

Markus Fruhmann: **Ein Modell zur Wahrnehmung der Ausgeprägtheit der Tonhöhe**

Markus Fruhmann

Ein Modell zur Wahrnehmung
der Ausgeprägtheit der Tonhöhe

Lehrstuhl für Mensch-Maschine-Kommunikation
der Technischen Universität München

Ein Modell zur Wahrnehmung der Ausgeprägtheit der Tonhöhe

Markus Fruhmann

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. G. Färber

Prüfer der Dissertation:

1. apl. Prof. Dr.-Ing., Dr.-Ing. habil. H. Fastl
2. Univ.-Prof. Dr.-Ing. K. Diepold

Die Dissertation wurde am 18.05.2006 bei der Technischen Universität München
eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik
am 20.11.2006 angenommen.

Die Neugier steht immer an erster Stelle
eines Problems, das gelöst werden will.
(Galileo Galilei)

Inhaltsverzeichnis

Einleitung	1
1 Grundlagen	5
1.1 Tonhöhenwahrnehmung	5
1.2 Die psychoakustische Größe “Ausgeprägtheit der Tonhöhe“ (AdT)	8
1.3 Physiologische und psychoakustische Grundlagen	9
1.3.1 Das menschliche Gehör	9
1.3.2 Frequenzgruppen und Tonheit	11
1.4 Die Fourier-t-Transformation	13
1.5 Modell der Lautheit nach Zwicker	14
1.6 Versuchsmethoden zur Untersuchung der Ausgeprägtheit der Tonhöhe	15
1.6.1 Größenschätzung mit Ankerschall	15
1.6.2 Random Access	16
1.6.3 Verfahren mit erzwungener Auswahl	16
1.7 Auswertung experimenteller Daten und Versuchsumgebung	18
2 Experimentelle Untersuchungen zur Ausgeprägtheit der Tonhöhe	19
2.1 Grundsätzliche Abhängigkeiten der Ausgeprägtheit der Tonhöhe	19
2.1.1 Frequenz eines Sinustons	19
2.1.2 Schalldauer	20
2.1.3 Schallpegel	21
2.1.4 Bandbreite von Bandpassrauschen	21
2.2 Amplitudenmodulation	22
2.3 Vergleich der Ausgeprägtheit der Tonhöhe verschiedener Schallarten	25
2.4 Ausgeprägtheit der Tonhöhe von harmonisch komplexen Tönen (HKT)	29
2.5 Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Residualtöne	31
2.5.1 Einfluß der Anzahl entfernter Harmonischer eines harmonisch komplexen Residualtons	31
2.5.2 Spektrale Dämpfung harmonisch komplexer Töne	32
2.6 Ausgeprägtheit der Tonhöhe beim Übergang von spektraler zu virtueller Tonhöhenwahrnehmung	34
2.7 Durch Kammfiltereffekte hervorgerufene Tonhöhen	35
2.7.1 Iterated Rippled Noise (IRN) als Anker- und Testschall	37
2.7.2 Iterated Rippled Noise im Vergleich zu Rauschen mit einfacher Blockwiederholung	38

2.7.3	Iterated Rippled Noise im Vergleich zu harmonisch komplexen Tönen	40
2.8	Weitere Untersuchungen zur Ausgeprägtheit der Tonhöhe von Bandpassrauschen	41
2.8.1	Unterschiedsschwelle der Ausgeprägtheit der Tonhöhe	42
2.8.2	Ausgeprägtheit der Tonhöhe von Bandpassrauschen verschiedener Mittenfrequenzen	43
2.8.3	Abhängigkeit von der Frequenzgruppenbreite	44
2.9	Zusammenfassung	46
3	Zeitliche und spektrale Merkmale der Ausgeprägtheit der Tonhöhe	47
3.1	Gehörgerechte Darstellung von Audiosignalen	47
3.1.1	Gewählte Transformationsparameter der Fourier-t-Transformation .	47
3.1.2	Berechnung von Mithörschwellenmustern	48
3.1.3	Modifikation des Lautheitsmodells nach Zwicker	50
3.1.4	Resultierende Spektraldarstellung	52
3.2	Zeitliche und spektrale Merkmale der Ausgeprägtheit der Tonhöhe	53
3.2.1	Sinustöne	53
3.2.2	Harmonisch komplexe Töne	54
3.2.3	Bandpassrauschen	55
3.2.4	Kammfilterrauschen	57
3.2.5	Harmonisch komplexe Residualtöne	59
3.3	Zusammenfassung	59
4	Ein Modell der Ausgeprägtheit der Tonhöhe	61
4.1	Übersicht über das Modell der Ausgeprägtheit der Tonhöhe	61
4.2	Spektraltransformation mittels FTT	64
4.3	Extraktion von Spektraltonhöhen	64
4.4	Berechnung und Gewichtung virtueller Tonhöhen	65
4.4.1	Gegenüberstellung von Ausgeprägtheit der Tonhöhe und “Pitch Sali- ence“	65
4.4.2	Rücktransformation von Tonhöhengewichten in das Pegelmaß	67
4.5	Berechnung von Mithörschwellenmustern (MHS)	68
4.6	Bestimmung der Grundfrequenz	68
4.7	Ermittlung des zeitlichen Verlaufs der Tonhöhe	69
4.8	Bildung des “Störanteils“	70
4.9	Berechnung der Lautheits-Zeit Verläufe von Signal- und Störanteil	71
4.10	Zusätzliche Gewichtungen des Signalanteils	72
4.10.1	Bandbreite	72
4.10.2	Amplitudenvariation	73
4.10.3	Grundfrequenz	75
4.11	Gewichtung aufgrund der Gesamtlautheit	75
4.12	Zusammenfassung	76

5	Verifikation des Modells der Ausprägtheit der Tonhöhe	77
5.1	Gütekriterien zur Beurteilung der Modellberechnungen	77
5.2	Versuchsergebnisse im Vergleich zu berechneter Ausprägtheit der Tonhöhe	79
5.2.1	Abhängigkeit der Ausprägtheit der Tonhöhe von Frequenz, Pegel und Dauer eines Sinustons	79
5.2.2	Ausprägtheit der Tonhöhe amplitudenmodulierter Sinustöne . . .	81
5.2.3	Ausprägtheit der Tonhöhe von harmonisch komplexen Tönen in Abhängigkeit von der Anzahl Harmonischer	83
5.2.4	Ausprägtheit der virtuellen Tonhöhe von harmonisch komplexen Residualtönen	84
5.2.5	Ausprägtheit der Tonhöhe von Bandpassrauschen	86
5.2.6	Ausprägtheit der Tonhöhe von Iterated Rippled Noise	87
5.2.7	Vergleich verschiedener Schallarten	88
5.3	Zusammenfassung	91
6	Anwendung der Ausprägtheit der Tonhöhe zur Geräuschbeurteilung	93
6.1	Bewertung der Klangqualität realer und synthetischer Geräusche	93
6.1.1	Bewertung der Klangqualität von Orchesterpauken	93
6.1.2	Beurteilung der Geräuschqualität in Hochgeschwindigkeitszügen . .	96
6.1.3	Beurteilung von Lüftergeräuschen	98
6.2	DIN 45 681: Ein Verfahren zur Bewertung tonaler Komponenten	100
6.2.1	Übersicht über DIN 45 681	100
6.2.2	Vergleich zur Ausprägtheit der Tonhöhe	101
6.3	Zusammenfassung	102
7	Zusammenfassung	103
	Danksagung	107
A	Abkürzungen und Formelzeichen	109
B	Weitere Daten und Versuchsergebnisse	113
B.1	Amplitudenmodulierte Sinustöne	113
B.2	Übergang von spektraler zu virtueller Tonhöhe	114
C	Versuchspersonen	115
	Literaturverzeichnis	117

Einleitung

Mit dem Begriff “Design“ verbindet man zunächst die optische Gestaltung von italienischen Sportwagen, französischer Mode oder außergewöhnlichen Möbeln. Aber auch Gegenstände des täglichen Lebens wie Fernseher, Staubsauger und PC-Gehäuse werden in der jüngeren Vergangenheit immer stärker an die optischen, haptischen und akustischen Wünsche des jeweiligen Kundenkreises angepasst. Das damit verbundene “Anfassgefühl“ stellt für viele Kunden ein kaufentscheidendes Argument dar, spiegelt es doch scheinbar die Qualität eines Produkts wider. So wird keine Firma, die sich gegenüber ihren Wettbewerbern behaupten möchte, diese Kundenanforderungen ignorieren.

Auch im akustischen Sinne steigt der Stellenwert von Design zunehmend. So ist der legendäre Klang einer Harley Davidson eng mit dieser Marke verbunden und auch ein Drucker in einem Büro sollte ein gewisses Klangbild besitzen, das sicher nicht dem aufdringlichen Geräusch eines Nadeldruckers entspricht. Oftmals ist man für Geräte des täglichen Lebens also auf der Suche nach Klangbildern, die den Eigenschaften und Funktionen entsprechen. Neben der Lautheit, die bei der erwähnten Harley eher hoch sein soll und beim Drucker in einem Büro gering, beinhaltet dies beispielsweise die richtige Abstimmung der im Geräusch auftretenden tonalen Komponenten.

Mit Hilfe der Psychoakustik, die die Zusammenhänge zwischen physikalischen Schallparametern und der damit verbundenen auditiven Wahrnehmung untersucht, ist es möglich, die Geräuschqualität über gehörgerechte Größen zu beschreiben. Beispielhaft seien hier die Wahrnehmung der Lautheit, Schwankungsstärke, Rauigkeit oder der Tonhöhe genannt [Zwi99]. Letzteres ist im Zusammenhang mit dem aktiven oder passiven Genuss von Musik und der damit verbundenen Notendarstellung ein geläufiger Begriff. Die Unterscheidung von Tonhöhen auf einer Skala von “tief“ nach “hoch“ ist also weit verbreitet und in vielen Fällen sinnvoll und zweckmäßig.

Gerade in der Musik sind Tonhöhen in Verbindung mit Melodien von essentieller Bedeutung. Musikinstrumente können anhand ihrer Klangfarbe und dem zeitlichen Ein- und Ausschwingverhalten auch von Laien sehr gut voneinander unterschieden werden. Ein Aspekt der Klangfarbe ist dabei die Deutlichkeit der jeweils gespielten Note bzw. Tonhöhe, also die Ausgeprägtheit der Tonhöhe des wahrgenommenen Klanges. Diese Ausgeprägtheit der Tonhöhe sollte bei Musikinstrumenten ein gewisses Maß nicht unterschreiten, um Melodien gut wahrnehmen zu können.

Bei den Betriebsgeräuschen von Geräten ist es dagegen sicherlich meist unerwünscht, daß diese in zu hohem Maße tonal klingen und so ist es erstrebenswert, die Ausgeprägtheit der Tonhöhe solcher Geräusche, z.B. des Reifen-Fahrbahn-Geräusches eines PKW, zu reduzieren. Bei der subjektiven Beurteilung von Geräuschen ist daher immer der Kontext, in dem das jeweilige Klangbild auftritt, zu berücksichtigen.

In der vorliegenden Arbeit wird nun eben diese Wahrnehmung der Ausgeprägtheit der Tonhöhe näher untersucht und mit physikalischen Parametern in Verbindung gebracht, mit dem Ziel ein Modell zur Beschreibung der beobachteten Daten zu erstellen. Dabei ist die Arbeit in folgende Abschnitte gegliedert:

Zunächst werden in **Kapitel 1** die Grundlagen der Psychophysik im Hinblick auf die Tonhöhenwahrnehmung diskutiert. Dabei werden die Begriffe "spektrale" und "virtuelle" Tonhöhe eingeführt sowie die Wahrnehmung der "Ausgeprägtheit der Tonhöhe" als psychoakustische Empfindungsgröße definiert. Zum Verständnis der im weiteren Verlauf der Arbeit beschriebenen Modellvorstellungen des menschlichen Gehörs werden daran anschließend einige physiologische und psychoakustische Grundlagen dargestellt. Insbesondere die sogenannte "Frequenz-Orts-Transformation" zur Umsetzung eingehender Schallwellen in Nervensignale und deren Interpretation als Filterbank werden dabei beschrieben. Zudem werden die verwendeten Versuchsmethoden und die Auswertung zugehöriger Ergebnisse erläutert, wie sie zur Erhebung psychoakustischer Daten im Rahmen dieser Arbeit verwendet wurden.

Die grundlegenden Abhängigkeiten der psychoakustischen Empfindungsgröße "Ausgeprägtheit der Tonhöhe" von physikalischen Schallparametern sind in **Kapitel 2** dargestellt. Dabei werden zunächst Abhängigkeiten für Sinustöne, diejenigen Schalle, welche die stärkste Ausgeprägtheit der Tonhöhe hervorrufen, gezeigt. Dies geschieht anhand teilweise bereits bekannter Untersuchungen zur Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Frequenz, vom Pegel, von der Dauer und von Amplitudenmodulationsparametern eines Sinustons. Darüber hinaus wurden Experimente für harmonisch komplexe Töne durchgeführt, deren Tonhöhe sowohl spektraler als auch virtueller Natur sein kann. Diese Schalle lassen sich alle über Linienspektren, also die Überlagerung einzelner Sinuskomponenten beschreiben. Gerade in der Geräuschbeurteilung hat man es jedoch meist mit Schallen zu tun, deren Spektren kontinuierlich aufgebaut sind und so eher rauschhaften Charakter besitzen. Als Übergang von Schallen mit Linienspektren zu solchen rauschhaften Signalen dienen Kammfilter-Rauschen. Diese lassen sich beispielsweise durch mehrfaches Durchlaufen identischer Kammfilter stark in ihrer Ausgeprägtheit der Tonhöhe variieren. Um weiteren Aufschluß über die Mechanismen und Zusammenhänge bei der Bildung der Ausgeprägtheit der Tonhöhe zu erhalten wird eingehend auf die Abhängigkeiten bei Schmal- und Bandpassrauschen eingegangen. Es werden also Einflüsse der physikalischen Parameter Frequenz, Pegel, Bandbreite und Amplitudenmodulation dargestellt. Aus den Versuchsergebnissen werden jeweils einfache Zusammenhänge für die Abhängigkeit der Empfindungsgröße vom jeweiligen Schallparameter abgeleitet.

Ausgehend von einer Spektrogrammdarstellung werden in **Kapitel 3** Merkmale hergeleitet, um die Zusammenhänge der Ausgeprägtheit der Tonhöhe mit Schallparametern qualitativ beschreiben zu können. Mit Hilfe der Fourier-t-Transformation und anschließender Mithörschwellenberechnung sowie Lautheitstransformation entsteht eine gehörge-rechte Darstellung. Diese zeigt zeitliche und spektrale Eigenschaften der experimentell untersuchten Schalle, welche mit den Versuchsergebnissen zur Ausgeprägtheit der Tonhöhe in kausalen Zusammenhang gebracht werden können.

Aufbauend auf diesen Merkmalen wird in **Kapitel 4** ein Algorithmus entwickelt, welcher die verschiedenen Einflüsse miteinander verknüpft, so daß sich die gewonnenen Daten beschreiben lassen. Dazu wird neben dem Berechnungsmodell der Lautheit nach Zwicker auf ein Verfahren zur Bestimmung virtueller Tonhöhen zurückgegriffen. Durch eine geeignete Nachverarbeitung der hieraus gewonnenen Daten wird ein Modell gebildet, welches die Ausgeprägtheit der Tonhöhe beschreibt. Die Anpassung an Versuchsergebnisse aus Kapitel 2 erfolgt dabei in Form von Gewichtungen, beispielsweise abhängig von der spektralen Lage, der Bandbreite oder auch zeitlichen Variationen des Signals.

Die Qualität des entstandenen Modells wird in **Kapitel 5** anhand von Vergleichen mit experimentellen Daten überprüft. Zur Qualitätsbeurteilung dienen dabei mathematische Kriterien wie die Rangfolge der experimentellen und berechneten Daten sowie die Lage der Berechnungen bezüglich der Interquartilbereiche der experimentell ermittelten Daten.

Kapitel 6 zeigt schließlich die Anwendbarkeit der Wahrnehmungsgröße “Ausgeprägtheit der Tonhöhe“ zur Beurteilung der Geräuschqualität anhand realer Schalle und solchen nachempfundenen synthetischen Schallen. Dabei werden neben einem Beispiel aus dem Bereich der musikalischen Akustik vor allem Schalle technischer Geräte herangezogen. Bei diesen ist die Bewertung tonaler Komponenten bereits Gegenstand von Normen und Richtlinien, die, aufbauend auf technischen Messungen, Bewertungskriterien und Pegelzuschläge definieren. Ein Vergleich mit diesen etablierten Verfahren zur Bewertung der Tonhaltigkeit, welche nicht immer mit der Ausgeprägtheit der Tonhöhe gleichzusetzen ist, positioniert schließlich die Ausgeprägtheit der Tonhöhe neben weiteren subjektiven und objektiven Verfahren als Möglichkeit zur Beurteilung von Schallen bezüglich der Deutlichkeit darin enthaltener tonaler Komponenten.

Abschließend wird ab Seite 103 in einer **Zusammenfassung** ein Überblick über die wesentlichen Ergebnisse der vorliegenden Arbeit gegeben.

Kapitel 1

Grundlagen

Im folgenden Kapitel werden Grundlagen, die zum weiteren Verständnis der Arbeit beitragen, erläutert. Zunächst wird ein Einstieg in die Tonhöhenwahrnehmung und deren Theorie gegeben. Ein kurzer Überblick physiologischer Aspekte des menschlichen Hörorgans und deren Bedeutung für die Wahrnehmung unterschiedlicher Tonhöhen folgt daran anschließend. Dabei sollen Grundzüge der Schallverarbeitung im Außen-, Mittel- und Innenohr dargestellt werden, sowie die Interpretation der sensorischen Schallumwandlung auf der Basilarmembran als Filterbank. Das Kapitel wird mit einem Überblick über die verwendeten Versuchsmethoden und deren Auswertung abgeschlossen.

1.1 Tonhöhenwahrnehmung

Der Begriff der Tonhöhe ist in Verbindung mit der Notendarstellung in der Musik weit verbreitet und geläufig. Meist handelt es sich bei der Tonhöhe um eine mit der Frequenz einer Schwingung korrelierte Wahrnehmung in dem Sinne, daß eine Erhöhung der Schwingungsfrequenz auch zu einer Erhöhung der wahrgenommenen Tonhöhe führt.

Bei Sinustönen hat beispielsweise eine innerhalb des hörbaren Bereichs niedrige Frequenz die Wahrnehmung einer niedrigen Tonhöhe zur Folge, eine hohe Frequenz führt entsprechend zu einer hohen Tonhöhe. Einheit der Tonhöhe H ist die sogenannte “pitch unit“ pu . Diese wird im Rahmen dieser Arbeit als diejenige Frequenz eines Sinustons f_T , mit einem Pegel von 60 dB definiert, dessen Tonhöhe mit der des Testschalls übereinstimmt [Ter79]. Also:

$$H = f_T / Hz [pu] \quad (1.1)$$

Diese Definition weicht von derjenigen in [Fle34] und [Ter72b] ab, wo die Normierung auf einen Lautstärkepegel von 40 phon anstelle eines Schallpegels von 60 dB erfolgte.

Es darf also keine generelle Identität von Frequenz und Tonhöhe angenommen werden. Verschiebungseffekte aufgrund der spektralen Lage, des Pegels oder gegenseitiger Beeinflussung verschiedener Schallkomponenten werden in der Literatur ausführlich behandelt [Wal69a, Ter70, Ter71, Ter76, Gru87]. Da jedoch die exakte Tonhöhe für die vorliegende Arbeit von untergeordneter Rolle ist, wird hier von der Tonhöhe eines Schalls gesprochen,

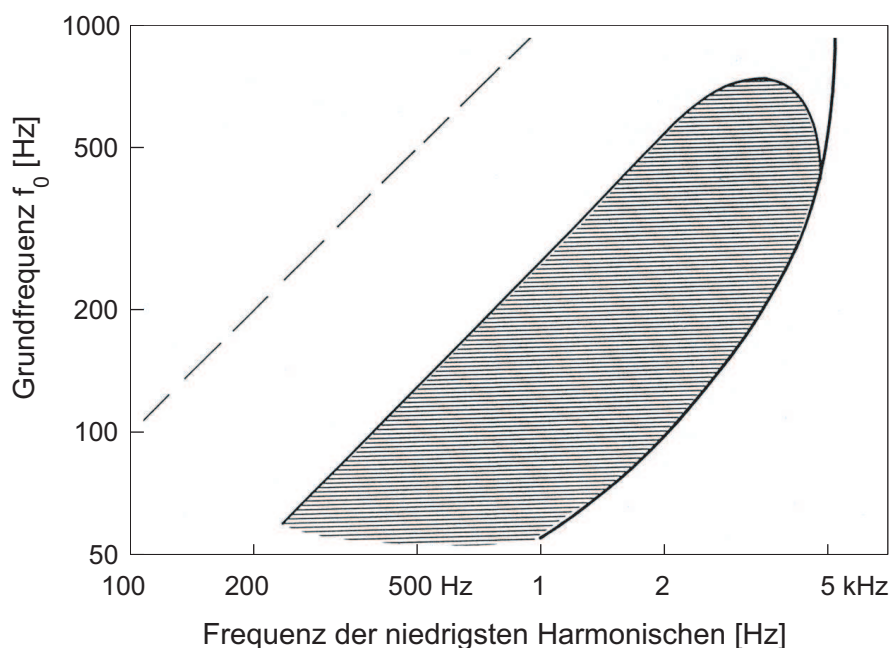


Abbildung 1.1: Existenzbereich der virtuellen Tonhöhe [Ter72b]. Harmonisch komplexe (Residual-) Töne deren spektrale Zusammensetzung Komponenten innerhalb der schraffierten Fläche enthält, führen zur Wahrnehmung virtueller Tonhöhen.

ohne deren genauen Wert kennen zu müssen. Im Folgenden wird anstelle der tatsächlich wahrgenommenen Tonhöhe lediglich die Grundfrequenz eines Schalls benannt, die die Tonhöhe maßgeblich bestimmt.

Spricht man von “spektraler Tonhöhe“, so ist die die Tonhöhe bestimmende Frequenz unmittelbar als Spektralkomponente im Schall enthalten. Dies ist beispielsweise bei Sinustönen der Fall und auch bei harmonisch komplexen Tönen wird die wahrgenommene Tonhöhe weitestgehend durch die Frequenz der ersten Harmonischen repräsentiert.

Schalle, die ihre Grundfrequenz nicht direkt, sondern nur als Vielfache in Form höherer Harmonischer enthalten, rufen trotz des Fehlens dieser Komponente eine Tonhöhe hervor, die eben dieser fehlenden Grundfrequenz-Komponente zuzuordnen ist. Man spricht in solchen Fällen von Residualklängen deren Tonhöhe durch virtuelle Tonhöhen bzw. Residualtonhöhen dominiert wird [Ter72b, Ter74, Pat76]. Dieser Effekt tritt beispielsweise bei der Kommunikation über den bandbegrenzten Sprechkanal eines Telefons auf. Trotz einer Bandbegrenzung auf einen Bereich von ca. 300 Hz bis 3,3 kHz läßt sich erkennen, ob ein männlicher oder weiblicher Sprecher gehört wird, obwohl die Grundfrequenz der Stimme mit ca. 100 bzw. 200 Hz jeweils außerhalb des übertragenen Bandes liegt.

Damit virtuelle Tonhöhen wahrnehmbar werden, muß die spektrale Zusammensetzung eines Schalls bestimmte Voraussetzungen erfüllen. Diese lassen sich über den in Abbil-

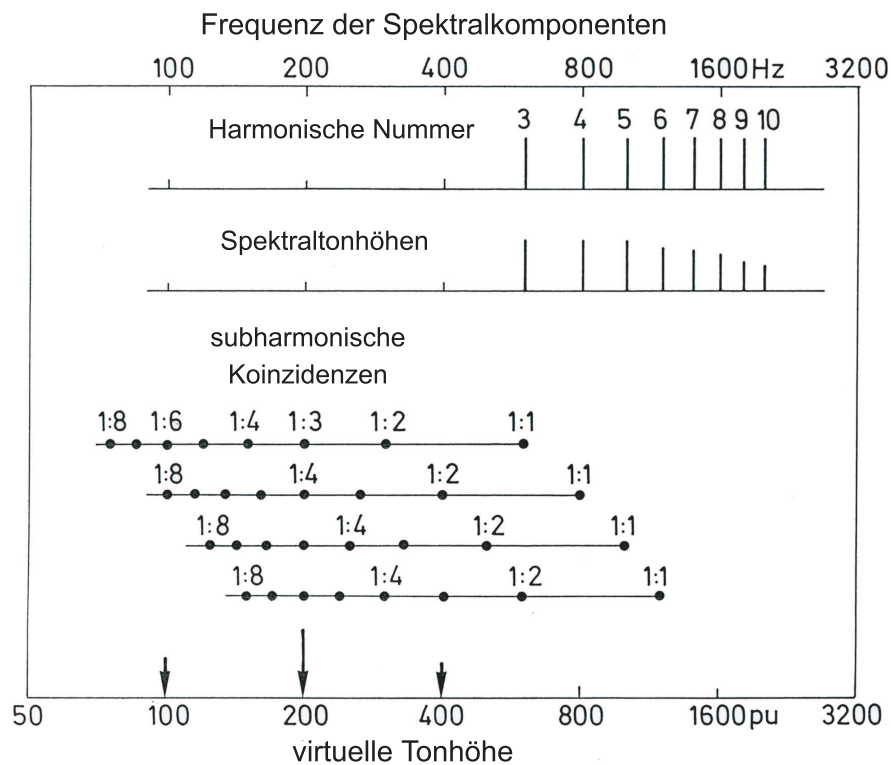


Abbildung 1.2: Schema des Modells zur Ermittlung virtueller Tonhöhen [Ter98].

dung 1.1 als schraffierte Fläche dargestellten Existenzbereich der virtuellen Tonhöhe veranschaulichen [Rit62, Ter72b]. Bei harmonisch komplexen (Residual-) Tönen muß die niedrigste im Schall enthaltene Spektralkomponente innerhalb dieses Bereichs liegen, um virtuelle Tonhöhen hervorzurufen. Ein Modell zur Bestimmung virtueller Tonhöhen und damit verbundener Verschiebungseffekte basiert auf dem Prinzip der “subharmonischen Koinzidenzen“ [Wal69b, Wal69c, Ter79].

Dabei werden für die spektral vorhandenen Teiltöne die Subharmonischen durch ganzzahlige Division gebildet. Über das Auffinden von Orten besonders vieler Übereinstimmungen (Koinzidenzen) dieser Subharmonischen ergeben sich die Frequenzen möglicher wahrnehmbarer virtueller Tonhöhen. Details des Verfahrens, das schematisch in Abbildung 1.2 dargestellt ist, bietet [Ter79].

Komplexe Schalle können mehrere Tonhöhenempfindungen hervorrufen. Die synthetische Wahrnehmung dominiert dabei im Allgemeinen die Tonhöhenwahrnehmung, so daß die Grundtonhöhe als die “Tonhöhe“ eines Schalls empfunden wird [Ter98]. Mit Hilfe von Akzentuierung oder als geübter Hörer läßt sich die analytische Wahrnehmung fördern und Teiltöne eines Schalls werden als Einzelkomponenten wahrnehmbar. In der vorliegenden Arbeit wird der Begriff “Tonhöhe“ jedoch stets im Sinne der synthetischen Wahrnehmung verwendet und bezeichnet somit die Grundtonhöhe eines Schalls [Ter72a, Sch98].

1.2 Die psychoakustische Größe “Ausgeprägtheit der Tonhöhe” (AdT)

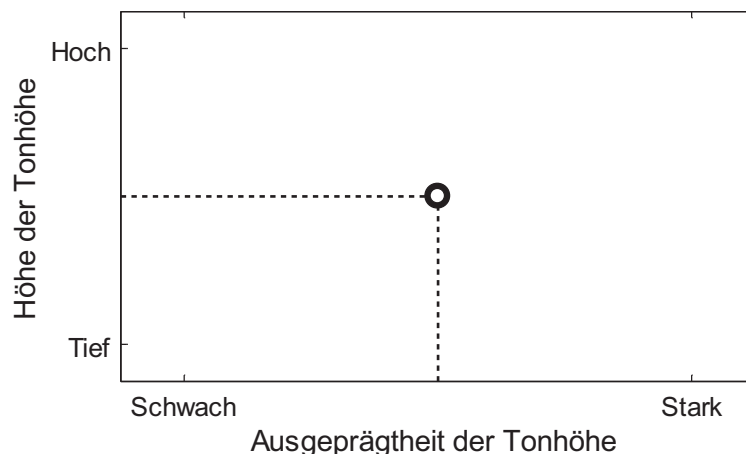


Abbildung 1.3: Zweidimensionale Wahrnehmung der Tonhöhe in den Dimensionen “Hoch“-”Tief“ zur Beschreibung der absoluten Tonhöhe und “Schwach“-”Stark“ für die Ausgeprägtheit der Tonhöhe.

Schalle, die eine wahrnehmbare Tonhöhe hervorrufen, lassen sich bezüglich dieser entlang einer Skala von “tief“ nach “hoch“ anordnen. Darüber hinaus ist es möglich, daß die Tonhöhe eines Schalls mehr oder weniger deutlich wahrgenommen wird. Man spricht dabei von mehr oder weniger stark ausgeprägten Tonhöhen, die zugehörige Empfindungsgröße wird als “Ausgeprägtheit der Tonhöhe“ bzw. im Englischen als “pitch strength“ bezeichnet [Plo75, Fas79c, Fas81]. Ein Schall läßt sich also bezüglich seiner Tonhöhe in einem mindestens zweidimensionalen Raum entsprechend Abbildung 1.3 positionieren.

In der Literatur werden zudem u.a. die Begriffe “Pitch Saliency“ [Ter82, Peg04], “Tonal Strength“ [Lee05] und “Tonhaltigkeit“ [Vor98, Dan02] zur Beschreibung der Deutlichkeit von Tonhöhen verwendet. Dabei werden jeweils einzelne Schallkomponenten bezüglich ihrer Wahrnehmbarkeit bewertet, beispielsweise durch die Genauigkeit, mit der Sinustöne auf die Frequenzen einzelner Komponenten harmonisch komplexer Töne eingeregelt werden können. Die “Ausgeprägtheit der Tonhöhe“ beschreibt jedoch diese Deutlichkeit in Form einer globalen Größe, also nicht eine Eigenschaft einzelner Komponenten des Schalls sondern die Deutlichkeit der nominellen Tonhöhe bzw. der Grundtonhöhe des Schalls.

Ein einfaches Beispiel ist die unterschiedliche Ausgeprägtheit der Tonhöhe von Sinustönen und harmonisch komplexen Tönen. Sinustöne besitzen eine sehr große Ausgeprägtheit der Tonhöhe, wohingegen harmonisch komplexe Töne, wie sie näherungsweise von Musikinstrumenten erzeugt werden, eine schwächere Ausgeprägtheit der Tonhöhe besitzen [Fas97].

Sinustöne stellen für eine bestimmte Tonhöhe jeweils das Maximum der Ausprägtheit der Tonhöhe dar, dagegen besitzen Signale wie weisses oder rosa Rauschen eine gegen Null gehende Ausprägtheit der Tonhöhe, da ihnen keine eindeutige Tonhöhe zuzuordnen ist.

Zur weiteren Illustration dient folgendes Beispiel: Eine Versuchsperson, der auf einer Klaviertastatur die Note C1 vorgespielt wird, wird nach der Tonhöhe gefragt, spontan "tief" antworten. Generiert man jedoch ein Tiefpassrauschen, dessen Bandkante eine vergleichbare Frequenz besitzt und so eine ähnliche Tonhöhe hervorruft [Fas71], wird die Versuchsperson spontan antworten, daß die Tonhöhe verglichen mit dem Klang des Klaviers, sehr schwach oder möglicherweise auf Anhieb gar nicht wahrnehmbar ist. Die Ausprägtheit der Tonhöhe kann also bei gleicher oder ähnlicher wahrgenommener Tonhöhe von Versuchspersonen deutlich unterschieden und so auch quantitativ bewertet werden.

1.3 Physiologische und psychoakustische Grundlagen

Zum weiteren Verständnis werden nun einige Grundlagen aus dem Bereich der Physiologie und der Psychoakustik dargestellt. Diese dienen der Erläuterung der Modellvorstellung der Schallverarbeitung im menschlichen Gehör, dessen Aufbau und Funktionsweise zunächst knapp beschrieben wird.

1.3.1 Das menschliche Gehör

Das menschliche Hörsystem läßt sich grob in drei Teile untergliedern:

(A) das Außenohr bzw. die Ohrmuschel, welches umgangssprachlich als "das Ohr" bezeichnet wird,

(B) das Mittelohr, in dem die Luftschwingungen zum lymphgefüllten Innenohr übertragen werden und

(C) das Innenohr, welches aus den eingehenden Schwingungen Nervensignale erzeugt, die wiederum im Gehirn weiterverarbeitet werden.

Das Außenohr ist der sichtbare Teil des Gehörs. Die Ohrmuschel dient hauptsächlich als Schalltrichter um den einfallenden Schall zu bündeln und an den Gehörgang weiterzugeben. Die Form des Außenohres (Pinna) und die Abmessungen des Gehörgangs haben maßgeblichen Einfluß auf das Richtungshören sowie die Empfindlichkeit des Hörsystems. Dies zeigt sich z.B. an einem charakteristischen Minimum der Ruhehörschwelle (RHS) bei Frequenzen um 3-4 kHz [Zwi55]. Diese Senke entspricht der Resonanzfrequenz eines $\lambda/4$ -Resonators mit einer Länge von 3 cm, wie sie beim Gehörgang vorzufinden ist. Das Außenohr wird vom Trommelfell, einer schwingungsfähigen Membran, abgeschlossen.

An diese Membran schließt das Mittelohr an, welches das Trommelfell mit dem ovalen Fenster verbindet. Da die Schallwellen, die das Außenohr empfängt in der Regel durch

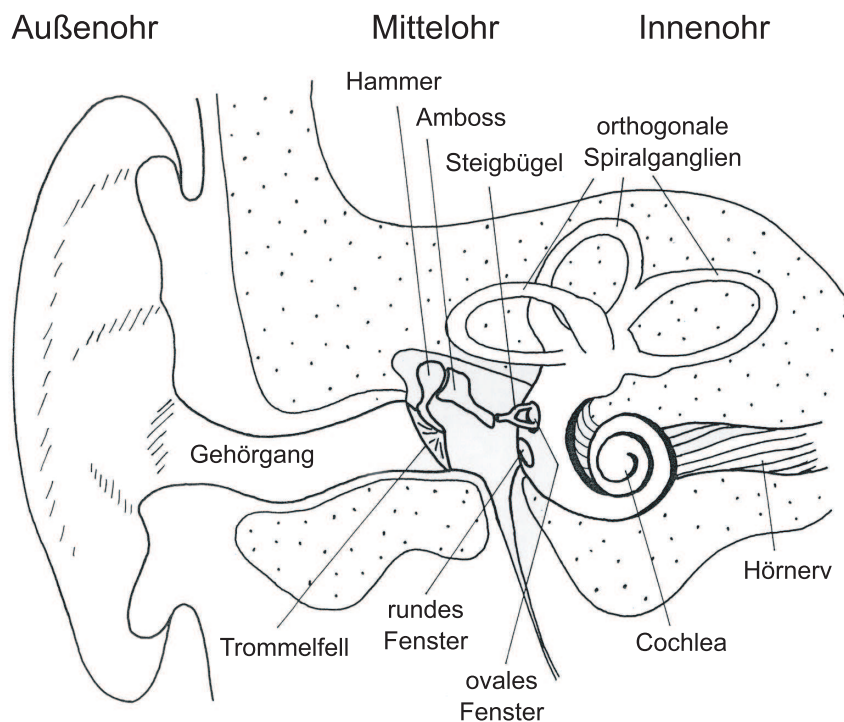


Abbildung 1.4: Schematischer Schnitt durch das menschliche Hörsystem. Darstellung nach [Zwi99].

das Medium Luft übertragen werden, muß eine Anpassung der unterschiedlichen Impedanzen von Luft und Lymphe, der Flüssigkeit, mit der das Innenohr gefüllt ist, erfolgen. Das zu dieser Impedanzanpassung benötigte Übersetzungsverhältnis entsteht durch die Kombination dreier Knöchelchen (Hammer, Amboss und Steigbügel) und dem Verhältnis ihrer Angriffsflächen an Trommelfell und ovalem Fenster.

Im Innenohr erfolgt die Umsetzung der Schwingungen in Nervensignale. Es besteht dabei nicht ausschließlich aus einem Organ, das dem Zweck des Hörens dient, sondern beinhaltet zudem das Gleichgewichtsorgan in Form dreier orthogonaler Spiralganglien. Die dreigeteilte Schnecke (Cochlea) wiederum beinhaltet, als Trennmembran zwischen den beiden Hauptkammern, die sogenannte Basilarmembran. Entlang dieser bewegt sich, ausgelöst durch das Schwingen des ovalen Fensters, eine Wanderwelle, welche die Basilarmembran zu Schwingungen anregt die von den inneren Haarzellen in afferente Nervenimpulse umgesetzt werden.

Der Ort der maximalen Auslenkung der Basilarmembran und damit der größten Anregung der Haarzellen, ist von der Frequenz eines Schalls abhängig. Die ortsabhängige Kombination aus festem Knochen und flexibler Basilarmembran führt zu unterschiedlichen Steifigkeiten und somit verschiedenen Resonanzfrequenzen. So werden unterschiedliche Schwingungsfrequenzen auf jeweils spezifische Anregungsorte auf der Basilarmembran abgebildet, was als "Frequenz-Orts-Transformation" bezeichnet wird.

1.3.2 Frequenzgruppen und Tonheit

Nach Zwicker läßt sich das Verhalten der Basilarmembran als Filterbank interpretieren [Zwi62, Zwi67]. Dazu ist es nötig, die zugehörigen Filter nicht nur anhand ihrer Mittenfrequenzen zu charakterisieren, sondern auch ihre jeweilige Bandbreite zu spezifizieren. Um diese Größen zu beschreiben muß zunächst die vorangehend erwähnte Frequenz-Orts-Transformation näher erläutert werden.

Die Frequenz-Orts-Transformation läßt sich durch das Verhältnis der weichen Basilarmembran und dem festen Knochen der Cochlea qualitativ beschreiben. Dieses Verhältnis ist von der Entfernung vom ovalen Fenster abhängig und weist entsprechend unterschiedliche Resonanzfrequenzen auf. An ihrem Eingang, besitzt die Schnecke den größten Durchmesser, die Basilarmembran ist hier im Gegensatz jedoch relativ schmal und steif. So ergibt es sich, daß hier die höchsten wahrnehmbaren Frequenzen (bis ca. 20 kHz) eine Anregung verursachen. An der Spitze der Schnecke (Helicotrema) ist dieses Verhältnis dagegen umgekehrt, die Basilarmembran ist relativ breit und weich und der Knochen nimmt nur noch einen kleinen Raum ein. Hier werden also die tiefen Frequenzen abgebildet. Ein Vorschlag zur Modellierung des Verhaltens der Basilarmembran findet sich beispielsweise in [Str85].

Diese Abbildung der Frequenzen auf die unterschiedlichen Orte der Basilarmembran läßt sich nicht durch einen streng linearen oder logarithmischen Zusammenhang beschreiben, sondern als Mischform hieraus [Fel55]. Wird die Basilarmembran linear unterteilt, so läßt sich eine Skala definieren, die die Orte der Abbildung bestimmter Frequenzen auf die Basilarmembran beschreibt. Diese Skala wird als "Tonheits"-Skala bezeichnet. Die Tonheit z erhält dabei die Einheit "Bark" nach dem deutschen Wissenschaftler Barkhausen.

Zur Beschreibung der Schall-Analyse durch die Basilarmembran existiert somit ein physiologischer Zusammenhang zwischen der "Lage" von Filtern auf der Basilarmembran und der erfassten Frequenz. Die Bandbreite der Filter läßt sich beispielsweise über Experimente zur Lautheitswahrnehmung von Dreitonkomplexen beschreiben. Erhöht man ausgehend von einer Mittenfrequenz f_c den Abstand zweier benachbarter Tonkomponenten und damit die Bandbreite Δf des resultierenden Schalls, so bleibt die Lautheit bis zu einer bestimmten, von der Mittenfrequenz abhängigen Bandbreite konstant. Erst bei weiterer Bandbreitenvergrößerung steigt die wahrgenommene Lautheit an [Bau56, Zwi57]. Frequenzbereiche bis zu dieser Grenzbandbreite Δf_G unterliegen offenbar einer gemeinsamen Verarbeitung durch das Gehör. Diese Frequenzbereiche werden als "Frequenzgruppen", die zugehörige Grenzbandbreite als "Frequenzgruppenbreite" bezeichnet [Zwi60b].

Ordnet man aneinander angrenzend frequenzgruppenbreite Filter an, so ergibt sich, daß sich eine Unterteilung des hörbaren Bereich in ca. 24 Bänder. Die Tonheit z wird also so definiert, daß sie diese 24 Frequenzgruppen repräsentiert. Ihre Abbildung auf die Basilarmembran ist linear, so daß beispielsweise bei der halben Länge der Basilarmembran (ca. 1,6 cm) eine Tonheit von 12 Bark entsprechend einer Frequenz von ca. 1700 Hz vorliegt

[Zwi60a]. Eine näherungsweise Umrechnungsmöglichkeit der physikalischen Frequenz f in die physiologisch orientierte Tonheit z ist über die Formel

$$z[\text{Bark}] = 13 \arctan(0,76f[\text{kHz}]) + 3,5 \arctan(f[\text{kHz}]/7,5)^2 \quad (1.2)$$

möglich [Zwi80]. Diese Skala stellt nun den Zusammenhang zwischen Frequenz und Tonheit her, was der Nummer der Frequenzgruppe in oben beschriebener Anordnung benachbarter Filter entspricht [Zwi61]. Sie ist also auch ein Maß für den Ort auf der Basilarmembran, der durch eine bestimmte Frequenz maximal angeregt wird und stellt somit eine physiologisch richtige Skalierung der Frequenz dar [Fel53, Bek60].

Aus Versuchsdaten läßt sich für die Frequenzgruppenbreite Δf_G der Zusammenhang

$$\Delta f_G[\text{Hz}] = 25 + 75[1 + 1,4(f[\text{kHz}])^2]^{0,69} \quad (1.3)$$

ableiten [Zwi80]. Näherungsweise ist die Frequenzgruppenbreite bis zu einer Mittenfrequenz von 500 Hz konstant bei 100 Hz und steigt darüber mit der Frequenz an, wobei sie einen Wert von 20 Prozent der Mittenfrequenz annimmt.

In einer Realisierung der Eigenschaften der Basilarmembran als Filterbank müssen nun die Mittenfrequenzen proportional zur Barkskala gewählt werden. Die Filterbandbreiten sind proportional zur Frequenzgruppenbreite bei der jeweiligen Mittenfrequenz zu wählen. Aus dieser Kombination ergeben sich Analysefähigkeiten, die denjenigen des menschlichen Gehörs Rechnung tragen und durch die geeignete Wahl der Proportionalitäten angepasst werden können. Ein Schema zur Veranschaulichung dieser Modellierung der Frequenz-Orts-Transformation ist in Abbildung 1.5 dargestellt.

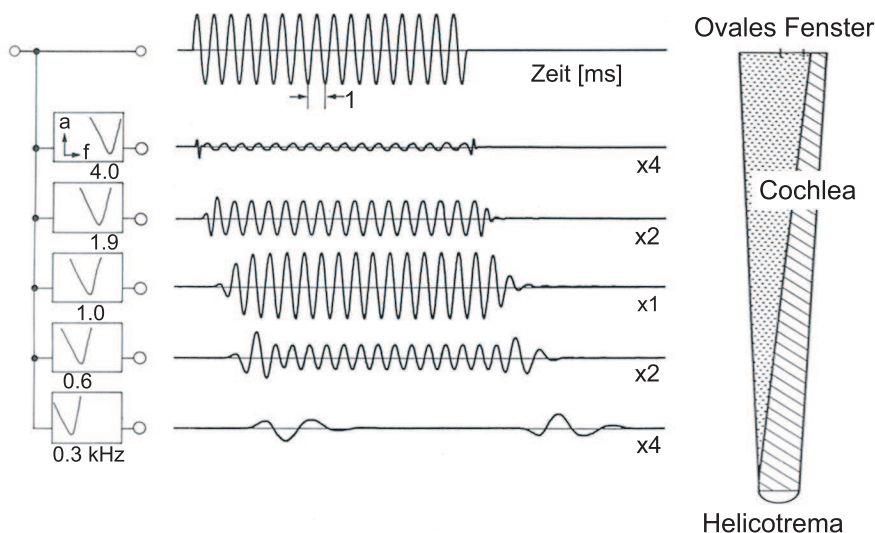


Abbildung 1.5: Modellierung der Frequenz-Orts-Transformation als Filterbank. Abbildung nach [Zwi99]

1.4 Die Fourier-t-Transformation

Neben der Darstellung in Form von Zeitsignalen haben sich in der Akustik Spektral- und Spektrogrammdarstellungen etabliert. Dazu bietet sich in der Informationstechnik eine Vielzahl möglicher Transformationen an, die am weitesten verbreitete ist sicherlich die Fourier-Transformation und ihre zeitvariante Version in Form der Kurzzeit Fourier-Transformation (STFT). In der Psychoakustik findet sich häufig die Interpretation des menschlichen Gehörs als Filterbank [Zwi62, Plo64, Zwi67]. Eine Spektraltransformation, die zur Modellierung einer psychoakustischen Größe beitragen soll, muß Eigenschaften besitzen, die sich im Gehör wiederfinden, das heißt Filterbandbreite und -einschwingverhalten sollten proportional zu den beim menschlichen Gehör gefundenen Eigenschaften sein.

Während die Fourier-Transformation [Opp89]

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \cdot e^{-j\omega t} dt \quad (1.4)$$

für alle Analysefrequenzen konstante Filter-Bandbreiten besitzt sind diese bei der über

$$Wf(u, s) = \int_{-\infty}^{+\infty} g(t) \frac{1}{\sqrt{s}} \Psi\left(\frac{t-u}{s}\right) dt \quad (1.5)$$

definierten Wavelet-Transformation [Vet92, Mal99] proportional zur aktuellen Mittenfrequenz. Beide Transformationen finden Anwendung in der Signalanalyse und auch in der Kodierung von Audiosignalen [Rio91, Ern00, Tza01]. Beide Arten von Transformationen lassen jedoch aufgrund der in Abschnitt 1.3.2 gezeigten Eigenschaften der Analyse durch die Basilarmembran keine gehörangepasste Spektralanalyse zu.

Als Transformation, die diese Eigenschaften berücksichtigt, wird von Terhardt die sogenannte Fourier-t-Transformation (FTT) vorgeschlagen [Ter85a, Ter85b]. Diese ist für einseitige Signale (Beginn bei $t=0$) über

$$F(\omega, t) = \int_0^t g(\lambda) e^{-a(t-\lambda)} e^{-\omega\lambda} d\lambda \quad (1.6)$$

definiert, wobei $e^{-a(t)}$ eine zeitliche Gewichtungsfunktion $h(\omega, t)$ darstellt. Diese wird so gewählt, daß der Eigenschaft natürlicher Spektralanalysatoren, zu denen das Gehör zu zählen ist, Rechnung getragen wird, weiter zurückliegende Signalanteile mit der Zeit „zu vergessen“.

Die Parametrierung dieser Gewichtungsfunktion bestimmt sowohl das zeitliche als auch das spektrale Auflösungsvermögen der Transformation. Zur Anpassung dieser Parameter ist es nötig Höreigenschaften wie das Frequenzunterscheidungsvermögen oder die Detektionsschwellen für Frequenzmodulationen zu berücksichtigen.

1.5 Modell der Lautheit nach Zwicker

Zahlreiche Untersuchungen zur Empfindungsgröße ‘Lautheit’ führten zunächst zu einem grafischen Bestimmungsverfahren ausgehend von gemessenen Terzbandpegeln [Zwi58, Zwi59a]. Später entstanden hieraus Berechnungsverfahren [Pau72], die schließlich zusammen mit dem grafischen Bestimmungsverfahren in die Norm DIN 45 631 [DIN91] und den internationalen Standard ISO 532 [ISO75] eingingen.

Das Verfahren besteht grundsätzlich aus drei Schritten: Zunächst werden gemessene Terz- oder Frequenzgruppenpegel des zu analysierenden Signals in die Kernlautheiten transformiert, was grafisch durch das Eintragen der gemessenen Pegelwerte in ein Diagramm entsprechend Abbildung 1.6 vorgenommen werden kann. Werden an diese Kernlautheiten die im Diagramm näherungsweise eingezeichneten, oberen Flanken der spektralen Verdeckung angehängt, so erhält man die spezifischen Lautheiten. Die Gesamtlautheit N des Schalls, angegeben in der Einheit ‘sone’, ist proportional zu der im Pegel-Diagramm entstehenden Fläche, die durch den Verlauf der spezifischen Lautheiten gebildet wird.

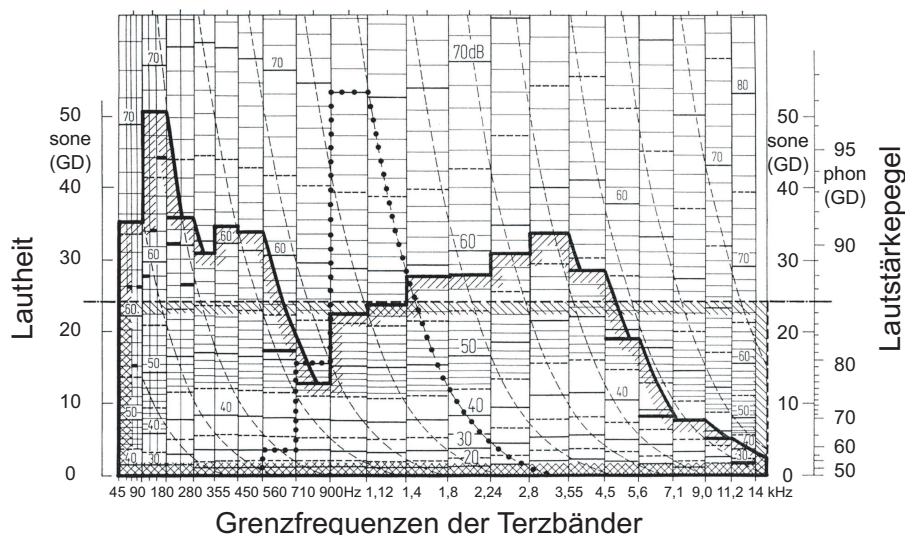


Abbildung 1.6: Pegel-Diagramm zur grafischen Bestimmung der Lautheit aus gemessenen Terzpegeln nach [Zwi59a, Zwi99].

In [Pau72] sowie DIN 45 631 [DIN91] ist ein Verfahren zur Berechnung der Lautheit aus Terz- bzw. Frequenzgruppenpegeln erläutert. Eine Erweiterung des Modells auf die veränderte Lautheitswahrnehmung bei Schwerhörenden war Teilaspekt der Arbeiten von Chalupper [Cha02a]. Gegenüber ursprünglich vorgeschlagenen technischen Realisierungen der automatischen Lautheitsbestimmung [Zwi77, Zwi78], sind die Verarbeitungsschritte ‘spektrale Maskierung’ und ‘Lautheitstransformation’ vertauscht, was in der speziellen Anwendung des Modells für Schwerhörende begründet ist [Cha02a]. Die spektrale Auflösung des Modells ist durch die Verwendung von 24 Bark Kanälen bestimmt, was für die Berechnung der Lautheit jedoch keine Einschränkung bedeutet.

1.6 Versuchsmethoden zur Untersuchung der Ausgeprägtheit der Tonhöhe

Die Erhebung experimenteller Daten der Ausgeprägtheit der Tonhöhe wurde im Rahmen dieser Arbeit mit drei unterschiedlichen psychophysikalischen Verfahren realisiert. Diese werden im Folgenden dargestellt und erläutert.

1.6.1 Größenschätzung mit Ankerschall

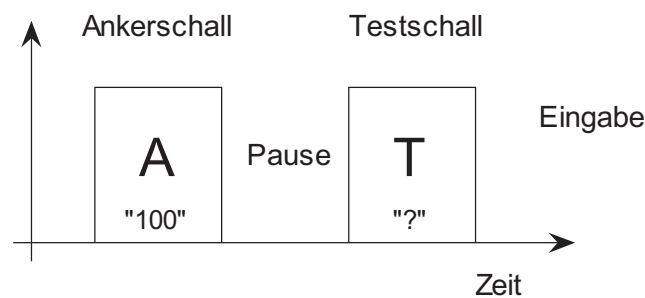


Abbildung 1.7: Ablauf eines Experiments mit Größenschätzung mit Ankerschall. Ein Schallpaar bestehend aus Anker- und Testschall wird dargeboten. Daraufhin gibt die Versuchsperson ihre Beurteilung beispielsweise über ein Computerterminal ein.

