

DIPLOMARBEIT

# Modell zur effizienten Kodierung von Signallaufzeiten für die binaurale Wiedergabe virtueller Schallquellen

Efficient modeling of the time-of-arrival in binaural reproduction of virtual sound sources

Harald Ziegelwanger

---

Institut für Elektronische Musik und Akustik  
Universität für Musik und darstellende Kunst Graz



Institut für Schallforschung  
Österreichische Akademie der Wissenschaften



Beurteiler : O.Univ.Prof. Mag.art. DI Dr.techn. Robert Höldrich

Betreuer : DI Dr.rer.nat Piotr Majdak

Wien, Jänner 2012



# Kurzfassung

Außenohrübertragungsfunktionen (engl. head-related transfer functions, HRTFs) werden üblicherweise als minimalphasige Systeme mit zugehöriger breitbandiger interauraler Zeitdifferenz approximiert. Dabei wird für eine räumliche Position die Zeitdifferenz direkt aus den HRTFs für das ipsi- und kontralaterale Ohr geschätzt. Im Falle multipler Positionen bleibt die monaurale Signallaufzeit (engl. time of arrival, TOA) für eine Position und somit die Information über die monaurale zeitliche Bezug zwischen den Positionen unberücksichtigt.

Basierend auf einem geometrischen Kopfmodell wird ein Verfahren zur robusten Schätzung von TOAs aus den HRTFs einer Person vorgestellt. Dabei werden Radius, Position der Ohren und die Kopfposition relativ zur Messanordnung als Modellparameter angepasst. Das Verfahren wird auf HRTFs aus der ARI-Datenbank angewandt und resultiert in plausiblen Ergebnissen. Die perzeptive Relevanz dieser TOA wird in einem vorläufigen psychoakustischen Experiment evaluiert.

Über die Parameter des TOA-Modells kann die Positionierung einer Person in der HRTF-Messung nachträglich ermittelt und somit die tatsächlich gemessenen Richtungen bestimmt werden. Die robuste Schätzung der TOA erlaubt weitere Analysen und Manipulationen von HRTFs. Zum Beispiel können die TOAs aus den HRTFs entfernt und die HRTFs in verkürzter Form abgespeichert werden. Für die binaurale Wiedergabe multipler virtueller Schallquellen können die TOAs mittels Modellparametern rekonstruiert werden, wobei die Synchronizität zwischen den Quellen gewährleistet ist.

## Abstract

Head-related transfer functions (HRTFs) are usually approximated as minimum-phase systems with associated broadband interaural time difference. For a single spatial location the time difference is estimated directly from HRTFs for the ipsilateral and contralateral ear. For multiple spatial locations this procedure does not take into account the monaural signal propagation time (time of arrival, TOA) between a sound source and one ear. Thus the information about the monaural synchronicity between multiple positions is disregarded.

A robust method for TOA estimation based on HRTFs measured for individuals is presented, combining TOAs for multiple positions in a geometric TOA model. Head radius, ear positions and the position of the head relative to the measuring configuration were adjusted as model parameters. Applying the method to HRTFs from the ARI database results in plausible parameter values. Perceptual relevance of this TOA is evaluated in a preliminary psychoacoustic experiment.

From the estimated parameters of the TOA model the positioning of a person in the HRTF measurement and thus the real measured directions can be determined. Robust TOA estimation allows further analysis and manipulation of HRTFs. As an example, the modelled TOAs can be removed from HRTFs and stored in abbreviated form. For binaural reproduction of multiple virtual sound sources the TOAs can be reconstructed using parameters from the TOA model. Thus synchronicity between multiple sound sources is maintained.

# Danksagung

Besonders bedanken möchte ich mich sowohl bei meinem Betreuer Piotr Majdak, für seine Unterstützung und sein großes Engagement für meine Diplomarbeit, als auch bei Michael Mihocic, für die technische Betreuung des Labors bei den psychoakustischen Experimenten.

Außerdem möchte ich mich bei allen Kollegen am Institut für Schallforschung für das freundliche Arbeitsumfeld und für die zahlreichen fachlichen Diskussionen bedanken, vor allem bei Bernhard Laback, Wolfgang Kreuzer, Christian Kasess und Peter Balasz.

Weiters möchte ich mich bei allen Versuchspersonen bedanken, die sich kurz vor Weihnachten für meine Experimente Zeit genommen haben.

Mein größter Dank geht aber an meine Eltern, die mir das Studium in Graz ermöglicht haben, und an meine Freundin Julia, die mich während meiner gesamten Studienzeit unterstützt und in anstrengenden Zeiten aufgebaut hat.

Die Arbeit wurde von der Österreichischen Akademie der Wissenschaften und durch ein Förderstipendium der Universität für Musik und darstellenden Kunst Graz unterstützt.

Harald Ziegelwanger  
Wien, Jänner 2012

## **EIDESSTATTLICHE ERKLÄRUNG**

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am .....

.....

# Inhaltsverzeichnis

<b>Kurzfassung</b>	<b>i</b>
<b>Danksagung</b>	<b>iii</b>
<b>Inhaltsverzeichnis</b>	<b>v</b>
<b>Abbildungsverzeichnis</b>	<b>ix</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Ziele der Arbeit . . . . .	2
1.2 Aufbau . . . . .	3
<b>2 Stand der Forschung</b>	<b>5</b>
2.1 Räumliches Hören . . . . .	5
2.1.1 Koordinatensysteme . . . . .	5
2.1.2 Lokalisation in der Horizontalebene (Duplex-Theorie) . . . . .	8
2.1.3 Lokalisation in der Vertikalebene . . . . .	10
2.2 Head-Related Transfer Functions . . . . .	11
2.2.1 Physik . . . . .	11
2.2.2 Messung . . . . .	12
2.2.3 Nachbearbeitung . . . . .	15
2.2.4 Minimalphasige Approximation . . . . .	17
2.3 Breitbandige Signallaufzeit . . . . .	19
2.3.1 ITD-Schätzung für eine Position . . . . .	20
2.3.2 Geometrische Modelle der ITD . . . . .	23
2.4 Datenreduzierung von HRTFs . . . . .	24

<b>3</b>	<b>Time of Arrival</b>	<b>25</b>
3.1	TOA-Schätzung . . . . .	26
3.2	Räumliches TOA-Modell . . . . .	28
3.2.1	Einfaches TOA-Modell . . . . .	30
3.2.2	Erweitertes TOA-Modell . . . . .	31
3.3	Einfluss der Parameter . . . . .	32
3.4	Parameteranpassung . . . . .	34
<b>4</b>	<b>Evaluierung</b>	<b>41</b>
4.1	Numerische Evaluierung . . . . .	41
4.1.1	Binauraler Vergleich . . . . .	41
4.1.2	Modellierter Kopfradius . . . . .	45
4.1.3	Getrennte Messung verschiedener Elevationen . . . . .	46
4.1.4	Erforderliche Anzahl an gemessenen Richtungen . . . . .	47
4.2	Vorläufige experimentelle Evaluierung . . . . .	49
4.2.1	Versuchspersonen und Equipment . . . . .	51
4.2.2	Stimuli . . . . .	51
4.2.3	Experimentelle Bedingungen . . . . .	53
4.2.4	Prozedur . . . . .	54
4.2.5	Resultate . . . . .	55
4.3	Diskussion . . . . .	58
<b>5</b>	<b>Anwendungen</b>	<b>59</b>
5.1	Richtungsabhängige Fensterung . . . . .	59
5.2	Phasenanalyse . . . . .	61
5.3	Bandbreitenerweiterung . . . . .	63
5.4	Nachbearbeitung nach der HRTF-Messung . . . . .	64
5.5	Datenreduktion . . . . .	65
5.6	Aufbereitung für räumliche Transformationen . . . . .	65
<b>6</b>	<b>Zusammenfassung</b>	<b>69</b>
6.1	Ausblick . . . . .	70

<b>A</b>	<b>Detaillierte Herleitung des TOA-Modells</b>	<b>71</b>
A.1	Einfaches TOA-Modell . . . . .	71
A.2	Erweitertes TOA-Modell . . . . .	72
<b>B</b>	<b>Abbildungen und Tabellen</b>	<b>77</b>
	<b>Referenzen</b>	<b>98</b>



# Abbildungsverzeichnis

1.1	Zusammenhang zwischen dem Stand der Forschung und der TOA-Bestimmung	3
2.1	Geodesisches Koordinatensystem (linkes Bild aus <a href="#">Majdak et al., 2007</a> ) . . . . .	6
2.2	Horizontal-polares Koordinatensystem (linkes Bild aus <a href="#">Majdak, 2002</a> ) . . . . .	7
2.3	Raumebenen ( <b>links</b> : Horizontalebene, <b>Mitte</b> : Frontalebene, <b>rechts</b> : Medianebene, aus <a href="#">Majdak, 2002</a> ) . . . . .	7
2.4	Schematische Darstellung der Laufzeitdifferenz und der Abschattung (aus <a href="#">Howard und Angus, 2009</a> ) . . . . .	8
2.5	JNDs der interauralen Zeitdifferenzen ( <b>a</b> ) aus <a href="#">Zwislocki und Feldman (1956)</a> und JNDs der lateralen Winkel ( <b>b</b> ) aus <a href="#">Mills (1958)</a> . . . . .	9
2.6	Lokalisation in der Vertikalebene, ( <b>a</b> ) aus <a href="#">Majdak (2011)</a> . . . . .	11
2.7	Schalldruck vor und im Ohrkanal (aus <a href="#">Møller et al., 1995</a> ) . . . . .	12
2.8	Messaufbau im Labor ( <b>a</b> ) und Mikrofon im Ohrkanal ( <b>b</b> ) . . . . .	13
2.9	Anordnung der 1550 Messpunkte . . . . .	13
2.10	Amplitudenspektren und ETCs der berechneten HRTFs und DTFs in der Horizontalebene und Medianebene des linken Ohres einer typischen Person .	14
2.11	Auswirkung der Bandbegrenzung auf die Natürlichkeit von Musik und Sprache (aus <a href="#">Moore und Tan, 2003</a> ) . . . . .	16
2.12	Spektrogramm eines Ausschnitts des Songs "One" von Metallica (aus <a href="#">Larsen und Aarts, 2004</a> ) . . . . .	16
2.13	Perzeptive Evaluierung der Lokalisationsleistung von minimalphasig approximierten HRTFs (aus <a href="#">Kistler und Wightman, 1992</a> ) . . . . .	18
2.14	Psychometrische Funktion einer Versuchsperson ( <b>a</b> ) und JNDs der Phasenänderung eines Teiltons in einem harmonischen Klang für drei verschiedene Versuchspersonen ( <b>b</b> ), aus <a href="#">Moore und Glasberg (1989)</a> . . . . .	19
2.15	Verschiedene ITD-Schätzmethode angewandt auf HRTFs und DTFs einer Person der ARI-Datenbank . . . . .	22

2.16	Frequenzabhängige ITD gemessen an einem Kunstkopf ohne Torso ( $\bullet$ ) mit blankem Torso ( $\times$ ) und bekleidetem Torso ( $\circ$ ) (aus <a href="#">Kuhn, 1977</a> ) . . . . .	23
3.1	TOA-Schätzung $\hat{\tau}$ ( <b>a</b> ) und Kreuzkorrelationskoeffizient $\rho$ ( <b>b</b> ) für DTFs eines Ohrs . . . . .	27
3.2	TOA-Schätzung bei $0^\circ$ Elevation für DTFs des linken Ohrs . . . . .	28
3.3	TOA-Modell (3D-Ansicht) . . . . .	29
3.4	Einfaches TOA-Modell (Skizze) . . . . .	30
3.5	Erweitertes TOA-Modell (Skizze) . . . . .	31
3.6	Variation der Parameter des einfachen Modells . . . . .	33
3.7	Variation der Parameter des erweiterten Modells . . . . .	33
3.8	Erkannte Ausreißer nach Schritt 1 . . . . .	35
3.9	TOAs nach Entfernung des lateral konstanten Delays aus $\Phi_G$ ( <b>a</b> ) und Anzahl der Richtungen in den lateralen Gruppen ( <b>b</b> ) . . . . .	36
3.10	Erkannte Ausreißer nach Schritt 2 . . . . .	36
3.11	Anpassung des einfachen Modells nach Schritt 3 ( <b>grün</b> : ohne Ausreißerererkennung, <b>blau</b> : mit Ausreißerererkennung) . . . . .	37
3.12	Erkannte Ausreißer nach der Verfeinerung der Detektion mithilfe des einfachen Modells . . . . .	39
3.13	Anpassung nach Schritt 4 ( <b>grün</b> : einfaches Modell aus Schritt 3, <b>blau</b> : erweitertes Modell) . . . . .	40
3.14	Modellierte TOA und die Abweichung von der geschätzten TOA jeweils in <i>ms</i> . . . . .	40
4.1	Modellierte Kugelradien für HRTFs der ARI-Datenbank ( <b>M1</b> : einfaches TOA-Modell, <b>M2</b> : erweitertes TOA-Modell) . . . . .	42
4.2	Kugelradien für 56 Versuchspersonen ( <b>blau</b> : linkes Ohr, <b>schwarz</b> : rechtes Ohr, <b>grün/rot</b> : Mittelwert beider Ohrradien) . . . . .	42
4.3	<b>oben</b> : Binaurale Differenz der Kugelradien ( <b>grün</b> : einfaches TOA-Modell, <b>rot</b> : erweitertes Modell), <b>unten</b> : Verschiebung in y-Richtung im erweiterten Modell ( <b>blau</b> : linkes Ohr, <b>schwarz</b> : rechtes Ohr, <b>rot</b> : Mittelwert der $y_M$ -Werte beider Ohren) . . . . .	43
4.4	Auswirkung der Verschiebung des Kopfes einer Versuchsperson entlang der Transversalachse für eine schlecht positionierte Person ( <b>a</b> ) und eine gut positionierte Person ( <b>b</b> ) . . . . .	44
4.5	Verschiebung in x-Richtung im erweiterten Modell ( <b>1</b> : linkes Ohr, <b>2</b> : rechtes Ohr) . . . . .	45

4.6	Vergleich der modellierten mittleren Kugelradien und der effektiven Kopfradien (oben) und Differenz und quadratischer Mittelwert der Differenz (unten)	46
4.7	Auswirkung der zeitlich und räumlich getrennten Messung von 5er und 10er Elevationen ( <b>oben</b> : Anpassung des Modells an alle Richtungen, <b>unten</b> : separate Anpassung an die beiden Elevationsgruppen)	47
4.8	Zufällige Reduktion der Messpunkte (10 mal für 3 Versuchspersonen)	48
4.9	Erwartungshaltung des Experiments	50
4.10	Zeitdarstellung und Amplitudenspektrum der 3 Stimuli ( <b>oben</b> : breitbandiges gaußverteiltes Rauschen, <b>Mitte</b> : nullphasiger komplexer Klang, <b>unten</b> : an Phasenspektrum der Referenz-HRTF angepasster komplexer Klang)	52
4.11	Unterschiedliche Amplitudenspektren nach der Faltung eines Signals mit einer HRTF und deren minimalphasigen Approximation (HRIR-Länge : 256 Samples)	53
4.12	Anordnung der Richtungen	54
4.13	Gemittelte Ergebnisse der experimentellen Evaluierung	56
4.14	Gemittelte Ergebnisse der experimentellen Evaluierung	57
5.1	ETCs nach Entfernung der TOA und Fensterung von HRIIRs für vier Elevationen)	60
5.2	Pol-Nullstellendiagramm in der z-Ebene ( <b>links</b> : gesamte HRIR, <b>rechts</b> : enthaltene Allpasskomponenten)	61
5.3	Phasenspektren der extrahierten Allpässe (Excessphase) für das linke Ohr ( <b>links</b> : ipsilaterale Richtungen, <b>rechts</b> : kontralaterale Richtungen)	62
5.4	Phasenspektren der extrahierten Allpässe (Excessphase) für das rechte Ohr ( <b>links</b> : kontralaterale Richtungen, <b>rechts</b> : ipsilaterale Richtungen)	62
5.5	Crossfade und Bandbreitenerweiterung	64
5.6	Amplituden- und Phasenspektren nach dem alten und neuen Postprocessing	65
5.7	Räumliche Phasenspektren bei ( <b>links</b> : originale HRTFs; <b>rechts</b> : TOA-reduzierte HRTFs)	66
5.8	Räumliche Phasenspektren bei ( <b>links</b> : originale HRTFs; <b>rechts</b> : TOA-reduzierte HRTFs)	67
A.1	Einfaches TOA-Modell (Skizze)	71
A.2	Erweitertes TOA-Modell (Skizze)	73
B.1	NH11, linkes Ohr	78
B.2	NH18, linkes Ohr	79

B.3	NH <sub>4</sub> , linkes Ohr . . . . .	80
B.4	Ergebnisse der experimentellen Evaluierung (VP 1) . . . . .	81
B.5	Ergebnisse der experimentellen Evaluierung (VP 1) . . . . .	82
B.6	Ergebnisse der experimentellen Evaluierung (VP 2) . . . . .	83
B.7	Ergebnisse der experimentellen Evaluierung (VP 2) . . . . .	84
B.8	Ergebnisse der experimentellen Evaluierung (VP 3) . . . . .	85
B.9	Ergebnisse der experimentellen Evaluierung (VP 3) . . . . .	86

# Kapitel 1

## Einleitung

Individuell gemessene Außenohrübertragungsfunktionen (engl. head-related transfer functions, HRTFs) ermöglichen die Positionierung von virtuellen Schallquellen über Kopfhörer (Binaural-Audio) oder Lautsprecher (Transaural-Audio). Gegenüber verallgemeinerten oder an Kunstköpfen gemessenen HRTFs bieten sie den Vorteil einer bestmöglichen Lokalisation und Externalisation von Hörereignissen für den Nutzer. Die Kombination von "gut" oder "natürlich" klingenden und zu genauer Lokalisation führenden HRTFs stellt eine Herausforderung in der virtuellen Akustik dar. Zum einen muss die Datenmenge der HRTFs reduziert werden, um Rechenleistung bei Echtzeitnutzung zu sparen, andererseits müssen alle Lokalisationsmerkmale erhalten bleiben und ein nahezu realistischer und vor allem natürlicher Klang ermöglicht werden.

Das wichtigste Lokalisationsmerkmal für die laterale Position einer Schallquelle ist hier die zeitliche Information zwischen den Ohren, welche interaurale Zeitdifferenz (engl. interaural time difference, ITD) genannt wird. Es wird angenommen, dass HRTFs unterhalb von 10 kHz durch Minimalphasensysteme und zugehöriger ITD perzeptiv ausreichend beschrieben werden können. Die ITD beschreibt Laufzeitunterschiede zwischen beiden Ohren. Für die binaurale Wiedergabe multipler Schallquellen müssen jedoch Laufzeitunterschiede zwischen verschiedenen Positionen an einem Ohr mitberücksichtigt werden, da auch die Signallaufzeit (engl. time of arrival, TOA) zwischen Schallquelle und einem Ohr eine Richtungsabhängigkeit aufweist. Die TOA hängt zum einen von Schalllaufwegen ab, zum anderen von Beugungseffekten an vom Kopf abgeschatteten Richtungen, welche zusätzlich eine Frequenzabhängigkeit aufweisen. Die räumliche Anordnung des Messaufbaus und die Genauigkeit der Positionierung einer Person bei der HRTF-Messung spielen ebenfalls eine Rolle.

Die korrekte und genaue Schätzung der TOA ist der erste Schritt und die Grundlage für die weitere genaue Untersuchung von HRTF-Phasenspektren und stellt ein nicht einfach zu lösendes Problem dar. Die TOA ist im Gegensatz zur interauralen Zeitdifferenz weder

psychoakustisch noch messtechnisch, mittels Signalverarbeitung aus HRTF-Datensätzen extrahiert, gut untersucht. Ein zu zufriedenstellenden Ergebnissen führender automatisierter Algorithmus zur TOA-Schätzung aus individuell gemessenen HRTFs wurde bis heute nicht publiziert.

## 1.1 Ziele der Arbeit

Das primäre Ziel dieser Arbeit ist die Entwicklung eines automatisierten Verfahrens zur robusten Berechnung der TOAs aus einem 3D-Satz individuell gemessener HRTFs.

Es wird ein geometrisches dreidimensionales TOA-Modell hergeleitet. Aus einem HRTF-Satz werden die TOAs für einzelne Positionen (Richtungen) geschätzt und an diese die Parameter des Modells für alle Richtungen angepasst. Ziel des Modells ist die Robustheit gegenüber Fehlschätzungen und Ungenauigkeiten im Messaufbau. Dabei sollen drei Aspekte bedacht werden:

- Bestimmung der breitbandigen TOA (ohne Einfluss der frequenzabhängigen Gruppenlaufzeit, verursacht von Allpasskomponenten)
- Perzeptive Relevanz dieser TOA (Übereinstimmung der geschätzten und wahrgenommenen TOA in Bezug auf die Referenz-HRTF)
- anthropometrische Relevanz des Modells (Übereinstimmung der Parameter des Modells mit physiologischen Abmessungen)

Das TOA-Modell kann mit ITD-Modellen aus der Literatur verglichen werden. Abbildung 1.1 zeigt den Zusammenhang diverser Themen, welche in Kapitel 2 genau beschrieben werden, in Bezug auf die TOA. Es werden die verschiedenen Studien zur ITD in Bezug auf die TOA diskutiert und zur TOA-Bestimmung kombiniert, um dadurch Probleme der einzelnen Ansätze zu lösen.

Das Modell wird angewandt, um die TOAs verschiedener Versuchspersonen zu schätzen. Neben der Aufrechterhaltung des zeitlichen Bezuges einzelner Schallquellen an einem Ohr bietet die Kenntnis der TOA den Vorteil der Darstellung und Untersuchung der verzögerungsfreien Phasenspektren von HRTFs. Die Ergebnisse der TOA-Schätzung und die Auswirkung der minimalphasigen Approximation werden sowohl numerisch als auch durch psychoakustische Experimente evaluiert. Es wird vor allem die Frage geklärt, ob die modellierte TOA der perzeptiven TOA entspricht. Über die TOA ist es aber auch erstmals möglich die Auswirkung von Allpasskomponenten an einem Ohr separat psychoakustisch zu untersuchen. Das Experiment soll Aufschluss darüber geben, ob es Signale oder Bedingungen gibt, bei

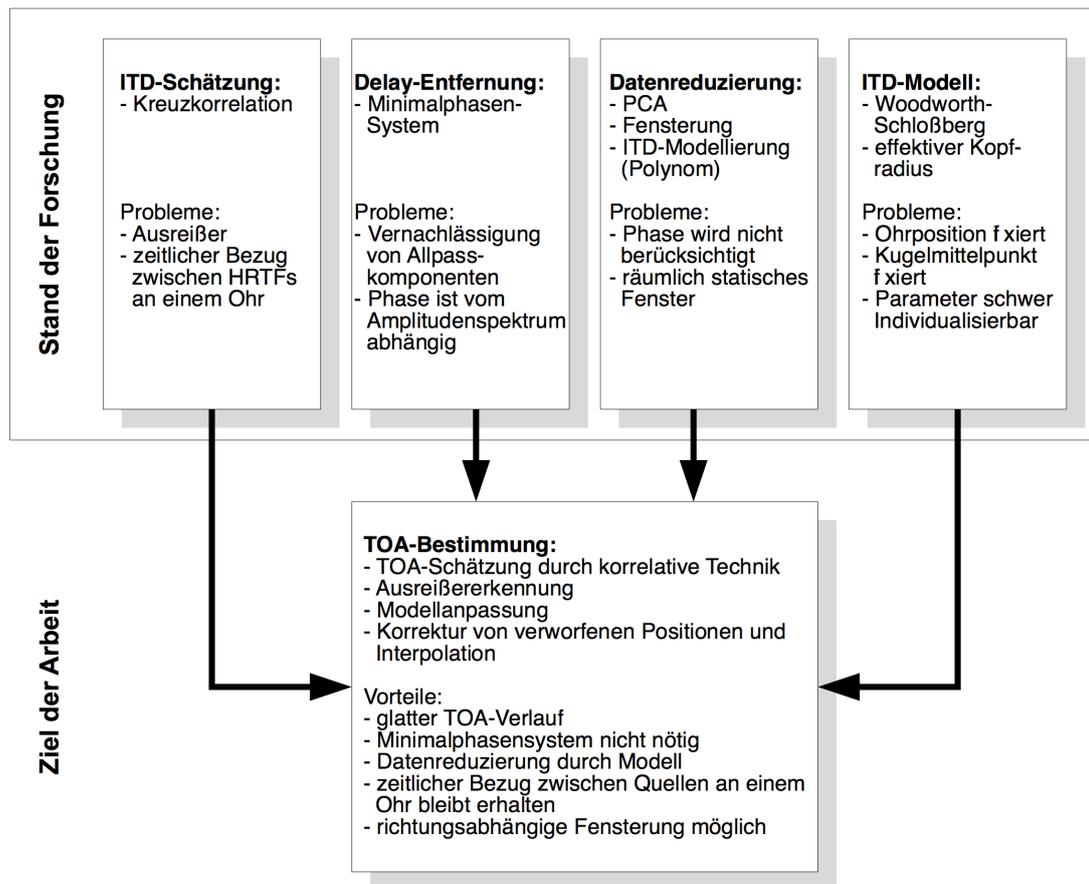


Abbildung 1.1: Zusammenhang zwischen dem Stand der Forschung und der TOA-Bestimmung

denen die bei Minimalphasenfiltern vernachlässigten Allpasskomponenten hörbar sind und trotz korrekter TOA-Modellierung Klangverfärbungen verursachen.

## 1.2 Aufbau

In Kapitel 2 werden die für diese Arbeit relevanten Grundlagen der räumlichen Wahrnehmung von Schallereignissen und die Einführung der wichtigsten Begriffe zusammengefasst und der Stand der Forschung aufgearbeitet. Dabei wird auf die Literatur über die interaurale Zeitdifferenz (engl. interaural time difference, ITD), sowohl aus der Sicht der Psychoakustik als auch der Signalverarbeitung und Messtechnik, zurückgegriffen.

Der Hauptteil der Arbeit liegt in Kapitel 3, welches das Verfahren zur TOA-Schätzung und Modellierung beschreibt. Das Modell wird in Kapitel 4 sowohl numerisch als auch experimentell evaluiert.

In Kapitel 5 werden einige Anwendungen der TOA-Schätzung auf HRTF-Datensätze vorgestellt.

Die Ergebnisse der Arbeit werden in Kapitel 6 zusammengefasst, welches auch einen Ausblick auf zukünftige Untersuchungen enthält.

# Kapitel 2

## Stand der Forschung

### 2.1 Räumliches Hören

#### 2.1.1 Koordinatensysteme

Bevor auf die Theorie des räumlichen Hörens genauer eingegangen wird, werden hier die in dieser Arbeit verwendeten Koordinatensysteme und zugehörigen Begriffe definiert. Es gilt hier zwei Bezugssysteme zu unterscheiden.

Für die Theorie des räumlichen Hörens ist es sinnvoll den Kopf einer Person als Bezugssystem zu definieren, also die Position von Schallquellen relativ zum Beobachter anzugeben. Der Ursprung der Koordinatensysteme liegt hier im Mittelpunkt zwischen dem linken und dem rechten Ohr einer Versuchsperson, des Beobachters. Deshalb wird die durch die Ohren verlaufende Achse auch interaurale Achse bezeichnet.

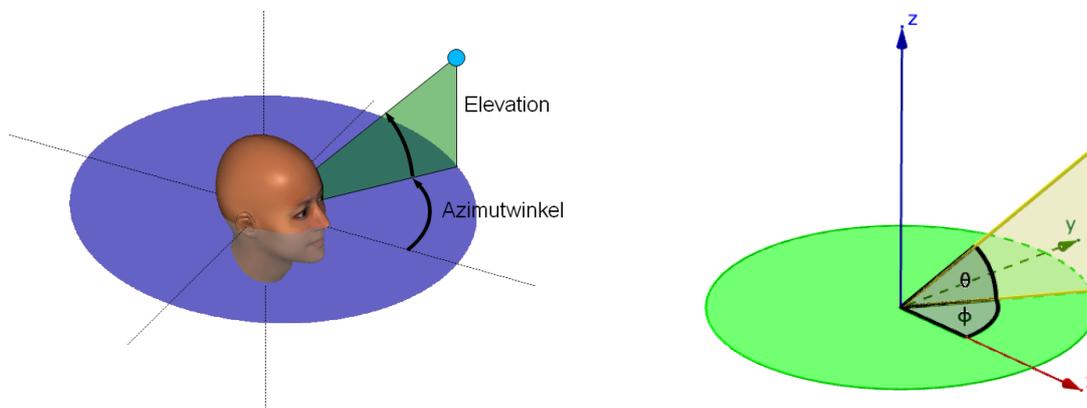
Für die Betrachtung von individuell gemessenen HRTFs macht es aber Sinn den Mittelpunkt der Messanordnung als räumlich statischen Bezugspunkt zu wählen (siehe Kapitel 2.2.2), da sich die Ausrichtung von Versuchspersonen interindividuell leicht unterscheiden kann (siehe Kapitel 4) und somit hier die auf die Person bezogenen Bezugssysteme bezüglich des Bezugssystems der Messanordnung interindividuell unterschiedlich geringfügig verschoben und gedreht sein können. Ist die Ausrichtung einer Person ideal, sind die Bezugssysteme identisch.

Da in dieser Arbeit nur das entfernungsunabhängige räumliche Hören, also nur die Richtung einer Schallquelle betrachtet wird, werden sphärische Koordinaten ohne Radius verwendet. Für die Berechnungen in Kapitel 3 werden auch die zugehörigen kartesischen Koordinatensysteme vorgestellt. Da die Unterschiede der Bezugssysteme gering sind, werden die Bezeichnungen für Winkel, Achsen und Raumebenen in beiden Bezugssystemen gleichbedeutend verwendet.

## Geodesisches Koordinatensystem

Das geodesische Koordinatensystem ist aus der Sicht des Beobachters das natürlichste Koordinatensystem. Es wird durch eine horizontale Auslenkung  $\phi$  (Azimut, von  $0^\circ$  bis  $360^\circ$ ) und den Höhenwinkel  $\theta$  (Elevation, von  $-90^\circ$  bis  $90^\circ$ ) beschrieben. Der Zenit befindet sich über dem Beobachter. Aufgrund der Rechtsdrehung ergibt sich ein zugehöriges kartesisches Koordinatensystem. Die Achsen entsprechen anatomischen Achsen, deren Koordinatenursprung allerdings im Körpermittelpunkt festgelegt ist, jedoch ohne Probleme zum Kopfmittelpunkt verschoben werden kann. Die x-Achse wird dorsoventrale Achse oder Sagittalachse genannt und verläuft ventral (vom Körpermittelpunkt zur Bauchdecke), die y-Achse mediolaterale Achse oder Transversalachse (entspricht der interauralen Achse) und verläuft sinistral (aus der Perspektive des Menschen vom Körpermittelpunkt nach links) und die z-Achse vertikale Achse oder Longitudinalachse und verläuft kranial (vom Körpermittelpunkt zum Kopf). Die beiden Koordinatensysteme lassen sich folgendermaßen ineinander umrechnen:

$$\begin{aligned}
 x &= r \cos(\phi) \cos(\theta) & \phi &= \arctan\left(\frac{y}{x}\right) \\
 y &= r \sin(\phi) \cos(\theta) & \theta &= \sin^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \\
 z &= r \sin(\theta) & r &= \sqrt{x^2 + y^2 + z^2}
 \end{aligned} \quad (2.1)$$

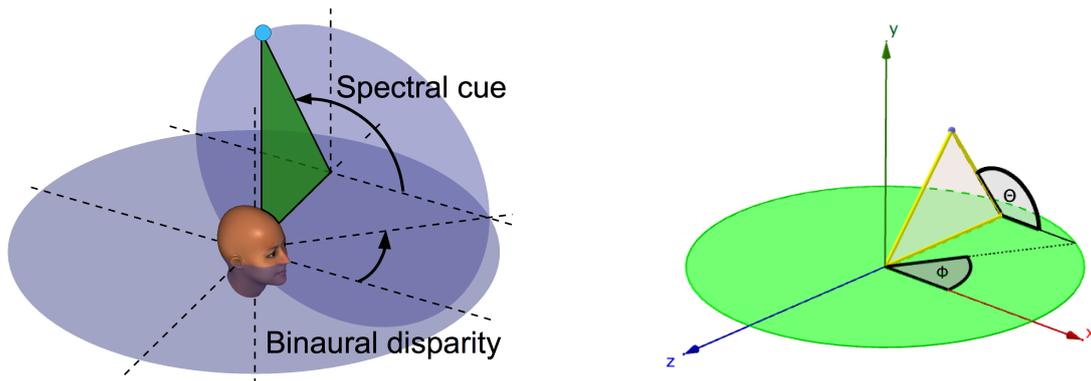


**Abbildung 2.1:** Geodesisches Koordinatensystem (linkes Bild aus Majdak et al., 2007)

## Horizontal-polares Koordinatensystem

Im Folgenden beschriebene Lokalisationsmerkmale (engl. cues) können aber durch das horizontal-polare Koordinatensystem geeigneter beschrieben werden. Die Pole dieses Systems liegen auf der interauralen Achse, wobei sich der Zenit rechts vom Beobachter befindet. Die horizontale Auslenkung  $\Phi$  (Lateralwinkel  $-90^\circ$  bis  $90^\circ$ ) wird zur Übereinstimmung mit den

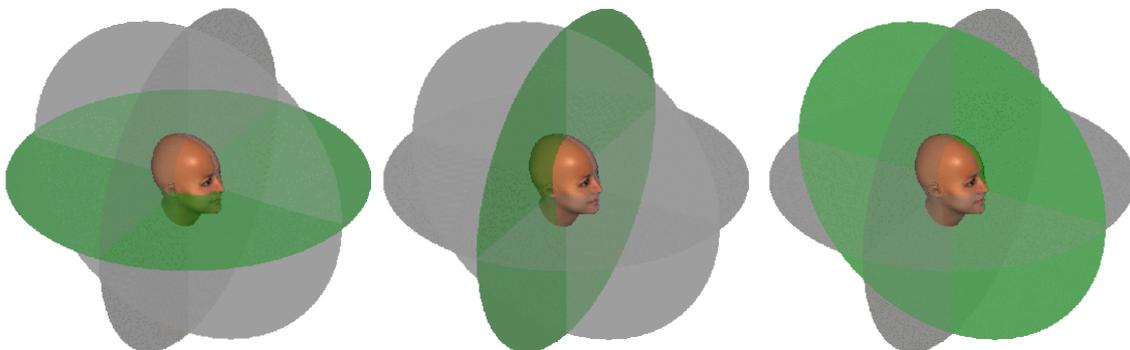
geodesischen Koordinaten, aber abweichend von der Literatur, aus der Sicht des Beobachters nach links positiv gezählt. Dieser Winkel beschreibt die Veränderung der binauralen Laufzeit- und Pegelunterschiede (siehe 2.1.2). Der Polarwinkel  $\Theta$  beschreibt die Richtung auf Sagittalebene (siehe Kapitel 2.1.3). Hier sind vor allem monaurale Merkmale für die Lokalisation verantwortlich. Die Koordinaten können analog zum geodesischen Koordinatensystem in kartesische Koordinaten umgerechnet werden. Die in dieser Arbeit verwendeten kartesischen Koordinaten beziehen sich aber immer auf das geodesische Koordinatensystem.



**Abbildung 2.2:** Horizontal-polares Koordinatensystem (linkes Bild aus Majdak, 2002)

### Raumebenen

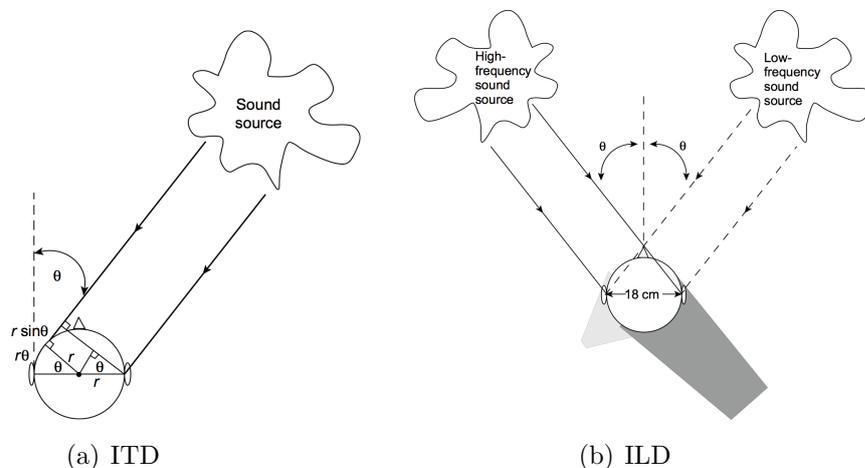
Der Beobachter wird durch drei normal aufeinander stehenden Ebenen geschnitten. Für die Lokalisation auf der Horizontalebene (Abb. 2.3, links) sind vor allem binaurale Cues verantwortlich. Aus Richtungen auf der Medianebene (Abb. 2.3, rechts) sind monaurale Cues (spektrale Färbungen) für die Höhenwahrnehmung und die Vorne- Hintenunterscheidbarkeit verantwortlich. Auf der Frontalebene (Abb. 2.3, Mitte) sind alle Cues von Bedeutung. Ebenen parallel zur Medianebene werden Sagittalebene genannt.



