

# Grundlagen der Akustik und Audiotechnik

Sprache, Musik und Geräusche - hiermit erleben wir unsere Umwelt, kommunizieren und bringen Gedanken, Gefühle und Ideen zum Ausdruck.

"Nicht sehen können trennt von den Dingen – nicht hören können trennt von den Menschen"



In diesem Skript finden Sie:

Was ist Schall???	3
Mathematische Beschreibung von Schallwellen	4
Schalldruck und Schalldruckpegel	5
Wertebereich des verarbeitbaren Schalldrucks	5
Definition des Schalldruckpegels	5
Bezugsschalldruck	5
Schalldruckpegel ist nicht „Lautstärke“	5
Hörphysiologie und Hörpsychologie	6
Hörphysiologie	6
Außenohr	6
Mittelohr	7
Innenohr	7
Wahrnehmung	8
Maskierungen	12
Räumliches Hören	13
Lokalisation von Schallquellen	13
Raumwahrnehmung	14
Auswirkung von Amplitude und Frequenz auf die Hörwahrnehmung	16
Hörbeispiele	16
Ton & Klang	17
Aufbau eines Klages	18
Bezeichnung der Klangkomponenten	18
Geräusche	19
Schwingungsverlauf von Geräuschen	19
Darstellung durch das Frequenzspektrum	20
Die Klangfarbe	21
Schalldruckpegel und Schallpegeladdition	22
Amplitudenmodulation	23
Frequenzmodulation	24
Audiotechnik	25
Qualitätskriterien	25
Frequenzgang, Bandbreite	25
Klirrfaktor (THD Total Harmonic Distortion):	26
Störabstand (Rauschabstand, „Signal/Noise-Ratio“ SNR, Fremdspannungsabstand)	27
Latenzen	27
Aufnahmetechnik	27
Mikrofon-Aufnahmetechnik: Elektromechanische Wandler	27
Richtcharakteristik	31
Weitere Tonabnehmer	31
Klangsynthese / Klangerzeuger	32
Wiedergabetechnik: Lautsprecher und Kopfhörer	33
Lautsprecherboxen	34
Kopfhörer	35
Ton-Aufnahmen in der Praxis	36
Popschutz	36
Popschirm	36
Windschutz	37
Maßnahmen gegen Körperschall	37
Aussteuerung und Anzeigeskalen	38
Raumakustik	39
Ton-Bearbeitung nach der Aufnahme:	39

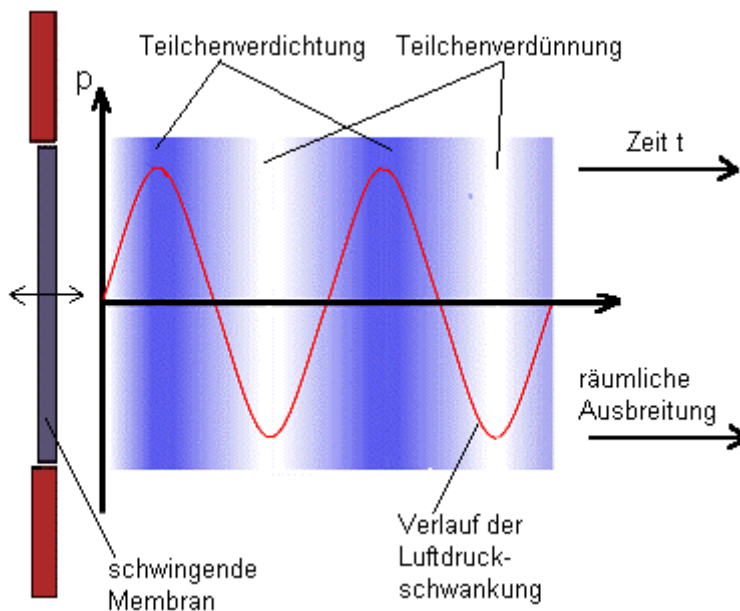
Sound-Effekte.....	40
Equalizer: .....	40
Flanger: .....	41
Phasing: .....	42
Ringmodulator:.....	42
Echo:.....	43
Literatur: .....	44

# Was ist Schall???

Die Grundlage jeder akustischen Wahrnehmung sind Schallwellen. Das sind Schwingungen, die sich durch die Luft als Druckwellen ausbreiten. Je nach Schwingungsform und Intensität nehmen wir diese in unterschiedlicher und vielfältiger Weise wahr. Damit Schall entstehen und sich ausbreiten kann, bedarf es also einer *Schallquelle* und eines *elastischen Mediums*, in dem sich der Schall fortpflanzen kann.

Steht die Schallquelle mit einem elastischen Medium in Verbindung, z.B. der Luft, so überträgt sie ihre Schwingungen auf die umgebenden (Luft-)moleküle. Die so angeregten Teilchen übertragen ihre Schwingungen wiederum auf ihre Nachbarmoleküle, so dass sich die von der Schallquelle ausgehende Erregung im gesamten Raum ausbreitet.

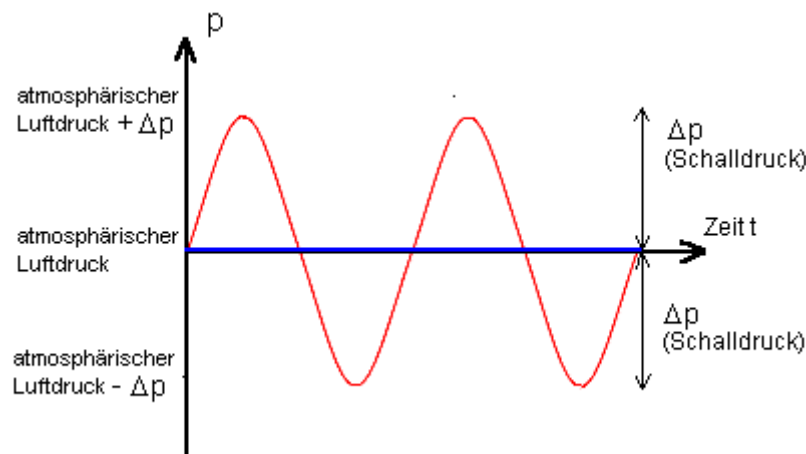
Um den Vorgang der Schallausbreitung zu veranschaulichen, werden in der nachfolgenden Animation die Luftmoleküle durch Masseklötzchen dargestellt. Die elastische Kopplung zwischen den Teilchen, wird durch die Federn angedeutet. (Anklicken zum animieren)



Durch die Anregung der Luftmoleküle kommt es zu Zonen mit Verdichtungen und Verdünnungen der Teilchenabstände, die aufeinander folgen und sich wellenartig im Raum ausbreiten. Die Verdichtung der Molekülabstände verursacht einen Luftdruckanstieg gegenüber dem schon vorhandenen atmosphärischen Luftdruck. Analog wird durch die Verdünnung der Teilchenabstände ein niedrigerer Luftdruck erzeugt.

Auf diese Weise entstehen Luftdruckschwankungen, die dem schon vorhandenen atmosphärischen Luftdruck überlagert und als Schalldruck  $[p]$  bezeichnet werden. Da sich dieser Vorgang wellenförmig ausbreitet, spricht man von der Entstehung einer Schallwelle.

Die obige Grafik zeigt den Verlauf der Schalldruckamplitude in Abhängigkeit von der Teilchendichte. Die untere Grafik zeigt die Überlagerung von Schalldruck und atmosphärischem Luftdruck:



**Schall ist die sich wellenartig ausbreitende räumliche und zeitliche Druckänderung eines elastischen Mediums**

(Anmerkung: Der Schalldruck  $p$  ist also der *Schallwechseldruck als Effektivwert*.)

## Mathematische Beschreibung von Schallwellen

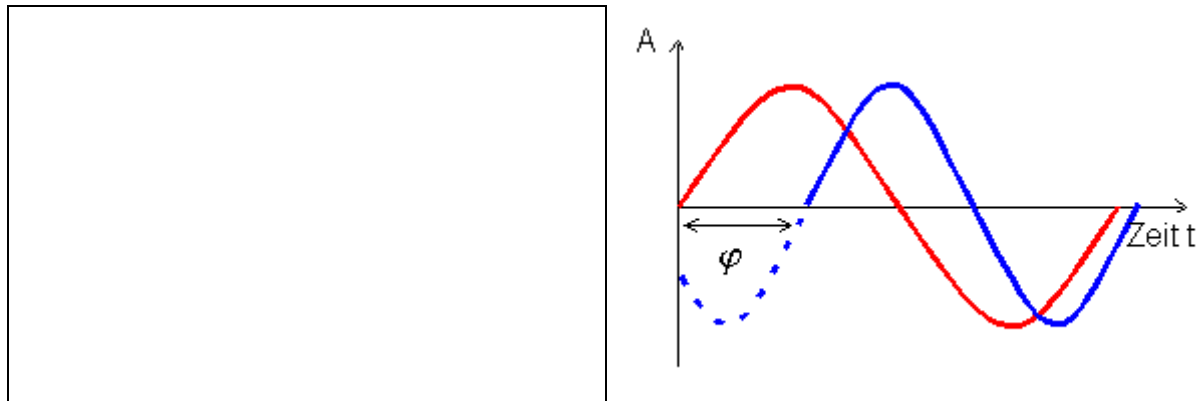
Da Schall auf Schwingungen beruht, lassen sich Schallereignisse durch ihren zeitlichen Schwingungsverlauf darstellen und beschreiben. Die einfachste Schwingungsform liegt mit der harmonischen oder sinusförmigen Schwingung vor, die man mit einem mathematischen Pendel oder mit bestimmten Schallquellen, z.B. einer Stimmgabel erzeugen kann.

Trägt man den Verlauf einer sinusförmigen Schwingung über der Zeit auf, so erhält man einen typischen Verlauf, der eindeutig bestimmt wird durch die Größen

- **Periodendauer T:** zeitliche Dauer eines Schwingungsverlaufs
- **Amplitude A:** Maximale Auslenkung der Schwingung
- **Phase  $\varphi$ :** zeitliche Verschiebung zweier Schwingungsverläufe zueinander

Aus Gründen der Verständlichkeit wird statt der Periodendauer **T** die Frequenz **f** zur Beschreibung der Schwingungsanzahl pro Sekunde verwendet. Diese ergibt sich direkt aus dem Kehrwert der Periodendauer:

- **Frequenz f:**  $f = 1/T$ : Schwingungsanzahl pro Sekunde. Die Einheit der Frequenz ist das Hertz [Hz]



Für die Ausbreitung der Schallwellen im Raum ist weiterhin die **Schallgeschwindigkeit c** und die **Dämpfung** im jeweiligen elastischen Medium interessant und für die räumliche Hörwahrnehmung wichtig. Die Schallgeschwindigkeit ist temperaturabhängig und beträgt in der Luft bei 20 °C  $c = 343 \text{ m/s}$  (=1234,8 km/h, 330 m/s bei 0 °C, Faustformel für Luft: 1 °C Temperaturerhöhung bewirkt eine um 60 cm/s erhöhte Schallgeschwindigkeit). (*c* kommt von lateinisch *celeritas*, zu deutsch: Eile, Schnelligkeit)

Die Schallgeschwindigkeit ist in dispersiven Medien (Flüssigkeiten, Wasser) frequenzabhängig, in nicht dispersiven Medien (Gasen, Luft) von der Frequenz unabhängig. Der statische, atmosphärische Grund-Luftdruck (ca. 95 – 104 kPa) hat keinen Einfluss auf die Schallgeschwindigkeit – bei gleicher Temperatur ist also der Schall „auf Meereshöhe“ genauso schnell wie auf der Zugspitze. Lediglich die Luftfeuchtigkeit hat noch eine – allerdings geringe, vernachlässigbare – Auswirkung auf die Schallgeschwindigkeit.

Zum Vergleich: Schallgeschwindigkeiten bei 20 °C in verschiedenen Materialien: *Eisen:* 5.170 m/s, *Diamant:* 18.000 m/s (!), *Wasser:* 1.484 m/s, *Helium:* 981 m/s, *Weich-PVC:* 80 m/s, *Gummi:* 100 m/s, *Holz:* ca. 3.000-5.000 m/s.

Praktische Auswirkungen: Sind an einer Tonerzeugung Luftsäulenschwinger, wie Holzbläser, Blechbläser oder Orgelpfeifen beteiligt, so ändert sich die Tonhöhe der Instrumente mit der Temperatur und wird als Verstimmung hörbar. Steigende Temperatur erwirkt dabei steigende Tonhöhe und umgekehrt. Beispielsweise ergibt eine Änderung der Temperatur um 1 °C etwa 0,75 Hz Frequenzänderung (Verstimmung) bei einer Tonhöhe von 440 Hz (Kammerton a').

Der Effektivwert des Schalldrucks  $\tilde{p}$  verhält sich im Freifeld (also ohne Schall-Reflexionen) umgekehrt proportional zur Entfernung *r* von einer (punktförmigen) Schallquelle (1/*r*-Gesetz, Abstandsgesetz):

$$\tilde{p} \sim \frac{1}{r} \qquad \frac{\tilde{p}_1}{\tilde{p}_2} = \frac{r_2}{r_1} \qquad \tilde{p}_1 = \tilde{p}_2 \frac{r_2}{r_1}$$

(Anmerkung: Die quadratischen Schallenergiegrößen, wie z. B. die Schallintensität, nehmen bei punktförmigen Schallquellen auch mit dem Quadrat der Entfernung ab.)

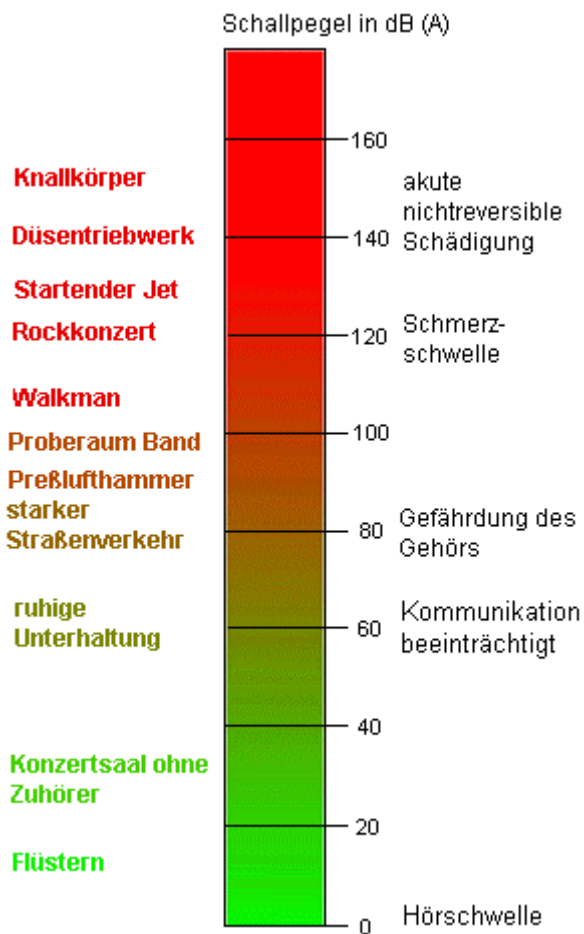
# Schalldruck und Schalldruckpegel

## Wertebereich des verarbeitbaren Schalldrucks

Als hörbarer Schall werden Schallereignisse im Frequenzbereich zwischen 20 Hz und 16000 Hz bezeichnet. Dabei kann das Gehör einen Schalldruckbereich von 20 µPa bis 20 Pa verarbeiten. Dies entspricht einem Verhältnis von 1:1.000.000.

Damit man diesen riesigen Wertebereich sinnvoll darstellen und beschreiben kann, wurde der *Schalldruckpegel* definiert. In dieser Darstellung beträgt dann der Wertebereich nur noch 0 dB – 120 dB, was eine wesentlich übersichtlichere Skalierung ergibt.

Ein weiterer Vorteil der Beschreibung des Schalldrucks durch den Schalldruckpegel liegt darin, dass die logarithmische Darstellung besser an die natürliche Schallverarbeitung des Gehörs angepasst ist. Er ist deshalb eine der zentralen Größen in der Akustik und wird zur Beschreibung der *Lautstärke* benötigt.



## Definition des Schalldruckpegels

Dieser ist als logarithmisches Maß für das Verhältnis zwischen dem *gemessenen Schalldruck* und einem *Bezugsschalldruck* definiert. Die Größe des Schalldruckpegels wurde eingeführt, damit man den Wertebereich des Schalldrucks, den das Gehör verarbeiten kann, mathematisch sinnvoll handhaben und darstellen kann. Eine Zuordnung zwischen Schalldruckpegel und entsprechender Lautstärkeempfindung soll durch die Grafik verdeutlicht werden. Aus der Grafik ist auch ersichtlich, dass das Gehör einen Wertebereich von 0 dB bis 120 dB verarbeiten kann.

Die Gleichung für den Schalldruckpegel lautet:

$$\text{Schalldruckpegel [dB]} = 20 \log \left( \frac{\text{Gemessener Schalldruck}}{\text{Bezugsschalldruck}} \right) \text{ dB}$$

(Häufig wird der Schalldruckpegel, obwohl dann physikalisch nicht eindeutig, auch einfach *Schallpegel* genannt.)

## Bezugsschalldruck

Der Bezugsschalldruck – also 0 dB – ist willkürlich, aber mit Absicht auf 20 µPa festgelegt. Ein sinusförmiger Ton von 1000 Hz ist bei diesem Pegel gerade noch hörbar.

## Schalldruckpegel ist nicht „Lautstärke“

Der Schalldruckpegel ist eine technische und keine psychoakustische Größe. Ein Rückschluss von Schalldruckpegel auf die wahrgenommene *Empfindung* ist nur eingeschränkt möglich. Generell gilt natürlich, dass eine Erhöhung bzw. Senkung des Schalldruckpegels auch eine lauter bzw. leiser wahrgenommene Hörempfindung hervorruft. Als Faustformel darf man annehmen, dass 10 dB Unterschied etwa als doppelte bzw. halbe Lautstärke wahrgenommen wird.

Wir müssen also noch einige Aspekte der Hörphysiologie und der Hörpsychologie untersuchen.

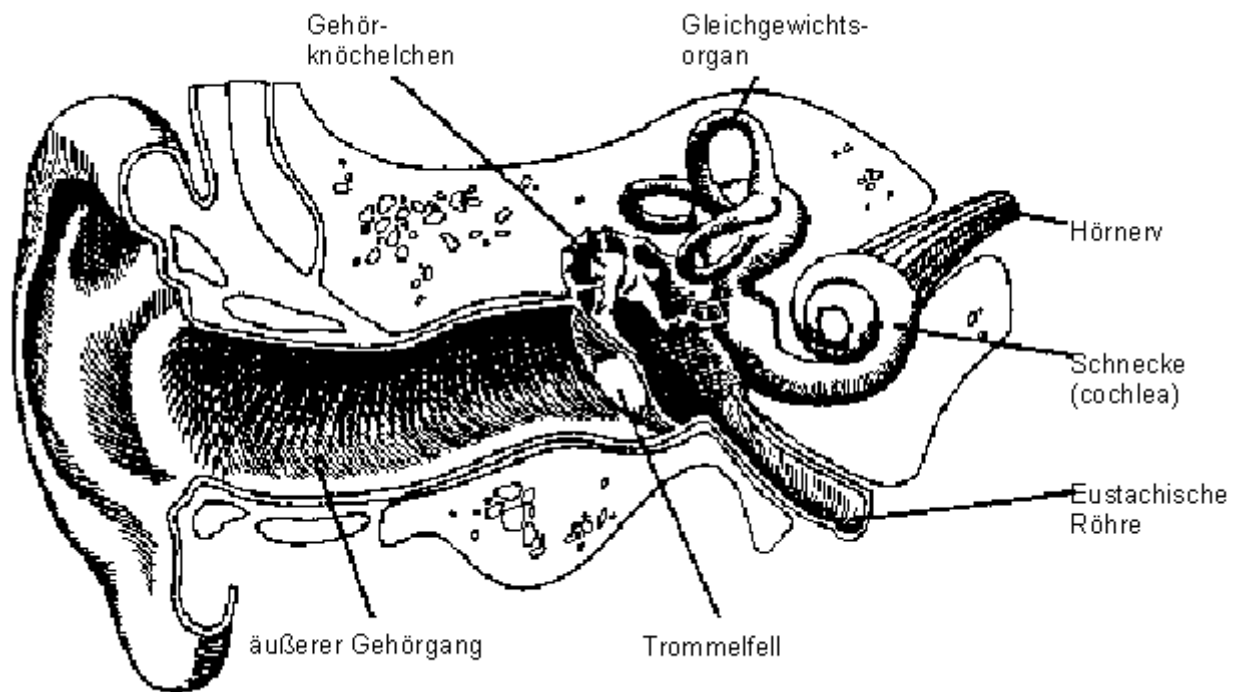
# Hörphysiologie und Hörpsychologie

## Hörphysiologie

Der Hörvorgang vollzieht sich in zwei Stufen. Zunächst wird der Schall aufgenommen, verstärkt, angepasst und weitergeleitet. Anschließend formen Sinneszellen die Schallinformation in elektrische Aktionspotentiale um, die über Neuronen dem Gehirn zugeleitet werden. Dort erfolgt dann die endgültige Auswertung, und es entsteht der Höreindruck.

Viele mit dem Hören zusammenhängende Fragen sind bis heute ungeklärt – dies gilt vor allem für die Verarbeitungsmechanismen im Gehirn. Erwiesenermaßen werden aber für die Auswertung der Höreindrücke neben dem aktuell einwirkenden Schallereignis auch die bisher erworbenen „Hörerfahrungen“ mit verwendet. Daraus erklärt sich beispielsweise die Tatsache, dass der Mensch auch noch Sprache versteht, die durch Lärm erheblich gestört ist. So genügt etwa in einer lärmgefüllten Maschinenhalle das Verstehen einiger Wortketten, um den Sinnzusammenhang des Gesprochenen zu erfassen. Ähnliches gilt auch für ein Telefongespräch über eine gestörte Leitung. Der Mensch ergänzt nicht verstandene Anteile – wie man sagt – aus dem *Kontext*. Dieses „Kontexthören“ ist so perfekt gelöst, dass wir mitunter schwören könnten, Laute oder Wörter wirklich gehört zu haben, die aber (physikalisch nachweisbar) gar nicht gesprochen wurden oder – beispielsweise durch Filterung – physikalisch nicht mehr im Sprachsignal vorhanden waren. Diese vom Menschen bis zur Perfektion entwickelten Erkennungsfähigkeiten lassen sich auch nicht im entferntesten durch technische Systeme erreichen.

Wir wollen einige anatomische Gegebenheiten des Ohres beschreiben und auf gewisse Hörphänomene eingehen. Dabei wird vor allem die frequenzabhängige Hörempfindlichkeit eine wichtige Rolle spielen. Bild 6 zeigt einen Schnitt durch das menschliche Ohr, das sich anatomisch und funktionsmäßig in drei Bereiche unterteilen lässt: das Außen-, Mittel- und Innenohr.



