



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Physiologische Grundlagen der Musikwahrnehmung

Verfasser

Saleh Siddiq

angestrebter akademischer Grad

Magister der Philosophie (Mag.phil.)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 316

Studienrichtung lt. Studienblatt:

Musikwissenschaft

Betreuer:

Univ.-Prof. Dr. Christoph Reuter

Vorwort

Im Jahr 2008 war ich mit meinem Studium der Musikwissenschaft soweit, daß es sinnvoll war, sich ernsthaft mit Diplomarbeitsthemen auseinanderzusetzen. Doch was im Wintersemester 2005 in einem (für Geisteswissenschaften) gutem Tempo begann, endete zunächst im Sommer 2007 am Institut im ersten Stock, letzte Türe rechts. Hätte ich auch nur ansatzweise geahnt, was auf mich zukommen würde, hätte ich mich vermutlich dagegen entschieden, mich als Studierendenvertreter zur Wahl zu stellen. Während der vier Jahre im oder nahe am Zentrum der *Strv Muvi* ist vieles passiert. Die Zeit stand keineswegs still, obwohl der Blick auf den zusehends verstaubenden Diplomandenschreibblock durchaus diesen Eindruck zu vermitteln vermochte. Es ist kaum zu glauben, daß sich unter der wachsenden Staubdecke 2010 tatsächlich so etwas wie ein – damals sehr praktisch anmutender – Themenwechsel vollziehen konnte.

Die Wahl des aktuellen Themas war aber nicht rein praktischer Natur. Mein Interesse an der Funktion des Gehörs trieb mich schon früh in die Lehrveranstaltungen Werner A. Deutschs, es aber endgültig zu meinem „Steckenpferd“ zu machen bedurfte der LV *Neuronale Verarbeitung akustischer Reize* von David Thallinger und meinem Betreuer Univ.-Prof. Christoph Reuter aus dem Wintersemester 08/09. Für mein Umfeld, so schien es mir, war es irgendwie ein logischer Schritt, die Arbeit dahingehend neu auszurichten. Retrospektiv schließe ich mich diesem Eindruck an. Daher wirkt es ein wenig wie Ironie, daß ausgerechnet der neurophysiologische Abschnitt abermals praktischen Überlegungen zum Opfer fiel.

Die gelegentlich auftauchende Frage, was Gehörphysiologie mit Musik zu tun habe, überrascht mich stets aufs neue, denn – auf den Punkt gebracht: ohne Gehör keine Musik. Angesichts dieses, so banalen wie basalen, Zusammenhangs ist es verwunderlich, wie stiefmütterlich ausgerechnet die Musikwissenschaft dieses ansonsten recht gut erforschte Themenfeld behandelt. Der Bedarf, gehörphysiologische Überlegungen in die Musikwissenschaft zu implementieren, ist jedenfalls noch lange nicht gedeckt.

Derart gut erforschte Felder wie die Gehörphysiologie haben den großen Nachteil, daß es praktisch unmöglich ist, im Rahmen von Diplomarbeiten neue wissenschaftliche Erkenntnisse zu generieren. Der große Vorteil ist, daß man aus einem reichen Fundus schöpfen und die Ausrichtung des eigenen Textes recht genau steuern kann. Mein Bestreben war es, möglichst anschauliche, leicht verständliche (und trotzdem detaillierte) Erklärungen zu liefern und dabei gerade solche Punkte in den Fokus zu rücken, die in den meisten aktuellen Darstellungen vereinfacht und/oder vernachlässigt werden, wie es z.B. bei der Hydromechanik oder der Schneckentrennwandschwingung im Innenohr der Fall ist.

Zu guter Letzt bleibt mir, mich bei all den „äußeren Haarzellen“ meiner Schaffenszeit zu bedanken, die meinen Fokus schärften und mich mit wertvollem Feedback versorgten. Besonders möchte ich meinem Betreuer Christoph Reuter danken, der mich einerseits mit einem immensen Vertrauensvorschuß sehr autonom arbeiten ließ, mir aber andererseits zu jeder Zeit motiviert und bereitwillig mit Rat und Tat zur Seite stand, sich von der obengenannten Staubdecke nicht entmutigen ließ, auch meinen Themenwechsel mit Begeisterung unterstützte und nicht zuletzt bei der Literaturbeschaffung eine enorme Hilfe war. Gesondert danken möchte ich an dieser Stelle auch meinen Eltern, deren bedingungsloser Unterstützung ich mir immer sicher sein kann: Es ist mir bewußt, daß dies keine Selbstverständlichkeit ist.

Wien, Dezember 2012

Inhaltsverzeichnis

Vorbereitung

o	Einleitung	I
o.1	Wie ist die Arbeit gegliedert?	I
o.2	Was kann die Arbeit leisten, was nicht?	2
o.2.1	Lesbarkeit	2
o.2.2	Einschränkungen	2
o.2.3	Voraussetzungen	3

Psychoakustik

I	Was hört das Ohr?	5
I.1	Hörbereich	6
I.1.1	Hörschwelle	6
	Verlauf der Hörschwelle	8
I.1.2	Obergrenzen und Dynamik	9
I.1.3	Frequenzbereich	10
I.2	Zeitauflösung des Gehörs	11
I.2.1	Integrationszeiten	12
	Erste Integrationszeit	13
	Zweite Integrationszeit	14
	Dritte Integrationszeit	15
I.2.2	Adaption und Ermüdung	16
	Adaption	16
	Ermüdung	17
I.3	Frequenzgruppen	18
I.3.1	Filterbreiten	19
I.3.2	Filterfunktion	20
I.4	Lautstärke	22
I.4.1	Reizgröße und Empfindungsgröße	22
I.4.2	Lautstärkewahrnehmung	25
I.5	Tonhöhe	26
I.5.1	Reizgröße und Empfindungsgröße	26
I.5.2	Tonhöhenwahrnehmung	28
	Unterschiedsschwellen	29
	Einfluß der Lautstärke	31
I.5.3	Tonhöhe von Klängen	31

1.6	Verdeckung	33
1.6.1	Mithörschwelle	33
1.6.2	Simultanverdeckung	36
1.6.3	Zeitliche Verdeckung	37
1.7	Lokalisation	39
1.7.1	Orientierung im Hörraum	40
	Lokalisationsunschärfe	41
1.7.2	Distanzhören	43
1.7.3	Binaurales Hören	44
	Binaurale Fusion	45
	Interaurale Differenzen	46
	Binaural Masking Level Difference	47

Anatomie

2	Wie ist das Ohr aufgebaut?	49
	<i>Exkurs: Zur Orientierung im Körper (anatomische Lagebezeichnungen)</i>	50
2.1	Schläfenbein	51
2.2	Außenohr	53
2.2.1	Ohrmuschel	53
2.2.2	Gehörgang	54
2.3	Mittelohr	55
2.3.1	Trommelfell	56
2.3.2	Mittelohrhöhle	59
	Paukenhöhle	59
	Warzenfortsatzzellen	61
	Eustachische Röhre	61
2.3.3	Gehörknöchelchen	63
	Hammer	63
	Amboß	64
	Steigbügel	65
2.3.4	Mittelohrmuskeln	66
2.4	Innenohr	67
2.4.1	Labyrinth	68
	Knöchernes Labyrinth	68
	Häutiges Labyrinth	70
2.4.2	Schnecke	71
	<i>Exkurs: Zur dargestellten und zur tatsächlichen Ausrichtung der Schnecke</i>	72
	Knöcherne Schnecke	72
	Häutige Schnecke	75
2.4.3	Basilarmembran	77

2.4.4	Cortiorgan	79
	Tektorialmembran	81
	Haarzellen	82
2.4.5	Lymphflüssigkeiten	85
	Kochleäre Potentiale	88

Physiologie

3	Wie arbeitet das Ohr?	91
3.1	Kopfbezogene Übertragungsfunktion	91
3.1.1	Kopf und Körper	92
	Interaurale Differenzen	93
3.1.2	Außenohr	97
	Monoaurales Hören	97
	Übertragungsfunktion des Außenohrs	98
3.2	Impedanzwandlung	100
	<i>Exkurs: Zum physikalischen Konzept der Impedanz</i>	101
3.2.1	Luftschalleitung (normaler Weg)	102
	Mittelohrimpedanz	103
	Druckausgleich	104
	Trommelfellbewegung	105
	Mechanik der Gehörknöchelchen	107
	Übertragungsfunktion des Mittelohrs	110
3.2.2	Stapediusreflex	111
	Aufgaben des Stapediusreflexes	113
	Funktion des Stapediusreflexes	114
3.2.3	Knochenschalleitung (alternativer Weg)	116
3.3	Mechanoelektrische Reizwandlung	117
3.3.1	Entstehung der Wanderwelle	118
3.3.2	Physik der Wanderwelle	121
	Frequenzverteilung auf der Basilarmembran	123
	Laufzeit der Wanderwelle	126
	Amplituden der Basilarmembran	127
3.3.3	Mechanik des Cortiorgans	129
3.3.4	Erregung der Haarzellen	131
	Reizwandlung	134
3.3.5	Kochleärer Verstärker	135
	Otoakustische Emissionen	136
	Elektromotilität der äußeren Haarzellen	137
	Funktion des kochleären Verstärkers	139

Fazit

4	Zusammenfassung	143
4.1	Auswirkungen auf die musikalische Praxis	144
4.2	Schlußbetrachtung	147

Anhang

5	Verwendete Literatur	149
---	----------------------------	-----

Vorbereitung

o Einleitung

Die vorliegende Arbeit beschäftigt sich mit den *physiologischen Grundlagen der Musikwahrnehmung*. Sie folgt akustischen Signalen auf ihrem Weg durch das *periphere auditive System* (ungeachtet des Hörnervs), beschreibt die anatomischen Strukturen der durchlaufenen Ohrabschnitte und versucht, die Vorbereitung und Bearbeitung akustischer Reize bis zum Transduktionsprozeß im Innenohr (wonach sie dann vom zentralen auditiven System „gehört“, d.h. ins Bewußtsein geführt werden) detailliert zu erklären. Der Fokus liegt auf den physiologischen Abläufen. Besonderes Interesse gilt der Impedanzwandlung durch die Strukturen des Mittelohrs und der Schallaufnahme durch die Sinneszellen mittels komplexer mechanischer und bioelektrischer Prozesse im Innenohr.

o.1 Wie ist die Arbeit gegliedert?

Der Hauptteil gliedert sich in drei Kapitel, wovon die beiden ersten (*Psychoakustik, Anatomie*) unabhängig voneinander auf das letzte (*Physiologie*) vorbereiten. Es schien mir sinnvoll, psychoakustische Erkenntnisse und anatomische Beschreibungen in eigene Kapitel auszulagern, um das Hauptthema – bildlich gesprochen – nicht durch belastende (aber notwendige) „Seitenbänder zu verzaubern“.

Das Kapitel *Psychoakustik* erörtert die Frage, was das Ohr überhaupt hört. Anders gesagt: Es stellt die wichtigsten ohrbezogenen Phänomene des Hörens vor. Von im Ohr begründeten Systemeinschränkungen (*Hörbereich*) entwickelt es sich über die wichtigsten Integrationswerkzeuge des Gehörs (Zeitkonstanten, Frequenzfilter) sowie dessen wichtigste Empfindungsgrößen (*Lautheit, Tonheit*) hin zu übergeordneten Phänomenen wie *Verdeckung* und *Lokalisation*.

Im Kapitel *Anatomie* wird – man ahnt es bereits – der anatomische Aufbau des Ohrs (ganz bewußt sehr detailreich) beschrieben. In diesem Kapitel weiche ich von meiner Praxis, soweit als möglich deutschen Begriffen den Vorzug zu geben ab, um mich der Eindeutigkeit anatomischer Lagebezeichnungen zu bedienen, die in einem kurzen Exkurs vorgestellt werden. Der weitere Textverlauf klärt grob die anatomische Makrosituation (Lage des Ohrs im *Schläfenbein*) und folgt danach dem Schallaufnahmepfad (in bewährter Weise) vom *Außenohr* übers *Mittelohr* zum *Innenohr*.

Das letzte Kapitel *Physiologie* geht der Hauptfrage nach, wie das Ohr arbeitet. Analog zum zweiten Kapitel dringt es von außen nach innen vor, indem es sich an den Aufgaben der Ohrabschnitte, richtungsbezogener Schallaufnahme (*kopfbezogene Übertragungsfunktion*), *Impedanzwandlung* und *mechanoelektrischer Reiztransformation* orientiert.

Im *Fazit* wird schließlich die Essenz aus den drei Hauptkapiteln gebildet (*Zusammenfassung*) und auf die natürlichen Hörsituationen des musikalischen Alltags bezogen. Die Angaben zur verwendeten Literatur, sowie alle von der Universität Wien geforderten – nicht textrelevanten – Pflichtbeiträge, befinden sich im *Anhang*.

o.2 Was kann die Arbeit leisten, was nicht?

Die Arbeit richtet sich an Musikwissenschaftler, die nicht vorhaben, die Grundlage aller Musik – das Ohr – außer Acht zu lassen. Ich habe mich daher bemüht, den Text möglichst frei von mathematischen Formeln zu halten und das „Medizinerlatein“ in Nebensätze und Klammern zu verbannen, ohne dabei die geläufigen Fachbegriffe zu unterschlagen. Darüber hinaus habe ich versucht, alle verständnisirrelevanten Herleitungen, Versuchsanordnungen, Experimente und Theorien auszublenden und den Text auf das Wesentliche zu reduzieren. Alles ist sozusagen dem Ziel untergeordnet, ein plausibles und verständliches Gesamtbild zu schmieden.

o.2.1 Lesbarkeit

Um den Text gut lesbar und verständlich zu halten, wurden vor allem zwei Maßnahmen getroffen. Erstens soll die terminologische Aufrüstung, wie sie in vielen einschlägigen Aufsätzen (sicher auch platzbedingt) zu beobachten ist, durchbrochen werden, indem weitestgehend deutschsprachige Fachbegriffe (die es zu fast allem gibt) kultiviert werden. Zweitens kommt der Text größtenteils ohne fachbegriffliche Abkürzungen aus. Nur sehr wenige, lange, sperrige und/oder extrem häufige Begriffe werden stellenweise abgekürzt, da dies dem Lesefluß m.E. förderlich ist. Um der Lesbarkeit abträgliche Konstruktionen zu vermeiden, sind außerdem alle im Text verwendeten Begriffe, die Personen bezeichnen, als geschlechtsneutral zu verstehen, sofern sie sich nicht auf konkrete Personen beziehen.

Im Text werden zwei verschiedene Auszeichnungen verwendet, um eine Reihe von Hervorhebungen abzudecken. *Kursive Schriftschnitte* markieren in erster Linie Fachbegriffe, sowie (im Text genannte) Titel und Überschriften. Ist ein Begriff *nicht* explizit als Fachterminus gemeint, entfällt die Auszeichnung. Ebenfalls kursiv stehen – wie im vorigen Satz zu sehen – betonte Worte um die Aufmerksamkeit zu lenken, außerdem mehrzeilige wörtliche Zitate. Alternativ werden die klassischen Anführungszeichen verwendet. Sie kommen zum Einsatz um dezidiert von Fachbegriffen zu unterscheiden, metaphorische, alltagssprachliche oder saloppe Formulierungen hervorzuheben und kürzere wörtliche Zitate zu kennzeichnen.

o.2.2 Einschränkungen

Meistens sind die Sachverhalte dargestellt, ohne ihre Erforschungsgeschichte und/oder Herleitung aufzurollen. Es war mir überdies wichtig, mich nicht in wissenschaftlichen Theorien zu verspinnen. Daher wurden Theorien und divergierende Gedanken auf gesicherte Gemeinsamkeiten reduziert und nur in den Text aufgenommen, wenn sie der Verständlichkeit dienen.

Ausschließlich aus Tierversuchen gewonnene Informationen werden im Text in aller Regel nicht berücksichtigt. Weniger wegen der schwierigen ethischen Frage – die zu diskutieren hier nicht der richtige Ort ist –, sondern weil man oft nicht abschätzen kann, in welchem Ausmaß die Ergebnisse auf menschliche Hören übertragbar sind. Generell habe ich solche Informationen nur dann herangezogen, wenn sie halfen, verständnisrelevante Lücken zu füllen.

Die vorliegende Arbeit ist auf das *periphere auditive System* exklusive Hörnerv (vgl. Viemeister & Plack 1993, 116), d.h. das Ohr beschränkt und läßt somit alle im zentralnervösen auditiven System ablaufenden Verarbeitungsschritte bewußt außer Acht. Bedenkt man zudem, daß die Schnecke bereits im 16. Jahrhundert entdeckt und erste hörtheoretische Überlegungen aus dem 17. Jahrhundert überliefert sind (vgl. Békésy 1941(b), 265), wird offensichtlich, daß kein Anspruch auf Vollständigkeit erhoben werden kann.

0.2.3 Voraussetzungen

Mit Blick auf den musikwissenschaftlichen Zielgruppenkern, werden bestimmte Grundlagen als bekannt vorausgesetzt. Dies betrifft: erstens, musikalische Grundkenntnisse, speziell bezüglich des „westlichen“ Tonsystems, wie es z.B. in der europäischen Kunstmusik und den meisten aktuellen Formen von Jazz und Populärmusik verwendet wird. Zweitens, Grundlagen der Akustik inklusive aller gängigen Maße und Skalen (z.B. Herleitung des Dezibels, Herleitung von Hertz, usw.). Drittens werden Begriffe aus der Klangforschung (Spektren, Formanten, Einschwingvorgang, etc.) nicht explizit erklärt.

Ich habe versucht, Kenntnisse aus allen drei Bereichen nicht all zu sehr zu strapazieren. Am wenigsten war dies im Feld der Akustik möglich, da akustische Signale die adäquaten Reize des Hörorgans und damit (implizit) unumgänglicher Teil dieser Arbeit sind. Einfach gehaltenene Erklärungen vieler Grundbegriffe findet man zur Not u.a. bei Fasold & Veres (2003, 15–26) Gelfand (2004, 1–35), Hellbrück und Ellermeier (2004, 18f., 53–70), Müller und Frings (2009, 443–448), Spitzer (2002, 23–48) und Yost (2008, 10–36).

Im Spannungsfeld verschiedener Disziplinen muß man sich immer im Klaren darüber sein, daß scheinbar eindeutige Begriffe fachabhängig ganz verschieden verwendet sein können. In dieser Arbeit sind davon besonders die Begriffe *Ton* und *Klang* betroffen. Musiker verstehen *Klang* normalerweise als Synonym für den Begriff *Akkord*, der seinerseits aus verschiedenen *Tönen* zusammengesetzt ist. Auch in der Physik setzt sich ein Klang aus verschiedenen Tönen zusammen, nur bezeichnet der Begriff *Ton* einen *Sinuston*. Mehrere Sinustöne (Teiltöne) schichten sich zu einem teiltonreichen Spektrum, einem *Klang* auf, der nun genau das darstellt, was Musiker als *Ton* bezeichnen (vgl. Dickreiter 1987, 53; Zwicker & Feldtkeller 3f.). Die folgende Tabelle soll den Sachverhalt verdeutlichen.

Physik	Musik
Ton, Sinuston (reiner Ton)	—
Klang (komplexer Ton)	Ton
Klanggemisch	Akkord

Tab. 0.1 **Physik und Musik** (Dickreiter 1987, Tab. 2.1, 54)

Da besonders die Psychoakustik viel mit Sinustönen arbeitet und nur die physikalische Terminologie über alle erforderlichen Begriffe verfügt, wird in dieser Arbeit grundsätzlich die physikalische Terminologie angewendet, solange dies nicht explizit anders angegeben ist.

Psychoakustik

1 Was hört das Ohr?

Empfindungen korrespondieren nicht 1:1 mit den auslösenden physikalischen Reizen. Zum Glück. Andernfalls wäre Fernsehen undenkbar, jedes mit Wechselstrom betriebene Licht würde flackern (nicht nur für Epileptiker problematisch) und auch die akustische Umwelt wäre unerträglich. Jede minimale Schwankung bis zu molekularem Rauschen wäre in einer derartigen Präzision hörbar, daß an (selbst einstimmige) Musik (Stichwort Ungenauigkeiten) nicht zu denken wäre (vgl. Gelfand 2004, 243; Spitzer 2002, 50, 78; Winckel 1952, 53). Die *Psychophysik*, Teildisziplin der Psychologie, untersucht die gesetzmäßige, funktionale Perzeption physikalischer Reize über die adäquaten Sinneskanäle. Jener Teil der Psychophysik, der sich auf Schallrezeption spezialisiert hat, wird *Psychoakustik* genannt (vgl. Bregman 1990, 1; Gelfand 2004, 243; Hellbrück & Ellermeier 2004, 18, 39f.; 2008, 52; Müller & Frings 2009, 417; Plattig in Keidel 1975, 359; Terhardt 1998, 11; Yost 2008, 13f., 275; Zwicker & Feldtkeller 1967, 42).

Da Empfindungen nicht direkt feststellbar sind, müssen Reiz und Empfindung voneinander getrennt werden. Stattdessen werden *Reiz-* und *Empfindungsgrößen* aufeinander bezogen. Als Reizgrößen werden die physikalischen Parameter des Reizes bezeichnet (z.B. Schallpegel, Frequenz). Sie definieren das *Schallereignis*. Als *Empfindungsgrößen* bezeichnet man voneinander trennbare, abstufbare sensorische Kontinua (z.B. *Lautheit*, *Tonheit*). Sie hängen immer – unterschiedlich stark – von *allen* Reizgrößen ab (z.B. Lautheit primär vom Schallpegel, sekundär aber auch von der Frequenz/Bandbreite). Empfindungsgrößen definieren die auditive Perzeption, das *Hörereignis* (vgl. Blauert 1947, 2; Hellbrück & Ellermeier 2004, 78f.; Müller & Frings 2009, 415, 417; Zwicker & Feldtkeller 1967, 43). Vergleiche (z.B. (wieviel) leiser/lauter als) und Schwellenmessungen liefern Informationen über die subjektiven Empfindungen. Sie gehören zu den Hauptaufgaben der Psychoakustik (vgl. Hellbrück & Ellermeier 2008, 52; Müller & Frings 2009, 417). Als Schwelle wird jene Reizstärke festgelegt, ab der Empfindungen überhaupt auftreten oder sich verändern. Sie werden entsprechend *Reizschwellen* oder *Unterschiedsschwellen* (*ebenmerklicher Unterschied*, EMU; just noticeable difference, jnd; difference limen, Δ^1) genannt (vgl. Gelfand 2004, 245, 248, 288; 2009, 89; Hellbrück & Ellermeier 2004, 78; Plattig in Keidel 1975, 360; Yost 2008, 150). Angaben von Grenz- bzw. Schwellwerten sind den Erkenntnissen der Psychoakustik also praktisch intrinsisch und ließen sich in diesem Kapitel folglich nicht vermeiden.

Um es auf den Punkt zu bringen: Die Psychoakustik befaßt sich mit Ursache (Reiz) und Wirkung (Empfindung) des Hörens. Das Ohr/Gehör selbst ist aus psychoakustischer Sicht eine „Blackbox“, über deren Inhalt (die physiologischen Zusammenhänge von Reiz und Empfindung) keine Aussagen getroffen werden.

Zunächst sollen kurz drei fundamentale Überlegungen zur schwellenbasierten Empfindungsmessung vorgestellt werden. Einerseits, weil sie zum Rüstzeug versierter Psychoakustiker (und Musikwissenschaftler) zählen, andererseits, weil es sich nicht vermeiden ließ, an einigen Stellen auf sie Bezug zu nehmen. Ernst Heinrich Weber (1795–1878) stellte fest, daß das Verhältnis von eben erkennbarem Reizunterschied (ΔI) und Ausgangsgröße des Reizes (I) konstant (k) ist. Die Formel $\Delta I : I = k$ ist heute

1 Der griechische Großbuchstabe *Delta* (Δ) wird normalerweise als Symbol für *Differenz* bzw. *Unterschied* verwendet (vgl. Gelfand 2009, 89), daher wird es in weiterer Folge in vielen unterschiedsbezogenen Formelzeichen auftauchen.

als *Webers Gesetz* bekannt (vgl. Gelfand 2004, 288f.; 2009, 89; Green 1993, 22; Guski 1996, 21; Hellbrück & Ellermeier 2004, 216; Müller & Frings 2009, 415; Plattig in Keidel 1975, 359f.; Yost 1993, 5; 2008, 150). Gustav Theodor Fechner (1801–1887) erkannte, daß gleichbleibende Reizverhältnisse (geometrische Reihe) gleichbleibende Empfindungsschritte (arithmetische Reihe) evozieren. Seine *psychophysische Fundamentalformel* (von Fechner selbst als „Weber’sches Gesetz“, später als *Fechners Gesetz* bezeichnet) $\log(I : I_0) \cdot k = E$ ist heute als *Weber-Fechner Gesetz* bekannt (vgl. Green 1993, 26; Guski 1996, 21, 23; Müller & Frings 2009, 415; Plattig in Keidel 1975, 361). Im 20. Jahrhundert stellte Stanley Smith Stevens die (seiner Ansicht nach gesetzmäßige) Verbindung zwischen Reizgröße und Empfindungsgröße über eine Potenzfunktion her. Er bezeichnete die Formel $E = k \cdot (I - I_0)^n$ (in verkürzter Form: $E = I^n$; zur Erklärung vgl. Plattig in Keidel 1975, 361) als *psychophysisches Potenzgesetz* (power law). Sie gilt für alle Intensitätsempfindungen (Helligkeit, Lautstärke, Temperatur, etc.) und wird mittels reizspezifischer Exponenten (die Größe der Hochzahl n gibt das Wachstumsverhältnis von Reiz und Empfindung wieder) an die Sinnesqualität angepaßt (vgl. Gelfand 2004, 352; 2009, 94; Hellbrück & Ellermeier 2004, 84; Plattig in Keidel 1975, 361; Yost 2008, 279).

1.1 Hörbereich

Schall wird nicht einfach gehört. Schall muß bestimmte Kriterien erfüllen, um auditive Empfindungen des Ohrs zu wecken. Schall muß eine gewisse Lautstärke erreichen, darf aber nicht zu laut werden und er muß innerhalb gewisser Frequenzgrenzen liegen, um vom Gehör erfaßt zu werden. Traditionellerweise wird in einem zweidimensionalen Koordinatensystem aus Frequenz und Pegel die Fläche zwischen Hörschwelle, Schmerzgrenze und den Frequenzgrenzen als *Hörbereich* oder *Hörfeld* bezeichnet (vgl. Fastl & Zwicker 2006, 18; Fricke & Louven 2008, 415; Hellbrück & Ellermeier 2004, 72; 2008, 55).

Nun wird das Gehör in der Praxis selten bis an seine Grenzen geführt. Die z.B. für Sprache und/oder Musik genutzten Bereiche fallen wesentlich kleiner als der gesamte Hörbereich aus und siedeln sich dort an, wo das Hören am besten aufgelöst, d.h. das Ohr am empfindlichsten ist. Wie Abbildung 1.1 zeigt, wird für Sprache nur in ein recht enger Pegel- und Frequenzbereich verwendet. Musik nutzt beide Spannen zwar weiter aus, stößt aber ebenfalls nicht an die Grenzen des Hörbereichs (vgl. Fricke & Louven 2008, 416).

1.1.1 Hörschwelle

Die Hörschwelle ist die Grenze, an der sich die unhörbaren von den hörbaren Pegeln scheiden (vgl. Fricke & Louven 2008, 416; Hellbrück & Ellermeier 2004, 67; Yost 2008, 144, 147; Zwicker & Feldtkeller 1967, 51). In stiller Umgebung, wenn nur ein Nutz- aber keine Störsignale vorhanden sind, wird die Hörbarkeitsgrenze des Nutzsignals als *absolute Hörschwelle* oder *Ruhehörschwelle* bezeichnet (vgl. Fricke & Louven 2008, 416; Hellbrück & Ellermeier 2004, 67, 69, 129; Zwicker & Feldtkeller 1967, 55). Die Ruhehörschwelle ist also eine *Reizschwelle*, d.h. eine Schwelle, ab der eine Empfindung überhaupt ausgelöst wird (vgl. Hellbrück & Ellermeier 2004, 70, 78; 2008, 53). Normale Hörsituationen

sind alles andere als ideal, Störgeräusche sind eigentlich immer vorhanden. Die dadurch zu höheren Pegeln verlagerte Schwelle wird als *Mithörschwelle* (siehe 1.6.1, S. 33) bezeichnet (vgl. Fricke & Louven 2008, 416; Hellbrück & Ellermeier 2004, 45, 67, 69, 129; Zwicker & Feldtkeller 1967, 55f.).

Hörschwellen sind interindividuell und intraindividuell verschieden, denn auch bei einer Person weichen die Empfindlichkeiten von linkem und rechtem Ohr i.d.R. zumindest minimal voneinander ab. Dasjenige Ohr mit der größeren Empfindlichkeit bestimmt die subjektive Hörschwelle. Sollten tatsächlich beide Ohren gleich empfindlich sein (was über weitere Frequenzbänder praktisch auszuschließen ist), würde die Hörschwelle noch etwas weiter absinken. Das Summationsverhältnis von monoauraler zu binauraler Beschallung liegt praktisch bei 2:1, wodurch die binaurale Hörschwelle um 2,5 bis 3 dB unter der monoauralen liegt (vgl. Blauert 1974, 209f.; Gelfand 2004, 398ff.; 2009, 104; Hellbrück & Ellermeier 2008, 61; Zwicker & Feldtkeller 1967, 55). Die gemessenen Werte, durch welche eine allgemeine Hörschwelle festgelegt wird, sind über viele (pathologisch unauffällige) Durchschnittsohren gemittelt. Individuelle Hörschwellen sehen meistens wesentlich unregelmäßiger aus (vgl. Gelfand 2004, 280, 282). Auch bei Mittelwertkurven ist es wichtig, das Alter der Versuchspersonen zu berücksichtigen, weil die Hörschwelle sich im Alter stark verändert. Für hohe Frequenzen werden alternde Ohren immer unempfindlicher, wohingegen sich die Hörschwelle im tieffrequenten Bereich unter 1 bis 2 kHz kaum verschiebt (vgl. Dickreiter 1987, 110; Fastl & Zwicker 2006, 22; Fricke & Louven 2008, 416; Hellbrück & Ellermeier 2004, 182ff.; Stevens & Hallowell 1938, 67; Zwicker & Feldtkeller 1967, 55).

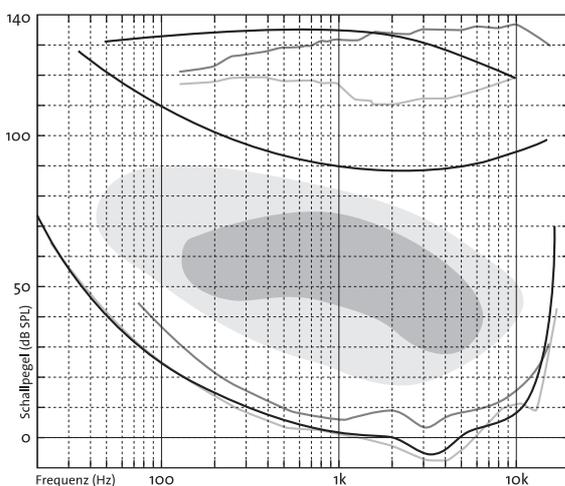


Abb. 1.1 **Hörraum und Hörschwellen** (modifiziert n. Fastl & Zwicker 2006, Abb. 2.1, 17; Gelfand 2004, 9.1, 281; Yost 2008, Abb. 10.2, 154; Zwicker & Feldtkeller 1967, Abb. 18.3, 54). Hörschwelle: Ruhehörschwelle (nach Zwicker & Feldtkeller, schwarz), MAF (hellgrau), MAP (dunkelgrau). Hörraum: Ungefährerer Nutzungsbereich von Musik (hellgrau) und Sprache (dunkelgrau). Obergrenzen: Schädigungsschwelle (in der Mitte um 90 dB, schwarz), Unbehaglichkeitsschwelle (hellgrau), Schmerzgrenze n. Yost (dunkelgrau) bzw. Fastl & Zwicker (schwarz).

Aus Abbildung 1.1 ist ersichtlich, daß man unterscheiden muß, ob die Hörschwelle im freien Schallfeld (über Lautsprecher) oder über Kopfhörer ermittelt wurde. Über Kopfhörer dargeboten, wird der Schalldruck monoaural effektiv zwischen Kopfhörermembran und Trommelfell gemessen, woraus die Bezeichnung *kleinster hörbarer Schalldruck* (minimal audible pressure, MAP) entstand. Bei Lautsprecherdarbietung wird zunächst die binaurale Hörschwelle der Versuchsperson und in einem zweiten Durchgang an Stelle der (nun abwesenden) Versuchsperson der Pegel des Freifeldschalls gemessen, weshalb vom *kleinsten hörbaren Schallfeld* (minimal audible field, MAF) gesprochen wird (vgl. Gelfand 2004, 280; 2009, 87; Hellbrück & Ellermeier 2004, 71; 2008, 53f.; Stevens & Hallowell 1938, 42f., 47; Yost 2008, 144). Daß die MAP-Schwelle immer durchschnittlich 6 bis 10 dB über der MAF-Schwelle lag, konnte lange Zeit nicht erklärt werden, was den Begriff der „fehlenden 6 dB“ (*missing 6 dB*) hervorbrachte (vgl. Gelfand 2004, 281; 2009, 87; Yost 2008, 145). Berücksichtigt

man allerdings die bei MAF-Messungen wirksame *kopfbezogene Übertragungsfunktion* (kobeÜF, siehe 3.1, S. 91), die Impedanzen des Ohrs und achtet auf korrekte und kompatible Kalibrierung bei den Messungen, kann die MAF/MAP-Differenz auf höchstens 2,5 dB reduziert werden, die sich ihrerseits leicht mit *binauraler Summation* (siehe S. 7) erklären lassen (vgl. Gelfand 2004, 281f.; 2009, 87; Hellbrück & Ellermeier 2008, 54; Yost 2008, 145, 155).

Verlauf der Hörschwelle

Schalldrücke werden meist in Pegeln angegeben. Pegel sind dimensionslos. D.h. sie sind keine absoluten, sondern auf bestimmte Normwerte bezogene Relativangaben (vgl. Müller & Frings 2009, 447). Als Bezugsschalldruck (Druck = N/m² = Pascal, Pa) für die physikalische bzw. absolute Hörschwelle wurden 20 µPa ($2 \cdot 10^{-5}$ Pa) festgelegt. Über diesen Bezugsschalldruck definierte Pegel werden als *Schalldruckpegel* (sound pressure level, SPL, oft auch kurz *Schallpegel*) bezeichnet und in *Dezibel SPL* (dB_{SPL}) angegeben. Der Bezugsschalldruck von 20 µPa wird oft als derjenige angegeben, bei dem ein 1 kHz Ton gerade eben wahrzunehmen ist, also überschwellig wird (vgl. Dickreiter 1987, 110; Fastl & Zwicker 2006, 18; Fricke & Louven 2008, 419; Gelfand 2004, 8f.; 2009, 28, 86f.; Hellbrück & Ellermeier 2004, 67f.; Lehnhardt & Laszig 2009, 4f.; Müller & Frings 2009, 447f.; Plattig in Keidel 1975, 371f.; Yost 2008, 146), wobei die tatsächliche Hörschwelle bei 1 kHz normalerweise ein wenig höher, bei 3 bis 4 dB_{SPL} liegt (vgl. Dickreiter 1987, 110; Fastl & Zwicker 2006, 21, 54; Plattig in Keidel 1975, 371).

Wie aus Abbildung 1.1 deutlich hervorgeht, ist die Hörschwelle nicht flach, d.h. das Gehör nicht für alle Frequenzen gleich empfindlich. Vielmehr nimmt die Hörschwelle einen U-förmigen Verlauf. Z.B. ist der notwendige Schalldruck, um einen Ton von 25 Hz hörbar zu machen 1000 mal höher, als der für einen 1 kHz Ton (vgl. Békésy 1960, 429; Fasold & Veres 2003, 47; Fricke & Louven 2008, 416; Gelfand 2009, 87; Plattig in Keidel 1975, 372). Am empfindlichsten ist das Gehör ungefähr von 2 bis 5 kHz, wo jeweils 0 dB_{SPL}, dazwischen sogar negative Dezibelwerte erreicht werden (vgl. Fastl & Zwicker 2006, 20; Gelfand 2004, 280; Zwicker & Feldtkeller 1967, 54). Wie sensibel das Ohr in diesem Bereich tatsächlich ist wird klar, wenn man sich vor Augen führt, daß es sich bei 0 dB_{SPL} um Amplituden im Größenbereich des Durchmessers von Wasserstoffatomen handelt (vgl. Gelfand 2004, 279). Wäre das Gehör noch empfindlicher, könnten wir molekulares Rauschen hören (vgl. Spitzer 2002, 50, 78; Stevens & Hallowell 1938, 357f.). Unter 1 kHz nimmt die Empfindlichkeit mit rund 6 dB/Oktave ab, über 4 kHz sogar mit 24 dB/Oktave (vgl. Yost 2008, 146). Im Bereich von 500 Hz bis 5 kHz ist das Gehör noch sehr empfindlich, im erweiterten Abschnitt von 100 Hz bis 10 kHz immerhin noch ausreichend sensibel. Erst darüber und darunter nimmt die Empfindlichkeit dramatisch ab, so daß die äußersten Flanken gleichzeitig praktisch die Grenzen der höchsten und tiefsten verwertbaren Frequenzen darstellen (vgl. Fricke & Louven 2008, 416; Gelfand 2004, 280; 2009, 87; Hellbrück & Ellermeier 2004, 67).

Es sollte nicht unerwähnt bleiben, daß die Richtigkeit der Hörschwellenkurve, speziell im Frequenzbereich unter 1 kHz, mancherseits angezweifelt wird. Die Gründe dafür sind zum einen, daß Verbesserungen der Meßtechnik abweichende Kurven hervorbrachten und zum anderen, daß es vor allem im tieffrequenten Bereich schwierig ist, ebene Schallfelder zu produzieren (vgl. Hellbrück & Ellermeier 2004, 72).

Nicht weiter verwunderlich ist, daß individuelle Hörschwellen frequenzabhängig streuen. Im empfindlichsten Frequenzbereich ist die Streuung minimal; 80% (der Mittelabschnitt einer Gauß'schen Normalverteilung) der Werte liegen in einer Spanne von nur 10 dB. Nach unten wird diese Streuung nur geringfügig weiter, während sich die Werte nach oben hin sehr deutlich streuen. Hier liegen 80% der Werte in einer Spanne von 25 dB. Dies mag auch daran liegen, daß individuelle Hörschwellen über 5 kHz mehrere Extrema aufweisen und zerklüfteter, variabler ausfallen, was seinerseits dadurch beeinflusst sein könnte, daß dieser hochfrequente Abschnitt deutlich empfindlicher auf Schallbelastung reagiert als tiefere Bereiche (vgl. Fastl & Zwicker 2006, 20f.; Zwicker & Feldtkeller 1967, 55).

Um die Handhabung hörbarer Pegel zu erleichtern, wird (v.a. in der Audiometrie) die krumme physikalische Hörschwelle (SPL) begradigt, indem man bei jeder Frequenz – unabhängig von physikalischen Schalldrücken – vom physiologischen Minimum ausgeht. Über viele Personen gemittelt ergibt sich so eine gerade *Nulllinie* bzw. *relative Hörschwelle*, die als *hearing level* (HL) gekennzeichnet wird. Der Bezugsschalldruck für 0 dB_{HL} ist für jede Frequenz der gerade eben hörbare Schalldruckpegel (SPL), also die absolute Hörschwelle. Anders gesagt: Die relative Hörschwelle liegt frequenzunabhängig immer bei 0 dB_{HL}, weil die unterschwelligeren Anteile der Schalldruckpegel (SPL) nicht berücksichtigt werden (vgl. Gelfand 2004, 285; 2009, 90, 113, 147; Hellbrück & Ellermeier 2004, 68f.; Lehnhardt & Laszig 2009, 4f.; Yost 2008, 146f.). Noch einen Schritt weiter geht die Referenzgröße *sensation level* (SL). Der Bezugsschalldruck ist variabel immer der Schalldruck, ab dem ein Schall in einer konkreten Hörsituation de facto zu hören ist. Dezibel SL (dB_{SL}) beziehen sich also auf die signal-, situations-, frequenz- und hörer-spezifische Hörschwelle; geben also sozusagen den aktuellen Signal-Rauschabstand wieder (vgl. Gelfand 2009, 90, 148; Hellbrück & Ellermeier 2004, 69; Lehnhardt & Laszig 2009, 4f.; Yost 2008, 27).

1.1.2 Obergrenzen und Dynamik

Im Gegensatz zur Hörschwelle ist die Obergrenze verhältnismäßig linear, also mehr oder minder frequenzunabhängig (vgl. Gelfand 2009, 88; Hellbrück & Ellermeier 2004, 67; Yost 2008, 147). Pegelobergrenzen sind theoretisch kein Ende der Empfindsamkeit wie die Frequenzobergrenze, sondern markieren die Grenze(n) des *Erträglichen*. Dennoch macht es durchaus Sinn, den Hörbereich hier als nach oben abgeschlossen zu betrachten. Der Abschluß ist jedoch kein klarer Schnitt, sondern stellt eine weiter gefaßte Art „Auslaufzone“ für verkraftbare Schallpegel dar und ist überdies davon abhängig, wie er definiert wird (vgl. Fricke & Louven 2008, 418; Gelfand 2004, 282; 2009, 88; Yost 2008, 147).

Ab grob 100 dB_{SPL} werden Schalle zu laut und zu unangenehm, um ihnen über eine sinnvolle Zeitspanne lauschen zu können. Man spricht in diesem Bereich von der *Unbehaglichkeitsschwelle* (threshold of discomfort). Ab etwa 120 bis 140 dB_{SPL} und mehr treten neben der rein auditiven Wahrnehmung auch Gefühle wie Druck, oder zunehmend starke Schmerzen auf. Hier ist dementsprechend die *Schmerzschwelle* oder *Fühlschwelle*² (threshold of pain, threshold of feeling) erreicht (vgl. Fricke & Louven 2008, 418; Gelfand 2004, 282; 2009, 88; Hellbrück & Ellermeier 2008, 55; Plattig in Keidel 1975, 375; Stevens & Hallowell 1938, 58; Yost 2008, 147; Zwicker & Feldtkeller 1967, 51). Sie markiert zwar die definierte Obergrenze des Hörbereichs, zu bleibenden Schäden kann es allerdings schon bei viel geringeren Schallpegeln kommen. Im empfindlichsten Band zwischen 1 und 5 kHz kann die „Schädigungsschwelle“ bis auf 90 dB_{SPL} absinken (vgl. Zwicker & Feldtkeller 1967, 51). Um 80 dB_{SPL} herum

2 Lehnhardt verwendet den Begriff *Fühlschwelle* „auf der anderen Seite“ (d.h. unterhalb) der Hörschwelle und ist von der Hörschwelle selbst zu unterscheiden, weil Vibrationen – besonders bei tiefen, über Knochenleitung dargebotenen Frequenzen – manchmal schon fühlbar sind, bevor hörbare Pegel erreicht werden (vgl. Lehnhardt & Laszig 2009(v), 7; Weinzierl 2008, 54).

liegt ein Grenzwert, der im englischen Sprachraum *effective quiet* genannt wird, weil Pegel bis 80 dB_{SPL} offenbar keinerlei Schäden und/oder Hörverlust verursachen (vgl. Gelfand 2004, 303). Belastung mit zu hohen Schallintensitäten – etwa durch zu laute Musik – kann aber zu *Ermüdung* und dadurch vorübergehendem Hörverlust (*temporary threshold shift*, TTS) führen. Normalerweise sinkt die Hörschwelle nach Ende der Belastung allmählich wieder auf ihr gewohntes Niveau zurück, was über 16 Stunden dauern kann. Nicht auszuschließen ist, daß die Hörschwelle aufgrund zu intensiver/langerer Belastung dauerhaft erhöht bleibt (siehe 1.2.2, S. 17; vgl. Gelfand 2004, 302ff.).

Der Dynamikbereich des Gehörs wird sehr unterschiedlich, traditionell aber mit 120 dB (bei 1 kHz) angegeben (vgl. Fricke & Louven 2002, 418; Hellbrück & Ellermeier 2004, 67; Spitzer 2002, 50; Winkel 1957, 86). Man muß allerdings bedenken, daß die Angaben zur Obergrenze (Schmerzschwelle) um bis zu 20 dB schwanken und die Untergrenze (Ruhehörschwelle) stark frequenzabhängig ist. Demnach ist auch die Dynamikspanne des Gehörs frequenzabhängig. Nach Yost (vgl. 1993, 4, 147) ist sie um 1 kHz 125 dB bis zu 135 dB weit, während sie um 100 Hz nur 80 bis 90 dB beträgt. Im empfindlichsten Frequenzband liegen 120 bis maximal 140 dB zwischen der Hörschwelle und der definitiven Schmerzgrenze. Das entspricht Verhältnissen von Eins zu einer (120 dB, 10⁶) bzw. zehn Millionen (140 dB, 10⁷) zwischen schwächstem hörbarem und stärkstem erträglichem Schalldruck³ (vgl. Gelfand 2004, 279; Hellbrück & Ellermeier 2004, 67; Yost 2008, 146f.).

1.1.3 Frequenzbereich

Schall pflanzt sich in Form von Druckschwankungen fort, deren Druckböue bei genügend zeitlichem Abstand wie einzelne Impulse erscheinen. Schalle können im weitesten Sinne also als Impulsfolgen betrachtet werden (vgl. Oehler 2008, 88; Spitzer 2002, 23, 77). Die Frequenz ist dabei die Antwort auf die Frage, wie schnell diese Einzelimpulse aufeinander folgen. Dies ist ausschlaggebend für die Tonhöhenempfindung, was gut z.B. an einem Motorrad nachzuvollziehen ist: Im Standgas sind die einzelnen Explosionen im Motor noch gut als solche zu erkennen, aber auch schon als tiefes Knattern interpretierbar. Gibt der Fahrer Gas, steigen Explosionsfrequenz und Tonhöhenempfindung (des Knatterns) im Ohr simultan an (vgl. Spitzer 2002, 77).

Ab einer bestimmten Intervalldauer ist das Ohr nicht mehr in der Lage, Impulse als zusammengehörig zu interpretieren. Damit ist die Untergrenze der Frequenzempfindlichkeit erreicht. Nach Gelfand sowie Stevens und Hallowell kann das menschliche Ohr (bei entsprechendem Pegel) bis zu 2 oder sogar 1 Hz tief hören, was sonderbar scheint (vgl. Gelfand 2004, 279; 2009, 94; Stevens & Hallowell 1938, 69), denn schon 2 Hz zu „hören“ würde praktisch bedeuten, z.B. die Schläge eines Metronoms bei 120 bpm (beats per minute) als ein einziges andauerndes akustisches Ereignis wahrzunehmen. Es sind rund 20 Hz erforderlich, um Einzelimpulse verlässlich zu Tonhöhenempfindungen zu verketteten. Manche Autoren geben aber auch bis zu 16 Hz an, was eventuell damit zusammenhängen könnte, daß extreme Schallpegel über 100 dB_{SPL}, die Grenze mitunter in den Infraschallbereich verlagern (vgl. Dickreiter 1987, 110; Fasold & Veres 2003, 16; Fastl & Zwicker 2006, 18; Fricke & Louven 2008, 416; Gelfand 2004, 279; 2009, 94; Hellbrück & Ellermeier 2008, 54; Müller & Frings 2009, 445, 448; Plattig in Keidel 1975, 375, 377; Stevens & Hallowell 1938, 69; Zwicker & Feldtkeller 1967, 51).

3 Manchmal ist von zwölf Zehnerpotenzen die Rede (vgl. Fricke & Louven 2002, 418; Hellbrück & Ellermeier 2004, 67; Spitzer 2002, 50). Dann wird nicht der Schalldruck (P), sondern die Schallintensität (I) berechnet. Da Intensität dem Quadrat des Schalldrucks proportional ist, kann die logarithmische Dezibelskala mit dem Faktor Zwei umgerechnet werden, was statt zwölf (I) eben sechs (P) Zehnerpotenzen ergibt (vgl. Fricke & Louven 2008, 419; Plattig 1975(c), 372; Yost 2008, 148; für eine kurze und einfache Darstellung der mathematischen Zusammenhänge von Druck und Intensität vgl. besonders Hellbrück & Ellermeier 2004, 68).

Nach oben nimmt die Frequenzempfindlichkeit praktisch sukzessive ab: Spätestens ab 1 kHz können einzelne Nervenfasern die Signalperiodizität nicht mehr korrekt abbilden, indem sie ihre Aktivität mit der Periodizität synchronisieren (*Phasenkopplung*, phase-locking). Ab 5 kHz tritt auch die gemeinsame Phasenkopplung mehrerer Nervenfaserbündel nicht mehr auf, so daß über 5 kHz vermutlich nur noch Ortsinformationen (*Tonotopie*, siehe 2.4.3; S. 77) den Tonhöhereindruck bestimmen (vgl. Moore 1993, 89, 94; Pollard & Jansson 1982, 251; Spitzer 2002, 70; Spreng in Keidel 1975, 131). Die absolute Obergrenze für hörbare Frequenzen wird meistens mit 20 kHz, mitunter aber auch schon mit Werten ab 16 kHz angegeben. Die Frequenzempfindlichkeit des Ohrs erstreckt sich also über drei Frequenzdekaden von rund 20 bis 20000 Hz bzw. musikalischer formuliert über rund zehn Oktaven (vgl. Dickreiter 1987, 110; Fasold & Veres 2003, 16; Fastl & Zwicker 2006, 18; Fricke & Louven 2008, 416; Gelfand 2004, 279; 2009, 94; Hellbrück & Ellermeier 2008, 55; Müller & Frings 2009, 445, 448; Plattig in Keidel 1975, 375, 377; Stevens & Hallowell 1938, 69; Zwicker & Feldtkeller 1967, 51). Dabei sollte bedacht werden, daß praktisch die Hälfte des Frequenzbereichs in die beiden obersten Oktaven (5–10 kHz, 10–20 kHz) fällt. In der musikalischen Praxis kommen selbst Grundtöne jenseits von 1 kHz (zum Vergleich, das hohe C (c³, C6) liegt bei etwa 1057 Hz) so gut wie nie vor (vgl. Hellbrück & Ellermeier 2004, 87; Zwicker & Feldtkeller 1967, 93f.). Dennoch ist dieses „obere Stockwerk“ des Frequenzraums keinesfalls nutzlos: Hier liegt das Reich der Obertonspektren, die nicht nur zentraler Bestandteil der Klangfarbe im engeren Sinn sind (vgl. Helmholtz 1896, 113ff.; Reuter 1995, 85ff.; Stumpf 1890, 30, 20ff.), sondern überdies stabilisierend auf die harmonischen Verhältnisse und Tonhöhereindrücke der musikalischen Grundtöne wirken (vgl. Zwicker & Feldtkeller 1967, 80).

1.2 Zeitauflösung des Gehörs

Vermutlich sind zeitliche Aspekte für keinen der fünf menschlichen Sinne von größerer Bedeutung als für das Hören (vgl. Viemeister & Plack 1993, 116). Zeitliche Aspekte sind eine – wenn nicht *die* – primäre Informationsquelle des Gehörs. Tonhöhe (Frequenz, Periodizität), Richtungshören (Laufzeit- und Phasenunterschiede) und Räumlichkeit (Reflexionen, Nachhallzeiten) sind alles Parameter, die auf Zeitbasis erfaßt werden (vgl. Winckel 1957, 89). Bedenkt man zudem, daß die meisten bedeutsamen Schalle kurzen und schnellen Veränderungen unterworfen, während länger andauernde, konstante Schalle eher bedeutungslos sind (man vergleiche nur die potentielle Bedeutung eines knackenden Zweiges oder des Blätterrauschens im Wind), verwundert nicht, daß das Gehör der zeitempfindlichste aller Sinne ist (vgl. Hellbrück & Ellermeier 2004, 132; Spitzer 2002, 78).

Kurze zeitliche Gegebenheiten werden besser gehört als gesehen. Das Gehör arbeitet um zwei bis drei Zehnerpotenzen schneller als der Gesichts- oder Geruchssinn (vgl. Spitzer 2002, 61, 78; Viemeister & Plack 1993, 116). Sprache ist z.B. sehr bewegter Schall, dessen kürzeste Fluktuationen unter 10 ms liegen (vgl. Hellbrück & Ellermeier 2004, 132). Mit „Lückentests“ (Pausen zwischen zwei Stimuli sollten als unterschiedlich lang erkannt werden) konnte die zeitliche Auflösung auf 2 bis 3 ms beziffert werden (vgl. Gelfand 2004, 300; 2009, 90; Hellbrück & Ellermeier 2004, 133; Viemeister & Plack 1993, 120f.). Schallquellen auf 1° bis 2° genau zu orten bedeutet, binaurale Laufzeitunterschiede in der Größenordnung von 10 bis 20 µs auszuwerten (vgl. Hellbrück & Ellermeier 2004, 132; 2008, 58; Spreng in Keidel 1975, 137; Yost 2008, 183). Dies hat zwar nicht zwingend mit der Schnelligkeit des Gehörs

zu tun, aber es zeigt, in welchen Dimensionen die zeitliche Empfindsamkeit des Ohrs operiert. Mittels Entdeckungen tonaler Anteile in Rauschen konnte zudem gezeigt werden, daß auch monoaural Auflösungen von 10 bis 20 μ s möglich sein könnten (vgl. Hellbrück & Ellermeier 2004, 132)

Die Anpassung des Gehörs an die konkreten Auswertungsanforderungen benötigt natürlich selbst auch Zeit, die gewissermaßen als Trägheit des Ohrs verstanden werden kann (vgl. Fasold & Veres 2003, 52; Fricke & Louven 2008, 433; Reuter 1995, 31). So ist ein bestimmtes Schallsignal nicht einfach nur da. Das Ohr braucht Zeit, um sich darauf einzustellen. Es dauert also seine Zeit, bis das Gehör durch den eintreffenden Schall sozusagen „eingeschwungen“ ist (vgl. Fasold & Veres 2003, 52; Winkel 1952, 52). Bis es soweit ist, hängen Wahrnehmbarkeit und Auflösung von der Signaldauer ab – und zwar in dem Sinne, daß sich beide mit zunehmender Signaldauer verbessern (vgl. Green 1993, 24, 25; Viemeister & Plack 1993, 143, 144). Innerhalb dieses bis zu mehreren hundert Millisekunden langen Zeitraums kommt es zu zeitlicher Integration (*temporal integration*) (vgl. Hellbrück & Ellermeier 2008, 60; Viemeister & Plack 1993, 143, 144; Yost 2008, 149). Sie wird sehr unterschiedlich mit Werten zwischen 250 und 500 ms angegeben (vgl. Gelfand 2004, 279; 2009, 88; Stevens & Hallowell 1938, 154; Yost 2008, 147). Darüber hinaus, d.h. nach Ende des Integrationsvorgangs, bringen längere Signaldauern (was ja mehr Gesamtenergie bedeutet) keinen weiteren Informationsgewinn, Energieverluste durch kürzere Dauern müssen allerdings durch entsprechend höhere Pegel ausgeglichen werden (vgl. Gelfand 2004, 286; 2009, 88f.; Yost 2008, 149). Gelfand (vgl. 2004, 287) bringt den treffenden Vergleich mit der Fotografie, genauer gesagt dem Verhältnis von Blendenöffnung zu Verschlusszeit: je kleiner die Blendenöffnung, desto länger muß der Verschluss offen sein, um die gleiche Lichtmenge durchzulassen.

Im Zeitbereich zwischen etwa 10 ms und ungefähr 100 bis 500 ms ist das benötigte Verhältnis von Signaldauer zu Schallenergie konstant indirekt proportional. Änderungen der Signaldauer um den Faktor zehn (eine Dekade) bewirken entsprechende Intensitätsänderungen um (8 bis) 10 dB (vgl. Gelfand 2004, 286, 355f.; Green 1993, 29; Hellbrück & Ellermeier 2004, 138; 2008, 60f.; Viemeister & Plack 1993, 144; Yost 2008, 147, 148). Unter 10 ms muß der Pegel sehr stark erhöht werden, um die Schallenergie konstant zu halten, weil das Spektrum bei dieser kurzen Dauer so breit wird, daß viel Energie in unhörbare Bereiche „ausgelagert“ wird und nicht mehr an der Signalerkennung partizipiert (*Küpfmüller'sche Beziehung*, siehe S. 13); ein Verlust, der nur durch immense Pegelsteigerungen ausgeglichen werden kann (vgl. Green 1993, 29f.; Viemeister & Plack 1993, 144; Yost 2008, 148). Umgekehrt können auf gleiche Art normalerweise unhörbare Spektren in den Hörbereich hinein erweitert und hörbar werden (vgl. Yost 2008, 148f.). Es bedeutet ebenfalls, daß breitbandige Signale vom „Mehr an Information“ früher profitieren als schmalbandige und generell schneller gehört werden (vgl. Hellbrück & Ellermeier 2004, 133). Die Erkennung eines Tons, dessen Spektrum schmal und dessen Energie über die Zeit verteilt ist, dauert etwa um die 300 ms, während die Erkennung eines Klicks, dessen Energie im breiten Spektrum angelegt ist, nach wenigen Millisekunden abgeschlossen ist (vgl. Yost 2008, 149).

1.2.1 Integrationszeiten

Die gute Zeitauflösung des Gehörs macht klar, daß das Resonanzsystem des Innenohrs sehr stark gedämpft sein muß (vgl. Winkel 1952, 53). Das Problem ist, daß Zeitauflösung und Frequenzauflösung

grundsätzlich gegeneinander arbeiten, weil schärfere Resonanzfrequenzen (schmalere Filter) längere Einschwingzeiten benötigen und auch länger ausklingen (vgl. Reuter 1995, 17; Viemeister & Plack 1993, 132; Winckel 1952, 53). Gesucht wird also ein Kompromiß zwischen schnellem Einschwingen und guter Frequenzauflösung des Systems (vgl. Winckel 1952, 54).

Das Ohr kommt dieser Anforderung nach dem Motto „alles zu seiner Zeit“ nach, indem es sich dem Stimulus anpaßt und entsprechend mit unterschiedlichen Verarbeitungsregistern operiert, also entweder die Zeitauflösung oder die Frequenzauflösung optimiert (vgl. Reuter 1995, 31). Vergleichbar hat sich das Spektrum der Musikinstrumente entwickelt, von denen die meisten entweder auf scharfe Einsätze (v.a. Perkussionsinstrumente) oder ihren Klang, ihre Teiltonstruktur ausgelegt sind (vgl. Winckel 1952 54). Umgekehrt sind die Musikinstrumente nun einmal „mit den Ohren gebaut“ (Fricke 1989, 1993 zit. nach Reuter 1995, 37; 1996, 17), so daß Ähnlichkeiten alles in allem nicht weiter verblüffen. Der Einschwingvorgang als Ausgleichsvorgang geht zwangsläufig mit einer hohen zeitlichen Fluktuation einher. Ist der eingeschwingene Zustand erreicht, nehmen die Fluktuationen ab und das Spektrum des Klangs tritt in den Vordergrund (vgl. Dickreiter, 1987, 69; Reuter 1995, 15–29, 37). Da sich das Ohr am Klang orientiert, ist auch die zeitliche Integration der Klangentwicklung entsprechend in verschiedene Verarbeitungsregister unterteilbar (vgl. Reuter 1995, 31; 1996, 15).

Erste Integrationszeit

Die erste Integrationszeit beschreibt den Zeitraum, den das Gehör braucht, um die *Frequenzgruppen* (siehe 1.3, S. 18) auf der Basilarmembran aufzubauen bzw. (bei Klangveränderungen) umzuorientieren (vgl. Fricke & Louven 2008, 433; Reuter 1995, 31; 1996, 15). Mit den Frequenzgruppen befindet sich auch die spektrale Auflösung noch im Aufbau, während die zeitliche schon sehr genau arbeitet. Alle spektralen Bewegungen, die innerhalb der ersten Integrationszeit passieren, werden demnach nur als Knack verstanden. Erst mit fertig ausgebildeten Frequenzgruppen ist das Ohr in der Lage, Klangspektren differenziert zu verarbeiten (vgl. Fricke & Louven 2008, 433; Reuter 1995, 31f.; 1996, 15). Der psychoakustische Richtwert liegt bei 10 ms, tatsächlich schwankt die erste Integrationszeit frequenz- und pegelabhängig zwischen 2 und 20 ms (vgl. Fricke & Louven 2008, 433; Pollard & Jansson 1982, 251; Reuter 1995, 31, 32; 1996, 15). So lange das Gehör braucht, die Frequenzgruppen aufzubauen, so lange sollten auch die Einschwingvorgänge dauern. Sind sie sozusagen „schneller als das Ohr erlaubt“, werden sie als Knack wahrgenommen, wobei höhere Frequenzen oder geringere Pegel prinzipiell kürzere knackfreie Einschwingzeiten ermöglichen (vgl. Fricke & Louven 2008, 433; Pollard & Jansson 1982, 251; Reuter 1995, 32; 1996, 15).

Die Erklärung dafür liegt im grundsätzlichen Zusammenhang von Bandbreite und Einschwingzeit, der in der *Küpfmüller'schen Beziehung* folgendermaßen formuliert ist: je größer die Bandbreite, desto kürzer die Einschwingzeit (vgl. Oehler 2008, 89; Reuter 1995, 32; Winckel 1952, 33f., 59; Zwicker & Feldtkeller 1967, 215). Mathematisch ausgedrückt bleibt das Produkt aus Bandbreite (Δf) und Einschwingdauer (τ) immer konstant (k): $\Delta f \cdot \tau = k$ (vgl. Reuter 1995, 17, 32). Letztlich ist dies nichts anderes als eine Übertragung der *Heisenberg'schen Unbestimmtheitskorrelation* auf die Akustik (vgl. Oehler 2008, 89; Winckel 1952, 34, 59).

Größere Bandbreiten erlauben kürzere knackfreie Einschwingzeiten (vgl. Reuter 1995, 32). Weil mit steigender Mittenfrequenz die Frequenzgruppenbreite proportional mitwächst (vgl. Hellbrück & Ellermeier 2004, 128; Moore 1993, 68; Viemeister & Plack 1993, 132), also sozusagen der zu einer Frequenzgruppe integrierte Frequenzbereich breiter wird, darf die knackfrei wahrgenommene Einschwingdauer mit steigender Frequenz immer kürzer werden (vgl. Reuter 1995, 32). Dementsprechend beanspruchen die (zeitlich besser aufgelösten; vgl. Gelfand 2004, 288; Hellbrück & Ellermeier 2004, 134) hohen Frequenzen eher die kürzeren Einschwingzeiten und tiefe Frequenzen eher die längeren (vgl. Fricke & Louven 2008, 433; Reuter 1995, 32; 1996, 15).

Zu kurze Signale, Einschwing-/Ausgleichsvorgänge oder plötzliche Frequenzänderungen werden aufgrund der nicht rechtzeitig angepaßten Filter nicht korrekt aufgelöst, sondern als indifferente Pegelsprünge gewertet, deren Spektrum sich kurzzeitig extrem weit fächert und dadurch als Knack wahrgenommen wird (vgl. Fricke & Louven 2008, 433; Gelfand 2004, 286; Reuter 1995, 32; Zwicker & Feldtkeller 1967, 66). Erst mit den fertig ausgebildeten Frequenzgruppen geht das Gehör vom „breitbandigen zum schmalbandigen Hörverhalten“ über (Scholl 1961, zit. nach Reuter 1995, 31). Gemessene knackfreie Einschwingzeiten liegen entsprechend, frequenz- und pegelabhängig mit insgesamt 3 bis 30 ms, knapp oberhalb der ersten Integrationszeit (vgl. Reuter 1995, 32; 1996, 15), wonach die meisten Musikinstrumente mit Einschwingzeiten über 10 ms auf der „sicheren Seite“ sind (vgl. Fricke & Louven 2008, 433; Winckel 1952, 97).

Während der ersten 1 bis 5 ms werden auch Richtungsinformationen sehr genau ausgewertet, indem *alle* eintreffenden Schallanteile zur Bildung des Hörereignisortes beitragen. Dieser zeitempfindliche Vorgang wird *Summenlokalisation* genannt (siehe 1.7.3, S. 45; vgl. Blauert 1974, 163ff., 177f.; Fasold & Veres 2003, 56; Guski 1996, 284).

Zweite Integrationszeit

Nach dem Aufbau der Frequenzgruppen nimmt die spektrale Auflösung ihre Arbeit auf, vor allem aber wird die Identität des Klangs etabliert, indem alle innerhalb der zweiten Integrationszeit liegenden Schallereignisse zu einem einzigen Hörereignis zusammengefaßt werden. Ihre Dauer schwankt signal- und frequenzabhängig zwischen rund 10 bis 100 ms, wobei der psychoakustische Richtwert grob bei 50 bis 55 ms liegt. Die zweite Integrationszeit (genau genommen ihr Grenzwert) wird auch *Verwischungsschwelle* (time perception smear) genannt (vgl. Fasold & Veres 2003, 52; Fricke & Louven 2008, 433; Reuter 1995, 33f.; 1996, 16; Winckel 1952, 35, 54).

Die Einschwingvorgänge der Musikinstrumente sind mit Werten zwischen 15 und 70 ms einerseits lang genug, um nicht als Einschaltknack wahrgenommen zu werden, andererseits kurz genug, um die Verwischungsschwelle nicht zu überschreiten. Wird sie überschritten, ist das Anschwellen des Klangs recht deutlich hörbar (vgl. Fricke & Louven 2008, 433; Reuter 1996, 17).

Ungefähr 5 bis 20 ms nach Signalbeginn, also nach Ende der ersten Integrationszeit, greift die spektrale Auflösung in Form der Fähigkeit, Klangfarben zu verwerten (vgl. Reuter 1995, 32; 1996 15; Winckel 1952, 54). Nach Winckel (vgl. 1952, 54) arbeitet sie schon nach nur 3 ms immerhin gut genug, um die Formanten eines Spektrums auszulesen. Etwas später, nach etwa 10 bis 50 ms, setzt auch die Fähigkeit ein, Tonhöhen zu erkennen (vgl. Reuter 1995, 33). Über 1 kHz sind Tonhöhereindrücke ab

rund 10 ms konstant (vgl. Gelfand 2004, 286). Es wirkt erstaunlich, daß Tonalität unterhalb von 1 kHz Messungen zufolge schon nach zwei, mit Lernerfahrung sogar nur einer Periode – zumindest vage – erkennbar sein soll (vgl. Gelfand 2004, 286; Reuter 1995, 33; 1996, 16). Trotz dieses Befunds kann man allgemein von einer, der Verwischungsschwelle entsprechenden (unteren) Tonhöhenerkennungsschwelle, ausgehen (vgl. Reuter 1995, 33; 1996, 16).

In den ersten 25 bis 50 ms nach dem Primärschall eintreffende Signalanteile unterliegen dem *Gesetz der ersten Wellenfront* (*Präzedenzeffekt*, Haas Effekt). In dieser Zeit bestimmt allein das Erstsinal (die erste Wellenfront) den Richtungseindruck und alle nachfolgenden Schallanteile werden ihm subsumiert (siehe 1.7.3, S. 45; vgl. Blauert 1974, 163, 177, 184ff.; vgl. Culling & Akeroyd 2010, 135f.; Fasold & Veres 2003, 56; Gelfand 2004, 412f.; 2009, 106; Guski 1996, 283, 285; Hellbrück & Ellermeier 2004, 152; Yost 2008, 180). Spätere Schallanteile bilden den Nachhall des Klangs und tragen so wesentlich zu seinem Charakter (seiner Identität) bei, da das zu einem Höreindruck verschmolzene Gesamtsignal auch die Raumcharakteristik transportiert. Die *Verwischungsschwelle* markiert dabei die Grenze, ab der aus nützlichen (identitätsstiftenden) störende (erkennbarkeitstrübende) Reflexionen werden (vgl. Fasold & Veres 2003, 52f., 56; Fricke & Louven 2008, 433f.; Hellbrück & Ellermeier 2004, 134). Die Grenze, ab der nachfolgende Schallanteile nicht mehr integriert, sondern als eigenständig aufgefaßt werden (z.B. Reflexionen als Echo), also Signaltrennung auftritt, wird *Echoschwelle* genannt (siehe 1.7.3, S. 46; vgl. Blauert 1974, 163, 179; Gelfand 2004, 413f.; Guski 1996, 284f.).

Darüber hinaus ist die Verwischungsschwelle auch für musikalisches Zusammenspiel von Bedeutung, weil sie die Toleranzgrenze für rhythmische Ungenauigkeiten festlegt. In Abhängigkeit von Faktoren wie Instrument, Spielweise und Richtungseindruck liegt sie zwischen etwa 2 und 70 ms (vgl. Fricke & Louven 2008, 433; Reuter 1996, 16).

Dritte Integrationszeit

Die dritte Integrationszeit entspricht dem Zeitraum, in dem die spektrale Auflösung ihre maximale Leistungsfähigkeit erreicht. Spektrale Bewegungen führen jetzt nicht mehr zu störenden, geräuschhaften Erweiterungen des Spektrums bis hin zu Knacken (Transienten), sondern werden gezielt mitverarbeitet (vgl. Reuter 1995, 34; 1996, 16f.). Die dritte Integrationszeit dauert frequenzabhängig minimal 100 bis 150 ms, maximal 200 bis 300 ms, wobei der Richtwert für die Psychoakustik bei 250 ms liegt. Ihr Grenzwert wird als *Klangverschmierungsschwelle* bezeichnet (vgl. Fricke & Louven 2008, 434; Reuter 1995, 34; 1996, 17).

Während der dritten Integrationszeit wird die Lautstärke aufsummiert und die Tonhöhenwahrnehmung gesteigert (vgl. Fricke & Louven 2008, 434). Nach etwa 100 ms nimmt die Frequenzempfindlichkeit des Gehörs nur noch langsam zu und hat mit der Klangverschmierungsschwelle, nach 200 bis 300 ms, ihr Maximum erreicht und den Tonhöhenindruck etabliert (vgl. Fricke & Louven 2008, 434; Reuter 1995, 34f.).

Bereits nach 65 bis 140 ms ist der Lautstärkeindruck endgültig festgeschrieben. Er ist pegelabhängig bei lauten Tönen früher erreicht als bei leisen (vgl. Fricke & Louven 2008, 434; Winkel 1952, 99), wobei vielleicht anzumerken ist, daß Winkel (vgl. 1952, 99) für 90 Phon 140 ms, für 20 Phon allerdings nur 20 ms, also genau umgekehrt gelagerte Werte angibt. Zudem streuen die Angaben zur Laut-

stärkenintegrationszeit weiter und variieren von Studie zu Studie (vgl. Gelfand 2004, 356). Anderen Angaben zufolge ist der Lautstärkeindruck bis etwa 100 bis 200 ms von der Signaldauer abhängig (vgl. Green 1993, 30; Hellbrück & Ellermeier 2004, 138).

Sicher ist, daß Signaldauern über 300 bis 500 ms (zeitliche Integration, siehe S. 12) keine Vorteile mehr bringen und im Sinne der Signalerkennung und -auswertung praktisch als unendlich lang angesehen werden können (vgl. Gelfand 2004, 279, 286; Yost 2008, 149). Mit anderen Worten: Mit der Klangverschmierungsschwelle (d.h. dem Ende der dritten Integrationszeit) hat das Gehör seine maximale Leistungsfähigkeit erreicht (vgl. Fricke & Louven 2008, 434; Reuter 1995, 34) und das Ohr ist quasi „eingeschwungen“ (Rösing zit. nach Reuter 1995, 35). Nach Ende der Integration setzen bereits Adaption und Ermüdung ein (vgl. Fricke & Louven 2008, 434; Reuter 1995, 34).

1.2.2 Adaption und Ermüdung

Grundsätzlich könnten Adaption und Ermüdung thematisch gleichermaßen sinnvoll der Hörschwelle zugeordnet sein, da sie – wie auch die Verdeckung – reizabhängige Hörschwellenverschiebungen darstellen (vgl. Hellbrück & Ellermeier 2004, 135). Aufgrund ihres zeitlichen Verhaltens sollen sie hier als die abfallende Flanke der zeitlichen Verarbeitungskurve des Gehörs verstanden werden.

Da Adaption auf elektrochemische Vorgänge in den Haarzellen und Ermüdung auf Effekte metabolischer Veränderungen der Endolymphe zurückgeführt werden (vgl. Hellbrück & Ellermeier 2004, 137; Lehnhardt & Laszig 2009, 89), ist eine Unterscheidung beider Phänomene, aufgrund der noch unklaren Beziehung zwischen Stoffwechselprozessen und elektrophysiologischen Vorgängen, alles andere als leicht (Blauert gibt den Übergang sogar als fließend an; vgl. 1974, 130; Hellbrück & Ellermeier 2004, 137).

Adaption

Als *Adaption* wird die schnelle Empfindlichkeitsabnahme des Gehörs kurz nach Schallbeginn bezeichnet, während ermüdungsbedingter Hörverlust erst *nach* Reizende eintritt (vgl. Blauert 1974, 130; 1997, 162; Gelfand 2004, 358). Da vor allem die Lautheitsempfindung betroffen ist, wird auch von *Lautheitsadaption* gesprochen (vgl. Hellbrück & Ellermeier 2004, 136).

Dauerhaft präsente, pegelkonstante Schalle nehmen nach einer gewissen Zeit in nennenswertem Maße in ihrer empfundenen Lautstärke ab. D.h. das Hörereignis wird leiser, obwohl das Schallereignis pegelkonstant ist (vgl. Gelfand 2004, 357). Vermutlich sollen Adaptionerscheinungen das Gehör vor Überlastung schützen, wobei Adaption interindividuell sehr stark verschieden ausgeprägt ist (vgl. Fricke & Louven 2008, 434; Gelfand 2004, 359). Interessanterweise erhöht Adaption zwar die Hörschwelle, senkt aber die Pegelunterschiedsschwellen des Gehörs (vgl. Lehnhardt & Laszig 2009, 89), was letztlich keinen Verlust, sondern eine Empfindlichkeitsverlagerung bedeutet. Außerdem wurde Adaption bisher nur bei recht geringen Pegeln nachgewiesen. Speziell bei Sinustönen kommt es offenbar nur bis etwa 30 dB_{SL} zu Adaption. Dabei sind Adaptionerscheinungen frequenzabhängig verschieden: Im hochfrequenten Bereich greift Adaption nicht nur stärker, sondern auch noch bei höheren Pegeln bis über

30 dB_{SL} (vgl. Gelfand 2004, 359; Hellbrück & Ellermeier 2004, 136). Die Frage, ob es bei hohen Pegeln zu Adaption kommt, ist vor allem deswegen unklar, weil noch keine befriedigende Trennlinie zwischen Adaption und Ermüdung gezogen werden konnte (vgl. Hellbrück & Ellermeier 2004, 136).

Bei Anfangslautstärken (Sinustöne) um und über 40 Phon ist die Lautstärke adaptionsbedingt nach 10 s um 3 dB und nach 100 s um 8 bis 10 dB gesunken. Für leisere Töne (unter 40 Phon) fällt die Adaption geringer aus und reduziert die Lautstärke zu den gleichen Zeitpunkten nur um 2 bzw. 5 dB (vgl. Fricke & Louven 2008, 434). Nach 2 bis 5 min. ist das Adaptionsmaximum erreicht. Die anschließende Readaption geht relativ schnell, so daß nach 1 bis 2 min. der Normalzustand wieder erreicht ist (vgl. Blauert 1974, 130; 1997, 162; Hellbrück & Ellermeier 2004, 136). Unabhängig von der Beschallung tritt Adaption generell binaural auf, wobei das „bessere“ Ohr (geringere Adaptionsstärke, größere Adaptionslatenz) das effektive Adaptionsverhalten des Gehörs bestimmt (vgl. Gelfand 2004, 359f.).

Adaption ist ein biochemischer Vorgang innerhalb der Haarzellen. Dabei kommt es nach Reizbeginn zum Abfall des Negativanteils im endokochleären Ruhepotential, wodurch die Empfindlichkeit der Haarzelle abnimmt. So strebt die Adaption „einer Sättigung zu, die einen Gleichgewichtszustand über längere Zeit gewährleistet“ (Lehnhardt & Laszig 2009, 89). Nach Reizende kehrt das Potential auf sein Ausgangsniveau zurück (vgl. Hellbrück & Ellermeier 2004, 136; Lehnhardt & Laszig 2009, 89).

Ermüdung

Als *Ermüdung* (auditory fatigue, *zeitweilige Hörschwellenverschiebung*, temporary threshold shift, TTS) werden langsamer wirkende Hörschwellenanhebungen nach lang andauernder intensiver Schallbelastung bezeichnet (vgl. Blauert 1974, 130; 1997, 162; Fasold & Veres 2003, 57; Gelfand 2004, 303; Hellbrück & Ellermeier 2004, 137; Stevens & Hallowell 1938, 218).

Im Gegensatz zur Adaption tritt Ermüdung nur bei hohen Schallpegeln über 80 dB_{SPL} auf (vgl. Hellbrück & Ellermeier 2004, 137). Das Ausmaß der Ermüdungserscheinungen hängt von Dauer und Intensität der Schallbelastung ab: Je länger und/oder intensiver die Belastung, desto größer der Hörverlust (vgl. Gelfand 2004, 303; Hellbrück & Ellermeier 2004, 137). Belastbarkeit und Belastungsdauer des Gehörs stehen indirekt proportional zueinander. D.h. theoretisch erlauben kürzere Beschallungszeiten *entsprechend* höhere Pegel (vgl. Fastl & Zwicker 2006, 18). Die ermüdungsbedingt verschobene Hörschwelle verhält sich ähnlich wie die verdeckungsbedingte *Mithörschwelle* (siehe 1.6.1, S. 33). Je intensiver die Belastung ist, desto breiter wird das belastete Frequenzband, wobei die hochfrequente Flanke flacher verläuft; d.h. der hochfrequente Bereich reagiert empfindlicher und wird früher und stärker mitbelastet als der tieffrequente. Eine Besonderheit ist, daß die maximale Ermüdung ungefähr eine halbe Oktave tiefer auftritt, als die Belastung stattgefunden hat (vgl. Gelfand 2004, 303; Hellbrück & Ellermeier 2004, 137).

Grundsätzlich wächst die Ermüdung mit zunehmender Belastungsdauer (vgl. Gelfand 2004, 303; Lehnhardt & Laszig 2009, 89). Nach acht bis 16 Stunden scheint dieser Zusammenhang aber aufgehoben, so daß darüber hinausgehende Belastungen den Hörverlust nicht weiter steigern (vgl. Gelfand 2004, 303). Ihr Maximum (höchste Schwelle) erreicht die Ermüdung etwa zwei Minuten nach Belastungsende, erst danach beginnt die Erholung (vgl. Gelfand 2004, 304; Hellbrück & Ellermeier 2004, 137). Der Rückfall der Hörschwelle auf ihr normales Niveau (Ruhehörschwelle) verläuft in aller Regel

linear und graduell. Allerdings braucht die Erholung viel Zeit und kann – je nach Dauer und Intensität der Belastung – bis zu 16 Stunden und länger dauern (vgl. Blauert 1974, 130; 1997, 163; Gelfand 2004, 302, 304f.; Hellbrück & Ellermeier 2004, 137). Nach besonders starker Belastung kann sich der Erholungsverlauf auch auf spätere Zeitpunkte verlagern, indem er zuerst stagniert und danach relativ schnell fortschreitet; sozusagen eine S-Kurve beschreibt. Allerdings kann es auch zu permanenten Schädigungen des Gehörs kommen, so daß die Hörschwelle nicht wieder auf ihr Normalniveau zurücksinkt, sondern dauerhaft erhöht bleibt (vgl. Gelfand 2004, 302, 305; Hellbrück & Ellermeier 2004, 137). In solchen Fällen kann natürlich nicht mehr von Ermüdung, sondern nur noch von *permanentem Hörverlust* (permanent threshold shift, PTS) die Rede sein (vgl. Fasold & Veres 2003, 57).

Durch die starken Schallbelastungen kommt es wahrscheinlich zu metabolischen Veränderungen im Endolymphhaushalt sowie in den Haarzellen selbst, die für Ermüdungserscheinungen verantwortlich sind. Beispiele für solche Veränderungen können Abfallen des Sauerstoffspiegels oder Ansammlungen von verschiedenen metabolischen Abfallstoffen sein (vgl. Hellbrück & Ellermeier 2004, 137).

1.3 Frequenzgruppen

Eine fundamentale Eigenschaft des Gehörs ist es, Schallsignale in Spektralbänder bestimmter Breite zu zerlegen und diese Frequenzbänder zu kleinen Einheiten zu integrieren. Diese Frequenzbänder werden *Frequenzgruppen* oder *kritische Bänder* (auditory filter) genannt (vgl. Hellbrück & Ellermeier 2004, 45, 87, 126; 2008, 57; Zwicker & Feldtkeller 1967, 73). Die auf die einzelnen Haarzellen entfallende Schalleistung wird innerhalb einer Frequenzgruppe integriert, sozusagen gekoppelt (vgl. Winckel 1957, 84), weshalb man die Integrationsbereiche auch *Kopplungsbreiten* nennt. Frequenzgruppen sind aller Wahrscheinlichkeit nach ein recht konstantes Phänomen und nur sehr marginalen interindividuellen Schwankungen unterworfen (vgl. Zwicker & Feldtkeller 1967, 57).

Es ist nicht auszuschließen, daß das Prinzip der Frequenzgruppen auf ein grundsätzliches Konzept des sensorischen neuronalen Apparats zurückzuführen ist. Immerhin konnten auch auf, als künstliche Basilarmembran genutzten Unterarmen (und anderen Hautregionen) von Versuchspersonen kritische Bandbreiten der Empfindungsintegration festgestellt werden (vgl. Békésy 1960, 569f.; Winckel 1957, 87). Bedenkt man, daß Frequenzgruppen ein Phänomen der Frequenzauflösung sind und alle physikalischen Reize auf irgendeine Weise aufgelöst werden müssen, um für das Nervensystem digitalisiert zu werden, ist dies zumindest keine vollkommen aus der Luft gegriffene Überlegung.

Die Frequenzgruppen sind mit ziemlicher Sicherheit im Innenohr anzusiedeln und als Filterstation auf so peripherer Ebene an so gut wie allen Hörphänomenen beteiligt und somit für den Hörvorgang von großer Bedeutung (vgl. Gelfand 2004, 322; Green 1993, 27; Hellbrück & Ellermeier 2004, 129; Moore 1993, 61, 81; Spreng in Keidel 1975, 142; Spitzer 2002, 100). Das Ohr arbeitet wie eine Filterbank mit sich mehr oder weniger überlappenden Bandpaßfiltern. Innerhalb eines Filters (d.h. innerhalb einer Frequenzgruppe) folgt die Signalverarbeitung anderen Gesetzmäßigkeiten als außerhalb. Die Frequenzgruppenbreite (kritische Bandbreite) ist demnach wenig überraschend jene Bandbreite, ab der abrupte Änderungen (der Verarbeitung und der Wahrnehmung) auftreten (vgl. Gelfand 2004, 322, 353; Guski 1996, 244f.; Hellbrück & Ellermeier 2004, 87; Zwicker & Feldtkeller 1967, 57).