**Abbildung 2.3:** Raumebenen (**links:** Horizontalebene, **Mitte:** Frontalebene, **rechts:** Medianebene, aus Majdak, 2002)

## 2.1.2 Lokalisation in der Horizontalebene (Duplex-Theorie)

1907 wurde die Duplex-Theorie formuliert, welche besagt, dass für die Richtungswahrnehmung interaurale Zeitdifferenzen (engl. interaural time differences, ITDs) und interaurale Pegeldifferenzen (engl. interaural level differences, ILDs) verantwortlich sind (John William Strutt, 1877). Und zwar nicht gleichberechtigt, sondern die ITD bei tiefen Frequenzen bis ca. 1.5 kHz und die ILD darüber. Er stellte diesen Umstand bei Lokalisationsexperimenten von Sinustönen fest. Sinustöne, deren Wellenlänge größer als der doppelte Kopfdurchmesser ist, werden durch den Kopf am kontralateralen Ohr nicht abgeschattet und es treten nur Laufzeitunterschiede zwischen den beiden Ohren auf, die über die Phasendifferenz ausgewertet werden. Liegt die Wellenlänge im Bereich des Kopfdurchmessers, wird diese Phasendifferenz mehrdeutig (Phasenambiguität) und es werden bis ca. 1.5 kHz Pegel- und Laufzeitunterschiede zur Lateralisation herangezogen. Darüber hinaus dient nur mehr die Pegeldifferenz, verursacht durch die Abschattung am kontralateralen Ohr, als Lateralisationsmerkmal. In Abbildung 2.4 ist eine einfache Skizze der ITD und ILD dargestellt. Da ITD und ILD als Unterschied zwischen den Ohren wahrgenommen werden, nennt man sie binaurale Lokalisationsmerkmale.



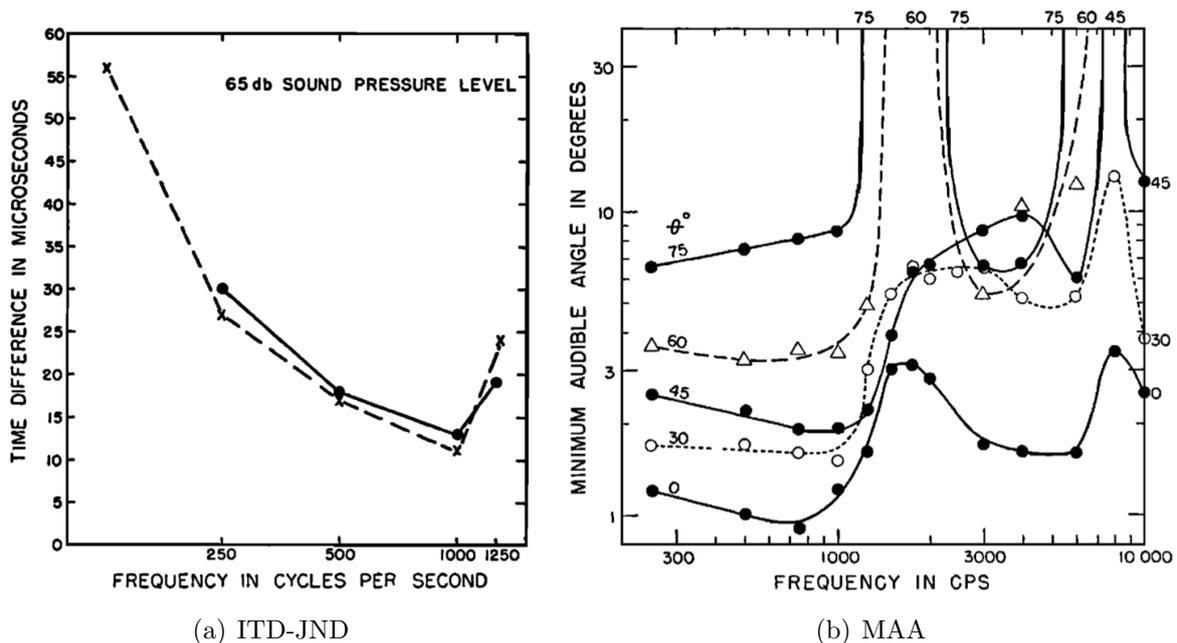
**Abbildung 2.4:** Schematische Darstellung der Laufzeitdifferenz und der Abschattung (aus Howard und Angus, 2009)

Ein Jahrhundert nach der Entdeckung von John William Strutt wurde die Duplex-Theorie in psychoakustischen Experimenten für Rauschsignale genauer untersucht (Macpherson und Middlebrooks, 2002). Mithilfe eines dimensionslosen Fehlers (Verhältnis von perzeptiv resultierender gemessener ITD/ILD-Abweichung zu eingestellter ITD/ILD-Abweichung) konnten ITD und ILD objektiv miteinander verglichen werden. So ist dieses Verhältnis für ITD bei Frequenzen unterhalb von 2 kHz nahe bei 1 und darüber nahe bei 0. Heute ist

bekannt, dass neben der Phasenambiguität auch die Refraktärzeit von Neuronen dafür verantwortlich ist. Bei der ILD sieht das Ergebnis wie erwartet genau umgekehrt aus.

Für breitbandiges Rauschen überwiegt im Experiment die ITD als Lokalisationsmerkmal. Dieses Ergebnis bestätigt die Dominanz der ITD gegenüber anderen konkurrierenden Merkmalen (Wightman und Kistler, 1992). So werden breitbandige Schallquellen auch bei extrem kontrahierenden Einstellungen von ITD, ILD und spektralen Färbungen eher an der Position verursacht durch die ITD wahrgenommen.

Die Schwellen des kleinsten wahrnehmbaren Unterschieds (engl. just noticeable difference, JND) für die ITD bei Sinustönen wurden bereits Mitte des letzten Jahrhunderts gemessen (Zwislocki und Feldman, 1956). Die höchste ITD-Sensitivität konnte bei 1 kHz gefunden werden, welche bei ca.  $13\mu s$  liegt (siehe Abb. 2.5a), das entspricht bei einer Abtastrate von 48 kHz ungefähr einem halben Abtastintervall. Betrachtet man die JNDs für den wahrnehmbaren Winkel (engl. minimum audible angle, MAA; Mills, 1958), korrelieren diese für die Richtung  $\phi = 0^\circ$ ,  $\theta = 0^\circ$  stark mit den ITD-JNDs (siehe Abb. 2.5b). Je weiter eine Schallquelle von der Medianebene ausgelenkt wird, desto größer ist der MAA und auch die ITD-JND.



**Abbildung 2.5:** JNDs der interauralen Zeitdifferenzen (a) aus Zwislocki und Feldman (1956) und JNDs der lateralen Winkel (b) aus Mills (1958)

Auch wenn Envelope- und Ongoing-ITD als Lokalisationsmerkmale in psychoakustischen Experimenten nachgewiesen werden konnten (Henning, 1974), spielen diese in der Natur nur eine untergeordnete vernachlässigbare Rolle. In natürlichen Situationen wie in Räumen

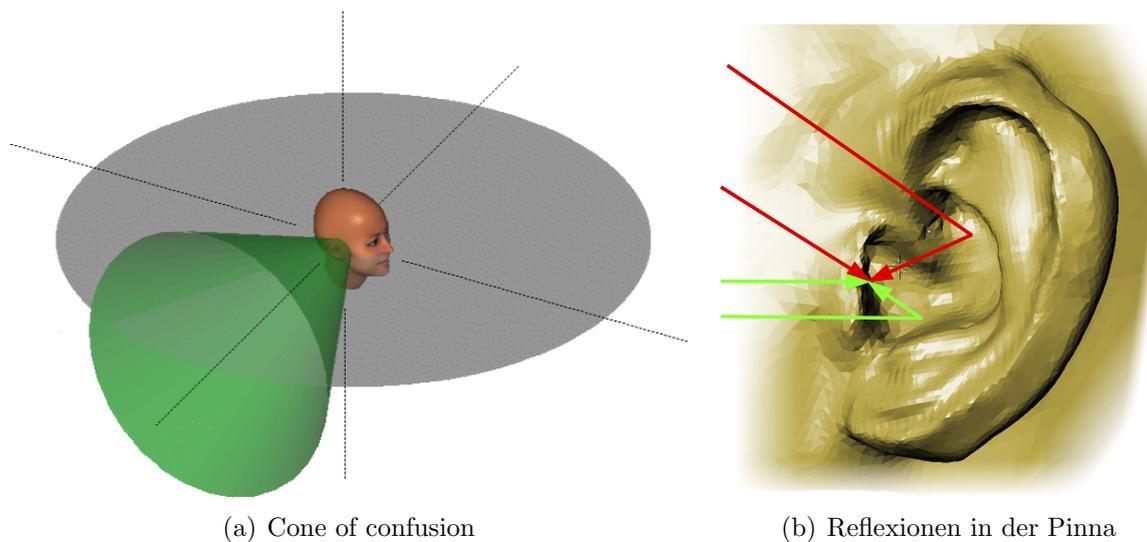
mit Nachhall ist die Kohärenz (Ähnlichkeit zweier Signale) zwischen den Signalen am linken und rechten Ohr nicht ausreichend, weshalb Versuchspersonen dieses Merkmal nicht ausnutzen (Constan und Hartmann, 2001). Eine weitere Begründung liegt in der Dominanz der ILD und spektralen Färbungen gegenüber kontrahierender Envelope-ITDs und in der geringen Modulationstiefe oder hohen Modulationsfrequenz natürlich auftretender Envelope-ITDs (Middlebrooks und Green, 1990).

Die ITD ist eine frequenzabhängige Größe. Die analytisch berechnete ITD zwischen zwei Punkten auf einer schallharten Kugel ( $r = 8.75\text{cm}$ ) weist für Frequenzen  $f < 625\text{Hz}$  einen um den Faktor 1.5 höheren Wert, verursacht durch Beugungseffekte am kontralateralen Ohr, auf (Kuhn, 1977). Diese Tatsache konnte auch mittels Messungen am Kunstkopf nachgewiesen werden. Allerdings ergeben sich bei Kombination des Kunstkopfes mit einem bekleideten oder blanken Torso leicht unterschiedliche Ergebnisse. Die Impedanz des Kopfes und Haare haben ebenfalls einen Einfluss auf diese Frequenzabhängigkeit (Treeby et al., 2007b,a). Ist eine mit der Position der restlichen Lokalisationsmerkmale übereinstimmende ITD vorhanden, hat die Feinstruktur des Phasenspektrums im Allgemeinen keinen Einfluss auf die Lokalisation von Schallquellen (Kulkarni et al., 1999). Es ist hier anzumerken, dass mit dieser Studie nicht gezeigt wird, dass Unterschiede im Phasenspektrum nicht hörbar sind. Die Ergebnisse dieser psychoakustischen Experimente ergeben einen hörbaren Unterschied für Tiefpass-gefilterte Stimuli. Eine mögliche Erklärung liegt in der Dispersion oder einer schlechten Schätzung der ITD. Der Anstieg der ITD bei tiefen Frequenzen ist aber für Übertragungsfunktionen, berechnet auf einer schallharten Kugel, nicht wahrnehmbar (Constan und Hartmann, 2003). Obwohl das Phasenspektrum nicht zur Lokalisation beiträgt, könnte es jedoch eine Auswirkung auf die Klangfarbe haben (siehe Kapitel 2.2.4).

### 2.1.3 Lokalisation in der Vertikalebene

Richtungen auf einem um die interaurale Achse rotationssymmetrischen Kegel weisen nahezu dieselben ITDs und ILDs auf und können daher alleine durch diese Merkmale nicht unterschieden werden. Daher wird dieser auch "Cone of confusion" genannt. Dieser ist in Abbildung 2.6a für einen lateralen Winkel dargestellt. Die Mehrdeutigkeit in der Horizontalebene kann zu Vorne-Hintenverwechslungen führen.

Um Schallquellen auf der Medianebene oder auf Sagittalebene unterscheiden zu können, wertet das auditorische System spektrale Informationen aus. Diese Merkmale entstehen durch den Einfluss von Torso, Kopf und Pinna. Bereits Mitte des letzten Jahrhunderts wurde die Duplex-Theorie um die reflektierende Wirkung der Ohrmuschel (Pinna) erweitert (Batteau, 1967), welche der einfallenden Schallwelle richtungsabhängig feine Echos hinzufügt und im Weiteren zu spektralen Färbungen führt (siehe Abb. 2.6b). So entstehen Überhöhungen (engl.



**Abbildung 2.6:** Lokalisation in der Vertikalebene, (a) aus Majdak (2011)

peaks) und Absenkungen (engl. notches), die vom Gehör ausgewertet werden können (Moore et al., 1989) und hauptsächlich oberhalb von 5 kHz auftreten. Da diese Klangfärbungen schon an einem Ohr wahrgenommen werden können, spricht man auch von monauralen Merkmalen für die Lokalisation (Wightman und Kistler, 1997). Die Messung dieser spektralen Informationen führt zu den HRTFs.

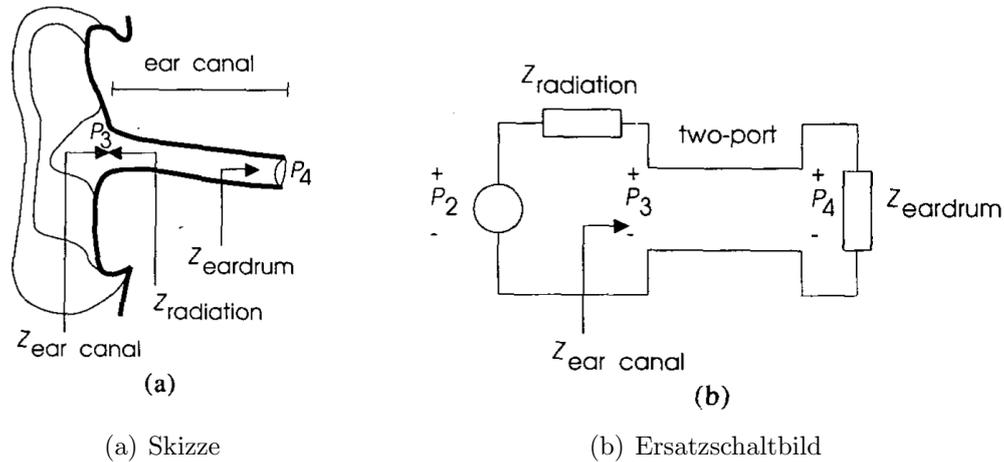
## 2.2 Head-Related Transfer Functions

HRTFs beschreiben die richtungsabhängige Filterwirkung von Pinna, Kopf und Torso zwischen einer Schallquelle und jeweils einem der beiden Gehörgänge (Wightman und Kistler, 1989). Sie sind die Fouriertransformierten der Außenohrimpulsantworten (engl. head-related impulse responses, HRIR). Diese werden üblicherweise in reflexionsarmen Räumen im Fernfeld gemessen. Dafür ist ein Messabstand von mindestens einem Meter (Brungart und Rabinowitz, 1999) erforderlich, ab welchem die Krümmung der auf das Messobjekt (menschliche Versuchsperson oder Kunstkopf) eintreffenden Schallwelle vernachlässigt werden kann und keine evaneszenten Wellen mehr auftreten. Auf Studien zu Nahfeld-HRTFs, welche die entfernungsabhängigen Eigenschaften der HRTF untersuchen, wird nicht näher eingegangen (siehe Brungart und Rabinowitz, 1999).

### 2.2.1 Physik

HRTFs bilden das Verhältnis von Schalldruck im Freifeld und am Eingang des Gehörgangs (Henrik und Møller, 1992). Abbildung 2.7 zeigt eine schematische Skizze der Pinna und

des Gehörgangs (links) und das dazu analoge Ersatzschaltbild (rechts).  $P_4$  repräsentiert den Schalldruck am Trommelfell und  $P_3$  den Schalldruck am Eingang des Gehörgangs. Über das Thèvenin-Theorem kann eine Quelle in zwei Teile aufgeteilt werden: zum einen in den "Leerlauf"-Schalldruck  $P_2$  und die Ersatzimpedanz  $Z_{radiation}$ .



**Abbildung 2.7:** Schalldruck vor und im Ohrkanal (aus Møller et al., 1995)

$P_2$  existiert unter realen Hörbedingungen physikalisch nicht, da hierfür die Schallschnelle am Eingang des Gehörgangs Null betragen müsste. Jedoch kann durch Messung des Schalldrucks am blockierten Gehörgang und einer Referenzmessung im Freifeld ( $P_1$ ) der gesamte richtungsabhängige Anteil der HRTFs beschrieben werden:

$$HRTF = \frac{P_2}{P_1}(\phi, \theta) = \frac{\text{Schalldruck am Eingang des geblockten Gehörgangs}}{\text{Schalldruck im Zentrum des Kopfes}}. \quad (2.2)$$

Der Schalldruck am offenen Eingang des Gehörgangs (richtungsunabhängiger Anteil) kann über die Ersatzimpedanz der Quelle  $Z_{radiation}$  und der Impedanz des Gehörgangs  $Z_{earcanal}$  berechnet:

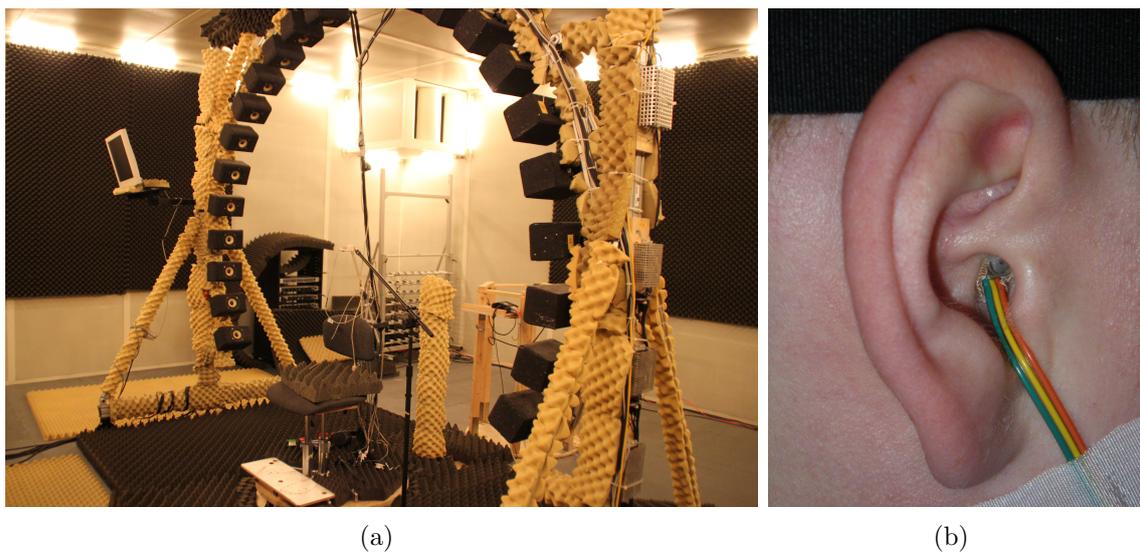
$$\frac{P_3}{P_2} = \frac{Z_{earcanal}}{Z_{earcanal} + Z_{radiation}}, \quad (2.3)$$

oder aber auch über ein Miniatur-Probenmikrofon gemessen werden. Der Schalldruck direkt am Trommelfell ( $P_4$ ) kann nur über ein solches Probenmikrofon gemessen werden.

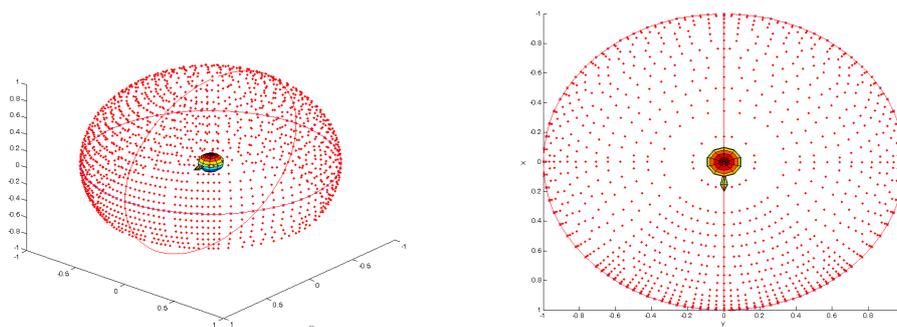
## 2.2.2 Messung

Die HRTF-Messung am Schallforschungsinstitut in Wien erfolgt in einem semianechoischen Raum durch platzieren von blockierenden Miniaturmikrofonen in den beiden Gehörgängen

einer Versuchsperson (siehe Abb. 2.8b), deren Kopf sich im Mittelpunkt eines senkrecht kreisförmig angeordneten Lautsprecherarrays befindet (Majdak et al., 2010). Die Versuchsperson wird dabei horizontal gedreht, um eine hohe räumliche Auflösung in der Horizontalebene zu erreichen. Abbildung 2.8a zeigt den Messaufbau. Das horizontale räumliche Abtastintervall liegt zwischen  $\pm 40^\circ$  bei  $2.5^\circ$  und von  $40^\circ$  bis  $90^\circ$  bei  $5^\circ$ . Elevationen werden von  $-30^\circ$  bis  $80^\circ$  in  $5^\circ$  Schritten gemessen. Die horizontale Messpunktdichte nimmt zu hohen Elevationen hin ab (für eine konstante räumliche Bogenlänge, siehe Abb. 2.9).

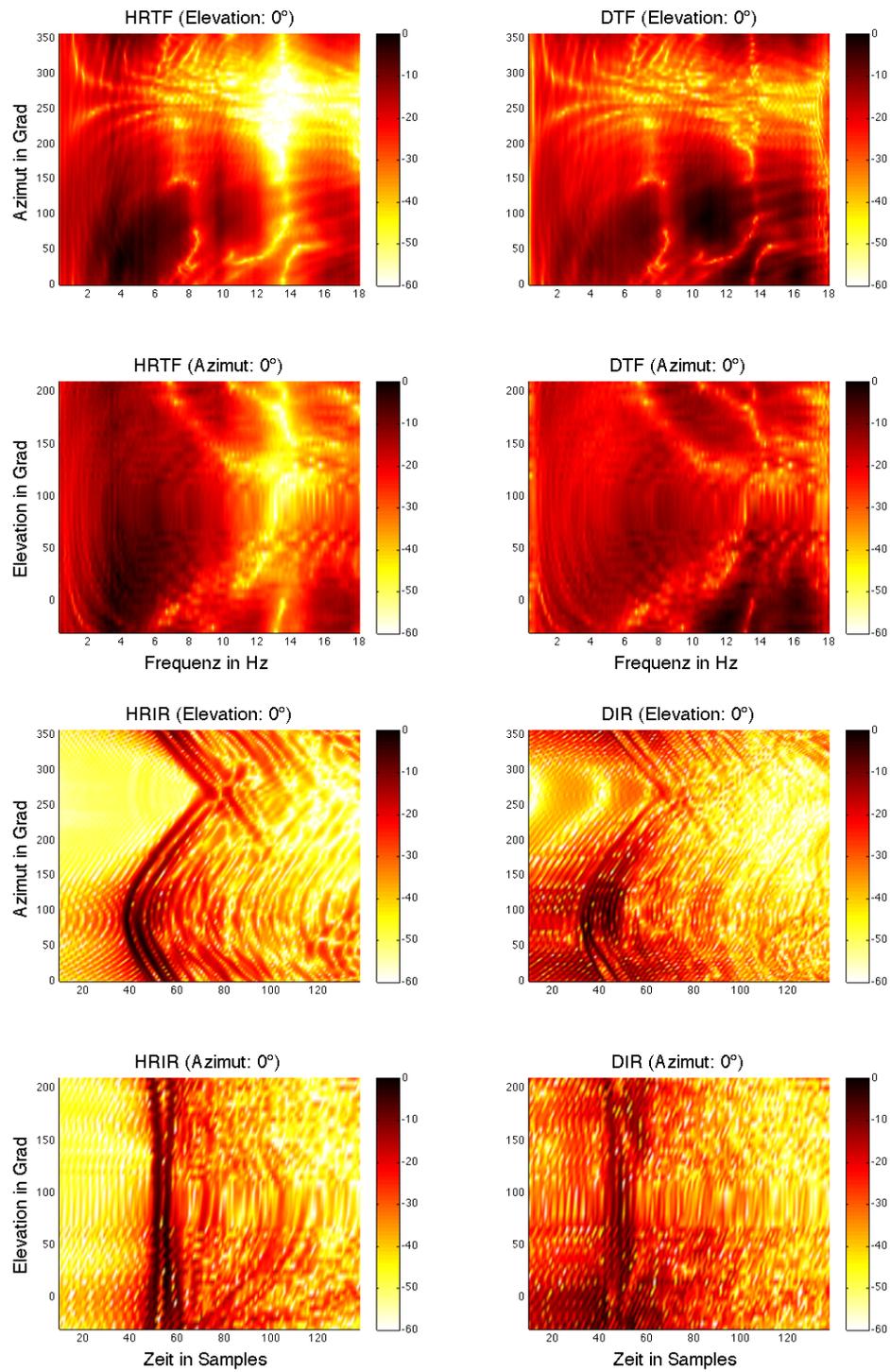


**Abbildung 2.8:** Messaufbau im Labor (a) und Mikrophon im Ohrkanal (b)



**Abbildung 2.9:** Anordnung der 1550 Messpunkte

Die Messung von Elevationen bei  $-30^\circ$ ,  $-20^\circ$ ,  $-10^\circ$ , ... und Elevationen bei  $-25^\circ$ ,  $-15^\circ$ ,  $-5^\circ$ , ... erfolgt räumlich und zeitlich getrennt, da somit unter optimaler Ausnutzung des Platzes auf dem Array Lautsprecher mit ausreichender Größe platziert werden können. Gemessen wird mithilfe der Multiple-Exponential-Sweep Methode (Majdak et al., 2007), welche eine Messung von 1550 Richtungen in nur 20 Minuten ermöglicht.



**Abbildung 2.10:** Amplitudenspektren und ETCs der berechneten HRTFs und DTFs in der Horizontalebene und Medianebene des linken Ohres einer typischen Person

### 2.2.3 Nachbearbeitung

Die gemessenen Rohdaten  $H_{roh}$  beinhalten neben den HRTFs  $H$  sowohl die Frequenzgänge von Mikrofon  $M$  und Lautsprecher  $L$  als auch den Einfluss des Messraums  $C$  (Majdak, 2010).

$$H_{roh} = C_{hrtf} \cdot L \cdot M \cdot H \quad (2.4)$$

Zur Berechnung der freifeldentzerrten HRTF wird eine Referenzmessung durchgeführt:

$$R_{roh} = C_{ref} \cdot L \cdot M = C_{ref} \cdot E \quad (2.5)$$

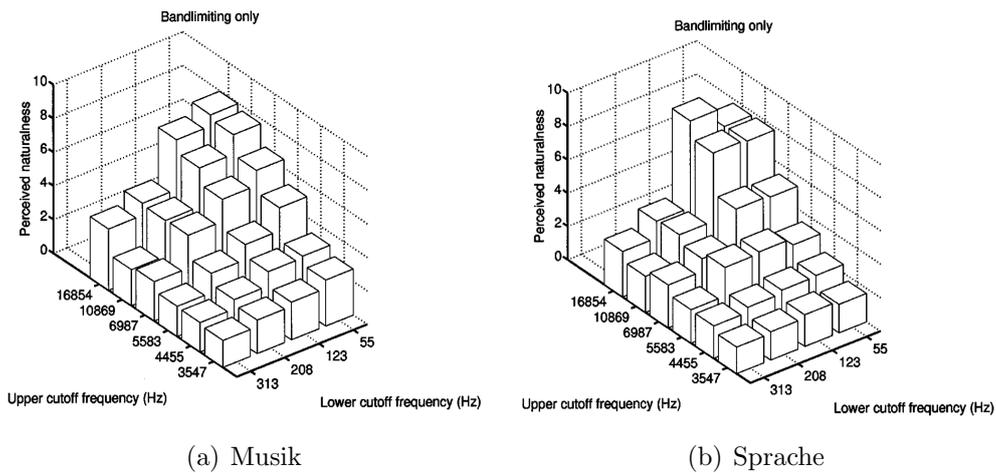
Aus der Referenzmessung wird der Raumeinfluss durch kepstrale Glättung mithilfe eines Lowpass-Lifters bei 3.8 ms entfernt. Die Phase des Equipments wird minimalphasig angenommen  $R_{post}$ .

Durch Division von  $H_{roh}$  durch  $R_{post}$  und derselben kepstralen Glättung im Frequenzbereich von 300 Hz bis 18 kHz werden die gewünschten Filter (HRTF) berechnet. Zur Berechnung der Richtungsübertragungsfunktionen (engl. directional transfer functions, DTFs) (Middlebrooks, 1999) werden die entzerrten HRTFs kepstral gemittelt und durch Berechnung des minimalphasigen Phasenspektrums über die Hilberttransformation die beiden richtungsunabhängigen Übertragungsfunktionen (engl. common transfer functions, CTFs) gebildet. Für die DTFs werden die entzerrten HRTFs durch die CTFs im Frequenzbereich von 300 Hz bis 18 kHz dividiert und noch einmal kepstral geglättet (Abb. 2.10 zeigt die Ergebnisse für eine typische Person). Abbildung 2.10 zeigt die Spektren der berechneten HRTFs und DTFs in der Horizontal- und der Medianebene. Weiters sind die Energie-Zeit-Kurven (engl. energy time curves, ETCs) von HRIRs  $h[n]$  derselben Ebenen dargestellt:

$$ETC[n] = 10 \cdot \log\{h[n]^2\}, \quad (2.6)$$

Ein Problem der Messung stellt die Bandbegrenzung des Equipments (Lautsprecher) dar. Das Messsignal erreicht erst ab einer Frequenz von 300 Hz die volle Energie. Diese Hochpass-Charakteristik spiegelt sich auch in den gemessenen HRTFs wider.

Der Frequenzbereich unterhalb dieser Grenzfrequenz ist aber von entscheidender Bedeutung für die Natürlichkeit von Audiosignalen, wie vor allem Musik (Moore und Tan, 2003). Bei einer Bandbegrenzung ab 200 Hz bei tiefen Frequenzen oder ab 7000 Hz zu hohen Frequenzen hin wird der Unterschied in der Natürlichkeit von Versuchspersonen deutlich wahrgenommen und der Höreindruck verschlimmert sich weiter bei weiterer Bandbreiteneinschränkung (siehe Abb. 2.11).

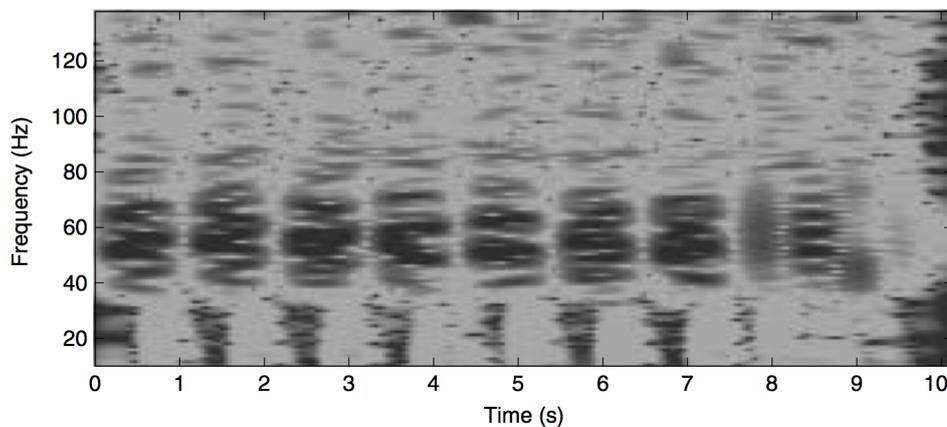


(a) Musik

(b) Sprache

**Abbildung 2.11:** Auswirkung der Bandbegrenzung auf die Natürlichkeit von Musik und Sprache (aus Moore und Tan, 2003)

Bei bandbegrenzten Übertragungs- oder Wiedergabesystemen muss die fehlende Information hinzugefügt oder psychoakustisch vorgetäuscht werden (Larsen und Aarts, 2004). Zum Beispiel reicht das Spektrum des Songs "One" von der Band "Metallica" bis zu 40 Hz hinunter (siehe Abb. 2.12).



**Abbildung 2.12:** Spektrogramm eines Ausschnitts des Songs "One" von Metallica (aus Larsen und Aarts, 2004)

Dieses Problem wirkt sich bei der reziproken Messmethode noch stärker aus (Zotkin et al., 2006). Bei diesem Messprinzip werden Miniaturlautsprecher im Gehörgang platziert und HRTFs mit einem Kugelmikrofonarray gemessen. Die Größe der Lautsprecher führt zu einer noch höheren Grenzfrequenz. Das Problem wird durch numerische Berechnung der HRTFs unterhalb der Grenzfrequenz kompensiert.

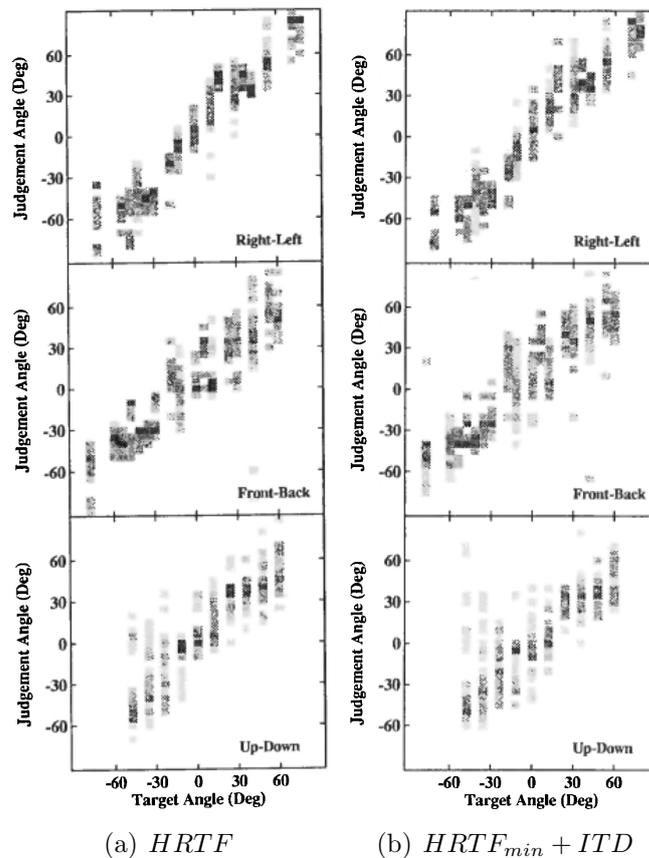
### 2.2.4 Minimalphasige Approximation

Jede Übertragungsfunktion kann in einen Minimalphasenfilter und einen Allpass zerlegt werden. Die Phasendifferenz zwischen originalem Phasenspektrum  $\varphi$  und minimalphasigem Anteil  $\varphi_{min}$  wird Excessphase  $\varphi_{exc}$  genannt, welche eine frequenzunabhängige Verzögerung  $\tau$  und eine Summe von Allpässen zweiter Ordnung  $\varphi_{ap}$  beinhaltet:

$$H = |H|e^{j\varphi} = |H|e^{j\varphi_{min}} e^{j\varphi_{exc}} = |H|e^{j\varphi_{min}} e^{-j\tau\omega} e^{j\varphi_{ap}}. \quad (2.7)$$

Auch HRTFs können in diese Phasenteile zerlegt werden (Mehrgardt und Mellert, 1977). Aus der visuellen Inspektion der Phasenspektren schlussfolgerten sie, dass HRTFs die Eigenschaft eines Minimalphasensystems bis zu einer Frequenz von 10 kHz besitzen und die Allpasskomponenten vernachlässigbar sind. Diese Aussage steht aber im Widerspruch mit der Entstehung der Filterwirkung. HRTFs entstehen durch Reflexionen in der Pinna (Verzögerung und Überlagerung), durch Reflexionen am Torso und an kontralateralen Richtungen zusätzlich durch Beugung (mehrere Ausbreitungswege) der Schallwellen um den Kopf. Eine perzeptive Überprüfung dieser Eigenschaft in Bezug auf die Lokalisation folgte 15 Jahre später als Kontrollexperiment für die Hauptkomponentenzerlegung von HRTFs (Kistler und Wightman, 1992). Dabei führen minimalphasig approximierten HRTFs zur selben Lokalisationsleistung wie die Originale (siehe Abb. 2.13). Links sind die Ergebnisse der Lokalisationsexperimente mit den empirisch gemessenen HRTFs und rechts die Ergebnisse mit den minimalphasig approximierten HRTFs dargestellt. Je dunkler ein Bereich in den Diagrammen eingefärbt ist, desto öfter wurde bei eingestelltem Winkel (engl. target angle) der wahrgenommene Winkel (engl. judgement angle) als Richtung von den Versuchspersonen identifiziert. Als ITD wurden für die psychoakustischen Experimente die Maxima der kreuzkorrelierten HRIRs beider Ohren verwendet (vgl. Kapitel 2.3).