**Bild 6:** Schnitt durch das menschliche Ohr (FEL84)

### Außenohr

Das Außenohr umfasst den Bereich zwischen Ohrmuschel und dem Trommelfell und besteht im wesentlichen aus dem Gehörgang – einem Rohr von ca. 2,5 cm Länge, das am Ende durch das Trommelfell abgeschlossen ist. Der Gehörgang wirkt als Hohlraum-Resonator, wobei der Resonanz- und damit auch der Verstärkungsbereich etwa zwischen 2 und 4 kHz liegt. In diesem Frequenzbereich ist das Ohr also am empfindlichsten.

Der Schall trifft schließlich auf das Trommelfell und bringt es zum Schwingen. Die Schwingungsamplituden sind dabei unvorstellbar klein: für einen 1000-Hz-Ton, der gerade wahrnehmbar ist, liegen die Amplituden bei etwa 0,01 nm (!) und damit noch unterhalb des Durchmessers eines Wasserstoffatoms.

### Mittelohr

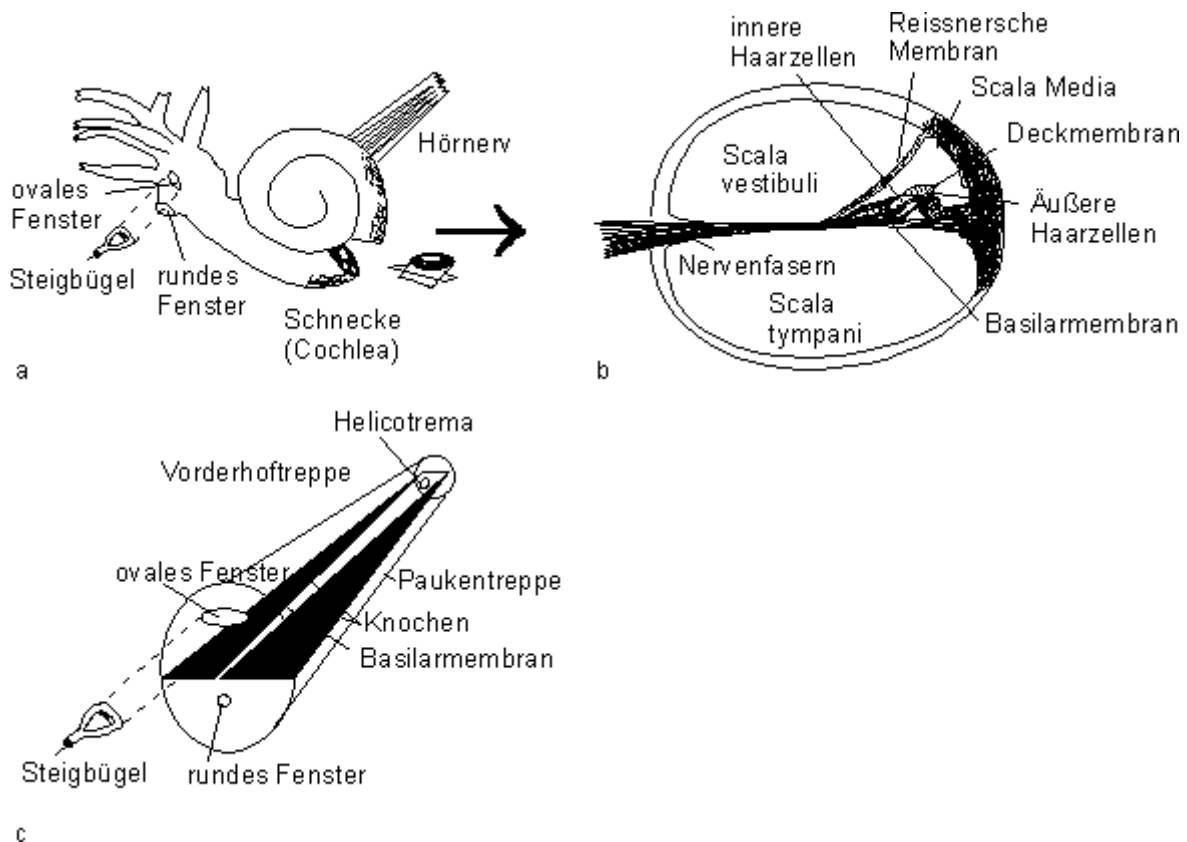
Das Mittelohr wird durch einen kleinen luftgefüllten Raum, die Paukenhöhle gebildet. Damit vor und hinter dem Trommelfell der gleiche (atmosphärische) Luftdruck herrscht, besteht eine Verbindung zwischen dem Mittelohr und dem Nasen-Rachenraum, die sogenannte *Eustachische Röhre*.

Im Mittelohr befinden sich die Gehörknöchelchen *Amboss*, *Hammer* und *Steigbügel*. Der Hammer ist am Trommelfell befestigt und überträgt die Trommelfell-Schwingungen auf den Amboss, der sie an den Steigbügel weitergibt. Der Steigbügel schließlich leitet die Schallschwingungen durch das *ovale Fenster* in das flüssigkeitsgefüllte Innere der *Schnecke*.

Die Gehörknöchelchen dienen in erster Linie der Impedanz-Anpassung zwischen der Luft und der Schnecken-Flüssigkeit. Sie haben außerdem eine - auf den Hebelgesetzen beruhende Verstärkungswirkung. Hinzu kommt schließlich noch eine Schutzfunktion: bei zu lauten Schallen wird die Bewegung der Knöchelchen durch Muskeln blockiert. Der Schutzmechanismus benötigt aber eine gewisse Reaktionszeit (ca. 60 bis 120 ms), so dass er bei plötzlich auftretenden Schalldrücken (Explosionen etc.) nicht wirksam werden kann.

### Innenohr

Das Innenohr besteht im wesentlichen aus der Schnecke (Cochlea), in der die Umsetzung der Schalldruck-schwankungen in Reize des Gehörnervs stattfindet. Bild 7 zeigt die Schnecke (Bild a) sowie einen Schnitt quer zur Schneckenwindung (Bild b). Man erkennt eine Unterteilung in drei Kanäle: die *Vorhof-treppe (scala vestibuli)*, die *Paukentreppe (scala tympani)* und einen kleinen Kanal in der Mitte, die *scala media*. Vorhof- und Paukentreppe sind an der Schneckenspitze durch eine kleine Öffnung, dem *Helicotrema*, miteinander verbunden. Am Ende der Paukentreppe, unterhalb vom ovalen Fenster, befindet sich noch eine durch eine feine Membran verschlossene Öffnung, das *runde Fenster*. Zur besseren Veranschaulichung ist in Bild (c) die Schnecke noch einmal ausgerollt dargestellt.



**Bild 7:** Schnecke (Cochlea). a) Gesamtansicht, b) Schnitt quer zur Schneckenwindung, c) Schnecke abgerollt (nach GRE76)

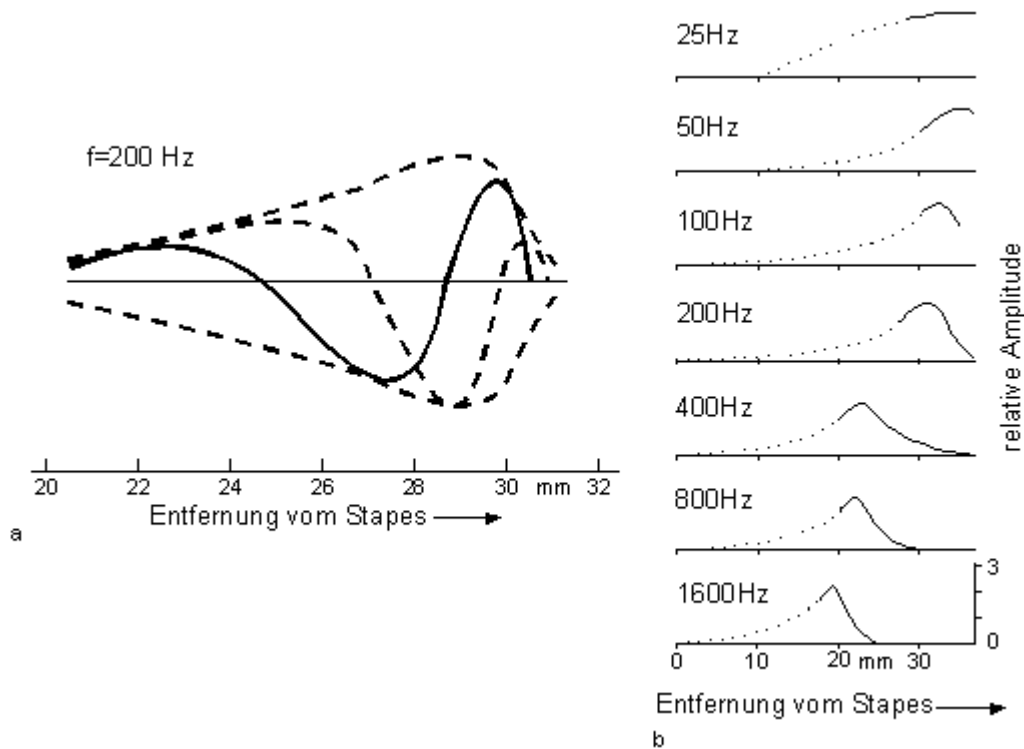
Das eigentliche Hörorgan befindet sich auf der Trennwand zwischen der scala media und der Paukentreppe. Diese Trennwand bezeichnet man als *Basilarmembran*. Auf dieser Membran sitzt das **Cortische Organ** mit seinen äußeren und inneren Haarzellen. Die Haarzellen stehen in Kontakt mit den Nervenfasern (enthalten aber selbst keine Nervenzellen!) und diese wiederum sind mit dem Gehörnerv verbunden, der im Zentral-Nervensystem endet.

Über dem Corti-Organ, auf der Spitze der Haarzellen aufliegend, ist die *Deckmembran* angeordnet.

Wird die Basilarmembran durch die Flüssigkeitswelle in Schwingungen versetzt, so entsteht an Stellen der Membrandurchbiegung eine Scherbewegung zwischen der Deckmembran und den Haarzellen. Hierdurch werden die zugehörigen Nervenzellen gereizt, der Reiz wird über den Hörnerv zum Gehirn geleitet und dort als Höreindruck wahrgenommen.

Wie bereits erwähnt, sind viele Hörphänomene bisher noch ungeklärt. Gesichert ist aber, dass die Flüssigkeitswelle je nach Frequenz unterschiedliche Gebiete der Basilarmembran durchbiegt. Da die Basilarmembran in der Nähe des Steigbügels schmal und straff ist, werden dort die hohen Frequenzen abgebildet, während am unteren Ende (in der Nähe des Helicotremas), wo die Basilarmembran breit und schlaff ist, tiefe Frequenzen lokalisiert werden.

Bild 8 zeigt den Wellenverlauf auf der Basilarmembran für zwei aufeinanderfolgende Zeitpunkte (Bild a) sowie die Ausprägung der Maxima in Abhängigkeit von der Frequenz (Bild b).



**Bild 8:** Verlauf der Flüssigkeitswelle im Innenohr. a) bei einer Frequenz für zwei Zeitpunkte, b) bei verschiedenen Frequenzen (nach BEK60)

## Wahrnehmung

Die Hörpsychologie befasst sich mit der Beschreibung von Hörphänomenen. Man betrachtet das Hörorgan als Nachrichtenempfänger und versucht, durch akustische Messungen, Höreigenschaften in reproduzierbarer Weise und objektiv zu beschreiben.

Für den Nachrichtentechniker ist vor allem der *Frequenzgang* von Interesse. Er gibt in unserem Fall die Abhängigkeit der Hörempfindlichkeit von der Frequenz an. Um die sogenannte **Hörschwelle** zu ermitteln, bestimmt man (mit Testpersonen) für jede Frequenz genau den Schallpegel, bei dem der zugehörige Ton gerade noch wahrnehmbar ist. Auf diese Weise erhält man die unterste Kurve in Bild 9.



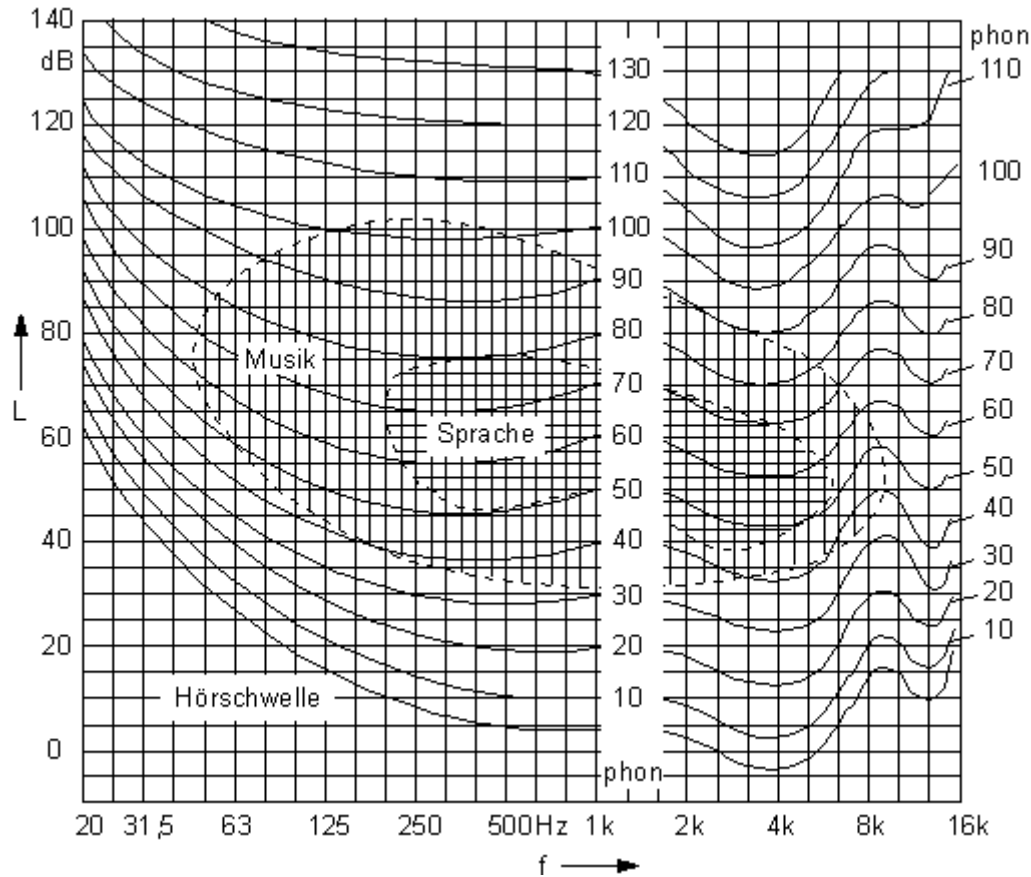


Bild 9: Hörfläche mit Kurven gleicher Lautstärke (FEL84)

So kann man generell feststellen, dass tiefe Töne mit einem wesentlich höheren Schalldruckpegel dargeboten werden müssen, um den gleichen Lautstärkeindruck zu erzeugen wie Töne im mittleren Tonhöhenbereich von etwa 500 bis 2000 Hz. Auch für die höheren Töne ist das Gehör im Vergleich zu den mittleren Tonlagen unempfindlicher, wenn auch nicht ganz so dramatisch. Dieser Effekt ist zudem bei kleinen Lautstärken viel deutlicher zu spüren als bei sehr lauten Darbietungen. Da diese Eigenschaft eine der wichtigsten des menschlichen Hörsinnes ist, hat man „Kurven gleicher Lautstärkepegel“ ermittelt, die abhängig von der Frequenz den Schalldruckpegel angeben, der die jeweils gleiche Lautstärkeempfindung erzeugt. Durchläuft also ein Sinuston den gesamten hörbaren Frequenzbereich, so muss für eine konstante Lautstärkeempfindung der Schalldruckpegel gemäß der in Bild 9 dargestellten Kurven variieren.

Mit den Kurven bekommt die Lautstärkeempfindung eine nach DIN 45 630 definierbare Einheit und heißt „phon“. Bei Sinustönen mit einer Frequenz von 1000 Hz stimmen der Schalldruckpegel in dB und der Lautstärkepegel in phon zahlenmäßig überein. Normal und beidohrig hörende Personen im Alter von etwa 18 bis 25 Jahren können Lautstärken von 4 phon gerade noch hören (Hörschwelle). Die größte Lautstärke bildet die Schwelle der Schmerzempfindung mit etwa 130 phon, Lautstärken in diesem Bereich verursachen schon bei relativ kurzer Dauer irreversible Hörschäden.

Die weiteren Kurven stellen die sogenannten *Isophonen* dar, das sind Kurven gleicher Lautstärke oder genauer: gleichen Lautstärke-Empfindens. Ihren Verlauf gewinnt man in der Weise, dass man dem Hörer einen Ton mit einem bestimmten Schallpegel bei einer Frequenz von 1 kHz anbietet und dann einen weiteren Ton bei einer anderen Frequenz, z.B. 500 Hz. Nun wird der 500-Hz-Ton solange in seinem Pegel verändert, bis der Hörer ihn als gleichlaut (wie den 1-kHz-Ton) empfindet. In gleicher Weise verfährt man mit anderen Frequenzen. Im Bild ist deutlich zu erkennen, dass alle Kurven bezüglich ihrer Hörempfindlichkeit über der Frequenz ähnlich verlaufen (die oberen Kurven sind nur etwas flacher). Für den Verlauf der Lautstärke längs einer Isophone hat man das Maß „phon“ eingeführt. Eine Angabe von z.B. 30 phon bedeutet daher sehr unterschiedliche Pegelwerte in Abhängigkeit von der Frequenz. Nur bei 1 kHz stimmen Pegelangabe (in dB) und phon-Angabe überein. Die oberste Isophone mit ca. 130 phon bezeichnet man als **Schmerzschwelle**. Damit soll ausgedrückt werden, dass die Schalle bereits schmerzhaft laut sind. Neben dem Schmerzempfinden tritt auch eine Schädigung der Haarzellen ein, die deshalb besonders gravierend ist, weil sich geschädigte Haarzellen nicht mehr regenerieren können. Treten die lauten Schalle daher über eine längere Zeit auf, so wird der Mensch unwiderruflich schwerhörig. Schallpegel an der Schmerzschwelle treten durchaus in der Praxis auf - beispielsweise dann, wenn man sich in einer Diskothek in die

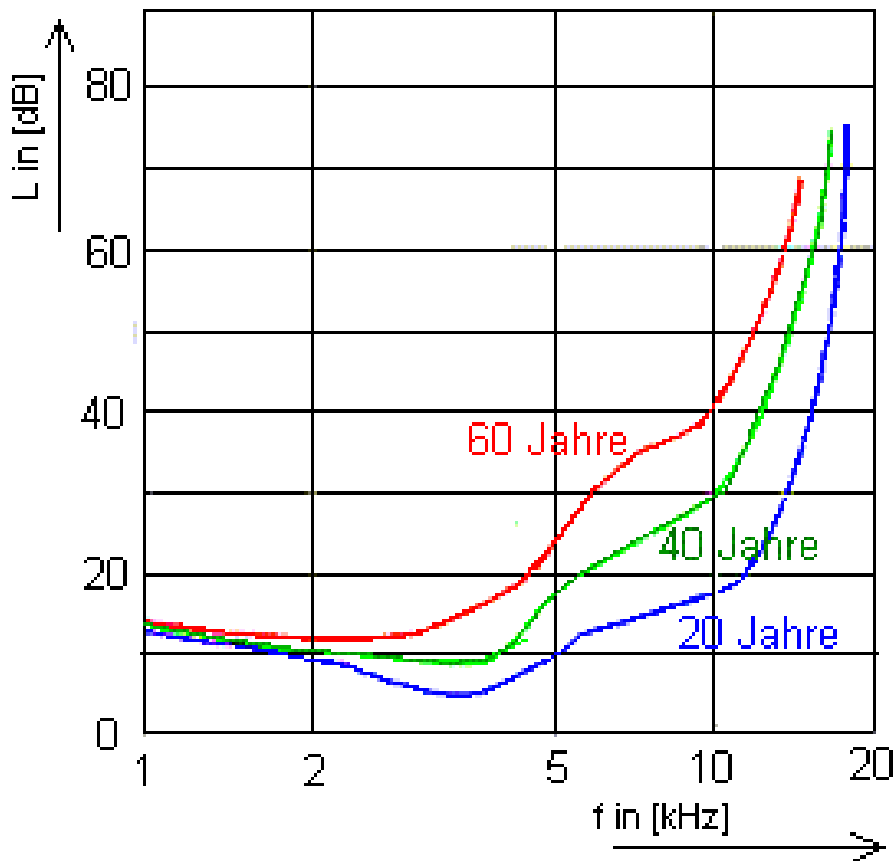
Nähe der Lautsprecher begibt. Aber auch die „Ohrstöpsel“ eines Walkmans oder MP3-Players können ohne weiteres schädigende Hörschalle produzieren; diese „Schallquellen“ sind also alles andere als harmlos!

Den gesamten Bereich zwischen Hör- und Schmerzschwelle bzw. im Frequenzumfang zwischen 20 Hz und ca. 16 kHz bezeichnet man als **Hörfläche**; in dieser spielt sich also das gesamte Hörgeschehen ab. Der Bereich der Sprache ist jedoch wesentlich begrenzter und im Bild schraffiert angegeben. Allerdings gibt es keine scharfe Begrenzung: Sprache reicht von etwa 60 Hz bis 8 kHz, in den Extrembereichen ist der Sprachpegel aber außerordentlich niedrig. Man kann daher den Frequenzumfang der Sprache auf etwa 300 bis 3400 Hz begrenzen, ohne dass die Sprachverständlichkeit darunter leidet. Beim Fernsprechen, wo man aus Kostengründen Bandbreite sparen muss, wird genau dieser Frequenzumfang verwendet. Für eine qualitativ hochwertige und damit natürliche Sprachwiedergabe sollte man aber den vorher genannten Bereich vorsehen.

Drei weitere Effekte sind zu beobachten bzw. zu erwähnen:

Zum einen verlaufen die Isophonen bei geringen Lautstärken wesentlich ausgeprägter mit einem steileren Kurvenverlauf sowohl in den Höhen als auch besonders in den Bässen. Dies bedeutet, dass die generelle Abnahme der Empfindlichkeit bei tiefen und hohen Tönen umso ausgeprägter ist, je leiser der Schall insgesamt ist. Bekannteste technische Konsequenz dieses Effekts ist übrigens der „Loudness“-Schalter bei den meisten heutigen Verstärkern. Diesen darf man bei leisen Darbietungen drücken, um die Bässe und Höhen ein wenig mehr anzuheben, damit sie gleich laut wie die mittleren Tonhöhen empfunden werden.

Zum anderen ist die Wahrnehmbarkeit gerade der hohen Frequenzen stark vom Lebensalter abhängig:



**Bild X:** unterschiedliche Hörschwellen je nach Lebensalter

Und drittens: Frauen hören besser als Männer, da der äußere Gehörgang besser geformt und kleiner ist. Daher haben Frauen eine um einige Prozent erhöhte Gehörgangsresonanz. Vor allem bei hohen Frequenzen liegen Hörschwellen bei Frauen niedriger als bei Männern. Und im Alter nimmt die Hörleistung weniger stark ab.