Die Versuchsmethode “Größenschätzung mit Ankerschall“, liefert quantitative Verhältnisse der abgefragten Empfindungsgröße [Ste56]. Der Versuchsperson werden dabei Schallpaare entsprechend Abbildung 1.7 dargeboten, die sie bezüglich der untersuchten Empfindungsgröße vergleichen soll. In der vorliegenden Arbeit soll jeweils die Wahrnehmungsgröße “Ausgeprägtheit der Tonhöhe“ zweier Schalle verglichen werden. Dazu wird dem jeweils ersten Schall (der Anker) eines Paares ein fester Wert zugeordnet, z.B. “100“. Der zweite Schall wird dann relativ zum Ersten beurteilt. Empfindet die Versuchsperson beispielsweise die Tonhöhe des Testschalls als doppelt so stark ausgeprägt wie diejenige des Ankerschalls, so antwortet sie mit dem doppelten Wert des Referenzschalls, also “200“.

Zu beachten ist bei dieser Versuchsmethode, daß durch den Bezug auf einen Ankerschall jeweils relative Werte der untersuchten Empfindungsgröße entstehen. Durch Normierung auf die Ergebnisse von Schallen, die in verschiedenen Experimenten dargeboten wurden, lassen sich Vergleiche der Ergebnisse unterschiedlicher Versuchsreihen anstellen. Um Abhängigkeiten vom jeweils verwendeten Ankerschall weitestgehend auszuschließen [Hel93], wurden mehrere Experimente für identische Fragestellungen durchgeführt. Dazu wurden in Hin- und Rückversuchen unterschiedliche Ankerschalle gewählt, die meist das Maximum oder das Minimum der Ausgeprägtheit der Tonhöhe der untersuchten Schalle markierten.

1.6.2 Random Access

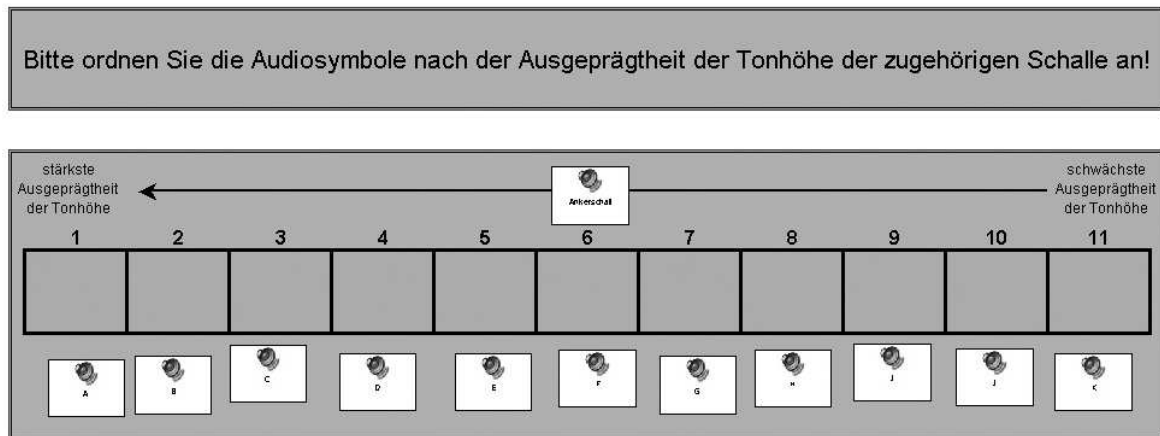


Abbildung 1.8: Grafische Oberfläche zur Versuchsmethode “Random Access“. Die Versuchsperson ordnet die Schalle bezüglich der untersuchten psychoakustischen Größe an, indem sie die zugehörigen Symbole am Bildschirm verschiebt. Der im Bildschirmmausschnitt eingetragene Ankerschall diente lediglich als Orientierung bezüglich der Tonhöhe, nicht aber als Referenz der Ausgeprägtheit der Tonhöhe.

Die Versuchsmethode “Random Access“ stellt eine Form von Rangfolgeexperiment dar [Fas00, Pat03a, Pat03b]. Sie gibt der Versuchsperson die Möglichkeit, über direkten Zugriff auf die zu beurteilenden Schalle, den Versuchsablauf selbst zu gestalten und die untersuchten Schalle beliebig oft anzuhören. Dazu sieht die Versuchsperson vor sich auf einem Bildschirm (ähnlich Abbildung 1.8) eine Reihe von Schallsymbolen, die durch anklicken den zugehörigen Testschall wiedergeben. Durch Anordnen der Schallsymbole entlang einer Skala entsteht eine Rangfolge der dargebotenen Signale, entsprechend der untersuchten Empfindungsgröße. Somit hat die Versuchsperson das Gefühl, selbst über den Versuchsablauf zu bestimmen, was die Motivation erhält und so Ermüdungseffekten vorbeugt [Hel93]. Anders als die Größenschätzung mit Ankerschall läßt diese Methode lediglich Urteile zu, welche Rangfolge die untersuchten Schalle bezüglich der untersuchten Größe bilden. Die quantitative Relation kann auf diese Weise nicht erfasst werden.

1.6.3 Verfahren mit erzwungener Auswahl

Die dritte verwendete Versuchsmethode basiert auf Verfahren, bei denen die Versuchsperson entscheiden muß, welcher Schall einer dargebotenen Gruppe, bezüglich einer Wahrnehmungsgröße den höheren Wert erreicht. Dabei wird die sogenannte 2-AFC Methode (“Two-Alternatives-Forced-Choice“) angewandt, über die wiederum ein adaptiver Einregelvorgang gesteuert werden kann. Um den Einregelvorgang zu beschleunigen werden Verfahren angewandt, die die Schrittweite der anzupassenden physikalischen Schallgröße adaptiv ändern. Dabei besteht sowohl die Möglichkeit, einfache Algorithmen zu verwen-

den [Dix48, Lev70], als auch statistisch aufwändige Verfahren wie den sogenannten PEST-Algorithmus (“Parameter Estimation by Sequential Testing“, [Wal66, Cre67, Kol88]).

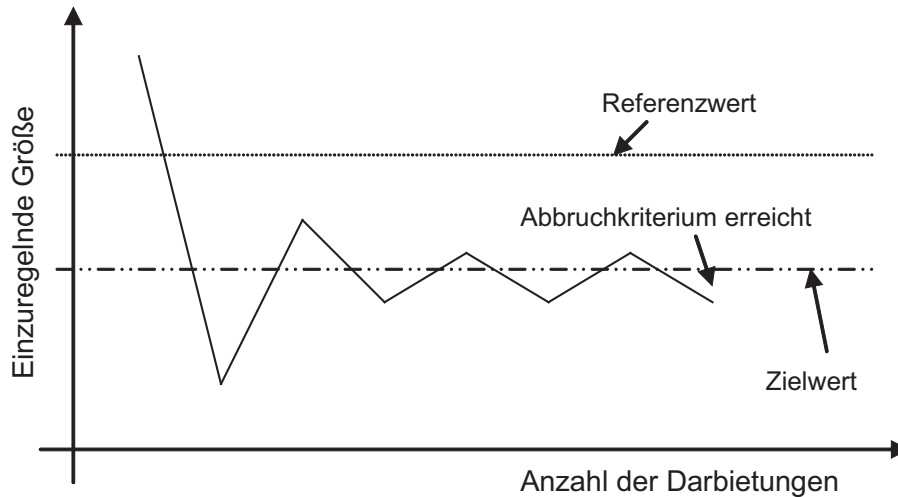


Abbildung 1.9: *Beispielhafter Einregelvorgang bei der Verwendung einer einfachen adaptiven Versuchsmethode. Bei Wendepunkten wird die Schrittweite der einzuregelnden Größe angepasst. Nach Erreichen einer minimalen Schrittweite wird nach einer festgelegten Anzahl von Wendepunkten das Abbruchkriterium erfüllt.*

Für ein einfaches adaptives Verfahren (“Adaptive Staircase“) entsprechend [Lev70] ist der Einregelvorgang für ein einzelnes Schallpaar in Abbildung 1.9 dargestellt. Je nach Urteil der Versuchsperson wird eine einzuregelnde physikalische Größe um einen bestimmten Wert verändert. Urteilt die Versuchsperson bei einer Darbietung anders als im vorangegangenen Durchlauf, muß die Größenänderung nun in umgekehrter Richtung erfolgen. Bis zu einer minimalen Schrittweite wird an solch einem Umkehrpunkt die Größenänderung (Schrittweite) verkleinert, beispielsweise halbiert. Ist die minimale Schrittweite erreicht, so läßt sich über eine festgelegte Anzahl n solcher Wendepunkte ein Abbruchkriterium definieren. Der Endwert der einzuregelnden Größe ergibt sich schließlich als Mittelwert oder Median der Antworten dieser letzten n Umkehrpunkte. Prinzipiell ist dieser Einregelvorgang bei Verwendung des PEST Algorithmus sehr ähnlich, jedoch basiert hier die Schrittweitenänderung sowie die Definition von Wendepunkten auf “Maximum Likelihood“- Kriterien und ist somit bei weitem komplexer.

Die Festlegung der minimalen Schrittweite hat unmittelbaren Einfluß auf die Durchführbarkeit dieser Experimente. Eine beliebig klein gewählte Endschrittweite würde zwar entsprechend genaue Ergebnisse ermöglichen, ist jedoch eine Versuchsperson nicht in der Lage solch kleine Schalländerungen wahrzunehmen, so erhöht sich die Versuchsdauer ohne zu genaueren Ergebnissen zu führen. Die minimale Schrittweite ist daher abhängig von der minimal wahrnehmbaren Schalländerung (“Just Noticeable Differences“, JND) und sollte dieser in Etwa entsprechen.

1.7 Auswertung experimenteller Daten und Versuchsumgebung

Mit den vorangehend beschriebenen Methoden durchgeführte Experimente liefern jeweils Ergebnisse für einzelne Schalle einer Versuchsreihe. Zur Auswertung der Ergebnisse einer einzelnen Versuchsperson werden nun Mediane und Interquartilbereiche über die Daten dieser Versuchsperson gebildet. Aufgrund seiner Skaleninvarianz und der Robustheit gegenüber Ausreißern wird der Median gegenüber dem arithmetischen Mittelwert bevorzugt [Bäu74]. Er ergibt sich für n gewonnene Ergebnisse x nach [Sac04] zu:

$$MED = \begin{cases} x_{(n+1)/2} & , n \text{ ungerade} \\ \frac{1}{2} (x_{n/2} + x_{n/2+1}) & , n \text{ gerade} \end{cases} \quad (1.7)$$

Der Median beschreibt also den Wert, über bzw. unter dem sich jeweils 50 % der Ergebnisse befinden. Der Interquartilbereich ergibt sich als Bereich zwischen dem 25 % und 75 % Perzentil der beobachteten Werte. So ergibt sich ein Maß für das mittlere Ergebnis sowie die intraindividuellen Schwankungen einer Versuchsperson.

Zur Ermittlung des Gesamtergebnisses eines Versuchs, aus den Ergebnissen der einzelnen Versuchspersonen, bieten sich mehrere Möglichkeiten an:

- (1) Mediane und Interquartile der Mediane der Versuchspersonen
- (2) Mediane der Mediane und Mediane der Interquartile der Versuchspersonen
- (3) Mediane und Interquartile über alle Einzelurteile aller Versuchspersonen

In Fall (1) entspricht die Darstellung dem mittleren Ergebnis der Versuchspersonen, wobei die dargestellten Streubereiche die interindividuellen Schwankungen der Mediane der einzelnen Versuchspersonen widerspiegeln. Fall (2) zeigt ebenfalls das mittlere Ergebnis der Versuchspersonen, die Streubereiche sind nun jedoch über die Interquartile der einzelnen Versuchspersonen definiert. Es wird also die mittlere intraindividuelle Streuung dargestellt. Schließlich lassen sich, wie Fall (3) andeutet, alle Versuchsergebnisse als Eingangsgröße für die Berechnung eines Mittelwerts und zugehöriger Streuungen ansehen. Da es sinnvoll erscheint das mittlere Ergebnis (Mediane) sowie die mittlere Schwankung der Urteile der Versuchspersonen verschiedener Versuchspersonen zu betrachten, werden sämtliche Auswertungen dieser Arbeit mit der unter (2) beschriebenen Variante erzeugt.

Sämtliche Experimente wurden in schallisolierten Versuchskabinen durchgeführt. Die Darbietung der Schalle erfolgte über freifeldentzerrte Kopfhörer vom Typ Beyerdynamic DT 48 [Zwi63b, Fas79a]. Die Schallerzeugung erfolgte dabei jeweils rechnergestützt mit Hilfe der Programmierumgebung MATLAB® [Mat14]. Außer bei den Versuchen, die mit der Methode "Random Access" durchgeführt wurden (hier wurden die Daten von Hand aus dem Programm Excel® übernommen und zur Weiterverarbeitung in Matlab®-Skripten eingetragen), wurden die Meßwerte vollautomatisch vom Rechner erfasst, gespeichert und mit dem geschilderten Verfahren ausgewertet.

Kapitel 2

Experimentelle Untersuchungen zur Ausgeprägtheit der Tonhöhe

Dieses Kapitel bietet einen Überblick über die Wahrnehmungsgröße “Ausgeprägtheit der Tonhöhe“ anhand experimenteller Untersuchungen. Dazu werden zunächst grundlegende Daten aus der Literatur zum Einfluß physikalischer Parameter wie Pegel, Dauer oder Frequenz von Sinustönen auf diese psychoakustische Größe dargestellt. Aufbauend auf diesen Experimenten wurden weitere Versuche durchgeführt. Diese zeigen einige Zusammenhänge der Ausgeprägtheit der Tonhöhe von harmonisch komplexen Tönen sowie Residualtönen, beispielsweise von der Gesamtanzahl Harmonischer oder auch von der Anzahl entfernter unterer Harmonischer. Darüber hinaus werden Ergebnisse von Experimenten mit rauschhaften Signalen wie Bandpassrauschen und Kammfilter-Rauschen vorgestellt. Diese Daten bilden ein Grundgerüst für das Verständnis der Ausgeprägtheit der Tonhöhe bezüglich ihrer Abhängigkeit von physikalischen Schallparametern.

2.1 Grundsätzliche Abhängigkeiten der Ausgeprägtheit der Tonhöhe

In [Fas89] sind die grundlegenden Abhängigkeiten der Ausgeprägtheit der Tonhöhe von Sinustönen aufgeführt. Diese sollen hier in einem kurzen Überblick dargestellt werden.

2.1.1 Frequenz eines Sinustons

Die Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Frequenz f_T eines Sinustons konstanten Darbietungspegels zeigt die in Abbildung 2.1 dargestellte Bandpasscharakteristik. Diese ist der Form der Kurven gleicher Lautheit [Zwi59b] bei vertikaler Spiegelung ähnlich. Tonhöhen von Sinustönen mittlerer Frequenzen (ca. 750 Hz bis 3 kHz) werden als sehr ausgeprägt wahrgenommen. Zu tieferen und höheren Frequenzen hin fällt die Ausgeprägtheit der Tonhöhe dagegen deutlich ab. Entgegen den Kurven gleicher Lautheit zeigen die Ergebnisse aus Abbildung 2.1 jedoch keine Bevorzugung von Frequenzen nahe der in Kapitel 1.3.1 beschriebenen Resonanzfrequenz bei 4 kHz.

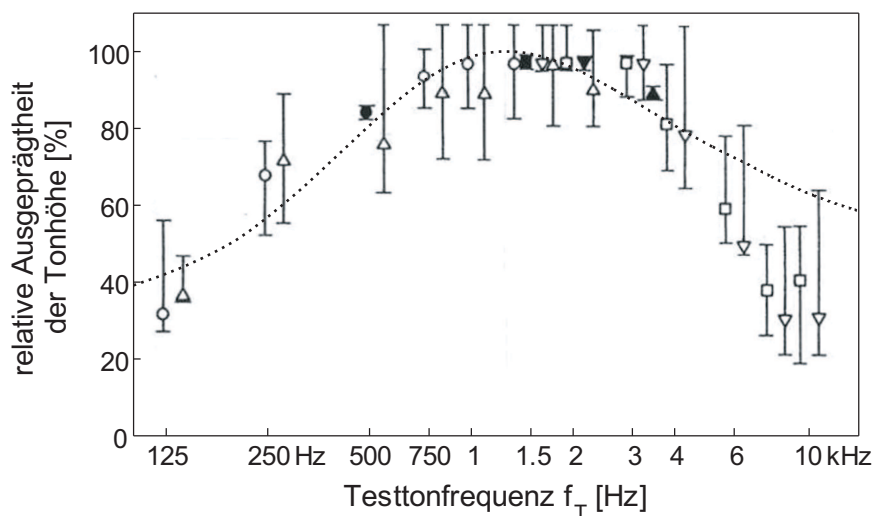


Abbildung 2.1: Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Frequenz eines Sinustons. Abbildung nach [Zwi99]. Gestrichelt eingezeichnet: Bandpass entsprechend Formel 2.1 mit den Parametern: $z_{ofs} = 6$, $z_c = 16$, $bp_{exp} = 3$.

Zur Beschreibung dieser Daten bietet sich also ein Bandpass an, wie er beispielsweise zur Näherung des spektralen Dominanzbereichs der Tonhöhenwahrnehmung verwendet wird [Plö67, Rit67]. Entgegen der Parametrierung des hierfür vorgeschlagenen Bandpasses bietet es sich jedoch an, einen Bandpass über der Tonheit anstelle der Frequenz anzusetzen, so daß sich für eine noch zu ermittelnde Mittentonheit z_c die Ausgeprägtheit der Tonhöhe eines Testtons der Tonheit z_T mit Hilfe eines Offsets z_{ofs} und eines Exponenten bp_{exp} über folgende Proportionalität beschreiben läßt:

$$AdT \sim \left[\frac{1}{\frac{z_T + z_{ofs}}{z_c} + \frac{z_c}{z_T + z_{ofs}}} \right]^{bp_{exp}} \quad (2.1)$$

2.1.2 Schalldauer

Die Abhängigkeit der Ausgeprägtheit der Tonhöhe eines Sinustons der Frequenz $f_T = 1$ kHz von dessen Dauer T_T ist in der linken Seite von Abbildung 2.2 dargestellt. Ab einer Schalldauer von $T_T > 300$ ms ist demnach die volle Ausgeprägtheit der Tonhöhe erreicht. Im logarithmischen Zeitmaßstab tritt für kürzere Dauern ein linearer Abfall der Ausgeprägtheit der Tonhöhe auf, vergleichbar mit der zeitlichen Integration der Lautheit [Zwi74, Buu96], was sich über einen Tiefpass geeigneter Dimensionierung beschreiben läßt [Zwi84]. Betrachtet man das Gehör als Netzwerk, das sich mit Hilfe elektrotechnischer Zusammenhänge beschreiben läßt, so entspricht dies der Einschwingzeit der verwendeten Filter. Für ein späteres Modell ist es also zweckmäßig, auf einem Modell der Wahrnehmungsgröße ‐Lautheit‐ aufzubauen, da so dieser Effekt der zeitlichen Integration bereits berücksichtigt wird.

2.1.3 Schallpegel

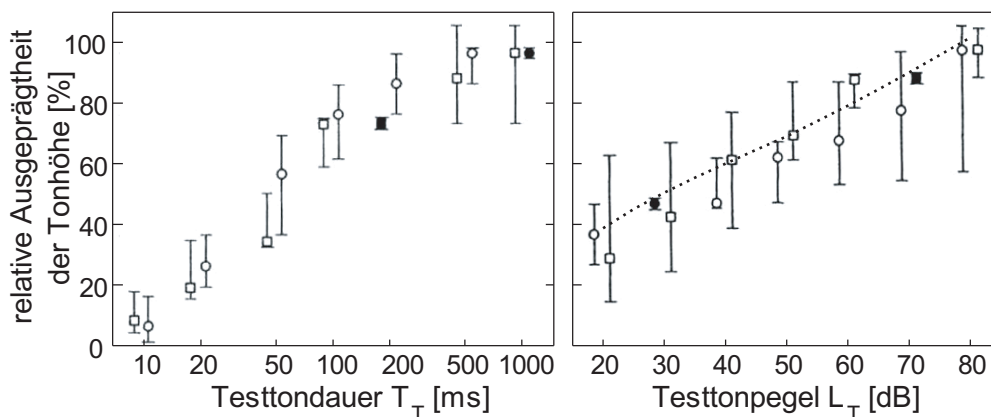


Abbildung 2.2: *Abhängigkeit der Ausprägtheit der Tonhöhe von der Dauer T_T (links) und vom Schallpegel L_T (rechts) eines Sinustons. Abbildungen nach [Zwi99]. Auf der rechten Seite gestrichelt eingezeichnet: Näherung entsprechend Formel 2.2.*

Wie Abbildung 2.2 (rechte Seite) zeigt, steigt die Ausprägtheit der Tonhöhe eines Sinustons im untersuchten Bereich in etwa linear mit dem Schallpegel L_T an. Vergleicht man diese Abhängigkeit mit der Lautheit eines Sinustons so zeigt sich für den untersuchten Pegelbereich $L_T > 20$ dB ein ähnliches Verhalten [Zwi58, Zwi59b]. Die Abhängigkeit der Ausprägtheit der Tonhöhe vom Schallpegel läßt sich so als eine Abhängigkeit von der Lautheit N_T des Testtons über

$$AdT \sim \frac{N_T^{0,13}}{[sone]} - 0,4 \quad (2.2)$$

beschreiben. Dies stützt die im vorangegangenen Abschnitt beschriebene These, zur Modellierung der Ausprägtheit der Tonhöhe von einem Modell der Lautheit auszugehen.

2.1.4 Bandbreite von Bandpassrauschen

Eine weitere grundlegende Abhängigkeit der Ausprägtheit der Tonhöhe läßt sich mit Hilfe der Bandbreite eines Schalls (Bandpassrauschen, $L = 50$ dB) darstellen. Wie in Abbildung 2.5 ersichtlich, nimmt die Ausprägtheit der Tonhöhe mit steigender Bandbreite Δf ab. Dies entspricht der in Kapitel 1 beschriebenen Eigenschaft, daß breitbandiges Rauschen keine Tonhöhe hervorruft und somit die Ausprägtheit der Tonhöhe für extreme Bandbreitenvergrößerung gegen Null gehen muß. Darüber hinaus läßt sich eine Abhängigkeit von der Mittenfrequenz f_c erkennen. Diese äußert sich darin, daß die Ausprägtheit der Tonhöhe für Bandpassrauschen niedriger Mittenfrequenz mit steigender Bandbreite schneller absinkt, als bei hohen Mittenfrequenzen.

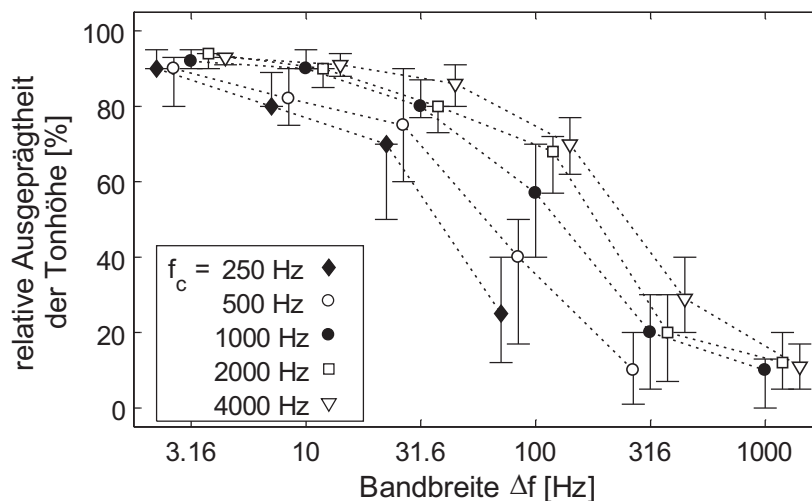


Abbildung 2.3: Relative Ausprägtheit der Tonhöhe von Bandpassrauschen der Mittenfrequenz f_c bei Variation der Bandbreite Δf . Abbildung nach [Zwi99].

Eine Erklärung dieses Effektes läßt sich über die in Abschnitt 1.3.2 beschriebene Bandbreite der Frequenzgruppen herleiten. Betrachtet man das Gehör als Filterbank bestehend aus Bandpassfiltern mit zur Frequenzgruppenbreite proportionalen Filterbandbreiten, so steigt diese Bandbreite ab einer Mittenfrequenz von ca. 500 Hz an. Ein Schall konstanter absoluter Bandbreite kann also bei höheren Mittenfrequenzen von einer solchen Filterbank besser erfasst bzw. aufgelöst werden und die Ausprägtheit der Tonhöhe ist größer für Schalle geringerer Bandbreite relativ zur aktuellen Frequenzgruppenbreite. Darüberhinaus ist aber auch für sehr geringe Bandbreiten ($\Delta f \ll \Delta f_G$) eine Abnahme der Ausprägtheit der Tonhöhe bei Bandbreitenerhöhung zu beobachten. Stochastische Analysen zeigen für Bandpassrauschen eine sogenannte Eigenmodulation f_{mod}^* , die sich über

$$f_{mod}^* = 0,64\Delta f \quad (2.3)$$

beschreiben läßt [Ric54, Fas75]. Somit läßt sich diese Abnahme der Ausprägtheit der Tonhöhe möglicherweise auf die wahrnehmbaren Amplitudenmodulationen und die damit verbundene Schwankungsstärke zurückführen.

2.2 Amplitudenmodulation

Bei amplitudenmodulierten Sinustönen (AM-SIN) müssen zwei Fälle unterschieden werden. Bei niedrigen Modulationsfrequenzen f_{mod} entspricht die wahrgenommene Tonhöhe derjenigen der Trägerfrequenz f_{Tr} . Hohe Modulationsfrequenzen sind dagegen selbst als virtuelle Tonhöhe wahrnehmbar. Der Übergang zwischen diesen Bereichen ist fließend und läßt sich mit Hilfe der Mithörschwellenmuster des Trägers beschreiben. Abbildung 2.4 zeigt die spektrale Zusammensetzung sowie die Mithörschwelle des Trägers eines amplitudenmodulierten Sinustons mit $f_{Tr} = 1000$ Hz für $f_{mod} = 25$ (links) und 250 Hz (rechts).

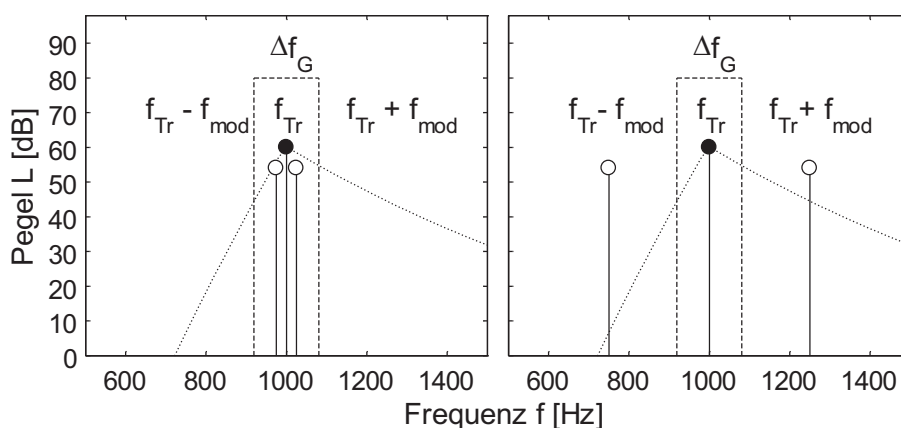


Abbildung 2.4: Mithörschwelle des Trägers $f_{Tr} = 1000$ Hz und Seitenlinien amplitudenmodulierter Sinustöne (Links: $f_{mod} = 25$ Hz, rechts: $f_{mod} = 250$ Hz).

Das Spektrum solch eines amplitudenmodulierten Tons besteht prinzipiell aus den drei Komponenten $f_{Tr} - f_{mod}$, f_{Tr} und $f_{Tr} + f_{mod}$. Die Pegeldifferenz des Trägers und der Seitenlinien bestimmt dabei die Modulationstiefe. Bei niedrigen Modulationsfrequenzen (Abbildung 2.4 links) werden die beiden Seitenlinien vom Träger maskiert und es ist lediglich eine Lautheitsschwankung des Trägers wahrnehmbar. Bei Erhöhung der Modulationsfrequenz treten die Seitenlinien aus der Mithörschwelle hervor und beginnen bezüglich der Tonhöhenwahrnehmung mit dem Träger zu konkurrieren, was als Rauigkeit wahrnehmbar ist. Bei Terhardt wird hierfür eine Grenze der Modulationsfrequenz von $f_{mod} \approx 25$ Hz angegeben [Ter68b]. Liegen die Seitenlinien schließlich außerhalb der Frequenzgruppe des Trägers ($f_{mod} > \Delta f_G/2$), so nimmt die Rauigkeit ab und die Modulationsfrequenz wird als virtuelle Tonhöhe wahrgenommen.

Im Folgenden wurden zunächst langsame Modulationen untersucht, bei denen lediglich Schwankungsstärke und Rauigkeit auftreten, jedoch keine virtuelle Tonhöhe wahrnehmbar ist. Die gefundenen Ergebnisse tragen auch zum Verständnis der Daten der Ausgeprägtheit der Tonhöhe sehr schmalbandiger Rauschen aus Abschnitt 2.1.4 bei.

Die Abhängigkeit der Ausgeprägtheit der Tonhöhe von Modulationsfrequenzen bis 32 Hz ist in Abbildung 2.5 dargestellt. Für eine steigende Modulationsfrequenz nimmt die Ausgeprägtheit der Tonhöhe ab, wobei sich eine leichte Abhängigkeit von der gewählten Grundfrequenz f_0 ($= f_{Tr}$) zeigt. So ist die Abnahme bei einem SIN-AM der Grundfrequenz $f_0 = 2000$ Hz etwas geringer als für eine Grundfrequenz $f_0 = 500$ bzw. 1000 Hz.

Für Modulationsfrequenzen $2 \leq f_{mod} \leq 8$ Hz ist ein leichtes Plateau auszumachen. Insbesondere für eine Grundfrequenz von 1000 Hz (Dreiecke) zeigt sich für eine Modulationsfrequenz von 4 Hz sogar ein leicht erhöhter Wert der Ausgeprägtheit der Tonhöhe gegenüber

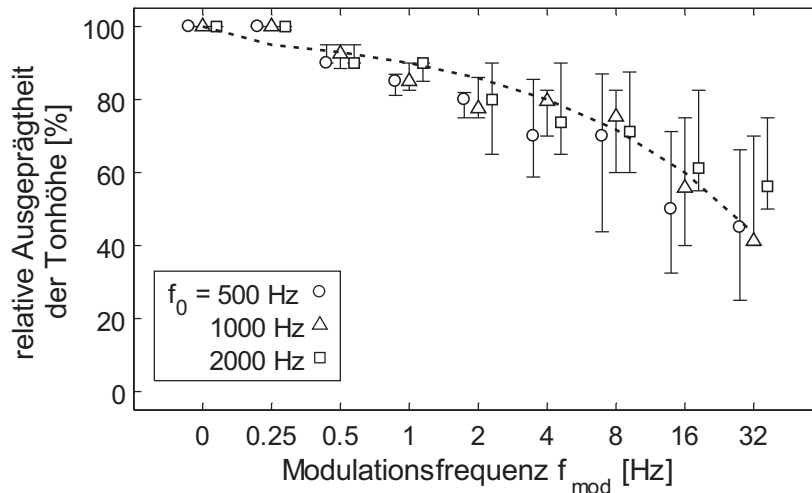


Abbildung 2.5: *Abhängigkeit der Ausgeprägtheit der Tonhöhe amplitudenmodulierter Sinustöne verschiedener Grundfrequenz von ihrer Modulationsfrequenz f_{mod} . Darbietungspegel $L_T = 70$ dB, Modulationsgrad 100 %, Darbietungsdauer $T_T = 1$ s. Anker: unmodulierter Sinuston. Punktiert eingezeichnet: Näherung entsprechend Formel 2.4*

den Ergebnissen bei Modulationsfrequenzen von 2 und 8 Hz. Diese Eigenschaft korreliert mit dem bei einer Modulationsfrequenz von 4 Hz zu findenden Maximum der Empfindungsgröße “Schwankungsstärke“, die in diesen Fällen möglicherweise für eine deutlichere Wahrnehmbarkeit der Tonhöhe sorgt. Da dieser Effekt nur bei zwei Versuchspersonen auftrat, sind die zugehörigen Daten in Anhang B dargestellt.

Die Tatsache, daß die Versuchsdaten bei einer Modulationsfrequenz von 0,25 Hz mit denen des Referenzschalls übereinstimmen, ist darauf zurückzuführen, daß die Schalldauer lediglich 1 Sekunde betrug, sodaß hier die Schwankung wesentlich langsamer erfolgt, als es bei dieser Schalldauer wahrnehmbar ist.

Für die in Abbildung 2.5 dargestellten Daten läßt sich der Abfall der Ausgeprägtheit der Tonhöhe im untersuchten Bereich der Modulationsfrequenz über

$$AdT \sim 100 - 10\sqrt{f_{mod}} \quad (2.4)$$

beschreiben. Dabei sind jedoch Einflüsse aufgrund der physikalischen und der wahrgenommenen Modulationstiefe nicht berücksichtigt, alle untersuchten Schalle besaßen einen mathematischen Modulationsgrad m von 100 %. In einem Modell, das eine der Schwankungsstärke ähnliche Größe zur Modellierung dieser Effekte heranzieht könnte die unterschiedliche Modulationstiefe ebenfalls Berücksichtigung finden.

Wird die Modulationsfrequenz einer Amplitudenmodulation hinreichend weit erhöht, so wird diese als eigenständige Tonhöhe wahrgenommen ($f_0 = f_{mod}$). Je nach spektraler Zusammensetzung kann es sich hierbei um eine spektrale oder virtuelle Tonhöhe handeln.

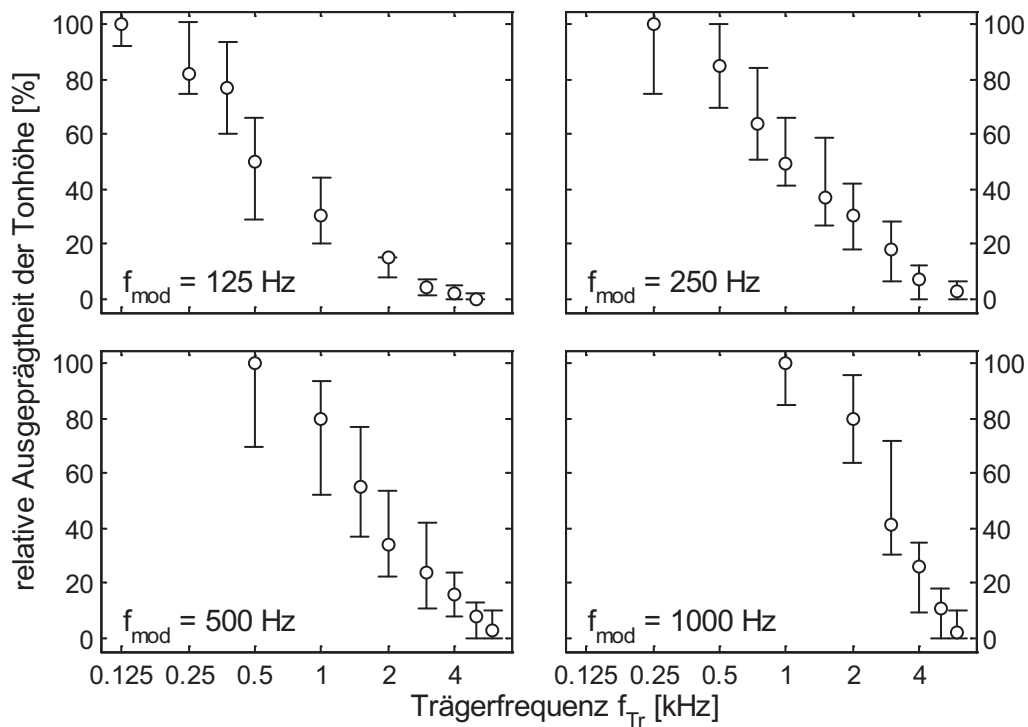


Abbildung 2.6: Relative Ausgeprägtheit der virtuellen Tonhöhe von amplitudenmodulierten Sinustönen variierender Trägerfrequenz f_{Tr} (nach [Fas90]). Es gilt $f_{mod} \leq f_0$, die Modulationsfrequenz f_{mod} ist als virtuelle Tonhöhe wahrnehmbar.

Ergebnisse für Modulationsfrequenzen $f_{mod} = 125$ bis 1000 Hz sind in Abbildung 2.6 dargestellt [Fas90]. Hier wird die Ausgeprägtheit der virtuellen Tonhöhe betrachtet, die durch die Modulation eines Trägers, mit der als Tonhöhe wahrnehmbaren Frequenz, entsteht. Die Entstehung dieser virtuellen Tonhöhen erfolgt analog zur virtuellen Tonhöhe von harmonisch komplexen Tönen. Diese amplitudenmodulierten Sinustöne lassen sich also als komplexe Töne, bestehend aus drei Komponenten, beschreiben.

Da die Trägerfrequenz der untersuchten Schalle variiert wird, verschiebt sich die spektrale Lage der Testschalle im Existenzbereich der virtuellen Tonhöhe aus Abbildung 1.1 immer weiter zu hohen Frequenzen hin. Dies führt zu einer schwächeren Ausbildung der virtuellen Tonhöhe und somit geringerer Ausgeprägtheit der virtuellen Tonhöhe.

2.3 Vergleich der Ausgeprägtheit der Tonhöhe verschiedener Schallarten

Das Beispiel von Klavierton und Tiefpassrauschen aus Kapitel 1.2 zeigt, daß sich die Ausgeprägtheit der Tonhöhe verschiedener Arten von Schallen bei ähnlicher Tonhöhe deutlich unterscheiden kann. Einen Überblick dazu bietet [Fas79c] für synthetische Signale und

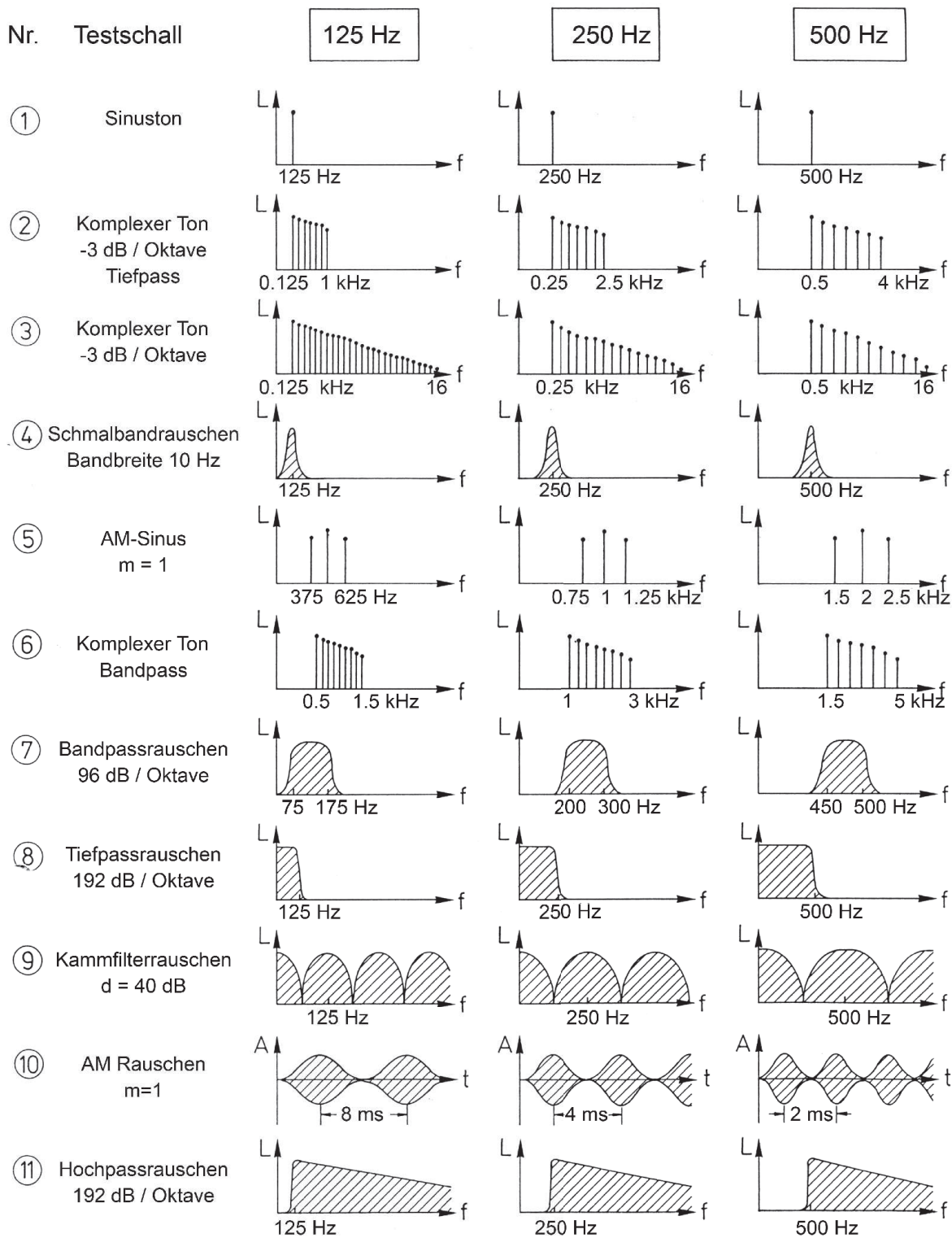


Abbildung 2.7: Qualitative Darstellung der von Fastl und Stoll untersuchten Schalle und deren Beurteilung bezüglich ihrer Ausprägtheit der Tonhöhe. Innerhalb einer Tabellenspalte sinkt die wahrgenommene Ausprägtheit der Tonhöhe jeweils von Oben nach Unten. Abbildung nach [Fas79c]

[Rak02] in Bezug auf musikalische Klänge. In Abbildung 2.7 sind die untersuchten synthetischen Schalle nach [Fas79c] tabellarisch angeordnet. In jeder Spalte der Tabelle sinkt die Ausgeprägtheit der Tonhöhe von oben nach unten. Vergleiche zwischen Schallen verschiedener Tonhöhe wurden nicht durchgeführt, weshalb die Anordnung der Spalten von links nach rechts lediglich nach aufsteigender Frequenz erfolgt, nicht aber in Abhängigkeit der Ausgeprägtheit der Tonhöhe zu verstehen ist.

Die in [Fas79c] dargestellten Streubereiche der Versuchsergebnisse sind relativ groß, weshalb insbesondere die Positionierung des Schalls 4 “Schmalbandrauschen“ in der entstandenen Reihenfolge nochmals untersucht werden soll. Mit der Methode “Random Access“ wurde ein Versuch zur Anordnung der Schalle nach [Fas79c] entlang einer Skala von niedriger zu hoher Ausgeprägtheit der Tonhöhe durchgeführt um die Rangfolge bezüglich ihrer Ausgeprägtheit der Tonhöhe.

Aus den Daten in Abbildung 2.8 läßt sich erkennen, daß die Streuungen der Positionsangaben teilweise auch hier recht groß ausfallen. Für den erwähnten Schall 4 reichen die Interquartilbereiche von Position 5 bis zu Position 2. Eine Erklärung weshalb manche Versuchspersonen die Tonhöhe harmonischer Töne als ausgeprägter bewerten als die Tonhöhe von Schmalbandrauschen läßt sich möglicherweise über das sogenannte analytische und synthetische Hören ableiten.

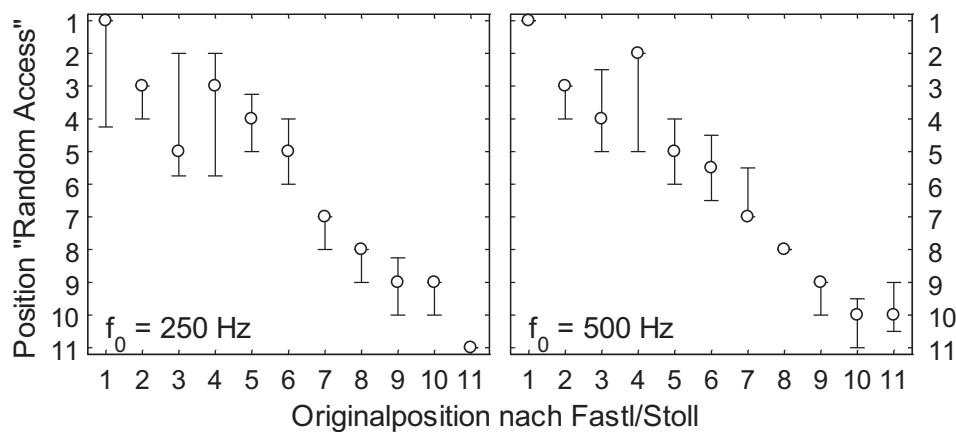


Abbildung 2.8: Vergleich der Positionen bei Random Access mit den Originalexperimenten, Links: Grundfrequenz 250 Hz, Rechts: 500 Hz

Als “analytischer Hörer“ wird dabei bezeichnet, wer bei der Wahrnehmung eines Schalls stärker auf die tatsächlich vorhandenen Spektralkomponenten achtet, also den Schall bei der Wahrnehmung “analysiert“. Für einen “synthetischen Hörer“ ist dagegen die Synthese der Schallkomponenten vorrangig, die verschiedenen Wahrnehmungsgrößen werden also aus der Verschmelzung der einzelnen Komponenten gebildet.

Mit Hilfe eines Experiments, entsprechend [Smo70], wurden die beteiligten Versuchspersonen in die beiden Gruppen analytische und synthetische Hörer unterteilt. Dabei werden jeweils Zweitonkomplexe nach ihrer absoluten Tonhöhe bewertet. Die Schallpaare bestehen dabei aus Schallen A und B ähnlicher spektraler Lage aber unterschiedlicher Grundfrequenz. Die Zusammensetzung der Signale ist Tabelle 2.3 zu entnehmen. Beispielsweise setzt sich das Schallpaar Nr. 1 aus einem Zweitonkomplex der Grundfrequenz $f_{0,A} = 250$ Hz bestehend aus der 3. und 4. Harmonischen (also 750 und 1000 Hz) gefolgt von einem Zweitonkomplex der Grundfrequenz $f_{0,B} = 200$ Hz, 4. und 5. Harmonische (800 und 1000 Hz), zusammen.

Signalnummer	Grundfrequenz $f_{0,A}$	Komponenten Schall A	Grundfrequenz $f_{0,B}$	Komponenten Schall B
1	250 Hz	750 + 1000 Hz	200 Hz	800 + 1000 Hz
2	250 Hz	1750 + 2000 Hz	200 Hz	1800 + 2000 Hz
3	500 Hz	1500 + 2000 Hz	400 Hz	1600 + 2000 Hz
4	500 Hz	3500 + 4000 Hz	400 Hz	3600 + 4000 Hz

Tabelle 2.1: Schallzusammensetzung der untersuchten Signale zur Zuordnung der Versuchspersonen zur Gruppe analytischer oder synthetischer Hörer. Darbietungspegel $L_T = 60$ dB, Schalldauer $T_T = 3$ s.

Die Signale sind also so gewählt daß innerhalb eines Signals bestehend aus Schall A und Schall B die spektrale Lage nach oben verschoben wird, die virtuelle Tonhöhe jedoch sinkt. Beurteilt nun eine Versuchsperson die Tonhöhe von Schall B innerhalb eines Paares als höher, so überwiegt der Einfluss der spektralen Lage auf die Tonhöhenwahrnehmung, es handelt sich also um einen analytischen Hörer. Umgekehrt handelt es sich um einen synthetischen Hörer wenn die Veränderung der virtuellen Tonhöhe die Wahrnehmung dominiert. Die Darbietungsreihenfolge innerhalb eines Schallpaares wird dabei zufällig gewählt, so daß Gewöhnungseffekte vermieden werden.

Entsprechend Abbildung 2.9 sind ca. 70 % der untersuchten Probanden der Gruppe “analytische Hörer“ zuzuordnen. Im Falle des Signalpaares 4 ist die Anzahl analytischer und synthetischer Hörer in Etwa gleich. Entsprechend Tabelle 2.3 handelt es sich dabei um Sinuspaare bestehend aus 3500/4000 Hz und 3600/4000 Hz, also Signale die am Rand des Existenzbereichs der virtuellen Tonhöhe liegen. Obwohl hier die virtuelle Tonhöhe nur noch sehr schwach ausgebildet ist, wird bei diesem Schallpaar der Anteil analytischer Hörer geringer als für die übrigen Schallkonfigurationen.

Die rechte Seite von Abbildung 2.9 zeigt die Versuchsergebnisse für analytische und synthetische Hörer speziell für das Signal 4 aus der ursprünglichen Untersuchung zu unterschiedlichen Schallarten. Die Tonhöhe von Schmalbandrauschen wird von synthetischen Hörern offenbar als ausgeprägter wahrgenommen als von analytischen Hörern, was sich in

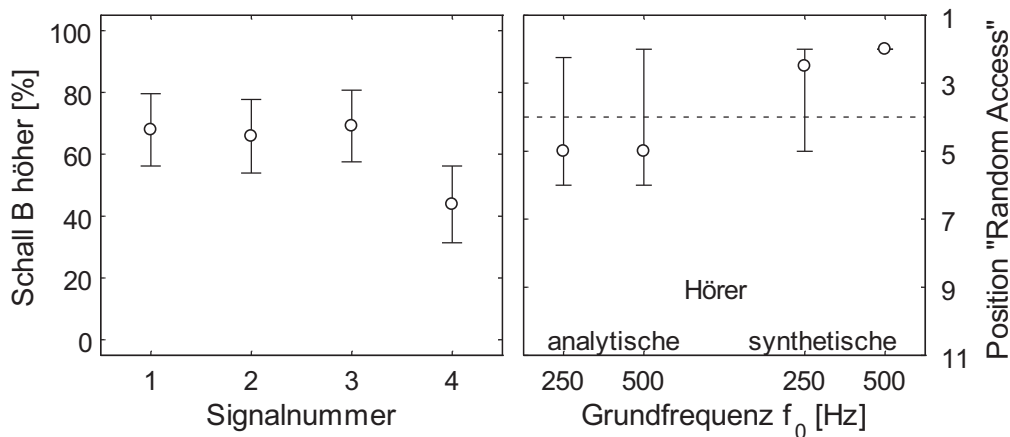


Abbildung 2.9: *Prozentsatz analytischer und synthetischer Hörer (Links) bei 26 Versuchspersonen. Rechts: Ergebnisse des Versuchs "Random Access" für das Signal 4 (Schmalbandrauschen, Rechts) für jeweils 13 Versuchspersonen getrennt nach analytischen und synthetischen Hörern.*

einer niedrigeren Positionsnummer widerspiegelt. Umgekehrt rufen harmonisch komplexe Töne bei synthetischen Hörern schwächer ausgeprägte Tonhöhen hervor. Grund hierfür könnte die "globalere" Hörweise sein, also die Tatsache, daß einzelne Spektralkomponenten (z.B. die Grundfrequenz) weniger stark zum Höreindruck beitragen als das Zusammenspiel aller Komponenten. Vielmehr ist bei synthetischen Hörern das Zusammenwirken aller Komponenten entscheidend und da Schmalbandrauschen nur aus einer einzelnen spektralen Komponente entsprechender Bandbreite besteht, wird diese durch keine weiteren Komponenten "gestört" und besitzt so subjektiv eine stärkere Ausgeprägtheit der Tonhöhe als harmonisch komplexe Töne, bei denen zusätzlich zur Grundfrequenz weitere Komponenten enthalten sind.

2.4 Ausgeprägtheit der Tonhöhe von harmonisch komplexen Tönen (HKT)

Harmonisch komplexe Töne (HKT) bestehen aus einer Reihe zueinander harmonischer Teiltöne, wobei die wahrgenommene Tonhöhe eines solchen Schalls weitestgehend mit der Grundfrequenz übereinstimmt. Die Grundfrequenz wird dabei als erste Harmonische bezeichnet, entsprechend wird die Komponente mit der n-fachen Frequenz als n-te Harmonische bezeichnet.

