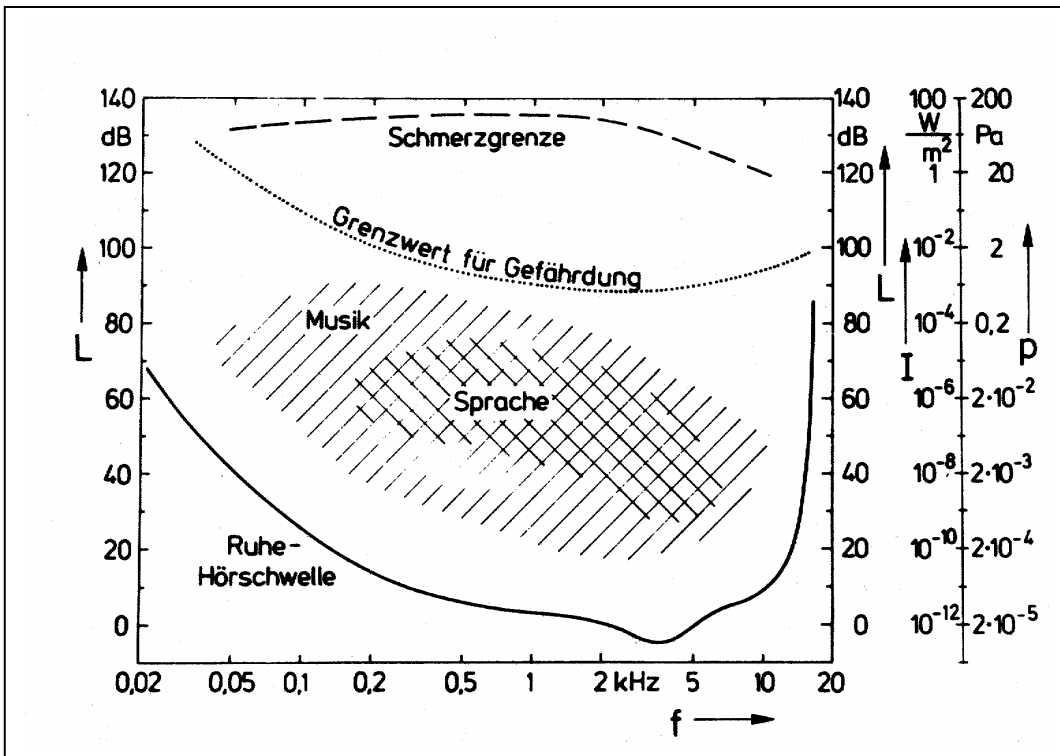


3.3 Hörfläche



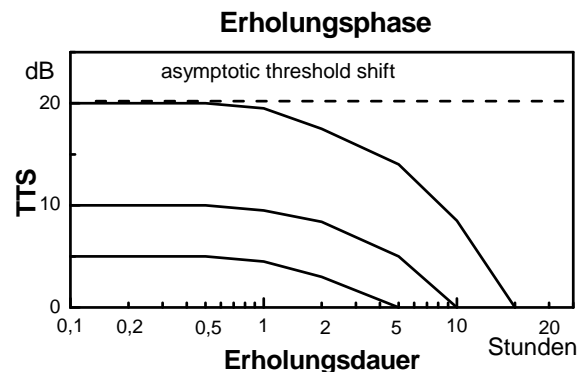
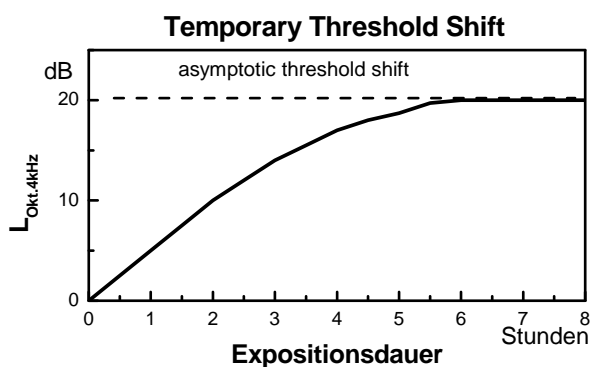
Ruhehörschwelle:
$$L / dB = 3,64 \left(\frac{f}{kHz} \right)^{-0,8} - 6,5 \exp \left(-0,6 \left(\frac{f}{kHz} - 3,3 \right)^2 \right) + 10^{-3} \left(\frac{f}{kHz} \right)^4$$

Männliche Stimme: ca. 100 Hz
 Weibliche Stimme: ca. 200 Hz

Sprache: 100 Hz bis 7000 Hz
 Musik: 40 Hz bis 10 kHz

Telefon: 300 Hz bis 3,4 kHz

Hörschwellenverschiebungen:

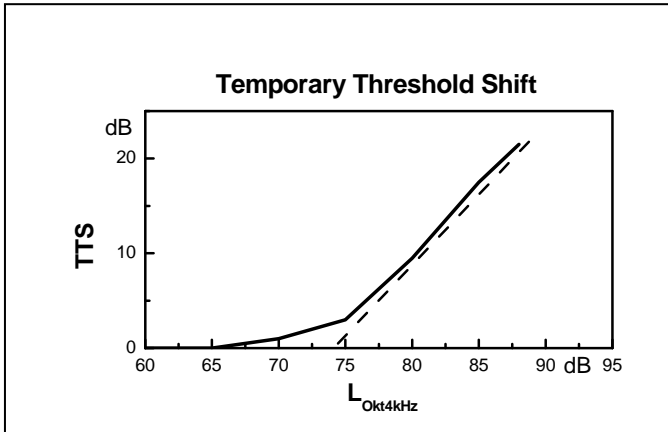


Die Temporary Threshold Shift (temporäre Hörschwellenverschiebung) gibt den gerade noch wahrnehmbaren Oktavpegel bei 4 kHz an, wenn ein Ohr entsprechend lange mit 90 dB(A) beschallt wird. Nach einer entsprechend langen lärmarmen Erholungszeit verläuft die TTS zu 0 dB. Die TTS beginnt ab ca. $L_{Okt.,4kHz}$ von 75 dB und ab einem Pegel $L_{Okt.,4kHz}$ von 80 dB steigt die TTS pro 1 dB Pegelerhöhung um 1,7 dB.

Bei zu hohen Pegeln oder zu geringer Erholungszeit geht die TTS in eine PTS (Permanent Threshold Shift, irreversibler Gehörschaden) über.

Beispiel: Wie hoch ist die Expositionsdauer und Erholungsphase bei einer TTS von 15 dB?

Anstieg der TTS mit 1,7 dB pro 1 dB Oktavpegel $L_{Okt4kHz}$ bei 4 kHz:



Äquivalenter Dauerschallpegel:

$$L_{Aeq} = 10 \log \left[\frac{1}{T} \int_0^T \left(\frac{p_A(t)}{p_0} \right)^2 dt \right] \text{ dB(A)}$$

Zeitkonstanten:

Impuls: 35 ms Anstieg / 1500 ms Abfall
 Fast: 125 ms
 Slow: 1000 ms

Schallpegel L_{Aeq} 4 kHz dB(A)	Expositionszeit Std.
90	8
93	4
96	2
99	1
102	0,5
105	0,25

Pro Intensitätsverdoppelung (Schallpegelerhöhung um 3 dB) halbiert sich die Expositionszeit. Schalle mit einem Pegel von 100 dB(A) im kritischen Frequenzbereich dürfen demnach nur eine Stunde gehört werden!

Jeder der oben angeführten Schallpegel erzeugt innerhalb der angegebenen Expositionszeit eine TTS von 20 dB, die wieder 16 Stunden Erholungszeit benötigt.

Lärminduzierter Gehörschaden:

1. Einzelereignis mit sehr hohem Schallpegel
 Spielzeugpistole $L_{max} = 130$ dB
2. Lärmbelastung über einen längeren Zeitraum

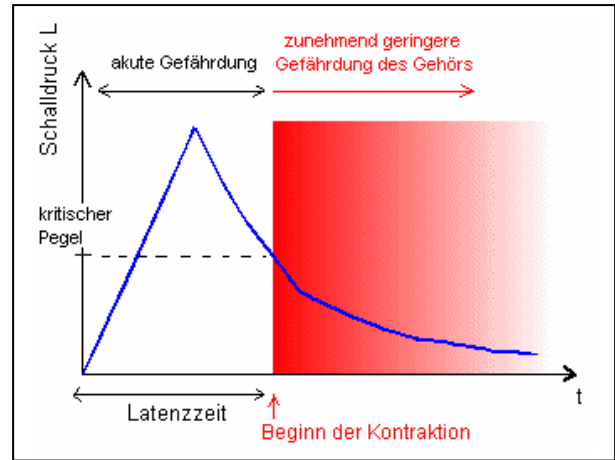
Unfallverhütungsvorschrift (UVV) Lärm: seit 6.3.2007 zurückgezogen

$85 \text{ dB(A)} \leq L \leq 89 \text{ dB(A)}$: Arbeitgeber muss Gehörschutz zur Verfügung stellen
 $L \geq 90 \text{ dB(A)}$: Arbeitnehmer muss Gehörschutz tragen

Lärm- und Vibrations-Arbeitsschutzverordnung (LärmVibrationsArbSchV): seit 6.3.2007 basierend auf EU-Richtlinie

unterer Auslösewert $L_{EX,8h} = 80$ dB(A): Arbeitgeber muss Gehörschutz anbieten
 oberer Auslösewert $L_{EX,8h} = 85$ dB(A): Arbeitnehmer muss Gehörschutz tragen

Stapediusreflex: Schutzmechanismus vor lauten Signalen
 Stapediusmuskel an Gehörknöchelchen kontrahiert sich bei lautem Schall und dämpft Übertragung um 10 bis 20 dB.
 Reaktionszeit: 100 ms



25 % der Bevölkerung hat Hörschaden:

Disco, Konzerte, Knall
 Tanzfläche: 95 - 106 dB(A)
 Anhebung: 2 dB/Std.

Niedersachsen: Ortsfester Arbeitsplatz (Thekenpersonal): 85 dB(A)
 Nicht ortsfester Arbeitsplatz (Kellner): 90 dB(A)

20 dB Hörverlust: Auffälligkeit bei Arzt

Maßnahmen am Einwirkungsort:

- Watte: 10-25 dB
- Stöpsel: 10-35 dB
- Kapsel: 15-40 dB
- Helm: 15-50 dB

Schallschutzfenster:

- Undichte Fenster bis 24 dB
- Isolierverglasung: 34 dB
- Verbundfenster: 44 dB
 (Deckglas 1,2 mm und 6 mm Scheibenabstand)

Technische Anleitung (TA) zum Schutz gegen Lärm: für alle gewerbliche Anlagen

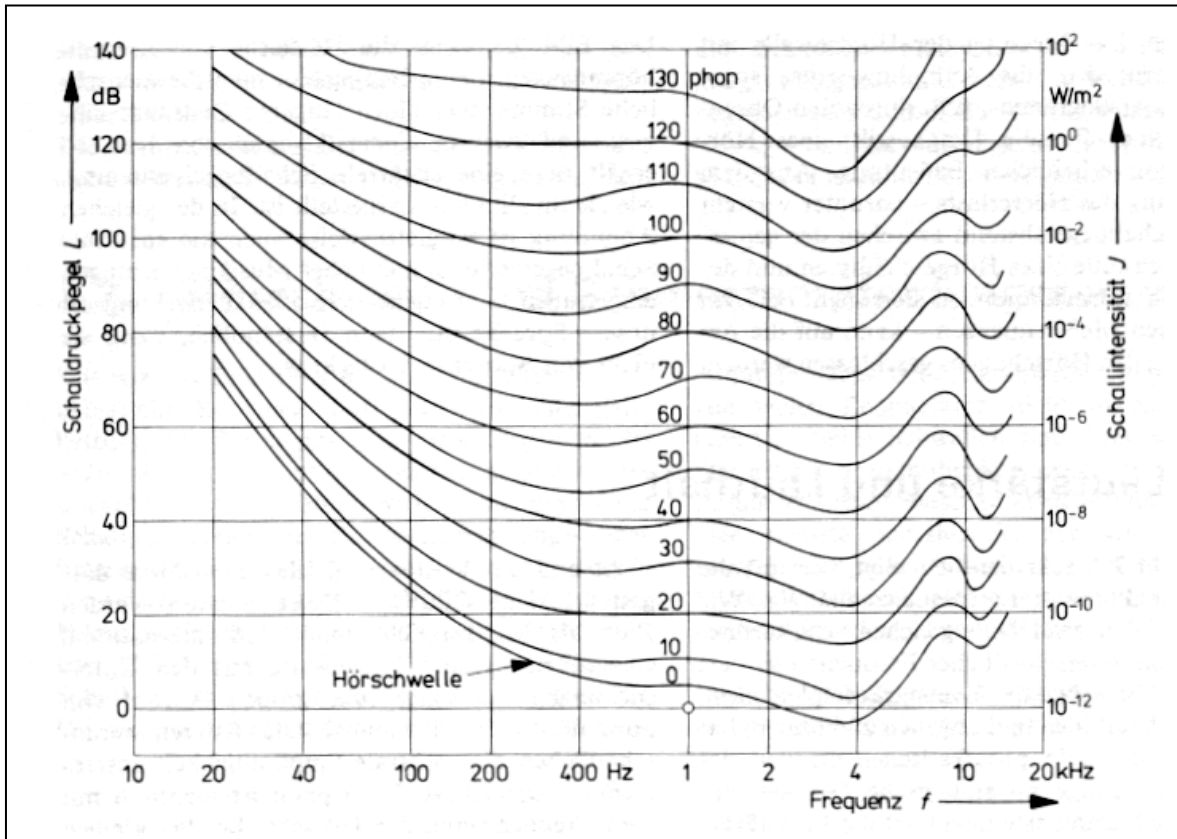
Immissionsrichtwerte für den Beurteilungspegel

In	tags	nachts
Industriegebieten	70 dB(A)	70 dB(A)
Gewerbegebieten	65	50
Kerngebieten, Dorfgebieten, Mischgebieten	60	45
Wohngebieten, Kleinsiedlungsgebieten	55	40
Reinen Wohngebieten	50	35
Kurgebiete, Krankenhäuser	45	35

Umweltbundesamt: Vorschlag: Discos 90-95 dB(A)
 Musikwiedergabegeräte per Kopfhörer 90 dB(A)

Altershörverlust: Verhärtung der Mechanik
 1 kHz: 10 dB, 4 kHz: 20 dB, hohe Frequenzen bei 10 kHz: 30 dB

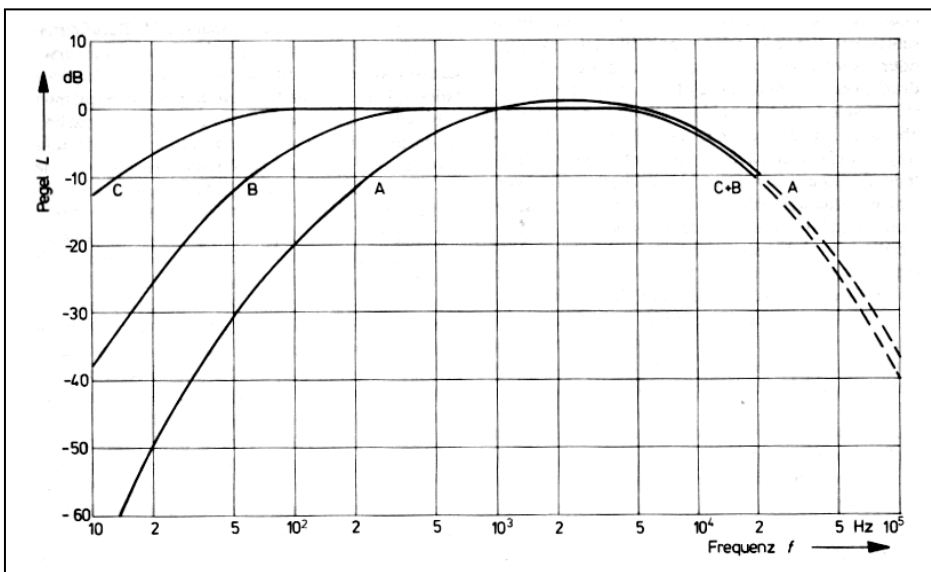
Kurven gleicher Lautstärke, **Isophone**: nach Robinson-Dadson (1956)



Fletcher-Munson führten 1933 erste Untersuchungen mit Sinustönen per **Kopfhörer** durch und veröffentlichten die Kurven. Etwas anderen Kurvenverlauf erhielten Robinson-Dadson 1956 bei frontalem Einfall von Sinustönen über **Lautsprecher** im reflexionsarmen Raum. Letztere sind in ISO R226 (R 454) und DIN 45630 (DIN 1318) genormt. Neuere Kurven sind in ISO 226:2003.

Der **Lautstärkepegel** L_N in phon, der eigentlich **Pegellautstärke** L_N heißen müsste, wird durch subjektiven Lautstärkevergleich mit einem 1-kHz-Sinuston ermittelt. Er gibt den Pegel des 1-kHz-Tones an, der genau so laut ist, wie ein gegebener Ton beliebiger Frequenz und kann aus den Isophonen ermittelt werden..

