図 3.36 の原波形のカップラレスポンスの周波数特性について、同様にピークを結ぶと、4.5kHz まで振幅曲線は下降し、また上昇して再び下降する。これは振幅が歪んでいることを意味する。次に、図 3.37 の逆フィルタ処理波形のカップラレスポンスの周波数特性について、同様にピーク値を追っていくとほぼきれいな反比例曲線が得られる。このことから、この実験系によるパワー歪みが、提案した逆フィルタによって補正可能であることが分かる。

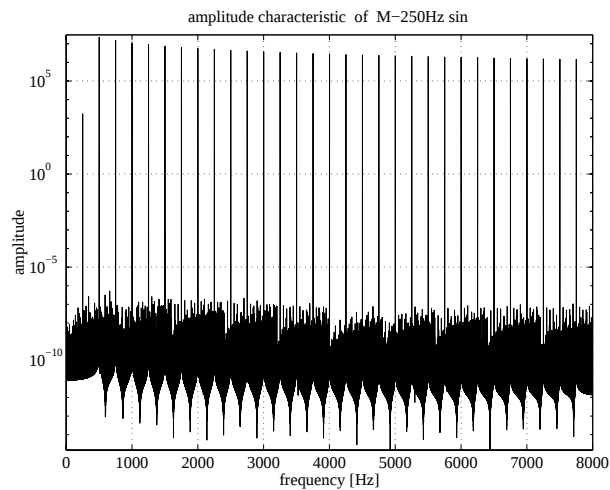


図 3.35: カットオフ周波数 300Hz のときの原波形の周波数特性

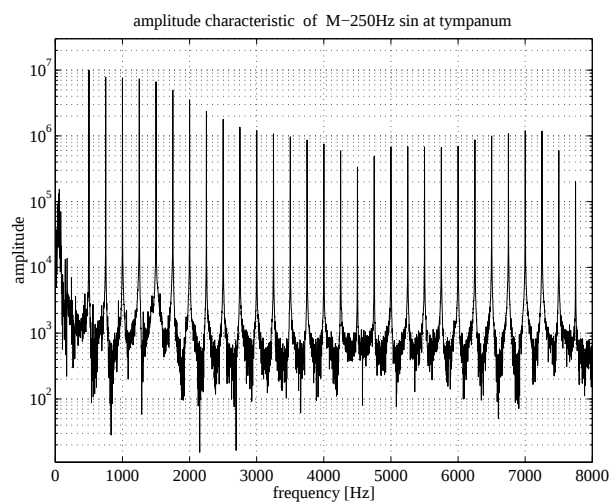


図 3.36: カットオフ周波数 300Hz のときの原波形のカップラレスポンスの周波数特性

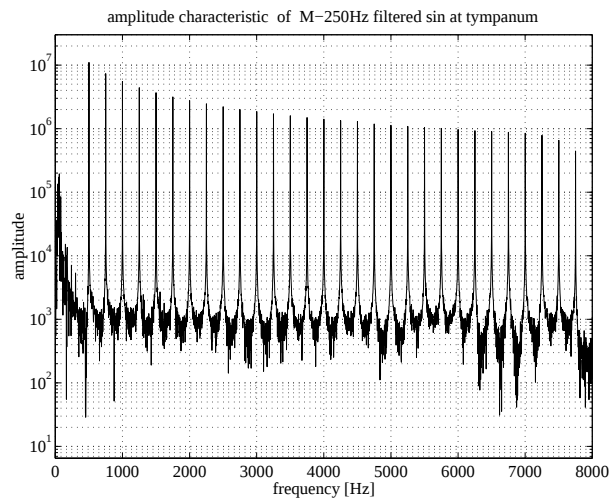


図 3.37: カットオフ周波数 300Hz のときの逆フィルタ処理波形のカップラレスポンスの周波数特性

3.5 考察

群遅延歪みの補正評価を周波数領域で行なうことは、計算誤差がさらにのってしまうために難しく、波形から判断するしかない。

原波形と原波形のカップラレスポンスとを比較すると、通過帯域が 60Hz ~ 9.3kHz、100Hz ~ 9.3kHz の場合、波形の概形や尖鋭度が全く異なり、低域における実験系の歪みの影響の大きさがみてとれる。また、通過帯域が 300Hz ~ 9.3kHz、500Hz ~ 9.3kHz の場合、概形は類似しているもののカップラレスポンスの方が尖鋭度がやや高くなっている。

次に原波形と逆フィルタ処理波形のカップラレスポンスとを比較する。通過帯域が 60Hz ~ 9.3kHz の場合、波形の概形は原波形のカップラレスポンスの場合ほどではないが異なっており、逆フィルタをかけても、歪みが大き過ぎて補正しきれないことを示している。通過帯域が 100Hz ~ 9.3kHz の場合、原波形のカップラレスポンスの場合よりも明らかに概形が類似しており、逆フィルタをかける効果は大きいといえるが、精度が安定しているとはいえない。通過帯域 300Hz ~ 9.3kHz、500Hz ~ 9.3kHz の場合、補正量が少なく、また時間変化も安定してくるため、原波形のカップラレスポンスの場合との概形の類似度に目立つ差は認められないが、尖鋭度が近い点で多少補正されているものと思われる。

また、この実験系でのパワー歪みも逆フィルタで補正できることが確認できた。

その帯域において、群遅延歪みの補正量が少なく時間安定性もあること、ヘッドホン Lambda Nova のカップラレスポンスと実耳レスポンスの差位が小さいこと、原波形作成時の利便性、などの理由から予備実験 2、本実験において、カットオフ周波数 300Hz の逆フィルタを使用する。

第 4 章

予備実験 1

位相変化が音色知覚に実際に影響するという過去の報告結果を確認するために、予備実験として Plomp らが行なった実験 [4] を逆フィルタをかけて追試した。なお、当時の実験状況を考慮して、安定化電源は使用していない。

