Deutsche Gesellschaft für Audiologie Deutsche Gesellschaft für Medizinische Physik Österreichische Gesellschaft für Medizinische Physik Schweizerische Gesellschaft für Strahlenbiologie und Medizinische Physik

# XVII. Winterschule für Medizinische Physik

# Medizinische Akustik und Audiologie

Birger Kollmeier/Oldenburg



Pichl/Steiermark 1. bis 6. März 2009 Deutsche Gesellschaft für Audiologie Deutsche Gesellschaft für Medizinische Physik Österreichische Gesellschaft für Medizinische Physik Schweizerische Gesellschaft für Strahlenbiologie und Medizinische Physik

# XVII. Winterschule

# Pichl/Steiermark 1. bis 6. März 2009

# Kurs 1

# Medizinische Akustik und Audiologie

**Prof. Dr. Dr. Birger Kollmeier** 

2009

# Inhalt

Liste	der Referenten	3
Kursprogramm		
Liste der Sponsoren		
Kurzfa	assungen der Vorträge	
1.	Bau- und Raumakustik	7
2.	Elektroakustik	31
3.	Das periphere Hörorgan – Funktionsweise und Modellbildung	53
4.	Psychoakustik und Sprachperzeption bei Normal- und	
	Schwerhörigen	63
5.	Hördiagnostik für die rehabilitative Audiologie	93
6.	Audiometrische Standardverfahren, Otoakustische Emissionen, Auditorisch evozierte Potentiale und Qualitätssicherung	103
7.	Physikalische Grundlagen der Akustik	121
8.	Akustische Messverfahren	141
9.	Signalverarbeitung für Hörhilfen und Audiologie	157
10.	Cochlea-Implantate	167
11.	Implantierbare Hörgeräte	179
12.	Versorgung und Rehabilitation mit technischen Hörhilfen	187
Anh	ang: Numerische Akustik	201

# Referenten

- Prof. Dr. Matthias Blau, FH Oldenburg-Ostfriesland-Wilhelmshaven, Institut für Hörtechnik und Audiologie Haus des Hörens, Marie-Curie-Straße 2, 26129 Oldenburg
- Dr. Dipl. Inform. Andreas Büchner, Hörzentrum Hannover an der MHH, Wissenschaftlicher Leiter, Karl-Wiechert-Allee 3, 30625 Hannover
- Prof. Dr. Norbert Dillier, Labor für experimentelle Audiologie, Universitäts-Spital Zürich, Ch-8091 Zürich
- **Dr.-Ing. Janina Fels,** Institut für Technische Akustik, RWTH Aachen 52056 Aachen
- Prof. Dr.-Ing. Herbert Hudde, Institut für Kommunikationsakustik, Ruhr-Universität Bochum 44780 Bochum
- Prof. Dr.-Ing. Thomas Janssen, HNO-Klinik der Technischen Universität, Experimentelle OtoRhino, Ismaningerstr. 22, 81675 München
- Prof. Dr. Jürgen Kießling, Universitätsklinikum Gießen und Marburg GmbH, Funktionsbereich Audiologie, Klinikstraße 29, 35385 Gießen
- Prof. Dr. Dr. Birger Kollmeier, Fakultät V/Institut für Physik, Abt. Medizinische Physik, Universität Oldenburg, Carl-von-Ossietzky-Straße 9-11, 26129 Oldenburg
- Prof. Dr. Thomas Lenarz, Medizinische Hochschule Hannover, Direktor der Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde, Carl-Neuberg-Straße 1, 30625 Hannover
- **Dr.-Ing. Klaus Welker** ehem. Radiolog. Klinik, Krankenhaus Berlin-Moabit Teplitzerstr.12, 14193 Berlin

# Kursprogramm

### Sonntag : Anreise

19.00 Empfang im Pichlmayrgut & Begrüßung (B. Kollmeier, R. Weise)

#### Montag: Grundlagen der Kommunikationsakustik und Signalverarbeitung

08.00 – 08.45 Bau- und Raumakustik 08.45 - 09.30 Binauraltechnik und virtuelle Akustik	Fels Fels
10.00 – 11.30 Elektroakustik 16.30 – 18.00 Modelle von Außen-, Mittel- und Innenohr 20.30 – 21:30 Abandvorapateltung: Akustiande Täusehungen und	Hudde Hudde
ihr Bezug zur Hörtheorie	Kollmeier
Dienstag: Hördiagnostik	
08.00 – 09.30 Psychoakustik und Sprachperzeption bei Normal- und Schwerhörigen 09.30 – 10.00 Pause	Kollmeier
10.00 – 10.45 Hördiagnostik für die rehabilitative Audiologie 10.45 – 11.30 Audiometrische Standardverfahren, Qualitätssicherung	Kollmeier
in der Hördiagnostik 16.30 – 18.00 Otoakustische Emissionen und auditorisch evozierte Potentiale	Janssen Janssen
Mittwoch : Akustische und physiologische Grundlagen	
08.00 – 09.30 Physikalische Grundlagen der Akustik 09.30 – 10.00 Pause	Blau
10.00 – 10.45 Numerische Methoden in der Akustik	Blau Blau
16.30 – 18.00 Signalverarbeitung für Hörhilfen und Audiologie	Dillier
Donnerstag: Rehabilitative Audiologie & Anwendungen	
08.00 – 09.30 Cochlea-Implantate und Implantierbare Hörgeräte 09.30 – 10.00 Pause	Dillier
10.00 – 11.30 Aktuelle Entwicklungen implantierbarer Hörsysteme 16.30 – 18.00 Versorgung und Rehabilitation mit technischen Hörhilfen	Büchner Kiessling
20.30 – 21.30 Ausbildungen und Berufsfelder in der Audiologie und Medizinischen Physik	Kiessling/ Welker/ Lenarz
Freiter, Drevie - Abeebluß	

### Freitag: Praxis + Abschluß

08.00 - 09.30 Kurzvortrag/ Einführung in die Praxisphase von 4 Firmen

- Standard-Audiometrie (Innoforce, C. Wille)
- Computergestützte Verfahren (HörTech, Michael Buschermöhle)
- Hörgeräte (Phonak, S. Kalis)
- Cochlear Implantate und implantierbare Hörgeräte (Med-El, Marcus Schmidt)

10.00 –10.45 Praxisphase I: Teilnehmer verteilt in 4 Kleingruppen führen praktische Demonstrationen bei den vier Stationen durch

10.45 –11.30 Praxisphase II

11.30-12.15 Praxisphase III

12.15–12.45 Praxisphase IV

Ausgabe der Teilnahmebescheinigungen: Kollmeier, Welker

# Sponsoren

# Die Winterschule wird unterstützt von den Firmen:

Sennheiser AG, Am Labor 1, 30900 Wedemark 2

Cochlear AG Basel Margarethenstraße 47, CH-4053 Basel, Schweiz

Phonak AG, Laubisrütistraße 28, CH-Stäfa, Schweiz

Widex Hörgeräte GmbH, Albstadtweg 6, 70567 Stuttgart

# Bau- und Raumakustik

Janina Fels / Aachen

# 1 Raumakustik

# 1.1 Einleitung

Es ist eine allgemeine Erfahrung, dass man in manchen Sälen die Sprachverständlichkeit sehr gut ist, in anderen dagegen nicht. Sieht man davon ab, wie deutlich der Redner spricht und wie laut der Lärmpegel ist, so bleibt noch ein sehr erheblicher Einfluss des Raumes auf die Sprachverständlichkeit aufgrund seiner akustischen Eigenschaften, oder wie man oft kurz sagt, aufgrund seiner Akustik.

Ähnlich liegen die Dinge in einem Opernhaus oder in einem Konzertsaal. Auch hier gibt es Beispiele für gute oder schlechte Akustik. Wir erwarten in einem Konzertsaal einen schönen, runden, räumlich wirkenden Orchesterklang als ein Ganzes, und es ist die Aufgabe der Raumakustik herauszufinden, was das eigentlich bedeutet und mit welchen greifbaren physikalischen Sachverhalten dies korreliert. Auch in einem Arbeitsraum, in einer Fabrikhalle oder einem Großraumbüro hängt der Lärmpegel von der Art und der Leistung der Schallquelle ab, zu einem erheblichen Teil aber auch davon, was der Raum aufgrund seiner akustischen Eigenschaften daraus macht.

Die Aufgabe der Raumakustik ist also zum Teil physikalischer Art, nämlich aus den baulichen Daten, also aus den Abmessungen, den verwendeten Materialien, der Anordnung des Publikums usw. das Schallfeld im Raum zu ermitteln. Zum anderen hat sie zu quantifizieren, welche objektiven Schallfelddaten oder Schallfeldparameter für gute oder weniger gute Hörbedingungen an den einzelnen Zuhörerplätzen verantwortlich sind. Der zweite Teil der Aufgabe ist nur durch Befragungen von Zuhörern oder durch systematische Hörversuche zu lösen. Schließlich werden von dem Raumakustiker - gewissermaßen durch Rückverfolgung dieser Aufgabenkette - Auskünfte darüber erwartet, wie ein Raum baulich gestaltet werden muss, bzw. welche Änderungen an einem geplanten oder existierenden Raum vorgenommen werden müssen, um die akustischen Erwartungen der Zuhörer zu erfüllen.

# 1.2 Nachhall und Nachhallzeit

Die physikalisch einwandfreie Beschreibung räumlicher Schallfelder geht von der Wellengleichung aus, deren Lösungen den durch Lage und Beschaffenheit der Wände bestimmten Randbedingungen anzupassen sind. Sie muss für praktische Zwecke ausscheiden, da sie auch mit den heutigen Mitteln rechnerisch nicht zu bewältigen ist und überdies weit mehr Informationen liefern würde, als für die akustische Beurteilung eines Raumes sinnvoll wäre. Eine für die raumakustische Praxis wesentlich brauchbarere Beschreibung geht im einfachsten Fall von einer Energiebilanz aus. Dem zufolge ist die zeitliche Energieänderung in einem Raum mit dem Volumen V gegeben durch die Leistung *P* der Schallquelle, vermindert um die Energie, die pro Zeit- und Flächeneinheit an der Raumwand absorbiert wird:

Es ist *w* die Schallenergiedichte, *c* die Schallgeschwindigkeit, *S* die gesamte Wandfläche;  $\alpha$  ist der Absorptionsgrad der Wände, definiert als das Verhältnis der nicht reflektierten Energie zur einfallenden Energie; Überstreichung kennzeichnet den gewogenen Mittelwert über alle Wandteile.

Für dw/dt = P = const. ist die stationäre Energiedichte *w* und der Schalldruckpegel *L* wird:

$$w = \frac{4P}{cS\overline{\alpha}} = \frac{4P}{cA}$$
 Gl. 1-2

bzw.

$$L = 10\log\frac{\tilde{p}^2}{p_0^2} = 10\log\frac{w}{w_0} = 10\log\frac{4P}{cAw_0}, \quad w_0 = 3.4 \cdot 10^{-10} \,\text{J/m}^3$$
Gl. 1-3

Die Größe  $A = S\alpha$  wird als äquivalente Absorptionsfläche bezeichnet. Sie ist eine ebenfalls wichtige Grundgröße der Bauakustik (s. Abschn. 2.2). Dagegen lautet die Lösung der Differentialgleichung (Gl. 1-1) für P = 0:

 $w(t) = w_0 \exp\left(\frac{cS}{4}\overline{\alpha}t\right)$  Gl. 1-4

 $(w_0 = Anfangswert der Energiedichte zur Zeit t = 0.)$  Die obigen Überlegungen gelten aber nur für den Fall des "diffusen Schallfelds", in dem alle Raumrichtungen (im Mittel) gleichermaßen an der Schallausbreitung beteiligt sind. In der Praxis ist die Voraussetzung oft nur näherungsweise erfüllt, wenn überhaupt. Erfahrungsgemäß gilt Gl. 1-4 auch dann einigermaßen, nicht aber die Gl. 1-2. Zur Berechnung des Lärmpegels in einem großen oder flachen oder langen Arbeitsraum muss man daher die Gl. 1-2 durch eine genauere Berechnungsmethode ersetzen.

Das ist für die Raumakustik besonders wichtig, da die Dauer des durch Gl. 1-4 beschriebenen Abklingvorgangs, d.h. des so genannten Nachhalls, bei den verschiedenen Frequenzen die wichtigste akustische Kenngröße eines Raumes darstellt. Man kennzeichnet sie durch die "Nachhallzeit" *T*. Das ist die Zeit, in der die Energiedichte auf das 10<sup>-6</sup>-fache ihres Anfangswertes (um 60 dB) abgesunken ist. Sie ergibt sich aus Gl. 1-4 nach Einsetzen der Konstanten und Näherung für kleine mittlere Absorptionsgrade zu

$$T = \left(0,16\frac{\mathrm{s}}{\mathrm{m}}\right) \cdot \frac{V}{S\overline{\alpha}} = \left(0,16\frac{\mathrm{s}}{\mathrm{m}}\right) \cdot \frac{V}{A}$$
Gl. 1-5

Diese Formel erlaubt bei bekannten Raumdaten eine meist hinreichend genaue Berechnung der Nachhallzeit.

Der Sprachverständlichkeit sind lange Nachhallzeiten grundsätzlich abträglich, da der Nachhall die Tendenz hat, die zeitliche und spektrale Struktur der Sprachlaute und der Silben zu verwischen. Da Räume mit völlig fehlendem Nachhall aber nur mit großem Kostenaufwand zu realisieren sind und auch andere Nachteile haben, gilt im Sinne eines Kompromisses eine Nachhallzeit von 0,6 s bis 1 s als optimal für Vortrags- oder Hörsäle, für Sitzungssäle usw. Bei Sprechtheatern lässt man in der Regel Nachhallzeiten bis etwa 1,2 s zu. Für die Sprachverständlichkeit ist weiterhin vorteilhaft, wenn die Nachhallzeit bei tiefen Frequenzen eher noch niedriger ist als im mittleren Frequenzbereich.

Günstige Bereiche der Nachhallzeit			
Art der Darbietung	mittlere Nachhallzeit in s		
Sprache (Hörsaal, Sitzungsraum)	< 1		
Sprechtheater	1,0 - 1,2		
Konzertsaal	1,7 - 2,1		
Oper	1,0 - 1,8		

# 1.3 Weitere raumakustische Kenngrößen

Eine einigermaßen korrekte Nachhallzeit ist zwar ein notwendiges, nicht aber ein hinreichendes Kriterium für gute Akustik. Außerdem charakterisiert sie eine allgemeine akustische Eigenschaft des Raumes, nicht aber die meist unterschiedlichen Hörbedingungen an verschiedenen Zuhörerplätzen. Weitere charakteristische Parameter lassen sich auf der Grundlage der so genannten Raumimpulsantwort gewinnen. Hierunter versteht man das Signal, das an einem Empfangsort auftritt, wenn der Schallsender einen sehr kurzen Impuls aussendet (s. Bild). In diesem Diagramm kennzeichnet die erste Spitze den so genannten Direktschall, das ist der Schallanteil, der sich auf dem kürzesten Weg von dem Schallsender zum Empfänger ausbreitet. Alle anderen Spitzen kennzeichnen nach Lage und Höhe den Beitrag der Schallanteile, die an der Decke und den Wänden des Raumes einmal, zweimal oder noch öfter reflektiert worden sind.

Da bei werden diese Anteile nach Maßgabe Absorptionsgrads des betreffenden der Wandteile abgeschwächt; eine weitere Abschwächung und außerdem eine mehr oder weniger große Verzögerung gegen-



über dem Direktschall ergibt sich daraus, dass diese reflektierten Schallanteile größere Wege zurückzulegen haben.

Hinsichtlich ihrer subjektiven Wirkung kann man im Allgemeinen in einer Raumimpulsantwort nach dem Bild drei Teile unterscheiden: Zum einen den Direktschall (a), der nach dem Gesetz der ersten Wellenfront die Richtung bestimmt, aus dem das Schallereignis gehört wird. Der zweite Bereich (b) umfasst die wenig verzögerten, noch weitgehend getrennt eintreffenden frühen Reflexionen. Schließlich folgen in immer kürzeren Abständen die zahlreichen schwächeren und länger verzögerten Reflexionen, die den eigentlichen Nachhall ausmachen (c). Da der letztere durch die Nachhallzeit meist ausreichend charakterisiert ist, sind die für eine bestimmte Hörposition im Raum maßgebenden und über die Nachhallzeit hinausgehenden akustischen Kennzeichen in dem zwischen 0 und 80 ms liegenden Bereich zu suchen.

Die Bemühungen der Raumakustiker konzentrierten sich in der zurückliegenden Zeit darauf, aus solchen gemessenen oder auch vorausberechneten Impulsantworten Parameter abzuleiten, welche - ergänzend zur Nachhallzeit - die akustische Qualität eines Saales oder auch nur bestimmter Plätze oder Platzgruppen in ihm kennzeichnen. Stellvertretend für viele andere sei hier die "Deutlichkeit" erwähnt, die als Maß für die Verständlichkeit von Sprache oder der Durchsichtigkeit von Musikdarbietungen dienen kann. Sie ist durch

$$D = \frac{\int_{0}^{50ms} E(t)dt}{\int_{0}^{\infty} E(t)dt}$$
Gl. 1-6

definiert. Dabei kennzeichnet E(t) die Energie in der Impulsantwort, die zur Zeit tam Empfangsort eintrifft. Dieser Definition liegt die Erkenntnis zugrunde, dass die um weniger als 50 ms gegenüber dem Direktschall verzögerten Schallreflexionen subjektiv dem Direktschall zugeschlagen werden, also dessen Lautstärke zu erhöhen scheinen. Diese Bestandteile der Impulsantwort werden daher oft als "nützliche Reflexionen" bezeichnet. Eine weitere, besonders für Konzertsäle wichtige Erkenntnis ist, dass die Richtungsverteilung der im Zeitbereich zwischen 0 und etwa 80 ms beim Hörer eintreffenden Reflexionen für den subjektiven Räumlichkeitseindruck maßgebend ist, der ein wesentliches Qualitätsmerkmal ist. Aufgrund entsprechender psychoakustischer Untersuchungen lässt sich dieser durch den so genannten "Seitenschallgrad" gut kennzeichnen:

$$LF = \frac{\int_{5ms}^{80ms} E(t)\cos^2 \Theta dt}{\int_{0}^{80ms} E(t)dt}$$
Gl. 1-7

Hierin ist *E* die Energie der Reflexionen und  $\Theta$  der Winkel, den Einfallsrichtung mit einer durch die beiden Ohren des Zuhörers gelegten Achse bildet. Die Integrale

sind über alle Reflexionen einschließlich des Direktschalls mit Verzögerungszeiten von weniger als 80 ms zu erstrecken.

Die Ergebnisse zahlreicher, von verschiedenen Forschern und Forschergruppen durchgeführter psychoakustischer Untersuchungen lassen sich dahingehend zusammenfassen, dass bei Sälen, die für irgendwelche akustischen Darbietungen bestimmt sind, die Nachhallzeit nach wie vor an erster Stelle unter den objektiven raumakustischen Kriterien zu sehen ist. Welche weiteren Kriterien von Bedeutung sind, hängt in erster Linie von der Art der Darbietung (Sprache oder Musik) ab.

# 1.4 Schallabsorption und Schallabsorber

Für die Schallabsorption an Wänden gibt es im Wesentlichen zwei Ursachen: die Bewegung von Luft in Poren des Wandmaterials (Verluste durch Reibung) und das Mitschwingen von Wänden im Schallfeld.

# 1.4.1 Poröse Absorptionsmaterialien

Der für das Absorptionsvermögen eines porösen Materials (Watte, Glaswolle, Filz u.ä.) charakteristische längenspezifische, äußere Strömungswiderstand  $\Xi$  wird durch einen Versuch mit konstanter Luftströmung gemäß folgender Formel bestimmt:

$$-\Delta p = \Xi \cdot \Delta x \cdot v_{\rm a}$$

wobei  $v_a$  die äußere Strömungsgeschwindigkeit ist. Im Inneren einer Probe herrscht die um 1/ $\sigma$  vergrößerte Strömungsgeschwindigkeit  $v_i$ , da sich die Stromlinien auf den kleineren Porenquerschnitt zusammenziehen müssen ( $\sigma$  = Porosität < 1):



Noppenschaumstoff zur Auskleidung von Raumbereichen



Steinwolle für raum- und bauakustische Zwecke

Da in praktischen Fällen das poröse Material sich meist als Schicht endlicher Dicke vor einer starren Wand befindet, wird ein Teil der eindringenden Wellen an dieser reflektiert und interferiert an der Grenzfläche mit der auftreffenden Welle. Dadurch ergeben sich häufig Schwankungen des Absorptionsgrades bei Variation der Frequenz. Eine merkliche Absorption wird erst erreicht, wenn der erste Schnellebauch der stehenden Welle vor der Wand im porösen Material liegt, dieses somit mindestens  $\lambda/4$  dick ist. Man spart also Absorptionsmaterial, wenn man dieses nicht direkt auf die Wand bringt, sondern dahinter einen Zwischenraum frei lässt.

# 1.4.2 Resonanzabsorber

Fällt eine Schallwelle auf eine bewegliche Wand, so wird diese im Allgemeinen zum Schwingen angeregt. Zugleich wird auf der abgewandten Seite eine Schallwelle abgestrahlt. Ist *m*'' die Flächenmasse der Wand (in kg/m<sup>2</sup>), dann ist ihr Absorptionsgrad bei senkrechtem Schalleinfall

$$\alpha = \frac{1}{1 + \left(\frac{\omega m''}{2\rho_0 c}\right)^2}$$

Er fällt mit wachsender Frequenz monoton ab - eine Folge der Massenträgheit. In der Praxis macht sich diese Art der Absorption nur bei dünnen Wänden (Fensterscheiben, Folien) bemerkbar und auch da nur bei tiefen Frequenzen. Wichtiger ist dagegen die Absorption einer schwingungsfähigen Wand, die im Abstand *d* vor einer schallharten Wand montiert ist. Der Luftraum der Dicke d wirkt wie eine Feder mit der flächenbezogenen Nachgiebigkeit

$$n'' = \frac{d}{\rho_0 c^2}$$

Bei der Frequenz

$$f_0 = \frac{1}{2\pi} \frac{1}{\sqrt{m''n''}} = \frac{1}{2\pi} \sqrt{\frac{\rho_0 c^2}{m'' d}}$$

wird die Schwingungsamplitude sehr groß, und unvermeidbare Verluste z.B. durch Biegung der Verkleidung (elastische Verluste) bilden einen Absorptionsmechanismus. Um eine höhere Absorption zu erhalten, kann man den Hohlraum zusätzlich ganz oder teilweise mit porösem Material mit der Verlustkonstante w (Reibungswiderstand) füllen. Derartige Resonanzabsorber spielen in der Raumakustik eine große Rolle, da man mit ihnen durch geeignete Dimensionierung Wände mit weitgehend vorgebbarer Absorption herstellen kann. Auch für Loch- oder Schlitzplatten kann man eine effektive mitschwingende Masse m'' berechnen. Diese Masse entspricht der Masse der Luft in den Löchern.

$$m'' = \frac{\rho_0}{\sigma} (b + 1, 6a)$$



Plattenresonator

Loch- / Schlitzabsorber

Helmholtz-Resonator

Hierin ist b die Plattendicke, a der Radius der als kreisförmig angenommenen Löcher und  $\sigma$  der "Perforationsgrad", d.h. der Bruchteil der Öffnungsfläche an der Gesamtfläche. Durch starke Perforation lassen sich also auch Schichten mit kleinen Flächenmassen herstellen, die zusammen mit einer porösen Schicht einen Resonator mit relativ hoher Resonanzfrequenz und wenig ausgeprägter Resonanz ergeben. Daher eignen sich dünne, hochperforierte Platten aus Blech, Gips oder Holz zur Abdeckung von Schluckstoffschichten (Mineralwolle oder ähnliche Stoffe), was praktisch aus Gründen des Berührungsschutzes, aber auch eines besseren Aussehens erforderlich ist. Diese Form findet oft Anwendung als sog. "Akustik-Decke". Dies ist typischerweise folgender Aufbau (von oben gesehen): massive Rohdecke - Luftraum - Mineralwolle - Lochplatte.



In der Vergangenheit fanden ebenfalls so genannte mikroperforierte Absorber zunehmend Beachtung. Diese bestehen aus einer Folie (z.B. transparenter Kunststoff) oder Platte (z.B. Acrylglas) mit sehr kleinen Löchern (d = 0,05 - 0,5 mm). In Abhängigkeit der Anwendung ist der Lochungsgrad im Bereich von 0,5 - 2 %. Bei solchen Anordnungen ist die Absorption hauptsächlich auf die viskose Reibung der Luftmoleküle an den Wänden der Mikro-Lochung zurückzuführen. Da mikroperforierte Absorber auch aus transparentem Material hergestellt werden können finden sie oft in Situationen Einsatz, in denen aus ästhetischen Gründen die Nutzung anderer Absorptionsmechanismen unerwünscht ist.

# 2 Bauakustik

# 2.1 Einleitung

Die Bauakustik beschäftigt sich mit der Schall- und Schwingungsausbreitung in Gebäuden sowie mit der Beurteilung des Schallschutzes in Gebäuden. Dabei ist Luftschall und Körperschall hinsichtlich der Anregung und der Fortpflanzung in Gebäude zu unterscheiden. Beispielsweise würde eine laut eingestellte Hifi-Anlage zunächst Luftschall im "Senderaum" abstrahlen. Die Schallübertragung in den sog. "Empfangsraum", der ein schutzbedürftiger Raum im Sinne des baulichen Lärmschutzes sein kann (z.B. ein Schlafzimmer in fremdem Wohnbereich) erfolgt dann über vielfältige Wege, z.B. als Luftschall über Lüftungskanäle, als Körperschall direkt über die Trennwand oder als Umwandlungen von Luft- in Körperschall und umgekehrt. Körperschall, der durch Gehen auf Decken erzeugt wird, nennt man "Trittschall". Ebenfalls von Interesse ist Körperschall aus haustechnischen Anlagen (Frisch- und Abwasserinstallation, Klimaanlagen, etc.) sowie die Schalldämmung gegen Außenlärm (Industrielärm, Straßen- und Luftverkehrslärm).

# 2.2 Luftschalldämmung

Eine wichtige Größe ist das Schalldämmmaß  $R = -10\log \tau$  mit dem Transmissionsgrad

 $\tau = I_d/I_0$ . In der Praxis werden die einfallende und die durchgelassene Schallintensität allerdings nicht unmittelbar gemessen, sondern die Messung des Schalldämmmaßes erfolgt zwischen zwei Räumen mit einem indirekten Verfahren. Dann gilt:

$$R = L_s - L_E + 10\log S / A \qquad \qquad \text{Gl. 2-1}$$

 $(L_{S} = Senderaum-Schalldruckpegel, L_{E} = Empfangsraum-Schalldruckpegel, S = Trennbauteilfläche, A = äquivalente Absorptionsfläche des Empfangsraumes)$ 





Messungen dieser Art werden in Gebäuden durchgeführt, um festzustellen, ob der gesetzlich vorgeschriebene Mindestschallschutz gewährleistet ist. Andererseits gibt es Labor-Prüfverfahren, die auf ähnlichen messtechnischen Grundlagen stehen und die dazu dienen, die Schalldämmung von Bauteilen oder Baukonstruktio-

nen zu ermitteln. Letztere Daten können verwendet werden, um bereits bei der Planung von Bauten die schalltechnische Eignung nachzuweisen.

Der wesentliche physikalische Unterschied dieser beiden Betrachtungsweisen liegt in der Tatsache, dass im Prüfstand nur jeweils eine Konstruktion beurteilt wird, dass aber im fertigen Gebäude alle Übertragungswege über alle trennenden und flankierenden Bauteile eine Rolle spielen. Letztere Tatsache kennzeichnet man dadurch, dass man das Schalldämmmaß mit einem Apostroph (*R*' statt *R*) versieht.

Die Ergebnisse der Schalldämmmaße werden im Frequenzbereich von 100 Hz – 3,15 kHz in Terzen dargestellt und beurteilt. Einfacher und letztlich relevant ist das sog. bewertete Schalldämmmaß  $R_w$  (oder  $R'_w$ ), welches durch einen Vergleich der gemessenen Schalldämm-Kurve mit einer international genormten Bezugskurve ergibt.



# 2.3 Berechnung der Schalldämmmaße einzelner Bauteilea) einschalige Bauelemente



Die Einfachwand und ihr elektrisches Analogon

$$R = 10\log(\frac{p_{\text{ein}}^2}{p_{\text{durch}}^2}) = 10\log\left|1 + \left(\frac{\omega m'}{2Z_0}\right)^2\right| \approx 20\log\left(\frac{\omega m'}{2Z_0}\right) \qquad \text{Gl. 2-2}$$

Dieses sog. "Massegesetz" gilt für biegeweiche Bauteile, wobei  $\omega$  die Kreisfrequenz, *m*' die flächenbezogene Masse und  $Z_0 = \rho_0 c$  die Kennimpedanz von Luft ist. Bei Schalleinfall unter dem Winkel  $\beta$  wird das Argument des Logarithmus mit dem Faktor "cos  $\beta$ " ergänzt. *R* steigt um 6 dB pro Oktav (doppelte Frequenz) oder pro Masseverdopplung an.

#### b) zweischalige Bauelemente



Die Doppelwand und ihr elektrisches Analogon

Bei zweischaligen Elementen wird das Massegesetz von einem wesentlich stärkeren Effekt überlagert, der auf das Zusammenwirken zweier Masse-Schalen auf einer Luftschicht, die als Feder wirkt, zurückzuführen ist.

$$R = 20 \log \left[ \frac{\omega^3 dm'_1 m'_2}{2\rho_0^2 c^3} \right], \quad \omega_0 = \sqrt{\frac{m'_1 + m'_2}{n' m'_1 m'_2}}$$
Gl. 2-3



Schalldämmmaß einer Doppelwand

Dabei tritt zwar bei der Resonanzfrequenz  $\omega_0$  eine Verschlechterung ein, aber im eigentlichen Nutzbereich steigt das Schalldämmmaß steil (18 dB/Oktav) an, bis die Wellenlänge so klein wird, dass statt der Federwirkung der Luftschicht stehende Wellen im Zwischenraum auftreten. In der Praxis findet man "Doppelwände" bei Haustrennwänden, Leichtbauwänden (z.B. Gipskartonplatten), im Fertighausbau, bei Fenstern und bei Vorsatzschalen zur Verbesserung der Schalldämmung. Wichtig sind eine gute Körperschallentkopplung der Schalen (keine Schallbrücken!) und eine geeignete Füllung des Hohlraums mit absorbierendem Material zur Dämpfung des Resonanzeinbruches.

# 3 Binaurales Hören

## 3.1 Dreidimensionale Schallwahrnehmung

Die Richtung eines Schallereignisses kann von normalhörenden Menschen aufgrund charakteristischer Merkmale identifiziert werden. Betrachtet wird zunächst das Schallfeld einer einzelnen Schallquelle im Freifeld, also ohne weitere Reflexionen von Wänden oder sonstigen Hindernissen. Ist in diesem Feld eine einzelne Person vorhanden, wird die sonst ungehinderte Ausbreitung des Schallfeldes durch Reflexionen und Beugungseffekte an Rumpf, Schulter, Kopf und Ohrmuschel gestört, das Schallsignal der Quelle wird auf dem Weg in die beiden Ohren dadurch linear verzerrt. Diese Verzerrungen sind insbesondere abhängig von der Einfallsrichtung des Schalls bezogen auf den Körper und dessen Abstand zur Schallquelle.

Der Schall einer seitlich angeordneten Schallquelle erreicht beide Ohren nach unterschiedlicher Laufzeit, weil der Schall zum abgewandten Ohr einen längeren Weg zurücklegen muss als zum zugewandten. Außerdem wird das abgewandte Ohr durch den Kopf akustisch abgeschattet. Die Schalldrücke an beiden Ohren unter-



scheiden sich also. Die beiden Abweichungen werden durch die "Interaurale Laufzeitdifferenz" und die "Interaurale Pegeldifferenz" quantifiziert. Bei tiefen Frequenzen ist der Einfluss auf die Schallausbreitung weniger ausgeprägt, da der Kopf gegenüber der Wellenlänge klein ist, die Pegelunterschiede sind dann eher gering. Daher wird für tiefe Frequenzen eher die Laufzeitdifferenz ausgewertet, bei höheren Frequenzen ist dann die Pegeldifferenz entscheidend.

Liegt die Schallquelle aber in der Medianebene, d.h. in der Ebene senkrecht durch den Kopf, auf der die Ohrachse senkrecht steht, dann treten die oben genannten Effekte nicht auf (unter Annahme eines symmetrischen Kopfes). Aber auch in dieser Ebene ist die Schallquellenlokalisation möglich.

Die Laufzeiten der Reflexionen von Kopf, Rumpf, Schulter und Ohrmuschel sind abhängig vom Elevationswinkel, dem Winkel zwischen der Einfallsrichtung und der horizontalen Ebene. Daraus ergibt sich ein charakteristisches Frequenzverhalten des Übertragungsweges von der Schallquelle bis zum Trommelfell. Dieses wird von Gehirn ausgewertet um auf den Elevationswinkel zu schließen.

Die oben beschriebenen linearen Verzerrungen auf dem Weg von der Schallquelle bis zum Trommelfell des linken und rechten Ohrs lassen sich für verschiedene Richtungen in reflexionsarmer Umgebung messen. Man erhält dann eine Sammlung von kopfbezogenen Übertragungsfunktionen, die man nach ihrer englischen Bezeichnung auch HRTF nennt (Head-Related Transfer Function). Wird das Signal einer Schallquelle mit der HRTF für eine bestimmte Einfallsrichtung gefiltert, und anschließen geeignet wiedergegeben (siehe weiter unten), wird die Schallquelle unter der entsprechenden Richtung wahrgenommen. Das funktioniert allerdings nur, wenn die Person mit den verwendeten HRTFs kompatibel ist.

# 3.2 Außenohr-Übertragungsfunktionen: HRTF Head-Related Transfer Function)



Das System "Außenohr" kann als LTI-System behandelt werden. Dementsprechend wird es mit Hilfe einer Stoßantwort oder einer komplexen Übertragungsfunktion eindeutig beschrieben. Die Außenohr-Übertragungsfunktion enthält die Information über die linearen Verzerrungen, die durch das Vorhandensein des Körpers, des Kopfes, der Ohrmuschel, des Gehörganges usw. hervorgerufen werden.



Einige Anteile sind richtungsabhängig, andere nicht. Als Bezugsschallfeld dient im einfachsten Fall eine ebene Schallwelle, die von vorne auf den Kopf einfällt. Zur exakten Angabe der Richtung muss man jedoch genauere Angaben machen, wobei man sich eines kopfbezogenen Koordinatensystems bedient. Die Horizontalebene wird von einem um den Kopf laufenden Kreis von  $\varphi = 0^{\circ}$  bis 360° und  $\vartheta = 0^{\circ}$  aufgespannt; die Medianebene entsprechend bei  $\varphi = 0^{\circ}$  und von  $\vartheta = 0^{\circ}$  bis  $\vartheta = 360^{\circ}$ . Zur vollständigen Beschreibung einer Schalleinfallsrichtung reicht somit die Angabe zweier Winkel aus. Die oben erwähnte "Vorne"-Richtung wird durch  $\varphi = 0^{\circ}$  und  $\vartheta = 0^{\circ}$  festgelegt.



Kopfbezogenes Koordinatensystem

Einige typische Außenohr-Übertragungsfunktionen (HRTF) von vier Versuchspersonen, gemessen für die "vorne"-Richtung.

#### 3.3 Kunstkopftechnik

Die Messung binauraler Signale muss nach dem zuvor Gesagten alle richtungsabhängigen Anteile einer Außenohr-Übertragungsfunktion beinhalten, denn richtungsunabhängige Anteile, die von Komponenten hinter dem Ohrkanal-Eingang stammen, können auch durch elektrische (typischerweise digitale) Filter nachgebildet werden. Für Laborzwecke ist es daher zunächst möglich, mit einer Versuchspersonengruppe (evtl. auch gezielt individuell) und Gehörgangs-Miniaturmikrofonen oder Sondenmikrofonen zu arbeiten. Hierbei werden typischerweise Mittelwerte über eine Anzahl von Versuchspersonen gewonnen und ausgewertet. Für die objektive binaurale Geräuschmesstechnik ist jedoch der Ersatz der Versuchspersonengruppen durch ein spezielles zweikanaliges Richtmikrofon wichtig, wodurch der Zeit- und Kostenaufwand sinkt und die Reproduzierbarkeit der Messergebnisse erhöht wird. Ein solches Richtmikrofon heißt "Kunstkopf" oder auch "Kopf- und Rumpfsimulator".

Die Entwicklung eines Kunstkopfes bedarf zahlreicher Detailuntersuchungen der Formen von Kopf, Rumpf, Ohrmuschel usw. Entweder versucht man, eine mittlere Geometrie zu finden, oder man stellt einen repräsentativen Kopf- und Rumpfsimulator her, der einer typischen Versuchsperson ähnelt. Die genaue Ausführung des Gehörganges und die Positionierung der Mikrofone sind für die richtungsabhängigen Anteile der Außenohr-Übertragungsfunktion nicht relevant. Eine geeignete Entzerrung der Ohrsignale ist für verschiedene Anwendungen ohnehin notwendig, sei es für Kopfhörerwiedergabe oder für messtechnische Zwecke in vorgegebenen Schallfeldern.



Drei Ausführungen standardisierter Kopf- und Rumpfsimulatoren

Geeignete Bezugsschallfelder sind speziell das "Freifeld" und das "Diffusfeld". Dementsprechend meint man bei einem freifeldentzerrten Kunstkopf, dass sein Übertragungsmaß bei Beschallung unter  $\varphi = 0^{\circ}$  und  $\vartheta = 0^{\circ}$  horizontal verläuft (d.h. unabhängig von der Frequenz konstant ist). Diese Eigenschaft ist daher auch ein wichtiges Prüfkriterium bei Eignungstests von Kunstköpfen. Unter Bezug auf die "Vorne-Richtung" können nämlich die Übertragungsfunktionen anderer Einfallsrichtungen oder auch das Übertragungsmaß für diffuse Beschallung beurteilt und für Qualifikationstests verwendet werden.

Die Geometrie und das Übertragungsverhalten von Kunstköpfen im freien und diffusen Schallfeld ist in internationalen Normen festgelegt (ISO TR 959 und ITU p.58). Beispiele für Anwendungsbereiche sind Messungen von Hörgeräten und von Telefonen (Hören und Sprechen) sowie die Raumakustik und der wichtige Bereich "Sound Quality", der in der Industrie mehr und mehr an Bedeutung gewinnt (z.B. im Fahrzeugbau und bei Haushaltsgeräten).

# 3.4 Binaurale Messgrößen

Binaurale Messungen haben oft nicht nur den Schalldruckpegel oder das Trommelfell-Schalldruckspektrum zum Ziel, sondern zunehmend auch komplexere Funktionen, die spezielle Eigenschaften des menschlichen Gehörs beschreiben sollen. Als Beispiele für binaurale Messgrößen seien hier nur einige Parameter erwähnt:

• ILD, ITD (Interaural Level Difference, Interaural Time Difference): werden allgemein als die wesentlichen Merkmale binauraler Schallreize angesehen. Die an den beiden Ohren auftretende Pegel und Laufzeitunterschiede werden im neuronalen System in verschiedenen Entscheidungsstufen ausgewertet. Spezielle Nervenzellen sind insbesondere in der Lage, "Gleichzeitigkeit" festzustellen bzw. die Zeitbezüge zwischen "links" und "rechts" zu messen, so dass aus ITD unmittelbar die azimutale Schalleinfallsrichtung ermittelt werden

- BMLD (Binaural Masking Level Difference): Pegelschwelle, bei welcher ein Testsignal unter binauralen Abhörbedingungen (evtl. mit einem Maskierer) unterschiedlich wahrgenommen wird. Die Variationen des Testschallfeldes können sehr vielfältig sein, z.B. können die Maskierer aus unterschiedlichen Richtungen einwirken, oder sie können einem Diffusfeld entsprechen.
- IACC (Interaural Cross Correlation Coefficient): Die interaurale Kreuzkorrelation wird in der Raumakustik als Maß dafür verwendet, zu quantifizieren wie viel Schallanteile am Ort des Zuhörers aus der Medianebene einfallen. Bei Schalleinfall aus der Medianebene ist bekanntlich die Korrelation (Ähnlichkeit) zwischen den Ohrsignalen sehr groß (im Idealfall ist die Korrelation gleich Eins bei τ = 0). Seitlicher Schalleinfall bedeutet eine Verkleinerung und Verschiebung des Korrelationsmaximums. Das Maximum der Korrelationsfunktion zwischen -1 ms < τ < 1 ms wird als Wert des IACC angegeben. Ein geringer I-ACC bedeutet, dass die Akustik eines Raumes durch einen hohen "Räumlichkeitseindruck" gekennzeichnet ist.

# 4 Raumakustische Auralisation

# 4.1 Binaurale Raumimpulsantworten

Impulsantworten, die für die Hörbarmachung des simulierten Schallfeldes geeignet sind, müssen eine Abtastrate von typischerweise 44,1 kHz aufweisen. Daher sind nur Spiegelschallquellenansätze oder gemischte Verfahren einsetzbar.

Natürlich muss "Auralisation" das zweichrige (binaurale) Hören berücksichtigen. Monophone Signale können zwar einen Eindruck über das verhallte Signal verschaffen, aber erst das räumliche Hören vermittelt den Klangeindruck in seiner Gesamtheit. Die Binauraltechnik erlaubt mit Hilfe der sog. "Außenohrübertragungsfunktionen" HRTF (siehe Abschn. 3.1) oder den zugehörigen HRIRs (head-related impulse responses) eine Simulation der einzelnen Schalleinfallsrichtungen. HRTF-Daten werden normalerweise an Versuchspersonen oder an Kunstköpfen gemessen.

Schließlich kann so eine binaurale Schalldruck-Impulsantwort aus der Summe der einzelnen Reflexionen (und dem Direktschall) erzeugt werden. Die binaurale Raumimpulsantwort setzt sich aus den Beiträgen des Direktschalls, falls dieser hörbar ist, und aus den Beiträgen der hörbaren Spiegelschallquellen zusammen. Die Formulierung der Superposition der Spiegelschallquellenbeiträge stellen wir zunächst im Frequenzbereich auf. Dazu wird der Begriff der Raumübertragungsfunktion verwendet. Die Erzeugung der Impulsantwort geschieht also mittelbar über die Erzeugung einer Frequenzfunktion.

Liegt von einem System die Impulsantwort h(t) vor, so kann diese Funktion mit der Fourier-Transformation

$$\boldsymbol{F}{h(t)} = \underline{H}(f) \text{ bzw. } h(t) \bigcirc \underline{H}(f)$$
Gl. 4-4

in die Übertragungsfunktion  $\underline{H}(f)$  überführt werden und mit der inversen fouriertransformation entsprechend umgekehrt. Die Transformationsgleichung ist durch das Fourierintegral gegeben. Somit können LTI-Systeme vollständig entweder im Zeitbereich oder im Frequenzbereich charakterisiert werden, und an mehreren Stellen ist mittels der Fouriertransformation ein Übergang zwischen den Bereichen in beide Richtungen möglich.

# 4.1.1 Übertragungsfunktion

Die Übertragungseigenschaften des Raumes bzw. der einzelnen Reflexionen lassen sich im Frequenzbereich durch sog. Frequenzfunktionen  $\underline{H}(f)$  darstellen, und

zwar entweder durch die Komponenten Real- und Imaginärteil (Re{*Hf*)} bzw. Im{*H*(*f*)} oder in der äquivalenten Form als Betrag und Phase (|*H*(*f*)| bzw.  $\varphi(f)$ )

$$\underline{H}(f) = \operatorname{Re}\{\underline{H}(f)\} + j\operatorname{Im}\{H(f)\} = |H(f)| \cdot e^{j\varphi(f)}$$
Gl. 4-5

Das Spektrum des Direktschalls wird dabei über folgende Formel gebildet:

$$\underline{H}_{\text{direkt}}(f) = H_{\text{Sender}}(f, \overline{r}_{S}) \cdot H_{\text{Luft}}(f, d) \cdot \underline{H}_{\text{Empf}}(f, \overline{r}_{E})$$
Gl. 4-6

Dabei ist  $H_{\text{Sender}}$  das Spektrum der Senderrichtcharakteristik und ist abhängig von der Richtung, in die der Schallstrahl vom Sender aus startet.  $H_{\text{Luft}}$  ist das Spektrum der Luftdämpfung, abhängig von der zurückgelegten Strecke des Schallstrahls, was beim Direktschall gleichbedeutend mit der Entfernung vom Sender ist.  $\underline{H}_{\text{Empf}}$  ist schließlich das komplexe Spektrum der jeweiligen binauralen Außenohrübertragungsfunktion und ist abhängig von der Richtung, aus welcher der Schallstrahl des Direktschalls beim Empfänger eintrifft. Diese Richtung ist beim Direktschall identisch mit der Schallstrahl am Ort des Senders startet.

Der Beitrag der j'ten gültigen Spiegelschallquelle ist:

$$\underline{H}_{j} = \frac{\mathrm{e}^{-jkr_{j}}}{r_{j}} \underline{H}_{\mathrm{Sender}} \underline{H}_{\mathrm{Empf}} \underline{H}_{\mathrm{Luft}} \prod_{i=1}^{n_{j}} \underline{H}_{i}$$
Gl. 4-7

wobei  $r_j$  den Abstand zwischen Spiegelschallquelle und Empfänger, *k* die Wellenzahl =  $2\pi f/c$ , <u>H</u><sub>Sender</sub> die Richtcharakteristik der Quelle, <u>H</u><sub>Empf</sub> die "HRTF" (des rechten oder des linken Ohres, siehe Abschn. 3.1), <u>H</u><sub>Luft</sub> die Luftabsorption und <u>H</u><sub>i</sub> die Reflexionsfaktoren der betroffenen Wände in der Spiegelschallquellenkette bedeutet. Die komplette binaurale Impulsantwort gewinnt man dann durch Addition aller Spiegelschallquellenbeiträge (r, I = rechtes, linkes Ohr) und durch inverse Fourier-Transformation:



$$h_{\text{total,r,l}}(t) = F^{-1} \left\{ H_{\text{direk}t} + \sum_{j=1}^{N} \underline{H}_{j,r,l} \right\}$$
Gl. 4-8

#### 4.2 Verarbeitung trockener Signale

Sobald die relevanten Funktionen, welche die Übertragung des Schalls im Raum beschreiben (Raumimpulsantwort) bekannt sind, und auch die Eigenschaften der

Empfänger in Form der Außenohr-Übertragungsfunktionen HRTF und HRIR richtungstreu angewendet wurden, ist das System Quelle – Raum – Empfänger vollständig repräsentiert. Dessen Übertragungseigenschaften können nun mit beliebigen Anregungssignalen verarbeitet werden. Wird der Raum als lineares zeitinvariantes "LTI-System" aufgefasst, so darf man für den Fall der rückwirkungsfreien Signalübertragung die Filterung des Signals durch die Raumeigenschaften durch eine Faltung vollziehen.

### 4.2.1 Eingangssignale

Das Eingangssignal *s*(*t*) sollte möglichst gut einem realistischen "Quellsignal" entsprechen. Ein Quellsignal kann beispielsweise von einem Musikinstrument, von einem Sprecher oder von einer lärmerzeugenden Maschine stammen. In diesem Zusammenhang bedeutet "Rückwirkungsfreiheit", dass die Leistungsabgabe der Quelle unabhängig von den Raumeigenschaften ist. In der Praxis ist das in der raumakustischen Auralisation immer erfüllt, ganz im Gegensatz zur Körperschallquellen und deren Auralisation.

Ein Quellsignal für die raumakustische Auralisation kann leicht in reflexionsfreier Umgebung aufgenommen und gespeichert werden. Es ist allerdings darauf zu achten, dass sich die Mikrofonposition bei der Aufnahme im Fernfeld befindet, damit keine Nahfeldeffekte und damit verbundene Verzerrungen in das Quellsignal gelangen. Derartige Aufnahmen sind für Musikstücke, für Sprache und auch für typische Lärmbeispiele verfügbar. Sie werden "trockene Signale" genannt, da sie vollkommen nachhallfrei (und somit unnatürlich) klingen. Aufnahmen in relativ gedämpften Raumumgebungen wie z.B. in Tonstudios können als Eingangssignale verwendet werden, sofern der Rest-Nachhall im Signal wesentlich geringer ist als der Hall des zu auralisierenden Raumes.

### 4.2.2 Der Begriff der Faltung

Wird ein LTI-System mit einem Eingangssignal s(t) gespeist, so kann ausgangsseitig ein Signal g(t) empfangen werden, für welches gilt:



$$g(t) = \int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau = s(t) * h(t)$$

wobei h(t) die Impulsantwort (oder Stoßantwort) des Systems genannt wird und das Integral eine Faltungsoperation ausdrückt. Diese allgemeine und sehr wichtige Formel ist die Grundlage für alle theoretischen Betrachtungen an LTI-Systemen. Sie erlaubt insbesondere die Konstruktion von Filtern und die digitale Filterung beliebiger Eingangssignale mit Impulsantworten h(t). Die Beschreibung von Filtern ist aber nach 4.1.1 ebenso im Frequenzbereich möglich. Sollen die Übertragungseigenschaften eines Systems bezüglich der Frequenz dargestellt werden, ist besonders die Information von Interesse, um wel-

$$\underline{S}(f) \circ \underline{H}(f) \longrightarrow \underline{G}(f)$$

ches Maß ein eingespeistes Signal einer bestimmten Frequenz von dem System gedämpft oder verstärkt wird. Die Faltung wird im Frequenzbereich durch eine Multiplikation ersetzt:

$$\underline{G}(f) = \underline{S}(f) \cdot \underline{H}(f)$$
 Gl. 4-10



Ob die Faltung im Zeitbereich "zu Fuß" oder mit Hilfe der Fouriertransformation im Frequenzbereich durchgeführt wird, ist für das Resultat unerheblich. Es kommt letztlich nur darauf, welche Methode schneller ist.

### 4.3 Freifeld-Verfahren: Binauraler Richtungsmischer

Schon unter Freifeldbedingungen gibt es interessante Anwendungen der Echtzeit-Auralisation. Eine elementare Aufgabe besteht darin, eine Schallquelle "in den Raum zu stellen", d.h. eine Quelle (mono) ohne Richtungs- und Entfernungsinformation im dreidimensionalen Raum außerhalb des Kopfes in einer gewissen Entfernung zu platzieren. Ein derartiger "Richtungsmischer" muss dem mehr Signal Merkmale aufprägen als einer einfacher Balance-Regler, der in der Stereobreite Verschiebungen hervorrufen kann. Vielmehr sind die aus Abschn. 3.1 bekannten HRTF zu verwenden, siehe Bild.



Ebenso können Richtungsinformationen in Signalen manipuliert, d.h. in der Richtung geändert werden.

Mit der Möglichkeit, Schallereignisse in den 3D-Raum zu platzieren und zu bewegen, ist das wichtigste Element der virtuellen Akustik bereits beschrieben. Alle akustischen und algorithmischen Details sind nun im Prinzip bekannt und müssen "nur" in Echtzeit realisiert werden.

Die notwendigen Blocklängen der HRTF-Filter liegen im Bereich von 128. Typischerweise reichen FIR-Filter (Transversalfilter) aus, auch ohne den Einsatz von DSP.

### 4.4 Erstellung der binauralen Raumimpulsantworten

Die binaurale Synthese ist auch für Schallfelder in Räumen einsetzbar, wobei die Anteile jeder Reflexion (Spiegelschallquelle) nach Gleichung 4.10 bzw. in der Summe nach 4.11 berechnet werden. Zu beachten ist dabei, dass die entsprechenden HRTF nicht pauschal, sondern für jede Reflexion spezifisch richtungsabhängig eingehen müssen.

# 5 Integration von Binauraltechnik in virtuelle Umgebungen

Die Computersimulation akustischer Szenen ist eine wichtige Grundlage für das Rendering. Die 3D-fähige Wiedergabeeinheit, sozusagen das "Audio-Frontend" oder die "Mensch-Audio-Schnittstelle", ist jedoch ebenfalls ein sehr kritisches Element, welches hohe Qualitätsansprüchen hinsichtlich Klangtreue und Lokalisation erfüllen muss, und zwar ähnlich hoch wie bei modernen Shutterbrillen und entsprechenden Stereo-Projektionsflächen mit Auflösungen von mindestens 2000 x 1600 Pixel. Ein derartiges Wiedergabesystem darf man durchaus als Super-HiFi-Anlage bezeichnen, wobei allerdings im Vergleich mit Techniken im Consumer-Hifi-Bereich völlig andere Synthese- (Aufnahme-) Techniken beachtet werden müssen. Steht bei einer Tonstudio- oder Liveaufnahme ein guter "Sound" und eine ansprechende räumliche Abbildung in der Stereobasis im Vordergrund, so muss bei der Virtuellen Realität (VR) nicht nur ein Stereo-Effekt, sondern eine korrekte 3-D-Abbildung erzielt werden. Eine im Frequenzgang neutrale Wiedergabe ist natürlich ohnehin vorausgesetzt.

# 5.1 Kopfhörer-basierte Wiedergabetechnik

Naheliegend ist die Verwendung von Kopfhörern, die auch weit verbreitet ist. Leider sind mit der Kopfhörerwiedergabe einige Nachteile verbunden. Es kann insbesondere zu "Im-Kopf-Lokalisation" kommen. Da die Wahrnehmung einer Schallquelle im Kopf in der Realität nicht vorkommt, kommt es in einem solchen Fall unweigerlich zu Diskrepanzen zwischen den Wahrnehmungen der verschiedenen Sinne.

Die Entzerrung von Kopfhörern ist weitaus schwieriger als die Entzerrung von Lautsprechern. Sowohl die Herstellung einer repräsentativen Strahlungsimpedanz in Form eines "künstlichen Ohres" (wie auch in der Audiometer-Kalibrierung verwendet) als auch eine reproduzierbare Anpassung des Kopfhörers an das künstliche oder an natürliche Ohren bergen große Messunsicherheiten. Selbst bei idealen Messbedingungen, guten Ergebnissen und wirkungsvoller digitaler Signalverarbeitung zur Entzerrung der linearen Übertragungseigenschaften ist nicht gewährleistet, dass die Bedingungen beim erneuten Aufsetzen des Kopfhörers oder beim Wechsel der Abhörperson bestehen bleiben.

# 5.2 Lautsprecher-basierte Wiedergabetechnik

Die Erzeugung räumlicher Schallfelder mit Lautsprechern kann prinzipiell auf zwei verschiedene Arten erfolgen. Zum einen kann versucht werden, mit Lautsprechern "ohrgerechte" Signale zu erzeugen, ohne auf die speziellen Eigenschaften des Schallfeldes im Abhörraum zu achten. Dies ist zulässig, denn das Hörempfinden hängt ausschließlich (jedenfalls sehr weitgehend) vom Schalldruck am Trommelfell ab und nicht vom großräumigen Wellenfeld. Zum anderen kann jedoch ver-

sucht werden, in einem bestimmten Raumvolumen das zu simulierende Schallfeld tatsächlich nachzubilden. Eine sehr einfache Variante, die auf die Abbildung von Schallquellen in der horizontalen Ebene beschränkt ist und keine echte Schallfeldnachbildung, sondern nur eine Näherung über "Phantomschallquellen", ist Dolby-Surround. Diesem Verfahren deutlich überlegen, aber auf mit extrem hohen Aufwand verbunden, ist die Wellenfeldsynthese (Wave Field Synthesis, WFS). Hier wird tatsächlich ein realistisches Schallfeld erzeugt, das bei Beschränkung auf eine agierende Person auch kopfnahe Schallquellen abbilden kann. Leider wird eine immense Zahl von Lautsprechern und DSP-Filterkanälen benötigt, so dass dieses Verfahren nur schwer große Verbreitung finden kann. Es gibt allerdings Ansätze für Kinobeschallung und High-End Audio. Ein guter Kompromiss für eine räumliche Wiedergabe ist das Verfahren "Ambisonics", eine Art Abbildung des Wellenfeldes aus Kugelfunktionen.

### 5.2.1 Kopfbezogene Verfahren

Mit deutlich weniger Aufwand kann eine dreidimensionale Schallfeldwahrnehmung mit binauralen Techniken (siehe Kapitel 7) erzeugt werden. Dieses Verfahren beruht auf der direkten Berechnung des <u>Schalldrucks am Trommelfell</u> unter Verwendung der oben bereits genannten HRTFs. Durch die Wiedergabeeinrichtung muss dann aber dafür gesorgt werden, dass das Signal für das linke Ohr auch nur beim linken Ohr ankommt und ebenso das für das rechte.



Alternativ zu Kopfhörern können binaurale Signale nämlich auch über zwei Lautsprecher wiedergegeben werden. Zunächst hat man dann aber das Problem, dass ohne weiteres keine hinreichende Kanaltrennung erreicht wird, weil beide Lautsprecher von beiden Ohren gehört werden können. Die Kanaltrennung muss erst durch eine geeignete Signalverarbeitung hergestellt werden. Dazu wird das Übersprechen vom abgewandten Lautsprecher durch ein geeignet gefiltertes Signal vom zugewandten Lautsprecher durch destruktive Interferenz am Ohr ausgelöscht. Die auf Interferenz beruhende Übersprechkompensation ist sehr empfindlich gegen Phasenfehler, die durch geringe Unterschiede in der Laufzeit der Signale von rechtem und linkem Lautsprecher resultieren. Sind die Übersprechkompensationsfilter statisch für einen Punkt eingemessen, dann muss sich die Person genau an dem Punkt befinden, für den eingemessen wurde. Bereits bei einer seitlichen Abweichung von unter zwei Zentimetern konnten von Versuchspersonen in Hörtests Unterschiede bei der Lokalisation oder der Klangfarbe wahrgenommen werden.