Man könnte noch einen vierten Effekt nennen: es kommt darauf an, in welchem Jahrzehnt man lebt... Die „Kurven gleicher Lautstärke“ werden mit verschiedenen Testpersonen und unter verschiedenen Testbedingungen neu vermessen und definiert – die letzte Erkenntnis und damit Änderung der internationalen Norm fand 2003 statt:

This International Standard specifies combinations of sound pressure levels and frequencies of pure continuous tones which are perceived as equally loud by human listeners. The specifications are based on the following conditions:

The sound field in the absence of the listener consists of a free progressive plane wave. The source of sound is directly in front of the listener. The sound signals are pure tones. The sound pressure level is measured at the position where the centre of the listener's head would be, but in the absence of the listener - listening is binaural. The listeners are ontologically normal persons in the age range from 18 years to 25 years inclusive.

The equal loudness level values are significantly higher than those of the old ISO 226, particularly below 1 kHz.

For the differences to the previous (**dashed blue line**) version of ISO 226:1987 curves look at the

**New (red line) revised curves of ISO 226:2003**

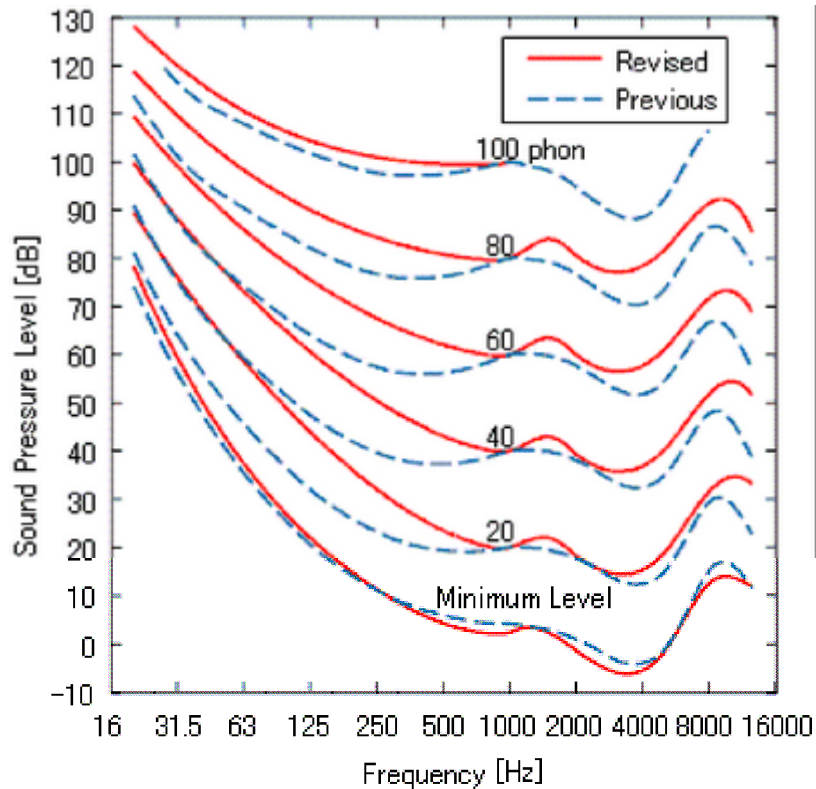


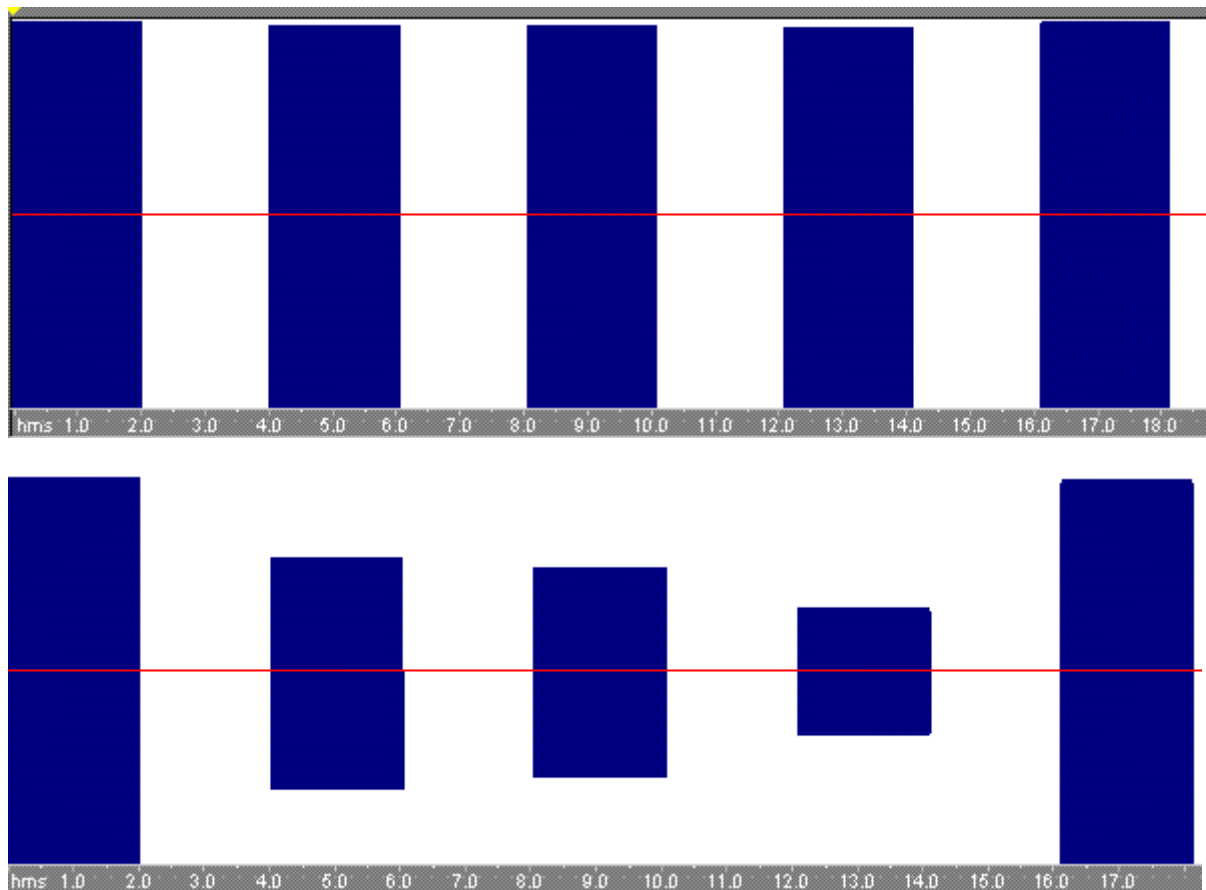
Bild XY: Neue Isophone in ISO 226:2003 Acoustics

### Audiobeispiel:

Wir wollen die frequenzabhängige Empfindlichkeit des Ohres anhand eines kleinen Experimentes demonstrieren. Hierzu hören wir zunächst eine Folge von Tonsignalen mit zunehmender Frequenz, jedoch mit konstantem Schallpegel. Bild 10a zeigt den eben erläuterten Verlauf.

 [PC-Version \(sin.wav\)](#) [Workstation-Version \(sin.au\)](#)

Es ist deutlich zu hören, dass zu Beginn (also bei tiefen Frequenzen) die Töne sehr leise sind. Mit steigender Frequenz empfinden wir die Töne als lauter - am lautesten im Bereich von etwa 2 bis 3 kHz. Anschließend nimmt das Lautstärkeempfinden wieder ab. Die Abnahme in Richtung hoher Frequenzen ist nicht so deutlich wahrnehmbar; da aus technischen Gründen bei unseren Experimenten das obere Frequenzband auf etwa 6 kHz begrenzt wurde.



**Bild 10:** Darbietung von verschiedenen Tönen, a) mit gleichem *Pegel* und b) mit gleicher *Lautstärke*. Die Töne erscheinen hier als Balken: die Breite entspricht der Darbietungsdauer, die Höhe dem Schallpegel.

In einem weiteren Experiment wollen wir die Schallpegel der Töne so einzustellen, dass bei allen Tönen der gleiche Lautstärkeindruck entsteht. Bild 10b zeigt das Ergebnis. Die unterschiedlichen Balkenhöhen sind dabei als Punkte auf einer Isophone interpretierbar.

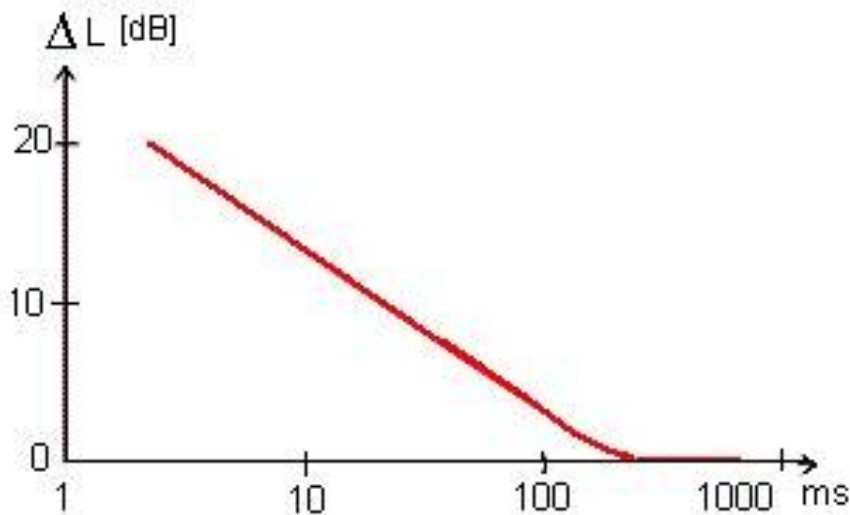
 [PC-Version \(sin-iso.wav\)](#) [Workstation-Version \(sin-iso.au\)](#)

## Maskierungen

Weitere interessante Phänomene sind die **Anpassung** und die **Verdeckung**.

Dank unserer menschlichen Evolution ist das Ohr trainiert, auf unerwartet eintretende Geräusche besonders empfindlich zu reagieren. Schon ein leises Knacken im Unterholz hat uns seinerzeit davor bewahrt, von einem plötzlich erscheinenden Löwen gefressen zu werden. Deshalb schrecken manche Menschen auch nachts bei kleinsten Geräuschen bereits aus dem Schlaf. Andererseits werden sogar relativ laute, aber gleichbleibende Geräusche, wie das Rauschen des Windes, Motorenlärm oder kontinuierliche Musikberieselung, bald als harmlos oder ermüdend eingestuft und vom Ohr nach einer gewissen Zeit als Hintergrundgeräusch nicht mehr wahrgenommen.

Diese über einen längeren Zeitraum erfolgende Anpassung erscheint uns aus eigener Erfahrung recht plausibel. Erstaunlich ist jedoch, dass die Ermüdung des Hörsinns bereits nach wenigen Millisekunden erfolgt.



**Bild X:** Schalldauer und Lautheitsempfinden

Diese Anpassung des Gehörs ist auch Ursache für eine weitere Eigenschaft, die **Verdeckung** genannt wird. Wenn bereits ein Schallereignis auf das Ohr einwirkt, wird die Empfindlichkeit für weitere akustische Reize vermindert. Wenn wir in einem ruhigen Zimmer noch das Ticken einer Uhr deutlich hören können, werden wir bei gleichzeitig geöffnetem Fenster und relativ dezenten Umweltgeräuschen von der Uhr wohl nichts mehr vernehmen. Ein Glück auch für Orchestermusiker, dass nicht jeder falsch gespielte Ton sofort auffällt, sondern gnädig verdeckt wird. Interessant ist, dass diese Verdeckung nicht nur exakt für die Dauer des verdeckenden Schallereignisses auftritt, sondern auch noch eine kleine Weile danach und sogar bereits davor! Je nach Lautstärke, Tonhöhe und Dauer zwar nur im Bereich von Sekundenbruchteilen, aber dieser Effekt reicht bereits aus, um die Datenmenge von Audioaufnahmen drastisch zu reduzieren.

Weiterer Inhalt zu Verdeckung folgt sowie Beispiele.

## Räumliches Hören

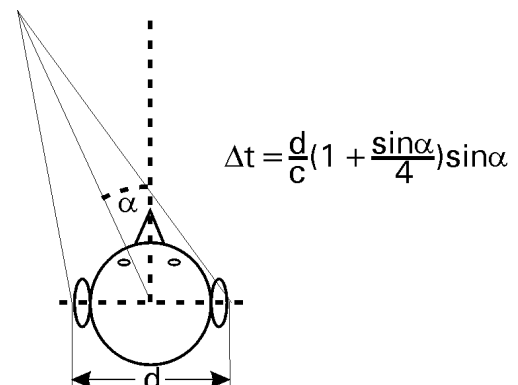
### Lokalisation von Schallquellen

Nicht zu vergessen ist natürlich die Fähigkeit des Hörsinnes, räumlich zu hören und damit Schallquellen zu lokalisieren: je nach Position der Schallquelle werden rechtes und linkes Ohr von Schallwellen unterschiedlich stimuliert („interaurale Differenzen“). Die Schallwellen, die von einer im Idealfall punktförmig angenommenen Schallquelle ausgehen, treffen am linken und am rechten Ohr mit unterschiedlicher Lautstärke („interaurale Pegeldifferenz“) und – wegen der unterschiedlichen Entfernung zur Schallquelle – mit einer gewissen Laufzeitverzögerung zu verschiedenen Zeiten ein („interaurale Zeitdifferenz“). Verbunden damit sind auch Phasendifferenzen und Unterschiede in der Klangfarbe, wobei alle Faktoren bei der Ortung von Schallquellen eine Rolle spielen.

Der Laufzeitunterschied ist von der Richtung des Schallerzeugers (Einfallswinkel) und dem Abstand der beiden Ohren ( $d$ ) abhängig. Er wurde in Versuchen ermittelt und beträgt (wobei  $c$  die Schallgeschwindigkeit mit ca. 344 m/s ist):

Der geringste Laufzeitunterschied, den der menschliche Hörsinn noch wahrnehmen kann, beträgt etwa 0,03 ms bei einem Einfallswinkel des Schalls von ungefähr  $3^\circ$ . Bei genau seitlichem Schalleinfall beträgt der damit verbundene maximale Laufzeitunterschied etwa 0,6 ms. Für impulshaltige Schallereignisse, die Sprache und Musik im allgemeinen darstellen, funktioniert die Ortung mit den Laufzeitunterschieden – besonders in Blickrichtung – sehr genau. Bei Dauerschall und Frequenzen über 800 Hz ist die Lokalisierung allerdings in der Regel weniger präzise oder sogar mehrdeutig.

Intensitätsunterschiede entstehen durch die abschattende Wirkung des Kopfes und sind erst ab Frequenzen über 500 Hz wahrnehmbar. Sie sind ebenfalls von der Einfallrichtung des Schalles, aber auch von der Frequenz abhängig. Die



**Bild 3.10:** Laufzeitunterschied beim räumlichen Hören

Zusammenhänge bei breitbandigen Schallereignissen sind sehr komplex und teilweise mehrdeutig. Eine Lokalisierung ausschließlich durch Intensitätsunterschiede ist daher nur eingeschränkt möglich.

Wichtig: Bei tiefen Frequenzen ( $f < 80$  Hz) versagt die Lokalisation; wir können die Richtung, aus der Basstöne kommen, nicht mehr bestimmen. Dies macht man sich oft zunutze, indem man für die Wiedergabe tiefer Frequenzen nur einen einzigen Basslautsprecher („Subwoofer“) einsetzt. Für die Richtungslokalisierung genügen dann zwei kleinere Lautsprecherboxen für mittlere und hohe Frequenzen. Kritiker bemängeln jedoch ein fehlendes „Räumlichkeitsgefühl“ beim Verzicht auf die stereofone (zweikanalige) Wiedergabe im Bassbereich.

Das Ohr lässt sich eher durch Laufzeitunterschiede als durch unterschiedliche Intensitäten von der Richtung der Schallquelle überzeugen. Dabei gilt das Gesetz der **ersten Wellenfront**: der Schall, der zuerst eintrifft, bestimmt das Richtungsempfinden, auch wenn auf den ersten Schall Reflexionen aus anderen Richtungen kommen. Bei Laufzeitdifferenzen bis zu 30 ms darf der später eintreffende Schall sogar bis zu 10 dB lauter sein als der zuerst eintreffende, ohne dass dadurch der Richtungseindruck beeinträchtigt wird.

Helmut Haas fand bereits 1951 heraus, dass für Verzögerungszeiten zwischen 10 ms und 30 ms gilt, dass eindeutig der zuerst einfallende Schall für die Lokalisation der Schallquelle maßgeblich ist, und zwar völlig unabhängig davon, aus welcher Richtung der verzögerte Schall eintrifft ( "Über den Einfluss eines Einfach-Echos auf die Hörsamkeit von Sprache", Acustica 1, 1951, S. 49). Man hört trotzdem nur **eine** Schallquelle. Bei Laufzeitdifferenzen  $\Delta t$  größer 40 ms wird langsam das Vorhandensein von getrennten Schallreflexionen bemerkt, doch lokalisiert man die Schallquelle nach wie vor aus der Richtung des zuerst einfallenden Schalls, wenn der Pegel der Reflexion = Pegel des Direktsignals ist.

Überschreitet die Laufzeitdifferenz eine kritische Zeitgrenze von  $\Delta t$ , die größer als 50 ms ist, so werden das Direktsignal und die Reflexion bzw. das verzögert abgestrahlte Signal bei gleichem Pegel als zeitlich und unter Umständen auch räumlich getrennte Signale empfunden. Man spricht dann von einem Echo.

Diesen Haas-Effekt kann man sich bei der Beschallung zunutze machen, indem man den Zuhörern ein über Lautsprecher verstärktes Signal anbietet, das gegenüber der ersten Wellenfront um ca. 20 ms (der „Haas-Effekt-Zeit“) verzögert und um bis zu 10 dB lauter ist. Im Ohr des Zuhörers verschmelzen erste Wellenfront und verzögertes Signal zu einem einzigen Höreindruck, als sei nur eine Originalschallquelle vorhanden, die jedoch zur besseren Verständlichkeit nun merkbar lauter ist:

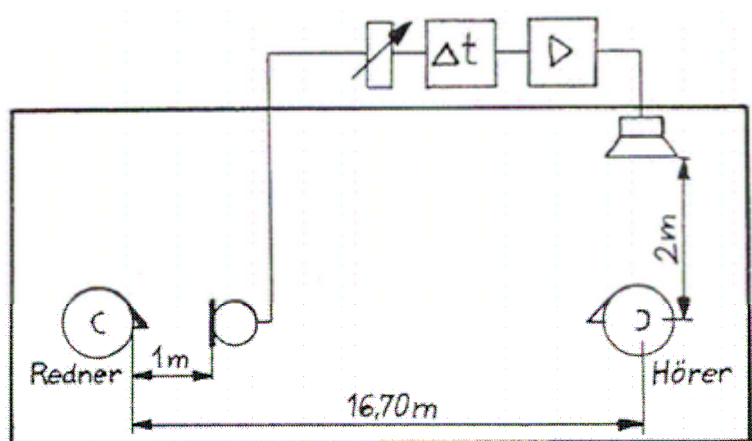


Bild y: Beschallung mit Haas-Effekt

## Raumwahrnehmung

Neben der Lokalisierung von Schallquellen ist die Orientierung („Wo befinde ich mich eigentlich? Wie ist die Umgebung beschaffen, in der ich mich gerade aufhalte?“) eine wichtige evolutionäre Errungenschaft des menschlichen Hörvermögens. Ausschlaggebend hierfür sind die frequenzabhängige Dämpfung des Schalls in unterschiedlichen Umgebungen und die Art der Reflexionen... (genauer ...)

Bild Y: Schallausbreitung in geschlossenen Räumen

Nach dem Eintreffen von direktem Schall und den ersten Reflexionen von Decke, Boden und Wänden beim Hörer, vermehren sich die Reflexionen und verdichten sich rasch zu dem, was Nachhall genannt wird.

Die Nachhallzeit ist als die Zeit definiert, in der der Schall nach dem Verstummen der Quelle um 60 dB abgenommen hat. In der Praxis bedeutet eine Abnahme um 60 dB, dass der Schall unhörbar geworden ist. Beachte: Je größer die Nachhallzeit, desto schlechter ist die Sprachverständlichkeit (Klatschtest).

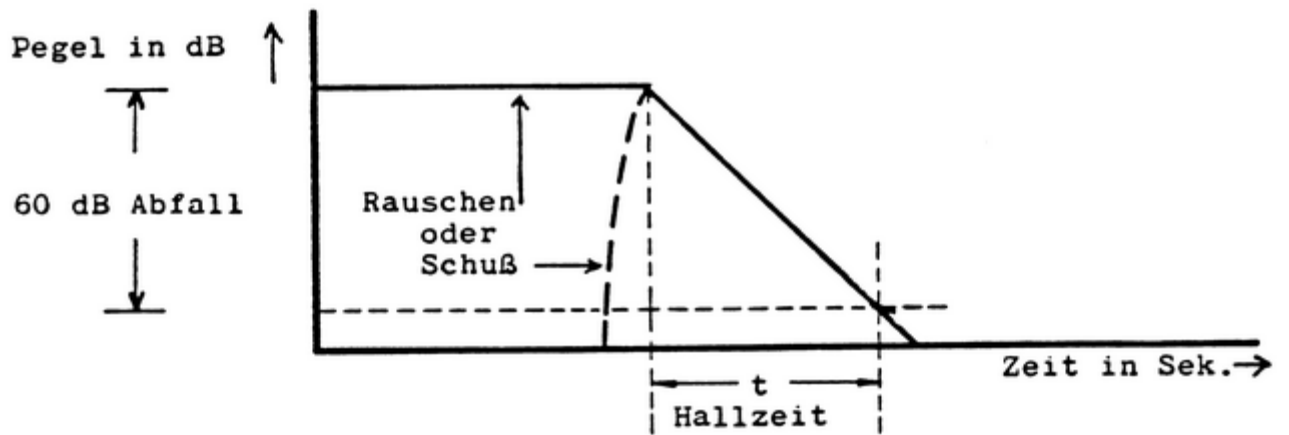


Bild xyz: Nachhallzeit



Bild Y: Schalltoter Raum an der TU Dresden

## Auswirkung von Amplitude und Frequenz auf die Hörwahrnehmung

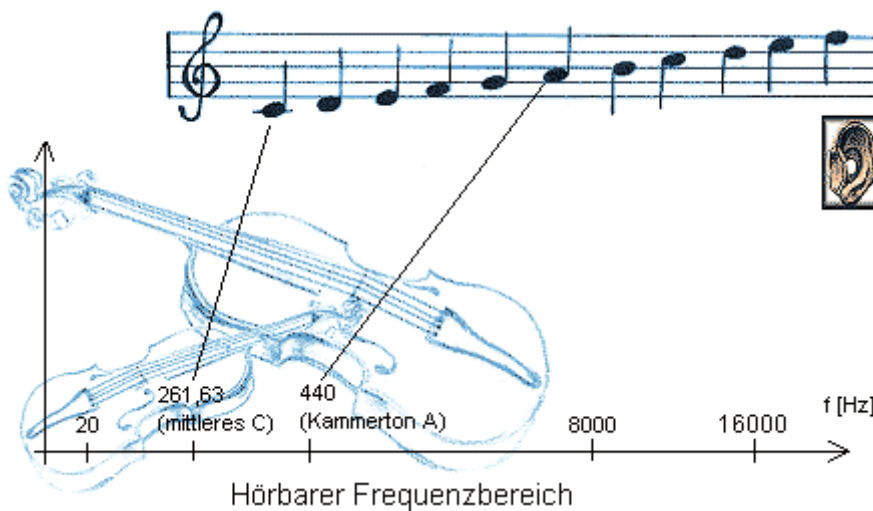
Werden sinusförmige Schwingungen mit einer harmonisch angeregten Lautsprechermembran hörbar gemacht, so haben die Größen Amplitude und Frequenz bestimmte Auswirkungen auf die Hörempfindung. Diese Parameter sind in erster Näherung für die Lautstärke und Tonhöhe eines Tones verantwortlich, was an folgenden Hörbeispielen verdeutlicht werden soll.



