

Modelle des peripheren Gehörs am Beispiel der menschlichen Lautheitsempfindung

Dissertation

zur Erlangung des akademischen Grades
Doktor der technischen Wissenschaften

vorgelegt von
Dipl.-Ing. Martin Pflüger

der
Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Graz

1. Begutachter: O.Univ.-Prof. Dipl.-Ing. DDr. Willibald Riedler
Vorstand des Instituts für Nachrichtentechnik und
Wellenausbreitung an der Technischen Universität Graz
2. Begutachter: O.Univ.-Prof. Dipl.-Ing. Dr. Reinhold Weiß
Vorstand des Instituts für Technische Informatik an der
Technischen Universität Graz

IEM

P 531

Graz, September 1997

Ich danke Herrn O.Univ.-Prof. Dipl.-Ing. DDr. Willibald Riedler für die Betreuung dieser Arbeit und für die hervorragenden Arbeitsbedingungen, die mir gewährt wurden.

Ich danke Herrn O.Univ.-Prof. Dipl.-Ing. Dr. Reinhold Weiß als Zweitbegutachter für die Durchsicht und Begutachtung der Arbeit.

Ich danke meinen Kollegen am Institut für Elektronische Musik der Hochschule für Musik und darstellende Kunst in Graz für anregende Diskussionen und vielseitige Unterstützung.

Mein besonderer Dank gilt Herrn HAss. Mag. Dipl.-Ing. Dr. Robert Höldrich für die ausgezeichnete Zusammenarbeit, für die vielen konstruktiven Diskussionen, für das große fachliche Wissen und für die freundschaftliche Motivation.

Nicht zuletzt möchte ich mich natürlich sehr herzlich bei meiner kleinen Familie bedanken: Verena, der besten aller Frauen und unseren Tieren Elpis, China, John und Marylin. Die haben mit ihrer Liebe, Geduld, Zerstreuung und Motivation einen wesentlichen Beitrag zu dieser Arbeit geleistet.

INHALT

1 Einleitung	1
2. Psychoakustik	3
2.1. Reize, Erregungen und Empfindungen	3
2.2. Aufbau und Funktionsweise des peripheren Gehörs	6
2.3. Das Konzept der kritischen Bandbreiten	10
2.4. Das Leistungsdichtespektrum-Modell	12
2.5. Formbestimmung auditiver Filter	13
2.5.1. Die 'Notched-Noise'-Methode	14
2.5.1.1. Symmetrische Einkerbungen	15
2.5.1.2. Unsymmetrische Einkerbungen	16
2.5.1.3. Die roex-Funktionen	19
2.5.1.4. Bestimmung der auditiven Filterformen mit variablen Maskierern	21
2.5.2. Die 'Rippled-Noise'-Methode	22
2.5.3. Psychophysikalische Abstimmkurven	23
2.5.4. Neuronale Abstimmkurven	24
2.5.5. Frequenzkorrekturen	24
2.6. Eigenschaften auditiver Filter	26
2.6.1. Energieäquivalente Rauschbandbreite	26
2.6.2. Auditive Filterformen	28
2.6.2.1. Auditive Filterformen bei konstanten Maskiererpegeln	28
2.6.2.2. Auditive Filterformen bei konstanten Testsignalpegeln	30
2.7. Das Erregungsmuster	31
2.7.1. Erregungsmuster auf Basis der Maskierungsfunktionen	32
2.7.2. Erregungsmuster auf Basis der auditiven Filter	32
2.8. Zusammenfassung und Diskussion	33
3. Berechnung der menschlichen Lautheitsempfindung	36
3.1. Der Lautstärkepegel	37
3.2. Kurven gleicher Lautstärkepegel	38

3.3. Kurven gleicher Reaktionszeit	41
3.4. Die Lautheit	43
3.5. Lautheitsberechnung auf Basis der Kurven gleicher Lautstärkepegel	44
3.5.1. Bewertete Schalldruckpegel	44
3.5.2. Dynamische Phonkurvenanpassung als Erweiterung des dB(A)	45
3.5.2.1. Beschreibung des Verfahrens	45
3.5.2.2. Vergleichende Analysen	48
3.6. Lautheitsberechnung auf Basis des Erregungsmusters	52
3.6.1. Berechnung des Erregungsmusters	53
3.6.2. Berechnung der Lautheit	54
3.7. Zusammenfassung und Diskussion	58
4. Modelle des peripheren Gehörs	63
4.1. Modelle des Außen- und Mittelohres	68
4.2. Modelle der Cochlea	70
4.2.1. Physiologisch motivierte Modelle	70
4.2.2. Funktional motivierte Modelle	71
4.2.3. Implementationsbeispiele	71
4.2.3.1. Kaskadenmodelle	72
4.2.3.2. Parallel - Filterbankmodelle	76
4.2.3.3. Frequenzbereichsmodelle	77
5. Gammatone Filter	79
5.1. Lineare Gammatone Filter	80
5.2. Nichtlineare Gammatone Filter	85
6. Ein nichtlineares Modell des peripheren Gehörs	89
6.1. Ein Modell des Außen- und Mittelohres	89
6.2. Ein nichtlineares Modell der Cochlea	95
6.2.1. Parametrisierung eingangspegelabhängiger nichtlinearer APGTF und OZGTF	95
6.2.2. Parametrisierung ausgangspegelabhängiger nichtlinearer APGTF und OZGTF	110
7. Diskussion	117
Anhang	123
dB(ELC) Berechnung	
Literaturverzeichnis	129

Kapitel 1

Einleitung

Die vorliegende Arbeit beschäftigt sich sowohl mit Modellen des peripheren Gehörs als auch mit Modellen der menschlichen Lautheitswahrnehmung. Eine Verknüpfung dieser beiden Themenbereiche ist insofern möglich, da das periphere Gehör als Vorstufe zur neuronalen Verarbeitung in seiner Funktionsweise einen wesentlichen Einfluß auf die Lautheitsempfindung ausübt. Den meisten Modellen der auditiven Wahrnehmung im allgemeinen und den Modellen der Lautheitswahrnehmung im speziellen ist daher die strukturelle Anlehnung an den anatomischen Aufbau des peripheren Gehörs gemeinsam.

Im peripheren Gehör wird bereits eine teilweise Analyse der Schallsignale und damit verbunden eine Datenreduktion durchgeführt. Dies bedeutet, daß nicht die gesamte akustische Information, die das Außenohr erreicht, auch tatsächlich an das Gehirn weitergeleitet wird. Das primäre Ziel eines Modells des peripheren Gehörs besteht also darin, diejenigen Informationen zu extrahieren, die das periphere Gehör dem Gehirn zur Verfügung stellt.

Die Berechnung der menschlichen Lautstärke- bzw. Lautheitsempfindung ist von besonderem Interesse, da die Lautheitsempfindung mit der Lärmempfindung eng korreliert und in der Lärmmeßtechnik beinahe ausschließlich Lautheitsmeßverfahren verwendet werden. Für die Lärmempfindung selbst gibt es weder eine allgemein anerkannte Definition noch eine allgemein anerkannte Meßgröße. Entsprechend den geltenden Normen, die sich zum größten Teil auf Daten der Lautheitsmeßtechnik stützen, müßte die Lärmmeßtechnik streng genommen in Lautheitsmeßtechnik umbenannt werden.

Die Arbeit ist folgendermaßen gegliedert:

Kapitel 2.: Psychoakustik

In diesem Kapitel werden grundlegende psychoakustische Begriffe definiert. Es ist ein kurzer Überblick über die Funktionsweise und den Aufbau des peripheren Gehörs dargestellt. Ausgehend vom Leistungsdichtespektrum-Modell werden die wesentlichsten Methoden zur Formbestimmung sowie die Eigenschaften auditiver Filter diskutiert.

Kapitel 3.: Berechnung der menschlichen Lautheitsempfindung

In diesem Kapitel werden Methoden zur Berechnung der menschlichen Lautheitsempfindung diskutiert. Es wird zwischen Lautheitsberechnungen auf Basis der Kurven gleicher Lautstärkepegel und auf Basis des Erregungsmusters unterschieden.

Kapitel 4.: Modelle des peripheren Gehörs

Dieses Kapitel beinhaltet einen Überblick über Modelle des peripheren Gehörs. Es wird zwischen Modellen des Außen-, Mittel- und Innenohres unterschieden. Implementationsbeispiele und mögliche Systematisierungen auditiver Modelle werden diskutiert.

Kapitel 5.: Gammatone Filter

In diesem Kapitel werden die Gammatone Funktion und deren Anwendungsmöglichkeiten auf Modelle des menschlichen Gehörs behandelt. Es wird zwischen linearen und nichtlinearen Gammatone Filterbänken unterschieden.

Kapitel 6.: Ein nichtlineares Modell des peripheren Gehörs

In diesem Kapitel wird ein nichtlineares funktionales Modell des peripheren Gehörs vorgestellt. Der lineare Amplitudengang des Außen- und Mittelohrfilters wird aus den Kurven gleicher Lautstärkepegel abgeleitet. Die Frequenzanalyse auf Ebene der Basilarmembran wird mit einer nichtlinearen Gammatone Filterbank (Parallel-Filterbank-Modell) durchgeführt.

Kapitel 7.: Diskussion

In diesem Kapitel werden die wichtigsten Ergebnisse der vorliegenden Arbeit zusammengefaßt.

Viele Modelle des peripheren Gehörs finden in speziellen Anwendungsgebieten wie beispielsweise der Lärmmeßtechnik, der automatisierten Sprachanalyse oder der Kompression von Audiodaten erfolgreich Verwendung. Dies sollte jedoch nicht darüber hinwegtäuschen, daß keinem einzigen Modell auch nur annähernd Allgemeingültigkeit attestiert werden kann. Die Vision einer 'Maschine', die gleich hört wie der Mensch, kann mit dem heutigen technischen Wissensstand nicht realisiert werden.

Kapitel 2

Psychoakustik

In diesem Kapitel werden grundlegende psychoakustische Begriffe definiert. Es ist ein kurzer Überblick über die Funktionsweise und den Aufbau des peripheren Gehörs dargestellt. Ausgehend vom Leistungsdichtespektrum-Modell werden die wesentlichsten Methoden zur Formbestimmung sowie die Eigenschaften auditiver Filter diskutiert.

Das primäre Ziel der Psychoakustik besteht im Entwurf einer 'Maschine', die das Hörempfinden eines 'durchschnittlichen' Menschen reproduziert und dadurch objektiviert¹. Diese Forschungsrichtung bewegt sich zwischen physikalischen, physiologischen und psychologischen Fachbereichen, die sich unterschiedlicher Termini bedienen. Die Akustik als Teilgebiet der Physik befaßt sich mit eindeutig definierten, mathematisch formulierbaren Ordnungsprinzipien, die Physiologie befaßt sich mit Funktionen des menschlichen Organismus und die Psychologie befaßt sich mit den seelisch-geistigen Prozessen des Menschen. Das Verständnis des komplexen Informationsverarbeitungssystems 'Gehör' kann also nur über eine multidisziplinäre Betrachtungsweise erreicht werden².

2.1. Reize, Erregungen und Empfindungen

Physikalische Größen, wie beispielsweise der Schalldruck, werden psychisch nicht als

¹ '... Gibt es Hörempfindungen, die unabhängig von der Versuchsperson allgemeingültigen Gesetzen gehorchen? Diese Frage ist berechtigt, denn im allgemeinen wird angenommen, jeder Mensch empfinde individuell und ganz persönlich. Experimente zeigen jedoch, daß selbst viele, ganz verschieden veranlagte Versuchspersonen über bestimmte Sinnesempfindungen übereinstimmende und damit vorhersagbare Aussagen machen. ...' Die oben gestellte Frage läßt sich also konkretisieren zu: '... Für welche Hörempfindungen erhält man auf Grund subjektiver Beobachtungen die Möglichkeit, Gesetze für den Zusammenhang von Empfindung und Schall so zu formulieren, wie die Naturwissenschaften, z.B. die Physik, auf Grund von objektiven Messungen die Naturgesetze formulieren? ...' ([Zwicker, Feldtkeller, 1967] S.V).

² [Hawkins et al., 1995] S.1

Schalldruck (im Sinne des physikalischen Schalldrucks) wahrgenommen. Genausowenig berechnet das auditive System eine Fourier Transformation, um Schallereignisse durch ihre spektrale Zusammensetzung zu unterscheiden. Vielmehr wird jede Empfindungsgröße¹ von vielen verschiedenen Reizgrößen² beeinflusst. Dabei ist meist eine Reizgröße von dominierender Bedeutung. Der Schalldruckpegel korreliert beispielsweise stark mit der Lautheitsempfindung. Aus diesem Grund werden üblicherweise spektral gewichtete Schalldruckpegel mit der Lautheitsempfindung in Beziehung gesetzt.³

Um menschliche Empfindungen, die mit Adjektiven wie beispielsweise 'laut-leise' beschrieben werden, mit physikalischen Größen in einen metrischen Zusammenhang bringen zu können, müssen sie quantitativ d.h. als Produkt aus Zahlenwert und Einheit erfaßt werden. Eine der wesentlichsten Aufgaben der Psychoakustik besteht also darin, phänomenologische Größen quantitativ zu erfassen. Ein phänomenologisches Absoluturteil wie beispielsweise 'sehr leise' beinhaltet weder den Nullpunkt einer 'laut-leise'-Empfindungsskala noch eine Quantisierungseinheit. Während sich in der Physik jede Meßgröße eindeutig einem Bezugssystem zuordnen läßt, müssen in der Psychologie entsprechend den gegebenen Randbedingungen stets neue Bezugssysteme festgelegt werden.

Verglichen mit physikalischen Systemen beinhalten experimentalpsychologische Meßdaten hohe Varianzen. Mittelwerte derartiger Meßergebnisse können nur als eine Art 'Modell' der tatsächlichen psychologischen Prozesse betrachtet werden (vgl. Abb. 2.1.).

Physiologische Reaktionen auf akustische Reize bezeichnet man als Erregungen⁴. Sie stehen in einem kausalen Verhältnis zu den menschlichen Empfindungen und dienen bei der Lautheitsberechnung beispielsweise als Zwischengrößen. Zur Ermittlung sowie Verifizierung der Erregungsgrößen stellt die physiologische Meßtechnik die zweite Grundlage der Psychoakustik dar. Auch in diesem Bereich begibt man sich auf ein Abstraktionsniveau, um die wesentlichen funktionellen Zusammenhänge begreifbar zu machen.

Sowohl bei psychologischen als auch bei physiologischen Messungen spielt die Akustik als drittes Standbein der Psychoakustik eine wesentliche Rolle, da die Ursache jeweils akustischer Natur ist. Akustische Schallfelder werden stark vereinfacht dargestellt, um die Beschreibungsform der Reizgrößen einfach zu halten. Üblicherweise werden nur ebene oder diffuse Schallfelder betrachtet. Die Abhängigkeitsbeziehungen zwischen Reizen, Erregungen und Empfindungen sind in Abb. 2.1. dargestellt. Die Unterscheidung zwischen tatsächlichen

¹ Unter einer Empfindungsgröße versteht man eine Empfindung, die getrennt von anderen Empfindungen wahrgenommen werden kann. Dabei wird absichtlich der Größenbegriff der Naturwissenschaften eingeführt (vgl. [Zwicker, 1982] S.2).

² Ein Schallereignis wird zu einem Schallreiz, sobald es auf das Hörorgan trifft. Die unterschiedlichen Darstellungsformen dieses Reizes bezeichnet man als Reizgrößen (Schallpegel, Spektrogramm, etc.) (vgl. [Zwicker, 1982] S.1-2).

³ [Zwicker, 1982] S.1-4

⁴ Unter einem Erregungsmuster versteht man beispielsweise das Muster neuronaler Aktivitäten, verursacht durch akustische Reize. Eine weitere Bedeutung des Erregungsmusters bezieht sich auf die Erregungsverteilung entlang der Basilarmembran (vgl. Kap. 2.2.). Diese Erregungsverteilung wird üblicherweise in dB angegeben, da diese Größe in der Akustik üblich ist.

Prozessen und Modellen ist grafisch hervorgehoben, da die jeweiligen Modelle starke Vereinfachungen der tatsächlichen Gegebenheiten beinhalten.

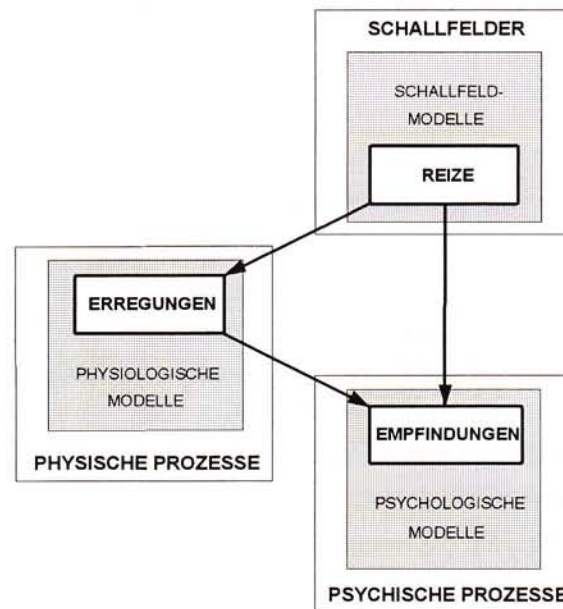


Abb. 2.1.: Abhängigkeitsbeziehungen zwischen Reizen, Erregungen und Empfindungen

Zur Modellierung des menschlichen Hörvermögens wird eine Kausalkette aus Reizen, Erregungen und Empfindungen postuliert (Abb. 2.2.).

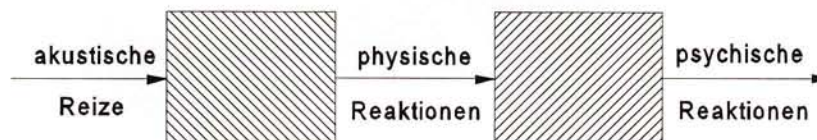


Abb. 2.2.: Modellierung des menschlichen Gehörs

Die gesuchten Zusammenhänge sind als 'black box' dargestellt. Eine Vereinfachung dieses Modells stellt eine direkte Verbindung zwischen Reizen und Empfindungen her und verzichtet auf physiologisch motivierte Zwischengrößen.

Aus signalverarbeitungstheoretischer Sicht besteht die Aufgabe darin, die physiologischen sowie psychologischen Reaktionen auf akustische Reize möglichst exakt und recheneffizient nachzubilden. Daraus ergeben sich Modelle des menschlichen Gehörs bzw. Modelle einzelner Eigenschaften des menschlichen Gehörs. Deren Entwicklung ist eng mit experimentellen Daten verbunden. Je mehr bzw. je bessere Daten über das Reaktionsverhalten gefunden werden, desto

besser werden die Modelle¹. Die ermittelten Reaktionen bzw. die zur Verfügung gestellten Daten sind jeweils nur unter bestimmten Randbedingungen gültig (vgl. Kap. 3.2.). Der Gültigkeitsbereich eines Modells läßt sich nicht auf beliebige Schallarten sowie Umgebungssituationen ausweiten. Keines der vorhandenen Modelle ist auch nur annähernd imstande, das komplexe Informationsverarbeitungssystem des menschlichen Gehörs in seiner Gesamtheit nachzubilden.

2.2. Aufbau und Funktionsweise des peripheren Gehörs

Die meisten Modelle der menschlichen Hörwahrnehmung setzten sich aus Funktionseinheiten zusammen, die einzelne physiologische Verarbeitungsstufen des Gehörs direkt nachbilden, oder sich in ihrer Funktionsweise einer physiologischen Verarbeitungsstufe zuordnen lassen. Das periphere Gehör setzt sich aus dem Außen-, Mittel- und dem Innenohr zusammen. Diese drei Stufen lassen sich sowohl funktionell als auch anatomisch voneinander eindeutig unterscheiden, was bei der zentralen (neuronalen) Verarbeitung von Schallreizen nicht in diesem Maße möglich ist. Man beschränkt sich häufig auf die Modellierung der peripheren Gehörs (vgl. Kap. 4), dessen Aufbau und Funktionsweise im folgenden kurz beschrieben ist.

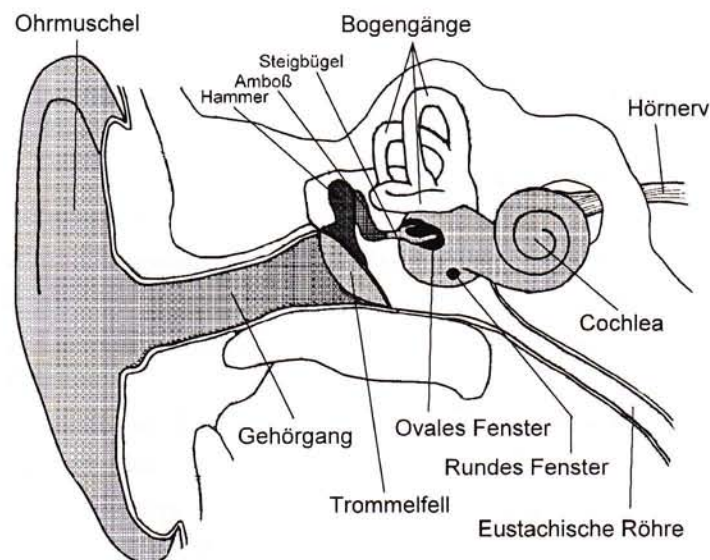


Abb. 2.3.: Aufbau des peripheren menschlichen Gehörs

¹ [Hawkins, et al., 1995] S.73-74

Das Außenohr¹

Das Außenohr besteht aus der Ohrmuschel und dem Gehörgang². Es nimmt die Schallenergie über die Ohrmuscheln auf und leitet diese im Gehörgang zum Trommelfell weiter (Luftleitung). Gleichzeitig setzt der einfallende Schall den gesamten Schädelknochen in Schwingung (Knochenleitung). Diese Schwingung wird ebenfalls an die Cochlea weitergeleitet und bewirkt eine Hörwahrnehmung³. Durch den Abstand zwischen den beiden Ohrmuscheln und Abschattungen am Kopf entstehen Laufzeit- und Intensitätsdifferenzen sowie Phasenverschiebungen, die eine Schalllokalisierung ermöglichen. Die asymmetrische Form der Ohrmuscheln ermöglicht eine Differenzierung zwischen 'vorne' und 'hinten'. Die Übertragungsfunktionen des Außenohres und des menschlichen Torso bei unterschiedlichen Schalleinfallrichtungen bezeichnet man als HRTF (Head Related Transfer Functions⁴) Kurven. Das Außenohr wirkt in grober Näherung wie eine gedeckte Orgelpfeife. Die Resonanzen liegen daher bei $\lambda/4$, $3\lambda/4$, $5\lambda/4$, etc.. Nimmt man für den Gehörgang eine Länge von 2,5 cm bei einer Schallgeschwindigkeit von $c = 340$ m/s an, so liegen diese Resonanzen bei 3,4 kHz, 10,2 kHz und 17 kHz. Die höhere Empfindlichkeit des Gehörs an 3,4 und 10,2 kHz ist aus den Kurven gleicher Lautstärkepegel (vgl. Kap. 3.2.) ersichtlich.

Das Mittelohr⁵

Die vom Trommelfell aufgenommene Luftschwingung wird über die Gehörknöchelchen Hammer, Amboß und Steigbügel auf das ovale Fenster des Innenohres übertragen. Diese drei Knöchelchen bilden das Mittelohr. Es hat die Aufgabe Luftschall in Flüssigkeitsschall zu transformieren. Diese 'Impedanzanpassung' wird sowohl mit der Hebelübersetzung der drei Mittelohrknöchelchen als auch mit der Flächentransformation vom Trommelfell zum ovalen Fenster erreicht. Im mittleren Frequenzbereich (um 1 kHz) ist das mit Gehörknöchelchen und Innenohr belastete Trommelfell genau an die Schallwellenimpedanz angepaßt. Ab einer Frequenz von etwa 2 bis 4 kHz wird die Masse dieses mechanischen Systems wirksam, was sich in einer Verschlechterung der Übertragungsfähigkeit äußert. Die Wirkung des Mittelohres kommt der eines Tiefpaßfilters gleich.

Der Mittelohrraum ist über die Eustachische Röhre mit dem Rachenraum verbunden. Diese Verbindung ist geschlossen und öffnet sich kurzzeitig beim Schlucken. Damit wird verhindert, daß eine Änderung des Atmosphärendrucks eine Verlagerung des Trommelfells sowie der Mittelohrkette aus dem 'Arbeitspunkt' bewirkt.

¹ [Keidel, 1975] S.44-63, [Zwicker, 1982] S.21, [Jahn, Santos-Sacchi, 1988] S.29-40

² Der Gehörgang besitzt durchschnittlich eine Länge von etwa 2,5 cm und einen Durchmesser von 0,7 cm (vgl. [Jahn, Santos-Sacchi, 1988] S.29).

³ Die Weiterleitung über das Schläfenbein an das Innenohr ist unter normalen Bedingungen von sekundärer Bedeutung ([Blauert, 1974] S.43)

⁴ [Moore, 1995], S.321-322

⁵ [Keidel, 1975] S.44-63, [Zwicker, 1982] S.21-22, [Jahn, Santos-Sacchi, 1988] S.103-124

Bei starker Schalleinwirkung wird die Spannung des Trommelfells vergrößert und die Mittelohrknöchelchen werden aus dem Arbeitspunkt verdreht. Dieser Schutzmechanismus stellt die erste wesentliche Nichtlinearität des peripheren Gehörs dar.

Das Innenohr¹

Das Innenohr oder die Cochlea nimmt eine zentrale Rolle bei der Hörwahrnehmung ein. Die Funktionsweise läßt sich in eine mechanische Spektralanalyse (Frequenz-, Ortstransformation) und in die Transformation von mechanischen Schwingungen zu neuronalen Impulsen einteilen. Durch die schneckenähnliche Form wird die Cochlea auch als Hörschnecke bezeichnet. Da diese Schneckenform keinen wesentlichen Einfluß auf die Funktionsweise der Cochlea ausübt und primär der Platzersparnis dient, wird sie meist abgerollt dargestellt (Abb. 2.4., schematische Darstellung der Cochlea).

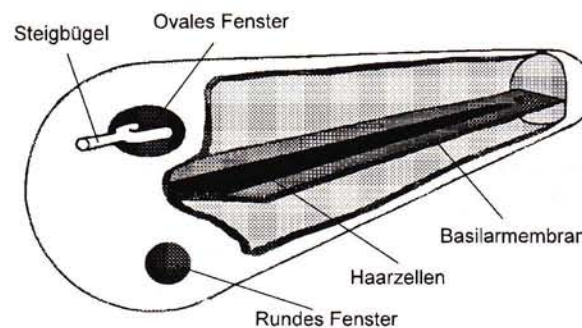


Abb. 2.4.: Schematische Darstellung der Cochlea (abgerollt)

Die Bewegung des Steigbügels (Mittelohr) wird über das ovale Fenster auf die Lymphflüssigkeit der Cochlea übertragen. Das runde Fenster dient als Druckausgleich, da die Lymphflüssigkeit inkompressibel ist. Die Basilarmembran gibt den am ovalen Fenster erzeugten Volumenschiebungen nach, und es bilden sich Wanderwellen entlang dieser Membran aus (Abb. 2.5.). Die Hüllkurve der Wanderwellen steigt bei einer sinusförmigen Erregung langsam an und verebbt nach Erreichung des Maximums verhältnismäßig schnell. Je höher die Frequenz, desto kürzer ist die Wanderwelle. Dadurch ergibt sich eine ortsabhängige spektrale Zerlegung der eintreffenden Schwingungen.

¹ [Keidel, 1975] S.6-35, [Zwicker, 1982] S.22-27, [Jahn, Santos-Sacchi, 1988] S.159-172

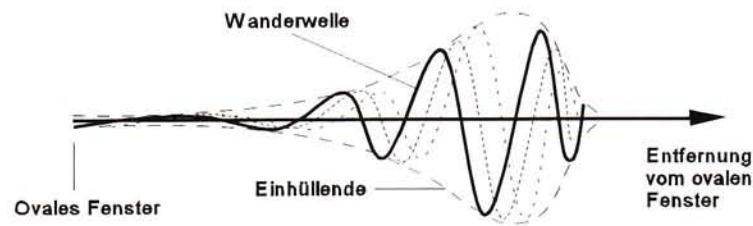


Abb. 2.5.: Wanderwelle entlang der Basilarmembran

Auf der Basilarmembran befinden sich die elektromechanischen Wandler (Haarzellen), die bei einer Auslenkung der Basilarmembran elektrische Impulse entlang des Hörnervs feuern.

Der Hörnerv¹

Der Hörnerv leitet die neuronalen Impulse Richtung Gehirn weiter. Er besteht aus etwa 20.000 bis 30.000 afferenten² Fasern und etwa 2000 efferenten Fasern. Die wahrgenommenen akustischen Frequenzen liegen gewöhnlich deutlich höher als die neuronalen Feuerfrequenzen. Die Dynamik einer einzelnen Faser liegt im Bereich von etwa 40 dB. Das Schwingungsverhalten der Basilarmembran wird von den efferenten Nervenfasern gesteuert. Dieser Mechanismus stellt die zweite wesentliche Nichtlinearität des peripheren Gehörs dar.

Die Funktion des Hörnervs beschränkt sich nicht nur auf die Weiterleitung der neuronalen Impulse. Die laterale Hemmung verursacht beispielsweise eine Kontrastverstärkung zwischen benachbarten Nervenfasern.

Die neuronalen Impulse des Hörnervs werden nicht direkt auf eine Gehirnregion abgebildet, sondern es gibt eine Vielzahl von Verschaltungen und Zwischenstationen, bevor diejenigen Hirnregionen erreicht werden, die man aufgrund erhöhter Aktivität bei Beschallung als Wahrnehmungszentren akustischer Reize bezeichnet. In diesen Zwischenstationen wird die eintreffende Information bereits verarbeitet, d.h. es finden sowohl partielle als auch totale Aussonderungen statt. Der Verlauf der Hörbahn bzw. die Funktionalität einzelner Zwischenstationen ist zur Zeit noch nicht so weit erforscht, daß sich aus den vorhandenen Hypothesen ein verifizierbares Modell entwickeln ließe.

¹ [Keidel, 1975] S.164-177, [Zwicker, 1982] S.27-30, [Jahn, Santos-Sacchi, 1988] S.359-384

² Afferente Fasern leiten in Richtung Gehirn, efferente Fasern leiten vom Gehirn weg.

2.3. Das Konzept der kritischen Bandbreiten

Die Funktionsweise des peripheren menschlichen Gehörs kann mit einer parallelen Bank von überlappenden Bandpaßfiltern (auditive Filter) simuliert werden. Dabei geht man davon aus, daß der subjektive Eindruck eines Schallereignisses mit den Ausgangskanälen dieser Filterbank in Zusammenhang gebracht werden kann. Dieses Konzept ist dem der kritischen Bandbreiten¹ ähnlich.

Das Konzept der kritischen Bandbreiten läßt sich auf ein Experiment von Fletcher aus dem Jahre 1940 zurückführen². Dabei wird die Wahrnehmbarkeitsschwelle (Hörschwelle) eines Sinustons als Funktion der Bandbreite eines Maskierungsrauschens bestimmt (Abb 2.6.).

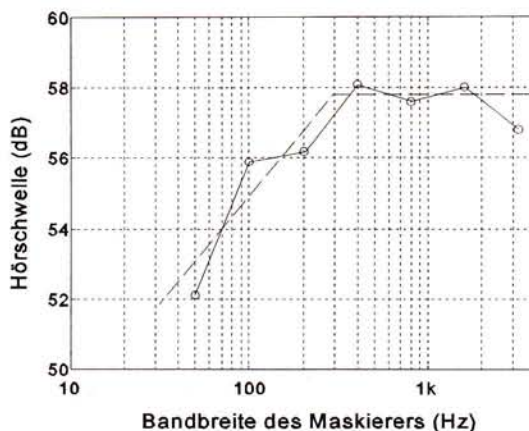


Abb. 2.6.: Hörschwelle eines 2 kHz Sinustons als Funktion der Bandbreite eines Maskierungsrauschens bei einer Mittenfrequenz von 2 kHz, o: gemessene Daten, --: Näherung, (vgl. [Moore, 1993])

Die Frequenz des Sinustons entspricht dabei stets der Mittenfrequenz des Schmalbandrauschens. Die Leistungsdichte des Rauschens wird konstant gehalten. Dadurch erhöht sich bei steigender Bandbreite die Gesamtleistung des Maskierers. Die Hörschwelle des Sinustons steigt anfangs mit steigender Bandbreite an, ändert sich jedoch ab einer bestimmten Bandbreite des Maskierers nicht weiter (Abb.2.6.).

Aus diesem Verhalten des auditiven Systems läßt sich die Erkenntnis ableiten, daß zur Detektion eines schmalbandigen Schallereignisses in Gegenwart eines Maskierers nur eine bestimmte, um das Testsignal zentrierte Bandbreite von Bedeutung ist. Stellt man sich diese Bandbreite als idealen Bandpaßfilter (mit

unendlichen Flankensteilheiten) vor, so erhöht sich die Maskierungsrauschleistung innerhalb dieser Bandbreite nur so lange, bis die Grenzfrequenzen erreicht sind. Übersteigt die Bandbreite des Maskierungsrauschens diese Filterbandbreite, so nimmt die Gesamtleistung des Maskierungsrauschens zwar weiter zu, die durch den Filter detektierte Leistung jedoch nicht. Mit dieser Theorie läßt sich das in Abb. 2.6 dargestellte Maskierungsverhalten erklären.

Fletcher nannte diese Bandbreite, die in etwa dem Knickpunkt der in Abb. 2.6. dargestellten strichlierten Geraden entspricht 'kritische Bandbreite' (CB, 'critical bandwidth'). Diese Bandbreite konnte in einer Vielzahl von Experimenten nachvollzogen werden³ und stellt eine der wesentlichsten Erkenntnisse über die Eigenschaften des menschlichen Gehörs dar.

¹ [Zwicker, 1982] S.46-53

² [Moore, 1993]

³ [Zwicker, 1982] S.46-53, [Zwicker, Fastl, 1990] S.133-155

Die kritische Bandbreite als Funktion der Mittenfrequenz wurde von Zwicker und Terhardt¹ 1980 angegeben mit:

$$CB(f) = 25 + 75 \left(1 + 1.4 \left(\frac{f}{1000} \right)^2 \right)^{0.69} \quad \text{Gl. 2.1.}$$

Aufgrund ihrer Bedeutung bei der Beschreibung von Hörwahrnehmungen dient diese Frequenzbreite (in der deutschsprachigen Literatur meist als 'Frequenzgruppenbreite' bezeichnet) als Grundlage einer gehörbezogenen Frequenzskala, der Frequenzgruppen-, Bark-, z-, CB-Raten- oder Tonheits-Skala. Die Frequenzgruppenskala erhält man, indem man die Frequenzgruppen so aneinanderreihet, daß jeweils die obere Grenze der unteren Frequenzgruppe mit der unteren Grenze der oberen Frequenzgruppe übereinstimmt. Dieser Skala wurde die Einheit Bark zugeteilt.

Unter Verwendung der kritischen Bandbreiten aus Gl. 2.1. wird die Transformation vom Hz- in den Bark-Bereich ausgedrückt durch:²

$$z(f) = 13 \arctan(0.00076f) + 3.5 \arctan\left(\frac{f}{7500}\right)^2 \quad \text{Gl. 2.2.}$$

Eine Frequenzbreite von $\Delta z = 1$ Bark entspricht dabei stets einer Frequenzgruppenbreite. Diese Transformation stimmt mit den psychoakustisch gemessenen Daten über den gesamten Hörbereich mit einer Genauigkeit von -0.25 bis +0.2 Bark überein. Eine Alternative zu Gl. 2.2. wurde 1990 von Traunmüller³ vorgeschlagen:

$$z(f) = \left(\frac{26.81f}{1960 + f} \right) - 0.53 \quad \text{Gl. 2.3.}$$

Diese vereinfachte Transformation entspricht den psychoakustisch gemessenen Daten im Frequenzbereich zwischen 0.2 und 6.7 kHz mit einer Genauigkeit von ± 0.05 Bark.

Die kritische Bandbreite ist, entsprechend ihrer Definition aus psychoakustischen Experimenten, ein empirisches Phänomen: '*... that bandwidth at which subjective responses rather abruptly*

¹ Mit dieser Gleichung konnten die in den 80er Jahren verwendeten kritischen Bandbreiten mit einer Genauigkeit von $\pm 10\%$ approximiert werden ([Zwicker, Terhardt, 1980]).

² [Zwicker, Terhardt, 1980]

³ [Traunmüller, 1990]

*change ...*¹. Der Zusammenhang zu physiologischen Messungen läßt sich insofern herstellen, als daß die Abbildung einzelner Frequenzen auf die Basilarmembran (Frequenz-Orts-Transformation) in etwa dieser Transformation entspricht.

2.4. Das Leistungsdichtespektrum-Modell

Aus der Idee der kritischen Bandbreiten entwickelte sich das Leistungsdichtespektrum-Modell ('Power-Spectrum-Model'). Dieses Modell bildet die Grundlage funktional motivierter auditiver Filterbänke.²

Das Leistungsdichtespektrum-Modell geht von folgenden Annahmen aus³:

- Das periphere menschliche Gehör kann mit einer parallelen Bank von linearen, überlappenden Bandpaßfiltern (auditive Filter) nachgebildet werden.
- Ein schmalbandiges Testsignal (vorzugsweise ein Sinuston) wird in einer Maskierungsumgebung mit demjenigen auditiven Filter wahrgenommen, der mit der Mittenfrequenz des Testsignals in etwa zusammenfällt. Dies bedeutet, daß nur diejenigen spektralen Komponenten des Maskierers für die Wahrnehmung des Testsignals von Bedeutung sind, die um das Testsignal zentriert sind. Die Gewichtung der einzelnen spektralen Maskiererkomponenten wird durch die Form der auditiven Filter bestimmt. Die Wahrnehmbarkeit des Testsignals stellt eine Funktion des SNR (Signal-Stör-Abstand) dieses auditiven Filters dar.
- Maskierer und Testsignal werden durch Leistungsdichtespektren repräsentiert, die mit einem, verglichen mit dem menschlichen Zeitaufklärungsvermögen, relativ langen Zeitfenster gebildet werden. Phasenbeziehungen sowie kurze zeitliche Änderungen der Einhüllenden werden dabei nicht berücksichtigt.

Genau genommen ist keine dieser Annahmen korrekt:

- Die auditiven Filter sind nicht linear.
- Zur Detektion eines schmalbandigen Testsignals kann die zur Wahrnehmung wesentliche Information von einem auditiven Filter stammen, der nicht um das Signal zentriert ist ('off-frequency-listening'). Weiters können die Informationen mehrerer auditiver Filter kombiniert werden. Das Verhalten des Maskierers außerhalb des um das Testsignal zentrierten auditiven Filters kann folglich die Detektion des Signals beeinflussen.
- Kurze zeitliche Änderungen des Maskierers können sich auf die Maskierung des Testsignals auswirken. Diese kurzen zeitlichen Änderungen werden durch die lange

¹ [Moore, Glasberg, 1983]

² Funktionale Modelle des peripheren Gehörs modellieren experimentalpsychologische Meßdaten ohne Berücksichtigung der zugrundeliegenden physikalischen bzw. physiologischen Vorgänge (vgl. Kap. 4).

³ vgl. [Moore, 1993]

Zeitfensterung nicht berücksichtigt.

Dieses Modell wird als Ausgangspunkt zur funktionalen Nachbildung des peripheren Gehörs weitgehend akzeptiert. Die zur Zeit durchgeführten Forschungen auf diesem Gebiet beziehen sich auf die Form der auditiven Filter sowie auf deren nichtlineares und zeitliches Verhalten.

2.5. Formbestimmung auditiver Filter

Die meisten Methoden zur Approximation auditiver Filter basieren auf den Annahmen des Leistungsdichtespektrum-Modells. Die unterschiedlichen Formen eines auditiven Filters an einer bestimmten Mittenfrequenz werden durch Messung der Hörschwellen eines Sinussignals, verdeckt durch einen Maskierer, bestimmt.

Bezeichnet man das Leistungsdichtespektrum eines Maskierers mit $N(f)$ und die Gewichtungsfunktion bzw. die Form des auditiven Filters mit $W(f)$, so wird das Leistungsdichtespektrum-Modell ausgedrückt durch:

$$P_s = K \int_0^{\infty} N(f) W(f) df \quad \text{Gl. 2.4.}$$

Der Wert P_s entspricht der Leistung des Testsignals an der Hörschwelle. Die Konstante K gleicht dem SNR am Ausgang des Filters $W(f)$, das notwendig ist, um die Hörschwelle zu erreichen. Sie bezieht sich auf die Effizienz des dem auditiven Filter folgenden Detektionsmechanismus. Je höher die Konstante K ist, desto schlechter ist die Detektion¹.

Durch Änderung des Maskierer-Spektrums und Messung der entsprechenden Änderungen von P_s ist es möglich $W(f)$ d.h. die auditive Filterform zu bestimmen². Das Maskierungssignal muß dabei so gewählt werden, daß die Annahmen des Leistungsdichtespektrum-Modells möglichst genau erfüllt werden. Setzt sich das Maskierungssignal beispielsweise aus mehreren Sinuskomponenten zusammen, so bilden sich Schwebungen (Amplitudenmodulation der Einhüllenden, primär mit der Differenzfrequenz aus zwei Sinus-Komponenten) aus, die die Hörbarkeit des Testsignals beeinflussen. Grundsätzlich sind Rauschmaskierer für Maskierungsexperimente besser geeignet, da sich dabei in der Regel weder Schwebungen innerhalb des Maskierers noch zwischen dem Testsignal (vorzugsweise Sinussignal) und dem Maskierer ausbilden können. Im Falle von Schmalbandrauschen (meist ausgeprägte

¹ Einer der wesentlichen Vorteile dieser Betrachtungsweise besteht darin, daß die Filterformen $W(f)$ bzw. die Frequenzselektivität getrennt von der Effizienz des Detektionsmechanismus behandelt werden kann. Liegen die Maskierungsdaten einer Versuchsperson beispielsweise um konstant 3dB höher als bei einer zweiten Versuchsperson, so besitzen beide Versuchspersonen zwar die selbe Frequenzselektivität $W(f)$, aber eine um 3 dB unterschiedliche Detektionseffizienz K .

² [Moore, Glasberg, 1987]

Schwankungen der Einhüllenden) sind ebenfalls Interaktionen zwischen Maskierer und Testsignal hörbar¹. Um die Hörbarkeit bzw. die Hörschwelle eines Testsignals ausschließlich auf die Maskierung zurückführen zu können, sollte also vorzugsweise breitbandiges Rauschen als Maskierer verwendet werden.

Die zweite wesentliche Voraussetzung für Maskierungsexperimente zur Bestimmung der Form von auditiven Filtern ist, daß möglichst nur ein auditiver Filter zur Hörwahrnehmung beiträgt. Die Zusammensetzung von Maskierer und Testsignal sollte daher so gestaltet werden, daß der auditive Filter, der um die Testsignalfrequenz zentriert ist, das höchste SNR besitzt. Würde man beispielsweise einen Maskierer verwenden, der ausschließlich über oder unter dem Signal zentriert ist (Hochpaß- oder Tiefpaßrauschen), so gäbe es eine Reihe von auditiven Filtern mit etwa gleich großem SNR. In diesem Fall würden Versuchspersonen die Informationen dieser Filter zur Schallwahrnehmung kombinieren. Da bei einem derartigen Versuch mehrere Filter an der Wahrnehmung beteiligt sind, kann daraus nicht auf die Form eines einzelnen auditiven Filters geschlossen werden.

2.5.1. Die 'Notched-Noise'-Methode

Die zur Zeit gebräuchlichste Methode zur Formbestimmung von auditiven Filtern ist die sogenannte 'Notched-Noise' (eingekerbtes Rauschen)-Methode². Bei dieser Methode wird die Wahrnehmbarkeitsschwelle eines Sinustons bestimmt, der sich in einer Einkerbung (Notch) eines breitbandigen Maskierers (Rauschen) befindet (vgl. Abb. 2.7.). Die Änderungen der Wahrnehmbarkeitsschwelle des Sinustons als Funktion der Einkerbungsbreite werden dafür verwendet, um die Form des auditiven Filters zu bestimmen, der um die Testsignalfrequenz zentriert ist. Gl. 2.4. kann in diesem Fall so interpretiert werden, daß die Leistung des Sinustons an der Hörschwelle in einem konstanten Verhältnis K zu demjenigen Rauschanteil steht, der unter die Flanken des auditiven Filters fällt (Abb. 2.7., schraffierter Bereich). Diese Annahme entspricht dem Konzept der kritischen Bandbreiten (vgl. Kap. 2.3.).

Durch die verhältnismäßig schmale Rausch-Einkerbung wird davon ausgegangen, daß das auditive System zur Detektion des Testsignals nur einen einzelnen auditiven Filter verwendet. Es werden die aus experimentalpsychologischen Maskierungsexperimenten ermittelten durchschnittlichen Eigenschaften des betrachteten Frequenzbereiches durch den Amplitudengang eines einzelnen Bandpaßfilters (des auditiven Filters) ausgedrückt.

¹ Eine Alternative zu Schmalbandrauschen mit ausgeprägten zeitlichen Fluktuationen stellt das sogenannte 'Low Noise Noise' dar. Diese Rauschsignale besitzen das gleiche Leistungsdichtespektrum wie 'gewöhnliches' Schmalbandrauschen, mit dem Unterschied, daß die zeitlichen Fluktuationen um ein Vielfaches schwächer ausgeprägt sind. Low Noise Noise wird zur Zeit in der Psychoakustik kaum verwendet ([Heijden, Kohlrausch, 1995]).

² [Patterson, 1976], [Weber, 1977], [Patterson, Nimmo-Smith, 1980], [Moore, Glasberg, 1983], [Glasberg, Moore, 1986], [Dubnko, Dirks, 1989]

2.5.1.1. Symmetrische Einkerbungen

In einer ersten Näherung werden die auditiven Filter mit einem Maskierungsrauschen bestimmt, dessen Einkerbung symmetrisch um das Testsignal zentriert ist. Diese Methode ist in Abb. 2.7. grafisch dargestellt. Die Leistungsdichte des Maskierers wird im Durchlaßband als konstant ($N(f) = N_0$) angenommen. Der vom Gehör verwendete auditive Filter wird als symmetrisch um das Testsignal zentriert gedacht d.h. die Signalfrequenz und die Filtermittenfrequenz stimmen überein ($f_0 = f_c$).

Die Einkerbungsbreite beträgt $2\Delta f$. Mit diesen Vereinfachungen wird aus Gl. 2.4:¹

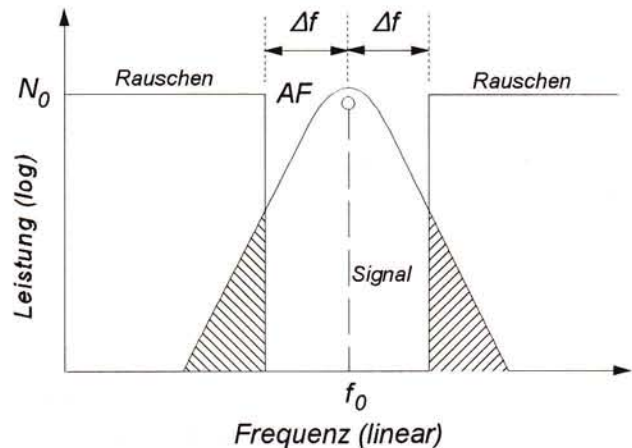


Abb. 2.7.: Grafische Darstellung der Notched-Noise-Methode mit symmetrischer Einkerbung

$$P_S = K N_0 \int_0^{f_c - \Delta f} W(f) df + K N_0 \int_{f_c + \Delta f}^{\infty} W(f) df \quad \text{Gl. 2.5.}$$

Die beiden Integrale aus Gl. 2.5. sind in Abb. 2.7. schraffiert dargestellt. Mit der normierten Frequenzvariable g

$$g = \frac{|f - f_c|}{f_c} \quad \text{Gl. 2.6.}$$

und einer Unterscheidung zwischen Δf_l als dem Abstand zwischen der Filtermittenfrequenz und dem unteren Rauschband und Δf_u als dem Abstand zwischen der Filtermittenfrequenz und dem oberem Rauschband wird aus Gl 2.5.:

$$P_S = K N_0 f_0 \int_0^{-\Delta f_l / f_c} W(g) dg + K N_0 f_0 \int_{\Delta f_u / f_c}^{\infty} W(g) dg \quad \text{Gl. 2.7.}$$

Da sowohl der auditive Filter als auch der Maskierer symmetrisch um das Signal zentriert sind,

¹ vgl. [Moore, Glasberg, 1987]

ergeben die beiden Integrale gleiche Werte. Für diesen speziellen Fall kann Gl. 2.7. vereinfacht werden zu:

$$P_s = 2K N_0 f_0 \int_0^{-\Delta f_i/f_c} W(g) dg \quad \text{Gl. 2.8.}$$

In psychoakustischen Versuchsreihen wird die Hörschwelle des Testsignals als Funktion der Einkerbungsbreite bestimmt (Bsp.: Abb. 2.8.)¹. Aus diesen Daten wird die Form des entsprechenden auditiven Filters $W(g)$ abgeleitet.

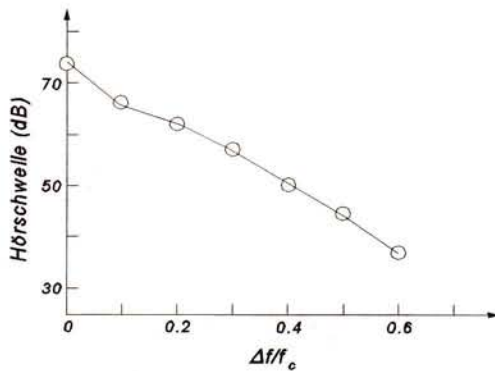


Abb. 2.8.: Psychoakustisch ermittelte Hörschwellen für eine Signalfrequenz von 200 Hz als Funktion der Einkerbungsbreite (symmetrisch um das Signal zentriert), (vgl. [Moore, 1995] S.174)

Dafür wird $W(g)$ aus Gl. 2.8. durch eine einfach parametrisierte Funktion wie beispielsweise die roex(p)-Funktion (vgl. Kap. 2.5.1.3.)

$$W(g) = (1 + pg)e^{-pg} \quad \text{Gl. 2.9.}$$

substituiert. Der Parameter p wird in einer iterativen Prozedur jeweils für eine Einkerbungsbreite so lange verändert, bis die berechneten Hörschwellen um einen minimalen mittleren quadratischen Fehler von den gemessenen Hörschwellen abweichen². Die mit symmetrischen Einkerbungen bestimmten auditiven Filter sind ausschließlich symmetrisch und für mittlere Schallpegel gültig.

2.5.1.2. Unsymmetrische Einkerbungen

Bei größeren Rauschpegeln sowie bei hörgeschädigten Probanden sind die auditiven Filter unsymmetrisch, d.h. die hochfrequenten und tieffrequenten Flanken besitzen auf einer linearen Frequenzachse unterschiedliche Steilheiten. Diese Unsymmetrie ist von wesentlicher Bedeutung, da nur mit unsymmetrischen Filtern das sogenannte 'upward spread of masking' beschreibbar ist³. Berücksichtigt man bei Notched-Noise Experimenten zusätzlich Einkerbungen, die

¹ [Moore, 1993]

² Die Detektionskonstante K ergibt sich aus dieser Prozedur durch Anpassung des errechneten Hörschwellenmittelwertes an den psychoakustisch gemessenen Hörschwellenmittelwert (vgl. [Patterson et al., 1982]).

³ Unter 'upward spread of masking' versteht man den nicht linearen Zusammenhang zwischen Maskiererpegel und Maskierungsschwelle. Aus diesem Effekt läßt sich eine erste Abschätzung der

unsymmetrisch um das Signal zentriert sind, so besteht die Möglichkeit, unsymmetrische auditive Filter zu bestimmen¹.

Bei unsymmetrisch zentrierten Einkerbungen wirkt sich der Effekt des off-frequency-listenings aus, d.h. die Position der Mittenfrequenz des auditiven Filters kann mit der Testsignalfrequenz nicht gleichgesetzt werden. Dieser Effekt ist in den beiden Grafiken Abb. 2.9. und 2.10. dargestellt.

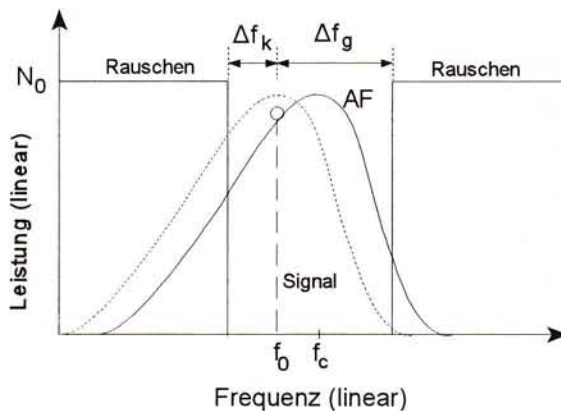


Abb. 2.9.: Off-frequency-listening mit $f_0 < f_c$

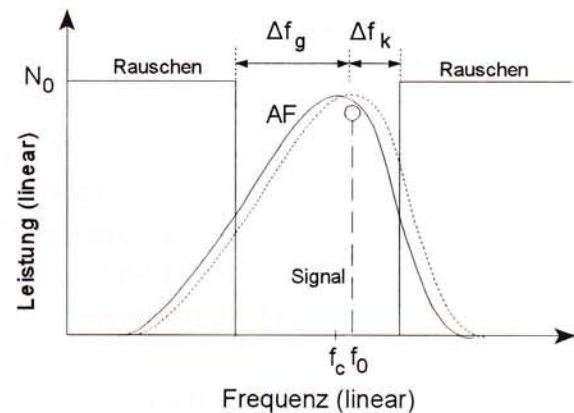


Abb. 2.10.: Off-frequency-listening mit $f_0 > f_c$

In Abb. 2.9. befindet sich das Testsignal näher an der unteren Rauschflanke, in Abb. 2.10. näher an der oberen Rauschflanke. Die punktiert dargestellten auditiven Filter sind jeweils um das Testsignal zentriert, die durchgezogenen auditiven Filter entsprechen in etwa den vom auditiven System tatsächlich zur Detektion verwendeten Filtern. Auch in diesem Fall wird dem auditiven System unterstellt, nur einen einzelnen auditiven Filter bei der Wahrnehmung des Signals zu verwenden.

Um die Form dieser unsymmetrischen auditiven Filter bestimmen zu können, wird in einem ersten Schritt eine Annahme darüber getroffen, an welcher Mittenfrequenz sich der verwendete Filter befindet. In Anlehnung an das Leistungsdichtespektrum-Modell wird davon ausgegangen, daß der auditive Filter mit dem höchsten SNR verwendet wird². Die durchgezogenen Filter in Abb. 2.9. und 2.10. besitzen das größte SNR. Aufgrund der Unsymmetrie der dargestellten Filter ist die Verschiebung bzw. der Effekt des off-frequency-listenings im Falle $f_0 < f_c$ stärker ausgeprägt³. Bei Verwendung eines symmetrischen auditiven Filters wäre die Verschiebung der Mittenfrequenzen in beiden Fällen gleich.

Unsymmetrie der auditiven Filter ableiten. Demzufolge sollte die untere Flanke des auditiven Filters mit steigendem Schallpegel flacher werden.

¹ [Patterson, Nimmo-Smith, 1980], [Glasberg, et al., 1984], [Glasberg, Moore, 1990]

² [Patterson, Nimmo-Smith, 1980]

³ Aus dieser Verschiebung ist ersichtlich, daß das untere Rauschband einen stärkeren Einfluß auf das Testsignal ausübt als das obere Rauschband.

Zur Bestimmung unsymmetrischer auditiver Filter werden die symmetrischen Daten der psychoakustischen Versuchsreihen um unsymmetrische Einkerbungsdaten erweitert. Üblicherweise wird das $\Delta f_g/f_0$ um den Faktor 0.2 größer als das $\Delta f_k/f_0$ gewählt.

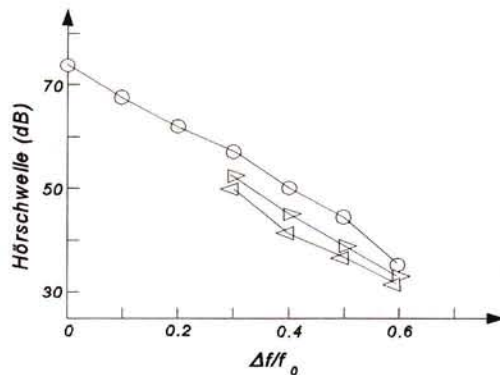


Abb. 2.11.: Hörschwellen für symmetrische (o) und unsymmetrische (\triangleright für $f_0 < f_c$, \triangleleft für $f_0 > f_c$) Einkerbungen, (vgl. [Moore, 1995] S.174)

Abb. 2.11. zeigt einen vollständigen Datensatz gemessener Hörschwellen¹. Die Kreise symbolisieren symmetrische Bedingungen. Die Pfeile nach rechts (\triangleright) beziehen sich auf eine Verschiebung der Einkerbung (bezogen auf die Signalfrequenz) nach rechts (vgl. Abb. 2.9.) und die Pfeile nach links (\triangleleft) symbolisieren die Einkerbungsverschiebung nach links (vgl. Abb. 2.10.). Das Verhältnis $\Delta f/f_0$ bezieht sich jeweils auf das Verhältnis $\Delta f_k/f_0$. Da sich die Maskierung durch das untere Rauschband stärker auswirkt, liegen die Hörschwellen im Falle $f_0 < f_c$ über den Hörschwellen der konträren Bedingung.

Bei Verwendung der roex(p) Funktion (Gl. 2.9.) als Beschreibung für $W(g)$ wird im Algorithmus zur Bestimmung der unsymmetrischen auditiven Filterformen zwischen p_l für die tieffrequente Filterflanke (Gl. 2.7. erstes Integral) und p_u für die hochfrequente Filterflanke (Gl. 2.7. zweites Integral) unterschieden. Diese Werte werden in einer iterativen Prozedur solange verändert, bis die berechneten Hörschwellen um einen minimalen mittleren quadratischen Fehler von den gemessenen Hörschwellen abweichen. Dabei werden die Mittenfrequenzen der auditiven Filter an den Frequenzen angenommen, die das höchste SNR liefern.²

Durch die numerische Lösung der Integrale aus Gl. 2.7. erhöht sich die Flexibilität des Algorithmus, d.h. die von Glasberg und Moore³ vorgeschlagene Berechnung läßt sich nicht nur auf Notched-Noise-Daten anwenden⁴. Grundsätzlich kann jedes Maskierer-Spektrum verwendet werden, vorausgesetzt es läßt sich spezifizieren und in Gl. 2.7. substituieren. Dabei sollten natürlich nur solche Maskierer verwendet werden, die den Effekt des off-frequency listenings weitgehend vermeiden und passende Informationen für die Berechnung der auditiven Filter zur Verfügung stellen⁵.