4.1 刺激音作成

原波形として、式 (4.1) であらわされる 4 種類の 10 成分調波複合音を、基本周波数 62.5Hz、125Hz、250Hz のそれぞれについて 20kHz サンプリングで作成した。

$$\begin{aligned} cc : & \cos 2\pi f_0 t + \frac{1}{2} \cos 2\pi 2f_0 t + \frac{1}{3} \cos 2\pi 3f_0 t + \cdots + \frac{1}{10} \cos 2\pi 10f_0 t \\ ss : & \sin 2\pi f_0 t + \frac{1}{2} \sin 2\pi 2f_0 t + \frac{1}{3} \sin 2\pi 3f_0 t + \cdots + \frac{1}{10} \sin 2\pi 10f_0 t \\ cs : & \cos 2\pi f_0 t + \frac{1}{2} \sin 2\pi 2f_0 t + \frac{1}{3} \cos 2\pi 3f_0 t + \cdots + \frac{1}{10} \sin 2\pi 10f_0 t \\ sc : & \sin 2\pi f_0 t + \frac{1}{2} \cos 2\pi 2f_0 t + \frac{1}{3} \sin 2\pi 3f_0 t + \cdots + \frac{1}{10} \cos 2\pi 10f_0 t \end{aligned} \quad (4.1)$$

f_0 : 基本周波数

この原波形に、通過帯域 60Hz ~ 9.3kHz の逆フィルタをかけたものを入力刺激音とした。

4.2 実験条件

- 逆フィルタはそれぞれの被験者について、実験を行なう直前に更新した。
- 基本周波数 62.5Hz、125Hz、250Hz のそれぞれについて、全 16 ペア¹を 3 回ずつ、計 48 ペアをランダムに片耳呈示し、対比較を行なった。
- 正常な聴力をもつ 10 人の被験者に同じ音に聞こえるか違う音に聞こえるか判断してもらった。
- 一つの刺激音の信号長 1 秒、ペア内の二つの刺激音の間隔 1 秒、ペアとペアの間隔 4 秒とした。
- 刺激音の立ち上がり、立ち下がり部には 5ms のリニアなテーパーをかけた。
- 音圧は被験者に聞きやすいレベルを選択してもらった。

4.3 実験結果

それぞれの組合せに対する類似度²を、図 4. 1に示す。なお、originalは逆フィルタをかけずに行なったときの類似度 (西)[13]、allは被験者 10 人全員による類似度、selected は判別率の高かった上位 3 人による類似度をあらわす。

4.4 考察

被験者の判別率には大きな個人差が見られ、明らかに二極化したため、判別率の高いグループ (selected) の結果を重視する。

ペアの組合せを、同じ刺激音の組合せ A { C-C, S-S, CS-CS, SC-SC } と作成条件が似ている刺激音の組合せ B { C-S, CS-S, C-CS, S-CS } とそれ以外の刺激音の組合せ C { C-CS, C-SC, S-CS, S-SC } のように分類すると selected は A の組合せの場合ほぼ判別が可能であり、B の組み合わせも基本周波数に関わらず 75 %前後の判別率があった。また all でも、C の組合せの場合、基本周波数 62.5Hz、125Hz の場合に約 80 %、250Hz の場合に、約 70 %の判別率が

¹ ペアの組合せ 10 通りで、そのうち異なる刺激音の組合せ 6 通りだけ順序を考慮

² ペアになった刺激音を同じと判断した割合

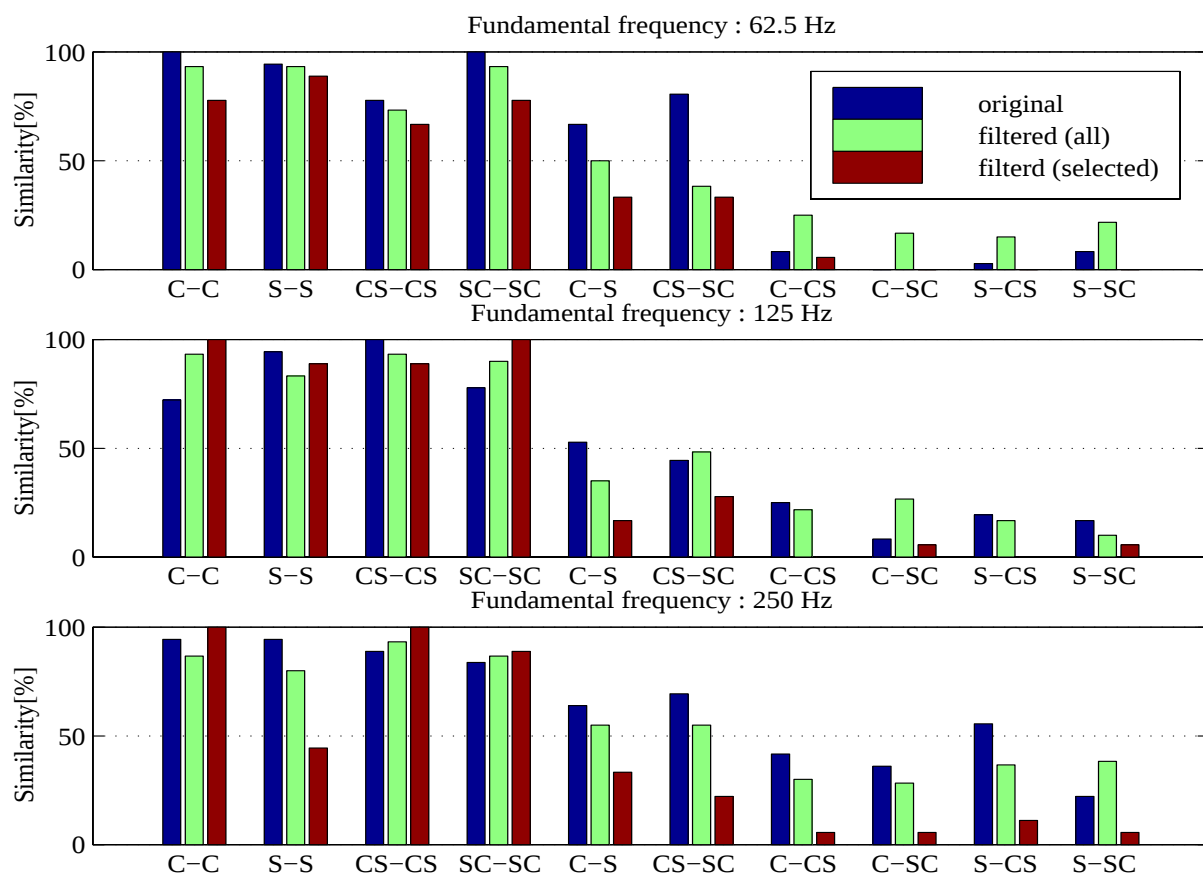


図 4.1: 10成分調波複合音使用時の類似度

あった。Bの組み合わせよりもCの組合せで判別率が高い理由として、sin と cos 両方から構成される音色では高い周波数帯域においてノイズが発生し、どちらか一方だけで構成される音色ではそれが発生しないことが挙げられる。