Die Abhängigkeit der Ausgeprägtheit der Tonhöhe harmonisch komplexer Töne von der Anzahl enthaltener Harmonischer n_h wurde bereits für Schalle mit konstantem Darbietungspegel untersucht [Fas97]. Hier werden nun Ergebnisse vorgestellt, die für Schalle gleicher spektraler Zusammensetzung, aber konstanter Darbietungslautheit erzielt wur-

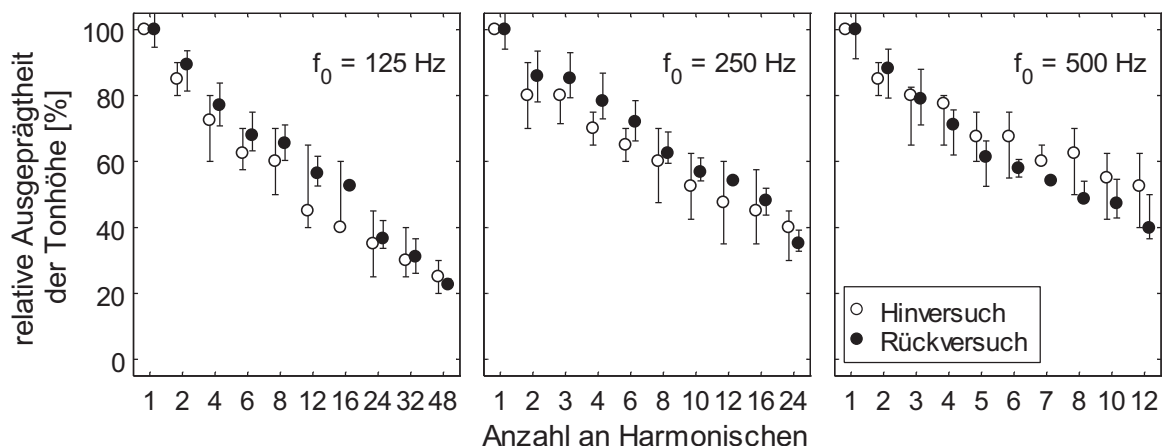


Abbildung 2.10: Ausprägtheit der Tonhöhe von Harmonisch komplexen Tönen mit variierender Anzahl Harmonischer, Versuchsergebnisse für Schalle gleicher Lautheit ($N_T = 5$ sone, $T_T = 0,7$ s). Nicht gefüllte Symbole: Anker Sinuston bei f_0 , Schwarz ausgefüllte Symbole: Anker harmonisch komplexer Ton mit $n_h = 16, 12$ bzw. 7 Harmonischen.

den. Durch die Lautheitsnormierung auf 5 sone mit Hilfe des Dynamic Loudness Models (DLM, [Cha02b]) können mögliche Einflüsse unterschiedlicher Lautheit auf die Ausprägtheit der Tonhöhe minimiert werden.

Abbildung 2.10 zeigt die Ergebnisse für die Experimente mit vollständigen harmonisch komplexen Tönen. Die Testschalle wurden dabei im Hinversuch jeweils mit dem Sinuston der zugehörigen Grundfrequenz verglichen. Im Rückversuch diente je nach Grundfrequenz ein anderer harmonisch komplexer Ton als Ankerschall (125 Hz: 16 Harmonische, 250 Hz: 12 Harmonische, 500 Hz: 7 Harmonische).

Ähnlich den Ergebnissen in [Fas97] sinkt die Ausprägtheit der Tonhöhe mit steigender Anzahl Harmonischer im Vergleich zum Sinus. Unabhängig von der Grundfrequenz sinkt die Ausprägtheit der Tonhöhe für eine gleiche Anzahl an Harmonischen auf ähnliche Werte ab. Für alle drei untersuchten Grundfrequenzen liegt die relative Ausprägtheit der Tonhöhe für eine Anzahl von 12 Harmonischen bei ca. 50 % der Ausprägtheit der Tonhöhe des Sinustons der Frequenz f_0 . Bei gleicher oberer Grenzfrequenz von 6 kHz (48, 24 bzw. 12 Harmonische bei Grundfrequenzen f_0 von 125, 250 und 500 Hz) fällt die Ausprägtheit der Tonhöhe mit steigender Grundfrequenz dagegen weniger stark ab.

Dieser Effekt ist auf die Breite der Frequenzgruppen bei den jeweiligen Grundfrequenzen zurückzuführen. Diese beträgt für alle drei Grundfrequenzen in Etwa 100 Hz, so daß die ersten beiden Harmonischen jeweils in getrennten Frequenzgruppen liegen. Bei einer Grundfrequenz von 125 Hz befinden sich bereits ab einer Tonheit von ca. 6 Bark benachbarte Harmonische innerhalb einer Frequenzgruppe, deren Breite hier ca. 140 Hz beträgt.

Bei einer Grundfrequenz von 500 Hz ist dies erst ab ca. 15 Bark der Fall ($\Delta f_G \approx 400$ Hz), so daß hier die einzelnen Teiltöne besser vom Gehör getrennt und analysiert werden können. So beeinträchtigen sie die Ausgeprägtheit der Tonhöhe weniger stark.

Wie in Abschnitt 2.1.4 beschrieben, besitzt die Breite der durch einen Schall hervorgerufenen Erregung der Basilarmembran, und damit in der Analogie einer Filterbank die Anzahl angeregter Filter, maßgeblichen Einfluß auf die Ausgeprägtheit der Tonhöhe. Schallkomponenten zusätzlich zur Grundfrequenz lassen sich als Störungen der Tonhöhe und somit als negativer Einfluß auf die Ausgeprägtheit der Tonhöhe interpretieren. Dabei ist dieser hemmende Einfluß größer, je mehr zusätzliche Komponenten enthalten sind.

Lassen sich die einzelnen Harmonischen vom Gehör bis zu einer höheren Tonheit getrennt auflösen, so ist die hemmende Wirkung auf die Ausgeprägtheit der Tonhöhe geringer, was sich in der geringeren Abnahme der Ausgeprägtheit der Tonhöhe für harmonisch komplexe Töne der Grundfrequenz $f_0 = 500$ Hz im Vergleich zu $f_0 = 125$ Hz bei identischer oberer Grenzfrequenz zeigt. Eine ähnliche Gesamtzahl an Harmonischen führt dagegen bei unterschiedlichen Grundfrequenzen zu einem vergleichbaren Rückgang der Ausgeprägtheit der Tonhöhe in Relation zum jeweiligen Ankerschall.

2.5 Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Residualtöne

Fehlt die Grundfrequenz oder mehrere untere Harmonische eines harmonisch komplexen Tons, so entsteht ein harmonisch komplexer Residualton (HKRT), dessen Tonhöhe als virtuelle Tonhöhe entsprechend Kapitel 1.1 wahrgenommen wird. Diese virtuelle Tonhöhe läßt sich ebenso wie spektrale Tonhöhen entsprechend ihrer Ausgeprägtheit der Tonhöhe skalieren.

2.5.1 Einfluß der Anzahl entfernter Harmonischer eines harmonisch komplexen Residualtons

Abbildung 2.11 zeigt die Ergebnisse eines Hörversuchs, bei dem eine unterschiedliche Anzahl unterer Harmonischer von einem harmonisch komplexen Ton entfernt wurde. Die Darbietungslautheit wurde dabei wiederum auf 5 sone eingestellt, die verwendeten Testschalle besaßen einen spektralen Pegelabfall von -3 dB/Oktave. Nicht gefüllte Symbole zeigen die Ergebnisse des Hinversuchs, bei dem der vollständige harmonisch Komplexe Ton als Anker diente, im Rückversuch (gefüllten Symbole) wurde der harmonisch komplexe Residualton mit vier entfernten unteren Harmonischen als Ankerschall dargeboten.

Es zeigt sich einerseits eine deutliche Abnahme der Ausgeprägtheit der Tonhöhe mit steigender Nummer der untersten enthaltenen Harmonischen, andererseits zeigt sich auch eine Abhängigkeit dieses Verhaltens von der Grundfrequenz. Beim Hinversuch ist mit steigender Grundfrequenz eine stärkere Abnahme der Ausgeprägtheit der Tonhöhe zu beobach-

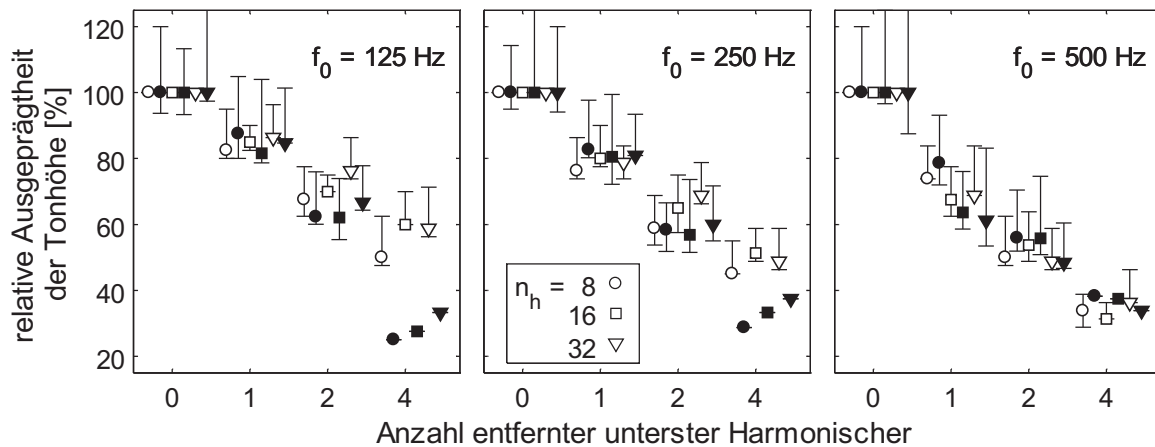


Abbildung 2.11: Ausprägtheit der Tonhöhe von harmonisch komplexen Tönen variierender Anzahl entfernter unterer Harmonischer ($N_T = 5$ sone, $T_T = 0,7$ s). Nicht gefüllte Symbole: Anker harmonisch komplexer Ton, Schwarz gefüllte Symbole: Anker harmonisch komplexer Residualton mit 4 entfernten Harmonischen.

ten. Ursache hierfür ist die spektrale Lage der Schalle im Existenzbereich der virtuellen Tonhöhe nach Abbildung 1.1, die sich mit wachsender Grundfrequenz immer weiter zu hohen Frequenzen hin verschiebt. Analog zu den Ergebnissen für amplitudenmodulierte Sinustöne aus Abschnitt 2.2 sinkt die Ausprägtheit der virtuellen Tonhöhe bei Verschiebung der Spektralkomponenten in Richtung der oberen Grenze des Existenzbereichs.

Die verschiedenen Symbole in Abbildung 2.11 repräsentieren Schalle mit unterschiedlicher nomineller Anzahl Harmonischer bzw. unterschiedlicher Nummer der höchsten enthaltenen Harmonischen. Diese Größe zeigt bei den untersuchten Schallen einen sehr schwachen Einfluß auf die Ausprägtheit der Tonhöhe. Für Schalle geringerer Anzahl Harmonischer sinkt die Ausprägtheit der Tonhöhe etwas stärker ab als bei einer höheren Anzahl. Dies spricht dafür, daß höhere Harmonische nur wenig zur Bildung der virtuellen Tonhöhe beitragen, wie dies auch in [Moo85] dargestellt wird. Ebenso ergeben sich daraus Auswirkungen auf die vorangehende Feststellung, daß zusätzliche Komponenten als Störungen der Tonhöhenwahrnehmung interpretiert werden können. Mit zunehmendem Abstand von der Grundfrequenz wird dieser störende Einfluß geringer.

2.5.2 Spektrale Dämpfung harmonisch komplexer Töne

Ein weiteres Experiment soll nun klären wie die Ausprägtheit der Tonhöhe von Schallen identischer spektraler Zusammensetzung, jedoch unterschiedlicher spektraler Pegeldämpfung ΔL_f bewertet wird. Der Pegel L_n der n -ten Harmonischen ergibt sich aus dem Pegel der ersten Harmonischen L_1 als

$$L_n = L_1 - \log_2(n) \cdot \frac{\Delta L_f}{\text{Oktave}} \quad [\text{dB}]. \quad (2.5)$$

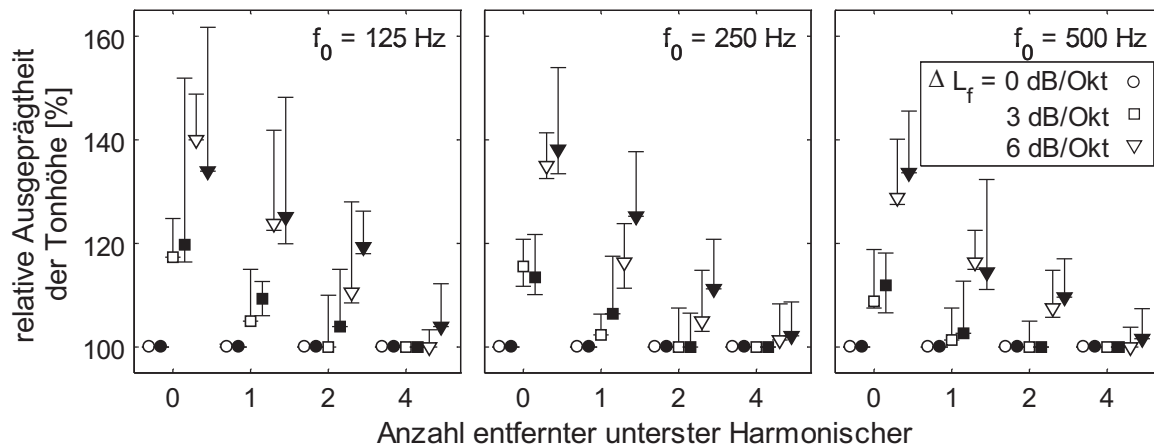


Abbildung 2.12: Relative Ausgeprägtheit der Tonhöhe von Harmonisch komplexen Tönen variierender Anzahl entfernter unterster Harmonischer mit verschiedenem spektralem Pegelabfall ($N_T = 5$ sone, $T_T = 0,7$ s). Nicht gefüllte Symbole: Anker 0 dB/Oktave, Schwarz gefüllte Symbole: Anker- und Testschall vertauscht.

Die Daten für drei verschiedene Grundfrequenzen f_0 sind in Abbildung 2.12 zu sehen. Die untersuchten harmonisch komplexen Töne besaßen jeweils 16 Harmonische von denen unterste Teiltöne bis zur 4. Harmonischen entfernt wurden. Im Hinversuch (nicht gefüllte Symbole) diente der Schall mit einem Pegelabfall von 0 dB/Oktave (gleicher Pegel aller Harmonischer) als Ankerschall. Im Rückversuch (schwarze Symbole) wurde die Reihenfolge aus Anker- und Testschall vertauscht. Die Urteile der Versuchsperson wurden über

$$AdT = 100 \cdot \frac{100}{x} \quad , x = \text{aktuelles Urteil der Versuchsperson} \quad (2.6)$$

auf den Testschall normiert. So wurde im Rückversuch also lediglich die Darbietungsreihenfolge aus Anker- und Testschall vertauscht. In der Darstellung treten aus diesem Grund sowohl für die Daten des Hin- als auch des Rückversuchs keine Interquartilbereiche für die Schalle ohne spektralen Pegelabfall auf.

Die Ergebnisse in Abbildung 2.12 zeigen für harmonisch komplexe Töne (0 entfernte unterste Harmonische) eine starke Abhängigkeit der Ausgeprägtheit der Tonhöhe vom spektralen Pegelabfall. So besitzen diese bei einem spektralen Pegelabfall von -6 dB/Oktave eine um bis zu 40 % größere Ausgeprägtheit der Tonhöhe als der Ankerschall mit einem Pegelabfall von $\Delta L_f = 0$ dB/Oktave. Wie in Abschnitt 2.5.1 beschrieben, haben hauptsächlich die untersten Harmonischen einen Einfluß auf die Bildung der Tonhöhe. Aufgrund gleicher Gesamtlautheit der untersuchten Schalle sind die Pegel dieser Komponenten deutlich höher, wenn ein spektraler Pegelabfall vorhanden ist und es tritt ein Effekt der Pegelabhängigkeit, ähnlich Abschnitt 2.1.3, ein. Die höheren Harmonischen besitzen bei größerem spektralem Pegelabfall pro Oktave einen gegenüber der Grundfrequenz deutlich verringerten Pegel und können so weniger zum “störenden“ Anteil beitragen (siehe 2.4). Die Ausgeprägtheit der Tonhöhe wird so weniger stark geschwächt.

Die Unterschiede der Ausprägtheit der Tonhöhe zwischen harmonisch komplexen Tönen verschiedener spektraler Dämpfung nehmen mit zunehmender Anzahl entfernter unterster Harmonischer ab und sind bei vier entfernten untersten Harmonischen kaum noch vorhanden. Aufgrund der Ergebnisse in Abbildung 2.11 ist bekannt, daß die Ausprägtheit der virtuellen Tonhöhe dieser Schalle insgesamt sehr gering ist. Die Tatsache, daß sich unterschiedliche spektrale Pegeldämpfungen kaum in unterschiedlichen Bewertungen der Ausprägtheit der Tonhöhe auswirken, ist auf die Tatsache zurück zu führen, daß der Pegelabfall je Oktave konstant gehalten wird und so bei höheren Harmonischen die Pegelunterschiede geringer ausfallen. Diese Signale unterscheiden sich also nur noch wenig in ihrer Zusammensetzung und rufen so auch eine sehr ähnliche Wahrnehmung der Ausprägtheit der Tonhöhe hervor.

2.6 Ausprägtheit der Tonhöhe beim Übergang von spektraler zu virtueller Tonhöhenwahrnehmung

Voranehend wurden Abhängigkeiten der Ausprägtheit der Tonhöhe sowohl für spektrale als auch für virtuelle Tonhöhen dargestellt. Für eine spätere Modellierung ist es sinnvoll, darüberhinaus den Übergang zwischen den Extremfällen "spektrale Tonhöhe dominant" und "virtuelle Tonhöhe dominant" zu untersuchen. Hierfür werden im Folgenden die Ergebnisse eines Experimentes vorgestellt, bei dem der Pegel der ersten Harmonischen eines harmonisch komplexen Tons unterschiedlich stark abgesenkt wurde [Fru05b]. Mit Hilfe der Größenschätzung mit Ankerschall wurden Schalle einer Lautheit von 5 sone dargeboten und beurteilt. Die vollständigen Ergebnisse für Schalle mit Grundfrequenzen 125, 250 und 500 Hz sind in Anhang B dargestellt, Abbildung 2.13 zeigt die Resultate für die Grundfrequenzen 125 und 500 Hz sowie spektrale Dämpfungen von 0 und 6 dB/Oktave.

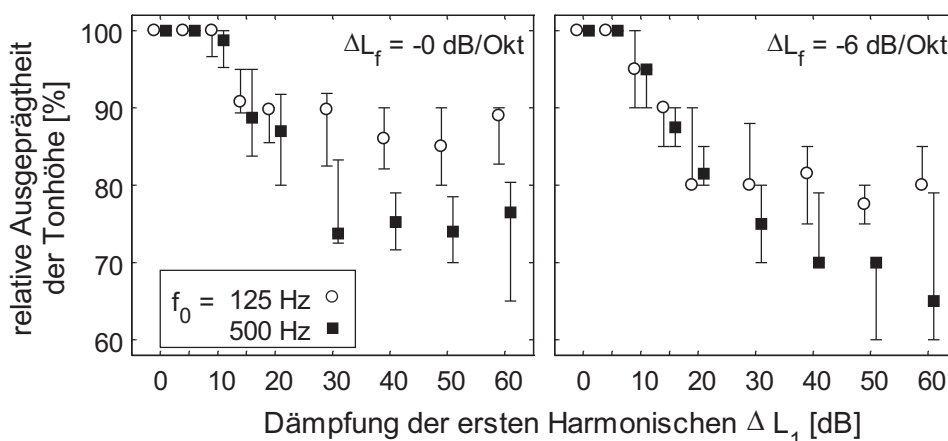


Abbildung 2.13: Ausprägtheit der Tonhöhe von harmonisch komplexen Tönen mit variierendem Pegel der ersten Harmonischen bei einer Lautheit N_T von 5 sone. Nicht gefüllte Symbole: $f_0 = 125$ Hz, Schwarz gefüllte Symbole: $f_0 = 500$ Hz.

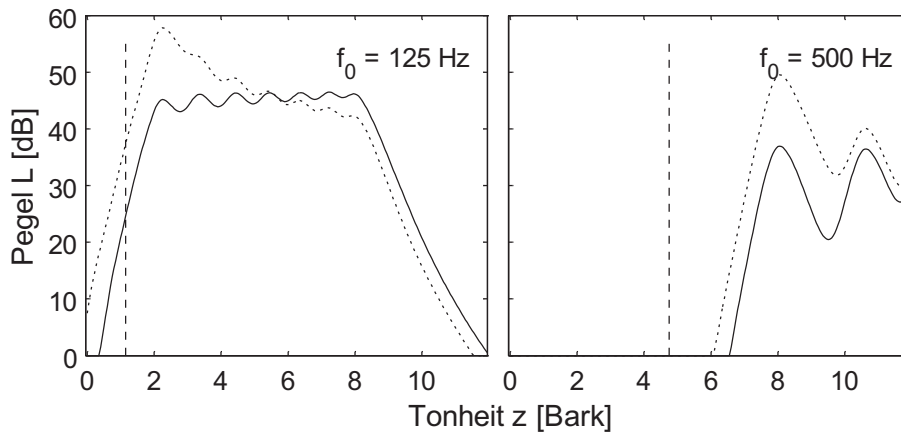


Abbildung 2.14: Mithörschwellen der spektralen Komponenten harmonisch komplexer Residualtöne. Durchgezogen: spektrale Pegeldämpfung $\Delta L_f = 0$ dB/Oktave, punktiert: -6 dB/Oktave. Die senkrechte gestrichelte Linie markiert jeweils f_0 .

Für eine Grundfrequenz von 125 Hz tritt für Schalle ohne spektrale Pegeldämpfung ab einer Dämpfung von f_0 von $\Delta L_1 \approx 15$ -20 dB eine Sättigung des Rückgangs der Ausgeprägtheit der Tonhöhe bei ca. 85 % der Ausgeprägtheit der Tonhöhe des Ankerschalls ein. Bei einer spektralen Dämpfung von -6 dB/Oktave tritt dieser Effekt erst ab Dämpfungen $\Delta L_1 > 20$ dB, und mit ca. 80 % auf einem etwas niedrigerem Niveau, als vorangehend beschrieben, auf. Bei $f_0 = 500$ Hz tritt die Sättigung jeweils erst bei um ca. 10 dB höherer Dämpfung ein und nimmt Werte von 70 % (-0 dB/Oktave) bzw. 75 % (-6 dB/Oktave) ein.

Für $f_0 = 125$ Hz ist dieses Verhalten näherungsweise in Einklang mit den in Abbildung 2.14 dargestellten Mithörschwellen (MHS) harmonisch komplexer Töne. Da die untere Flanke der MHS eine nahezu frequenz- und pegelunabhängige Steigung von -27 dB/Bark besitzt, ist die MHS der zweiten Harmonischen an der Stelle der Grundfrequenz um ca. 30 dB abgesunken, im Fall einer spektralen Dämpfung von $\Delta L_f = -6$ dB/Oktave ist diese Differenz entsprechend größer. So läßt sich erklären, daß die Sättigung der Ausgeprägtheit der Tonhöhe erst für höhere Dämpfungen von f_0 eintritt. Die nötige Dämpfung bei Erreichen der Sättigung ist aber im Fall der Grundfrequenz von 500 Hz nicht über die MHS-Muster zu erklären. Vielmehr ist hier die MHS bei f_0 bereits unter die Ruhehörschwelle (RHS) abgesunken, die virtuelle Tonhöhe wird also dominant, obwohl die spektrale Komponente nicht aufgrund von Maskierungseigenschaften verdeckt wird.

2.7 Durch Kammfiltereffekte hervorgerufene Tonhöhen

Gerade in der Aufnahme und Wiedergabe von Musik im Bereich von Tonstudios und Live-Konzerten spielen Kammfiltereffekte, hervorgerufen durch Laufzeitunterschiede, eine entscheidende Rolle für die resultierende Aufnahmequalität. Diese durch akustische, elektrische oder digitale Verzögerungen hervorgerufenen Klangcolorierungen sind in der Regel

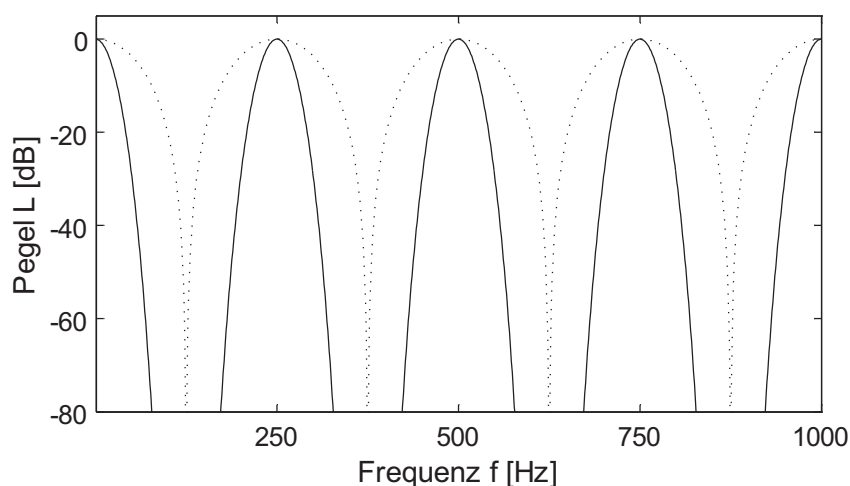


Abbildung 2.15: Qualitative Betragsspektren von Iterated Rippled Noise (IRN) nach Gleichung 2.8. Durchgezogen: $g=1$, $n=16$. Gestrichelt: $g=1$, $n=2$. Jeweils identische Verzögerungszeit $d=4$ ms.

unerwünscht und werden durch entsprechend angepasste Laufzeitglieder ausgeglichen. In Form von Flanger- und Phaser-Effekten sind sie in der modernen Popmusik jedoch oftmals erwünscht, da sie eine interessante Möglichkeit bieten, die Klangfarbe eines Instruments zu variieren. In allen Fällen lassen sich diese Effekte mit Hilfe von Verzögerungsschaltungen beschreiben. Die resultierende zusätzliche Tonhöhe entspricht dem Kehrwert des auftretenden Laufzeitunterschieds bzw. der eingestellten Verzögerungszeit d (“delay“) im Rückkopplungsweig.

Diese Effekte wurden bereits sehr früh von Huygens als Klangcolorierung bei durch Treppen reflektiertes Brunnenrauschen am französischen Hof erkannt. Später erkannte Minnaert die Verbindung mit der Wahrnehmung von Tonhöhen. Eine Übersicht zu historischen und neueren Erkenntnissen zu Wiederholungstonhöhen (“Repetition Pitch“) ist beispielsweise bei Bilsen und Ritsma zu finden [Bil70].

Die Stärke der Kammfilterwirkung kann mit Hilfe einer Dämpfung im Verzögerungsteil (Parameter g , “gain“) oder durch mehrfaches Hintereinanderschalten identischer Kammfilter (Parameter n “Iterationen“) reguliert werden. Abhängig von diesen beiden Parametern läßt sich die Deutlichkeit der Wahrnehmbarkeit der erwünschten oder unerwünschten Tonhöhe, also ihre Ausprägtheit der Tonhöhe, variieren.

Diejenigen Schalle, die durch mehrfaches Durchlaufen identischer FIR oder IIR-Filter generiert werden, werden in der Liteartur als “Iterated Rippled Noise“ (IRN) bezeichnet [Yos96]. Diese Schalle besitzen je nach Dämpfung und Zahl der Schleifendurchläufe (Iterationen) mehr oder weniger stark ausgeprägte spektrale Maxima bei ganzzahligen Vielfachen des Kehrwerts der Verzögerungszeit. Die wahrgenommene Tonhöhe entspricht

also weitestgehend der eines harmonisch komplexen Tones, dessen spektrale Komponenten jeweils bei diesen Maxima liegen.

Das Spektrum der durch n-faches Durchlaufen einer Verzögerungsschaltung der Verzögerung d und Dämpfung g entstehenden Schalle läßt sich durch

$$H(\omega) = (1 + ge^{-j\omega d})^n \quad (2.7)$$

beschreiben, das Betragsspektrum ergibt sich zu:

$$|H(\omega)|^2 = [1 + g^2 + 2g \cos(\omega d)]^n \quad (2.8)$$

Sowohl g als auch n haben dabei Einfluss auf die in Abbildung 2.15 dargestellte Tiefe und Breite der spektralen Einschnitte. Die Abstände der Maxima, also die Tonhöhe des entstehenden Signals wird, abgesehen von möglichen Verschiebungseffekten, ausschließlich durch die Verzögerungszeit d festgelegt.

2.7.1 Iterated Rippled Noise (IRN) als Anker- und Testschall

In Anlehnung an Experimente aus der Literatur (siehe auch [Yos78, Fas88, Wie98, Wie00]) wurden IRN-S (IRN "Add-Same") entsprechend [Yos96] mit dem "Add-Same" Algorithmus erstellt. Hierbei entspricht das Eingangssignal einer Filterstufe jeweils dem Ausgangssignal der vorangegangenen Stufe.

Die spektrale Lage der Schalle wurde dabei mit einem Bandpassfilter der Bandbreite 1000 Hz bei verschiedener Mittenfrequenz variiert, wodurch je nach Mittenfrequenz spektrale und virtuelle Tonhöhenwahrnehmung auftreten kann. Durch verschiedene Verzögerungszeiten (4 bis 32 ms) standen vier verschiedene Tonhöhen zur Verfügung ($f_0 = 250, 125, 62,5$ und $31,25$ Hz). Innerhalb einer Filterbedingung diente jeweils das Signal mit einer Verzögerungszeit von 4 ms als Ankerschall, es wurden also Schalle unterschiedlicher Tonhöhe miteinander verglichen.

Die Ergebnisse in Abbildung 2.16 zeigen mit steigender Verzögerungszeit d ein Absinken der Ausgeprägtheit der Tonhöhe. Dies ist in Einklang mit den Ergebnissen für Sinustöne unterschiedlicher Frequenz aus Abbildung 2.1 die für Frequenzen bis 1000 Hz mit der Tonhöhe ansteigen. Dabei muß man die extrem niedrigen Grundfrequenzen von 31,25 und 62,5 Hz bereits als Grenzfälle der Tonhöhenwahrnehmung betrachten, da diese Schalle eine sehr starke Rauigkeitsempfindung hervorrufen, die die Schallwahrnehmung dominiert [Ter68a]. Die verschiedenen spektralen Filterbedingungen haben deutlich unterschiedliche Ergebnisse zur Folge. So ist der Rückgang der Ausgeprägtheit der Tonhöhe bei der höchsten Mittenfrequenz ($f_c = 4000$ Hz) von 100 auf 80 % nur etwa halb so groß wie bei der niedrigsten Mittenfrequenz ($f_c = 500$ Hz), wo die Ausgeprägtheit der Tonhöhe von 100 auf ca. 65 % abfällt.

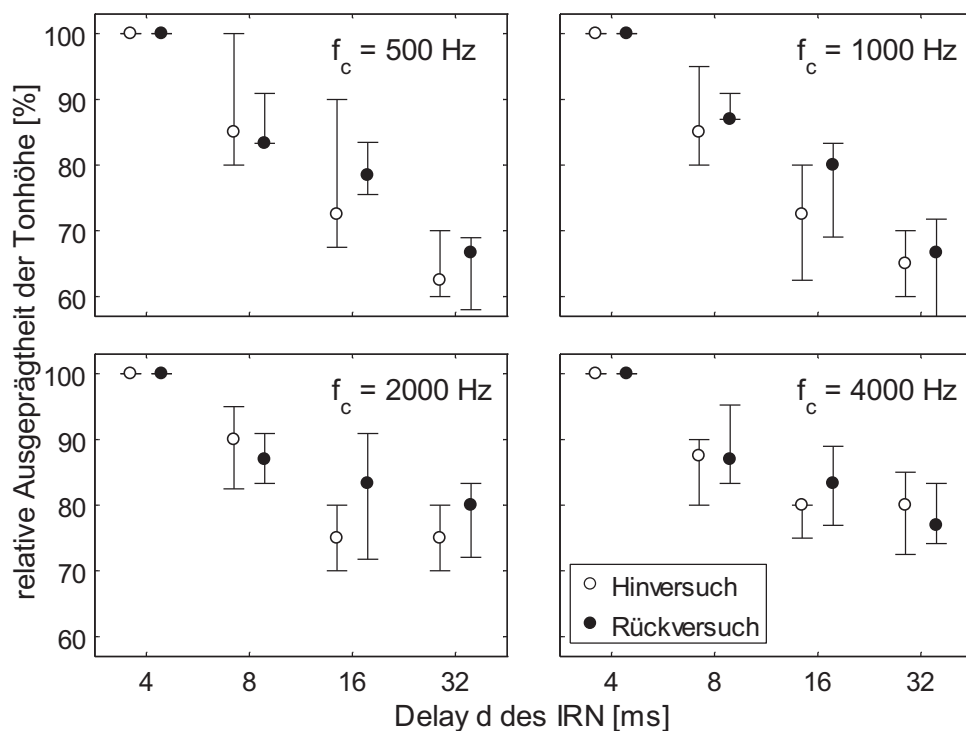


Abbildung 2.16: Relative Ausprägtheit der Tonhöhe von Iterated Rippled Noise. Anker- und Testschalle sind Iterated Rippled Noises bei identischer Filterbedingung und variierender Verzögerungszeit. Daten nach [Fru02].

Diese Ergebnisse lassen sich über die Prinzipien der Frequenzgruppen erklären. Da die zur geringsten Verzögerungszeit resultierende Grundfrequenz den höchsten Wert annimmt, ist diese von den Frequenzgruppenfiltern am besten erfassbar und ruft so die jeweils am stärksten ausgeprägte Tonhöhe hervor. Die unterschiedlichen Resultate bei verschiedenen Filterbedingungen sind dagegen auf die Lage der spektralen Komponenten im Existenzbereich der virtuellen Tonhöhe zurückzuführen. In der höchsten Filterbedingung wird der in Abbildung 1.1 dargestellte Existenzbereich der virtuellen Tonhöhe bereits wieder verlassen, weshalb die virtuelle Tonhöhe, vergleichbar mit den Ergebnissen für amplitudenmodulierte Sinustöne (Abschnitt 2.2) und harmonisch komplexe Residualtöne (Abschnitt 2.5), nur noch sehr schwach wahrnehmbar ist. Die Tonhöhenwahrnehmung ähnelt bei dieser spektralen Lage sehr stark derjenigen, wie sie ein Bandpassrauschen bei gleicher Mittenfrequenz und Bandbreite hervorruft.

2.7.2 Iterated Rippled Noise im Vergleich zu Rauschen mit einfacher Blockwiederholung

Als AABB-Rauschen bezeichnete Wiederholungsrauschen entstehen durch die einfache Wiederholung gleichlanger Segmente (A, B, ...) eines Rauschsignals. Aufgrund der jeweils einfachen Wiederholung (AA, BB, CC, ...) entstehen Tonhöhen die, wie bei Iterated Ripp-

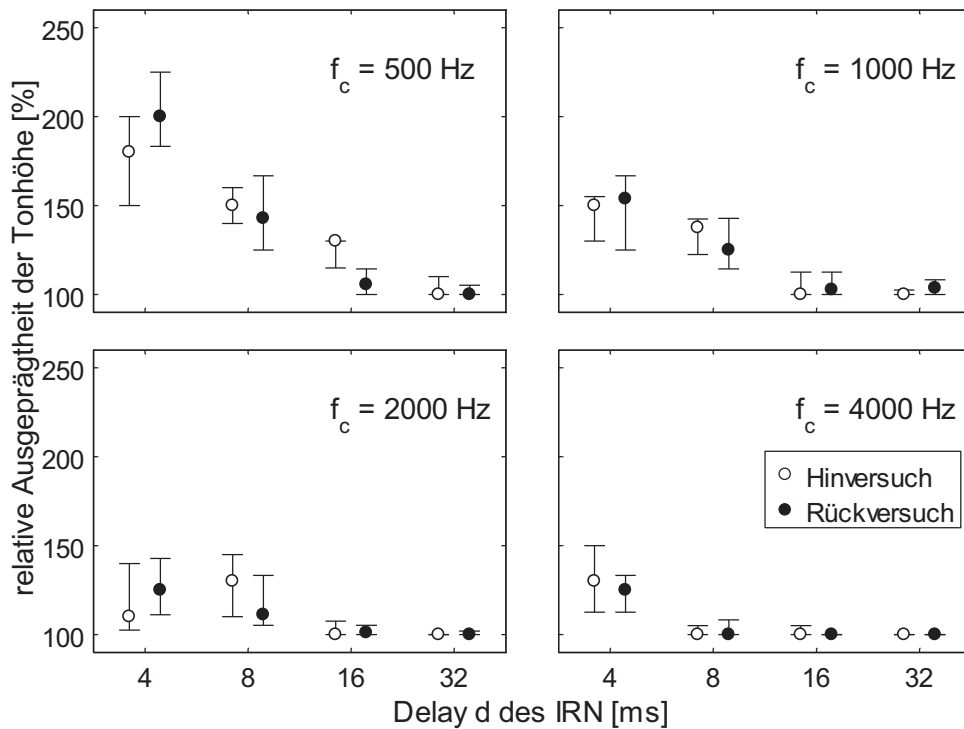


Abbildung 2.17: Relative Ausprägtheit der Tonhöhe von Iterated Rippled Noise, Anker AABR-Rauschen. Identische Filterbedingung von Anker- und Testschall, Segmentlänge entsprechend der Verzögerungszeit des untersuchten Iterated Rippled Noise. Daten nach [Fru02].

led Noise, dem Kehrwert der Verzögerungszeit, hier also der Segmentlänge, entsprechen. Abbildung 2.17 zeigt die Ausprägtheit der Tonhöhe von Iterated Rippled Noise im Vergleich zu AABR-Rauschen entsprechend [Fru02]. Als Anker diente jeweils AABR-Rauschen dessen Segmentlänge mit der Verzögerungszeit des Iterated Rippled Noise übereinstimmt, wodurch die zu vergleichenden Schalle eine ähnliche absolute Tonhöhe hervorrufen. Als Eingangssignal der Algorithmen zur Erzeugung der Iterated Rippled Noise und AABR-Rauschen diente weisses Rauschen. Die verwendeten Iterated Rippled Noises wurden als IRN-S mit $n = 16$ Iterationen und einer Dämpfung von $g = 0$ dB im erzeugt.

Es zeigt sich, daß für niedrige Verzögerungszeiten ($d = 8$ bzw. 4 ms entsprechend 125 und 250 Hz) die Ausprägtheit der Tonhöhe von Iterated Rippled Noise deutlich höher ist, als die von AABR-Rauschen. Aufgrund des Auflösungsvermögens des Gehörs werden auch hier die spektralen Maxima der untersuchten Schalle besser erfasst. Diese sind jedoch bei AABR-Rauschen deutlich weniger stark ausgeprägt als bei Iterated Rippled Noise mit der hier verwendeten relativ hohen Anzahl an Iterationen. Die einzelnen “Harmonischen“ des Iterated Rippled Noise werden also deutlicher wahrgenommen als die recht schwach ausgeprägten Maxima des AABR-Rauschens und tragen so stärker zur Ausbildung der Tonhöhenwahrnehmung bei. Für die Tonhöhen von 31,25 und 62,5 Hz ist dagegen na-

hezu kein Unterschied der Ausprägtheit der Tonhöhe wahrnehmbar. Aufgrund dieser sehr niedrigen Tonhöhen fallen viele spektrale Maxima in eine Frequenzgruppe, was eine schlechte Auflösbarkeit und so eine geringe Ausprägtheit der Tonhöhe zur Folge hat. Spektrale Maskierungseigenschaften führen hier also zu einer untergeordneten Rolle der spektralen Maxima.

2.7.3 Iterated Rippled Noise im Vergleich zu harmonisch komplexen Tönen

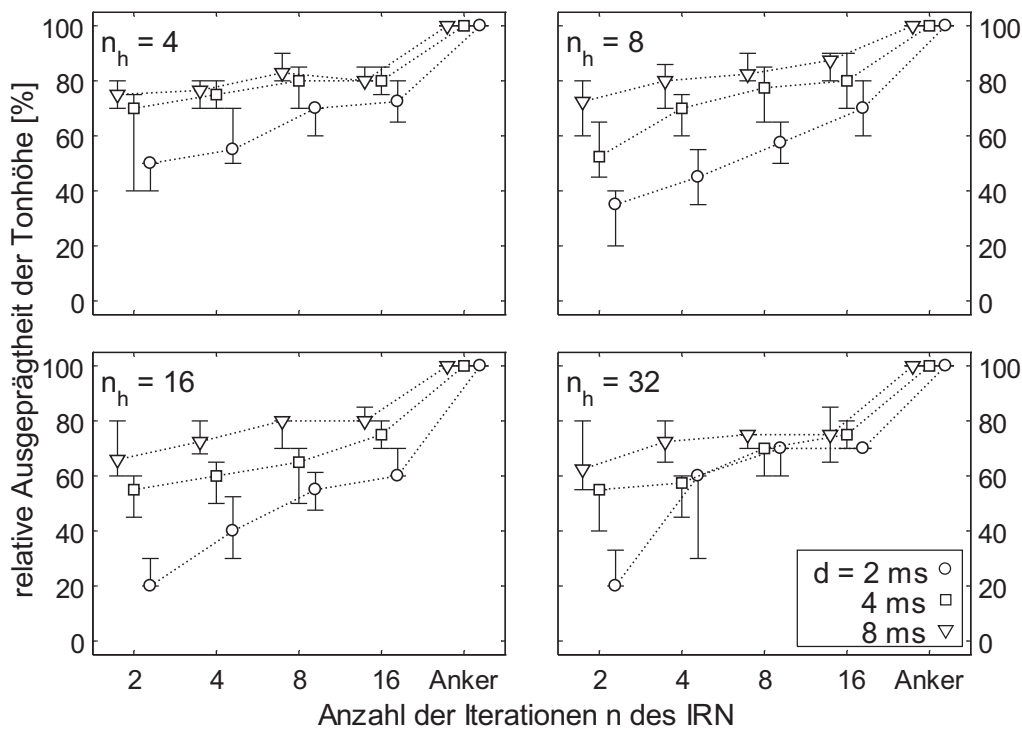


Abbildung 2.18: Relative Ausprägtheit der Tonhöhe von Iterated Rippled Noise bei Bandpassfilterung entsprechend der Anzahl Harmonischer n_h des Ankerschalls (harmonisch komplexer Ton). Daten nach [Fru04b].

Die Spektren von Iterated Rippled Noise sind, wie in Abschnitt 2.7 beschrieben, vergleichbar mit denen von harmonisch komplexen Tönen. Unterscheidungsmerkmal ist die ‘Breite’ der einzelnen Spektralkomponenten, die bei harmonisch komplexen Tönen lediglich aus einer einzelnen Sinuslinie bestehen und bei Iterated Rippled Noise eher den Charakter von Schmalbandrauschen besitzen.

Durch Erhöhung der Anzahl an Iterationen werden diese Maxima schmaler und heben sich gegenüber den dazwischen liegenden Minima deutlicher ab. Dieser Effekt wurde in Abbildung 2.15 dargestellt. Im Extremfall unendlich vieler Iterationen nähert sich das Spektrum von Iterated Rippled Noise dem Spektrum eines entsprechend parametrisierten

harmonisch komplexen Tons an und die beiden Schalle lassen sich nur noch aufgrund verbliebener zeitlicher Effekte unterscheiden. Diese entstehen bei Iterated Rippled Noise aufgrund des statistischen Eingangssignals des Synthesealgorithmus in Form von rosa oder weissem Rauschen, bei harmonisch komplexen Tönen sind aufgrund der Harmonischen in Form von Sinustönen keine zeitlichen Variationen vorhanden.

Abbildung 2.18 zeigt die Ausgeprägtheit der Tonhöhe von Iterated Rippled Noise im Vergleich zu harmonisch komplexen Tönen bei ähnlicher spektraler Zusammensetzung (Verzögerungszeit $d = 1/f_0$, Bandbegrenzung entsprechend der Harmonischen des harmonisch komplexen Tons). Der Ankerschall (harmonisch komplexer Ton) besitzt jeweils die stärkste Ausgeprägtheit der Tonhöhe, wobei Iterated Rippled Noises mit 16 Iterationen bereits 80 bis 90 % der Ausgeprägtheit der Tonhöhe dieses harmonisch komplexen Tons erreichen können. Bei Reduzierung der Anzahl der Iterationen sinkt aufgrund des rauschartigeren Charakters des entstehenden Spektrums die Ausgeprägtheit der Tonhöhe deutlich ab. In Einklang mit vorangehend beschriebenen Versuchsergebnissen sinkt die Ausgeprägtheit der Tonhöhe für höhere Grundfrequenzen (niedrigere Verzögerungszeiten) stärker als für niedrige Grundfrequenzen. Dies ist auf die Auflösbarkeit der einzelnen Teiltöne zurückzuführen, die dazu führt, daß die MHS-Muster von harmonisch komplexen Tönen und Iterated Rippled Noises bei niedrigeren Tonhöhen bereits bei weniger Iterationen ähnliche Verläufe zeigen als bei hohen Frequenzen.

Abhängig von der Gesamtanzahl an Harmonischen bzw. der Bandbreite der untersuchten Schalle sinkt die Ausgeprägtheit der Tonhöhe bei Verwendung einer höheren Anzahl Harmonischer stärker ab als bei der Verwendung von beispielsweise lediglich vier Harmonischen. Hier besitzt die Zusammensetzung der Schalle bei höheren Frequenzen einen störenden Einfluß. Rauschähnliche Komponenten des Iterated Rippled Noise wirken sich bei vergrößerter Bandbreite im Vergleich mit dem jeweiligen harmonischen komplexen Ton stärker negativ auf die Wahrnehmung der Ausgeprägtheit der Tonhöhe aus als bei Iterated Rippled Noise geringerer Bandbreiten bzw. harmonisch komplexer Töne niedrigerer Anzahl an Harmonischen.

2.8 Weitere Untersuchungen zur Ausgeprägtheit der Tonhöhe von Bandpassrauschen

Die Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Bandbreite von Bandpassrauschen wurde beispielsweise in [Wie91] untersucht. Die dabei gefundenen Ergebnisse sind in Abbildung 2.3 dargestellt und zeigen eine deutliche Abhängigkeit von den Parametern Mittenfrequenz sowie Bandbreite des Bandpassrauschens. Dabei führt bei konstanter Mittenfrequenz eine Erhöhung der Bandbreite zu einer Abnahme der Ausgeprägtheit der Tonhöhe. Für eine spätere Modellierung der Ausgeprägtheit der Tonhöhe und Umsetzung in einen Algorithmus sind gerade hinsichtlich der Zusammenhänge mit Frequenzgruppen und Frequenzgruppenbreite weiterführende Untersuchungen nötig.

2.8.1 Unterschiedsschwelle der Ausprägtheit der Tonhöhe

Die Größe der Änderung einer psychoakustischen Wahrnehmung läßt sich über das Äquivalenzgesetz beschreiben [Ter68c]. Dabei zugrunde liegen die sogenannten Unterschiedsschwellen, deren mehrfaches Überschreiten wiederum zu einer dementsprechend größeren Änderung der Empfindung führt. Wie in Abschnitt 1.6.3 beschrieben, müssen bei Experimenten mit Hilfe adaptiver Verfahren minimale Schrittweiten definiert werden, um möglichst rasch ein Abbruchkriterium zu erreichen. Hierfür wurde in einem Experiment die eben wahrnehmbare Unterschiedsschwelle der Bandbreitenänderung von Bandpassrauschen untersucht. Das Beurteilungskriterium war dabei, ob ein Unterschied bezüglich der Ausprägtheit der Tonhöhe der dargebotenen Bandpassrauschen feststellbar sei.

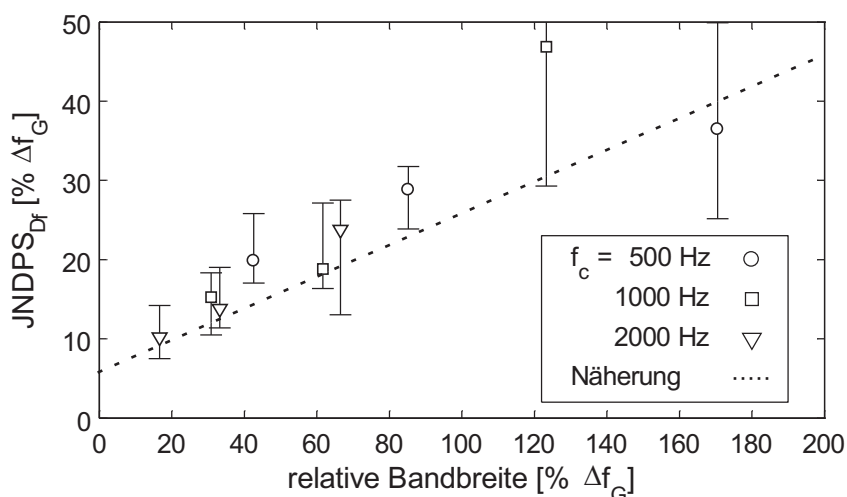


Abbildung 2.19: Unterschiedsschwellen für die Ausprägtheit der Tonhöhe von Bandpassrauschen bezogen auf die jeweilige Frequenzgruppenbreite. Darbietungsspegel der Testschalle entsprechend einer Lautheit von $N_T = 5$ sone, Testschalldauer $T_T = 0,5$ s. Gestrichelt eingezeichnet: Näherung aus Formel 2.9.

Die Ergebnisse von Experimenten zur Bestimmung der Unterschiedsschwelle der Ausprägtheit der Tonhöhe bei Bandpassrauschen sind in Abbildung 2.19 für verschiedene Mittenfrequenzen und Bandbreiten dargestellt. Dabei ist jeweils die benötigte Bandbreitenänderung aufgetragen, die zu einer wahrnehmbaren Veränderung der Ausprägtheit der Tonhöhe führt.

Aufgrund der herausragenden Bedeutung der Frequenzgruppen wird diese Darstellung jeweils relativ zur Frequenzgruppenbreite bei der Mittenfrequenz des untersuchten Schalls gewählt. Die Ergebnisse zeigen einen Anstieg der Unterschiedsschwelle mit steigender relativer Bandbreite der untersuchten Bandpassrauschen. Die in Abbildung 2.19 eingetragene

Näherung folgt dem linearem Zusammenhang

$$JNDPS_{\Delta f} = \frac{1}{5} \frac{\Delta f}{\Delta f_G(f_c)} + 8 \quad [\% \Delta f_G] \quad (2.9)$$

und liegt für die untersuchten Schalle stets im Interquartilbereich der Versuchsergebnisse. Zwei der untersuchten Schalle besitzen Bandbreiten, die jeweils größer sind als die Frequenzgruppenbreite. Für beide Schalle ($f_c = 1000$ Hz, $\Delta f = 120 \% \Delta f_G$ und $f_c = 500$ Hz, $\Delta f = 170 \% \Delta f_G$) weisen die Ergebnisse sehr große Streubereiche auf und die Mediane weichen deutlich von der angegebenen Näherung ab, welche dennoch innerhalb der Interquartilbereiche liegt.

2.8.2 Ausprägtheit der Tonhöhe von Bandpassrauschen verschiedener Mittenfrequenzen

Für die Ermittlung der bei unterschiedlichen Mittenfrequenzen benötigten Bandbreite zur Erzielung gleicher Ausprägtheit der Tonhöhe wurde ein adaptives Verfahren gewählt. Mit Hilfe pendelnden Einregelns wird dabei die Bandbreite eines Vergleichsschalls variiert, bis das Abbruchkriterium erreicht ist. Die Ergebnisse dieser Experimente entsprechend [Fru04a] sind in Abbildung 2.20 dargestellt. Der Vergleichsschall konstanter Bandbreite $\Delta f_{ref} = 20 / 50 / 100$ Hz besitzt hierbei eine Mittenfrequenz von 500 Hz. Testschalle bei Mittenfrequenzen von 250, 1000 und 2000 Hz wurden adaptiv eingeregelt bis gleiche wahrgenommene Ausprägtheit der Tonhöhe erreicht wurde. Zur Überprüfung der Versuchsmethode dienten Testschalle bei einer Mittenfrequenz von 500 Hz.