¹ [Moore, 1993]

² Diese Verschiebungen sind für gewöhnlich klein. Um jedoch Anomalien der berechneten Filter zu vermeiden (tieffrequenter Bereich, hörgeschädigte Probanden), schlagen [Glasberg, Moore, 1990] eine Limitierung dieser Verschiebung auf max. 0.1 oder 0.2 g im Algorithmus vor.

³ [Glasberg, Moore 1990]

⁴ Beispielsweise können mit diesem Verfahren auch aus Datensätzen der Rippled-Noise Methode die Formen der auditiven Filter bestimmt werden.

⁵ vgl. [Glasberg, Moore, 1990]

2.5.1.3. Die roex - Funktionen

Die beschriebene Formbestimmung der auditiven Filter erfordert eine Annahme über die zu erwartenden Filterformen (Substitution von $W(g)$ in Gl. 2.7.). Patterson¹ schlägt dafür eine einfache Gruppe von Funktionen vor. Diesen sogenannten roex-Funktionen ist die abgerundete Spitze und der exponentielle Verlauf gemeinsam, woraus sich ihr Name ableitet (*r*ounded *e*xponential). Üblicherweise werden die roex-Funktionen als Funktion der normierten Frequenzvariable g angegeben. Die Wahl des Funktionstyps ist insofern von wesentlicher Bedeutung, da die zu bestimmenden auditiven Filter nur diejenigen Formen annehmen können, die mit diesen Funktionen vorgegeben werden. Die vorhandenen auditiven Filterformen leiten sich also nicht nur aus psychophysikalischen Versuchsreihen ab.

Bei der einfachsten roex-Funktion, der roex(p) Funktion (Gl. 2.9.) werden die auditiven Filterformen durch Änderung eines einzelnen Parameters approximiert. Dieser Parameter (p) gibt sowohl die Bandbreite als auch die Form der Flanken an. Je größer der Wert von p ist, desto schärfer ist der Filter. In Abb. 2.12. sind roex(p) Filterformen für verschiedene Werte von p dargestellt. Die Ordinatenkalierung bezieht sich auf die Leistungsdämpfung².

Bei der roex(p,r) Funktion (Abb. 2.13.)

$$W(g) = (1-r)(1+pg)e^{-pg} + r \quad \text{Gl. 2.10.}$$

als Erweiterung der roex(p)-Funktion wird ebenfalls zwischen p_l und p_u für die tieffrequente und hochfrequente Filterflanke unterschieden. Der Parameter r , üblicherweise für beide Flanken gleich, flacht die äußeren Flankenteile ab und stellt damit eine

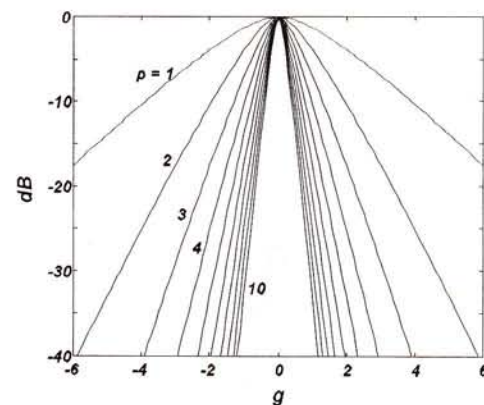


Abb. 2.12.: roex(p)-Funktion für $p = 1$ bis 10

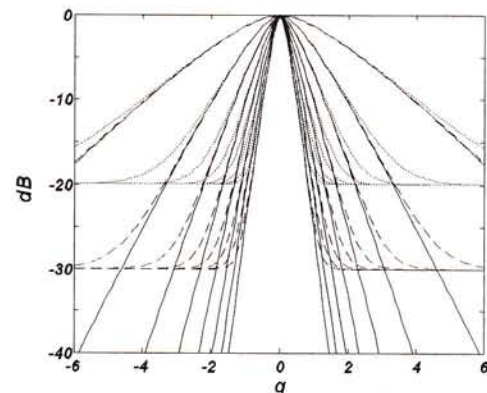


Abb. 2.13.: roex(p,r)-Funktion für $p=1$ bis 8 ($r=0$: durchgezogene Linien, $r=0.001$: strichlierte Linien und $r=0.01$ punktierte Linien)

¹ [Patterson, et al., 1982]

² In frühen Arbeiten wurden die Filterformen mit $|H(g)|^2$ bezeichnet, da sich die Leistungsdämpfung auf das Quadrat des Amplitudenganges bezieht. Aus dem Zusammenhang mit den Leistungsdichtespektrum-Modell bzw. der Notched-Noise-Methode ist jedoch klar, daß es sich ausschließlich um Leistungsbetrachtungen handelt. Aus diesem Grund werden die roex-Funktionen der Einfachheit halber als $W(g)$ bezeichnet und als Leistungsgrößen interpretiert. Für eine Darstellung in dB ergibt sich daher der zehnfache Zehnerlogarithmus (vgl. [Patterson, et al., 1982]).

Dynamikeinschränkung des Filters dar. Diese Dynamikeinschränkung wird in einigen Arbeiten mit der Ruhehörschwelle in Verbindung gebracht.¹

Die $roex(p,w,t)$ -Funktion (Abb. 2.14. und 2.15.) beschreibt das Verhalten des Filters außerhalb des Durchlaßbandes durch einen zusätzlichen Term²:

$$W(g) = (1-w)(1+pg)e^{-pg} + w(1+tg)e^{-tg} \quad \text{Gl. 2.11.}$$

Der Parameter p entspricht wiederum der Filtersteilheit, t bestimmt die Steigung der Flankenspitzen, und w gibt die Lage des Knickpunktes an. Für $t=0$ gleicht diese Form der $roex(p,r)$ -Funktion mit $w = r$.

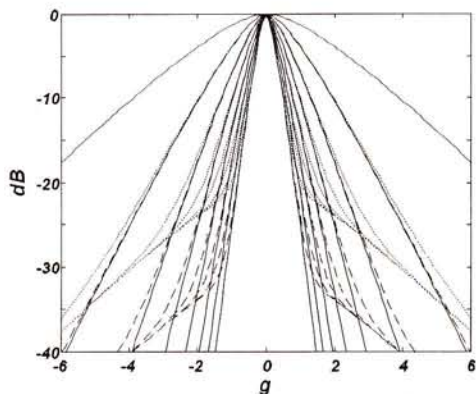


Abb. 2.14.: $roex(p,w,t)$ -Funktion für $p=1$ bis 8 und 3 unterschiedlichen w -Parametern: $w=0$ durchgezogene Linien, $w=0.001$ strichlierte Linien und $w=0.01$ punktierte Linien. Der Parameter t ist konstant ($t = 1$)

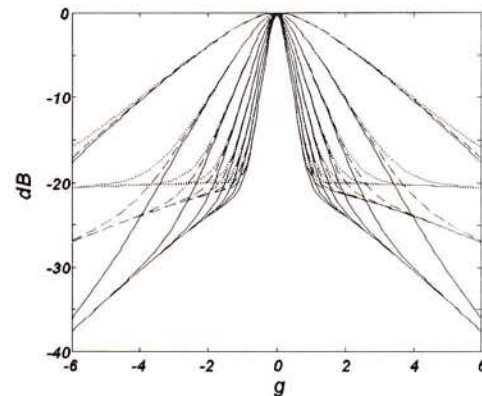


Abb. 2.15.: $roex(p,w,t)$ -Funktion für $p=1$ bis 8 und 3 unterschiedlichen t Parametern: $t=1$ durchgezogene Linien, $t=0.5$ strichlierte Linien und $t=0.1$ punktierte Linien. Der Parameter w ist konstant ($w = 0.01$)

Mit dieser Funktion können die psychoakustischen Datensätze am genauesten approximiert werden. Der Nachteil dabei ist die größere Komplexität. Die $roex(p)$ und $roex(p,r)$ - Funktionen werden zumeist als ausreichend erachtet³.

¹ [Nielsen, 1993]

² [Patterson, 1982], [Zhou, 1995]

³ [Moore, 1995] S.161-206

2.5.1.4. Bestimmung der auditiven Filterformen mit variablen Maskierern¹

Bei der bisherigen Betrachtungsweise wurde eine auditive Filterform mit einer konstanten Leistungsdichte N_0 des Maskierers durch Änderung des Signalpegels bestimmt. Auf dieser Methode basieren beinahe alle Arbeiten². Das Problem bei dieser Methode ist, daß der berechnete auditive Filter d.h. eine einzelne Filterform, unterschiedliche Leistungen des zu detektierenden Sinustons beinhaltet. Da der Sinuston innerhalb der Rauscheinkehlung den überwiegenden Leistungsanteil liefert, ist es naheliegend, eine einzelne auditive Filterform mit einer konstanten Signalleistung und unterschiedlichen Maskierungsleistungsdichten zu bestimmen. Diese Methode ist in [Rosen, Baker, 1994] beschrieben.

Für symmetrische Einkerbungen, eine Substitution von $W(g)$ durch die $roex(p)$ Funktion und eine Darstellung von K und P_s in dB, läßt sich Gl. 2.8. umschreiben zu:

$$P_s = K + N_0 + 10 \log_{10}(2f_0) - 10 \log_{10}(p) + 10 \log_{10}[(2 + pg)e^{-pg}] \quad \text{Gl. 2.12.}$$

Die zu approximierenden Parameter aus Gl. 2.12. werden durch Funktionen (Polynome) der Rauschleistungsdichte N_0 ersetzt. Für Polynome 1. Ordnung wird aus Gl. 2.12.:

$$P_s = (K_0 + K_1 N_0) + N_0 + 10 \log_{10}(2f_0) - 10 \log_{10}(p_0 + p_1 N_0) + 10 \log_{10}[(2 + (p_0 + p_1 N_0)g)e^{-(p_0 + p_1 N_0)g}] \quad \text{Gl. 2.13.}$$

Im Unterschied zu einer direkten Approximation der Parameterwerte p und K werden die entsprechenden Polynomparameter approximiert. Daraus ergibt sich direkt die Abhängigkeit dieser Parameterwerte von der Rauschleistungsdichte N_0 . Die iterative Anpassungsprozedur bleibt grundsätzlich unverändert, wobei innerhalb eines Durchlaufes die Meßdaten sowohl basierend auf fixen Signalpegeln als auch auf fixen Maskiererpegeln verwendet werden können. Bei der ursprünglichen Methode wird die Anpassung nur für eine konstante Rauschleistungsdichte durchgeführt.

Der Nachteil dieses Modells besteht darin, daß in Abhängigkeit der Polynomordnung relativ viele Parameter notwendig sind, um die gemessenen Daten zu approximieren. Rosen und Baker konnten zeigen, daß für ein Testsignal an 2 kHz die auditiven Filter mit Polynomen 1. Ordnung approximiert werden können, ohne die Anpassungsgenauigkeit wesentlich zu verringern.

¹ [Rosen, Baker, 1994]

² [Leek, Summers, 1993], [Lufti, Patterson, 1984], [Glasberg et al., 1983], [Glasberg et al., 1984], [Glasberg, Moore, 1986], [Glasberg, Moore, 1990], [Moore, Glasberg, 1983], [Moore, Glasberg, 1987], [Moore et al., 1990], [Moore et al., 1995], [Moore et al., 1996], [Patterson, 1976], [Patterson, Nimmo-Smith, 1980], [Patterson, Moore, 1986], [Shailer et al., 1990], [Tyler et al., 1984], [Wright, 1996], [Zhou, 1995]

2.5.2. Die 'Rippled-Noise'-Methode

Bei dieser Methode wird im Unterschied zur Notched-Noise-Methode sogenannter 'rippled-noise' (kammgefiltertes weißes Rauschen) als Maskierer verwendet. Eine Möglichkeit dieses Signal zu erzeugen besteht darin, zu weißem Rauschen dasselbe um τ verzögerte Rauschen zu addieren. Die spektralen Spitzen dieses Signals besitzen dadurch einen Abstand von $1/\tau$ Hz. Addiert man das verzögerte Signal in Phase, so befindet sich das erste Maximum bei 0 Hz. Dieses Rauschen wird 'cosine+' oder 'normales' Rauschen genannt. Dreht man die Polarität des verzögerten Rauschens um, so befindet sich das erste Maximum bei $0.5/\tau$ Hz. Dieses Signal wird 'cosine-' oder 'invertiertes' Rauschen genannt (Abb. 2.16.). Der Faktor τ bestimmt die Anzahl der innerhalb der halben Abtastrate befindlichen Maxima und wird deshalb als Welligkeitsdichtefaktor bezeichnet. Dieser Faktor stellt die unabhängige Variable in der psychoakustischen Versuchsreihe dar.

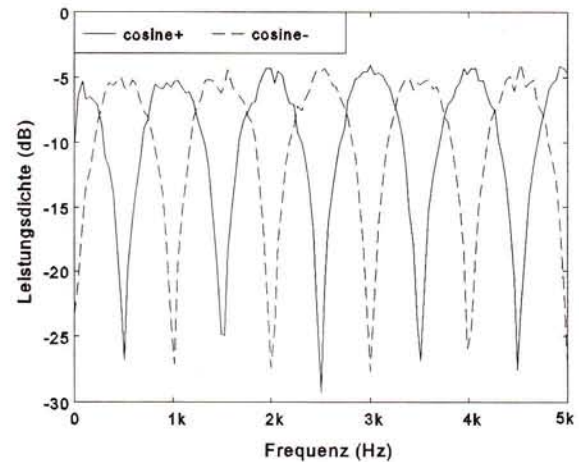


Abb. 2.16.: Kammgefiltertes weißes Rauschen, Länge 1 s, Abtastrate 10kHz, Zeitverzögerung $T = 1/1000$ s, (Welch-Periodogramm mit Hanning-Fenster, Fensterlänge 256 Punkte, Overlap: 1/2)

Das Leistungsdichtespektrum $I(f)$ (Abb. 2.16) ist gegeben mit:

$$I(f) = I_m (1 \pm m \cos(2\pi f \tau)) \quad \text{Gl. 2.14.}$$

mit I_m als der mittleren Leistungsdichte und m als der Modulationstiefe des Rauschens (typischer Wert: $m = 0.998$)¹. Bei dieser Methode wird bei konstantem τ der Hörschwellenunterschied zwischen normalem und invertiertem Maskierer bestimmt. Dieser Vergleich wird für mehrere τ durchgeführt (Je größer der Faktor τ , desto mehr Maxima befinden sich innerhalb der halben Abtastrate). Das sinusförmige Testsignal befindet sich dabei stets an einer festen Frequenz. Durch das Zusammenrücken der Maxima im Frequenzbereich verschwindet die Differenz zwischen normalem und invertiertem Rauschen mit steigendem τ . Die auditive Filterform erhält man entweder durch Approximation der auditiven Filter als Fourier-Reihe² oder mit der in Kap. 2.5.1. beschriebenen iterativen Anpassung.³

Diese Filterformen unterscheiden sich von den Notched-Noise Filtern durch einen etwas

¹ [Pick, 1980]

² [Houtgast, 1977], [Pick, 1980]

³ [Glasberg, Moore, 1990]

breiteren Verlauf und einer etwas flacheren Spitze¹. Mit dieser Methode ist es nicht möglich, die Formen von auditiven Filtern über einen größeren Dynamikbereich zu bestimmen. Pick² beschränkt sich beispielsweise auf einen Dynamikbereich von 10 dB.

2.5.3. Psychophysikalische Abstimmkurven

Eine weitere Möglichkeit, die Frequenzselektivität des menschlichen Gehörs zu beschreiben, stellen die sogenannten psychophysikalischen Abstimmkurven (Psychophysical Tuning Curves)³ dar.

Dabei werden die Frequenz und der Pegel eines Sinustons konstant gehalten. In einer Versuchsreihe wird derjenige Schmalbandmaskierpegel (in einigen wenigen Fällen werden Sinus-Maskierer verwendet) als Funktion der Mittenfrequenz des Maskierers bestimmt, der gerade noch Maskierung erzeugt. Ein derartiges Beispiel ist in Abb. 2.17. dargestellt. Der Signalpegel wird klein gewählt (z.B.: 10 dB), sodaß angenommen werden kann, daß nur ein einzelner auditiver Filter (zentriert um die Signalfrequenz) Verwendung findet. Die Abstimmkurve kann also interpretiert werden als diejenige Leistung des Maskierers, die an diesem einen auditiven Filter stets eine konstante Ausgangsleistung erzeugt.

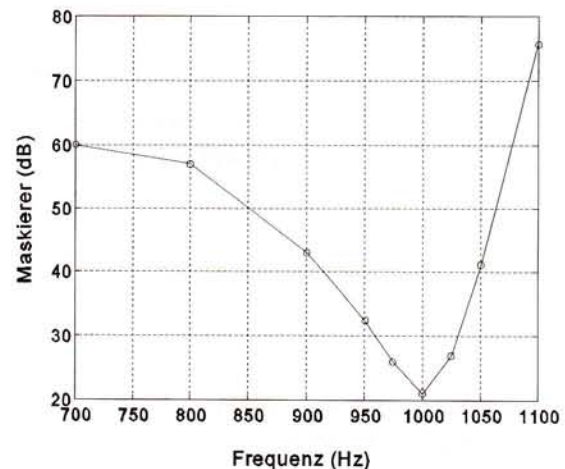


Abb 2.17.: Psychophysikalische Abstimmkurve für eine Signalfrequenz von 1 kHz (vgl. [Moore, 1995] S.168)

Das off-frequency-listening läßt sich dadurch vermeiden, daß ein zusätzliches Breitbandrauschen mit kleinem Pegel und einer Einkerbung an der Signalfrequenz hinzugefügt wird. Unter der Annahme von linearen auditiven Filtern ergibt sich die Form der Filter durch Invertierung der Abstimmkurven.⁴

Die Annahme, daß bei diesen Versuchen nur ein einzelner Filter vom auditiven System verwendet wird, ist kritisch. Befindet sich der Maskierer beispielsweise über dem Sinussignal, so ist es wahrscheinlich, daß der verwendete auditive Filter (sofern es nur ein Filter ist) sich zu tieferen Frequenzen hin verschiebt. Es gibt starke Anzeichen dafür, daß das auditive System, sofern sich die Möglichkeit bietet, auf das off-frequency-listening zurückgreift.⁵

¹ [Moore, 1995] S.172

² [Pick, 1980]

³ vgl. [Moore, 1989] S.96-98, [Moore, 1995] S.166-168

⁴ [Moore, 1995]

⁵ vgl. [Moore, 1989] S. 98

2.5.4. Neuronale Abstimmkurven¹

Die neuronalen Abstimmkurven (Neural Tuning Curves) entsprechen in ihrer Form etwa den psychophysikalischen Abstimmkurven. Sie werden so bestimmt, daß derjenige Pegel eines Testsignals (z.B.: Sinuston) an unterschiedlichen Frequenzen gesucht wird, der an einer einzelnen Faser des Hörnervs ein konstantes Signal erzeugt. Die neuronalen und psychophysikalischen Abstimmkurven lassen sich jedoch insofern nicht vergleichen, da zur Bestimmung der neuronalen Abstimmkurven nur ein Testsignal präsent ist, während bei der Bestimmung der psychophysikalischen Abstimmkurven das auditive System mit zwei Signalen erregt wird. Weiters beziehen sich die neuronalen Abstimmkurven auf eine einzelne Nervenfasern, während die psychophysikalischen Kurven eine Gruppe Nervenfasern zusammenfassen. In Hinblick auf eine Modellierung des menschlichen Gehörs mit einer parallelen Bandpaßfilterbank, deren Amplitudengänge die Form der auditiven Filter nachbilden, erscheinen die neuronalen Abstimmkurven von geringerer Relevanz, da sie einzelne Nervenfasern berücksichtigen und eine parallele Filterbank demgegenüber jeweils über einen bestimmten Frequenzbereich bzw. über eine Gruppe von Nervenfasern integriert.

2.5.5. Frequenzkorrekturen

Entsprechend der Annahme, daß auditive Filter sich aus Prozessen der Cochlea sowie der anschließenden neuronalen Verarbeitung erklären lassen, werden die den Versuchspersonen dargebotenen Schalle entsprechend den Amplitudengängen des Kopfhörers (sofern die Versuchsschalle über Kopfhörer dargeboten werden) sowie des Außen- und Mittelohres entzerrt, um die Cochlea mit einem ungewichteten Amplitudengang zu erregen.

Kopfhörer-Entzerrung

Kopfhörer-Entzerrungsfunktionen werden entweder mit psychoakustischen Versuchsreihen oder mit Miniatur-Meßmikrofonen ermittelt. Diese Übertragungsfunktionen stellen einen eindeutigen Zusammenhang zwischen dem Schalldruck eines ebenen ungestörten Schallfeldes und der elektrischen Größe am Eingang des zu kalibrierenden Kopfhörers dar.

Das Verfahren nach DIN 45 619 Teil 1 besteht aus einem Vergleich der Lautstärken, die abwechselnd von einer senkrecht von vorne einfallenden ebenen Schallwelle und von einem Kopfhörer erzeugt werden. Die elektrische Erregung am Kopfhörer wird dabei so eingestellt, daß beide Schalle (Freifeld - Kopfhörer) als gleich laut empfunden werden². Als Testsignale werden Sinus-Töne oder Terzbandrauschen verwendet.

Eine weitere Möglichkeit ergibt sich aus der Messung der Ruhehörschwelle. Dabei wird die Freifeld-Übertragungsfunktion aus der Differenz zwischen der gemessenen Ruhehörschwelle und

¹ [Moore, 1995] S.75-79

² [Fastl, Zwicker, 1983], [DIN 45619]

der nach ISO 266 genormten Ruhehörschwelle bestimmt¹. Diese Methode ist nicht allgemein anerkannt.

Neben diesen beiden Methoden, die sich auf psychoakustische Versuchsreihen beziehen, besteht die Möglichkeit, Freifeld-Übertragungsfunktionen mit Miniatur-Meßmikrofonen zu bestimmen, die entweder am Eingang des Gehörganges oder im Gehörgang selbst plaziert werden.

Außen- und Mittelohrentzerrung

Unter der Annahme, daß die Cochlea für alle Frequenzen gleich empfindlich ist, kann die Ruhehörschwelle (vgl. Kap. 3.1.2.) als erste Näherung für die Übertragungsfunktion des Außen- und Mittelohres verwendet werden. Glasberg und Moore² schlagen dafür zwei Funktionen vor:

- MAF-Korrektur (minimum audible field), Ruhehörschwelle nach ISO R 226
- MAP-Korrektur (minimum audible pressure at the eardrum)³

Würde der Amplitudengang des Außen- und Mittelohres tatsächlich der Ruhehörschwelle entsprechen, so bliebe die Frage ungeklärt, warum sich die Kurven gleicher Lautstärkepegel (vgl. Kap. 3.1.2.), besonders im hohen Phon Bereich, von der Ruhehörschwelle unterscheiden. Eine mögliche Erklärung für dieses Phänomen läßt sich aus dem internen Rauschen der Cochlea ableiten. Dabei geht man davon aus, daß dieses interne Rauschen bei tiefen Frequenzen stärker ausgeprägt ist als bei hohen Frequenzen⁴. Unter dieser Voraussetzung ergibt sich bei tiefen Frequenzen (unter 1 kHz) sowie bei größeren Schallpegeln ein flacherer Phon-Kurvenverlauf als an der Ruhehörschwelle. Demzufolge würden höhere Phon-Kurven den Amplitudengängen des Außen- und Mittelohres besser entsprechen. Dafür schlagen Glasberg und Moore eine dritte Korrektur vor:

- ELC-Korrektur (equal loudness contour)⁵, inverse 100 phon Kurve nach ISO R 226

Da in der ISO Norm für 25 Hz, 10 kHz und 12,5 kHz keine Angaben über die 100 phon Kurve vorhanden sind, verwendet Agerkvist⁶ die inverse 90 Phon Kurve.

Die nach ISO 226 genormten Phon-Kurven beziehen sich auf eine Versuchsreihe von Robinson und Dadson aus dem Jahre 1956⁷. Neuere Versuchsreihen zur Bestimmung der Phon-Kurven weichen jedoch wesentlich von diesen Kurven ab. Aus diesem Grund ist eine Revision des vorhandenen Standards seit einigen Jahren in Arbeit, die große Varianz neuester Daten

¹ [Moore et al., 1990]

² [Glasberg, Moore, 1990]

³ [Killion, 1978], [Green et al., 1987]

⁴ [Nedzelnitzky, 1980]

⁵ [Glasberg, Moore, 1990]

⁶ [Agerkvist, 1994]

⁷ Die nach ISO genormten Phon-Kurven werden häufig als Fletcher-Munson Kurven bezeichnet. Die Phon-Kurven von Fletcher und Munson (1933) entsprechen nicht den genormten Phon-Kurven, womit diese Bezeichnung falsch ist.

verhinderte aber zur Zeit eine Revision dieser Norm. Einzig die Form der Ruhehörschwelle und die wesentlichen Abweichungen vom vorhandenen Standard sind den neueren Versuchsreihen gemeinsam¹ (vgl. Kap. 3.2.).

Die bis dato veröffentlichten auditiven Filter basieren auf unterschiedlichen Außen- und Mittelohrentzerrungen. Glasberg und Moore verwenden beispielsweise unter 1 kHz die ELC-Korrektur, über 1 kHz die MAF- oder MAP-Korrektur². Zhou³ verwendet demgegenüber unter 8 kHz die MAP-Korrektur und über 8 kHz die mittlere Ruhehörschwelle seiner Probanden. Diese unterschiedlichen Korrekturannahmen wirken sich besonders im tiefen und hohen Frequenzbereich aus. Bei mittleren Frequenzen sind die drei Korrekturen ähnlich.

Für die Implementation der aus psychoakustischen Versuchsreihen zur Verfügung gestellten auditiven Filter ist es also notwendig, die auditiven Filter an die jeweils verwendete Außen- und Mittelohrübertragungsfunktion anzupassen.

2.6. Eigenschaften auditiver Filter

2.6.1. Energieäquivalente Rauschbandbreite

Die ERB-Breiten (equivalent rectangular bandwidth) sind die energieäquivalenten rechteckigen Bandbreiten der auditiven Filter. Sie entsprechen qualitativ den kritischen Bandbreiten (CB-Breiten, vgl. Kap. 2.3.), unterscheiden sich jedoch quantitativ von diesen. Da in der Literatur für die Berechnung der ERB-Breiten und ERB-Raten unterschiedliche Angaben vorhanden sind, werden diese Funktionen im folgenden zusammengefaßt.

1983 wurde die ERB-Breite (in Hz) als Funktion der auditiven Filter-Mittenfrequenz mit

$$ERB(f) = 6.23 \left(\frac{f}{1000} \right)^2 + 93.39 \left(\frac{f}{1000} \right) + 28.52 \quad \text{Gl. 2.15.}$$

angenommen⁴. Diese Funktion bezieht sich auf symmetrische auditive Filter, sie besitzt daher nur für mittlere Schallpegel Gültigkeit. Gl. 2.15. wurde 1990 unter Berücksichtigung unsymmetrischer auditiver Filter auf

$$ERB(f) = 24.7 (0,00437 f + 1) \quad \text{Gl. 2.16.}$$

korrigiert⁵.

¹ [Suzuki, Sone, 1993], [Gabriel et al., 1993], [Fastl et al., 1990], [ISO 226, 1987], [Sørensen, 1994]

² [Glasberg, Moore, 1990]

³ [Zhou, 1995]

⁴ [Moore, Glasberg, 1983]

⁵ [Glasberg, Moore, 1990]

Die ERB-Breiten sind über den gesamten Frequenzbereich schmaler als die CB-Breiten. Besonders stark ist dieser Unterschied im Bereich unter 500 Hz ausgeprägt (Abb. 2.18.). In diesem Bereich sind die CB-Breiten konstant 100 Hz, während die ERB-Breiten deutlich darunter liegen. Die Validität der CB-Breiten unter 500 Hz ist anzuzweifeln, da diese Daten auf einigen wenigen psychoakustischen Versuchen beruhen¹.

In Anlehnung an die Frequenzgruppenskala läßt sich durch Integration der reziproken ERB-Breiten-Funktionen eine Transformation vom linearen Frequenzbereich (in Hz) in einen gehörbezogenen Frequenzbereich angeben (vgl. Bark-Transformation Gl. 2.2.).

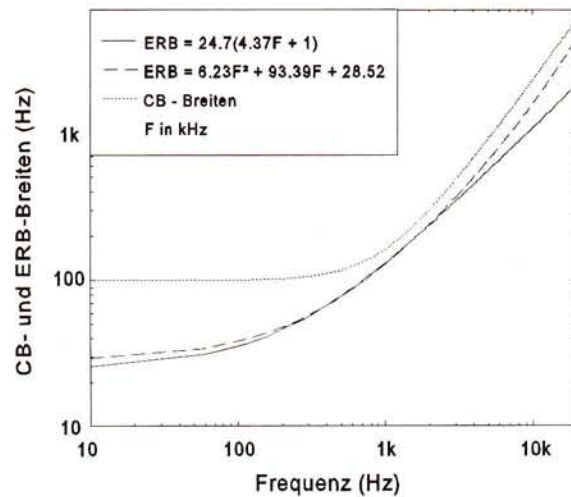


Abb. 2.18.: Vergleich zwischen den CB-Breiten (Gl. 2.1.) und den ERB-Breiten (Gl. 2.15., 2.16.)

Diese Skala entspricht qualitativ der Bark-Skala und wird in ERB-Raten, E oder e skaliert. Gleich der Frequenzgruppenskala läßt sie sich als 'Abstand entlang der Basilarmembran' interpretieren. Die ERB-Rate auf Basis der ERB-Breiten aus dem Jahre 1983 (Gl. 2.15.) wird angegeben als²:

$$e(f) = 11.17 \ln \left| \frac{f+312}{f+14675} \right| + 43 \quad \text{Gl. 2.17.}$$

Die Konstante des unbestimmten Integrals ist so gewählt, daß die ERB-Rate an $f = 0$ Hz den Wert $e = 0$ annimmt ($C = 43$). Gl. 2.17. ist im Frequenzbereich zwischen 0,1 und 6,5 kHz gültig. Die ERB-Rate aus Gl. 2.16. errechnet sich analog (Gl. 2.18.a)³:

$$\begin{aligned} e(f) &= 9.264 \ln(0.108f + 24.7) - 29.7 \\ e(f) &= 9.264 \ln(0.00437f + 1) \\ e(f) &= 21.4 \log_{10}(0.00437f + 1) \end{aligned} \quad \text{Gl. 2.18. a,b,c}$$

¹ [Moore, 1993]

² [Moore, Glasberg, 1983]

³ Gl. 2.18b ist eine Umformung von Gl. 2.18a bei der die Konstante des unbestimmten Integrals in den Logarithmus vorgezogen wurde. Gl. 2.18.c (Angabe im Zehnerlogarithmus) als weitere Ableitung von Gl. 2.18.a weicht mit steigenden ERB-Raten von Gl. 2.18.a ab (max. Abweichung: 0.13 ERB-Raten bei 20 kHz) ([Glasberg, Moore, 1990]).

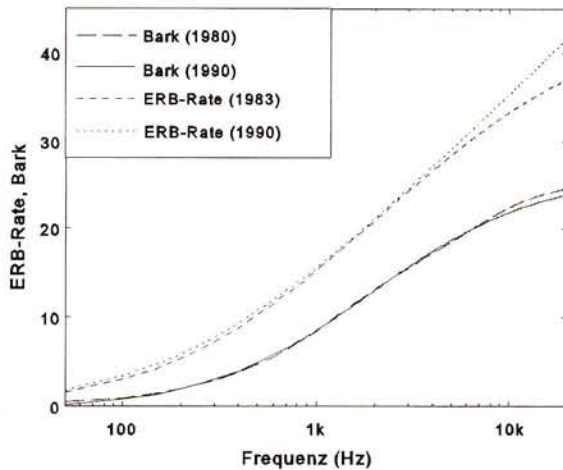


Abb. 2.19.: Frequenz-Bark-Transformationen (Gl. 2.2., 2.3.) und Frequenz-ERB-Raten-Transformationen (Gl. 2.17., 2.18a.)

In Abb. 2.19. sind die beiden ERB-Raten Transformationen (Gl. 2.17, 2.18a) im Vergleich zu den in Kap. 2.3. angegebenen Bark-Transformationen (Gl. 2.2., 2.3.) dargestellt. Die ERB-Raten attestieren, entsprechend den geringeren ERB-Breiten, dem Gehör eine bessere Frequenzauflösung. Die Bark-Skala teilt den hörbaren Frequenzbereich in 24 Bark, die ERB-Raten-Skala in etwa 40 ERB-Breiten ein.

2.6.2. Auditive Filterformen

Im Bereich der psychoakustischen (d.h. experimentalpsychologischen) Versuchsreihen ist die Notched-Noise Methode am weitesten entwickelt. Aus diesem Grund werden im folgenden nur diejenigen auditiven Filter betrachtet, die auf Notched-Noise Experimenten beruhen. Diese Experimente basieren auf zwei unterschiedlichen Annahmen und führen zu unterschiedlichen Ergebnissen. Die erste Annahme setzt ein konstantes Leistungsdichtespektrum des Maskierers bei unterschiedlichen Einkerbungsbreiten voraus ('fixed-masker paradigm', vgl. Kap. 2.5.1.1. bis 2.5.1.3.). Bei der zweiten Methode wird der Pegel des Testsignales bei unterschiedlichen Einkerbungsbreiten konstant gehalten ('fixed-probe paradigm', vgl. Kap. 2.5.1.4.)¹. Diese beiden Annahmen führen sowohl zu unterschiedlichen Filterformen als auch zu unterschiedlichen Pegelabhängigkeiten. Sie sind in den Kap. 2.6.2.1. und 2.6.2.2. getrennt dargestellt. Alle dargestellten Filter sind insofern vereinfacht, da die Pegelabhängigkeiten der hochfrequenten Flanken nicht berücksichtigt werden.

2.6.2.1. Auditive Filterformen bei konstanten Maskiererpegeln

- Ein auditiver Filter an einer Mittenfrequenz von $f_c = 1$ kHz ist auf einer linearen Frequenzachse näherungsweise symmetrisch, wenn der Pegel des Eingangssignals etwa 51 dB/ERB beträgt. Dies entspricht einem Rausch-Leistungsdichtespektrum (weißes Rauschen) von etwa 30 dB im Durchlaßband des Filters. Die auditiven Filter anderer Mittenfrequenzen sind dann symmetrisch, wenn der Eingangspegel dieser Filter dem Pegel von 51dB/ERB an 1 kHz entspricht. Der äquivalente Eingangsrauschpegel (dB/ERB) berechnet sich nach der Außen- und

¹ [Rosen, Baker, 1994]

Mittelohrkorrektur¹.

Unter Verwendung der roex(p) Funktion wird der Parameter p als Funktion der Filtermittenfrequenz angegeben mit:

$$p(f_c) = \frac{4f_c}{ERB(f_c)} = \frac{4f_c}{24.7(0.0043f_c + 1)} \quad \text{Gl. 2.19.}$$

Diese Filterform entspricht einem linearen auditiven Filter ($p_l = p_u = p$). Der Gültigkeitsbereich erstreckt sich über mittlere Schallpegel.

- Mit steigendem Pegel wird die tieffrequente Flanke der auditiven Filter flacher. Dies wird als der primäre Effekt der Pegelabhängigkeit angesehen, da die Änderungen der oberen Flanken sowohl weniger stark ausgeprägt als auch nicht konsistent sind. Die hochfrequente Filterflanke wird konstant auf $p_u(f_c) = p(f_c)$ gesetzt. Die Pegelabhängigkeit der tieffrequenten Filterflanken wird ausgedrückt durch²:

$$p_l(f_c) = p(f_c) \left(1 - \frac{0.38}{p(1\text{kHz})} (L_{ERB} - 51\text{dB}) \right) \quad \text{Gl. 2.20.}$$

$p(1\text{kHz})$ entspricht dem Parameter p für $f_c = 1\text{ kHz}$ und einer äquivalenten Rauschbandbreite von 51 dB/ERB. L_{ERB} entspricht dem äquivalenten Eingangsruschpegel in dB/ERB an der jeweiligen Filtermittenfrequenz.

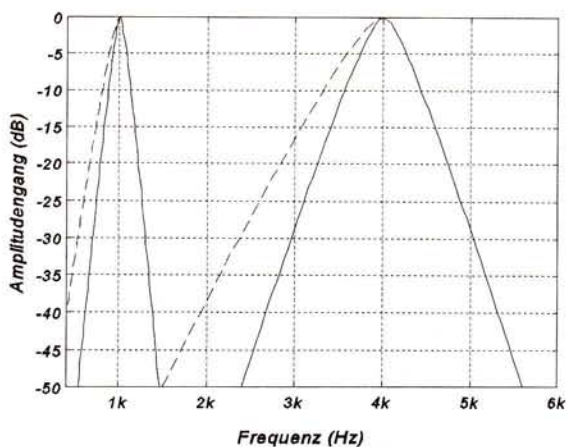


Abb. 2.20.: Auditive roex(p)-Filter an den Mittenfrequenzen $f_c = 1\text{ kHz}$ und 4 kHz jeweils bei zwei unterschiedlichen äquivalenten Rauschpegeln: $L_{ERB} = 51\text{ dB/ERB}$ (—) und 80 dB/ERB (---), lineare Frequenzachse

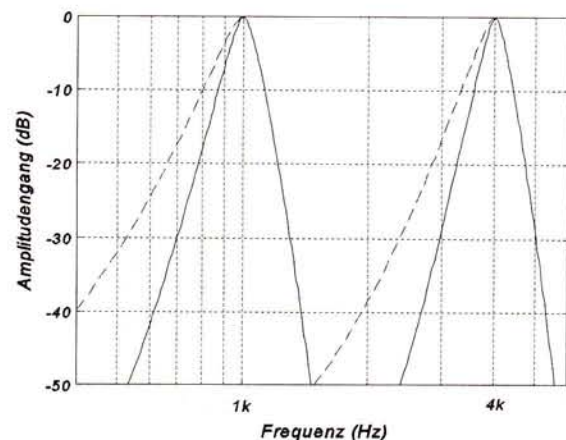


Abb. 2.21.: Auditive roex(p)-Filter an den Mittenfrequenzen $f_c = 1\text{ kHz}$ und 4 kHz jeweils bei zwei unterschiedlichen äquivalenten Rauschpegeln: $L_{ERB} = 51\text{ dB/ERB}$ (—) und 80 dB/ERB (---), log. Frequenzachse

¹ vgl. [Moore, 1993]

² [Avergvist, 1994], [Glasberg, Moore, 1990]

In Abb. 2.20. und 2.21. sind pegelabhängige Amplitudengänge entsprechend Gl. 2.19. und 2.20. für $f_c = 1\text{k}$ und 4kHz und zwei unterschiedliche äquivalente Rauschpegel ($L_{\text{ERB}} = 51\text{ dB/ERB}$ (—) und 80 dB/ERB (---)) dargestellt.

Bei der Bestimmung von Erregungsmustern aus pegelabhängigen auditiven Filterbänken ist zur Zeit noch unklar, aus welchem Signal sich der Steuerungspegel für die nichtlineare Steuerung der tieffrequenten Filterflanken errechnet. Unter der Voraussetzung, daß der Steuerungspegel für einen Frequenzkanal d.h. für eine Filtermittenfrequenz aus diesem Kanal ausgekoppelt werden kann und kanalübergreifende Steuerungspegel ausgeschlossen werden, läßt sich diese Frage auf die Notched-Noise-Methode zurückführen.

Ein fixer Maskierer-Pegel bei der Notched-Noise-Methode bedingt die Pegelabhängigkeit vom Eingang der auditiven Filter. Auf dieser Annahme basiert die oben dargestellte nichtlineare Steuerung (Gl. 2.19. und 2.20.)¹. In einer vergleichenden Analyse zwischen eingangs- und ausgangspegelabhängigen Filtern konnten Moore und Glasberg² zeigen, daß bei Verwendung der roex(p)-Funktion ein eingangspegelabhängiges Modell den 'ubward spread of masking' genauer nachbildet als ein ausgangspegelabhängiges Modell. Als Alternative dazu ergibt sich die Möglichkeit, die Pegelabhängigkeit der Filterformen mit dem Ausgangssignal der Filter zu steuern³. Die Auskopplung des Steuerungssignals nach dem Filter bezieht sich auf Notched-Noise Experimente mit konstantem Testsignalpegel und variablen Maskiererpegeln.

2.6.2.2. Auditive Filterformen bei konstanten Testsignalpegeln⁴

Die nichtlinearen Änderungen eines Filters an 2 kHz werden als Funktion des Testsignalpegels P_s (in dB) für die roex(p,r) Funktion angegeben mit:

$$\begin{aligned} p_l &= 39 - 0,42P_s \\ p_u &= 27,1 \\ r_l &= 10^{-4,67 + 0,042P_s} \\ r_u &= 0 \end{aligned} \quad \text{Gl. 2.21.}$$

Gl. 2.21. ist für Testsignalpegel von $P_s = 30$ bis 60 dB gültig. Diese roex(p,r) Filterformen sind in Abb. 2.22. für $P_s = 30, 40, 50$ und 60 dB auf einer linearen Frequenzachse dargestellt. Die Abflachung des $P_s = 30\text{ dB}$ Filters unter $1,2\text{ kHz}$ wird als Artefakt der roex(p,r) Filterformen interpretiert. Die nichtlinearen Änderungen der hochfrequenten Filterflanken sind auch in diesem Fall nicht signifikant und werden vernachlässigt.

¹ [Lufti, Patterson, 1984]

² [Moore, Glasberg, 1987]

³ [Rosen, Baker, 1994]

⁴ [Rosen, Baker, 1994]

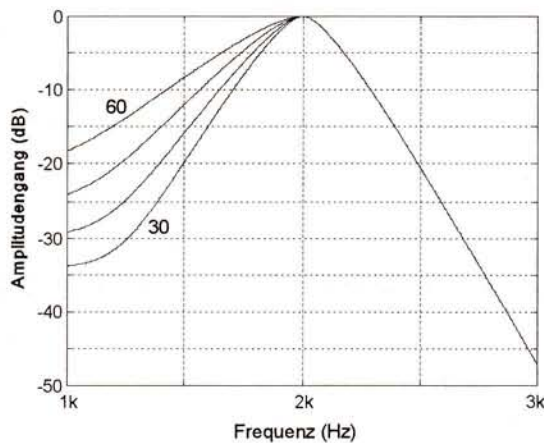


Abb. 2.22.: roex(p,r) Filterformen nach Gl. 2.21. mit $P_s = 30, 40, 50, 60$ dB

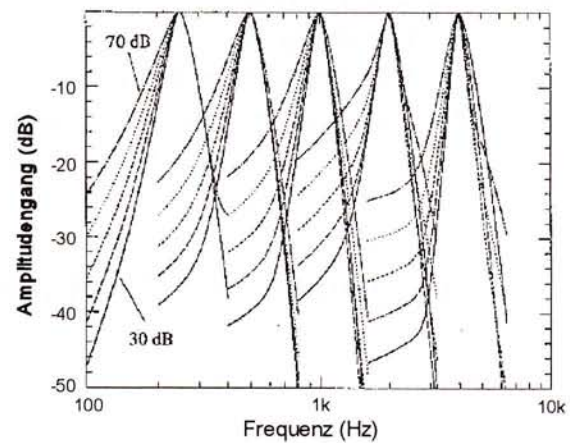


Abb. 2.23.: roex(p,w,t) Filterformen mit $P_s = 30, 40, 50, 60, 70$ dB an 250, 500, 1k, 2k, 4k Hz, (vgl. [Rosen et al., 1997])

Weiterführende Untersuchungen für unterschiedliche Filtermittenfrequenzen und einen größeren Testsignalpegelbereich ($P_s = 30 - 70$ dB) sind in Abb. 2.23. unter Verwendung der roex(p,w,t)-Funktionen angeführt¹. Diese Filterformen unterscheiden sich an $f_c = 2$ kHz von den in Abb. 2.22. dargestellten roex(p,r) Filtern. Die Abweichungen lassen sich auf die unterschiedlichen roex-Funktionen zurückführen, da die psychoakustischen Versuchsreihen unter den gleichen Bedingungen durchgeführt wurden.

Die mit konstanten Testsignalpegeln ermittelten auditiven Filter sind grundsätzlich breitbandiger als die in Kap. 2.6.2.1. dargestellten Filter. Zur Zeit liegen jedoch keine vollständigen Datensätze über den gesamten hörbaren Frequenzbereich vor.

2.7. Das Erregungsmuster

Unter dem Begriff des Erregungsmusters werden Frequenz- oder Zeit-Frequenzdarstellungen verstanden, die in einem Bezug zu den Übertragungseigenschaften des menschlichen Gehörs stehen. Aus der Bezeichnung 'Erregungs'-muster läßt sich ein Zusammenhang zu einer physiologisch motivierten Größe ableiten. Für physiologische Modelle (vgl. Kap. 4) trifft diese Bezeichnung zu. Bei funktionalen Modellen werden jedoch keine physiologischen Erregungsgrößen nachgebildet, sondern experimentalpsychologisch ermittelte Übertragungseigenschaften simuliert. Das physiologische Pendant eines funktionalen Erregungsmusters ist die Schwingungsamplitude der Basilarmembran des Innenohres. An dieser Stelle des peripheren Gehörs wird eine erste Zeit-Frequenzanalyse der eintreffenden Schallsignale durchgeführt. Die physiologische Analyse (Wanderwellen entlang der Basilarmembran, vgl. Kap. 2.2.) unterscheidet sich grundsätzlich von den Annahmen des Leistungsdichtespektrum-Modells (vgl. Kap. 2.4.) bzw. eines funktionalen Modells.

¹ [Rosen et al. 1996]

Für stationäre Schalle sind Frequenzdarstellungen des Erregungsmusters, die über einen längeren Zeitraum integriert werden, insofern ausreichend, als den Langzeitadaptionen des menschlichen Gehörs sekundäre Bedeutung zukommt. Im folgenden werden zwei unterschiedliche Arten von funktionalen Erregungsmustern angeführt: Erregungsmuster auf Basis der Maskierungsfunktionen und Erregungsmuster auf Basis der auditiven Filter.

2.7.1. Erregungsmuster auf Basis der Maskierungsfunktionen

Unter der Annahme, daß Erregungsmuster eine ähnliche Form wie Maskierungsmuster besitzen und nur um einige wenige dB abweichen, läßt sich das Erregungsmuster aus Maskierungsmustern bestimmen¹. Diese Annahme ist insofern problematisch, da die Bestimmung der Maskierungsmuster von Nebeneffekten, wie der Detektion von Kombinationstönen, Schwebungen und dem 'off-frequency listening', beeinflußt werden. Die Detektion von Kombinationstönen und Schwebungen kann von geübten Versuchspersonen zwar weitgehend ignoriert werden, der Effekt des off-frequency listenings jedoch nicht, da dieses Phänomen auf psychischer Ebene nicht beeinflußbar ist.

Die Ableitung des Erregungsmusters aus Maskierungsmustern beinhaltet daher eine Diskrepanz zu den Annahmen des Leistungsdichtespektrum-Modells und sollte vermieden werden.

2.7.2. Erregungsmuster auf Basis der auditiven Filter

Die auditiven Filter repräsentieren die Frequenzselektivität des menschlichen Gehörs. Durch die entwickelten Methoden zur Formbestimmung dieser Filter (vgl. Kap. 2.5.) wird die Detektion von Schwebungen zwischen Maskierer und Testsignal sowie von Kombinationsbändern weitgehend vermieden. Zusätzlich wird das off-frequency listening berücksichtigt. Diese Methoden entsprechen dadurch den Annahmen des Leistungsdichtespektrum-Modells. Die Berechnung des Erregungsmusters aus einer linearen d.h. pegelunabhängigen auditiven Filterbank ist in Abb. 2.24. dargestellt.

Nach einer Außen- und Mittelohrfilterung werden von den Ausgangssignalen der auditiven Filter die Einhüllenden gebildet und das interne Rauschen der Cochlea wird hinzuaddiert (vgl. Kap. 4 und 6). Das Zeitverhalten des Erregungsmusters bzw. die Einhüllenden der Impulsantworten der auditiven Filter (vgl. Kap. 6) können in der Verarbeitungsstufe 'Bestimmung der Einhüllenden' an das Gehör adaptiert werden. Voraussetzung dafür ist jedoch, daß die Impulsantworten der auditiven Filter kürzer sind als die Zeitauflösung des Gehörs, da bei der Bestimmung der Einhüllenden die Impulsantworten nur verlängert, nicht jedoch verkürzt werden können.

¹ [Zwicker, 1982] S. 112-120

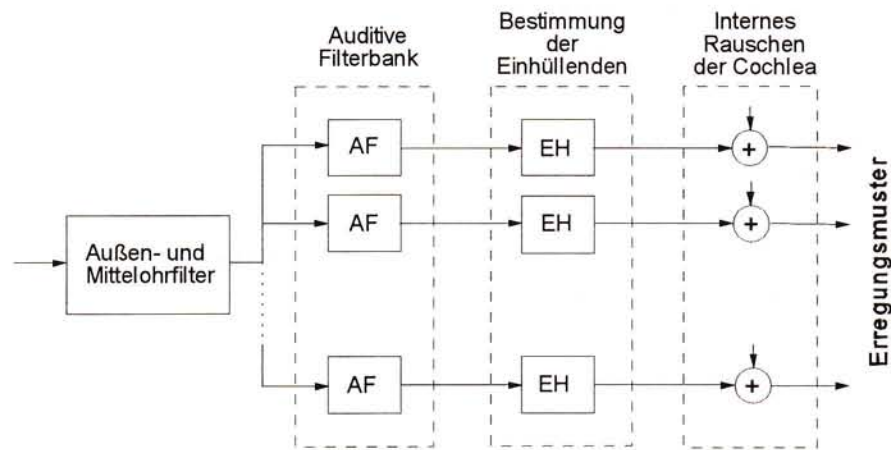


Abb. 2.24.: Berechnung des Erregungsmusters auf Basis der auditiven Filter

Abb. 2.25 zeigt ein Beispiel für eine Zeit-Frequenzdarstellung eines Erregungsmusters. Üblicherweise werden die Erregungsmuster in absoluten dB skaliert. Die Frequenzauflösung ergibt sich aus den auditiven Filterformen bzw. Amplitudengängen, den Abständen der Filtermittenfrequenzen und der Anzahl der verwendeten Filter. Eine genauere Betrachtung des Zeitverhaltens ist in den Kap. 5. und 6. angeführt.

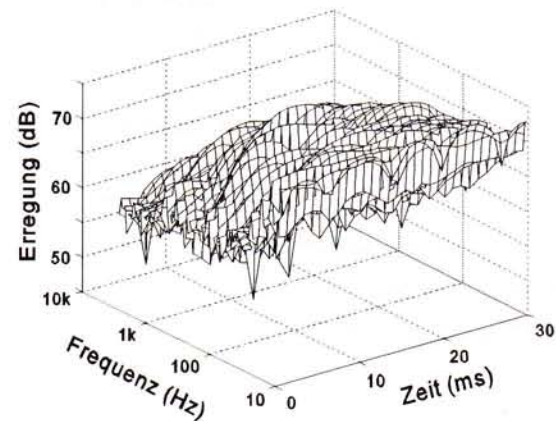


Abb. 2.25.: Erregungsmuster (Zeit-Frequenzdarstellung)

2.8. Zusammenfassung und Diskussion

Die Arbeit zur Modellierung der menschlichen Hörwahrnehmung teilt sich in zwei Teile. In die Analyse der Hörwahrnehmung und in die Entwicklung der daraus erhaltenen Erkenntnisse zu einem Modell (Synthese). In diesem Kapitel wurde primär die Analyse betrachtet. Die Beschäftigung mit den psychoakustischen Grundlagen der verwendeten Daten ist in Hinblick auf die Validität des angestrebten nichtlinearen funktionalen Modells des peripheren Gehörs von wesentlicher Bedeutung.

Den meisten auditiven Modellen ist die strukturelle Anlehnung an den anatomischen Aufbau des Gehörs gemeinsam (vgl. Kap. 4.). Das Erregungsmuster oder das neuronale Aktivitätsmuster bilden eine Analogie zu den physiologischen Erregungsgrößen der Cochlea und damit der Signalverarbeitung im peripheren Gehör. Die definierte Schnittstelle der Erregung

(hydromechanisch oder neuronal) ermöglicht eine getrennte Betrachtungsweise der beiden Komplexe 'periphere Schallverarbeitung' und 'zentrale Schallverarbeitung'.

Um die gewünschten Eigenschaften des Gehörs (z.B.: Lautheitsberechnung, Spracherkennung, etc.) nachzubilden, werden diese Erregungen weiteren Verarbeitungsstufen unterzogen. Während das periphere Gehör in einer Art 'Bottom-Up'-Entwicklung modelliert wird, gestaltet sich die Arbeit an einer praktischen Anwendung meist als 'Top-Down' Entwicklung, d.h. man geht von den zu simulierenden Meßdaten aus und versucht einen Zusammenhang zu den Erregungsgrößen herzustellen. Diese Vorgangsweise ist insofern sinnvoll, da die Kenntnisse über die Funktionsweise des Gehirns bzw. die zentrale Schallverarbeitung zur Zeit nicht ausreichen, um eine durchgehend physiologische Simulation zu entwickeln.

Erregungsmuster, die auf auditiven Filtern basieren, stellen ein brauchbares Konzept dar, um wesentliche Aspekte der auditiven Wahrnehmung zu simulieren. Die auditiven Filter repräsentieren, im Idealfall, sowohl die Frequenzselektivität als auch die Zeitauflösung des peripheren Gehörs. Im allgemeinen läßt sich ein derartig funktionales Modell leichter an praktische Anwendungen adaptieren als ein physiologisches Modell.

Die Notched-Noise-Methode ist eine von vielen Möglichkeiten, die Frequenzselektivität des Gehörs zu analysieren. Diese Methode wurde gewählt, da sie zur Zeit am weitesten entwickelt ist. Die Vor- und Nachteile der Notched-Noise Methode sind im folgenden kurz zusammengefaßt¹. Sie stellen die grundlegenden psychoakustischen Randbedingungen dar.

- Im Unterschied zu Methoden, die Sinussignale als Maskierer verwenden, ändern sich die Interaktionen zwischen Maskierer und Testsignal bei einer Veränderung der Einkerbungsbreite nicht.
- Das off-frequency listening wird sowohl minimiert als auch bei der Analyse mitberücksichtigt. Im Unterschied dazu muß bei schmalbandigen Maskierern das off-frequency listening mit zusätzlichen Maskierern verhindert werden. Diese Methode bedeutet eine Einschränkung des Dynamikbereiches².
- Bei Verwendung eines eingekerbten Maskierungsrauschens ist das SNR am Ausgang eines einzelnen auditiven Filters am höchsten. Es kann davon ausgegangen werden, daß nur ein einzelner Filter in den Detektionsprozeß involviert ist³. Verwendet man demgegenüber einen einzelnen schmalbandigen Maskierer, so entstehen mehrere auditive Filter mit etwa gleich großem SNR. In diesem Fall läßt sich nicht feststellen, welcher Filter bzw. welche Kombination von Filtern bei der Wahrnehmung tatsächlich verwendet wird.
- Die Notched-Noise-Methode ermöglicht eine getrennte Betrachtung der Frequenzselektivität sowie der Detektionseffizienz und ermöglicht die Formbestimmung von auditiven Filtern über einen weiten Dynamikbereich.

¹ vgl. [Moore et al., 1995], [Glasberg, Moore, 1996]

² [Johnson-Davies, Patterson, 1979], [O'Loughlin, Moore, 1981], [Moore et al., 1984]

³ [Moore et al., 1992]

- Der wesentlichste Nachteil der Notched-Noise-Methode besteht darin, daß zur Zeit nicht klar ist, ob innerhalb eines Experiments die Leistungsdichte des Maskierungsrauschens oder der Testsignalpegel konstant gehalten werden sollte. In dieser Arbeit wird die zweite Methode favorisiert, wenngleich dafür zur Zeit keine vollständigen Meßdaten vorhanden sind.
- Die Ergebnisse von Notched-Noise Experimenten können durch Kombinationsbänder¹ beeinflußt werden, die sich durch Interaktionen innerhalb des oberen Rauschbandes oder durch Interaktionen zwischen dem Testsignal und einem der beiden Rauschbänder ausbilden². Diese Kombinationsbänder können detektierbar sein, auch wenn das Testsignal selbst nicht wahrgenommen wird³. Der Einfluß dieser Kombinationsbänder auf die Bestimmung der auditiven Filter bzw. der ERB-Breiten gilt als vernachlässigbar⁴.
- Die Meßergebnisse beziehen sich auf junge normalhörende Personen. Die Varianzen sind, verglichen mit technischen Meßgrößen, verhältnismäßig hoch. Dieser Unsicherheitsfaktor ist integraler Bestandteil experimentalpsychologischer Versuchsreihen und läßt sich nicht vermeiden.

Die Amplitudengänge der auditiven Filter werden mit steigenden Pegeln flacher. Diese Abflachung wird primär durch die tieffrequente Filterflanke verursacht. Die meisten Probanden zeigen keine signifikanten Änderungen der hochfrequenten Filterflanken.

Bei der Berechnung von Erregungsmustern aus nicht linearen (d.h. pegelabhängigen) auditiven Filtern stellt sich die Frage, aus welchem Signal die pegelabhängige Steuerung der Filter bestimmt werden muß. Vernachlässigt man die Möglichkeit, Signale mehrerer Frequenzkanäle zur Bildung eines Steuerungspegels zusammenzufassen, so ergeben sich grundsätzlich zwei Möglichkeiten. Entweder man verwendet das Eingangssignal⁵ oder das Ausgangssignal eines Filters zur Steuerung der Nichtlinearität.

Der Streitpunkt, ob die Eingangs- oder Ausgangsleistung die nichtlineare Steuerung liefert, wird mit der Notched-Noise-Methode insofern in Verbindung gebracht, als eine konstante Maskiererleistungsdichte eine eingangspegelabhängige Steuerung und ein konstanter Testsignalpegel eine ausgangspegelabhängige Steuerung bedingt⁶.