Filterkurven: A-, B-, C-Bewertung:



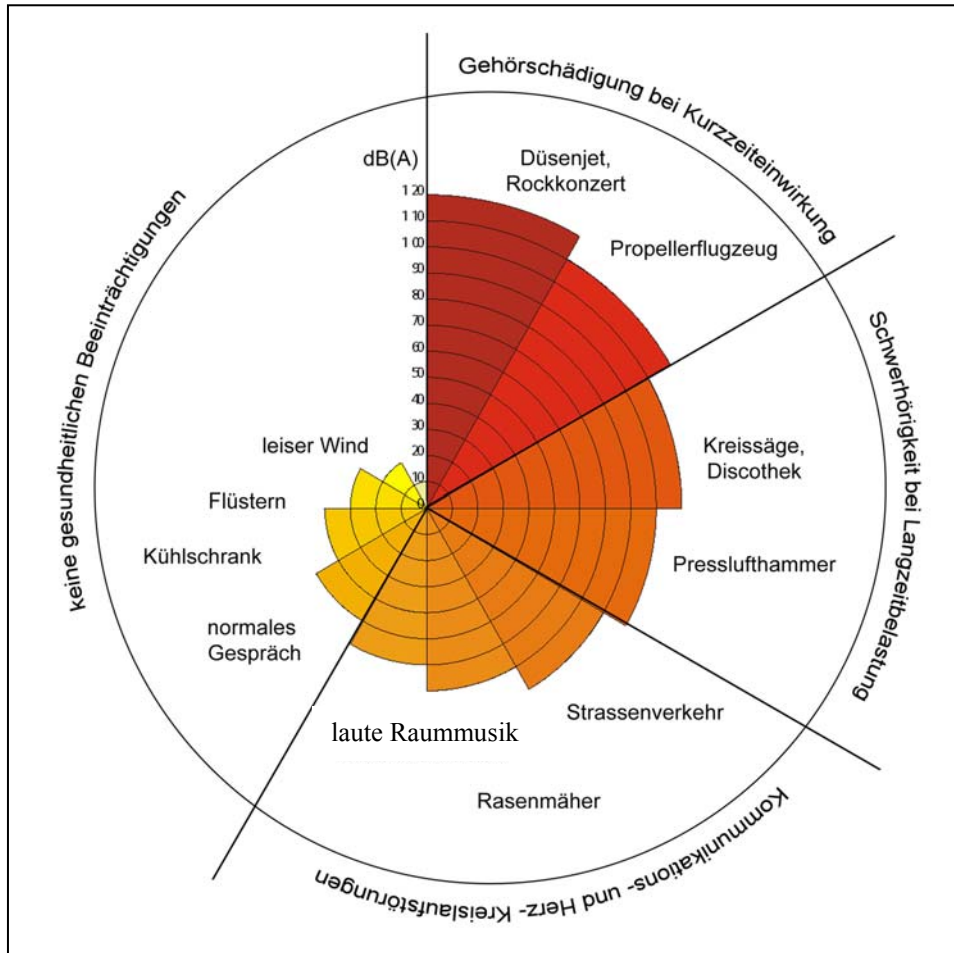
Der Schalldruckpegel in dB(A) ist kein direktes Maß für die Lautheit!

z.B. 6 dB mehr Schalldruckpegel bedeutet **nicht** doppelte Lautheit!

A-Pegel:

- Kein Maß für Lautheit bei hohen Pegeln
bei lauten Schallen zu geringe Pegelanzeige bei tiefen und hohen Frequenzen
- Die angegebenen Pegelwerte lassen nicht auf Lautheitsverhältnisse schließen
40 phon → 60 phon: Empfindung nicht um 50 % gesteigert, sondern um Faktor 4
- Verdeckungseffekte werden nicht berücksichtigt
Die A-Bewertung lässt nur einkanalige Messungen zu

Lärmpegel:



Beispiel: Welche Pegellautstärke hat ein 50-Hz-Ton mit 90 dB(A)?
Wie lange darf er ohne Hörschädigung angehört werden?

3.4 Frequenzgruppe

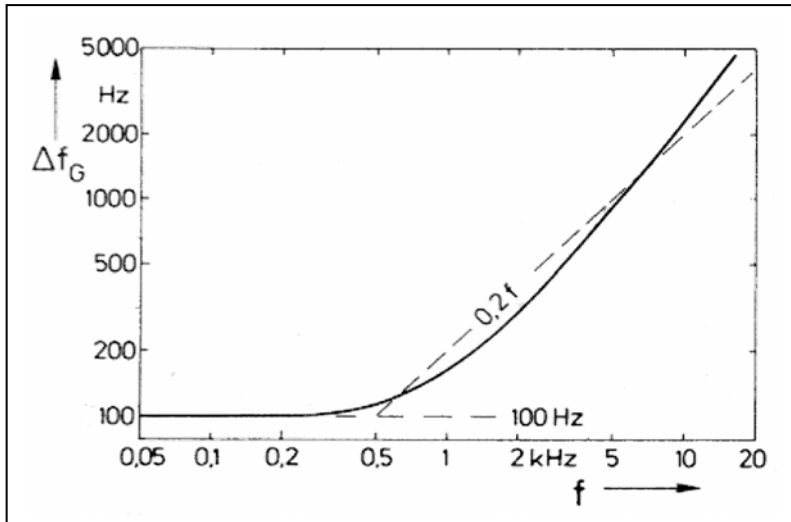
Die Frequenzgruppe Δf_G spielt bei der Lautstärkeempfindung und bei der Richtungsempfindung eine wichtige Rolle. Das Gehör analysiert den Schall in charakteristischen Frequenzbändern, den sog. Frequenzgruppen.

$$f < 500 \text{ Hz: } \Delta f_G = 100 \text{ Hz}$$

$$f \geq 500 \text{ Hz: } \Delta f_G \approx \text{Terz} \approx 0,2 \cdot f$$

Demnach wächst die Frequenzgruppenbreite von 500 Hz bis 5 kHz um den Faktor 10 an und beträgt bei 5 kHz etwa 1000 Hz. Bei sehr hohen Frequenzen am Ende des Hörbereiches ist die Frequenzgruppe etwa 3 kHz breit.

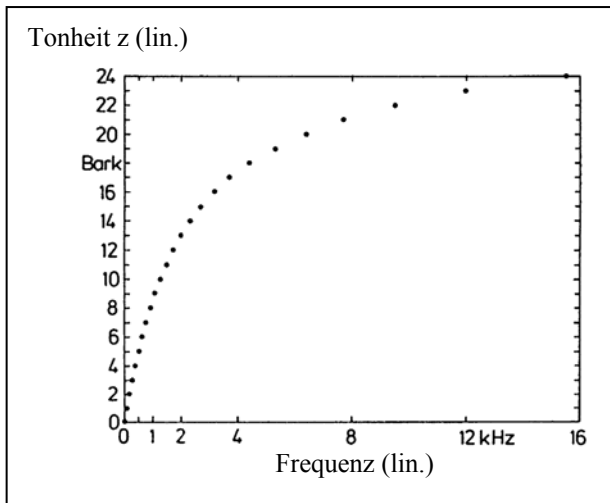
Breite Δf_G der Frequenzgruppe als Funktion der Frequenz f :



Der Hörbereich umfasst 30 Terzen und 24 Frequenzgruppen.

Der Frequenzbereich des Telefons von 300 Hz bis 3,4 kHz umfasst 13 Frequenzgruppen.

Die Anzahl der bis zur Frequenz f aneinanderreihbaren Frequenzgruppen ergibt den Zusammenhang zwischen der Tonheit z in Bark und der Frequenz in Hertz:



z	f_l, f_u	f_c	z	Δf_G	z	f_l, f_u	f_c	z	Δf_G
Bark	Hz	Hz	Bark	Hz	Bark	Hz	Hz	Bark	Hz
0	0				12	1720			
		50	0.5	100			1850	12.5	280
1	100	150	1.5	100	13	2000	2150	13.5	320
2	200	250	2.5	100	14	2320	2500	14.5	380
3	300	350	3.5	100	15	2700	2900	15.5	450
4	400	450	4.5	110	16	3150	3400	16.5	550
5	510	570	5.5	120	17	3700	4000	17.5	700
6	630	700	6.5	140	18	4400	4800	18.5	900
7	770	840	7.5	150	19	5300	5800	19.5	1100
8	920	1000	8.5	160	20	6400	7000	20.5	1300
9	1080	1170	9.5	190	21	770	8500	21.5	1800
10	1270	1370	10.5	210	22	9500	10500	22.5	2500
11	1480	1600	11.5	240	23	12000	13500	23.5	3500
12	1720				24	15500			

Die Umrechnung von Frequenz f in Tonheit z und umgekehrt kann mit sehr guter Näherung mit folgenden Formeln erfolgen:

$$z = \frac{26,81 f / \text{Hz}}{1960 + f / \text{Hz}} - 0,53 \text{ Bark}$$

$$f = \frac{1960 (z / \text{Bark} + 0,53)}{26,28 - z / \text{Bark}} \text{ Hz}$$

Rauschsignale, bei denen die Bandbreite Δf kleiner ist als die zur entsprechenden Mittenfrequenz f_m gehörige Frequenzgruppenbreite Δf_G , werden allgemein als **Schmalbandrauschen** (SBR) bezeichnet.

Beispiel: Welche Frequenzgruppenbreite und Tonheit liegt bei 6 kHz vor?

3.5 Verdeckung

3.5.1 Spektrale Verdeckung

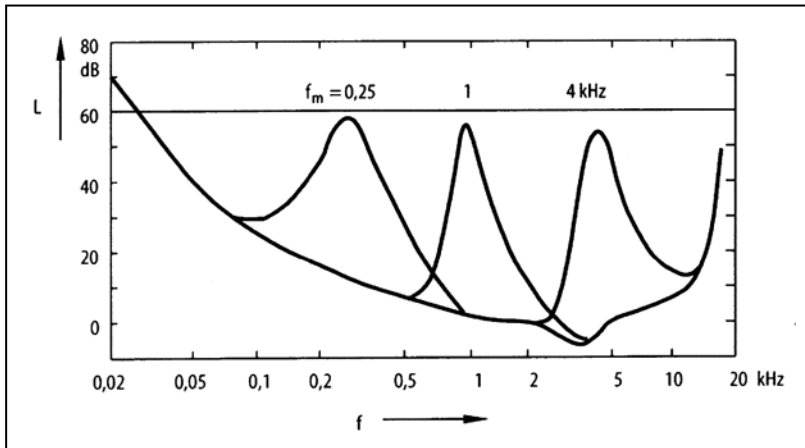
Ein stationäres Schallereignis (z.B. Musik) bzw. ein Testton, der in ruhiger Umgebung gut hörbar ist, wird in Gegenwart eines Störschalls (z.B. Presslufthammer) bzw. Maskierers unhörbar. Diesen Effekt nennt man Verdeckung. Damit der Nutzschall trotz des Maskierers hörbar bleibt, muss sein Pegel erheblich höher sein als in ruhiger Umgebung.

Zur quantitativen Beschreibung der Verdeckungseffekte dienen die **Mithörschwellen**, die als derjenige Schalldruckpegel eines Testschalles definiert sind, den dieser haben muss, damit er neben dem Störschall (Maskierer) gerade noch wahrgenommen werden kann.

Die Zeitstruktur kann zunächst vernachlässigt werden, wenn der Schall länger als 200 ms dauert.

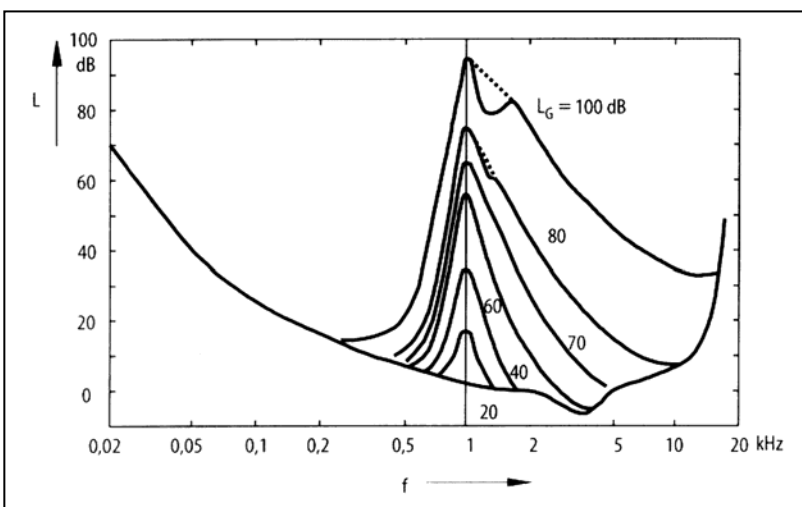
Zur Bestimmung der Mithörschwellen wird den Versuchspersonen neben dem Schmalbandrauschen als Maskierer ein Sinuston dargeboten, dessen Pegel so einzustellen ist, dass er gerade noch hörbar ist, während sich seine Frequenz langsam ändert. Der von mehreren Versuchspersonen gemittelte Schalldruckpegel des Testtones über der Frequenz ergibt die Mithörschwelle. Die Mithörschwelle ohne Maskierer ist die Ruhehörschwelle.

Mithörschwelle bei Schmalbandrauschen unterschiedlicher Mittenfrequenz bei 60 dB:



Mithörschwellen eines maskierenden Schmalbandrauschens mit einem Pegel von 60 dB bei einer Mittenfrequenz von 250 Hz und einer Bandbreite von 100 Hz, bei einer Mittenfrequenz von 1 kHz und einer Bandbreite von 160 Hz sowie bei einer Mittenfrequenz von 4 kHz und einer Bandbreite von 700 Hz.

Mithörschwelle bei Schmalbandrauschen unterschiedlichen Pegels L_G :



Die Mithörschwelle eines Schmalbandrauschens bei der Mittenfrequenz 1 kHz mit einer Bandbreite von 160 Hz weist insbesondere bei den Flanken zu höheren Frequenzen hin eine Pegelabhängigkeit auf. Diese geringere Steigung der Mithörschwelle nennt man **nichtlineare Auffächerung**.

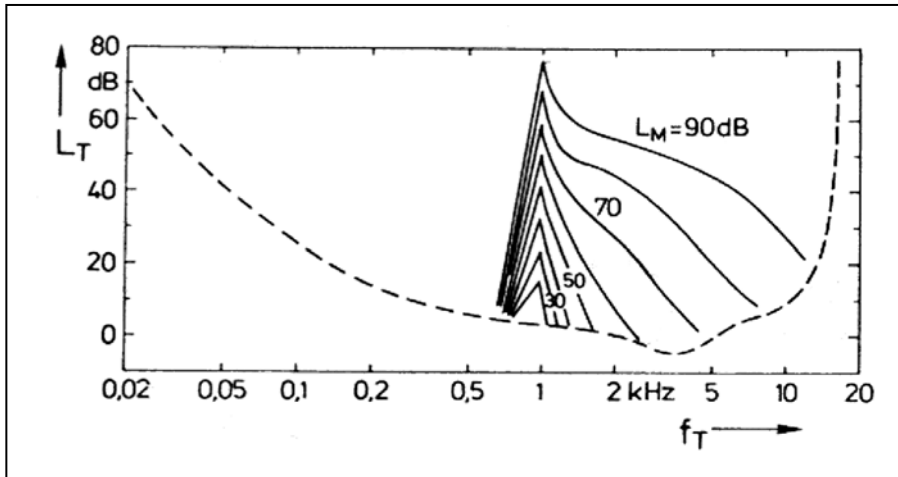
Steigungen der Flanken:

$$S_1 = 27 \text{ dB/Bark}$$

$$S_2 = 24 + \frac{0,23 \text{ kHz}}{f_m} - 0,2 \frac{L_G}{\text{dB}} \text{ dB/Bark}$$

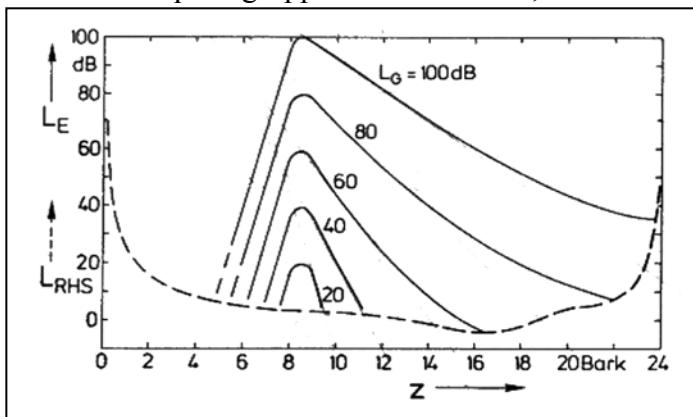
Bei Sinustönen als Maskierer besitzen die zugehörigen Mithörschwellen eine etwas veränderte Form, die wesentlichen Verläufe bleiben aber gleich. Auch hier ist eine nichtlineare Auffächerung der oberen Flanke mit zunehmenden Maskiererpegel L_M erkennbar, während die Steigung der unteren Flanke abgesehen von sehr niedrigen Maskiererpegeln weitgehend gleich bleibt. Wenn der Testton eine ähnliche Frequenz (hier: 1 kHz) aufweist wie der Maskiererton, so ist die Bestimmung der Mithörschwelle sehr schwierig, da beide spektral benachbarten Töne Schwebungen verursachen, die von Verdeckungseffekten schwer zu unterscheiden sind.