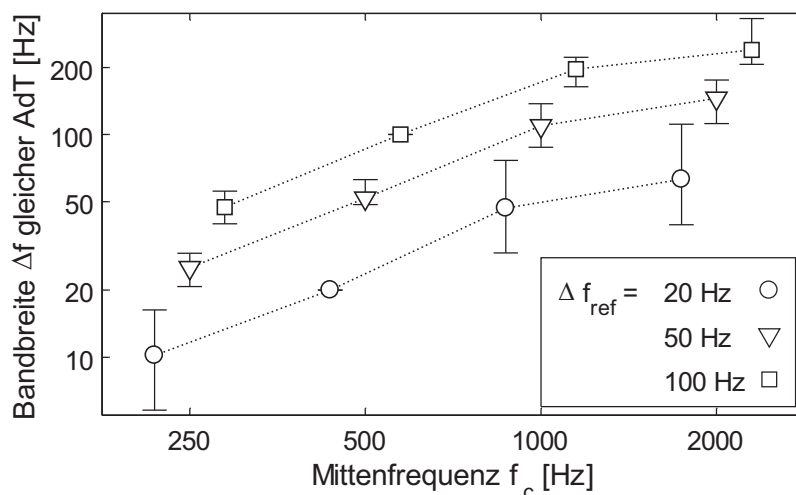


Abbildung 2.20: Bandbreite von Bandpassrauschen unterschiedlicher Mittenfrequenz bei gleicher Ausprägtheit der Tonhöhe wie ein Referenz-Bandpassrauschen bei einer Mittenfrequenz von 500 Hz. Darbietungspegel der Testschalle entsprechend einer Lautheit von $N_T = 5$ sone, Testschalldauer $T_T = 0,6$ s.

Für steigende Mittenfrequenzen steigt die benötigte Bandbreite, um eine gleiche Ausgeprägtheit der Tonhöhe hervorzurufen. Die Ergebnisse zeigen bei doppelt logarithmischer Darstellung eine deutlich schwächere Zunahme der Bandbreite ab einer Mittenfrequenz von 1000 Hz. Die Ergebnisse verschiedener Referenzbandbreiten lassen sich dagegen entlang der Ordinate verschieben und ineinander überführen.

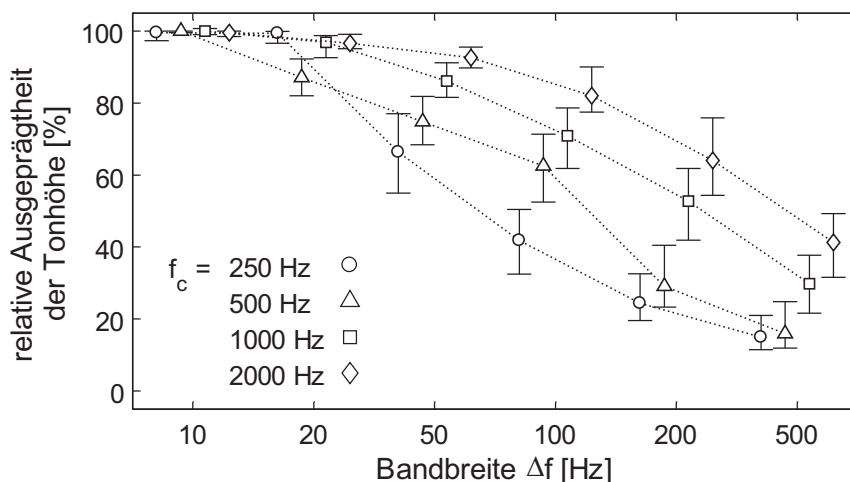


Abbildung 2.21: Relative Ausgeprägtheit der Tonhöhe von Bandpassrauschen in Abhängigkeit von der absoluten Bandbreite Δf . Anker: Schmalbandrauschen, $\Delta f = 10$ Hz. Daten nach [Fru04a].

Eine Beschreibung der gewonnenen Daten, einzig basierend auf dem Konzept der Frequenzgruppen, ist nicht möglich, da die Ergebnisse bei Normierung auf die Frequenzgruppenbreite Δf_G keine, konstanten Verläufe aufweisen. Da für Mittenfrequenzen von 250 sowie 500 Hz die Frequenzgruppenbreite ca. 100 Hz beträgt, müßte auch in der in Abbildung 2.20 gewählten Darstellung ein horizontaler Verlauf zwischen den zugehörigen Werten entstehen, was nicht der Fall ist. Zieht man zusätzlich die Frequenzabhängigkeit der Ausgeprägtheit der Tonhöhe, wie sie in Abschnitt 2.1.1 dargestellt ist, in Betracht, so läßt sich eine qualitative Erklärung der Ergebnisse formulieren. Bei identischen Bandbreite besitzt folglich ein Bandpassrauschen niedrigerer Mittenfrequenz eine geringere Ausgeprägtheit der Tonhöhe als bei höherer Mittenfrequenz. Um nun die selbe Ausgeprägtheit der Tonhöhe hervorzurufen muß die Bandbreite verringert werden, um eine Erhöhung der Ausgeprägtheit der Tonhöhe auf den Wert des Referenzschalls zu bewirken.

2.8.3 Abhängigkeit von der Frequenzgruppenbreite

Zur weiteren Untersuchung der Ausgeprägtheit der Tonhöhe von Bandpassrauschen wurden zwei Experimente ähnlich den in Abschnitt 2.1.4 dargestellten Untersuchungen durchgeführt. Dabei wurde zunächst die Abhängigkeit der Ausgeprägtheit der Tonhöhe von der absoluten Bandbreite für einen Ankerschall der Bandbreite 10 Hz untersucht. Die Pegel

der Schalle wurden wiederum so gewählt, daß sämtliche untersuchten Signale eine Lautheit von 5 sone aufwiesen. Die Ergebnisse in Abbildung 2.21 zeigen die aus Abbildung 2.3 bekannte Abhängigkeit von der Mittenfrequenz. Bei niedrigerer Mittenfrequenz ($f_c = 250$ Hz) sinkt die Ausgeprägtheit der Tonhöhe bei Erhöhung der Bandbreite schneller ab, als bei höheren Mittenfrequenzen. Die zugehörigen Experimente unterscheiden sich dabei lediglich in der abweichenden Wahl des Ankerschalls ($\Delta f = 10$ Hz anstelle von 3,16 Hz) und in der Normierung auf eine konstante Lautheit N_T anstelle konstanten Darbietungspegels L_T . Es tritt also nahezu keine Abhängigkeit der Ergebnisse vom gewählten Anker auf. Eine Pegelabhängigkeit der Ergebnisse ist im Rahmen der Pegeländerungen aufgrund der Anpassung an die konstante Darbietungslautheit ebenfalls nicht zu beobachten.

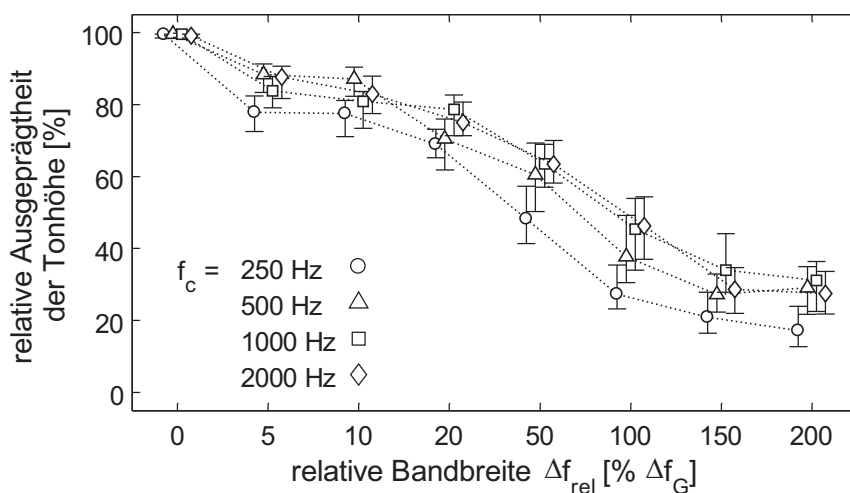


Abbildung 2.22: Relative Ausgeprägtheit der Tonhöhe von Bandpassrauschen bei Variation der Bandbreite relativ zur Frequenzgruppenbreite nach [Fru04a], Anker: Sinus.

In Abbildung 2.22 sind Ergebnisse dargestellt, bei denen die Bandbreite der Testschalle proportional zur Frequenzgruppenbreite gewählt wurde. Als Ankerschall diente ein Sinuston (relative Bandbreite: $\Delta f_{rel} = 0$ % der Frequenzgruppenbreite), der bei der entsprechenden Mittenfrequenz lag. Die resultierenden Werte liegen deutlich dichter beieinander als im vorangehend beschriebenen Experiment. Besonders bei einer Bandbreite die der Frequenzgruppenbreite entspricht ($\Delta f_{rel} = 100$ %) lassen sich dennoch deutliche Unterschiede erkennen. So ist für eine Mittenfrequenz von 2000 Hz die Ausgeprägtheit der Tonhöhe auf die Hälfte der Ausgeprägtheit des zugehörigen Sinustones abgefallen, während sie bei der niedrigsten Mittenfrequenz von 250 Hz bereits auf ca. 20 % gesunken ist. Die Ursache für diese Unterschiede liegt entweder in der zugrundeliegenden Vorstellung von der Bandbreite der Frequenzgruppen oder in einer Interaktion mit weiteren Hörwahrnehmungen, beispielsweise der Rauigkeit. Schalle mit einer Bandbreite von 100 Hz besitzen aufgrund ihrer effektiven Modulationsfrequenz $f_{mod}^* = 0,64 \Delta f = 64$ Hz eine sehr große Rauigkeit, was möglicherweise die Wahrnehmung der Ausgeprägtheit der Tonhöhe beeinträchtigt.

2.9 Zusammenfassung

Zusammenfassend sollen hier nochmals die wichtigsten Versuchsergebnisse zur Wahrnehmung der Ausgeprägtheit der Tonhöhe aufgeführt werden:

- Sinustöne weisen Abhängigkeiten auf, die der Lautheitswahrnehmung ähnlich sind. Steigender Pegel sowie steigende Dauer (bis $T_T \approx 300$ ms) führen zu einem Anstieg der Ausgeprägtheit der Tonhöhe. Die Frequenz eines geht in Form einer Bandpasscharakteristik in die Ergebnisse der Ausgeprägtheit der Tonhöhe ein. Amplitudenmodulation wirkt sich bis zu $f_{mod} = 32$ Hz mit steigender Modulationsfrequenz hemmend auf die Ausgeprägtheit der Tonhöhe aus.
- Die Ausgeprägtheit der Tonhöhe harmonisch komplexer Töne sinkt bei Erhöhung der Anzahl verwendeter Harmonischer. Eine identische Anzahl Harmonischer führt dabei bei unterschiedlichen Grundfrequenzen f_0 zu einem ähnlichen Rückgang der Ausgeprägtheit der Tonhöhe bei Vergleich mit dem Sinuston der Frequenz f_0 .
- Die Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Töne sinkt mit steigender Anzahl entfernter unterster Harmonischer. Harmonisch komplexe Residualtöne besitzen eine stärker ausgeprägte Tonhöhe, wenn ihr Spektrum einen konstanten Pegelabfall ΔL_f pro Oktave aufweist.
- Wiederholungsrauschen (Iterated Rippled Noise und AABB-Rauschen) verhalten sich bezüglich der Ausgeprägtheit der Tonhöhe ähnlich wie harmonisch komplexe Töne deren spektrale Zusammensetzung den Maxima im Spektrum der Wiederholungsrauschen entspricht. Wiederholungsrauschen besitzen generell eine schwächer ausgeprägte Tonhöhe als harmonisch komplexe Töne, diese lässt sich bei Iterated Rippled Noise durch Erhöhung der Anzahl verwendeter Iterationen anheben.
- Die Ausgeprägtheit der Tonhöhe von Bandpassrauschen verringert sich bei Erhöhung der Bandbreite. Bei niedrigen Mittenfrequenzen erfolgt dieser Rückgang der Ausgeprägtheit der Tonhöhe schneller als bei höheren Mittenfrequenzen. Eine Normierung der Bandbreiten auf die Frequenzgruppenbreite bei der jeweiligen Mittenfrequenz führt für unterschiedliche Mittenfrequenzen zu ähnlichen, aber nicht identischen Ergebnissen.

Die in diesem Kapitel dargestellten Versuche dienten der Erfassung grundsätzlicher Abhängigkeiten der Ausgeprägtheit der Tonhöhe von verschiedensten Schallparametern. Somit wurde für eine Vielzahl synthetischer Schalle ein Datensatz erstellt, der zukünftigen Untersuchungen als Ausgangspunkt dienen kann und darüberhinaus insbesondere zur Modellierung der Ausgeprägtheit der Tonhöhe Anwendung finden soll.

Kapitel 3

Zeitliche und spektrale Merkmale der Ausgeprägtheit der Tonhöhe

Die Ausgeprägtheit der Tonhöhe ist, entsprechend den in Kapitel 2 dargestellten Daten, von den physikalischen und psychoakustischen Parametern des jeweils untersuchten Schalls abhängig. Im folgenden Kapitel wird zunächst eine gehörgerechte Darstellungsform hergeleitet mit deren Hilfe sich charakterisierende Einflußgrößen der Ausgeprägtheit der Tonhöhe darstellen lassen. Für die vorangehend untersuchten Schallarten werden anhand dieser Darstellung Merkmale ermittelt, die die jeweiligen Einflußgrößen über spektrale und zeitliche Eigenschaften repräsentieren.

3.1 Gehörgerechte Darstellung von Audiosignalen

Bereits in Kapitel 1.4 wurde die Fourier-t-Transformation als gehörangepasste Spektraltransformation eingeführt. Die Anpassung an psychoakustische Eigenschaften, wie die Trennbarkeit einzelner Spektraltonhöhen sowie die Detektierbarkeit von Tonhöhenschwankungen, erfolgt dabei über die geeignete Parameterwahl. Darüber hinaus muß eine Modellierung der Maskierungseffekte zur Unterdrückung nicht wahrnehmbarer Schallkomponenten erfolgen. Eine Transformation in die Lautheit stellt zusätzlich eine psychoakustisch richtige "Pegeldarstellung" sicher. Die zu dieser Darstellung notwendigen Einzelschritte werden nachfolgend erläutert.

3.1.1 Gewählte Transformationsparameter der Fourier-t-Transformation

Über eine entsprechende Parameterwahl kann die FTT so angepasst werden, daß die Transformationsergebnisse eine wahrnehmungsgerechte optische Repräsentation des Schalls darstellen. Zum Einen ist die geeignete Wahl der verwendeten Fensterfunktion von großer Bedeutung um die Voraussetzungen für eine adäquate Darstellung der zeitlichen und spektralen Eigenschaften des Gehörs zu ermöglichen [Sch90], andererseits werden Anzahl und Bandbreite der verwendeten Filter durch eben wahrnehmbare Frequenz- und Amplitudenänderungen vorgegeben.

Um diese Eigenschaften zu erhalten schlägt Mummert, ausgehend von den von Heinbach für das Teiltonzeitmuster verwendeten Parametern, die Verwendung von $NFI = 481$ Kanälen mit einem Abstand von $\Delta z_A = 0,051$ Bark vor [Hei88b, Mum98]. Dies stellt sicher, daß benachbarte Kanäle ausreichend dicht beieinander liegen um Einzeltöne trennen zu können und entspricht der Detektionsschwelle für Frequenzmodulationen von ca. 0,04 Bark [Zwi67]. So sind also die kleinsten im Transformationsergebnis sichtbaren Frequenzänderungen vergleichbar mit den durch das Gehör eben wahrnehmbaren Frequenzunterschieden.

Die Analysebandbreite (3dB-Bandbreite) der Filter ergibt sich aus der Wahrnehmbarkeit zweier getrennter Tonhöhen. Dieser Abstand beträgt 0,2 bis 0,5 Bark [Plo64, Ter68b], was zu einer Parameterwahl von $\Delta z = 0,3$ Bark führt. Im Transformationsergebnis sind also getrennte Linien bzw. Maxima nur für solche Schallkomponenten sichtbar, die auch tatsächlich vom Gehör als getrennte Tonhöhen erfasst werden können.

Die Parameter Bandbreite und Abstand der Filter ergeben sich also nicht durch direktes Umskalieren der Barkskala über einen festen Wert sondern werden getrennt, aufgrund einzelner Gehöreigenschaften, angepasst.

3.1.2 Berechnung von Mithörschwellenmustern

Der Effekt, daß laute Geräusche leisere Geräusche verdecken, also unhörbar machen können, wird in der Psychoakustik als "Maskierung" bezeichnet. Diejenige Pegelschwelle, die ein Testschall überschreiten muß, um in Anwesenheit eines maskierenden Schalls wahrnehmbar zu sein, wird als Mithörschwelle (MHS) dieses maskierenden Schalls bezeichnet. Dabei können sowohl spektrale als auch zeitliche Verdeckungseffekte auftreten. Diese Effekte werden beispielsweise in der datenreduzierten Audiokodierung angewandt [Wie90, Bra94] und müssen im Sinne einer gehörgerechten Spektraldarstellung auch im hier zu entwickelnden Modell berücksichtigt werden.

Spektrale Maskierungseffekte, die zu eingeschränkter oder ganz veränderter Wahrnehmung weiterer Schallkomponenten führen können, wurden beispielsweise für Sinustöne von Schöne [Sch77, Sch79] und Sonntag [Son82] untersucht. Effekte die durch schmalbandige und komplexere Maskierer hervorgerufen werden, sind darüberhinaus beispielsweise in den Arbeiten von Fastl zu finden [Fas75, Fas76a, Fas76b, Fas76c, Fas79b].

Bei der vorangegangenen beschriebenen Parametrierung der FTT ist diese in der Lage, das spektrale und zeitliche Auflösungsvermögen des menschlichen Gehörs aufgrund der gewählten Filterbandbreiten und -abstände weitestgehend korrekt zu reproduzieren. Da jedoch die Flankensteilheit der beobachteten Mithörschwellen vom Pegel und der Zusammensetzung des maskierenden Schalls abhängt [Sch77, Sch79, Son82], die FTT jedoch mit konstanten, pegelunabhängigen Filterprototypen arbeitet, ist eine zusätzliche Nachbildung der Mithörschwellen nötig.

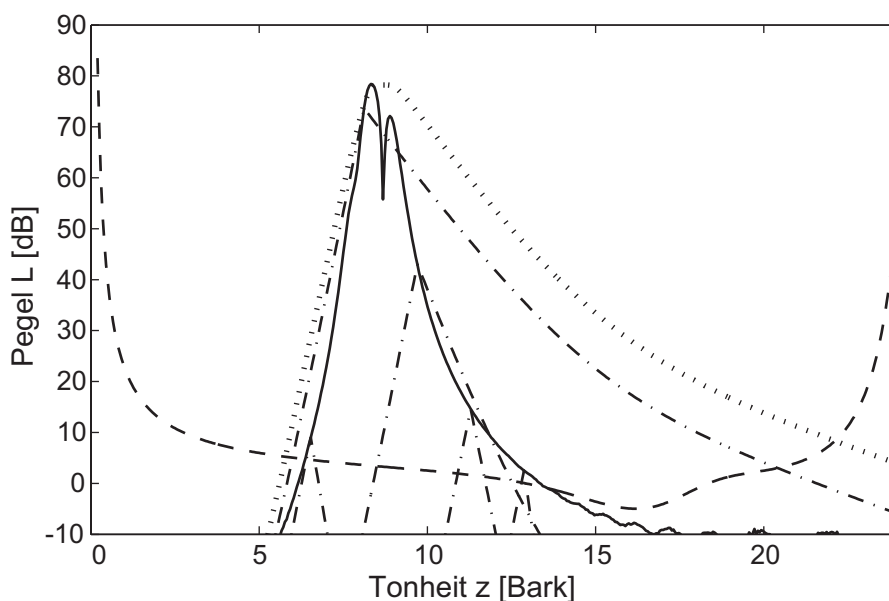


Abbildung 3.1: Ausschnitt aus dem FTT-Spektrogramm eines Schmalbandrauschens mit $f_c = 1$ kHz und $L_M = 80$ dB (durchgezogene Linie) mit exemplarischen Erregungsmustern einzelner Analysepunkte (punktirt-gestrichelt) und resultierender Mithörschwelle (gepunktet)

Eine einfache Näherung zur Beschreibung der Form spektraler Mithörschwellen beschreibt die MHS eines Maskierers der Frequenz f_M und des Pegels L_M über zwei Geradenstücke. Die Steigung der Flanke zu tieferen Frequenzen wird dabei mit s_l und die Steigung der Flanke zu Frequenzen oberhalb f_M wird als s_u bezeichnet. Die Steigungen s_l und s_u lassen sich dabei über

$$s_l = -27 \text{ dB/Bark} \quad (3.1)$$

$$s_u(L_M, f_M) = \left(-24 - \frac{230 \text{ Hz}}{f_M} + 0,2 \frac{L_M}{\text{dB}} \right) \text{ dB/Bark} \quad (3.2)$$

beschreiben [Mai67, Ter79]. Die Flanke zu Frequenzen unterhalb der Maskiererfrequenz f_M ist mit einem Wert von 27 dB/Bark näherungsweise frequenz- und pegelunabhängig. Die Flanke s_u besitzt sowohl eine Abhängigkeit von der Maskiererfrequenz f_M als auch vom Maskiererpegel L_M .

Um aus den Näherungen für Sinus-Maskierer zu den tatsächlichen Mithörschwellen beliebiger Schalle zu gelangen, lassen sich die einzelnen Spektralanteile der Fourier-t-Transformation als Teilkomponenten betrachten [vR99]. Für jede dieser Teilkomponenten wird die beschriebene Näherung berechnet und so die Teilmithörschwellen gebildet.

Die resultierende Mithörschwelle des Gesamtschalls ergibt sich über eine Intensitätsaddition der einzelnen Teilmithörschwellen. Beispielhaft ist dieses Vorgehen in Abbildung 3.1

für Schmalbandrauschen dargestellt. Auch Additionseffekte der MHS mehrerer Schallkomponenten werden auf diese Weise nachgebildet [Zwi75, Lum84b, Lum84a].

Neben dieser Näherung existieren auch aufwändigere Beschreibungen zur Berechnung von Mithörschwellenmustern [Gla90]. Da jedoch die Form der MHS aufgrund der gewählten Intensitätsaddition sehr gut nachvollzogen werden kann, wurde für die Berechnung der elementaren MHS lediglich die erwähnte Geradennäherung mit überlagerten Sigmoidfunktionen verwandt.

Zusätzlich zu den spektralen Verdeckungseigenschaften treten zeitliche Effekte auf, die zu einer Verdeckung aufeinander folgender Schalle führen können. Dieses Verhalten läßt sich über einen Tiefpass realisieren, dessen Zeitkonstante über die Schalldauer adaptiv angepasst wird [Zwi84, Wid98]. Dazu wird über das Eingangssignal ein Kondensator geeigneter Kapazität aufgeladen, wodurch für länger andauernde Schalle ein langsamerer Signalabfall am Ausgang realisiert wird, als bei kurzen Signalen, bei denen der Kondensator noch nicht geladen ist. Dieser in [Vog75] beschriebene Tiefpass wird in algorithmischer Form in das Modell der Mithörschwellenberechnung übernommen.

3.1.3 Modifikation des Lautheitsmodells nach Zwicker

Bis zu diesem Schritt erfolgte die Verarbeitung der Signale ausschließlich im Pegelmaß. Gemessene Schallpegel sind jedoch kein eindeutiges Maß für die wahrgenommene Lautstärke. So werden Schalle unterschiedlicher Bandbreite bei identischem Schallpegel als unterschiedlich laut wahrgenommen [Zwi83] und auch die Anwendung von Bewertungskurven (beispielsweise A- oder C-Bewertungen [DIN94]) kann hier zu irreführenden Ergebnissen führen und der Lautheitsempfindung entgegengesetzte Werte liefern [Hel87]. Für die gehörgerechte Darstellung von Schallsignalen muß also die Darstellung der Magnituden im psychoakustischen Maßstab "Lautheit" erfolgen. Die Bestimmung dieser Werte wird nachfolgend dargestellt.

Aufgrund der in Abschnitt 3.1.1 beschriebenen Parameterwahl für die Fourier-t-Transformation, mußte eine Erhöhung der Kanalanzahl des Berechnungsverfahrens der Lautheit erfolgen. Das algorithmisch vorliegende Lautheitsmodell bestimmt aufgrund der Anwendung für Schwerhörende die spektrale Maskierung in der Lautheitsdomäne [Cha02b]. Da die Daten der in Abschnitt 3.1.2 beschriebenen Modellierung der Mithörschwellen überwiegend im Pegelmaß vorliegen, wurde für das modifizierte Berechnungsverfahren die Reihenfolge aus Lautheitstransformation und Maskierungsberechnung wieder vertauscht. Das so entstehende, modifizierte Verfahren zur Lautheitsberechnung entspricht also wieder den Gegebenheiten, wie sie in [Pau72] vorgeschlagen wurden. Das zugehörige Blockschaltbild ist in Abbildung 3.2 dargestellt.

Da das modifizierte Modell mit einer wesentlich höheren Kanalanzahl arbeitet als dies im ursprünglichen Modell der Fall war, mußte eine Anpassung der Summation der Einzelkanäle

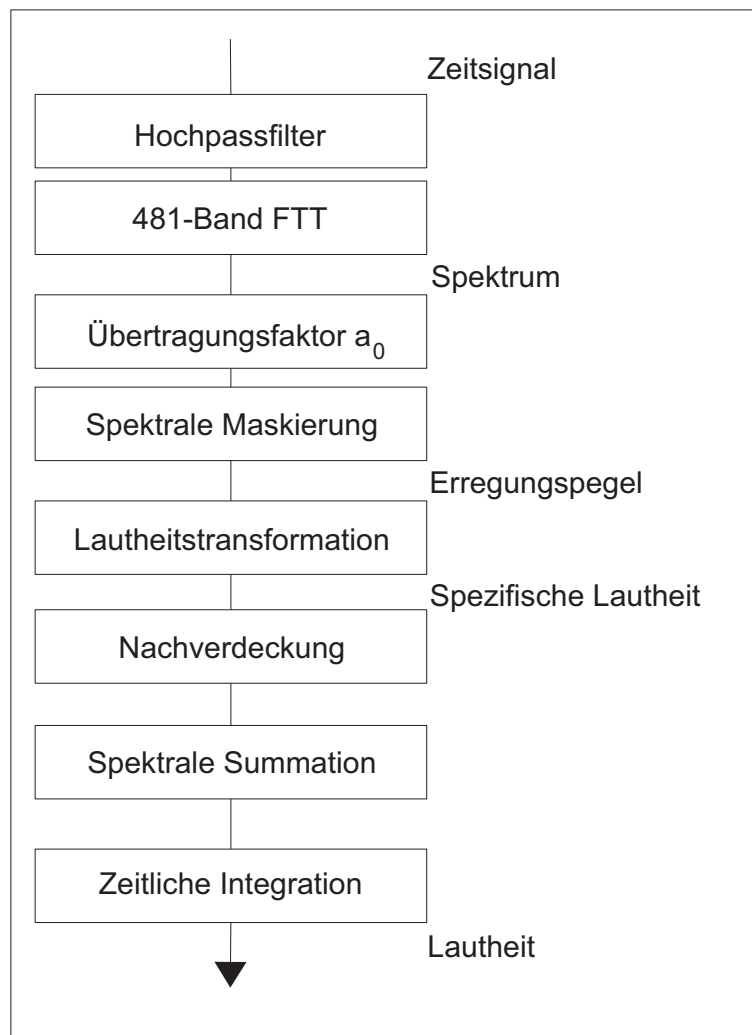


Abbildung 3.2: Blockdiagramm des modifizierten Lautheitsmodells mit erweiterter Frequenzauflösung und adaptiver spektraler Lautheitssummation.

le vorgenommen werden. Vergleiche mit experimentellen Daten der Lautheit zeigten, daß eine schalladaptive Summation, der für die einzelnen Kanäle ermittelten Lautheiten, die beste Übereinstimmung erbrachte. Die Werte spezifischer Lautheit liegen mit 481 Kanälen weit höher aufgelöst vor, als im ursprünglichen Modell, welches 24 Bark-breite Kanäle verwendet. Es müssen also zunächst Kanäle zu Frequenzgruppen zusammengefasst werden. Diese Zusammenfassung zu Frequenzgruppen erfolgt adaptiv in einer Weise, daß zunächst die dominante Frequenzgruppe gebildet wird, die den höchsten Wert spezifischer Lautheit beinhaltet. Symmetrisch um diesen Kanal größter spezifischer Lautheit werden 20 Kanäle zu einer Frequenzgruppe zusammengefasst. Die weiteren Frequenzgruppen ergeben sich durch jeweils benachbarte Zusammenfassung der nächsten 20 Kanäle. Innerhalb der so gebildeten Frequenzgruppen wird jeweils der Mittelwert der spezifischen Lautheiten als spezifische Lautheit der jeweiligen Frequenzgruppe gebildet. Die Gesamtlautheit ergibt sich wie im ursprünglichen grafischen Verfahren als Summe der Kanallautheiten der einzelnen Frequenzbänder.

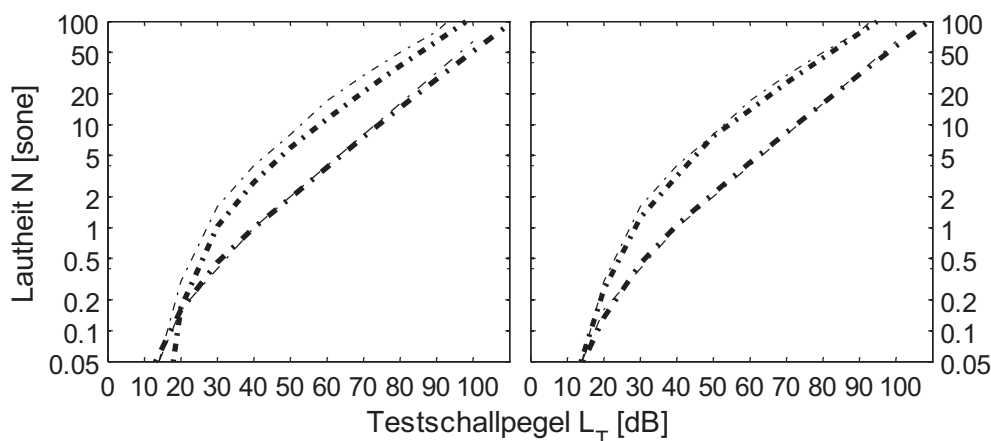


Abbildung 3.3: Berechnete Lautheit von gleichmäßig anregendem Rauschen (UEN, punktgestrichelte Linien) und Sinustönen der Frequenz 1 kHz (gestrichelt). Die dünnen Linien geben Vergleichsdaten aus [Zwi99] wieder, die dicken Linien entsprechen berechneten Daten der Lautheit, Links: Dynamic Loudness Model nach [Cha02b], Rechts: modifizierte Implementierung mit adaptiven Frequenzgruppengrenzen.

Die Lautheitstransformation wurde aus dem ursprünglichen Verfahren nach [Cha02b] übernommen. Ihre Parametrierung wurde jedoch hinsichtlich der Übereinstimmung der berechneten Ergebnisse mit Daten aus der Literatur optimiert. Die resultierenden berechneten Lautheiten für Sinustöne der Frequenz 1 kHz sowie gleichmäßig anregendes Rauschen [Zwi99] zeigt Abbildung 3.3.

Durch die Anpassung der Lautheitstransformation in Verbindung mit der adaptiven Lautheitssummutation konnte gegenüber dem ursprünglichen Algorithmus (Abbildung 3.3 links) eine Verbesserung insbesondere für breitbandige sowie sehr leise Schalle erreicht werden.

3.1.4 Resultierende Spektraldarstellung

Das Resultat der vorangehend beschriebenen Spektraltransformation mit anschließender Berechnung der Mithörschwellen und Übergang in die Lautheit ist eine Spektrogrammdarstellung, welche die wichtigsten Schalleigenschaften gehörgerecht repräsentiert. Dabei ist die Skalierung gehörgerecht im Sinne der Lautheitswahrnehmung und die Frequenzaufteilung orientiert sich an der bereits beschriebenen Bark-Skala.

Im folgenden Abschnitt werden anhand dieser Darstellung sowohl Spektrogramme als auch Lautheits-skalierte Spektren sowie Lautheits-Zeit Verläufe für die verschiedenen untersuchten Schalle dargestellt. Aus diesen Diagrammen werden qualitative Merkmale zur Modellierung der Eigenschaften der Ausgeprägtheit der Tonhöhe abgeleitet, die später in den Berechnungsalgorithmus zur Bestimmung der Ausgeprägtheit der Tonhöhe eingehen werden.

3.2 Zeitliche und spektrale Merkmale der Ausprägtheit der Tonhöhe

3.2.1 Sinustöne

Ein Mithörschwellenmuster von Sinustönen ist exemplarisch in Abbildung 3.4 für einen Ton mit einer Frequenz von 1000 Hz und einem Pegel von 60 dB dargestellt. Im linken Teil der Abbildung ist ein Spektrogramm dargestellt, welches das Ergebnis der Fourier-t-Transformation mit anschließender Maskierungsberechnung und Lautheitstransformation in der Tonheits-Zeit Ebene darstellt. Hier stellen dunklere Bereiche Werte höherer Lautheit dar, während helle Bereiche weniger große Lautheit repräsentieren.

Es ergibt sich für einen Ton mit einer Frequenz von 1000 Hz erwartungsgemäß ein Maximum bei 8,5 Bark, zu höheren Tonheiten hin erkennt man die gegenüber der unteren Flanke langsamer abfallende obere Flanke der spektralen Maskierung. Die mittlere Abbildung zeigt, wie in den folgenden Abschnitten, einen beliebigen Zeitausschnitt in Tonheits-Richtung durch das Spektrogramm. Es zeigt den aktuellen Verlauf der berechneten Mithörschwelle im Lautheitsmaß, wobei hier aufgrund der Lautheitsdarstellung der Abfall der oberen Flanke wesentlich steiler als in einer Pegeldarstellung erscheint. Auf der rechten Seite ist schließlich der zeitliche Verlauf der Lautheit des Signals dargestellt. Entsprechend bekannter Versuchsergebnisse ergibt sich eine Lautheit von ca. 4 sone für den untersuchten Sinuston mit einem Schallpegel von 60 dB [Zwi63a].

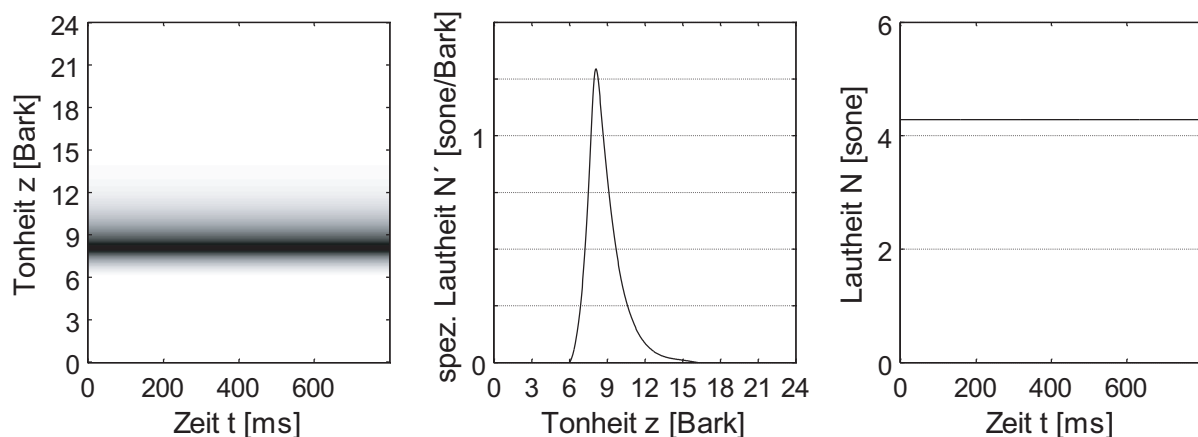


Abbildung 3.4: *Lautheitsskaliertes FTT-Spektrogramm, instantanes Spektrum sowie Lautheits-Zeit Verlauf eines Sinustons der Frequenz 1000 Hz, Pegel: 60 dB.*

Es ist nun ersichtlich, daß dieser Schall nur ein einziges spektrales Maximum hervorruft, dessen Position durch die verwendete Grundfrequenz bzw. deren Tonheit bestimmt ist. Weder die Tonhöhe noch die Lautheit des Schalls ändern sich über der Zeit, was in Einklang mit der Hörempfindung, die solch ein stationärer Sinuston hervorruft, ist. Da Sinus-

töne die jeweils höchste Ausgeprägtheit der Tonhöhe besitzen, bedeuten die beobachteten Merkmale für die Modellierung der Ausgeprägtheit der Tonhöhe, daß die Abwesenheit zeitlicher Schwankungen sowohl im Sinne spektraler Variationen (z.B. Frequenzmodulation) als auch Schwankungen der Lautheit (z.B. Amplitudenmodulation) zu einer Maximierung des berechneten Ergebnisses führen müssen.

3.2.2 Harmonisch komplexe Töne

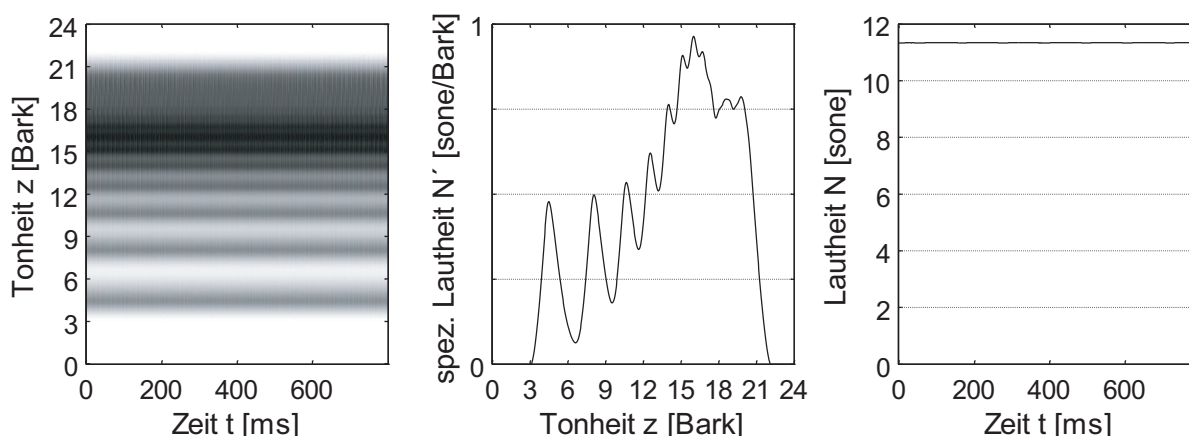


Abbildung 3.5: *Lautheitsskaliertes FTT-Spektrogramm, instantanes Spektrum sowie Lautheits-Zeit Verlauf eines harmonisch komplexen Tons. Grundfrequenz: 500 Hz, 1. bis 16. Harmonische, Pegel 60 dB.*

Abbildung 3.5 zeigt, in gleicher Weise wie Abbildung 3.4, das Lautheits-Tonheits-Zeitmuster, das Mithörschwellen-Muster sowie den Lautheits-Zeit Verlauf eines harmonisch komplexen Tons (Gesamtpegel $L = 60$ dB, gleicher Pegel aller Harmonischen). Bei der gewählten Grundfrequenz f_0 von 500 Hz sind neben der Grundfrequenz-Komponente in etwa die ersten acht Harmonischen als getrennte Maxima erkennbar. Darüber ist aufgrund des spektralen Auflösungsvermögens des Gehörs keine eindeutige Trennung der Harmonischen mehr möglich. Bei einer Tonheit von 18 Bark beträgt die Frequenzgruppenbreite in etwa 900 Hz, was deutlich breiter ist als der Abstand der einzelnen Harmonischen des dargestellten Signals. Vorangehend wurde die Wahrnehmbarkeitsschwelle für zwei getrennte Tonhöhen mit einem Abstand von 0,2-0,5 Bark angegeben. Dieses Verhalten spiegelt sich hier wieder, liegt doch der Abstand der Harmonischen mit 500 Hz im Bereich von $0,5 \cdot \Delta f_G$, entsprechend 450 Hz.

Der Verlauf der Lautheit bleibt, genau wie in Abschnitt 3.2.1 für Sinustöne gezeigt wurde, über der Zeit konstant. Die Grundfrequenz des harmonisch komplexen Tons wird als spektrale Komponente wiedergegeben. Zudem treten weitere Komponenten auf, die teilweise als Einzeltöne sichtbar sind, teilweise jedoch zu einer rauschähnlichen Komponente verschmelzen. In Form von Schwebungen (“Beats“) bleibt die Grundfrequenz für diese

nichtaufgelösten Spektralkomponenten als Periodentonhöhe erhalten [Mum98, Hor06], was im Rahmen dieser Arbeit jedoch nicht weiter berücksichtigt werden soll. Zwar ist auch der Lautheitsverlauf der nicht aufgelösten Harmonischen über der Zeit konstant, dennoch bilden diese eine Störkomponente undefinierter Tonhöhe.

Die Ergebnisse der Hörversuche in Kapitel 2.4 zeigten eine unmittelbare Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Anzahl der verwendeten Harmonischen. Mit steigender Anzahl Harmonischer zeigte sich hier eine sinkende Ausgeprägtheit der Tonhöhe. Zur Modellierung der Versuchsergebnisse läßt sich dieses Merkmal jedoch nicht heran zu ziehen. Zwar lassen sich durch Akzentuierung einzelne Teiltöne hervorheben [Sch97, Sch98, Cha97], im Allgemeinen werden jedoch Teiltöne nicht als Einzeltöne durch das menschliche Gehör wahrgenommen. Im Normalfall überwiegt die Wahrnehmung der virtuellen Tonhöhe und lediglich geübte Hörer sind, aufgrund analytischen Hörens, in der Lage, Teiltöne in komplexen Signalen wahrzunehmen [Ter92].

Die Parametrierung der verwendeten Fourier-t-Transformation hat zur Folge, daß auch die Transformationsergebnisse nur begrenzt Aufschluß über die enthaltenen Teiltöne liefern. Über eine Fourier-Transformation genügend hoher Auflösung ließen sich zwar alle Teiltöne ermitteln und so dem Algorithmus zuführen, im Sinne eines psychoakustisch motivierten Modells ist dieser Ansatz jedoch nicht zu verfolgen, da auf diese Weise Merkmale einfließen würden, die dem menschlichen Hörsystem nicht zur Verfügung stehen. Vielmehr ist zu vermuten, daß die wahrgenommene Tonhöhe durch zusätzliche Komponenten gestört wird, was sich im negativen Einfluss der höheren Harmonischen auf die Wahrnehmung der Tonhöhe und so auch der Ausgeprägtheit der Tonhöhe zeigt. Zur Beschreibung der Ausgeprägtheit der Tonhöhe von harmonisch komplexen Tönen muß also die (gedrosselte) Lautheit des Grundtons (bzw. der zu bewertenden Tonhöhe) mit dem umliegenden “Störgeräusch“ auf geeignete Art und Weise ins Verhältnis gesetzt werden.

3.2.3 Bandpassrauschen

Die Transformationsergebnisse für Schmal- und Bandpassrauschen sind in Abbildung 3.6 dargestellt. Der obere Teil zeigt dabei ein Schmalbandrauschen der Bandbreite 10 Hz, darunter ist ein Signal identischer Mittenfrequenz von 1000 Hz und Pegel von 60 dB, jedoch einer Bandbreite von 200 Hz dargestellt. Die jeweils linke und mittlere Abbildung sind dabei sehr ähnlich, die unterschiedliche Bandbreite geht also in diesen Fällen aus der Spektraldarstellung nicht unmittelbar hervor. Zudem fällt auf, daß insbesondere im, beliebig gewählten Ausschnitt des FTT-Spektrogramms das Rauschsignal nicht von einem Sinuston zu unterscheiden ist. Die Spektrogrammdarstellung und der Lautheits-Zeit Verlauf zeigen jedoch deutliche zeitliche Fluktuationen.

Aufgrund ihres statistischen Charakters weisen diese Signale eine zeitliche Variabilität auf. Der jeweilige Transformationsausschnitt (spektrales Mithörschwellen-Muster) stellt also ein instantanes Spektrum dar, welches für beliebige andere Zeitausschnitte zwar

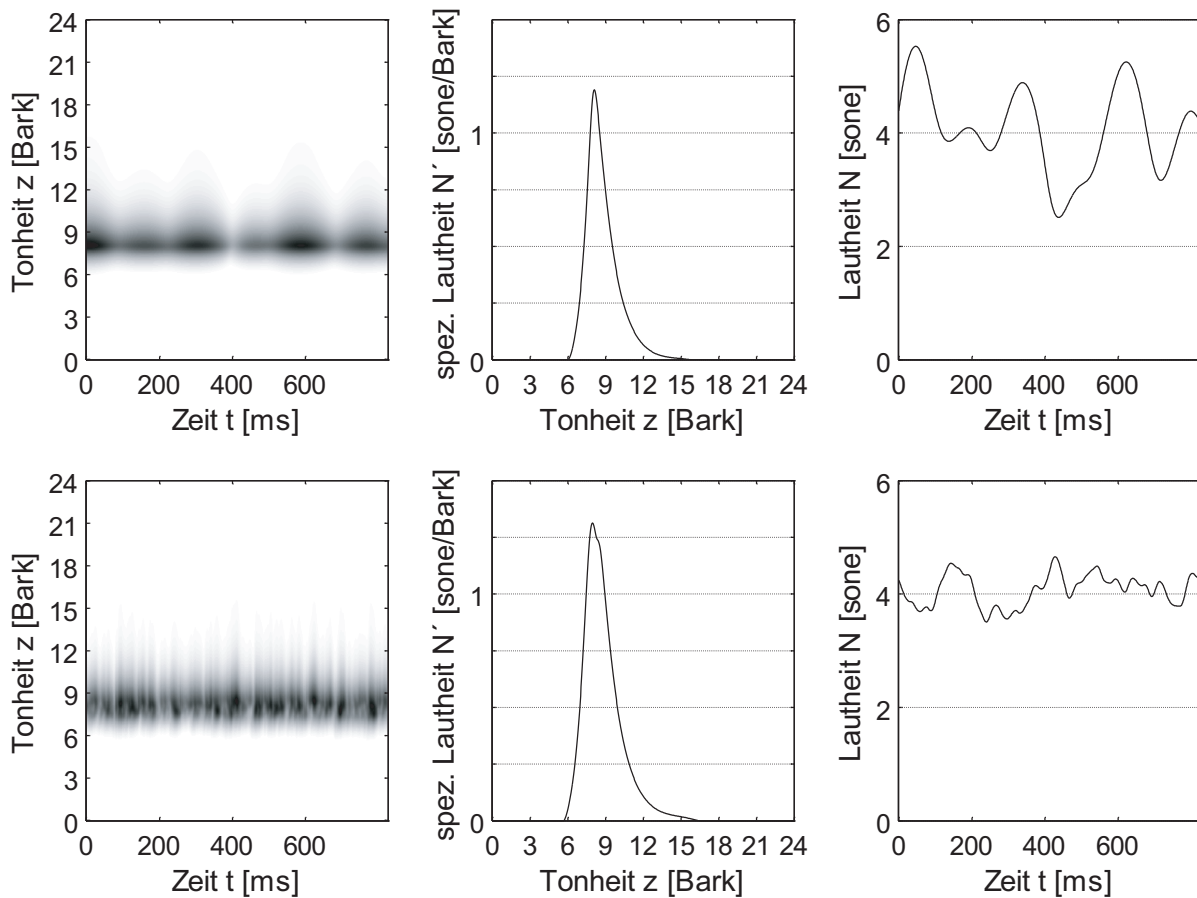


Abbildung 3.6: Lautheitsskaliertes FTT-Spektrogramm, instantanes Spektrum sowie Lautheits-Zeit Verlauf von Bandpassrauschen. Mittenfrequenz: 1000 Hz, Pegel 60 dB, Bandbreite Δf : Oben 10 Hz, Unten 200 Hz.

grundsätzlich ähnliche Eigenschaften aufweist, sich jedoch im Detail (Lage und Höhe der Maxima) deutlich unterscheiden kann. Es handelt sich hierbei also nicht um das Langzeitspektrum des dargestellten Schalls.

Die Eigenmodulation zeigt sich in deutlichen Lautheitsschwankungen, sichtbar in den jeweils rechten Teilbildern von Abbildung 3.6. Für das Signal mit einer Bandbreite von 10 Hz ergibt sich entsprechend Formel 2.3 eine Modulationsfrequenz von $f_{mod}^* = 6,4$ Hz, das untere Signal schwankt in etwa mit einer Modulationsfrequenz von 128 Hz. Die Modulation ist dabei jedoch nicht so regelmäßig wie bei deterministisch amplitudenmodulierten Signalen, die angegebene Modulationsfrequenz steht vielmehr für die mittlere Modulationsfrequenz über einen längeren Signalabschnitt. Gerade der Lautheitsverlauf der dargestellten Bandpassrauschen zeigt nun eine für die Modellierung auszuwertende Schalleigenschaft: Bei geringer Bandbreite findet sich eine langsame Schwankung hoher Modulationstiefe, was eine hohe Ausgeprägtheit der Tonhöhe zur Folge hat. Größere Bandbreiten führen dagegen zu schnellen Schwankungen mit geringer Tiefe. In Einklang mit [Fas92a] nimmt hier die

wahrgenommene Schwankungsstärke ab und die Wahrnehmungsgröße ‘‘Rauhigkeit‘‘ überwiegt, was für diese Schalle zu einer geringeren Ausgeprägtheit der Tonhöhe führt. Es müssen also Größen in die Modellierung eingehen, die die wahrnehmbare Schwankungsstärke und Rauhigkeit eines Schalls widerspiegeln.

Eine direkte Modellierung der Bandbreite eines Schalles wäre beispielsweise über ein 3 dB Kriterium oder ein generelles Kriterium der spezifischen Lautheit N' über einen Faktor x möglich. So wäre beispielsweise ein Ausschlußkriterium der Form $N' > xN'(f_0)$ denkbar. Hier würden nur Anteile des Spektrums Berücksichtigung finden, deren spezifische Lautheit N' gegenüber der spezifischen Lautheit an der Stelle der Mittenfrequenz bzw. der wahrgenommenen Tonhöhe $N'(f_0)$ um weniger als einen Faktor x abgesunken ist.

Genauere Betrachtungen der Spektrogramme in Abbildung 3.6 zeigen, daß bei Schmal- und Bandpassrauschen die Position des Maximums innerhalb eines Zeitschritts nicht konstant ist und daß zudem mehrere Maxima auftreten können. Wird über ein einfaches lineares Abstandsmaß dasjenige Maximum ermittelt, welches der zu bewertenden Tonhöhe f_0 am Nächsten liegt, so zeigt sich, daß die Position dieses Maximums POS_{f_0} nun ebenfalls schwankt. Diese Schwankung läßt sich als Maß der Bandbreite heranziehen, da die möglichen Positionen dieses Maximums mit steigender Bandbreite eine größere Variabilität aufweisen. Da die Schwankungen nicht regelmäßig erfolgen, wie dies bei frequenzmodulierten Sinustönen der Fall wäre, läßt sich keine Modulationsfrequenz im Sinne einer Frequenzmodulation ermitteln. Vielmehr bietet es sich an, die Standardabweichung der Menge der Positionen der zu f_0 gehörenden Maxima $[POS_{f_0}]$ als Maß der Bandbreite des Bandpassrauschens heranzuziehen. Es gilt also der Zusammenhang

$$\Delta f \sim \sigma([POS_{f_0}]) \quad (3.3)$$

Zu den bereits für Sinustöne und harmonisch komplexen Tönen beschriebenen Merkmalen Lautheitsschwankung der f_0 -Komponente und Lautheit zusätzlicher aufgelöster und nicht aufgelöster Spektralkomponenten, sind nun also noch die Varianz der Grundfrequenzkomponente sowie die zeitliche Schwankung des Lautheitsverlaufs hinzuzuzählen.

3.2.4 Kammfilterrauschen

Wiederholungsrauschen, die durch die mehrfache Verwendung von Kammfiltern entstehen (siehe Kapitel 2.7), zeigen ähnliche spektrale Eigenschaften wie harmonisch komplexe Töne, jedoch sind die spektralen Maxima aufgrund des statistischen Eingangssignals stärkeren zeitlichen Schwankungen unterworfen. Dies führt dazu, daß die wahrgenommene Tonhöhe dieser Schalle zwar mit der eines harmonisch komplexen Tons vergleichbarer spektraler Zusammensetzung nahezu übereinstimmt, jedoch nicht so stark ausgeprägt ist.