Für die Lokalisation macht es auch keinen Unterschied ob eine HRTF ein minimalphasiges, linearphasiges oder sogar ein in der Zeit umgedrehtes Phasenspektrum aufweist (Kulkarni et al., 1999), solange das Amplitudenspektrum und die breitbandige ITD korrekt abgebildet wird. Jedoch sind Unterschiede bei Richtungen nahe der interauralen Achse und Tiefpass-gefiltertem Stimulus binaural klar hörbar. Eine mögliche Begründung wäre die Erhöhung der ITD bei tiefen Frequenzen. Der Anstieg der ITD zu tiefen Frequenzen hin unter 800 Hz hat jedoch perzeptiv keine Auswirkung (Constan und Hartmann, 2003). Eine andere Möglichkeit wäre die Hörbarkeit von monauralen oder binauralen Phasenunterschieden, verursacht durch die minimalphasige Approximation (Vernachlässigung der Allpasskomponenten). Bei komplexen Tönen ist bereits ein Phasenunterschied von wenigen Grad eines Teiltons detektierbar (Moore und Glasberg, 1989). In Abbildung 2.14 (links) ist die ermittelte psychometrische Funktion für einen nullphasigen Tonkomplex mit einer Grundfrequenz

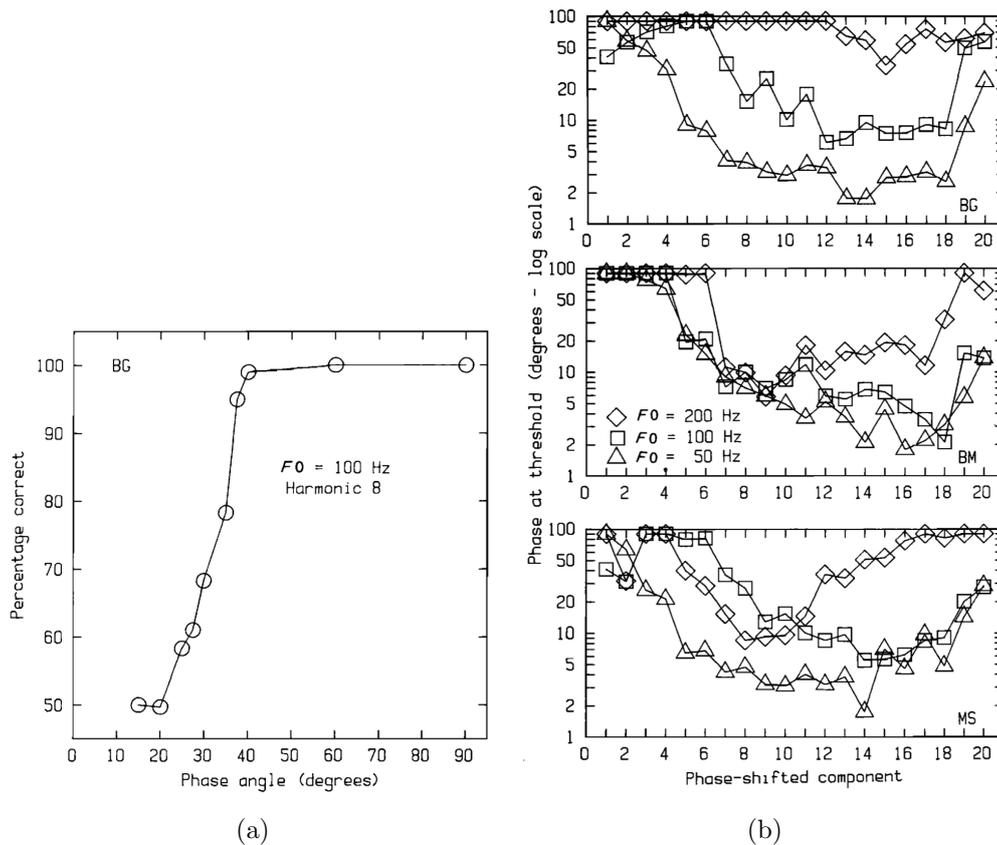


**Abbildung 2.13:** Perzeptive Evaluierung der Lokalisationsleistung von minimalphasig approximierten HRTFs (aus Kistler und Wightman, 1992)

von 100 Hz und Phasenvariation des achten Teiltons dargestellt. In Abbildung 2.14 (rechts) sind die JNDs dreier Versuchspersonen für Tonkomplexe mit Grundfrequenzen von 50, 100 und 200 Hz als Funktion der variierten Teiltöne dargestellt. Für Grundfrequenz 50 Hz sind alle Versuchspersonen ab einer Phasenänderung von 5 Grad zwischen achtem und achtzehntem Teilton sensitiv.

Die ITD ist als Verzögerung zweier HRIRs bei der minimalphasigen Approximation nicht klar definiert (Katz et al., 2005). Die breitbandige ITD unterscheidet sich von der perzeptiven ITD, welche auch die Verzögerung von Allpasskomponenten im wahrnehmbaren Frequenzbereich der ITD inkludiert. Es ist ebenfalls nicht klar, ob diese frequenzabhängigen Verzögerungsanteile der Allpasskomponenten durch eine breitbandige ITD, also eine Art mittlere ITD bezüglich der Frequenz, perzeptiv korrekt approximiert werden können.

In den Studien zur Hörbarkeit von Allpässen wurde die Auswirkung deren Vernachlässigung nicht getrennt von der Genauigkeit der Laufzeitmessung untersucht. Zum Beispiel sind Allpässe vernachlässigbar, wenn die ITD nur auf  $30\mu\text{s}$  genau geschätzt wird (Plogsties et al., 2000). Dieses Ergebnis gibt jedoch nur Auskunft über die Relevanz der genauen



**Abbildung 2.14:** Psychometrische Funktion einer Versuchsperson (a) und JNDs der Phasenänderung eines Teiltons in einem harmonischen Klang für drei verschiedene Versuchspersonen (b), aus Moore und Glasberg (1989)

ITD-Messung. Weiters wurde geschlossen, dass Allpasskomponenten hauptsächlich an Frequenzen die einen Notch aufweisen entstehen und deshalb durch das zugehörige Amplitudenspektrum unterdrückt werden. Das trifft aber auf Allpässe, die als Folge der Disperion der ITD bei tiefen Frequenzen entstehen, nicht zu. HRTF-Allpässe sind separat betrachtet, also ohne HRTF-Amplitudenspektrum, deutlich wahrnehmbar. Das ist kein verwunderliches Ergebnis, da Allpässe zweiter Ordnung bei ihrer Mittenfrequenz im Gegensatz zu einem reinen Einheitsimpuls ein starkes Ein- und Ausschwingverhalten aufweisen können (Minnaar et al., 1999).

## 2.3 Breitbandige Signallaufzeit

Um HRTFs minimalphasig approximieren zu können besteht also grundsätzlich die Herausforderung darin, die Signallaufzeiten (engl. time of arrival, TOA)  $\tau_l$  und  $\tau_r$  möglichst genau aus  $\varphi_{exc}$  zu schätzen.

$$e^{j\varphi_{exc}} = e^{-j\tau\omega} e^{j\varphi_{ap}} \quad (2.8)$$

Es gibt wenig Literatur zur expliziten TOA-Schätzung aus HRTFs. Es gibt aber viele ITD-Studien, die, um die ITD schätzen zu können, implizit die TOAs geschätzt haben. Es folgt eine kurze Beschreibung der für die TOA relevanten ITD-Studien. Die Schätzung der breitbandigen Signallaufzeit stellt, anders als man meinen würde, ein nicht triviales Problem dar.

Theoretisch lässt sich die TOA durch Mittelung der durch Ableitung von  $\varphi_{ges}$  nach der Frequenz  $\omega$  ermittelten Gruppenlaufzeit berechnen. Die praktikable Schätzung im Fourier-Raum erweist sich jedoch als schwierig. Um die frequenzabhängige Gruppenlaufzeit zu berechnen, wird ein stetiges Phasenspektrum benötigt. Diese Bedingung trifft allerdings aufgrund der  $2\pi$ -Periodizität per Definition der Phase nicht zu. Die Phase muss daher vor der Ableitung mittels eines Algorithmus "abgewickelt" werden. Die Aufrechterhaltung der breitbandigen Signallaufzeit im Spektrum ist hier aber nicht gewährleistet.

Bei der Schätzung im Zeitbereich besteht die Schwierigkeit im Finden eines allgemeinen direkten mathematischen Zusammenhangs einer frequenzunabhängigen Verzögerung zwischen Frequenz- und Zeitdomäne, weil das Phasenspektrum über eine nichtlineare Funktion, den Arkustangens, berechnet wird.

### 2.3.1 ITD-Schätzung für eine Position

Als ITD wird in dieser Arbeit die Differenz der TOAs an den beiden Ohren unter Vernachlässigung der Allpässe definiert:

$$ITD = \tau_l - \tau_r. \quad (2.9)$$

Es gibt zwei Ansätze zur ITD-Schätzung. Entweder wird die interaurale Kreuzkorrelationsfunktion (engl. interaural cross-correlation, IACC) berechnet und daraus direkt die frequenzunabhängige Verzögerung als ITD geschätzt. Bei dieser direkten ITD-Schätzung wird meistens der Maximalwert der IACC-Funktion im Zeitbereich als ITD definiert (Wightman und Kistler, 1992; Kulkarni et al., 1999).

$$ITD = \arg \max_k |\mathcal{F}^{-1}\{H_l(\omega)H_r^*(\omega)\}[k]| \quad (2.10)$$

$$= \arg \max_k \left| \sum_{n=0}^{N-1} h_l[n]h_r[n-k] \right|, \quad (2.11)$$

wobei  $H_l$  und  $H_r$  HRTFs und  $h_l$  und  $h_r$  HRIR des linken und rechten Ohres,  $\omega$  die Frequenz,  $n$  die diskrete Zeit in Abtastintervallen und  $k$  eine Laufvariable repräsentieren. Beim zweiten Ansatz wird versucht die TOA der linken und rechten HRIR separat zu bestimmen und über deren Differenz die ITD zu berechnen. Dabei erfolgt die Schätzung der TOA auf unterschiedlichen Wegen. Die folgenden vier gebräuchlichsten Methoden können sowohl zur TOA-Schätzung einer HRIR oder zur ITD-Schätzung aus der IACC-Funktion zweier HRIRs genutzt werden.

### Schätzung im Zeitbereich

Als TOA kann das Überschreiten einer bestimmten Pegelschwelle der HRIR definiert werden (Møller et al., 1995).

$$\tau = \arg \min_n \{ |h[n]| - \arg \max_n |h[n] \cdot 10^{threshold/20}| \} \quad (2.12)$$

Die Schwelle (engl. threshold) wird hier relativ zum Maximalwert der HRIR angegeben (z.B.  $-20dB$ ). Diese kann auch beim Maximalwert selbst gewählt werden, wodurch sich diese Methode auf eine Peak-Detection reduziert. Diese Methode ist nicht robust gegenüber Impulsantworten mit mehreren Peaks oder schwach ausgeprägtem Peak, wie bei kontralateralen Richtungen nahe der interauralen Achse.

Die Signallaufzeit lässt sich auch als zeitliche Position des Maximalwerts der Kreuzkorrelationsfunktion einer HRTF und deren minimalphasigen Version berechnen (Wightman und Kistler, 2005; Nam et al., 2011). Durch die Multiplikation der beiden gleichen Amplitudenspektren werden Bereiche mit geringer Energie (Notches) zusätzlich abgeschwächt und fallen somit bei der Schätzung im Zeitbereich weniger stark ins Gewicht.

$$\tau = \arg \max_k |\mathcal{F}^{-1}\{H(\omega)H_{min}^*(\omega)\}[k]| \quad (2.13)$$

### Schätzung im Frequenzbereich

Jot (1995) verwenden zur ITD-Schätzung eine lineare Regression der HRTF-Excessphase  $\varphi_{exc}$ . Als nutzbaren Frequenzbereich geben sie 1 bis 5 kHz an. Sie gehen von der Minimalphasigkeit von HRTFs, wie in Kapitel 2.2.4 beschrieben, aus.

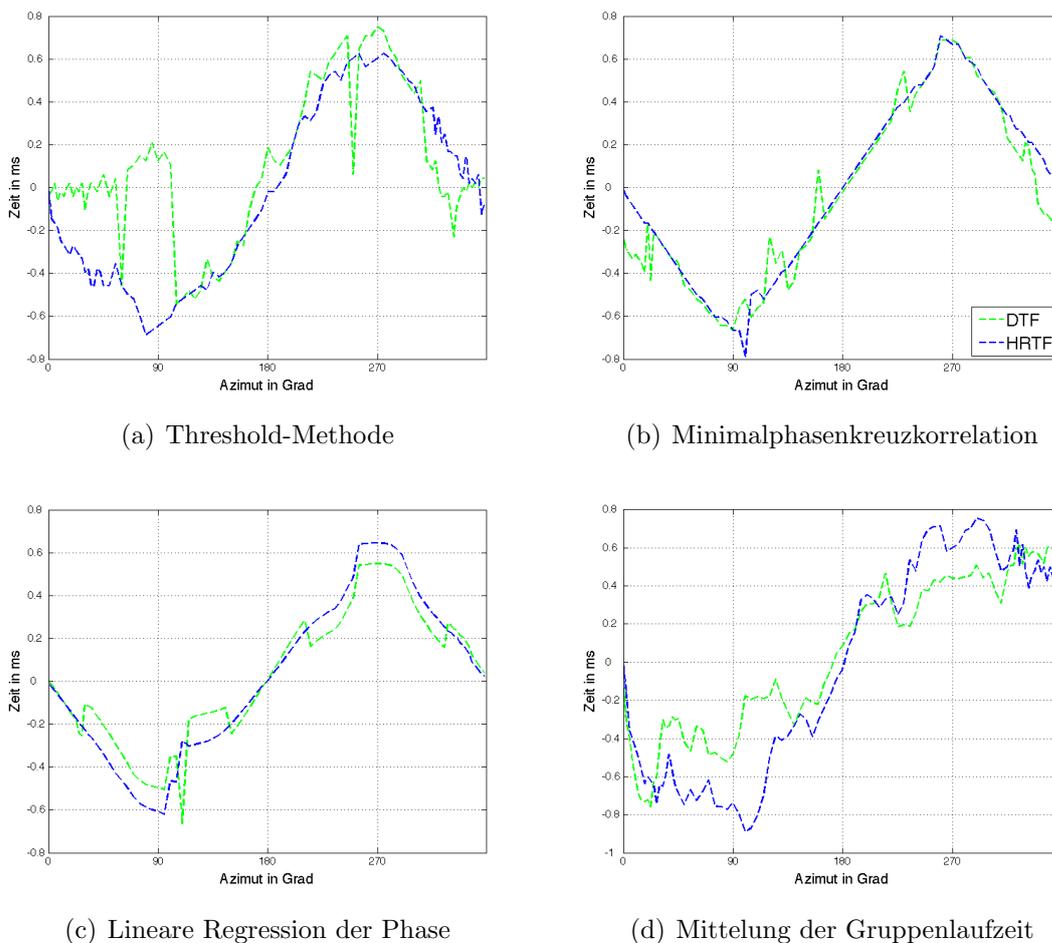
$$\tau = \arg \min_k \sqrt{\sum_{\omega} [\varphi(\omega) + \omega k]^2} \quad (2.14)$$

Bei der Gruppenlaufzeitschätzung wird entweder die aus dem Phasenspektrum durch Ableitung berechnete Gruppenlaufzeit über einen bestimmten Frequenzbereich gemittelt (Jot,

1995) oder die Gruppenlaufzeit bei 0 Hz herangezogen (Minnaar et al., 2000). Jedoch sind HRTFs in den meisten Fällen bandbegrenzt (siehe 2.2) und in diesem Frequenzbereich die Gruppenlaufzeit nicht auswertbar.

$$\tau = \frac{1}{\omega_2 - \omega_1} \sum_{\omega=\omega_1}^{\omega_2} -\frac{d\varphi(\omega)}{d\omega} \quad (2.15)$$

In Abbildung 2.15 sind die verschiedenen Schätzmethoden beispielhaft anhand der HRTFs und DTFs einer Versuchsperson dargestellt. Aus visueller Betrachtung der ermittelten ITDs scheint die minimalphasige Kreuzkorrelation die besten Ergebnisse zu liefern, das heißt die ITD weist hier den glattesten Verlauf auf. Dennoch sind Ausreißer sichtbar, welche sich in Anwendungen virtueller Akustik negativ auswirken könnten.



**Abbildung 2.15:** Verschiedene ITD-Schätzmethoden angewandt auf HRTFs und DTFs einer Person der ARI-Datenbank

### 2.3.2 Geometrische Modelle der ITD

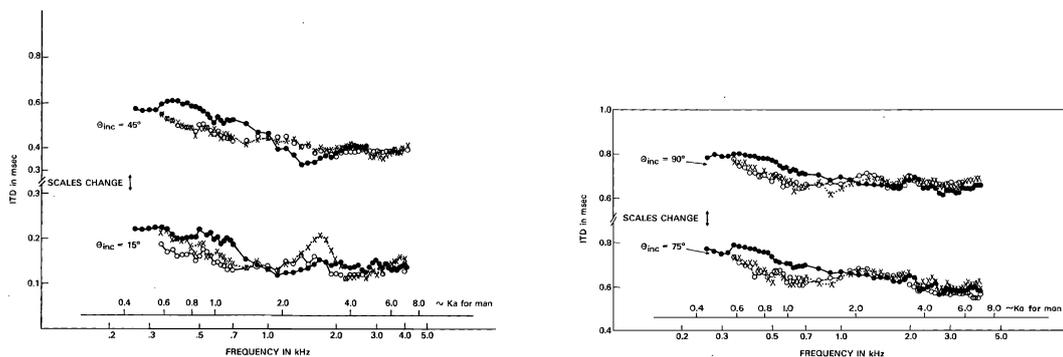
Ein TOA-Modell ist nicht bekannt. Allerdings sind einige ITD-Modelle aus der Literatur verfügbar. Jedoch ist hier wichtig anzumerken, dass die ITD eine Symmetrie bezüglich der Medianebene aufweist, welche auf die TOA nicht übertragbar ist. Eine einfache Formel zur Modellierung der ITD wurde bereits Mitte des 20. Jahrhunderts von [Woodworth und Schlossberg \(1962\)](#) basierend auf einem Ray-Tracing-Modell auf einer schallharten Kugel definiert. Sie bezieht nur einen dem Kopf angenäherten Radius  $r$  und die Schallgeschwindigkeit  $c$  mit ein:

$$ITD = \frac{r}{c} [\sin(\phi) + \phi]. \quad (2.16)$$

[Savioja \(1999\)](#) erweiterten diese Formel zur Berechnung der ITD bei unterschiedlichen Elevationen:

$$ITD = \frac{r}{c} [\sin(\phi) + \phi] \cos(\theta). \quad (2.17)$$

Die frequenzabhängige ITD, analytisch berechnet auf einer schallharten Kugel, ist bei Frequenzen unterhalb von ca. 625 Hz um den Faktor 1.5 größer als bei höheren Frequenzen ([Kuhn, 1977](#)). Diese Tatsache stimmt auch gut mit Messungen am Kunstkopf ohne Torso überein (siehe Abb. 2.16, ●). Wird allerdings der Torso mitberücksichtigt, ist dieser Effekt nicht so deutlich ausgeprägt, vor allem beim bekleideten Torso (siehe Abb. 2.16, × und ○). Es stellt sich die Frage, wie dieser ITD-Verlauf für echte Versuchspersonen mit Haaren und nicht schallharter Oberfläche aussieht. Die Impedanz einer Kugel hat nämlich einen Effekt auf die ITD ([Treeby et al., 2007a](#)) und ([Treeby et al., 2007b](#)).



**Abbildung 2.16:** Frequenzabhängige ITD gemessen an einem Kunstkopf ohne Torso (●) mit blankem Torso (×) und bekleidetem Torso (○) (aus [Kuhn, 1977](#))

Da die Form des menschlichen Kopfes nicht einer Kugel entspricht, wurde auch versucht ein ITD-Modell für einen Ellipsoiden zu entwickeln ([Duda et al., 1999](#)). Da es aber keine

analytische Lösung für den kürzesten Weg zwischen zwei Punkten auf einem Ellipsoiden gibt, wurde hier nur ein Verfahren mit numerischen Zwischenberechnungen zur ITD-Berechnung gefunden. Auch eine Approximation von HRTFs als Kombination von zwei Kugeln für Kopf und Torso wurde durchgeführt (Algazi et al., 2002). Mit dieser Studie konnte der Einfluss von Schulterreflexionen auf HRIRs geklärt werden, aber sie warf unter anderem die Frage der Relevanz der Ohrpositionen für die HRTF-Modellierung bezüglich des modellierten Kopfmittelpunkts auf.

Aus allen HRTF-Informationen kann die Kombination aus Laufzeitschätzung und Laufzeitmodellierung am meisten Auskunft über die Beziehung der HRTF-Messdaten zu den physiologischen Abmessungen (anthropometrische Daten) einer Person geben. Aus Formel 2.16 ist ersichtlich, dass sich die ITD proportional zum Radius verhält. Wird die Woodworth-Schlossberg-Formel an ITD-Schätzung angepasst, kann über lineare Regression von anthropometrischen Daten und Kugelradius aus dem ITD-Modell eine Formel zur Berechnung eines effektiven Kopfradius  $r_{eff}$  ermittelt werden (Algazi et al., 2001):

$$r_{eff} = \frac{0.51b + 0.019h + 0.18t}{2} + 3.2cm, \quad (2.18)$$

wobei  $b$  die Kopfbreite,  $h$  die Kopfhöhe und  $t$  die Kopftiefe darstellt. Auffallend ist der geringe Einfluss der Kopfhöhe. Die Konstante von  $3.2cm$  kann ebenfalls physiologisch nicht erklärt werden.

## 2.4 Datenreduzierung von HRTFs

Es gibt verschiedene Ansätze zur Reduzierung der Datenmenge von HRTFs. Die erste Möglichkeit besteht in der Hauptkomponentenanalyse der HRTF-Amplitudenspektren (Kistler und Wightman, 1992). Allerdings wurde bei dieser Methode das Phasenspektrum minimalphasig angenommen und die ITD geschätzt. Zweitens kann die Länge der HRIRs durch Glättung im Frequenzbereich verkürzt werden (Kulkarni und Colburn, 1998). Eine weitere Möglichkeit liegt in der Modellierung der ITD (Wang et al., 2009). Allerdings wurde auch bei diesen Lösungen von minimalphasigen HRTFs ausgegangen. Die einfachste Form der Datenreduzierung besteht in der Reduzierung der gemessenen Richtungen. Jedoch muss hierfür zwischen gemessenen HRTFs interpoliert werden können, ein nicht einfach zu lösendes Problem (Begault, 2000; Matsumoto, 2004).

# Kapitel 3

## Time of Arrival

Die interaurale Zeitdifferenz ist das wichtigste binaurale Merkmal für die laterale Auslenkung einer Schallquelle, deshalb werden Studien über die zeitliche Information von HRTFs auf die ITD beschränkt. Sie ist allerdings eine perzeptive Größe und physikalisch messtechnisch nicht unmittelbar greifbar. Die ITD ist als Verzögerung in der interauralen Phasendifferenz eines HRTF-Paars enthalten. Sie wird aus der Differenz der beiden TOAs eines HRIR-Paars berechnet. HRIRs beinhalten aber nicht nur die zeitliche Information zwischen den Ohren (ITD), sondern auch die zeitliche Ausrichtung aller HRIRs zueinander, weshalb die Reduzierung der zeitlichen Information auf die ITD bei mehreren Schallquellen, ungeachtet der Bedeutung der restlichen Phaseninformation für die Lokalisation oder für die Natürlichkeit, schon einen Informationsverlust verursacht.

Auch für die Signalverarbeitung einzelner HRTFs bringt die ITD keinen Nutzen. Mit ihr kann nur nach der Entfernung der Gruppenlaufzeit, durch Ersetzen einer HRTF mit einem Minimalphasensystem, die Signallaufzeit zwischen den beiden Ohren rekonstruiert werden. Wird jedoch anstatt der ITD die TOA selbst berechnet, kann einerseits die TOA ohne minimalphasige Approximation der HRTFs entfernt werden und andererseits der zeitliche Bezug aller HRIRs zueinander aufrechterhalten werden. Die Entfernung der Signallaufzeit aus allen HRTFs bildet die Grundlage zur weiteren genauen Untersuchung der nicht abgewickelten Phasenspektren in zukünftigen Studien.

Die TOAs werden wie die ITDs für alle Richtungen separat geschätzt. Bezogen auf FIR-Filter entspricht die TOA einer mittleren Gruppenlaufzeit, die im gesamten Frequenzbereich auftritt. Im abgewickelten Phasenspektrum entspricht sie einer fallenden Geraden, welche vom minimalphasigen Anteil der Phase und Phasendrehungen, verursacht durch Allpasskomponenten, überlagert wird. Aufgrund dieser Phasendrehungen kann die Schätzung der TOA fehlerhaft sein. Durch die Verwendung eines TOA-Modells können alle TOAs eines Ohres in Beziehung zueinander gesetzt und die Schätzung verbessert werden.

### 3.1 TOA-Schätzung

Aus allen vorgestellten ITD-Schätzmethoden erscheint die Minimalphasenkreuzkorrelation (vgl. Kapitel 2.3.1) als Grundlage für die Bestimmung der TOA am besten geeignet. Die Methode beruht auf der Annahme der minimalphasigen Eigenschaft von HRTFs. Als TOA wird die zeitliche Position des Maximalwerts der Kreuzkorrelationsfunktion zwischen einer HRIR  $h[n]$  und ihrer minimalphasigen Variante  $h_{min}[n]$  definiert:

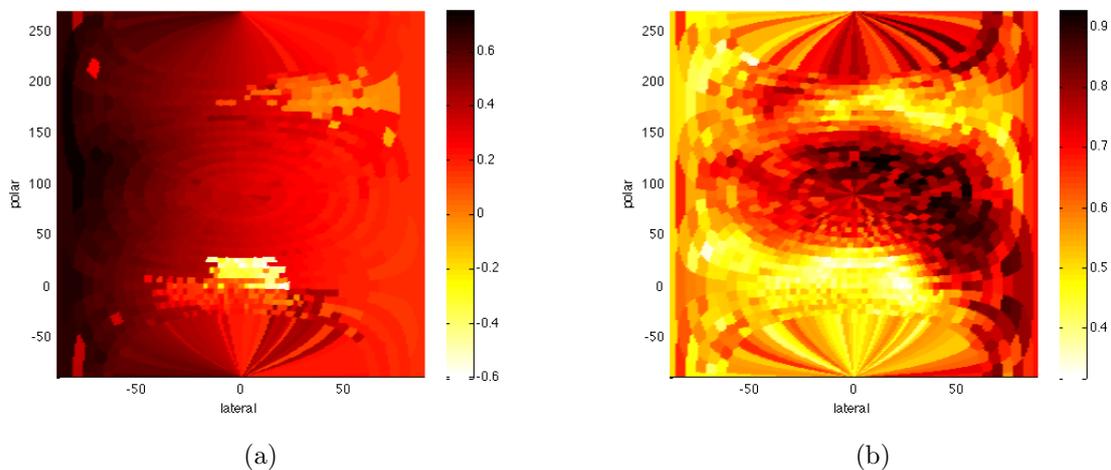
$$\begin{aligned}
 \hat{\tau} &= \arg \max_k \left| \sum_{n=-\infty}^{\infty} h[n] h_{min}[n - k] \right| \\
 &= \arg \max_k \left| \mathcal{F}^{-1} \{ H(\omega) H_{min}^*(\omega) \} [k] \right| \\
 &= \arg \max_k \left| \mathcal{F}^{-1} \{ |H(\omega)|^2 e^{j[-\tau\omega + \varphi_{ap}(\omega)]} \} [k] \right|,
 \end{aligned} \tag{3.1}$$

wobei  $\varphi_{ap}$  die Phase der Allpasskomponenten,  $\tau$  die tatsächliche TOA und  $\hat{\tau}$  die geschätzte TOA darstellt. In Abb. 3.1a ist diese für alle HRTFs eines Ohres im horizontal-polaren Koordinatensystem dargestellt (ein positiver Lateralwinkel repräsentiert hier ipsilaterale Richtungen und ein negativer Winkel kontralaterale Richtungen). Die Schätzung erfolgt im Zeitbereich, wodurch die Problematik der Phasenambiguität keine Rolle spielt. Auch Phasendrehungen beeinflussen hier, im Gegensatz zum linearen Phasen-Fit, nicht direkt das Ergebnis. Durch die Multiplikation einer HRTF mit ihrer konjugiert komplexen minimalphasigen Variante wird das Amplitudenspektrum quadriert. Frequenzbereiche mit niedriger Amplitude (z.B. Notches) werden gegenüber Bereichen mit hohem Pegel im Verhältnis zusätzlich abgeschwächt. Somit fallen bei der Schätzung im Zeitbereich Frequenzen mit korrekt gemessener Phase stärker ins Gewicht und die breitbandige Verzögerung wird nur gering beeinflusst, da minimalphasige Filter eine minimale Gruppenlaufzeit aufweisen (Oppenheim et al., 2004, 364).

Angenommen das minimalphasig approximierte Phasenspektrum einer HRTF entspricht näherungsweise dem verzögerungsfreien Anteil des Phasenspektrums einer HRTF, dann muss die Kreuzkorrelationsfunktion im Zeitbereich der Impulsantwort eines linearphasigen Filters entsprechen, dessen Maximalwert an der Position der TOA zu finden wäre. Diese Annahme kann jedoch nicht verallgemeinert werden (vgl. Kapitel 2.2.4). Besonders an HRTF-Positionen, an denen das Ohr der Schallquelle abgewandt ist, können  $\varphi(\omega)$  und  $\varphi_{min}(\omega)$  stark voneinander abweichen. Ein Gütemaß für die Ähnlichkeit (Kohärenz) einer HRTF mit ihrer minimalphasigen Approximation ist der normalisierte Kreuzkorrelationskoeffizient  $\rho$  (Kulkarni et al., 1999):

$$\rho(\hat{\tau}) = \frac{\sum_{n=0}^{N-1} h[n]h_{min}[n - \hat{\tau}]}{\sum_{n=0}^{N-1} h[n]^2} \quad (3.2)$$

Durch die Normalisierung auf die Energie ergibt die Kreuzkorrelation einen Allpass, welcher einen verzögerten Einheitsimpuls ergeben müsste, falls die HRTF ein Minimalphasensystem wäre. Abbildung 3.1b zeigt  $\rho$  für geschätzte TOA  $\hat{\tau}$  geschätzt mit Formel 3.1 an einem Ohr. Ist das Ohr der Schallquelle zugewandt, liegt er durchgehend über 0.7. Das lässt darauf schließen, dass HRTFs aus diesen Richtungen gut approximiert werden können. Doch größten Teils liegt er unter 0.6, vor allem in der Horizontalebene. Hier wird der Maximalwert der Kreuzkorrelation unter Umständen nicht der TOA entsprechen.<sup>1</sup>

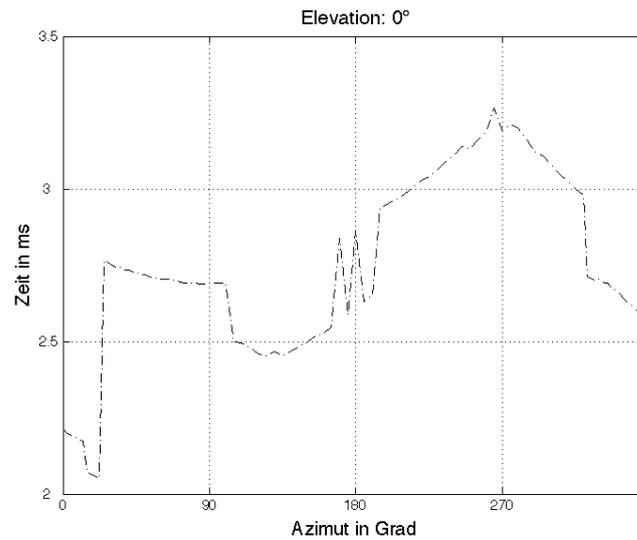


**Abbildung 3.1:** TOA-Schätzung  $\hat{\tau}$  (a) und Kreuzkorrelationskoeffizient  $\rho$  (b) für DTFs eines Ohrs

Vergleicht man Abb. 3.1a und Abb. 3.1b im Bereich um  $0^\circ$  lateral,  $0^\circ$  polar und  $60^\circ$  lateral,  $180^\circ$  polar wird ein Zusammenhang zwischen Kreuzkorrelationskoeffizienten und TOA-Schätzung ersichtlich. Die TOA weist in diesen Bereichen keinen glatten Verlauf auf und die Kreuzkorrelationskoeffizienten sind sehr niedrig (ca. 0.4). Hier ist besonders mit Fehlschätzungen zu rechnen.

In Abb. 3.2 ist der TOA-Verlauf in der Horizontalebene dargestellt. Eine als korrekt angenommene Kurve lässt sich erahnen und es ist erkennbar, dass die abweichenden TOAs hauptsächlich zu früh geschätzt werden. Die Abweichung vom angenommenen Kurvenverlauf scheint nicht normal verteilt zu sein. Bei einem Least-Square-Fit für eine modellierte TOA

1. In allen folgenden Abbildungen sind Ergebnisse der Berechnungen aus DTFs dargestellt, da DTFs für die TOA-Bestimmung den Worst-Case darstellten. Im Text wird jedoch immer allgemein von HRTFs gesprochen.



**Abbildung 3.2:** TOA-Schätzung bei 0° Elevation für DTFs des linken Ohrs

ist deshalb mit einer starken Abweichung zu rechnen. Deshalb müssen Ausreißer gesondert erkannt und entfernt werden.

## 3.2 Räumliches TOA-Modell

Die in Kapitel 3.1 vorgestellte Schätzung wird für jede Richtung separat durchgeführt. Es handelt sich um eine blinde Schätzung. Bei der HRTF-Messung ist der Messaufbau allerdings bekannt, ebenso wie die Position des Messobjektes, welches die richtungsabhängigen Weg-, Laufzeitdifferenzen hervorruft. Die klare geometrische Form des Messaufbaus lässt auf einen Zusammenhang zwischen den TOAs einzelner Richtungen schließen. Die TOAs von HRTFs können wie Schalllaufzeiten um eine, an den Kopf angenäherte, Kugel modelliert werden. Das heißt, die TOA wird als eine Funktion der sphärischen Winkel und weniger Parameter angenommen:

$$\tilde{\tau} = f(\phi, \theta) \quad (3.3)$$

Es ist davon auszugehen, dass sich die TOAs aufgrund der feinen räumlichen Abtastung von Messposition zu Messposition nur gering ändern und einen glatten Verlauf aufweisen.<sup>2</sup> Deshalb kann dieses a priori Wissen über den richtungsabhängigen TOA-Verlauf zur Verbesserung der TOA-Bestimmung herangezogen werden.

<sup>2</sup> Ausgenommen von dieser Bedingung ist die Richtung der maximalen Abschattung (kontralaterale Richtung). Hier tritt eine Unstetigkeit in der Änderung der TOA auf.

Die Grundlage für das vorgestellte TOA-Modell bildet das ITD-Modell nach [Woodworth und Schlossberg \(1962\)](#). Dieses ITD-Modell ist durch die fixe Position der Ohren bei  $\pm\frac{\pi}{2}$  auf einen Freiheitsgrad, den Radius, beschränkt. Im ersten Schritt wird dieses ITD-Modell auf die TOA verallgemeinert und die Verschiebung der Ohrposition erlaubt. Dadurch erhält man zwei weitere Freiheitsgrade  $\phi_o$  (Azimut der Ohrposition) und  $\theta_o$  (Elevation der Ohrposition) und ein einfaches TOA-Modell. Mit einer weiteren Verallgemeinerung wird auch eine geringe Verschiebung  $|\vec{M}| < r$  des Mittelpunkts der Kugel  $\vec{M}$  zum Mittelpunkt der Messanordnung  $\vec{O}$  zugelassen und im Weiteren erweitertes TOA-Modell bezeichnet.

In [Abbildung 3.3](#) ist eine Skizze des TOA-Modells dargestellt. Die gelben Vektoren stellen die einfallende ebene Schallwelle dar. Die dunkelgraue Kugel um den Mittelpunkt  $\vec{M}$  mit Radius  $r$  modelliert den menschlichen Kopf. Die dreidimensionale Darstellung soll veranschaulichen, dass die Großkreise (blau und grün), auf die im Weiteren verwiesen wird, auf richtungsabhängig unterschiedlich liegenden Ebenen liegen.

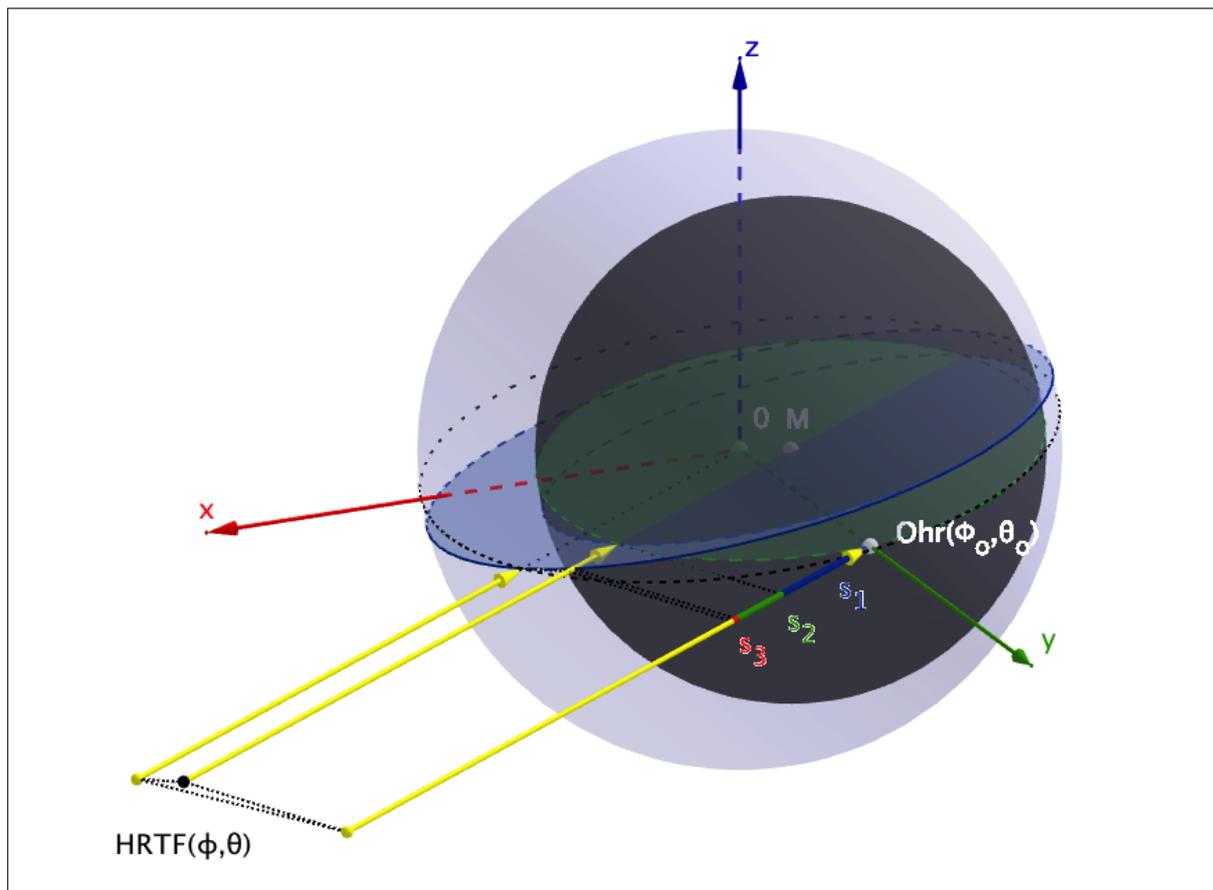


Abbildung 3.3: TOA-Modell (3D-Ansicht)

### 3.2.1 Einfaches TOA-Modell

Im Sonderfall  $\vec{M} = \vec{0}$  (die beiden Kugeln in Abb. 3.3 wären identisch) spannen der Vektor der Schalleinfallrichtung  $\vec{HRTF}$  und der Vektor der Ohrposition  $\vec{Ohr}$  eine Ebene auf, die die Einheitskugel per Definition immer in einem Großkreis schneiden.