Mithörschwelle L_T verdeckt durch 1 kHz-Sinustöne verschiedenen Pegels L_M :

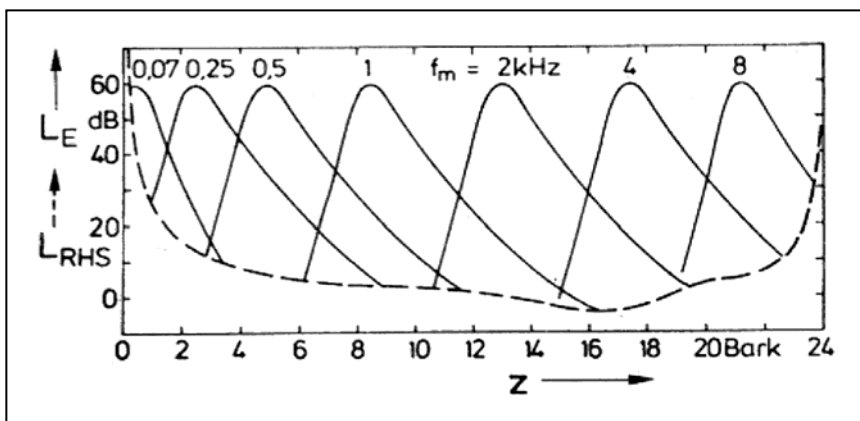


Die Darstellung der Mithörschwellen in der Tonheit führt zu einheitlichen Kurven, die auf der Tonheitsskala z verschoben werden können.

MHS für Frequenzgruppenrauschen bei 8,5 Bark bei versch. Maskiererpegeln L_G :



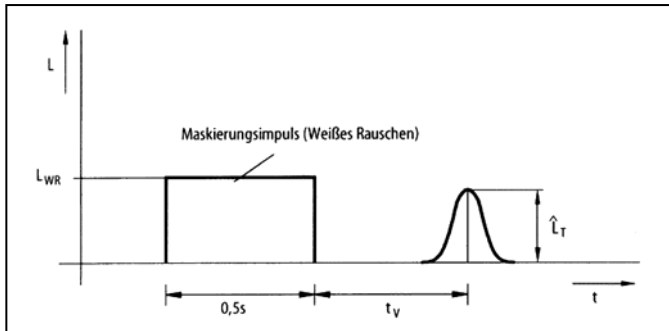
MHS für Schmalbandrauschen versch. Mittenfrequenz f_m und 60 dB Maskiererpegel:



3.5.2 Zeitliche Verdeckung

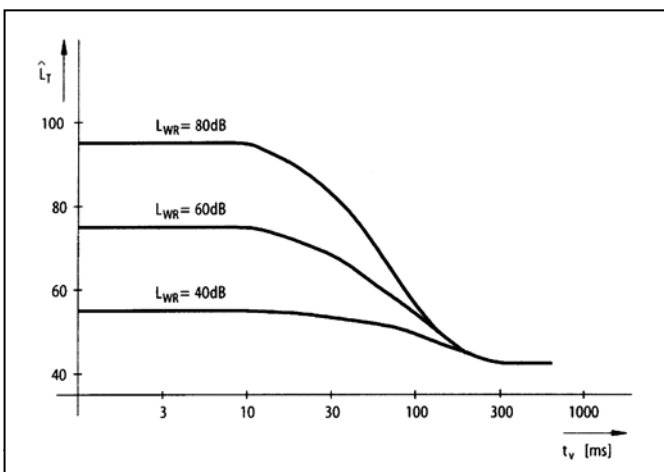
Zeitliche Verdeckungseffekte werden dadurch ermittelt, dass den Versuchspersonen per Kopfhörer Weißes Rauschen (alle Frequenzanteile mit gleicher spektraler Dichte vorhanden) über eine Dauer von 0,5 sec. dargeboten wird. Auf diesen Maskierungsimpuls folgt mit einem zeitlichen Abstand von t_v ein 20 μ s langer Testtonimpuls, den die Personen im Pegel so einstellen müssen, dass er gerade noch hörbar ist.

Messung der Nachhörschwelle:

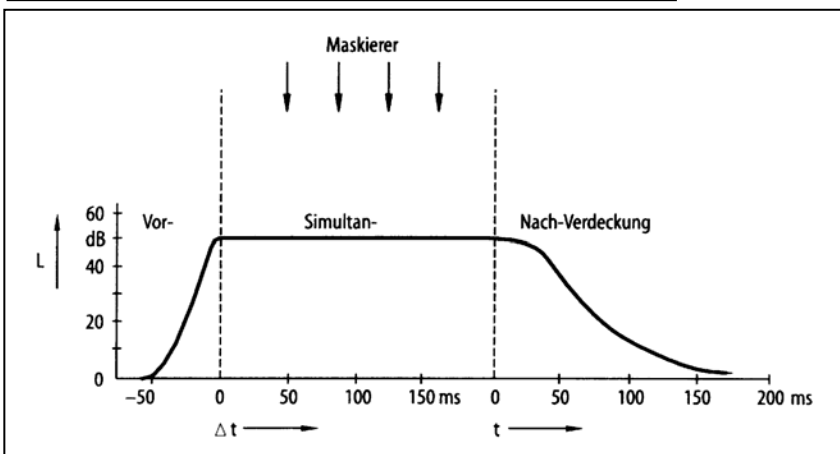


Befindet sich der Testtonimpuls nach dem Maskiererimpuls erhält man die **Nachverdeckung** bzw. Nachhörschwelle, befindet er sich zeitlich gesehen vor dem Maskierer erhält man die **Vorverdeckung**.

Nachhörschwelle bei verschiedenen Amplituden des Maskierers:



Die Nachverdeckung hält selbst 10 ms nach Abschalten des Maskierers noch an und klingt dann mit zunehmender zeitlicher Verzögerung t_v des Testtonimpulses langsam ab bis nach 200 ms keine Nachverdeckung mehr registriert wird.



Neben den zeitlichen Verdeckungseffekten besitzt das Ohr auch eine „Einschwingzeit“, d.h. Schalle die kürzer als 100 ms andauern, werden leiser wahrgenommen, da bei kurzen Signalen das Ohr über der Zeit integriert. **Das Produkt aus Intensität und Zeitdauer ist daher konstant:**

$$I \cdot T = \text{const. für } T < 100 \text{ ms}$$

Deshalb verwendet man bei Schallpegelmessungen oft die Zeitkonstante „fast“ (125 ms).

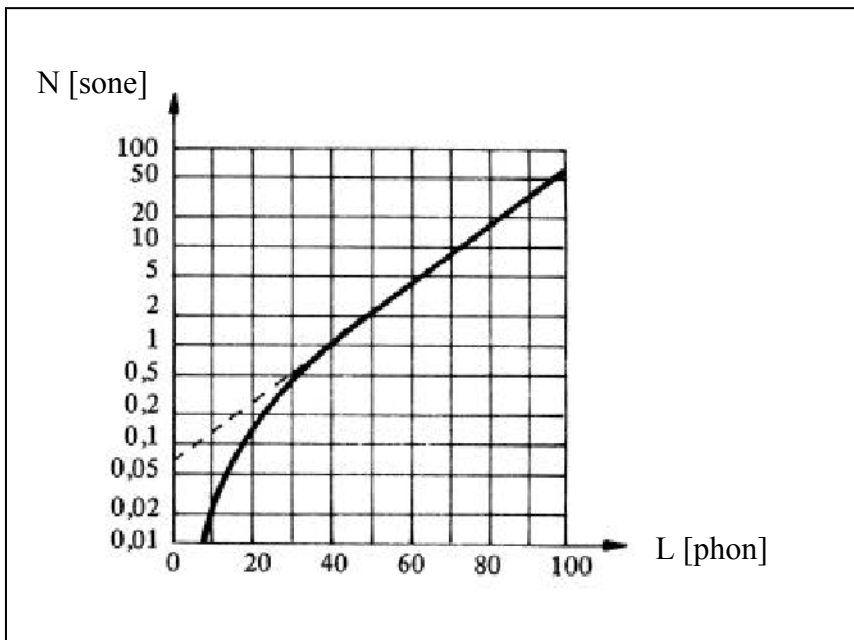
3.6 Lautheit

Die Lautstärkeempfindung ist eine Intensitätsempfindung. Ihre Empfindungsfunktion geht also nicht unmittelbar aus den eben wahrnehmbaren Pegelstufen hervor. Dennoch spielen eben wahrnehmbare Schalldruck- oder Schallpegeländerungen eine wichtige Rolle. Die Empfindungsfunktion für die Lautstärkeempfindung, die Lautheit, lässt sich nur aus Verhältnismessungen ableiten.

Die **gedrosselte Lautheit** liegt zwischen der normalen Lautheit, die ein Testschall ohne Störschall hervorruft, und der Lautheit Null des Testschalles, die bei der Mithörschwelle, bei der die Störung den Testschall gerade verdeckt, erreicht wird.

Mit Hörversuchen wurde festgestellt, dass die Lautheit nur ab einem Pegel von ca. 40 dB linear mit dem Darbietungspegel ansteigt. Bezogen auf die Schallintensität gehorcht die Lautheit N somit einem Potenzgesetz. Das Berechnungsverfahren der stationären Lautheit ist in DIN 45631 bzw. in ISO 532B genormt.

Zusammenhang zwischen Schalldruckpegel L und Lautheit N in sone für Sinustöne:



$$\frac{N_{1kHz}}{\text{sone}} = \frac{1}{16} \left(\frac{I_{1kHz}}{I_0} \right)^{0,3} \approx 2^{\frac{L_{1kHz} / \text{dB} - 40}{10}} \quad \text{für : } L \geq 40 \text{ phon}$$

$$L_{1kHz} = 33,22 \log \frac{N}{\text{sone}} + 40 \text{ phon} \quad \text{für : } N \geq 1 \text{ sone}$$

Als **Lautstärkepegel** eines beliebigen Schalles wird derjenige Pegel eines 1-kHz-Tones bezeichnet, der bei frontalem Einfall auf die Versuchsperson in einer ebenen Welle die gleiche Lautstärkeempfindung hervorruft wie der zu messende Schall.

Per Definition besitzt ein 1-kHz-Ton mit einem Pegel von 40 dB eine Lautheit von 1 sone (lat. sonare = klingen). Durch Hörversuche wurde ermittelt, dass eine Pegelerhöhung des 1-kHz-Sinustones um 10 dB eine Lautheitsverdoppelung bzw. eine Pegelabsenkung um 10 dB eine Lautheitshalbierung hervorruft.

Es ergibt sich damit folgender Zusammenhang:

Pegel L	Lautheit N
30 dB	~ 0,5 sone
40	1
50	2
60	4
70	8
80	16
90	32
100	64

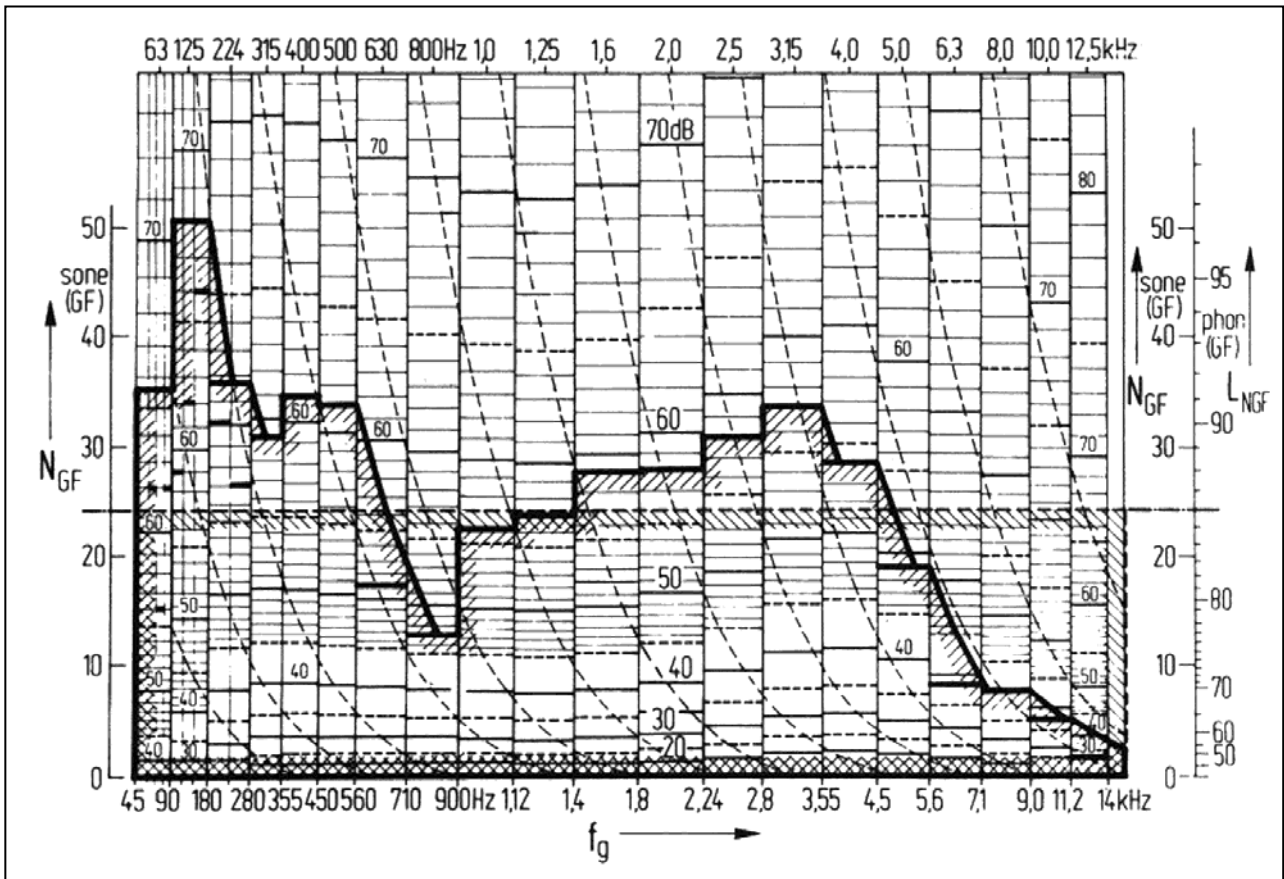
40 phon → 60 phon: Faktor 4
 40 phon → 100 phon: Faktor 64

Für **Sinustöne** beliebiger Frequenz kann die Lautheit leicht bestimmt werden. Aus den Kurven gleicher Lautheit wird der zugehörige äquivalente Pegel des 1 kHz-Tones in phon graphisch ermittelt. Dieser Pegel L_{1kHz} kann dann mit obiger Formel in die Lautheit N in sone umgerechnet werden. Die eben wahrnehmbare Pegeländerung liegt bei ca. 1 dB.

Beispiel: Welche Lautheit besitzt ein 50-Hz-Ton mit 90 dB(A)?

Für **Rauschen** trifft dieser einfache Zusammenhang zwischen Pegel und Empfindung nicht mehr zu. Hier entspricht in etwa eine Pegelerhöhung von 13 dB einer Lautheitsverdoppelung. Bei der Lautheitsberechnung von breitbandigen Schallen müssen auch die Verdeckungseffekte berücksichtigt werden. Handelt es sich nicht um stationäre, sondern um zeitlich verändernde Signale müssen auch Einschwingzeiten des Gehörs beachtet werden.

Graphische Ermittlung der Gesamtlautheit von Schall mit mehreren Frequenzkomponenten:



Die Lautheit innerhalb einer Frequenzgruppe bezeichnet man als *spezifische Lautheit N'*. Die Flächenanteile aller spezifischen Lautheiten addieren sich zur *Gesamtlautheit*.

Die Lautheiten unterschiedlicher Schalle können zueinander ins Verhältnis gesetzt werden. Die Lautheit ist abhängig von Schallpegel, Frequenz, Bandbreite und Zeitdauer des Signals.

Für die Ermittlung der Lautheit von breitbandigen Schallen werden in der Norm DIN 45631 fünf Kurvenblätter für das freie Schallfeld (N_{GF}) und fünf Blätter für das diffuse Schallfeld (N_{DF}) zur Verfügung gestellt. Hiermit kann graphisch die spezifische Lautheit, die Gesamtlautheit sowie der zugehörige Lautstärkepegel L_{NGF} bzw. L_{NGD} ermittelt werden. Es sind damit Lautstärkepegel bis 65, 80, 98, 115 und 130 phon bestimmbar.