Die Diagramme in Abbildung 3.7 zeigen das Verhalten für unterschiedlich häufiges Durchlaufen der Filterstrukturen bei ansonsten identischen Parametern ($d = 2$ ms, 1. bis 8. Harmonische). In der oberen Reihe werden zwei Iterationen verwendet, in der unteren Reihe

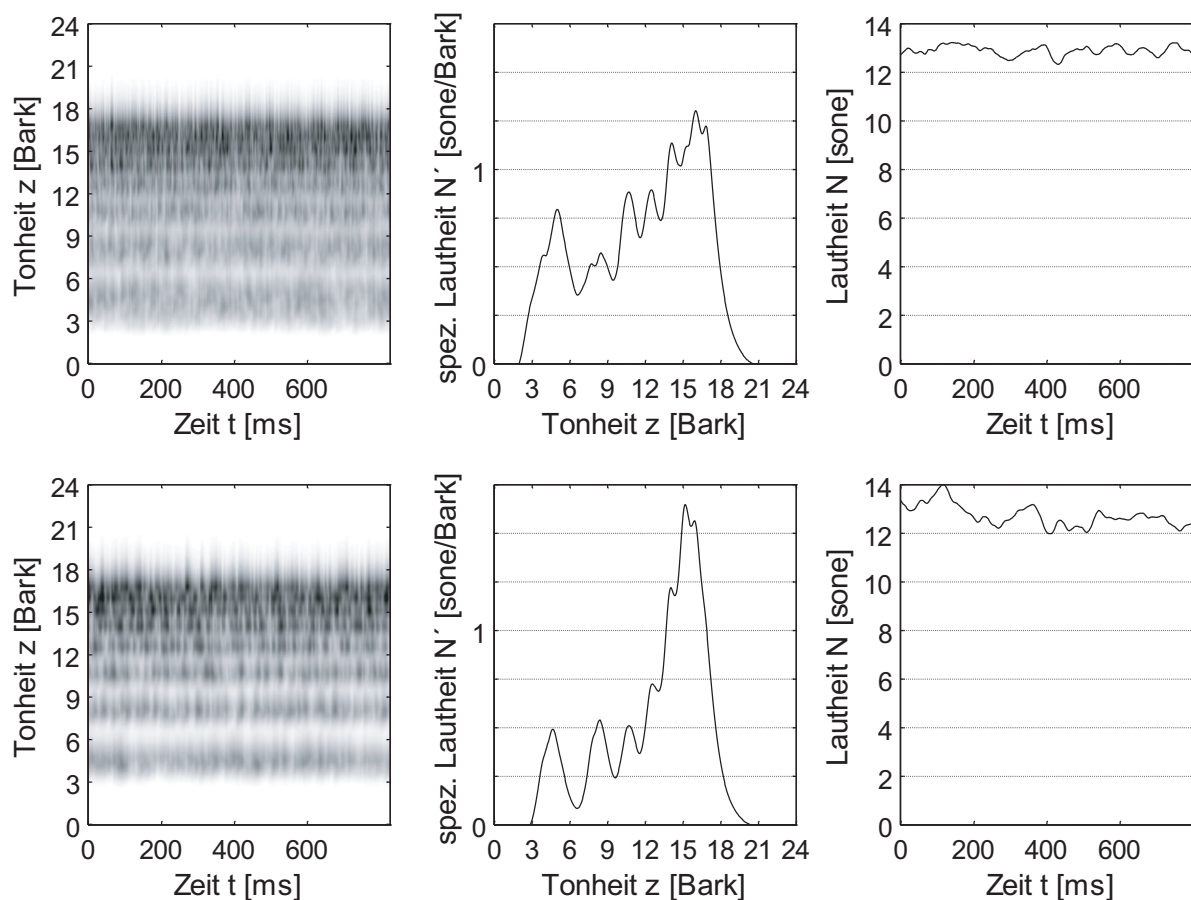


Abbildung 3.7: Lautheitsskaliertes FTT-Spektrogramm, instantanes Spektrum sowie Lautheits-Zeit Verlauf von Iterated Rippled Noise. Verzögerungszeit: 2 ms, 1. bis 8. Harmonische, Oben: 2 Iterationen, Unten: 16 Iterationen.

dagegen 16. Die zugehörigen spezifischen Lautheits-Tonheits Muster zeigen ein deutlicheres Hervortreten der spektralen Maxima bei Erhöhung der Anzahl der Iterationen. Vergleichbar mit den Transformationsergebnissen von harmonisch komplexen Tönen in Abschnitt 3.2.2 werden höhere Harmonische unabhängig von der verwendeten Anzahl an Iterationen nicht mehr als getrennte Komponenten aufgelöst, sondern verschmelzen zu einem Störanteil. Die zeitliche Schwankung des Lautheitsverlaufs ist dagegen in beiden dargestellten Fällen recht ähnlich und trägt offenbar keine Information, die zur Modellierung der Ausgeprägtheit der Tonhöhe dieser Signale beitragen könnte.

Die bereits vorangehend eingeführten Merkmale der Ausgeprägtheit der Tonhöhe sind also auch bei Wiederholungsrauschen zu finden und tragen so zur Beschreibung der zugehörigen experimentellen Daten innerhalb eines Modells bei. Darüber hinaus ist eine Beschreibung der Güte der spektralen Maxima nötig, die mit steigender Anzahl an Iterationen deutlicher werden. Dadurch wird das Spektrum von Iterated Rippled Noise dem harmonisch komplexer Töne ähnlicher und die Ausgeprägtheit der Tonhöhe steigt an.

3.2.5 Harmonisch komplexe Residualtöne

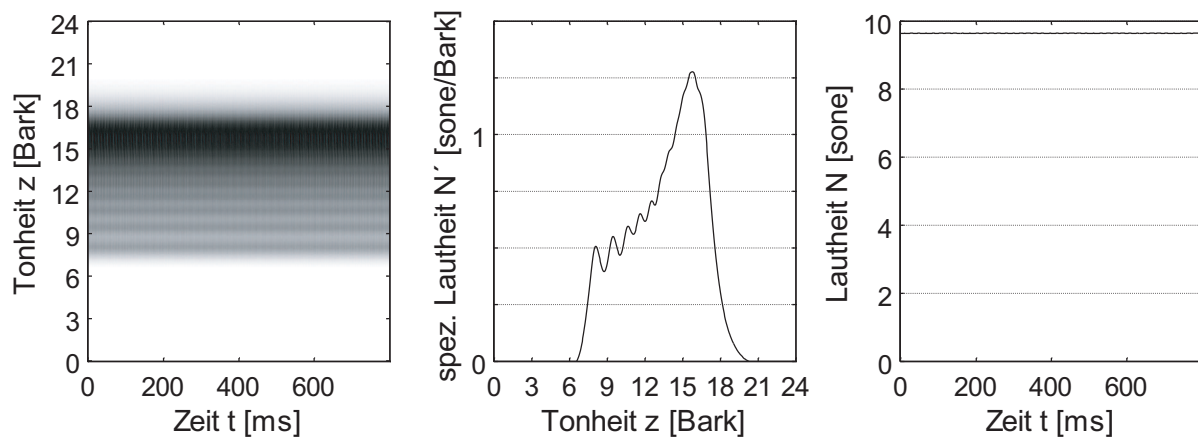


Abbildung 3.8: Lautheitsskaliertes FTT-Spektrogramm, instantanes Spektrum und Lautheits-Zeit Verlauf eines harmonisch komplexen Residualtons, dessen Tonhöhe als virtuelle Tonhöhe wahrgenommen wird. $f_0 = 250$ Hz, 4. bis 16. Harmonische.

Das Spektrogramm eines harmonisch komplexen Residualtons ist in Abbildung 3.8 für drei fehlende untere Harmonische dargestellt. Das Signal und auch die Transformationsergebnisse enthalten keinerlei Komponente bei dieser fehlenden Grundfrequenz. Für die Modellierung der Ausprägtheit der Tonhöhe muß für diese Schalle also die fehlende Grundfrequenz-Komponente errechnet werden.

Dazu schlägt Terhardt [Ter79] einen Algorithmus vor, der aus spektralen Daten mögliche virtuelle Tonhöhen und deren "Gewicht" berechnet. Diese Gewichte beziehen sich auf das Hervortreten einzelner Teiltöne und werden als "Pitch Saliency" bezeichnet. Neben den bislang aufgeführten Merkmalen werden also zur Modellierung der Ausprägtheit der virtuellen Tonhöhe die errechneten Tonhöhengewichte virtueller Tonhöhen herangezogen.

3.3 Zusammenfassung

Aus den vorangehend dargestellten Daten wurden jeweils Merkmale bestimmt, die sich zur qualitativen Beschreibung der gewonnenen Daten der Ausprägtheit der Tonhöhe eignen. Die Merkmale, die sich für die untersuchten Schallarten aus den Spektrogrammdarstellungen ermitteln lassen, sind in Tabelle 3.1 zusammenfassend dargestellt. Sie umfassen physikalische Größen wie die Frequenz und Bandbreite von Schallen, aber auch psychoakustische Merkmale in Form der Lautheit des Signal- und Störanteils, sowie der Schwankung der Amplitude und Frequenz der wahrnehmbaren Tonhöhe. Für jede Schallart ergibt sich dabei eine charakteristische Zuordnung von Merkmalen, die die jeweiligen Versuchsergebnisse qualitativ beschreiben.

Signal	Tonhöhe bzw. f_0	Bandbreite Δf	Lautheit		Schwankung f_0	
			f_0	Störanteil	N(t)	f(t)
SIN	X	O	X	O	O	O
AM-SIN	X	X	X	O	X	O
FM-SIN	X	X	X	O	O	X
HKT	X	X	X	X	O	O
IRN	X	X	X	X	X	X
BPR	X	X	X	X	X	X
HKRT	X	X	X	X	O	O

Tabelle 3.1: Zusammenfassung der charakterisierenden Merkmale verschiedener Schallarten bezüglich der Wahrnehmung Ausgeprägtheit der Tonhöhe. "X": Merkmal vorhanden, "O": Merkmal nicht vorhanden.

Daraus läßt sich nun ein Modell der Ausgeprägtheit der Tonhöhe ableiten, das unter zuhilfenahme von Fourier-t-Transformation und virtueller Tonhöhe die prinzipiellen Zusammenhänge der experimentellen Daten nachvollziehen kann. Tabelle 3.2 stellt die Merkmale und mögliche Ansätze ihrer Modellierung dar. Die weitere Modellbildung wird nachfolgend in Kapitel 4 detailliert beschrieben.

Merkmal	Umsetzung
Tonhöhe bzw. f_0	Detektion der Grundfrequenz
Bandbreite Δf	Ermittlung der Anzahl angeregter Kanäle im Spektrum
Lautheit des tonalen Schallanteils	Berechnung der Lautheit des Schallanteils, der die Tonhöhe repräsentiert
Lautheit des störenden Schallanteils	Berechnung der Lautheit des Schallanteils zusätzlich zur Grundton-Komponente
Zeitliche Schwankung der Lautheit	Ermittlung der Schwankung des Lautheitsverlaufs des Signals über der Zeit
Zeitliche Schwankung der Tonhöhe	Bestimmung der Grundfrequenz-Komponente in jedem Analyse Zeitschritt und daraus Bestimmung der Schwankungen
virtuelle Tonhöhe	Berechnung mittels Algorithmus zur Beschreibung der Tonhöhenwahrnehmung
Gewicht der virtuellen Tonhöhe	Berechnete "Pitch Saliency" aus dem Algorithmus zur Beschreibung der Tonhöhenwahrnehmung

Tabelle 3.2: Mögliche Ermittlung der charakterisierenden Merkmale bezüglich der Ausgeprägtheit der Tonhöhe nach Tabelle 3.1.

Kapitel 4

Ein Modell der Ausgeprägtheit der Tonhöhe

Voranehend wurden Merkmale bestimmt, die die Abhängigkeiten der Ausgeprägtheit der Tonhöhe über Plausibilitäten beschreiben können. Um nun einen Algorithmus zur Berechnung der Ausgeprägtheit der Tonhöhe zu erstellen müssen diese Merkmale aus einem Schallsignal extrahiert und in geeigneter Weise miteinander verrechnet werden. Mit Hilfe bereits etablierter Algorithmen und Modelle lassen sich Teile dieser Merkmale ermitteln [Fru05a]. Darüberhinaus sind weitere physikalisch und psychoakustisch motivierte Größen nötig um ein Modell zur Beschreibung der Wahrnehmung “Ausgeprägtheit der Tonhöhe“ zu erstellen [Fru06].

4.1 Übersicht über das Modell der Ausgeprägtheit der Tonhöhe

Wie bereits in Abschnitt 2.4 erläutert wurde, stellen zusätzlich zur Grundfrequenz vorhandene Schallkomponenten einen hemmenden Einfluß auf die wahrgenommene Ausgeprägtheit der Tonhöhe dar. Eine prinzipielle Modellierung der Ausgeprägtheit der Tonhöhe erfolgt also über ein geeignetes Verhältnis der Lautheit des Signalanteils N_0 zur Lautheit des Störanteils N_N der Form

$$AdT \sim \frac{N_0}{N_N} \cdot W_i , \quad (4.1)$$

also durch eine mit einem Signal/Rausch Verhältnis vergleichbare Größe. Darüberhinaus sind weitere Gewichtungen W_i aufgrund der ermittelten Einflußgrößen nötig.

Über diesen einfachen Zusammenhang wird sichergestellt, daß Sinustöne jeweils die maximale Ausgeprägtheit der Tonhöhe hervorrufen, da hier Signal- und Störanteil identisch sind. Zudem wird das Verhalten harmonisch komplexer Töne nachgebildet, daß eine steigende Anzahl an Komponenten zu sinkender Ausgeprägtheit der Tonhöhe führt, da hier der Störanteil wächst, während die f_0 -Komponente nahezu konstant bleibt.

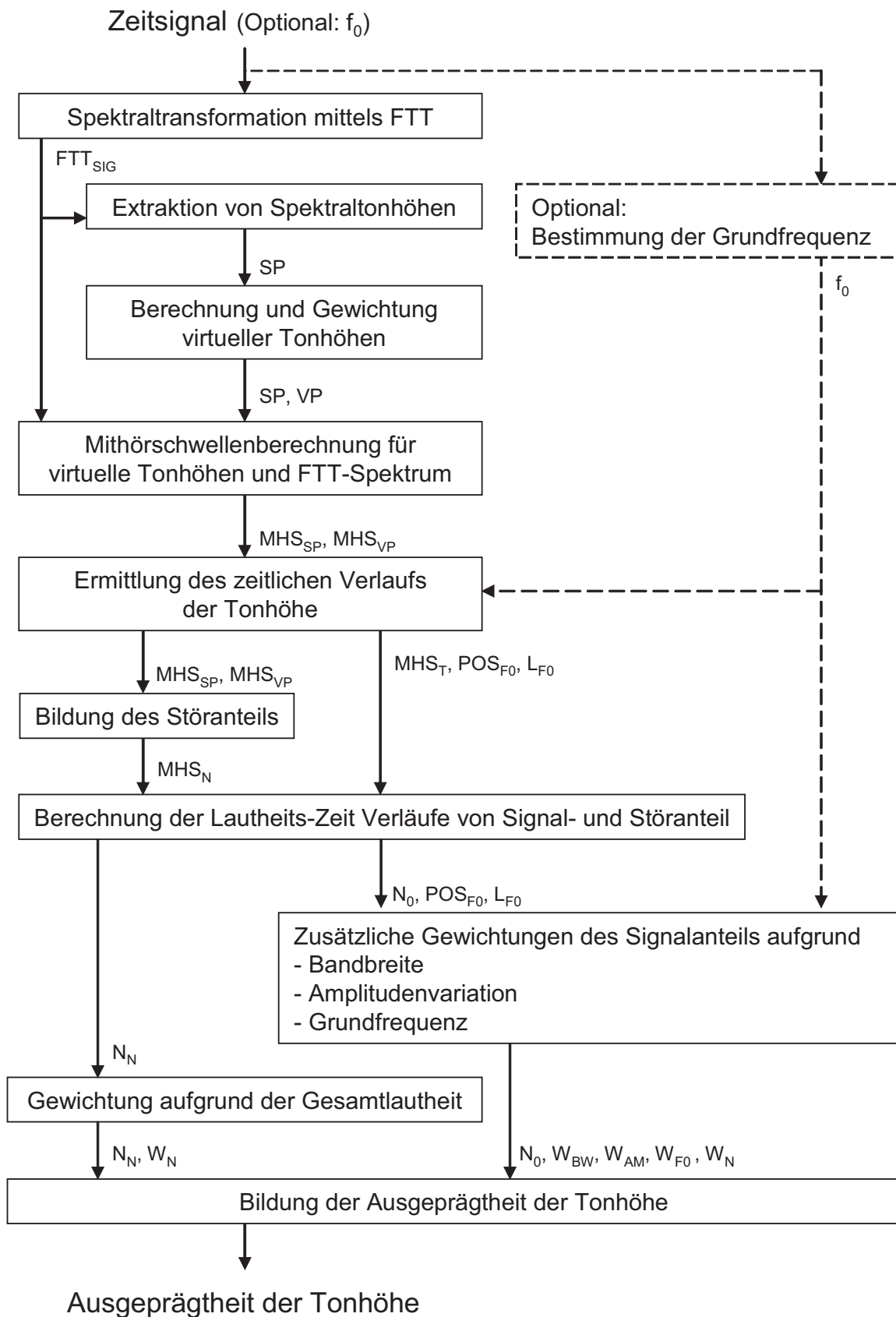


Abbildung 4.1: Blockschaltbild des ausgehend von den experimentellen Untersuchungen erstellten Modells der Ausprägtheit der Tonhöhe.

Um die experimentell gewonnenen Daten zu beschreiben wurden in Kapitel 2 Erklärungsansätze aufgrund psychoakustischer und physiologischer Eigenschaften aufgezeigt. Die unterschiedlichen Einflußgrößen werden nun zu einem Modell kombiniert, das über verschiedene Schallarten hinweg Vorhersagen über die Ausgeprägtheit der Tonhöhe zuläßt. Eine Übersicht über dieses Modell ist in Abbildung 4.1 gezeigt.

Die ersten Blöcke des Modells bestehen dabei aus Verfahren die im Rahmen der Arbeiten von Terhardt entstanden sind [Ter82, Ter85a, Ter86]. Die zur gehörgerechten Spektralanalyse verwendete Fourier-t-Transformation (FTT) wurde bereits in Abschnitt 1.4 erläutert. Sie liefert Spektrogramme, die in Tonheits-skalierten Zeitschritten jeweils Pegelspektren des analysierten Schalls enthalten. Die nachfolgende Extraktion von Spektraltonhöhen bildet über Maximumskriterien aus den FTT-Spektren einzelner Zeitschritte die darin enthaltenen Teiltöne unter Berücksichtigung von Verschiebungseffekten der wahrgenommenen Tonhöhe nach. Mithilfe des Modells der virtuellen Tonhöhe aus Abschnitt 1.1 werden aus den Werten spektraler Tonhöhen virtuelle Tonhöhen gewonnen.

Über die Bildung eines Anteils, der die Grundfrequenzkomponente repräsentiert, lassen sich die Daten in den Störanteil sowie den Signalanteil aufspalten. Nachdem für beide Anteile die korrespondierenden Mithörschwellenmuster berechnet wurden können über das modifizierte Berechnungsverfahren der Lautheit aus Kapitel 3.1.3 die beiden Größen N_0 und N_N bestimmt werden.

Zur weiteren Anpassung des Modells an experimentell ermittelte Werte werden zusätzliche Gewichtsfunktionen bestimmt, die sich an psychoakustischen Empfindungsgrößen (Lautheit, Schwankungsstärke, Rauigkeit, subjektive Dauer) und deren Modellierung orientieren. Dazu wird ein Gewicht W_{BW} zur Beschreibung der Bandbreite, ein Gewicht W_{AM} , welches Einflüsse der psychoakustischen Größen "Schwankungsstärke" und "Rauigkeit" beschreibt, sowie ein Gewicht W_{F0} zur Anpassung an die spektrale Lage der zu bewertenden Tonhöhe eingeführt.

Insgesamt ergibt sich die Ausgeprägtheit der Tonhöhe über den folgenden Zusammenhang:

$$AdT \sim N^y W_{BW} W_{AM} W_{F0} \frac{N_0}{N_N}. \quad (4.2)$$

Es handelt sich also um ein Verhältnismaß, das durch zusätzliche Gewichtungen weitere Einflußgrößen der Schallwahrnehmung berücksichtigt. Hervorzuheben ist dabei nochmals die psychoakustische Motivation der einzelnen Gewichte. Ein Modell, das all diese Einflußgrößen berücksichtigt wird zunächst in der Lage sein, die experimentell gewonnenen Daten nachzuvollziehen, aber auch Vorhersagen für weitere Schalle sollten sinnvolle Ergebnisse liefern. Auf diese Weise entsteht also nicht lediglich ein Rechenschema zur Berechnung bekannter Daten sondern ein Modell zur allgemeinen Beschreibung der Ausgeprägtheit der Tonhöhe verschiedenster Schalle und Geräusche. Nachfolgend werden die einzelnen Blöcke des in Abbildung 4.1 dargestellten Modells beschrieben und erläutert.

4.2 Spektraltransformation mittels FFT

Das Eingangssignal des Modells bildet ein kalibrierter Vektor, welcher das Zeitsignal des zu untersuchenden Schalls beinhaltet. Darüberhinaus kann optional die Grundfrequenz des Schalls übergeben werden, falls diese nicht automatisch detektiert werden soll. Die Transformation des Zeitsignals in Spektrogrammdateien erfolgt über die beschriebene Fourier-t-Transformation unter Verwendung der in Abschnitt 3.1.1 aufgeführten Parameter. Das Resultat der Transformation sind Pegelwerte $FTT_{SIG}(z, t)$ die entsprechend der Tonheits- und Zeitaufösung in einer Matrix angeordnet sind.

4.3 Extraktion von Spektraltonhöhen

Spektraltonhöhen sind tonale Schallanteile, die als physikalische Komponenten im Schallsignal enthalten sind und unter bestimmten Voraussetzungen wie Akzentuierung [Sch98] als Teilton wahrnehmbar sind. Sie lassen sich in Spektraldarstellungen als lokale Überhöhungen erkennen und aufgrund einer Maximumsdetektion extrahieren. Die Spektraldarstellung kann dabei prinzipiell verschiedensten Transformationen entstammen, so wurde beispielsweise bei Heinbach zur Ermittlung des Teilton-Zeit Musters (TTZM) eine Fourier-Transformation zur Spektralanalyse herangezogen um den Verlauf einzelner Teiltöne über der Zeit zu bestimmen [Hei88a, Hei88b].

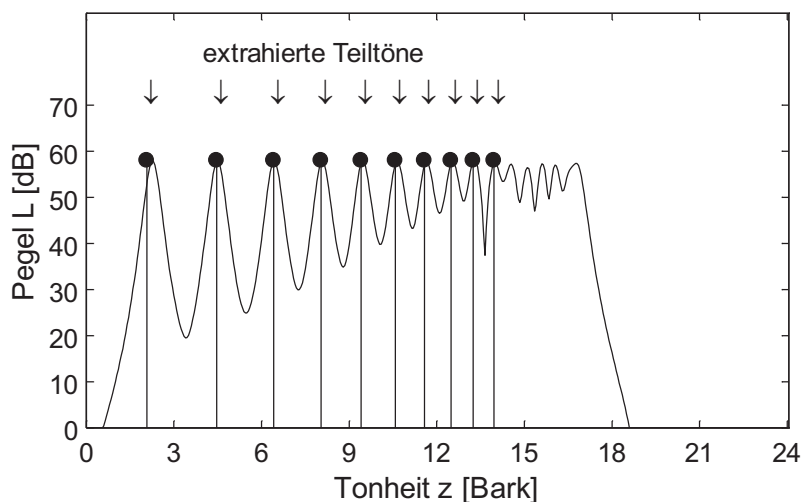


Abbildung 4.2: Extrahierte Teiltöne eines harmonisch komplexen Tons der Grundfrequenz 250 Hz, 1. bis 16. Harmonische.

Die verwendete Implementierung der Fourier-t-Transformation stellt diese Maximumsextraktion zur Verfügung. Aus dem Ergebnis der Fourier-t-Transformation des Eingangssignals $FTT_{SIG}(z, t)$ werden für jeden Zeitschritt der Spektraltransformation die ermittelten Maxima in Form von Frequenz f_{SP} und Pegel L_{SP} ausgegeben. Dabei wird über

die Geradennäherungen aus Abschnitt 3.1.2 ein einfaches Maskierungskriterium benutzt, um unter der Mithörschwelle liegende Maxima des FTT-Spektrogramms zu verwerfen.

Die auf diese Weise extrahierten Teiltonkandidaten werden dem in [Ter82] beschriebenen Modell zugeführt. Dabei handelt es sich um ein etabliertes Modell der Tonhöhenwahrnehmung, das in der Lage ist, die vielfältigen Effekte der Tonhöhenwahrnehmung einschließlich zahlreicher Verschiebungseffekte zu beschreiben [Ter70]. Es entstehen also ausgehend von den Frequenzen und Pegeln der physikalischen Teiltöne Daten über die Tonhöhe dieser Teiltöne. Das Ergebnis des verwendeten Algorithmus ist beispielhaft in Abbildung 4.2 für einen harmonisch komplexen Ton dargestellt. Von den 16 enthaltenen Harmonischen werden 11 als Teiltöne extrahiert.

Es entsteht so die Menge $SP(f_{SP}, L_{SP}, t_{SP})$ der Frequenzen und Pegel extrahierter Teiltöne in den einzelnen Zeitschritten. Diese repräsentieren die Spektraltonhöhen, die vom Gehör als einzelne Teiltöne wahrgenommen werden können.

4.4 Berechnung und Gewichtung virtueller Tonhöhen

Über den Algorithmus zur Ermittlung der “Pitch Saliency“ aus [Ter82] werden für die extrahierten Spektraltonhöhen $SP(f_{SP}, L_{SP}, t_{SP})$ Tonhöhengewichte W_{SP} berechnet. Darüberhinaus werden aufbauend auf dem Modell der “Subharmonischen Koinzidenzen“ virtuelle Tonhöhen berechnet, für die entsprechend ihrer spektralen Lage und Anzahl der Koinzidenzen als “Pitch Saliency“ bezeichnete Gewichte W_{VP} berechnet werden. Die Berechnung dieser Gewichte wird aus den zur Verfügung stehenden Programmteilen übernommen [Mum97], wodurch die Menge $VP(f_{VP}, W_{VP}, t_{VP})$ der Tonhöhenkandidaten und Gewichte der berechneten virtuellen Tonhöhen gebildet wird.

Zusätzlich zu den spektralen Tonhöhen stehen also die Kandidaten virtueller Tonhöhen in Form von Tonhöhen und Gewichten zur Verfügung. Die ermittelte “Pitch Saliency“ beschreibt dabei zwar die Bedeutung eines Teiltons für die Wahrnehmung eines komplexen Tones, unterscheidet sich im Detail jedoch deutlich von der Wahrnehmung “Ausgeprägtheit der Tonhöhe“. Zur Verdeutlichung dieser Diskrepanz werden nun einige Daten der Ausgeprägtheit der Tonhöhe im Vergleich zu berechneten Tonhöhengewichten dargestellt.

4.4.1 Gegenüberstellung von Ausgeprägtheit der Tonhöhe und “Pitch Saliency“

In der englischsprachigen Literatur werden die beiden Begriffe “Pitch Saliency“ und “Pitch Strength“, also “Ausgeprägtheit der Tonhöhe“ gelegentlich als synonym behandelt [Peg04], was jedoch insbesondere für die durch Terhardt’s Algorithmus ermittelten Werte nicht korrekt ist. Diese Vermischung der beiden Begriffe wird bei Terhardt jedoch stets vermieden.

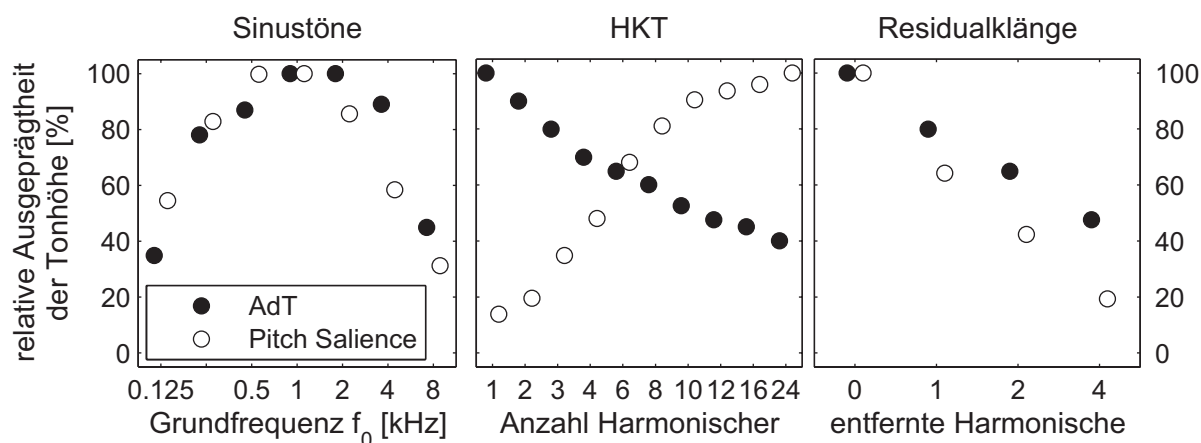


Abbildung 4.3: Berechnete “Pitch Saliency“ im Vergleich zur Ausgeprägtheit der Tonhöhe verschiedener Schallarten ($L_T = 60$ dB, $T_T = 1$ s): Sinus: Abhängigkeit von der Grundfrequenz (links), Harmonisch komplexer Ton: Abhängigkeit von der Anzahl Harmonischer (Mitte) und von der Anzahl entfernter unterster Harmonischer (rechts).

Um die Unterschiede und Gemeinsamkeiten der Größen “Ausgeprägtheit der Tonhöhe“ und “Pitch Saliency“ hervorzuheben werden in Abbildung 4.3 berechnete Daten der “Pitch Saliency“ mit experimentellen Ergebnissen der Ausgeprägtheit der Tonhöhe gegenüber gestellt. Die Abhängigkeiten für die Spektraltonhöhen von Sinustönen und der virtuellen Tonhöhe von harmonisch komplexen Tönen lassen sich über das Modell der “Pitch Saliency“ in Grundzügen nachvollziehen. Die Ausgeprägtheit der Spektraltonhöhe von harmonisch komplexen Tönen weicht jedoch von den berechneten Werten der “Pitch Saliency“ ab. So steigt diese bei Erhöhung der Anzahl Harmonischer an, während die Versuchsergebnisse zur Ausgeprägtheit der Tonhöhe entsprechend Abschnitt 2.4 zeigen, daß eine höhere Anzahl Harmonischer zu einer Verringerung der Ausgeprägtheit der Tonhöhe führt.

Dieser Unterschied läßt sich über das zugrundeliegende Modell und Berechnungsschema nachvollziehen: Zur Berechnung der “Pitch Saliency“ werden die Subharmonischen jeder einzelnen Komponente gebildet und mit denjenigen der anderen Teiltöne überlagert. Bei harmonisch komplexen Tönen entstehen demzufolge bei der tatsächlichen Grundfrequenz sehr viele Subharmonische, deren Gewichte näherungsweise aufsummiert werden. So tragen weitere Harmonische zu einem Anstieg der “Pitch Saliency“ bei, während die experimentellen Daten zur Ausgeprägtheit der Tonhöhe eine deutliche Abnahme dieser Empfindungsgröße bei Zunahme der Teiltöne zeigen.

Für das erstellte Modell der Ausgeprägtheit der Tonhöhe werden aus diesem Grund zwar die berechneten Gewichte der Tonhöhen weiter verwendet, zur Modellierung der Ausgeprägtheit der Tonhöhe sind jedoch weitere Verarbeitungsschritte notwendig. Zunächst müssen die Gewichte der virtuellen Tonhöhen in eine Pegeldarstellung zurücktransformiert werden, um später mit den spektralen Komponenten verglichen werden zu können.

4.4.2 Rücktransformation von Tonhöhengewichten in das Pegelmaß

Zur Berechnung der Ausprägtheit der Tonhöhe soll wie in Abschnitt 4.1 ein Verhältnis der Lautheit der zu bewertenden Tonhöhe zum Störanteil eines Schalls herangezogen werden. Im Fall virtueller Tonhöhen liegen jedoch keine Pegelwerte für die Komponente der Tonhöhe vor, da diese nicht als physikalische Komponente im Schall enthalten ist. Es liegt für virtuelle Tonhöhen die Menge $VP(f_{VP}, W_{VP}, t_{VP})$ vor, die für alle Kandidaten virtueller Tonhöhen Frequenz f_{VP} und Gewicht W_{VP} entsprechend dem Analysezeitschritt t_{VP} enthält. Um nun spektrale und virtuelle Tonhöhen gleichermaßen bewerten zu können müssen die Gewichte virtueller Tonhöhen in das Pegelmaß zurück transformiert werden.

Wird vorausgesetzt, daß sich die Gewichte spektraler Komponenten W_{SP} in einen Pegel L_{SP} zurückrechnen lassen, kann unter Zuhilfenahme eines Referenzgewichts $W_{0,SP}$ und der frequenzabhängigen Korrekturgröße C_{f_0} der Pegel einer spektralen Komponente über

$$L_{SP} = 20 \cdot \log_{10} \left(\frac{W_{SP}}{W_{0,SP}} \right) \cdot C_{f_0}(f_0) \quad (4.3)$$

berechnet werden. Das Referenzgewicht $W_{0,SP}$ entspricht dabei der zur Umrechnung von Schalldruck in Schallpegel benötigten Größe p_0 (Referenzschalldruck). Es ergibt sich aus dem Maximum des Pegelspektrums des aktuellen Zeitschritts $\max(FTT)$ und der noch zu definierenden Grundfrequenz-abhängigen Korrekturgröße C_{f_0} zu

$$W_{0,SP} = \frac{\max(W_{SP})}{10^{\left(\frac{\max(FTT)}{20 \cdot C_{f_0}}\right)}} \quad (4.4)$$

Durch Anpassung berechneter Daten an die Ergebnisse der Ausprägtheit der virtuellen Tonhöhe ergibt sich der Faktor C_{f_0} zu

$$C_{f_0}(f_0) = \frac{3}{2500} \cdot \frac{f_0}{[Hz]} + \frac{3}{5} \quad (4.5)$$

Aus diesen Größen werden die Gewichte virtueller Tonhöhen W_{VP} , analog zu Formel 4.3, über den Zusammenhang

$$L_{VP}^* = 20 \cdot \log_{10} \left(\frac{W_{VP}}{W_{0,SP}} \right) \cdot C_{f_0}(f_0) \quad (4.6)$$

in Pegelwerte L_{VP}^* umgerechnet. Diese Umrechnung erfolgt für alle Werte virtueller Tonhöhen $VP(f_{VP}, W_{VP}, t_{VP})$ aller Zeitschritte der ermittelten Daten virtueller Tonhöhen.

Die virtuellen Tonhöhen lassen sich nun über die Pegelwerte L_{VP}^* und die vorangehend ermittelten Tonhöhen in ein Spektrogramm virtueller Tonhöhen FTT_{VP} eintragen und repräsentieren hier sinusartige Schallanteile, welche die entsprechenden Tonhöhen hervorrufen und aufgrund des ermittelten Pegels eine ähnliche Lautheit besitzen sollten wie die tatsächlich wahrgenommenen virtuellen Tonhöhen.

Es liegen nun also Tabellen der virtuellen und spektralen Tonhöhen vor, sowie das FTT-Pegelspektrogramm des Eingangssignals, als auch ein Spektrogramm, das die virtuellen Tonhöhen in Form von Pegelwerten an diskreten Stellen beinhaltet. Durch Zusammenfassen der Tonhöhenwerte f_{VP} und Pegel L_{VP}^* der einzelnen Zeitschritte t entsteht für die virtuellen Tonhöhen ebenfalls eine Menge der Tonhöhenkandidaten und zugehöriger Pegelwerte $VP(f_{VP}, L_{VP}^*, t_{VP})$.

4.5 Berechnung von Mithörschwellenmustern (MHS)

Sowohl für die Spektrogramme spektraler als auch virtueller Tonhöhen werden die Mithörschwellenmuster über das in Abschnitt 3.1.2 beschriebene Verfahren berechnet. Die Komponenten des FTT-Pegelspektrogramms FTT_{SIG} besitzen aufgrund der verwendeten Analysefilter bereits eine gewisse Bandbreite, während die virtuellen Tonhöhen VP nur als diskrete Positionen bestimmten Pegels vorliegen. Da im Allgemeinen zahlreiche Kandidaten virtueller Tonhöhen vorliegen, ist aufgrund der Intensitätsaddition der Teilmithörschwellen keine genauere Modellierung der MHS virtueller Komponenten erforderlich. Resultat sind zwei Pegelspektrogramme MHS_{SP} und MHS_{VP} , welche die Mithörschwellenmuster der spektralen und virtuellen Schallanteile in einer Tonheits-Zeit-Ebene beinhalten.

4.6 Bestimmung der Grundfrequenz

Wird die Grundfrequenz des aktuellen Schalls nicht als Parameter an das Modell übergeben, so erfolgt eine automatische Ermittlung mithilfe eines Autokorrelationsdetektors. Dieser berechnet für die Maxima der AKF des Eingangssignals über ein Schwellwertkriterium die Abstände dieser Maxima. Die lokalen Maxima $MAX_{AKF,loc}$ der AKF werden dabei berücksichtigt, wenn Sie gegenüber dem globalen Maximum $MAX_{AKF,glob}$ um weniger als 75 % abgefallen sind, so daß für die Menge der Maxima MAX_{AKF} gilt:

$$MAX_{AKF} = \{MAX_{AKF,loc} \mid MAX_{AKF,loc} > 0,25 \cdot MAX_{AKF,glob}\} \quad (4.7)$$

Die Grundfrequenz des untersuchten Schalls ergibt sich als Kehrwert des Mittelwerts der Abstände benachbarter Maxima. Die Abstände der Maxima ΔMAX_{AKF} entsprechen der Steigung der Positionen der Maxima, trägt man diese in aufsteigender Reihenfolge auf.

$$f_0 = \frac{1}{mean(\Delta MAX_{AKF})} = \frac{1}{mean(\frac{d}{dt}(MAX_{AKF}))} \quad (4.8)$$

Für einen Schall wird somit ein einzelner Wert der Grundfrequenz f_0 ermittelt. Zeitlich variante Tonhöhen, beispielsweise bei Sinus-Sweeps, können so vom Modell nicht korrekt verarbeitet werden. Eine zeitliche Fensterung der Grundfrequenzdetektion würde hier eine Verbesserung schaffen, ließe aber nur eine blockweise, statt einer kontinuierlichen Anpassung der untersuchten Tonhöhe innerhalb des Modells zu.

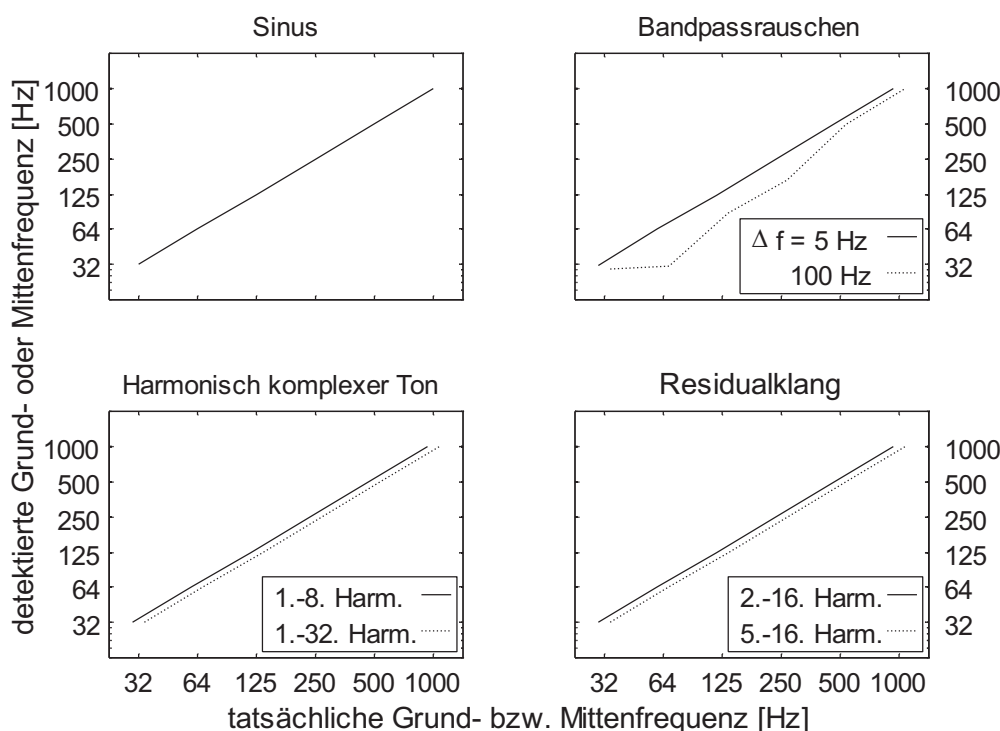


Abbildung 4.4: Ergebnisse der Grundfrequenzdetektion für verschiedene Schallarten. Bei harmonisch komplexen Tönen muß zwischen spektraler (SP) und virtueller Tonhöhe (VP) unterschieden werden.

Die Zweckmäßigkeit des Verfahrens zeigt Abbildung 4.4 für Sinustöne, Bandpassrauschen sowie harmonisch komplexe Töne und Residualtöne. Für die dargestellten Schalle stimmen detektierte Grundfrequenz und tatsächliche Grundfrequenz überein. Lediglich bei Bandpassrauschen treten Abweichungen auf. Im Beispiel werden für eine Bandbreite Δf von 100 Hz etwas zu geringe Werte bei Mittenfrequenzen von 64 bis 250 Hz berechnet. Da die exakte Tonhöhe für die vorliegende Arbeit von untergeordneter Rolle ist, reicht die Genauigkeit der erzielten Ergebnisse für die weitere Berechnung aus.

4.7 Ermittlung des zeitlichen Verlaufs der Tonhöhe

Der Verlauf des Pegels und der Position der zur Grundfrequenz f_0 gehörenden Tonhöhe innerhalb der gebildeten Spektrogramme dient der späteren Beschreibung der zeitlichen Variationen von Tonhöhe und Lautheit.

Die Tonhöhe eines Schalls kann dabei dem Algorithmus in Form eines Parameters übergeben werden, ist dies nicht der Fall, erfolgt eine automatische Detektion über das in Abschnitt 4.6 beschriebene Verfahren.

Zur Bestimmung des Verlaufs der Tonhöhe wird zunächst die Kanalnummer CH_{f_0} des Analysefilters bestimmt, der die nominelle Tonhöhe des Schalls widerspiegelt. Dieser ergibt

sich aus der Tonheit der Grundfrequenz z_{f_0} und dem Abstand der Analysefilter der Fourier-Transformation Δz_A zu

$$CH_{f_0} = \frac{z_{f_0}}{\Delta z_A} . \quad (4.9)$$

Über einen linearen Abstandsklassifikator [Rus88] läßt sich in den Spektrogrammen virtueller und spektraler Tonhöhen für jeden Zeitschritt das Maximum bestimmen, das CH_{f_0} am Nächsten liegt. Somit werden jeweils zwei Kandidaten $POS_{f_0,SP}$ und $POS_{f_0,VP}$ der möglichen Position der Tonhöhe in den Spektrogrammen ermittelt. Um nun zu unterscheiden, ob die Tonhöhe eines Schalls als spektrale oder virtuelle Tonhöhe wahrgenommen wird, wird zur weiteren Unterscheidung der FTT-Kanal, welcher der zweiten Harmonischen CH_{2f_0} des zu bewertenden Schalls entspricht, benötigt. Dieser ergibt sich entsprechend Formel 4.9 über deren Tonheit z_{2f_0} zu

$$CH_{2f_0} = \frac{z_{2f_0}}{\Delta z_A} . \quad (4.10)$$

Das Maximum spektraler Pegel $POS_{f_0,SP}$ wird beibehalten, wenn es genügend nah bei der Grundfrequenz liegt. Andernfalls wird $POS_{f_0,VP}$ als Position des zur Tonhöhe gehörenden Maximums angenommen. Die Position der Tonhöhe POS_{f_0} ergibt sich also als

$$POS_{f_0} = \begin{cases} POS_{f_0,SP} & , POS_{f_0,SP} \leq 0,9 \cdot CH_{2f_0} \\ POS_{f_0,VP} & , POS_{f_0,SP} > 0,9 \cdot CH_{2f_0} \end{cases} \quad (4.11)$$

Die Begrenzung über den Faktor 0,9 stellt eine Abgrenzung zur zweiten Harmonischen sicher und wurde empirisch aufgrund des Verhaltens des Modells bei virtuellen Tonhöhen angepasst. Der zugehörige Pegelwert L_{f_0} wird entsprechend analog zu diesem Unterscheidungskriterium aus den Daten FTT_{SP} oder FTT_{VP} entnommen:

$$L_{f_0} = \begin{cases} L_{f_0,SP} & , POS_{f_0,SP} \leq 0,9 \cdot CH_{2f_0} \\ L_{f_0,VP} & , POS_{f_0,SP} > 0,9 \cdot CH_{2f_0} \end{cases} \quad (4.12)$$

Diese Ermittlung von POS_{f_0} und L_{f_0} erfolgt für jeden Zeitschritt der FTT-Spektrogramme. So werden zwei Vektoren an die weiteren Verarbeitungsschritte des Modells weitergegeben, welche die Werte dieser Positionen und Pegel beinhalten.

4.8 Bildung des “Störanteils“

Zur Bildung des Verhältnisses N_0/N_N wird ein “Signalanteil“, der die Tonhöhe repräsentiert, und ein “Störanteil“, der die hemmenden Schallanteile beinhaltet, benötigt. Der “Signalanteil“ ergibt sich aus den Daten des Pegel- und Frequenzverlaufs der Tonhöhe entsprechend dem Verfahren aus Abschnitt 4.7. Durch Eintragen dieser Daten in eine Matrix, der Form der FTT-Spektrogramme, und Berechnung des Mithörschwellenmusters entsteht das Pegelspektrogramm der zu beurteilenden Tonhöhe MHS_T .

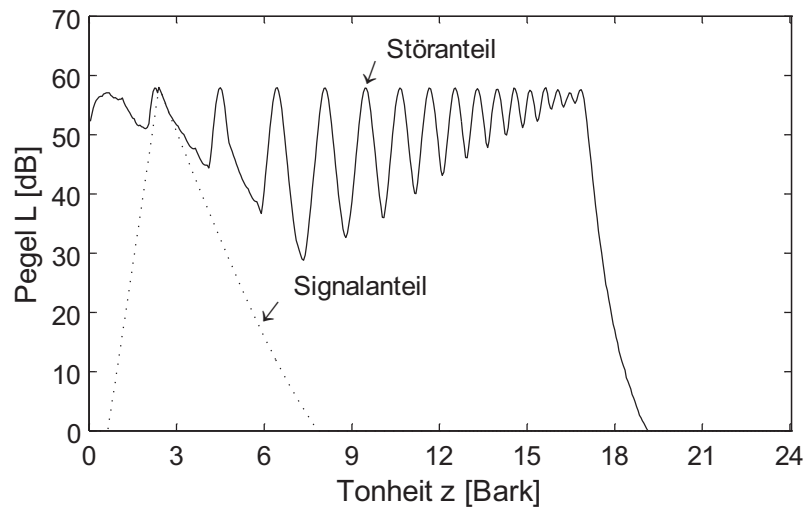


Abbildung 4.5: Ergebnis der Aufteilung in Signal- und Störanteil am Beispiel eines harmonisch komplexen Tons ($L_T = 60$ dB, $f_0 = 250$ Hz, 16 Harmonische). Durchgezogene Linie: Gesamtsignal bzw. Störanteil, gestrichelt: f_0 -Komponente bzw. Signalanteil.

Der “Störanteil“ MHS_N ergibt sich als Überlagerung der Pegelspektrogramme virtueller und spektraler Komponenten zu

$$MHS_N = \max(MHS_{SP}, MHS_{VP}) . \quad (4.13)$$

Bei Sinustönen sind MHS_N und MHS_T identisch. Das Verhältnis dieser beiden Pegelspektrogramme (und auch der zugehörigen Lautheiten) ist also gleich 1. Dadurch erreichen Sinustöne die maximale Ausprägtheit der Tonhöhe, da der Störanteil MHS_N für alle anderen Schallarten zusätzliche Anteile enthält. Das Resultat der Bildung des Signal- und Störanteils ist in Abbildung 4.5 für einen Zeitschnitt der resultierenden FTT-Pegelspektrogramme dargestellt. Die erste Harmonische des dargestellten harmonisch komplexen Tons wird als Signalanteil behandelt, der Störanteil ergibt sich als Pegelspektrum des Gesamtschalls.

4.9 Berechnung der Lautheits-Zeit Verläufe von Signal- und Störanteil

In Kapitel 3.1.3 wurde die Modifikation eines Algorithmus zur Berechnung der Lautheit dargestellt. Dieser Algorithmus benutzt zur Spektralanalyse ebenfalls die Fourier-t-Transformation und Mithörschwellenberechnung, wodurch als Zwischengröße FTT-Pegelspektrogramme bzw. Mithörschwellenmuster entstehen. Die in Abschnitt 4.8 bestimmten Spektrogramme MHS_T und MHS_N können also diesem Algorithmus zugeführt werden indem sie an der entsprechenden Stelle übergeben werden.

Resultat der Lautheitsberechnung sind Lautheits-Zeit Verläufe für den Signal- und Störanteil ($N_0(t)$ bzw. $N_N(t)$), die die Abhängigkeit der Lautheit der beiden Anteile von der Zeit

beschreiben. Bei der Berechnung der Lautheit wird dabei auf die Berücksichtigung von Effekten gedrosselter Lautheit verzichtet [Zwi63a, Deu88]. Zwar wird die Lautheit der ersten Harmonischen eines harmonisch komplexen Tons durch die höheren Harmonischen beeinflusst, bei der Erstellung des Algorithmus zeigte sich jedoch, daß die erzielten Ergebnisse auch bei Verwendung der ungedrosselten Lautheit mit den Versuchsdaten in Einklang gebracht werden können.

4.10 Zusätzliche Gewichtungen des Signalanteils

Das Verhältnis N_0/N_N liefert ein grobes Maß für die Ausgeprägtheit der Tonhöhe eines Schalls. Zur Anpassung an experimentell gewonnene Daten sind Anpassungen aufgrund zusätzlicher Einflußgrößen nötig. Diese Einflüsse lassen sich weitestgehend auf psychoakustische Empfindungen wie Schwankungsstärke, Rauigkeit und Lautheit zurückführen. Die Bildung und Motivation dieser Gewichtsfunktionen werden nachfolgend vorgestellt.

4.10.1 Bandbreite

Die in den Abschnitten 2.1.4 und 2.8 vorgestellten Experimente zeigen eine starke Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Bandbreite von Bandpassrauschen. Diese äußert sich im Lautheits-Tonheits-Zeit Muster durch verschieden breitbandige Anregung, darüber hinaus in Lage und Pegel des Maximums $\{POS_{f_0}, L_{f_0}\}$ zur Beschreibung der Tonhöhe. Das in Formel 4.1 gebildete Verhältnis der Lautheiten von Signal- und Störanteil (N_0 bzw. N_N) führt bei Bandbreitenvergrößerung zu einer Reduktion der Ausgeprägtheit der Tonhöhe. Die Lautheit des Rauschanteils steigt dabei mit wachsender Bandbreite, während die Lautheit der f_0 -Komponente nahezu konstant bleibt.

Eine Anpassung des Modells an dieses Verhalten wird über das Gewicht W_{BW} realisiert. Dazu wird, wie in Abschnitt 3.2.3 erwähnt, die Standardabweichung der Position des zur Grundfrequenz gehörenden Maximums als Maß der Bandbreite herangezogen. Diese Position liegt als Nummer des Kanals des FTT-Spektrums vor, der der wahrgenommenen Tonhöhe zugeordnet wird.

Über ein erstes Gewicht W_{STD} wird so die Schwankung der wahrgenommenen Tonhöhe nachgebildet. Dieses wird über eine gewichtete Cosinus-Funktion umgesetzt, wobei das Argument der Cosinus-Funktion auf π limitiert wird, um ein Ansteigen des Gewichts für höhere Werte des Arguments ($> \pi$) zu verhindern.

Mit Hilfe eines weiteren Gewichts W_H wird eine Anpassung an die Daten für sehr geringe Bandbreiten erreicht. Diese führen in den experimentellen Daten aufgrund starker, nicht regelmässiger, Amplitudenschwankungen bereits zu einer Reduzierung der Ausgeprägtheit der Tonhöhe. An dieser Stelle erfolgt also über die Berücksichtigung von Tonhöhen-schwankungen eine Anpassung an Amplitudenvariationen, welche nicht ausschließlich auf Schwankungen der wahrgenommenen Lautheit zurückzuführen sind.