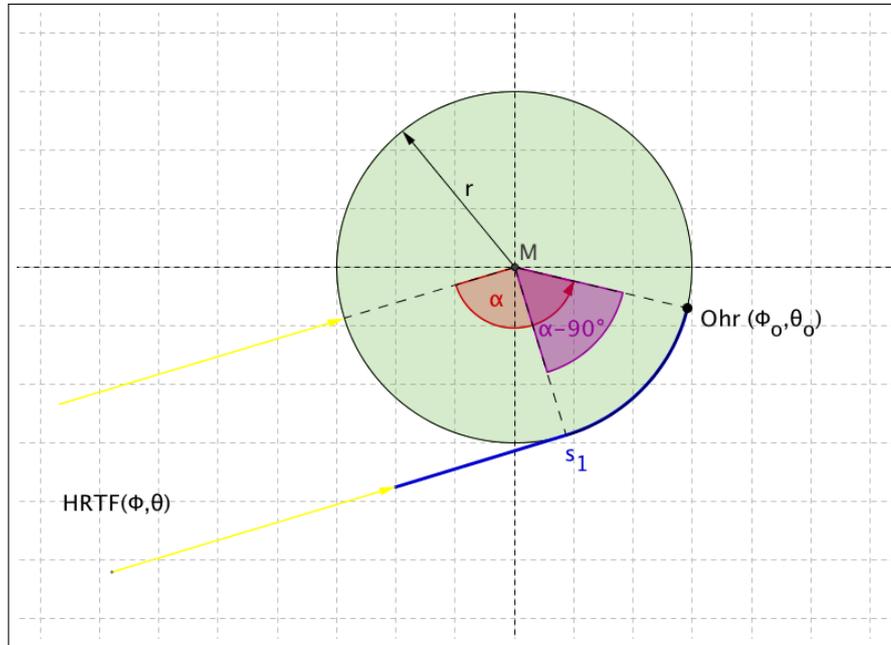


Abbildung 3.4: Einfaches TOA-Modell (Skizze)

Der eingeschlossene Winkel zwischen den beiden Vektoren lässt sich aus deren Skalarprodukt berechnen. Der eingeschlossene Winkelbogen stellt die kürzeste Entfernung zwischen zwei Punkten auf der Kugel dar.  $s_1$  beschreibt den Weg vom ersten Auftreffen der ebenen Welle auf die Kugel bis zur Ohrposition. Ist das Ohr der einfallenden Schallwelle zugewandt ( $\alpha \leq \pi$ ), ergeben sich nur direkte Wegdifferenzen. Im Gegensatz dazu kommen im kontralateralen Fall ( $\alpha > \pi$ ) zusätzlich Schalllaufzeiten aufgrund der Beugung um die Kugel hinzu. Um die Lautsprecherentfernung in der Messanordnung miteinzubeziehen wird noch ein richtungsunabhängiges Delay  $\bar{\tau}$  eingeführt.

$$\tilde{\tau}(\phi, \theta) = \begin{cases} \frac{r}{c} [1 - \sin(\theta_o) \sin(\theta) - \cos(\theta_o) \cos(\theta) \cos(\phi_o - \phi)] + \bar{\tau} & \text{für } \alpha \leq \frac{\pi}{2} \\ \frac{r}{c} [1 + \cos^{-1}(\sin(\theta_o) \sin(\theta) + \cos(\theta_o) \cos(\theta) \cos(\phi_o - \phi)) - \frac{\pi}{2}] + \bar{\tau} & \text{für } \alpha > \frac{\pi}{2} \end{cases} \quad (3.4)$$

### 3.2.2 Erweitertes TOA-Modell

Als Erweiterung zum einfachen TOA-Modell wird nun eine Verschiebung der Kugel zum Mittelpunkt der Messanordnung zugelassen. Im Gegensatz zum einfachen Modell beginnt nach der Verschiebung des Kugelmittelpunktes der richtungsabhängige Anteil der TOA nicht beim ersten Auftreffen der ebenen Schallwelle auf die Modellkugel, sondern auf einer Kugel im Ursprung mit Radius  $r + |\vec{M}|$  (vgl. Abb. 3.3: große Kugel). Dadurch treten zwei zusätzliche richtungsabhängige direkte Schallwege  $s_1$  und  $s_2$  (ohne Beugung) auf.

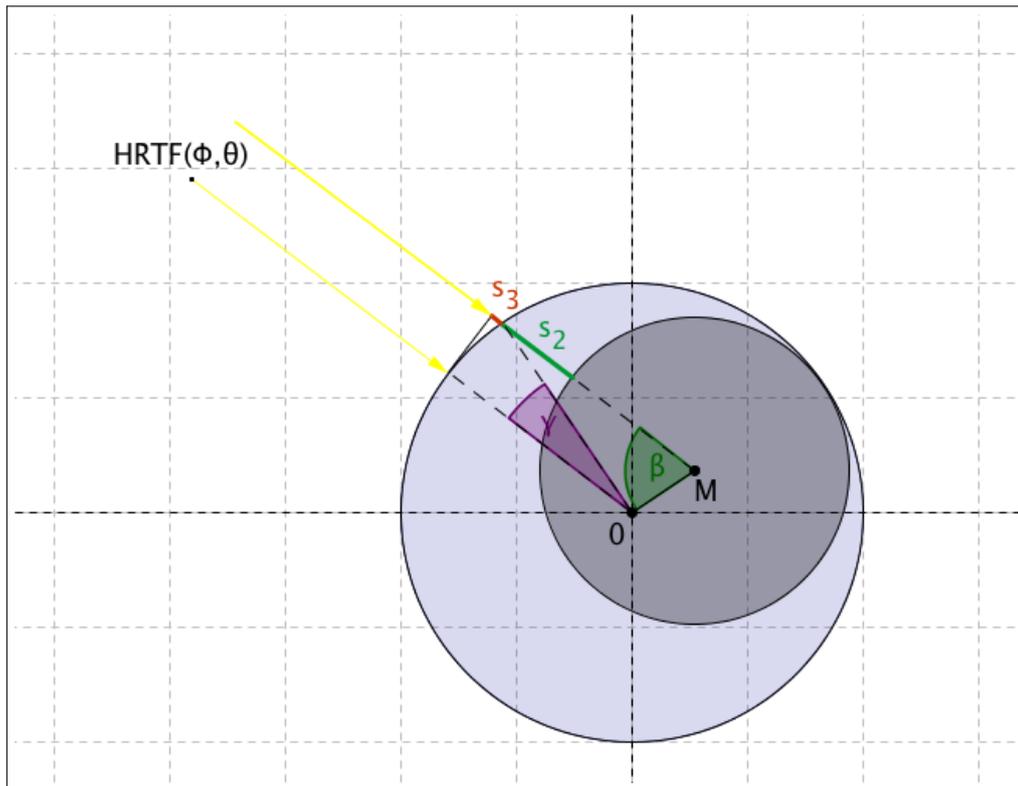


Abbildung 3.5: Erweitertes TOA-Modell (Skizze)

In Abbildung 3.5 ist eine Skizze des erweiterten Modells zu sehen. Es zeigt zwei Großkreise, die durch den Ortsvektor des Mittelpunkts der Kugel und die HRTF-Richtung aufgespannt werden (Achtung: der kleine Großkreis entspricht nicht dem Großkreis beim einfachen Modell). Über geometrische Beziehungen können  $\beta$ ,  $\gamma$ ,  $s_2$  und  $s_3$  berechnet werden (genaue Herleitung siehe Anhang A):

$$\beta = \cos^{-1}\{-\cos(\theta)[x_M \cos(\phi) + y_M \sin(\phi)] - z_M \sin(\theta)\} \quad (3.5)$$

$$s_2 = -r + |\vec{M}| \cos(\beta) + \sqrt{r^2 + |\vec{M}|^2 \cos^2(\beta) + 2|\vec{M}|r} \quad (3.6)$$

$$\gamma = \pi - \beta - \cos^{-1} \left\{ \frac{2|\vec{M}|^2 + 2|\vec{M}|r - 2rs_2 - s_2^2}{2|\vec{M}|^2 + 2|\vec{M}|r} \right\} \quad (3.7)$$

$$s_3 = \frac{|\vec{M}|\sin(\beta)}{2(|\vec{M}| + r)\tan(\frac{\gamma}{2})} \quad (3.8)$$

Der restliche Schallweg um die Kugel  $s_1$  lässt sich analog zum einfachen Modell berechnen, wobei  $\phi_o$  und  $\theta_o$  von  $\vec{M}$  ausgehend die Position eines Ohrs auf der Modellkugel beschreiben. Für das gesamte erweiterte Modell erhält man als Ergebnis:

$$\tilde{\tau}_e(\phi, \theta) = \begin{cases} \frac{1}{c} \left\{ -r\sin(\theta_o)\sin(\theta) - r\cos(\theta_o)\cos(\theta)\cos(\phi_o - \phi) \right. \\ \quad + |\vec{M}|\cos(\beta) + \sqrt{r^2 + |\vec{M}|^2\cos^2(\beta) + 2|\vec{M}|r} \\ \quad \left. + \frac{|\vec{M}|\sin(\beta)}{2(|\vec{M}| + r)\tan(\frac{\gamma}{2})} \right\} + \bar{\tau} & \alpha \leq \frac{\pi}{2} \\ \frac{1}{c} \left\{ r[\cos^{-1}(\sin(\theta_o)\sin(\theta) + \cos(\theta_o)\cos(\theta)\cos(\phi_o - \phi)) - \frac{\pi}{2}] \right. \\ \quad + |\vec{M}|\cos(\beta) + \sqrt{r^2 + |\vec{M}|^2\cos^2(\beta) + 2|\vec{M}|r} \\ \quad \left. + \frac{|\vec{M}|\sin(\beta)}{2(|\vec{M}| + r)\tan(\frac{\gamma}{2})} \right\} + \bar{\tau} & \alpha > \frac{\pi}{2} \end{cases} \quad (3.9)$$

### 3.3 Einfluss der Parameter

Abbildung 3.6 und 3.7 zeigen TOAs in der Horizontalebene für unterschiedlichste Parameter-einstellungen beim einfachen und erweiterten Modell. In Abbildung 3.6 wird für das einfache Modell bei den Kurven 1 bis 5 der Kugelradius von  $6\text{cm}$  bis  $10\text{cm}$ , in Kurve 6 bis 10  $\phi_o$  von  $70^\circ$  bis  $110^\circ$  und in Kurve 11 bis 13  $\theta$  von  $-20^\circ$  bis  $0^\circ$  variiert. Hier ist deutlich erkennbar, dass der Radius den Abstand zwischen Minimum und Maximum und die Position des Ohrs deren Ausrichtung bestimmt. In Abbildung 3.7 wird für das erweiterte Modell bei den Kurven 14 bis 18 der Kugelradius von  $6\text{cm}$  bis  $10\text{cm}$ , in Kurve 19 bis 23  $x_M$ , in Kurve 24 bis 28  $y_M$  und in Kurve 29 bis 33  $z_M$  jeweils von  $-2\text{cm}$  bis  $2\text{cm}$  variiert, wobei  $\phi_o$  und  $\theta_o$  jeweils so eingestellt werden, dass die Ohrposition immer auf der Transversalachse liegt. Der Unterschied zum einfachen Modell liegt einerseits im Abstand zwischen Minimum und Maximum, welcher größtenteils von zwei Parametern ( $r$  und  $y_M$ ) abhängt, und andererseits in der Positionen von Minimum und Maximum, welche nicht  $180^\circ$  voneinander entfernt sein müssen.

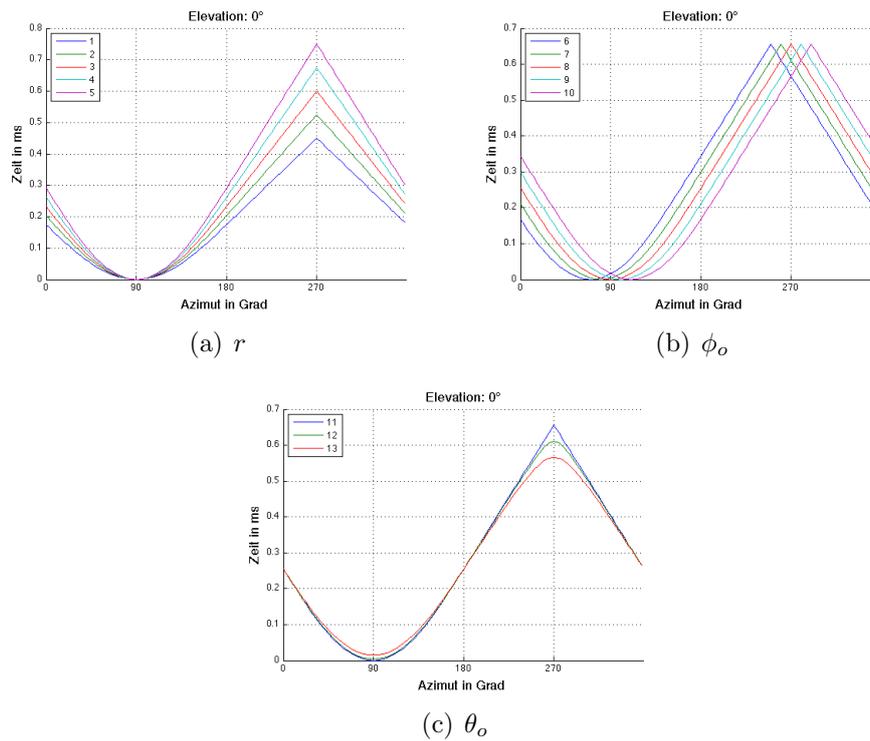


Abbildung 3.6: Variation der Parameter des einfachen Modells

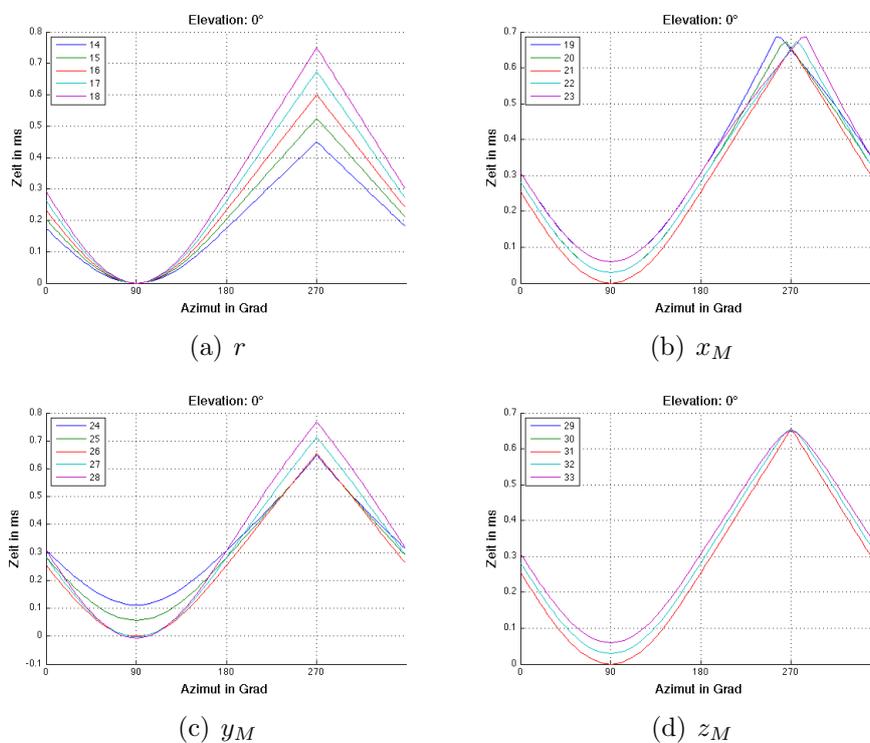


Abbildung 3.7: Variation der Parameter des erweiterten Modells

### 3.4 Parameteranpassung

Um die Parameter der beiden Modelle  $r$ ,  $\phi_o$ ,  $\theta_o$ ,  $\vec{M}$  und  $\bar{\tau}$  an die TOA-Schätzungen anzupassen, wird der quadratische Fehler  $e^2$  zwischen Modell  $\tilde{\tau}(\phi, \theta)$  und Schätzung  $\hat{\tau}_{\phi, \theta}$  (für alle  $\phi, \theta$  aus der Menge aller als korrekt angenommenen Schätzungen  $\hat{\tau}_K$ ) minimiert.

$$e^2 = \frac{1}{M} \sum_{\theta} \sum_{\phi} [\hat{\tau}_{\phi, \theta} - \tilde{\tau}(\phi, \theta)]^2 \quad \forall(\phi, \theta) | \hat{\tau}_{\phi, \theta} \in \hat{\tau}_K \quad (3.10)$$

$\hat{\tau}_K$  repräsentiert die Menge aller TOA-Schätzungen, welche im folgenden Abschnitt als korrekte Schätzungen bezeichnet werden, also keine Ausreißer sind.  $M$  ist die Anzahl der Elemente in dieser Menge. Die Ausreißererkennung ist notwendig, da diese den quadratischen Fehler stark beeinflussen. Auch die Wahl der Startwerte der Modellparameter spielt eine Rolle. Da das einfache TOA-Modell weniger Freiheitsgrade aufweist und leichter anzupassen ist, weil zum Beispiel der Abstand zwischen Minimum und Maximum nur vom Radius abhängt (siehe Kapitel 3.3), wird es verwendet um Startwerte für das erweiterte Modell zu berechnen.

#### Schritt 1:

Um kostengünstig eine hohe räumliche Auflösung bei der Messung von HRTFs zu ermöglichen, wird meistens die Versuchsperson innerhalb einer fixen Anordnung weniger Lautsprecher horizontal gedreht. Deshalb befinden sich die Messpositionen auf einigen Elevationsstufen mit hoher lateraler Auflösung. Am ARI werden HRTFs auf 22 Elevationsstufen (von  $-30^\circ$  bis  $80^\circ$ ), mit einer horizontalen Auflösung von bis zu  $2.5^\circ$  gemessen. Deshalb eignet sich als erstes Kriterium zur Ausreißererkennung die Glattheit der richtungsabhängigen TOA bezüglich des Azimuts für jede Elevationsstufe. Dafür wird die horizontale TOA-Änderung berechnet:

$$\hat{\tau}'_{\phi, \theta} = \frac{\hat{\tau}_{\phi, \theta} - \hat{\tau}_{\phi + \Delta\phi, \theta}}{\Delta\phi} \quad (3.11)$$

Als Ausreißer werden alle Richtungen definiert, deren Azimut-abhängige Änderung  $\hat{\tau}'_{\phi, \theta}$  größer als die mittlere quadratische Änderung aller Richtungen ist. Alle anderen Richtungen werden vorläufig als korrekte Schätzungen  $\hat{\tau}_K$  bezeichnet.

$$|\hat{\tau}_{\phi, \theta}| < \sqrt{\frac{1}{M} \sum_{\phi, \theta} (\hat{\tau}'_{\phi, \theta})^2} \Rightarrow \hat{\tau}_{\phi, \theta} \in \hat{\tau}_K \quad (3.12)$$

Abb. 3.8 zeigt die geschätzte TOA bei Elevation  $5^\circ$  und markierte erkannte Ausreißer (rote Punkte). Es ist einerseits deutlich zu sehen, dass viele Ausreißer korrekt erkannt werden,

aber auch Schätzungen, die bei visueller Kontrolle als richtige TOA-Werte empfunden werden würden, markiert werden, da die Steigung zwischen zwei Punkten definiert ist und keine Aussage darüber getroffen werden kann, welcher von beiden der Ausreißer ist. Andererseits gibt es Stellen, an denen die TOA-Kurve glatt verläuft aber einen Offset zu den restlichen Schätzungen aufweisen und diese nicht als Ausreißer erkannt werden. Diese Fälle werden im 2. Schritt erfasst.

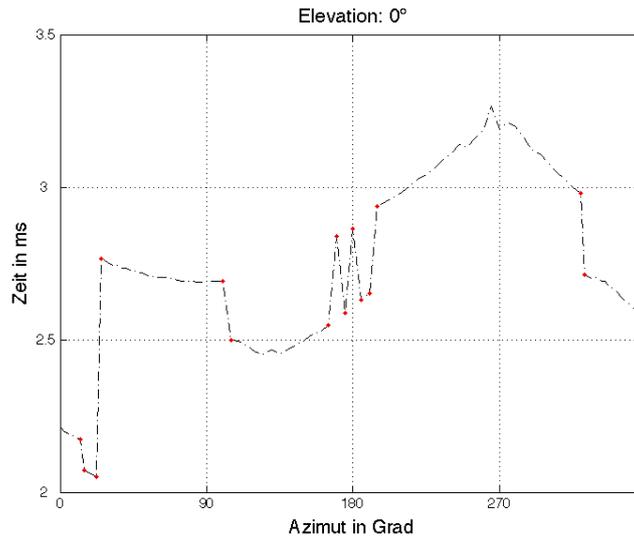


Abbildung 3.8: Erkannte Ausreißer nach Schritt 1

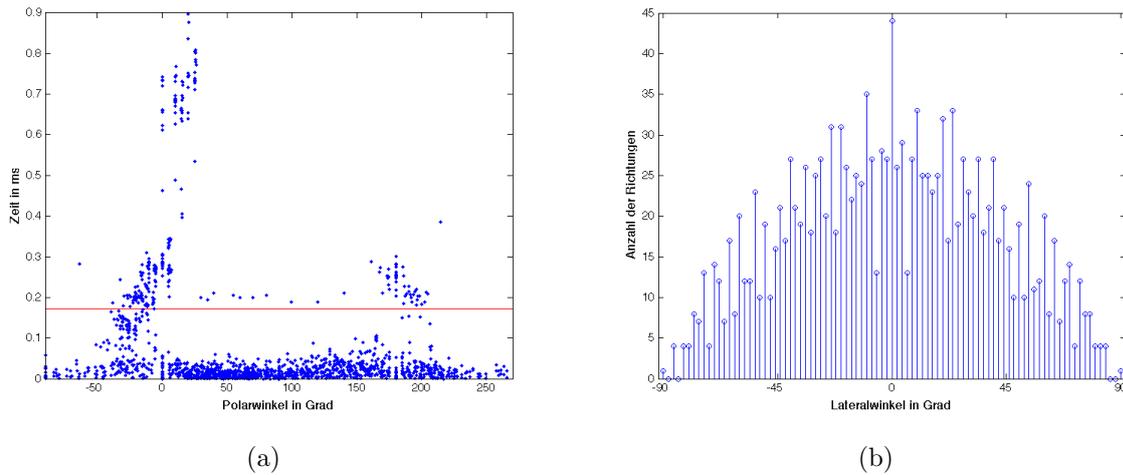
### Schritt 2:

Im zweiten Schritt wird die Eigenschaft ausgenutzt, dass die TOAs aus Richtungen mit gleichem Lateralwinkel (Sagittalebene) näherungsweise konstant sind. Aufgrund des Messaufbaus müssen die Messrichtungen aber nicht auf einer Sagittalebene liegen, was bei den ARI-HRTFs der Fall ist. Deshalb werden HRTFs zwischen jeweils zwei Sagittalebene im Abstand  $\epsilon$  zu einer Gruppe  $G$  konstanter lateralen Auslenkung  $\Phi_G$  zusammengefasst. In mehreren Durchläufen wird in jeder Gruppe der Mittelwert  $\mu_G$  aus  $\hat{\tau}_{K,G}$  berechnet und die Varianz dieser Gruppe  $\sigma_G^2$  berechnet:

$$\overline{\sigma_G^2} = \frac{1}{G} \sum_G \sigma_G^2 = \frac{1}{G} \sum_G \left( \sum_i^N [\hat{\tau}_{\Phi_G \pm \epsilon, \Theta_i}] - \underbrace{\frac{1}{N} \sum_j^N [\hat{\tau}_{\Phi_G \pm \epsilon, \Theta_j}]}_{\mu_G} \right)^2. \quad (3.13)$$

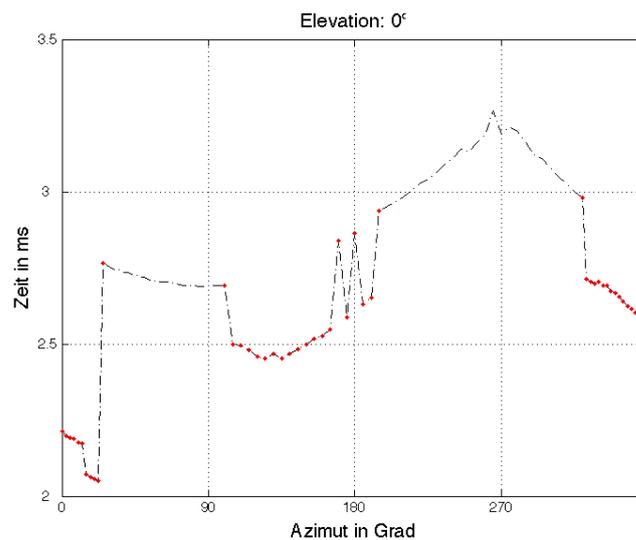
Ausreißer sind Richtungen deren TOAs außerhalb der mittleren Varianz  $\overline{\sigma_G^2}$  liegen:

$$[\hat{\tau}_{\phi, \theta} - \mu_G]^2 > \overline{\sigma_G^2} \quad \Rightarrow \quad \hat{\tau}_{\phi, \theta} \notin \hat{\tau}_K. \quad (3.14)$$



**Abbildung 3.9:** TOAs nach Entfernung des lateral konstanten Delays aus  $\Phi_G$  (a) und Anzahl der Richtungen in den lateralen Gruppen (b)

In Abbildung 3.9 sind die quadratischen Abweichungen der TOAs von den mittleren TOAs ihrer lateralen Gruppen  $\mu_G$  als Funktion vom Polarwinkel zusammengefasst. In Rot ist die mittlere Varianz  $\overline{\sigma_G^2}$  eingezeichnet. Gemeinsam mit der Bedingung aus Schritt 1 bildet diese Methode eine sehr effiziente Fehlererkennung als Grundlage für die Modellanpassung (siehe Abb. 3.10).



**Abbildung 3.10:** Erkannte Ausreißer nach Schritt 2

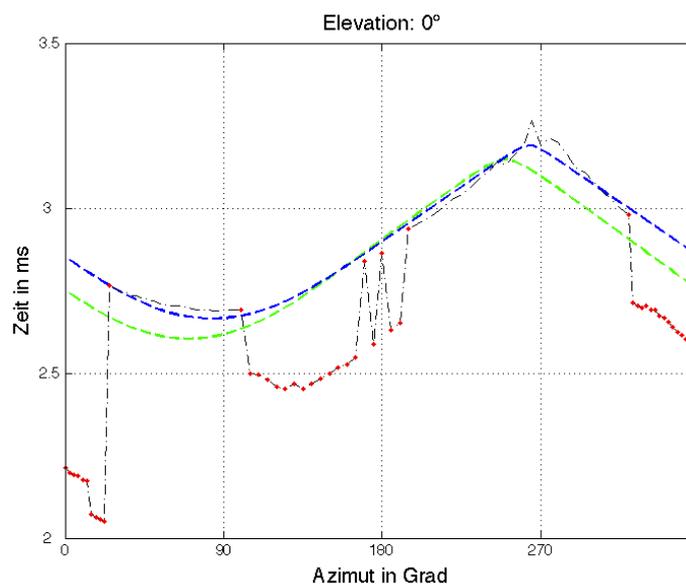
**Schritt 3:**

Als nächstes wird das einfache Modell an die Menge der korrekten TOAs  $\hat{\tau}_K$  angepasst, um daraus Startwerte für das folgende erweiterte Modell zu bestimmen.

Als Startwerte der Modellparameter für das einfache Modell werden für die Ohrposition  $\phi_o = \pm 90^\circ$  und  $\theta = 0^\circ$  verwendet. Für den Radius wird der durchschnittliche Kopfradius von  $r = 8.75 \text{ cm}$  angenommen (Hartley und Fry, 1921). Für die richtungsunabhängige Signallaufzeit wird das Minimum aus  $\hat{\tau}_K$  als Startwert genutzt:

$$\bar{\tau} = \min\{\hat{\tau}_K\} \quad (3.15)$$

Abbildung 3.11 zeigt die Anpassung des einfachen Modells. Die grüne Kurve entspricht der Anpassung ohne Ausreißerererkennung (Menge  $\hat{\tau}$ ), die blaue Kurve der Anpassung mit Ausreißerererkennung (Menge  $\hat{\tau}_K$ ). Es ist eine deutliche Verbesserung zu erkennen. Weiters ist zu sehen, dass das Maximum des Modells nicht exakt beim Maximum der Schätzung liegt. Tabelle 3.1 zeigt die gewonnenen Parameter nach separater Anpassung des Modells an beide Ohren. Es stechen die unterschiedlichen Radien der beiden Anpassungen heraus. Diese Unterschiede sind physiologisch schwer erklärbar und liegen wahrscheinlich an zu großer Vereinfachung im einfachen Modell.



**Abbildung 3.11:** Anpassung des einfachen Modells nach Schritt 3 (**grün:** ohne Ausreißerererkennung, **blau:** mit Ausreißerererkennung)

	$\mathbf{r}$	$\phi_o$	$\theta_o$	$\bar{\tau}$
linkes Ohr	7.39 cm	82.52°	-4.86°	2.653 ms
rechtes Ohr	8.76 cm	-87.87°	-14.10°	2.631 ms

**Tabelle 3.1:** Parameter des einfachen Modells

#### Schritt 4:

Zuerst werden die Resultate des einfachen Modells verwendet, um noch einmal die Ausreißerererkennung zu verfeinern. Dafür wird der Abstand zwischen modellierten TOAs und der Menge aller geschätzten TOAs herangezogen. Liegt eine Schätzung außerhalb des  $k$ -fachen Bereichs um der mittleren quadratischen Differenz zwischen Modell und Schätzung, so wird die Schätzung als Ausreißer gekennzeichnet und aus der Menge der korrekten Schätzungen entfernt.

$$[\hat{\tau}_{\phi,\theta} - \tilde{\tau}(\phi, \theta)]^2 > k^2 \frac{1}{M} \sum_{\phi,\theta} [\hat{\tau}_{\phi,\theta} - \tilde{\tau}(\phi, \theta)]^2 \Rightarrow \hat{\tau}_{\phi,\theta} \notin \hat{\tau}_K \quad (3.16)$$

Liegt eine Schätzung innerhalb des  $\frac{1}{k^2}$ -fachen Bereichs um der mittleren quadratischen Differenz zwischen Modell und Schätzung, so wird die Richtung zur Menge der korrekten Schätzungen hinzugefügt.

$$[\hat{\tau}_{\phi,\theta} - \tilde{\tau}(\phi, \theta)]^2 < \frac{1}{k^2} \frac{1}{M} \sum_{\phi,\theta} [\hat{\tau}_{\phi,\theta} - \tilde{\tau}(\phi, \theta)]^2 \Rightarrow \hat{\tau}_{\phi,\theta} \in \hat{\tau}_K \quad (3.17)$$

In dieser Arbeit erwies sich  $k^2 = 4$  als guter Wert. Somit wird die Anzahl der plausiblen Positionen als Datenmenge für die Parameteranpassung maximiert.

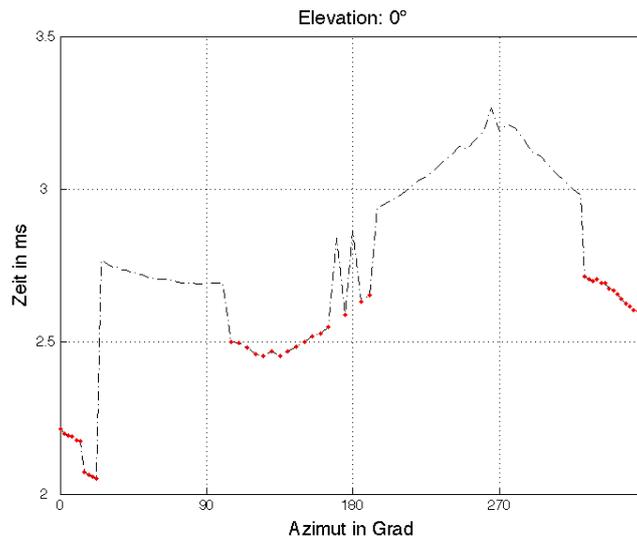
Die gewonnenen Parameter aus Schritt 3 werden nun zur Berechnung der Startwerte für das erweiterte Modell herangezogen. Radius und seitliche Verschiebung ergeben sich aus den beiden Kugelradien:

$$r = \frac{r_l + r_r}{2}, \quad y_M = \frac{r_l - r_r}{2} \quad (3.18)$$

Die Startwerte für  $x_M$  und  $z_M$  werden aus den beiden Winkeln der Ohrposition berechnet:

$$x_M = -r \cos(\phi_o) \cos(\theta_o), \quad z_M = -r \sin(\theta_o) \quad (3.19)$$

Im erweiterten Modell werden für HRTFs der ARI-Datenbank die Ohrpositionen auf der Transversalachse festgehalten, da diese Positionierung mittels Ausrichtung an Laserpointern



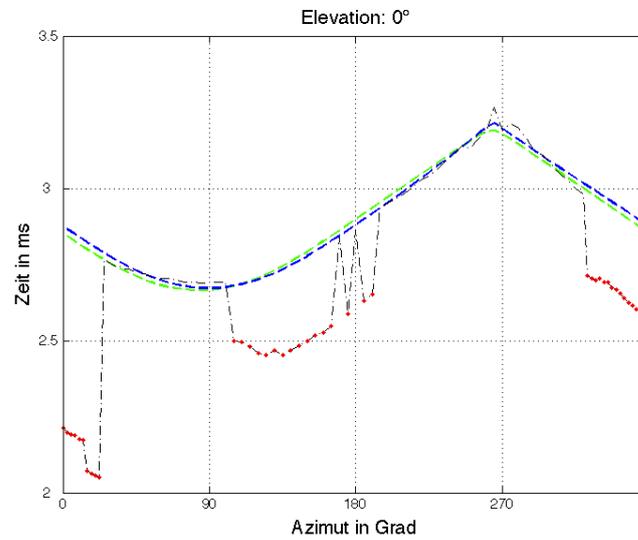
**Abbildung 3.12:** Erkannte Ausreißer nach der Verfeinerung der Detektion mithilfe des einfachen Modells

bei der Messung garantiert ist. Abbildung 3.13 zeigt das Ergebnis der Anpassung für die Horizontalebene (die Ergebnisse aller Richtungen sind in Anhang B Abbildung B.1 abgebildet). Als Modellparameter ergeben sich für die gezeigte Versuchsperson:

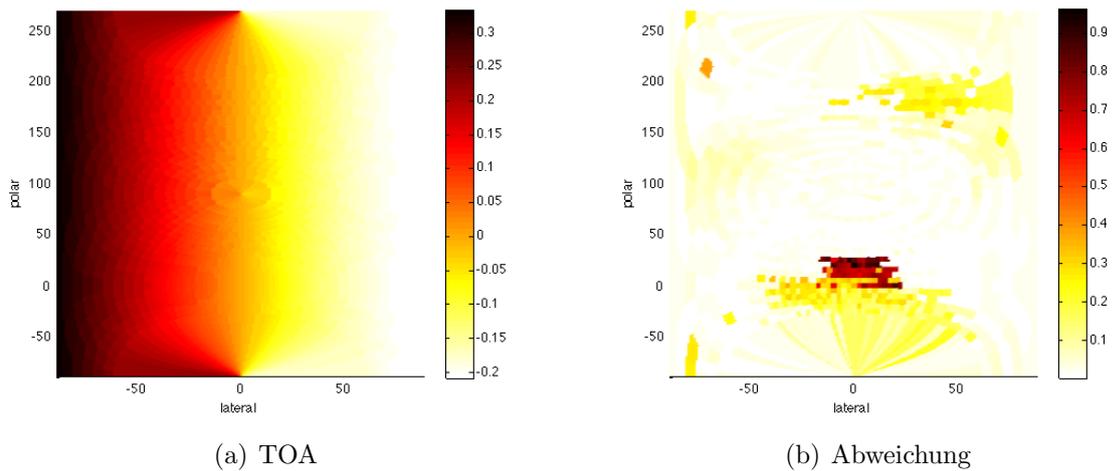
	$r$	$x_M$	$y_M$	$z_M$	$\bar{\tau}$
linkes Ohr	7.7 cm	-0.14 cm	-0.55 cm	0 cm	2.64 ms
rechtes Ohr	7.8 cm	-0.18 cm	-0.55 cm	0 cm	2.66 ms

**Tabelle 3.2:** Parameter des erweiterten Modells

Abbildung 3.14a zeigt die erhaltene TOA geplottet über Lateral- und Polarwinkel. In Abb. 3.14b ist die Differenz zwischen der erhaltenen TOA und der mit Gleichung 3.1 geschätzten TOA zu sehen. Hier ist noch einmal eine große Abweichung im Bereich von  $0^\circ$  lateral,  $0^\circ$  polar und  $60^\circ$  lateral,  $180^\circ$  polar zu erkennen (vgl. Abb. 3.1a und 3.1b).



