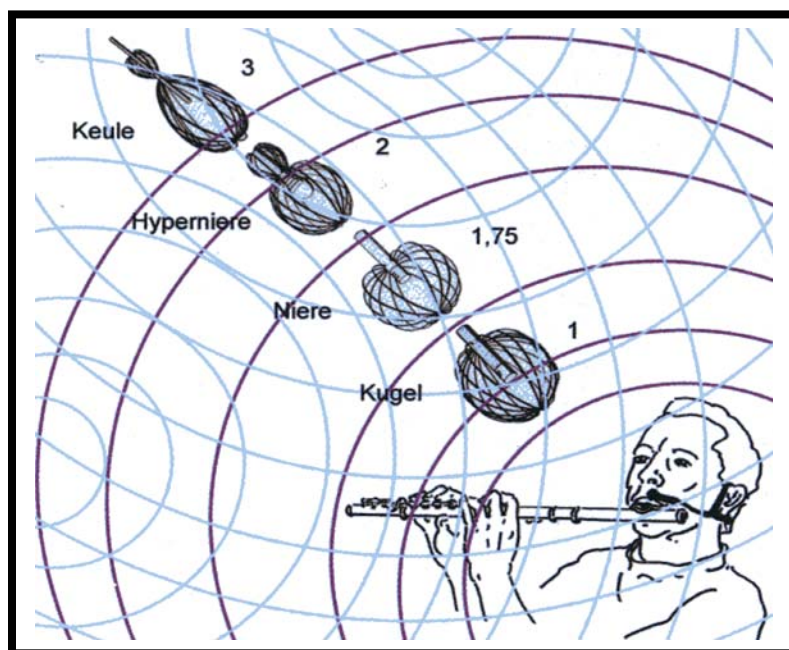


Skript zur Vorlesung

Tontechnik



1. Akustische Kommunikation

Akustik ist die Lehre von Schallwellen in Gasen, Fluiden und Festkörpern.

Akustik: ακουω ich höre (griech.)
Audio: audio ich höre (lat.)

„**Audio**“ umfasst im weitesten Sinne die Phänomenologie aller, d.h. sowohl bewusst erzeugter (z.B. Musik, Alarmsignale, ...) als auch auf natürlichem Wege (Vogelgezwitscher, Bachrauschen, ...) oder ungewollt (Straßenlärm, ...) entstandener Sinnesreize der menschlichen Umwelt, die primär mit Hilfe des Gehörs wahrgenommen wird.

Unsere Umwelt nehmen wir hauptsächlich visuell wahr.
Dennoch: Wegsehen ist kein Problem, aber Weghören ...?

Das Ohr ist ein Achtungsorgan, das immer eingeschaltet ist. Mangelhaftes akustisches Design stört auch (oder gerade) dann, wenn es nur unbewusst wahrgenommen wird!

Musik eignet sich am besten, um Emotionen mitzuteilen: Tanzmusik, Musik zu spirituellen Handlungen, Trinklieder, Schlachtgesänge usw. Die meisten Filme wären leblos, wenn Musik und Geräusche nicht die zum Bild passende Stimmung vermitteln würden.

Oder wissenschaftlich: Auditive Sinnesreize werden zuerst im Stamm- und Zwischenhirn (also den für die Körpergrundfunktionen, Hormonhaushalt, Gefühle etc. zuständigen Hirnteilen) verarbeitet und dann erst ins Großhirn für die rationale Auswertung weitergeleitet. Aus diesem Grund vermag Musik auch unmittelbare Körperreaktionen wie z.B. Änderungen von Pulsfrequenz oder Blutdruck hervorzurufen.

Emotion ist die Domäne auditiver Wahrnehmung.

Eine Vielzahl von menschlichen Handlungen, natürlichen Vorgängen und Bewegungen bewirkt stets auch charakteristische akustische Ereignisse. Die Handlung als Ursache und das akustische Ereignis als deren Wirkung sind dabei fest aneinander gebunden.

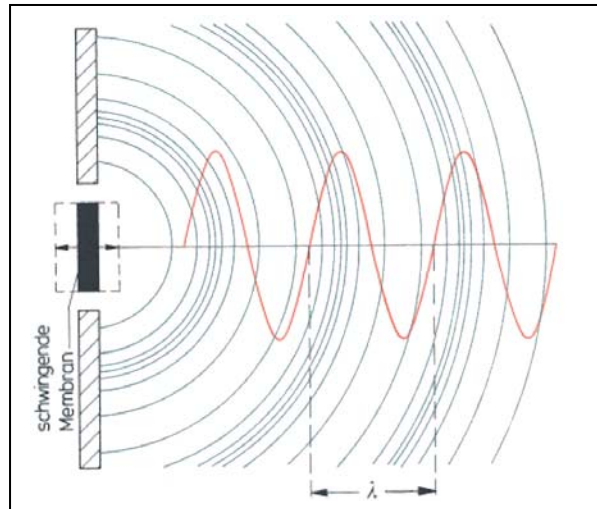
Dem Faktor Zeit kommt hierbei eine besondere Bedeutung zu, da hierin ein wesentlicher Unterschied zur visuellen Wahrnehmung liegt. Akustische Ereignisse sind grundsätzlich flüchtig und nicht dauerhaft, sie benötigen die zeitliche Veränderung, um überhaupt existieren zu können. Das Ohr nimmt zusätzlich Eindrücke aus dem gesamten umgebenden Raum wahr. Räumliche Wahrnehmung geschieht daher sehr wesentlich über das Hören.

Audiodesign:

- Weghören ist nicht möglich
- Ich hör, was Du nicht siehst
- Das Ohr ist manchmal schneller als das Auge
- Akustische Ereignisse verbinden
- Akustische Ereignisse können ein objektiveres Bild der Wirklichkeit vermitteln
- Akustische Ereignisse sind nicht austauschbar
- Musik weckt Emotionen
- Der Ton macht die Musik
- Musik kann die Gedächtnisleistung steigern
- Laute erregen Aufmerksamkeit

Akustik ist die Wissenschaft vom Schall. Schall sind mechanische Schwingungen und Wellen eines elastischen Mediums. Die Akustik wendet die allgemeine Wellenlehre auf mechanische Longitudinalwellen im Hörbereich an. Schall ist an Materie gebunden, im Vakuum gibt es keinen Schall.

Erzeugung von Schallwellen:



Wichtige Eigenschaften akustischer Ereignisse sind Zeit und räumliche Richtung. Im Gegensatz zu Bildern ist bei Luftschall der dreidimensionale Ortsbereich und das Zeitverhalten zu berücksichtigen.

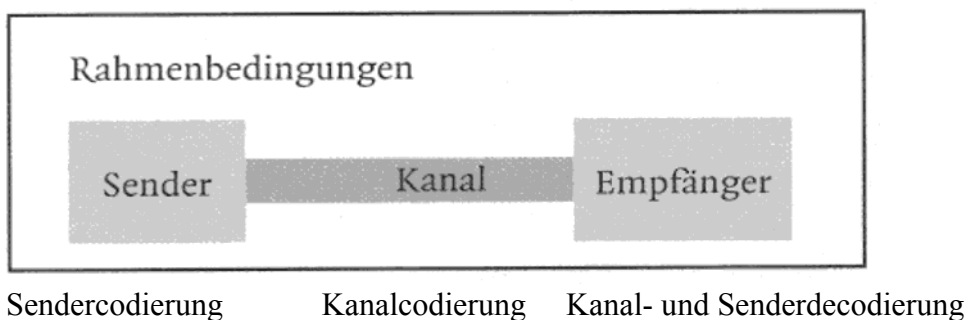
Schallempfindung:

Die physikalischen Schallgrößen werden in Empfindungen umgewandelt und vom Gehirn entsprechend interpretiert. Die wesentlichen Entsprechungen sind:

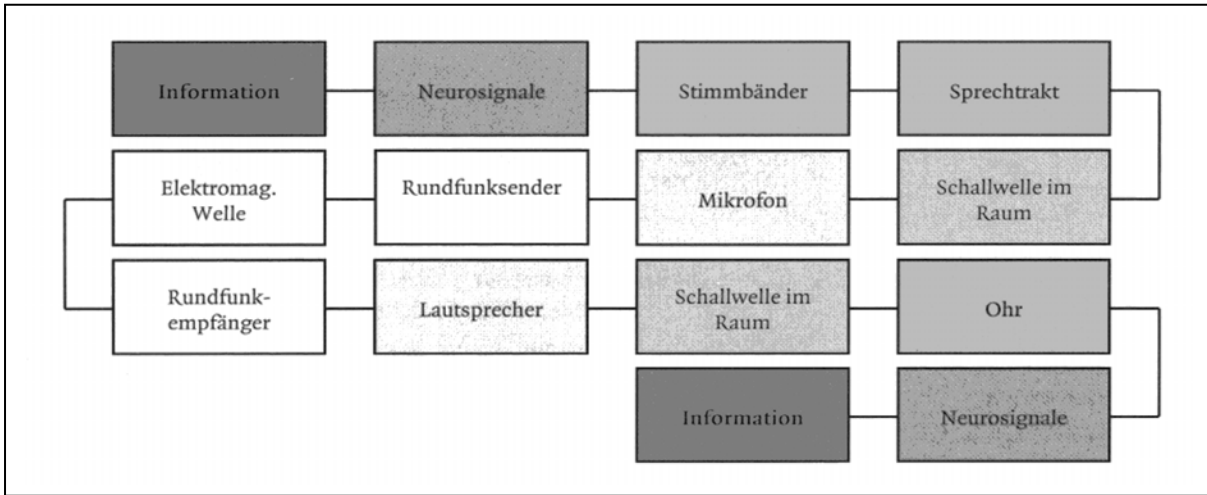
Reiz:		Empfindung:
Frequenz	→	Tonhöhe
Schallintensität	→	Lautstärke
Spektrum	→	Klangfarbe

Die Zusammenhänge sind allerdings sehr komplex. So ist z.B. die Tonhöhenempfindung geringfügig auch abhängig von der Schallintensität. Das Ohr kann eine Grundtonhöhe aus einem Spektrum ermitteln, welches diese Frequenz gar nicht enthält (virtuelle Tonhöhe). Es kann beim Telefonieren beispielsweise zwischen männlicher (100 Hz) und weiblicher (200 Hz) Stimme unterscheiden, obwohl das Telefonsignal nur Frequenzen zwischen 300 und 3000 Hz enthält. Auch die Klangfarbe wird nicht nur vom Spektrum, sondern auch von An- und Abklingvorgängen beeinflusst.