¹ [Zwicker, Bubel, 1977], [Lufti, 1983], [Glasberg, Moore, 1994]

² [Moore et al. 1995]

³ [Greenwood, 1972]

⁴ [Lufti, Patterson, 1982]

⁵ Bei der eingangsabhängigen Steuerung wird der Steuerungspegel innerhalb der jeweiligen ERB-Breite ($f_c \pm 1/2$ ERB-Breite) gebildet ([Glasberg, Moore, 1990]).

⁶ [Rosen, Baker, 1994]

Kapitel 3

Berechnung der menschlichen Lautheitsempfindung

In diesem Kapitel werden Methoden zur Berechnung der menschlichen Lautheitsempfindung diskutiert. Es wird zwischen Lautheitsberechnungen auf Basis der Kurven gleicher Lautstärkepegel und auf Basis des Erregungsmusters unterschieden.

Die Berechnung der menschlichen Lautstärke- bzw. Lautheitsempfindung ist insofern von besonderem Interesse, da die Lautheitsempfindung mit der Lärmempfindung eng korreliert¹. Man kann davon ausgehen, daß ein lauterer Schallereignis mit größerer Wahrscheinlichkeit als Lärm empfunden wird als ein leiseres.

Die Lautheit läßt sich als getrennte Empfindungsgröße näherungsweise in einer physikalischen Metrik quantisieren. Dies ist die Voraussetzung für die Entwicklung eines technisch realisierbaren Modells. Demgegenüber läßt sich die Lärmempfindung um vieles schwieriger quantisieren, da die psychischen Randbedingungen einen wesentlich stärkeren Einfluß auf das Antwortverhalten von Versuchspersonen ausüben. Der Informationsgehalt bzw. die subjektive Interpretation des Informationsgehaltes eines Schallsignals wirkt sich beispielsweise bei Versuchsreihen zur quantitativen Erfassung der Lärmempfindung deutlich stärker aus als bei der Lautheitsempfindung².

Man beschränkt sich daher in der Lärmmeßtechnik primär auf die Bestimmung der Lautheit eines Schallsignals. Erst bei einer deutlichen Diskrepanz zwischen den Lautheitsmeßwerten und der tatsächlichen Lärmempfindung werden die Lautheitsmeßwerte um weitere Empfindungsmeßgrößen wie der Schärfe, Schwankungsstärke, Rauigkeit oder Tonhaltigkeit erweitert³. Für die Lärmempfindung selbst gibt es weder eine allgemein anerkannte Definition noch eine

¹ [Genuit, 1991], [Zwicker, 1991]

² [Fastl, 1988]

³ [Schiffbänker et al., 1991], [Zwicker, 1991]

allgemein anerkannte Meßgröße¹. Entsprechend den geltenden Normen, die sich zum größten Teil auf Daten der Lautheitsmeßtechnik stützen, müßte die Lärmmeßtechnik streng genommen in Lautheitsmeßtechnik umbenannt werden.

Diese Normen entsprechen nicht dem neuesten wissenschaftlichen Stand. Da jedoch keinem der vorhandenen Verfahren zur Berechnung der Lautheitsempfindung Allgemeingültigkeit attestiert werden kann, scheiterten bislang alle Versuche, die als problematisch geltenden gesetzlichen Bestimmungen des akustischen Umweltschutzes zu revidieren.²

3.1. Der Lautstärkepegel

Der Lautstärkepegel³ L_N in phon eines beliebigen Schallsignals stimmt zahlenmäßig mit dem Schalldruckpegel eines gleich laut empfundenen 1 kHz Signals überein. Ein Schallsignal besitzt also beispielsweise einen Lautstärkepegel von 60 phon, wenn es gleich laut empfunden wird wie ein 1 kHz Ton mit einem absoluten Schalldruckpegel⁴ von 60 dB.

Mit dieser Referenz ist es grundsätzlich möglich, jedem beliebigen Schallsignal einen Einzahlwert für die empfundene Lautstärkeempfindung zuzuordnen. Dafür müßte allerdings für jedes Schallsignal eine Gruppe 'normalhörender' Versuchspersonen einen 1 kHz Ton im Pegel so einstellen, daß der Testschall gleich laut empfunden wird wie der Referenzton. Diese Vorgangsweise ist ineffizient, da die Meßgrößen-Erfassung stets mit einer Versuchsreihe gekoppelt ist.^{5 6}

Die vorhandenen Lautstärkemaße entstammen ausschließlich Laborversuchen, die sich - verglichen mit Feldversuchen - einfacher durchführen und reproduzieren lassen. Laborversuche sind allerdings keinesfalls unumstritten.

'... In der Psychologie machen sich immer wieder Richtungen gegen Laborexperimente stark,

¹ [Zwicker, 1991]

² [Fastl, 1988], [Zwicker, 1991]

³ Im deutschsprachigen Raum wird der Lautstärkepegel auch als Pegellautstärke bezeichnet. Diese Bezeichnung geht auf Zwicker zurück (vgl. [Zwicker, 1982] S. 73). Sie konnte sich international nicht durchsetzen.

⁴ Verwendet man bei der Schallpegelberechnung einen beliebigen Bezugsschalldruck, so handelt es sich um einen *relativen Schalldruckpegel*. Der *absolute Schalldruckpegel* bezieht sich auf einen festgelegten Bezugsschalldruck von $p_0 = 20 \mu\text{Pa}$ [DIN 45 630]. Die Bezeichnungen Schallpegel oder Schalldruckpegel beziehen sich in dieser Arbeit stets auf den absoluten Schalldruckpegel:

$$L_p = 20 \log \frac{p}{p_0}$$

⁵ [Neumann, 1993] S.38-42

⁶ Abgesehen von der unrealistischen Ineffizienz ist diese Methode ebenso mit einer hohen methodischen Varianz behaftet wie psychophysische Wahrnehmungsexperimente im allgemeinen.

*weil sie diesen Experimenten grundsätzlich die Allgemeingültigkeit und Validität absprechen. So bezeichnet Kaminski (1988) die Laborsituation "als eine durch mancherlei 'Künstlichkeiten' ausgezeichnete Untermenge von Lebensumständen". Da diese ökologische Psychologie die Forderung nach Untersuchungen unter natürlichen Bedingungen berechtigterweise auf ihre Banner geschrieben hat, liegt eine solche Kritik auf der Hand. ...'*¹

Selbst bei Laborexperimenten können geringe Änderungen der Versuchsdurchführung zu stark unterschiedlichen Ergebnissen führen. Ein Beispiel dafür ist in Kap. 3.2. angeführt.

Um die Ermittlung von Lautstärkepegeln reproduzierbar (und damit objektiv) zu gestalten, müssen die Randbedingungen der Versuchsreihen bis ins kleinste Detail vorgegeben werden.

Abgesehen von der grundsätzlichen Problematik Laborversuche - Feldversuche, schränken diese strikten methodologischen Vorgaben die Allgemeingültigkeit der erhaltenen Ergebnisse weiter ein.

Es kann also festgehalten werden, daß der Lautstärkepegel eine Empfindungsgröße darstellt, die sich auf experimentalpsychologische Laborexperimente stützt und im Sinne einer physikalischen Meßgröße quantifiziert ist². Das Phonmaß stellt damit einen ersten wesentlichen Schritt zur Entwicklung von Lautstärke- bzw. Lautheitsmeßverfahren dar. Alle Lautheits-, Lautstärke- oder Lärmmeßverfahren stehen in irgendeiner Form mit dem Phonmaß in Verbindung. Die Auseinandersetzung mit den theoretischen Formulierungen und daraus resultierenden Gültigkeitsbereichen des Phon-Maßes ist dadurch von grundlegender Bedeutung.

3.2. Kurven gleicher Lautstärkepegel (KGL, Isophone)

Bestimmt man für Sinussignale oder Schmalbandrauschen unterschiedlicher Mittenfrequenzen denjenigen Schalldruckpegel, der zur gleichen Lautstärkeempfindung führt wie ein 1 kHz Signal, so erhält man die Kurven gleicher Lautstärkepegel. Sie geben einen ersten Einblick in die spektrale Lautstärkegewichtung des menschlichen Gehörs. Zeitliche Effekte werden dabei nicht berücksichtigt.

Die KGL wurden in den letzten 60 Jahren wiederholt Änderungen unterzogen. Die ersten vollständigen KGL wurden 1933 von Fletcher und Munson veröffentlicht. Der zweite vollständige Datensatz aus dem Jahr 1956 von Robinson und Dadson wurde in die internationale Norm der ISO unter ISO R 226³ aufgenommen. Sowohl die Fletcher-Munson als auch die Robinson-Dadson Kurven wurden teilweise interpoliert und lassen sich über weite Bereiche meßtechnisch nicht nachvollziehen.⁴

Aus diesem Grund wird seit einigen Jahren an der Revision des ISO Standards gearbeitet. Die

¹ [Schick, 1990], S. 29

² [Zwicker, 1982] S.1-15

³ [ISO 226]

⁴ [Gabriel, et al. 1993], [Gabriel et al., 1994], [Gabriel et al., 1995], [Suzuki, Sone, 1993]

Replikationsstudien Ende der 80er und Anfang der 90er Jahre weisen überraschenderweise relativ hohe Unterschiede auf. Da die Ursachen für diese Unterschiede nach wie vor nicht vollständig geklärt sind, konnte man sich bis heute auf keinen neuen Standard einigen.¹

In folgender Abbildung sind Kurven gleicher Lautstärkepegel unterschiedlicher Versuchsreihen dargestellt.

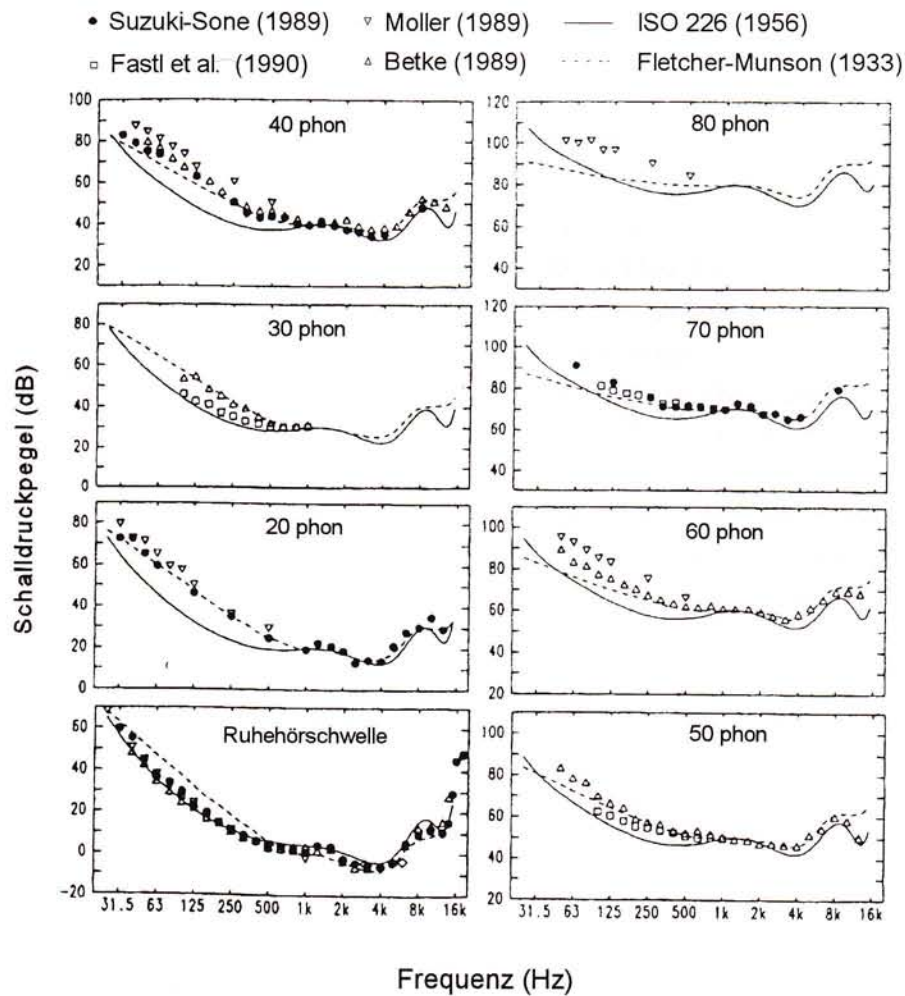


Abb. 3.1.: Vergleiche zwischen Kurven gleicher Lautstärkepegel unterschiedlicher Versuchsreihen, vgl. [Suzuki, Sone, 1993]

Abgesehen von den Fletcher-Munson Kurven stimmen die Ruhehörschwellen² bei allen Versuchsreihen überein. Die Unterschiede der anderen Phonkurven sind mitunter beträchtlich. Setzt man für alle dargestellten Versuchsreihen einen validen Versuchsaufbau voraus, so ergeben

¹ [Sorensen, 1994]

² Die Ruhehörschwellen werden üblicherweise mit einer eigenen Methode, der Methode des pendelnden Einregels bestimmt (vgl. [Zwicker, 1982] S. 31-33).

sich, im Bereich unter 1 kHz methodisch bedingte Abweichungen von bis zu 20 dB.

Der Frequenzbereich unter 1 kHz ist zur Bildung der Lautheit insofern wichtig, da bei natürlich vorkommenden Schallen in diesem Frequenzbereich die überwiegende Leistung konzentriert ist. Als weitere Schlußfolgerung aus diesem Vergleich läßt sich die Notwendigkeit der Revision des ISO 226 Standards ableiten.

*'... Die neuesten Untersuchungen der KGL zeigen erhebliche Unterschiede zum ISO 226 Standard. Daraus ergibt sich die Notwendigkeit diesen Standard zu revidieren. Auf der anderen Seite zeigen die neuen Daten untereinander starke Unterschiede. Daraus folgt, daß methodische Unterschiede einen starken Einfluß ausüben. ...'*¹

Anhand zweier sehr ähnlicher Versuchsreihen zur Bestimmung der KGL wird im folgenden gezeigt, welchen signifikanten Einfluß methodische Randbedingungen auf die gemessenen KGL haben können. Sowohl in [Fastl et al., 1990] als auch in [Betke, 1989] werden die KGL mit der Konstanzmethode² bestimmt. Trotzdem treten bei diesen beiden Versuchsreihen Unterschiede von bis zu 13 dB auf³. Die beiden Versuchsreihen unterscheiden sich im wesentlichen durch zwei Versuchsparameter. Erstens wird bei Betke der Testton aus einem 30 dB Intervall entnommen, bei Fastl aus einem 16 dB Intervall. Zweitens sind bei Betke die Testtonpegel über den Testtonpegelbereich gleichverteilt, während bei Fastl die Pegel häufiger in der Mitte des Testtonpegelbereichs dargeboten werden.

In einer ersten Analyse dieser Problematik fanden [Gabriel et al., 1994] heraus, daß die Pegelhäufigkeit keinen wesentlichen Einfluß auf die Lage der KGL hat, während die Lage des Pegelintervalls bzw. Testtonpegelbereichs sich nachhaltig auf die KGL auswirkt.

Weiters kommen [Gabriel et al., 1995] zu folgenden Ergebnissen:

'... Die hier beschriebenen Meßergebnisse haben gezeigt, daß bei der Konstanzmethode die Lage des Pegelintervalls einen signifikanten Einfluß auf die Lage der Isophonen hat, auch wenn der Testtonpegelbereich 50 dB umfaßt oder in einer Messung mehrere Frequenzen angeboten werden. Dabei ist der maximal dargebotene Testtonpegel nicht allein maßgeblich für den Kontexteffekt⁴. Werden in einer Messung verschieden laute Referenzöne präsentiert,

¹ *'... Recent results on the reexamination of the equal-loudness contours show significant differences from the present standard ISO 226. This indicates a need for the present standard to be entirely revised. The data, on the other hand, shows considerable discrepancy among researchers. Our results suggested that methodological difference could cause such discrepancy. ...'* ([Suzuki, Sone, 1993]).

² Bei der Konstanzmethode werden Tonpaare dargeboten, die vor der Messung festgelegt wurden. Jedes Tonpaar besteht aus dem 1-kHz Referenzton und dem Testton. Die Dauer der Signale beträgt für gewöhnlich 1s, die Pause zwischen Referenz- und Testton 0,5 s. Zur Bestimmung einer Phon-Kurve wird der Referenzton stets mit festem Pegel, der Testton mit verschiedenen Pegeln, die aus einem vorgegebenen Intervall gewählt werden, angeboten. Sowohl die Reihenfolge Testton - Referenzton als auch die Darbietung der Testpaare sind zufällig. Nach Darbietung eines Paares beurteilt die VP, welcher der beiden Töne lauter empfunden wurde. Aus dem 50% Punkt der psychometrischen Funktion ergibt sich der Pegel des Testtons, der gleich laut wie der Referenzton empfunden wird.

³ [Gabriel et al., 1994]

⁴ Die Abhängigkeit des Lautstärkeurteils von den gewählten Testtonpegeln wird als Kontexteffekt bezeichnet.

verschwindet der Kontexteffekt. Es ist allerdings nicht auszuschließen, daß auch hier die Ergebnisse wieder von den speziellen Parametern dieser Messung abhängen.

Die hier beschriebenen Ergebnisse lassen keinen eindeutigen Schluß zu, ob der Kontexteffekt auf dem Antwortverhalten der VP beruht, oder ob die in einer Messung dargebotenen Pegel die Wahrnehmung der VP im Sinne einer Adaption ändern. Generell deuten die Ergebnisse darauf hin, daß Töne mit unterschiedlichen Frequenzen bezüglich ihrer Lautstärke nicht kontextfrei beurteilt werden können. ...¹

Obige Untersuchungen der KGL setzen ein ebenes Schallfeld sowie eine Schalleinfallrichtung von 0° (d.h. 'von vorne') voraus. Durch die Richtcharakteristik des menschlichen Gehörs unterscheiden sich die KGL für freie und diffuse Schallfelder. Der in der ISO 226 festgehaltene Freifeld-Diffusfeld Differenzpegel ΔL bezieht sich auf objektive Experimente d.h. auf die Unterschiede der Übertragungsfunktionen des menschlichen Kopfes (HRTF ... head related transfer function). Diese Differenzpegel konnten von [Takeshima et al., 1994] mit einer subjektiven d.h. mit einer psychoakustischen Versuchsreihe für die Ruhehörschwelle und die 40 phon Kurve näherungsweise nachgebildet werden.

'... Es konnte bestätigt werden, daß die Formen der subjektiven (mit Versuchsreihen bestimmten) und der objektiven (Unterschiede der HRTF-Kurven) Differenzpegel ΔL ähnlich sind. Wir fanden, daß die subjektiven Differenzpegel 1-3 dB kleiner sind als die objektiven Differenzpegel. Diese Unterschiede können wir zur Zeit nicht erklären. ...²

3.3. Kurven gleicher Reaktionszeit (Isochrone)

Die Lautheitswahrnehmung für Sinustöne ist der Reaktionszeit auf die dargebotenen Schallsignale äquivalent. Zu diesem Ergebnis kommen [Müller, Fichtl, 1994] in einer Untersuchung, bei der sie sowohl Isophone (KGL) als auch Isochrone (Kurven gleicher Reaktionszeit) unter gleichen Versuchsbedingungen ermittelten.

In einer ersten Untersuchung wurden KGL mit dem Kategorien-Unterteilungsverfahren³ (KU) bestimmt. Ein Vergleich zwischen diesen KU-KGL und den KGL nach ISO 226 ist in Abb. 3.2. dargestellt.

¹ [Gabriel et al., 1995]

² *'... It was confirmed that the shape of the subjective ΔL and the objective ΔL are similar. It is known that the head transfer function is a main factor in ΔL . However, we found that Δ threshold, i.e., the subjective difference, was approximately 1 - 3 dB lower than Δ HRTF, the objective value. We have not been successful in explaining this difference yet. ...'* ([Takeshima et al., 1994]).

³ Bei dem Kategorien-Unterteilungsverfahren werden die dargebotenen Schallereignisse in unterschiedliche Kategorien wie leise, mittel, laut eingeteilt. [Müller, Fichtl, 1994] verwendeten ein zweistufiges Verfahren. In einer ersten Stufe wurden die VPs angewiesen nach 5 Kategorien zu unterteilen (sehr leise, leise, mittel, laut, sehr laut). In einer zweiten Stufe wurde innerhalb einer dieser 5 Kategorien differenziert. Dafür wurde jede Kategorie in 10 Intervalle eingeteilt (nichts gehört: 0, sehr leise: 1 - 10, leise: 11 - 20, etc.).

Die Übereinstimmung zwischen diesen beiden Datensätzen ist, trotz der unterschiedlichen Bestimmungsmethoden¹, relativ gut. Die Unterschiede werden von [Müller, Fichtl, 1994] auf die unterschiedlichen Eigenschaften der Versuchsräume und auf methodische Einflüsse zurückgeführt.

Da die Eigenschaften der bei diesen Versuchen verwendeten Schallkabinen äußerst schlecht waren (30 - 40 dB(A) Grundgeräuschpegel), ist die Validität der Ergebnisse bei tieffrequenten und niederpegeligen Schallsignalen nicht gegeben. Auffallend an Abb. 3.2. ist, daß die mit der KU-Methode bestimmten KGL zwischen etwa 300 Hz und 1 kHz ein ähnliches Verhalten aufweisen wie die ISO KGL. Beide Untersuchungen attestieren dem menschlichen Gehör in diesem Bereich eine stärkere Lautheits-Empfindlichkeit als an 1 kHz. Dieser Effekt ist bei keiner anderen Versuchsreihe vorhanden².

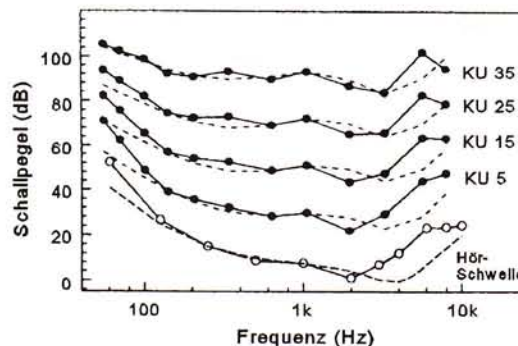


Abb. 3.2.: Vergleich zwischen den KU-KGL (durchgezogene Linien) und den KGL nach ISO 226 (strichlierte Linien), vgl. [Müller, Fichtl, 1994]

In einem zweiten Versuch wurden die Reaktionszeiten der Versuchspersonen auf Sinussignale unterschiedlicher Frequenz und Pegel bestimmt. Die damit ermittelten Kurven gleicher Reaktionszeit zeigen ähnliches Verhalten wie die mit der Kategorien-Unterteilungsmethode bestimmten KGL (Abb. 3.3.).

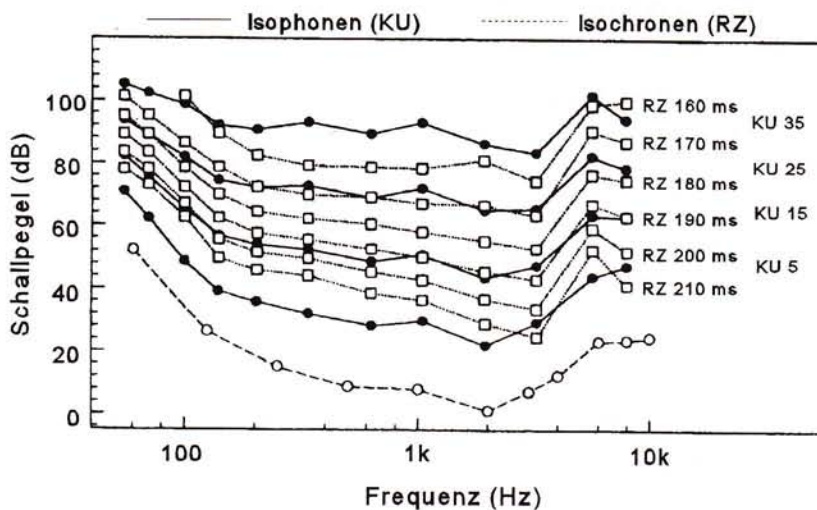


Abb. 3.3.: Vergleich zwischen den KU-KGL und den Kurven gleicher Reaktionszeit, vgl. [Müller, Fichtl, 1994]

¹ Die ISO 226 Phonkurven wurden mit der Konstanzmethode bestimmt.

² [Suzuki, Sone, 1993]

Die Standardabweichungen der Reaktionszeiten werden mit 33,1 ms (intraindividuell) und 35 ms (zwischen VP) angegeben¹. Ein wesentlicher Unterschied besteht u.a. im Bereich zwischen etwa 300 Hz und 1 kHz. Die Isochronen stimmen in diesem Bereich besser mit den neuesten Versuchsreihen, basierend auf der Konstanzmethode, überein als die KU-KGL. Ein weiterer Beweis dafür, daß die KGL nach ISO 226 im Bereich unter 1 kHz falsch sind.

Die Verwendung der Reaktionszeit als Zwischengröße zur Ermittlung der KGL stellt einen neuen Ansatz dar. Diese Methode ist insofern von Interesse, da die seit langem ungeklärten komplexen methodischen Einflüsse der Konstanzmethode vermieden werden können. Dabei ergeben sich allerdings neue Randbedingungen, deren Einflüsse in weiterführenden Untersuchungen ermittelt werden müssen.

3.4. Die Lautheit

Umgangssprachlich werden die beiden Begriffe Lautstärke und Lautheit als gleichwertig betrachtet. Bei der quantitativen Erfassung der Lautstärke- bzw. Lautheitsempfindung unterscheidet man jedoch diese beiden Begriffe durch unterschiedliche Einheiten. Der Lautstärkepegel L_N wird in phon angegeben und die Lautheit N in sone.

Einem Sinussignal mit 1 kHz und 40 dB absolutem Schalldruckpegel (dies entspricht einem Lautstärkepegel von $L_N = 40$ phon) wird eine Lautheit von $N=1$ sone zugeordnet. Die Skalierung in sone wurde so gewählt, daß eine Änderung der Lautstärkeempfindung um einen bestimmten Faktor zahlenmäßig die gleiche Änderung des Lautheitsmaßes bewirkt. Verdoppelt man beispielsweise die empfundene Lautstärke eines 1 kHz Signals mit $L_p = 40$ dB ($L_N = 40$ phon, $N = 1$ sone), so ergibt sich ein absoluter Schalldruckpegel von $L_p = 50$ dB ($L_N = 50$ phon) und eine Lautheit von $N = 2$ sone. Damit ist aus der quantifizierten Empfindungsgröße der Lautheit direkt ablesbar, um wieviel mal lauter ein Schallereignis im Vergleich zu einem anderen Schallereignis ist. Daraus wird die größere Nähe dieser Empfindungsgröße zur tatsächlichen Lautstärkeempfindung abgeleitet.²

Die Ermittlung der Lautheitsskalierung beruht auf psychoakustischen Messungen der Verhältnislautheit; d.h. man konzentrierte sich bei den Messungen auf die Fragestellung, um welchen Schallpegel ein 1 kHz Sinussignal verändert werden muß, um ein gewisses Verhältnis (beispielsweise eine Verdoppelung) der ursprünglichen Lautheit zu erreichen. Damit läßt sich zwischen dem absoluten Schalldruckpegel eines 1 kHz Signals bzw. dem Lautstärkepegel L_N und der Lautheit N eine eindeutige Beziehung definieren. Dieser Zusammenhang ist in Abb. 3.4. dargestellt.

¹ [Müller, Fichtl, 1994]

² [Zwicker, Feldtkeller, 1967] S.118-149, [Zwicker, 1982] S.68-83, [Zwicker, Fastl, 1990] S.181-214, [Moore, Glasberg, 1996]

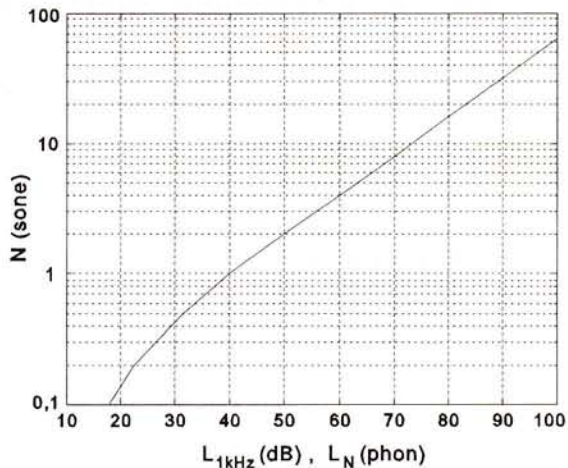


Abb 3.4.: Zusammenhang zwischen der Lautheit N und dem Lautstärkepegel L_N , vgl. [Moore, Glasberg, 1996]

Mit diesem Zusammenhang können, unabhängig von der Art des Schallsignals, Lautstärkepegel und Lautheiten eindeutig ineinander umgerechnet werden. Dies ist insofern erforderlich, da diese beiden Größen die gleiche Qualität mit unterschiedlichen Quantitäten ausdrücken. Aufgrund der größeren Nähe der Lautheit zur tatsächlichen Lautstärkeempfindung wird sie in der Psychoakustik dem Lautstärkepegel für gewöhnlich vorgezogen.

3.5. Lautheitsberechnung auf Basis der Kurven gleicher Lautstärkepegel

3.5.1. Bewertete Schalldruckpegel

Das A-bewertete Dezibel dB(A) bildet die Empfindungsgröße Lautheit durch Filterung mittels einer annähernd inversen 40-Phon-Kurve nach. Daraus ergeben sich zwei systematische Fehler. Erstens sind die Phon-Kurven nur für Schallereignisse gültig, deren Bandbreite kleiner als die einer Frequenzgruppe ist, wodurch die erhöhte Lautheit breitbandiger Schallsignale keine Berücksichtigung findet. Zweitens kann aus der Filterung mit einer einzelnen inversen Phon-Kurve kein Urteil über die Gewichtungsunterschiede bei verschiedenen Pegeln abgeleitet werden. Die in [Pflüger et al., 1996a] vorgestellte Meßgröße dB(ELC) (Equal Loudness Curves) stellt eine detailliertere Implementation der Kurven gleicher Lautstärkepegel dar und ermöglicht durch eine dynamische Phonkurvenanpassung die Berücksichtigung des gesamten Hörfeldes. Diese Methode war in den dreißiger Jahren bereits bekannt, stellte die damaligen Akustiker jedoch vor technisch nicht lösbare Probleme. Man einigte sich daher auf 3 repräsentative Phonkurven: dB(A), dB(B) und dB(C), von denen jeweils eine den akustischen Gegebenheiten entsprechend ausgewählt wurde. Aus diesen drei Bewertungskurven entwickelte sich Ende der sechziger Jahre das dB(A) als die 'vorzugsweise zu verwendende' Meßgröße¹. Die Berechnung des dB(ELC) wird deshalb bewußt unter Einbeziehung des A-bewerteten Dezibels durchgeführt, um eine Kompatibilität zu vorhandenen Meßgeräten herzustellen.

Die detaillierte Implementation dieses Verfahrens sowie die Beschreibungen der für dieses Verfahren entworfenen Filter sind im Anhang A angeführt.

¹ [Schick, 1990] S.16-22

3.5.2. Dynamische Phonkurvenanpassung als Erweiterung des dB(A)

3.5.2.1. Beschreibung des Verfahrens¹

Ein ELC-Filter bildet, in Abhängigkeit eines Lautstärkepegel-Schätzwertes, die inverse 40, 60, 80 oder 100 Phon-Kurve nach (Abb. 3.5.). Zur Differenzierung zwischen freiem und diffusem Schallfeld wird das Ausgangssignal des ELC - Filters zusätzlich mit einem Diffusfeld-Filter gewichtet. Dadurch ergeben sich zwei Meßgrößen, das dB(ELC)f für freie Schallfelder und das dB(ELC)d für diffuse Schallfelder.

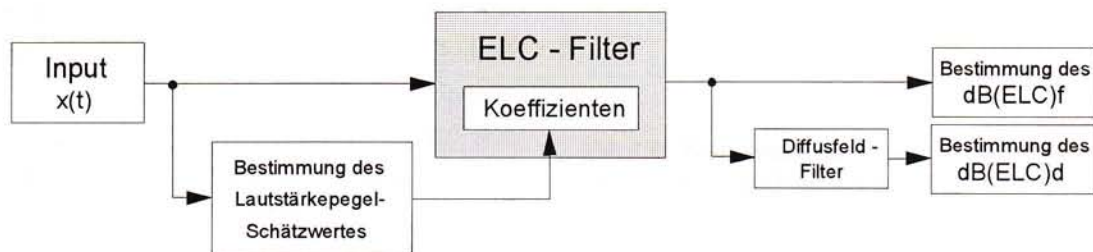


Abb. 3.5.: Filterung eines Schallsignals entsprechend den Kurven gleicher Lautstärkepegel (ELC-Filter)

Der Lautstärkepegel-Schätzwert als Steuerungsparameter des ELC-Filters errechnet sich aus dem unbewerteten sowie dem A-bewerteten Schallpegel (Abb. 3.6.).

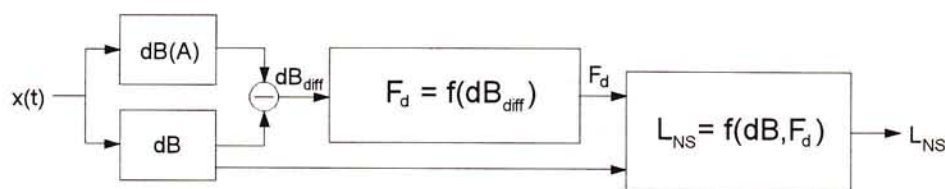


Abb. 3.6.: Bestimmung des Lautstärkepegel-Schätzwertes L_{NS}

In einem ersten Schritt wird aus der Differenz zwischen unbewertetem und A-bewertetem Schallpegel dB_{diff} und der Übertragungsfunktion des A-bewertenden Filters $A(f)$ eine sogenannte 'dominierende Frequenzkomponente' F_d bestimmt (vgl. Abb. 3.6.: $F_d = f(dB_{diff})$).

Diese Vorgangsweise unterstellt eine einzelne Sinuskomponente, die entsprechend dem A-Filter in ihrer Amplitude gewichtet wurde. Für eine Schätzung des Lautstärkepegels erscheint diese Vereinfachung ausreichend, da damit lediglich eine ungefähre Abschätzung des Phon-Bereichs

¹ [Pflüger et al., 1996a], [Pflüger et al., 1996b]

erreicht werden soll (Abb. 3.7.).

Aufgrund der negativen Steigung des A-Filters ab etwa 2.5 kHz können nur dominierende Frequenzen bis zu 2.5 kHz eindeutig bestimmt werden. Alle höheren Frequenzen werden auf den Frequenzbereich unter 2.5 kHz abgebildet.

Unter Verwendung eines Bewertungsfilters mit einem, im Unterschied zur A-Bewertung, monotonen Amplitudengang könnte dieser Fehler vermieden werden. Die vorteilhafte Kompatibilität zu vorhandenen dB(A)-Meßgeräten ginge dadurch jedoch verloren. Mit der dominierenden Frequenzkomponente F_d und dem unbewerteten Dezibel wird der Lautstärkepegel-Schätzwert aus einer Tabelle bestimmt (vgl. Abb. 3.6.: $L_{NS} = f(\text{dB}, F_d)$).

Dabei wird zwischen vier Lautstärkepegelbereichen: 40, 60, 80 und 100 Phon unterschieden. Dies bedeutet, daß der Lautstärkepegel-Schätzwert lediglich einen dieser vier Werte annehmen kann.

Als Beispiel für die Bildung des Lautstärkepegel-Schätzwertes L_{NS} aus der dominierenden Frequenzkomponente und dem unbewerteten Schallpegel sind in Abb. 3.8. die Schätzwerte für $F_d = 26$ Hz und 1055 Hz unter Verwendung der Kurven gleicher Lautstärkepegel nach ISO 226 grafisch dargestellt.

Bei einer dominierenden Frequenzkomponente von 26 Hz mit 80 dB ergibt sich beispielsweise ein Schätzwert von $L_{NS} = 40$ Phon, der den Koeffizientensatz der inversen 40-Phon Kurve am ELC-Filter auswählt¹.

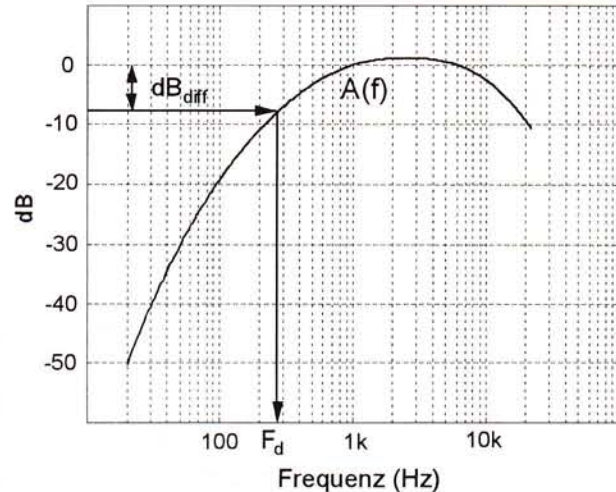


Abb.3.7.: Bestimmung der dominierenden Frequenzkomponente F_d als Funktion der Pegeldifferenz dB_{diff}

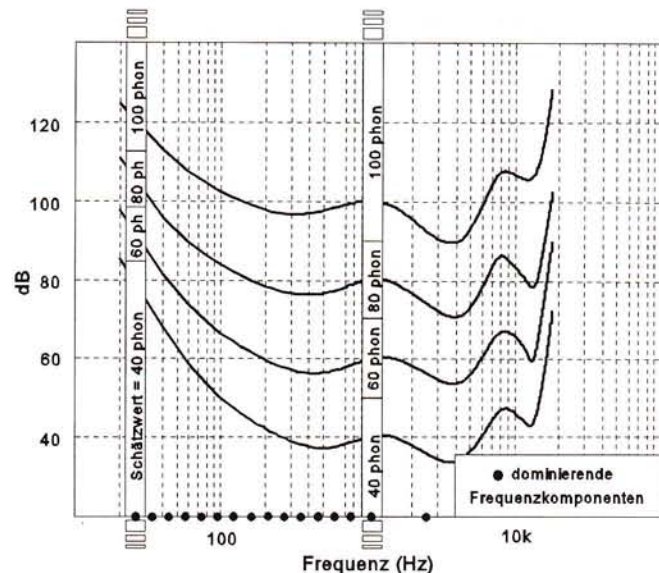


Abb. 3.8.: Lautstärkepegel-Schätzwerte L_{NS} als Funktion der dominierenden Frequenzkomponenten F_d und dem unbewerteten Schallpegel

¹ Auch in dieser Verarbeitungsstufe, dem 'Kern' der dB(ELC)-Berechnung, wird davon ausgegangen, daß es sich um ein Signal handelt, dessen lautheitsbestimmender Spektralanteil die Breite einer Frequenzgruppe nicht übersteigt.

Da sich die ISO-Phon Kurven wesentlich von neueren Forschungsergebnissen¹ unterscheiden, wurden für die Berechnung des dB(ELC) zusätzlich neue Phon Kurven für 40, 60, 80 und 100 phon entwickelt (Abb. 3.9.).

Diese Modifikationen sind eine Mittelung über die mitunter stark divergierenden Ergebnisse neuerer Versuchsreihen. Sie stellen keine endgültigen Kurven gleicher Lautstärkepegel dar. Ihre Funktion beschränkt sich vielmehr darauf, tendenzielle Unterschiede zwischen den ISO-Phon-Kurven und neueren Forschungsergebnissen zu dokumentieren. Besonders im Bereich unter 1 kHz unterscheiden sich die ISO Phon Kurven wesentlich von allen neuen Untersuchungen. Aus diesem Grund erscheint es sinnvoll, das Hauptaugenmerk auf die modifizierten Phon-Kurven zu legen.

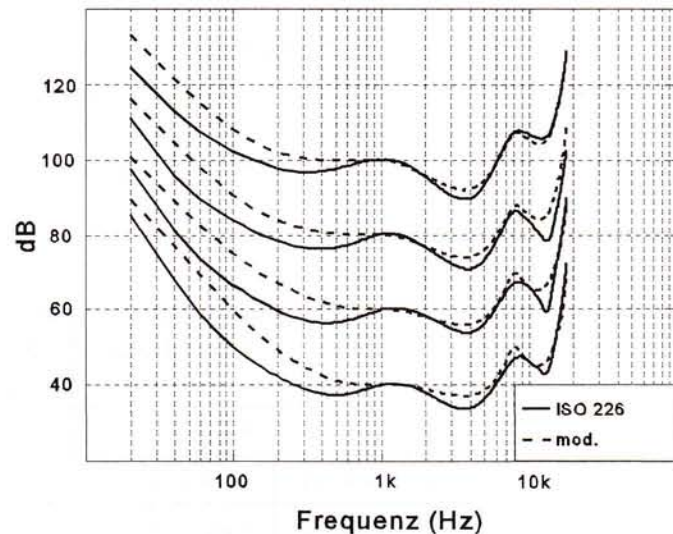


Abb. 3.9.: Vergleich zwischen den ISO 226 und den modifizierten Phon Kurven

Zur Differenzierung zwischen freiem und diffusem Schallfeld (dB(ELC)f und dB(ELC)d) wird das Ausgangssignal des ELC-Filters zusätzlich mit einem Diffusfeld-Filter (vgl. Kap. 3.2.) gefiltert (Abb. 3.10.).

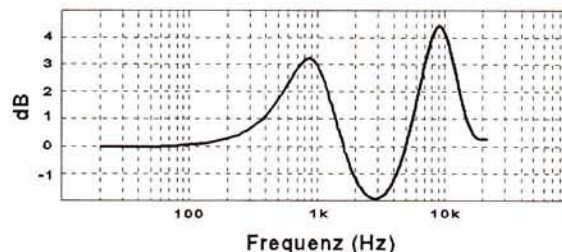


Abb. 3.10.: Amplitudengang des Diffusfeld-Filters

Für eine vergleichende Analyse ergeben sich vier unterschiedliche Meßwerte:

- dB(ELC)fI ... Freifeld, ISO 226
- dB(ELC)dI ... Diffusfeld, ISO 226
- dB(ELC)fm ... Freifeld, mod. Phonkurven
- dB(ELC)dm ... Diffusfeld, mod. Phonkurven

¹ [Gabriel et al., 1993], [Suzuki, Sone, 1993], [Sorensen, 1994], [Fastl et al., 1990]

3.5.2.2. Vergleichende Analysen

Zur Bestimmung der Unterschiede zwischen dem dB(A) und dem dB(ELC) wurden aus einem Geräuscharchiv¹ aus den Bereichen Industrie, Großstadt, Transport und Verkehr 344 Ausschnitte zu je 10 sek. entnommen. Diese Ausschnitte wurden blockweise (Blocklänge $N = 2048$ Abtastwerte) mit einer Abtastrate von 44,1 kHz und zwei unterschiedlichen Referenzfaktoren RF analysiert². Daraus ergaben sich etwa 150.000 Signalblöcke zu je 46,4 ms, für die jeweils der dB-, dB(A)-, dB(ELC)fl-, dB(ELC)dI-, dB(ELC)fm- und der dB(ELC)dm-Wert bestimmt wurden. Zur übersichtlicheren Darstellung wurden die analysierten Signalblöcke in 8 dB(A)-Gruppen unterteilt (Tab. 3.1.).

Bezeichnungen		dB(A)-Bereich	Signalblöcke pro Gruppe	Bezeichnungen		dB(A)-Bereich	Signalblöcke pro Gruppe
G1	Gruppe 1	$30 \leq \text{dB(A)} < 40$	5591	G5	Gruppe 5	$70 \leq \text{dB(A)} < 80$	19356
G2	Gruppe 2	$40 \leq \text{dB(A)} < 50$	13604	G6	Gruppe 6	$80 \leq \text{dB(A)} < 90$	20217
G3	Gruppe 3	$50 \leq \text{dB(A)} < 60$	23172	G7	Gruppe 7	$90 \leq \text{dB(A)} < 100$	27771
G4	Gruppe 4	$60 \leq \text{dB(A)} < 70$	28451	G8	Gruppe 8	$100 \leq \text{dB(A)} < 110$	8002

Tab. 3.1.: Zusammenfassung des dB(A)-Bereichs in 8 Gruppen G1 - G8

Die beiden Grafiken Abb. 3.11. und 3.12. zeigen die mittleren Pegeldifferenzen dB(ELC) - dB(A) sowie die entsprechenden Standardabweichungen in den dB(A)-Gruppen G1-G8.

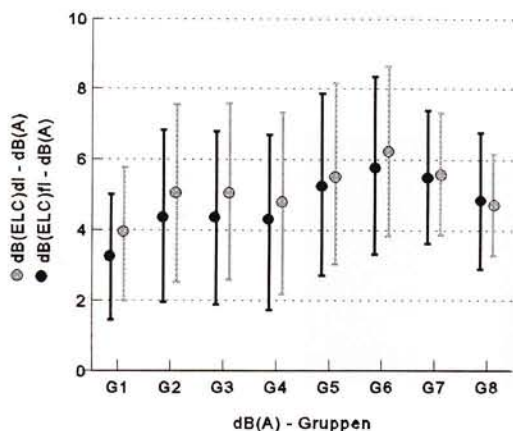


Abb. 3.11.: Differenz zwischen den dB(ELC)-ISO-Werten und dem dB(A) als Funktion der dB(A)-Gruppen

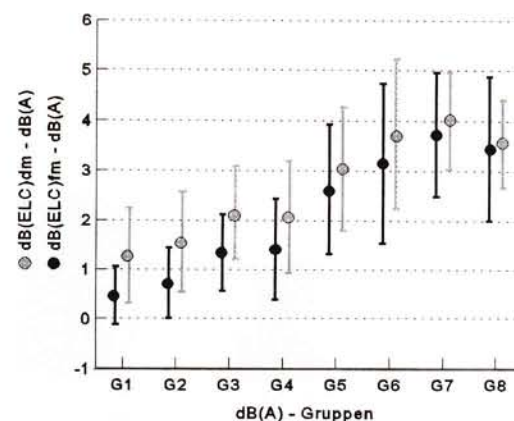


Abb. 3.12.: Differenz zwischen den dB(ELC)-mod.-Werten und dem dB(A) als Funktion der dB(A)-Gruppen

¹ [Digifffects]

² Mit dem Referenzfaktor RF kann das Eingangssignal unterschiedlichen Lautstärkepegelbereichen zugeordnet werden (vgl. Anhang, [Pflüger et al., 1996a]).

Die ELC Filterung bewertet die analysierten Schalle höher als der A-bewertende Filter. Die Mittelwerte bewegen sich für die ISO-Kurven im Bereich von 3,23 bis 6,24 dB, für die modifizierten Phon-Kurven im Bereich von 0,46 bis 4 dB. Die erwartete Differenzzunahme mit steigendem dB(A)-Wert ist ersichtlich, wobei ein im Mittel höherer dB(A)-Wert nicht zwingend eine höhere mittlere dB(ELC)-Pegeldifferenz bedingt. Der Abfall ab Gruppe 6 (ISO) bzw. Gruppe 7 (mod.) läßt sich auf die spektralen Eigenschaften der zu Gruppen zusammengefaßten Blöcke zurückführen. Der stärkere Abfall des dB(ELC)I wird durch die geringere Verstärkung des ISO-100 Filters im Vergleich zu den übrigen ISO Filtern im Bereich zwischen 300 Hz und 1 kHz verursacht. Dieser Effekt unterstreicht die Bedeutung dieses Frequenzbereiches zur Bildung bewerteter Schallpegel.

Aufgrund der im Verfahren verwendeten ELC-Filter, entsprechend den inversen 40, 60, 80 und 100 phon Kurven, ergibt sich im Bereich zwischen 40 und 100 phon eine Auflösung von 20 phon. Im Unterschied dazu ist der dB(A)-Bereich in 10 dB(A)-Schritten unterteilt. Dadurch entsteht die Möglichkeit, daß benachbarte Gruppen größtenteils mit demselben ELC-Filter gefiltert wurden. Für die beiden Gruppen G7 und G8 wird beispielsweise angenommen, daß beide Gruppen mit demselben Filter (ELC-100) gefiltert wurden. Die unterschiedlichen mittleren Pegeldifferenzen ergeben sich dadurch einzig aus den spektralen Unterschieden der beiden Gruppen.

Zur näheren Betrachtung der dB(ELC) Abweichungen bei gleichen dB(A)-Werten wurden jeweils innerhalb einer Gruppe Paare mit gleichen dB(A)-Werten gebildet.¹ Die zugehörigen dB(ELC)-Differenzen sind in den beiden Abb. 3.13. und 3.14. für das dB(ELC)_{fl} und das dB(ELC)_{fm} in Form von Histogrammen dargestellt. Aufgrund der verhältnismäßig geringen Unterschiede zwischen Freifeld- und Diffusfeldwerten wurde in dieser zweiten Analyse auf die Berücksichtigung der Diffusfeldwerte verzichtet.

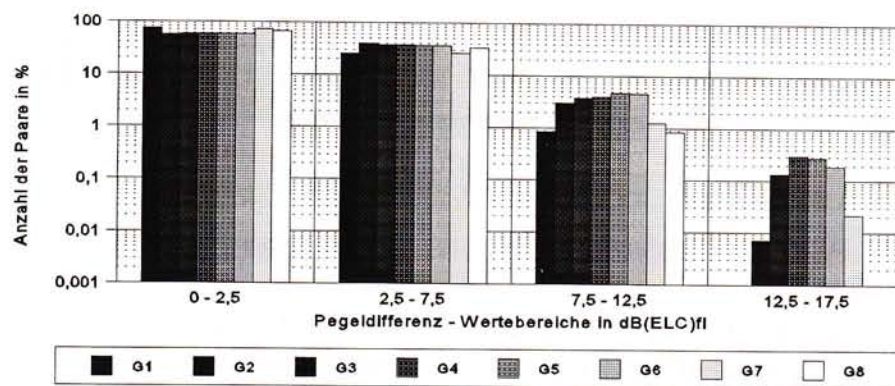


Abb. 3.13.: dB(ELC)_{fl} Differenzen bei gleichen dB(A)-Werten getrennt nach den 8 dB(A) Gruppen G1 bis G8

¹ [Pflüger et al., 1996b]

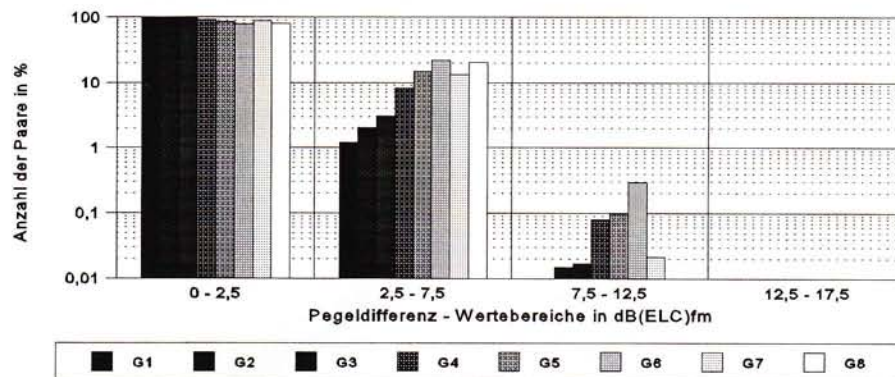


Abb. 3.14.: dB(ELC)fm Differenzen bei gleichen dB(A)-Werten getrennt nach den 8 dB(A) Gruppen G1 bis G8

Entsprechend den acht dB(A)-Gruppen enthält jede der beiden Grafiken 8 farblich voneinander getrennte Histogramme. Für das dB(ELC)fl (Abb. 3.13.) ergaben sich zum Beispiel für die Gruppe G4 (60-70 dB(A), 28450 Paare mit gleichem dB(A)-Wert = 100%) folgende Werte:

dB(ELC)fl - Differenz	0 - 2,5	2,5 - 7,5	7,5 - 12,5	12,5 - 17,5
Anzahl der Paare in %	57,8	37,8	4	0,3

Dies bedeutet, daß innerhalb der Gruppe G4 57,8 % der Paare mit gleichem dB(A)-Wert eine dB(ELC)fl-Differenz aufweisen, die sich im Bereich von 0 bis 2,5 dB(ELC)fl bewegt.

Die maximalen Differenzen der zu einem dB(A)-Paar gehörenden dB(ELC)fl-Werte fallen in den Bereich zwischen 12,5 und 17,5 dB. Diese relativ hohen Unterschiede treten in den Gruppen G3 bis G6 immerhin noch mit Wahrscheinlichkeiten zwischen 0,1 und 0,3 % auf.

Beim dB(ELC)fm entstanden maximale Differenzen zwischen 7,5 und 12,5 dB. Diese Abweichung zwischen den genormten und den modifizierten Kurven läßt sich zum überwiegenden Teil auf die unterschiedlichen Amplitudengänge der ELC-Filter unter 1 kHz zurückführen, da bei den gewählten Schallsignalen die tieffrequenten Signalanteile dominieren.

Aufgrund der relativ großen Abweichungen zwischen dem dB(ELC) und dem dB(A) wurden diese beiden Meßgrößen in einer dritten Analyse den beiden zur Zeit leistungsfähigsten Lautheitsberechnungsverfahren gegenübergestellt.¹

Dafür wurden 15.000 Signalausschnitte aus den Bereichen Industrie, Großstadt, Transport und Verkehr zu je 125ms analysiert. Die Schallausschnitte wurden wiederum willkürlich ausgewählt, wobei sich aufgrund der gewählten Schallarten eine gewisse Tendenz in Richtung breitbandiger Schalle mit starken geräuschhaften Anteilen ergab.

¹ [Pflüger et al., 1996b]

Für jeden Signalausschnitt wurde das dB(A), dB(ELC)fl, dB(ELC)fm, phonGF¹ und das phonStR² berechnet.

Die Verteilungen der einzelnen Meßgrößen sind in den Abb. 3.15. - 3.18. jeweils in Bezug auf das phonGF dargestellt.

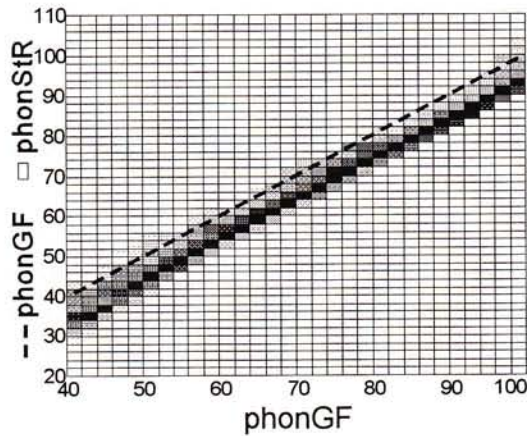


Abb. 3.15.: Vergleich phonGF - phonStR

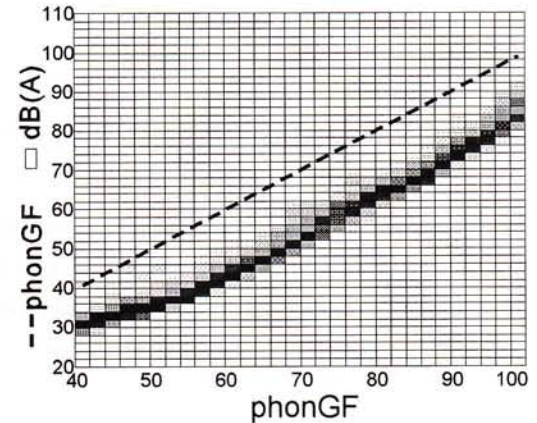


Abb. 3.16.: Vergleich phonGF - dB(A)

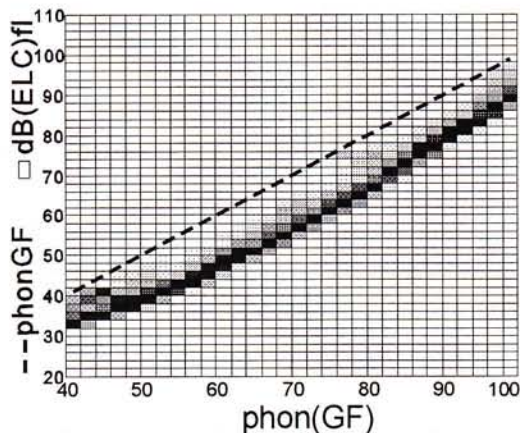


Abb. 3.17.: Vergleich phonGF - dB(ELC)fl

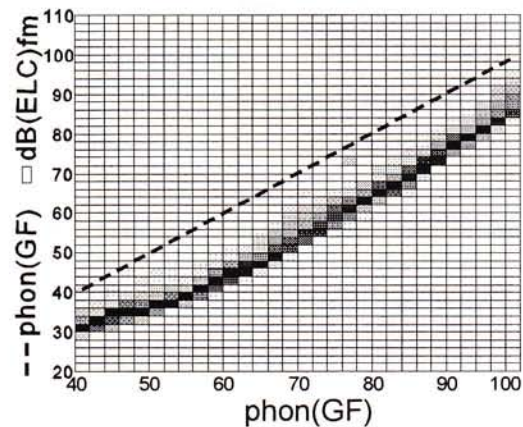


Abb. 3.18.: Vergleich phonGF - dB(ELC)fm

Spalte 1 in Abb. 3.15. symbolisiert beispielsweise die Verteilung des phonStR für diejenigen Signalblöcke, deren phonGF-Werte zwischen 40 und 42 phonGF liegen. Die phonStR-Werte sind um durchschnittlich 5 phon geringer als die phonGF-Werte und ließen sich durch eine generelle Anhebung um 5 phon in phonGF umrechnen.

¹ Berechnungsverfahren nach Zwicker, Freifeldmeßwerte, [Zwicker, 1982] S. 138-145

² Berechnungsverfahren nach Stevens mit Terzpegeln und Modifikation nach Robinson ([Schick, 1990] S. 50-52).

Zwischen dem phonGF und den dB-Werten (Abb. 3.16.-3.17.) besteht kein linearer Zusammenhang. Die Abstände sind darüber hinaus größer als zwischen dem phonGF und dem phonStR, da die Bandbreite der Schalle keine Berücksichtigung findet. Die Streuungen der grafisch dargestellten Verteilungen sind für alle 4 Meßgrößen etwa gleich groß.