Kontinuierliche Erhöhung der Amplitude von 0% auf 100%, Frequenz konstant



Erhöhung der Frequenz von 440 Hz ( $T_1=2,2$  ms) auf 2000 Hz ( $T_n=0,5$  ms), Amplitude konstant



In der nebenstehenden Grafik ist der Zusammenhang zwischen dem Zahlenwert der Frequenzangabe und der empfundenen Tonhöhe grafisch und durch ein Hörbeispiel dargestellt. Die hier angedeutete Relation zwischen steigender Tonhöhe und steigender Frequenz findet ihr Äquivalent in der Notenschreibweise. Auch hier sind die Noten von Tönen steigender Tonhöhe aufsteigend im Notensystem angeordnet.

### Hörbeispiele

Das menschliche Gehör kann einen Frequenzbereich zwischen 20 Hz und 16000 Hz verarbeiten. Im ersten Beispiel wird der gesamte hörbare Frequenzbereich vorgestellt. Im zweiten Beispiel sollen einige Töne bestimmter Frequenz dargeboten werden. dies soll helfen, die Frequenzangabe eines Tones mit der Tonhöhe, die diese hervorruft, zu verbinden

1. Hörbeispiel:

Es wird der gesamte hörbare Frequenzbereich zwischen 20 Hz bis 16000 Hz durchfahren.



2. Hörbeispiel:

Nacheinander werden die Tonhöhen der Frequenzen 261,63 Hz (mittleres C), 440 Hz (Kammerton A), 1000 Hz und 5000 Hz vorgestellt.



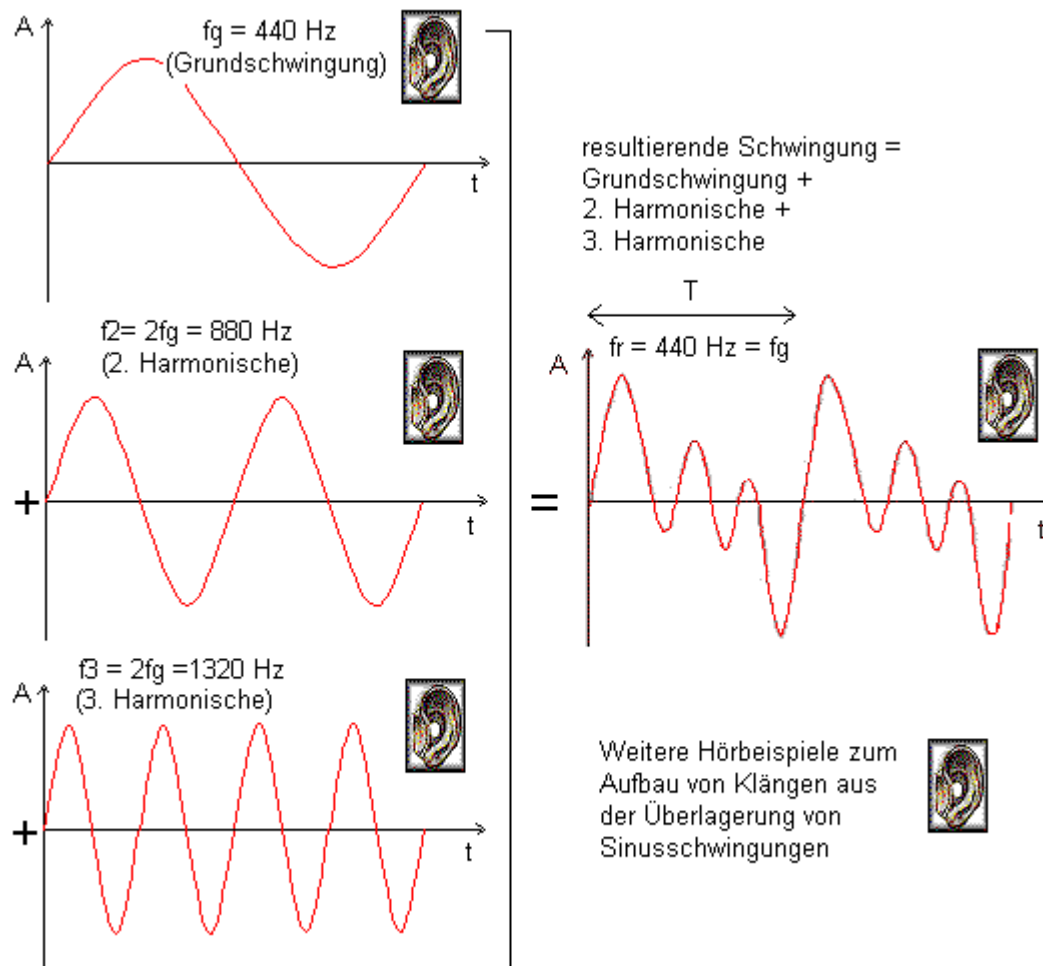


## Ton & Klang

Neben Tönen, die durch Sinusschwingungen erzeugt werden, existiert noch eine unendliche Anzahl an Schallereignissen, die einen von der Sinusform abweichenden Schwingungsverlauf aufweisen. Ein Schallereignis, das zwar einen periodischen Schwingungsverlauf aufweist, dessen Amplitudenverlauf aber von der Sinusform abweicht, entsteht durch die Überlagerung mehrerer Sinusschwingungen verschiedener Amplituden und Frequenzen. Die Frequenzen der Einzelschwingungen müssen hierbei in einem ganzzahligen Verhältnis zueinander stehen. Schallereignisse die diese Bedingung erfüllen, werden als **Klang** bezeichnet. Die Komponenten eines Klanges haben spezielle Namen, die in der Akustik zur physikalischen Beschreibung von Schallen häufig benutzt werden.






Bei Klängen bezeichnet man die tiefste vorhandene Frequenz als Grundschwingung. Die darauf aufbauenden Teilschwingungen mit ganzzahligen Vielfachen der Grundfrequenz werden als Oberschwingungen oder Harmonische bezeichnet. Hieraus geht hervor, dass die von Musikinstrumenten erzeugten Töne, physikalisch gesehen, Klänge darstellen.

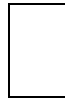
Dieser Sachverhalt soll durch die Grafik demonstriert werden. Der Aufbau von Klängen durch die Überlagerung von Sinusschwingungen soll durch ein Hörbeispiel verdeutlicht werden. Achten Sie darauf, dass die resultierende Schwingung, die aus der Überlagerung hervorgeht, den gleichen Tonhöhereindruck hervorruft, wie die Grundschwingung.



### Aufbau eines Klanges

Im folgenden Hörbeispiel wird der Aufbau eines Klanges durch die Überlagerung einer Grundschwingung der Frequenz  $f = 200$  Hz (1. Harmonische) mit der 2. bis 5. Harmonischen dargestellt.

zugefügte Harmonische	resultierender Klang
1. Harmonische (Grundschwingung) $f = 200$ Hz	
2. Harmonische $f = 400$ Hz	
3. Harmonische $f = 600$ Hz	
4. Harmonische $f = 800$ Hz	
5. Harmonische $f = 1000$ Hz	



Hörbeispiel:  
Aufbau des Klanges aus den ersten 5 Harmonischen

Beginnend mit der Grundschwingung wird ein Klang aufgebaut, indem jede Sekunde eine neue Harmonische überlagert wird.

Nach fünf Sekunden entsteht somit ein Klang, in dem die ersten fünf Harmonischen mit gleicher Ausprägtheit vorhanden sind.

Beachten Sie, dass die resultierende Schwingung die gleiche Tonhöhe aufweist, wie die Grundschwingung, obwohl die überlagerten Obertöne mit steigender Ordnung eine ansteigende Tonhöhe aufweisen. Um dies zu demonstrieren, wird erneut der Aufbau des Klanges durch die Überlagerung der ersten fünf Harmonischen vorgestellt. Nun wird nach jedem neu hinzugefügtem Oberton ein Sinuston der Frequenz  $f = 200$  Hz abgespielt. Sie werden feststellen, dass sich durch das Hinzufügen der Obertöne nichts an der wahrgenommenen Tonhöhe ändert.

Die Frequenz der Grundschwingung ist somit bestimmend für die empfundene Tonhöhe.



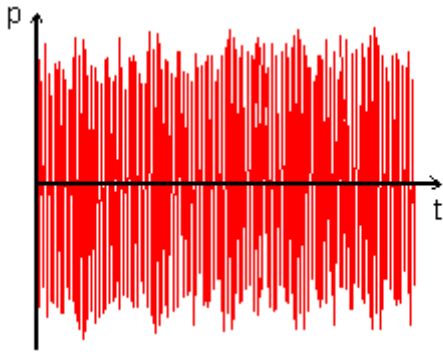
Hörbeispiel zur konstanten Tonhöhe bei der Überlagerung der Harmonischen eines Klanges.

### Bezeichnung der Klangkomponenten

Die einzelnen Klangkomponenten werden als Harmonische, Obertöne oder Partialtöne bezeichnet. Den Zusammenhang zwischen den einzelnen Bezeichnungen gibt die folgende Tabelle wieder.

Vielfache der Grundfrequenz $f_0$	Harmonische	Obertöne	Partial- oder Teiltöne
1	1. Harmonische oder Grundwelle	Grundton	1. Partialton
2	2. Harmonische	1. Oberton	2. Partialton
3	3. Harmonische	2. Oberton	3. Partialton
4	usw.	usw.	usw

# Geräusche



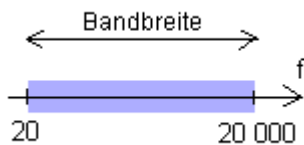
Schallereignisse, die aus theoretisch unendlich vielen Einzelschwingungen bestehen, deren Frequenzabweichungen (Frequenzabstände) beliebig klein sind, werden als Geräusch bezeichnet.

Geräusche weisen i.a. einen nichtperiodischen Schwingungsverlauf auf. Das bedeutet, dass ein zu einem beliebigen Zeitpunkt betrachteter Amplitudenverlauf nur mit Hilfe statistischer Aussagen angegeben werden kann.

Ein Ausschnitt eines charakteristischen Schwingungsverlaufes eines Geräusches ist in der nebenstehenden Grafik dargestellt. Es gibt unendlich viele Geräusche, die sich durch ihre Intensität, ihre Klangfarbe und ihren Rhythmus unterscheiden. Zwei für die Akustik wichtige Geräusche sind

das sogenannte weiße Rauschen und das Bandpassrauschen (Schmalbandrauschen). Durch ein Hörbeispiel soll weißes Rauschen und Bandpassrauschen vorgestellt sowie die Entstehung von Bandpassrauschen aus weißem Rauschen demonstriert werden.

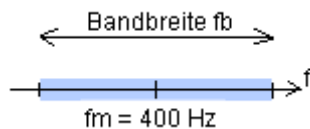
## Schwingungsverlauf von Geräuschen



Weißes Rauschen besitzt keine Tonhöhe, keinen Rhythmus und weist zu jedem Zeitpunkt die gleichen Eigenschaften im Schwingungsverlauf auf. Es enthält alle Frequenzen zwischen 20 Hz und 20 kHz (s. Grafik).



Die Zeitfunktion kann nur mit statistischen Methoden beschrieben werden. Im nebenstehenden Hörbeispiel wird 3 Sekunden lang weißes Rauschen dargeboten.



Filtert man aus weißem Rauschen alle Frequenzanteile bis auf einen bestimmten Frequenzbereich heraus, so erhält man das Bandpassrauschen. Werden aus weißem Rauschen alle Frequenzanteile außerhalb des Bereichs von  $400 \text{ Hz} \pm 50 \text{ Hz}$  entfernt, so spricht man von einem Bandpassrauschen der Mittenfrequenz 400 Hz und der Bandbreite 100 Hz. Da Bandpassrauschen aus weißem Rauschen erzeugt wird, gilt auch hier, dass die Zeitfunktion nur mit Hilfe statistischer Mittel beschrieben werden kann. Verkleinert man den Frequenzbereich der nicht

herausgefilterten Frequenzanteile, also die Bandbreite, so geht das Bandpassrauschen in das sogenannte Schmalbandrauschen über, welches oft eine eindeutige Tonhöhe aufweist.

Im nachfolgenden Hörbeispiel wird die Bandbreite eines Bandpassrauschens mit der Mittenfrequenz 400 Hz und der Bandbreite 100 Hz schrittweise um jeweils 30 Hz verkleinert. Aus dem Bandpassrauschen entsteht somit ein Schmalbandrauschen der Mittenfrequenz 400 Hz und der Bandbreite  $f_b = 10 \text{ Hz}$ , welches eine eindeutige Tonhöhe aufweist. Je schmaler die Bandbreite wird, desto ausgeprägter wird sich eine Tonhöhenwahrnehmung einstellen, die der Mittenfrequenz  $f = 400 \text{ Hz}$  entspricht.

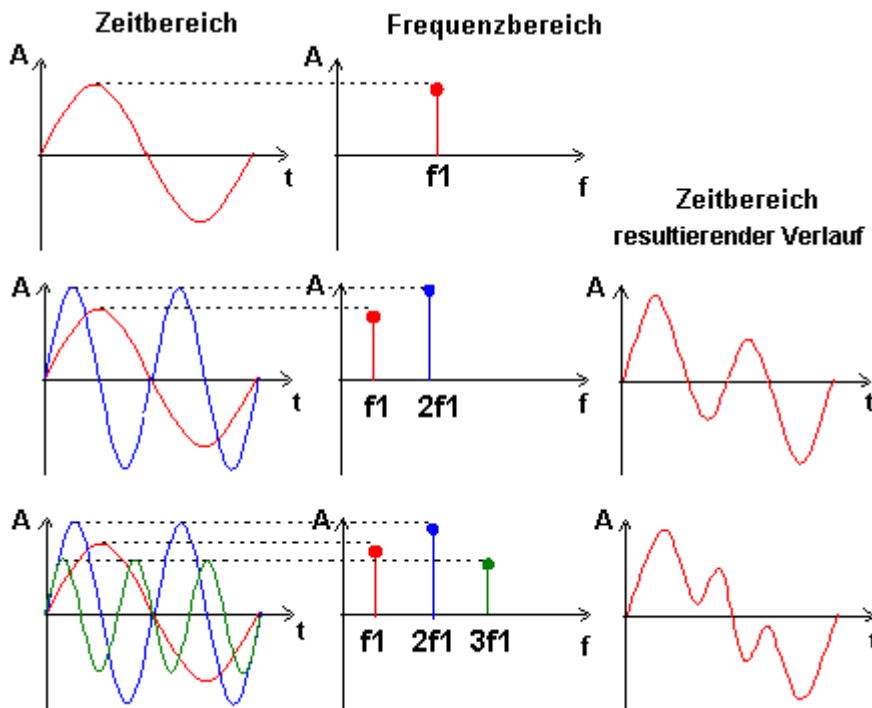
**Mittenfrequenz  $f_m = 400 \text{ Hz} = \text{const.}$**

Bandbreite $f_b = 100 \text{ Hz}$	Bandbreite $f_b = 70 \text{ Hz}$	Bandbreite $f_b = 40 \text{ Hz}$	Bandbreite $f_b = 10 \text{ Hz}$

## Darstellung durch das Frequenzspektrum

Jede periodische Schwingung kann als eine Überlagerung von Sinusschwingungen unterschiedlicher Amplituden und Frequenzen dargestellt werden.

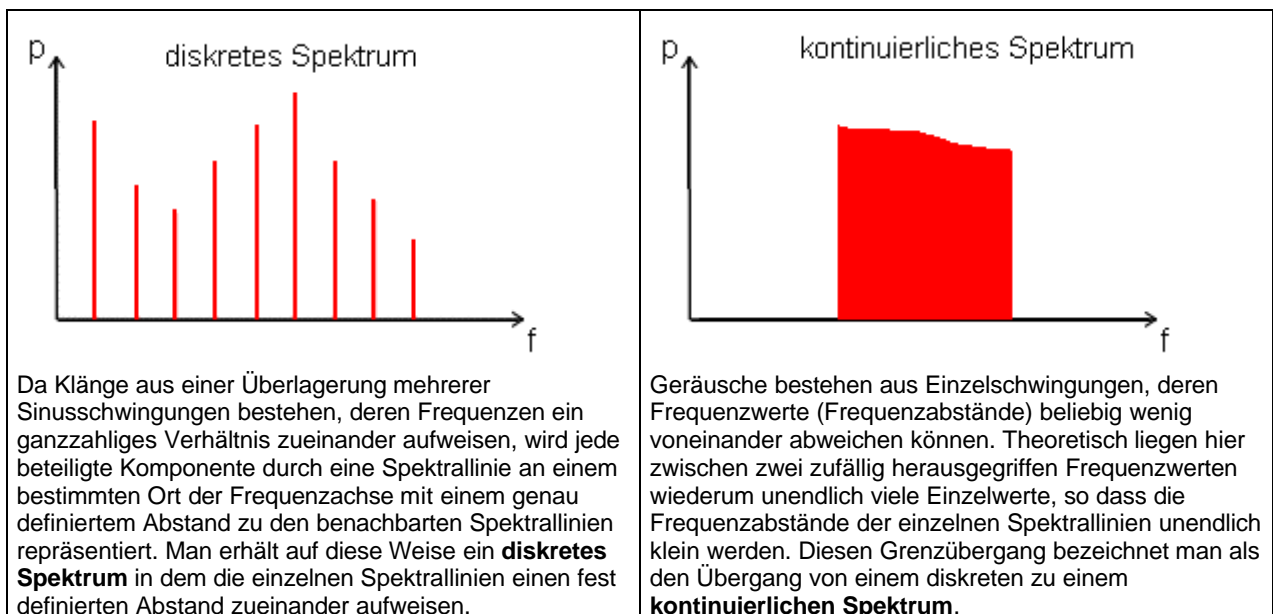
Trägt man die Amplituden der beteiligten Schwingungen in Abhängigkeit der Frequenz auf, so erhält man eine Darstellung, die als Frequenzspektrum (kurz: Spektrum) bezeichnet wird. Im Spektrum wird die Amplitude jeder Teilschwingung durch eine Linie entsprechender Höhe in Abhängigkeit ihrer Frequenz repräsentiert. Der Zusammenhang zwischen dem zeitlichen Schwingungsverlauf und dem Spektrum soll durch die nebenstehende Zeichnung verdeutlicht werden.



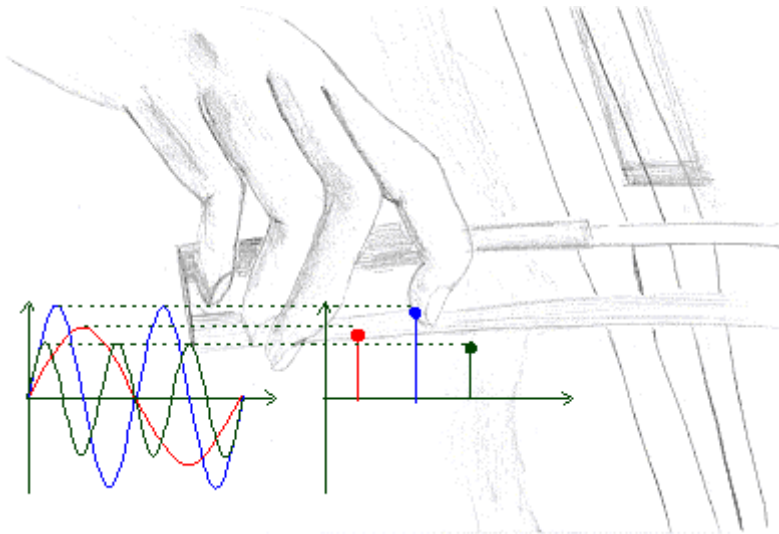
Die Darstellung von Schwingungen durch das Spektrum nennt man **Darstellung im Frequenzbereich**.

Analog bezeichnet man die Darstellung als zeitlichen Verlauf als **Darstellung im Zeitbereich**.

Reine Töne und Klänge können als diskretes Spektrum, d.h. als Spektrum mit einzelnen Linien dargestellt werden. Geräusche hingegen besitzen ein kontinuierliches Spektrum. Die Darstellung durch das Spektrum ist ein wichtiges Hilfsmittel zur Beschreibung von Schallen, da jedes Schallereignis ein charakteristisches Spektrum besitzt:



# Die Klangfarbe



Hört man einen Klang oder ein Geräusch, so empfindet man zusätzlich zur Tonhöhe und Lautstärke noch etwas anderes: Die **Klangfarbe**.

Die Klangfarbe wird dadurch definiert, dass die Amplituden der verschiedenen Oberschwingungen mit unterschiedlicher Ausprägtheit vorhanden sind. Klänge weisen eine zunehmend "härtere, brillantere" Färbung auf, wenn die Zahl der Oberschwingungen und deren Ausprägtheit zunimmt. Je stärker die Ausprägtheit der beteiligten Obertöne in einem Klang ist, um so "weicher" klingt er.

## Hörbeispiele zur Klangfarbe

Spektren der vorgestellten Klänge	Beispiele
	<input type="text"/>
	<input type="text"/>
	<input type="text"/>
	<input type="text"/>
	<input type="text"/>

Nachfolgend werden einige Klänge, bestehend aus den ersten 5 Harmonischen vorgestellt. Die Frequenz der Grundschwingung (1. Harmonische) beträgt bei allen Beispielen  $f = 400$  Hz. Nacheinander wird die Ausprägung der einzelnen Oberschwingungen entsprechend den dargestellten Spektren verändert.


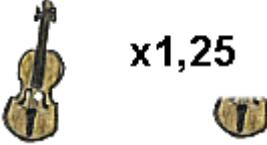
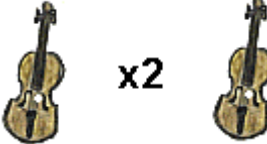


Beachten Sie, dass die Tonhöhe bei allen Beispielen konstant bleibt und der Tonhöhe der 1. Harmonischen (Grundschwingung) entspricht. Die Töne werden mit zunehmender Ordnung und Ausprägtheit der Obertöne härter klingen.

## Schalldruckpegel und Schallpegeladdition

Im folgenden Hörbeispiel soll der Zusammenhang zwischen der Erhöhung des Schalldruckpegels und der damit verbundenen Lautstärkeänderung demonstriert werden.

Durch die Definition des Schalldruckpegels als logarithmisches Maß verdoppelt sich die empfundene Lautstärke bei einer Erhöhung des Schalldruckpegels um 10 dB.