Kommunikationsmodell nach Shannon:



Kommunikationskette (z.B. Rundfunkübertragung):



Einflüsse auf Kommunikationskette:

- Raum
- Rauschen
- Wandlerfehler
- Übertragungsfehler
- Verzerrungen

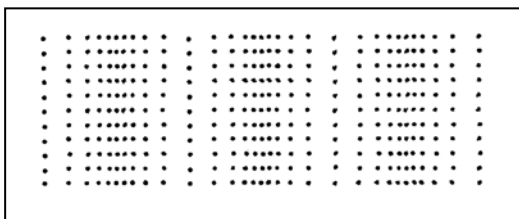
2. Akustische Grundbegriffe

2.1 Wellenausbreitung

Wellentypen in festen Körpern:

- Dichte- oder Longitudinalwelle: höchste Ausbreitungsgeschw., einzige Form in Gasen
- Schub- oder Transversalwelle
- Torsionswelle
- Biegewelle

Longitudinal- oder Dichtewelle:



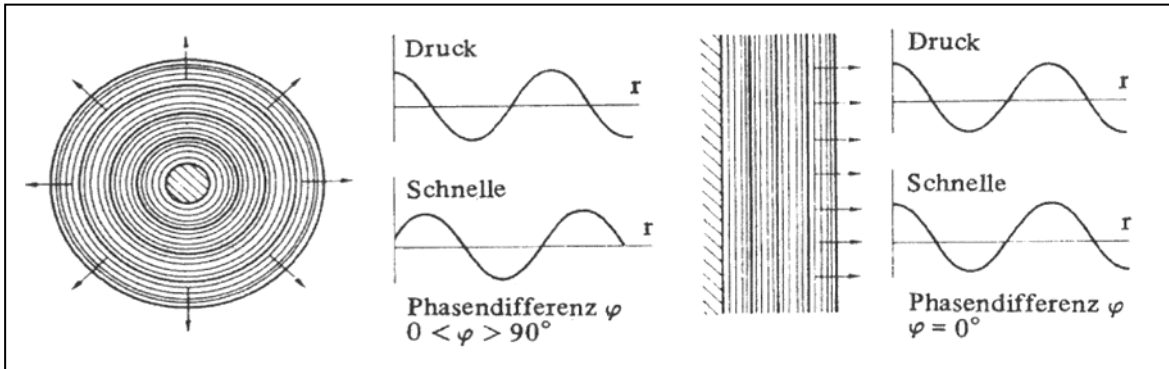
Die Longitudinalwelle breitet sich mit der für das betreffende Material größten Schallgeschwindigkeit aus. Sie tritt in solchen Körpern und Medien auf, deren Ausdehnung nach allen Richtungen als unendlich groß oder zumindest sehr groß gegenüber der Wellenlänge angesehen wird.

Die meisten technischen und natürlichen Schallquellen emittieren Kugelwellen. Zweckmäßig ist es, zur Beschreibung dieser Wellenfelder in Kugelkoordinaten überzugehen. Für die Entfernung des Aufpunktes von der Quelle steht der Radius r der Kugelwelle. Weiterhin verwendet man den Azimutwinkel φ (Längengrad) und den Deklinationswinkel θ (Breitengrad).

Bei einer symmetrischen Kugelwelle unterliegt nur die Koordinate r Veränderungen. Sie breitet sich mit der Schallgeschwindigkeit c von einem bei $r = 0$ gelegenen Punkt in Richtung wachsender Entfernung r aus und zwar gleichmäßig nach allen Raumrichtungen. Dabei behält sie ihre Form bei, ihre Stärke nimmt aber proportional $1/r$ ab.

Schalldruck p und Schallschnelle v sind nahe der Quelle (Nahfeld der Quelle) nicht in Phase.

Kugelwelle und ebene Welle:

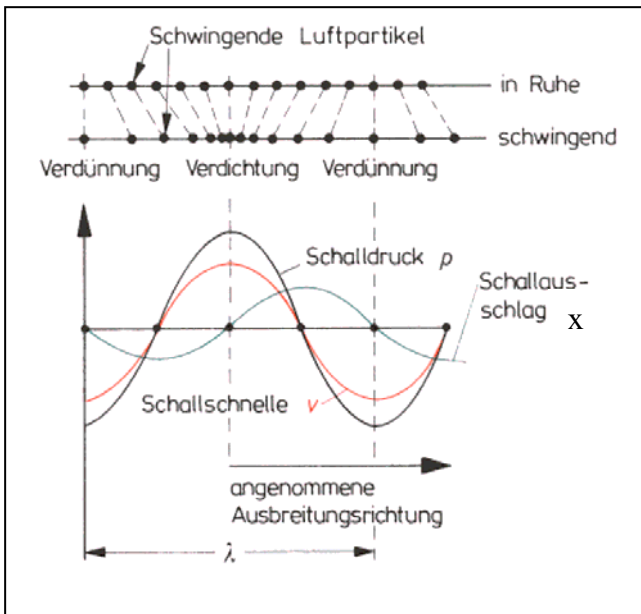


In der Kugelwelle sind die sich ausbreitenden Wellenfronten Kugelschalen, die sich mit zunehmendem Abstand von der Quelle quasioebenen Flächen, d.h. für ein betrachtetes Flächenelement ebenen Wellenflächen, nähern. Bei einer ebenen Welle sind Schalldruck und Schallschnelle in Phase und die beschreibenden Gleichungen vereinfachen sich erheblich.

Hängt der momentane Zustand in einer Schallwelle also nur von einer einzigen Richtung, etwa der x-Richtung ab, so spricht man von einer ebenen Welle. Näherungsweise erhält man diese Wellenform in einem großen Abstand zur Kugelwelle. Als einfache Grenze, ab der mit einer ebenen Welle gerechnet werden kann, gilt:

$$\text{Wellenlänge } \lambda < 2\pi r$$

Ebene Welle: Schalldruck und Schallschnelle sind in Phase



$$p \approx \frac{f}{r} \sin 2\pi \left(ft - \frac{r}{\lambda} \right) = \hat{p} \sin (\omega t - kr)$$

$$v \approx \frac{f}{r} \sin 2\pi \left(ft - \frac{r}{\lambda} \right) = \hat{v} \sin (\omega t - kr)$$

mit:
$$kr = \frac{\omega r}{c} = \frac{2\pi r}{\lambda}$$

r = Abstand von der Quelle [m]

p = Schalldruck [Pa]

v = Schallschnelle [m/s]

k = Kreiswellenzahl

Nahfeld: $2\pi r < \lambda$

Übergang: $kr = 1$

Fernfeld: $2\pi r > \lambda$



2.2 Schallfeldparameter

Gleichfeld und Wechselfeld:

$$\begin{aligned} \rho &= \rho_0 + \rho_{\sim} \\ p &= p_0 + p_{\sim} \\ v &= v_0 + v_{\sim} \end{aligned}$$

- **Schalldruckpegel L:** [dB]
 Bezugsschalldruck: 20 μPa
 SPL: Sound Pressure Level, Bezugsschalldruck immer 20 μPa

$$L = 20 \cdot \log \frac{p}{p_0} \text{ dB}$$

- **Schallintensitätspegel L:** [dB]
 Bezugsintensität: $I_0 = 10^{-12} \text{ W/m}^2$

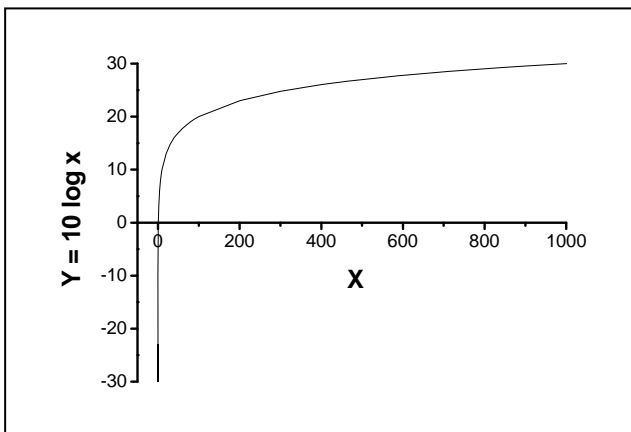
$$L = 10 \cdot \log \frac{I}{I_0} \text{ dB}$$

Logarithmische Gesetzmäßigkeiten:

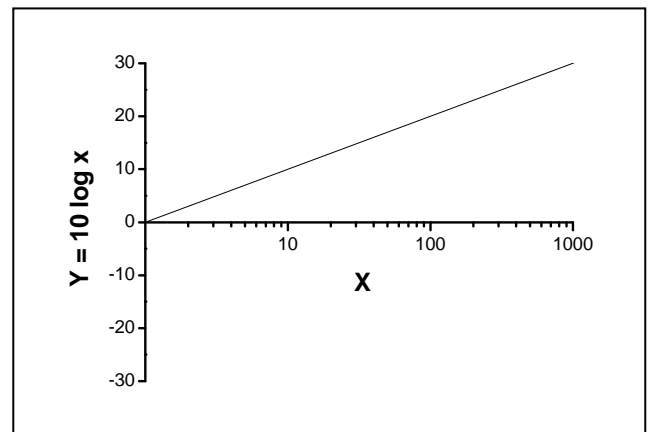
$$y = \log x^a = a \log x$$

$$y = \log(a \cdot b) = \log a + \log b$$

$$y = 10 \cdot \log x \quad \longrightarrow \quad x = 10^{\frac{y}{10}}$$



x-Achse linear



x-Achse logarithmisch

Pegelrechnung:

$$I = \frac{P}{Z} \quad P = \frac{U^2}{R}$$

Pegel [dB]	p/p ₀ bzw. U/U ₀	I/I ₀ bzw. P/P ₀
40	10 ²	10 ⁴
20	10	10 ²
10	√10 = 3,16	10
6	2	4
3	√2 = 1,41	2
1	1,12	1,26
0	1	1
-1	1/1,12 = 0,89	1/1,26 = 0,79
-3	1/√2 = 0,707	1/2 = 0,5
-6	1/2 = 0,5	1/4 = 0,25
-10	1/√10 = 0,316	1/10 = 0,1
-20	1/10 = 0,1	1/100 = 0,01
-40	1/10 ² = 0,01	1/10 ⁴ = 0,0001

„Absolute“ Pegel bei festen Bezugsgrößen:

„Studiopegel“ = „Rundfunknormpegel“ = „Funkhausnormpegel“:

$$L = 20 \log (u / 1,55 \text{ V}) \text{ dB}$$

andere Bezugsgrößen:

$$\text{dBm} \rightarrow L = 10 \log (P / 1 \text{ mW}) \text{ dBm}$$

$$\text{dBu} \rightarrow L = 20 \log (U / 0,775 \text{ V}) \text{ dBu}$$

$$\text{dB}_V \rightarrow L = 20 \log (U / 1 \text{ V}) \text{ dB}_V$$

$$\text{dB}_{\text{SPL}} \rightarrow L = 20 \log (p / 20 \text{ } \mu\text{Pa}) \text{ dB}_{\text{SPL}}$$

Beispiel: Welche Ausgangsspannung liefert eine D/A-Wandlerkarte, bei der am Ausgang max. 19 dBu anliegen?