**Abbildung 3.13:** Anpassung nach Schritt 4 (**grün:** einfaches Modell aus Schritt 3, **blau:** erweitertes Modell)



**Abbildung 3.14:** Modellierte TOA und die Abweichung von der geschätzten TOA jeweils in *ms*

# Kapitel 4

## Evaluierung

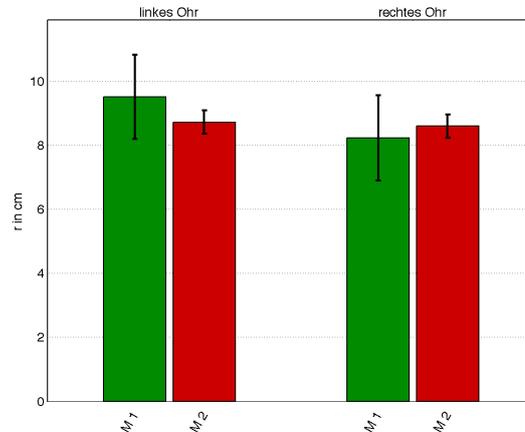
### 4.1 Numerische Evaluierung

Zur numerischen Evaluierung der TOA-Modellierung stehen 56 Datensätze der ARI-HRTF-Datenbank (10 mit anthropometrischen Daten) und 25 Datensätze der CIPIC-Datenbank (alle mit anthropometrischen Daten) zur Verfügung. Die Bewertung der TOA-Modellierung durch ein objektives Fehlermaß macht hier keinen Sinn, da es keinen Referenzwert gibt. Es kann aber versucht werden die Modellparameter mit der Physiologie der Versuchspersonen zu vergleichen und in Beziehung zu setzen. Zum einen wird die Schwankung der Parameter zwischen den beiden Ohren einer Versuchsperson, zum anderen die Plausibilität der Ergebnisse in Bezug auf die anthropometrischen Daten untersucht. Die ermittelten Werte der Parameter des einfachen und erweiterten Modells sind in den Tabellen B.1 bis B.4 geordnet nach den Versuchspersonen im Anhang B angefügt.

#### 4.1.1 Binauraler Vergleich

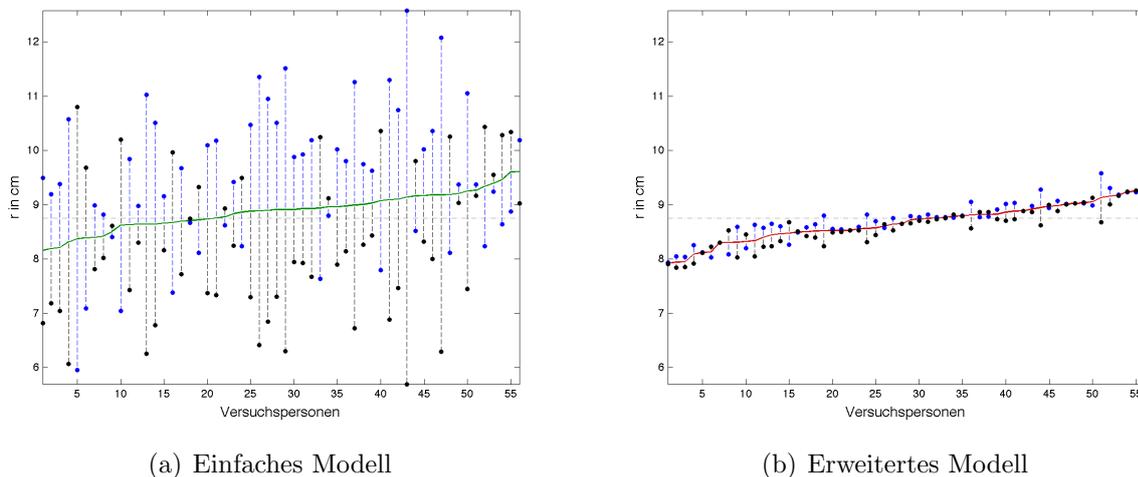
Beim einfachen Modell befindet sich der Kugelmittelpunkt wie beim ITD-Modell nach Woodworth und Schlossberg (Algazi et al., 2001) im Mittelpunkt der Messanordnung. Dabei wird allerdings angenommen, dass der Kopf der Versuchspersonen bei der HRTF-Messung korrekt ausgerichtet ist. Anders ausgedrückt befindet sich bei idealen Messbedingungen der Mittelpunkt der interauralen Achse einer Versuchsperson am Mittelpunkt der Messanordnung. Wäre diese Bedingung erfüllt, müsste das einfache Modell der menschlichen Physiologie entsprechende Ergebnisse liefern. Zum Beispiel sollte sich bei der statistischen Auswertung der erhaltenen Kugelradien für linkes und rechtes Ohr ungefähr der gleiche Mittelwert ergeben. Aus Abbildung 4.1 (grüne Balken) ist leicht ersichtlich, dass diese Annahme für die ARI-

HRTF-Datenbank nicht zutrifft. Der Radius bei der Modellanpassung ist beim rechten Ohr im Mittel kleiner.<sup>1</sup>



**Abbildung 4.1:** Modellierte Kugelradien für HRTFs der ARI-Datenbank (**M1:** einfaches TOA-Modell, **M2:** erweitertes TOA-Modell)

Bei der genaueren Betrachtung der einzelnen Kopfradien in Abb. 4.2a (die Versuchspersonen sind aufsteigend nach dem Mittelwert der beiden Kugelradien geordnet) wird aber klar, dass die beiden unterschiedlichen Kopfradien bei einer Versuchsperson voneinander abhängig sind. Der Mittelwert zwischen jeweils linkem und rechtem Ohr liegt im erwarteten Bereich.



(a) Einfaches Modell

(b) Erweitertes Modell

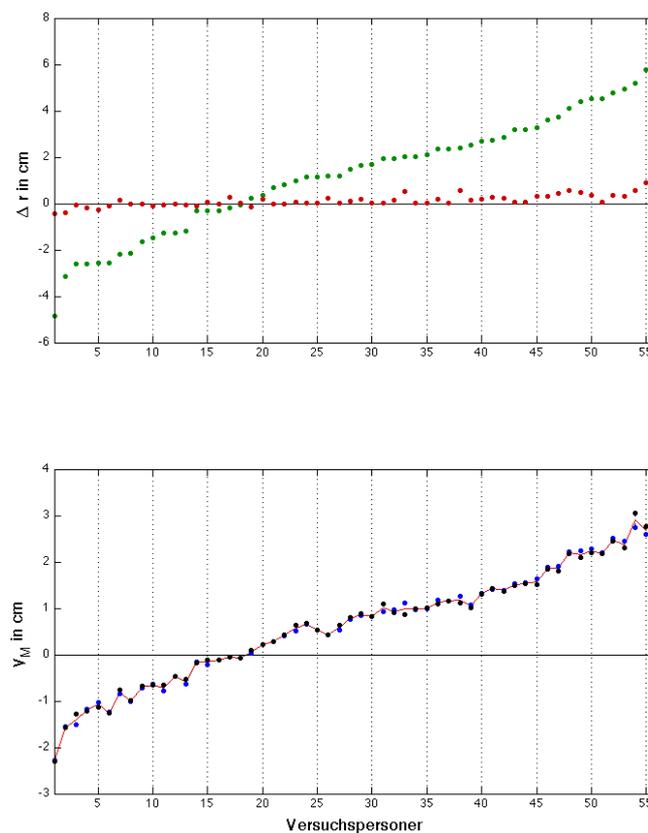
**Abbildung 4.2:** Kugelradien für 56 Versuchspersonen (**blau:** linkes Ohr, **schwarz:** rechtes Ohr, **grün/rot:** Mittelwert beider Ohrradien)

Eine Erklärung hierfür findet man in der Position des Kopfmittelpunkts. Ist die Ausrichtung der Versuchspersonen bei der Messung nicht ideal, resultiert dies in einer Fehlschätzung

1. Diese Annahme trifft ebenso wenig auf die HRTF-Sätze der CIPIC-Datenbank zu, woraus geschlossen werden kann, dass allgemein Positionierungsschwierigkeiten bei der HRTF-Messung zu erwarten sind.

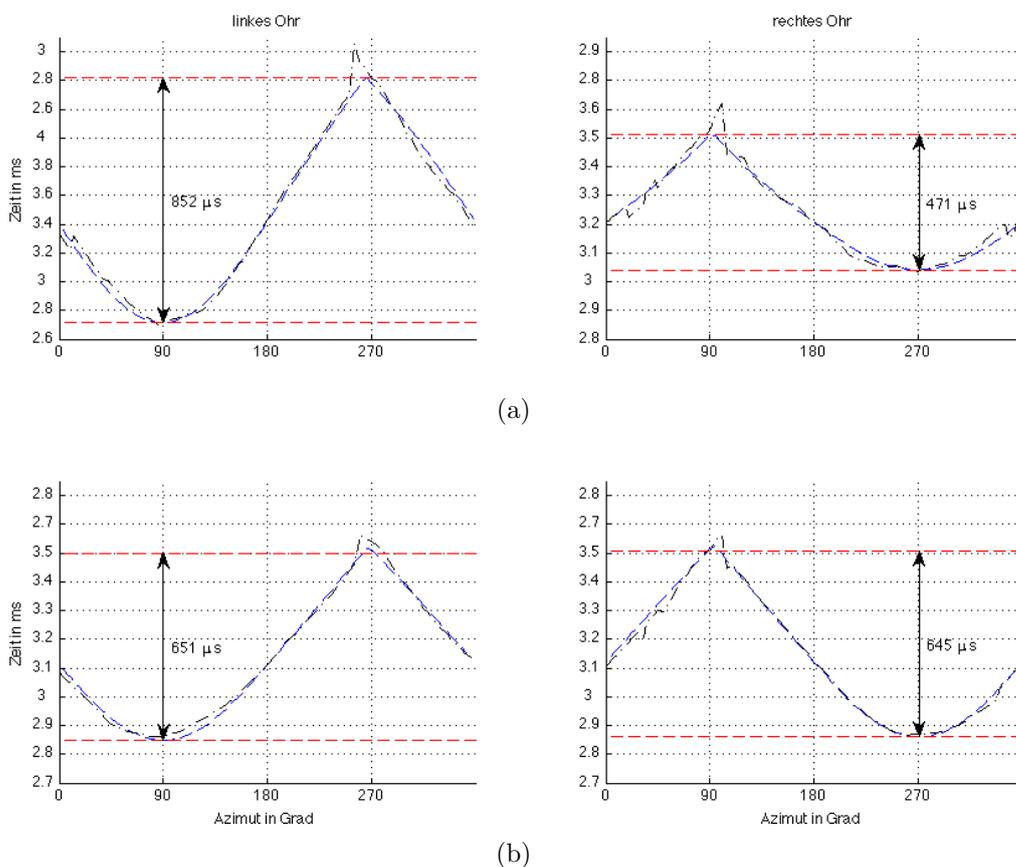
des Kopfradius im einfachen Modell. Wird diese Verschiebung auch im Modell (erweitertes Modell) zugelassen, führen die erhaltenen Modellparameter der beiden Ohren einer Versuchsperson zu plausiblen Ergebnissen (siehe Abb. 4.1 und Abb. 4.2b). Ein weiteres plausibles Ergebnis liefern die Mittelwerte der resultierenden Kugelradien über alle Versuchspersonen für linkes und rechtes Ohr, denn diese entsprechen annähernd dem mittleren Kopfradius erwachsener Menschen ( $r = 8.75\text{cm}$ ; Hartley und Fry, 1921).

Betrachtet man die Differenz der beiden Radien der beiden Modelle (Abbildung 4.3 oben) und die berechnete Verschiebung in y-Richtung beim erweiterten Modell (Abbildung 4.3 unten) wird klar, dass der fehlende Freiheitsgrad im einfachen Modell durch unterschiedliche Radien ausgeglichen wird. Bei sinistraler Verschiebung des Kopfes ( $y_M < 0$ ) wird die zeitliche Ausdehnung zwischen Minimum und Maximum der TOA am rechten Ohr geringer, der Radius beim einfachen Modell kleiner und somit die Differenz der Radien negativ.

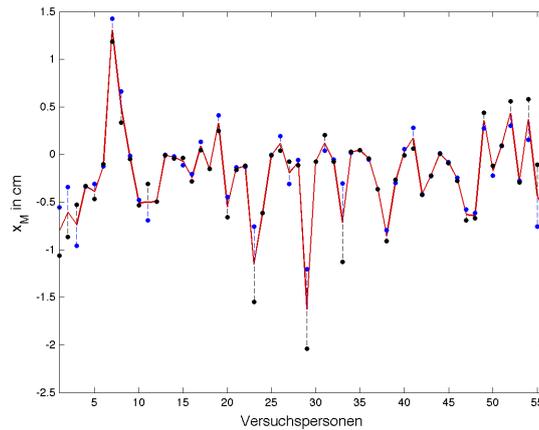


**Abbildung 4.3:** oben: Binaurale Differenz der Kugelradien (**grün:** einfaches TOA-Modell, **rot:** erweitertes Modell), unten: Verschiebung in y-Richtung im erweiterten Modell (**blau:** linkes Ohr, **schwarz:** rechtes Ohr, **rot:** Mittelwert der  $y_M$ -Werte beider Ohren)

Abbildung 4.4 illustriert das Problem der mediolateralen Verschiebung einer Versuchsperson genauer. Es wird die TOA zweier Versuchspersonen für linkes und rechtes Ohr in der Horizontalebene dargestellt. Die HRTFs der ersten Person weisen eine hohe mediolaterale Verschiebung auf ( $y_M = 2.9\text{cm}$  bei  $r = 9.1\text{cm}$ ), wodurch ein großer Unterschied im Abstand der Minima und Maxima der TOA zwischen den beiden Ohren entsteht. Die zweite Versuchsperson ist gut ausgerichtet ( $y_M = 0.07\text{cm}$  bei  $r = 9\text{cm}$ ) und es resultiert ein annähernd gleicher TOA-Verlauf für beide Ohren. Für die sagittale Verschiebung erhält man ein ähnliches Ergebnis wie für Radien und Verschiebung auf der Transversalachse, annähernd gleiche Ergebnisse für linkes und rechtes Ohr. Die erhaltenen Verschiebungen auf der Longitudinalachse sind näherungsweise Null und werden hier nicht weiter angegeben.



**Abbildung 4.4:** Auswirkung der Verschiebung des Kopfes einer Versuchsperson entlang der Transversalachse für eine schlecht positionierte Person (a) und eine gut positionierte Person (b)



**Abbildung 4.5:** Verschiebung in x-Richtung im erweiterten Modell (**1:** linkes Ohr, **2:** rechtes Ohr)

### 4.1.2 Modellierter Kopfradius

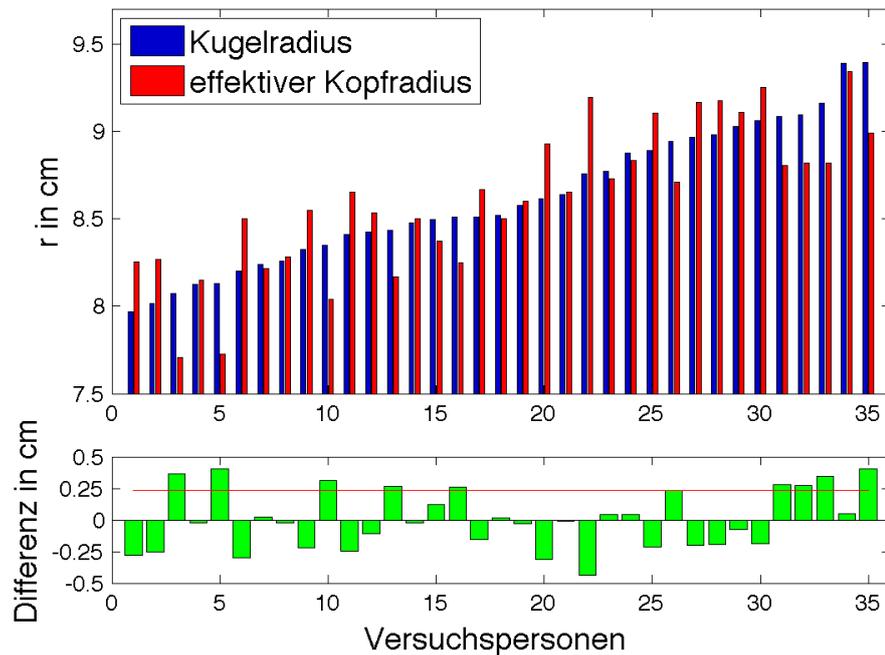
Die Betrachtung der Laufwege um den Kopf und die Tatsache, dass der modellierte Kugelradius von der maximalen TOA abhängt, führt zur Annahme, dass der Kugelradius zur Hälfte von der Kopfbreite und aus einer Kombination von Höhe und Tiefe zusammengesetzt ist.

Um diese Annahme zu überprüfen wurden anthropometrische Daten (Kopfbreite  $b$ , Kopftiefe  $t$  und Kopfhöhe  $h$ ) und der aus dem Modell erhaltene Kugelradius durch lineare Regression miteinander in Beziehung gesetzt. Die anthropometrischen Daten stammen aus der ARI-Datenbank für 10 Versuchspersonen und aus der CIPIC-Datenbank für 25 Versuchspersonen. Folgende lineare Regression approximiert am besten die modellierten Kugelradien:

$$\tilde{r} = \frac{0.53b + 0.22h + 0.24t}{2}, \quad (4.1)$$

und liefert einen Korrelationskoeffizienten von  $R = 0.83$  und einen Signifikanzwert von  $p < 10^{-9}$ . Abbildung 4.6 zeigt nach mittlerem Kugelradius aufsteigend geordnet die modellierten Kugelradien. Für den Pool der Versuchspersonen scheint die Annahme bestätigt zu sein. Auch für einzelne Versuchspersonen beträgt die Differenz maximal 5 mm.

Interessanterweise kommen [Algazi et al. \(2001\)](#) zu einem anderen Schluss. In Tabelle 4.1 werden die Koeffizienten verglichen. Die in [Algazi et al. \(2001\)](#) kaum relevante Kopfhöhe ist hier ebenso wichtig wie die Kopftiefe. Jedoch beträgt die Konstante in der hier vorgestellten Formel Null. Eine Konstante kann allerdings physikalisch oder physiologisch nicht begründet werden.



**Abbildung 4.6:** Vergleich der modellierten mittleren Kugelradien und der effektiven Kopfradien (oben) und Differenz und quadratischer Mittelwert der Differenz (unten)

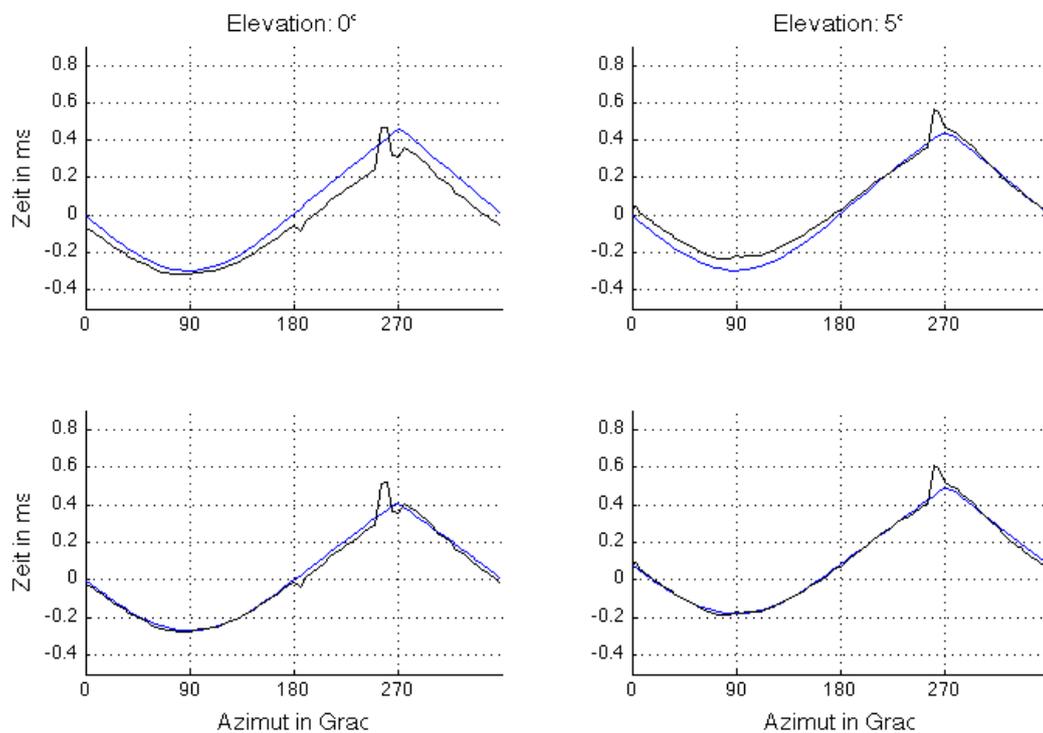
Koeffizient	Algazi	diese Arbeit
Kopfbreite $b$	0.51	0.53
Kopfhöhe $h$	0.019	0.22
Kopftiefe $t$	0.18	0.24
Konstante	3.2	0

**Tabelle 4.1:** Vergleich der Koeffizienten für den Kopfradius

### 4.1.3 Getrennte Messung verschiedener Elevationen

Eine weitere mögliche Fehlerquelle stellt die zeitlich und räumlich getrennt Messung von Elevationen bei  $10^\circ * k$  und  $5^\circ * (2k - 1)$  dar (vgl. 2.2). Abhängig vom Rotationspunkt der Versuchspersonen in der Messung kann ein Unterschied im Abstand der Lautsprecher zum Kopf zwischen diesen beiden Elevationsstufen entstehen. Werden separat Modelle an 5er und 10er Elevationen angepasst, kann eine Fehlschätzung vermieden werden. Es resultieren gleiche modellierte Kugelradien, aber unterschiedliche Kugelmittelpunkte. Abbildung 4.7 zeigt die TOAs einer Versuchsperson, bei der dieser Effekt deutlich ausgeprägt ist. In den oberen beiden Diagrammen wurde das Modell an alle Richtungen gleichzeitig angepasst. Dadurch stimmt die modellierte TOA mit keiner geschätzten TOA überein. Bei den unteren wurden die beiden Elevationsstufen getrennt voneinander modelliert und die Übereinstimmung ist

deutlich besser.



**Abbildung 4.7:** Auswirkung der zeitlich und räumlich getrennten Messung von 5er und 10er Elevationen (**oben:** Anpassung des Modells an alle Richtungen, **unten:** separate Anpassung an die beiden Elevationsgruppen)

#### 4.1.4 Erforderliche Anzahl an gemessenen Richtungen

Es stellt sich die Frage, wieviele Messpositionen für die TOA-Modellierung notwendig sind. Zur Beantwortung wurde erstens die Anzahl der zur Verfügung stehenden Richtungen für drei HRTF-Sets jeweils 10 Mal zufällig reduziert. Die Ergebnisse der daraus gemittelten Modellparameter (blau) sind in Abbildung 4.8 logarithmisch bezogen auf die Anzahl dargestellt. Bei bis zu ca. 250 – 350 Positionen sind die Abweichungen in den Parametern minimal, danach steigt die Varianz (rot) an.

Die Anzahl der Positionen könnte durch eine besondere Auswahl an Positionen weiter reduziert werden. Es könnten zum Beispiel einige Positionen auf der Horizontalebene und auf Sagittalebene genügen.

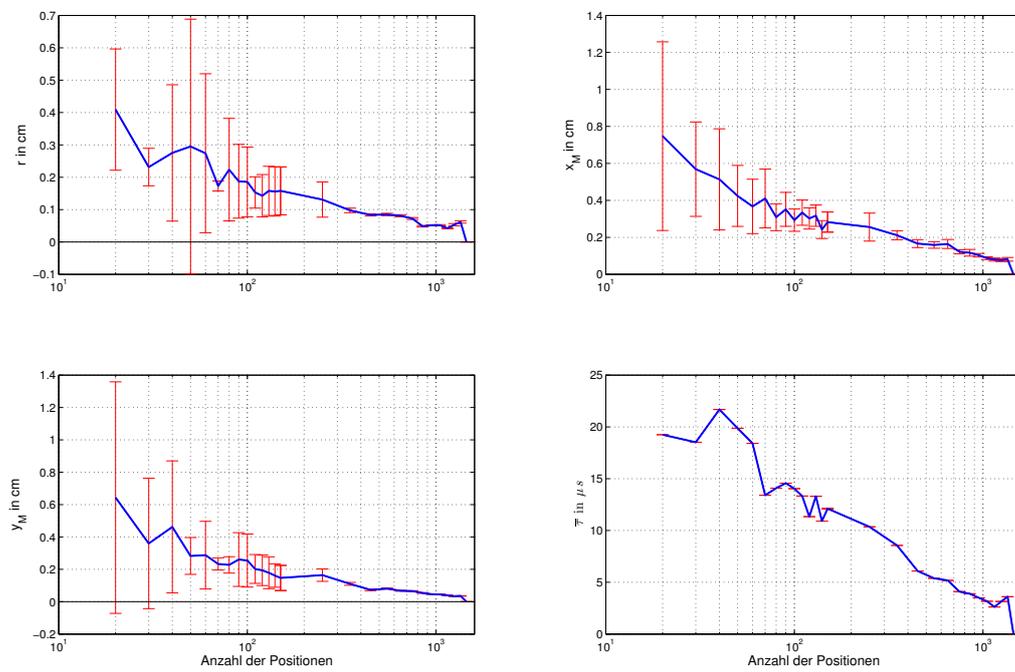


Abbildung 4.8: Zufällige Reduktion der Messpunkte (10 mal für 3 Versuchspersonen)

## 4.2 Vorläufige experimentelle Evaluierung

Die TOA kann verlustlos aus HRTFs entfernt und wieder aufgeprägt werden. Daher ist es nicht möglich die modellierte TOA direkt psychoakustisch zu evaluieren. Um die psychoakustische Relevanz der modellierten TOA zeigen zu können, werden originale HRTFs mit deren minimalphasig approximierten Varianten verglichen. Dabei werden aber zwei Fragestellungen in einem Experiment überprüft:

- Wieweit muss eine minimalphasige HRTF verzögert werden, um vom Original nicht unterschieden werden zu können?
- Wie wirkt sich die Vernachlässigung der Allpasskomponenten aus?

Diese überschneiden sich gewissermaßen. Es ist möglich, dass die frequenzunabhängige breitbandige TOA die perzeptive TOA, deren relevante Anteile vor allem von Allpasskomponenten unterhalb von  $1.5\text{kHz}$  abhängen können, unterschätzt. Daher muss zum Finden der perzeptiven TOA die im Experiment eingestellte TOA in einem Bereich um die modellierte TOA variiert werden, aber auch die Phasenspektren mitberücksichtigt werden. In der Literatur wurden folgende Stimuli in ITD- und Phasenexperimenten verwendet:

- Weißes Rauschen (50 Hz bis 18 kHz),
- Schmalbandiges Rauschen (600 bis 800 Hz), da die ITD hier die größte Änderung (simuliert auf einer schallharten Kugel) aufweist (Constan und Hartmann, 2003),
- komplexer Ton (50 Hz, 20/30 Teiltöne) in Phase und mit zufälliger Phase (Moore und Glasberg, 1989).

Es wird erwartet, dass es eine TOA gibt, bei der eine Versuchsperson keinen Unterschied zwischen Referenzsignal und Target hören kann und hier die Ergebnisse bei Chance-Performance liegen sollten. Dafür eignet sich breitbandiges Rauschen, denn hier sollte ein eventueller Phasenunterschied zwischen Referenz und Target, verursacht durch die vernachlässigten Allpasskomponenten, nicht wahrgenommen werden können. Nimmt der Abstand zur perzeptiven TOA zu, wird der Unterschied des Targets zum Referenzsignal immer deutlicher und die Versuchsperson kann diesen mit höherer Wahrscheinlichkeit erkennen.

Sind allerdings Phasenunterschiede zwischen Referenz und Target aufgrund der Beschaffenheit des Stimulus auch bei der perzeptiv korrekten TOA (ermittelt mit breitbandigem Rauschen) wahrnehmbar, wird die Trefferwahrscheinlichkeit der Versuchspersonen auch hier nicht im Zufallsbereich liegen. Dafür wird ein komplexer Ton entworfen, der am Targetkanal nach der HRTF-Filterung ein lineares Phasenspektrum aufweist. Als Merkmal wird sich

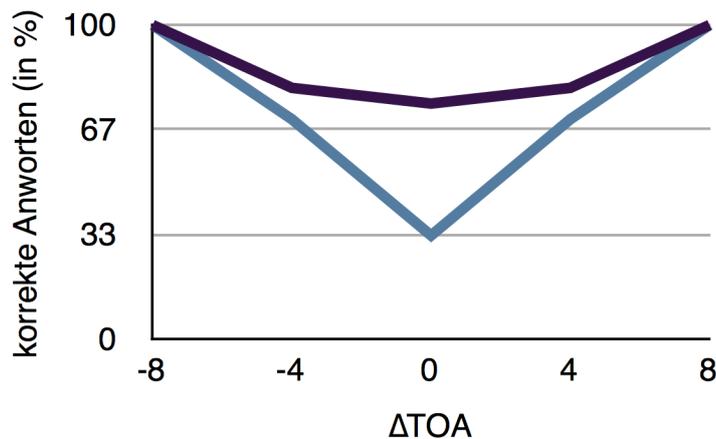


Abbildung 4.9: Erwartungshaltung des Experiments

hier ein Klangunterschied zwischen Referenz und Target bemerkbar machen aufgrund der unterschiedlichen Phasenlage der Teiltöne.

Nam et al. (2008) haben gezeigt, dass Allpass-Komponenten in Kombination mit den zugehörigen HRTF-Amplitudenspektren nicht hörbar sind. Dabei wurden allerdings originale HRTFs und approximierte HRTFs bei beiden Ohren gleichzeitig verglichen. Wird nur die ipsilaterale Richtung ersetzt, sollte bei korrekter TOA kein Unterschied hörbar sein. Beim kontralateralen Ohr ist ein anderes Ergebnis zu erwarten. Kuhn (1977) und Constan und Hartmann (2003) zeigten den Einfluss der Beugung auf die frequenzabhängige ITD. Die ITD, verursacht durch eine schallharte Kugel, ist bei tiefen Frequenzen größer. Constan und Hartmann (2003) zeigten, dass die Änderung der frequenzabhängigen ITD bei Schmalbandrauschen nicht hörbar ist. Da diese Frequenzabhängigkeit der TOA am kontralateralen Ohr in einem Unterschied der Phasenspektren von Referenz-HRTF und der minimalphasig approximierten HRTF resultiert, sollte der komplexe Ton als Stimulus zum Erkennen des Phasenunterschieds ausreichen. Daher kann auf das Schmalbandrauschen im Experiment verzichtet werden.

Es ist zu erwarten, dass ipsilaterale HRTFs durch minimalphasige Filter und zugehörige TOA approximierbar sein sollten und kein Unterschied hörbar sein sollte, da hier keine Beugung und somit keine frequenzabhängige TOA auftritt. Approximiert man die kontralaterale HRTF, so kann es durchaus sein, dass immer ein Unterschied hörbar sein wird, weil die Approximation mit dem Original im Phasenspektrum nicht übereinstimmt. Das heißt die Kurve wird hier vielleicht nie auf Chance-Performance hinabsinken. Weiters ist zu erwarten, dass die modellierte TOA die perzeptive TOA für kontralaterale Richtungen unterschätzt.

### 4.2.1 Versuchspersonen und Equipment

Die Evaluierung wurde mit drei Versuchspersonen im Alter von 22, 35 und 40 Jahren durchgeführt. Alle drei nahmen schon an früheren Experimenten teil und weisen einen Hörverlust (HV) von maximal  $20dB_{HL}$  (engl. hearing level, HL) in höchstens einem Frequenzband auf. Die Versuchspersonen wurde für den Zeitaufwand finanziell entschädigt. Tests mit weiteren 3 Personen sind am ARI geplant.

Die Experimente wurden in einer von zwei akustisch isolierten Kabinen durchgeführt, wobei keine Versuchsperson die Kabine wechselte. Die Durchführung des Experiments und die Signalverarbeitung erfolgen auf PC-Systemen mit dem am ARI eigens entwickelten Softwarepaket ExpSuite.<sup>2</sup> Die Stimuli werden über 24-bit D/A-Konverter (ADDA 2402, Digital Audio Denmark und E-MU 0404 USB, Creative Systems) bei einer Abtastfrequenz von 48 kHz ausgegeben. Die analogen Signale werden über Kopfhörerverstärker (HB6, TDT und G93, Lake People) und Kopfhörer (HDA 200, Sennheiser) wiedergegeben. Das Equipment wird über ein mit einem künstlichen Ohr (4153, Brüel & Kjaer) verbundenen Schallpegelmessgerät (2260, Brüel & Kjaer) kalibriert.

### 4.2.2 Stimuli

Drei Stimuli wurden verwendet. Der erste Stimulus war auf gleiche Energie normalisiertes weißes Gauß'sches Rauschen. Im Experiment wurde das selbe Rauschen in einer Testsequenz verwendet. Für jede Testsequenz wurde das Rauschsignal aus einem Pool von 100 Signalen zufällig ausgewählt.

Signal 2 war ein komplexer Ton (CT0) mit einer Grundfrequenz ( $f_0$ ) von 50 Hz und 29 Obertöne gleicher Amplitude:

$$x = \sum_{i=1}^{30} \cos(2\pi f_0 \cdot i \cdot n). \quad (4.2)$$

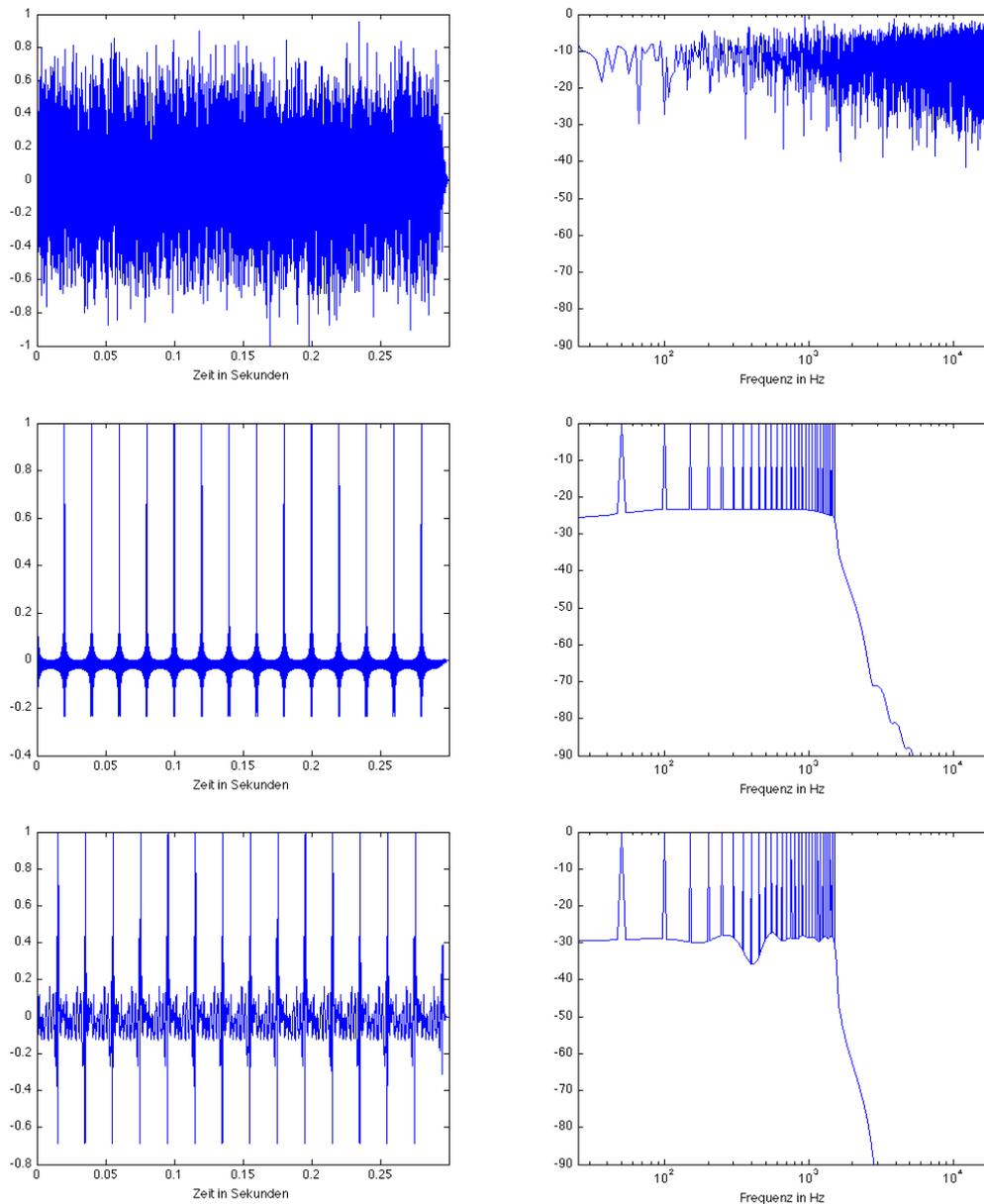
Stimulus 3 (CTHRTF) basiert auf CT0, die Phasenlage der einzelnen Teiltöne wurde aber an das Phasenspektrum der Referenz-HRTF so angepasst, dass nach der Faltung alle Teiltöne nullphasig zueinander lagen:

$$x = \sum_{i=1}^{30} \cos(2\pi f_0 \cdot i \cdot n - \varphi(f_0 i)). \quad (4.3)$$

Die Stimuli wurden mit einem kurzen Start- und langen Endfenster (Hann-Fenster mit 50 und 500 Samples Länge) gefiltert, um die Onset-ITD nicht zu unterdrücken. In Abbildung

2. <http://sourceforge.net/projects/expsuite/>

4.10 sind 3 typische Stimuli sowohl in der Zeitdarstellung als auch als Amplitudenspektrum dargestellt. Alle Stimuli wurden bei einer Abtastrate von 48 kHz generiert. Alle Stimuli wurden mit einer Länge von 300 ms (16000 Samples) entworfen. Der Schalldruckpegel (engl. sound pressure level, SPL) bei der Wiedergabe betrug ca.  $67dB_{SPL}$  und es wurde Level-roving in der Höhe von  $\pm 0,5dB$  verwendet.