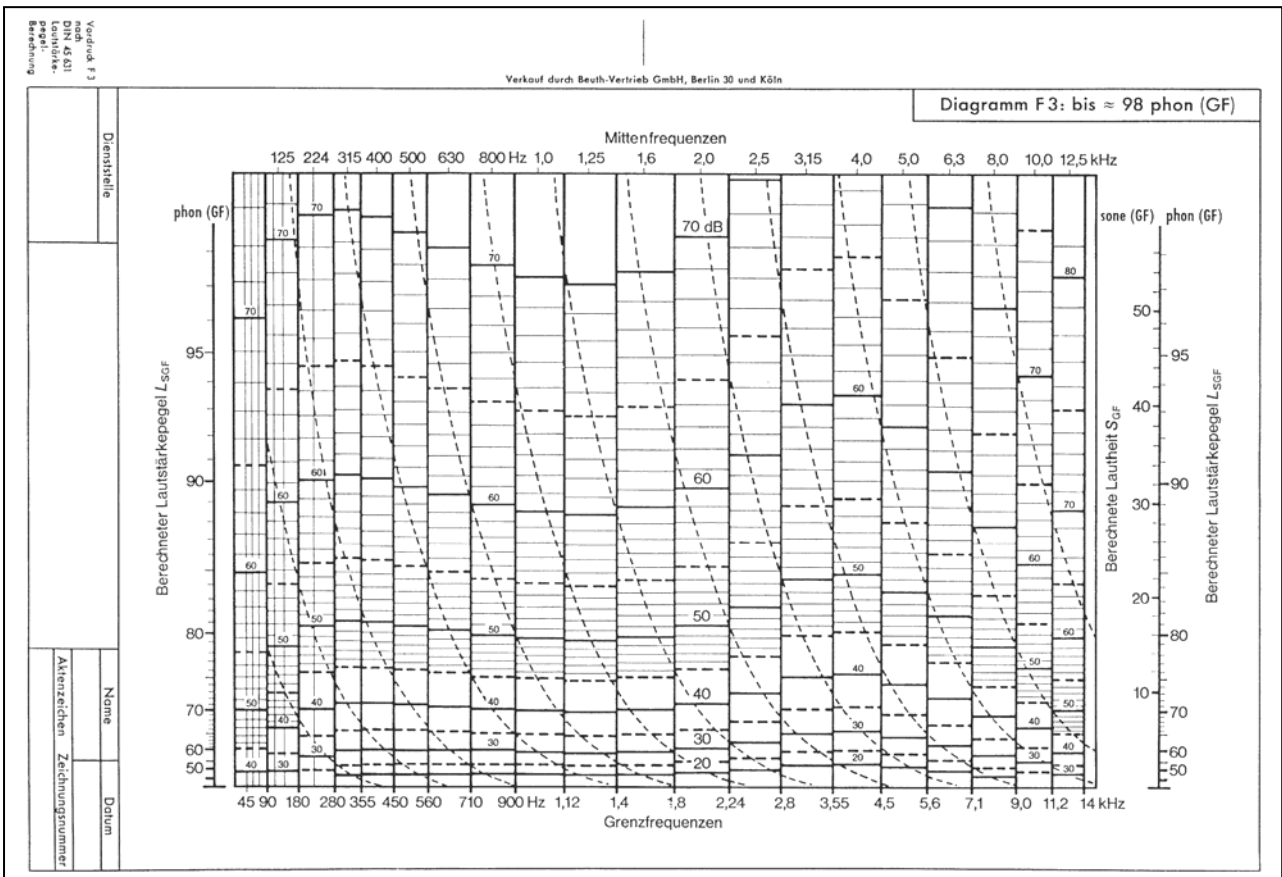
Die Anwendung der Kurvenblätter erfolgt folgendermaßen:

Man wählt das zur Schallfeldform passende Blatt aus, das den höchsten vorkommenden Terzpegel gerade noch enthält. Die gemessenen Terzpegel werden als horizontale Strecken eingetragen und notfalls, wenn der benachbarte Terzpegel sehr viel geringer ist, ihre obere Flanke parallel zu den gestichelten Kurven abfallend dargestellt, um die Verdeckungseffekte zu berücksichtigen. Bei Frequenzen unter 280 Hz müssen mit Hilfe des Nomogramms mehrere Terzpegel zusammengefasst werden, um der konstanten Bandbreite der Frequenzgruppe von 100 Hz Rechnung zu tragen.

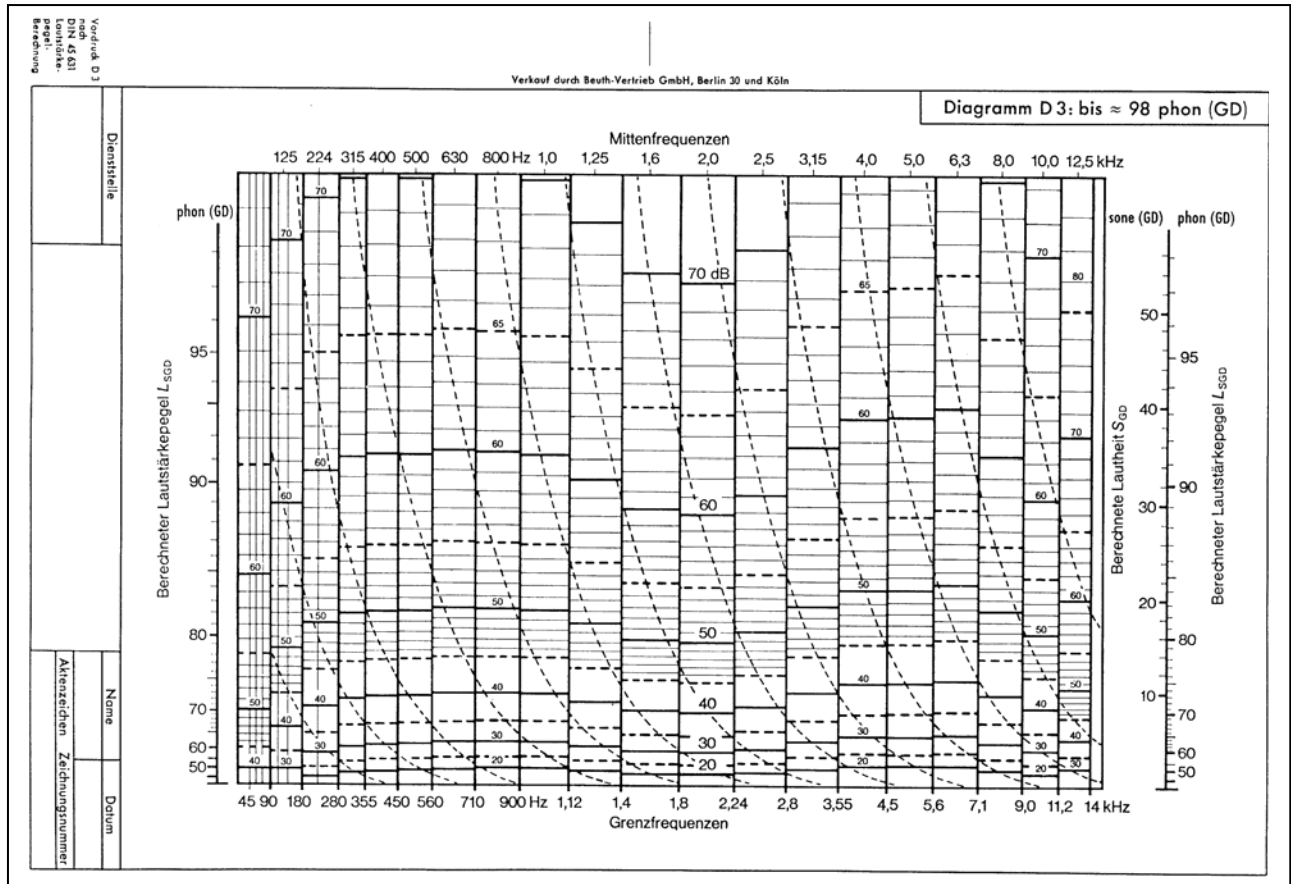
Auf diese Weise entsteht zusammen mit der Abszisse eine stark umrandete Fläche, die der Gesamtlautheit entspricht. Die Höhe eines flächengleichen Rechtecks über der Abszisse entspricht der Gesamtlautheit in sone. Ebenso kann der äquivalente Lautstärkepegel L_N eines gleichlauten 1-kHz-Tones abgelesen werden.

Im obigen Beispiel erhält man somit eine berechnete Lautheit N_{GF} von 24 sone(GF) und einen Lautstärkepegel L_{NGF} von 86 phon(GF).

Lautheitsberechnungsblatt für freies Schallfeld:

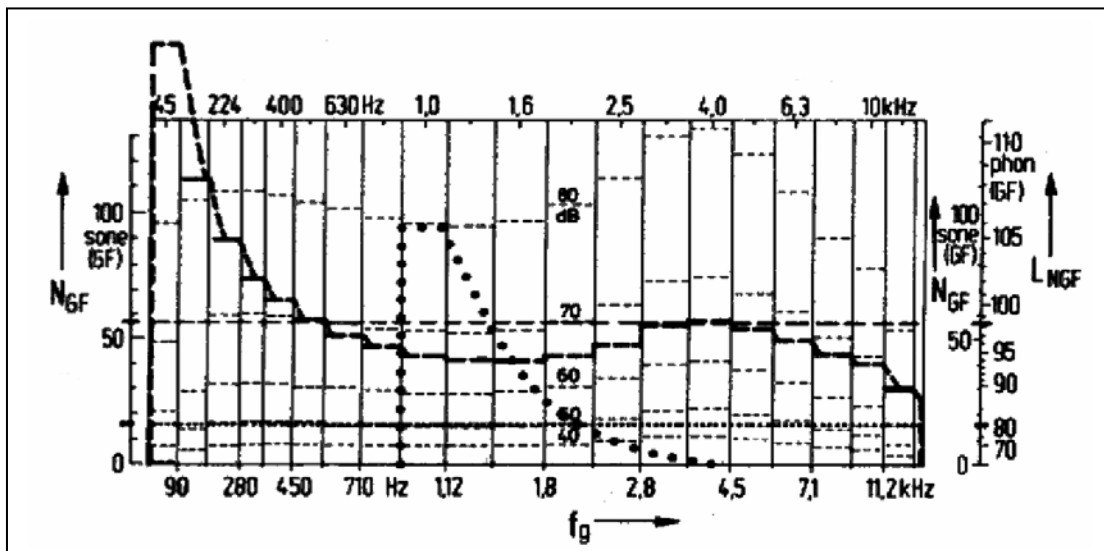


Lautheitsberechnungsblatt für diffuses Schallfeld:



Beispiel:

Ein breitbandiger Schall spezieller spektraler Zusammensetzung (gestrichelt) ist trotz gleichen Wertes für den A-bewerteten Schallpegel um den Faktor 3,5 (!) lauter als ein separat dargebotener schmalbandiger Schall (gepunktet).

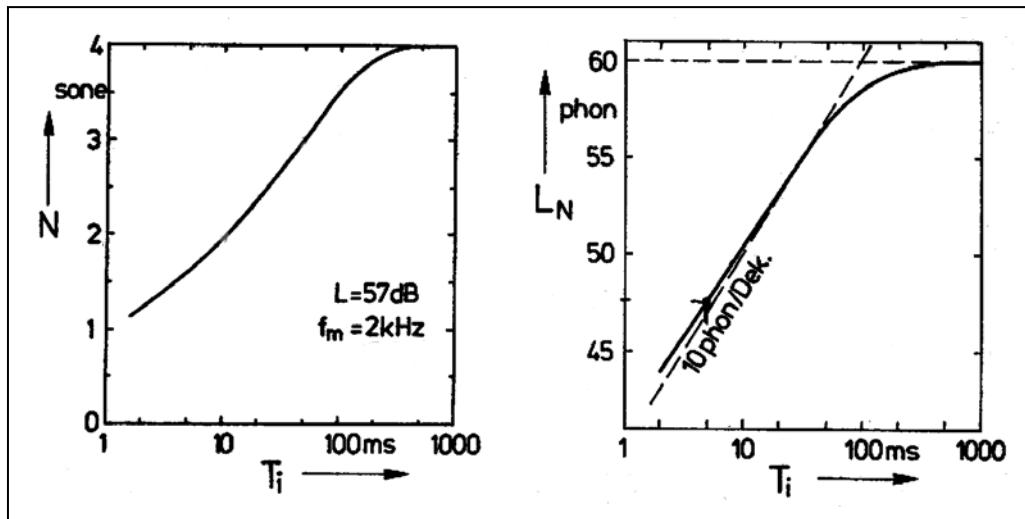


Die gedrosselte Lautheit entspricht der nicht-verdeckten Lautheit und damit Fläche des Signals. Das schmalbandige Signal (gepunktet) wird im obigen Beispiel bei gleichzeitiger Darbietung beider Schalle leiser gehört und zwar entsprechend seiner unverdeckten Fläche. Das breitbandige Signal drosselt also infolge seiner Maskierung die Lautheit des schmalbandigen Signals. Gleiches gilt, wenn man die Lautheit des breitbandigen Signal für sich betrachtet. Diese wird durch die maskierende Wirkung des schmalbandigen Signals ebenfalls gedrosselt.

Abhängigkeit der Lautheit von der Signaldauer:

Wird die Dauer T_i eines eine Sekunde langen Signals mit einer bestimmten Lautheit laufend verkürzt, so stellt man fest, dass ab 100 ms geringer werdende Signaldauern zu deutlich geringerer Lautheit führen. Das Gehör integriert also über das Signal mit einer Zeitkonstante von etwa 100 ms. Um diese Eigenschaft auch auf physikalischer Seite zu berücksichtigen, wird der Pegel vielfach mit Messgeräten mit Zeitkonstante fast (125 ms) gemessen.

Abhängigkeit der Lautheit N und des Lautstärkepegels L_N eines 2-kHz-Tonimpulses, der aus einem Dauerton mit dem Schallpegel $L = 57$ dB ausgeschnitten ist, abhängig von seiner Dauer T_i :



Beispiel: Um welches Lautheitsverhältnis bzw. Schalldruckverhältnis verringert sich ein von 100 ms auf 10 ms verkürzter 2-kHz Tonimpuls?

3.7 Tonhöhe

Die Tonhöhenempfindung ist eine **Positionsempfindung**, da versch. Frequenzen über die Frequenz-Orts-Transformation auf versch. Orte der Basilarmembrane abgebildet werden. Daher kann man die Empfindungsfunktion der Tonhöhe aus den Reizstufen ermitteln. Eine zweite Möglichkeit der Bestimmung der Empfindungsfunktion besteht in der Messung von Verhältniswerten.

Eben wahrnehmbare Frequenzänderungen:

Die eben wahrnehmbaren Änderungen der Frequenz von Sinustönen, auch Unterschiedsschwellen der Frequenz genannt, können durch plötzliche Änderungen der Frequenz erzeugt werden. Solche Frequenzänderungen haben allerdings den Nachteil, dass sie Knacke hervorrufen. Diese Knacke stören die Bestimmung der eben wahrnehmbaren Frequenzänderung. Daher wird sie häufig mit Hilfe einer sinusförmigen Frequenzmodulation bestimmt. Die Frequenzänderung ist nachrichtentechnisch bei Frequenzmodulation über den Frequenzhub Δf definiert. Die Frequenz f schwankt bei solch einer Modulation zwischen $f_1 - \Delta f$ und $f_1 + \Delta f$. Die Größe $2\Delta f$ ist die Reizstufe, in unserem Fall die gesuchte Frequenzstufe. Das Gehör ist für sinusförmige Frequenzänderungen bei Modulationsfrequenzen um 4 Hz am empfindlichsten.

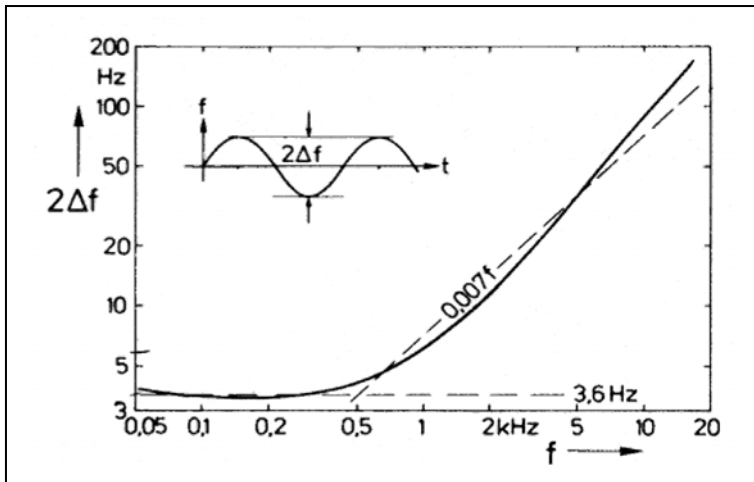
Am Wichtigsten ist die Abhängigkeit der Frequenzstufe $2\Delta f$ von der Frequenz f . Bei konstanter Lautstärke (z.B. 60 phon) und f_{mod} gleich 4 Hz ergeben sich folgende Unterschiedsschwellen:

$$\begin{aligned}
 f < 500 \text{ Hz: } & 2\Delta f = 3,6 \text{ Hz} \\
 f \geq 500 \text{ Hz: } & 2\Delta f \approx 0,007 \cdot f
 \end{aligned}$$

Bei höheren Frequenzen kann also eine Änderung der Frequenz von 0,7 % gerade wahrgenommen werden, während die Frequenzstufe z.B. bei 100 Hz bei 3,6 % liegt.