Bei der dynamischen Übersprechkompensation wird das System auf die jeweilige Position der Ohren adaptiert. Dies erfordert eine aufwendige und schnelle Datenverarbeitung. Die Positionsdaten werden von einem Head-Tracker-System übermittelt, welches in typischen VR-Systemen für die stereoskopische Projektion ohnehin integriert ist. Interessant ist dieses Verfahren, weil es die Vorteile eines binauralen Verfahrens ohne "Im Kopf Lokalisation" und ohne die Verwendung von Kopfhörern erlaubt, die einen unnatürlichen Einflussfaktor auf die Immersion in die virtuelle Umgebung darstellen. Sofern diese Technik für die Wiedergabe binauraler Szenen aus Kunstkopfaufnahmen verwendet wird, arbeitet sie nicht anders als ein Kopfhörer: Bei Kopfdrehung dreht sich das Schallereignis mit.

Ein kombinierter Ansatz, welcher schließlich die head-getrackte Wiedergabeeinheit mit der binauralen Synthese koppelt, stellt das Ideal dar. Somit lassen sich raumfeste Quellen realisieren. Erst dann ist gewährleistet, dass trotz vollständiger Bewegungsfreiheit die akustische Szene mit der visuellen Szene gekoppelt und im 3D-Raum fixiert ist.

# Elektroakustik

Herbert Hudde / Elektroakustik

1. Einleitung

2. Elektrodynamischer Lautsprecher

2.1 Aufbau und Grundprinzip

2.2 Ersatzschaltbild unter Verwendung der elektromechanischen und elektroakustischen Analogien

2.3 Schallabstrahlung, Kugelstrahler nullter Ordnung

2.4 Betriebsverhalten des eingebauten Lautsprechers

3. Kondensatormikrofon

3.1 Aufbau und Ersatzschaltbild

3.2 Betriebsverhalten

4. Richtwirkung

4.1 Richtwirkung durch Interferenz

4.2 Druck- und Druckgradientenempfänger

Anhang: Kleine Schallsender und -empfänger für Hörgeräte

#### 1. Einleitung

Grundlage der Elektroakustik sind Komponenten, mit denen (a) Schall aufgrund einer elektrischen Anregung erzeugt werden kann oder (b) akustische Messgrößen - meist Schalldrücke - in proportionale elektrische Signale verwandelt werden. In das Fachgebiet der Elektroakustik gehören auch alle Anwendungen solcher Komponenten zur Schallaufnahme, Speicherung und Schallwiedergabe. Dies umfasst die Erfassung von Schallfeldern mit vielen gemeinsam ausgewerteten Mikrofonsignalen (Mikrofon-Arrays) oder etwa Kunstköpfen und die Schallwiedergabe über Surround-Systeme oder die Beschallung von Räumen mit einer Vielzahl geeignet angesteuerter Lautsprecher oder. Die Elektroakustik ist daher ein sehr umfangreiches Fachgebiet.

In diesem Kurs können nur einige Grundlagen besprochen werden. Die Basis der meisten elektroakustischen Komponenten sind elektromechanische Wandler, die über eine schwingungsfähige Membran an ein Schallfeld angekoppelt sind. In den meisten Fällen arbeiten die so entstehenden elektroakustischen Wandler annähernd linear, zeitinvariant und reziprok. Die Reziprozität beinhaltet, dass die Wandler in beide Richtungen, also als Schallsender und empfänger, betrieben werden können, und dass zwischen beiden Betriebsrichtungen wohldefinierte Beziehungen bestehen. In aller Regel wird ein elektroakustischer Wandler aber selbstverständlich für eine gewünschte Betriebsrichtung entworfen und dimensioniert.

Neben zwei speziellen Wandlern, dem elektrodynamischen Lautsprecher und dem Kondensatormikrofon, wird die Erzeugung von Richtcharakteristiken betrachtet.

#### 2. Elektrodynamischer Lautsprecher

Zur Realisierung von Lautsprechern wird mit Abstand am häufigsten ein elektrodynamischer Wandler eingesetzt. Bei einem solchen Wandler bewegt sich eine Spule im Luftspalt eines Magnetkreises. Neben dem dynamischen Wandlerprinzip gibt es noch ein zweites Wandlerprinzip, das auf Wirkungen von Magnetfeldern beruht. Dieses "elektromagnetische Wandlerprinzip", bei dem die stromdurchflossene Spule unbewegt bleibt, wird hier nicht besprochen.

#### 2.1 Aufbau und Grundprinzip

In Abb. 2-1 ist ein Schnitt durch einen rotationssymmetrischen dynamischen Lautsprecher dargestellt. Das Kernstück ist ein Magnetkreis, der einen Permanentmagneten und Weicheisen zur Führung der magnetischen Feldlinien enthält. Auf diese Weise wird im Luftspalt eine magnetische Induktion *B* erzeugt, die eine elektrische Spule der Länge *l* senkrecht durchsetzt. Die Grundgleichungen, die den dynamischen Wandler charakterisieren, beschreiben die Lorentzkraft auf einen stromdurchflossenen Leiter und das Induktionsgesetz.



Abb. 2-1: Schnitt durch einen elektrodynamischen Lautsprecher

Verhindert man eine Bewegung der Schwingspule durch "Festbremsen", so gelten die aus der Elektrotechnik bekannten Gleichungen. Ein Strom *i* erzeugt eine proportionale Lorentzkraft

$$\mathbf{F} = i \cdot (\mathbf{I} \times \mathbf{B}), \quad bzw. \quad F = Bli = Mi \tag{2-1}$$

Man überzeuge sich anhand der Abb. 1 und der Vektorbeziehung in Gl. (1), dass die erzeugte Kraft die Spule tatsächlich in die gewünschte Richtung bewegt.

Die Größe M = Bl charakterisiert die Stärke des Wandlers. Sie wird als Wandlerkonstante bezeichnet. Die Luftspaltinduktion *B* und die Leiterlänge *l* sollten also prinzipiell möglichst groß sein, wobei große Leiterlängen zu großen Schwingspulenmassen und zu großen Luftspalten führen, die die Luftspaltinduktion wiederum verkleinern. Man benötigt also vor allem eine große Luftspaltinduktion *B* an, so dass die Leiterlänge nicht zu groß gemacht werden muss. Die erreichbare Größenordnung der Wandlerkonstanten beträgt einige N/A.

Versetzt man nun umgekehrt die Schwingspule durch eine externe Kraft in Bewegung, so entsteht an den Enden der elektrisch leerlaufenden Spule eine induzierte Spannung

$$u = \mathbf{l} \cdot (\mathbf{v} \times \mathbf{B}), \quad bzw. \quad u = Blv = Mv \tag{2-2}$$

Man überzeuge sich wieder davon, dass die Richtung der beteiligten Vektoren so ist, dass tatsächlich eine Spannung induziert wird.

#### 2.2 Ersatzschaltbild unter Verwendung der elektromechanischen und elektroakustischen Analogien

Mit den beiden Gleichungen wird der so genannte "innere Wandler" bereits erfasst. Er

enthält das idealisierte Grundprinzip, ohne die bei der Realisierung zwangsläufig auftretenden elektrischen und mechanischen Zusatzeffekte zu erfassen. Der reale Wandler besitzt darüber hinaus weitere elektrische und mechanische Eigenschaften: die Spule hat elektrisch gesehen eine Induktivität und einen Verlustwiderstand; die Schwingspule zusammen mit der Schall abstrahlenden Membran hat eine Masse und muss federnd aufgehängt werden. Dies kann durch Hinzufügen einer elektrischen und einer mechanischen Impedanz im Ersatzschaltbild (Abb. 2-2) approximiert werden.



Abb. 2-2: Ersatzschaltbild eines elektrodynamischen Wandlers. Der Index iw kennzeichnet die Größen am inneren Wandler. Die Schnelle  $v_{iw}$  am inneren Wandler ist gleichzeitig die Membranschnelle  $v_M$ .

Die zusätzlichen *elektrischen* Zusatzeigenschaften ergeben sich bei Festbremsung der Spule. Dann lässt sich eine komplexe elektrische Eingangsimpedanz  $Z_{el}$  messen, die näherungsweise durch die Reihenschaltung eines ohmschen Widerstandes R und einer Induktivität L beschrieben werden kann.

$$Z_{el} = R + j\omega L \tag{2-3}$$

Auch die zusätzlichen *mechanischen* Eigenschaften lassen sich mit Netzwerkelementen beschreiben, da die mechanischen Grundeigenschaften, gegeben durch Masse, Steife und Reibung zu Gleichungen führen, die zu elektrischen Elementgleichungen analog sind. Hierzu benötigt man die Beziehungen zwischen Elongation  $\xi$ , Schnelle (Schwinggeschwindigkeit) v und Beschleunigung a:

$$v = \frac{\mathrm{d}\xi}{\mathrm{d}t} \quad , \quad a = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} \tag{2-4}$$

bzw. in komplexer Schreibweise

$$v = j\omega\xi$$
,  $a = j\omega v = -\omega^2\xi$  (2-5)

Die gewünschten analogen Beziehungen ergeben sich, wenn man die mechanischen Gesetze durch Kräfte und Schnellen beschreibt. Aus dem Newtonschen Gesetz F = m a wird

$$F = ma = j\omega m \cdot v \tag{2-6}$$

Man kann somit eine mechanische Impedanz  $Z_m = F / v$  definieren, die bei der Masse j $\omega m$  beträgt.

Die naheliegende Wahl, dass man bei der elektromechanischen Analogie entsprechend der Impedanz die Kraft F als Analogon der elektrischen Spannung u und die Schnelle v als Analogon de Stroms i ansieht, ist zwar möglich, führt aber zu der Schwierigkeit, dass die Struktur der elektrischen Ersatzschaltung dual zu der mechanischen Struktur ist. Das bedeutet, dass in Netzwerken Knoten- und Maschengleichungen vertauscht werden, also z. B. Reihenschaltungen und Parallelschaltungen gegenüber der originalen mechanischen Struktur vertauscht werden.



Abb. 2-3: Zur Wahl der translatorischen mechanischen Analogiebeziehungen

Abb. 2-3 soll veranschaulichen, dass man die Wahl der Analogiepaare gemäß  $F \Leftrightarrow i$ ,  $v \Leftrightarrow u$  treffen muss, wenn man strukturelle Übereinstimmung zwischen analogen mechanischen und elektrischen Netzwerken erreichen will. Will man eine Kraft messen, die auf ein mechanisches System wirkt, so muss man die Verbindung auftrennen und einen Kraftmesser in den Zweig schalten. Die Kraft entspricht strukturell also einem elektrischen Strom. Die Schnelle einer schwingen Masse muss relativ zu einem Bezugspunkt gemessen werden. Dies entspricht einer Spannung zwischen zwei Klemmen, also einem Potenzial mit Bezug auf eine Bezugsklemme. Die Komplikation dieser Wahl besteht darin, dass nun elektrische Impedanzen mechanischen Admittanzen und umgekehrt entsprechen, was etwas gewöhnungsbedürftig ist.

Nach dieser strukturtreuen Wahl der Analogie wird eine Masse nicht durch nach Gl. (2-6) nahe liegende Induktivität, sondern durch eine Kapazität dargestellt.

Die Elongation einer idealen Feder mit der Steife *s* ist kraftproportional ( $\xi = F/s$ ). Drückt man dies wieder durch die Schnelle *v* aus und geht von der Steife zum Kehrwert "Nachgiebigkeit" n = 1/s über, so erhält man bei der gewählten Analogie eine mechanische Induktivität.

$$F = \frac{1}{n}\xi = \frac{1}{j\omega n} \cdot v \tag{2-7}$$

Schließlich ist mechanische Reibungskraft häufig annähernd der Schnelle proportional, was auf einen mechanischen Wirkwiderstand *w* führt.

$$F = w \cdot v \tag{2-8}$$

Im Fall der Schwingspule ist die Schnelle bzw. Auslenkung für die beteiligten Elemente Masse, Nachgiebigkeit und Reibung identisch. Die Gesamtkraft in Abb. 2 ist somit

$$F = F_{iw} = \left(w + j\omega m + \frac{1}{j\omega m}\right) \cdot v = Z_{mech}v_M$$
(2-9)

Danach ist die mechanische Gesamtimpedanz die Summe aller Teilimpedanzen. Dies bestätigt die genannte Dualität zwischen elektrischen und mechanischen Netzwerken: bei mechanischen Netzwerken sind bei Parallelschaltung die Teilimpedanzen zu addieren!

Die drei Grundelemente der strukturtreuen elektromechanischen Analogie sind in der folgenden Abbildung zusammen gefasst.



Abb. 2-4: Masse, Feder und Reibungswiderstand und ihre Entsprechungen gemäß der strukturtreuen mechanischen Analogie

Aus der Wahl der strukturtreuen Analogie folgt, dass die Grundgleichungen des inneren Wandlers einem elektrischen Übertrager entsprechen, der die Gln. (2-2) repräsentiert. Dieser wurde bereits in Abb. 2-2 verwendet.

Strahlt der dynamische Wandler Schall ab, so addiert sich noch eine der akustischen Belastung entsprechende Impedanz zu  $Z_{mech}$ . Um im Ersatzschaltbild nach Abb. 2 auch die akustische Last zu berücksichtigen, muss man die akustische Impedanz  $Z_{ak} = p/q$  auf die mechanische Seite transformieren. Die akustische Impedanz ist definiert als Verhältnis von Schalldruck *p* zu Schallfluss *q*. Die Umrechnung der akustischen Größen in mechanische ergibt sich aus der wirksamen Membranfläche *A* gemäß

$$F = pA, \quad v = q/A \tag{2-10}$$

Da man bei der akustischen Analogie sinnvollerweise die struktur- und impedanztreuen Analogiepaare  $(p \Leftrightarrow u)$  und  $(q \Leftrightarrow i)$  benutzt, während ja mechanisch die Paare  $(v \Leftrightarrow u)$ und  $(F \Leftrightarrow i)$  angewandt wurden, stellt sich das verkoppelnde Element nicht als Übertrager, sondern als Gyrator heraus. Formal muss man also das Ersatzschaltbild nach Abb. 2-2 wie folgt ergänzen (Abb. 2-5).



Abb. 2-5: Vollständiges Ersatzschaltbild des dynamischen Wandlers inklusive einer akustischen Belastung  $Z_{rad}$  (Abstrahlimpedanz, radiation impedance).

Die ins Mechanische umgerechnete akustische Belastung lässt sich einfach angeben

$$Z_{m,rad} = \frac{F_{rad}}{v_M} = \frac{p_{rad}A}{q_{rad}/A} = A^2 Z_{rad}$$
(2-11)
Die gesamte wirksame mechanische Impedanz, die den Wandler belastet, ist also

$$Z_{mech,ges} = Z_{mech} + Z_{m,rad} = Z_{mech} + A^2 Z_{rad} , \qquad (2-12)$$

wobei  $Z_{ak}$  die so genannte Abstrahlimpedanz des Lautsprechers ist. Allerdings spielt die akustische Lastimpedanz bei Lautsprechern hinsichtlich der Wandlerbelastung meist keine wesentliche Rolle. Sie nimmt jedoch einen erheblichen Einfluss auf den abgestrahlten Schall. Um das Betriebsverhalten eines elektrodynamischen Lautsprechers untersuchen zu können, benötigt man zunächst eine Abschätzung der wirksamen Abstrahlimpedanz.

#### 2.3 Schallabstrahlung, Kugelstrahler nullter Ordnung

Die Abstrahlimpedanz eines Lautsprechers nach Abb. 2-1 hängt in komplizierter Weise von den geometrischen Details der Anordnung ab. Für den nicht eingebauten Lautsprecher müsste man berücksichtigen, dass Schall auf Vorder- und Rückseite der Membran abgestrahlt wird; man müsste also zumindest zwei getrennte Abstrahlimpedanzen annehmen. Eine einfache Abschätzung ergibt sich nur, wenn der Lautsprecher in eine Box eingebaut wird. Die rückwärtige Schallabstrahlung wird dadurch unterbunden. Damit erhält man ein einseitig Schall abstrahlendes Element. Solche Elemente arbeiten bei niedrigen Frequenzen fast wie ein Kugelstrahler, der Schall völlig gleichmäßig in alle Richtungen abstrahlt. Die Entstehung angenäherter Kugelwellen ergibt sich aus den Grundgleichungen der linearen Akustik bzw. aus dem Huygens-Fresnelschen Prinzip. In idealisierter Form entspricht danach der eingebaute Lautsprecher bei niedrigen Frequenzen einem Kugelstrahler nullter Ordnung, den man sich als eine "atmende Kugel" vorstellen kann.

Für diese Idealisierung erhält man bei Abstrahlung ins freie Schallfeld einen mit dem Abstand r gemäß

$$p(r) = \frac{p_{+}}{r} e^{-j\beta_{0}r}.$$
(2-13)

abklingenden Schalldruck. In der Gleichung bedeutet  $p_+$  eine zur Amplitude der abgestrahlten Schalldruckwelle proportionale Konstante und  $\beta_0$  die Wellenzahl, für die folgende Zusammenhänge mit Wellenlänge  $\lambda$ , Schallgeschwindigkeit c, Frequenz f und Kreisfrequenz  $\omega$  gelten

$$\beta_0 = \frac{2\pi}{\lambda} = \frac{2\pi f}{c} = \frac{\omega}{c}.$$
(2-14)

Die Schallschnelle ergibt sich nach der akustischen Eulergleichung gemäß ( $\rho_{=}$  bezeichnet die Ruhedichte der Luft)

$$\mathbf{v}(r) = -\frac{1}{j\omega\rho_{\pm}}\frac{\partial p}{\partial r} = -\frac{p_{\pm}}{j\omega\rho_{\pm}}(-\frac{1}{r^{2}}e^{-j\beta_{0}r} - \frac{j\beta_{0}}{r}e^{-j\beta_{0}r})$$
(2-15)

Um auf akustische Impedanzen überzugehen, benötigt man den zugehörigen Schallfluss an der Kugeloberfläche, der sich durch Multiplikation mit der Kugeloberfläche  $A(r_0) = 4\pi r_0^2$ 

$$q(r_0) = A(r_0)v_r(r_0) = A(r_0)p_+ \left[\frac{1}{r_0 \cdot \rho_{=}c} + \frac{1}{j\omega\rho_{=}r_0^2}\right]e^{-j\rho_0 r_0} = \frac{p_+}{r_0}A(r_0)\left[\frac{1}{\rho_{=}c} + \frac{1}{j\omega\rho_{=}r_0}\right]$$
(2-16)

Damit erhält man für die akustische Impedanz an der Kugeloberfläche, also die Abstrahlimpedanz der Kugel

$$Z_{rad}(r_0) = \frac{p(r_0)}{q(r_0)} = \frac{1}{\frac{A(r_0)}{\rho_{=}c} + \frac{A(r_0)}{j\omega\rho_{=}r_0}} = \frac{\rho_{=}c}{A(r_0)} \frac{1}{1 + \frac{c}{j\omega r_0}} = \frac{\rho_{=}c}{A(r_0)} \frac{1}{1 + \frac{1}{j\beta_0 r_0}}.$$
(2-17)

Um dieses Ergebnis zu interpretieren, benötigen wir die Elemente der elektroakustischen Analogie. Die sind in Abb. 2-6 dargestellt.



Abb. 2-6: Elektroakustische Analogie: Ein abgeschlossenes Volumen bewirkt eine Nachgiebigkeit, die einem elektrischen Kondensator entspricht. Die akustische Masse in einem (engen) Röhrchen entspricht einer Induktivität, ein absorbierendes Element kann im einfachsten Fall durch einen Widerstand gekennzeichnet werden.



Abb. 2-7: Abstrahlimpedanz eines Kugelstrahlers nullter Ordnung.

Die Impedanz nach Gl. (2-17) lässt sich somit als Parallelschaltung einer Masse (Induktivität) und eines Widerstandes (akustische Feldimpedanz des freien Schallfeldes) interpretieren. Bei niedrigen Frequenzen "schließt die Induktivität den Widerstand fast kurz". Die Abstrahlimpedanz ist dann fast nur durch die mitschwingende akustische Masse  $m_{ak}$  (Induktivität) mit der akustischen Impedanz j $\omega m_{ak}$  gegeben.

#### 2.4 Betriebsverhalten des eingebauten Lautsprechers

Nun liegen alle Elemente vor, um das Betriebsverhalten des Lautsprechers berechnen zu können. Um von der elektrischen Spulenimpedanz unabhängig zu werden, betrachten wir den Fall einer Stromspeisung des Lautsprechers. Der von einem Spulenstrom i hervorgerufene Schalldruck p lässt sich aus dem Ersatzschaltbild nach Abb. 2-5 leicht berechnen.

$$p = Z_{rad} q = Z_{rad} Av = \frac{Z_{rad} AF_i}{Z_{mech} + A^2 Z_{rad}} = \frac{BlAZ_{rad}}{Z_{mech} + A^2 Z_{rad}} \cdot i \approx \frac{BlAZ_{rad}}{Z_{mech}} \cdot i.$$
(2-18)

Dies ist der tatsächlich vor der Membran auftretende Druck. Offenbar ist es günstig, die Wandlerkonstante M = Bl und die abstrahlende Fläche A groß zu wählen.

Meist interessiert man sich nicht für den unmittelbar an der Membran auftretenden Schalldruck, sondern für die ins Fernfeld abgestrahlte Leistung bzw. den dort auftretenden Schalldruck. Dazu muss man die akustische Wirkleistung betrachten. Für diese ist nicht die gesamte Abstrahlimpedanz, sondern nur ihr Realteil verantwortlich.

$$P = \frac{1}{2} |q|^2 \operatorname{Re} \{Z_{rad}\} = \frac{1}{2} |q|^2 \operatorname{Re} \left\{ \frac{\rho_{=}c/A(r)}{[1+1/j\beta_0 r]} \right\} = \frac{1}{2} |q|^2 \frac{\rho_{=}c}{A(r)} \cdot \frac{1}{1+1/(\beta_0 r)^2}$$
(2-19)

Bei niedrigen Frequenzen ist die abgestrahlte Wirkleistung nur sehr gering. Aus Gl. (2-19) ergibt sich die Näherung

$$P \approx \frac{1}{2} |q|^2 \frac{\rho_{=}c}{A(r)} \cdot \left(\beta_0 r\right)^2 = \frac{1}{2} |q|^2 \frac{\rho_{=}}{4\pi c} \omega^2 .$$
(2-20)

Die Schallabstrahlung ist also bei niedrigen Frequenzen sehr schlecht. Die abgestrahlte Wirkleistung steigt proportional mit dem Quadrat der Frequenz. Die schlechte Schallabstrahlung bei niedrigen Frequenzen lässt sich physikalisch so interpretieren, dass hier vor allem nur eine akustische Masse ohne Transport von Wirkleistung hin und her bewegt wird. Die Abstrahlimpedanz ist fast rein imaginär. Erst mit wachsender Frequenz wird auch der Realteil größer. Trotz eines hohen Schalldrucks vor der Membran kommt niederfrequent kaum Schallabstrahlung zustande.

Offenbar besteht die einzige Möglichkeit, die Schallabstrahlung zu vergrößern, darin, den erzeugenden Schallfluss q so groß wie möglich zu machen. Das bedeutet: man braucht große Membranflächen und große Auslenkungen der Membranen, um niedrige Frequenzen abzustrahlen ("Tieftöner").

Für das Fernfeld wird nur der Teil des Membrandrucks wirksam der mit der Membranschnelle in Phase ist. Dies lässt sich über den Realteil der Abstrahlimpedanz ausdrücken. Der im Fernfeld wirksame Anteil des Membranschalldrucks ist demnach

$$p_{M,Fern} = q \cdot \operatorname{Re}\left\{Z_{rad}\right\} = A \frac{Bli}{Z_{mech,ges}} \operatorname{Re}\left\{Z_{rad}\right\} \approx BlA \frac{\left(\rho/4\pi c\right) \cdot \omega^2}{w + j\omega m + 1/(j\omega n)}i$$
(2-21)

Bei sehr niedrigen Frequenzen steigt dieser Druckanteil mit der dritten Potenz der Frequenz, oberhalb der mechanischen Resonanz des Lautsprechers linear mit der Frequenz. Unterhalb der mechanischen Resonanz ist die Schallabstrahlung also denkbar schlecht. Der Lautsprecher muss also so tief abgestimmt werden, dass der gewünschte Frequenzbereich oberhalb der Resonanzfrequenz liegt. Dies bedeutet eine relativ weiche Lagerung der Membran, da die Resonanzfrequenz gemäß

$$f_0 = \frac{1}{2\pi \cdot \sqrt{m \cdot n}} \tag{2-22}$$

von der Nachgiebigkeit abhängt. Die Masse wird man nicht unnötig groß machen, da dies die Schallabstrahlung bei hohen Frequenzen verschlechtern würde.

Die Zusammenhänge werden in Abb. 2-8 veranschaulicht. Die Schnelle der Membran folgt der mechanischen Admittanz. Der massebestimmte Abfall proportional  $1/\omega$  oberhalb der Resonanz führt wegen des Anstiegs des Realteils der Abstrahlimpedanz mit  $\omega^2$  zu einem ansteigenden Druckanteil für das Fernfeld. Erst wenn die Abstrahlimpedanz allmählich zunehmend reell wird, fällt der Druck. Man beachte die große Differenz zwischen dem tatsächlich vor der Membran auftretenden Druck und dem Fernfeldanteil. Ferner erkennt man, dass die abgestrahlte Schallleistung im mittleren Frequenzbereich in etwa frequenzunabhängig ist.



Abb. 2-8: Schallabstrahlung beim elektrodynamischen Lautsprecher bei frequenzunabhängigem Speisestrom.

Die schlechte Schallabstrahlung bei niedrigen Frequenzen kann durch verschiedene Maßnahmen verbessert werden. Hier sind neben den bereits genannten großen Membranschallflüssen vor allem eine tiefe mechanische Resonanzabstimmung und Bassreflexöffnungen zu nennen, bei denen der rückwärtige Schallanteil der Membran durch geeignete Phasendrehung zu einer Schallverstärkung genutzt werden kann.

Strahlt der Schallsender nicht in den freien Raum, sondern z. B. direkt in den Gehörgang, so wird er durch eine völlig andere Impedanz als ein Kugelstrahler nullter Ordung belastet. Ein Kopfhörer oder ein Hörgeräte-Wandler arbeitet auf ein mehr oder minder geschlossenes Volumen. Dies hat den positiven Effekt, dass die Schallabstrahlung bei den tiefen Frequenzen kein großes Problem ist. In einem kleinen Volumen lassen sich mit geringen Flüssen bereits große Schalldrücke erzeugen. Ein Volumen V ist im Rahmen der elektroakustischen Analogie (Abb. 2-6) durch eine Nachgiebigkeit  $n_{ak}$  (Kondensator) gegeben, der entsprechend der Gleichung

$$p = \frac{q}{j\omega n_{ak}}, \quad n_{ak} = \frac{V}{\rho_{=}c^{2}}$$
(2-23)

bei geringen Volumina und niedrigen Frequenzen hohe Schalldrücke erzeugt.

#### 3. Kondensatormikrofon

#### 3.1 Aufbau und Ersatzschaltbild

Bei dielektrischen Wandlern wird die Änderung eines *elektrischen* Feldes ausgenutzt. Das elektrische Feld befindet sich in einem Dielektrikum zwischen zwei metallisch leitenden Elektroden. Es wird also ein elektrischer Kondensator wirksam. Wie später gezeigt wird, ist das dielektrische Wandlerprinzip grundsätzlich nichtlinear. Man benötigt daher eine Linearisierung für kleine Signale um einen einzustellenden Arbeitspunkt herum. Die Spannung am Kondenstor muss also einen Gleichanteil enthalten, der viel größer als die Signalspannung ist. Den Gleichanteil nennt man Polarisationsspannung. Die Polarisationsspannung kann extern durch eine geeignete Gleichspannungsquelle erzeugt werden. Alternativ kann man Elektretfolien verwenden, die vorpolarisiert wurden. Die Polarisationsspannung ist dann sozusagen "bereits eingebaut".

Die Wandlerkonstante *N* des dielektrischen Mikrofons erhält man durch Berechnung der Kraft im Kondensator. Diese wiederum erhält man am einfachsten aus der gespeicherten Energie

$$W = \frac{1}{2}Cu^2 = \frac{1}{2}\frac{Q^2}{C}.$$
(3-1)

Im leerlaufenden Kondensator (keine Änderung der Ladung) mit einem Dielektrikum der Dielektrizitätszahl  $\varepsilon$  entsteht aufgrund der elektrischen Feldstärke eine Kraft *F*, die vom Plattenabstand *x* abhängt.

$$F(x) = \frac{\mathrm{d}W}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} \left\{ \frac{1}{2} \frac{Q^2}{C(x)} \right\} = \frac{Q^2}{2} \frac{\mathrm{d}}{\mathrm{d}x} \left\{ \frac{x}{\varepsilon A} \right\} = \frac{Q^2}{2\varepsilon A} \approx \frac{C_0^2}{2\varepsilon A} u^2 = \frac{C_0}{2x} u^2 \tag{3-2}$$

Der Zusammenhang zwischen Spannung und Kraft ist offenbar nicht linear, sondern quadratisch, annähernd gültig für den Kondensator  $C_0$  in Ruheposition der Platten (x = d). Für eine Spannung mit Gleichanteil  $U_p$  (Polarisationsspannung) wird daraus

$$F = \frac{C_0}{2d}u^2 = \frac{C_0}{2d} \cdot \left(U_P + u_{\sim}\right)^2 \quad \Rightarrow \quad F_{\sim} \approx \frac{C_0}{2d} \cdot 2U_P u_{\sim} = \frac{C_0 U_P}{d} u_{\sim} = N u_{\sim}$$
(3-3)

Mit der Polarisationsspannung  $U_p$  kann der Arbeitspunkt gut eingestellt werden. Die Wandlerkonstante des linearisierten inneren Wandlers ist

$$N = \frac{C_0 U_P}{d} \,. \tag{3-4}$$

Für den verlust- und speicherfreien inneren Wandler ist das Verhältnis von Kraft zu Spannung mit dem von Strom zu Schnelle identisch. Zusammengefasst gilt also für den inneren dielektrischen Wandler

$$F = Nu$$
,  $v = i/N$ , mit  $N = \frac{C_0 U_P}{d}$ . (3-5)

Der Begriff "Kondensatormikrofon" ist eigentlich mit dem Begriff "dielektrisches Mikrofon" identisch. Man spricht jedoch meist nur dann von Kondensatormikrofonen, wenn hochwertige Mikrofone gemeint sind. Messmikrofone werden aus Gründen der Langzeitstabilität vorwiegend mit externer Polarisationsspannung, jedoch auch mit Elektretfolie hergestellt. Im letzteren Fall wird die Elektretfolie ausschließlich auf die feste Gegenelektrode geklebt ("Rückplatten-Elektretmikrofone"). Hochwertige Studiomikrofone sind häufig als Elektretmikrofone ausgeführt. Dielektrische Mess- und Studiomikrofone zeichnen sich vor allem durch ihren extrem glatten Frequenzgang aus. Diese ergibt sich - wie im Verlauf dieses Abschnitts deutlich werden wird - aus der typischerweise sehr geringen Membranmasse, wodurch Trägheitseffekte erst bei sehr hohen Frequenzen relevant werden. Der glatte Frequenzgang ist auch der Grund für die Herstellung von besonders hochwertigen dielektrischen Kopfhörern.

Unter dem Begriff "Elektretmikrofone" werden meist kleine, wenig aufwändig hergestellte, relativ billige Mikrofone verstanden. Derartige Mikrofone kosten inklusive des integrierten Impedanzwandlers ca. 100  $\in$ . Hochwertige Kondensatormikrofon-Kapseln kosten dagegen 500-1500  $\in$ , hinzu kommt der Mikrofonvorverstärker und ein Mikrofonspeisegerät, das die Betriebsspannung für den Vorverstärker und evtl. die Polarisationsspannung liefert. Elektretmikrofone werden als Industriemikrofone, aber auch als Labormikrofone vielfach verwendet, wenn die Anforderungen an die Qualität (Glätte des Frequenzgangs, Unempfindlichkeit gegenüber Umwelteinflüssen wie Temperatur und Luftfeuchtigkeit, Langzeitstabilität, Klirrfaktor, etc.) nicht sehr hoch sind. Bei geeigneter Betriebsweise und mit sorgfältiger Kalibrierung lassen sich mit den Billig-Mikrofonen durchaus sehr genaue Messungen durchführen. Da Elektretmikrofone sich gut miniaturisieren lassen, werden sie auch in Hörgeräten verwendet.

Kondensatormikrofone kann man auch in der so genannten Hochfrequenzschaltung betreiben. Dabei bildet die Kondensatormikrofon-Kapsel den Kondensator eines Schwingkreises. Man wertet bei dieser Betriebsart die Amplitudenschwankungen eines hochfrequenten Trägers durch die Kapazitätsänderungen aus. Die Hochfrequenzschaltung hat insbesondere dann Vorteile, wenn man sehr niederfrequente Schallsignale erfassen will oder auch unter Bedingungen, bei denen elektrische Störungen eine Rolle spielen. Diese können durch die Trägerfrequenztechnik weitgehend ausgefiltert werden. Hier wird aber nur die wesentlich häufiger verwendete NF-Schaltung besprochen. Eine typische Kapselausführung zeigt die nächste Abbildung.

Die Membran bildet eine der beiden Elektroden (Abb. 3-1). Sie muss also metallisch und sehr dünn sein. Als Dielektrikum dient die Luft im Raum zwischen den beiden Elektroden. Der Abstand muss sehr schmal sein, um eine hinreichend große Kapazität  $C_0$  zu erhalten, er beträgt z.B. 20 µm. Damit das Kondensatormikrofon linear arbeitet, muss neben dem Vorhandensein einer Polarisationsspannung sichergestellt sein, dass die Membranauslenkungen nicht in die Größenordnung des Abstands kommen, sie sind also auf etwa 1 µm begrenzt. Dieser Wert entspricht dem Maximalpegel. Bei einer typischen Dynamik von etwa 120 dB und mehr liegt also die Auslenkung an der unteren Messgrenze bei 1 pm!



Abb. 3-1: Aufbau einer Kondensatormikrofon-Kapsel

Die Druckausgleichsöffnung dient dem statischen Druckausgleich. Ohne sie würden Schwankungen des statischen Drucks zu einer Vorspannung der Membran und damit zu einer Reduzierung der Empfindlichkeit führen. Allerdings wirkt sich die Druckausgleichsöffnung auch auf sehr niederfrequenten Schall aus. Die Öffnung wird daher äußerst klein gemacht, um nicht ungewollt ein Druckgradienten-Mikrofon zu bauen. Sie bewirkt eine untere Grenzfrequenz des Mikrofons, die man meist etwas unter die elektrische Grenzfrequenz legt, die durch die elektrischen Eigenschaften der Mikrofonkapsel und des Vorverstärkers bestimmt wird.

#### 3.2 Betriebsverhalten



Abb. 3-2: Verschaltete Mikrofonkapsel (Ersatzschaltbild)

Die elektrische Verschaltung der Mikrofonkapsel inklusive der unvermeidlichen Ersatzelemente zeigt Abb. 3-2. Die maßgebliche Kapazität  $C(t) = C_0 + \Delta C(t)$  liegt in der Größenordung von mehreren pF. Sie ist also extrem klein, so dass der Anschluss eines Kabels die Kapazität stark verändert. Daher ist es notwendig, einen Impedanzwandler unmittelbar hinter die Kapsel zu schalten. Dieser drückt sich im Ersatzschaltbild als Eingangsimpedanz des Vorverstärkers aus ( $C_V$ ,  $R_V$ ). Eine unerwünschte Streukapazität  $C_s$ , die zwar kleiner als C(t)ist, aber durchaus die Größenordnung von einem pF erreicht, ist unvermeidlich.

Der Isolationswiderstand  $R_0$  kann mit großem Aufwand auf etwa  $10^{16}\Omega$  gebracht werden. Der Vorwiderstand  $R_P$  der Polarisationsspannungsquelle und der Eingangswiderstand des Verstärkers liegen typischerweise im G $\Omega$ -Bereich. Die Polarisationsspannung  $U_P$  weist in der Regel Werte zwischen 30 und 500 V auf.

Der Kondensator  $C_{\rm K}$  ist eine "Koppelkapazität", die der Auskopplung der Polarisationsspannung dient. Die gewünschte Ausgangsspannung ist ja nur die gegenüber der Polarisationsspannung  $U_{\rm P}$  äußerst kleine Wechselspannung  $u_{\sim}$  (gute Linearisierung). Die Koppelkapazität kann so groß gewählt werden, dass sie nicht stört (sinnvolle untere Grenzfrequenz).

Die Kapselkapazität ändert sich mit der Auslenkung gemäß

$$C(t) = \frac{\varepsilon A}{d + \xi(t)} = \frac{\varepsilon A}{d} \frac{1}{1 + \xi(t)/d} \approx C_0 \left( 1 - \frac{\xi(t)}{d} \right)$$
(3-6)

wobei die Näherung für hinreichend kleine Auslenkungen gilt. Eine Analyse des Ersatzschaltbildes nach Abb. 3-2 ergibt für die Ausgangsspannung bei harmonischer Anregung, genähert für kleine Kapazitätsänderungen

$$u \approx \frac{j\omega C_0 \frac{R_0}{R_0 + R_p} U_p \frac{\xi}{d}}{\frac{1}{R_0} + \frac{1}{R_p} + \frac{1}{R_V} + j\omega (C_0 + C_s + C_V)} \approx \frac{j\omega C_0 U_p \frac{\xi}{d}}{\frac{1}{R} + j\omega C} = \frac{j\omega N\xi}{\frac{1}{R} + j\omega C}$$
(3-7)

mit den Abkürzungen  $R = R_0 ||R_P||R_V$  und  $C = C_0 + C_s + C_V$ . Offensichtlich tritt durch alle elektrischen Elemente zusammen eine untere Grenzfrequenz  $f_u = 1/2\pi RC$  auf.

Man legt die elektrische Grenzfrequenz des Mikrofons natürlich so tief wie möglich (*R* möglichst groß). Oberhalb dieser Grenzfrequenz gilt

$$u \approx \frac{j\omega N\xi}{j\omega C} = \frac{N}{C}\xi = \frac{U_P C_0}{(C_0 + C_s + C_V)}\frac{\xi}{d}$$
(3-8)

Das Mikrofon arbeit dann als Elongationswandler  $(u \Box \xi)$ . Die Gleichung zeigt ferner, dass die Streukapazität und die Kapazität des Vorverstärkers die Empfindlichkeit reduzieren. Man darf also nicht etwa die Grenzfrequenz durch Vergrößerung von  $C_V$  absenken, sondern durch großes *R* (der niedrigste Teilwiderstand ist maßgeblich!). Wenn man bei sehr niedrigen Frequenzen im Infraschallbereich messen muss, benötigt man Mikrofone in Hochfrequenzschaltung. Damit können sogar statische Auslenkungen gemessen werden.

Gl. 3-7 lässt sich auf der elektrischen Seite des inneren Wandlers als Beschaltung mit einer Parallelschaltung aus R und C deuten. Insgesamt ergibt sich damit folgendes Ersatzschaltbild des Kondensatormikrofons.



Abb. 3-3: Ersatzschaltung eines dielektrischen Mikrofons. Rechts: Vereinfachung für den Fall einer bereits bekannten Schnelle-Anregung.



#### Abb. 3-4: Kondensatormikrofon-Kapseln mit Vorverstärker.

Wie bereits gesagt, muss der Vorverstärker (als Impedanzwandler) direkt hinter die Mikrofonkapsel geschaltet werden. Ein längeres Anschlusskabel würde durch seine parallel geschaltete Kapazität die Empfindlichkeit des Mikrofons nach Gl. 3-8 stark herabsetzen. Ferner würden Störfelder aufgrund des hohen Impedanzniveaus leicht in den Vorverstärker eindringen können. Die Verbindung vom Mikrofon zum Vorverstärker wird daher stets gut gekapselt und abgeschirmt ausgeführt. Die Baugröße des Mikrofons wird meist im Wesentlichen durch den Vorverstärker (mit Heizung zur Vermeidung des Temperatureinflusses) bestimmt (Abb. 3-4).

Aus dem Ersatzschaltbild liest man für elektrischen Leerlauf ab

$$u = Z_{el} \cdot i_i = Z_{el} \cdot N \cdot v_i = \frac{N \cdot Z_{el}}{Z_{mech} + N^2 \cdot Z_{el}} \cdot F = \frac{A \cdot N \cdot Z_{el}}{Z_{mech} + N^2 \cdot Z_{el}} \cdot p.$$
(3-9)

Bei realen Ausführungen ist im Nenner die transformierte elektrische Impedanz  $N^2 \cdot Z_{el}$  gegenüber der mechanischen Impedanz vernachlässigbar. Dann lässt sich die Gleichung oberhalb der elektrischen Grenzfrequenz schreiben

$$u = \frac{A \cdot N \cdot (1/j\omega C)}{w + j\omega m + (1/j\omega n)} \cdot p = \frac{A \cdot N \cdot n/C}{1 + j\omega nw - (f/f_0)^2} \cdot p \quad \text{mit} \quad f_0 = \frac{1}{2\pi\sqrt{mn}}$$
(3-10)

Diese Gleichung zeigt, dass man mit Hochabstimmung, also mit einer Resonanzfrequenz  $f_0$  oberhalb des gewünschten Frequenzbereichs einen äußerst glatten Frequenzgang erreichen kann. Da auch der Reibungswiderstand w problemlos gering gehalten werden kann, ist die elektrische Spannung u im Messfrequenzbereich nur durch den konstanten Zähler bestimmt. Die Hochabstimmung liegt wegen der geringen Membranmasse von Kondensatormikrofonen konstruktiv ohnehin nahe. Die zu einer Hochabstimmung notwendige steife Membraneinspannung (geringe Nachgiebigkeit n) steht allerdings im Widerspruch zu einer hohen Empfindlichkeit (Abb. 3-5).



Abb. 3-5: Frequenzgänge eines Kondensatormikrofons mit veränderlicher Steife der Membran. Eine Erniedrigung der Membransteife erhöht zwar die Mikrofonempfindlichkeit, jedoch wird die mechanische Resonanzfrequenz herabgesetzt. Bei den drei gezeigten Verläufen liegt jeweils ein Faktor 10 zwischen den Membransteifen. Die angenommene Güte der mechanischen Resonanz beträgt 50.

Die Umwandlung eines Schalldrucks in eine proportionale elektrische Spannung ist offenbar mit einem Kondensatormikrofon nahezu ideal realisierbar. Größere Fehler treten jedoch durch die Rückwirkung des Mikrofons auf das Schallfeld auf. Das Mikrofon bewirkt durch seine Anwesenheit Beugungs- und Streuungseffekte, die wesentlich von der Baugröße abhängen.

Bei Messungen im freien Schallfeld bzw. in einem reflexionsarmen Raum, bei denen der Schall von nur einer Schallquelle auf das Mikrofon gelangt, kann man daher eine deutliche Richtungsabhängigkeit des resultierenden Frequenzgangs beobachten (Abb. 3-6). Bei dem betrachteten Mikrofon handelt es sich um ein so genanntes "Druckmikrofon", weil es eine druckproportionale Ausgangsspannung mit flachem Frequenzgang liefert. Bei Schalleinfall von vorne in der Mittelachse der Mikrofonmembran entsteht ein deutlicher Druckstau-Effekt bis zu etwa 12 dB. Dies ist weit mehr als der Druckstau vor einer schallharten Wand (Druckverdopplung entspricht 6 dB).



Abb. 3-6: Frequenzgänge eines "Druckmikrofons" für verschiedene Schalleinfallsrichtungen. (Verhältnis zwischen dem Schalldruck an der Mikrofonmembran und dem Druck, der an der gleichen Stelle ohne Vorhandensein des Mikrofons auftreten würde.)

Neben den "Druckmikrofonen" werden auch "Freifeld-Mikrofone" angeboten, deren Empfindlichkeit bei hohen Frequenzen durch akustomechanische Maßnahmen so abgesenkt wird, dass der durch Beugung und Streuung hervorgerufene Anstieg möglichst gut kompensiert wird. Dies lässt sich selbstverständlich nur für eine Richtung, nämlich für frontalen Schalleinfall realisieren. Freifeld-Mikrofone werden hauptsächlich für Messungen im freien Schallfeld (z. B. im reflexionsarmen Raum) eingesetzt, wenn man eindeutig eine Schallausbreitungsrichtung definieren kann. Natürlich muss man darauf achten, dass das Mikrofon tatsächlich in Richtung der Ausbreitungen bei diffusem Schalleinfall relativ gering. Man wird also bei Messungen in diffusen Schallfeldern (z. B. im Hallraum) auch Druck-Mikrofone verwenden.

#### 4. Richtwirkung

#### 4.1 Richtwirkung durch Interferenz

Als einfachsten Fall betrachten wir Wandler, die entlang einer Linie angebracht sind. Bei Schallsendern bedeutet das eine Lautsprecherzeile, bei Empfängern ein entlang einer Linie angeordnetes Mikrofonarray. Im Folgenden wird die Richtwirkung anhand einer kontinuierlichen Belegung der Linie mit Schallstrahlern betrachtet. Aus der Reziprozität akustischer Systeme lässt sich folgern, dass die Richtcharakteristik einer Anordnung erhalten bleibt wenn man zwischen Sende- und Empfangsbetrieb wechselt. Die hier zunächst für den Sendebetrieb erhaltenen Richtcharakteristiken gelten also genauso für den Schallempfang.

Die Belegung der Linie mit Schallstrahlern beschreibt man am besten durch Punktstrahler mit gegebenen Quellschallflüssen, die man auch als Ergiebigkeit der Strahler bezeichnet. Durch Reduzierung des Kugelstrahlers nullter Ordnung auf einen Punkt erhält man aus der hier wiederholten Gleichung für den Kugelstrahler nullter Ordnung

$$p(r) = \frac{p_{+}}{r} e^{-j\beta_{0}r}.$$
(2-13)

den Schalldruckverlauf für einen Punktstrahler nullter Ordnung mit der Ergiebigkeit  $q_0$ .

$$p(r) = \frac{j\omega\rho_{\pm}}{4\pi} \frac{q_0}{r} e^{-j\beta_0 r}.$$
(4-1)

Abb. 4-1: Zur Interferenz beim Linienstrahler in einem weit entfernten Aufpunkt. Der Linienstrahler hat die Gesamtlänge L und besitzt einen längenbezogenen Ergiebigkeitsbelag  $q'_0(x)$ . Der Ausdruck  $x\cos\theta$  bezeichnet die ortsabhängige Abweichung der Weglänge von der mittleren Weglänge.

Betrachtet man nun eine Linienstrahler endlicher Länge L, der einen längenbezogenen , i. A. ortsabhängigen Ergiebigkeitsbelag  $q'_0(x)$  besitzt, so kann man an einem weit entfernten Aufpunkt durch Integration die Beiträge aller differentiell kleinen Punktstrahler überlagern ("Huygens-Fresnelsches Prinzip").

$$p(r_0,\theta) = \frac{j\omega\rho_{=}}{4\pi} \int_{-L/2}^{L/2} q_0'(x) \frac{e^{-j\beta_0(r_0 - x\cos\theta)}}{r_0 - x\cos\theta} dx = \frac{j\omega\rho_{=}}{4\pi} e^{-j\beta_0 r_0} \int_{-L/2}^{L/2} q_0'(x) \frac{e^{j\beta_0 x\cos\theta}}{r_0 - x\cos\theta} dx$$
(4-2)

Da wir einen Aufpunkt im Fernfeld betrachten, sind die Amplitudenänderungen aufgrund der unterschiedlichen Weglängen nur gering (also der Nenner im Integranden). Hingegen bleiben die Laufzeitdifferenzen trotzdem wichtig, weil sich die Laufzeit in der Phase widerspiegelt, die modulo  $2\pi$  wirksam wird. Die lange Grundstrecke mit der zugehörigen Phase  $\beta_0 r_0$  entspricht einer konstanten Laufzeit und ist ist für die Richtcharakteristik uninteressant. Für die Interferenz werden nur die Phasendifferenzen  $\beta_0 x \cos \theta$  wirksam. Mit der genannten Näherung erhalten wir

$$p(r_0,\theta) = \frac{j\omega\rho_{=}}{4\pi r_0} e^{-j\beta_0 r_0} \int_{-L/2}^{L/2} q_0'(x) e^{j\beta_0 x\cos\theta} dx = \frac{j\omega\rho_{=}}{4\pi r_0} e^{-j\beta_0 r_0} \int_{-L/2}^{L/2} q_0'(x) e^{jkx} dx$$
(4-3)

Die Größe

$$k = \beta_0 \cos \theta = 2\pi \cos \theta / \lambda . \tag{4-4}$$

bezeichnet man als Ortswellenzahl oder "Raumfrequenz".

Die zuletzt genannte Bezeichnung ergibt sich daraus, dass Gl. 4-3 zeigt, dass der entstehende Schalldruck aus dem Ergiebigkeitsbelag durch eine Fourier-Transformation hervorgeht. Der Ergiebigkeitsbelag  $q'_0(x)$  spielt darin die Rolle der Zeitfunktion. Der Druck ist pro-

portional zu einer Fourier-Transformierten, wobei die Rolle der Frequenz von der Ortswellenzahl (Raumfrequenz) übernommen wird. Die Berechnung von Schalldruckverteilungen aus Ergiebigkeitsbelägen mit Hilfe der Fourier-Transformation wird als "Fourier-Akustik" bezeichnet.

$$f(t) \Leftrightarrow F(j\omega) \iff q'_0(x) \Leftrightarrow p(-jk)$$

$$p(k) \Box \int_{-L/2}^{L/2} q'_0(x) e^{+jkx} dx \quad , \quad k = \beta_0 \cos \theta = \frac{2\pi}{\lambda} \cos \theta$$
(4-5)

Um daraus den Richtfaktor zu erhalten, muss man auf den Druck in der Richtung der größten Schallabstrahlung beziehen, so dass der Maximalwert 1 wird und alle Vorfaktoren herausfallen. Für den Linienstrahler mit konstantem Ergiebigkeitsbelag  $q'_0$  ergibt das Integral entsprechend der Fourier-Transformierten des Rechteckpulses eine Spaltfunktion ("Sinus cardinalis" si(x) oder auch sinc(x)), so dass man für den Betrag des Richtfaktors erhält

$$\left|D(\theta)\right| = \left|\frac{\sin(kL/2)}{kL/2}\right| = \left|\frac{\sin\left[\beta_0 \cdot (L/2) \cdot \cos\theta\right]}{\beta_0 \cdot (L/2) \cdot \cos\theta}\right| = \left|\sin\left\{\frac{\beta_0 L}{2} \cdot \cos\theta\right\}\right| = \left|\sin\left\{\pi\frac{L}{\lambda}\cos\theta\right\}\right|$$
(4-6)



Abb. 4-2: Richtcharakteristik eines Linienelements mit konstantem Ergiebigkeitsbelag  $q'_0$  der Länge L für den Fall  $L = 2\lambda$ .

Mit der Lautsprecherzeile bzw. dem Mikrofonarray kann man also eine deutliche Richtwirkung erzielen. Mit längeren Linienelementen wird die "Hauptkeule" schärfer, aber es entstehen immer mehr "Nebenkeulen".

Eine wichtige Anwendung kann man sich durch Anwendung einer bekannten Regel der Fouriertransformation überlegen:

$$f(t)e^{j\omega_0 t} \iff F(j\omega - j\omega_0) \quad bzw. \quad q'_0(x)e^{-jk_0 x} \iff D(k - k_0) = D(\cos\theta - \cos\theta_0)$$
(4-8)

Durch Belegung der Lautsprecherzeile mit einer linearen Phase gemäß Gl. 4-8 kann also der Bezugswinkel  $\theta_0$  verändert werden. Die Richtcharakteristik wird kann somit durch das Hinzufügen des Phasenverlaufs kontinuierlich geschwenkt werden, was für die Beschallungstechnik sehr hilfreich ist. Das "elektronische Schwenken" wird durch entsprechend verzögerte Ansteuerung von Lautsprechern realisiert, so dass der Ergiebigkeitsbelag durch die Teilstrahler den gewünschten Phasenverlauf über der Länge der Lautsprecherzeile hat. Da die Abhängigkeit  $D(\cos\theta - \cos\theta_0)$ , nicht jedoch  $D[\cos(\theta - \theta_0)]$  ist, ändert die Richtcharakteristik dabei etwas ihr Aussehen.



Abb. 4-3: Schwenken der Richtcharakteristik einer Lautsprecherzeile durch Laufzeitketten.

Die entsprechenden Möglichkeiten zum Schwenken der Richtcharakteristik kann man auch für linienförmig angeordnete Mikrofonarrays ("Beamforming") anwenden. Hierfür werden bei Hörgeräten bis zu drei Mikrofone verwendet.

Die hier nur für Linienstrahler- und empfänger dargestellten Möglichkeiten lassen sich auf flächenhafte Elemente übertragen. Dies läuft auf eine räumliche Fouriertransformation hinaus. Richtcharakteristiken lassen sich damit beliebig im Raum schwenken.

#### 4.2 Druck- und Druckgradientenempfänger

Einfache Richtcharakteristiken kann man auch dadurch erzielen, dass man Schall nicht nur auf die Vorderseite einer Membran auftreffen lässt, sondern in konstruktiv definierter Weise auch auf die Rückseite.

Zunächst betrachten wir aber den Fall, dass der Schall tatsächlich nur von einer Seite auf eine Mikrofonmembran treffen kann. Damit hat man einen einseitig Schall aufnehmenden Empfänger realisiert, der dem einseitig Schall abstrahlenden Sender entspricht. Wie dieser hat er niederfrequent eine kugelförmige Richtcharakteristik nullter Ordnung, die keine Richtung bevorzugt ("omnidirektionales" Mikrofon). Ein solches Mikrofon nennt man auch Druckempfänger.

Lässt man nun den Schall auch von der Rückseite auf das Mikrofon treffen, so wird das Ausgangssignal des Mikrofons  $x_{mic}$  proportional zur Druckdifferenz  $\Delta p$ . Geht man vereinfachend davon aus, dass das Schallfeld durch die Anwesenheit des Mikrofons nicht gestört wird, ergibt sich bei einer endlichen effektiven Länge  $\Delta x$  des Mikrofons folgender Zusammenhang

$$x_{mic} \Box \Delta p \approx \frac{\partial p}{\partial x} \cdot \Delta x .$$
(4-9)

Es wird also der räumliche Druckgradient als akustisches Eingangssignal wirksam. Daher bezeichnet man solche Mikrofone als Druckgradientenmikrofone.

In einem rein fortschreitenden ebenen Schallfeld bzw. im Fernfeld eines Kugelstrahlers, das aus einem Winkel  $\mathcal{G}$  gegenüber der Mittelachse durch die Membran einfällt, erhält man den Druckgradienten wie folgt

$$p = p_{+} e^{-j\beta_{0}x\cos\vartheta} \implies \frac{\partial p}{\partial x} = -j\beta_{0}\cos\vartheta \cdot p \Box j\omega\cos\vartheta, \qquad (4-10)$$

Man erhält eine dem Term  $\cos \vartheta$  entsprechende Richtcharakteristik. Diese hat die Form einer

liegenden Acht und wird deswegen auch als Achtercharakteristik bezeichnet. Räumlich entspricht die "Acht " zwei Kugeln, die mit entgegengesetzter Phasenlage empfangen.



Abb. 4-4: Achter-Charakteristik eines Druckgradientenempfängers (Kugelschallempfänger erster Qrdnung)

Gl. (4-10) zeigt, dass die Empfindlichkeit bei niedrigen Frequenzen sehr gering ist. Dies erklärt sich dadurch, dass die Drücke auf beiden Seiten der Membran fast identisch sind. Sie unterscheiden sich niederfrequent ja fast nur durch die (kleine) Phase.

Durch Kombination eines Druck- und eines Druckgradienten-Empfängers lassen sich weitere Richtcharakteristiken einstellen. Man kann eine Linearkombination beider Anteile bilden, indem man entweder elektrisch die Ausgangssignale eines Druck- und eines Druck- gradienten-Empfängers addiert oder indem man nur ein einziges Mikrofon verwendet, bei dem der rückwärtige Schalleinlass nur teilweise wirksam wird.

Dadurch erhält man Ausgangssignale, die einen Druck- und einen Druckgradientenanteil besitzen

$$x_{mic} = c_1 p + c_2 \frac{\partial p}{\partial r} \,. \tag{4-11}$$

Damit erzeugt man die Richtcharakteristik

$$D(\mathcal{G}) = \frac{x_{mic}(\mathcal{G})}{\left\langle x_{mic}(\mathcal{G}) \right\rangle_{\max}} = \frac{c_1 + c_2 \cos \mathcal{G}}{c_1 + c_2}$$
(4-12)

So lässt sich ein bevorzugt aus einer Seite empfangendes Mikrofon realisieren. Für gleiche Gewichtung, also  $c_1 = c_2$  erhält man die "Nierencharakteristik" ("Kardioide")

$$D = 0.5 (1 + \cos \theta) \,. \tag{4-13}$$



Abb. 4-53: Bildung einer Nierencharakteristik aus Kugel- und Achtercharakteristik

Durch Variation der Koeffizienten lassen sich auch andere Nierenformen bilden. So unterscheidet man z. B. die Superniere und die Hyperniere. In der Realität lassen sich solche Idealformen allerdings nur approximieren. Reale Richtcharakteristiken sind immer deutlich frequenzabhängig.



#### Abb. 4-6: Verschiedene Nierencharakteristiken.

Verschiedenartige Richtcharakteristiken bei Verwendung einer einzigen Membran werden durch akustische Filter erzeugt, die den rückwärtigen Schalleintritt frequenzabhängig variieren. Ein Beispiel ist in Abb. 4-7 dargestellt. Allerdings ist diese Darstellung eher als Prinzipschema zu sehen, praktische Ausführungen sind meist aufwändiger.