selected がB、Cの組合わせのほとんどで、ori gi nalの判別率を上回ったことから、逆フィルタ処理を行なった場合にも、位相変化が音色の知覚に影響することが確認できた。また基本周波数が 62.5Hz、125Hz の場合の判別率に明確な差はみられなかったものの、250Hz がもっとも判別が難しいという結果は、”基本周波数が 200Hz 以下なら音色の違いを判別しやすい”という過去の報告に矛盾するものではない。また C-S の組合わせで明らかに順序効果が見られ、呈示順が C- Sのときの類似度は allが 17.5%、selecteが 55.5%、S- Cのときは allが 82.5%、selecteが 55.5%であった。

音質の違いに対する被験者の感想は、大きく聞こえた、高く聞こえた、こもって聞こえたなどのことばにあらわされた。また逆フィルタ処理波形と、原波形を入力信号として呈示したところ、ほとんどの人がその音色の違いを感じることができた。

selected に属する三人の被験者のうち二人は、聴感訓練の経験者とバイオリン演奏などで音楽に携わっていた人で、このことは訓練効果の有効性を示唆している。

第 5 章

予備実験 2

予備実験 1 では、ラフな実験条件下で、原波形に逆フィルタをかけて実験を行ない、結果が過去の報告結果と矛盾しないことを確認した。ここでは、厳密に補正を行なった場合に、位相変化がどの程度音色知覚に影響するのかおよその検討を行なうために、実験を行なう。また実験結果から、本実験でどのような、原波形を使用するのかを決定する。

そこで原波形として第 3 1 高調波までもつ調波複合音を作成し、高い精度で位相補正ができるとおもわれる通過帯域をもつ逆フィルタをかけて入力刺激音を作成した。また、安定化電源を使用することで系の群遅延特性の安定化を図った。これらより実験結果の信頼性は予備実験 1 よりも高いものと思われる。

実験結果は予備実験 1 の結果と比較し、逆フィルタの効果を検討する。

5.1 刺激音作成

式 (5.1)、式 (5.2)、式 (5.3) はそれぞれ基本周波数 62.5Hz, 125Hz, 250Hz の場合の原波形をあらわす。逆フィルタの通過帯域は基本周波数 62.5Hz のとき 300 ~ 1945Hz、125Hz のとき 300 ~ 3880Hz、250Hz のとき 300 ~ 7755Hz であり、通過帯域以下の周波数を持つ低次の調波成分は、取り除いてある。

精度の高い逆フィルタを作成するために、300 ~ 1945Hz、300 ~ 3880Hz、300 ~ 7755Hz の各帯域における群遅延の標準偏差の閾値を、それぞれ 215ms 、 $200\mu\text{s}$ 、 $310\mu\text{s}$ と設定し、各帯域における群遅延が全て閾値以下となるときに周波数応答を用いて、逆フィルタを作成する。

これら 4 種類の調波複合音に、逆フィルタをかけて入力刺激音を作成する。

$$f_0 = 62.5\text{Hz}$$

$$\begin{aligned}
 cc : & \quad \frac{1}{5} \cos 2\pi 5 f_0 t + \frac{1}{6} \cos 2\pi 6 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t \\
 ss : & \quad \frac{1}{5} \sin 2\pi 5 f_0 t + \frac{1}{6} \sin 2\pi 6 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 cs : & \quad \frac{1}{5} \cos 2\pi 5 f_0 t + \frac{1}{6} \sin 2\pi 6 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 sc : & \quad \frac{1}{5} \sin 2\pi 5 f_0 t + \frac{1}{6} \cos 2\pi 6 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t
 \end{aligned} \tag{5.1}$$

$$f_0 = 125\text{Hz}$$

$$\begin{aligned}
 cc : & \quad \frac{1}{3} \cos 2\pi 3 f_0 t + \frac{1}{4} \cos 2\pi 4 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t \\
 ss : & \quad \frac{1}{3} \sin 2\pi 3 f_0 t + \frac{1}{4} \sin 2\pi 4 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 cs : & \quad \frac{1}{3} \cos 2\pi 3 f_0 t + \frac{1}{4} \sin 2\pi 4 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 sc : & \quad \frac{1}{3} \sin 2\pi 3 f_0 t + \frac{1}{4} \cos 2\pi 4 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t
 \end{aligned} \tag{5.2}$$

$$f_0 = 250\text{Hz}$$

$$\begin{aligned}
 cc : & \quad \frac{1}{2} \cos 2\pi 2 f_0 t + \frac{1}{3} \cos 2\pi 3 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t \\
 ss : & \quad \frac{1}{2} \sin 2\pi 2 f_0 t + \frac{1}{3} \sin 2\pi 3 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 cs : & \quad \frac{1}{2} \sin 2\pi 2 f_0 t + \frac{1}{3} \cos 2\pi 3 f_0 t + \cdots + \frac{1}{31} \sin 2\pi 31 f_0 t \\
 sc : & \quad \frac{1}{2} \cos 2\pi 2 f_0 t + \frac{1}{3} \sin 2\pi 3 f_0 t + \cdots + \frac{1}{31} \cos 2\pi 31 f_0 t
 \end{aligned} \tag{5.3}$$

5.2 実験条件

- 逆フィルタはそれぞれの被験者について、実験を行なう直前に更新した。
- 基本周波数 62.5Hz、125Hz、250Hz のそれぞれについて、全 20 ペア¹を 3 回ずつ、計 60 ペアをランダムに片耳呈示し、対比較を行なった。
- 6 人の被験者に、同じ音に聞こえるか違う音に聞こえるか判断してもらった。
- 一つの刺激音の信号長 1 秒、ペア内の二つの刺激音の間隔 1 秒、ペアとペアの間隔 4 秒とした。
- 刺激音の立ち上がり、立ち下がり部には 5ms のリニアなテーパーをかけた。
- 安定化電源を使用した。
- 音圧は固定した。なお、入力刺激音のカップラにおける音圧は、基本周波数 62.5Hz、125Hz、250Hz でそれぞれ $77.2 \pm 0.8 \text{ dB(A)}$ 、 $78 \pm 0.4 \text{ dB(A)}$ 、 $78 \pm 0.4 \text{ dB(A)}$ であった。