Ein musikalischer Halbton entspricht einer Frequenzabweichung von etwa 6 Prozent. Dies heißt, dass eine Änderung um zwei Frequenzstufen bereits der Änderung um einen musikalischen Halbton entspricht. Frequenzänderungen tiefer Frequenzen können daher nur sehr schlecht wahrgenommen werden. Da die musikalischen Töne, insbesondere die mit tiefer Grundfrequenz, aus vielen Harmonischen aufgebaut sind, wird die Frequenzänderung solcher Töne an der Änderung der hohen Harmonischen erkannt, ein für die musikalische Interpretation wichtiges Ergebnis. Die eben wahrnehmbaren Frequenzänderungen sind bei mittleren und hohen Frequenzen mit einem Wert von 0,7 % sehr klein. Das Gehör ist demnach gegen Frequenzänderungen sehr empfindlich.

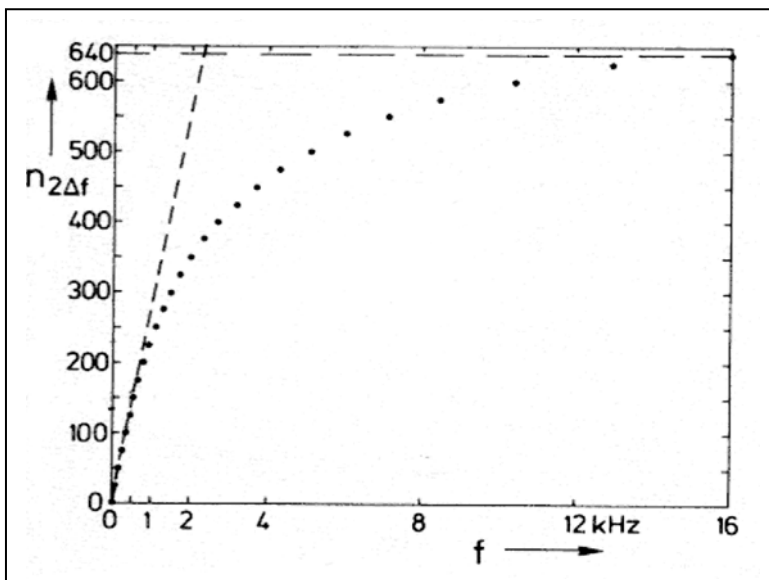
Abhängigkeit der Frequenzstufe $2\Delta f$ von der Frequenz f ($f_{\text{mod}} = 4 \text{ Hz}$, $L = 60 \text{ phon}$):



Vergleiche Funktion Frequenzgruppe Δf_G in Abhängigkeit von der Frequenz f !

Im folgenden Bild ist die Anzahl $n_{2\Delta f}$ der einzelnen Reizstufen dargestellt, die bis zu einer bestimmten Frequenz f aneinandergereiht werden können. Da die Reizstufen sehr klein sind, sind von Punkt zu Punkt jeweils 25 Reizstufen zusammengefasst. Von Null beginnend steigt die Zahl der Frequenzstufen zunächst proportional mit der Frequenz an. Oberhalb 500 Hz ist der Anstieg geringer und deutet auf ein logarithmisches Fortschreiten hin. Eine Oktave entspricht dann in etwa 100 Frequenzstufen. Bis zur oberen Frequenzgrenze von 16 kHz können 640 vom menschlichen Gehör unterscheidbare Frequenzstufen aneinandergereiht werden. Diese Auflösung ist sehr fein. Sie wird verständlich, wenn man berücksichtigt, dass etwa 3600 Sinneszellen längs des Cortischen Organs aneinandergereiht sind. Dort haben sie einen Abstand von nur $9 \mu\text{m}$. Einer Frequenzstufe würde ungefähr eine Verschiebung um 6 Sinneszellen entsprechen. Wegen der ähnlichen Kurvenformen hängen Frequenzgruppe und Frequenzstufe eng miteinander zusammen.

Anzahl $n_{2\Delta f}$ der bis zur Frequenz f aneinanderreihbaren Frequenzstufen. Von Punkt zu Punkt sind 25 Stufen zusammengefasst. Lineare Skalierung!



Verhältnistönhöhe:

Die Empfindungsfunktion der Tonhöhenempfindung kann direkt nur durch die Messung von Verhältniswerten bestimmt werden. Versuchspersonen erhalten dabei die Aufgabe, den Reiz (Frequenz) so zu verändern, dass im Vergleich zu einem Ausgangsreiz der doppelte oder auch der halbe Wert der Empfindung hervorgerufen wird.

Bei tieferen Frequenzen findet man wie in der Musik, dass sich die Frequenz, der die halbe Tonhöhe zugeordnet wird, und die Ausgangsfrequenz wie 0,5 zu 1 verhalten.

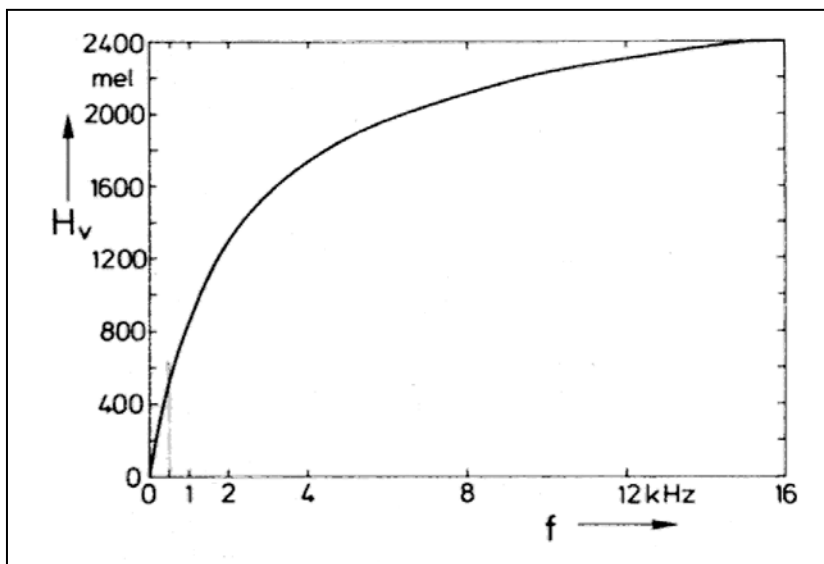
Bei hohen Frequenzen ist das Ergebnis jedoch überraschend: Wird beispielsweise eine Ausgangsfrequenz von 8 kHz gewählt, so geben die Versuchspersonen im Mittel für die halbe Tonhöhenempfindung eine Frequenz an, die bei etwa 1300 Hz liegt. Die Versuche zeigen also, dass von 220 Hz nach 440 Hz, aber auch von 1300 Hz nach 8000 Hz jeweils eine Verdoppelung der Tonhöhenempfindung stattfindet. Aus vielen Verhältniswertmessungen und Festlegen eines Fixpunktes kann die zugehörige Empfindungsfunktion gebildet werden. Als Fixpunkt wird dabei die Normalfrequenz 125 Hz aus der Oktavreihe gewählt.

Als Einheit für die Empfindungsgröße wurde das „mel“ gewählt, da die Verhältnistönhöhe eng mit dem melodischen Empfinden zusammenhängt. Man kann damit festhalten, dass ein reiner Ton der Frequenz 125 Hz die Verhältnistönhöhe 125 mel besitzt. Auch der 440-Hz-Ton (das eingestrichene a) besitzt eine Verhältnistönhöhe von 440 mel. Je höher die Frequenz wird, umso mehr weichen der Zahlenwert der Frequenz in Hz und der Zahlenwert der Verhältnistönhöhe in mel voneinander ab. Ein Ton der Frequenz 1300 Hz besitzt 1050 mel, ein 8-kHz-Ton muss dementsprechend 2100 mel entsprechen.

Das starke Abweichen der Verhältnistönhöhe von der Proportionalität zur Frequenz mag zunächst überraschen, weil man in der Musik mit Oktaven arbeitet. Wenn man jedoch auf dem Klavier im mittleren Frequenzbereich um das eingestrichene a eine Melodie spielt und dieselbe Melodie jetzt ganz rechts außen auf der Klaviatur bei hohen Frequenzen wiederholt, lernt man sofort einen mit der Verhältnistönhöhe zusammenhängenden Effekt kennen: Die Weite der in der Umgebung des eingestrichenen a gespielten Melodie ist viel größer als die Weite derselben Melodie, die auf der höchsten Oktave des Klaviers gespielt wird. Dabei wird also deutlich, dass bei extremer Tonlage das harmonische Empfinden und das melodische Empfinden voneinander divergieren. Die in der Notenschrift angegebenen Grundfrequenzen überschreiten nur selten 1500 Hz.

Unter 500 Hz ist die Verhältnistönhöhe in mel identisch mit der Frequenz in Hz, über 500 Hz verringert sich die Verhältnistönhöhe mehr und mehr. Bei 16 kHz erreicht sie einen maximalen Wert von 2400 mel.

Verhältnistönhöhe H_v in mel als Funktion der Frequenz f . Lineare Skalierung:



Beispiel: Ermitteln Sie die zu 12 kHz gehörige Frequenz halber Tonhöhenempfindung.

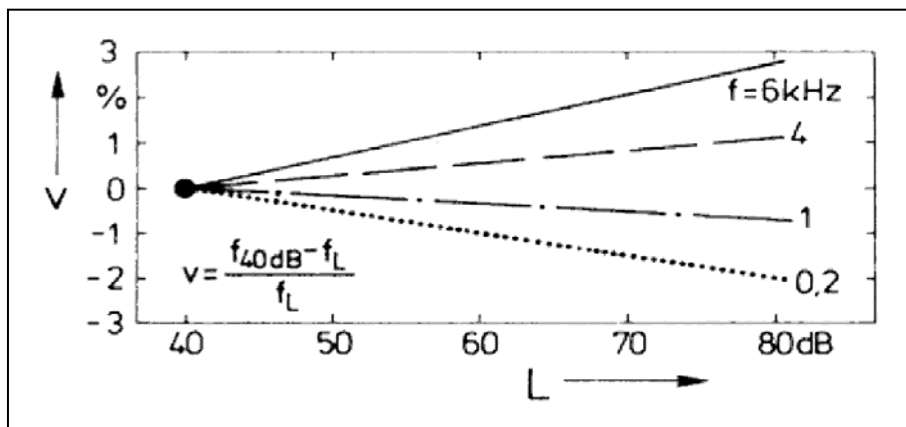
Spektrale Tonhöhe:

Die Tonhöhe eines Sinustones hängt nicht nur von der Frequenz, sondern auch von anderen Parametern wie z.B. dem Pegel ab. Durch Vergleich der Tonhöhen von Tönen verschiedenen Pegels kann dieser Effekt quantitativ bestimmt werden. Die Tonhöhe eines Tones mit dem Pegel L und der Frequenz f_L wird durch Vergleich mit der Tonhöhe gemessen, die ein Ton mit dem Pegel 40 dB und einer Frequenz f_{40dB} hervorruft. Dabei wird f_{40dB} so gewählt, dass gleicher Tonhöhen Eindruck bei den alternierend dargebotenen Tönen verschiedenen Pegels entsteht. Die Frequenz des Vergleichsschalls wird als Tonhöhenfrequenz bezeichnet, weil sie das Kriterium gleicher Tonhöhe erfüllt, aber in Einheiten der Frequenz angegeben wird.

Wird ein 200-Hz-Ton alternierend mit einem Pegel von 80 dB und dann mit einem Pegel von 40 dB dargeboten, so erzeugt der laute 200-Hz-Ton eine tiefere Tonhöhe als der leisere 200-Hz-Ton. Macht man dasselbe Experiment bei 6 kHz, so findet man ein umgekehrtes Verhalten. Der 6-kHz-Ton mit 80 dB erzeugt eine höhere Tonhöhenempfindung als der Ton gleicher Frequenz, jedoch mit dem Pegel von nur 40 dB.

Diese Pegelabhängigkeit ist jedoch sehr gering. Die für gleiche Tonhöhenempfindung notwendigen relativen Frequenzänderungen v erreichen für den Pegelbereich zwischen 40 und 80 dB Maximalwerte von nur 3 %.

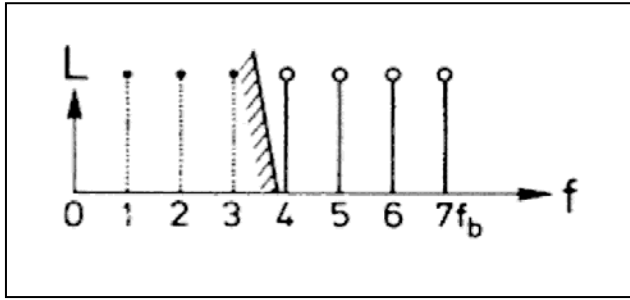
Relative Frequenzänderung v reiner Töne, die für vier Frequenzbeispiele (Parameter) nötig ist, um bei versch. Pegel L gleiche Tonhöhenempfindung zu erzeugen wie ein Ton mit $L = 40$ dB:



Virtuelle Tonhöhe:

Während den bisher beschriebenen Tonhöheneffekten stets eine physikalisch vorhandene Komponente (nämlich ein Sinuston bestimmter Frequenz) zugeordnet werden konnte, ist dies bei der Virtuellen Tonhöhe nicht mehr der Fall. Die Virtuelle Tonhöhe entsteht, wenn bei einem breitbandigen Linienspektrum nur ein höherfrequenter Teil übertragen wird. Dies ist z.B. beim Telefonieren der Fall, da nur Frequenzen zwischen 300 Hz und 3000 Hz übertragen werden. Dennoch kann das Ohr zwischen männlicher (Grundfrequenz: 100 Hz) und weiblicher (Grundfrequenz: 200 Hz) unterscheiden, obwohl physikalisch die zugehörige Frequenz nicht dargeboten wird. Offenbar sucht das Gehör aufgrund von Erfahrungen, die es beim Hören von Sprache und Musik gewonnen hat, aus den übertragenen Spektrallinien die zugehörige Grundtonhöhe. Das Gehör leitet diese Grundtonhöhe aus den spektralen Tonhöhen ab, so dass hier wieder die geringe Pegelabhängigkeit der Tonhöhenempfindung eine Rolle spielt und daher die Virtuelle Tonhöhe nicht vollständig identisch sein muss mit der physikalisch berechenbaren Grundfrequenz bzw. dem Spektrallinienabstand. Die Virtuelle Tonhöhe ist keine einzelne Empfindungsgröße, sondern eine komplexe Empfindung aus mehreren Empfindungsgrößen.

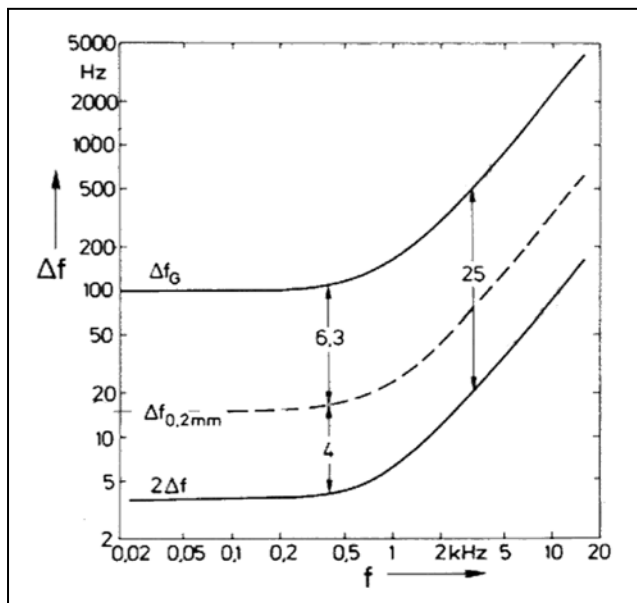
Erzeugung der Virtuellen Tonhöhe:



Zusammenhänge zwischen den Skalen der Tonhöhenempfindung:

Die Kurvenverläufe der Frequenzgruppenbreite Δf_G und der eben wahrnehmbaren Frequenzstufe $2\Delta f$ in Abhängigkeit von der Frequenz f gleichen einander im prinzipiellen Verlauf. Man kann nun wegen der Frequenz-Orts-Transformation eine Frequenzstufe $\Delta f_{0,2mm}$ festlegen, welche dadurch definiert ist, dass auf der Basilarmembrane die Abstände maximaler Auslenkung 0,2 mm entfernt sind. Werden alle drei Frequenzstufen (Frequenzgruppenbreite als Stufe betrachtet) in Abhängigkeit von der Frequenz f in eine Graphik eingetragen, so sind die Zusammenhänge gut erkennbar: Die Frequenzänderung $\Delta f_{0,2mm}$, die einer Veränderung des Ortes maximaler Auslenkung um 0,2 mm auf der Basilarmembrane entsprechen würde, ist viermal so groß wie die eben wahrnehmbare Frequenzstufe $2\Delta f$ und um den Faktor 6,3 kleiner als die Breite Δf_G einer Frequenzgruppe. Dies bedeutet, dass die eben wahrnehmbare Frequenzstufe $2\Delta f$ unabhängig vom Ort auf der Basilarmembrane 0,05 mm bzw. 5,6 Haarzellen ausmacht. Eine Frequenzgruppenbreite Δf_G entspricht einer örtlichen Breite auf der Basilarmembrane von etwa 1,3 mm bzw. umfasst 150 aneinandergereihte Haarzellen.