1.3.1 Filterbreiten

Die Beobachtung, daß die Mithörschwelle von *weißem Rauschen* (Rauschen mit gleichmäßigem, frequenzunabhängigem Leistungsspektrum über alle hörbaren Frequenzen) über 500 Hz mit steigender Frequenz konstant um 3 dB/Oktave (10 dB/Dekade) zunimmt, konnte dadurch erklärt werden, daß der Hörbereich in schmale Frequenzbänder unterteilt ist, in denen Schalleistung verarbeitet wird (vgl. Fast & Zwicker 2006, 149; Hellbrück & Ellermeier 2004, 127; Zwicker & Feldtkeller 1967, 57, 68, 70). Da die Dichte von weißem Rauschen frequenzunabhängig ist, muß die Leistung in gleichen Frequenzbändern unabhängig von der Mittenfrequenz gleich sein. Wächst die Bandbreite proportional zur Mittenfrequenz, fällt logischerweise auch mehr Rauschen (also mehr Leistung) in jedes weitere Frequenzband (vgl. Hellbrück & Ellermeier 2004, 127; Zwicker & Feldtkeller 1967, 57, 68). Die Frequenzbänder entsprechen den Frequenzgruppen, deren Bandbreite so ermittelt werden kann (vgl. Zwicker & Feldtkeller 1967, 57). Mittlerweile gibt es unzählige Methoden, die Frequenzgruppenbreiten zu ermitteln, auf die hier nicht weiter eingegangen werden soll (vgl. dazu z.B. Fastl & Zwicker 2006, 150–158).

Die traditionelle Darstellung ist, daß die Frequenzgruppenbreiten unter 500 Hz konstant 90 bis 100 Hz, und über 500 Hz konstant etwa 17 bis 20% der Mittenfrequenz, also ungefähr eine Terz, betragen (vgl. Fastl & Zwicker 2006, 150; Gelfand 2004, 323; Guski 1996, 245; Hellbrück & Ellermeier 2004, 88, 128; Moore 1993, 67f., 81; Reuter 1995, 33, 1996, 15; Spitzer 2002, 101; Spreng in Keidel 1975, 140; Winckel 1957, 85; Zwicker & Feldtkeller 1967, 57, 68, 72f.; Pollard & Jansson (vgl. 1982, 250) ziehen die Grenze bei 450 Hz). Die Darstellung ist jedoch nur eine Annäherung an tatsächliche Sachverhalte: Die Frequenzgruppenbreiten wachsen bis etwa 3 kHz ein wenig langsamer als proportional zur Mittenfrequenz an, darüber hingegen ein wenig schneller, so daß die Wachstumsspanne eher annähernd konstante 10 bis 20% betragen dürfte (vgl. Hellbrück & Ellermeier 2008, 57; Fast & Zwicker 2006, 158). Dies harmoniert auch mit der Feststellung, daß die Frequenzgruppenbreite um die Mittenfrequenz 1 kHz mit 160 bis 165 Hz eher im Bereich von 17% der Mittenfrequenz, statt der anzunehmenden 200 Hz bei 20% liegt (vgl. Hellbrück & Ellermeier 2004, 128; Moore 1993, 77; Zwicker & Feldtkeller 1967, 72). Nach Moore fußt die traditionelle Beschreibung überhaupt auf dünner, fragwürdiger Datenbasis und neuere, mit zuverlässigeren Methoden durchgeführte Messungen deuten an, daß die Frequenzgruppenbreiten auch unterhalb von 500 Hz logarithmisch abnehmen (vgl. Gelfand 2004, 322, 323; Hellbrück & Ellermeier 2004, 128; Moore 1993, 68; Viemeister & Plack 1993, 132).

Das Spannende ist, daß die offenbar kritische 500 Hz-Grenze zwischen linearer und logarithmischer Frequenzskala bei verschiedenen (frequenzbezogenen) Hörphänomenen auftaucht, wie etwa der Tonhöhenwahrnehmung (vgl. Fricke & Louven 2008, 423; Hellbrück & Ellermeier 2004, 86f.; Zwicker & Feldtkeller 1967, 80, 82, 93), den Frequenzunterschiedsschwellen (vgl. Fricke & Louven 2008, 424; Hellbrück & Ellermeier 2004, 85f.; Reuter 1995, 33; 1996, 15f.; Zwicker & Feldtkeller 1967, 68, 92) und natürlich der schon erwähnten Mithörschwelle weißen Rauschens. Der Schluß liegt nahe, daß die 500 Hz-Grenze mit der Besonderheit, darunter linear und darüber logarithmisch zu arbeiten eine fundamentale Eigenschaft des Gehörs ist (vgl. Zwicker & Feldtkeller 1967, 68). Da scheint es verwunderlich, daß ausgerechnet ein ebenfalls grundlegendes Instrument wie die Frequenzgruppen die Ausnahme der Regel sein soll.

Auf der Basilarmembran haben Frequenzgruppen keine fixen Positionen, sondern können signalabhängig an jeder beliebigen Stelle gebildet werden (vgl. Gelfand 2009, 94; Hellbrück & Ellermeier 2004, 128; Moore 1993, 71; Zwicker & Feldtkeller 1967, 73, 89, 133). Legt man sie (rein hypothetisch) jedoch so an, daß sich ein Band an das nächste reiht, die obere Grenzfrequenz der einen also die untere Grenzfrequenz der anderen ist, kommt man traditionellerweise auf 24 Frequenzgruppen verteilt über die Länge der Basilarmembran (vgl. Fastl & Zwicker 2006, 158; Hellbrück & Ellermeier 2004, 88, 128; Pollard & Jansson 1982, 250; Winckel 1957, 84; Zwicker & Feldtkeller 1967, 73). Auf Basis der 24 Frequenzgruppen wurde die sogenannte *Frequenzgruppenskala* (critical-band rate scale) mit der Einheit *Bark* (nach Heinrich Barkhausen) entwickelt. Eine Frequenzgruppe entspricht also 1 Bark (vgl. Fastl & Zwicker 2006, 158f.; Hellbrück & Ellermeier 2004, 88, 128; 2008, 57; Zwicker & Feldtkeller 1967, 90). Mancherseits wurde die Basilarmembran allerdings auch schon in 30 aneinandergereihte Frequenzgruppen unterteilt (vgl. Reuter 1996, 15; Zwicker & Feldtkeller 1967, 78).

Frequenzgruppen sind keine linearen Filter. Bandbreite und Filterform sind nicht nur frequenzabhängig, sondern verändern sich auch in Abhängigkeit von der Schallintensität, d.h. praktisch mit dem *Frequenzgruppenpegel* (Pegel der Teilintensität, die innerhalb einer Frequenzgruppe wirksam ist; vgl. Fastl & Zwicker 2006, 165–172; Zwicker & Feldtkeller 1967, 109–115, 133–139, 187). Mit steigender Schallintensität wird die tieffrequente Flanke zunehmend flacher, die hochfrequente steiler. Insgesamt nimmt die Bandbreite dadurch zu (vgl. Gelfand 2004, 322; Moore 1993, 69f.). Nun weisen Filter grundsätzlich keine senkrechten Flanken und daher auch keine sauber bestimmbare Rechtecksform auf. Um „handlichere“ Rechteckfilter für theoretische Überlegungen zu simulieren, können idealisierte, äquivalente Rechteckfilter (*equivalent rectangular bandwidth*, ERB) errechnet werden (vgl. Gelfand 2004, 322; Hellbrück & Ellermeier 2008, 57; Moore 1993, 63f.). Allerdings muß man bedenken, daß tatsächliche Filter und ihre ERB unterschiedliche Bandbreiten haben, wenn beide die gleiche Energiemenge passieren lassen. Während die Frequenzgruppe zentriert bei 1 kHz grob etwa 160 Hz weit ist, beträgt die korrespondierende ERB nur zwischen 128 und 133 Hz (vgl. Hellbrück & Ellermeier 2008, 58; Moore 1993, 77).

1.3.2 Filterfunktion

Frequenzgruppen sind für verschiedene Hörphänomene verantwortlich. Besonders trifft dies auf die minimale Frequenzauflösung, die Bildung von Lautheitseindrücken, speziell komplexer Signale und *Verdeckung* zu (vgl. Gelfand 2004, 353; Hellbrück & Ellermeier 2004, 89; Moore 1993, 81; Spitzer 2002, 100; Spreng in Keidel 1975, 142). Aber auch viele – hier nicht behandelte – Phänomene des Streamings, bzw. der auditiven Szenenanalyse lassen sich auf den Einfluß der Frequenzgruppen zurückführen (vgl. Reuter 1996, 36f., 67, 293).

Die Frequenzauflösung des Gehörs ist nicht unbegrenzt gut. Zwei Töne sehr ähnlicher Frequenz liegen nahe beieinander und werden nicht als zwei Töne, sondern als Schwebung wahrgenommen. Vergrößert man den Frequenzabstand zwischen beiden Tönen, wandelt sich die Schwebung zu Rauigkeit (ungefähr ab $\Delta f = 15$ Hz). Erst bei einem Abstand von einer Frequenzgruppe, d.h. wenn beide Töne

schließlich in unterschiedliche Frequenzgruppen fallen, werden tatsächlich zwei eigenständige Töne wahrgenommen (vgl. Hellbrück & Ellermeier 2004, 125, 140; Pollard & Jansson 1982, 251; Spitzer 2002, Abb. 3.9, 60, 94, 99, 100).

Lautstärke wird nicht insgesamt (über den kompletten hörbaren Frequenzraum), sondern nur innerhalb der Frequenzgruppen integriert. Die Gesamtintensität wird also durch das Gehör in Teilintensitäten gespalten, die separat verarbeitet werden. Dies bedeutet zugleich, daß das Gehör nicht eine breite Mithörschwelle bildet, sondern mit separaten, frequenzgruppenbreiten (Mit-)Hörschwellen operiert (vgl. Zwicker & Feldtkeller 1967, 57, 68, 73, 109). Allgemein gilt: Innerhalb einer Frequenzgruppe wird die Schallintensität addiert. Zwei pegelgleiche Töne klingen also nur so laut, wie ein alleiniger Ton, dessen Pegel um 3 dB angehoben wurde. Bei konstanter Intensität sind außerdem Veränderungen der Signalbandbreite (Tonabstand, Δf Rauschen) im Bezug auf die Lautheit irrelevant. Außerhalb einer Frequenzgruppe werden die Lautheiten addiert. Die beiden Töne klingen nun also viel lauter, als innerhalb der Frequenzgruppe. Auch Erweiterungen der Signalbandbreite führen, trotz weiterhin physikalisch konstanter Intensität, zu proportionalen Lautheitsanstiegen (vgl. Gelfand 2004, 355; 2009, 94; Guski 1996, 244f.; Hellbrück & Ellermeier 2004, 140, 141; 2008, 61; Moore 1993, 77; Winckel 1957, 84; Zwicker & Feldtkeller 1967, 123ff., 133–139).

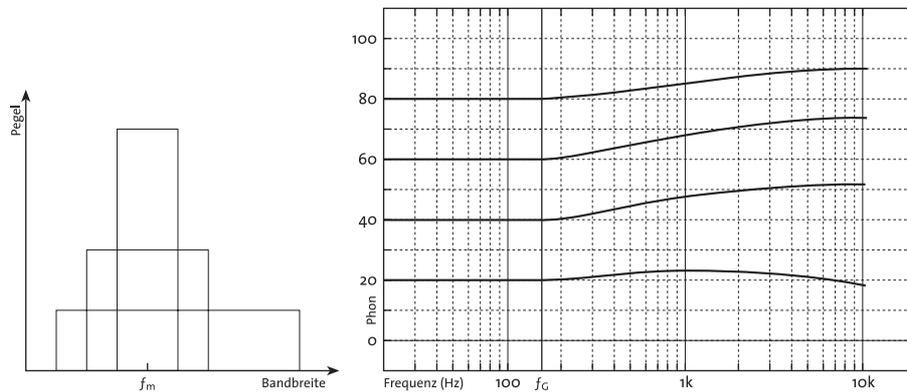


Abb. 1.2 **Lautheitsänderung durch Bandbreitenerweiterung** (Zwicker & Feldtkeller 1967, Abb. 48.1 & 48.2, 123f.). Links: Alle Rechtecke haben die gleiche Fläche, repräsentieren also die gleiche Schallenergie. Erhöht man die Bandbreite und reduziert entsprechend den Druckpegel, bleibt die Schallintensität konstant. Rechts: Übersteigt die Bandbreite die Frequenzgruppenbreite (f_c), wächst die Lautheit trotz konstanter Intensität an.

Der Einfluß von Frequenzgruppen auf Signalverdeckung ist auf die weiter oben bereits angeführte Teilung der Hörschwelle in unabhängige schmalbandige Abschnitte (Δf = Frequenzgruppenbreite) zurückzuführen und läßt sich (in aller Kürze) wie folgt beschreiben: Wie schon erwähnt wird die Gesamtintensität in frequenzgruppenbreite Teilintensitäten zerlegt. Es werden also nur die Energieanteile des Maskierers innerhalb der fraglichen Frequenzgruppe zur Bildung der Mithörschwelle verwendet. Außerhalb des Bandes liegende Schallanteile sind nicht verdeckungsrelevant. Liegt der Testschallpegel über dem Frequenzgruppenpegel des Maskierers, wird der Testschall wahrgenommen – unabhängig von aller weiteren Schallenergie des Maskierers außerhalb der Frequenzgruppe (vgl. Guski 1996, 244; Hellbrück & Ellermeier 2004, 126; Zwicker & Feldtkeller 1967, 57, 73, 109).

1.4 Lautstärke

Ein großer Komplex der Psychoakustik und der Physiologie, mit nach wie vor vielen offenen Fragen, ist die Wahrnehmung von Lautstärke. Einer der Gründe dafür ist wahrscheinlich, daß die Lautstärke nur schwer systematisch kategorisiert und gemessen werden kann (vgl. Fricke & Louven 2008, 421).

Im allgemeinen Sprachgebrauch wird die Lautstärke gerne mit dem Schallpegel gleichgesetzt. Physikalisch ist allein der Schalldruck, in einer logarithmischen Skala als Schalldruckpegel angegeben, relevant. Die korrespondierende Empfindung subjektiver Lautstärke ist die *Lautheit* (vgl. Fricke & Louven 2008, 419; Gelfand 2004, 347; Zwicker & Feldtkeller 1967, 1f., 44). Praktisch wird die Empfindungsgröße Lautheit in die Größen *Lautstärkepegel* und *Verhältnislautheit* aufgeteilt. Der Lautstärkepegel trägt der Lautheitsabhängigkeit vom Schalldruckpegel, allerdings auch ihrer Frequenzabhängigkeit, Rechnung; die Verhältnislautheit gibt Auskunft darüber, in welchem Empfindungsverhältnis zwei verschiedene Lautheiten zueinander stehen (vgl. Fricke & Louven 2008, 419f.; Gelfand 2004, 351; Hellbrück & Ellermeier 2004, 82, 138).

Über Pegel und Frequenz hinaus ist die Lautheitsbildung allerdings noch von weiteren Faktoren abhängig. Die Abhängigkeit von der Signaldauer wurde bereits im Rahmen der *Zeitauflösung des Gehörs* (siehe 1.2, S. 11) dargestellt, die Abhängigkeit von der Bandbreite im Rahmen der *Frequenzgruppen* (siehe 1.3, S. 18).

1.4.1 Reizgröße und Empfindungsgröße

Der Schalldruck ist im Grunde der adäquate Reiz des Hörens insgesamt (vgl. Zwicker & Feldtkeller 1967, 42). Die Menge an Schall(druck), die das Gehör auf einmal zu verarbeiten hat, macht sich in der Sinnesqualität der Lautstärke bemerkbar. Da es sich bei Schalldruck zunächst um nichts anderes als extrem winzige Luftdruckschwankungen handelt, liegt es nahe, Schalldruck in derselben Einheit, also in (Mikro-)Bar (μb) bzw. (Mikro-)Pascal (μPa) zu berechnen (vgl. Zwicker & Feldtkeller 1967, 1). Um die extreme Dynamikspanne des Gehörs halbwegs übersichtlich zu halten, werden Schalldrücke in logarithmischen Pegelskalen angegeben. Wie weiter oben schon dargestellt (siehe 1.1.1, S. 8), sind Pegel an sich dimensionslos, können also nicht als Zahl mal Einheit dargestellt werden, sondern bedürfen eines Bezugswerts (vgl. Müller & Frings 2009, 447; Stevens & Hallowell 1938, 110; Zwicker & Feldtkeller 1967, 1). Das Ergebnis ist der *Schalldruckpegel* (L) mit dem Bezugswert $20 \mu\text{Pa}$ (vgl. Fastl & Zwicker 2006, 18; Lehnhardt & Laszig 2009, 4; Müller & Frings 2009, 447; Yost 2008, 146). Der Schalldruckpegel ist als Reizgröße unabhängig von der physiologischen Qualität der Empfindung (vgl. Fricke & Louven 2008, 419).

Die Lautstärke (engl. loudness) ist prinzipiell die korrespondierende Empfindungsgröße zur physikalischen Reizgröße Schalldruck. Der Schalldruck wird in einem Pegelverhältnis angegeben und auch die Lautstärke wird, auf den Schalldruckpegel eines gleichlauten Referenzschalls bezogen, als *Lautstärkepegel* (L_N) angegeben (vgl. Gelfand 2004, 347; Hellbrück & Ellermeier 2004, 72; Stevens & Hallowell 1938, 111; Zwicker & Feldtkeller 1967, 118).

Der *Lautstärkepegel* wird in der von Barkhausen eingeführten Einheit *Phon* gemessen. Als Bezugspunkt zwischen Phon und Schalldruckpegel wurde die Frequenz 1 kHz definiert. Hier sind Phon- und Schallpegelwerte identisch (vgl. Fastl & Zwicker 2006, 159; Fricke & Louven 2008, 420; Gelfand 2004, 348; 2009, 92; Hellbrück & Ellermeier 2004, 72; 2008, 55; Plattig in Keidel 1975, 374f.; Stevens & Hallowell 1938, 111; Zwicker & Feldtkeller 1976, 118). Signale unterschiedlicher Frequenz und gleicher Lautstärke bekommen, trotz unterschiedlicher Pegel, den gleichen Phonwert zugeschrieben. So entsteht eine Linie, die gleich laute Pegel als Funktion der Frequenz miteinander verbindet. Bezieht man Signale unterschiedlicher Frequenz auf verschieden starke Referenzpegel (Phon bei 1 kHz), bekommt man eine Reihe von Kurven, die gleiche Lautstärken angeben und folglich als *Kurven gleicher Lautstärkepegel* (Isophonkurven, Isophonen) bezeichnet wurden (vgl. Fasold & Veres 2003, 47; Fricke & Louven 2008, 420; Gelfand 2004, 348; 2009, 91f.; Hellbrück & Ellermeier 2004, 72; 2008, 55). Sie verdeutlichen, daß die subjektiv empfundene Lautstärke primär pegelabhängig, sekundär aber auch frequenzabhängig ist

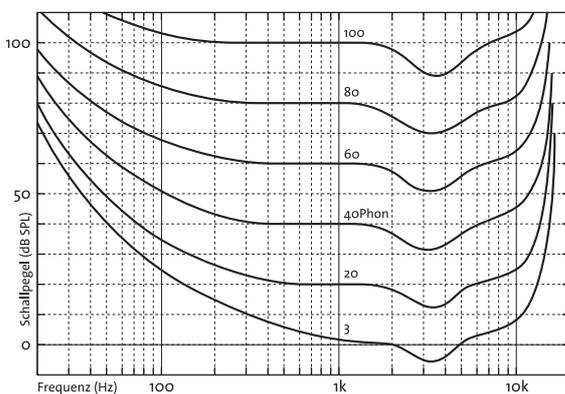


Abb. 1.3 **Kurven gleicher Lautstärkepegel** (Zwicker & Feldtkeller 1967, 46.1, 121). Per Definition stimmen Schallpegel- und Phonwerte bei 1 kHz überein. Töne, die gleich laut wie der Bezugston bei 1 kHz erscheinen, werden durch eine Phonkurve verbunden. Alle Töne entlang dieser Kurve werden, trotz unterschiedlicher Schallpegel, als gleich laut empfunden.

(vgl. Hellbrück & Ellermeier 2004, 138). Weil Fletcher und Munson (1933) als erste allgemein anerkannte Isophonkurven veröffentlichten, werden sie auch *Fletcher-Munson-Kurven* genannt (vgl. Gelfand 2004, 349; 2009, 92).

Verschiedene Studien brachten immer wieder leicht divergente Phonkurven hervor. Im Kern sind sie sich jedoch sehr ähnlich und zeigen stets denselben Grundverlauf (vgl. Gelfand 2004, 349). Durch solche Schwankungen sind die Phonkurven prinzipiell derselben Kritik wie die Hörschwelle ausgesetzt: Unter verbesserten Bedingungen durchgeführte Messungen wichen zum Teil um mehr als 10 dB von den traditionellen Fletcher-Munson-Kurven ab (vgl. Hellbrück & Ellermeier 2004, 72). Dennoch haben sie sich weitestgehend durchgesetzt.

Die Phonkurven strukturieren den Hörbereich oberhalb der Hörschwelle. Bei geringen Lautstärken folgen die Kurven gleicher Lautstärke dem Verlauf der Hörschwelle (MAF). An den Rändern des Hörbereichs, wo die Hörschwelle bei höheren Pegeln liegt, sind die Phonkurven in kleinere Pegelbereiche komprimiert. Zu hohen Pegeln hin verlaufen sie zunehmend gerader (vgl. Fricke & Louven 2008, 420; Gelfand 2004, 349; 2009, 92; Hellbrück & Ellermeier 2008, 55; Plattig in Keidel 1975, 375). Folglich müssen die Pegelschritte für einen gleichen Lautstärkezuwachs frequenzabhängig differieren. An den Rändern des Hörbereichs ist der Lautstärkezuwachs bei gleichem Pegelzuwachs deutlich größer als im mittelfrequenten Bereich (vgl. Fricke & Louven 2008, 420; Gelfand 2004, 349; 2009, 92). Lautstärkeempfindungen sind daher auch frequenzabhängig (vgl. Fricke & Louven 2008, 419; Gelfand 2004, 349; 2009, 91).

Um der Frequenzabhängigkeit der Lautstärkeempfindung Rechnung zu tragen, wurden frequenzgewichtete Skalen entwickelt, die praktisch umgekehrte und vereinfachte Hörschwellen bzw. Phonkurven

darstellen. Sie werden als L_A , L_B bzw. L_C bezeichnet und in (den Einheiten) dB_A , dB_B bzw. dB_C gemessen (vgl. Fasold & Veres 2003, 48; Gelfand 2004, 349; Hellbrück & Ellermeier 2004, 76, 77; 2008, 56). Die Skalen orientieren sich an der Frequenzspanne des Gehörs bei verschiedenen Pegeln. Die A-Skala orientiert sich dabei an den Phonkurven niedriger Pegel bis 30 Phon, vernachlässigt also den tieffrequenten Bereich. Die B-Skala orientiert sich an den Phonkurven mittlerer Pegel bis 60 Phon, verläuft also relativ gerade. Die C-Skala ist dem Frequenzgang des Ohrs bei hohen Lautstärkepegeln über 60 Phon nachempfunden, demnach praktisch durchgängig flach (vgl. Fasold & Veres 2003, 16; Gelfand 2004, 349; Hellbrück & Ellermeier 2004, 77; 2008, 56). Kurioserweise bedient man sich bei Lärmmessungen vereinbarungsgemäß ausgerechnet der A-Skala, obwohl Lärm i.d.R. recht hohe Pegel aufweist, die Skala aber für leise Pegel geschaffen wurde. Die übliche Rechtfertigung argumentiert damit, daß tiefe Frequenzen weniger schädlich seien (vgl. Fasold & Veres 2003, 50f.; Hellbrück & Ellermeier 2004, 77; 2008, 56).

Der Lautstärkepegel kann kein reines Empfindungsmaß sein. Wie die Zusammensetzung aus *Lautstärke* und *Pegel* schon andeutet, handelt es sich dabei (nur) um ein halb subjektives und halb objektives Maß (vgl. Fricke & Louven 2008, 420; vgl. Hellbrück & Ellermeier 2004, 46; Zwicker & Feldtkeller 1967, 126). Die subjektiv empfundene Lautstärke, die Empfindungsgröße, wird mit dem Begriff *Lautheit* (engl. ebenfalls loudness) gekennzeichnet. Die Lautheit als Empfindungsgröße wird allerdings oft mit der *Verhältnislautheit* (N) gleichgesetzt (vgl. Hellbrück & Ellermeier 2004, 82; 2008, 59; Zwicker & Feldtkeller 1967, 44). Die Verhältnislautheit, gemessen in der Einheit *Sone*, gibt Auskunft über das empfundene Verhältnis zweier Lautstärken, ist also die Antwort auf die Frage, um wieviel ein Signal lauter als ein anderes empfunden wird (vgl. Fasold & Veres 2003, 48; Fricke & Louven 2008, 420; Gelfand 2004, 351; Hellbrück & Ellermeier 2004, 82; Stevens & Hallowell 1938, 119).

Ein Sone wurde einem Schalldruckpegel von $40\text{ dB}_{\text{SPL}}$ bei 1 kHz gleichgesetzt. Die Soneskala orientiert sich damit ebenfalls am Bezugspunkt 1 kHz . D.h. 40 Phon entsprechen 1 Sone . Die Soneskala wurde auf 40 Phon geeicht, weil unterhalb von 40 Phon das *Potenzgesetz* (siehe 1, S. 6) nicht mehr greift (vgl. Fasold & Veres 2003, 48; Fricke & Louven 2008, 420; Gelfand 2004, 351; 2009, 93; Hellbrück & Ellermeier 2004, 84; 2008, 59; Stevens & Hallowell 1938, 117f.). Oberhalb von 40 Phon gehen Pegelerhöhungen um 10 dB meist mit Lautheitsverdoppelungen einher. Hier entsprechen gleiche Pegelverhältnisse offenbar gleichen Wahrnehmungsverhältnissen. Ausgehend von 1 Sone werden Lautheitsverdoppelungen durch doppelte Sonewerte (2 Sone) angegeben und umgekehrt (vgl. Fricke & Louven 2008, 420; Gel-

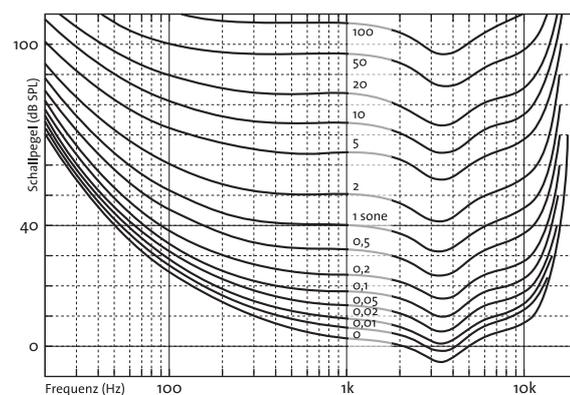


Abb. 1.4 **Kurven gleicher Lautheit** (Zwicker & Feldtkeller 1967, 49.3, 129). Unterschiedliche Lautheiten werden in der Einheit Sone gemessen. Töne werden zu einem Referenzton bei 1 kHz und 40 Phon in Bezug gesetzt. Gleich laut erscheinende Töne werden zur Referenzkurve bei 1 Sone verbunden, alle anderen Kurven stellen Lautheitsverhältnisse (meist Verdoppelungen oder Halbierungen) zu dieser Linie dar. In der Abbildung ist zu erkennen, daß bei 1 kHz über 40 Phon Lautheitsverdoppelungen meist ungefähr einem Schallpegelzuwachs von 10 dB entsprechen.

fand 2004, 348, 351; 2009, 93; Hellbrück & Ellermeier 2004, 82, 84; Stevens & Hallowell 1938, 117). Die Verhältnislautheit ist eine Funktion der Phonkurven. Da Phonwerte bereits „frequenzbereinigte“ Angaben liefern, können Sone praktisch als frequenzunabhängig angesehen werden. Aus gemessenen Sonewerten ergeben sich – analog zu den Phonkurven – Sonekurven, die *Kurven gleicher Lautheit*. Sie sind den Kurven gleicher Lautstärke ziemlich ähnlich (vgl. Fricke & Louven 2008, 420f.; Gelfand 2004, 351).

1.4.2 Lautstärkewahrnehmung

Lautheitsänderungen von leise nach laut werden, trotz jeweils gleicher Reizpegeländerung, deutlich stärker wahrgenommen als umgekehrt. Einem nicht unumstrittenen Erklärungsversuch nach wird auf einen zukommenden, also lauter werdenden Schallquellen größere Bedeutung beigemessen als sich entfernenden, also leiser werdenden Schallquellen (vgl. Hellbrück & Ellermeier 2004, 146).

Bei höheren Schallpegeln ist das Gehör auch empfindlicher für Amplitudenmodulationen (vgl. Zwicker & Feldtkeller 1967, 94f.). Bei 20 dB_{SPL} können Modulationen von 10% wahrgenommen werden, bei 40 dB_{SPL} schon Modulationen von 6%, bei 60 dB_{SPL} 4% und bei 100 dB_{SPL} sogar Modulationen von 1% des Pegels. Zwischen 20 und 100 dB_{SPL} verringert sich die Modulationsschwelle also um den Faktor zehn (vgl. Hellbrück & Ellermeier 2004, 81; Zwicker & Feldtkeller 1967, 95; andere Werte aber tendenziell die gleiche Richtung auch bei Fasold & Veres 2003, 51).

Die Unterschiedsschwelle bei 40 dB_{SPL} liegt bei 0,4 dB, bei 100 dB_{SPL} bei 0,2 dB. Das Gehör ist für diskrete Pegeländerungen also prinzipiell empfindlicher als für Modulationen (vgl. Hellbrück & Ellermeier 2004, 81). Die Ausnahme bilden extrem hohe Pegel über 100 dB_{SPL}. Hier ist die Modulationsempfindlichkeit größer (vgl. Hellbrück & Ellermeier 2004, 81). Daß die Unterschiedsempfindlichkeit zu höheren Pegeln hin besser wird, anstatt (wie das Weber'sche Gesetz voraussagt) konstant zu bleiben, wird auch als *near miss to Weber's law* bezeichnet (vgl. Green 1993, 22; Hellbrück & Ellermeier 2004, 81; Yost 1993, 5).

Welche physiologischen Prozesse an der Lautstärkewahrnehmung beteiligt sind, ist noch nicht ganz geklärt (vgl. Hellbrück & Ellermeier 2004, 111). An dieser Stelle ist ein Ausflug in die neuronale Schallkodierung unvermeidlich: Steigt die Schallintensität, wird der Auslenkungsbereich auf der Basilarmembran breiter, wodurch auch die Anzahl der involvierten Haarzellen größer wird. Dies legt zumindest den Schluß nahe, daß ein Konzept der Lautstärkenkodierung die Masse der Nervenimpulse sein muß (vgl. Hellbrück & Ellermeier 2004, 139; Spreng in Keidel 1975, 132).

Eine weitere Feststellung ist, daß die Feuerrate bestimmter Nervenfasern in Abhängigkeit zur Reizintensität steht. Je stärker der Reiz, desto höher die Feuerrate. Allerdings wächst die Feuerrate bei einem Reizintensitätszuwachs um den Faktor 100 nur um den Faktor 4. Grund dafür ist die Begrenzung der Feuerrate einzelner Neurone durch ihre Refraktärzeit (vgl. Hellbrück & Ellermeier 2004, 139). Dies begrenzt die Dynamikspanne einer isolierten Nervenfaser auf etwa 30 bis 40 dB (vgl. Guski 1996, 246; Hellbrück & Ellermeier 2004, 111; Yost 1993, 5). Ein Neuron ist so bei steigendem Reizpegel gesättigt, bevor die maximal wahrnehmbare Dynamikspanne von etwa 120 bis 130 dB durchlaufen wurde (vgl. Guski 1996, 94; Hellbrück & Ellermeier 2004, 111f., 134; Yost 1993, 4). Vermutlich sind einzelne Ner-

venfasern nicht nur auf Frequenzen, sondern auch auf Dynamikbereiche spezialisiert, d.h. unterschiedlich empfindlich, so daß die volle Dynamikspanne wohl in verschiedenen Registern verarbeitet wird (vgl. Hellbrück & Ellermeier 2004, 139).

Die Synchronisierung der Feuerraten wird mit steigendem Reizpegel vermutlich immer besser (vgl. Hellbrück & Ellermeier 2004, 139). Die Lautstärkewahrnehmung bzw. Intensitätskodierung funktioniert wahrscheinlich als Kombination aus der verbesserten Synchronisation, der größeren Masse an erregten Fasern und der steigenden Erregung (alle drei abhängig vom steigendem Reizpegel) sowie der Spezialisierung verschiedener Fasern auf bestimmte Pegelbereiche (vgl. Hellbrück & Ellermeier 2004, 139).

1.5 Tonhöhe

„Die Tonhöhe gehört neben dem Rhythmus zu den strukturbildenden Parametern der Musik, die im Vergleich zu Lautstärke und Klangfarbe in viel höherem Grade systemfähig sind“ (Fricke & Louven 2008, 421). Der Grund dafür ist, daß Tonhöhe (wie auch Rhythmus) durch zeitliche Faktoren bestimmt ist und das neuronale System passenderweise auf der zeitlichen Ebene agiert. Das liegt zum einen besonders daran, daß Tonhöhen (wie auch der Rhythmus) durch zeitliche Faktoren bestimmt sind, was der Arbeitsweise des neuronalen Systems entgegenkommt. Zum anderen liegt es daran, daß die Tonhöhenwahrnehmung nicht kontinuierlich, sondern kategorial funktioniert. Dadurch werden die vielen (unvermeidbaren) musikalischen Ungenauigkeiten – innerhalb gewisser Grenzen – durch das Gehör auf Sollwerte quantisiert, die sich ihrerseits gehörmäßig gut in Gefüge ordnen lassen (vgl. Fricke & Louven 2008, 421).

Tonhöhe ist ein ungenauer Begriff, der in der Physik, Musik oder Wahrnehmungspsychologie jeweils anders verwendet wird (vgl. Zwicker & Feldtkeller 1967, 78). Die Physik setzt den Tonhöhenbegriff im Grunde mit der Frequenz gleich. In der Musik wird unter Tonhöhe eine konstruierte Harmonie verstanden, deren Intervalle im Prinzip nichts anderes als Frequenzverhältnisse sind. Die Psychoakustik versteht die Tonhöhe als Antwort auf die Frage, wie hoch etwas gehört wird, bzw. welcher Stimulus (um wie viel) höher oder tiefer als der andere ist (vgl. Zwicker & Feldtkeller 1967, 78, 81). Die letztgenannte Interpretation wird auch *Verhältnistonhöhe* genannt. Um ihren Status als Empfindungsgröße zu dokumentieren und sie zugleich von den alltäglichen und reizbezogenen Tonhöhenkonzepten zu unterscheiden, wurde ihr der Begriff *Tonheit* zugesprochen (vgl. Hellbrück & Ellermeier 2004, 87; Zwicker & Feldtkeller 1967, 82).

1.5.1 Reizgröße und Empfindungsgröße

Die Tonhöhenwahrnehmung hängt mit der Frequenz zusammen; genaugenommen mit der Grundtonfrequenz (Periodizität) von Klängen und der Frequenz von Sinustönen (vgl. Fricke & Louven 2008, 424; Gelfand 2004, 365, 377ff.; Moore 1993, 88). Die Frequenz, mit der Einheit *Hertz*, ist die physikalische Größe des Reizes, also eine *Reizgröße* (vgl. Stevens & Hallowell 1938, 70; Zwicker & Feldtkeller 1967,

81f.). In der Musik bezeichnet die Tonhöhe die Höhe oder Lage eines Tons im Tonsystem bzw. der entsprechenden Tonleiter (vgl. Gelfand 2004, 371; Zwicker & Feldtkeller 1967, 78). Wenn die Frequenz in einem bestimmten Verhältnis erhöht (erniedrigt) wird, ändert sich die musikalische Tonhöhe in einem bestimmten Intervall. Eine Verdopplung (Halbierung) ist z.B. immer eine Oktave. Die musikalische Tonhöhe (von Zwicker und Feldtkeller als *harmonische Tonhöhe* bezeichnet) ist also nichts anderes als ein Logarithmus der Frequenz (vgl. Zwicker & Feldtkeller 1967, 81). Die Einheit der harmonischen Tonhöhe ist die *Oktave*. Die Oktavperiodizität bildet die Grundlage der westlichen Tonsysteme (vgl. Gelfand 2004, 373; Zwicker & Feldtkeller 1967, 81). Klänge in Oktavlage tendieren bei simultaner Darbietung dazu, zu verschmelzen. Frauen- und Männerstimmen werden beim Singen der gleichen Melodie als einstimmig betrachtet, obwohl sie naturgemäß zueinander oktaviert sind (vgl. Zwicker & Feldtkeller 1967, 81). Der Begriff *unisono* ist also praktisch ein oktavunempfindlicher Begriff für Einstimmigkeit. Das wissenschaftliche Bedürfnis, die harmonische Tonhöhe absolut zu bestimmen führte dazu, daß Bezugstöne definiert wurden. Zwicker und Feldtkeller (vgl. 1967, 82) legten den Ton c_0 (131 Hz, kleines C, C_3) als Bezugston fest. Wenn man wollte, könnte man ihn als „wissenschaftlichen oder psychoakustischen Kammerton“ bezeichnen. Der Index gibt – der englischen Formatierung vergleichbar – die Oktavlage an, wobei tiefere Oktaven mit negativen Indizes versehen werden (vgl. Zwicker & Feldtkeller 1967, 82). Es läßt sich also festhalten: Frequenz und harmonische Tonhöhe (Logarithmus der Frequenz) sind die Reizgrößen (vgl. Zwicker & Feldtkeller 1967, 81f.).

Wie schon erwähnt, wurde der Verhältnistonheit, um sie vom allgemeinsprachlichen Bedeutungsstrauß des Begriffs „Tonhöhe“ zu isolieren, der Begriff *Tonheit* (z) zugeordnet (vgl. Hellbrück & Ellermeier 2004, 87; 2008, 65; Zwicker & Feldtkeller 1967, 82). Die Tonheit stellt die subjektiv empfundene Tonhöhe, also die *Empfindungsgröße* dar (vgl. Gelfand 2004, 373; Zwicker & Feldtkeller 1967, 81). Als ihre Einheit führte Stevens das – aufgrund der engen Verbindung von Tonheit und Melodieempfindung – vom englischen Wort ‚melody‘ abgeleitete *mel* ein (vgl. Fastl & Zwicker 2006, 112, Gelfand 2004, 365; 2009, 94; Hellbrück & Ellermeier 2004, 87; 2008, 65; Stevens & Hallowell 1938, 81; Zwicker & Feldtkeller 1967, 82). Auf Stevens geht auch der erste Bezug zur Frequenzskala (1000 Hz = 1000 mel, bei 40 Phon) zurück, der besonders im englischen Sprachraum immer noch Verwendung findet (vgl. Gelfand 2004, 366f.; 2009, 94; Stevens & Hallowell 1938, 81). Zwicker und Feldtkeller richteten die Mel-Skala dagegen am weiter oben schon erwähnten Bezugston c_0 (131 Hz) aus und schrieben ihm 131 mel zu (vgl. Zwicker & Feldtkeller 1967, 82). Letzterer genießt m.E. den Vorteil, sich gleichzeitig adäquat auf musikalische Tonhöhen zu beziehen (ungeachtet der wenigen Hertz Unterschied, die bei verschiedenen Frequenzuteilungen der harmonischen Tonhöhen immer wieder auftauchen).

Der adäquate Reiz für die Empfindung Tonheit ist die Frequenz. So wie die Frequenz von einer logarithmischen Tonhöhenkala (harmonische Tonhöhe) begleitet wird, läßt sich auch die Tonheit in einer logarithmischen Tonhöhenkala, der *melodischen Tonhöhe*, ausdrücken. Sie funktioniert analog zur harmonischen Tonhöhe, indem bestimmte Tonheitsverhältnisse in bestimmten *melodischen* Intervallen wahrgenommen werden (vgl. Zwicker & Feldtkeller 1967, 81, 82f., 90, 93). An dieser Stelle ist also festzuhalten: Tonheit und melodische Tonhöhe (Logarithmus der Tonheit) sind die Empfindungsgrößen (vgl. Fricke & Louven 2008, 423; Zwicker & Feldtkeller 1967, 81).

1.5.2 Tonhöhenwahrnehmung

Unter 500 Hz entsprechen sich Frequenz und Tonhöhenempfindung, also auch *harmonische* und *melodische Tonhöhe*. Jede Oktave enthält doppelt so viele Tonheitsstufen (mel) wie die nächst tiefere. D.h. Empfindungsgröße und Reizgröße stimmen überein, was sich auch zahlenmäßig niederschlägt. Hertz- und Melskala verlaufen deckungsgleich: 131 Hz sind auch 131 mel, auch für den gegenwärtig im westlichen Tonsystem gebräuchlichsten Kammerton a' (440 Hz) trifft dies im Prinzip noch zu (vgl. Fastl & Zwicker 2006, 112f.; Gelfand 2004, 373; Hellbrück & Ellermeier 2004, 86f.; 2008, 65; Zwicker & Feldtkeller 1967, 82, 93). Ab etwa 500 Hz werden die empfundenen Intervalle immer enger. Sie bleiben hinter den tatsächlichen Frequenzschritten zurück. D.h. es sind immer größere Frequenzschritte notwendig, um die Tonheitsschritte konstant zu halten (vgl. Fastl & Zwicker 2006, 113; Fricke & Louven 2008, 423; Gelfand 2004, 367; 2009, 96; Hellbrück & Ellermeier 2004, 86; Stevens & Hallowell 1938, 83f.; Zwicker & Feldtkeller 1967, 80). Anders gesagt: „Transponiert man eine Melodie [aus Sinustönen] aus tiefer Lage harmonisch korrekt in eine hohe Tonlage, so wird sie melodisch zu eng. Transponiert man sie aber melodisch korrekt, so wird sie harmonisch unbrauchbar“ (Zwicker & Feldtkeller 1967, 93). Melodien werden in höheren Lagen also immer enger gehört (Zwicker & Feldtkeller 1967, 93; Hellbrück & Ellermeier 2004, 86f.; 2008, 65). Die Tonhöhenempfindung kann nicht mit den musikalischen Intervallen (Frequenzverhältnissen) schritthalten. Ein tatsächliches Oktavverhältnis ($\Delta f = 2:1$) wird mitunter nur noch als Quinte ($\Delta z = 3:2$) wahrgenommen (vgl. Fricke & Louven 2008, 423; Zwicker & Feldtkeller 1967, 79). Verdoppelt man einen 1050 mel Ton (1,3 kHz), wird man feststellen, daß 2100 mel nicht bei den zu erwartenden 4 kHz, sondern bei 8 kHz liegen (vgl. Fastl & Zwicker 2006, 111, 112f.; Gelfand 2004, 367; Hellbrück & Ellermeier 2004, 86; 2008, 65; Zwicker & Feldtkeller 1967, 80). Über 500 Hz entspricht jede Oktave nur noch konstant 100 mel (vgl. Zwicker & Feldtkeller 1967, 93).

Legt man die Mel-Skala mit dem Bezugston c_0 zugrunde, entsprechen diese 100 mel genau 1 Bark. Die Tonheitsempfindung steht offenbar in Zusammenhang mit den Frequenzgruppenbreiten. Zumindest ist die Barkskala mit der Melskala verknüpft. In diesem Sinne hat man für die Stufe von 100 mel den Begriff *Tongruppe* geschaffen (vgl. Fastl & Zwicker 2006, 164; Gelfand 2004, 367f.; Hellbrück & Ellermeier 2004, 88, 128; 2008, 65; Zwicker & Feldtkeller 1967, 90). Innerhalb einer Frequenzgruppe kann das Ohr 27 verschiedene Tonhöhen, d.h. 27 Frequenzstufen unterscheiden. Die Tonhöhenauflösung einer Frequenzgruppe beträgt also 1 : 27 Bark, was rein rechnerisch etwa 3,7 mel entspricht (vgl. Hellbrück & Ellermeier 2004, 128; Moore 1993, 91). Das hieße natürlich, daß erst eine Veränderung der Tonheit um 3,7 mel die empfundene Tonhöhe verändert. Zwicker und Feldtkeller gehen von 3,9 mel aus und kommen so auf 620 hörbare Tonhöhenstufen, die sie in 140 gleich weite Stufen unter 500 Hz und 480 proportional zur Frequenz weiter werdende Stufen zwischen 500 Hz und 16 kHz unterteilen (vgl. Zwicker & Feldtkeller 1967, 92f.). Hellbrück und Ellermeier (vgl. 2004, 128) geben mit insgesamt 640 ähnlich viele erkennbare Tonstufen an, während Fasold und Veres (vgl. 2003, 51) 850 Stufen angeben und Stevens und Hallowell sogar 1500 erkennbare Stufen hochrechnen (vgl. 1938, 95).

Die empfundenen Tonhöhen bleiben ab 500 Hz immer weiter hinter den entsprechenden Frequenzstufen zurück. Während das hörbare Frequenzspektrum bis 16 bzw. 20 kHz reicht, steigt die Tonheit nur bis 2400 mel (vgl. Fastl & Zwicker 2006, 113; Fricke & Louven 2008, 423; Zwicker & Feldtkeller

1967, 79, 82). Mel sind – wenn man so will – die „gehörte Frequenz“. Hat ein Ton z.B. 1 kHz, was ungefähr 800 mel ausmacht, wird eine Tonhöhe empfunden, die bei einem linearen Frequenz/mel Verhältnis etwa 0,8 kHz entsprechen würde. Anders gesagt: Ein Ton, der seine Periode 1000 mal in einer Sekunde durchläuft (1 kHz), hört sich an, als hätte er 0,8 kHz.

Zu dem „Rückstand“ auf die Frequenz kommt ab ungefähr 5 kHz eine gewisse Unbestimmtheit hinzu. Sinustonfolgen können keinen klaren Melodieindruck mehr erzeugen. Die Oktavwahrnehmung gelingt offenbar auch nur, wenn beide Intervallfrequenzen unter 5 kHz liegen. Vermutlich ist dafür ein Prinzipienwechsel in der physiologischen Tonhöhenenerkennung um 5 kHz verantwortlich: Periodische Muster (die zeitliche Struktur des Signals, also auch die Frequenz) werden aufgrund neuronaler System-einschränkungen nicht mehr präzise verarbeitet, so daß die Tonhöhe nur noch tonotopisch, d.h. den Erregungsort auf der Basilarmembran, bestimmt – oder, gemessen an der mangelnden Präzision, eher „geschätzt“ – werden kann (vgl. Moore 1993, 94).

Unterschiedsschwellen

Um die kleinste hörbare Veränderung, den *ebenmerklichen Unterschied* (EMU, just noticeable difference, jnd), zu ermitteln, gibt es prinzipiell zwei Möglichkeiten. Entweder man sucht nach der *Modulationsschwelle* (frequency modulation difference limen, FMDL), d.h. der Weite, die eine Signalschwankung haben muß, um erkannt zu werden; oder man sucht nach der *Unterschiedsschwelle* (difference limen, DL, Δf), d.h. dem Unterschied, den zwei diskrete Signale aufweisen müssen, um als verschieden erkannt zu werden (vgl. Fastl & Zwicker 2006, 175; Gelfand 2004, 288, 365; 2009, 89; Hellbrück & Ellermeier 2004, 80f., 216; Moore 1993, 89; Plattig in Keidel 1975, 359f.; Stevens & Hallowell 1938, 84).

Frequenzunterschiedsschwellen sind im Verhältnis zur Mittenfrequenz (des Unterschieds) im mittelfrequenten Bereich am kleinsten und werden zu den sehr hohen und sehr tiefen Frequenzen hin größer, wobei die Schwankungen nicht unbedingt linear mit der Frequenz verlaufen. Im Vergleich fallen die frequenzabhängigen Schwankungen der Modulationsschwellen weniger stark aus (vgl. Gelfand 2004, 293; Moore 1993, 89; Winckel 1952, 57; Yost 2008, 151). Zu hohen und tiefen Frequenzen hin nimmt das Auflösungsvermögen generell ab. Nach Winckel ist es im Bereich zwischen 0,8 und 3 kHz maximal, mit einem ebenmerklichen Frequenzunterschied von 3 Hz (vgl. Gelfand 2004, 293; Winckel 1952, 57; Yost 2008, 151). Daneben hängt die Tonhöhendiskrimination auch von der Lautstärke ab. Generell nimmt sie mit steigendem Schallpegel geringfügig zu (d.h. Modulations- und Unterschiedsschwellen sinken). Bei tiefen Frequenzen scheint der Einfluß des Pegels größer zu sein als bei hohen (vgl. Fricke & Louven 2008, 423f.; Gelfand 2004, 293, 300; Green 1993, 25; Moore 1993, 89; Winckel 1952, 66f.). Als Mittelwert ebenmerklicher Frequenzunterschiede geben Fasold und Veres (vgl. 2003, 51) 2% der Ausgangsfrequenz an. Die Minimalwerte liegen zum Teil deutlich darunter.

Unter 500 Hz entspricht die Modulationsschwelle einem frequenzunabhängigen *Frequenzhub* (Δf) von etwa 1,8 Hz (vgl. Hellbrück & Ellermeier 2004, 85f.; Reuter 1995, 33; 1996, 15f.; Zwicker & Feldtkeller 1967, 68, 92). Es gilt aber zu bedenken, daß der Frequenzhub die Differenz aus maximaler Modulationsfrequenz und (unmodulierter) Trägerfrequenz (f_1) angibt. Die maximale Modulationsfrequenz entspricht also $f_1 + \Delta f$, die minimale $f_1 - \Delta f$. Die hörbare Reizstufe beträgt demnach $2 \cdot \Delta f$, bei $\Delta f = 1,8$ Hz also 3,6 Hz (vgl. Hellbrück & Ellermeier 2004, 85f.; Zwicker & Feldtkeller 1967, 66). Über

500 Hz beträgt der minimale hörbare Frequenzhub 0,35% der Ausgangsfrequenz (vgl. Hellbrück & Ellermeier 2004, 86; Reuter 1995, 33; 1996, 16; Zwicker & Feldtkeller 1967, 68, 92). Die ebenmerkliche Reizstufe beträgt somit 0,7% der Ausgangsfrequenz.

Die Unterschiedsschwellen betragen für tiefe Frequenzen unter 500 Hz bei minimal 0,8 bis 1 Hz, oberhalb von 500 Hz etwa 0,2 bis 0,3% der Ausgangsfrequenz (vgl. Fasold & Veres 2003, 51; Fricke & Louven 2008, 424; Gelfand 2004, 294; Hellbrück & Ellermeier 2004, 86; Terhardt 1998, 351). Die Unterschiedsschwellen der Frequenzauflösung sind offensichtlich generell niedriger als die Modulationsschwellen. Für beide (Modulations- und Unterschiedsschwellen) bedeuten in der Frequenz konstante Schwellenwerte, daß die Trennschärfe unter 500 Hz mit sinkender Frequenz stetig schlechter wird (vgl. Stevens & Hallowell 1938, 88; Terhardt 1998, 353).

Die Auflösungsunschärfe bietet im Umkehrschluß natürlich eine gewisse Toleranzschwelle an. Die genaue Einhaltung der zahlenmäßig angegebenen Intervallbeziehungen ist nicht zwingend erforderlich. Im Gegenteil: Geringe Verstimmungen innerhalb des Klangaufbaus werden vom Gehör durchaus positiv bewertet. Vermutlich werden auch die geringen Unterschiede diverser Stimmungssysteme überbewertet. Läßt man ein Legato von einer zu tiefen kleinen Terz bis zu einer zu hohen großen Terz durchlaufen, ist das Ohr aufgrund der Frequenzunschärfe nicht in der Lage, eine klare Grenze zwischen kleiner und großer Terz zu ziehen (vgl. Winckel 1952, 60f., 85). Eher springt der Eindruck während des Legatoverlaufs irgendwann um. Dies geschieht je nach Verlaufsrichtung jedoch nicht zum selben Zeitpunkt, weil die Tonhöhenwahrnehmung nicht kontinuierlich, sondern kategorial funktioniert (vgl. Fricke & Louven 2008, 422). Es muß also erst die etablierte Kategorie sozusagen erwiesenermaßen widerlegt sein, bevor eine neue herangezogen wird.

Die Trennschärfe aufeinanderfolgender Signale unterscheidet sich von der Auflösung simultaner Klangkomplexe. Wie schon beschrieben sind hier die *Frequenzgruppen* maßgebend (siehe 1.3, S. 18). Während ein 1000 Hz-Ton von einem darauffolgenden 1003 Hz Ton sauber unterschieden werden kann (vgl. Fricke & Louven 2008, 424), erwecken entsprechende Simultanklänge Schwebungseindrücke, die auch bei erweitertem Frequenzabstand anhalten, bei 15 bis 20 Hz in Rauigkeit übergehen, aus der sich erst mit Abständen jenseits einer Frequenzgruppenbreite in saubere Tontrennung auflösen (vgl. Hellbrück & Ellermeier 2004, 125, 140).

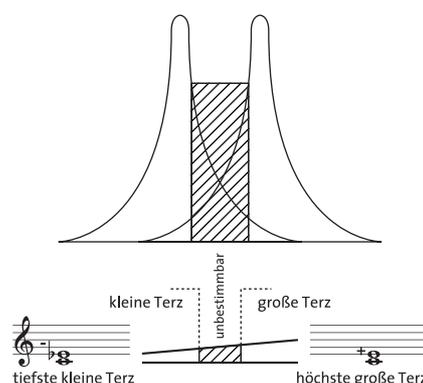


Abb. 1.5 **Intervallunschärfe** (Winckel 1952, Abb. 39, 61). Oben: Unschärfereich auf der Basilarmembran ($\pm 8\%$ der Sollfrequenz, d.h. Mitten- bzw. Resonanzfrequenz). Unten: Kontinuierliche Intervallspreizung von einer zu engen kleinen Terz zu einer zu weiten großen Terz. In einem Bereich dazwischen läßt sich die Terz auditiv nicht als groß oder klein definieren.

Einfluß der Lautstärke

Die Lautstärke beeinflusst nicht nur die Tonhöhendiskrimination, sondern auch die Tonhöhen selbst. Hohe Klänge werden bei größerer Lautstärke höher, tiefe Klänge tiefer wahrgenommen. Betroffen ist aber nicht der Reiz, er bleibt stabil, sondern allein die Empfindung (vgl. Fricke & Louven 2008, 424; Gelfand 2004, 376; Moore 1993, 94; Stevens & Hallowell 1938, 72; Terhardt 1998, 355; Winckel 1952, 66f.; 1960, 82f.).

Der mittelfrequente Bereich ist grundsätzlich eher lautstärkeunabhängig. Zwischen 1 und 2 kHz hat die Lautstärke mit Schwankungen unter 1% praktisch keinen Einfluß. Darüber hinaus kann man, nach Winckel, eine erweiterte Zone von etwa 0,8 bis 3 kHz ebenfalls als ziemlich klangkonstant ansehen (vgl. Fricke & Louven 2008, 424; Moore 1993, 94; Stevens & Hallowell 1938, 72; Terhardt 1998, 355; Winckel 1960, 84, 85). Allerdings streuen die Angaben der Grenzfrequenzen, ab denen lautstärkeabhängige Tonhöhenschwankungen auftreten, teils beträchtlich und stehen durchaus im Widerspruch zu expliziten Angaben der pegelunabhängigen Mittelspanne. Nach oben nimmt der Effekt ab etwa 3 bis 4 kHz merklich zu (vgl. Gelfand 2004, 376; Moore 1993, 94). Nach unten sind die Angaben weniger kohärent. Für Fricke und Louven ist der Lautstärkeeinfluß um 200 Hz nur gering, nach anderen Angaben nimmt er bereits unter 1 bis 2 kHz deutlich zu (vgl. Fricke & Louven 2008, 424; Gelfand 2004, 376; Moore 1993, 94; Terhardt 1998, 355). Töne sind im Allgemeinen ein wenig lautstärkeempfindlicher. Sie werden mit steigender Lautstärke tiefer, wenn ihre Frequenz unter 1,5 kHz liegt und höher, wenn sie über 2 kHz liegt (vgl. Terhardt 1998, 355). Je nach Pegel kann die empfundene Tonhöhe um bis zu 5 bis 10% von der physikalischen abweichen. Die Abweichung hoher Töne scheint mit bis zu 10% stärker ausgeprägt als jene tiefer Töne mit bis zu 6% (vgl. Fricke & Louven 2008, 424; Moore 1993, 94).

1.5.3 Tonhöhe von Klängen

Daß die geschilderten Phänomene der Tonhöhenveränderung in der musikalischen Praxis nicht so sehr auffallen, läßt sich auf die stabilisierende Wirkung der teiltonreichen Spektren der meisten natürlichen Klänge, bzw. damit verbundene ergänzende Tonhöhengewinnungsstrategien zurückführen (vgl. Hellbrück & Ellermeier 2008, 65; Winckel 1960, 84f.; Zwicker & Feldtkeller 1967, 80).

Tonhöhen von Klängen werden kaum aus der Frequenz des ersten Teiltens (Grundton) abgeleitet, sondern aus der Periodizität des Signals gewonnen. Zum einen wird dafür statt der Grundtonfrequenz die zeitliche Struktur des gesamten Spektrums, sozusagen die Hüllkurvenfrequenz analysiert (Periodizitätsanalyse). Jeder Teilton moduliert die Gesamtschwingung. Sind alle Teiltöne ganzzahlige vielfache derselben Frequenz (harmonisches Spektrum), ergibt sich daraus eine Modulation in eben dieser Frequenz (vgl. Fricke & Louven 2008, 424ff.; Gelfand 2004, 378; Moore 1993, 96; Yost 2008, 193f.). Zum anderen werden auch die (harmonischen) Teiltonverhältnisse selbst ausgewertet, indem die Grundfrequenz als größter gemeinsamer Teiler des Spektrums herausdividiert wird. Dementsprechend könnte man Grundtöne von Klängen als „Quotiententöne“ bezeichnen⁴ (vgl. Fricke & Louven 2008, 425f., 427; Gelfand 2004, 379; Meyer-Eppler et al. 1959, 73; Reuter 1996, 18; Stevens & Hallowell 1938, 99; Winckel spricht hingegen von spektraler Differenztonbildung; vgl. 1960, 85).

4 Dringt man konsequent zum Grund der Sache durch, muß man natürlich feststellen, daß eigentlich keine *neuen* Strategien an der Tonhöhenauswertung von Klängen beteiligt sind, sondern nur ein Mehr an Information. Im Grunde handelt es sich bei Sinustönen um extrem dünne Spektren mit nur einem einzigen Teilton, deren Hüllkurven der Sinusschwingung und deren Periodizität ihrer Frequenz entspricht.

Die Konsequenz daraus lautet: Klänge brauchen keinen physikalisch anwesenden Grundton, um auf ihrer Grundtonhöhe gehört zu werden. Die im (harmonischen) Spektrum enthaltene Periodizität genügt dem Gehör, um eine Art *virtueller Tonhöhe* (virtual pitch) zu erzeugen. Diese virtuelle Grundtonhöhe ist eigentlich eine schon sehr alte Entdeckung Seebecks (1843), die allerdings lange im Dornröschenschlaf lag, bevor sie schließlich von Schouten wiederentdeckt und im Jahr 1940 als *Residue* (engl. „Rest“, lat. „Residuum“, *Residualton*) bezeichnet wurde (vgl. Fricke & Louven 2008, 426; Gelfand 2004, 378; 2009, 96; Hellbrück & Ellermeier 2008, 66; Meyer-Eppler et al. 1959, 71; Moore 1993, 95f.; Reuter 1996, 18; Schouten 1940, 290; Seebeck 1843, 449ff.; 1844, 360f.; Stevens & Hallowell 1938, 99; Terhardt 1998, 347; Yost 2008, 193; für eine kurze Zusammenfassung der Forschungsgeschichte vgl. Reuter 1996, 21f.).

Residuen als virtuelle Töne sind von physikalisch präsenten Tönen (Grundtöne, Obertöne, Kombinationstöne) kategorisch zu unterscheiden, da sie natürlich zu keinen physikalischen Interaktionen mit ihrer akustischen Umwelt, wie beispielsweise Schwebung oder Verdeckung, fähig sind (vgl. Gelfand 2004, 378; Meyer-Eppler et al. 1959, 72; Moore 1993, 96; Reuter 1996, 18; Schouten 1938, 1090f.). Die Physik des Residualtons ist prinzipiell nichts anderes als die Physik seiner Teiltöne, an deren Verhalten und Schicksal er zwangsläufig gebunden ist (vgl. Reuter 1996, 20, 32). Gleichermassen sind Residualtöne kategorisch von physikalisch vorhandenen Grundtönen zu unterscheiden. Nicht zuletzt, da Residualtöne kein Wellenmaximum auf der Basilarmembran erzeugen können, die tonotope Grundtonhöhenbestimmung also ausfällt (vgl. Gelfand 2009, 96; Moore 1993, 96).

So gewöhnungsbedürftig das Prinzip der virtuellen Tonhöhe anfangs auch wirken mag: Residualtöne sind keinesfalls außergewöhnlich oder selten. Sie sind eher der Normalzustand. Ein alltägliches Beispiel bietet das schmale Übertragungsband des Telefons, dem die Grundtöne menschlicher Stimmen rigoros zum Opfer fallen. Übertragen wird lediglich ein spektraler Ausschnitt, den das Ohr nutzt, die Stimme vom anderen Ende der Leitung in ihrer normalen Tonhöhe (allerdings durch das beschnittene Spektrum in anderer Klangfarbe) zu hören (vgl. Gelfand 2004, 378; Hellbrück & Ellermeier 2008, 66; Moore 1993, 96; Schouten 1940, 287f.; Terhardt 1998, 315). Doch selbst wenn der Grundton physikalisch vorhanden ist, werden Tonhöhenindrücke von Klängen in erster Linie durch die Periodizität des Klangs generiert (vgl. Moore 1993, 96).

Da Klänge aufgrund ihres Spektrums über mehr Information verfügen, die das Ohr zur Tonhöhenbestimmung nutzt, ist es eigentlich nur konsequent, daß auch die Unterschiedsschwellen für Klangtonhöhendiskrimination niedriger sind als für Töne (vgl. Moore 1993, 96). Ändert sich die Periodizität des Teiltongemischs, ändert sich auch die Residualtonhöhe. Die Unterschiedsschwellen hängen also von jenen der Teiltöne ab. Die minimale Unterschiedsschwelle beträgt (zwischen 100 und 400 Hz) nur 0,2%. Zur Erinnerung: Die Unterschiedsschwelle eines Tons unter 500 Hz beträgt grob 1 Hz und damit z.B. bei 200 Hz 0,5% (vgl. Moore 1993, 96; Reuter 1996, 18; Terhardt 1998, 351, 353).

Neben der Periodizitätsanalyse soll noch die Ortsanalyse genannt sein. Die räumliche Spreizung des hörbaren Spektrums über die Länge der Basilarmembran als grundlegendes Prinzip der Frequenzauflösung ermöglicht es dem Gehör, allein durch den Erregungsort auf der Basilarmembran eine grobe Tonhöheninterpretation durchzuführen (vgl. Boer & Nuttall 2010, 145; Hellbrück & Ellermeier 2004, 102; Moore 1993, 88f.; Spreng in Keidel 1975, 133). Für periodische Signale dürfte die Ortsauswertung als solche eher zweitrangig sein (vgl. Moore 1993, 96). Vermutlich wird die Ortsinformation dann tragend, wenn

andere Auswertungsmuster versagen. Dies ist erstens ab ungefähr 5 kHz der Fall, wenn die Periodizitätsanalyse aufgrund von Systemeinschränkungen (siehe 3.3.4, S. 131) versagt und zweitens bei aperiodischen Signalen wie etwa Rauschen. Schmalbandiges Rauschen kann relativ klare Tonhöhenindrücke vermitteln, die allerdings mit zunehmender Bandbreite (und somit zwangsläufig diffuser werdender Ortsinformation) nachläßt und ab etwa einer Frequenzgruppenbreite verloren geht (vgl. Gelfand 2004, 377; Terhardt 1998, 324; Zwicker & Feldtkeller 1967, 80).

1.6 Verdeckung

Von *Verdeckung* (*Maskierung*) spricht man, wenn die Hörschwelle eines beliebigen Schallsignals durch ein zweites, unabhängiges Signal angehoben und ersteres dadurch in seiner Erkennbarkeit beeinträchtigt wird (vgl. Fricke & Louven 2008, 416; Gelfand 2004, 313; 2009, 98; Hellbrück & Ellermeier 2004, 135; Pollard & Jansson 1982, 250; Reuter 1996, 28; Spitzer 2002, 67; Stevens & Hallowell 1938, 127; Stumpf 1890, 220; Zwicker & Feldtkeller 1967, 55f.). Bei Schallen „regiert das Gesetz des Stärkeren“ (Fricke & Louven 2008, 417) und jeder muß sehen, wo er bleibt. Das lautere Signal bleibt in seiner Lautheit unbeeinflusst, während das andere ganz (Verdeckung) oder teilweise (*partielle Verdeckung*) unter seiner Mithörschwelle verschwindet (vgl. Gelfand 2004, 313; Hellbrück & Ellermeier 2004, 135; Reuter 1996, 28).

Der verdeckende Schall wird meistens *Maskierer* (Verdecker, *Störschall*) genannt, der verdeckte Schall hingegen *Nutz-* oder *Testschall*. Bei gleichzeitiger physikalischer Anwesenheit von Nutz- und Störschall ist von *Simultanverdeckung* (simultaneous masking) die Rede, treten Nutz- und Störschall in Folge auf, kommt es zu *nichtsimultaner* oder *zeitlicher Verdeckung*. Je nach dem, ob Nutz- oder Störschall zuerst erklingt, handelt es sich um *Vor-* oder *Nachverdeckung* (vgl. Fricke & Louven 2008, 416, 417; Gelfand 2004, 313, 331; Hellbrück & Ellermeier 2004, 135; Pollard & Jansson 1982, 251; Reuter 1996, 28; Viemeister & Plack 1993, 139; Zwicker & Feldtkeller 1967, 209, 211).

Voraussetzung, damit es zu Verdeckung kommen kann ist allerdings, daß beide Schalle sich spektral überschneiden oder zumindest nahe beieinander liegen (vgl. Spitzer 2002, 68; Zwicker & Feldtkeller 1967, 56). Welches Ausmaß die Verdeckung annimmt, hängt im weiteren von der störschallbedingten Mithörschwelle des Nutzschalls ab.

1.6.1 Mithörschwelle

Die Ruhehörschwelle wird durch Störschall künstlich angehoben. Diese durch die Anwesenheit eines Störgeräuschs zu höheren Pegeln verschobene Hörschwelle wird als *Mithörschwelle* (*Verdeckungsschwelle*⁵) bezeichnet und ist im Spektralbereich des Störschalls wirksam (vgl. Fricke & Louven 2008, 416; Reuter 1996, 28; Zwicker & Feldtkeller 1967, 51, 56). Wie die Ruhehörschwelle Auskunft darüber gibt, ab welchem Pegel ein Schallsignal in ruhiger Umgebung hörbar wird, gibt die Mithörschwelle an, welchen Schallpegel ein Nutz- oder Testsignal haben muß, um trotz der Anwesenheit des Störschalls hörbar zu werden/bleiben (vgl. Blauert 1974, 206; Zwicker & Feldtkeller 1967, 55f.). Dabei sollte man im Hinterkopf behalten, daß nur die Störschallenergie innerhalb der Frequenzgruppe(n) des Nutzschalls

5 Die Begriffe *Mithörschwelle* und *Verdeckungsschwelle* bezeichnen dieselbe Schwelle, einmal bezogen auf den Nutz- und einmal auf den Störschall bezogen. *Mithörschwelle*: Hörschwelle des verdeckten Schalls (Nutzschall) in Anwesenheit eines Verdeckers (Störschall); *Verdeckungsschwelle*: Hörschwelle, die den durch Störschall belegten Hörbereich begrenzt.

verdeckungsrelevant ist (siehe 1.3, S. 21; vgl. Guski 1996, 244; Hellbrück & Ellermeier 2004, 126; Zwicker & Feldtkeller 1967, 73). Liegen alle spektralen Komponenten des Nutzsignals innerhalb einer einzigen Frequenzgruppe, werden die Intensitäten aller Teiltöne zugunsten der Hörbarkeit des Signals summiert. Verteilt sich das Spektrum aber über mehrere Frequenzgruppen, wird die Hörbarkeit des Gesamtsignals in der energiereichsten Frequenzgruppe ermittelt (vgl. Green 1993, 30; Moore 1993, 77f.). Der Nutzschall muß den Frequenzgruppenpegel des Störschalls allerdings nicht übertreffen, um hörbar zu sein. Die Mithörschwellen liegen immer relativ konstant knapp 4 dB darunter (vgl. Reuter 1996, 29; Zwicker & Feldtkeller 1967, 60, 73).

Die auffälligen interindividuellen Variationen im Detailverlauf natürlicher Hörschwellen (Ruhehörschwellen) sind im Verlauf der Mithörschwellen nicht zu finden. Dies ist erstens darin begründet, daß Nutz- und Störsignal auf den gleichen Abschnitt der Ruhehörschwelle (dieselbe Ohrempfindlichkeit) treffen und es daher ausschließlich auf ihr gegenseitiges Lautheitsverhältnis ankommt. Zweitens sind auch Frequenzgruppen und Verdeckungsmaß (Differenz von Frequenzgruppenpegel des Störsignals und Mithörschwelle des Nutzschalls) keinen nennenswerten interindividuellen Schwankungen unterworfen (vgl. Fastl & Zwicker 2006, 63; Zwicker & Feldtkeller 1967, 56, 57). Jeder Störschall belegt einen seinem Spektrum gemäßen Bereich des Hörfelds, woraufhin die Mithörschwellen aller anderen Schalle über diese besetzte Fläche verlagert werden. Der Mithörschwellenverlauf hängt demnach von der Form des Störsignals ab: Jeder Schall formt eine ihm eigene, charakteristische Mithörschwelle (vgl. Zwicker & Feldtkeller 1967, 56).

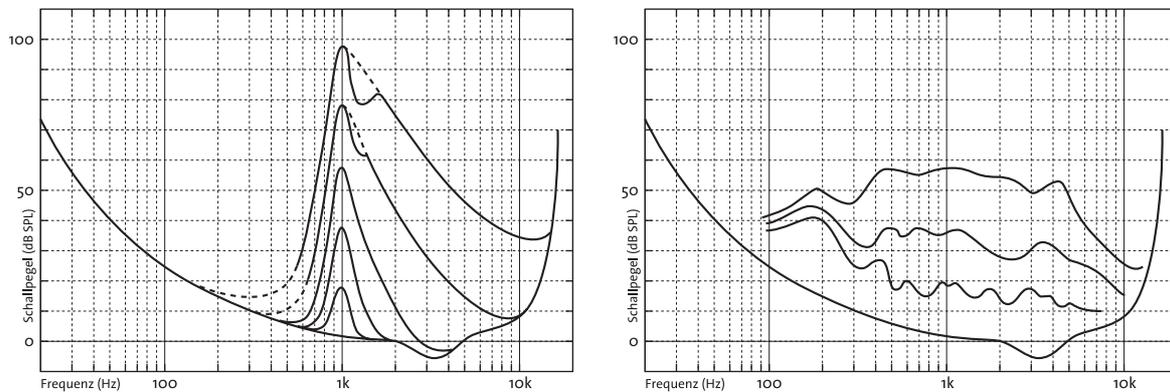


Abb. 1.6 **Mithörschwellen** (Zwicker & Feldtkeller 1967, Abb. 20.1 & 23.3, 60, 65). Rechts: Mithörschwellen bei schmalbandigem Verdeckter. Mit steigendem Verdeckterpegel wird die hochfrequente Flanke zunehmend flacher. Bei hohen Pegeln bildet sich eine zweite Verdeckungsspitze oberhalb der Maskierfrequenz und der tieffrequente Übergang in die Ruhehörschwelle wird sehr flach. Links: Verdeckungskurven von Geigenklängen verschieden stark intoniert. Mit stärkerer Dynamik wachsen die höheren Teiltöne stärker an, weshalb die Kurven zunehmend flacher werden.

Praktisch vollkommen flache Mithörschwellen erhält man, wenn man die Schallintensitätsdichte von weißem Rauschen über 500 dB um 3 dB/Oktave abschwächt. Das so generierte Rauschen wurde von Zwicker und Feldtkeller *gleichmäßig verdeckendes Rauschen* genannt (vgl. Zwicker & Feldtkeller 1967, 58f.); nicht zu verwechseln mit *rosa Rauschen* (wie rotes Licht im hohen Spektralbereich schwächer), dessen Leistungsdichte ebenfalls mit 3 dB/Oktave abfällt, allerdings nicht erst ab 500 Hz (vgl. Fastl &

Zwicker 2006, 62; Weinzierl 2008, 12). Solche flachen Mithörschwellen können für Versuchszwecke nützlich sein, befinden sich aber fernab jeglicher akustischer Realität. Natürliche Spektren sind mehr oder weniger breitbandig und äußerst individuell zerklüftet. Die Mithörschwellen passen sich den spektralen Amplitudenverläufen i.d.R. recht gut an (vgl. Fastl & Zwicker 2006, 71f.; Reuter 1996, 31; Zwicker & Feldtkeller 1967, 56, 64f.).

Ist der Verdeckter ein schmalbandiges Signal, liegt das Maximum der Mithörschwelle bei der Frequenz des Verdeckers ungefähr 4 dB unterhalb seines Pegels (vgl. Zwicker & Feldtkeller 1967, 60). Bei schwachen Pegeln fallen die Flanken zu beiden Seiten relativ symmetrisch ab (vgl. Gelfand 2004, 315; 2009, 100; Reuter 1996, 28). Dies ändert sich mit wachsendem Pegel. Die tieffrequente Flanke verläuft mit 80 bis 240 dB/Oktave bei Sinustönen und mit etwas flacheren 55 bis 190 dB/Oktave bei schmalbandigem Rauschen pegelunabhängig immer recht steil (vgl. Gelfand 2004, 315; 2009, 100; Moore 1993, 70; Reuter 1996, 28; vgl. Zwicker & Feldtkeller 1967, 60). Die hochfrequente Flanke wird mit wachsendem Pegel immer flacher. Wächst der Pegel des Verdeckers um 20 dB, hebt sich die Mithörschwelle in Verdeckernähe ebenfalls um 20 dB, bei höheren Frequenzen jedoch um 30 dB. Höhere, über dem Maskierer liegende Frequenzen, werden in ihrer Lautheit also viel stärker beeinträchtigt als tiefere Frequenzen (vgl. Fasold & Veres 2003, 51; Gelfand 2009, 100; Moore 1993, 70f.; Reuter 1996, 29; Zwicker & Feldtkeller 1967, 60). Bei extrem starken Pegeln zeigen die Mithörschwellen zwei Besonderheiten. Erstens bildet sich ein zweites Maximum, daß etwas höher als die eigentliche Störschallfrequenz liegt. Liegt das Störsignal bei 1 kHz, bildet sich das Nebenmaximum bei zirka 1,7 Hz (vgl. Zwicker & Feldtkeller 1967, 60). Zweitens wird die tieffrequente Flanke beim Übergang in die Ruhehörschwelle extrem flach und verläuft ein kleines Stück verhältnismäßig weit über der Ruhehörschwelle (vgl. Zwicker & Feldtkeller 1967, 60). Die grundsätzliche Form der Mithörschwelle ist frequenzabhängig. Verschiebt man die Mittenfrequenz des Verdeckers von tief nach hoch durch das Spektrum, wird die anfangs noch recht breite Mithörschwelle zunehmend schmalbandiger, was m.E. vor allem am immer enger werdenden Raum für die hochfrequente Flanke liegt (vgl. Gelfand 2004, 315; Zwicker & Feldtkeller 1967, 61).

Werden Töne durch ihresgleichen verdeckt, gestaltet sich die Mithörschwelle wesentlich komplizierter, da es zu Rauigkeit und Schwebung zwischen verdecktem Ton und Verdeckter kommen kann, wenn ihre Frequenzen zu nahe beieinander liegen. In solchen Fällen ist oft unklar, ob der verdeckte Ton tatsächlich selbst gehört wurde, oder ob nur die Schwebung mit dem Verdeckter erkannt wurde (vgl. Fastl & Zwicker 2006, 67f.; Gelfand 2004, 316, 317; Moore 1993, 70; Zwicker & Feldtkeller 1967, 63f.).

Die Verdeckungsschwellen natürlicher periodischer Signale, also von Klängen, erstreckt sich über den gesamten Spektralbereich und folgen dabei – wie schon erwähnt – recht präzise den spektralen Hüllkurven der Verdeckter, weshalb man sie m.E. als quasi-Summation der schmalbandigen Verdeckungsschwellen einzelner Teiltöne verstehen kann (vgl. Reuter 1996, 31; Zwicker & Feldtkeller 1967, 64f.). Da sie den Teiltonamplituden folgen, verschieben sie sich, zumindest bei Instrumentenklängen, zwangsläufig entsprechend der Dynamikregeln für Spektren (der *Schumann'schen Klangfarbengesetze*; vgl. dazu Mertens 1975, 31f.; Oehler 2008, 80–84; Reuter 1995, 76f.; 1996, 110f.; Schumann 1929, 15–18, 98ff.). So wird die Mithörschwelle bei steigenden Pegeln beispielsweise flacher, da die höheren Teiltöne stärker zulegen als die tieferen (siehe Abb. 1.6; vgl. Zwicker & Feldtkeller 1967, 65). Angemerkt sei noch, daß es natürlich auch zwischen verdeckenden Teiltönen und verdeckten (Teil-)Tönen zu Schwebungen kommen kann (vgl. Zwicker & Feldtkeller 1967, 65).

Die Verdeckungsschwellen von Hoch- bzw. Tiefpaßrauschen verlaufen im Bereich des Rauschens flach, der Leistungsdichte entsprechend, und fallen an der Grenzfrequenz in gleicher Weise wie die entsprechende schmalbandige Mithörschwellenflanke ab (vgl. Zwicker & Feldtkeller 1967, 62).

1.6.2 Simultanverdeckung

Wie sich an den Mithörschwellen deutlich ablesen läßt, ist ein lauter tiefer Ton eher in der Lage höhere Töne zu verdecken als umgekehrt. Hohe Töne können zwar prinzipiell auch tiefe verdecken, was aber angesichts der steil abfallenden Mithörschwelle nur in sehr begrenztem Umfang möglich ist (vgl. Fricke & Louven 2008, 416; Gelfand 2004, 314; 2009, 100; Hellbrück & Ellermeier 2004, 129; Reuter 1996, 29; Spitzer 2002, 67; Stumpf 1890, 227f.; Zwicker & Feldtkeller 1967, 66). Dieses Phänomen wird auch *Asymmetrie der Maskierung* (engl. auch: *upward spread of masking*) genannt (vgl. Gelfand 2009, 100; Hellbrück & Ellermeier 2004, 129). Diese asymmetrische Verschleppung der Mithörschwellen zu hohen Frequenzen dürfte darin begründet sein, daß die Wanderwellen auf der Basilarmembran von den hohen zu den tiefen Frequenzorten laufen (siehe 3.3.2, S. 123; vgl. Gelfand 2004, 316; 2009, 100; Spitzer 2002, 67). Versuche, die *charakteristische Frequenz*⁶ einzelner Nervenfasern des Hörnervs durch frequenznahe Stimuli zu verdecken, ergaben Maskierungskurven, die den Erregungsschwellen der Nervenfasern (*Abstimmungskurven*, *frequency tuning curves*, FTC) verblüffend ähnlich sehen (vgl. Gelfand 2009, 102f.; Hellbrück & Ellermeier 2004, 129f.; Moore 1993, 61). Um dieser Analogie Rechnung zu tragen, werden die Verdeckungskurven *psychophysische Abstimmungskurven* (*psychophysical tuning curves*, PTC) genannt (vgl. Gelfand 2009, 103; Hellbrück & Ellermeier 2004, 129; Moore 1993, 60f.). Die psychophysischen Abstimmungskurven zeigen eine extrem steile hochfrequente Flanke, jedoch einen flacheren tieffrequenten Flankenverlauf. Auch hier zeigt sich also, daß tiefere Frequenzen nur recht geringe Intensitäten benötigen, um verdeckungsrelevant zu werden, während höhere Frequenzen praktisch kaum verdeckungsfähig sind (vgl. Hellbrück & Ellermeier 2004, 130).

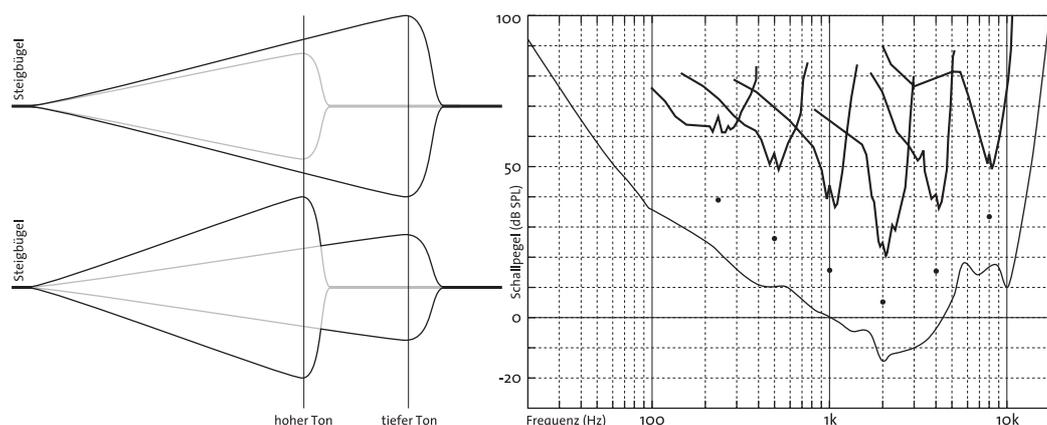


Abb. 1.7 Links: **Einfluß des Wanderwellenverlaufs auf die Verdeckung** (modifiziert n. Gelfand 2004, Abb. 10.3, 316). Leise hohe Töne werden durch laute tiefe Töne verdeckt (oben), umgekehrt ist dies nicht möglich (unten). Rechts: **Psychophysische Abstimmungskurven** (Hellbrück & Ellermeier 2004, Abb. 4.4, 130). Jeder Punkt symbolisiert eine charakteristische Frequenz, die jeweilige Kurve darüber ihre Mithörschwelle.

6 Als *charakteristische Frequenz* (Bestfrequenz, BF; characteristic frequency, critical bzw. center frequency, CF) bezeichnet man jene Frequenz, für die eine bestimmte Faser des Hörnervs am empfindlichsten ist (vgl. Gelfand 2004, 141; Guski 1996, 93; Hellbrück & Ellermeier 2004, 104; Yost 2008, 100). Es handelt sich also um die effektive Mittenfrequenz (ungeachtet der asymmetrischen Flanken) des Erregungsbandes einer Faser.