Nomogramm zur Addition von inkohärenten Schallen:

$L_1 - L_2:$	ΔL zu $L_1:$	$L_1 > L_2$
0	$L_1 + 3$ dB	
1 bis 3 dB	$L_1 + 2$ dB	
4 bis 9 dB	$L_1 + 1$ dB	
≥ 10 dB	$L_1 + 0$ dB	

Faktor: Verhältnis linearer Größen
Schalldruck, Schallschnelle

Grad: Verhältnis quadratischer Größen
Schallintensität, Schall-Leistung

Maß oder Pegel: logarithmisches Verhältnis
Übertragungsmaß, Schalldruckpegel, Schallintensitätspegel

Kohärente Quellen: voneinander abhängige Quellen \longrightarrow Schalldruckpegel
gleichfrequente Töne *Schalldrücke addieren sich*

Inkohärente Quellen: voneinander unabhängige Quellen \longrightarrow Schallintensitätspegel
verschiedenfrequente Töne, Geräusche *Intensitäten addieren sich*

Frequenzintervalle:

$f_u/f_o = 1 : 1,26$	Terz (= 1/3 Oktave)
$f_u/f_o = 1 : 2$	Oktave
$f_u/f_o = 1 : 10$	Dekade

1 dB/Terz = 3 dB/Okt = 10 dB/Dek

Vorzugsreihe für Terz- und Oktavband-Mittenfrequenzen nach der ISO-Empfehlung R 266:

Mittenfrequenz in Hz	Terzfilter	Oktavfilter	Mittenfrequenz in Hz	Terzfilter	Oktavfilter
12,5	+		800	+	
16	+	+	1000	+	+
20	+		1250	+	
25	+		1600	+	
31,5	+	+	2000	+	+
40	+		2500	+	
50	+		3150	+	
63	+	+	4000	+	+
80	+		5000	+	
100	+		6300	+	
125	+	+	8000	+	+
160	+		10000	+	
200	+		12500	+	
250	+	+	16000	+	+
315	+		20000	+	
400	+		25000	+	
500	+	+	31500	+	+
630	+		40000	+	

$$\Delta f = f_o - f_u = (\sqrt[6]{2} - 1/\sqrt[6]{2}) f_m \approx 0,23 f_m \qquad f_m = \sqrt{f_u f_o}$$

2.3 Signaldarstellung

Der **Effektivwert** (RMS-Wert, root mean square) eines Zeitsignals beschreibt die Leistung des Signals über einen längeren Zeitabschnitt T_0 :

$$u_{RMS}(t) = \sqrt{\frac{1}{T_0} \int_0^{T_0} u^2(t) dt}$$

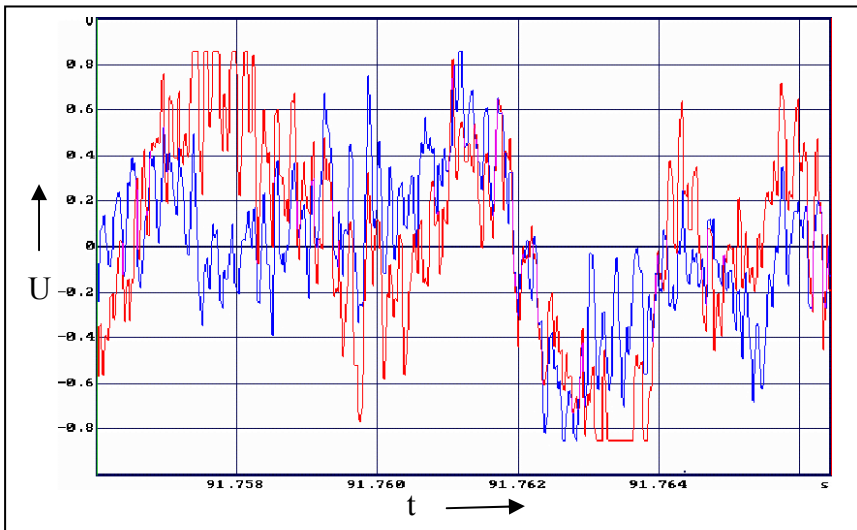
Crestfaktor:

$$\text{Crestfaktor} = \frac{\text{Spitzenwert}}{\text{Effektivwert}}$$

Typische Crestfaktoren von Signalen: wichtig für Aussteuerung von Signalen

Signal	Crestfaktor	Pegeldifferenz in dB
Rechteck	0	0
Sinus	1,41	3
Weißes Rauschen	3,3	≈ 10
Sprache	6...10	16...20
Musik	3...30	10...30

Digitale Übersteuerung (Clipping) auf einer CD:



Harmonische Schwingung:

$$x(t) = A \cos(\omega t + \varphi)$$

mit: $\omega = 2 \pi f$ und $T = 1/f$

oder komplex: $\underline{x}(t) = A e^{j(\omega t + \varphi)}$

Beobachtbar ist nur Realteil oder Imaginärteil:

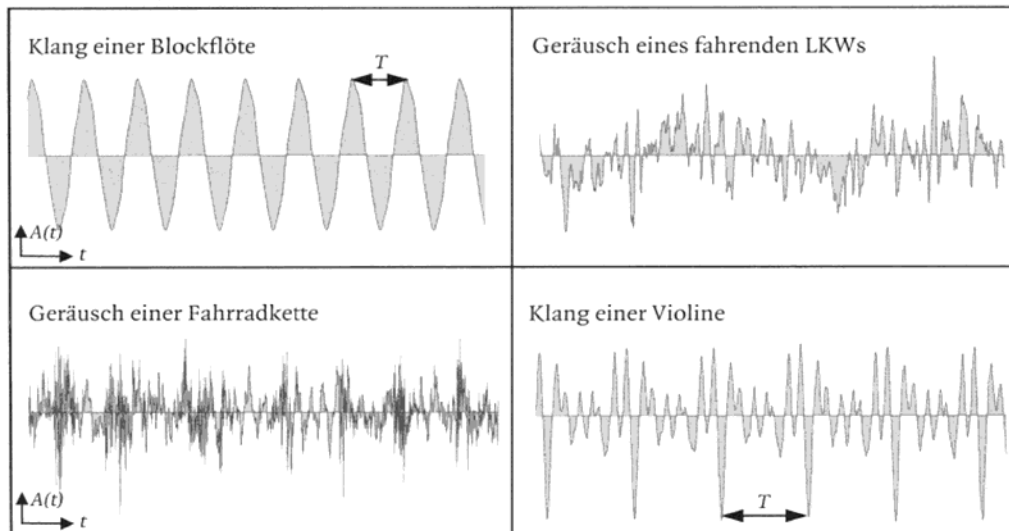
$$\text{Re} \{A e^{j(\omega t + \varphi)}\} = A \cos(\omega t + \varphi)$$

Vorteil der komplexen Schreibweise:

$$\frac{d\underline{x}}{dt} \equiv j\omega \underline{x}$$

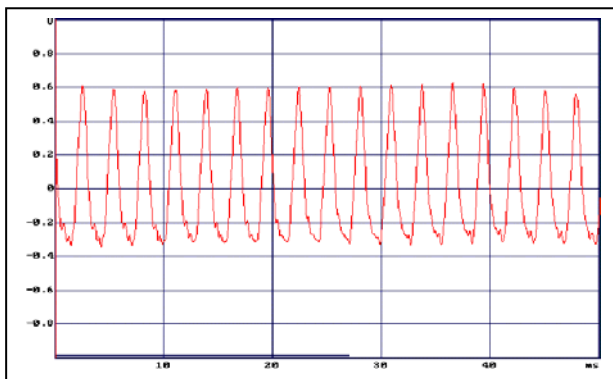
$$\int \underline{x} dt \equiv \frac{\underline{x}}{j\omega}$$

Signalformen von Klängen und Geräuschen:

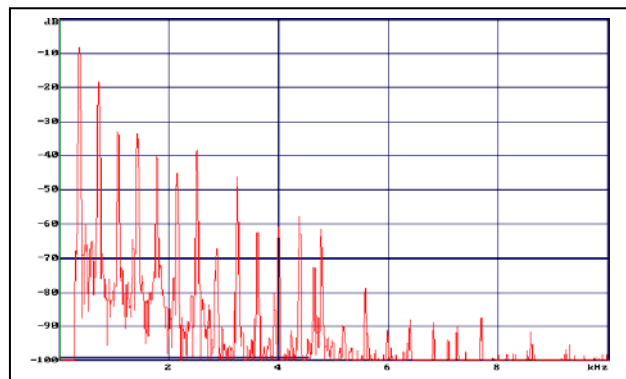


Klang: periodisches Signal → diskretes Spektrum
 Geräusch: aperiodisches Signal → kontinuierliches Spektrum