Frequenzgruppenbreite Δf_G , Frequenzstufe $2\Delta f$ und Frequenzänderung $\Delta f_{0,2mm}$ als Funktion der Frequenz f . Doppelpfeile geben die Faktoren an, um welche die Kurven gegeneinander verschoben sind:



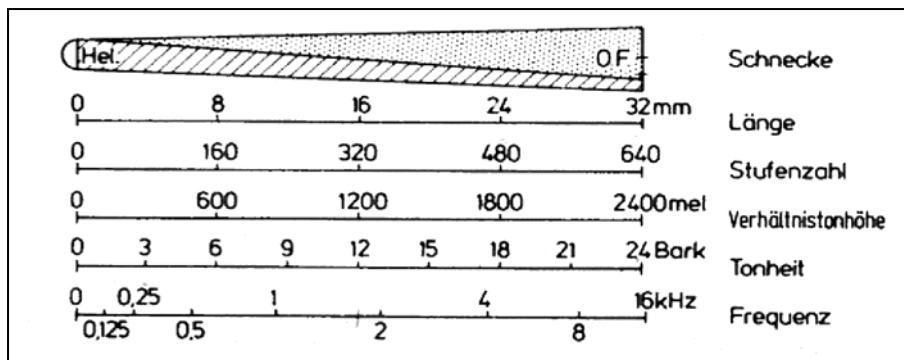
Nimmt man die anatomischen Daten der Haarzellen (Abstand von Haarzelle zu Haarzelle etwa 9 μm , 3600 Haarzellen auf der gesamten Länge der Basilarmembrane) zu Hilfe und berücksichtigt außerdem, dass die Verhältnistönhöhe die gleiche Funktion ergeben hat wie die Aufsummierung der Frequenzgruppenbreiten bzw. die Aufsummierung der Frequenzstufen, so können die in der Tabelle dargestellten sehr wichtigen Beziehungen angegeben werden.

Skalen der Basilarmembrane:

<u>Länge BM</u> \cong	24 Bark \cong	32 mm \cong	640 Stufen \cong	2400 mel \cong	3600 Haarzellen
	<u>1 Bark</u> \cong	1,3 mm \cong	27 Stufen \cong	100 mel \cong	150 Haarzellen
	0,7 Bark \cong	<u>1 mm</u> \cong	20 Stufen \cong	75 mel \cong	110 Haarzellen
	0,04 Bark \cong	50 μ m \cong	<u>1 Stufe</u> \cong	3,8 mel \cong	5,6 Haarzellen
	0,01 Bark \cong	13 μ m \cong	0,26 Stufen \cong	<u>1 mel</u> \cong	1,5 Haarzellen
	0,007 Bark \cong	9 μ m \cong	0,18 Stufen \cong	0,7 mel \cong	<u>1 Haarzelle</u>

Wird das Innenohr in seine Gesamtlänge aufgerollt, so können die Zusammenhänge über die Länge der Basilarmembrane wie folgt dargestellt werden. Die Basilarmembrane besitzt vom Helicotrema (tiefe Frequenzen) bis zum ovalen Fenster (hohe Frequenzen) eine Gesamtlänge von 32 mm, was in der ersten Skale dargestellt ist. Die zweite Skale zeigt die Frequenzstufenanzahl, von der insgesamt 640 gleichmäßig auf der Basilarmembrane verteilt sind. In der dritten Skale ist linear die bis 2400 mel reichende Verhältnistonhöhe aufgeführt. Die vierte Skale zeigt die 24 Bark umfassende lineare Tonheitsskalierung. Als einzige nichtlineare Skale ist schließlich die Frequenz aufgetragen. Unter 500 Hz ist auch diese Skale linear und wird über 500 Hz logarithmisch. In der Mitte der Basilarmembrane ist eine Frequenz von 1800 Hz (= 12 Bark) anzusiedeln. Eine rein physikalische Darstellung wie die Frequenzskale ist daher für die Beschreibung der Vorgänge im Innenohr nicht geeignet.

Skalen der Tonhöhe transformiert auf die Länge der Basilarmembrane. Alle Skalen sind bis auf die der Frequenz linear geteilt:



Intervalle:

Bei der Wahrnehmung der Tonhöhe wird nicht nur jede Verdoppelung der Grundfrequenz eines Klanges stets als Oktave wahrgenommen, sondern es wird auch ganz allgemein jede Multiplikation mit einem konstanten Faktor als gleich bleibender Abstand empfunden. Feststehende Frequenzverhältnisse werden als Intervalle bezeichnet und kennzeichnen einen bestimmten, stets als konstant empfundenen Abstand der Tonhöhe zweier Klänge.

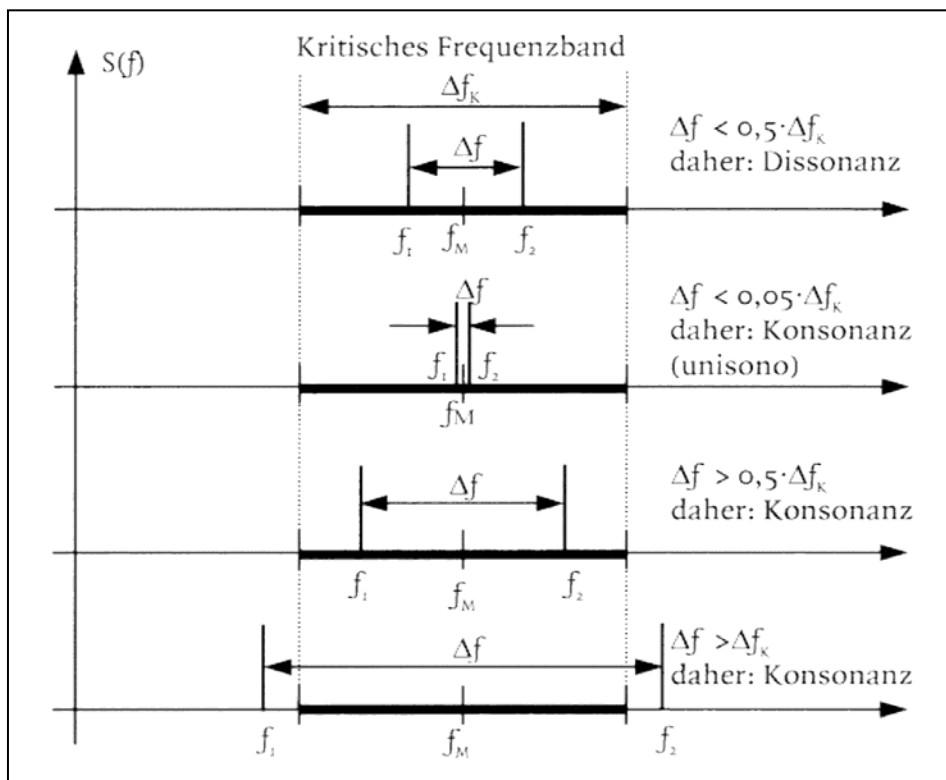
Die Oktave ist das Intervall mit dem einfachsten Frequenzverhältnis von 1:2 und sie entspricht auch dem Verhältnis von Grundton zu erstem Teilton eines Klanges. Jene Intervalle, die in der westlich geprägten Musik von Bedeutung sind, ergeben sich aus den Frequenzverhältnissen der weiteren Teiltöne eines Klanges. Da diese Teiltonfrequenzen stets Vielfache der Grundfrequenz sind, müssen Grundfrequenzen und erste Teiltonfrequenz im Verhältnis 1:2, erste und zweite Teiltonfrequenz im Verhältnis 2:3, zweite und dritte Teiltonfrequenz im Verhältnis 3:4 usw. stehen.

Intervalle und ihr Frequenzverhältnis:

Intervall	Oktave	Quint	Quart	Große Terz	Kleine Terz	Große Sekunde	Große Sexte	Kleine Sexte
Frequenzverhältnis	1 : 2	2 : 3	3 : 4	4 : 5	5 : 6	8 : 9	5 : 3	8 : 5

Intervalle können als konsonant und dissonant bezeichnet werden. Konsonante Intervalle werden als wohlklingend, glatt und angenehm, dissonante hingegen tendenziell als spannungsvoll, rau, eher unangenehm und „auseinander klingend“ empfunden. Diese Einteilung ist für das Audiodesign insofern von Relevanz, als durch das Wechselspiel von Intervallqualitäten formale und dramaturgische Gestaltung möglich wird. Der Übergang zwischen Konsonanz und Dissonanz ist fließend und stark von Kultur und individueller Hörerfahrung geprägt. Schon Pythagoras stellte die sog. Proportionstheorie auf, dass ein Intervall umso konsonanter empfunden wird, je einfacher sein Frequenzverhältnis ist. Nach der Klangverwandtschaftstheorie von Helmholtz ist ein Intervall dann konsonant, wenn ein oder mehrere Obertöne der beiden Klänge zusammenfallen. Neuer Erkenntnisse über das Empfinden von Konsonanz und Dissonanz zeigen, dass zwei reine Sinussignale dann als dissonant wahrgenommen werden, wenn sie innerhalb eines kritischen Frequenzbandes der Basilmembrane liegen und ihr Frequenzunterschied zwischen 5 und 50 Prozent dieses Frequenzbandes beträgt. Die Einteilung in konsonante und dissonante Intervalle ist daher auch von der absoluten Tonhöhe abhängig.

Wahrnehmung von Konsonanz und Dissonanz bei Sinussignalen ($\Delta f_G = \Delta f_K$):



Bei zwei komplexen Klängen muss untersucht werden, wie viele Teiltöne des höheren Klanges in den kritischen Frequenzbändern der ersten sieben Teiltöne des tieferen Klanges liegen. Je mehr in ein Frequenzband fallende Teiltöne nach obiger Erklärung als dissonant gelten, desto dissonanter wird das Intervall empfunden. Das Empfinden von Konsonanz und Dissonanz hängt daher auch davon ab, welche Teiltöne eines Klanges wie stark ausgeprägt sind. Bei der Instrumentation und beim Arrangieren von Musik muss daher beachtet werden, dass manche Klangkombination besser und andere schlechter zusammenklingen.

Beispiel: Klingen 2 kHz und 2,1 kHz konsonant?

3.8 Schwankungsstärke

Ändert sich die die Amplitude (Modulationsgrad m) oder die Frequenz (Modulationsfrequenz f_{mod}) nur langsam, so kann das Gehör der Schwankung von empfundener Lautstärke und Tonhöhe noch folgen. Man spricht von der Empfindung Schwankungsstärke. Beispiele sind die musikalischen Klangeffekte Vibrato und Tremolo, die technisch durch Amplituden- oder Frequenzmodulation erzeugt werden können und vom Menschen auch richtig als Änderung der Lautheit bzw. der Amplitude empfunden werden. Derartige Effekte werden z.B. als Klappern, Rattern, Schlagen, Rumpeln oder Flattern bezeichnet. Auch Schwebungen, die durch Überlagerung zweier harmonischer Schwingungen entstehen, können durch die Schwankungsstärke erfasst werden, solange die von der Differenz der beiden Schwingungen bestimmte Schwebungsfrequenz unterhalb von 20 Hz liegt. Attribute wie pulsierend, brodelnd, hämmernd oder zwitschernd beruhen maßgeblich auf der Schwankungsstärke. Die max. Empfindlichkeit des Gehörs ist bei einer Schwankung um ca. 4 Hz zu finden, welche auch bei Sprache der mittleren Silbenfrequenz entspricht. Die Schwankungsstärke ist abhängig von Modulationsfrequenz, Modulationsgrad, Schallpegel und Frequenzlage. **Sie gibt also die Auffälligkeit von langsamen zeitlichen Schwankungen im Signal an.** Ihr Zeichen ist F , die Einheit vacil. 1 vacil entspricht der Schwankungsstärke eines 1-kHz-Tones mit 60 dB, Modulationsfrequenz 4 Hz und Modulationsgrad 1.

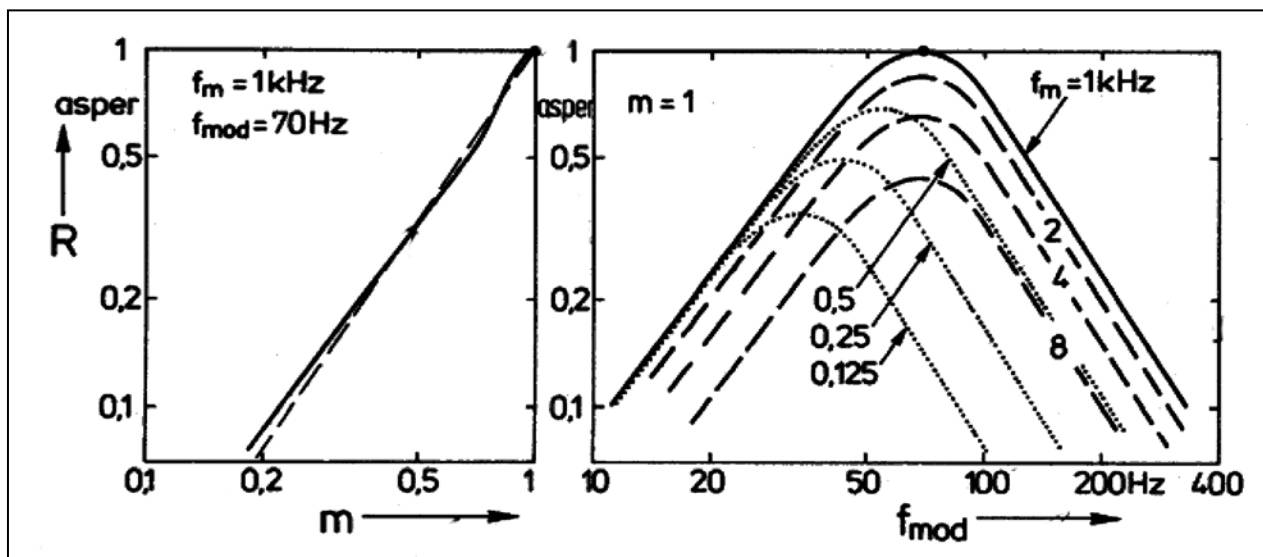
$$F, 1 \text{ vacil: } f = 1 \text{ kHz, } L = 60 \text{ dB, } f_{mod} = 4 \text{ Hz, } m = 1$$

3.9 Rauhigkeit

Ändern sich Amplitude und Frequenz zu rasch, um als Schwankungen von Lautstärke oder Tonhöhe empfunden zu werden, so führen sie zur Wahrnehmung von Rauhigkeit. **Die Rauhigkeit gibt also die Auffälligkeit von schnellen zeitlichen Veränderungen im Signal an**, z.B. die Überlagerung zweier harmonischer Töne, deren Frequenzdifferenz zwar kleiner als eine Frequenzgruppe aber größer als 20 Hz ist. Die Rauhigkeit ist bei einer Schwankungsfrequenz von 70 Hz am größten, die Seitenlinien der Amplitudenmodulation werden dabei noch nicht getrennt hörbar. Zur Festlegung der Empfindungsfunktion für die Rauhigkeit R wird einem 1-kHz-Ton, der mit $m = 1$ und $f_{mod} = 70 \text{ Hz}$ sinusförmig in der Amplitude moduliert ist und einen Pegel von 60 dB besitzt, die Rauhigkeit von 1 asper zugeordnet. Derartige Effekte werden als Brummen, Krächzen, Knirschen und Surren beschrieben.

$$R, 1 \text{ asper: } f = 1 \text{ kHz, } L = 60 \text{ dB, } f_{mod} = 70 \text{ Hz, } m = 1$$

Rauhigkeit R eines sinusförmig amplitudenmodulierten Tones abhängig vom Modulationsgrad m für eine Trägerfrequenz von 1 kHz und eine Modulationsfrequenz von 70 Hz (links) bzw. abhängig von der Modulationsfrequenz f_{mod} mit der Trägerfrequenz f_m als Parameter (rechts):



Dadurch, dass alle Tonparameter bei hohem Rauigkeitseindruck zur Definition von 1 asper herangezogen wurden, stellt 1 asper eine hohe Rauigkeit dar, die nur noch von stark amplitudenmoduliertem Rauschen (4 asper) überschritten wird. Eine Veränderung des Modulationsgrades m um den Faktor 2 auf 0,5 bewirkt beispielsweise bereits eine starke Reduzierung der Rauigkeit auf 0,3 asper. Wird der Modulationsgrad nochmals um Faktor 2 auf 0,25 reduziert, so erreicht die Rauigkeit bereits so kleine Werte, dass ein Teil der Versuchspersonen diesen Schall als glatt und nicht mehr als rau bezeichnet. Der Rauigkeitseindruck geht dann in einen Ton über.