Zwei wichtige psychoakustische Phänomene sind *Drosselung* und *Aufholtendenz*. Sie hängen stark mit dem Verlauf der Kurven gleicher Lautheit oberhalb der Mithörschwelle zusammen. Der durch den Störschall belegte Bereich wird aus dem Hörfeld nicht einfach „herausgeschnitten“. Vielmehr wird es – und mit ihm die strukturgebenden Lautheitskurven – darüber gestaucht. Es bleibt also der komplette Dynamikbereich erhalten, nur verläuft er wegen der Stauchung über eine viel kürzere Schalldruckpegelspanne (vgl. Fricke & Louven 2008, 417; Reuter 1996, 28f.; Zwicker & Feldtkeller 1967, 56). Signale, die nur knapp über die Mithörschwelle ragen, werden als sehr leise wahrgenommen. Sie wirken deshalb so leise, weil der Großteil ihres Pegels (dB_{SPL}) unterschwellig, und nur ein kleiner Teil davon überschwellig (dB_{SL}), also hörbar, ist. Das Signal wird in seiner Lautheit *gedrosselt* (vgl. Fricke & Louven 2008, 417; Reuter 1996, 29). Aufgrund des gestauchten Hörbereichs liegt es ja auch bei den niedrigsten Phonwerten. Erhöht man den Pegel des gedrosselten Signals um nur wenige Dezibel, erklingt es sehr bald in fast ungestörter Lautheit. Diese Lautheitsrückgewinnung geht ebenfalls der gestauchten Lautheitskurven wegen so schnell. Mitunter liegen bis zu 20 Phon innerhalb von nur 2 dB_{SPL} . Die Lautheiten partiell verdeckter Signale haben also die Tendenz, den Vorsprung ihrer Schalldruckpegel schnell *aufzuholen* (vgl. Fricke & Louven 2008, 417; Reuter 1996, 29). Der Lautheitsverlust aufgrund der Teilverdeckung wird *Drosselung* genannt, die Fähigkeit diesen Rückstand unglaublich schnell aufzuholen wird als *Aufholtendenz* (recruitment) bezeichnet (vgl. Fricke & Louven 2008, 417; Reuter 1996, 29).

Die „Aufholjagd“ beginnt schon, bevor der Schall überhaupt hörbar sein dürfte, nämlich rund 4 dB unterhalb des effektiven Störschallpegels. Nur weitere 6 dB darüber ist die Anwesenheit des Störschalls bereits bedeutungslos. In Zahlen: Wenn ein Verdecker mit 60 dB_{SPL} bei 1 kHz wirkt, ist der Nutzschall bereits mit einem Pegel von 56 dB_{SPL} erkennbar, obwohl er bei -4 dB_{SL} liegt. Schon bei einem Pegel von 66 dB_{SPL} (6 dB_{SL}) ist er in quasi voller Lautheit, d.h. mit 60 Phon zu hören. In nur 10 dB (von -4 dB_{SL} bis 6 dB_{SL}) wurden (von 0 bis 60 Phon) so viele Lautheitsstufen bewältigt, wie normalerweise zwischen 0 und 60 dB_{SPL} (vgl. Reuter 1996, 29; Zwicker & Feldtkeller 1967, 60).

Trotz der eindrucksvollen Aufholtendenz sind wir niemals in der Lage, alles von dem zu hören, was gleichzeitig auf unsere Ohren einwirkt. Vieles davon wäre aber auch im Sinne des Höreindrucks nur ineffektiver Ballast. Die Verdeckung wirkt dem sehr effizient entgegen, da verworfen wird, was ohnehin nicht genutzt werden kann. Nichts anderes als dieses Prinzip der rigorosen Informationsreduktion liegt den Komprimierungsverfahren des sogenannten mp3-Formats zugrunde (vgl. Fricke & Louven 2008, 417; Spitzer 2002, 68).

1.6.3 Zeitliche Verdeckung

Wie zu Beginn schon angedeutet, kann es auch zu Verdeckung kommen, wenn Stör- und Nutzsignal nicht gleichzeitig, sondern nacheinander erklingen. Dies wird *zeitliche Verdeckung* (temporal masking) oder *nichtsimultane Verdeckung* (nonsimultaneous masking) genannt (vgl. Gelfand 2004, 331; 2009, 103; Viemeister & Plack 1993, 139). Die genauen physiologischen Mechanismen der zeitlichen Verdeckungen sind noch nicht geklärt. Sicher ist aber, daß der Verdeckungseffekt umso größer ist, je kürzer die Zeitabstände zwischen und je ähnlicher die Frequenzen und/oder Spektren von Nutz- und Störschall sind (vgl. Gelfand 2004, 332, 334).

Das Phänomen der auch nach Ende des Störschalls wirksamen Verdeckungsschwelle wird *Nachverdeckung* (post-masking) oder *Vorwärtsverdeckung* (forward masking) genannt (vgl. Fricke & Louven 2008, 416; Gelfand 2004, 331; 2009, 103; Hellbrück & Ellermeier 2004, 135; Viemeister & Plack 1993, 139). Entsprechend wird die nach Ende des Störschalls noch beeinflusste (erhöhte) Hörschwelle *Nachhörschwelle* genannt (vgl. Zwicker & Feldtkeller 1967, 205). Sie bleibt noch etwa 2 bis 5 ms auf dem Niveau der Mithörschwelle und sinkt dann in 100 bis 300 ms – der dritten Integrationszeit entsprechend – zur Ruhehörschwelle ab (vgl. Fricke & Louven 2008, 416f.; Gelfand 2004, 333; Reuter 1996, 31; Zwicker & Feldtkeller 1967, 205ff.). Wie der Rückgang zur Ruhehörschwelle im Detail verläuft und wie lange er dauert, hängt von den beteiligten Signalen ab (vgl. Reuter 1996, 31).

Eine mögliche Erklärung der Nachverdeckung bieten die neuronalen Erregungsmuster. Zu Beginn eines Stimulus feuern die Neuronen erst heftig, lassen dann nach (siehe 1.2.2, S. 17) und stellen ihre Aktivität mit Stimulusende kurzzeitig vollkommen ein, bevor die normale Spontanaktivität wieder aufgenommen wird. Möglicherweise ist die kurze Erholungspause der Grund für die Nachverdeckung (vgl. Viemeister & Plack 1993, 140). Die Hemmungsphase dauert etwa 150 ms, was für Viemeister und Plack (vgl. 1993, 140) gut in einer Nachverdeckungsspanne von 200 ms aufgeht. Darauf, wie gut die Erholungsphasen mit den verdeckterabhängig variablen Nachverdeckungszeiten bis 300 ms korrespondieren, wurde nicht eingegangen. Hellbrück & Ellermeier (vgl. 2004, 135) geben aber an, daß die ausgelöste Erregung noch bis zu maximal 500 ms nachwirken kann, was in jedem Fall ein ausreichender Zeitrahmen wäre.

Neben der Nachverdeckung gibt es auch das „paradox anmutende Phänomen“ (Hellbrück & Ellermeier 2004, 135) der *Vorverdeckung* (pre-masking) oder *Rückwärtsverdeckung* (backward masking), bei dem ein Reiz durch einen späteren, stärkeren Reiz verdeckt wird (vgl. Fricke & Louven 2008, 417; Gelfand 2004, 331; 2009, 103; Hellbrück & Ellermeier 2004, 135; Viemeister & Plack 1993, 139; Zwicker & Feldtkeller 1967, 209). Es fällt auf, daß die Begriffe der zeitlichen Verdeckung mitunter mißverständlich sein können. Während Vor- und Nachverdeckung vom Verdeckter, vom Störschall ausgehen (der Nutzschall war vorher: Vorverdeckung und umgekehrt), orientieren sich Vorwärts- und Rückwärtsmaskierung an der Zeitachse. Vorwärts in der Zeit zu gehen, bedeutet in die Zukunft zu gehen: der Störschall verdeckt „die Zukunft“, d.h. er geht dem Nutzschall voraus.

Wie kann ein Stimulus durch einen erst nachfolgend auftretenden Störschall verdeckt werden? Ein Erklärungsansatz ist – vereinfacht gesagt – daß das Gehör bei der Verarbeitung Prioritäten setzt. Natürlich kann ein noch nicht vorhandenes Signal kein schon vergangenes Signal verdecken. Aber das Gehör braucht eine gewisse Zeit, bis ein Reiz zu einer Empfindung führt. Dabei ist durchaus denkbar, daß lautere Schalle höhere Priorität genießen und schneller verarbeitet werden als leisere, wodurch der zweite den ersten unterwegs einholt (vgl. Hellbrück & Ellermeier 2004, 135; Zwicker & Feldtkeller 1967, 209). D.h. ein nachfolgendes starkes Signal kann ein unmittelbar vorausgehendes schwaches, unter der Mithörschwelle liegendes Signal in seiner Wirksamkeit unterdrücken, indem seine Hörschwelle beeinflusst wird (vgl. Fricke & Louven 2008, 417; Zwicker & Feldtkeller 1967, 208).

Die Ruhehörschwelle wird schon 10 bis 20 ms vor Beginn des Verdeckers angehoben. Der Schwellenverlauf richtet sich nach der spektralen Ähnlichkeit von Nutz- und Störsignal. Sind beide Signale gleich, wird die Ruhehörschwelle völlig glatt zur Mithörschwelle angehoben. Sind die Spektren aber stark unterschiedlich, hebt sich die Hörschwelle erst weit über die Mithörschwelle hinweg an, erreicht

kurz *nach* dem Einsatz des Störschalls ihr Maximum und sinkt dann innerhalb von rund 25 ms auf die Mithörschwelle ab (vgl. Gelfand 2004, 333; Reuter 1996, 31; Zwicker & Feldtkeller 1967, 209). Der kurze Buckel in der Vorhörschwelle manifestiert vermutlich die Tatsache, daß die Frequenzgruppen umorientiert und auf die neue spektrale Situation angepaßt werden. An dieser Stelle lohnt auch die Feststellung, daß die Dauer der Vorhörschwelle der ersten Integrationszeit entspricht (vgl. Reuter 1996, 31; Zwicker & Feldtkeller 1967, 210).

Es soll noch das Phänomen genannt werden, daß, bei zu kurzen Signalpausen unter 200 ms, Vor- und Nachverdeckungsschwelle mehr oder weniger direkt ineinander übergehen können (d.h. ohne auf die Ruhehörschwelle abzusinken). Wie der Zeitrahmen zeigt, sinkt die Nachverdeckungsschwelle zwar etwas schneller als normalerweise (vgl. Reuter 1996, 31), ihr Einfluß ist aber trotzdem größer als jener der Vorverdeckungsschwelle. Diese neue entstandene „Zwischenhörschwelle“ liegt allerdings höher als durch die theoretische Addition von Vor- und Nachhörschwelle vorausgesagt (vgl. Reuter 1996, 31, 34; Viemeister & Plack 1993, 141; Zwicker & Feldtkeller 1967, 210).

1.7 Lokalisation

Anhand der akustischen Merkmale des von einer Schallquelle abgesonderten Signals wird das dadurch entstandene Hörereignis im (virtuellen) Hörraum positioniert. Dieser Vorgang wird als *Lokalisation* bezeichnet. *Lokalisation* stellt also den Bezug zwischen der faktischen Schallquellenposition und der empfundenen Hörereignisposition her (vgl. Blauert 1974, 30; Yost 2008, 173).

Um einen guten Richtungs- und Räumlichkeitseindruck zu bekommen, stehen dem Gehör verschiedene Möglichkeiten zur Verfügung. Zum einen kann es die filternden Einflüsse von Körper, Kopf und Außenohr auf das Spektrum verwerten, zum anderen kann es aus der Kopfposition relativ zur Schallquelle Informationen gewinnen, die sich in unterschiedlichen Schallsignalen an beiden Ohren (unterschiedliche Ohrsignale, interaurale Differenzen) äußern (vgl. Blauert 1974, 30; Guski 1996, 274; Hellbrück & Ellermeier 2004, 151; Yost 2008, 173).

In natürlichen Hörsituationen wirken alle Merkmale zusammen, so daß man ziemlich sicher davon ausgehen kann, daß der Lokalisationsprozeß auch von allen Gebrauch macht. Die interauralen Differenzen bieten vor allem horizontale Richtungsinformationen, das gefilterte Spektrum vertikale Richtungsinformationen (vgl. Blauert 1974, 132; 1997, 165; Rosowski 2010, 63f.). Reflexionen des Raumes, die – je nach wegbedingter Laufzeitverzögerung – früher (Nachhall) oder später (Echo) nach dem Direktschall eintreffen, legen die informelle Basis des Räumlichkeitseindrucks (vgl. Fasold & Veres 2003, 56; Fricke & Louven 2008, 433; Hellbrück & Ellermeier 2004, 134). Das Mischverhältnis von Direktschall und Reflexionen gibt (zusammen mit dem Lautheitseindruck) zudem Aufschluß über die Entfernung von der Schallquelle (vgl. Culling & Akeroyd 2010, 130; Gelfand 2004, 410; Hellbrück & Ellermeier 2004, 145; Spitzer 2002, 64; Yost 2008, 179).

Durch den Einfluß des Kopfs fallen die Signale an beiden Ohren praktisch immer unterschiedlich aus, es sei denn, ihre Quelle liegt genau auf der Medianebene (vgl. Blauert 1974, 56; Gelfand 2004, 85; Yost 2008, 71). Diese Unterschiede werden als *Interaurale Differenzen* bezeichnet. Genaugenommen wird

zwischen *interauralen Zeit-* und *Pegeldifferenzen* unterschieden (vgl. Blauert 1974, 113ff., 126ff., 132; 1997, 138; Culling & Akeroyd 2010, 123; Wightman & Kistler 1993, 159, 164; Yost 2008, 173f.).

Der Vorgang, der beide Ohrsignale zu einer gemeinsamen Wahrnehmung vereint, wird *binaurale Fusion* genannt (vgl. Gelfand 2004, 391) und ist nicht mit der *binauralen Summation*, einem Phänomen der Lautheit zu verwechseln (siehe 1.1.1, S. 7; vgl. Hellbrück & Ellermeier 2004, 146). Räumlichkeit und Richtungseindruck werden so mitgestaltet. Signalanteile (Reflexionen), die innerhalb einer gewissen Zeit nach dem ersten Direktschall eintreffen, fallen unter das *Gesetz der ersten Wellenfront*, demzufolge alle Energie dem ersten Signal samt Richtungseindruck subsumiert wird (vgl. Blauert 1974, 163, 177, 184ff.; vgl. Culling & Akeroyd 2010, 135f.; Fasold & Veres 2003, 56; Gelfand 2004, 412f.; Guski 1996, 285; Hellbrück & Ellermeier 2004, 152; Yost 2008, 180). Innerhalb dieser Zeit bilden Reflexionen den Nachhall (vgl. Fricke & Louven 2008, 433; Hellbrück & Ellermeier 2004, 134). Wird die *Echoschwelle* (siehe 1.2.1, S. 15) überschritten, werden alle eintreffenden Signale samt eigener Einfallsrichtung identifiziert (vgl. Blauert 1974, 163; Fasold & Veres 2003, 53; Gelfand 2004, 413; Guski 1996, 284f.). Wahrscheinlich bieten Reflexionen den wichtigsten Hinweis, um die Schallquellenentfernung in einen Distanzeindruck zu übersetzen. Der Direktschallpegel nimmt mit zunehmender Schallquellenentfernung logarithmisch ab. Der Raumhallpegel bleibt jedoch relativ gleich, so daß das Mischverhältnis über die Distanz zur Schallquelle Aufschluß geben kann (vgl. Culling & Akeroyd 2010, 130; Fasold & Veres 2003, 65, 116; Gelfand 2004, 410, 411; Hellbrück & Ellermeier 2004, 145; Spitzer 2002, 64; Yost 2008, 179).

1.7.1 Orientierung im Hörraum

Um sich leichter zurechtzufinden, ist der Hörraum in drei im Verhältnis zum Kopf fixierten Raumebenen geteilt, deren Bezeichnungen sich auf die Körperebenen beziehen (siehe *Exkurs: Zur Orientierung am Körper*, S. 50). Die *Horizontalebene* befindet sich auf Höhe der *Frankfurter Horizontalen*, die *Mediane Ebene* liegt längs in der Schädelmitte und die *Frontalebene* steht senkrecht zu beiden anderen (vgl. Blauert 1974, 11f.; Gelfand 2009, 47; Schünke et al. 2009, 134 b; Yost 2008, 173).

Auf den Ebenen wird die Position des Schallereignisses (tatsächliches Signal) bzw. des Hörereignisses (äquivalente Empfindung) in Winkeln angegeben. Der horizontale Einfallswinkel wird dabei (in Anlehnung an die Geographie) als *Azimut*, der vertikale als *Erhebungswinkel* (elevation) bezeichnet (vgl. Gelfand 2004, 85; Yost 2008, 173). Der Nullpunkt (0°) beider Ebenen ist konvergent und liegt auf der Sagittalachse vorne. Der Azimut wird im Uhrzeigersinn angegeben, so daß 90° rechts, 180° hinten und 270° links liegen. Mitunter wird auch ein hemisphärisches System verwendet, in dem links -90° bzw. „ 90° links“ wären. Der Erhebungswinkel wird nach oben abgetragen, so daß 90° oben liegen. Nach hinten nimmt der Erhebungswinkel wieder ab, so daß er bei 180° Azimut auf der Sagittalachse wieder 0° beträgt (vgl. Blauert 1974, 11, Abb. 4, 11; Gelfand 2004, 85; 2009, 47).

Die seitliche Auslenkung des Hörereignisses wird als *Lateralisation* bezeichnet. Dabei wird der Azimut (wie der Panoramaregler an der Stereoanlage) eindimensional betrachtet, indem man ihn auf die Horizontalachse (bzw. Ohrenachse) projiziert (vgl. Blauert 1974, 112; Gelfand 2004, 396; Yost 2008, 181). Sie tritt auf, wenn die Ohrsignale keine spektralen Hinweise auf die Einfallsrichtung enthalten. Dies ist

bei identischen Ohrsignalen, meistens durch Kopfhörerbeschallung verursacht, der Fall. Das Hörereignis ist dann auf der Achse zwischen den Ohren, also im Kopf lokalisiert (vgl. Blauert 1974, 97, 107, 110; 1997, 133; Gelfand 2004, 204, 396; Hellbrück & Ellermeier 2004, 150; Spitzer 2002, 66; Wightman & Kistler 1993, 184; Yost 2008, 181, 183).

Lokalisationsunschärfe

Die kleinstmögliche Verrückung des Hörereignisses im Hörraum, hervorgerufen durch richtungsspezifische Merkmale des Schallereignisses, wird als *Lokalisationsunschärfe* bezeichnet (vgl. Blauert 1974, 30). Dieses räumliche Auflösungsvermögen ist nicht überall im Hörraum gleich. Da die Ohren in der Waagerechten angeordnet sind, ist die horizontale Ebene zweiohrig und damit wesentlich besser aufgelöst als die vertikale (vgl. Blauert 1974, 35; Culling & Akeroyd 2010, 127; Yost 2008, 178). Auch wenn man den Hörraum zur „Hörfläche“ macht, indem man ihn auf die Horizontalebene reduziert, ist die Auflösung nicht homogen. Im vorderen Bereich des Hörfelds ist die Lokalisationsunschärfe geringer als im hinteren Bereich, am größten ist sie jedoch seitlich des Kopfes, jeweils direkt vor den Ohren (vgl. Blauert 1974, 31f.; Culling & Akeroyd 2010, 127; Gelfand 2004, 409; Guski 1996, 273; Spitzer 2002, 71; Stevens & Hallowell 1938, 176; Yost 2008, 176, 182).

Nach einer Zusammenstellung von Blauert pendeln die signalabhängigen Werte direkt vorne (0° Azimut und Erhebung) zwischen $0,75(!)$ und $11,8^\circ$. Die minimale vordere Lokalisationsunschärfe (minimaler hörbarer Winkelunterschied, *minimal audible angle*, MAA) wird aber meist als zwischen 1 bis 3° schwankend angegeben, wobei 1° als der absolut beste erreichbare Wert gilt (vgl. Blauert 1974, Tab. 2, 31, 31f.; Culling & Akeroyd 2010, 127; Gelfand 2004, 408f.; 2009, 105; Guski 1996, 273; Hellbrück & Ellermeier 2004, 150; Keidel 1975, 383; Spitzer 2002, 71). Dementsprechend kann das Ohr Zeitdifferenzen (diesenfals aus unterschiedlichen Laufzeiten herrührend) in der Größenordnung von ungefähr 10 bis $20 \mu\text{s}$ wahrnehmen (vgl. Hellbrück & Ellermeier 2004, 132, 150; Spreng in Keidel 1975, 137;

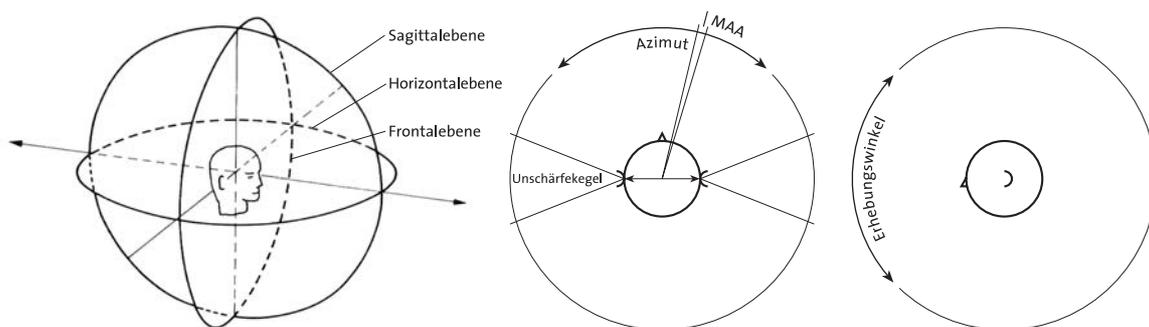


Abb. 1.8 Hörraum und Lokalisationsunschärfe (modifiziert n. Blauert 1974, Abb. 4, 11; Yost 2008, Abb. 12.1, 174). Links: Einteilung des Hörraums basierend auf den Körperebenen. Mitte: Horizontaler Lokalisationswinkel (Azimut). Kleinster hörbarer Winkelunterschied (MAA). Der Unschärfekegel um die Ohrenachse symbolisiert den Bereich der geringsten Auflösung. Zwischen beiden Ohren (auf der Ohrenachse) ist die Lateralisationsachse angedeutet. Rechts: Die vertikale Raumauflösung wird als Erhebungswinkel bezeichnet.

Yost 2008, 183). Seitlich (Azimut 90 bzw. 270°) beträgt die Lokalisationsunschärfe das drei- bis zehnfache des vorwärtsbezogenen Wertes, die Angaben schwanken zwischen 6 und 20° (vgl. Blauert 1974, 32; Guski 1996, 273; Spitzer 2002, 71). Seitlich, direkt vor den Ohren, liegt ein sehr diffuser, praktisch unaufgelöster, als „Unschärfekegel“ (*cone of confusion*) bezeichneter Bereich (vgl. Gelfand 2004, 409; Stevens & Hallowell 1938, 176; Wightman & Kistler 1993, 164). Hinten ist der Hörraum etwa halb so gut wie vorne aufgelöst (vgl. Blauert 1974, 32). Wegen der bezüglich des Erhebungswinkels praktisch identischen Ohrsignale ist die vertikale Auflösung generell schlechter als die horizontale. Die Werte schwanken signalabhängig zwischen 4 und 17° (vgl. Blauert 1974, 35; vgl. Culling & Akeroyd 2010, 127; Guski 1996, 273f.).

Die räumliche Auflösung des Gehörs wird nicht nur von den richtungsabhängigen Merkmalen des Schalls beeinflusst, sondern auch von einigen anderen Eigenschaften wie Dauer, Pegel und Spektrum (vgl. Blauert 1974, 35, 125; Gelfand 2004, 402f.; Wightman & Kistler 1993, 181f.).

Ein wichtiger Faktor ist die Breite des Spektrums. Generell läßt sich sagen, daß breitbandige Signale räumlich besser aufgelöst sind als schmalbandige. Die Richtungsinformationen schmalbandiger Signale sind begrenzt und mißverständlich, während breitbandige Signale mehr verwertbare Information bieten und so die Fehlerquote, speziell bei axialen Verwechslungen, niedrig halten (vgl. Blauert 1974, 35; Gelfand 2004, 403; Wightman & Kistler 1993, 181f.; Yost 2008, 174f.). Nicht nur die Frage wieviele, sondern auch welche Bänder ist von Bedeutung. Am genauesten funktioniert die räumliche Auflösung unter 1 und über 4 kHz. Letzteres gilt besonders für vorne/hinten Axialfehler. Dazwischen, speziell zwischen 2 und 4 kHz ist die Fehlerquote bei Versuchspersonen am höchsten (vgl. Gelfand 2004, 402f.; Yost 2008, 174f.).

Änderungen der Lokalisationsunschärfe wurden auch einhergehend mit Pegeländerungen beobachtet. Steigt der Pegel von niedrigen zu mittleren Intensitäten, nimmt die Lokalisationsunschärfe ab um im weiteren konstant zu bleiben bzw. zu extremen Pegeln hin wieder leicht zuzunehmen. Ähnliches läßt sich im Zusammenhang mit der Signaldauer feststellen. Mit wachsender Dauer nimmt die Unschärfe ab (vgl. Blauert 1974, 125, 130).

Des weiteren ist die Lokalisationsunschärfe durch die Signalbandbreite bedingt. Je breitbandiger das Signal, desto besser ist seine Raumauflösung. Schmale Bänder führen in Ermangelung spektraler Richtungsinformationen zu Lokalisations-, speziell Axialfehlern. Breite Spektren bieten mehr verwertbare Information, weshalb sie generell nur mit geringen Fehlerquoten einhergehen (vgl. Blauert 1974, 35; Gelfand 2004, 403; Wightman & Kistler 1993, 181f.).

Obwohl auch die Annahme existiert, Kopfbewegungen wären für das Richtungshören eher bedeutungslos (vgl. Wightman & Kistler 1993, 172), muß man intuitiv vermuten, daß Peilbewegungen des Kopfs die räumliche Auflösung verbessern. Immerhin wird durch Ausrichtung des Kopfs das Schallsignal in eine optimale, meist mehr oder weniger vorwärtsgerichtete Position gebracht und so die Lokalisationsunschärfe minimiert (vgl. Blauert 1974, 144, 146; 1997, 179; Stevens & Hallowell 1938, 180; Yost 2008, 176).

Bei ausreichender Peilbewegung entsteht das Hörereignis früher oder später immer in Richtung der Schallquelle. Peilbewegungen verbessern also die räumliche Auflösung und sind überdies geeignet, Axialfehler zu vermeiden (vgl. Blauert 1974, 35, 143ff.; 1997, 178ff.; Gelfand 2004, 404; Guski 1996, 275;

Wightman & Kistler 1993, 171; Yost 2008, 176). Dies geschieht wohl über einen Informationszuwachs. Dreht man den Kopf zur (angenommenen) Schallquellenposition, wird der Schall vermutlich zweimal richtungsspezifisch ausgewertet. Einmal beim unmittelbaren Eintreffen in Einfallrichtung und noch einmal nach Ausrichtung des Kopfs in (dann) quasi Vorwärtsrichtung (vgl. Blauert 1974, 148; Guski 1996, 278; Wightman & Kistler 1993, 171f.). Detailliertere Informationen zu Kopfbewegungen beim räumlichen Hören (sogenannte *Drehtheorien*) finden sich bei Blauert (vgl. 1974, 143–154; 1997, 178–190).

1.7.2 Distanzhören

Darüber, wie das Gehör die Entfernung von Schallquellen tatsächlich ermittelt, ist nur wenig bekannt. Fakt ist jedoch, daß die Entfernungsbestimmung wesentlich schlechter funktioniert, als die sonstige Lokalisation (vgl. Yost 2008, 179). Im Gegensatz zu anderen Lokalisationsparametern steht der Distanzerkennung kein primärer Mechanismus zur Verfügung, so daß auf „Hilfsinformationen“ zurückgegriffen werden muß (vgl. Keidel 1975, 390).

Ist die Entfernung des Hörereignisses kleiner als der Kopfradius, also das Hörereignis scheinbar im Kopf, spricht man von *Im-Kopf-Lokalisiertheit* oder *intrakranialer Lokalisation* (vgl. Blauert 1974, 95; 1997, 117). Sie tritt vor allem bei identischen Ohrsignalen und/oder Kopfhörerbeschallung auf, weil das Spektrum nicht durch kopfbezogene Übertragungsfunktionen gefiltert und so mit den üblichen Richtungsinformationen versehen wurde. Jedes Hörereignis ist so auf der Achse zwischen den Ohren lokalisiert (vgl. Blauert 1974, 97, 107, 110; 1997, 133; Gelfand 2004, 204, 396; 2009, 104; Hellbrück & Ellermeier 2004, 150; Spitzer 2002, 66; Wightman & Kistler 1993, 184; Yost 2008, 181, 183). Einerseits, um die Eindimensionalität zu beschreiben und andererseits, um von *extrakranialer Schallortung* (Lokalisation) zu unterscheiden, wird die Ortung interkranialer Hörereignisse auf der Ohrenachse als *Lateralisation* bezeichnet (vgl. Blauert 1974, 112; Gelfand 2004, 396; Yost 2008, 181). Um die Hörereignisse trotz Kopfhörerdarbietung außerhalb des Kopfs zu platzieren, müssen sie künstlich durch kopfbezogene Übertragungsfunktionen gefiltert werden. Durch diesen, binaurale Synthese genannten Vorgang, wird der Kopfhörerschall mit jenen richtungsabhängigen spektralen Feinheiten angereichert, die Freifeldschall aufweist (vgl. Hellbrück & Ellermeier 2004, 152; Yost 2008, 183). Ist das Hörereignis nun extrakranial, d.h. außerhalb des Kopfs lokalisiert (vgl. Gelfand 2004, 396), weist es zwangsläufig eine gewisse Distanz zu diesem auf.

Aus der Signallautstärke lassen sich Hinweise auf die Schallquellenentfernung ziehen. Intuitiv würde man wohl sagen: Laute Signale wirken näher als leise. Der Schallpegel halbiert sich näherungsweise mit jeder Entfernungsverdoppelung, sinkt also jeweils um 6 dB (vgl. Blauert 1974, 96; 1997, 118; Fasold & Veres 2003, 34; Gelfand 2004, 410; Hellbrück & Ellermeier 2004, 144; Yost 2008, 179; Keidel 1975, 390). Um allerdings auch dem Hörereignis eine Distanzverdoppelung mitzuteilen, müßte der Schallpegel um bis zu 20 dB reduziert werden (vgl. Blauert 1974, 99; 1997, 122; Gelfand 2004, 411). Die Hörereignisentfernung nimmt weniger stark zu als die Schallquellenentfernung. Es gibt also offenbar einen Trend des Zurückbleibens, wenn die Entfernung alleine aus der Lautstärke eingeschätzt werden soll. Für Sprache gilt dies allerdings nicht. Nur bei Flüsterlauten läßt sich ein ähnlicher Trend beobachten.

Schreien hingegen wird tendenziell eher weiter entfernt gehört, als die Schallquelle tatsächlich verortet ist. (vgl. Blauert 1974, 99f.; 1997, 100, 122f.; Guski 1996, 279).

Dies könnte ein Hinweis darauf sein, daß Erfahrungswerte beim Abschätzen von Entfernungen bedeutsam sind. Weiß man, wie eine Schallquelle klingt (klingen soll), lassen sich Parameter wie Lautstärke und Spektrum (Klangfarbe) besser einordnen. Kurzum: Bekanntes wird sicherer und richtiger „distanziert“ als Unbekanntes (vgl. Blauert 1974, 37; Culling & Akeroyd 2010, 130; Gelfand 2004, 411; Spitzer 2002, 63; Yost 2008, 179). Da das Gehör an Sprachschalle in aller Regel außerordentlich gut gewöhnt ist, überrascht nicht, daß Extreme wie Flüstern und Schreien den Entfernungseindruck dehnen bzw. stauchen. Die intime Sprachsituation des Flüsterns suggeriert räumliche Nähe, während die distanzierte Sprachsituation, verbunden mit der dynamischen Aufhellung des Klangspektrums von Geschrei durchaus größere Räumlichkeit vermittelt.

Klangspektren sind tatsächlich distanzabhängig, weil Luft ab Entfernungen von 15 m nicht mehr als lineares Medium angesehen werden kann. Allerdings werden hohe Frequenzen stärker gedämpft als tiefe⁷, da die kurzen Wellenlängen viel eher „auf der Strecke bleiben“ (vgl. Blauert 1974, 96, 102; 1997, 118; Gelfand 2004, 411; Guski 1996, 276f.; Hellbrück & Ellermeier 2004, 144; Spitzer 2002, 64; Wightman & Kistler 1993, 172). Ein gutes Beispiel ist Donnerrollen, das aus weiter Ferne nur noch als dumpfes Grollen, in nächster Nähe jedoch als relativ helles Krachen gehört wird. Das Beispiel verdeutlicht allerdings auch, daß die Luft als Medium Klangspektren eher über sehr große Distanzen beeinflusst. Zwar kann der Einfluß auf Frequenzen über 10 kHz bereits ab 15 m deutlich hörbar sein, wovon „normale“ Hörsituationen aber nicht betroffen sein dürften (vgl. Blauert 1974, 103; 1997 126, 127; Hellbrück & Ellermeier 2004, 144; Keidel 1975, 390; Spitzer 2002, 64). Entscheidender ist womöglich, daß Spektren durch den Kopf distanzabhängig verzerrt werden. Das Gehör wertet besonders distanzabhängige Verzerrungen aus Entfernungen unter 3 m, weil sich noch keine ebene Wellenfront bilden konnte. Allerdings sind die Verzerrungen bereits ab Entfernungen von 25 cm relativ gering (vgl. Blauert 1974, 96, 107; 1997, 118, 131).

Die meisten Informationen bezüglich der Schallquellenentfernung gewinnt das Gehör wahrscheinlich aus dem Verhältnis von Direktschall zu Reflexionen. Analog zum Schallpegel nimmt der Direktschallgehalt indirekt proportional zum Quadrat der Entfernung ab, während die Reflexionen (Raumhallpegel) halbwegs konstant bleiben. Vereinfacht gesagt: Das Verhältnis aus Direktschall und Reflexionen verändert sich mit wachsender Schallquellenentfernung zugunsten der Reflexionen (vgl. Culling & Akeroyd 2010, 130; Fasold & Veres 2003, 65, 116; Gelfand 2004, 410, 411; Guski 1996, 276; Hellbrück & Ellermeier 2004, 145; Spitzer 2002, 64; Yost 2008, 179).

1.7.3 Binaurales Hören

Zweiohrigkeit (Binauralität) bildet die Grundlage des räumlichen Hörens (vgl. Blauert 1974, 56, 59, 77; Gelfand 2004, 85; 2009, 49; Hellbrück & Ellermeier 2004, 44; Yost 2008, 71; Young 2007, 431), und somit auch die Grundlage für generelle, auf Räumlichkeit basierende Mustererkennungsprozesse der auditiven Szenenanalyse (vgl. Bregman 1990, 73–83, 297; Hellbrück & Ellermeier 2004, 152; Reuter 1996, 51ff.). Mit nur einem Ohr verschlechtert sich die Raumauflösung merklich, wobei das Gehör grundsätzlich auch monauraler Lokalisation fähig ist. Mit zwei Ohren profitiert die Lokalisations-

7 Es gilt zwischen Lautstärke und Dynamik zu unterscheiden: Der U-förmige Verlauf der Hörschwelle lehrt uns, daß breite Spektren mit steigendem Pegel besonders im tieffrequenzen Bereich dazugewinnen – also tiefenlastiger werden (vgl. Blauert 1974, 97f.; 1997, 120). Ein Blick auf die *Klangfarbengesetze* macht aber deutlich, daß sich spektrale Energie mit stärkerer Dynamik zu den hohen Teiltönen verlagert (vgl. dazu Mertens 1975, 31f.; Oehler 2008, 80–84; Reuter 1995, 76f.; 1996, 110f.; Schumann 1929, 15–18, 98ff.).

fähigkeit jedoch (bei entsprechendem Schalleinfallswinkel) nachvollziehbarerweise vom Informationsgewinn aus *binauralen Differenzen* (vgl. Hellbrück & Ellermeier 2004, 149).

Studien lassen vermuten, daß sowohl die Intensitäts- als auch die Frequenzauflösung binaural besser sind als monoaural (vgl. Gelfand 2004, 391). Zumindest bei geringen Lautheiten ist die binaurale Frequenzunterschiedsschwelle niedriger (vgl. Gelfand 2004, 391). Ob die Auflösung sich aber tatsächlich verbessert, oder nur von der summationsbedingt besseren Lautheit profitiert, ist unklar (vgl. Gelfand 2004, 391), denn binaurales Hören beeinflusst auch die Lautheit: Verschließt man ein Ohr, nimmt die Lautheit ein wenig ab. Entsprechend ist auch die binaurale Hörschwelle etwas niedriger (vgl. Gelfand 2004, 389; Hellbrück & Ellermeier 2004, 146; 2008, 61). Im Frequenzbereich von etwa 1,5 bis 2 kHz ist die (binaurale) Raumauflösung am schlechtesten, was sich auch in der Uneindeutigkeit von Richtungsinformationen in diesem Frequenzbereich ausdrückt (vgl. Gelfand 2004, 409; Spitzer 2002, 71).

Binaurale Fusion

Unter *binauraler Fusion* versteht man die Tatsache, daß *zwei* Ohren *eine* akustische Szenerie hören, die Ohrsignale also zu einem einzigen Höreindruck verschmelzen. Die Bildung eines gemeinsamen Höreignisorts (Phantomschallquelle, Stereoeffekt) aus den unterschiedlichen Ohrsignalen wird als *Summenlokalisierung* bezeichnet (vgl. Blauert 1974, 163ff., 177f.; Gelfand 2004, 391f.; 2009, 104; Guski 1996, 284). Gänzlich ungleiche Ohrsignale werden zwar nicht fusioniert, Ähnlichkeiten reichen aber bereits aus. Naheliegende Frequenzen, dichotisch dargeboten, fusionieren beispielsweise zu Schwebungseindrücken. Vermutlich nutzt der Fusionsprozeß die spektrale Struktur der Ohrsignale, um Ähnlichkeiten zu ermitteln (vgl. Blauert 1974, 173; Gelfand 2004, 392f.; 2009, 104).

Binaurale Fusion ist ein zeitabhängiger Vorgang. Verzögerungen unter 5 ms führen zu *Summenlokalisierung*. Bei längeren Verzögerungen, bis etwa 50 ms, greift das sogenannte *Gesetz der ersten Wellenfront*. Noch längere Verzögerungen führen zu perzeptiven Trennungen in Primärschall und Echo. Darüber hinaus verzögerte Schallanteile werden als vollkommen eigenständige Signale empfunden (vgl. Blauert 1974, 163ff., 188; Fasold & Veres 2003, 56; Gelfand 2004, 413; Guski 1996, 283f.).

Eine kurze Zeit, etwa 1 bis maximal 5 ms nach Signalbeginn, tragen alle Schallanteile zur Bildung des Hörereignisorts bei, weshalb man von *Summenlokalisierung* spricht. Befinden sich zwei Schallquellen im gleichen Abstand vom Hörer und beschallen diesen diotisch, entsteht das Hörereignis in der Mitte zwischen beiden Schallquellen. Wird eines der beiden Quellsignale verzögert oder in seiner Lautstärke verändert, verschiebt sich der Hörereignisort entsprechend (vgl. Blauert 1974, 163ff., 177f.; Fasold & Veres 2003, 56; Gelfand 2004, 391f.; Guski 1996, 284).

Werden die Verzögerungen größer, fallen alle Schallanteile, die dem Erstschall in weniger als 25 bis 50 ms folgen, unter das *Gesetz der ersten Wellenfront* (law of first wavefront) oder *Präzedenzeffekt* (precedence effect, Haas Effekt). Dies besagt, daß nur das Erstsinal (eben die erste Wellenfront) bezüglich seiner Richtungsinformation verarbeitet wird. Alle folgenden Signalanteile werden dem Erstsinal subsumiert, selbst wenn sie bis zu 10 dB höhere Schallpegel aufweisen. Anders gesagt: Es entsteht ein einziges Hörereignis, dessen Hörereignisort durch die Einfallrichtung des ersten Signalanteils bestimmt

wird (vgl. Blauert 1974, 163, 177, 184ff.; Culling & Akeroyd 2010, 135f.; Fasold & Veres 2003, 56; Gelfand 2004, 412f.; 2009, 106; Guski 1996, 285; Hellbrück & Ellermeier 2004, 152; Keidel 1975, 388; Yost 2008, 180). Laborversuche zeigten, daß der Fusionseffekt kurzzeitig aufgehoben werden kann. Dazu läßt man die beiden aufeinanderfolgenden Signalanteile zeitlich ihre Plätze tauschen (das erste wird zum zweiten Signal). Kurze Zeit werden beide getrennt wahrgenommen, bevor sich der Fusionsprozeß auf die neue Situation eingestellt hat und die Arbeit wieder aufnimmt (vgl. Gelfand 2004, 415).

Perzeptive Integration findet nur in einem begrenzten Bereich statt, geht jedoch mit wachsender Verzögerung zwischen den Signalanteilen mehr oder weniger schrittweise in vollkommene Signaltrennung über. Die Fusionsgrenze wird von Guski (vgl. 1996, 283) bei 40 ms angegeben, wobei Gelfand (vgl. 2004, 413) zufolge bereits ab 35 ms uneigenständige Zweitsignale (sozusagen „Delay“ Effekte) wahrnehmbar sind, deren Richtungseindruck allerdings unter dem Einfluß des Erstsignals steht. Die Schwelle, ab der tatsächliche Signaltrennung auftritt, wird *Echoschwelle* genannt (vgl. Blauert 1974, 179; Gelfand 2004, 413). Der Zeitpunkt der Echoschwelle variiert signalabhängig stark, weshalb sehr unterschiedliche Zeiten, meist zwischen 2 und 50 ms, angeführt werden. Für Klicks werden mit 2 bis 10 ms im Schnitt recht frühe Werte angegeben, während die Angaben für komplexere Signale wie Sprache und Musik in eine Spanne von 20 bis 100 ms fallen (vgl. Blauert 1974, 184; Culling & Akeroyd 2010, 135; Fasold & Veres 2003, 52f.; Gelfand 2004, 413; Guski 1996, 283f.; Hellbrück & Ellermeier 2004, 152; Keidel 1975, 388f.). Abgesehen von der grundsätzlichen Natur des Schallsignals können auch einzelne Parameter den Zeitpunkt der Signaltrennung beeinflussen. Weichere Wellenfronten erlauben z.B. weitere Zeitabstände, während unterschiedliche Einfallsrichtungen die erlaubten Zeitabstände verkürzen (vgl. Fricke & Louven 2008, 434). Jenseits der Echoschwelle werden vollkommen eigenständige Hörereignisse mit eigenem Richtungseindruck gebildet, die mitunter sogar als Störschall (z.B. Echo) empfunden werden (vgl. Blauert 1974, 163; Gelfand 2004, 414; Guski 1996, 284f.).

Interaurale Differenzen

Der Kopf ist ein beeinflussendes Hindernis für freie Schallausbreitung, wodurch es in den meisten Hörsituationen zunächst zu unterschiedlichen Ohrsignalen kommt (vgl. Blauert 1974, 56; Gelfand 2004, 85; Yost 2008, 71). Behindert durch den Kopf erreicht der Schall eines der beiden Ohren schwächer und, aufgrund des längeren Weges um den Kopf herum, auch später (vgl. Blauert 1974, 59; 1997, 72; Culling & Akeroyd 2010, 123; Keidel 1975, 382f.; Yost 2008, 174). Dementsprechend wird zwischen *interauren Zeit-* bzw. *Pegeldifferenzen* unterschieden (vgl. Blauert 1974, 113ff., 126ff., 132; 1997, 138; Culling & Akeroyd 2010, 123; Wightman & Kistler 1993, 159, 164; Yost 2008, 173f.). Die sogenannte *Duplextheorie* postuliert die Annahme, daß Laufzeitunterschiede bei tieferen Frequenzen und Pegelunterschiede bei höheren die Basis des räumlichen Hörens bilden (vgl. Culling & Akeroyd 2010, 123; Gelfand 2004, 396; 2009, 105; Spitzer 2002, 70f.; Wightman & Kistler 1993, 168, 182; Yost 2008, 176, 183). Ihre Stärken liegen in der horizontalen Raumauflösung; für vertikale Lokalisation und vorne/hinten Unterscheidungen sind sie praktisch ohne Bedeutung (vgl. Blauert 1974, 131; Culling & Akeroyd 2010, 127). Interaurale Differenzen sind frequenzabhängige Erscheinungen. Um 1,5 kHz herum scheinen sie nur unklare Richtungseindrücke zu vermitteln, wodurch die Lokalisation fehleran-

fälliger wird. Zeitdifferenzen wirken besonders unterhalb dieses Frequenzbereichs, Pegeldifferenzen darüber (vgl. Blauert 1974, 126, 134; Gelfand 2004, 398, 402, 407; Wightman & Kistler 1993, 164).

Liegt die Schallquelle nicht in der Medianebene, ist sie näher zu einem der beiden Ohren. Dorthin ist der Schall kürzer unterwegs als zum anderen. Diese Laufzeitunterschiede werden *interaurale Zeitdifferenzen* (IZD) genannt (vgl. Blauert 1974, 114f.; 1997, 141ff.; Gelfand 2004, 396ff.; 2009, 49; Spitzer 2002, 69; Wightman & Kistler 1993, 159). Auch bei Laufzeitunterschieden kann zwischen Phasenlaufzeiten (Phasendifferenzen von Sinustönen/Teiltönen) und Gruppenlaufzeiten (Hüllkurvenunterschiede) differenziert werden. Erstere sind unter 1,6 kHz von primärer Bedeutung, darüber nur noch letztere (vgl. Blauert 1974, 122f.; 1997, 165). Demnach sind sie über einen größeren Spektralbereich wirksam. Traditionellerweise geht die Psychoakustik aber davon aus, daß Laufzeitunterschiede überwiegend im tieffrequenten Bereich bis grob 1,5 kHz bedeutsam sind (vgl. Gelfand 2004, 398, 402, 407; 2009, 105; Hellbrück & Ellermeier 2004, 149; Wightman & Kistler 1993, 169, 181f.).

Möglicherweise geht ihre Wirkung im höherfrequenten Bereich unter, weil über 1,5 bis 1,6 kHz *interaurale Pegelunterschiede* (IPD) den Richtungseindruck wesentlich stärker beeinflussen. Im Gegenzug sind Pegelunterschiede unter 1,5 bis 1,6 kHz praktisch bedeutungslos (vgl. Blauert 1974, 132, 134; 1997, 165; Culling & Akeroyd 2010, 126; Wightman & Kistler 1993, 169; Spitzer 2002, 65; Yost 2008, 72). Sie entstehen durch den Schallschatten des Kopfs, der besonders hohe Frequenzen beeinträchtigt (vgl. Culling & Akeroyd 2010, 123, 126; Gelfand 2004, 397f.; 2009, 48; Rosowski, 2010, 62; Spitzer 2002, 65; Yost 2008, 72).

Blauert diskutiert Versuchsergebnisse, nach denen Spektren ohne Anteile über 1,6 kHz alleine durch IZD geortet werden, Spektren mit Anteilen über 1,6 kHz (unabhängig von eventuell vorhandenen Anteilen darunter) durch IZD und IPD geortet werden (vgl. Blauert 1974, 136ff.; 1997, 170ff.). Die Frage, inwiefern IPD isoliert wirken, wurde beim Komponieren der Stimuli offenbar vernachlässigt (vgl. Blauert 1974, 137).

Binaural Masking Level Difference

Der Umstand, daß Nutz- und Störsignale besser unterschieden werden, wenn sie räumlich (bzw. binaural) getrennt sind, wurde *Binaural masking level difference*, BMLD oder *masking level difference* MLD benannt (vgl. Gelfand 2004, 418; 2009, 104; Hellbrück & Ellermeier 2004, 153; Lehnhardt & Hesse in Lehnhardt & Laszig 2009, 192). Genau genommen wird die Differenz aus monotonischer Mithörschwelle und jeweils untersuchter Vorspielbedingung, in Dezibel angegeben, als BMLD bezeichnet. BMLD hängen demnach besonders mit räumlichem Hören, Verdeckung und Streamingphänomenen wie dem Cocktailpartyeffekt zusammen (vgl. Blauert 1974, 206; Gelfand 2004, 417; 2009, 104f.; Hellbrück & Ellermeier 2004, 155f.; Yost 2008, 184f.).

Die Mithörschwelle (binaurale Verdeckungsschwelle) sinkt, wenn Nutz- und Störsignal beiden Ohren auf unterschiedliche Weise präsentiert werden. D.h. bei monotonischer⁸ Nutz- aber diotonischer Störbeschallung und umgekehrt sowie dichotonischer Beschallung, ist die Mithörschwelle niedriger als bei diotonischer oder monotonischer Nutz- und Störbeschallung (vgl. Hellbrück & Ellermeier 2004, 154f.; Lehnhardt & Hesse 2009, 192; Yost 2008, 184f.). Die Mithörschwelle ist in dichotonischer Darbietung niedriger als in monotonischer oder diotonischer Darbietung. Am niedrigsten ist sie jedoch, wenn Nutz- oder

8 *monotonisch*: nur ein Ohr beschallt; *diotonisch*: beide Ohren gleich beschallt; *dichotonisch*: beide Ohren unterschiedlich beschallt (vgl. Blauert 1974, 77, Fußnote; Yost 2008, 184).

Störsignal dichotisch zueinander phasenverdrehet präsentiert werden (vgl. Gelfand 2004, 418; Yost 2008, 184f., 186). Mit anderen Worten: Das Gehör profitiert davon, beide Ohrsignale miteinander zu vergleichen und die „Mechanismen des binauralen Hörens einzusetzen“ (Hellbrück & Ellermeier 2004, 155).

Nach Hellbrück und Ellermeier (vgl. 2004, 156) gibt es im wesentlichen zwei Erklärungsmodelle: erstens das *Webster-Jeffress-Modell* (Webster 1951; Jeffress 1948), nach dem Nutz- und Störsignal bei unterschiedlicher interauraler Phasenbeziehung als unterschiedlich lateralisiert behandelt werden können; und zweitens das *Equalization-and-cancellation-Modell* (Kock 1950; Durlach & Colburn 1978), das davon ausgeht, daß die Ohrsignale so transformiert werden, daß die Störsignale binaural dichotisch sind (equalization). Wenn dann beide Ohrsignale voneinander abgezogen werden, bleibt nur das Nutzsignal übrig (cancellation). Letzteres ist allerdings eher ein Computermodell, das die Physiologie kaum einbezieht (vgl. Hellbrück & Ellermeier 2004, 156).

Anatomie

2 Wie ist das Ohr aufgebaut?

Es entspricht meiner Überzeugung und Erfahrung, daß detaillierte Kenntnisse und gutes *Verständnis* der Ohranatomie fundamentale Voraussetzungen sind, um die Physiologie des Gehörs – soweit dies überhaupt möglich ist – richtig zu verstehen. Zu diesem Zweck soll nun die „psychoakustische Blackbox“ geöffnet und der strukturelle Aufbau des Gehörs beleuchtet werden. Allerdings ist der Begriff *Gehör* im anatomischen Sinne problematisch. Er ist funktionell geprägt und erfaßt „das Hören“ als einen der fünf menschlichen Sinne, an dessen physiologischer Umsetzung verschiedene anatomische Strukturen – unterteilt in das *periphere auditive System* (das Ohr inklusive Hörnerv) und das *zentrale auditive System* (der mit dem Hören befaßte Teil des zentralen Nervensystems) – beteiligt sind.

Daß der Mensch über zwei Ohren verfügt, mag banal klingen, tatsächlich steckt darin aber bereits das erste anatomische Detail. Zweiohrigkeit (Binauralität) befähigt uns Menschen, analog zum binokularen Sehen, räumlich zu hören (vgl. Blauert 1974, 56, 59, 77; Gelfand 2004, 85; Yost 2008, 71; Young 2007, 431). Ist vom „Ohr“ die Rede, bezieht man sich zumeist nur auf den von außen sichtbaren Teil, also die Ohrmuschel (vgl. Spitzer 2002, 55), die jedoch nur „die Spitze des Eisbergs“ und damit einen kleinen Teil des Ohrs darstellt (vgl. Gelfand 2004, 34). Beide Muscheln sind seitlich relativ starr und nach vorne offen am – von knapp 13,5 cm (bei Frauen) bzw. 15 cm bis rund 18,5 cm (bei Männern) breiten – Kopf befestigt (vgl. Blauert 1997, 301; Guski 1996, 88; Spitzer 2002, 55; Terhardt 1998, 53). Auf Einzelheiten der generellen *Kopfanatomie* wird in weiterer Folge nur noch bei Bedarf eingegangen (für äußerst detaillierte Darstellungen der gesamten Anatomie des Kopfs vgl. Fanghänel et al. 2003(a); Schünke et al. 2009).

Das Ohr wird in drei Abschnitte gegliedert: *Außenohr* (*Ohrmuschel* und *Gehörgang*), *Mittelohr* (*Trommelfell*, „Mittelohrhöhle“ (für eine Begriffserklärung siehe 2.3.2, S. 59) und *Gehörknöchelchen*) und *Innenohr* (vgl. Gelfand 2004, 37; 2009, 34; Reiss 2003, 591; Schünke et al. 2009, 126ff.; Spitzer 2002, 55; Terhardt 1998, 53). Das Innenohr stellt einen anatomischen Sonderfall dar: Als strukturelle Einheit zerfällt es physiologisch mit dem *Hörorgan* (*Schnecke*) und dem *Gleichgewichtsorgan* (*Vorhof*, *Bogengänge*) in zwei funktionell unabhängige Systeme (vgl. Gelfand 2004, 34; Reiss 2003, 591; Schünke et al. 2009, 134; Terhardt 1998, 56).

Außen- und Mittelohr werden als *Schalleitungsapparat* bezeichnet, da sie der Schnecke, dem als *Schallempfindungs-* oder *Schallaufnahmeapparat* bezeichneten eigentlichen Sinnesorgan, als Zulieferer dienen (vgl. Reiss 2003, 591; Schünke et al. 2009, 126; Spitzer 2002, 58; Terhardt 1998, 56). Genau genommen ist letztlich das mit den Sinneszellen (Haarzellen) bestückte *Cortiorgan* im Innenohr mit der Transformation der Schallreize beauftragt, die in der Schnecke über das Cortiorgan verteilt werden. Dies veranlaßte Ranke, eine feiner abgestimmte Nomenklatur vorzuschlagen, wonach Außen- und Mittelohr als „Reiztransportorgan“, die Schnecke, auf die spektrale Dispersion entlang der Basilarmembran bezogen (siehe 2.4.3, S. 77; 3.3.2, S. 123), als „Reizverteilungsorgan“ und nur der tatsächliche Reizwandler als „Reiztransformationsorgan“ bezeichnet würde (vgl. Ranke 1953, 33ff.).



Abb. 2.1 **das Ohr** (modifiziert n. Schünke et al. 2009, 126 b, 130 A). Links: Frontalschnitt, Ansicht von vorne. Rechts: Ansicht von oben. Das Außenohr ist jeweils gelb eingefärbt, das Mittelohr türkis und das Innenohr grün.

Exkurs: Zur Orientierung am Körper (anatomische Lagebezeichnungen)

Um Lagen im Bezug zum Körper eindeutig zu bestimmen und Mißverständnissen vorzubeugen ist es notwendig, dafür eindeutige Begriffe zu verwenden. Die Anatomie verfügt über ein entsprechendes Vokabular (vgl. Pera et al. 2003, 9, 12). Das einzige Bezugssystem dieser Termini ist der Körper selbst, dessen Lage im Raum unberücksichtigt bleibt (vgl. Pera et al. 2003, 13). Diffusionen durch alltägliche, bezugsflexible Bezeichnungen wie „oben“ oder „über“ werden auf diese Weise vermieden. Allein aus dieser, rein praktischen Überlegung heraus, werden im Folgenden *überwiegend* Begriffe der anatomischen Nomenklatur benutzt. Die wichtigsten und häufigsten sollen an dieser Stelle kurz erklärt werden.

Das Grundgerüst für eine Orientierung am Körper bilden die jeweils senkrecht zueinander stehenden Körperachsen, aus denen sich auch die entsprechenden Schnittebenen ableiten (vgl. Pera et al. 2003, 12; Szunyoghy & Fehér 2006, 11).

Longitudinale Achse. Jene Achse, die man intuitiv wohl als „Längsachse“ durch den Körper bezeichnen würde, wird *vertikale* bzw. *longitudinale Achse* genannt. Sie verläuft sozusagen von „Kopf bis Fuß“ (genauer gesagt zwischen die Füße) bzw. von *kranial* („kopfwärts“) nach *kaudal* („schwanzwärts“, „steißwärts“) durch den Körper – quasi wie ein Grillspieß durchs Spanferkel.

Sagittale Achse. Kippt man die Achse um 90° nach vorne (anterior), so daß sie nun von vorne (*ventral*, „bauchwärts“) nach hinten (*dorsal*, „rückenwärts“) durch den Körper führt, handelt es sich um die *sagittale Achse*.

Horizontale Achse. Dreht man die sagittale Achse 90° um die Longitudinalachse, so daß sie nun von links (*sinistral*) nach rechts (*dextral*) – vergleichbar mit der Haltestange einer Tischfußballfigur – durch den Körper führt, erhält man die *transversale* bzw. *horizontale Achse* (vgl. Gelfand 2004, 38; Pera et al 2003, 12f.).

Sagittalebene. Die longitudinale Achse und die sagittale Achse bilden zusammen eine Schnittebene, die den Körper in eine linke und eine rechte Hälfte teilt. Diese Ebene wird *Sagittalebene* genannt. Liegt die Ebene exakt in der Körpermitte, so daß sich zwei symmetrische Hälften ergeben, wird sie auch als *Medianebene* (Median-Sagittalebene) bezeichnet.

Horizontalebene. Kippt man die Sagittalebene um 90° nach links oder rechts, erhält man eine Ebene, die sich aus der sagittalen und der horizontalen Achse bildet und den Körper in eine obere (*superiore*) und eine untere (*inferiore*) Hälfte teilt. Sie wird als *Horizontalebene* oder *Transversalebene* bezeichnet.

Frontalebene. Kippt man die Horizontalebene um 90° nach vorne (oder dreht die Sagittalebene 90° um die Longitudinalachse), erhält man schließlich die *Frontalebene*. Sie teilt den Körper in eine vordere (anteriore) und eine hintere (posteriore) Hälfte (vgl. Gelfand 2004, 38; 2009, 35; Pera et al. 2003, 12f.; Szunyoghy & Fehér 2006, 11f.).

Aus den genannten Achsen und Ebenen leiten sich die anatomischen Richtungsbezeichnungen ab. Die hauptsächlich verwendeten sind in Tabelle 2.1 kurz zusammengefaßt:

Begriff	am Körper	Ansicht
anterior	(weiter) vorne	von vorne
apikal	zur Spitze (Apex)	von der Spitze
basal	zur Basis	von der Basis
dorsal	rückenwärts	rückseitig, von hinten
inferior	(weiter) unten	von unten
kaudal	schwanz-, steißwärts	steißseitig, von unten
lateral	nach außen	von der Seite, seitlich
medial	zur Mitte, mittig	von der Mitte
posterior	(weiter) hinten	von hinten
superior	(weiter) oben	von oben
ventral	bauchwärts	bauchseitig, von vorne

Tab. 2.1 **Anatomische Lagebezeichnungen.** Die lateinischen Termini können sowohl als Lagebezeichnungen (Mitte), als auch als Perspektiven (links) verstanden werden.

2.1 Schläfenbein

Die inneren Teile des Ohrs liegen im *Schläfenbein* (Os temporale) der Schädelbasis. Das Schläfenbein entwickelt sich aus drei Anlagen (deren Strukturen teilweise sogar noch weiter differenziert werden; vgl. Gelfand 2004, 40f.; 2009, 36), die zu einem einheitlichen Knochen verwachsen (vgl. Fanghänel et al. 2003(b), 196; Schünke et al. 2009, 32; Spee 1896, 148) und an der Einkapselung des Ohrs beteiligt sind. Eine sehr detaillierte Beschreibung des Schläfenbeins und aller seiner Bestandteile findet sich bei Spee (vgl. 1896, 148–222).

Überblicksartig geschildert sieht das Schläfenbein folgendermaßen aus: Der Schuppenteil (*Schläfenbeinschuppe*, Pars squamosa) liegt im Schläfenbein superior, bildet nach ebendort die großflächige Schläfenwand des Schädels und nach kaudal einerseits die Dachstruktur des Gehörganges, vor allem aber anterior davon die Angriffsfläche des Kiefergelenks (*Kiefergelenksgrube*, Fossa mandibularis). Das *Felsenbein* (Felsenteil, Pars petrosa) liegt inferoposterior an der Schuppe und bildet u.a. die massive anteromedial gerichtete, *Felsenbeinpyramide* genannte Struktur des Schläfenbeins, welche das empfindliche Innenohr beherbergt. Der *Paukenteil* (Pars tympanica) bildet im Dreiviertelkreis die anteriore und inferoposteriore Wand des äußeren Gehörganges und dockt von unten, anterolateral des Felsenbeins, an den Schuppenteil an (vgl. Fanghänel et al. 2003(b), 197ff., 208; Gelfand 2004, 42; 2009, 36ff.; Schünke et al. 2009, 32f.). Anterior des Paukenteils ist die Kiefergelenksgrube durch eine hauchdünne Platte des Felsenbeins vom Schuppenteil getrennt (vgl. Fanghänel et al. 2003(b), 208). Die Naht zwischen Paukenteil und dem Felsenbeinspan (*Glaserpalte*, *Fissura petrotympanica*; vor dem Span zum Kiefergelenk liegt die *Fissura petrosquamosa*) bietet Schlupflöcher für einige Komponenten des Mittelohrs (vgl. Schünke et al. 2009, 121 C, 133). Unmittelbar posterior des Paukenteils entwickelt das Felsenbein (und damit auch Teile der Schuppe) durch den Zug der *großen Kopfwender* (Musculi sternocleidomastoidei) in den ersten Lebensjahren nach inferior den *Warzenfortsatz* (Processus mastoideum) (vgl. Fanghänel et al. 2003(b), 199f.; Schünke et al. 2009, 32f.). Er ist, durch in Zahl, Form und Größe variierende Zellen (Warzenfortsatzzellen, Cellulae mastoideae), innerlich pneumatisiert und über die Warzenfortsatzhöhle mit der Paukenhöhle verbunden (vgl. Fanghänel et al. 2003(b), 199f.; Gelfand 2004, 42; 2009, 38; Reiss 2003, 610; Schünke et al. 2009, 32f.).



Abb. 2.2 **Schläfenbein** (modifiziert n. Schünke et al. 2009, 33 a & b). Links: Ansicht von Lateral. Rechts: Ansicht von kaudal. Der Paukenteil ist jeweils türkis gefärbt, das Felsenbein grün, der Schuppenteil ist grau.

2.2 Außenohr

Das *Außenohr* (Auris externa) setzt sich aus der *Ohrmuschel* und dem *äußeren Gehörgang* zusammen (vgl. Gelfand 2004, 37; 2009, 38; Reiss 2003, 591; Schünke et al. 2009, 128; Schwalbe 1897, 113; Spitzer 2002, 56). Das Außenohr funktioniert wie ein Trichter, der den Schall bündelt und ans Trommelfell leitet. Die Ohrmuschel liegt seitlich am Kopf an und ist dort unbeweglich nach vorne gewandt. Sie besteht aus knorpeligem Gewebe und ist – von einer groben Grundarchitektur abgesehen – sehr individuell geformt. Zur Mitte hin bildet sie eine trichterförmige, vertiefende Öffnung zum Gehörgang (vgl. Gelfand 2004, 42f.; Reiss 2003, 593; Terhardt 1998, 55). An dessen Ende sitzt das Trommelfell, „eine schrägliegende Hautmembran, welche am Rande der Mittelohrhöhle angewachsen ist“ (Terhardt 1998, 55; vgl. Gelfand 2004, 44; Reiss 2003, 593).

2.2.1 Ohrmuschel

Die *Ohrmuschel* (Auricula) ist eine unbewegliche, seitlich am Kopf gelegene, sehr variabel geformte Hautfalte, der durch eine Knorpelplatte Stabilität und eine grobe Grundstruktur verliehen wird (vgl. Guski 1996, 88; Ranke 1953, 35; Reiss 2003, 591; Schwalbe 1897, 113; Terhardt 1998, 55). Bis auf das rein häutige Ohrläppchen ist die gesamte Ohrmuschel durch das Knorpelskelett gestützt. Vorne und hinten am Knorpel sitzen verschiedene, beim Menschen stark zurückgebildete und vollkommen funktionslose Muskeln (vgl. Gelfand 2004, 42; 2009, 38; Reiss 2003, 594; Schünke 2009, 127; Schwalbe 1897, 143).

Die Struktur der Ohrmuschel ist im Detail sehr individuell, eine einheitliche Grundstruktur ist allerdings gut zu erkennen (vgl. Guski 1996, 88; Reiss 2003, 591; Terhardt 1998, 55). Sie lässt sich im Grunde als ein mehr oder weniger halbkreisförmiger Trichter beschreiben, der von zwei Bögen umlaufen wird. Vom Ohrläppchen radial um die gesamte Muschel herum verläuft mit der *Ohrleiste* (Helix) der äußere Bogen, eine nach innen geschlagene, die Ohrmuschel säumende Krempe. Anterior der Muschel zieht sie

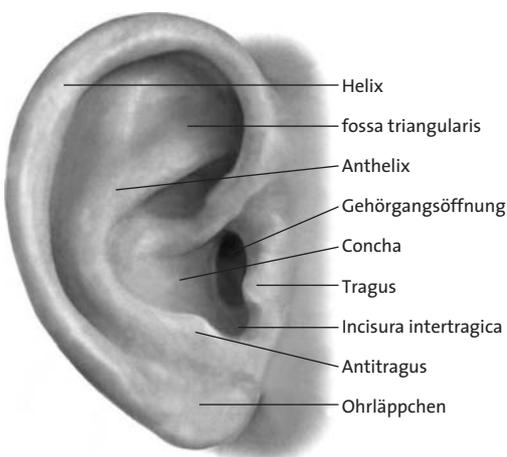


Abb. 2.3 **Ohrmuschel** (modifiziert n. Schünke et al. 2009, 127 B).

fast bis zur Mitte herab, um dann mit einer scharfen Kurve als quasi horizontaler Grat (*Crus helix*) in die halbkreisförmige, trichterartige Vertiefung der Muschel (*Concha*, *Cavum conchae*) hineinzulaufen. Der Trichter mündet unmittelbar unterhalb des anterioren Helixknicks im Gehörgang (vgl. Gelfand 2004, 43; Reiss 2003, 594; Schünke et al. 2009, 127; Schwalbe 1897, 115, 119; Terhardt 1998, 55). Direkt unter dessen Öffnung liegt eine auffangbeckenartige Vertiefung (*Incisura intertragica*). Sie wird durch zwei Knorpelvorsprünge, den direkt (schützend) über die Gehörgangsöffnung (also lateral davon) ragenden *Tragus* und den, am inferioren Rand der

Concha sitzenden *Antitragus* geformt. Aus dem Antitragus geht der innere Muschelbogen, die *Gegenleiste* (Anthelix, Antihelix⁹) hervor. Er verläuft wulstartig innerhalb der Helix um den Trichter herum. Im oberen Muschelbereich, superior der *Crus heliis*, spaltet er sich in zwei Arme (*Crus anthelicis superioris* bzw. *inferius*), die mit der *Fossa triangularis* eine dreieckige Vertiefung umschließen (vgl. Gelfand 2004, 43; 2009, 39; Reiss 2003, 594; Schünke et al. 2009, 127; Schwalbe 1897, 113, 118f.).

So wie die Form, fällt auch die Größe der Ohrmuschel von Mensch zu Mensch sehr unterschiedlich aus, wobei sie sich im Laufe des Lebens auch intrapersonell ändert und mit dem Alter größer wird (vgl. Guski 1996, 88; Tan et al. 1997). Beim Erwachsenen beträgt sie sagittal ungefähr 30 bis 40 mm (Guski (vgl. 1996, 88) mißt sogar nur 20 mm) und longitudinal zwischen 50 und 70 mm (vgl. Guski 1996, 88; Terhardt 1998, 55).

2.2.2 Gehörgang

Der *äußere Gehörgang* (Meatus acusticus externus) ist ein schlauchartiger, ins Schläfenbein führender, S-gewundener, in medialer Richtung leicht konisch zulaufender, im Querschnitt unregelmäßiger Kanal (vgl. Gelfand 2004, 44; 2009, 39; Ranke 1953, 36; Reiss 2003, 596; Schünke et al. 2009, 129; Schwalbe 1897, 164). Derart überblicksmäßig betrachtet, ist der Gehörgang recht übereinstimmend beschrieben. Geht man jedoch tiefer ins Detail, ist stellenweise überraschend, wie weit die Beschreibungen einer vergleichsweise gut zugänglichen Struktur voneinander abweichen. Reiss beschreibt den Verlauf der Gehörgangswindung folgendermaßen:

Der knorpelige Teil ist von innen gesehen nach oben, hinten und lateral gerichtet und verengt sich trichterförmig gegen den knöchernen Teil. An der Verbindungsstelle ist er am engsten [...], um sich im knöchernen Teil abwärts, vorwärts, medial zu wenden. Gegen das Trommelfell wird er wieder weiter. (Reiss 2003, 596).

An irgendeiner Stelle seiner Beschreibung muß ein Irrtum vorliegen, denn wie man sie auch dreht und wendet: sie beschreibt eine gerade Linie. Andere Autoren bezeichnen den Gehörgang immerhin als leicht, bzw. konvex gebogene Röhre (vgl. Hellbrück & Ellermeier 2004, 91; Plattig in Keidel 1975, 48). Wahrscheinlich handelt es sich bei Reiss um einen Orientierungsfehler, denn an anderer Stelle (vgl. 203, 596) liegt er mit dem Gros der (hier verwendeten) Literatur in Übereinstimmung. Faßt man diese zusammen, führt der Gehörgang von außen Richtung Trommelfell zunächst nach anterosuperior, macht nach etwa einem Drittel eine Kurve nach superoposterior und ungefähr vor dem letzten Drittel nach antero-inferior. Vereinfacht gesagt vollzieht die Windung des Gehörgangs zwei Kurven, deren Verlauf man als dreidimensionale S-Kurve bezeichnen könnte (vgl. Gelfand 2004, 44; 2009, 39; Reiss 2003, 596; Schünke et al. 2009, 129; Schwalbe 1897, 164–169).

Gemessen an seiner Windung und seiner dermalen, bzw. subkutanen Struktur, könnte man den Verlauf des Gehörgangs in drei Abschnitte gliedern. Von der Eintrittsöffnung in der Concha verläuft er zunächst in anterosuperiorer Richtung (vgl. Gelfand 2004, 44; 2009, Abb. 2.3, 36; Schünke et al. 2009, 129 a & b, 130 A). In diesem äußeren Abschnitt ist er unter der Haut knorpelig ausgekleidet. Der Knorpel ist die Fortsetzung des Ohrmuschelknorpels. Durch seine löchrige, mit Bindegewebe aufgefüllte

9 Bei Fanghänel et al. (2003, 594) wird der *Antitragus* alternativ als *Antihelix* bezeichnet, was verwirrend sein kann, da der Antitragus fließend in die *Anthelix* (d.h. der Innere Bogen der Ohrmuschel) übergeht.

Struktur bildet er die Grundlage für die Flexibilität des äußeren Gehörgangs. Die Haut der Concha setzt sich ebenfalls in der Haut des Gehörgangs fort. Hier sitzt die überwiegende Mehrheit der Talk- und Zeruminaldrüsen, welche für die Absonderung des reinigenden Ohrenschmalz (*Cerumen*) verantwortlich sind (vgl. Reiss 2003, 596; Schünke 2009, 129; Schwalbe 1897, 176). Im mittleren Drittel verliert sich die subkutane Knorpelschicht und der Gehörgang führt direkt durch den Knochen. Mit ihr nimmt auch die Haar- und Drüsendichte in der Haut des Gehörgangs ab. (vgl. Gelfand 2004, 44; Reiss 2003, 596; Schünke 2009, 129; Schwalbe 1897, 158, 173). Das innerste Drittel verläuft schließlich nach anteroinferior. Die obere Wandung des Gehörgangs wendet sich jedoch nur zaghaft nach inferior, was, bei großzügiger Betrachtung, als leichte Erweiterung des Gehörgangs entgegen seiner medial gerichteten Verjüngung interpretiert werden kann (vgl. Helmholtz 1896, 211; Reiss 2003, 596; Schwalbe 1897, 167). Am Ende des Gehörgangs sitzt das *Trommelfell*, „eine schrägliegende Hautmembran, welche am Rande der Mittelohrhöhle angewachsen ist“ (Terhardt 1998, 55; vgl. Gelfand 2004, 44; 2009, 39; Stevens & Hallowell 1938, 249) und das Außenohr vom Mittelohr abgrenzt.

Durch die Schräglage des Trommelfells (siehe 2.3.1, S. 56) ist die hintere, obere Gehörgangswand rund 6 mm kürzer als die vordere untere (vgl. Bürkner 1900, 7; Reiss 2003, 596; Schwalbe 1897, 163; Gelfand (2004, 44) mißt hier sogar einen Längenunterschied von fast 100%). Die durchschnittliche Länge des Gehörgangs wird zwischen 23 bis 35 mm angegeben, sein Volumen zwischen 0,8 und 1 cm³ (vgl. Békésy 1960, 135; Békésy & Rosenblith 1951, Tab. 1, 1079, 1079; Gelfand 2009, 39; Guski 1996, 89; Hellbrück & Ellermeier 2004, 91; Plattig in Keidel 1975, 48; Ranke 1953, 36; Reiss 2003; 596; Schünke 2009, 129; Stevens & Hallowell 1938, 249; Spitzer 2002, 56; Terhardt 1998, 55). Diese, angesichts der Dimensionen bemerkenswerte Spanne ist möglicherweise auf den genauen Meßort zurückzuführen. Nach Angaben von Schwalbe (vgl. 1897, 169f.) ist die posteriore Gehörgangswand ungefähr 24 mm. die anteriore dagegen rund 35 mm lang.

Während die in den Gehörgang führende Concha als annähernd halbkreisförmig beschrieben werden kann (siehe 2.2.1, S. 53, vgl. Terhardt 1998, 55), stellt sich der Gehörgangquerschnitt eher als oval dar (vgl. Reiss 2003, 596). Dadurch, und aufgrund seiner pseudokonischen Verjüngung Richtung Trommelfell, variieren die Querschnittmaße teils erheblich. Während Gelfand (2004, 44; 2009, 39) für einen ausgewachsenen Gehörgang eine durchschnittliche Höhe (9 mm) und Breite (6,5 mm) angibt, wird in anderer Literatur diesbezüglich kaum differenziert. Hier schwanken die Angaben zwischen 6 und 8 mm (vgl. Békésy & Rosenblith 1951, Tab.1, 1079; Guski 1996, 89; Hellbrück & Ellermeier 2004, 91; Plattig in Keidel 1975, 48; Ranke 1953, 36; Schünke et al. 2009, 129; Stevens & Hallowell 1938, 249; Terhardt 1998, 55).

2.3 Mittelohr

Zwischen dem das Außenohr begrenzenden *Trommelfell* einerseits und den das Innenohr abschließenden Membranen (*ovales*, bzw. *rundes Fenster*) andererseits liegt das *Mittelohr* (*Auris Media*) (vgl. Guski 1996, 90; Reiss 2003, 599; Schünke et al. 2009, 130; Terhardt 1998, 55). Es ist ein Höhlenkomplex im Felsenbein, der sich aus der *Paukenhöhle*, den Hohlräumen des posterior des Ohrs gelegenen *Warzenfortsatzes* sowie der Mittelohr und Rachenraum verbindenden *Eustachischen Röhre* zusammensetzt

(vgl. Guski 1996, 90; Reiss 2003, 599; Schünke et al. 2009, 130; Stepp & Voss 2005, 861). In der Paukenhöhle befinden sich die drei Gehörknöchelchen *Hammer*, *Amboß*, und *Steigbügel*. „Es handelt sich dabei um sehr feine und leichte Kalziumkörperchen, welche durch feine Sehnen und Muskeln miteinander verbunden“ (Terhardt 1998, 55) und im Hohlraum des Mittelohrs flexibel gelagert sind (vgl. Gelfand 2004, 47f.; 2009, 42; Reiss 2003, 603). Sie schaffen „eine gelenkige Verbindung zwischen Trommelfell und ovalem Fenster“ (Schünke et al. 2009, 132), deren Hauptaufgabe es ist, die übers Außenohr ans Trommelfell geleiteten Luftschwingungen den Innenohrflüssigkeiten zu vermitteln (vgl. Gelfand 2004, 87ff.; Guski 1996, 90; Helmholtz 1896, 211; Müller & Frings 2009, 430; Plattig in Keidel 1975, 50; Reiss 2003, 591, 603; Schünke et al. 2009, 126, 132; Spitzer 2002, 58; Terhardt 1998, 56; Zwicker & Feldtkeller 1967, 83). Darüber hinaus stellt es eine wichtige Schutzinstanz des empfindlichen Innenohrs dar (vgl. Plattig in Keidel 1975, 50; Spitzer 2002, 58; Terhardt 1998, 56).

2.3.1 Trommelfell

Das *Trommelfell* (Membrana tympani) grenzt Außenohr und Mittelohr voneinander ab (vgl. Gelfand 2004, 37; Hellbrück & Ellermeier 2004, 91; Plattig in Keidel 1975, Abb. 2, 47; Ranke 1953, 37; Reiss 2003, 597; Spitzer 2002, 57). Die Zuordnung des Trommelfells selbst fällt bei dieser Grenzziehung unterschiedlich aus. Obwohl es nach Gelfand normalerweise als Teil des Mittelohrs verstanden (vgl. Gelfand 2004, 37; 2009, 38) und entsprechend subsumiert wird (vgl. Spitzer 2002, 57), findet man es häufig – wenn auch vielfach nur implizit – dem Außenohr zugeordnet (vgl. Guski 1996, 89f.; Reiss 2003, 593; Schünke et al. 2009, 128f.; Terhardt 1998, 55). Mitunter wird auch zwischen Außenohr, Trommelfell und Mittelohr differenziert, und das Trommelfell so implizit im „Niemandland“ zwischen den beiden Ohrteilen angesiedelt (vgl. Plattig in Keidel 1975, 46).

Anhand seiner Schichtung aus Gehörgangshaut, funktioneller Membran (Grundschicht) und Paukenhöhlenschleimhaut (siehe S. 58), läßt sich das Trommelfell nicht eindeutig zuordnen. Physiologisch muß man es m.E. eher dem Mittelohr zurechnen, da die Gehörknöchelchen erst *gemeinsam* mit dem ovalen Fenster als Empfänger und dem Trommelfell, dessen Eingangssignal weitergeleitet wird, ein vollständiges Übertragungssystem bilden können (vgl. Gelfand 2009, 42; Hellbrück & Ellermeier 2004, 92; Plattig in Keidel 1975, 52; Reiss 2003, 603f.; Schünke et al. 2009, 132; Spitzer 2002, 57). Neben seiner Hauptaufgabe als akustischer Druckempfänger (vgl. Hellbrück & Ellermeier 2004, 92), beteiligen sich Teile des Trommelfells (hier ist die *Shrapnellmembran* (siehe 3.2.1, S. 104) hervorzuheben) auch am Druckausgleich zwischen Außen- und Innendruck der Paukenhöhle (vgl. Didyk et al. 2007).

Das Trommelfell ist eine näherungsweise ovale, pseudokonisch geformte, perlgrau erscheinende, mehrschichtige Membran, die in einer schrägen Stellung am Ende des Gehörgangs sitzt (vgl. Bürkner 1900; Gelfand 2004, 44; 2009, 40; Schünke et al. 2009, 129; Terhardt 1998, 55). Seine laterale Seite ist nach anteroinferior gerichtet. Wegen seines konischen Einzugs wird die Lage des Trommelfells immer mittels einer imaginären, durch seinen Rand (den *Paukenring*) gelegten Ebene definiert (vgl. Politzer 1896, 4). Im Horizontalschnitt bildet es mit der Sagittalachse einen nach hinten offenen Winkel von etwa 50°, im Frontalschnitt mit der Longitudinalachse einen nach oben offenen Winkel von rund 35 bis 45°, wobei auch Weiten bis 55° angegeben wurden (vgl. Parent & Allen 2010, Abb. 1 (a), 154; Politzer 1896, 4f.;

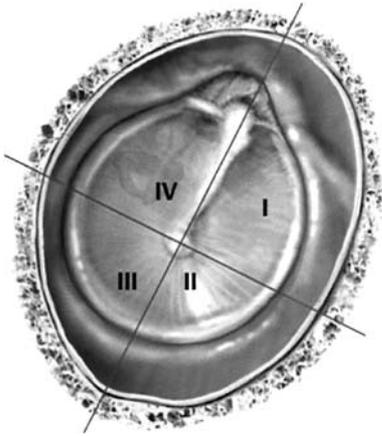


Abb. 2.4 **Trommelfell** (modifiziert n. Schünke et al. 2009, 129 E). Nummeriert sind die vier Quadranten.

Ranke 1953, 37; Reiss 2003, 598; Siebenmann 1897, 248; Schwalbe 1897, 163). Die vertikale Achse der Trommelfellfläche ist aus lateraler Sicht um etwa 30° nach vorne geneigt, so daß sie von anterosuperior nach posteroinferior verläuft (vgl. Gelfand 2004, 44; 2009, 40).

Die Form des Trommelfells ergibt sich aus seiner Aufhängung zwischen Gehörgang und Mittelohrknochen. Seine Grundschicht, die *Lamina propria*, ist am Rand zu einem Bindegewebsring (*Paukenring*, *Anulus fibrocartilagineus*) ausgebildet, der in einer Rinne des Schläfenbeins (*Sulcus tympanicus*) am Ende des annähernd ovalen Gehörgangs verankert ist (vgl. Helmholtz 1896, 211; Reiss 2003, 598f.; Siebenmann 1897; 255). Rosowski berichtet von Muskelgewebefunden und der darauf basierenden Möglichkeit, im Paukenring könnte ein weiterer (siehe 2.3.4, S. 67) Mittelohrmuskel schlummern (vgl. Rosowski 2010, 54). In diesem Haltering ist das Trommelfell aufgespannt und wird durch den *Hammergriff*, mit dem es verwachsen ist (vgl. Hellbrück & Ellermeier 2004, 92; Ranke 1953 39; Reiss 2003, 603; Schünke et al. 2009, 129; Stevens & Hallowell 1938, 249), näherungsweise konisch Richtung Mittelohr gezogen. Der untere Ansatz des Hammergriffs liegt ein kleines Stück unterhalb des imaginären Trommelfellmittelpunktes und rund 2 mm medial einer durch den Bindegewebsring gelegten Ebene, wo er den Nabel (*Umbo*) der konischen Vertiefung bildet (vgl. Bürkner 1900, 7; Gelfand 2004, 44; 2009, 40; Helmholtz 1896, 214; Politzer 1896, 5; Reiss 2003, 597ff.; Schünke et al. 2009, 129; Siebenmann 1897, 249; Stevens & Hallowell 1938, 249). Blickt man durch den Gehörgang¹⁰ aufs Trommelfell, sieht man den Hammergriff als helle Knochenstruktur (*Stria mallearis*) durchschimmern. Er verläuft in anterosuperiorer Richtung ungefähr bis zum Rand der ovalen Grundform des Trommelfells und endet dort in einem hervorstehenden Knopf (*Prominentia mallearis*), der durch den *seitlichen Hammerfortsatz* gebildet wird (vgl. Bürkner 1900, 7f.; Gelfand 2004, 44; 2009, 40; Helmholtz 1896, 214; Reiss 2003, 597; Schünke et al. 2009, 129). Sein Neigungswinkel entspricht im rechten Ohr ungefähr der Ein-Uhr-Zeigerposition, im linken Ohr ungefähr jener von elf Uhr (vgl. Gelfand 2004, 44; 2009, 40). Daraus läßt sich die Achsneigung des Trommelfells von ca. 30° ableiten. Axial oberhalb der *Prominentia mallearis* ist die Trommelfellfläche auffällig erweitert – ähnlich dem Längsquerschnitt einer Birne, oder eines unregelmäßigen, an der Spitze abgerundeten Herzens (vgl. Politzer 1896, 2; Siebenmann 1897, 249). In diesem Bereich ist der Paukenring quasi wie ein Hufeisen unterbrochen (*Incisura tympanica*) und das Trommelfell schlaff (vgl. Bürkner 1900, 7; Politzer 1896, 1f.; Reiss 2003, 598). Die beiden Endpunkte des Paukenrings werden durch kurze Knochenvorsprünge verdeutlicht, von denen der anteriore als *Spina tympanica major* und der posteriore als *Spina tympanica minor* bezeichnet wird. Der Paukenring selbst geht hier nahtlos in die Beinhaut (*Periost*) des Schläfenbeins über (vgl. Politzer 1896, 1, 13). Vom Hammerfortsatz zu den beiden Endpunkten verlaufen zwei mehr oder weniger auffällige Falten, die den gespannten Hauptteil des Trommelfells (*Pars tensa*) vom schlaffen Nebenteil (*Shrapnellmembran*, *Pars flaccida*) im Bereich zwischen *Incisura tympanica* und *Hammerfortsatz* gut sichtbar trennen (vgl. Bürkner 1900, 9; Gelfand 2004, 44; 2009; Politzer 1896, 6; Reiss 2003, 597; Siebenmann 1897, 282).