5.3 実験結果

被験者 6 人全員 (all) による類似度を図 5.1 に、そのうち判別率の高かった選抜被験者 5 人 (selected) による類似度を図 5.2 に示す。

5.4 考察

今回は被験者として、予備実験 1 において判別率が高かった人を中心に選択したため、selected (5 人) の割合が高く、all (6 人) との類似度にほとんど差はなかった。ここでも selected の結果を重視する。

予備実験 1 と同様にペアの組合せを、同じ刺激音の組合せ A { C-C, S-S, CS-CS, SC-SC } 作成条件が似ている刺激音の組合せ B { C-S, CS-SC } それ以外の刺激音の組合せ C { C-CS, C-SC, S-CS, SC-CS } のように分類すると、図 5.2 より、C の組合わせならば基本周波数 62.5Hz、125Hz の場合に 100 % 判別でき、250Hz の場合でも予備実験 1 と同程度の高い

¹ ペアの組合せ 10 通りで、全ての組み合わせの順序を考慮

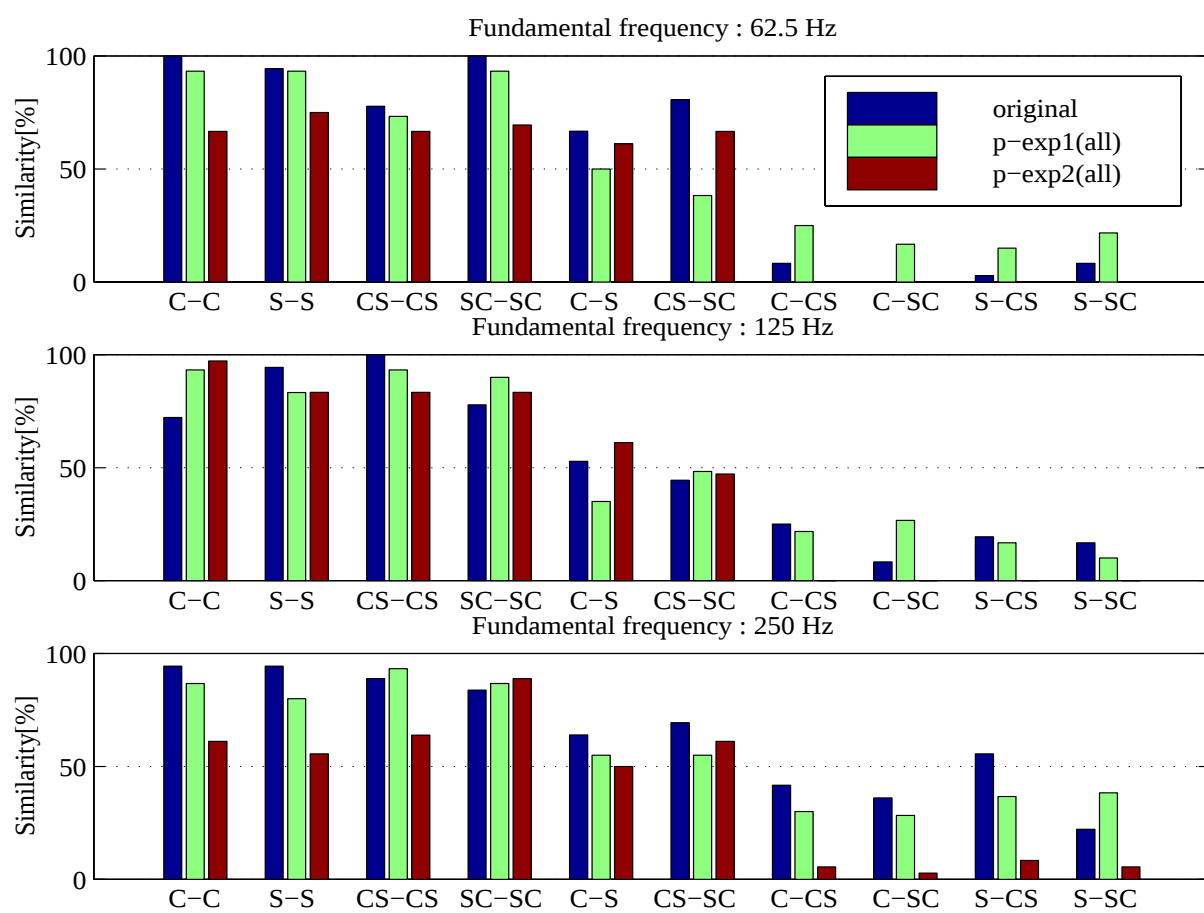


図 5.1: 3 1 成分調波複合音使用時の類似度 : 全被験者

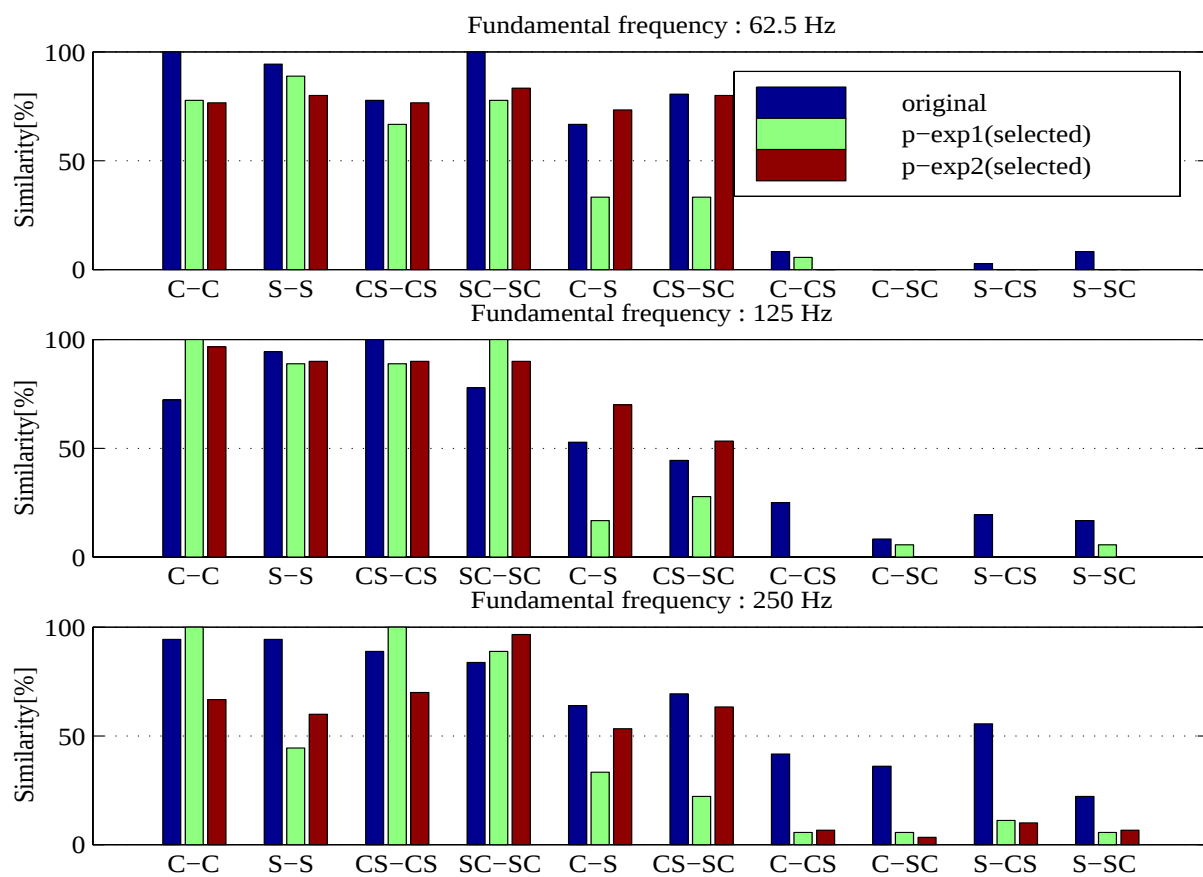


図 5.2: 3 1 成分調波複合音使用時の類似度 : 選抜被験者

表 5.1: 原波形と逆フィルタ処理波形との判別率 [%]

刺激音	c	s	cs	sc
62.5Hz	100	100	75	100
125Hz	100	75	75	50
250Hz	100	100	50	50

判別率が得られ、音色の差が一層明確になったことが分かる。逆に B の組み合わせでは、すべての基本周波数で類似度が 50 % を越えており、予備実験 1 に比べて音色の違いが判別しにくくなっている。逆フィルタによる影響だけならば、B、C の組合せで、ともに判別率が高くなるか低くなるかのいずれかであるはずなので、この判別率の変化は、原波形の高調波の次数が上がったことの影響であると思われる。そこで、予備実験 1 の結果がある程度信頼できるものと仮定すると、B の組合せは位相を変化させてないものと、全ての調波の位相を $\pi/2$ 変化させたものとの組合せ、C の組合せは位相を変化させてないものと、ひとつおきに調波の位相を $\pi/2$ 変化させたものとの組合せ、と見ることで、音色知覚機構が次の傾向をもつ可能性がある。

- a 位相を変化させる調波の密度が小さい場合、低域よりも高域の調波の影響が大きい。