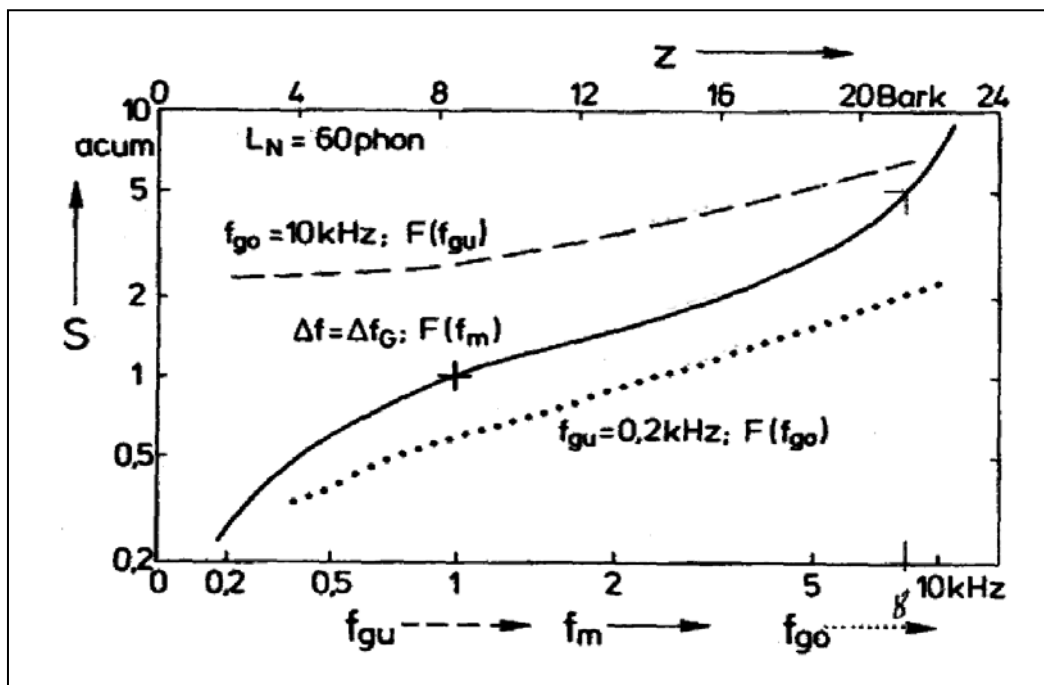
Für verschiedene Trägerfrequenzen f_m hängt die Rauigkeit eines amplitudenmodulierten Sinustones stark von der Modulationsfrequenz f_{mod} ab. Ein modulierter 1-kHz-Ton erreicht die größte Rauigkeit, bei tieferen und höheren Trägerfrequenzen f_m nimmt die Rauigkeit ab. Bei tieferen Trägerfrequenzen verschiebt sich das Maximum zu tieferen Modulationsfrequenzen f_{mod} .

3.10 Schärfe

Untersuchungen, die über die Entstehung der Klangfarbe durchgeführt wurden, haben gezeigt, dass die Schärfe ein wesentlicher Teil der Klangfarbe ist. Sie kann als eine Empfindungsgröße angesehen werden, auf die wir getrennt achten können. Als Hörattribute verbergen sich dahinter Begriffe wie Dichte, Härte, Helligkeit und Brillanz. Es gibt Schalle, die man als schrill, hart oder scharf klingend bezeichnet und andererseits solche, die als sanft oder ausgeglichen empfunden werden. Scharf klingende Schalle sind Quietschen einer Bremse oder Geräusch einer Kreissäge. Die Schärfe S wird stark von der spektralen Zusammensetzung, insbesondere dem Höhenanteil geprägt. Der Schärfe eines Schmalbandrauschens ($\Delta f \leq \Delta f_G$) mit der Mittenfrequenz 1 kHz und einem Pegel von 60 dB wird die Schärfe von 1 acum zugeordnet.

$$S, 1 \text{ acum: } f_m = 1 \text{ kHz, } L = 60 \text{ dB, } \Delta f \leq \Delta f_G$$

Schärfe S von SBR (durchgezogen), von TP-Rauschen (punktirt) und von HP-Rauschen (gestrichelt) abh. von Mittenfrequenz f_m bzw. oberer Grenzfrequenz f_{go} bzw. unterer Grenzfrequenz f_{gu} :



Demnach wächst die Schärfe mit ansteigender Mittenfrequenz des Schmalbandrauschens kontinuierlich, dies bei kleinen Frequenzen weniger als bei hohen Frequenzen. Die Schärfe eines Schmalbandrauschens bei 8 kHz ist etwa fünfmal größer als diejenige eines Schmalbandrauschens bei 1 kHz. Von tiefen nach hohen Mittenfrequenzen wächst die Schärfe von Schmalbandrauschen fast um den Faktor 20.

Beispiel: Welches Schärfeverhältnis hat ein HPR zu einem TPR mit 2 kHz Eckfrequenz?

3.11 Klangfarbe

Es besteht ein sehr enger Zusammenhang zwischen Klangfarbe und dem Spektrum des Signals. Weißes Rauschen beinhaltet beispielsweise alle Frequenzen gleichermaßen und weist keine Klangfarbe auf. Erst durch Filterung oder Hervorheben einzelner spektraler Anteile entsteht eine Klangfarbe. In Analogie zur Optik wird weißes Licht erst durch spektrale Anteile farbig und bunt. Der Farbe des Lichtes ist daher die Tonhöhe analog, der Buntheit die Klangfarbe. Ebenso wie monochromatisches Licht eines Lasers die reinste und ausgeprägteste Farbe hat, repräsentiert ein Sinuston die reinste und ausgeprägteste Klangfärbung.

Ganz wesentlich für die Klangfarbe sind die Formanten, also Frequenzbereiche, die aufgrund der Schwingungseigenschaften eines Resonanzkörpers im Spektrum verstärkt werden. Da diese Bereiche völlig unabhängig von der Grundfrequenz bzw. der wahrnehmbaren Tonhöhe sind, können sie als wesentliches Merkmal der Klangfarbe angesehen werden und zu deren Beschreibung verwendet werden. Aber auch der zeitliche Verlauf eines Signals trägt zu Klangfarbe bei, so z.B. der Anzupfvorgang eines Saiteninstrumentes. Wird bei einem Instrumentalklang die Einschwingphase weggeschnitten, so können selbst erfahrene Musiker das Schallereignis nicht mehr zweifelsfrei einem Instrument zuordnen.

Im Gegensatz zur Lautheit und Tonhöhe ist die Klangfarbe keine eindimensionale Eigenschaft und lässt sich daher nicht auf einer Skala zwischen laut und leise oder hoch und tief erfassen. Wichtige psychoakustische Parameter, die zum Parameter Klangfarbe zusammenwirken, sind Schwankungsstärke, Rauigkeit, Volumen, Dichte und Schärfe. Tonhöhe und Klangfarbe sind eng miteinander verbunden, in gewisser Hinsicht sogar identisch.

Kurzübersicht der Psychoakustikparameter:

Parameter	Beschreibung	Definition	Zeichen	Einheit	übliche Skalierung
Lautheit:	Hörkurven, Frequenzgruppenbildung, spektrale Verdeckungen, zeitliche Verdeckungen,	1-kHz-Ton mit 40 dB => 1 sone, linearer Zusammenhang zwischen Lautheitsempfindung und sone-Wert	N	sones	[N] = sones
Schärfe:	Anteil an hohen Frequenzen	frequenzgruppenbreites Schmalbandrauschen für Mittenfrequenz 1 kHz, 60 dB => 1 acum (acum = scharf)	S	acum	[S] = deziacum
Schwankungsstärke:	langsame Schwankungen der spezifischen Lautheit (bis ca. 10 Hz Modulationsfrequenz)	1-kHz-Ton 60 dB mit 4 Hz moduliert Modulationsgrad 1 => 1 vacil (vacillare = schwanken)	F	vacil	[F] = centivacil
Rauhigkeit:	schnelle Schwankungen der spezifischen Lautheit (ca. 10 Hz bis ca. 250 Hz Modulationsfrequenz)	1-kHz-Ton 60 dB mit 70 Hz moduliert Modulationsgrad 1 => 1 asper (asper = rau, uneben)	R	asper	[R] = centiasper
Tonhaltigkeit:	Ausgeprägtheit von einzelnen Tönen in einem Signalgemisch	1-kHz-Ton, 60 dB => 100 %	AdT	%	[AdT] = %

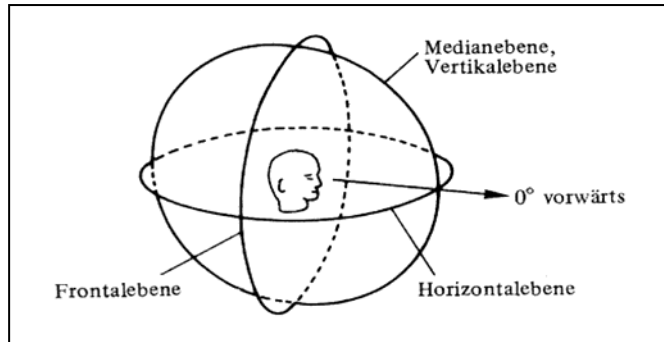
3.12 Räumliche Wahrnehmung

Um einschätzen zu können, wo sich - relativ zum Hörer – eine Schallquelle befindet, muss das Gehör aus dem ankommenden Signal Informationen über Entfernung und Richtung gewinnen können.

Das menschliche Gehör kann vor allem die räumliche Tiefe – also nah und fern – und die räumliche Richtung – also links und rechts – unterscheiden. Etwas schwieriger ist die Unterscheidung zwischen oben und unten, vor allem dann, wenn entsprechende Vergleichsmöglichkeiten fehlen.

Richtungswahrnehmung ist die Zuordnung der vom Gehör zum Bewusstsein gebrachten Hörereignisrichtung zur Schallereignisrichtung. Zusammen mit der Entfernungswahrnehmung bildet sie die räumliche Wahrnehmung des Gehörs. Richtungs- und Entfernungswahrnehmung schließen auch die Wahrnehmung der räumlichen Ausdehnung einer Schallquelle ein. Die sog. Räumlichkeit einer Schallquelle ist nicht ihre tatsächliche Ausdehnung, sondern die Empfindung, dass der Raum um die Schallquelle mit Schall erfüllt ist, die Räumlichkeit nimmt mit der Lautstärke und mit der Stärke seitlicher Reflexionen zu.

Kopfbezogenes Koordinatensystem:



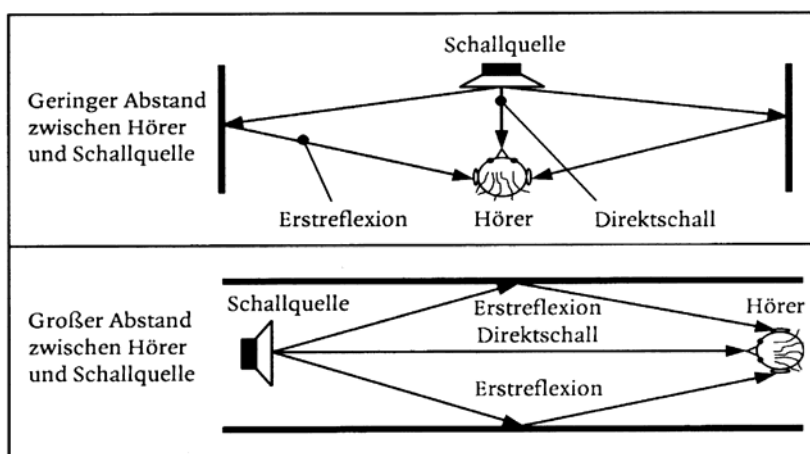
3.12.1 Wahrnehmung der räumlichen Tiefe

Eine Entfernungsbestimmung einfach aus der Lautstärke des Schallsignals könnte aus der Erkenntnis gewonnen werden, dass bei kugelförmiger Schallausbreitung die Schallintensität mit dem Quadrat des Abstandes bzw. der Schalldruck linear mit dem Abstand abnimmt. Dies scheidet jedoch daran, dass die Ausgangslautstärke des Signals nicht bekannt ist.

Wichtig für die Wahrnehmung der Entfernung von der Schallquelle sind daher sowohl das Pegelverhältnis von Direktschall und Erstreflexionen bzw. Erstreflexionen und Nachhall als auch die Verzögerungszeit zwischen Direktschall und Erstreflexionen bzw. Erstreflexionen und Nachhall (siehe 8.1 Raumakustik). Ist der Abstand zwischen Schallquelle und Hörer nur gering, so kommt der Direktschall mit nur geringer Abschwächung und Verzögerung beim Hörer an. Erstreflexionen und Nachhall müssen hingegen zusätzlich zu den Dämpfungen bei der Reflexion im Vergleich zum Direktschall wesentlich längere Wegstrecken zurücklegen. Ihre Verzögerungen sind daher vergleichsweise groß und ihre Pegel gering. Je schwächer also die Erstreflexionen im Vergleich zum Direktschall sind, desto näher wird das Schallsignal empfunden.

Ist hingegen der Hörer weit von der Schallquelle entfernt, so trifft auch schon der Direktschall mit größerer Verzögerung und abgeschwächtem Schallpegel ein. Die Wegstrecken, die die Erstreflexionen zurücklegen müssen, werden nur unwesentlich länger sein. Verzögerung und Pegelunterschied zwischen Direktschall und Erstreflexion sind also im Vergleich zu einer nahen Schallquelle gering.

Wahrnehmung der räumlichen Tiefe:



3.12.2 Wahrnehmung der räumlichen Richtung

Die Lokalisationsunschärfe, d.h. die Änderung in der Schallquellenrichtung, die nötig ist, um die Hörempfindungsrichtung zu ändern, beträgt bei frontaler Beschallung etwa 3° , entsprechend $\tau \approx 30 \mu\text{s}$ (entspricht Schallwegunterschied von ca. 1 cm), bei seitlicher Beschallung ergeben sich etwa 10° . Für stereophone Übertragungsanlagen bedeutet dies, dass in beiden Kanälen zur richtungstreuen Wiedergabe Laufzeitfehler unter $30 \mu\text{s}$ Laufzeitdifferenz liegen sollten.

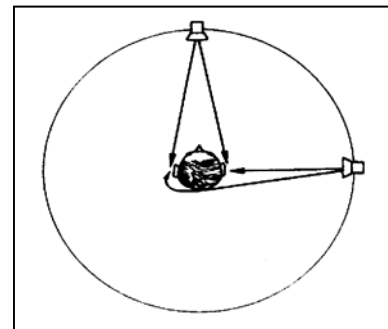
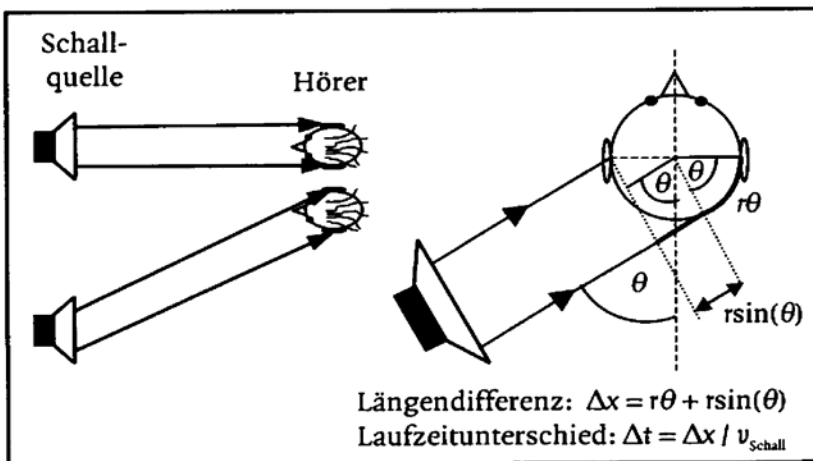
Tiefe Frequenzen unter 100 Hz sind vom Gehör nicht ortbar, so dass entsprechend tief abgestimmte Subwoofer theoretisch hinsichtlich Ortbarkeit im Raum beliebig aufgestellt werden können.

a, Laufzeitunterschied:

Für die Bestimmung der seitlichen Position einer Schallquelle wird zuerst der Laufzeitunterschied der **ersten Wellenfront** zwischen linkem und rechtem Ohr ausgewertet. Eine von rechts kommende Schallwelle muss einen um etwa 20 cm längeren Weg (um den Kopf herum) zurücklegen, um zum linken Ohr zu gelangen. Daraus ergibt sich ein Laufzeitunterschied τ_{max} von etwa 0,6 ms. Ob sich die Quelle vorne oder hinten bzw. oben oder unten befindet, lässt sich aufgrund des Laufzeitunterschiedes nicht ermitteln. Der Laufzeitunterschied kann nur für die erste beim Hörer eintreffende Wellenfront ausgewertet werden und ist insbesondere im Falle von percussiven, kurzen und somit impulsartigen Schallereignissen von Bedeutung. Die Wanderung der Lokalisation aus der Mitte erfolgt linear mit der interauralen Zeitdifferenz.