Abb. 4-7: Richtmikrofon mit akustischen Filtern

#### Anhang: Kleine Schallsender und -empfänger für Hörgeräte

Im Folgenden sind für einige besonders kleine Schallsender und -empfänger, wie sie in Hörgeräten verwendet werden, die erreichbaren Empfindlichkeiten angegeben. Neben den hier gezeigten Eigenschaften sind selbstverständlich weitere Daten von Interesse. Dazu zählt insbesondere das Rauschverhalten.

#### **Elektromagnetische Sender**



Knowles electromagnetic receiver BK (Dimensions: 7.87mm x 5.59mm x 4.04mm) Angaben in dB re  $20\mu$ Pa, also als sound pressure level (SPL) für Stromspeisung mit 1 mA



Knowles electromagnetic mini receiver FFH (Dimensions: 5.09mm x 2.80mm x 2.59mm)



Bauform eines miniaturisierten elektromagnetischen Hörgeräte-Schallsenders (Hörer). Die beiden Spulen erzeugen gegenphasige Induktionen im oberen und unteren Teil, wodurch eine Wechselkraft auf den Anker ausgeübt wird. Diese wird über ein Koppelelement auf die Membran übertragen. Der Schallauslass ist rechts zu sehen.

#### Elektretmikrofone



Knowles electret microphone EA (5.56 mm x 3.98 mm x 2.21 mm) -60 dB re 1V/0.1 Pa = -60 dB re  $10^4$  mV/Pa entspricht 10 mV/Pa



Knowles electret microphone FG (dimensions: 2.59 mm diameter x 2.59 mm)

# Das periphere Hörorgan – Funktionsweise und Modellbildung

Herbert Hudde / Bochum

#### Außenohr

#### Akustische Modellierung des Gehörgangs

Der Gehörgang bildet eine akustische Leitung. Dadurch kann in bestimmten Frequenzbereichen eine (passive) Schalldruckverstärkung realisiert werden. Nur bei niedrigen Frequenzen (unter 1 kHz) kann der Gehörgang als Volumen angesehen werden, in dem überall fast der gleiche Schalldruck herrscht.



Abb. 1: Grundmodell des Gehörgangs.  $p_{ecE}$ : Quellschalldruck des äußeren Schallfelds am Eingang E des Gehörgangs.  $Z_{ecE}$ : Innenimpedanz der äußeren Quelle, identisch mit der Abstrahlimpedanz des Gehörgangs aus Ebene E.  $Z_D$ : "Trommelfellimpedanz" in der Ebene D, **K**: akustische Kettenmatrix des Gehörgangs.

#### Impedanzen, Übertragungsfunktionen, Resonanzen



Abb 2: Übertragungsfunktionen verschiedener Gehörgänge von einer äußeren Schallquelle zum Trommelfell. Die Anhebungen entstehen aufgrund des harten Trommelfells bei Frequenzen, bei denen der Gehörgang etwa eine bzw. drei viertel Wellenlängen lang ist.

Die Übertragungsfunktion Quelle bis Trommelfell zeigt Gehörgangsresonanzen:

$$\frac{p_{\rm D}}{p_{\rm ecE}} = \frac{1}{\left(1 + Y_{\rm D}' Z_{\rm ecE}'\right) \cdot \cos\left(\beta_0 l_{\rm ec}\right) + j\left(Y_{\rm D}' + Z_{\rm ecE}'\right) \cdot \sin\left(\beta_0 l_{\rm ec}\right)}$$
  
mit  $Z_{\rm ecE}' = Z_{\rm ecE} / Z_{\rm L}$   $Y_{\rm D}' = 1/(Z_{\rm D} / Z_{\rm L}) = Y_{\rm D} \cdot Z_{\rm L}$ .  
Hauptresonanzen bei:  $\cos\left(\beta_0 l_{\rm ec}\right) = 0$ ,  $\left|\sin\left(\beta_0 l_{\rm ec}\right)\right| = 1$ 

Das sind die  $\lambda/4$ -Resonanz und (schwächer) die  $3\lambda/4$ -Resonanz. Sie treten auf, weil der Gehörgang am Ende ziemlich hart (Trommelfellimpedanz) und am Eingang ziemlich weich (Abstrahlimpedanz) abgeschlossen ist.

#### Modellierung von Außenohr-Übertragungsfunktionen

Man benötigt allgemeine numerische Verfahren für Schallfeldberechnungen im Außenraum eines Kopfes. Da sich HRTFs auf Beschallung aus dem Fernfeld beziehen, ist die Verwendung von Randelementen (Boundary Element Method, BEM) besonders vorteilhaft.

#### Zusammenfassung Außenohr

- Funktional gesehen gehören nicht nur die Pinna, sondern der gesamte Kopf und sogar der Rumpf (zumindest die Schultern) zum Außenohr, weil alle genannten Elemente die richtungsabhängige Schalleinkopplung in den Gehörgang beeinflussen.
- Der Gehörgang bietet Schutz vor Verletzungen des Mittelohres.
- Der Gehörgang bildet einen Leitungsresonator, der im für die Sprachkommunikation wichtigen Frequenzbereich zwischen 2 und 5 kHz eine Anhebung des Schalldrucks um bis zu etwa 15 dB bewirkt.
- Die Übertragungsfunktionen des Gehörgangs werden vor allem durch seine Länge, aber auch durch seine gesamte Form individuell beeinflusst.

## Mittelohr

#### **Eindimensionale Modellierung**



Abb. 3: Einfachste Grundstruktur eines Mittelohrmodells ohne Berücksichtigung von elastischen Elementen und Massen: Mittelohr als Hebelübertrager mit Hebelverhältnis h. Die Impedanz  $Z_{cav}$  beschreibt die (geringe) Wirkung der Paukenhöhle. Hingegen bestimmt die akustische Eingangsimpedanz  $Z_c$  der Cochlea, auf die die Stapesfußplatte arbeitet, wesentlich die Mittelohrübertragungsfunktionen mit.

Wenn man annimmt, dass das Mittelohres mechanisch durch einen einfachen Hebelübertrager mit dem Hebelverhältnis *h* charakterisiert werden kann, so ergibt sich aus dem Ersatzschaltbild folgender Zusammenhang zwischen den akustischen Ein- und Ausgangsgrößen (am Trommelfell bzw. hinter der Stapesfußplatte).

$$\begin{pmatrix} p_D \\ q_D \end{pmatrix} = \begin{pmatrix} 1/A_D & 0 \\ 0 & A_D \end{pmatrix} \cdot \begin{pmatrix} 1/h & 0 \\ 0 & h \end{pmatrix} \cdot \begin{pmatrix} A_F & 0 \\ 0 & 1/A_F \end{pmatrix} \cdot \begin{pmatrix} p_C \\ q_C \end{pmatrix} = \begin{pmatrix} \frac{A_F}{h \cdot A_D} & 0 \\ 0 & \frac{h \cdot A_D}{A_F} \end{pmatrix} \cdot \begin{pmatrix} p_C \\ q_C \end{pmatrix}$$

Das gesamte wirksame Übertragungsverhältnis ist danach durch das Flächenverhältnis Trommelfell zu Stapesfußplatte und das Hebelverhältnis gegeben. Da das Hebelverhältnis jedoch nahe bei 1 liegt, wird hauptsächlich das Flächenverhältnis wirksam.

#### Übertragungsverhalten

Die Annahme, dass das Mittelohr mechanisch ein Hebelübertrager sei, ist viel zu grob, um die tatsächlichen Verhältnisse richtig zu beschreiben. Die elastischen Elemente (Trommelfell, Bänder, Muskeln) und die vor allem durch ihre Massenträgheit wirksamen Ossikel führen dazu, dass die vereinfachte Betrachtung eines Hebelübertragers allenfalls im Bereich der Hauptresonanz des Mittelohres näherungsweise anwendbar ist. Eine realistischere Beschreibung des Mittelohres führt auf folgende Übertragungsfunktionen.



Abb. 4: Schalldruckübertragungsfunktion von einer äußeren Quelle ( $p_{ecE}$ ) zur Cochlea ( $p_C$ ) für verschiedene Fälle im Vergleich zu einem theoretischen Optimum, das der (unrealistischen) Annahme entspricht, dass dem äußeren Schallfeld die maximale Leistung entnommen und vollständig in die Cochlea übertragen wird.

Die tatsächlich auftretenden Schwingungen können von den üblichen eindimensionalen Netzwerkmodellen des Mittelohres nicht erfasst werden. Mit verallgemeinerten Netzwerkmodellen oder mit finiten Elementen können realistischere Vorhersagen gemacht werden. Dabei zeigt sich, dass die Ossikel in verschiedenen Frequenzbereichen recht unterschiedlich schwingen. Trotzdem entstehen ähnliche Übertragungsfunktionen wie bei eindimensionalen Modellen.



Abb. 5: Stapesauslenkung bei Anregung mit einem Schalldruck von 1 Pa als Funktion der Frequenz. Dieser Frequenzgang wurde mit einem verallgemeinerten Netzwerkmodell berechnet, das räumliche Schwingungen der Ossikel erfasst. Die Schwingungsformen ändern sich stark mit der Frequenz.

#### Besondere Merkmale und Mechanismen des Mittelohrs

- Existenz einer besonders empfindlichen Vorzugsrichtung.
- Sehr flexibles Incudostapedial-Gelenk.
- Hohe Flexibilität der gesamten Ossikelkette.
- Seitlich aus dem Übertragungsweg vom Manubrium zum ovalen Fenster ausgelagertes Massenzentrum (Malleus-Kopf und Incus-Hauptkörper).
- Konisch eingezogene Form des Trommelfells.
- Überlastschutz im Incudomallear-Gelenk.
- Breitbandige Luftschall-Übertragung auch unter ungünstigen Bedingungen.
- Allgemein geringe Parameter-Empfindlichkeit.
- Erhöhte Stoß-Isolation.

#### Zusammenfassung Mittelohr:

- Das Mittelohr bildet im Bereich seiner Hauptresonanz bei etwa 800-1000 Hz einen Anpassungsübertrager zur Verbesserung der Energieübertragung von der Luft im Gehörgang in die Flüssigkeiten des Innenohres. Das theoretische Optimum wird allerdings um etwa 20 dB verfehlt.
- Das Übertragerverhältnis wird vorwiegend durch das Flächenverhältnis Trommelfell/Stapes-Fußplatte gebildet.
- Bei niedrigen und bei hohen Frequenzen bestimmen allerdings weitgehend die elastischen Elemente (Trommelfell und Bänder) und die Massen (Ossikel) das Übertragungsverhalten. Daher ist die Abweichung vom Optimum dort viel größer (60 dB bei 160 Hz, 50 dB bei 16 kHz).
- In Kombination mit der Druckverstärkung durch den Gehörgang entsteht im Frequenzbereich zwischen 800 Hz und 4 kHz eine gute Annäherung an das Optimum (bis auf etwa 20 dB).
- Das "biomechanische Design" des menschlichen Mittelohrs macht aus der Not eine Tugend: Die mangelnden Möglichkeiten, präzise Lager und Führungen zu realisieren, werden ausgeglichen durch einen ausgeklügelten Einsatz von Flexibilität und Massenverteilung.
- Dadurch werden u. A. eine gute Isolation vor Stößen auf den Kopf und eine insgesamt geringe Parameterempfindlichkeit erzielt.



Abb. 6: Das Innenohr besteht aus dem Gleichgewichtsorgan und der Cochlea. Nur die Cochlea ist für das Hören zuständig. Im Inneren der Cochlea befindet sich ein flüssigkeitsgefülltes System mit drei von einander getrennten Kanälen (Scalen).

Das Kanalsystem in der Cochlea windet sich in etwa zweieinhalb Windungen von der Basis bis zum Apex. Für die mechanische Funktionsweise spielt die Reissnersche Membran nur eine untergeordnete Rolle. Daher bilden Scala vestibuli und Scala media zusammen einen oberen Kanal und die Scala tympani allein einen unteren Kanal. Die beiden Kanäle werden getrennt durch die Basilarmembran (BM) mit dem darauf befindlichen Corti-Organ. Die mechanische Abstimmung der inhomogenen Basilarmembran führen zu einer Wanderwelle entlang dieser Membran von der Basis zum Apex. Die Abstimmung erfolgt im Wesentlichen durch die mechanische Nachgiebigkeit der BM, die von Apex bis zur Basis immmer weicher wird. Die Wanderwelle bildet an einer von der Anregungsfrequenz abhängigen Stelle ein Maximum aus. Zu jedem Ort gehört also eine Bestfrequenz. Bei Anregung mit breitbandigen Signalen findet dadurch eine spektrale Aufspaltung in der Weise statt, dass bestimmte Frequenzbereiche bestimmten Abschnitten auf der Basilarmembran entsprechen. In diesem Sinne arbeitet die Cochlea als eine Art Fourieranalysator.



Abb. 7: Querschnitt (A) durch die gesamte Cochlea und (B) durch eine Windung. Im rechten Teilbild erkennt man die Lage der Basilarmembran (BM) mit dem darauf liegenden Corti-Organ.

# Eindimensionale Modellierung der in das Kanalsystem eingebetteten Basilarmembran (passives Modell)



Abb. 8: Boxmodell der zwischen flüssigkeitsgefüllten Räumen eingebetteten abgewickelten Basilarmebran.

Setzt man voraus, dass die BM wie folgt durch einen auf die Membranfläche bezogenen Impedanz-Z''(z) beschrieben werden kann

$$p(z) = \Delta p_{BM}(z) = Z''(z) \cdot v_{BM}(z)$$

so findet man folgende Differenzialgleichung

$$\frac{\partial^2 p(\omega, z)}{\partial z^2} + 2j\omega\rho \frac{b}{A}Y''(\omega, z) \cdot p(\omega, z) = 0.$$

Hierin bedeuten  $p(\omega, z)$  Schalldruckdifferenz zwischen Ober- und Unterseite der BM bei einer Kreisfrequenz  $\omega$  und an einem Ort z auf der BM,  $Y''(\omega, z)$  flächenbezogener Admittanzbelag der BM (Kehrwert des Impedanzbelags),  $\rho$  die Dichte, b die Breite der BM und A die Querschnittsfläche beiden Kanäle. Gemäß Frequenz und Ort auf der BM bestimmt entweder der nachgiebigkeitsbestimmte oder der reibungsbestimmte Teil des Admittanzbelags die Differenzialgleichung. Man findet für die beiden Teile sehr große Unterschiede im Verhalten. Nachgiebigkeitsbestimmter Bereich:

$$\frac{\partial^2 p}{\partial z^2} - 2\rho \frac{b}{A} n_0'' \omega^2 \mathrm{e}^{\varepsilon z} p = 0$$

Reibungsbestimmter Bereich:

$$\frac{\partial^2 p}{\partial z^2} + 2j\rho \frac{b}{A} \frac{1}{w''} \omega p = 0$$

Im letzteren Fall ist der auftretende Koeffizient imaginär, im Gegensatz zum rellen Koeffizienten in der oberen Gleichung. Dies führt auf die beiden Lösungsanteile, die dem Anstieg und dem raschen Abfall der folgenden Kurven entsprechen.



Abb. 9: Örtliche Schnelleverteilung entlang der BM bei Anregung mit sinusförmigen Stapesschnellen. Die als Parameter angegebenen Zahlenwerte sind Frequenzen in kHz.

Im nachgiebigkeitsbestimmten Bereich wächst die Schwingungsamplitude der einlaufenden Wanderwelle an, wobei die Welle abgebremst wird. Gemäß der Differenzialgleichung entsteht am Übergang in den reibungsbestimmten Bereich ein Maximum, danach ein steiler Abfall Richtung Apex. Der charakteristische Ort (maximaler Erregung bei einer gegebenen Frequenz) tritt danach am Übergang vom nachgiebigkeits- zum reibungsbestimmten Bereich auf. Wie man sieht, bilden sich die hohen Frequenzen am Apex und die niedrigeren Frequenzen näher am Apex statt (die niederfrequentere Wanderwelle dringt tiefer ein). Die ungefähre Lage der Bestfrequenzen geht aus folgender Abbildung hervor.



Abb. 4.1-4: Die Basilarmembran nimmt in ihrer Breite von der Basis zum Apex hin zu, obwohl die räumlichen Verhältnisse immer beengter werden. Die Bestfrequenzen (in Hz) entlang der Basilarmembran nehmen wegen der immer größer werdenden Steife entsprechend zu.

#### Kontinuumsmodell der Basilarmembran

Das eindimensionale Modell ist ein Langwellenmodell: es setzt voraus, dass die Wellenlängen so groß (die Frequenzen so niedrig) sind, dass nur die Schwingungen senkrecht zur Basilarmembran beachtet werden müssen. Berechnungen mit einem Finite-Elemente-Modell der Basilarmembran zeigen jedoch, dass tatsächlich Quermoden auf der Basilarmembran eine wesentliche Rolle spielen, weil dadurch die Basilarmembran effektiv noch nachgiebiger wird. Tatsächlich wird die Lage des charakteristischen Ort nicht durch den Übergang zu einem reibungsbestimmten Bereich, sondern durch den Übergang von eindimensionaler Wellenausbreitung zu Ausbreitung mit Quermoden gebildet wird.



Abb. 10: Betrag der Auslenkung in nm bei Anregung mit 1 Pa im Vestibulum bei 2 kHz

#### Corti-Organ und aktive Cochlea

Das auf der BM aufliegende Corti-Organ enthält die inneren und die äußeren Haarzellen (IHC, OHC). Der Mensch besitzt ca. 3500 IHCs und mehr als dreimal so viele OHCs (die OHC sind wie in Abb. 11 ersichtlich in Dreierreihen angebracht). Von den 25.000 afferenten Nervenfasern, die Information von der Cochlea in die Hörbahn einspeisen, sind 95% an den IHC angeschlossen. Offensichtlich sind also die IHC die eigentlichen Sensoren in der Cochlea, die die bereits spektral aufgespaltenen Signale in Form von Aktionspotenzialen weiterleiten. Die weitaus größere Zahl der OHCs dient als eine Art Hilfsmotor. Der bis heute nicht genau verstandene Mechanismus führt zu einer Empfindlichkeitssteigerung des Gehörs bei niedrigen Pegeln. Er wird daher auch als cochleärer Verstärker bezeichnet. Gleichzeitig bewirkt er ein deutlich schärferes Tuning. Das bedeutet, dass die in Abb. 11 gezeigten örtlichen Schnelleverteilungen ein noch schärferes Maximum aufweisen. Dies führt zu entsprechend verschärften Frequenzgängen (Tuningkurven), also zu einer Erhöhung der Frequenzauflösung bei niedrigen Pegeln.



Abb. 11: Corti-Organ mit inneren (IHC) und äußeren (OHC) Haarzellen. Die inneren Haarzellen bilden die eigentlichen Sensoren, obwohl die Anzahl der in Dreierreihen angeordneten äußeren Haarzellen viel größer ist.

Der Hilfsmotor der OHC beruht auf der Fähigkeit dieser Haarzellen, aktiv ihre Länge zu ändern, und zwar gesteuert durch das Rezeptorpotenzial, also die elektrischen Potenzialänderungen im Inneren der Haarzelle. Ohne im Einzelnen zu verstehen, wie dies zu einer Zusatzanregung führt, kann man die gemessenen Eigenschaften des cochleären Verstärkers in etwa durch eine sehr einfache Annahme erfassen. Nach dieser schon 1986 veröffentlichten Idee von Neely & Kim bewirken die Schwingungen der OHC am Ort z der Haarzellen eine additive Zusatzanregung  $p_{OHC}(z)$ . Anstelle der "passiven" Beziehung

 $p(z) = \Delta p_{BM}(z) = Z''(z) \cdot v_{BM}(z)$ 

zwischen Druckdifferenz an der BM und der Schnelle der BM, die durch den Impedanzbelag gegeben war, fordert man nun

 $p(z) + p_{OHC}(z) = Z''(z) \cdot v_{BM}(z)$ 

wobei der zusätzlich eingespeiste, die BM also zusätzlich antreibende Druckdifferenz durch

die Schnelle der Stereocilien der OHC gesteuert wird. Wenn dieser zusätzliche Anteil mit der richtigen Phasenlage eingespeist wird, führt dies zu einer Entdämpfung des Impedanzbelags der BM, was die gewünschte Verschärfung der Tuningkurve erklären kann.



Abb. 12: Verschärfung der Tuningkurven der Cochlea bei niedrigen Anregungspegeln. Dargestellt ist die an einem bestimmten Ort auf der BM gemessene BM-Auslenkung als Funktion der Frequenz. Man beachte, dass sich mit abnehmendem Pegel nicht nur die Tuningschärfe verändert, sondern auch die Frequenz verschiebt, bei der das Maximum auftritt.

Die Wirkung des cochleären Verstärkers ist auf niedrige Pegel begrenzt, weil die aktiven Längenänderungen bereits bei einem Pegel von 40-50 dB über der Hörschwelle in die Sättigung gehen. Bei einer weiteren Steigerung des Pegels wird daher der relative Anteil des cochleären Verstärkers immer geringer. Dies wird durch die folgende Abbildung bestätigt.



Abb. 13: Auslenkung der Basilarmembran als Funktion des Eingangspegels in der Nähe der jeweiligen Bestfrequenz (Messergebnisse und Interpretation von Johnstone et al. 1986)

Offensichtlich wird durch diese Kennlinie beschrieben, dass die Ausgangsdynamik kleiner

als die Dynamik der eingespeisten Signale ist. Der cochleäre Verstärker liefert also einen wesentlichen Beitrag zur Kompression der Signaldynamik im Verlauf der Signalverarbeitung durch unser Gehör. Die richtige Nachbildung dieser Kompression spielt auch in Modellen, die die auditive Wahrnehmung beschreiben, eine wichtige Rolle.

#### Zusammenfassung Cochlea

- Die Cochlea realisiert eine spektrale Zerlegung der vom Mittelohr kommenden Signale durch die abgestimmte Basilarmembran (Erhöhung der Elastizität und der Breite entlang des Wegs von der Basis zum Apex).
- Der im Vestibulum erzeugte Schalldruck löst Wanderwellen auf der inhomogenen Basilarmembran aus. Wellen in den Flüssigkeiten spielen dagegen nur eine untergeordnete Rolle, weil die Wellenlängen in den Flüssigkeiten, außer bei den höchsten Audiofrequenzen, viel länger als die Kanäle sind.
- Die Wanderwelle wird auf dem Weg von der Basis zum Apex durch die zunehmende Nachgiebigkeit abgebremst, so dass sich an einem durch die anregende Frequenz bestimmten "charakteristischen Ort" zeitlich aufeinander folgende Wellenzüge überlagern.
- Dadurch entsteht am charakteristischen Ort bei der zugehörigen "Bestfrequenz" eine maximale Antwort auf die Erregung.
- Zu einem Ort auf der Basilarmembran gehört daher ein bestimmtes Filter, das eine "Tuningkurve" mit Maximum bei der Bestfrequenz besitzt.
- Das auf der Basilarmembran aufgesetzte Corti-Organ trägt ca. 3500 innere und etwa 3.5 mal so viele äußere Haarzellen (IHC und OHC), die durch ihren Ort bestimmten Frequenzen zugeordnet sind.
- Die IHC sind Schwingungssensoren, die ihre Signale an das zentrale Hörorgan weiterleiten.
- Die äußeren Haarzellen (OHC) bewirken eine Verstärkung, die auf niedrige Pegel beschränkt ist (cochleärer Verstärker). Sie erhöhen somit die Hörschwelle.
- Die OHC bewirken bei niedrigen Pegeln gleichzeitig eine Verschärfung der Tuningkurven.

# Psychoakustik und Sprachperzeption bei Normal- und Schwerhörigen

Birger Kollmeier / Oldenburg

# 1. Wahrnehmungs-Grundgrößen

Das menschliche Gehör ist in faszinierender Weise optimal an das Empfangen von akustischen Signalen aus der Umwelt angepaßt, insbesondere an das Verstehen von Sprache. Weil das akustische Sprachsignal sich zeitlich stark ändert und unterschiedliche Frequenzanteile aufweist, benötigt Ohr für die das Sprachwahrnehmung die Fähigkeit, zu jedem Zeitpunkt die Intensität wahrzunehmen. mit der jede Frequenz momentan vorliegt. Wesentliche Grundgrößen der Wahrnehmung sind daher die Umsetzung verschiedener Schallintensitäten in subjektiv empfundene Lautheit, die Umsetzung verschiedener Frequenzen in subjektiv empfundene Tonhöhen, die Umsetzung verschiedener Zeitdauern und Rhythmen in subjektiv empfundene Zeitmuster, die Umsetzung von akustischen Signalen in subjektiv empfundene Klänge, und das Trennen verschiedener Klänge (z. B. Nutzsignal vom störenden Hintergrundsignal). Diese Beziehungen zwischen dem akustischen Reiz und der subjektiven Wahrnehmung, die durch psychologische Meßmethoden erfaßt werden können, werden in der Psychoakustik untersucht.

# 1.1 Intensitätsabbildung (Lautheit) und Intensitätsauflösung

Als **Ruhehörschwelle** wird der Pegel eines soeben wahrnehmbaren Sinustons bei einer vorgegebenen Frequenz in Ruhe (d. h. ohne Vorliegen störender Signale) bezeichnet. Bei niedrigen und bei hohen Frequenzen muß ein wesentlich höherer Schalldruckpegel erzeugt werden als bei mittleren, damit der Ton von normalhörenden Menschen schwellenhaft wahrgenommen werden kann (vgl. Abbildung 1). Wenn ein Sinus-Ton im Pegel kontinuierlich erhöht wird, nimmt die Versuchsperson nach Überschreiten der Ruhehörschwelle einen immer lauteren Ton wahr. Der Grad, in dem diese Lautheitswahrnehmung mit zunehmenden Tonpegel ansteigt, hängt von der Frequenz ab (vgl. Abbildung.1): Ausgehend von einem Sinus-Ton bei 1000 Hz kann man bei jeder Frequenz denjenigen Tonpegel bestimmen, der zum gleichen Lautheitseindruck führt. Diese Kurve wird als Isophone oder Kurve gleicher Pegellautstärke bezeichnet. Während die 10 Phon-Isophone bei 1000 Hz genau 10 dB über der Ruhehörschwelle ist und zu niedrigen und hohen Frequenzen ähnlich steil ansteigt wie die Ruhe-Hörschwelle, flachen die zu höheren Pegeln gehörenden Isophonen zunehmend ab, so daß die 120 Phon-Isophone, die ungefähr der Schmerzschwelle bei Normalhörenden für alle Frequenzen entspricht, etwa den gleichen Schalldruck für alle Frequenzen angibt.



Abbildung 1: Ruhehörschwelle und Isophonen

Um den Lautheitsanstieg mit zunehmendem Pegel zu bestimmen, d. h. beispielsweise die Pegelerhöhung, die für eine Verdopplung der empfundenen Lautheit notwendig ist, bedient man sich verschiedener subjektiver Skalierungsmethoden. Bei der **freien Größenschätzung** nach Stevens und Mitarbeitern (Stevens, 1957) soll die Versuchsperson die jeweils empfundene Lautheit mit einer beliebigen Zahl kennzeichnen. Trägt man nun die von der Versuchsperson gegebenen Zahlenwerte über die angebotenen Pegel auf, erhält man die nach Stevens benannte **Potenzfunktion** 

$$N[sone] = \left(\frac{I}{I_0}\right)^{\alpha} \quad . \tag{8.1}$$

Dabei bezeichnet N die Lautheit in der Einheit "sone", I die Intensität des Testtons und I<sub>0</sub> die Intensität des Referenz-Sinustons bei 1 kHz mit einem Pegel von 40 dB SPL, der genau der Lautheit 1 sone entspricht. Der Exponent  $\alpha$  nimmt im Mittel über viele Versuchspersonen und Experimente den Wert von etwa 0,3 an. Das bedeutet, daß bei einer Erhöhung der Schallintensität um 10 dB die Lautheit ungefähr um den Faktor 2 ansteigt (da 10 dB eine Verzehnfachung der Intensität bedeutet und  $10^{0,3} \cong 2$  ist). Die Lautheit hängt damit in anderer Weise von der Schallintensität ab als der Schallpegel. Die Lautheit in sone darf auch nicht mit der Lautstärke (in Phon) verwechselt werden, da der erste Begriff die Größe einer subjektiven Empfindung kennzeichnet und der zweite Begriff einen Pegel bezeichnet (d. h. eine physikalische Reizgröße), die zu einer vorgegebenen Lautheitsempfindung führt (nämlich zu der Lautheit, die für alle Lautstärken auf einer Isophonen gleich ist). Daher wird der zweite Begriff in der Normung auch als "Pegellautstärke" bezeichnet.

Eine andere Möglichkeit zur Bestimmung der subjektiven Lautheitsempfindung besteht bei der kategorialen Skalierung darin, daß der Versuchsperson eine gewisse Zahl von Lautheitskategorien (z. B. "sehr leise", "leise", "mittel", "laut" und "sehr laut") als Antwortmöglichkeiten vorgegeben wird. Bei dem Verfahren der "Würzburger Hörfeldskalierung" nach Heller (1985) wird jede dieser groben Kategorien in einem zweiten Schritt in 10 feinere Kategorien unterteilt. Bei der "Oldenburger Hörflächenskalierung" (Hohmann und Kollmeier, 1995), die sich insbesondere für werden schwerhörigen Patienten eignet. Messungen mit neben den fünf Hauptkategorien vier Zwischenkategorien und die Grenzfälle "nicht gehört" und "zu laut" vorgegeben, so daß die Gesamtzahl der Antwortkategorien nicht 50, sondern 11 beträgt. Trägt man die Antwort der Versuchsperson in Kategorial-Einheiten über dem Pegel des Testschalls auf, erhält man im mittleren Pegelbereich eine Gerade, d. h. die empfundene Kategorial-Lautheit ist dort proportional zum Logarithmus der Intensität (vgl. Abbildung 2). Abweichungen zu der im logarithmischen Maßstab aufgetragenen Lautheit in sone treten bei sehr niedrigen und bei sehr hohen Pegeln auf:



Abbildung 2: Sone-Skala und Kategorial-Skala über den Pegel

Bei sehr niedrigen Pegeln stehen für die relativ grobe Kategorial-Skalierung nicht genügend Antwortalternativen zur Verfügung, so daß die Kurve erst bei mittleren Pegeln ansteigt, während die Lautheit in sone gegen Null geht und ihr Logarithmus dementsprechend steil abfällt. Bei sehr hohen Pegeln ist der Antwortbereich der Kategorial-Lautheit ebenfalls begrenzt, so daß hier ein Sättigungseffekt auftritt, während der Logarithmus der Lautheit in sone weiter linear ansteigt. Dieser Effekt ist auf den "Logarithmic Response Bias" (Poulton, 1989) zurückzuführen, d. h. die Versuchspersonen beurteilen bei der Erhöhung einer großen Zahl nicht mehr den absoluten Zuwachs dieser Zahl, sondern das Verhältnis der neuen Zahl zur alten Zahl. Damit tendieren die Versuchspersonen bei hohen Pegeln bei der absoluten Größenschätzung dazu, ihre Empfindungsgröße mit dem Logarithmus der angenommenen, großen Zahlen zu beschreiben. Dieser Effekt tritt bei der Kategorial-Skalierung mit einer wesentlich kleineren Zahl von Antwortalternativen nicht auf, so daß die ungefähre Proportionalität zwischen dem Logarithmus der Lautheit in sone und dem absoluten Wert der Kategorial-Skala plausibel erscheint (paralleler Verlauf der Kurven in Abbildung 2 über einen relativ großen Pegelbereich). Eine logarithmische Beziehung zwischen der Reizgröße (in unserem Fall die Intensität I) und der Empfindungsgröße E folgt auch aus dem **Weber-Fechner'schen Gesetz**, das eines der ersten und wichtigsten Gesetze der Psychophysik darstellt. Weber und Fechner konnten nachweisen, daß der Reizunterschied  $\Delta I$ , der zu einer eben wahrnehmbaren Änderung der Empfindung  $\Delta E$  führt, proportional zu der absoluten Größe I des Reizes ist, so daß der sogenannte **Weber-Bruch**  $\frac{\Delta I}{I}$  für alle Werte von I annähernd konstant sein muß und proportional der kleinsten wahrnehmbaren Empfindung  $\Delta E$  ist:

$$\Delta \mathbf{E} = \mathbf{k} \cdot \frac{\Delta \mathbf{I}}{\mathbf{I}} \tag{8.2}$$

Durch Aneinanderreihen von eben merklichen Unterschieden ∆E (was mathematisch der Aufintegration von Gleichung 8.2 entspricht) läßt sich nun eine Beziehung zwischen der Empfindungsgröße E und der Reizgröße I herleiten:

$$E = k' \cdot \log_{10} I + k''$$
 (8.3)

k' und k" sind Konstanten, die aus den Randbedingungen festgelegt werden. Diese logarithmische Abhängigkeit der Wahrnehmungsgröße von der Reizgröße entspricht gerade dem Zusammenhang zwischen Kategorial-Lautheit und Tonpegel, d. h. dem Logarithmus der Intensität. Dieses Gesetz entspricht auch der sone-Skala unter der Annahme, daß die empfundene Lautheit durch den Logarithmus der von der Versuchsperson angegebenen Zahl, und nicht durch die Zahl selber wiedergegeben wird.

Neben der bisher behandelten absoluten Wahrnehmung der Schallintensität interessiert auch die **differentielle Wahrnehmung**, d. h. die kleinste wahrnehmbare Änderung einer Schallintensität. Sie wird auch als JND ("just noticeable difference") bezeichnet und tritt als  $\Delta I$  im Weber-Fechner'schen Gesetz (Gleichung 8.2) auf. Für die meisten Schalle beträgt die JND etwa 1 dB, d. h.

$$\frac{\Delta I + I}{I} = 10^{\frac{1dB}{10}} \Longrightarrow \frac{\Delta I}{I} = 0,259$$
(8.4)

In der Tat mußte die erste logarithmische Pegelskala, deren Einheit das Bel ist (nach Alexander Graham Bell, der unabhängig von Philip Reis das Telefon erfand und zu einem kommerziellen Erfolg führte, 1 Bel =  $\log_{10} (I/I_0)$ ) in eine 10-fach feinere Skala, das **Dezibel (dB)** unterteilt werden. Damit entspricht eine Einheit dieser Skala gerade einem kleinsten hörbaren Pegelunterschied. Während bei breitbandigen Signalen über einen

weiten Pegelbereich Webers Gesetz gültig ist und der kleinste wahrnehmbare Pegelunterschied 1 dB beträgt, ist bei schmalbandigen Signalen (wie Sinustönen) dieser Unterschied bei kleinen Pegeln größer als 1 dB und nimmt mit zunehmendem Pegel stetig ab. Diese Abweichung wird als "near miss to Weber's law" bezeichnet und hängt mit der speziellen Anregung der Basilarmembran bei Sinustönen zusammen.

## 1.2 Tonhöhe und Frequenzauflösung

Wird die Frequenz eines Sinustons erhöht, führt dies zu der Wahrnehmung einer Tonhöhe. Wenn diese Tonhöhenwahrnehmung Weberansteigenden dem Fechner'schen Gesetz entsprechend proportional zum Logarithmus der Frequenz wäre. entspräche dies genau der musikalischen Tonhöhenempfindung: Eine Oktave (Frequenzverhältnis 2:1) würde als doppelt so hoch empfunden werden, unabhängig von der jeweiligen Frequenz. Tatsächlich folgt die Tonhöhenwahrnehmung dieser Gesetzmäßigkeit nur in einem begrenzten Bereich mittlerer Frequenzen, der bei musikalisch ausgebildeten Versuchspersonen sich noch etwas erweitern läßt. Läßt man Versuchspersonen jedoch unabhängig von musikalischen Intervallen die empfundene Tonhöhe im Vergleich zum Referenzton bei 1 kHz angeben, so erhält man die Mel-Skala als subjektive Tonhöhenempfindungskala. Sie entspricht ungefähr der Frequenz-Orts-Transformation auf der Basilarmembran, bei der die verschiedenen Frequenzen zu einer maximalen Auslenkung an unterschiedlichen Orten führen (gepunktete Linie in Abb. 3). Die gleiche Beziehung zwischen subjektiver Tonhöhe und objektiver Frequenz liefert die in Abb. 3 dargestellte Bark-Skala (benannt nach dem deutschen Physiker und Akustiker Heinrich Barkhausen), die auf dem Konzept der Frequenzgruppe beruht (Zwicker und Mitarb. 1957).



Abbildung 3: Mel-Skala, Bark-Skala, ERB-Skala

Sie basiert auf der Frequenzabhängigkeit der Lautheitswahrnehmung und der Maskierung: Bei der Lautheitssummation wird die empfundene Lautheit eines schmalbandigen Signals (z. B. eines Schmalbandrauschens) bei gleicher Leistung, aber verschiedener Bandbreite bestimmt (vgl. Abbildung 4). Wenn dieselbe Leistung auf

einen größeren Frequenzbereich verteilt wird, steigt die wahrgenommene Lautheit auf etwa das Doppelte an (entspricht ungefähr 10 dB Lautstärkengewinn). Bleibt die Leistung jedoch in einem Frequenzbereich, der kleiner als die **Frequenzgruppenbreite** ist, hängt die wahrgenommene Lautheit nicht von der Bandbreite ab, so daß aus dieser Bandbreitenabhängigkeit auf die Größe der Frequenzgruppe geschlossen werden kann. Die Frequenzgruppenbreite beträgt etwa 100 Hz unterhalb von 500 Hz und etwa 1/5 der Frequenz oberhalb von 500 Hz (vgl. Abbildung 3).



Abbildung 4: Bandbreitenabhängigkeit der Lautheit

Anschaulich versteht man unter der Frequenzgruppe diejenige Bandbreite im Gehör, innerhalb derer sämtliche Signale gemeinsam verarbeitet und zu einem "Erregungspegel" zusammengefaßt werden, der der Lautheitswahrnehmung zugrunde liegt. Wenn das Signalspektrum in verschiedene Frequenzgruppen fällt, wird nicht mehr die Leistung aufsummiert, sondern die Teil-Lautheiten in den verschiedenen Frequenzgruppen.



Abbildung 5: Erregungspegelmuster

Eine Vorstellung zum Zustandekommen der Frequenzgruppe und der Frequenzauflösung im auditorischen System ist die Erregungsverteilung auf der Basilarmembran (vgl. Abbildung 5). Bei einem Sinus-Ton mit  $f_1$  wird die Basilarmembran nicht nur bei dieser Frequenz, sondern im schwächerem Maße auch bei den darüber- bzw. darunterliegenden Frequenzen angeregt. Wenn auf der x-Achse die Frequenz in Bark aufgetragen ist, kann diese Verbreiterung der Erregung als ein dreieckförmiges Muster dargestellt werden, dessen Flanken zu tiefen Frequenzen mit etwa 25 dB pro Bark ansteigen und zu hohen Frequenzen mit etwa 10 dB pro Bark abfallen. Bei hohen Pegeln werden diese Flanken flacher. Wird nun ein weiterer Sinus-Ton mit einer Frequenz  $f_2$  mit einem Pegel angeboten, der unterhalb dieses

Erregungspegels bei der Frequenz  $f_2$  liegt, wird dieser zusätzliche Ton vom ersten Ton vollständig **maskiert**, d. h. der zweite Ton wird im Beisein des ersten Tons nicht mehr gehört. Durch derartige Verdeckungsexperimente kann die Form und Steilheit der Erregungspegel-Verteilung über der Frequenz ausgemessen werden.

Um aus derartigen Experimenten auf die Frequenz-Gruppe zu schließen, wird meist ein **Notched-Noise** verwendet, bei dem ein Rauschen nur unterhalb einer Grenzfrequenz  $f_u$  und oberhalb einer Frequenz  $f_o$  vorliegt. Je schmaler die zwischen  $f_u$  und  $f_0$  liegende spektrale Lücke ("notch") ist, desto höher liegt die **Mithörschwelle** eines in diesem Frequenzbereich liegenden Tons, d. h. der Pegel des Tons muß relativ hoch sein, damit er gehört wird. Aus der Abnahme dieser Mithörschwelle mit zunehmender Breite der spektralen Lücke kann auf die Frequenzgruppenbreite geschlossen werden, die in Einheiten einer "equivalent rectangular bandwidth"(ERB) nach Moore und Glasberg (1987) gemessen wird. Darauf beruht die ebenfalls in Abb. 3 angegebene ERB-Skala, die sehr ähnlich der Bark-Skala ist.

Die Verteilung der Erregungspegel über der Frequenz kann für die Berechnung der werden. sich Lautheit verwendet wobei die überlappenden Bereiche des Erregungspegel-Verlaufs für verschiedene Freguenzen nur einmal berücksichtigt werden. Dazu wird bei jeder Frequenzgruppe der Erregungspegel gemäß dem Potenzgesetz (Gleichung 8.1) umgeformt. Die Gesamt-Lautheit wird dann durch Integration (Aufsummation) der spezifischen Lautheiten über sämtliche Frequenzen gebildet. Diese Berechnung entspricht dem nach Iso 532 B genormten Lautheitsmodell nach Zwicker, das für stationäre Schallsignale die subjektiv empfundene Lautheit in sone sehr gut berechnen kann. Dabei werden sowohl spektrale Maskierungseffekte als auch Lautheitssummations-Effekte und Kombinationen dieser Effekte richtig vorhergesagt. Eine Anwendung auf zeitlich stark schwankende Schall (wie z. B. Sprache) ist jedoch problematisch, da das Modell für stationäre Signale entwickelt wurde.

Die Frequenzgruppe läßt sich auch mit einer Reihe weiterer Maskierungsexperimente bestimmen, bei denen sich der maskierende Schall (meistens ein Rauschen) stark im Frequenzbereich ändert, z. B. sinusförmige Änderung im Frequenzbereich ("ripple-noise"), Tiefpaßrauschen oder Hochpaßrauschen. Für alle Fälle werden ähnliche Frequenzgruppenbreiten gemessen, wobei sich neben der Breite dieser Frequenzgruppen auch die effektive Form der Filter im Gehör angeben läßt, mit denen jeweils ein begrenzter Frequenzbereich analysiert und zu einer Frequenzgruppe zusammengefaßt wird. Die einfachste Form ist rechteckförmig, eine realistischere Filterform ist ein gerundetes Exponentialfilter dessen Flanken die Verhältnisse im Erregungspegelmuster wiederspiegeln. Man darf sich eine Frequenzgruppe nicht als eine starre, bei einer festen Frequenz gelegene rechteckförmige Filterung des Schallsignals vorstellen, sondern eine kontinuierliche, bei jeder Mittenfrequenz auftretende Filterung mit einer abgerundeten, kontinuierlichen Filterform.

relativ geringen Im Gegensatz zu der Frequenzauflösung durch die Frequenzgruppenfilter bei der Maskierung steht die sehr hohe Frequenzauflösung bei der **Diskrimination** verschiedener Frequenzen, die nicht gleichzeitig, sondern in Folge angeboten werden. Dies bei der Musikwahrnehmung (z. B. beim Stimmen von Instrumenten) wichtige Phänomen der Tonhöhenunterscheidung ermöglicht einen kleinsten hörbaren Frequenzunterschied von etwa 3 Hz für Frequenzen unterhalb von 500 Hz und etwa 0,6 % für Frequenzen über 1000 Hz. Dies entspricht ungefähr 1/30 Bark, d.h. die Tonhöhenunterscheidung ist wesentlich feiner als die Maskierung im Frequenzbereich. Dies liegt an der massiven Parallelverarbeitung im Gehör: Wenn die Frequenz eines einzelnen Sinustons um wenige Herz verschoben wird, verschiebt sich das ganze Erregungsmuster auf der Basilarmembran, so daß im Mittel über sämtliche beteiligten Nervenfasern selbst ein kleiner Unterschied in der Verschiebung des Erregungsmusters feststellbar ist, obwohl das von einem einzelnen Ton hervorgerufene Erregungsmuster selbst relativ breit sein kann. Daher mißt die Frequenzauflösung (Diskrimination) eine andere Leistung des Gehörs als die Maskierung im Frequenzbereich, die auch als **spektrale Integration** bezeichnet werden kann.

# 1.3 Zeitliche Verarbeitung im Hörsystem

Neben der spektralen Verschmierung (Integration) und Frequenzauflösung spielt die zeitliche Verarbeitung von akustischen Signalen im auditorischen System eine wichtige Rolle, die ebenfalls durch eine zeitliche Verschmierung (Zeitintegration) und eine Zeitauflösung gekennzeichnet werden kann. Die zeitliche Integration bezeichnet dabei die Fähigkeit, einen lang andauernden, stationären Klang bei gleichem Pegel als lauter wahrzunehmen als einen kurzen Klang. Diese Eigenschaft kann mit einem Maskierungsexperiment demonstriert werden (vgl. Abbildung 6), bei der die Mithörschwelle für einen Signalton bestimmt wird, d. h. der Pegel, bei dem der Ton im Rauschen soeben noch hörbar ist. Sie nimmt mit zunehmender Dauer T des Testtons ab und erreicht für Werte von etwa 200 ms einen stabilen Wert, der durch weitere Verlängerungen des Tons nicht mehr verändert wird. Diese Eigenschaft kann durch eine einen 200 ms Energieintegration über Bereich von erklärt werden. Bei Innenohrschwerhörenden ist diese Energie-Summation gestört, so daß die Schwelle bei zunehmender Dauer T nicht weiter abnimmt.



Abbildung 6: Zeitliche Integration (Schema, s. Text)

Die zeitliche Verschmierung der internen Repräsentation akustischer Signale läßt sich durch die psychoakustisch meßbaren Phänomene der **Nachverdeckung** und **Vorverdeckung** beschreiben. Bei der Nachverdeckung wird ein Testsignal (z. B. ein kurzer Testtonpuls) zeitlich nach dem Abschalten eines Maskierungssignals (z. B. Rauschen) angeboten und der Pegel des Tons wird solange variiert, bis er soeben nicht mehr hörbar ist (Mithörschwelle). Diese Mithörschwelle nimmt mit zunehmendem Abstand des Testtons vom Maskierende ab und erreicht bei etwa 200 ms die Ruhehörschwelle. Dabei ist der Verlauf dieser Nachverdeckungskurve abhängig von der Maskiererdauer, d. h. vom Adaptationszustand des auditorischen Systems (vgl. Abbildung 7). Bei der Vorverdeckung wird der Testton dagegen nicht nach dem Abschalten, sondern zeitlich vor dem Anschalten eines Maskierers angeboten, wobei sich eine ähnliche Verschmierung der Mithörschwelle als Funktion der Verzögerungszeit ergibt wie bei der Nachverdeckung. Allerdings ist der Zeitbereich, über den sich die Vorverdeckung erstreckt, wesentlich kürzer (etwa 10 ms) und es ergibt sich keine vergleichbar starke Abhängigkeit von der Dauer des Maskierers.

Einen kombinierten Effekt von Vor- und Nachverdeckung beobachtet man bei der Lückendetektion (Englisch: Gapdetection), bei der die kleinste in einem Rauschen wahrnehmbare Pausendauer gemessen wird (vgl. Abb. 8). Da aufgrund der zeitlichen Verschmierung im auditorischen System kleine Lücken im Rauschen nicht wahrgenommen werden können, liegt bei Normalhörenden bei der Verwendung von breitbandigen Rauschen die minimal detektierbare Lückendauer bei etwa 8 ms. Da sie ein einfach und schnell zu bestimmendes Maß für die zeitliche Verarbeitung ist, wird sie oft zur Charakterisierung der gestörten zeitlichen Verarbeitung bei Schwerhörenden eingesetzt, wobei aufgrund starker interindividueller Schwankungen Werte zwischen 8 und 100 ms auftreten können.




Abbildung 8: Gap-detection

Eine weitere Eigenschaft der zeitlichen Verarbeitung akustischer Signale ist die Modulationswahrnehmung, d. h. die Wahrnehmung von Schallen mit einer aufgeprägten (z. B. sinusförmig variierenden) Einhüllenden (vgl. Abb. 9). Als Trägersignale können Sinussignale oder Rauschen verwendet werden, während als Modulationssignale (d. h. Verformung der Einhüllenden) zumeist sinusförmige Signale mit einer bestimmten Modulationsfrequenz fmod verwendet werden. Eine derartige Signalform ist sehr ähnlich der Form von Sprache, weil bei der Sprachartikulation ein Trägersignal (z. B. die Schwingung der Stimmritze) durch den Vokaltrakt in der Intensität zeitlich stark verändert wird. Im Modulationsfrequenzbereich zwischen 0 und etwa 10 Hz werden Modulationen als Schwankungen der Lautstärke wahrgenommen, die Empfindung wird als Schwankungsstärke bezeichnet. Bei Frequenzen zwischen etwa 10 und 20 Hz entsteht ein "knarrender" Klangeindruck, der auch als R-Rauhigkeit bezeichnet wird, während Modulationen zwischen etwa 20 Hz und 80 Hz als Rauhigkeit des Klanges wahrgenommen werden können. Bei noch höheren Modulationsfreguenzen treten im Frequenzspektrum des akustischen Signals Seitenlinien auf, die aufgrund der spektralen Verfärbung als Klangänderungen wahrgenommen werden können.



Abbildung 9: Moduliertes Signal, PMTF für breitbandige Träger, d. h. der kleinste detektierbare Modulationsgrad als Funktion der Modulationsfrequenz

Als Maß für die Zeitauflösung kann die psychoakustische Modulationstransfer-Funktion bestimmt werden. d. h. für jede Modulationsfrequenz derjenige Modulationsgrad m, der soeben noch wahrgenommen werden kann. Sie zeigt für Modulationen von etwa 4 Hz ein Maximum (dies entspricht in etwa der Sprachsilbenfrequenz und ist daher gut an die zu empfangene Sprache angepaßt) und für zunehmende Modulationsfrequenzen wird die Empfindlichkeit geringer. Diese Modulations-Tiefpaßcharakteristik des auditorischen Systems korrespondiert zu der zeitlichen Verschmierung, die in den Vor- und Nachverdeckungsexperimenten aufgezeigt wurde.



Abbildung 10: Verarbeitungsmodell nach Dau et al.

Um die verschiedenen zeitlichen Eigenschaften des Hörsystems (zeitliche Integration, Vor- und Nachverdeckung, Modulationswahrnehmung) mit möglichst wenig Annahmen beschreiben zu können, wurden Hörmodelle zur zeitlichen Verarbeitung aufgestellt. Bei dem in Abbildung 10 schematisch dargestellten Verarbeitungsmodell nach Dau et al. (1997) wird die Verarbeitung unterschiedlicher Frequenzen zunächst durch eine Filterbank beschrieben, die das Eingangssignal in verschiedene Frequenzgruppenbreite Bandpaß-Signale aufteilt. In jedem Bandpaßsignal wird eine Halbwellengleichrichtung mit Tiefpaßfilterung durchgeführt, das ungefähr der Aktion der Haarzellen entspricht und eine Extraktion der Einhüllenden bewirkt. Die zeitlichen Verarbeitungs-Eigenschaften werden durch die nachgeschalteten Adaptations-Schleifen nachgebildet, bei denen das Eingangssignal durch das tiefpaßgefilterte Ausgangssignal geteilt wird. Dadurch wird eine gewisse Adaptation an den Mittelwert des Eingangssignals ermöglicht, während schnelle Änderungen (Fluktuationen) ohne weitere Beeinflussung durchgelassen werden. Durch Hintereinanderschalten mehrerer dieser Nachregelschleifen mit unterschiedlichen Zeitkonstanten zwischen 1ms und 500 ms wird approximativ eine logarithmische Kompression für langsame Änderungen im Eingangssignal erreicht und eine unveränderte Durchlässigkeit für schnelle Änderungen im Eingangssignal. Mit dieser Modellstruktur lassen sich Vor- und Nachverdeckung sowie die Lückendetektion (Gap detection) bereits sehr gut beschreiben. Die zeitliche Integration von Testtönen und die Modulationswahrnehmung kann durch Einfügen einer Modulationsfilterbank erreicht werden, die für jede Mittenfrequenz die Einhüllenden-Fluktuationen in verschiedene Modulationsfrequenzen aufspaltet. Am Ausgang dieser Modulationsfilterbank steht ein bestimmtes, zeitlich variables Muster zur Verfügung, das die interne Repräsentation von akustischen Signalen darstellen soll. Mit dieser internen Repräsentation lassen sich verschiedene Experimente einfach dadurch beschreiben, daß eine bestimmte, minimale Änderung in dieser Repräsentation zur Wahrnehmung eines Unterschieds im Klang führt, der gerade zur Detektion des "gesuchten" akustischen Stimulus führt.

## 1.4 Binaurale Interaktion

Neben den bisher beschriebenen Leistungen des Gehörs bei der Intensitäts-, Frequenz-, und Zeitverarbeitung im auditorischen System ist die binaurale Interaktion, d. h. die im Gehirn stattfindende vergleichende Verarbeitung der an beiden Ohren anliegenden Signale von besonderer Wichtigkeit für das Hören in natürlicher Umgebung. Es trägt signifikant zur Hallunterdrückung, zur Ortung (Lokalisation) von Schallquellen im Raum und zur Unterdrückung von "unerwünschten" Störgeräuschquellen in realen akustischen Situationen bei.



Abbildung 11: Physikalische Merkmale für die Lokalisation

Wenn eine Schallquelle so im Raum angeordnet ist, daß der von ihr ausgesendete Schall mit einem bestimmten Einfallswinkel auf den Kopf des Hörers fällt, bewirkt dieser (winkelabhängige) Schalleinfall eine **interaurale Zeitverzögerung** (d. h. der Schall erreicht das der Schallquelle zugewandte Ohr eher als das abgewandte Ohr), einen **interauralen Intensitätsunterschied** (d. h. das der Schallquelle zugewandte Ohr empfängt eine höhere Schallintensität), sowie eine von der Einfallsrichtung abhängige **spektrale Verfärbung** (d. h. der Frequenzgehalt des empfangenen Schalls wird je nach Einfallsrichtung unterschiedlich verändert vgl. Abb. 11.). Aufgrund dieser akustischen Merkmale ist der normalhörende Mensch bei breitbandigen Signalen in der Lage, den Ort einer Schallquelle aus der Vorne-Richtung mit einer Ungenauigkeit von etwa 1 Grad bzw. bei seitlichem Einfall oder Schalleinfall von oben mit einer Ungenauigkeit von etwa 5 Grad aufzulösen. Diese erstaunlich hohe Ortungsleistung wird allerdings bei schmalbandigen Signalen und bei Vorliegen von Nachhall bzw. Echo eingeschränkt.

Dieser Effekt ist darauf zurückzuführen, daß die bei einer bestimmten Frequenz anliegenden interauralen Intensitäts- und Zeitunterschiede nicht eindeutig zu einer bestimmten Einfallsrichtung gehören, sondern daß mehrere Schalleinfallsrichtungen bei derselben Frequenz zu denselben interauralen Intensitäts- und Zeitunterschieden führen können. Beispielsweise führen sämtliche Einfallsrichtungen in der Medianebene zu einer interauralen Zeit- und Laufzeitdifferenz von ungefähr 0, so daß bei schmalbandigen Signalen nicht sicher zwischen der Vorne-, Oben-, und Hinten-Einfallsrichtung unterschieden werden kann und es zu **Richtungsverwechselungen** kommt. Ein Schalleinfall von vorne links kann in ähnlicher Weise mit einem Schalleinfall von hinten links und Einfallsrichtungen verwechselt werden, die auf den sogenannten "**Cone of Confusion**" liegen. Bei schmalbandigen akustischen Signalen werden zudem bestimmte Frequenzen zu bestimmten Schalleinfallsrichtungen zugeordnet ("**richtungsbestimmende Bänder**" nach Blauert, 1974), so daß hier die spektrale Information teilweise die Information aus den interauralen Zeit- und Intensitätsunterschieden überschreibt. Erst bei breitbandiger Schalldarbietung kann durch den Vergleich über mehrere Frequenzen hinweg eine eindeutige, sämtliche Mehrdeutigkeiten bei schmalbandiger Signaldarbeitung vermeidende Lokalisation durchgeführt werden.

Die Rolle der interauralen Zeit- und Intensitätsdifferenzen für die Lokalisation und die binaurale Verarbeitung kann mit Kopfhörer-Experimenten erforscht werden, bei denen jeder dieser physikalischen Parameter einzeln variiert werden kann. So stellt sich heraus, daß die interauralen Zeitdifferenzen insbesondere bei niedrigen Frequenzen (unter 1,5 kHz) eine dominierende Rolle für die Lokalisation spielen und der kleinste wahrnehmbare interaurale Zeitunterschied etwa 20 µs beträgt, während er oberhalb von 1,5 kHz etwa 50 µs beträgt. Die interauralen Intensitäts-Unterschieds-Schwellen betragen dagegen unterhalb von 1,5 kHz etwa 3 dB und oberhalb von 1,5 kHz etwa 1 dB, so daß sie besonders bei hohen Frequenzen für die Lokalisation dominant sind. Diese bereits von Lord Rayleigh (1877) formulierte Duplex-Theorie des binauralen Hörens hat ihre physikalische Begründung darin, daß bei niedrigen Frequenzen aufgrund der (relativ zum Kopfdurchmesser) großen Schallwellenlänge nur geringe zwischen Intensitätsunterschiede beiden den Ohren auftreten. SO daß Intensitätsunterschiede erst bei hohen Frequenzen veläßlich ausgewertet werden können. Bei niedrigen Frequenzen können die Neuronen im auditorischen System der Schallwellenform noch sehr gut folgen und daher interaurale Zeit-Unterschiede gut detektieren, während bei hohen Frequenzen interaurale Zeit-Differenzen nur im Einhüllenden-Verlauf der Signale, nicht jedoch in der (von den Neuronen nicht mehr aufgelösten) Feinstruktur der übertragenden Signale erfaßt werden können.