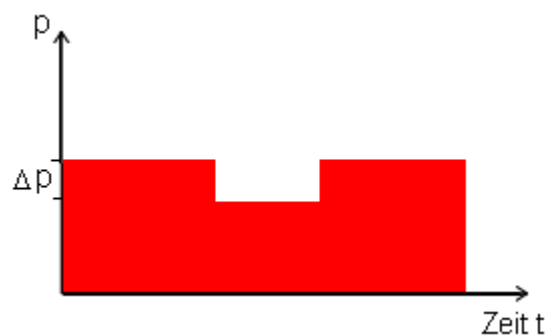
Die Grafik verdeutlicht den Zusammenhang zwischen der Pegeldifferenz in dB, der Schalldruckänderung und der dazugehörigen Lautstärkeänderung.

Lautstärkeänderung	Schallpegel-unterschied	Anzahl gleicher Schallquellen	Schalldruck-änderung	Hörbeispiele
Grundlautstärke	-	 x1	x 1	
gerade noch hörbar lauter	+ 1 dB	 x1,25	x 1,1	<input type="checkbox"/>
deutlich hörbar lauter	+ 3 dB	 x2	x 1,4	<input type="checkbox"/>
doppelt so laut wie die Grundlautstärke	+ 10 dB	 x10	x 3	<input type="checkbox"/>
viermal so laut wie die Grundlautstärke	+ 20 dB	 x100	x 10	<input type="checkbox"/>

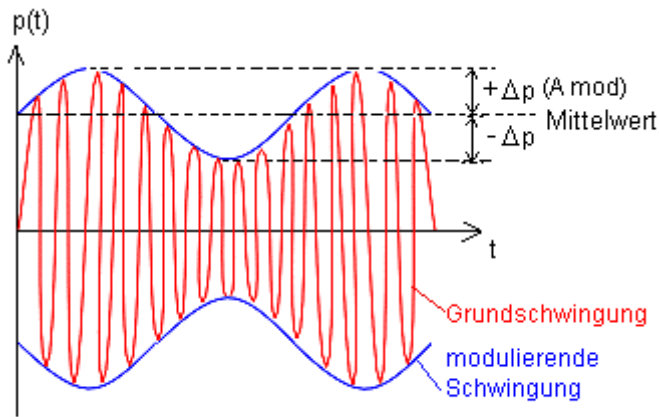
### Hörbeispiele zu Pegeldifferenzen

Vorge stellt werden soll die Auswirkung einiger Pegeldifferenzen auf die Hörwahrnehmung.

Hierzu werden nacheinander drei Sinustöne einer Frequenz von 440 Hz präsentiert. Der zweite Ton wird um einen bestimmten Pegel abgeschwächt, anschließend wieder in der ursprünglichen Lautstärke vorgespielt. Auf diese Weise ist ein direkter Vergleich zwischen dem vollen und dem abgeschwächten Signal möglich. Der zeitliche Verlauf der Hörbeispiele ist in der nebenstehenden Grafik verdeutlicht.



# Amplitudenmodulation



Als Amplitudenmodulation bezeichnet man die periodische Änderung der Amplitude eines Schallereignisses in Abhängigkeit der Zeit. Ändert man die Amplitude eines reinen Tones periodisch über der Zeit, so weist das Resultat zwar eine konstante Tonhöhe aber eine sich periodisch verändernde um einen Mittelwert schwankende Lautstärke auf (s. Grafik).

Amplitudenmodulierte Töne stellen eine wichtige Klasse in der technischen Akustik dar.

Die zu modulierende Schwingung wird Grundschiwingung (Trägerschwingung, rot) genannt. Die Grundschiwingung bestimmt die Tonhöhe und die Lautstärke des Tones.

Die Schwiwingung, welche die Schwankung der Lautstärke um den Mittelwert verursacht, wird modulierende Schwiwingung (blau), ihre Frequenz Modulationsfrequenz genannt. Die Modulationsfrequenz bestimmt die Schnelligkeit, die Amplitude der modulierenden Schwiwingung die Größe der Lautstärkeänderung.

Die Auswirkung einer Änderung der Parameter Modulationsfrequenz und Modulationsamplitude auf das Hörereignis werden durch ein Hörbeispiel demonstriert.

## Hörbeispiele zur Amplitudenmodulation

Amplitudenänderung der resultierenden Schwiwingung in dB	Entsprechende Änderung der Amplitude $A_{mod}$ um den Faktor	Beispiele
6	2	<input type="text"/>
15	5,6	<input type="text"/>
20	10	<input type="text"/>

Im ersten Beispiel soll die Modulationsfrequenz konstant gehalten und die Amplitude der modulierenden Schwiwingung verändert werden.

Hierzu wird ein Sinuston der Frequenz  $f = 500$  Hz mit einer konstanten Modulationsfrequenz von  $f_m = 1$  Hz moduliert. Die Amplitude  $A_{mod}$  der modulierenden Schwiwingung wird schrittweise geändert, so dass sich eine periodische Änderung der Amplitude der Grundschiwingung ergibt. In den nebenstehenden drei Beispielen wird eine Schwankung der Lautstärke um 6 dB, 15 dB und 20 dB um den Mittelwert eingestellt.

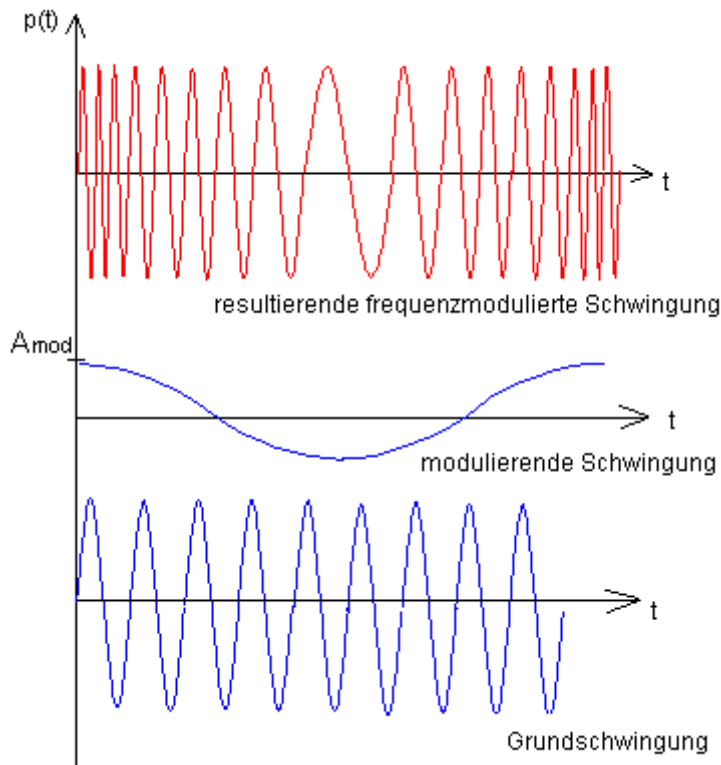
Die Lautstärke der vorgestellten Beispiele ändert sich mit konstanter Geschwindigkeit. Die Stärke der Lautstärke-schwankung ist von der Amplitude  $A_{mod}$  der modulierenden Schwiwingung abhängig.

Modulationsfrequenz $f_{mod}$ in [Hz]	Beispiele
6	<input type="text"/>
3	<input type="text"/>
10	<input type="text"/>

Im zweiten Beispiel wird erneut ein Sinuston der Frequenz  $f = 500$  Hz moduliert. Die Amplitude der Modulierenden Schwiwingung wird so eingestellt, dass sich eine konstante Lautstärkeänderung von 25 dB um den Mittelwert der Grundschiwingung ergibt.

Die Modulationsfrequenz  $f_{mod}$  wird nacheinander von 1 Hz auf 3 Hz und 10 Hz erhöht.

Die Lautstärkeänderung bleibt in allen drei Fällen konstant, da an der Amplitude der modulierenden Frequenz nichts verändert wird. Durch die Erhöhung der Modulationsfrequenz  $f_{mod}$  stellt sich aber eine höher werdende Schwankungsgeschwindigkeit ein.



## Frequenzmodulation

Als Frequenzmodulation bezeichnet man die periodische Änderung der Frequenz eines Schallereignisses. Hierdurch weisen frequenzmodulierte Schallereignisse eine sich zeitlich ändernde Tonhöhe auf.

Ändert man die Frequenz eines reinen Tones periodisch über der Zeit, so erhält man einen Ton konstanter Lautstärke, aber mit sich periodisch ändernder Tonhöhe.

Eine frequenzmodulierte Schwingung weist eine konstante Amplitude bei sich periodisch ändernder Frequenz auf. Dieser Zusammenhang ist in der nebenstehenden Abbildung dargestellt. Die zu modulierende Schwingung wird Grundschiwingung genannt, diese bestimmt die Grundtonhöhe und die Lautstärke des Tones. Die Frequenz der modulierenden Schwingung, welche die Tonhöhenchwankung verursacht, wird Modulationsfrequenz genannt. Die Modulationsfrequenz bestimmt die Schnelligkeit, die Amplitude der modulierenden Schwingung die Größe der Frequenzänderung. Hörbeispiele: Die Auswirkungen dieser Parameter auf das

Hörereignis sowie einige typische Vertreter frequenzmodulierter Schallereignisse.

### Hörbeispiele zur Frequenzmodulation

In den folgenden beiden Hörbeispiele soll der Zusammenhang zwischen Modulationsfrequenz und Amplitude der modulierenden Schwingung und der frequenzmodulierten Schwingung verdeutlicht werden.

Beispiel 1	Amplitudenwert $A_{mod}$ der modulierenden Schwingung	Beispiel 2	Frequenz $f_{mod}$ in [Hz] der modulierenden Schwingung (entspricht der Anzahl der Tonhöhenchwankungen pro Sekunde)
<input type="checkbox"/>	0 (keine Modulation)	<input type="checkbox"/>	0 (keine Schwankung)
<input type="checkbox"/>	0,1	<input type="checkbox"/>	1
<input type="checkbox"/>	0,3	<input type="checkbox"/>	5
<input type="checkbox"/>	0,5	<input type="checkbox"/>	10
<p>Die Modulationsfrequenz (Frequenz der modulierenden Schwingung) wird konstant auf <math>f_{mod} = 2</math> Hz gehalten. Nun wird schrittweise die Amplitude <math>A_{mod}</math> der modulierenden Schwingung erhöht. Durch die Amplitudenerhöhung wird sich eine stärker werdende Schwankung der Tonhöhe einstellen.</p> <p>Da die Modulationsfrequenz in allen Beispielen <math>f_m = 2</math> Hz beträgt, schwankt die Tonhöhe 2 mal periodisch pro Sekunde. Die Grundtonhöhe entspricht der Frequenz <math>f = 500</math> Hz.</p>		<p>In diesem Beispiel wird die Amplitude <math>A_{mod}</math> der modulierenden Schwingung konstant gehalten und die Modulationsfrequenz <math>f_{mod}</math> geändert. Die Schwankung der Tonhöhe der frequenzmodulierten Schwingung bleibt durch die konstant gehaltene Amplitude der modulierenden Schwingung während der einzelnen Beispiele gleich.</p> <p>Durch die Änderung der Modulationsfrequenz <math>f_{mod}</math> tritt aber ein Anstieg der Tonhöhenänderungen pro Sekunde auf. Auch hier beträgt die Frequenz der Grundschiwingung 500 Hz.</p>	



# Audiotechnik

Die technische Aufnahme, Speicherung, Bearbeitung und Wiedergabe von Schall bezeichnen wir allgemein als Audiotechnik.

Am Anfang der Kette von Aufnahme, Speicherung, Bearbeitung und Wiedergabe von Ton steht das Mikrofon. Für die Speicherung stehen uns die analoge oder digitale Bandaufzeichnung zur Verfügung. Geschickt ist jedoch die Verwendung eines digitalen Harddisk-Recording-Systems auf einem „normalen“ PC oder einer eigens dafür optimierten „Workstation“, auf dem wir die Audioaufnahmen auch gleich bearbeiten können. Zur Langzeit-Speicherung (Archivierung, Sammlung) benutzen wir – nach der Ära von Schallplatten, Tonbändern und Musikkassetten – CDs, DVDs und zunehmend elektronische Medien wie Festplatten, Speicherkarten oder „Webpace“. Schließlich folgen am Schluss der Kette im Studio die Monitoranlage, die HiFi-Anlage zu Hause, die kleinen PC-Lautsprecher neben dem Computer des Anwenders oder Kopfhörer verschiedenster Bauart.

## Qualitätskriterien

Abgesehen von gewollten, weil künstlerisch gestaltenden Veränderungen einer Audioaufnahme sollten alle Glieder der Audio-Übertragungskette, also vom aufnehmenden Schallwandler über die passiven und aktiven elektronischen Komponenten bis zu den Lautsprechern das Audiomaterial gar nicht oder zumindest nur in „erträglichen Grenzen“ verfälschen.

Stichworte:

## Frequenzgang, Bandbreite

Der Frequenzgang beschreibt das Verhalten eines „linearen zeitinvarianten Systems“ (etwa eines Mikrofons, einer Lautsprecherbox oder eines Klangfilters) in Abhängigkeit der Frequenz. Zur Ermittlung des Frequenzgangs wird das System mit einer sinusförmigen Eingangsgröße und variabler Frequenz angeregt. Die Ausgangsgröße des Systems ist ebenfalls wieder sinusförmig mit gleicher Kreisfrequenz, jedoch mit einer anderen Amplitude und einer Phasenverschiebung gegenüber dem Eingangssignal. Das Ergebnis dieses Versuchs wird über der Frequenz für das jeweilige Amplitudenverhältnis und die jeweilige Phasenlage aufgetragen. Sie werden als **Amplitudengang** (Betragsfrequenzgang) bzw. als **Phasengang** (Phasenfrequenzgang) bezeichnet. Die kombinierte Darstellung ist das **Bode-Diagramm**, hier am Beispiel eines Tiefpass-Filters, das Signale mit niedrigen Frequenzen passieren lässt und Signale mit höheren Frequenzen zunehmend dämpft.

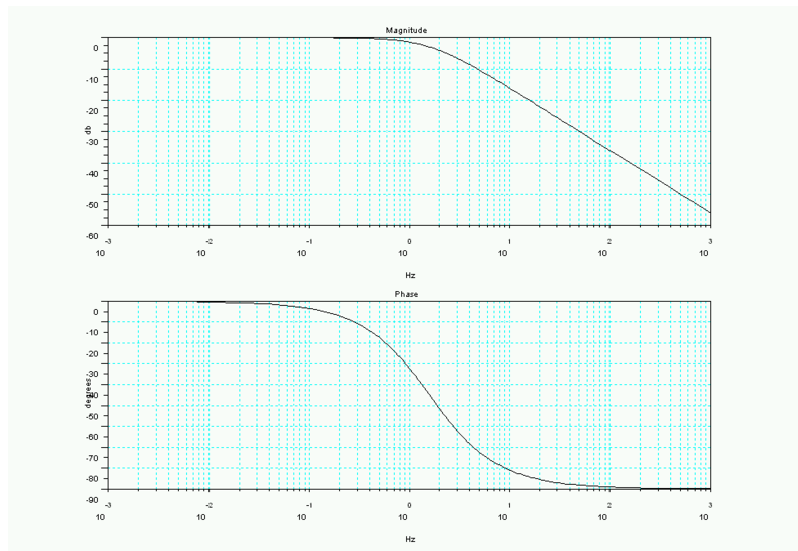


Bild xyz: Bode-Diagramm eines Tiefpasses

Alle Übertragungssysteme haben in der Regel eine obere und eine untere Grenzfrequenz, das heißt Signale mit Frequenzen darunter bzw. darüber werden nicht übertragen. Reale Übertragungssysteme weisen allerdings keine unendlich steilen Flanken des Frequenzganges auf, sondern eine zunehmende Dämpfung des Signals. Neben der Angabe der absoluten Grenzfrequenzen erfolgt in der Praxis deshalb häufig die Angabe der Frequenzen bei Erreichen von 3 dB Dämpfung, der dazwischen liegende Frequenzbereich wird als Bandbreite bezeichnet.

Für eine originalgetreue Audiowiedergabe – egal, ob bei der Beschallung von Konzerten oder der Wiedergabe von Audioaufnahmen – sollte ein linearer Frequenzgang innerhalb unseres Hörbereiches (Hörfeld, s.o.) vorhanden sein. Da unser eigener Ohr-Frequenzgang jedoch bereits abhängig von der dargebotenen Lautstärke (Schalldruck) sehr nichtlinear ist, leben wir hier ausschließlich von Kompromissen. Der „Loudness-Schalter“ zur Anpassung des Frequenzganges an unser Gehör ist da das beste Beispiel.

In der Elektroakustik sollten wir jedoch erreichen, dass von der Aufnahme (mit dem Mikrofon oder anderen Tonaufnehmern) bis zum Lautsprecher oder Kopfhörer ein linearer Frequenzgang anzutreffen ist, um eine größtmögliche Kontrolle über das Klanggeschehen zu erhalten. Hochwertige Audio-Komponenten weisen in ihren Technischen Datenblättern deshalb immer den Frequenzgang in Form eines Messprotokolls oder Bode-Diagramms aus.

In der Audiotechnik gilt ein Frequenzgang dann als „ausgewogen“, das heißt dass die Bandbreite für das menschliche Hörorgan als angenehm empfunden wird (wieder einmal das statistische Mittel von subjektiven Äußerungen der Probanden...), wenn das Produkt aus oberer und unterer Grenzfrequenz 400.000 beträgt, also zum Beispiel den Frequenzbereich von 20 – 20.000 Hz umfasst ( was der „HiFi“-Norm [HiFi = High Fidelity] DIN xyz entspricht) – oder auch einer Bandbreite von 40 – 10.000 Hz oder 80 – 5.000 Hz.

Im Umkehrschluss bedeutet das aber auch, dass bei einer Audioaufnahme mit fehlenden Bässen nicht der Bass-Klangregler auf maximalen Anschlag zu stellen ist (denn wo nichts ist, da kann man auch nichts verstärken!), sondern die hohen Frequenzen beschnitten oder wenigstens gedämpft werden können, um ein ausgewogenes Klangbild zu erreichen.

Andererseits müssen oder wollen wir den Frequenzgang der elektroakustischen Übertragungskette absichtlich verbiegen. Bei Beschallungen von Konzertsälen oder anderen Veranstaltungsräumen finden wir durch unterschiedliche Oberflächenbeschaffenheiten (z.B. auch die Bestuhlung) einen „verbogenen“ Frequenzgang des Raumes vor. Unebenheiten im Frequenzgang können wir mit einem sogenannten „Equalizer“ wieder geradebiegen. Außerdem kann zwischen Mikrofon und Lautsprechern ein Rückkopplungseffekt auftreten, der je nach Raumbeschaffenheit sich in einem „Pfeifen“ bei einer bestimmten Frequenz bemerkbar macht. Auch diese Rückkopplungsfrequenz können wir in einem „Equalizer“ dämpfen und somit die Gesamtverstärkung wieder erhöhen. Näheres zum Equalizer im Kapitel „Effekte / Audiobearbeitung“.

(evtl. dazu Bild Equalizer mit „Pfeif-Frequenz“)

### **Klirrfaktor (THD Total Harmonic Distortion):**

Idealerweise folgt das Ausgangssignal eines elektroakustischen Bauteils linear dem Eingangssignal, d.-h.: es wird völlig unverzerrt übertragen und dabei verstärkt oder abgeschwächt. Durch Nichtlinearitäten bei den verwendeten (aktiven) Komponenten (Röhren, Transistoren, PC) wird allerdings die Kurvenform des Ausgangssignals insbesondere bei hohen Signalamplituden verformt (gestaucht). Das ist vergleichbar mit einer Addition von Obertönen und macht sich als unangenehme Verzerrung und „schepperndes“ oder „klirrendes“ Zusatzgeräusch bemerkbar.

Das menschliche Gehör ist, in Abhängigkeit von der Frequenz, für Verzerrungen (Klirr) empfindlich. 5 % Klirrfaktor im Bassbereich sind meist nicht wahrnehmbar. Im Präsenz- bzw. Brillanzbereich (1 bis 4 kHz), wo das Gehör am empfindlichsten ist, können unter bestimmten Bedingungen Verzerrungen auch noch unter 0,5 % hörbar sein. Die Hörbarkeit von Klirr in der elektroakustischen Übertragung (HiFi) hängt jedoch auch stark von der Beschaffenheit des Nutzsignals (Musik, Sprache) und dessen Spektrum ab. Mehrere sinusähnliche Klänge gelten als am empfindlichsten gegen Klirr. Etwa beim Zusammenspiel mehrerer Flöten kann Klirr schon ab 0,5 % gehört werden, da hier sehr obertonarme Klänge vorliegen. Bei Sprache oder anderen spektral „dichten“ Klängen und Geräuschen, wie z. B. Schlagzeug, ist Klirr erst bei deutlich größeren Klirrfaktoren hörbar.

Elektroakustische Geräte erzeugen unterschiedlich starken Klirr. HiFi-Verstärker sind heute meistens so konstruiert, dass der von ihnen erzeugte Klirrfaktor in weiten Bereichen völlig unterhalb der Wahrnehmungsschwelle liegt (Klirrfaktoren unter 0,1 %), es sei denn, man betreibt den Verstärker nahe seiner Leistungsgrenze. Klirr entsteht zumeist bei der Schallwandlung im Lautsprecher. Diese erzeugen frequenzabhängig, gerade bei höheren Pegeln (>95 dB SPL) oft hörbaren Klirr. Auch Tonabnehmer für Langspielplatten sowie die Rille selbst klirren mit Werten oberhalb der Wahrnehmungsschwelle.

Der Klirrfaktor  $k$  gibt an, wie stark die Oberschwingungen, die bei der Verzerrung eines sinusförmigen Signals entstehen, im Vergleich zum Gesamtsignal sind. Es ist das Verhältnis des Oberschwingungs- zum Gesamteffektivwert einschließlich Grundschwingungsanteil. Für ein harmonisches Signal ergibt sich der Gesamtklirrfaktor (Klirrfaktor) zu:

$$k = \sqrt{\frac{\bar{U}_2^2 + \bar{U}_3^2 + \bar{U}_4^2 + \dots}{\bar{U}_1^2 + \bar{U}_2^2 + \bar{U}_3^2 + \bar{U}_4^2 + \dots}}$$

Der Klirrfaktor  $k$  ist immer kleiner oder gleich 1 und wird deshalb meistens in Prozent angegeben.

### Dynamik, Störabstand (Rauschabstand, „Signal/Noise-Ratio“ SNR, Fremdspannungsabstand)

Als Dynamik bezeichnet man das Verhältnis von größtem (noch unverzerrten) Signal zum kleinsten Signal, das eine Übertragungskette übertragen kann. Nach „oben“ ist das Signal durch die Aussteuerungsgrenzen z.B. der Verstärker begrenzt – siehe „Klirrfaktor“. Bei kleinen Signalen werden dagegen evtl. eingestreute Störsignale hörbar, die sich als Brummen oder Rauschen bemerkbar machen können.

Das Signal-Rausch-Verhältnis (auch Störabstand  $a$  bzw. (Signal-)Rauschabstand  $a_R$ , oft auch abgekürzt als SRV beziehungsweise SNR oder S/N vom Englischen signal-to-noise ratio) ist ein Maß für die Qualität eines aus einer Quelle stammenden Nutzsignals, das von einem Rauschsignal überlagert ist. Es ist definiert als das Verhältnis der mittleren Leistung des Nutzsignals der Signalquelle zur mittleren Rauschleistung des Störsignals der gleichen Signalquelle.