Bei dauerhaften Signalen wird der **Phasenunterschied** $\Delta\phi = 2\pi f \Delta t$ zwischen den beiden Ohren für Frequenzen **kleiner als 750 Hz** ausgewertet. Der Phasenunterschied ergibt sich analog zum Laufzeitunterschied dadurch, dass die Schallquelle von einem Ohr weiter entfernt ist als vom anderen. Die Phasendifferenz $\Delta\phi$ darf einen Winkel von 180° nicht übersteigen, da andernfalls das Ergebnis nicht mehr eindeutig interpretiert werden kann. Für hohe Frequenzen werden die Phasenunterschiede jedoch so gering, dass sie von den Hörnerven nicht mehr ausgewertet werden können.

Laufzeitunterschied:

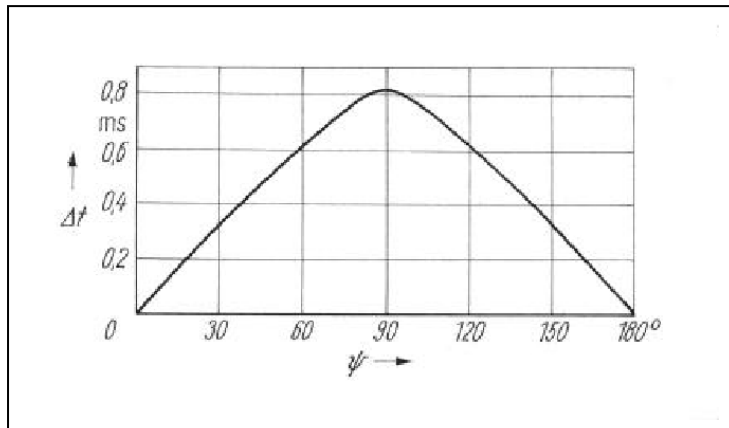


Die für die Lokalisation maßgebliche Größe ist also die Einfallrichtung der ersten Wellenfront (**Precedence- oder Haas-Effekt**), d.h. der vom Sprecher bzw. Lautsprecher ausgehende Direktschall. Spätere Schallanteile (Lautsprecherdirektschall, Reflexionen) beeinflussen die Lokalisation nicht, wenn sie etwa 10-30 ms nach dem Direktschall eintreffen und ihr Pegel denjenigen des Direktschalls um nicht mehr als 10 dB überschreitet.

Der Haas-Effekt findet praktische Bedeutung beim Einsatz von Stützlautsprechern, die Sprache verzögert abstrahlen, so dass bis $t_v = 50 \text{ ms}$ der Sprecher bzw. Hauptlautsprecher als Schallquelle geortet wird. Laufzeiten von 50 bis 100 ms führen zu einem verwischten Schalleindruck, ab etwa

100 ms werden separate Echos als Schallwiederholungen gehört, sofern sie ausreichenden Pegel aufweisen und nicht von Hall- oder Störgeräuschen verdeckt werden. Durch eine Höhenabsenkung beim Echo (poröse Wandabsorber) wird die Störwirkung reduziert.

Interauraler Laufzeitunterschied Δt in Abhängigkeit vom Schalleinfallswinkel:



b, Intensitätsunterschied:

Ferner wird der interaurale Intensitätsunterschied, das ist der Pegelunterschied den ein Schallergebnis zwischen linkem und rechtem Ohr hervorruft, ausgewertet. Für diesen Pegelunterschied ist weniger der Abstand zwischen den beiden Ohren verantwortlich, als vielmehr die Tatsache, dass der Kopf als Schallabsorber wirkt. Im Falle tiefer Frequenzen ist der Kopf klein im Verhältnis zur Wellenlänge und stellt somit kein nennenswertes Hindernis für die eintreffende Schallwelle dar. Der Intensitätsunterschied gewinnt daher erst **ab Frequenzen von etwa 1,6 kHz** an Bedeutung. Intensitätsunterschiede tragen aber insgesamt betrachtet weniger zur Lokalisationsbestimmung einer Quelle bei als Laufzeitunterschiede.

Frequenzabhängigkeit der Abschattungsämpfung a_A bei genau seitlicher Beschallung:

f =	100	200	500	1k	2k	5k	10k	Hz
a_A =	0	1	3	6	8	20	10-20	dB

Vor allem hohe Frequenzanteile werden vom Außenohr richtungsabhängig gebündelt und verstärkt. Die relative Stärke hochfrequenter Schallanteile scheint sich daher zu ändern, wenn eine Schallquelle von hinten nach vorne wandert. Außerdem werden i.a. hohe Frequenzen stärker gedämpft als tiefe, weshalb weiter entfernte Schallereignisse meist dumpfer klingen.

Zusätzlich werden auch Bewegungen des Kopfes ausgeführt, um zu hören, wie sich dabei der Schall ändert. Kurze perkussive Klänge sind deshalb in der Regel schwerer zu lokalisieren, als lang andauernde Schallereignisse. Beim Tragen von Kopfhörern fällt diese Möglichkeit der Raumwahrnehmung weg, weshalb der Raumeindruck in diesem Fall auch immer etwas verfälscht ist.

Viele audioteknische Geräte ermöglichen die Einstellung der räumlichen Richtung mit Hilfe eines sogenannten Panorama-Reglers. Die Panoramastellung bestimmt die Pegeldifferenz zwischen den beiden Stereokanälen. Es wird also auf diese Weise ausschließlich ein Intensitäts- und kein Laufzeit- bzw. Phasenunterschied erzeugt, weshalb die Panorama-Regelung nur für den höheren Frequenzbereich wirksam werden kann.

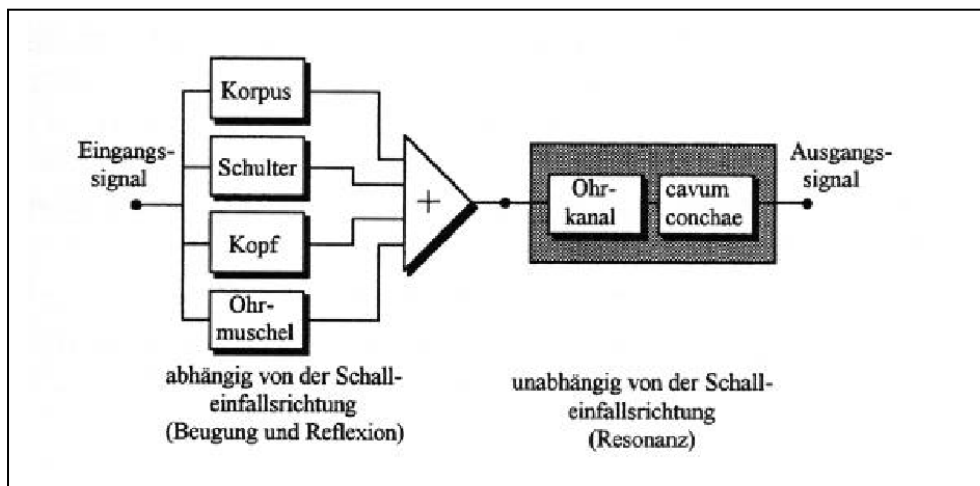
c, Spektrale Unterschiede:

Auch in der vertikalen Ebene, der Medianebene, ist eine Lokalisierung der Schallquelle möglich. Durch die Form und Beschaffenheit von Kopf und Ohren werden je nach Einfallsrichtung richtungsbestimmende Frequenzbänder angehoben, so dass aufgrund von Erfahrungswerten bei breitbandigen Signalen eine Richtungsbestimmung möglich ist. Weißes Rauschen kann hierbei im günstigsten Fall mit einer Lokalisationsunschärfe von $\pm 4^\circ$ bestimmt werden. Das Richtungshören in der Medianebene gewinnt besondere Bedeutung beim Übertragungsverfahren der Kunstkopfstereophonie.

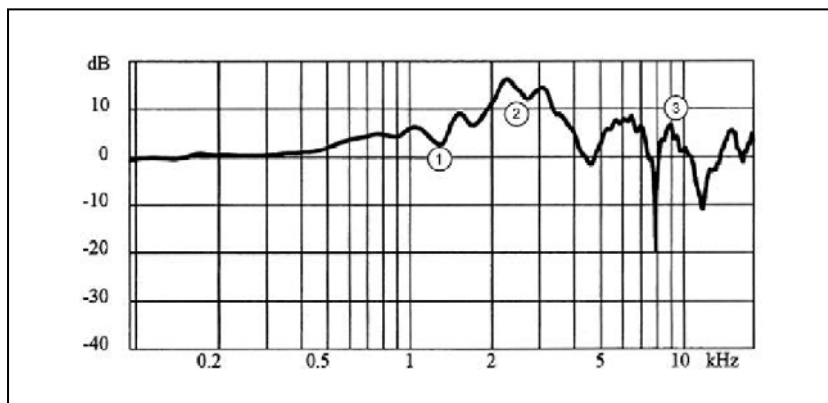
Die Unterscheidung, ob eine Schallquelle vor oder hinter einem Hörer positioniert wurde, ist aufgrund der Form des Außenohres und der Ohrmuschel möglich, die bewirken, dass vor allem Frequenzen oberhalb von 5 kHz richtungsabhängig gebündelt und verstärkt bzw. abgeschwächt werden. Wandert also eine Schallquelle von vorne nach hinten, so scheint sich ihr Signalanteil im hohen Frequenzbereich nach und nach zu verringern. Die Unterscheidung zwischen hinten und vorne ist dadurch zumindest in einem beschränkten Ausmaß möglich. Um eine Schallquelle genauer lokalisieren zu können, sind geringfügige Kopfbewegungen notwendig.

Die hierfür relevante, frequenzabhängige Außenohrübertragungsfunktion wird als HRTF (Head Related Transfer Function) bezeichnet.

Beiträge von richtungsabhängigen und richtungsunabhängigen Komponenten zur Außenohrübertragungsfunktion:

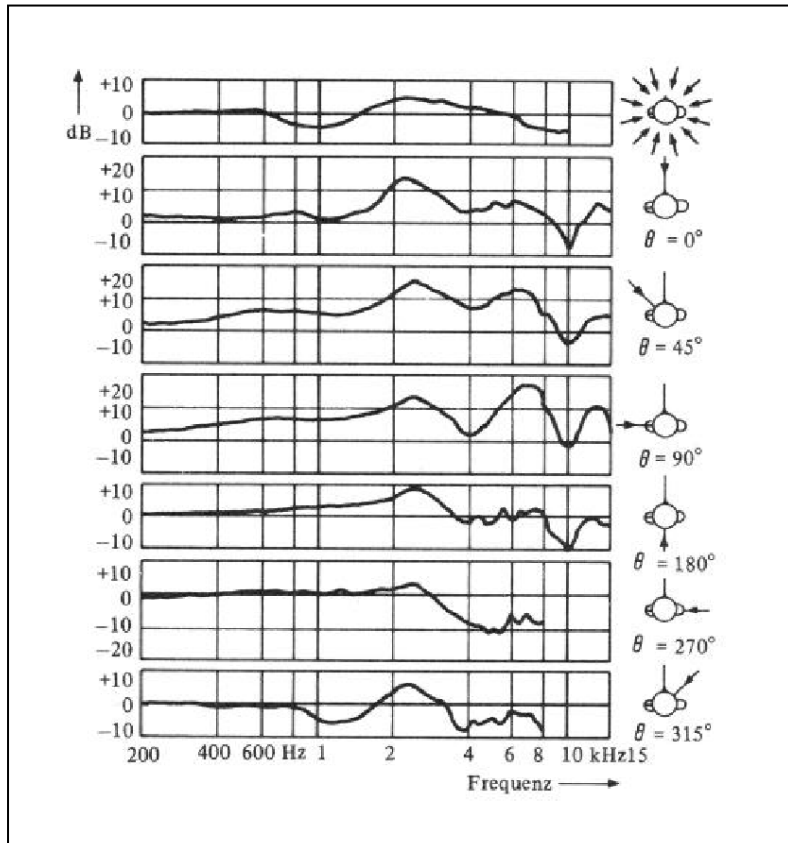


Typische HRTF mit den Charakteristika:



- 1, Reflexion an der Schulter
- 2, $\lambda/4$ -Resonanz des Ohrkanals
- 3, Ohrmuschel-Interferenzen

HRTF in Abhängigkeit vom horizontalen Einfallswinkel:



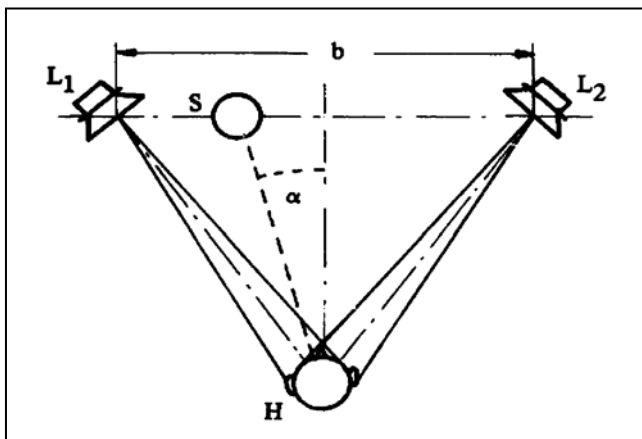
Zusammenfassung:

- links/rechts: Laufzeitunterschiede, erste Wellenfront (Phasenunterschiede $f < 750$ Hz)
Intensitätsunterschiede $f > 1,6$ kHz
- oben/unten: Spektrale Unterschiede, HRTF
- vorne/hinten: Spektrale Unterschiede, HRTF
- tief/weit: Zeitl. Abstand Direktschall - Erstreflexion

3.12.3 Phantomquelle

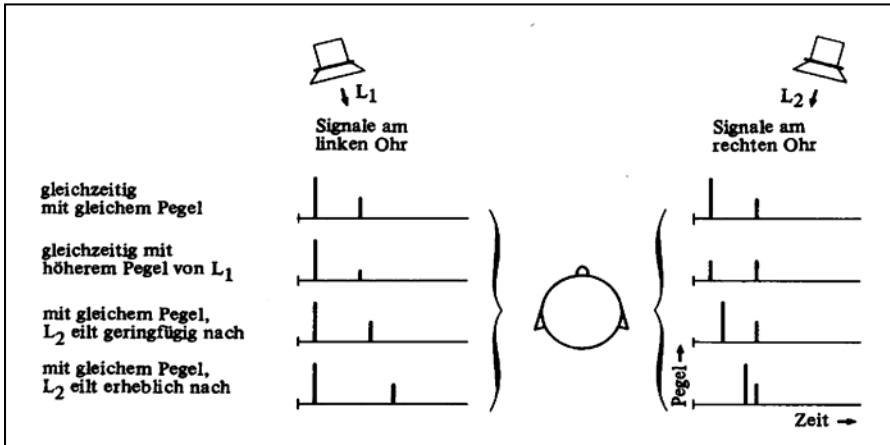
Während bei der Lokalisation einer realen Schallquelle der Richtungseindruck aus einem einzigen Schallsignal gewonnen wird, überlagern sich bei der stereophonen Lautsprecherwiedergabe zwei Schallsignale. Hierbei erhält jedes Ohr von jedem der beiden Lautsprecher einen bestimmten Schallanteil, woraus der Richtungseindruck einer fiktiven Schallquelle entsteht.

Lautsprecheranordnung für Stereowiedergabe:



Voraussetzung für eine einwandfreie Lokalisation ist, dass die Lautsprecher von derselben Schallquelle stammende Signale ohne Phasenumkehrung abstrahlen, dass die Pegel- und/oder Laufzeitdifferenzen innerhalb bestimmter Grenzen bleiben und dass sich der Hörer in einer bestimmten geometrischen Anordnung zu den Lautsprechern befindet, nämlich innerhalb der Stereohörfläche. Hierbei erhält jedes Ohr von jedem der beiden Lautsprecher einen bestimmten Schallanteil, woraus der Richtungseindruck einer fiktiven Schallquelle entsteht.

Signalkonfigurationen an den Ohren bei Stereowiedergabe:



Bei Zustandekommen der Phantomschallquelle ist Voraussetzung, dass sich der Hörer auf der Mittelachse zwischen den beiden Lautsprechern befindet. Jede Abweichung bringt zusätzliche Pegel- und Laufzeitdifferenzen und damit ein Auswandern der fiktiven Schallquelle. Die Stereofläche ist sehr schmal. Bei einer tolerierten Verschiebung der Mittenschallquelle um ± 50 cm bei einer Basisbreite von 3 m ist die Hörfläche in 3 m Hörabstand nur etwa 38 cm. Bei größerer Basisbreite ist die Hörfläche sogar noch schmalere. Lautsprecher mit einem großen Abstrahlwinkel (Kugel-, Kalottenlautsprecher) verbreitern die Stereohörfläche etwa um den Faktor 1,5. Sie haben jedoch den Nachteil geringerer Lokalisationsschärfe.

Bei Laufzeitunterschieden zwischen den Lautsprechern gilt das Gesetz der ersten einfallenden Wellenfront. Bei reinen Pegeldifferenzen führt ein Pegelunterschied von 15-20 dB bereits zur Lokalisation am lauterem Lautsprecher.

Stereohörfläche:

- a, Begrenzung bei Lautsprechern mit normaler Richtcharakteristik
- b, mit breiter Richtcharakteristik