Abbildung 12: Binaurale Störgeräuschbefreiung

Zur Erfassung der **Störgeräuschbefreiung** im binauralen System kann ein psychoakustisches Experiment verwendet werden, bei dem die Detektion von Testtönen im Rauschen zwischen der binauralen und der monauralen Situation verglichen werden (vgl. Abb. 12). Zunächst wird der Versuchsperson ein Ton im Rauschen angeboten, dessen Pegel solange verringert wird, bis die Versuchsperson ihn nicht mehr hört (Mithörschwelle). Wenn anschließend dasselbe Rauschen am anderen Ohr gleichzeitig angeboten wird, wird die wahrgenommene Lautheit des Rauschens zwar höher, der Ton wird jedoch wieder hörbar, so daß eine Verringerung der Mithörschwelle gemessen werden kann, die als binaurale **Masking Level Difference** (MLD) bezeichnet wird. Ihre Größe hängt von der interauralen Phasenlage des Testtons und der zwischen den beiden Ohren bestehenden Beziehung zwischen dem Rauschen ab. Beispielsweise wird die MLD maximal (ca. 14-18 dB), wenn dasselbe Rauschen und ein um 180 Grad (Phase  $\pi$ ) verschobener Sinuston bei etwa 500 Hz verwendet wird. Keine MLD tritt dagegen auf, wenn für das Rauschen und das Signal dieselbe Beziehung an beiden Ohren vorliegt (z. B. bei identischem Rauschen und identischem Ton an beiden Ohren oder bei Phaseninvertierung zwischen beiden Ohren).

In einer guten Näherung kann diese Störgeräuschunterdrückung des binauralen Systems durch eine (mit Restfehlern behaftete) **Subtraktion** der beiden Ohrsignale aufgefaßt werden, die im Gehirn nach einer Anpassung der vom rechten und linken Ohr stammenden Signale (z. B. durch entsprechende Zeitverzögerung und Amplituden-Anpassung) vorgenommen wird. Mit dieser als "Equalization and Cancellation Theory" (EC-Theorie) nach Durlach (1972) bezeichneten Modellvorstellung kann bereits eine

Vielzahl der binauralen Maskierungsexperimente quantitativ gedeutet werden. Ein eher auf physiologischen Vorstellungen basierendes Modell der binauralen Signalverarbeitung wurde dagegen von Jeffress (1948) aufgestellt. Es sieht eine Art neuronales Kreuzkorrelations-Netzwerk vor, bei dem die an den beiden Ohren ankommenden Signale in einer Laufzeitkette sukzessiv zeitverzögert werden und durch Koinzidenz-Detektoren die Laufzeitunterschiede zwischen den beiden Ohren jeweils abgegriffen werden. Auf der Basis dieses Modells wurden in der neueren Literatur eine Reihe weiterer Modelle entwickelt, die die binaurale Signalverarbeitung einschließlich einer Reihe von Effekten richtig vorhersagen (z. B. die Lateralisation bzw. Lokalisation von Schallereignissen mit bestimmter interauraler Zeit- und Pegeldifferenz, das Gesetz der ersten Wellenfront, die MLD für verschiedene Signale und Maskierer, sowie zeitliche und spektrale Eigenschaften dieser Effekte). Da auf diesem Gebiet noch weiter intensiv geforscht wird, sei auf die weiterführende Literatur von Blauert (1983) und Colburn (1996) verwiesen.

## 2. Psychoakustik des pathologischen Gehörs

## -Intensitätsabbildung

Bei sämtlichen Formen der Schwerhörigkeit tritt ein Hörverlust auf, d. h. eine Verringerung der Sensitivität des Hörorgans für Schall. Der Pegel eines akustischen Signals muß daher gegenüber Normalhörenden angehoben werden, um eine d. h. Empfindung hervorzurufen, die Hörschwelle angehoben. ist Die Frequenzabhängigkeit der Hörschwelle wird im Tonschwellenaudiogramm erfaßt. Bei weiterer Erhöhung des Pegels ergibt sich ein Anstieg der Lautheit, der bei Innenohrschwerhörigen in der Regel steiler ist als bei Normalhörenden (Recruitment). Dieser Effekt führt zu dem bei vielen Schwerhörenden beobachteten Phänomen, daß sie bei niedrigen Sprachlautstärken nichts verstehen. Wenn der Gesprächspartner aufgefordert wird, doch etwas lauter zu reden, beklagen sich die Schwerhörigen darüber, daß es schnell zu laut ist und die Sprecher nicht gleich schreien sollen. Dieses Phänomen kann mit überschwelligen Tests erfaßt werden. In der Routine-Audiometrie wird bei seitendifferentem Hörverlust dazu der Fowler-Test verwendet.

Den Lautheitsanstieg mit zunehmenden Pegel kann man für jedes Ohr separat mit der Hörflächenskalierung ermitteln, bei der sich bei Patienten mit Recruitment ein steilerer Anstieg der Kategoriallautheit ergibt. Die Steigung dieser Lautheitsfunktion ist jedoch individuell sehr unterschiedlich und kann möglicherweise von dem Verhältnis der Schädigung innerer und äußerer Haarzellen abhängen (vgl. 3.2). Daher ist der Kompressionsverlust des auditorischen Systems, der sich in der Steigung der Hörflächenskalierung ausdrückt, der als eine weitere, vom Hörverlust weitgehend unabhängige Komponente zu erfassen. Der Kompressionsverlust nimmt mit zunehmenden Hörverlust zwar tendenziell zu, läßt sich aufgrund der großen individuellen Schwankungen aber nicht daraus vorhersagen (vgl. Kießling, 1995, Launer et al, 1996). Bei Schalleitungs-Schwerhörigkeit ist die Pegel-Lautheitsfunktion der Hörflächenskalierung höheren verschoben. zu Pegeln hin Aufgrund des Adaptationseffektes ist diese Verschiebung jedoch nicht so groß wie der an der Schwelle meßbare Hörverlust. Außerdem kann auch die Steigung leicht variieren, so

daß die Hörflächenskalierung keine sichere Unterscheidungsmöglichkeit zwischen Schalleitungs- und Schallempfindungs-Schwerhörigkeit bietet.



Abbildung 13: Intensitäts-JND für verschiedene Schwerhörigkeitsformen

Für die Schalleitungs-Schwerhörigkeit verschiebt sich die Pegelabhängigkeit der Intensitätsauflösung (d. h. der kleinste hörbare Intensitätsunterschied, Intensitäts-Just Noticeable Difference, JND) für Sinustöne zu höheren Pegeln (Abb 13). Bei Innenohrschwerhörigen mit Recruitment geschieht die Abnahme der Intensitäts-JND dagegen in einem wesentlich kleineren Pegelbereich. Sie haben bei derselben wahrgenommenen Lautheit ungefähr die gleiche bzw. eine geringgradig erhöhte Intensitäts-JND wie Normalhörende. Die früher verbreitete Annahme, daß Innenohr-Schwerhörende eine kleinere Intensitäts-JND besitzen, weil sie einen steileren Lautheitsanstieg haben, hat sich als falsch herausgestellt (Hohmann, 1993). Wenn man die Intensitäts-Diskrimination allerdings knapp oberhalb der Hörschwelle mißt (z. B. 20 dB SISI-Test in der überschwelligen Audiometrie), wie beim können Innenohrschwerhörige etwas niedrigere Diskriminations-Schwellen erreichen, weil sie dort eine größere Lautheit wahrnehmen als Normalhörende. Allerdings ist diese Diskriminationsschwelle sehr variabel, so daß der SISI-Test ein sehr indirektes Maß für ein Recruitment darstellt, was sich als sehr unzuverlässig herausgestellt hat. Bei breitbandigen Signalen (z. B. Rauschen oder Sprache) hängt die Intensitäts-Diskrimination sehr wenig vom Pegel ab. Innenohrschwerhörige zeigen daher in ihrem Dynamikbereich eine leicht erhöhte Intensitäts-JND gegenüber Normalhörenden, wobei jedoch die gleiche (geringere) Pegelabhängigkeit auftritt. Bei hochgradig Schwerhörigen mit einer sensorineuralen Schwerhörigkeit bzw. bei Patienten mit einer neuralen Schwerhörigkeit kann die Intensitäts-JND relativ stark erhöht sein, so daß es diesen Patienten schwerfällt, akustische Signale (wie z. B. Sprache) in ihrem auditorischen System richtig abzubilden und zu verstehen.

# -Frequenz- und Zeitabbildung

Bei Schalleitungsschwerhörigkeit treten keine nennenswerten Abweichungen in den überschwelligen Funktionen gegenüber Normalhörenden auf, sofern der Darbietungspegel auf die individuelle Hörschwelle bezogen wird. Im folgenden sollen daher nur die bei sensorineuralen Schwerhörigkeiten auftretenden Veränderungen der Frequenz- und Zeitabbildung betrachtet werden:

Aufgrund der bei Schädigung der äußeren Haarzellen verringerten Frequenzselektivität des Gehörs nimmt auch die psychoakustisch meßbare Frequenzauflösung des Gehörs ab. Dies macht sich in einer Verbreiterung der Frequenzgruppe bemerkbar, d. h. bei Innenohrschwerhörenden wird ein größerer spektraler Bereich Gehör im zusammengefaßt, um die Lautheit und die Maskierung zu bestimmen (s. Abb. 14). Die Verbreiterung der Frequenzgruppe führt zu einem Anstieg der Mithörschwelle im Rauschen, d. h. ein in einem Rauschen versteckter Testton muß stärker im Pegel erhöht werden als bei Normalhörenden, um detektiert werden zu können. Dies liegt daran, daß mehr Energie von benachbarten Frequenzen zur Maskierung des Ziel-Testtons beiträgt. Allerdings hängt diese Verbreiterung vom Signalpegel ab: Auch bei Normalhörenden tritt eine Verbreiterung der Frequenzgruppe mit hohem Pegel auf. Daher sollte man die Frequenzgruppe von Schwerhörenden nicht bei gleichem Pegel über der individuellen Hörschwelle, sondern bei gleichem absoluten Pegel vergleichen. Dabei ist die Schwerhörenden nur geringgradig Frequenzgruppe von gegenüber der von Normalhörenden verbreitert. Bei einer entsprechenden Anhebung des Pegels über die Hörschwelle des Innenohrschwerhörigen (z. B. durch ein Hörgerät) wird daher der Effekt der verbreiterten Frequenzgruppe geringer.



Abbildung 14: Frequenzgruppenbreite als Funktion des Pegels für Normal- und Innenohrschwerhörende

Die Zeitauflösung bei Innenohrschwerhörigen ist ebenfalls verringert, d. h. man mißt eine verlängerte Vor- und Nachverdeckung und eine Verlängerung der kleinsten detektierbaren Pausendauer im Rauschen. Allerdings ist hier genau wie bei der Frequenzauflösung die Pegelabhängigkeit zu berücksichtigen, da bei den Innenohrschwerhörenden ein wesentlich kleinerer Pegelbereich zur Verfügung steht als bei Normalhörenden. Wenn man die Experimente bei dem gleichen Pegel oberhalb der Ruhehörschwelle (bzw. bei der gleichen wahrgenommenen Lautheit) durchführt und auch den Effekt des Recruitments berücksichtigt, ist die Zeitauflösung nicht wesentlich gegenüber der von Normalhörenden gestört. Dies ist auch der Grund, wieso die Modulations-Transferfunktion psychoakustisch meßbare ungefähr den aleichen Tiefpaßverlauf wie bei Normalhörenden aufweist. Allerdings ist die Schwelle insgesamt

erhöht, so daß Schwerhörende nicht so geringe Modulationstiefen detektieren können wie Normalhörende.

Im Gegensatz zu den relativ wenig geänderten Zeitkonstanten bei der Zeitauflösung steht das aeänderte Verhalten bei der zeitlichen Integration bei Innenohrschwerhörigen (s.o.): Während bei Normalhörenden die Wahrnehmbarkeit und die Lautstärke eines Tones mit zunehmender Tondauer bis zu einer Maximaldauer von etwa 200 ms zunimmt, tritt dieser Effekt bei Innenohrschwerhörigen nicht in gleicher Weise auf. Dies ist ein Grund dafür, daß Signale, die im Rauschen versteckt sind (maskierte Signale), von Innenohrschwerhörenden erst bei wesentlich höheren Signal-Rauschverhältnissen detektiert werden können als von Normalhörenden.

## -Weitere auditorische Funktionen

Generell zeigt sich bei Innenohrschwerhörenden eine allgemeine Verschlechterung sämtlicher psychoakustisch meßbarer auditorischer Funktionen im Vergleich zu Normalhörenden. Zu diesen eingeschränkten Funktionen gehören neben der oben erwähnten Intensitäts-, Frequenz- und Zeitauflösung auch andersartig gemessene Detektions- und Diskriminationsleistungen im Zeit- und Frequenzbereich (z. B. Tonhöhenunterscheidung), komplexere Erkennungs- und Unterscheidungsleistungen, sowie die binaurale Interaktion (d. h. die Extraktion von Unterschieden zwischen den Signalen, die an den beiden Ohren anliegen), die zur Ortung von Schallquellen und zur Trennung zwischen Nutzsignalen und Störsignalen beiträgt. Dabei stellt man fest, daß sämtliche dieser auditorischen Funktionen primär vom Hörverlust beeinflußt werden. Im Mittel über alle Patienten verschlechtern sich alle Funktionen mit zunehmendem Hörverlust. Abgesehen von diesem generellen Effekt gibt es jedoch nur sehr geringe gegenseitige Abhängigkeiten zwischen den Schädigungen in der Hörfunktion. Beispielsweise kann die Frequenzauflösung stark gestört sein und das binaurale Hören relativ intakt (oder umgekehrt). Dies führt zu der Vorstellung, daß sich Innenohrdurch die zwei Hörverlust Hörstörungen primär Komponenten und Kompressionsverlust charakterisieren lassen. Die darüber hinausgehenden Einschränkungen unterliegen diffusen. unterschiedlichen einem an Orten in verschiedener, nicht vorhersehbarer Weise wirkenden Schädigungsprozeß.

## 3. Sprachwahrnehmung bei Normal- und Schwerhörigen

Die sprachliche Kommunikation zwischen Menschen ist eine hochspezialisierte und faszinierende Leistung, die letztendlich die Grundlage unserer Kultur darstellt. Um die beim Verstehen von Sprache (und ihrer möglichen Störungen) beteiligten Prozesse verstehen zu können, muß der gesamte Übertragungsweg vom Sender (sprechender Mensch) über die akustische Darstellung von Sprache bis zum Empfänger (hörender Mensch) genau betrachtet werden.

### 3.1 Sprachakustik

Wenn das akustische Sprachsignal z. B. mit einem Mikrophon aufgenommen wird, läßt sich aus dem Zeitverlauf des Signals nur sehr wenig Information über das geäußerte Sprachelement ableiten. Beispielsweise kann gesehen werden, ob es sich um einen Vokal (periodisches Signal aufgrund der Stimmlippen-Schwingung) oder einen stimmlosen Konsonanten mit Rauschanteilen handelt (irreguläres Zeitsignal). Weiterhin lassen sich mit hoher Zeitauflösung zeitliche Übergänge feststellen (vgl. Abbildung 15). Eine Möglichkeit, mehr Informationen über die akustische Filterung des Sprachsignals im Vokaltrakt zu erhalten, bietet die Analyse des Frequenzgehalts des Schallsignal (Leistungsspektrum). Dabei wird für jede Frequenz über die gesamte Dauer des Signals gemittelt die Energie errechnet, die in das jeweilige Frequenzband fällt (vgl. Abbildung 15 rechts). Bei einem stationären Signal, das sich nicht zeitlich ändert (z. B. bei gehaltenen Vokalen) kann man in diesem (Leistungs-)Spektrum sehr gut die Grundfrequenz (Schwingungsfrequenz der Glottis) mit ihren Obertönen (ganzzahlige Vielfache der Grundfrequenz) erkennen, sowie die Formanten, d. h. die Frequenzbereiche, bei denen der Vokaltrakt eine besonders hohe Verstärkung des akustischen Signals bewirkt. Die Formanten stellen sich damit als Spitzen im Spektrum dar und sind charakteristisch für den jeweils artikulierten Vokal (s. unten).



Abbildung 15: Zeit-Signal (unten), Spektrum (rechts) und Spektrogramm (links oben) des Wortes "Hörakustik"

Da das Spektrum über eine lange Zeit gebildet wird, ist es ungeeignet zur Beurteilung der sich zeitlich stark ändernden akustischen Anteile von fließender Sprache. Es hat allerdings eine große Bedeutung als **mittleres Sprachspektrum**, d. h. die Energieverteilung im Frequenzbereich, die im Mittel bei einem männlichen bzw. weiblichen Sprecher auftritt (vgl. Abbildung 16). Daraus läßt sich ablesen, daß im

Bereich der Grundfrequenzen zwischen etwa 300 und 800 Hz am meisten Sprachenergie vorliegt (Maximum des Spektrums), während zu niedrigen Frequenzen ein Abfall von etwa 20 dB pro Oktave erfolgt und zu hohen Frequenzen von etwa 10 dB pro Oktave. Die Langzeitspektren von männlichen und weiblichen bzw. kindlichen Sprechern unterscheiden sich durch die Lage des Maximums im Frequenzbereich, das bei Frauen und Kindern deutlich höher liegt als bei Männern. Mit zunehmender **Sprechanstrengung** verschiebt sich das spektrale Maximum zu hohen Frequenzen, so daß sich die Stimme aus dem Hintergrund (mit vorwiegend tieffrequenter Leistung) deutlicher abgrenzen läßt, ohne daß die akustische Gesamtleistung der Stimme in gleichem Maß ansteigt.



Abbildung 16: Mittleres Spektrum von männlichen bzw. weiblichen Sprechern

Aufgrund des Spektrums von gehaltenen Vokalen bzw. ausgeschnittenen Vokalen läßt sich ein Zusammenhang zwischen der Lage des ersten Formantens (d. h. spektrales Maximum im Bereich zwischen etwa 200 und 800 Hz), dem zweiten Formanten (zweites spektrales Maximum zwischen etwa 600 und 2.500 Hz), der Stellung der Artikulationsorgane und dem jeweils geäußerten Vokal aufstellen (**Vokaldreieck** bzw. Vokaltrapez, vgl. Abbildung 17).



Abbildung 17: Vokaldreieck. Aufgetragen ist die Frequenz des ersten Formanten F<sub>1</sub> (Abzisse) und des zweiten Formanten F<sub>2</sub> (Ordinate) für unterschiedliche Vokale (schematisch)

# 3.2 Sprachverständlichkeit

Beim Verstehen von Sprache spielt einerseits die **Sprachverständlichkeit** eine Rolle, d. h. die Eigenschaft des Sprachmaterials, von einem durchschnittlichen, normalhörenden Probanden verstanden zu werden. Andererseits spielt die individuelle **Sprachperzeptionsleistung** eine Rolle, d. h. die von dem individuellen Patienten erbrachten Voraussetzungen zum Sprachverstehen, die mehr oder weniger stark gestört sein können. In der **Sprachaudiometrie** mißt man die Verständlichkeit von standardisiertem Sprachmaterial bei dem individuellen Patienten, um Rückschlüsse auf dessen Sprachperzeptionsleistungen führen zu können.

# 3.2.1 Methoden zur Bestimmung der Sprachverständlichkeit

Das für Verständlichkeitsmessungen verwendete Sprachmaterial sollte möglichst repräsentativ sein für die Sprache und für die zu betrachtende Kommunikationssituation. Dabei tritt als grundlegendes Problem die hohe Zahl von Variablen auf, für die eine möglichst vernünftige Festlegung erfolgen sollte. Zu diesen Variablen gehört die Art des Tests: **"offene" Tests**, bei dem ein Test-Item (z. B. ein Wort oder Satz) dem Probanden dargeboten wird, das er möglichst korrekt wiederholen soll, und **"geschlossene" Tests**, bei dem der Proband das richtige Test-Item aus einer Liste von möglichen Antworten bezeichnet. Weitere Variablen sind die Länge der Test-Items (z. B. einsilbige oder mehrsilbige Wörter oder Sätze), die Auswerte-Methode (Bewertung richtig erkannter Phoneme, Silben, Wörter oder Sätze), der Sprecher (männlicher oder weiblicher Sprecher, geschulter oder ungeschulter Sprecher oder etwa synthetische Sprache), sowie die Wahl eines Störgeräusches und einer räumlichen Anordnung von Nutzschall-und Störgeräuschquelle. Jedes der gebräuchlichen Sprachtestverfahren besitzt eine eigene Kombination dieser Variablen, so daß die Test-Ergebnisse sich nicht leicht miteinander vergleichen lassen.

Außerdem ist die Festlegung der Variablen abhängig vom Einsatzzweck des Tests. Beispielsweise steht bei der **Diagnostik** von Hörstörungen die analytische Fähigkeit des Tests im Vordergrund (d. h. die Möglichkeit, aus den auftretenden Phonem-Verwechslungen Rückschlüsse auf das gestörte Hörsystem zu ziehen), während bei der **Begutachtung** eher eine hohe Reproduzierbarkeit des Tests und eine hohe Repräsentanz des Sprachmaterials für alltägliche Kommunikation im Vordergrund steht. Bei der **Hörgeräteanpassung** ist dagegen eine hohe Sensitivität gegenüber kleinen Änderungen der Einstellparameter des Hörgerätes wichtig.

Der derzeit in der Standard-Audiometrie am häufigsten eingesetzte Sprachtest, der **Freiburger Wörtertest**, ist ein "offener" Test mit Einsilbern bzw. mit Mehrsilbern (Zahlen). Als Alternative dazu wurde in jüngerer Zeit das **Reimtestverfahren** mit

geschlossenen Antwortalternativen eingeführt, die sich in nur einem Phonem unterscheiden und sich daher reimen (z. B. die Alternativen "Sinn", "Hin", "bin", "Kinn", "Zinn"). Dieses Verfahren vermeidet Fehlerquellen bei der Testbewertung durch den Testleiter und ist automatisierbar. Die Einsilber-Tests haben den Vorteil einer geringen Redundanz (d. h. aus einem richtig erkannten Teil des Wortes kann nicht auf den anderen, unverständlichen Teil des Wortes geschlossen werden) und bieten eine hohe analytische Aussagekraft bei den auftretenden Phonem-Verwechslungen. Da im Deutschen Zweisilber jedoch häufiger als Einsilber sind, sind Zweisilber-Tests eher repräsentativ für die deutsche Sprache, so daß auch ein Reimtest-Verfahren für Zweisilber entwickelt wurde. Sprachtests mit Sätzen bieten dagegen eine sehr Kommunikations-Situation. Diskriminationsfunktion realitätsnahe Ihre (d. h. der Prozentsatz richtig erkannter Wörter als Funktion des Sprachpegels) ist sehr steil, so daß sie eine hohe Sensitivität gegenüber Änderungen im Sprachpegel oder im Signal-Rauschabstand aufweisen. Für die deutsche Sprache gibt es den standardisierten Marburger Satztest und den neueren Göttinger Satztest. Ein Überblick über moderne Verfahren der deutschsprachigen Sprachaudiometrie findet sich bei Kollmeier (1992).

## 3.2.2 Methoden zur Berechnung und Modellierung der Sprachverständlichkeit

Um das Verstehen von Sprache guantitativ zu erfassen und die Sprachverständlichkeit für eine vorgegebene akustische Situation oder einen bestimmten Hörverlust vorhersagen zu können, wurden verschiedene Methoden zur Berechnung bzw. Vorhersage der Sprachverständlichkeit aus dem zugrundliegenden Sprachmaterial und (z. B. Nachhall, den akustischen Gegebenheiten Störgeräusch, angeborene Ruhehörschwelle) klassische entwickelt. Die Methode der Sprachverständlichkeitsvorhersage ist der Artikulations-Index (AI) und der Speech Transmison Index (STI). Sie beruhen auf der Annahme, daß die gesamte Sprach-Information auf die verschiedenen Frequenzbänder des akustischen Signals verteilt ist, und daß jedes Band einen gewissen Beitrag zur Gesamt-Sprachverständlichkeit liefert. Die Breite jedes dieser Bänder orientiert sich dabei an der Frequenz-Gruppenbreite (Bark-Skala oder in erster Näherung Terz-Bandbreite). In jedem dieser Frequenzbänder kann nun ein "effektiver" Signal-Rauschabstand ermittelt werden, d. h. das Verhältnis zwischen der Nutzenergie des zu verstehenden Sprachsignals und den Störanteilen. Diese werden durch Nachhall oder durch ein Störrauschen verursacht, oder dadurch, daß die Energie in diesem Band unterhalb der Ruhehörschwelle des jeweiligen Patienten liegt, was durch ein angenommenes "internes Rauschen" nachgebildet wird. Wenn der Signal-Rauschabstand in dem Band größer als 15 dB ist, geht man davon aus, daß dieses Band vollständig zur Verständlichkeit beiträgt, während bei einem Signal-Rauschabstand von kleiner als -15 dB sämtliche Sprachinformation in diesem Band maskiert ist, so daß das Band nicht zur Gesamt-Sprachverständlichkeit beiträgt. Der Signal-Rauschabstand im Band j (SNR (j)) wird also auf einen Bereich von -15dB bis +15dB begrenzt. Zur Ermittlung des Artikulationsindex wird nun eine gewichtete Mittelung über die Signal-Rauschabstände in den einzelnen Bändern durchgeführt, bei der jedes Band i mit einem Gewichtungsfaktor Wi multipliziert wird:

AI oder STI=
$$\sum_{j} W_{j} \cdot [SNR(j) + 15dB]$$
 (1)

Die Gewichtungsfaktoren Wi sind dabei so gewählt, daß der STI oder AI nur Werte zwischen 0 (entsprechend einer Sprachverständlichkeit von 0, wenn in jedem Band nur Rauschen und keine signifikanten Sprachanteile vorliegen) und 1 variiert (vollständige Sprachverständlichkeit, da in jedem Frequenzband die Sprache von dem Rauschen kaum beeinflußt wird). Damit stellt der AI bzw. STI ein Maß für die Sprachverständlichkeit das direkt mit der meßbaren mittleren dar. Sprachverständlichkeit zusammenhängt. Diese Beziehung hängt allerdings von dem verwendeten Sprachmaterial und Test ab (vgl. Abbildung 18). Bei Sprachtests mit geringer Redundanz (z. B. bei sinnlosen oder sinnbehafteten Einsilbern) wächst die Sprachverständlichkeit mit zunehmendem AI bzw. STI nur langsam an, während die Sprachverständlichkeit bei Sätzen mit hoher Redundanz schon bei relativ kleinem Al sehr hoch ist. Dies liegt an der Möglichkeit, den Satz aufgrund des Satzzusammenhanges schon dann vollständig richtig zu verstehen, wenn nur einige Teile des Satzes verstanden werden.



Abbildung 18: Sprachverständlichkeit für verschiedenes Sprachmaterial als Funktion des AI bzw. STI

Mit Hilfe der in Abbildung 18 angegebenen Kurven ist es daher prinzipiell möglich, die Verständlichkeit für verschiedene Sprachmaterialien bei vorgegebenen akustischen Bedingungen ineinander umzurechnen. Es muß jedoch betont werden, daß der Artikulations Index und der Speech Transmission Index nur für den Mittelwert von Sprachverständlichkeiten über ein normalhörendes Versuchspersonenkollektiv und eine große Anzahl von Tests gelten. Für den individuellen Patienten und für die Verständlichkeit eines einzelnen Wortes oder eines einzelnen Satzes treten große Schwankungen in der Verständlichkeit auf, so daß diese Berechnungsmethode für diesen Fall nicht angewendet werden kann (vgl. Kollmeier, 1990).

Der STI unterscheidet sich vom AI durch die Vorgehensweise zur Berechnung des "effektiven" Signal-Rauschabstandes SNR (j) in jedem Band j. Während beim AI das Leistungsspektrum des Sprachmaterials und das Leistungsspektrum des Störsignals getrennt voneinander bekannt sein müssen und ins Verhältnis gesetzt werden, wird beim STI der Signal-Rauschabstand anhand des Sprachsignals und der Mischung aus Sprachsignal und Störsignal mit Hilfe der **Modulations Transfer Funktion** (MTF) bestimmt. Dazu werden in jedem Frequenzbereich die im Original-Signal auftretenden Modulationen gemessen und die in der Mischung verbleibenden Modulationen werden ins Verhältnis zu den ursprünglichen Modulationen gesetzt. Wenn die ursprünglichen Modulationen vollständig erhalten sind (Modulations-Transfer-Funktion von 1), spricht dies für ein sehr hohes Signal-Rauschverhältnis, während die Anwesenheit von Störgeräusch oder von Nachhall die im gemischten Signal verbleibenden Modulationen im gemischten Signal für die Sprachverständlichkeit von entscheidender Bedeutung ist und errechnet den Signal-Rauschabstand aus diesem Erhalt der Modulationen. Dadurch können der Einfluß von Nachhall und der Einfluß von Störrauschen in gleicher Weise behandelt werden, so daß sich der STI besonders für den Einsatz in der Raumakustik bewährt hat. Die zeitlichen Eigenschaften eines Sprachübertragungssystems (z. B. des gestörten Gehörs) werden ebenfalls besser berücksichtigt als beim AI (vgl. Houtgast und Steneken, 1973).



Abbildung 19: Perzeptionsmodell zum Sprachverstehen nach Holube und Kollmeier (1996)

Neben den auf Eigenschaften Sprachmaterials beruhenden den des STI Berechnungsverfahren des ΑI und aibt es Modelle zur Sprachverständlichkeitsvorhersage, die sich eher an den Eigenschaften des Gehörs orientieren und die Verarbeitungsschritte bei der Sprachwahrnehmung explizit nachbilden. Als Beispiel für ein derartiges Perzeptionsmodell soll das in Abbildung 19 skizzierte Modell nach Holube und Kollmeier (1996) erläutert werden: Das gestörte Sprachsignal wird zunächst einer peripheren Vorverarbeitungs-Stufe zugeführt, die die "effektive" Signalverarbeitung im auditorischen System möglichst gut nachbildet. Die Elemente dieser Vorverarbeitungsstufe entsprechen dabei genau den anhand von psychoakustischen physiologischen Experimenten und gewonnenen Verarbeitungseinheiten. Zunächst wird eine Aufspaltung des Signals in verschiedene Frequenzbereiche vorgenommen (entspricht der Frequenz-Orts-Transformation der Basilarmembran), anschließend wird eine Extraktion der Einhüllenden und eine Adaptation mit zeitlicher Kontrastierung in jedem Frequenzkanal vorgenommen (dies entspricht ungefähr der Aktion der Haarzellen und des Hörnervs). Die statistischen Ubertragungsfehler durch das Nervensystem und die Begrenzung der Empfindlichkeit durch die Ruhehörschwelle wird durch ein zusätzliches "internes" Rauschen nachgebildet. Am Ausgang der peripheren Vorverarbeitungsstufe liegt dann zu jedem Zeitpunkt für jede Frequenz ein komprimierter und zeitlich kontrastierter Intensitätswert vor, der ungefähr der **internen Repräsentation** der akustischen Signale entspricht und eine ähnliche Darstellung der Sprachinformation wie das Spektrogramm ermöglicht.

Dieselben Vorverarbeitungsschritte werden auch für ein Referenzsignal (ungestörtes Sprachsignal) durchgeführt. In einer zentralen Verarbeitungsstufe (Mustererkenner) kann dann überprüft werden, inwiefern die interne Repräsentation des Eingangssignals mit derjenigen eines vorgegebenen Vergleichsignals übereinstimmt. Diese zentrale Mustererkennung entspricht ungefähr dem Prozeß der Sprachwahrnehmung, bei dem die aktuelle akustische Wahrnehmung mit der möglichen Wahrnehmung von allen möglichen bekannten Wörtern verglichen wird. Als "erkanntes Wort" wird dasjenige ausgewählt, dessen interne Repräsentation am ähnlichsten mit der Repräsentation des Eingangssignals ist. Dieser zentrale Verarbeitungs-Vorgang kann daher mit einem Spracherkennungsalgorithmus durchgeführt werden und bildet das "Weltwissen" des Hörers nach.

Mit einer derartigen Modellstruktur kann untersucht werden, inwiefern sich Verarbeitungsfehler in der peripheren Vorverarbeitung (z. B. Änderung der spektralen und zeitlichen Auflösung oder eine Anhebung der Ruhehörschwelle) auf die veränderte Erkennungsleistung beim Sprachverstehen auswirkt. Der Vorteil eines derartigen Modells besteht darin, daß die Sprachverständlichkeitsvorhersage nicht global, im Mittel über viele Versuchspersonen und viele Wörter erfolgen muß, sondern daß für jede iedes zu erkennende individuelle Versuchsperson und Wort eine eiaene Verständlichkeitsberechnung durchgeführt werden kann. Die Vorhersagen mit diesem Modell stimmen für die Sprachverständlichkeit in Ruhe und insbesondere für verschiedene Störgeräusche sehr gut mit den gemessenen Verständlichkeitswerten bei Normal- und Schwerhörenden überein (vgl. Holube und Kollmeier, 1996). Aufgrund des hohen Rechenzeitaufwandes für dieses Modell und die noch nicht in allen Einzelheiten geklärten Verarbeitungsprozesse sind jedoch noch weitere Forschungsarbeiten zu seiner Validierung notwendig.

# 3.3 Sprachperzeption bei pathologischem Gehör

Bei Schwerhörigkeit ist das Verstehen von Sprache in Ruhe und unter Störgeräuscheinfluß reduziert. In der **Sprachaudiometrie** wird diese Reduktion quantitativ erfaßt, in dem die Sprachverständlichkeit (d. h. der Prozentsatz korrekt verstandener Test-Items einer vorgegebenen Testliste von Wörtern, Zahlen oder Sätzen) für verschiedene Sprachpegel und für verschiedene Testbedingungen bestimmt wird (z. B. in Ruhe oder unter Störgeräusch mit unterschiedlicher räumlicher Anordnung von Störschall und Nutzschall). Das Ziel dieser Messungen ist entweder die differenzierte **Diagnostik** von Hörstörungen, die **Begutachtung** eines Hörschadens oder die **Anpassung von Hörhilfen** (d. h. das Ausmessen der Sprachverständlichkeit ohne Hörhilfe und den Gewinn an Sprachverständlichkeit, der durch die Hörhilfe erzielt wird). Für die Bestimmung der Sprachverständlichkeit gibt es mehrere Testverfahren (vgl. 1.2) von denen der Freiburger Einsilber- bzw. Zahlentest und der Marburger Satztest standardisiert sind. Neuere und weniger fehleranfällige Verfahren wie der Einsilber- oder Zweisilber-Reimtest und der Göttinger Satztest befinden sich im Erprobungsstadium.



Abbildung 20: Diskriminationsfunktionen bei unterschiedlichen Pathologien (schematisch)

Für den Sprachtest in Ruhe findet man bei den verschiedenen Schwerhörigkeitsformen prinzipiell die in Abbildung 20 dargestellten Diskriminationsfunktionen (d. h. Prozent Sprachverständlichkeit als Funktion des Sprachpegels). Bei Normalhörenden hängt die Steigung der Diskriminationsfunktion vom verwendeten Sprachmaterial ab: Die Diskriminationsfunktion ist relativ flach bei Einsilbern mit offenem Antwortformat, etwas steiler beim Reimtestverfahren und extrem steil (ca. 20 % Verständlichkeit pro dB) bei Sätzen wie z. B. beim Göttinger Satztest. Bei Schalleitungs-Schwerhörigkeiten ist die Diskriminationsfunktion zu hohen Pegeln hin verschoben. Bei Schallempfindungs-Schwerhörigkeiten ist die Diskriminationsfunktion ebenfalls verschoben, was auf die "Abschwächungswirkung" des Hörverlusts zurückgeführt wird. Die darüber hinausgehende "Verzerrungswirkung" des Hörverlusts ("Fehlhörigkeit") macht sich dagegen in einer Abflachung (Verringerung der Steigung) der Diskriminationsfunktion und in einem Diskriminationsverlust bemerkbar, d. h. die Diskriminationskurve erreicht bei hohen Pegeln nicht 100 % sondern bleibt um einen bestimmten Wert, der als Diskriminationsverlust bezeichnet wird, unterhalb von 100 %. Bei sehr hohen Pegeln kann bei einem Patienten sogar eine Verschlechterung der Sprachverständlichkeit mit zunehmendem Pegel beobachtet werden ("roll-over"-Phänomen oder R-Kurve), das im Zeitalter verbesserter audiologischer Technik (verzerrungsfreie Verstärker und Kopfhörer, CD-Wiedergabe des Sprachmaterials) allerdings selten geworden ist. Die beiden erstgenannten Effekte lassen sich bei Innenohrschwerhörigen darauf zurückführen, daß sie bestimmte Wörter erst bei sehr hohen Pegeln oder überhaupt nicht verstehen können, während andere Wörter schon bei relativ niedrigen Pegeln verständlich sind. Der Übergangsbereich der Diskriminationsfunktion zwischen "unverständlich" und "vollständig verständlich" überdeckt daher einen größeren Pegelbereich mehr als bei Normalhörenden bzw. Schalleitungsschwerhörigen.

Ziel der Sprachaudiometrie in Ruhe ist es nun, die wesentlichen Eigenschaften der Diskriminationsfunktion des individuellen Patienten zu erfassen. d. h. die Verständlichkeitsschwelle (der zu 50 % Sprachverständlichkeit gehörende die Steigung der Diskriminationsfunktion und den Sprachpegel), Diskriminationsverlust bzw. das Vorliegen einer R-Kurve. In der Routine-Audiometrie wird für die einsilbigen Wörter des Freiburger Sprachtest daher die Sprachverständlichkeit bei 65 dB bestimmt und bei um jeweils 15 dB erhöhtem Sprachpegel bis entweder die Erkennungsrate 100 % erreicht wird oder die Unbehaglichkeits-Schwelle überschritten wird. Daraus ergibt sich die Diskriminationsfähigkeit bei 65 dB (in %), der Pegel bei dem die höchste Erkennungsrate erzielt wird (dB Opt.), die Differenz zu 100 % bei diesem Pegel (Diskriminationsverlust in Prozent) und die Gesamtwortverständlichkeit (GWV in %), die sich aus der Summe der Erkennungsraten in Prozent bei 60, 80 und 100 dB ergibt. Bei den mehrsilbigen Zahlwörtern des Freiburger Tests wird die Erkennungsrate für einen oder zwei verschiedene Pegel bestimmt. Unter Berücksichtigung der Steigung der Diskriminations-Normkurve für Normalhörende wird anschließend durch Interpolation derjenige Pegel bestimmt, bei dem 50 % der Zahlen verstanden werden. Die Differenz zu dem Wert für Normalhörende wird als Hörverlust (HV) für Zahlwörter in dB sprachaudiometrische Testverfahren die angegeben. Neuere ermitteln Sprachverständlichkeitsschwelle, den Diskriminationsverlust und die Steigung der Diskriminationsfunktion durch adaptive Messungen, bei denen der Sprachpegel in Abhängigkeit von den Antworten des Probanden variiert wird. Dadurch wird mit wenigen Versuchsschritten der Pegel maximalen Sprachverstehens ermittelt und anschließend die Sprachverständlichkeitsschwelle. Durch das Anpassen einer Diskriminationsfunktion an die Antworten des Probanden läßt sich anschließend die Steigung der Diskriminationsfunktion angeben. Diese Verfahren werden vom Computer gesteuert bzw. ausgewertet, so daß sie bei zunehmender Verbreitung der computergesteuerten Audiometrie zunehmende Bedeutung für die Praxis gewinnen werden (vgl. Kollmeier, 1996).

Bei der **Sprachaudiometrie unter Störgeräusch** interessiert vorwiegend die Sprachverständlichkeits-Schwelle, d. h. derjenige Signal-Rauschabstand (Sprachpegel in Relation zum Störgeräuschpegel), bei dem eine Sprachverständlichkeit von 50 % erzielt werden kann. Dieser Wert ist bei sensorineural Schwerhörenden in der Regel deutlich erhöht gegenüber Normalhörenden und Schalleitungs-Schwerhörigen. Dies ist ebenfalls Ausdruck der "Verzerrungswirkung" des Hörverlusts und entspricht den Beschwerden der Schwerhörenden, daß sie unter Umgebungsgeräusch mehr Schwierigkeiten haben, einer Unterhaltung zu folgen, als in Ruhe ("Cocktail-Party-Effekt" oder **Gesellschaftsschwerhörigkeit**). Da bei zusammenhängenden Sprachmaterialien wie einer Unterhaltung oder Sätzen eine Zunahme der Sprachverständlichkeit um etwa 20 % pro Verbesserung des Signal-Rauschabstands um 1 dB erfolgt, ist eine hohe Meßgenauigkeit zur Erfassung dieser Komponente des Hörverlusts notwendig. Einen besonders großen Unterschied in der Sprachverständlichkeitsschwelle zwischen Normal- und Innenohrschwerhörigen beobachtet man bei der Verwendung von

fluktuierenden Störgeräuschen, welche zeitliche Einhüllenden-Schwankungen in ähnlicher Weise aufweisen, wie ein einzelner, störender Sprecher (Fastl, 1987). Während Normalhörende in den "Lücken" des Störgeräuschs noch einen Teil der Sprache gut verstehen können und daher durch ein derartig fluktuierendes Störgeräusch relativ wenig gestört sind, ist bei Innenohrschwerhörigen diese Fähigkeit stark gestört.

Ein weiteres Handicap beim Sprachverstehen im Störgeräusch haben Innenohrschwerhörende gegenüber Normalhörenden in **räumlichen Nutzschall-Störschallsituationen**. Während Normalhörende sowohl durch monaurale Verarbeitung als auch durch binaurale Signalverarbeitung (d. h. durch den Vergleich zwischen den an beiden Ohren jeweils anliegenden Signalen) eine gewisse Störgeräuschunterdrückung durchführen können und sich auf den Nutzsprecher konzentrieren können, ist dieser Effekt bei Schwerhörenden in sehr unterschiedlichem Maße gestört.

#### Literatur.

- Blauert, J. (1983): Spatial Hearing: The psychophysics of human sound localization, (Cambridge, Massachussets, MIT Press).
- Böhme, G., Welzl-Müller, K.: Audiometrie Hörprüfungen im Erwachsenen- und Kindesalter. Huber, Berlin 1993.
- Dau, T., Kollmeier, B., Kohlrausch, A. (1996). Modeling modulation perception: modulation low-pass filter or modulation filterbank. In: B. Kollmeier (Ed.), Psychoacoustics, Speech and Hearing Aids. World Scientific, Singapur, pp. 45-48.
- Durlach, N.I. (1972). Binaural signal detection: Equalization and cancellation theory. In: J.V. Tobias (Ed.) Foundations of Modern Auditory Theory, Academic Press, New York, Vol. 2, 363-462.
- Fastl, H. (1987). Ein Störgeräusch für die Sprachaudiometrie. Audiologische Akustik 26, 2-13.
- Festen, J.M. (1996): Temporal resolution and the importance of temporal envelope cues for speech perception. In: B. Kollmeier (Ed.), Psychoacoustics, Speech and Hearing Aids. World Scientific, Singapur, pp.95-102.
- Gelfand, S.A.: Hearing, an Introduction. M. Deccer, 1990.
- Hellbrück, J.: Hören Physiologie, Psychologie und Pathologie. Hogrefe Verlag, Göttingen 1993.
- Heller, O. (1985): Hörfeldaudiometrie mit dem Verfahren der Kategorienunterteilung (KU). Psychologische Beiträge 27, 478-493.
- Helmholtz, H.L.F.: Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik, Vieweg, Braunschweig, 1863.
- Hohmann, V. (1993). Dynamikkompression für Hörgeräte Psychoakustische Grundlagen und Algorithmen. Fortschritt-Berichte VDI, Reihe 17 Nr. 93, Düsseldorf. Zugl.: Dissertation, Universität Göttingen.
- Hohmann, V., Kollmeier, B. (1995). Weiterentwicklung und klinischer Einsatz der Hörfeldskalierung. Audiol. Akustik 34 (2), 48-59.
- Holube, I., Kollmeier, B. (1996). Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model. J. Acoust. Soc. Am. (im Druck)
- Houtgast, T., Steeneken, H.J.M. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. Acustica 28, 66-73.
- Jakobsen, R., Halle, M., Fant, G. (1951). Preliminaries to Speech Analysis: The distinctive Features and their Correlate. MIT Press, Cambridge, MA.
- Jeffress, L.A. (1948): A place theory of sound localization. J. Comp. Physiol. Psych. 61, 468-486.
- Kießling, J. (1995). Zum überschwelligen Lautheitsanstieg bei Schallempfindungsschwerhörigen -Konsequenzen für die Hörgeräte-Entwicklung und - Anpassung. Audiol. Akustik 34 (2), 82-89.
- Kießling, J., Kollmeier, B. Diller, G. Versorgung und Rehabilitation mit Hörgeräten. Thieme-Verlag
- Launer, S., Hohmann, V., Kollmeier, B. (1996). Modeling loudness growth and loudness summation in hearing-impaired listeners. In: W.
- Jestaedt (Ed.), Modeling Sensorineural Hearing Loss. Lawrence Erlbaum & Assoc., Hillsdale, NJ (im Druck).

Kollmeier, B. (1990): Meßmethodik, Modellierung und

Verbesserung der Verständlichkeit von Sprache. Habilitationsschrift, Universität Göttingen.

Kollmeier, B. (Ed.): Moderne Verfahren der Sprachaudiometrie. Median-Verlag, Heidelberg, 1992

- Kollmeier, B. (1996): Computer-controlled speech audiometric techniques for the assessment of hearing loss and the evaluation of hearing aids. In: B. Kollmeier (Ed.), Psychoacoustics, Speech and Hearing Aids. World Scientific, Singapur, pp.57-68.
- Kollmeier, B.(2001)Die Oldenburger Hörforschung. Einblicke Forschungsmagazin der Carl von Ossietzky Universität Oldenburg, 33, p. 3-6.
- Kollmeier, B.: Skript zur Vorlesung "Physiologische, psychologische und audiologische Akustik, Universität Oldenburg http://www.physik.uni-

oldenburg.de/Docs/medi/html/teaching\_download.html.

- Kollmeier, B. (2002) Cocktail-Partys und Hörgeräte: Biophysik des Gehörs. Physik Journal, 4, p. 39-45
- Moore, B.C.J.: An Introduction to the Psychology of Hearing, Academic Press, New York, 1989.
- Moore, B.C.J.: Perceptual Consequences of Cochlear Damage. Oxford University Press, Oxford 1995.
- Moore, B.C.J., Glasberg, B.R. (1987). Formulae describing frequency selectivity in the perception of loudness and level and their use in calculating excitation patterns. Hear. Res. 28, 209-225.
- Pickles, J.A.: Hearing, an Introduction to the Physiology of Hearing, Academic Press 1982.
- Poulton, E.C. (1989). Bias in Quantifying Judgements. Lawrence Erlbaum Assoc., Hillsdale, N.J.
- Rayleigh, Lord. Theory of Sound. London, Macmillan & Co. Ltd., 1877, Nachdruck 1929.
- Stevens, S.S. (1957). On the psychphysical law. Psychol. Rev. 64, 153 181.

Zwicker, E., Fastl, H. Psychoacoustics - Facts and Models. Springer Verlag, Berlin 1990.

# Hördiagnostik für die rehabilitative Audiologie

## Birger Kollmeier / Oldenburg

Der folgende Beitrag soll einen kurzen Einblick in den Entwicklungsstand neuer audiologischer Messverfahren als Ausgangsbasis für eine Hörgeräte-Versorgung aus der Sicht des Autors vermitteln. Dabei wird kein Anspruch auf Vollständigkeit oder Ausgewogenheit des präsentierten Materials erhoben, da dies in Lehrbüchern der Audiologie (z. B. Laszig, 2000, Kießling et al., 2008, Böhme u. Welzl-Müller, 1998) in weitaus besserem Maß bereits existiert.

Einen Überblick über das Spektrum der in der Audiologie eingesetzten Verfahren gibt Abbildung 1. Ebenfalls angemerkt sind Neu- bzw. Weiterentwicklung von Verfahren, die in der Arbeitsgruppe des Autors entwickelt worden sind und zum Teil soweit ausgereift sind, daß sie den Sprung von der Grundlagenforschung in die klinischaudiologische Diagnostik und HNO-ärztliche Praxis bzw. Anwendung in der Hörgeräte-Anpassung in Kürze vollziehen können. Diese Verfahren sollen kurz vorgestellt werden.



Abbildung 1: Funktionelle Gliederung des Gehörsystems mit Funktionen und möglichen Störungen (aus Kollmeier, 1997). Zusätzlich angegeben sind die neueren audiologischen Testverfahren, die in der Arbeitsgruppe des Autors für die praktische Anwendbarkeit entwickelt wurden.

## Oldenburger Hörflächenskalierung

Physiologische Erkenntnisse der letzten 10 Jahre weisen auf einen schnellen, aktiven Verstärkungsmechanismus in der Cochlea hin, der zu einer Dynamik-Kompression der die Basilarmembran anregenden Signale im auditorischen System führt. Ein Ausfall dieser "aktiven Prozesse" führt zu einem Wegfall der Kompressions-Wirkung auf der Ebene der Cochlea ("Kompressionsverlust") und damit zu dem als "Recruitment" bezeichneten Phänomen des pathologischen Lautheitsanstiegs. Dieser Kompressionsverlust ist (vorwiegend) durch die Schädigung äußerer Haarzellen bedingt und ist weitgehend unabhängig vom Sensitivitätsverlust (vorwiegend durch Schädigung der inneren Haarzellen und teilweise durch

Schädigung der äußeren Haarzellen bedingt), da Patienten mit gleichem Audiogramm (d. h. gleichem Sensitivitätsverlust) ein sehr stark unterschiedliches überschwelliges Lautheitsempfinden aufweisen können (Launer et al., 1996, Kießling et al., 1995). Daher ist es sinnvoll und notwendig, den Kompressionsverlust zusätzlich zu dem Sensitivitätsverlust auszumessen. Als einfache und robuste klinische Meßmethode bietet sich die kategoriale Lautheitsskalierung an, die in Form der "Oldenburger Hörflächenskalierung" optimiert wurde (Kollmeier, 1997). Wichtige Merkmale dieses Verfahrens sind:

- 10+1 Kategorien, d. h. 5 verbale Hauptkategorien "sehr leise, leise, mittellaut, laut, sehr laut" plus vier Zwischenstufen und den zwei Begrenzungskategorien "unhörbar" und "zu laut".
- Terzbandrauschen von 2 sec Länge als Stimulus.
- Dynamikbereich der angebotenen Stimuli entweder in einer Vormessung individuell angepaßt oder in einem adaptiven Verfahren zu Beginn der Messung so aufgespannt, daß der gesamte Hör-Dynamikbereich des individuellen Patienten tatsächlich abgetastet wird (Brand et al. 1997).
- Die Form der Lautheitsfunktion ist abhängig von der Bandbreite des Stimulus (Brand, 1999).
- Als Zielfunktion wird eine aus 2 Geradenstücken zusammengesetzte Funktion gewählt, die in der Mitte durch einen Bezier-Fit geglättet ist (Brand, 1999). Diese Zielfunktion weist von allen möglichen Funktionen die geringste Abweichung von den Daten über einen weiten Pegelbereich und die gewünschte Flexibilität der Form bei nicht zu großer Parameterzahl auf.

Dieses als ACALOS (adaptive categorical loudness scaling) bezeichnete Verfahren wurde inzwischen vom EU-Projekt Hearcom als Vorschlag für die Europäische Harmonisierung audiologischer Meßverfahren in das empfohlene Inventar aufgenommen (weitere Informationen unter www.hearcom.eu).

## Neue Verfahren der Sprachaudiometrie

Aufgrund der einhelligen Kritik an der in der deutschsprachigen Routine-Audiometrie benutzten sprachaudiometrischen Verfahren (z. B. mögliche Manipulierbarkeit des Listenunterschiede, Freiburger Sprachtests aufgrund starker Vorkommen ungebräuchlicher Wörter, Verwendung des Einsilbers als im Deutschen relativ ungebräuchliche Wortform) und aufgrund der hohen Nachfrage nach entsprechend aussagefähigen zuverlässigen und Sprachtests wurde eine Reihe von Sprachtestverfahren in der Arbeitsgruppe des Autors entwickelt und z. T. für den praktischen Einsatz vorbereitet. Einen Überblick über die im deutschen Sprachraum für die Sprachaudiometrie anwendbaren Verfahren gibt Tabelle 1.

Test -Items	Name des Tests	Test Material pro Liste	Autoren
Logatome	Kieler Logatomtest	CVC	Müller-Deile (pers. Comm.)
(sinnleere	OLLO – Oldenburg	150 VCV und CVC, 40 Sprecher,	Meier et al., 2005
Einsilber)	logatome corpus	6 Variabilitäten	
sinnbehaftete	Freiburger	20 gebräuchliche Wörter	Hahlbrock, 1953
Einsilberl	Einsilbertest		
	Dreinsilber - Test	3 wiederholte Einsilber	Döring& Hamacher, 1992
	Einsilber Reimtest	33+33+34 Wörter pro Liste, 6 reimende Alternativen	Sotscheck, 1982
	Einsilber Reimtest	33+25+14 Wörter pro Liste,	v.Wallenberg & Kollmeier,
	(WAKO)	s reimende Minimalpaare als Alternativen	1989
	Verkürzter Reimtest	25 Reimtest-Trials pro Liste	Brand &Wagener, 2005
sinnbehaftete Zweisilber	Zweisilber -Reimtest	24+24+24 Wörter	Kliem & Kollmeier, 1994
		4 reimende Minimalpaare	
	Oldenburger Kinder	12 Wörter, je 3 reimende Bild-Paare	Kliem & Kollmeier, 1995
	Reimtest	100	1
	AAST-Test	6 Zweisilber, Bild-Antwort-Alternative	e Coninx, 2005
Mehrsilber	Freiburger Zahlentest	10 Zahlwörter 4-5 silbig	Hahlbrock, 1953
	Zahlentripel - Test	10 3-digit strings	Wagener et al.,2005
Sätze	Marburger Satztest	10 kurze sinnvolle Sätze	Niemeyer, 1967
	Basler Satztest	15 hoch vorhersagbare & 15 niedrig Vorhersagbare Sätze	Tschopp & Ingold 1992
	Göttinger Satztest	10 kurze sinnvolle Sätze	Wesselkamp & Kollmeier, 1994
	HSM-Satztest	20 kurze sinnvolle Sätze	Hochmair et al., Schmidt et al., 1997
	Oldenburger Satztest	10 syntaktisch feste, inhaltleere,	Wagener et al., 1999
		Sätze	

**Tabelle 1**: Überblick über die verschiedenen Sprachtest-Materialien für die deutsche Sprache, geordnet nach der Länge der Test-Items. Die verschiedenen Hintergrundfarben geben den Einsatzbereich wieder, der durch entsprechende Publikationen belegt wurde: weiß: Verwendung in Ruhe, hellgelb: Verwendung im Störschall, dunkelgelb: Verwendung in Ruhe und im Störschall.

Die Testverfahren unterscheiden sich nicht nur in der Länge der jeweiligen Test-Items, sondern auch in ihrem intendierten und validierten Verwendungszweck: Für die Verwendung mit Kindern wurde beispielsweise neben dem in Ruhe anzuwendenden Mainzer und Göttinger Kinder-Sprachtest der Oldenburger Kinder-Reimtest (OLKI, Einsatzgebiet im Störschall und in Ruhe) konzipiert. Für die Verwendung mit Cochlea-Implantat-Nutzern sieht der HSM-Satztest eine besonders geringe Sprechgeschwindigkeit vor. Während die älteren Sprachverständlichkeits-Tests (z. B. der Freiburger-Einsilber-Test) primär für den Einsatz in Ruhe konzipiert und validiert wurden, ist die Mehrzahl der in Tab. 1. aufgeführten jüngeren Testverfahren für den Einsatz unter Störschall konzipiert. Einige der moderneren Testverfahren wurden zusätzlich für den Einsatz in Ruhe validiert (Einsilber-Reimtest, verkürzter Reimtest, Oldenburger Kinder-Reimtest, Göttinger Satztest und Oldenburger Satztest). Sie weisen für den Einsatz in Ruhe keine signifikant andere Steigerung der Diskriminationsfunktion und keine signifikant größeren Listen-Unterschiede auf als für den Einsatz in Störschall. Sinnbehaftete Testmaterialien (insbesondere Sätze wie vom Göttinger Satztest oder HSM-Satztest) besitzen den Nachteil, dass sie nach einmaligem Gebrauch für einen bestimmten Patienten längere Zeit nicht wieder verwendet werden können, um einen WiedererkennungsEffekt des gesamten Satzes anhand eines einzelnen Wortteils auszuschließen. Dies wird beim Oldenburger Satztest ausgeschlossen: Eine (scheinbar) zufällige Kombination aus Name, Verb, Zahlwort, Adjektiv und Objekt (z. B. "Peter kauft 17 nasse Sessel") mit jeweils 10 Alternativen pro Wort erschwert die Wiedererkennung des spezifischen Satzes, so daß dieses Material beliebig häufig eingesetzt werden kann. Aufgrund der Kompatibilität dieses Verfahrens mit entsprechenden Tests in anderen Sprachen wird zudem eine Länder-übergreifende Normierung von Sprachverständlichkeits-Tests möglich (vgl. Kollmeier, 2007).