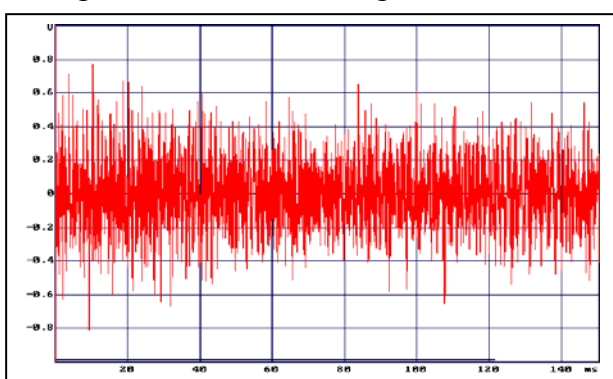
Zeitsignal eines Klaviertones:



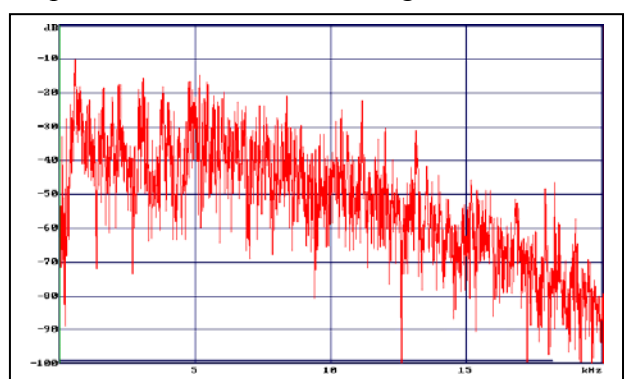
Spektrum eines Klaviertones:



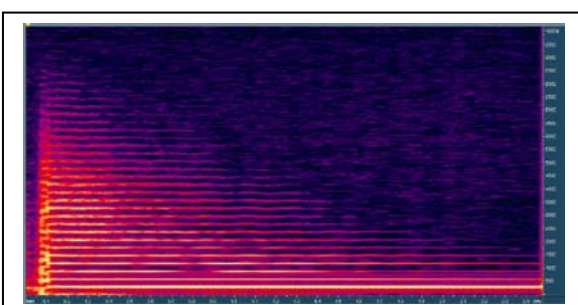
Zeitsignal eines Beckenschlages:



Spektrum eines Beckenschlages:



Spektrogramm eines E-Gitarren-Tones:



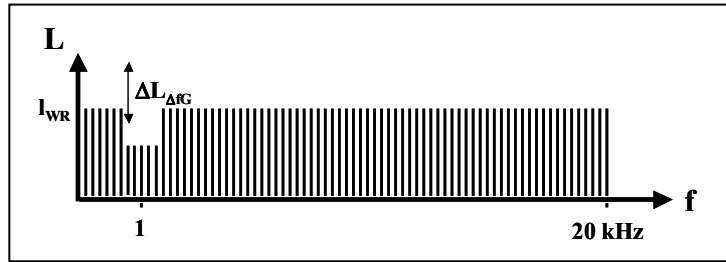
x-Achse: Zeit
 y-Achse: Frequenz
 Farbe: Pegel

Komplexe Töne bzw. Klänge können aus mehreren Teiltönen zusammengesetzt werden, da sie periodisch sind:

$$s(t) = \sum_{n=1}^N A_n \cos(2\pi n f t + \varphi_n)$$

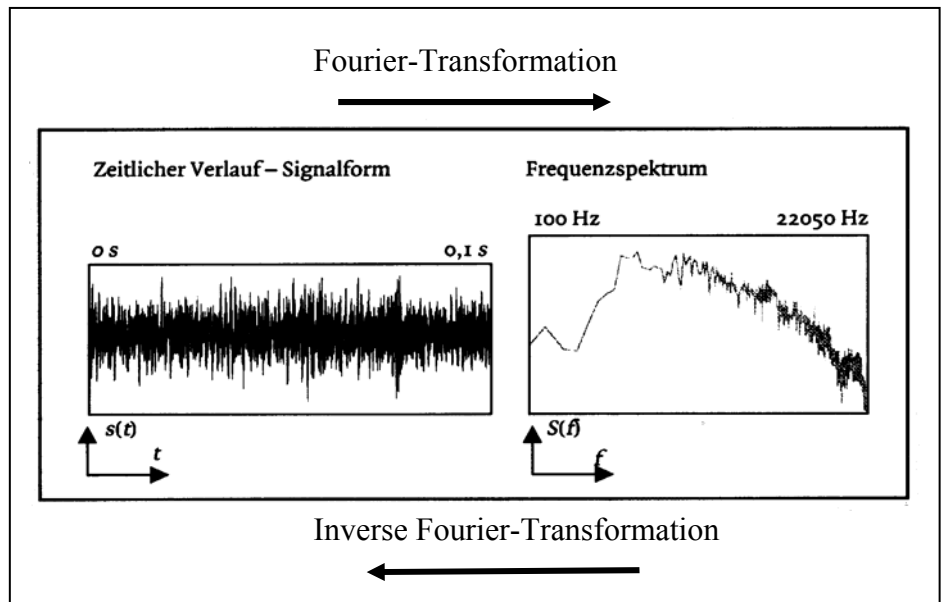
Linienpektrum mit konstanter Amplitude (I_{WR} konstant):

Addition vieler Cosinusschwingungen



Mit Hilfe der Fourier-Transformation ist es möglich, für jedes beliebige Signal das entsprechende Spektrum zu ermitteln. Zeitbereich und Frequenzbereich eines Signals stehen gleichwertig nebeneinander. Die inverse Fourier-Transformation ermöglicht umgekehrt, aus dem Spektrum des Signals den entsprechenden Zeitverlauf zu ermitteln. Mit der Fast Fourier-Transformation steht ein effizienter Algorithmus zur Verfügung, der in einigen Softwarepaketen für die Audiotbearbeitung implementiert ist.

Zeit- und Frequenzbereich eines Bachrauschens:



Rauschen:

Während Sinustöne und Klänge aus einzelnen Spektrallinien bestehen, gibt es sehr viele Schallereignisse, die ein kontinuierliches Spektrum enthalten. Ein typischer Schall dieser Art ist das Rauschen, das als ständiger, aber nicht periodisch verlaufender stochastischer Vorgang mit Hilfe statistischer Kenngrößen beschrieben werden kann. Beispiele sind das Rauschen eines Wasserfalls, Strömungsgeräusche oder ein gesprochener f-Laut. Die wichtigste Größe zur Kennzeichnung von Rauschvorgängen ist die (frequenzabhängige) spektrale Schallintensitätsdichte $dI(f)/df$. Aus ihr kann der Schallintensitätsdichtepiegel oder kurz der Dichtepiegel $l(f)$ bestimmt werden:

$$l(f) = 10 \log \frac{dI(f)/df}{I_0/1\text{Hz}} \text{ dB}$$

Der Bezugswert der Schallintensitätsdichte ist die Bezugsintensität I_0 pro Hertz Bandbreite.

Der Dichtepegel $l(f)$ kann aber auch folgendermaßen ausgedrückt werden:

$$l(f) = 10 \log \frac{dI(f)/I_0}{df/1\text{Hz}} \text{ dB}$$

Aus der Schallintensitätsdichte $dI(f)/df$ wird durch Integration über das betrachtete Frequenzintervall von f_u bis f_o die in dieses Intervall fallende Gesamtschallintensität bestimmt:

$$I = \int_{f_u}^{f_o} \frac{dI(f)}{df} df$$

Für **Weißes Rauschen**, das eine frequenzunabhängige Schallintensitätsdichte $dI(f)/df$ und damit frequenzunabhängigen Dichtepegel $l(f)$ aufweist, vereinfacht sich die Gleichung zu:

$$I_{WR} = \frac{dI(f)}{df} (f_o - f_u)$$

Daraus folgt für den Gesamtpegel:

$$L_{WR} = 10 \log \left[\frac{dI(f)/df}{I_0} (f_o - f_u) \right] \text{ dB} = l_{WR} + 10 \log \frac{f_o - f_u}{1 \text{ Hz}} \text{ dB}$$

Der Intensitätsdichtepegel l_{WR} ist über der Frequenz konstant:

Der Dichtepegel l_{WR} ist somit:

$$l_{WR} = \left(10 \log \frac{I}{I_0} - 10 \log \frac{f_o - f_u}{1\text{Hz}} \right) \text{ dB} = L_{WR} - \left(10 \log \frac{f_o - f_u}{1 \text{ Hz}} \right) \text{ dB}$$

Bei bekanntem Dichtepegel l_{WR} eines Weißen Rauschens, das sich von **20 Hz bis 20 kHz** erstreckt, lässt sich daher der Gesamtpegel L_{WR} leicht berechnen zu:

$$L_{WR} = l_{WR} + 10 \log (20000 - 20) \text{ dB} = l_{WR} + 43 \text{ dB}$$

Beispiel: Berechnen Sie den Dichtepegel l_{WR} bei einem Gesamtpegel L_{WR} von 80 dB in einem Frequenzbereich von 1 bis 4 kHz.

Wieviele Terzen sind darin enthalten?

Wird Weißes Rauschen mit Filtern **konstanter** Bandbreite (z.B. $\Delta f = 100$ Hz) analysiert, so erhält man bei verschiedenen Filtermittenfrequenzen immer die gleiche Filterausgangsspannung bzw. den gleichen Pegel. Wird hingegen Weißes Rauschen mit Filtern **konstanter relativer** Bandbreite wie z.B. Terzfiltern gemessen, so erhält man wegen der mit steigender Mittenfrequenz zunehmenden Filterbandbreite keinen frequenzunabhängigen Ausgangspegel mehr, sondern bei Verwendung von Terzfiltern einen Ausgangspegel, der mit 10 dB pro Frequenzdekade zunimmt, da die Bandbreite einer Terz bei 1 kHz beispielsweise zehnmal so groß ist wie bei 100 Hz. Dementsprechend fällt bei Weißem Rauschen in den 1-kHz-Terzfilter die 10fache Leistung eines 100-Hz-Terzfilters, was einer Pegelzunahme um 10 dB entspricht.