¹⁰ Durch die Windung des Gehörgangs ist die Sicht aufs Trommelfell verdeckt. Um die Windung auszugleichen und den Blick aufs Trommelfell freizumachen, muß die Ohrmuschel nach hinten und oben gezogen werden (vgl. Gelfand 2009, 39; Plattig 1975(a), 48; Reiss 2003, 596).

Das Trommelfell hat eine sehnige Struktur. Die Lamina propria besteht aus einer Ebene radial verlaufender, steifer Fasern (*Stratum radiale*) und einer medial davon gelegenen Ebene zirkulär laufender, elastischer Fasern (*Stratum circulare*) (vgl. Békésy 1941(a), 9; Gelfand 2004, 44; 2009, 40; Helmholtz 1896, 220; Politzer 1896, 11; Reiss 2003, 597f.; Siebenmann 1897, 259f.). Insgesamt ergibt sich daraus eine spinnennetzartige Faserstruktur, die zumeist zu einer Schicht zusammengefaßt wird, jedoch auch als doppelschichtig aufgefaßt werden kann (vgl. Gelfand, 2004, 44; 2009, 40). Funktionell bilden beide Ebenen eine Einheit, da sie dem Trommelfell erst im Verbund seine Stabilität und sein Spannungsverhalten verleihen. Die zirkulären Fasern finden sich vor allem am Rand, d.h. zum Paukenring hin gelegen. In der Nähe des Nabels fehlen sie ganz. Die radialen Fasern sind nur in der „Kurve“ inferior des Nabels tatsächlich radial ausgerichtet. Auf den „Geraden“ oberhalb des Nabels verlaufen sie mehr oder weniger horizontal vom Paukenring zum Hammergriff (vgl. Politzer 1896, 12; Siebenmann 1897, 260). Nach Helmholtz sind die Radialfasern nicht vollkommen gerade, sondern durch den Zug der aus seiner Sicht elastischen Ringfasern leicht konvex (Richtung Gehörgang) gespannt (vgl. Békésy 1941(a), 9f.; Helmholtz 1896, 220; Ranke 1953, 38). Beides wird von anderen Autoren nicht erwähnt. Am unteren äußeren Rand des Trommelfells beschreibt Békésy eine nur selten erwähnte¹¹ Falte des Trommelfells, die sich quasi aus der konvexen Wölbung heraus bildet. Ihr Krümmungsradius beträgt etwa 0,5 bis 0,8 mm. Sie verleiht dem Trommelfell genug Spielraum, um unter bestimmten Bedingungen kolbenartig zu schwingen (vgl. Békésy 1941(a), 13f.; Ranke 1953, 37). Lateral ist die Lamina propria mit einer dünnen, haar- und drüsenfreien Hautschicht bedeckt (*Stratum cutaneum*), die eine unmittelbare Fortsetzung der Haut des Gehörgangs darstellt. Medial ist sie mit einer verdünnten, einschichtigen Fortsetzung der Paukenhöhlenschleimhaut (*Stratum mucosum*) ausgekleidet (vgl. Gelfand 2004, 44; 2009, 40; Reiss 2003, 598). Während die *Lamina propria* der *Pars tensa* sozusagen intrinsisch ist, sich hier der dreischichtige Aufbau des Trommelfells also eindeutig festschreiben läßt, ist die Schichtung der Shrapnellmembran weniger eindeutig. Da ihr die faserige Struktur der Lamina propria fehlt, wird sie oft (vgl. Gelfand 2004, 44; 2009, 40; Politzer 1896, 3f.; Siebenmann 1897, 259) als zweischichtig (*Stratum cutaneum* und *mucosum*) beschrieben. Reiss findet jedoch auch hier eine „locker aufgebaute Lamina propria ohne radiärer und zirkulärer Faseranordnung“ (2003, 598).

Die Dicke des Trommelfells wird zumeist mit Werten um 0,1 mm angegeben (vgl. Békésy & Rosenblith 1951, Tab. 1, 1079; Gelfand 2004, 44; 2009, 40; Reiss 2003, 598; Siebenmann 1897, 247), wobei nach Békésy nur 0,05 mm zutreffen sollen (vgl. Békésy 1960, 195). Diese Werte wurden vermutlich im frei gespannten Teil der Membran gemessen, da sie in ihren Randbereichen nahe des fibrösen Halterings bzw. des Hammergriffs mit bis zu 0,4 mm deutlich dicker sein dürfte (vgl. Politzer 1896, 9). In der Literatur wird der Durchmesser des Trommelfells für gewöhnlich mit Werten von ungefähr 8 bis 10 mm angegeben (vgl. Gelfand 2004, 44; 2009, 40; Hellbrück & Ellermeier 2004, 91; Ranke 1953, 37; Terhardt 1998, 55). Vertikal gemessen liegen die Werte mit 9 bis 10 mm meist knapp 1 mm über den horizontal gemessenen (vgl. Gelfand 2004, 44; 2009, 40; Helmholtz 1896, 219; Politzer 1896, 2), wobei Erstere die Strecke vom unteren Scheitelpunkt des Paukenrings bis zur Spina tympanica minor abdecken (vgl. Politzer 1896, 2). Die Ausdehnung der halbmondförmigen Shrapnellmembran wird mit einer radialen Höhe von 1,75 bis 3 mm und einer Breite von rund 2 bis 3 mm (entspricht der Incisura tympanica) angegeben (vgl. Politzer 1896, 6; Siebenmann 1897, 294).

11 „In vielen anatomischen Zeichnungen der Handbücher fehlt diese Falte, und sie war auch an manchen Präparaten nur sehr schwach ausgebildet“ (Békésy 1941(a), 14). Möglicherweise liegt darin der Grund, warum diese Trommelfellfalte in der Literatur kaum zu finden ist.

Angaben der Trommelfellfläche schwanken zwischen 50 und 90 mm² (vgl. Békésy & Rosenblith 1951, Tab. 1, 1079; Gelfand 2004, 90; 2009, 51; Müller & Frings 2009, 452; Schünke et al. 2009, 129; Stevens & Hallowell 1938, 260; Terhardt 1998, 55; bei Lehnhardt (vgl. Lehnardt & Laszig 2009, 27) sind 6,4 mm² genannt), wobei die wirksame Fläche zumeist mit rund 55 bis 60 mm² angegeben wird (vgl. Békésy 1960, 102; Békésy & Rosenblith 1951, 1082; Müller & Frings 2009, 452; Plattig in Keidel 1975, 52; Yost 2008, 76).

Der Vollständigkeit halber sei erwähnt, daß das Trommelfell in vier Quadranten unterteilt wird, die nach anterior abwärts (also im rechten Ohr mit, im linken gegen den Uhrzeigersinn) beziffert werden (vgl. Bürkner 1900, 8; Politzer 1896, 17; Schünke et al. 2009, 129; Siebenmann 1897, 250). Diese Teilung ist „klinisch wichtig, da sie der Beschreibung der Lokalisation von krankhaften Veränderungen dient“ (Schünke et al. 2009, 129). Für die eigentliche Funktion ist sie jedoch bedeutungslos (für detaillierte Beschreibungen der Quadranten vgl. Politzer 1896, 17–19; Schünke et al. 2009, 129 E).

2.3.2 Mittelohrhöhle

Die *Paukenhöhle* ist der Schnittpunkt des Hohlraumkomplexes im Felsenbein. Nach posterior ist sie über die *Warzenfortsatzhöhle* mit den *Warzenfortsatzzellen* zu einem Höhlenkomplex verbunden (vgl. Gelfand 2004, 42, 46; 2009, 38, 42; Reiss 2003, 600; Schünke et al 2009, 130; Stepp & Voss 2005, 861; Voss et al. 2000, 59). Während *Trommelfell*, *ovales* und *rundes Fenster* die Mittelohrhöhle verschließen, ist sie durch die *Eustachische Röhre* mit dem Rachenraum verbunden, um notwendige Anpassungen des Innendrucks an den Umgebungsluftdruck zu ermöglichen (vgl. Gelfand 2004, 47; Guski 1996, 91; Helmholtz 1896, 211f.; Schünke et al 2009, 131; Terhardt 1998, 56; Zwicker & Feldtkeller 1967, 86).

Paukenhöhle

Die *Paukenhöhle* (Cavitas tympani) liegt unmittelbar medial des Trommelfells und ist im Horizontalschnitt parallel zu diesem, d.h. von posterolateral nach anteromedial, ausgerichtet (vgl. Schünke et al 2009, 130 A). Nach vorne kommuniziert sie durch die Eustachische Röhre mit dem Rachenraum (vgl. Békésy & Rosenblith 1951, 1075; Gelfand 2004, 46f.; 2009, 42; Guski 1996, 91; Helmholtz 1896, 212; Reiss 2003, 591, 600; Schünke et al 2009, 130; Terhardt 1998, 56) und nach hinten durch einen schmalen Zugang (Aditus) mit der Warzenfortsatzhöhle (vgl. Gelfand 2004, 42, 46; 2009, 38, 42; Reiss 2003, 600; Schünke et al 2009, 130; Stepp & Voss 2005, 861; Voss et al. 2000, 59). Gemessen an der Lage des Trommelfells wird die Paukenhöhle in drei Stockwerke unterteilt. Aus lateraler Sicht direkt hinter dem Trommelfell liegt das mittlere Stockwerk (Mesotympanon), der Kern der Paukenhöhle mit der Übertragungskette aufs ovale Fenster. Nach superoposterior und lateral entwickelt sich daraus der sogenannte *Paukenhöhlenkuppelraum* (Epitympanon, Recessus epitympanicus). Er verengt sich dorsal zum Eingang in den Warzenfortsatz. Überdies beherbergt er die massiven Körper der beiden Gehörknöchelchen *Hammer* und *Amboß* (vgl. Gelfand 2009, 40; Reiss 2003, 600; Schünke et al. 2009, 133). Unterhalb des Trommelfells entwickelt sich in anteromedialer Richtung der *Paukenhöhlenkeller* (Hypo-

tympanon, Recessus hypotympanicus). Dort sitzt die Austrittsöffnung zur Eustachischen Röhre, die das Mittelohr mit dem Rachenraum verbindet (vgl. Reiss 2003, 600; Schünke et al. 2009, 133). Wie die Abbildungen 2.1, S. 50, und 2.5 verraten, liegt der Paukenhöhlenheller allerdings eher auf Höhe des Trommelfells.

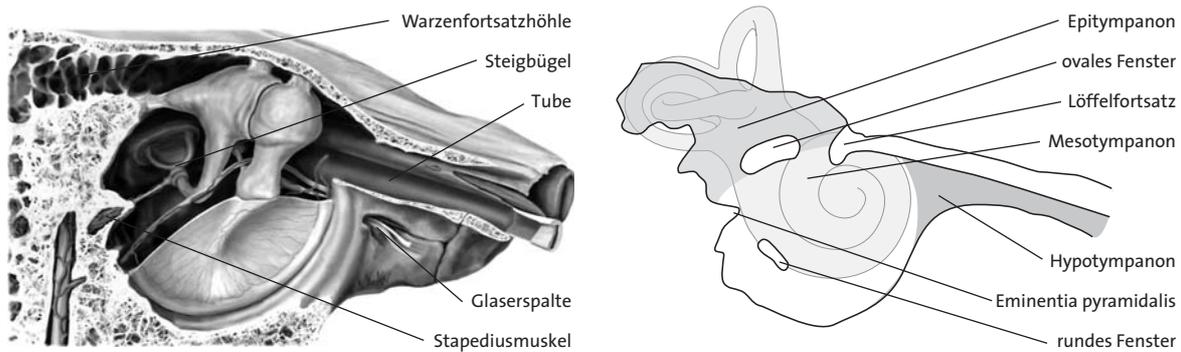


Abb. 2.5 **Paukenhöhle** eines rechten Ohrs von lateral gesehen (modifiziert n. Schünke et al. 2009, 131 C, 133 C).

Die Paukenhöhle ist ein – durch das Wachstum des Schläfenbeins bedingt passiv modellierter (vgl. Fritscher et al. 2004, 244) – unregelmäßig geformter, hoher, schmaler, sechswandiger Raum (vgl. Gelfand 2004, 44; 2009, 41; Reiss 2003, 600; Schünke et al. 2009, 130). Von lateral erinnert die Paukenhöhle an eine nach posterior gekippte, zerkrantschte Acht. Ihre mehr oder weniger deutliche Tailenverengung teilt sie in zwei Kammern, wobei die untere Kammer (im wesentlichen Mesotympanon und Kellerraum) im Schnitt deutlich größer als die obere (Kuppelraum) ist (vgl. Fritscher et al. 2004, Abb. 1, 2, 245; Schünke et al. 2009, 131 C). Der Kuppelraum ist relativ geräumig. Zwischen Trommelfell und basaler Schneckenwindung liegt ihre engste Stelle, während sie sich in den Keller hinein wieder erweitert (vgl. Plattig in Keidel 1975, Abb. 2, 47; Reiss 2003, 600).

Dünne Knochenplatten schließen die Paukenhöhle ins Schläfenbein und trennen sie von der Fossa jugularis (Boden) bzw. der mittleren Schädelgrube (Decke), die laterale Wand bildet hauptsächlich das Trommelfell (vgl. Gelfand 2004, 44; Reiss 2003, 600; Schünke et al. 2009, 130). Medial liegt das Innenohr, dessen basale Schneckenwindung sich in einer großflächigen Erhebung (*Promontorium*) gegenüber dem Trommelfell in den Raum wölbt (vgl. Gelfand 2004, 44, 46; 2009, 42; Reiss 2003, 601f.; Schünke et al. 2009, 130 B, 131 C). Ventral öffnet sich die Mündung der Eustachischen Röhre. Unmittelbar superior, nur durch eine dünne Knochenlamelle von der Tube getrennt, verläuft der Kanal des Trommelfellspanners und endet ventromedial im *Löffelfortsatz* (*Processus cochleaformis*), einem dornartigen Vorsprung (vgl. Gelfand 2004, 46; 2009, 42; Reiss 2003, 601f., 610; Schünke et al. 2009, 130, 131 C). Lateral der Tube liegt die *Glaserspalte*, eine Passage durchs Schläfenbein, die unter anderem vom vorderen Hammerband und der (fürs Gehör unbedeutenden) *Paukensaiten* (*Chorda tympani*) genutzt wird (vgl. Schünke et al. 2009, 121; 131 C; Reiss 2003, 601, 603). Dorsal, auf Höhe des ovalen Fensters, findet sich ein pyramidenartiger Vorsprung (*Eminentia pyramidalis*), durch dessen offene Spitze die Sehne des *Steigbügelmuskels* in die Paukenhöhle tritt (vgl. Gelfand 2004, 46, 49; 2009, 42f.; Reiss 2003, 600; Schünke et al. 2009, 130ff.).

Warzenfortsatzzellen

Die *Paukenhöhlenkuppel* verengt sich dorsal zum Zugang (Aditus) zur posterior der Paukenhöhle gelegenen kurzen *Warzenfortsatzhöhle* (Antrum mastoideum). Ihr schließt sich nach posteroinferior ein in Form, Zahl und Größe variabler Hohlraumkomplex aus vielen kleinen Kammern, den *Warzenfortsatzzellen* (Cellulae mastoideae) an. Sie bilden mit der Paukenhöhle einen gemeinsamen mit Schleimhaut ausgekleideten Hohlraum (vgl. Gelfand 2004, 42; 2009, 38; Reiss 2003, 610; Schünke et al. 2009, 32; Stepp & Voss 2005, 861; Voss et al. 2000, 59). Der *Warzenfortsatz* (Processus mastoideum) wird durch die Halsmuskulatur sozusagen aus dem Schläfenbein herausgezogen (vgl. Fanghänel et al. 2003, 199f.; Schünke et al. 2009, 32f.). Dabei wird er, parallel zu seinem Wachstum, von der Warzenfortsatzhöhle aus pneumatisiert (vgl. Fanghänel et al. 2003(b), 199f.; Reiss 2003, 610; Schünke et al. 2009, 32). Die so entstehende Struktur aus Warzenfortsatzzellen läßt sich mit dem Hohlraummuster eines groben Schwamms vergleichen. Die Durchströmung des Warzenfortsatzes mit Luft (Pneumatisierung) fällt jedoch sehr variabel aus, kann mitunter gehemmt sein oder sich auf andere Bereiche des Schläfenbeins ausweiten (vgl. Reiss 2003, 610). Dies hat zum Teil erstaunliche Volumenunterschiede zur Folge; mitunter treten sogar intrapersonelle Schwankungen von linkem zu rechtem Ohr auf (vgl. Stepp & Voss 2005, Endnote 6, 870).

Während Maße und Masse der Mittelohrkomponenten entscheidenden Einfluß auf das Hörvermögen haben (vgl. Hemilä et al. 1995), ist der Mittelohrraum aus physiologischer Sicht besonders aufgrund seines Volumens interessant, da sein Innendruck die Impedanz des Trommelfells bestimmt (vgl. Stepp & Voss 2005, 869). Die Paukenhöhle mißt an ihrer engsten Stelle, zwischen Trommelfell und Promontorium, knapp 2 bis 3 mm (vgl. Reiss 2003, 600). Stieger et alii geben für den Abstand zwischen Spitze des Hammergriffs (er fällt mit dem Nabel des Trommelfells zusammen) und Promontorium sogar einen gemessenen Minimalwert von 1,3 mm an (vgl. 2006, 378), was die Winzigkeit des Raumes verdeutlicht. Angaben zum Volumen der Paukenhöhle allein schwanken zwischen 0,5 bis 2 cm³ (vgl. Békésy & Rosenblith 1951, Tab. I, 1079; Stepp & Voss 2005, 861, 868; Stevens & Hallowell 1938, 249; Voss et al. 2000, 59). Wie bereits dargestellt ist das Volumen der Warzenfortsatzzellen deutlich größeren Schwankungen unterworfen. Für die gesamte Mittelohrhöhle (Paukenhöhle, Aditus, Warzenfortsatzhöhle und -zellen) schwanken die gemessenen Werte von 1 bis 22 cm³ (vgl. Molivaer et al. zit. nach Stepp & Voss 2005, 861; Voss et al. 2000, 59). Im Durchschnitt beträgt das Volumen des vollständigen Hohlraumkomplexes rund 6,5 cm³ (vgl. Molivaer et al. zit. nach Voss et al. 2000, 59).

Eustachische Röhre

Die *Eustachische Röhre* (Tuba eustachii, Tuba auditiva), benannt nach ihrem Entdecker Bartolomeo Eustachi (vgl. Plattig in Keidel 1975, 49), ist ein länglicher, aus der Paukenhöhle nach anteroinferior und medial gerichteter Schlauch (daher auch *Tube*), der aufgrund seines sich rachenwärts (pharyngeal) trompetenförmig erweiternden Radius auch *Ohrtrumpete* genannt wird (vgl. Gelfand 2004, 47; Helmholtz 1896, 212; Reiss 2003, 611; Schünke et al. 2009, 131). Offen stellt die Eustachische Röhre eine druckausgleichende Verbindung zwischen Mittelohrhöhle und Rachenraum her, geschlossen dient sie

dem Mittelohr als Schutz vor Druckschwankungen im Rachenraum, die z.B. durch Sprechen, Atmung oder Husten erzeugt werden (vgl. Plattig in Keidel 1975, 48; Ranke 1953, 36f.; Reiss 2003, 591, 600; Schünke et al. 2009, 131; Stevens & Hallowell 1938, 249).

Neben ihrer Funktion als Luftschacht dient die Tube auch als Abfluß (Drainage). Überdies stellt ihre Verschließbarkeit gegenüber der notwendigen Verbindung zum Rachenraum eine wichtige Schutzinstanz gegen mögliche Infektionsübergriffe auf das Mittelohr dar (vgl. Di Martino et al. 2004, 1030, 1038; Didyk et al. 2007, 27; Lehnhardt & Laszig 2009, 27; Schünke et al. 2009, 131; Terhardt 1998, 56).

Die Eustachische Röhre teilt sich der Länge nach in zwei Abschnitte. Das laterale Drittel, direkt von der Paukenhöhle ausgehend, ist ein Knochenkanal im Felsenbein (vgl. Gelfand 2004, 47; 2009, 45; Helmholtz 1896, 212; Reiss 2003, 610; Schünke et al. 2009, 131), wobei es vom oberhalb verlaufenden Trommelfellspannerkanal nur durch eine dünne Knochenlamelle getrennt ist (vgl. Reiss 2003, 610). Richtung Rachen verengt sich der knöcherne Abschnitt trichterförmig. Der Übergang zum medialen Abschnitt (*Isthmus tubae auditivae*) bildet die engste Passage der Tube (vgl. Reiss 2003, 611; Gelfand 2004, 47; 2009, 45). Die rachenseitigen Zweidrittel der Tube bestehen aus einem flexiblen Knorpelmantel (*Pars cartilaginea*). Er umhüllt die Röhre nicht vollumfänglich, sondern ist im lateralen Bereich gespalten (vgl. Gelfand 2004, Abb. 2.10, 48; 2009, 45; Reiss 2003, 610; Helmholtz (1896, 212) verortet den Spalt im superioren Teil des Knorpels), was durch eine dünne Membran (*Lamina membranacea*) kompensiert wird (vgl. Helmholtz 1896, 212; Reiss 2003, 610; Schünke et al. 2009, 131). Diese Konstruktion ermöglicht der Muskulatur des Gaumensegels die beiläufige Öffnung der Röhre. Da der Knorpel (hakenartig) eingerollt ist sowie aufgrund des umliegenden Gewebedrucks und der Adhäsionskräfte der Schleimhaut, bleibt die Röhre im Ruhezustand geschlossen. Lediglich oberhalb des Hakens besteht eine konstante Öffnung, das sogenannte *Sicherheitsrohr* (vgl. Gelfand 2004, 47; 2009, 45; Di Martino et al. 2004, 1030; Reiss 2003, 610; Schünke et al. 2009, 131; Stevens & Hallowell 1938, 249).

Wie die Paukenhöhle ist auch die Eustachische Röhre mit einer schützenden Schleimhaut ausgekleidet. Ihre Flimmerhärchen sind schlundwärts gerichtet, was den Ausfluß aus der Paukenhöhle erleichtert und gleichzeitig Keimen das Eindringen erschwert, da diese der Tube sozusagen „gegen den Strich“ gehen (vgl. Reiss 2003, 611; Schünke et al. 2009, 131).

Die Gaumensegelmuskulatur besteht aus dem lateral der Röhre gelegenen *Gaumenspanner* (*Musculus tensor veli palatini*) und dem medialen *Gaumenheber* (*Musculus levator veli palatini*) sowie dem von der Tube nach inferior ziehenden Tuben-Rachenmuskel (*Musculus salpingopharyngeus*) (vgl. Reiss 2003, 611; Schünke et al. 2009, 131, 169, 181; Gelfand 2004, Abb. 2.10, 48; 2009, Abb. 2.13, 45). Der Gaumenspanner ist durch entsprechende Fasern des *Nervus mandibularis* (V₃, dritter Hauptast des V. Hirnnervs, *Nervus trigeminus*) innerviert, die restliche Gaumenmuskulatur durch den *Plexus pharyngeus*, ein Innervationsgeflecht bestehend aus Fasern des IX. (*Nervus glossopharyngeus*) und X. Hirnnervs (*Nervus vagus*) (vgl. Schünke et al. 2009, 114ff.).

Insgesamt ist die Eustachische Röhre 3,5 bis 4 cm lang (vgl. Gelfand 2004, 47; 2009, 45; Reiss 2003, 610). Am Engpaß (Isthmus) zwischen knöchernem und knorpeligem Abschnitt ist sie 1 bis 2 mm dünn, während die knöcherne Öffnung zur Paukenhöhle durchschnittlich 3 bis 6 mm mißt (vgl. Gelfand 2004, 47; 2009, 45).

2.3.3 Gehörknöchelchen

Das funktionelle Herzstück des Mittelohrs ist die Schalltransmission vom Außenohr zum Innenohr über die drei Gehörknochen *Hammer*, *Amboß* und *Steigbügel* (vgl. Gelfand 2004, 47; 2009, 42; Spitzer 2002, 57; Stevens & Hallowell 1938, 249, 259). Sie sind die kleinsten Knochen des menschlichen Körpers und bilden gemeinsam eine Übertragungskette, die lediglich mit Trommelfell und ovalem Fenster feste Verbindungen eingeht, ansonsten über Sehnen und Bänder vollkommen frei und flexibel in der Paukenhöhle gelagert ist (vgl. Gelfand 2004, 47f.; 2009, 42). Untereinander sind die Gehörknöchelchen über (trotz ihrer Winzigkeit) komplett ausgebildete Gelenke miteinander verbunden (vgl. Helmholtz 1896; 213ff.; Reiss 2003, 603f.; Schünke et al. 2009, 132).



Abb. 2.6 **Gehörknöchelchen** eines linken Ohrs von medial gesehen. (modifiziert n. Schünke et al. 2009, 132 g). Vom Hintergrund nach vorne: Hammer, Amboß, Steigbügel.

Hammer

Der *Hammer* (Malleus) ist – von außen nach innen gesehen – der erste Gehörknochen. Er besteht aus Kopf, Hals, Griff und zwei Fortsätzen (vgl. Gelfand 2004, 48; 2009, 42; Helmholtz 1896, 213ff.; Reiss 2003, 603). Er wiegt um die 23 bis 25 mg und ist mit 5 bis 6 mm, bzw. 8 bis 9 mm von Kopf bis Hammerspitze der längste der drei Gehörknochen (vgl. Békésy & Rosenblith 1951, Tab.1, 1079; Gelfand 2004, 49; 2009, 42). Der *Hammerkopf* (Caput mallei) ist näherungsweise kugelförmig und zugleich der massivste Teil des Knochens. Er ist in der Paukenhöhlenkuppel gelagert und trägt die Gelenksfläche für den dorsal ansetzenden Amboß (vgl. Gelfand 2004, 48; 2009, 42; Helmholtz 1896, 213ff.; Reiss 2003, 603). Darunter sitzt der schmale, kurze Hals des Hammers (*Hammerhals*, Collum mallei). Nach unten hin verbreitert er sich zum *Hammergriff* (Manubrium mallei), welcher – zum Hals um etwa 30° nach medial schräggestellt (vgl. Siebenmann 1897, 249) – sich zu seiner Spitze hin verjüngt. Direkt oberhalb des Hammergriffansatzes entspringt der,

dornartig nach anterior ragende, lange, *vordere Hammerfortsatz* (Processus mallearis anterior; nach Helmholtz (1896, 213) ist er beim Erwachsenen zu einem kurzen Stumpf abgetragen) (vgl. Gelfand 2004, 48f. 2009, 42; Reiss 2003, 603), der Ansatz des Hammergriffs selbst entwickelt sich nach lateral zum kurzen, *seitlichen Hammerfortsatz* (Processus mallearis lateralis) (vgl. Gelfand 2009, 42; Helmholtz 1896, 213f.; Reiss 2003, 603).

Der Kopf des Hammers ist durch das *obere Hammerband* (Ligamentum mallei superius) in der Paukenhöhlenkuppel befestigt (vgl. Gelfand 2004, 48; Helmholtz 1896, 215; Reiss 2003, 603; Siebenmann 1897, 282). Das *vordere Hammerband* (Ligamentum mallei anterius) befestigt den langen Hammerfortsatz axial nach anterior in der Glaserspalte (vgl. Gelfand 2004, 49; 2009, 42f. Reiss 2003, 603; Schünke et al. 2009, 133 C). Gemäß einiger älterer Beschreibungen fächert es sich nach lateral auf und verbindet so den langen Hammerfortsatz mit der Innenkante jenes Knochenvorsprungs, der als Spina tympanica major die Incisura tympanica anterior begrenzt (vgl. Helmholtz 1896, 214f.; Siebenmann

1897, 282). Das *seitliche Hammerband* (Ligamentum mallei laterale) verbindet den Hammerhals mit der lateralen Paukenwand oberhalb des Trommelfells. Hier verläuft es von der Spina tympanica major in einer bogenförmigen Ansatzlinie entlang der Incisura tympanica bis zur Spina tympanica minor (vgl. Gelfand 2009, 43; Helmholtz 1896, 215; Politzer 1896, 6; Reiss 2003, 603; Siebenmann 1897, 282; Schünke et al. 2009, 133 D). Die hintersten Fasern stehen offenbar in einem recht weiten Winkel zum vorderen Hammerband (die *Kernstränge* des vorderen und seitlichen Hammerbandes dürften eher einen Winkel nahe 90° bilden), so daß Helmholtz sich veranlaßt sah, sie als „Ligamentum mallei posterius“ (hinteres Hammerband) zu bezeichnen und durch beide Bandmassen die Rotationsachse des Hammers zu verlegen (vgl. Helmholtz 189, 215; Siebenmann 1897, 282f.). Tatsächlich verläuft diese anterior durch das vordere Hammerband, posterior jedoch durch den kurzen Amboßschenkel und dessen Aufhängung in der Paukenhöhle (vgl. Gelfand 2009, Abb. 2.51, 52; Reiss 2003, 604; Schünke et al. 2009, 132). Der Hammergriff schließlich ist seiner gesamten Länge nach mit dem Trommelfell verwachsen (vgl. Gelfand 2004, 44, 48; 2009, 40, 42; Hellbrück & Ellermeier 2004, 92; Helmholtz 1896, 214; Ranke 1953, 39; Reiss 2003, 597, 603). Seine Spitze bildet den Nabel des Trommelfellkonus, und der seitliche Hammerfortsatz zeichnet sich auf der lateralen Trommelfellseite als hervorstehender Knopf (*Prominentia mallearis*) ab (vgl. Gelfand 2004, 44, 48; 2009, 40, 42; Helmholtz 1896, 214; Reiss 2003, 597, 603).

In neuerer Zeit zeichnet sich ab, daß das obere Hammerband „keine eigentlichen ligamentösen Bestandteile“ (Siebenmann 1897, 282) enthält, sondern nur Teil der Schleimhautdecke des Hammers ist. Passend dazu konnten obere (superiore) Bänder z.B. bei Rennmäusen nicht nachgewiesen werden (vgl. Dirckx & Decraemer 2001, 135; Rosowski 2010, 52; Siebenmann 1897, 282). Gegen obere Befestigungsbänder spricht auch das Rotationsverhalten von Hammer und Amboß (siehe 3.2.1, S. 107). Nach Rosowski ist nur das vordere Hammerband tatsächlich bestätigt, was auch die Existenz eines seitlichen Hammerbands in Frage stellen würde (vgl. Rosowski 2010, 52).

Amboß

Der *Amboß* (Incus), mittlerer der drei Gehörknöchelchen, erinnert seiner Form nach eher an einen zweiwurzigen Backenzahn denn an einen Amboß. Die Kaufläche entspricht der Gelenkfläche Richtung Hammer, die beiden gespreizten Wurzeln bilden den *kurzen* bzw. *langen Amboßschenkel* (vgl. Gelfand 2004, 49; 2009, 42; Helmholtz 1896, 216; Reiss 2003, 603). Mit rund 27 bis 30 mg ist er der massivste unter den Gehörknochen (vgl. Békésy & Rosenblith 1951, Tab.I, 1079; Gelfand 2004, 49; 2009, 42). Über den kurzen Schenkel ist er etwa 5 mm, über den langen Schenkel 7 mm lang (vgl. Gelfand 2004, 49; 2009, 42). Der *Amboßkörper* (Corpus incudis) liegt in der Paukenhöhlenkuppel posterior des Hammerkopfs und trägt die sattelförmige Gelenkfläche für den Hammer (vgl. Gelfand 2004, 49; Helmholtz 1896, 216; Reiss 2003, 603). Der *kurze Amboßschenkel* (Crus breve incudis) gleicht eher einer nasenförmigen Verjüngung des nach posterior gestreckten Amboßkörpers. Der dünne, *lange Amboßschenkel* (Crus longum incudis) strebt dagegen, parallel zum Hammergriff, nach inferior. Die Spitze vollzieht einen Knick, wodurch das ihr aufsitzende Köpfchen für das Steigbügelgelenk nach medial gerichtet ist (vgl. Gelfand 2004, 49; 2009, 42; Helmholtz 1896, 216; Reiss 2003, 603). Zusammen mit dem etwas längeren Hammergriff bildet er einen effektiven Hebel. Das Armlängenverhältnis

bestimmt die Hebelwirkung, die meistens mit 1,3:1 angegeben wird (siehe 3.2.1, S. 110; vgl. Békésy 1941(a), 9; 1960, 95, 102; Békésy & Rosenblith 1951, 1082; Rosowski 2010, 60; Gelfand 2004, 94; 2009, 51; Reiss 2003, 604; Schünke et al. 2009, 132; Yost 2008, 76).

Der Körper des Amboß ist durch das *obere Amboßband* (Ligamentum incudis superius), posterior des Hammers, ebenfalls in der Paukenhöhlenkuppel aufgehängt (vgl. Reiss 2003, 603). Ihm haften bezüglich seiner Existenz prinzipiell dieselben Zweifel wie dem oberen Hammerband an (vgl. Dirckx & Decraemer 2001, 135; Rosowski 2010, 52). Der kurze Schenkel ist durch das *hintere Amboßband* (Ligamentum incudis posterius) direkt unterhalb der Warzenfortsatzhöhle an der hinteren Paukenhöhlenwand verankert (vgl. Gelfand 2004, 49; 2009, 43; Helmholtz 1896, 216; Reiss 2003, 603). Er stellt die posteriore Axialaufhängung des Hammer-Amboß-Gesamtkörpers dar.

Der Amboß liegt hinter dem Hammer, und zwar so, daß sein kurzer Fortsatz eine Achse mit dem langen Hammerfortsatz bildet und sein langer Schenkel parallel zum Hammergriff verläuft (vgl. Gelfand 2004, 49; 2009, 42; Reiss 2003, 603f.). Beide Knöchelchen sind insgesamt derart arrangiert, daß ihr gemeinsamer Masseschwerpunkt auf der durch vorderen Hammerfortsatz und hinteren Amboßschenkel nebst fibröser Aufhängung determinierten Rotationsachse liegt (vgl. Békésy 1941(a), 14; 1960, 102; Gelfand 2004, 93f.; Lehnhardt & Laszig 2009, 27f.; Reiss 2003, 604; Rosowski 2010, 65).

Das zugehörige Gelenk, eine großflächige aber enge Verbindung mit straffer und dennoch flexibler Gelenkkapsel, sorgt dafür, daß beide sich prinzipiell als Einheit bewegen, allerdings in gewissem Rahmen gegeneinander scheren können (vgl. Gelfand 2004, 49; 2009, 42; Reiss 2003, 604; Rosowski 2010, 67; Stevens & Hallowell 1938, 250). Dazu greifen die beiden sattelförmigen Gelenksflächen (vgl. Gelfand 2009, 42; Helmholtz 1896, 216) derart ineinander, daß bei Auswärtsbewegungen des Hammers (z.B. durch Druckschwankungen) ein gewisser Spielraum herrscht (vgl. Helmholtz 1896, 217; Rosowski 2010, 67). Schwingt das Trommelfell, drehen sich Hammer und Amboß mehr oder minder gemeinsam, wie ein zweiarmiger Hebel mit ungleichen Armen (vgl. Helmholtz 1896, 217; Reiss 2003, 604) um die oben beschriebene Achse.

Steigbügel

Das letzte, bzw. innerste Glied der Gehörknöchelchenkette ist der *Steigbügel* (Stapes), dessen Form durch seinen Namen bereits sehr treffend beschrieben ist. Er besteht aus zwei Schenkeln, die seinen Kopf mit seiner Fußplatte verbinden (vgl. Gelfand 2004, 49; 2009, 42; Helmholtz 1896, 218; Reiss 2003, 603). Mit der Fußplatte „steht“ der Steigbügel medial im ovalen Fenster. Die Schenkel liegen hintereinander, wobei der *vordere Steigbügelschenkel* (Crus stapedis anterior) etwas kürzer, dafür gerader als der *hintere Steigbügelschenkel* (Crus stapedis posterius) ist (vgl. Gelfand 2004, 49; Reiss 2003, 603). Lateral, an der Schenkelwurzel, sitzt der *Steigbügelkopf* (Caput stapedis), der seinerseits – ebenfalls lateral – die Gelenkspfanne für das Amboßgelenk trägt. Medial, an den gespreizten Enden der Schenkel, sitzt die *Steigbügelfußplatte* (Basis stapedis), eine ovale Knochenplatte deren Rand wulstartig verstärkt ist (vgl. Gelfand 2004, 49; 2009, 42; Helmholtz 1896, 218; Reiss 2003, 603). Im Zwischenraum von Schenkeln und Fußplatte (Obturator foramen (stapedis); vgl. Gelfand 2009, 42) ist die *Stapediusmembran* (Membrana stapedialis) aufgespannt (vgl. Gelfand 2004, 49).

Die Steigbügelfußplatte sitzt quasi paßgenau im ovalen Fenster des Innenohrs. Ein kleiner, um die Steigbügelplatte freibleibender Spalt, wird ringsherum durch das *Steigbügelband* (*Ringband*, Ligamentum anulare stapedis) abgedichtet, welches gleichzeitig den Steigbügel im ovalen Fenster befestigt (vgl. Békésy & Rosenblith 1951, 1077f.; Gelfand 2004, 49; 2009, 42; Reiss 2003, 603; Stevens & Hallowell 1938, 249). Da es anterior etwas weiter und flexibler ist, hat es nach Békésy Einfluß auf die Kolbenbewegung des Steigbügels im ovalen Fenster (vgl. Békésy 1960, 112, Ranke 1953, 39; Yost 2008, 78). Langer Amboßschenkel und Steigbügelkopf sind lateral über ein Kugelgelenk miteinander verbunden (vgl. Gelfand 2004, 49; 2009, 42; Helmholtz 1896, 218; Reiss 2003, 603; Stevens & Hallowell 1938, 250). Da sich der Amboß um seine Rotationsachse dreht, wird der Steigbügel nicht durch eine kolbenartige Druckübertragung, sondern eher eine Schaukelbewegung des Amboßschenkels angeregt. Weil der Steigbügelkopf dadurch auch vertikal ausgelenkt wird, kippt die Fußplatte mehr ins ovale Fenster, als daß sie hineingedrückt wird (vgl. Helmholtz 1896, 218; Reiss 2003, 604; Schünke et al. 2009, 132).

Der Steigbügel ist der kleinste Knochen im menschlichen Körper. Seine Fußplatte ist gerade einmal 3 mm breit und knapp 1,3 bis 1,5 mm (meist 1,4 mm) hoch, was ungefähr der meistens postulierten Fläche von 3,2 mm² entspricht (vgl. Békésy & Rosenblith 1951, Tab. 1, 1079, 1082; Gelfand 2004, 42, 49, 90; Hellbrück & Ellermeier 2004, 92; Helmholtz 1896, 219; Ranke 1953, 39; Schünke et al. 2009, 132, Stevens & Hallowell 1938, 260; Terhardt 1998, 55; Yost 2008, 76). Insgesamt variieren die Flächenangaben zwischen 2,2 und 3,5 mm² (vgl. Lehnhardt & Laszig 2009, 27; Müller & Frings 2009, 452). Vom Kopf bis zur Fußplatte ist er knapp 3,5 mm lang und wiegt ganze 2,5 bis 4 mg (vgl. Békésy & Rosenblith 1951, Tab. 1, 1079; Gelfand 2004, 49; 2009, 42). Das Steigbügelband ist am posterioren Ende der Fußplatte etwas länger als anterior. Maximal ist es knapp 0,1 mm und minimal etwa 15 µm weit (vgl. Békésy 1960, 197).

2.3.4 Mittelohrmuskeln

Bereits das Mittelohr greift aktiv in den Hörprozeß ein. Dies geschieht über zwei Muskeln, deren Spannung die Schwingungsfähigkeit der Gehörknöchelchen und damit die Schalltransmission beeinflusst (vgl. Gelfand 2004, 50; Müller & Werner 2998, 425; Spitzer 2002, 58; Terhardt 1998, 55). Der *Trommelfellspanner* (*Musculus tensor tympani*) zieht den Hammergriff in die Paukenhöhle und steuert dadurch die Spannung des Trommelfells. Der *Steigbügelmuskel* (*Stapediusmuskel*, *Musculus stapedius*) zieht das Amboß-Steigbügelgelenk aus seiner Bewegungsebene und beeinträchtigt so dessen Übertragungsfähigkeit (vgl. Gelfand 2004, 50; Helmholtz 1896, 215f.; Reiss 2003, 604f; Schünke et al. 2009, 133; Stevens & Hallowell 1938, 250).

Um Verzerrungen durch Muskelvibrationen vorzubeugen, sind beide Muskeln praktisch im Knochen versenkt und vom Klangraum Paukenhöhle entkoppelt (vgl. Békésy 1960, 114; Stevens & Hallowell 1938, 250; Yost 2008, 77). Der Kanal des Trommelfellspanners (*Semicanalculus muscoli tensoris tympani*) begleitet die Eustachische Röhre parallel und ist von dieser nur durch eine dünne Knochenlamelle getrennt. Seine Sehne greift über den *Löffelfortsatz* nach lateral das obere Ende des Hammergriffs an (vgl. Gelfand 2004, 49; 2009, 43; Reiss 2003, 604f.; Schünke et al. 2009, 133). Er ist durch den *Nervus muscoli tensoris tympani*, einen Ast des *Unterkiefernerve* (*Nervus mandibularis*, wiederum der dritte

Hauptast (V_3) des V. Hirnnervs, *Nervus trigeminus*), innerviert (vgl. Gelfand 2009, 43; Reiss 2003, 605; Russolo & Semeraro 1977, 377; Schünke et al. 2009, 104f., 107, 133, 215 C).

Der Steigbügelmuskel liegt in einem pyramidenartigen Vorsprung (*Eminentia pyramidalis*) verborgen, welcher an der Rückwand der Paukenhöhle, ungefähr auf Höhe des ovalen Fensters sitzt (vgl. Gelfand 2004, 46, 49; 2009, 42f.; Reiss 2003, 600; Schünke et al. 2009, 130ff.). Von der Pyramiden spitze aus greift die Sehne des Muskels den Steigbügelkopf an (vgl. Gelfand 2004, 49; 2009, 43; Reiss 2003, 605; Schünke et al. 2009, 133). Der Steigbügelmuskel ist durch den *Nervus stapedius*, einen Ast des *Fazialisnervs* (VII. Hirnnerv, *Nervus facialis*) innerviert (vgl. Gelfand 2009, 43; Reiss 2003, 605; Schünke et al. 2009, 108ff., 133).

Der Vollständigkeit halber sei noch erwähnt, daß über die Möglichkeit eines dritten Muskelsystems im Mittelohr diskutiert wird. Derartige Überlegungen fußen auf Funden vermeintlichen Muskelgewebes im *Paukenring*, das offenbar eng mit den radialen Fasern des Trommelfells verknüpft ist (vgl. Rosowski 2010, 54).

Wie die meisten Komponenten des Mittelohrs, sind auch die Muskeln die kleinsten ihrer Art im menschlichen Körper. Der Trommelfellspanner ist nur ungefähr 25 mm lang und weist einen Querschnitt von rund 5,85 mm² auf. Mit einer Länge von knapp 6,3 mm und einem mittleren Querschnitt von ungefähr 4,9 mm², ist der Steigbügelmuskel sogar noch wesentlich kleiner (vgl. Wever & Lawrence 1954 zit. nach Gelfand 2004, 49; 2009, 43).

2.4 Innenohr

Jenseits des ovalen Fensters liegt schließlich das *Innenohr* (*Auris interna*). Das winzige, kaum 2 cm lange Gebilde, erstreckt sich in der Felsenbeinpyramide mehr oder weniger parallel zu deren Längsachse. Im Horizontalschnitt ist es von posterolateral nach anteromedial in einem Winkel von 45° zur Sagittalachse ausgerichtet, im Sagittalschnitt liegt es annähernd horizontal (vgl. Schünke et al. 2009, 134 a & b).

Die Aufgabe des Innenohrs als *Schallempfindungsapparat* ist im Transduktionsprozeß zu suchen (siehe 2, S. 49; vgl. Reiss 2003, 591; Gelfand 2004, 121; Graumann & Sasse 2005, 90; Guski 1996, 91; Spitzer 2002, 58; Terhardt 1998, 56; Yost 2008, 83), wobei dies nicht nur für das Hören gilt. Das Innenohr enthält mit dem *Hörorgan* (*Organum cochleare*) und dem *Gleichgewichtsorgan* (*Organum vestibulare*) zwei unabhängig voneinander funktionierende Organe (vgl. Gelfand 2004, 34; Graumann & Sasse 2005, 110; Reiss 2003, 591; Schünke et al. 2009, 134; Terhardt 1998, 56), die aufgrund ihrer Lage und anatomischen Homogenität begrifflich auch zum *Organum vestibulocochleare* zusammengefaßt werden (vgl. Reiss 2003, 591). Da der Gleichgewichtssinn nicht Teil dieser Arbeit ist, werden dessen Strukturen größtenteils vernachlässigt und im Folgenden nur am Rande betrachtet (für detaillierte Darstellungen und Ausführungen vgl. Graumann & Sasse 2005, 125–138; Müller & Frings 2009, 434–441; Reiss 2003, 615–618; Schünke et al. 2009, 134f., 138f., 440f.).

2.4.1 Labyrinth

Aufgrund seiner komplizierten Figur wird das Innenohr auch *Labyrinth* genannt (vgl. Helmholtz 1896, 221; Reiss 2003, 591; Siebenmann 1897, 297f.; Spitzer 2002, 58; Stevens & Hallowell 1938, 268). Es setzt sich grob aus *Vorhof*, *Schnecke* und den drei *Bogengängen* zusammen (vgl. Gelfand 2004, 50f.; Graumann & Sasse 2005, 110; Helmholtz 1896; 220; Reiss 2003, 612; Schünke et al. 2009, 134; Yost 2008, 83). Der Vorhof zählt (als Namensgeber¹²) ebenso wie die Bogengänge zum Gleichgewichtsorgan, während die Schnecke das Hörorgan beheimatet (vgl. Graumann & Sasse 2005, 110; Reiss 2003, 615; Schünke et al. 2009, 134; Spitzer 2002, 58; Yost 2008, 83).

Das Labyrinth ist innerlich ein membranöses Gebilde, im Felsenbein einem äußeren „Abdruck“ seiner selbst eingelagert. Dieser äußere Hohlraumkomplex im Felsenbein wird *knöchernes Labyrinth* genannt. Das darin gelagerte membranöse System wird entsprechend als *häutiges Labyrinth* bezeichnet. Es ist mit *Endolympe* gefüllt, von *Perilymphe* umgeben (vgl. Gelfand 2004, 51; 2009, 50, 54; Hellbrück & Ellermeier 2004, 93; Dieffenbach et al. 1837, 53; Reiss 2003, 591, 612; Schünke et al. 2009, 134; Siebenmann 1897, 298; Stevens & Hallowell 1938, 268, 270) und enthält die tatsächlichen Sinnesorgane des Hörens, bzw. des Gleichgewichts (vgl. Gelfand 2004, 51; Reiss 2003, 612; Schünke et al. 2009, 134). Obwohl häutiges und knöchernes Labyrinth im Detail bauliche Unterschiede aufweisen, prägen die Strukturen des häutigen Labyrinths „bereits während der Embryonalentwicklung die Form des Felsenbeins“ (Reiss 2003, 615) und fungieren somit als Vorlage für die Entwicklung des knöchernen Labyrinths (vgl. Graumann & Sasse 2005, 91).

Knöchernes Labyrinth

Das *knöcherne Labyrinth* (*Labyrinthus osseus*) besteht aus einer zentralen Höhlung, dem *Vorhof*, welchem superoposterior drei Kanäle, die *knöchernen Bogengänge*, sowie anteromedial die Schnecke entspringen (vgl. Gelfand 2004, 50f.; Graumann & Sasse 2005, 125; Helmholtz 1896; 220; Reiss 2003, 612; Schünke et al. 2009, 134; Spee 1896, 190f.; Stevens & Hallowell 1938, 268f.). Hinzu kommen zwei dünne Kanälchen im Felsenbein, die das Labyrinth in inferiorer bzw. posteromedialer Richtung verlassen (vgl. Fanghänel et al. 2003(b), 199; Reiss 2003, 615; Schünke et al. 2009, 134). Sie sind am Abfluß der Lympflüssigkeiten beteiligt (vgl. Graumann & Sasse 2005, 129; Schünke et al. 2009, 134).

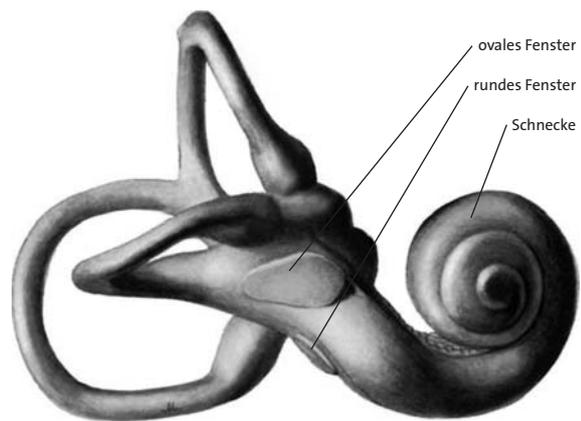


Abb. 2.7 **knöchernes Labyrinth** eines rechten Ohrs von lateral gesehen (modifiziert n. Reiss 2003, Abb. 7.30, 624).

Die zentrale Kammer des knöchernen Labyrinths, den *Vorhof* (Vestibulum), könnte man als Schnittstelle des Mikrosystems Innenohr bezeichnen (vgl.

Yost 2008, 83). Über ihn stehen die Lymphräume von Gehör- und Gleichgewichtsorgan untereinander

¹² Dem *Vorhof* (Vestibulum) hat das Gleichgewichtsorgan (Organum vestibulare, Vestibularapparat, bzw. -organ) seinen lateinischen Namen zu verdanken (vgl. Reiss 2003, 612).

und mit umliegenden Flüssigkeitsräumen (zerebrospinaler Raum) in Verbindung (vgl. Graumann & Sasse 2005, 129; Helmholtz 189, 221f.; vgl. Reiss 2003, 612; Schünke et al. 2009, 134).

Der Vorhof ist ein kleiner, annähernd ellipsoider Raum (vgl. Graumann & Sasse 2005, 127; Reiss 2003, 612; Spee 1896, 190, 196). Seine mediale Wand ist durch einen von anterosuperior nach postero-inferior verlaufenden Grat, die *Crista vestibuli*, in zwei Ausbuchtungen, den oberhalb (anteroinferior) des Grats gelegenen *Recessus sphaericus* und den unterhalb (posteriosuperior) gelegenen *Recessus ellipticus*, geteilt. Im Bereich beider Ausbuchtungen ist die mediale Wand siebartig durchlöchert, um Fasern des Gleichgewichtsnerfs Durchschlupf zu gewähren (vgl. Graumann & Sasse 2005, 127; Reiss 2003, 612; Spee 1896, 198f.). Ohne auf Details einzugehen sei zudem erwähnt, daß am *Recessus ellipticus* die – gemäß ihrer Lage als *vorderer*, *hinterer* bzw. *seitlicher Kanal* (*Canalis semicircularis anterior*, *posterior* bzw. *lateralis*) bezeichneten – knöchernen Kanäle der Bogengänge entspringen und nach ungefähr C-förmigem Verlauf wieder eintreffen, wobei vorderer und hinterer Kanal einseitig einen gemeinsamen Schenkel (*Crus osseum commune*) bilden (vgl. Graumann & Sasse 2005, Abb. 1–72/B, 111, 128; Reiss 2003, 612f.; Schünke et al. 2009, 134; Spee 1896, 190, 196).

Medial neben den Schenkeln der Bogengänge verläßt der *Aquaeductus vestibuli* (*Canaliculus vestibuli*) den Vorhof. Es verläuft zunächst parallel zum *Crus commune* nach posterior, verjüngt sich unterwegs, wendet dann an seinem engsten Punkt (*Isthmus*) in einem Winkel von 90 bis 135° nach inferomedial, erweitert sich jetzt trichterförmig und tritt schließlich durch eine unscheinbare Öffnung (*Apertura canaliculi vestibuli*) an der dorsalen Außenwand des Felsenbeins aus (vgl. Fanghänel et al. 2003(b), 199; Graumann & Sasse 2005, 127, 129; Helmholtz 1896, 221f.; Reiss 2003; 615; Schünke et al. 2009, 134; Spee 1896, 200).

Nach antero-inferior prägt sich mit dem *Recessus cochlearis* eine weitere Vertiefung des Vorhofs ins Felsenbein. Superior dieser Ausbeulung, gegenüber der *Crista vestibuli*, trägt die laterale Wand des Vorhofs das ovale Fenster. Nach anterior entwickelt sich aus dem *Recessus cochlearis* die *Vorhoftreppe* der Schnecke, wobei der Boden des *Recessus cochlearis* ebenfalls in die Schnecke zieht und sich dort zur „Schnecken-trennwand“ (*Lamina spiralis*) (quasi dem „Boden“ der Vorhoftreppe) wandelt (vgl. Graumann & Sasse 2005, 128; Reiss 2003, 612; Spee 1896, 197). Der rückläufige Kanal (sozusagen die „Gegenfahrbahn“ zur Vorhoftreppe), die *Paukentreppe*, liegt „unterhalb“ der *Lamina spiralis ossea* und endet unterhalb des *Recessus vestibuli* mit dem *runden Fenster* (vgl. Graumann & Sasse 2005, 113; Reiss 2003, 612).

Unweit des runden Fensters zieht der *Perilymphgang* (*Ductus perilymphaticus*) durch das *Aquaeductus cochleae* (*Canaliculus cochleae*) nach kaudal und verbindet den Perilymphraum des knöchernen Labyrinths mit dem Subarachnoidalraum¹³ (vgl. Fanghänel et al. 2003(b), 199; Graumann & Sasse 2005, 129; Ranke 1953, 64; Reiss 2003; 615; Schünke et al. 2009, 134).

Die Gesamtlänge des Labyrinths über die äußersten Scheitelpunkte von Schnecke und hinterem Bogengang beträgt knapp 17,5 bis 18 mm (vgl. Spee 1896, 191). Der Durchmesser des Vorhofs wird, je nach Achse, zwischen 3 und 7 mm angegeben (vgl. Gelfand 2004, 50f.; Graumann & Sasse 2005, 127; Reiss 2003, 612; Spee 1896, 196; Yost 2008, 83). Basierend auf den Angaben von Graumann/Sasse und Reiss (6 × 4 × 6 mm) kommt der Vorhof auf ein Volumen von knapp 1,4 cm³. Die Schnecke hat ungefähr einen basalen Durchmesser von 0,9 bis 1 cm und eine Höhe (sozusagen eine *Modioluslänge*) von rund 5 mm (vgl. Gelfand 2004, 51, 53; 2004, 55; 2009, 55; Terhardt 1998, 56).

13 Raum zwischen der mittleren Schicht der Hirnhaut (Spinnwebhaut, Arachnoidea) und der inneren Schicht (*Pia mater*). Dieser Zwischenraum (Subarachnoidalraum) ist von Gefäßen und Nerven durchzogen und ansonsten mit Hirnwasser (*Liquor cerebrospinalis*) gefüllt (für detailliertere Beschreibungen vgl. Schünke et al. 2009, 258, 263 C; Trepel 2008, 289).

Häutiges Labyrinth

Das *häutige Labyrinth* (Labyrinthus membranaceus) formiert sich zentral aus zwei Bläschen, dem *Utriculus* (kleiner Schlauch) und dem *Sacculus* (Säckchen), die beide vom knöchernen Vorhof umschlossen werden. Analog zur Knochenhöhle entspringen dem Utriculus drei Schläuche, die *häutigen Bogengänge* (Ductus semicirculares). Anteromedial vor dem Sacculus verläuft, seinem knöchernen Kanal gemäß, der Schneckengang (vgl. Helmholtz 1896, 222f.; Reiss 2003, 591; Schünke et al. 2009, 134; Stevens & Hallowell 1938, 271; Yost 2008, 83). Alle Strukturen des häutigen Labyrinths sind durch winzige Kanäle miteinander verbunden, mit *Endolymphe* gefüllt und von *Perilymphe* umgeben (vgl. Gelfand 2004, 51; 2009, 54f.; Graumann & Sasse 2005, 111; Dieffenbach et al. 1837, 53; Reiss 2003, 591, 612; Schünke et al. 2009, 134).

Der längliche Utriculus des häutigen Labyrinths liegt schräg (superoventral nach inferodorsal) im hinteren Bereich des Vorhofs und verursacht den Recessus ellipticus. Ihm gehen nach dorsal die drei *häutigen Bogengänge* (Ducti semicirculares) ab. Analog zu den Knochenkanälen nehmen sie einen C-förmigen Verlauf (vgl. Helmholtz 1896, 223; Reiss 2003, 612, 616f.; Schünke et al. 2009, 134).

Anterior des Utriculus liegt der kleinere, flachere Sacculus im Recessus sphaericus des Vorhofs (vgl. Helmholtz 1896, 223; Reiss 2003, 612, 616f.; Schünke et al. 2009, 134). Sacculus und Utriculus enthalten mit den *Maculae (utriculi* bzw. *sacculi*) die auf Linearbeschleunigung spezialisierten Sinnesorgane (vgl. Graumann & Sasse 2005, 133; Reiss 2003, 618; Schünke et al. 2009, 138f.).

Beide Vorhofsäckchen sind über ein Schlauchsystem miteinander verbunden. Die vom entsprechenden Säckchen ausgehenden *Ductus saccularis* und *Ductus utricularis* vereinen sich zum *Endolymphgang* (Ductus endolymphaticus) (vgl. Graumann & Sasse 2005, 129; Reiss 2003, 615; Schünke et al. 2009, 134). Demnach führt er zunächst parallel zum Crus commune nach dorsal, verjüngt sich auf diesem Abschnitt, wendet an seiner engsten Stelle nach inferomedial und erweitert sich trichterförmig und tritt als *Endolymphsack* (Saccus endolymphaticus) aus dem Felsenbein. Der *Endolymphsack* ist eine an der rückwärtigen Felsenbeinaußenwand gelegene Tasche, in der die Endolymphe resorbiert wird (vgl. Gelfand 2009, 55; Graumann & Sasse 2005, 129; Reiss 2003, 615, 618; Schünke et al. 2009, 134).

Über einen dünnen Schlauch, den *Ductus reuniens*, sind Gleichgewichtsanteil und Höranteil des Endolymphraums, genaugenommen Sacculus und Schneckengang, miteinander verbunden (vgl. Gelfand 2004, 53; 2009, 54f.; Graumann & Sasse 2005, 129; Helmholtz 1896, 223, 224; Reiss 2003, 615, 617; Schünke et al. 2009, 134).

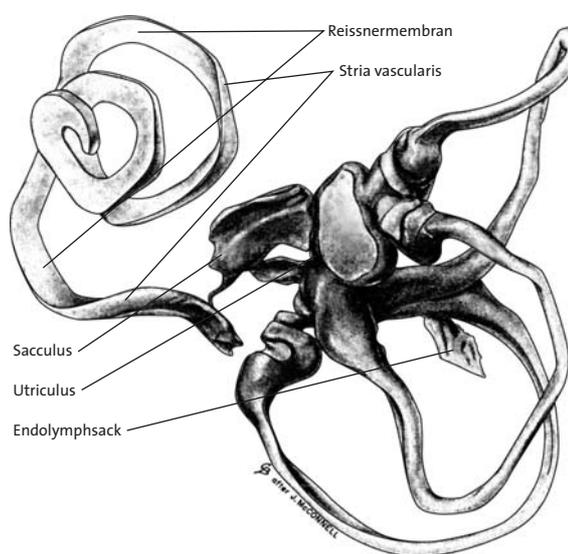


Abb. 2.8 **häutiges Labyrinth** eines linken Ohrs von lateral gesehen (modifiziert n. Gelfand 2004, Abb. 2.13, 51).

Nach Schünke et alii (vgl. 2009, 134) „schwimmt“ das häutige quasi im knöchernen Labyrinth, was allerdings nur zum Teil richtig sein dürfte. Zwischen Utriculus und Sacculus ist der Perilymphraum durch eine feine Grenzmembran (*Membrana limitans*) geteilt (vgl. Reiss 2003, 623). Posterior dieser Membran, bei den Bogengängen und dem Utriculus, ist er von einem Netzwerk zarter Bindegewebsfasern und Zellfortsätzen durchzogen. Dadurch sind die Strukturen des häutigen Labyrinths weich in den Perilymphraum eingelagert. Anterior der Grenzmembran ist das häutige Labyrinth nicht derartig gestützt, sondern, speziell in der Schnecke, direkt im knöchernen Kanal aufgespannt (vgl. Graumann & Sasse 2005, 129; Reiss 2003, 623).

2.4.2 Schnecke

Die *Schnecke* (Cochlea) ist nach ihren tierischen Namensvettern benannt, an deren Behausungen ihre Form erinnert (vgl. Békésy & Rosenblith 1951, 1076; Helmholtz 1896, 223; Stevens & Hallowell 1938, 269). Besonders deutlich erkennbar ist dies bei Abgüssen des knöchernen Labyrinths (siehe Abb. 2.7, S. 68; vgl. Helmholtz 1896, 222f.; Reiss 2003, 624).

Schneckenabgüsse sind Negativabdrücke des knöchernen *Schneckenkanals*. Er verjüngt sich vom Vorhof zur Spitze (Apex) hin kontinuierlich und windet sich in ungefähr zweieinhalb bis zweidreiviertel Wicklungen spiralförmig – eben wie ein Schneckenhaus – um eine zentrale Achse (vgl. Gelfand 2004, 53; 2009, 55; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003, 613f.; Spee 1896, 201; Spitzer 2002, 58; Stevens & Hallowell 1938, 273; Winckel 1960, 80; Yost 2008, 84f.; Zwicker & Feldtkeller 1967, 84). Der Länge nach ist er durch zwei Membranen dreigeteilt (vgl. Fuchs 2010, 4; Gelfand 2004, 54). Der zwischen beiden Membranen liegende *Schneckengang* ist als Teil des häutigen Labyrinths mit Endolymphe gefüllt, von Perilymphe umgeben und beheimatet mit dem Cortiorgan das eigentliche Hörorgan (vgl. Fuchs 2010, 6; Gelfand 2004, 51ff.; 2009, 53f.; Graumann & Sasse 2005, 111; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003, 591; Terhardt 1998, 57; Yost 2008, 83, 86). In der *Schneckenspitze* (Apex) enden beide Membranen etwas verfrüht, beschließen den Schneckengang und lassen eine kleine Verbindung zwischen den außen liegenden Treppen, das *Schneckenloch*, frei (vgl. Békésy 1941b, 265; Gelfand 2004, 54f.; 2009, 54; Graumann & Sasse 2005, 113; Helmholtz 1896, 223, 225; Ranke 1953, 64; Reiss 2003, 614; Stevens & Hallowell 1938, 269; Terhardt 1998, 57; Yost 2008, 85). Oberhalb des Schneckengangs verläuft die dem Vorhof entspringende *Vorhof-treppe*, unterhalb die am *runden Fenster* schließende *Paukentreppe* (vgl. Gelfand 2004, 55; 2009, 54; Helmholtz 1896, 224; Reiss 2003, 614, 623; Yost 2008, 85f.).

Der Schneckenkanal zieht vom Vorhof nach anteromedial und wickelt sich in superolateraler Richtung auf. Nach Reiss (vgl. 2003, 613) wickelt sich die Schnecke rechts mit und links gegen den Uhrzeigersinn. Reiss muß dies aus medialer Perspektive festgestellt haben (siehe Abb. 2.1, S. 50 und Abb. 2.7, S. 68). Die Achse der Wicklung richtet sich, mehr oder weniger parallel zum vorderen Bogengang, in einem Winkel von etwa 45° an Horizontal- und Sagittalachse aus, so daß die Schneckenspitze (Cupula cochleae) nach anterolateral zeigt (vgl. Schünke et al. 2009, 134).

Exkurs: Zur dargestellten und zur tatsächlichen Ausrichtung der Schnecke

Traditionellerweise gehen anatomische Beschreibungen der Schnecke von einem senkrecht gestellten Modiolus aus. Daher sind auch die meisten Querschnitte durch die Schnecke mit dem Apex bzw. der Vorhoftreppe oben und der Basis bzw. der Paukentreppe unten abgebildet (vgl. u.a. Gelfand 2004, 56, 60; Graumann & Sasse 2005, Abb. 1–74, 113; Helmholtz 1896, Abb. 46., 225; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, Abb. 5.1, 35; Reiss 2003, 619; Reiss et al. 1989; Schünke et al. 2009, 136; Siebenmann 1897, 305; Spee 1896, Abb. 44, 114; Stevens & Hallowell 1938, 271; Yost 2008, 86). Diese Darstellung zeigt nur die relativen Lagen innerhalb der Schnecke ohne übergeordnetes Bezugssystem. Sie ist eigentlich nur deshalb eine sinnvolle Darstellung, weil sie sich mittlerweile in den meisten Köpfen festgesetzt hat. Nimmt man den Schädel als Bezugssystem, sind die tatsächlichen Lageverhältnisse im Innenohr jedoch andere.

Zu Beginn, d.h. am Vorhof, wo die Vorhoftreppe dem Recessus cochlearis entspringt und die Paukentreppe inferior des Vorhofs endet, liegen beide Treppen tatsächlich noch übereinander (vgl. Siebenmann 1897, 305; Spee 1896, 207). Im weiteren Verlauf, bevor der Schneckenkanal sich spiralig einrollt, verdreht er sich um seine eigene Achse, so daß die Vorhoftreppe nach inferolateral und die Paukentreppe nach superomedial rutscht, wodurch auch der Schneckengang nicht mehr waagrecht zwischen beiden Perilymphschläuchen liegt, sondern sich eher senkrecht, von außen, zwischen beide quetscht (vgl. Siebenmann 1897, 305f.; Spee 1896, 207f.).

Die „obere“ Vorhoftreppe (vgl. u.a. Gelfand 2004, 55; Yost 2008, 85) ist in der Schneckenspirale also nur in senkrechter Darstellung die obere (vgl. Reiss 2003, 614). Tatsächlich ist sie die äußere, zur anterolateralen Außenfläche der Felsenbeinpyramide orientiert. Dementsprechend liegt die Paukentreppe in der Schnecke innen, zur „Hinterseite“ der Pyramide (bzw. zur Schneckenbasis, zum inneren Gehörgang) hin gelegen (vgl. Helmholtz 1896, Abb. 45, 224; Siebenmann 1897, 305).

Knöcherner Schnecke

Der gewundene *Schneckenkanal* (Canalis cochlearis) bildet im Felsenbein einen Hohlraum, in dessen Mitte ein poröser, spindelartiger Knochenkegel, der *Modiolus* (Spindel) zurückbleibt (vgl. Gelfand 2004, 53; 2009, 55; Reiss 2003, 614; Yost 2008, 85; Spee 1896, 201; Stevens & Hallowell 1938, 269). Die Längsachse des Modiolus entspricht der Wicklungsachse des Kanals, wobei dessen Windungen nicht exakt konzentrisch verlaufen, sondern ein wenig nach inferior versetzt sind (vgl. Gelfand 2004, 53; 2009, 55; Graumann & Sasse 2005, 111f.; Reiss 2003, 614). Gleichzeitig liegen sie nicht exakt nebeneinander. Mit jeder Windung verkürzt sich der Krümmungsradius stark, während sich die Spiralhöhe nur sehr zögerlich nach (antero-)lateral entwickelt. Dies hat zur Folge, daß der Anfang der zweiten Windung an der konkaven Seite der ersten und wenigstens seine Paukentreppe noch ganz auf der Ebene der Basalwindung liegen. Im Verlauf der Mittelwindung dehnt sich die Spirale am stärksten lateralwärts, während sie dies in der Spitzenwindung nicht mehr tut und ihre letzte Halbwindung an die konkave Seite der dorsalen Hälfte der Mittelwindung legt (vgl. Spee 1896, 208).

Seine Spindelform gewinnt der Modiolus nicht zuletzt durch zwei Knochenlamellen, die ihn – vergleichbar mit dem Gewinde einer Schraube – normal zu seiner Längsachse umlaufen. Die größere

davon, das *Septum cochleae*, wirkt eigentlich nur bei abgetragener Außenwand (der Schnecke) wie eine Lamelle, da sie den Schneckenhohlraum gewöhnlich komplett durchtrennt und auf diese Weise die einzelnen Schneckenwindungen voneinander abgrenzt (vgl. Gelfand 2004, 53; 2009, 55; Graumann & Sasse 2005, 112; Helmholtz 1896, 223f.; Reiss 2003, 614; Spee 1896; 201). Folgt man dem Verlauf des Schneckenkanals, entsteht sie aus der basalen Wand der ersten Windung. Ungeachtet der Schneckenlage im Kopf, könnte man sie als „Boden“ der Basalwindung bezeichnen. Sobald diese voll um den Modiolus herumgelaufen ist und sich als Mittelwindung (lateral versetzt) an die basale legt, wird der „Boden“ der Mittelwindung gleichzeitig „Dach“ der Basalwindung (vgl. Siebenmann 1897, 201). Am Ende der letzten Halbwindung (Spitzenwindung) wendet sich das Septum zunächst nach lateral, stellt sich also von ventral einem weiteren Kanalverlauf in den Weg und umkapselt den Kanal nach superior soweit, bis es in der Längsachse des Modiolus endet. Diese freie Septumendkante stellt sozusagen den Modiolus der Spitzenwindung dar (vgl. ders 1897, 204f.). Aufgrund des sich verkürzenden Radius der Windungen steht das Septum nicht tatsächlich normal zur Längsachse des Modiolus, sondern schräg Richtung Apex, also nach lateral gekippt. Gleichzeitig ist es wegen der Windungsverlagerung nach inferior in der superioren Hälfte dicker.

Die kleinere Lamelle, die *knöchernen Trennwand* (Lamina spiralis ossea), entsteht aus dem Boden des Recessus cochlearis (vgl. Graumann & Sasse 2005, 128; Spee 1896, 197, 207). Sie zieht auf ungefähr mittlerer Höhe in den Schneckenkanal. Während sie „an der Basis fast die ganze Kanalbreite einnimmt, läßt sie zur Spitze hin einen Spalt zunehmender Breite offen“ (Terhardt 1998, 57f.; vgl. Zwicker & Feldtkeller 1967, 84). Sie ragt demnach quasi wie eine frei tragende Wendeltreppe in den Schneckenkanal hinein (vgl. Gelfand 2004, 53; 2009, 55; Graumann & Sasse 2005, 112; Helmholtz 1896, 223f.; Reiss 2003, 614). Apikal endet die Trennwand etwas früher als der Schneckenkanal. Ihr axial freiliegendes Ende verläuft sichel- oder hakenförmig (*Hamulus*) um die (nun den Modiolus bildende) Kante des Septums herum (vgl. Gelfand 2004, Abb. 2.16, 54; Helmholtz 1896, 224; Reiss 2003, 614; Siebenmann 1897, 206f.; Spee 1896, 207f.). Durch diese Hakenform winden sich beide Knochenblätter (Septum und knöchernen Trennwand) konkav umeinander und lassen so in ihrer Mitte das *Schneckenloch* (Helicotrema) frei (vgl. Helmholtz 1896, 223f.; Siebenmann 1897, 206f.).

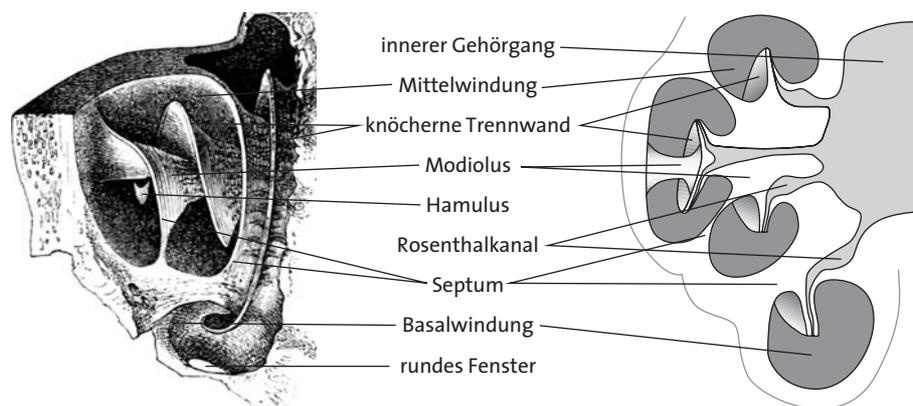


Abb. 2.9 **knöchernen Schnecke** eines rechten Ohres von vorne gesehen. Links: der intakte Modiolus mit dem Spiralblatt (Lamina spiralis), das Septum erscheint nur bei abgetragener Außenwand als freie Lamelle (modifiziert n. Helmholtz 1896, Abb. 45, 224). Rechts: Querschnitt durch den Modiolus (modifiziert n. Gelfand 2004, Abb. 2.14, 53). Gut zu erkennen ist die tatsächliche Lage der Schnecke.

Genaugenommen besteht die knöcherne Trennwand aus zwei dünnen Knochenblättern (Gelfand 2004, 55; 2009, 55; Graumann & Sasse 2005, 113; Helmholtz 1896, 228; Reiss 2003, 614; Spee 1896, 205). Feine Knochenstege stemmen beide Blätter voneinander ab und lassen so radial verlaufende Kanäle und Öffnungen (*Habenulae perforata*) frei, die den Fasern des Hörnervs Zugang zum Schneckenkanal gewähren. Auch der Modiolus ist für den Faserdurchlauf mit feinen Kanälen durchzogen und basal, für den Austritt des Hörnervs in den *inneren Gehörgang* (Meatus acusticus internus), siebartig perforiert (vgl. Gelfand 2004, 53, 55; 2009, 55f.; Helmholtz 1896, 227f.; Reiss 2003, 614f. Spee 1896, 204f.). An der Wurzel der knöchernen Trennwand verläuft im Modiolus der *Rosenthalkanal* (Canalis spiralis modioli), wo die Neurone des Spiralganglions liegen (vgl. Gelfand 2004, Abb. 2.14, 53, 66; Reiss 2003, Abb. 7.33, 626; Spee 1896, 205; Stevens & Hallowell 1938, 269).

Die knöcherne Trennwand schneidet das Lumen des Schneckenkanals ein und deutet dessen Spaltung in zwei separierte Treppen an (vgl. Békésy 1941b; 265; Graumann & Sasse 2005, 110, 112; Helmholtz 1896, 223f.; Reiss 2003, 614; Yost 2008, 85). In der Basalwindung wird sie dabei durch ein, ihr an der äußeren Schneckenwand gegenüberliegendes, zunächst breiteres, rasch dünner werdendes Knochenblättchen (*Lamina spiralis secunda*) unterstützt (vgl. Graumann & Sasse 2005, 112; Spee 1896, 207). Die „obere“ Treppe zieht vom ovalen Fenster aus dem Recessus cochlearis des Vorhofs in die Schnecke. Dementsprechend wird sie als *Vorhof-treppe* (Scala vestibuli) bezeichnet. In der Schneckenspitze kehrt sie über das Schneckenloch um und läuft nun „unter“ der Trennwand zurück bis zum runden Fenster inferior des Vorhofs, welches sie von der Paukenhöhle trennt. Entsprechend wird dieser Abschnitt als *Pauken-treppe* (Scala tympani) bezeichnet (vgl. Gelfand 2004, 54f.; 2009, 54; Graumann & Sasse 2005, 113; Helmholtz 1896, 224; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003, 614, 623; Yost 2008, 85f.). Am blinden Ende der Paukentreppe, unweit des runden Fensters liegt mit dem (membranös verschlossenen) *Schneckenkanälchen* quasi der „Abfluß“ der Perilymphe aus dem Innenohr (vgl. Gelfand 2004, 52; Graumann & Sasse 2005, 113, 129; Schriever 1967, 36).

Die Länge des Schneckenkanals wird zwar mit 30 bis 35 bis 38 mm sehr variabel angegeben (vgl. Békésy & Rosenblith 1951, Tab. 1, 1079; Graumann & Sasse 2005, 111; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003,

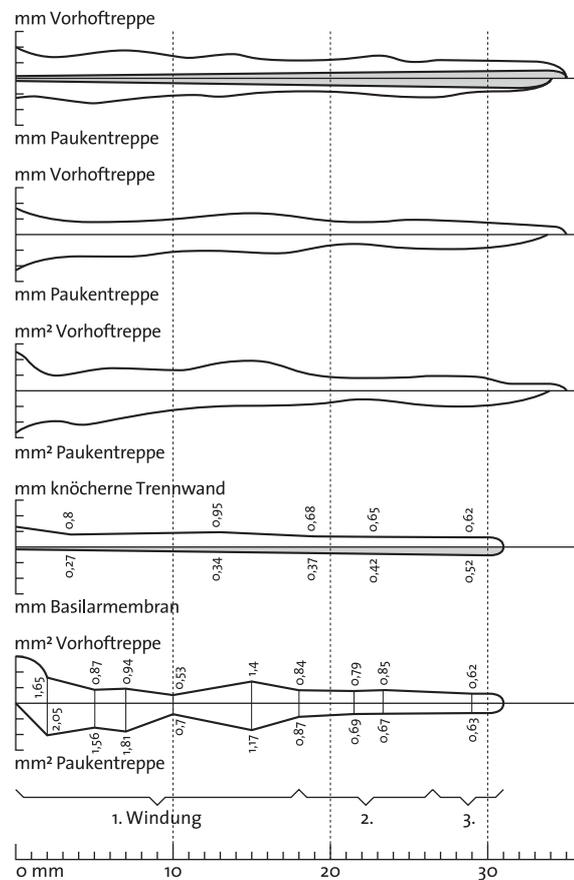


Abb. 2.10 **Breitenverhältnisse der Schnecke** (modifiziert n. Ranke 1953, Abb. 49, 65; Yost 2008, 74, 87). Von oben nach unten: Kanalbreiten im Verhältnis zur BM (grau), Kanalhöhen, approximative Kanalquerschnitte (Ranke, nach Zwislocki), Breite der Schneckentrennwand, Kanalquerschnitte, Schneckenwindungen (Yost, Daten von Wrightson und Keith).

613; Schünke et al. 2009, 136; Siebenmann 1897, 306), dennoch kann man von einer ungefähren Länge von knapp 35 mm ausgehen (vgl. Békésy 1960, 197; Gelfand 2004, 53; 2009, 55; Guski 1996, 91; Terhardt 1998, 57; Yost 2008, 85).

Beide Treppen verjüngen sich von den Fenstern zum Schneckenloch im Querschnitt um ungefähr 1 bis 2 mm². Basal ist die Weite der Vorhoftreppe mit knapp 1,6 mm angegeben und jene der Paukentreppe mit knapp 2 mm etwas weiter. Apikal messen beide Treppen rund 0,6 mm. Dabei ist zu bemerken, daß die Schnecke insgesamt, wie Abbildung 2.10 veranschaulicht, keine besonders kontinuierliche Verjüngung erfährt, sondern vor allem mitten in der basalen Windung eine bemerkenswerte Engstelle aufweist (vgl. Graumann & Sasse 2005, 113; Ranke 1953, 64). Das Schneckenloch selbst ist zwischen 0,25 und 0,4 mm² groß (vgl. Rinsdorf in Keidel 1975, 64).

Häutige Schnecke

Die knöcherne Schnecke ist der Länge nach durch zwei Membranen dreigeteilt. Der mittig liegende *Schneckenengang* (Ductus cochlearis) bildet den häutigen Abschnitt der Schnecke (vgl. Békésy 1941(b), 265; Graumann & Sasse 2005, 113; Helmholtz 1896, 224; Müller & Frings 2009, 452ff.; Reiss 2003, 616). Analog zu den äußeren Treppen (Scala vestibuli, Scala tympani) wird der Schneckenengang auch *Scala media* genannt (vgl. Gelfand 2004, 54; Müller & Frings 2009, 452; Stevens & Hallowell 1938, 271; Yost 2008, 86). Er lagert sich der konkaven Wand des Schneckenkanals an und ist medial an der knöchernen Trennwand aufgehängt (vgl. Reiss 2003, 620; Gelfand 2004, 55; 2009, 55; Helmholtz 1896, 224). Von der Vorhoftreppe trennt ihn die *Reissnermembran*, von der Paukentreppe die *Basilar-membran* (vgl. Reiss 2003, 620; Gelfand 2004, 55; Helmholtz 1896, 224f.; Müller & Frings 2009, 453; Ranke 1953, 65).

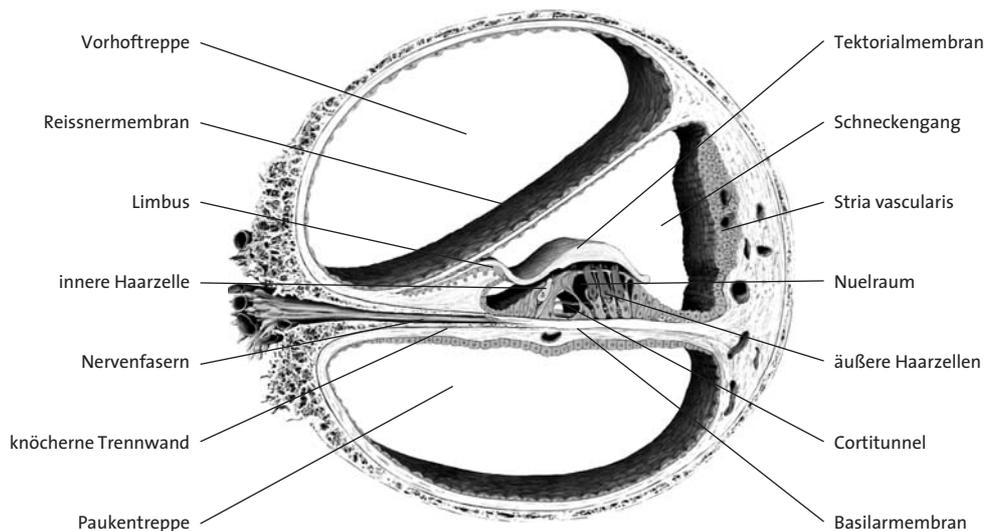


Abb. 2.11 **Querschnitt durch die Schnecke** (modifiziert n. Schünke et al. 2009, 136 c). Darstellung in traditioneller Ansicht mit der Vorhoftreppe oben und der Paukentreppe unten.