Das Signal-Rausch-Verhältnis ist definiert als das Verhältnis der vorhandenen mittleren Signalleistung zur vorhanden mittleren Rauschleistung, wobei der Ursprung der Rauschleistung nicht berücksichtigt wird.

Als Verhältnis von Größen gleicher Maßeinheit ist das Signal-Rausch-Verhältnis dimensionslos. Es ist also:

$$\text{SNR} = \frac{\text{Nutzsignalleistung}}{\text{Rauschleistung}} = \frac{P_{v,S}}{P_{v,N}}$$

Da aber die Signalleistung bei vielen technischen Anwendungen um mehrere Größenordnungen größer ist als die Rauschleistung, wird das Signal-Rausch-Verhältnis oft im logarithmischen Maßstab dargestellt. Man benutzt dazu die Pseudoeinheit Bel (B) beziehungsweise Dezibel (dB).

$$\text{SNR}|_{\text{dB}} = 10 \lg \left( \frac{\text{Nutzsignalleistung}}{\text{Rauschleistung}} \right) = 10 \lg \left( \frac{P_{v,S}}{P_{v,N}} \right)$$

### Latenzen

Als Latenzen bezeichnet man Laufzeitverzögerungen, also die gegenüber einer Schallaufnahme zeitlich verzögerte Wiedergabe dieses Schallereignisses. Dieses kann einerseits gestalterisch gewollt sein, zum Beispiel zum Erzeugen einer künstlichen Raumakustik durch das Hinzufügen von Echos und Nachhall zum Originalsignal.

Andererseits bringt eine Latenz bei der digitalen Audibearbeitung auch unangenehme Effekte und Störungen hervor. Die Digitalisierung, Speicherung auf Festplatte und anschließende Analogisierung benötigt nun mal etwas Zeit, und so hören wir beispielsweise die Aufnahme am Mikrofoneingang eines PCs erst etliche Millisekunden später am PC-Lautsprecher. Ein Mithören bei der Aufnahme ist für die beteiligten Sprecher oder Musiker so überhaupt nicht möglich!

(evtl. Bild:

2 Wege: analog ohne Zeitverzögerung in Echtzeit – digital mit Zeitverzögerung durch A/D- und D/A-Umsetzer)

## Aufnahmetechnik

### Mikrofon-Aufnahmetechnik: Elektromechanische Wandler

Mikrofone wandeln mechanische in elektrische Energie. Dabei gibt es jedoch unterschiedliche Wandlungsprinzipien und Richtcharakteristiken, die sich in der Bauform und vielen weiteren Eigenschaften der Mikrophone niederschlagen. Im Prinzip sind auch piezoelektrische Wandler und Kohlekörnermikrofone für Tonaufnahmen denkbar, für die Tonstudiopraxis von besonderer Bedeutung sind jedoch ausschließlich das elektrostatische und das elektrodynamische Wandlungsprinzip.

### Kohlekörner-Mikrofone

Das schon 1860 entwickelte Kohlemikrofon wird auch heute noch gelegentlich zur Sprachübertragung verwendet. Treffen Schallwellen auf die Membran, so werden die einzelnen Kohlekörnchen in der „Sprechkapsel“ analog dazu mehr oder weniger fest aneinander gepresst – damit ändert sich der Übergangswiderstand  $\Delta R$  benachbarter Kohlekörner (Kohlegrieß).

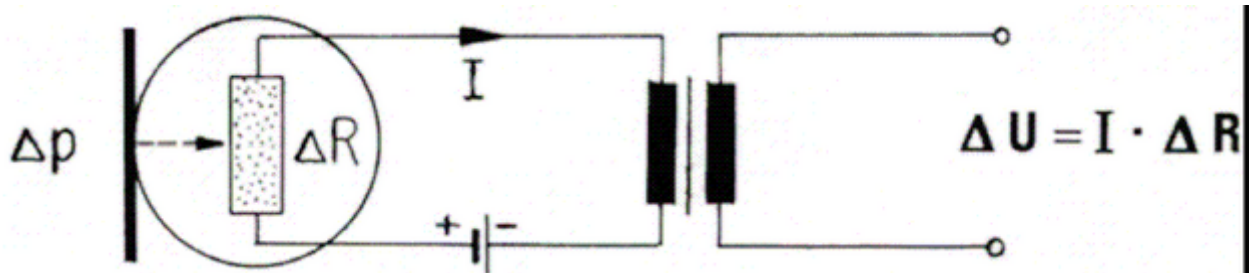


Bild FB 3.1: Prinzip eines Kohlemikrofons

Wird eine Gleichspannungsquelle in den Mikrofonkreis gelegt, so bewirkt diese Widerstandsänderung auch eine Spannungsänderung. Durch einen Übertrager wird der Wechselspannungsanteil  $\Delta U$  (Signalspannung) abgetrennt.

Die Spannungsänderungen sind den Widerstandsänderungen nicht exakt proportional, darum ist der Klirrfaktor von 10 % recht hoch. Auch wenn hier das primitive Prinzip des gesteuerten "Wackelkontakts" angewendet wird, so ist die Sprachverständlichkeit ausgesprochen gut.

### Piezelektrische Mikrofone

Viele kristalline und keramische Materialien weisen die Eigenschaft auf, dass bei Verformungen – sprich: mechanischer Beanspruchung – auf der Kristalloberfläche elektrische Ladungen entstehen. Diese Ladungsverschiebungen lassen sich mit einem Ladungsverstärker in eine analoge Spannung umwandeln. Als Piezelektrisches Material wird meist eine Piezokeramik (Blei-Zirkonat-Titanat PZT) verwendet.

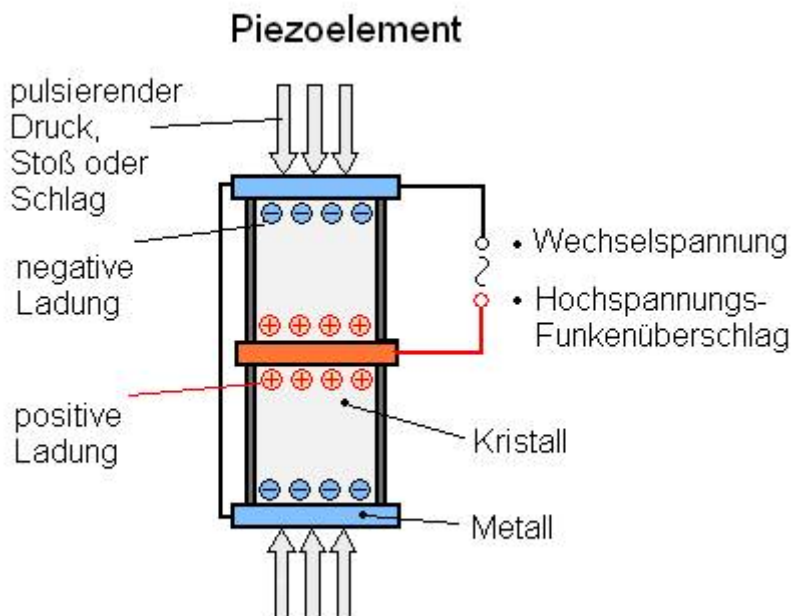


Bild FB 3.1: Prinzip eines piezoelektrischen Wandlers (Bild aus Wikipedia)

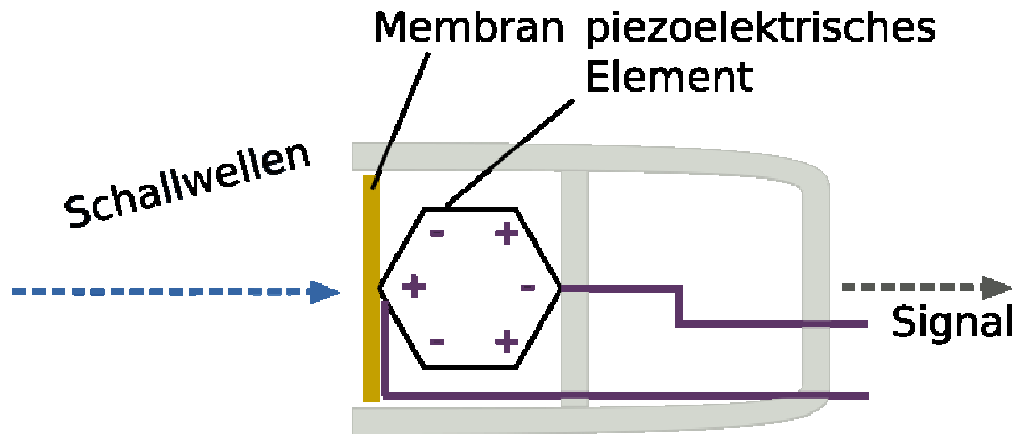


Bild FB 3.1: Prinzip eines piezoelektrischen Mikrofons (Kristallmikrofon, Bild aus Wikipedia)

Piezoelektrische Mikrofone, auch Kristallmikrofone genannt, sind mechanisch robust und haben Vorteile durch ihre simple Bauweise. Der Klirrfaktor und der Frequenzgang in den tiefen Frequenzen lässt jedoch arg zu wünschen übrig. Piezoelektrische Wandler finden wir deshalb vorwiegend als „Kontaktmikrofone“ bei bestimmten Musikinstrumenten (siehe „Tonabnehmer“) und als Ultraschallwandler – oder durch Umkehrung des physikalischen Prinzips (Anregung des Materials durch elektrische Spannung zu Schwingungen) als Hochtonlautsprecher. Nicht zu vergessen der Einsatz der Quartzsteine als Taktgeber in elektronischen Schaltungen, z.B. allen PCs und Hochfrequenz-Empfangsgeräten wie Radios, Fernsehgeräten und Handys.

### Elektrostatische Mikrofone

Vom Physikunterricht in der Schule wissen wir, dass sich die Kapazität eines Plattenkondensators verändert, wenn sich der Abstand der Platten zueinander verändert.

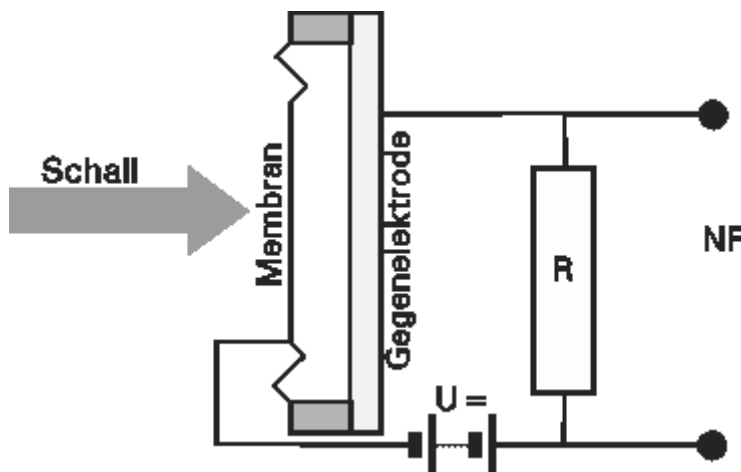


Bild FB 3.1: Prinzip eines Kondensatormikrofons

Bei Kondensatormikrofonen haben wir statt einer Platte eine etwa 1-10  $\mu\text{m}$  dünne, den elektrischen Strom leitende, schwingungsfähige Membran. Diese ist planparallel – in einem Abstand von etwa 5-10  $\mu\text{m}$  – gegenüber einer feststehenden Gegenelektrode eingespannt. Durch eine Polarisationsspannung werden die beiden Elektroden mit einer festen Ladungsmenge  $Q$  aufgeladen. Trifft nun Schall auf die schwingungsfähige Membran, so verändert sich die Kapazität dieser Anordnung und es fließt ein schwacher Strom mit der Frequenz der Schallschwingung am Ort der Mikrofonmembran.

Der Hauptvertreter für das elektrostatische Wandlungsprinzip ist das Kondensatormikrofon. Es bietet hohe Qualität und hat einen sehr guten Frequenzgang, benötigt aber eine zusätzliche Betriebsspannung. Diese wird oft als

„Phantomspannung“ von 48 V vom anschließenden Vorverstärker geliefert und über das Mikrofonanschlusskabel eingespeist. Bei sogenannten „Elektret- Mikrofonen“ reicht jedoch eine geringe Vorspannung von 1,5 bis 3 V, so dass handelsübliche Batterien oder Akkus im Mikrofongehäuse selbst untergebracht werden können. Mit dieser Spannungsversorgung lässt sich auch ein im Mikrofon integrierter Impedanzwandler betreiben.

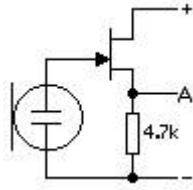


Bild xyz: Elektretmikrofon mit Impedanzwandler

Dreipolige Elektret-Kapseln werden vorzugsweise in Drainschaltung betrieben (siehe Bild) und sichern einen geringen Klirrfaktor. Zweipolige Elektret-Kapseln werden in Sourceschaltung betrieben – damit kann sogar die Spannungsversorgung entfallen und es reicht eine zweipolige Verbindung zum nachfolgenden Verstärker. Deshalb findet man diese Variante als Standard bei PC-Soundkarten mit ihren 3,5 mm-Klinkenstecker-Anschlüssen.

### Elektrodynamische Mikrofone

Ähnlich wie beim Dynamo am Fahrrad entsteht hier eine induzierte Spannung durch die relative Bewegung eines elektrischen Leiters in einem magnetischen Feld. Beim elektrodynamischen Mikrofon ist eine Membran schwingungsfähig mit einer Spule verbunden. Trifft eine Schallwelle auf die Membran, so bewegt sie die mit ihr verbundene Spule im Luftspalt eines Topfmagneten. Dabei wird in der Spule eine Spannung mit der Frequenz der Schallschwingung induziert.

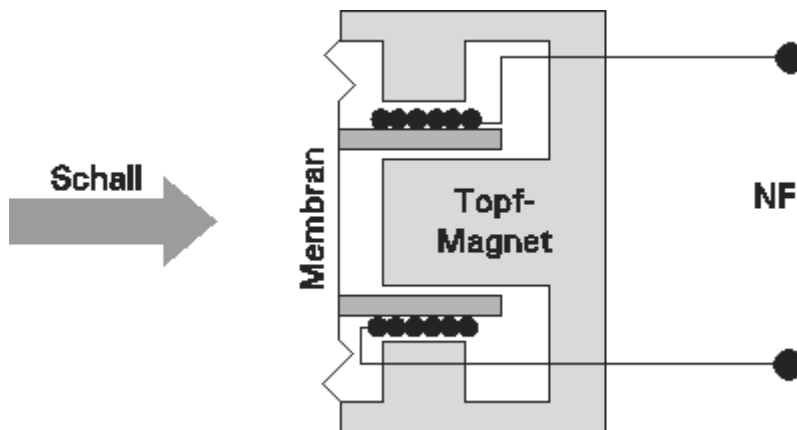


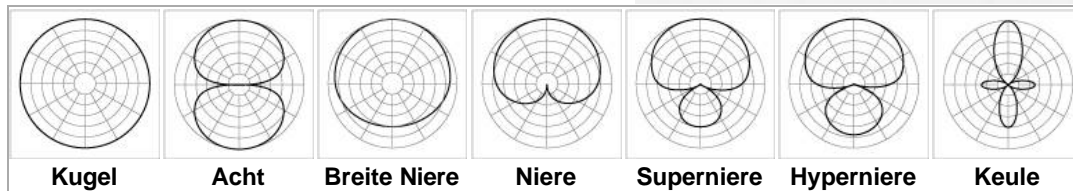
Bild FB 3.2: Prinzip eines Tauchspulenmikrofons

Der Hauptvertreter des elektrodynamischen Wandlungsprinzips ist das Tauchspulenmikrofon. Es benötigt keine Speisespannung, ist sehr robust und arbeitet auch noch bei hohen Lautstärken einwandfrei. Sein Frequenzgang ist jedoch im Vergleich zum Kondensatormikrofon unregelmäßiger.

## Richtcharakteristik

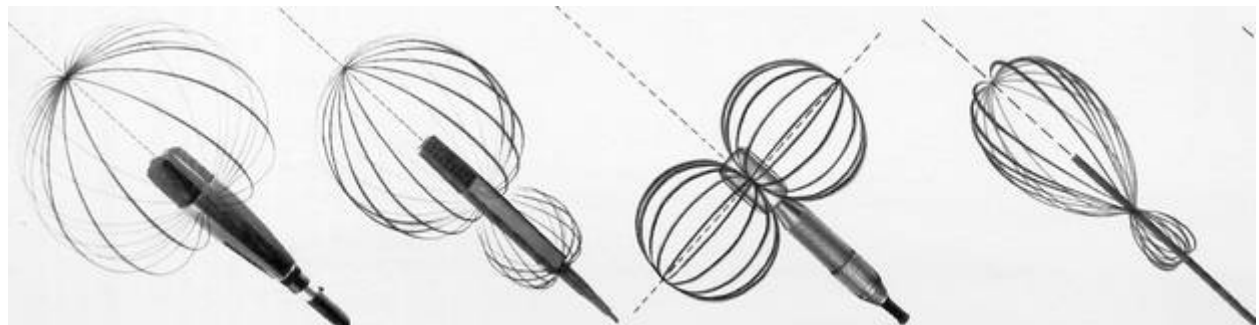
Um den Einfluss eines Raumes auf den Klang oder von Störgeräuschen aus der Umgebung schon bei der Aufnahme zu unterdrücken, stattet man Mikrofone mit einer *Richtwirkung* aus. Sie ist um so größer, je „unempfindlicher“ das Mikrofon gegenüber dem Schall von den Seiten oder von hinten ist, im Verhältnis zu dem Schall, der direkt von vorn auf die Membran trifft. Um dies für den Anwender zu verdeutlichen, wählt man eine Darstellung der Richtcharakteristik in Polarkoordinaten, die jedoch nur in einer Ebene dargestellt zu werden braucht.

Nach dem Grad und der Ausprägung der Richtwirkung abgestuft unterscheiden wir folgende Charakteristiken: Kugel, Acht (wobei die „Acht“ wie die Form der Ziffer 8 sowohl nach vorne und nach hinten weist, gegensätzliche Polarität), Niere, Superniere, Hyperniere und Keule:



**Bild X:** Typische Richtcharakteristiken von Mikrofonen

Niere und Superniere sind Standardcharakteristiken. Bedingung für den Einsatz einer Kugelcharakteristik ist eine gute Raumakustik, Atmos (typische Umgebungsgeräusche) lassen sich damit gut einfangen. Bei Reportageeinsätzen unter ungünstigen störenden Umgebungsgeräuschen oder beim „Angeln“ von O-Ton beim Film werden gerne Mikrofone mit Hypernieren- oder Keulencharakteristik verwendet.



Bisweilen fragen Musiker und Amateure, was denn das in den Mikrofonprospekten so aufdringlich erscheinende geheimnisvolle Wort "Druckgradienten-Empfänger" bedeutet. Nun: Einmembran-Mikrofone mit Kugel-Richtcharakteristik sind reine Schalldruck-Wandler und werden als „Druck-Empfänger“ bezeichnet. Einmembran-Mikrofone mit Achter-Richtcharakteristik wandeln den Schalldruckunterschied vor und hinter der Membran und sind somit sind reine „Druckgradienten-Empfänger“. Die dazwischen liegenden Nieren-Typen sind Mischformen mit unterschiedlichen Druck- und Druckgradienten-Anteilen. Im technischen Sprachgebrauch werden alle Mikrofone mit Richtwirkung durchweg als "Druckgradienten-Empfänger" bezeichnet. Es ist also wenig informativ, wenn in den Mikrofonprospekten an fast allen Mikrofonen die Bezeichnung "Druckgradientenmikrofon" steht, denn das heißt dann ganz einfach "Richtmikrofon" und sagt nichts über die Ausprägung der Richtwirkung aus.

Eine spezielle Bauform sind Mikrofone, die – vergleichbar mit optischen Spiegelteleskopen – im Brennpunkt eines (akustischen) Parabolspiegels angeordnet sind und somit eine sehr starke Bündelung der Richtwirkung ermöglichen – ideal für Abhöraktionen aus großen Entfernungen...

## Weitere Tonabnehmer

Folgt: Z.B. Spulentonabnehmer für Saiteninstrumente (Stahlsaiten), evtl. Luftschallwandler / Flüssigkeitsschallwandler (Hydrofon, Sonar) / Kontaktwandler

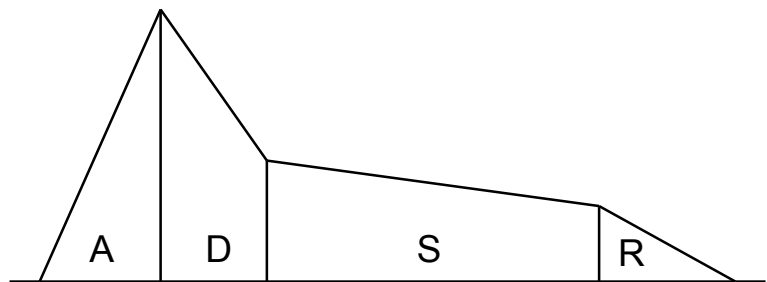
## Klangsynthese / Klangerzeuger

Wenn nach dem Fourier-Theorem jede periodische Schwingung als Addition von reinen Sinusschwingungen analysiert werden kann, so lassen sich beliebige Audiosignale auch umgekehrt aus der Addition einzelner Sinusschwingungen synthetisieren. Das ist auch das Grundprinzip aller analogen Synthesizer. Charakteristisch ist jedoch auch der zeitliche Verlauf von Lautstärke der Grundschwingung und der beteiligten harmonischen Obertöne sowie evtl. Geräuschanteile. Bei Blasinstrumenten ist zum Beispiel zu Beginn jeden Tones nur der Luftstrom in das Instrument zu hören (Rauschen), bis sich die Luftschwingung auf die (durch Luftöffnungen oder Ventile variable) Resonanzfrequenz des Instruments einschwingt.

- Einschwingzeit: Anfangsphase, in der sich ein Ton (mit Obertönen) aufbaut
- Ausschwingzeit: Phase, in der ein Ton ausklingt
- Einschwing- und Ausschwingverhalten (etwa 0,2 Sek.) der Töne sind charakteristisch für einzelne Instrumente

Wir verwenden deshalb eine Amplituden- (und Frequenz- bzw. Filter-) Modulation nach dem ADSR-Modell:

- Anstiegszeit (attack time)
- Abfallzeit (decay time)
- Haltezeit (sustain time)
- Abklingzeit (release time)



Eine andere Möglichkeit ist natürlich, real aufgenommene Klänge zu speichern und auf Abruf zur Verfügung zu stellen. Ein berühmtes Instrument aus der Analogzeit ist das „Mellotron“, das die Klänge auf Tonbändern konservierte, die beim Tastenanschlag der Klaviatur einmal abgespielt wurden (je Keyboardtaste ein eigenes Tonband mit eigenem Wiedergabekopf, max. Tondauer etwa sieben Sekunden!) und dann über einen Motor wieder in die Startposition gezogen wurden.

Im Digitalzeitalter sind die Klänge natürlich digitalisiert und werden aus einem Speicher heraus in beliebiger Länge abgerufen. Die experimentelle Erzeugung völlig neuer Klänge und „Instrumente“, die noch nicht gesampelt (digitalisiert) vorliegen, gestaltet sich mit diesen Geräten allerdings meistens schwieriger.