Um bei Messungen mit Terzfiltern einen frequenzunabhängigen Ausgangspegel zu erhalten, muss daher ein Rauschen verwendet werden, dessen Dichtepiegel I_{RR} zu hohen Frequenzen hin mit 10 dB/Dekade abnimmt. Dieses Rauschen wird als Rosa Rauschen bezeichnet. Neben dem häufig verwendeten Rosa Rauschen gibt es noch weitere farbige Rauscharten.

Der Intensitätsdichtepiegel von Rosa Rauschen, Rotem Rauschen bzw. Blauem Rauschen fällt mit 10 dB/Dek., 20 dB/Dek. bzw. steigt mit 10 dB/Dek.:

Für viele Messungen wird in der Elektroakustik zur Frequenzgangbestimmung anstelle eines Gleitsinus bandbegrenztes Rosa Rauschen mit Terzfiltern analysiert. Der Hörbereich umfasst 30 Terzen. Für Rosa Rauschen im Bereich 20 Hz – 20 kHz und gleicher Leistung von Sinuston und Rauschen ergibt sich die Pegeldifferenz von:

$$L_{\text{Terz}} - L_{\text{Sinus}} = 10 \log 1/30 \text{ dB} \approx -14,8 \text{ dB}$$

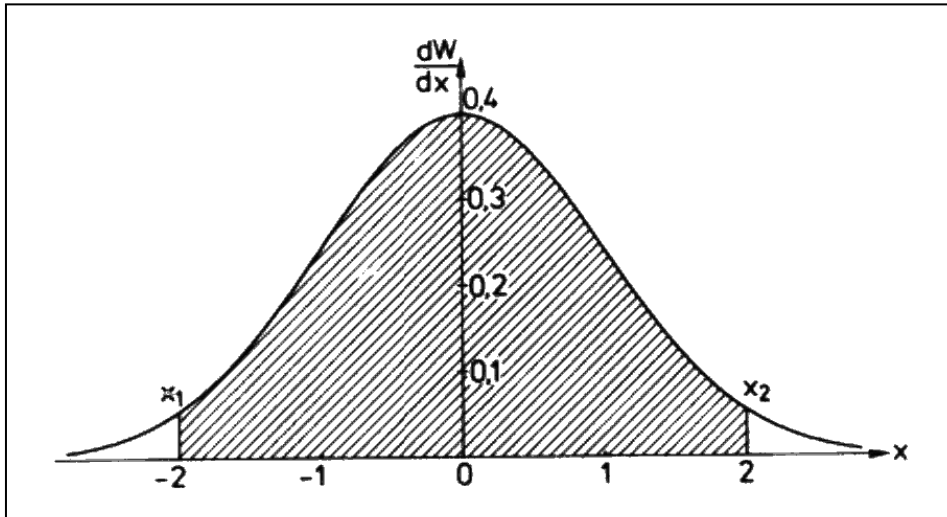
Beachte: In eine Terz fällt nur dann 1/30 der Gesamtintensität, wenn sich der Frequenzbereich auch von 20 Hz bis 20 kHz erstreckt. Ist der Frequenzbereich geringer, so können in diesen auch nur weniger Terzen untergebracht werden.

Der Augenblickswert einer Rauschspannung $u(t)$ oder eines Rauschschalldruckes $p(t)$ kann nicht vorherbestimmt werden, sondern es lassen sich nur statistische Größen angeben. Die Wahrscheinlichkeit dW dafür, dass der Augenblickswert $p(t)$ des Weißes Rauschens zwischen $p-dp/2$ und $p+dp/2$ liegt, ist:

$$dW = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \quad \text{mit:} \quad x = \frac{p(t)}{p_{\text{eff}}}$$

Die Wahrscheinlichkeitsdichte dW/dx für Weißes Rauschen besitzt also eine Gaußverteilung.

Wahrscheinlichkeitsdichte für Weißes Rauschen:



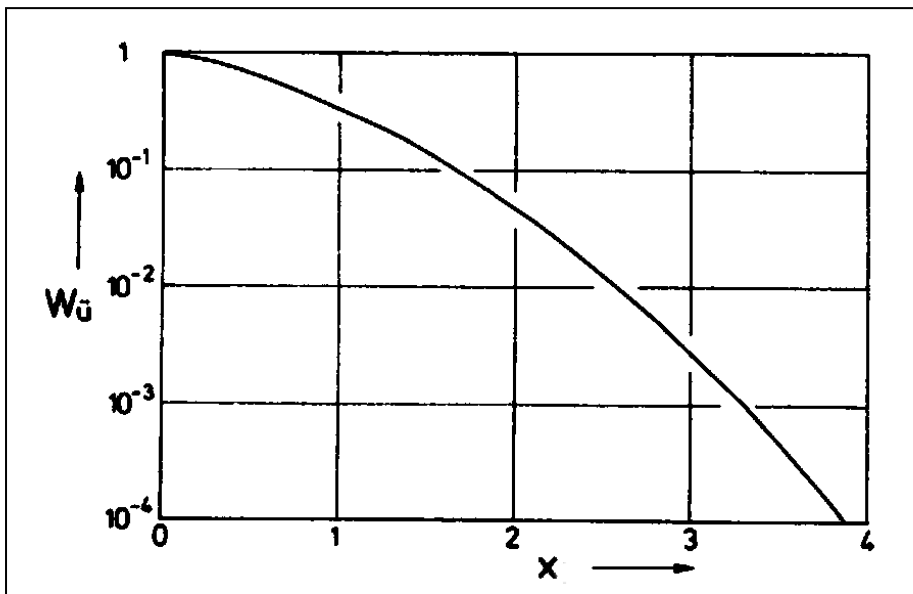
Die Wahrscheinlichkeit W dafür, dass x zwischen zwei Werten x_1 und x_2 liegt, ist gleich der Fläche unter der Kurve zwischen diesen beiden Abszissenwerten.

$$W = \int_{x_1}^{x_2} \frac{dW}{dx} dx = \int_{x_1}^{x_2} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

Die Überschreitungswahrscheinlichkeit $W_{\ddot{u}}$ dafür, dass der Betrag des Schalldruckes bei Weißem Rauschen einen bestimmten Wert p überschreitet, ist:

$$W_{\ddot{u}} = 2 \int_x^{\infty} \frac{dW}{dx} dx = 1 - \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{x^2}{2}} dx \quad \text{mit: } x = \frac{p(t)}{p_{eff}}, p \geq 0$$

Überschreitungswahrscheinlichkeit für Weißes Rauschen:



Um eine Überschreitungswahrscheinlichkeit von 1/1000 sicherzustellen, muss gelten:

$$x = 3,3 \text{ bzw. } p = 3,3 p_{eff}$$

Üblicherweise wird bei der Übertragung von Weißem Rauschen gegenüber den Effektivwert, der von einem Spannungsmesser angezeigt wird, eine Aussteuerungsreserve ΔL von etwa $20 \log 3,3 \text{ dB} \approx 10 \text{ dB}$ eingehalten. Dadurch kann zumindest innerhalb einer gewissen Wahrscheinlichkeit eine Übersteuerung der Aufnahme vermieden werden.

Beispiel: Mit welcher Wahrscheinlichkeit W_u wird eine Aufnahme von WR mit 6 dB Pegelreserve übersteuert?

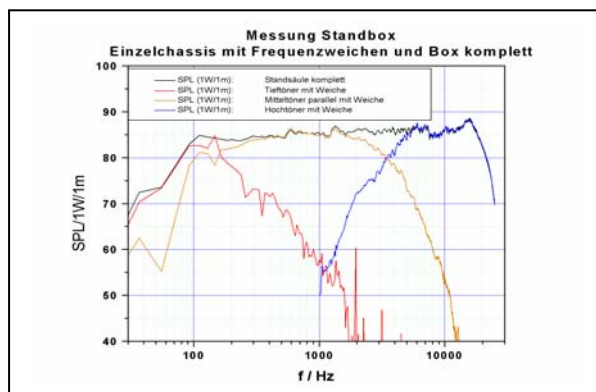
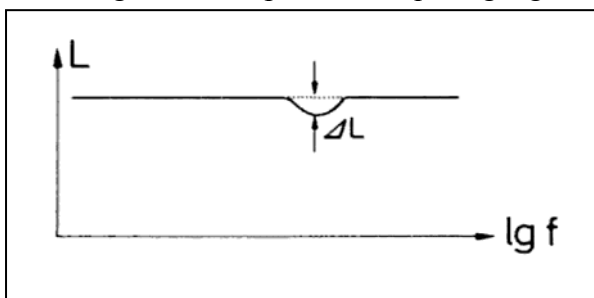
2.4 Verzerrungen

2.4.1 Lineare Verzerrungen

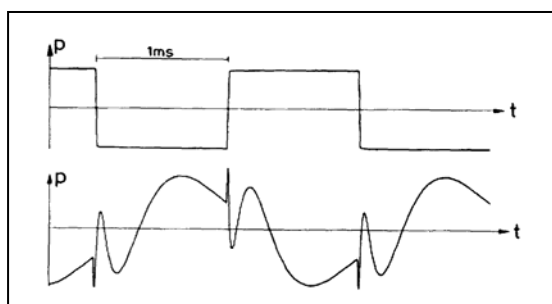
Für eine ideale Übertragung sollte für jede Amplitude das Übertragungsmaß von der Frequenz unabhängig sein. Dies ist nicht erreichbar, weil insbesondere die Schallwandler und die Räume, in denen sie aufgestellt sind, im zu übertragenden Frequenzbereich Resonanzen aufweisen. Wir müssen daher bei allen elektroakustischen Übertragungen mit Veränderungen der spektralen Zusammensetzung des Schalls rechnen.

Die Grenze der Hörbarkeit solcher Änderungen ist durch die Frequenzselektivität und die Pegelunterschiedsschwelle des Gehörs bestimmt. Demnach wird bei breitbandigen Schallen eine unmittelbare Änderung des Amplitudenfrequenzganges im direkten Vergleich dann bemerkt, wenn sich der Pegel in irgendeiner Frequenzgruppe um mehr als 1 dB ändert. Ohne Vergleich und bei nur auf Erfahrung oder Erinnerung basierenden Klangeindrücken können Pegelunterschiede ΔL von ca. 2-3 dB bemerkt werden.