- Einsilber-Reimtest nach von Wallenberg und Kollmeier (WAKO-Test) Aufbauend auf dem Einsilber-Reimtest nach Sotscheck (1982) wurde eine für die Audiologie verwendbare Version mit phonetischen Minimalpaaren. gebräuchlichen Wörtern und aufgrund umfangreicher Validierungsmessungen als äguivalent anzusehenden Testlisten zusammengestellt (von Wallenberg und Kollmeier, 1989, Müller, 1992, Kollmeier et al., 1992). Dieses Testverfahren besitzt eine Reihe von Vorteilen gegenüber dem Freiburger Einsilbertest (z. B. geschlossene Ankündigungssatz, Wortauswahl, Analysierbarkeit von Phonemverwechselungen, hohe Homogenität des Testmaterials). Es konnte sich aufgrund der notwendigen technischen Voraussetzungen bisher (Computersteuerung zur Darstellung der Reim-Alternativen, Länge der Testlisten) nicht in der audiologischen Praxis durchsetzen. Da dieses Verfahren jedoch ein hohes Potential hat, sollten in Zukunft verkürzte Listen für den praktischen Einsatz in der Audiologie zusammengestellt werden.
- Zweisilber-Reimtest nach Kliem und Kollmeier (1994) Dieses Verfahren bietet sich aufgrund der größeren Häufigkeit von Zweisilbern als Einsilbern in der deutschen Sprache ebenfalls als Standard-Verfahren an. Sein praktischer Einsatz scheiterte jedoch bisher an der Länge der Testlisten und des relativ hohen technischen Aufwandes/geringen Verfügbarkeit von apparativen Voraussetzungen zur Durchführung des Tests. Im Zuge der voranschreitenden Automatisierung der Audiometrie sollte jedoch auch dieser Test zukünftig Einzug in die Praxis finden.
- Göttinger Satztest

Der von Wesselkamp et al. (1992) und Kollmeier und Wesselkamp (1997) vorgestellte Göttinger Satztest enthält 20 Listen von je 10 kurzen, sinnvollen Sätzen. Das Testmaterial ist hinsichtlich der Homogenität zwischen den Sätzen und zwischen den Testlisten ebenso ptimiert wie bezüglich der Anzahl der Wörter und Silben in jeder Testliste und der ungefähren phonetischen Äguivalenz der Testlisten. Das Verfahren hat in Form des BIRD-Sprachtest (Firma Starkey, Norderstedt) weite Verbreitung und Akzeptanz gefunden, insbesondere zur Hörgeräte-Anpassungen Überprüfung von und zur Durchführung der Sprachaudiometrie unter Störgeräuschen. Ein Nachteil sind jedoch die begrenzte Zahl von verfügbaren Listen und die schlechte Wiederholbarkeit des Tests mit derselben Versuchsperson, da die Testsätze leicht gemerkt werden können. Dies motivierte die Entwicklung des Oldenburger Satztests (s. u.).

Oldenburger Satztest:

Dieser Test wurde in Anlehnung an den schwedischen Satztest nach Hagermann (1984) konstruiert und bietet sich als ein Kandidat für die europäische Harmonisierung von Satztestverfahren an, da eine Übertragbarkeit in andere Sprachen verhältnismäßig leicht gegeben ist. Der von Wagener et al. (1998, 1999) vorgestellte Oldenburger Satztest benutzt als Inventar 5 mal 10 verschiedene Wörter, die jeweils in einem pseudozufällig zusammengestellten

sinnleeren Satz der Form "Name-Verb-Zahlwort-Eigenschaft-Objekt" angeordnet sind (z. B. Nina bekommt vier rote Blumen (vgl. Abb. 2)).



Abbildung 2: Testwortinventar und Konstruktionsprinzip des Oldenburger Satztests nach Wagener et al. (1998, 1999).

Bei der Aufsprache und der Resynthese wurde auf eine hohe Natürlichkeit der zusammengesetzten Sätze geachtet. Außerdem wurden die resultierenden Testsätze so zusammengestellt, daß eine möglichst hohe Homogenität zwischen den Testsätzen und den Testlisten resultiert (Wagener et al. 1999). Die einzelnen Testlisten weichen in ihrer mittleren Schwelle sehr gering voneinander ab (empirische Standardabweichung von 0,16 dB). Bei den Messungen zur Evaluierung des Satztests, die unabhängig von den Messungen zur Erstkonstruktion des Tests waren, ergab sich eine mittlere Schwelle (Signal-Rauschabstand für 50 % Verständlichkeit der Sätze) von minus 7,1dB und eine mittlere Steigung von 17,1 %/dB. Während diese mittlere Steigung aufgrund der Testkonstruktion erwartet wurde (Erwartungswert 17,2 % pro dB), ist die bei den Evaluationsmessungen ermittelte Schwelle um 1,3 dB höher als aufgrund der Testkonstruktion (Messungen mit trainierten Versuchspersonen) erwartet wurde. Dies ist auf den Trainingseffekt zurückzuführen: Wagener et al. (2000) konnten nachweisen, daß sich im Laufe der ersten 30 Testsätze bei naiven Versuchspersonen die Schwelle um bis zu 2 dB im Signal-Rauschabstand nach unten verschieben kann. Dieser Trainingseffekt fällt allerdings geringer aus, wenn der Sprachpegel festgehalten wird und der Störgeräuschpegel variiert wird (im Gegensatz zu dem für die Evaluationsmessungen angenommenen Prinzip). Ein derartiges Vorgehen wurde auch von Hagermann (1984) vorgeschlagen.

Für die praktische Durchführung des Oldenburger Satztests wurde eine manuelle Version erstellt, die vom Hörzentrum Oldenburg vertrieben wird. Die apparative Voraussetzung ist ein Audiometer mit CD-Spieler und einer Schrittweite von 1 dB. Bei dem manuellen adaptiven Verfahren wird je nach Anzahl der pro Satz richtig erkannten Wörter der Signal-Rauschabstand verändert. Bei der Verwendung von 30 Testsätzen ist damit eine Genauigkeit von 0,5 dB erreichbar, so daß Schwellenunterschiede von 2 dB sicher nachgewiesen werden können. Damit ist Β. der Nachweis eines binauralen Gewinns bei der beidohrigen z. Hörgeräteversorgung leicht nachzuweisen.

## • Oldenburger Kinder-Reimtest

Kliem und Kollmeier (1995) entwickelten ein Zweisilber-Reimtestverfahren mit Wörtern, die für den Einsatz von Vorschul- und Schulkindern geeignet sind. Für dieses Wortmaterial wurden von Achtzehn et al. (1997) Bildkarten entwickelt, in denen die jeweils 3 sich "reimenden" Antwortalternativen kindgerecht bildlich dargestellt sind (Abbildung 3).



Beule - Keule - Eule Tanne - Tasse - Tasche Brote - brüte - brate

Abbildung 3: Beispiele für die Antwortkarten des Oldenburger Kinder-Reimtest nach Achtzehn et al. (1997).

Evaluationsmessungen mit Vorschul-Ferner wurden und Schulkindern durchgeführt, bei denen die Gebräuchlichkeit der verwendeten Testwörter und die Verwendbarkeit der entwickelten Bildkarten ebenso getestet wurde, wie die Sprachverständlichkeitsschwelle für die einzelnen Wörter im Rahmen des Meßverfahrens. Aufgrund dieser **Evaluations-Messungen** konnten 10 gleichwertige Testlisten zu jeweils 12 Wörtern zusammengestellt werden, für die die 50%-Verständlichkeitsschwelle in Ruhe bei 23 dB SPL liegt (Brand et al., 1999). Die Steigung der Diskriminationsfunktion ist mit 6 % pro dB vergleichbar mit derjenigen des Freiburger Sprachtests, so daß sich der Oldenburger Kinderreimtest für die Sprachaudiometrie in Ruhe bei Kindern ähnlich einsetzen läßt wie der Freiburger Sprachtest bei Erwachsenen (Brand et al., 1999). Eine für die Praxis anwendbare, manuelle Version mit dem Bildkartenmaterial ist beim Hörzentrum Oldenburg erhältlich.

#### Räumliche Sprachtests

Zur Abschätzung des Vorteils binauralens (beidohrigen) Hörens gegenüber dem monauralen Hören und speziell zur Überprüfung des Versorgungsgewinns durch eine beidseitige Hörgeräte-Versorgung bietet sich die Durchführung von räumlichen Sprachtests an, bei denen der Nutzsprecher und der Störschall aus verschiedenen räumlichen Richtungen dem Patienten angeboten werden. Für den klinischen Alltag bietet sich dabei die in Abbildung 4 dargestellte räumliche Position an, die in Verbindung mit jedem Sprachverständlichkeitstest unter Störgeräusch (z. B. Oldenburger Satztest) benutzt werden kann.



Abbildung 4: Räumliche Anordnung für den räumlichen Sprachtest. Die obere Situation gibt die Bestimmung der ILD (Intelligibility Level Difference) an, die untere Situation die Ermittlung der Binaural Intelligibility Level Difference (BILD).

Zunächst wird in der Referenz-Situation S<sub>0</sub> N<sub>0</sub> (d. h. Sprache und Störsignal von vorne) die Sprachverständlichkeitsschwelle (Speech Reception Threshold, SRT) gemessen. Sie ist ein Maß für mögliche Einschränkungen des Sprachverstehens in Rauschen ohne räumlichen Detektionsvorteil. Im Vergleich dazu wird die Sprachverständlichkeitsschwelle bei seitlichem Einfall des Störgeräuschs (Situation bzw. SoN90  $S_0N_{270}$ ) gemessen. Die Differenz beider Sprachverständlichkeitsschwellen gibt die ILD (Intelligibility Level Difference) an, d. h. den räumlichen Vorteil bei der Sprachverständlichkeit, der sowohl aus einem monauralen Anteil (Kopfabschattungseffekt, d. h. besseres Signalrauschverhältnis an dem der Störquelle abgewandten Ohr) und einem binauralen Anteil (bessere Diskrimination durch Vergleich der Ohrsignale auf zentralem Niveau) zusammengesetzt ist. Um den binauralen Anteil getrennt vom monauralen Anteil zu bestimmen, wird in der zuletzt genannten Situation zusätzlich die Situation bei Verstöpselung/Vertäubung des "schlechteren" Ohres (d. h. der Störschallquelle zugewandten Ohres) ausgemessen. Der Unterschied zwischen der monauralen und der binauralen Situation wird als Binaural Intelligibility Level Difference (BILD) bezeichnet. Abb. 5 gibt für verschiedene Versuchspersonengruppen die gemittelten Werte für ILD und BILD wieder (aus Kollmeier et al., 1999). Obwohl tendenziell mit zunehmendem Hörverlust und mit zunehmender Asymmetrie des Hörverlusts die binaurale Leistungsfähigkeit abnimmt, läßt sich für den individuellen Patienten generell keine Vorhersage des binauralen Gewinns aus dem Audiogramm oder anderen audiologischen Parametern ableiten (vgl. Kinkel et al., 1992). Daher ist die binaurale Leistungsfähigkeit als unabhängige Komponente anzusehen, die separat ausgemessen werden sollte.



Abbildung 5: ILD (Kreise) und BILD (Quadrate) für unterschiedliche Probandengruppen.

## Literatur

- Achtzehn, J., V. Kühnel, B. Kollmeier and R. Schönfeld (1997). Zur Entwicklung eines Zweisilber-Kinder-Reimtests für die Audiologie. <u>Aktuelle phoniatrisch-</u> pädaudiologische Aspekte. M.Gross (Ed.). Heidelberg, median verlag. **5**.
- Böhme, G., Welzl-Müller, K. (1998). <u>Audiometrie. Hörprüfungen im Erwachsenen-</u> und Kindesalter. Bern, Verlag Hans Huber.
- Brand, T., V. Hohmann and B. Kollmeier (1997). Adaptive categorical loudness scaling. <u>Seventh Oldenburg Symposium on Psychological Acoustics</u>. A. Schick and M. Klatte. Oldenburg, BIS, Universität Oldenburg. **7:** 603-610.
- Brand, T. (1999). Analysis and Optimization of Psychophysical Procedures in Audiology. Dissertation FB Physik, Oldenburg, Universität Oldenburg.
- Brand, T., J. Achtzehn and B. Kollmeier (1999). Erstellung von Testlisten für den Oldenburger Kinder-Reimtest. <u>Z.f. Audiologie Supplement II</u>. K. Schorn (Ed.). Heidelberg, median verlag: 50-51.
- Dau, T., O. Wegner, V. Mellert and B. Kollmeier (2000). "Auditory brainstem responses (ABR) with optimized chirp signals compensating basilar-membrane dispersion." <u>J. Acoustical Soc. Am.</u> **107(3)**: 1530-1540.
- Grätz, K., M. Mauermann, R. Schönfeld and B. Kollmeier (1999). Niedrigpegel-Reflexaudiometrie - eine neue Methode zur Messung des Stapediusreflexes.

Zeitschrift für Audiologie, Suppl. 2. K. Schorn (Ed.). Heidelberg, median-Verlag: 125-126.

- Hagerman, B. (1984). "Clinical measurements of speech reception threshold in noise." <u>Scand-Audiol</u> **13**(1): 57-63.
- Kemp, D. T. (1978). "Stimulated acoustic emissions from within the human auditory system." J. Acoust. Soc. Am. 64: 1386-1391.

Kießling, J. (1995). "Zum überschwelligen Lautheitsanstieg bei Schallempfindungsschwerhörigen -- Konsequenzen für die Hörgeräte--Entwicklung und --Anpassung." <u>Audiologische Akustik(</u>2): 82--89.

- Kießling, J., B. Kollmeier and G. Diller (2008). <u>Versorgung und Rehabilitation mit</u> <u>Hörgeräten</u>. Stuttgart, Georg Thieme Verlag, 2. Aufl..
- Kinkel, M. and B. Kollmeier (1992). "Binaurales Hören bei Normal- und Schwerhörigen II: Analyse der Ergebnisse." <u>Audiologische Akustik</u> **31**: 22-33.
- Kliem, K. and B. Kollmeier (1994). "Entwicklung und Evaluation eines Zweisilber-Reimtestverfahrens für die deutsche Sprachaudiometrie." <u>Audiologische Akustik</u> 33: 4-15.
- Kliem, K. and B. Kollmeier (1995). "Überlegungen zur Entwicklung eines Zweisilber-Kinder-Reimtests für die klinische Audiologie." <u>Audiol.Akustik</u> **34**: 6-11.
- Kollmeier, B., M. C., M. Wesselkamp and K. Kliem (1992). Weiterentwicklung des Reimtests nach Sotscheck. <u>Moderne Verfahren der Sprachaudiometrie</u>. B. Kollmeier (Ed.). Heidelberg, Median-Verlag: 216-237.
- Kollmeier, B. (1997). Grundlagen. <u>Versorgung und Rehabilitation mit Hörgeräten</u>. J. Kießling, B. Kollmeier and G. Diller (Ed.). Stuttgart, Georg Thieme Verlag: 1-48.
- Kollmeier, B., Ed. (1997). <u>Hörflächenskalierung Grundlagen und Anwendung der</u> <u>kategorialen Lautheitsskalierung für Hördiagnostik und Hörgeräteversorgung</u>. Buchreihe Audiologische Akustik. Heidelberg, Median-Verlag.

Kollmeier, B. and M. Wesselkamp (1997). "Development and Evaluation of a German Sentence Test for objective and subjective Speech Intelligibility Assessment." Journal of the Acoustical Society of America **102**(4): 2412-2421.

Kollmeier, B., V. Hohmann and V. Kühnel (1999). Erprobung und Validierung von sprachaudiometrischen und anderen computer-gesteuerten Verfahren für die klinische Audiometrie. Zeitschrift für Audiologie, Supplement I. B. Kollmeier (Ed.). Heidelberg, median verlag: 73-78.

Launer, S., I. Holube, V. Hohmann and B. Kollmeier (1996). "Categorical loudness scaling in hearing-impaired listeners - Can loudness growth be predicted from the audiogram." <u>Audiologische Akustik</u> **35**(4): 156-163.

Laszig, R. (2000). Praxis der Audiometrie,. Stuttgart, Georg Thieme Verlag.

Müller, C. (1992). Perzeptive Analyse und die Weiterentwicklung eines Reimtestverfahrens für die Sprachaudiometrie, Universität Göttingen.

Neumann, J., S. Uppenkamp and B. Kollmeier (1994). "Chirp Evoked Otoacoustic Emissions." <u>Hear. Res.</u> **79**: 17-25.

Neumann, J., S. Uppenkamp and B. Kollmeier (1996). "Detection of the Acoustic Reflex below 80 dB HL." <u>Audiology and Neuro-Otology</u> **1**(6): 359-369.

- Sotscheck, J. (1982). "Ein Reimtest für Sprachverständlichkeitsmessungen mit deutscher Sprache als ein verbessertes Verfahren zur Bestimmung der Sprachübertragungsgüte." <u>Der Fernmeldeingenieur</u> **36**: Heft 5/1, 1--83.
- v. Wallenberg, E. L. and B. Kollmeier (1989). "Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten." <u>Audiologische Akustik</u> **28(2)**: 50-65.

- Wagener, K., V. Kühnel and B. Kollmeier (1999). "Entwicklung und Evaluation eines Satztests in deutscher Sprache I: Design des Oldenburger Satztests." <u>Zeitschrift</u> <u>für Audiologie/Audiological Acoustics</u> **38**(1): 4-15.
- Wagener, K., T. Brand and B. Kollmeier (1999). "Entwicklung und Evaluation eines Satztests in deutscher Sprache II: Optimierung des Oldenburger Satztests." Zeitschrift für Audiologie/Audiological Acoustics **38**(2): 44-56.
- Wagener, K., T. Brand and B. Kollmeier (2000). Internationale Vergleichbarkeit sprachaudiometrischer Verfahren: Einfluß von Störgeräusch und Meßprozedur auf die Satzverständlichkeit. <u>Fortschritte der Akustik - DAGA 2000</u>. Oldenburg, DEGA, Oldenburg: (in press).
- Wegner, O., T. Dau and B. Kollmeier (2000). Frequenzspezifische Messung früher akustisch evozierter Potentiale (FAEP) mit optimierten Chirp-Signalen. Fortschritte der Akustik - DAGA 2000. Oldenburg, DEGA: (im Druck).
- Wesselkamp, M., Kliem, K. and B. Kollmeier (1992). Erstellung eines optimierten Satztests in deutscher Sprache. <u>Moderne Verfahren der Sprachaudiometrie</u>. B. Kollmeier (Ed.). Heidelberg, Median-Verlag: 330-343.

# Audiometrische Standardverfahren, Otoakustische Emissionen, Auditorisch evozierte Potentiale und Qualitätssicherung

Thomas Janssen / München

## Psychoakustische (subjektive) Hörtests

Stimmgabeltests (Rinne/Weber) Hörweitenprüfung Reintonschwelle in Luft- und Knochenleitung (LL, KL) → Tonschwellenaudiometrie Überschwellige Tests → Fowler, SISI, Carhart, Langenbeck Lautheit → Kategoriale Lautheit (Hörfeldskalierung) Freiburger Sprachverständlichkeitstest → Hörverlust für Zahlen in dB, Einsilberverständlichkeit in % Sprachtest im Störgeräusch → z.B. Oldenburger Satztest

## Physiologische (objektive) Hörtests

*Trommelfell-Impedanz*<sup>-1</sup> in Abhängigkeit vom statischen Druck → Tympanometrie *Trommelfell-Impedanz*<sup>-1</sup> Änderung bei ipsi/contralateraler Stimulation → Stapediusreflex Otoakustische Emissionen → TEOAE, DPOAE Auditorisch evozierte Potentiale → FAEP, ASSR

### Pädaudiologische Hörtests

Reflexaudiometrie (Auropalpebralreflex, Moro-Reflex) Verhaltensaudiometrie (Zuwendungsreaktionen der Augen oder des Kopfes) Ablenkaudiometrie (Konditionierung durch Belohnung, Visuell Re-Enforcement) Spielaudiometrie (Aktive Mitarbeit) Sprachtests (Mainzer Kindersprachtest, Göttinger Kindersprachverständnistest, Heidelberger Konsonant-Vokal-Konsonant Test, Oldenburger Zweisilber-Kinderreimtest (OLKI), Würzburger Kindersprachtest, AAST)

Zum besseren Verständnis der Testsverfahren soll im Folgenden kurz auf die Funktion und die Funktionsstörungen des Hörorgans eingegangen werden. Das Hörorgan kann man sich als eine Kette hinter einander geschalteter Schall verarbeitender Systeme vorstellen. Das äußere Ohr dient der Schallaufnahme und Schallverstärkung (Trichter, Gehörgangsresonanz), das Mittelohr der Anpassung der unterschiedlichen Impedanzen der Luft und der Innenohrlymphe (Druckerhöhung durch unterschiedliche Flächen des Trommelfells und Steigbügelfußplatte, Kraftverstärkung durch Hebelwirkung der Gehörknöchelchen), die Cochlea zur mechanischen Verstärkung des Schalls (Äußere Haarzellen als cochleäre Verstärker) sowie der Umsetzung des mechanischen Reizes (Stereozilienauslenkung der Inneren Haarzellen) in elektrische Impulse (Aktionspotentiale auf den Hörnervenfasern der neuralen Hörbahn (Cochleariskern, oberer Olivenkomplex, Lemniscus lateralis, Colliculus inferior, Subcortex, Cortex).



Die genannten Verfahren dienen dazu, Ort und Grad einer Schwerhörigkeit zu bestimmen. Die Tonschwellenaudiometrie in Luft- und Knochenleitung kann nur zwischen Schalleitungsschwerhörigkeit (äußeres Ohr. Mittelohr) und Schallempfindungsschwerhörigkeit (cochleär, neural) unterscheiden. Die überschwelligen subjektiven Tests haben das Ziel, die Art der Schallempfindungsschwerhörigkeit zu bestimmen. Deren Aussagekraft ist aber beschränkt. Die OAE, die die cochleäre Funktion, und die AEP, die die neurale Aktivität erfassen können, müssen in Hinblick auf eine genauere Topodiagnostik Einsatz Wenn auch über die zum kommen. Tonschwellenaudiometrie eine Mitteohrschwerhörigkeit erkannt werden kann, so muss die Tympanometrie bemüht werden, um die Art der Mittelohrschwerhörigkeit (Otosklerose, Paukenerguß, Tubenfunktionsstörung, Kettenluxation) zu bestimmen.

Im Besonderen für die Hörgeräteanpassung ist die Bestimmung des Recruitments (Lautheitsausgleich) wichtig. Dies gelingt mit Hilfe der Hörfeldskalierung bei Erwachsenen und älteren Kindern. Bei kleinen Kindern können nur die Stapediusreflexschwelle und die DPOAE-Wachstumsfunktionen Anwendung finden.

Die Tonschwellenaudiometrie ist die am häufigsten durchgeführte Untersuchungsmethode, da sie die Funktionsstörung frequenzspezifisch und quantitativ erfassen kann. Prinzipiell sollen alle möglichen Verfahren zur Diagnostik eingesetzt werden. Je mehr Tests, umso besser die Diagnostik. Je genauer die Kenngrößen der gestörten Hörfunktion bestimmt werden können, umso besser gelingt die Anpassung einer Hörhilfe. Bei kleinen Kindern ist der Einsatz subjektiver Tests problematisch. Hier müssen die objektiven Tests verstärkt zum Einsatz kommen.

# Psychoakustische (subjektive) Hörtests

## Stimmgabeltests (Rinne/Weber)

Rinnetest: Stimmgabel (440Hz) wird angeschlagen und vor das Ohr gehalten (LL), dann auf das Mastoid gesetzt (KL). Frage: Wo ist der Ton lauter? Vor dem Ohr oder hinter dem Ohr? Antwort: vor dem Ohr  $\rightarrow$  Keine Schallleitungsschwerhörigkeit Antwort: hinter dem Ohr  $\rightarrow$  Schallleitungsschwerhörigkeit Webertest: Stimmgabel auf Vertex gesetzt. Frage: Wo wird der Ton gehört? Antwort: Mitte  $\rightarrow$  Normale Hörfunktion oder seitengleiche Hörminderung Antwort: Links oder rechts  $\rightarrow$  Innenohrschwerhörigkeit rechts oder links oder Mittelohrschwerhörigkeit links oder rechts



# Hörweitenprüfung

Nachsprechen zweistelliger Zahlen in Flüstersprache oder Umgangssprache in unterschiedlichem Abstand zwischen Patient und Untersucher. Abstand, bei dem gehört und richtig nachgesprochen wird, gibt groben Hinweis auf Grad der Hörstörung.



# Hörweitenprüfung

## Reintonschwelle

in Luft- und Knochenleitung (LL, KL) → Tonschwellenaudiometrie.

Prüfschall über Kopfhörer (LL) (Einschluss des Mittelohres) oder Knochenleitungshörer (KL) (Ausschluss des Mittelohres). Frequenzbereich 125 bis 10 kHz in Oktavschritten, Pegelbereich -10 dB HL bis 110 dB (abhängig von Frequenz und Schallzufuhr (LL, KL). Messvorschrift: Beginn mit LL des besseren Ohres (aus Vortests Stimmgabel, Sprachabstandsprüfung), dann LL am Gegenohr, dann KL. Ermittelung des Schwellenpegels durch Aufsteigende Methode oder Eingabelungsmethode. Symbole: LL rechts "o" verbunden durch (rote) Linien, LL links "x" verbunden durch (blaue) Linien, KL rechts ">" verbunden durch gestrichelte (rote) Linien, KL links "<" verbunden durch gestrichelte (blaue) Linien.

Sowohl bei LL als auch bei KL wird die Hörschwelle Normalhörender bei jeder Frequenz als Bezugswert (0 dB) herangezogen. Im Tonschwellenaudiogramm wird der Hörverlust in dB HL gegenüber diesem Bezugswert eingetragen.



# Überschwellige Tests

Fowler (Lautheitsausgleichstest). Vergleich des Lautheitseindrucks beider Ohren. Bedingung: Seitendifferenz mehr als 30 dB. Bei Lautheitsausgleich → Recruitment SISI (short increment sensitivity index): Wieviele von 20 dargebotenen 1-dB Pegelsprüngen eines Dauertons (20 dB über Hörschwelle) werden erkannt; > 60% → Recruitment Carhart (Schwellenschwundtest): Wenn bei einem Dauerton die Schwelle um mehr als 30 dB abwandert → neurale Schädigung

Geräuschaudiometrie nach Langenbeck (Mithörschwelle): Erhöhung der Mithörschwelle  $\rightarrow$  neurale Schädigung.



Fowlertest bei Hochtonhörverlust rechts. Lautheitsvergleich bei 4 kHz. Hörverlust rechts 50 dB, links 20 dB. Die Werte der linken Seite sind rechts auf der benachbarten Linie (eigentlich 6 kHz) eingetragen, aber durch das x für den Pogel links (4 kHz) gekennzeichnet. Eis 90 dB ist die Differenz von 30 dB ausgeglichen. *Lautheitsausgleich rechts*: Innenchrschädigung.

### Schwellenschwundtest (Carhart)



#### Geräuschaudiometrie



#### Lautheit

Kategoriale Lautheit (Hörfeldskalierung)

Bewertung der empfundenen Lautheit auf einer vorgegebenen Kategorialskala (z.B. nicht gehört, leise, mittellaut, laut, sehr laut, zu laut). Bei Innenohrschwerhörigkeit mit Recruitment gibt der Patient bei hohen Reizpegeln ein normales Lautheitsempfinden an. Bei kleinen Reizpegeln wird entsprechend dem Hörverlust nicht gehört → Versteilung der Lautheitsfunktion. Bei der Schallleitungschwerhörigkeit und der neuralen Schwerhörigkeit erfolgt eine Verschiebung der Lautheitsfunktion entlang der X-Achse. Hier gibt es auch bei hohen Reizpegeln keinen Lautheitsausgleich.


#### Sprachtests

Freiburger Sprachtest: Zahlentest (10 Gruppen zu je 10 zweistelligen Zahlen) + Einsilbertest (20 Gruppen zu je 20 Einsilbern)

Ergebnis:

1. Hörverlust für Zahlen in dB = Sprachpegel, bei dem der Patient 50% der Zahlen versteht. Startpegel 20 dB über Hörschwelle bei 500 Hz, nächster Pegel 5 + Startpegel.

2. Sprachverstehen der Einsilber in % bei Sprachpegel 65, 80, 95, 110 dB SPL.

#### Sprachaudiogramm



Oldenburger Satztest: Ermittlung der Sprachverständlichkeitsschwelle für Sätze im Störgeräusch = Ermittlung des Signal-Rausch-Verhältnisses, bei dem 50% der Sätze einer Testliste (40 Listen mit je 30 Sätzen) verstanden werden.



#### Physiologische (objektive) Hörtests

#### Tympanometrie

Das Tympanometer misst die Nachgiebigkeit (Compliance = Impedanz <sup>-1</sup>) des Trommelfells in Abhängigkeit des im Gehörgang mittels einer Pumpe eingeleiteten Über - und Unterdrucks (Tympanogramm). Beim Überdruck (bis +200 daPa) wird das Trommelfell in die Paukenhöhle hineingedrückt, beim Unterdruck (bis- 400 daPa) wird das Trommelfell in den Gehörgang hineingezogen. Mit zunehmendem Über- bzw. Unterdruck nimmt die Beweglichkeit des Trommelfells immer mehr ab. Bei normaler Mittelohrfunktion erhält man eine Glockenkurve mit maximaler Compliance um den Atmosphärendruck, und minimaler Compliance bei den Extremwerten +200 und -400 daPa (Glockenkurve mit großer Compliance). Bei der Otosklerose ist wegen der eingeschränkten Beweglichkeit des Steigbügels die Beweglichkeit des Trommelfells eingeschränkt (Glockenkurve mit kleiner Compliance). Beim Paukenerguss ist wegen der mitschwingenden Sekretflüssigkeit die Beweglichkeit stark eingeschränkt (Flache Kurve, keine Gipfelbildung). Bei der Tubenfunktionsstörung ist der Gipfel der Glockenkurve in den negativen Druckbereich entsprechend dem in der Paukenhöhle herrschenden Unterdruck verschoben. Der Unterdruck entsteht durch den Sauerstoffverbrauch des (entzündlichen) Gewebes bei verschlossener Tube als Folge vergrößerter Rachenmandeln (Adenoide).

#### Tympanometrie



Mittelohr-Erguss

Tubenfunktionsstörung



Otosklerose



#### Stapediusreflex

Sowohl mit ipsi- als auch mit contralateral applizierten Tönen (1, 2, 4 kHz) oder Rauschen (Breitband, Schmalband) lässt sich der Stapediusreflex auslösen. Bei Kontraktion des Stapediusmuskels ist die Nachgiebigkeit der Gehörknöchelchenkette eingeschränkt (Schutzfunktion). Das Tympanometer registriert die veränderte Nachgiebigkeit während der Reflexauslösung. Die Reflexschwelle eines Normalhörenden liegt bei etwa 80 dB HL. Dies ist auch der Fall bei Patienten mit Funktionsstörungen der äußeren Haarzellen bei normaler Funktion der inneren Haarzellen (Recruitment). Bei der Mittelohrschwerhörigkeit und der neuralen Schwerhörigkeit ist die Reflexschwelle etwa dem Hörverlust entsprechend verschoben (oder nicht messbar).

#### Otoakustische Emissionen

Otoakustische Emissionen (OAE) sind Schallaussendungen des Innenohres. OAE können spontan vorhanden sein oder mit akustischen Reizen erzwungen werden. Die verblüffende Tatsache, dass Schall aus dem Innenohr emittiert wird, setzt die Existenz einer *Energiequelle in der Kochlea* voraus. Seit dem Nachweis der *Motilität der äußeren Haarzellen* wird angenommen, dass die äußeren Haarzellen als *mechanische Verstärker* fungieren, die die Sensitivität und Trennschärfe des Hörorgans erhöhen. OAE sind Epiphänomene des normalen Hörvorgangs. Sie sind Ausdruck des nicht-linearen Verstärkungsprozesses in der Kochlea. OAE schließen eine bisher bestehende Lücke, da sie die direkte Prüfung der Innenohrfunktion erlauben.

Wegen ihrer einfachen Handhabbarkeit und der Tatsache, dass OAE (i.G. zu den AEP) bereits mit der Geburt voll entwickelt sind, sind sie die Methode der Wahl beim Neugeborenen Hörscreening. In der pädaudiologischen Diagnostik sind sie eine wertvolle Hilfe zur frequenzspezifischen und quantitativen Bestimmung kochleärer Funktionsstörungen. Da sie die kompressive Verstärkung der Kochlea direkt widerspiegeln, können sie auch eine Methode zur Bestimmung des Rekruitments sein. Mit Hilfe der OAE können kleinste Funktionsstörungen erfasst werden. Sie sind daher eine höchst sensitive Methode zur Erkennung beginnender Hörstörungen und Bestimmung ihres Verlaufs bei Lärmexposition oder Einnahme von ototoxischen Medikamenten.



Man unterscheidet:

**spontane OAE (SOAE),** die in der Form tonaler Schallsignale ohne Einwirkung eines akustischen Reizes fortwährend emittiert werden,

evozierte OAE, die durch Einwirkung eines äußeren Schallereignisses entstehen. Die evozierten otoakustischen Emissionen werden je nach Art der Schallreizung unterteilt in

*transitorisch* evozierte OAE (**TEOAE**), *simultan* evozierte OAE (**SEOAE**),

Distorsionsprodukt-OAE (**DPOAE** oder **DP**).

Durch einen Schallreiz werden die nach ihrem Standort in der Cochlea auf unterschiedliche Frequenzen abgestimmten äußeren Haarzellen zu Schwingungen angeregt. Bei *transienter* Anregung mit Klicks oder Tonimpulsen werden die äußeren Haarzellen *angestoßen*. Ihr Schwingungsverhalten entspricht einer Stoßantwort. Frequenz und Dauer der Stoßantwort sind abhängig von der Eigenfrequenz der äußeren Haarzelle und damit vom Standort der Haarzelle in der Kochlea. Bei *stationärer* Anregung des Ohres mit Tönen werden die im schmalen Bereich der Basilarmembranausbauchung liegenden Haarzellen in *periodische* Schwingungen versetzt, deren Frequenz der des anregenden Tones entspricht. Auf welche Weise die äußeren Haarzellen auch in Kontraktionen versetzt werden - sei es wie bei den *spontanen OAE* durch eine *interne Energiequelle* oder wie bei den *evozierten* OAE durch Stimulation mittels Schallenergie von außen - jede Kontraktion der mikromechanischen Strukturen des Corti-Organs setzt Schwingungsenergie in der Cochlea frei, die als retrograde Schallaussendung mit einem empfindlichen Mikrophon im äußeren Gehörgang gemessen werden kann.

### TEOAE



TEOAE werden mit kurzen Schallimpulsen (Klick, Tonimpuls) ausgelöst und weisen ein breitbandiges Spektrum auf. Sie sind die Summe der Emissionen aus einem weiten Bereich der Cochlea. Bei Anregung des Ohres mit einem transienten Schallreiz (Klick) breitet sich eine Welle auf der Basilarmembran aus, die von basal nach apikal läuft. Auf ihrem Weg zur Schneckenspitze stößt sie die äußeren Haarzellen hintereinander an. Die je nach ihrem Standort in der Kochlea auf unterschiedliche Frequenzen abgestimmten äußeren Haarzellen erzeugen dabei infolge der Filtereigenschaften der Basilarmembran und des Cortischen Organs Stoßantworten, deren Schwingungen – abhängig von ihrem Standort in der Kochlea - unterschiedliche Frequenz und Dauer haben. TEOAE sind die Summe der Stoßantworten der äußeren Haarzellen. Die aus dem basalen Bereich der Kochlea stammenden TEOAE-Komponenten tiefere Frequenzen. Die basalen TEOAE-Komponenten treten früh, die medialen und apikalen TEOAE-Komponenten treten entsprechend der Laufzeit in der Cochlea später auf (Latenz). Der Vorteil der TEOAE ist der, dass mit *einem* 

Reiz nahezu alle Haarzellen in der Cochlea angestoßen werden und ihre Fähigkeit zur Kontraktion in einem Messvorgang erfasst werden kann. Sind die äußeren Haarzellen schwingungsfähig, so emittieren sie Stoßantworten mit einer ihrem Standort entsprechenden Frequenz. Sind äußere Haarzellen in bestimmten Cochleaabschnitten in ihrer Funktion gestört, so senden diese keinen Schall aus, und es fehlt in der Summe der emittierten Stoßantworten mit der jeweiligen Frequenz und Latenz.

TEOAE sind bei Neugeborenen die Hörscreening-Methode der Wahl, weil sie schnell und einfach messbar sind und durch den Einsatz statistischer Tests eine automatische Befunderhebung möglich ist: 1. Verdacht auf Schwerhörigkeit, weitere Untersuchungen notwendig (fail), 2. kein Verdacht auf das Vorliegen einer Schwerhörigkeit (pass). Bei Fruchtwasserresten in der Paukenhöhle in den ersten Lebensstunden und –tagen sind TEOAE nicht nachweisbar. Es wird daher ein mehrstufiges Hörscreening angewendet: Weiderholungsmessung der TEOAE ab dem 3. Lebenstag, zusätzliche Messung der FAEP. Es ist zu beachten, dass retrocochleäre Hörstörungen und eine auditorische Neuropathie (Störungen der Synapsenfunktion der Inneren Haarzellen) unerkannt bleiben, wenn nur TEOAE gemessen werden.



Während TEOAE eher eine qualitative Messgröße zur Bestimmung von kochleären Funktionsstörungen sind, sind die DPOAE in der Lage, guantitative Informationen über den Sensitivitäts-, Kompressions- und Trennschärfeverlust des kochleären Verstärkers zu geben. DPOAE werden mit 2 Tönen (Primärtöne) benachbarter Frequenz ausgelöst und entstehen als direkte Folge der Nichtlinearität des kochleären Verstärkungsmechanismus. Sie geben die Emissionen aus der schmalen Überlappungszone der Wanderwellen der Primärtöne wieder. Die DPOAE haben eine Frequenz, die sich aus der Kombination der Primärtonfrequenzen f<sub>1</sub> und f<sub>2</sub> zusammensetzt. Beim Menschen haben die Distorsionsproduktemissionen mit der Frequenz  $2f_1-f_2$  die größte Schalldruckamplitude. Bei Verschiebung des Überlappungsbereichs entlang der Cochlea können Verzerrungen an unterschiedlichen Cochleaorten gemessen werden. Das Ergebnis ist dann ein DPOAE-Gramm. Mit einem DPOAE-Gramm kann man grob nachschauen, in welchem Bereich der Cochlea eine Funktionsstörung vorliegt. Hierbei ist aber zu beachten, dass die DPOAE schwellennah ausgelöst werden müssen. Viele Leute wundern sich, warum DPOAE-Gramm und Tonschwellenaudiogramm nicht übereinstimmen. Sie haben dann einfach überschwellige Reizpegel und nicht schwellennahe Reizpegel zur Auslösung der DPOAE verwendet.



Wenn man die Frequenzen der Primärtöne konstant hält und ihre Pegel ändert, erhält man eine *DPOAE I/O-Funktion*. Bei der Einstellung der Primärtonpegel ist die kompressive Nicht-Linearität der Cochlea am Ort f2 zu beachten. Die Primärtöne müssen daher in ihrem Pegel unterschiedlich eingestellt werden, z.B. nach der "Pegelschere" (L1 = 0,4 L2 -39). Wenn man so verfährt, dann spiegelt die DPOAE I/O-Funktion die nicht-lineare Verstärkung der äußeren Haarzellen wide. Mit *extrapolierten DPOAE I/O-Funktionen* kann eine relativ genaue Schätzung des Hörverlustes in der Form eines *DPOAE-Kochleogramms* vorgenommen werden. Mit den DPOAE-Kochleogrammen lässt sich bei Neugeborenen und Kindern eine frequenzspezifische und quantitative Aussage über den Hörverlust treffen. Mit Hilfe des Steigungsprofils kann der Kompressionsverlust der cochleären Verstärkung bestimmt werden. So lassen sich aus den DPOAE I/O-Funktionen Anpassparameter für Hörgeräte ableiten. Dies allerdings nur bis zu Hörverlusten von 50 dB HL. Bei Hörverlusten größer als 50 dB HL müssen die frequenzspezifischen auditorischen steady-state Antworten (ASSR) weiterhelfen.





Die DPOAE kann man unterdrücken, wenn man einen Suppressorton appliziert, den man in Frequenz und Pegel so variiert, dass sich die DPOAE-Amplitude um einen bestimmten Wert ändert. Man erhält dann eine *DPOAE Iso-Suppressionstuningkurve*. Aus den DPOAE Iso-Suppressionstuningkurven können Informationen über die mechanische Abstimmung der Cochlea gewonnen werden. Die DPOAE Iso-Suppressionstuningkurven sind deckungsgleich mit den neuralen Tuningkurven. Das bedeutet die Trennschärfe des Hörorgans findet in der Cochlea statt. Neugeborene haben gleiche DPOAE Iso-Suppressionstuningkurven wie Erwachsene. Die Cochlea ist damit schon von Geburt an reif.

Mit Hilfe der DPOAE lässt sich auch die Reflexstärke des efferenten Hörsystems bestimmen. Entweder mit der DPOAE-Adaptation oder der contralateralen DPOAE-Suppression. Die DPOAE-Adaptation spiegelt das Einschwingverhalten der äußeren Haarzellen wider. Beim Einschalten der Primärtöne kommt es zunächst zu einem raschen Anstieg des DPOAE Pegels innerhalb von etwa 200 ms. Danach nimmt die Amplitude um etwa 1 dB ab und bleibt konstant. Das bedeutet, dass das efferente Hörsystem die Schwingung der äußeren Haarzellen dämpft. Auch die akustische Stimulation des Gegenohres führt zu einer Dämpfung der Haarzellschwingung. Das Ergebnis ist dann eine contralaterale DPOAE-Suppression. Die contralaterale DPOAE-Suppression kann Werte von über 10 dB annehmen, wenn bestimmte Primärtonpegelkombinationen verwendet werden. Damit ist die contralaterale DPOAE-Suppression eine geeignete Messgröße zur Untersuchung der Funktion des efferenten Hörsystems. Das efferente Hörsystem gibt noch viele Rätsel auf. Man weiß nicht, ob es der Verbesserung der Signalerkennung im Störgeräusch oder zum Schutz bei zu lauter Schalleinwirkung dient. Die DPOAE könnten ein Prädiktor der Vulnerabilität der Cochlea sein, wenn man davon ausgeht, dass Ohren mit einer großen Reflexstärke besser vor Lärm geschützt sind als Ohren mit kleiner Reflexstärke.

#### Auditorisch evozierte Potentiale → FAEP, ASSR

Auditorisch evozierte Potentiale repräsentieren die an der Kopfhaut über Klebeelektroden gemessene reizsynchrone elektrische Aktivität auf verschiedenen Stufen der Hörbahn. Die an der Kopfhaut abgegriffene Spannung liegt je nach Entstehungsort zwischen 100 Nanovolt und 1 Mikrovolt. Zur Extraktion der reizsynchronen Antwort aus dem EEG ist eine Mittelung einzelner Antworten notwendig (n=2000). Der Signal/Störabstand nimmt mit der Wurzel aus der Anzahl n der Mittelungen zu. Bei kleinen Potentialamplituden muss die Anzahl der Mittelungen erhöht werden (n=4000). Man unterteilt die Potentialkomponenten nach ihrem zeitlichen Auftreten nach Reizbeginn (Latenz) in frühe akustisch evozierte Potentiale (FAEP), in Potentiale mittlerer und später Latenz (MLR, SAEP).

Die FAEP bilden drei prominente, mit römischen Ziffern (I, III, V) gekennzeichnete Wellen aus. Mit abnehmendem Reizpegel nimmt die Latenz der Potentialwellen zu (Welle V bei 80 dBnHL = 5,5ms, bei 10 dB nHL = 8ms bei einem Normalhörenden). Bei einer Mittelohrschwerhörigkeit liegt die Potentialschwelle in etwa dem Hörverlust entsprechend höher (gilt streng genommen nur für den pantonalen Hörverlust). Die Latenz ist verschoben entsprechend dem Hörverlust; d.h. bei einer Mittelohrschwerhörigkeit von 40 dB beispielsweise wird bei einem Reizpegel von 60 dB eine Antwort erzielt, die der Antwort eines Normalhörenden bei einem Reizpegel von 20 dBnHL entspricht. Bei einer pancochleären Schwerhörigkeit ist die Potentialschwelle entsprechend dem Hörverlust verschoben, die Latenz ist jedoch normal. Bei einer basocochleären Schwerhörigkeit ist die Latenz entsprechend der Laufzeit der Wanderwelle über den gestörten Sinneszellabschnitt verlängert. Die Verschiebung der Potentialschwelle entspricht dabei nicht unbedingt dem Hörverlust. Bei einem Hochtonabfall mit begrenztem Hörverlust ist die Latenz bei kleinen Reizpegeln verschoben, bei hohen Reizpegeln normalisiert sich die Latenz.

Die Amplitude der FAEP gibt keine direkte Information über den Hörverlust. Allerdings fällt beim Recruitment eine sprunghafte Änderung der Potentialamplitude mit steigendem Reizpegel auf. Bei einer retrocochleären Hörstörung ist die Inter-Peak-Latenz (Zeit zwischen Welle I und V) entsprechend der verlangsamten neuralen Transmission auf den direkt (Neurofibromatose) oder indirekt (Druck durch Raumforderung) geschädigten Hörnervenfasern. Die Inter-Peak-Latenz beträgt im Mittel bei Normalhörenden 3,8 ms (Frauen) bzw. 4,2 ms (Männer). Bei einer verlängerten Inter-Peak-Latenz von 0,2 ms besteht Verdacht auf eine retrocochleäre Störung.

Die FAEP werden üblicherweise mit Klick-Reizen ausgelöst. Wegen der hohen Anfangsgeschwindigkeit (260 km/h) der durch den Klick in der Cochlea ausgelösten Welle ist die Synchronisation der Entladungen (Aktionspotentiale auf den 30 000 Hörnervenfasern) im basalen Bereich der Cochlea am größten. Die basalen Sinneszellen und Hörnervenfasern tragen somit am meisten zur Bildung der FAEP bei. Wegen der transienten Anregung können keine frequenzspezifischen Antworten erzeugt werden. Durch Maskierung gegrenzter Cocheabereiche mittels Hochpass-Rauschen (derived response technique) oder Notched-Noise ist eine indirekte frequenzspezifische Aussage über den Hörverlust möglich. Wegen der geringen Anzahl der zur Potentialbildung beitragenden Hörnervenfasern ist die Potentialamplitude sehr klein und so die Aussagekraft der Antworten eingeschränkt. Die mit Tonimpulsen ausgelösten FAEP können nur dann eine frequenzspezifische Aussage über den Hörverlust treffen, wenn sie mit kleinen Reizpegeln ausgelöst werden. Bei hohen Reizpegeln, die ja bei Patienten mit größeren Hörverlusten zur Anwendung kommen müssen, ist wegen der basalen Miterregung (auch bei tieffrequenten Tonimpulsen werden basale Sinneszellen aktiviert, obwohl die Wanderwelle nach der Frequenz-Orts-Transformation im apikalen Bereich der Cochlea ihre maximale Auslenkung hat) keine frequenzspezifische Erfassung des Hörverlustes möglich. Dies gilt auch für den Chirp.

In der klinischen Diagnostik werden die FAEP zur groben Erfassung des Hörverlustes bei Kindern und zur Erkennung von Akustikusneurinomen und Hirnstammtumoren herangezogen. Bei einem hohen Vigilanzniveau (Wachheit) und bei abgeschlossener Hörbahnreifung können die frequenzspezifischeren MLR und SAEP eingesetzt werden. Da diese Forderungen bei kleinen Kindern nicht erfüllt sind, spielen die MLR und SAEP in der Pädaudiologie nur eine untergeordnete Rolle.

Ereigniskorrelierte Potentiale (event-related-potentials, ERP) dienen zur Erfassung von zentralen Störungen; die Mismatch Negativity (MMN), die die unbewusste Detektion seltener Ereignisse (Devianten) innerhalb einer Serie von Standardreizen erkennt, und die Welle P300 bzw. P400 (late positive complex), die bei einer Reaktion auf einen unerwarteten Reiz bzw. bei semantischer Verarbeitung von Sprache nachweisbar sind. Diese Verfahren können nur bei ältern Kindern z.B. zur Objektivierung von auditiven Verarbeitungs- und Wahrnehmungsstörungen und Sprachentwicklungsstörungen eingesetzt werden.

#### Akustisch evozierte Potentiale



Die auditorischen Steady State Antworten (ASSR) bieten die Möglichkeit zur frequenzspezifischen Erfassung des Hörverlusts. Allerdings sind zu ihrer Ableitung lange Messzeiten erforderlich.



Die ASSR werden mit amplitudenmodulierten Dauertönen ausgelöst. Typische Modulationsfrequenzen (fm) sind 40 Hz, 80 Hz und 120 Hz. Je niedriger die Modulationsfrequenz, umso höher der Entstehungsort der Potentiale auf der aufsteigenden Hörbahn und umso größer die Potentialamplitude (und damit bessere Messbedingungen). Wegen der Vigilanzabhängigkeit können bei sedierten Kindern nur die 80Hz oder 120Hz ASSR mit der kleineren Amplitude eingesetzt werden. In den letzten Jahren sind eine Reihe von Veröffentlichungen über die Methode der ASSR erschienen (Überblick s. Picton et. al). Es gibt allerdings vergleichsweise wenige Veröffentlichungen über ihren klinischen Einsatz.

Der Vorteil gegenüber den FAEP ist der, dass die Sinneszellen in der Cochlea entsprechend der Trägerfrequenz (fc= 125 Hz – 8 kHz) in gezielten Arealen erregt werden können, die synchronisierte Aktivität jedoch im Hirnstamm (Modulationsfrequenz 80 Hz) oder Subcortex (Modulationsfrequenz 40 Hz) stattfindet.



40 Hz ASSR vs. 80 Hz ASSR

Ausblick: DPOAE lassen sich nur bei Hörverlusten bis 50 dB HL messen. Unterhalb 1 kHz sind wegen des hohen Störgeräuschs DPOAE nur schwer messbar. Bei hohen Frequenzen > 6 kHz tritt das Problem der stehenden Wellen auf. Das bedeutet, dass der an der Sondenspitze gemessene Schalldruck nicht dem Schalldruck am Trommelfell entspricht. Um eine Schätzung des Hörverlusts im gesamten Frequenz- und Dynamikbereich des Hörens vornehmen zu können, müssen DPOAE und ASSR gemessen werden: DPOAE wo möglich, ASSR wo nötig. DPOAE und ASSR lassen sich auch simultan messen, z.B. können während der langen ASSR-Messzeiten DPOAE ohne zusätzlichen Zeitaufwand registriert werden.







#### Pädaudiologische Hörtests

In der pädaudiologischen Diagnostik werden psychoakustische Prüfverfahren und physiologische Messverfahren zur Lokalisation und quantitativen Erfassung einer Schwerhörigkeit angewendet. Die audiometrischen Tests liefern Kenngrößen der gestörten Hörfunktion. Physiologische Tests (Tympanometrie, OAE und AEP) sind essentielle Bestandteile der audiometrischen Testbatterie, da sie nicht auf die aktive Mitarbeit des Kindes angewiesen sind, setzen aber ein ruhiges, in der Regel spontan schlafendes oder sediertes Kind voraus. Sie bieten darüber hinaus die Möglichkeit, den Hörschaden zu lokalisieren.

Die unverzichtbare Basis für eine erfolgreiche Hörgeräte-Versorgung ist die Ermittlung einer seitengetrennten und frequenzspezifischen Hörschwelle. Um die Befunde untereinander abzusichern, bedarf es der Zusammenschau der Ergebnisse verschiedener audiometrischer Verfahren, sowohl der psychometrischen, als auch der physiologischen Verfahren.

Zur Diagnose frühkindlicher Hörstörungen stehen altersabhängig verschiedene subjektive Hörprüfverfahren zur Verfügung. Im Rahmen subjektiver Hörprüfungen werden vom Untersucher die stark altersabhängigen Reflex- und Reaktionsschwellen ermittelt, die als Antwort auf akustische Reize nachweisbar sind. Die Schwellenwerte zur Auslösung von Reflexen oder Verhaltensänderungen nach akustischer Reizung unterliegen jedoch einem altersbedingten Reifungsprozess; die korrekte Bewertung dieser Antworten gelingt nur einem erfahrenen Untersucher. Als subjektive Hörprüfverfahren kommen im ersten und zweiten Lebensjahr die Reflex- und Verhaltensaudiometrie zum Einsatz, ab ca. dem dritten Lebensjahr die Spielaudiometrie.

Die subjektiven Hörprüfmethoden beim Säugling und Kleinkind lassen sich in folgende Kategorien einteilen:

#### Reflexaudiometrie

Bei der Reflexaudiometrie werden die in den ersten Lebensmonaten noch nachweisbaren unbedingten Reflexe des Kindes, wie z. B. das Zucken der Augenlider (Auropalpebralreflex), das reflexartige Anziehen der Arme und Beine (Moro-Reflex, Schreck-Reflex), Änderungen der Atmung, der Mimik und Gestik als Antwort auf akustische Reize (Töne, Rauschen, Kinderlieder etc.) beurteilt. Die Reaktionen des Kindes auf Bewegungen und visuelle Reize können jedoch dabei oft als Hörreaktionen fehl gedeutet werden.

#### Verhaltensaudiometrie

Mit zunehmendem Alter lassen die unbedingten Reflexe der Kinder nach, so dass ab dem 3.-4. Lebensmonat eine Verhaltensaudiometrie eingesetzt wird. Dabei werden Verhaltensänderungen des Kindes, wie z.B. Zuwendungsreaktionen der Augen oder des Kopfes zur Schallquelle, Änderungen der Mimik, Gestik oder der Atmung sowie Aktivitätsänderungen als Antwort auf akustische Reize im Freifeld sowie über Knochenleitungshörer geprüft.

#### Ablenkaudiometrie

Ab dem 6.-7. Lebensmonat lässt sich die sog. Ablenkaudiometrie durchführen. Dieses Verfahren beruht darauf, dass auf einem Bildschirm kindgerechte Bilder gleichzeitig mit der Präsentation eines lauten, überschwelligen akustischen Signals gezeigt werden. In der Konditionierungsphase lernt das Kind, dass visueller und auditiver Reiz immer gleichzeitig erscheinen. In der anschließenden Untersuchungsphase lässt man die visuelle Darbietung etwas verzögert gegen die auditive Stimulation auftreten. Das Kind wird für die richtige Reaktion auf einen auditiven Stimulus mit einem neuen Bild belohnt. Der Reizpegel wird dabei bis zur Reaktionsschwelle erniedrigt (d.h., man tastet sich langsam an die "Hörschwelle" heran).

#### Spielaudiometrie

Die Spielaudiometrie kann ca. ab dem 3. Lebensjahr eingesetzt werden. Im Gegensatz zur Verhaltensaudiometrie arbeitet das Kind aktiv mit. Auch hier bedarf es einer Konditionierungsphase. Das Kind führt nach Gabe eines akustischen Reizes eine Spielhandlung aus (z.B. es steckt ein Holzklötzchen in ein Steckbrett).

#### Sprachtests

Sprachtests untersuchen das Diskriminationsverhalten für komplexe Signale mit semantischer Information. Im deutschsprachigen Raum werden bei Vorschul- und Schulkinder in der Regel der Mainzer Kindersprachtest und der Göttinger Kindersprachverständnistest I und II verwendet. Dazu gibt es den Heidelberger Konsonant-Vokal-Konsonnat Test, den Oldenburger Zweisilber-Kinderreimtest (OLKI) und den Würzburger Kindersprachtest. Neu ist der adaptive auditive Sprachtest (AAST), wo sechs Zweisilber (Flugzeug, Eisbär, Schneemann, Lenkrad,

Handschuh, Fußball) verwendet werden. Das Kind muss – wenn es das Wort verstanden hat - das entsprechende Bild auf dem Touchscreen eines PDAs auswählen. Das Testergebnis ist eine Zahl in dB, die aussagt, bei welcher Intensität (dB) die Hälfte der Worte verstanden wird.



#### DPOAE-Kochleogramme

Bei der Verhaltens- oder Spielaudiometrie wird eine hohe Diskrepanz zwischen der Reaktionsschwelle und der tatsächlichen Hörschwelle beobachtet. Je nach Alter kann die Differenz zwischen 60 dB (4 Monate altes Kind) und 25 dB (2 Jahre altes Kind) betragen. Das DPOAE-Kochleogramm erlaubt eine genauere Schätzung des Hörverlustes und bietet sich daher als ergänzende oder alternative Methode an. Weitere Vorteile sind seitenselektive Erfassung des Hörverlustes und schnellere Messzeiten. In Hinblick auf die Sprachentwicklung sind bei Kindern Hörgeräte schon bei cochleären Hörverlusten ab 25 dBHL indiziert. Das bedeutet, dass DPOAE-Kochleogramme auch bei der Hörgeräteanpassung wertvolle Dienste leisten könne.

#### DPOAE-Kochleogramme



#### Qualitätssicherung

#### Subjektive Tests

Hörprüfraum: schallarm, maximal zulässiger Pegel des Störschalls in ISO 389-1

Audiometer: geeicht, messtechnische Überprüfung des Audiometres inclusive Hörer mindestens 2 x jährlich.

Durch Einhaltung der Standards (Tonschwelle: IEC 60645-1, ISO 389-1-2-3-4, ISO 8253-1; Sprachtests: IEC 60645-2, ISO 8253-3) ist gewährleistet, dass die an unterschiedlichen Orten erhobenen Daten vergleichbar sind.

#### Fehlerquellen:

Raum:

Störschall zu groß

Geräte:

Falsche Kalibrierung, schlechter Funktionszustand müssen vermieden werden. Untersucher:

1. Falsches Anbringen der elektroakustischen Wandler. Verwechselung der Seite kann bei einer anstehenden OP katastrophal sein! falsche Ankoppelung an den Gehörgang kann die Hörschwellenbestimmung erheblich verschlechtern!

2. Schlechte Instruktion des Patienten und schnelles Untersuchungstempo können ebenfalls zu falschen Ergebnissen führen.

3. Nicht-Beachten der Fühlschwelle bei der Knochenleitung und Nicht-Beachten der Vertäubungsregeln zur Vermeidung des Überhörens.