Der von der knöchernen Trennwand zur lateralen Schneckenwand freigelassene Spalt wird durch die *Basilarmembran* überbrückt (vgl. Graumann & Sasse 2005, 112; Stevens & Hallowell 1938, 269; Terhardt 1998, 57f.; Yost 2008, 85). Medial geht sie mehr oder weniger fließend in die knöcherne Trennwand (vgl. Békésy 1941(b), 265f., 276; Helmholtz 1896, 223) und auf deren „vestibularer“ Seite (d.h. zur Vorhoftrappe, bzw. im Schneckenang gelegen) oberflächlich in den *Limbus cochlearis*, eine wulstartige Verdickung der *Knochenhaut* der Trennwand, über (vgl. Gelfand 2004, 55; 2009, 55; Siebenmann 1897, 309). Lateral endet sie nicht direkt an der Knochenhaut, sondern im *Ligamentum spirale cochleae*, einer dicken, der Knochenhaut aufsitzenden, gefäßversetzten fibrösen Struktur, welche die laterale Wand des Schneckenkanals großzügig auskleidet (vgl. Gelfand 2004, 55; 2009, 55; Reiss 2003, 620; Spee 1896, 307; Stevens & Hallowell 1938, 271). Der Limbus vervollständigt medial die hauptsächlich aus der Basilarmembran bestehende „tympanale“ (d.h. zur Paukentreppe hin gelegene) Wand des Schneckengangs.

Die *Reissnermembran* (*Membrana vestibularis*, *Membrana reissneri*) entspringt einerseits (medial) dem Limbus cochlearis und ist andererseits (lateral), ebenso wie die Basilarmembran, am Ligamentum spirale aufgehängt. Sie bildet die „vestibulare“ Wand des Schneckengangs. Mit der Basilarmembran formt sie einen lateral offenen Winkel zwischen 30 und 45°, wodurch der Schneckenang im Querschnitt dreieckig erscheint (vgl. Békésy 1941(b), 265; Gelfand 2004, 55; 2009, 55; Graumann & Sasse 2005, 113f.; Helmholtz 1896, 224f.; Müller & Frings 2009, 453; Ranke 1953, 64; Reiss 2003, 616, 620; Siebenmann 1897, 306; Stevens & Hallowell 1938, 271; Yost 2008, 85; Zwicker & Feldtkeller 1967, 84; Zwislocki 1948, 20).

Zwischen den Aufhängungspunkten beider Membranen ist das Ligamentum spirale seinerseits mit einem mehrschichtigen Epithel, der *Stria vascularis*, ausgekleidet, welches die äußere, konkave Wand des Schneckengangs bildet und gleichzeitig für die Produktion der Endolymphe verantwortlich gemacht wird (vgl. Fuchs 2010, 6; Gelfand 2004, 59; 2009, 56; Graumann & Sasse 2005, 114; Marcus & Wagemann 2010; Reiss 2003, 618, 620; Yost 2008, 86f.).

Im Gegensatz zu den äußeren Treppen ist der Schneckenang geschlossen und endet auf beiden Seiten praktisch blind (vgl. Gelfand 2004, 54; Graumann & Sasse 2005, 113). Reissner- und Basilarmembran folgen seiner Länge (bei unterschiedlichen Winkeln zueinander) vergleichsweise konstant parallel. Erst in der Spitze der Schnecke, kurz vor dem Schneckenloch, treffen sie zusammen (vgl. Yost 2008, 86) und bilden sein apikales Ende, das *Caecum cupulare* (vgl. Graumann & Sasse 2005, 113; Reiss 2003, 620). Basal, im Vorhof inferior des ovalen Fensters, endet er im *Caecum vestibulare* (vgl. Reiss 2003, 620) praktisch blind, de facto steht er hier über den *Ductus reuniens* mit dem Sacculus in Verbindung (vgl. Gelfand 2004, 53; Graumann & Sasse 2005, 129; Helmholtz 1896, 223, 224; Reiss 2003, 615, 617; Schünke et al. 2009, 134; Siebenmann 1897, 306). Durch seine sackartigen Enden büßt der Schneckenang an beiden Enden seinen dreieckigen Querschnitt ein (vgl. Siebenmann 1897, 307).

Die Länge des Schneckengangs wird in einer weiten Spanne von 31 bis über 35 mm angegeben. Geht man von einer Kanallänge von 35 mm aus, sind Längen im Bereich von 31 bis 34 mm naheliegende Werte, da Platz für das Schneckenloch bleiben muß. Basal beträgt der Durchmesser des Schneckengangs rund 0,5 und apikal 0,35 mm, seine Breite hingegen basal ungefähr 0,45 und apikal 0,8 mm (vgl. Graumann & Sasse 2005, 113; Ranke 1953, 64; Siebenmann 1897, 306; Yost 2008, Abb. 7.4, 87, 88).

2.4.3 Basilarmembran

Während die *Reissnermembran* als simple, ihrer gesamten Länge nach gleich elastische Trennhaut zwischen Peri- und Endolympe fungiert, ist die *Basilarmembran* von besonderer Beschaffenheit und Bedeutung für die mechanischen Abläufe im Innenohr (vgl. Békésy & Rosenblith 1951, 1094; Gelfand 2004, 55; Graumann & Sasse 2005, 112). Die *Basilarmembran* (Membrana basilaris, Lamina basilaris, Lamina spiralis membranacea, BM) ist der Länge nach zwischen knöcherner Trennwand (medial) und Ligamentum spirale (lateral) aufgehängt (vgl. Békésy 1941(b), 265; Gelfand 2004, 55; Graumann & Sasse 2005, 114; Reiss 2003, Siebenmann 1897, 309; 620; Stenfelt et al. 2003, 132; Yost 2008, 85). Medial geht sie dabei mehr oder weniger fließend in die knöcherne Trennwand, bzw. deren Knochenhaut (speziell den Limbus cochlearis) über (vgl. Békésy 1941(b), 265f., 276; Helmholtz 1896, 223; Siebenmann 1897, 309), so daß beide eigentlich als Einheit zu verstehen sind (vgl. Békésy 1941(b)). In der basalen Windung wird das Ligamentum spirale durch die Lamina spiralis secunda unterstützt (vgl. Graumann & Sasse 2005, 112; Siebenmann 1897, 309; Spee 1896, 207).

Dem Spalt zwischen knöcherner Trennwand und lateraler Schneckenwand entsprechend (siehe 2.4.2, S. 73) ist die BM basal besonders schmal – der knöcherne Abschnitt der Trennwand ist hier gut und gerne mehr als doppelt so breit – dafür jedoch sehr steif. Richtung Apex wird die BM zunehmend breiter und elastischer. Insgesamt nimmt die Elastizität vom Steigbügel bis zum Schneckenloch um den Faktor 100 zu (vgl. Békésy 1941(b), 276; Boer & Nuttall 2010, 145, 154; Fuchs 2010, 4; Gelfand 2004, 55; 56, 2009, 54; 127; Spitzer 2002, 59; Stenfelt et al. 2003, 141; Stevens & Hallowell 1938, 273; Winckel 1960, 80; Yost 2008, 87f., 94; Zwislocki 1948; 14). Mit der Elastizität verändert sich auch die Frequenzempfindlichkeit entlang der BM, wodurch jeder Punkt eine für ihn spezifische Eigenresonanz aufweist (*charakteristische Frequenz*; siehe 3.3.2, S. 123). Da Steifheit hohe und Elastizität tiefe Frequenzen begünstigt, werden die örtlichen Resonanzfrequenzen von der Basis zum Apex immer geringer (vgl. Békésy 1941b, 266; Békésy & Rosenblith 1951, 1097; Boer & Nuttall 2010, 139; Fuchs 2010, 4; Gelfand 2009, 66; Hellbrück & Ellermeier 2004, 100; Müller & Frings 2009, 458; Nilsen & Russell 1999, 642; Spitzer 2002, 59f.; Stevens & Hallowell 1938, 356; Terhardt 1998, 57; Winckel 1960, 80; Yost 2008, 94f.; Zwicker & Feldtkeller 1967, 85). Bereits auf der BM zeigt sich also das tonotopische¹⁴ Ordnungsprinzip des Gehörs, nach dem Frequenzen (bzw. Frequenzbänder) ortsabhängig kodiert sind (vgl. Oxenham & Wojtczak 2010, 6; Plattig in Keidel 1975, 370). Zu erwähnen ist vielleicht noch, daß Ranke und Schriever den Sachverhalt umgekehrt darstellen und der BM eine vom Steigbügel zum Schneckenloch mit zunehmender Breite um den Faktor 100 *abnehmende* Elastizität zuschreiben (vgl. Ranke 1953, 82; Schriever 1967, 21).

Anders als Helmholtz sie im Rahmen seiner *Resonanztheorie* beschrieben hat (vgl. 1863, 215ff.; 1896, 235ff.) und man aufgrund des (einschlägigen) Stichworts „Resonanzfrequenz“ vielleicht erwarten würde, ist die BM keine vibrationsempfindliche, gespannte Membran, deren Querfasern einer Reihe von Resonanzseiten gleichen. Dagegen spricht nicht zuletzt, daß sie sowohl in longitudinaler, als auch in radialer Richtung vollkommen gleich elastisch zu sein scheint (vgl. Békésy 1941(b), 273; Békésy & Rosenblith 1951, 1093; Gelfand 2009, 65; Schriever 1967, 21). Viel eher muß man sie als spannungsfreies (wenn überhaupt, nur basal geringfügig gespanntes) Gebilde, am ehesten als gallertartige, von einer dünnen homogenen Faserschicht überlagerte, biegesteife Platte verstehen (vgl. Békésy 1941(b), 266ff., 274; Békésy & Rosenblith 1951, 1090, 1094; Gelfand 2004, 122; Winckel 1960, 80; Yost 2008,

14 *Tonotopie* (Tonos = gr. Spannung, sprich: Klang, Ton; Topos = gr. Ort): Bestimmte Frequenzen sind an bestimmte Orte (auf der Basilarmembran, der Hörrinde, etc.) gebunden.

88; Zwislocki 1948, 14). Ähnlich wie beim Trommelfell wäre ständige Spannung sogar störend, da bereits kleinste Veränderungen nennenswerte Auswirkungen hätten. Überdies kann Dispersion allein durch die ortsvariable, breitenabhängige Membransteifheit erreicht werden (vgl. Békésy 1941(b), 268; Békésy & Rosenblith 1951, 1094).

Die BM besteht nach Békésy aus eher starren, radial verlaufenden Bindegewebsfasern (vgl. Békésy 1941(b), 267; Müller & Frings 2009, 454; Reiss 2003, 620; Siebenmann 1897, 311). Während Stumpf (vgl. 1890, 94) nur im lateralen Abschnitt (*zona pectinata*, s.u.) isolierbare Radialfasern finden konnte, stellt Reiss (vgl. 2003, 620) Faserlängen fest, die ziemlich genau der gesamten Membranbreite entsprechen (basal zirka 0,1 bis 0,2 mm und apikal um die 0,5 mm). Eine besondere, für Spannung nötige und von Helmholtz (vgl. 1896, 225, 240) festgestellte, Stärke dieser Querfasern kann Békésy nicht finden:

Denn zerrupft man unter dem Mikroskop die Basilmembran, so erweist sie sich als locker im Verhältnis etwa zu den Fasern des in die Schneckenachse eindringenden Hörnerven, während sich z. B. die Faserhaut, die auf der Gehirnseite des Felsenbeins liegt, kaum zerreißen läßt (Békésy 1941(b), 267).

Sowohl von tympanaler, als auch von vestibularer (im Schneckenang) Seite ist die BM einschichtig verkleidet (vgl. Békésy 1941(b), 267; Graumann & Sasse 2004, 114; Reiss 2003, 620). Auf tympanaler Seite ist dies die tympanale Belegschicht, welche quasi die gesamte Paukentreppe in sehr unterschiedlicher Schichtdicke auskleidet (vgl. Reiss 2003, 620f.; Siebenmann 1897, 310). Ihre Längsfasern stützen die BM und verleihen ihr in Längs- und Querichtung vergleichbare elastische Eigenschaften (vgl. Békésy 1941(b), 267). Auf vestibularer Seite erhebt sich aus der Beschichtung eine verdickende Häufung diverser Stütz- und Sinneszellen, das Cortiorgan (vgl. Fuchs 2010, 6; Gelfand 2004, 56; 2009, 55; Helmholtz 1896, 231f.; Müller & Frings 2009, 454; Reiss 2003, 621; Reiss et al. 1989, 49; Schriever 1967, 39; Siebenmann 187, 310; Yost 2008, 88).

In radialer Richtung lassen sich zwei Abschnitte der BM unterscheiden. Der mediale Bereich, ungefähr dem *Cortitunnel* (siehe 2.4.4, S. 80) entsprechend (*Zona arcuata*), ist dünner, während der laterale Abschnitt (*Zona pectinata*) dicker geschichtet ist (vgl. Gelfand 2004, 56, Abb. 2.19, 57; Siebenmann 1897, 310f.).

Folgt man Békésys und Helmholtz' Ausführungen, wird die zunächst knöcherne Schneckenkannwand nach lateral zunehmend elastischer und geht schließlich – mehr oder weniger fließend – in einen membranösen Abschnitt, die BM, über (vgl. Békésy 1941(b), 265f., 276; Helmholtz 1896, 223; Siebenmann 1897, 310). Die weiter oben noch als unvollständig aufgefaßte Spaltung des Schneckenkanals durch die knöcherne Trennwand (siehe 2.4.2, S. 74), muß demnach als vollständige Teilung durch eine teils knöcherne, teils häutige *Schneckenkannwand* verstanden werden (vgl. Békésy 1941(b), 265; Békésy & Rosenblith 1951, 1078, 1095; Helmholtz 1896, 223; Stevens & Hallowell 1938, 269).

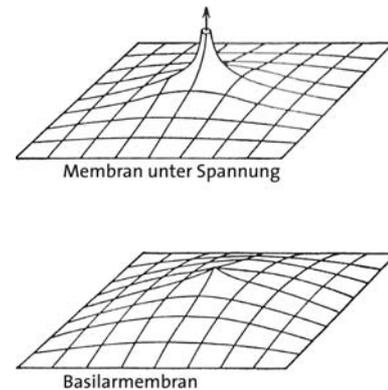


Abb. 2.12 Die Spannungsverhältnisse der Basilmembran wie bei Békésy dargestellt (1941(b), Abb. 2, 267).

Die legitime Zusammenfassung von BM und knöcherner Trennwand scheint sich auch aus der Bewegung zu ergeben. Die BM schwingt nicht alleine. Die gesamte Trennwand schwingt phasengleich, wenn auch die Amplituden medial (im knöchernen Abschnitt) geringer sind als am lateralen Ende der BM (vgl. Békésy 1941(b), 275; Békésy & Rosenblith 1951, 1095; Stenfelt et al. 2003, 131f., 137, 141ff.). Im Grunde würde es sogar genügen, nur die Reissnermembran zu beobachten, deren mediales, am Limbus verwachsenes Ende die größten Amplituden zeigt (vgl. Békésy 1941(b), 275; Ranke 1953, 105; Schriever 1967, 21), was bei einer völlig starren Lamina spiralis ossea nicht denkbar wäre.

Vielfach werden derlei Sachverhalte in jüngerer Literatur nicht (mehr) berücksichtigt. Sowohl strukturell, als auch physikalisch wird (mitunter implizit) eine klare Demarkationslinie zwischen knöcherner Trennwand und BM gezogen, wobei letztere als alleine schwingendes Element benannt wird (vgl. Fuchs 2010; Gelfand 2004, 2009, Graumann & Sasse 2005, Guski 1996; Hellbrück & Ellermeier 2004; Müller & Frings 2009; Reiss 2003, Schünke et al. 2009, Spitzer 2002, Terhardt 1998, Yost 2008; Zwicke & Feldtkeller 1967, 84).

Die BM ist um die 32 mm lang, wobei die Angaben von 31,5 bis 35 mm reichen (vgl. Békésy 1941(b), 273; 1960, 29; Fastl & Zwicker 2006, 162; Gelfand 2004, 55; Rinsdorf in Keidel 1975, 64; Yost 2008, 88; Winckel 1960, 80; Zwicker & Feldtkeller 1967, 91). Basal ist sie ungefähr 0,08 bis 0,2 mm breit, während sie apikal zwischen 0,5 und 0,6 mm mißt (vgl. Békésy 1960, 405; Gelfand 2004, 55; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35f.; Reiss 2003, 620; Rinsdorf in Keidel 1975, 64; Terhardt 1998, 58; Yost 2008, 87f.; Zwicker & Feldtkeller 1967, 84). Siebenmann gibt nach Retzius sogar Werte von 0,45 (basal) bis 0,8 mm (apikal) an (vgl. 1897, 309). Im Gegensatz zur Schnecke, deren Verjüngung höchst unregelmäßig verläuft, nimmt die Breitenzunahme der BM einen äußerst regelmäßigen, glatten und „systematischen“ (Stevens & Hallowell 1938, 273) Verlauf (siehe Abb. 2.10, S. 74).

2.4.4 Cortiorgan

Bei senkrechter Betrachtung „auf“ der Basilarmembran, d.h. auf ihrer eigentlich lateralen, vestibularen Seite im Schneckengang, liegt mit dem *Cortiorgan* (Organum spirale cochleae, Organum corti) der eigentliche sensorische „Schallaufnahmeapparat“ des Gehörs (vgl. Békésy 1941b, 265; Gelfand 2004, 56; Graumann & Sasse 2005, 114; Guski 1996, 91; Hellbrück & Ellermeier 2004, 94; Müller & Frings 2009, 454; Ranke 1953, 65; Reiss 2003, 620; Schünke et al. 2009, 136; Spitzer 2002, 58; Yost 2008, 88; Zwicker & Feldtkeller 1967, 84). Das Cortiorgan trägt den Namen seines Entdeckers, Alfonso Giacomo Gaspare, Graf von Corti zu San Stefano Belbo (vgl. Helmholtz 1896, 228; Plattig in Keidel 1975, 44; Ranke 1953, 65; Reiss et al. 1989, 44). Es verläuft im Grenzbereich von BM und knöcherner Trennwand.

Das Cortiorgan besteht oberflächlich betrachtet aus einer verdickenden Häufung diverser Stütz- und Pfeilerzellen und spezieller Sinneszellen, den *Haarzellen* (vgl. Fuchs 2010, 6; Gelfand 2004, 56; 2009, 56; Helmholtz 1896, 231f.; Müller & Frings 2009, 454; Reiss 2003, 621; Reiss et al. 1989, 49; Schriever 1967, 39; Siebenmann 187, 310; Stevens & Hallowell 1938, 271f.; Yost 2008, 88). Im radialen Schnitt findet sich medial gelegen eine Reihe *innerer Haarzellen* (IHZ), während lateral drei Reihen

äußerer Haarzellen (ÄHZ) liegen (vgl. Fuchs 2010, 6; Gelfand 2004, 56; 2009, 56; Graumann & Sasse 2005, 114, 116; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 94; Müller & Frings 2009, 456; Reiss 2003, 622; Siebenmann 1897, 313; Stevens & Hallowell 1938, 272; Terhardt 1998, 58; Yost 2008, 88; Zwicker & Feltdtkeller 1967, 85f.). Es wurden auch schon vier oder fünf Reihen ÄHZ beschrieben (vgl. Gelfand 2004, 56; Reiss 2003, 622; Terhardt 1998, 58; Siebenmann 1897, 311; Stevens & Hallowell 1938, 272; Yost 2008, 88), was möglicherweise eine Frage des Meßpunkts ist, denn ab der Mittelwindung Richtung Apex verliert sich die Ordnung im Arrangement der ÄHZ, wofür nach Siebenmann eine mitunter unvollständige vierte (bisweilen gar eine fünfte) Reihe auftritt (vgl. Gelfand 2004, 56; Siebenmann 1897, 313). Möglich ist auch, daß die verlorene Ordnung mehrfach ÄHZ aus der Reihe tanzen läßt, was den Eindruck einer vierten Reihe vermittelt. In jedem Fall kann man von normalerweise drei Reihen ausgehen.

Nimmt man das Cortiorgan genauer unter die Lupe, zeigt sich, daß es sich keinesfalls um ein massives Gebilde, sondern um eine Stütz- und Isolationskonstruktion für die Haarzellen handelt. Den Kern bilden die stabilen *inneren* und *äußeren Pfeilerzellen* (Epitheliocyti pilares interni bzw. externi). Die inneren Pfeilerzellen stehen auf der Kante der knöchernen Trennwand, die äußeren auf der Basilar-membran. Sie lehnen sich schräg gegeneinander, so daß sie unter sich einen im Querschnitt dreieckigen Raum, den *Cortitunnel* (innerer Tunnel, Cuniculus internus) freilassen (vgl. Gelfand 2004, 57; 2009, 56; Graumann & Sasse 2005, 114ff.; Reiss 2003, 621; Siebenmann 1897, 311f.; Stevens & Hallowell 1938, 271f.; Yost 2008, 88). Er bildet zwar nicht die geometrische Mitte des Cortiorgans, wohl aber dessen „Gravitationszentrum“, dem die umliegenden Strukturen zugeneigt sind. Die pfannenartigen Köpfe der inneren Pfeilerzellen umschließen jene der äußeren Pfeilerzellen wie bei einem Kugelgelenk (vgl. Graumann & Sasse 2005, 116; Siebenmann 1897, 312), wodurch das cortische Grundgerüst bei aller Stabilität flexibel und für die Schwingungen der Schnecken-trennwand gerüstet bleibt.

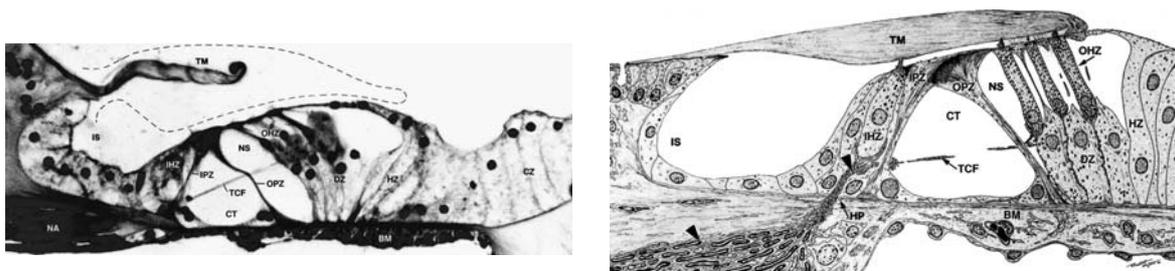


Abb. 2.13 **Querschnitt durch das Cortiorgan** (Reiss et al. 1989, Abb. 81, 49). BM Basilmembran, CT Cortitunnel, DZ Deiterzellen, HP Habenula perforata, IHZ/OHZ innere/äußere Haarzellen, IPZ/OPZ innere/äußere Pfeilerzellen, IS interner Sulzus, NS Nuelraum (mittlerer Tunnel), TFC Nerven-faserstrang durch Cortitunnel, TM Tektorialmembran.

Medial der inneren Pfeilerzellen sitzen die kurzen, runden, zum Cortitunnel geneigten IHZ auf den sattelförmigen Enden ihrer Stützzellen, den *inneren Phalangenzellen* (Epitheliocyti phalangei interni) (vgl. Gelfand 2004, 56; Graumann & Sasse 2005, 114). Direkt medial davon markiert eine Reihe *innerer Grenzzellen* (Epitheliocyti limitantes interni) den medialen Abschluß des Cortiorgans und bildet einen fließenden Übergang in die einschichtige Verkleidung des Schnecken-gangs (vgl. Graumann & Sasse 2005, 114; Reiss 2003, Abb. 7.25, 621). Generell läßt sich medial der Pfeilerzellen ein enger und dichter Zellverbund erkennen, der die IHZ fest umschließt und über die inneren Grenzzellen schnell auf

die umliegende Epitheldicke zurückgeht. Lateral ist das Cortiorgan weniger dicht und gerüstartiger strukturiert. Die *äußeren Phalangenzellen* (Epitheliocyti phalangei externi) oder *Deiterzellen* übernehmen hier die Stützfunktion, wofür sie – wie die Pfeilerzellen – durch ein intrazelluläres Filamentgerüst verstärkt sind (vgl. Graumann & Sasse 2005, 117). Als Träger der ÄHZ sind sie in drei Reihen organisiert, die wie Sitzreihen im Theater zueinander verschoben sind und dabei recht lose, ohne seitlichen Kontakt nebeneinander stehen (vgl. Gelfand 2004, 56; Graumann & Sasse 2005, 116). Dadurch entstehen medial und lateral der Deiterzellen und ÄHZ größere Räume, die als *Nuelraum* (mittlerer Tunnel, Cuniculus medium) und *äußerer Tunnel* (Cuniculus externus) bezeichnet werden (vgl. Graumann & Sasse 2005, Abb. 1.76, 116, 117). Alle diese Tunnelräume stehen weiträumig miteinander in Verbindung und können als, speziell die ÄHZ umschwemmender, Lymphraum des Cortiorgans aufgefaßt werden (vgl. Siebenmann 1897, 312). Die Deiterzellen bilden, ähnlich wie ihre inneren Gegenstücke, Sattel, auf denen die ÄHZ sitzen. Lateral verfügen sie jedoch über Fortsätze, die sich apikal zu rautenförmigen Kopfplatten ausweiten, welche die Köpfe der ÄHZ erfassen und sich untereinander sowie mit den umliegenden Zellen fest verbinden (vgl. Gelfand 2004, 56f.; Graumann & Sasse 2005, 116; Reiss 2003, 621; Siebenmann 1897, 312). Aus den Kopfplatten der Deiterzellen, den Pfeilerzellen und den Kutikularplatten der Haarzellen entsteht mit der *Membrana reticularis* eine dichte Barriere, die den cortischen Lymphraum von der Endolymphe des Schneckengangs isoliert (vgl. Fukazawa 2002, 55; Gelfand 2004, 52, 57; 2009, 56; Graumann & Sasse 2005, 118; Iwasa 2010, 180; Reiss 2003, 621f.; Siebenmann 1897, 312). Die ÄHZ sind nicht nur zum Cortitunnel, sondern auch längs der BM in basaler Richtung geneigt. Die Fortsätze der Deiterzellen lehnen sich dagegen nach apikal, und treffen in der Membrana reticularis auf die Kutikularplatten der apikal benachbarten ÄHZ (vgl. Fukazawa 2002, 55; Gelfand 2004, Abb. 2.21, 58). Lateral der Deiterzellen schließt eine Reihe *äußerer Grenzzellen* (Epitheliocyti limitantes externi) das Cortiorgan ab (vgl. Graumann & Sasse 2005, 117). Über Reihen höherer Stützzellen (*Hensenzellen*) und niedrigerer Stützzellen (*Claudiuszellen*) geht das Gebilde Richtung Sulcus externus allmählich in indifferente Epithelzellen über (vgl. Gelfand 2004, 57; Graumann & Sasse 2005, 117; Reiss 2003, 621; Siebenmann 1897, 313). Dadurch, daß die BM von der Basis bis zur Spitze immer breiter wird, verändern sich die mechanischen Voraussetzungen für das Cortiorgan. Basal nimmt es praktisch die gesamte Membranbreite ein, während es apikal nur den medialen Abschnitt (teilweise) belegt und über mehr Bewegungsfreiheit verfügt (vgl. Fukazawa 2002, 60).

Im Übergangsbereich zwischen knöcherner Trennwand und BM treten die Nervenfasern durch die Zona perforata in das Cortiorgan ein (vgl. Reiss et al. 1989, 49; Siebenmann 1897, 310). Von hier aus verlaufen dicke Faserbündel direkt zu den IHZ, andere Fasern ziehen sowohl radial als auch leicht nach apikal gerichtet durch bzw. unter dem Cortitunnel zu den ÄHZ (vgl. Siebenmann 1897, 312, 315; Reiss et al. 1989, 49).

Tektorialmembran

Über dem Cortiorgan liegt die *Tektorialmembran* (Membrana tectoria). Sie entspringt gemeinsam mit der Reissnermembran am Limbus cochlearis (vgl. Gelfand 2009, 55f.; Graumann & Sasse 2005, 114; Siebenmann 1897, 313; Stevens & Hallowell 1938, 273), mit dessen oberer Lippe sie fest verbunden ist (vgl. Gelfand 2004, 59, Siebenmann 1897, 310, 313) und dessen *Interdentalzellen* ihre Entstehung

zumindest teilweise mit zugeschrieben wird (vgl. Graumann & Sasse 2005, 114). Strukturell wird die Tektorialmembran nicht als Teil des Cortiorgans, sondern diesem aufliegend beschrieben (vgl. Keidel 1975, 11; Schriever 1967, 34; Siebenmann 1897, 313). Da beide jedoch erst gemeinsam funktionieren – quasi gegenseitig sinnstiftend sind – schien es mir gerechtfertigt, sie an dieser Stelle einzugliedern.

Die Tektorialmembran kann im Radialschnitt in drei Längszonen – eine innere (limbische), mittlere (cortische) und äußere (Rand) Zone – unterteilt werden (vgl. Gelfand 2004, Abb. 2.19, 57; Siebenmann 1897, 313). Im cortischen Abschnitt wird die Tektorialmembran merklich dicker. Insgesamt wird sie meist als weiche, band- oder gallertartige, elastische Masse beschrieben und besteht vorwiegend aus Kollagen Typ-II Fasern, wodurch sie bei geringer Dehnbarkeit hohe Zugfestigkeit gewinnt (vgl. Boer & Nuttall 2010, 164; Fukazawa 2002, 54; Gelfand 2004, 59; 2009, 56; Hackney & Furness 2010, 231; Müller & Frings 2009, 454; Reiss 2003, 622; Siebenmann 1897, 313; Spitzer 2002, 60; Yost 2008, 88). Auf cortischer Seite zeichnet sich ein kleiner Grat, der *Hensenstreifen*, ab. Er verläuft mehr oder weniger parallel zu und über den frei stehenden Stereozilien der IHZ und soll in diesem Bereich den Endolymphfluß optimieren (vgl. Gelfand 2004, Abb. 2.19, 57; Hackney & Furness 2010, 231f.; Lukashkin et al. 2010, 33; Siebenmann 1897, 313). Das scheinbar strategische anatomische Verhältnis zwischen Tektorialmembran und Haarzellen deutet an, daß erstere bei der Stimulation letzterer eine Schlüsselrolle spielt (vgl. Ghaffari et al. 2007, 16510).

Schon „über“ dem Cortitunnel nimmt die Membrandicke wieder ab. Im Randabschnitt „über“ den Hensenzellen endet sie in der Basalwindung mit einem leichten Saum, weiter apikal in einer netzartigen Struktur, dem *Löwenberger'schen Randfasernetz* (vgl. Gelfand 2004, 59; 2009, 56; Schriever 1967, 34; Siebenmann 1897, 313). Darüber, ob das Netz mit den Hensen-, Grenz- oder anderen cortischen Zellen verbunden ist, herrscht in der Literatur Uneinigkeit (vgl. Gelfand 2004, 59; Siebenmann 1897, 313; Schriever 1967, 34; Yost 2008, 88).

Die Tektorialmembran ist nicht einfach nur elastisch. Es lohnt sich, einen genaueren Blick auf die Verhältnisse zu werfen. Während Békésy und Rosenblith (vgl. 1951, 1095) ihr nur nahe des Steigbügels formbedingt eine gewisse Steifheit zusprechen, ist sie nach neueren Erkenntnissen viskoelastisch, wobei Elastizität und Masse (analog zur BM) längs variieren (vgl. Boer & Nuttall 2010, 164f.; Ghaffari et al. 2007, 16510). Nach Meinung von Ghaffari et alii (vgl. 2007, 16510) ist die Viskoelastizität zur mechanischen Vorkopplung über weite Distanzen geeignet. Da die Tektorialmembran nach allgemeinem Verständnis keinen Druckdifferenzen (wie die Schneckenrennwand) ausgesetzt ist, sind Wanderwellen nur durch elastische Vorkopplung innerhalb der Membran möglich (vgl. Lukashkin et al. 2010, 32).

Haarzellen

Die Haarzellen tragen ihren Namen aufgrund der haarähnlichen Fortsätze an ihrem „Kopffende“ (vgl. Gelfand 2004, 59; 2009, 56; Schriever 1967, 39; Stevens & Hallowell 1938, 272f.). Prinzipiell setzen sich innere und äußere Haarzellen ähnlich zusammen, verfügen darüber hinaus jedoch über entscheidende Unterschiede (vgl. Gelfand 2009, 57).

Die Zellkörper beider Typen enden apikal mit einer verdickten Kopfplatte, *Kutikularplatte* genannt (vgl. Gelfand 2004, 59; 2009, 56; Yost 2008, 88). Allerdings machen die Kutikularplatten nicht die gesamte Fläche der Kopfplatten aus. Stattdessen sind lateral Überbleibsel eines einzigen rudimentären

Kinozilium, sogenannte *Basalkörperchen*, erkennbar (vgl. Gelfand 2004, 59; 2009, 57; Reiss 2003, 622; Yost 2008, 88). Auf den Kutikularplatten befinden sich vollendete Sinneshärchen, die Stereozilien (vgl. Yost 2008, 88). Diese sind in mehreren, meist drei Reihen angeordnet (vgl. Gelfand 2004, 59), wobei die Haarzellen sozusagen einen „Stufenschnitt“ tragen: axial (d.h. zum Modiolus hin gelegen) sind die Stereozilien kürzer als lateral (vgl. Fuchs 2010, 6; Gelfand 2004, 61; 2009, 57; Graumann & Sasse 2005, 118; Hackney & Furness 2010, 231). Generell sind die Stereozilien im Verlauf der Schnecke nicht gleich lang, sondern basal durchschnittlich etwas kürzer als apikal (vgl. Gelfand 2004, 61). An ihrer Wurzel auf der Kutikularplatte sind sie erkennbar dünner als am restlichen Körper, weshalb sie an sich steif bleiben und nur basal umgeknickt werden (vgl. Gelfand 2004, 60f.; Schriever 1967, 39).

Die Stereozilien der IHZ und der ÄHZ sind jeweils miteinander verbunden (vgl. Fuchs 2010, 8; Gelfand 2004, 62f.; Hackney & Furness 2010, 231, 242; Ranke 1953, 106). Durch knapp unter den „Haarspitzen“ wurzelnde *Cross-Links* ist jedes Stereozilium an seine unmittelbaren Reihennachbarn gebunden. Noch etwas tiefer verlaufen auch radiale *Cross-Links*, die so alle *Reihen* miteinander verbinden (vgl. Gelfand 2004, 62f.; Hackney & Furness 2010, 244). Darüber hinaus sind jeweils die Spitzen der kürzeren mit den Seiten der längeren Stereozilien über sogenannte *Tip-Links* gekoppelt (vgl. Gelfand 2004, 63; Hackney & Furness 2010, 235). Für IHZ werden sogar Verbindungen zwischen Stereozilien benachbarter Zellen beschrieben (vgl. Gelfand 2004, 62). Den *Tip-Links* wird die entscheidende Rolle für die Aktivierung der Haarzellen zugeschrieben (vgl. Fuchs 2010, 8; Gelfand 2004, 63; Yost 2008, 90). Ebenso entscheidend ist, daß die über *Tip-Links* verbundenen Stereozilien unterschiedlich lang sind (vgl. Fuchs 2010, 6). An den Wurzeln der *Tip-Links* liegen die Ionenkanäle (*mechanoelektrische Transduktionskanäle*, MET-Kanäle) der Haarzellen. Erstere verschließen letztere wie ein Stopfen, wobei im Ruhezustand einige Ionenkanäle (nach Fuchs etwa 10%) offen bleiben (vgl. Fuchs 2010, 8; Hackney & Furness 2010, 242ff.; Iwasa 2010, 180; Spitzer 2002, 60). Werden die Stereozilien Richtung der längsten geknickt, spannen die *Tip-Links* und die Ionenkanäle öffnen sich komplett, kommt es zur Auslenkung Richtung der kürzeren, werden die Kanäle ob der entspannten *Tip-Links* komplett verschlossen (vgl. Gelfand 2009, 58f.; Iwasa 2010, 180; Spitzer 2002, 60), was letztlich zur De- (auf, Erregung) bzw. Hyperpolarisation (zu, Hemmung) der Haarzelle führt (vgl. Gelfand 2009, 59; Graumann & Sasse 2005, 119; Hellbrück & Ellermeier 2004, 101).

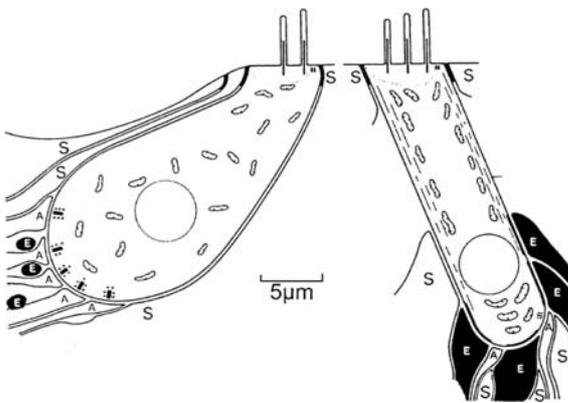


Abb. 2.14 **Haarzellen** (Hellbrück & Ellermeier 2008, Abb. 2.3, 44). Links: innere Haarzelle (IHZ), Rechts: äußere Haarzelle (ÄHZ). A afferente Synapse, E efferente Synapse, S Stützzellen.

Die *inneren Haarzellen* (Epitheliocyti sensorii pilosi interni, IHZ) haben eine rundliche oder birnenartige Form (vgl. Gelfand 2004, 63; 2009, 57; Graumann & Sasse 2005, 118). Sie tragen 40 bis 50 Stereozilien, die in mindestens zwei, normalerweise drei bis vier Reihen organisiert (vgl. Graumann & Sasse 2005, Tab. 1–9., 117; Reiss 2003, 622; Yost 2008, 88) und in einem nach axial konkaven Bogen, bzw. einer leichten U-Form, aufgestellt sind (vgl. Gelfand 2004, 60; Graumann & Sasse 2005, 118; Terhardt 1998, 58; Yost 2008, 88). Von rund 4 bis 5 µm an der Schneckenbasis werden die Stereozilien Richtung Apex,

wo sie 7 bis 8 und sogar 12 μm erreichen, mit annähernd linearem Verlauf immer länger (vgl. Gelfand 2004, 61; Ranke 1953, 106). Die Länge der Sinneshärchen korreliert mit ihrer Frequenzempfindlichkeit: je länger sie sind, desto tiefer liegt ihr bevorzugtes Frequenzband (vgl. Ranke 2004, 62). Man geht davon aus, daß die Stereozilien der IHZ keinen (nennenswerten) Kontakt zur Tektorialmembran haben (vgl. Gelfand 2004, 61; 2009, 57; Graumann & Sasse 2005, 118; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 94; Lukashkin et al. 2010, 33; Müller & Frings 2009, 456; Reiss 2003, 622).

Normalerweise werden ungefähr 3500 in einer Reihe angeordnete IHZ gezählt (vgl. Gelfand 2004, 56, 59; 2009, 57; Graumann & Sasse 2005, Tab. 1–9, 117; Guski 1996, 92; Müller & Frings 2009, 456; Ranke 1953, 108; Rinsdorf in Keidel 1975, 65; Schriever 1967, 39; Spitzer 2002, 58; Stevens & Hallowell 1938, 274; Zwicker & Feldtkeller 1967, 86), wobei auch schon 3600 bis sogar 4500 IHZ gezählt wurden (vgl. Fastl & Zwicker 2006, 162; Hellbrück & Ellermeier 2004, 128; Terhardt 1998, 58). Nach Gelfand (vgl. 2004, 59) verteilen sie sich mit einer Dichte von zirka 86 IHZ pro mm auf das Cortiorgan. Weil dieser rein rechnerische Haarzellenabstand von 11,6 μm bei 3500 IHZ eine Strecke von über 40 mm erfordern würde, dürften die IHZ mit einem mittleren Abstand von 9 μm sogar noch dichter stehen (vgl. Fastl & Zwicker 2006, 162; Zwicker & Feldtkeller 1967, 86, 92).

Die IHZ sind die eigentlichen Transduktoren, die das Signal an die afferenten Bahnen des neuronalen auditiven Systems leiten (vgl. Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 101; Iwasa 2010, 179; Müller & Frings 2009, 456). Jede einzelne IHZ hat bis zu 20 afferente Synapsen, wobei jede Faser nur eine bestimmte IHZ innerviert. Die wenigen efferenten Fasern, die „sich zu den IHZ verirren“ greifen diese dagegen nicht direkt ab, sondern setzen an einer afferenten Faser, nahe dem synaptischen Pol der Haarzelle, an (vgl. Fuchs 2010, 8; Gelfand 2004, 69, 71; 2009, 60; Graumann & Sasse 2005, Abb. 1–77, 117, 121; Guski 1996, 92; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 109; Müller & Frings 2009, 456; Ranke 1953, 108; Terhardt 1998, 58).

Die *äußeren Haarzellen* (Eptheliocyti sensorii pilosi externi, ÄHZ) haben einen schlanken, länglichen Zellkörper (vgl. Gelfand 2004, 63; 2009, 57). Sie tragen ungefähr 100 bis 150 Stereozilien in drei bis sieben Reihen (vgl. Gelfand 2004, 60; Graumann & Sasse 2005, Tab. 1–9, 117; Reiss 2003, 622; Yost 2008, 88), die in einer axial offenen V- oder W-Form angeordnet sind (vgl. Gelfand 2004, 60; 2009, 57; Graumann & Sasse 2005, 118; Iwasa 2010, 180; Terhardt 1998, 58; Yost 2008, 88). Auch die Stereozilien der ÄHZ werden im Schneckenverlauf Richtung Apex länger, jedoch nicht so linear wie die der IHZ. Ihre Längen werden mit 4 bis 8 μm angegeben. Im Gegensatz zu denen der IHZ sind die längsten jeder ÄHZ fest in der Tektorialmembran verankert (vgl. Fuchs 2010, 6; Gelfand 2004, 61ff.; Graumann & Sasse 2005, 118; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 94, 109, 112; Iwasa 2010, 180; Müller & Frings 2009, 456; Ranke 1953, 106; Reiss 2003, 622; Yost 2008, 88). Manche Autoren sehen die Stereozilien der ÄHZ über eine Art Schleimbeschichtung in Verbindung stehen (vgl. Graumann & Sasse 2005, 118; Ranke 1953, 106).

Die ÄHZ sind meist in drei zueinander versetzten Reihen aufgestellt, wobei auch bis zu fünf Reihen beschrieben werden. Die vierte und fünfte Reihe sind oft nur apikal und unvollständig entwickelt (vgl. Fuchs 2010, 6; Gelfand 2004, 56; Graumann & Sasse 2005, 116; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 94, 112; Müller & Frings 2009, 456; Ranke 1953, 108; Reiss 2003, 622; Rinsdorf in Keidel 1975, 65; Schriever 1967, 39; Terhardt 1998, 58; Yost 2008, 88). Darüber, wie viele ÄHZ im Cortiorgan zu finden sind, gehen die Beschreibungen auseinander. Während meistens von rund

12000 Zellen berichtet wird (vgl. Gelfand 2004, 56, 59; 2009, 57; Guski 1996, 92; Ranke 1953, 108; Spitzer 2002, 62; Stumpf 1890, 94), werden seltener auch 10500 bis 20000 ÄHZ angegeben (vgl. Graumann & Sasse 2005, Tab. 1–9, 117; Müller & Frings 2009, 456; Rinsdorf in Keidel 1975, 65; Stevens & Hallowell 1938, 274; Stumpf 1890, 94). Gelfand zählt zirka 343 ÄHZ pro mm (vgl. 2004, 59), was rechnerisch am ehesten für eine Gesamtzahl von um die 12000 ÄHZ (eine Zahl, die ihrerseits wunderbar in drei Reihen aufginge) sprechen würde.

Die ÄHZ dienen nicht der Schalltransduktion, sondern sind in erster Linie dafür verantwortlich, die Frequenzselektivität und die Empfindlichkeit des Gehörs zu optimieren (vgl. Boer & Nuttall 2010, 157; Hellbrück & Ellermeier 2004, 103, 108; Janssen in Lehnhardt & Laszig 2009, 113; Müller & Frings 2009, 460; Yost 2008, 113). Sie sind folglich überwiegend efferent direkt innerviert. Jede efferente Faser innerviert dabei mehrere, mindestens fünf bis sechs ÄHZ auf einer Strecke von 0,6 mm. Im Extremfall können einzelne Fasern aber auch Bereiche von einer Viertel- bis sogar einer Halbwindung abgreifen (vgl. Gelfand 2004, 71; Graumann & Sasse 2005, 121; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 109, 112; Müller & Frings 2009, 456, 460; Ranke 1953, 108; Stevens & Hallowell 1938, 275; Schriever 1967, 39f.; Terhardt 1998, 59). Mehr efferente Fasern enden in der medialen Reihe, weniger in der lateralen. Auffälliger ist aber, daß die Anzahl efferenter Fasern entlang der BM von der Basis in die Spitze stetig abnimmt. Möglicherweise ist die Dynamikkontrolle im hochfrequenten Bereich kritischer als bei tieferen Frequenzen (vgl. Gelfand 2004, 71; Hellbrück & Ellermeier 2004, 112f.). Allerdings ist noch unklar, welchen Einfluß die efferente Innervierung der ÄHZ auf den Transduktionsprozeß tatsächlich hat (vgl. Iwasa 2010, 205; Yost 2008, 110).

Den Einfluß der ÄHZ selbst sieht man klarer. Sie können zu diesem Zweck ihre Körperlänge in (an die Schneckenschwingung angepaßten) Oszillationen verändern (vgl. Fuchs 2010, 10; Gelfand 2004, 64; 2009, 58; Hellbrück & Ellermeier 2004, 107f.; Müller & Frings 2009, 460f.; Schünke et al. 2009, 137; Yost 2008, 109f.). Als Treiber dieser sogenannten *Elektromotilität* wurde das neu entdeckte Protein *Prestin* identifiziert. Es ist zu tausenden in die laterale Zellwand der ÄHZ eingelagert. Es reagiert extrem empfindlich auf Wechselspannungen mit eigenen, sogenannten *Konformationsänderungen* (die Fähigkeit, die eigene Körperstruktur zu verändern), die sich insgesamt zur Größenänderung der Zellmembran (und damit der Zelle) aufsummieren (vgl. Boer & Nuttall 2010, 162; Gelfand 2004, 166; Hellbrück & Ellermeier 2004, 108; Iwasa 2010, 206; Müller & Frings 2009, 461; Strenzke et al. 2008, 28; Yost 2008, 110; für detaillierte Ausführungen zur lateralen Zellenwand vgl. Iwasa 2010, 181f.).

2.4.5 Lymphflüssigkeiten

Das Innenohr besteht mit dem *Perilymphraum* und dem *Endolymphraum* aus zwei Flüssigkeitsräumen, deren mechanische und elektrochemische Interaktion entscheidend für den Transduktionsprozeß im Innenohr ist (vgl. Marcus & Wangemann 2010, 213).

Der Perilymphraum entspricht weitestgehend dem knöchernen Labyrinth abzüglich des häutigen Labyrinths, welches seinerseits den Endolymphraum bildet (vgl. Gelfand 2004, 51; Hellbrück & Ellermeier 2004, 93; Reiss 2003, 591, 612; Schünke et al. 2009, 134; Siebenmann 1897, 298). Die Perilymphe ist eine natriumreiche und kaliumarme Flüssigkeit. Im Gegensatz dazu ist die Endolympe

bemerkenswert kaliumreich und natriumarm. Die Kaliumionen sind der entscheidende Ladungsträger für das endokochleäre Transduktionspotential (vgl. Gelfand 2004, 52; 2009, 54; Hellbrück & Ellermeier 2004, 93; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Marcus & Wangemann 2010, 213, 215; Yost 2008, 87).

Auch mechanisch spielen die Lymphflüssigkeiten eine tragende Rolle, nämlich bei der Übertragung der Steigbügelbewegungen ans Cortiorgan durch die Bildung einer Wanderwelle entlang der BM. Die physikalischen Eigenschaften der Lymphflüssigkeiten werden gewöhnlich mit jenen von Wasser verglichen, ihre Viskosität ist demnach äußerst gering (vgl. Gelfand 2004, 87; Terhardt 1998, 56). Ranke gibt nach Békésy für die Perilymphe mit 0,0197 jedoch ungefähr die doppelte Viskositätskonstante von Wasser bei 37°C (0,0101) an (vgl. Ranke 1953, 99).

Die Endolymphe gleicht nach Breschet beim Menschen „ganz der Perilymphe, ist so klar, wie das reinste Wasser und ebenso dünnflüssig“ (zit. nach Dieffenbach et al. 1837, 56). Erwähnt werden muß, daß man zu Zeiten Breschets die Schnecke noch in zwei Treppen geteilt sah und seine Endolympfwerte im Gleichgewichtsorgan gewonnen wurden (vgl. Dieffenbach et al. 1837, 53). Ansonsten wird die Endolymphe, sofern beide Lymphen (rein mechanisch) überhaupt differenziert betrachtet werden, oft als hoch viskos, fast gallertartig, beschrieben (vgl. Békésy 1941(b), 275; 1960, 407; Békésy & Rosenblith 1951, 1077; Spitzer 2002, 60). Nach Ranke ist sie ungefähr 2,9 mal so zäh wie Wasser (vgl. Ranke 1953, 107). Überdies ist anzunehmen, daß Békésy gute Gründe hatte, die Endolymphe in Modellversuchen mittels hoch viskoser Flüssigkeiten wie Glycerin oder Honig zu simulieren (vgl. Békésy 1960, 423; Zwislocki 1948, 15f.).

Elektrolyte	Perilymphe	Cortilymphe	Endolymphe
Kalium (K ⁺)	10 mval/l	15 mval/l	144 mval/l
Natrium (Na ⁺)	140 mval/l	135 mval/l	5 mval/l
Proteine	200–400 mg%	200–300 mg%	126 mg%

Tab. 2.2 **Zusammensetzung der Lymphflüssigkeiten**

(nach Graumann & Sasse 2005, Tab. 1.8, 115)

Die Produktion der *Perilymphe* ist noch umstritten, wahrscheinlich entsteht sie aber an Blut-Perilymph-Schranken im Bereich des Limbus und der Stria vascularis als Ultrafiltrat (Plasmafiltrat) des Blutes (vgl. Gelfand 2004, 52; Graumann & Sasse 2005, 113; 129; Schriever 1967, 36; Schünke et al. 2009, 134). Sie zeichnet sich in ihrer Zusammensetzung – wie die meisten extrazellulären Flüssigkeiten – durch eine hohe Natriumionenkonzentration (Na⁺) und eine geringe Kaliumionenkonzentration (K⁺) aus (vgl. Fuchs 2010, 6; Gelfand 2004, 52; 2009, 54; Hellbrück & Ellermeier 2004, 93; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Plattig in Keidel 1975, 93; Schriever 1967, 35; Yost 2008, 87).

Auch das Cortiorgan ist mit Lymphe gefüllt. Hierzu finden sich in der Literatur unterschiedliche Annahmen: Manche Autoren finden eine Flüssigkeit, deren Zusammensetzung der perilymphatischen stark ähnelt aber eben nicht völlig gleicht, so daß von *Cortilymphe* die Rede ist (siehe Tab. 2.2; vgl. Graumann & Sasse 2005, 119; Reiss 2003, 621; Schriever 1967, 38; Yost 2008, 88). Andere fassen die Cortilymphe hingegen als („reinrassige“) Perilymphe auf (vgl. Fuchs 2010, 6; Gelfand 2004, 52; 2009, 57). Gelfand weist sogar darauf hin, daß bezüglich der Flüssigkeit im Cortiorgan in der *Vergangenheit* von

Cortilymphe gesprochen wurde (vgl. Gelfand 2004, 52). Da erst die *Membrana reticularis* eine galvanische Barriere zwischen Peri- und Endolymphraum bietet, kann das Cortiorgan als Teil des Perilymphraums verstanden werden (vgl. Gelfand 2004, 52, 57; 2009, 56f.; Graumann & Sasse 2005, 118; Reiss 2003, 62ff.; Siebenmann 1897, 312). Ob nun Corti- oder Perilymphe ist funktionell gar nicht so entscheidend, weil klar ist, daß die cortische Lymphflüssigkeit als Gegenpol zur Endolymph am Haarzellenkörper und aufgrund der blank durch den Cortitunnel führenden Nervenfasern prinzipiell perilymphartig (also extrazellulären Flüssigkeiten entsprechend) zusammengesetzt sein muß (vgl. Fuchs 2010, 6; Schriever 1967, 38).

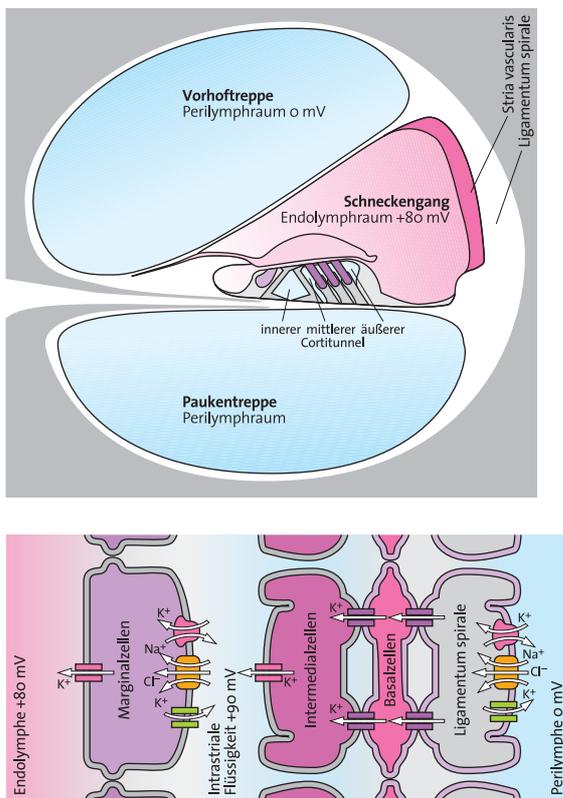


Abb. 2.15 **Lymphflüssigkeitssystem der Schnecke.** Oben: Lymphräume. Der Schneckenkanal abzüglich des Cortiorgans entspricht dem Endolymphraum, da die *Membrana reticularis* und nicht die Basalmembran die galvanische Barriere zwischen den Lymphräumen ist. Unten: Generierung des endokochleären Potentials in der *Stria vascularis* (modifiziert n. Marcus & Wangemann 2010, Abb. 7.3, 218). Die Zellen des *Ligamentum spirale* und die Schichten der *Stria vascularis* transportieren Kaliumionen aus der Peri- zur Endolymph. Die Basalzellen bilden die Grenzmembran, an der das endokochleäre Potential zwischen Perilymphraum und interstrialem Lymphraum (endolymphähnlich) entsteht.

Über den unweit des runden Fensters durch das *Schneckenkanälchen* (*Canaliculus cochleae*, *Aqueductus cochleae*) austretenden *Perilymphgang* (*Ductus perilymphaticus*) fließt die Perilymphe schließlich in den Subarachnoidalraum ab (vgl. Fanghänel et al. 2003(b), 199; Graumann & Sasse 2005, 122, 129; Helmholtz 1896, 22ff.; Reiss 2003, 615; Schünke et al. 2009, 134; Stevens & Hallowell 1938, 269). Dort mischt sie sich entweder mit dem Hirnwasser (*Liquor cerebrospinalis*) oder wird in der äußeren Hirnhaut (*Dura mater*) resorbiert (vgl. Graumann & Sasse 2005, 129). Für Resorption spricht, daß im Schneckenkanälchen mittels einer feinen Membran keine direkte Flüssigkeitskommunikation möglich ist (vgl. 2004, 52; Schriever 1967, 36).

Die *Endolymph* stellt ein Filtrat der Perilymphe dar, wird wahrscheinlich hauptsächlich, aber nicht ausschließlich, in der *Stria vascularis* produziert und im Endolymphsack resorbiert (vgl. Fuchs 2010, 6; Gelfand 2004, 59; Graumann & Sasse 2005, 129; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003, 615, 618, 620; Schriever 1967, 37; Yost 2008, 87).

Die konkave Wand des Schneckenkanals ist mit dem *Ligamentum spirale cochleae* großzügig ausgekleidet (vgl. Gelfand 2004, 55; Reiss 2003, 620; Spee 1896, 307). Dabei handelt es sich um eine der Knochenhaut aufsitzende, gefäßdurchzogene, schwammige Struktur, die mit quasi-Perilymphe durchschwemmt ist (vgl. Marcus & Wangemann 2010, 219; Reiss 2003, 620).

Längs zwischen Reissner- und Basilarmembran differenziert sich vom Ligamentum spirale ein mehrschichtiges Epithel, die *Stria vascularis*, und bildet die laterale Wand des Schneckengangs (vgl. Fuchs 2010, 6; Gelfand 2004, 59; Graumann & Sasse 2005, 114; Reiss 2003, 620; Yost 2008, 86f.). Sie ist mit Basal-, Intermediär- und Marginalzellen dreischichtig aufgebaut und von zahlreichen Kapillaren penetriert (vgl. Marcus & Wangemann 2010, 219; Reiss 2003, 620; Yost 2008, 87). Basal- und Marginalzellen formen die Diffusionsbarrieren, welche den intrastrialen Flüssigkeitsraum gegen das perilymphdurchschwämmte Ligamentum spirale (Basalzellen) bzw. den mit Endolymphe gefüllten Schnecken gang (Marginalzellen) abdichten. Die basale Membran bildet dabei jene Barriere, an der das endokochleäre Potential gebildet wird (vgl. Gelfand 2004, 59; Marcus & Wangemann 2010, 219; Reiss 2003, 620; Schriever 1967, 41). Die Basalzellen absorbieren K^+ -Ionen aus der Perilymphe und führen sie an die intermediäre Schicht zur „Zwischenlagerung“ ab. Von hier wird Kalium an den interstitialen Flüssigkeitsraum abgegeben (dementsprechend endolymphähnlich; vgl. Marcus & Wangemann 2010, Abb. 7.3, 218), woraus sich nun die Marginalzellen bedienen, um K^+ letztlich in die Endolymphe zu sekretieren (vgl. dies. 2010, 219).

Die Endolymphe fließt über den mit Sacculus und Utriculus verbundenen Endolymphgang (Ductus endolymphaticus) ab, wo sie im Endolymphsack (Saccus endolymphaticus) – einer an der rückwärtigen Felsenbeinaußenwand epidural (zwischen den Blättern der Dura mater) gelegenen, trichterförmigen Tasche – ähnlich wie von der Perilymphe angenommen, resorbiert wird (vgl. Gelfand 2004, 52; 2009, 55; Graumann & Sasse 2005, 129; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Reiss 2003, 615, 618; Schünke et al. 2009, 134).

Kochleäre Potentiale

Das Ruhepotential der Schnecke entsteht aus dem ungefähr 80 mV starken Spannungsgefälle zwischen Endolymphe und Perilymphe. Nomenklatorisch wird auch der Begriff „Bestandstrom“ verwendet, obwohl es sich nicht um Strom, sondern um Spannung handelt. Die positive Ladung der Endolymphe (Kaliumionen) wird als *endokochleäres Potential* bezeichnet (vgl. Gelfand 2004, 135ff.; 2009, 57; Hellbrück & Ellermeier 2004, 94; Iwasa 2010, 180; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Müller & Frings 2009, 460f.; Ranke 1953, 110; Schriever 1967, 41). Es entsteht an der basalen Membran der *Stria vascularis* (vgl. Gelfand 2004, 59; Marcus & Wangemann 2010, 219; Reiss 2003, 620; Schriever 1967, 41) und weist entlang des Schneckengangs Schwankungen im Bereich von rund 3 mV auf (vgl. Ranke 1953, 112).

Ein weiteres Ruhepotential findet sich im Spannungsverhältnis von Perilymphe zum Inneren der Haarzellen. Dieses *intrazelluläre Potential* beträgt grob -40 bis -90 mV, jedoch für IHZ und ÄHZ unterschiedlich stark. Während in IHZ Werte um die -40 bis -50 mV gemessen werden, scheinen die ÄHZ mit um die -70 bis -90 mV deutlich stärker negativ geladen zu sein (vgl. Gelfand 2004, 137; 2009, 57; Graumann & Sasse 2005, 119; Iwasa 2010, 180; Plattig in Keidel 1975, 91; Schriever 1967, 40).

Békésy hat die Potentiale innerhalb der Schnecke zuerst gemessen, indem er eine Elektrode eines Voltmeters in die Perilymphe (0 mV) gesetzt hat und mit der anderen durch die Schichten der Schnecke gefahren ist (vgl. Gelfand 2004, 137). In seinem Windschatten hat eine Vielzahl von Messungen statt-

gefunden, die zwar einen reichen Strauß an Zahlen hervorbrachten, aber trotzdem ein recht klares Bild erkennen lassen: Die Endolymphe weist, wie bereits dargestellt, ein positives Potential um 80 mV auf, dessen Werte insgesamt zwischen 50 und 100 mV angegeben werden (vgl. Gelfand 2004, 135ff.; Hellbrück & Ellermeier 2004, 94; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 35; Müller & Frings 2009, 460f.; Plattig in Keidel 1975, 97).

Im Cortiorgan wurden negative Werte zwischen -50 und -70 mV gemessen (vgl. Gelfand 2004, 137; Plattig in Keidel 1975, 97). Daß die Stereozilien der Haarzellen in Endolymphe baden und die Zellkörper jenseits der Membrana reticularis von perilymphähnlicher Flüssigkeit umschwemmt sind, ist die entscheidende Voraussetzung für die *elektromechanische Reizwandlung* (siehe 3.3.4, S. 131; vgl. Graumann & Sasse 2005, 119; Stevens & Hallowell 1938, 273). Dementsprechend ist das intrazelluläre Potential bezogen auf die Endolymphe natürlich – im wahrsten Sinne des Wortes – spannender als auf die Perilymphe bezogen. Intrazelluläre Potentiale zur Endolymphe werden mit Werten zwischen -100 und -130 mV angegeben (vgl. Gelfand 2004, 137; Ranke 1953, 112). Diese indifferenten Werte beziehen sich vermutlich auf IHZ, da deren Potential mit etwa -127 mV im passenden Bereich liegen (vgl. Graumann & Sasse 2005, 119), während Potentiale der ÄHZ um -155 mV liegen dürften (vgl. Graumann & Sasse 2005, 119). Bezogen auf die Perilymphe leiten sich daraus Ruhepotentiale von ungefähr -40 mV (IHZ) bzw. -70 mV (ÄHZ) ab (vgl. Gelfand 2004, 135ff.; Graumann & Sasse 2005, 119; Schriever 1967, 41).

Wenngleich sich ein klares Bild der Spannungslage innerhalb der Schnecke abzeichnet, ist es doch wenig sinnvoll, Ansprüche mathematischer Genauigkeit zu erheben, weil Schwankungen innerhalb eines Pols (wie des endokochleären Potentials entlang der Schnecke) genauso vorkommen können wie interindividuelle Unterschiede und nicht zuletzt auch Meßungenauigkeiten aufgrund der Winzigkeit der zu messenden Elemente.

Physiologie

3 Wie arbeitet das Ohr?

Die Standarddarstellung des Hörvorgangs liest sich in etwa so: Schalldruckschwankungen des Mediums (meistens Luft) regen am Ende des Gehörgangs das Trommelfell zum Schwingen an, wodurch die Gehörknöchelchenkette im Mittelohr angestoßen wird. Sie überträgt die luftkonformen Trommelfellschwingungen auf die Innenohrflüssigkeiten. Dort, am Ende der Schalleitung, sorgen komplexe mechanische Vorgänge dafür, daß die mit (rein sensorischer) Schallempfindung betrauten Haarzellen erregt werden und den Transduktionsprozeß in Gang setzen (vgl. Fuchs 2010, 2, 4; Spitzer 2002, 78; Winckel 1952, 55f.; 1960, 79f.; Zwicker & Feldtkeller 1967, 83, 85).

Das Ohr (*peripheres auditives System*) ist aber nicht einfach nur ein Mikrofon, das mit einem Kabel (Hörnerv) an einen Prozessor (Gehirn) angeschlossen ist. Es ist ein kompliziertes und raffiniertes Organ. Wer das Gehör verstehen will, muß im Grunde schon mit dem Oberkörper anfangen. Der Hals beeinflusst das Schwingverhalten des Kopfs (vgl. Békésy 1960, 164), Kopf und Torso produzieren Schallabschattungen bzw. -reflexionen und die Anordnung der Ohren am Kopf ist (zusammen mit den genannten Dämpfungen und Reflexionen) für räumliche Schallwahrnehmung von tragender Bedeutung (vgl. Blauert 1974; Gelfand 2004, 85; Guski 1996, 88; Yost 2008, 71f.; Young 2007, 431). Die Ohren (von den tatsächlichen Rezeptoren im Cortiorgan ganz zu schweigen) bekommen bereits gehörig manipulierten (allerdings dadurch mit verwertbaren Informationen angereicherten) Schall „zu hören“.

Im Mittelohr wird der Schall an die Bedingungen im Innenohr angepaßt und dadurch ebenfalls verzerrt (vgl. Lehnhardt & Laszig 2009, 27; Spitzer 2002, 57; Terhardt 1998, 56; Yost 2008, 75; Young 2007, 429, 432). Mit dem *akustischen Reflex* steht dem Mittelohr ein dem Pupillenreflex vergleichbarer Mechanismus zur Verfügung (vgl. Spitzer 2002, 58), über den in die Schalleitung eingegriffen werden kann (vgl. Békésy & Rosenblith 1951, 1085; Blauert 1974, 43; Gelfand 2004, 112; Schünke et al. 2009, 133; Yost 2008, 77, 79). Das Innenohr ist ein harmonischer Analysator, der Schall spektral zerlegt und in Nervenimpulse übersetzt (vgl. Blauert 1974, 116; Gelfand 2004, 121; Pollard & Jansson 1982, 251; Spitzer 2002, 58). Auch die Spektralanalyse des Innenohrs wird über den kochleären Verstärker aktiv gesteuert (vgl. Hellbrück & Ellermeier 2004, 103; Janssen in Lehnhardt & Laszig 2009, 113f.; Müller & Frings 2009, 460; Spitzer 2002, 62f.; Yost 2008, 111f.). Hören ist also bereits ein aktiver Prozeß, bevor der Schall das neuronale System erreicht hat (vgl. Spitzer 2007, 63).

3.1 Kopfbezogene Übertragungsfunktion

Schon bevor also Schall die Ohren erreicht, wird er durch Kopf und Körper erheblich manipuliert, indem diese den Schall richtungs- und frequenzabhängig dämpfen. Diese linearen Verzerrungen hängen auch mit binauralem Hören zusammen, da der Kopf beide Ohren räumlich voneinander trennt und zudem gegeneinander abschattet.

Binaurales Hören ist mit binokularem Sehen vergleichbar. Divergente Perspektiven beider Ohren resultieren in unterschiedlichen Ohrsignalen (*interaurale Differenzen*), woraus bereits Richtungsinfo-

mationen gewonnen werden (vgl. Blauert 1974, 56, 59, 77; Gelfand 2004, 85; Guski 1996, 88; Hellbrück & Ellermeier 2004, 151; Müller & Frings 2009, 464; Yost 2008, 71, 180; Young 2007, 431). Da interaurale Differenzen naturgemäß bereits vor dem Eintreffen an den Ohren entstehen müssen, spiegeln sie vor allem den verzerrenden Einfluß von Kopf und Körper wider (vgl. Blauert 1974; Rosowski, 2010, 62; Yost 2008, 72, 178).

Einen weiteren Filter bildet das Außenohr selbst, wo zusätzliche richtungs-, entfernungs- und frequenzabhängige Verzerrungen stattfinden (vgl. Békésy & Rosenblith 1951, 1078; Blauert 1974, 50f., 55; Yost 2008, 178). Faßt man die Übertragungsfunktionen von Kopf, Körper und Außenohr zusammen, erhält man ein richtungsempfindliches Filtersystem, das als *kopfbezogene Übertragungsfunktion* (KobeÜF, head-related transfer function, HRTF) bezeichnet wird (vgl. Blauert 1974, 59f.; Culling & Akeroyd 2010, 128; Gelfand 2009, 47; Hellbrück & Ellermeier 2004, 152; Rosowski 2010, 49, 61f.; Yost 2008, 72f., 178f., 183).

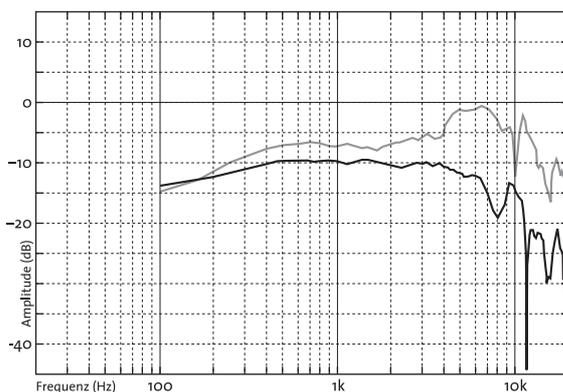


Abb. 3.1 **kopfbezogene Übertragungsfunktion** (Yost 2008, Abb. 6.5, 72). Die Kurven zeigen die verschiedenen Ohrsignale von linkem (graue Linie) und rechtem Ohr im Gehörgang gemessen bei Beschallung von 270° Azimut.

schwächer. Wegen des Größenverhältnisses von Hindernis (Kopf) und Wellenlänge werden mittlere und hohe Frequenzen stärker gedämpft als tiefere (vgl. Blauert 1974, 59; Gelfand 2004, 85, 87; Müller & Frings 2009, 449; Rosowski, 2010, 62; Yost 2008, 72, 178).

3.1.1 Kopf und Körper

Der Kopf durchmißt von Ohr zu Ohr zwischen 13,5 und 18,5 cm (vgl. Blauert 1997, 301; Guski 1996, 88; Spitzer 2002, 65; Terhardt 1998, 53). In vielen Überlegungen ist der Kopf durch eine starre Kugel mit Durchmesser (Ohrabstand) zwischen 22 und 23 cm ersetzt (vgl. Gelfand 2004, 397). Kugelmodelle sind geeignet um tendenzielle und vereinfachte Kenntnisse des Kopfeinflusses zu gewinnen, versagen jedoch – wenig überraschend – bei genauer Beschreibung der kopfbezogenen Verzerrungen (vgl. Gelfand 2004, 399). Dies hat mehrere Gründe. Zunächst ist der Kopf keine Kugel. Außerdem ist er in keiner Richtung vollkommen symmetrisch, so daß auch die Medianebene keine Symmetrieebene dar-

Die Filterwirkung eines Übertragungssystems ergibt sich aus der Differenz von Eingangs- und Ausgangssignal. Es wird also das direkte Signal der Schallquelle (Freifeldschall, im Idealfall an der Stelle des Trommelfells, wenn der Hörer nicht da wäre) mit dem Schall im Gehörgang abgeglichen (vgl. Blauert 1974, 62; Hellbrück & Ellermeier 2004, 151f.; Yost 2008, 71f.; Young 2007, 429). Dabei ergeben sich z.B. für ein von 270° Azimut kommendes Schallsignal Kurven, wie sie in der nebenstehenden Abbildung zu sehen sind. Durch den Abstand beider Ohren zueinander und den Schallschatten des Kopfs dazwischen erreicht der Schall das abgewandte Ohr (außer er kommt direkt von 0 bzw. 180° Azimut) später und

stellt. Auch die Ohren liegen nicht auf der Horizontalachse, sondern etwas weiter posterior (vgl. Blauert 1974, 145; 997, 305; Culling & Akeroyd 2010, 127; Gelfand 2004, 399). Die links/rechts-Asymmetrie des Kopfs spielt keine erhebliche Rolle, da individuelle anatomische Nuancen (auch des Außenohrs selbst) eine erlernte akustische Normalität darstellen, die erst dann ihren Status als (selbstverständlich subjektive) „absolute Wahrheit“ verlieren, wenn sich die effektive Anatomie ändert (z.B. durch Unfälle, Krankheiten oder gewollte Manipulation wie etwa mit Knetmasse) und man durch ungewohnte oder fremde kobeÜF (dann über Kopfhörer) gefilterte Signale lokalisieren muß, wobei auch hier Lerneffekte zu beobachten sind (vgl. Culling & Akeroyd 2010, 128; Gelfand 2004, 407; Spitzer 2002, 66; Wightman & Kistler 1993, 168, 187).

Die Anordnung der Ohren hat direkten Einfluß auf das räumliche Auflösungsvermögen. Da sich das Leben an Land normalerweise knapp über dem Boden abspielt, ist das Gehör des Menschen (sowie der meisten hörfähigen Landlebewesen) für diese Situation optimiert. Beide Ohren sind auf der Horizontalachse voneinander distanziert und nach vorne gerichtet. Dadurch ist erstens die horizontale Raumauflösung (hier wirken die interauralen Differenzen) wesentlich besser als die vertikale (vgl. Blauert 1974, 35; Culling & Akeroyd 2010, 127; Guski 1996, 273f.; Yost 2008, 178, 180) und zweitens die Horizontalebene ihrerseits vorne besser als hinten (die Ohren sind nach vorne gerichtet) und seitlich am schlechtesten (weil die Ohren von hier gesehen hintereinander liegen) aufgelöst (vgl. Blauert 1974, 31f.; Culling & Akeroyd 2010, 127; Gelfand 2004, 409; Guski 1996, 273; Spitzer 2002, 71; Yost 2008, 176). Die schwache Wirkung interauraler Unterschiede seitlich des Kopfs drückt sich in einem „Unschärfekegel“ (*cone of confusion*) aus, innerhalb dessen veränderte Schalleinfallswinkel keinen wahrnehmbaren Einfluß auf den Ort des Hörereignisses haben (vgl. Gelfand 2004, 409; Wightman & Kistler 1993, 164).

Interaurale Differenzen

Trifft Schall direkt von der Seite ein, erreicht er das gegenüberliegende Ohr später und schwächer (vgl. Blauert 1974, 59; 1997, 72; Culling & Akeroyd 2010, 123; Hellbrück & Ellermeier 2004, 149; Müller & Frings 2009, 449, 464; Yost 2008, 174). Später, weil beide Ohren um den Kopfdurchmesser voneinander entfernt liegen und der Schall zum abgelegenen Ohr somit entsprechend länger unterwegs ist und schwächer, weil der Kopf zwischen beiden Ohren als akustisches Hindernis fungiert, das den Schall frequenzabhängig mehr oder weniger stark abschattet. Je nach Wirkung werden die resultierenden Unterschiede der Ohrsignale als *interaurale Zeit-* (IZD) respektive *Pegeldifferenzen* (IPD) klassifiziert (vgl. Blauert 1974, 113ff., 126ff., 132; 1997, 138; Culling & Akeroyd 2010, 123; Wightman & Kistler 1993, 159, 164; Yost 2008, 173f.).

Da Schallwellen unter 1,5 kHz länger sind, als der Kopf groß ist, werden tiefe Frequenzen durch den Kopf nicht mehr effektiv gedämpft, sondern um ihn gebeugt. Die resultierenden Wegdifferenzen werden als interaurale Phasenunterschiede in der Richtungsbeurteilung verarbeitet (vgl. Gelfand 2004, 397f., 402; 2009, 49; Stevens & Hallowell 1938, 171f.; Wightman & Kistler 1993, 164). Sie erreichen ihr Maximum, wenn die Schallquelle direkt links oder rechts des Kopfs liegt. Die maximale Differenz ist

der Quotient aus Kopfdurchmesser und Schallgeschwindigkeit (vgl. Spitzer 2002, 70). Da sowohl Dividend als auch Divisor innerhalb gewisser Grenzen variieren, sind unterschiedliche Messungen und/oder Berechnungen der maximalen IZD nicht verwunderlich. Naturgemäß ist sie bei rund 90° bzw. 270° Azimut erreicht. Die hierzu angegebenen Werte schwanken von 0,5 bis 0,8 ms (vgl. Blauert 1974, 115; 1997, 143; Culling & Akeroyd 2010, 123; Gelfand 2004, 398; Wightman & Kistler 1993, 159, 169; Spitzer 2002, 70; Yost 2008, 174). Für hohe Frequenzen beträgt die maximale Laufzeitdifferenz allerdings nur 0,7 ms. Erst ab 500 Hz und tiefer werden Unterschiede bis 0,82 ms erreicht. Dies deutet – vereinfacht interpretiert – eine Frequenzabhängigkeit der Laufzeitunterschiede an (vgl. Gelfand 2004, 399). Künstliche Vergrößerungen darüber hinaus vermögen die natürlichen Grenzen zwar ein wenig zu strecken, haben aber keinen nennenswerten Einfluß auf den Azimut des Hörereignisses (vgl. Blauert 1974, 115; 1997, 143).

Wichtig ist es, zu bedenken, daß Phasenverschiebungen zweideutig sein können. Ist eine Periode von 1 ms z.B. linksöhrig um 0,7 ms verzögert (entspricht etwa 90° Azimut), kann dies so, oder als rechtsöhrige Verzögerung um 0,3 ms (rund 310° Azimut) verstanden werden. Da das Gehör normalerweise die kürzere Latenz bevorzugt, sollte eine halbe Signalperiode die maximale IZD überdauern (vgl. Blauert 1974, 119; Stevens & Hallowell 1938, 171f.; Wightman & Kistler 1993, 178f.; Yost 2008, 177).

Zieht man Kopfdurchmesser von 17 bis 23 cm und eine Schallgeschwindigkeit von 343 m/s heran, kommt man auf Frequenzen zwischen 0,8 und 1 kHz, bei denen die halbe Periode Verzögerungen entspricht, die Phasenverschiebungen doppeldeutig machen. Ab rund 1,3 bis 2 kHz, wo die Gesamtperiode der maximalen IZD entspricht, sind Phasenverschiebungen praktisch unbrauchbar. Statt der mikrostrukturellen Phasenverschiebung einzelner Teiltöne wird ab hier nur noch der reelle Laufzeitunterschied der Signalhüllkurve, sozusagen die Makrostruktur, verwertet (vgl. Blauert 1974, 119ff., 132; 1997, 149, 164f., 311f.; Culling & Akeroyd 2010, 123; Gelfand 2004, 398, 402, 405; Spitzer 2002, 70; Spreng in Keidel 1975, 137; Stevens & Hallowell 1938, 172; Yost 2008, 182). Dabei ist nicht unbedingt die Gesamthüllkurve, sondern eher jene einzelner „Frequenzgruppen“ (hier: Gruppenthüllkurven, nicht mit den kritischen Bändern der Psychoakustik zu verwechseln; siehe 1.3, S. 18) ausschlaggebend (vgl. Blauert 1974, 122).

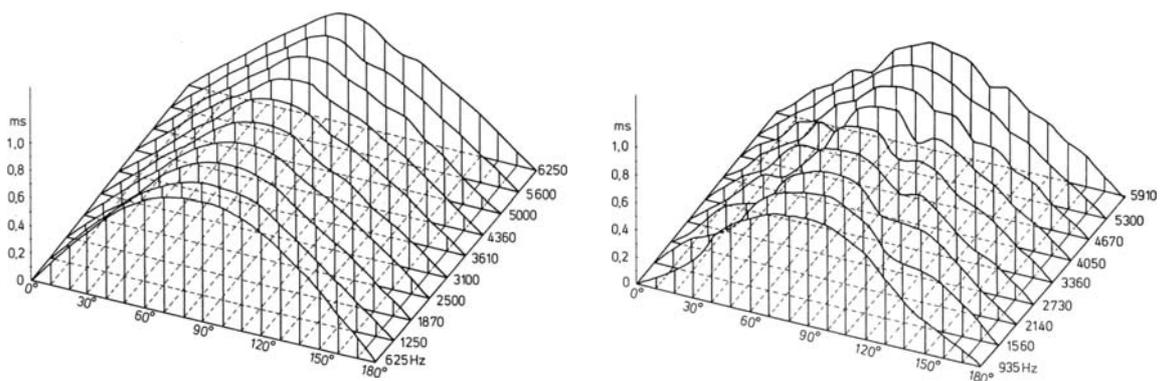


Abb. 3.2 interaurale Laufzeitdifferenzen (Blauert 1974, Abb. 44 & 45, 58). Auf der Ordinate (Y-Achse) ist die Laufzeit abgetragen, auf der Abszisse (X-Achse) der Seitenwinkel (Azimut) und auf der Applikate (Z-Achse) die Frequenz. Links: interaurale Phasenlaufzeitdifferenzen, rechts: interaurale Gruppenlaufzeitdifferenzen.

Enthält das Spektrum sowohl hohe als auch tiefe Anteile, werden Phasen- *und* Gruppenlaufzeit ausgewertet. Im hochfrequenten Bereich, wo praktisch alle Wellenlängen kleiner als der Kopf sind, haben die Hüllkurven breiterer Frequenzbänder größere Bedeutung, die sie aber mit sinkender Trägerfrequenz verlieren, weil die einzelnen Wellen länger als der Kopf sind und so bereits Einzelfrequenzen verwertbare Information transportieren. In aller Regel laufen Phasen und Gruppen nicht gleich, wobei im Konfliktfall der durch die Hüllkurve bestimmte Anteil dominiert (vgl. Blauert 1974, 123, 132; 1997, 164f., 311f.; Culling & Akeroyd 2010, 123; Gelfand 2004, 407).

Neben Zeitdifferenzen stehen dem Gehör auch *interaurale Pegeldifferenzen* zur Verfügung. Sie entstehen durch den Schallschatten des Kopfs (vgl. Culling & Akeroyd 2010, 123, 126; Gelfand 2004, 397f.; 2009, 48; Guski 1996, 88; Rosowski, 2010, 62; Stevens & Hallowell 1938, 169; Yost 2008, 72, 173f.). Diese IPD sind stark frequenzabhängig und betreffen besonders hohe Frequenzen. Sie nehmen mit steigender Frequenz zu und erreichen um 6 kHz ein Maximum von 20 bis 30 dB, wobei in natürlichen Hörsituationen maximal 10 dB erreicht werden (vgl. Blauert 1974, 126; Gelfand 2004, 398; Hellbrück & Ellermeier 2004, 150; Wightman & Kistler 1993, 164, 169).

Dies liegt daran, daß hohe Frequenzen, deren Wellen kürzer als der Kopfradius sind, von diesem wirkungsvoll abgeschattet werden, wodurch die Schallintensität am entfernt liegenden Ohr signifikant abnimmt (vgl. Blauert 1974, 59; 1997, 72; Culling & Akeroyd 2010, 123; Gelfand 2004, 397f.; 2009, 48; Hellbrück & Ellermeier 2004, 149; Spitzer 2002, 65; Yost 2008, 174, 177). Frequenzen unter 1,5 bis 1,6 kHz übertreffen mit ihren Wellenlängen den Kopfdurchmesser. Dadurch fallen Pegeldifferenzen nach unten immer geringer aus, bis sie unter 200 Hz praktisch verschwinden. Folglich klingt Schall am abgewandten Ohr dumpfer (vgl. Gelfand 2004, 85, 398; 2009, 48; Guski 1996, 88; Wightman & Kistler 1993, 169).