- b 位相を変化させる調波の密度が大きい場合、高域よりも低域の調波の影響が大きい。

この傾向は基本周波数が 250Hz の場合よりも、62.5Hz、125Hz の場合に、やや顕著にみられ、”200Hz 以下の基本周波数の方が判別しやすい” という報告に矛盾しない。

ここで使用した原波形の高調波成分のパワーは、次数と反比例しているため、耳が絶対的にこの帯域のどこに敏感であるかは、一概にいけない。

参考までに、逆フィルタ処理波形と原波形を入力信号として対比較で呈示したときの、各基本周波数における判別率を表 5. 1 に示す。

第 6 章

本実験

予備実験 2 の結果より、知覚のおよその傾向がわかった。本実験ではより細かく位相変化の影響を見ることを目的として刺激音を作成し、その結果に系統的な検討を加える。

そこで細かく音色知覚との相関を調べるために、31 成分調波複合波の一つの高調波だけを π ずらしたものを刺激音として使用した。予備実験 1、2 では、原波形の高調波成分のパワーは、その次数に反比例して減衰していたが、本実験では、高調波成分のパワーを一定にすることで、高域の位相変化の影響をより大きくした。また、位相の変化幅を $\pi/2$ から π にすることで時間領域での波形変化を大きくした。

また傾向 a とほぼ同じことを表している、パルスリボンモデルの妥当性についても検証した。

パルスリボンモデルとは、Patterson が位相変化の知覚モデルとして提案したもので、基底膜を模擬したフィルタバンクにより周波数分析された波形をもとにパルスを生成して、生成されたパルス列の様相の違いにより、知覚特性を説明しようとしたものである。このモデルによると、知覚特性は聴覚フィルタ間の時間差に依存しない。すなわち、聴覚フィルタのバンド幅に含まれる調波の数と、その調波間の位相差だけが問題となる。例えば、ある聴覚フィルタ内に複数の調波が含まれるとき、そのうちの 1 本の位相を変化させると、同じ聴覚フィルタ内の他の調波との位相差により音色の違いは判別しやすくなる。逆に聴覚フィルタ内に 1 本の調波しか含まれない場合は、その位相を変化させても音色の違いは判別しにくい。

6.1 刺激音作成

式 (6.1)、式 (6.2) はそれぞれ基本周波数 125Hz、250Hz の原波形をあらわす。原波形は各周波数ごとに、1 本の高調波成分だけ位相が π ずれたものが 8 個、位相が揃っているものが 1 個、8 本の高調波成分の位相が π ずれたものが 1 個の計 10 個である。