Bei auf der Bühne spielbaren Instrumenten sind Klangerzeugung und in der Regel eine Klaviatur („Keyboard“) im selben Gehäuse untergebracht. Klangerzeugung und Spieltisch lassen sich jedoch auch getrennt betreiben und über eine sogenannte MIDI-Schnittstelle miteinander verbinden (MIDI = Musical Instrument Digital Interface). Dazu sendet das Keyboard (oder ein PC) Informationen über den zu erzeugenden Ton an den Klangerzeuger, beispielsweise „Note-on“ („Schalte Ton an“) mit „Tonhöhe“ und „Velocity“ („Anschlagsstärke“) und „Note-off“ („Schalte Ton aus“).

Im Prinzip ist MIDI eine unidirektionale Schnittstelle zur seriellen Datenübertragung, die Übertragungsgeschwindigkeit beträgt feste 31250 Bits pro Sekunde.





## Wiedergabetechnik: Lautsprecher und Kopfhörer

Der Lautsprecher ist auch im Zeitalter der digitalen Höchstleistungen immer noch das schwächste Glied in der elektroakustischen Übertragungskette. Er hat die schwere Aufgabe, sehr schnelle elektrische Impulse in Luftbewegung umzuwandeln. Ein Lautsprecher hat dann einen guten Wirkungsgrad, wenn er das möglichst ohne Verluste schafft. Einen solchen Lautsprecher muss man jedoch erst noch erfinden. Heutige Modelle wandeln leider 9/10 der eingegebenen Leistung in Wärme um.

Auch hier werden die von den Mikrofonen bekannten elektromechanischen Wandlerprinzipien verwendet:

Der **elektrodynamische Wandler** besteht aus einem Permanentmagneten sowie einer schwingfähigen Membran, in deren Mitte sich eine Spule befindet, die vom Strom des Musiksignals durchflossen wird und dadurch als Elektromagnet wirkt. Die beiden Magneten ziehen sich je nach Musiksinal unterschiedlich stark an, bzw. stoßen sich ab. Durch diese Bewegung wird die Membran in Schwingung versetzt und erzeugt so Schalldruckwellen (=Töne). Da die Bewegung nur durch die Spule erzeugt wird, breitet sich die Schwingung der Membran ringförmig von der Mitte nach außen aus. Dieses Prinzip entspricht genau der umgekehrten Wirkungsweise von dynamischen Mikrofonen. Und in der Tat lässt sich jeder Lautsprecher nicht nur theoretisch auch als Mikrofon einsetzen und umgekehrt – allerdings darf man dabei natürlich keine klanglichen Genüsse erwarten...

Die Membran versetzt die angrenzende Luft in Schwingungen. Sie sollte leicht und steif sein (zwei sich widersprechende Forderungen!). Als Membran wird u.a. langfaseriges Papier, Kunststoff (Polypropylen, Kevlar), Hartschaum (Polystyrol) oder Aluminium (Wabenstruktur) verwendet. Die Papiermembran wurde in den letzten Jahrzehnten so verbessert, dass sie sich gegenüber den modernen Materialien ausgezeichnet behaupten konnte.

Anders sieht es beim **elektrostatischen Wandler** aus. Hier besteht die Membran aus einer dünnen Polymerfolie, die zwischen zwei Flächenelektroden aufgespannt ist. Dadurch wird die gesamte Membran bei einer Signaländerung gleichmäßig in Schwingung versetzt und zwar so schnell, dass diese Schallerzeugung das Ursprungssignal fast völlig originalgetreu und frei von Verzerrungen und Fehlern umwandelt. Das liegt auch daran, dass die Masse des Wandlers wesentlich kleiner ist, als Spule und Magnet mit ihrer mechanischen Konstruktion im dynamischen System. Da elektrostatische Lautsprecher (und Kopfhörer) durch höheren Produktionsaufwand auch deutlich teurer sind, werden sie fast ausschließlich im absoluten Profisektor eingesetzt.



Die Grundform des elektrodynamischen Lautsprechers ist wohl der **Konuslautsprecher** (Konus = Kegel ohne Spitze), mit der in einen Alu- oder Druckgusskorb eingeklebten Membran und dem dahinter hängenden Magneten. Konuslautsprecher gibt es in allen Durchmessern. Faustregel: Je größer der Durchmesser, desto mehr verschieben sich die Übertragungsfrequenzen des Lautsprechers nach unten, also zu den Tieftönen.

Konuslautsprecher mit geringem Durchmesser sind daher vorzugsweise für die Wandlung von hohen Frequenzen geeignet, mittlere Durchmesser weisen auf Mitten- oder Breitband- Lautsprecher hin, Basslautsprecher haben in der Regel möglichst große Durchmesser. Das ist auch logisch, weil nach den physikalischen Grundgesetzen die tiefen Töne sehr lange Wellen aufweisen. Der Lautsprecher muss erst mal gewaltige Luftmassen in Bewegung setzen, bevor ein tiefer Ton zustande kommt. Das ist so wie beim Paddelboot: Mit einem Eislöffel kann man kein Wasser verdrängen, erst mit einem schaufelartigen Ruder geht es vorwärts.

Spezielle Konuslautsprecher haben in der Mitte noch einen zweiten Konus eingeklebt, der die dort auftretenden Mitten und Höhen besonders verstärkt. Sie sind sehr geeignet für zusammengeschaltete Systeme mit vielen solcher

Lautsprecher, aber auch für Monitorboxen. Vorteil: die Frequenzweiche entfällt, man hat trotzdem ein relativ breites Frequenzband.

Andere Konuslautsprecher haben eine Alukalotte (das ist eine silberne Halbkugel in der Mitte), um besonders die Mittenabstrahlung zu unterstützen. Diese Lautsprecher werden gerne für Gitarrenboxen verwendet.

Einen anderen Typus des Lautsprechers haben wir im Kalottenlautsprecher, der keinen Konus hat, sondern nur die Kalotte, die fest eingespannt ist. Kalottenlautsprecher strahlen obere Mitten und Höhen ab. Aus HiFi-Boxen kennt man sie schon lange, es gibt sie auch für Musiker. Durch die sich vorwölbende Kalotte haben sie einen großen Abstrahlwinkel (180 Grad).

Schließlich sind da noch die **Hörner**, die aus zwei Teilen bestehen: dem eigentlichen Horn und dem so genannten Treiber. Der Treiber arbeitet nach dem Druckkammerprinzip. Eine dünne Metallfolie "schwimmt" auf einem Luftdruckpolster und wird von dem Magneten mehr oder weniger angezogen.



Die Hörner müssen sehr genau gearbeitet sein, die erzeugte Druckwelle soll ja nahtlos an die Luft weitergegeben werden. Die Krümmungen der Trichterwände berechnen sich daher nach komplizierten exponentiellen Funktionen.

Hörner gibt es für Hoch- und Mittelton Einsatz in vielen Größen. An manchen können sogar mehrere Treiber gleichzeitig angeflanscht werden. Hörner strahlen den Schall sehr gerichtet ab, sodass er sich erst auf bestimmte Entfernung entfalten kann. Das kann Vor- und Nachteile haben, je nach räumlichen Gegebenheiten. Hörner verfälschen den Klang auch irgendwie, weil da immer eine Art Megaphoneffekt mitwirkt. In sehr großen Bühnenanlagen haben sie sicher ihre Berechtigung und Wirkung. Viele Musiker verzichten aber in ihren PA-Anlagen auf Hörner.

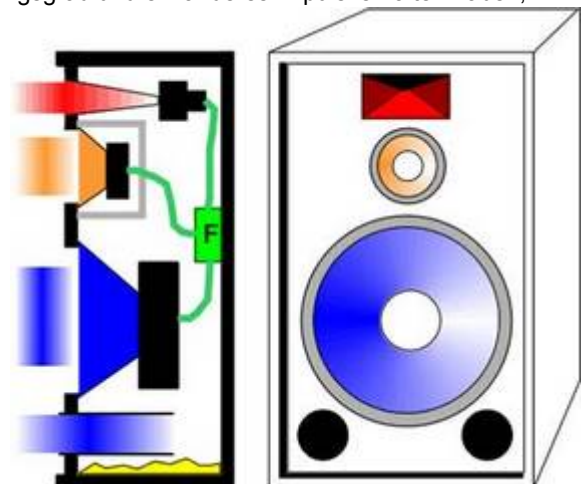
Eine neue Sorte Hörner wurde Mitte der 70er Jahre entwickelt. Es sind die piezokeramischen Hörner, kurz **Piezos**, die sehr klein sind, aber bedenkenlos und in Massen in Reihe oder parallel geschaltet werden können - wohlgemerkt ohne Frequenzweiche! Sie haben auch kaum Grenzen in der Belastbarkeit. Bei ihnen verformt sich die keramische Membran bei Anlegung einer Wechselspannung (Musiksignal) und erzeugt akustische Schwingungen, die ab 2000 Hz bis 25 KHz reichen.



Die Piezos werden aber mittlerweile kaum mehr eingesetzt. Wohl, weil man erkannt hat, dass sie den Sound nicht verbessern, sondern ihn eher schrill machen. Bei Rückkopplungen können sie sich im Ultraschallbereich selbst zerstören.

HiFi-Lautsprecher eignen sich schlecht für Bandzwecke (und umgekehrt), da sie – konstruktionsbedingt durch weiche Aufhängung und lange Schwingspule – einen schlechten Wirkungsgrad und ein anderes Impulsverhalten haben, dafür aber einen ausgeglichenen Frequenzgang.

Musikerlautsprecher sind dagegen hart aufgehängt, haben eine speziell gewickelte kurze Schwingspule und können große Impulse und damit Membran-Auslenkungen aushalten.



## Lautsprecherboxen

Die Mehrwegbox ist ein breitbändiges Beschallungssystem mit zwei oder drei Lautsprecherwegen und Reflexrohr. Die Lautsprecher werden dabei über eine Frequenzweiche geführt. Der Mitteltöner ist abgedeckt, damit seine Membran nicht von den Basswellen beeinflusst wird. Diese Box ist als HiFi-, Studiomonitor-, Gesangs- oder PA-Box geeignet, weil sie alle Frequenzen gut überträgt, durch das Reflexrohr auch in den Bässen Einiges zu bieten hat und durch das Hochtontorn weit reicht. Solche Boxen gibt es heute in sehr kompakter Form mit hoher Leistung.

## Kopfhörer

Während Sie es tun, nehmen Sie die Welt um sich herum nicht mehr wahr. Und nachdem Sie es getan haben, könnten Ihnen die Ohren klingeln. Jugendliche Musikfans bevorzugen meist die federleichten Ohrstöpsel, doch für Musiker und Studientechniker sind solche Exemplare denkbar ungeeignet.

**Offene Systeme** sind für das natürliche, originale Hören die ideale Lösung. Technisch ist dieses Ideal aber nicht vollkommen zu erreichen, weil eine Konstruktion immer das ungehinderte Schwingen und Strahlen der Membran verhindert. Durch eine 'luftige' Bauweise und 'weiche' Materialien wird von den Entwicklern versucht, diesem Ideal möglichst nahe zu kommen, wobei der Erfolg (im sog. akustischen Abdeckungskoeffizienten beschrieben) bei etwas 30 bis maximal 55 % Durchlässigkeit besteht. Offene Systeme liegen locker auf dem Kopf auf und schirmen das Ohr nur wenig nach außen ab, so dass auch Geräusche aus der Umgebung voll wahrgenommen werden, bzw. auch die Umgebung Geräusche aus dem Kopfhörer wahrnehmen kann. Daher ist der Einsatz solcher Kopfhörersysteme im Aufnahmebereich eines Studios eher ungeeignet – dafür sind sie auch bei langem Einsatz bequem zu tragen.



**Halboffene Systeme** sind dann meist das Ergebnis der Kopfhörerentwicklung. Hier werden die hohen Frequenzen durch die Kopfhörerkonstruktion nach außen weitgehend abgeschirmt, während die tiefen Frequenzen nach außen durchgelassen werden. Entsprechend können die tieffrequenten Umgebungsgeräusche störend in Erscheinung treten. Dafür werden die tiefen Signalanteile naturgetreuer wiedergegeben als bei geschlossenen Systemen. Vor allem im Heimbereich werden halboffene Kopfhörer überwiegend gekauft, weil der Tragekomfort besser als bei geschlossenen Systemen ist und die Störung durch tiefe Frequenzen dort meist ohne Bedeutung ist.

**Geschlossene Systeme** dichten den Bereich zwischen Ohr und Schallerzeuger für bessere Basswiedergabe und gegen Störeinflüsse von außen ab. Mit ringförmigen Polstern wird die Ohröffnung möglichst komplett umschlossen und von Bügel fest angepresst. Dadurch kann das Tragen nach längerer Zeit leider unbequem werden, außerdem wird bei längerem Tragen der Hitzestau um die Ohren lästig. Dafür schirmt jedoch das Gehäuse und die Ringpolster das Ohr besser vor unerwünschten Umgebungsgeräuschen ab und auch vom erzeugten Schalldruck geht wenig nach außen verloren, so dass vor allem die Basswiedergabe eindrucksvoll druckvoll, wenn auch gepresst bleibt.



## Ton-Aufnahmen in der Praxis

Alle Aufnahmen sollen zunächst in bester Qualität erstellt werden. Denn reduzieren lässt sich die Qualität jederzeit, aber was einmal entfernt wurde oder im Aufnahmematerial gar nicht vorhanden ist, lässt sich niemals wieder hinzufügen. Umgekehrt bedeutet das Entfernen von Störgeräuschen aller Art und das Korrigieren von Aufnahme Fehlern auch mit modernsten Rechner-Techniken einen enormen manuellen Nachbearbeitungsaufwand, den man mit einer klugen Aufnahmetechnik hätte vermeiden können – wenn überhaupt möglich.

1. Akustische Umgebung beachten und evtl. Maßnahmen gegen unerwünschte Reflexionen und Neben- / Störgeräusche treffen.
2. Geeignete Mikrofone auswählen:  
 Richtcharakteristik (Kugel, Acht, Niere, Superniere, Hypernieren, Keule), evtl. mit Windschutz, Ploppschutz  
 Wandlerprinzip: elektrodynamisch / elektrostatisch (Kondensator)  
 Ausrichtung und Abstand des Mikrofons / der Mikrofone zur Schallquelle, evtl. Abschattungen und Kammfiltereffekte beachten  
 (Kleiner Tipp am Rande für Außeneinsätze: für Elektret-Kondensatormikrofone immer einen passenden Batterie-Satz mitnehmen... Und überhaupt kann der LötKolben in Verbund mit einem NF-tauglichen Multimeter und einer handvoll Stecker, Buchsen und Kabel in einem Tontechniker-Koffer kein Schaden sein...)
3. Aussteuerung:  
 nicht untersteuern (analog: Rauschen, digital: Quantisierungsrauschen)  
 nicht übersteuern (analog: Verzerrungen, digital: Clipping)

*Hörbeispiel: Clipping durch Übersteuerung bei der Aufnahme*

### Popschutz

Poplaute entstehen durch die kleinen Windstöße, die der Mensch produziert, wenn er Konsonanten wie P, B, D und T spricht. Man nennt diese Konsonanten auch „Verschlusslaute“, „Plosivlaute“, manchmal auch „Explosivlaute“. Und wie eine kleine Explosion hören sie sich auch an, wenn der Windstoß direkt auf die Membran des Mikros trifft.

Redet man von einem Popschutz, meint man meist die Schaumstoffüberzüge für Gesangs- und Reportermikrofone. Viele Mikros sind bereits durch die Konstruktion des Mikrofonkorbs vor Poplauten geschützt, hier bringt ein zusätzlicher Schaumstoffüberzug nur noch wenig Verbesserung. Anders sieht es aus bei Mikrofonen, die primär für Instrumentenabnahme gedacht sind, aber auch für Gesang oder Sprache benutzt werden sollen. Das Beyerdynamic M201 wird z.B. wegen seines exzellenten Klanges auch gerne für Interviews benutzt. Hier hilft ein Schaumstoffüberzug sehr.

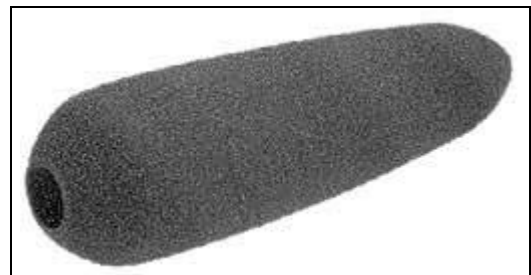


Bild xyz: Für fast jedes Mikrofon gibt es einen passenden Windschutz

Die Schaumstoffüberzüge, die der Volksmund Popschutz nennt, werden von den Herstellern normalerweise als Windschutz bezeichnet. Das hat seinen Grund, denn Windgeräusche bei Außenaufnahmen oder Liveauftritten halten solche Schaumstoffüberzüge viel besser ab als Poplaute.

Ein Schaumstoff-Windschutz kann trotzdem auch im Innenbereich nützlich sein. Viele Sänger können sich nicht angewöhnen, etwas Abstand zum Mikrofon einzuhalten. Ein Windschutz dämpft die Atemgeräusche etwas und schützt den Sänger vor rauen Lippen durch das Drahtgeflecht des Mikrofonkorbs.

### Popschirm

Der Popschirm hat viele Namen: Ploppkiller, Popstopper oder oft auch Popschutz – Schirm deutet aber schon an, dass etwas aufgespannt ist.

Der Popschirm ist historisch gesehen noch gar nicht **so alt**. In den 1970er Jahren arbeitete man entweder ohne jeden Schutz vor Poplauten, was einen etwas größeren Mikrofonabstand bedingt, oder man benutzte einen Schaumstoff-Windschutz. Wer sich Platten aus jener Zeit anhört, kann durchaus den ein oder anderen „Plop“ auf der fertigen Aufnahme ausmachen. Anfang der 1980er Jahre war man um einen transparenteren, **höhenreicheren Sound bemüht**. Man bemerkte, dass Schaumstoff-Windschütze Höhen klauen. Da inzwischen kürzere



Ein Popschirm aus Metall

Mikrofonabstände bevorzugt wurden, weil sie einen **direkteren Sound ergaben**, suchte man nun nach neuen Lösungen. Irgend jemand entdeckte, dass ein Damenstrumpf – über einen Kleiderbügel gezogen und vor dem Mikrofon montiert – ganz ausgezeichnet Plosivlaute ausbremste und dabei die Höhenwiedergabe nur wenig bedämpfte. Wenn Sie beim Video zum Benefizsong „We Are the World“ von 1984 genau aufpassen, sehen Sie solche rustikalen Damenstrumpf-Konstruktionen. Da nur wenige Künstler beim Singen gerne an Damenstrümpfen schnüffeln, kamen bald Lösungen auf den Markt, die diese Konstruktion in ein salonfähiges Design überführten. Et voilà: der Popschirm.

Popschirme gibt es inzwischen in verschiedensten Ausführungen. Der beliebte K&M Popschirm besteht aus zwei Lagen Gewebe, manch andere nur aus einer Lage, wieder andere gestatten durch Klappmechanismen die Zahl der „Filterstufen“ zu variieren. Seit jüngerer Zeit gibt es auch Popschirme aus Metall, die den Luftstrom nicht stoppen, sondern nur umleiten, sodass er die Mikrofonmembran nicht trifft. Metallpopschirme gelten als besonders transparent klingend. Gewebepopschirme haben sich als sehr sicher bewährt; sie schützen auch besser vor ungewollten Spuckattacken auf die Mikrofonmembran.

## Windschutz

Auch Windgeräusche können sehr lästig werden und das aufzunehmende Nutzsignal übertönen. Da der eigentliche Wandler (Mikrofonkapsel) zum Schutz vor mechanischer Beschädigung in einen Mikrofonkopf eingebaut ist, „fängt“ sich der Wind in dieser Konstruktion und verursacht Störgeräusche, die eine saubere Aufnahme manchmal unmöglich machen.

Bei Film, TV und Rundfunk benützt man bei Außenaufnahmen deshalb oft einen speziellen Fellwindschutz, im Fachjargon auch „Windjammer“ oder scherzhaft „Tote Katze“ genannt. Ein solcher Fellwindschutz ist noch effektiver als ein Windschutz aus Schaumstoff. Der Wind fängt sich in den feinen Haaren und wird auf diese Weise verwirbelt und ausgebremst, bevor er die Mikrofonmembran erreichen kann. Schon an schwach windigen Tagen ist ein Fellwindschutz eigentlich unerlässlich für saubere Aufnahmen. Ein Fellwindschutz ist der Regel recht teuer und passt nur auf wenige Mikrofontypen.



Bild xyz: RODE DeadCat Windschutz

## Maßnahmen gegen Körperschall

Mikrofone sind leider nicht nur über die Membran für Luftschall empfindlich, sondern über das Gehäuse auch über jede Form von Körperschall. Das können die Hände des Reporters sein, der beim Interview dem Gesprächspartner das Mikro unter die Nase hält – oder einfach bereits das Scheuern des Kabels in Steckverbindungen. Das Mikrofon ist also am besten gegen direkte Berührungen zu schützen und wird deshalb – wo immer möglich – auf ein Stativ geschraubt. Wird das Mikrofon in der Hand gehalten, so ist die „Kabelschleife“ das erste, was angehende Reporter für ihre Einsätze lernen: Das vom Mikrofon abgehende Kabel hängt nicht einfach herab, sondern der Reporter schlingt das Kabel einmal um seine Hand – quasi als „Zugentlastung“ für das Kabel.

Aber auch bei der Verwendung von Stativen können sich Vibrationen über den Boden und das Stativ bis zum empfindlichen Mikrofongehäuse ausbreiten, weshalb hier spezielle elastische Mikrofonhalterungen (oft auch als „Spinne“ bezeichnet) eingesetzt werden.



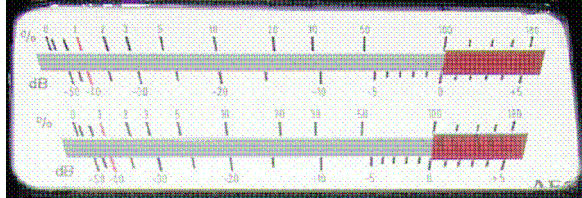
Bild xyz: RODE SM4 Mikrofonhalterung

Anders als Bühnenmikros übertragen Studiomikrofone auch das leiseste Rumpeln – und noch dazu oft lauter als es eigentlich war. Erschütterungen könnten daher eine ansonsten tolle Aufnahme komplett ruinieren. Elastische Aufhängungen entkoppeln das Mikrofon vom Mikrofonständer; Trittschall kann nicht mehr zum Mikrofon gelangen.

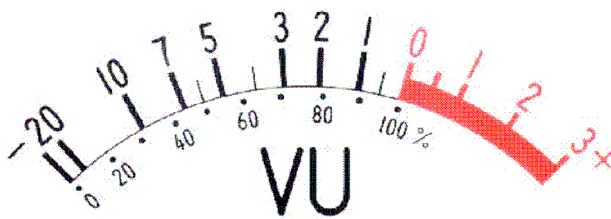
## Aussteuerung und Anzeigeskalen

Alle Aussteuerungsmesser zeigen den effektiven Spannungswert an und repräsentieren somit den Schalldruck.