Änderungen des Amplitudenfrequenzganges:



Änderungen im Phasenfrequenzgang sind im Allgemeinen kaum hörbar. Beide Zeitfunktionen erzeugen denselben Klangeindruck, obwohl die Phase des zweiten Signals sehr stark verzerrt wurde:



2.4.2 Harmonische nichtlineare Verzerrungen

Wenn die Übertragungskennlinie eines Systems von der idealen Geraden abweicht, kann sie durch eine Polynomdarstellung beschrieben werden, deren höhere Koeffizienten ungleich Null sind.

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots$$

Wird für die Eingangsgröße x eine Sinusschwingung $x = x_0 \cos(\omega_1 t)$ eingesetzt und die Polynomdarstellung entsprechend ausmultipliziert, so entstehen zusätzliche Harmonische bei der doppelten und dreifachen bzw. bei Vielfachen der Grundfrequenz.

$y = (a_2 x_0^2 / 2 + \dots)$	Gleichdruck!
$+ (a_1 x_0 + 3 a_3 x_0^3 / 4 + \dots) \cos \omega_1 t$	Grundschwingung bei f_1
$+ (a_2 x_0^2 / 2 + \dots) \cos 2 \omega_1 t$	2. Harmonische bei $2f_1$
$+ (a_3 x_0^3 / 4 + \dots) \cos 3 \omega_1 t$	3. Harmonische bei $3f_1$
$+ (\dots)$	

Der Klirrfaktor k_{ges} ist folgendermaßen definiert:

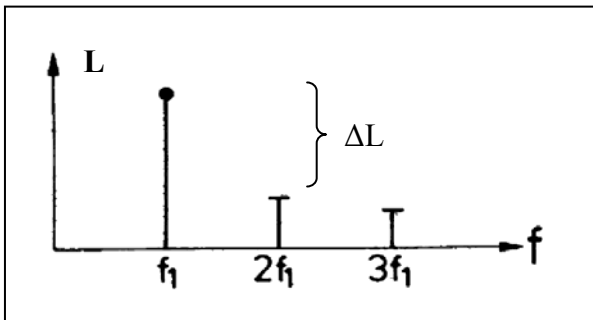
Näherungsweise kann daraus der quadratische Klirrfaktor k_2 bzw. der kubische Klirrfaktor k_3 abgeleitet werden:

Unter der Annahme, dass $a_1 = 1$ und $a_2 \ll 1$ sowie $a_3 \ll 1$ ist, ergibt sich folgende Abhängigkeit von k_2 bzw. k_3 :

$$k_2 \approx \frac{1}{2} a_2 x_0 \qquad k_3 \approx \frac{1}{4} a_3 x_0^2$$

Der kubische Klirrfaktor ist damit quadratisch von der Amplitude des Eingangssignals abhängig. Kubische Verzerrungen werden zudem aufgrund ihrer spektralen Komponente besser wahrgenommen als quadratische Verzerrungen, da die 3. Harmonische nicht in unsere auf Oktaven beruhende Harmonielehre passt.

Verzerrungsprodukte bei Übertragung eines Sinustones über eine nichtlineare Kennlinie:



$$\Delta L = L_1 - L_n = 20 \log \frac{p_1}{p_n} \text{ dB} = 20 \log \frac{1}{k_n} \text{ dB}$$

Klirrdämpfung a_n :

$$a_n = 20 \log \frac{1}{k_n} \text{ dB}$$

Beispiel: Berechnen Sie die quadratische Verzerrung k_2 bei einem Pegelabstand ΔL von 40 dB.

2.4.3 Intermodulations-Verzerrungen

Diese Art der Verzerrungen tritt auf, wenn mindestens zwei Sinustöne über eine nichtlineare Kennlinie übertragen werden. Für das Eingangssignal wird damit $x = x_1 \cos(\omega_1 t) + x_2 \cos(\omega_2 t)$ gleichgesetzt.

Für rein quadratische Verzerrungen ergeben sich damit folgende Frequenzen:

$$y = a_1 x + a_2 x^2 \quad \text{und mit} \quad a_2 \ll a_1$$

$$y \approx 0,5 a_2 (x_1^2 + x_2^2)$$

Gleichdruck!

$$+ a_1 (x_1 \cos \omega_1 t + x_2 \cos \omega_2 t)$$

Frequenzen f_1, f_2

$$+ 0,5 a_2 (x_1^2 \cos 2\omega_1 t + x_2^2 \cos 2\omega_2 t)$$

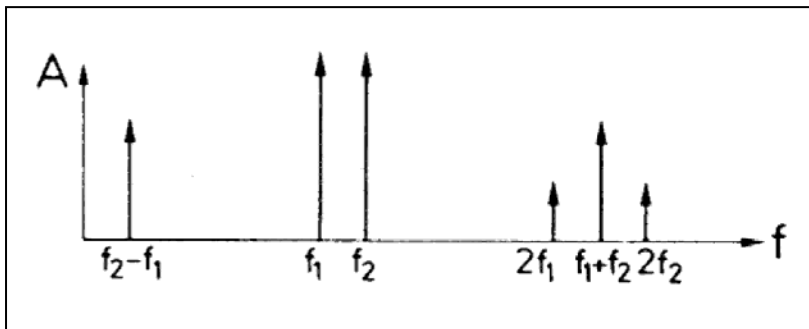
Frequenzen $2f_1, 2f_2$

$$+ a_2 x_1 x_2 [\cos(\omega_2 + \omega_1)t + \cos(\omega_2 - \omega_1)t]$$

Frequenzen $f_2 + f_1, f_2 - f_1$

Zu beachten ist, dass die Amplitude $a_2 x_1 x_2$ der Summen- und Differenzschwingung für den häufig auftretenden Fall $x_1 = x_2$ doppelt so groß ist wie die Amplituden $0,5 a_2 x_1^2$ bzw. $0,5 a_2 x_2^2$ der beiden zweiten Harmonischen. Die Summenschwingung bei der Frequenz $f_1 + f_2$ liegt zwischen den Frequenzen $2f_1$ und $2f_2$. Die Differenzschwingung dagegen liegt völlig allein bei tiefen Frequenzen. Da sie aus einer Differenz zweier Frequenzen entsteht und daher im allgemeinen Fall unharmonisch liegt, darüber hinaus durch Verdeckung kaum beeinflusst wird, weil die Verdeckung durch die starken Primärtöne vor allem nach höheren Frequenzen wirkt, ist diese Differenzschwingung in fast allen Fällen der am meisten störende Anteil der Verzerrungsprodukte.

Verzerrungsprodukte bei der Übertragung zweier Sinustöne über eine Kennlinie mit quadratischen Verzerrungen:



Für rein kubische Verzerrungen gilt:

$$y = a_1 x + a_3 x^3 \quad \text{und mit} \quad a_3 \ll a_1:$$

$$y \approx a_1 (x_1 \cos \omega_1 t + x_2 \cos \omega_2 t)$$

Frequenzen f_1, f_2

$$+ 0,25 a_3 (x_1^3 \cos 3\omega_1 t + x_2^3 \cos 3\omega_2 t)$$

Frequenzen $3f_1, 3f_2$

$$+ 0,75 a_3 [x_1^2 x_2 \cos(2\omega_1 \pm \omega_2)t$$

Frequenzen $2f_1 + f_2, 2f_1 - f_2$

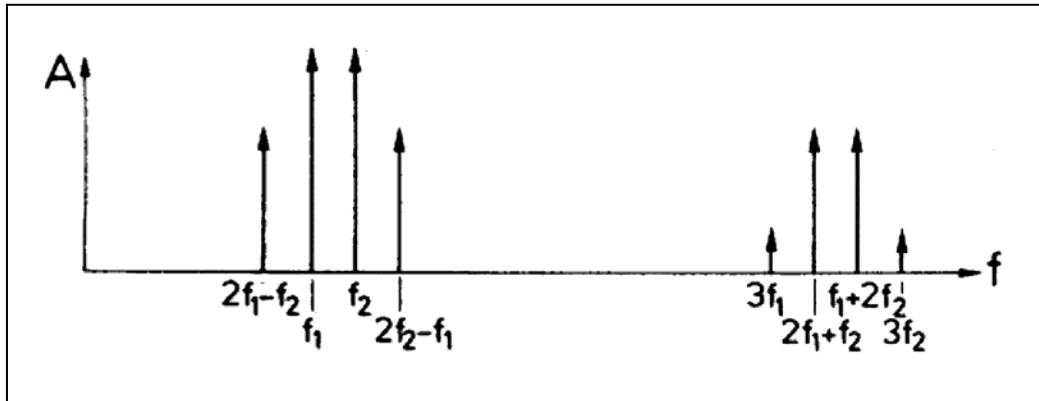
$$+ x_1 x_2^2 \cos(\omega_1 \pm 2\omega_2)t]$$

$2f_2 + f_1, 2f_2 - f_1$

Für den Fall $x_1 = x_2$ werden die Kombinationsschwingungen in ihrer Amplitude dreimal größer als die dritten Harmonischen, d.h. ihre Pegel unterscheiden sich um fast 10 dB. Dies ist mit ein Grund dafür, dass auch bei kubischen Verzerrungen der tiefe Differenzton (bei $2f_1 - f_2$) das am häufigsten hörbare Verzerrungsprodukt ist. Bei tiefer Frequenzlage der Primärtöne ergeben sich allerdings bei mittleren Abhörpegeln auch Fälle, bei denen die Summentöne stören, weil sie wegen größerem relativen Frequenzabstand von den Primärtönen (verglichen mit den Verhältnissen bei quadratischen Verzerrungen) nicht immer verdeckt werden.

Wie die kubischen harmonischen Verzerrungen hängt auch bei den kubischen Summen- und Differenztönen deren Amplitude kubisch bzw. deren Klirrfaktor (einmal durch Eingangsamplitude teilen) quadratisch von der Amplitude des Eingangssignals ab.

Verzerrungsprodukte zweier Sinustöne über eine Kennlinie mit kubischen Verzerrungen:



Beachte: Quadratische Verzerrungen wachsen linear mit der Aussteuerung, die kubischen Verzerrungen sogar quadratisch mit der Aussteuerung an.