Die erwartete stärkere Korrelation zwischen dem Verfahren nach Zwicker und dem Verfahren nach Stevens konnte nicht bewiesen werden. Ebenso wenig stellt das dB(ELC) für die untersuchten Schallereignisse eine wesentliche Verbesserung bezüglich des dB(A) dar.

3.6. Lautheitsberechnung auf Basis des Erregungsmusters

Die Grundlage für die auf dem Erregungsmuster basierenden Verfahren zur Lautheitsberechnung bildet das Verfahren nach Zwicker¹. Dieses Verfahren ist bis zur heutigen Zeit gültig, da die vorgeschlagenen Verbesserungen die gleiche Grundstruktur besitzen und keine qualitativen Neuerungen beinhalten. Diese Struktur ist in Abb. 3.19. dargestellt:

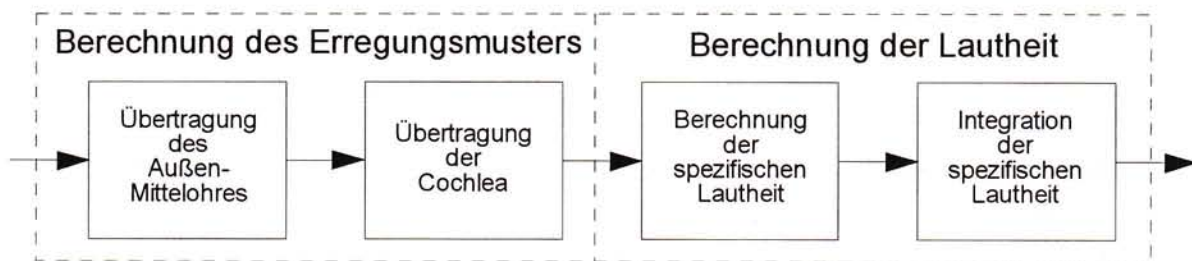


Abb. 3.19.: Lautheitsberechnung auf Basis des Erregungsmusters

Das Verfahren teilt sich in zwei Teile. Im ersten Teil wird das Erregungsmuster berechnet, im zweiten Teil wird aus dem Erregungsmuster die Lautheit als Einzahlwert bestimmt. Das Erregungsmuster (Zeit-Frequenzdarstellung des Schalles unter Berücksichtigung der Übertragungseigenschaften des Gehörs, vgl. Kap. 2.7.) kann unabhängig von der Lautheit betrachtet werden.

Die hier behandelten Lautheitsberechnungsverfahren basieren auf zwei unterschiedlichen Erregungsmustern. Bei dem Verfahren nach Zwicker sowie bei der Verbesserung nach Zollner² wird das Erregungsmuster aus dem Maskierungsmuster berechnet (vgl. Kap. 2.7.1.). Bei den Verbesserungen nach Glasberg & Moore³ und Agerkvist⁴ wird das Erregungsmuster mit Hilfe von auditiven Filtern bestimmt (vgl. Kap. 2.7.2.).

¹ [Zwicker, Feldtkeller, 1967] S.184-203, [Zwicker, 1982] S.138-145, [Zwicker, Fastl, 1990] S.283-291, [DIN 45 631], [ISO 532]

² [Zollner, 1996]

³ [Glasberg, Moore, 1996]

⁴ [Agerkvist, 1994]

3.6.1. Berechnung des Erregungsmusters

Übertragung des Außen- und Mittelohres

In der ersten Verarbeitungsstufe wird das Übertragungsverhalten des Außen- und Mittelohres nachgebildet. Das Ausgangssignal dieser Stufe wird physiologisch der Erregung am ovalen Fenster zugeordnet (vgl. Kap. 2.5.5.). Zwicker verwendet als Amplitudengang über 2kHz die inverse Ruhehörschwelle, unter 2 kHz wird 0 dB Verstärkung angenommen. Glasberg und Moore verwenden als Amplitudengang über 1 kHz die inverse Ruhehörschwelle und unter 1 kHz die inverse 100 phon Kurve nach ISO 226. Agerkvist verwendet die inverse 90 phon Kurve nach ISO 226 über den gesamten Frequenzbereich, da der Unterschied zur Ruhehörschwelle über 1 kHz nicht signifikant ist.

Alle drei Methoden sind problematisch, da der ISO 226 Standard mit großer Wahrscheinlichkeit, besonders im Bereich unter 1 kHz, falsch ist (vgl. Kap. 3.2.). Um im Erregungsmuster die Ruhehörschwelle nachbilden zu können, wird die Differenz aus dem jeweiligen Außen-Mittelohr Amplitudengang und der Ruhehörschwelle zum Erregungsmuster addiert (internes Rauschen der Cochlea, vgl. Abb. 2.24.).

Übertragungsverhalten der Cochlea

Im Verfahren nach Zwicker wird das Erregungsmuster aus Maskierungsfunktionen abgeleitet. Diese Methode ist problematisch, da die Bestimmung der Maskierungsmuster von Nebeneffekten wie der Detektion von Kombinationstönen und Amplitudenschwankungen sowie dem 'off-frequency listening' beeinflusst werden (vgl. Kap. 2.7.1.).

Die Frequenzskalierung bezieht sich bei Zwicker auf die kritische Bandbreite. Diese Skalierung beruht auf psychoakustischen Meßdaten, die besonders im tieffrequenten Bereich (unter 500 Hz) von Faktoren beeinflusst werden, die nicht mit der Frequenzselektivität in Verbindung stehen (vgl. Kap. 2.3.). Weiters wird (ähnlich einem Terzpegelspektrum) eine starre Barkzuteilung getroffen d.h. das Frequenzspektrum wird in vorgegebene Frequenzbänder unterteilt, deren Breiten jeweils einem Bark entsprechen. Es besteht also die Möglichkeit, daß ein schmalbandiges spektrales Maximum, dessen Breite kleiner als ein Bark ist, auf zwei Barkgruppen aufgeteilt wird. Tatsächlich richten sich die Frequenzgruppen nach den höchsten Pegeln ¹.

In den Modellen von Glasberg/Moore und Agerkvist wird das Erregungsmuster aus den Ausgangssignalen einer auditiven Filterbank bestimmt. Die verwendeten auditiven Filter beziehen sich auf die Ergebnisse von Notched-Noise Experimenten mit konstanten Maskiererpegeln d.h. es werden eingangspegelabhängige auditive Filter verwendet. Diese Methode ist seit 1994

¹ Auf die detaillierte Beschreibung des Erregungsmusters nach Zwicker wird verzichtet, da dieses grafische Verfahren als obsolet betrachtet werden kann. Das Prinzip dabei ist die Bestimmung des Erregungsmusters aus dem Maskierungsmuster. Diese Methode wurde [Zollner, 1996] übernommen und um die Nachverdeckung erweitert.

umstritten (vgl. Kap. 2.6.). Die Frequenzauflösung läßt sich bei diesen Modellen durch die Anzahl der verwendeten Filter verändern. Das Problem einer starren ERB-Breiten Zuteilung existiert nicht. Um auf das endgültige Erregungsmuster zu kommen, werden von den Ausgangssignalen der auditiven Filterbank die Einhüllenden gebildet und das interne Rauschen der Cochlea wird addiert.

3.6.2. Berechnung der Lautheit

Berechnung der spezifischen Lautheit

Die Grundannahme des Funktionsschemas zur Bildung der Lautheit besteht darin, daß die Lautheit N nicht aus der spektralen Zusammensetzung eines Schallsignals direkt entsteht, sondern daß die Gesamtlautheit aus dem Integral der spezifischen Lautheit N' über die Tonheit z gebildet wird¹. Es werden in einem ersten Schritt 'Teillautheiten' d.h. Lautheiten innerhalb der kritischen Bandbreiten² bestimmt. In einem zweiten Schritt werden diese Teillautheiten zu einer Gesamtlautheit integriert. Diese Vorgangsweise ist insofern sinnvoll, da sich das menschliche Lautheitsempfinden innerhalb einer kritischen Bandbreite nach eigenen Gesetzmäßigkeiten verhält.

Die Transformation vom Erregungsmuster zur spezifischen Lautheit läßt sich physiologisch etwa dem Transduktionsprozeß der Haarzellen zuordnen. Daraus ergibt sich eine Korrelation zwischen dem neuronalen Aktivitätsmuster und der empfundenen Lautheit.

Die spezifische Lautheit N' wird über ein Potenzgesetz mit der Erregung E in Beziehung gesetzt (Gl. 3.1.). Üblicherweise werden die Erregungen relativ zu $E_0 = 10^{-12}$ W angegeben.

$$N' \text{ prop. } \left(\frac{E}{E_0} \right)^k \quad \text{Gl. 3.1.}$$

Aus diesem Potenzgesetz leitet sich die spezifische Lautheit nach Zwicker ab³:

¹ vgl. [Zwicker, 1982] S. 130

² Die Begriffe 'kritische Bandbreite' und 'ERB-Breite' werden in den folgenden Ausführungen als gleichwertig behandelt, da sie in Hinblick auf die Berechnung der spezifischen Lautheit gleiche Qualitäten besitzen.

³ [Zwicker, Feldtkeller, 1967] S. 138

$$N'(z) = N'_0 \left(\frac{E_{RHS}(z)}{s(z) E_0} \right)^k \left[\left(1 - s(z) + \frac{s(z) E(z)}{E_{RHS}(z)} \right)^k - 1 \right] \quad \text{Gl. 3.2.}$$

Die Konstante N'_0 wird so gewählt, daß die Lautheit eines 1 kHz Signals mit einem absoluten Schallpegel von 40 dB einer Lautheit von 1 sone entspricht. E_{RHS} entspricht der Erregung an der Ruhehörschwelle. Der Exponent $k = 0,23$ wird aus den Lautheitsfunktionen eines 1 kHz Sinussignals und gleichmäßig anregendem Rauschen bestimmt¹.

Unter dem Schwellenfaktor (Detektionsverhältnis, Detektionseffizienz)

$$s = \frac{E_{\text{Testton}}}{E_{\text{Maskierer}}} \quad \text{Gl. 3.3.}$$

versteht man das Verhältnis aus der eben wahrnehmbaren Erregung eines Testtones zur Erregung eines Stör- oder Maskierungsrauschens innerhalb einer kritischen Bandbreite. Dieses Verhältnis ist sowohl pegel- als auch frequenzabhängig.

Für den Frequenzbereich um 1 kHz wird der Schwellenfaktor auf $s = 0,5$ gesetzt. Mit $N'_0 = 0,068$ wird Gl. 3.2. zu²:

$$N'(z) = 0,08 \left(\frac{E_{RHS}(z)}{E_0} \right)^{0,23} \left[\left(0,5 + \frac{0,5 E(z)}{E_{RHS}(z)} \right)^{0,23} - 1 \right] \quad \text{Gl. 3.4.}$$

Dieses Modell wurde von Moore und Glasberg³ für Erregungsmuster basierend auf auditiven Filtern adaptiert. Die ursprüngliche Frequenzskalierung in z (Bark-Skala) ändert sich dadurch zur ERB-Raten Skala e . Die Parameter $k = 0,23$ und $s = 0,5$ werden zunächst beibehalten, der Normierungsfaktor N'_0 ändert sich auf 0,084. Für den Frequenzbereich um 1 kHz wird aus Gl. 3.4.:

$$N'(e) = 0,084 \left(\frac{E_{RHS}(e)}{E_0} \right)^{0,23} \left[\left(0,5 + \frac{0,5 E(e)}{E_{RHS}(e)} \right)^{0,23} - 1 \right] \quad \text{Gl. 3.5.}$$

¹ vgl. [Zwicker, Fastl, 1990] Abb. 8.4. und 8.5.

² [Zwicker, Feldtkeller, 1967] S. 138

³ [Moore, 1986] S.256

In den beiden Abbildungen Abb. 3.20. und 3.21. sind einige Übergänge von Erregungsmustern zu spezifischen Lautheitsmustern grafisch dargestellt. Diese spezifischen Lautheitsmuster gelten für stationäre Schalle und wurden mit Gl. 3.5. berechnet.

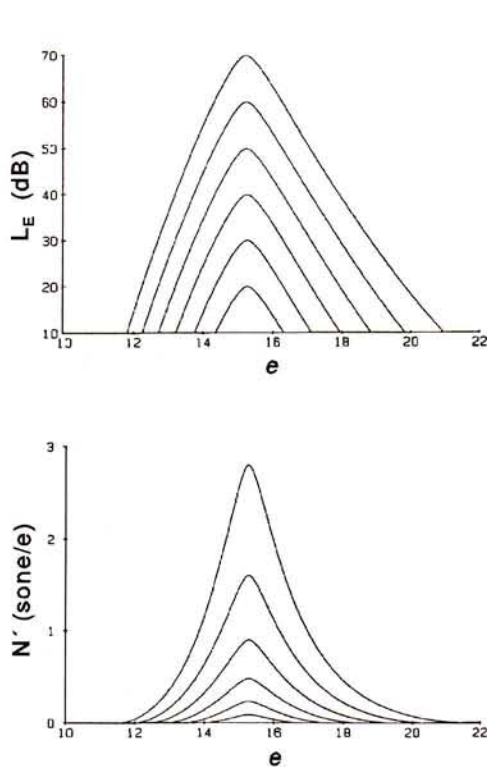


Abb. 3.20.: Obere Grafik: Erregungsmuster eines 1 kHz Signals mit 20 bis 70 dB; untere Grafik: entsprechende spezifische Lautheiten (aus [Moore, 1986] S. 258)

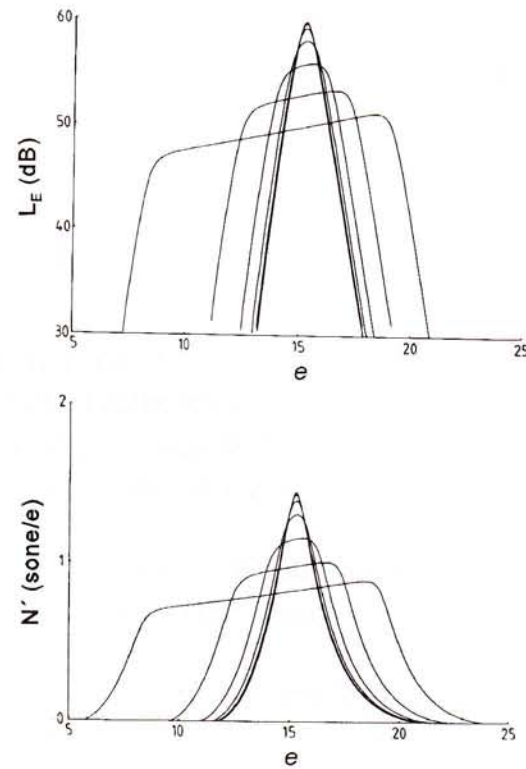


Abb. 3.21.: Obere Grafik: Erregungsmuster eines Signals mit 1, 2, 4, 8, 12, 16, 32, 64 Sinus-Komponenten; der Abstand der Komponenten beträgt jeweils 20 Hz; die Komponenten sind immer um 1 kHz zentriert. Die Gesamtleistung ist konstant d.h. unabhängig von der Komponenten-Anzahl. Untere Grafik: entsprechende spezifische Lautheiten (aus [Moore, 1986] S. 260)

Ein etwas unterschiedlicher Ansatz wurde 1996 von Glasberg und Moore¹ vorgestellt. In diesem Modell wird das zu analysierende Schallsignal zweigeteilt in Signal und Maskierer: $E = E_{\text{SIG}} + E_{\text{MASK}}$. Diese Unterteilung ermöglicht die Berücksichtigung gedrosselter Lautheiten.

Die spezifische Lautheit eines Testsignals (Erregung: E_{SIG}) in Ruhe d.h. ohne ein Stör- bzw. Maskierungssignal wird definiert als²:

¹ [Glasberg, Moore, 1996]

² [Stone et al., 1997]

$$N' = c \left[\left(\frac{E_{SIG}}{E_0} \right)^k - \left(\frac{E_{RHS}}{E_0} \right)^k \right] \quad \text{für } E_{SIG} > E_{RHS}$$

Gl. 3.6.

$$N' = 0 \quad \text{für } E_{SIG} < E_{RHS}$$

Die Parameter c und k sind konstant. Die spezifische Lautheit eines durch E_{MASK} teilweise maskierten Schalles E_{SIG} ist¹:

$$N'_{SIG} = c \left[1 - \left(\frac{sE_{MASK}}{E_{SIG}} \right) \right] \left[\left(\frac{E_{SIG}}{E_0} \right)^k - \left(\frac{E_{RHS}}{E_0} \right)^k \right]$$

Gl. 3.7.

Für die Berechnung der spezifischen Lautheit nach Gl. 3.7. benötigt man die getrennten Erregungsmuster E_{SIG} und E_{MASK} . Negative spezifische Lautheiten werden Null gesetzt (vgl. Gl. 3.6.); d.h. es werden nur spezifische Lautheiten mit $(E_{SIG} > E_{RHS})$ & $(E_{SIG} > sE_{MASK})$ summiert. Entsprechend wird also die spezifische Lautheit an der Ruhehörschwelle und an der Mithörschwelle des Maskieres gleich 0.

Das Detektionsverhältnis s wird bei Zwicker zwischen $s = 0,63$ (-2 dB) für tiefe Frequenzen, $s = 0,5$ (-3 dB) um 1kHz und $s = 0,25$ (-6 dB) an hohen Frequenzen interpoliert. Glasberg und Moore verwenden für Gl. 3.7. weit höhere Detektionsverhältnisse (bis zu $s = 5$ (+ 7 dB)). Diese Daten beruhen auf Notched-Noise Experimenten. Sie sind abhängig von der Versuchsperson, der Signallänge, der psychoakustischen Meßmethode und der verwendeten Hörschwelle.

Zwei Parameter der spezifischen Lautheit nach Gl. 3.7. müssen geschätzt werden, die Parameter k und c . Eine Möglichkeit den Exponenten k zu schätzen besteht in der Approximation der Pegeldifferenzen zwischen einem 1kHz Signal und gleichmäßig anregendem Rauschen (GAR) bei gleicher Lautheitsempfindung. Diese Methode wird von Zwicker² verwendet ($k = 0,23$). Glasberg und Moore approximieren den Exponenten k entsprechend der Lautheitsfunktion eines 1 kHz Sinussignals ($k = 0,2106$). Die Konstante c wird so gewählt, daß die Lautheit eines 1 kHz Sinussignals mit 40 dB absolutem Schallpegel eine Gesamtlautheit von 1 sone erhält ($c = 0.0806$).

Diese Verfahren gelten primär für stationäre Schalle, da sie auf experimentalpsychologischen Versuchsreihen beruhen, die mit stationären Schallen durchgeführt wurden. Eine Erweiterung des Gültigkeitsbereiches auf instationäre Schalle wird bei kurzen Ereignissen durch die Nachverdeckung und bei länger andauernden Schallereignissen durch Perzentilwerte

¹ [Glasberg, Moore, 1996]

² [Zwicker, Feldtkeller, 1967] S. 137

berücksichtigt¹.

Integration der spezifischen Lautheit

Die Gesamtlautheit N als Einzahlwert wird durch Integration der spezifischen Lautheit N' über die Frequenz bestimmt. Entsprechend den beiden unterschiedlichen Frequenzskalierungen z und e ergibt sich entweder

$$N = \int_0^{24} N'(z) dz \quad \text{oder} \quad N = \int_0^{40} N'(e) de \quad . \quad \text{Gl. 3.8.}$$

3.7. Zusammenfassung und Diskussion

Im Gegensatz zu den erhofften einfachen Zusammenhängen zwischen menschlichen Reaktionen und physikalischen Parametern muß man in der Psychoakustik meist von einer multivariablen Abhängigkeit ausgehen d.h. die menschlichen Reaktionen sind eine Funktion mehrerer physikalischer Parameter sowie der persönlichen Zustände einzelner Personen.

Die gesuchten Zusammenhänge werden in Experimenten mit Versuchspersonen gefunden, deren Eindrücke, Gedanken, Wahrnehmungen, Empfindungen und Verhaltensweisen erfaßt werden². Psychologische Faktoren wie die momentane psychische Verfassung der Probanden, ihre individuelle Empfindlichkeit sowie eine bestimmte Einstellung zum Informationsgehalt der Schallquellen bilden also ein wesentliches Kriterium.

Für den akustischen Umweltschutz bzw. für die Lärmmeßtechnik als dem zentralen Anwendungsgebiet der Lautheitsmeßtechnik geht man von folgenden Forderungen aus:

*' ... Durch die Emission von Schallquellen entsteht in der Nähe oder in größerer Entfernung von der Quelle eine Immission, die, soweit sich dort Personen aufhalten, die Personen belästigt, stört oder schädigen kann. ... Ein Schallbewertungsverfahren, das für die Lärmbekämpfung brauchbar ist, muß einerseits eine Geräuschkenngröße ergeben, die näherungsweise mit den möglichen Wirkungen durch Lärm gut korreliert, andererseits muß diese Geräuschkenngröße der Immission auf einfache Weise aus den Emissionskenngrößen der Quellen bestimmbar bzw. abschätzbar sein. ... '*³

Ein für die Lärmmeßtechnik adaptiertes Lautheitsbewertungsverfahren muß demzufolge drei

¹ *' ... Die wahrgenommene äquivalente Dauerlautheit ist wesentlich größer als der arithmetische Mittelwert der wahrgenommenen zeitvarianten Lautheit. Dies bedeutet, daß die wahrgenommene äquivalente Dauerlautheit von den lauten Ereignissen innerhalb der Lärmimmission stark beeinflusst wird. Die wahrgenommene äquivalente Dauerlautheit von Straßenverkehrslärm entspricht der Perzentillautheit N_4 , also derjenigen Lautheit, die in 4% der Meßzeit überschritten wird. ... '* ([Fastl, 1991]).

² vgl. [Schick, 1990] S. 16

³ [Lazarus, 1990]

Kriterien erfüllen. Es muß sowohl ein Maß für die Stör- als auch für die Schädigungswirkung sein und es müssen sich Immissionen aus Emissionen der Quellen berechnen lassen. Letzteres ist nach wie vor ein häufiges Argument für die bewerteten Schallpegel, da sich deren energetische Addition in der Praxis bewährt¹. Eine ähnlich einfache Addition von unterschiedlichen Lärmemitteln existiert für das Phon und das Sone zur Zeit nicht.

Die Möglichkeit, extraaurale physiologische Reaktionen auf Schall² als Maß für empfundenen Lärm zu verwenden, ist ebenso problematisch wie experimentalpsychologische Untersuchungen, da die Reaktionen des Organismus auf Lärm erhebliche individuelle Varianzen aufweisen, diese Reaktionen auch durch andere Stressoren ausgelöst werden können und der Informationsgehalt eine wesentliche Rolle spielt³.

Die von Zwicker vorgeschlagene 'unbeeinflusste Lästigkeit' (UBA ... 'unbiased annoyance')⁴ stellt einen Ansatz dar, die Lautheit als primäres Maß für die Lärm- bzw. Lästigkeitsempfindung abzulösen. Die UBA faßt die Lautheit (10% Perzentilwert), die Schärfe und Schwankungsstärke zu einem Einzahlwert zusammen. Dieser Versuch wurde sehr stark kritisiert⁵. Es ist tatsächlich

¹ '... Im Verfahren nach Zwicker ist folgender Effekt vorhanden: Werden Schallenergien innerhalb einer Frequenzgruppe hinzugefügt, wächst die Fläche und damit die Lautheit nur gering, da die vertikale Achse log. skaliert ist. Diese Addition ist energetisch wie beim dB(A). Treten jedoch zu einer einzelnen Frequenzgruppe weitere Frequenzgruppen hinzu, wächst die Fläche überproportional im Vergleich zur energetischen Addition. Dieser Effekt ist von Zwicker so gewollt und wird als Lautheitsaddition bezeichnet. Lautheiten verschiedener Frequenzgruppen addieren sich nach Zwicker im Prinzip linear und nicht energetisch-logarithmisch. Dies betont die Bandbreite eines Schalles. ... Die praktische Anwendung dieses Verfahrens führt zu unbrauchbaren Ergebnissen bei praktischen Maßnahmen der Lärminderung. So kann bei gleichbleibendem Schallenergieoutput eines Systems Lärminderung durch Konzentration auf eine oder wenige Terzbänder rechnerisch laut Zwicker Methode erreicht werden, obwohl eine solche Emission eine größere Reichweite und eine stärkere Durchschlagskraft gegenüber Dämmmaßnahmen besitzt. ...' ([Grigo, 1994]).

² '...extraaurale physiologische Lärmwirkungen, deren wichtigste kurz zusammengefaßt werden sollen:
 - Wirkungen auf das endokrine System ...
 - Wirkungen auf das Herz- Kreislaufsystem mit Änderungen der Herzfrequenz, Blutdruckveränderungen, Konstriktion peripherer Gefäße mit Herabsetzung der Hautdurchblutung und Hauttemperatur sowie Verkleinerung der Fingerpulsamplitude
 - Atemfrequenzänderung
 - Steigerung des Stoffwechsels
 - Abnahme der Tätigkeit der Verdauungsorgane
 - Änderung des Muskeltonus
 - Einfluß auf andere Sinnesorgane - wie Störung des Tiefen- und Farbsehens, Störung des Gleichgewichtes (Übelkeit) ...' ([Scheidt et al., 1986]).

³ vgl. [Scheidt et al., 1986]

⁴ '...Unbeeinflusste Lästigkeit wird definiert als Reaktion einer Person, die im Laborversuch unter beschreibbaren akustischen Bedingungen ausschließlich der Lästigkeit von Schall ausgesetzt ist und die keine Beziehung zur Schallquelle hat. Der Einfluß der Lautheit, der Tageszeit, der Schärfe des Schalls, der zeitlichen Variationen und der tonalen Komponenten wird in verschiedenen Untersuchungen bestimmt, ...' ([Zwicker, 1991]).

⁵ '... Das Konzept der 'unbeeinflussten Lästigkeit' von Zwicker wird auf der Grundlage begriffsanalytischer, methodologischer und empirischer Aspekte kritisiert. Es wird gezeigt, daß Zwickers Lästigkeitsbegriff gerade jene Komponenten der Lästigkeit wegdefiniert, die für das Leben außerhalb des Labors entscheidend sind - so z.B. die Beziehung zwischen Schallquelle und betroffener Person. ... Darüber hinaus muß bezweifelt werden, daß Zwickers Lästigkeitsmaß jemals valide Vorhersagen subjektiver

kritisch eine 'unbeeinflusste' Lästigkeit zu definieren, da der Begriff der Lästigkeit stark mit individuellen Faktoren verknüpft ist und sich kaum objektivieren läßt. Für eine 'unbeeinflusste' Meßgröße erscheint die Bezeichnung 'Lästigkeit' als unpassend. Versteht man diesen Begriff jedoch als abstrakte Bezeichnung für eine Erweiterung des Lautheitsmaßes in Richtung Lärmempfindung, so können damit in speziellen Anwendungsfällen durchaus genauere Lärmvorhersagen getroffen werden als mit einer reinen Lautheitsbewertung.

Der von der Grazer Firma AVL entwickelte 'AVL annoyance index'¹ zur objektiven Beurteilung von Motorengeräuschen ähnelt der 'unbeeinflussten Lästigkeit' von Zwicker insofern, als daß beide Verfahren das Hauptgewicht auf die Lautheit nach Zwicker legen und andere Empfindungsgrößen als Erweiterung verwenden. Beim AVL annoyance index wird die Lautheit mit 80 - 85% gewichtet. Die restlichen Empfindungsgrößen wie die wahrgenommene Periodizität, Rauigkeit, Schärfe und Impulshaltigkeit gehen mit 15 - 20 % in den annoyance Index ein. An diesem Beispiel ist die Dominanz der Lautheit ersichtlich.

Das A-bewertete Dezibel (dB(A)) ist als weltweiter Standard anerkannt und wird je nach Anwendungsgebiet mit entsprechenden Zu- bzw. Abschlägen verwendet. Diese Vorgangsweise wird sich in den nächsten Jahren voraussichtlich nicht ändern, wenngleich aufwendigere Verfahren wie die Verfahren nach Zwicker oder nach Stevens sich in speziellen Anwendungsgebieten wie der Flugzeug- oder Autoindustrie zur akustischen Optimierung weiter verbreiten werden. Es ist bis heute nicht gelungen, eine allgemein anerkannte Methode für die Lärmmessung zu entwickeln. Abgesehen von wirtschaftlichen und wirtschaftspolitischen Überlegungen ist dies der wesentlichste Grund, warum die als problematisch geltenden gesetzlichen Bestimmungen des akustischen Umweltschutzes bis zur heutigen Zeit nicht grundlegend revidiert wurden. Solange sich die Wissenschaft nicht auf die eine oder andere Methode zur Lärmmessung einigt, ist eine generelle Ablösung des dB(A) nicht zu erwarten.

Eine dynamische Phonkurvenanpassung ist durchaus sinnvoll, da diese Meßgröße im Unterschied zum dB(A) das gesamte Hörfeld berücksichtigt. Die Unterschiede zwischen dem dB(A) und dem hier vorgeschlagenen dB(ELC) sind teilweise sehr groß. Bei gleichen dB(A)-Werten bilden sich dB(ELC)-Differenzen von bis zu 20 dB(ELC) aus (vgl. Abb. 3.13.). Die teilweise großen Unterschiede zwischen dem dB(A) und dem dB(ELC) würden eine Revision des dB(A) durchaus rechtfertigen.

Als einziger Nachteil bleibt beim dB(ELC) die Nichtberücksichtigung der Lautheitszunahme breitbandiger Schalle gegenüber schmalbandigen Schallen aufrecht. Eine Differenzierung zwischen Freifeld- und Diffusfeld-dB(ELC)-Werten erscheint aufgrund der um vieles größeren Unterschiede zwischen dem dB(A) und dem dB(ELC) von sekundärer Bedeutung. Für die analysierten Schalle ist unter Verwendung einer Korrekturfunktion (Anhebung der Erwartungswerte auf das phonGF-Niveau) die Übereinstimmung zwischen dem phonStR und dem phonGF nicht besser, als zwischen den dB-Werten und dem phonGF (vgl. Abb. 3.15. bis 3.18.). Dieses Ergebnis überrascht, da das Verfahren nach Stevens eine ähnliche spektrale Behandlung wie das Zwicker-Verfahren vorsieht.

Lästigkeit im Feld gestattet und in irgendeiner Hinsicht Vorteile gegenüber derzeit gebräuchlichen Schalldosismaßen bietet. ...' ([Guski, Bosshardt, 1992]).

¹ [Schiffbänker et al., 1991], [Ronacher, Stücklschwaiger, 1994], [Ronacher, 1996]

Das dB(ELC) wurde in Hinblick auf die Lärmmeßtechnik entwickelt. Aus diesem Grund wurde das dB(A) in den Berechnungsablauf des dB(ELC) integriert. Die Frage, inwieweit sich eine Substitution des dB(A) durch das dB(ELC) in der praktischen Lärmmeßtechnik auswirken würde, kann aufgrund der durchgeführten Analysen nicht endgültig beantwortet werden.

Die Grundlage aller Lautheitsmeßverfahren bilden die Kurven gleicher Lautstärkepegel oder Daten vergleichbarer Versuchsreihen¹. Die Meßergebnisse verschiedener Versuchsreihen unterscheiden sich jedoch wesentlich. So konnte der ISO Standard 226 beispielsweise mit neueren Versuchsreihen nicht nachvollzogen werden. Um Replikationsstudien durchführen zu können, müssen sämtliche prozeduralen Details der Versuchsreihen übereinstimmen. Eine derartig engmaschige Definition der Versuchsdurchführung schränkt die Allgemeingültigkeit der ermittelten Daten wesentlich ein.

Es stellt sich die Frage, ob Laborversuche, die bei einer geringfügigen Änderung der Versuchsdurchführung (Bsp.: Änderung des Pegelintervalls in der Konstanzmethode vgl. Kap. 3.2.) zu stark unterschiedlichen Ergebnissen führen, sich für Freifeldmessungen überhaupt adaptieren lassen². Aus diesem Problem ergibt sich eine Unsicherheit die allen Lautheits- und dadurch auch Lärmmeßgrößen gemeinsam ist. Die Hoffnung auf eine endgültige Gestalt der KGL ist also eine Illusion, da eine 100%ige Bedingungskontrolle psychologisch nicht möglich ist. Die KGL können, ebenso wie die darauf beruhenden Lautheitsmeßverfahren, streng genommen nur Tendenzen anzeigen und keine exakten Meßdaten liefern.

Abgesehen vom Anwendungsgebiet der Lärmmeßtechnik besteht das primäre Ziel einer Lautheitsbewertung natürlich in der Nachbildung subjektiver Lautheitsurteile. Der Vorteil der Lautheitsempfindung ist, daß im Unterschied zur Lärmempfindung dem Informationsgehalt des Schalles weit weniger Gewicht zukommt und sich dadurch persönliche Einflüsse weniger stark auswirken. Das am schwierigsten zu realisierende Kriterium beim Entwurf eines Lautheitsbewertungsverfahrens ist die Allgemeingültigkeit. Für eine oder mehrere Gruppen von Schallarten können, abgesehen von den psychophysikalisch bedingten Varianzen, genaue Vorhersagen getroffen werden. Es ist aber bis heute nicht gelungen eine Methode zu entwickeln, die für alle möglichen Schallarten und -pegel valide Vorhersagen liefert.

¹ '... Die Phon-Kurven sind nicht mehr nur ein physikalisches Maß, sondern ein Maß, in dem physikalische Parameter mit Ergebnissen von Höruntersuchungen integriert sind; oder anders: Physikalische Größen werden mit Hilfe nicht-physikalischer Daten bewertet. Das Phon-Maß stellt deshalb auch einen entscheidenden Schritt zur Konstruktion weiterer Verfahren zur Lärmmessung dar. Dabei halten wir jedoch fest: Phon Kurven sind reine Lautstärkekurven, verwendet man sie als Bewertungskurven in Situationen, wenn der Schall stört und lästig ist, so sind sie nicht unmittelbar ein Maß für die Lästigkeit. ...' ([Schick, 1990]).

² '... Da die Ermittlung der KGL mit Wahrnehmungsuntersuchungen verbunden ist, bedeutet dies: Eine Kritik an diesem Verfahren macht meist auch eine Auseinandersetzung mit jenen psychophysischen Wahrnehmungsexperimenten sowie derer theoretischer Formulierungen erforderlich, welche den Bewertungsverfahren selbst zugrunde liegen. Denkt man beispielsweise an den Grundsatz der Methodenabhängigkeit von Phänomenen, so wird man begründet vermuten, daß je nach VP Stichprobe, Instruktionen, Darbietungsmethoden (auf- absteigende Verfahren, Konstanzverfahren, Herstellungsmethoden) und anderen experimentellen Anordnungen oft unterschiedliche Ergebnisse zu erwarten sind. ...' ([Schick, 1990]).

Das Anfang der 60er Jahre entstandene Lautheitsbewertungsverfahren nach Zwicker findet in zunehmendem Maße für spezielle Anwendungsgebiete wie der Entwicklung von geräuscharmen Verkehrsmitteln Verwendung; nicht zuletzt aufgrund der seit einigen Jahren verfügbaren echtzeitfähigen Meßstationen.

Eine grundlegende Neuerung dieser Methode wurde 1996 von Glasberg und Moore¹ vorgeschlagen. In dieser Revision wurde in erster Linie das Erregungsmuster verändert. Zwicker's Erregungsmuster ist insofern problematisch, da die Bestimmung der Maskierungsmuster von Nebeneffekten beeinflusst werden und es sich um eine starre Barkzuteilung handelt (vgl. Kap. 3.6.2.). Das Erregungsmuster nach Glasberg und Moore basiert auf Gammatone Filtern, wobei die Richtigkeit deren Amplitudengänge seit der Arbeit von Rosen und Baker² bezweifelt werden muß (vgl. Kap. 2.8.).

Ebenso wie eine valide Lärmmeßgröße eine valide Lautheitsmeßgröße bedingt, ist für die Berechnung der Lautheit das Erregungsmuster von entscheidender Bedeutung. Man muß also in erster Linie ein Erregungsmuster finden, dessen Zeit- Frequenz- Eigenschaften möglichst genau dem menschlichen Gehör entsprechen. Erst darauf aufbauend erscheint die Behandlung der Lautheitsberechnung aus dem Erregungsmuster sinnvoll. Erregungsmuster ergeben sich aus Modellen des peripheren Gehörs. Diese werden im folgenden behandelt.

¹ [Glasberg, Moore, 1996]

² [Rosen, Baker, 1994]

Kapitel 4

Modelle des peripheren Gehörs

Dieses Kapitel beinhaltet einen Überblick über Modelle des peripheren Gehörs. Es wird zwischen Modellen des Außen-, Mittel- und Innenohres unterschieden. Implementationsbeispiele und mögliche Systematisierungen auditiver Modelle werden diskutiert.

Die Entwicklung auditiver Modelle wird durch die Komplexität der einzelnen Verarbeitungsstufen des menschlichen Gehörs und deren Interaktionen stark erschwert. Man beschränkt sich daher meist darauf, einzelne Eigenschaften des Gehörs nachzubilden. Aufgrund unterschiedlicher Zielsetzungen wie beispielsweise automatisierte Sprachanalyse, Maskierungsmodelle zur Datenreduktion, Modelle der Lautheits-, Rauigkeits-, Schärfe- und Tonhöhenwahrnehmung, ergeben sich unterschiedliche Modelle. Dabei ist man bestrebt eine Darstellungsform zu erreichen, die der gesuchten Eigenschaft möglichst nahe kommt und das Ergebnis in einfach erfaßbarer Form wiedergibt. Bei der in Kapitel 3 behandelten Berechnung der menschlichen Lautheitsempfindung wird beispielsweise versucht, die Lautheitsempfindung auf einen objektivierten Einzahlwert zu reduzieren.

Den meisten auditiven Modellen ist die strukturelle Anlehnung an den anatomischen Aufbau des peripheren Gehörs gemeinsam¹. Dies läßt sich darauf zurückführen, daß sich die einzelnen Verarbeitungsstufen des peripheren Gehörs näherungsweise nachbilden lassen. Dies ist im Falle der zentralen Verarbeitung nicht möglich. Im peripheren Gehör wird bereits eine teilweise Analyse der Schallsignale und damit verbunden eine Datenreduktion durchgeführt. Dies bedeutet, daß nicht die gesamte akustische Information, die das Ohr erreicht, auch tatsächlich an das Gehirn weitergeleitet wird².

In Abb. 4.1. sind die Verarbeitungsstufen eines vereinfachten Modells der peripheren

¹ [Hawkins, et al., 1995] S.1-14

² Diese Erkenntnis macht man sich beispielsweise bei Datenreduktionsverfahren in Anlehnung an die Übertragungseigenschaften des peripheren auditiven Systems zunutze (z.B.: Mini Disk, DCC, MPEG Audio) ([Zollner, Zwicker, 1993] S.305-343).

Schallverarbeitung in Zusammenhang zum anatomischen Aufbau dargestellt:

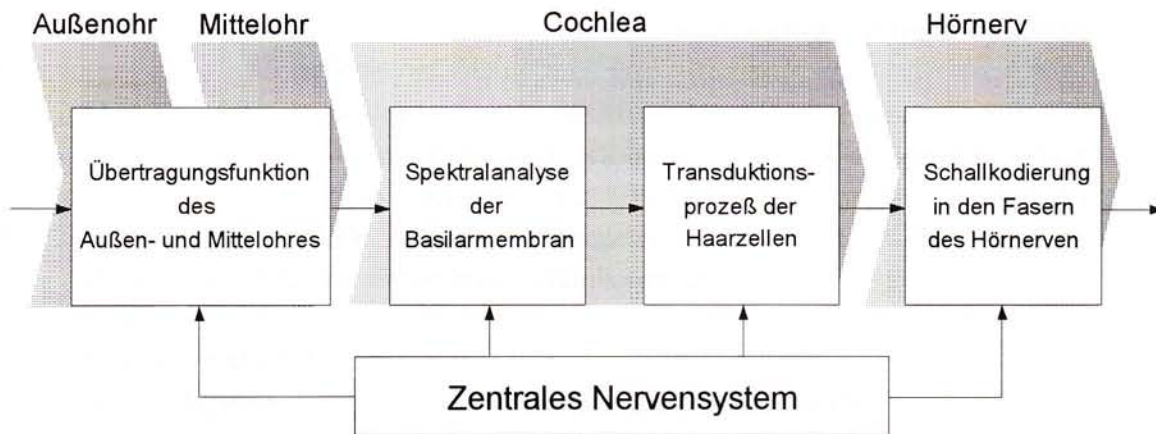


Abb. 4.1.: Verarbeitungsstufen zur Modellierung des peripheren Gehörs im Vergleich zur Anatomie

Interaktionen zwischen den einzelnen Verarbeitungsstufen sind in diesem Modell nicht dargestellt. Die Außen- und Mittelohrübertragungseigenschaften werden meist gemeinsam betrachtet. Die Spektralanalyseeigenschaften werden vorwiegend der Basilarmembran zugeordnet. Der Transduktionsprozeß der Haarzellen sowie die Schallkodierung in den Fasern des Hörnervs werden in vielen Modellen nicht berücksichtigt, da die Funktionsweisen dieser Verarbeitungsstufen nicht vollständig geklärt sind und sich deren Implementation äußerst aufwendig gestaltet.

Jede einzelne dieser Verarbeitungsstufen ist genau genommen nicht linear d.h. die Übertragungseigenschaften dieser Stufen werden vom zentralen Nervensystem und von Interaktionen zwischen den einzelnen Stufen beeinflusst. Die genaue Wirkungsweise dieser Rückkopplungszweige ist ebenfalls noch unklar. Für viele Modelle erscheint es nicht zielführend diese Regelschleifen mitzuberechnen, da der Aufwand dafür in keiner Relation zu den erreichten Verbesserungen stünde.

Unterscheidung auditiver Modelle

Die am häufigsten verwendete Unterscheidung auditiver Modelle differenziert zwischen funktionalen und physiologisch motivierten Modellen. Funktionale Modelle simulieren gewisse Eingangs - Ausgangs Eigenschaften des Gehörs, ohne auf die physiologischen Vorgänge des peripheren auditiven Systems einzugehen. Physiologisch motivierte Modelle versuchen die biophysikalischen Vorgänge möglichst exakt nachzubilden.

Weitere Unterscheidungsmöglichkeiten auditiver Modelle sind:

- Biophysikalisch - signalverarbeitungstheoretische Modelle¹

Diese Unterscheidung entspricht in etwa der Einteilung in funktionale und physiologische Modelle. Biophysikalische Modelle entsprechen den physiologischen Modellen und signalverarbeitungstheoretische Modelle entsprechen etwa den funktionalen Modellen.
- Analoge - digitale Modelle

Die Verwendung analoger Schaltungsstrukturen hat den Vorteil, daß die gesamte Signalverarbeitung a priori parallel erfolgt und sie deshalb im Gegensatz zu den Digitalrechnermodellen leichter für Echtzeitanwendungen verwendbar sind. Digitalrechnermodelle besitzen demgegenüber den Vorteil, daß sie eine größere programmtechnische Flexibilität ermöglichen und Simulationen leichter durchführbar sind.
- Aktive - passive Modelle

Aktive Modelle beziehen sich primär auf die Funktionsweise der äußeren Haarzellen. Diese Haarzellen stehen mit den efferenten Fasern des Hörnervs in Verbindung und beeinflussen aktiv das mechanische Verhalten der Basilarmembran.
- Lineare - nicht lineare Modelle

Das Verhalten des peripheren auditiven Systems ist von starker Nichtlinearität gekennzeichnet. Diese Nichtlinearität wirkt sich sowohl im Mittelohr als auch in der Cochlea aus. Während das nichtlineare Verhalten des Mittelohres für gewöhnlich vernachlässigt wird, stellen aufwendigere Modelle der Cochlea deren Nichtlinearität ansatzweise dar.
- Phänomenologische - rationale Modelle²

Diese, von Kates verwendete Unterscheidung nach eher philosophischen Termini, läßt sich natürlich auf auditive Modelle anwenden. Sie würde sich aber besser in eine wissenschaftstheoretische Arbeit integrieren.
- Echtzeitfähige- nicht echtzeitfähige Modelle

Die Echtzeitfähigkeit auditiver Modelle ist insofern von besonderem Interesse, da die Reaktionszeiten auf akustische Reize dem menschlichen Verhalten entsprechen sollten. Die dafür erforderliche Hard- und Software ist aufwendig. Aus diesem Grund sind die meisten Implementationen nach wie vor nicht echtzeitfähig.

Die Problematik dieser Einteilungen besteht darin, daß sich nur im Falle der technischen Unterscheidungskriterien (Bsp.: analoge - digitale Modelle) die Grenzen eindeutig bestimmen lassen. Diese Unterscheidungskriterien beinhalten jedoch keinerlei Bezug zur Psychoakustik und könnten für beliebige Modelle angewendet werden. Aus diesem Grund erscheint eine Einteilung in beispielsweise physiologische und funktionale Modelle passender.

Funktionale Modelle dürften allerdings im strengen Sinne der Definition keine separate Berücksichtigung des Außen- und Mittelohres beinhalten. Trotz dieser Unschärfe erscheint diese

¹ [Hawkins et al., 1995]

² [Kates, 1993]

weit verbreitete Einteilung durchaus sinnvoll und wird in dieser Arbeit aufrechterhalten. Ein Beispiel für die Unterscheidung in funktionale und physiologische Modelle ist die von Patterson entwickelte Software AIM (Auditory Image Model)¹. Sie stellt einzelne Module zur Verfügung, mittels derer die Zeit-Frequenz Darstellung auf der Ebene des Hörnervs simuliert werden kann. Dieses Modell wurde primär zur Spracherkennung entwickelt. Es setzt sich aus 3 grundsätzlichen Verarbeitungsstufen zusammen, die in einen funktionalen und einen physiologischen Zweig getrennt sind (Abb. 4.2.).

AIM - Modell

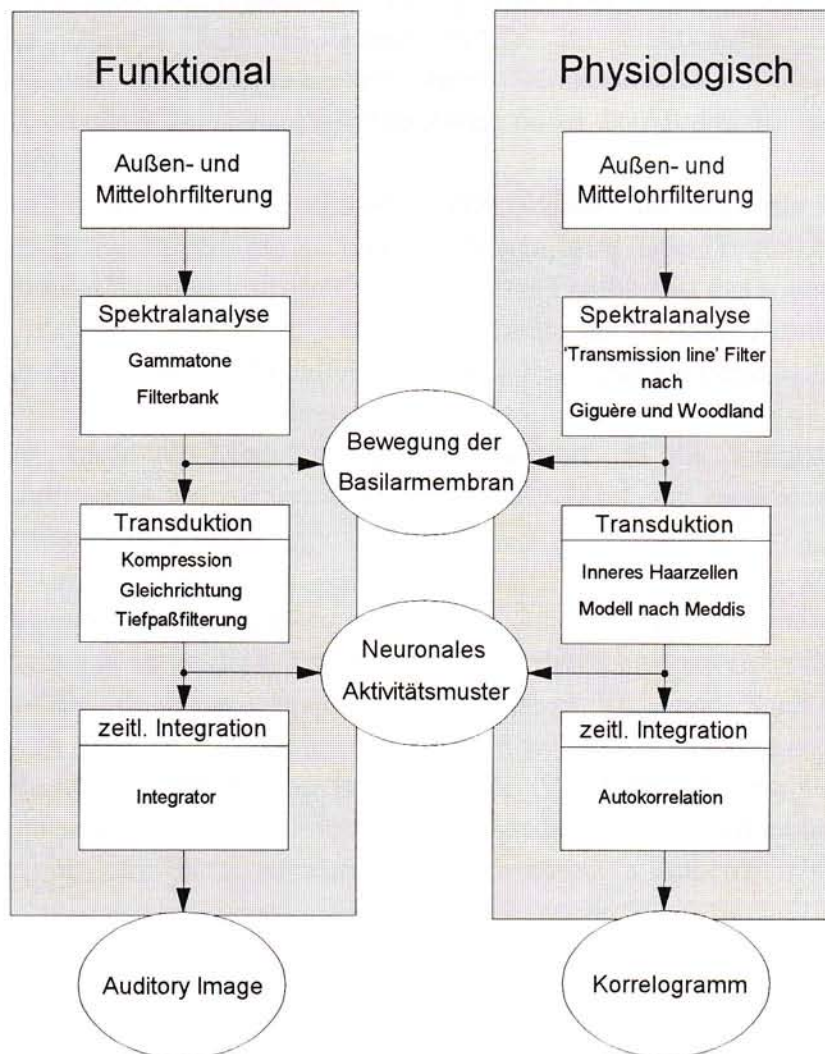


Abb. 4.2.: 'Auditory Image' Model (AIM) nach [Patterson et al. 1995]

¹ [Patterson et al., 1995]

Die Spektralanalysestufen konvertieren das Schallsignal in eine Darstellung der Basilarmembranbewegung. Diese Darstellung findet in einer Vielzahl von Modellen Verwendung. Die meisten auditiven Modelle beinhalten sogenannte 'frequenzbezogene Kanäle', wobei jeder einzelne Kanal das Verhalten des auditiven Systems an einem bestimmten Punkt der Basilarmembran modelliert.

Die Gammatonefilterbank als eine einfache Approximation der Basilarmembranbewegung ist linear ausgeführt, d.h. die Filter sind pegelunabhängig¹. Die physiologisch motivierte Spektralanalyse bezieht sich auf einen nichtlinearen 'transmission-line' Filter von Giguère und Woodland². Dieser Filterbank simuliert die hydrodynamischen Eigenschaften der Cochlea.

Die Transduktionsstufen simulieren den Übergang vom mechanischen in den neuronalen Bereich, der den inneren Haarzellen zugeordnet wird. Als Ergebnis dieser Stufen erhält man das neuronale Aktivitätsmuster (Neural Activity Pattern, NAP). Im funktionalen Zweig wird die Funktionsweise der inneren Haarzellen durch Komprimierung, Gleichrichtung und Tiefpaßfilterung simuliert. Im physiologischen Zweig findet das Modell der inneren Haarzellen von Meddis³ Verwendung.

Bei periodischen Signalen, wie musikalischen Tönen oszilliert das neuronale Aktivitätsmuster. Um der statischen Hörwahrnehmung Rechnung zu tragen, wird diese Darstellung durch ein sich zeitlich verschiebendes Integrationsfenster⁴ geglättet. Das Ergebnis ist das sogenannte 'Auditory Image'. Eine Alternative dazu stellt das Korrelogramm dar⁵. Durch Autokorrelation jedes einzelnen Kanals des neuronalen Aktivitätsmusters bleiben dabei die Periodizitätsinformationen erhalten.

Diese Software wurde als Modulkonzept implementiert, da die Validität der einzelnen Verarbeitungsschritte nach wie vor unklar ist. Patterson⁶ schlägt eine vergleichende Analyse der einzelnen Darstellungsformen vor.

Wesentliche Fragen bei der Entwicklung auditiver Modelle sind nach wie vor ungeklärt. Weder für die physiologisch noch für die funktional motivierte Betrachtungsweise existieren allgemein anerkannte Modelle, da jedes einzelne Modell gewisse Vor- und Nachteile besitzt. Die wissenschaftlichen Arbeiten konzentrieren sich häufig auf die Entwicklung neuer Modelle. Für viele dieser Modelle existieren nur unzureichende Vergleiche zu psychoakustisch oder physiologisch gemessenen Daten. Die Frage, welches Modell für eine bestimmte Anwendung am besten geeignet ist, kann nicht eindeutig geklärt werden. In Anbetracht der vorhandenen

¹ [Patterson et al., 1992]

² [Giguère, Woodland, 1994a], [Giguère, Woodland, 1994b], Die 'transmission line' Modelle stellen eindimensionale Simulationen der Wanderwellenausbreitung entlang der Basilarmembran dar. Diese Modelle werden in Kaskadenform realisiert (vgl. Kap. 4.2.3.1.)

³ [Meddis, 1988], [Meddis, Hewitt, 1991]

⁴ [Patterson, Holdsworth, 1991]

⁵ [Slaney, et al., 1994], [Slaney, 1996]

⁶ [Patterson, et al., 1995]

Varianzen und der nicht modellierbaren zentralen Verarbeitung erscheint beispielsweise die Simulation der Lautheit auf Ebene der Basilarmembran als ausreichend. Aus Gründen der Recheneffizienz und der größeren Nähe zu psychoakustisch ermittelten Lautheitsurteilen wird dafür meist ein funktionaler Ansatz verwendet.

4.1. Modelle des Außen- und Mittelohres

Über das Außen- und Mittelohr wird die eintreffende Schallenergie an die Cochlea weitergeleitet. Obwohl das Mittelohr durch die Mittelohrmuskeln aktive Komponenten besitzt, wird dessen Funktion von den passiven Eigenschaften dominiert. Aus diesem Grund wird das Mittelohr, ebenso wie das Außenohr, zumeist als passives Modell nachgebildet.^{1 2}

Die funktionale Nachbildung des Außen- und Mittelohres wird mit einem breitbandigen Bandpaßfilter durchgeführt. Der Amplitudengang dieses Filters wird entweder von den Kurven gleicher Lautstärkepegel oder vom 'gerade noch' wahrnehmbaren Schalldruck am Trommelfell abgeleitet (vgl. Kap. 2.5.5., 3.6.1. und 6.1.)³. Die Kurven gleicher Lautstärkepegel werden unter der Annahme eines freien Schallfeldes (Einfallrichtung 'von vorne', $\theta = \phi = 0^\circ$)⁴ bestimmt. Für diffuse Schallfeldbedingungen wird der Amplitudengang des Außen- Mittelohrfilters mit einer Diffusfeld-Korrektur adaptiert (vgl. Kap. 3.2.)⁵.

Diese Modelle können keinerlei Richtungsinformationen verarbeiten, sie sind nur für ein freies Schallfeld mit $\theta = \phi = 0^\circ$ oder für ein diffuses Schallfeld gültig. Trotzdem werden diese Modelle, wie beispielsweise in der Lärmmeßtechnik, für beliebige Schallfelder verwendet.

Die physiologisch motivierte Nachbildung der Außen- Mittelohrübertragung geht von einer Kaskade unabhängiger Prozesse aus. Die physiologischen Vorgänge werden dafür, in einem ersten Abstraktionsschritt, als mechanische Vorgänge beschrieben. In einer zweiten Stufe werden diese mechanischen Modelle, unter Verwendung der elektromechanischen Analogien, mit dem Instrumentarium der elektrischen Netzwerkanalyse und -synthese behandelt. Daraus ergibt sich der Vorteil, daß das Betriebsverhalten komplizierter mechanischer Schwingungsgebilde mit einfacher handzuhabenden elektrischen Schaltungen studiert werden kann^{6 7}.

¹ [Hawkins et al., 1995] S. 15-61

² '... Eine Vereinfachung der Modelle besteht darin, daß das Außen- und Mittelohr als linear betrachtet werden. Diese Vereinfachung besitzt für Schallpegel kleiner als 140 dB SPL Gültigkeit (Guinan and Peaks, 1967; Nedzelnitsky, 1980). Dieser Dynamikbereich beinhaltet den überwiegenden Anteil der natürlich vorkommenden Schalle. ...' ([Rosowski, 1991a]).

³ [Glasberg, Moore, 1990], [Killion, 1978], [Green et al., 1987]

⁴ Elevation = θ , Azimut = ϕ

⁵ [Takeshima et al., 1994]

⁶ [Zollner, Zwicker, 1993] S.123-148

⁷ '... Es liegt nahe, das hochentwickelte Instrumentarium der elektrischen Netzwerkanalyse und -synthese auch auf mechanische Netzwerke anzuwenden und die Bewegungsabläufe mechanischer Systeme mit entsprechenden elektrischen Schaltungen zu simulieren. Für den Elektrotechniker ergibt sich hierbei

Das Kriterium bei dieser Vorgangsweise ist der Übergang vom physiologischen in den mechanischen Bereich. Dabei müssen große Vereinfachungen in Kauf genommen werden, und die Allgemeingültigkeit dieser Modelle geht verloren.¹

Das erste Glied der physiologisch motivierten Kaskade bildet die Nachbildung des Körpers, des Kopfes und der Ohrmuscheln. Die akustische Wirkung dieser drei Komponenten beruht auf unterschiedlichen physikalischen Phänomenen wie Reflexion, Abschattung, Streuung, Beugung, Interferenz und Resonanz. Der Kopf und die Ohrmuscheln sind für das räumliche Hören von entscheidender Bedeutung. Am Eingang des äußeren Gehörganges ist die Transformation der räumlichen Schallfeldinformationen in die zeitliche Ebene vollständig durchgeführt. Diese Funktionalität läßt sich durch ein lineares Filter modellieren, dessen Übertragungsfunktion primär von der Richtung der Schallquelle abhängt (HRTF-Kurven, vgl. Kap. 2.2.).²

Der äußere Gehörgang ist durch das Trommelfell abgeschlossen und entspricht in seiner Wirkung etwa einer gedeckten Orgelpfeife (vgl. Kap. 2.2.). Durch die Schallausbreitung im Gehörgang wird das Trommelfell zum Schwingen angeregt. Üblicherweise ist die Eustachische Röhre geschlossen, d.h. der Mittelohrraum ist mit dem äußeren Schalldruck nicht verbunden. Aus diesem Grund kann das Trommelfell näherungsweise als ein Schalldruckempfänger betrachtet werden³. Die vom Trommelfell aufgenommene Luftschwingung wird über das Mittelohr an die Cochlea weitergegeben. Durch die mechanische Trägheit der Mittelohrknöchelchen entspricht der Amplitudengang dieser Verarbeitungsstufe einem Tiefpaß (vgl. Kap. 2.2.)⁴.

der Vorteil, daß das Betriebsverhalten komplizierter mechanischer Schwingungsgebilde anhand der ihm vertrauteren elektrischen Schaltung studiert werden kann. Darüber hinaus sind Änderungen leichter durchzuführen, wenn z.B. nicht ein Dämpfungselement auszuwechseln ist, sondern nur ein elektrischer Widerstand verändert werden muß. Allerdings kann nicht jeder mechanische Aufbau elektrisch simuliert werden: Bevor z.B. eine in vielen Moden schwingende Platte simuliert werden kann, müssen alle Schwingungsmoden bekannt sein. Es ist nicht möglich, eine durch Masse, Biegesteifigkeit und Dämpfung bestimmte Platte mit drei elektrischen Komponenten zu simulieren in der Hoffnung, daß dadurch höhere Schwingungsmoden der Platte ermittelt werden könnten. ...' ([Zollner, Zwicker, 1993] S. 123).