Bei analogen Aussteuerungsanzeigen gibt es zwei unterschiedliche Skalen: Da gibt es eine in dB und eine in Prozent. Bei der Dezibel-Skala handelt es sich um die Spannung und den dazu proportionalen Schalldruck. Die Prozent-Skala erinnert den Tontechniker an den logarithmischen Zusammenhang von Schalldruck und Schalldruckpegel.



**Bild X:** Lichtzeigerinstrument



**Bild Y:** VU-Meter

Der analoge Aussteuerungs-Spannungspegel wird auf 0,775 Volt entsprechend 0 dB<sub>u</sub> bezogen, wobei der europäische Rundfunk-Studiopiegel +6 dB<sub>u</sub> = 1,55 Volt beträgt.

Der "internationale" USA-Studio-Pegel ist +4 dB<sub>u</sub> = 1,228 Volt. Dieser setzt sich immer mehr durch. In USA gibt es häufig den Bezug zur Spannung 1 Volt, entsprechend dem Pegel von 0 dB-V als Bezugsspannung.

Bei der digitalen Anzeige findet man dagegen nur noch eine dBFS-Skala.

Ein Bezugspegel bezieht sich auf eine Spannung von 1 Volt:	<b>1 V = 0 dB</b>
Der gerade genannte Bezugspegel 1 V = 0 dB wird zur Abgrenzung mit einem Zusatz versehen:	<b>0 dBV</b>
Ein weiterer - heute gebräuchlicher - Bezugspegel für den Nullpunkt, der sich besonders in der Tonstudioteknik durchgesetzt hat, bezieht sich auf eine Spannung von 0,775 Volt, die sich aus einer elektrischen Leistung von 1 mW (Milliwatt) gemessen an einem Widerstand von 600 Ohm errechnet:	<b>0,775 V = 0 dB</b>
Auf diesen Nullpunkt sind nun alle anderen Spannungen bezogen, unabhängig vom Widerstand. Auch für ihn gilt bei Verdoppelung der Spannung eine Zunahme um jeweils 6dB. Zur Unterscheidung wird der Buchstabe U (früher m) angehängt:	<b>0,775 V = 0 dBU</b> <b>1,55 V = +6 dBU</b>
+6dB ist der so genannte Studionormpegel für elektrische Signale. Danach werden z.B. Bandmaschinen eingemessen, weil professionelle Mischpulte diesen Pegel bei Vollaussteuerung abgeben. Man muss also aufpassen, dass man die Nullpunkte nicht verwechselt, weil sich daraus andere dB-Werte ergeben. Noch mal:	<b>1 V = 0 dBV</b> <b>0,775 V = 0 dBU</b>
In der digitalen Welt gibt es ebenfalls eine eigene Pegelnorm: "fs" bedeutet "full scale" und umfasst den 16-Bit-Zahlenwert von -32768 bis +32768.	<b>dBfs</b>

## Raumakustik

Bei Mikrofon-Aufnahmen in geschlossenen Räumen tritt stets auch die Raumakustik in Erscheinung – gleichgültig, ob es sich um ein kleine Kabine für Sprachaufnahmen oder um eine Kathedrale handelt. Das Mikrofon nimmt sowohl den direkten Schall, aber auch den Schall auf, der an den Begrenzungsflächen des Raumes oder an Gegenständen reflektiert, gebeugt, gebündelt oder zerstreut wird.

Besonders bei Sprachaufnahmen ist darauf zu achten, dass keine störenden Reflexionen an das Mikrofon dringen können. Denn bei Laufzeitunterschieden zwischen direktem und indirektem Schall von unter 1 ms bis etwa 20 ms können sogenannte *Kammfiltereffekte* mit unangenehmen Klangfärbungen auftreten.

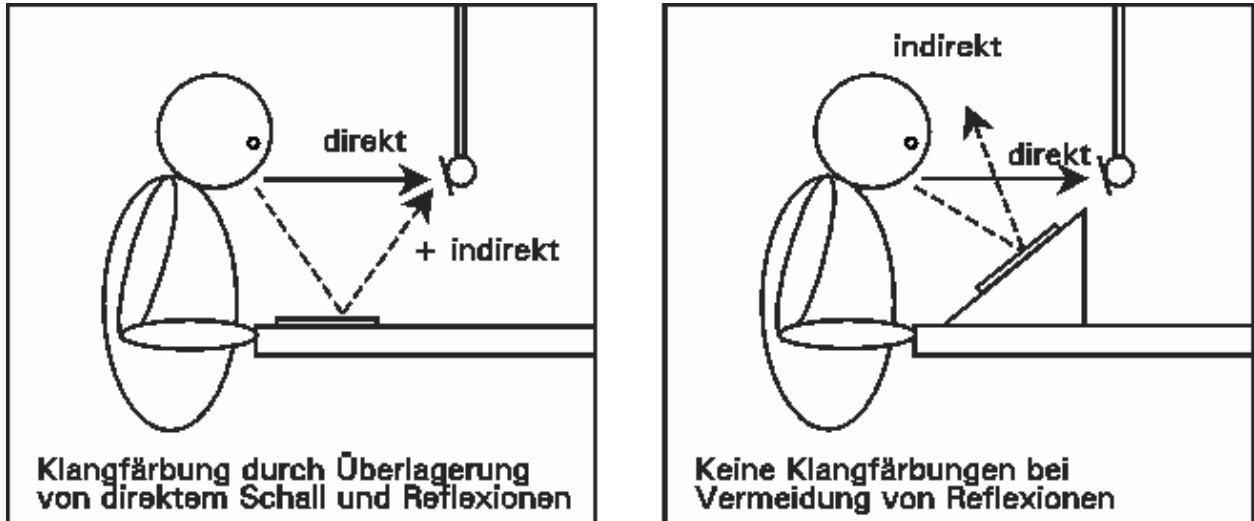


Bild XYZ: Klangfärbungen durch Kammfiltereffekte und Schallreflexionen

## Ton-Bearbeitung nach der Aufnahme:

Wie bereits erwähnt, lassen sich auf dem Computer die aufgenommen Töne nicht nur abhören, sondern auch auf dem Computerbildschirm anschauen. Dabei ist die grafische Darstellung der Wellenform nichts anderes als die Abbildung des Spannungsverlaufes über der Zeit. Die grafische Darstellung ist besonders komfortabel, um geeignete Stellen für den Tonschnitt aufzufinden. Dazu braucht es freilich etwas Übung.



Bild xyz: Tonaufnahme mit Aussteuerungskontrolle (Soundforge)

Die folgende Tabelle zeigt einige typische Schritte, um Tonaufnahmen direkt nach der Aufnahme für eine spätere Mischung und Bearbeitung aufzubereiten. Diese Arbeitsschritte reichen auch in der Regel bereits aus, um Sprache für interaktive Anwendungen auf dem Computer zu bearbeiten:

DC-Offset	Tonsignal symmetrieren: Es kann vorkommen, dass dem Nutzsignal ein Gleichspannungsanteil überlagert ist. Das kann durch Wandlerungenauigkeiten oder bereits auf der Übertragung zum Wandler geschehen sein. Beim Symmetrieren wird der Gleichspannungsanteil wieder aus dem Signal herausgerechnet.
De-Clipping	soll die durch Übersteuerung entstandenen Clipping-Effekte (obertonreiche Verzerrungen) mildern.
Normalisieren	Systemdynamik nutzen: Mit der Funktion Normalize wird die Aufnahme auf einen maximalen Spannungspegel gebracht. Dazu wird zunächst die lauteste Stelle in einem Signalverlauf gesucht, die an dieser Stelle vorhandene Aussteuerungsreserve ermittelt, und anschließend die gesamte Datei um diesen „Reservefaktor“ lauter gerechnet. Automatisch die Lautstärke verschiedener „Takes“ anpassen, in der Regel durch lineare Multiplikation mit errechnetem Faktor auf den Maximalpegel (Obacht: Rundungsfehler)
De-Esser	Zischlaute reduzieren: Bei der Aufnahme von Sprache kann es vorkommen, dass die Stimme einen runden sonoren Klang hat, jedoch die Zischlaute zu stark im Klangbild vertreten sind. Wie stark diese Laute bei Sprechern vertreten sind, ist von Natur aus unterschiedlich. Deshalb ist es nötig, diese Zischlaute abzuschwächen, um ein ausgewogeneres Klangbild der Stimme zu erreichen. Der De-esser ist quasi ein frequenzabhängiger Kompressor.
De-Noiser	Aufnahmen entrauschen: Bei einigen Programmen wird ein sogenannter „Noiseprint“ – ähnlich einem Fingerabdruck – von einer Stelle abgenommen, die lediglich das für eine Aufnahme typische Rauschen enthält. Dieses Rauschen wird analysiert, das Ergebnis dient als Grundlage für den Vorgang des Entrauschens. Das Rauschen wird schlichtweg aus der gesamten Aufnahme herausgerechnet.
Compressor	Aufnahme komprimieren: Der Kompressor hat die Aufgabe, die Dynamik der Aufnahme zu verringern, das heißt das Verhältnis zwischen Laut und Leise zu verkleinern. Laute und leise Stellen rücken in ihrer Lautstärke dichter zusammen. Auf diese Weise werden die Auswirkungen unterschiedlicher Abstände des Sprechers zum Mikrofon ausgeglichen (soweit dies unter klanglichen Aspekten vertretbar ist) und die Aufnahme kann insgesamt in ihrer Lautstärke angehoben werden. Das erhöht die durchschnittliche Lautstärke der Sprachaufnahme. Dieser Effekt tritt auch bei gleichbleibendem Abstand des Sprechers zum Mikrofon ein. Die Folge ist damit eine Erhöhung der Durchsetzungskraft der Stimme und die Systemdynamik wird möglichst gut genutzt. Damit ist sowohl einem gestalterischen wie technischen Anspruch genüge getan. <i>Obacht:</i> Details und Dynamik werden dabei oft „herausgebügelt“, siehe „Mastering“
Fade in/out	Blenden: Um sicher zu gehen, dass eine Aufnahme nicht mit „knacksenden“ Spannungssprüngen anfängt und endet, können kurze Blenden im Bereich von Millisekunden am Anfang und Ende in die Aufnahme gerechnet werden.
Cut	Sprechttext durch Schneiden für die Anwendung vorbereiten: Mundgeräusche und Geräusche des Luftholens werden besonders laut aufgenommen, wenn der Sprecher nahe am Mikrofon spricht. Man kann sie aber in ihrer Lautstärke verringern (siehe Fade in) oder gar herausschneiden. Ebenso können unnötige Pausen am Anfang und Ende von Textpassagen verschwinden. <i>Obacht:</i> das Fehlen und Wegschneiden von Atmungsgeräuschen (insbesondere das Luftholen vor Sätzen) wird oft auch als unnatürlich empfunden! In vielen Anwendungen werden die gesprochenen Texte zum Teil interaktiv abgerufen. Deshalb sind die Sprechtexte oftmals in die entsprechenden Bestandteile zu zerschneiden und in Abschnitten als getrennte Audiodateien zu speichern.

## Sound-Effekte

### Equalizer:

Ein Equalizer setzt sich aus mehreren Filtern zusammen, mit denen das Spektrum des Audiosignals bearbeitet werden kann. Üblicherweise wird ein Equalizer verwendet, um lineare Verzerrungen eines Signals zu korrigieren (siehe: „Qualitätskriterien / Frequenzgang“). Man unterscheidet zwischen verschiedenen Bauarten und Bedienkonzepten.

**Grafischer Equalizer:** Hier ist jedem beeinflussbaren Frequenzband ein eigener Regler zugeordnet (er hat 26 bis 33, typischerweise 31 Frequenzbänder von je 1/3 Oktave Breite), so dass der Verlauf der Frequenzkorrektur „grafisch“ durch die Regler dargestellt wird. Diese Art von Equalizern wird auch x-Band-Equalizer genannt. Das x steht hier für die Anzahl der Regler (bei 31 Reglern als 31-Band-Equalizer).



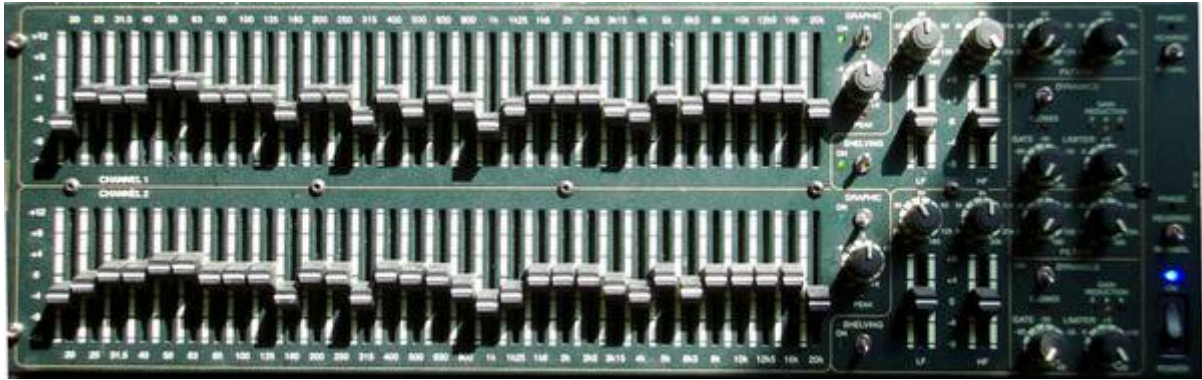
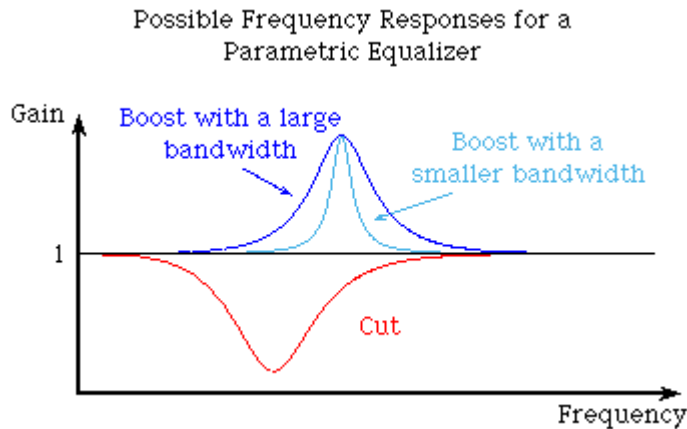


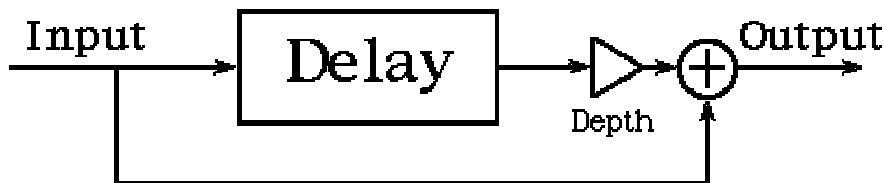
Bild xyz: Professioneller, grafischer 31-Band-Equalizer

**Parametrischer Equalizer:** Hier können für ein oder mehrere Frequenzbänder die Mittenfrequenz und die Amplitudenänderung (semiparametrischer Equalizer) sowie häufig auch die Filtergüte Q (entsprechend der Bandbreite) eingestellt werden (vollparametrischer Equalizer). Diese Bauform findet man vor allem in Mischpulten und Effektgeräten.

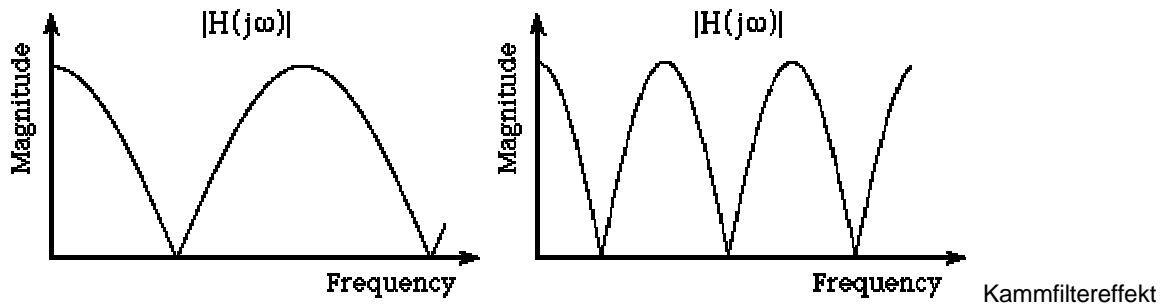


**Flanger:**

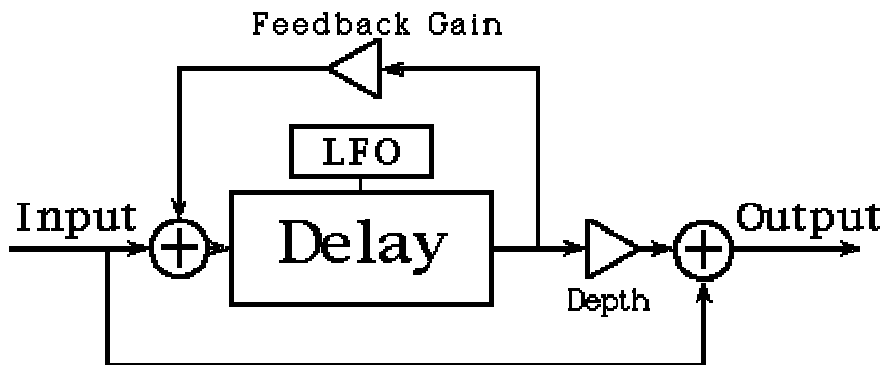
Beim Flanger wird das Eingangssignal zunächst verzweigt in zwei Signalzweige. Der eine Zweig führt das Eingangssignal unverändert zu einer Mischstufe, das Signal des anderen Zweigs wird zeitlich verzögert ebenfalls der Mischstufe zugeführt. Die Mischstufe bildet daraus das spätere Ausgangssignal.



Die Zeitverzögerung wird in einem kleinen Bereich (etwa 1 bis 20 Millisekunden) laufend variiert, wodurch sich kleine Schwankungen der Tonhöhe nach oben und unten ergeben. Durch die Überlagerung mit dem unveränderten Originalsignal ergeben sich Interferenzen (Kammfiltereffekt). Der interessant klingende Effekt des Flangers beruht darauf, dass die Interferenzen aufgrund der variierenden Zeitverzögerung »wandern« und dadurch für Dynamik im Klangbild sorgen. Die Musik bekommt damit einen etwas synthetischen, „spacigen“ Klang.



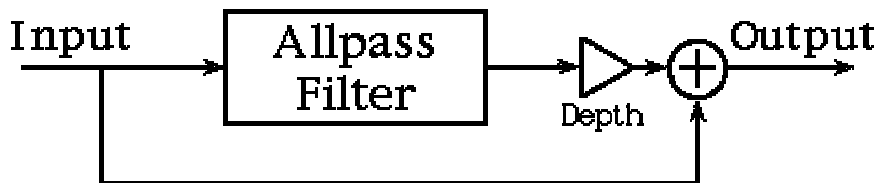
Den klanglichen Hauptunterschied zum sogenannten Phaser-Effekt bewirkt die Rückkopplung des zeitverzögerten Signals zum Eingang. Die Stärke der Rückkopplung kann in der Regel fein eingestellt werden und hat großen Einfluss auf den Klangeindruck.



Mit dem Flanger-Effekt lässt sich auch ein pseudo-Stereosignal erzeugen, indem die anfänglichen Zeitverzögerungen für linken und rechten Kanal gegeneinander phasenverschoben werden.

### Phasing:

Der Effekt beruht auf der periodischen Veränderung einer phasenverschobenen Kopie des Originalsignals. Es bildet sich ein kammartiges Frequenzspektrum mit regelmäßigen Auslöschungen und Anhebungen diskreter Frequenzen. Diese Auslöschungen und Anhebungen werden im Takt einer niedrigen Frequenz (etwa 0,5 bis 10 Hz) im Spektrum verschoben. Dadurch bildet sich der charakteristische Klangeindruck.

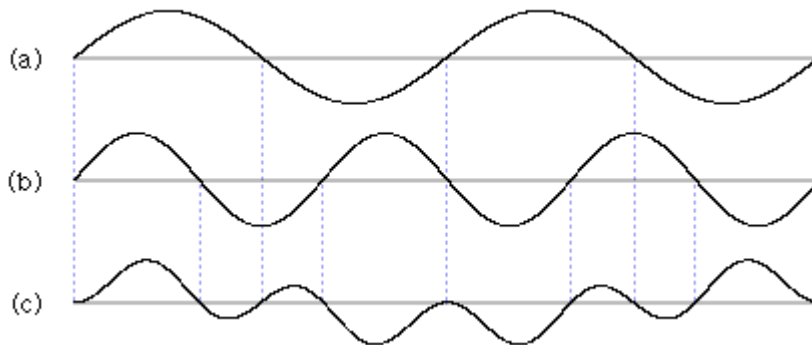
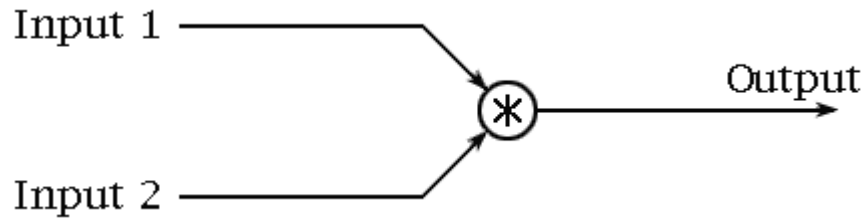


Der Effekt wurde angeblich zufällig entdeckt, als die gleiche Aufzeichnung gleichzeitig von zwei Tonbandgeräten abgespielt wurde. Als die Wiedergabe unsynchron wurde (durch geringe Unterschiede der Bandgeschwindigkeit) stellte sich der Phaser-Effekt ein. Der amerikanische Komponist Steve Reich entdeckte den Phaser-Effekt gegen Mitte der 1960er Jahre für seine Musik. Einige Rockbands der Endsechziger-Psychedelic-Ära setzten Phasing in ihren Songs ein, um ihnen einen schwingenden, "abgehobenen" Sound zu geben.

Der Effekt klingt ähnlich dem Flanger, der die Phasenverschiebung allerdings mit einer zeitverzögerten Kopie des Originalsignals erreicht.

### Ringmodulator:

Ein Ringmodulator multipliziert zwei Audiosignale, wobei als Ausgangssignal die Summe und die Differenz zweier Frequenzen entstehen.



Im Beispiel wird aus einer 400-Hz-Sinuskurve und einem 600-Hz-Signal ein Ausgangssignal, das als Überlagerung einer 200-Hz- und einer 1000-Hz-Sinuskurve zu interpretieren ist.

Dieser Effekt erzeugt – bei entsprechender Hüllkurve des Signals – Gong-artige Klänge.

**Echo:**

Nachhall:

Vorhall:

Kompressor / Expander:

Harmonizer:

Vocoder:

mp3 / aac / mp4 (siehe hierzu eigene Unterlagen zum Download)

## Literatur:

Michael Dickreiter: Handbuch der Tonstudioteknik. 6. Auflage, K.G. Saur Verlag KG, München, 1997, ISBN 3-598-11320-x

Wird fortgesetzt...