位相変化を行なう高調波の次数と、周波数との対応を、表 6.1 に示す。周波数を対数表示したときに、逆フィルタの通過帯域内でおおよそ等間隔になるような 8 本の高調波を選択して、位相を変化させる高調波を決定した。また位相を変化させる高調波の周波数を中心周波数とする聴覚フィルタのバンド幅を式 (6.3) より計算し、基本周波数からその聴覚フィルタに含まれる調波成分の本数を求めた。

$$f_0 = 125\text{Hz}$$

$$\begin{aligned} & \sin 2\pi 3f_0 t + \cdots + \sin(\pi h f_0 t) + \cdots + \sin 2\pi 31f_0 t \quad (h = 4, 5, 7, 10, 13, 17, 23, 30) \\ & \sin 2\pi 3f_0 t + \sin 2\pi 4f_0 t + \cdots + \sin 2\pi 31f_0 t \quad (6.1) \\ & \sin 2\pi 3f_0 t + \sin(\pi 4f_0 t) + \cdots + \sin(\pi 5f_0 t) + \cdots + \sin(\pi 7f_0 t) + \cdots \\ & \quad + \sin(\pi 10f_0 t) + \cdots + \sin(\pi 13f_0 t) + \cdots + \sin(\pi 17f_0 t) + \cdots \\ & \quad + \sin(\pi 23f_0 t) + \cdots + \sin(\pi 30f_0 t) + \sin 2\pi 31f_0 t \end{aligned}$$

$$f_0 = 250\text{Hz}$$

$$\begin{aligned} & \sin 2\pi 2f_0 t + \cdots + \sin(\pi h f_0 t) + \cdots + \sin 2\pi 31f_0 t \quad (h = 3, 4, 5, 8, 11, 16, 22, 30) \\ & \sin 2\pi 2f_0 t + \sin 2\pi 3f_0 t + \cdots + \sin 2\pi 31f_0 t \quad (6.2) \\ & \sin 2\pi 2f_0 t + \sin(\pi 3f_0 t) + \cdots + \sin(\pi 4f_0 t) + \cdots + \sin(\pi 6f_0 t) + \cdots \\ & \quad + \sin(\pi 8f_0 t) + \cdots + \sin(\pi 11f_0 t) + \cdots + \sin(\pi 16f_0 t) + \cdots \\ & \quad + \sin(\pi 22f_0 t) + \cdots + \sin(\pi 30f_0 t) + \sin 2\pi 31f_0 t \end{aligned}$$

$$ERB[Hz] = 24.7(437f[kHz] + 1) \quad (6.3)$$

この原波形に逆フィルタをかけて、入力刺激音を作成した。逆フィルタの通過帯域は基本周波数 125Hz のとき 300 ~ 3880Hz、250Hz のとき 300 ~ 7755Hz であり、通過帯域にない調波成分は、原波形作成時に取り除いてある。

表 6.1: 位相を変化させる高調波の次数とその周波数 : () 内の数字はその高調波が含まれる聴覚フィルタ内の高調波の数

	刺激音の実験番号	1	2	3	4	5	6	7	8
62.5Hz	調波の次数	6(2)	8(2)	10(2)	12(2)	15(3)	19(3)	24(4)	30(4)
	周波数 [Hz]	375	500	625	750	937.5	1187.5	1500	1875
125Hz	調波の次数	4(1)	5(1)	7(1)	10(2)	13(2)	17(3)	23(3)	30(4)
	周波数 [Hz]	500	625	875	1250	1625	2125	2875	3750
250Hz	調波の次数	3(1)	4(1)	6(1)	8(1)	11(2)	16(2)	22(3)	30(4)
	周波数 [Hz]	750	1000	1500	2000	2750	4000	5500	7500

6.2 実験条件

- 逆フィルタは各周波数ごとに共通のものを使用した。
- 刺激音の立ち上がり、立ち下がり部には 5ms のリニアなテーパーをかけた。
- 基本周波数 125Hz、250Hz のそれぞれについて、1 本の高調波成分の位相を変化させたものと位相変化を行っていないものの組合せ 27 ペアと、1 本の位相を変化させたものと 8 本すべての位相を変化させたものの組合せ 27 ペアの、全 54 ペアを 3 回ずつ計 162 ペアを、ランダムに片耳呈示し対比較を行なった。
- 正常な聴力をもつ 9 人の被験者に判断してもらった。
- 一つの刺激音の信号長 1 秒、ペア内の二つの刺激音の間隔 1 秒、ペアとペアの間隔 4 秒である。
- 安定化電源を使用した。
- 入力刺激音のカップラにおける音圧は、基本周波数 125Hz のとき $81.6\text{dB} \pm 1.4\text{dB(A)}$ 、基本周波数 250Hz のとき $80.5 \pm 1\text{dB(A)}$ であった。