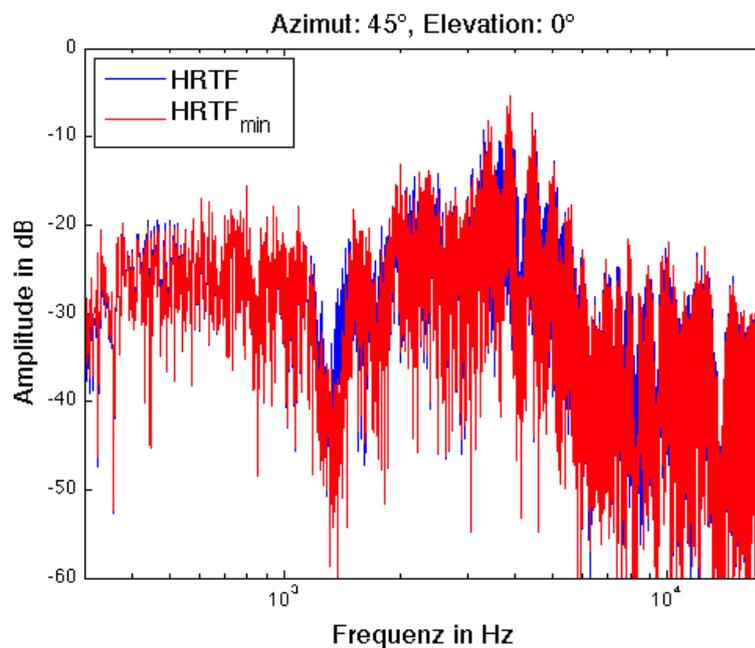


**Abbildung 4.10:** Zeitdarstellung und Amplitudenspektrum der 3 Stimuli (**oben:** breitbandiges gaußverteiltes Rauschen, **Mitte:** nullphasiger komplexer Klang, **unten:** an Phasenspektrum der Referenz-HRTF angepasster komplexer Klang)

Als Referenz-HRTFs wurden die nach dem Postprocessing (siehe Kapitel 2.2.3) entzerrten HRTFs mit einer Länge von 256 Samples verwendet. Als Target-HRTFs wurden mi-

nimalphasig approximierten Versionen der auf 4096 Punkte durch Zero-Padding erweiterten Referenz-HRTFs ohne TOA berechnet. Die Stimuli wurden mit den HRTFs gefaltet und im Falle der Target-HRTFs nach der Faltung mit Hilfe eines zirkulären Zeitshifts die TOA aufgeprägt. Um hier keine Artefakte am Beginn der Stimuli zu erzeugen, wurden zuvor die Stimulusignale um 300 Samples durch Zero-Padding erweitert.

Der Grund für die Länge der HRTFs im Experiment liegt in unterschiedlichen Amplitudenspektren nach der Faltung eines Stimulus mit originaler und approximierter HRTF kürzerer Länge als 4096 Samples. Eine Übertragungsfunktion und die minimalphasige Variante weisen nur bei gleicher Länge das gleiche Amplitudenspektrum auf. Werden beide durch Zero-Padding erweitert, weichen sie voneinander ab. Wird zum Beispiel weißes Gauß'sches Rauschen einer halben Sekunde Länge mit einer 256 Samples langen HRTF und deren minimalphasigen Approximation gefaltet, so ergeben sich unterschiedliche Langzeitamplitudenspektren (siehe Abb. 4.11). Dieser Unterschied ist im direkten Vergleich klar hörbar.



**Abbildung 4.11:** Unterschiedliche Amplitudenspektren nach der Faltung eines Signals mit einer HRTF und deren minimalphasigen Approximation (HRIR-Länge: 256 Samples)

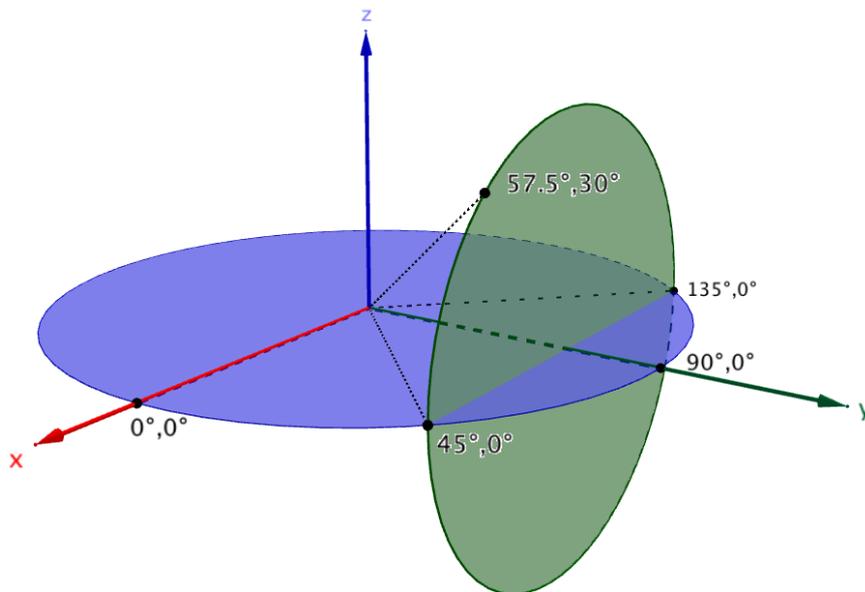
### 4.2.3 Experimentelle Bedingungen

In Tabelle 4.2 sind die verschiedenen Bedingungen aufgelistet. Die TOA wird sowohl ipsi- als auch kontralateral mit den drei vorgestellten Stimuli überprüft. Die unabhängige variable  $\Delta TOA$  wird im Bereich  $\pm 8$  Samples variiert, wobei die Werte  $\pm 6$  nicht evaluiert werden, um Zeit beim Experiment zu sparen.

Target	Stimuli	$\Delta TOA$	Positionen
ipsilateral kontralateral	Gauß'sches Rauschen CT0 CTHRTF	-8	
		-4	$0^\circ, 0^\circ$
		-2	$45^\circ, 0^\circ$
		0	$57.5^\circ, 30^\circ$
		2	$90^\circ, 0^\circ$
		4	$135^\circ, 0^\circ$
		8	

**Tabelle 4.2:** Experimentelle Faktoren und Faktorstufen

Das Experiment wird für fünf Richtungen durchgeführt, wobei vier auf der Horizontalebene und drei auf einer Sagittalebene liegen (siehe Abb. 4.12). Daraus ergibt sich eine Dauer von 12 bis 14 Stunden pro Versuchsperson, abhängig von der Reaktionszeit jeder Versuchsperson.



**Abbildung 4.12:** Anordnung der Richtungen

#### 4.2.4 Prozedur

Das Experiment wird als 3-Interval/3-Alternative Forced Choice Methode (3AFC) durchgeführt. Dabei besteht eine Testsequenz aus drei binauralen Stimuli, welche über Kopfhörer präsentiert werden, wobei einer der drei Stimuli an einem Ohr den Target enthält. Der Target beinhaltet im Gegensatz zur Referenz eine mit der TOA verzögerte minimalphasig approximierten HRTF. Die Signale wurden immer binaural präsentiert, wobei je nach Bedingung die ipsi- oder kontralaterale HRTF approximiert (Target) wurde. Die jeweils andere HRTF

wurde nicht verändert (Referenz). Die Versuchsperson wählt mittels Joypad das Targetintervall. Die Versuchspersonen bekommen nach jedem Item eine Rückmeldung und bei falscher Antwort das Targetintervall angezeigt.

Zu Beginn wurde der Versuchsperson erklärt, wie das Experiment abläuft, auf welche Merkmale sie hören kann (nicht auf Lautheit!) und mit welchen Stimuli sie getestet wird. Beim ersten Experiment wird der Raum ( $\Delta TOA$  geschätzt  $\pm 12$  Samples) der TOA um den geschätzten Wert manuell abgetastet, um zu überprüfen ob der perzeptive TOA-Wert im experimentellen Rahmen liegt. Dafür werden der Versuchsperson Testsequenzen vorgespielt, wobei die TOA mit  $\Delta TOA$  im Target variiert wird. Die Versuchsperson kennt die Position des Targets und wird befragt ob sie den Unterschied hört, um auf den Target sensibilisiert zu werden. Dieser Test wird mit Breitbandrauschen und einem an die Referenz-HRTF angepassten Tonkomplex durchgeführt. Anschließend wird der Test wiederholt, jedoch wird zusätzlich die Position des Targets in der Testsequenz variiert. Die Versuchsperson nennt das Intervall, in dem sie einen Unterschied hört. Es wird der TOA-Bereich gesucht, in dem die Versuchsperson nur raten kann. Das entspricht einer manuell durchgeführten adaptiven Prozedur. Erwähnenswert ist ein hoher Lerneffekt der Versuchspersonen. Hier liegen alle Versuchspersonen im Bereich von zwei Samples um die modellierte TOA.

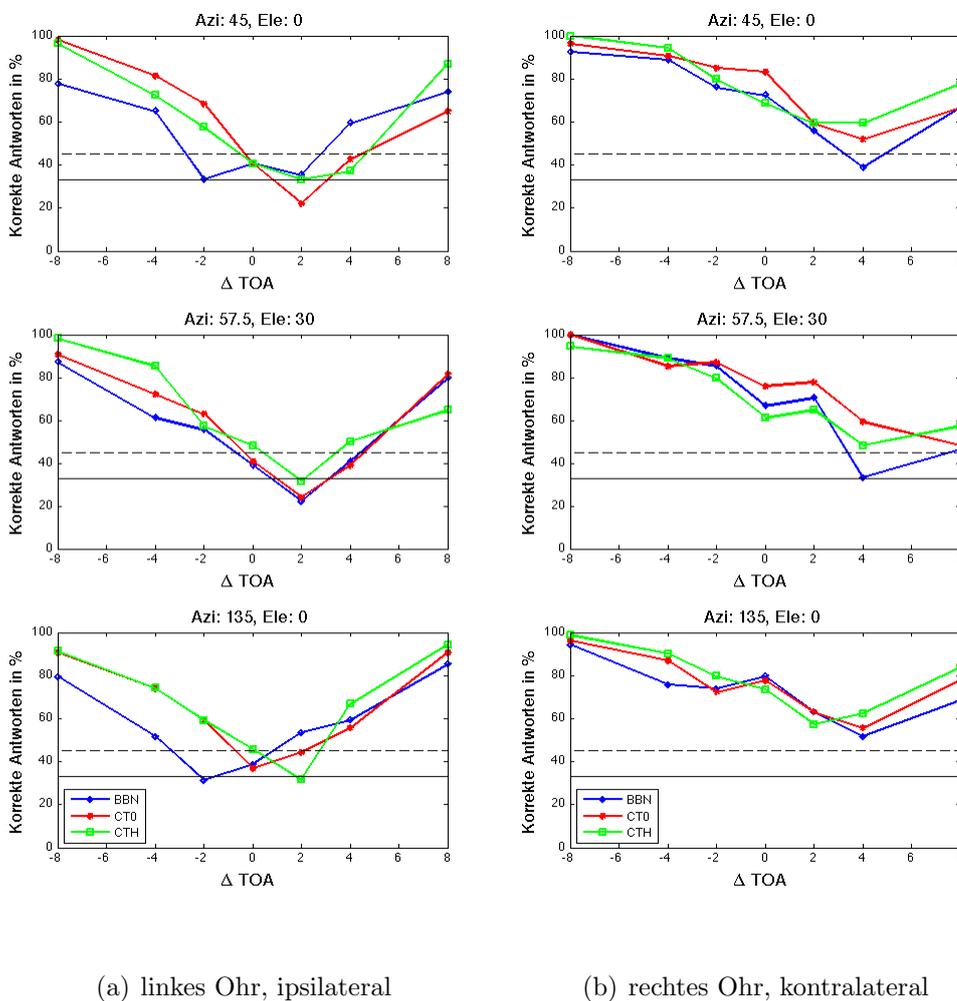
Im nächsten Schritt werden automatisierte Vortests für alle drei Stimuli durchgeführt. Jeder Stimulus wird in einem Block für die Richtung  $\phi = 0^\circ$ ,  $\theta = 0^\circ$  bei  $\Delta TOA -8,0,8$  Samples mit 15 Wiederholungen und ein zweites Mal mit  $\Delta TOA -8,-4,0,4,8$  mit 15 Wiederholungen und Target an beiden Ohren getestet (entspricht 90 bzw. 150 Testsequenzen pro Block). Anhand dieser Resultate wird entschieden ob eine Versuchsperson für die Studie geeignet ist oder nicht. Liegt ihre Performance bei  $\Delta TOA -8$  und  $8$  nicht mindestens über 80 Prozent, wird sie nicht weiter teilnehmen.

Daraufhin folgt das Hauptexperiment. Die drei verschiedenen Stimuli werden getrennt voneinander getestet und auf zwei Blöcke mit neun Wiederholungen pro Testsequenz aufgeteilt. Innerhalb eines Blockes werden die fünf Richtungen gruppiert und pro Richtung die einzelnen Testsequenzen randomisiert. Die Reihenfolge der Richtungen wird ebenfalls zufällig ausgewählt. Nach jeder Richtung werden der Versuchsperson Testsequenzen der nächsten Testrichtung präsentiert, um ihr den Übergang zu erleichtern.

### 4.2.5 Resultate

In Abbildung 4.14 und 4.13 sind die über die drei Versuchspersonen gemittelten vorläufigen Ergebnisse dargestellt. Über die variierte TOA ist die Trefferquote der Versuchspersonen in Prozent dargestellt. Bei 33% (durchgezogene schwarze Linie) liegen die Antworten der Versuchspersonen im Zufallsbereich. Die Signifikanzgrenze von  $p < 5\%$  liegt bei 45% (strichlierte

schwarze Linie) der Trefferquote. Die einzelnen Resultate der Versuchspersonen sind in Anhang B abgebildet. Abbildung 4.14 zeigt die Ergebnisse in der Horizontalebene für ipsi- und kontralateralem Target am linken Ohr, wobei die kontralateralen Richtungen am rechten Ohr evaluiert wurden und hier für eine bessere Darstellung für ein Ohr zusammengefasst wurden. Abbildung 4.13 zeigt die Ergebnisse für ipsi- (a) und kontralateralen (b) Target der evaluierten Richtungen auf einer Sagittalebene. Für alle Bedingungen mit ipsilateralem Target liegt die perzeptive TOA im Bereich von zwei Samples um die modellierte TOA. Bei kontralateralem Target ist der Unterschied zwischen perzeptiver und modellierter TOA größer und liegt bei ca. 4 Samples.



**Abbildung 4.13:** Gemittelte Ergebnisse der experimentellen Evaluierung

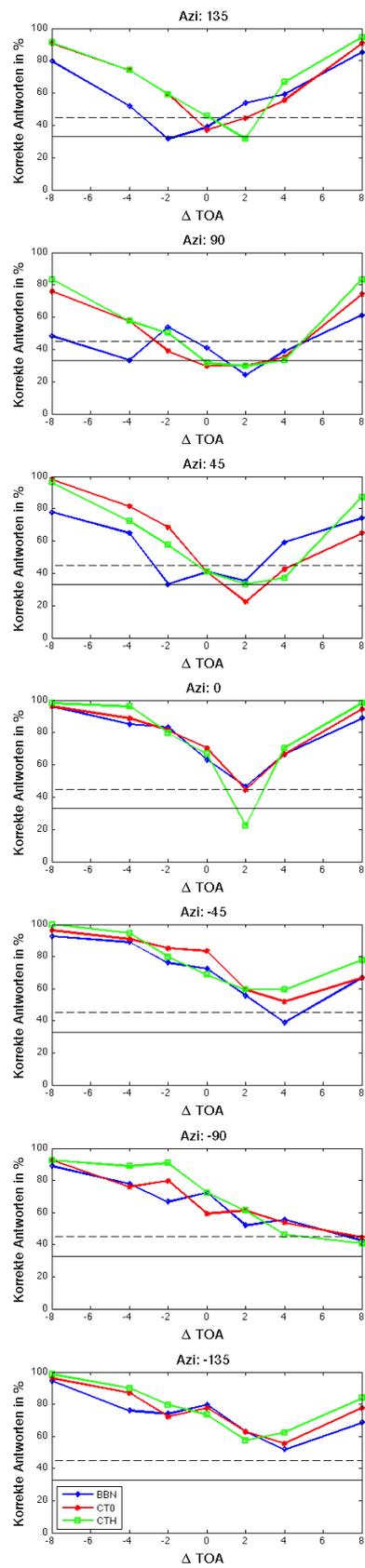


Abbildung 4.14: Gemittelte Ergebnisse der experimentellen Evaluierung

### 4.3 Diskussion

Die vorgestellte Methode zur TOA-Bestimmung aus individuell gemessenen HRTFs erweist sich als sehr robust. Unter Ausnutzung der Glattheit auf Horizontalebene und der geringen Varianz auf Sagittalebene der TOA werden Ausreißer in der TOA-Schätzung erkannt. Mithilfe des einfachen TOA-Modells können Startwerte für das erweiterte Modell gefunden werden, welches wie gezeigt alle nötigen Freiheitsgrade aufweist um plausible Ergebnisse zu liefern. Somit können TOA-Schätzungen, welche als Ausreißer erkannt wurden, über das Modell korrekt rekonstruiert und Richtungen zwischen den Messpunkten beliebig interpoliert werden. Die Parameter des Modells liefern ebenfalls Auskunft über die Ausrichtung einer Versuchsperson während der HRTF-Messung. Am bedeutsamsten ist die Auswirkung der Verschiebung des Kopfes einer Versuchsperson entlang der Transversalachse während der HRTF-Messung. Hier wirken sich schon geringe Abweichungen  $y_M < 1\text{cm}$  in einer Asymmetrie der TOA zwischen dem linken und dem rechten Ohr aus. Da, abgesehen von Kunstköpfen, Objekte bei der HRTF-Messung nicht genau ausgerichtet werden können, scheint die Korrektur der TOA eine Notwendigkeit zu sein. Da Köpfe realer Versuchspersonen immer eine Asymmetrie aufweisen und der Kopfmittelpunkt nicht klar definiert werden kann, ist es gar nicht möglich diese im Messaufbau korrekt zu positionieren.

Die modellierte TOA scheint die perzeptive TOA generell leicht zu unterschätzen. Da die modellierte TOA die frequenzabhängige Gruppenlaufzeit, verursacht durch Allpasskomponenten, nicht berücksichtigt, war dieses Ergebnis zu erwarten. Für ipsilaterale Richtungen stimmen modellierte und perzeptive TOA bis auf 2 Samples überein. Bei kontralateralen Richtungen  $\phi = 45^\circ$  beträgt der Unterschied bereits ca. 4 Samples. Ein Ergebnis, welches wieder auf perzeptiv relevante Allpasskomponenten zurückzuführen ist. Bei größtmöglicher Abschattung durch den Kopf  $\phi = 90^\circ$  weichen modellierte und perzeptive TOA bereits um mindestens 8 Samples voneinander ab. Da die perzeptive TOA generell um mindestens 2 Samples unterschätzt wird, handelt es sich hierbei um eine richtungsunabhängige Abweichung. Deshalb würde die ITD in der Nähe der Medianebene korrekt in den TOAs abgebildet werden. Die ITD-Abweichung bei  $\pm 45^\circ$  würde bei ca.  $40\mu\text{s}$  und somit für diese laterale Auslenkung nicht im wahrnehmbaren Bereich liegen.

Der erwartete hörbare Unterschied der Phasenspektren zwischen empirisch gemessenen und minimalphasig approximierten HRTFs zeigt sich für kontralaterale Targets aus Richtungen auf der Sagittalebene  $\Phi = 45^\circ$ . Die Trefferquote für die komplexen Töne liegt hier im signifikanten Bereich. Für diese Richtungen gibt es keine TOA, mit deren Verzögerung eine minimalphasige HRTF von der originalen HRTF nicht unterschieden werden kann.

# Kapitel 5

## Anwendungen

Die modellierte TOA kann für verschiedene Anwendungen genutzt werden. Auch die Bestimmung der genauen Position und Ausrichtung einer Person bei der HRTF-Messung kann verwendet werden, um HRTFs nach der Messung zu korrigieren. Die Anwendungsvielfalt wird hier und im Ausblick in Kapitel 6 aufgezeigt.

### 5.1 Richtungsabhängige Fensterung

Erstens ist es möglich die genaue zeitliche Ausrichtung aller HRTFs robust zu bestimmen. Durch dieses Wissen kann eine automatische richtungsabhängige Fensterung durchgeführt werden, welche Raumreflexionen aus den Messdaten ausblenden und die Länge der HRIRs kürzen soll. Diese richtungsabhängige Fensterung bietet den Vorteil, dass die spektralen Informationen aller HRTFs gleich lang erfasst werden. Denn bei der Fensterung aller HRTFs durch ein räumlich statisches Fenster würden HRTFs mit großer TOA gegenüber HRTFs mit kleiner TOA früher abgeschnitten werden. Abbildung 5.1 zeigt ETCs gemessener HRIRs (links) und gefensterter HRIRs (rechts). Das Fenster weist eine Länge von 5.3 ms (256 Samples bei 48 kHz Abtastrate) auf, beginnt 32 Samples vor der geschätzten TOA mit einem halben ansteigenden Hann-Fenster und endet mit einem halben fallenden Hann-Fenster 4-facher Länge. Dadurch wird die TOA aus allen HRIRs entfernt (vgl. Abbildung 5.1 links und rechts). Der Grund für die geringfügige Abweichung von Abbildung 5.1 (Elevation 20° rechts) von einer exakten senkrechten schwarzen Linie bei ca. 40 Samples liegt darin, dass die TOA nicht über das Maximum einer HRIR definiert wurde (siehe Kapitel 3). Die ETCs der Horizontalebene bei 40° und 60° lassen den Schluss zu, dass hier die Minimalphasenkreuzkorrelation und die Maximalwertdetektion zum selben Ergebnis führen. Die Ausrichtung des Fensters wird umso wichtiger, je kürzer das Fenster gewählt wird.

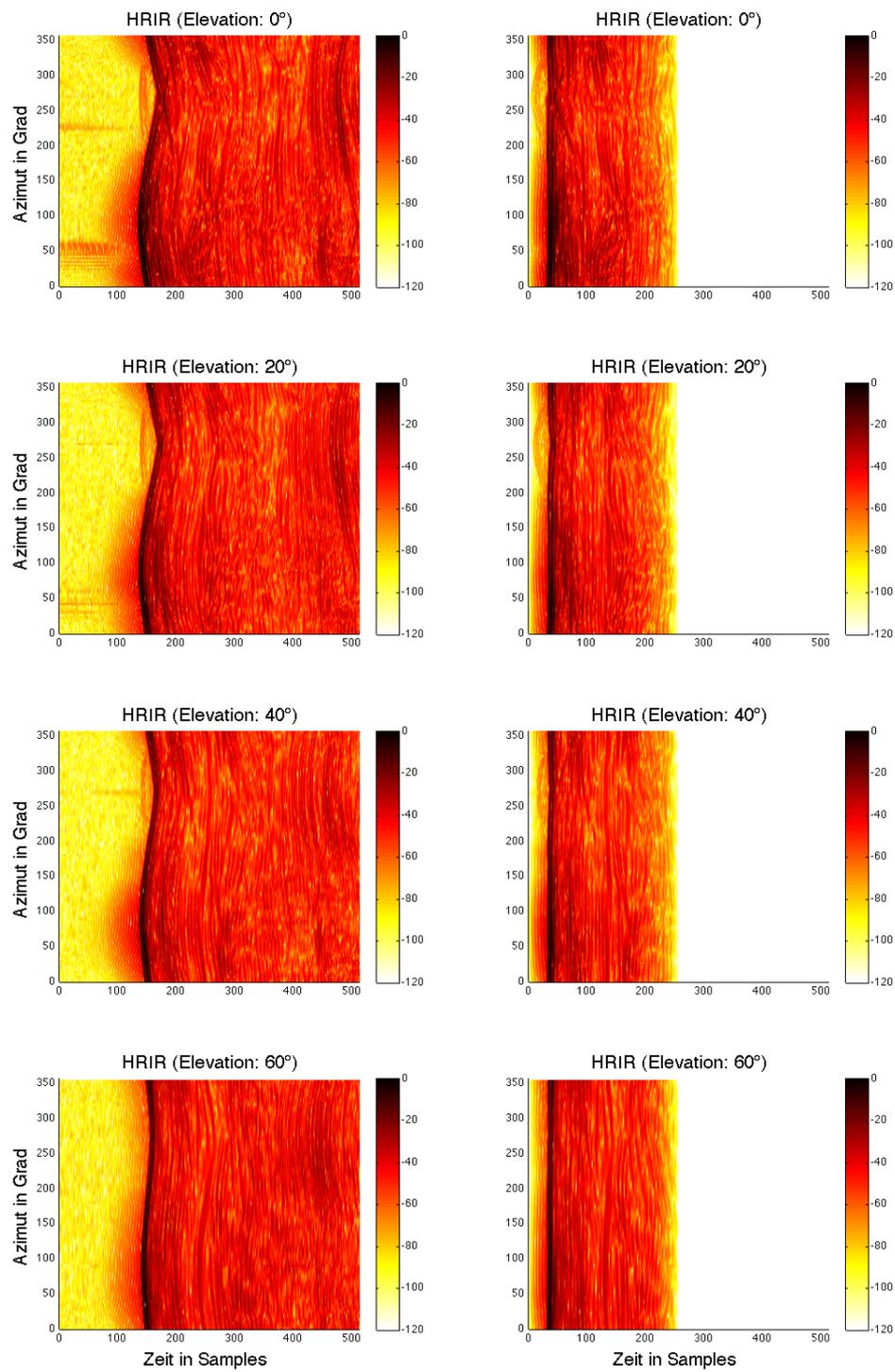
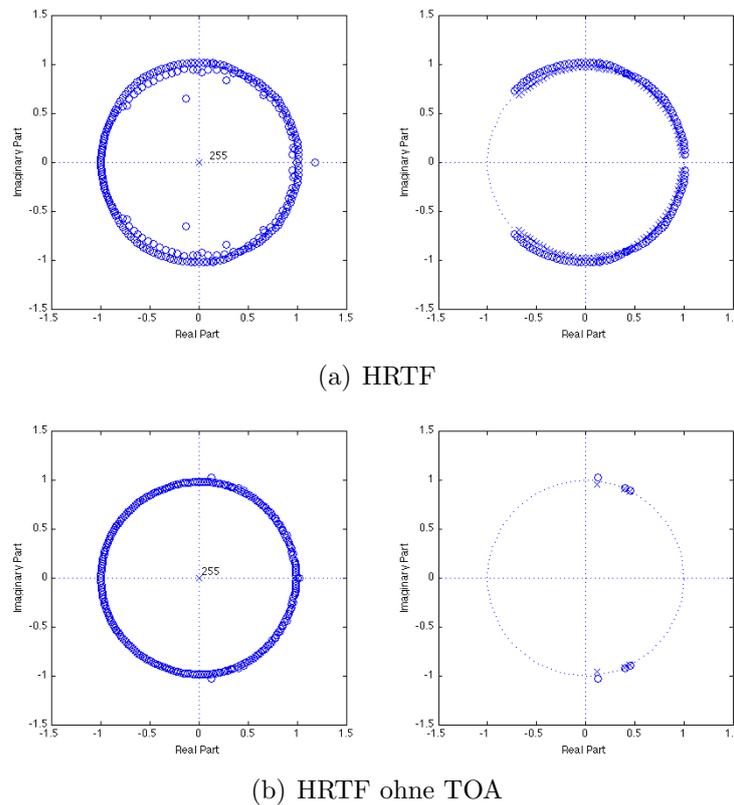


Abbildung 5.1: ETCs nach Entfernung der TOA und Fensterung von HRIIRs für vier Elevationen)

## 5.2 Phasenanalyse

Die TOA-bereinigten HRTFs können die Grundlage für die Analyse der Phasenspektren in HRTFs bilden. In Abbildung 5.2 links sind die Nullstellen für die originale und die TOA-reduzierte HRTF in der  $z$ -Ebene dargestellt.



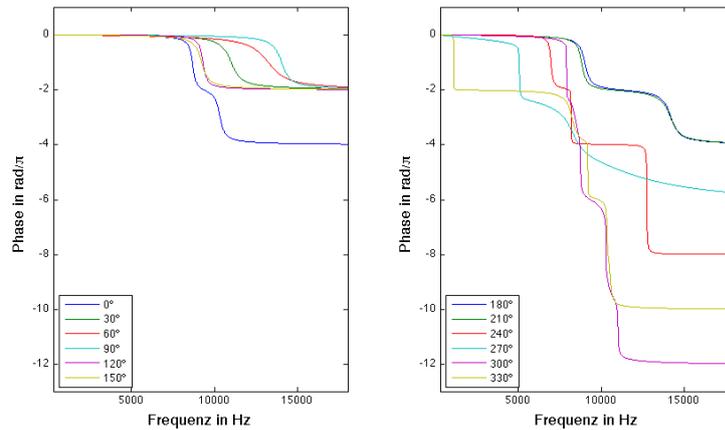
**Abbildung 5.2:** Pol-Nullstellendiagramm in der  $z$ -Ebene (**links:** gesamte HRIR, **rechts:** enthaltene Allpasskomponenten)

Zum Beispiel können die verbleibenden Allpässe extrahiert werden, weil sie keine breitbandige Gruppenlaufzeit mehr beinhalten. Dafür werden alle Nullstellen außerhalb des Einheitskreises mit einem Winkel von

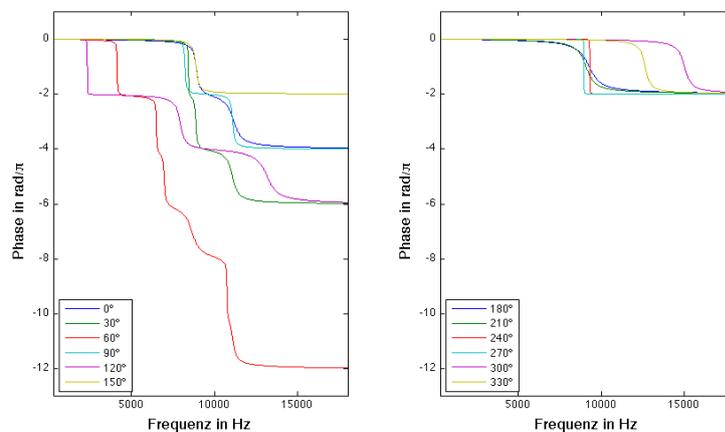
$$\frac{1}{80}\pi < |\angle\{NS\}| < \frac{3}{4}\pi$$

(das entspricht bei einer Abtastrate von 48 kHz einem Frequenzbereich von 300 Hz bis 18 kHz) in den Einheitskreis gespiegelt und an die ursprüngliche Position eine Polstelle gesetzt (siehe Abb. 5.2 rechts). Im Gegensatz zu originalen HRTFs ist die Anzahl der enthaltenen Allpässe zweiter Ordnung deutlich reduziert und somit das Rückrechnen von Pol-Nullstellen auf Polynome numerisch durchführbar.

Mithilfe dieser mathematischen Darstellung wäre eine direkte Implementation der TOA-reduzierten Excessphase als IIR-Filter möglich. Abbildung 5.3 und 5.4 zeigen die Phasenspektren dieser Allpässe für 12 Positionen in der Horizontalebene für linkes und rechtes Ohr.



**Abbildung 5.3:** Phasenspektren der extrahierten Allpässe (Excessphase) für das linke Ohr (**links:** ipsilaterale Richtungen, **rechts:** kontralaterale Richtungen)



**Abbildung 5.4:** Phasenspektren der extrahierten Allpässe (Excessphase) für das rechte Ohr (**links:** kontralaterale Richtungen, **rechts:** ipsilaterale Richtungen)

Es ist deutlich erkennbar, dass alle Richtungen mindestens einen Allpass 2-ter Ordnung aufweisen, was darauf deutet, dass diese HRTFs keine Minimalphasensysteme sind. Das gilt vor allem für kontralaterale Richtungen (siehe Abb. 5.3 rechts und Abb. 5.4 links). Weiters ist zu erkennen, dass die Phasenspektren bis auf die Phasendrehungen von  $360^\circ$  keine Steigung aufweisen und dass die Allpässe nicht symmetrisch sind, also die Phasenspektren zwischen dem linken und dem rechten Ohr unterschiedlich sind.

## 5.3 Bandbreitenerweiterung

Die am ARI gemessenen HRTFs weisen eine Hochpass-Charakteristik auf, weil das Anregungssignal erst ab 300 Hz die volle Energie erreicht und die Lautsprecher unterhalb von 300 Hz das Signal deutlich abschwächen. Da diese Frequenzen einen substanziellen Effekt auf die Natürlichkeit von Audiosignalen, vor allem bei Musik (Moore und Tan, 2003), haben, sollte dieser Bereich in den HRTFs angehoben werden. Zur Anhebung des Amplitudenspektrums unterhalb der unteren Grenzfrequenz der Lautsprecher wird dieses durch den Mittelwert des tiefsten gemessenen Frequenzbandes extrapoliert. Um einen glatten Verlauf zu erhalten, wird ein Crossfade zum gemessenen Amplitudenspektrum durchgeführt. Diese Erweiterung des Amplitudenspektrums entspricht im wesentlichen der Bandbreitenerweiterung aus Zotkin et al. (2006) für die reziproke HRTF-Messung, jedoch wird hier nicht mit einer Simulation, sondern mit einem konstanten Wert extrapoliert und es wird nicht ein linearer Crossfade, sondern eine Sigmoid-Funktion gewählt, um einen glatten Übergang zu gewährleisten. Bei TOA-bereinigten HRTFs ist diese Bandbreitenerweiterung besonders leicht durchzuführen, weil das Phasenspektrum im relevanten Bereich auf 0 gesetzt werden kann und nach erneuter Verzögerung auch in diesem Frequenzbereich die TOA und somit auch die ITD aufgeprägt werden kann.

$$|HRTF(\omega)| = \begin{cases} \bar{H}_{\omega_1, \omega_2} & \text{für } \omega < \omega_1 \\ \bar{H}_{\omega_1, \omega_2} + \frac{|H(\omega)| - \bar{H}_{\omega_1, \omega_2}}{1 + e^{-6 \frac{(\omega - \omega_1)}{\omega_2 - \omega_1}}} & \text{für } \omega_1 \leq \omega \leq \omega_2 \\ |H(\omega)| & \text{für } \omega_2 \leq \omega \end{cases} \quad (5.1)$$

$$\bar{H}_{\omega_1, \omega_2} = \frac{1}{\omega_2 - \omega_1} \sum_{\omega=\omega_1}^{\omega_2} |H(\omega)|$$

Abbildung 5.5 zeigt den linearen und den verwendeten Crossfade (links) und originalen und korrigierten Amplitudenfrequenzgang für eine HRTF. Diese Methode liefert einen Amplitudenfrequenzgang, der oberhalb der Bandbreitenerweiterung den selben Frequenzgang liefert wie das bisherige Postprocessing (siehe Abb. 5.6 oben), jedoch ohne kepstrale Glättung.

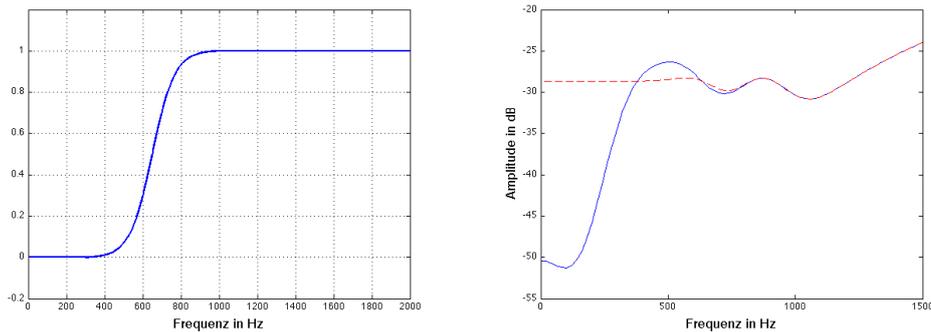


Abbildung 5.5: Crossfade und Bandbreitenerweiterung

## 5.4 Nachbearbeitung nach der HRTF-Messung

Ziel des Post-Processings ist die Erstellung eines HRTF-Datensatzes, der sowohl eine genaue Lokalisation von Schallquellen als auch einen natürlichen Klang nach der Filterung ermöglicht. Alle zuvor erwähnten Methoden können dafür herangezogen werden. Zuerst werden die HRTFs mithilfe der TOA-Schätzung gefenstert, die TOAs entfernt und alle Richtungen um die Länge des Startfensters zirkulär in der Zeit gedreht. Somit weisen alle Richtungen ein flaches Phasenspektrum, bis auf Phasendrehungen von  $360^\circ$  an den Positionen von Allpässen, auf. Das Amplitudenspektrum kann durch den vorgestellten Crossfade und das Phasenspektrum durch Nullsetzen erweitert werden. Weicht der ermittelte Modellparameter  $y_M$  von 0 ab, kann diese Fehlpositionierung durch Korrektur dieses Parameters behoben werden. Da der Radius unangetastet bleibt, wird die ITD nicht verändert. Es müssen aber alle Richtungen geringfügig angepasst werden.