Patient: Patient kann simulieren oder aggravieren.

Tinnitus: Durch Verdeckung des Prüftons durch den Tinnitus kann es zu widersprüchlichen Ergebnissen zwischen den physiologischen und psychoakustischen Tests kommen.

#### Objektive Tests

OAE: Die bei der Registrierung der OAE verwendeten Ohrsonden sind sehr stoßempfindlich. Die wöchentliche Messung der Übertragungsfunktion in einem künstlichen Ohrsimulator (B&K 5147) ist zu empfehlen. Die Röhrchen der Ohrsonden können durch Cerumen verstopft werden. Es sollten daher Sonden zur Anwendung kommen, die zerlegbar und damit leicht zu reinigen sind. Über das Sondenkabel können Störsignale entstehen. Um dies zu vermeiden sollte die Sonde am Kabel über eine Vorrichtung über den Kopf des Patienten schwebend angebracht werden.

Um schwellenahe Messungen der OAE zu gewährleisten oder Messungen an Patienten mit größeren Hörverlusten durchführen zu können, müssen Mikrophone verwendet werden, die ein geringes Eigenrauschen haben. Nach 200 Mittelungen sollte das Störgeräusch bei den DPOAE im mittleren Frequenzbereich besser als -30 dB SPL sein.

Eine Überprüfung des Sondensitzes während der Messung sollte automatisch durch Messung der Schalldruckübertragungsfunktion erfolgen. Es sollten einheitliche Kriterien zur Sicherstellung der Validität der Antwort verwendet werden (z.B. Korrelation der beiden TEOAE-Aufnahmepuffer > 65%, Signal-Störabstand von > 6 dB bei den DPOAE).

AEP: s. Empfehlungen der ADANO zur Durchführung der ERA mit transienten Potentialen

#### Literatur:

Mrowinski und Mitarbeiter: Audiometrie, Thieme Verlag (Kleines Lehrbuch über audiometrische Standardverfahren der subjektiven Audiometrie. Objektive Audiometrie und Pädaudiologie wird nur kurz dargestellt.

Lehnhardt und Laszig (Hrsg): Praxis der Audiometrie (Großes Lehrbuch der subjektiven und objektiven Verfahren der Audiometrie). Thieme Verlag. Eine neue Ausgabe erfolgt in 2009.

Hoth und Lenarz: Elektrische Reaktions-Audiometrie, Springer Hoth und Lenarz: Otoakustische Emissionen, Thieme

#### Zu den OAE:

Janssen und Müller: Otoacoustic emissions as a diagnostic tool in a clinical context. In: Active processes and otoacoustic emissions. Springer Handbook of Auditory Research 30. Springer New York

#### Zu den ASSR:

Pethe J, Mühler R, von Specht H (2001). Zur Abhängigkeit der AMFR von der Vigilanz. HNO 49:188-193 Pethe J (2003) Amplitude modulation following responses (AMFR) – ein neuer Weg zur objektiven Bestimmung der Hörschwelle. HörBericht 72, Geers Hörakustik

Picton TW, Skinner CR, Champagne SC, Kellett AJC, Maistre A (1987) Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. J Acoust Soc Am 82: 165-178

Picton TW, Dimitrijevic A, John MS, von Roon P (2001) The use of phase in the detection of auditory steady-state responses. Clinical Neurophysiology 112: 1698-1711



Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

### Physikalische Grundlagen der Akustik

Prof.Dr.-Ing. Matthias Blau Institut für Hörtechnik und Audiologie FH Oldenburg/Ostfriesland/Wilhelmshaven

XXI. Winterschule der Deutschen Gesellschaft für Medizinische Physik Pichl 2009



## Übersicht

- Was ist Schall? Grundgrößen: p, v, .... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt ... Diffuses Schallfeld
- 1. Was ist Schall?
- 2. Grundgrößen: p,  $\vec{v}$ ,  $\vec{I}$ , P
- 3. Schallausbreitung 1
- 4. Schallausbreitung 2
- 5. Ebene Wellen
- 6. Kugelwellen
- 7. Echt dreidimensionale Wellen
- 8. Diffuses Schallfeld



1.

### Was ist Schall?

 Was ist Schall?

 Grundgrößen: p, v, ...

 Schallausbreitung 1

 Schallausbreitung 2

 Ebene Wellen

 Kugelwellen

 Echt...

Diffuses Schallfeld

- Schall: mechanische Schwingungen und Wellen in einem elastischen Medium
  - elastisches Medium: Gas, Flüssigkeit, Festkörper im weiteren: Luft
  - Hörschall: Beschränkung auf (16...20) Hz  $\leq f \leq (16...20)$  kHz

Institut für Hörtechnik + Audiologie IHA

Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

### 1. Was ist Schall?

 1D-Betrachtung: Fortpflanzung lokaler Druckschwankungen (Verdichtungen, Verdünnungen) als Longitudinalwelle

λ.

t=0			
t=0.05T			
t=0.1T			
t=0.15T			
t=0.2T			
t=0.25T			
t=0.3T	•		~ *
	$\mathbf{x} = \mathbf{x}$	)	- Λ





## Grundgrößen: p, $\vec{v}$ , $\vec{I}$ , P

- Was ist Schall? Grundgrößen: p, v, . Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...
- Diffuses Schallfeld
- Schallschnelle  $\vec{v}$ :
  - Wechselgeschwindigkeit, mit der die Luftteilchen schwingen
  - Richtungssinn = Ausbreitungsrichtung der Welle
  - $-1\mu$ m/s (sehr leise) ... einige cm/s (sehr laut)

### • Schalldruck *p*:

- dem statischen Luftdruck ( $\approx 100~{\rm kPa}$ ) überlagerte Druckschwankungen
- ungerichtet
- < 1 mPa (sehr leise) . . . einige 10 Pa (sehr laut)



Was ist Schall? Grundgrößen: p, v,.

Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

Schallausbreitung 1 Schallausbreitung 2

## **2.** Grundgrößen: p, $\vec{v}$ , $\vec{I}$ , P

- Schallleistung P:
  - reine Quellgröße, kennzeichnet Energiefluss der Quelle

$$-P\propto ilde{p}^2$$





Institut für rtechnik + Audiologie

Was ist Schall? Grundgrößen: p, v,...

Schallausbreitung 1

Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

### 3. Schallausbreitung 1

- Schallgeschwindigkeit *c*:
  - adiabatische Zustandsänderung  $c^2 \approx \kappa p_{-}/\rho_{-}$ oder  $c^2 \approx \kappa RT_{-}$ .

$$\begin{array}{c} - \ c \propto \sqrt{T} \\ c \mid_{0^{\circ}\mathsf{C}} = 331 \text{ m/s} \\ c \mid_{20^{\circ}\mathsf{C}} = 343 \text{ m/s} \\ - \ \lambda = c/f \end{array}$$





Was ist Schall? Grundgrößen: p, v,...

Ebene Wellen Kugelwellen Echt...

Schallausbreitung 1 Schallausbreitung 2

Diffuses Schallfeld

## 3. Schallausbreitung 1

 Reflexion, Beugung, Streuung: Die Wellenlänge ist das Maß aller Dinge

(REICHARDT)







Institut für echnik + Audiologie

Was ist Schall?

Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

Grundgrößen: p, v,... Schallausbreitung 1 Schallausbreitung 2

### 4. Schallausbreitung 2

• Wellengleichung (D'ALAMBERT, EULER, LAGRANGE 1747-1762) 1  $\partial^2 n(\vec{r} t)$ 

$$\Delta p(\vec{r},t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r},t)}{\partial t^2} = 0$$

- Laplace-Operator (in kartesischen Koordinaten)

$$\Delta(\vec{r} = x, y, z) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

- Voraussetzungen: adiabatische Zustandsänderung, keine Quellen im betrachteten Volumen,  $|\vec{v}| \ll c$
- Annahme sinusförmiger Zeitverläufe  $p(t) = \hat{p}\cos(\omega t + \varphi_p) = \Re\{\hat{p}(\omega) e^{j\varphi_p(\omega)} e^{j\omega t}\}$   $= \Re\{\hat{\underline{p}}(\omega) e^{j\omega t}\}$   $\Delta \underline{\hat{p}}(\vec{r}, \omega) + k^2 \underline{\hat{p}}(\vec{r}, \omega) = 0 \quad \text{Helmholtz-Gleichung}, k = \frac{\omega}{c}$ Seite 126



### 4. Schallausbreitung 2

Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

IHÁ

Was ist Schall? Grundgrößen: p, v,...

Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

- Beziehung zwischen Schalldruck und Schallschnelle
  - eindimensional: Luftvolumen Sdx wird durch Kraft Sp(x) Sp(x + dx) in x-Richtung beschleunigt,

$$S\left\{p(x) - p(x + dx)\right\} = \varrho_{-}Sdx\frac{dv_{x}}{dt}$$
  
:  
$$-\frac{\partial p}{\partial x} = \varrho_{-}\frac{\partial v_{x}}{\partial t}$$

- dreidimensional -grad  $p = \varrho_{-} \frac{\partial \vec{v}}{\partial t}$ 

$$-\operatorname{grad} \underline{\hat{p}}(\omega) = j\omega \varrho \underline{\hat{\vec{v}}}(\omega)$$
  
(sinusförmige Anregung)

## 5. Ebene Wellen

- Wo kommen ebene Wellen vor?
  - in Rohren bei tiefen Frequenzen (ka < 1.84)





- Lösung:  $\underline{\hat{p}}(x,\omega) = \underline{\hat{p}}_e(\omega) \, \mathbf{e}^{-\mathbf{j}kx} + \underline{\hat{p}}_r(\omega) \, \mathbf{e}^{+\mathbf{j}kx}$  $\underline{\hat{v}}_x(x,\omega) \underline{\varrho}_{-}c = \underline{\hat{p}}_e(\omega) \, \mathbf{e}^{-\mathbf{j}kx} - \underline{\hat{p}}_r(\omega) \, \mathbf{e}^{+\mathbf{j}kx}$ 
  - $\underline{\hat{p}_{e}}(\omega)$  und  $\underline{\hat{p}_{r}}(\omega)$  aus Randbedingungen



### 5. Ebene Wellen

• Beispiel: Rohr der Länge L, rechts mit spezifischer Impedanz  $Z_{\rm W}(\omega)$  abgeschlossen, links mit kolbenförmiger Schnelle  $\hat{v}_0(\omega)$  angeregt

$$\begin{split} \underline{\hat{p}}(x,\omega) &= \underline{\hat{p}}_{e}(\omega) \, \mathbf{e}^{-\mathbf{j}kx} + \underline{\hat{p}}_{r}(\omega) \, \mathbf{e}^{+\mathbf{j}kx} \\ \underline{\hat{v}}_{x}(x,\omega) \underline{\rho}_{-}c &= \underline{\hat{p}}_{e}(\omega) \, \mathbf{e}^{-\mathbf{j}kx} - \underline{\hat{p}}_{r}(\omega) \, \mathbf{e}^{+\mathbf{j}kx} \\ Z_{\mathsf{W}}(\omega) &= \frac{\underline{\hat{p}}(x=0,\omega)}{\underline{\hat{v}}_{x}(x=0,\omega)} = \varrho_{-}c \, \frac{\underline{\hat{p}}_{e}(\omega) + \underline{\hat{p}}_{r}(\omega)}{\underline{\hat{p}}_{e}(\omega) - \underline{\hat{p}}_{r}(\omega)} \\ \hat{v}_{0}(\omega) &= \underline{\hat{v}}_{x}(x=-L,\omega) = \frac{1}{\varrho_{-}c} \bigg\{ \underline{\hat{p}}_{e}(\omega) \, \mathbf{e}^{+\mathbf{j}kL} - \underline{\hat{p}}_{r}(\omega) \, \mathbf{e}^{-\mathbf{j}kL} \bigg\} \\ \mathbf{f} \ddot{\mathsf{u}}\mathsf{hrt} \, \mathsf{zu} \quad \underline{\hat{p}}_{r}(\omega) &= \underline{\hat{p}}_{e}(\omega) \frac{Z_{\mathsf{W}}(\omega) - \varrho_{-}c}{Z_{\mathsf{W}}(\omega) + \varrho_{-}c} = \underline{\hat{p}}_{e}(\omega) \underline{r} \\ \underline{\hat{p}}_{e}(\omega) &= \frac{\hat{v}_{0}(\omega)\varrho_{-}c}{\mathbf{e}^{+\mathbf{j}kL} - \underline{r}\mathbf{e}^{-\mathbf{j}kL}} \quad \underline{r} \dots \text{Reflexionsfaktor} \end{split}$$

### 5. Ebene Wellen



Hörtechnik

Was ist Schall?

Grundgrößen:  $p, \vec{v}, \dots$ Schallausbreitung 1

Schallausbreitung 2

Diffuses Schallfeld

Ebene Wellen Kugelwellen

Echt...

IHA





Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

### 6. Kugelwellen

- Wo kommen Kugelwellen vor?
  - im Fernfeld von Strahlern, die klein gegen  $\lambda$  sind





$$\Delta(r,\alpha,\vartheta) = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2}{\partial \alpha^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \vartheta^2} + \frac{1}{r^2} \cot \vartheta \frac{\partial}{\partial \vartheta}$$

• allgemeine Lösung der Wellengleichung für Abstrahlung ins freie Schallfeld:

$$\underline{\hat{p}}(r,\alpha,\vartheta,\omega) \propto \frac{1}{r}\,\Gamma(\alpha,\vartheta)\,\mathbf{e}^{-\mathbf{j}kr}$$

Institut für hnik + Audiologie

### 6. Kugelwellen

Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt ... Diffuses Schallfeld • Monopol (Punktstrahler)



– keine Abhängigkeit von der Richtung, nur von der Entfernung zu Quelle – 1D-Betrachtung in r-Richtung

$$\Delta p = \frac{\partial^2 p}{\partial r^2} + \frac{2}{r} \frac{\partial p}{\partial r} = \frac{1}{r} \frac{\partial^2 (pr)}{\partial r^2}$$



(z.B. Schlauchenden)

# IHÁ

Was ist Schall?

Ebene Wellen

Kugelwellen Echt...

#### Kugelwellen **6**.









Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

## 6. Kugelwellen

• Strahlerzeile (cont'd)

- Schalldruck  

$$\underline{\hat{p}}(r,\vartheta,\omega) \Big|_{2\pi r \gg \lambda} = \frac{\mathbf{j}\omega \,\varrho}{4\pi r} \underline{\hat{q}} N \Gamma(\vartheta,\omega) \, \mathrm{e}^{-\mathbf{j}k \left(r + \frac{N-1}{2}d\cos\vartheta\right)}$$

- Richtfaktor

$$\Gamma(\vartheta,\omega)|_{2\pi r \gg \lambda} = \frac{1}{N} \frac{\sin\left[(N\pi \, d/\lambda)\cos\vartheta\right]}{\sin\left[(\pi \, d/\lambda)\cos\vartheta\right]}$$



Was ist Schall?			
Grundgrößen: $p, \vec{v},$			
Schallausbreitung 1			
Schallausbreitung 2			
Ebene Wellen			
Kugelwellen			
Echt			
Diffuses Schallfeld			

### 6. Kugelwellen

- Strahlerzeile (cont'd)
  - Beispiele: gerichtete Beschallung in Kirchen, auf Bahnsteigen; gleiche Prinzipiem gelten reziprok auch f
    ür Richtmikrofone!



(Quelle: Microtech Gefell)



- Richtfaktor

$$\Gamma(\vartheta,\omega)\big|_{2\pi r\gg\lambda} = 2 \frac{J_1\left[(2\pi a/\lambda)\sin\vartheta\right]}{(2\pi a/\lambda)\sin\vartheta}$$



Was ist Schall? Grundgrößen:  $p, \vec{v}, \ldots$ Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt. Diffuses Schallfeld

#### Echt dreidimensionale Wellen 7.

- Lösung der Wellengleichung durch Produktansatz und Koordinatensystem-angepasste Randbedingungen (lokal reagierende Ränder)
- z.B. Eigenschwingungen im Quaderraum mit schallharten Wänden
  - Helmholtz-Gleichung

$$\frac{\partial^2 \underline{\hat{p}}}{\partial x^2} + \frac{\partial^2 \underline{\hat{p}}}{\partial y^2} + \frac{\partial^2 \underline{\hat{p}}}{\partial z^2} + k^2 \underline{\hat{p}} = 0$$

- Randbedingungen: Normalenschnelle den an Wänden gleich Null

$$\frac{\partial \underline{\hat{p}}}{\partial x}\Big|_{x=0,L_x} = 0, \quad \frac{\partial \underline{\hat{p}}}{\partial y}\Big|_{y=0,L_y} = 0, \quad \frac{\partial \underline{\hat{p}}}{\partial z}\Big|_{z=0,L_z} = 0.$$

#### Echt dreidimensionale Wellen 7.

Produktansatz

$$\underline{\hat{p}}(x,y,z) = \underline{\hat{p}}_x(x) \cdot \underline{\hat{p}}_y(y) \cdot \underline{\hat{p}}_z(z)$$

führt zu

$$\frac{1}{\underline{\hat{p}}_x} \frac{\partial \underline{p}_x}{\partial x^2}$$

1 බ2 ක

unabhängig  $= -k_x^2 = -k_y^2 = -k_z^2$ 

unabhängig von x, y, z von x, y, z von x, y, z

 $+ \frac{1}{\underline{\hat{p}}_{y}} \frac{\partial^{2} \underline{\hat{p}}_{y}}{\partial y^{2}} + \frac{1}{\underline{\hat{p}}_{z}} \frac{\partial^{2} \underline{\hat{p}}_{z}}{\partial z^{2}} + k^{2} = 0$ unabhängig

Was ist Schall? Grundgrößen:  $p, \vec{v}, \ldots$ Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen

IHO

Diffuses Schallfeld



Was ist Schall? Grundgrößen: p, v, .. Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt...

Diffuses Schallfeld

## 7. Echt dreidimensionale Wellen

- damit 3 gewöhnliche Differenzialgleichungen

$$\begin{aligned} \frac{\partial^2 \underline{\hat{p}}_x}{\partial x^2} + k_x^2 \underline{\hat{p}}_x &= 0 \quad \rightarrow \quad \underline{\hat{p}}_x = A \mathbf{e}^{-\mathbf{j}k_x x} + B \mathbf{e}^{+\mathbf{j}k_x x} \\ \frac{\partial^2 \underline{\hat{p}}_y}{\partial y^2} + k_y^2 \underline{\hat{p}}_y &= 0 \quad \rightarrow \quad \underline{\hat{p}}_y = C \mathbf{e}^{-\mathbf{j}k_y y} + D \mathbf{e}^{+\mathbf{j}k_y y} \\ \frac{\partial^2 \underline{\hat{p}}_z}{\partial z^2} + k_z^2 \underline{\hat{p}}_z &= 0 \quad \rightarrow \quad \underline{\hat{p}}_z = E \mathbf{e}^{-\mathbf{j}k_z z} + F \mathbf{e}^{+\mathbf{j}k_z z} \end{aligned}$$

#### insgesamt:

$$\underline{\hat{p}}(x, y, z) = \left(A \mathbf{e}^{-\mathbf{j}k_x x} + B \mathbf{e}^{+\mathbf{j}k_x x}\right) \left(C \mathbf{e}^{-\mathbf{j}k_y y} + D \mathbf{e}^{+\mathbf{j}k_y y}\right) \left(E \mathbf{e}^{-\mathbf{j}k_z z} + F \mathbf{e}^{+\mathbf{j}k_z z}\right)$$

mit  $k_x^2 + k_y^2 + k_z^2 = k^2$ .

#### Institut für ik + Audiologie

## 7. Echt dreidimensionale Wellen

- Randbedingungen:  $\begin{aligned}
\frac{\partial \hat{p}}{\partial x} \Big|_{x=0,L_x} &= 0 \\
\left(-jk_x A e^{-jk_x x} + jk_x B e^{+jk_x x}\right) \Big|_{x=0,L_x} \hat{\underline{p}}_y(y) \hat{\underline{p}}_z(z) &= 0 \\
x &= 0: \qquad A = B \\
x &= L_x: \qquad e^{-jk_x L_x} - e^{+jk_x L_x} &= 0 \\
-2j \sin(k_x L_x) &= 0 \\
k_x &= (n_x \pi)/L_x \\
analog: \qquad C = D \\
k_y &= (n_y \pi)/L_y \\
E &= F \\
k_z &= (n_z \pi)/L_z
\end{aligned}$ 

Was ist Schall? Grundgrößen: p, v, .. Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen

Diffuses Schallfeld



Was ist Schall? Grundgrößen: p, v, . Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen

Diffuses Schallfeld

7. Echt dreidimensionale Wellen

– aus  $k_x^2 + k_y^2 + k_z^2 = k^2 = \omega^2/c^2$  folgen die Eigenfrequenzen

$$f_n = \frac{c}{2} \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2}$$

- bei den Eigenfrequenzen ergibt sich für  $\underline{\hat{p}}_n(x,y,z)$ 

$$\underline{\hat{p}}_{n}(x, y, z, f_{n}) = \operatorname{const} \cdot \cos\left(\frac{n_{x}\pi}{L_{x}}x\right) \cos\left(\frac{n_{y}\pi}{L_{y}}y\right) \cos\left(\frac{n_{z}\pi}{L_{z}}z\right) \operatorname{Das} \operatorname{Produkt} \left\{\cos\left(\cdot\right)\cos\left(\cdot\right)\cos\left(\cdot\right)\right\} \text{ heißt zu}$$

Das Produkt  $\left\{ \cos(\cdot) \cos(\cdot) \cos(\cdot) \right\}$  heißt zur *n* ten Eigenfrequenz zugehörige Eigenfunktion  $\Psi_n(x, y, z)$ 



Was ist Schall? Grundgrößen: p, v,... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

## 7. Echt dreidimensionale Wellen

• erzwungene Schwingungen durch Punktstrahler  $(\hat{q}_{s}(\omega) \text{ an der Stelle } x_{s}, y_{s}, z_{s})$ , kleine Dämpfung

$$\frac{\underline{\hat{p}}(x,y,z,\omega)}{\underline{\hat{q}}_{\mathsf{s}}(\omega)} = -\frac{\mathsf{j}\omega\varrho_{\_}c^2}{V}\sum_{n}\frac{\Psi_n(x,y,z)\Psi_{\mathsf{s},n}(x_{\mathsf{s}},y_{\mathsf{s}},z_{\mathsf{s}})}{\omega^2 - \omega_n^2 + \delta_n^2}$$

Summe von "Moden" mit Eigenfrequenzen  $\omega_n$ (inkl. Dämpfung  $\delta_n$ ) und dazugehörigen Empfängerund Sender-Eigenfunktionen  $\Psi_n(x, y, z)$  und  $\Psi_n(x, y, z)$ 

In der Praxis werden die  $\omega_n$ ,  $\delta_n$ ,  $\Psi_n(x, y, z)$  und  $\Psi_n(x, y, z)$  durch Messung von  $\underline{\hat{p}}(x, y, z, \omega)/\underline{\hat{q}}_{s}(\omega)$  bestimmt (experimentelle Modalanalyse)



#### Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

## 7. Echt dreidimensionale Wellen

• Modendichte in Räumen

$$\frac{\mathrm{d}N}{\mathrm{d}f}\approx 4\pi V\frac{f^2}{c^3}$$

z.B. typischer Wohnraum ( $V = 40 \text{ m}^3$ ): 0.12 Moden/Hz @ f = 100 Hz12 Moden/Hz @ f = 1 kHz



## 8. Diffuses Schallfeld

Was ist Schall?		
Grundgrößen: $p, \vec{v}, \ldots$		
Schallausbreitung 1		
Schallausbreitung 2		
Ebene Wellen		
Kugelwellen		
Echt		
Diffuses Schallfeld		

• ÜF des Schalldrucks in typischem Wohnraum



- Idee: statistische Betrachtung, d.h. Mittelung (über Ort, Zeit, Frequenz)
  - $\rightarrow$  Schallenergiedichte  $w = dW/dV = \overline{\tilde{p_B}^2}/(\varrho_c^2)$
- ideal diffuses Schallfeld: 1.) w an allen Punkten gleich groß, 2.) an jedem Punkt alle Schalleinfallsrichtungen gleich wahrscheinlich
- Grundgleichung  $V \frac{dw}{dt} + \frac{cA}{4}w = P$



Was ist Schall? Grundgrößen: p, v,...

Schallausbreitung 1

Schallausbreitung 2 Ebene Wellen

Diffuses Schallfeld

Kugelwellen

Echt..

- Anwendbarkeit:
  - Mittelung über ausreichend viele Moden (Betrachtung in Frequenzbändern mit ausreichen hoher Modendichte)
    - ightarrow untere Grenzfrequenz für Terzen
      - $f_{\sf g}=700\,{\sf Hz}\big/\sqrt[3]{V/{\sf m}^3}$  (20 Moden pro Terz)
    - $\rightarrow$  alternativ:  $f_{\rm g} = 2000 \, {\rm Hz} \frac{T/{\rm s}}{V/{\rm m}^3}$ 
      - (3 Moden pro mittlerer Moden-Bandbreite)
  - genügend weit weg von Quellen und Wänden
  - (möglichst zufällige Verteilung der Wandimpedanz)

Institut für chnik + Audiologie

Was ist Schall? Grundgrößen: p, v,.. Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen

Echt.

Diffuses Schallfeld

### 8. Diffuses Schallfeld

• Spezialfall zeitlich stationäre Anregung (dw/dt = 0)

$$\frac{A}{4} \frac{\overline{\tilde{p_{\mathsf{B}}}^2}}{\varrho_{\_}c} = P$$

-A äquivalente Schallabsorptionsfläche

$$A = \sum_{i} \alpha_{i} S_{i}$$
 (SABINE)  
$$A = -S \ln \left( 1 - \frac{1}{S} \sum_{i} \alpha_{i} S_{i} \right)$$
 (Eyring)

- $S_i$ ,  $\alpha_i$  Flächeninhalt und Absorptionsgrad der *i*ten Teilfläche
- $S = \sum_i S_i$  Gesamtfläche
- als Pegel ausgedrückt

$$L_W = L_{p,\mathsf{m}} + 10 \, \lg \, \frac{A}{4 \, \mathsf{m}^2} \, \mathsf{dB}$$



Was ist Schall? Grundgrößen: p, v, ... Schallausbreitung 1 Schallausbreitung 2 Ebene Wellen Kugelwellen Echt... Diffuses Schallfeld

### 8. Diffuses Schallfeld

• Spezialfall Abschalten einer stationären Quelle  $V\frac{\mathrm{d}w}{\mathrm{d}t} + \frac{cA}{4}w = 0$ - Lösung:  $w(t) = w_0 e^{-t/\tau} \text{ mit } \tau = 4V/(cA)$ – als Pegel ausgedrückt  $L_{p,\mathsf{m}} = L_0 - \frac{2.5cA}{V \ln 10} t \,\mathsf{dB}$  $L_{\rho}/dB$   $\wedge$ Nachhallzeit T 0  $T = 24 \ln 10 \frac{V}{cA}$  $= \frac{55.3}{c} \frac{V}{A}$ -20 -60 dB -40 -60 Ò 1 2 3 t/s



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionen

### **Akustische Messverfahren**

Prof.Dr.-Ing. Matthias Blau Institut für Hörtechnik und Audiologie FH Oldenburg/Ostfriesland/Wilhelmshaven

XXI. Winterschule der Deutschen Gesellschaft für Medizinische Physik Pichl 2009



## Übersicht

- Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionen
- 1. Das Wesen des Messens
- 2. Instrumentierung
- 3. Schallpegel
- 4. Spektren
- 5. Übertragungsfunktionen



### 1. Das Wesen des Messens

Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktioner

- Messen bedeutet Parametrisieren von Modellen
- im Zusammenhang mit akustischen Messungen häufig verwendete Modelle:
  - a) wenn die Anregung nicht kontrolliert werden kann: stationäre Zufallsprozesse
    - $\rightarrow$  aufgenommener Schalldruck ist Stichprobe eines bandbegrenzten  $G\rm{AUSS}\mathchar`-Prozesses$

$$p \sim \mathcal{N}(0, \widetilde{p}^2)$$
 mit  $\widetilde{p}^2 = \int_{-\infty}^{+\infty} \Gamma_{pp}(f) \, \mathrm{d}f$ 

- $\rightarrow$  Messaufgabe: schätze  $\tilde{p}^2$  oder  $\Gamma_{pp}(f)$
- b) wenn die Anregung kontrolliert werden kann: LTI-System
  - $\rightarrow$  Messaufgabe: schätze H(f) oder h(t)
- wenn Bedingung b) erfüllt ist, immer Übertragungsfunktionen / Impulsantworten messen!



### 2. Instrumentierung

• Schalldruck *p*: Mikrofone



- bei Messmikrofonen unterschiedliche Entzerrungen (Druck vs. Freifeld vs. Diffusfeld) beachten
- Signalkonditionierung: integrierte Vorverstärker, externe Speiseteile (Polarisationsspannung bei Messmikrofonen!)
- Kalibrierung: in der Regel in kleinen abgeschlossenen Volumina

Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktioner



### 2. Instrumentierung

- Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionen
- Schallschnelle  $\vec{v}$ : Mikrofonpaare oder Hitzdraht-Anemometrie





- Signalkonditionierung: siehe Messmikrofone bzw.
   spezielle Speisung bei Hitzdrahtsonden
- Kalibrierung: in speziellen Kupplern (Mikrofonpaare) bzw. in Rohren oder im Freifeld



Das Wesen des...

Instrumentierung Wer misst, misst Mist

Übertragungsfunktionen

Schallpegel Spektren

### 2. Instrumentierung

• Schwingbeschleunigung  $\vec{a}$  bzw. Schwingschnelle  $\vec{v}$ : Beschleunigungsaufnehmer bzw. Laser-Vibrometrie





- Signalkonditionierung: Ladungsverstärker bzw.
   Laser-Controller
- Kalibrierung: Kalibrierschwingtische, ggf. mit Kalibriermassen
  - Seite 143


Das Wesen des.

Schallpegel Spektren

Wer misst, misst Mist

Übertragungsfunktioner

## 2. Instrumentierung

- Signalanalyse
  - für Standardaufgaben entsprechende Kiste kaufen
  - für Forschung/Entwicklung PC-basierte Systeme (offen, flexibel), z.B. PureMeasurement

(www.hoertechnik-audiologie.de/downloads)



Institut für lörtechnik + Audiologie

Übertragungsfunktioner

Das Wesen des.. Instrumentierung

Schallpegel

Spektren

- 3. Wer misst, misst Mist
- empfohlene Strategie: AAA

• Anhören:

Rauschen?, Verzerrungen?, Brummen?, ....?

• Anschauen: Oszillogramme, Spektrogramme,



• Analysieren: Kohärenz, ...



Das Wesen des. Instrumentierung

Schallpegel Spektren

Wer misst, misst Mist

Übertragungsfunktioner

#### **Schallpegel** 4.

- Definition Schalldruckpegel  $L_p = 10 \lg (\tilde{p}^2/p_0^2) dB$
- $\tilde{p}^2$ : p(t) quadrieren, zeitlich mitteln (integrieren)

$$\tilde{p}^2(t_x) = \frac{1}{\tau} \int_{t=-\infty}^{t_x} p^2(t) \, \mathbf{e}^{-(t_x-t)/\tau} \, \mathrm{d}t$$



$$\tau_{\rm SLOW} = 1$$
s,  $\tau_{\rm FAST} = 125$ m

IHÁ

#### **Schallpegel** 4.

- zeitliche Mittelung
  - -äquivalenter Dauerschallpegel  $L_{eq}$

$$L_{\rm eq} = 10 \log \left( \frac{1}{T_{\rm r}} \sum_{i} T_i \operatorname{Mittelwert}_{t_x \text{ in } T_i} \left\{ \frac{\tilde{p}^2(t_x)}{p_0^2} \right\} \right) {\rm dB}$$

oder

$$L_{\rm eq} = 10 \log \left( \frac{1}{T_{\rm r}} \int_{t=0}^{T_{\rm r}} p^2(t) \,\mathrm{d}t \right) \left/ p_0^2 \,\mathrm{dB} \right.$$

- Taktmaximal-Mittelungspegel  $L_{T,m}$  oder  $L_{T,eq}$ 

$$L_{\rm T,m} = 10 \log \left( \frac{1}{T_{\rm r}} \sum_{i} T_{i} \max_{t_{x} \text{ in } T_{i}} \left\{ \frac{\tilde{p}^{2}(t_{x})}{p_{0}^{2}} \right\} \right) \text{dB}$$
  
Seite 145

Das Wesen des.. Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionen



Seite 146



### 5. Spektren

- Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktioner
- Spektren über Filter (cont'd)

- Statistik von Effektivwertschätzern

- \* Modell:  $p_{\mathsf{B}} \sim \mathcal{N}(0, \tilde{p_{\mathsf{B}}}^2)$
- \* Schätzer 1 (zeitkontinuierlich, Rechteckfenster):

 $\widehat{\tilde{p}}_{\mathsf{B}}^{2} = \frac{1}{T} \int_{T} p_{\mathsf{B}}^{2}(t) \, \mathsf{d}t$   $\mathsf{E}\left\{\widehat{\tilde{p}}_{\mathsf{B}}^{2}\right\} = \widetilde{p}_{\mathsf{B}}^{2}, \quad \mathsf{Var}\left\{\widehat{\tilde{p}}_{\mathsf{B}}^{2}\right\} = \frac{\widetilde{p}_{\mathsf{B}}^{4}}{BT}$ \* Schätzer 2 (zeitdiskret, Rechteckfenster):  $\widehat{\tilde{p}_{\mathsf{B}}^{2}} = \frac{1}{T} \sum_{n}^{N} p_{\mathsf{B}}^{2}(n/F)$ 

$$\mathsf{E}\left\{\widehat{\tilde{p}_{\mathsf{B},N}^{2}}\right\} = \widetilde{p}_{\mathsf{B}}^{2}, \quad \mathsf{Var}\left\{\widehat{\tilde{p}_{\mathsf{B},N}^{2}}\right\} = \frac{1}{N}\frac{\tilde{p}_{\mathsf{B}}^{4}}{B/F_{\mathsf{s}}}$$

Institut für echnik + Audiologie

Das Wesen des.. Instrumentierung

Schallpegel Spektren

Wer misst, misst Mist

Übertragungsfunktioner

### 5. Spektren

• Spektren über Filter (cont'd)

\* Schätzer 3 (zeitdiskret, Exponentialfenster):

$$\widehat{\tilde{p}_{\mathsf{B},\alpha}^2}(n) = \alpha p_\mathsf{B}^2(n) + (1-\alpha)\widehat{\tilde{p}_{\mathsf{B},\alpha}^2}(n-1)$$

äquivalente Zeitkonstante  $\tau = -\frac{1}{F}\ln(1-\alpha)$ 

$$\mathsf{E}\left\{\widehat{\tilde{p}_{\mathsf{B},\alpha}^{2}}\right\} = \tilde{p}_{\mathsf{B}}^{2}, \quad \mathsf{Var}\left\{\widehat{\tilde{p}_{\mathsf{B},\alpha}^{2}}\right\} = \frac{\alpha}{2-\alpha}\frac{\tilde{p}_{\mathsf{B}}^{4}}{B/F_{\mathsf{s}}}$$

- bei allen Schätzern ist die Varianz umgekehrt proportional zu Filterbandbreite  ${\cal B}$  und der jeweiligen Mittelungszeit
- daher bei Terz-/Oktavfiltern besonders bei niedrigen Mittenfrequenzen hohe Varianz aufgrund kleiner Filterbandbreite

Seite 147



Das Wesen des.

Instrumentierung Wer misst, misst Mist

Übertragungsfunktioner

Schallpegel Spektren

### 5. Spektren

- Spektren über FFT
  - Modell:  $p \sim \mathcal{N}(0, \tilde{p}^2), \quad \tilde{p}^2 = \int_{-\infty}^{+\infty} \Gamma_{pp}(f) \, \mathrm{d}f$
  - Messaufgabe: Schätzen der Leistungsdichte

$$\Gamma_{pp}(f) = \mathcal{F}\left\{\gamma_{pp}(\tau)\right\} = \lim_{T \to \infty} \frac{1}{T} \mathsf{E}\left\{P_T^*(f) P_T(f)\right\}$$

- Realisierung: Methode nach WELCH (1967)

$$\begin{split} \widehat{S}_{pp}(f = kF_{\rm s}/M) &= \frac{1}{R} \sum_{r=1}^{R} \frac{1}{MU_M} P_M^*(k,r) P_M(k,r) \\ P_M(k) &= {\rm FFT} \big\{ p_M(m) w_M(m) \big\} \\ MU_M &= \sum_{m=1}^{M} w_M^2(m) \quad {\rm Fensterenergie} \end{split}$$





Das Wesen des.. Instrumentierung

Schallpegel Spektren

Wer misst, misst Mist

Übertragungsfunktioner

#### 5. **Spektren**

- Spektren über FFT (cont'd)
  - Vergleich des WELCH-Schätzers mit wahrer Leistungsdichte
    - \* Zeitabtastung  $\rightarrow$  *aliasing* im Frequenzbereich Lösung: analoge Anti-Alias-Filter
    - \* M endlich  $\rightarrow$  *leakage* im Frequenzbereich Lösung: HANN-Fenster mit  $M > 4F_s/B_r$
    - \* Mittelung über endlich viele (R) Blöcke  $\rightarrow$  zufällige Schwankungen des Schätzers Lösung: R groß genug, überlappende Blöcke

$$\begin{split} & \mathsf{Var}\big\{\widehat{S}_{pp}(k)\big\}\Big|_{\text{kein overlap}} \approx \frac{1}{R}S_{pp}^2(k) \\ & \mathsf{Var}\big\{\widehat{S}_{pp}(k)\big\}\Big|_{\text{50\% overlap}} \approx \frac{11}{9R}S_{pp}^2(k) \end{split}$$

IHA

Das Wesen des... Instrumentierung

Schallpegel

Spektren

Wer misst, misst Mist

Übertragungsfunktione

#### 5. **Spektren**

- Spektren über FFT (cont'd)
  - Nachbemerkung zu Skalierungsfaktoren

anala

analog:  

$$\tilde{p}^{2} = \int_{-\infty}^{+\infty} \Gamma_{pp}(f) \, \mathrm{d}f$$

$$\downarrow \text{Zeitabtastung}$$

$$\left(S_{pp} = F_{s}\Gamma_{pp}\right)$$
zeitdiskret:
$$\frac{1}{N} \sum_{n=1}^{N} p^{2}\left(\frac{n}{F_{s}}\right) = \frac{1}{F_{s}} \int_{-\frac{1}{2}}^{+\frac{1}{2}} S_{pp}\left(\frac{f}{F_{s}}\right) \, \mathrm{d}\left(\frac{f}{F_{s}}\right)$$

$$\downarrow \text{Frequenzabtastung}$$

 $=\frac{F_{\rm s}/M}{F_{\rm s}}\sum_{k=1}^{M}S_{pp}(k)$ 





Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktioner

# 6. Übertragungsfunktionen

- punktuellen Schätzung deterministische Testsignale
  - Testsignale: Sinus, stepped sine, Multiton, Sweeps, Maximalfolgen
  - Schätzer:

a) 
$$\widehat{H}_{xy}(f) = \frac{\sum_{r=1}^{R} Y_M(f,r)}{\sum_{r=1}^{R} X_M(f,r)}$$
  
b)  $\widehat{H}_{xy}(f) = \frac{1}{R} \sum_{r=1}^{R} \frac{Y_M(f,r)}{X_M(f,r)}$ 

– Messrauschen führt nur zu zufälligen Fehlern



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktione

# 6. Übertragungsfunktionen

- mittlere Schätzung zufällige Testsignale
  - Testsignale: weißes oder gefärbtes Rauschen, Sprache, Musik
  - Modell: Messrauschen, Nichlinearitäten und Zeitvarianzen werden als unkorrelierte Rauschvorgänge auf Eingang und Ausgang betrachtet

 $\rightarrow$  für jede betrachtete Frequenz f gilt

$$\begin{pmatrix} \Gamma_{yy} & \Gamma_{xy}^* \\ \Gamma_{xy} & \Gamma_{xx} \end{pmatrix} \begin{pmatrix} 1 \\ -H_{xy} \end{pmatrix} = \begin{pmatrix} \Gamma_{nn} & 0 \\ 0 & \Gamma_{mm} \end{pmatrix} \begin{pmatrix} 1 \\ -H_{xy} \end{pmatrix}$$

 verschiedene Schätzer sind jeweils optimal f
ür bestimmte Annahmen zu den Rauschvorg
ängen auf Eingang und Ausgang



Hörtechnik + Audiologie

6. Übertragungsfunktionen

Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionen • mittlere Schätzung - zufällige Testsignale (cont'd)

- $-H_2$ -Schätzer
  - \* Annahme:  $\Gamma_{nn} = 0$

(kein Rauschen am Ausgang)

\* Schätzer:

$$H_{xy} = \frac{\Gamma_{yy}}{\Gamma_{xy}^*}$$

\* was wenn ...?  $\rightarrow$  Überschätzung



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktione

## 6. Übertragungsfunktionen

- mittlere Schätzung zufällige Testsignale (cont'd)
  - $-H_3$ -Schätzer
    - \* Annahme:  $\Gamma_{xx}/\Gamma_{mm} = \Gamma_{yy}/\Gamma_{nn}$ (gleiche SNRs an Ein- und Ausgang)

\* Schätzer:

$$H_{xy} = \sqrt{\frac{\Gamma_{xy}}{\Gamma_{xx}}\frac{\Gamma_{yy}}{\Gamma_{xy}^*}}$$

\* was wenn . . . ?  $\rightarrow$  Unter- oder Überschätzung



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktionei

# 6. Übertragungsfunktionen

- mittlere Schätzung zufällige Testsignale (cont'd)
  - $-H_V$ -Schätzer

\* Annahme:  $\Gamma_{mm} = \Gamma_{nn}$  (Rauschen auf Eingang und Ausgang gleich groß, aber unkorreliert)

\* Schätzer:

$$H_{xy} = \frac{1}{2\Gamma_{xy}} \left( \Gamma_{yy} - \Gamma_{xx} + \sqrt{\left(\Gamma_{yy} - \Gamma_{xx}\right)^2 + 4\left|\Gamma_{xy}\right|^2} \right)$$

\* was wenn . . . ?  $\rightarrow$  Unter- oder Überschätzung



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktione

## 6. Übertragungsfunktionen

- mittlere Schätzung zufällige Testsignale (cont'd)
  - Problem zeitliche Verzögerung zwischen Eingangsund Ausgangssignal (z.B. durch Schallausbreitung)
     \* führt zu systematischen Fehlern

Kreuzleistungsdichte  $\widehat{S}_{xy} \approx \left(1 - \frac{\tau}{M/F_s}\right) S_{xy}$ Übertragungsfunktion  $\widehat{H}_{xy} \approx \left(1 - \frac{\tau}{M/F_s}\right) H_{xy}$ Kohärenz  $\widehat{\gamma}_{xy}^2 \approx \left(1 - \frac{\tau}{M/F_s}\right)^2 \gamma_{xy}^2$ 

\* Abhilfe: a) Zeitfenster lang genug b) Verzögerung um au im gemessenen x(t) kompensieren

Institut für lörtechnik + Audiologie

Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktion

# 6. Übertragungsfunktionen

- mittlere Schätzung zufällige Testsignale (cont'd)
  - zufällige Fehler
    - \* H-Schätzer beruhen auf Schätzern von Leistungsdichten, die ihrerseits zufällige Fehler aufweisenm, siehe Abschnitt 5
    - \* aber: zufällige Fehler der Leistunsdichteschätzer sind häufig stark korreliert, da sie in der Regel aus *einem* Prozess hervorgegangen sind
       \* wichtige Varianzformeln:

$$\begin{aligned} & \mathsf{Var}\left\{\left|\widehat{H}_{xy}\right|\right\} \approx \frac{1 - \gamma_{xy}^2}{2R\gamma_{xy}^2} \left|H_{xy}\right|^2 \\ & \mathsf{Var}\left\{\mathsf{arg}\left(\widehat{H}_{xy}\right)\right\} \approx \frac{1 - \gamma_{xy}^2}{2R\gamma_{xy}^2} \mathsf{rad}^2 \end{aligned}$$



Das Wesen des... Instrumentierung Wer misst, misst Mist Schallpegel Spektren Übertragungsfunktione

# 6. Übertragungsfunktionen

- praktische Vorgehensweise
  - 1. AAA
  - 2. mit Rauschen beginnen
  - 3. keine Übertragungsfunktion ohne Kohärenz!

 $\gamma_{xy}^2 = |\Gamma_{xy}|^2 / (\Gamma_{xx}\Gamma_{yy})$  ist wichtiger Indikator für viele Fehler (schlechtes SNR, leakage Nichtlinearitäten, Zeitvarianzen)

Achtung: bei periodischen Anregungssignalen reagiert  $\gamma_{xy}^2$  nicht auf Übersteuerungen!

- 4. Vergleich mit Multitonsignal, Fensterlänge optimieren (ggf. x(t) verzögern)
- 5. Entscheidung für ein Anregungssignal, Anregungspegel optimieren

### Signalverarbeitung für Hörhilfen und Audiologie

Norbert Dillier / Zürich

#### Einführung

Ausgehend von der Feststellung und diagnostischen Eingrenzung einer Hörstörung schliesst sich wenn möglich eine Behandlung dieser Hörstörung durch medizinische und technische Mittel an. Medizinische Behandlungsmöglichkeiten beinhalten medikamentöse Therapien und chirurgische Eingriffe. Technische Lösungen bestehen aus der Anpassung von Hörinstrumenten bzw. implantierbaren Systemen (Mittelohr-, Cochlea- und Hirnstammimplantate). In diesem Beitrag beschränken wir uns auf die technisch-apparativen Rehalibitationsmöglichkeiten und dabei insbesondere auf Aspekte der Signalverarbeitung.

Ziele der Signalverarbeitung für Hörinstrumente, Hörhilfen, Sprachprozessoren oder Kommunikationsgeräte ist es, die zur sprachlichen Verständigung erforderlichen Schallsignale für die hörgeschädigte Person wahrnehmbar ("Hören") und unterscheidbar ("Verstehen") zu machen und wenn möglich einem natürlichen Klangempfinden ("Hörqualität") anzunähern.

Traditionelle Hörgeräte mussten sich aufgrund der begrenzten Möglichkeiten der Analogschaltungstechnik auf einfache Verstärkung, Filterung und Pegelbegrenzung beschränken. Moderne Systeme mit digitaler Signalverarbeitung erlauben vielfältigere und neuartige Transformationen der Schallsignale, sodass vielleicht in nicht allzu ferner Zukunft die Vision eines "Verstehgerätes" im Unterschied zu den bisherigen "Hörgeräten" erfüllt sein könnte.

Die Algorithmen und Verfahren der digitalen Signalverarbeitung können in zwei grosse Klassen unterteilt werden. Die erste Gruppe ist relativ eng an die konkreten Ausprägungen eines Hörschadens gekoppelt und versucht, aufgrund möglichst genauer individueller Angaben über Art und Ausmass dieses Hörschadens Signalverarbeitungsparameter so zu wählen, dass Grundfunktionen der gestörten Hörwahrnehmung wie Lautheits- und Tonhöhenempfindung, Detektion zeitlicher Modulationen und weitere Klangeigenschaften möglichst gut kompensiert werden. Die zweite Gruppe betrifft globalere Funktionen, welche als Vorverarbeitung für praktisch beliebige nachgeschaltete Systeme benutzt werden können. Reduzierung von Störlärmeinfluss durch gesteuerte Richtmikrofone, automatische Klassifikation von Klangsituationen zur Wahl unterschiedlicher Verarbeitungsprogramme, Integration von Kommunikationssystemen wie Telefon, Raumbeschallungsanlagen, Internetanbindung usw. Im folgenden werden zu diesen beiden Signalverarbeitungsklassen einige Beispiele vorgestellt.

#### Signalverarbeitung für Innenohrschwerhörigkeit

Die Signalverarbeitung im Gehör und die bei einer Innenohr-Schwerhörigkeit möglicherweise auftretenden Störungen können vereinfacht wie folgt beschrieben werden. Das Schallsignal gelangt über Außen- und Mittelohr in das Innenohr, dessen Wirkung grob durch eine Filterbank mit anschließender einhüllenden Extraktion und Dynamikkompression (kompressive Nichtlinearität z.B. durch aktive Prozesse im Innenohr) modelliert werden kann. Die weitere Signalanalyse im Hirnstamm kann durch eine Zerlegung nach Modulationsfrequenzen (Modulationsfilterbank) und einen binauralen Vergleich mit Störschallunterdrückung charakterisiert werden. Am Ausgang dieser Vorverarbeitungsstufen stehen die akustischen Merkmale des Eingangsschallsignals in Form einer "internen Repräsentation" der kognitiven Verarbeitung im Hörkortex zur Verfügung, wo eine optimale Verarbeitung und Mustererkennung der eingehenden, transformierten Schallsignale vorgenommen wird. Limitiert wird dabei die "Schärfe" der internen Repräsentation durch das "interne Rauschen". Als Störungen ist an erster Stelle die Abschwächungswirkung zu nennen, die durch Schallleitungs-Schwerhörigkeit und Ausfall der inneren sowie teilweisen Ausfall der äußeren Haarzellen bedingt ist, sowie als zweiter Faktor der Kompressionsverlust (Ausfall der äußeren Haarzellen), an dritter Stelle der binaurale Verlust (Reduktion der binauralen Störschall-Unterdrückung) und an vierter Stelle eine Erhöhung des internen Rauschens als "zentraler Hörverlust".

Ausgehend von diesem grob-schematischen Modell sind folgende Möglichkeiten zur Kompensation der Verarbeitungsstörung denkbar:

- 1. Die Abschwächungswirkung lässt sich durch eine lineare Verstärkung kompensieren.
- Der Kompressionsverlust lässt sich durch eine Dynamikkompression kompensieren, die modellbasiert erfolgen sollte, um gemäß einem möglichst validierten Verarbeitungsmodell zu einer optimalen Kompensation dieses Funktionsausfalls zu gelangen.
- 3. Der binaurale Verlust sollte durch eine "echt" binaurale Signalverarbeitung (z. B. Verstärkung der Unterschiede zwischen beiden Ohren) erfolgen, die normalerweise im Gehirn durch binaurale Signalverarbeitung durchgeführt wird.
- 4. Der zentrale Hörverlust kann durch eine Stör-Reduktion und eine Anpassung der jeweiligen Signalverarbeitung an die akute aktuelle akustische Situation erfolgen.

Die Wiederherstellung der Dynamik- und Lautheitsempfindung in unterschiedlichen Frequenzbereichen erfolgt sinnvollerweise durch einen Lautheitsmodell-gesteuerten Ansatz, bei dem das jeweilige Eingangssignal einerseits mit einem Lautheitsmodell für Normalhörende bewertet wird und zugleich mit einem Lautheitsmodell für den individuellen Schwerhörigen. Aus dem Vergleich kann die Modifikation des Eingangssignals im Hörgerät so durchgeführt werden, daß am Ausgang beider Lautheitsmodelle möglichst der gleiche (frequenzabhängige) Wert resultiert (vgl. Hohmann, 1993, Launer, 1995, Appell, 2002). Da dieser Ansatz wesentlich von der Güte und Validierung des verwendeten Lautheitsmodells abhängt, gilt es, hier eine an die Gegebenheiten von Schwerhörenden und an praktische Messungen angepasste und validierte Version des Lautheitsmodells zu verwenden.



Abb. 1:Modell zur Berechnung der Lautheit für Normal- und Schwerhörige (modifiziert nach Zwicker).

Das in Abb.1 gezeigte modifizierte Modell berücksichtigt die kochleären Defizite Schwerhöriger. Die erhöhte Hörschwelle dämpft direkt die Anregung der Cochlea, die Steilheit der Lautheitsfunktion beeinflusst die Potenzfunktion und die Lautheitsbildung wird damit im Vergleich zu Normalhörenden stark verändert. Die optimalen Parameter für das Lautheitsmodell liefern das Audiogramm und die Lautheitsskalierung.

Die Zielverstärkung in einem solchen Signalverarbeitungsalgorithmus ist abhängig vom Signal. In jedem einzelnen Filterband wird die Verstärkung so gewählt, dass die entsprechende Bandlautheit (spezifische Lautheit) derjenigen eines Normalhörenden bei gleichem Signal entspricht; die Lautheit wird normalisiert.

Die Umwandlung des Erregungspegel-Musters in eine spezifische Lautheit geschieht durch eine Lautheitstransformation, die die Form der Isophonen von Normal- und Schwerhörenden insbesondere bei tiefen Frequenzen und dabei auch neuere Messungen zum Verlauf dieser Isophonen berücksichtigt (Gabriel et al., 1997). Nach der sich anschließenden spektralen Aufintegration der spezifischen Lautheit und Bestimmung der Gesamt-Lautheit in sone erfolgt eine Transformation in kategoriale Einheiten, so daß die Kategorial-Lautheit (KU) vorhergesagt werden kann, die sich unmittelbar mit Hilfe der kategorialen Lautheitsskalierung experimentell nachmessen lässt

Herkömmliche Hörgeräte modellieren die Cochlea als lineare, pegelunabhängige Filterbank und berechnen die Zielverstärkung als Funktion des Schalldruckpegels. Mit einem solchen Multiband-Konzept können jedoch maskierte Signale nicht erkannt werden, daher werden auch maskierte Komponenten verstärkt. Diese Fehlsignale äussern sich als Rauschen und Verzerrungen und beeinträchtigen die Klangqualität des Hörgerätes. Klassische Multiband-Systeme sind auch nicht in der Lage, schmal- und breitbandige Signale zu differenzieren und können damit auch die verminderte Lautheitssummation Schwerhöriger nicht kompensieren.

Eine gehörgerechte Signalverarbeitung hingegen bezieht die elementaren Funktionen der normalen und geschädigten Cochlea wie Maskierung, Lautheit und Lautheitssummation mit ein.

#### Möglichkeiten der Störgeräusch-Reduktion

Die derzeitige Lage von Störgeräusch-Unterdrückungs-Algorithmen bei Hörsystemen stellt sich wie folgt dar:

- Die Annahmen verschiedener Störunterdrückungs-Algorithmen sind in der Realität nie vollständig erfüllt, so daß in modernen Hörgeräten verschiedene Algorithmen kombiniert bzw. je nach vorliegender akustischer Umgebungssituation aktiviert und deaktiviert werden müssen.
- Akustisch "einfache" Situationen sind mit derzeitiger Technologie beherrschbar (dazu zählt stationäres Störgeräusch mit einer großen spektralen Differenz zwischen Sprache und Störgeräusch und stabilen räumlichen Differenzen).
- Probleme bereiten dagegen nach wie vor "schwierige Situationen", die sich durch instationäre Störschallquellen, durch Nachhall und durch mehrere gleichzeitig aktive Sprecher von verschiedenen Richtungen auszeichnen. Hier besteht noch ein großer Forschungsbedarf.

Die neueste Entwicklung in diesem Bereich stellen adaptive Mehr-Mikrofon-Systeme dar (Spriet et al., 2005). Bei ihnen wird die Richtcharakteristik in Abhängigkeit von der Einfallsrichtung des Störgeräusches variiert (siehe Abb. 2)



Abb. 2: Adaptives Richtmikrofonsystem. Die Richtcharakteristik wird in Abhängigkeit von der Einfallsrichtung des Störgeräusches variiert.

Zur Reduzierung von Störgeräuschen können in digitalen Hörgeräten die Eingangssignale analysiert und dadurch Nutzsignale von Rauschen unterschieden werden. Die frequenzabhängige Verstärkung des Hörgerätes kann dann in Abhängigkeit von der Umgebungssituation geregelt werden. Um Sprache von Störgeräuschen zu unterscheiden, müssen spezielle Kenntnisse über die Eigenschaften von Sprache herangezogen werden. Mit deren Hilfe können z. B. Sprachpausen erkannt werden. In solchen Pausen kann das Rauschsignal näherungsweise berechnet werden und vom Eingangssignal subtrahiert werden. Auf diese Weise wird das Rauschsignal reduziert. Da die Subtraktion im Frequenzbereich erfolgt, nennt man dieses Verfahren spektrale Subtraktion (s. z.B. Marcincik und Kollmeier, 1999).

Bei einem anderen Algorithmus werden z. B. die langsamen Schwankungen der Einhüllenden der Signale berechnet. Die Änderungsgeschwindigkeit der Einhüllenden wird als Modulationsfrequenz bezeichnet. Die Einhüllende zeigt im Frequenzbereich einen für Sprache typischen Verlauf mit einem Maximum im Bereich von 2-8 Hz. Diese Frequenzen sind durch den Rhythmus der Sprache (die Dauer der Wörter, Silben und Phoneme) bestimmt und sind sowohl vom Klang der Stimme als auch von der gesprochenen Sprache unabhängig. Störgeräusche dagegen beinhalten zumeist andere Modulationsfrequenzen. Ein Algorithmus kann nun die Einhüllenden der Eingangssignale analysieren. Wenn die typischen Modulationsfrequenzen von Sprache im Bereich von bis zu 10 Hz nicht vorhanden sind, wird die Verstärkung in dem entsprechenden Frequenzbereich und damit das Störgeräusch reduziert.