Pegelunterschiede sind bei entsprechend breitem Spektrum also zwingendermaßen gleichzeitig auch *interaurale Klangfarbenunterschiede* (IKD), die dem Gehör zusätzlich zur Verfügung stehen (vgl. Lehnhardt & Hesse 2009, 188f.). Ist das Signal schmalbandig, so daß alle Teiltonamplituden (im Extremfall Sinuston nur eine) gedämpft werden, wirkt allein der Pegelunterschied. Verfügt das Spektrum auch über Anteile unterhalb der kritischen 1,5 bis 1,6 kHz, resultiert die Dämpfung der höheren Teiltöne einerseits in IPD; andererseits, wegen des manipulierten Amplitudenverhältnisses von hohen zu tiefen Spektralbereichen, auch in IKD.

Da der Kopf unter 1,6 kHz keine ausreichende Abschattung mehr bietet, sind IPD an sich nur für höhere Frequenzen von Bedeutung (vgl. Culling & Akeroyd 2010, 126; Spitzer 2002, 65; Wightman & Kistler 1993, 169; Yost 2008, 72). Die Feststellung, daß IPD über das gesamte hörbare Spektrum wirksam, aber bei größeren Bandbreiten mit signifikanten Spektralanteilen über 1,6 kHz größer und am verlässlichsten sind (vgl. Blauert 1974, 132, 134, 139; 1997, 165; Wightman & Kistler 1993, 170), müßte demnach eher (wenn nicht ausschließlich) für IKD gelten, da IPD bei hohen und tiefen Spektralanteilen immer auch IKD produzieren.

IPD sind auch richtungsabhängig. Wie alle interauralen Differenzen erreichen sie bei zirka 90° bzw. 270° Azimut ihr (frequenzabhängiges) Maximum von 20 bis 30 dB, wobei sie in natürlichen Hörsituationen maximal 6 bis 12 dB erreichen (vgl. Gelfand 2004, 398f.; Hellbrück & Ellermeier 2004, 150; Keidel 1975, 384; Wightman & Kistler 1993, 164, 169). Bei IPD oberhalb von 20 dB nimmt die Lokalisationsunschärfe (man bedenke den Unschärfekegel) zu, so daß der Grenzwert „ganz seitlich“ nur

schwer meßbar ist (vgl. Blauert 1974, 176). Interessanterweise wurde bei maximalem Winkel (90° bzw. 270° Azimut) am abgewandten Ohr ein Verstärkungseffekt gemessen. Der Schall war dort zwar immer noch deutlich schwächer als vor dem anderen Ohr, aber stärker als im freien Schallfeld. Dies ist vermutlich auf Beugungsinterferenzen zurückzuführen und folglich am dem Körper aufsitzenden Kopf

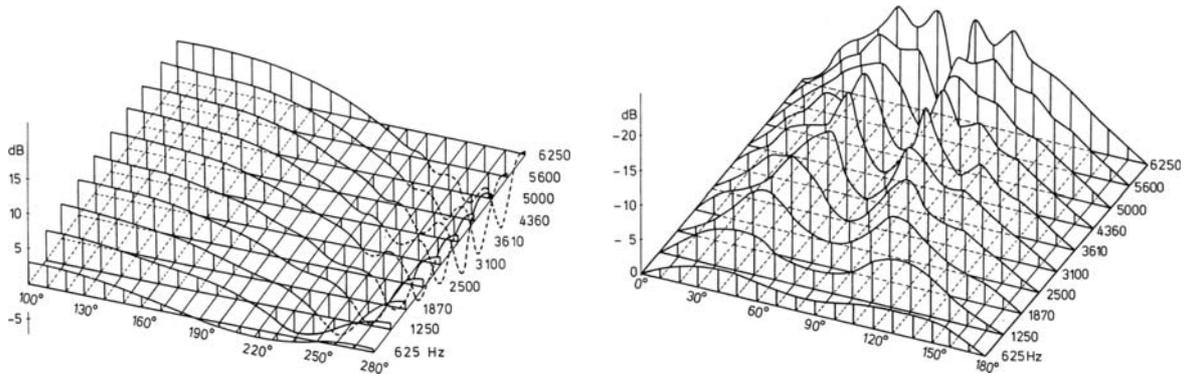


Abb. 3.3 **interaurale Pegeldifferenzen** (Blauert 1974, Abb. 42 & 43, 57). Auf der Ordinate (Y-Achse) ist der Pegelunterschied abgetragen, auf der Abszisse (X-Achse) der Seitenwinkel (Azimut) und auf der Applikate (Z-Achse) die Frequenz. Links: Schalldruck am linken Ohr bezogen auf den Schalldruck am Kugelmittelpunkt bei Abwesenheit der Kugel (Ohr bei 100° Azimut). Rechts: Pegeldifferenz zwischen linkem und rechtem Ohr (Ohren bei 100° und 260° Azimut).

wohl weniger ausgeprägt als an der freien Kugel. Mitunter tragen jedoch auch Körperreflexionen zu einem solchen Verstärkungseffekt bei (vgl. Blauert 1974, 59; 1997, 72; Culling & Akeroyd 2010, 126). Höchstwahrscheinlich läßt sich das Zusammenwirken verschiedener interauraler Differenzen folgendermaßen zusammenfassen: Liegt das gesamte Klangspektrum unterhalb von 1,6 kHz, sind lediglich die interauralen Phasenlaufzeiten (IZD) ausschlaggebend. Liegen nennenswerte Anteile des Spektrums oberhalb von 1,6 kHz, sind spektrale Unterschiede in Form von Hüllkurvenverschiebungen (IZD) und Intensitätsunterschieden (IPD) ausschlaggebend (vgl. Blauert 1974, 137ff.) sowie – bei Spektren beiderseits von 1,6 kHz – auch Klangfarbenunterschiede. In „natürlichen“ Hörsituationen steht dem Gehör in aller Regel die gesamte Palette interauraler Signaldifferenzen zur Verfügung (vgl. Blauert 1974, 132; Lehnhardt & Hesse in Lehnhardt & Laszig 2009, 189). Besonders im Frequenzbereich von ungefähr 1 bis 4 kHz bzw. 1,5 bis 2 kHz um die Grenzfrequenz 1,6 kHz sind IZD und IPD unklar und uneindeutig, wodurch es vermehrt zu Lokalisationsfehlern, speziell Axialfehlern kommt (vgl. Gelfand 2004, 398, 402, 409; Spitzer 2002, 71). Dies ist auf schmalbandige Schalle wie Sinustöne (die Lieblingsstimuli der Psychoakustiker) oder quasi-Sinustöne (wie sie von Stimmgabeln erzeugt werden) zurückzuführen (vgl. Hellbrück & Ellermeier 2004, 44). Unter normalen Umständen fallen Hörern etwaige Ortungsschwierigkeiten kaum auf, weil (automatische) Peilbewegungen des Kopfs dabei helfen, Lokalisationsfehler zu vermeiden (vgl. Blauert 1974, 35, 143ff.; Gelfand 2004, 404; Guski 1996, 275; Wightman & Kistler 1993, 171).

Der Kopf vor allem ein Hindernis für freie Schallausbreitung dar. Er bewirkt Dämpfungen und Beugungen, die das Schallfeld stören. Besonders deswegen, und weil die Ohren an unterschiedlichen Seiten des Kopfs sitzen, werden sie so gut wie immer mit unterschiedlichen Schallsignalen versorgt (vgl. Blau-

ert 1974, 56; Gelfand 2004, 85; Guski 1996, 88; Yost 2008, 71). Quasi diotische Ohrsignale, d.h. Signale ohne verwertbare interaurale Differenzen, entstehen nur durch Schallquellen in der Medianebene. Dementsprechend werden solche (diotischen) Hörereignisse direkt vorne oder hinten (bzw. in der Medianebene liegend) lokalisiert (vgl. Blauert 1974, 35, 80; Gelfand 2004, 396).

3.1.2 Außenohr

Das Außenohr leitet nun das von Kopf und Körper gefilterte Schallgeschehen ans Mittelohr weiter, dem es gleichzeitig als Schutzvorrichtung gegen Fremdkörper und klimatische Einflüsse dient (vgl. Békésy & Rosenblith 1951, 1079; Blauert 1974, 50f.; Rosowski 2010, 49; Yost 2008, 73; Young 2007, 431). Doch nicht nur binaural, auch monaural – sprich mit nur einem Ohr – werden noch Richtungsinformationen gewonnen, da auch das Außenohr, speziell die Ohrmuschel als linearer Filter arbeitet, dessen Übertragungsfunktion von Richtung und Entfernung der Schallquelle abhängt (vgl. Blauert 1974, 50f.; 1997, 63; vgl. Gelfand 2004, 86; 2009, 49; Spitzer 2002, 56; Wightman & Kistler 1993, 165f.; Young 2007, 431).

Obwohl Schall von der „Schattenseite“ des Kopfs selbstverständlich auch für ein einzelnes Ohr anders als ungestört eintreffender Schall klingt, ist leicht nachvollziehbar, daß monaurales Hören nicht mit der horizontalen Raumauflösung binauralen Hörens schritthalten kann. Daher ist es wenig verwunderlich, daß horizontales Richtungshören vor allem über interaurale Unterschiede betrieben wird. Monoaurale Verzerrungsprodukte dienen fast ausschließlich vertikaler, bzw. sagittaler (vorne/hinten) Lokalisation sowie Distanzhören (vgl. Blauert 1997, 304; Culling & Akeroyd 2010, 127; Rosowski 2010, 63f.).

Monoaurales Hören

Monoaurale Lokalisation wird also über die Filterfunktion der Muschel betrieben (vgl. Blauert 1997, 307; Gelfand 2004, 400; Spitzer 2002, 66). Des weiteren ist die Muschel kritisch für extra- bzw. intrakraniale Lokalisation, also Hörorte außerhalb bzw. innerhalb des Kopfradius; ein Umstand, der leicht mit Kopfhörern überprüfbar ist, da Kopfhörer die Funktion der Muschel umgehen und i.d.R. Hörereignisse innerhalb des Kopfs evozieren (vgl. Blauert 1974, 86; 1997, 118; Gelfand 2004, 400; Wightman & Kistler 1993, 184; Yost 2008, 181, 183).

Am wichtigsten ist jedoch die Feststellung, daß die Ohrmuschel dort eingreift, wo binaurales Hören versagt, sprich, wenn binaurale Differenzen praktisch nicht vorhanden sind. Dies betrifft Einfallswinkel von vorne und hinten (0° bzw. 180° Azimut) sowie jegliche Erhebungswinkel, die man – der Lateralisation ähnlich – auf die Medianebene projizieren könnte (vgl. Blauert 1974, 35, 80; Gelfand 2004, 398; 2009, 49; Guski 1996, 89; Spitzer 2002, 69). Neben monoauraler und extrakranialer ist die Muschel somit auch an vertikaler und sagittaler Lokalisation, d.h. der Feststellung von Erhebungswinkeln und vorne/hinten-Unterscheidungen beteiligt (vgl. Blauert 1974, 82; 1997, 99f.; Culling & Akeroyd 2010, 127; Gelfand 2004, 86, 400; Guski 1996, 274; Spitzer 2002, 56; Wightman & Kistler 1993, 181, 185). Besonders letzteres läßt sich überprüfen, indem man verschiedene Muschelstrukturen z.B. mit Knet-

gummi ausfüllt, was die allgemeine, aber speziell die vorne/hinten-Lokalisation beeinträchtigt (vgl. Békésy & Rosenblith 1951, 1079; Gelfand 2004, 403; Spitzer 2002, 66). Verschlussgrad (Knetgummimasse) und Lokalisationsfähigkeit scheinen in indirektem Verhältnis zueinander zu stehen: je mehr Knetgummi, desto schlechter die Lokalisation. Konkrete Muschelbereiche, die mit speziellen Einbrüchen der Lokalisationsfähigkeit korrelieren, wurden dabei nicht gefunden (vgl. Spitzer 2002, 66).

Dies wirft die Frage auf, wie die Muschel tatsächlich funktioniert. Standardmodelle der Ohrmuschelfunktion besagen im allgemeinen, daß die Muschel den Schall trichterartig einfängt und auf direktem bzw. mittels Reflexionen indirektem Weg ans Mittelohr weiterleitet (vgl. Békésy & Rosenblith 1951, 1078; Müller & Frings 2009, 450; Reiss 2003, 591, 594; Rosowski 2010, 49; Spitzer 2002, 56; Young 2007, 431). Da Reflexionen längere Wege bedeuten, finden richtungsabhängig Verzögerungen statt, die den Einfallswinkel in zeitliche Information übertragen (vgl. Blauert 1974, 50f.; 1997, 63; Guski 1996, 89; Spitzer 2002, 56).

Tatsächlich ist die Trichterfunktion der Ohrmuschel (zumindest beim Menschen) vernachlässigbar (vgl. Gelfand 2004, 85f.; Graumann & Sasse 2005, 97). Blauert zweifelt auch die reflexionsbasierte Funktion der Muschel an, weil die Strukturen der Ohrmuschel dafür zu klein sind. Für gerichtete Rückwürfe müßten sie größer als die in Frage kommenden Wellenlängen sein, was schon für 4 kHz-Wellen mit rund 8,5 cm Länge (selbst 20 kHz-Wellen sind noch rund 1,7 cm lang) offensichtlich nicht mehr der Fall ist (vgl. Blauert 1974, 51; 1997, 64). Blauert (1974) zufolge finden keine gerichteten Reflexionen, sondern Beugungen und Streuungen (also Dispersion) statt, wobei die Höhlungen der Muschel „geführte Schallausbreitung“ (Blauert, 1974, 51) zulassen, wo – je nach Wellenlänge – auch Resonanzen auftreten können (vgl. Blauert 1974, 51; 1997, 64).

Interessanterweise wird in neuerer Literatur immer wieder von Reflexionen gesprochen (vgl. Guski 1996, 89; Müller & Frings 2009, 452; Reiss 2003, 594; Spitzer 2002, 56; Young 2007, 432; und auch Blauert selbst nennt Reflexionen: 1997, 63). Inwiefern es sich dabei um einfache Übernahmen des oben genannten Standardmodells, Simplifizierungen des Streuungsverhaltens oder tatsächlich neue Erkenntnisse handelt, war im Rahmen dieser Arbeit nicht nachzuvollziehen.

Einige mehr oder weniger isolierte Reflexionen wurden allerdings beschrieben. Vermutlich finden besonders Rückwürfe von Anthelix und Concha statt. Reflexionen der Concha-“Rückwand“ sind möglicherweise speziell für vertikale Lokalisation von Bedeutung (vgl. Culling & Akeroyd 2010, 127; Guski 1996, 274). Ein von Young (vgl. 2007, 432) erwähnter Einschnitt im Frequenzbereich zwischen 3 und 10 kHz entsteht ihm zufolge durch reflexionsbedingte Phasenverschiebungen, die Auslöschungen verursachen.

Übertragungsfunktion des Außenohrs

Wie die Richtungsabhängigkeit der Ohrmuschelübertragung genau zustande kommt, ist offenbar noch nicht detailliert nachgewiesen (vgl. Blauert 1974, 55). Eine Möglichkeit, diese Frage zu umgehen ist, nicht das Zustandekommen, sondern die Resultate der Übertragungsfunktion, d.h. den Einfluß der Ohrmuschel auf das Spektrum des Schalls am Trommelfell, zu betrachten (vgl. Blauert 1974, 52ff.; 1997, 65ff.). Dabei konnten einige Eigenresonanzen für die Filtercharakteristik des Außenohrs verant-

wortlich gemacht werden. Muschel und Gehörgang funktionieren als Einheit. Sie bilden ein Resonatorsystem, dessen Resonanzfrequenzen den Maxima ihrer Übertragungsfunktion entsprechen (vgl. Blauert 1974, 35ff., 53ff.).

Verfolgt man die spektrale Filterkurve der Außenohrübertragungen, fallen Maxima auf, die sich mit Eigenresonanzen des Außenohrs erklären lassen. Für die Form der Filterkurve hoher Frequenzen wird wegen ihrer kleinen Strukturen vor allem die Muschel verantwortlich gemacht (vgl. Gelfand 2004, 400, 403; Yost 2008, 72), während Concharesonanzen mittel- bis hochfrequente Minima zugeschrieben werden. Ähnlich dem obengenannten Reflexionsmodell sind sie möglicherweise Erkennungsmerkmale des Erhebungswinkels der Schallquelle (vgl. Rosowski 2010, 63).

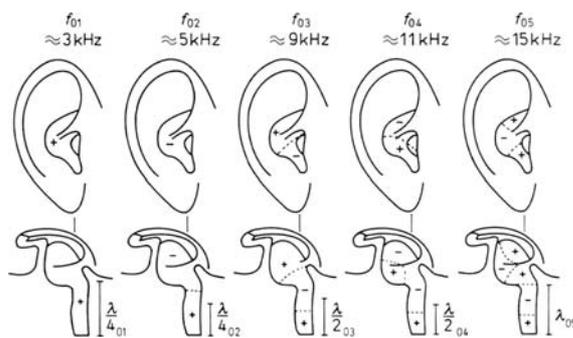


Abb. 3.4 **Außenohrresonanzen** (Blauert 1974, Abb. 40, 54). Schalldruckverteilung bei unterschiedlichen Eigenresonanzen eines künstlichen Außenohrs. Die gepunkteten Linien deuten Knotenflächen des Schalldrucks an (nach Shaw und Teranishi 1986). Die Linien verdeutlichen auch, daß mit steigender Frequenz immer kleinere Muschelteile involviert sind (siehe S. 98)

1993, 180). Zwischen 1,5 bzw. 2 und 7 kHz sind aufgrund von Resonanzen in Concha und Gehörgang Verstärkungen um 10 bis 15 dB möglich (vgl. Gelfand 2004, 87; Yost 2008, 73, 75; Young 2007, 430). Um 5 kHz liegt die Resonanz der gesamten Concha. Die Eingangsimpedanz des Gehörgangs schließt die Concha akustisch ab, so daß sie wie ein einseitig geschlossenes Gefäß resoniert. Die Wellenlänge (λ) der Resonanz beträgt (wie es bei einseitig geschlossenen Röhren der Fall ist) das vierfache der Conchatiefe, wobei hier die halbe Rohrtiefe Mündungseffekten zuzuschreiben ist (vgl. Blauert 1974, 50, 53f.; 1997, 67; Yost 2008, 73).

Die am besten untersuchte Resonanz ist mit Sicherheit die des Gehörgangs. Er ist mit einem einseitig geschlossenen (gedackten) Rohr vergleichbar, weshalb es sich ebenfalls um eine $\lambda/4$ -Resonanz handelt (vgl. Blauert 1974, 44, 53; 1997, 67; Gelfand 2004, 86; 2009, 46). Die Resonanzfrequenz des Gehörgangs liegt je nach Länge um die 3 kHz. Die Angaben schwanken zwischen 2 und 5,5 kHz, wobei das Gros der Werte auf ein Frequenzband zwischen rund 2,5 und 4 kHz konzentriert ist (vgl. Békésy & Rosenblith 1951, 1079; Blauert 1974, 53; 1997, 67; Fricke & Louven 2008, 418; Gelfand 2004, 86; 2009, 46f.; Guski 1996, 89; Hellbrück & Ellermeier 2004, 98; Rossing 2007, 4; Spitzer 2002, 56; Yost 2008, 73). Innerhalb dieses Resonanzbereichs finden Verstärkungen um 5 bis 10 dB statt, weshalb die

Eigenresonanzen mit korrespondierenden Erhebungen der Filterkurve wurden für die Bänder um 9, 11 und 13 kHz beobachtet. Sie haben zwar geringere Bandbreiten, verschmelzen aber zu einem insgesamt eher breiteren Plateau in der Übertragungskurve (vgl. Blauert 1974, 54; 1997, 68). Resonanzen in diesem Frequenzbereich sind offenbar sehr individuell, wodurch sie sich bei Messungen an natürlichen Ohren herausmitteln und so nur schwer isoliert nachzuweisen sind (vgl. Blauert 1974, 54; Young 2007, 430). Für ihre Bedeutung bezüglich der Erhebungswinkel mag auch sprechen, daß das obere und das stärkere hintere Blauert'sche Band mit 8 bzw. 14 kHz in zumindest ähnliche Bereiche fallen (vgl. Blauert 1974, 91; Abb. 66., S. 89; Wightman & Kistler

Hörschwelle hier am niedrigsten ist, das Gehör seinen empfindlichsten Bereich hat (vgl. Békésy & Rosenblith 1951, 1079; Rossing 2007, 4; Spitzer 2002, 56). Wenig überraschend fällt auch das stärker ausgeprägte vordere Blauert'sche Band in denselben Frequenzbereich (vgl. Blauert 1974, Abb. 66., S. 89). Rosowski (vgl. 2010, 70) erwähnt auch Resonanzen bezüglich der Gehörgangsquerschnitte, unter der Bedingung, daß die Wellenlängen dem Durchmesser entsprechen bzw. darunter liegen. Bedenkt man jedoch, daß die kleinste hörbare Wellenlänge grob 1,7 cm beträgt aber Gehörgangsquerschnitte im Millimeterbereich liegen, sind Querschnittresonanzen wohl als irrelevant zu bewerten. Unter etwa 3 kHz sinkt die Resonanzkurve des Außenohrs ab, da die Wellenlängen für Resonanzen bzw. gute Absorption des Ohres zu groß und die Impedanz des Mittelohres zu ungünstig wird (vgl. Young 2007, 431, 433).

Die Gehörgangsresonanz ist selbstverständlich richtungsunabhängig. Für Resonanzen der Muschel kann dies natürlich nicht gelten. Die 5 kHz-Resonanz ist im vorderen Schallfeld (0° bis 90° Azimut) konstant. Im hinteren Schallfeld findet jedoch bis zirka 110° eine Absenkung um 15 bis 20 dB statt. Der damit erreichte Minimalwert bleibt bis 180° konstant. Bei der 10 kHz-Resonanz ist bei Beschallung von vorne im, Gegensatz zu Beschallung von hinten, keinerlei Welligkeit festzustellen (vgl. Blauert 1974, 55; 1997, 68f.). Beide Resonanzen sind offensichtlich an vorne/hinten-Unterscheidungen beteiligt.

3.2 Impedanzwandlung

Der akustische Widerstand, die *Impedanz*, von Luft ist wesentlich geringer als der von Wasser. In Luft können schon mit wenig Kraftaufwand weite Molekülbewegungen herbeigeführt werden, während im Wasser viel Kraft für vergleichsweise geringe Bewegungen erforderlich ist. Das Ergebnis dieses Unterschieds ist eine Impedanzdiskrepanz (*impedance mismatch*) zwischen Luft und Wasser in der Größenordnung von knapp 60 dB (vgl. Gelfand 2004, 87f.; 2009, 50; Helmholtz 1896, 219; Spitzer 2002, 57; Yost 2008, 75f.; Young 2007, 432). Um Schall aber effizient zu übertragen, müßten Kraft und Bewegung an der Impedanzgrenze mehr oder weniger gleich sein. Andernfalls führen Reflexionen an der Übertrittsgrenze vom nachgiebigeren ins dichtere Medium (z.B. an der Wasseroberfläche, oder an Fensterscheiben; vgl. Lehnhardt & Laszig 2009, 63; Ranke 1953, 20; Spitzer 2002, 57) zu hohen Energieverlusten (vgl. Guski 1996, 90; Lehnhardt & Laszig 2009, 63; Müller & Frings 2009, 452; Reiss 2003, 604; Spitzer 2002, 57; Stevens & Hallowell 1938, 260; Young 2007, 432).

Nach dem evolutionären Schritt des Lebens aus dem Wasser an Land mußte also zwangsläufig die Entwicklung eines Apparates folgen, um den neuen Gegebenheiten Rechnung zu tragen. Mit dem Mittelohr hat die Natur ein geniales Übertragungssystem erschaffen, das die Luftschallimpedanz an die Verhältnisse des (immer noch) flüssigkeitsgefüllten Innenohrs anpaßt (vgl. Gelfand 2004, 88; 2009, 50; Guski 1996, 90; Helmholtz 1896, 219; Lehnhardt & Laszig 2009, 27; Schünke et al. 2009, 132; Spitzer 2002, 57; Stevens & Hallowell 1938, 259; Terhardt 1998, 56; Yost 2008, 75ff., 79; Young 2007, 429, 432; Zwicker & Feldtkeller 1967, 83; Zwislocki 1948, 15). Neben seiner Hauptaufgabe, der Impedanzwandlung, dient das Mittelohr dem Innenohr überdies als Schutzinstanz vor zu lauten Schallen, Schmutz und Krankheitserregern (vgl. Spitzer 2002, 58; Terhardt 1998, 56).

Exkurs: Zum physikalischen Konzept der Impedanz

Einen kurzen Blick auf das Konzept der *Impedanz* zu werfen hilft dabei zu verstehen, wie ein System (z.B. das Mittelohr) Schall transportiert, bzw. Widerstand leistet und warum es auf manche Frequenzen eher anspricht als auf andere (vgl. Gelfand 2004, 32; Yost 2008, 28). Die *Impedanz* ist das Gegenstück, sozusagen der Widerstand des freien Energieflusses durch ein System. Anders gesagt ist Impedanz (Z) die Antwort auf die Frage, wieviel Kraft (F) nötig ist, um eine bestimmte Geschwindigkeit (v) zu produzieren ($Z = F : v$): je größer der Kraftaufwand, umso größer die Impedanz (vgl. Gelfand 2004, 32; 2009, 23, 26). *Akustische Impedanz* entspricht übrigens dem heute kaum noch gebrauchten Begriff des *Schallwellenwiderstands* (vgl. Hellbrück & Ellermeier 2004, 66).

Impedanz ist eine komplexe Größe, die sich aus zwei Komponenten zusammensetzt: einer „realen“ Komponente, der *Resistanz*, und einer „imaginären“ Komponente, der (in der Elektrotechnik als *Blindwiderstand* bezeichneten) *Reaktanz* (vgl. Gelfand 2004, 32; 2009, 24, 205; Lehnhardt & Laszig 2009, 63; Yost 2008, 78). Ein Teil der einem System zugeführten Energie wird durch Reibung als Wärme abgegeben. Dieser Reibungsverlust wird *Resistanz* genannt. Die *Resistanz*, also die Reibung, die in einem System herrscht, tritt dementsprechend frequenzunabhängig immer auf (vgl. Gelfand 2004, 32; 2009, 23f., 25; Yost 2008, 28, 78). Die *Reaktanz* bezeichnet die Energieaufnahme (im Gegensatz zum Reibungsverlust) durch das System und wird durch Masse und Steifheit (bzw. ihren Kehrwert, der Nachgiebigkeit) des Systems gebildet (vgl. Gelfand 2004, 32; Lehnhardt & Laszig 2009, 63). Es gibt also zwei Arten von Reaktanzen: Die *positive* oder *Massereaktanz* und die *negative* oder *Feder- bzw. Steifheitsreaktanz* (vgl. Gelfand 2004, 32; 2009, 24, 205; Yost 2008, 28, 78). Beides sind frequenzabhängige Größen (vgl. Gelfand 2004, 33f.; 2009, 25; Yost 2008, 28, 78).

Als *Massereaktanz* wird der massebedingte Widerstand eines Systems, also sozusagen dessen Trägheit bezeichnet (vgl. Gelfand 2004, 33; 2009, 24). Die *Massereaktanz* ist direkt proportional zur Frequenz. Je höher die Frequenz, desto größer die *Massereaktanz* (vgl. Gelfand 2004, 33; 2009, 26). Oder anders gesagt: Masse bremst. Kleine Körper schwingen schnell, große Körper langsam, weshalb große *Massereaktanz* hohe Frequenzen stärker beeinträchtigt als tiefe (vgl. Lehnhardt & Laszig 2009, 28; Yost 2008, 78). Die Resonanzfrequenz nimmt mit zunehmender Masse ab. Die Rückstellkraft, die ein elastisches Element im Falle der Auslenkung aus der Ruhelage entwickelt, wird als *Federreaktanz* bezeichnet. Sie entspricht demnach der Steifheit (Schwierigkeit des Zusammendrückens) oder Nachgiebigkeit (Komplianz, Leichtigkeit des Auseinanderziehens) des Systems (vgl. Gelfand 2009, 24; Lehnhardt & Laszig 2009, 63; Yost 2008, 78). Die *Federreaktanz* ist indirekt proportional zur Frequenz. Je höher die Frequenz, desto kleiner die *Federreaktanz* (vgl. Gelfand 2004, 34; 2009, 26). Oder anders gesagt: Steifheit beschleunigt. Feste Körper klingen hell, weiche Körper klingen dumpf, weshalb tiefe Frequenzen durch große *Federreaktanz* stärker beeinträchtigt werden als hohe (vgl. Gelfand 2004, 34; Lehnhardt & Laszig 2009, 28; Yost 2008, 78). Die Resonanzfrequenz nimmt also mit wachsender Steifheit zu.

Weil *Massereaktanz* direkt und *Federreaktanz* indirekt proportional zur Frequenz sind, sind beide bei einer bestimmten Frequenz gleich stark und heben sich gegenseitig auf. Dies ist dann die Resonanzfrequenz des Systems. In diesem Fall wird die Impedanz nur noch durch die Reibung, d.h. die *Resistanz* gebildet (vgl. Gelfand 2004, 34f.; 2009, 26; Yost 2008, 78). Man kann Impedanz also vereinfachend als eine Art „Sammelwiderstand“ betrachten, der sich aus Reibung, Steifheit und Masse (Trägheit) zusammensetzt (vgl. Gelfand 2009, 23).

Die Schalltransmission ans Innenohr erfolgt grundsätzlich auf zwei verschiedenen Wegen: Erstens auf dem „normalen“ Weg über die Gehörknöchelchenkette. Luftschwingungen im Gehörgang werden vom Trommelfell aufgenommen und über die Gehörknöchelchen angepaßt auf die Perilymphe im Innenohr übertragen. Zweitens durch Knochenschall. Direkt oder indirekt herbeigeführte Schädel-schwingungen wirken sich zwangsläufig aufs Innenohr aus und lösen dort entsprechende Hörempfindungen aus (vgl. Békésy & Rosenblith 1951, 1108; Ranke 1953, 35; Rosowski 2010, 59; Spitzer 2002, 57; Stevens & Hallowell 1938, 252; Yost 2008, 75).

3.2.1 Luftschalleitung (normaler Weg)

Auf dem „normalen“ Weg über die Gehörknöchelchen wird die relativ niedrige Gehörgangsimpedanz an die vergleichsweise hohe Innenohr-impedanz angepaßt, indem die schwachen aber weiten Trommelfellschwingungen auf kleine aber kräftige Steigbügelbewegungen „untersetzt“¹⁵ werden (vgl. Békésy 1960, 197; Békésy & Rosenblith 1951, 1082; Helmholtz 1896, 219; Müller & Frings 2009, 450ff.; Rosowski 2010, 60; Spitzer 2002, 57f.). Die nötige Verstärkung wird durch Druckanpassung und Hebelwirkung erreicht. Da Druck Kraft pro Fläche ist, bewirkt die Konzentration der relativ großen Trommelfellfläche auf die relativ kleine Steigbügelfußplatte bereits eine Verstärkung. Durch den im Vergleich zum (den Steigbügel antreibenden) vorderen Amboßschenkel etwas längeren Hammergriff wird zusätzlich Kraft gewonnen (vgl. Békésy 1941(a), 9, 14; Gelfand 2004, 90, 94; 2009, 50; Hellbrück & Ellermeier 2004, 99; Ranke 1953, 44; Schünke et al. 2009, 132; Spitzer 2002, 57; Yost 2008, 76; Young 2007, 432).

Als weiterer Verstärker kommt die Krümmung des Trommelfells in Frage. Dieser zuerst von Helmholtz (vgl. Gelfand 2004, 90f.; 2009, 50; Helmholtz 1896, 219ff.; Rosowski 2010, 60f.) beschriebene Mechanismus wurde u.a. von Békésy nicht übernommen (vgl. 1941(a), 9; 1960, 96; Gelfand 2004, 91; Rosowski 2010, 61). In jüngerer Literatur wird er aber durchaus noch beschrieben oder erwähnt (vgl. Gelfand 2004, 89f.; Lehnhardt & Laszig 2009, 27; Rosowski 2010, 61; Spitzer 2002, 57; Yost 2008, 76) und neuere Untersuchungen der Trommelfellbewegung (vgl. Gelfand 2004, 92; Rosowski 2010, 61, 65f.) sind in Helmholtz' Sinne interpretierbar.

Mechanisch ist das Mittelohr kein einfacher Schwingungskörper wie z.B. eine Stimmgabel, sondern ein gedämpftes System, dessen physikalische Eigenschaften durch Federkraft (Rückstellkraft), Masse (Trägheit) und Reibung definiert sind (vgl. Békésy 1941(a), 15; 1960, 104; Lehnhardt & Laszig 2009, 27). Die Dämpfung reduziert bzw. verhindert Verzerrungen wie z.B. Transienten¹⁶ in Sprachlauten (vgl. Békésy & Rosenblith 1951, 1084; Lehnhardt & Laszig 2009, 27).

Die Rückstellkräfte werden hauptsächlich durch das Luftpolster der Mittelohrhöhle und die fixierenden Bandmassen der Gehörknöchelchen bestimmt. Steigbügel und Perilymphe des Innenohrs bilden Masse und Trägheit. Da Hammer und Amboß so um ihre Rotationsachse gelagert sind, daß diese mit ihrem Massenschwerpunkt zusammenfällt, leistet nicht ihre Masse, sondern nur ihre Rotationsträgheit einen nennenswerten Beitrag (vgl. Békésy 1960, 102; Békésy & Rosenblith 1951, 1084; Gelfand 2004, 94; Lehnhardt & Laszig 2009, 27f.).

¹⁵ Grundsätzlich werden unterschiedliche Werte innerhalb einer physikalischen Größe durch *Übersetzung* erreicht (ein typisches Beispiel wären Getriebedrehzahlen). Übersetzungsverhältnisse größer als Eins (z.B. Eingangsamplitude zu Ausgangsamplitude = 2:1) werden (umgangssprachlich) auch als *Untersetzung* bezeichnet.

¹⁶ Als *Transienten* bezeichnet man kurzfristige impulshafte Verbreiterungen des Spektrums aufgrund schneller Ausgleichsvorgänge.

Die Reibung entsteht vor allem in der Perilymphe, zu geringen Teilen aber auch in der Luft im Gehörgang und der Paukenhöhle, mitunter sogar durch die Bandmassen (vgl. Békésy 1960, 104; Lehnhardt & Laszig 2009, 28).

Mittelohrimpedanz

Wie gut das Trommelfell den Gehörgangsschalldruck aufnimmt, hängt von seiner Eingangsimpedanz ab. Diese wird im Gehörgang als Steifheit bzw. Nachgiebigkeit des Trommelfells gemessen (vgl. Lehnhardt & Laszig 2009, 63f.; Yost 2008, 78). Da Messungen der Trommelfellsteifheit im Prinzip das gesamte Mittelohr betreffen, sind die Begriffe *Trommelfellimpedanz* und *Mittelohrimpedanz* offenkundig synonym zu verstehen.

Interessanterweise scheint die Trommelfellmembran selbst, samt ihrer Elastizität, keinen entscheidenden Einfluß auf die Impedanz zu haben. Die Messwerte entsprechen ziemlich genau der gemessenen Impedanz eines Luftpolsters, dessen Volumen dem der Mittelohrhöhle entspricht. Man könnte also sagen: die Trommelfellimpedanz ist in erster Linie durch das Luftvolumen in der Paukenhöhle bestimmt (vgl. Békésy 1960, 110, 173). Dabei muß auch das Verhalten der Shrapnellmembran (Pars flaccida) berücksichtigt werden (vgl. Békésy 1960, 104).

Unterhalb von 500 Hz verläuft die Impedanzkurve sehr einheitlich, während sie darüber mehrere Extrema aufweist. Bis etwa 1 kHz hängt die Kurve vom Gesamtvolumen des Mittelohrraums ab, während sie darüber durch die individuelle Detailanatomie, also die Mittelohrhöhlenform, gezeichnet wird (vgl. Stepp & Voss 2005, 863f., 868f.). In diesem Sinne scheint es bemerkenswert, daß Voss et alii (vgl. 2000, 67) unter 2 kHz nur einen geringen Einfluß des Mittelohrluftpolsters feststellen.

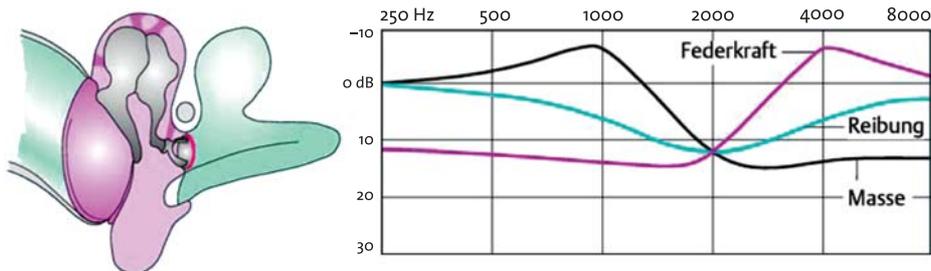


Abb. 3.5 **Mittelohrimpedanz und ihre Komponenten** (Lehnhardt & Laszig 2009, Abb. 4.3 & 4.4, 29). Reibung (türkis), Masse (grau) und Federkraft (rosa). Links: Die beitragenden Komponenten. Rechts: der Einfluß von Massereaktanz (Masse) und Federreaktanz (Federkraft) auf die Eigenfrequenz des Mittelohrs. Änderungen der Resistanz (Masse) beeinflussen die Eigenfrequenz kaum.

Da Steifheit hohe Frequenzen eher begünstigt als tiefe, sind tiefe Frequenzbereiche, speziell unter 1 kHz, durch Federreaktanzen dominiert (also Trommelfellsteifheit, Bandmassen etc.), während in höheren Frequenzbereichen die Massereaktanz (also vor allem die Eingangsimpedanz des Innenohrs) dominiert (vgl. Rosowski 2010, 61; Stepp & Voss 2005, 863f., 869; Voss et al. 2000, 67; Yost 2008, 78). In einem Mittelbereich um 1 bis 2 kHz neutralisieren sich die Reaktanzen und resistive Komponenten (praktisch

Reibung) bestimmen die Impedanz. Zugleich erreicht das Trommelfell ab etwa 1,5 kHz einen nahezu optimalen Absorptionsgrad, was eine minimale Impedanz bedeutet (vgl. Békésy & Rosenblith 1951, 1080; Yost 2008, 78).

Mit steigendem Druck im Gehörgang nimmt die Steifheit des Trommelfells zu, es reagiert dann zunehmend wie eine gespannte Membran, was zur Abnahme der Absorptionsfähigkeit und somit der Übertragung ans Mittelohr führt (vgl. Békésy 1960, 100, 141).

Druckausgleich

Das Mittelohr verfügt über Mechanismen, die konstanten, an den Umgebungsdruck angepaßten Binnenluftdruck gewährleisten (vgl. Di Martino et al. 2004; Didyk et al. 2007; Schünke et al. 2009, 131) und so druckbedingte Trommelfellspannungen und daraus resultierende Übertragungsverluste und sogar Schmerzen verhindern sollen (vgl. Helmholtz 1896, 212; Müller & Frings 2009, 452; Stepp & Voss 2005, 869). Größere Druckschwankungen werden über die *Eustachische Röhre* ausgeglichen, während die *Shrapnellmembran* kleinere und schnellere Fluktuationen abfängt.

Die *Shrapnellmembran* wird als Barorezeptor verstanden, der u.a. für die Öffnung der Eustachischen Röhre bedeutsam ist (vgl. Di Martino et al. 2004, 1030; Didyk et al. 2007, 27). Zudem werden ihr druck- und volumenstabilisierende Aufgaben zugeschrieben, wodurch sie das Trommelfell vor statischer Druckbelastung (und möglicherweise sogar Lautstärke) schützen soll (vgl. Didyk et al. 2007, 21ff., 27; Rosowski 2010, 52).

Sie ist fähig, auf kleinste Druckänderungen zu reagieren, indem sie sich in gleicher Frequenz wie die Druckschwankungen, je nachdem, ein- oder auswärts bewegt. Allerdings fehlt es ihr an Spannung, um schallempfindlich zu sein (vgl. Didyk et al. 2007, 21f.). Am Punkt maximaler Auslenkung verharrt die Shrapnellmembran kurz, bevor sie „den Rückweg antritt“. Daraus ergibt sich eine gewisse, druckabhängige Verzögerung (Hysterese). Mit wachsender Druckamplitude wächst auch die Verzögerung (vgl. Didyk et al. 2007, 23, 26).

Ebenso nimmt die Impedanz der Pars flaccida zu höheren Druckamplituden hin zu, wohingegen ihre eigene Bewegungsamplitude abnimmt. Diese größere Nachgiebigkeit bei geringen Amplituden spricht dafür, daß die Shrapnellmembran für kleine Druckschwankungen optimiert ist (vgl. Didyk et al. 2007, 22, 27). Volumenänderungen im Mittelohr durch Shrapnellmembranbewegungen betragen nur wenige Mikroliter bzw. weniger als 0,2–0,5% des Höhlenvolumens. Einwärts können bis zu 40 Pa ausgeglichen werden, auswärts das Doppelte. Dieser winzige Spielraum reicht für kleinere atmosphärische Druckfluktuationen, die normalerweise wenige -zig Pascal betragen (vgl. Didyk et al. 2007, 21, 27).

Um die Schalltransmission durchs Mittelohr effizient zu gestalten, muß gleicher Luftdruck beiderseits des Trommelfells gewährleistet sein. Dies zu bewerkstelligen ist die Hauptaufgabe der *Eustachischen Röhre* (vgl. Di Martino et al. 2004, 1030, 1038; Didyk et al. 2007, 27; Guski 1996, 90f.; Hellbrück & Ellermeier 2004, 92; Müller & Frings 2009, 452; Ranke 1953, 36f.; Reiss 2003, 611; Lehnhardt & Laszig 2009, 27; Stevens & Hallowell 1938, 249; Terhardt 1998, 56). Gäbe es die Tube nicht, würde das Trommelfell auf Schwankungen des äußeren Luftdrucks wie die Membran eines Dosenbarometers

reagieren (vgl. Müller & Frings 2009, 452). Der Mittelohrinnendruck wäre dabei keineswegs konstant. Unabhängig vom Außendruck würde er mit der Zeit fallen, da Gase im Körper über den Blutkreislauf absorbiert werden (vgl. Békésy & Rosenblith 1951, 1075). Die kleinen druckausgleichbedingten Korrekturen der Lage des Trommelfells machen sich oft als leises Knistern oder Knacken bemerkbar (vgl. Zwicker & Feldtkeller 1967, 86). Neben ihrer Funktion als Ausgleichsrohr dient die Tube dem Mittelohr auch als Abschirmung gegen selbstinduzierte Druckschwankungen etwa durch Sprechen oder Husten (vgl. Ranke 1953, 37).

Knorpelform und -elastizität, Schleimhautadhäsion und umliegender Gewebedruck halten die Eustachische Röhre – abgesehen von einem dünnen Sicherheitsrohr superior des Knorpelhakens – im Ruhezustand geschlossen (vgl. Di Martino et al. 2004, 1030f.; Gelfand 2004, 47; 2009, 45; Reiss 2003, 611; Schünke et al. 2009, 131). Die Tube wird durch aktive Muskelbewegung beim Schlucken, Gähnen und Trinken geöffnet (vgl. Di Martino et al. 2004, 1030; Lehnhardt & Laszig 2009, 27; Stevens & Hallowell 1938, 249; Zwicker & Feldtkeller 1967, 86). Passive Öffnungen sind dagegen unphysiologisch und treten meistens druckbedingt auf (vgl. Di Martino et al. 2004, 1030f.). Als möglicher Auslöser des Tubenreflexes werden *Chemorezeptoren* in der Paukenhöhlenschleimhaut, die Shrapnellmembran sowie *Druckrezeptoren* der Tuben-Rachenöffnung in Betracht gezogen (vgl. Di Martino et al. 2004, 1030; Didyk et al. 2007, 27), deren sensorische Fasern mit den motorischen Fasern des Öffnungsmechanismus verknüpft sind (vgl. Didyk et al. 2007, 27).

Geöffnet wird die Tube hauptsächlich durch die beidseitig neben ihr liegenden Gaumenmuskeln, den lateral gelegenen *Gaumenspanner* (*Musculus tensor veli palatini*) und den medial gelegenen *Gaumenheber* (*Musculus levator veli palatini*). Beim Schlucken kontrahieren beide. Dabei ziehen die Fasern des Gaumenspanners an der Lamina membranacea und öffnen dadurch den Knorpelteil der Röhre (vgl. Di Martino et al. 2004, 1030; Reiss 2003, 611; Schünke et al. 2009, 131; schematisch dargestellt bei Gelfand 2004, Abb. 2.10, 48; 2009, Abb. 2.13, 45).

Trommelfellbewegung

Das Trommelfell (genau genommen die Pars tensa) wird durch den Schalldruck im Gehörgang angeregt, wobei die aufs Trommelfell einwirkende Kraft aus der Differenz des Drucks vor und hinter der Membran kommt (vgl. Blauert 1974, 43; 1997, 54; Voss et al. 2000, 51). Weil der Druck in der Paukenhöhle – zumindest in den Zeitspannen hörbarer Frequenzen – nur durch das Trommelfell selbst an den äußeren Schalldruck gekoppelt ist, sind beide proportional zueinander (vgl. Blauert 1974, 43f.). Schallempfänger, die nur vom Druck auf einer Seite abhängen, werden als *Druckempfänger* bezeichnet (vgl. Blauert 1974, 44; 1997, 54). Daß das Trommelfell (zumindest unter 1,5 kHz) tatsächlich ein reiner Druckempfänger ist, untermauern Békésy und Rosenblith mit der Feststellung, daß das Trommelfell, in einem Druckknoten gelegen, keine Hörempfindung auslöst, während ein Druckbauch sehr wohl zu Hörempfindungen führt (vgl. Békésy 1941(a), 9; 1960, 95; Békésy & Rosenblith 1951, 1080).

Das Schwingungsverhalten des Trommelfells folgt komplexen Mustern und wird nach wie vor diskutiert. Es hat diesbezüglich bereits sehr unterschiedliche Ideen gegeben (vgl. Gelfand 2004, 90ff; Rosowski 2010, 60f., 64f.). Generelle Übereinstimmung findet man besonders im frequenzabhängigen

Verhalten: Einfachere Bewegungsmuster brechen mit steigender Frequenz auf und weichen zunehmend komplizierteren Formen (vgl. Békésy & Rosenblith 1951, 1080; Cheng et al. 2010, 66f., 75f.; Gelfand 2004, 92; Rosowski 2010, 65; Yost 2008, 74).

Bei tiefen Frequenzen unter 2 kHz finden sich nur wenige (ein bis drei) Maxima, wobei die größten Amplituden posterior des Hammergriffs erreicht werden (vgl. Cheng et al. 2010, 67, 71f., 75; Gelfand 2004, 92). Mit steigender Frequenz verkomplizieren sich die Muster durch weitere Maxima (vgl. Békésy & Rosenblith 1951, 1080; Cheng et al. 2010, 67, 71f., 75; Gelfand 2004, 92; Rosowski 2010, 61, 64f.; Yost 2008, 74). Ab 4 kHz zeichnen sich Wellenbewegungen ab, die über 8 kHz deutlich erkennbar sind (vgl. Cheng et al. 2010, 72; Rosowski 2010, 65). Bezüglich der Phasen läßt sich beobachten, daß sich bei tiefen Frequenzen unter 1 kHz mehr oder weniger das gesamte Trommelfell homogen bewegt, während die sich mit steigender Frequenz ausbildenden Maxima gleichsam wachsende Phasendifferenzen aufweisen (vgl. Cheng et al. 2010, 71f., 75; Rosowski 2010, 61, 64f.). Die Verteilung der Maximalamplituden konzentriert sich bei tiefen Frequenzen ausschließlich auf die posteriore und darin speziell auf den darin superioren Quadranten des Trommelfells. Mit steigender Frequenz verteilen sich die Maximalamplituden mehr und mehr über die gesamte Trommelfellfläche (vgl. Cheng et al. 2010, 71ff., 75f.; Rosowski 2010, 61, 64f.). Auf-

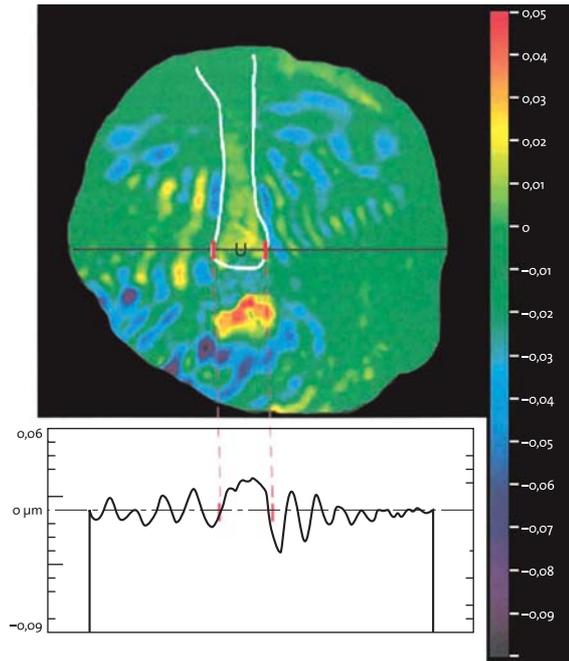


Abb. 3.6 **Trommelfellbewegung** (Rosowski 2010, Abb. 3.13, 66). Im Diagramm unten sind deutlich die Wellenbewegungen zu erkennen. H Umbo (Ende des Hammergriffs).

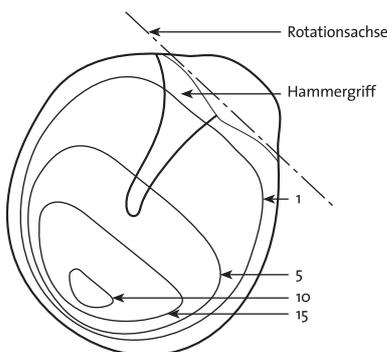


Abb. 3.7 **Trommelfellbewegung nach Békésy** (1941(a), Abb. 13, 13). Die Zahlen geben die Amplitudenverhältnisse der umrissenen Amplitudenbereiche wieder.

fallend ist, daß die Amplituden des Hammergriffs (in abnehmender Weise bei steigender Frequenz) zum Teil deutlich hinter den Maximalamplituden zurückbleiben (vgl. Cheng et al. 2010, 71; Gelfand 2004, 92).

Alles in allem werden Trommelfellschwingungen in jüngster Literatur als mit steigender Frequenz komplexer werdende Wellenbewegungen und/oder (aufgrund verschiedener knotenartiger Minima) stationäre Moden beschrieben (vgl. Cheng et al. 2010, 75f.; Rosowski 2010, 65). Inwiefern die Phasengleichheit bei tiefen Frequenzen die von Békésy beobachtete Kolbenbewegung tatsächlich widerspiegelt, welcher die komplexen Muster (nur) überlagert ist, ist nicht geklärt (vgl. Békésy 1941(a), 14; 1960, 100; Cheng et al. 2010, 76; Gelfand 2004, 91; Rochefoucauld & Olson 2010, 10f.).

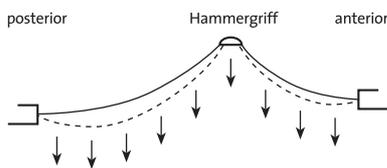


Abb. 3.8 Das Prinzip gekrümmter Membranen nach Helmholtz (Gelfand 2004, 3.5, 91).

Angesichts der im Sinne Helmholtz'scher Membrankrümmung interpretierbaren kleinen Amplituden im Nabelbereich ist zu überlegen, ob die Kraftübertragung auf die Gehörknöchelchen nicht tatsächlich in erster Linie durch Krümmungseffekte geschieht (vgl. Cheng et al. 2010, 71; Gelfand 2004, 92).

Helmholtz war der Meinung, daß eine leichte Krümmung der Trommelfellmembran (nicht zu verwechseln mit der grundsätzlich konischen Anlage des Trommelfells) nennenswerte Schallverstärkungen mit sich bringt. Das Prinzip basiert auf der Tatsache, daß gespannte Fäden einerseits mit einem Bruchteil der für die Spannung benötigten Kraft mittig aus der Spannungsachse ausgelenkt werden können und andererseits bereits minimale Auslenkung große Kraft auf die gespannten Enden ausübt (vgl. Békésy 1941(a), 9; 1960, 95f.; Gelfand 2004, 90f.; 2009, 50; Helmholtz 1896, 119ff.). Weite Membranbewegungen würden sich demzufolge in Hammergriffbewegungen mit geringerer Amplitude aber größerer Kraft übersetzen.

Da Trommelfellbewegungen durch das Luftkissen der Paukenhöhle bestimmt sind und die Reaktanz des Luftpolsters größer als die des Trommelfells ist, fallen mögliche Verzerrungen durch die Membran selbst nicht ins Gewicht (vgl. Békésy & Rosenblith 1951, 1088). Weil Luft (zumal in diesen Dimensionen) ein äußerst lineares Medium ist, verwundert dann auch nicht, daß auch das Trommelfell als quasi linearer Überträger fungiert (vgl. Békésy 1960, 110; Békésy & Rosenblith 1951, 1087).

Neben den beiden bereits erwähnten Übertragungswegen (siehe 3.2, S. 102) besteht noch eine dritte Möglichkeit, Schall ans Innenohr zu leiten, die hier der Vollständigkeit halber erwähnt sei. Trommelfellbewegungen produzieren nebenbei Volumenschwankungen in der Paukenhöhle, wodurch wiederum die Innenohrfenster bewegt werden. Wegen der großen Impedanz und weil ovales und rundes Fenster gleichzeitig angeregt werden, ist diese Art der Luftstimulation äußerst ineffizient und kann bei Ohren mit gesunden Gehörknöchelchen vernachlässigt werden (vgl. Rosowski 2010, 59; Yost 2008, 75, 79; Stevens & Hallowell 1938, 252; Zwislocki 1948, 15).

Mechanik der Gehörknöchelchen

Die Gehörknöchelchenkette überträgt die Trommelfellschwingungen auf die Perilymphe im Innenohr. Durch das Flächenverhältnis von Trommelfell zu Steigbügel Fußplatte und das Längenverhältnis von Hammergriff zum langen Amboßschenkel werden die übertragenen Schwingungen erheblich verstärkt (vgl. Békésy 1941(a), 9, 14; Békésy & Rosenblith 1951, 1082; Gelfand 2004, 90, 94; Helmholtz 1896, 220; Lehnhardt & Laszig 2009, 32; Reiss 2003, 604; Rosowski 2010, 60; Schünke et al. 2009, 132; Stevens & Hallowell 1938, 256, 260; Yost 2008, 76; Young 2007, 432).

Wie praktisch alle Mittelohrkomponenten, reagieren auch die Gehörknöchelchen frequenzabhängig. Das Standardmodell der Gehörknöchelchenbewegung beschreibt simple Achsrotationen der Hammer-Amboßeinheit sowie daraus resultierende Kolbenbewegungen des Steigbügels im ovalen Fenster. Auf das Schwingerverhalten bei tiefen Frequenzen paßt die Beschreibung relativ gut (vgl. Békésy 1960, 112;

Békésy & Rosenblith 1951, 1078, Abb. 14, 1086; Gelfand 2004, 92f.; Helmholtz 1896, 217; Reiss 2003, 604; Rosowski 2010, 65; Yost 2008, 77f.). Bei höheren Frequenzen bilden sich jedoch kompliziertere Muster aus. Dabei wird die Rotation des Hammers, überlagert von Translationsbewegungen, vom Amboß auf den Steigbügel übertragen, wo sie sich als im Kern kolbenartige, jedoch von Schaukelbewegungen überlagerte Schwingung manifestiert (vgl. Aibara et al. 2001, 101; Békésy & Rosenblith 1951, 1084; Gelfand 2004, 94; Rosowski 2010, 52, 61, 65ff.).

Die Gehörknöchelchen sind durch verschiedene Bänder locker in der Paukenhöhle eingehängt (vgl. Békésy 1941(a), 1; 1960, 102; Békésy & Rosenblith 1951, 1077). Die axiale Aufhängung von Hammer und Amboß sorgt dafür, daß beide um diese Achse rotieren (vgl. Gelfand 2004, 93f.; Reiss 2003, 604; Rosowski 2010, 65; Stevens & Hallowell 1938, 250). Ihre Masse ist so ausbalanciert, daß ihr Schwerpunkt mit dem fibrösen Drehlager zusammenfällt. Dadurch werden Masseträgheit und Lagerkräfte quasi umgangen und die effektive Trägheit auf das Trägheitsmoment reduziert (vgl. Békésy 1941(a), 14; 1960, 102; Gelfand 2004, 94; Lehnhardt & Laszig 2009, 27f.; Ranke 1953, 40; Reiss 2003, 604). So ein Arrangement minimiert Verzerrungen (z.B. durch translatorische Bewegungen) und optimiert die Übertragungsbedingungen. Bei Frequenzen unter 200 Hz sind die Massekräfte so gering, daß allein die Bandachse die Rotation zu bestimmen vermag. Für die bei hohen Frequenzen auftretenden Kräfte ist sie jedoch zu schwach. Die Knöchelchen werden allein dadurch in Position gehalten, daß ihr Masse-schwerpunkt auf der Bandachse liegt (vgl. Békésy 1941(a), 14; 1960, 103; Békésy & Rosenblith 1951, 1084); ein Umstand, den Békésy experimentell bestätigte indem er zeigte, daß Durchtrennungen der Bänder keinen Einfluß auf die Rotationsbewegungen haben (vgl. Békésy 1941(a), 14; 1960, 102).

Wie schon weiter oben erwähnt, folgt der Hammer dem Trommelfell nur teilweise (siehe S. 106) und zeigt frequenzabhängig verschiedene Bewegungsmuster. Bis ca. 2,5 kHz schwingt er mehr oder weniger eindimensional um seine Achse. Jenseits dieses Frequenzbereichs lagern sich der Rotation Translationen an, die ihm insgesamt dreidimensionale elliptische Bewegungen verpassen (vgl. Gelfand 2004, 94; Rosowski 2010, 66).

Aufgrund des flexiblen Gelenks sind Gleitbewegungen zwischen Hammer und Amboß möglich (vgl. Helmholtz 1896, 217; Reiss 2003, 604; Rosowski 2010, 67; Stevens & Hallowell 1938, 256), die dazu führen, daß der Amboß die Rotationsgeschwindigkeiten des Hammers über den gesamten Hörbereich nicht erreicht. Unter 1 kHz bewegt er sich ungefähr halb so weit wie der Hammer, oberhalb von 3 kHz nimmt das Verhältnis sogar noch weiter ab (vgl. Rosowski 2010, 67; Stevens & Hallowell 1938, 256).

Erwähnenswert ist außerdem, daß neben Gleit- (zwischen Hammer und Amboß), Translations- und sonstigen Schaukelbewegungen auch Verbiegungen der Gehörknöchelchen selbst beobachtet wurden. Solche Unsteifheiten in der Übertragungskette sind der Hebelwirkung und damit der Schalltransmission abträglich (vgl. Rosowski 2010, 61).

Die Hebelwirkung entsteht, weil der Hammergriff ein wenig länger als der lange Amboßschenkel ist. Weitere Amplituden werden so in größere Kraft übersetzt (vgl. Ranke 1953; 39). Die Hebelwirkung geht nach Helmholtz (vgl. 1896, 220) mit einem Faktor von 1,5 in die Mittelohrverstärkung ein, wird aber zumeist als Faktor von 1,3 angegeben (vgl. Békésy 1941(a), 9; 1960, 95, 102; Békésy & Rosenblith 1951, 1082; Rosowski 2010, 60; Gelfand 2004, 94; 2009, 51; Reiss 2003, 604; Schünke et al. 2009, 132; Yost 2008, 76).

Interessanterweise spiegeln sich die komplexen ungerichteten Bewegungsanteile der Hammer-Amboßbewegungen nur sehr bedingt in den Steigbügelbewegungen wider (vgl. Rosowski 2010, 66). Diese sind im niederfrequenten Bereich und bei moderaten Intensitäten grundsätzlich kolbenartig angelegt (vgl. Aibara et al. 2001, 101, 109; Békésy 1960, 112; Békésy & Rosenblith 1951, 1078, Abb. 14, 1086; Gelfand 2004, 92; Rosowski 2010, 66f.; Voss et al. 2000, 51; Yost 2008, 77f.), wobei die Amplitude maximal 0,1 mm beträgt (vgl. Helmholtz 1896, 218f.). Theoretisch entsteht die Kolbenbewegung rein translatorisch entlang der Steigbügellängsachse, praktisch weist sie wohl aber – anatomisch bedingt – zumindest anteilig (um das posteriore Ende der Fußplatte gelagerte) Rotationskomponenten auf (vgl. Békésy 1960, 112; Békésy & Rosenblith 1951, 1078, Abb. 14, 1086; Yost 2008, 77f.). Auch bei steigender Frequenz bleibt die Kolbenbewegung erhalten, wird allerdings

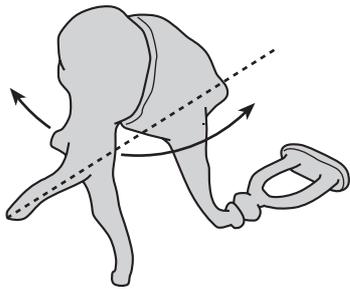


Abb. 3.9 **Bewegung der Gehörknöchelchen** (Schünke et al. 2009, 132 B.d). Die unterbrochene Linie markiert die Rotationsachse durch den vorderen Hammerfortsatz und den hinteren Amboßschenkel. Durch das Trommelfell wird der Hammer einwärts getrieben. Der Amboß überträgt dies auf den Steigbügel und drückt ihn annähernd kolbenartig ins ovale Fenster.

ab ungefähr 1 kHz zunehmend von Schaukelbewegungen überlagert (vgl. Aibara et al. 2001, 101, 108; Voss et al. 2000, 51). Bei hohen Intensitäten ändert sich das Bewegungsschema offenbar. Statt einer Kolbenbewegung beschreibt Békésy Schaukelbewegungen um die Längsachse der Fußplatte. Weil die Fußplatte sehr klein ist, kommt es zu Umflüssen der Innenohrlymphe um die Plattenbewegung, wodurch das Innenohr vor allzu großen Pegeln geschützt wird (vgl. Békésy 1960, 112f.; Békésy & Rosenblith 1951, 1085; Gelfand 2004, 92f.; Reiss 2003, 604; Stevens & Hallowell 1938, 255f.; Yost 2008, 78). Ob und, wenn ja, welchen Einfluß Schaukelbewegungen auf die Schallübertragung ans Innenohr haben, ist nach wie vor unklar (vgl. Rosowski 2010, 67).

Zerlegt man die Steigbügelbewegung in ihre räumlichen Komponenten, zeigt die Kolbenbewegung die sanfteste Frequenzkurve. Extrema der Frequenzkurven anderer Bewegungskomponenten sind im Innenohrschall nicht auffindbar, was den Schluß nahelegt, daß vor allem (wenngleich nicht ausschließlich) die Kolbenbewegung bedeutend für die Schalltransmission ist (vgl. Rosowski 2010, 79).

Der auf die große Trommelfellfläche wirkende Druck wird durch die Gehörknöchelchen auf die kleine Steigbügel Fußplatte konzentriert (vgl. Yost 2008, 76). Dieses Flächenverhältnis ist der Hauptfaktor der Mittelohrverstärkung. Wie die Angaben der Trommelfellfläche (siehe 2.3.1, S. 59) schwanken auch die Angaben des Verstärkungsfaktors, der sich als Quotient von Trommelfellfläche und Steigbügel Fußplatte berechnet. Die Werte liegen meist um 17, insgesamt zwischen 17 und 26,6 (vgl. Békésy 1941(a), 14; 1960, 102; Békésy & Rosenblith 1951, 1082; Gelfand 2004, 94; 2009, 51; Helmholtz 1896, 219; Lehnhardt & Laszig 2009, 27, 32; Müller & Frings 2009, 452; Reiss 2003, 604; Rosowski 2010, 60; Schünke et al. 2009, 132, Yost 2008, 76).

Grundsätzlich ist der schallinduzierten Knochenbewegung keine Frequenzgrenze gesetzt, dennoch nimmt die Steigbügelbewegung bei hohen Frequenzen ab. Im Bereich zwischen 0,9 und 2 kHz, etwa um 1 kHz, erreicht die Steigbügelgeschwindigkeit ihr Maximum und fällt zu höheren Frequenzen hin ab (vgl. Rosowski 2010, 78; Voss et al. 2000, 49, 63). Eine mögliche Erklärung bietet das Hammer-

Amboßgelenk, das für die Geschwindigkeiten hoher Frequenzen zu nachgiebig reagiert, wodurch die Übertragungsfähigkeit abnimmt. Die Abnahme der Steigbügelbewegung fällt offenbar mit der Obergrenze des Hörbereichs zusammen, so daß sich die Frage stellt, ob die Knöchelchenübertragung eine Rolle bei der Limitierung des Hörbereichs spielt (vgl. Rosowski 2010, 67, 78).

Übertragungsfunktion des Mittelohrs

Die Bezifferung der effektiven Verstärkung des Mittelohrmechanismus ist gewaltigen Schwankungen unterworfen. Grundsätzliche Einigkeit besteht darin, daß das Flächenverhältnis von Trommelfell zu Steigbügelfußplatte sowie der Hebelmechanismus der Gehörknöchelchen eine Verstärkung bewirken (vgl. Békésy 1941(a), 9, 14; Gelfand 2004, 90, 94; 2009, 50; Hellbrück & Ellermeier 2004, 98f.; Ranke 1953, 44; Schünke et al. 2009, 132; Stevens & Hallowell 1938, 260; Yost 2008, 76; Young 2007, 432).

Das Flächenverhältnis wird insgesamt in einer sehr weiten Spanne zwischen 17:1 und 26,6:1 angegeben, mehrheitlich jedoch als Verhältnis von 17:1. Der Mittelohrhebel wird normalerweise mit einem Verhältnis von 1,3:1 angegeben (vgl. Békésy 1941(a), 9, 14; 1960, 95, 102; Békésy & Rosenblith 1951, 1082; Gelfand 2004, 94; 2009, 51; Helmholtz 1896, 219; Lehnhardt & Laszig 2009, 27, 32; Müller & Frings 2009, 452; Ranke 1953, 44; Reiss 2003, 604; Rosowski 2010, 60; Schünke et al. 2009, 132, Yost 2008, 76). Rein rechnerisch läge die Verstärkung dann zwischen $17 \cdot 1,3$ und $26,6 \cdot 1,3$, also 22,1 bzw. 34,6. In der Literatur wird die Verstärkung (nachvollziehbarerweise passend zum oben favorisierten Flächenverhältnis) überwiegend mit Werten um 22 beziffert (vgl. Békésy 1941(a), 14; 1960, 102; Békésy & Rosenblith 1951, 1082; Gelfand 2009, 51; Lehnhardt & Laszig 2009, 32; Ranke 1953, 44; Reiss 2003, 604; Schünke et al. 2009, 132). Dies würde Verstärkungen um $(20 \cdot \log_{22,1})$ rund 26,8 dB bedeuten. Es ist allerdings anzunehmen, daß auch die Trommelfellkrümmung zur Gesamtverstärkung beiträgt (vgl. Gelfand 2004, 92; 2009, 50; Helmholtz 1896, 220f.; Yost 2008, 76). Der Faktor wird in der Literatur jedoch recht stiefmütterlich behandelt und mancherseits nicht berücksichtigt (vgl. Békésy 1941(a), 9; 1960, 96; Gelfand 2004, 91; 2009, 51). Wird der Krümmungseffekt berücksichtigt, geht er meist als Kraftverdoppelung in die Rechnung ein (vgl. Gelfand 2004, 94; Yost 2008, 76). Dementsprechend hätte das Mittelohr ein Verstärkungsverhältnis zwischen 44,2 und 69,2. Dies würde bedeuten, der Schalldruck am ovalen Fenster wäre 33 bis 36,9 dB größer als am Trommelfell (vgl. Gelfand 2004, 94; 2009, 51; Yost 2008, 76).

Der tatsächliche Schalldruck am ovalen Fenster ist frequenzabhängig (vgl. Gelfand 2004, 97; Hellbrück & Ellermeier 2004, 99; Yost 2008, 76). Im Frequenzbereich zwischen 1 und 2 kHz neutralisieren sich Feder- und Massereaktanz, so daß um 1,4 bis 1,5 kHz die Trommelfellimpedanz am geringsten und seine Aufnahmefähigkeit am größten ist (vgl. Békésy & Rosenblith 1951, 1080; Lehnhardt & Laszig 2009, 63; Rinsdorf in Keidel 1975, 85; Yost 2008, 78). Die Impedanzschneise stellt zugleich den Resonanzbereich des Mittelohrapparats dar (vgl. Aibara et al. 2001, 108; Békésy 1941(a), 15; 1960, 104; Békésy & Rosenblith 1951, 1084; Lehnhardt & Laszig 2009, 27; Gelfand 2004, 96). Gelfand (vgl. 2004, 96) nennt einen wesentlich weiteren reaktanzminimalen Aufnahmebereich bis etwa 5 bzw. 6 kHz und zudem bei 800 Hz sogar eine zweite, kleinere Resonanzspitze. Die Verstärkungskurve des Mittelohrs hat also Bandpaßcharakteristik mit einer Spitze (Resonanzbereich) um 1 bis 2 kHz. Darunter ver-

läuft die Filterkurve eher flach, wobei sie ab 100 Hz mit rund 6 dB pro Oktave ansteigt. Zu den hohen Frequenzen hin fällt die Filterkurve mit ähnlicher Flanke (grob 6 dB/Oktave) ab (vgl. Aibara et al. 2001, 108; Lehnhardt & Laszig 2009, 27). Diese impedanz- und resonanzbedingte Verstärkungskurve des Mittelohrs trägt maßgeblich zum analogen Verlauf der U-förmigen Hörschwelle bei (siehe 1.1.1, S. 6; vgl. Gelfand 2004, 96; Hellbrück & Ellermeier 2004, 99; Yost 2008, 77, 146). Darüber hinaus erreicht das Zusammenwirken von Mittelohrverstärkung und Außenohrresonanzen um 2 kHz Verstärkungen um etwa 60 dB, wodurch ein Großteil des Impedanzunterschieds von Luft zu Wasser ausgeglichen wird. Luftschalldruckwechsel führen hier also zu äquivalenten Innenohrschalldrücken, weshalb es zu keinen impedanzbedingten Empfindlichkeitsverlusten kommt (vgl. Yost 2008, 77).

Wenngleich auch die Mittelohrübertragungsfunktion nicht ideal ist, arbeitet das Mittelohr doch sehr linear (vgl. Gelfand 2004, 89; Rosowski 2010, 61). Bis rund 130 dB_{SPL} treten so gut wie keine Obertöne auf. Erst oberhalb der Schmerzschwelle treten plötzlich Obertöne auf, wobei ein 30%iger Pegelanstieg einen 8%igen Obertonzuwachs bedingt (vgl. Békésy 1941(a), 12; Cheng et al. 2010, 68; Voss et al. 2000, 49, 67).

Wäre der Gehörknöchelchenapparat tatsächlich ein starrer Körper, ginge die Übertragung vom Außenohr ans Innenohr ohne jeden Zeitverlust vonstatten. Tatsächlich sind jedoch zwischen Gehörgang und ovalem Fenster relativ lineare Verzögerungen von rund 20 µs festzustellen. Unklar ist aber, welche Komponenten im Detail an der Latenzentwicklung beteiligt sind. Als Kandidaten gelten neben der Wellenbewegung im Trommelfell vor allem die Nachgiebigkeit der Gelenke sowie der Gehörknöchelchen selbst (vgl. Rosowski 2010, 69f.).

Entgegen der Annahme, die ein einfacher Übertragungsmechanismus zulässt, kann man die Übertragungsfunktion des Mittelohrs nicht einfach umkehren, so daß die Übertragung vom Innenohr ans Außenohr (*otoakustische Emissionen*; siehe 3.3.5, S. 136) äquivalent zur Vorwärtsübertragung wäre (vgl. Rosowski 2010, 73). Obwohl das Mittelohr als inaktives, zeitinvariantes, lineares System ziemlich sicher reziprok ist, hängt die Übertragungsfunktion doch stark von den Impedanzen ab, die in beide Richtungen sehr unterschiedlich ausfallen, was letztlich die Erklärung für die Nicht-Reziprozität des Mittelohrs ist (vgl. Rosowski 2010, 73).

3.2.2 Stapediusreflex

Das Gehör ist in der Lage, die physikalisch determinierte Übertragungsfunktion des Mittelohrs physiologisch zu manipulieren. Hören ist also bereits ab dem Mittelohr ein aktiver Prozeß. Dies geschieht mittels eines efferenten Kontrollmechanismus, welcher über den Tonus der Mittelohrmuskeln die Übertragungsfunktion reguliert und so die Mittelohrimpedanz optimiert (vgl. Gelfand 2004, 103; Hellbrück & Ellermeier 2004, 99; Müller & Frings 2009, 452; Rosowski 2010, 52; Schünke et al. 2009, 133; Terhardt 1998, 55f; Yost 2008, 77, 79).

Beide Muskeln haben in entspanntem Zustand keinen Einfluß auf die Mittelohrübertragung (vgl. Dirckx & Decraemer 2001, 134; Voss et al. 2000, 44). Der Steigbügelmuskel zieht bei Kontraktion den Steigbügel nach posterior, wodurch dieser eine entsprechende Kippbewegung vollzieht, die Fußplatte –

posterior leicht versenkt – im ovalen Fenster verkeilt und eine sanfte Auswärtsbewegung des Trommelfells bewirkt (vgl. Gelfand 2004, 102; Békésy & Rosenblith 1951, 1085; Lehnhardt & Laszig 2009, 76; Reiss 2003, 605). Der Trommelfellspanner hat in entspanntem Zustand keinen Einfluß auf die Trommelfellposition. Bei Kontraktion zieht er das Trommelfell in die Paukenhöhle (vgl. Békésy & Rosenblith 1951, 1085; Hellbrück & Ellermeier 2004, 99; Gelfand 2004, 102).

Durch ihre Kontraktion versteifen die Muskeln die Gehörknöchelchenkette. Dies hat zur Folge, daß sich erstens die Beweglichkeit der Gehörknöchelchen und damit ihre Schalleitfähigkeit verschlechtert und zweitens die Eingangsimpedanz des Mittelohrs angehoben wird. Insgesamt nimmt die Ohrempfindlichkeit dadurch ab (vgl. Békésy & Rosenblith 1951, 1085; Blauert 1974, 43; Gelfand 2004, 103, 112; Müller & Frings 2009, 452; Schünke et al. 2009, 133; Yost 2008, 77, 79), so daß die Mittelohrmuskeln praktisch als Abschwächer arbeiten.

Die erhöhte Mittelohrsteifigkeit durch reflexartige Kontraktion der Mittelohrmuskeln wird *akustischer Reflex* oder *Stapediusreflex* genannt (vgl. Gelfand 2004, 102; Hellbrück & Ellermeier 2004, 99; Lehnhardt & Laszig 2009, 70; Yost 2008, 79). Er ist mit dem Pupillenreflex vergleichbar, der das Auge durch Verengen der Pupille vor zu starkem Lichteinfall schützt (vgl. Békésy 1960, 112; Spitzer 2002, 58; Stevens & Hallowell 1938, 265). Da sie durch Anspannung die Schwingfähigkeit des Systems reduzieren und so das Innenohr vor zu lauten Schallen schützen, arbeiten sie also praktisch als Limiter oder Abschwächer (vgl. Gelfand 2004, 115; Schünke et al. 2009, 132f.; Spitzer 2002, 58; Terhardt 1998, 56). Blauert beschreibt den Reflex so:

Der akustische Reflex stellt eine (trägheitsbehaftete) Kontraktion der Mittelohrmuskeln dar, die zu einer Versteifung der Schalleitungskette des Mittelohres und damit zu einer Verminderung der Ohrempfindlichkeit führt. Gleichzeitig ist damit eine Erhöhung der Trommelfellimpedanz um 100% verbunden (Blauert 1974, 49f.).

Ob allerdings tatsächlich beide Mittelohrmuskeln am akustischen Reflex beteiligt sind, ist unklar (vgl. Gelfand 2004, 102). Rein mechanisch würden Trommelfellspannerkontraktionen das Trommelfell einwärts ziehen. Otoskopische Beobachtungen zeigen meistens nur Auswärtsversetzungen, was der Steigbügelmuskelkontraktion entspricht (vgl. Gelfand 2004, 102). Die Innervierung des Trommelfellspanners, die nicht (oder nicht offensichtlich) mit auditiven Nervenbahnen verschaltet ist (vgl. Gelfand 2009, 43; Reiss 2003, 605; Russolo & Semeraro 1977, 377; Schünke et al. 2009, 104f., 107, 133, 215 C), spricht eher gegen eine prominente Rolle des Trommelfellspanners bei akustischen Reflexen. Läsionsstudien zufolge legen Deaktivierungen (Betäubungen, Verletzungen, Lähmungen, etc.) des Steigbügelmuskels den akustischen Reflex zwingend lahm, was bei Läsionen des Trommelfellspanners nicht der Fall ist (vgl. Gelfand 2004, 103; Russolo & Semeraro 1977, 374). Schallinduzierte Reaktionen des Trommelfellspanners sind beim Menschen äußerst selten (vgl. Russolo & Semeraro 1977, 374). Immerhin sind Reaktionen auf intensive Schalle bekannt (vgl. Gelfand 2004, 103; Schünke et al. 2009, 133). Ansonsten reagiert er vor allem auf nicht-akustische Reize (wie z.B. Augenlidhebungen am oder Luft-

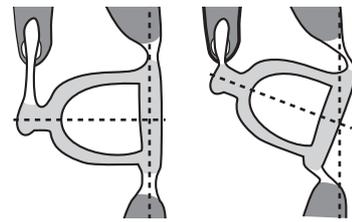


Abb. 3.10 **Zug des Stapediusmuskels** (Schünke et al. 2009, 132 B.a,b). Durch den nach posterior gerichteten Zug am Kopf des Steigbügels verkeilt sich dessen Fußplatte im ovalen Fenster.

strahlen ins Auge, elektrische Stimulation der Zunge). Derartige Reaktionen sind aber wohl eher Teil fazialer Schreckreaktionen (vgl. Gelfand 2004, 103; Russolo & Semeraro 1977, 374).

Békésy stellte fest, daß selbst der leichenstarre Trommelfellspanner das Trommelfell nur zu rund einem Drittel dämpft, woraus er schließt, daß der Muskel nur geringen Einfluß auf den akustischen Reflex hätte (vgl. Békésy 1960, 112). Man kann also davon ausgehen, daß es sich (zumindest beim Menschen) um einen reinen *Stapediusreflex* handelt (vgl. Gelfand 2004, 103; Russolo, Semeraro 1977, 374; Schünke et al. 2009, 133).

Der Stapediusreflex ist ein neuronal gesteuerter, motorischer Reflex, dessen afferenter (sensorischer) Schenkel über das Innenohr und den Hörnerv in den oberen Olivenkern verläuft, wo über den Fazialiskern (Nucleus nervi facialis) auf efferenten (motorischen) Schenkel umgeschaltet wird. Dieser verläuft über den Nervus facialis und den Stapediusnerv (vgl. Lehnhardt & Laszig 2009, 70; Schünke et al. 2009, 439; Yost 2008, 78). Der genaue Verlauf im Hirnstamm ist nicht bekannt. Da der Reflex aber möglichst schnell greifen muß, ist nicht davon auszugehen, daß mehrere Umschaltstationen beteiligt sind (vgl. Lehnhardt & Laszig 2009, 70).

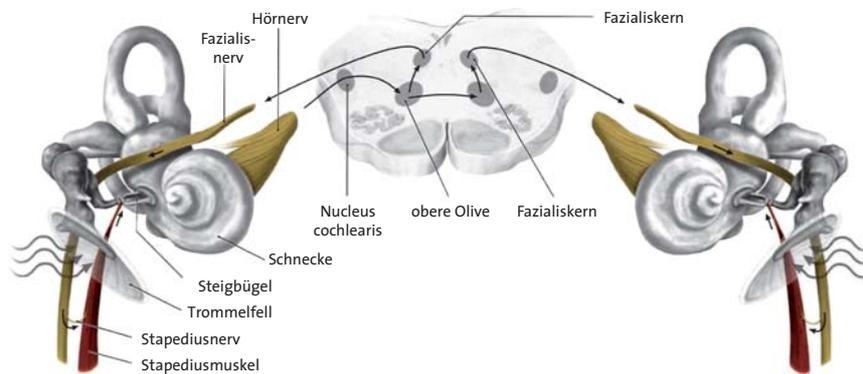


Abb. 3.11 **Verlauf des Stapediusreflexes** (modifiziert n. Schünke et al. 2009, 439 B). Der afferente Schenkel verläuft über die sensorischen Leitungen der Hörbahn. Im Hirnstamm wird beidseitig vom Olivenkernkomplex auf den Fazialiskern umgeschaltet. Der efferente Schenkel verläuft über die motorischen Bahnen des Fazialisnervs.

Aufgaben des Stapediusreflexes

Über die Aufgaben des Stapediusreflexes gibt es verschiedene Theorien (vgl. Gelfand 2004, 114). In jedem Fall spricht er auf intensive Schalle an. Seine Reaktionsschwelle ist relativ kompliziert signal- und frequenzabhängig. Vereinfacht gilt: je tieffrequenter der Stimulus, umso verlässlicher; je breitbandiger, umso niederschwelliger wird der Reflex ausgelöst (vgl. Lehnhardt & Laszig 2009, 70; Yost 2008, 77). Die Reflexschwelle wird allgemein mit Pegeln zwischen 60 und 100 dB_{SPL} bzw. mit rund 80 dB_{SL} angegeben (vgl. Blauert 1974, 43, 50; Guski 1996, 91; Hellbrück & Ellermeier 2004, 99; Rosowski 2010, 52; Spitzer 2002, 58; Yost 2008, 77), wobei er seine volle Wirksamkeit oberhalb von 100 dB erreicht (vgl. Blauert 1974, 50).

Für Töne zwischen 0,25 und 4 kHz schwanken die Schwellwertangaben zwischen 70 und 100 dB_{SPL}, für breitbandige Signale liegt die Schwelle etwa 20 dB tiefer (vgl. Gelfand 2004, 103; Lehnhardt & Laszig 2009, 70), allerdings sinkt die Schwelle nicht kontinuierlich, sondern erst ab einer gewissen Bandbreite ab. Gleiches gilt für wachsende Tonintervalle. Dies legt die Vermutung nahe, daß die Stapediusreflexschwelle mit einer Art kritischer Bandbreite (nicht zu verwechseln mit der psychoakustischen Frequenzgruppenbreite) verknüpft ist (vgl. Gelfand 2004, 104).

Als *Prävokalisationsreflex* wird das Auslösen der Steigbügelmuskelkontraktion unmittelbar vor selbsterzeugtem Schall wie vor allem Sprechen, Schreien und weiteren Vokalisationen bezeichnet. Es werden allerdings auch andere Eigengeräusche wie Kauen als Auslöser genannt (vgl. Guski 1996, 91; Hellbrück & Ellermeier 2004, 100; Rosowski 2010, 52; Spitzer 2002, 58).