¹ [Zollner, Zwicker, 1993] S.123-148, [Hawkins et al., 1995] S.1-61

² [Blauert, 1974] S.50-51, [Rosowski, 1991a], [Rosowski, 1991b]

³ [Blauert, 1974] S.43-44

⁴ *'... Die Ergebnisse einer Leistungsflußanalyse zeigen, daß ein großer Anteil der Leistung am Trommelfell vom Mittelohr nicht absorbiert wird. Dies gilt hauptsächlich für tiefe Frequenzen. Zweitens konnte gezeigt werden, daß weniger als die Hälfte der auf das Mittelohr auftreffenden Leistung die Cochlea erreicht. ... Das Modell zeigt bei einer breitbandigen Anregung, eine starke Betonung des mittleren Frequenzbereiches. ...'* ([Rosowski, 1991a]).

4.2. Modelle der Cochlea

Die Spektralanalyseeigenschaften des menschlichen Gehörs werden zu einem großen Teil der Cochlea zugeordnet. Die mechanisch erzeugte Frequenzerlegung auf der Basilarmembran wird an den Hörnerv weitergegeben. Daraus ergeben sich drei mögliche Betrachtungsebenen der auditiven Spektralanalyse:

Spektralanalyse auf Ebene

- der Basilarmembran,
- des Hörnervs oder
- des Gesamtsystems (periphere und zentrale Verarbeitung).

Die Berücksichtigung des Transduktionsprozesses d.h. der elektro-mechanischen Transformation der Haarzellen ist aufwendig und wird in vielen Modellen der Cochlea nicht durchgeführt. Dies bedeutet, daß die meisten Simulationen der auditiven Spektralanalyse sich auf die Ebene der Basilarmembran beziehen.

Die hier behandelten Modelle der Cochlea werden in physiologisch und funktional motivierte Modelle eingeteilt. Entsprechend der gewählten Realisierungsform ergeben sich Abweichungen vom ursprünglichen Modell. Je stärker sich die Eigenschaften der Implementationsform auswirken, desto eher werden die Modelle den sogenannten 'Ingenieurlösungen' (engineering solutions) oder den 'Signalverarbeitungsmodellen' zugeordnet. Diese Einteilung, entsprechend Hubbard und Mountain¹, läßt sich mit der funktional-physiologischen Einteilung nicht vereinbaren, da Signalverarbeitungsmodelle sowohl funktionaler als auch physiologischer Natur sein können. Die in Kap. 4.2.3. dargestellten Implementationsbeispiele stellen durchwegs Signalverarbeitungsmodelle dar, wobei die Kaskadenmodelle sich vorwiegend auf physiologische Modelle stützen und die parallelen Filterbank- und Frequenzbereichsmodelle vorwiegend funktionalen Charakter aufweisen.

Der momentane Schwerpunkt der wissenschaftlichen Arbeiten auf dem Gebiet der Cochlea-Simulationen liegt im Bereich der nichtlinearen und aktiven Modelle. Weder die Nichtlinearität noch das aktive Verhalten der Cochlea sind physiologisch vollständig geklärt.

4.2.1. Physiologisch motivierte Modelle

Physiologische Modelle beziehen sich ausschließlich auf die biophysikalische Funktionsweise der Cochlea. Dabei unterscheidet man zwischen makromechanischen und mikromechanischen Modellen².

Bei den makromechanischen Modellen (1-, 2- oder 3- dimensional) wird das Verhalten der

¹ [Hawkins, et al., 1995] S.62-120

² Bsp.: [Lyon, 1996], [Lyon, Mead, 1988a], [Lyon, Mead, 1988b], [Kates, 1993], [Kates, 1995], [Ahn, Westerkamp, 1990], [Hawkins, et al., 1995], [Bollerott, et al., 1996], [Carney, 1993], [Zwicker, 1986], [Giguere, Woodland, 1994a], [Giguere, Woodland, 1994b]

Cochlea auf die hydrodynamischen Eigenschaften der Cochlea Flüssigkeit und die Basilarmembran reduziert. Die einzelnen Bestandteile des Cortischen Organs werden dabei nicht beachtet. Im Unterschied dazu ist man bei den mikromechanischen Modellen bestrebt, die Funktionsweise einiger anatomischer Details des Cortischen Organs zu berücksichtigen. Die physiologisch motivierten Modelle werden üblicherweise in Form von elektrischen Ersatzschaltbildern dargestellt.¹

4.2.2. Funktional motivierte Modelle

Funktional motivierte Modelle beziehen sich in erster Linie auf die Nachbildung experimentalpsychologischer Versuchsreihen. Im Unterschied zu den physiologischen Modellen wird hier aus dem Gesamtverhalten der auditiven Wahrnehmung auf die Eigenschaften der Cochlea geschlossen². Die Interpretation der psychoakustischen Daten wird meist unter Berücksichtigung physiologischer Erkenntnisse, wie beispielsweise der Wanderwellenausbreitung, entlang der Basilarmembran durchgeführt. Das wesentlichste Merkmal dieser Modelle ist, daß die detaillierten mechanischen Vorgänge nicht von Interesse sind.

Aufgrund der Ausgangssituation dieser Simulationen, die sich auf das auditive Verhalten in seiner Gesamtheit beziehen, werden teilweise Funktionalitäten in die Cochlea projiziert, die auf physiologischer Ebene nicht vorhanden sind. Ein funktionales Cochleamodell stellt also genau genommen keine Approximation der Cochlea, sondern eine Approximation der im auditiven System durchgeführten Spektralanalyse dar.

Die meisten Modelle der Cochlea sind funktionaler Natur. Sie sind für gewöhnlich recheneffizienter als physiologische Modelle und finden in den meisten Anwendungsbereichen der Psychoakustik Verwendung. In diesen Bereich fallen die sogenannten 'Gammatone Filter', die in den Kap. 5 und 6 separat besprochen werden.

4.2.3. Implementationsbeispiele

Die hier angeführten Implementationsbeispiele werden in drei Gruppen:

- Kaskaden-
- Parallelfilter- und
- Frequenzbereichsmodelle eingeteilt.

¹ [Giguere, Woodland, 1994a]

² Bsp.: [Patterson, et al., 1987], [Patterson, et al., 1992], [Patterson, Holdsworth, 1991], [Slaney, 1988], [Slaney, 1993], [Slaney, et al., 1994], [Slaney, 1996], [Cooke, 1991], [Cooke, et al., 1993], [Hawkins, et al., 1995], [Colomes, et al., 1995], [Agerkvist, 1994], [Moore, 1993]

Die Kaskaden-Modelle setzen sich meist aus einer Reihe von Tiefpaßfiltern zusammen, wobei jeder Filter eine etwas tiefere Grenzfrequenz besitzt als der vorherige. Die Gruppenlaufzeiten dieser Filter simulieren dabei die Ausbreitung der Wanderwellen entlang der Basilarmembran. Diese Kaskadenrealisierungen beziehen sich für gewöhnlich auf physiologische Modelle, weichen von diesen jedoch aus implementationstechnischen Gründen ab und werden deshalb oft den Ingenieurslösungen oder den Signalverarbeitungsmodellen zugeordnet.

Die Parallelfilter-Modelle bestehen aus einer parallelen Bandpaß Filterbank. Die Zeitverzögerungen der Cochlea werden ebenfalls durch die Gruppenlaufzeiten der Filter nachgebildet. Diese Modelle sind funktionaler Natur und werden ebenfalls den Ingenieurslösungen oder den Signalverarbeitungsmodellen zugeordnet.

Diejenigen Modelle, die sich keiner speziellen Filterstruktur bedienen, werden unter dem Überbegriff 'Frequenzbereichsmodelle' zusammengefaßt. In diesen Bereich fallen neben den üblichen Zeit-Frequenztransformationen, wie der short-time-Fourier oder der Wavelet Transformation, auch bilineare oder adaptive Analysen.

4.2.3.1. Kaskadenmodelle

Die Struktur eines Tiefpaß-Kaskadenmodells ist in Abb. 4.3. dargestellt:

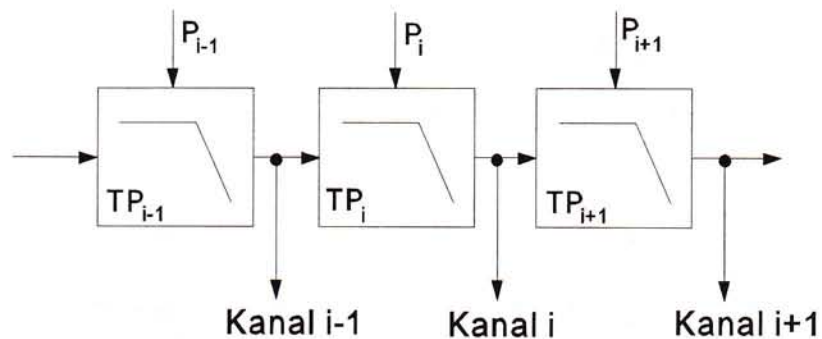


Abb. 4.3.: Cochleamodell in Form einer Tiefpaßkette

Die Tiefpaß-Kaskade simuliert die Wanderwellenausbreitung entlang der Cochlea. Eine Rechtfertigung für dieses Modell wird von Lyon und Mead¹ aus einer Lösung der 2-dimensionalen Wellengleichung der Cochlea abgeleitet. Demgegenüber bezeichnet Kates² diese Implementation als phänomenologisches Modell.

Die Grenzfrequenzen der Tiefpaßfilter verringern sich mit dem laufenden Filterindex i . Auf diese

¹ [Lyon, Mead, 1988a]

² [Kates, 1993]

Weise ergeben sich zwischen dem Eingang und den einzelnen Ausgangskanälen Bandpaßcharakteristika. Die Amplitudengänge dieser Bandpässe entsprechen den inversen Abstimmkurven. Durch die separate Parametrisierung jedes einzelnen Tiefpasses besteht die Möglichkeit das nichtlineare sowie aktive Verhalten der Cochlea zu simulieren. Für diese Parametrisierung, die sich äußerst komplex gestaltet, gibt es keinen allgemein anerkannten Lösungsvorschlag. Ein weiterer Nachteil dieser Modelle ist, daß sie von den zugrundeliegenden bio-physikalischen Grundlagen abweichen. Aus diesem Grund können nicht alle Schlußfolgerungen der Ausgangsmodelle für diese Implementation übernommen werden.

Sowohl das von Lyon vorgestellte 'Passive-Long-Wave-Model' (PLWM)¹ als auch dessen Weiterentwicklung unter dem Namen 'Active-Short-Wave-Model' (ASWM)² besitzen eine derartige Tiefpaß-Kaskade. Beide Modelle simulieren die Feuerungsraten auditiver Nervenfasern. Das PLWM basiert auf einer eindimensionalen, das ASWM auf einer zweidimensionalen hydrodynamischen Approximation der Cochlea. Die Pole und Nullstellen der Tiefpässe ergeben sich aus Vereinfachungen der Wellengleichungen. Die wesentlichste Neuerung des ASWM besteht in einer zusätzlichen Simulation der aktiven äußeren Haarzellen.

Abb. 4.4. zeigt die Modellierung der Basilarmembranbewegung des ASWM. Die Tiefpaß-Kaskade modelliert das Wanderwellenverhalten. Die AGC-Filter (Automatic Gain Control) integrieren die Energie des jeweiligen Ausgangskanals und steuern die Dämpfung der Basilarmembran. Die Parametrisierung der Tiefpaß-Filter bezieht sich auf die Polgüten.

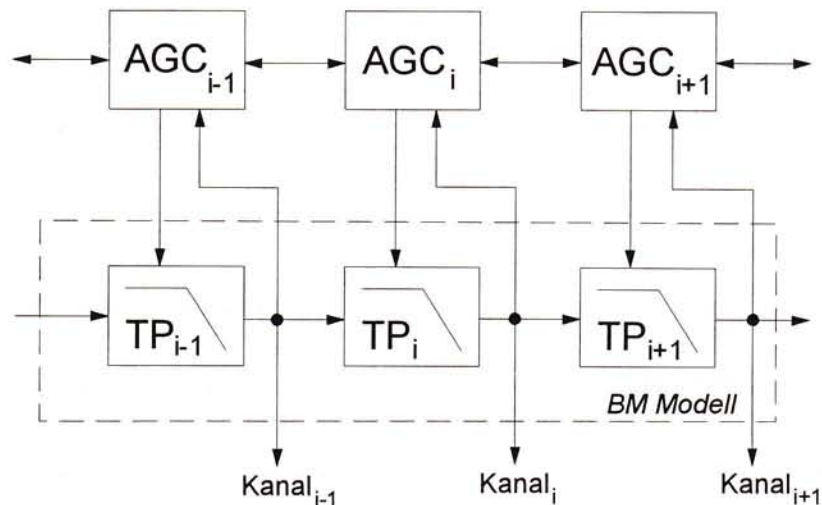


Abb. 4.4.: ASWM, Modellierung der Basilarmembran, vgl. [Cooke, et al., 1993], S.103

Das bekannteste Beispiel für eine ASWM Realisierung ist die sogenannte 'Silicon Cochlea' von

¹ [Lyon, 1982]

² [Cooke, et al., 1993], S. 95-116

Lyon und Mead¹. Diese analoge VLSI Realisierung setzt sich aus einer Kette von 480 aktiven Tiefpaß-Filtern 2ter Ordnung zusammen. Problematisch an diesem Modell ist, daß die Gruppenlaufzeiten zu groß sind.

Das Tiefpaßkaskadenmodell von Kates² basiert auf einer Kaskade von 112 Tiefpaß-Filtern (100 Hz bis 16kHz). Die Grundstruktur dieses Modells ist den Modellen von Lyon/Mead³ und Ahn/Westerkamp⁴ ähnlich. Im Unterschied zum Modell von Lyon/Mead kommt dieses Modell den gemessenen Gruppenlaufzeiten der Cochlea näher. Dies wird durch die Integration der Nullstellen in die Übertragungsfunktionen der Tiefpaß-Filter erreicht. Durch die zusätzlichen Nullstellen werden die Amplitudengänge jedoch breiter als die inversen Abstimmkurven. Der Aufbau einer einzelnen Verarbeitungsstufe dieses Modells ist, unter Einbeziehung der Transduktionsstufe, in Abb. 4.5. dargestellt:

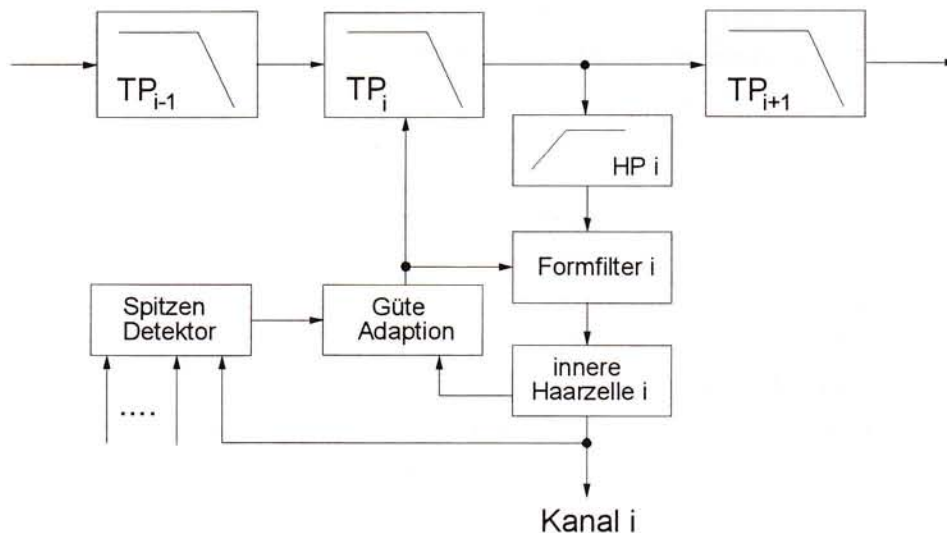


Abb. 4.5.: Kaskadiertes Tiefpaß-Modell nach Kates vgl. [Bollerott, et al., 1996]

Der Ausgang des Tiefpaßfilters wird mit einem Hochpaß gefiltert. Dieser Filter simuliert die Schallschnelletransformation⁵. Der zweite Filter (Formfilter) wirkt auf die 'Schallschnelle' und ergibt den mechanischen Ausgang der Cochlea (Spektralanalyse auf Ebene der Basilarmembran). Das Rückkopplungssystem als Approximation der Wirkungsweise der äußeren Haarzellen steuert

¹ [Lyon, Mead, 1988b]

² [Kates, 1993], [Kates, 1995]

³ [Lyon, Mead, 1988a]

⁴ Das Modell von Ahn und Westerkamp [Ahn, Westerkamp, 1990] bezieht sich auf einen Entwurf von Seneff aus dem Jahr 1988 [Seneff, 1988] und geht ebenfalls von einer Kette von Tiefpaßfiltern aus. Insgesamt werden 20 Kanäle im Bereich von 200 bis 3500 Hz verwendet.

⁵ [Tobias, 1970]

die Güten des Tiefpaß- und des Formfilters. Der 'Spitzen Detektor' als eine der beiden Steuerungen der Güte-Adaption berücksichtigt alle vorgeschalteten Filterstufen. Die Güte-Adaption wirkt als Pegelregelung und verleiht dem System eine Dynamikkompression von 2,5:1 entsprechend der gemessenen Kompression der Cochlea. Dieses System ist nicht echtzeitfähig. Bollerott¹ verwendet diese Grundstruktur zur Modellierung eines 16-kanaligen Cochlea-Implantats.

Eine weitere Simulation der Wanderwellenausbreitung entlang der Basilarmembran bietet das sogenannte 'transmission-line'-Modell. Dieses Modell basiert auf einer 1-dimensionalen tieffrequenten Approximation der Cochlea-Wellengleichung². Erweiterungen dieses Modells bezüglich des aktiven Verhaltens der Cochlea (Mikromechanische Resonanzen, Rückkopplungsmechanismus der äußeren Haarzellen) finden sich in den Arbeiten von Neely/Kim³ und Geisler⁴.

Der Nachteil der transmission-line Modelle wird von Kates folgendermaßen zusammengefaßt⁵:

- Das Bandpaßverhalten der einzelnen Kanäle ist, im Vergleich zu den inversen Abstimmkurven, an den Spitzen zu scharf, während die untere Flanke zu flach ist. Die damit erreichten Amplitudengänge entsprechen eher den neuronalen Abstimmkurven.
- Diese Modelle erfordern für gewöhnlich viele Filterstufen. Deng und Geisler verwenden beispielsweise 1400 Stufen bei einer Abtastrate von 100 kHz. Reduziert man die Stufen der transmission-line Modelle, so können sich bei niedrigpegeligen Signalen Instabilitäten ergeben.

Eine Alternative zu den Tiefpaß-Kaskaden- und transmission-line Modellen ist das 'modified transmission line'-Modell von Kates⁶. Dieses Modell stellt genauere Amplitudengänge und Gruppenlaufzeiten zur Verfügung. Die Tiefpaßfilter einer Tiefpaß-Kaskade werden durch die einzelnen Sektionen eines 1-dimensionalen transmission-line Modells ersetzt. Jede Stufe dieses Modells arbeitet weitgehend unabhängig⁷.

Die 'Zweitton-Unterdrückung'⁸ (two-tone suppression) kann selbst mit diesem Modell nicht

¹ [Bollerott et al., 1996]. In dieser Arbeit bezieht sich Bollerott auf die 'genaue' Approximation der inversen Abstimmkurven des Kates Modells. Kates selbst bezeichnet die Approximationen seines Modells hingegen als nicht genau ([Kates, 1993]).

² [Schroeder, 1973], [Zweig et al., 1976], [Zwicker, 1986], [Deng, Geisler, 1987]

³ [Neely, Kim, 1986]

⁴ [Geisler, 1991]

⁵ vgl. [Kates, 1993]

⁶ [Kates, 1993]

⁷ Jede einzelne Stufe wird ausschließlich von der vorhergehenden gesteuert. Als Abschlußwiderstand findet die charakteristische Impedanz der folgenden Stufe Verwendung.

⁸ Der Begriff Zweitton-Unterdrückung bezieht sich auf eine nichtlineare Eigenschaft der Cochlea, bei der die Feuerungsrate der einem Testton entsprechenden Hörnervfaser durch einen zweiten Testton unterschiedlicher Frequenz reduziert wird.

ausreichend nachgebildet werden¹.

4.2.3.2. Parallel - Filterbankmodelle

Bei diesen Modellen wird die Cochlea durch eine Bank paralleler Bandpaßfilter simuliert (Abb. 4.6.). In diesen Bereich fallen die seit langem verwendeten Terzbandfilter². Da die Terzbreite über 500 Hz etwa der Barkbreite entspricht, können Terzfilter als psychoakustisch motiviert bezeichnet werden.

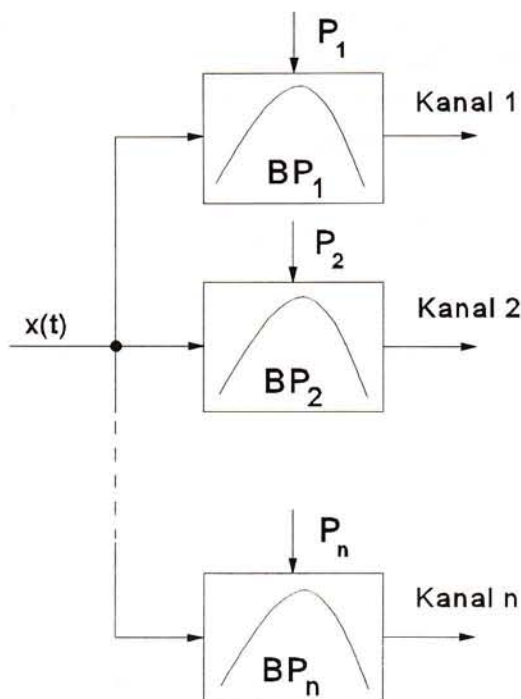


Abb. 4.6.: Paralleles Filterbank Modell

Die zur Zeit versiertesten Modelle dieser Art basieren auf den Gammatone Filtern. Diese Filter werden in den Kap.5 und 6 ausführlich diskutiert. Der wesentliche Unterschied dieser Modelle zu den Kaskadenmodellen besteht darin, daß das Wanderwellenverhalten der Cochlea nicht explizit nachgebildet wird. Dadurch ist der Bezug dieser Modelle zur physiologischen Ebene nicht in dem Maße vorhanden, wie bei den Kaskadenmodellen. Sie können daher eindeutig den funktionalen Modellen zugeordnet werden.

Das nichtlineare Verhalten der Cochlea wird durch die Steuerung der Filterparameter P (Abb. 4.6.) erreicht. Dieser Steuerparameter wirkt auf die tieffrequente Flanke der Bandpässe. Die Parametrisierung erfolgt meist in der s -Ebene. In den Arbeiten von Lyon werden die unterschiedlichsten Möglichkeiten zur Parametrisierung in der s -Ebene diskutiert³.

Parallel-Filterbankmodelle finden beispielsweise in den Bereichen der psychoakustisch orientierten Datenreduktionsverfahren⁴, bei der Analyse von musikalischen Schallsignalen⁵ und bei der Qualitätsbestimmung von Audiogeräten⁶ Verwendung. Diese Modelle stellen keine

¹ [Kates, 1995]

² [Zollner, 1996]

³ [Lyon, 1982], [Lyon, 1990], [Lyon, 1996]

⁴ [Colomes, et al., 1995]

⁵ [Eckel, 1989]

⁶ [Beerends, Stemerding, 1992]

exakten Nachbildungen der peripheren Schallverarbeitung dar, sie bilden jedoch wesentliche Eigenschaften des Gehörs nach.

4.2.3.3. Frequenzbereichsmodelle

Unter dem Überbegriff 'Frequenzbereichsmodelle' werden hier diejenigen auditiven Modelle zusammengefaßt, die im Unterschied zu den Filtermodellen in einer ersten Verarbeitungsstufe eine Zeit-Frequenz Analyse beinhalten, die nicht dem auditiven Modell selbst zugeordnet wird (vgl. Abb. 4.7.). Die dafür verwendeten Transformationen werden meist unter Berücksichtigung des menschlichen Zeit- und Frequenzauflösungsvermögens implementiert.

Diese Modelle sind rein funktionaler Natur, da sie vom anatomischen Aufbau des peripheren Gehörs grundsätzlich abweichen, das Wanderwellenverhalten der Cochlea nicht nachbilden und sich das Zeit- Frequenzauflösungsvermögen auf experimentalpsychologische Daten bezieht.

In der Arbeit von Agerkvist¹ werden die Anwendungsmöglichkeiten unterschiedlicher Zeit-Frequenz Analysemethoden auf auditive Modelle diskutiert. Es werden lineare, bilineare und adaptive Analysemethoden gegenübergestellt. Jede dieser drei Methoden besitzt bestimmte Vor- und Nachteile. Agerkvist verwendet die lineare 'wavelet-packet' Analyse, da sich diese Analyse recheneffizient implementieren läßt und das Zeit-Frequenzauflösungsvermögen an das menschliche Gehör angenähert werden kann.

Abb. 4.7. zeigt das Blockschaltbild eines eingangspegelabhängigen Frequenzbereichsmodells.

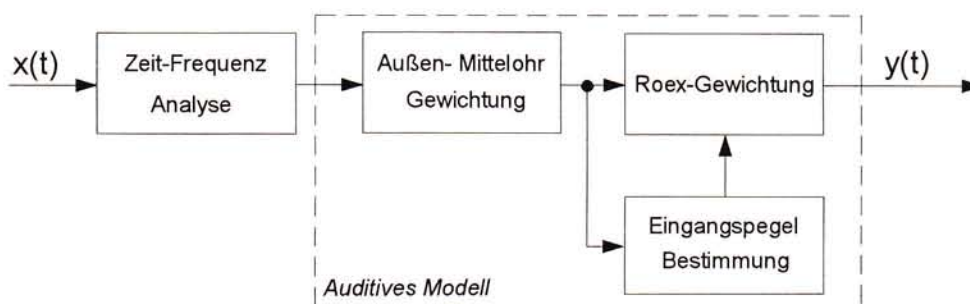


Abb. 4.7.: Eingangspegelabhängiges nichtlinearers Frequenzbereichsmodell vgl. [Agerkvist, 1994]

Die drei wesentlichen Vorteile dieses Modells bestehen darin, daß jede beliebige Zeit-Frequenzanalyse verwendet werden kann, die roex-Funktionen (vgl. Kap. 2.5.1.3.) exakt nachgebildet werden können und sich das nichtlineare Verhalten einfach implementieren läßt.

¹ [Agerkvist, 1994]

Der Nachteil dieses Modells ist, daß die Zeit-Frequenzanalyse aus dem auditiven Modell ausgekoppelt wird und eine echtzeitfähige Implementation aufwendig ist.

Die roex-Gewichtung bezieht sich auf eine kanalbezogene Betrachtungsweise der peripheren auditiven Schallverarbeitung (vgl. Kap. 2.4.). Die Kanäle entsprechen den Ausgängen der auditiven Filter. Diese Modelle sind insofern den parallelen Filterbank Modellen ähnlich.¹

¹ [Eckel, 1989], [Colomes, et al., 1995], [Beerends, Stemerdink, 1992], [Sporer, et al., 1995], [Moore, 1993]

Kapitel 5

Gammatone Filter

In diesem Kapitel wird die Gammatone Funktion und deren Anwendungsmöglichkeiten auf Modelle des menschlichen Gehörs diskutiert. Es wird zwischen linearen und nichtlinearen Gammatone Filterbänken unterschieden.

Der Ursprung der Gammatone-Funktion geht auf die 70er Jahre zurück¹. Der Name leitet sich aus der Form dieser Funktion ab. Ein Cosinus-Signal ('tone') ist mit einer *Gamma*-Funktion amplitudenmoduliert (Gl. 5.1.).

$$g(t) = at^{n-1} e^{-2\pi bt} \cos(2\pi f_c t + \phi) \quad (t > 0) \quad \text{Gl. 5.1.}$$

Diese Funktion stellt die Impulsantwort eines Gammatone Filters (GTF) dar. Der Parameter b bestimmt größtenteils die Länge der Impulsantwort und damit die Bandbreite des Filters. Die Filterordnung n bestimmt zum überwiegenden Teil die Flankensteilheiten. Die Filtermittenfrequenz wird mit dem Parameter f_c festgelegt.

Für Filterordnungen zwischen $n = 3 - 5$ entsprechen die Amplitudengänge dieser Filter etwa den psychoakustisch ermittelten roex(p)-Filtern (vgl. Kap. 2.5.)².

Da die Gammatone Funktion sich aus physiologisch gemessenen Daten entwickelte und der Amplitudengang dieses Filters die experimentalpsychologisch ermittelte Frequenzauflösung wiedergibt, stellen die Gammatone Filter ein Verbindungsglied zwischen der Physiologie und der Psychoakustik dar. Durch die übersichtliche Beschreibung dieses Filters in Termen der

¹ Die Entwicklungsgeschichte der Gammatone Funktion ist in [Cooke, 1991] S.16 und [Lyon, 1996] S. 9 detailliert dargestellt. Ursprünglich diente diese Funktion zur Beschreibung von 'revcor' Daten, die eine Näherung der Impulsantwort einer einzelnen Faser des Hörnervs darstellen.

² [Patterson et. al., 1987] zeigen, daß mit einem Gammatone Filter 4ter Ordnung roex(p)-Formen über einen Dynamikbereich von etwa 60 dB nachgebildet werden können.

Impulsantwort kann dieser Filter leicht an gemessene Daten angepaßt werden.

Der GTF ist minimalphasig. Seine Impulsantwort besitzt kurze Anstiegs- und lange Abfallzeiten. Dies wird als ein wesentlicher Vorteil gegenüber linearphasigen Filtern betrachtet¹.

Die beiden Nachteile des GTF sind sein annähernd symmetrischer Amplitudengang und eine schwierige Parametrisierung für pegelabhängige Änderungen des Amplitudenganges.

Die Ausgänge (bzw. Kanäle) einer Gammatone Filterbank (Parallel-Filterbank-Modell, vgl. Kap. 4.2.3.2.) entsprechen einer Zeit- Frequenzdarstellung auf der Basilmembranebene. Die Leistungen werden als Erregungsmuster interpretiert. Sie bilden das Frequenzauflösungsvermögen des menschlichen Gehörs als rein funktionales Modell nach. Die übliche Bezeichnung dieser Filterbänke als Cochlea Modelle besitzt daher nur bedingt Gültigkeit (vgl. Kap. 4.2.)

5.1. Lineare Gammatone Filter

Da der Parameter b aus Gl. 5.1. zum überwiegenden Teil die Bandbreite des Gammatone Filters bestimmt, wird dieser Parameter aus den ERB-Breiten berechnet. Üblicherweise werden lineare Gammatone Filterbänke mit Filtern 4ter Ordnung ($n = 4$) und $b = 1.019 \cdot \text{ERB}(f_c)$ verwendet (vgl. Kap. 2.6.). Die 3 dB Bandbreite dieser Filter entspricht 0.887 mal der ERB-Breite².

Die verwendeten ERB-Breiten unterscheiden sich nicht wesentlich voneinander. Am häufigsten wird auf die ERB-Breiten nach [Glasberg, Moore, 1990] (vgl. Kap. 2.6.1.) referenziert.

Aus einer allgemeinen ERB-Breiten Gleichung

$$\text{ERB}(f, A, B, x) = \left(\left(\frac{f}{A} \right)^x + B^x \right)^{\frac{1}{x}} \quad \text{Gl. 5.2.}$$

lassen sich diese ERB-Breiten mit: $A = 9.264$, $B = 24.7$ und $x = 1$ bestimmen. Im Vergleich dazu verwenden Lyon $A = 8$, $B = 125$, $x = 2$ und Greenwood $A = 7.238$, $B = 22.85$, $x = 1$.³ Diese drei Bandbreiten sind in Abb. 5.1. dargestellt. Die von Lyon verwendeten Bandpaßfilter beziehen sich offensichtlich auf die kritischen Bandbreiten (vgl. Kap. 2.3.), während die beiden anderen Autoren sich auf die ERB-Breiten (vgl. Kap. 2.6.1.) beziehen. In dieser Arbeit werden die Parameter von Glasberg und Moore verwendet.

¹ [Patterson et al., 1987]

² [Patterson, 1994]

³ [Slaney, 1993]

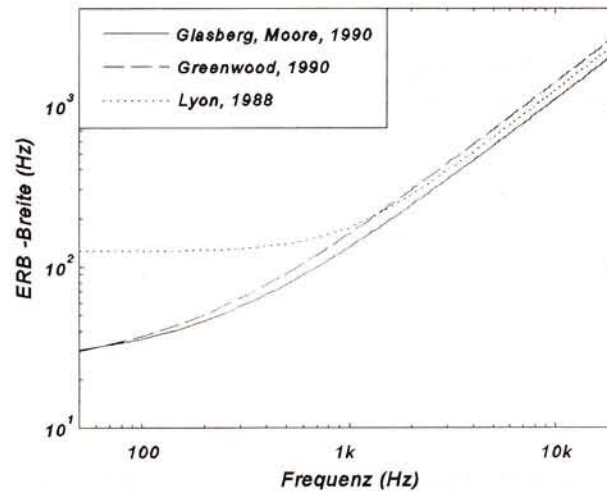


Abb. 5.1.: Vergleich unterschiedlicher ERB-Breiten - Glasberg/Moore, Greenwood, Lyon. vgl. [Slaney, 1993]

Aufteilung der Gammatone Filter (Mittenfrequenzen) über den betrachteten Frequenzbereich:

Unter Bezugnahme des Energieflusses vom Eingang der Cochlea (hohe Frequenzen) hin zum Helicotrema (tiefe Frequenzen) gibt Slaney die Aufteilung der Filtermittenfrequenzen nach folgender Funktion vor:¹

$$f_c(m) = -(AB) + e^{\frac{m(-\ln(f_s/2+AB) + \ln(f_{c,\min}+AB))}{k}} (f_s/2 + AB)$$

f_c ...Filtermittenfrequenzen

$f_{c,\min}$...tiefste Filtermittenfrequenz

f_s ...Abtastrate

k ...Anzahl der Filter im Bereich zwischen $f_{c,\min}$ und $f_s/2$

m ...Filterindex (Kanalnummer $m = 0$ entspricht der halben Abtastrate)

$AB = 9.264 * 24.7 = 228.82$

Gl. 5.3.

Diese Gleichung läßt sich vereinfachen zu:

$$f_c(m) = -AB + \left(\frac{f_{c,\min} + AB}{f_s/2 + AB} \right)^{\frac{m}{k}} (f_s/2 + AB)$$

Gl. 5.4.

Gleichungen 5.3. (bzw. 5.4.) kann mit folgendem Ansatz nachvollzogen werden:

¹ [Slaney, 1993]

- 1) Bestimmung der ERB-Raten für die niedrigste Filtermittenfrequenz $e(f_{c,\min})$ und die halbe Abtastrate $e(f_{s/2})$ mit Gl. 2.18.b.
- 2) Teilung des ERB-Raten Intervalls zwischen $e(f_{c,\min})$ und $e(f_{s/2})$ in k gleiche Teile. Daraus ergeben sich die gesuchten Filtermittenfrequenzen im ERB-Raten-Bereich:

$$E_c(m) = e(f_{c,\min}) + m \left(\frac{e(f_{s/2}) - e(f_{c,\min})}{k} \right)$$

E_c ... gesuchte Filtermittenfrequenzen im ERB-Raten-Bereich
 m ... Filterindex ($m=0$ bis $k-1$)
 $e(f)$... ERB-Rate der Frequenz f nach Gl. 2.18.b.

Gl. 5.5.

Diese Mittenfrequenzen lassen sich mit Gl. 2.18.b. in den Hz-Frequenzbereich zurückrechnen. Das Ergebnis lautet:

$$f_c(m) = \frac{(0.00437f_{c,\min} + 1) \left(\frac{0.00437f_{s/2} + 1}{0.00437f_{c,\min} + 1} \right)^{\frac{m}{k}} - 1}{0.00437}$$

Gl. 5.6.

Gl. 5.6. liefert exakt dieselben Ergebnisse wie Gl. 5.3. (bzw. 5.4.). Der einzige Unterschied besteht darin, daß sich bei Gl. 5.6. die kleinste Filtermittenfrequenz nicht bei $m = k$, sondern bei $m = 0$ befindet.

Ein Zusammenhang zum 'Energiefluß vom Eingang der Cochlea hin zum Helicotrema' kann nicht nachvollzogen werden.

Abb. 5.2. zeigt die Filtermittenfrequenzen (in Hz, log. Ordinaten-skalierung) nach Gl. 5.4. für $f_s = 32\text{kHz}$, $f_{c,\min} = 30\text{ Hz}$ und $k = 40$.

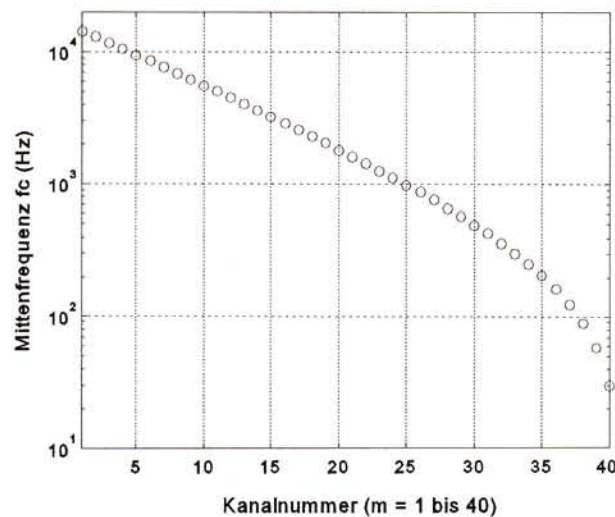


Abb.: 5.2.: Mittenfrequenzen der Gammatone Filter nach Gl. 5.4. mit $f_s = 32\text{kHz}$, $f_{c,\min} = 30\text{ Hz}$ und $k = 40$

Die Realisierung der Gammatone Filter mit digitalen rekursiven Filterstrukturen wird üblicherweise in zwei Schritten durchgeführt:

- 1) Bestimmung der analogen Übertragungsfunktion $H(s)$ mit der Laplacetransformation
- 2) Transformation vom s - in den z -Bereich

Cooke¹ diskutiert 3 mögliche Transformationen vom analogen in den digitalen Bereich, die Pol-Nullstellenübertragung², die Bilineartransformation³ und die Impulsinvarianztransformation⁴. Bei allen drei Methoden setzt er eine Frequenzverschiebung um die Mittenfrequenz ω_c auf $\omega_c = 0$ voraus. Dadurch reduziert sich der ursprüngliche Bandpaßfilter auf einen Tiefpaßfilter, der sich leichter in die z-Ebene transformieren läßt. Das auf diese Art gefilterte Signal muß anschließend wieder auf die ursprüngliche Filtermittenfrequenz zurückmoduliert werden. Cooke kommt zu dem Ergebnis, daß die Impulsinvarianztransformation der Impulsantwort und damit dem analogen Amplituden- und Phasengang am besten entspricht.

Die von Slaney⁵ verwendete Methode stützt sich ebenfalls auf die Impulsinvarianztransformation. Im Unterschied zu Cooke (Filter 4. Ordnung mit komplexen Koeffizienten) verwendet Slaney Filter 8. Ordnung mit reellen Koeffizienten, wodurch der Rechenaufwand für die Filterung selbst gleich bleibt. Die Methode von Slaney ist insofern effizienter, da bei dieser Methode die Modulation (Frequenzverschiebung um ω_c) vor und nach der Filterung wegfällt.

Die von Slaney entworfenen digitalen Filter werden in dieser Arbeit als Basis für eine nichtlineare Gammatone Filterbank verwendet. Aus diesem Grund sind die Koeffizienten dieser Filter im folgenden angeführt. Jeder Gammatone Filter setzt sich aus einer Kaskade von 4 Filtern jeweils zweiter Ordnung zusammen. Diese 4 Filter realisieren also einen Gammatone Filter 4ter Ordnung mit einer digitalen Filterstruktur 8ter Ordnung.

Die allgemeine Form der Gammatone Filter nach Slaney lautet:

$$H(z) = \prod_{i=1}^4 H_i(z) = \prod_{i=1}^4 \frac{b_{0,i} + b_{1,i} z^{-1} + b_{2,i} z^{-2}}{1 + a_{1,i} z^{-1} + a_{2,i} z^{-2}} \quad \text{Gl. 5.7.}$$

Die 4 Filterstufen unterscheiden sich lediglich durch die Koeffizienten $b_{1,i}$ (d.h. durch ihre

¹ [Cooke, 1991] S.9-32

² Pol- Nullstellenübertragung: Alle Polstellen und endlichen Nullstellen der analogen Übertragungsfunktion werden mit $z = e^{sT}$ in den z-Bereich übertragen. Alle Nullstellen im Unendlichen werden auf $z = -1$ übertragen. Der fehlende Verstärkungsfaktor wird durch Anpassung an einer wesentlichen Frequenz bestimmt. Bei dieser Methode ist es meist der Fall, daß eine der zwei Filtercharakteristika (Amplitudengang, Phasengang) eine starke Abweichung vom analogen Filter aufweist.

³ Die Bilineartransformation ist definiert als: $s = 2/T \left((1 - z^{-1}) / (1 + z^{-1}) \right)$. Die imaginäre Achse der s-Ebene wird auf den Einheitskreis der z-Ebene komprimiert. Die linke Hälfte der s-Ebene fällt in den Einheitskreis. Durch die Komprimierung der unendlich langen Frequenzachse aus der s-Ebene auf den Einheitskreis ($f = 0 - fs/2$) der z-Ebene treten bei dieser Methode Frequenzverzerrungen auf.

⁴ Bei der Impulsinvarianztransformation wird die analoge Impulsantwort abgetastet und anschließend in den z-Bereich transformiert. Dabei werden nach einer Partialbruchzerlegung die einzelnen Terme von $H(s)$ durch die Substitution sp Der Nachteil dieser Methode besteht darin, daß durch die Abtastung der Impulsantwort bei nicht hinreichend bandbegrenzten Filtern Frequenzaliasing entsteht.

⁵ [Slaney, 1993]

Nullstellen).

$$\begin{aligned}
 b_{0,i} &= T & b_{2,i} &= 0 \\
 a_{1,i} &= \frac{-2 \cos(2\pi f_c T)}{e^{BT}} & a_{2,i} &= e^{-2BT} \\
 b_{1,i} &= \left(\frac{-T \cos(2\pi f_c T)}{e^{BT}} + (-1)^i \frac{\sqrt{3+m_i} 2^{1.5} T \sin(2\pi f_c T)}{e^{BT}} \right) & & \text{Gl. 5.8.}
 \end{aligned}$$

$T = 1/f_s \dots$ Abtastperiode
 $B = 2\pi \cdot 1.019 \text{ERB}(f_c) \dots$ Filterbandbreite
 $m_i = 1, 1, -1, -1$
 $i = 1, 2, 3, 4$

Die Gesamtverstärkung dieser Filterkaskade wird an f_c auf 0 dB normiert. Eine Zusammenfassung dieser 4 Filterstufen 2ter Ordnung zu einem Gesamtfiler 8ter Ordnung ist insofern kritisch, da die erforderliche Koeffizientengenauigkeit stark zunimmt. Selbst bei Verwendung von doppelter Fließkommagenauigkeit für die Koeffizienten werden die Filter bei tiefen Filtermittenfrequenzen instabil.

Durch die reellen Filterkoeffizienten der einzelnen biquad-Stufen und $b_{2,i} = 0$ befinden sich die Nullstellen dieser Filter stets auf der reellen Achse.

Lyon¹ diskutiert zwei mögliche Vereinfachungen dieser GTF:

- 1) Verschiebung der Nullstellen auf den Koordinatenursprung der z-Ebene. Diese Filter werden All-Pol-Gammatonefilter (APGTF) genannt².
- 2) Verschiebung einer einzelnen Nullstelle auf $z = 1$. Die anderen Nullstellen werden in den Koordinatenursprung verschoben. Damit erreicht man, verglichen mit den APGTF, eine steilere tieffrequente Filterflanke nahe der Frequenz 0 Hz. Diese Filter werden One-Zero-Gammatonefilter (OZGTF) genannt. Sie approximieren die psychoakustisch gemessenen Daten genauer als die APGTF, sind jedoch analytisch nicht so einfach handhabbar.

Die Vorteile der APGTF und OZGTF gegenüber den GTF faßt Lyon folgendermaßen zusammen:³

- Die Unsymmetrie der Amplitudengänge entspricht den experimentalpsychologisch gemessenen Daten besser als der Amplitudengang des GTF.
- APGTF und OZGTF besitzen eine einfachere Parameterisierung.
- APGTF und OZGTF lassen sich einfacher und recheneffizienter implementieren.
- APGTF und OZGTF stellen eine Verbindung zu Filterkaskadenmodellen der Cochlea dar.

¹ [Lyon, 1996]

² Slaney ([Slaney, 1993]) verwendete diese Filter bereits 1993 aus Gründen der Rechenzeiterparnis.

³ [Lyon, 1996]

5.2. Nichtlineare Gammatone Filter

Für breitbandige Schalle wie Sprache oder Musik sind lineare GTFB bei mittleren Schallpegeln durchaus ausreichend. Im hohen Pegelbereich wirkt sich, selbst bei breitbandigen Schallen, die Unsymmetrie der auditiven Filter aus. Bei schmalbandigen Schallen ist die exakte Form der auditiven Filter stets von grundlegender Bedeutung, da die Flankensteilheiten in diesem Fall einen stärkeren Einfluß auf das Erregungsmuster ausüben. Eine universell verwendbare auditive Filterbank sollte also die Nichtlinearität der auditiven Filter mitberücksichtigen.

Die zur Zeit am häufigsten verwendeten auditiven Filterformen zur Modellierung der Nichtlinearität des peripheren Gehörs beziehen sich auf [Glasberg, Moore, 1990] (vgl. Gl. 2.19. und 2.20.). Diese Gleichungen stellen eine Approximation über psychoakustische Daten aus Notched Noise Experimenten dar (vgl. Kap. 2.5.1.).

Während lineare GTFB seit einigen Jahren in auditiven Modellen Verwendung finden¹, befinden sich die nichtlinearen GTFB zur Zeit noch in der Entwicklungsphase. Das Problem bei den nichtlinearen Filterbänken besteht darin, daß sich die Parametrisierung relativ komplex gestaltet.

Implementationsbeispiele nichtlinearer Gammatone Filter

Bei den folgenden Implementationsbeispielen handelt es sich sowohl um fertige Implementationen als auch um Ideen bzw. Ansätze zur nichtlinearen Parametrisierung der Gammatone Filter.

Die Übertragungsfunktion des analogen Gammatone Filters gibt Lyon folgendermaßen an:²

$$H_{GTF}(s) = \frac{aF(n)}{2} \left[\frac{e^{j\phi} [s - (-b - j\omega_c)]^n + e^{-j\phi} [s - (-b + j\omega_c)]^n}{[(s+b)^2 + \omega_c^2]^n} \right] \quad \text{Gl. 5.9.}$$

Die APGTF werden mit der Konstanten k zu

$$H_{APGTF}(s) = \frac{k}{[(s+b)^2 + \omega_c^2]^n} \quad \text{Gl. 5.10.}$$

und die OZGTF zu

¹ [Moore, 1993]

² [Lyon, 1996]

$$H_{OZGTF}(s) = \frac{sk}{[(s+b)^2 + \omega_c^2]^n} \quad \text{Gl. 5.11.}$$

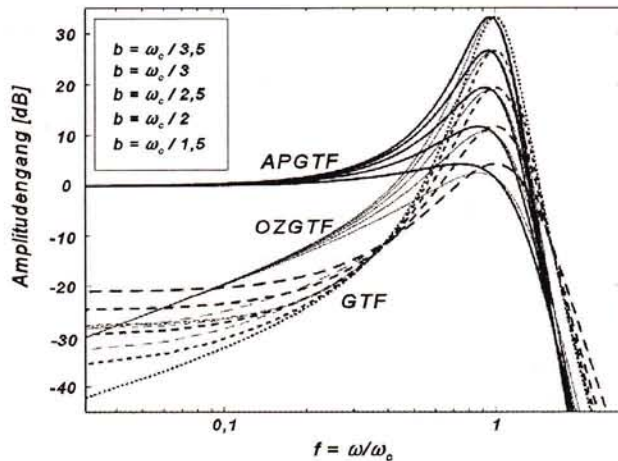


Abb. 5.3.: Beispiele für parametrisierte APGTF (durchgezogene Linien, fett), OZGTF (durchgezogene Linien, dünn), cos-phasige GTF (strichlierte Linien, fett) und sin-phasige GTF (strichlierte Linien, dünn) vgl. [Lyon, 1996]

Lyon¹ diskutiert unterschiedliche Möglichkeiten der Parametrisierung dieser drei Filtertypen in der s-Ebene. Eine etwaige Transformation dieser Parametrisierungen in die z-Ebene wird nicht behandelt.

In Abb. 5.3. sind einige Beispiele einer kartesischen Parametrisierung² nach Lyon für Filter 6ter Ordnung dargestellt. Die sin- und cos-phasigen GTF unterscheiden sich nur im Bereich $f \ll 1$. Die OZGTF stellen, verglichen mit den APGTF eine genauere Approximation der GTF und damit der psychoakustischen Daten dar.

Die 'Gammachirp' Funktion (Gl. 5.12.) stellt eine Modifikation der Gammatone Funktion (Gl. 5.1.) dar³. Im Unterschied zur ursprünglichen Funktion befindet sich dabei ein frequenzmodulierter Cosinus innerhalb der Einhüllenden. Aus diesem cos-phasigen 'chirp' Signal leitet sich der Name dieser Funktion ab.

$$g(t) = at^{n-1} e^{-2\pi bt} \cos(2\pi f_c t + c \ln t + \phi) \quad (t > 0) \quad \text{Gl. 5.12.}$$

Die Änderung des Amplitudenganges wird mit dem Parameter c erreicht.

In Abb. 5.4. sind 5 Amplitudengänge einer Gammachirp Funktion mit $f_c = 2\text{kHz}$, $n = 4 + 0.8c$, $b = 1.019 \cdot \text{ERB-Breite (an 2 kHz)}$ und $c = -1.2$ bis $+0.4$ dargestellt. Im Vergleich dazu sind 4

¹ [Lyon, 1996]

² Unter einer kartesischen Parametrisierung wird eine lautstärkepegelabhängige Verschiebung der Polstellen in der s-Ebene verstanden. Eine zweite Parametrisierung nach [Lyon, 1996] basiert auf einem polaren Koordinatensystem.

³ [Iriño, 1995], [Iriño, 1996], [Iriño, Patterson, 1996]

Funktionen nach Glasberg und Moore (Gl. 2.19. und 2.20.) für $L = 30, 50, 70$ und 90 dB/ERB strichliert dargestellt.

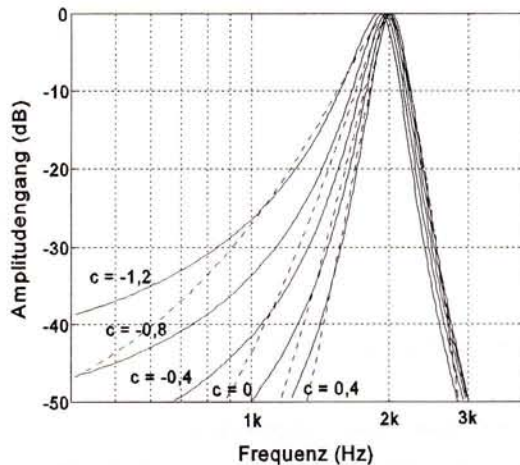


Abb. 5.4.: Durchgezogene Linien: Amplitudengänge eines Gammatone Filters mit: $f_c = 2$ kHz, $n = 4 + 0.8c$, $c = -1,2$ bis $+0,4$. Strichlierte Linien: AF Formen nach Glasberg und Moore (Gl. 2.19. und 2.20.) für $L = 30, 50, 70$ und 90 dB/ERB

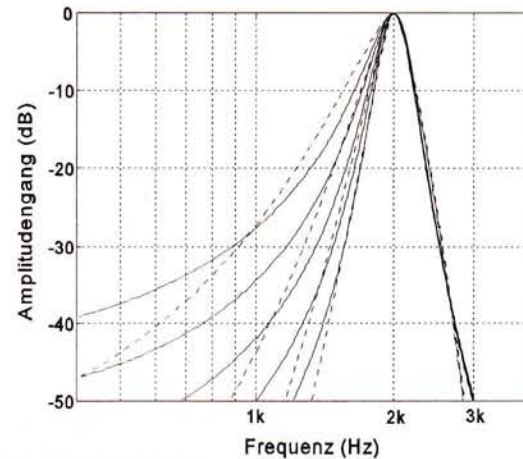


Abb 5.5.: Durchgezogene Linien: mittenfrequenz-korrigierte Amplitudengänge eines Gammatone Filters mit: $f_c = 2$ kHz, $n = 4 + 0.8c$, $c = -1,2$ bis $+0,4$. Strichlierte Linien: AF Formen nach Glasberg und Moore (Gl. 2.19. und 2.20.) für $L = 30, 50, 70$ und 90 dB/ERB.

Aus Abb 5.4. ist ein wesentliches Problem dieser Funktion ersichtlich. Die Mittenfrequenzen verschieben sich mit der pegelabhängigen Parametrisierung (c). Für das dargestellte Beispiel bewegt sich dieser Fehler im Bereich zwischen etwa -100 und $+25$ Hz. Bezogen auf die Filtermittenfrequenz von 2 kHz beträgt diese Verschiebung immerhin -5% bis $+1,25\%$. Im Erregungsmuster wirkt sich diese Verschiebung der Filtermittenfrequenz so aus, daß ein amplitudenmoduliertes Signal als frequenzmoduliertes Signal erscheint. Diese Frequenzverschiebung müßte also in einer Implementation zusätzlich korrigiert werden.

In Abb. 5.5. sind dieselben Amplitudengänge wie Abb. 5.4. - jedoch mit korrigierten Mittenfrequenzen dargestellt. Die Gammatone Filter sind an ihrer Spitze zu schmalbandig und an den tieffrequenten Flankenenden zu breitbandig.

Ein weiteres wesentliches Problem des Gammatone Filters besteht darin, daß sich die Filterordnungen pegelabhängig mitändern. Es ist mit einer digitalen Filterstruktur nicht möglich diese Ordnungsänderungen nachzubilden. Für das in Abb. 5.4. angeführte Beispiel ändert sich die Filterordnung beispielsweise von $n = 3,04$ bis $4,32$.

Das Modell von Carney¹ wurde ursprünglich zur Simulation tieffrequenter Hörnervenfasern von Katzen entworfen. In der Eingangsstufe dieses Modells befindet sich eine nichtlineare Filterstruktur, die die mechanische Selektion der Basilarmembran nachbildet. In Abb 5.6. ist das Blockschaltbild eines einzelnen Kanals dargestellt.

¹ [Carney, 1993]

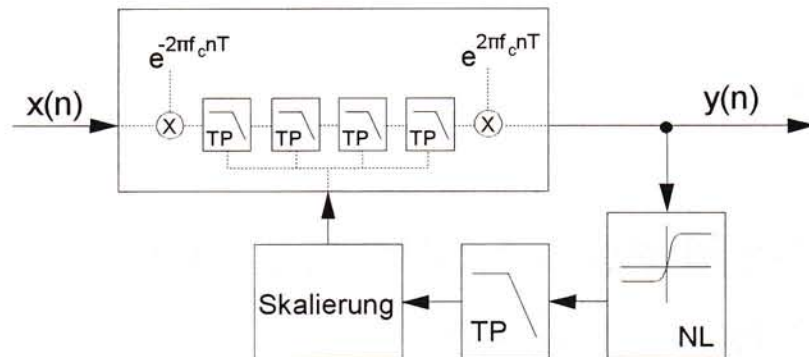


Abb. 5.6.: Modellierung eines einzelnen Cochlea Kanals nach [Carney, 1993]

Der Steuerungsparameter wird mit einer verschobenen *tanh*-Kennlinie (Erzeugung eines Gleichanteiles) und einem TP Filter (Extraktion des Gleichanteiles) bestimmt. Das nichtlineare Verhalten der Filter beschränkt sich auf einen Dynamikbereich von 30 - 90 dB des Ausgangssignals $y(n)$. In Anlehnung an die Gammatone Filter von [Patterson et al., 1987] und [Cooke, 1991] besteht der eigentliche Filter aus einer Frequenzverschiebung um die Mittenfrequenz f_c (Multiplikation des Eingangssignals mit $e^{-2\pi f_c nT}$), 4 Tiefpaßfiltern 1ter Ordnung und einer Frequenzverschiebung zurück auf f_c . Durch die Steuerung der Nichtlinearität an der Mittenfrequenz $f_c = 0$ können mit diesem Modell nur symmetrische Filter realisiert werden. Diese Einschränkung widerspricht den psychoakustisch ermittelten auditiven Filterformen, die besonders im hohen Pegelbereich stark unsymmetrisch sind (vgl. Kap. 2.6.2.).

Das 'Dual Resonance Model' einer nichtlinearen Gammatone Filterbank von Meddis¹ setzt sich aus zwei getrennten Signalwegen - einem schmalbandigen und einem breitbandigen - zusammen (Abb. 5.7.). Die speicherlose Nichtlinearität befindet sich im schmalbandigen Signalweg zwischen zwei kaskadierten Gammatonefiltern 2ter Ordnung. Durch Addition der beiden Ausgangssignale werden die nichtlinearen Amplitudengänge der auditiven Filter approximiert. Der Verstärkungsfaktor α_w steuert das Verhältnis der beiden Signalwege. Dieses Modell besitzt relativ viele Parameter und ist aus diesem Grund schwierig zu handhaben.

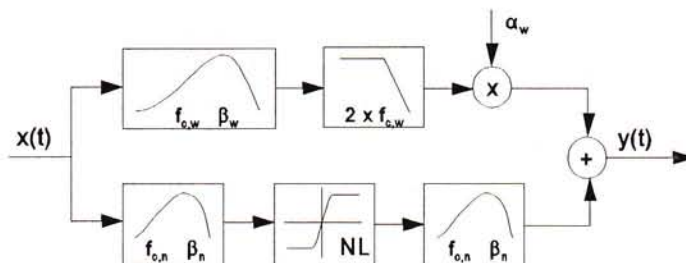


Abb 5.7.: Dual Resonanz Bandpaßmodell nach [Meddis, 1996]

¹ [Meddis, 1996]

Kapitel 6

Ein nichtlineares Modell des peripheren Gehörs

In diesem Kapitel wird ein nichtlineares funktionales Modell des peripheren Gehörs behandelt. Der lineare Amplitudengang des Außen- und Mittelohrfilters wird aus den Kurven gleicher Lautstärkepegel abgeleitet. Die Frequenzanalyse auf Ebene der Basilarmembran wird mit einer nichtlinearen Gammatone Filterbank (Parallel-Filterbank-Modell) durchgeführt.