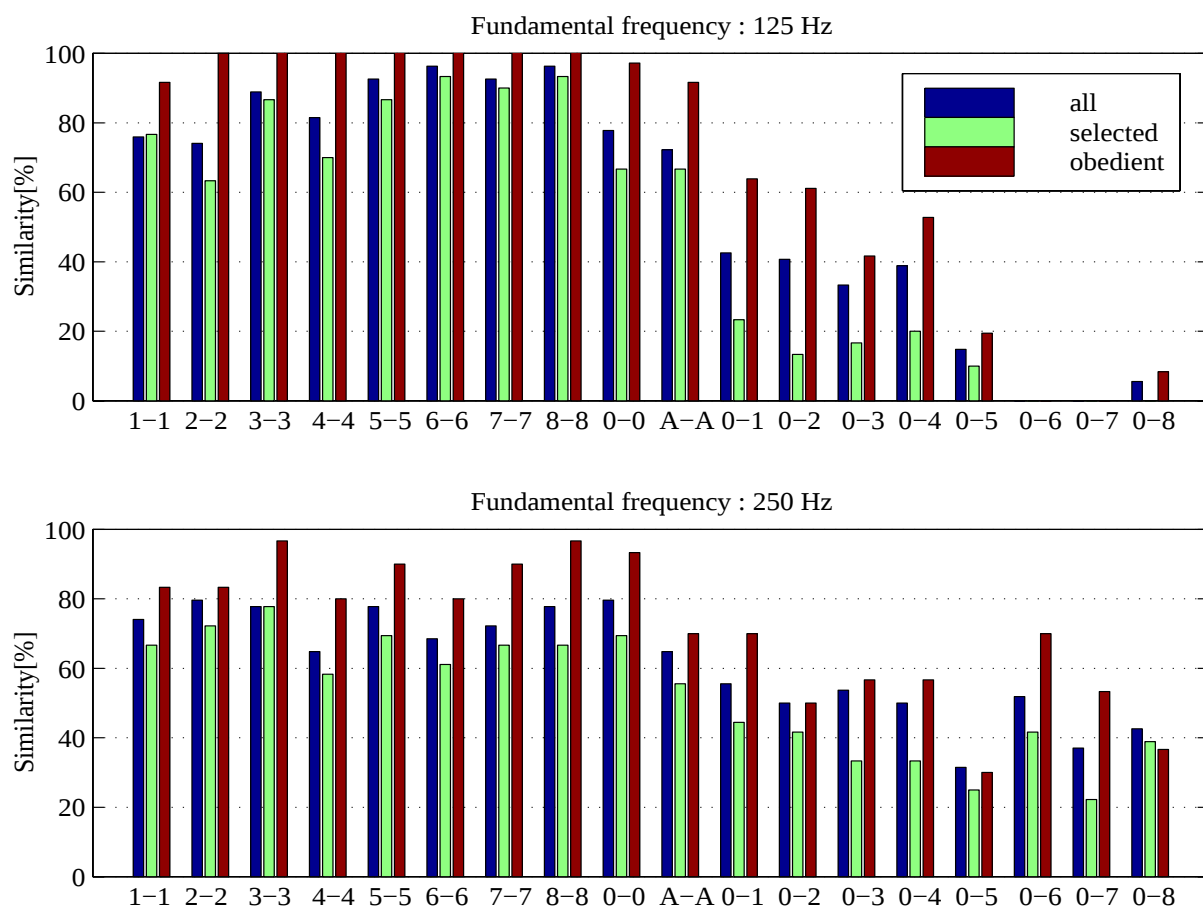


図 6.1: 位相変化のない原波形から作成した刺激音との類似度 (all:全被験者、selected: 判別率の高い被験者、obedient: 同じ刺激音の組合せの判別率が高い被験者)

6.3 実験結果

1 本の高調波の位相を π 変化させた刺激音と位相を変化させていない刺激音との類似度を図 6. 1に、8 本の高調波の位相を変化させた刺激音との類似度を図 6. 2に示す。

グラフの数字は、1～8 が位対応する高調波 (表 6. 1参照) の位相を変化させた刺激音、0 は位相を変化させてない刺激音、A は 8 本すべての位相を変化させた刺激音を表す。

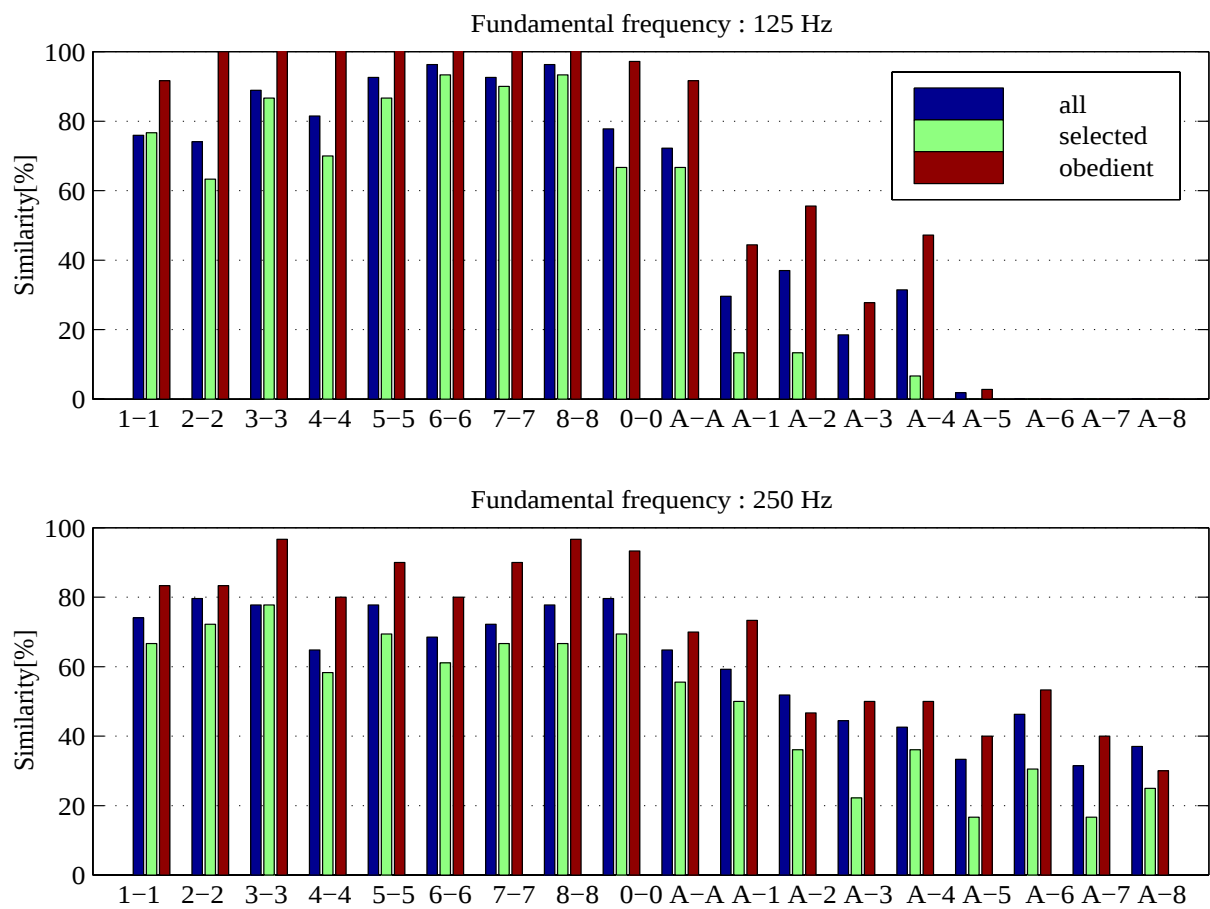


図 6.2: 8 本の高調波の位相を π 変化させた原波形から作成した刺激音との類似度 (all: 全被験者、selected: 判別率の高い被験者、obedient: 同じ刺激音の組合せの判別率が高い被験者)