#### Feedbackreduktion

Ein wichtiges Qualitätsmerkmal von Hörgeräten ist die Unterdrückung der akustischen Rückkopplungen (Feedback). Sie wird meist durch eine Reduzierung der Verstärkung bei der Anpassung des Hörgerätes vermieden. Eine andere Möglichkeit sind schmalbandige Lückenfilter (Notch-Filter), die bei den kritischen Feedbackfrequenzen positioniert werden und die Verstärkung bei der betreffenden Frequenz damit selektiv reduzieren. Solche statischen Anpassungen sind jedoch wirkungslos, falls die Rückkopplungen als Reaktion auf veränderte äußere akustische Bedingungen plötzlich bei einer anderen Frequenz auftreten. Dieses Problem können adaptive Algorithmen lösen. Sie analysieren ständig das Eingangssignal und erkennen Rückkopplungen. Ein Notch-Filter oder ein Kompensationsfilter eliminieren den Feedbackanteil im Signal und werden ständig an die aktuellen Bedingungen angepaßt. Der Kompensationsfilters-Ansatz vermeidet Rückkopplungen durch die Einstellung eines adaptiven Filters, so dass sein Ausgangssignal die rückgekoppelten Anteile im Eingangssignal gerade auslöscht.

#### Auditorische Objekterkennung und Klangklassifizierung

Hörgeräte ermöglichen es, für verschiedene akustische Situationen unterschiedliche Programme zu wählen, um den Frequenzgang und Kompressionsparameter zu ändern, oder Richtmikrofon, Störgeräusch-Reduktion oder Feedback-Unterdrückung zu aktivieren. Der Hörgeräteträger hat dabei die nicht immer leichte Aufgabe, die akustische Hörsituation zu beurteilen und dann per Schalter am Hörgerät oder über eine Fernbedienung das entsprechende Programm zu wählen. Ein automatisches Erkennen der aktuellen akustischen Situation und automatisches Umschalten in das geeignetste Programm könnten den Hörgerätekomfort verbessern. Ein System zur Geräuschklassifizierung wird im folgenden vorgestellt, welches die vier Klassen 'Sprache', 'Sprache im Störgeräusch', 'Störgeräusch' und 'Musik' robust erkennen soll.

Es sind bereits einige Hörgeräte im Handel, die eine Geräuschklassifizierung vornehmen. In einem Gerät von Widex wird die Amplitudenstatistik des Signals ausgewertet, um zwischen impulshaften und kontinuierlichen Klängen zu unterscheiden und damit ein Störgeräusch-Unterdrückungssystem zu steuern (Ludvigsen, 1993). Im Zusammenhang mit Störgeräusch-Unterdrückung wird auch von ReSound eine Geräuschklassifizierung vorgenommen (Edwards et al., 1998). Anhand einer Analyse der Amplitudenmodulationen wird Sprache im Signal detektiert. Von Phonak ist ein Gerät im Handel, das abhängig von der akustischen Umgebung zwischen zwei Programmen umschaltet (Phonak, 1999). Die Klassifizierung in die Klassen 'Sprache im Störgeräusch' und 'Andere' erfolgt aufgrund einer Analyse der zeitlichen Fluktuationen und der Form des Spektrums, wie es von Kates (1995) vorgeschlagen wurde. Ostendorf et al. (1998) unterscheiden basierend auf einer Modulationsfrequenz-Analyse die drei Klassen 'Sprache', 'Sprache im Störgeräusch', und 'Störgeräusch'. Von Nordqvist (2000) wird ein Verfahren vorgeschlagen, mit dem reine Sprache und verschiedene Hintergrund-Geräusche mittels einer LPC-Analyse und HMMs klassifiziert werden. Ein Ansatz von Feldbusch (1998) erkennt reine Sprache, Stimmengewirr und Verkehrslärm, indem verschiedene zeitliche und spektrale Merkmale mittels neuronalen Netzen ausgewertet werden.

Alle genannten Ansätze erlauben ein robustes Erkennen von reiner Sprache; Musik hingegen kann bis anhin nicht identifiziert werden, und es ist nur teilweise möglich, Störgeräusch von Sprache im Störgeräusch zu trennen.

#### Auditorische Szenenanalyse

Auditorische Szenenanalyse beschreibt die Mechanismen und Strategien, die das auditorische System verwendet, um die akustische Umgebung zu analysieren. Obwohl dieser Prozess noch nicht vollständig verstanden wird, ist bekannt, dass das auditorische System charakteristische Merkmale aus dem akustischen Signal extrahiert. Diese auditorischen Merkmale umfassen spektrale Trennung, spektrales Profil, Harmonizität, Onsets und Offsets, kohärente Amplituden- und Frequenz-Modulationen sowie räumliche und zeitliche Trennung.

#### Auditorische Merkmale

Im folgenden Ansatz für die Geräuschklassifizierung werden bis anhin vier Merkmalsgruppen verwendet: Amplituden-Modulationen, spektrales Profil, Harmonizität und Onsets. Für die Modellierung der Amplituden-Modulationen werden die genannten bisherigen Ansätze von Ludvigsen (1993), Ostendorf et al. (1998) und Kates (1995) verwendet. Fig. 3 zeigt als Beispiel, wie das Amplituden-Histogramm mittels Perzentilen modelliert wird, und der Abstand zwischen den Perzentilen Aufschluss über die Modulationstiefe gibt.



Fig. 3: Amplitudenhistogramm und Perzentile eines Sprachsignals.Das Merkmal 'Breite' ist definiert als Differenz zwischen der 90 % und der 10 % Perzentile.

Das spektrale Profil wird auf eine rudimentäre Weise durch zwei Merkmale modelliert, den spektralen Schwerpunkt und die zeitlichen Schwankungen des spektralen Schwerpunktes (Fig. 4). Diese beiden Merkmale wurden bereits im Klassifizierer von Phonak (1999) verwendet.



Fig. 4: Zur Beschreibung des spektralen Profils werden der spektrale Schwerpunkt und dessen zeitliche Schwankungen berechnet.

Für die Beschreibung der Harmonizität wird normalerweise der Pitch eines Klanges benutzt (Fig. 5). In unserem Ansatz wird die Harmonizität durch zwei Merkmale charakterisiert: die Tonalität und die Pitch-Varianz. Die Tonalität ist definiert durch das Verhältnis der harmonischen zu den unharmonischen Anteilen im Zeitsignal.



Fig. 5: Typischer zeitlicher Pitch-Verlauf eines Sprachsignals. Das Merkmal Tonalität ist das Verhältnis von harmonischen zu unharmonischen (hier durch 0 Hz angezeigten) Anteilen.

Um Amplituden-Onsets zu modellieren, wird das Auftreten hoher und schneller Amplitudenanstiege im Bark-Spektrum analysiert. Fig. 6 zeigt ein typisches Onsetmuster für Popmusik. Da im Onsetmuster auch sehr schön der Rhythmus eines Signals sichtbar wird, wird zusätzlich ein Merkmal für den Takt berechnet.



Fig. 6: Onsetmuster von Pop-Musik. Starke Onsets sind durch Punkte dargestellt. Der Rhythmus der Musik ist deutlich erkennbar.

Die genannten auditorischen Merkmale wurden zusammen mit Klassifizierern unterschiedlicher Komplexität evaluiert. Fünf verschiedene Klassifizierer-Typen kamen zur Anwendung: Regelbasierter Ansatz, Minimum-Distance Klassifizierer, Bayes Klassifizierer, neuronales Netz, und Hidden Markov Modell (HMM).

Der Vorteil der ersten zwei Klassifizierer ist, dass diese etwa viermal weniger Rechenzeit und Speicherplatz benötigen als die übrigen Ansätze; dafür ermöglichen die letzteren eine genauere Unterteilung des Merkmalsraums.

Die Performance der einzelnen Klassifizierer mit unterschiedlichen Merkmalssets wurde mit etwa 300 verschiedenen Alltagsklängen getestet. Die Klang-Datenbank enthält unterschiedlichste Beispiele zu den vier Klassen; für die Klasse 'Sprache' zum Beispiel verschiedene Sprecher, die unterschiedliche Sprachen mit unterschiedlichen Geschwindigkeiten und Tonlagen in unterschiedlich halligen Räumen sprechen, oder Signale der Klasse 'Störgeräusch' aus Haushalt, Verkehr, Restaurant, Industrie usw., 'Sprache im Störgeräusch' bei verschiedenen SNR, und 'Musik' von Klassik über Pop und Rock zu einzelnen Instrumenten und Gesang.

Von jeder Klasse wurden 80 % der Klänge fürs Training und 20 % fürs Testen der Klassifizierer verwendet. Viele Klänge wurden sehr robust erkannt, insbesondere reine und schwach verhallte Sprache, Verkehrs- und Stimmen-Geräusche, klassische Musik, Instrumente und Gesang. Probleme bereiten vor allem stark verhallte Sprache, fluktuierende Störgeräusche und komprimierte Pop-Musik. Dafür sollten weitere Merkmale aus dem Signal extrahiert werden, wie zum Beispiel ein Mass für die Stärke des Nachhalls im Signal.

Merkmale zur auditorischen Objekterkennung

Einige charakteristischen spezifischen Merkmale (Yost & Sheft, 1993), welche für die Bildung akustischer Objekte wichtig sind, werden im folgenden aufgelistet.

• Spektrale Trennung

Die bemerkenswerte Fähigkeit des auditorischen Systems, den spektralen Inhalt eines Signals zu bestimmen, ist eine wichtige Basis für die weitere Trennung in die nachfolgenden Merkmale.

• Spektrales Profil

Die meisten Klangquellen produzieren ein bestimmtes Amplitudenspektrum, dessen Profil relativ konstant bleibt wenn der Gesamtpegel verändert wird. Verschiedene Klangklassen, wie z.B. Sprache, stationäres Störgeräusch und Musik, haben oft unterschiedliche spektrale Profile, was zur Unterscheidung von Quellen beitragen kann. Das spektrale Profil ist auch für die Klangfarbe (Timbre) von Musikinstrumenten und Stimmen wesentlich. Das Timbre wird jedoch ebenso durch temporale Faktoren bestimmt, wie Anstiegs- und Abfallzeiten von gespielten Noten eines Instruments.

• Harmonische Struktur

Die Teiltöne vieler natürlicher Klangquellen haben eine harmonische oder annähernd harmonische Struktur, die ihren Ursprung in einem schwingenden Medium hat. Die Harmonischen formen ein einzelnes auditorisches Bild, beschrieben als Pitch. Dies ist ein hervorragendes Beispiel von Fusion, denn es ist in der Regel unmöglich, einzelne Teiltöne herauszuhören. Da unterschiedliche Quellen meist verschiedene harmonische Serien bilden, ist mit dem Pitch eine Trennung von Quellen möglich.

#### • Gemeinsame Ein / Ausschwingzeit

Eines der stärksten Merkmale für die Gruppierung von Teiltönen ist der gemeinsame Anstieg der Amplitude. Gleichzeitig einsetzende Teiltöne werden in der Regel zu einem auditorischen Objekt fusioniert. Kleinere Asynchronitäten bestimmen zu einem wesentlichen Teil das Timbre der Quelle, grössere asynchrone Amplitudenanstiege einzelner Teiltöne tragen andererseits stark zur Trennung in separate Quellen bei, auch wenn die Teiltöne untereinander harmonische Verhält-nisse aufweisen. Messungen haben ergeben, dass Musiker in einem Orchester um sich selbst zu hören bis zu 50 ms versetzt spielen (Mellinger & Mont-Reynaud, 1996). Auch das Publikum kann dadurch einzelne Instrumente heraushören, ohne aber den Eindruck eines gleichzeitigen Einsatzes zu verlieren.

• Kohärente Amplituden- und Frequenzmodulation

Amplitudenmodulation ist charakteristisch für viele natürliche Klangquellen wie z.B. Sprache. Unter kohärenter Amplitudenmodulation versteht man die gleichzeitige Amplitudenänderung von mehreren Teiltönen. Das auditorische System vermag spektrale Komponenten zu gruppieren, die mit demselben zeitlichen Muster moduliert sind, oder zu trennen, wenn die Muster unter-schiedlich sind. Dies funktioniert für Modulationsfrequenzen unter 50 bis 100 Hz.

Weit weniger stark scheint die gemeinsame Frequenzmodulation mehrerer Partialtöne zur Fusion oder Trennung beizutragen. Trotzdem kann sie gerade bei Musik zur Trennung von Instrumenten beitragen (Vibrato).

#### Lokalisation

Das auditorische System verrichtet eine schwierige Aufgabe, wenn es entscheidet, woher im Raum ein Signal kommt. Interaurale Zeit- und Pegeldifferenzen und Aspekte der Übertragungs-funktion von Kopf / Ohrmuschel sind entscheidend für die Lokalisation. Für die Trennung von Klangquellen ist die räumliche Lokalisation zwar signifikant aber nicht unerlässlich.

#### Gruppierungsverfahren

Unter Gruppierung versteht man die auf den vorgängig beschriebenen physikalischen Merkmalen basierende Zuordnung der Ereignisse in einem akustischen Signal zu verschiedenen akustischen Objekten (Fusion oder Trennung).

Primitive Gruppierung und Gruppierung nach Schema

Bregman (1990) unterscheidet zwischen zwei Mechanismen um zu entscheiden, welche Komponenten zu einer bestimmten Klangquelle gehören.

- Primitive Gruppierungsmechanismen teilen das Eingangssignal auf der Basis einfacher physikalischer Signalmerkmale auf, wie dies im vorhergehenden Abschnitt beschrieben wurde. Diese Mechanismen funktionieren ohne vorheriges Wissen.
- Bei der Gruppierung nach Schema geschieht die Auswahl und Kombination von Komponenten des komplexen Signals aufgrund von Erfahrungen und gelernten Mustern, also aufgrund der Eigenschaften von Klängen oder Klangklassen, die einem schon früher begegnet sind.

Die Erfahrung zeigt, dass die kontextunabhängige, primitive Gruppierung viel einfacher modellierbar und anzuwenden ist als die Gruppierung nach Schema. Allerdings ist nicht zu erwarten, dass eine vergleichbare Klassifizierungs- und Quellentrennleistung wie beim auditorischen System erreicht werden kann, wenn die höher liegenden Gruppierungsmechanismen nicht berücksichtigt werden.

#### Gestaltprinzipien

Am Anfang des zwanzigsten Jahrhunderts hat die Forschung der visuellen Wahrnehmung die sogenannten Gestaltprinzipien formuliert, welche nun auch für auditorische Empfindungen Anwendung finden. Bregman (1990) nennt dazu zwei Grundregeln (allocation und accounting):

- Jedes Element eines komplexen Signals kann nur einer Quelle zugeordnet werden.
- Alle Elemente des akustischen Ereignisses müssen einer Quelle zugeordnet werden. Kann ein Element keiner existierenden Quelle zugeordnet werden, so wird es selbst zu einer Quelle.

Für die Bildung von auditorischen Objekten werden diese Gestaltprinzipien auf verschiedenen Ebenen angewendet.

- Kontinuität: Amplitudenmoduliertes Rauschen wird als einzelnes Objekt wahrgenommen, falls die Modulation glatt (zum Beispiel sinusförmig) ist. Bei Rechteckmodulation hingegen hört man zwei Objekte. Pfeifen, zusammenhängend oder abgesetzt, erzeugt eine oder zwei Quellen.
- Nähe: Aufeinanderfolgende Noten, die spektral nahe beieinander liegen (Pitch), werden zur gleichen Quelle gezählt. Es braucht viel, bis eine Melodie durch Störsignale zerfällt.
- Ähnlichkeit: Ähnliche Klangkomponenten werden gruppiert; Noten mit demselben Timbre werden meistens derselben Quelle zugeordnet.
- Gemeinsames Schicksal: Ändern sich Teiltöne eines Klangs auf gleiche Weise, werden sie gruppiert. Gemeinsamer Amplitudenanstieg und gemeinsame Modulationen sind starke Gruppierungsprinzipien. Werden in einem harmonischen Klang einige Teiltöne moduliert, heben sie sich von den anderen ab, und man hört zwei Klänge.
- Geschlossenheit: Geschlossenheit hilft, Kontinuität trotz abrupter Änderungen zu erhalten. Lücken in einem Signal trennen Ereignisse, während Rauschen in den Lücken hilft, sie wieder zu fusionieren. Ein Beispiel dazu wäre Klatschen beim Sprechen. Auch die Trennung von mehreren gleichzeitigen Sprechern wird dadurch unterstützt.
- Gute Fortsetzung: Die meisten Klänge ändern nicht plötzlich den Charakter; ein Klavier wird nicht plötzlich wie eine Violine tönen. Das heisst, dass feine, glatte Änderungen einer einzigen Quelle zugeordnet werden, plötzliche Wechsel erzeugen eine neue Quelle.

Wenn wir uns die Prozesse zur akustischen Szenenanalyse nochmals vor Augen führen und mit den Methoden aktueller Hörgeräte-Automaten vergleichen so erkennen wir, dass die heutigen Lösungsansätze sehr technisch inspiriert sind und noch wenig mit unserem eigenen Wahrnehmungssystem zu tun haben. Wir haben aber die Grenzen der heutigen Lösungen erkannt und neue Wege entdeckt, die noch fehlenden Klassifizierungsleistungen einzubringen. Die primitive Gruppierung auditorisch basierter Merkmale mit Hilfe der Gestaltprinzipien ist ein vielversprechender Lösungsansatz zur besseren Klassifizierung der akustischen Umgebung. Es dürfte möglich sein, einige Prinzipien zu formulieren und in Hörsystemen einzusetzen.

#### Schlussfolgerungen

Die eingangs genannten Ziele der rehabilitativen Audiologie (Hören, Verstehen, Klangqualität) sind durch digitale Signalverarbeitung heute nur unzureichend erfüllbar, obwohl in den letzten Jahren ein deutlicher Fortschritt erzielt worden ist und auch mit weiteren Fortschritten in nächster Zeit zu rechnen ist. Aus audiologischer Sicht ist dem Versuch eines Ausgleichens des Hörschadens eine besondere Bedeutung beizumessen, wobei es um eine Verbesserung der Diagnostik und des Verständnisses von Verarbeitungsdefiziten bei Innenohr-Schwerhörigkeit primär geht. Dabei erscheint der Modell-basierte Ansatz erfolgversprechend, da durch Verarbeitungsmodelle unser derzeitiges Wissen über die Funktion und etwaige Fehlfunktion des Hörsystems in für die Hörgeräte-Signalverarbeitung nutzbarer Form charakterisiert werden kann. Komplexe Computermodelle zur auditorischen Szenenanalyse sind allerdings derzeit zur Implementation in Hörgeräten noch zu aufwendig. Wenn nicht nur rein technisch inspirierte Konzepte zur Klassifizierung akustischer Situationen und Objekte verwendet werden sollen, lassen sich den zusammengefassten Theorien folgend aber Modelle und Regeln zur Verwendung in Hörinstrumenten ableiten.

#### Literatur

- Appell, Jens-Ekkehardt, Loudness models for rehabilitative Audiology, Dissertation, Universität Oldenburg, 2002 (http://docserver.bis.uni-oldenburg.de/publikationen/dissertation/2002/applou02/applou02.html)
- Büchler, M. et al. (2000). "Klassifizierung der akustischen Umgebung für Hörgeräte-Anwendungen", DAGA 2000.

Dau, T. et al. (1998). "Psychophysics, Physiology and Models of Hearing". World Scientific Publishing.

- Edwards, B. W., Hou, Z., Struck, C. J., Dharan, P. (1999). "Signal-processing algorithms for a new softwarebased, digital hearing device," The Hearing Journal, Vol. 51, No. 9, 44-52.
- Feldbusch, F. (1998). "Geräuscherkennung mittels Neuronaler Netze," Zeitschrift für Audiologie 1/1998, 30-36.
- Gabriel, B., B. Kollmeier and V. Mellert (1997). "Influence of individual listener, measurement room and choice of test tone levels on the shape of equal-loudness level contours." Acustica united with acta acustica 83(4): 670-683.
- Hohmann, V. (1993). Dynamikkompression für Hörgeräte- Psychoakustische Grundlagen und Algorithmen. Düsseldorf, VDI-Verlag.
- Kates, J. M. (1995). "Classification of background noises for hearing-aid applications", Journal of the Acoustical Society of America, 97 (1), 461-470.
- Launer, S. (1995). Loudness perception in listeners with sensorineural hearing loss. Fachbereich Physik. Oldenburg, Carl-von-Ossietzky Universität Oldenburg.( http://www.physik.unioldenburg.de/Docs/medi/pub\_down.html)
- Ludvigsen, C. (1997). "Schaltungsanordnung für die automatische Regelung von Hörhilfsgeräten", Europäische Patentschrift, EP 0 732 036 B1.
- Marzinzik, M. and Kollmeier, B. (**1999**). "Development and Evaluation of Single-Microphone Noise Reduction Algorithms for Digital Hearing Aids," in Psychophysics, Physiology and Models of Hearing, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, Singapore), pp. 279-282.
- Mellinger, D. K., Mont-Reynaud, B. M. (1996). "Scene Analysis," in Auditory Computation, edited by H. L. Hawkins et al., Springer, New York.
- Nordqvist, P. (2000). "Automatic Classification of Different Listening Environments in a Generalized Adaptive Hearing Aid," International Hearing Aid Research Conference, IHCON 2000, Lake Tahoe, CA.
- Ostendorf, M. et al. (1997). "Empirische Klassifizierung verschiedener akustischer Signale und Sprache mittels einer Modulationsfrequenzanalyse", DAGA 97, 608-609.
- Phonak Hearing Systems (1999). "Claro AutoSelect", company brochure no. 028-0148-02.
- Spriet, A., Moonen, M., and Wouters, J. (2005). "Stochastic gradient-based implementation of Spatially Preprocessed Speech Distortion Weighted Multichannel Wiener filtering for noise reduction in hearing aids," IEEE Trans. Signal Process. 53, 911-925.
- Yost, A. W., Sheft, S. (1993). "Auditory Perception," in Human Psychophysics, edited by W. A. Yost et al., Springer, Berlin.
- Zwicker, E., Fastl, H. (1990). "Psychoacoustics, Facts and Models", Springer, Berlin.

### **Cochlea-Implantate**

#### Norbert Dillier / Zürich

#### Einführung

Patienten, welche aufgrund einer Innenohr-Schädigung ertaubt oder hochgradig schwerhörig geworden sind, sind heute vielfach in der Lage, mit einem Cochlear Implant (CI) wieder Sprache zu hören und zu verstehen. Auch frühertaubte oder taub geborene Kleinkinder können mit Hilfe der elektrischen Hörnervstimulation nicht nur eine direkte Verbindung zur akustischen Umwelt aufnehmen, sondern auch eine dem Alter entsprechende Sprache entwickeln.

Ein CI-System besteht aus Implantat, Sprachprozessor und Mikrofon-Sendeeinheit. Die Stimulationselektroden zur elektrischen Reizung des Hörnerven werden in einer Operation in die Hörschnecke eingelegt. Durch den hinter dem Ohr getragenen digitalen Sprachprozessor werden Schallwellen in elektrische Impulse umgewandelt und über eine Sendespule durch die Haut zum Implantat und zu den Hörnervfasern übertragen. Der Hörnerv leitet dieses Reizmuster in ähnlicher Art und Weise wie beim natürlichen Hören zum Gehirn weiter, wo es als Hör- und Klangempfindung wahrgenommen wird. Je nach Alter und Vorgeschichte ist nach der Operation eine mehrwöchige bis mehrmonatige Hörtrainings- und Rehabilitationsphase zum Erlernen und Unterscheiden der künstlichen Hörsignale angezeigt.

Das Cochlear Implant als erster und bislang einziger in der Praxis funktionierender Ersatz für ein Sinnesorgan ist ein erfolgreiches Beispiel für die gegenseitige Befruchtung von Biologie, Medizin und Technik. Weltweit wurden schätzungsweise bereits nahezu 100'000 Patienten mit diesen Implantaten versorgt.

Die aktuelle Forschung und Entwicklung zielt auf die verbesserte Programmierung und Anpassung der Sprachprozessoren. Dazu werden auch elektrophysiologische Messungen der Hörnervantworten bei verschiedenen Reizsignalen durchgeführt. Mit neuartigen Elektrodenanordnungen sowie zunehmend höheren Reizraten wird ein natürlicheres Klangbild sowie eine differenziertere Wahrnehmungsleistung angestrebt.

#### Komponenten des Cochlear Implants

Cochlear Implants bestehen aus Komponenten, welche ausserhalb des Körpers getragen werden und dem eigentlichen Implantat. Ueber ein hinter dem Ohr befestigtes Mikrophon mit Richtcharakteristik werden Schallsignale aufgenommen und über ein Kabel dem Sprachprozessor zugeleitet, welcher die Schallreize in geeignete elektrische Stimulationsmuster umwandelt und wiederum über ein Kabel als Hochfrequenzsignale einer Sendespule zuführt. Die Information gelangt drahtlos über induktive Kopplung zum Implantat mit seiner Empfangsspule sowie der Decodierungs- und Stimulatorelektronik. Die in der Cochlea plazierten Elelektroden vermitteln das nunmehr elektrische Signal zum Hörnerv und lösen einen Höreindruck aus. Für telemetrische Messungen dienen beide Spulen als Sender und Empfänger.

Der Sprachprozessor enthält die digitale Reizcodierungselektronik mit zwei bis acht gespeicherten Programmen und eine oder mehrere (aufladbare) Batterien. Derzeit sind Systeme von vier kommerziellen Anbietern verfügbar (Tabelle 1). Bereits sind Sprachprozessoren soweit miniaturisiert worden, dass sie als Hinter-dem-Ohr-Einheit getragen werden können. HdO-Sprachprozessoren haben nahezu die gleiche Leistungsfähigkeit wie die in Hemd- oder Blusentasche oder bei Kleinkindern in Brustbeutel oder Rucksäckchen getragenen grösseren Taschengerät-Sprachprozessoren.

	Nucleus, Cochlear	Advanced Bio-	MedEl	Neurelec MXM
	(Australien)	nics	(Oesterreich)	(Frankreich)
		(USA)		, ,
Bezeichnung	CI24R/CI24RE/	HiRes 90k	Combi40+/ Pul-	Digisonic
	Freedom	Harmony	sar/Sonata	
Abmessung (mm)	27x18x6.4	31x25x6	33.4x23.4x4	28Øx6.8
Gewicht (Gramm)	9.5	8	9	15
Material Gehäuse	Titan	Titan	Keramik Titan	Keramik
Max. Pulsrate (pps)	14'500/31'600	83'000	18'180/?	7'800
Stimulationskanäle	22/(virtuell 43)	16	12	15
Stimulationsmodus	Bipolar Monopolar Common ground	Bipolar Monopolar	Monopolar	Common ground
Sprachprozessoren	SPrint ESPrit, Esprit-3G, Freedom (HdO)	Clarion Platinum PSP und BTE, HiRes Harmony	CIS PRO+ TEMPO+, Opus (HdO)	Digisonic, DigiSP K TG und HdO
Anzahl Programme	8/2	3	3	2
Strategien	SPEAK (NofM) CIS ACE (NofM) MP3000	SAS CIS PPS	CIS NofM FSP	CIS/ASR NofM
Information	www.cochlear.com	www.bionicear.c	www.medel.com	www.neurelec.c

*Tabelle 1: Technische Daten der kommerziell verfügbaren Implantate: CI22M und CI24M (Nucleus, Cochlear), Clarion (Advanced Bionics), Combi 40+ (MedEl), Digisonic (MXM).* 

#### Funktionsweise des Cochlear Implants

Die Aufgabe eines Cochlear Implants ist es, im Falle einer Innenohr-Taubheit den Hörnerven direkt elektrisch zu stimulieren und dadurch Hörempfindungen zu erzeugen. Die Information, welche im akustischen Schallsignal enthalten ist, muss durch geeignete Umwandlung für jeden Patienten individuell in die adäquaten elektroneuralen Erregungsmuster übersetzt werden.

Die heutigen CI-Systeme sind Meilensteine auf dem Weg einer fortschreitenden technologischen Entwicklung und Verbesserung. Sie stellen aber ohne Ausnahme im Vergleich mit dem zu ersetzenden biologischen System technische Kompromisslösungen dar, welche immer nur Teilfunktionen des peripheren Gehörs ersetzen können (Patrick u. Evans 1995).

Der Hörnerv kann mit unterschiedlichen elektrischen Reizformen angeregt werden. Die für einen effektiven Reiz wirksame elektrische Grösse ist der Strom, welcher während einer gewissen Zeit in das Nervengewebe hinein- oder herausfliesst. Zur Vermeidung von schädlichen elektrochemischen Reaktionen am Elektroden-Gewebe-Uebergang muss das elektrische Stimulationssignal gleichstromfrei angekoppelt werden (eine aktuelle Uebersicht über Modelle der elektrischen Hörnervstimulation ist zu finden in Hamacher, 2004).

Als Reizformen kommen sinusoidale (analoge) oder pulsatile Signale zur Anwendung.Bei der analogen Stimulation wird das Signal ähnlich wie bei einem konventionellen Hörgerät gefiltert und komprimiert und dadurch an den eingeschränkten elektrischen Dynamikbereich angepasst. Sobald allerdings mehrere nahe beeinander liegende Reizorte vorhanden sind, ergeben sich Lautheitssummationen durch Stromüberlagerungen. Biphasische Pulse erlauben demgegenüber eine zeitlich besser kontrollierbare Reizung als analoge Sinussignale. Bei der Stimulierung von mehreren Kanälen können Interferenzen vermindert werden, indem die Kanäle zeitlich versetzt ("interleaved") aktiviert werden. Nachteilig wirkt sich unter Umständen bei der sequentiellen Stimulation die verminderte Zeitauflösung aus, da pro Reizpuls immer nur die Information eines Frequenzbandes übertragen wird, die anderen also inaktiv bleiben.

Voraussetzung für die elektrische Stimulation ist eine genügende Anzahl elektrisch erregbarer Nervenzellen. Durch die Lage der Stimulations- und Referenzelektroden kann der aktivierte Bereich beeinflusst werden. Die bipolare Elektrodenansteuerung zielt darauf ab, einen möglichst geringen Anteil von Nervenfasern und damit ein schmales Frequenzband zu aktivieren. Bei weiter auseinanderliegenden Elektroden werden bei gleicher Reizstärke mehr Neuronen aktiviert als bei nahe beisammenliegenden. Die monopolare Reizkonfiguration stellt diesbezüglich einen Grenzfall dar, indem vor allem die Hörnervfasern in der Nähe der aktiven intracochleären Elektrode aktiviert werden, während die im Muskelgewebe plazierte Referenzelektrode keine Empfindungen auslöst.

Für die Sprachprozessor-Einstellung muss für jede Elektrodenkonfiguration die Wahrnehmungsschwelle (Threshold, T-level) sowie die Reizstärke für eine angenehme, erträgliche Lautheitsempfindung (Comfort, C-level, manchmal auch MCL, most comfortable level genannt, bzw. durch UCL, upper comfortable level, ersetzt) bestimmt werden. Der Bereich zwischen diesen Werten (dynamic range, dynamischer Bereich) liegt in den meisten Fällen unterhalb von 6 dB (Ladung/Phase auf 1 nanoCoulomb bezogen) und ist damit wesentlich kleiner als die bis zu 120 dB Dynamik bei der akustischen Stimulation.

Die Hörempfindungen bei elektrischer Stimulation werden durch die Reizrate und den Ort der Stimulation bestimmt. Die Reizrate eines periodischen Stimulationssignals kann bis zu einer Frequenz von etwa 300 Pulsen pro Sekunde der Grundfrequenz der menschlichen Stimme (Pitch) zugeordnet werden. Die Stimulation an verschiedenen Reizorten innerhalb der Cochlea ruft Empfindungen hervor, die mit der ortsabhängigen Tonhöhenempfindung Normalhörender verglichen werden kann. Die Reizelektroden können somit wie die Tasten eines Klaviers unterschiedliche Klangspektren wiedergeben. Die wahrgenommenen charakteristischen Frequenzen hängen von der Eindringtiefe des Elektrodenträgers ab. Mit den über 17 mm verteilten 22 Elektroden des Nucleus-Implantats beispielsweise kann bei einer Eindringtiefe von ca. 20 mm der Frequenzbereich zwischen etwa 700 und 11'000 Hz angesprochen

werden.

In jüngster Zeit sind von mehreren Implantatherstellern neue Arten von Elektrodenformen vorgestellt worden, welche einerseits tiefer in die Hörschnecke einzudringen erlauben und andererseits gezielter das elektrische Stromfeld in Richtung auf die Hörnervfasern im Zentrum der Schnecke ausrichten sollen. Dadurch wird eine bessere Trennschärfe der einzelnen Elektrodenkanäle sowie eine Reduktion der Stromstärke angestrebt. Ein Beispiel einer vorgeformten Elektrode, welche zur Einführung in die Cochlea mit einem sogenannten Stilett (einem dünnen Draht) gerade gehalten wird, ist in Abb. 1 dargestellt. Inwiefern die angestrebten Ziele mit diesen neuen Elektroden erreicht werden können, ist Gegenstand laufender Studien.



Abb. 1 Beispiel einer vorgeformten Elektrode (Nucleus 24 Contour), welche sich bei der Implantation nach Entfernen eines dünnen Drahtes (Stilett) eng an die innere Schneckenwand (Modiolus) anschmiegt.

Bezüglich Sprachcodierung haben die Forschungsarbeiten der letzten zehn Jahre zu einer beträchtlichen qualitativen Verbesserung und ebenso zu einer Vereinheitlichung der zugrundeliegenden Konzepte und Strategien geführt. Die Merkmalsextraktion gemäss dem Schema eines Formantvocoders, wie sie in den Sprachprozessor der ersten Generation realisiert war und bei welchem nur die zwei oder drei Maximalwerte des Frequenzspektrums ermittelt und auf die Reizelektroden abgebildet wurden, wurde zugunsten immer detaillierterer zeitlich besser aufgelöster Spektralinformation aufgegeben. Bei der SPEAK-Strategie beispielsweise werden laufend aus 20 Frequenzbändern die sechs bis zehn grössten Werte bestimmt und mit einer Reizrate von etwa 250 Pulsen pro Sekunde (pps) auf die

entsprechenden Elektroden abgebildet (Skinner et al. 1994). Die moderneren ACE-Strategien (Advanced Combination Encoder, manchmal auch mit dem Begriff NofM bezeichnet) erlauben die Kombination von höheren Reizraten auf gleichzeitig mehr Kanälen bis zur derzeit maximalen Gesamtrate von 14'400 pps (mit dem CI24R-Implantat) bzw. 31'600 (mit dem CI24RE-Implantat) (also zum Beispiel 16 Kanäle mit je 900 bzw. 1975 pps, 12 Kanäle mit je 1200 bzw. 2600 pps oder 9 Kanäle mit je 1800 bzw. 3500 pps). Abb. 2 zeigt beispielhaft, wie ein gesprochenes Wort spektral in verschiedene Frequenzbänder zerlegt wird, von denen zu jedem Zeitpunkt diejenigen Bänder mit der grössten Signalenergie ausgewählt und den entsprechenden Elektroden zugeordnet werden. Das zeitlich und räumlich verteilte Reizmus-

ter für ein Elektroden-Bündel ist in Abb. 3 dargestellt.

Die früher häufig verwendete Breitband-Analog-Stimulation, welche heute nur noch als SAS-Strategie (Simultaneous Analog Stimulation) im Clarion-System angeboten wird, wurde von mehreren Forschungsgruppen durch pulsatile Reizung möglichst hoher Raten (CIS, continuous interleaved sampling) ersetzt. Besonders bei Elektrodenkonfigurationen mit geringem Abstand zwischen benachbarten Elektroden bietet die nichtsimultane Pulsreizung gegenüber simultaner Analogreizung gewichtige Vorteile. Die hohen Reizraten können insbesondere die zeitliche Feinstruktur der Sprachsignale besser reproduzieren (Wilson et al. 1991).



Abb.2 Frequenzanalyse des Sprachsignals gemäss der Advanced Combination Encoder (ACE) Strategie des Nucleus-Sprachprozessors. Die obere Kurve stellt das gesprochene Wort "Fall" dar, welches mit einer Filterbank in maximal 22 Frequenzbänder aufgeteilt wird. Die zu einem Zeitpunkt energiereichsten Bandfilter werden zur Stimulation ausgewählt und einer Reizelektrode zugeordnet. Die unteren beiden Grafiken stellen die Frequenzspektren und die jeweils aktiven Elektroden zum Zeitpunkt 100 msec (links, Konsonant) und 240 msec (rechts, Vokal) dar.

Wie gut die durch die Signalcodierung reduzierte und auf die implantierten Elektroden übertragene Sprachinformation im Einzelfall zur Diskrimination von Wörtern und Sätzen ausreicht, kann nur durch Evaluationsexperimente mit grösseren Patientengruppen ermittelt werden. Solche Experimente sind naturgemäss aufwendig und von vielen Variablen abhängig (Dillier et al. 1995; Brill et al. 1997).

Zusammenfassend lässt sich jedoch sagen, dass mit den modernen digitalen Strategien spontan ein besseres Konsonantenverständnis erzielt wurde als mit den von den Patienten früher benutzten Sprachprozessoren. Die besten Ergebnisse wurden mit Verfahren erzielt, welche höhere Reizraten und differenziertere Abbildungen spektraler Feinstrukturen enthielten



Abb. 3 Zeitlicher Verlauf der Reizsignale an den einzelnen Elektroden. Dargestellt ist eine Elektrodogramm-Analyse des gesprochenen Wortes "Fall". Mit einer Wiederholrate von 900 Hertz werden jeweils maximal 12 Elektroden zeitlich versetzt aktiviert ("interleaved stimulation"). Die Auswahl der Kanäle erfolgt wie in Abb. 3 beschrieben gemäss der spektralen Energieverteilung.Horizontal ist die Zeit von 0 bis 400 Millisekunden dargestellt, vertikal die Elektrodenposition von 22 (apikal, tiefste Ortsfrequenz) bis 1 (basal, höchste Ortsfrequenz). Die Länge und Schwärzung der Striche ist proportional zur jeweiligen Reizamplitude. Die unteren drei Grafiken zeigen zeitlich gedehnte Ausschnitte von jeweils 20 msec Dauer: stimmloser Konsonant /f/ (100 - 120 msec), Vokal /a/ (240 - 260 msec), stimmhafter Konsonant /l/ (350 - 370 msec).

#### Audiologische Indikationen bei Erwachsenen und Kindern

Für die Versorgung mit einem Cochlear Implant sind Patienten geeignet, die eine beidseitige überwiegend sensorische Taubheit bis hin zur an Taubheit grenzenden Schwerhörigkeit haben.

Dieser Indikationsbereich ist jedoch nicht an starre Grenzen der Tonschwellenaudiometrie gebunden, sondern vielmehr an die tatsächliche Sprachperzeption. Patienten, die trotz optimaler Versorgung mit konventionellen Hörgeräten kein Sprachverstehen haben, sind Kandidaten für ein Cochlear Implant. Ausnahme sind ältere Kinder, Jugendliche und Erwachsene, die taub geboren sind.

Erwachsene müssen in der präoperativen Evaluation ton- und sprachaudiometrisch untersucht werden. Hinzu kommen die Impedanzaudiometrie sowie die Hirnstammaudiometrie und Elektrocochleographie. Das Resultat einer solchen Untersuchung kann eine Neuversorgung mit optimierten Hörgeräten sein. Liegt das Einsilberverstehen im Freiburger Sprachtest unter 40 % - bei Freifeldmessung mit Hörgeräten -, so ist heute die Indikation zum Cochlear Implant gegeben (Fraysse et al. 1998). Im übrigen dürfen Patienten mit kurzer Ertaubungsdauer ein besseres Sprachverstehen nach Implantation erwarten als Langzeitertaubte.

Bei Kindern haben wir aus audiometrischer Sicht mit sehr ähnlichen Bedingungen zu rechnen. Die Indikation wird dann als gegeben angesehen, wenn im Vorschulalter mit Hörgeräten eine ausreichende Sprachperzeption ausbleibt und somit eine adäquate Sprachproduktion nicht zu erwarten ist. In Abhängigkeit vom Alter des Kindes sind objektive audiometrische Testverfahren einzusetzen, ergänzt durch Spiel- und verhaltensaudiometrisch erfasste Tonschwellen. Entscheidend für den Erfolg einer CI-Versorgung beim Kleinkind ist die frühe Erkennung des Hörschadens. Die rechtzeitige Versorgung mit Hörgeräten und optimale hörund sprachgerichtete sonderpädagogische Förderung sollten 6 Monate nach Diagnosestellung eine Entscheidung für oder gegen ein Cochlear Implant erbringen (Dillier u. Laszig 2001)

#### Intra- und postoperative objektive Messungen

Die modernen Cochlear Implants enthalten alle eine Möglichkeit zur Ueberprüfung der Funktionsfähigkeit des Implantats sowie zur Messung der Elektrodenimpedanzen. Die Uebermittlung der gemessenen Werte erfolgt telemetrisch über die Sende/Empfangsspule und wird in einem Testgerät oder Messprogramm ausgewertet.

Die elektrische Impedanzmessung ermöglicht die Identifikation von kurzgeschlossenen oder unterbrochenen Elektrodenzuleitungen innerhalb weniger Sekunden. Für die Messung werden Testreize mit geringer Stromstärke verwendet, welche normalerweise von den Patienten nicht wahrgenommen werden können. Aus dem Verhältnis der gemessenen Spannung über den Reizelektroden und der aufgeprägten Stromamplitude wird ein Mass für die Elektrodenimpedanz ermittelt (Swanson et al. 1995; Zierhofer et al. 1997; Schulman 1995).

Die intra- oder post-operativen objektiven audiologischen Verfahren verwenden zur Stimulation die implantierten intracochleären sowie allenfalls extracochleäre Referenz-Elektroden (bei monopolarer Stimulation). Intraoperativ stehen einerseits die Ableitung der elektrisch evozierten Reizantworten über Hautelektroden bzw. über das Implantat selbst, andererseits die visuelle Registrierung des Stapediusreflexes zur Verfügung (Shallop 1993).

Die intraoperative Beobachtung des Stapediusreflexes unter operationsmikroskopischer Sicht ist wenig aufwendig, wenig störanfällig und erfolgt während der Stimulation des Hörnerven über das Implantat. Insofern dient die Stapediusreflexmessung auch der Funktionsprüfung des Hörnerven und damit letztendlich auch der des Implantates. Die Stapediusreflexschwelle liegt etwa im Uebergang zwischen mittlerem und oberem Drittel des späteren Dynamikbereichs (Battmer et al. 1990) und liefert Richtgrössen zur Vermeidung von Ueberstimulation bei der ersten Anpassung. Mit den Telemetrieoptionen der heutigen Implantate hat sich die Bedeutung der Stapediusreflexbeobachtung zur Sicherstellung der Implantatsfunktion relativiert.

Neben den elektrisch evozierten Stapediusreflexen sind die elektrisch evozierten Hirnstammpotentiale und die telemetrisch übermittelten elektrisch evozierten Aktionspotentiale (neurale Reizantwort-Telemetrie, NRT) weitere objektive audiologische Messverfahren. Nach intracochleärer Stimulation über das Implantat werden Hirnstammpotentiale über Fernelektroden wie bei der BERA am Kopf abgeleitet oder Aktionspotentiale des Ganglion spirale über intracochleäre Elektroden und Telemetrie gesendet und registriert.



Abb. 4 NRT-Messaufbau. Ein Sprachprozessor (SPrint), welcher mit dem klinischen Anpasscomputer verbunden ist, sendet Reizpulse über die Sendespule zum Implantat (Nucleus CI24). Die neuralen Antworten werden durch die implantierten Elektroden abgeleitet, im Implant verstärkt und über die Sende/Empangsspulen zum Sprachprozessor und PC zurückgesendet.

Die neuste Generation von Implantaten erlaubt neben der mehrkanaligen Stimulation auch die Messung von Elektroden- und Implantatseigenschaften sowie die Ableitung neuraler Reizantworten (NRT). Die Uebermittlung der Messdaten erfolgt telemetrisch durch die intakte Haut (Abbas et al. 1999). Zur Messung werden ausser dem standardmässigen Programmierungssystem (PC mit Spezialinterface) und einem Sprachprozessor nur das NRT-Programm benötigt. Abb. 4 zeigt den Aufbau der verschiedenen Komponenten des Messsystems. Ein vom NRT-Programm ausgewähltes Reizsignal (Stimulus) wird über die ausgewählte Elektrode ausgegeben. Die resultierende neurale Antwort wird von einer anderen Elektrode in der Nähe abgeleitet. Das Messergebnis wird verstärkt, codiert und über die Sendeeinheit an den Sprachprozessor zurückgeschickt.



Abb. 5 Beispiel einer intraoperativen NRT-Messung. Links: Antwortkurven bei Stimulation an allen 22 intracochleären Elektroden (monopolare Reizung und Ableitung, 1.5 mm Abstand der Messelektrode von der aktiven Reizelektrode, Stromstärke jeweils 210 CL-Einheiten, Pulsbreite 25 µs, Reizrate 80 Pulse pro Sekunde, 50 Mittelungen). Rechts unten: Einzelkurve mit negativer (N1) und positiver (P1) Welle. Rechts oben: Antwortamplituden (P1 - N1) bei 3 Stromstärken (210, 200, 190 CL) für alle 22 intracochleären Elektroden.

Die entsprechenden neuralen Reizantworten (oder Nervenantworten) werden am Bildschirm dargestellt und können anschliessend ausgewertet werden. Die Form und Grösse der Antwort sind abhängig von den Reiz- und Messbedingungen. NRT-Messungen können ohne grösseren Aufwand intraoperativ durchgeführt werden. Sie bieten neben dem Nachweis der neuralen Aktivität auch den Vorteil der kompletten Funktionsüberprüfung des Implantats.Ein Beispiel einer intraoperativen Serie von NRT-Messungen über alle 22 intracochleären Elektroden zeigt Abb. 5. Häufig lässt sich eine typische Antwortkurve des Summenaktionspotentials gewinnen mit einer negativen Welle (N1), welche von einer positiven Welle (P1) gefolgt wird (s. Abb. 5 rechts unten). Die Zunahme des neuralen Summenaktionspotentials mit steigender Reizstärke (SL, Stimulus-Level) korrespondiert mit der postoperativen subjektiven Lautheitswahrnehmung (von sehr leise bis angenehm laut). Aus den Amplitudendifferenzen der beiden Wellen für drei verschiedene Reizstärken (210, 200, 190) kann eine NRT-Schwelle extrapoliert werden (Abb. 5 rechts oben), welche meist etwa in der Mitte zwischen den späteren T- und C-Werten liegt.

#### **Postoperative Basis- und Folgetherapie**

Der Erfolg einer CI-Versorgung hängt von vielen Faktoren ab, nicht zuletzt von der Anpassung des Sprachprozessors. Ein gut angepasster Sprachprozessor ermöglicht es dem Patienten, die Klangeindrücke in ihrer Vielfalt wahrzunehmen und schneller voneinander unterscheiden zu lernen. Hauptziel einer Anpassung ist es, für jede aktive Elektrode die Hörschwellen und den Pegel der angenehmen Lautheit zu bestimmen. Bei Erwachsenen ist diese Anpassung relativ problemlos, da sie ihr Lautheitsempfinden in der Regel gut beschreiben und skalieren können. Die Arbeit mit Kindern erweist sich als schwieriger, da deren subjektive Angaben nicht immer so zuverlässig (und reproduzierbar) erfolgen. Zur Verifizierung der subjektiven Angaben können objektive Messverfahren beigezogen werden.

Die Anpassung des Sprachprozessors erfolgt nach abgeschlossener Wundheilung. Alle modernen Cochlear Implants werden heute mittels eines Computers über eine spezielle Software angepasst. Grundsätzlich müssen für alle Implantatstypen die optimalen Stimulationsparameter bestimmt und im Prozessor gespeichert werden. Die Ermittlung der minimalen Stromstärken, die Höreindrücke auslösen (Hörschwelle, Threshold bzw. T-level) und der maximalen Stromstärken, die noch angenehm laut empfunden werden (Comfort level bzw. most comfortable level, abgekürzt C-level oder MCL) für alle verfügbaren Kanäle ist bei ertaubten Erwachsenen einfach und zuverlässig.

Die postoperative Basis- und Folgetherapie ertaubter Erwachsen beinhaltet eine bis auf die Erstanpassung weitgehend ambulante Begleitung, Beratung über technische Fragen, Ersatz von defekten Teilen, Veränderung der Anpassungsdaten und Aktualisierung der individuellen Hard- und Software sowie die Dokumentation der Fortschritte mittels verschiedener Hör- und Sprachtests.

Die CI-Anpassung bei Kleinkindern bedarf einer engen Zusammenarbeit von Pädagogen und Audiologen und ähnelt in mancher Hinsicht der pädagogischen Verhaltensbeobachtung zur Hörprüfung oder zur Hörgeräteanpassung (Bertram 1998). Das "Hören" der Kleinkinder, insbesondere taubgeborener, ist zunächst nur an sehr diskreten Reaktionen der Augen oder der Mimik zu erkennen. Sie zu entdecken und richtig zu werten, ist das hauptsächliche Ziel der ersten Sitzungen zur Anpassung des Sprachprozessors. Erfahrene und eingespielte Teams sind in der Lage, auch mit Kindern unter 2 Jahren zu arbeiten und zuverlässige Anpasswerte zu erreichen (Lenarz et al. 1994).

Parallel zur Vervollständigung und wiederholten Kontrolle des Sprachprozessor-Programms werden Hörübungen und Hörtests zur Förderung und Verlaufsdokumentation der Sprachentwicklung durchgeführt Hörtest (Bertram 1998) (McCormick et al. 1994). Für die Verlaufskontrolle bei Kleinkindern vor dem Spracherwerb bietet sich der LiP-Test (Listening Progress) dar, welcher in einer Uebersichtstabelle die sprachlichen Teilleistungen auflistet (Archbold et al. 1995). Der Test of Auditory Perception of Speech (TAPS, (Reid u. Lehnhardt 1994)) ist ebenfalls auf CI-Kinder abgestimmt und zwar schon ab dem dritten Lebensjahr. Als weitere Testbatterie zu erwähnen ist auch der EARS test (Evaluation of Auditory Responses, MedEl), der aus ursprünglich englischen Tests zusammengesetzt ist und in verschiedenen Europäischen Sprachversionen vorliegt. Eine Zusammenstellung von weiteren Tests für Sprachperzeption und -Produktion von CI-Kindern findet sich in (Dyar 1994).

Die technologische Entwicklung ist heute so weit fortgeschritten, dass auch bei Kleinkindern

eine sichere, zuverlässige und erfolgreiche Versorgung mit einem Cochlear Implant vorgenommen werden kann (Lehnhardt et al. 1992). Neben der Hörentwicklung sind speziell bei Kindern vor allem die Fortschritte in der Sprachentwicklung hervorzuheben (Robbins et al. 1995; Waltzman et al. 1997).

#### Literatur

- Abbas PJ, Brown CJ, Shallop JK, Firszt JB, Hughes ML, Hong SH, Staller SJ. Summary of results using the nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear 1999;20: 45-59
- Archbold S, Lutman ME, Marshall DH. Categories of Auditory Performance. Ann Otol Rhinol Laryngol [Suppl] 1995;166: 312-314
- Battmer RD, Laszig R, Lehnhardt E. Electrically elicited stapedius reflex in cochlear implant patients. Ear Hear 1990;11: 370-374
- Bertram B. Rehabilitationskonzept bei Kindern. In: Lenarz T (ed) Cochlea-Implantat. Ein praktischer Leitfaden für die Versorgung von Kindern und Erwachsenen. Springer-Verlag, Berlin, Heidelberg, 1998: 108-121
- Brill SM, Gstoettner W, Helms J, von Ilberg C, Baumgartner W, Muller J, Kiefer J. Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+. Am J Otol 1997;18: S104-6
- Dillier N, Battmer RD, Döring WH, Müller-Deile J. Multicentric field evaluation of a new speech coding strategy for cochlear implants. Audiology 1995;34: 145-159
- Dillier N, Laszig R. Audiometrie und Cochlear Implant. In: Lehnhardt E, Laszig R (eds) Praxis der Audiometrie. Georg Thieme-Verlag, Stuttgart, 2001: 293-309
- Dyar D. Monitoring progress: the role of a speech and language therapist. In: McCormick B, Archbold S, Sheppard S (eds) Cochlear Implants for young children. Whurr Publishers Ltd, London, 1994: 237-268
- Fraysse B, Dillier N, Klenzner T, Laszig R, Manrique MJ, Perez CM, Morgon A, Müller-Deile J, Ramos AM. Cochlear Implants for Adults Obtaining Marginal Benefit from Acoustic Amplification. Am J Otol 1998;19: 591-597
- Hamacher, V. (2004). Signalverarbeitungsmodelle des elektrisch stimulierten Gehörs. Wissenschaftsverlag Mainz in Aachen, ISBN 3-928493-48-5
- Lehnhardt E, Gnadeberg S, Battmer RD, von Wallenberg EL. Experience with the cochlear miniature speech processor in adults and children together with a comparison of unipolar and bipolar modes. ORL 1992;54: 308-313
- Lenarz T, Lehnhardt E, Bertram B. Cochlear Implant bei Kindern. Thieme Verlag, Stuttgart, New York, 1994: 1-158.

McCormick B, Archbold S, Sheppard S. Cochlear Implants for Young Children. Whurr Pub-

lisher Ltd., London, 1994: 1-291.

- Patrick JF, Evans AR. Implant designs for future coding strategies. Ann Otol Rhinol Laryngol [Suppl] 1995;166: 137-138
- Reid J, Lehnhardt M. Speech perception test results for European children using the Nucleus Cochlear Implant. Adv in Cochlear Implants, Proc 3rd Int Cochlear Implant Conf (Innsbruck AU) 1994522-527
- Robbins AM, Osberger MJ, Miyamoto RT, Kessler KS. Language development in young children with cochlear implants. Adv Otorhinolaryngol 1995;50: 160-166
- Schulman JH. Using impedance telemetry to diagnose cochlear electrode history, location, and functionality. Ann Otol Rhinol Laryngol [Suppl] 1995;166: 85-87
- Shallop JK. Objective electrophysiological measures from cochlear implant patients. [Review]. Ear Hear 1993;14: 58-63
- Skinner MW, Clark GM, Whitford LA, Seligman PM, Staller SJ, Shipp DB, Shallop JK, Everingham C, Menapace CM, Arndt PL, Antognelli T, Brimacombe JA, Pijl S, Daniels P, George CR, McDermott HJ, Beiter AL. Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. Am J Otol Suppl 1994;15: 15-27
- Swanson B, Seligman P, Carter P. Impedance measurement of the Nucleus 22-electrode array in patients. Ann Otol Rhinol Laryngol [Suppl] 1995;166: 141-144
- Waltzman S, Cohen NL, Gomolin R, Green J, Shapiro W, Brackett D, Zara C. Perception and production results in children implanted between 2 and 5 years of age. Adv Otorhinolaryngol 1997;52: 177-180
- Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. Nature 1991;352: 236-238
- Zierhofer CM, Hochmair IJ, Hochmair ES. The advanced Combi 40+ cochlear implant. Am J Otol 1997;18: S37-S38

### Implantierbare Hörgeräte

### Andreas Büchner / Hannover

Neben den konventionellen Hörgeräten gibt es inzwischen eine Vielzahl an teil- und vollimplantierbaren Hörsystemen. Es gibt auch eine Reihe von medizinischen und audiologischen Indikationen, bei denen die Anwendung von implantierbaren Hörsystemen ein besseres Ergebnis für die Patienten bringen kann. Allerdings hängt der Erfolg der Hörversorgung solcher Therapiemaßnahmen davon ab, dass die Auswahl der Patienten unter Betrachtziehen aller medizinischen und audiologischen Aspekte in ausgewiesenen Kliniken mit ausreichender Erfahrung auf diesem Gebiet durchgeführt wird. Einige von den Indikationen für implantierbare Hörsysteme sind in folgenden genannt:

- Rezidivierende Gehörgangentzündungen infolge konv. HG
- Malformationen des Gehörganges (z. B. Verengung, Exostose usw.)
- Ausgeprägter Hochtonhörverlust mit geringer Hörminderung in Tief- und Mitteltonbereich
- Limitierte Sprachverständlichkeit bei Nebengeräuschen mit konv. HG
- Schallleitung- und kombinierte Schwerhörigkeiten infolge von verschiedene Krankheiten (z. B. fehlender Gehörgang oder Mittelohrstrukturen, Chronische Mittelohrentzündungen mit oder ohne Trommelfellperforation, Radikalhöhle, Otosklerose usw.)
- Tragekomfort und soziale Aspekte

Die implantierbaren Hörgeräte werden in teil- und vollimplantierbare Hörsysteme eingeteilt. Die teilimplantierbaren Hörgeräte werden wiederum in passive und aktive Hörsysteme gegliedert. Untere Tabelle zeigt die Einteilung der implantierbaren Hörgräte mit dazugehörigen Hörsystemen.


## Vibrant<sup>®</sup> Soundbridge<sup>®</sup>

Dieses teilimplantierbare aktive Hörgerät Vibrant Soundbridge (VSB) von Fa. MED-EL besteht aus einem internen Teil (genant als VORP, S. Bild) und einem externen Teil als Audioprozessor, welcher die Signalverarbeitung und Stromversorgung des Implantates beinhaltet. Die HNO-Klinik an der MHH verfügt über jahrelange Erfahrung (seit 1997) mit über 100 Implantationen an Patienten.





Die VSB eignet sich für gering- bis mittelgradige, grenzend an hochgradige Schallleitungs-, Schallempfindung- und kombinierte Schwerhörigkeiten. Die Bilder zeigen die Operationsbzw. Ankopplungsmethode für verschiedene Indikationen.



Ankopplung am Amboss für Innenohrschwerhörigkeiten Ankopplung am runden Fenster für kombinierte Schwerhörigkeiten



Die Auswahlkriterien für die VSB je nach Indikationen sind in den Diagrammen dargestellt.

Um die Eignung dieses Systems präoperativ zu bestimmen bzw. dem Patienten einen Höreindruck zu vermitteln, wird ein Simulationssystem (das sog. DDS-System) mit dem Trommelfell in Kontakt gebracht und Hörsamples (Sprache und Musik) vorgespielt (S. Bild).