Die beiden Gewichte W_{STD} und W_H ergeben sich zu:

$$W_{STD} = 0,5 \cdot (1,05 + 0,95 \cdot \cos[\min(\pi, \frac{\sigma(POS_{f_0})}{2,5})]) \quad (4.14)$$

$$W_H = \max\left(0,8, \sqrt{\frac{1}{1 + \sigma(POS_{f_0})}}\right) \quad (4.15)$$

Das Gewicht W_{BW} ergibt sich nun als Produkt der beiden vorangehend beschriebenen Bewertungsfunktionen zu:

$$W_{BW} = W_{STD} \cdot W_H \quad (4.16)$$

Der Verlauf der Gewichtsfunktionen W_{STD} , W_H sowie des resultierenden Gewichts W_{BW} in Abhängigkeit von der Standardabweichung des Kanals, der das zur Tonhöhe resultierende Maximum beinhaltet, $\sigma(POS_{f_0})$ ist in Abbildung 4.6 gezeigt.

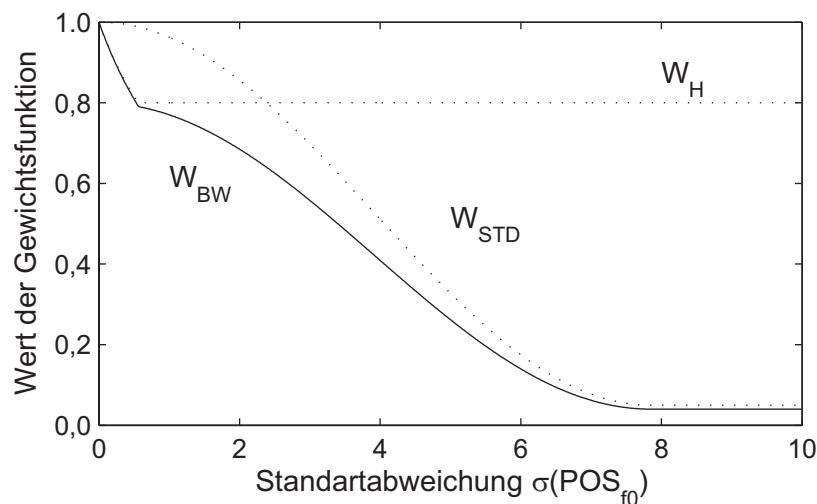


Abbildung 4.6: Gewichtsfunktionen W_{STD} , W_H und W_{BW} in Abhängigkeit der Standardabweichung der Kanalnummer, welche der Grundtonhöhe zugeordnet wird.

4.10.2 Amplitudenvariation

Aus Sicht der Psychoakustik äußern sich Amplitudenvariationen eines Schallsignals in Form der Wahrnehmungen Schwankungsstärke und Rauigkeit. Bei langsamen Modulationsfrequenzen f_{mod} überwiegt dabei die Schwankungsstärke, ab Modulationsfrequenzen $f_{mod} > 20\text{-}30$ Hz wird die Wahrnehmung der Rauigkeit dominant. Beide Wahrnehmungen zeigen bandpassähnliche Abhängigkeiten von der Modulationsfrequenz und wurden bereits in Form etablierter Modelle beschrieben [Sch78, Aur85]. Diese basieren grundsätzlich auf der Verrechnung von Modulationsfrequenz und Modulationstiefe.

Die Modulationsfrequenz ergibt sich durch Bestimmung der Nulldurchgänge der zeitlichen Ableitung des ermittelten Lautheitsverlaufs $N_0(t)$. Sie geht in ein Gewicht zur Nachbildung der Abhängigkeit der Ausprägtheit der Tonhöhe von Schwankungen der Amplitude

ein und repräsentiert somit den Einfluß von Schwankungsstärke und Rauigkeit aufgrund von Amplitudenmodulation auf die wahrgenommene Ausgeprägtheit der Tonhöhe. Diese Gewichtung erfolgt abhängig von der Modulationsfrequenz f_{mod} mit Hilfe des frequenzabhängigen Exponenten $E_{AM}(f_0)$ über den Zusammenhang

$$W_{AM} = \left(\frac{0.5}{(f_{mod}/5) + 1} + 0.5 \right)^{E_{AM}(f_0)}. \quad (4.17)$$

Der Verlauf von W_{AM} ist für eine Grundfrequenz von 1000 Hz in Abbildung 4.7 dargestellt. Wie zu sehen ist erfolgt lediglich eine Dämpfung mit steigender Modulationsfrequenz, die ab $f_{mod} \approx 25$ Hz in Sättigung geht. Somit wird hauptsächlich der Einfluß langsamer Modulationsfrequenzen über dieses Gewicht modelliert, die bei höheren Modulationsfrequenzen dominante Wahrnehmung ‘‘Rauigkeit‘‘ wird nicht gesondert berücksichtigt.

Der Exponent $E_{AM}(f_0)$ passt die Abnahme der Gewichtsfunktion mit zunehmender Modulationsfrequenz an die Daten der Hörversuche aus Abschnitt 2.2 für langsame Modulationen ($f_{mod} < 32$ Hz) an. Über

$$E_{AM}(f_0) = -\frac{0,5}{1750} \cdot f_0 + \frac{11}{7} \quad (4.18)$$

wird dieser Exponent frequenzabhängig über die Grundfrequenz des zu bewertenden Schalls gebildet. Auf eine Berücksichtigung des Amplituden- oder Lautheitshubs wird an dieser Stelle verzichtet. Zwar liegt es Nahe, daß verschieden stark wahrnehmbare Schwankungen zu unterschiedlichen Ergebnissen führen, jedoch liegen hier derzeit noch keine experimentellen Daten vor, so daß keine Berücksichtigung eines entsprechenden Einflusses vorgenommen wird.

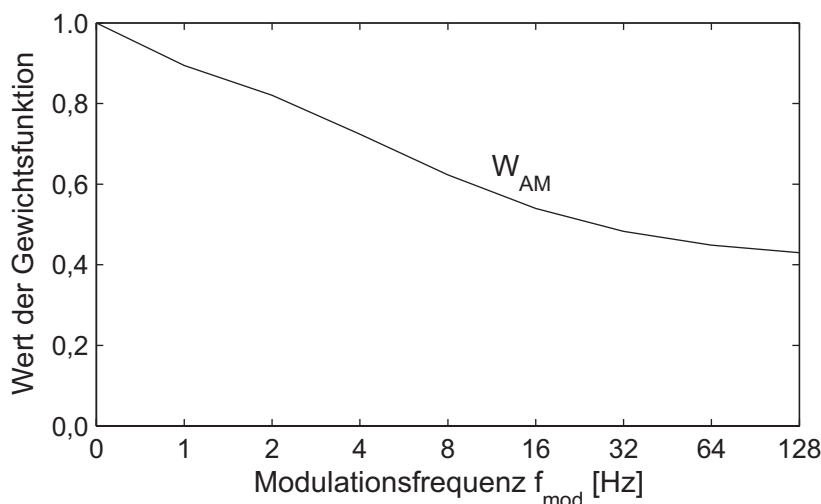


Abbildung 4.7: Verlauf der Gewichtsfunktionen W_{AM} bei festem Exponenten $E_{AM}(f_0)$ bei Modulationsfrequenzen $f_{mod} < 128$ Hz.

4.10.3 Grundfrequenz

Die in Kapitel 2.1.1 dargestellte Abhängigkeit der Ausprägtheit der Tonhöhe von der Frequenz eines Sinustons läßt sich in Form eines Bandpasses beschreiben. Durch eine Realisierung dieses Bandpasses abhängig von der Tonheit der zu bewertenden Tonhöhe z_0 ergibt sich das Gewicht W_{F0} zu:

$$W_{F0} \sim \left(\frac{2}{\frac{14}{z_0+5} + \frac{z_0+5}{14}} \right)^2 \quad (4.19)$$

Der Exponent legt die Stärke des Abfalls des Bandpasses bei Entfernen von der Mitten-tonheit $z_c = 14$ Bark fest. Abbildung 4.8 (links) zeigt den Verlauf dieser Gewichtsfunktion.

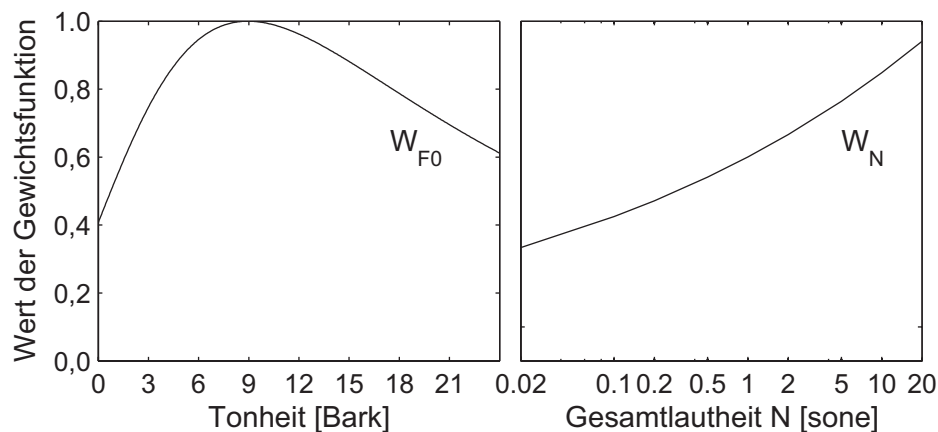


Abbildung 4.8: Werte der Gewichtsfunktion W_{F0} zur Anpassung des Modells an die Grundfrequenz des untersuchten Schalls (links) sowie W_N zur Anpassung an die Abhängigkeit der Ausprägtheit der Tonhöhe von der Lautheit eines Schalls (rechts).

4.11 Gewichtung aufgrund der Gesamtlautheit

In Kapitel 2.1.3 wurde die Abhängigkeit der Ausprägtheit der Tonhöhe vom Schallpegel über einen einfachen exponentiellen Zusammenhang mit der wahrgenommenen Lautheit beschrieben. Das Gewicht W_N zur Beschreibung dieser Abhängigkeit besitzt also die Form

$$W_N = \left(\frac{N_5}{30} \right)^{0,15} \quad (4.20)$$

Der Exponent 0,15 dient zur Einstellung der Steigung dieses Zusammenhangs und wurde für die Darstellung in Abbildung 4.8 mit Hilfe der Daten aus Abschnitt 2.1.2 optimiert. Der Wert N_5 wird als Perzentillautheit bezeichnet und repräsentiert, entgegen der in der Mathematik üblichen Schreibweise, denjenigen Wert, unterhalb dem 95 % der ermittelten Lautheitswerte liegen.

Für die Anpassung an die Daten zur Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Schalldauer ist aufgrund der Verwendung eines Lautheitsmodells innerhalb des Modells keine gesonderte Anpassung vorgesehen, da diese entsprechend Abbildung 2.2 nahezu identisch mit der Abhängigkeit der Lautheit von der Schalldauer ist. Da das verwendete Lautheitsmodell diesen Zusammenhang korrekt wiedergibt, ist zu erwarten, daß auch die Modellierung der Ausgeprägtheit der Tonhöhe für diese Abhängigkeit weitestgehend korrekt sein wird.

4.12 Zusammenfassung

In diesem Kapitel wurde ein Modell der Ausgeprägtheit der Tonhöhe vorgestellt. Ausgehend von etablierten Algorithmen (FTT, Modell der Tonhöhenwahrnehmung, Berechnung der “Pitch Saliency“) werden dazu Merkmale extrahiert, die weitergehend verarbeitet werden um die Ausgeprägtheit der Tonhöhe bestimmen zu können.

Aus den Gewichten virtueller Tonhöhen W_{VP} werden dazu Pegelwerte L_{VP}^* zur Repräsentation dieser Tonhöhen in einer Spektrogrammdarstellung ermittelt. Für die Pegelspektrogramme der zu beurteilenden Tonhöhe FTT_{f_0} , der virtuellen Tonhöhe FTT_{VP} sowie der spektralen Tonhöhe FTT_{SP} werden jeweils die Mithörschwellenmuster berechnet. Zur Transformation der Mithörschwellenmuster MHS_N und MHS_T in die Lautheit wird ein modifiziertes Lautheitsmodell benutzt. So liegen für die Komponenten, die zur Beschreibung der Ausgeprägtheit der Tonhöhe weiter verwendet werden, die jeweiligen Lautheitsverläufe N_0 und N_N vor.

Über zusätzliche Gewichtungsfunktionen zur Beschreibung der Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Bandbreite W_{BW} , Amplitudenmodulation W_{AM} , Grundfrequenz W_{F_0} und Lautheit W_N wird eine Anpassung an Abhängigkeiten von weiteren psychoakustischen Größen erreicht. Die Abhängigkeit von der Schalldauer wird dagegen aufgrund des verwendeten modifizierten Lautheitsmodells bereits erfasst.

Aus den so ermittelten Größen ergibt sich schließlich die Ausgeprägtheit der Tonhöhe über den Zusammenhang

$$AdT \sim W_N \cdot W_{BW} \cdot W_{AM} \cdot W_{F_0} \cdot \frac{N_0}{N_N}. \quad (4.21)$$

Kapitel 5

Verifikation des Modells der Ausgeprägtheit der Tonhöhe

Das über die experimentell gewonnenen Daten und Spektralanalysen erstellte Modell der Ausgeprägtheit der Tonhöhe wird im Folgenden auf seine Anwendbarkeit hin untersucht. Dazu werden Versuchsergebnisse aus Kapitel 2 den, mit Hilfe des Modells, berechneten Daten gegenübergestellt. Zur Bewertung der berechneten Daten werden objektive Bewertungskriterien verwandt, welche zunächst motiviert und erläutert werden sollen.

5.1 Gütekriterien zur Beurteilung der Modellberechnungen

Um die durch das Modell berechneten Werte der Ausgeprägtheit der Tonhöhe AdT_M zu beurteilen, bedarf es objektiver Bewertungskriterien. Die Ergebnisse psychoakustischer Experimente können individuell sehr unterschiedlich ausfallen, sollten in ihrer Tendenz jedoch keine interindividuellen Widersprüche aufweisen. Beispielsweise sollten für eine Abhängigkeit der Form stetig steigender Empfindungsgröße bei entsprechender Variation einer physikalischen Größe alle Versuchspersonen zumindest diesem Trend folgen. Das heißt die Rangfolge der Ergebnisse sollte interindividuell identisch sein. Als Kriterium läßt sich also die Rangfolge der berechneten Ergebnisse ermitteln, was mit Hilfe des Rangkorrelationskoeffizienten RANKCORR nach Spearman [Sac04] geschieht, welcher die prinzipielle Übereinstimmung experimentell gewonnener und berechneter Daten zeigt. Unterschiede in der Skalierung der Empfindungsgröße werden dabei nicht berücksichtigt. Zur Überprüfung der Hypothese H_0 , daß Versuchsergebnisse und Modellberechnungen der gleichen Grundgesamtheit entstammen, dienen die Schranken des Rangkorrelationskoeffizienten für ein Signifikanzniveau von 5 % [Zar99, Lee06]. Auf diese Weise kann die Hypothese positiver Korrelationen zweier Rangfolgen und somit im vorliegenden Fall die grundsätzlich geeignete Modellierung der Versuchsdaten nachgewiesen werden.

Um nun den Abstand der berechneten Daten zu den Ergebnissen der Experimente zu bewerten, werden zwei Kriterien verwandt: Die Anzahl berechneter Daten, die sich innerhalb des Interquartilbereichs der Versuchsergebnisse befindet, sowie der mittlere Abstand vom Median der Versuchsergebnisse, normiert auf den jeweiligen Interquartilbe-

reich. Der Prozentsatz der Modellberechnungen AdT_M innerhalb der Interquartilbereiche ($PCTINTERQ$) der Versuchsergebnisse ergibt sich über einen einfachen Vergleich der Berechnungen mit den Interquartilbereichen zu:

$$PCTINTERQ = 100 \cdot \frac{n(Q_{1,Exp} < AdT_M < Q_{3,Exp})}{n_{Schalle}} [\%], \quad (5.1)$$

wobei $Q_{1,Exp}$ und $Q_{3,Exp}$ das erste und dritte Quartil der Versuchsergebnisse darstellen. $n(Q_{1,Exp} < AdT_M < Q_{3,Exp})$ bezeichnet somit die Anzahl der berechneten Werte innerhalb der zugehörigen Interquartilbereiche und $n_{Schalle}$ entspricht der Gesamtzahl untersuchter Signale.

Um zusätzlich den Abstand der Berechnungen zu den Versuchsergebnissen zu bewerten, ist ein Gütemaß in Form eines Abstandsmasses erforderlich. Aufgrund der geringen Stichprobenanzahl kann der Nachweis normalverteilter Versuchsergebnisse nicht erbracht werden [Sac04] und so für die Ergebnisse weder Standardabweichung noch Varianz ermittelt werden. Aus diesem Grund wird der Abstand vom Median der Versuchsergebnisse je nach Lage oberhalb oder unterhalb dieses Wertes mit dem 1. oder 3. Quartil bewertet. So ergibt sich für den gewichteten Abstand $DIST$ vom Median der Versuchsergebnisse:

$$DIST = \begin{cases} \frac{MED_{Exp} - AdT_M}{MED_{Exp} - Q_{1,Exp}} [\%], & AdT_M < MED_{Exp} \\ \frac{AdT_M - MED_{Exp}}{Q_{3,Exp} - MED_{Exp}} [\%], & AdT_M > MED_{Exp} \\ 0 & [\%], AdT_M = MED_{Exp} \end{cases} \quad (5.2)$$

Der mittlere gewichtete Abstand ergibt sich als arithmetisches Mittel der einzelnen Abstände:

$$MEANDIST = \frac{\sum_{i=1}^{n_{Schalle}} DIST_i}{n_{Schalle}} [\%] \quad (5.3)$$

Die somit definierten Größen werden in der Folge zur Bewertung der Modellberechnungen im Vergleich zu experimentell gewonnenen Daten herangezogen. Dabei werden die Versuchsergebnisse wie in Abschnitt 1.7 beschrieben dargestellt. Bei Experimenten, für die sowohl Hin- als auch Rückversuche (mit unterschiedlichem Anker) durchgeführt wurden, wird jeweils der Mittelwert der Ergebnisse aus diesen beiden Versuchsreihen verwendet. Daten aus der Literatur werden so interpretiert, daß die Versuchsergebnisse in vergleichbarer Form als Mediane mit Interquartilbereichen oder arithmetische Mittelwerte mit Varianzen ermittelt wurden. Aufgrund teilweise fehlender Angaben hierzu sind Fehler durch diese Interpretation nicht ausgeschlossen, insbesondere das Kriterium ‘‘Rangkorrelation‘‘ bleibt hiervon aber unberührt.

5.2 Versuchsergebnisse im Vergleich zu berechneter Ausgeprägtheit der Tonhöhe

Um die Richtigkeit der Überlegungen zu zeigen, die dem erstellten Modell der Ausgeprägtheit der Tonhöhe zugrunde liegen, sollen nun die Versuchsergebnisse aus psychoakustischen Experimenten den mit Hilfe des Modells berechneten Daten gegenübergestellt werden. Dabei werden, falls nicht anders vermerkt, die Daten der Hörversuche als nicht gefüllte Symbole in Form von Median und Interquartilbereichen dargestellt, die berechneten Daten werden jeweils als ausgefüllte Symbole eingetragen.

Die Daten der Experimente, für die mehrere Versuche mit verschiedenen Ankerschallen durchgeführt wurden, werden als Mittelwerte der Mediane und Interquartilbereiche der verfügbaren Versuchsreihen dargestellt. Zur Beurteilung der berechneten Daten im Vergleich zu den experimentell ermittelten Ergebnissen, dienen die in Abschnitt 5.1 eingeführten Kriterien Rangkorrelation und normierter Abstand zu den Medianen.

5.2.1 Abhängigkeit der Ausgeprägtheit der Tonhöhe von Frequenz, Pegel und Dauer eines Sinustons

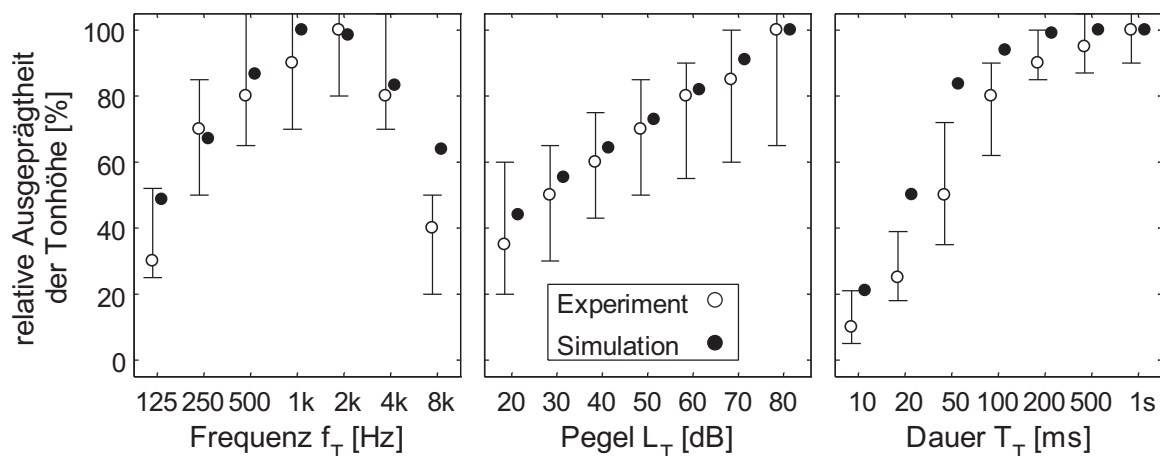


Abbildung 5.1: Experimentell ermittelte und berechnete Daten zur Ausgeprägtheit der Tonhöhe von Sinustönen in Abhängigkeit von Frequenz, Pegel und Dauer. Experimentelle Daten entsprechend Kapitel 2.1.

Abbildung 5.1 zeigt die Ergebnisse der Modellberechnungen im Vergleich zu experimentell ermittelten Daten aus Kapitel 2.1 für die Ausgeprägtheit der Tonhöhe von Sinustönen in Abhängigkeit von deren Frequenz, Pegel und Dauer. Die Versuchsergebnisse werden durch Mittelwerte der verfügbaren Daten, welche durch Experimente mit verschiedenen Ankerschallen (Kapitel 2.1) gewonnen wurden, repräsentiert.

Für die Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Frequenz f_T eines Sinustons (Abbildung 5.1, linke Seite) wird die resultierende Bandpasscharakteristik prinzipiell nachvollzogen. Abgesehen vom Ergebnis für die Testtonfrequenz $f_T = 8$ kHz liegen alle durch das Modell berechneten Werte im Bereich der Interquartile der Daten aus den Experimenten. Der mittlere, auf die Quartile normierte Abstand beträgt nur ca. 61 %. Mit einem Wert von 0,95 des Rangkorrelationskoeffizienten kann die Hypothese geeigneter Modellierung der Versuchsergebnisse auf einem 5 %-Niveau nicht abgelehnt werden.

Ähnliches gilt für die in der Mitte von Abbildung 5.1 dargestellte Abhängigkeit der Ausgeprägtheit der Tonhöhe eines Sinustons der Frequenz $f_T = 1$ kHz von dessen Darbietungspegel L_T . Das Modell ermittelt für steigende Darbietungspegel analog zu den Versuchsergebnissen ansteigende Werte der Ausgeprägtheit der Tonhöhe. Die berechneten Werte liegen sehr dicht bei den Medianen der Versuchsergebnisse, der mittlere normierte Abstand beträgt nur 25 % des jeweiligen Quartils. Der Rangkorrelationskoeffizient zeigt mit einem Wert von 1,00 identische Rangfolge der Ergebnisse der Hörversuche und der Modellberechnungen an.

Auf der rechten Seite von Abbildung 5.1 sind Berechnungen und Versuchsergebnisse für die Ausgeprägtheit der Tonhöhe von Sinustönen der Frequenz $f_T = 1$ kHz in Abhängigkeit von der Dauer dargestellt. Unterhalb einer Schalldauer T_T von 100 ms zeigen sich deutliche Abweichungen der berechneten Werte von den Versuchsergebnissen. Der Anstieg der berechneten Ausgeprägtheit der Tonhöhe erfolgt für sehr kurze Dauern mit steigender Dauer zu schnell, so daß hier eine leichte Überbewertung auftritt.

	Frequenz	Pegel	Dauer
RANKCORR	0,95	1,00	1,00
Signifikant ($\alpha = 0,05$)	ja	ja	ja
PCTINTERQ [%]	85,7	100	57,1
MEANDIST [%]	61	25	102

Tabelle 5.1: Berechnete Gütekriterien für die Modellierung der Versuchsdaten aus den Experimenten zur Ausgeprägtheit der Tonhöhe von Sinustönen in Abhängigkeit von deren Frequenz, Pegel und Dauer.

Tabelle 5.1 faßt die Ergebnisse der Modellberechnungen für die Abhängigkeiten bei Sinustönen nochmals zusammen. Dargestellt ist der berechnete Rangkorrelationskoeffizient nach Spearman [Sac04] und dessen Signifikanz auf einem Niveau von 5 %. Darüber hinaus sind der Prozentsatz an Modellberechnungen, die sich innerhalb des Interquartilbereichs der Versuchsergebnisse befindet und der mittlere normierte Abstand vom Median der Versuchsergebnisse eingetragen. Aufgrund der Rangkorrelationskoeffizienten kann für alle drei Datensätze die Nullhypothese positiver Korrelation, der Versuchsergebnisse und der Modellberechnungen, auf einem Signifikanzniveau von 5 % nicht abgelehnt werden.

5.2.2 Ausgeprägtheit der Tonhöhe amplitudenmodulierter Sinustöne

Abbildung 5.2 zeigt die Ergebnisse des Hörversuchs aus Abschnitt 2.2 für Sinustöne mit langsamer Amplitudenmodulation im Vergleich zu berechneten Daten. Die in Tabelle 5.2 aufgeführten Werte belegen, daß die Abhängigkeiten der Ausgeprägtheit der Tonhöhe amplitudenmodulierter Sinustöne vom Modell nachvollzogen werden können.

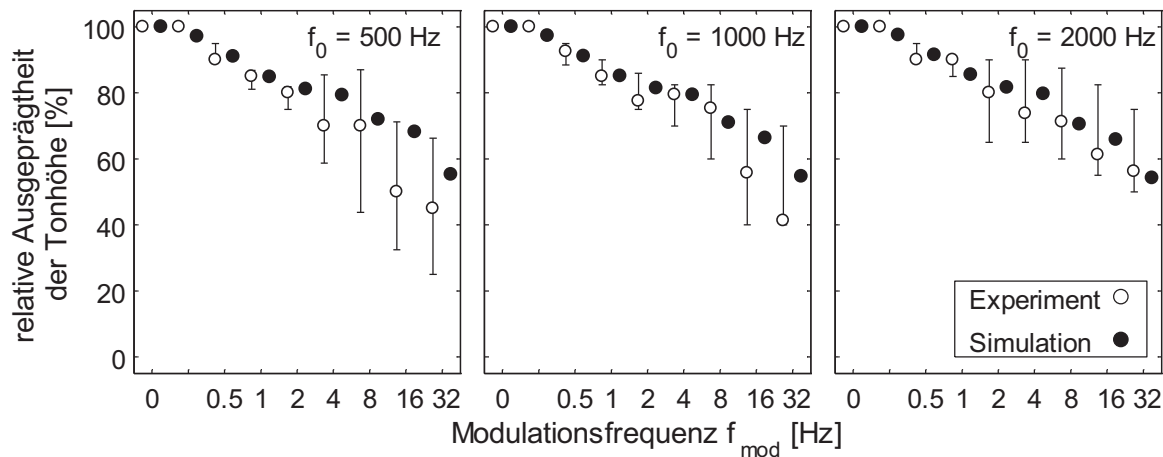


Abbildung 5.2: *Abhängigkeit der Ausgeprägtheit der Tonhöhe und berechnete Werte für amplitudenmodulierte Sinustöne unterschiedlicher Grundfrequenz bei Variation der Modulationsfrequenz f_{mod} entsprechend Kapitel 2.2.*

Die Anzahl der Berechnungen innerhalb der Interquartilbereiche der Versuchsergebnisse ist durchwegs größer als 75 %. Der normierte Abstand der Berechnungen von den Medianen MEANDIST ist sehr gering (< 65 % der Quartile), was die Gültigkeit des Modells für diese Daten bestätigt.

Daten nach [Fas90] für höhere Modulationsfrequenzen sind in Abbildung 5.3 zusammen mit den Modellberechnungen dargestellt. Die berechneten Gütemaße in Tabelle 5.3 zeigen aufgrund der Rangkorrelation erneut eine sehr gute Übereinstimmung mit den experimentell ermittelten Daten (RANKCORR = 1).

$f_0 =$	500 Hz	1000 Hz	2000 Hz
RANKCORR	0,99	0,98	0,99
Signifikant ($\alpha = 0,05$)	ja	ja	ja
PCTINTERQ [%]	78	89	78
MEANDIST [%]	51	59	64

Tabelle 5.2: *Berechnete Gütekriterien der Modellierung der Versuchsdaten aus den Experimenten zur Ausgeprägtheit der Tonhöhe von AM Sinustönen mit langsamer Modulation ($f_{mod} < 32$ Hz) bei drei verschiedenen Mittenfrequenzen.*

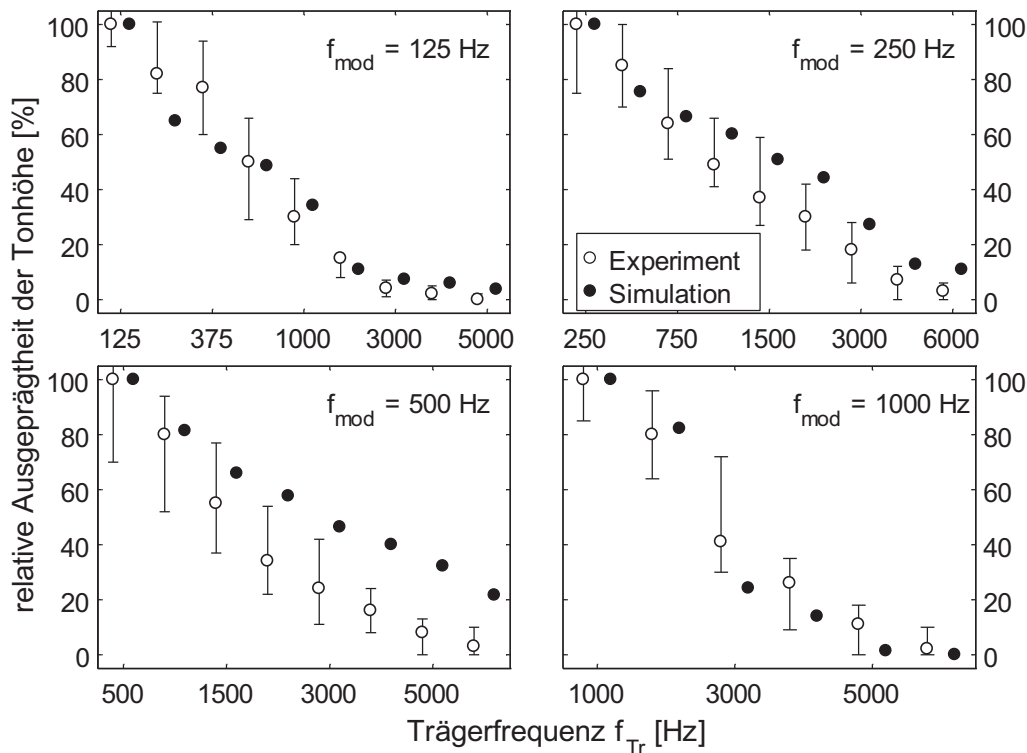


Abbildung 5.3: Abhängigkeit der Ausprägtheit der virtuellen Tonhöhe von amplitudenmodulierten Sinustönen für vier verschiedene Modulationsfrequenzen. Daten entsprechend Kapitel 2.2.

Abbildung 5.3 sowie die mittleren Abstände vom Median der Versuchsergebnisse (MEANDIST) zeigen, daß zwar nicht alle Ergebnisse auch quantitativ nachvollzogen werden können, die berechneten Ergebnisse aber sehr nahe bei den Ergebnissen aus den Hörversuchen liegen. Insbesondere ist zu berücksichtigen, daß die Signale bei Trägerfrequenzen über 4000 Hz bereits den Existenzbereich der virtuellen Tonhöhe verlassen und diese nur noch sehr schwach ausgeprägte Tonhöhen hervorrufen. Dennoch zeigen die Modellberechnungen auch in diesem Bereich eine übereinstimmende Reihenfolge mit den Versuchsergebnissen.

$f_{mod} =$	125 Hz	250 Hz	500 Hz	1000 Hz
RANKCORR	1,00	1,00	1,00	1,00
Signifikant ($\alpha = 0,05$)	ja	ja	ja	ja
PCTINTERQ [%]	56	78	38	83
MEANDIST [%]	95	80	162	71

Tabelle 5.3: Gütekriterien der Modellierung der Versuchsdaten nach [Fas90] zur Ausprägtheit der Tonhöhe von AM Sinustönen.

5.2.3 Ausprägtheit der Tonhöhe von harmonisch komplexen Tönen in Abhängigkeit von der Anzahl Harmonischer

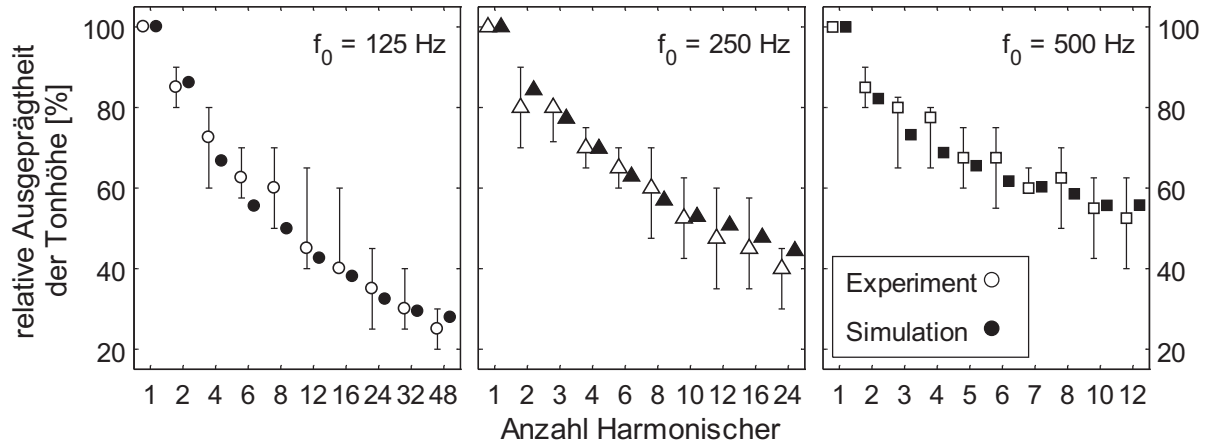


Abbildung 5.4: Experimentell bestimmte und berechnete Daten zur Ausprägtheit der Tonhöhe von harmonisch komplexen Tönen variabler Anzahl Harmonischer bei drei verschiedenen Grundfrequenzen. Daten entsprechend Kapitel 2.4.

Die Daten des Versuchs zur Ausprägtheit der Tonhöhe harmonisch komplexer Töne mit variierender Anzahl Harmonischer aus Kapitel 2.4 sind in Abbildung 5.4 im Vergleich zu den Modellberechnungen dargestellt. Der Effekt sinkender Ausprägtheit der Tonhöhe mit steigender Anzahl Harmonischer kann vom Modell qualitativ und quantitativ sehr gut nachvollzogen werden. Die Werte der berechneten Gütemaße, entsprechend Tabelle 5.4, zeigen mit Rangkorrelationskoeffizienten nahe 1,00 eine sehr gute Übereinstimmung der Berechnungen mit den Versuchsergebnissen. Auch der mittlere Abstand zu den Medianen der Experimente ist jeweils kleiner als die Hälfte der ermittelten Quartile ($\text{MEANDIST} < 50\%$). Für die Grundfrequenzen $f_0 = 250$ und 500 Hz befinden sich alle berechneten Werte innerhalb der Interquartilbereiche der Versuchsergebnisse ($\text{PCTINTERQ} = 100\%$). Für die niedrigste Grundfrequenz $f_0 = 125$ Hz liegen die Berechnungen für drei Schalle außerhalb dieses Bereichs. So ist insgesamt eine sehr gute Übereinstim-

$f_0 =$	125 Hz	250 Hz	500 Hz
RANKCORR	1,00	1,00	0,97
Signifikant ($\alpha = 0,05$)	ja	ja	ja
PCTINTERQ [%]	70	100	100
MEANDIST [%]	46	29	32

Tabelle 5.4: Gütekriterien der Modellierung der Ausprägtheit der Tonhöhe harmonisch komplexer Töne variabler Anzahl Harmonischer bei drei verschiedenen Grundfrequenzen f_0 .

mung der berechneten Daten mit den Versuchsergebnissen festzustellen.

5.2.4 Ausgeprägtheit der virtuellen Tonhöhe von harmonisch komplexen Residualtönen

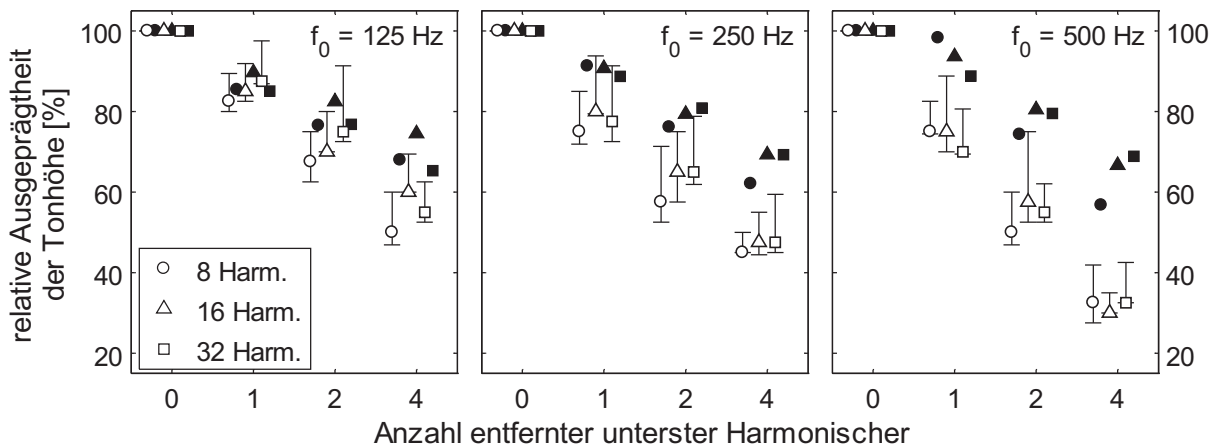


Abbildung 5.5: Ergebnisse aus Kapitel 2.5.1 und berechnete Daten der Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Residualtöne bei drei Grundfrequenzen. Nicht gefüllte Symbole: subjektive Daten, gefüllte Symbole: Berechnung.

Die Daten aus dem Versuch zur Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Töne mit variierender Anzahl entfernter unterer Harmonischer aus Abschnitt 2.5.1

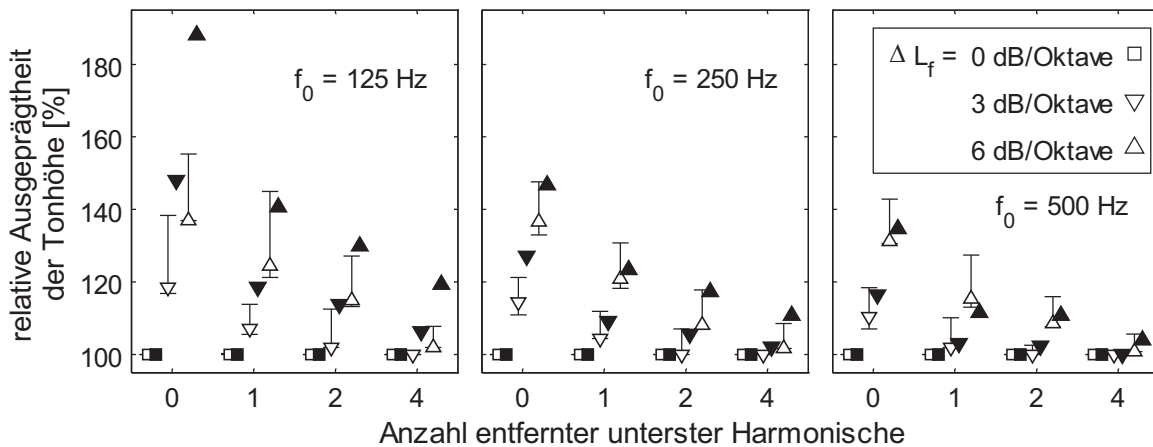


Abbildung 5.6: Versuchsergebnisse entsprechend Kapitel 2.5.2 sowie berechnete Daten der Ausgeprägtheit der virtuellen Tonhöhe harmonisch komplexer Töne variabler Anzahl entfernter unterster Harmonischer bei drei verschiedenen Grundfrequenzen und variierender spektraler Dämpfung. Nicht gefüllte Symbole: subjektive Daten, gefüllte Symbole: Modellberechnungen.

sind in Abbildung 5.5 als Mittelwerte der Ergebnisse von Hin- und Rückversuch dargestellt. Zur Bewertung der ebenfalls dargestellten Ergebnisse der Modellberechnungen sind in Tabelle 5.5 die Gütekriterien der Modellierung dieser Daten aufgeführt.

Diese zeigen, wie auch aus Abbildung 5.6 zu erkennen ist, daß die Eigenschaft fallender Ausgeprägtheit der Tonhöhe bei Erhöhung der Anzahl entfernter unterer Harmonischer nachvollzogen werden kann. Die Abweichung von den experimentell ermittelten Versuchsergebnissen ist jedoch teils sehr hoch. Bei einer Grundfrequenz von 500 Hz wird eine zu geringe Abnahme der Ausgeprägtheit der Tonhöhe prognostiziert. Die berechneten Daten liegen, abgesehen vom Ankerschall, ausschließlich außerhalb der Interquartilbereiche. Zudem wird die Abhängigkeit von der Grundfrequenz, daß bei höherer Grundfrequenz ein stärkerer Rückgang der Ausgeprägtheit der Tonhöhe auftritt, nur unzureichend nachvollzogen.

Anders verhält es sich für die in Abbildung 5.6 dargestellten Ergebnisse der Versuche und Modellberechnungen für virtuelle Tonhöhen harmonisch komplexer Töne bei verschiedener spektraler Dämpfung nach Kapitel 2.5.2. Wie Tabelle 5.6 zeigt werden die prinzipiellen Abhängigkeiten nachvollzogen. Da innerhalb des Versuchs diejenigen Schalle mit einer spektralen Dämpfung von 0 dB als Anker dienten, ist hier die Übereinstimmung zwischen Modell und Simulation bei 100 %, da die Modellberechnungen ebenfalls auf diese Werte normiert wurden.

Lediglich die Ergebnisse der Berechnungen für eine Grundfrequenz von 500 Hz liegen überwiegend innerhalb der Interquartilbereiche. Die Verläufe der experimentell gewonnenen Daten können dennoch sehr gut nachvollzogen werden, berücksichtigt man die

	$f_0 =$	125 Hz	250 Hz	500 Hz
$n_h = 8$	RANKCORR Signifikant ($\alpha = 0,05$)	1,00	1,00	1,00
	PCTINTERQ [%]	50	25	25
	MEANDIST [%]	85	159	140
$n_h = 16$	RANKCORR Signifikant ($\alpha = 0,05$)	1,00	1,00	1,00
	PCTINTERQ [%]	50	50	25
	MEANDIST [%]	86	128	95
$n_h = 32$	RANKCORR Signifikant ($\alpha = 0,05$)	1,00	1,00	1,00
	PCTINTERQ [%]	50	50	25
	MEANDIST [%]	140	250	223

Tabelle 5.5: Berechnete Gütekriterien für die Modellierung der Versuchsdaten zur virtuellen Tonhöhe harmonisch komplexer Töne.

	$f_0 =$	125 Hz	250 Hz	500 Hz
$\Delta L = 0$ dB	RANKCORR	1,00	1,00	1,00
	Signifikant ($\alpha = 0,05$)	Angabe nicht möglich (n = 4)		
	PCTINTERQ [%]	100	100	100
	MEANDIST [%]	0	0	0
$\Delta L = 3$ dB	RANKCORR	1,00	0,95	0,95
	Signifikant ($\alpha = 0,05$)	Angabe nicht möglich (n = 4)		
	PCTINTERQ [%]	0	50	75
	MEANDIST [%]	110	83	46
$\Delta L = 6$ dB	RANKCORR	1,00	1,00	1,00
	Signifikant ($\alpha = 0,05$)	Angabe nicht möglich (n = 4)		
	PCTINTERQ [%]	25	75	75
	MEANDIST [%]	195	86	73

Tabelle 5.6: Gütekriterien für die Modellierung der Versuchsdaten zur virtuellen Tonhöhe harmonisch komplexer Töne verschiedener spektraler Dämpfung.

Schwierigkeit der Bewertung virtueller Tonhöhen. Bei vier entfernten unteren Harmonischen tritt durchwegs eine leichte Überschätzung der Ausgeprägtheit der Tonhöhe bei spektraler Dämpfung von 6 dB/Oktave gegenüber 0 dB/Oktave auf. Die Versuchsergebnisse deuten zwar auf eine solche Bewertung hin, vor allem bei einer Grundfrequenz von 125 Hz liegt der berechnete Wert jedoch deutlich über der experimentell ermittelten Ausgeprägtheit der Tonhöhe.

5.2.5 Ausgeprägtheit der Tonhöhe von Bandpassrauschen

Wie in Abschnitt 2.8.3 beschrieben, sinkt die Ausgeprägtheit der Tonhöhe von Bandpassrauschen bei Erhöhung der Bandbreite. Darüberhinaus erfolgt diese Absenkung für niedrige Mittenfrequenzen schneller als bei Hohen. Bei gleicher Bandbreite ist also die Ausgeprägtheit der Tonhöhe bei einer höheren Mittenfrequenz, bezogen auf den Ankerschall bei gleicher Mittenfrequenz, größer, als bei niedrigen Mittenfrequenzen.

	$f_c =$	250 Hz	500 Hz	1000 Hz	2000 Hz
RANKCORR		1,00	1,00	1,00	1,00
Signifikant ($\alpha = 0,05$)		ja	ja	ja	ja
PCTINTERQ [%]		50	67	33	67
MEANDIST [%]		241	71	97	104

Tabelle 5.7: Gütekriterien für die Modellierung der Versuchsdaten zur Ausgeprägtheit der Tonhöhe von Bandpassrauschen.

Abbildung 5.7 zeigt diese Ergebnisse im Vergleich zu den berechneten Daten. Anhand der in Tabelle 5.7 dargestellten Rangkorrelationskoeffizienten ist ersichtlich, daß die Daten für eine feste Mittenfrequenz, in Einklang mit den Ergebnissen der Experimente, mit steigender Bandbreite abfallen. Die Abstände von den Medianen der Versuchsdaten sind jedoch relativ groß, abgesehen von einer Mittenfrequenz von 500 bzw. 1000 Hz liegen die berechneten Werte im Mittel außerhalb des Interquartilbereichs (MEANDIST > 100 %).

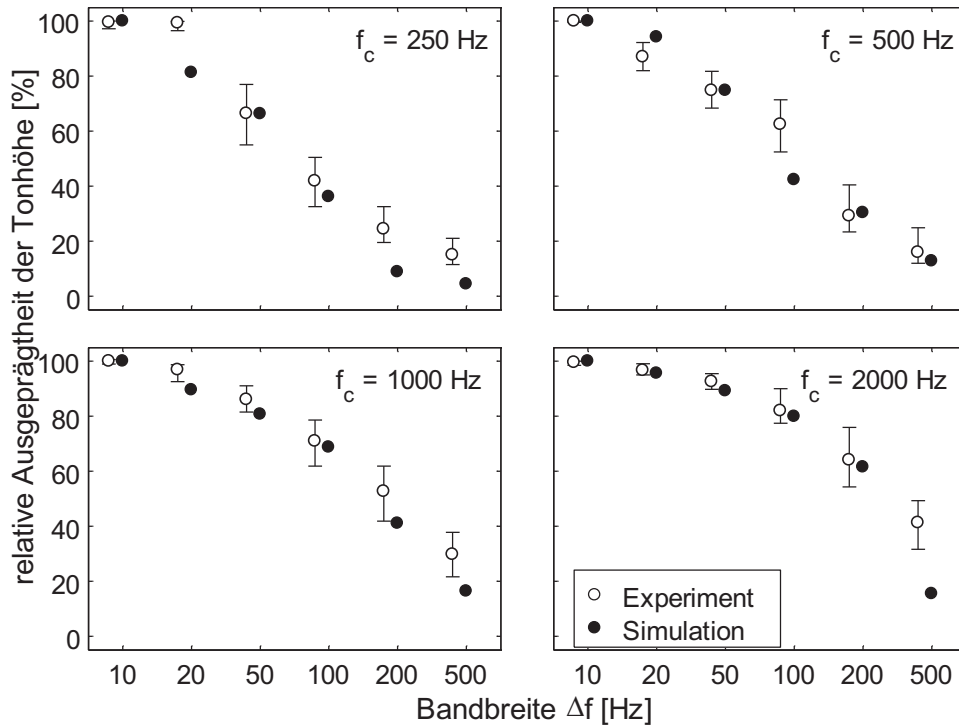


Abbildung 5.7: Versuchsergebnisse und Berechnete Daten zur Ausprägtheit der Tonhöhe von Bandpassrauschen verschiedener Mittenfrequenz. Nicht gefüllte Symbole: subjektive Daten, gefüllte Symbole: Simulation. Daten entsprechend Kapitel 2.8.2.

5.2.6 Ausprägtheit der Tonhöhe von Iterated Rippled Noise

Kammfilterrauschen besitzen ein Spektrum mit harmonischen Anteilen. So verhalten sich auch die Werte der Ausprägtheit der Tonhöhe weitestgehend analog zu den für harmonisch komplexe Töne gefundenen Werten. Abbildung 5.8 zeigt Versuchsergebnisse und Simulationen für die Ausprägtheit der virtuellen Tonhöhe bei Anwendung verschiedener Filterbedingungen in einer Gegenüberstellung.

Die Versuchsergebnisse werden vom Modell lediglich qualitativ nachvollzogen. Die Rangfolge der berechneten Werte entspricht der der Versuchsergebnisse, jedoch läßt sich aufgrund der geringen Stichprobengröße keine Aussage über die Signifikanz machen. Mit mittleren Abständen vom Median größer als 300 % liegen die berechneten Werte sehr

weit von den Versuchsergebnissen entfernt, die Ausgeprägtheit der Tonhöhe wird vom Modell für hohe Verzögerungszeiten der Iterated Rippled Noises deutlich schwächer bewertet als von den Versuchspersonen. Eine Anpassung der frequenzabhängigen Gewichtung W_{f_0} könnte hier für eine Verbesserung sorgen. Da jedoch eine Verzögerungszeit von 32 ms einer Frequenz von 31,25 Hz entspricht, dominiert hier die psychoakustische Wahrnehmung "Rauhigkeit", die Wahrnehmung der Tonhöhe tritt in den Hintergrund. Offenbar werden hier die Urteile der Versuchspersonen durch die stark wahrnehmbare Rauhigkeit beeinflusst.

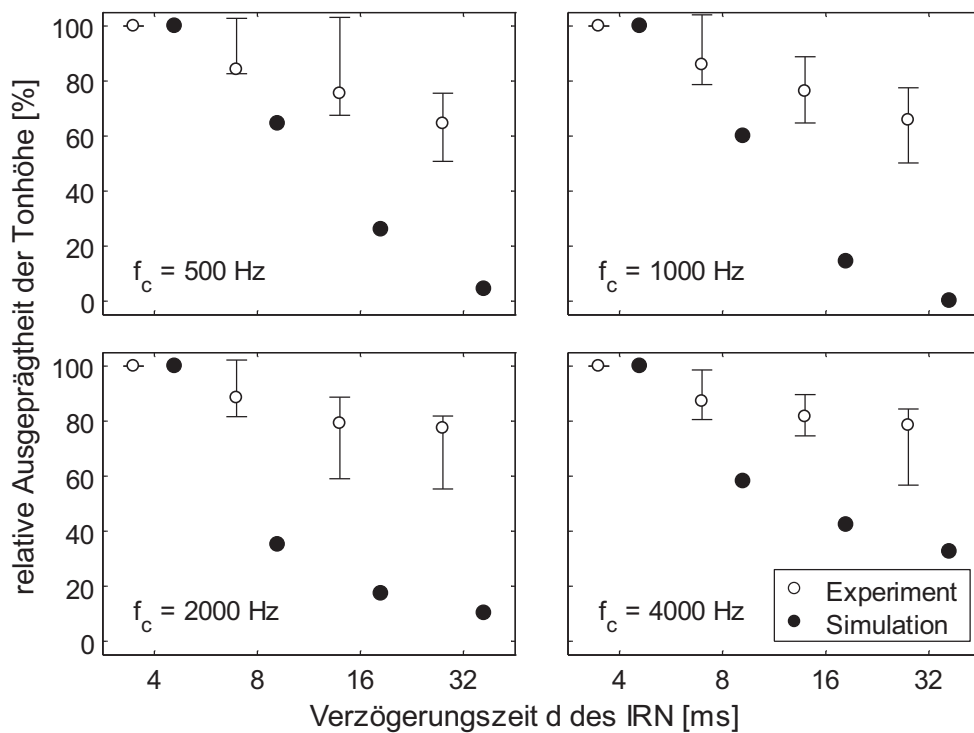


Abbildung 5.8: Versuchsergebnisse und Berechnete Daten zur Ausgeprägtheit der Tonhöhe von Iterated Rippled Noise. Daten entsprechend Kapitel 2.7.1.