Möglichkeiten zur Reduzierung der Hörbarkeit nichtlinearer Verzerrungen:

- Vor dem Bereich mit den stärksten Verzerrungen kann die Aussteuerung abgesenkt und danach wieder angehoben werden. Allerdings wird die Dynamik, d.h. die Pegeldifferenz zwischen maximalem Signalpegel und dem Geräuschpegel dadurch ebenfalls geringer.
- Der Hörer selbst kann lediglich den Abhörpegel verringern. Dabei werden auch die Verzerrungsprodukte (insbesondere kubische) verkleinert, so dass sie sich der Hörschwelle annähern oder sogar unter ihr verschwinden.

Hörbarkeit: Musik: $k = 0,5 \%$ (Flötenduett, Klavierstücke) bis 1% (Orchester)

Sprache: $k = 5 \%$ bis 10% , ab 20% störend

Bei Kopfhörer sind diese Grenzen etwas tiefer, da höherer Abhörpegel und geringerer Störpegel.

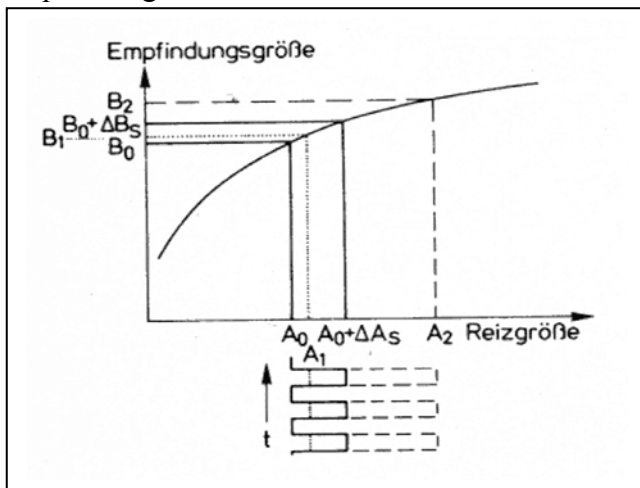
3. Psychoakustik

3.1 Reiz und Empfindung

Jeder physikalische Reiz löst eine psychische Empfindung aus, sofern er von einem Sinnesorgan wahrgenommen wird. Die Fähigkeit, sich auf einzelne Komponenten der Empfindungen zu konzentrieren und sie getrennt beurteilen zu können, nützen wir aus und bezeichnen die Empfindungskomponenten als Empfindungsgrößen. So wie Reizgrößen als Produkt aus Zahlenwert und Einheit definiert sind, werden auch Empfindungsgrößen Zahlenwert und Einheit zugeordnet. Jede Empfindungsgröße wird von allen Reizgrößen beeinflusst, im Allgemeinen dominiert jedoch eine Reizgröße. Der Zusammenhang zwischen einer dominierenden Reizgröße und der zugehörigen Empfindungsgröße wird als Empfindungsfunktion bezeichnet.

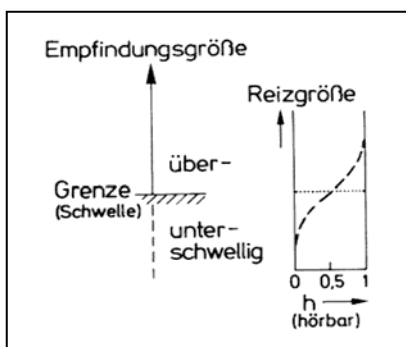
Die Zusammenhänge zwischen den verschiedenen Reizgrößen und den verschiedenen Empfindungsgrößen in eindeutiger Form anzugeben, ist die wichtigste Aufgabe der Psychoakustik.

Empfindungsfunktion:

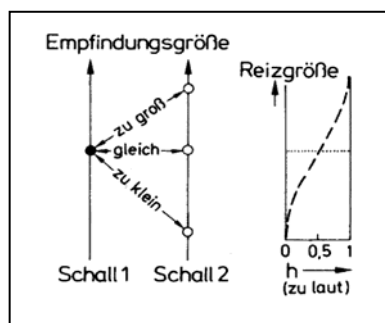


Wenn die Reizgröße sich nur um einen sehr kleinen Betrag ändert, kann die Empfindung dieselbe bleiben. Eine gerade wahrnehmbare Reizgrößenänderung wird als Reizstufe ΔA_s bezeichnet, die eine Empfindungsstufe ΔB_s hervorruft. Die Empfindungsgrößen, die anwachsen können, werden als Intensitätsempfindungen, diejenigen, die mit Orten zusammenhängen, an denen sie wahrgenommen werden, werden als Positionsempfindungen bezeichnet. Während die Empfindungsstufen für die Intensitätsempfindung von der Empfindungsgröße selbst abhängig sind, ist die Empfindungsstufe bei den Positionsempfindungen konstant. Die Reizstufen sind bei beiden von der Reizgröße selbst stark abhängig. Diejenige Reizgröße, bei der eine zu untersuchende Empfindungsgröße überhaupt erst ausgelöst wird, bezeichnet man als Grenzwert oder Schwellenwert. Durch Hörversuche können Grenzwerte, Vergleichswerte oder Verhältniswerte ermittelt werden, um eine Empfindungsfunktion zu konstruieren.

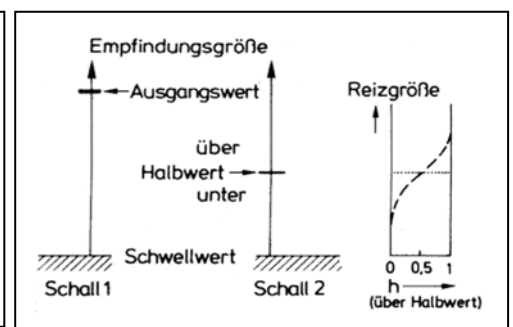
Grenzwert:



Vergleichswert:

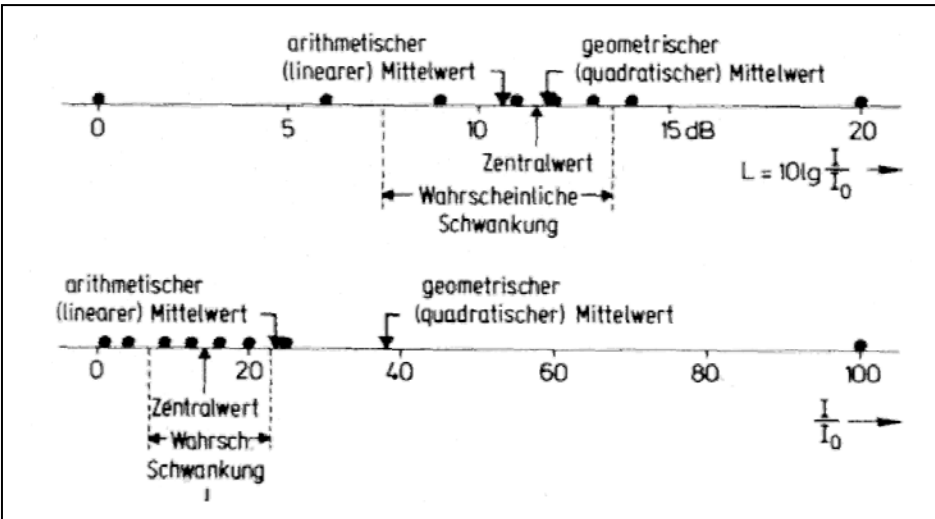


Verhältniswert:



Die Werte der Reizgrößen können meist sehr genau angegeben werden, während die Genauigkeit, mit der eine einzelne Versuchsperson eine Empfindungsgröße bestimmen kann, meist sehr viel schlechter ist. Die Messungen müssen daher mehrmals wiederholt werden, um Schwankungen derselben Versuchsperson (= intraindividuelle Schwankungen) ermitteln bzw. ausmitteln zu können. Abweichungen der Mittelwerte verschiedener Versuchspersonen (= interindividuelle Schwankungen) werden durch eine entsprechend große Anzahl von Testpersonen (mindestens acht) minimiert. Da die Angaben der Versuchspersonen nicht immer linear sind, sollte eine gegen nichtlineare Transformation invariante Mittelwertbildung verwendet werden. Hierzu ist der Zentralwert (Median) mit Wahrscheinlicher Schwankung gut geeignet.

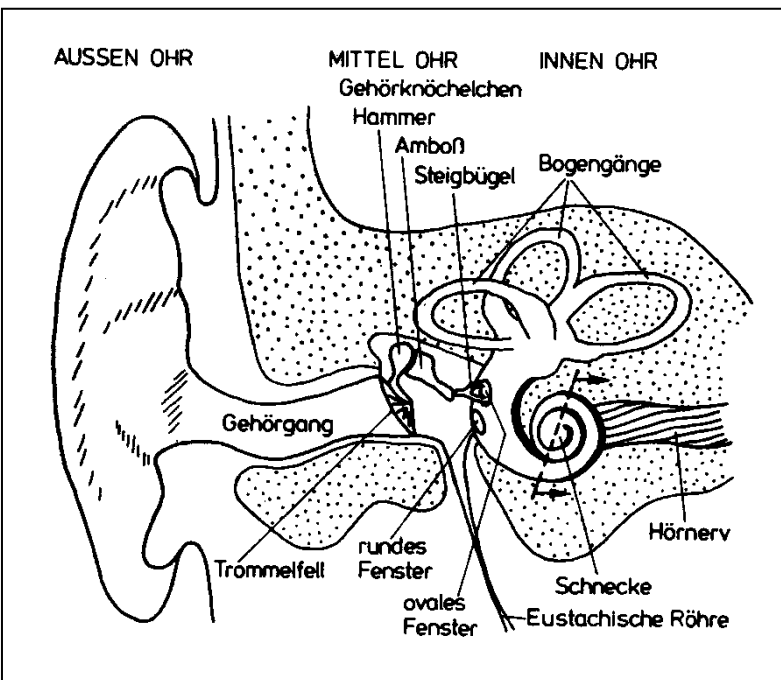
Zentralwert im Vergleich zu arithm. und geom. Mittelwert:



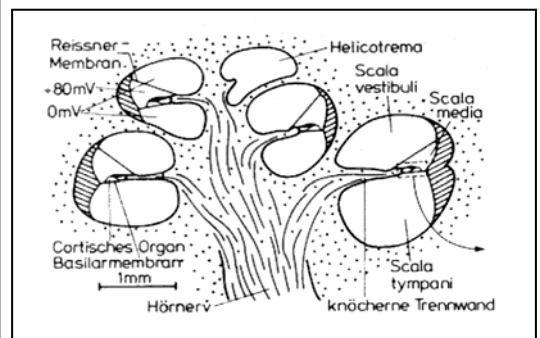
Beispiel: Ermitteln Sie ZW und W. Schw. der Werte:
900, 350, 1000, 900, 1300, 800, 1100, 1150 Hz

3.2 Hörphysiologie

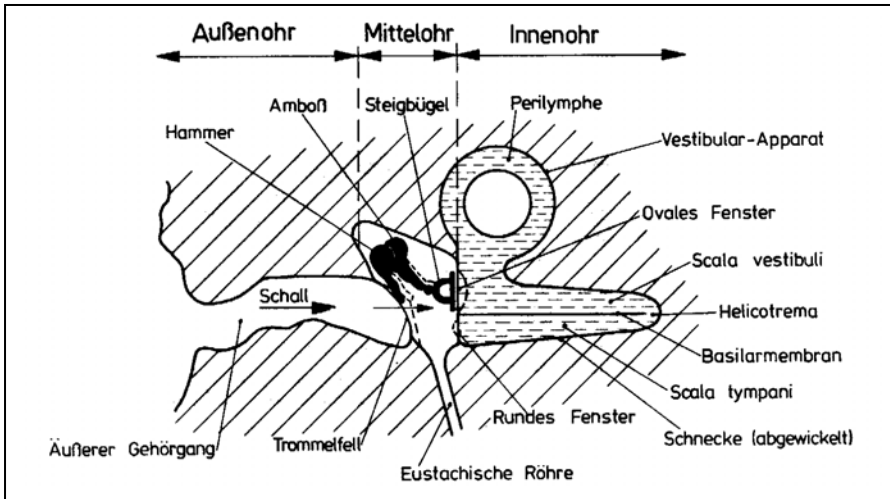
Aufbau des menschlichen Ohres:



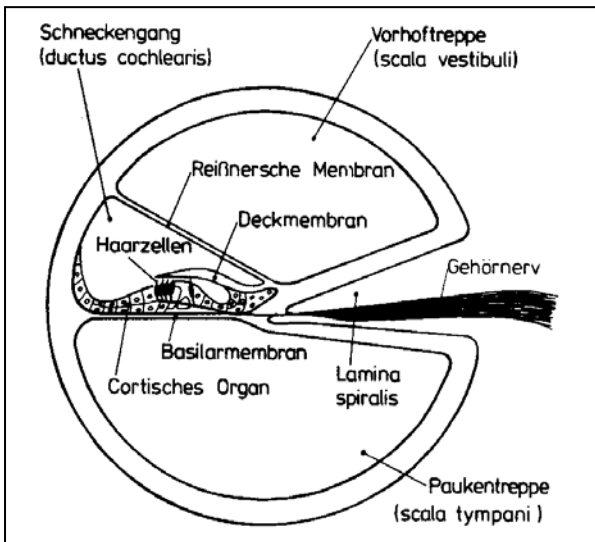
Schnitt durch Innenohr:



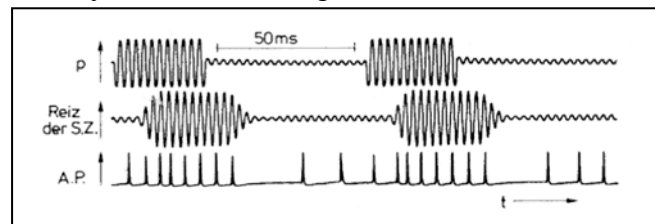
Schematische Darstellung von Außen-, Mittel- und Innenohr:



Schnitt durch die ausgerollte Schnecke:



Reizsynchrone Aktionspotentiale:



Das **Außenohr** besteht aus der Ohrmuschel (Pinna), welche eine charakteristische Außenohrübertragungsfunktion erzeugt, die zur Richtungswahrnehmung verwendet wird. Der Gehörgang leitet wie ein einseitig geschlossenes Rohr (Länge: 2 cm) Frequenzen zwischen 3 und 5 kHz besonders gut zum Trommelfell ($\lambda/4$ -Resonanz). Der Schalldruck bringt das Trommelfell stark zum Schwingen. Die Gehörknöchelchen Hammer, Ambos und Steigbügel leiten im **Mittelohr** die Schallwelle zum ovalen Fenster, hinter dem sich in der Scala vestibuli Lympheflüssigkeit befindet. Die Gehörknöchelchen bewirken wie eine hydraulische Presse eine Anpassung des geringen Schallwellenwiderstandes in der Luft ($Z_0 = 414 \text{ Ns/m}^3$) an den hohen Wellenwiderstand der Flüssigkeit ($Z_{Fl} = 1,48 \cdot 10^6 \text{ Ns/m}^3$) durch:

1. Flächenübersetzung: große Amplitude des Trommelfells \rightarrow kleine Ampl. des ovalen Fensters
2. Hebel: kleine Kraft des Trommelfells \rightarrow große Kraft am ovalen Fenster

Die Masse der Gehörknöchelchen wirkt bereits ab 2 kHz, wird aber durch die Gehörgangresonanz ($\lambda/4$) bei mittleren Frequenzen ausgeglichen, bewirkt aber bei 20 kHz eine Dämpfung um 60 dB.

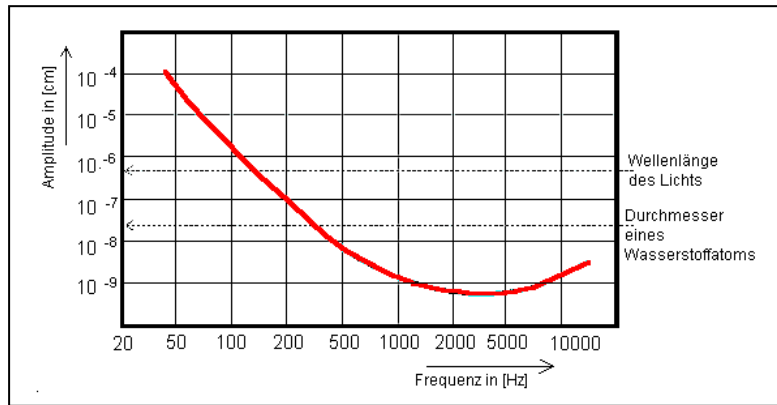
Die Eustachische Röhre ermöglicht im Innenohr einen Druckausgleich: 20 m \rightarrow 140 dB!

Die Umwandlung der Schallwellen in Nervenimpulse erfolgt im **Innenohr**. Die Cochlea besitzt beim Menschen etwa $2 \frac{1}{2}$ Windungen und besteht aus drei parallelen Kanälen, auch Skalen ge-

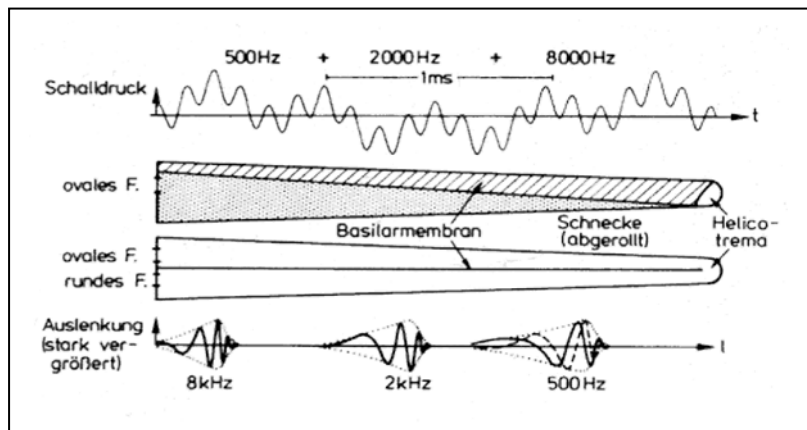
annt. Die Scala vestibuli, welche durch das ovale Fenster angeregt wird, ist von der Scala media nur durch eine dünne Membrane, die sog. Reissner-Membrane getrennt. Für hydromechanische Betrachtungen kann diese Membran als nicht vorhanden angesehen werden. Die Basilarmembrane grenzt wiederum die Scala media zur Scala tympani ab. Auf ihr sitzen die eigentlichen Sinneszellen. Die Basilarmembrane ist in der Nähe des ovalen Fensters breiter und wird in Richtung Helicotrema schmaler. Das Helicotrema verbindet die Scala vestibuli mit der Scala tympani und lässt so mit dem runden Fenster einen Druckausgleich zu.

Auf der Basilarmembrane befindet sich das Cortische Organ mit einer Reihe innerer Haarzellen und drei Reihen äußerer Haarzellen. Pro Reihe sind etwa 3600 Sinneszellen aneinandergereiht. Beim Schwingen der Basilarmembrane entstehen Scherkräfte (Reiz) auf die Haarzellen, die in Aktionspotenziale umgewandelt werden und durch die 30000 Nervenfasern zur Auswertung ins Gehirn weitergeleitet werden. Hohe Frequenzen erregen die Basilarmembrane in der Nähe des ovalen Fensters, tiefe Frequenzen nahe am Helicotrema. Es entsteht also eine Frequenz-Orts-Transformation. Die Tonhöhenempfindung ist daher eine Positionempfindung. Die Sinneszellen an den verschiedenen Positionen auf der Basilarmembrane sind nur für die ihnen aufgrund ihrer Position zugeordneten Frequenzen empfindlich (charakteristische Frequenz). Eine Schallschwingung wandert auf der Basilarmembrane entlang und verursacht nur bei einer bestimmten Position einen Maximalausschlag des Basilarmembrane. Man spricht von einer Wanderwelle.

Amplitude des Trommelfells:



Frequenz-Orts-Transformation:
Wanderwelle



Frequenz-Orts-Transformation:
Distanz vom ovalen Fenster