**DDS (Direct Drive Stimulation)** 

Darstellung der Simulationsmethode

### **Otologics MET**<sup>TM</sup>

Das teilimplantierbare Hörsystem MET (Middle Ear Transducer) von Fa. Otologics eignet sich für mittel- bis hochgradige Schallempfindungsschwerhörigkeiten. Dieses System besteht auch aus einem internen Implantat und einem externen Audioprozessor (s. Bild). Das Diagramm rechts zeigt die Auswahlkriterien für MET.



### **Otologics CARINA**<sup>TM</sup>

Das vollimplantierbare Hörsystem CARINA von Fa. Otologics hat seit 2006 eine CE-Markt Zulassung und wurde bis jetzt weltweit an 450 Patienten implantiert. Dieses System eignet sich für die mittelbis hochgradige Schallempfindungs- und kombinierten Schwerhörigkeiten.





Ankopplung am Amboss für Innenohrschwerhörigkeiten

Bei diesem System werden alle Teile (Mikrofon, Batterie, Audioprozessor und der Wandler) implantiert. Die Programmierung und das Aufladen der Batterie erfolgt transkutan (über die Haut) wie in Abb. dargestellt. Die Aufladedauer beträgt etwa 1 Std. für je etwa 16 Std. Nutzungszeit.







Ankopplung am Stapes für kombinierte Schwerhörigkeiten



Aufladen von CARINA über das Gurtsystem

Frequency [kHz]

1 1.5 2 3 4 6

Die Abb. zeigt die Auswahlkriterien für das CARINA.



0.25

0

0.5

#### Esteem®

Das Funktionsprinzip des vollimplantierbaren Hörsystems Esteem von Fa. Envoy Medical basiert auf die Schallaufnahme am Amboß durch einen Piezosensor, welcher nach entsprechender Signalverarbeitung und Verstärkung durch einen zweiten Piezowandler, welcher am Stapes implantiert wird, an das Innenohr weitergegeben wird. Allerdings, um Verzerrungen bzw. Rückkopplungen zu vermeiden, muss die normale Gehörknöchelchenkette unterbrochen werden. Die Energieversorgung dieses Systems erfolgt über spezielle Batterien, die sonst bei Herzschrittmachern eingesetzt werden. Die Batterien haben eine Lebensdauer von ca. 3 Jahren und müssen dann über einen kleinen chirurgischen Eingriff gewechselt werden.





### Versorgung und Rehabilitation mit technischen Hörhilfen

Jürgen Kießling / Gießen

#### 1. Aufbau und Funktion von Hörgeräten

Unter der Bezeichnung *technische Hörhilfen* versteht man grundsätzlich sämtliche Formen konventioneller Hörgeräte und implantierbarer Hörsysteme sowie Zusatzeinrichtungen und Übertragungsanlagen (FM-Anlagen, Bluetooth), die im Englischen unter dem Begriff *Assistive Listening Devices (ALD)* zusammengefasst werden. Da implantierbare Hörsysteme an anderer Stelle abhandelt werden, beschränkt sich der vorliegende Beitrag auf die Behandlung nichtimplantierbarer Hörhilfen. Heutige Hörgeräte basieren grundsätzlich auf digitaler Signalverarbeitung. Das grundlegende Funktionsprinzip von Digitalhörgeräten basiert auf den in Abb. 1 schematisch dargestellten Komponenten: Schallempfänger (Mikrofon bzw. Induktionsspule), Vorverstärker, Analog-Digital-Wandler, Signalprozessor mit n Frequenzbändern, Digital-Analog-Wandler, Nachverstärker und Schallsender (Hörer).

Im normalen Betrieb handelt es sich beim Schallempfänger um ein Mikrofon, das den Schall am Hörgeräteeingang aufnimmt und für die weitere Verarbeitung in eine elektrische Wechselspannung umwandelt. Neben dem Mikrofon sind zahlreiche Hörgeräte zusätzlich mit einem Empfänger für den induktiven Betrieb (Induktionsspule, Telefonspule) ausgestattet, um optional auch elektromagnetische Wellen aufnehmen zu können, die z. B. von Induktionsschleifen oder Telefonhörern abgestrahlt werden. Über eine weitere Möglichkeit der Signaleinspeisung verfügen Hörgeräte, die mit Audioeingang ausgestattet sind. Bei diesen Geräten kann zum Anschluss externer Schallquellen, wie Klassenverstärkeranlagen, Funkanlagen, Radio- oder Fernsehgeräten, CD- oder MP3-Playern sowie anderer Zusatzeinrichtungen, das Eingangssignal direkt als elektrische Spannung über den Audioeingang ins Hörgerät eingekoppelt werden. Insbesondere für die Versorgung schwerhöriger Kinder hat sich der Audioeingang als unverzichtbar erwiesen, wohingegen er bei der Versorgung Erwachsener bisher nur begrenzte Bedeutung erlangt hat.



Abb. 1: Funktionsdiagramm eines mehrkanaligen Hörgerätes mit digitaler Signalverarbeitung

Das Mikrofon (bzw. die Induktionsspule oder der Audioeingang) liefert das elektrische Eingangssignal, das in den nächsten Schritten verstärkt und mittels Analog-Digital-Wandler für die weitere Verarbeitung digitalisiert wird. Das Digitalsignal wird im Signalprozessor entsprechend den Bedürfnissen des Nutzers bearbeitet, d.h. verstärkt, komprimiert, gefiltert oder in anderer Weise modifiziert, wobei auch bei Digitalhörgeräten die Verstärkungsfunktion im Vordergrund steht. Angesichts der Vielzahl von innenohrschwerhörigen Hörgerätekandidaten mit Recruitment sind in der Regel nicht-lineare Verstärkungssysteme angezeigt, bei denen das erforderliche frequenzabhängige Verstärkungs- und Kompressionsverhalten durch Aufteilung des Eingangssignals in mehrere Frequenzbänder verwirklicht wird. Nach digitaler Verarbeitung wird das Signal digital-analog gewandelt und nachverstärkt, um den nachfolgenden Wandler zu betreiben. Bei Luftleitungshörgeräten hat ein Miniaturhörer die Funktion des Schallsenders, der das Ohr direkt oder über ein Ohrpassstück beschallt. Knochenleitungshörgeräte sind mit Knochenleitungshörern ausgestattet, die eine vibratorische Anregung des Schädelknochens bewirken und damit Wanderwellen in der Cochlea direkt auslösen.

Neben den in Abb. 1 beschriebenen Hauptkomponenten benötigen Hörgeräte eine Energiezelle (Batterie oder Akku) und, sofern es sich nicht um Automatikgeräte handelt, Bedienungselemente für den Hörgeräteträger (Ein-Ausschalter, Programmschalter, Verstärkungsregler). Zudem verfügen Hörgeräte über eine Kabelanschlussmöglichkeit (Buchse, Adapter für Batteriefach) an das Programmiersystem, mit dem der Akustiker die Hörgerätewiedergabe an die Gehöreigenschaften des Nutzers individuell anpassen kann. Solche Programmiersysteme bestehen aus einem hörgerätespezifischen Interface und einem Personal Computer, auf dem die Anpassprogramme der Hörgeräthersteller installiert sind. Zusammen mit der Softwareplattform NOAH haben sich universelle kabelgebundene bzw. kabellose Schnittstellen zwischen Hörgerät und Computer als de facto Industriestandard zur Programmierung von Hörgeräten durchgesetzt, die von den Hörgeräte in Verbindung mit herstellerspezifischer Anpasssoftware programmiert und Messgeräte für die Hörgerätanpassung gesteuert werden. Daneben bieten die Hersteller auch Stand-alone-Versionen ihrer Anpassprogramme an.

#### 2. Signalverarbeitung in Hörgeräten

Die primäre Aufgabe von Hörgeräten besteht darin, den einfallenden Schall, insbesondere Sprache und andere relevante Schallereignisse, in den Restdynamikbereich des Hörgeräteträgers passend zu übertragen, so dass in etwa eine lautheitsgerechte Wahrnehmung erfolgen kann: leise Eingangssignale sollen leise und laute Eingangssignale sollen laut gehört werden. Dazu ist eine individuelle Frequenz- und Dynamikanpassung erforderlich. Sofern die Restdynamik ausreichend breit ist, kann die Verstärkung linear erfolgen, d. h. unabhängig vom Eingangspegel liefert das Hörgerät frequenzspezifisch eine einheitliche Verstärkung. Da die Mehrzahl der Hörgerätekandidaten jedoch unter Innenohrschwerhörigkeit mit Recruitment, also unter einer eingeengten Restdynamik, leiden, bedarf es in der Regel nicht-linearer Verstärkungssysteme.

Im Falle von stark frequenzabhängigen Hörstörungen verbunden mit frequenzspezifisch eingeengtem Restdynamikbereich, gelingt es mit einkanaligen, frequenzunabhängigen Kompressionssystemen nicht, Sprache adäquat in den Restdynamikbereich zu übertragen. In diesem für Hörgeräteversorgungen typischen Fall sind mehrkanalige Verstärkungs- und Kompressionssysteme notwendig. Digitalhörgeräte zerlegen das Eingangssignal grundsätzlich in mehrere Frequenzbänder, so dass frequenzspezifische Verstärkung und Kompression realisiert werden und deren Parameter individuell an das jeweilige Gehör angepasst werden können. In Abb. 2 ist eine frequenzabhängige Verstärkungs- und Kompressionswirkung für ein 12-kanaliges System dargestellt, mit dem das Sprachsignal komplett und dynamikgerecht in den Restdynamikbereich verlagert werden kann, ohne die Unbehaglichkeitsschwelle zu überschreiten. Neben der Möglichkeit der differenzierten Frequenz- und Dynamikanpassung bieten aktuelle Hörgeräte zusätzliche Funktions- und Ausstattungsmerkmale, die auf komplexer Signalverarbeitung beruhen. In diesem Bereich der digitalen Signalverarbeitung konnten bei der Bekämpfung akustischer Rückkopplungen deutliche Fortschritte erzielt werden. Zur Lösung dieses Problems haben sich Verfahren etabliert, die auf einer Auslöschung der Rückkopplung durch Addition eines gegenphasigen Signals beruhen. Zu diesem Zweck wird der akustische Rückkopplungspfad kontinuierlich analysiert, ein zum Rückkopplungssignal gegenphasiges Signal erzeugt und dem Hörgeräteeingang überlagert, so dass es nach einer kurzen Adaptationszeit zur Auslöschung der Rückkopplung kommt. Durch die Adaptivität passt sich das System eventuellen Änderungen des Rückkopplungspfades ständig an, so dass man derzeit eine um 15-20 dB höhere Verstärkung rückkopplungsfrei erreichen und im täglichen Gebrauch nutzen kann, was der offenen Versorgung in den letzten Jahren zum Durchbruch verholfen hat.

Ein anderes Problem von großer praktischer Bedeutung ist das eingeschränkte Sprachverstehen im Störgeräusch, speziell im Stimmengewirr. Auch diesem Problem kann man mit modernen Signalverarbeitungsalgorithmen entgegenwirken. Der heute gängigste Lösungsansatz beruht auf der Annahme, dass es sich beim Nutzsignal meist um Sprache handelt und das Nutzsignal damit durch eine sprachtypische Amplitudenmodulation gekennzeichnet ist. Zur Detektion und Reduktion des Störschalls nutzt man die Aufteilung des Eingangssignals in mehrere Frequenzbänder (siehe Abb. 2), in denen jeweils



Abb. 2: Grundprinzip der Signalverarbeitung in mehrkanaligen Hörgeräten veranschaulicht anhand eines Tonaudiogramms in Relativdarstellung.

eine Modulationsanalyse vorgenommen wird. Findet das System in einem Band eine ausgeprägte sprachtypische Modulation, ist davon auszugehen, dass in diesem Frequenzband Sprachanteile in erheblichem Maße auftreten, die erhalten werden müssen. Dementsprechend erfolgt in diesem Band keine Abschwächung. In den Bändern, in denen dagegen keine oder keine wesentliche sprachtypische Modulation auftritt, wird abhängig vom Modulationsgrad eine Abschwächung des Signals vorgenommen. Dann werden die bearbeiteten Teilsignale der einzelnen Frequenzbänder wieder zusammengeführt und dem Hörgerätenutzer über den Hörer hörbar gemacht.

Wenn in den einzelnen Frequenzbändern entweder überwiegend Sprachanteile oder Störschall auftreten, funktioniert diese Form der Störschallunterdrückung gut. Liegt dagegen, wie es in realen Störschallsituationen häufig der Fall ist, eine weitgehende spektrale Überlappung von



Abb. 3: Störschallunterdrückung auf der Basis einer Modulationsanalyse: Bänder mit Sprachmodulation werden verstärkt, Bänder ohne Sprachmodulation abgeschwächt.

Nutz- und Störschall vor, so kann diese Form der Störschallunterdrückung lediglich begrenzte Wirkung entwickeln, da in den Frequenzbändern des Überlappungsbereichs ein ähnlicher Modulationsgrad vorliegt, deshalb alle relevanten Bänder in gleichem Umfang abgeschwächt werden und somit nicht zur Verbesserung des Signal-Störschall-Abstands beitragen können.

Handelt es sich um eine Störschallsituation, bei der der Nutzschall von vorn einfällt, also der typischen Situation zweier Gesprächspartner im Störschall, kann mit Richtmikrofonen eine signifikante Verbesserung des Sprachverstehens erreicht werden. Das Grundprinzip moderner Richtmikrofonhörgeräte basiert auf dem Zusammenwirken zweier Mikrofone, die in ausreichendem Abstand in Vorn-Hinten-Richtung angeordnet sind und deren Signale voneinander subtrahiert werden. Dabei wird das Signal des hinteren Mikrofons zeitlich so verzögert, dass es bei der nachfolgenden Subtraktion der Mikrofonsignale zu einer Auslöschung des rückwärtigen oder seitlichen Schalls kommt.

Typische Richtmikrofonsysteme verfügen über zwei Einzelmikrofone, die in der beschriebenen Weise zusammenwirken, doch kann die Richtwirkung durch Mikrofonarrays bestehend aus drei oder mehr zusammen geschalteten Mikrofonen noch gesteigert werden. Das hier beschriebene Mikrofonprinzip bietet eine universelle Richtcharakteristik oder durch Wahl einer anderen Zeitverzögerung die Möglichkeit, die gewünschte Richtung maximaler Störschallunterdrückung (häufig die Hinten-Richtung) im Rahmen der Hörgeräteanpassung individuell einstellen zu können. Heutige Richtmikrofonhörgeräte wirken vielfach über diese Form der fest eingestellten Richtwirkung hinaus und verfügen über eine adaptive Richtwirkung, die sich automatisch auf die Unterdrückung der Hauptstörschallquelle einstellt.

Basierend auf den hier behandelten Formen der Signalverarbeitung (Verstärkung, Dynamikompression, Rückkopplungsunterdrückung, Störschallunterdrückung, Richtwirkung) kann ein Set von Grundeinstellungen für typische, immer wiederkehrende Hörsituationen wie z. B. Sprache in Ruhe, Sprache im Störschall, Störschall allein, Musik, Telefonieren etc., definiert und im Hörgerät programmiert werden. Diese so genannten Hörprogramme können zudem entsprechend den persönlichen Wünschen und Bedürfnissen des Hörgerätenutzers optimiert, auf verschiedenen Programmplätzen des Hörgerätes abgelegt und situationsbedingt vom Nutzer angewählt werden. In dieser Form können typischerweise zwei bis vier Hörprogramme zur Verfügung gestellt werden. Die Programmwahl erfolgt entweder mit Hilfe eines Tippschalters am Gerät, mit dem die einzelnen Programme zyklisch durchgeschaltet werden können, oder mit Hilfe einer Fernbedienung. Wird in ein anderes Hörprogramm umgeschaltet, bestätigt das Hörgerät die Umschaltung mit einer akustischen Programmkennung (Tonfolge, Sprachansage), die wahlweise deaktiviert und bei manchen Produkten auch individualisiert werden kann.

Die praktische Erfahrung mit Multiprogramm-Hörgeräten zeigt, dass zahlreiche Hörgerätenutzer ihre Hörprogramme häufig nicht entsprechend ihrer Bestimmung nutzen. So zeigen Studien mit Datalogging-Hörgeräten (siehe unten), dass das Basisprogramm, das beim Einschalten der Hörgeräte aktiviert wird, zu 70-80 % der Nutzungsdauer beibehalten wird, auch in Situationen in denen andere Hörprogramme zweckmäßiger gewesen wären. Deshalb bieten die Hörgerätehersteller zunehmend die Möglichkeit der automatischen Programmeinstellung auf der Basis einer Situationserkennung und -klassifizierung (akustische Szenenanalyse) an. Dazu führt das Hörgerät eine kontinuierliche Analyse des Umgebungsschalls durch und bewertet eine Reihe von Schallfeldparametern (Pegel, Frequenz, Modulation, zeitliche Veränderungen, Links-Rechts-Vergleich etc.). Auf dieser Grundlage erfolgt eine Klassifikation der jeweiligen Hörsituation und nachfolgend die Einstellung des zugeordneten Hörprogramms. Neben Hörgeräten mit Festprogrammautomatik gibt es solche, die eine situationsbedingte Optimierung einzelner Ein-

stellparameter vornehmen, also nicht auf einem vorgegebenen Satz von Einstellungen beruhen.

Zahlreiche Hörgerätetypen bieten die Möglichkeit, das Nutzungsverhalten der Geräte intern aufzuzeichnen und zu speichern (Datalogging). In diesen Fällen kann der Hörgeräteakustiker beim nächsten Besuch des Kunden z. B. auslesen, wie lange die einzelnen Hörprogramme in welchen Hörsituationen genutzt worden sind, wie oft manuelle Aktionen (Ein-Ausschalten, Programmwahl) durchgeführt wurden und in welchem Umfang die Verstärkung nachgeregelt wurde. Diese Information kann der Akustiker in zweierlei Hinsicht zur Optimierung der Versorgung nutzen: Zum einen kann er den Hörgeräteträger gezielt beraten und nachschulen, wenn erkennbar ist, dass die Hörgeräte nicht adäquat genutzt worden sind. Ferner kann eine gezielte Nachanpassung durchgeführt werden, wenn die aufgezeichneten Daten zeigen, dass der Hörgerätenutzer systematische Abweichungen von der vorgewählten Einstellung vornimmt.

Daneben sind auch so genannte selbstlernende Hörgeräte auf dem Markt, die auf der Grundlage des Nutzungsverhaltens automatische Korrekturen der Hörgeräteeinstellung vornehmen. Korrigiert der Hörgeräteträger die eingestellte Verstärkung oder den Klang nach jedem Einschalten immer wieder tendenziell in die gleiche Richtung, so regelt ein selbstlernendes System die Verstärkungseinstellung oder den Klang langsam über mehrere Tage nach bis die Einstellung auf den gewünschten Wert konvergiert ist. Dieser Lernprozess erfolgt separat in jedem Hörprogramm, so dass eine programmspezifische Optimierung möglich ist.

### 3. Hörgerätebauformen

#### 3.1 HdO-Hörgeräte

Die am stärksten verbreitete Bauform sind die so genannten Hinter-dem-Ohr-Hörgeräte (HdO), deren Marktanteil in Deutschland bei etwa 90 % liegt. Bei HdO-Geräten ist die Schalleintrittsöffnung des Mikrofons an der Gehäuseoberseite platziert, um einen möglichst ungehinderten Schalleinfall zu ermöglichen. Sofern das Gerät über eine zuschaltbare Richtcharakteristik verfügt, sind mehrere Mikrofonöffnungen vorhanden. Bei klassischen HdO-Geräten ist der Hörer im Gehäuse integriert und der Schall wird über einen Hörgerätewinkel, einen Schallschlauch und eine individuell gefertigte (Abb. 4, links) oder universelle (Abb. 4, Mitte) Otoplastik in den äußeren Gehörgang des Hörgeräteträgers geleitet.

Ferner ist der Signalprozessor im HdO-Gehäuse untergebracht und kann über ein Programmierinterface mit einem Computer verbunden und programmiert werden (siehe Abb. 1). Die

Batterie ist in

der



Abb. 4: Hinter-dem Ohr-Hörgeräte mit individueller Otoplastik (links, transparent zur Darstellung der Komponenten), offenem Standardohrstück (Mitte) und ausgelagertem Hörer im Gehörgang (rechts).

Regel im unteren Teil des Gehäuses platziert, während die Bedienungselemente (Lautstärkesteller, ggf. Programmwahlschalter) auf der Gehäuseoberseite angeordnet sind. Optional kann die Bedienung auch mittels Fernbedienung erfolgen. Automatikgeräte kommen ohne mechanische Bedienungselemente aus.

Für die Wiedergabeeigenschaften von HdO-Hörgeräten ist die akustische Ankopplung, bestehend aus Hörgerätewinkel, Schallschlauch, Otoplastik und Gehörgangsrestvolumen, von be-

sonderer Bedeutung. So kann die Hörgerätewiedergabe durch gezielte Modifikationen am akustischen System (Dimensionierung, Dämpfungselemente, Zusatzbohrungen, Resonanzräume) an die Bedürfnisse des Nutzers angepasst werden. Die wichtigsten und wirkungsvollsten Möglichkeiten, die Wiedergabekurve in definierter Weise zu modifizieren, bieten Zusatzbohrungen (Vents) in der Otoplastik und total offene Versorgungen. Bei Zusatzbohrungen handelt es sich um Bohrungen, meist parallel zum Schallkanal, die das Gehörgangsrestvolumen nach außen eröffnen, so dass Schallkanal, Residualvolumen und Zusatzbohrung gemeinsam einen Hohlraumresonator bilden. Zusatzbohrungen stellen einen akustischen Tiefpass dar, der für niedrige Frequenzen durchlässig und für hohe Frequenzen weitgehend undurchlässig ist. So fließen tieffrequente Schallanteile aus dem Gehörgangsrestvolumen nach außen ab, was zu einer Absenkung der Wiedergabekurve im Frequenzbereich unterhalb 1000 Hz führt.

Heute kann die Anpassung von Hörgeräten in vielen Fällen mit offenem Gehörgang erfolgen. Durch die Verwendung offener Ohrpassstücke wird das Auftreten des Okklusionseffekts (Verschlusseffekt) vermieden, der sich primär durch Unnatürlichkeit der eigenen Stimme äußert, was wiederum häufig zu Akzeptanzproblemen mit den Hörgeräten führt. Grund dafür ist die Tatsache, dass der Schall der eigenen Stimme sowohl auf dem Luftleitungsweg über den äußeren Gehörgang als auch körpergeleitet über die Gehörgangswand ans eigene Ohr gelangt. Im unversorgten Fall fließt ein Teil des körpergeleiteten Schalls nach außen ab und der so entstehende Klangeindruck wird als natürlich empfunden. Wird dagegen ein Hörgerät mit geschlosser Otoplastik angepasst, so fängt sich der körpergeleitete Schallanteil im Gehörgang, weil ein Schallabfluss nicht möglich ist.

Möglich wurden offenen Versorgung durch Fortschritte in der Signalverarbeitung. Effiziente Rückkopplungsunterdrückungssysteme (siehe 2.) erlauben bei offener Versorgung 15-20 dB höhere Verstärkung. Daneben hat die schnelle Signalverarbeitung in modernen Hörgeräten dazu beigetragen, dass die zeitliche Verzögerung des im Hörgerät verarbeiteten Schalls gegenüber dem Direktschall durch das Vent so gering geworden ist (< 5 ms), dass sie vom Hörgeräteträger nicht mehr störend wahrgenommen wird. So konnten sich zwei neue Produktgruppen etablieren, nämlich HdO-Geräte mit:

- dünnem Schallschlauch und offenem Silikon-Standardohrstück (Abb. 4, Mitte) bzw.
- ausgelagertem Hörer im Gehörgang (RIC: Receiver-in-Canal; Abb. 4, rechts).

Geräte mit dünnem Schallschlauch (Mikroschallschlauch) und offenem Silikon-Standardohrstück zeichnen sich durch Unauffälligkeit und hohen Tragekomfort aus. Der wesentliche Vorzug dieser Kopplungssysteme besteht aber darin, dass sie bezüglich der Okklusionswirkung dem offenen Gehörgang sehr nahe kommen. Diesen Vorzügen stehen nur unbedeutende Nachteile gegenüber. So bieten Silikon-Standardohrstücke nicht immer einen sicheren Sitz im Gehörgang. In diesen Fällen können Mikroschlauchsysteme auch mit individuell gefertigten, offenen Otoplastiken ausgestattet werden. Ferner muss bei Mikroschlauchsystemen beachtet werden, dass der Schallschlauch mit abnehmendem Durchmesser eine zunehmende Höhenabsenkung verursacht, die durch entsprechende Verstärkungsanhebung kompensiert werden muss.

Bei HdO-Geräten mit externem Hörer entfällt der Schallschlauch und so werden störende Schlaucheffekte, wie Tiefpasswirkung und Resonanzspitzen vermieden. Durch die Auslagerung des Hörers können die Geräte noch kleiner gestaltet werden. Die offene Anpassung erfolgt bei den Ex-Hörersystemen mit ähnlichen Standardohrstücken wie sie für Mikroschlauchsysteme üblich sind. Auch Systeme mit externem Hörer können mit offenen Maßotoplastiken angepasst werden, um eine bessere Fixierung im Gehörgang zu erreichen. Da der Hörer im Gehörgang

Zerumen- und Feuchtigkeitseinflüssen direkt ausgesetzt ist, muss die Hörer-Kabeleinheit bei Defekten leicht austauschbar sein. Daneben haben Ex-Hörersysteme tendenziell den Nachteil, nicht ganz so offen angepasst werden zu können wie Mikroschlauchsysteme, da der Hörer einen Teil des Gehörgangsquerschnitts verlegt und damit den Grad der Offenheit reduziert.

#### 3.2 Andere Hörgerätebauformen

Im-Ohr-Hörgeräte (IO-Geräte), die in Deutschland derzeit einen Marktanteil von etwa 10 % haben, werden meist als Gehörgangsgeräte ausgeführt (Abb. 5; Mitte und rechts). Nur noch selten kommen größere Bauformen, die die Concha in Teilen ausfüllen (Abb. 5, links), zum Einsatz. Die akustischen Vorteile von IO-Geräten beruhen vornehmlich auf der Platzierung des Mikrofons am Gehörgangseintritt, also der Schallaufnahme am natürlichen Ort, und darauf, dass der Schallschlauch mit den oben genannten Nachteilen entfällt.

Negativ wirkt sich bei IO-Geräten aus, dass komplett offene Anpassungen nicht möglich sind, da die Hörgerätetechnik auch bei extrem kleiner Bauform den Gehörgangsquerschnitt zu großen Teilen ausfüllt. Doch können abhängig von den



Abb. 5: Im-Ohr-Hörgeräte-Bauformen: Conchagerät (links), Gehörgangsgerät (Mitte), Gehörgangsgerät mit speziellem Vent zur Reduktion des Okklusionseffekts (rechts).

Gehörgangsdimensionen auch mit Im-Ohr-Hörgeräten tendenziell offenere Versorgungen realisiert werden, um den Okklusionseffekt zu reduzieren. Das geling mit möglichst großen und kurzen Vents (Abb. 5, rechts), da die akustische Masse des Vents für das Ausmaß des Okklusionseffekts maßgeblich ist. Wenn der Gehörgang für derartige Lösungen nicht genügend groß ist, kann man auf IO-Varianten ausweichen, bei denen entweder das Mikrofon in die Cymba ausgelagert wird, oder das eigentliche Hörgerät in der Cymba platziert ist und über ein Kabel mit externem Hörer im Gehörgang verbunden ist, um den Gehörgang möglichst offen zu halten.

Luftleitungshörbrillen stellen de facto keine separate Form der Versorgung dar, sondern verkörpern HdO-Technologie, die in einem Brillenbügel an Stelle eines HdO-Gehäuses integriert ist. Das kann durch Anbringung von HdO-Geräten mittels Adapter an eine Brille erfolgen, was nur noch sehr selten zum Einsatz kommt. Daneben werden eigens gefertigte Hörbrillenbügel angeboten, die herstellerseitig alle Hörgerätekomponenten enthalten und an nahezu jede Brille angebracht werden können.

In Fällen von Schallleitungsstörungen, bei denen eine Gehör verbessernde Operation nicht indiziert oder nicht gewünscht ist, können Knochenleitungshörhilfen in Erwägung gezogen werden. In ihrer herkömmlichen Form werden Knochenleitungshörgeräte als beidseitige Hörbrille ausgeführt, bei der beide Brillenbügel mit Knochenleitungshörern ausgestattet sind. Bei Kleinkindern, für die eine Brillenlösung noch nicht in Frage kommt, kann die Knochenleitungsversorgung vorübergehend auch mit Hilfe eines Kopfbügels erfolgen. Ein Vorzug von Knochenleitungsversorgungen ist der unverschlossene Gehörgang, der eine gute Belüftung des Gehörgangs sicherstellt, die aber auch mit offener Luftleitungsversorgung erreicht werden kann. Ein kosmetischer Vorzug ergibt sich daraus, dass es bei Knochenleitungsbrillen keines Schallschlauchs bedarf. Dem steht der erforderliche Andruck der Knochenleitungshörer bzw. des Brillenbügels gegenüber, der gelegentlich Kopfschmerz verursacht und in manchen Fällen zur Knochenresorption führen kann. Auch reicht die Verstärkungsleistung von Knochenleitungshörbrillen oft nicht aus.

#### 4. Hörgeräteanpassung und -kontrolle

#### 4.1 Hörgeräteanpassung

Am Beginn einer Hörgeräteanpassung muss die Erfassung des persönlichen Bedarfsprofils stehen. Das persönliche Bedarfsprofil zielt darauf ab, die Nutzungsgewohnheiten, den Anspruch und die Erwartungen zu ermitteln, die der Nutzer an seine Hörgeräte richtet. Auf dieser Grundlage können die Ausstattungsmerkmale der Hörgeräte entsprechend dem persönlichen Bedarf ausgewählt und angepasst werden. Die Erhebung des Bedarfsprofils kann informal im Rahmen eines Beratungsgesprächs erfolgen, aber grundsätzlich ist zum Abschluss der informalen Beratung eine strukturierte Befragung anhand eines Fragebogens zweckmäßig. Das in diesem Kontext wohl am häufigsten eingesetzte Befragungsinstrument ist der so genannte COSI-Fragebogen (Client Oriented Scale of Improvement), mit dem in dieser Phase zunächst die für den Nutzer wichtigsten Kommunikations- bzw. Hörsituationen erfasst werden. Mit strukturierten Befragungsinstrumenten dieser Art kann das Bedarfsprofil eindeutig festgehalten, im Verlauf der Anpassung als Referenz heran gezogen und mit dem Hörgerätenutzer immer wieder besprochen und mit den persönlichen Zielen abgeglichen werden.

Neben der Erfassung von grundsätzlichen Zielen und Wünschen müssen die persönlichen Präferenzen für bestimmte Ausstattungsmerkmale und die Nutzungsgewohnheiten besprochen werden soweit sie für die Typenauswahl relevant sind, so z. B. Kostenaspekte (Anschaffung, Betrieb, Service), Bedienbarkeit, Bauform (Funktionalität, Sichtbarkeit), Stellmöglichkeiten (Verstärkung, Hörprogramme, Klang), Automatikfunktionen (Verstärkungsregelung, Situationserkennung), Fernbedienung, Anschlussmöglichkeiten von Zusatzgeräten (FM, Telefon, TV, HiFi usw.).

An die Erfassung des Bedarfsprofils schließt sich eine Basisanpassung der Hörgeräte an. Wie bereits an anderer Stelle dargelegt, ist die Wiederherstellung der Hörbarkeit das primäre Ziel der gängigen Anpassformeln und -prozeduren, die im Rahmen der Basisanpassung verwendet werden. Damit basiert die Hörgeräteanpassung grundsätzlich auf einem modifizierten Lautheits-Mapping. Trotzdem haben sich lautheitsgestützte Anpassprozeduren für die Basisanpassung nicht durchgesetzt sondern hörschwellenbasierte Formeln, da diese schnell und unkompliziert angewendet werden können. Die schwellenbasierten Verfahren zur Bestimmung des frequenzbezogenen Verstärkungsbedarfs gehen von einem funktionalen Zusammenhang zwischen der Hörschwelle und dem Bereich angenehmen Hörens (MCL: Most Comfortable Level) aus. Auf dieser Grundlage wurden verschiedene generische, schwellenbasierte Anpassformeln für nichtlineare Hörgeräte entwickelt (NAL-NL, DSL [i/o]). Diese Anpassformeln liefern Zielfrequenzgänge für niedrige, mittlere und hohe Eingangspegel und ggf. den maximal zulässigen Ausgangsschalldruckpegel (MPO: Maximum Power Output). Die aktuellen Versionen dieser Anpassformeln (NAL-NL-2 und DSLm [i/o] v5) erlauben eine stärker personalisierte Form der Anpassung (z. B. Alter, Geschlecht, Erfahrung mit Hörgeräten) als die Formeln der ersten Generation.

Um der persönlichen Hörpräferenz gerecht zu werden, die im Einzelfall entweder klang- oder verständlichkeitsorientiert sein kann, bedarf es einer auf der Basisanpassung aufsetzenden Feinanpassung. Auf der Grundlage der Präferenzen des Hörgerätenutzers kann die Feinanpassung der Hörgeräteparameter entweder informal auf der Basis von Erfahrungswerten erfolgen oder mit Hilfe so genannter Anpassmanager (Anpassassistent), wie sie in kommerzieller Anpassoftware angeboten werden. Daneben ist eine Reihe von generischen, also produktunabhängigen Verfahren zur interaktiv-adaptiven Feinanpassung von Hörgeräten entwickelt worden.

Um im Rahmen der Feinanpassung auf die situationsspezifischen Bedürfnisse des Hörgerätenutzers individuell eingehen zu können, kann die Feinanpassung in virtuellen akustischen Situationen durchgeführt werden. Im einfachsten Fall werden dazu Schallbeispiele mit einem Lautsprecher, stereophon oder in Surroundtechnik abgespielt, um für den Nutzer wichtige Hörsituationen im Anpassraum zu simulieren. Durch das zusätzliche Angebot von visueller Information in Form von Videosequenzen kann eine noch realitätsnähere Situation erreicht werden.

Da sich die meisten Hörstörungen schleichend entwickeln und bei den Betroffenen zudem die Neigung besteht, die Hörgeräteversorgungen möglichst lange hinauszuzögern, ist das Gehör zum Zeitpunkt der Versorgung in starkem Maße vom normalen Hören entwöhnt. Deshalb ist es meist nicht sinnvoll, von Anfang an eine komplette Kompensation der Hörstörung anzustreben, da dies von vielen Hörgeräteträgern nicht toleriert wird. So wird zunächst eine reduzierte Gesamtverstärkung, und im Falle typischer Hochtonhörverluste eventuell eine überproportionale Absenkung der Höhenwiedergabe, eingestellt, um eine gute Spontanakzeptanz und ein möglichst regelmäßiges Trageverhalten zu erreichen. Aufbauend darauf wird in den ersten Wochen und Monaten die Verstärkung schrittweise erhöht, um den Hörgeräteträger an die bestmögliche Hörgeräteeinstellung heran zuführen. Diesen Prozess bezeichnet man als gleitende Anpassung, der für Erstversorgungen unverzichtbar ist, aber auch bei Wiederversorgungen eine Rolle spielen kann, wenn der Klang der neuen Hörgeräte vom Nutzer nicht sofort akzeptiert wird. Wenn im Einzelfall erforderlich, sollte die gleitende Anpassung durch Audiotherapie ergänzt werden, um das Hören- und Verstehenlernen mit Hörgeräten systematisch zu fördern. In Deutschland ist die Audiotherapie noch nicht flächendeckend verfügbar, aber das audiotherapeutische Angebot wird ständig ausgeweitet, so dass man zunehmend darauf zurückgreifen kann.

Die praktische Umsetzung der beschriebenen Anpassungsund Feinanpassungsschritte erfolgt durch Programmierung der Hörgeräte mit Hilfe der Anpasssoftware des betreffenden Hörgeräteherstellers (Abb. 6). In der Regel wird die herstellerspezifische Software über die NOAH-Plattform gestartet, über die auch alle gängigen Audiometer und Hörgerätemesssysteme in die Anpassumgebung eingebunden werden können. Von dort aus kann auf alle Kundendaten zuaeariffen werden kann, die in vorher-



Abb. 6: Programmieroberfläche zur Feinanpassung eines mehrkanaligen Hörgerätes.

gehenden Sitzungen gesammelt und gespeichert wurden (Audiogramme, Messungen, Hörgeräteeinstellungen). Aufbau der Anpasssoftware und Gestaltung der Bedienungsoberfläche sind von Hersteller zu Hersteller unterschiedlich, doch gliedert sich die Anpasssoftware typischer Weise in drei Arbeitsphasen (Voreinstellung, Basisanpassung und Feinanpassung), die vom Hörgeräteakustiker sequenziell abgearbeitet werden.

#### 4.2 Hörgerätekontrolle

Ein wichtiges Element im Rahmen der Hörgeräteversorgung stellt die Hörgerätekontrolle im Sinne der Erfassung der Hörgerätewirkung dar. Wichtig ist die Hörgerätekontrolle zunächst für

den Hörgeräteakustiker, der wissen möchte, ob sich die Hörgeräte wie gewünscht verhalten. Ferner ist die Quantifizierung des Versorgungserfolgs von großer Bedeutung für den Hörgeräteakustiker, den Nutzer und den Kostenträger, die sich vom Nutzen der Hörgeräteversorgung überzeugen möchten.

Zum besseren Verständnis ordnet man die Verfahren der Hörgerätekontrolle am besten synoptisch entsprechend ihres Ansatzpunktes entlang des Signalwegs bzw. des aufsteigenden auditorischen Systems:

- Messung des Frequenzgangs am Ohrsimulator (oder Kuppler)
- Messung des Frequenzgangs mittels Sondenmikrofon am Ohr des Nutzers
- Hörschwellenbestimmung mit angepasstem Hörgerät (meist nur bei Kindern)
- Lautheitsskalierung mit angepassten Hörgeräten
- Sprachverständlichkeitsmessungen mit angepassten Hörgeräten
- Frageninventare und Hörtagebücher zu Akzeptanz, Zufriedenheit und Performance.

Die Messung des Hörgerätefreguenzgangs am Ohrsimulator (Messmikrofon in einer Kavität zur Simulation des Gehörgangs) oder mittels Sondenmikrofon am Ohr des Hörgeräteträgers (Abb. 7) und der Vergleich den Zielfremit quenzgängen (siehe 4.1) dient zunächst der Verifikation der erwarten Wirkung des Hörgeräts. Das aleiche ailt für die im peripheren Teil des auditorischen Systems ansetzenden



Abb. 7: Sondenmikrofonmessung (Insitu-Messung) des Frequenzgangs der wirksamen akustischen Verstärkung (REIG: Real Ear Insertion Gain): REAG (links) – REUG (Mitte) = REIG (rechts). Oben ist jeweils die Messsituation, unten der gemessene bzw. berechnete Frequenzgang dargestellt.

Verfahren (Hörschwellenbestimmung, Lautheitsskalierung). Mit der Verifikation können Defizite gegenüber dem beabsichtigten Verstärkungs- und Kompressionsverhalten frequenzbezogen aufgedeckt und korrigiert werden.

Aussagen über die Qualität der Versorgung sind auf dieser Ebene noch nicht möglich. Das ist Gegenstand der Validierung, die sich an die Verifikation anschließt. Validierungsverfahren, wie z. B. Sprachaudiometrie in Ruhe und im Störschall, Bewertung von Klangqualität und Höranstrengung, ermöglichen in der Anpassumgebung eine Über-Alles-Kontrolle (Hörgerät, akustische Kopplung, auditorische Verarbeitung und Wahrnehmung) und geben Hinweise auf den zu erwartenden Hörgerätenutzen im Alltag. Von noch höherem Wert ist die Validierung des Hörgerätenutzens im täglichen Leben auf der Basis informaler oder formaler Befragung (Fragebögen, Hörtagtagebuch) des Hörgeräteträgers und nahe stehender Personen (Partner, Familie). Sofern die Hörgerätekontrolle Mängel erkennen lässt, ist eine weitere Feinanpassung und gleitende Nachanpassung zur Steigerung des Hörgerätenutzens erforderlich. Dabei ist ein begleitendes Monitoring der Feinanpassung mit Hilfe der Sondenmikrofonmessung zweckmäßig. Ergänzende Informationen liefert eine Darstellung der Sprachdynamik (Pegelhäufigkeitsstatistik) am Ausgang des Hörgerätes in Relation zur Hör- und zur Unbehaglichkeitsschwelle (Abb. 8).



Abb. 8: Pegelhäufigkeitsstatistik am Hörgeräteeingang (links) und am Hörgeräteausgang (rechts) für ein Sprachsignal mit Eingangspegel von 65 dB. Dargestellt sind die Perzentilgrenzen für das 30%-, 65%- und 99%-Perzentil in Relation zur Hör- und zur Unbehaglichkeitsschwelle in Absolutdarstellung (links Norm; rechts Fallbeispiel).

#### 5. Hörgeräteindikation und Ablauf der Hörgeräteversorgung

Die Indikationskriterien gründen sich in erster Linie auf den tonaudiometrischen Hörverlust und das Einsilberverstehen ohne Hörgeräte und sind in den Hilfsmittel-Richtlinien niedergelegt (http://www.g-ba.de/downloads/62-492-66/RL-Hilfsmittel-2004-10-19.pdf). Unter tonaudiometrischem Aspekt ist eine Versorgung mit Hörgeräten möglich, wenn in mindestens einer der Prüffrequenzen im Frequenzbereich 500 - 3000 Hz der Hörverlust 30 dB übersteigt. Sofern ein Sprachtest möglich ist, muss sich die Hörgeräteindikation auch an der Sprachverständlichkeit ohne Hörgerät orientieren. Entsprechend den Hilfsmittel-Richtlinien ist das sprachaudiometrische Grenzkriterium überschritten, wenn Einsilber des Freiburger Tests über Kopfhörer bei 65 dB zu 80 % oder weniger verstanden werden. Neben den audiometrischen Kriterien ist es natürlich erforderlich, dass Hörgeräte vom Nutzer akzeptiert werden und bedient werden können.

Im deutschsprachigen Raum erfolgt die Hörgeräteversorgung in der Regel im Zusammenwirken von Hals-Nasen-Ohrenärzten und Hörgeräteakustikern. Dieses duale Versorgungsmodell auf der Basis der klassischen Aufgabenteilung zwischen beiden Berufsgruppen ist für Deutschland in der AWMF-Leitlinie "Hörgeräteversorgung" Nr. 017/065 (www.uni-duesseldorf.de/AWMF/II-na/017-065.htm) festgelegt. Im dualen Versorgungsmodell liegen die Indikationsstellung Verordnung von Hörgeräten in Händen des HNO-Arztes. Im Anschluss daran nimmt der Hörgeräteakustiker die Anpassung der Hörgeräte in mehreren Sitzungen vor (Hörgewöhnung). Neben der individuellen Anpassung der Hörgeräte an die Hörstörung gehören die Einweisung des Hörgeräteträgers in die Benutzung der Geräte, Nachbetreuung, Beratung bezüglich zusätzlicher Kommunikationshilfen, Service- und Reparaturleistungen zu den Aufgaben des Hörgeräteakustikers. Im Falle einer kassenärztlichen Versorgung überzeugt sich der verordnende HNO-Arzt nach Abschluss der Hörgeräteanpassung davon, dass mit den angepassten Hörgeräten eine ausreichende Verbesserung des Kommunikationsvermögens erzielt wird und die Hörhilfen zweckmäßig sind. Die HNO-ärztliche Abschlussuntersuchung stützt sich auf die sprachaudiometrische

Untersuchung mit Hörgeräten sowie auf den persönlichen Eindruck des Arztes vom Versorgungserfolg und der Fähigkeit des Patienten, die Hörgeräte zu handhaben.

# Numerische Akustik

### Ennes Sarradj, Gesellschaft für Akustikforschung Dresden mbH

# 1 Einleitung

Akustischen Messungen und Berechnungen sind mittlerweile in vielen Fällen nicht ohne Einsatz eines Computers vorstellbar. Aus diesem Grund existiert eine Vielzahl so genannter numerischer Verfahren, von denen im folgenden ein kleine Auswahl vorgestellt werden soll. Die Darstellung bleibt dabei auf Verfahren zur Berechnung von Schallfeldern beschränkt. Bild 1 zeigt die verallgemeinerte Aufgabenstellung für diesen Fall. In einem Gebiet, das auch unbeschränkt ausgedehnt sein kann, lasse sich die Physik der Schallausbreitung durch eine Differentialgleichung und durch Randbedingungen beschreiben (Bsp. siehe Bild 1). Aufgabe der Berechnung des Schallfeldes ist es, eine Funktion f zu finden, die sowohl die Differentialgleichung erfüllt als auch den Randbedingungen genügt (Randwertproblem). Zur Lösung lassen sich prinzipiell folgende zwei Wege gehen:

analytisch: exakte Lösung des Randwertproblems:

- nur möglich, wenn der Typ der Funktion *f* bekannt ist (z.B. bei einfachen Geometrien wie Rechteck und Kreis)
- für praktische Anwendungen meist nicht möglich

numerisch: näherungsweise Lösung des Randwertproblems:

- Überführung analytischer Ausdrücke und Zusammenhänge (Ableitungen, Integrale, Differentialgleichungen, ...) in algebraische (Addition, Multiplikation, Gleichungen, ...)
- Lösung algebraischer Gleichungen führt auf Näherungslösung für gesuchte Funktion f

Die Näherungslösung für f wird dabei durch einzelne Funktionswerte gegeben, die an Stützstellen (*Knoten*) berechnet werden. Alle Knoten zusammen bilden ein *Gitter*, mit dem das gesamte Gebiet bzw. nur ein bestimmter Teil davon versehen wird. Somit ergibt sich eine Diskretisierung von f und damit auch eine veränderte Aufgabenstellung. An Stelle der Funktion f sind nun einzelne Funktionswerte gesucht.

Prinzipiell bieten sich folgende Möglichkeiten für den Übergang von Analysis zu Algebra:

1. numerische Differentiation:



Beispiel: DGL: akustische Wellengleichung

 $\Delta p + k^2 p = 0$ 

RB: "schallharter" Rand

$$\frac{\partial p}{\partial \overrightarrow{n}} = 0$$

Bild 1: verallgemeinerte Aufgabenstellung bei der Feldberechnung

- alle Ableitungen werden durch Differenzenquotienten ersetzt
- Differentialgleichung wird zur Gleichung
- Gebiet wird mit einem regelmäßigen (strukturierten) Gitter versehen
- Werte von *f* an den Gitterknoten werden als Lösung eines Gleichungssystems erhalten
- Bsp.: Finite-Differenzen-Methode (FDM)
- 2. numerische Integration:
  - Differentialgleichung und Randbedingungen werden in Gleichungen umgewandelt, die bestimmte Integrale enthalten (Integralgleichungen)
  - (auch unregelmäßiges) Gitter im Gebiet oder auf Teilen davon
  - Integrale werden durch Summen über Funktionswerte von *f* in den Gitterpunkten angenähert
  - Werte von f an den Gitterknoten werden als Lösung eines Gleichungssystems erhalten
  - Bsp.: Finite-Elemente-Methode (FEM), Randelemente-Methode (BEM)

Nicht jedes Verfahren ist für jedes Problem einsetzbar. Welches numerische Verfahren für eine bestimmte Aufgabenstellung zu Einsatz kommt, hängt darüber hinaus auch davon ab, welcher Aufwand mit der Berechnung verbunden ist.

Zwei Verfahren, die für Akustik von besonderer Bedeutung sind, werden im folgenden vorgestellt.



**Bild 2:** Zwei mögliche FEM-Berechnungsgitter zur Berechnung des Schallfeldes in einem vereinfachten Pkw-Innenraum. Links: Hexaeder-Elemente, Rechts: Tetraeder-Elemente

### 2 Finite-Elemente-Methode(FEM)

Bei der FEM wird das Gebiet durch das Gitter in einzelne Teilgebiete, die finiten Elemente, zerlegt (Beispiele siehe Bild 2). Die gesuchte Funktion, z.B. die Schalldruckverteilung  $p(\mathbf{x})$  wird durch Interpolation angenähert und durch die (unbekannten) Werte  $p_i$  in den zum Element gehörenden Knoten ausgedrückt:

$$p(\mathbf{x}) = \sum_{i} N_i(\mathbf{x}) p_i = \mathbf{N}^{\mathrm{T}} \mathbf{p}.$$
 (1)

Die Interpolationsfunktionen  $N_i$  sind die so genannten Formfunktionen, z.B. lineare Funktionen, die so gewählt werden, dass sie jeweils nur in einem Knoten den Wert 1 haben und in allen anderen Knoten verschwinden.

Der nächste Schritt ist die Umformung der Differentialgleichung in eine Integralgleichung, hier am Beispiel der akustischen Wellengleichung (in komplexer Schreibweise):

$$\Delta p + k^2 p = 0. \tag{2}$$

Zunächst wird mit einer beliebigen (beschränkten) Hilfsfunktion *w* multipliziert und dann über das Elementvolumen integriert:

$$\int_{V} w\left(\Delta p + k^{2} p\right) dV = 0 \tag{3}$$

Anschließend wird die erste Greensche Formel zur Umformung dieses Integrals angewendet und außerdem die in (1) angegebene Interpolation für p und ebenso für w eingeführt, so dass sich:

$$\int_{V} \mathbf{w}^{\mathrm{T}} \nabla \mathbf{N} \nabla \mathbf{N}^{\mathrm{T}} \mathbf{p} - k^{2} \mathbf{w}^{\mathrm{T}} \mathbf{N} \mathbf{N}^{\mathrm{T}} \mathbf{p} dV = \int_{S} \mathbf{w}^{\mathrm{T}} \mathbf{N} \nabla \mathbf{p} d \overrightarrow{n}$$
(4)

ergibt. Nun kann die Hilfsfunktion wieder entfernt werden:

$$\left(\int_{V} \nabla \mathbf{N} \nabla \mathbf{N}^{\mathrm{T}} dV - k^{2} \int_{V} \mathbf{N} \mathbf{N}^{\mathrm{T}} dV\right) \mathbf{p} = \int_{S} \mathbf{N} \nabla \mathbf{p} d\,\overrightarrow{n}$$
(5)

Damit ist ein Gleichungsystem für **p** (enthält die Knotenwerte  $p_i$ ) aufgestellt. Die auftretenden Integrale, die nur die Formfunktionen und deren Ableitungen enthalten, werden numerisch berechnet. Die rechte Seite ist als Randbedingung nutzbar und der auftretende Ausdruck  $\nabla$ **p** muss nicht ausgerechnet werden. Da (5) jeweils nur innerhalb eines Elementes gilt, wird die rechte Seite zur Verknüpfung des Gleichungsystems mit dem anderer Elemente genutzt (Die Details dazu sind hier fortgelassen). Somit entsteht ein großes Gleichungsystem, dass gelöst werden kann. Dabei sind in der Akustik prinzipiell zwei Fälle zu unterscheiden:

- keine Berücksichtigung äußerer Anregung: Eigenfrequenzen (Resonanzen) und zugehörige Schwingungszustände werden berechnet (Modalanalyse, Bild 3)
- mit Berücksichtigung äußerer Anregung: Schallfeld infolge einer oder mehrerer Quellen wird berechnet (harmonische bzw. transiente Analyse, Bild 4)

Damit die mit der FEM erhaltene Näherungslösung noch brauchbar ist, dürfen die Elemente nicht zu groß gegenüber der Schallwellenlänge sein. Als Faustregel gilt, dass mindestens 6 Elemente pro Wellenlänge vorgesehen werden müssen. Mit steigender Frequenz bzw. größeren Abmessungen des Berechnungsgebietes steigt der Aufwand stark an und wird bald nicht mehr beherrschbar. Derzeit ist die Behandlung von Problemen mit etwa 10<sup>7</sup> Unbekannten auf größeren Rechnern möglich. Übliche Rechenzeiten liegen für praktische Probleme dabei im Bereich von vielen Stunden bis zu einigen Tagen.

Der für die Akustik wichtige Fall der Schallausbreitung über große Entfernungen (viele Wellenlängen bzw. bis ins Unendliche) ist nicht ohne weiteres mit der FEM zu behandeln, da das gesamte Gebiet mit einem Rechengitter versehen werden müsste und die Anzahl der Unbekannten viel zu groß werden würde. Eine Alternative bietet das im folgenden Abschnitt beschriebene Verfahren.

### **3** Randelemente-Methode (Boundary Element Method, BEM)

Zum Verständnis der diesem Verfahre zugrunde liegenden Idee sei noch einmal die allgemeine Aufgabenstellung nach Bild 1 betrachtet. Oft genügt es, die gesuchte Funktion f (z.B. Schalldruckverteilung **p**) nur in einem Teil des Gebietes zu kennen. In diesem Fall reicht es aus, f nur auf dem Rand des Gebietes zu bestimmen und daraus dann die gesuchten Funktionswerte im Innern des Gebietes zu berechnen.

Ausgangspunkt für ein solches Verfahren ist eine Integralgleichung der Akustik, die so genannte Kirchhoff-Helmholtz-Integralgleichung (KHI). Sie besagt, dass der Schalldruck an einem Punkt  $\overrightarrow{Y}$  innerhalb, auf dem Rand und außerhalb des Gebietes V bestimmt werden kann, wenn



**Bild 3:** Ergebnisse einer Modalanalyse eines Pkw-Innenraums: Schalldruckverteilungen bei vier verschiedenen Eigenfrequenzen



**Bild 4:** Berechnung des Übertragungsverhaltens eines Schalldämpfers Links: verwendetes Berechnungsgitter, Rechts: Übertragungsmaß in dB als Ergebnis einer harmonischen Analyse



**Bild 5:** Außenraumproblem: Durch einen dünnen Schlauch und eine Hüllfläche in unendlicher Entfernung wird das außerhalb des Randes liegende Gebiet zu einem innerhalb des Randes liegenden

die Schalldruckverteilung (1.Term) und die Normalenableitung des Schalldrucks (=Schallschnelle, 2.Term) auf dem Rand des Gebietes *S* bekannt sind:

$$\int_{S} \left( p(\overrightarrow{X}) \frac{\partial G(\overrightarrow{X}, \overrightarrow{Y})}{\partial \overrightarrow{n}} - G(\overrightarrow{X}, \overrightarrow{Y}) \frac{\partial p(\overrightarrow{X})}{\partial \overrightarrow{n}} \right) dS(\overrightarrow{X}) = p(\overrightarrow{Y}) \cdot \begin{cases} 0, & \overrightarrow{Y} \notin V; \\ \frac{1}{2}, & \overrightarrow{Y} \in S; \\ 1, & \overrightarrow{Y} \in V \cap \overrightarrow{Y} \notin S. \end{cases}$$
(6)

*G* ist die Greensche Funktion der akustischen Wellengleichung (für den dreidimensionalen Fall:  $e^{-jkr}/4\pi r$ ).

Bevor aber die Gleichung zur Berechung des gesuchten Schalldrucks verwendet werden kann, müssen der Schalldruck und die Schallschnelle auf dem Rand bekannt sein.

Liegt auch der Punkt  $\overline{Y}$  auf dem Rand, so taucht der gesuchte Schalldruck auf dem Rand auf beiden Seiten der Integralgleichung auf. Wird nun der Rand ähnlich wie bei der FEM in einzelne Knoten und (Rand-)Elemente zerlegt und p mit Hilfe von Formfunktionen interpoliert, dann kann das Integral als Summe ausgedrückt werden. Der Schalldruck an einem bestimmten Randpunkt(=Knoten) ergibt sich dann aus der Summe, in der die Schalldrücke und Schallschnellen an allen anderen Knoten enthalten sind. Werden so alle N Knoten betrachtet ergibt sich ein Gleichungsystem, das N Gleichungen und 2N Unbekannte enthält, nämlich p und  $\partial p/\partial \vec{n}$  an allen Knoten. Mit Hilfe von N Randbedingungen für p oder  $\partial p/\partial \vec{n}$  erhält man die fehlenden N Gleichungen und kann das Gleichungsystem und damit die gestellte Berechnungsaufgabe lösen.

Mit einem kleinen gedanklichen Trick kann man sich klar machen, dass das Verfahren auch funktioniert, wenn das Gebiet *außerhalb* des Randes liegt (Aussenraum-Problem, Bild 5). Dazu wird der Rand um einen kleinen, dünnen Schlauch mit vernachlässigbar kleiner Oberfläche mit einer zweiten Berandung verbunden, die sich unendlich weit weg befindet. Somit befindet sich das Gebiet innerhalb des so neu definierten Randes, jedoch tragen Schlauch als auch unendlich weit entfernte Oberfläche nichts zum Integral in (6) bei.



**Bild 6:** Frontaler Einfall einer ebenen Schallwelle auf ein Modell eines menschlichen Kopfes und Torso. Links: verwendetes Modell und Visualisierungsfläche, Mitte: Berechnungsgitter, Rechts: berechnete Schalldruckverteilung bei 1600 Hz (Quelle: Jan Rejlek "Modellierung der Schallstreuung mit Hilfe der Randelementemethode" Diplomarbeit CVUT Praha 2004)

Besonders die Möglichkeit, Schallfelder außerhalb einer Randfläche mit beherrschbarem Aufwand zu berechnen, macht die BEM zu einem attraktiven Werkzeug der Akustik. Ein Beispiel zeigt Bild 6, wo die Schalldruckverteilung an einem Torso mit Kopf berechnet wurde.

Bei der BEM müssen deutlich weniger Unbekannte bestimmt werden als bei der FEM. Allerdings ist der Aufwand zur Lösung des Gleichungssystem höher. Deshalb sind nur etwa 10<sup>5</sup> Unbekannte praktisch beherrschbar. Um die Vorteile beider Verfahren zu nutzen, lassen sich FEM und BEM auch gekoppelt anwenden.