Auch bei monotonischer Beschallung setzt der Stapediusreflex beidseitig ein, wobei die ipsilaterale Schwelle durchschnittlich 5 dB unter der kontralateralen liegt (vgl. Békésy & Rosenblith 1951, 1085; Gelfand 2004, 103; Lehnhardt & Laszig 2009, 72).

Funktion des Stapediusreflexes

Die Dämpfung durch den Stapediusreflex ist pegelabhängig, indem sie mit steigendem Pegel zunimmt (vgl. Gelfand 2004, 105f.; Lehnhardt & Laszig 2009, 71; Stevens & Hallowell 1938, 265; Yost 2008, 77). Dies geschieht für Töne bis etwa 120 dB_{SPL} und für Breitbandstimuli bis etwa 110 dB_{SPL} linear mit rund 0,6 bis 0,7 dB pro überschwelligem Dezibel (vgl. Gelfand 2004, 106; Yost 2008, 77). Erst bei noch höheren Pegeln ist der Reflex gesättigt, da die Muskeln ihre maximale Kontraktion erreicht haben (vgl. Gelfand 2004, 106; Lehnhardt & Laszig 2009, 71). Grob verallgemeinert reduziert der Stapediusreflex laute Schalle um 10 bis 30 dB, wobei die Wirkung am kontralateralen Ohr bei monotonischer Auslösung nur rund 5 bis 10 dB beträgt (vgl. Békésy & Rosenblith 1951, 1085; Guski 1996, 91; Hellbrück & Ellermeier 2004, 99; Spitzer 2002, 58; Yost 2008, 77).

Mit Dauer der Beschallung nimmt die Stärke des Reflexes ab. Diese Ermüdung ist als Adaption bekannt. Das Adaptionsverhalten ist frequenzabhängig. Mit steigender Frequenz verstärkt sich der Adaptionseffekt nicht nur, er tritt auch früher ein (vgl. Gelfand 2004, 108f.; Guski 1996, 91; Lehnhardt & Laszig 2009, 71f.). Nach Hellbrück und Ellermeier (vgl. 2004, 99) tritt jedoch kaum ein Adaptionseffekt auf:

Die Muskelfasern seien viel mehr sehr reich an Mitochondrien, sarkoplasmatischem Retikulum und Transversaltubuli. Sie verfügen somit reichhaltig über Energieresourcen und das für die Muskelkontraktion notwendige Kalziumionen-Konzentrat (Hellbrück & Ellermeier 2004, 99).

Wegen des Umwegs über die neuronalen Kanäle des Reflexes, der mechanischen Mittelohrträgheit und der Reaktionszeit des Steigbügelmuskels vergeht zwischen Stimuluseinsatz und effektiver Mittelohrversteifung eine gewisse Zeit (vgl. Gelfand 2004, 104; Spitzer 2002, 58; Terhardt 1998, 56). Die tatsächliche Latenzdauer ist pegel- und frequenzabhängig. Mit steigendem Pegel nimmt die Latenz ab. Bei einer Stimulusfrequenz von 1 kHz sinkt die Latenzzeit von 150 ms bei 80 dB_{SL} auf nur noch 40 ms bei 100 dB_{SL} (vgl. Gelfand 2004, 104f.; Lehnhardt & Laszig 2009, 71).

Konkrete Angaben zur Latenz des Stapediusreflexes sind aufgrund ihrer Signalabhängigkeit nur schwer zu erreichen und gehen – wenig verwunderlich – weit auseinander. Meistens werden Werte um 10 ms angegeben (vgl. Békésy & Rosenblith 1951, 1085; Spitzer 2002, 58; Yost 2008, 77), obwohl gut und gerne 150 bis 200 ms bis zur vollen Muskelkontraktion vergehen können (vgl. Guski 1996, 91; Hellbrück & Ellermeier 2004, 99; Stevens & Hallowell 1938, 265; Yost 2008, 77). Möchte man einen Durchschnittswert bilden, könnte man eventuell von einer Größenordnung um 100 ms ausgehen (vgl. Terhardt 1998, 56).

Werte unter 10 ms dürften in jedem Fall ausgeschlossen sein und selbst 10 ms scheinen optimistisch, wenn man elektromyographische Antwortzeiten des Steigbügelmuskels von 12 ms zugrunde legt und bedenkt, daß die elektrische Muskelantwort logischerweise *vor* der effektiven Impedanzänderung liegen muß (vgl. Gelfand 2004, 104f.).

Wegen seiner Latenz kann der Stapediusreflex vor Schallen wie Klicks, Knacks und Knallen etc., nicht effizient schützen, weil deren Einschwingvorgänge unterhalb der minimalen (genaugenommen der signalabhängigen) Reflexlatenz liegen (vgl. Békésy & Rosenblith 1951, 1085; Gelfand 2004, 114; Hellbrück & Ellermeier 2004, 99; Spitzer 2002, 58; Terhardt 1998, 56). Trotzdem bietet der Stapediusreflex dem Innenohr einen gewissen Schutz vor hochintensiven Schallen, der wohl aber eher als Anpassung an vorübergehend hohe Pegel zu verstehen ist (vgl. Békésy 1960, 112; Békésy & Rosenblith 1951, 1085; Gelfand 2004, 114; Guski 1996, 91; Spitzer 2002, 58; Terhardt 1998, 55f.; Yost 2008, 77). Das führt zu der Frage, ob Schallschutz überhaupt als eigentliche Aufgabe des Stapediusreflexes anzusehen ist, oder ob andere, übertragungsoptimierende Funktionen dessen Hauptaufgaben sind (vgl. Gelfand 2004, 114; Hellbrück & Ellermeier 2004, 99).

Eine solche Funktion, die den Mittelohrmuskeln zugeschrieben wird, ist das Unterdrücken von Verzerrungsprodukten bei zu großen Geschwindigkeiten der Knöchelchen, d.h. zu hohen Signalfrequenzen (vgl. Gelfand 2004, 114; Schünke et al. 2009, 132; Yost 2008, 77). Reiss zufolge ist der Trommelfellspanner an derartigen regulierenden Maßnahmen sehr wohl (aus akustischer Sicht sogar hauptsächlich) beteiligt (vgl. Reiss 2003, 605).

Ist der Kontaktdruck zwischen zwei Schwingungsüberträgern geringer als der Anregungsdruck, lösen sich die Überträger voneinander. Dadurch ist keine saubere Übertragung mehr möglich und Verzerrungen treten auf (vgl. Békésy 1960, 113; Yost 2008, 77). Bei kleinen Amplituden und Geschwindigkeiten können die Bandmassen im Mittelohr verzerrungsfreie Übertragungen gewährleisten. Bei zu heftigen Schwingungen ist dies nicht mehr der Fall. Dann greifen die Mittelohrmuskeln ein. Sie ziehen die Gehörknöchelchenkette mehr oder weniger antagonistisch zusammen, um sie „fester zu schnüren“ und verzerrungsfrei zu halten. Je höher Frequenz und Pegel sind, desto stärker agieren die Muskeln (vgl. Békésy 1960, 114; Yost 2008, 77).

Möglicherweise muß man ihnen überdies die Funktion eines „Noise gates“ bzw. Hochpaßfilters zuschreiben. Da sie durch Anspannung die Steifheit des Mittelohrs erhöhen, dämpfen sie so vor allem tieffrequente Schalle ab (vgl. Gelfand 2004, 114; Hellbrück & Ellermeier 2004, 99; Schünke et al. 2009, 133). Dies würde Maskierungseffekte reduzieren und die Dynamikspanne des Gehörs optimieren. Da auch die meisten (physiologischen) Körpergeräusche tieffrequent sind, ist eine signifikante Absenkung des internen Pegels und eine damit verbundene Erweiterung des Signal-Rauschabstands sehr wahrscheinlich (vgl. Gelfand 2004, 114f.; Hellbrück & Ellermeier 2004, 99f.; Müller & Frings 2009, 452).

3.2.3 Knochenschalleitung (alternativer Weg)

Neben dem klassischen Weg über das Mittelohr kann Schall auch, wesentlich leiser, über den Schädelknochen direkt ans Innenohr geleitet werden (vgl. Békésy 1960, 128; Békésy & Rosenblith 1951, 1108ff.; Blauert 1974, 43, 154; Rinsdorf in Keidel 1975, 85; Rosowski 2010, 59; Spitzer 2002, 57; Stevens & Hallowell 1938, 291; Yost 2008, 75, 79).

Normalerweise ist Knochenschall von sekundärer Bedeutung, weil er wesentlich leiser als Luftschall ist. Da der Impedanzunterschied von Luft zu Knochen größer ist als jener von Luft zu Wasser (Innenohr), liegt die Knochenleitungshörschwelle deutlich über der „normalen“ Hörschwelle (vgl. Blauert 1974, 43f.; Gelfand 2004, 98; Müller & Frings 2009, 452; Stevens & Hallowell 1938, 292). D.h., Schall muß stark genug sein, effektive Schädelbewegungen auszulösen. Die Angaben bei Knochenanregung direkt aus der Luft bewegen sich zwischen 40 und 60 dB oberhalb der Luftleitungshörschwelle (vgl. Békésy 1960, 178; Blauert 1974, 43, 154; Gelfand 2004, 98; Stevens & Hallowell 1938, 292). Wird der Schädel direkt angeregt, gewinnt der Knochenschall an Bedeutung, da nur noch die ausreichend geringe Impedanz zwischen Knochen und Wasser überbrückt werden muß (vgl. Yost 2008, 79).

Anregungen des Schädels lösen wegen seiner Dichte und Elastizität immer stehende Wellen im Knochen aus. Dabei folgt der Unterkiefer den Schädelbewegungen mit einer gewissen Verzögerung. Da der Hals nachgiebiger als der Kopf einzuschätzen ist, kann man, zumindest bei tiefen Frequenzen, von reinen Translationsbewegungen ausgehen (vgl. Békésy 1960, 129, 143, 164; Gelfand 2004, 98; Lehnhardt & Laszig 2009, 34). Bei größeren Geschwindigkeiten aufgrund höherer Frequenzen kommt es zu trägheitsbedingten Schädeldeformationen. Infolgedessen treten mit weiter steigender Frequenz Knotenlinien auf. Die Grundordnung der Knotenlinien richtet sich am Anregungspunkt des Schädels aus. Mit steigender Frequenz „wandern“ die Knotenlinien über die Schädeloberfläche und bilden frequenzabhängige Schwingungsmuster (vgl. Békésy 1960, 130ff., Abb. 6–4., 131, 164, 167f.; Békésy & Rosenblith 1951, 1109; Lehnhardt & Laszig 2009, 34).

Die frequenzabhängigen Muster lassen sich grob in etwa wie folgt beschreiben. Unter 200 Hz schwingen alle Schädelteile parallel in die gleiche Richtung. Der Schädel schwingt also als Einheit (vgl. Békésy 1960, 164; Békésy & Rosenblith 1951, 1109; Gelfand 2004, 98; Lehnhardt & Laszig 2009, 34). Über 200 Hz bildet sich ein zweipoliges Schwingungsmuster heraus (vgl. Békésy 1960, Abb. 6–2, 130), das ab 800 Hz Stirn und Hinterkopf gegenphasig schwingen läßt (vgl. Békésy 1960, Abb. 6–2 & 6–3, 130f.; Békésy & Rosenblith 1951, 1109; Gelfand 2004, 98). Zwischen 1 und 1,5 kHz liegt ein Übergangsbereich, in dem sich die Phasen von Stirn und Hinterkopf sowie die Knotenlinien zueinander verschieben (vgl. Békésy 1960, 167). Um 1,6 kHz bzw. 1,8 kHz ist die Schwingung nun vierpolig, indem neben Stirn und Hinterkopf auch die Schläfen gegenphasig zueinander schwingen (vgl. Békésy 1960, Abb. 6–2 & 6–3, 130f.; Gelfand 2004, 98). Die Amplituden der Schwingungspole sind gleich groß, was nach Békésy (vgl. 1960, 167f.) die erste Resonanz des Schädels markiert (Békésy & Rosenblith sprechen allerdings schon beim 800 Hz-Muster von Resonanz; vgl. 1951, 1109). Zu bedenken gilt, daß die Schwingungen des Schädels durch die Haut beeinträchtigt werden (vgl. dazu Békésy 1960, 156ff.).

Knochenschall kann auf zwei unterschiedliche Arten das Innenohr anregen, an denen prinzipiell alle drei Ohrabschnitte beteiligt sind (vgl. Gelfand 2004, 99). Beim *direkten Knochenschall* vibriert das Schlä-

fenbein und überträgt diese Schwingung ans Innenohr, indem die Labyrinthkapsel in konzentrische, rhythmische Kompressions- und Dilatationsschwankungen (Expansionsschwankungen) und so die Perilymphe in Bewegung versetzt wird (vgl. Békésy 1960, 177; Gelfand 2004, 99, Abb. 3.16, 100; Lehnhardt & Laszig 2009, 33; Stevens & Hallowell 1938, 293; Yost 2008, 79).

Weil die Perilymphe praktisch inkompressibel ist, weicht sie den Labyrinthkompressionen zu den flexibelsten Teilen ihrer Wandung, den Fenstern hin aus (vgl. Gelfand 2004, 99; Rinsdorf in Keidel 1975, 86f.). Da der Großteil des Innenohrs an die Vorhoftrappe gekoppelt ist, verfügt diese über eine deutlich größere Oberfläche als die Paukentreppe, was zu deutlich größeren Kompressionen bei gleichzeitig mehr verdrängter Masse (Lymphe) führt. Schon allein dadurch entsteht ein Lymphfluß Richtung Paukentreppe. Verstärkt wird der Fluß durch die höhere Elastizität des runden Fensters, da das ovale Fenster durch die Mittelohrmasse belastet ist (vgl. Békésy 1960, 145; Békésy & Rosenblith 1951, 1109f.; Gelfand 2004, 99; Lehnhardt & Laszig 2009, 33f.; Rinsdorf in Keidel 1975, 87; Stevens & Hallowell 1938, 293). Faszinierenderweise entstehen so aus konzentrischen Labyrinthkompressionen die gleichen vom ovalen Fenster zur Schneckenspitze laufenden Wanderwellen wie bei Schwingungen der Steigbügelfußplatte. Es ist sogar möglich, Knochenschall durch gegenphasigen Luftschall komplett auszulöschen (vgl. Békésy 1960, 127f.; Békésy & Rosenblith 1951, 1108; Gelfand 2004, 98; Lehnhardt & Laszig 2009, 33; Rinsdorf in Keidel 1975, 86; Stenfeld et al. 2003, 131; Stevens & Hallowell 1938, 293).

Der zweite Weg wird als *osteotympanaler Knochenschall* bezeichnet. Schädelschwingungen übertragen sich nicht nur ans Innenohr, sondern regen auch Lufträume zu Schwingungen an. Dies versetzt vor allem die Luftsäule des Gehörgangs in Schwingung, wodurch zwangsläufig die Mittelohrmechanik angetrieben wird. Dadurch wird die durch direkten Knochenschall angeregte Lymphbewegung unterstützt, weil der Steigbügel funktionell (zur Lymphe phasengedreht) fixiert wird (vgl. Békésy 1960, 177; Gelfand 2004, 101; Lehnhardt & Laszig 2009, 33f.; Rinsdorf in Keidel 1975, 88; Stevens & Hallowell 1938, 293). Die Fixierung ist abhängig vom Phasenversatz zwischen direktem und osteotympanalem Knochenschall, der wiederum frequenzabhängig ist. Im mittelfrequenten Bereich sind die Umstände offenbar am besten (vgl. Lehnhardt & Laszig 2009, 34).

Überdies lösen die Kopfschwingungen trägheitsbedingte Schwingungen der Gehörknöchelchen aus (vgl. Békésy 1960, 177; Békésy & Rosenblith 1951, 1109; Lehnhardt & Laszig 2009, 34). Da die Knöchelchen mehr oder weniger entlang der Horizontalachse aufgehängt sind, reagieren sie auf laterale Schädelschwingungen besser als auf sagittale (vgl. Békésy 1960, 145; Gelfand 2004, 100; Lehnhardt & Laszig 2009, 34). Die Knöchelchen schwingen wie Pendel relativ zum Kopf und erzeugen so Steigbügelbewegungen im ovalen Fenster. Aus Trägheitsgründen ist die resultierende *tieffrequente* Steigbügelbewegung annähernd gegenphasig zu den Luftschwingungen im Gehörgang (vgl. Békésy 1960, 145; Békésy & Rosenblith 1951, 1109; Gelfand 2004, 100f.).

3.3 Mechanoelektrische Reizwandlung

Im Innenohr werden die herangeführten Signale nach dem tonotopischen Prinzip bereits grob spektral analysiert und schließlich in Nervenimpulse umgewandelt (mechanoelektrische Reizwandlung, *Transduktion*), die vom neuronalen auditiven System weiterverarbeitet werden (vgl. Blauert 1974, 116; Fuchs

2010, 4; Gelfand 2004, 121; Graumann & Sasse 2005, 90; Guski 1996, 91; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37; Müller & Frings 2009, 452; Spitzer 2002, 58; Yost 2008, 83).

Weil der Zugang zum Innenohr schwierig und die dort ablaufenden Prozesse klein und komplex sind, wurde anfangs vieles nur über Theorien, sogenannte *Hörtheorien*, also spekulativ und theoretisch erarbeitet. Auch heute sind die Vorgänge im Innenohr noch nicht hundertprozentig entschlüsselt (vgl. Hellbrück & Ellermeier 2004, 109; Schriever 1967, 18). Die allgemein vertretene theoretische Vorstellung der Abläufe im Innenohr formulierte Georg von Békésy in seiner *Wanderwellentheorie*. Schwingungen des Steigbügels (sowie jede andere Art der Anregung) versetzen die Innenohrlymphe in Bewegung, woraus eine Wanderwelle auf der Basilmembran (BM) resultiert. Die spezielle Struktur der BM entzieht der Wanderwelle ihre spektrale Energie in Laufrichtung von hoch nach tief (*Tonotopie*), die von den Haarzellen abgegriffen und ins neuronale System gespeist wird (vgl. Gelfand 2009, 65; Hellbrück & Ellermeier 2004, 100f.; Kemp 2010, 93ff.; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 36ff.; Pollard & Jansson 1982, 251; Rinsdorf in Keidel 1975, 68ff.; Yost 2008, 90, 92).

3.3.1 Entstehung der Wanderwelle

Schwingungen der Hörknöchelchenkette sorgen für kolbenartige Pumpbewegungen der Steigbügelfußplatte im ovalen Fenster, wodurch Perilymphbewegungen angetrieben werden (vgl. Hellbrück & Ellermeier 2004, 93; Yost 2008, 75; Zwicker & Feldtkeller 1967, 84). Schallwellen in Flüssigkeiten setzen sich aus einer starken Druckwelle und vergleichsweise geringen Flüssigkeitsbewegungen zusammen. Im Rahmen der physikalischen Gegebenheiten des Innenohrs können die Lymphflüssigkeiten als praktisch inkompressibel angesehen werden. Demnach muß ihr Volumen allerdings konstant sein und die vom Steigbügel verdrängte Flüssigkeit augenblicklich woanders Platz finden (vgl. Ranke 1953, 80; Kemp 2010, 105, 114; Schriever 1967, 35; Terhardt 1998, 56). Weil der Aquaeductus cochleae wegen seines kleinen Durchmessers als dämpfender Resonator (Schallabfluß) keine Rolle spielt (vgl. Ranke 1953, 64), können Steigbügelbewegungen nur am runden Fenster ausgeglichen werden (vgl. Békésy & Rosenblith 1951, 1077; Hellbrück & Ellermeier 2004, 93; Kemp 2010, 105; Müller & Frings 2009, 453; Stevens & Hallowell 1938, 255; Winckel 1960, 80). Wegen des größeren Widerstands auf Seiten des Vorhofs breiten sich Druckwellen in der Perilymphe unabhängig vom Anregungsort *immer* in dieselbe Richtung, vom ovalen zum runden Fenster aus (siehe S. 117; vgl. Békésy 1960, 145; Békésy & Rosenblith 1951, 1109f.; Gelfand 2004, 99; Lehnhardt & Laszig 2009, 33f.).

Einwärtsbewegungen des Steigbügels bewirken minimale Druckanstiege in der Perilymphe, die sich als Druckwellen (Volumenverschiebung) entlang der Vorhoftreppe entladen, welche die gesamte Schnecke mit Schallgeschwindigkeit durchlaufen (vgl. Müller & Frings 2009, 453; Ranke 1953, 81). Bei Wasserschallgeschwindigkeiten zwischen rund 1,45 und 1,5 km/s und einer Strecke von zirka 60 bis 70 mm (beide Schneckenotreppen) beträgt die Latenzzeit des runden Fensters um die 40 µs (vgl. Boer & Nuttall 2010, 152; Ranke 1953, 80f.). Während das ovale Fenster noch eingedellt ist, beult sich das runde Fenster schon aus (vgl. Müller & Frings 2009, 453; Ranke 1953, 80). Eine solche Latenz ist im Hörbereich praktisch bedeutungslos. 40 µs entsprechen erst bei 10 kHz einer halben Periode, wobei die korrespondierenden Wellen bereits sehr basal enden, also zu kurze Laufzeiten haben (vgl. Ranke 1953, 80f.).

Verschiebungen der Perilymphe zum runden Fenster werden auch von der nachgiebigen *Schnecken-trennwand* aufgefangen. Diese wird gegen die Paukentreppe ausgebeult, wodurch auch dort Perilymphe verdrängt wird. Gemeinsam leisten die verdrängte Perilymphe und die ihr Equilibrium anstrebende Schnecken-trennwand Widerstand, so daß letztere in die Vorhoftreppe zurück schwingt, wo sich das Spiel umkehrt. Auf diese Weise entstehen aus den Volumenverschiebungen der Perilymphe Schwingungen der Schnecken-trennwand (vgl. Boer & Nuttall 2010, 152; Müller & Frings 2009, 453; Kemp 2010, 114). Hat sich aus der Schwingung eine ausgewachsene Welle gebildet, wird am (noch zu bestimmenden) Punkt ihrer maximalen Amplitude die gesamte Energie aus der Perilymphe durch die Trennwand absorbiert und an die Paukentreppe durchgereicht und am runden Fenster an widerstandsarme Luft in der Paukenhöhle abgestrahlt. Apikal dieses Punktes kommt die Lymphe zur Ruhe und ist in weiterer Folge nicht mehr am Energietransfer vom ovalen zum runden Fenster beteiligt (vgl. Békésy 1960, 419; Békésy & Rosenblith 1951, 1077; Kemp 2010, 114; Müller & Frings 2009, 453; Ranke 1953, 80). Man könnte hier in gewisser Weise von einer Art „akustischem Kurzschluß“ sprechen (wenngleich dieser Begriff eigentlich ein etwas anderes Phänomen der physikalischen Akustik bezeichnet). Die Druckwelle muß

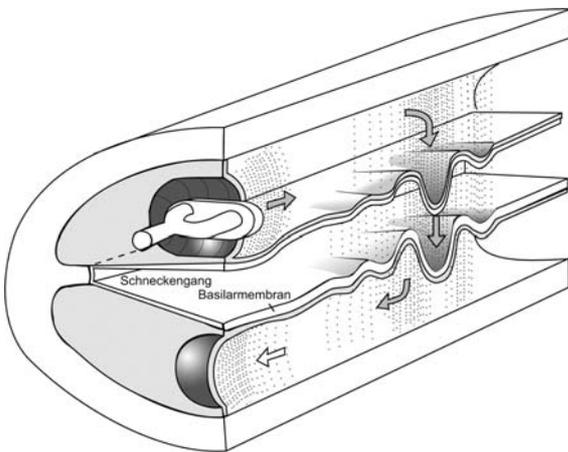


Abb. 3.12 **Druckausgleich** (Müller & Frings 2009, Abb. 19.5, 454).

gegen die Flüssigkeitssäule der Perilymphe in der Vorhoftreppe arbeiten. Da die Eigenfrequenz mechanischer Systeme umso höher ist, je kleiner die beteiligten Massen sind, wird die Druckwelle mit steigender Anregungsfrequenz immer früher auf die Schnecken-trennwand ausweichen (siehe *Exkurs: Zum physikalischen Konzept der Impedanz*, S. 101; vgl. Lehnhardt & Laszig 2009, 28; Stevens & Hallowell 1938, 276f.; Yost 2008, 78). Dies ist ein erster physikalischer Hinweis darauf, wie sich der Ort des „akustischen Kurzschluß“, unabhängig von der Beschaffenheit der Basilarmembran, frequenzbedingt in der Schnecke verschieben kann.

Wellenbewegungen innerhalb der Schnecke werden immer recht schnell auf die BM projiziert. Dabei bewegt sich diese nicht allein. Sie schwingt gemeinsam und prinzipiell phasengleich mit der keinesfalls steifen knöchernen Trennwand, wobei Letztere in Geschwindigkeit und Amplitude um 5 bis 6 dB unter den Werten der BM bleibt (vgl. Békésy 1941(b), 276f.; Stenfelt et al. 2003, 131f., 137, 141, 143). Darüber hinaus bewegen sich auch Reissnermembran und BM praktisch synchron und quasi phasengleich (vgl. Müller & Frings 2009, 453; Schriever 1967, 27). Demnach schwingen alle Elemente des Schneckengangs, also zwischen den Grenzmembranen auch Cortiorgan und Tectorialmembran, in Längsrichtung gleichartig und gleichphasig (vgl. Békésy & Rosenblith 1951, 1095, 1098; Zwislocki 1948, 16), weshalb sie bei Betrachtung der übergeordneten Schneckenphysik zunächst als Einheit aufgefaßt werden können.

Békésy hat dies im Versuch bestätigt. In seinen Modellen hat er den Hohlraum zwischen Reissnermembran und BM (Schneckengang) mit Glycerin oder Honig befüllt, um die Endolymphe zu simulieren, was keinerlei Einfluß auf die festgestellten Bewegungen der Basilarmembran hatte (vgl. Békésy 1960,

423; Zwislocki 1948, 15f.). Daraus läßt sich außerdem schließen, daß die mechanischen Eigenschaften des Schneckengangs hauptsächlich durch die Basilarmembran bestimmt sind (vgl. Schriever 1967, 21f.).

Resultat dieser Erkenntnis ist die vereinfachende Zusammenfassung des Schneckengangs unter den Begriffen *Basilarmembran* oder (*Schnecken-*)*Trennwand* (cochlea partition) (vgl. Békésy 1960, 407f.; Schriever 1967, 22), was leider oft zu Unterscheidungsschwierigkeiten führt, wenn nicht klar ist, ob auf die Basilarmembran als solche, oder den gesamten Schneckenang Bezug genommen wird. Hier soll daher in weiterer Folge der zusammengefaßte Komplex als *Schneckentrennwand* (knöcherne Trennwand und Schneckenang) oder *Schneckenang* (Trennwand ohne Lamina spiralis ossea) bezeichnet werden.

Druckbewegungen in der Vorhoftreppe lösen Transversalschwingungen der Schneckentrennwand aus. Diese sind mit den Transversalbewegungen von Wasseroberflächenwellen vergleichbar, nur daß nicht Gravitation, sondern – wie schon gezeigt – die ihre Balance suchende Trennwand sowie der Flüssigkeitswiderstand als Rückstellkräfte dienen (vgl. Boer & Nuttall 2010, 152; Kemp 2010, 114; Müller & Frings 2009, 453f.; Zwislocki 1948, 16). Im Wasser erzeugen Wellenberge an der Oberfläche Druck nach unten. Dieser will den Wellenberg absenken und deshalb die darunterliegende Wassersäule verdrängen. Wegen der trägheitsbedingten Verzögerung des Verdrängungs- oder Ausgleichsvorgangs beginnt die Welle, sich entlang der Oberfläche auszubreiten. Gleiches gilt im Innenohr, wo sich Perilymphe und Schneckentrennwand gegenseitig abwechselnd verdrängen (vgl. Boer & Nuttall 2010, 152; Kemp 2010, 114). Die Fortpflanzungsrichtung der entstehenden Welle wird durch die besondere Struktur der Basilarmembran mit ihrer vom Steigbügel zum Schneckenloch abnehmenden Steifheit bestimmt. Klarerweise will die verdrängte Lymphe lieber mit elastischeren (apikalen) als steiferen (basalen) Trennwandabschnitten wechselwirken. Wellen wandern also immer von den steifen zu den elastischen Bereichen (vgl. Gelfand 2004, 127; Spitzer 2002, 59). So lösen Druckbewegungen in der Vorhoftreppe laufende Wellen, sogenannte *Wanderwellen*, entlang des Schneckentrennwand aus (vgl. Békésy 1960, 419; Békésy & Rosenblith 1951, 1091; Hellbrück & Ellermeier 2004, 100; Kemp 2010, 114; Reiss 2003, 622).

Nicht nur die Laufrichtung, sondern praktisch alle Eigenschaften der Wanderwelle sind durch die Beschaffenheit der BM determiniert. Zunächst ist allerdings nur interessant, daß jede Welle an einem bestimmten, frequenzabhängigen Punkt auf der BM ihr Maximum erreicht und danach relativ schnell verebbt (vgl. Békésy & Rosenblith 1951, 1106; Gelfand 2004, 128; Hellbrück & Ellermeier 2004, 100; Reiss 2003, 622; Schriever 1967, 23; Spitzer 2002, 59; Zwislocki 1948, 16).

Drückt der Steigbügel das ovale Fenster herein, drückt die Lymphe die Schneckentrennwand Richtung Paukentreppe und umgekehrt. Gleichzeitig entstehen in beiden Treppen gegensinnige Längsströmungen der Perilymphe zwischen Fenster und Auslenkungsort auf der BM (vgl. Schriever 1967, 18, 35). Aufgrund der komplizierten Volumenverhältnisse im Innenohr werden die Transversalbewegungen der Schneckentrennwand von Ausgleichsbewegungen in den Perilymphkanälen begleitet, die sich beim Wellenmaximum zu Flüssigkeitswirbeln ausbilden (vgl. Békésy 1960, 420; Békésy & Rosenblith 1951, 1102, 1105; Ranke 1953, 84; Stevens & Hallowell 1938, 279, 286). Diese Wirbel manifestieren die lokalen Volumenverschiebungen in der Perilymphe. Ein Wellenberg treibt sozusagen eine Bugwelle vor sich her. Die verdrängte Flüssigkeit wird vom „Platzüberangebot“ des nachfolgenden Wellentals angesaugt. So entstehen gegensinnige Strömungen in beiden Treppen, die an der Trennwand in Ausbreitungsrichtung (zum Schneckenloch), und an der Knochenwand gegen die Laufrichtung (zu den Fenstern) fließen (vgl.

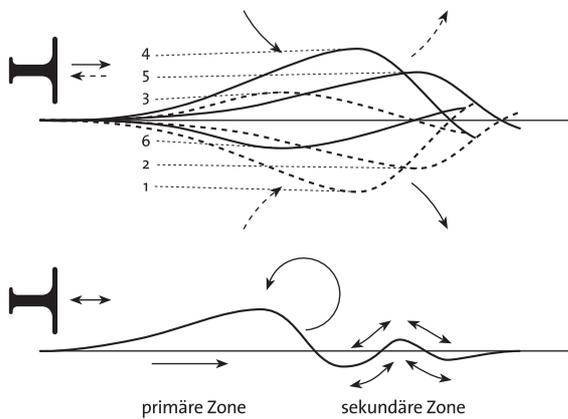


Abb. 3.13 **Wirbel** (modifiziert n. Békésy 1960, Abb. 11.21 & 11.22, 421f.). Oben: Steigbügel-, Lymph- und BM-Bewegung während einer ganzen Periode (Phasen der Welle aufsteigend nummeriert). Unten: Wellenbewegung der BM und die korrespondierenden Lymphbewegungen.

Wirbel mehr zu finden. Hier fließt die Lymphe nur noch an den verebbenden Ausschlägen entlang (vgl. Schriever 1967, 19, 35).

Interessanterweise stellte sich im Versuch heraus, daß es so gut wie unmöglich ist, durch alle möglichen Veränderungen der Schneckenform die Wirbelposition (und entsprechend des Wellenmaximums) zu beeinflussen. Aus diesem Grund ist anzunehmen, daß die sehr unregelmäßige Verjüngung der Schnecke relativ unbedeutend und der Verlauf der Wanderwelle hauptsächlich auf die mechanischen Eigenschaften der BM zurückzuführen ist (vgl. Békésy & Rosenblith 1951, Abb. 46., Abb. 47, 1106, 1105; Ranke 1953, 84; Yost 2008, 92; Zwislocki 1948, 16f., 32f.).

Rinsdorf hält es für unmöglich, daß die von Békésy beobachteten Wirbel bei physiologischen Pegeln überhaupt auftreten können. Rinsdorf führt in diesem Sinne Berechnungen von Zwislocki ins Feld (vgl. Rinsdorf in Keidel 1975, 70). Bei Zwislocki wird (an anderer Stelle, als der von Rinsdorf zitierten) die Existenz der Wirbel m.E. jedoch klar bestätigt und ihre Bedeutung (mehr Wirkung als Ursache) im Einklang mit Békésy bewertet (vgl. Zwislocki 1948, 57ff., 72).

3.3.2 Physik der Wanderwelle

Schneidet man den gedanklich entrollten Schneckenkanal der Länge nach durch und betrachtet seinen Längsschnitt, stellt sich die Basilarmembran (BM) als simpler Strich dar, an dem sich die Entwicklung der Wanderwelle entlang der Schnecke sehr gut beobachten läßt. Es fallen ganz offensichtliche Grundeigenschaften der Welle auf: Sie läuft frequenzabhängig unterschiedlich tief in die Schnecke hinein und ihre Amplitude wächst zunehmend an, während Geschwindigkeit und Wellenlänge gleichzeitig abnehmen. Am Ort ihres Amplitudenmaximums bricht die Welle ziemlich abrupt ab und läßt alle weiter apikal verorteten Schneckenbereiche in Ruhe (vgl. Békésy & Rosenblith 1951, 1099, 1101; Boer & Nuttall 2010,

Békésy 1960, 421; Békésy & Rosenblith 1951, Abb. 39., 1101; Schriever 1967, 19; Zwislocki 1948, 57).

Die Wirbeldurchmesser entsprechen immer der Flüssigkeitssäule über der Membran und ihre Rotationsgeschwindigkeiten der Steigbügelamplitude (vgl. Békésy 1960, 420; Schriever 1967, 18f.). Wirbel sind immer am Ort des Wellenmaximums zu finden, das bei hohen Frequenzen eher zum Steigbügel und bei tiefen Frequenzen eher zum Schneckenloch rückt (vgl. Békésy 1960, 420; Békésy & Rosenblith 1951, 1105; Schriever 1967, 19). Bei besonders starken Schwingungen können Wirbel auch nahe am Steigbügel auftreten, jenseits (apikal) des Wellenmaximums, wo nur noch schwache Ausschläge das Absterben der Welle bezeugen, sind aber offenbar grundsätzlich keine

153, 167; Gelfand 2004, 127f.; 2009, 66; Hellbrück & Ellermeier 2004, 100; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 36f.; Zwicker & Feldtkeller 1967, 85, 88; Zwislocki 1948, 16, 32).

Basal ist die BM steifer als apikal (vgl. Békésy & Rosenblith 1951, 1101). Dies hat Folgen. In Systemen mit ähnlichen Elastizitätseigenschaften schwingt der steife vor dem elastischen Teil. Dadurch ist die Laufrichtung der Wanderwelle vom steifen, impedanzreichen, basalen Ende zum nachgiebigen, impedanzarmen, apikalen Ende der BM determiniert (vgl. Gelfand 2004, 127; Spitzer 2002, 59). Zusätzlich wird auch die Form der Welle durch die BM beeinflusst. Mit abnehmender Steifheit wird die Welle immer kürzer, weil die Membran biegsamer wird. In anderen Worten: mit zunehmender Distanz vom Steigbügel nimmt die Wellenlänge kontinuierlich ab (vgl. Békésy & Rosenblith 1951, 1091, 1099, 1101; Boer & Nuttall 2010, 153, 167; Kemp 2010, 115; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37; Rinsdorf in Keidel 1975, 69; Zwislocki 1948, 16, 32). Geht man davon aus, daß die Reizfrequenz konstant bleibt, muß, bei schrumpfender Wellenlänge, gemäß des Zusammenhangs $\lambda = c : f$ auch die Ausbreitungsgeschwindigkeit abnehmen, was mit zunehmender Steigbügeldistanz tatsächlich der Fall ist. Von basal etwa 100 m/s bleiben apikal noch etwa 2 bis 3 m/s übrig (vgl. Békésy & Rosenblith 1951, 1101; Gelfand 2004, 127; Rinsdorf in Keidel 1975, 69, 72; Schriever 1967, 28; Spreng in Keidel 1975, 130).

Während Geschwindigkeit und Länge einer Welle auf der BM bei konstanter Frequenz abnehmen, wächst ihre Amplitude kontinuierlich an, da Elastizitätszuwachs und Verlangsamung die Wellenenergie auf immer kürzere Membranabschnitte konzentrieren, bis die Welle ihren Resonanzpunkt erreicht. Dort ist ihre Amplitude maximal und die Membranimpedanz minimal, weil sich die abnehmende Federreaktanz (Rückstellkraft) und die zunehmende Massereaktanz (Trägheit) frequenzspezifisch neutralisieren (vgl. Gelfand 2004, 128; Kemp 2010, 114f.; Nielsen & Russell 2000, 11751; Spitzer 2002, 59). Weiter apikal wird die Belastung durch die weiterhin zunehmende Massereaktanz zu groß und die Welle kommt zum Erliegen (vgl. Nielsen & Russell 2000, 11751; Zwicker & Feldtkeller 1967, 85, 88). Anders gesagt schluckt die BM an diesem Punkt größter Nachgiebigkeit praktisch alle Energie und gibt sie an die Paukentreppe ab. Dadurch kommt die Welle zum Stillstand. Da sich die Lymphbewegungen an diesem Punkt ausgleichen, findet zwischen Resonanzpunkt und Schneckenloch keine Lymphbewegung mehr statt. Folglich kann man Reflexionen in der Schnecken Spitze ausschließen, die für stehende Wellen in der Schnecke sorgen könnten (vgl. Békésy 1960, 419; Gelfand 2004, 128; Ranke 1953, 80; Spitzer 2002, 59). Apikal des Resonanzpunktes fällt die Hüllkurve der Welle steiflankig ab und Amplitude und Wellenlänge verringern sich rasch, so daß die Welle jenseits des Maximums praktisch verebbt (vgl. Békésy 1943, 68; Gelfand 2004, 128; Hellbrück & Ellermeier 2004, 100; Schriever 1967, 19; Zwislocki 1948, 16). Der Verlauf von hohen zu tiefen Frequenzen verhindert die Entstehung von Untertönen, weil die gesamte Energie bereits am Resonanzpunkt des Grundtons ausgeglichen ist und tiefere Frequenzbereiche in Ruhe bleiben (vgl. Békésy 1941(b), 276).

Man nimmt an, daß die transversale Schwingung der Wanderwellen zu radialen Scherbewegungen führt, durch welche die Stereozilien ausgelenkt und die Haarzellen erregt werden (siehe 3.3.3, S. 129; vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 131ff.; Nilsen & Russell 1999, 642; Schriever 1967, 33). Diese Scherbewegungen sind aber nicht überall entlang der Wanderwelle gleich, vielmehr ist die Welle in eine *primäre* und eine *sekundäre Zone* teilbar (siehe Abb. 3.13, S. 121). Der Energietransfer vom ovalen zum runden Fenster findet in der primären Zone statt. Ein sehr kleiner Teil der Energie „überlebt“ und unterhält den sekundären Abschnitt der Welle (vgl. Gelfand 2004, 131; Schriever 1967, 35). Nur in unmittel-

barer Umgebung des Wellenmaximums dominieren transversale Bewegungen die Schwingung. Basal davon finden hauptsächlich radiale Scherbewegungen statt, apikal finden sich, im sehr kurzen Bereich der sekundären Zone, longitudinale Bewegungen (vgl. Gelfand 2004, 131; Schriever 1967, 33; Winkel 1957, 84). Da die radialen Scherbewegungen Konsequenzen der transversalen Schwingung sind, muß man radiale und transversale Bewegungen m.E. als zwei Komponenten *einer* Bewegung auffassen, deren Amplitude hauptsächlich durch den transversalen Anteil bestimmt wird. Somit besteht die Wanderwelle bis zum Maximum aus einem anwachsenden Wellenberg, während die longitudinalen Schwingungen in der zweiten Zone mehrere Wellenberge und -täler formen. Kurz gesagt: *Vor* dem Maximum bewegt sich die gesamte Welle in Phase, *nach* dem Maximum jedoch mehrphasig (vgl. Békésy 1960, 42ff.; Schriever 1967, 19).

Frequenzverteilung auf der Basilarmembran

Da jeder Ort auf der BM andere Elastizitätseigenschaften aufweist, spricht auch jeder Punkt auf eine andere Frequenz besonders stark an, wodurch Wanderwellen frequenzabhängig an verschiedenen Orten ihr Maximum ausbilden (vgl. Békésy & Rosenblith 1951, 1091, 1097, 1106; Boer & Nuttall 2010, 154; Gelfand 2004, 128; Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 100; Kemp 2010, 114; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37f.; Müller & Frings 2009, 458; Nilsen & Russell 1999, 642; Reiss 2003, 622; Schriever 1967, 23ff., 31). Weil Elastizität tiefe Frequenzen eher begünstigt als hohe, nimmt die Eigenresonanz (jene Frequenz, auf welche die Membran mit größter Nachgiebigkeit reagiert) vom Steigbügel zum Schneckenloch kontinuierlich ab. Demnach resonieren hohe Frequenzen an der Basis, die tiefsten in der Schneckenspitze (vgl. Békésy 1941b, 266; Békésy & Rosenblith 1951, 1097; Boer & Nuttall 2010, 139; Fuchs 2010, 4; Hellbrück & Ellermeier 2004, 100; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37f.; Müller & Frings 2009, 458; Nilsen & Russell 1999, 642; Rinsdorf in Keidel 1975, 68f.; Spitzer 2002, 59f.; Stevens & Hallowell 1938, 356; Stumpf 1890, 90f.; Terhardt 1998, 57; Winkel 1960, 80; Yost 2008, 94f.; Zwicker & Feldtkeller 1967, 85, 87). Die Resonanzfrequenz der BM wird in Analogie zur Lieblingsfrequenz der korrespondierenden Nervenfasern auch gerne als *charakteristische Frequenz* (CF) bezeichnet.

Ihre Anatomie befähigt die Schnecke, bis zu einem gewissen Grad „Schwingungen des Steigbügels verschiedener Frequenz längs der Schneckenwand räumlich auseinanderzuziehen“ (Békésy 1943, 66; vgl. Hellbrück & Ellermeier 2004, 102; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37). Läuft ein (in der Natur, also auch in der Musik, üblicherweise) breitbandiges Signal die BM hinab, resonieren die kurzwelligen Anteile des Spektrums schon recht nah hinter dem ovalen Fenster, während die langwelligen Komponenten noch tiefer in die Schnecke hinein „die Welle reiten“. Das Spektrum erfährt sozusagen eine räumliche Spreizung. Dieses Phänomen, *Dispersion* genannt, tritt immer dann auf, wenn die Ausbreitungsgeschwindigkeit frequenzabhängig ist (vgl. Kemp 2010, 110, 115; Ranke 1953, 86; Schriever 1967, 31). Durch Dispersion werden Wellen auf ihrem Weg zum Schneckenloch in ihre Einzelteile (Teiltöne) zerlegt, wobei tiefere Teiltöne weiter weg vom Steigbügel verarbeitet werden (vgl. Boer & Nuttall 2010, 139, 143; Hackney & Furness 2010, 231; Nilsen & Russell 1999, 642). Zum besseren Verständnis ließe sich die Frequenzverteilung der Schnecke mit einer Klaviatur vergleichen, auf der

bekanntlich jeder Taste eine bestimmte Grundtonhöhe zugeordnet ist. Frequenzen sind an Tasten und über deren festgelegte Anordnung an einen bestimmten Ort gebunden (vgl. Müller & Frings 2009, 457; Winckel 1960, 80). Analog dazu korreliert jede Reizfrequenz mit einem bestimmten Ort auf der BM. Allein der Ort des Wellenmaximums enthält Informationen zur Signaltonhöhe, es findet also eine Ortskodierung der Frequenzen statt. Dieses *Tonotopie* genannte Ordnungsprinzip bildet die Grundlage der Frequenzselektivität des Gehörs (vgl. Boer & Nuttall 2010, 145; Hellbrück & Ellermeier 2004, 102; Spreng in Keidel 1975, 133; Stevens & Hallowell 1938, 277f.).

Allerdings beeinflussen tiefe Frequenzen höhere Frequenzen immer mit. Dies äußert sich vor allem darin, daß tiefe Töne hohe Töne in weitaus höherem Maß verdecken, als es umgekehrt möglich wäre (siehe 1.6.2, S. 36; vgl. Fricke & Louven 2008, 416; Gelfand 2004, 314; Hellbrück & Ellermeier 2004, 129; Reuter 1996, 29; Spitzer 2002, 67; Zwicker & Feldtkeller 1967, 66). Die Begründung dafür ist in der Laufrichtung der Wanderwelle von hohen zu tiefen Frequenzen zu suchen. Während hohe Töne bereits nahe des Steigbügels stehen bleiben, laufen tiefe Frequenzen weit in die Schnecke hinein. Jeder tiefe Ton „überrollt“ die höheren Frequenzorte sozusagen, während die hohen Frequenzen stoppen, bevor sie tiefe Frequenzen ernsthaft beeinflussen können. Tiefe Frequenzen interferieren folglich mit der Wahrnehmung hoher Frequenzen, heben z.B. deren Mithörschwelle an, was umgekehrt nicht möglich ist (vgl. Gelfand 2004, 316; Guski 1996, 246; Spitzer 2002, 67; Ranke 1953, 95f.). Für den Klaviaturvergleich bedeutet das, jeden Ton mit einem Glissando von der tiefsten Taste aus zu intonieren.

Für Frequenzen unter 25 Hz ist keine mechanische Frequenzanalyse festzustellen. Der Spitzenbereich wird vollständig ausgelenkt, bis 30 Hz ist die gesamte apikale Windung betroffen (vgl. Békésy 1943, 67; Békésy & Rosenblith 1951, 1106f.; Schriever 1967, 25). Tiefe Frequenzen regen die gesamte BM an, wobei sich ihr Maximum mit steigender Frequenz Richtung Basis verschiebt. Erst ab 50 Hz ist eine Spitze erkennbar, ab 150 Hz können Wanderwellen beobachtet werden (vgl. Békésy & Rosenblith 1951, 1099; Gelfand 2004, 128; Schriever 1967, 25). Ab etwa 800 Hz ist die apikale Windung nicht mehr am Schwingungsgeschehen beteiligt (vgl. Békésy 1943, 67). Zwischen 300 und 1000 Hz werden sowohl mechanische Frequenzdiskrimination, als auch Tonhöhenwahrnehmung merklich besser (vgl. Békésy & Rosenblith 1951, 1107). Wie Abbildung 3.15 zeigt, werden die Wellenspitzen mit steigender Frequenz spitzer, wobei sich im mittelfrequenten Bereich zwischen etwa 0,2 und 3 kHz nach Békésy (vgl. 1943, 72) sehr ähnliche bis deckungsgleiche Resonanzkurven ausbilden.

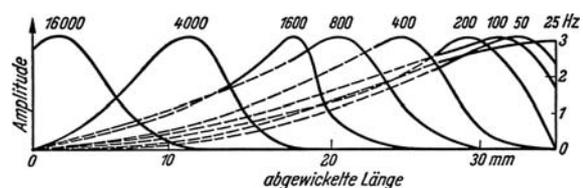


Abb. 3.14 **frequenzspezifische Wellenspitzen** (Winckel 1960, Abb. 45, 81). Je tiefer die Frequenz ist, umso weiter apikal (rechts) liegt das Wellenmaximum (umso länger ist auch die Welle) und umso flacher ihr Kurvenverlauf.

Älteren Überlegungen zufolge könnte sich das Maximum mit steigendem Schallpegel verlagern. Höhere Pegel führen zu weiteren Auslenkungen der BM. Der Elastizitätsmodul nimmt bei lauterem Schall zu (vereinfacht heißt das: je weiter die BM-Auslenkung, desto geringer ist ihre Reserve, auf weitere Auslenkung elastisch zu reagieren). Stärkere Auslenkungen versteifen die BM also, wodurch die Wellengeschwindigkeit steigt, ihre Länge anwächst und ihr Maximum etwas weiter läuft (vgl. Ranke 1953, 93f., 102; Winckel 1960, 84). Als Erklärung für pegelabhängige Verschiebungen empfundener Tonhöhen (siehe 1.5.2, S. 31) kommt das Modell allerdings weniger in Frage (vgl. Moore 1993, 94f.).

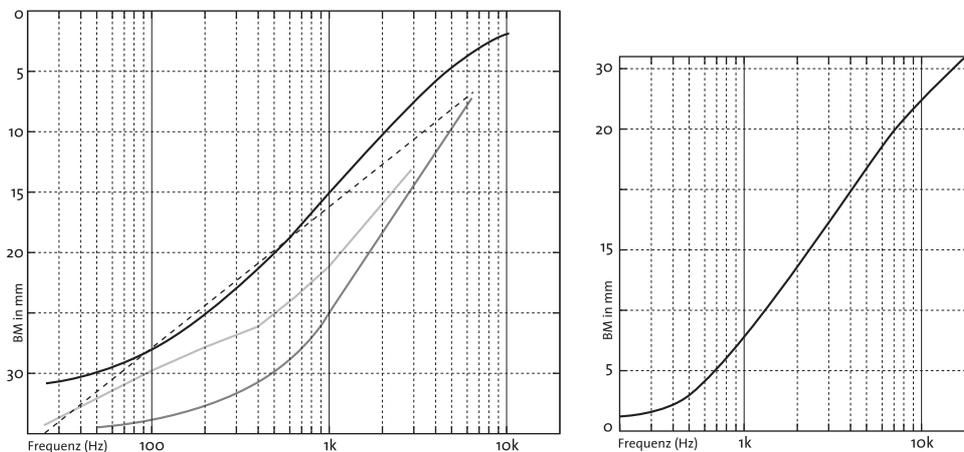


Abb. 3.15 **Frequenzverteilung auf der Basilarmembran.** Grundsätzlich gilt: je steiler die Kurve, umso besser die Frequenzauflösung. Links: Verteilung nach Békésy (hellgrau) bzw. Wegel & Lane (schwarz). Streng logarithmische Frequenzverteilung würde der unterbrochenen Linie entsprechen. In dunkelgrau: „Ort der Haarzellenschädigung für Ausfälle im Audiogramm“ (Ranke 1953, Abb. 71, 92). Rechts: Verteilung nach Winckel (1960, Abb. 46, 82). Die Kurve verdeutlicht den Wechsel von linearer zu logarithmischer Frequenzverteilung um 500 Hz. Interessanterweise deutet die auf der Ordinate (Y-Achse) abgetragene Steigbügelentfernung an, daß tiefe Frequenzen basal und hohe apikal verortet sind.

Ausgehend von Békésys Meßergebnissen (siehe Abb. 3.15) wurden diverse Kurven errechnet, die keine lineare Frequenzverteilung entlang der BM vorhersagen. Speziell im mittleren Frequenzbereich deuten steiler verlaufende Kurven bessere Frequenzauflösung und somit weiter auseinander liegende Wellenmaxima auf der BM an (vgl. Ranke 1953, 91f.). Heute nimmt man an, daß gleiche Frequenzverhältnisse (z.B. Oktaven) ungefähr gleiche Abstände auf der BM einnehmen. Mit anderen Worten: Die Frequenzverteilung auf der BM ist logarithmisch organisiert. Von einer 32 mm langen Membran entfielen somit 3,2 mm auf jede der zehn Oktaven des Hörbereichs (vgl. Kemp 2010, 115; Müller & Frings 2009, 458; Winckel 1960, 81f.). Zwar ist die Funktionsweise der *Frequenzgruppen* (siehe 1.3, S. 18) weiterhin unklar, doch kann man mit großer Sicherheit annehmen, daß ihre physiologischen Grundlagen im Innenohr liegen und eng mit dem Frequenzauflösungsvermögen der BM verknüpft sind (vgl. Hellbrück & Ellermeier 2004, 127; Moore 1993, 61, 81; Spreng in Keidel 1975, 142). Die Frequenzgruppenbreiten werden auf der BM mit Weiten zwischen ungefähr 0,9 bis 2 mm, meistens jedoch mit 1,3 mm angegeben. Die korrespondierenden ERB betragen zirka 0,9 mm (vgl. Fastl & Zwicker 2006, 162; Gelfand 2004, 322; Hellbrück & Ellermeier 2004, 128; Kemp 2010, 115; Moore 1993, 81; Spreng in Keidel 1975, 142; Winckel 1957, 84). Die Frequenzgruppenbreite definiert die Grenze, bis zu der saubere Tontrennung noch möglich ist. Liegen zwei Töne zu nahe beisammen, beginnen sie zu „verschwimmen“, so daß letztlich nur *ein* rauher oder schwebender Ton ($f = f_1 - \Delta f_{1,2} : 2$) gehört wird (vgl. Spitzer 2002, Abb. 3.9, 60, 94, 100). Geht man von Wellenspitzenbreiten um 4 mm (vgl. Zwicker & Feldtkeller 1967, 89) aus, die sich zwangsläufig ineinander schieben wenn ihre Mittenfrequenzen zusammenrücken, wird deutlich, warum die Signale miteinander verschwimmen. Sobald die zur Differenzierung bedeutsame Zwischenrinne verschwindet, verschmelzen beide Wellen zu einer gemeinsamen Membranauslenkung, aus der ein einziges Hörereignis entsteht. Beide Töne werden nun durch dieselben Nervenfasern „abgenommen“, so

daß Differenzierungen kaum noch möglich sind (vgl. Spitzer 2002, Abb. 3.9, 60, 100). Aus dem gleichen Grund verändert sich auch die Lautheitsaddition: Innerhalb einer Frequenzgruppe fallen die evozierten Nervenimpulse größtenteils zusammen. Dies entspricht praktisch einem einzigen Hörereignis, dessen Lautstärke durch die doppelte Anzahl von Nervenimpulsen kodiert ist, was einer Intensitätsverdopplung gleich kommt (vgl. Hellbrück & Ellermeier 2004, 88; 140; Stevens & Hallowell 1938, 130f.).

Es ist naheliegend, daß auch die Mithörschwelle innerhalb der Frequenzgruppe auf diese Weise gebildet wird. Eine Erklärung ist die Hypothese der *besetzten Leitung*. Hiernach belastet der Maskierer durch Erregung einen eng begrenzten Bereich von Nervenfasern und hindert diese daran, durch den gesuchten (maskierten) Nutzschall effektiv erregt zu werden (vgl. Guski 1996, 245; Hellbrück & Ellermeier 2004, 135; Spitzer 2002, 68; Stevens & Hallowell 1938, 127). Allerdings müßte man die Hypothese erweitern, um dem Problem zu begegnen, daß Nervenfasern nur Dynamikspannen von etwa 30 bis 40 dB abdecken (vgl. Guski 1996, 246; Hellbrück & Ellermeier 2004, 111; Pollard & Jansson 1982, 251; Yost 1993, 5), während aber wesentlich weitere Spannen verdeckungsrelevant sein können. Eine Lösung dafür wäre die Annahme, daß der Maskierer die durch das Nutzsignal erzeugte Synchronisierung der Fasern untergräbt (vgl. Guski 1996, 246).

Laufzeit der Wanderwelle

Da sich die Welle entlang der BM fortbewegt ist klar, daß Zeit vergeht, bevor sie das Schneckenloch erreicht (vgl. Békésy 1943, 73; Békésy & Rosenblith 1951, 1100). Die Zeit, in der eine Frequenz ihren Resonanzpunkt erreicht, läßt sich errechnen, indem man die Periodenzahl (n) mit der Periodendauer ($T : f$) multipliziert ($n : f = s$). Dies ist möglich, weil nach Kemp frequenzunabhängig alle Wellen in etwa gleich viele Perioden durchlaufen, bis sie ihren Resonanzpunkt erreichen. Anhand der Latenzen otoakustischer Emissionen (OAE, siehe 3.3.5, S. 136), belaufen sich Schätzungen auf 3 bis 9 Perioden (vgl. Kemp 2010, 115) bzw. nur 2,5 bis 3,5 Perioden (vgl. Robles & Ruggero 2001, zit. nach Lukashkin et al. 2010, 32). Die von Kemp dargestellte Rechnung widerspricht der postulierten Laufzeit einer Welle von Steigbügel zum Schneckenloch von rund 5 ms (vgl. Békésy & Rosenblith 1951, 1101; Schriever 1967, 28). Geht man von 20 Hz als tiefstmöglicher Frequenz aus, würden 5 ms Laufzeit lediglich einer Zehntelperiode entsprechen. Rechnet man bei gleicher Frequenz mit drei Perioden, kommt man auf eine Laufzeit von 150 ms.

Weil sich die Phase laufender Wellen mit der Zeit (quasi von Ort zu Ort) verschiebt, kann auch aus der Phasenbeziehung von Steigbügel und fraglichem Wellenort auf der BM die Wellenlaufzeit errechnet werden (vgl. Schriever 1967, 27; Yost 2008, 97f.). In unmittelbarer Steigbügelnähe schwingt die BM grundsätzlich quasi gleichphasig (vgl. Békésy & Rosenblith 1951, 1101). Die Laufzeiten zwischen Steigbügel und Schneckenspitze sind relativ konstant. Da Wellen höherer Frequenz schneller schwingen (kürzere Periodendauer), ist das Ausmaß der Phasenverschiebung pro Laufzeit frequenzabhängig (vgl. Boer & Nuttall 2010, 146; Gelfand 2004, 128f; Schriever 1967, 27; Yost 2008, 98).

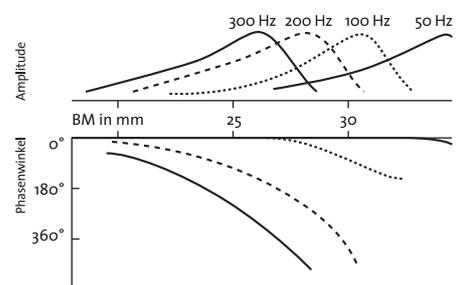


Abb. 3.16 Phasenlaufzeit der Wanderwelle (Yost 2008, Abb. 7.17, 98).

Dabei sollte man aber im Auge behalten, daß die Ausbreitungsgeschwindigkeit entlang der BM abnimmt, was die Phase zusätzlich verzögert. Eine 200 Hz-Welle hat eine Periode von 5 ms. Etwa 27 mm vom Steigbügel entfernt, beträgt der gemessene Phasenunterschied zum Steigbügel 180°. Das bedeutet, die Welle hat zum Meßort eine Laufzeit von 2,5 ms benötigt (vgl. Yost 2008, 98).

Die Laufzeit der Wanderwelle kann hörbar gemacht werden, indem man die tonotopische Organisation des Gehörs ausnutzt. Durch dichotische Beschallung mit einem hochfrequenten bzw. tieffrequenten Dauerton „maskiert“ man die entsprechenden Frequenzbereiche. Sendet man nun einen diotischen Klick, entsteht das korrelierende Hörereignis nicht erwartungsgemäß in der Mitte, sondern zum tieffrequent maskierten Ohr hin verschoben (vgl. Békésy 1943, 73; Békésy & Rosenblith 1951, 1101). Wer sich hier spontan an das psychoakustische *Gesetz der ersten Wellenfront* (siehe 1.7.3, S. 45; vgl. Blauert 1974, 163, 177, 184ff.; Culling & Akeroyd 2010, 135f.; Gelfand 2004, 412f.; Guski 1996, 285) erinnert fühlt, liegt vermutlich nicht ganz falsch. Békésys Versuch legt den Schluß nahe, daß völlig egal ist, an welcher Stelle sich die bilateralen Signallaufzeiten zueinander verzögern, da sie erst im neuronalen System gegeneinander verrechnet werden (vgl. Müller & Frings 2009, 464; Spitzer 2002, 74ff).

Die Laufzeit der Wanderwelle ist nicht nur hörbar, sondern auch aus musikalischer Sicht von Bedeutung. Daß die hohen Töne vor den tiefen erregt werden sorgt dafür, daß Klangspektren, deren Teiltöne gleichzeitig einsetzen, sich im Ohr von hohen zu tiefen Teiltönen kaskadierend aufbauen, was erklären könnte, warum die Ansprache von Musikinstrumenten besser beurteilt wird, wenn hohe Teiltöne möglichst früh einsetzen (vgl. Reuter 1995, 32).

Amplituden der Basilarmembran

Nach Békésy bewegen sich die BM-Amplituden in der Größenordnung von Elementarteilchen (vgl. Békésy & Rosenblith 1951, 1113; Boer & Nuttall 2010, 143; Keidel 1975, 31; Ranke 1953, 102; Schriever 1967, 35). Békésy konnte außerdem pegelbezogen ein relativ lineares, mit Lautsprechersystemen vergleichbares Schwingungsverhalten der BM beobachten (vgl. Békésy 1943, 67; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37). Dabei stellte er generell recht flache Resonanzkurven fest (vgl. Békésy 1943, 72). Man muß allerdings berücksichtigen, daß Békésy seine Messungen nur mit (unphysiologisch) hohen Schallintensitäten an Leichenohren durchführen und nur visuell ablesen konnte (vgl. Hellbrück & Ellermeier 2004, 102f.; Rinsdorf in Keidel 1975, 69).

Spätere Messungen *in vivo* konnten Amplituden der BM im Nanometerbereich nachweisen (vgl. Hellbrück & Ellermeier 2004, 103; Müller & Frings 2009, 454; Zwicker & Feldtkeller 1967, 87). Nach Hellbrück und Ellermeier liegen die kleinsten meßbaren Amplituden in der Größenordnung von rund 1 nm (10^{-9} m), was in etwa dem 30-fachen Mittelwert von Atomdurchmessern entspricht (vgl. Hellbrück & Ellermeier 2004, 103; Zwicker & Feldtkeller 1967, 87). Nach Zwicker und Feldtkeller kommen die Amplituden von Schalldrücken um $3 \cdot 10^{-2}$ μ b (3000 μ Pa, 43,5 dB_{SPL}, ruhige Umgangssprache) sogar der Größenordnung von Atomdurchmessern gleich. Damit wäre nicht nur Békésys Annahme prinzipiell bestätigt, sondern es verdeutlicht auch die enorme, bis an die Grenzen des Sinnvollen ausgereizte Empfindlichkeit des Gehörs (vgl. Gelfand 2004, 279; Spitzer 2002, 78; Zwicker & Feldtkeller 1967, 87). Auf der anderen Seite stehen Auslenkungen der BM des lebenden Ohrs in *keinem* linearen Verhältnis zum

Schalldruck (vgl. Hellbrück & Ellermeier 2004, 103; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37; Ranke 1953, 93). Békésys mittels hoher Intensitäten gewonnene Daten können für kleine Pegel nicht einfach „runtergerechnet“ werden (vgl. Boer & Nuttall 2010, 151). Nur bei niedrigen Pegeln reagiert die BM linear. Ab etwa 30 bis 40 dB nimmt die Membranamplitude im Verhältnis zum Signalpegel immer weniger zu. Beträgt das Verhältnis von stärkstem zu schwächstem Stimulus 80 dB, liegt das Verhältnis der entsprechenden Membranamplituden lediglich bei rund 55 dB. Je höher der Reizpegel wird, desto geringer fällt der Amplitudenzuwachs auf der BM aus (vgl. Boer & Nuttall 2010, 149ff.; Hellbrück & Ellermeier 2004, 103; Nilsen & Russell 1999, 646; Ranke 1953, 93).

Wenngleich die tonotopische Frequenzverteilung gutes Startkapital für die feine Frequenzauflösung des Gehörs ist, kann gerade diese Feinheit nicht allein durch die physikalisch erzeugte Wellenspitze erklärt werden. Die Flanken der Wellenspitze sind (speziell bei Sinustönen) zu flach, die Resonanzkurve zu weit. Nach Zwicker und Feldtkeller (vgl. 1967, 89) sind die physikalischen Wellenspitzen etwa 4 mm breit. Rund 2 mm weiter apikal kann die Welle als praktisch erloschen gelten. Teilweise werden die Ausbuchtungsbereiche der Wellenmaxima mit rund zwei bis drei Oktaven sogar noch weiter bemessen (vgl. Békésy & Rosenblith 1951, 1107; Boer & Nuttall 2010, 157; Hellbrück & Ellermeier 2004, 102; Lehnhardt & Janssen in Lehnhardt & Laszig 2009 38; Ranke 1953, 87, 103; Rinsdorf in Keidel 1975, 72; Winckel 1960, 80; Zwicker & Feldtkeller 1967, 109). Dennoch entspricht die Frequenzauflösung der BM den fein abgestuften neuronalen Abstimmungskurven und verläuft nur abseits der Mittenfrequenz (charakteristischen Frequenz) etwas flacher (vgl. Boer & Nuttall 2010, 147f.; Hellbrück & Ellermeier 2004, 103; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37).

Sowohl das Kompressionsverhalten, als auch die Frequenzauflösung werden erst erklärbar, wenn man das aktive biomechanische Instrument des Innenohrs, den sogenannten *kochleären Verstärker*, bzw. die Aktivität der äußeren Haarzellen mit einbezieht (siehe 3.3.5, S. 135; vgl. Boer & Nuttall 2010, 158; Gelfand 2004, 166; Hellbrück & Ellermeier 2004, 103; Janssen in Lehnhardt & Laszig 2009, 113f.; Müller & Frings 2009, 460; Yost 2008, 113). Kompression wird durch Abschwächung starker und Anhebung schwacher Signale erreicht; die grobe mechanische Spektralzerlegung mittels der BM wird durch aktiven Eingriff in die Physik der Wanderwelle (letztlich auch Amplitudenmanipulation) verfeinert (vgl. Boer & Nuttall 2010, 151; Hellbrück & Ellermeier 2004, 103; Lehnhardt & Janssen in Lehnhardt & Laszig 2009, 37; Müller & Frings 2009, 460; Nilsen & Russell 1999, 646).

Bei geringen Pegeln kann das Verhalten gleichzeitig klingender Frequenzen, wie z.B. Teiltöne eines Breitbandsignals, als mehr oder weniger linear angesehen werden, d.h. alle Spektralkomponenten werden wie alleinstehende Töne behandelt. Steigen die Pegel aber über 30 dB, treten Nichtlinearitäten auf, die rein mechanisch nicht zu erklären sind. Die Amplituden anliegender Signalkomponenten beeinflussen sich gegenseitig (vgl. Boer & Nuttall 2010, 149ff.; Ranke 1953, 94f.). Neben der schon beschriebenen frequenzabhängiger Verdeckung (siehe S. 125; 1.6.2, S. 36) entstehen Verzerrungsprodukte, Kombinations-töne, u.a. durch Massekopplung. Dies geschieht, wenn zwei verschiedene Signalkomponenten dasselbe Masseteilchen bewegen (vgl. Ranke 1953, 96f.; Winckel 1960, 85; für eine weiterführende Darstellung der physikalischen Grundlagen vgl. Ranke 1953, 95ff.).

3.3.3 Mechanik des Cortiorgans

Nun muß man sich vergegenwärtigen, daß die Welle nicht in zweidimensionaler Weise allein auf der Basilmembran (BM) dahin rollt, sondern die gesamte Schneckentrennwand erfaßt. Sie entwickelt sich längs der BM, löst dabei aber effektiv quergerichtete Membranschwingungen aus, weil diese durch ihre Aufhängung (am Ligamentum spirale und der knöchernen Trennwand) seitlich gehemmt ist (siehe

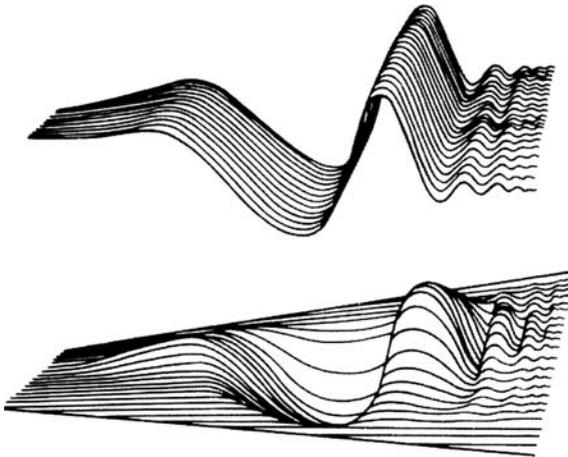


Abb. 3.17 **Schwingung der seitlich fixierten Basilmembran** (Yost 2008, Abb. 7.13, 96). Oben: Schwingung einer seitlich freien Membran. Unten: Schwingung einer seitlich fixierten Membran. Darstellung nach Tonndorf (1960).

ausgelenkt werden (siehe Abb. 3.18; vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 131ff.; 2009, 67; Nilsen & Russell 1999, 642; Schriever 1967, 33). Neuerdings nimmt man außerdem an, daß selbige Bewegungen die Endolymphe im subtektorialen Raum hin und her pumpen, wodurch die Stereozilien der IHZ ausgelenkt werden (vgl. Graumann & Sasse 2005, 119; Gummer et al. 1996, 8727; Lukashkin et al. 2010, 33; Ranke 1953, 106).

Tatsächlich ist die Darstellung mit dem klar definierten Ankerpunkt der BM zu einfach, da auch die knöcherne Trennwand mitschwingt. Im Querschnitt erscheint sie wie ein einseitig fixierter, am freien Ende durch die BM abgedichteter Stab (vgl. Békésy 1941(b), Abb. 12, 275, 276; Békésy &

Abb. 3.17), während sie mittig frei schwingen kann. Diese radiale Schwingung der BM ist entscheidend für die Erregung der Haarzellen (vgl. Gelfand 2004, 131ff.; 2009, 66f.; Hellbrück & Ellermeier 2004, 100f.; Keidel 1975, 11; Müller & Frings 2009, 454, 461; Reiss 2003, 622).

Man geht davon aus, daß Auslenkungen der an der knöchernen Trennwand aufgehängten BM zu Scherbewegungen mit der am Limbus verankerten, sich gleichsam bewegenden Tektorialmembran führen, wodurch die Stereozilien der ÄHZ

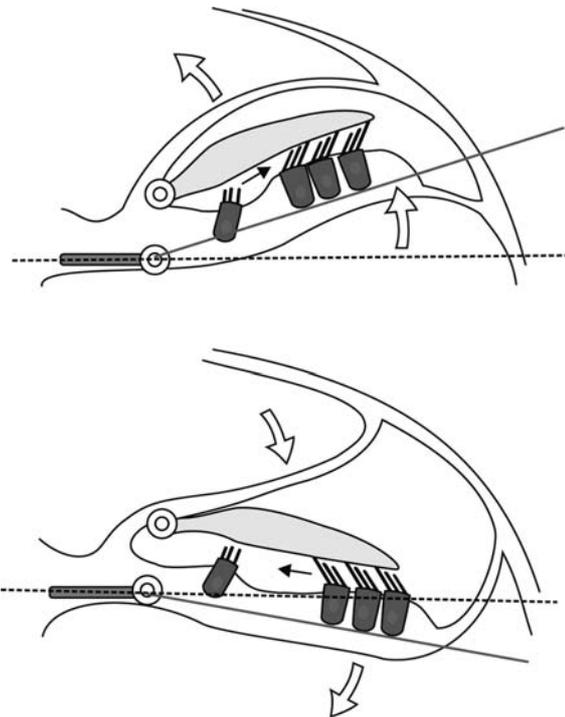


Abb. 3.18 **Scherbewegung zwischen Tektorialmembran und Basilmembran** (Müller & Frings 2009, Abb. 19.7, 456). Wegen unterschiedlichen Ankerpunkten verschieben sich TM und BM bei Auf- und Abwärtsbewegungen radial zueinander, wodurch die Stereozilien der Haarzellen (speziell der ÄHZ) ausgelenkt werden.

Rosenblith 1951, Abb. 2, 1076; Stenfelt et al. 2003, 132). Von medial nach lateral nimmt die Steifheit der Schneckentrennwand kontinuierlich ab, indem die knöcherne Struktur zunehmend membranös wird. Insgesamt schwingt sie radial bei oberflächlicher Betrachtung einheitlich, aber asymmetrisch, weil die Amplituden mit der Elastizität wachsen, so daß die maximalen Auslenkungen nahe der lateralen Aufhängung (also auf der BM) auftreten (vgl. Békésy 1941(b), Abb. 12, 275, 265f., 276; Békésy & Rosenblith 1951, Abb. 2, 1076; Ranke 1953, 105; Schriever 1967, 21; Stenfelt et al. 2003, 131, 137, 141ff.).

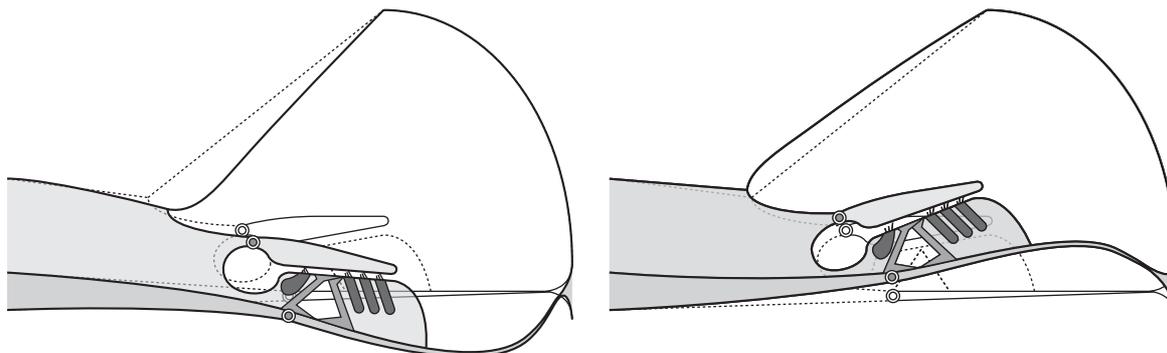


Abb. 3.19 **Radiale Schneckenschwingung**. Bezieht man alle Faktoren mit ein, müßte die Schneckenschwingung im Querschnitt wie dargestellt aussehen. Durch die Bewegung der knöchernen Trennwand schwingt die Reissnermembran medial, die Basilarmembran lateral am stärksten (vgl. Békésy 1974(b), Abb. 12, 275). Das gesamte Cortiorgan wird auf der schwingenden Schneckentrennwand bewegt. Da aber die Amplitude der BM größer als die der knöchernen Trennwand ist, kommt es trotzdem zur in Abb. 3.17 dargestellten Scherbewegung.

D.h. natürlich auch, daß die mediale BM-„Kante“ kein ruhender Ankerpunkt, sondern selbst bereits bewegt ist. Da allerdings der Limbus auch auf der knöchernen Trennwand liegt, gilt für die Aufhängung der Tektorialmembran grundsätzlich selbiges, so daß die oben vereinfacht vorgestellte Scherbewegung ziemlich sicher zutreffend ist.

Daß auch die Tektorialmembran entscheidenden Anteil am Hörprozeß hat, wird durch Tierversuche gestützt. Mäuse, denen durch Genmanipulation die Ausbildung einer funktionstüchtigen Tektorialmembran unterschlagen wurde, zeigten Hörverluste von bis zu 35 dB. Ähnliche Versuche mit genetisch manipulierten Tektorialmembranen zeigten Auffälligkeiten der allgemeinen Hörempfindlichkeit und speziell der Frequenzauflösung des Gehörs (vgl. Boer & Nuttall 2010, 164; Ghaffari et al. 2007, 16510).

Vermutlich ist die Tektorialmembran also an der Frequenzauflösung des Gehörs beteiligt. Obwohl direkte Beobachtungen in vivo noch nicht gelangen, konnten in situ und/oder indirekt Resonanzen der Tektorialmembran nachgewiesen werden (vgl. Boer & Nuttall 2010, 164; Gummer et al. 1996, 8727; Lukashkin et al. 2010, 26, 28–32). Sie zeigen einen transversalen und einen radialen (bis zu 30 dB stärkeren) Vektor, so daß die Resonanzbewegung im Querschnitt elliptisch erscheint, mit der Hauptachse parallel zur BM (vgl. Gummer et al. 1996, 8730f.).

Nach neuesten Auffassungen ist die Tektorialmembran ebenfalls tonotop organisiert (vgl. Müller & Frings 2009, 458) und Messungen in vitro deuten an, daß sogar Wanderwellen auf der Tektorialmembran entstehen, was ja schon aufgrund ihrer Struktur naheliegt (siehe 2.4.4, S. 81; vgl. Ghaffari et al. 2007, 16510; Lukashkin et al. 2010, 32). Wanderwellen auf der Tektorialmembran können sich aller-

dings *nur* durch mechanische Vorkopplung innerhalb der Membran ausbreiten, weshalb sie tiefer in die Schnecke laufen, als die Basilmembranresonanz vom Steigbügel entfernt liegt (vgl. Lukashkin et al. 2010, 32). Um dieses Problem zu lösen und die Tektorialmembran effektiv in die Schneckenmechanik einzubinden, ist ihre Resonanzfrequenz rund eine halbe Oktave tiefer als die BM-Resonanz. Nur so ist es möglich, daß beide (unabhängigen) Wanderwellen durch die ÄHZ effektiv gekoppelt werden (vgl. Fukazawa 2002, 55; Ghaffari et al. 2007, 16513; Gummer et al. 1996, 8729; Lukashkin et al. 2010, 26, 32; Nilsen & Russell 1999, 646). Ein weiteres Problem ist, daß die selbstgetriebene Wellenausbreitung auf der Tektorialmembran den Resonanzen notwendige Energie entzieht. Auf diese Weise, und mit dem analog zur Membranbreite wachsenden Energiebedarf der Welle, wird erklärt, weshalb apikal schwächere Resonanzen auftreten (vgl. Lukashkin et al. 2010, 32f.).

Die von Békésy dargestellte homogene, gleichphasige Bewegung der Schneckentrennwand spiegelt die Verhältnisse post mortem wieder (vgl. Nilsen & Russell 1999, 647). In vivo sind die radialen Schwingungsmuster, speziell der BM, komplex und inhomogen. In vitro konnten um bis zu 15 mal stärkere Bewegungsverhältnisse zwischen Tektorialmembran und Membrana reticularis als zwischen Tektorialmembran und BM beobachtet werden (vgl. Boer & Nuttall 2010, 163). Aktive Eingriffe der ÄHZ in die Schneckenschwingung sorgen für Torsionen des Cortiorgans und damit für die komplizierten radialen Schwingungsmuster in der lebenden Schnecke (vgl. Nilsen & Russell 1999, 646; 2000, 11751). Dabei werden die ÄHZ vermutlich durch die Resonanz der Tektorialmembran gesteuert (vgl. Boer & Nuttall 2010, 164; Gummer et al. 1996, 8732; Lukashkin et al. 2010, 26; Nilsen & Russell 1999, 646; 2000, 11755), wobei die kausalen Zusammenhänge hier noch nicht ganz klar sind (siehe 3.3.5, S. 140; vgl. Ghaffari et al. 2007, 16513; Lukashkin et al. 2010, 33). Das Schwingungsverhalten von BM und Cortiorgan in lebenden Ohren ist in praktisch allen Belangen massiv durch die Aktivität der ÄHZ beeinflusst. Bevor aber die aktiven Mechanismen des Innenohrs (kochleärer Verstärker) genauer beleuchtet werden, soll der Transduktionsvorgang im allgemeinen in den Fokus rücken.

3.3.4 Erregung der Haarzellen

Es gilt als gesichert, daß tatsächlich die Haarzellen im Innenohr als neuronale Reizwandler dienen (vgl. Gelfand 2004, 130; Ghaffari et al. 2007, 16510; Hackney & Furness 2010, 231; Pollard & Jansson 1982, 251; Ranke 1953, 103; Schriever 1967, 38; Stumpf 1890, 91; Winckel 1957, 84; Yost 2008, 108). Die IHZ sind die eigentlichen Sinneszellen, die das Signal dem afferenten Nervensystem zuleiten, während die ÄHZ die Schneckenschwingung aktiv manipulieren, wodurch sowohl die Frequenzselektivität, als auch die Empfindlichkeit des Gehörs verbessert wird (vgl. Hackney & Furness 2010, 231; Hellbrück & Ellermeier 2004, 101; Iwasa 2010, 179; Müller & Frings 2009, 456; Nielsen & Russell 1999, 642; 2000, 11751; Reiss 2003, 622f.; Yost 2008, 110ff.). Weiter oben wurde bereits die Scherbewegung zwischen BM und Tektorialmembran als mechanische Grundlage der Haarzellenstimulation ausgemacht (siehe 3.3.3, S. 129; vgl. Hackney & Furness 2010, 232; Yost 2008, 108). Die dafür notwendigen radialen Scherkräfte treten vor allem im basalen Wellenabschnitt zwischen Steigbügel und Wellenmaximum auf. Apikal sind die Amplituden so klein, daß es nur noch zu longitudinalen (Scher-)Bewegungen kommt (vgl. Gelfand 2004, 131; Schriever 1967, 33). Der Auslenkungsgrad der Stereozilien hängt von den

Amplituden der BM ab, während die „Wackelfrequenz“ der Stereozilien mit der Schwingungsfrequenz der Membran einher geht (vgl. Müller & Frings 2009, 454, 461).

Die längsten Stereozilien der ÄHZ sind fest in der Tektorialmembran verankert und somit „an ihr Schicksal gebunden“. Tektorialemembran und Membrana reticularis formen also gewissermaßen ein Hebelsystem, indem sie sich zueinander verschieben und die Stereozilien so entsprechend auslenken (vgl. Boer & Nuttall 2010, 161; Hellbrück & Ellermeier 2004, 100f.; Müller & Frings 2009, 454, 461; Reiss 2003, 622; Yost 2008, 108). Der Auslenkungsgrad der Stereozilien der ÄHZ korreliert also direkt mit der Auslenkung der „Scher“, so daß es wenig verwunderlich ist, daß Antworten der ÄHZ analog zur Amplitude der BM ausfallen (vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 133).

Die Auslenkung der Stereozilien der IHZ ist komplizierter als bei den ÄHZ und wird nach wie vor kontrovers gesehen (vgl. Boer & Nuttall 2010, 161). Die mittlerweile veraltete Auffassung, daß die Stereozilien im Zuge der Scherbewegung gegen die Tektorialmembran gestoßen werden, deswegen abknicken und die Haarzelle stimulieren, scheint zu einfach zu sein (vgl. Schriever 1967, 31). Der heute gängigen Theorie zufolge werden sie durch die viskose Kraft des aus der Scherbewegung entstehenden subtektorialen Lymphstroms erfaßt und ausgelenkt (vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 133f.; Lukashkin et al. 2010, 33; Müller & Frings 2009, 461f.; Ranke 1953, 105f.; Reiss 2003, 622; Schriever 1967, 34; Yost 2008, 108). Sie sind demnach auf die Bewegung der Endolymph angewiesen und antworten analog zur Frequenz bzw. Geschwindigkeit der BM (vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 133f.).