6.4 考察

all は被験者全員、selected は判別率の高い被験者、obedient は同じ組合せの判別率が高い被験者を表す。

予備実験 2 でみられた以下の傾向をもとに、考察する。

a 位相を変化させる調波の密度が小さい場合、低域よりも高域の調波の影響が大きい。

b 位相を変化させる調波の密度が大きい場合、高域よりも低域の調波の影響が大きい。

基本周波数が 125Hz の場合、高次の高調波の位相を変化させるにしたがって、刺激音 0、および A との判別率は高くなる。これは傾向 a を満たす。特に聴覚フィルタ内に調波が 3 本以上含まれる、刺激音 6、7、8、の場合、判別率はほぼ 100 % に近い。この刺激音を実際に聞くと、音が分極しているように聞こえた。

また刺激音 0、および A との判別率をみると、A との比較の方が判別率が少しずつ高い。これは、傾向 b を満たしている。0 と A の類似度は all で基本周波数 125Hz のとき、67 %、250Hz のとき、63 % であった。

基本周波数 250Hz の場合も、125Hz ほどあきらかではないが、傾向 a、b とともにみられる。このことから、“基本周波数が 200Hz 以下の場合音色の違いが判別しやすい” という報告はほぼ確かめられた。

絶対的な周波数との関係は特に見られないが、“2kHz 以上の高域では音色の違いを知覚することが難しい” という報告は、基本周波数を低くとり、位相を変化させる調波を 1 本に限定すれば、4kHz 以上の高域でもほぼ音色の違いを知覚できるはずであり、完全に誤りである。傾向 a、b は少なくとも基本周波数 250Hz 以下の場合には、基本周波数に関係なくみられる。

傾向 a とパルスリボンモデルは似ているが、異なる聴覚フィルタ間の関連性が傾向 a にはある点で異なり、パルスリボンモデルはその点でやや妥当性を欠く。

基本周波数が 62.5Hz の場合、準備段階で異なる組合せのほとんどが、判別できたため、実験は省略した。この違いは基本周波数 125Hz の場合よりも大きく、よりはっきりと分離して聞こえた。このことは一つの聴覚フィルタ内に複数本の調波が入ることからパルスリボンモデルでも説明できる。

これらより、ある基本周波数以下で、少数の高調波の位相を変化させる場合には、パルスリボンモデルは十分な妥当性をもつが、まばらな多数の高調波の位相を変化させる時

は、やや妥当性を欠くことが分かった。また、調波複合音の位相変化による音色の知覚には、傾向 a、傾向 b がみられ、傾向 a の方が傾向 b よりも強くみられた。

聞き直しが可能な Macintosh の実験処理システムを使用したところ、計算機とのラインの影響のためか、安定化電源を使用しても群遅延の標準偏差が閾値を越えてしまい、使用を断念したことを付け加えておく。

第 7 章

おわりに

7.1 まとめ

本研究では、実験系の位相の歪みを実用帯域で補正する方法を提案した。また、位相歪みの影響は AC 電源、ヘッドホンの影響が大きいことが分かった。

実験系の歪みをできる限り除去したのちに、過去の実験結果の妥当性を確認した。また、位相と音色知覚との間には明確な相関があり、音色の違いは聴覚フィルタ内に含まれる調波間の位相差に起因することが分かった。この結果は楽音の位相差の音色知覚に対する基礎的研究となり得る。また位相に対する聴覚特性究明の一助となるものである。

7.2 今後の課題

位相補正の改良の条件として以下のことがあげられる。

- 実耳レスポンスの測定
- インパルス応答の測定結果に含まれる誤差の検討

また、音色知覚機構の検討として以下のことが必要である。

- 位相の変化幅との相関。
- 音圧との相関。
- 多次元尺度構成法による分析。

最後に位相知覚モデル構築のために、以下の必要がある。

- 音色の知覚特性の定式化。

謝辞

本論文の作成にあたり、熱心に御指導して頂いた赤木 正人 助教授、同講座の岩城 護 助手に厚く御礼申し上げます。

また、本研究を進める過程において、多大なアドバイスを下さった鵜木 祐司 氏、水町 光徳 氏をはじめ、実験に御協力いただいた赤木研究室の皆様にご感謝致します。

最後に、本研究に対する議論とアドバイスを頂いた、すべての皆様に謝意を表します。

参考文献

- [1] von Helmholtz, H. L. F.: "Ueber die Klangfarbe der Vocale,"
Ann. Phys. Chem., 18, 280-290 (1859).
- [2] Fletcher, H.: "Loudness, Pitch and the Timbre of Musical Tones and Their Relation to the Intensity, the Frequency and the Overtone Structure," J. Acoust. Soc. Am., 6, 2, 59-69, (1934).
- [3] Licklider, J. C. R.: "Effects of Changes in the Phase Pattern upon the Sound of a 16-Harmonic Tone," J. Acoust. Soc. Am., 29, Suppl. CB, 780, 53rd meeting, (1957).
- [4] Plomp, R. and Steeneken, H. J. M.: "Effects of Phase on the Timbre of Complex Tones," J. Acoust. Soc. Am., 46, 2409-421 (1969).
- [5] Schroeder, M. R.: "New Results Concerning Monaural Phase Sensitivity," J. Acoust. Soc. Am. Suppl. J5, 1579, 58th meeting, (1959).
- [6] Patterson, R. D.: "A pulse ribbon model of monaural phase perception," J. Acoust. Soc. Am., 82, 5, 1560-1586 (1987).
- [7] 平原達也: "聴覚実験に用いられるヘッドホンの物理特性," 日本音響学会誌, vol. 53, no. 10, pp. 778-806 (1987).
- [8] 鈴木陽一、浅野太、曾根敏雄: "音響系の伝達関数の模擬をめぐって(その2)," J. Acoust. Soc. Jpn., 45, 44-50, (1989).
- [9] Blaert, J. and Law, P.: "Group delay distortions in electroacoustic systems," J. Acoust. Soc. Am., 63, 5, 1478-1483 (1978).

- [10] 河原、津崎、パターソン：” オールパスフィルタの位相操作による時間構造制御とその知覚への影響について,” 音響学会聴覚研資,H-96-74, (1996).
- [11] 鈴木、浅野、曽根：” 音響系の伝達関数の模擬をめぐって (その2),” 日本音響学会誌, vol. 45, no. 1, pp44-50 (1989).
- [12] 鈴木、浅野、金、曽根：” 時間引き延ばしパルスの設計法に関する考察,” 信学技報,EA 9286 (1992).
- [13] 西：” 副テーマ報告書,” 副テーマ報告書,(1994).