Während der psychoakustischen Experimente erwähnte eine Versuchsperson, dass sie die Stimuli, positioniert in Richtung  $0^\circ, 0^\circ$ , eigentlich von rechts wahrnimmt. Es stellte sich heraus, dass die linke HRIR eine um 1 Sample höhere TOA aufwies als die rechte. Ein Umstand, der verdeutlicht, wie wichtig diese Korrektur für die virtuelle Akustik mit individuellen HRTFs ist, sei sie auch noch so gering.

In Abbildung 5.6 (unten) sind die Phasenspektren nach altem und neuem Postprocessing und deren minimalphasigen Varianten dargestellt. Es ist zu erkennen, dass sich diese zwischen 2 und 10 kHz nur äußerst gering unterscheiden. Ein weiterer Beleg für die messtechnisch korrekt ermittelte und entfernte Signallaufzeit. Unterhalb von 2 kHz liegt der Phasenunterschied sowohl zwischen altem und neuem Postprocessing als auch zwischen originalem und approximiertem Phasenspektrum durchaus im wahrnehmbaren Bereich.

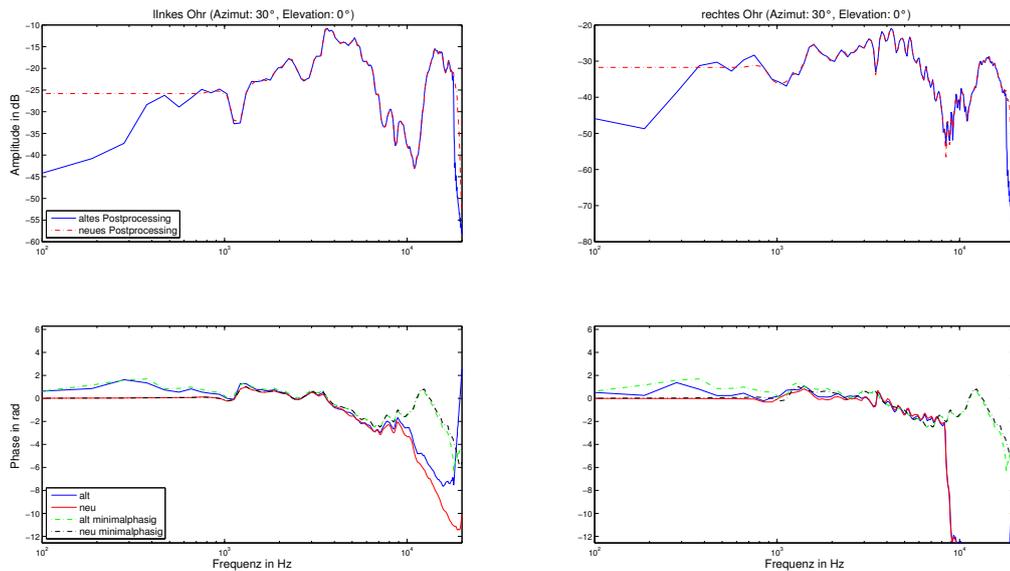


Abbildung 5.6: Amplituden- und Phasenspektren nach dem alten und neuen Postprocessing

## 5.5 Datenreduktion

Soll die Datenmenge weiter reduziert werden, könnte die minimalphasige Approximation der HRTFs verwendet werden, wenn die Allpässe für den Verwendungszweck vernachlässigbar sind. Hierfür werden aus den Amplitudenspektren über die Hilberttransformation die zugehörigen Minimalphasenspektren berechnet. Anschließend können diese durch einseitige Fensterung weiter verkürzt werden. Die Auswirkung auf die Lokalisation ist minimal (vgl. Kapitel 2), jedoch ist eine Veränderung auf den Klang nicht auszuschließen (Untersuchung siehe Kapitel 3) und hier nur als mögliche Methode zur Filtergenerierung für recheneffiziente Echtzeitimplementationen angeführt. Der Vorteil liegt darin, das Amplitudenspektrum beliebig verändern zu können, ohne selbst das Phasenspektrum manipulieren zu müssen.

## 5.6 Aufbereitung für räumliche Transformationen

Zaar (2011) untersuchte die Zerlegung von HRTF-Amplituden- und HRTF-Phasenspektren in Kugelflächenfunktionen. Die räumliche Phase muss hierfür allerdings einen glatten Verlauf aufweisen. Um die Zerlegung zu ermöglichen wurde ein räumlicher Unwrap-Algorithmus entwickelt. Es ist davon auszugehen, dass die räumliche Phase nach der TOA-Entfernung ohne Phasenabwicklung bis zu einer bestimmten Frequenz einen glatten Verlauf ohne Phasensprünge zwischen  $-\pi$  und  $\pi$  aufweist. Abbildung 5.7 und 5.8 zeigen die räumlichen Pha-

senspektren für die originalen HRTFs (links) und für die um die TOA zirkulär in der Zeit gedrehten HRTFs (rechts). Nach der TOA-Entfernung weisen diese einen deutlich flacheren Verlauf und bis ca. 3 kHz keine Unstetigkeiten auf.

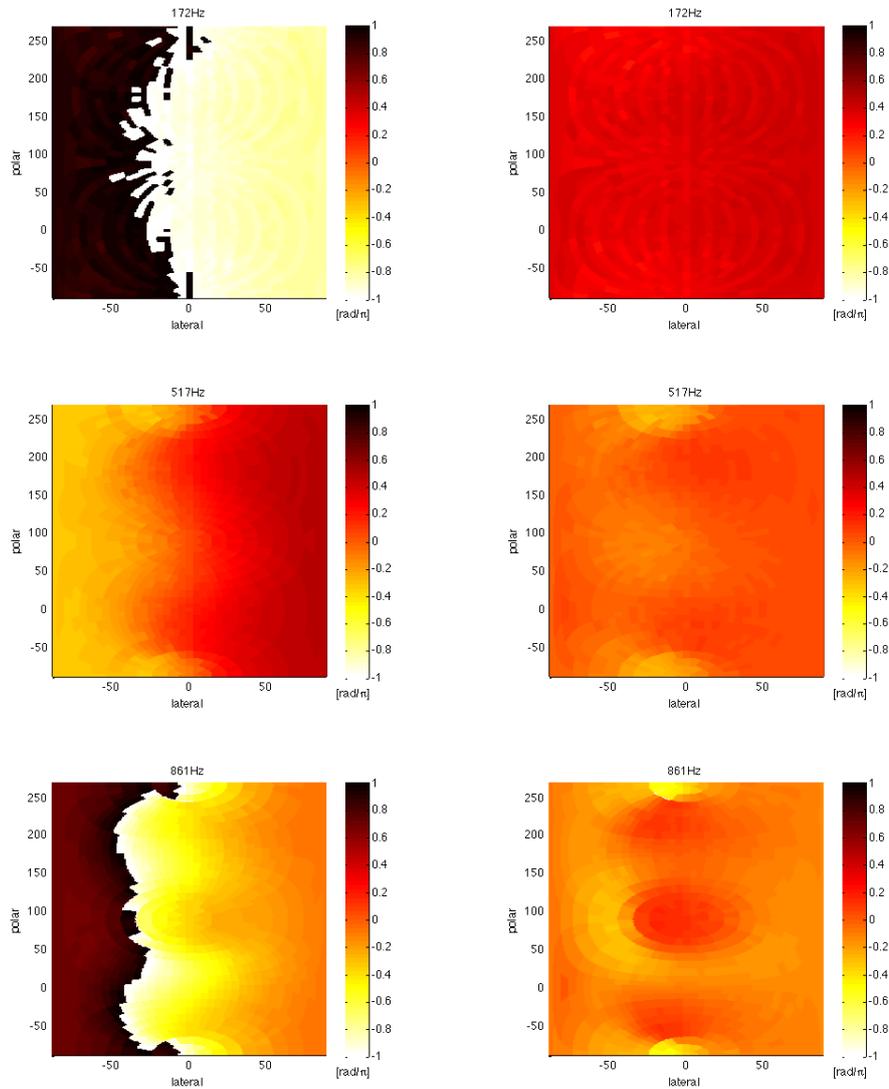


Abbildung 5.7: Räumliche Phasenspektren bei (**links:** originale HRTFs; **rechts:** TOA-reduzierte HRTFs)

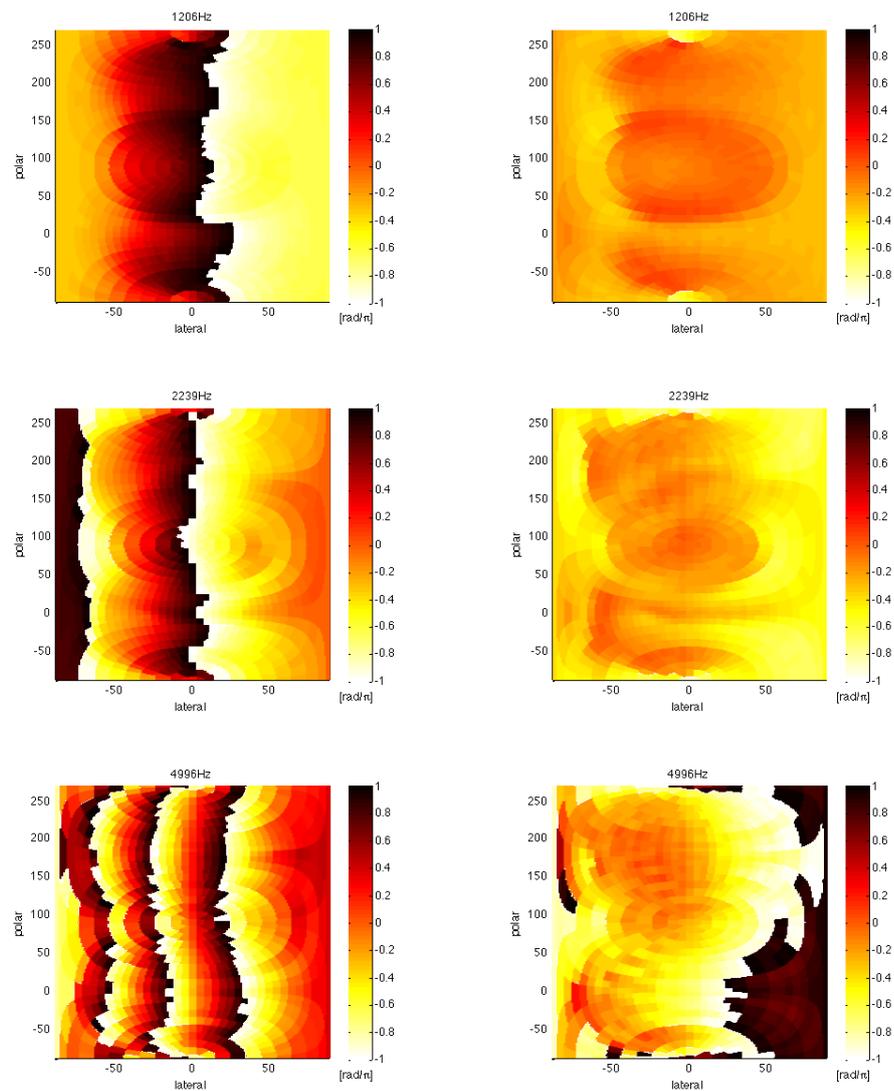


Abbildung 5.8: Räumliche Phasenspektren bei (links: originale HRTFs; rechts: TOA-reduzierte HRTFs)



# Kapitel 6

## Zusammenfassung

Ein automatisches Verfahren zur Bestimmung der TOA aus individuell gemessenen HRTFs wurde vorgestellt. Es kombiniert TOA-Schätzungen einzelner Richtungen durch ein TOA-Modell, welches einerseits die Schätzung verbessert und andererseits Auskunft über die Ausrichtung der Versuchsperson bezüglich des Messaufbaus gibt. Dieses Verfahren ist äußerst robust, einerseits aufgrund der Erkennung und Entfernung von Ausreißern in der Schätzung und andererseits aufgrund der Kontinuität des Modells, wodurch verworfene Richtungen korrekt interpoliert werden können.

Die numerische Evaluierung konnte zeigen, dass das TOA-Modell bezogen auf die Physiologie plausible Werte liefert. Weiters ist gezeigt worden, dass sich eine geringe seitliche Verschiebung ( $y_M < 1\text{cm}$ ) der Versuchspersonen während der HRTF-Messung auf die TOA auswirkt. Da auch das Modell diese Verschiebung mitberücksichtigt, ergibt auch hier die TOA-Modellierung für linkes und rechtes Ohr annähernd gleiche Kugelradien. Durch lineare Regression von anthropometrischen Daten und dem modellierten Kugelradius konnte eine Formel für den effektiven Kopfradius gefunden werden, die sich erheblich in den Koeffizienten von [Algazi et al. \(2001\)](#) unterscheidet, aber mit der Erwartungshaltung annähernd übereinstimmt.

Aus den Ergebnissen der vorläufigen psychoakustischen Experimente ist zu sehen, dass die geschätzte TOA für ipsilaterale Richtungen mit der perzeptiven TOA gut übereinstimmt. Wie erwartet wird die perzeptive TOA an kontralateralen Richtungen unterschätzt. Ein hörbarer Unterschied zwischen originalen und minimalphasig approximierten Phasenspektren konnte für gewisse Richtungen nachgewiesen werden. Es zeigte sich beim experimentellen Design, dass minimalphasig approximierte HRTFs eine gewisse Länge aufweisen müssen, um nach der Faltung mit einem Signal das selbe Amplitudenspektrum zu erzeugen wie originale HRTFs kürzerer Länge.

## 6.1 Ausblick

Die in Kapitel 5 vorgestellten Anwendungen der TOA müssen genauer untersucht und evaluiert werden. Die perzeptiv notwendige HRIR-Länge nach richtungsabhängiger Fensterung muss in einem psychoakustischen Experiment ermittelt werden. Weiters stellt sich die Frage, ob es einen Zusammenhang der Allpasskomponenten zwischen benachbarten Richtungen gibt und auch diese modelliert werden können. Die Bandbreitenerweiterung sowohl im Amplituden- als auch Phasenspektrum muss ebenfalls evaluiert und gegebenenfalls angepasst werden. Die TOA-bereinigten HRTFs bilden das Ausgangsmaterial für das am ARI durchgeführte Projekt "Time Frequency Virtual Acoustics", in welchem eine effiziente Echtzeitmethode zur Faltung von Audiosignalen mit HRIRs mit geringer Latenz mittels Subbandmethode implementiert wird. Die Ergebnisse der Diplomarbeit werden auf der DAGA 2012 in Darmstadt präsentiert.

# Anhang A

## Detaillierte Herleitung des TOA-Modells

### A.1 Einfaches TOA-Modell

Beim einfachen TOA-Modell sind Kugelmittelpunkt und Schnittpunkt aller eintreffenden ebenen Schallwellen koinzident positioniert. Es treten nur Schallwege um die Kugel auf und sonst keine richtungsabhängigen Wegdifferenzen. Diese werden an einem Punkt auf der Kugeloberfläche ausgewertet, welcher als Ohrposition interpretiert werden kann. Die Schallwellen aus allen Richtungen sind beim ersten Auftreffen auf die Kugel synchron. Der Ortsvektor dieses Punktes und die Schalleinfallrichtung spannen immer eine Ebene auf, die die Kugel in einem Großkreis schneidet (siehe Abbildung A.1).

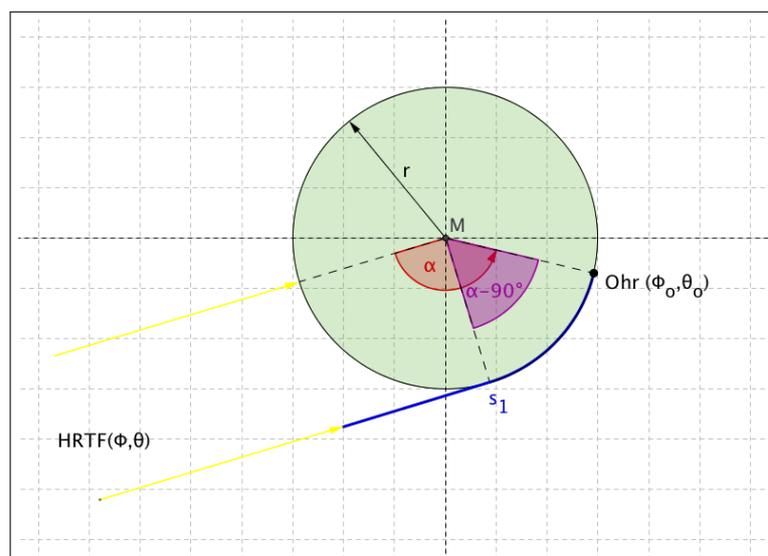


Abbildung A.1: Einfaches TOA-Modell (Skizze)

Der Winkel  $\alpha$  wird mittels Skalarprodukt berechnet und beschreibt im folgenden immer den Winkel zwischen Auswertungspunkt auf der Kugel und der Schalleinfallrichtung.

$$\begin{aligned}
\cos(\alpha) &= \left\langle \begin{pmatrix} \cos(\varphi)\cos(\vartheta) \\ \sin(\varphi)\cos(\vartheta) \\ \sin(\vartheta) \end{pmatrix}, \begin{pmatrix} \cos(\varphi_o)\cos(\vartheta_o) \\ \sin(\varphi_o)\cos(\vartheta_o) \\ \sin(\vartheta_o) \end{pmatrix} \right\rangle \\
&= \cos(\varphi)\cos(\vartheta)\cos(\varphi_o)\cos(\vartheta_o) + \sin(\varphi)\cos(\vartheta)\sin(\varphi_o)\cos(\vartheta_o) + \sin(\vartheta)\sin(\vartheta_o) \quad (\text{A.1}) \\
&= \cos(\vartheta)\cos(\vartheta_o) \underbrace{[\cos(\varphi)\cos(\varphi_o) + \sin(\varphi)\sin(\varphi_o)]}_{\cos(\varphi-\varphi_o)} + \sin(\vartheta)\sin(\vartheta_o) \\
&= \sin(\vartheta_o)\sin(\vartheta) + \cos(\vartheta_o)\cos(\vartheta)\cos(\varphi_o - \varphi)
\end{aligned}$$

Aus dem Winkel  $\alpha$ , dem Kugelradius  $r$  und der Schallgeschwindigkeit  $c$  kann die richtungsabhängige Schalllaufzeit um die Kugel berechnet werden. Ist der Auswertungspunkt der Schallwelle zugewandt ergeben sich nur direkte Schallwege.

$$\tilde{\tau} = \frac{r}{c}[1 - \cos(\alpha)] \quad \text{für } \alpha \leq \pi \quad (\text{A.2})$$

Ist dieser der Schallwelle abgewandt, tritt neben dem einfachen Schallweg  $r$  noch ein Beugungsweg um die Kugel auf.

$$\tilde{\tau} = \frac{r}{c}\left[1 + \alpha - \frac{\pi}{2}\right] \quad \text{für } \alpha > \pi \quad (\text{A.3})$$

$$\tilde{\tau}(\varphi, \vartheta) = \begin{cases} \frac{r}{c}[1 - \sin(\vartheta_o)\sin(\vartheta) - \cos(\vartheta_o)\cos(\vartheta)\cos(\varphi_o - \varphi)] & \text{für } \alpha \leq \frac{\pi}{2} \\ \frac{r}{c}\left[1 + \cos^{-1}(\sin(\vartheta_o)\sin(\vartheta) + \cos(\vartheta_o)\cos(\vartheta)\cos(\varphi_o - \varphi)) - \frac{\pi}{2}\right] & \text{für } \alpha > \frac{\pi}{2} \end{cases} \quad (\text{A.4})$$

## A.2 Erweitertes TOA-Modell

Das erweiterte Modell entspricht im wesentlichen dem einfachen Modell, mit dem Unterschied der möglichen Freiheitsgrade der Kugel. Diese kann hier beliebig vom Nullpunkt verschoben werden. Die ebenen Schallwellen sind dadurch beim ersten Kontakt mit der Kugel nicht mehr

synchron. Es tritt ein zusätzlicher richtungsabhängiger Schallweg zum einfachen Modell auf. Der Ortsvektor des Mittelpunkts der Kugel und die Schalleinfallrichtung spannen eine Ebene auf, die die Modellkugel und die synchrone äußere Kugel in zwei Großkreise schneidet (siehe Abb. A.2).

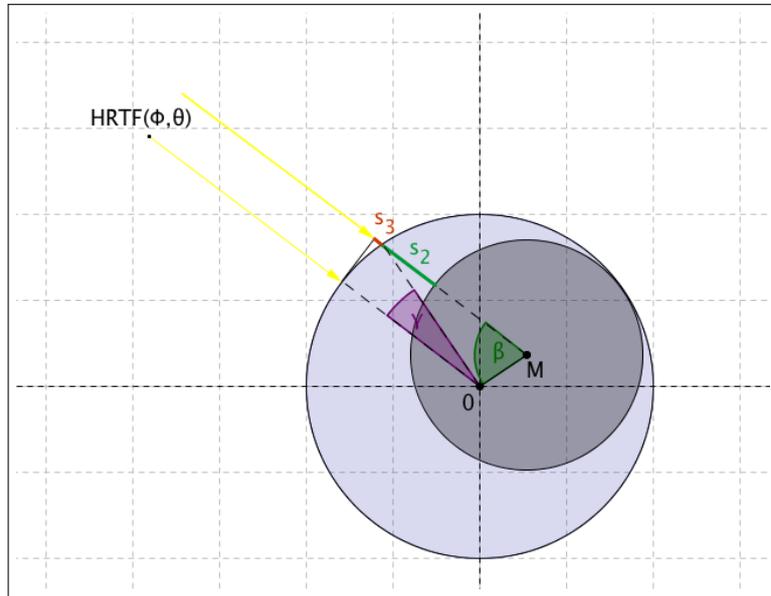


Abbildung A.2: Erweitertes TOA-Modell (Skizze)

Nach der Bestimmung des Winkels  $\beta$  zwischen Ortsvektor  $|\vec{M}|$  und Schalleinfallrichtung kann mit dem Kosinussatz der Schalllaufweg  $s_2$  berechnet werden.

$$\begin{aligned} \cos(\beta) &= \left\langle \begin{pmatrix} \cos(\varphi)\cos(\vartheta) \\ \sin(\varphi)\cos(\vartheta) \\ \sin(\vartheta) \end{pmatrix}, - \begin{pmatrix} x_M \\ y_M \\ z_M \end{pmatrix} \right\rangle \\ &= -\cos(\vartheta)[x_M\cos(\varphi) + y_M\sin(\varphi)] - z_M\sin(\vartheta) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} (s_2 + r)^2 - 2|\vec{M}|\cos(\beta)(s_2 + r) + |\vec{M}|^2 - (|\vec{M}| + r)^2 &= 0 \\ s_2^2 + 2rs_2 + r^2 - 2|\vec{M}|\cos(\beta)(s_2 + r) + |\vec{M}|^2 - |\vec{M}|^2 - 2|\vec{M}|r - r^2 &= 0 \\ s_2^2 + (2r - 2|\vec{M}|\cos(\beta))s_2 - 2|\vec{M}|r(1 + \cos(\beta)) &= 0 \end{aligned}$$

$$\begin{aligned} s_2 &= -r + |\vec{M}|\cos(\beta) \pm \sqrt{r^2 - 2r|\vec{M}|\cos(\beta) + |\vec{M}|^2\cos^2(\beta) + 2|\vec{M}|r(1 + \cos(\beta))} \\ &= -r + |\vec{M}|\cos(\beta) \pm \sqrt{r^2 + |\vec{M}|^2\cos^2(\beta) + 2|\vec{M}|r} \end{aligned} \quad (\text{A.6})$$

Aufgrund der quadratischen Gleichung erhält man zwei Lösungen. Setzt man jedoch  $|\vec{M}| = -|H\vec{R}TF|$ , so führt nur die Lösung mit positivem Vorzeichen der Wurzel zu einer wahren Behauptung.

$$\begin{aligned} 2|\vec{M}| &= -r + |\vec{M}| \pm \sqrt{r^2 + 2|\vec{M}|r + |\vec{M}|^2} \\ 2|\vec{M}| &= -r + |\vec{M}| \pm (r + |\vec{M}|) \end{aligned} \quad (\text{A.7})$$

$$s_2 = -r + |\vec{M}|\cos(\beta) + \sqrt{r^2 + |\vec{M}|^2\cos^2(\beta) + 2|\vec{M}|r} \quad (\text{A.8})$$

Der Winkel  $\gamma$  wird über Umweg, durch Berechnung von  $\delta$  mit Hilfe des Kosinussatzes, berechnet.  $\delta$  ergänzt  $\beta + \gamma$  auf  $180^\circ$ .

$$\begin{aligned} (r + s_2)^2 &= |\vec{M}|^2 + (|\vec{M}| + r)^2 - 2|\vec{M}|(|\vec{M}| + r)\cos(\delta) \\ \cos(\delta) &= \frac{|\vec{M}|^2 + |\vec{M}|^2 + 2|\vec{M}|r + \cancel{r^2} - \cancel{r^2} - 2rs_2 - s_2^2}{2|\vec{M}|^2 + 2|\vec{M}|r} \end{aligned} \quad (\text{A.9})$$

$$\gamma = \pi - \beta - \cos^{-1}\left\{\frac{2|\vec{M}|^2 + 2|\vec{M}|r - 2rs_2 - s_2^2}{2|\vec{M}|^2 + 2|\vec{M}|r}\right\} \quad (\text{A.10})$$

Zur Bestimmung von  $s_3$  wird zunächst der Normalabstand  $d$  zwischen dem durch die Punkte  $\vec{0}$  und  $\vec{M}$  parallel verschobenen Vektoren  $H\vec{R}TF$  ermittelt.

$$\begin{aligned} d &= (|\vec{M}| + r)^2 + (|\vec{M}| + r)^2 - 2(|\vec{M}| + r)^2\cos(\gamma) \\ &= 2(|\vec{M}| + r)^2\cos(\gamma) \end{aligned} \quad (\text{A.11})$$

Daraus lässt sich der Winkel  $\rho$  und im Weiteren der Weg  $s_1$  berechnen.

$$\begin{aligned} \frac{d}{\sin(\gamma)} &= \frac{|\vec{M}| + r}{\sin(\rho)} \\ 2(|\vec{M}| + r) \underbrace{\frac{1 - \cos(\gamma)}{\sin(\gamma)}}_{\tan(\frac{\gamma}{2})} &= \frac{|\vec{M}| + r}{\sin(\rho)} \\ \sin(\rho) &= \frac{1}{2(|\vec{M}| + r)\tan(\frac{\gamma}{2})} \end{aligned} \quad (\text{A.12})$$

$$s_3 = \frac{|\vec{M}| \sin(\beta)}{2(|\vec{M}| + r) \tan(\frac{\gamma}{2})} \quad (\text{A.13})$$

Die richtungsabhängigen Schalllaufzeiten des erweiterten Modells  $\tilde{\tau}_e$  ergeben sich nun aus der Summe von einfachem Modell  $\tilde{\tau}$ ,  $s_2$  und  $s_3$ .

$$\tilde{\tau}_e(\varphi, \vartheta) = \begin{cases} \frac{1}{c} \left\{ -r \sin(\vartheta_o) \sin(\vartheta) - r \cos(\vartheta_o) \cos(\vartheta) \cos(\varphi_o - \varphi) \right. \\ \quad \left. + |\vec{M}| \cos(\beta) + \sqrt{r^2 + |\vec{M}|^2 \cos^2(\beta) + 2|\vec{M}|r} \right. \\ \quad \left. + \frac{|\vec{M}| \sin(\beta)}{2(|\vec{M}| + r) \tan(\frac{\gamma}{2})} \right\} + \bar{\tau} & \alpha \leq \frac{\pi}{2} \\ \frac{1}{c} \left\{ r [\cos^{-1}(\sin(\vartheta_o) \sin(\vartheta) + \cos(\vartheta_o) \cos(\vartheta) \cos(\varphi_o - \varphi)) - \frac{\pi}{2}] \right. \\ \quad \left. + |\vec{M}| \cos(\beta) + \sqrt{r^2 + |\vec{M}|^2 \cos^2(\beta) + 2|\vec{M}|r} \right. \\ \quad \left. + \frac{|\vec{M}| \sin(\beta)}{2(|\vec{M}| + r) \tan(\frac{\gamma}{2})} \right\} + \bar{\tau} & \alpha > \frac{\pi}{2} \end{cases} \quad (\text{A.14})$$



# Anhang B

## Abbildungen und Tabellen

In Abbildung B.1 bis B.3 werden die modellierten TOAs für DTF des linken Ohrs von drei Versuchspersonen in Horizontalebene von  $\theta = -30^\circ$  links oben bis  $\theta = 80^\circ$  rechts unten gezeigt.

In Abbildung B.4 bis B.9 sind die Ergebnisse der einzelnen Versuchspersonen der vorläufigen experimentellen Evaluierung dargestellt.

In Tabelle B.1 und B.2 beinhalten die ermittelten Parameter des einfachen Modells aller untersuchten HRTF-Datensätze der ARI-Datenbank für das linke und rechte Ohr. Tabelle B.3 und B.4 zeigen die resultierenden Parameter für das erweiterte Modell.



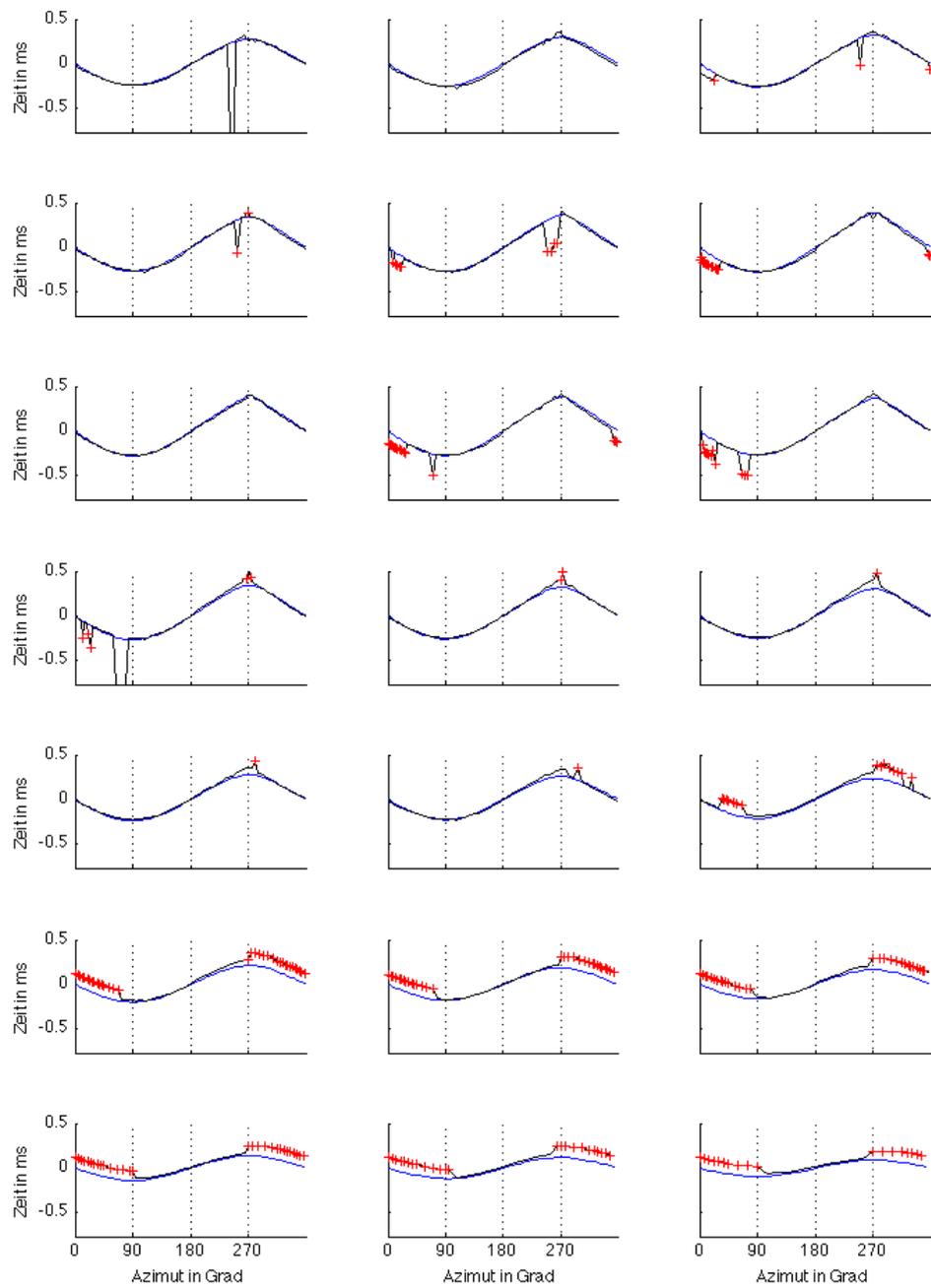


Abbildung B.2: NH18, linkes Ohr

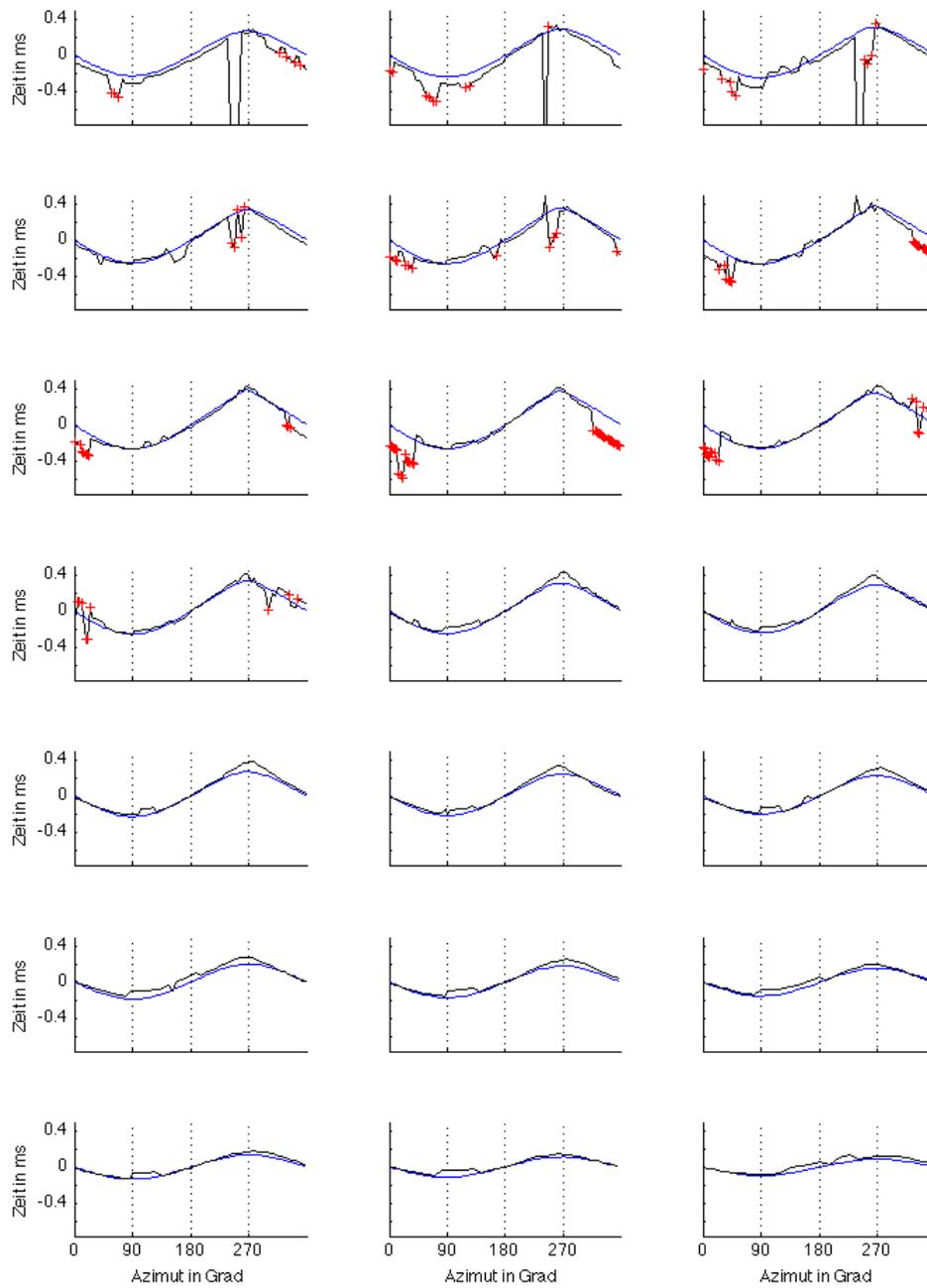


Abbildung B.3: NH4, linkes Ohr

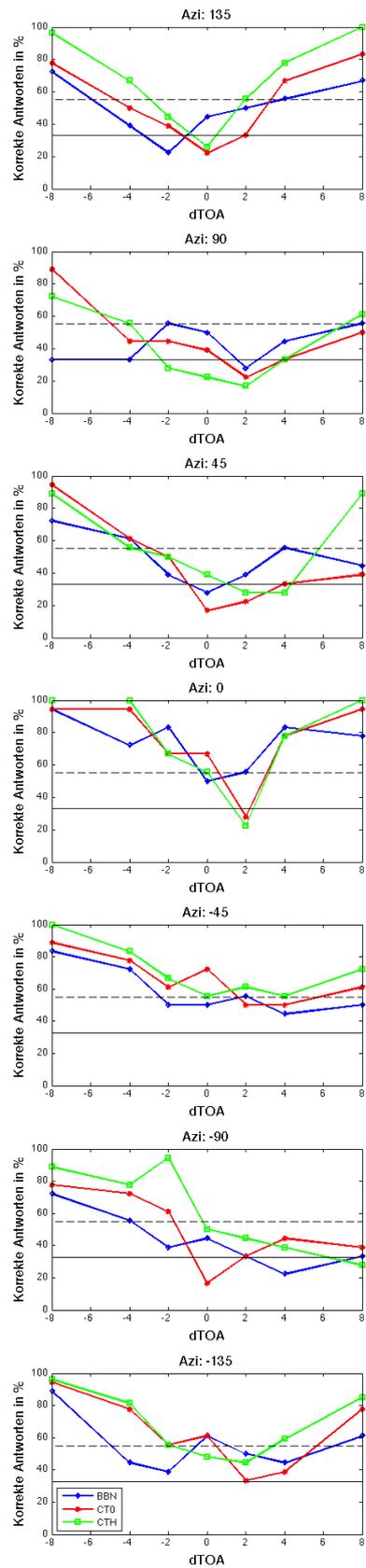
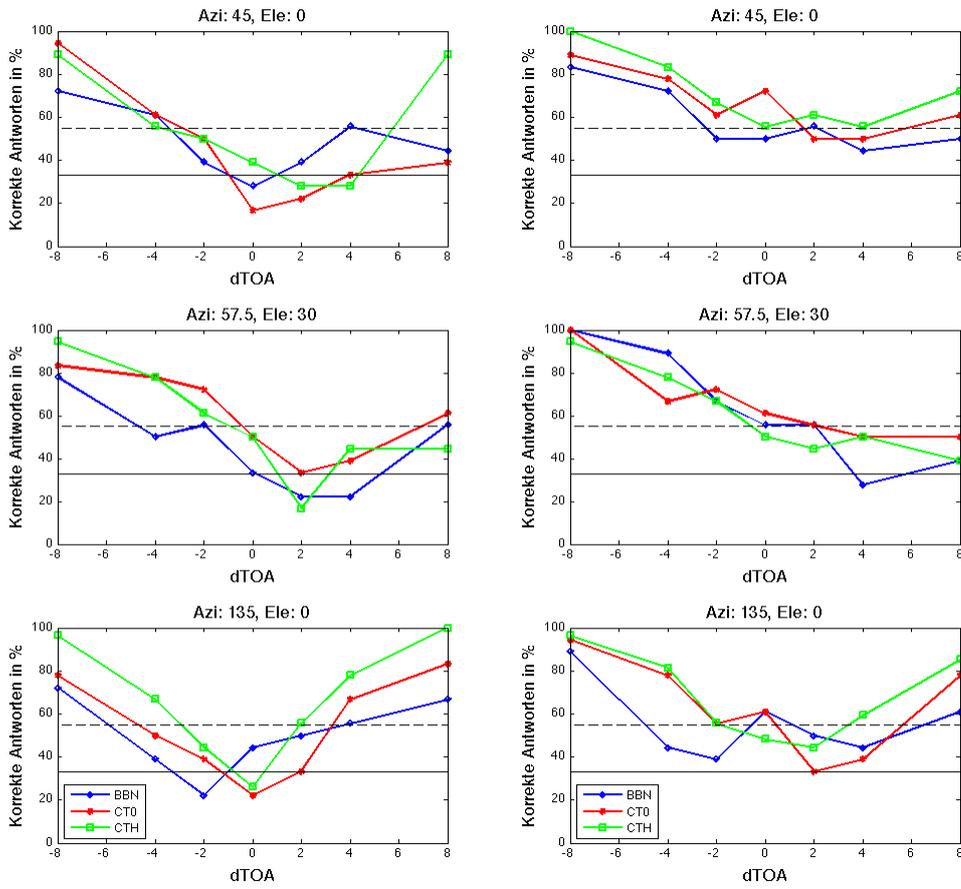