Man nimmt an, daß die Öffnungsmechanismen der Ionenkanäle an den Stereozilien durch die Tip-Links moderiert werden (vgl. Hackney & Furness 2010, 248; Yost 2008, 108f.). Diese sind federartig, so daß die Ionenkanäle auch ohne Auslenkung öffnen und schließen, bzw. im Ruhezustand halb offen stehen (vgl. Gelfand 2004, 135; Spitzer 2002, 61). Werden die Stereozilien ausgelenkt, werden die Tip-Links gebzw. entspannt, wodurch die Ionenkanäle sich länger bzw. weiter öffnen/schließen, was den Ioneneinfluß intensiviert bzw. komplett unterdrückt (vgl. Gelfand 2004, 135; Hackney & Furness 2010, 242; Yost 2008, 109).

Die nach lateral gerichtete Auslenkung der Stereozilien, hin zu ihrem längsten, dehnt die Tip-Links, öffnet so die Ionenkanäle und führt zu Depolarisation (Erregung, Exzitation) der Haarzelle (vgl. Gelfand 2004, 130, 134f.; Graumann & Sasse 2005, 119; Hellbrück & Ellermeier 2004, 101; Hackney & Furness 2010, 233; Müller & Frings 2009, 461; Nilsen & Russell 1999, 642; Reiss 2003, 622). Die geöffneten Ionenkanäle lassen positiv geladene Kaliumionen (K^+) aus der Endolymph ins Zellinnere strömen (vgl. Graumann & Sasse 2005, 119; Hackney & Furness 2010, 233; Hellbrück & Ellermeier 2004, 101; Müller & Frings 2009, 461; Reiss 2003, 622). Die durch Stereozilienauslenkung geöffneten Ionenkanäle schließen (bei konstantem Auslenkungsgrad) nach kurzer Zeit wieder, was den Ioneneinstrom abreißen läßt (vgl. Hackney & Furness 2010, 236). Der Transduktionsprozeß unterliegt also einer gewissen, vermutlich kalziumabhängigen Adaption (vgl. Boer & Nuttall 2010, 162; Hackney & Furness 2010, 236f., 248ff.). Weitere (d.h. noch intensivere) Auslenkung der Stereozilien öffnet die Ionenkanäle jedoch wieder (vgl. Hackney & Furness 2010, 236). Nach Ranke „bedeutet die Adaption keinen Verlust an Erregbarkeit, sondern stellt eine Anpassung des Erregungsniveaus an die jeweils länger vorherrschende mittlere Reizintensität dar“ (zit. nach Spreng in Keidel 1975, 117).

Werden die Stereozilien in Richtung ihrer Ausgangslage (bzw. nach medial) zurückgelenkt, entspannen sich die Tip-Links und die Ionenkanäle werden sukzessive geschlossen (vgl. Gelfand 2004, 130, 134f.; Hackney & Furness 2010, 233; Hellbrück & Ellermeier 2004, 101). Dabei werden die K^+ -Ionen lateral durch Kanäle in der Zellwand aus den Haarzellen „herausgepumpt“ (vgl. Graumann & Sasse 2005, 119; Hellbrück & Ellermeier 2004, 101; Reiss 2003, 622). Offene Kanäle an der Zellenbasis lassen parallel zum K^+ -Ausfluß an die kaliumarme Cortilymphe Chlorionen (Cl^-) einfließen, wodurch es zur Repolarisation der Haarzelle kommt. Dieser Prozeß dauert 1 bis 2 ms (vgl. Blauert 1974, 119; Graumann & Sasse 2005, 119; Hellbrück & Ellermeier 2004, 10). Die Repolarisationsphase ist in zwei Abschnitte zu unterteilen, die *absolute* und *relative Refraktärzeit* genannt werden. Während der absoluten Refraktärzeit (etwa 1 ms) ist die Haarzelle für jegliche Stimulation unempfindlich. Danach, während der relativen Refraktärzeit hängt es umso mehr von der Reizintensität ab, ob eine überschwellige Erregung der Haarzelle erreicht wird, weil die Erregungsschwelle der Haarzelle mit fortschreitender Repolarisation auf Normalniveau absinkt (vgl. Gelfand 2004, 123; Spreng in Keidel 1975, 116, 123, 131). Dadurch sind die möglichen Feuerraten der Nervenfasern auf maximal 1 kHz limitiert, obwohl die Ionenkanäle Reaktionszeiten von lediglich 10 μ s aufweisen (vgl. Gelfand 2004, 123; Hellbrück & Ellermeier 2004, 139; Spitzer 2002, 61; Spreng in Keidel 1975, 137). Dies ist eine wichtige Systemeinschränkung weil bekannt ist, daß die Impulssalven der Nervenfasern mit der Stimulusphase synchronisiert, sozusagen gekoppelt (*Phasenkopplung*), sind (vgl. Moore 1993, 89; Spitzer 2002, 70). In Gruppenarbeit können mehrere Fasern die kritische Frequenzgrenze nach oben ausdehnen, indem sie sich nicht nur mit dem Stimulus, sondern auch untereinander synchronisieren. Dabei kann jede Faser Perioden auslassen, die von anderen abgebildet werden. So wird die Gesamtperiode neuronal korrekt repräsentiert, während jede Faser ihre eigene Aktivitätsfrequenz sogar reduzieren kann. Diese Art der Phasenkopplung tritt nur bei Frequenzen bis etwa 5 kHz auf, darüber kommt es zu keiner – wie auch immer gearteten – Phasenkopplung mehr (vgl. Moore 1993, 89).

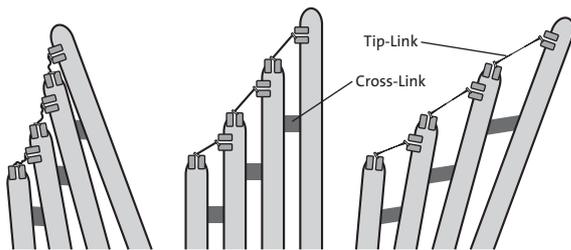


Abb. 3.20 **Auslenkung der Stereozilien und Öffnung der Ionenkanäle** (modifiziert n. Gelfand 2004, Abb. 4.11, 134; Müller & Frings 2009, Abb. 19.9, 459). Mitte: Ruhezustand. Die Ionenkanäle an den Enden der Tip-Links sind teilweise offen (ermöglicht Spontanaktivität). Links: Hyperpolarisation. Die Tip-Links entspannen, die Ionenkanäle schließen vollständig, kein Kaliumeinfluß ist möglich: die Haarzelle hyperpolarisiert (Hemmung). Rechts: Depolarisation. Die Tip-Links werden straff gespannt, was die Ionenkanäle vollständig öffnet. Kaliumeinfluß löst Einflüsse weiterer Stoffe aus, wodurch die Haarzelle depolarisiert wird (Aktivität).

Durch Auslenkung in die Gegenrichtung, also zum Modiolus nach medial, entspannen die Tip-Links völlig und die Ionenkanäle schließen vollständig, wodurch der Ioneneinfluß in die Haarzelle über einen längeren Zeitraum unterbunden und die Haarzelle so hyperpolarisiert (Unterdrückung, Inhibition) wird (vgl. Gelfand 2004, 134; Hackney & Furness 2010, 242).

Die (funktionelle) Kopplung der Stimulation mit den Tip-Links muß noch endgültig bewiesen werden (vgl. Hackney & Furness 2010, 245). Es ist nicht sicher, ob die Elastizität der Tip-Links ausreicht, um als Federsystem die Öffnungsmechanismen der Ionenkanäle zu betreiben. Es ist nicht einmal sicher, daß direkte Verbindungen zwischen den Ionenkanälen und den Tip-Links bestehen. Andererseits wird die Theorie u.a. dadurch

gestützt, daß die Tip-Links offenbar auslenkungsabhängig Steifheitsänderungen der Haarbündel hervorrufen können, und nicht zuletzt dadurch, daß in Entwicklungsstudien Tip-Links und Transduktionsfähigkeit der Haarzellen im Gleichschritt auftauchen (vgl. Hackney & Furness 2010, 235, 244). Trotz aller Unsicherheiten ist die Idee, daß die Öffnung der Ionenkanäle durch die Tip-Links gesteuert werden, die zur Zeit gängige und allgemein anerkannte Theorie (vgl. Hackney & Furness 2010, 248).

Reizwandlung

Die Auslenkung der Stereozilien kann gewissermaßen als die Schnittstelle zwischen dem mechanischen und dem neuronalen Sinnesapparat betrachtet werden, hier findet die mechanoelektrische Reizwandlung statt. Mechanische Aktivitäten lenken die Stereozilien aus, was die Stimulation der Haarzellen (speziell der IHZ) bewirkt, die wiederum das afferente auditive System aktiviert (vgl. Boer & Nuttall 2010, 161; Gelfand 2004, 136; Yost 2008, 108, 110ff.). Dieser Prozeß der Umwandlung von sinnesadäquaten Reizen in Nervenimpulse wird *Transduktion* genannt (vgl. Graumann & Sasse 2005, 119).

Von großer Bedeutung für die gesamte Transduktion ist, daß die apikalen Enden der Haarzellen und ihre Stereozilien von der positiv geladenen, kaliumreichen und natriumarmen Endolymphe umschwemmt sind, während die basolateralen Enden in negativ geladener, natriumreicher und kaliumarmer Peri- bzw. Cortilymphe baden (vgl. Graumann & Sasse 2005, 119; Iwasa 2010, 180).

Bei Reizbeginn nimmt das Ruhepotential ab. Diese Abnahme ist größer als das Wechsellpotential des Reizfolgestroms. Nach Signallende baut sich die Ruhespannung langsam wieder auf (vgl. Ranke 1953,

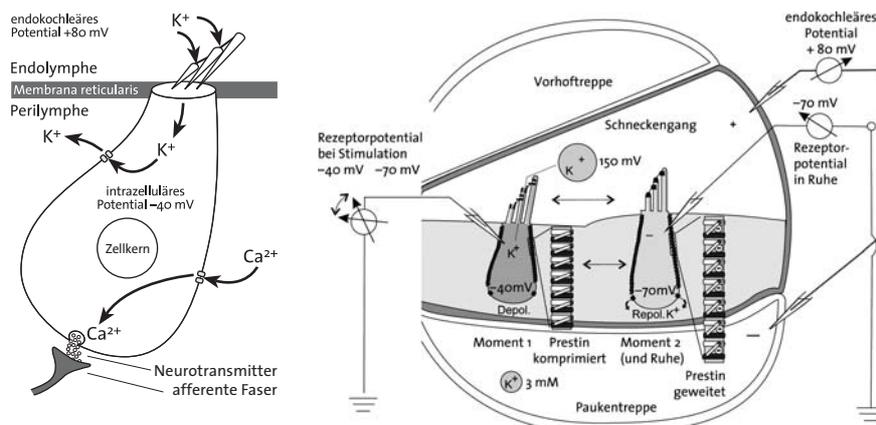


Abb. 3.21 **Depolarisation der Haarzellen**. Links: Depolarisation einer IHZ (Gelfand 2004, Abb. 4.12, 135). Durch die offenen Ionenkanäle an den Stereozilien strömt Kalium ein. Der Kaliumeinfluß öffnet die Ionenkanäle in der Zellwand, wodurch Kalzium einströmt, die Haarzelle depolarisiert. Wird die Depolarisation überschwellig (*Aktionspotential*), werden am synaptischen Spalt Botenstoffe freigesetzt. Rechts: Depolarisation einer ÄHZ (Müller & Frings 2009, 2009, Abb. 19.9, 459). Die Depolarisation funktioniert auf gleiche Weise wie bei der IHZ (nur Kaliumfluß abgebildet). Sie verändert das *intrazelluläre Potential* (*Rezeptorpotential*) und dadurch das *Membranpotential*, welches die *Konformation* der Prestinmoleküle in der Zellenwand steuert (siehe 3.3.5).

112). Von der Endolymphe führt ein K^+ -Ionenstrom durch die Haarzelle zur Perilymphe (Cortilymphe). Weil die Stereozilien in der Stimulusfrequenz hin und her wackeln und sich die Ionenkanäle entsprechend öffnen und schließen, entstehen oszillierende K^+ -Ströme, die zu frequenzabhängigen Wechselströmen in der Haarzelle führen (vgl. Müller & Frings 2009, 461). Diese Wechselströme werden *Reizfolgestrom* oder *kochleäres Mikrophonpotential*¹⁷ (cochlear microphonics) genannt (vgl. Hellbrück & Ellermeier 2004, 101; Plattig in Keidel 1975, 95; Ranke 1953, 110; Yost 2008, 103).

Der tatsächliche Erregungszeitpunkt liegt ungefähr mittig zwischen der von positiv zu negativ wechselnden Phase. Bis zum Auftreten des Aktionspotentials vergehen bei hohen bis mittleren Reizintensitäten in etwa 0,5 ms. Bei sinkender Intensität nimmt die Latenz leicht zu (vgl. Spreng in Keidel 1975, 130). Während dieser Zeit wird die Haarzelle durch den K^+ -Einfluß teilweise polarisiert, was basale Ionenkanäle öffnet, durch die positiv geladene Kalziumionen (Ca^{2+}) einströmen, wodurch erst das *Rezeptorpotential* (IHZ = Rezeptor) entsteht, die Ausschüttung der Neurotransmitter am synaptischen Spalt angeregt und ein Aktionspotential in der zugehörigen afferenten Nervenfasern ausgelöst wird (vgl. Gelfand 2004, 135; Graumann & Sasse 2005, 119; Hackney & Furness 2010, 234; Hellbrück & Ellermeier 2004, 101). Auf die Frage, wie das Schallsignal neuronal abgebildet wird, soll hier nicht eingegangen werden.

3.3.5 Kochleärer Verstärker

Békésy selbst erkannte bereits, daß die Spitzen der Wanderwellen bei weitem zu stumpf sind, um die feine Frequenzauflösung des Gehörs nur mit den mechanischen Vorgängen im Innenohr zu erklären (vgl. Békésy & Rosenblith 1951, 1107; Boer & Nuttall 2010, 157; Hellbrück & Ellermeier 2004, 102; Ranke 1953, 87, 103; Winckel 1957, 85). Aktive biologische Prozesse manipulieren die Schneckenschwingung.

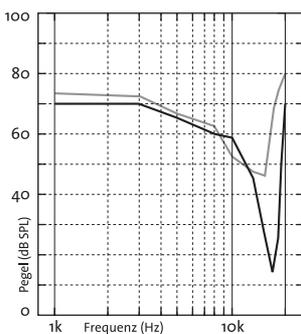


Abb. 3.22 Einfluß des kochleären Verstärkers (Gelfand 2009, Abb. 2.40, 68). Schwarz: die extreme Spitze verdeutlicht die feine Abstimmung einer gesunden Schnecke. Grau: die scharfe Spitze fehlt, bei Beeinträchtigungen der ÄHZ oder post mortem.

Dadurch wird die gute Frequenzselektivität des Gehörs schon im Innenohr erreicht, zudem die Empfindlichkeit und folglich die Dynamik verbessert. Die aktiven Vorgänge hinter diesen Verbesserungen sind unter dem Begriff *kochleärer Verstärker* (cochlear amplifier) bekannt (vgl. Gelfand 2004, 165; 2009, 68; Hellbrück & Ellermeier 2004, 103; Kemp 2010, 93; Müller & Frings 2009, 460; Yost 2008, 112).

Der kochleäre Verstärker ist allerdings keine besondere Funktion im „Stand-by-Betrieb“, die bei Bedarf zugeschaltet wird, sondern otophysiologische Normalität. Auch wenn in der Literatur davon die Rede ist, daß das System die Frequenzselektivität und Sensitivität des Gehörs *verbessert* (vgl. Boer & Nuttall 2010, 157, 168; Gelfand 2004, 165f.; Hellbrück & Ellermeier 2004, 108; Janssen in Lehnhardt & Laszig 2009, 113; Yost 2008, 111, 113), gilt dies nur auf passive Wanderwellen bezogen. Zu solchen kommt es aber nur in massiv geschädigten oder toten Ohren. Tatsächlich stellt der kochleäre Verstärker die sensitive Normalität erst her. Unter demselben Licht ist auch die Feststellung zu sehen, daß es im Innenohr zu Nichtlinearitäten kommt (vgl. Iwasa 2010, 179; Janssen in Lehnhardt & Laszig 2009, 114; Yost 2008, 110). Darunter werden

17 Der Begriff *cochlear microphonics* geht auf einen äußerst entwürdigenden Tierversuch zurück, in dem Wever und Bray (1930) feststellten, daß sie ihre Stimmen hören konnten, wenn sie den Hörnerv einer Katze über einen Verstärker an einen Lautsprecher schlossen und dann ins Ohr der Katze sprachen (vgl. Gelfand 2004, 137; Plattig in Keidel 1975, 93; Wever & Bray 1930, 344).

Abweichungen vom voraussichtlichen Verhalten der passiven Wanderwelle aufgrund des aktiven Eingriffs der ÄHZ verstanden. Ein Teil der Energie, die aus dem aktiven System an die Schnecke zurückfließt, „sickert“ bis zum Trommelfell durch und macht sich im Gehörgang als schwache Schallabsonderung bemerkbar (vgl. Boer & Nuttall 2010, 144; Gelfand 2004, 161, 165f.; 2009, 68; Janssen in Lehnhardt & Laszig 2009, 113; Kemp 2010, 93, 107).

Otoakustische Emissionen

Das Ohr ist tatsächlich auch ein Lautsprecher, der über das Trommelfell schwache Schallsignale an den Gehörgang abgeben kann (vgl. Boer & Nuttall 2010, 144; Gelfand 2004, 161; Janssen in Lehnhardt & Laszig 2009, 113; Kemp 2010, 107). Kurz nachdem ein Schallreiz ins Ohr geschickt wurde, kann im Gehörgang ein zweites Schallereignis, eine Art Echo registriert werden (vgl. Gelfand 2004, 161; Hellbrück & Ellermeier 2004, 107; Yost 2008, 112). Nach Kemp, dem als erstem der meßtechnische Nachweis gelang, wurde das Phänomen zunächst als *Kemp Echo* bezeichnet (vgl. Boer & Nuttall 2010, 144; Gelfand 2004, 161; Hellbrück & Ellermeier 2004, 107; Janssen in Lehnhardt & Laszig 2009, 113). Allerdings handelt es sich nicht um passive Echos, sondern um durch aktive Prozesse generierte Signale, die als direktes Nebenprodukt des aktiven Hörvorgangs auftreten und u.a. nach dem Tod nicht mehr nachzuweisen sind. Daher werden sie heute *otoakustische Emissionen* (OAE) genannt (vgl. Boer & Nuttall 2010, 144; Gelfand 2004, 161; 2009, 68; Hellbrück & Ellermeier 2004, 107; Kemp 2010, 93, 107; Müller & Frings 2009, 460; Yost 2008, 112).

OAE können Antworten des Innenohrs auf Schallreize (meistens Klicks) oder spontaner Natur sein. Zur Unterscheidung werden sie *evozierte otoakustische Emissionen* (EOAE, *klick-evozierte OAE*) bzw. *spontane otoakustische Emissionen* (SOAE) genannt (vgl. Boer & Nuttall 2010, 144, 166; Gelfand 2004, 161; Hellbrück & Ellermeier 2004, 107; Janssen in Lehnhardt & Laszig 2009, 113f.; Kemp 2010, 107; Müller & Frings 2009, 460; Yost 2008, 112; für eine Übersicht über die verschiedenen OAE-Typen vgl. Janssen in Lehnhardt & Laszig 2009, 114–117 (Kurzfassung); Kemp 2010, 107–113 (ausführlich); Yost 2008, 113).

Wird ein Klick ins Ohr geschickt, kann man nach einer Verzögerung von 5 bis 15 ms nach Abklingen des Klicks ein „Echo“ aus dem Innenohr aufnehmen (vgl. Boer & Nuttall 2010, 165f.; Gelfand 2004, 161; Hellbrück & Ellermeier 2004, 107; Yost 2008, 112). OAE haben extrem niedrige Pegel und liegen meist nahe oder unter der Hörschwelle. In ruhiger Umgebung können OAE vom verursachenden Ohr allerdings mitunter gehört werden (vgl. Kemp 2010, 109; Müller & Frings 2009, 460; Yost 2008, 112). Die

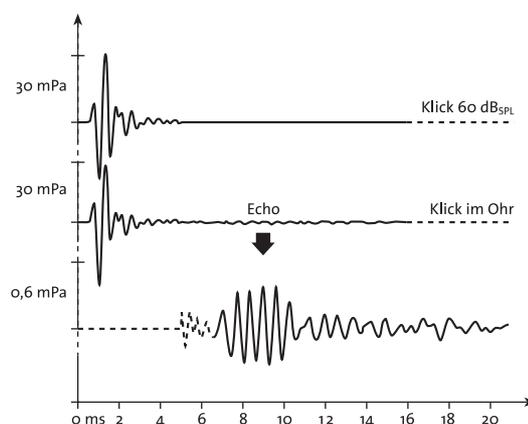


Abb. 3.23 **Evozierte otoakustische Emission** (Gelfand 2004, Abb. 4.34, 162). Oben: gedämpfte Schwingung des Senders (Zwislocki Kuppler). Mitte: Repräsentation der Schwingung im Ohr gefolgt vom Echo (EOAE). Unten: vergrößerte Darstellung des Echos der Ohrantwort. Darstellung nach Johnson und Eberling (1982).

Latenzzeit, nach der sie auftreten, ist stimulusabhängig. Höhere Reizpegel und/oder höhere Frequenz verkürzen die Latenzzeit. Letzteres ist gut nachvollziehbar, wenn man bedenkt, daß hohe Frequenzen näher zum ovalen Fenster auf der BM liegen und daher in beide Richtungen (in die Schnecke hinein und heraus) kürzere Laufzeiten erfordern (vgl. Gelfand 2004, 163). Dabei kann fast jeder beliebige Stimulus OAE erzeugen (vgl. Kemp 2010, 109). EOAE sind weder ein Abbild des erzeugenden Reizes, noch haben sie ein pauschales spektrales Grundmuster (vgl. Kemp 2010, 109). Sie reflektieren generell eher die Aktivität (aktiven Bereiche) der Schnecke, weniger das Signal selbst (vgl. Gelfand 2004, 163; Janssen in Lehnhardt & Laszig 2009, 114; Yost 2008, 113). Sie konzentrieren sich um die Mittenfrequenz des Reizes, wobei sie bei Sinustönen auch recht schmalbandig, also praktisch tonal ausfallen können. Bei Dauertonbeschallung gleichen sie sich der Reizfrequenz an (vgl. Gelfand 2004, 163, 165; Yost 2008, 112).

Spontane otoakustische Emissionen sind – das sollte nicht unerwähnt bleiben – keine Seltenheit, sondern eher Normalität. Bei Erwachsenen sind in jedem dritten Ohr SOAE feststellbar, bei Säuglingen sogar noch öfter (vgl. Kemp 2010, 107). Das Ohr ist in ständiger Bereitschaft und hält seine Auslöseschwelle durch andauernde Spontanaktivität der Haarzellen niedrig. Wahrscheinlich reflektieren SOAE diese Spontanaktivität des Innenohrs (vgl. Müller & Frings 2009, 460). SOAE bestehen oft aus einem oder mehreren Sinustönen, wobei ihre Spektren meist mehr oder weniger periodisch ausfallen und die einzelnen Teiltöne nie nahe zusammen liegen (vgl. Boer & Nuttall 2010, 166; Kemp 2010, 107).

OAE sind ein Ausdruck der nichtlinearen Verstärkung im Innenohr. Ihr Ursprung konnte bis zu den äußeren Haarzellen (ÄHZ) zurückverfolgt werden. Bei Schädigungen der ÄHZ und/oder Innenohrschwerhörigkeit sind OAE nicht mehr nachweisbar (vgl. Gelfand 2009, 68; Hellbrück & Ellermeier 2004, 107; Janssen in Lehnhardt & Laszig 2009, 113f.; Müller & Frings 2009, 460; Yost 2008, 112).

Elektromotilität der äußeren Haarzellen

Die äußeren Haarzellen sind zu speziellen aktiven Reaktionen fähig. Ihre Stereozilien nehmen wie Antennen die Schneckenschwingung auf und bewirken zur Schwingung synchronisierte Längenänderungen des Zellkörpers. Sie kontrahieren und expandieren in axialer Richtung, wenn sie stimuliert werden (vgl. Boer & Nuttall 2010, 162; Fukazawa 2002, 54; Gelfand 2004, 166; Hellbrück & Ellermeier 2004, 108; Müller & Frings 2009, 460f.; Yost 2008, 109f.). Die ÄHZ sind einzigartige, von chemischer Energie unabhängige elektromechanische Energiewandler, die fähig sind Wechselstrom auf piezoelektrische Weise in mechanische Bewegung transformieren (vgl. Iwasa 2010, 209; Müller & Frings 2009, 461).

Depolarisation (steigende Spannung) bewirkt Kontraktionen, Hyperpolarisation (sinkende Spannung) dagegen Elongationen. Die Kontraktionen fallen generell stärker aus als die Elongationen. An unbelasteten, isolierten ÄHZ wurden insgesamt Längenänderungen bis zu 5% der Gesamtlänge beobachtet. Gelfand gibt Änderungen bis zu 4% an, allerdings nicht, unter welchen Umständen sie zustande kamen (vgl. Fukazawa 2002, 54; Gelfand 2004, 166; Iwasa 2010, 183; Nilsen & Russell 1999, 642, 646; Yost 2008, 110). Die axiale Steifheit der ÄHZ ist – wie für elastische Materialien üblich – indirekt proportional zur Länge, d.h. die Steifheit pro Längeneinheit ist nicht von der Gesamtlänge abhängig (vgl. Boer & Nuttall 2010, 162; Iwasa 2010, 185).

Die Fähigkeit, ihre Länge synchron zu allen hörbaren Frequenzen zu ändern erfordert es, daß die ÄHZ mit Frequenzen bis zu 20 kHz oszillieren können (vgl. Yost 2008, 110). Schwingungen der BM sor-

gen für Stereozilienauslenkungen in beide Richtungen, wodurch die Ionenkanäle abwechselnd geöffnet und geschlossen werden, was zu Wechselströmen durch die Haarzellen führt (vgl. Müller & Frings 2009, 461). Wird eine ÄHZ durch externe Kraft gestreckt (z.B. durch die Tektorialmembran), kann ein Ionenfluß (durch die Haarzelle) beobachtet werden, der nachläßt, wenn die Streckung nachläßt. Der Stromfluß – quasi das *Membranpotential* (die Spannungsdifferenz innerhalb und außerhalb der Zelle) – ist also proportional zur Geschwindigkeit der Längenänderung (vgl. Fukazawa 2002, 54; Iwasa 2010, 186). Auf dieser Tatsache basierend konnte man bei isolierten, mit elektrischen Wechselfeldern stimulierten ÄHZ von Meerschweinchen Längsoszillationen bis zu 30 kHz (vgl. Hellbrück & Ellermeier 2004, 108) bzw. 80 kHz (vgl. Müller & Frings 2009, 461) erzeugen.

Unter elektrischem Gleichstrom, oder durch Kaliumzufuhr (K^+), konnten hingegen nur langsame Bewegungen des Zellkörpers ausgelöst werden. Dies nährt die Vermutung, daß man zwischen schnellen und langsamen Bewegungen unterscheiden muß (vgl. Hellbrück & Ellermeier 2004, 108; Iwasa 2010, 179; Yost 2008, 110). Nach Yost findet die schnelle Bewegung in einem Zeitfenster von 0,15 bis 0,2 ms nach Stimulation statt, während die langsame Bewegung erst wenige Sekunden danach einsetzt (vgl. Yost 2008, 110). Iwasa beschreibt stattdessen zwei verschiedene Rückkopplungsmechanismen (reverse transduction): einen im Stereozilienbündel, einen in der lateralen Zellenwand. Der schnellere (*fast adaption*), wirkt direkt auf die Verschlusszeiten der Ionenkanäle (*mechanoelectric transducer (MET) channels*). Diese reagieren in einem Zeitfenster von weniger als 1 ms, bzw. sogar 10 μ s (vgl. Iwasa 2010, 179; Spitzer 2002, 61). Der langsamere (*slow adaption*) hat ein Zeitfenster von rund 10 ms. Er wird indirekt durch den schnelleren Mechanismus gesteuert, weil sich die spannungsabhängige Motilität der lateralen Zellenwand am Rezeptorpotential orientiert, das mittels Ionenfluß generiert wird (vgl. Iwasa 2010, 179). Man geht davon aus, daß Rezeptorpotential bzw. Motilität konstant 90° phasenverzögert, und dadurch mit den *Bewegungen* der BM synchronisiert sind (vgl. Boer & Nuttall 2010, 164; Lukashkin et al. 2010, 27f.; Nilsen & Russell 1999, 646; 2000, 11755).

Es ist bekannt, daß ÄHZ in ihrer Kutikularplatte einige für Muskelkontraktion verantwortliche Proteine (Aktine) wie Myosin, etc. besitzen. Diese könnten für die langsamen Bewegungen verantwortlich sein (vgl. Iwasa 2010, 179; Hackney & Furness 2010, 238ff.; Hellbrück & Ellermeier 2004, 108; Müller & Frings 2009, 461; Yost 2008, 110). Möglicherweise sind langsame Bewegungen Vorbereitungsmechanismen, um die Empfindlichkeit des Gehörs auf die Hörsituation (leisen Schallen lauschen, laute Pegel hören) abzustimmen (vgl. Hellbrück & Ellermeier 2004, 113). Die durch Wechselstrom angeregten Bewegungen sind zu schnell, um neuronaler Natur zu sein (vgl. Yost 2008, 110). Auch die bekannten Motorproteine aus der Kutikularplatte scheiden nach entsprechenden Kontrollversuchen als Treiber für die schnellen synchronisierten Längsoszillationen der ÄHZ aus (vgl. Hellbrück & Ellermeier 2004, 108; Müller & Frings 2009, 461). Die Basis der Beweglichkeit liegt in Konformationsänderungen der Zellmembran (vgl. Gelfand 2004, 166). Entscheidender Treiber der schnellen Längsoszillationen ist das Protein *Prestin*, das ausschließlich in den Zellwänden der ÄHZ in größerer Zahl vorkommt und mit entsprechenden Oszillationen hoch empfindlich auf elektrische Wechselfelder reagiert (vgl. Boer & Nuttall 2010, 162; Gelfand 2004, 166; Hellbrück & Ellermeier 2004, 108; Iwasa 2010, 206; Müller & Frings 2009, 461; Yost 2008, 110). Prestinmoleküle sind zahlreich in den Zellwänden der ÄHZ gestapelt. De- und Hyperpolarisationen führen zu Konformationsänderungen der Moleküle, die sich in der Zellwand zu einer Gesamtbewegung aufsummieren, was letztlich zur Längenänderung der ÄHZ führt (vgl. Boer & Nuttall 2010, 162; Gelfand 2004, 166; Müller & Frings 2009, 461; Strenzke et al. 2008, 28).

Durch ihre Elektromotilität verändern die äußeren Haarzellen nicht nur ihre eigene Gestalt, sondern auch ihre unmittelbare Umgebung, weil Membrana reticularis und Tektorialmembran eben über die ÄHZ mechanisch aneinander gekoppelt sind. Regional eng begrenzte, aktive Eingriffe der ÄHZ in die Synchronisation der Schwingungskomponenten manipulieren die mechanischen Schwingungsmuster des Innenohrs, wodurch Empfindlichkeit und Frequenzauflösung des Gehörs enorm verbessert werden (siehe Abb. 3.22, S. 135; vgl. Müller & Frings 2009, 460; Nilsen & Russell 1999, 642; 2000, 11751; Yost 2008, 110f.).

Da die ÄHZ überwiegend efferent vernetzt sind, ist naheliegend, daß sie durch ein efferentes System gesteuert werden. Man nimmt an, daß efferente Kontrolle die Aktivität der ÄHZ unterdrückt (Inhibition) und so die generelle Empfindlichkeit des Gehörs dämpft. Allerdings sind efferente Einflüsse eher langsam, da der Befehlsverlauf aus der Schnecke in den Hirnstamm (vermutlich zum Olivenkernkomplex) und zurück führen würde (vgl. Hellbrück & Ellermeier 2004, 112; Iwasa 2010, 181, 206; Müller & Frings 2009, 461f.; Yost 2008, 110). Die efferente Synapse steuert durch Modifikationen des Kalziumbestands die Polarität der Haarzelle und kann diese durch Hyperpolarisation effektiv dämpfen. Vermutlich deshalb wird auch von kalziumbezogenem Verhalten des kochleären Verstärkers gesprochen (vgl. Boer & Nuttall 2010, 162; Hackney & Furness 2010, 236; Hellbrück & Ellermeier 2004, 109; Iwasa 2010, 206). Nach Hellbrück und Ellermeier greift efferente Dämpfung erst ab Signalpegeln von 40 bis 50 dB ein, um zu gewährleisten, daß es zwischen IHZ und Tektorialmembran zu keinem Kontakt kommt (vgl. Hellbrück & Ellermeier 2004, 112). De facto liegen die Aufgaben und Einflüsse der efferenten Kontrolle auf die ÄHZ aktuell noch weitestgehend im Dunkeln und gehören zu den am wenigsten verstandenen Komponenten des kochleären Verstärkers (vgl. Iwasa 2010, 205; Yost 2008, 110).

Funktion des kochleären Verstärkers

Die Längsoszillationen der ÄHZ sind der Mechanismus hinter dem kochleären Verstärker. Sie sind in der Lage, sich auf die Bewegungen der BM einzustellen und diese zu manipulieren (vgl. Gelfand 2004, 166; Hellbrück & Ellermeier 2004, 103, 108; Iwasa 2010, 179; Janssen in Lehnhardt & Laszig 2009, 113; Müller & Frings 2009, 460f.; Nilsen & Russell 1999, 642; Yost 2008, 113). Kontraktionen der ÄHZ ziehen Membrana reticularis und Tektorialmembran zueinander (vgl. Nilsen & Russell 1999, 642). Da die ÄHZ nicht senkrecht zur BM stehen, sondern zum Cortitunnel geneigt sind, besteht ihre Bewegung – analog zu den elliptischen Schwingungen der Tektorialmembran – neben der transversalen auch in einer radialen Komponente. Dadurch sind die ÄHZ optimal ausgerichtet, um die radialen Strömungen der subtektorialen Endolympe zu verstärken (vgl. Boer & Nuttall 2010, 163; Ghaffari et al. 2007, 16513).

Ob die ÄHZ effektiv auf die Schneckenschwingung einwirken, hängt entscheidend von der Phasenbeziehung zwischen passiver Basilarmembranschwingung und ÄHZ-Aktivität ab. Neben der Motilitätslatenz von 90° wird der effektive Einfluß der ÄHZ auf die Schneckenschwingung wahrscheinlich besonders durch die Resonanz der Tektorialmembran gesteuert (vgl. Boer & Nuttall 2010, 164; Gummer et al. 1996, 8732; Lukashkin et al. 2010, 26; Nilsen & Russell 1999, 642, 646; 2000, 11751, 11755). Einem Modell von Gummer et alii zufolge reagiert die Tektorialmembran im Bereich der charakteristischen Frequenz träge, wodurch sie sich der physikalischen Scherbewegung widersetzt, was eine Phasenumkehr der ÄHZ-Stimulation bedeutet und letztlich Verstärkungen der Schneckenschwingung

bewirkt. Apikal der Schneckenresonanz (CF) reagiert die Tektorialmembran nachgiebig, wodurch die Phasenumkehr aufgehoben wird und die ÄHZ bremsend auf die BM einwirken (vgl. Boer & Nuttall 2010, 165; Gummer et al. 1996, 8732; Lukashkin et al. 2010, 27).

Es ist zu betonen, daß Details der physikalischen und physiologischen Zusammenwirkung von Tektorialmembran und ÄHZ nach wie vor größtenteils unbekannt sind. Es ist ohne weiteres möglich, daß nicht die Tektorialmembran die ÄHZ beeinflusst, sondern umgekehrt (vgl. Ghaffari et al. 2007, 16513; Lukashkin et al. 2010, 33).

Die Bewegungen der ÄHZ sorgen für Torsionen des Cortiorgans (vgl. Nilsen & Russell 1999, 646; 2000, 11755). Membrana reticularis und BM werden durch Kontraktionen der ÄHZ zusammengezogen bzw. durch Elongationen auseinandergesprengt. Dadurch werden die inkompressiblen äußeren Pfeilerzellen, deren Sockel sich mit der BM bewegen, nach medial bzw. lateral gekippt, was wiederum die inneren Pfeilerzellen affiziert und im Endeffekt den Cortitunnel, vor allem aber die Membrana reticularis entsprechend um ihre Längsachse neigt (vgl. Nilsen & Russell 1999, 647; 2000, 11751, 11755). Die Relativbewegungen von Membrana reticularis und Tektorialmembran beeinflussen vermutlich auch den Fluß der Endolymphe im subtektorialen Raum. Bewegt sich die Basilar membran Richtung Endolymphraum, sorgt ihre lateral weitere Amplitude dafür, daß die Endolymphe unter die Tektorialmembran strömt, während die entgegengesetzte Bewegung Endolymphe aus dem inneren Sulcus herauspreßt (vgl. Boer & Nuttall 2010, 165; Graumann & Sasse 2005, 119; Lukashkin et al. 2010, 33; Nielsen & Russell 2000, 11755; Ranke 1953, 105f.).

Generell sind die Amplituden im Bereich des Cortiorgans (zwischen 30 und 120 µm von der medialen Kante entfernt) größer als im Übergangsbereich zur knöchernen Trennwand. Ein Umstand, der sich mit steigendem Pegel verschärft, jedoch auch durch stärkeres Kompressionsverhalten begleitet ist, als dies im weiten Membranabschnitt lateral des Cortiorgans der Fall ist. Wie Abbildung 3.25 verdeutlicht, weist die radiale BM-Schwingung generell zwei Maxima auf, wovon eines beim Cortitunnel, das andere grob basal der ÄHZ liegt. Dazwischen, „unter“ den äußeren Pfeilerzellen, liegt ein Bewegungsknoten, der bis zu Pegeln von 80 bis 90 dB_{SPL} konstant vorhanden ist (vgl. Nilsen & Russell 1999, 643). Die schwingungsarmen Bereiche nahe Trennwand, Ligamentum spirale und mittigem Knoten zeigen sehr ähnliche Phasenlagen zueinander, dürften zu den Maxima basal des Tunnels und der ÄHZ aber um grob 90 bis 135° verzögert sein (vgl. Nilsen & Russell 1999, 645f.).

In jedem Fall operiert der kochleäre Verstärker nur in einem sehr engen Aufnahmebereich um die charakteristische Frequenz. Die Spanne des Aufnahmebereichs wird unterschiedlich angegeben. Die Angaben reichen von einer Oktave (rechnerisch etwa 3,2 mm) bis etwa 1,2 mm um die charakteristische

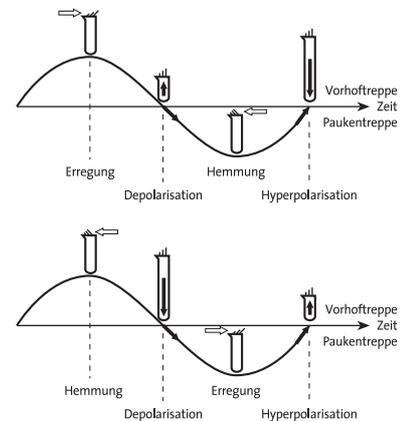


Abb. 3.24 **Einfluß der Tektorialmembranresonanz auf die Aktivität der ÄHZ** (Gummer et al. 1996, Abb. 4, 8731). Oben: Tektorialmembran reagiert nachgiebig, d.h. ÄHZ Schwingungen dämpfen BM. Unten: Im Bereich der CF reagiert die Tektorialmembran träge, was eine Phasenumkehr der ÄHZ bewirkt, die nun die BM-Schwingung verstärken.

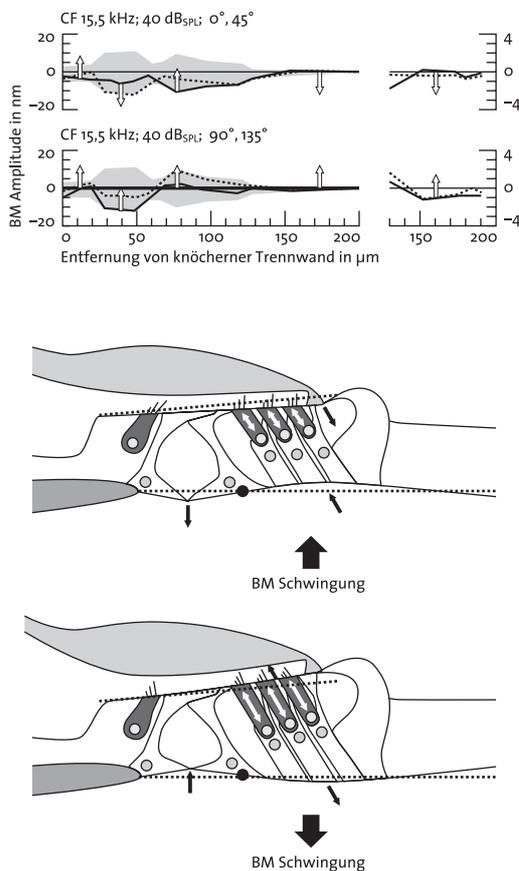


Abb. 3.25 **radiale Schwingungen der aktiven Schnecke** (Nielsen & Russell 2000, Abb. 5, 11755). Oben: Die radiale Schwingung der BM zu 0 und 90° (durchgezogene Kurven) sowie 45 und 135° (gepunktete Kurven). Die Graue Fläche stellt die Hüllkurve dar, die Pfeile deuten die Bewegungsrichtung an. Deutlich zu erkennen sind die beiden Maxima im Bereich des Cortitunnels (30–50 μm) und der ÄHZ (80–120 μm). Mitte und Unten: Torsionen des Cortiorgans durch die Motilität der ÄHZ. Der schwarze Punkt markiert den Angelpunkt und damit das in der Hüllkurve erkennbare Schwingungsminimum.

dieser Stelle könnten efferente Kontrollmechanismen eingreifen, um zu hohe Pegel abzuschwächen (vgl. Boer & Nuttall 2010, 158; Hellbrück & Ellermeier 2004, 112; Müller & Frings 2009, 461; Nilsen & Russell 1999, 642, 646).

Frequenz. Letztere Strecke würde gerundet in etwa 400 ÄHZ involvieren. Bei klarer Tonhöhe sind möglicherweise sogar nur ein bis drei Querreihen ÄHZ involviert (vgl. Ghaffari et al. 2007, 16513f.; Müller & Frings 2009, 460; Nielsen & Russell 1999, 642; 2000, 11757f.).

Durch den schmalen, scharf begrenzten Aktivierungsbereich wird die physikalisch stumpfe Wanderwelle sozusagen auf die charakteristische Frequenz „zugeschnitten“ (vgl. Janssen in Lehnhardt & Laszig 2009, 113; Müller & Frings 2009, 460; Yost 2008, 113). „Evans (1983) verglich diese [...] Funktion [mit] dem Scharfstellen eines Radiosenders auf den optimalen Empfang“ (vgl. Hellbrück & Ellermeier 2004, 109). Innerhalb des schmalen Bandes optimieren die oszillierenden ÄHZ mittels mechanischer Kopplung die Synchronisation von Membrana reticularis und Tektorialmembran (vgl. Hellbrück & Ellermeier 2004, 109; Yost 2008, 110f.). Mit ihrer Bewegung führen die ÄHZ Energie an die Schneckenschwingung zurück. Durch diese Energierückgewinnung wird die Empfindlichkeit des Innenohrs verbessert, weil die IHZ nun aus verstärkten Schwingungen „lesen“ können (vgl. Iwasa 2010, 179; Kemp 2010, 93; Yost 2008, 112). Die otoakustischen Emissionen sind ein Nebenprodukt der aktiven Energierückgewinnung des Innenohrs (vgl. Kemp 2010, 107). Durch den Einfluß der ÄHZ werden die mechanischen Schwingungen schwacher Signale bis zu 100-fach (40 dB) verstärkt (vgl. Hellbrück & Ellermeier 2004, 108f.; Müller & Frings 2009, 460f.; Nilsen & Russell 1999, 642). Bei geringeren Pegeln funktioniert die Übertragung (oder Transduktion) noch verzerrungsfrei. Höhere Pegel, ab etwa 40 dB_{SPL} werden mit reduzierter Amplitude ins neuronale System gespeist, also praktisch komprimiert. An

Fazit

4 Zusammenfassung

Schallpegel und Frequenz müssen innerhalb bestimmter, als *Hörbereich* bezeichneter Grenzen liegen, damit Schall nicht „durch den Raster“ fällt. Des Weiteren ist für adäquate Schallverarbeitung eine gewisse Mindestdauer vonnöten. Sind diese grundlegenden Erkennungskriterien erfüllt, wird das *Schallereignis* vom Gehör aufgenommen und in ein entsprechendes Perzept, das *Hörereignis* übersetzt.

Wird das Ohr „angeregt“, bilden sich zunächst die *Frequenzgruppen* (ca. 10 ms). Erst danach kann das Spektrum analysiert werden. Der Charakter des Schalls kann trotzdem schon vorher durchdringen (vgl. Winckel 1952, 52). Er wird aus allen Schallanteilen unterhalb der *Verwischungsschwelle* (ca. 50 ms) generiert. Das Ohr gewinnt weiterhin Information aus der zunehmenden Signaldauer, bis die *Klangverschmierungsschwelle* (ca. 250 ms) erreicht ist. Dann ist das Signal optimal ausgewertet und es setzt ein Gewöhnungseffekt (Adaption) ein.

Körper, Kopf und *Ohrmuschel* filtern Schall abhängig von ihrer relativen Position zur Schallquelle. Beide Ohren empfangen meistens leicht unterschiedlich gefilterte Signale. Aus der spektralen Filterkurve und den Unterschieden beider Ohrsignale errechnet das Gehör den Einfallswinkel des Schallsignals. Der mit dem *Trommelfell* einseitig verschlossene *Gehörgang* wirkt wie ein gedacktes Resonanzrohr und verstärkt den Schall zwischen ungefähr 2 und 5 kHz, was die auffällige Senke der *Ruhehörschwelle* in diesem Bereich erklärt.

Das große und leicht bewegliche Trommelfell nimmt die Schwingungen der Luft auf. Über die Gehörknöchelchenkette des Mittelohrs werden die Schwingungen komprimiert und verstärkt. Auf diese Weise wird der Impedanzunterschied zwischen Luft und Innenohrflüssigkeiten überwunden. Zwischen 1 und 2 kHz überträgt das Mittelohr am besten, was neben der Gehörgangsresonanz ebenfalls im Verlauf der Hörschwelle reflektiert ist.

Der *Steigbügel* am Ende der Knöchelchenkette versetzt die *Perilymphe* des Innenohrs in Bewegung, wodurch sich auf der *Schneckenwand* (für eine Begriffserklärung siehe 3.3.1, S. 120; 2.4.3, S. 78) Wanderwellen Richtung Schneckenspitze ausbilden. Die besondere Struktur der *Basilarmembran* entzieht den Wellen erst ihre höchsten und im weiteren Verlauf immer tiefere Spektralanteile, bis die Welle mit dem tiefsten Teilton „aufgebraucht“ ist und verebbt. Jeder Teilton resoniert an einer für seine Frequenz spezifischen Stelle der Basilarmembran (*Tonotopie*).

Im *Cortiorgan* auf der Basilarmembran befinden sich die Sinneszellen, die *inneren* und *äußeren Haarzellen* (IHZ, ÄHZ). Durch die Membranschwingungen wird die feine Mikromechanik des Cortiorgans angestoßen (siehe 3.3.3, S. 129), was wiederum die Haarzellen erregt und den Transduktionsprozeß in Gang setzt. Die IHZ, die eigentlichen Transduktoren, werden durch die ÄHZ unterstützt, deren Aufgabe es ist, aktiv in die Schneckenphysik einzugreifen und die Membranschwingung zu optimieren. Wahrscheinlich, aber keinesfalls sicher ist, daß die ÄHZ durch die Mikromechanik des Cortiorgans zur Verstärkung angeregt und durch efferente Kontrolle aus dem Hirnstamm zu Dämpfungszwecken gehemmt werden.

4.1 Auswirkungen auf die musikalische Praxis

Das Ohr ist nicht einfach ein Mikrofon, das über ein Kabel (Hörnerv) an einen Prozessor (Gehirn) angeschlossen ist. Es ist ein kompliziertes und raffiniertes Organ. Seine Funktionsweise greift massiv in jenes akustische Weltbild ein, das es vermitteln soll. Daraus entstehen mehr oder weniger schwerwiegende Konsequenzen für Musikkonsum, (gemeinsames) Musizieren und (nicht zuletzt) die Musikwissenschaft.

Die Zeitkonstanten des Gehörs (siehe 1.2.1, S. 12) machen Hören überhaupt möglich. Würde das Ohr nicht verschiedene Toleranzschwellen setzen, unterhalb derer Schallanteile zusammengefaßt werden, dränge das tatsächliche, schier unendliche akustische Bombardement ungefiltert und ungeordnet ins Bewußtsein, wo die Informationsüberladung erst aussortiert werden müßte. Die zeitliche Strukturierung begegnet diesem Umstand mit einigen praxisrelevanten Vorkehrungen. Mit der *Verwischungsschwelle* (siehe 1.2.1, S. 14) definiert es die im musikalischen Zusammenspiel bedeutsame Toleranzgrenze für rhythmische Ungenauigkeiten. Unter die Verwischungsschwelle fallende Einsatzunterschiede werden noch als synchron gehört (vgl. Reuter 1996, 16).

In der Beschallungstechnik macht man sich das *Gesetz der ersten Wellenfront* (siehe 1.2.1, S. 15; 1.7.3, S. 45) zunutze, um „den Schalldruckpegel vor allem im hinteren Zuhörerbereich durch Lautsprecher, die mit entsprechender Verzögerung betrieben werden, zu erhöhen, ohne daß diese geortet werden“ (Fasold & Veres 2003, 56).

Ein der *Verwischungsschwelle* ähnliches Zeitfenster hat die Bildung des räumlichen Klangeindrucks. Schallanteile unterhalb der *Echoschwelle* (siehe 1.2.1, S. 15; 1.7.3, S. 45) werden zur Konstruktion des Nachhalls verwendet (vgl. Fricke & Louven 2008, 433; Culling & Akeroyd 2010, 135). Spätere Schallanteile werden als eigenständig aufgefaßt und bilden Echoeindrücke. In der Audiotechnik hat man das Prinzip umgekehrt und durch exakt bemessene Verzögerungen von Schallanteilen dem Ohr Räumlichkeitsverhältnisse vorgegaukelt. Heute greift man dafür auf Impulsantworten von Räumen zurück (vgl. Weinzierl 2008, 16), das Prinzip ist im Grunde aber dasselbe.

An dieser Stelle soll auch noch einmal daran erinnert werden, daß die Einschwingvorgänge der meisten (klammert man die Klangerzeugungsmöglichkeiten der digitalen Welt aus, eigentlich aller) Musikinstrumente ziemlich gut dem Zeitraster des Ohrs entsprechen (vgl. Fricke & Louven 2008, 433; Winckel 1952, 97). Sonderlich verwundern kann dieser Umstand nicht. In Film und Fernsehen haben sich Bildraten entwickelt, die vor den Zeitkonstanten des Auges bestehen müssen. Entsprechendes gilt für Musikinstrumente, deren Referenz nur das Ohr sein kann.

Die frequenzabhängigen Hörschwellen haben einen nicht zu unterschätzenden Einfluß. Die Lautstärke-regler der meisten Wiedergabegeräte haben relativ lineare Frequenzgänge. D.h. alle Frequenzen werden durch Drehen des Knopfes gleichermaßen beeinflusst. Aufgrund der U-förmigen *Ruhehörschwelle* (siehe 1.1.1, S. 6) wird Musik dumpfer (tiefenlastiger), wenn man sie lauter stellt und heller (brillanter), wenn man sie leiser stellt (vgl. Gelfand 2004, 349; 2009, 93; Hellbrück & Ellermeier 2008, 55). Zudem kann es wegen der frequenz- und pegelabhängigen *Mithörschwellen* (siehe 1.6.1, S. 33) vorkommen, daß tiefe Töne frequenznahe hohe Töne plötzlich verdecken, wenn die Wiedergabelautstärke inadäquat hoch ist. Die frequenzabhängige Empfindlichkeit des Ohrs beeinflusst die auditive Wahrnehmung also nicht nur quantitativ, sondern auch qualitativ (vgl. Zwicker & Feldtkeller 1967, 66).

„Unser Hörsystem ist zum Hören vieler Töne gar nicht in der Lage, wenn zugleich andere, lautere Töne klingen“ (Spitzer 2002, 68). „Verdeckung ist ein rigoroses Instrument der Informationsreduktion, und beim Schall regiert das Gesetz des Stärkeren“ (Fricke & Louven 2008, 417). Zwei Sätze, die das Wesen der *Verdeckung* (siehe 1.6, S. 33) sehr schön beschreiben und einfangen, warum Verdeckung auch für Übertragung und Archivierung von (akustischen) Daten so interessant ist. Des Verdeckungsprinzips bedient man sich u.a. bei der Komprimierung im *MPEG-1 Layer 3* Format (besser bekannt als *mp3*): Alle für das Ohr versteckten Schallanteile werden verworfen, ohne daß sich das Klangbild ernsthaft verändert. Dadurch kann die Datenmenge erheblich reduziert werden (vgl. Fricke & Louven 2008, 417; Spitzer 2002, 68; für Näheres zum Komprimierungsalgorithmus vgl. Werwein & Schick 2008, 870ff.).

Verdeckung ist aber nicht nur Freund, sondern auch Feind. Angesichts der rasanten Entwicklung von mobilen Wiedergabegeräten, haben sich auch die Abhörbedingungen aus dem wohlarrangierten Klangraum der klassischen Stereoanlage zunehmend auf ungünstigere, störräuschbelastete Hörsituationen verlagert (vgl. Maempel et al. 2008, 781):

Diesen Umständen wird oft durch eine stärkere Monokompatibilität, vereinfachte Lokalisationsverteilung, geringere Dynamik bzw. höhere Lautheit, klangfarbliche Aggressivität und Vermeidung der Nutzung der Außenoktaven für musikalisch wesentliche Information begegnet“ (Maempel et al. 2008, 781).

Beim Arrangieren und in der Aufführungspraxis trägt man auch dem Phänomen der *asymmetrischen Verdeckung* (siehe 1.6.2, S. 36) Rechnung, indem man Instrumente bzw. Stimmen, die im Frequenzraum höher angesiedelt sind (z.B. Frauen- gegenüber Männerstimmen, Geigen gegenüber Celli, etc.), zahlenmäßig stärker besetzt (vgl. Hellbrück & Ellermeier 2004, 129; Zwicker & Feldtkeller 1967, 66).

Verdeckung ist in der Musik allgegenwärtig. Spielen mehrere Instrumente zusammen, leiden ihre Klänge unter gegenseitiger *partieller Verdeckung*, aus der oft nur ihre, in unterschiedliche Frequenzabschnitte fallenden, Formanten hervorstehen (vgl. Reuter 1996, 79). Vollständig hörbare Klänge sind Ausnahmefälle. In aller Regel stehen die physikalischen Grundtöne im akustischen Abseits. Hier kommen die *Residualtöne* (siehe 1.5.3, S. 32) ins Spiel. Residualtöne machen mehrstimmiges Musizieren überhaupt erst möglich, weil die hörbaren Formanttöne dem Ohr genügen, um daraus die Periodizitätstonhöhe zu erschließen.

Die *Frequenzgruppen* des Gehörs (siehe 1.3, S. 18; 3.3.2, S. 125) sind fundamental in die Rauigkeitsempfindung von Tonintervallen eingebunden. Gleiches gilt freilich auch für Spektraltöne von Klängen, woraus oft Erklärungen für das Konsonanz-/Dissonanzverständnis westlicher Tonsysteme konstruiert werden (vgl. Spitzer 2002, 98ff.). Oft bleibt dabei unbeachtet, daß das „westliche System“ kein einheitliches Musikverständnis impliziert. Spitzer bezeichnet die große Septime aufgrund ihrer vielen „ungünstig“ eng zusammenliegenden Teiltöne als dissonant (vgl. Spitzer 2002, 98). Die meisten Jazzmusiker werden dem kaum zustimmen. Außerdem sind Intervalle, je nach ihrer Frequenzlage, sehr unterschiedlich zu bewerten. Unter 500 Hz (zur besseren Einschätzung: h' liegt bei rund 495 Hz) werden die integrierenden Filter (Frequenzgruppen) verhältnismäßig immer breiter (siehe 1.3.1, S. 19; 3.3.2, S. 125; vgl. Spitzer 2002, 100). Unter 500 Hz ist auch die als harmonisch konsonant angesehene

Terz ein akustisch/auditiv rauhes Intervall. Hier greift die Frequenzgruppe in Theorie und Praxis des Komponierens ein: Während es für hohe Klänge keine solche Einschränkung gibt, sollten tiefe Klänge stets in weiter Lage geführt werden (vgl. Reuter 1996, 72; Spitzer 2002, 100).

An dieser Stelle möchte ich kurz auf ein Streamingphänomen eingehen, das ansonsten nicht Teil dieser Arbeit ist. Tougas und Bregman stellten mittels Hörversuchen fest, daß es dem Ohr unmöglich ist, sich kreuzenden Tonleitern über den Kreuzungspunkt hinweg zu folgen, was als psychologischer Beleg für das musikalische Verbot der Stimmkreuzung angesehen wurde (vgl. Reuter 1996, 47f.; Tougas & Bregman 1985, 788–792). Dagegen spricht, daß intuitives, enkulturiertes Wissen Erwartungshaltungen zu Ungunsten der Stimmkreuzung wecken könnte (vgl. Reuter 1996, 48). Andererseits wurden akute Erwartungshaltungen (auch und besonders zugunsten der Stimmkreuzung) erzeugt, indem jedem Versuch Vergleichsstimuli vorausgingen, die vorspielten, welches Muster herausgehört werden sollte (vgl. Tougas & Bregman 1985, 790f.). Sobald beide Tonleitern in derselben Frequenzgruppe angekommen waren, wurden sie nicht mehr als getrennte Ereignisse interpretiert. Unterhielten sie wieder eigene Frequenzgruppen, wurden sie auch perceptiv getrennt. Nachdem das *Prinzip der Nähe* (vgl. Bregman 1990, 196f.; Reuter 1996, 38; Tougas & Bregman 1985, 788) ein hochrangiges Ordnungsmuster des Gehörs zu sein scheint, ist es nicht unplausibel, daß eher zwei „Bögen“ als ein „Kreuz“ gehört werden (vgl. Reuter 1996, 48; Tougas & Bregman 1985, 792f.).

Richtet man den Blick – oder besser gesagt: das Ohr – auf die Tonhöhenempfindung, findet man, daß die Psychoakustik Phänomene zutage gefördert hat, die zwar auffallen, wenn man danach sucht, in der musikalischen Praxis allerdings weniger bedeutsam sind. Die empfundene Tonhöhe ist ab ungefähr 500 Hz immer tiefer, als sie gemäß des auslösenden Reizes, der Frequenz, sein sollte (siehe 1.5.2, S. 28). Transponiert man eine Melodie aus tieferen Oktaven zu höheren, wird sie gefühlt enger (Zwicker & Feldtkeller 1967, 93). Frequenzverdoppelungen, die in tiefen Frequenzregionen die erwarteten Oktaven bilden, werden in hohen Frequenzbereichen mitunter nur noch als Quinte empfunden (vgl. Fricke & Louven 2008, 423).

In der Praxis sind diese, immer etwas komisch anmutenden Behauptungen weniger auffällig, weil: Erstens normalerweise obertonreiche Klänge produziert werden, die harmonische Beziehungen für das Gehör sehr stark stabilisieren (vgl. Zwicker & Feldtkeller 1967, 80). Im Umkehrschluß heißt das: Mit Sinustönen stellt man die oben beschriebenen Verstimmungen leichter fest, was eine Erklärung dafür sein könnte, daß sie vor den sinustonverliebten Psychoakustikern nicht lange verborgen blieben. Zweitens hört das Ohr aus tonalen Strukturen sehr schnell die ihm vertrauten harmonischen Verhältnisse heraus (vgl. Zwicker & Feldtkeller 1967, 79). Dies muß ausdrücklich auf jedes erdenkliche Tonsystem der Welt bezogen werden. Drittens findet der Bruch im Frequenzskalensystem erst bei 500 Hz statt. Grundtöne oberhalb von 1 kHz werden in der Musik allerdings kaum bis gar nicht verwendet und oberhalb von 5 kHz von kaum einem Klangerzeuger überhaupt produziert (vgl. Hellbrück & Ellermeier 2004, 87; Moore 1993, 89; Zwicker & Feldtkeller 1967, 93f.). Auch hier liegt der Schluß nahe, daß sich die historisch gewachsenen Frequenzlagen der Tonumfänge von Musikinstrumenten an den Einschränkungen des Gehörs orientiert haben.

Die Tonhöhenwahrnehmung wird zu höheren Frequenzen nicht nur enger, sondern auch schlechter. Beides ist vermutlich darin begründet, daß die Fähigkeit zur Phasenkopplung einzelner Nervenfasern auf

1 kHz und die Synchronisationsfähigkeit ganzer Faserbündel auf 5 kHz beschränkt ist (siehe 1.1.3, S. 11; 3.3-4, S. 133). Klare Tonhöhenindrücke sind werden demnach ab 1 kHz schwächer und verschwinden um 5 kHz mehr oder weniger vollständig. Mit anderen Worten: über 5 kHz ist die Tonhöhenwahrnehmung massiv beeinträchtigt und praktisch unbrauchbar. Mit Sicherheit ist diess physiologische Einschränkung der Hauptgrund dafür, warum sich das tonale Spektrum der meisten Musiksysteme quasi ausschließlich auf den Frequenzbereich unter 1 kHz konzentriert.

Wie beschrieben, kann auch die Lautstärke Einfluß auf die Tonhöhe nehmen (siehe 1.5.2, S. 31). Man kann das Phänomen in ruhiger Umgebung recht gut an lange klingenden Tönen (z.B. schwingende Saiten) beobachten, da sie sich aufgrund der Pegelabhängigkeit beim Ausklingen leicht verstimmen. Beim Musizieren fallen derartige Verstimmungen kaum auf, weil sie sich weniger im vertikalen Aufbau, sondern eher in der Folge von Klängen zeigen (vgl. Winckel 1960, 84f.). Allerdings kann man beobachten, „daß eine Dreiklangfolge in tiefer Lage im Crescendo oder Decrescendo, die anschließend durch einen Dreiklang vertikal bestätigt wird, verstimmt erscheint“ (Winckel 1960, 84). Auch dieses Phänomen wird durch obertonreiche Spektren untergraben (vgl. Stevens & Hallowell 1938, 75; Winckel 1960, 84).

4.2 Schlußbetrachtung

Ich mache es kurz: Das Gehör beeinflusst das Hören. Der Empfänger beeinflusst sozusagen die Wahrnehmung. Dies ist natürlich nichts spezielles, sondern vollkommen normal. Viele psychoakustische Phänomene wurden jedoch isoliert, in idealisierten Hörsituationen mit angepaßten Stimuli ermittelt. Sie verdeutlichen die enorme Leistungsfähigkeit des Gehörs und zeigen, woran sich Entwickler von z.B. Wiedergabegeräten (hier wird mitunter sogar übertrieben: vgl. Hellbrück & Ellermeier 2008, 55), Hörgeräten oder Implantaten orientieren sollten.

In natürlichen Hörsituationen gelten aber oft andere Grenzwerte, weil Schalle in anderer, nicht idealisierter, natürlicher Hörumgebung dem Ohr mehr (Spektren) bzw. weniger (Verdeckung) Information bieten, der Organismus aber auch anders auf seine akustische Umwelt reagieren kann (z.B. Peilbewegungen, optische Unterstützung) und diese auch besser kennt, als die sterile und künstliche Hörsituation im Labor. Doch vor allem in natürlichen Hörsituationen gilt, daß die Verarbeitungsmechanismen des Ohrs die uns bekannte auditive Szenerie erst zu dem machen, was sie ist. Die Manipulation setzt sich im auditiven Nervensystem fort, wo die Höreindrücke erst Gestalt annehmen und des weiteren vor allem inhaltlich ausgewertet werden (vgl. Koelsch 2005, 365; Koelsch & Schröger 2008, 393). Alles in allem ist der Hörvorgang ein aktiver und manipulativer Prozeß, so daß es sicher keine allzu kühne Behauptung ist zu sagen: „Das Ohr macht die Musik“.

Anhang

5 Verwendete Literatur

- Aibara, Ryuichi et alii 2001. *Human middle-ear sound transfer function and cochlear input impedance*. In: *Hearing Research*. Ausg. 152, S. 100–109
- Békésy, Georg von 1941(a). *Über die Messung der Schwingungsamplitude der Gehörknöchelchen mittels einer kapazitiven Sonde*. In: *Akustische Zeitschrift*. Bd. 6(1), S. 1–16
- Békésy, Georg von 1941(b). *Über die Elastizität der Schneckentrennwand des Ohres*. In: *Akustische Zeitschrift*. Bd. 6(5), S. 265–278
- Békésy, Georg von 1943. *Über die Resonanzkurve und die Abklingzeit der verschiedenen Stellen der Schneckentrennwand*. In: *Akustische Zeitschrift*. Bd. 8, S. 66–76
- Békésy, Georg von 1960. *Experiments in Hearing*. New York, McGraw-Hill
- Békésy, Georg von; Rosenblith, Walter A. 1951. *The Mechanical Properties of the Ear*. In: Stevens, Stanley S. *Handbook of experimental Physiology*. New York, Wiley & Sons, S. 1075–1115
- Blauert, Jens 1970. *Ein neuartiges Präsenzfilter*. In: *Fernseh- und Kinotechnik*. Ausg. 3, S. 75–78
- Blauert, Jens. 1974. *Räumliches Hören*. Stuttgart, Hirzel
- Blauert, Jens. 1997. *Spatial Hearing. The Psychophysics of Human Sound Localization*. Cambridge, MIT Press
- Boer, Egbert de; Nuttall, Alfred L. 2010. *Cochlea mechanics, tuning, non-linearities*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 139–177
- Bürkner, Kurd 1900. *Atlas von Beleuchtungsbildern des Trommelfells*. Jena, Fischer
- Cheng, Jeffrey Tao et alii 2010. *motion of the surface of the human tympanic membrane measured with stroboscopic holography*. In: *Hearing Research*. Ausg. 263, S. 66–77
- Culling, John F.; Akeroyd, Michael A. 2010. *Spatial Hearing*. in: Plack, Christopher J. (Hrsg.). *Hearing. The Oxford Handbook of Auditory Science*. Bd. 3, New York, Oxford Univ. Press, S. 123–144
- Dickreiter, Michael 1987. *Handbuch der Tonstudioteknik*. Bd. 1, München, Saur
- Didyk, Lyudmyla A. et alii 2007. *The mechanical reaction of the pars flaccida of the eardrum to rapid air pressure oscillations modeling different levels of atmospheric disturbances*. In: *Hearing Research*. Ausg. 223, S. 20–28
- Dieffenbach, Johann F.; Fricke, Johann C. G.; Oppenheim, Friedrich W. (Hrsg.) 1837. *Zeitschrift für die gesammte Medicin, mit besonderer Rücksicht auf Hospitalpraxis und ausländische Literatur*. Bd. 4(1), Hamburg, Perthes, Besser & Mauke
- Dirckx, Joris J. J., Decraemer, Willem F. 2001. *Effect of middle ear components on eardrum quasi-static deformation*. In: *Hearing Research*. Ausg. 175, S. 124–137
- Di Martino, Ercole et alii 2004. *Funktionsuntersuchungen der Tuba Eustachii*. In: *HNO*. Ausg. 52(11), S. 1029–1040
- Fanghänel, Jochen et alii (Hrsg.) 2003(a). *A. Waldeyer – Anatomie des Menschen*. Berlin, de Gruyter
- Fanghänel, Jochen et alii 2003(b). *Kopf, Cranium, und Hals, Collum*. In: Fanghänel, Jochen et alii (Hrsg.) 2003. *A. Waldeyer – Anatomie des Menschen*. Berlin, de Gruyter, S. 178–353
-

- Fasold, Wolfgang; Veres, Eva ²2003. *Schallschutz und Raumakustik in der Praxis. Planungsbeispiele und konstruktive Lösungen*. Berlin, Huss-Medien
- Fastl, Hugo; Zwicker, Eberhard ³2006. *Psychoacoustics. facts and models*. Heidelberg, Springer
- Fricke, Jobst P.; Louven, Christoph 2008. *Psychoakustische Grundlagen des Musikhörens*. In: Bruhn, Herbert; Lehmann, Andreas; Kopiez, Reinhard (Hrsg.). *Musikpsychologie. Das neue Handbuch*. Reinbek, Rowohlt, S. 413–436
- Fritscher, Karl D. et alii 2004. *Analyzing inter-individual shape variations of the middle ear cavity by developing a common shape model based on medial representation*. In: *International Congress Series*. Ausg. 1268, S. 243–248
- Fuchs, Paul A. (Hrsg.) 2010. *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press
- Fukazawa, Tatsuya 2002. *How can the cochlea amplifier be realized by the outer hair cells which have nothing to push against?* In: *Hearing Research*. Ausg. 172, S. 53–61
- Gelfand, Stanley A. ⁴2004. *Hearing. An introduction to psychological and physiological acoustics*. New York, Dekker
- Gelfand, Stanley A. 2009. *Essentials of Audiology*. New York, Thieme
- Ghaffari, Rozbeh; Aranyosi, Alexander J.; Freeman, Dennis M. 2007. *Longitudinally propagating traveling waves of the mammalian tectorial membrane*. In: *PNAS*. Ausg. 104(24), S. 16510–16515
- Graumann, Walther; Sasse, Peter (Hrsg.) 2005. *Sinnessysteme, Haut, ZNS, periphere Leitungsbahnen. Compact Lehrbuch Anatomie*. Bd. 4. Stuttgart, Schattauer
- Green, David M. 1993. *Auditory Intensity Discrimination*. In: Yost, William A.; Popper, Arthur N.; Fay, Richard R. (Hrsg.). *Human Psychophysics. Springer Handbook of Auditory Research*. Bd. 3, New York, Springer, S. 13–34
- Gummer, Anthony W.; Hemmert, Werner; Zenner, Hans-Peter 1996. *Resonant tectorial membrane motion in the inner ear: its crucial role in frequency tuning*. In: *PNAS*. Ausg. 93, S. 8727–8732
- Guski, Rainer 1996. *Wahrnehmen – ein Lehrbuch*. Stuttgart, Kohlhammer
- Hackney, Carole M.; Furness, David N. 2010. *Hair bundle structure and mechanotransduction*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 231–256
- Hemilä, Simo; Nummela, Sirpo; Reuter, Tom 1995. *What middle ear parameters tell about impedance matching and high frequency hearing*. In: *Hearing Research*. Ausg. 85, S. 31–44
- Hellbrück, Jürgen; Ellermeier, Wolfgang ²2004. *Hören. Physiologie, Psychologie und Pathologie*. Göttingen, Hogrefe
- Hellbrück, Jürgen; Ellermeier, Wolfgang 2008. *Hören – Psychoakustik – Audiologie*. In: Weinzierl, Stefan (Hrsg.) *Handbuch der Audiotechnik*. Berlin, Springer, S. 41–85
- Helmholtz, Hermann von 1863 und ⁵1896. *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig, Vieweg
- Iwasa, Kuni H. 2010. *Electromotility of outer hair cells*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 179–212
- Jang, Chul Ho et alii 2009. *The effect of increased inner ear pressure on tympanic membrane vibration*. In: *IJPORL*. Ausg. 73, S. 371–375
-

- Keidel, Wolf D. (Hrsg.) 1975. *Physiologie des Gehörs. Akustische Informationsverarbeitung. Einführung für Ärzte, Biologen, Psychologen und Nachrichtentechniker*. Stuttgart, Thieme
- Kemp, David T. 2010. *Otoacoustic emissions and evoked potentials*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 93–137
- Koelsch, Stefan 2005. *Ein neurokognitives Modell der Musikperzeption*. In: *Musiktherapeutische Umschau*. Ausg. 26(4), Göttingen, Vandenhoeck & Ruprecht, S. 365–381
- Koelsch, Stefan; Schröger, Erich 2008. *Neurowissenschaftliche Grundlagen der Musikwahrnehmung*. In: Bruhn, Herbert; Lehmann, Andreas; Kopiez, Reinhard (Hrsg.). *Musikpsychologie. Das neue Handbuch*. Reinbek, Rowohlt, S. 393–412
- Lehnhardt, Ernst; Laszig, Roland (Hrsg.) 2009. *Praxis der Audiometrie*. Stuttgart, Thieme, S. 89–95
- Lukashkin, Andrei N.; Richardson, Guy P.; Russell, Ian J. 2010. *Multiple roles for the tectorial membrane in the active cochlea*. In: *Hearing Research*. Ausg. 266, S. 26–35
- Maempel, Hans-Joachim; Weinzierl, Stefan; Kaminski, Peter 2008. *Audiobearbeitung*. In: Weinzierl, Stefan (Hrsg.) *Handbuch der Audiotechnik*. Berlin, Springer, S. 719–784
- Marcus, Daniel C.; Wangemann, Philine 2010. *Inner ear fluid homeostasis*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 213–230
- Mertens, Paul-Heinrich 1975. *Die Schumannschen Klangfarbengesetze und ihre Bedeutung für die Übertragung von Sprache und Musik*. Frankfurt/M, Bochinsky
- Meyer-Eppler, Werner; Sendhoff, Hanswilly; Rupprath, R. 1959. Residualton und Formantton. In: *Gravesaner Blätter*, Ausg. 14, S. 70–83
- Moore, Brian C.J. 1993. *Frequency Analysis and Pitch Perception*. In: Yost, William A.; Popper, Arthur N.; Fay, Richard R. (Hrsg.). *Human Psychophysics. Springer Handbook of Auditory Research*. Bd. 3, New York, Springer, S. 56–115
- Müller, Werner; Frings, Stephan 2009. *Tier- und Humanphysiologie. Eine Einführung*. Berlin, Springer
- Nilsen, K.E.; Russell, Ian J. 1999. *Timing of cochlear feedback: spatial and temporal representation of a tone across the basilar membrane*. In: *Nature Neuroscience*. Ausg. 2(7), S. 642–648
- Oehler, Michael 2008. *Digitale Impulsformung als Werkzeug für die Analyse und Synthese von Blasinstrumenten*. Frankfurt/M, Lang
- Oxenham, Andrew J.; Wojtczak, Magdalena 2010. *Frequency selectivity and masking*. In: Plack, Christopher J. (Hrsg.). *Hearing. The Oxford Handbook of Auditory Science*. Bd. 3, New York, Oxford Univ. Press, S. 5–44
- Parent, Pierre; Allen, Jont B. 2010. *Time-domain "wave" model of the human tympanic membrane*. In: *Hearing Research*. Ausg. 263, S. 152–167
- Pera, Franz et alii 2003. *Gegenstand und Arbeitsgebiete der Anatomie – Orientierung am menschlichen Körper*. In: Fanghänel, Jochen et alii (Hrsg.) 172003. *A. Waldeyer – Anatomie des Menschen*. Berlin, de Gruyter, S. 1–14
- Politzer, Adam 1896. *Atlas der Beleuchtungsbilder des Trommelfells im gesunden und kranken Zustande. Für praktische Ärzte und Studierende*. Wien, Braunmüller
- Pollard, H. F.; Jansson E. V. 1982. *Analysis and Assessment of Musical Starting Transients*. In: *Acustica*. Ausg. 51(5), S. 249–262
- Ranke, Otto F. 1953. *Physiologie des Gehörs*. In: Ranke, Otto F.; Lullies, Hans. *Gehör. Stimme. Sprache*.
-

- Berlin, Springer, S. 3–162
- Reiss, Gebhard et alii 1989. *Das Stato-akustische Organ. Ein Bildatlas zur Evolution, Physiologie und Morphologie*. Hannover, Duphar
- Reiss, Gebhard 2003. *Gehör- und Gleichgewichtsorgan, Organum vestibulocholeare*. In: Fanghänel, Jochen et alii (Hrsg.) 172003. *A. Waldeyer – Anatomie des Menschen*. Berlin, de Gruyter, S. 591–627
- Reuter, Christoph 1995. *Der Einschwingvorgang nichtperkussiver Musikinstrumente. Auswertung physikalischer und psychoakustischer Messungen*. Frankfurt/M, Lang
- Reuter, Christoph 1996. *Die auditive Diskrimination von Orchesterinstrumenten. Verschmelzung und Heraushörbarkeit von Instrumentalklangfarben im Ensemblespiel*. Frankfurt/M, Lang
- Rochefoucauld, Omerline de la; Olson, Elizabeth S. 2010. *A sum of simple and complex motions on the eardrum and manubrium in gerbil*. In: *Hearing Research*. Ausg. 263, S. 9–15
- Rosowski, John J. 2010. *External and middle ear function*. In: Fuchs, Paul A. (Hrsg.). *The Ear. The Oxford Handbook of Auditory Science*. Bd. 1, New York, Oxford Univ. Press, S. 49–91
- Rossing, Thomas D. (Hrsg.) 2007. *Springer Handbook of Acoustics*. New York, Springer
- Russolo, Mario, Semeraro, A. 1977. *Value of the Isolated Tensor Tympani Muscle Reflex Elicited by Electric Lingual Stimulation*. In: *IJA*, Ausg. 16, S. 373–379
- Schouten, Jan F. 1938. *The perception of subjective tones*. In: *Koninklijke nederlandse Akademie van Wetenschappen*. Nachdruck von: *Proceedings*. Ausg. 41(10), S. 1086–1094
- Schouten, Jan F. 1940. *The perception of pitch*. In: *Philips technical review*. Ausg. 5(10), S. 286–294
- Schriever, Hans 1967. *Physiologie des akustischen Organs*. Berlin, de Gruyter
- Schumann, K. Erich, 1929. *Physik der Klangfarben*. Bd. 2, Berlin, Habil.
- Schünke, Michael; Schulte, Erik; Schumacher, Udo 2009. *Prometheus. Kopf, Hals und Neuroanatomie*. Stuttgart, Thieme
- Schwalbe, Gustav A. 1897. *Das äussere Ohr*. In: Bardeleben, Karl von (Hrsg.). *Sinnesorgane. Handbuch der Anatomie des Menschen*. Bd. 5, Jena, Fischer, S. 113–192
- Seebeck, August 1843. *Über die Sirene*. In: *Annalen der Physik und Chemie*. Bd. 60, S. 449–481
- Seebeck, August 1844. *Über die Definition des Tones*. In: *Annalen der Physik und Chemie*. Bd. 63, S. 353–368
- Siebenmann, Friedrich 1897. *Mittelohr und Labyrinth*. In: Bardeleben, Karl von (Hrsg.). *Sinnesorgane. Handbuch der Anatomie des Menschen*. Bd. 5, Jena, Fischer, S. 193–324
- Spee, Ferdinand Graf von 1896. *Kopf*. In: Bardeleben, Karl von (Hrsg.). *Skelettehre. Handbuch der Anatomie*. Bd. 1, Jena, Fischer, S. 93–372
- Spitzer, Manfred 2002. *Musik im Kopf. Hören, Musizieren, Verstehen und Erleben im neuronalen Netzwerk*. Stuttgart, Schattauer
- Stenfelt, Stefan et alii 2003. *Basilar membrane and osseous spiral lamina motion in human cadavers with air and bone conduction stimuli*. In: *Hearing Research*. Ausg. 181, S. 131–143
- Strenzke, Nicola et alii 2008. *Update zur Physiologie und Patophysiologie des Innenohrs. Pathomechanismen der sensorinauralen Schwerhörigkeit*. In: *HNO*. Ausg. 56, S. 27–36
- Stepp, Cara E.; Voss, Susan E. 2005. *Acoustics of the human middle-ear air space*. In: *JASA*. Ausg. 118(2), S. 861–871
- Stevens, Stanley S.; Hallowell, Davis 1938. *Hearing. Its Psychology and Physiology*. New York, Wiley

- Stumpf, Carl 1890. *Tonpsychologie*. Bd. 2, Stuttgart, Hirzel
- Szunyogh, András; Fehér, György 2006. *Menschliche Anatomie für Künstler*. Potsdam, Ullmann/Tandem
- Viemeister, Neal F.; Plack, Christopher J. 1993. *Time Analysis*. In: Yost, William A.; Popper, Arthur N.; Fay, Richard R. (Hrsg.). *Human Psychophysics. Springer Handbook of Auditory Research*. Bd. 3, New York, Springer, S. 116–154
- Voss, Susan E. et alii 2000. *Acoustic responses of the human middle ear*. In: *Hearing Research*. Ausg. 150, S. 43–69
- Weinzierl, Stefan (Hrsg.) 2008. *Handbuch der Audiotechnik*. Berlin, Springer
- Werwein, Martin; Schick, Matthias 2008. *Wandler, Prozessoren, Systemarchitektur*. In: Weinzierl, Stefan (Hrsg.) *Handbuch der Audiotechnik*. Berlin, Springer, S. 885–944
- Wever, Ernest G.; Bray, Charles W. 1930. *Action current in the auditory nerve in response to acoustical stimulation*. In: *PNAS*. Ausg. 16, S. 344–350
- Winckel, Fritz 1952. *Klangwelt unter der Lupe. Ästhetisch-Naturwissenschaftliche Betrachtungen. Hinweise zur Aufführungspraxis in Konzert und Rundfunk*. Berlin, Hesse
- Winckel, Fritz 1957. *Das Ohr als Zeitmeßorgan. Übersicht über neuere Hörtheorien*. In: *Gravesaner Blätter*, Ausg. 9, S. 83–92
- Winckel, Fritz 1960. *Phänomene des musikalischen Hörens. Ästhetisch-naturwissenschaftliche Betrachtungen, Hinweise zur Aufführungspraxis in Konzert und Rundfunk*. Berlin, Hesse
- Wightman, Frederic L.; Kistler, Doris J. 1993. *Sound Localization*. In: Yost, William A.; Popper, Arthur N.; Fay, Richard R. (Hrsg.). *Human Psychophysics. Springer Handbook of Auditory Research*. Bd. 3, New York, Springer, S. 155–192
- Tan, Robert S. et alii 1997. *Ear size as predictor of chronological age*. In: *Archives of Gerontology and Geriatrics*. Ausg. 25, S. 187–191
- Terhardt, Ernst 1998. *Akustische Kommunikation: Grundlagen mit Hörbeispielen*. Heidelberg, Springer
- Trepel, Martin 2008. *Neuroanatomie. Struktur und Funktion*. München, Urban&Fischer
- Yost, William A. 1993. *Overview: Psychoacoustics*. In: Yost, William A.; Popper, Arthur N.; Fay, Richard R. (Hrsg.). *Human Psychophysics. Springer Handbook of Auditory Research*. Bd. 3, New York, Springer, S. 1–12
- Yost, William A. 2008. *Fundamentals of hearing. An introduction*. San Diego, Elsevier Acad. Press
- Young, Eric D. 2007. *Physiological Acoustics*. In: Rossing, Thomas D. (Hrsg.). *Springer Handbook of Acoustics*. New York, Springer, S. 429–457
- Zwislocki, Josef 1948. *Theorie der Schneckenmechanik. Qualitative und quantitative Analyse*. Solothurn, Gassmann
- Zwicker, Eberhard; Feldtkeller, Richard 1967. *Das Ohr als Nachrichtenempfänger*. Stuttgart, Hirzel
-