6.1. Ein Modell des Außen- und Mittelohres

Da sich die Übertragungsfunktion des Außen- und Mittelohres (AM) physiologisch nicht eindeutig messen läßt, wird sie aus indirekten Messungen abgeleitet. Unter der Annahme, daß die Cochlea für alle Frequenzen gleich empfindlich ist, läßt sich der Amplitudengang der Außen- und Mittelohrfilterung aus der Ruhehörschwelle oder dem eben wahrnehmbaren Schalldruckpegel am Trommelfell bestimmen¹.

Unter dieser Annahme bleibt jedoch die Frage ungeklärt, warum die Kurven gleicher Lautstärkepegel (KGL) im Bereich unter 1 kHz mit steigender Lautstärke deutlich flacher werden. Geht man davon aus, daß das interne Rauschen der Cochlea bei tiefen Frequenzen stärker ausgeprägt ist als bei hohen Frequenzen, so läßt sich dieser Effekt erklären. Das niederpegelige Rauschen der Cochlea hebt zwar die Ruhehörschwelle an, bei den höheren KGL wirkt sich das Grundrauschen der Cochlea jedoch nicht aus.

Glasberg und Moore² verwenden daher bei den Notched-Noise Experimenten zur Bestimmung der pegelabhängigen $roex(p)$ Filter die inverse 100 Phon Kurve nach ISO 226 als Außen- und

¹ [Glasberg, Moore, 1990] bezeichnen diese beiden Außen- und Mittelohrfilter als MAF (minimum audible field, inverse Ruhehörschwelle nach ISO) und MAP (minimum audible pressure, minimal hörbaren Schalldruckpegel am Trommelfell), vgl. Kap. 2.5.5. .

² [Moore, Glasberg, 1990]

Mittelohrfilter, da diese Kurve als vom internen Rauschen der Cochlea unabhängig betrachtet wird. Über 1 kHz entspricht diese Kurve etwa der Ruhehörschwelle. Aus diesem Grund wird über 1 kHz kein internes Rauschen zum Erregungsmuster hinzuaddiert. Diese Art der AM-Filterung wird auch in der Erweiterung des Zwicker-Lautheitsmodells von Glasberg und Moore verwendet¹. Agerkvist² verwendet zur Berechnung des Erregungsmusters die inverse 90 Phon ISO Kurve, da für die 100 Phon ISO Kurve für 25Hz, 10kHz und 12,5kHz keine Daten existieren.

Die Anforderungen an den Amplitudengang des Außen- und Mittelohrfilters sind folgende:

- Der Amplitudengang über 1 kHz entspricht der inversen Ruhehörschwelle.
Über 1 kHz besitzen alle Kurven gleicher Lautstärkepegel eine sehr ähnliche Form. Daraus resultiert die Annahme, daß über 1 kHz das innere Rauschen der Cochlea die Ruhehörschwelle nicht beeinflusst. Da die Ruhehörschwelle die zur Zeit einzige Kurve gleichen Lautstärkepegels ist, deren Form außer Zweifel steht, wird über 1 kHz deren Form nachgebildet.
- Der Amplitudengang unter 1 kHz wird aus Daten neuerer Versuchsreihen approximiert.
Es gibt zur Zeit keine allgemein anerkannte Form der Kurven gleicher Lautstärkepegel. Die nach ISO 226 genormten Kurven sind mit großer Wahrscheinlichkeit falsch, neuere Versuchsreihen zeigen untereinander starke Abweichungen (vgl. Kap. 3.2.). Aus diesem Grund bildet der verwendete Außen-Mittelohramplitudengang eine Mittelung über die unterschiedlichen Datensätze.
- Eine nachträgliche Adaption des Außen- und Mittelohramplitudengangs an zukünftige Formen sollte auf einfache Weise möglich sein, ohne den gesamten Filter neu entwerfen zu müssen.
Es wird zur Zeit an einer Revision des ISO 226 Standards (Kurven gleicher Lautstärkepegel) gearbeitet. Aus diesem Grund sollte sich das Modell an diese, in den nächsten Jahren erwartete Neuerung ohne grundlegenden Neuentwurf adaptieren lassen (vgl. Kap. 3.2.).
- Das innere Rauschen der Cochlea, d.h. die Differenz aus Außen- Mittelohramplitudengang und Ruhehörschwelle läßt sich in geschlossener Form berechnen.
Um das innere Rauschen der Cochlea nicht, wie zur Zeit üblich, aus Tabellen entnehmen zu müssen, ist eine geschlossene Lösung dieser Stufe wünschenswert.

¹ [Moore, Glasberg, 1996], vgl. Kap. 3

² [Agerkvist, 1994]

Die Übertragungsfunktion des Außen-Mittelohrfilters $H_{AM}(z) = H_{TP}(z) * H_{HP}(z)$ setzt sich aus einem Hochpaß- und einem Tiefpaßfilter zusammen (Abb. 6.1.). Diese Zweiteilung ermöglicht die separate Steuerung der Filterflanke unter 1 kHz mit dem Parameter R (Einzahlwert).

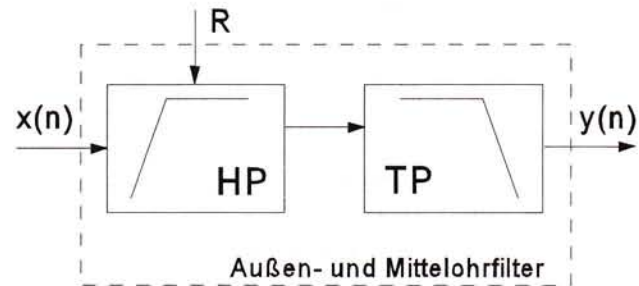


Abb. 6.1.: Außen- und Mittelohrfilter

Der rekursive Tiefpaßfilter 8ter Ordnung

$$H_{TP}(z) = 0,109 \frac{z^7(z+1)}{z^8 - 2,5359z^7 + 3,9295z^6 - 4,7532z^5 + 4,7251z^4 - 3,5548z^3 + 2,1396z^2 - 0,9879z + 0,2836} \quad \text{Gl. 6.1.}$$

bildet die beiden Resonanzen der inversen Ruhehörschwelle an 4 kHz und 12 kHz nach (vgl. Abb. 6.2 und 6.3.). Unter 1 kHz ist der Amplitudengang dieses Filters konstant 0 dB. Dadurch können mit dem parametrisierten Hochpaßfilter $H_{HP}(z)$, dessen Amplitudengang über 1 kHz von der Parametrisierung unabhängig 0 dB beträgt, die Änderungen des Amplitudengangs unter 1 kHz unabhängig modelliert werden. Der Steuerungsparameter R gibt die Lage der beiden Polstellen auf der reellen Achse wieder.

Die Übertragungsfunktion des Hochpaßfilters lautet:

$$H_{HP}(z) = \frac{z^2 - 2z + 1}{z^2 - 2Rz + R^2} \quad \text{Gl. 6.2.}$$

Für den verwendeten Außen- und Mittelohrfilter wird ein Steuerungsparameter $R = R_{AM} = 0,989$ vorgeschlagen. Der damit erhaltene Amplitudengang entspricht bei etwa 30 Hz der inversen 100 Phon ISO Kurve (Abb. 6.2.). Die leichte Resonanz des 100 Phon ISO Filters bei etwa 300 Hz wird nicht nachgebildet, da diese Resonanz bei keiner der neueren Versuchsreihen zur Bestimmung der Kurven gleicher Lautstärkepegel vorhanden ist¹.

In Abb 6.3. ist die Änderung des AM-Filter Amplitudengangs als Funktion des Parameters R für $R = 0,94, 0,95, 0,96, 0,97, 0,98, 0,99$ dargestellt. Aus dieser Abbildung ist ersichtlich, welche

¹ Der in Abb. 3.1. dargestellte Vergleich zwischen den Kurven gleicher Lautstärkepegel nach ISO und den Kurven gleicher Lautstärkepegel aus 4 neueren Datensätzen beinhaltet keine Angaben über neuere Forschungsergebnisse für die 100 Phon Kurve. Da bei den verglichenen Kurven gleicher Lautstärkepegel (Ruhehörschwelle, 20, 30, 40, 50, 60, 70 und 80 phon) jedoch die Resonanz bei etwa 300 Hz ausschließlich bei den ISO Phon Kurven vorhanden ist, läßt sich daraus schließen, daß es sich dabei mit großer Wahrscheinlichkeit um ein Artefakt der Versuchsreihe von Robinson und Dadson handelt (vgl. Kap. 3.2.).

Amplitudengänge sich mit dieser einfachen Parametrisierung realisieren lassen.

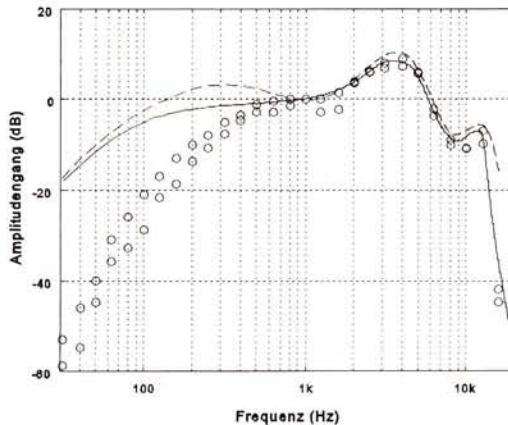


Abb. 6.2.: Inverse 100 Phon ISO Kurve (strichliert) und inverse Ruheshwellen nach ISO und [Fastl et al., 1990] (Kreise) verglichen mit dem Außen- und Mittelohrfilter (durchgezogene Linie, $R_{AM} = 0,989$)

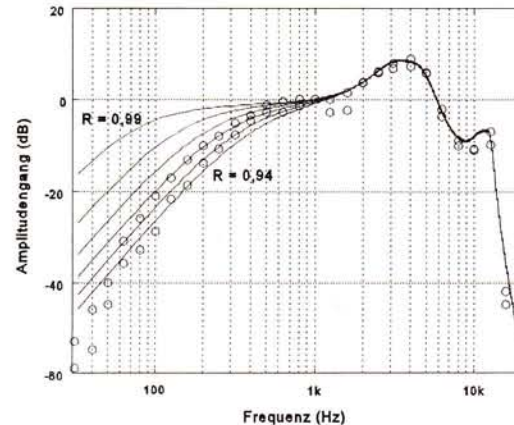


Abb. 6.3.: Inverse Ruheshwellen nach ISO und [Fastl et al., 1990] (Kreise) sowie Außen- und Mittelohrfilter für ($R = 0,94, 0,95, 0,96, 0,97, 0,98, 0,99$, durchgezogene Linien)

Ein wesentlicher Vorteil dieses Filters besteht darin, daß sich die Ruheshwelle ebenfalls mit dieser Filterstruktur nachbilden läßt (Vorschlag: $R = R_{RH} = 0,957$) und sich dadurch die Differenz aus Außen- Mittelohrfilter und Ruheshwelle (Abb. 6.4.) in geschlossener Form lösen läßt.

Diese Differenz ergibt sich einzig aus der Übertragungsfunktion des Hochpaßfilters $H_{HP}(z)$, da der Tiefpaßfilter unter 1 kHz eine Verstärkung von 0 dB aufweist.

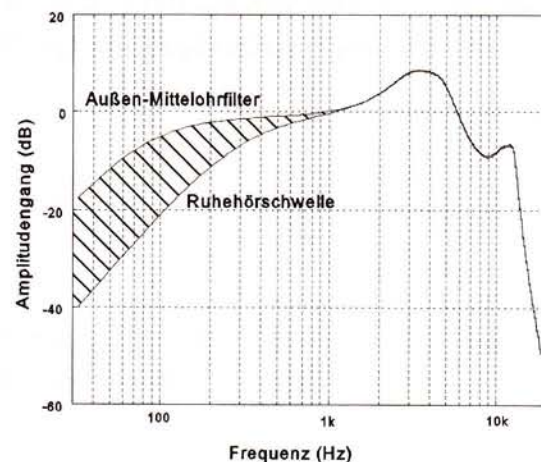


Abb. 6.4.: Differenz aus der Ruheshwelle und dem Außen- Mittelohrfilter (schraffiert)

Die Übertragungsfunktion des Hochpaßfilters (Gl. 6.2.) läßt sich umschreiben zu (vgl. Abb. 6.5.):

$$H_{HP}(e^{j\Theta}) = \frac{N_1(\Theta)e^{j\alpha_1} N_2(\Theta)e^{j\alpha_2}}{D_1(\Theta)e^{j\beta_1} D_2(\Theta)e^{j\beta_2}}$$

Gl. 6.3.

$$\text{mit } \Theta = 2\pi f_c T$$

Da sowohl die beiden Polstellen als auch die beiden Nullstellen dieses Filters jeweils an der gleichen Stelle liegen, kann Gl. 6.3. mit $N = N_1 = N_2$, $\alpha = \alpha_1 = \alpha_2$, $D = D_1 = D_2$, $\beta = \beta_1 = \beta_2$ vereinfacht werden zu Gl. 6.4.:

$$H_{HP}(e^{j\Theta}) = \frac{N^2(\Theta)}{D^2(\Theta)} e^{j(2\alpha - 2\beta)} \quad \text{Gl. 6.4.}$$

Durch die alleinige Betrachtung des Amplitudengangs kann die Phase aus Gl. 6.4. vernachlässigt werden:

$$|H_{HP}(e^{j\Theta})| = \frac{N^2(\Theta)}{D^2(\Theta)} \quad \text{Gl. 6.5.}$$

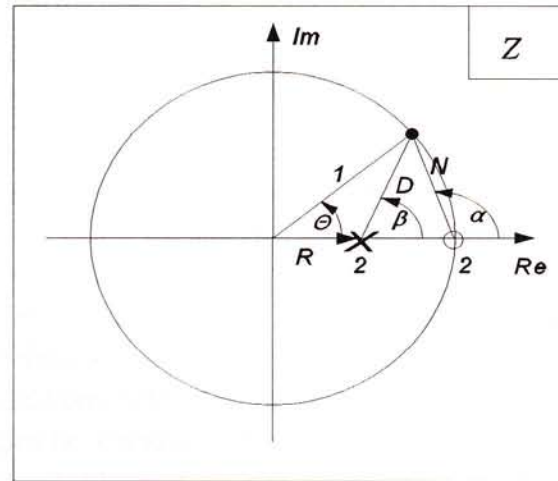


Abb. 6.5.: Doppelpol- und Doppelnullstellen des Hochpaßfilters

Setzt man $D^2 = 1 + R^2 - 2R\cos(\Theta)$, so wird die Differenz $A_{Diff}(\Theta)$ aus dem AM-Amplitudengang (Parameter: $N = N_{AF}$, $D = D_{AF}$, $R = R_{AF}$) und dem Ruhehörschwellenfilter (Parameter: $N = N_{RH}$, $D = D_{RH}$, $R = R_{RH}$) zu:

$$A_{Diff}(\Theta) = \frac{N_{AM}^2(\Theta) D_{RH}^2(\Theta)}{D_{AM}^2(\Theta) N_{RH}^2(\Theta)} = \frac{D_{RH}^2(\Theta)}{D_{AM}^2(\Theta)} = \frac{1 + R_{RH}^2 - 2R_{RH}\cos(\Theta)}{1 + R_{AM}^2 - 2R_{AM}\cos(\Theta)} \quad \text{Gl. 6.6.}$$

Setzt man in Glg. 6.6. für $\Theta = 2\pi fT$, $R_{AM} = 0,989$ und für $R_{RH} = 0,957$, so ergibt sich die Differenz zwischen dem Außen- Mittelohrfilter und dem Ruhehörschwellen Filter als Funktion der Frequenz mit:

$$A_{Diff}(f) = \frac{1,915849 - 1,914\cos(2\pi fT)}{1,978121 - 1,978\cos(2\pi fT)} \quad \text{Gl. 6.7.}$$

Es besteht also die Möglichkeit, die Form des Außen- Mittelohrfilters und damit verbunden das innere Rauschen der Cochlea mit einem einzigen Parameter R_{AM} zu ändern.

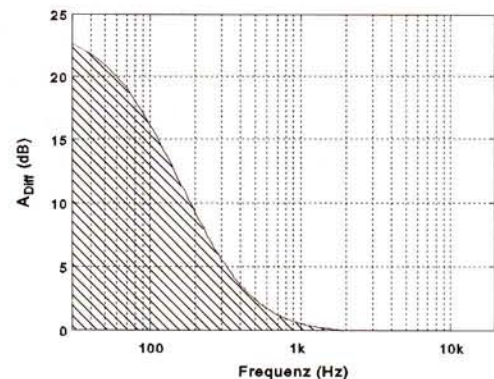


Abb. 6.6.: Differenz zwischen dem Außen- Mittelohrfilter nach Gl. 6.7.

Für die Simulation diffuser Schallfeldbedingungen wird ein Diffusfeldfilter $H_{FD}(z)$ (Gl. 6.8., Abb. 6.7.) zum Außen- Mittelohrfilter in Serie geschaltet.

$$H_{FD}(z) = \frac{0,097z^7 + 0,9368z^6 - 2,5392z^5 + 2,1854z^4 - 0,802z^3 + 0,2112z^2 - 0,102z + 0,0293}{z^4 - 2,2709z^3 + 1,9829z^2 - 0,9877z + 0,2922} \quad \text{Gl. 6.8.}$$

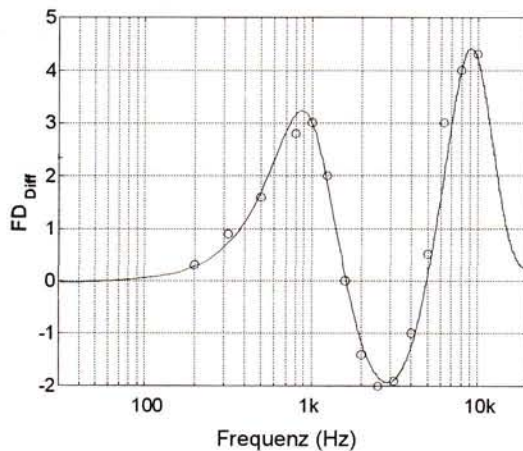


Abb. 6.7.: Diffusfeldfilter (durchgezogene Linie), Daten aus ISO 226 (Kreise)

Damit werden die Unterschiede der Kurven gleicher Lautstärkepegel bei Freifeld- und Diffusfeldbedingungen nachgebildet.

Bei einer Frequenz von 1 kHz bedeutet dies beispielsweise, daß der Schallpegel im diffusen Schallfeld um 3 dB im Pegel geringer sein muß, um die gleiche Lautstärkeempfindung hervorzurufen wie ein 1 kHz im ebenen Schallfeld (vgl. Kap. 3.5.2.). Das zum Erregungsmuster hinzuaddierte interne Rauschen der Cochlea wird von dieser zusätzlichen Diffusfeldfilterung nicht beeinflusst.

Es ergeben sich zwei mögliche Außen- und Mittelohrfiltertypen:

AM_F ... Außen- und Mittelohrfilter für freie Schallfeldbedingungen

AM_D ... Außen- und Mittelohrfilter für diffuse Schallfeldbedingungen

Mit diesem einfach parametrisierten Außen- und Mittelohrfilter läßt sich eine dem dB(ELC) (vgl. Kap. 3.5.2.) qualitativ gleichwertige Meßgröße um vieles recheneffizienter implementieren. Der Nachteil einer derartigen dB(ELC) Berechnung besteht darin, daß durch die einfache Parametrisierung (Verschiebung eines reellen Doppelpols) einige Details der Kurven gleicher Lautstärkepegel unter 1 kHz nicht berücksichtigt werden können. Eine Beantwortung der Frage, inwieweit sich dieser Nachteil auf die Vorhersage von Lautheits- bzw. Lärmbelastungen tatsächlich auswirkt, kann aus theoretischen Überlegungen allein nicht abgeleitet werden.

6.2. Ein nichtlineares Modell der Cochlea

Das im folgenden beschriebene funktionale Modell der Cochlea wird aus der linearen Gammatone Filterbank nach Slaney (vgl. Kap. 5.1.) abgeleitet. Es werden sowohl eingangspegelabhängige als auch ausgangspegelabhängige auditive Filter nachgebildet (vgl. Kap. 2.6.2.).

6.2.1. Parametrisierung eingangspegelabhängiger nichtlinearer APGTF und OZGTF¹

Die linearen Gammatone Filter nach Slaney (Gl. 5.7. und 5.8.) werden in nichtlineare Filter übergeführt, indem bei den drei Filterstufen $H_2(z)$ bis $H_4(z)$ (Abb. 6.8.) die Polstellen verschoben werden.

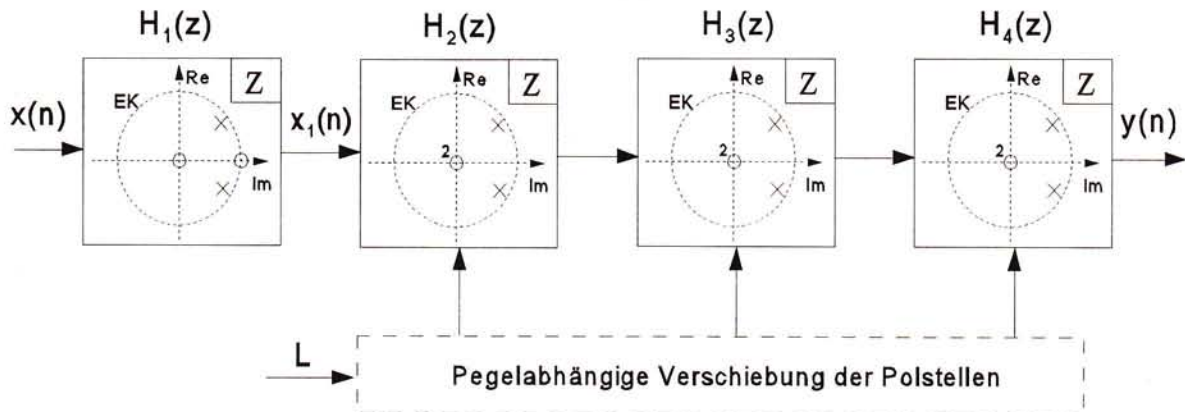


Abb. 6.8.: Kaskadenstruktur eines nichtlinearen OZGTF

Die nichtlinearen All-Pol (nlAPGTF) und One-Zero (nlOZGTF) Gammatone Filter unterscheiden sich nur durch die erste Stufe $H_1(z)$. Bei den nlAPGTF befinden sich alle Nullstellen an $z = 0$, während bei den nlOZGTF eine Nullstelle an $z = 1$ liegt. Da die erste Filterstufe $H_1(z)$ nicht verändert wird, besteht die Möglichkeit den Ausgang dieser Stufe $x_1(n)$ zur Bestimmung des energieäquivalenten Eingangspegels L (vgl. Gl. 2.20.) zu verwenden. Auf diese Weise kann eine externe Filterung zur eingangspegelabhängigen Steuerung der Nichtlinearität vermieden werden.

Die Polstellenpaare einer einzelnen Filterstufe ergeben sich aus Gl. 5.8. zu:

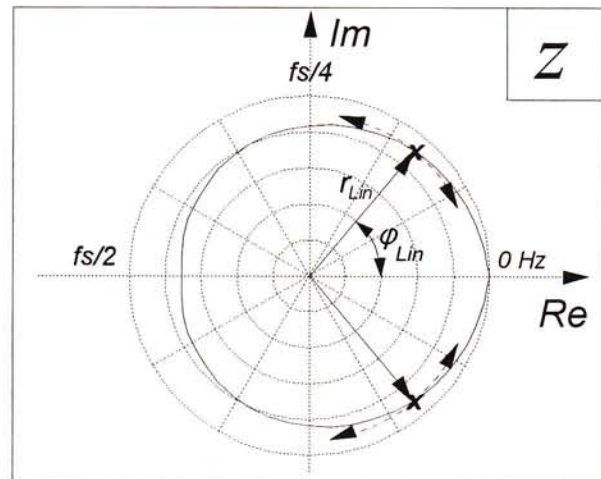
$$p_{1,2} = r_{Lin} e^{\pm j\varphi_{Lin}} = \frac{1}{e^{BT}} e^{\pm j2\pi f_c T} \quad \text{Gl. 6.9.}$$

¹ [Pflüger et al., 1997a], [Pflüger et al., 1997b]

Die Positionen dieser Polstellen sind in Abb. 6.9. als Funktion der Filtermittenfrequenz in der z-Ebene dargestellt (durchgezogene Linie, $f_s = 44,1$ kHz).

Die pegelabhängige Verschiebung der Polstellen wird unter Beibehaltung der Beträge ($r_{Lin} = \text{konst.}$) durchgeführt (Abb. 6.9. strichlierte Linien).

Dadurch ändert sich bei den Übertragungsfunktionen $H_2(z)$ bis $H_4(z)$ jeweils nur ein Koeffizient ($a_{1,i}$, vgl. Gl. 5.8.).



Die Winkelverschiebung der Polstellen wird durch eine Multiplikation des Polstellenwinkels $\varphi_{Lin} = \pm 2\pi f_c T$ mit einem pegelabhängigen Faktor K erreicht (Gl. 6.10.a).

Abb. 6.9.: Positionen der Polstellen als Funktion der Filtermittenfrequenz (durchgezogene Linie), pegelabhängige Verschiebung der Polstellen bei konst. Beträgen (strichlierte Linien)

Durch diese Multiplikation ist die Winkelverschiebung umso größer, je größer der Winkel φ_{Lin} ist. Dies ist erforderlich, da die Abstände der Polstellen zum Einheitskreis mit steigendem φ_{Lin} größer werden (vgl. Abb. 6.9.) und sich dadurch die Empfindlichkeit des Amplitudengangs bezüglich einer Polstellenverschiebung verringert.

$$\varphi_{NLin} = K \varphi_{Lin} \quad K(f_c, L) = \frac{k_\varphi(L) - 1}{r_{Lin}(f_c)} + 1 \quad \text{Gl. 6.10.a,b}$$

Der Multiplikator K berechnet sich aus dem steuerungspegelabhängigen Parameter k_φ und dem Betrag r_{Lin} der entsprechenden Polstelle (Gl. 6.10.b). Diese Funktion bewirkt bei sehr hohen Frequenzen ($f_c \gg 1$ kHz bei $f_s = 44,1$ kHz) eine zusätzliche Verstärkung der Polstellenverschiebung. Für $r_{Lin} \approx 1$ gilt $K \approx k_\varphi$ (durchgezogene Gerade in Abb. 6.10.).

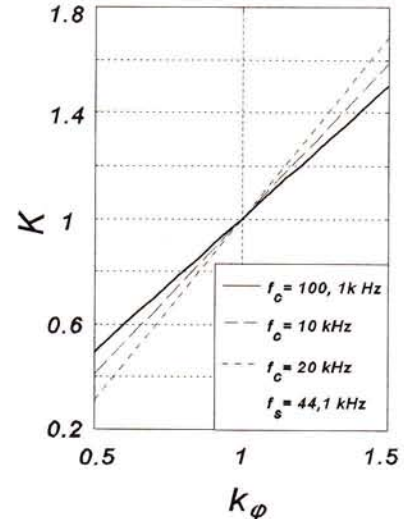


Abb. 6.10.: K als Funktion von k_φ bei 4 unterschiedlichen Mittenfrequenzen $f_c = 100, 1k, 10k$ und $20k$ Hz

Die Polstellen der Filter aus Gl.6.9. werden mit Gl. 6.10a,b zu:

$$p_{1,2} = r_{Lin} e^{\pm j\varphi_{NLin}} = \frac{1}{e^{BT}} e^{\pm j2\pi f_c TK} = \frac{1}{e^{BT}} e^{\pm j2\pi f_c T(e^{BT}(k_\varphi - 1) + 1)} \quad \text{Gl. 6.11.}$$

Die Pegelabhängigkeit der Parameter k_φ wird mit arctan-Funktionen approximiert (Gl. 6.12. und 6.13.). Die Indizes 1,2,3,4 beziehen sich auf die jeweilige Filterstufe.

- k_φ für nlAPGTF (vgl. Abb. 6.11.):

$$\begin{aligned} k_{\varphi 1}(L) &= 1 \\ k_{\varphi 2}(L) &= 0,9898 - 0,0066 \arctan(0,6(L-85)) \\ k_{\varphi 3}(L) &= 0,893 - 0,075 \arctan(0,1(L-70)) \\ k_{\varphi 4}(L) &= 1,0749 + 0,0165 \arctan(0,4(L-55)) \end{aligned} \quad \text{Gl. 6.12.}$$

- k_φ für nlOZGTF (vgl. Abb. 6.12.):

$$\begin{aligned} k_{\varphi 1}(L) &= 1 \\ k_{\varphi 2}(L) &= 0,9789 - 0,014 \arctan(0,2(L-80)) \\ k_{\varphi 3}(L) &= 0,8673 - 0,093 \arctan(0,1(L-70)) \\ k_{\varphi 4}(L) &= 1,0606 + 0,025 \arctan(0,1(L-68)) \end{aligned} \quad \text{Gl. 6.13.}$$

Der Parameter k_φ ist unabhängig von der Filtermittenfrequenz und kann daher für alle Filterbankstufen verwendet werden.

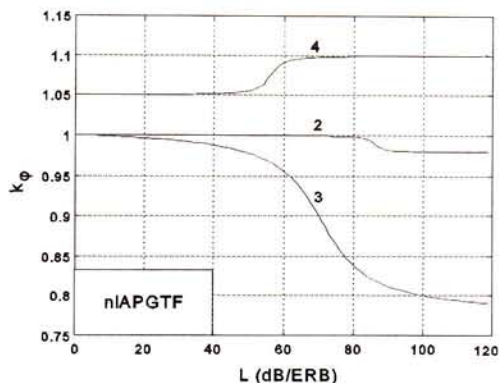


Abb. 6.11.: Steuerungspegelabhängiger Parameter k_φ für nlAPGTF

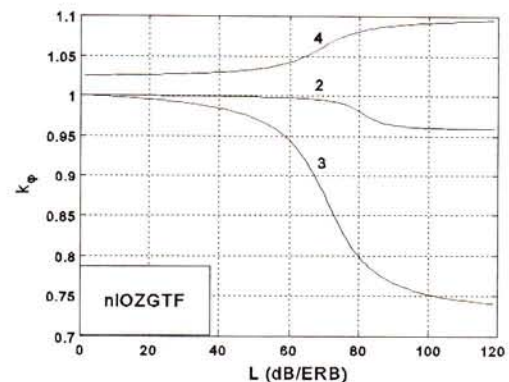


Abb. 6.12.: Steuerungspegelabhängiger Parameter k_φ für nlOZGTF

Aus den Abbildungen 6.11. und 6.12. ist ersichtlich, daß bei kleinen Steuerungspegeln L die Polstellenpositionen in etwa den linearen GTF aus Kap. 5 entsprechen. Der einzige Unterschied besteht in der 4ten Filterstufe, da der Parameter $k_{\varphi 4}$ für $L = 0$ nicht 1 ist. Dieser Unterschied

wirkt sich auf den Amplitudengang der Filterkaskade nicht wesentlich aus.

Der Zusammenhang zwischen der Polstellenverschiebung und den Koeffizienten des Nennerpolynoms einer einzelnen Filterstufe (Gl. 5.7.) ist in den beiden Gleichungen Gl. 6.14. und 6.15. dargestellt.

Der Koeffizient $a_{2,i}$ ist für alle Filterstufen pegelunabhängig:

$$a_{2,i} = r_{Lin}^2 = \frac{1}{e^{2BT}} \quad \text{Gl. 6.14.}$$

$$i = 1,2,3,4$$

Der Koeffizient $a_{1,i}$ enthält den pegelabhängigen Parameter k_{φ} . Dieser Koeffizient läßt sich beliebig interpolieren d.h. es ergeben sich immer konjugiert komplexe Polstellen mit konstanten Beträgen.

$$a_{1,i} = -2 r_{Lin} \cos(\varphi_{NLin,i}) = -2 \frac{1}{e^{BT}} \cos(2\pi f_c T (e^{BT} (k_{\varphi,i} - 1) + 1)) \quad \text{Gl. 6.15.}$$

$$i = 1,2,3,4$$

Für die Koeffizienten des Zählerpolynoms gilt:

nlAPGTF:

$$b_{0,i} = 1, \quad b_{1,i} = 0, \quad b_{2,i} = 0 \quad \text{Gl. 6.16.}$$

$$i = 1,2,3,4$$

nlOZGTF:

$$b_{0,1} = 1, \quad b_{1,1} = -1, \quad b_{2,1} = 0 \quad \text{Gl. 6.17.}$$

$$b_{0,i} = 1, \quad b_{1,i} = 0, \quad b_{2,i} = 0, \quad i = 2,3,4$$

Die Abtastperiode T des Zählerpolynoms aus Gl. 5.8. ist in den Gl. 6.16. und 6.17. nicht angeführt, da dieser Wert in den Verstärkungsfaktoren (Gl. 6.18. und 6.19.) berücksichtigt wird.

Der Ausgang der ersten Filterstufe $x_1(n)$ muß an der Filtermittenfrequenz f_c auf 0 dB normiert sein, da dieses Signal optional zur Bestimmung des Steuerungspegels L vorgesehen ist. Dies wird mit dem Verstärkungsfaktor V_1 in der ersten Stufe erreicht. Da bei der ersten Filterstufe der Maximalwert des Amplitudengangs mit Sicherheit an f_c liegt (bei der ersten Filterstufe werden die Polstellen nicht verschoben), ergibt sich dieser Verstärkungsfaktor direkt aus der

Übertragungsfunktion $H_1(z)$:

$$\begin{aligned} \text{nIAPGTF: } V_1 &= \left| 1 + a_{1,1} e^{-j2\pi f_c T} + a_{2,1} e^{-j4\pi f_c T} \right| \\ \text{nIOZGTF: } V_1 &= \left| \frac{1 + a_{1,1} e^{-j2\pi f_c T} + a_{2,1} e^{-j4\pi f_c T}}{1 - e^{-j2\pi f_c T}} \right| \end{aligned} \quad \text{Gl. 6.18.}$$

Eine Normierung jeder einzelnen Filterstufe (0 dB an f_c) ist nicht zielführend, da durch die Verschiebung der Polstellen die Amplitudengangsmaximalwerte nicht an der Mittenfrequenz des Gesamtfilters liegen. Faßt man jedoch die drei Filterstufen $H_2(z)$ bis $H_4(z)$ zusammen, so ergibt sich wiederum ein Amplitudengangsmaximum, das sich etwa an f_c befindet. Daraus läßt sich der Verstärkungsfaktor V_{com} (Gl. 6.19.) bestimmen, der zusammen mit dem Faktor V_1 die gesamte Filterkaskade auf 0 dB an f_c normiert.

$$\begin{aligned} V_{\text{com}} &= \left| 1 + A e^{-j2\pi f_c T} + B e^{-j4\pi f_c T} + C e^{-j6\pi f_c T} + D e^{-j8\pi f_c T} + E e^{-j10\pi f_c T} + F e^{-j12\pi f_c T} \right| \\ A &= a_{1,2} + a_{1,3} + a_{1,4} \\ B &= 3 a_{2,2} + a_{1,2} a_{1,3} + a_{1,3} a_{1,4} + a_{1,2} a_{1,4} \\ C &= 2 a_{1,2} a_{2,2} + 2 a_{2,2} a_{1,3} + 2 a_{2,2} a_{1,4} + a_{1,2} a_{1,3} a_{1,4} \\ D &= 3 a_{2,2}^2 + a_{1,2} a_{2,2} a_{1,3} + a_{1,2} a_{2,2} a_{1,4} + a_{2,2} a_{1,3} a_{1,4} \\ E &= a_{2,2}^2 (a_{1,2} + a_{1,3} + a_{1,4}) \\ F &= a_{2,2}^3 \\ \text{mit } a_{2,2} &= a_{2,3} = a_{2,4} \end{aligned} \quad \text{Gl. 6.19.}$$

Der Faktor V_{com} kann für nIAPGTF und nIOZGTF verwendet werden, da sich die beiden Filtertypen nur durch die erste Stufe unterscheiden.

Aus der Darstellung der Filterkaskade in Form von 4 Pol-Nullstellendiagrammen (Abb. 6.8.) wird bei Verwendung der transponierten Direktform II die digitale Filterstruktur in Abb. 6.13. . Die Steuerungseinheit für diese Filterkaskade ist in Abb. 6.14. separat dargestellt.

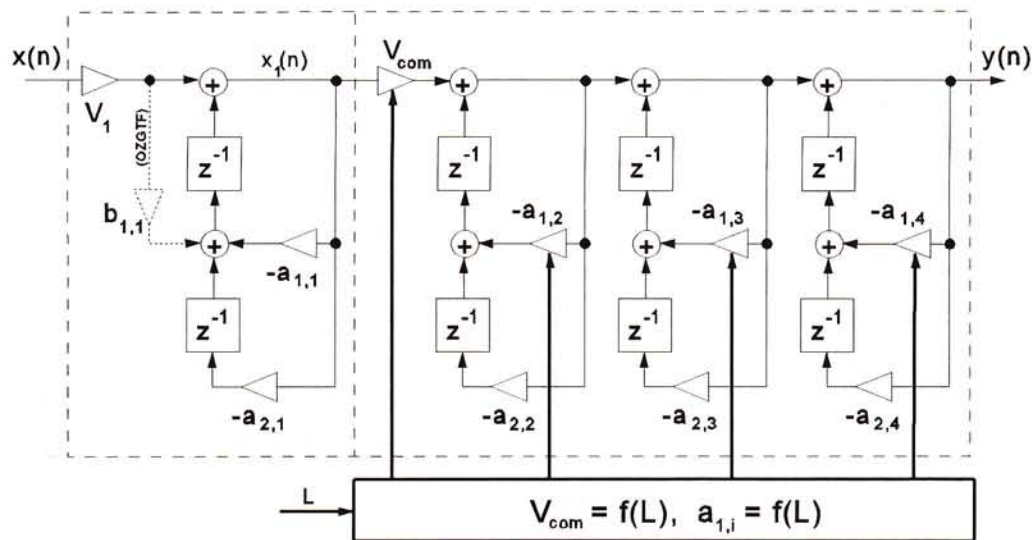


Abb. 6.13.: Nichtlineare nAPGTF und nOZGTF

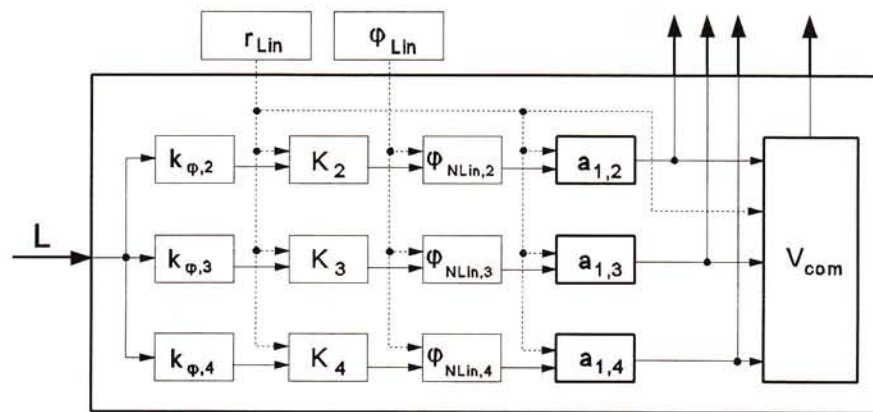


Abb. 6.14.: Steuerungseinheit für nAPGTF and nOZGTF

Eigenschaften der eingangspegelabhängigen nAPGTF und nOZGTF:

Die Gleichungen Gl. 6.10. bis 6.19. wurden so gewählt, daß sich die Amplitudengänge für $L = 0 - 50$ dB/ERB kaum ändern und etwa den linearen APGTF und OZGTF entsprechen. Die Änderungen der tieffrequenten Filterflanken finden im Bereich zwischen $L = 50 - 90$ dB/ERB statt. Dadurch besteht die Möglichkeit, bei gleichen Steuerungsparametern die resultierenden Amplitudengänge direkt mit den roex(p) Filterformen nach Glasberg und Moore¹ (vgl. Kap. 2.6. Gl. 2.19. und 2.20.) zu vergleichen.

¹ [Glasberg, Moore, 1990]

In Abb. 6.15. sind roex(p) Filterformen für 5 unterschiedliche Mittenfrequenzen (100, 300, 1k, 3k und 10k Hz) und Steuerungspegel ($L = 50, 60, 70, 80$ und 90 dB/ERB) nach Gl. 2.19. und 2.20. dargestellt.

Im Vergleich dazu sind die Amplitudengänge der nAPGTF und der nOZGTF für dieselben Mittenfrequenzen und Steuerungspegel in den Abb. 6.16. und 6.17. dargestellt.

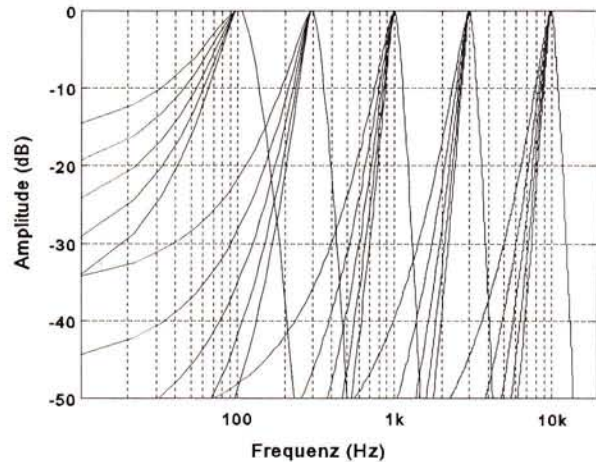


Abb. 6.15.: Auditive Filterformen nach [Glasberg, Moore, 1990] für $f_c = 100, 300, 1k, 3k, 10k$ Hz und $L = 50, 60, 70, 80, 90$ dB/ERB

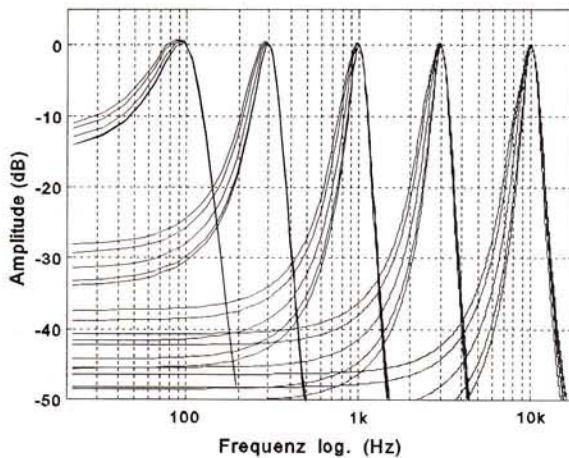


Abb. 6.16.: nAPGTF Amplitudengänge an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz für jeweils $L = 50, 60, 70, 80$ und 90 dB/ERB

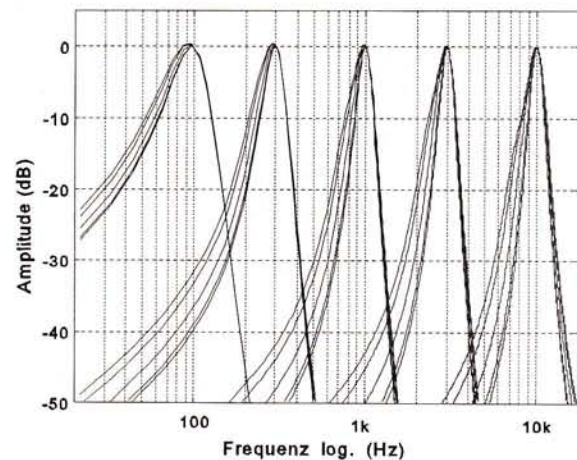


Abb. 6.17.: nOZGTF Amplitudengänge an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz für jeweils $L = 50, 60, 70, 80$ und 90 dB/ERB

Bei Frequenzen von $f \ll f_c$ ist die Änderung der tieffrequenten Flanke bei Glasberg und Moore deutlich stärker ausgeprägt. Dies bedeutet, daß beispielsweise das 'upward spread of masking' bei diesen Amplitudengängen stärker ausgeprägt ist.

Durch die Positionierung der Nullstellen an $z = 0$ sind im Falle der nAPGTF die tieffrequenten Flanken ab einem bestimmten Pegel waagrecht (Abb. 6.16.). Dadurch wird das 'upward spread of masking' ab einem bestimmten Frequenzabstand vom Maskierungssignal überbewertet. Bei den OZGTF ist dieser Effekt durch die Nullstelle an $z = 1$ nicht vorhanden (Abb. 6.17.).

Nahe der Filtermittenfrequenz sind sowohl die nAPGTF (Abb. 6.18.) als auch die nOZGTF (Abb. 6.19.) breiter als die roex(p)-Funktionen nach Glasberg und Moore.

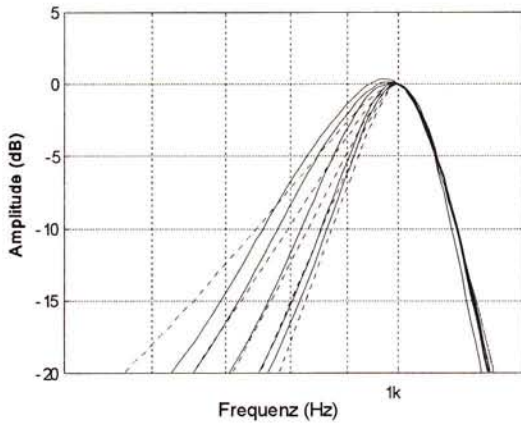


Abb. 6.18.: nAPGTF Amplitudengänge für $L = 50, 60, 70, 80$ und 90 dB/ERB an $f_c = 1$ kHz (durchgezogene Linien) im Vergleich zu [Glasberg, Moore, 1990] (strichlierte Linien)

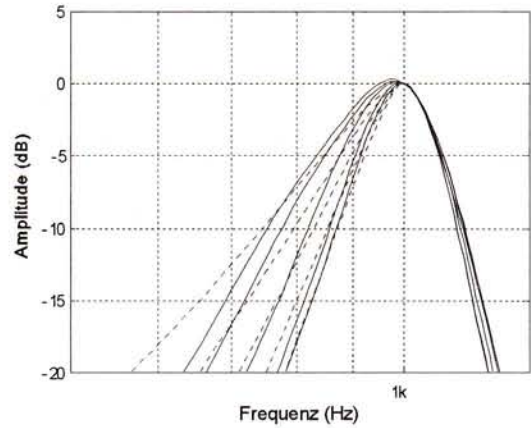
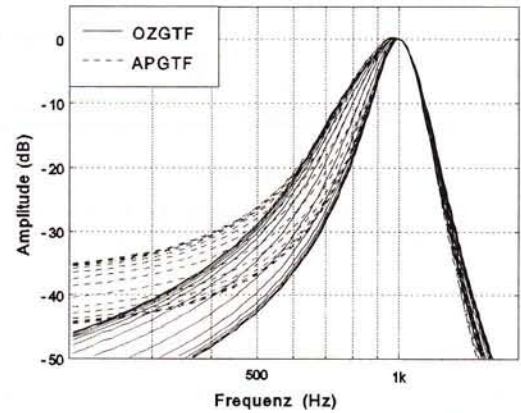


Abb. 6.19.: nOZGTF Amplitudengänge für $L = 50, 60, 70, 80$ und 90 dB/ERB an $f_c = 1$ kHz (durchgezogene Linien) im Vergleich zu [Glasberg, Moore, 1990] (strichlierte Linien)

In Abb. 6.20. sind die maximal möglichen Änderungen der tieffrequenten Flanke für 1 kHz bei Pegelwerten von $L = 40 - 120$ dB/ERB dargestellt. Mit einer entsprechenden Korrekturfunktion bei der Berechnung des Steuerungspegels können die hier dargestellten Flanken auch anderen Pegeln zugeordnet werden.



Die Impulsantworten der nAPGTF und nOZGTF sind für $f_c = 1$ kHz und $L = 50, 60, 70, 80$ und 90 dB/ERB in den beiden Grafiken Abb. 6.21. und 6.22. dargestellt.

Abb. 6.20.: Nichtlinearitäten des nAPGTF und nOZGTF bei einer Mittenfrequenz von $f_c = 1$ kHz und $L = 40, 45, 50, \dots, 115, 120$ dB/ERB

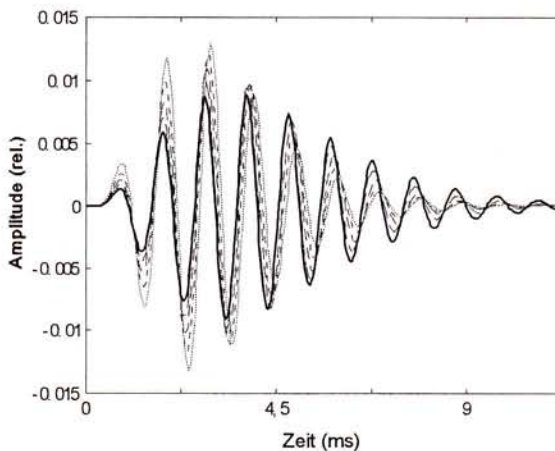


Abb. 6.21.: nAPGTF Impulsantworten für $L = 50$ (dicke Linie), $60, 70, 80$ und 90 (punktiert) dB/ERB an $f_c = 1$ kHz

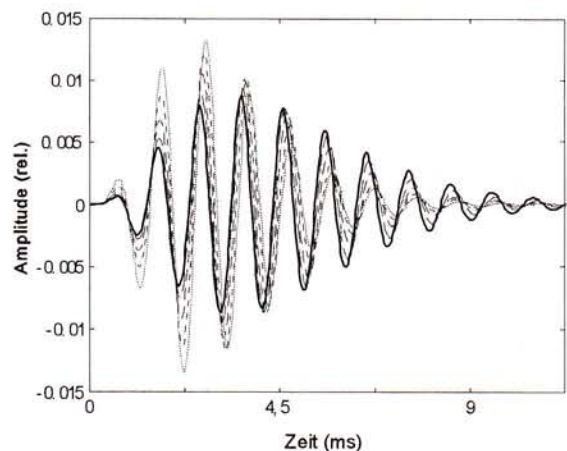


Abb. 6.22.: nOZGTF Impulsantworten für $L = 50$ (dicke Linie), $60, 70, 80$ und 90 (punktiert) dB/ERB an $f_c = 1$ kHz

Die Impulsantworten für $L = 50$ dB/ERB entsprechen den ursprünglichen Gammatone Funktionen. Diese sind in den beiden Grafiken dick dargestellt. Je höher der Steuerungspegel bzw. je flacher die tieffrequente Flanke der Filter ist, desto kürzer werden die Impulsantworten. Dies entspricht insofern der Hörwahrnehmung, als daß laute Schalle schneller wahrgenommen werden als leise.

Die Gruppenlaufzeiten der Filter sind in den Abb. 6.23 und 6.24. für 5 unterschiedliche Frequenzen ($f_c = 100, 300, 1k, 3k$ und $10k$ Hz) und 5 unterschiedliche Pegel ($L = 50, 60, 70, 80$ und 90 dB/ERB) dargestellt.

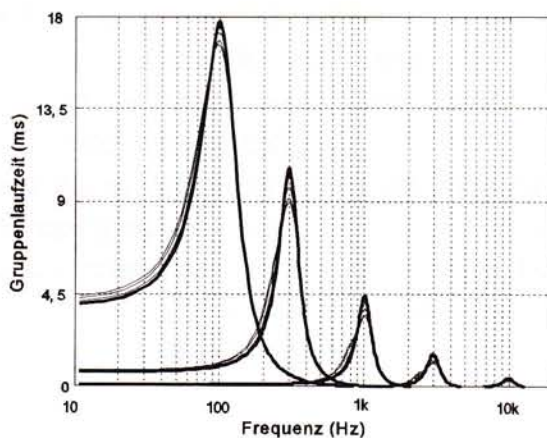


Abb. 6.23.: nAPGTF Gruppenlaufzeiten für $L = 50$ (dicke Linie), 60, 70, 80 und 90 dB/ERB an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz

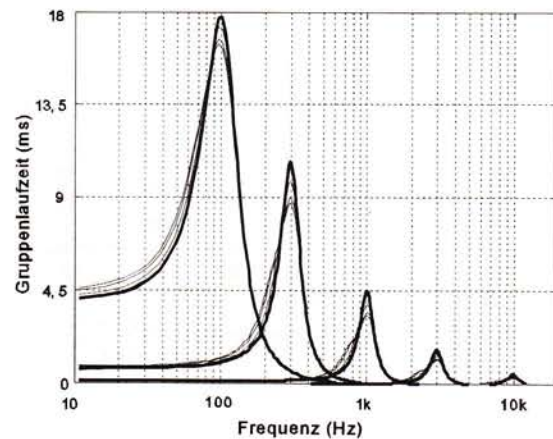


Abb. 6.24.: nOZGTF Gruppenlaufzeiten für $L = 50$ (dicke Linien), 60, 70, 80 und 90 dB/ERB an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz

Durch die Verschiebung der Polstellen ergeben sich geringfügige Verschiebungen der Amplitudengangsmaximalwerte. Diese Verschiebungen sind abhängig von der Filtermittenfrequenz und wirken sich besonders im tiefen Frequenzbereich aus (Abb. 6.25. und 6.26.).

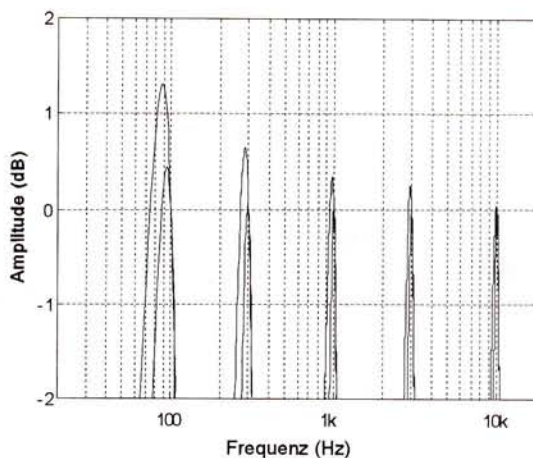


Abb. 6.25.: nAPGTF: Maximaler Frequenzverschiebungs- und Verstärkungsfaktorfehler an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz

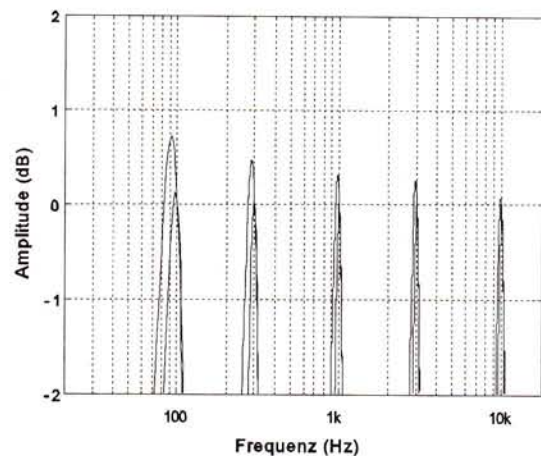


Abb. 6.26.: nOZGTF: Maximaler Frequenzverschiebungs- und Verstärkungsfaktorfehler an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz

Aus diesen Frequenzverschiebungen ergibt sich ein weiteres Problem bezüglich der Normierung auf 0 dB, da der Normierungsfaktor V_{com} sich immer auf die ursprüngliche Mittenfrequenz f_c bezieht. Dieser Fehler ist ebenfalls im tiefen Frequenzbereich am stärksten ausgeprägt (Abb. 6.25. und 6.26.). Bei Verwendung einer stufenweisen Abtastratenreduktion können diese beiden Fehler verringert werden.

Für die eingangspegelabhängige Berechnung des Steuerungsparameters L wird das Signal nach der ersten Filterstufe ausgekoppelt ($x_1(n)$, vgl. Abb. 6.8.). Die Amplitudengänge der ersten Filterstufe sind in Abb. 6.27. für $f_c = 100, 300, 1k, 3k$ und $10k$ Hz dargestellt. Diese Amplitudengänge sind vom Steuerungsparameter L unabhängig.

Für die in Abb. 6.28. und 6.29. dargestellten Erregungsmuster wurden 100 der oben angeführten nichtlinearen Gammatone Filter verwendet. Im Vergleich dazu sind Datensätze aus [Moore, Glasberg, 1987]¹ und [Zwicker, 1982]² dargestellt.

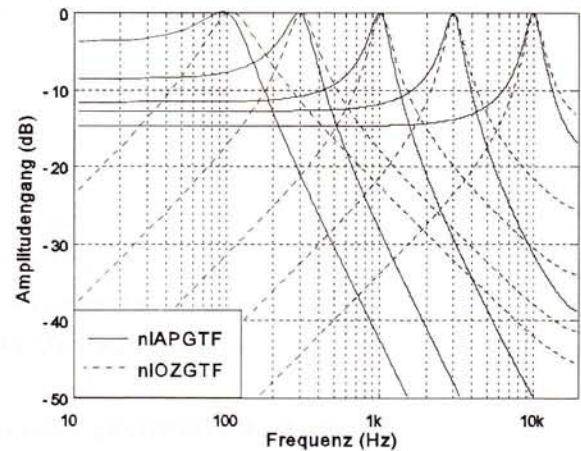


Abb. 6.27.: Amplitudengänge des $x_1(n)$ an $f_c = 100, 300, 1k, 3k$ und $10k$ Hz

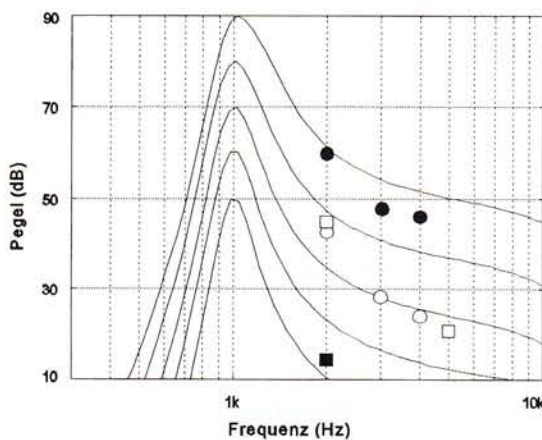


Abb. 6.28.: nIAPGTF Erregungsmuster für einen 1 kHz Sinus im Vergleich zu einigen Daten aus [Moore, Glasberg, 1987] (\circ : $L = 80$ dB/ERB, \bullet : $L = 90$ dB/ERB) und [Zwicker, 1982] (\square : $L = 80$ /ERB, \blacksquare : $L = 60$ /ERB)

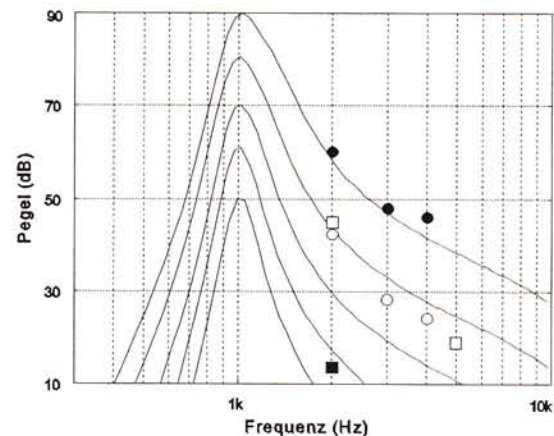


Abb. 6.29.: nIOZGTF Erregungsmuster für einen 1 kHz Sinus im Vergleich zu einigen Daten aus [Moore, Glasberg, 1987] (\circ : $L = 80$ dB/ERB, \bullet : $L = 90$ dB/ERB) und [Zwicker, 1982] (\square : $L = 80$ /ERB, \blacksquare : $L = 60$ dB/ERB)

¹Die aus [Moore, Glasberg, 1987] entnommenen Daten beziehen sich auf das Erregungsmuster eines 1 kHz Sinussignals. Die Nichtlinearität der AF wurde dabei mit dem Eingangssignal gesteuert.