Die berechneten Werte der Ausgeprägtheit der Tonhöhe sind aufgrund der spektralen Lage im Existenzbereich der virtuellen Tonhöhe insgesamt sehr klein. So ist die korrekt ermittelte Rangfolge trotz der deutlichen quantitativen Abweichungen als Bestätigung des Modells zu werten.

5.2.7 Vergleich verschiedener Schallarten

Als Überblick über die Leistungsfähigkeit des erstellten Modells zum Vergleich der Ausgeprägtheit der Tonhöhe verschiedenster Schalle sollen an dieser Stelle nochmals die Versuche von Fastl und Stoll herangezogen werden [Fas79c]. Die dabei verwendeten Schalle besitzen stark unterschiedlich ausgeprägte Tonhöhen, Sinustöne beispielsweise besitzen

$f_c =$	500 Hz	1000 Hz	2000 Hz	4000 Hz
RANKCORR	1,00	1,00	1,00	1,00
Signifikant ($\alpha = 0,05$)	Angabe nicht möglich (n = 4)			
PCTINTERQ [%]	25	25	25	25
MEANDIST [%]	614	330	349	304

Tabelle 5.8: Gütekriterien der Modellierung der Versuchsdaten zur Ausprägtheit der Tonhöhe von Iterated Rippled Noise als Anker und Testschall.

eine sehr deutlich ausgeprägte Tonhöhe und werden mit Bandpass- und Wiederholungsrauschen verglichen, welche relativ schwach ausgeprägte Tonhöhen besitzen.

Voranehend wurde bereits gezeigt, daß das Modell innerhalb einer Schallklasse (z.B. Bandpassrauschen) in der Lage ist, die Versuchsergebnisse weitestgehend nachzuvollziehen. Unklar blieb dabei jedoch die quantitative Bewertung unterschiedlicher Schallarten bei ähnlicher nomineller Tonhöhe.

Abbildung 5.9 zeigt die Daten der Ausprägtheit der Tonhöhe verschiedener Schalle im Vergleich zu den berechneten Ergebnissen. Wie ersichtlich und Tabelle 5.9 zu entnehmen ist, stimmen die ermittelten Rangfolgen der Ausprägtheit der Tonhöhe (RANKCORR > 0,85) mit den Versuchsergebnissen überwiegend gut überein. Die Abstände von den Medianen der Versuchsergebnisse zeigen für eine Grundfrequenz von 125 Hz eine schlechtere Modellierung der Versuchsergebnisse im Vergleich zu höheren Grundfrequenzen (MEANDIST = 172 % im Vergleich zu 47 % bei $f_0 = 500$ Hz), was auf den stark

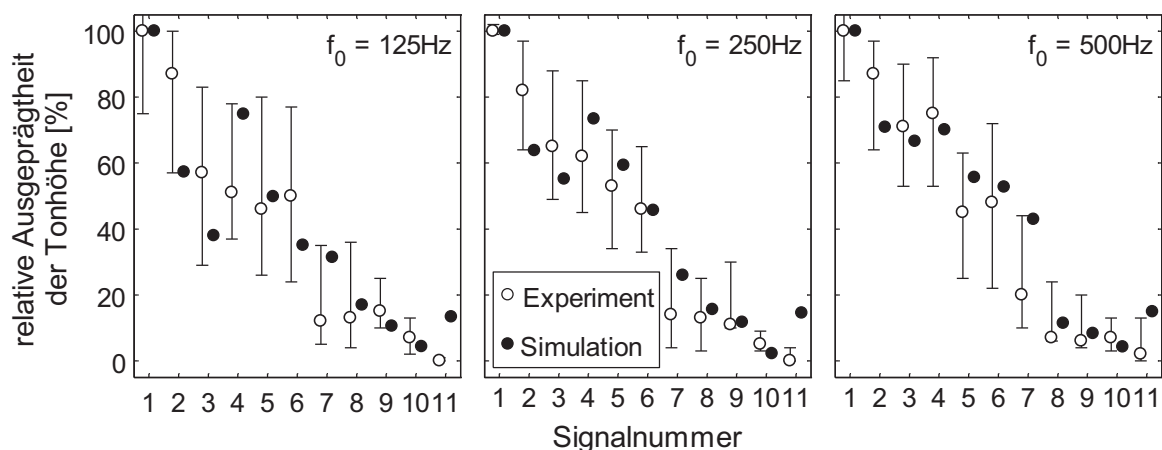


Abbildung 5.9: Berechnete Werte der Ausprägtheit der Tonhöhe für die von Fastl und Stoll in [Fas79c] verwendeten Schalle im Vergleich zu deren Versuchsergebnissen. Die Reihenfolge der Schalle (Signalnummer) entspricht einer Sortierung nach ihrer Ausprägtheit der Tonhöhe von stark nach schwach.

von den Versuchsergebnissen abweichenden Wert des Schalls Nummer 11 (Hochpassrauschen) zurückzuführen ist. Die Werte der Schalle 1-10 liegen für alle drei Grundfrequenzen innerhalb bzw. sehr nahe bei den Interquartilbereichen der Versuchsergebnisse. Da Hochpassrauschen nicht zur Ermittlung der Merkmale der Ausgeprägtheit der Tonhöhe in Kapitel 3 herangezogen wurde, könnte diesbezüglich noch eine Optimierung des Modells durchgeführt werden, beispielsweise durch Einführung der Wahrnehmungsgröße ‘‘Schärfe‘‘ als weiteren, die Ausgeprägtheit der Tonhöhe beeinflussenden, Faktor.

$f_0 =$	125 Hz	250 Hz	500 Hz
RANKCORR	0,85	0,93	0,92
Signifikant ($\alpha = 0,05$)	ja	ja	ja
PCTINTERQ [%]	91	73	91
MEANDIST [%]	172	76	47

Tabelle 5.9: Gütekriterien zur Bewertung der Modellierung der Versuchsdaten zur Ausgeprägtheit der Tonhöhe von Fastl und Stoll nach [Fas79c].

Das Modell ermittelt insbesondere für $f_0 = 125$ bzw. 250 Hz für das Signal 4 ‘‘Schmalbandrauschen‘‘ einen etwas höheren Wert der Ausgeprägtheit der Tonhöhe als die Experimente andeuten. Dies führt zu einer leicht abweichenden Reihenfolge der Ausgeprägtheit der Tonhöhe der untersuchten Schalle. Der ermittelte Wert, für diesen Schall, liegt aber stets innerhalb der Interquartilbereiche der Versuchsergebnisse und so kann von einer geeigneten Modellierung gesprochen werden. Für eine Grundfrequenz von 500 Hz ist zudem zu sehen, daß die Tonhöhe von Schall 4 im Experiment ebenfalls als etwas stärker ausgeprägt bewertet wird als bei Schall 3. Diese Bewertung des Schalls Nummer 4 ‘‘Schmalbandrauschen‘‘ führte zu einer erneuten Untersuchung der Ausgeprägtheit der Tonhöhe der in [Fas79c] verwendeten Schalle mit Hilfe der Methode ‘‘Random Access‘‘. Abbildung 5.10 zeigt den Vergleich der über dieses Experiment ermittelten Rangfolge mit der Rangfolge der Modellberechnungen. Die Tonhöhe von Schall 4 wurde in diesem Experiment jeweils als stärker ausgeprägt bewertet als in den Versuchen aus [Fas79c], was zu einer leicht abweichenden Rangfolge der Ergebnisse gegenüber den ursprünglichen Versuchen führt.

Für eine Grundfrequenz von 250 Hz ist eine sehr gute Übereinstimmung der berechneten Werte mit den Ergebnissen des Hörversuchs zu erkennen. Entsprechend Tabelle 5.10 liegen neun der elf berechneten Werte innerhalb der Interquartilbereiche der Versuchsergebnisse (PCTINTERQ = 82 %). Der mittlere Abstand zum Median der Daten aus den Experimenten entspricht mit 64 % in Etwa der Hälfte der zur Normierung herangezogenen Interquartilbereiche. Bei einer Grundfrequenz von 500 Hz zeigen die Gütekriterien in Tabelle 5.10 ähnliche Werte. Der mittlere normierte Abstand MEANDIST von den Medianen der Versuchsergebnissen ist mit 62 % etwas geringer als bei einer Grundfrequenz von 250 Hz und 73 % der Werte liegen innerhalb der Interquartilbereiche. Somit bestätigen die Gütekriterien eine gute Übereinstimmung, die Hypothese positiver Korrelation kann

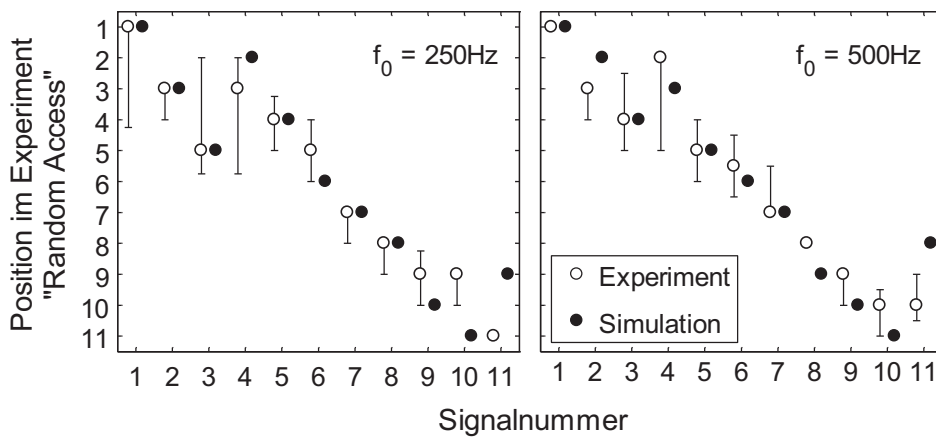


Abbildung 5.10: Berechnete Werte der Ausprägtheit der Tonhöhe im Vergleich zu den Ergebnissen mithilfe der Versuchsmethode “Random Access“ für verschiedene Schallarten entsprechend Kapitel 2.3.

erneut auf einem Signifikanzniveau von 5 % nicht abgelehnt werden.

$f_0 =$	250 Hz	500 Hz
RANKCORR	0,97	0,95
Signifikant ($\alpha = 0,05$)	ja	ja
PCTINTERQ [%]	82	73
MEANDIST [%]	64	62

Tabelle 5.10: Gütekriterien für die Modellierung der Rangfolge der Ausprägtheit der Tonhöhe verschiedener Schalle für die mittels “Random Access“ gewonnenen Versuchsdaten.

So wurde also über verschiedene Schallarten hinweg die Zuverlässigkeit des erstellten Modells nachgewiesen. Das Modell ist nicht nur in der Lage, Versuchsergebnisse für Vergleiche innerhalb einer Schallart nachzuvollziehen, sondern auch Vergleiche unterschiedlicher Schallarten sind möglich und zeigen korrekte Ergebnisse.

5.3 Zusammenfassung

In diesem Kapitel wurden die Berechnungen des in Kapitel 4 hergeleiteten Modells der Ausprägtheit der Tonhöhe mit Versuchsergebnissen aus Abschnitt 2 verglichen. Zur objektiven Bewertung der Modellberechnungen wurden Gütekriterien in Form des Rangkorrelationskoeffizienten nach Spearman, einem Abstandsmaß mit Normierung auf die jeweiligen Interquartilbereiche sowie dem Kriterium, ob berechnete Werte innerhalb des Interquartilbereichs der Ergebnisse aus den Experimenten liegen, angewandt. Die Signi-

fikanz des verwendeten Rangkorrelationskoeffizienten wurde dabei auf einem Niveau von $\alpha = 0,05$ untersucht, was eine Aussage über die Zuverlässigkeit der Übereinstimmung der gewonnenen Rangfolgen zulässt.

Die Vergleiche mit grundlegenden Experimenten zur Ausgeprägtheit der Tonhöhe von Sinustönen, aber auch mit Daten für spektrale Tonhöhen harmonisch komplexer Töne, zeigen eine sehr gute Übereinstimmung des Modells mit den experimentell ermittelten Daten. Die Ergebnisse der Modellberechnungen liegen für diese Daten weitestgehend innerhalb der Interquartilbereiche der experimentell gewonnenen Daten.

Die Versuchsergebnisse zur virtuellen Tonhöhe harmonisch komplexer Töne zeigen teilweise eine Überschätzung der Ausgeprägtheit der Tonhöhe durch das Modell. Dagegen unterschätzt das Modell die Ausgeprägtheit der Tonhöhe der untersuchten Kammfilterrauschen bei sehr niedrigen absoluten Tonhöhen. Die Rangfolge innerhalb einer Schallklasse wird dabei jedoch immer korrekt wiedergegeben.

Bei amplitudenmodulierten Sinustönen werden auch für Schalle, deren spektrale Lage im Existenzbereich der virtuellen Tonhöhe auf eine sehr schwache Ausbildung der virtuellen Tonhöhe hinweist, gute Ergebnisse erzielt, wobei insbesondere die Rangfolge der Versuchsergebnisse durch das Modell sehr gut nachvollzogen wird. Bei niedrigen Modulationsfrequenzen bestätigt sich die Einführung des Gewichts aufgrund von Amplitudenschwankungen W_{AM} in Abschnitt 4. Bei schnellen Modulationen überwiegt die Wahrnehmung der Modulationsfrequenz als virtuelle Tonhöhe, was wiederum die Zweckmässigkeit der Rückrechnung der Tonhöhengewichte virtueller Tonhöhen in Pegelwerte (Abschnitt 4.4.2) zur Repräsentation der wahrgenommenen Tonhöhe bestätigt.

Der abschließende Vergleich mit dem Experiment zur Ausgeprägtheit der Tonhöhe verschiedener Schallarten nach Fastl und Stoll [Fas79c] zeigt eine gute Übereinstimmung, sowohl der ermittelten Rangfolgen als auch der quantitativen Werte der berechneten Ausgeprägtheit der Tonhöhe mit den zugehörigen Versuchsergebnissen. Lediglich Hochpassrauschen werden, verglichen mit den Versuchsergebnissen, als zu stark ausgeprägt bewertet.

Das Modell ist also nicht nur in der Lage innerhalb einer bestimmten Kategorie von Schallen diese miteinander zu vergleichen, sondern ebenso Schalle unterschiedlichster Zusammensetzung bezüglich ihrer Ausgeprägtheit der Tonhöhe zu bewerten und zueinander in Relation zu setzen. Da die bislang untersuchten Ergebnisse ausschließlich von Experimenten mit synthetischen Signalen stammen werden im anschließenden Kapitel reale Schalle bezüglich ihrer Bewertung durch Versuchspersonen und das Modell verglichen um die praktische Anwendbarkeit des Modells zu verifizieren.

Kapitel 6

Anwendung der Ausgeprägtheit der Tonhöhe zur Geräuschbeurteilung

Bisher wurde die Ausgeprägtheit der Tonhöhe ausschließlich verwendet um die Tonhöhe synthetischer Schalle zu bewerten. Untersuchungen dieser Art können zwar Aufschluß und Hinweise über die Arbeitsweise des menschlichen Gehörs bieten und so indirekt beispielsweise zur Entwicklung von Hörgeräten beitragen, eine unmittelbare Anwendung in der Ingenieurs-Praxis fehlt jedoch.

Zahlreiche Veröffentlichungen zeigen aber, daß insbesondere in der jüngeren Zeit die Frage nach Bewertungskriterien zur Beurteilung tonaler Komponenten in Produktgeräuschen an Bedeutung gewinnt. In diesem Zusammenhang wird bereits teilweise die Ausgeprägtheit der Tonhöhe als subjektives Maß verwendet. Auch hier stellt sich wiederum die Frage, ob ein Algorithmus in der Lage ist, die Aussagen von Probanden zu prognostizieren um so möglicherweise als objektives Kriterium eingesetzt werden zu können.

Im Folgenden werden einige Untersuchungen der Ausgeprägtheit der Tonhöhe realer Schalle sowie realen Schallen nachempfundenen Signalen berechneten Daten. Zur Beurteilung der Modellvorhersagen werden dabei die auch in Abschnitt 5.2 verwendeten Gütekriterien herangezogen. Abschließend erfolgt beispielhaft ein Vergleich mit dem in DIN 45 681 beschriebenen Verfahren zur Beurteilung tonaler Geräuschanteile.

6.1 Bewertung der Klangqualität realer und synthetischer Geräusche

6.1.1 Bewertung der Klangqualität von Orchesterpauken

Die eingehende Untersuchung des Klanges von Orchesterpauken ergab einerseits, daß eine steigende Tonhöhe mit einer steigenden Ausgeprägtheit der Tonhöhe einhergeht, was mit den grundlegenden Untersuchungen für Sinustöne aus Abschnitt 2.1 übereinstimmt. Zudem wurde aber festgestellt, daß die Benutzung von Naturfellen gegenüber Kunstfellen stärker ausgeprägte Tonhöhen hervorruft, was wiederum von den befragten Berufsmusikern bevorzugt wird [Fas92b, Fle91, Fle94].

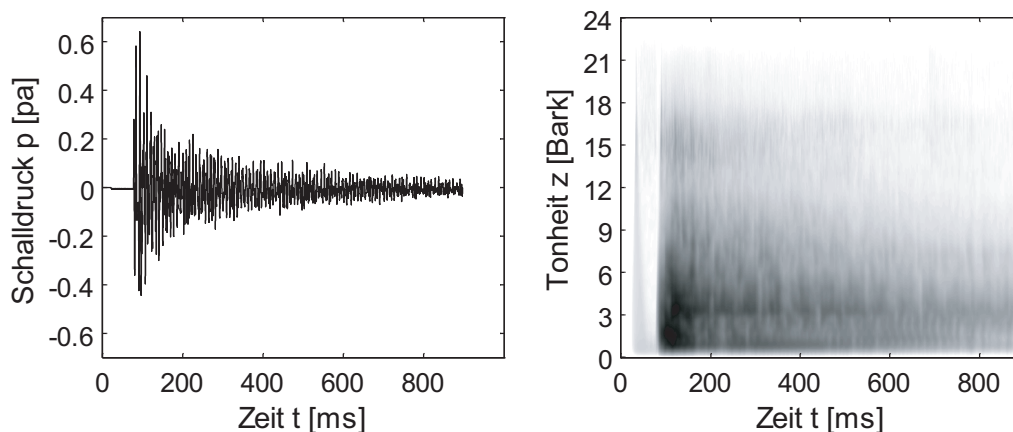


Abbildung 6.1: Beispiel eines Paukenklangs. Links: Zeitsignal, rechts: Lautheits-Tonheits-Zeit Muster.

Versuchspersonen sind also in der Lage, die perkussiven Klänge von Pauken bezüglich ihrer Ausgeprägtheit der Tonhöhe zu beurteilen, obwohl diese Tonhöhen nur schwach ausgeprägt sind. Zudem geben die Bewertungen einen Hinweis auf die Präferenz bei unterschiedlicher Bespannung des Instruments derart, daß die stärker ausgeprägte Tonhöhe des Naturfells gegenüber dem Klang des Kunstfells von den Musikern bevorzugt wird.

In Abbildung 6.2 sind einige Ergebnisse aus [Fle94] im Vergleich zu den Berechnungen mit dem hier entwickelten Algorithmus dargestellt. Die berechneten Daten entstammen dabei jeweils den Berechnungen für 4 Einzelschalle. Diese sind den Stereo-Aufnahmen zweier Paukenanschläge entnommen. Zur Bewertung durch das Modell wurden jeweils Ausschnitte der Länge 1 s extrahiert und unter Vorgabe der nominellen Tonhöhe dem Modell zugeführt. Wie Abbildung 6.1 zeigt, unterscheidet sich das Lautheits-Tonheits-Zeit Muster solch eines Schalls sehr deutlich von den Schallen, die in den grundlegenden Untersuchungen in Kapitel 2 verwendet wurden. Es weist ein deutliches Abklingen über der Zeit auf und die spektralen Komponenten sind lediglich schwach erkennbar.

Dennoch ist der Verlauf der berechneten Werte sehr ähnlich mit denen der Hörversuche. Für das Kunstfell wird jedoch für eine Stimmung von 110 Hz (der musikalische Ton a_2) als Median der vier Berechnungen ein geringerer Wert, als für 83 Hz (e_2) ermittelt. Aufgrund der in Abschnitt 4.10 beschriebenen grundfrequenzabhängigen Gewichtung sollte hier ein Anstieg ähnlich den Daten des Hörversuchs auftreten. Die Interquartile der Modellvorhersagen zeigen aber eine sehr große Schwankung der Berechnungen für die vier zugrunde liegenden Schallausschnitte. Die Eigenschaft höherer Ausgeprägtheit der Tonhöhe des Naturfells gegenüber dem Kunstfell kann ebenfalls nachvollzogen werden. Da beide Bespannungen sehr schwach ausgeprägte Tonhöhen hervorrufen ist diese Tatsache besonders hervorzuheben, zeigt sie doch erneut die Anwendbarkeit des aus Daten synthetischer Schalle entstandenen Modells auf reale Schalle.

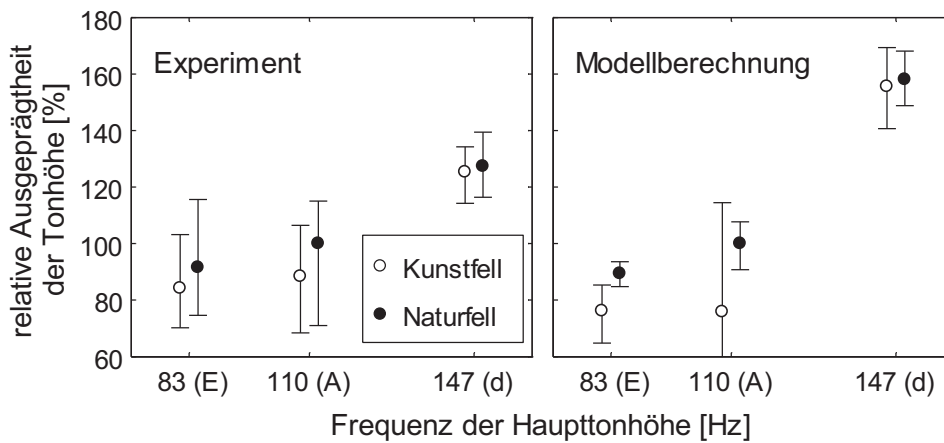


Abbildung 6.2: Vergleich der experimentellen Daten nach [Fle94] zur Ausgeprägtheit der Tonhöhe von Orchesterpauken mit den berechneten Werten der Schlagtonhöhen E, A und d. Offene Symbole: Kunstfell, schwarz ausgefüllte Symbole: Naturfell. Links: experimentelle Daten, Rechts: berechnete Werte.

Die Werte der Gütekriterien zeigen, daß die quantitative Übereinstimmung der Berechnungen mit den Versuchsergebnissen schlechter als beispielsweise für die in Abschnitt 5.2.3 zur Verifikation herangezogenen harmonisch komplexen Töne ist. Jeweils zwei der drei ermittelten Mediane der Berechnungen liegen innerhalb der Interquartilbereiche der Versuchsergebnisse, der mittlere Abstand von den Medianen der Versuchsergebnisse ist mit 85 (Naturfell) bzw. 145 % (Kunstfell) relativ groß. Eine Selektion der berechneten Werte bzw. eine Einschränkung auf beispielsweise zwei geeignete Schallausschnitte hätte hier sicherlich eine Verbesserung zur Folge. Es wurde aber bewusst auf eine solche Optimierung verzichtet, um die Ergebnisse weder zu schönen noch zu verfälschen. Das Modell ist also prinzipiell in der Lage auch für solch perkussive Klänge sinnvolle Vorhersagen zur Ausgeprägtheit der Tonhöhe zu treffen. Die quantitative Übereinstimmung zeigt noch Anpassungsbedarf im Detail, ist möglicherweise aber in der Wahl der verwendeten Schallausschnitte begründet, da nicht mehr bekannt ist, welche konkreten Ausschnitte für die ursprünglichen Experimente verwendet wurden.

	Naturfell	Kunstfell
RANKCORR	1,00	0,50
Signifikant ($\alpha = 0,05$)	Angabe nicht möglich (n = 3)	
PCTINTERQ [%]	67	67
MEANDIST [%]	85	145

Tabelle 6.1: Gütekriterien für die Modellierung der Versuchsdaten zur Ausgeprägtheit der Tonhöhe von Paukenklängen.

6.1.2 Beurteilung der Geräuschqualität in Hochgeschwindigkeitszügen

Im Innenraum von Personenzügen spielt die Geräuschqualität eine besondere Rolle für die Fahrgäste. Da hier die Reisenden selbst nicht durch die Fahraufgabe abgelenkt werden, wie dies bei PKW zumindest für den Fahrer der Fall ist, erhält das Hintergrundgeräusch eine besondere Bedeutung. Dabei besteht eine gewisse Diskrepanz zwischen der Lautheit des Geräusches und der subjektiv empfundenen Privatsphäre. Ein Reisegast möchte nicht das Gefühl haben, jeder Mitreisende kann seine Konversation mitverfolgen, will aber ebenso nicht angestrengt und übertrieben laut sprechen müssen, wenn er ein Gespräch führt. Zudem ist aber auch die Klangfarbe des Geräusches von entscheidender Bedeutung. Starke tonale Komponenten wirken, wie bereits mehrfach erwähnt, sehr belästigend und würden bei mehrstündiger Fahrt das Komfortempfinden der Fahrgäste stark beeinträchtigen.

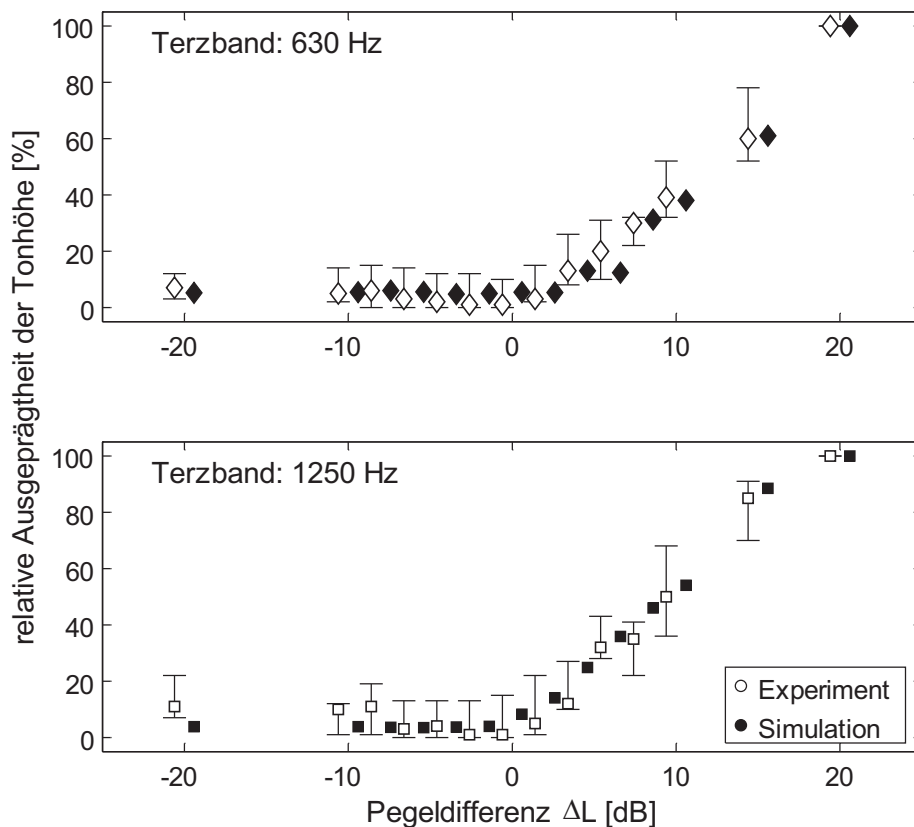


Abbildung 6.3: Vergleich der experimentellen Daten (offene Symbole) zur Ausgeprägtheit der Tonhöhe des Innengeräusches von Hochgeschwindigkeitszügen nach [Pat00] mit den berechneten Werten (gefüllte Symbole) des Modells der Ausgeprägtheit der Tonhöhe. Oben: Pegelvariation der Terz bei 630 Hz, Unten: Pegelvariation der Terz bei 1250 Hz.

In [Pat00] wurden synthetisch erstellte Innengeräusche von Personenzügen untersucht, die jeweils durch Anheben oder Absenken des Pegels innerhalb eines Terzbandes (Mittenfrequenzen 630 bzw. 1250 Hz) variiert wurden. Solche Variationen der Innengeräusche treten in der Praxis aufgrund von äußeren Einflüssen (z.B. verriffelte Gleise) durchaus häufig auf. Die Versuchspersonen sollten mit Hilfe der freien Größenschätzung (also ohne Ankerschall) diese Schalle bezüglich ihrer Ausgeprägtheit der Tonhöhe bewerten.

In Abbildung 6.3 sind die Ergebnisse der Hörversuche im Vergleich zu den Berechnungen mit dem hier entwickelten Modell dargestellt. Die Eigenschaft, daß eine Pegelanhebung innerhalb eines Frequenzbandes zu einer Verstärkung der Ausgeprägtheit der Tonhöhe führt, kann selbst bei geringen Pegelanhebungen sehr gut nachvollzogen werden. Die Beobachtung, daß eine Pegelabsenkung, also eine Bandsperre, ebenfalls zu einer leichten Erhöhung der Ausgeprägtheit der Tonhöhe führt, kann mit dem implementierten Algorithmus derzeit nicht nachvollzogen werden.

Grund hierfür ist das Fehlen von Bandsperrenrauschen bei den grundsätzlichen Untersuchungen aus Kapitel 2. Bei der Entstehung des Modells spielten diese Schalle keine Rolle und so wurde auch nicht explizit auf die Modellierung der Ergebnisse solcher Signale Wert gelegt. Die Überhöhung der Bandkanten (spektrale Inhibition), wie sie beispielsweise in [vR99] zur Beschreibung der Tonhöhenwahrnehmung vorgeschlagen wird, könnte möglicherweise auch im Modell der Ausgeprägtheit der Tonhöhe Anwendung finden und so zu besseren Ergebnissen in diesen Fällen führen.

	$f_{Terz} = 630 \text{ Hz}$	$f_{Terz} = 1250 \text{ Hz}$
RANKCORR	0,89	0,77
Signifikant ($\alpha = 0,05$)	ja	ja
PCTINTERQ [%]	100	86
MEANDIST [%]	26	60

Tabelle 6.2: Gütekriterien für die Modellierung der Versuchsdaten nach [Pat00] für die Geräuschqualität im Inneren von Zügen.

Die Werte der Gütekriterien in Tabelle 6.1.2 zeigen vor allem für die Variation des Terzbandes bei einer Mittenfrequenz von 630 Hz einen sehr geringen mittleren Abstand von den Medianen der experimentellen Daten (MEANDIST = 26 %). Zudem liegen alle berechneten Werte innerhalb der Interquartilbereiche. Für das Terzband bei 1250 Hz ist die Übereinstimmung geringfügig schlechter (PCTINTERQ = 86 %, MEANDIST = 60 %), dennoch erreicht der Rangkorrelationskoeffizient in beiden Fällen Werte über 0,75, was zum Erfüllen des gewählten Signifikanzkriteriums ausreicht.

6.1.3 Beurteilung von Lüftergeräuschen

Viele technische Geräte benötigen zur Unterstützung der Kühlung Ventilatoren, die aufgrund ihrer Rotation Drehklänge erzeugen. Genannt seien an dieser Stelle beispielhaft Personal Computer, deren Prozessoren (CPUs) aufgrund steigender Taktfrequenzen mehr und mehr elektrische Leistung aufnehmen, die überwiegend in Wärme umgesetzt wird. So stieg die Leistungsaufnahme von Desktop CPUs in den vergangenen zehn Jahren von ca. 12 W (1995, Taktfrequenz 120 MHz) auf ca. 120 W (2005, Taktfrequenz 3000 MHz) um den Faktor 10 an [tom05].

Auch LCD- und DLP-Projektoren zur Präsentation stehender und bewegter Bilder besitzen schon allein aufgrund der verwendeten lichtstarken Lampen eine Leistungsaufnahme von ca. 200 W. Diese muß zum überwiegenden Teil in Form von Abwärme aus dem Gehäuse transportiert werden. Auch hier erzeugen die dabei verwendeten Lüftungssysteme Geräusche, die als störend empfunden werden, insbesondere wenn sie wahrnehmbare tonale Komponenten enthalten.

Zahlreiche Veröffentlichungen mit industriellem Hintergrund zeigen, daß man sich auch im kommerziellen Bereich bewußt ist, daß dieses Problem besteht. Geeignetes Sound Design gilt dabei zunehmend als Kaufargument für den Kunden. So sind mittlerweile Prozessorhersteller selbst dabei, geeignete Richtlinien und Kriterien zur Bewertung von Lüftergeräuschen zu erstellen [Tra04, Bel03, Tra05] und die PC-Industrie versucht dem Problem durch neue Richtlinien für den Gehäusaufbau von Personal Computern, beispielsweise in Form des neuen BTX-Standarts zu begegnen [Win04]. Auch die Hersteller von Projektoren sind bemüht die Geräusche ihrer Geräte zu optimieren [Whi05]. Dabei wird zudem der Veränderung dieser Geräusche im Laufe des Lebenszyklus' eines Geräts besondere Beachtung geschenkt [ANS95].

Um Lüftergeräusche zu simulieren bietet es sich an, harmonisch komplexe Töne und rosa Rauschen zu überlagern [Hua05]. Abbildung 6.4 zeigt Versuchsergebnisse für Schalle mit drei verschiedenen Grundfrequenzen. Der Gesamtpegel wurde dabei so gewählt, daß die Lautheit der Schalle einen Wert von 5 sone erreicht. Die tonale Komponente des Schalls wurde mit Hilfe von harmonisch komplexen Tönen, bestehend aus 9 Komponenten mit 3 dB Pegelabfall je Harmonischer simuliert. Das Pegelverhältnis ΔL aus rosa Rauschen und harmonisch komplexem Ton wurde in einem Bereich von -16 bis +16 dB variiert. Zur Untersuchung wurde die Methode Größenschätzung mit Ankerschall verwendet, als Anker diente der Schall mit einem Pegelverhältnis $\Delta L = 0 \text{ dB}$.

Die Ergebnisse zeigen eine steigende Ausgeprägtheit der Tonhöhe mit zunehmendem relativen Pegel des harmonisch komplexen Tons. Dieses Verhalten zeigt dabei keine Abhängigkeit von der verwendeten Grundfrequenz. Bei den Signalen mit Grundfrequenzen von 250 und 500 Hz ist die Tendenz einer Sättigung der Ausgeprägtheit der Tonhöhe ab relativen Pegeln von $\Delta L > +12 \text{ dB}$ zu beobachten. Bei starker Reduzierung des Pegels

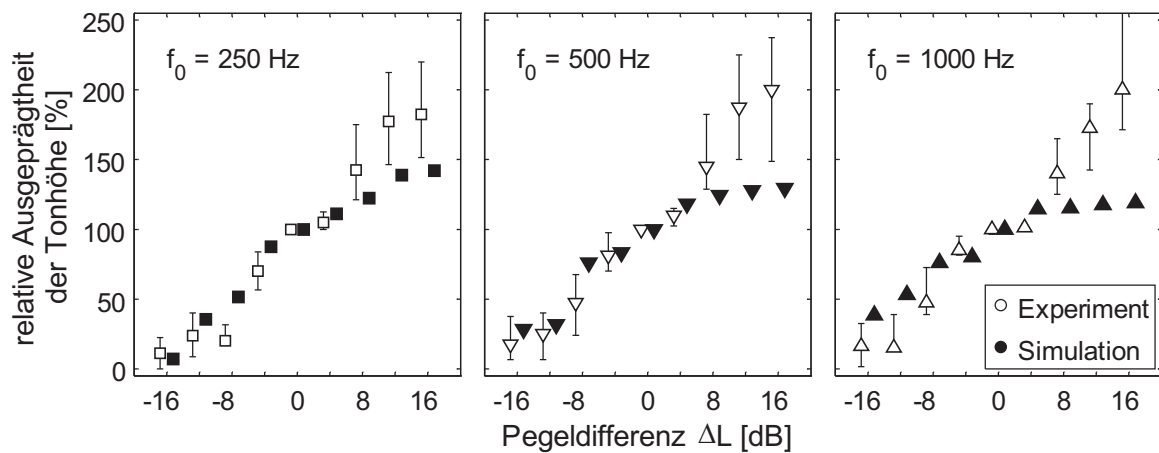


Abbildung 6.4: Ausgeprägtheit der Tonhöhe von simulierten Lüftergeräuschen bei drei unterschiedlichen Grundfrequenzen. Wiedergabelautheit $N_T = 5$ sone, Schalldauer $T_T = 0,5$ s.

des harmonisch komplexen Tons ist das untersuchte Signal von rosa Rauschen nur noch schwach zu unterscheiden, da hier die tonalen Komponenten nahezu vollständig durch das Rauschen verdeckt werden. So lassen sich die Versuchsergebnisse sehr geringer Werte der Ausgeprägtheit der Tonhöhe erklären, welche auch vom Modell nachvollzogen werden können.

	$f_0 = 250\text{Hz}$	$f_0 = 500\text{Hz}$	$f_0 = 1000\text{Hz}$
RANKCORR	0,98	1,00	0,98
Signifikant ($\alpha = 0,05$)	ja	ja	ja
PCTINTERQ [%]	56	44	11
MEANDIST [%]	105	93	190

Tabelle 6.3: Gütekriterien für die Modellierung der Versuchsdaten zur Ausgeprägtheit der Tonhöhe simulierter Lüftergeräusche.

Wie die Gütekriterien in Tabelle 6.1.3 veranschaulichen, ist die ermittelte Rangfolge in Einklang mit den Versuchsergebnissen. Die quantitativen Ergebnisse liegen jedoch vor allem für sehr hohe Pegeldifferenzen ΔL relativ weit von den Versuchsergebnissen entfernt.

Auffällig ist, daß die Modellberechnungen bei hohen Pegelüberschüssen nur noch einen geringen Anstieg der Ausgeprägtheit der Tonhöhe zeigen. Der in den Versuchen beobachtete Sättigungseffekt tritt hier also ebenfalls, jedoch bereits bei geringerem Pegelüberschuß auf.

6.2 DIN 45 681: Ein Verfahren zur Bewertung tonaler Komponenten

Eng mit der Ausgeprägtheit der Tonhöhe verbunden ist der Begriff "Tonhaltigkeit". Während die Ausgeprägtheit der Tonhöhe ein globales Urteil für einen Schall erlaubt, beschreibt die Tonhaltigkeit das Vorhandensein und die Deutlichkeit einzelner spektraler Komponenten, sie ist also der in Abschnitt 4.4.1 beschriebenen "Pitch Saliency" ähnlich.

Nachfolgend wird stellvertretend für technische Verfahren zur Bewertung der Tonhaltigkeit das in DIN 45 681 dargestellte Verfahren beschrieben [DIN05]. Dabei werden lediglich die Grundgedanken aufgeführt und Details der jeweiligen Umsetzung ausgespart. Ähnliche Prinzipien finden beispielsweise auch in ANSI S1.13-1995 und ISO 7779-1999 in Form von "Tone-to-Noise Ratio" und "Prominence Ratio" Anwendung [ANS95, ISO99].

Abschließend wird ein Vergleich mit der psychoakustischen Größe "Ausgeprägtheit der Tonhöhe" durchgeführt, um ausgehend von Gemeinsamkeiten und Unterschieden Ausblicke auf mögliche zukünftige Bewertungsverfahren zu geben.

6.2.1 Übersicht über DIN 45 681

Ziel des in DIN 45 681 beschriebenen Verfahrens ist die Ermittlung eines Pegelzuschlags K_T um den gemessenen Schallpegel eines Geräusches bei Vorhandensein tonaler Komponenten zu erhöhen und damit der störenden Wirkung Rechnung zu tragen. Bislang erfolgte die Vergabe dieses Tonzuschlags aufgrund des subjektiven Urteils des jeweiligen Gutachters. Hintergrund ist hierbei die Lästigkeit tonaler Komponenten in Schallen, die eigentlich keine oder nur sehr schwach wahrnehmbare Tonhöhen enthalten sollten, was beispielsweise bei den Schallimmissionen großer Industrieanlagen der Fall ist.

In gemessenen Spektren der zu bewertenden Schalle wird zur Ermittlung des Tonzuschlags K_T für jede hervortretende Spektralkomponente innerhalb einer Frequenzgruppe der Pegelüberschuß ΔL über den Geräuschanteil in der umgebenden Frequenzgruppe gebildet. Die Tonfrequenz f_T wird dabei als geometrischer Mittelwert der relevanten Frequenzgruppe mit den Grenzfrequenzen f_1 und f_2 interpretiert, so daß gilt:

$$f_T = \sqrt{f_1 f_2} \text{ Hz} \quad (6.1)$$

Die zu bewertende tonale Komponente läßt sich über die Frequenz f_T und den Pegel L_T beschreiben. Über den Pegel L_G der durch f_1 und f_2 begrenzten Frequenzgruppe und das Verdeckungsmaß a_v ergibt sich der Pegelüberschuß ΔL zu:

$$\Delta L = L_T - L_G - a_v \quad [dB] \quad (6.2)$$

Abhängig von den so ermittelten Pegelüberschüssen wird ein Tonzuschlag, entsprechend Abbildung 6.5, ermittelt. Der mittlere Pegelüberschuß wird schließlich zum gemessenen

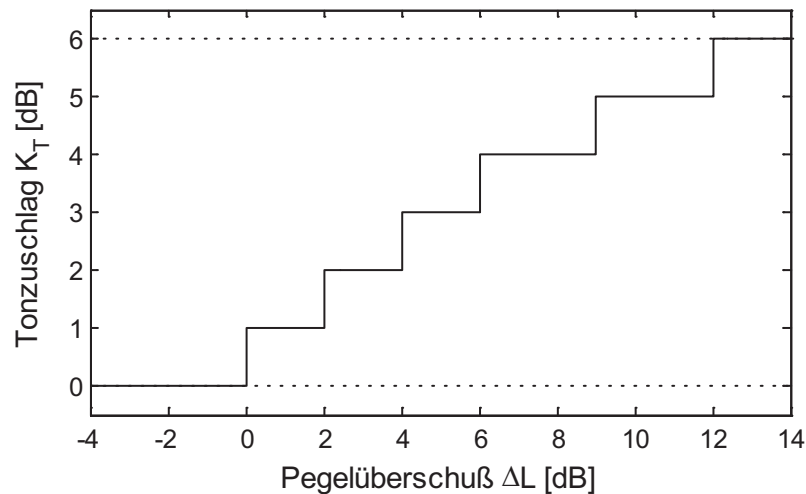


Abbildung 6.5: *Tonzuschlag K_T in Abhängigkeit des Pegelüberschusses ΔL gemäß DIN 45 681.*

Summenpegel hinzuaddiert. Dieser Tonzuschlag beträgt für einzelne Teiltöne maximal 6 dB und summiert sich für den Gesamtschall durch Bewertung aller tonalen Komponenten zu maximal 10 dB. Während in [Sei96] und [Bec96] nur eine geringe Übereinstimmung subjektiver Beurteilungen und berechneter Werte festgestellt wurde, zeigte sich in [Pom98] die Praktikabilität dieses rechnerischen Tonzuschlags. Die von Gutachtern bestimmten Pegelzuschläge konnten hier durch eine geringfügige Anpassung der Umrechnung des Pegelüberschusses in den Tonzuschlag gut nachvollzogen werden.

Die Richtlinie wird derzeit laufend um weitere Einflußgrößen erweitert und so den subjektiven Ergebnissen und neuen Erkenntnissen angepasst. Insbesondere im Hinblick auf die Teils sehr schlechte Übereinstimmung mit subjektiven Daten ist eine zukünftig weiter reichendere Berücksichtigung psychoakustischer Hintergründe nicht auszuschließen.

6.2.2 Vergleich zur Ausprägtheit der Tonhöhe

Das beschriebene Verfahren zur Bewertung der Tonhaltigkeit von Schallen bestimmt für einzelne Tonkomponenten einen Pegelzuschlag, der ein gewisses Maximum nicht übersteigen kann. Verglichen mit der Ausprägtheit der Tonhöhe ist dieses Vorgehen sinnvoll, da diese durch die maximale Ausprägtheit der Tonhöhe von Sinustönen ebenfalls begrenzt ist. Die Pegelabhängigkeit der Ausprägtheit der Tonhöhe wird in gewissen Grenzen durch den gemessenen Gesamtpegel erfasst, bei festem Pegel werden die Schalle bezüglich ihrer Tonhaltigkeit mit Hilfe der Tonzuschläge zusätzlich bewertet.

Was die technischen Verfahren jedoch bisher nicht berücksichtigen, ist die Interaktion einzelner Teiltöne. Werden mehrere Teiltöne detektiert, so wird für jeden Ton entsprechend definierter Kriterien ein Tonzuschlag ermittelt und diese Einzelwerte aufsummiert. Die

Daten der Ausgeprägtheit der Tonhöhe zeigen jedoch, daß vor allem im Fall harmonischer Teiltöne die Ausgeprägtheit der Tonhöhe abnimmt, wenn mehrere Komponenten vorhanden sind. Zwar ist in den jeweiligen Bestimmungsvorschriften der maximale Tonzuschlag begrenzt, ein Anstieg des Pegelzuschlags mit zunehmender Anzahl tonaler Komponenten kann so aber nicht vermieden werden.

Insgesamt können die technischen Verfahren als sinnvoller Ansatz gesehen werden, die belästigende Wirkung tonaler Komponenten in einen vergleichbaren Maßstab umzusetzen. Aus Sicht der Psychoakustik wäre jedoch die Berücksichtigung der Abhängigkeiten der Ausgeprägtheit der Tonhöhe wünschenswert um eine wahrnehmungsgerechte Bewertung des tonalen Charakters eines Schalles zu erzielen.

6.3 Zusammenfassung

Der Vergleich experimenteller und berechneter Daten für Schalle aus dem Bereich der Geräuschbeurteilung zeigte grundsätzliche Übereinstimmung mit den Versuchsergebnissen. Anhand von Paukenklängen konnte die Anwendbarkeit des Algorithmus auf reale Schalle gezeigt werden. Hier wurden sowohl die grundsätzlichen Abhängigkeiten der Ausgeprägtheit der Tonhöhe als auch die Präferenz von Berufsmusikern für Naturfelle anstelle künstlicher Paukenbespannungen korrekt nachvollzogen.

Stellvertretend für Lüftergeräusche und Innengeräusche von Personenzügen wurden Geräusche untersucht, die diesen realen Schallen synthetisch nachempfunden sind. Die berechneten Werte stimmen auch hier prinzipiell mit den experimentell ermittelten Daten überein. Für die untersuchten Schalle gilt jeweils, daß die Ausgeprägtheit der Tonhöhe relativ gering ist, so daß es für ein technisches Verfahren schwierig ist, tonale Komponenten korrekt zu ermitteln. Daß das in dieser Arbeit eingeführte Modell der Ausgeprägtheit der Tonhöhe dennoch gute Übereinstimmung mit den experimentellen Daten zeigt, spricht für die in Kapitel 4 dargestellte Umsetzung des Algorithmus.

Ein Vergleich mit technischen Verfahren zur Beurteilung tonaler Komponenten zeigte schließlich Gemeinsamkeiten und Unterschiede der Begriffe "Tonhaltigkeit" und "Ausgeprägtheit der Tonhöhe" auf. Durch die in DIN 45 681 beschriebene technische Realisierung des bisher subjektiv vergebenen Tonzuschlags ist hierfür ein objektives Verfahren verfügbar. Unterschiede zur Ausgeprägtheit der Tonhöhe zeigen aber, daß in ein wahrnehmungsgerechtes Bewertungsverfahren tonaler Komponenten weitere psychoakustische Größen, insbesondere die Ausgeprägtheit der Tonhöhe, einfließen sollten.

Kapitel 7

Zusammenfassung

Die vorliegende Arbeit stellt Ergebnisse aus psychoakustischen Experimenten zur Wahrnehmung der Ausgeprägtheit der Tonhöhe sowie ein daraus abgeleitetes Modell zur Beschreibung und Berechnung dieser psychophysikalischen Größe vor.

In Kapitel 1 werden dazu benötigte Grundlagen vorgestellt. Dabei handelt es sich sowohl um physiologische Grundlagen zum Verständnis des menschlichen Gehörs als auch um Begriffserklärungen und Beschreibungen der verwendeten Versuchsmethoden. Insbesondere die beschriebene Frequenz-Orts-Transformation stellt eine elementare Basis für die später entwickelten Grundlagen und Zusammenhänge zum Verständnis der gewonnenen Versuchsergebnisse dar.

Kapitel 2 stellt die Ergebnisse grundlegender Untersuchungen zur psychoakustischen Wahrnehmung der Ausgeprägtheit der Tonhöhe dar. Ausgehend von bekannten Daten aus der Literatur, die grundsätzliche Abhängigkeiten für Sinustöne von deren Dauer, Frequenz und Pegel darstellen, werden weitere Experimente erläutert und analysiert, um die weitreichenden Einflußgrößen zur Bestimmung der Ausgeprägtheit der Tonhöhe zu erfassen. So werden Experimente für spektrale und virtuelle Tonhöhen von harmonisch komplexen Tönen dargestellt. Dabei zeigt sich eine Abnahme der Ausgeprägtheit der (spektralen) Tonhöhe bei Erhöhung der Anzahl enthaltener Harmonischer unabhängig davon, ob die Schalle auf gleichen Wiedergabepiegel oder gleiche Lautheit kalibriert werden. Die Ausgeprägtheit virtueller Tonhöhen zeigt bei Entfernen unterer Teiltöne eine Abnahme mit steigender Anzahl dieser entfernten untersten Harmonischen. Ein Experiment mit Pegelvariation der ersten Harmonischen (also des Grundtons) kann nicht nur qualitativ sondern auch quantitativ den Übergang von spektraler zu virtueller Tonhöhe nachvollziehen.

Um nicht nur Schalle zu untersuchen, die sich über eine Reihe von Sinuskomponenten beschreiben lassen, sondern auch eher rauschhafte Signale mit in die Modellbildung einzubeziehen, wurden daran anschließend Kammfiltereffekte untersucht. Dabei wurden sogenannte Iterated Rippled Noises generiert, indem eine einzelne Kammfilterschaltung mehrfach durchlaufen wird. Dabei zeigt sich eine Erhöhung der Ausgeprägtheit der Tonhöhe mit zunehmender Anzahl dieser Durchläufe. Zudem wurden Tonhöhen von Bandpassrauschen untersucht um mögliche Abhängigkeiten von deren Bandbreite und spektraler Lage zu erfassen. Diese zeigen sich sowohl in einer generellen Abnahme der Ausgeprägtheit der

Tonhöhe bei Vergrößerung der Bandbreite und auch in einem langsameren Abfall bei höheren als bei niedrigen Mittenfrequenzen.

Die Ergebnisse dieser Versuche werden jeweils mit Hilfe der Modellvorstellung beschrieben, daß die Basilarmembran als vorverarbeitendes Element des Hörsystems mit einer Filterbank bestehend aus Bandpässen vergleichbar ist. Die Ausgeprägtheit der Tonhöhe ist in dieser Vorstellung dann maximal, wenn der Schall durch diese Filterbank optimal aufgelöst werden kann, wenn also möglichst wenige "Elementarfilter" durch das Signal angeregt werden. Insbesondere für Sinustöne ist dies anschaulich und die Abnahme der Ausgeprägtheit der Tonhöhe von Bandpassrauschen läßt sich so ebenfalls unmittelbar erklären. Bei harmonisch komplexen Tönen und Kammfilterrauschen muß man jeweils die Auflösbarkeit der einzelnen Komponenten betrachten. Ein Problem hierbei ist die korrekte Skalierung der Filterbandbreiten- und Mittenfrequenzen, die dem Modell zugrunde gelegt wird. Da die Untersuchungen keine Hinweise auf eine abweichende Parametrierung geben, wird die etablierte Skalierung entsprechend der Frequenzgruppenbreite und Tonheit angewandt.