Abbildung B.4: Ergebnisse der experimentellen Evaluierung (VP 1)



(a) linkes Ohr, ipsilateral

(b) rechtes Ohr, kontralateral

Abbildung B.5: Ergebnisse der experimentellen Evaluierung (VP 1)

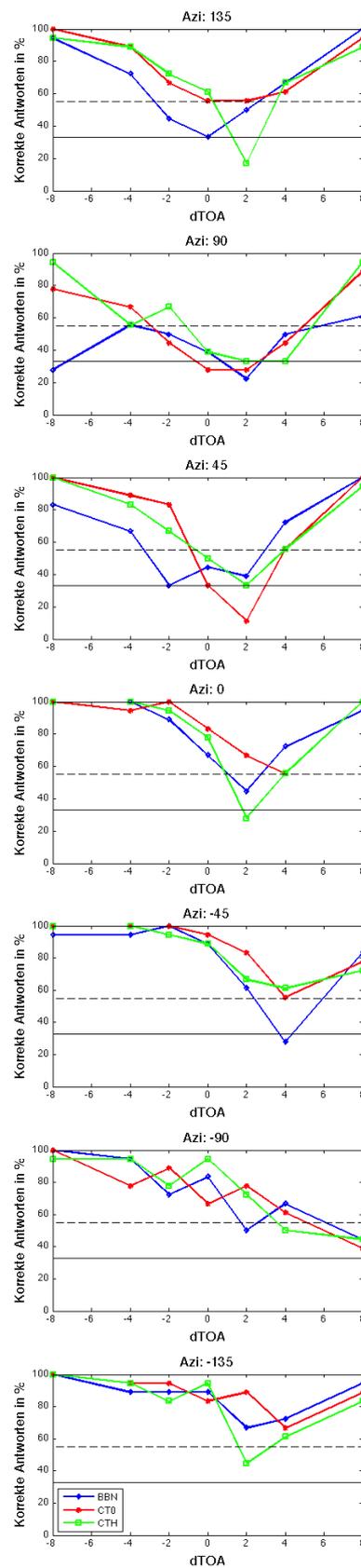
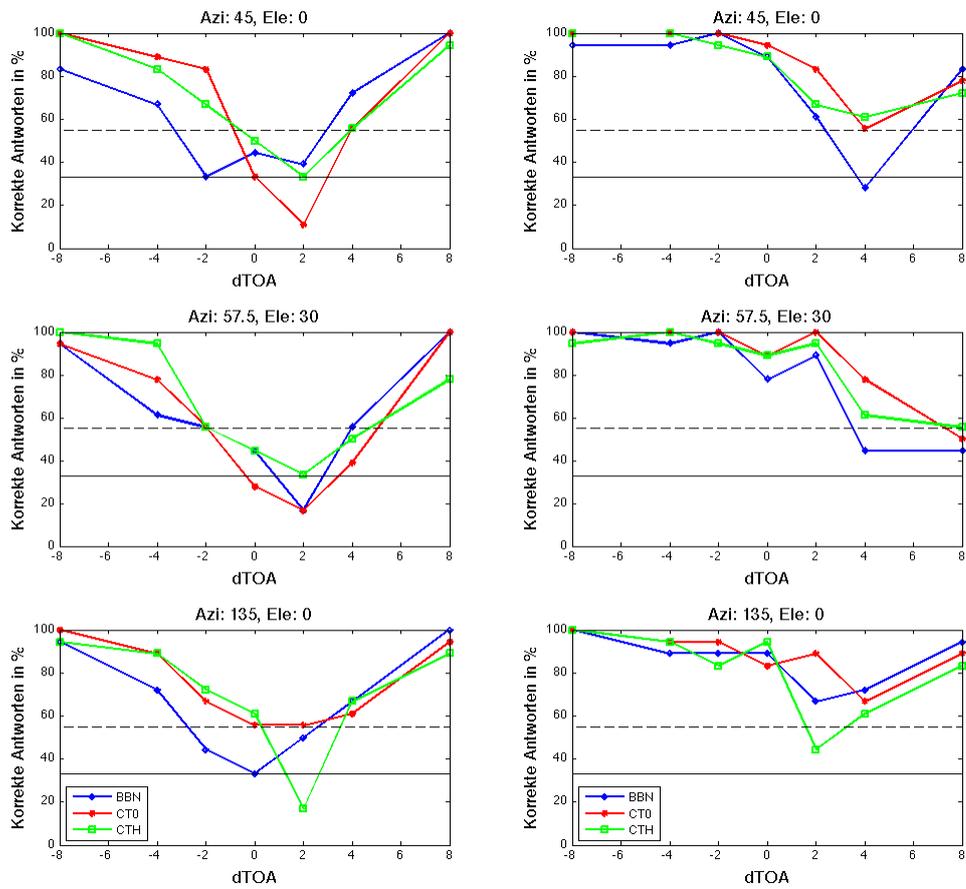


Abbildung B.6: Ergebnisse der experimentellen Evaluierung (VP 2)



(a) linkes Ohr, ipsilateral

(b) rechtes Ohr, kontralateral

Abbildung B.7: Ergebnisse der experimentellen Evaluierung (VP 2)

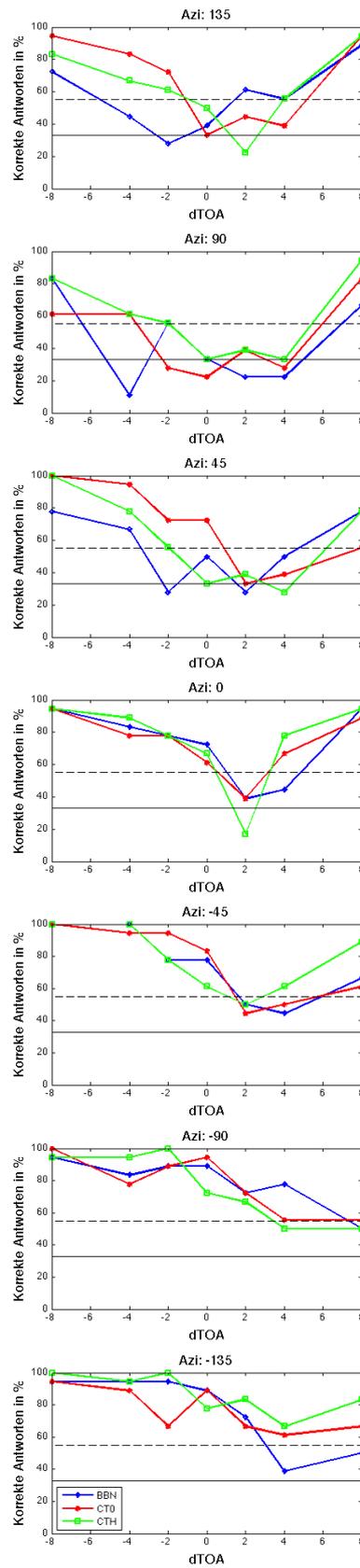
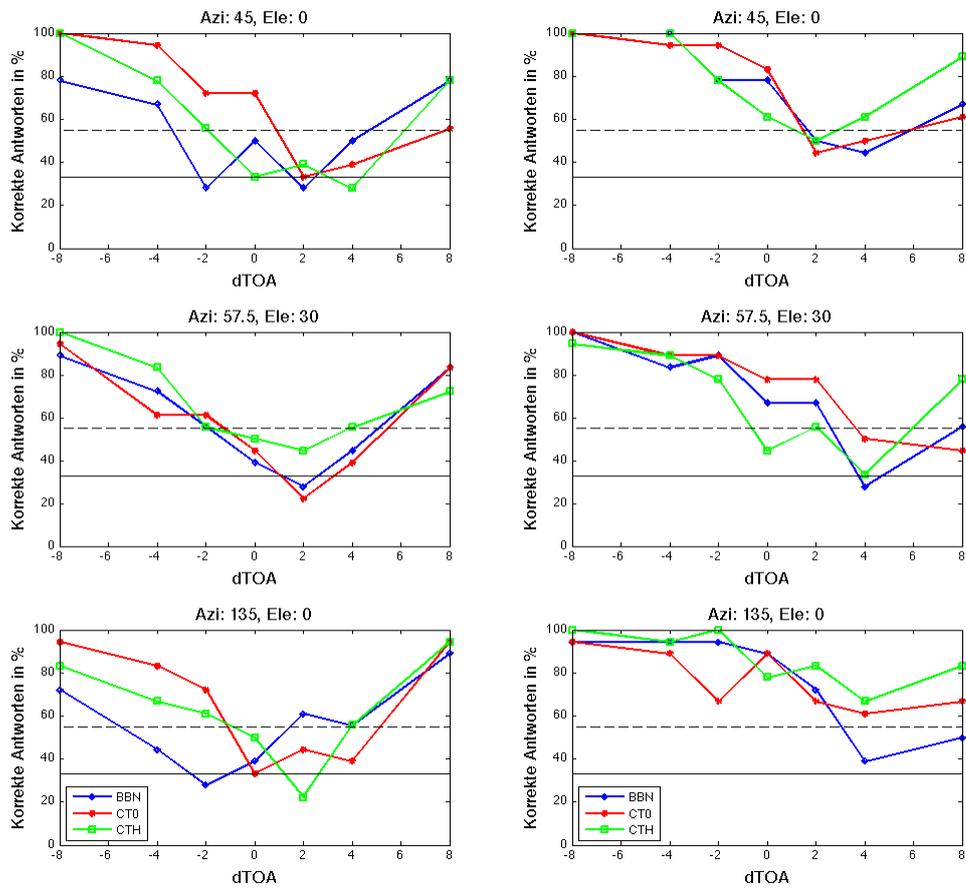


Abbildung B.8: Ergebnisse der experimentellen Evaluierung (VP 3)



(a) linkes Ohr, ipsilateral

(b) rechtes Ohr, kontralateral

Abbildung B.9: Ergebnisse der experimentellen Evaluierung (VP 3)

**Tabelle B.1:** Parameter für das einfache Modell der einzelnen Versuchspersonen (linkes Ohr)

VP	$r$ [mm]	$\phi_o$ [°]	$\theta_o$ [°]	$\bar{\tau}$ [ms]
NH2	92.39	84.481	-2.7643	2.7197
NH4	98.525	89.699	-5.5432	2.7316
NH5	85.941	85.877	-11.628	2.7393
NH8	108.43	86.822	-3.4724	2.6565
NH10	98.637	87.377	-3.963	2.7181
NH12	84.533	85.548	-1.8893	2.75
NH13	83.885	94.935	-6.459	2.7505
NH14	97.578	83.142	-8.2832	2.7129
NH16	76.177	87.692	-9.5476	2.7587
NH17	76.326	87.483	-2.595	2.7899
NH18	98.8	90.767	-5.7221	2.6951
NH19	82.241	96.692	-2.6518	2.7329
NH21	101.69	92.457	-2.7363	2.6925
NH22	102.5	86.317	-7.4987	2.6941
NH25	95.038	88.992	-6.607	2.6772
NH27	98.344	87.859	-4.6446	2.7158
NH28	70.766	86.757	-2.6538	2.8015
NH29	88.692	90.197	-8.261	2.7433
NH30	69.891	87.477	-3.3219	2.7921
NH32	94.044	91.296	-2.8689	2.7139
NH33	92.687	91.343	-7.4133	2.7174
NH34	105.24	88.788	-4.9548	2.6996
NH35	94.229	91.591	-5.2136	2.7317
NH36	80.392	90.01	-4.9636	2.7655
NH39	86.689	87.009	-6.7505	2.6935
NH40	77.651	87.461	-9.0906	2.7352
NH41	90.232	90.683	-6.7907	2.6914
NH42	114.03	94.102	-6.554	2.609
NH43	110.84	92.933	-8.994	2.6314
NH46	92.651	88.105	-3.2589	2.7392
NH47	104.96	90.857	-8.4361	2.6484
NH52	92.285	94.31	-7.4712	2.7013

Fortsetzung nächste Seite

**Tabelle B.1** – Fortsetzung

VP	$r$ [mm]	$\phi_o$ [°]	$\theta_o$ [°]	$\bar{\tau}$ [ms]
NH53	103.61	87.783	-7.1945	2.6733
NH54	78.457	90.73	-7.727	2.742
NH55	104.39	86.336	-4.1973	2.6918
NH56	85.798	92.799	-6.0822	2.7454
NH57	103.4	81.841	-6.7079	2.6888
NH58	101.96	81.413	-8.7558	2.6933
NH61	97.684	88.609	-4.492	2.7238
NH62	81.619	82.727	-12.564	2.7585
NH63	110.87	86.985	-8.0343	2.6813
NH64	86.105	88.571	-5.6283	2.7546
NH65	88.271	88.407	-7.3138	2.7458
NH66	118.36	87.95	-2.0151	2.664
NH68	59.338	88.929	-9.8413	2.8092
NH70	100.55	86.383	-5.1525	2.7174
NH71	113.33	88.334	-7.1255	2.6738
NH72	81.158	94.131	-1.645	2.7727
NH74	92.751	90.122	-5.1185	2.7304
NH89	108.07	85.109	-5.0076	2.7153
NH90	88.308	78.551	-5.2919	2.7585
NH92	99.056	78.963	-6.5383	2.7136
NH93	86.807	90.35	-3.5098	2.7381
NH102	95.42	84.093	-5.6363	2.7427
NH103	109.57	89.164	-7.5208	2.6798
NH104	103.99	89.968	-7.9334	2.708

**Tabelle B.2:** Parameter für das einfache Modell der einzelnen Versuchspersonen (rechtes Ohr)

VP	$r$ [mm]	$\phi_o$ [°]	$\theta_o$ [°]	$\bar{\tau}$ [ms]
NH2	81.528	-81.365	-3.3235	2.7573
NH4	83.467	-86.511	-6.9723	2.784
NH5	89.19	-88.79	-5.8068	2.7198
NH8	75.717	-85.738	-6.4828	2.7499
NH10	81.662	-73.518	-3.695	2.7644

Fortsetzung nächste Seite

Tabelle B.2 – Fortsetzung

VP	$r$ [mm]	$\phi_o$ [°]	$\theta_o$ [°]	$\bar{\tau}$ [ms]
NH12	97.532	-86.937	-5.1028	2.7154
NH13	81.373	-92.625	-7.7836	2.7537
NH14	77.433	-75.511	-12.116	2.7737
NH16	99.164	-79.028	-10.176	2.7074
NH17	102.42	-84.459	-11.791	2.7065
NH18	79.005	-97.558	-4.9271	2.7636
NH19	100.76	-103.41	-9.1621	2.6865
NH21	71.928	-80.565	-12.938	2.8049
NH22	75.453	-87.051	-11.855	2.773
NH25	77.101	-87.082	-3.3896	2.7268
NH27	81.423	-75.712	-6.7785	2.7668
NH28	95.191	-87.486	-5.7491	2.7328
NH29	77.629	-85.69	-5.5305	2.779
NH30	101.87	-86.318	-4.1582	2.7
NH32	81.25	-87.913	-11.601	2.7481
NH33	89.67	-81.14	-5.0443	2.7271
NH34	69.628	-76.4	-12.735	2.8027
NH35	67.012	-88.855	-6.8292	2.8199
NH36	92.93	-89.262	-3.3895	2.7227
NH39	78.778	-95.932	-13.52	2.7196
NH40	101.73	-90.724	-9.8541	2.6617
NH41	70.285	-94.972	-5.9308	2.7586
NH42	58.443	-107.58	-16.506	2.8048
NH43	65.492	-102.54	-18.884	2.7624
NH46	93.432	-87.52	-7.4906	2.7383
NH47	73.199	-87.253	-15.636	2.749
NH52	90.927	-85.521	-8.3193	2.6982
NH53	72.054	-90.434	-12.549	2.7616
NH54	101.63	-91.672	-0.76248	2.679
NH55	60.389	-94.016	-11.592	2.8162
NH56	100.72	-89.094	-6.3547	2.6975
NH57	80.889	-89.265	-7.3368	2.7583
NH58	87.358	-75.806	-15.006	2.742
NH61	79.239	-90.074	-6.9394	2.7832

Fortsetzung nächste Seite

Tabelle B.2 – Fortsetzung

VP	$r$ [mm]	$\phi_o$ [°]	$\theta_o$ [°]	$\bar{\tau}$ [ms]
NH62	95.074	-81.519	-12.715	2.7213
NH63	70.772	-84.34	-7.3857	2.799
NH64	87.582	-85.992	-11.731	2.7449
NH65	83.147	-88.006	-9.2114	2.7646
NH66	62.461	-90.787	-3.0846	2.8343
NH68	106.81	-83.586	-2.4665	2.6731
NH70	72.906	-86.883	-6.2894	2.7941
NH71	69.076	-95.275	-7.4801	2.7969
NH72	99.944	-90.599	-5.5761	2.7193
NH74	69.922	-90.094	-10.963	2.7971
NH89	67.964	-76.659	-10.309	2.8278
NH90	103.04	-83.325	-4.7481	2.7184
NH92	83.251	-78.923	-4.6149	2.7635
NH93	89.821	-85.673	-9.6495	2.7355
NH102	84.809	-90.61	-2.0107	2.7712
NH103	61.523	-101.74	-7.4561	2.8108
NH104	74.253	-88.708	-8.9	2.7956

Tabelle B.3: Parameter für das erweiterte Modell der einzelnen Versuchspersonen (linkes Ohr)

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH2	85.165	-10.54	4.5606	1.2058e-06	2.706
NH4	88.61	-3.6848	6.6004	4.4083e-07	2.7445
NH5	86.778	-5.4269	-1.7439	-7.2328e-05	2.727
NH8	92.498	-7.7813	16.003	-3.9614e-06	2.658
NH10	89.234	-14.883	8.3212	-0.00016021	2.7011
NH12	91.631	-8.8763	-8.9563	-9.1608e-06	2.6925
NH13	81.789	4.8756	1.3155	8.0394e-06	2.7519
NH14	85.067	-15.631	9.4637	-0.00031106	2.7083
NH16	83.37	-9.7476	-9.4444	-1.0553e-05	2.7059
NH17	88.08	-8.4724	-13.654	-5.0138e-05	2.7091
NH18	87.165	5.584	9.407	-0.00027737	2.7064
NH19	87.965	14.903	-7.8145	-0.00010007	2.6737

Fortsetzung nächste Seite

Tabelle B.3 – Fortsetzung

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH21	85.349	3.1944	15.736	-4.2967e-05	2.7038
NH22	87.04	-6.0796	11.891	-0.00013381	2.7108
NH25	84.486	-0.72343	8.6597	-5.9906e-05	2.6898
NH27	89.392	-12.797	8.8479	-0.00012354	2.7037
NH28	79.45	-4.8242	-11.076	-0.00010809	2.7415
NH29	80.792	-2.095	4.642	-0.00010251	2.7602
NH30	83.833	-5.4578	-15.537	1.9589e-06	2.707
NH32	87.136	1.971	6.4619	-2.9579e-05	2.7182
NH33	91.239	-11.36	0.57522	-2.9252e-05	2.6984
NH34	87.127	-7.8732	18.84	-0.00011037	2.7064
NH35	79.91	1.4586	13.948	-6.8488e-05	2.744
NH36	84.498	-0.68134	-5.8714	-5.9361e-06	2.7396
NH39	80.622	0.79405	3.694	4.1235e-06	2.7044
NH40	86.742	-3.0804	-10.715	-2.4586e-05	2.683
NH41	77.947	2.7797	8.8606	-1.4058e-05	2.7085
NH42	86.039	9.4674	29.377	-9.5137e-06	2.6281
NH43	83.734	7.7366	21.581	-0.053486	2.6661
NH46	91.381	-3.7092	-0.30911	-4.3092e-05	2.7333
NH47	86.093	-2.0085	15.137	-1.7938e-05	2.6745
NH52	90.643	2.0447	1.0871	9.9757e-06	2.7086
NH53	85.09	-4.6906	15.498	-0.00014744	2.6914
NH54	87.232	1.4721	-9.8641	-0.00017124	2.694
NH55	81.17	-3.2116	22.439	4.0711e-06	2.7028
NH56	90.961	1.776	-6.762	-4.035e-06	2.7166
NH57	89.295	-8.7725	9.6918	-5.3083e-05	2.6952
NH58	91.78	-17.511	6.5246	-5.3592e-06	2.6767
NH61	85.946	-1.3595	8.0502	-9.0666e-06	2.7387
NH62	86.316	-12.51	-5.8391	-1.1371e-05	2.7176
NH63	88.534	-6.434	19.128	-3.4069e-05	2.7022
NH64	85.591	-4.3028	-0.44777	-3.4557e-05	2.7478
NH65	83.79	-2.4	2.1555	-2.973e-05	2.7555
NH66	92.18	-8.8618	27.174	-0.00022636	2.6694
NH68	79.338	-6.7743	-22.15	-4.6759e-05	2.6937
NH70	85.502	-5.6791	14.031	-0.00012792	2.7244

Fortsetzung nächste Seite

Tabelle B.3 – Fortsetzung

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH71	89.517	1.1474	22.006	-0.00017032	2.6927
NH72	88.554	4.3207	-8.7964	-4.058e-05	2.7241
NH74	78.66	0.10737	10.53	-0.014734	2.749
NH89	88.176	-12.15	21.204	-0.00017823	2.7153
NH90	92.348	-14.961	-5.9728	-4.9388e-05	2.7035
NH92	89.134	-15.237	7.0331	-1.5757e-05	2.6986
NH93	86.591	-3.7414	-1.2273	-0.00020852	2.7304
NH102	89.217	-6.2028	5.4328	-0.00029561	2.7416
NH103	85.544	3.1677	24.733	-6.3517e-05	2.6978
NH104	86.803	-0.77475	14.229	-3.4606e-05	2.7305

Tabelle B.4: Parameter für das erweiterte Modell der einzelnen Versuchspersonen (rechtes Ohr)

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH2	85.179	-11.14	5.3688	-3.7455e-06	2.7082
NH4	88.54	-4.3885	7.3187	-0.00021556	2.7478
NH5	86.623	-4.9784	-1.5961	-4.1474e-05	2.7186
NH8	88.826	-5.9177	14.37	-0.00014763	2.6691
NH10	87.82	-25.713	8.7309	-5.6864e-05	2.668
NH12	90.245	-6.127	-6.3712	-6.1593e-05	2.7151
NH13	81.557	4.3529	1.1482	-0.0001605	2.7467
NH14	84.359	-15.841	9.0651	1.1232e-06	2.7069
NH16	86.188	-12.282	-10.571	-1.8053e-05	2.71
NH17	89.012	-8.6217	-13.452	3.4855e-08	2.7142
NH18	86.901	8.3232	10.798	-1.8292e-05	2.7094
NH19	89.773	17.321	-8.2666	4.8378e-07	2.6853
NH21	84.376	-4.0574	15.654	-1.1676e-05	2.7308
NH22	86.699	-6.0386	11.585	-1.5858e-05	2.7106
NH25	82.339	-1.2427	7.7328	4.2723e-06	2.6889
NH27	86.516	-13.712	7.2932	-0.0023752	2.7069
NH28	82.007	-5.1286	-12.255	-0.00026179	2.7416
NH29	80.7	-2.1189	4.5586	-5.1002e-05	2.758
NH30	84.852	-5.8751	-15.966	-1.6064e-05	2.7083

Fortsetzung nächste Seite

Tabelle B.4 – Fortsetzung

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH32	84.888	-0.11846	5.7202	-1.5055e-05	2.7279
NH33	89.85	-12.695	0.76024	-0.0057324	2.6896
NH34	82.966	-9.2405	15.852	0.00019744	2.7193
NH35	77.508	0.75053	11.93	-1.8888e-05	2.7593
NH36	84.822	-0.77865	-5.4881	-7.9324e-05	2.7337
NH39	81.873	1.6793	5.5397	-7.8452e-05	2.7071
NH40	86.926	-3.1	-10.858	-0.00026258	2.686
NH41	78.078	3.6532	10.136	-5.7262e-07	2.7107
NH42	81.972	12.499	30.07	-5.0141e-05	2.673
NH43	83.007	7.922	20.89	1.0166e-06	2.6695
NH46	91.853	-4.2708	-0.22349	-0.00023688	2.7364
NH47	85.291	-2.098	14.729	-0.00029197	2.6839
NH52	91.086	-0.85056	1.8134	-8.0596e-05	2.6962
NH53	83.909	-4.2245	14.435	1.2644e-05	2.6941
NH54	90.652	3.2104	-10.62	-0.00020088	2.6827
NH55	77.97	-1.9623	19.881	-0.0004065	2.718
NH56	90.96	1.4189	-6.2759	-0.00033704	2.714
NH57	89.995	-9.3676	11.34	-4.8156e-05	2.6944
NH58	91.778	-17.618	6.5285	-2.4325e-05	2.6863
NH61	85.775	-1.2941	8.1089	-0.0001755	2.7443
NH62	86.324	-12.518	-5.8432	-0.00017145	2.7208
NH63	85.81	-6.2619	18.788	-0.00019513	2.7028
NH64	85.591	-4.3025	-0.447	-5.2021e-05	2.7473
NH65	83.802	-2.4344	2.1996	2.8305e-06	2.7599
NH66	87.324	-1.486	28.357	-0.00017416	2.6882
NH68	83.035	-11.459	-24.592	-8.4747e-06	2.6726
NH70	82.547	-5.1441	12.204	7.3841e-06	2.7303
NH71	86.255	2.7809	21.169	-1.5931e-05	2.6918
NH72	88.969	3.2278	-8.6895	-5.642e-05	2.73
NH74	78.272	0.082937	10.176	6.2634e-06	2.7512
NH89	84.345	-12.169	19.092	-3.191e-05	2.7192
NH90	95.449	-14.271	-7.5168	-0.00016437	2.6987
NH92	89.133	-15.237	7.0326	-7.6743e-05	2.6987
NH93	86.858	-4.4398	-1.2452	-1.0637e-05	2.7371

Fortsetzung nächste Seite

Tabelle B.4 – Fortsetzung

Versuchsperson	$r$ [mm]	$x_M$ [mm]	$y_M$ [mm]	$z_M$ [mm]	$\bar{\tau}$ [ms]
NH102	88.587	-3.1205	6.6122	-0.00010903	2.7396
NH103	80.954	6.0257	23.317	-4.0455e-05	2.6974
NH104	85.515	-0.89008	13.18	1.5346e-05	2.7313

# Referenzen

- Algazi, V. R., Avendano, C., und Duda, R. O. (2001). Estimation of a spherical-head model from anthropometry. *J. Audio Eng. Soc.*, 49(6):472–479.
- Algazi, V. R., Duda, R. O., Duraiswami, R., Gumerov, N. A., und Tang, Z. (2002). Approximating the head-related transfer function using simple geometric models of the head and torso. *J. Acoust. Soc. Am.*, 112(5):2053–2064.
- Batteau, D. W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society of London Series B Biological Sciences*, 168(1011):158–180.
- Begault, D. R. (2000). *3-D Sound for Virtual Reality and Multimedia*. Morgan Kaufmann, second edition.
- Brungart, D. S. und Rabinowitz, W. M. (1999). Auditory localization of nearby sources. head-related transfer functions. *J. Acoust. Soc. Am.*, 106(3):1465–1479.
- Constan, Z. A. und Hartmann, W. M. (2001). Sound localization by interaural time differences at high frequencies. *J. Acoust. Soc. Am.*, 109(5):2485–2485.
- Constan, Z. A. und Hartmann, W. M. (2003). On the detection of dispersion in the head-related transfer function. *J. Acoust. Soc. Am.*, 114(2):998–1008.
- Duda, R. O., Avendano, C., und Algazi, V. R. (1999). An adaptable ellipsoidal head model for the interaural time difference. In *Proceedings of the Acoustics, Speech, and Signal Processing, 1999. on 1999 IEEE International Conference - Volume 02, ICASSP '99*, pages 965–968, Washington, DC, USA. IEEE Computer Society.
- Hartley, R. V. L. und Fry, T. C. (1921). The binaural location of pure tones. *Phys. Rev.*, 18:431–442.
- Henning, G. B. (1974). Detectability of interaural delay in high-frequency complex waveforms. *J. Acoust. Soc. Am.*, 55(1):84–90.

- Henrik und Møller (1992). Fundamentals of binaural technology. *Applied Acoustics*, 36(3-4):171 – 218.
- Howard, D. M. und Angus, J. A. S. (2009). *Acoustics and Psychoacoustics*. Focal Press, fourth edition edition.
- John William Strutt, r. B. R. (1877). *The theory of sound*. Macmillan and Co. London.
- Jot, Jean-Marc; Larcher, V. W. O. (1995). Digital signal processing issues in the context of binaural and transaural stereophony. In *Audio Engineering Society Convention 98*.
- Katz, B., Nicol, R., und Busson, S. (2005). Subjective investigations of the interaural time difference in the horizontal plane. In *Audio Engineering Society Convention 118*.
- Kistler, D. J. und Wightman, F. L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91(3):1637–1647.
- Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.*, 62(1):157–167.
- Kulkarni, A. und Colburn, H. S. (1998). Role of spectral detail in sound-source localization. *Nature*, 396(6713):747–749.
- Kulkarni, A., Isabelle, S. K., und Colburn, H. S. (1999). Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105(5):2821–2840.
- Larsen, E. und Aarts, R. M. (2004). *Audio Bandwidth Extension*. John Wiley and Sons Ltd.
- Macpherson, E. A. und Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *J. Acoust. Soc. Am.*, 111(5):2219–2236.
- Majdak, P. (2002). Cues on hrtfs. interne Präsentation.
- Majdak, P. (2010). Measurement of head-related transfer functions: post processing.
- Majdak, P. (2011). Algorithmen in akustik und computermusik. Skriptum zur Vorlesung.
- Majdak, P., Balazs, P., und Laback, B. (2007). Multiple exponential sweep method for fast measurement of head-related transfer functions. *J. Audio Eng. Soc.*, 55(7/8):623–637.
- Majdak, P., Goupell, M., und Laback, B. (2010). 3-d localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attention, Perception and Psychophysics*, 72:454–469. 10.3758/APP.72.2.454.

- Matsumoto, Mitsuo; Yamanaka, S. T. M. N. H. (2004). Effect of arrival time correction on the accuracy of binaural impulse response interpolation–interpolation methods of binaural response. *J. Audio Eng. Soc.*, 52(1/2):56–61.
- Mehrgardt, S. und Mellert, V. (1977). Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, 61(6):1567–1576.
- Middlebrooks, J. C. (1999). Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *J. Acoust. Soc. Am.*, 106(3):1493–1510.
- Middlebrooks, J. C. und Green, D. M. (1990). Directional dependence of interaural envelope delays. *J. Acoust. Soc. Am.*, 87(5):2149–2162.
- Mills, A. W. (1958). On the minimum audible angle. *J. Acoust. Soc. Am.*, 30(4):237–246.
- Minnaar, P., Christensen, F., Møller, H., Olesen, S. K., und Plogsties, J. (1999). Audibility of all-pass components in binaural synthesis. In *Audio Engineering Society Convention 106*.
- Minnaar, P., Plogsties, J., Olesen, S. K., Christensen, F., und Møller, H. (2000). The interaural time difference in binaural synthesis. In *Audio Engineering Society Convention 108*.
- Møller, H., Sørensen, M. F., Hammershøi, D., und Jensen, C. B. (1995). Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43(5):300–321.
- Moore, B. C. J. und Glasberg, B. R. (1989). Difference limens for phase in normal and hearing-impaired subjects. *J. Acoust. Soc. Am.*, 86(4):1351–1365.
- Moore, B. C. J., Oldfield, S. R., und Dooley, G. J. (1989). Detection and discrimination of spectral peaks and notches at 1 and 8 khz. *J. Acoust. Soc. Am.*, 85(2):820–836.
- Moore, B. C. J. und Tan, C.-T. (2003). Perceived naturalness of spectrally distorted speech and music. *J. Acoust. Soc. Am.*, 114(1):408–419.
- Nam, J., Abel, J. S., und Smith III, J. O. (2011). A method for estimating interaural time difference for binaural synthesis. In *Audio Engineering Society Convention 125*.
- Nam, J., Kolar, M. A., und Abel, J. S. (2008). On the minimum-phase nature of head-related transfer functions. In *Audio Engineering Society Convention 125*.
- Oppenheim, A. V., Schaffer, R. W., und Buck, J. R. (2004). *Zeitdiskrete Signalverarbeitung*. Pearson Studium.

- Plogsties, J., Minnaar, P., Olesen, S. K., Christensen, F., und Møller, H. (2000). Audibility of all-pass components in head-related transfer functions. In *Audio Engineering Society Convention 108*.
- Savioja, Lauri; Huopaniemi, J. L. T. V. R. (1999). Creating interactive virtual acoustic environments. *J. Audio Eng. Soc.*, 47(9):675–705.
- Treeby, B. E., Pan, J., und Paurobally, R. M. (2007a). The effect of hair on auditory localization cues. *J. Acoust. Soc. Am.*, 122(6):3586–3597.
- Treeby, B. E., Paurobally, R. M., und Pan, J. (2007b). The effect of impedance on interaural azimuth cues derived from a spherical head model. *J. Acoust. Soc. Am.*, 121(4):2217–2226.
- Wang, L., Yin, F., und Chen, Z. (2009). A hybrid compression method for head-related transfer functions. *Applied Acoustics*, 70(9):1212–1218.
- Wightman, F. und Kistler, D. (2005). Measurement and validation of human hrtfs for use in hearing research. *Acta Acustica united with Acustica*, 91(3):429–439.
- Wightman, F. L. und Kistler, D. J. (1989). Headphone simulation of free-field listening. i: Stimulus synthesis. *J. Acoust. Soc. Am.*, 85(2):858–867.
- Wightman, F. L. und Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am.*, 91(3):1648–1661.
- Wightman, F. L. und Kistler, D. J. (1997). Monaural sound localization revisited. *J. Acoust. Soc. Am.*, 101(2):1050–1063.
- Woodworth, R. S. und Schlossberg, H. (1962). *Experimental Psychology*. Holt.
- Zaar, J. (2011). Phase unwrapping for spherical interpolation of head-related transfer functions. Master’s thesis, Universität für Musik und darstellende Kunst, Graz.
- Zotkin, D. N., Duraiswami, R., Grassi, E., und Gumerov, N. A. (2006). Fast head-related transfer function measurement via reciprocity. *J. Acoust. Soc. Am.*, 120(4):2202–2215.
- Zwislocki, J. und Feldman, R. S. (1956). Just noticeable differences in dichotic phase. *J. Acoust. Soc. Am.*, 28(5):860–864.