²Die aus [Zwicker, 1982] entnommenen Daten ergeben sich aus den Mithörschwellen von frequenzgruppenbreitem Rauschen zentriert um 1 kHz. Die Mithörschwellen wurden um den Differenzbetrag aus Maskierer- und Mithörschwellenpegel an der Mittenfrequenz des Maskierers angehoben.

Für diese Erregungsmuster wurde weder eine eingangs- noch eine ausgangspiegelabhängige Steuerung verwendet. Bei einem 1 kHz Pegel von beispielsweise 90 dB wurde angenommen, daß alle auditiven Filter eine untere Flankensteilheit entsprechend einem Pegel von $L = 90$ dB/ERB besaßen. Diese beiden Grafiken stellen damit eine erste grundsätzliche Abschätzung der erreichbaren Erregungsmuster eines 1 kHz Sinus Signals dar. Durch die verwendete Steuerung der Filter wird das Erregungsmuster des Maskierers überbewertet.

Die nichtlineare Auffächerung der hochfrequenten Erregungsmusterflanke ('upward spread of masking') ist bei den nIOZGTF nicht so deutlich ausgeprägt wie bei den nIAPGTF. Die nIOZGTF entsprechen sowohl den Erregungsmustern nach [Moore, Glasberg, 1987] als auch den aus den Mithörschwellen berechneten Erregungsmustern deutlich besser als die nIAPGTF. Dies gilt vorwiegend im Bereich $f \gg 1$ kHz. Die nIAPGTF Erregungsmuster sind im Bereich $f \gg 1$ kHz zu hoch. Dies gilt jedoch nur für den hier betrachteten Fall, daß alle AF eine tieffrequente Flankensteilheit entsprechend $L = 90$ dB/ERB besitzen.

Zur Abschätzung der Amplitudengangsänderungen einer gesamten Filterbank wird zunächst eine nichtlineare Gammatone Filterbank mit 40 Filtern ($f_{\min} = 40$ Hz, $f_s = 44,1$ kHz) betrachtet (Abb. 6.30. und 6.31.).

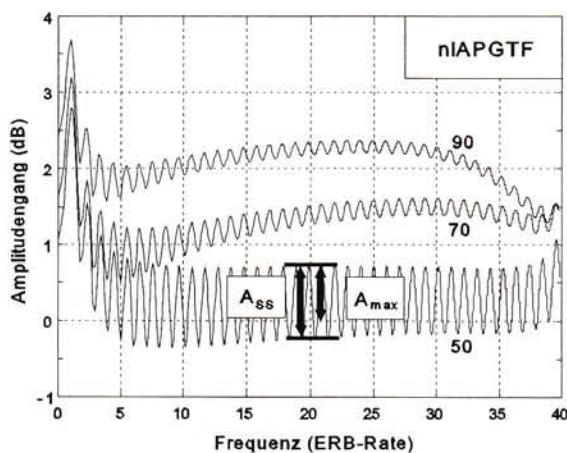


Abb. 6.30.: Amplitudengänge einer nIAPGTF Filterbank mit 40 Filtern bei $L = 50, 70$ und 90 dB/ERB.

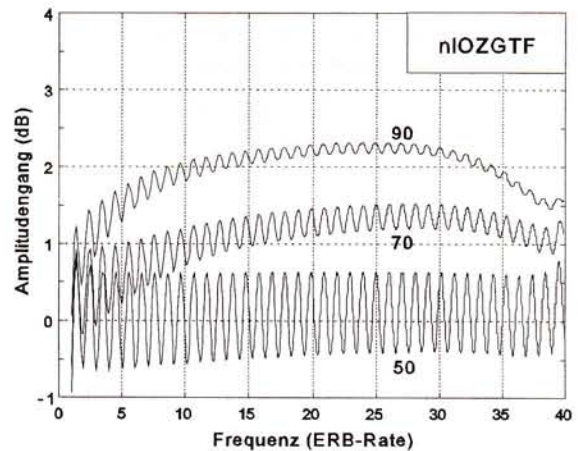


Abb. 6.31.: Amplitudengänge einer nIOZGTF Filterbank mit 40 Filtern bei $L = 50, 70$ und 90 dB/ERB.

Der Normierungsfaktor zur Bestimmung der Amplitudenpegel wurde so gewählt, daß ein einzelner nIGTF 0 dB an f_c hätte. Die jeweils 3 Kurven innerhalb der beiden Grafiken Abb. 6.30. und 6.31. beziehen sich auf $L = 50, 70, 90$ dB/ERB Filter. Die $L = 50$ dB/ERB Kurven bedeuten beispielsweise, daß alle 40 Filter eine untere Flankensteilheit entsprechend $L = 50$ dB/ERB besitzen¹.

Bei den breiteren Filtern verringert sich die Welligkeit A_{ss} (Abb. 6.30.) der Gesamtübertragung.

¹ Da sich die hier betrachteten nIGTF bei einem Steuerungspegel von $L \leq 50$ dB/ERB kaum ändern, wird davon ausgegangen, daß die $L = 50$ dB/ERB Filter die schmalbandigste Realisierungsform darstellen.

Der überproportional große Anstieg der nIAPGTF Amplitudengänge bei kleinen ERB-Raten (0-3 ERB-Raten, Abb. 6.30.) ergibt sich daraus, daß die nIAPGTF bei niederen Filtermittenfrequenzen in eine Tiefpaßcharakteristik übergehen. Dieser Effekt ist bei den nIOZGTF aufgrund der Polstelle an $z = 1$ nicht vorhanden (Abb. 6.31.).

Bei mehr als 40 Filtern verringert sich die Welligkeit der Amplitudengänge (die Filter rücken näher zusammen). In den beiden Abb. 6.32. und 6.33. sind die Amplitudengänge einer Filterbank mit 80 Filtern dargestellt. Die Welligkeit der Amplitudengänge ist nicht mehr vorhanden, dafür ist die Verstärkung (A_{\max} , vgl. Abb. 6.30.) dieser Gammatone Filterbank deutlich höher.

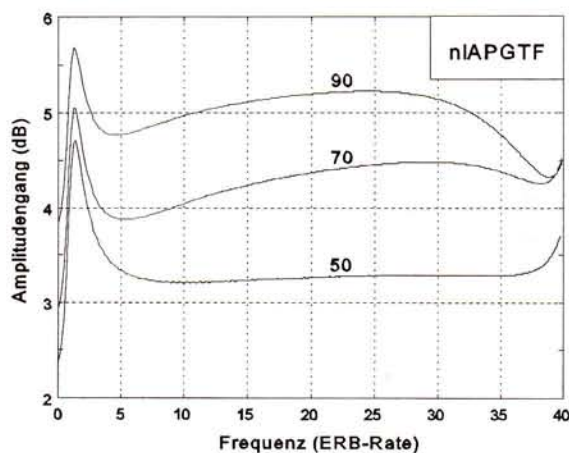


Abb. 6.32.: Amplitudengänge einer nIAPGTF Filterbank mit 80 Filtern bei $L = 50, 70$ und 90 dB/ERB.

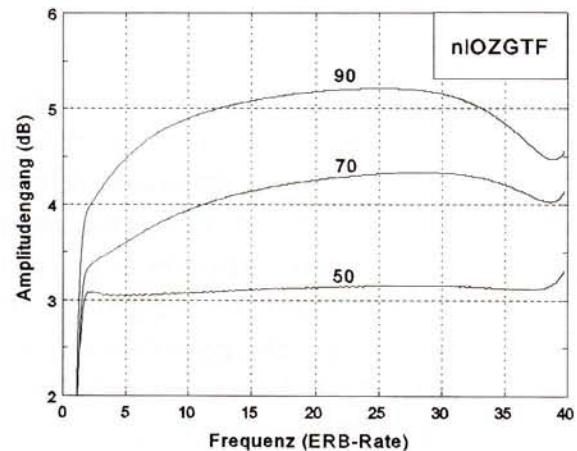


Abb. 6.33.: Amplitudengänge einer nIOZGTF Filterbank mit 80 Filtern bei $L = 50, 70$ und 90 dB/ERB.

Die Welligkeit einer Filterbank ist für die Steuerungsparameter $L \leq 50$ dB/ERB am stärksten ausgeprägt. Zur Abschätzung der maximalen Welligkeit A_{ss} und des Verstärkungsfaktors A_{\max} (bezogen auf die 0 dB eines einzelnen Filters) einer Filterbank sind in Abb. 6.34. diese beiden Parameter als Funktion der Filteranzahl für $L = 50$ dB/ERB dargestellt. Die beiden Kennwerte A_{ss} und A_{\max} wurden dafür jeweils aus dem Bereich um 20 ERB-Raten entnommen.

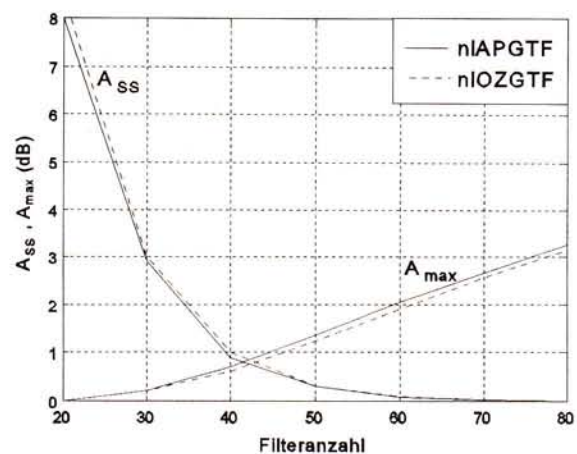


Abb. 6.34.: Maximale Welligkeit A_{ss} und Verstärkungsfaktor A_{\max} für $L = 50$ dB/ERB als Funktion der Filteranzahl.

Die Simulation der Gammatone Filterbank wurde stets mit einer Abtastrate von 44,1 kHz, der transponierten Direktform II und doppelter Fließkommazahlengenauigkeit durchgeführt. Der Vorteil der Direktform liegt in der geringsten Anzahl von Multiplikationen und Additionen. Durch die hohe Abtastrate von 44,1 kHz befinden sich dabei jedoch viele Polstellen im ersten Quadranten der z-Ebene nahe dem Einheitskreis. Dies erfordert eine hohe Koeffizientengenauigkeit, um zu vermeiden, daß durch die Koeffizientenquantisierung die Polstellen auf den Einheitskreis springen¹. Für eine effiziente Implementation sollte demgegenüber die Wortbreite der Koeffizienten möglichst gering sein. Andere Filterstrukturen wie beispielsweise die Kingsbury Struktur² oder die Verwendung einer Abtastratenreduktion erlauben Filterkoeffizienten mit geringerer Wortbreite.

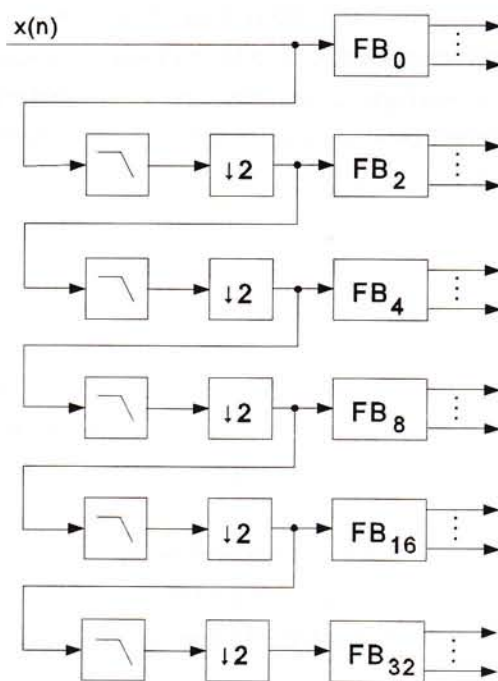


Abb. 6.35.: Abtastratenreduzierte Implementation einer nIGTFB

Im folgenden werden die Eigenschaften einer 5-stufig abtastratenreduzierten Implementation (jeweils um den Faktor 2, Abb. 6.35.³) einer Realisierungsform mit konstanter Abtastrate gegenübergestellt. Als Anti-Aliasing Filter wird ein Butterworth-Filter 5ter Ordnung (Grenzfrequenz = $f_s * 0,18$ Hz) verwendet.

Um Verzerrungen der Amplitudengänge zu vermeiden, sollten sich die Polstellen nicht in der Nähe der Frequenz 0 Hz oder der halben Abtastrate befinden.

Die Filtermittenfrequenzbereiche der nächst folgenden Abtastrate ergeben sich durch Halbierung des vorhergehenden Frequenzbereichs. Dadurch befinden sich die Polstellen der verwendeten Filter (mit Ausnahme der ersten und letzten Stufe) stets im gleichen Bereich der z-Ebene ($\varphi_{Lin} = 20$ bis 50 Grad).

In der Tab. 6.1. sind die verwendeten Filtermittenfrequenzbereiche in Abhängigkeit der Abtastrate angeführt:

¹ Befindet sich mindestens eine Polstelle auf oder außerhalb des Einheitskreises, so ist der Filter instabil.

² [Zölzer, 1994]

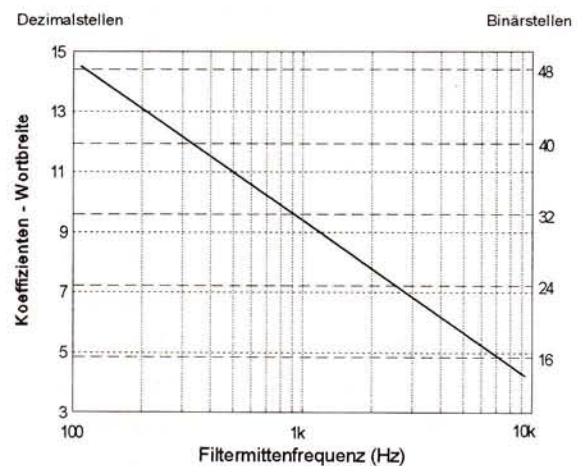
³ Die Bezeichnung FB_i in Abb. 6.35. steht für eine um den Faktor i abtastreduzierte Filterbank.

Abtastrate (Hz)	44100	22050	11025	5512,5	2756,25	1378,125
Abtastratenreduktion		↓2	↓4	↓8	↓16	↓32
verwendeter Filtermittenfrequenzbereich (Hz)	22k ↑ 3k	3k ↑ 1,5k	1,5k ↑ 750	750 ↑ 375	375 ↑ 187	187 ↑ 20

Tab. 6.1.: Filtermittenfrequenzbereiche bei unterschiedlichen Abtastraten

In Abb. 6.36. ist die mindestens erforderliche Koeffizientengenauigkeit¹ bei einer Abtastrate von $f_s = 44,1$ kHz als Funktion der Filtermittenfrequenz angegeben. Dieser Zusammenhang ist logarithmisch und gilt näherungsweise für nLPGTF und nOZGTF. Bei einer Abtastratenreduktion um den Faktor i wird davon ausgegangen, daß sich die Filtermittenfrequenzen dieser Grafik um diesen Faktor dividieren lassen.

Aus dieser Grafik ist ersichtlich, daß für die oben angeführte abtastratenreduzierte Implementation eine Koeffizientenwortbreite von 24 Bit ausreichend ist.

Abb. 6.36.: Abschätzung der mindestens erforderlichen Koeffizientengenauigkeit bei einer Abtastrate von $f_s = 44,1$ kHz

Nimmt man die Zeitkomplexität einer einzelnen Filterkaskade mit N pro Abtastwert an, so ergibt sich für eine nicht abtastratenreduzierte Filterkaskade pro Sekunde eine Zeitkomplexität von $ZK = N * 44100$. Für eine Filterbank von beispielsweise 40 Filtern ergibt sich $ZK = 40 * N * 44100$.

Bei einer stufenweisen Abtastratenreduktion (vgl. Abb. 6.35.) reduziert sich die Zeitkomplexität pro Sekunde auf

$$ZK = F_0 * N * f_s + F_2 * N * \frac{f_s}{2} + F_4 * N * \frac{f_s}{4} + F_8 * N * \frac{f_s}{8} + F_{16} * N * \frac{f_s}{16} + F_{32} * N * \frac{f_s}{32}$$

Gl. 6.20.

mit F_i ... Anzahl der Filtermittenfrequenzen im entsprechenden Abtastratenfrequenzbereich

¹ Für den Zusammenhang zwischen der Filtermittenfrequenz und der mind. erforderlichen Koeffizientenwortbreite wurden die Koeffizienten mit doppelter Fließkommagenauigkeit berechnet und anschließend quantisiert.

Wird den Anti-Aliasing Filtern der Einfachheit halber ebenfalls eine Zeitkomplexität N zugeordnet, so ergibt sich für das oben angeführte Beispiel einer 5-stufigen Abtastratenreduktion eine Rechenaufwandsersparnis von etwa 43 % gegenüber einer nichtabtastratenreduzierten Implementation. Dieser Wert ist von der Größe der Filterbank weitgehend unabhängig.

Die in Abb. 6.25. und 6.26. qualitativ dargestellten maximalen Frequenzverschiebungs- und Verstärkungsfaktorfehler eines einzelnen Filters sind in den beiden folgenden Grafiken (Abb. 6.37. und 6.38.) für Filtermittenfrequenzen zwischen 50 Hz und 1kHz (in diesem Bereich sind diese Fehler am stärksten ausgeprägt) quantifiziert. Dabei wird zwischen einer 5-stufig abtastreduzierten und einer nicht abtastreduzierten Implementation unterschieden.

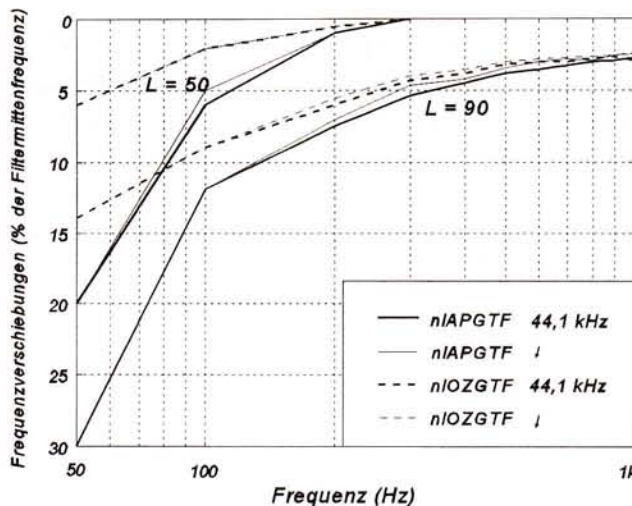


Abb. 6.37.: Maximaler Frequenzverschiebungsfehler für nIAPGTF und nIOZGTF

In Abb. 6.38. sind zusätzlich die Filtermittenfrequenzbereiche der einzelnen abtastreduzierten Stufen (132, 116, ...) dargestellt. Die Bezeichnungen $L = 50$ und $L = 90$ beziehen sich auf den verwendeten Steuerungsparameter und gelten jeweils für die 4 nächst liegenden Kurven.

Die durch die Abtastratenreduktion erreichte Verringerung dieser beiden Fehler ist relativ gering. Als Hauptargumente zur Verwendung einer Abtastratenreduktion bleiben daher die geringe Anforderung an die Koeffizientengenauigkeit und die Rechenzeitersparnis.

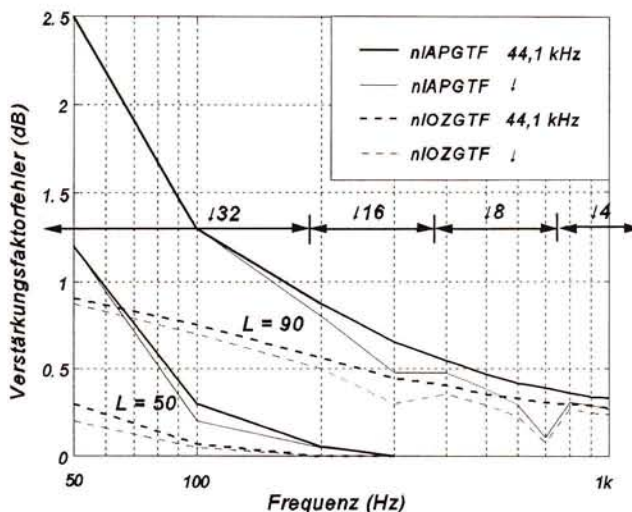


Abb. 6.38.: Maximaler Verstärkungsfaktorfehler für nIAPGTF und nIOZGTF

6.2.2. Parametrisierung ausgangspegelabhängiger nichtlinearer APGTF und OZGTF

Im Unterschied zu der in Kap. 6.2.1. angeführten Parametrisierung eingangspegelabhängiger Filter wird bei dieser Parametrisierung der Radius der Polstellen (r_{Lin}) nicht konstant gehalten. Es wird sowohl die Phase φ_{Lin} als auch der Radius r_{Lin} pegelabhängig verschoben d.h. mit einem Faktor k_φ bzw. k_r multipliziert.

$$p_{1,2} = k_{r_i} r_{Lin} e^{\pm j k_{\varphi_i} \varphi_{Lin}} \quad \text{Gl. 6.21.}$$

$$i = 2,3,4$$

Die Koeffizienten des Zählerpolynoms ändern sich, im Vergleich zur eingangspegelabhängigen Parametrisierung, nicht (Gl. 6.16. und 6.17.). Die pegelabhängigen Koeffizienten der Filterstufen $H_2(z) - H_4(z)$ werden zu (nlAPGTF und nlOZGTF):

$$a_{1,i} = -2 k_{r_i} r_{Lin} \cos(k_{\varphi_i} \varphi_{Lin}) = -2 \frac{k_{r_i}}{e^{BT}} \cos(k_{\varphi_i} 2\pi f_c T) \quad a_{2,i} = (k_{r_i} r_{Lin})^2 = \frac{k_{r_i}^2}{e^{2BT}} \quad \text{Gl. 6.22.}$$

$$i = 2,3,4$$

Die Normierung der ersten Filterstufe $H_1(z)$ auf 0 dB an f_c bleibt ebenfalls gleich (Gl. 6.18.), da die erste Filterstufe nicht parametrisiert wird. Die Normierung der drei pegelabhängigen Filterstufen V_{com} wird zu:

$$V_{com} = \left| 1 + E e^{-j2\pi f_c T} + F e^{-j4\pi f_c T} + G e^{-j6\pi f_c T} + H e^{-j8\pi f_c T} + I e^{-j10\pi f_c T} + J e^{-j12\pi f_c T} \right|$$

$$\begin{aligned} A &= a_{1,2} + a_{1,3} & C &= a_{1,2} a_{2,3} + a_{2,2} a_{1,3} \\ B &= a_{2,2} + a_{2,3} + a_{1,2} a_{1,3} & D &= a_{2,2} a_{2,3} \end{aligned} \quad \text{Gl. 6.23.}$$

$$\begin{aligned} E &= A + a_{1,4} & H &= B a_{2,4} + C a_{1,4} + D \\ F &= A a_{1,4} + B + a_{2,4} & I &= C a_{2,4} + D a_{1,4} \\ G &= A a_{2,4} + B a_{1,4} + C & J &= D a_{2,4} \end{aligned}$$

Der Verstärkungsfaktor V_{com} kann wiederum für beide Filtertypen (nlAPGTF oder nlOZGTF) verwendet werden.

Diese Parametrisierung ist unter Verwendung der transponierten Direktform II als digitale Filterstruktur in Abb. 6.39. dargestellt. Im Unterschied zur eingangspegelabhängigen

Parametrisierung müssen durch die Änderung der Polstellenradien auch die Koeffizienten $a_{2,i}$ ($i = 2,3,4$) verändert werden.

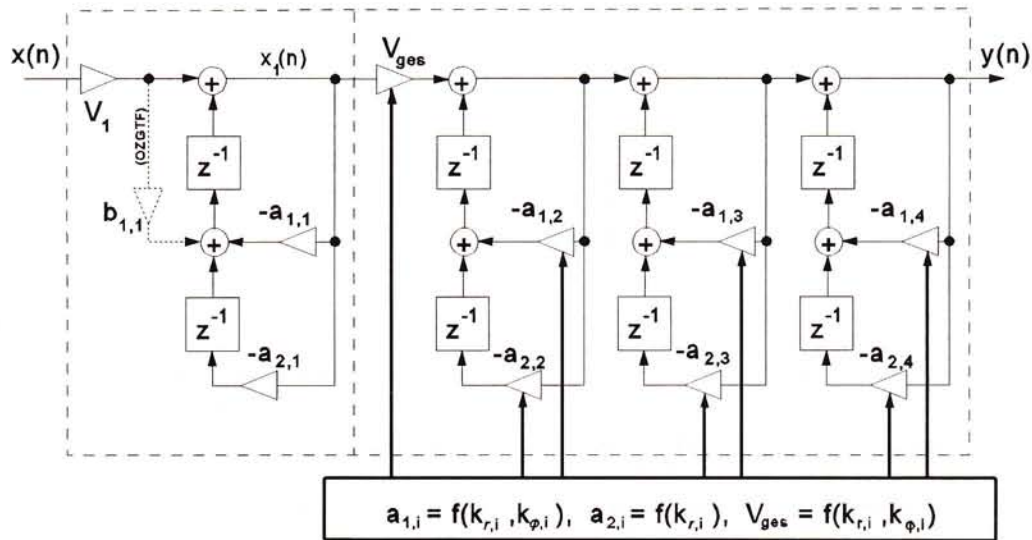


Abb. 6.39.: Nichtlineare nIAPGTF und nIOZGTF

Anpassung der Amplitudengänge mit 5 freien Parametern

Um die tieffrequente Filterflanke pegelabhängig zu heben, werden zwei Polstellenpaare bei Verringerung ihres Radius zu tiefen Frequenzen hin verschoben ($k_{r2} \leq 1, k_{r3} \leq 1, k_{\phi2} \leq 1, k_{\phi3} \leq 1$). Mit einem Polstellenpaar wird die Filtermittenfrequenz konstant gehalten ($k_{r4} = 1, k_{\phi4} \geq 1$). D.h. von den 6 möglichen pegelabhängigen Parametern werden in dieser ersten Anpassung 5 tatsächlich verwendet.

In den beiden Tabellen 6.2. und 6.3. sind der mittlere quadratische Fehler (rms) und der maximale Fehler zwischen den erreichten Amplitudengängen und den roex(p,r) Filterformen von [Rosen, Baker, 1994] angeführt. Die Amplitudengänge wurden dafür mit einer iterativen Prozedur unter Minimierung des mittleren quadratischen Fehlers angepaßt.

Tab. 6.2.: rms and max. Fehler (dB) zwischen den auditiven Filtern nach [Rosen, Baker, 1994] und den Amplitudengängen des nIAPGTF bei 5 veränderlichen Parametern.

	Steuerungspegel P_s (dB)			
	30	40	50	60
rms Fehler	0,39	0,51	0,30	0,23
max. Fehler	1,23	1,23	0,88	0,83

Tab. 6.3.: rms and max. Fehler (dB) zwischen den auditiven Filtern nach [Rosen, Baker, 1994] und den Amplitudengängen des nIOZGTF bei 5 veränderlichen Parametern.

	Steuerungspegel P_s (dB)			
	30	40	50	60
rms Fehler	0,56	0,23	0,15	0,14
max. Fehler	1,25	0,56	0,61	0,50

In den beiden Abbildungen 6.40. und 6.41. sind diese Filter im Vergleich zu den auditiven Filtern nach [Rosen, Baker, 1994] dargestellt.

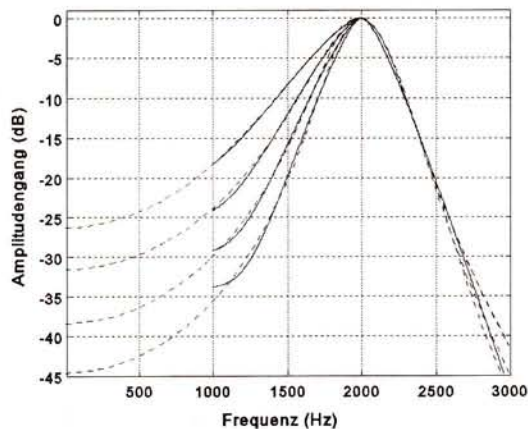


Abb. 6.40.: Amplitudengänge des 5 parametrischen nLPGTF (strichliert) im Vergleich zu den auditiven Filtern nach [Rosen, Baker, 1994] für $P_s = 30, 40, 50, 60$ dB

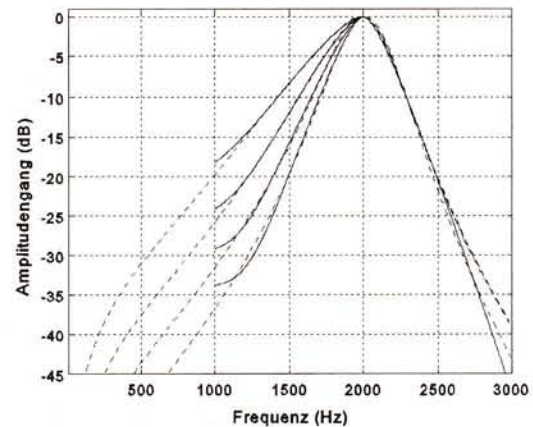


Abb. 6.41.: Amplitudengänge des 5 parametrischen nOZGTF (strichliert) im Vergleich zu den auditiven Filtern nach [Rosen, Baker, 1994] für $P_s = 30, 40, 50, 60$ dB

Der Nachteil einer Anpassung mit 5 freien Parametern besteht darin, daß im Betrieb einer derartigen Filterbank diese 5 Parameter pegelabhängig verändert werden müßten, was sehr rechenaufwendig ist.

Anpassung der Amplitudengänge mit 2 freien Parametern

Um die Steuerungseinheit der Filter möglichst einfach bzw. recheneffizient implementieren zu können, werden 2 Polstellenpaare mit identischen Parametern betrieben: $k_r = k_{r2} = k_{r3}$, $k_\varphi = k_{\varphi2} = k_{\varphi3}$. Das Polstellenpaar der 4. Stufe ($H_4(z)$) wird auf konstant $k_{r4} = 1$ und $k_{\varphi4} = 1,039$ gesetzt. Diese vereinfachte Parametrisierung wird nur für die nOZGTF durchgeführt, da diese Filter für ausgangspegelabhängige Bedingungen besser geeignet zu sein scheinen¹.

Das Verhältnis der beiden pegelabhängigen Parameter k_r und k_φ zum Steuerungspegel P_s wird angegeben mit²:

$$k_r = 0,9178 + 0,2131 * 10^{-\frac{P_s}{50}}, \quad k_\varphi = 1,0772 - 0,0171 * 10^{\frac{P_s}{45}} \quad \text{Gl. 6.24.}$$

¹ Diese Annahme gründet auf einem Schriftverkehr mit Dr. Roy Patterson.

² Die Bezeichnung für den ausgangspegelabhängigen Steuerungspegel P_s wurde, ebenso wie die Bezeichnung für den eingangspegelabhängigen Steuerungspegel L , aus der Literatur übernommen ([Rosen, Baker, 1994]).

Gl. 6.24. ist für Pegel von $P_s = 30$ bis 60 dB gültig. In der praktischen Anwendung wird innerhalb der Steuerungseinheit der Pegel nicht berechnet, um Rechenzeit zu sparen. Es ist also erstrebenswert, die Beziehung zwischen der Leistung und den Steuerungsparametern in einem direkten, einfachen Verhältnis auszudrücken.

Durch den exponentiellen Zusammenhang in Gl. 6.24. vereinfacht sich das Verhältnis zwischen den Parametern k_r , k_ϕ und der Steuerungsleistung p_s zu:

$$k_r = 0,9178 + 0,2131 * p_s^{-0,2}, \quad k_\phi = 1,0772 - 0,0171 * p_s^{0,22} \quad \text{Gl. 6.25.}$$

In der Tab. 6.4. sind der mittlere quadratische Fehler (rms) und der maximale Fehler zwischen den erreichten Amplitudengängen des 2 parametrischen nOZGTF und den roex(p,r) Filterformen von [Rosen, Baker, 1994] angeführt. Durch die vereinfachte Parametrisierung ergeben sich wiederum leichte Verschiebungen der Filtermittenfrequenzen. Daraus resultiert, ebenso wie bei der eingangspegelabhängigen Parametrisierung, ein geringfügiger Verstärkungsfehler. Diese beiden Fehler sind in Tab. 6.5. dargestellt.

Tab. 6.4.: rms and max. Fehler (dB) zwischen den auditiven Filtern nach [Rosen, Baker, 1994] und den Amplitudengängen des nOZGTF bei 2 veränderlichen Parametern.

	Steuerungspegel P_s (dB)			
	30	40	50	60
rms Fehler	0,57	0,30	0,20	0,25
max. Fehler	1,30	0,84	0,49	0,60

Tab. 6.5.: Verschiebungen der Filtermittenfrequenzen (f_c Versch. in % bezogen auf f_c) und Verstärkungsfehler (Amp. Fehler in dB) des nOZGTF bei 2 veränderlichen Parametern.

	Steuerungspegel P_s (dB)			
	30	40	50	60
f_c Versch.	0,75	0,10	-0,50	-1
Amp. Fehler	0,05	0,01	0,02	0,07

Aus diesen beiden Tabellen ist ersichtlich, daß die pegelabhängige Parametrisierung eines Gammatone Filters mit nur 2 freien Parametern möglich ist. Die Fehlerzunahme im Vergleich zu einem 5 parametrischen Gammatone Filter ist minimal.

In Abb. 6.42. sind die auditiven Filter nach [Rosen, Baker, 1994] im Vergleich zu den Amplitudengängen der 2 parametrischen nOZGTF dargestellt. Die Impulsantworten dieses Filters sind in Abb. 6.43. dargestellt. Sie werden, ebenso wie bei eingangspegelabhängigen Filtern mit steigendem Steuerungspegel kürzer.

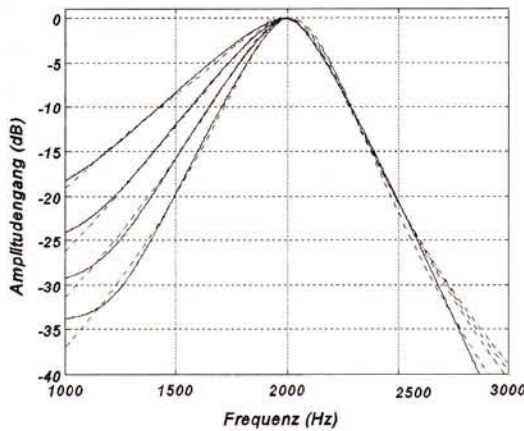


Abb. 6.42.: Amplitudengänge des 2-parametrischen nOZGTF (strichliert) im Vergleich zu den auditiven Filtern nach [Rosen, Baker, 1994] für $P_s = 30, 40, 50, 60$ dB

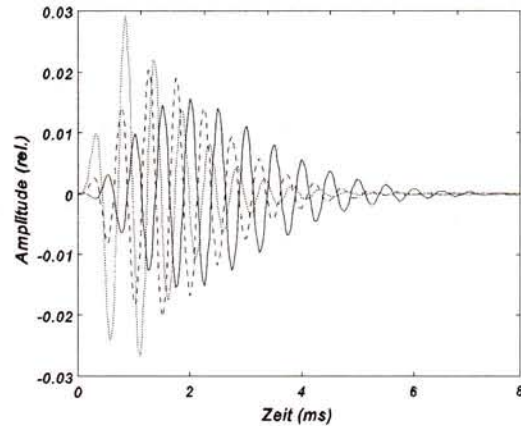


Abb. 6.43.: Gammatone Funktion (durchgezogene Linie), nOZGTF ($P_s = 30$ dB, strichlierte Linie) und nOZGTF ($P_s = 60$ dB, punktierte Linie)

Die pegelabhängigen Filterkoeffizienten werden als Funktion der beiden Steuerungsparameter k_r und k_φ zu:

$$a_{2,2} = a_{2,3} = \frac{k_r^2}{e^{2BT}}$$

$$a_{1,2} = a_{1,3} = -2 \frac{k_r}{e^{BT}} \cos(k_\varphi 2\pi f_c T)$$

Gl. 6.26.

Diese einfache Parametrisierung gilt, im Unterschied zur eingangspegelabhängigen Parametrisierung, nur für eine Filtermittenfrequenz von $f_c = 2$ kHz. Um eine vollständige Filterbank aufbauen zu können, ist eine Transposition zu anderen Filtermittenfrequenzen notwendig.

Transposition zu anderen Filtermittenfrequenzen¹

Die im folgenden angeführte Transposition geht von folgender Annahme aus: Die qualitativen Eigenschaften der auditiven Filter sind unabhängig von der Filtermittenfrequenz d.h. die tieffrequenten Filterflanken steigen mit steigendem Steuerungspegel, während die hochfrequenten Filterflanken pegelunabhängig sind.

Unter dieser Voraussetzung kann die kombinierte Abhängigkeit der Steuerungsparameter k_r und k_φ von der Filtermittenfrequenz und dem Steuerungspegel ($k_i = f(f_c, P_s)$) aufgespalten werden in

¹ Diese Transposition wird in einer allgemeinen Form angegeben, da für ausgangspegelabhängige auditive Filter nicht genügend Notched-Noise Daten existieren, um eine Filterbank über den gesamten hörbaren Frequenzbereich aufbauen zu können.

Basierend auf einem Schriftverkehr mit Dr. Stuart Rosen und Dr. Richard Baker wird davon ausgegangen, daß diese Transposition in weiteren Experimenten bestätigt wird.

zwei unabhängige Funktionen, in eine Funktion der Filtermittenfrequenz und eine Funktion des Steuerungspegels.

Ausgehend von einem Prototyp-Filter (als Prototyp kann beispielsweise der oben angeführte 2 kHz Filter verwendet werden) wird dessen nichtlineares Verhalten an eine beliebige Mittenfrequenz transponiert.

Die Polstellenpaare des nichtlinearen Prototyps an $f_{c,pro}$ sind (vgl. Gl. 6.21.):

$$P_{1,2,i,pro} = k_{r_i,pro} r_{Lin,pro} e^{\pm j k_{\varphi_i,pro} \varphi_{Lin,pro}} \quad \text{Gl. 6.27.}$$

Um diesen Filter an eine beliebige andere Mittenfrequenz f_c' verschieben zu können, wird ein Bandbreitenverhältnis $R_{bw}(f_c')$ definiert. Dieses Bandbreitenverhältnis setzt die Bandbreite des schmalbandigsten Filters an f_c' (d.h. den Filter mit dem niedersten Steuerungspegel ($P_{s,min}$)) und die Bandbreite des schmalbandigsten Prototyp-Filters (d.h. wiederum mit dem niedersten Steuerungspegel ($P_{s,min}$)) in Beziehung.

$$R_{bw}(f_c') = \frac{B_{f_c', P_{s,min}}}{B_{pro, P_{s,min}}} \quad \text{Gl. 6.28.}$$

Mit Gl. 6.28. werden die Radien der Polstellenpaare an f_c' zu:

$$r_{i, f_c'} = (k_{r_i,pro} r_{Lin,pro})^{R_{bw}(f_c')} = k_{r_i,pro}^{R_{bw}(f_c')} r_{Lin,pro}^{R_{bw}(f_c')} \quad \text{Gl. 6.29.}$$

Mit dem Bandbreitenverhältnis $R_{bw}(f_c')$ und dem Transpositionsfaktor $trans(f_c')$ werden die Polstellenwinkel an f_c' zu:

$$\varphi_{i, f_c'} = \varphi_{Lin,pro} R_{bw}(f_c') \left[k_{\varphi_i,pro} + \left(\frac{trans(f_c')}{R_{bw}(f_c')} - 1 \right) \right] \quad \text{Gl. 6.30.}$$

mit: $trans(f_c') = \frac{f_c'}{f_{c,pro}}$

Aus dieser Transposition ergeben sich neue Filter an f_c' . Diese Filter besitzen einen vorgegebenen Amplitudengang für den kleinsten Steuerungspegel. Die Steuerungsparameter $k_{i,pro}$ sind unabhängig von der jeweiligen Mittenfrequenz. Die Nachbildung auditiver Filter an beliebigen Filtermittenfrequenzen vereinfacht sich daher zu einer Beziehung zwischen dem Steuerungspegel P_s' und der Zwischengrößen $P_{s,pro}$.

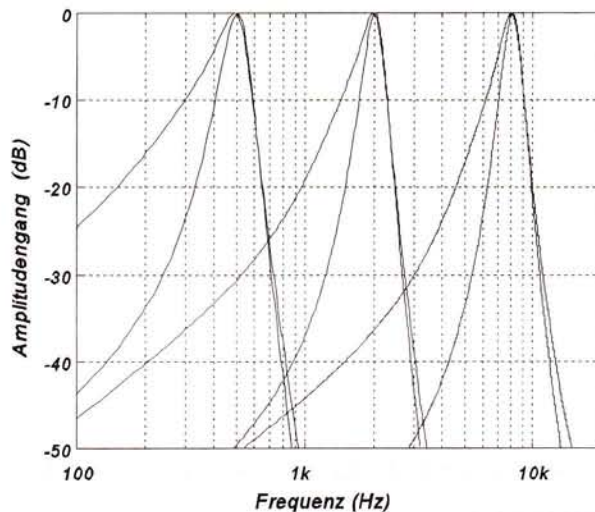


Abb. 6.44.: Amplitudengänge des 2 parametrischen nOZGTF mit: $f_c' = 500, 2k, 8k$ Hz und $P_{s,pro} = 30, 60$ dB

In Abb. 6.44. ist diese Transposition mit dem 2 parametrischen nOZGTF an 2 kHz als Prototype-Filter für $f_c' = 500, 2k, 8k$ Hz und $P_{s,pro} = 30, 60$ dB dargestellt. Das Bandbreitenverhältnis $R_{bw}(f_c')$ wurde mit der ERB-Breiten Gleichung nach [Glasberg, Moore, 1990] berechnet (Gl. 6.31.).

$$R_{bw}(f_c') = \frac{ERB(f_c')}{ERB(2kHz)} \quad \text{Gl. 6.31.}$$

$$ERB(f) = 24,7 + 0,108f$$

Aufgrund von fehlenden ausgangspiegelabhängigen auditiven Filterformen basierend auf Notched-Noise Experimenten kann diese Transposition zur Zeit nicht verifiziert werden.

Kapitel 7

Diskussion

In diesem Kapitel werden die Randbedingungen und Ergebnisse der vorliegenden Arbeit zusammengefaßt.

Modelle der menschlichen Hörwahrnehmung

Die Psychoakustik läßt sich in zwei Themenbereiche gliedern, in die Analyse der Hörwahrnehmung und in die Synthese der daraus gewonnenen Ergebnisse zu einem Modell der Wahrnehmung. Während man bei der Analyse versucht, bestimmte Eigenschaften zu extrahieren und zu quantifizieren, so beschäftigt man sich bei der Synthese damit die physiologischen und psychologischen Reaktionen auf akustische Reize möglichst exakt und recheneffizient nachzubilden. Die Entwicklung auditiver Modelle ist also mit experimentellen Daten eng verknüpft. Je mehr bzw. je genauer die zur Verfügung stehenden Daten sind, desto genauer werden die Vorhersagen der entwickelten Modelle. Umgekehrt können aus Modellen weiterführende Analysen abgeleitet werden (vgl. Kap. 2).

Ein wesentliches Kriterium ist der Gültigkeitsbereich der Modelle. In erster Linie erstreckt sich die Gültigkeit eines Modells über diejenigen Versuchsbedingungen, aus denen sich das Modell ableitet (vgl. Bestimmung der Kurven gleicher Lautstärkepegel, Kap. 3.2.). Besonders in der Anwendung ist die Erweiterung dieses Gültigkeitsbereiches jedoch von besonderem Interesse. Ein Lautheitsbewertungsverfahren sollte beispielsweise in der Lage sein, für alle möglichen Arten von Schallreizen das durchschnittliche Antwortverhalten einer Gruppe normalhörender Personen zu reproduzieren und dadurch zu objektivieren (vgl. Kap. 3.7.).

Die Wahrnehmung akustischer Reize ist keine passive Projektion sondern ein aktiver Prozeß, in dem der Hörer aktiv entscheidet, welche Teile der eintreffenden Informationen wahrgenommen bzw. als 'wahr' angenommen werden. Das Konzept des Leistungsdichtespektrum-Modells (vgl. Kap. 2.4.) ist zwar kein vollständiger Ersatz für diesen komplexen Prozeß, es eignet sich jedoch

sehr gut als Ausgangspunkt für ein 'Herantasten' an die Hörwahrnehmung in ihrer Gesamtheit.

Den meisten Modellen der menschlichen Hörwahrnehmung ist die strukturelle Anlehnung an den anatomischen Aufbau des peripheren Gehörs gemeinsam. Dies läßt sich darauf zurückführen, daß sich die einzelnen Verarbeitungsstufen des peripheren Gehörs sowohl anatomisch als auch physiologisch voneinander unterscheiden lassen. Bei der zentralen d.h. neuronalen Verarbeitung von Schallreizen ist dies nicht möglich (vgl. Kap. 4).

Modelle des peripheren Gehörs

Im peripheren Gehör wird bereits eine teilweise Analyse der Schallsignale und damit verbunden eine Datenreduktion durchgeführt. Dies bedeutet, daß nicht die gesamte akustische Information, die das Ohr erreicht, auch tatsächlich an das Gehirn weitergeleitet wird. Die Aufgabe eines Modells des peripheren Gehörs besteht also in erster Linie darin, diejenigen Schallanteile zu extrahieren, die das periphere Gehör dem Gehirn zur Verfügung stellt (vgl. Kap. 4).

Unter dem Begriff des Erregungsmusters werden Frequenz- oder Zeit-Frequenzdarstellungen verstanden, die unter Berücksichtigung der Übertragungseigenschaften des menschlichen Gehörs berechnet wurden. Dabei handelt es sich in erster Linie um physiologisch motivierte Größen, wobei jedoch üblicherweise auch die Frequenz- oder Zeit-Frequenzdarstellungen basierend auf funktionalen Modellen als Erregungsmuster bezeichnet werden (vgl. Kap. 2.7.).

Modelle des peripheren Gehörs bzw. damit berechnete Erregungsmuster lassen sich für beinahe alle Anwendungen nutzen, in denen ein Zusammenhang zum menschlichen Hörvermögen vorhanden ist. Einige Anwendungsbeispiele sind:

- Akustischer Umweltschutz - Lärmmeßtechnik

In der Lärmmeßtechnik werden meist Lautheitsmeßverfahren verwendet. Das zur Zeit genaueste Lautheitsmeßverfahren (Verfahren nach Zwicker, vgl. Kap. 3.6.) beinhaltet eine explizite Berechnung des Erregungsmusters.

- Schallwirkungsforschung

Für die Analyse der unterschiedlichsten Wirkungen von Schall auf den Menschen werden häufig Erregungsmuster als Beschreibungen der Schallreize verwendet.

- Vorverarbeitungsstufe zur automatisierten Spracherkennung

Die automatisierte Spracherkennung ist insofern von Interesse, da sie eine wesentliche Voraussetzung für die akustische Kommunikation zwischen Mensch und Maschine ist. Ein Anwendungsgebiet für eine automatisierte Spracherkennung ist beispielsweise die Steuerung eines Rechners mit akustischen Befehlen.

- Qualitätsbeurteilung von Audiogeräten

Für die Herstellung von Audiogeräten ist eine automatisierte Qualitätskontrolle unter Berücksichtigung der Eigenschaften des menschlichen Gehörs von Interesse. Die Entwicklung von Audiogeräten wird mit Hilfe von Modellen des peripheren Gehörs

dahingehend optimiert, daß sich die Qualitäten der Audiogeräte in zunehmendem Maße an die Eigenschaften des Gehörs anlehnen. Es wäre unwirtschaftlich eine zu hohe Qualität zu produzieren d.h. Audiogeräte zu entwickeln, deren Übertragungsqualitäten weit über den Wahrnehmungsbereich des Menschen hinausgehen.¹

- Datenkompression in Anlehnung an die Eigenschaften des peripheren Gehörs
In der Audiotechnik finden in zunehmendem Maße Datenkompressionsverfahren Verwendung, die sich in ihrer Funktionsweise an den Übertragungseigenschaften des peripheren Gehörs orientieren. Ziel dieser Datenreduktionen ist die bessere Ausnutzung von Übertragungsbandbreiten und Speichermedien.²
- Optische Darstellung akustischer Signale
Optische Darstellungen des Erregungsmusters werden beispielsweise als Therapieunterstützung in der Logopädie (Sprachheilkunde) oder für die musikwissenschaftliche Analyse von Musikaufzeichnungen verwendet.
- Entwicklung von Hörhilfen
Für alle Arten von Hörhilfen (akustische Verstärker am Eingang des Hörkanals, Ersatz für die mechanische Übertragung des Mittelohres und Cochlea-Implantate) ist die Modellierung des peripheren Gehörs bzw. ein Verständnis über die Funktionsweise des peripheren Gehörs von wesentlicher Bedeutung. Bei Cochlea-Implantaten wird beispielsweise ein Schlauch mit z.B. 23 Elektroden in die Cochlea implantiert. Diese Elektroden stimulieren direkt die einzelnen Fasern des Hörnervs. Die akustische Information wird dafür über ein Mikrofon aufgenommen und das periphere Gehör wird mit einem digitalen Modell simuliert.³

Funktionale Modelle des peripheren Gehörs simulieren bestimmte Eingangs - Ausgangs Eigenschaften, ohne auf die physiologischen Vorgänge im Detail einzugehen. Im allgemeinen läßt sich ein funktionales Modell leichter an praktische Anwendungen adaptieren als ein physiologisches Modell. Nicht zuletzt, da sich funktionale Modelle meist aus konkreten Anwendungen ableiten (vgl. Kap. 4.2.2.).

Das Leistungsdichtespektrum-Modell wird als Ausgangspunkt zur funktionalen Nachbildung des peripheren Gehörs weitgehend akzeptiert. Dieses Modell basiert auf der Hypothese, daß sich das periphere Gehör mit einer parallelen Bank von überlappenden Bandpaßfiltern (auditive Filter) nachbilden läßt. Es wird zur Zeit in Richtung einer pegelabhängigen Nichtlinearität und der Berücksichtigung instationärer Schalle erweitert (vgl. Kap. 2.4.).

Die auditiven Filter repräsentieren im Idealfall sowohl die Frequenzselektivität als auch die

¹ [Beerends, Stemerding, 1992]

² [Zollner, Zwicker, 1993]

³ [Implants, 1997]

Zeitauflösung des peripheren Gehörs.

Eine Möglichkeit, den Amplitudengang von auditiven Filtern und damit die Frequenzselektivität des Gehörs zu bestimmen, ist die Notched-Noise-Methode (vgl. Kap. 2.5.1.). Die auf dieser Methode basierenden auditiven Filter sind für ein funktionales Parallel-Filterbankmodell besonders gut geeignet, weil

- der Effekt des 'off-frequency listenings' minimiert wird,
- sich die Frequenzselektivität getrennt von der Detektionseffizienz bestimmen läßt,
- Interaktionen zwischen Maskierer und Testsignal gering sind,
- die Filterformen über einen großen Frequenz- und Pegelbereich bestimmbar sind und
- aufgrund der Vielzahl an bereits durchgeführten Notched-Noise Experimenten eine gewisse Ausgereiftheit dieser Versuche gegeben ist (vgl. Kap. 2.5.1.).

Der Nachteil der Notched-Noise Methode besteht darin, daß zur Zeit nicht geklärt ist, ob innerhalb eines Notched-Noise Experiments die Leistungsdichte des Maskierungsrauschens oder der Testsignalpegel konstant gehalten werden sollten. Damit verbunden ist die Unklarheit darüber, ob die nichtlinearen Filter eingangspegelabhängig oder ausgangspegelabhängig gesteuert werden sollten. Ungeklärt ist weiters die Frage, ob Signale mehrerer Frequenzkanäle zur Bildung eines Steuerungspegels zusammenzufassen sind (vgl. Kap. 2.5.1.4.). Dieser Diskrepanz wird in dieser Arbeit insofern Rechnung getragen, als daß sowohl eingangspegelabhängige als auch ausgangspegelabhängige auditive Filter nachgebildet werden (vgl. Kap. 6.2.1. und 6.2.2.).

Das entwickelte nichtlineare Modell des peripheren Gehörs (vgl. Kap. 6.) teilt sich in einen linearen Außen- und Mittelohrfilter und in eine nichtlineare Bandpaßfilterbank als funktionales Modell der Cochlea. Der Amplitudengang des Außen- und Mittelohrfilters wird aus den Kurven gleicher Lautstärkepegel abgeleitet (vgl. Kap. 3.2.). Da an einer Revision dieser Kurven, besonders im Bereich unter 1kHz, zur Zeit gearbeitet wird, ist dieser Filter parametrisiert, d.h. der Amplitudengang kann mit einem einzelnen Parameter leicht an neuere Datensätze angepaßt werden (vgl. Kap. 6.1.). Die nichtlineare Steuerung der Gammatone Filter (All-Pol oder One-Zero Gammatone Filter) ist in einer allgemeinen Form gegeben, um unterschiedliche Steuerpegelberechnungen verwenden zu können. Vermutlich wird zukünftig die Art der Steuerpegelberechnung mit unterschiedlichen Anwendungen differieren.

Ein wesentlicher Vorteil der hier vorgestellten nichtlinearen Parametrisierung von Gammatone Filtern besteht darin, daß die Filtermittenfrequenzen pegelunabhängig sind. Vergleichbare Modelle (vgl. Gammachirp Funktion, Kap. 5.2.) ändern die Filtermittenfrequenzen pegelabhängig. Berechnet man ein Erregungsmuster mit auditiven Filtern, deren Mittenfrequenzen sich pegelabhängig verschieben, so erscheint ein amplitudenmoduliertes Eingangssignal als frequenzmoduliertes Signal im Erregungsmuster. Diese Mehrdeutigkeit erschwert die Interpretation der Erregungsmuster. Es erscheint daher für Parallel-Filterbankmodelle nicht sinnvoll die Filtermittenfrequenzen pegelabhängig zu ändern, obwohl das als Bandpaßfilterbank interpretierte Gehör keine 'fixen' Filtermittenfrequenzen kennt und die Mittenfrequenzen der auditiven Filter den jeweiligen Eingangssignalen anpaßt.

Eine Filtermittenfrequenzverschiebung ließe sich mit der in dieser Arbeit behandelten Verschiebung der Polstellenpaare ebenfalls nachbilden (vgl. Kap. 6.2.).

Ein weiterer Vorteil der hier behandelten Parametrisierung ist die pegelunabhängige Filterordnung. Die Ordnung der Gammachirp Funktion beispielsweise ändert sich pegelabhängig (vgl. Kap. 5.2.). Dadurch läßt sich diese Funktion mit einer digitalen Filterstruktur nicht realisieren.

Eine mögliche Anwendung für die nichtlinearen Gammatone Filter ist die Berechnung der menschlichen Lautheitsempfindung. Lautheitsbewertungsverfahren sind insofern von besonderem Interesse, da die Lautheitsempfindung mit der Lärmempfindung eng korreliert und Lautheitsbewertungsverfahren aus diesem Grund in der Lärmmeßtechnik häufig verwendet werden.

Berechnung der menschlichen Lautheitsempfindung

Unter der Vielzahl an Modellen zur Berechnung der menschlichen Lautheitsempfindung findet der Aufbau des Modells nach Zwicker die breiteste Akzeptanz (vgl. Kap. 3.6.). Dieses Verfahren teilt sich in zwei Teile, in eine Berechnung des Erregungsmusters als erste Stufe und, darauf aufbauend, in die Berechnung der Lautheit. Für die Berechnung der Lautheit ist das Erregungsmuster offensichtlich von entscheidender Bedeutung. Man muß also in erster Linie ein Erregungsmuster finden, dessen Zeit- Frequenzeigenschaften möglichst genau dem menschlichen Gehör entsprechen. Erst darauf aufbauend erscheint eine valide Lautheitsberechnung möglich.

Im Unterschied zu einer möglichst exakten Nachbildung der Lautheitsempfindung sind bei der Verwendung eines Lautheitsbewertungsverfahrens in der Lärmmeßtechnik andere Randbedingungen gegeben. Ein für die Lärmmeßtechnik adaptiertes Lautheitsbewertungsverfahren muß drei Kriterien erfüllen. Es muß ein Maß für die Störwirkung und die Schädigungswirkung sein und es müssen sich Immissionen aus Emissionen der Quellen berechnen lassen. Letzteres ist nach wie vor ein häufiges Argument für die bewerteten Schallpegel, da sich deren energetische Addition in der Praxis bewährt hat. Eine ähnlich einfache Addition von unterschiedlichen Lärmemittenten existiert für das phon-Maß und das sone-Maß zur Zeit nicht. Diese Kritik ist meßtechnisch motiviert und hat mit der Validität eines Lautheitsbewertungsverfahrens selbst nichts zu tun (vgl. Kap. 3.7.).