Um Merkmale zur Beschreibung der experimentell gewonnenen Daten zu erhalten, wird in Kapitel 3 eine lautheitsskalierte Spektrogrammdarstellung eingeführt. Diese beinhaltet sowohl ein an die hohe spektrale Auflösung der verwendeten Fourier-t-Transformation angepasstes Lautheitsmodell als auch ein Berechnungsverfahren der resultierenden Mithörschwellenmuster. So ergibt sich eine Signaldarstellung, welche die prinzipiellen Schalleigenschaften in einer gehörgerechten Form darstellt. Über die gewonnenen Spektrogramme (Lautheits-Tonheits-Zeit Muster), zeitliche Ausschnitte daraus (instantanes Spektrum) und dem Lautheits-Zeit Verlauf lassen sich Eigenschaften ableiten, die für die untersuchten Schalle Merkmale zur Beschreibung ihrer Ausgeprägtheit der Tonhöhe darstellen. Dabei handelt es sich um die spektrale Lage der zu bewertenden Tonhöhe, die Bandbreite des untersuchten Schalls sowie die Lautheit des Signal- und Störanteils. Desweiteren lassen sich zeitliche Schwankungen der Tonhöhe als auch des Verlaufs der Lautheit erkennen. Im Fall von Residualklängen, deren Tonhöhe als virtuelle Tonhöhe wahrgenommen wird, liefert ein bereits etablierter Algorithmus ein Gewicht zur Beschreibung der Bedeutung der untersuchten Tonhöhe für den Gesamtschall.

Die Extraktion und weitere Verarbeitung dieser Merkmale in Form eines Modells zur Beschreibung der Ausgeprägtheit der Tonhöhe werden in Kapitel 4 geschildert. Das Berechnungsverfahren baut auf der Verwendung vorhandener und etablierter Verfahren auf. So werden die bereits erwähnte Fourier-t-Transformation, das Lautheitsmodell nach Zwicker sowie Verfahren zur Bestimmung der Tonhöhe und virtuellen Tonhöhe nach Terhardt eingesetzt. Die mit diesen Mitteln errechneten Werte werden zusätzlich aufgrund der vorangehend gefundenen Schallmerkmale gewichtet. Dabei werden die Gewichte psychoakustisch motiviert und beschreiben Einflüsse beispielsweise der Wahrnehmungsgrößen Schwankungsstärke und Lautheit.

Die Qualität der mittels des Modells berechneten Daten wird in Kapitel 5 anhand dreier Gütekriterien beurteilt. Dazu werden Daten aus Hörversuchen mit zugehörigen Berechnungen verglichen. Als Kriterium dient dabei der Rangkorrelationskoeffizient nach Spearman, der die prinzipielle Tendenz der Daten miteinander vergleicht. Da dieses Kriterium keine Aussage über die quantitativen Ergebnisse zuläßt, werden zusätzlich der Prozentsatz berechneter Daten innerhalb des Interquartilbereichs der Versuchsergebnisse, sowie der mittlere Abstand vom Median (bezogen auf die Perzentile der Versuchsergebnisse) als Kriterien herangezogen. Beim Rangkorrelationskoeffizienten zeigt sich dabei jeweils eine sehr gute Übereinstimmung der Berechnungen mit den experimentellen Daten. Das Modell beschreibt also die Abhängigkeiten in Grundzügen korrekt. Die schärferen Kriterien zeigen bei grundlegenden Untersuchungen von Sinustönen und harmonisch komplexen Tönen sehr geringe Abweichungen von den Versuchsergebnissen. Für komplexere Schalle insbesondere bei sehr geringer Ausprägtheit der Tonhöhe steigen die quantitativen Abweichungen von den Versuchsergebnissen etwas an, wobei die Rangfolge stets korrekt wiedergegeben wird.

Als Anwendung der Ausprägtheit der Tonhöhe wird in Kapitel 6 auf die Bewertung der Klangqualität von Orchesterpauken, sowie die Geräuschbeurteilung von technischen Geräten eingegangen. Während in der musikalischen Akustik im aufgeführten Beispiel eine höhere Ausprägtheit der Tonhöhe (Naturfell gegen Kunstfell) wünschenswert ist und von Berufsmusikern bevorzugt wird, ist die Präferenz in der Geräuschbeurteilung in der Regel entgegengesetzt. Hier besitzt die Ausprägtheit der Tonhöhe einen starken Einfluß auf die Lästigkeit eines Schalls. Beispiele für das Innengeräusch von Personenzügen sowie das Lüftergeräusch von Personal Computern zeigen, anhand der genannten Kriterien, erneut eine prinzipielle Übereinstimmung der berechneten Daten mit den Versuchsergebnissen und somit eine Anwendbarkeit des Modells auf reale und realen Schall nachempfundene Signale.

Ein kurzer Vergleich mit technischen Bewertungsverfahren tonaler Komponenten anhand des Verfahrens aus DIN 45 681 zeigt, daß die Ausprägtheit der Tonhöhe in Form eines Modells zur objektiven Schallbeurteilung beitragen kann. Insbesondere die Interaktion harmonischer Komponenten, wie sie bei harmonisch komplexen Tönen auftritt, wird durch die Ausprägtheit der Tonhöhe besser beschrieben als mithilfe des in DIN 45 681 geschilderten Prinzips eines Pegelzuschlags.

Die große Anzahl veröffentlichter Untersuchungen und Berichte zur Beurteilung tonaler Komponenten zeigt das industrielle Interesse an einer sinnvollen Bewertung von Schall bezüglich dieser Eigenschaften. Die Ausprägtheit der Tonhöhe als psychoakustische Größe und damit die Schallwahrnehmung des Menschen widerspiegelnde Empfindung kann hier noch wichtige Beiträge leisten.

Danksagung

An dieser Stelle möchte ich meinen Dank an alle Personen aussprechen, die zum Gelingen dieser Arbeit auf ihre Weise beigetragen haben.

Besonderer Dank geht an Prof. Dr.-Ing. H. Fastl, der nicht nur die Arbeit initiierte sondern auch selbst durch zahlreiche experimentelle Untersuchungen im Laufe seiner Karriere mir hilfreiche Beiträge zu den dargestellten Forschungsarbeiten lieferte. Hervorzuheben ist der große Freiraum und die gleichzeitige Diskussionsbereitschaft in allen Phasen meiner wissenschaftlichen Tätigkeit.

Den Kollegen am Lehrstuhl für Mensch-Maschine-Kommunikation möchte ich für das angenehme Arbeitsklima und die fortwährende Unterstützung in Form von Espresso und “guten“ Ratschlägen danken. Die Kollegen aus der AG Technische Akustik, vor allem Dr.-Ing. C. Patsouras, für die nette Zimmernachbarschaft, und Dipl.-Ing. S. Kerber, für die schöne Zusammenarbeit, möchte ich dabei besonders erwähnen.

Ein ewiger Quell akustischer Weisheit darf hier auch nicht unerwähnt bleiben: Der Stammtisch zur Kontaktpflege mit ehemaligen Kollegen erwies sich mit mal mehr, mal weniger fachgerechtem Diskussionsstoff als Ratgeber in allen Lagen. Dabei möchte ich vor allem Dr.-Ing. M. Mummert für die ausdauernde Diskussion seiner Forschungsarbeiten und Dipl.-Ing. T. Horn für weitsichtige und inspirierende Ratschläge danken.

Nicht zu vergessen sind natürlich die zahlreichen Versuchspersonen, die geduldig an Experimenten teilnahmen und teilweise im Rahmen von Bachelor-, Studien- und Diplomarbeiten zu den gesammelten Ergebnissen beitrugen. Hervorheben möchte ich hier Dipl.-Ing. D. Menzel, Dipl.-Ing. F. Klüber, Dipl.-Ing. J. Putner, Dipl.-Ing. G. v. Boogaart sowie Dipl.-Ing. S. Kilger und Dipl.-Ing. N. Köhler.

Prof. Dr.-Ing. K. Diepold danke ich für die Übernahme des Mitberichts.

Meiner Freundin Uta-Marie danke ich für ihre große Geduld mit meiner immerwährenden Ungeduld, vor allem in der anstrengenden Schlußphase solch einer Arbeit.

Ein Teil dieser Arbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen der Forschergruppe “Hörobjekte“ FA 140/1 gefördert.

Anhang A

Abkürzungen und Formelzeichen

AFC	Alternative Forced Choice
AABB	Rauschen zusammengesetzt aus wiederholten Segmenten
AdT	Ausgeprägtheit der Tonhöhe
AKF	Autokorrelationsfunktion
AM	Amplitudenmodulation
AM-SIN	amplitudenmodulierter Sinuston
BBR	Breitbandrauschen
BPR	Bandpassrauschen
DIST	Abstand der Modellberechnung zu den Versuchsergebnissen
DLM	Dynamic Loudness Model
FM	Frequenzmodulation
FM-SIN	frequenzmodulierter Sinuston
FTT	Fourier-t-Transformation
HKT	Harmonisch Komplexer Ton
HKRT	Harmonisch Komplexer Residualton
IRN	Iterated Rippled Noise
IRN-O	Iterated Rippled Noise, "Add Original"
IRN-S	Iterated Rippled Noise, "Add Same"
JND	Just-Noticeable-Difference, Unterschiedsschwelle
JNDPS	Unterschiedsschwelle der Ausgeprägtheit der Tonhöhe
MEANDIST	mittlerer Abstand der Modellberechnung zu den Versuchsdaten
MHS	Mithörschwelle
PCTINTERQ	Prozentsatz innerhalb Interquartilbereich
PEST	Parameter Estimation by Sequential Testing
RANKCORR	Rangkorrelationskoeffizient nach Spearman
R-FTT	Rücktransformation der FTT
RHS	Ruhehörschwelle
SBR	Schmalbandrauschen
SIN	reiner Sinuston
STFT	Kurzzeit Fourier Transformation
TTZM	Teiltonzeitmuster
AdT_M	Modellberechnung der Ausgeprägtheit der Tonhöhe

a_v	psychoakustisches Verdeckungsmaß
C_{f_0}	Korrektur der Umrechnung von Tonhöhengewichten in Pegelwerte
CH_{f_0}	Zur f_0 gehörender Analysekanal der FTT
CH_{2f_0}	Analysekanal der zweiten Harmonischen
$E_{AM}(f_0)$	Exponent zur Anpassung von W_{AM} an die Grundfrequenz
Δf	Bandbreite
Δf_G	Frequenzgruppenbreite
Δf_{rel}	relative Bandbreite
f	Frequenz
f_c	Mittelfrequenz
f_0	Grundfrequenz
f_{mod}	Modulationsfrequenz
f_{mod}^*	Eigenmodulationsfrequenz von Bandpassrauschen
f_M	Frequenz eines maskierenden Schalls
f_{SP}	Frequenz einer Spektraltonhöhe
f_T	Testtonfrequenz
f_{Tr}	Trägerfrequenz bei AM-SIN
f_{VP}	Frequenz einer virtuellen Tonhöhe
FTT_N	Pegelspektrogramm des Störanteils
FTT_{SP}	Pegelspektrogramm der spektralen Tonhöhen
FTT_{SIG}	Pegelspektrogramm des Eingangssignals
FTT_T	Pegelspektrogramm des Signalanteils
FTT_{VP}	Pegelspektrogramm der virtuellen Tonhöhen
K_T	Pegelzuschlag nach DIN 45 681
H	Tonhöhe in Pitch Units (pu)
ΔL	Pegelüberschuß über benachbarte Spektralanteile
ΔL_1	Pegeldämpfung der ersten Harmonischen eines HKT
ΔL_f	spektrale Pegeldämpfung in dB / Oktave
L_M	Pegel eines maskierenden Schalls
$L_{f_0,SP}$	Pegel der spektralen Tonhöhe f_0
$L_{f_0,VP}$	“Pegel“ der virtuellen Tonhöhe f_0
L_G	Frequenzgruppenpegel
L_{SP}	Pegel einer spektralen Tonhöhe
L_{VP}^*	Äquivalenter Pegel einer virtuellen Tonhöhe
$MAX(FTT)$	Maximalwert des aktuellen FTT Spektrums
$MAX(AKF)$	Maximum der Autokorrelationsfunktion
$MAX(AKF, glob)$	Globales Maximum der Autokorrelationsfunktion
$MAX(AKF, loc)$	Lokales Maximum der Autokorrelationsfunktion
MED	Median einer Menge von (Versuchs-) ergebnissen
MED_{EXP}	Median der Versuchsergebnisse
MHS_N	Mithörschwellenmuster des Störanteils
MHS_{SP}	Mithörschwellenmuster der spektralen Tonhöhen

MHS_T	Mithörschwellenmuster des Signalanteils
MHS_{VP}	Mithörschwellenmuster der virtuellen Tonhöhen
$MULT(f_0)$	Grundfrequenz-Anpassung der ‘‘Pegel‘‘ virtueller Tonhöhen
ΔN	Lautheitshub bei Amplitudenschwankung
$n_{Schalle}$	Anzahl untersuchter Schalle
n_h	Anzahl Harmonischer eines HKT
N_5	Perzentillautheit (entsprechend 95 % Perzentil des Lautheits-Zeit Verlaufs eines Schalls)
N_{95}	Perzentillautheit (entsprechend 5 % Perzentil des Lautheits-Zeit Verlaufs eines Schalls)
N_0	Lautheit des Signalanteils
N_N	Lautheit des Störanteils
N_T	Lautheit des Testschalls
$p_{0,SP}$	Referenz zur Umrechnung von Tonhöhengewichten in Pegelwerte
POS_{f_0}	Kanalnummer des für f_0 ermittelten Maximums der FTT
$POS_{f_0,SP}$	Kanalnummer der für f_0 ermittelten spektralen Tonhöhe
$POS_{f_0,VP}$	Kanalnummer der für f_0 ermittelten virtueller Tonhöhe
$Q_{1,Exp}$	Wert des ersten Quartils der Versuchsergebnisse
$Q_{3,Exp}$	Wert des dritten Quartils der Versuchsergebnisse
s_l	Flankensteilheit der MHS für Frequenzen unterhalb f_M
$s_u(L_M, f_M)$	Flankensteilheit der MHS für Frequenzen oberhalb f_M
$SP(f_{SP}, L_{SP}, t_{SP})$	Menge der spektralen Tonhöhen
σ	Standartabweichung einer Datenmenge
t_{SP}	Zeitschritt einer Spektraltonhöhe
t_{VP}	Zeitschritt einer virtuellen Tonhöhe
T_T	Schalldauer
$VP(f_{VP}, L_{VP}^*, t_{VP})$	Menge der virtuellen Tonhöhen
W_{AM}	Gewichtung der AdT aufgrund von Amplitudenschwankungen
W_{BW}	Gewichtung der AdT aufgrund der Bandbreite eines Schalls
W_{F_0}	Gewichtung der AdT aufgrund der Tonhöhe eines Schalls
W_H	Gewichtung aufgrund geringer Standartabweichung von POS_{f_0}
W_N	Gewichtung der AdT aufgrund der Gesamtlautheit eines Schalls
W_{SP}	Gewicht (Pitch Saliency) einer spektralen Tonhöhenkomponente
W_{STD}	Gewichtung aufgrund der Varianz von POS_{f_0}
W_{VP}	Gewicht (Pitch Saliency) einer virtuellen Tonhöhenkomponente
z	Tonheit
Δz_A	Abstand der Analysefilter der FTT
z_c	Mittentonheit
z_{f_0}	Tonheit der Frequenz f_0
z_{2f_0}	Tonheit der zweiten Harmonischen
z_T	Tonheit eines Testtones

Anhang B

Weitere Daten und Versuchsergebnisse

B.1 Amplitudenmodulierte Sinustöne

Untenstehend in Abbildung B.1 sind Daten der Ausgeprägtheit der Tonhöhe amplitudenmodulierter Sinustöne entsprechend Abschnitt 2.2 dargestellt und entsprechend der erwähnten Aufteilung der Versuchspersonen getrennt dargestellt.

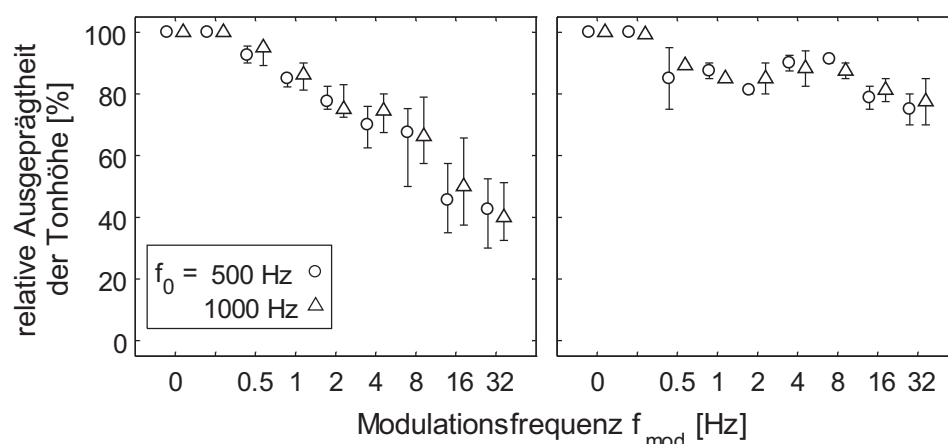


Abbildung B.1: Ausgeprägtheit der Tonhöhe von amplitudenmodulierten Sinustönen. Versuchsdaten getrennt nach Versuchspersonen die eine Überhöhung bei $f_{mod} = 4$ Hz zeigen (rechte Seite) bzw. bei denen diese Überhöhung nicht auftritt (linke Seite).

Die Daten zeigen, daß die Versuchspersonengruppe, die für eine Modulationsfrequenz von 4 Hz eine Überhöhung der Ausgeprägtheit der Tonhöhe zeigt (rechte Seite), generell die Abnahme der Ausgeprägtheit der Tonhöhe mit steigender Modulationsfrequenz recht gering bewertet. Die Überhöhung bei $f_{mod} = 4$ Hz ist jedoch im Umfeld dieser geringen Gesamtabnahme mit einem Anstieg von ca. 80 % auf 90 % recht deutlich.

Der Effekt trat lediglich bei zwei der zehn teilnehmenden Probanden auf. Aufgrund der Übereinstimmung mit dem Maximum der wahrnehmbaren Schwankungsstärke liegt hier möglicherweise eine entsprechende Abhängigkeit der Ausgeprägtheit der Tonhöhe von der Schwankungsstärke vor, was Ausgangspunkt weiterer Untersuchungen sein könnte.

B.2 Übergang von spektraler zu virtueller Tonhöhe

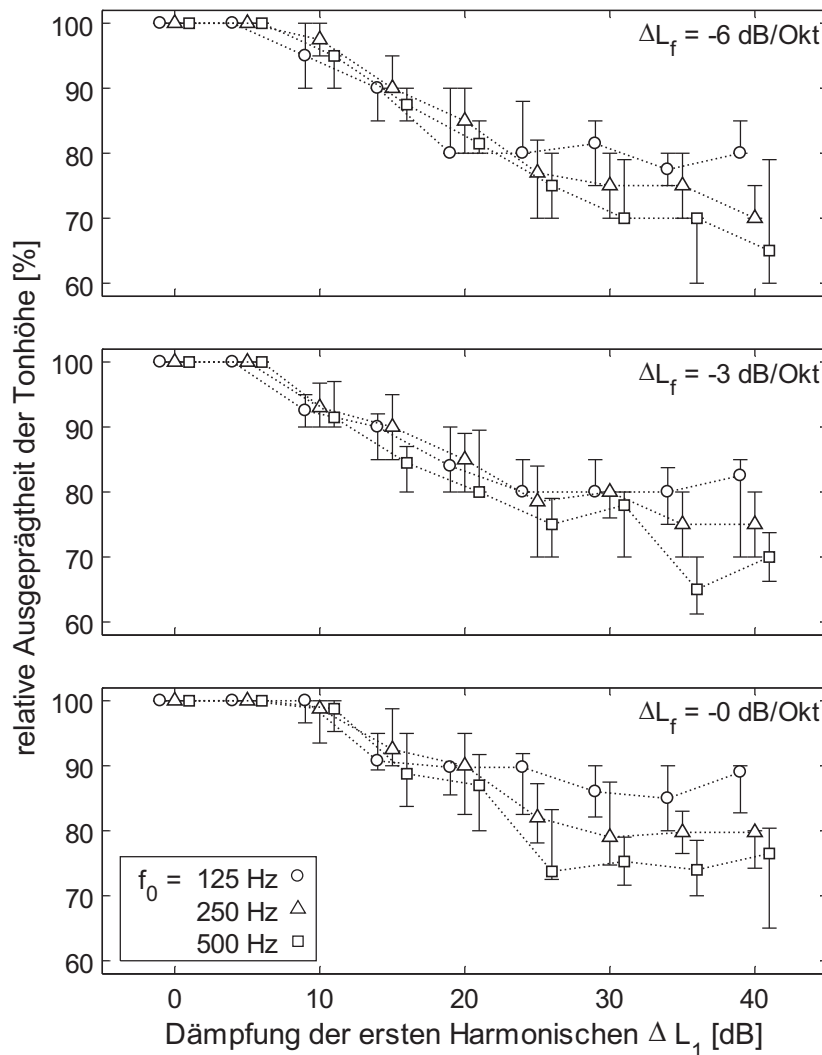


Abbildung B.2: Ausgeprägtheit der Tonhöhe von harmonisch komplexen Tönen bei Pegeldämpfung der ersten Harmonischen. Grundfrequenzen: $f_0 = 125, 250$ und 500 Hz. Zusätzliche spektrale Dämpfung von $0, 3$ und 6 dB/Oktave.

Abbildung B.2 zeigt die vollständigen Daten des in Kapitel 2.6 vorgestellten Versuchs zur Ausgeprägtheit der Tonhöhe harmonisch komplexer Töne mit unterschiedlicher Dämpfung der ersten Harmonischen.

Die zusätzlichen Daten stehen in Einklang mit den im Hauptteil beschriebenen Beobachtungen. Eine gesonderte Diskussion ist somit nicht nötig.

Anhang C

Versuchspersonen

An den Versuchen nahmen jeweils mindestens acht Personen teil. Um sicher zu stellen, daß es sich dabei ausschließlich um Normalhörende handelt, wurde mit Hilfe des Bekesy Tracking die Ruhehörschwelle (RHS) jeder Versuchsperson ermittelt [Kai52]. Es zeigte sich bei der Mehrzahl der Probanden normales Hörvermögen im Bereich ± 10 dB um die mittlere Ruhehörschwelle im Bereich 20 Hz - 16 kHz [Zwi55]. Bei wenigen Versuchspersonen lagen die Werte im erweiterten Bereich von ± 20 dB, vor allem im Bereich 2-5 kHz traten Überhöhungen der RHS auf ("Disco-Loch").

Bei den Probanden handelte es sich überwiegend um Mitarbeiter (wissenschaftliche Mitarbeiter, Diplomanden und Bachelorarbeiter) des Lehrstuhls für Mensch-Maschine-Kommunikation der TU München. Jede Versuchsperson nahm an mehreren Experimenten teil, es ist also davon auszugehen, daß man die Versuchspersonen als geübte Hörer einstufen kann. Ein Großteil der Probanden ist darüberhinaus selbst musikalisch aktiv und hat somit weitere Erfahrungen mit dem Begriff "Tonhöhe".

Literaturverzeichnis

- [ANS95] **ANSI.** *ANSI S1.113-1995: American National Standard Methods for the Measurement of Sound Pressure Levels in Air.* American National Standards Institute, 1995.
- [Aur85] **Aures, W.** *Ein Berechnungsverfahren der Rauigkeit.* *Acustica*, 58: 268–281, 1985.
- [Bau56] **Bauch, H.** *Die Bedeutung der Frequenzgruppe für die Lautheit von Klängen.* *Acustica*, 6: 40–45, 1956.
- [Bec96] **Beckenbauer, T., Stemplinger, I., Seiter, A.** *Basics and use of DIN 45681: Detection of tonal components and determination of a tone adjustment for the noise assessment.* In *Inter-Noise 96.* Liverpool, UK, 3271–3276, 1996.
- [Bek60] **Bekesy, G. v.** *Experiments in Hearing.* McGraw-Hill Book Co., New York-Toronto-London, 1960.
- [Bel03] **Beltman, W. M.** *The airflow and acoustic performance of current 92 mm system fans.* In *Noise-Con 2003.* Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2003.
- [Bil70] **Bilsen, F. A., Ritsma, R. J.** *Repetition Pitch and Its Implication for Hearing Theory.* *Acustica*, 22: 63–73, 1970.
- [Bra94] **Brandenburg, K., Stoll, G.** *ISO-MPEG-1 Audio: A Generic Standard for Coding of High-Quality Digital Audio.* *Journal of the Audio Engineering Society*, 42: 780–792, 1994.
- [Bäu74] **Bäuerle, R.** *Modal-, Zentral, oder arithmetischer Mittel-Wert? Eine Betrachtung zur Auswertung von Meßergebnissen.* *Archiv für technisches Messen, Blatt JO21-22*: 33–34, 1974.
- [Buu96] **Buus, S., Florentine, M., Poulsen, T.** *Temporal integration of loudness, loudness discrimination and the form of the loudness function.* *J. Acoust. Soc. Am.*, 101: 669–680, 1996.
- [Cha97] **Chalupper, J., Schmid, W.** *Akzentuierung und Ausgeprägtheit von Spektraltonhöhen bei harmonischen Komplexen Tönen.* In *Fortschritte der Akustik - DAGA '97.* Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 357–358, 1997.

- [Cha02a] **Chalupper, J.** *Perzeptive Folgen von Innenohrschwerhörigkeit: Modellierung, Simulation und Rehabilitation.* Shaker Verlag, Aachen, 2002.
- [Cha02b] **Chalupper, J., Fastl, H.** *Dynamic loudness model (DLM) for normal and hearing-impaired listeners.* ACUSTICA / Acta Acustica, 88: 378–386, 2002.
- [Cre67] **Creelman, C. D., Taylor, M.M.** *PEST: Efficient Estimates on Probability Functions.* J. Acoust. Soc. Am., 41: 782–787, 1967.
- [Dan02] **Daniel, P.** *Vergleich von subjektiven Daten mit Modellvorhersagen zur Tonhaltigkeit von Geräuschen.* Persönliche Kommunikation, 2002.
- [Deu88] **Deuter, K.** *Gedrosselte Lautheit bei tiefen, mittleren und hohen Frequenzen.* In Fortschritte der Akustik - DAGA '88. Verl.: DPG- GmbH, Bad Honnef, 577–580, 1988.
- [DIN91] **DIN.** *Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach E. Zwicker.* DIN45631, Deutsches Institut für Normung e.V., 1991.
- [DIN94] **DIN.** *Schallpegelmesser, Teil 1: Anforderungen.* DIN EN 61 672-1, Deutsches Institut für Normung e.V., 1994.
- [DIN05] **DIN.** *Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschmissionen.* DIN45681, Deutsches Institut für Normung e.V., 2005.
- [Dix48] **Dixon, W.J., Mood, A.M.** *A Method for Obtaining and Analyzing Sensitivity Data.* J. of the American Statistical Association, 43: 109–1269, 1948.
- [Ern00] **Erne, M.** *Signal adaptive audio coding using Wavelets and rate optimization.* Hartung-Gorre Verlag Konstanz, 2000.
- [Fas71] **Fastl, H.** *Über Tonhöhenempfindungen bei Rauschen.* Acustica, 25: 350–354, 1971.
- [Fas75] **Fastl, H.** *Loudness and Masking Patterns of Narrow Noise Bands.* Acustica, 33: 266–271, 1975.
- [Fas76a] **Fastl, H.** *Masking Patterns of Subcritical versus Critical Band-Maskers at 8.5 kHz.* Acustica, 34: 167–171, 1976.
- [Fas76b] **Fastl, H.** *Temporal Masking Effects: I Broad Band Noise Masker.* Acustica, 35: 287–301, 1976.
- [Fas76c] **Fastl, H.** *Temporal Masking Effects: II. Critical Band Noise Masker.* Acustica, 36: 317–330, 1976.

- [Fas79a] **Fastl, H.** *Comparison of DT 48, TDH 49 and TDH 39 earphones.* J. Acoust. Soc. Am., 66: 702–703, 1979.
- [Fas79b] **Fastl, H.** *Temporal Masking Effects: III. Pure Tone Masker.* Acustica, 43: 282–293, 1979.
- [Fas79c] **Fastl, H., Stoll, G.** *Scaling of Pitch Strength.* Hearing Research, 1: 293–301, 1979.
- [Fas81] **Fastl, H.** *Ausgeprägtheit der Tonhöhe pulsmodulierter Breitbandrauschen.* In Fortschritte der Akustik - DAGA '81. VDE-Verlag, Berlin, 725–728, 1981.
- [Fas88] **Fastl, H.** *Pitch and pitch strength of peaked ripple noise.* In 8th Intern. Symp. on Hearing,. Academic Press, London, 370–379, 1988.
- [Fas89] **Fastl, H.** *Pitch Strength of pure tones.* In 13th International congress on acoustics (ICA), 11–14, 1989.
- [Fas90] **Fastl, H., Wiesmann, N.** *Ausgeprägtheit der virtuellen Tonhöhe von AM- und QFM-Tönen.* In Fortschritte der Akustik - DAGA '90. Verl.: DPG- GmbH, Bad Honnef, 759–762, 1990.
- [Fas92a] **Fastl, H.** *Fluctuation Strength of Narrow Band Noise.* Advances in the Biosciences, 83: 331–336, 1992.
- [Fas92b] **Fastl, H., Fleischer, H.** *Über die Ausgeprägtheit der Tonhöhe von Paukenklängen.* In Fortschritte der Akustik - DAGA '92. Verl.: DPG- GmbH, Bad Honnef, 237–240, 1992.
- [Fas97] **Fastl, H.** *Pitch strength and frequency discrimination for noise bands or complex tones.* 11th International Symposium on Hearing, Grantham, UK, 1st-6th August, 1997.
- [Fas00] **Fastl, H.** *Sound Quality of Electric Razors - Effects of Loudness.* inter-noise 2000, CD-ROM, 2000.
- [Fel53] **Feldtkeller, R., Zwicker, E.** *Die Größe der Elementarstufen der Tonhöhenempfindung und der Lautheitsempfindung.* Acustica, 3: 97–100, 1953.
- [Fel55] **Feldtkeller, R.** *Über die Zerlegung des Schallspektrums in Frequenzgruppen durch das Gehör.* Elektron. Rundschau, 9: 387–389, 1955.
- [Fle34] **Fletcher, H.** *Loudness, pitch and timbre of musical tones and their relations to the intensity, the frequency and the overtone structure.* J. Acoust. Soc. Am., 6: 59–69, 1934.
- [Fle91] **Fleischer, H.** *Akustische Untersuchungen an Orchesterpauken.* In Forschungsbericht aus dem Institut für Mechanik. Universität der Bundeswehr München, 02/91 1991.

- [Fle94] **Fleischer, H.** *Zur Tonhöhe von Paukenklängen.* In Forschungsbericht aus dem Institut für Mechanik. Universität der Bundeswehr München, 02/94 1994.
- [Fru02] **Fruhmann, M., Fastl, H.** *Zur Ausgeprägtheit der Tonhöhe bei Wiederholungsrauschen.* In Fortschritte der Akustik - DAGA '02. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 472–473, 2002.
- [Fru04a] **Fruhmann, M.** *On the pitch strength of bandpass noises.* In 18th International congress on acoustics (ICA), 1791–1794, 2004.
- [Fru04b] **Fruhmann, M.** *On the pitch strength of harmonic complex tones and comb-filter noises.* In Fortschritte der Akustik - CFA/DAGA '04, 757–758, 2004.
- [Fru05a] **Fruhmann, M.** *Introduction and practical use of an algorithm for the calculation of pitch strength.* J. Acoust. Soc. Am., 118: 1894, 2005.
- [Fru05b] **Fruhmann, M., Klüber, F.** *On the pitch strength of complete and incomplete harmonic complex tones.* In Fortschritte der Akustik - DAGA '05. Verl.: Dt. Gesell. für Akustik e. V., Berlin, 467–468, 2005.
- [Fru06] **Fruhmann, M.** *Ein Algorithmus zur Beschreibung der Ausgeprägtheit der Tonhöhe in Theorie und Praxis.* In Fortschritte der Akustik - DAGA '06. Verl.: Dt. Gesell. für Akustik e. V., Berlin, 2006.
- [Gla90] **Glasberg, B. R., Moore, B. C. J.** *Derivation of auditory filter shapes from notched-noise data.* Hearing Research, 47: 103–138, 1990.
- [Gru87] **Grubert, A.** *Tonhöhen harmonischer und inharmonischer komplexer Töne.* In Fortschritte der Akustik - DAGA '87. Verl.: DPG- GmbH, Bad Honnef, 573–576, 1987.
- [Hei88a] **Heinbach, W.** *Aurally adequate signal representation: The Part-Tone-Time-Pattern.* Acustica, 67: 113–121, 1988.
- [Hei88b] **Heinbach, W.** *Gehörgerechte Repräsentation von Audiosignalen durch das Teiltonzeitmuster.* Dissertation, Technische Universität München, 1988.
- [Hel87] **Hellman, R., Zwicker, E.** *Why can a decrease in $dB(A)$ produce an increase in loudness?* J. Acoust. Soc. Am., 82: 1700–1705, 1987.
- [Hel93] **Hellbrück, J.** *Hören - Physiologie, Psychologie und Pathologie.* Hogrefe - Verlag für Psychologie, Göttingen, 1993.
- [Hor06] **Horn, T.** Persönliche Kommunikation, 2001-2006.
- [Hua05] **Huang, L., Lu, H. Z., Wang, J., So, R. M. C.** *Measurement and analysis of noise from a ventilation fan.* In Noise-Con 2005. Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2005.

- [ISO75] **ISO.** *Acoustics - Method for calculating loudness level.* International Organization for Standardization, 532, 1975.
- [ISO99] **ISO.** *Acoustics - Measurement of airborne noise emitted by information technology and telecommunications equipment.* International Organization for Standardization, 7779, 1999.
- [Kai52] **Kaiser, W.** *Das Bekesy-Audiometer der Technischen Hochschule Stuttgart.* Acustica, Akust. Beihefte, AB235–AB238, 1952.
- [Kol88] **Kollmeier, B., Gilkey, R.H., Sieben, U.K.** *Adaptive staircase techniques in psychoacoustics: a comparison of human data and a mathematical model.* J. Acoust. Soc. Am., 83: 1852–1862, 1988.
- [Lee05] **Lee, K. H., Surprenant, A. M.** *Tonal strength of harmonic complex tones in machinery noise.* In Noise-Con 2005. Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2005.
- [Lee06] **Lee, P. M.** *Statistical Tables.* www.york.ac.uk/depts/maths/tables/welcome.html, 2006.
- [Lev70] **Levitt, H.** *Transformed Up-Down Methods in Psychoacoustics.* J. Acoust. Soc. Am., 49: 467–477, 1970.
- [Lum84a] **Lumer, G.** *Addition von Mithörschwellen.* In Fortschritte der Akustik - DAGA '84. Verl.: DPG- GmbH, Bad Honnef, 753–756, 1984.
- [Lum84b] **Lumer, G.** *Überlagerung von Mithörschwellen an den unteren Flanken schmalbandiger Schalle.* Acustica, 54: 154–160, 1984.
- [Mai67] **Maiwald, D.** *Ein Funktionsschema des Gehörs zur Beschreibung der Erkennbarkeit kleiner Frequenz- und Amplitudenänderungen.* Acustica, 18: 81–92, 1967.
- [Mal99] **Mallat, S.** *A wavelet tour of signal processing.* Academic Press, 2nd edition, 1999.
- [Mat14] **Mathworks, The.** *MATLAB - the language of technical computing.* www.mathworks.de, Release 14.
- [Moo85] **Moore, B.C.J., Glasberg, B.R., Peters, R.W.** *Relative dominance of individual partials in determining the pitch of complex tones.* J. Acoust. Soc. Am., 77: 1853–1860, 1985.
- [Mum97] **Mummert, M.** *Software-Paket zur Dissertation von M. Mummert (SWMUMDIS).* <http://mum.homepage.t-online.de/>, 1997.

- [Mum98] **Mummert, M.** *Sprachcodierung durch Konturierung eines gehörangepassten Spektrogramms und ihre Anwendung zur Datenreduktion.* In Fortschrittberichte VDI. VDI Verlag GmbH, Düsseldorf, 1998.
- [Opp89] **Oppenheim, A. V., Willsky, A. S.** *Signale und Systeme.* VCH Verlagsgesellschaft, 1989.
- [Pat76] **Patterson, R. D., Wightman, F. L.** *Residue pitch as a function of component spacing.* J. Acoust. Soc. Am., 59: 1450–1459, 1976.
- [Pat00] **Patsouras, C., Fastl, H., Widmann, U., Hölzl, G.** *Privacy versus Sound Quality in High Speed Trains.* In inter-noise 2000. Multi-Science Publishing Co Ltd, CD-ROM, 2000.
- [Pat03a] **Patsouras, C.** *Geräuschqualität von Fahrzeugen - Beurteilung, Gestaltung und multimodale Einflüsse.* In Berichte aus der Akustik. Shaker Verlag, Aachen, 2003.
- [Pat03b] **Patsouras, Ch., Böhm, M., Fastl, H., Patsouras, D., Pfaffelhuber, K.** *Methodenvergleich zur Beurteilung der Geräuschqualität: Random Access versus Größenschätzung mit Ankerschall.* In Fortschritte der Akustik - DAGA '03. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 240–241, 2003.
- [Pau72] **Paulus, E., Zwicker, E.** *Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln.* Acustica, 27: 253–264, 1972.
- [Peg04] **Peganos, H., Melcher, J.R., Oxenham, A.J.** *A Neural Representation of Pitch Salience in Nonprimary Human Auditory Cortex Revealed with Functional Magnetic Resonance Imaging.* The Journal of Neuroscience, 24: 6810–6815, 2004.
- [Plo64] **Plomp, R.** *The Ear as a Frequency Analyzer.* J. Acoust. Soc. Am., 36: 1628–1636, 1964.
- [Plo67] **Plomp, R.** *Pitch of Complex Tones.* J. Acoust. Soc. Am., 41: 1526–1533, 1967.
- [Plo75] **Plomp, R.** *Auditory psychophysics.* Annu. Rev. Psychol., 26: 207–232, 1975.
- [Pom98] **Pompetzki, W.** *Vergleich der Tonhaltigkeit nach DIN 45 681 mit subjektiven Bewertungen.* In Fortschritte der Akustik - DAGA '98. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 224–225, 1998.
- [Rak02] **Rakowski, A.** *From acoustics to psychology: Pitch strength of sounds.* Music Practitioner, 2002.

- [Ric54] **Rice, S. O.** *Mathematical analysis of random noise*. Noise and Stochastic Processes, Dover Publications, New York, 1954.
- [Rio91] **Rioul, O., Vetterli, M.** *Wavelets and Signal Processing*. IEEE Signal Processing Magazine, 8: 14–38, 1991.
- [Rit62] **Ritsma, R. J.** *Existence Region of the Tonal Residue. I*. J. Acoust. Soc. Am., 34: 1224–1229, 1962.
- [Rit67] **Ritsma, R. J.** *Frequencies dominant in the perception of the pitch of complex sounds*. J. Acoust. Soc. Am., 42: 191–198, 1967.
- [Rus88] **Ruske, G.** *Methoden der Klassifikation*. In Automatische Spracherkennung. Oldenburg Verlag: München, Wien, 23–66, 1988.
- [Sac04] **Sachs, L.** *Angewandte Statistik*. 11. Auflage. Springer Verlag, Berlin, 2004.
- [Sch77] **Schöne, P.** *Nichtlinearitäten im Mithörschwellen-Tonheitsmuster von Sinustönen*. Acustica, 37: 37–44, 1977.
- [Sch78] **Schöne, P.** *Vergleich dreier Funktionsschemata der akustischen Schwankungsstärke*. Biol. Cybernetics, 29: 57–62, 1978.
- [Sch79] **Schöne, P.** *Mithörschwellen-Tonheitsmuster maskierender Sinustöne*. Acustica, 43: 197–204, 1979.
- [Sch90] **Schlang, M., Mummert, M.** *Die Bedeutung der Fensterfunktion für die Fourier-t-Transformation als gehörgerechte Spektralanalyse*. In Fortschritte der Akustik - DAGA '90. Verl.: DPG- GmbH, Bad Honnef, 1043–1046, 1990.
- [Sch97] **Schmid, W.** *Akzentuierung und Ausgeprägtheit von Spektraltonhöhen bei harmonisch Komplexen Tönen*. In Fortschritte der Akustik - DAGA '97. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 357–358, 1997.
- [Sch98] **Schmid, W.** *Akzentuierende Wirkung von Zeigertönen auf Spektraltonhöhen Komplexer Töne*. In Fortschritte der Akustik - DAGA '98. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 468–469, 1998.
- [Sei96] **Seiter, A., Stemplinger, I., Beckenbauer, T.** *Untersuchungen zur Tonhaltigkeit von Geräuschen*. In Fortschritte der Akustik, DAGA 96. Dt. Gesell. für Akustik e. V., Oldenburg,, 238–239, 1996.
- [Smo70] **Smoorenburg, G. F.** *Pitch Perception of Two-Frequency Stimuli*. J. Acoust. Soc. Am., 48: 924–942, 1970.
- [Son82] **Sonntag, B.** *Zur Abhängigkeit der Mithörschwellen-Tonheitsmuster maskierender Sinustöne von deren Tonheit*. Acustica, 52: 95–97, 1982.

- [Ste56] **Stevens, S. S.** *The direct estimation of sensory magnitudes - loudness.* American Journal of Psychology, 69: 1–25, 1956.
- [Str85] **Strube, H. W.** *A Computationally Efficient Basilar-Membrane Model.* Acustica, 58: 207–214, 1985.
- [Ter68a] **Terhardt, E.** *Über akustische Rauigkeit und Schwankungsstärke.* Acustica, 20: 215–223, 1968.
- [Ter68b] **Terhardt, E.** *Über die durch amplitudenmodulierte Sinustöne hervorgerufene Hörempfindung.* Acustica, 20: 210–214, 1968.
- [Ter68c] **Terhardt, E.** *Untersuchungen über die Datenreduktion durch das menschliche Gehör.* Kybernetik, 383–395, 1968.
- [Ter70] **Terhardt, E.** *Oktavspreizung und Tonhöhenverschiebung bei Sinustönen.* Acustica, 22: 345–351, 1970.
- [Ter71] **Terhardt, E.** *Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon.* Seventh international congress on acoustics, Budapest 1971, 621–624, 1971.
- [Ter72a] **Terhardt, E.** *Tonhöhenwahrnehmung und harmonisches Empfinden.* In Akustik und Schwingungstechnik. VDI-Verlag, Düsseldorf, 59–68, 1972.
- [Ter72b] **Terhardt, E.** *Zur Tonhöhenwahrnehmung von Klängen. I. Psychoakustische Grundlagen.* Acustica, 26: 173–186, 1972.
- [Ter74] **Terhardt, E.** *Pitch, consonance and harmony.* J. Acoust. Soc. Am., 55: 1061–1069, 1974.
- [Ter76] **Terhardt, E.** *Frequenz und Tonhöhe.* Instrumentenbau, 30: 232–235, 1976.
- [Ter79] **Terhardt, E.** *Calculating Virtual Pitch.* Hearing Research, 1: 155–182, 1979.
- [Ter82] **Terhardt, E., Stoll, G., Seewann, M.** *Algorithm for extraction of pitch and pitch salience from complex tonal signals.* J. Acoust. Soc. Am., 71: 679–688, 1982.
- [Ter85a] **Terhardt, E.** *Fourier Transformation of time signals: conceptual revision.* Acustica, 57: 242–256, 1985.
- [Ter85b] **Terhardt, E.** *Verfahren zur gehörbezogenen Frequenzanalyse.* In Fortschritte der Akustik - DAGA '85. Verl.: DPG- GmbH, Bad Honnef, 811–814, 1985.
- [Ter86] **Terhardt, E.** *Pitch perception and frequency analysis.* In 6. FASE, 221–228, 1986.

- [Ter92] **Terhardt, E.** *From speech to language: On auditory information processing.* The Auditory Processing of Speech: From Sounds to Words, Mouton de Gruyter, Berlin, 363–380, 1992.
- [Ter98] **Terhardt, E.** *Akustische Kommunikation - Grundlagen mit Hörbeispielen.* Springer Verlag, Berlin / Heidelberg, 1998.
- [tom05] **tomshardware.** *Benchmark-Marathon: CPUs von 1993 bis 2005.* <http://www.tomshardware.de>, 2005.
- [Tra04] **Trautman, M. A., Beltman, W. M.** *Experimental setups for quiet fan design.* In Noise-Con 2004. Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2004.
- [Tra05] **Trautman, M. A., Beltman, W. M.** *Sound quality parameters of an axial fan.* In Noise-Con 2005. Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2005.
- [Tza01] **Tzanetakis, G., Essl, G., Cook, P.** *Audio Analysis using the Discrete Wavelet Transform.* In WSES Int. Conf. on Acoustics and Music: Theory and Applications (AMTA 2001). WSES Press, 2001.
- [Vet92] **Vetterli, M.** *Wavelets and Filter Banks: Theory and Design.* IEEE Transactions on signal processing, 40: 2207–2232, 1992.
- [Vog75] **Vogel, A.** *Ein gemeinsames Funktionsschema zur Beschreibung der Lautheit und der Rauigkeit.* Biol. Cybernetics, 18: 31–40, 1975.
- [Vor98] **Vormann, M., Schick, A., Meis, M., Klatte, M., Mellert, V.** *Tonhaltigkeit aus psychophysikalischer Sichtweise: Untersuchung zur Versuchspersoneninstruktion bei der Bewertung der Tonhaltigkeit.* In Fortschritte der Akustik - DAGA '98. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 456–457, 1998.
- [vR99] **von Rücker, C.** *Ein Verfahren zur Tonhöhenanalyse unter Berücksichtigung zeitlich spektraler Kontrasteffekte.* Herbert Utz Verlag, München, 1999.
- [Wal66] **Wald, A.** *chapter A.7. Sequential Analysis,* Wiley, 196–199, 1966.
- [Wal69a] **Walliser, K.** *Über die Abhängigkeiten der Tonhöhenempfindung von Sinustönen vom Schallpegel, von überlagertem drosselndem Störschall und von der Darbietungsdauer.* Acustica, 21: 211–221, 1969.
- [Wal69b] **Walliser, K.** *Über ein Funktionsschema für die Bildung der Periodentonhöhe aus dem Schallreiz.* Kybernetik, 6: 65–72, 1969.
- [Wal69c] **Walliser, K.** *Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe.* Acustica, 21: 319–329, 1969.

- [Whi05] **White, M. S.** *Acoustic emissions of digital data/video projectors - Investigating noise sources and their change during product aging.* In Noise-Con 2005. Ames, IA: Institute of Noise Control Engineering of the USA, CD-ROM, 2005.
- [Wid98] **Widmann, U., Lippold, R., Fastl, H.** *Ein Computerprogramm zur Simulation der Nachverdeckung für Anwendungen in akustischen Meßsystemen.* In Fortschritte der Akustik - DAGA '98. Verl.: Dt. Gesell. für Akustik e. V., Oldenburg, 96–97, 1998.
- [Wie90] **Wiese, D., Stoll, G.** *Bitrate Reduction of High Quality Audio Signals by Modelling the Ears Masking Thresholds.* 89th AES Convention, Preprint No. 2970, 1990.
- [Wie91] **Wiesmann, N., Fastl, H.** *Ausgeprägtheit der Tonhöhe und Frequenzunterschiedsschwellen von Bandpass-Rauschen.* In Fortschritte der Akustik - DAGA '91. Verl.: DPG- GmbH, Bad Honnef, 505–508, 1991.
- [Wie98] **Wiegrebe, L., Patterson, R. D., Demany, L., Carlyon, R. P.** *Temporal dynamics of pitch strength in regular-interval noises.* J. Acoust. Soc. Am., 104: 2307–2313, 1998.
- [Wie00] **Wiegrebe, L., Hirsch, H. S., Patterson, R. D., Fastl, H.** *Temporal dynamics of pitch strength in regular-interval noises: effects of listening region and auditory model.* J. Acoust. Soc. Am., 107: 3343–3350, 2000.
- [Win04] **Windeck, C.** *Agenda 2010/PC: Von ATX zu BTX.* CT Magazin für Computertechnik, 10: 130–137, 2004.
- [Yos78] **Yost, Y. A., Hill, R.** *Strength of the Pitches associated with rippled noise.* J. Acoust. Soc. Am., 64: 485–492, 1978.
- [Yos96] **Yost, W. A.** *Pitch of iterated rippled noise.* J. Acoust. Soc. Am., 100: 511–518, 1996.
- [Zar99] **Zar, J. H.** *Biostatistical Analysis, 4th Edition.* Prentice-Hall, Inc., Upper Saddle River, NJ, 1999.
- [Zwi55] **Zwicker, E., Heinz, W.** *Zur Häufigkeit der menschlichen Hörschwelle.* Acustica, 5: 75–80, 1955.
- [Zwi57] **Zwicker, E., Flottorp, G., Stevens, S. S.** *Critical band width in loudness summation.* J. Acoust. Soc. Am., 29: 548–557, 1957.
- [Zwi58] **Zwicker, E.** *Über psychologische und methodische Grundlagen der Lautheit.* Acustica, 1: 237–258, 1958.

- [Zwi59a] **Zwicker, E.** *Ein grafisches Verfahren zur Bestimmung der Lautstärke und der Lautheit aus dem Terzpegeldiagramm.* Frequenz, 13: 234–238, 1959.
- [Zwi59b] **Zwicker, E.** *Lautstärke und Lautheit.* In Cremer, L., editor, 3rd International congress on acoustics (ICA), 63–78, 1959.
- [Zwi60a] **Zwicker, E.** *Über die Rolle der Frequenzgruppe beim Hören.* Erg. Biol., 23: 187–203, 1960.
- [Zwi60b] **Zwicker, E.** *Zur Unterteilung des hörbaren Frequenzbereichs in Frequenzgruppen.* Acustica, 10: 185, 1960.
- [Zwi61] **Zwicker, E.** *Subdivision of the audible frequency range into critical bands.* J. Acoust. Soc. Am., 33: 248, 1961.
- [Zwi62] **Zwicker, E.** *Über ein einfaches Funktionsschema des Gehörs.* Acustica, 12: 22–28, 1962.
- [Zwi63a] **Zwicker, E.** *Über die Lautheit von ungedrosselten und gedrosselten Schallen.* Acustica, 13: 194–211, 1963.
- [Zwi63b] **Zwicker, E., Maiwald, D.** *Über das Freifeldübetragsmaß des Kopfhörers DT 48.* Acustica, 13: 181–182, 1963.
- [Zwi67] **Zwicker, E., Feldtkeller, R.** *Das Ohr als Nachrichtenempfänger.* Hirzel-Verlag, Stuttgart, 1967.
- [Zwi74] **Zwicker, E.** *Die Zeitkonstanten (Grenzdauern) des Gehörs.* Zeitschrift für Hörgeräte-Akustik, 13: 82–102, 1974.
- [Zwi75] **Zwicker, E., Herla, S.** *Über die Addition von Verdeckungseffekten.* Acustica, 34: 89–97, 1975.
- [Zwi77] **Zwicker, E.** *Procedure for calculating loudness of temporally variable sounds.* J. Acoust. Soc. Am., 62: 675–682, 1977.
- [Zwi78] **Zwicker, E.** *Erratum: Procedure for calculating loudness of temporally variable sounds.* J. Acoust. Soc. Am., 63: 283, 1978.
- [Zwi80] **Zwicker, E., Terhardt, E.** *Analytical expressions for critical-band rate and critical bandwidth as a function of frequency.* J. Acoust. Soc. Am., 68: 1523–1525, 1980.
- [Zwi83] **Zwicker, E., Yamada, Y.** *Lautstärke von Tonkomplexen in Abhängigkeit von der Bandbreite, vom Schallpegel und von zugefügtem Breitbandrauschen.* Acustica, 53: 26–30, 1983.
- [Zwi84] **Zwicker, E.** *Dependence of post-masking on masker duration and its relation to temporal effects in loudness.* J. Acoust. Soc. Am., 75: 219–223, 1984.

- [Zwi99] **Zwicker, E., Fastl, H.** *Psychoacoustics: Facts and Models*. Springer Verlag, Heidelberg, 2nd edition, 1999.