Die in Kap. 3.5.2. behandelte dynamische Phonkurvenanpassung dB(ELC) stellt einen praxisorientierten Ansatz dar, den momentanen Standard der Lärmmeßtechnik, das dB(A), zu erweitern. Im Unterschied zum dB(A) berücksichtigt das dB(ELC) das gesamte Hörfeld. Anhand von 2 praktischen Analysen (Kap. 3.5.2.2.) wird gezeigt, daß sich bei gleichen dB(A)-Werten dB(ELC)-Differenzen von bis zu 20 dB(ELC) ergeben können. Trotz dieser großen Unterschiede kann aus den durchgeführten Analysen nicht eindeutig darauf geschlossen werden, daß diese

Verbesserungen eine Ablöse des dB(A) rechtfertigen würden.

Die Kurven gleicher Lautstärkepegel zählen zu den bekanntesten und am weitesten verbreiteten psychoakustischen Daten. Sowohl für die Modellierung des Außen- und Mittelohrfilters als auch für die Modellierung der Lautheitsempfindung sind diese Kurven von wesentlicher Bedeutung. Aus der in Kap. 3.2. angeführten Unsicherheit bei der Bestimmung der Kurven gleicher Lautstärkepegel (geringfügige Änderungen der Versuchsdurchführung führen zu stark unterschiedlichen Ergebnissen) wird die Problematik der Modellierungsversuche menschlicher Empfindungen deutlich.

Anhang

Implementation der dB(ELC) Berechnung Verwendete Filter zur dB(ELC) Berechnung

Implementation der dB(ELC) Berechnung

Die Software zur Berechnung des dB(ELC) ist unter MatLab für Windows implementiert. Das Audiosignal¹ $x(n)$ wird blockweise bearbeitet (Abb. A.1.). Die Blocklänge N und damit die Zeitauflösung der Berechnung ist variabel². Um eine kontinuierliche Filterung des Eingangssignals zu gewährleisten, werden die internen Zustände der Filter beim Übergang zum nächst folgenden Block übernommen. Alle verwendeten Filter sind rekursiv und in der transponierten Direktform II ausgeführt. Die angegebenen Zeitverzögerungen Δt_1 bis Δt_3 kompensieren die durch die einzelnen Verarbeitungsstufen verursachten Zeitverschiebungen.

Die Übergänge von den Zeitsignalen $s(n)$ ($x(n)$, $a(n)$, $b(n)$, $c(n)$, $d(n)$, $e(n)$) zu den entsprechenden Pegelwerten werden von folgenden Gliedern durchgeführt:

$$dB(\dots) = 10 \log \left(\frac{\frac{1}{N} \sum_{n=1}^N s^2(n)}{RF} \right) \quad \text{Gl. A.1.}$$

Mit dem Referenzfaktor RF werden die Zeitfunktionen $s(n)$ kalibriert. Durch Änderung dieses Faktors besteht die Möglichkeit das Eingangssignal unterschiedlichen Lautstärkepegelbereichen zuzuordnen.

¹ Mono, Windows PCM (*.wav), $f_s = 44,1$ kHz

² Übliche Zeitkonstanten in der Lärmmeßtechnik sind: 'Slow': 1 s, 'Fast': 125 ms, 'Impulse': Einschwingvorgang 35 ms, Ausschwingvorgang 1,5 s.

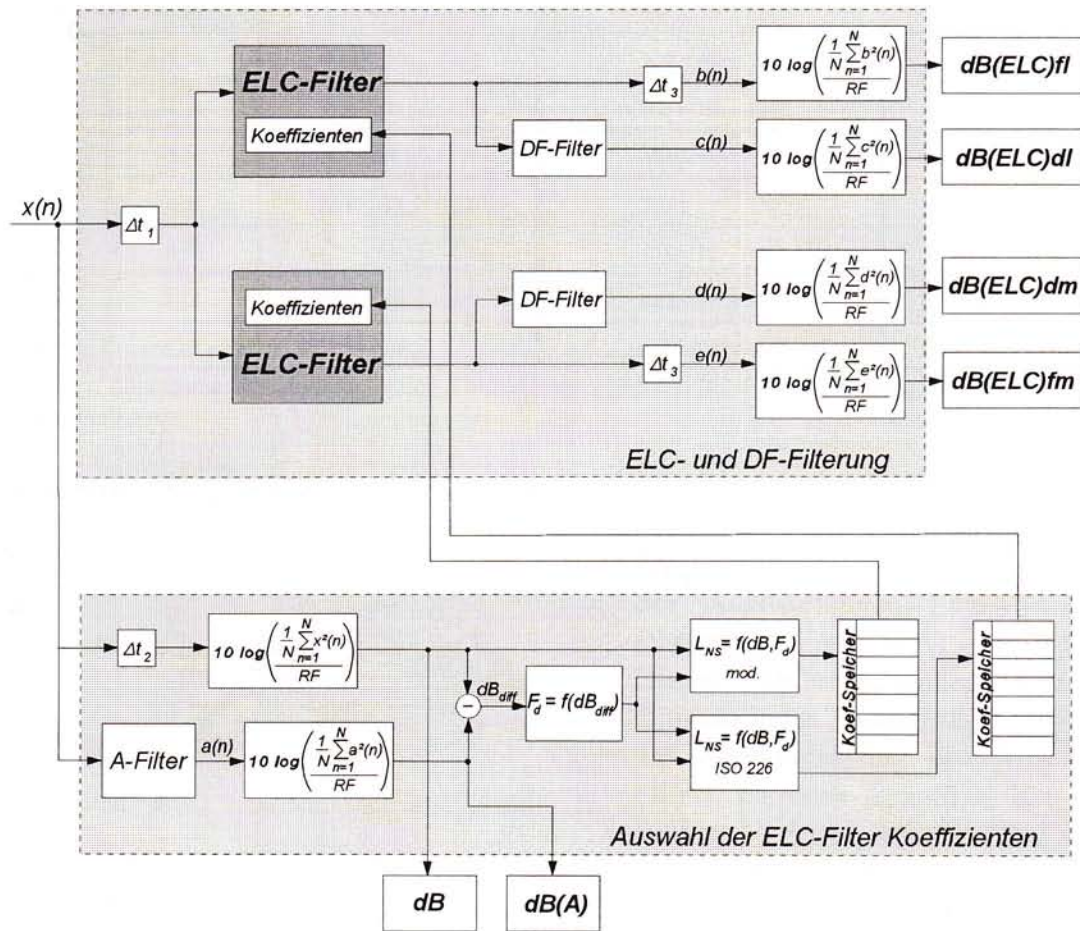


Abb. A.1.: Berechnung des dB(ELC)

Berechnung der dominierenden Frequenzkomponenten-Schätzwerte:

Der Zusammenhang zwischen der Differenz aus unbewertetem und A-bewertetem Schallpegel (dB_{diff}) und der dominierenden Frequenzkomponente F_d (vgl. Kap. 3.5.2.) ist in der Tabelle A.1. angeführt:

F_d (Hz)	dB_{diff} an F_d	Zugeordneter dB-Bereich	Zugeordneter Frequenzbereich (FB)	Zugeordneter FB über 2,5 kHz
26	-43,89	$dB_{diff} \leq -41,37$	$FB \leq 29$	
34	-37,88	$-41,37 < dB_{diff} \leq -35,57$	$29 < FB \leq 38$	
44	-32,67	$-35,57 < dB_{diff} \leq -30,28$	$38 < FB \leq 50$	
57	-27,93	$-30,28 < dB_{diff} \leq -25,7$	$50 < FB \leq 65$	
74	-23,6	$-25,7 < dB_{diff} \leq -21,66$	$65 < FB \leq 84$	
96	-19,72	$-21,66 < dB_{diff} \leq -17,97$	$84 < FB \leq 109$	
124	-16,29	$-17,97 < dB_{diff} \leq -14,71$	$109 < FB \leq 141$	

161	-13,17	-14,71 < dB _{Diff} ≤ -11,72	141 < FB ≤ 184	
209	-10,4	-11,72 < dB _{Diff} ≤ -9,14	184 < FB ≤ 238	19684 ≤ FB
271	-7,94	-9,14 < dB _{Diff} ≤ -6,8	238 < FB ≤ 309	16146 ≤ FB < 19684
352	-5,74	-6,8 < dB _{Diff} ≤ -4,7	309 < FB ≤ 401	13136 ≤ FB < 16146
458	-3,82	-4,7 < dB _{Diff} ≤ -2,99	401 < FB ≤ 521	10719 ≤ FB < 13136
593	-2,23	-2,99 < dB _{Diff} ≤ -1,55	521 < FB ≤ 676	8620 ≤ FB < 10719
770	-0,95	-1,55 < dB _{Diff} ≤ -0,39	676 < FB ≤ 887	6794 ≤ FB < 8620
1055	0,16	-0,39 < dB _{Diff} ≤ 0,71	887 < FB ≤ 1339	4643 ≤ FB < 6794
2513	1,27	0,71 < dB _{Diff}	1339 < FB < 4643	

Tab. A.1.: Zuordnung der dominierenden Frequenzkomponenten zu den entsprechenden Pegeldifferenzen und Frequenzbereichen

Der hörbare Frequenzbereich wird entsprechend Tab. A.1. in 23 Frequenz-bänder (FB)

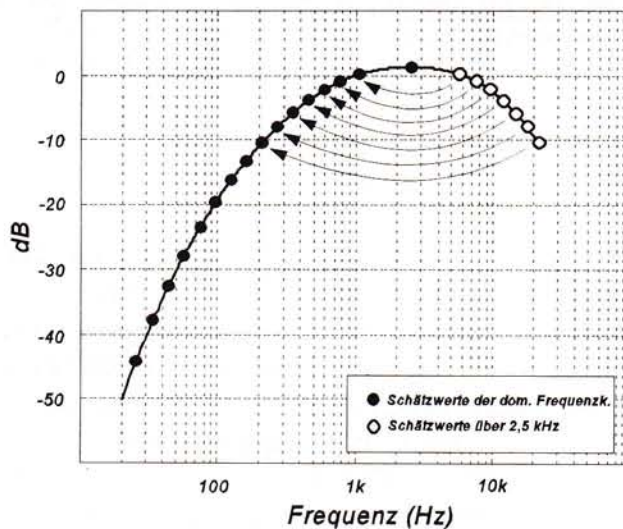


Abb. A.2.: Schätzwerte der dominierenden Frequenzkomponenten und Amplitudengang des A-bewertenden Filters

eingeteilt. Aus den Grenzfrequenzen dieser Frequenzbänder und dem Amplitudengang des A-bewertenden Filters werden 16 Differenzbänder (dB(A) - dB) bestimmt. Diese Differenzbänder werden 16 dominierenden Frequenzen zugeordnet. Da der Amplitudengang des A-bewertenden Filters ab etwa 2,5 kHz eine negative Steigung aufweist, können den ursprünglich 23 Frequenzbändern nur 16 Differenzbänder zugeordnet werden. Alle Frequenzen über 2,5 kHz werden auf den Bereich unter 2,5 kHz abgebildet.

In Abb. A.2. sind die 16 Schätzwerte der dominierenden Frequenzkomponenten unter 2,5 kHz (•) und über 2,5 kHz (o) sowie die Übertragungsfunktion des A-bewertenden Filters dargestellt.

Berechnung des Schallpegelschätzwertes:

Die Berechnung des Schallpegelschätzwertes aus der dominierenden Frequenzkomponente und dem unbewerteten Schallpegel ($L_{NS} = f(dB, F_d)$) ist in der Tabelle A.2. angeführt:

F _d (Hz)	40 Phon ISO 226	60 Phon ISO 226	80 Phon ISO 226	100 Phon ISO 226	40 Phon mod.	60 Phon mod.	80 Phon mod.	100 Phon mod.
26	dB ≤ 86	86 < dB ≤ 99	99 < dB ≤ 113	113 < dB	dB ≤ 91	91 < dB ≤ 104	104 < dB ≤ 121	121 < dB
34	dB ≤ 80	80 < dB ≤ 94	94 < dB ≤ 108	108 < dB	dB ≤ 86	86 < dB ≤ 100	100 < dB ≤ 117	117 < dB
44	dB ≤ 73	73 < dB ≤ 87	87 < dB ≤ 103	103 < dB	dB ≤ 82	82 < dB ≤ 96	96 < dB ≤ 113	113 < dB

57	$\text{dB} \leq 67$	$67 < \text{dB} \leq 83$	$83 < \text{dB} \leq 99$	$99 < \text{dB}$	$\text{dB} \leq 78$	$78 < \text{dB} \leq 91$	$91 < \text{dB} \leq 108$	$108 < \text{dB}$
74	$\text{dB} \leq 63$	$63 < \text{dB} \leq 79$	$79 < \text{dB} \leq 96$	$96 < \text{dB}$	$\text{dB} \leq 73$	$73 < \text{dB} \leq 87$	$87 < \text{dB} \leq 104$	$104 < \text{dB}$
96	$\text{dB} \leq 59$	$59 < \text{dB} \leq 76$	$76 < \text{dB} \leq 94$	$94 < \text{dB}$	$\text{dB} \leq 69$	$69 < \text{dB} \leq 84$	$84 < \text{dB} \leq 101$	$101 < \text{dB}$
124	$\text{dB} \leq 55$	$55 < \text{dB} \leq 73$	$73 < \text{dB} \leq 92$	$92 < \text{dB}$	$\text{dB} \leq 65$	$65 < \text{dB} \leq 81$	$81 < \text{dB} \leq 97$	$97 < \text{dB}$
161	$\text{dB} \leq 53$	$53 < \text{dB} \leq 71$	$71 < \text{dB} \leq 90$	$90 < \text{dB}$	$\text{dB} \leq 62$	$62 < \text{dB} \leq 78$	$78 < \text{dB} \leq 95$	$95 < \text{dB}$
209	$\text{dB} \leq 50$	$50 < \text{dB} \leq 69$	$69 < \text{dB} \leq 88$	$88 < \text{dB}$	$\text{dB} \leq 58$	$58 < \text{dB} \leq 76$	$76 < \text{dB} \leq 93$	$93 < \text{dB}$
271	$\text{dB} \leq 49$	$49 < \text{dB} \leq 67$	$67 < \text{dB} \leq 87$	$87 < \text{dB}$	$\text{dB} \leq 56$	$56 < \text{dB} \leq 74$	$74 < \text{dB} \leq 91$	$91 < \text{dB}$
352	$\text{dB} \leq 47$	$47 < \text{dB} \leq 66$	$66 < \text{dB} \leq 86$	$86 < \text{dB}$	$\text{dB} \leq 54$	$54 < \text{dB} \leq 73$	$73 < \text{dB} \leq 91$	$91 < \text{dB}$
458	$\text{dB} \leq 47$	$47 < \text{dB} \leq 66$	$66 < \text{dB} \leq 87$	$87 < \text{dB}$	$\text{dB} \leq 52$	$52 < \text{dB} \leq 71$	$71 < \text{dB} \leq 90$	$90 < \text{dB}$
593	$\text{dB} \leq 47$	$47 < \text{dB} \leq 67$	$67 < \text{dB} \leq 88$	$88 < \text{dB}$	$\text{dB} \leq 51$	$51 < \text{dB} \leq 71$	$71 < \text{dB} \leq 90$	$90 < \text{dB}$
770	$\text{dB} \leq 48$	$48 < \text{dB} \leq 69$	$69 < \text{dB} \leq 89$	$89 < \text{dB}$	$\text{dB} \leq 50$	$50 < \text{dB} \leq 71$	$71 < \text{dB} \leq 90$	$90 < \text{dB}$
1055	$\text{dB} \leq 50$	$50 < \text{dB} \leq 70$	$70 < \text{dB} \leq 91$	$91 < \text{dB}$	$\text{dB} \leq 50$	$50 < \text{dB} \leq 70$	$70 < \text{dB} \leq 90$	$90 < \text{dB}$
2513	$\text{dB} \leq 46$	$46 < \text{dB} \leq 65$	$65 < \text{dB} \leq 84$	$84 < \text{dB}$	$\text{dB} \leq 47$	$47 < \text{dB} \leq 66$	$66 < \text{dB} \leq 84$	$84 < \text{dB}$

Tab. A.2.: Bestimmung des Lautstärkepegel-Schätzwertes aus der dominierenden Frequenzkomponente F_d und dem unbewerteten Pegel

Verwendete Filter zur dB(ELC) Berechnung

Mit Ausnahme des dB(A) Filters wurden alle verwendeten Filter zur Berechnung des dB(ELC) mit einer Methode entworfen, die für die geforderten Amplitudengänge minimale Filterordnungen ermöglicht. Diese Methode ist empirisch und unterscheidet sich in ihrer Logik grundsätzlich von vorhandenen signalverarbeitungstheoretischen Ansätzen. Der geforderte sowie der tatsächlich erreichte Amplitudengang werden in der gleichen Grafik dargestellt nach optischen Kriterien angenähert. Dabei wird unter Veränderung der Pol- und Nullstellenpositionen die Anzahl der Pole und Nullstellen solange erhöht, bis die Amplitudengänge einander entsprechen. Mit dieser Methode können Filter entwickelt werden, deren Ordnung etwa um den Faktor 5 geringer ist als die vergleichbarer Filter die mit dem Yule-Walker¹ Approximationsalgorithmus berechnet wurden. In Tab. A.4. sind die entworfenen Filter dargestellt.

Abkürzungen:

A	A-bewertender Filter	100	inverser 100-phon Filter nach ISO 226
FD	Diffusfeld-Filter	40m	modifizierter inverser 40-phon Filter nach ISO 226
40	inverser 40-phon Filter nach ISO 226	60m	modifizierter inverser 60-phon Filter nach ISO 226
60	inverser 60-phon Filter nach ISO 226	80m	modifizierter inverser 80-phon Filter nach ISO 226
80	inverser 80-phon Filter nach ISO 226	100m	modifizierter inverser 100-phon Filter nach ISO 226

Tab. A.3.: Abkürzungen der Filter

¹ [Kammeyer, Kroschel, 1992]

Pole, Nullstellen und Verstärkungsfaktoren der Filter:

A	k = 0,255	N	4x1 ¹		2x-1				
		P	0,9971	0,9848	0,9001	2x0,07			
FD	k = 0,097	N	-12		$0,847e^{\pm j0,08}$	$0,36e^{\pm j2}$	$0,52e^{\pm j0,41}$		
		P	$0,932e^{\pm j0,12}$		$0,58e^{\pm j1,2}$				
40	k = 0,2	N	3x1		$0,9e^{\pm j0,14}$	$0,8e^{\pm j0,16}$	$0,75e^{\pm j1,13}$	$e^{\pm j2,8}$	-1
		P	0,84		$0,945e^{\pm j0,05}$	$0,9905e^{\pm j0,006}$	$0,7e^{\pm j0,3}$	$0,75e^{\pm j0,55}$	$0,8e^{\pm j1,9}$
60	k = 0,2	N	3x1		$0,9e^{\pm j0,14}$	$0,77e^{\pm j0,16}$	$0,75e^{\pm j1,1}$	$e^{\pm j3}$	-1
		P	0,85		$0,947e^{\pm j0,037}$	$0,9925e^{\pm j0,0034}$	$0,7e^{\pm j0,3}$	$0,78e^{\pm j0,59}$	$0,87e^{\pm j1,9}$
80	k = 0,2	N	3x1		$0,92e^{\pm j0,14}$	$0,77e^{\pm j0,16}$	$0,82e^{\pm j1,1}$	$1,35e^{\pm j3,1}$	-1
		P	0,85		$0,942e^{\pm j0,023}$	$0,995e^{\pm j0,0036}$	$0,67e^{\pm j0,32}$	$0,79e^{\pm j0,59}$	$0,82e^{\pm j1,95}$
100	k = 0,2	N	2x1	1,007	$0,91e^{\pm j0,1}$	$0,86e^{\pm j0,21}$	$0,78e^{\pm j1,1}$	$e^{\pm j3,1}$	-1
		P	0,85		$0,955e^{\pm j0,015}$	$0,9935e^{\pm j0,004}$	$0,82e^{\pm j0,32}$	$0,82e^{\pm j0,59}$	$0,7e^{\pm j1,95}$
40m	k = 0,2	N	2x1	1,01	$0,878e^{\pm j0,14}$	$0,68e^{\pm j0,3}$	$0,86e^{\pm j1,13}$	$e^{\pm j2,8}$	-1
		P	0,7		$0,91e^{\pm j0,05}$	$0,98e^{\pm j0,006}$	$0,55e^{\pm j0,39}$	$0,58e^{\pm j0,55}$	$0,5e^{\pm j1,9}$
60m	k = 0,2	N	2x1	1,03	$0,89e^{\pm j0,14}$	$0,68e^{\pm j0,3}$	$0,86e^{\pm j1,13}$	$e^{\pm j2,8}$	-1
		P	0,7		$0,905e^{\pm j0,05}$	$0,98e^{\pm j0,006}$	$0,6e^{\pm j0,4}$	$0,58e^{\pm j0,55}$	$0,5e^{\pm j1,9}$
80m	k = 0,2	N	2x1	1,09	$0,889e^{\pm j0,14}$	$0,68e^{\pm j0,3}$	$0,82e^{\pm j1,13}$	$e^{\pm j2,8}$	-1
		P	0,7		$0,878e^{\pm j0,05}$	$0,978e^{\pm j0,006}$	$0,62e^{\pm j0,4}$	$0,7e^{\pm j0,55}$	$0,65e^{\pm j1,9}$
100m	k = 0,2	N	2x1	1,15	$0,9e^{\pm j0,14}$	$0,68e^{\pm j0,3}$	$0,82e^{\pm j1,13}$	$e^{\pm j2,8}$	-1
		P	0,7		$0,85e^{\pm j0,05}$	$0,98e^{\pm j0,006}$	$0,65e^{\pm j0,4}$	$0,73e^{\pm j0,55}$	$0,65e^{\pm j1,9}$

Tab. A.4.: Verstärkungsfaktoren k, Pole P und Nullstellen N der verwendeten Filter

Amplitudengänge:

Die Amplitudengänge des Diffusfeld- und A-bewertenden Filters sind in den Abb. 3.7. und Abb. 3.10. (Kap. 3.5.2.) dargestellt. Die Amplitudengänge der ELC Filter sind in den beiden Abb. A.3. und A.4. dargestellt.

¹ Der Ausdruck 4x1 bedeutet beispielsweise 4 Nullstellen an der Stelle $z = 1$

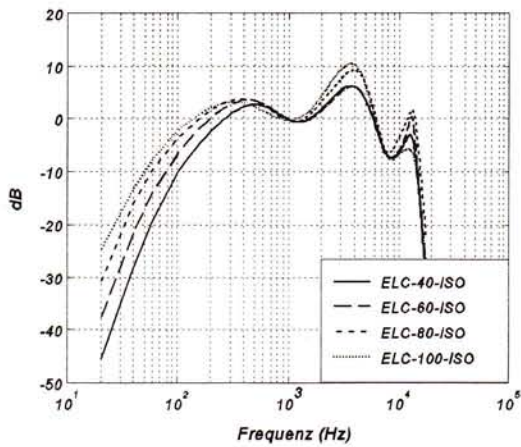


Abb. A.3.: Amplitudengänge der ELC-ISO-Filter

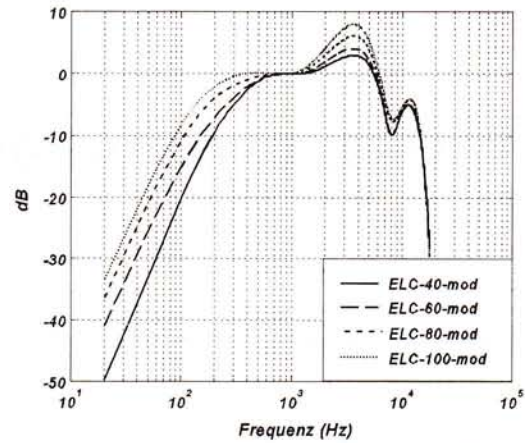


Abb. A.4.: Amplitudengänge der ELC-mod.-Filter

Impulsantworten:

Die in Abb. A.5. dargestellten Impulsantworten der Filter bewegen sich im Bereich von etwa 0,5 - 2 ms. Mögliche Transiente beim Umschalten der Filterkoeffizienten sind dadurch gering. Dies ist vor allem dann von Bedeutung, wenn ein Block mit kleiner Leistung auf einen mit großer Leistung folgt. Für die übliche Zeitauflösung von 125 ms ('fast') bleibt das Übersprechen beispielsweise auf etwa 2% der Blocklänge beschränkt.

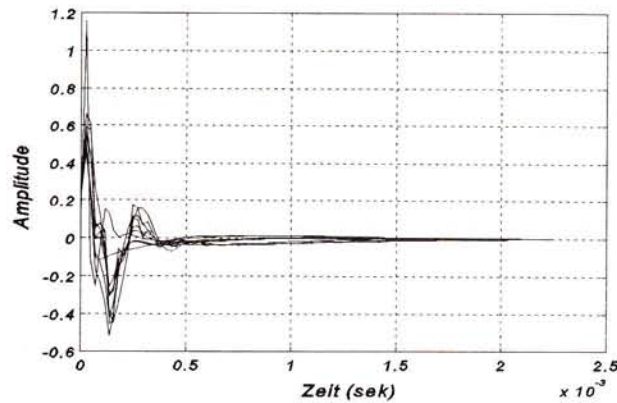


Abb. A.5.: Impulsantworten der verwendeten Filter

Literaturverzeichnis

[Ahn, Westerkamp, 1990]

Ahn S.-G., Westerkamp J.J.: 'Cochlear Modeling Using a General Purpose Digital Signal Processor', IEEE CH2881-1/90/0000-0057, 1990, 57-63

[Agerkvist, 1994]

Agerkvist F.T.: 'Time-Frequency Analysis and Auditory Modells', Dissertation, Technische Universität Dänemark, 1994

[Agerkvist, 1996]

Agerkvist F.T.: 'A Time-Frequency Auditory Model Using Wavelet Packets', J. Audio Eng. Soc. 44 (1/2), 1996, 37-50

[Beerends, Stemerding, 1992]

Beerends J.G., Stemerding J.A.: 'A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation', J. Audio Eng. Soc., Vol. 40, No. 12, 1992, 963-978

[Berglund et al., 1996]

Berglund B., Hassmen P., Soames Job R.F.: 'Sources and effects of low-frequency noise', J. Acoust. Soc. Am. 99(5), 1996, 2985-3002

[Betke, 1989]

Betke K.: 'Kurven gleicher Pegellautstärke und Hörschwelle im ebenen Schallfeld', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 89, DPG-Verlag, Bad Honnef, 1989

[Blauert, 1974]

Blauert J.: 'Räumliches Hören', S. Hirzel Verlag Stuttgart, 1974

[Bollerott et al., 1996]

Bollerott M., Despang H.G., Kluge W., Schwarz A.: 'Softwaremodell der natürlichen Cochlea', Acustica-acta acustica, Vol. 82, 1996, 102-113

[Carney, 1993]

Carney L.H.: 'A model for the responses of low-frequency auditory-nerve fibres in cat', J. Acoust. Soc. Am. 93, Jänner 1993, 401-417

[Colomes et al., 1995]

Colomes C., Lever M., Rault J.B., Dehery Y.F.: 'A Perceptual Model Applied to Audio Bit-Rate Reduction', J. Audio Eng. Soc., Vol. 43, No. 4, 1995, 233-239

[Cooke, 1991]

Cooke M.: 'Modelling Auditory Processing and Organisation', Dissertation, University Sheffield, 1991, Cambridge University Press, 1993

[Cooke et al., 1993]

Cooke M., Beet S., Crawford M. (Eds.): 'Visual Representations of Speech Signals', John Wiley & Sons Ltd., West Sussex, England, 1993

[Cooper, 1996]

Cooper N.P.: 'Two-tone suppression in cochlear mechanics', J. Acoust. Soc. Am. 99(5), 1996, 3087-3098

[Dau et al., 1996]

Dau T., Püschel D., Kohlrausch A.: 'A quantitative model of the 'effective' signal processing in the auditory system. I. Model structure', J. Acoust. Soc. Am. 99(6), 1996, 3615-3631

[de Boer, 1975]

de Boer E.: 'Synthetic whole-nerve action potentials for the cat', J. Acoust. Soc. Am. 58, 1030-1045

[de Boer, de Jongh, 1978]

de Boer E., de Jongh H.R.: 'On cochlea encoding: Potentials and limitations of the reverse-correlation technique', J. Acoust. Soc. Am. 63, 115-135

[Deng, Geisler, 1987]

Deng L., Geisler C.D.: 'A composite auditory model for processing speech sounds', J. Acoust. Soc. Am. 82(6), 1987, 2001-2012

[DIGIFFFECTS]

Geräuscharchiv 'DIGIFFFECTS', Sonoton - recorded music library, München

[DIN 45 630]

Deutsche Normen, 'Grundlagen der Schallmessung: Normalkurven gleicher Lautstärkepegel', DIN 45630, Blatt 2, 1967

[DIN 45 631]

Deutsche Normen, 'Berechnung des Lautstärkepegels aus dem Geräuschspektrum: Verfahren nach E. Zwicker', DIN 45631, 1967

[DIN 45619]

Deutsche Normen, 'Kopfhörer - Bestimmung des Freifeld-Übertragungsmaßes durch Lautstärkevergleich mit einer fortschreitenden Welle', DIN 45619, Teil 1, 1975

[Dubno, Dirks, 1989]

Dubno J.R., Dirks D.D.: 'Auditory filter characteristics and consonant recognition for hearing-impaired listeners', J. Acoust. Soc. Am. 85, 1989, 1666-1675

[Eckel, 1989]

Eckel G.: 'Ein Modell der Mehrfachverdeckung für die Analyse musikalischer Schallsignale', Dissertation, Universität Wien, 1989

[Fastl, Zwicker, 1983]

Fastl H., Zwicker E.: 'A free-field equalizer for THD 39 earphones', J. Acoust. Soc. Am. 73, 1983, 312-314

[Fastl, 1988]

Fastl H.: 'Gehörbezogene Lärmessverfahren', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 88, 111-124

[Fastl et al., 1990]

Fastl H., Jaroszewski A., Schorer E., Zwicker E.: 'Equal Loudness Contours between 100 and 1000 Hz for 30, 50, and 70 phon', *Acustica* Vol. 70, 1990, 197-201

[Fastl, 1991]

Fastl H.: 'Beurteilung und Messung der wahrgenommenen äquivalenten Dauerlautheit', *Zeitschrift für Lärmbekämpfung* 38, 1991, 98-103

[Gabriel et al., 1993]

Gabriel B., Kollmeier B., Mellert V.: 'Influence of various measurement procedures on the equal-loudness level contours, In Sixth Oldenburg Symposium on Psychological Acoustics, 1993, 223-231

[Gabriel et al., 1994]

Gabriel B., Kollmeier B., Mellert V.: 'Einfluß verschiedener Meßmethoden auf die Kurven gleicher Pegellautstärke', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 94, Technische Universität Dresden, 1994, 1085-1088

[Gabriel et al., 1995]

Gabriel B., Kollmeier B., Mellert V.: 'Kontexteffekte bei der Bestimmung der Kurven gleicher Pegellautstärke', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 95, Fraunhofer-Institut für zerstörungsfreie Prüfverfahren, Universität Saarbrücken, 1995, 899-902

[Geisler, Cai, 1996]

Geisler C.D., Cai Y.: 'Relationships between frequency-tuning and spatial-tuning curves in the mammalian cochlea', *J. Acoust. Soc. Am.* 99(3), 1996, 1550-1555

[Genuit, 1991]

Genuit K.: 'Gehörgerechte Lärmbewertung', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 91, 75-91

[Giguère, Woodland, 1994a]

Giguère C., Woodland P.C.: 'A computational model of the auditory periphery for speech and hearing research. I. Ascending path', *J. Acoust. Soc. Am.* 95, 331-342

[Giguère, Woodland, 1994b]

Giguère C., Woodland P.C.: 'A computational model of the auditory periphery for speech and hearing research. II. Descending paths', *J. Acoust. Soc. Am.* 95, 343-349

[Glasberg et al., 1984]

Glasberg B.R., Moore B.C.J., Patterson R. D., Nimmo-Smith I.: 'Dynamic range and asymmetry of the auditory filter', *J. Acoust. Soc. Am.* 76, 1984, 419-427

[Glasberg, Moore, 1986]

Glasberg B.R., Moore B.C.J.: 'Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments', *J. Acoust. Soc. Am.* 79(4), 1986, 1020-1033

[Glasberg, Moore, 1990]

Glasberg B.R., Moore B.C.J.: 'Derivation of auditory filter shapes from notched-noise data', *Hear. Res.* 47, 1990, 103-138

[Glasberg, Moore, 1994]

Glasberg B.R., Moore B.C.J.: 'Growth-of-masking functions for several types of maskers', *J. Acoust. Soc. Am.* 96, 1994, 134-144

[Glasberg, Moore, 1996]

Glasberg B.R., Moore B.C.J.: 'A Revision of Zwicker's Loudness Model', *Acustica united with acta acoustica*, 1996, 335-345

[Green et al., 1987]

Green D.M., Kidd G., Stevens K.N.: 'High frequency audiometric assessment of a young adult population', J. Acoust. Soc. Am. 81, 1987, 485-494

[Greenwood, 1972]

Greenwood, D.D.: 'Masking by combination bands : Estimation of the levels of the combination bands $(n+1)f_1 - nf_n$ ', J. Acoust. Soc. Am. 52, 1972, 1144-1154

[Greenwood, 1990]

Greenwood D.D.: 'A cochlea frequency-position function for several species 29 years later', J. Acoust. Soc. Am. 87 (6), 1990, 2592-2605

[Grigo, 1994]

Grigo R.: 'dB(phon) und das Lautheitsverfahren nach Zwicker', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 94, Technische Universität Dresden, 1994, 1093-1096

[Guski, Bosshardt, 1992]

Guski R., Bosshardt H.-G.: 'Gibt es eine 'unbeeinflusste' Lästigkeit?', Zeitschrift für Lärmbekämpfung 39, 1992, 67-74

[Hall, 1980]

Hall J.L.: 'Cochlea models: Evidence in support of mechanical nonlinearity and second filter (a review)', Hearing Research 2, 1980, 455-464

[Hawkins et al., 1995]

Hawkins H.L., McMullen T.A., Popper A.N., Fay R.R. (Eds.): 'Auditory Computation', Springer Handbook of Auditory Research, Series Editors: Fay R.R., Popper A.N., Springer, 1995

[Heijden, Kohlrausch, 1995]

van der Heijden M., Kohlrausch A.: 'The role of envelope fluctuations in spectral masking', J. Acoust. Soc. Am. 97(3) March 1995, 1800-1807

[Houtgast, 1977]

Houtgast T.: 'Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker', J. Acoust. Soc. Am. 62, 1977, 409-415

[Implants, 1997]

Cochlear Implants, Internet www-Seite: http://weber.u.washington.edu/~otoweb/cochlear_implants.html, Department of Otolaryngology, University of Washington, Seattle, Washington, 1997

[Irino, 1995]

Irino T.: 'An optimal auditory Filter', IEEE Workshop on Applications of Signal Processing to Audio and Acoustics', Oktober 1995

[Irino, 1996]

Irino T.: 'A Gammachirp Function as an Optimal Auditory Filter with the Mellin Transform', IEEE, ICASSP96, Mai 1996, Atlanta

[Irino, Patterson, 1996]

Irino T., Patterson R.D.: 'Temporal asymmetry in the auditory system', J. Acoust. Soc. Am. 99(4), 1996, 2316-2331

[ISO 226]

'Acoustics - Normal Equal-Loudness Level Contours', ISO 226 (E), 1987, 20-27

[ISO 532]

'Acoustics - Method for calculating loudness level', ISO 532 (E), 1975, 39-55

[Jahn, Santos-Sacchi, 1988]

Jahn A.F., Santos-Sacchi J.: 'Physiology of the Ear', Raven Press Ltd., New York, 1988

[Johnson-Davies, Patterson, 1979]

Johnson-Davies, D., and Patterson, R. D. (1979). "Psychophysical tuning curves: restricting the listening band to the signal region," J. Acoust. Soc. Am. 65, 765-770

[Kammeyer, Kroschel, 1992]

Kammeyer K., Kroschel K.: 'Digitale Signalverarbeitung', 2. Auflage, B.G. Teubner Stuttgart, 1992

[Kates, 1993]

Kates J.M.: 'Accurate Tuning Curves in a Cochlear Model', IEEE Trans. on Speech and Audio Processing, Vol. 1, No. 4, 1993, 453-462

[Kates, 1995]

Kates J.M.: 'Two-Tone Suppression in a Cochlear Model', IEEE Trans. on Speech and Audio Processing, Vol. 3, No. 5, 1995, 396-406

[Keidel, 1975]

Keidel W.D.: 'Physiologie des Gehörs', Georg Thieme Verlag Stuttgart, 1975

[Killion, 1978]

Killion M.C.: 'Revised estimate of minimum audible pressure. Where is the 'missing 6 dB'? ', J. Acoust. Soc. Am. 63, 1978, 1501-1508

[Lazarus, 1990]

Lazarus H.: 'Läßt sich das Lautheitsbewertungsverfahren nach Zwicker für die praktische Lärmbekämpfung nutzen?', Zeitschrift für Lärmbekämpfung 37, 1990, 54-55

[Leek, Summers, 1993]

Leek M.R., Summers V.: 'Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise', J. Acoust. Soc. Am. 94(6), 1993, 3127-3137

[Lufti, Patterson, 1982]

Lufti R. A., Patterson R. D.: 'Combination bands and the measurement of auditory filter', J. Acoust. Soc. Am. 71, 1982, 421-423

[Lufti, 1983]

Lufti R. A.: 'Simultaneous masking and unmasking with bandlimited noise', J. Acoust. Soc. Am. 73, 1983, 899-905

[Lufti, Patterson, 1984]

Lufti R.A., Patterson R.D.: 'On the growth of masking asymmetry with stimulus intensity', J. Acoust. Soc. Am. 76(3), 1984, 739-745

[Lyon, 1982]

Lyon R.F.: 'A computational model of filtering, detection, and compression in the cochlea', Proc. ICASSP 1982, Paris, 1282-1285

[Lyon, Mead, 1988a]

Lyon R.F., Mead C.A.: 'Cochlear hydrodynamics demystified', California Inst. Tech. Comp. Sci. Tech. Rep. Caltech-CS-TR-88-4, 1988

[Lyon, Mead, 1988b]

Lyon R.F., Mead C.A.: 'An analog electronic cochlea', IEEE Trans. Acoust., Speech, and Signal Processing, vol. 36, no. 7, pp. 1119-1134, 1988

[Lyon, 1990]

Lyon R.F.: 'Automatic gain control in cochlear mechanics', reprint from 'the Mechanics and Biophysics of Hearing', P.Dallos et al. (Eds.), Springer Verlag, 1990, 395-402

[Lyon, 1996]

Lyon R.F.: 'The All-Pole Gammatone Filter and Auditory Models', Internet, temporäre Unterlagen für 'Computational Models of Signal Processing in the Auditory System', Forum Acusticum 1996, Antwerpen

[Marks, 1996]

Marks L.E.: 'Recalibrating the perception of loudness: Interaural transfer', J. Acoust. Soc. Am., 100(1), 1996, 473-480

[Meddis, 1988]

Meddis R.: 'Simulation of auditory-neural transduction: Further studies', J. Acoust. Soc. Am. 83, 1988, 1056-1063

[Meddis, Hewitt, 1991]

Meddis R., Hewitt M.J.: 'Modeling the perception of concurrent vowels with different fundamental frequencies', J. Acoust. Soc. Am. 91, 1991, 233-245

[Meddis, 1996]

Meddis R.: 'Dual resonance nonlinear filter (DRNL)', Internet WWW:
<http://info.lut.ac.uk/departments/hu/groups/speechlab/drnlp.html>

[Møller, Nilsson, 1979]

Møller A.R., Nilsson H.G.: 'Inner Ear Impuls Response and Basilar Membrane Modelling', Acustica, Vol. 41, 1979, 258-262

[Moore, Glasberg, 1983]

Moore B.C.J., Glasberg B.R.: 'Suggested formulae for calculating auditory-filter bandwidths and excitation patterns', J. Acoust. Soc. Am. 74(3), 1983, 750-753

[Moore et al., 1984]

Moore B.C.J., Glasberg B.R., Roberts B.: 'Defining the measurement of psychophysical tuning curves', J. Acoust. Soc. Am. 76(4), 1984, 1057-1066

[Moore, 1986]

Moore B.C.J. (Ed.), 'Frequency Selectivity in Hearing', Academic, London, 1986

[Moore, Glasberg, 1987]

Moore B.C.J., Glasberg B.R.: 'Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns', Hear. Res. 28, 1987, 209-225

[Moore, 1989]

Moore B.C.J.: 'An Introduction to the Psychology of Hearing', Academic Press, 1989

[Moore et al., 1990]

Moore B.C.J., Peters R.W., Glasberg B.R.: 'Auditory filter shapes at low center frequencies', J. Acoust. Soc. Am. 88(1), 1990, 132-140

[Moore et al., 1992]

Moore B.C.J., Glasberg B.R., Simpson A.: 'Evaluation of a method of simulating reduced frequency selectivity', J. Acoust. Soc. Am. 91, 1992, 3402-3423

[Moore, 1993]

Moore B.C.J.: 'Characterisation of Simultaneous, Forward and Backward Masking', In: The Proceedings of the AES 12th International Conference - The Perception of Reproduced Sound, 1993, June 28-30

[Moore, 1995]

Moore B.C.J. (Ed.): 'Hearing-Handbook of Perception and Cognition', 2nd Edition, Academic Press, 1995

[Moore et al., 1995]

Moore B.C.J., Glasberg B.R., van der Heijden M., Houtsma A.J.M., Kohlrausch A.: 'Comparison of auditory filter shapes obtained with notched-noise and noise-tone maskers', *J.Acoust.Soc.Am.* 97(2), 1995, 1175-1182

[Moore et al., 1996a]

Moore B.C.J., Hafter E.R., Glasberg B.R.: 'The probe-signal method and auditory-filter shape: Results from normal- and hearing-impaired subjects', *J.Acoust.Soc.Am.* 99(1), 1996, 542-552

[Moore et al., 1996b]

Moore B.C.J., Wojtczak M., Vickers D.A.: 'Effect of loudness recruitment on the perception of amplitude modulation', *J. Acoust. Soc. Am.*, 100(1), 1996, 481-489

[Moore, Glasberg, 1996]

Moore B.C.J., Glasberg B.R.: 'A Revision of Zwicker's Loudness Model', *Acustica - acta acustica*, Vol. 82, 1996, 335-346

[Moore et al., 1997]

Moore B.C.J., Glasberg B.R., Baer T.: 'A Model for the Prediction of Thresholds, Loudness, and Partial Loudness', *Journal of the Audio Engineering Society*, Vol. 45, Nr. 4, 1997, 224-240

[Moore, 1997]

Moore B.C.J.: 'Psychoacoustic Consequences of Compression in the Peripheral auditory System', in Schick A., Klatt M. (Eds.), *Contributions to Psychological Acoustics, Results of the Seventh Oldenburg Symposium on Psychological Acoustics*, Carl von Ossietzky Universität Oldenburg, 1997, 565-586

[Müller, Fichtl, 1994]

Müller F., Fichtl E.: 'Kategorien-Unterteilungs-Isophonen und Isochronen der Lautheit', *Konferenz der Deutschen Gesellschaft für Akustik, DAGA 94*, Technische Universität Dresden, 1994, 1101-1104

[Nedzelnitzky, 1980]

Nedzelnitzky V.: 'Sound pressures in the basal turn of the cat cochlea', *J.Acoust.Soc.Am.* 68, 1980, 1676-1689

[Neely, Kim, 1986]

Neely S.T., Kim D.O.: 'A model for active elements in the Cochlear mechanics', *J. Acoust. Soc. Am.* 79(5), 1986, 1472-1480

[Neumann, 1993]

Neumann J.: 'Lärmmeßpraxis am Arbeitsplatz und in der Nachbarschaft', Expert Verlag, 6. Auflage, 1993

[Nuttall, 1996]

Nuttall A.L., Dolan D.F.: 'Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig', *J. Acoust. Soc. Am.* 99(3), 1996, 1556-1565

[O'Loughlin, Moore, 1981]

O'Loughlin B. J., Moore B. C. J.: 'Improving psychoacoustical tuning curves', *Hear. Res.* 5, 1981, 343-346

[Ostergaard, 1990]

Ostergaard P.B.: 'Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse', *J.*

Acoust. Soc. Am. 87(4), 1990, 1813-1816

[Patterson, 1976]

Patterson R. D.: 'Auditory filter shapes derived with noise stimuli', J. Acoust. Soc. Am. 59, 1976, 640-654

[Patterson, Nimmo-Smith, 1980]

Patterson R. D., Nimmo-Smith I.: 'Off-frequency listening and auditory filter asymmetry', J. Acoust. Soc. Am. 67, 1980, 229-245

[Patterson et al., 1982]

Patterson R. D., Nimmo-Smith I., Weber D.L., Milroy R.: 'The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold', J. Acoust. Soc. Am. 72 (6), 1982, 1788-1803

[Patterson, Moore, 1986]

Patterson R.D., Moore B.C.J.: 'Auditory Filters and excitation patterns as representations of frequency resolution', In: Moore B.C.J. (Ed.), Frequency Selectivity in Hearing, Academic, London, 1986, 123-177

[Patterson et al., 1987]

Patterson R.D., Nimmo-Smith I., Holdsworth J., Rice P.: 'An efficient auditory filterbank based on the gammatone function', meeting of the IOC Speech Group on Auditory Modelling at RSRE, Dezember 1987

[Patterson, 1987]

Patterson R.D.: 'A pulse ribbon model of monaural phase perception', J. Acoust. Soc. Am., 82(5), 1987, 1560-1586

[Patterson, Holdsworth, 1991]

Patterson R.D., Holdsworth J.: 'A functional model of neural activity patterns and auditory images', In Advances in Speech, Hearing and Language Processing, (W.A. Ainsworth, ed.) Vol 3. JAI Press, London

[Patterson et al., 1992a]

Patterson R.D., Holdsworth J., Allerhand M.: 'Auditory Modells as Preprocessors for Speech Recognition', In The Auditory Processing of Speech, Schouten (ed.), Mouton de Gruyter, 1992, 67-83

[Patterson et al., 1992b]

Patterson R.D., Robinson K., Holdsworth J., McKeown D., Zhang C., Allerhand M.H.: 'Complex sounds and auditory images', In Auditory Physiology and Perception, (Eds.) Cazals Y., Demany L., Horner K., Pergamon, Oxford, 1992, 429-446

[Patterson, 1994]

Patterson R.D.: 'The sound of a sinusoid: Spectral models', J. Acoust. Soc. Am. 96 (3), 1994, 1409-1418

[Patterson et al., 1995]

Patterson R.D., Allerhand M.H., Giguere C.: 'Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform', J. Acoust. Soc. Am. 98 (4), 1995, 1890-1894

[Pflüger et al., 1996a]

Pflüger M., Höldrich R., Riedler W.: 'dB(ELC) - dynamische Phonkurvenanpassung als Erweiterung des dB(A)', Konferenz der Deutschen Gesellschaft für Akustik, DAGA 96, Bonn, 1996

[Pflüger et al., 1996b]

Pflüger M., Höldrich R., Riedler W.: 'dB(ELC) - A proposal to modify the dB(A)', INTER•NOISE, Liverpool, 1996

[Pflüger et al., 1997a]

Pflüger M., Höldrich R., Riedler W.: 'A nonlinear functional model of the spectral analysis performed in the peripheral auditory system', ICSPAT, 8th International Conference on Signal Processing, San Diego, 1997

[Pflüger et al., 1997b]

Pflüger M., Höldrich R., Riedler W.: 'A nonlinear model of the peripheral auditory system', AES 103rd Convention, New York, 1997

[Pick, 1980]

Pick G.F.: "Level dependence of psychophysical frequency resolution and auditory filter shape", J. Acoust. Soc. 68(4), 1980, 1085-1095

[Plack, Moore, 1990]

Plack C.J., Moore B.C.J.: 'Temporal window shape as a function of frequency and level', J. Acoust. Soc. Am., 87(5), 1990, 2178-2187

[Plack, 1996a]

Plack C.J.: 'Temporal factors in referential intensity coding', J. Acoust. Soc. Am. 100(2), 1996, 1031-1042

[Plack, 1996b]

Plack C.J.: 'Loudness enhancement and intensity discrimination under forward and backward masking', J. Acoust. Soc. Am. 100(2), 1996, 1024-1030

[Ronacher, Stücklschwaiger, 1994]

Ronacher A., Stücklschwaiger W.: 'Entwicklung eines objektiven Beurteilungsmaßstabes für die subjektive Lästigkeit von Motorengeräuschen', Beitrag zur Fachtagung: 'Soundengineering / kundenbezogene Akustikentwicklung in der Fahrzeugtechnik', Haus der Technik, Essen, 1994

[Ronacher, 1996]

Ronacher A.: 'Praktische Anwendung der Psychoakustik; Geräuschgestaltung: Über die Psychoakustik zur Kundenzufriedenheit', Beitrag zum FVV-Workshop, VDMA-Haus, Frankfurt am Main, 1996

[Rosen, Baker, 1994]

Rosen S., Baker R.J.: 'Characterising auditory filter nonlinearity', Hear. Res. 73, 1994, 231-243

[Rosen et al., 1996]

Rosen S., Baker R.J., Darling A.M.: 'Level dependence of auditory filtering at 2 kHz in normal and hearing-impaired listeners', <http://www.phon.ucl.ac.uk/home/richard/aro/poster.htm>

[Rosowski, 1991a]

Rosowski J.: 'The effect of external- and middle-ear filtering on auditory threshold and noise-induced hearing loss', J. Acoust. Soc. Am. 90 (1), 1991, 124-135

[Rosowski, 1991b]

Rosowski J.: 'Erratum: The effect of external- and middle-ear filtering on auditory threshold and noise-induced hearing loss', J. Acoust. Soc. Am. 90 (6), 1991, 3373

[Schäfer, 1978]

Schäfer P.: 'Vergleichende Analyse von Lärmbewertungs-Verfahren', Umweltforschungsplan des Bundesministers des Inneren, Institut für Ergonomie, TU München, 1979

[Scheidt et al., 1986]

Scheidt R., Bartsch R., Brückner Chr.: 'Die extraaurale Wirkung von Lärm - eine Literaturübersicht', Zeitschrift für Lärmbekämpfung 33, 1986, 151-159

[Schick, 1990]

Schick A.: 'Schallbewertung - Grundlagen der Lärmforschung', Springer Verlag, 1990

[Schiffbänker et al., 1991]

Schiffbänker H., Brandl F.K., Thien G.E.: 'Development and Application of an Evaluation Technique to Assess the Subjective Character of Engine Noise', Noise and Vibration Conference and Exposition, Traverse City, Michigan, 1991, SAE Technical Paper Series, 911081

[Schlauch, Morison, 1995]

Schlauch R.S., Morison L.M.: 'The Spectral Spread of Loudness Adaption for Tones that Decrease Continuously in Level', *Acustica* 81, 1995, 279-280

[Schouten, 1992]

Schouten M.E.H. (Ed.): 'The Auditory Processing of Speech', Mouton de Gruyter, Berlin, New York, 1992

[Schroeder, 1973]

Schroeder M.R.: 'An integrable model for the basilar membrane', *J. Acoust. Soc. Am.* 53(2), 1973, 429-434

[Schöder, 1981]

Schröder E.: 'Lärm - Entwurf einer Definition', *Zeitschrift für Lärmbekämpfung* 28, 1981, 154-155

[Seneff, 1988]

Seneff S.A.: 'A joint synchrony/mean-rate model of auditory speech processing', *Journal of Phonetics* No. 16, 1988, 55-76

[Shailer et al., 1990]

Shailer M.J., Moore B.C.J., Glasberg B.R., Watson N., Harris S.: 'Auditory filter shapes at 8 and 10 kHz', *J. Acoust. Soc. Am.* 88(1), 1990, 141-148

[Slaney, 1988]

Slaney M.: 'Lyon's Cochlea Model', Apple Computer Technical Report #13, 1988

[Slaney, 1993]

Slaney M.: 'An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank', Apple Computer Technical Report #35, 1993

[Slaney et al., 1994]

Slaney M., Naar D., Lyon R.F.: 'Auditory model inversion for sound separation', *Proc. Of IEEE ICASSP*, Vol. II, 1994, 77-80

[Slaney, 1996]

Slaney M.: 'Pattern playback in the 90's', In: Tesauro G., Touretzky D., Leed T. (Eds.): *Advances in Neural Information Processing Systems 7*, MIT Press, Cambridge, MA, 1996

[Sørensen, 1994]

Sørensen M.F.: 'Equal Loudness Level Contours for Frequencies 20-1000 Hz at 10, 20, 40, 60, 80 and 100 Phon Levels', *Inter•Noise 94*, Yokohama-Japan, 1994, 1073-1075

[Sporer et al., 1995]

Sporer T., Gbur U., Herre J., Kapust R.: 'Evaluating a Measurement System', *J. Audio Eng. Soc.*, Vol. 43, No. 5, 1995, 353-363

[Stone et al., 1997]

Stone M.A., Moore B.C.J., Glasberg B.R.: 'A Real-time DSP-Based Loudness Meter', in Schick A., Klatt M. (Eds.), *Contributions to Psychological Acoustics, Results of the Seventh Oldenburg Symposium on Psychological Acoustics*, Carl von Ossietzky Universität Oldenburg, 1997, 587-601

[Straßen, 1979]

Straßen T.: 'Subjektive Lästigkeit von Lärm und ihre Erfassung durch Lärmmeßverfahren - Eine vergleichende

- Untersuchung mit Hilfe der Hauptkomponentenanalyse', Dissertation, Technische Hochschule Aachen, 1979
- [Suzuki, Sone, 1993]
Suzuki Y., Sone T.: 'Frequency Characteristics of Loudness Perception: Principles and Applications, In Sixth Oldenburg Symposium on Psychological Acoustics, 1993, 193-221
- [Takeshima et al., 1994]
Takeshima H., Suzuki Y., Kumagai M., Sone T., Fujimori T.: 'Subjective and objective transformation level between free-field and diffuse-field in equal-loudness level contours', INTER NOISE 94, Yokohama, Japan, 1994, 1077-1080
- [Thomassen, 1989]
Thomassen H.G.: 'Die Lautheitsbewertung nach Zwicker: Für und Wider', Zeitschrift für Lärmbekämpfung 36, 1989, 49-50
- [Tobias, 1970]
Tobias J.V.: 'Foundations of Modern Auditory Theory', Academic Press, Inc., London, 1970
- [Traunmüller, 1990]
Traunmüller H.: 'Analytical expression for the tonotopic sensory scale', J. Acoust. Soc. Am. 88(1), 1990, 97-100
- [Tyler et al., 1984]
Tyler R.S., Hall J.W., Glasberg B.R., Moore B.C.J., Patterson R.D.: 'Auditory filter asymmetry in the hearing impaired', J. Acoust. Soc. Am., 76(5), 1984, 1363-1368
- [Voss et al., 1996]
Voss S.E., Rosowski J.J., Peake W.T.: 'Is the pressure difference between the oval and round windows the effective acoustic stimulus for the cochlea?', J. Acoust. Soc. Am. 100(3), 1996, 1602-1616
- [Wan, Yu, 1995]
Wan W., Yu X.: 'Transfer Function, Cochlear Map and Bandwidth Characteristics', Acustica 81, 1995, 89-93
- [Wang, Shamma, 1995]
Wang K., Shamma S.A.: 'Spectral Shape Analysis in the Central Auditory System', IEEE Transactions on Speech and Audio Processing, Vol. 3, No.5, 1995, 382-395
- [Watts, 1993]
Watts L.: 'Cochlear Mechanics', Dissertation, California Institute of Technology, Pasadena, Kalifornien, 1993
- [Weber, 1977]
Weber, 1977, J. Acoust. Soc. Am. 62, 424-429
- [Wong, Norwich, 1995]
Wong W., Norwich K.H.: 'Obtaining equal loudness contours from Weber fractions' J. Acoust. Soc. Am., 97(6), 1995, 3761-3767
- [Wright, 1996]
Wright B.A.: 'Auditory filter asymmetry at 2000 Hz in 80 normal-hearing ears', J. Acoust. Soc. Am., 100(3), 1996, 1717-1721
- [Yang, Fan, 1992]
Yang J., Fan C.: 'Modelling of nonlinear Cochlear Mechanics', Acustica 77, 1992, 100-105
- [Zhou, 1995]
Zhou B.: 'Auditory filter shapes at high frequencies', J. Acoust. Soc. Am. 98(4), 1995, 1935-1942

[Zollner, Zwicker, 1993]

Zollner M., Zwicker E.: 'Elektroakustik', 3. Aufl., Springer Verlag, 1993

[Zollner, 1996]

Zollner M.: 'Psychoakustische Meßtechnik', Cortex electronic GmbH, Erzb.-Buchberger-Allee 14, D-93051 Regensburg

[Zölzer, 1994]

Zölzer U.: 'Roundoff Error Analysis of Digital Filters', J. Audio Eng. Soc., Vol. 42, No. 4, 1991, 232-244

[Zweig et al., 1976]

Zweig G., Lipes R., Pierce J.R.: 'The cochlear compromise', J. Acoust. Soc. Am. 59(4), 1976, 975-982

[Zwicker, Feldtkeller, 1967]

Zwicker E., Feldtkeller R.: 'Das Ohr als Nachrichtenempfänger', S. Hirzel Verlag, Stuttgart, 1967

[Zwicker, Bubel, 1977]

Zwicker E., Bubel G.: 'Einfluß nichtlinearer Effekte auf die Frequenzselektivität des Gehörs', Acustica 38, 1977, 67-71

[Zwicker, Terhardt, 1980]

Zwicker E., Terhardt E.: 'Analytical expressions for critical-band rate and critical bandwidth as a function of frequency', J. Acoust. Soc. Am., 68(5), 1980, 1523-1525

[Zwicker, 1982]

Zwicker E.: 'Psychoakustik', Springer Verlag, Berlin, Heidelberg, 1982

[Zwicker, 1986]

Zwicker E.: 'A hardware cochlear nonlinear preprocessing model with active feedback', J. Acoust. Soc. Am. 80(1), 1986, 146-153

[Zwicker, Fastl, 1990]

Zwicker E., Fastl H.: 'Psychoacoustics Facts & Models', Springer Verlag, 1990

[Zwicker, 1991]

Zwicker E.: 'Ein Vorschlag zur Definition und zur Berechnung der unbeeinflussten Lästigkeit', Zeitschrift für Lärmbekämpfung 38, 1991, 91-97